

エンジニアズ



ライブラリ

工場における
サンプリング

石川 馨 著

東京
丸善株式会社

まえがき

最近工業に非常な勢で普及しつつある「新しい品質管理」——「統計的品質管理」——の進展に伴い、技術者、経営者、事務的関係者など工業関係者が、統計的にもものを見たり、考えたりするようになってきた。

しかるにここで逢着した問題は、従来から行っていたサンプリング法、測定法、試験法など、JES、JIS、仕様書の不合理という問題である。これを合理化しなければならぬという気運が起つている。これは世界的な現象で、1940年以降、特に1945年以降、米英においても大いに研究、改良されつつあり、それが改訂されつつある最近の米英の規格にも示されている。

ところが、新しい統計学を用いた「標本調査法」(Sample Survey)、「市場調査」(Market Research)に関する著書は、大分でているが、工業関係の原材料、製品などのものを対象にした、工程管理、測定法などを統計学と結びつけたサンプリングに関する著書は、いわゆる「抜取検査」関係以外にはほとんどない。私が日本科学技術連盟の品質管理講習会で、サンプリングに関する入門的なプリントをつくったのは、1950年秋であり、その後何回かプリントを改訂したのも、これらの要望があり、私が燃料関係の研究室におりこの問題に困却していたからであった。その後各方面からのおすすりもあり、前記プリントをもとにして多くの筆を加えたのが本書である。私は化学技術者であり、数学を最も不得意とするところで、統計的な色彩の濃いサンプリングについて本書の筆をとることは不適當かも知れないが、入門書という意味において、わが国産業の合理化の一助にでもなればと思つて、敢えて筆をとつた。

現在問題になっているサンプリング調査に関する問題は、大別して

- 1) 社会政治面における調査のため
- 2) 市場調査のため
- 3) 工業における原材料、製品などの工業材料、工程管理や調査のため

と3方面があるが、本書では3)の工業、特に工場におけるサンプリングをとりあつかった。

工業におけるサンプリング法の合理化研究、ならびに決定には、その専門の技術と新しい統計学との縦横の密接な結びつきが必要である。特に数理統計学を縦横に使わねばならぬが、本書では原則として、できるだけ簡単な方法を用いた。数理統計学や統計的方法については、本書では到底詳細にはのべられぬので、原則として必要最小限にとどめ、その使い方の一部をのべたにすぎない。さらにサンプリングの研究をすすめられたい方は、数理統計学、標本調査法、実験計画法など数学者の書かれた本と併行して、読んでいただきたい。本書では管理図法と統計的方法の入門的な部分を修得されていることを前提として、わかりやすく書いたつもりである。しかしサンプリングにおいて、道具として数理統計学は重要であるが、実際面になると考え方の問題の方がより大切である。この点については、拙文をも恥じず各所に相当くどく書いたつもりである。考え方を理解していただければ、あとは基礎的な統計的な方法を縦横に組合せ、あるいは研究し、実験し、経験すれば、道は自らひらけてくるであろう。そして日本の工業から不合理なサンプリングを追放していただく第1歩ともなれば、私の望外の喜びとするところである。

貿易によらなければ、自立して行けぬわが国の工業としては、統計的品質管理の考え方とその実施と合理化による均一な製品とともに、合理的なサンプリングを確立して、世界の信用をうることが必須条件であると信ずる。また合理的なサンプリング法を確立しておかぬと、これらの取引にお

いて、意外に大きな経済的損失を蒙ることを、工業関係者は銘記しておいていただきたい。特に経営者には、合理的なサンプリングの必要性、その研究の必要性を認識していただきたい。

私の数学的並に技術的知識の不足と、私の技術者としての立場から、数学的に厳密を欠くところや、数学的な間違い、あるいは大学の研究室にいる技術者として、技術的に認識不足の点があると思う。また化学関係者として本書の内容にカタヨリがあるかも知れない。これらの点については、是非諸先輩並に読者諸兄から、積極的な御注意をうけたまわりたいと思っている¹⁾。

終りに私のこの方面への知識につき、多くの御指導や御援助を賜った、Dr. W. E. Deming、日本科学技術連盟リサーチグループ、特に増山元三郎博士、坂元平八氏、浅井晃氏、並に日本科学技術連盟事務局に厚く御礼申上げる。また本書の出版についていろいろ御尽力いただいた、丸善出版株式会社、また研究室で校正その他いろいろ労をおかけした、今泉益正君（現日本鋼管）、勝百合子さんに対し感謝の意を表する次第である。

1952年5月

東京大学工学部応用化学科燃料工学研究室にて

著者

1) 連絡先 東京都（日本橋局区内）中央区日本橋江戸橋2丁目9番地 丸善出版株式会社編集部
東京都（調布局区内）北多摩郡調布町上石原578番地 石川 肇

科学は道具である。

統計学も道具である。

管理図ももちろん道具である。

サンプリングも一つの道具である。

物理も、化学も、数学も、力学も道具である。

道具だけでは何の役にも立たない。

道具は使わなければ何にもならない。

これを使って、役立てる人が必要である。

この道具を総合して使って、役立てる責任者が技術者である。

したがって技術者にはこれを使いこなす責任がある。

しかし道具を使うには、道具がなければならぬし、

道具の使い方を知らなければならぬ。

道具の使い方を知っただけでは何にもならない。

実際に使って役立てることである。

道具の進歩がなければ使い方の進歩もない。

目 次

第 1 章 緒 論

1.1	本書の目的	1
1.2	本書の内容	1
1.3	合理的なサンプリングの重要性	2
1.4	サンプリングの手順	7
1.5	サンプル 母集団 サンプリング	9
1.6	サンプリングとその目的など	16
1.7	対象物による分類	21
第 2 章	ランダムサンプリング	24
2.1	ランダムサンプリングであるための条件	24
2.2	ランダムサンプリングのやりかた	25
2.3	ランダムサンプリングの数理	40
第 3 章	サンプリング及び測定法の誤差	71
3.1	まえがき	71
3.2	サンプリング誤差と測定誤差の関係	71
3.3	信頼性, 精度, 正確さ	78
3.4	測定誤差	85
3.5	サンプリングによる誤差	104
第 4 章	サンプリングの種類と簡単な理論	108
4.1	まえがき	108
4.2	工場におけるサンプリングと社会標本調査法の比較	109
4.3	層別サンプリング	113
4.4	集落サンプリング	125
4.5	副次サンプリング・2段サンプリング・多段サンプリング	131
4.6	有意サンプリング	138

目 次

4・7	ある期間中の平均サンプリング	139
4・8	反覆サンプリング	141
4・9	比推定	142
第5章	サンプリングと予備知識	151
5・1	予備知識と予備調査	151
5・2	ロットの歴史と工程の管理状態	153
5・3	各種分散の推定法	156
5・4	級内分散, 級間分散推定の例	169
第6章	単位体のサンプリング	187
6・1	計量値のサンプリング	187
6・2	計数値のサンプリング	197
第7章	集合体のサンプリング	211
7・1	まえがき	211
7・2	流体のサンプリング	212
7・3	ケーキ状, 泥状体のサンプリング	233
7・4	連続体のサンプリング	236
7・5	同じ大きさの副ロットにわけられている集合体の サンプリング	240
第8章	粉塊混合物のサンプリング	243
8・1	まえがき	243
8・2	バラッキをあたえる原因	244
8・3	サンプリングの面から見たロット	247
8・4	大口試料(1次試料)の集め方	256
8・5	縮分および分析誤差	271
8・6	コークスのサンプリング	287
第9章	総 括	291
9・1	品質管理とサンプリング	291

目 次

9・2 サンプルング問題の解決手順……………303
9・3 むすび……………306

附 録

A 1 分散分析の構造……………307
A 1・1 母数模型と変量模型……………307
A 1・2 一元配量法……………310
A 1・3 二元配量法（繰り返しのない場合）……………313
A 1・4 二元配量法（繰り返しのある場合）……………315
A 1・5 多元配量法……………321
A 1・6 ラテン方格法……………331

附 表……………334

第1章 緒 論

1.1 本書の目的

本書の目的は工場における原材料から、半製品、製品にいたるまでの工業材料並びに工程からのサンプリングを合理化するための特に、数理統計的、技術的、管理的に見たサンプリングについてのべた入門書である。

したがって数理統計学の知識をあまり持っていない技術者でも、理解して戴けるつもりであるが、出来れば統計的品質管理、あるいは数理統計学の勉強と併行して読んで戴ければ幸である。

社会経済面における、あるいは市場調査におけるサンプリング法については、具体的にはふれなかったが、考えかたとしてはその方面にも参考になる点が多いものと思う。

1.2 本書の内容

本書は内容を9章と附録にわけた。

第1章は主としてサンプリングに対する考えかたや手順などをのべた。この章は特に注意して、考えながらよんで戴きたい。特にサンプリングの手順にのべた考えかたは本書を通じて述べたかかった考えかたである。

第2章はサンプリングにおいて必ず考えなければならぬ、ランダムサンプリング、及びこれを実際に現場で実施するときの注意や、ランダムサンプリングしたときの、推定や精度の基本的な数理についてのべた。数理については、本書を通じて、現場で使用しやすい式を主として使うという立場からのべたので、途中の式の証明など面倒と思われる式などは飛ばして読んで戴いてよい。

第3章はサンプリングと密接な関係にある測定法、測定誤差との関係のべた。測定、誤差というものに対する考え方を理解する上において、以下の章を読まれるまえに読んでいただきたい。

第4章は、単純ランダムサンプリングでは、非常に能率がわるいので、これに対するいろいろの方法の代表的なものをのべた。この章では、その

式を導く証明は数理的に相当厄介なものが多いので、その部分は出来るだけ省略し、もっぱらその結果についてのべた。実際のサンプリングにおいてこれらの式を使用する際には、その式を導いた前提条件があるので、その前提条件が、実際の場合に如何になっているかは十分注意していただきたい。

第5章以後の章は主として、実際にサンプリング法を計画実施する際に注意しなければならぬことをのべた。特に第5章はサンプリングについて、予備知識のえかた、予備調査のやりかた、などについてのべたので、この章にかかれたことを、十分に考え、注意して、予備調査を行えばあとのサンプリングは非常に容易に行えるようになるので、重要な章である。

第6章、第7章は単位体、集合体について実際サンプリングを行う上の方法や技術的な注意事項についてのべた。特にガス体、液体などのサンプリングでは、統計的な考え方も必要であるが技術的な注意事項がより以上大切であるので、これに主眼をおいてのべた。集合体中粉塊状物のサンプリングについては、特に考え方や、手法に問題が多いので、第8章に別にのべた。

第9章、むすびとして、サンプリング法標準作成上注意すべき点、品質管理を行い、技術的改良を行う場合に必要と思われるサンプリングに一寸ふれた。

附録には、サンプリング予備調査の際に必要な特殊の解析法をのべた。これは統計的予備知識をもった方によんで戴きたい。

本書の内容は以上のように常識的なもの許りであるが、読者はこれを軽視せずに十分に通読して、実際面に活用していただきたい。日本の工業はサンプリングの理論家ではなく、実行家を要望している。

1・3 合理的なサンプリングの重要性

われわれは、工場で多くのデータをとっているが、その各々のデータを何のためにとるかということを考えて見ると、そこには必ず目的、したがって、そのデータの結果を見て、必ずそれに伴う行動 (action) がある筈である。

たとえば原材料についていえば、その量が契約量通りあるか否か、そ

の品質について契約した純度や強度が保証できるか否かをしらべて、その結果、受入を拒否する、取引上の値段の交渉をする、クレームをつける、工程における原材料の処理、配合条件をきめる等の行動をとる。工程においてデータをとるのは、その工程を作業標準通りに行うため、その工程が作業標準通りに行われているか、異状が起っているかをチェックして、工程に対して行動をとるためである。

したがってサンプリングする前には、あらかじめ、目的をはっきり認識し、行動の基準をはっきり決めておかなければならない。たとえば、鉛石類の購入に際しては、どのような目的に対し、どのようなサンプリングをし、どのように測定した結果が、どのようであれば値段をどのようにするかを、あらかじめはっきり決めておかなければならない。すなわちはっきりした契約書——行動の基準——を決めておかなければならない。

そこでわれわれがデータを得るまでの工程¹⁾を考えて見よう。たとえばある1日に入荷した石炭の平均灰分を知りたいときには、その1日に入荷した石炭からサンプルをとり、これを縮分し、分析して灰分をうるのである。またAという組がある1時間に生産した合成樹脂製品からサンプルをとり、測定してその強度を調べる。すなわちデータを得るためには、考えているもとの集団の中からサンプルをとるという工程を経なければならない。

しかもこのデータをとる目的は何かというと、サンプルの値そのものを知りたいためではなくて、もとの集団—母集団—の性質を正しく推定して処置—行動—をとるためである。このもとの集団とサンプルとを結びつけてくれるのがサンプリングである。したがってサンプリングは、この目的に適するように、

正しく、

信頼性があり、

迅速に、

経済的に、いいかえると、最小の費用で所要精度をうるように、

-
- 1) ここで工程 (process) という言葉を使ったのは、以下に記したことを読まればわかるように、われわれのデータは、サンプリング、測定計算等の人間や機械装置による多くの作業の結果えられるので、工程として見なければならぬからである。

要するに合理的に行われなくてはならない。

ところが、われわれが工業において取扱っているものは、材料でも製品でも必ずバラツキや変化をもっており、それからサンプルをとるので、そのサンプル自身の性質も、繰返し同じ方法で、同じようにサンプリングしても必ずバラツキしている。したがってこのサンプルをとるという工程——サンプリング——は技術的には勿論のこと、統計的な見地、また管理という見地から考えなければならぬものである。

われわれはサンプリングしなければデータが得られず、しかもサンプリングの仕方が悪いと、たとえば1万トンの鉄鉱石が船で入ったときに、サンプリング法が悪く、鉄分が真の値より、1%多く出るようなサンプリング法を行っていわば、非常に大きな損失となる。また中間製品などのサンプリング法が、目的に対して不合理なサンプリング法であるとする、工程に対して間違った判断をし、処置をとったり、あるいは次の工程で誤った配合をしてしまい、製品の品質や収率に大きな損失をあたえる。

サンプリングはこのように大きく経済や技術と関係のある重要な工程であるにもかかわらず、従来はよい加減に、あるいは全く考えられずに行われていた場合が多い。したがって、サンプリングや測定に対する信用がなく、おかしなデータが出ると、自分に都合が悪いときには、サンプリングが悪いのだと簡単に逃げてしまったりしている。これでは一体何のためにデータをとっているのかわからない。わが国の多くの工場などには目的のはっきりしない、行動を伴わない、ただデータをとるためのデータが多く、これがただ過去の記録として、工場の一隅に死蔵されている場合が多い。あるいはサンプリングにより検査するよりは、全数検査の方が正しいというので、その検査の誤差や過誤や非常に多くの経費をかけていることを忘れて、サンプリングによる調査や検査は信用出来ない、必要でないなどいっているのが日本の多くの現状である。しかし全数検査が不可能な場合、かえって信頼性の低い場合、あるいは不経済である場合が多いということは今やあらゆる分野において過去の経験より証明されている。たとえば、化学製品の純度、不純度などのように化学分析を伴う場合、布、鉄板、合成樹脂製品などの強度試験、電球の寿命試験などのように破壊試験では、全数検査は不可能である。ボルト、ナット類のように、非常に多量のもの

が生産され、あるいは取引される時、その全数を測定したのでは不経済であり、間違いも多い。

またわが国に多量輸入されている、綿花、砂糖、羊毛などが何千俵も入って来たときに、その目方を全部測定する場合など、測定者の測定は乱暴になり、秤量機の調整の不十分などにより測定誤差が大きくなり、かえってサンプリングによる計量よりも不正確になる。Dr. W. E. Deming のアメリカの税関における経験でも、「全数調査をやっていた1947年の税関における輸入検査費用及び輸入量を100とすると、1950年にはサンプリング調査を応用したために、輸入量は250~300%に増大したにもかかわらず、経費は20%となった。すなわち順当りの経費になおすと10%以下になり、しかもかえって正確な結果を得ており、荷下しも早く出来るようになったので、待船料の節約も莫大なものである」といわれている。

すなわち、われわれは、あらゆる面で合理的なサンプリングを考えざるを得ない状況になっている。これは世界的な状況で、現在欧米において盛んに研究され、特に米国においては1945年以降その規格を次々と合理的に改良しつつある。

ところが従来わが国で行われているサンプリングはJESやJIS取引の契約書、仕様書、規格など、あるいは工程中や最後の製品のサンプリングにおいても、不明瞭、不明朗、アイマイなものが多い。サンプリング法の不明瞭な規格などはノンセンスである。

また最近普及してきた統計的品質管理に於ても、ただ現在とっているデータから機械的に管理図に書いただけでは、そのサンプリング法を十分反省して見なければ、何を管理するための管理図をかいているのか、あるいはどのような処置をとったらよいか見当のつかないこともある。われわれが統計的な方法や考え方を利用して、データからいろいろな知識をえて、行動にうつろうとするときに、データの歴史、すなわち、いかなる工程を通った品物を、いかなるサンプリング法でサンプルをとり、いかなる性能をもつ測定法で測定されたデータであるかを統計的に十分検討して見なければ、間違った行動をとってしまうことがある。わるいサンプリングより得たデータを統計的に解析をしても、いみのないことが多い。理論的な数理統計学を応用した、管理とか、検定とか、いわゆる抜取検査などに

においても、これが**実際の工業の場との結びつきはサンプリングによりおこなわれているのである**。したがってサンプリングについては、**数理統計学の理論的な面と実際の面——技術的および経済的¹⁾に考えて——との両者を併せ考えて、合理的なサンプリング法——サンプリング法の選定、推定式の決定、サンプルの数、所要精度、信頼度など——をきめなければならない。**

“Sampling is the central problem in statistics.”

たとえば**一定の費用で最良の精度をうるためのサンプリング法、一定の精度を最小の費用でうるための方法、理論的に必要なサンプル数と実際にとりうるサンプル数や分析回数、従来の技術的知識の利用によるサンプル数の減小など、技術と密接な関係がある。**

したがってこのサンプリング法の合理化は、これからの技術者の一つの大きな責務であり、特に数理統計学に理解のある技術者——わが国の現状では特に品質管理技術者——に課せられた問題であろう。数理統計学者の理論のみでも、又統計的知識のない古い技術者のみでも、工場における合理的なサンプリング法を決めることは困難である。

工場における合理的なサンプリング法を決めるのに非常に参考になるのは、統計的品質管理と同様に最近わが国においても急速に普及しつつあるいわゆる**標本調査法 (Sample Survey) の理論と手法である**。したがって本書ではこの方面の手法も相当応用してある。社会科学において、社会現象を対象としてサンプリングして調査する場合と、われわれが工場において材料をサンプリングする場合は類似点も極めて多いが、やや異っている点もある。これについては§4に詳細にのべるが

- 1) 社会現象も1つのプロセスとして考えられているが、工場における材料、製品の方が**はるかにその変化がはげしく、**
- 2) また日常作業として頻繁にサンプリングを行い、
- 3) サンプリングを行うときに**カタヨリが入りやすい (たとえばよいところをえらぶ)。**
- 4) 面接調査のように面倒な、誤差の入りやすい工程はないが、

1) 技術的といったら必ず経済的という意味を含んでいるので、今後は単に技術的という。技術者は科学者ともがう。 Engineer is an economist.

5) 化学分析のように混合試料として縮分や測定が行われる。

などことなつた点もあるので必ずしも標本調査法の技術全部をそのままは取り入れられない。しかし考えかたは非常に役に立つから、更に研究をすすめたいかたは参考書^{1, 2, 3, 4)}を参照されたい。

最後にまとめると合理的なサンプリングを行うためには、

- 1) 目的の明確化 2) 技術的な知識 3) 統計的な知識
- 4) 予備データ 5) 管理という考え方
- 6) 合理的なサンプリング法を確立し、実行しようというポリシーと関係者全員の協力

などが必要である。わが国の場合には特に第6項が大切である。

1.4 サンプリングの手順

われわれがサンプリング法を設計し、サンプリングを行い、データをもとめ、その結果により処置をとるという手順は、これを各段階にわけて考えると次のようになる。

1. 目的をきめる。 “Specify what is wanted.”

われわれがサンプリングを行うときにまず第1に考えなければならぬのはその目的である。いいかえると、われわれがそれからデータをとるのは必ず何か目的があり、その結果により処置、行動をとるためにサンプリングを行っている筈である。したがって何を知らたいためにサンプリングをするのかを十分に考えなければならぬ。従来工場ととられているサンプルには、目的がアイマイで、どのような目的でとられているのか関係者全員にはっきり徹底していないために、ただ漠然とサンプルをとり、測定している場合が多い。無駄な、あるいは不足な、あるいは目的がはっきりしていないために不合理なサンプリングが相当に多い。製品を保証したいためにサンプリングしているのか、工程を管理したいために行っているのか混乱している。現在行われているサンプリング法についてもう一度よく反省して、目的をはっきりさせて、サンプリング法を設計しなおさなければ

- 1) W. E. Deming ; Some Theory of Sampling; John Wiley 1950.
- 2) 齊藤・淺井；標本調査の設計、培風館、1951.
- 3) W. G. Cochran ; Sample Survey Techniques ; Institute of Statistics, North Carolina State College.
- 4) 林；サンプリング調査はどう行ふか、東京大学出版部、1951.

ならぬ場合が相当にある。これは最も大切な段階である。

2. 予備知識を整理して、これからサンプルをとろうとする集団について最大の情報をつかむ。

その集団についての技術的、経験的な予備知識を十分整理しなければ、合理的なサンプリング法を決定することは出来ない。工場には、その集団について過去に何らかのデータや技術的経験的知識がある筈であるからこれを集め、統計的方法を用いて十分解析する。そしてその集団について統計的な模型を考える。また従来のサンプリングや測定が、実際にどのように行われていたかを調査する。できれば過去に行っていたサンプリング法及び測定法の信頼性、精度、正確さ等を検討する。

この場合もっとも必要なのは、その集団の平均値、バラツキの様子、その集団をいくつかの小さい集団にわけたときの平均値やそのバラツキ具合あるいはその材料、製品の管理状態を知ることである。

もし過去に有効な知識やデータがないときには、簡単な予備調査が必要である。この予備調査は、あまり経費のかからぬ、しかし長期にわたって材料や製品やサンプリング法の変化を十分チェック出来るように行った方がよい。

3. サンプリング法—並びに測定法—を決定する。

予備知識を利用して、目的に対して技術的に合理的になるように、統計的にはっきり検討したサンプリング法を決定する。決定したならば出来るだけ詳細な、実施上の細かい注意もつけた、サンプリング法標準を書いたものにしておく。これは必ず書いたものにしておかなければならない。また後にいろいろ改訂することがあるから、ルーズリーフ式などにしておいた方がよい。この標準を関係者一同に徹底させておくことが必要である。いわゆる標本調査の設計とことなり、工業においてはサンプリングを多く現場員が行うので、サンプリング自身にカタヨリが入りやすい。したがってサンプリング法の決定にあたっては、できればサンプリングを管理できるようにチェックサンプルをとるようにしておいた方がよい。

この際に目的などを考えてサンプリングによる誤差を最大どのくらいまで許しうるかを技術的にきめておかなければならない。

4. サンプリングを実施する。

標準サンプリング法によりサンプリングを行い、サンプルをとる。このサンプリングは一つの工程と考えて、十分に管理すること、すなわちサンプリングの実施は管理状態になければならない。

5. えたサンプルを別に決められた測定法により測定する。

この測定法も勿論標準を決めて書いたものにしておくこと。測定法とサンプリングの関係については後章で詳細にのべる。

6. えたデータより結論を出す。

この結論を出すに当っては、技術的に直感的にはっきり結論の出せぬときはデータを統計的に解析し、解釈し結論を出す。

7. えた結論により処置をとる。

最初に立てた目的と結論とを比較して処置をとる。

8. 更にその処置の結果をチェックする。

われわれが目的を決めてから、処置をとるまでには以上の段階を経なければならない。しかしこれだけで終るのではなくて、以上のようにして得た知識をもととして、目的、サンプリング法、測定法、及び処置などについて再検討を加えて、たえず、その目的や状態に適した、また、もっとも合理的な経済的な方法を研究し作成して行かなければならない。たとえば 3)→4)→5)→6)→3)……, 1)→2)→……→8)→1)→2)……等を循環してその状況に応じたサンプリング法を設計し、改良して行くのである。従来の日本の技術者の欠点は、たとえば一度サンプリング法標準を作成してしまると、もうそのままにしておいて、絶えずこれを改訂してゆくという努力にかけることである。または一度に理想的なサンプリングを決めようとして、非常に不経済な努力をしたり、あるいは真に合理的なサンプリング法はきめられぬといって途中で投げ出してしまうことである。

1・5 サンプル 母集団 サンプリング

工業にはいろいろのデータがある。たとえば、原材料、中間製品、製品などの純度、不純度、目方、長さ、キズの数、温度、圧力、時間、収率、コストなどその他いろいろな特性についてのデータがある。われわれはこれらの真の特性の値—特性値—を、すなわち、真の品質を知りたいのである。あるいはこれを管理したり、受入れるか出荷するか否かを決めたりし

たいのである。

しかも、これらのものが完全に同じものであるとか、完全に均一なものであれば、その一部を調査すれば全体の品質などを推定することはできるが、工場におけるデータを見ればわかるように、たとえばわれわれの使用する原料製品の特性には必ずバラツキがある。したがって、その真の値を知るためには、その特性をすべて調査しなければならない。しかし、これには、測定誤差があり、全部を調査することは時間的にも、経済的にも不経済であり、まして破壊試験をとまなうときには、たとえば強度試験や化学分析をとまなう場合には不可能である。

またたとえ全数調査が出来たとしても、また経済ということを見捨てても、かえって不正確な場合が多いということは、過去の経験より証明されている。そして大量の全数をしらべるよりは、その一部を抜きとって、これについて、間違いなく、なるべく精密な調査を行った方が、そのもとの集団の品質の特性につき、より正しい多くの情報を得られる場合が多い。たとえば、入荷した石炭 100 トンの真の発熱量、入荷した 10,000 トンの鉄鉱石の真の鉄分を正しく知るには、これら全部を、正しく測定誤差のない測定法で測定しなければならぬことになるが、これでは利用出来る石炭や鉄鉱石もなくなってしまい、また非常に不経済で到底実行できない。あるいは入荷した 10,000 トンの砂糖の真の目方を知るためには、全部の袋を一つずつ正確な秤りで正しく測定すれば、真の目方がわかるように考えられ勝であるが、たとえば数万袋を一袋ずつ測れたとしても、このような測定を正しく行うことは実際問題として、測定する人間が人夫などであり荷上げの時間が制限されている場合があり、測定も非常に乱暴になり、また各袋の風袋などということを見ると、不正確な値しかえられず、かえってその中の何十袋かを抜きとって、精密に測定した方が遙に経済的に正しい値のえられることが経験されている。

そこでわれわれは、目的とする集団の品質特性を知るためには、その一部を抜きとって、測定を行うということになる。

この場合、われわれが目的とする品質特性の集団を、またはその集団そのものを、母集団（仕切、ロット、バッチ、population, universe, batch, lot, bulk）という。たとえば、今日一日に生産された 1 万コの電球の光

度や寿命、今日入荷した一山の鉱石の成分、あるいはその電球、一山の石炭すべてを母集団という。これに対して、われわれが試験や測定などの調査をするために、その母集団の一部を抜き取ったもの、あるいはその品質特性をサンプル¹⁾ (試料, 標本, sample) という。このサンプルを抜き取ることをサンプリング (試料抜取, sampling) という。

われわれは母集団から、出来るだけ合理的にサンプリングし、サンプルを得て、そのサンプルを測定して、その特性値よりもとの母集団について、出来るだけ信頼性のある、正しい知識を経済的に得たいのである。すなわちわれわれはサンプルの特性値を知りたいためにデータを取っているのではなく、母集団の特性値を知りたいためにデータを取っているのである。

したがって、サンプルのもつべき条件としては、「われわれが知りたい目的に対して、そのサンプルの特性値が、もとの母集団の特性を出来るだけ正しくあらわすようなものでなければならない。しかもこのサンプルの取りかた、サンプリング法が経済的に行われなければならない」ことになる。このためには、サンプリングをいかに行って行くべきかということ論ずることになる。

備考

1) 工場においては、母集団の上に更に大母集団 (供給源, grand population, grand lot, supply) を考え、区別してよぶこともある。たとえばある1日の製品を母集団と考えると、毎日母集団が次々出てくるわけであるが、この母集団が出てくるもとの供給源、すなわち大母集団というものが考えられる。すなわちある1日と技術的に考えて同じ条件で生産したときでも、たとえばいろいろな不良率をもつた製品が生産されてくると考えられる。この技術的に考えて、同じ条件で生産されてくると考えられる製品の集団を大母集団という。したがって、母集団もこの大母集団、供給源に対しては一つのサンプルということになる。この意味において、全数調査を行った場合でもそれは一つのサンプルで

1) 管理図法の場合には、試料 (sample) とは、数コのデータをまとめて1つの群 (sub-group) としたものをいうが、本書では母集団から一部を抜き取ったものを一般にサンプルということにする。

ある。しかしこの場合には、実際問題としてサンプリング法が問題になることはない。本書では母集団からの現実のサンプリングのみを考えていくことにする。

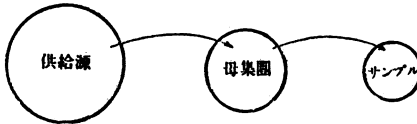


図 1・1 供給源・母集団・サンプルの関係

2) 研究室とか工場で実験を行うとき、たとえば同じ条件で繰返し 3 回実験を行ったとき、このデータはこれと同一条件で無限回実験を行ったときに考えられるある母集団からの 3 コのサンプルであると考えられるが、この場合にも実際問題として、サンプリングの問題も起らないので、本書では省略する。

1・5・1 サンプルのもつべき条件

われわれはある母集団から、一つあるいはそれ以上の特性を知りたいために、サンプリングを行い、サンプルをとるのである。たとえば、ある 1 日に生産されたガラス瓶の母集団から、その容量、目方、成分などを知りたいためにサンプルをとるときを考えると、従来はただ漠然とその中から何本かサンプルをとって来て、これを試験している。一体これではよいであろうか。たとえば、生産に従事しているものがサンプリングして来ると、熟練工になればなるほど、経験から良い瓶と思われるものを意識的に、あるいは無意識的にサンプリングしてはいないか、あるいは熟練した検査工が他からサンプルを取りに行くときには、悪い瓶を多くサンプリングしてはいないか。これではわれわれがサンプルから母集団の特性を推定したいという目的に反することは、明らかであろう。そこでわれわれがサンプルをとるときに、そのサンプルはどのような条件をもっていればよいかを考えて見よう。

1. サンプルの一つまたは 1 以上の特性値が、もとの母集団のそれらの特性値の真の値に出来るだけ近づくようにサンプルをとること。これには 2 以下にのべるような条件を満足することが必要である。

2. 同じサンプリング法を何回か繰返し行ったときに、そのサンプルの特性値は通常バラツキているが、その結果が繰返し回数が増えるに従って、母集団の真の特性値にいくらかでも近づいて行くようなサンプリング法でなければならない。いいかえると、よい品物を常に多くとったり、悪い品物を常に多くとったりするような**カタヨリ** (bias) があってはならない。これにはサンプルが全く**ランダム** (random) にとられていなければならない。

3. サンプルより母集団の特性値を推定した場合、その誤差が合理的にきめられるようなものでなければならない。母集団のある特性の平均値、たとえば今日入荷した石炭灰分の平均値は、信頼度99%で、サンプルの平均値 $\pm 0.5\%$ 以内にあるというように、統計的にはっきりしたものでなければならない。

従来行われている多くのサンプリング法は、ただ漠然とあまり統計的な検討を経ていないので、精度がどのくらいあるか、その信頼度がどのくらいあるかなどということは殆んど論じられていない。まして、調査の目的やそのデータの用途に応じて、必要な一定の精度に応じて、サンプルの数をきめられているような、合理的なサンプリング法は殆んどない。

4. サンプルは信頼性のあるサンプリングより得られたものでなければならない。たとえばよく現場でいわれることは、妙なデータがでると、「あれはサンプリングが悪いのだ」と。これはそのサンプリング法に信頼性がないからである。このように信頼性のないサンプリングならば、サンプルをとるということは気やすめにすぎないといっても過言ではないであろう。われわれが従来から工業でとっているデータには案外気やすめのデータが多い。

5. サンプルをとり測定をするのは、母集団についての知識を得たいためであるが、この得られた知識の価値が、サンプルをとり測定する費用よりも少ないならば、技術的に考えてこのサンプルはとらぬ方がよい。すなわちサンプリングは、経済性を考えて、技術的に必要な精度と結びつけて、もっとも実用的なサンプリング法がきめられるようなものでなければならない。

本書では、このような条件を満足させるには、どのようなサンプリング

法を設計し、どのようなサンプリングを実施したらよいかを論ずることになる。

数理統計学と現実の世界との結びつきがサンプリングにあることを前に述べたが、数理統計学においては、サンプルがある一つの母集団から、ランダムサンプリング(無作為抜取, random sampling) されていることが、前提条件になっている。したがってサンプリング法には、後述のようにいろいろな方法があるが、最終の面においては常に抽象的な無作為化 (randomize) という概念をいかに、現実の世界に適用して、その無作為性を保持するかということに帰する。これによって、その目的に応じたサンプリング法を設計し、それを実施した結果を、統計的に解析することが出来るようになるのである。

サンプリングの問題は、これらのことを考慮に入れて、決定される一つの技術上の問題である。

1・5・2 母集団とサンプルの大きさ サンプルの単位

たとえば、入荷した2,000 コのペヤリングから、10コ of サンプルをとって試験するとき、このとったサンプルの箇数のことを**サンプルの大きさ**(サンプル数 sample size) という。これに対して、入荷したペヤリングの母集団に含まれるペヤリングの箇數 2,000 コのことを**母集団の大きさ**(ロットの大きさ, size of population, lot size 等) という。

われわれが取り扱う母集団には、このように**単位体** (unit) を有限箇含んでいる**有限母集団**¹⁾ (finite population) と、無限箇含んでいると考えられる**無限母集団** (infinite population) とがある。通常われわれが工場で接する母集団には、有限母集団の場合が多いが、サンプリングにより推定を行う場合には、実用的には、統計的に無限母集団と見なしてよい場合が多い²⁾。

この例のように、母集団の単位体が1コ1コ明確に識別出来る場合はよいが、たとえば山積や吹詰した肥料の母集団を考えると、サンプルの単位に何をとるかということが問題になる。

1) 有限箇から成り立っている場合でも、サンプルを取り出して測定し、これをもとにもどして、再びよく攪拌して、サンプルを取るという操作を繰返すときは、(たとえばチップ実験) 無限母集団とみなすべきものである。

2) § 2・3 参照。

サンプルの単位 (sample unit) とは、サンプルの一つに何をとりかということである。電球1コとか、電柱1本、鉄板1枚、薬品1瓶、肥料1匁、糸1束、石炭1貨車などを単位として問題とするとき、すなわち単位体が明確になっているときには、この単位体をサンプルの単位にとればよいのであるから、別に問題はない。

ところが、糸や紙のような連続体の品質特性を問題にするときには、どの位の長さを1単位とするか、流体においては、何ccを、何kgを1単位とするか、あるいは石炭や鉱石のように、大小の塊粉状体の混合しているときには、これらの混合体何g、何kg、をサンプルの単位にとるかが問題になるので、これを明確にしておかなければならない。たとえば、石炭をサンプリングするときに、ある一定の大きさのスコップでとるサンプルの量をサンプルの単位として、サンプルの大きさ、すなわち何杯をいかにしてとるかということを決める必要がある。

このように集合体の場合にとるサンプルの単位を、**インクリメント (increment)** といい、そのスコップ一杯のサンプルの単位量のことを**インクリメントの大きさ (size of increment)** という。糸、針金、紙などのように、長い連続体の場合には、サンプルの単位の長さを**試長 (sample length)** という。

たとえば鉱石や肥料のサンプリングにおいては、インクリメントの大きさを1kgに、500gにとるとか、液体のサンプリングにおいては、インクリメントの大きさを100ccにとるとかいう。また糸などにおいては試長を1mにとるとかいう。この場合1インクリメントの石炭とは、1回の運動でとった量をいうのであって、たとえば500gのインクリメントとは、500g入るスコップで1回でとったサンプルの量をいうのであって、100gずつ5回とり、これを一緒にしたものをいうのではない。後者の場合にはインクリメントの大きさは100gである。

このインクリメントの大きさや試長をどの位にとるかということは、最小の費用で所要精度をうるように対象物の性質、技術的条件、工程の変化状況などを考慮してきめられるもので、そのきめかたについては§7、§8でのべる。

母集団の大きさ(N)はインクリメントなどをサンプルの単位にとったと

きには、インクリメントの特性値は取り方により無限にかわりうるので、一般に無限母集団とみなした方が合理的である。して母集団の大きさ N に相当するものをもとめるときは

$$N = \frac{\text{母集団の全量}}{\text{インクリメントの大きさ}}$$

1.6 サンプリングとその目的など

ここで目的という言葉を用いるのは一寸不適當かも知れないが、サンプリングを必要とする場合の分類及びサンプリングによりもとの集団の何を知りたい場合があるかということについてのべよう。

われわれが工場において、サンプルをとりデータをとるのは、もとの製品の母集団や工程という母集団を知るためであるが、母集団のいかなる性質を知りたい——推定する——ためにサンプリングを行うかを分類して見ると

1. 母集団のある品質特性の分布状態を知るために、
これには多くの場合、平均値とバラツキを知れば十分な場合が多い。時には分布型のカタヨリ具合—歪度、ゆがみ (skewness) — やトガリかた—尖度、扁平度、とがり (peakedness, flatness)—を問題にすることもある。
2. 母集団のある品質特性の平均値や総計を知るために。
この場合が非常に多い。

サンプリングによりえたサンプルより母集団のある品質特性の平均値や総計を合理的に推定するためには、その分布状態を知らなければならぬので、1) は 2) の予備調査として行うことも多い。

また見かたをかえて工場を管理して行く立場にたつて、サンプリングをいかなるときに行うかを考えて見ると、

1. 管理図を作成するために。

通常工程を管理して行くために、工程を母集団と考えて、製品などのサンプルをとり、その測定値をいくつかの合理的な群へわけて、管理図を作成している。この場合のサンプリング法は、その管理の目的、工程の管理状態、群のつくりかたよりきまってくる。この場合は、特に目的によりサンプリング法が非常にかわってくるから注意を要する。本書でも少し触れ

るが、詳細は品質管理の参考書を参照されたい。

2. 受入検査をするために。

工場に原材料や半製品を受入れる場合に、あるいは前の工程から半製品を受入れるときに、サンプルングを行う場合で、原材料のロット (lot, batch) を合格にするか不合格にするかを決定したり、あるいはその品位を知って値段をきめたり、あるいは次の工程においてあらかじめ処置をとったりするために行うものである。この場合、合格品と不合格品の数や率によって合格、不合格を判定する場合、すなわち計数値、計量値による抜取検査については、他に多くの良書が出ているので本書では省略するが、この場合のサンプルング法としてはロットの概念をはっきりさせて、統計的な危険率などを決めて、あとは必要なサンプル数をランダムに、または一定間隔でとることになる。本書では、たとえば石炭、羊毛、鉄板、電球ボルト等が入ったときに、その灰分、洗上純毛率、硫黄や磷の含有率、寿命時間、不良数や不良率などの品位をもっとも経済的に、または正しく推定するためのサンプルング法についてのべる。

これらの受入検査について、従来わが国で混乱している1、2の点について一寸ふれておこう。

第1に受入検査について、仕様書、規格がはっきりきまっていない場合が多いことである。購入品や材料、半製品を受入れる際には、必ず、技術部や生産部などからの要求による品位がある筈である。特に統計的品質管理をおこなって行くと、必ずこのような仕様や社内規格が必要になってくる筈である。検査である以上、そこには必ず判定基準になる標準法がなければならぬ。従来あっても書かれたものになっていなかったり、たとえ書かれたものになっていても、皆にそれが徹底していなかったり、不合理なものであったりしている。これは資材の購入に際して、販売者と購入者が契約をする際にもあとでもしっかりした正しい行動をとるために、是非はっきりさせておくべきことで、販売側及び購入側両者の技術的、経済的事情を考えて互に協定した仕様書がなければならない。ところがこの仕様書がない場合、あるいはあっても不合理な、極めてアイマイなものが多い。この仕様書には必ず具体的なサンプルング法、測定法、標準値、及び標準値より過不足あった場合の処置などを、明確にきめておかななくてはならぬ。

い。たとえばある原料を購入するに際しては、いかなるサンプリングを行って、いかなる測定を行った結果が標準の品位と比較して、その原料を受入れるか、返却するか、値引するか、全数検査するか（そのときの費用をいずれでもつか）などの行動の基準をはっきりさせておくことである。これらのことがはっきり決められていなかったり、不合理であるために、いろいろ不明朗な問題を起している。

第2に受入仕様と設計仕様との混乱である。この両者は同じものではない。設計仕様によりはじめて受入仕様をつくるのが出来るのである。たとえば、設計課で設計仕様としてある材料が強度 110 kg/cm^2 以上あることが必要であると決められたとする。資材課では現在市場にあるその材料の強度のパラツキは標準偏差で 4 kg/cm^2 であること、及びA社からの製品は管理状態にあることも知っていたとする。そこで資材課は技術課と打ち合せて、不満足な品物が工場へ入ったときの損失、その材料の品位による価格の差、受入検査に要する費用などを総合的に考えて、たとえばサンプルを何個、どのようにしてとって平均品位 118 kg/cm^2 以上あるいは、 120 kg/cm^2 以上なら合格、そのときの危険率¹⁾とはどのように決定する。しかしこのように平均値を大きくすると非常に不経済であるという場合にはその材料の用途により、危険率（信頼度）や設計仕様をかえなければならぬ場合も起ってこよう。またB社では製品のバラツキが大きく、標準偏差 6 kg/cm^2 もあるようならば、B社については平均強度 122 kg/cm^2 以上とか、 125 kg/cm^2 以上としなければなるまい。この関係を図示すると図 1・2 のようになる。またC社では製品が管理状態になれば、サンプル数を余程増加しても、その結果を保証することは危険であろう。これらのことをいろいろと考慮して、受入仕様を決定し契約を結ぶのである。また各生産者により管理水準が異なっていれば、生産者別に、その管理状態に応じて、厳格、普通、緩和等の受入仕様を作成すべきである。このように受入仕様は設計仕様をある信頼度で、保証するように決定されるべきものである。よく資材課において、設計仕様をそのまま受入仕様に使っている例を見るが、これではわれわれがサンプリングにより検査をする以上、受入仕様で合格しても必ずしも設計仕様を満足しているものとはいえない。

1) この危険率は技術的条件とポリシーによりきめられるものである。

3. 出荷検査をするために.

2. と立場がことなるだけで考えかたはよく似ている. 工場内において各工程毎に生産者, 消費者という立場から考えると, 2. と同様にして仕様が決められ, それに従ってサンプルングや測定を行えばよい.

ここでも注意すべき1, 2の点につきのべよう.

第1に, 生産者は自分の製品が仕様にあっているということを保証する義務があり, 消費者は自分を保護するために, これをチェックする意味でサンプルング検査を行うということである. —わが国では生産者が製品を保証する義務, 責任の観念が, 特に販売関係者, 経営者にかけていることが多いのではあるまいか. 反省の要がある. —したがって消費者は, そのデータによる管理図と, 販売側の管理図とに有意な差がなく, しかも管理状態を示しているときには, 消費者の方はサンプルング回数を非常にへらすことが出来るということである.

第2に販売仕様と出荷検査の仕様とは, 別のものであるということである. これは設計仕様と受入仕様の関係とよく似ている. たとえばある薬品会社の仕様として, Aという薬品は各1錠に純分70%以上であるとして売り出しているときに, 各錠剤間には純分のバラツキがあるから, 何錠かの錠剤をとって, その平均純分が70%以上あったからといって, または10錠を別々に分析して, それがいずれも70%以上あるからといって, それだけでは全錠剤があるいはその真の平均値が純分70%以上あるという保証にはならない. ところが販売仕様そのままに出荷検査をしている場合が多い. 会社のポリシーとして販売仕様が決定すれば, 出荷仕様は, 工程の管理状態や, 製品のバラツキの大きさとにらみ合せて統計的に, 合理的にきめられなければならないものである. この場合平均純度70%という保証では, 消費者はその中どの位成分の悪いものが入っているかわからぬから, あまり保証の意味のない場合が多い.

いま設計仕様と受入仕様, 販売仕様と出荷仕様との関係を簡単に図示してみると次のようになる.

図1・2において, D.S. は設計仕様, S.S. は販売仕様を示し, たとえばD.S.の線以上のものを要求しているとする. ベル型は材料や製品の分布状況, 管理状況を示すものとする. a) の状態のものを受入れまたは出

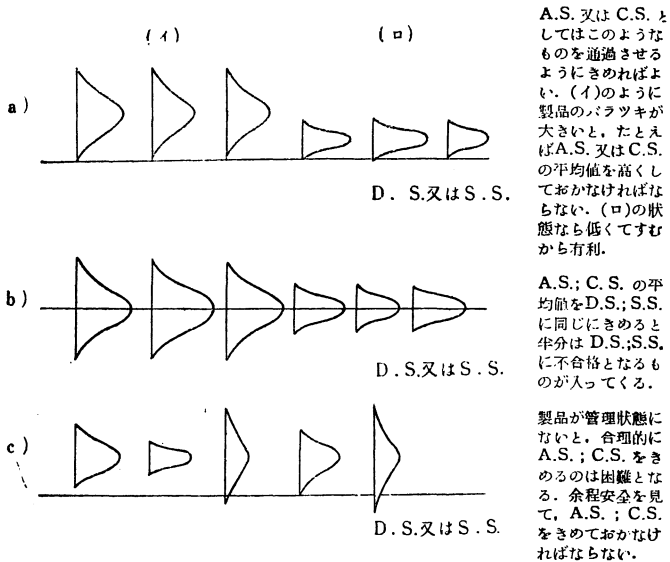


図1・2 設計仕様(D.S.)と受入仕様(A.S.);
販売仕様(S.S.)と出荷仕様(G.S.)
ベル型：つぎつぎと入ってくるロットの分布状況を示す。

荷するような検査方式をとっておけば、設計または消費者に対して製品を十分に保証することができる。この場合バラツキが大きいと平均値を大きくしておかなければならず、バラツキが小さくなれば平均値はあまり大きくなくてよいから経済的である。たとえばガソリンを5ガロンに各缶18リットル入っているということを販売契約したときに、充填作業が乱暴でバラツキが大きいときには、平均0.5リットル余分に入れるようにしておかなければ、どの缶も18リットル以上入っているという保証にはならないが、充填作業が管理されていて、バラツキが小さければ、入れ目は平均0.1リットルで十分保証できるようになる。これで判るようにバラツキを小さくすることは入れ目において非常に経済的になる。b)の状態のものであれば、半分は不合格となってしまうので、もしこれを基準に受入または出荷検査を行っていたとすれば不適當である。c)は生産工程が管理されてい

ない場合で、イ)の状態では設計仕様を満足していても、ロ)の状態のときには不合格品が相当に入る。したがって最悪の状態を想定して受入仕様をきめ、厳重な検査をしなければならず、これでもなお安心はできない。すなわち合理的な受入仕様をきめるには、生産者側の管理状態をよく監視して、仕様書やサンプリング法をきめなければならない。

4. 工場において検査や工場実験を行うために.

この場合にはその目的に応じて、サンプリング法を設計し、これに従って、サンプリングを行うことになる。本書にのべるいろいろな方法を理解されれば、自ら決定されると思う。

1.7 対象物による分類

サンプリング法を決定するとき、実際問題として最も問題になるのはその対象となる集団の状況や、その特性がどのようなものであるかということである。いまこれを統計的、技術的に考えて分類すると次のようにわけられる。

1. 単位体の場合 (discrete materials) (§ 6 参照)

電球、煉瓦、鉄板、ボルト類、錠剤を1粒ずつ考えたとき、ラジオや自動車を一車一台単位に考えたとき、容器に入った錠剤や粉末、または一束ずつにした糸のカセ、貨車で入った石炭(1台ずつの目方)、畝に入った肥料類のように、1個1個殆んど同じような単位¹⁾から成り立っている場合で、その単位体1個1個の品質特性を問題とし、測定や試験が行えるような場合である。これを更に次のように分類できる。

a) 定性的——計数的——な調査をする場合 (Sampling of attribute)

その単位体1個1個について良品であるか、不良品であるかの判定を行う場合または1個1個を1級品、2級品、3級品等とわかるような場合。いわゆる抜取検査法に出てくる大部分のものはこの型のものである。

b) 定量的な一計量的な調査をする場合 (Sampling of variables)

1個1個ではあるが、その単位体の長さ、重さ、純度、強度などのように量的な値で測られる場合。

2. 集合体の場合 (bulk materials)

1) 厳密な意味における同じものはないので、殆んど同じようなという表現を用いた。

1個1個別々に測定するのではなく、集合した状態のものから、その一部分の小さな集合部分を測定する場合で、この小さな集合部分の大きさは必要に応じていろいろの大きさにかえられる。

この場合は計数的な場合は殆んどなく、計量的な場合が殆んどである。したがってこの場合は、その集合体の状況によりいろいろに分類する。

a) 連続体の場合

i) 紡糸中の糸、圧伸中の線、製紙中の紙、などのように、固体が連続して動いていて、途中から切断してサンプルをとることが経済的に困難な場合。(§7.4)

ii) ボビンに巻いた糸とか、1束の針金や糸、1巻の紙やフープのような止っている連続体の中からサンプルをとる場合。(§7.4)

iii) 管の中、炉の中等を流れている流体、すなわちガス体及び液体を流れている途中からサンプリングするような場合、これは1)と状況はよく似ているが流体であるので自由にサンプリングできるところがことなっている。(§7.2)

iv) ガス体、液体等の流体で連続体ではあるが、タンク、ドラム管、ポンプ等の容器に入っている場合。(§7.2)

v) 不均一な流体の場合。たとえば微粒子を含んだガス、泥状体、不均一な溶液、粉状体、コロイド性の流体、などがこれに相当する。(§7.3)

b) 粉塊状体の場合。(§8)

石炭、鉱石、結晶等のように、塊状、粒状、及び粉状等各種の粒度のもの混合体の灰分、鉄分、不純分等を知ろうとする場合。

c) 羊毛、綿花等の繊維状体の集合体の場合。(§7.5)

梱包で入って来たあるいは工程中の羊毛や綿花からのサンプリングを考えるとき。

b) 及び c) はいずれも、仄毎とか、梱包毎とか、貨車毎、あるいは船倉毎、などのように、いくつかの単位体に分けられている場合には、その単位体毎の、平均純度や、成分を知りたい場合と、これら全体の平均値、あるいは一日の工程を流れる原材料などの平均純度などを知りたい場合とに分けられる。

3. 温度、圧力、流量、その他の測定値の場合。

この場合には、現実にはサンプルをとり測定をするという段階はなく、いつのデータをサンプリングするかという問題である。

4. その他

このような分類は必ずしも決定的なものではなく、その対象物の状態により、またその調査目的の品質特性に何をとるかにより、他の分類にうつることがある。たとえば、糸は紡糸中ポビンにまきとっているときは 2-a-i) であるが、ポビンにまきとってしまったものから糸を何米かサンプリングするときには、2-a-ii) であり、束にしてその級別をつけるときは 1-a であり、その一束の目方や長さを測るときには 1-b とみなせる。またガソリンが大きなタンクに入っているときは 2-a-iv) であり、パイプを流れているときにサンプリングするときには 2-a-iii) であり、ガロン缶につめたときにその缶毎の性質を問題にするときには 1-a または 1-b に属する。また硫安等も、臥に入っているときも、山に積んであるときも、その窒素分などを問題にするときには、2-b として取り扱うが、臥に入っている硫安の目方を問題にするときには 1-b として取り扱うべきである。石炭の灰分を問題にするときも中塊炭や切込炭ならば 2-b として取り扱うべきであるが、これを微粉碎して微粉炭としたときには 2-a-v) とみなせる。このように同じものでも、その状態の変化や調査の目的により、どの分類として取り扱うかは変化してくるから、サンプリングの計画をたてるときには、その目的、及び対象物の状況や特性をよく研究して、技術的、統計的な知識によりいかに取り扱うかを決定しなければならない。

1) 2) の各場合については、後章において具体的に論をすすめるが、3) の測定値のサンプリング法は大体連続体のサンプリングに類似しているのでその項を参照されたい。工場におけるその他のサンプリングとしてたとえば社員の健康状況や、家計状況調査のため、予算管理のため、経済上の統計データを早くまとめるためなどいろいろの調査のためのサンプリングの問題があるが、この点についても考えかたはよく似ているので、本書または標本調査法の手法を研究されれば自ら解決される問題と思う。市場調査を対象とした場合については、本書ではふれない。

第 2 章 ランダムサンプリング (Random Sampling)

2・1 ランダムサンプリングであるための条件

従来のサンプルという概念では、ややもすると、生産者、納入者は意識的あるいは無意識的に良い部分を選んでとるので、サンプルは良いが、実際に注文して製品が入荷して見ると、サンプルとは大分異なった悪い製品である場合が多い。日本の製品が海外で信用されないのは、このような点も大分影響していると思われる。また生産者、納入者によるサンプルの測定値と、消費者、購入者による測定値の食い違いにより、いろいろなトラブルや不明朗な事件をおこしている。われわれは母集団の品位を推定するためにサンプルをとるのであるから、このような有意なサンプリング法、意識を働かせたサンプリング法では不適當であることは明らかである。そこで、ランダムサンプリングを行うことが必要になってくるが、ランダムサンプルはいかに行うべきであろうか。これについては、サンプルのもつべき条件として、§1・5・1 でのべたが、これを抽象化したのべると次のようになる。

「母集団のどの部分も、その外観や位置などに関係なく、目的とする特性について、すべての部分が同じチャンスで、サンプル中にあらわれるようなサンプリング法でなければならない。いいかえると、いずれの部分も同じ確率でサンプル中にとられなければならない。」

しかしこのように母集団からランダムサンプルをとるということは、いはやすくして、実際に母集団全体から真にランダムにとるということは仲々厄介で、困難、あるいは不可能な場合がある。たとえば、倉庫に山積にした吹入の肥料から、サンプルをランダムにとることは非常に困難であり、サンプリングに多額の経費を要することになり、また何百トンも山積になっている石炭、1万トンも船積してあるバラ積の鉄鉱石から、そのままランダムサンプリングを行うことは不可能である。たとえ広い場所にひろげてある、あるいは運搬中の鉄石類でも、機械的にサンプリングでき

る場合にはまだランダムサンプリングできるが（しかし機械にも癖があることが多いことに注意）、人間がとることになると、どうしても選びやすい。たとえばランダムにとるのが厄介なので端や上の方からよい加減にとったり、その製品の生産に従事する現場関係者がサンプリングすると良さそうなところを選びやすい。私もある現場で、製品に影響をあたえると技術者が考えていた要因について、いろいろ有意差の有無の検定を行って見たが、どうしても有意な差があるとでてこなかったのに、組別に検討したところ、各組間に非常に有意な差があることがわかり、しかもその原因が、各組毎にサンプルの選び方に差があるためであったことを経験している。

更に具合の悪いことには、得たサンプルがランダムサンプルであるということを証明することが、実際問題として、仲々厄介であり、また不可能な場合もある。サンプルをとった人からランダムにとりましたといわれると、あとからでは手のつけられない場合が多い。

そこでこれらの困難を少なくするために、いろいろな注意や工夫が行われるのである。たとえば、母集団を技術的、統計的に考えて、いくつかの小さな集団にわけて、その小さい集団からランダムにサンプルをとるようにすれば、とりやすくなる。

しかしサンプリングの実施面の最後の段階では、すなわち工夫して小さい集団にわけた場合も、その各小さい集団の中からは必ずランダムサンプリングを行うことになる。

このように母集団からランダムサンプリングを行うと、必ずそのデータにはサンプリングによるバラツキが入ってくる。このバラツキを合理的に解析するのが数理統計学であり、サンプリングを統計的に考えなければならぬ理由はここにある。数理統計学では、サンプリングがランダムに行われているという前提条件があるので、ランダムサンプルの条件を満足してはじめて、そのデータを数理統計学の対象とすることができるのである。

2・2 ランダムサンプリングのやりかた

まず世界共通のサンプリングの原則をのべよう。

原則1. 「その製品の生産に従事したものにサンプリングさせてはいけない」

サンプリングはその生産に直接従事しない、たとえば検査員、分析員などに行わせるべきである。その製品を生産したものは、どうしても良いところをえらびやすい。したがってそのようなサンプルは、よい方へカタヨリを生ずる。このことは熟練工になる程、たとえ意識的でなくも、前におかれた二つの製品の中一つをとるときには、無意識的によいと思われる方へ手がのびてしまう。場合によっては、よい方をえらんでもよいこともあるが、どれがよいかという判断が主観的になるので、人により、時により、同じように真によいものをえらぶことは、熟練工でも困難である。したがってこのような有意サンプリング (purposive selection) はさけるべきである。有意サンプリングではサンプリングによるカタヨリを生じやすく、数理統計学の対象にもならず、したがって誤差もわからず、従来の経験からも間違いの多いことが証明されている。しかしこの方法によれば、時間や経費が非常に節約できるので、従来からよく用いられているが、例外的な場合を除いては、これは早急に中止しなければならない。

原則2. 「サンプリングを行うときには、必ず責任あるものが立会しなければならない」

サンプルはとってしまえば、あとでどのようにサンプリングしたか、不明になってしまう場合が多い。また検査員がサンプリングするときでも、悪いところをサンプリングしやすい。したがって、責任者が必ず立会ってサンプリング標準通りに、サンプリングを行っているかチェックしなければならない。またチェックできるようにサンプリング標準をきめておくことよい。従来わが国では、殆んど責任ある人が立会していないのではあるまいか。サンプリングの責任者の責任は正しいサンプリングが行われているか否かを管理することにある筈である。

原則3. 「試料採取関係者に、サンプリングの目的と重要性を十分徹底させなければならない」

合理的なサンプリング標準をつくり、いくらこれを教育しても、特にわが国においては、試料採取者がサンプリングの目的や重要性を認識せず、自分に都合のよいように、よい加減にサンプリングしている。この教育の問題が、わが国のサンプリングを合理的に改善する際にもっとも問題、障害になる点であろう。サンプリングの問題は人の問題であることが極めて多い。

原則4. 「サンプリングは、その対象物——ロット——の移動中に行うべきである。静止中に行うことはさけるべきである」

倉庫に山積になった製品や、瓶詰して包装し箱詰してしまった薬品類、山積にした石炭、船に積んでしまった、あるいは荷下し前の鉱石類などから、ランダムにサンプルをとることは、不可能か、あるいは非常に不経済である。液体の場合にはできる場合もあるが、固体の場合には、技術的に考えて不可能である。したがって、たとえば瓶詰したときに、石炭が入荷して荷下し中に、鉱石を荷積中にサンプリングすべきである。このことは特に石炭や鉱石類のように粉塊状物の混合物の場合に注意しなければならないことである。

以上の4原則はいずれも、ランダムサンプリングを確実にに行わせるための注意である。

しからばランダムサンプリングはいかに行うべきであろうか。次に2, 3の方法をのべよう。ここでは母集団全体からランダムサンプリングを行う方法についてのべる。これをいろいろに層別したり、集落化したりして工夫した場合については、§4においてのべるが、§4にのべるような方法を用いる場合でも、何れも最終段階のサンプリングを行うときには、ここにのべるような方法を用いなければならない。

- 1) 単純ランダムサンプリング (unrestricted random sampling)
 - 2) 系統的サンプリング (systematic sampling)
 - 3) ゼグザグサンプリング (zigzag sampling)
- 2), 3)はある意味において、1) の良い代用品だといってよい。

2.2.1 単純ランダムサンプリング (unrestricted random sampling)

調査しようという母集団全体から全くランダムにサンプリングする方法である。§4にのべるような工夫をしたサンプリングの場合でも、最終段階においては、ランダムサンプリングを行うので、これは特に何も制限をつけずに全体からのランダムサンプリングであるので、単純ランダムサンプリングと名付けている。単純ランダムサンプリングは母集団について、技術的にも、統計的にも、全く予備知識をもっていないときに用いられる方法である。予備知識を少しでももっているときには、これ以外の方法を用いた方がよい。

この方法はあくまで、面倒であっても、考えている母集団全体からランダムにとるのであって、たとえば、倉庫に入っている梱包した綿花があるときに、表面に出ている梱包からランダムに梱包をえらんだのでは、ランダムサンプリングとはいえないし¹⁾、山積になった石炭の表面だけから、あるいは1日の製品の流れの中から30分間にできた製品の中だけから、ランダムにとっても、そのサンプルは石炭山全体からのランダムサンプル、あるいは1日の製品全体からのランダムサンプルとはいえない。この点工業用原材料、製品の場合にはうっかりすると、取りやすい表面などからのみランダムにとっていることが多いから注意を要する。このような場合には、倉庫に入れてしまってから、山積にしてしまってからでは、ランダムサンプルをとることは経済的に不可能に近いので、原則4. にのべたように、移動中に、倉庫に入れるときに、山へ積むときにサンプリングしなければならないことになる。

ランダムサンプリングを行うには、次に §2.2.1 の1でのべるような乱数を利用するのが一般である。単位体の場合には、それに番号をうてるときには、番号を1から順にうち、番号をうつのが面倒あるいは困難なときには、ただ1, 2, ……とかぞえて、もとめた乱数に相当する番号の製品をサンプルとして抜き取る。このとき意識的あるいは無意識的なまちがいがないように注意しなければならない。たとえば1日に1,000コの製品が生産されるときに、その中から20コ製品をとるときには、乱数として、

162, 732, 853, 315, 990, 378, 767,
175, 438, 962, 224, 408, 138, 688,
374, 347, 600, 849, 049, 426

の20コの3桁の数字がえらばれたときには、はじめからかぞえて、この番に相当する製品をサンプルとしてとればよい。すなわち

049, 138, 162, 175, 224, 315, 347,
374, 378, 408, 426, 438, 600, 688,
732, 767, 849, 853, 962, 990

番目を間違なくサンプルとしてとる。

1) 倉庫への積みかたがランダムになつており、入れてすぐならランダムとみなせる場合もある。順に積んだ場合は多くの場合不可。

単位体で小さい部品が非常に沢山あるような場合には、番号をうつこともかぞえることも困難である。このようなときには、これを十分に攪拌するとか、特別なランダムサンプリングマシン¹⁾を作成するとよい。しかし完全に攪拌するということが非常に困難であるので、攪拌が十分であるということをチェックしておく必要があるが、ボルト、ナット、その他特に小さい多量の単位体や、流体、粉体などを常に取り扱うような場合に適当である。

連続体、粉塊状体の場合には、特殊な技術的考慮が必要な場合が多く、原則として移動中にサンプルをとらなければならぬ場合が多いが、一般には、移動中に、その移動する量に応じて、サンプリングの間隔を乱数により決定すればよい。たとえば、2,000トンの工業用塩を荷下し中にサンプリングするときに、インクリメント30コをとる必要のあるときは、2,000の中から30コの乱数をえらび、たとえば0151……と出れば、151トン目の中からランダムに1インクリメントをとればよい。もし荷下し量が一定量ずつ行われているときには、その荷下し時間中、たとえば10時間荷下しにかかるとすれば、600分であるから、600の中から30コの数字をランダムにえらび、その数字の時間に来た岩塩から、1インクリメントをランダムにとればよい。

しかし母集団全体からランダムサンプリングすることは、非常に面倒で間違いを起しやすく、サンプルの量も多量となるので、できるだけ無作為性を失わないように簡単に行える系統的サンプリング (§2.2.2) およびデックザグサンプリング (§2.2.3)、あるいはサンプル量が少なくてすむような工夫 (§4) が行われる。

実際、工場において、サンプリング法を計画し、設計する際には、母集団の特性、形状、経費、サンプリングの難易、その他すべての条件を技術的、統計的に考えて、決定しなければならない。場合によっては、経費を節約するために、無作為化をある程度無視して、誤差のある程度の増加を認めなければならぬ場合もある。

1. 乱数のつくりかた ランダムサンプリングを行うには乱数を必要

1) たとえば F. L. Fay (John Tractor Co.) Purchased Part Quality Control ; American Machinist Dec. (1947) p. 98.

とする。

乱数 (random number) とは、数字が無秩序に並んでいるもので、もとめたい範囲の数が、同じ確率でえられるような数列をいう。たとえば、6コの数の乱数は正しいサイコロを、ふってえられた順にならべると、2, 6, 3, 4, 1, 5となった、これは乱数である。

この乱数のつくりかたには、乱数表、サイコロ型式、乱数器式などいろいろあるのでこれについてのべよう。

これらいろいろの方法の中、その現場にもっとも適した方法を使うようにしないと、たとえば折角乱数表が現場に渡してあっても、その使い方が面倒なために、かえて使えぬ、あるいは使わぬ場合があるから、技術者が相当具体的に、間違の少ない、使いやすい方法を、場合によっては、とるべきサンプルの番号を示してやらなければならない。

2. 乱数表 (tables of random sampling numbers) 乱数表とは、1桁、2桁、3桁……等の数が同じような確率で無秩序に現われるように、0から9までの数字を並べたもので、すでに次のようなものが発表されている。本書の巻末の附表5に、その一例を示す。

Fisher-Yates の乱数表 (数字の数 15,000コ)

R. A. Fisher and F. Yates; "Statistical Tables for biological, agricultural and medical research" Oliver and Boyd. Edinburgh 1938.

Tippett の乱数表 (数字の数 41,600コ)

L. H. C. Tippett; "Random Sampling Number" University of London Tracts for computers No. 15. Cambridge Univ. Press London 1927.

Kendall-Smith の乱数表 (数字の数 100,000コ)

M. G. Kendall and B. B. Smith; "Tables of Random Sampling Numbers" University of London Tracts for computers No. 24. Cambridge Univ. Press. London 1939.

Snedecor の乱数表 (数字の数 10,000コ)

G. W. Snedecor; "Statistical Methods" Iowa State College Press. Iowa U. S. A. 4th ed. table 1, 2. 1946.

これらの乱数表はわが国の多くの統計学の本に掲載されているが、Fisher-Yates のものが比較的多い。しかし実用的にはいずれを用いても大差

ない。

乱数表の使いかたには、いろいろあるが、ここでは単純ランダムサンプリングを行う場合だけをのべる。

例 1) 87コの製品からサンプル5コをランダムに抜きとる方法。

i) 乱数表の何頁を用いるかをランダムに選ぶ。附表5の乱数表の何頁を用いるかをサイコロでえらぶ。この場合3頁あるから1, 4, が出たら1頁, 2, 5, が出たら2頁, 3, 6, が出たら3頁をえらぶことにしておく。4の目が出たので1頁をえらぶ。乱数表は全体として数字の並びかたがランダムになっているので、1頁だけを常に用いるとカタヨリを生ずる。行列についても同じ。

ii) 1頁の4つの群の中いずれを用いるかをランダムにきめる。サイコロで4の目までを群に相当させ、5, 6, が出たら捨てることにする。2の目が出たので、左から2群目をえらぶ。

iii) 2群中5列の2桁の数字の中、いずれを用いるかをランダムにきめる。5の目が出たので、最右列の2桁の数字を用いることにする。

iv) この場合、対象物の数が87コであるから、1から87までの数字の中、この列を上からとって、5コえらぶ。この場合

61, 37, 46, 77, 26,

すなわち、26番目、37番目、46番目、61番目、77番目の製品をサンプルとしてとればランダムサンプルとなる。

注意 1) 00 及び 88~99 の数字が出て来たたら、これは捨てる。

もしロットの大きさ100のときは、00は100とする。

2) 01~87の数字が2回以上出て来たたら、これは捨てる。

3) 頁, 群, 行列などはランダムにえらぶ。

4) もし1列で不足の際には、次の列に引続き移って差支えない。ただしはじめから次の列へ移るといふ原則を決めておかねばならない。または別の列をランダムにえらぶ。

例 2) 30コの製品から10コをランダムに抜きとる方法。

i) 例 1) と同様に頁, 群, 列をランダムにえらんだら、2頁、第1群の第3列となった。

ii) 例 1) と同様に行えば、

80, 78, (16), 40, 60, 76, (22), 59, 39, 40, 50,
 (06), 44, 32, (13), 44, (07), 18, (24), 90, 74, (08),
 72, (02), 94, 38, (27), 18, 57, 06, 87, (21)

すなわち

2, 6, 7, 8, 13, 16, 21, 22, 24, 27

の10コをサンプルとしてとればよい。

このようなえらびかたをすると、非常に多くの数字を拾い出さなければならぬ。そこで次のような方法を用いる。

iii) 1~30, 31~60, 61~90, と30の n 倍を考え, 1, 31, 61, が出たら1番目, 2, 32, 62, が出たら2番目をとることにする。前例と同様に00および91~99が出たら捨てる。その他例1)の注意に同じ。

60	76	16	40	00	76	(22)	59	39	40	50
↓	↓		↓				↓	↓		
(30)	(16)		(10)				(29)	(9)		
(06)	44	32	(13)							
	↓	↓								
	(14)	(2)								

すなわち, 2, 6, 9, 10, 13, 14, 16, 22, 29, 30, をとればよい。

注意 1) この方法は計算は厄介であるが数列を引く数は少なくすむので, ロットの大きさの最初の桁の数が小さく, サンプルの大きさが大きいときには便利である。たとえば159コから20コをとるとき, 例1)の方法によると附表乱数表2頁の最初の3桁をつかっても, 118, 155, 136, 104, と4コの数字しかえらべぬ。しかしこのようなときに上の方法をそのまま用いると計算が厄介であるから1~129, 201~329, 401~529, 601~729, 801~829, をそれぞれ1~129に対応させて用いると簡単でよい。この例では, 同じ列を用いて

77, 34, 13, 107, 24, 118, 101, 24,
 78, 67, 73, 129, 101, 36, 86, 49,
 103, 78, 21, 35, 32, 92,

となり、比較的簡単にもとめられる。

2) 乱数表の使いかたには、このように列のみを用いず、行を用いたり、斜めに用いたり、いろいろな使いかたがある。

3. サイコロ型式のもの サイコロは、小型で、廉価であるので、簡単な乱数をうるには便利である。しかしサイコロに癖があったり、振りかたに癖があることが多いから、サイコロの選択、使用法には注意しなければならない。

a) 通常の1～6の目の正6面体のサイコロ 6までの乱数をうるのに便利である。あるいは2つの中1つをえらぶというようなサンプリングのときは、偶数、奇数によりわければよい。使いかたについては、1の乱数表のえらびかたの例題をみられたい。しかし2コ以上を振って、その和を用いるという方法は、この和の出る確率がことなる、すなわち2項分布あるいは多項分布をなすので、不適當である。1コずつ別々に、独立に用いなければならぬ。

b) 正20面体サイコロ (東芝石田保士氏考案) 正20面体は、正三角形がちょうど20面あるので、これに0から9までの数字を二つずつ記入できる。したがってこのサイコロを用いれば、1コ振れば、1桁の乱数、2回あるいは色のちがうものを2コ振れば2桁の乱数がえられる。現場で用いるには便利なものである。これを乱数サイともいう。

注意 1) 正20面体の各面にカタヨリがないように作成することは困難であるから、目のうちかたを統計的に検討した方がよい。

2) 10までの数字をえらぶとき、0が出れば10とする。100までのときは、2コが00と出たら100とする。これは乱数表を使用するときと同様である。

4. 乱数器 最近工業における数理統計学の応用にしたがって、いろいろな乱数器が考案されている。小型のものもあり、カタヨリも比較的少なく、便利なものが出来ている。カタヨリはむしろサイコロの場合よりも少ない。

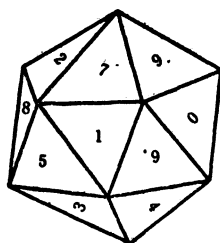


図 2.1
正20面体サイコロ

オミクジ式、回転盤式、風車式などいろいろある。宝クジなどの当りクジの番号も、一つの乱数である。一般の乱数器は多くの場合、その数の範囲は0～9、オミクジ式のものでも00～99までである。その使いかたは、正20面体および乱数表の場合と同じ。

2・2・2 系統的サンプリング (systematic sampling)

母集団全体から、全くランダムに、単純ランダムサンプリングを、実際に工場で行うのは、困難あるいは厄介な場合が多い。このときには、サンプルを、時間的、あるいは空間的に、一定間隔でサンプリングすればよい場合が多い。

一定間隔でサンプリングする方法を系統的サンプリングという。

この方法は単純無作為サンプリングに比し、間違いも少なく、簡単であるので、従来から工業において多く用いられている。

1時間毎にサンプルをとるとか、30袋毎にサンプルをとるとか、いうのは、すべてこのサンプリング法である。

母集団が N コからなり、サンプルの大きさ n コとすると

$$a \equiv \frac{n}{N} \quad (2 \cdot 1)$$

を**抜取比 (sampling ratio)** といい、その逆数

$$\frac{1}{a} \equiv \frac{N}{n} \quad (2 \cdot 2)$$

を**サンプリング間隔 (sampling interval, 抜取間隔)** という。

したがって、系統的サンプリングでは、製品 $1/a$ コごとにサンプルをとることになる。

注意 1) $1/a \equiv N/n$ において、 N を n で割り切れぬときは剰余ができる。この剰余による、カタヨリなども考えられるが、われわれが工場で接するように、 N が n に対して、大きければ、あまり問題にしなくともよい。最後の間隔にとるべき番号が入っていれば、それととり、入っていなければ最後の間隔からはとらないこともある。

2) ランダムネスをできるだけ確保するために、ランダムスタートですなわちはじめの1コはその間隔の中でランダムに選ばなければならない。

1. 系統的サンプリングのやりかた 単位体の場合には2つのやりかたがある。

第1の方法は、単位体を(なるべくランダムに)ならべて、その中から一定間隔でサンプルをとって行く方法で、たとえば抜取比 1/10 の場合には、図2.2のように各行に順にならべて、その中1列をえらぶ。どの列をえらぶかはランダムに、たとえば乱数表で行う。今乱数表より6が出れば、6列をサンプルとしてとる。

第2の方法は、単位体に番号をうち、(なるべくランダムに¹⁾)、その番号を一定間隔でサンプルにとる。たとえば、抜取比 1/15 のときには、

1, 2, 3, 4, 5, 6, 7, 8, 9,
10, 11, 12, 13, 14, 15, 16, 17,
18, 19, 20, 21, ……

と番号をうち、乱数表より1から15までの数をひき、たとえば07を得たら $7, 15+7=22, 37, 52, 67, ……$ 番目をサンプルとしてとる。番号はなるべくランダムに打った方がよいが、通常は生産順、出荷順に打ってかまわない。たとえば荷下し中の綿花の梱包からサ

ンプリングするには、抜取比 1/50 ならば、ランダムスタートをするためにまず50までの乱数をひき、たとえば27と出たらば、荷下し中、梱包の数を数えながら、27, 77, 127, ……番目をとればよい。

単位体以外の場合、および単位体でも数が非常に多いときには、上にのべたような方法を厳密に実施することは困難となるので、大体一定量ごと、あるいは大体一定時間ごとにサンプルをとって行けばよい。たとえば、工程管理(たとえば管理図)を行うときなどに通常よくやっているように1

1) 完全にランダムに番号を打てば、単純ランダムサンプリングと全く同じ結果となる。

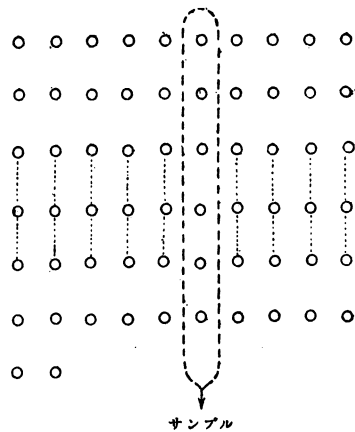


図 2.2

系統的サンプリングのやりかた

時間ごとにそのときの製品をサンプルにとるとか、荷下し中の鉱石をクレーンで5杯ごとに1インクリメントをとるなどこれである。

しかし系統的サンプリングでは、カタヨリの入る恐れがあるので、次のべるようなことに注意しなければならない。

2. 系統的サンプリング実施上の注意 系統的サンプリングは、実施が簡単であり、容易であるので工業において、多く用いられているが、単純ランダムサンプリングに比し、

- 1) カタヨリが入りやすい。
- 2) サンプル数の少なくてすむことが多い。
- 3) サンプル数を多く必要とする場合もある。しかしこれは非常に特殊な場合である。

以上のことは、対象物の問題としている特性の時間的あるいは位置的、空間的の変化、特に周期性、傾向により、起ってくる問題である。

i) 特性の変化が全くランダムであるとき。

品質管理において、真の管理状態に達したときはこの状態であるが、このときには、一定間隔でとったサンプルは、単純ランダムサンプリングのときと全く同様になる(図2・3a参照)。

ii) 特性の変化に周期性があっても、その周期がサンプリング間隔よりも短い場合¹⁾。

この場合にも、一定間隔でとったサンプルは大體ランダムサンプルとみなしてよい(図2・3b, c参照)。

iii) 特性の変化に周期性があり、その周期がサンプリング間隔に等しい場合。

この場合には、サンプル間のバラツキは小さくなるが、カタヨリが入ってくる。系統的サンプリングを行うときにもっとも注意しなければならない場合である。したがってサンプリング間隔をかえなければならない(図2・3d参照)。

たとえば、工程に1時間、半日、1日などの周期があるときに、1時間ごと、半日ごと、1日ごとにサンプルをとっているとカタヨリが入る。あ

1) ただし整数分の1になっていないこと。厳密にいうと位相のことも考えなければならぬが、実用的には考える必要のない場合が多い。

るいは毎朝8時にサンプルをとるとか、毎時ごとにサンプルをとるとか、きめておくと、現場ではそのときの製品は、一生懸命よいものをつくるとか、製品の検査を、容器の上のものばかりサンプリングすることがわかっていると、容器の上の方によいものをえらんでいれておくなどという場合

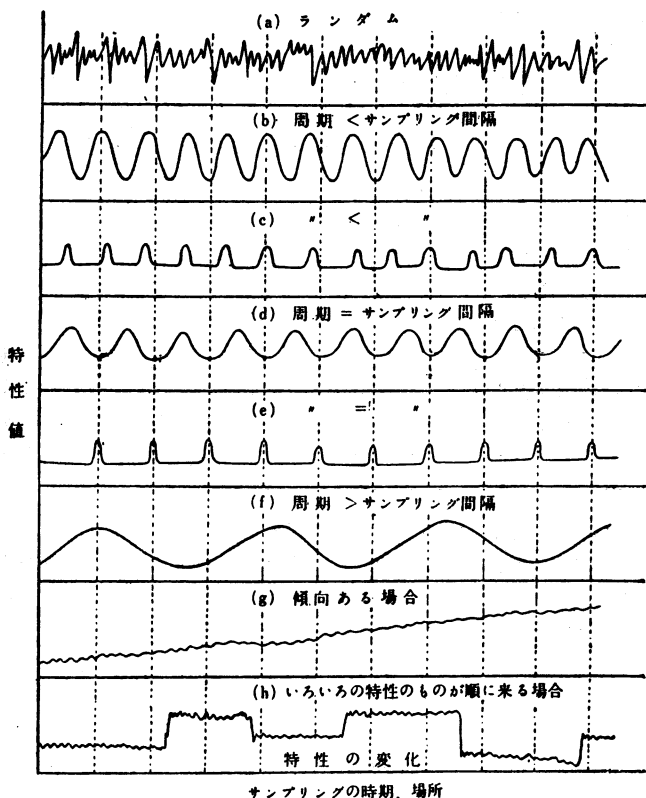


図2.3 系統的サンプリングと特性の変化の関係
点線はサンプリングの時期、場所を示す

もこれであり、よい製品のみをサンプルとしてとってしまうことになる。あるいは糸のコーンの端の方はいつも悪くなる癖のあるときに、いつもコーンの端からだけサンプルをとっていると結果は真の値より悪い方へカタヨってしまう(図2・3e参照)。

iv) 特性の変化の周期がサンプリング間隔より長い場合(図2・3f参照)。

v) 特性の変化に傾向のある場合(図2・3g参照)。

vi) 特に異なった特性の変化のある場合(図2・3h参照)。

iv) v) vi) の場合は単純ランダムサンプリングにはならず、後にのべるような層別サンプリング、あるいは副次サンプリングの効果を発揮して、平均値を推定する場合の精度はかえってよくなる。カタヨリも殆んど入らず系統的サンプリングの効果を発揮する場合が多い。

たとえば、iv) は化学工業における連続反応工程のように、大波のある工程などからサンプリングするとき、v) は船積して来た粉塊状の混合体を荷上げするとき、はじめは塊が多く、底に行くにしたがって粉が多くなり品質が次第に悪くなってゆくと、vi) は別のロットと見なされるものが次々とながれてくる場合で、たとえば船積の鉱石が各ハッチごとに相当品質の異なったものが入っているときに、これを次々と順に荷下しするとき、あるいは、1, 2, 3号炉があるときに、製品がはじめは1号炉から、次に2号炉から、最後に3号炉から出てくるときに、それにかまわず一定間隔でサンプリングしているときがこれに相当する。

この場合の数理や実例は、§4においてのべる。

系統的サンプリングを行ったために、かえって多くのサンプルを必要とする場合は、サンプリングの周期がちょうど、特性のパラッキの大きい時期に相当した場合で、たとえば中塊炭を貨車からおろしはじめだけからサンプルをとったために、塊の多い部分を取り、インクリメント間のパラッキが大きくなったり、ポビンの端は紡糸のはじめにあたるために、デニールラが特に大きくなるような場合がこれで、この場合は多く、カタヨリが入っており、技術的に考えて、合理的なサンプリングにかえなければならぬ場合が多い。

以上でわかるように、系統的サンプリングを行う際には、特にカタヨリが入ることがあるから特性の変化に対する予備知識が必要で、その周期性

に注意しなければならない。これらの周期性をさけるためには、はじめに番号をうつときには、順をなるべくランダムにうっておけばよい。

場合によっては、技術的に周期性をさけることが困難な場合がある。たとえば、金属製品の成分を知るためのサンプリングのように、破壊的であるのでサンプルの採取場所に非常な制限をうけ、あるきまったところからサンプルをとるために、——これも一種の系統的サンプリングである——カタヨリを生じやすい。あるいはポビン、巻き取った紙、フープなどではその中からランダムにとると、その1巻を無駄にしてしまうので、サンプルは先端からとらなければならぬこともある。

このように系統的なカタヨリが必然的に入って来るような場合には、あらかじめ予備実験によりカタヨリの大きさを相関関係などよりもとめておき、カタヨったサンプルの結果から、予備知識を利用して全体の平均値などを推定する場合もある。

2.2.3 チックザグサンプリング (zigzag sampling)

系統的サンプリングの周期性によるカタヨリなどの入る欠点を少なくするために考案された方法で、この方がランダムサンプリングと同じと見なせる場合が多い。しかし実際のサンプルのとりかたは、系統的サンプリング法よりも稍厄介である。やりかたは

1. 系統的サンプリングのときと同様に番号をうつ。

たとえば採取比 $1/7$ とすると、

1, ②, 3, 4, 5, 6, 7, | 8, 9, 10, 11, 12, ⑬,
14, | 15, ⑯, 17, 18, 19, 20, 21, | 22, ……………

はじめは乱数表により、ランダムスタートではじめる。2と出れば、2, $2 \times 7 - 1 = 13$, $2 \times 7 + 2 = 16$, $4 \times 7 - 1 = 27$, ……ととって行く。

2. 番号のうちかたを、たとえば製造順に1, 2, 3, 4, 5, 6, 7とはじめの採取間隔では同じだが、次に来たものから逆に14, 13, 12, 11, 10, 9, 8, とうつ。3番目の採取間隔では、普通に15, 16, ……21とうち、番号のうちかたを順逆交互に繰り返す。どれをとるかはランダムスタートで、たとえば5がでれば、5, $7 + 5 = 12$, $2 \times 7 + 5 = 19$, ……と番号は系統的にとって行く。

1, 2, 3, 4, ⑤, 6, 7, | 14, 13, ⑱, 11, 10, 9, 8, | 15,

16, 17, 18, ⑩, 20, 21, | 28, 27, ⑳, …… …

このようなとりかたをすると、とる番号はきまってくるので、一々乱数をもとめる必要がなく、比較的規則的にサンプルがとれる。サンプリング間隔は、たとえば 1) の場合には、 $10-4-10-4$ …… となり一定でないために、工程や製品の周期性によるカタヨリをさけることができ、層別サンプリングの一種と考えるときも、各層（抜取間隔）内でのとる時期が一定でないので、系統的サンプリングよりも安心して使用することができ。しかし、周期性に対する注意はわすれてはならない。

系統的サンプリングよりも、そのとりかたが稍厄介であるが、この方法が容易に行えるときには、カタヨリのない場合が多いから用いるとよい。

注意 1) 上の例で、はじめに 1, 4, 7, をとると系統的サンプリングと同じようになる。

2) その他の注意は、系統的サンプリングの項参照。

2.3 ランダムサンプリングの数理¹⁾

ランダムサンプリングの数理は、数理統計学入門そのものであるが、ここでは簡単にのべ、詳細な点は数理統計学の専門書を参照されたい。

2.3.1 データの性質、統計量、母数

われわれがある母集団、たとえば石炭の山から 1 スコップずつランダムにサンプルをとり、その灰分を分析したデータを見れば、それは必ずバラツキをもっている。これは母集団を 1 スコップずつにわけて考えたときにバラツキをもっており、それからのサンプルであるので、バラツキが入ってくるのである。同様に、石炭山のあちこちから 30 杯ずつランダムにサンプルをとり一緒にして分析したデータも、これを繰り返すと、やはりバラツキをもっている。また 1 スコップの石炭から分析サンプルをとる際にも、バラツキが入ってくる。すなわちわれわれが生産するものは、科学的に考えられるいろいろな原因により、その品質特性にバラツキをもっており、それを更にサンプリングするのであるから、われわれが得たサンプルについてのデータは必ずバラツキをもっている。この必ずバラツクという事実があるので、われわれはサンプリングを、数理統計的に考えなければな

1) 数理統計学をある程度御存知の方は本章の注意或は備考以外は飛ばして読んで戴いてよい。

らないのである。また同じ製品を何回か繰り返し精密に測定したり分析したりすると、必ずそのデータもバラツクものである。したがって測定や分析誤差なども統計的な考えかたで見なければならない。

たとえば、2つの石炭のロットからサンプルを各々とり分析した結果が8.62%と8.76%と出た。この場合に前者の方が灰分が少ないといえるであろうか。従来のようにサンプリングの誤差のはっきりしていない方法では、何ともいえないであろう。

従来われわれは何でもデータを平均して、その平均値の差だけを主観的直観的に比較していた。平均値の世界でデータを見たり、サンプリングを考えていた場合が多い。

表2.1に示したデータは、石炭を同じロットから100回繰り返しサンプリングしたときの灰分の分析値である。これだけではデータがバラツキについてそのサンプリング法について何もわからぬが、これを表2.2のようにある一定の間隔(級の巾, class interval)でいくつかの級(class, cell)にわけて、その中に入るデータの度数をチェックして表にすると、データが大体どのように分布しているかがわかる。この表を度数表(frequency table)という。

表 2.1 石炭の灰分 (%)

8.42	8.03	8.16	7.79	8.70	8.32	8.34	7.81
8.07	7.80	7.96	8.69	9.60	8.74	7.63	8.02
8.66	8.06	8.04	7.30	8.20	8.31	8.00	8.28
8.07	8.52	8.35	8.77	8.99	8.27	8.30	7.66
8.93	8.52	7.89	8.20	8.86	7.57	9.60	8.56
9.05	7.63	8.04	8.31	8.71	8.90	8.81	7.94
7.82	8.04	8.15	8.41	9.10	8.05	9.56	8.31
9.40	7.79	8.09	8.26	8.44	7.98	8.41	8.02
8.77	7.57	8.04	8.22	8.60	9.11	8.02	7.91
8.22	7.92	8.35	7.98	7.96	8.17	8.71	8.05
8.46	8.01	8.50	8.14	8.31	8.91	8.60	8.17
8.54	7.57	9.68	7.97	8.01	7.40	7.98	7.49
7.85	8.70	8.21	7.88				
max. 9.68				min. 7.30			

表 2・2 石炭灰分の度数表

級の番号	級の限界(%)	中心値(%)	チエック	度数 f_i
1	7.255~7.455	7.355	/	1
2	7.455~7.655	7.555	+++ //	7
3	7.655~7.855	7.755	+++ //	7
4	7.855~8.055	7.955	+++ +++ +++ +++ ////	24
5	8.055~8.255	8.155	+++ +++ ////	14
6	8.255~8.455	8.355	+++ +++ +++ /	16
7	8.455~8.655	8.555	+++ ///	8
8	8.655~8.855	8.755	+++ +++	10
9	8.855~9.055	8.955	+++ /	6
10	9.055~9.255	9.155	//	2
11	9.255~9.455	9.355	/	1
12	9.455~9.655	9.555	///	3
13	9.655~9.855	9.755	/	1
計				100

度数表を図にあらわしたものが図 2・4 である。この図を柱状図¹⁾(ヒストグラム, histogram)という。この柱の頭の中点を結んでえた多角形を度数多角形(frequency polygon)という。データをこのような表や図にすると、サンプリングしたときのデータのバラツキ状況がよくわかる。このデータの値と、そのあらわれる度数のバラツキの関係を、**度数分布**(frequency distribution)という。以上のことは、一定個数のデータについて行ったのであるが、これではデータの総数により、各柱の高さがかわり、形もかわってくるので、各級に入るデータの度数を、全体のデータの数で割って**相対度数**(relative frequency)とし、**相対度数分布**であらわすことがある。この相対度数分布において、データの数を無限に増加し、級の巾をせまくして行くと、相対度数分布を示す度数多角形は曲線型となる。この曲

1) 度数表およびヒストグラムの縦わけの仕方にはいろいろな注意が必要である。数理統計学の書籍参照。

線を度数分布曲線 (frequency distribution curve) あるいは単に分布曲線という。

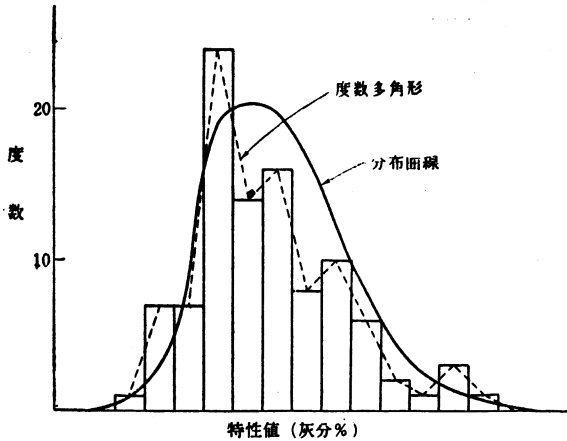


図 2・4 ヒストグラム 石炭サンプリングのデータ

度数分布をあらわす方法として、以上の方法のほかに、ある値以下の値の度数の分布をもって示すことがある。たとえば灰分7.45, 7.65, …%以下の度数がいくつあるかを示すと便利なことがある。この度数分布を累積度数分布 (cumulative frequency distribution) という。累積(度数)分布

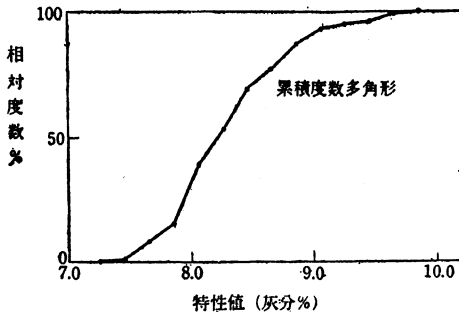


図 2・5 累積度数分布 石炭サンプリングのデータ

においても、相対累積度数分布、累積度数多角形、累積度数曲線 (cumulative frequency curve または og ve [ógudʒaiv]) などがある。

これらの分布曲線は、そのサンプリング法の性能、一般にはデータの分布状況をよく示すものである。このような分布曲線にしなくとも、ヒストグラムを見れば、多くの場合十分である。このようにわれわれのデータはバラツキをもっている以上必ずそこに分布というものが考えられる。図を見ればそのデータの分布状況は大体わかるが、これをもっとまとめた形で数的にあらわす方法を考えてみよう。

この山の型をもった分布をあらわすには、いろいろな量、測度 (measure) を用いる。

- 1) 中心がどの辺にあるかを示す測度…………… 中心的傾向 (central tendency)
- 2) 山の巾、データのバラツキの大きさを示す測度…………… バラツキ (dispersion)
- 3) 分布のスソが、すなわち分布の山がどのくらい非対称になっているかを示す測度…………… ゆがみ (歪度, skewness)
- 4) 山のとがり具合、すなわち扁平さを示す測度…………… とがり (尖度, peakedness, 平らたさ, flatness)

以上4種類の測度がきまれば、分布型を殆んど完全にあらわすことができる。しかしサンプリングにおいては、大部分の場合、中心的傾向とバラツキの測度を知れば十分な場合が多く、特別な場合にゆがみまで考えればよい。しかしわれわれはサンプルの値からこれらの値をもとめるので、サンプリングによるバラツキを考えると、ゆがみをもとめるには、少なくとも100コ以上、できれば250コ以上のデータがなければあまり意味のない場合が多い。

サンプルからもとめたこれらの測度を統計量 (statistics) という。母集団についてのこれらの測度を母数 (parameter) という。中心的傾向の測度をあらわすには、平均値、メジアン、モード、中点値、幾何平均などがある。この中算術平均 (平均値 arithmetic mean, mean) がもっともよく用いられる。これは測定値を全部加えて、全測定個数で割ったもので、サンプルについては一般に次の式であらわされる。

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_i + \cdots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} \quad (2.3)$$

ただし \bar{x} : サンプルの平均値 (試料平均) = 統計量

x_i : 各測定値

n : 測定値の数 (サンプルの大きさ)

母集団についても同様に

$$\mu = \frac{x_1 + x_2 + \cdots + x_i + \cdots + x_N}{N} = \frac{\sum_{i=1}^N x_i}{N} \quad (2.4)$$

ただし μ : 母集団の平均値 (母平均) = 母数

N : 母集団の大きさ

注意 算術平均はサンプリングの際にもっともよく用いられ、これより母平均を知ろうという場合が多い。算術平均の計算は (2.3) (2.4) 式により簡単に行えるように見えるが、数理統計学では、多くの数値計算を行うので、計算をできるだけ簡単に、しかも間違いなく行えるように計算方式を工夫することが常に行われる。

計算簡単化の方針は、1) 取り扱う桁数をできるだけ小さくする、2) 小数点あるいは最後に0の多くつくときはこれを取る。このためには、

1) まず技術的、統計的に考えて、不必要な数字は4捨5入する。

2) データーから大体平均値に近い一定数 a (これを假りの平均といふ) を引いて計算し、あとで加える。

$$\bar{x} = a + \frac{\sum (x_i - a)}{n} \quad (2.5)$$

たとえば、窒素分が 20.96, 20.89, ……とあれば、20 を引いて、0.96, 0.89, ……として計算し、あとから20を加えてやる。

3) 10^n (一般には h) を掛けて計算し、最後に 10^{-n} (一般には h^{-1}) を掛ける。

$$\bar{x} = \frac{1}{10^n} \sum \frac{10^n x_i}{n} \quad (2.6)$$

たとえば 0.96, 0.89, ……を100倍して 96, 89 として計算し、あとで100で割る。2) 3) を組合せると

$$\bar{x} = a + \frac{1}{h} \frac{\sum (x_i - a)h}{n} \quad (2.7)$$

上の例では20を引き、100倍して計算し、あとで100で割り、20を加える。

4) 同じ値のデータが多くでてくるときは

$$\bar{x} = \frac{\sum x_i f_i}{n} = a + \frac{1}{h} \frac{\sum (x_i - a)h f_i}{n} \quad (2.8)$$

5) データの数が非常に多いときには、度数表をつくり、各級に入る値の平均が、その級の中心値に等しいとみなして簡略算を行う(表2.3参照)。

以上のことは、母集団についても全く同様に行える。

メジアン (中央値 median) 測定値を最大値から最小値まで順にならべたときに、 n コの測定値がそれより大きい値の数が $n/2$ 、小さい値の数が $n/2$ にちょうど2分されるような値。

モード (最多値 mode) 度数の極大となる値、すなわち分布曲線の山の極大となる値、通常は1つであるが、2つ以上のモードを持つ分布もある。

バラツキの測度としては、偏差二乗和、分散、標準偏差、範囲、変動係数などがよく用いられる。

偏差二乗和 (平方和 sum of squares) 各測定値と平均値との差(偏差 deviation)の二乗の和である。二乗和を $S.S.$ とするとサンプルについては

$$S.S. = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 = \sum (x_i - \bar{x})^2 \quad (2.9)$$

$$= \sum x_i^2 - n\bar{x}^2 = \sum x_i^2 - \frac{T^2}{n} \quad (2.10)$$

ただし T : 測定値の総和 $T = \sum x_i$

計算には、(2.10)式の方が便利である。二乗和は、測定値の数により、大きくなったり、小さくなったりするので、バラツキの大きさを比較するのに不便であるので、分布のバラツキの大きさをあらわすには通常分散、標準偏差などを用いる。

分散 (variance) 偏差二乗の平均値、すなわちサンプル分散(試料分

散) を s^2 とすると

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{S.S.}{n} \quad (2.11)$$

母集団については、母分散 σ^2 は (または V , $\text{Var. } x$ と書く)

$$\sigma^2 \equiv V = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N} \quad (2.12)$$

標準偏差 (standard deviation) 分散の平方根。分散は、単位が測定値の二乗になっているので、不便なことがあるので、測定値と同じ単位の標準偏差を用いる。サンプルの標準偏差 (試料標準偏差) を s とすると

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} = \sqrt{\frac{S.S.}{n}} = \sqrt{s^2} \quad (2.13)$$

母集団の標準偏差 (母標準偏差) を σ とすると

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}} = \sqrt{V} \quad (2.14)$$

注意 偏差二乗和、分散、標準偏差の計算も相当厄介であり、大きい数字になるので、簡単化を行って計算しないと、間違いも多く、時間的にも不経済である。簡単化の原則は

- 1) まず技術的、統計的に考えて unnecessary な数字は 4 捨 5 入する。
- 2) 平均値をもとめるのは厄介であるので、(2.9) 式より (2.10) 式を用いる。
- 3) データから大体平均値に近い値 a (分散などの計算の際には特にできるだけ平均値に近い値の方がよい) を引いて計算する。この場合は偏差をとっているのであるから、後で加えてやる必要はない。

$$S.S. = \sum (x_i - a)^2 - \frac{[\sum (x_i - a)]^2}{n} \quad (2.15)$$

a のえらびかたがうまければ (2.15) 式の第一項は、通常二乗表を利用して行方が 300 までの二乗表があれば十分計算できる。この第 1 項を未修正二乗和といい、第 2 項を修正項 (correction factor, C.F.) という、

表 2.3 度数表による計算
石炭サンプリングのデータ (表 2.1 のデータ)

級の番号	中心値 $x\%$	度数 f	新しい 単位 u	u^2	fu	fu^2	$u+u^2$	$f(u+u^2)$	相対 度数 %
1	7.355	1	-4	16	-4	16	12	12	1
2	7.555	7	-3	9	-21	63	6	42	7
3	7.755	7	-2	4	-14	28	2	14	7
4	7.955	24	-1	1	-24	24	0	0	24
5	8.155= x_0	14	0	0	-63	0	0	0	14
6	8.355	16	1	1	16	16	2	32	16
7	8.555	8	2	4	16	32	6	48	8
8	8.755	10	3	9	30	90	12	120	10
9	8.955	6	4	16	24	96	20	120	6
10	9.155	2	5	25	10	50	30	60	2
11	9.355	1	6	36	6	36	42	42	1
12	9.555	3	7	49	21	147	56	168	3
13	9.755	1	8	64	8	64	72	72	1
計	(級の間隔 $h=0.2\%$) $\Sigma f = n$	100	—	—	131 -63 68+	662	—	730	100

計算法 1) 大体平均値に近い級を 0 とし、この中心値を x_0 とする。それより大きい方へ級を単位として、1, 2, ……、小さい方へ -1, -2, …… と新しい単位 u をとる。 u は最小の級を 0 とすることもある。計算機を用いるときはこの方が便利。

2) fu , fu^2 , $f(u+u^2)$ およびそれらの和をもとめる。

3) $fu+fu^2=f(u+u^2)$ が成立すれば計算は合っている。

4) 平均値の計算

$$\bar{x} = x_0 + \frac{\Sigma fu}{\Sigma f} \times h = 8.155 + \frac{68}{100} \times 0.2 = 8.291 \quad (2.16)$$

5) 標準偏差の計算

$$s = h \times \sqrt{\left(\frac{\Sigma fu^2 - \frac{(\Sigma fu)^2}{\Sigma f}}{\Sigma f} \right)} = 0.2 \times \sqrt{\left(\frac{662 - \frac{68^2}{100}}{100} \right)} \quad (2.17)$$

$$= 0.2 \times \sqrt{(662 - 46.24) / 100} = 0.2 \times \sqrt{6.1576} \approx 0.2 \times 2.48 = 4.96$$

注意 1) 級の数の多い程、精度はよくなる。

2) 標準偏差の計算には級をわけ、その各級に入る値をその中心値により代表させたことによる補正 (Sheppard の補正) をすることもあるが、実用的にはなくともよい。

4) 10^n (一般には h) を掛けて計算し、最後に、偏差二乗和および分散の場合には 10^{-2n} (一般には h^{-2})、標準偏差では 10^{-n} を掛ける。

たとえば、0.25, 0.23, 0.22 というデータがあれば、まず 0.23 を引き、0.02, 0, -0.01 とし、100 倍して 2, 0, -1 として S.S. を計算すると

$$S.S. = \frac{1}{(100)^2} \left\{ (4+0+1) - \frac{(2+0-1)^2}{3} \right\} = 4.67 \times 10^{-4}$$

5) 同じ値のデータが多く出てくるときには

$$S.S. = \sum x_i^2 f_i - \frac{T^2}{n} \quad (2.18)$$

6) データの数が非常に多いときには、度数表をつくり、簡略算を行う (表 2.3 参照)。

範囲 (レンジ range) 測定値の最大値と最小値の差の絶対値である。

$$R = |x_{\max} - x_{\min}|$$

範囲は標準偏差にくらべて計算が非常に簡単であるので、精度はやや悪いがよく用いられる。一般に測定値の数が 10 以下のときに用いてよく、10 以上になったら使わぬ方がよい。工場におけるサンプリングのように、毎日データがもとめられるような場合には、サンプリングの精度をもとめるのによく用いられる。

変動係数 (coefficient of variation) 標準偏差を平均値で割ったものである。したがって単位に関係のない測度である。これは平均値を単位として、相対的なバラツキの大きさをあらわすのに便利であり、市場調査などの標本調査法ではよく用いられる。工場におけるサンプリングでも用いることがある。

$$\text{サンプルについて、} c \text{ とすれば} \quad c = s/\bar{x} \quad (2.19)$$

$$\text{母集団について、} C \text{ とすれば} \quad C = \sigma/\mu \quad (2.20)$$

また、これと逆に各測定値などのバラツキの状況を標準偏差を単位としてあらわすことを**規準化 (standardize)** するといひ、偏差を標準偏差で割ったもので示す。サンプルについては

$$u_i = \frac{x_i - \bar{x}}{s} \quad (2.21)$$

この u_i を規準測度(standard measure) といひ、測定の単位に関係なく各データが相対的にどのくらいバラツキているかを示すのに便利である。

このほか、バラツキをあらわす測度として、サンプリングにおいて従来よく用いられている平均偏差(mean deviation) という量もあるが、新しい数理統計学では殆んど用いられない。

$$\text{平均偏差} = \frac{\sum |x_i - \bar{x}|}{n}$$

ゆがみやとがりをあらわす測度は、殆んど用いないが、式だけを記しておこう。ゆがみを k であらわすと、サンプルについては

$$k = \sqrt{\beta_1} = \frac{\frac{1}{n} \sum (x_i - \bar{x})^3}{s^3} \quad (2.22)$$

または、簡単に、近似的に次式であらわすこともある。

$$k = \frac{\text{平均値} - \text{モード}}{\text{標準偏差}}$$

モードが平均値より左に、すなわち負の方にカタヨっている分布は、右に長く正の方へスソ(tail)を引いているので $k > 0$ となり、これを正に歪んでいる、右に歪んでいるといひ、この逆を負または左に歪んでいるといふ。

とがりを β_2 とすると(α_4 とあらわすこともある)、サンプルについては

$$\beta_2 \equiv \alpha_4 = \frac{\frac{1}{n} \sum (x_i - \bar{x})^4}{s^4} \quad (2.23)$$

β_2 の大きいことは分布がとがっていることを示し、小さいことは扁平なことを示している。(2.22)式および(2.23)式はこのままでは計算が厄介であるので、通常簡便法を用いる。

注意 以上のべて来たように、バラツキをもったデータをあらわす測度にはいろいろあるが、この測度には、母集団そのものに関するものと、サンプリングしてえたサンプルについての場合とは、はっきり区別しなければならぬ。本書では、後述の各章においても記号を区別してあるから、その式が母集団についての式であるか、サンプルについての

1) 統計的には $\sqrt{\beta_1}$ であらわす場合が多い。

式であるか、よく認識して、使っていただきたい。

2・3・2 分布、確率

われわれのデータは、バラツキをもっているから、必ず分布をもっている。その相対度数を理想化した形で分布曲線がえられる。われわれの現場のデータから、分布曲線そのものは得られないが、近似的には大体そのデータの示す分布状況を分布曲線であらわすことができる。われわれの得られる分布曲線にはいろいろな分布がある。

大きく分けると、連続分布と離散的分布に分けられる。

連続分布 (continuous distribution) とは、連続的な値の分布である。たとえば、長さ、目方、温度、圧力、収率などのように連続的な値をとるもの、すなわち計量的な値、計量値(variable)の分布はこれに属する。この分布には統計的に大切な正規分布、カイ二乗分布、 t 分布、 F 分布等がある。

離散的分布 (discontinuous distribution) は、不連続的な値の分布である。たとえば、電球何コ中不良何コとか、繊維の1級、2級、3級、4級品が各何束とか、鉄板中のキズが何コとか、工場における事故回数、あるいはある製品の非使用者の数であるとか、一般に、0, 1, 2, ……という整数値しかとらず、不連続な値である。あるいは製品 n コ中不良 k コというときに、不良率をとる場合でも、不良1コ増加するにしたがって、不良率は飛び飛びに不連続な値をとる。このように不連続な値を、計数値(attribute)というが、この分布は離散的分布をなす。すなわち不連続分布である。この中2項分布、ポアソン分布などが大切な分布である。

これらの分布の中、特に統計的に大切な分布は**正規分布 (normal distribution)**である。正規分布は、昔の誤差論などにおいても、ガウスの誤差曲線として知られている分布である。サンプリングや測定誤差の分布、製品の品質特性の分布も、正規分布に近いものが殆んどである。正確にいうと、少し歪んでいるものもあるが、実用的には正規分布と見なしてよい場合が多い。もし非常に歪んだ分布でも、簡単な変数変換、たとえば \sqrt{x} 、 x^2 、 $1/x$ 、 $\log(a+x)$ など、を行くと正規分布とみなせるようになる。しかしサンプリングや測定の誤差を論ずる場合には、そのまま正規分布とみなして殆んど間違いない。

正規分布は図2・6に示したような左右対称のベル型の分布である。したがって、平均値はモードと一致し、ゆがみはなく、 $k=0$ 、とかりかたは、 $\beta_2=3.0$ である。 k や β_2 がきまつているので、正規分布を規定するには、平均値と分散をきめればよい。平均値 μ 、分散 σ^2 の正規分布を $N(\mu, \sigma^2)$ とあらわす。この分布を平均値を中心にして、図のように標準偏差で切っ

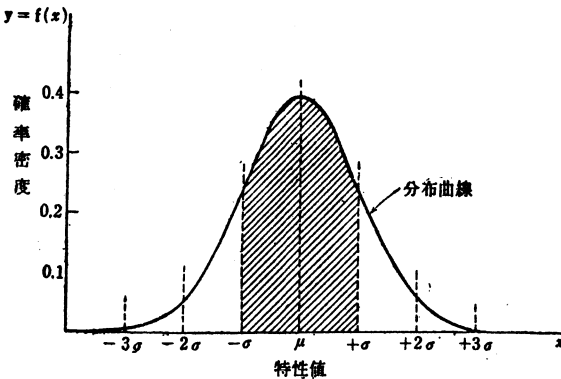


図 2・6 正規分布

て見ると、 $\mu \pm \sigma$ 、 $\mu \pm 2\sigma$ 、 $\mu \pm 3\sigma$ の内側に入る、あるいは外に出るデータの度数がわかる。附表1にこの相対度数が各データの平均値からの偏差を標準偏差で割った値、規準化した値 u に対して示してある。すなわち曲線にかこまれた全面積と、斜線の部分の面積の割合が $\mu \pm \sigma$ の間に入るデータの割合、相対度数が68.3%あることを示している。これはいいかえると、正規分布にしたがうデータを1つとったときに、 $\mu \pm \sigma$ の中に入ると期待される割合は68.3%である。すなわちデータを1つずつ、繰り返し100回とると、その中68.3%が $\mu \pm \sigma$ の中に、31.7%が $\mu \pm \sigma$ より外にでることが期待される。この期待される割合のことを確率 (probability) という。たとえば、正しいサイコロを正しくふって1の目が出る確率は1/6であることが容易にわかるであろう。確率は0.683とあらわすことと、100倍して68.3%とあらわすことがある。

同様に $\mu \pm 2\sigma$ より外に出るようなデータの確率は、4.6% ($P_r =$

0.046), $\mu \pm 3\sigma$ 外に出る確率は約 0.3% ($P_r = 0.003$) である。

すなわち、データが分布をもつということは、ある値以上または以下のデータが出る確率というものが考えられることになる。

この確率という概念が入って来た点が、新しい数理統計学の一つの大きな特徴である。前にのべたすべての分布において、確率ということを考えなければならない。

正規分布曲線に関する数式的問題は、本書の範囲を越えるので、次に代表的な式を参考のためにあげるにとどめる。平均値 μ 、分散 σ^2 の正規分布曲線は次式であらわされる。

$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

規準化した形では

$$y = f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

ただし y は縦軸の値で、確率密度という。

規準化した正規分布は平均値 0、分散 1 の分布 $N(0, 1)$ をなす。附表 1 に正規分布に関するいろいろの値を示してある。

2.3.3 統計量の分布

測定値が分布をもっているのと同様に、われわれがある分布をもった 1 つの母集団から、サンプルをとり、そのサンプルの平均値、分散、標準偏差などの統計量をもとめると、この統計量は一定の値とはならず、やはりある分布をもってデータがバラツク。たとえば、0 から 9 までの特性をもった 10 コの製品から、ランダムに 3 コのサンプルをとり (1 コとり測定したらもとへもどし、よく攪拌してまた 1 コ取り出すという操作を繰り返して)、その平均値を計算し、これを 5 回繰り返したら次のようになった。

測定値	平均値
5, 9, 8	7.33
2 0 6	2.67
0 2 4	2.00
6 9 5	6.67
0 7 7	4.67

すなわちサンプルの各測定値がサンプリングによりバラツクので、その統計量も、必ずバラツキ、分布をもっている。たとえばサンプルの平均値は、ある一定の平均値や標準偏差をもった分布をなす。この統計量の標準偏差のことを標準誤差 (standard error) といい、分散を誤差分散 (error variance) といふ。

この統計量の分布は、サンプルがランダムにとられていれば、ランダムサンプリングしたサンプルのデータからの統計量であれば、その分布法則が数理統計学によりもとめられる。サンプリングを統計的に解析して行くには、ランダムサンプリングをなすことが前提条件になっているのはこの点にある。

次に重要な統計量の分布法則をのべよう。

平均値の分布 サンプルの平均値 \bar{x} は、サンプルがある1つの母集団からランダムサンプリングされていけば、次の法則に従って分布する¹⁾。

1. \bar{x} の平均値 $\mu_{\bar{x}}$ は、母集団の個々のサンプル単位の平均値 μ_x に等しい²⁾。

$$\mu_{\bar{x}} = \mu_x \quad (2.24)$$

ある分布をもつ母集団から、一定数のサンプルをとり、その平均値を繰り返しもとめると、その各サンプルからの平均値は、もとの母集団分布の平均値 μ_x のまわりにバラツク。母集団の大きさ N 、サンプルの大きさ n には関係しない。

2. \bar{x} の標準誤差 $\sigma_{\bar{x}}$ は、母集団の大きさ N 、サンプルの大きさ n 、もとのサンプル単位の母集団の標準偏差 σ_x とすると

$$\sigma_{\bar{x}}^2 = \frac{N-n}{N-1} \frac{\sigma_x^2}{n} \quad (2.25)$$

もし $N \gg 1$ ならば

$$\sigma_{\bar{x}}^2 \approx \left(\frac{1}{n} - \frac{1}{N} \right) \sigma_x^2 = \left(1 - \frac{n}{N} \right) \frac{\sigma_x^2}{n} \quad (2.26)$$

1) ここではすべて母数についてのべる。ある1組のサンプルについて、統計量の実現値を考えると、たとえば平均値は、そのような平均値の分布という母集団からの1つのサンプルと考えられる。 $\mu_{\bar{x}}$ 、 $\sigma_{\bar{x}}^2$ はこの母集団の母数である。

2) 統計量の期待値 (expected value) とはその分布の平均値のことである。たとえば $E(\bar{x}) = \mu$ とかく。

もし $n/N \leq 0.1$ ならば (無限母集団と見なせる場合)

$$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n} \quad (2.27)$$

サンプリングして、材料などの平均値をなるべくバラツキを小さくして知りたい場合が多いので、上の3式はサンプリングにおいて非常に重要な式である。

注意 1) (2.25) 式は有限母集団からサンプリングしたときの、平均値のバラツキをあらわす式で、 $(N-n)/(N-1)$ を有限母集団修正 (有限修正; finite population correction; finite multiplier) という。

市場調査とか、社会調査の場合には、抜取比 $a = n/N$ が大きく $a > 0.1$ のときには (2.25) 式あるいは (2.26) 式を用いなければならないが工場におけるサンプリングのときには、 $a \leq 0.1$ すなわち 10分の1もサンプルをとることは少ないので、(2.27) 式で十分な場合が多い。

すなわち無限母集団からのサンプリングと見なせる場合となる。管理図の場合には、工程を母集団と見ているから、無限母集団として(2.27)式を用いる。

最近では $a < 0.15$ なら (2.27) 式で十分な場合が多いともいわれている。後章においてのべるいろいろなサンプリング法においても、有限修正がついていると、非常に式が複雑になるが、幸に工場では、これの必要な場合が少ないので、今後の式もなるべく使いやすいうように簡略化してある。

2) (2.27) 式は非常に興味ある性質をもっている。

a) 平均値 \bar{x} の分散は、もとのサンプル単位間の分散の n 分の1になる。したがって、平均値のバラツキをできるだけ小さくするためには、

i) サンプルの大きさ n を大きくする。サンプルを多くとる。

ii) サンプル単位間の分散 σ_x^2 を小さくするようにする。これは前にのべたサンプル単位が選定できる場合、たとえばインクリメントや試長をきめるときに、インクリメント間の分散を小さくするように考えなければならぬことを示している。

b) $\sigma_{\bar{x}}^2$ はサンプルの大きさ n には反比例するが、母集団の大きさ N には関係しない。すなわち、母集団のバラツキ σ_x^2 が同じであれば、母

集団の大きさ 10,000 コの場合でも、1,000 コの場合でも、同じサンプルの大きさ、たとえば $n=50$ コ、では、平均値の精度（バラツキ）はかわりないことを示している。たとえば (2.26) 式より

$N = 10,000$ のとき

$$\sigma_{\bar{x}}^2 = 0.995 \frac{\sigma_x^2}{50} \quad \sigma_{\bar{x}} = 0.997 \frac{\sigma_x}{\sqrt{50}}$$

$N = 1,000$ のとき

$$\sigma_{\bar{x}}^2 = 0.95 \frac{\sigma_x^2}{50} \quad \sigma_{\bar{x}} = 0.975 \frac{\sigma_x}{\sqrt{50}}$$

すなわち標準偏差で約 2% 位しかかわらぬので、殆んど問題にする必要はない。この場合サンプルの大きさを 1 コかえればよい程度である。すなわち統計的には工場におけるサンプリングでは、サンプルの大きさは、母集団の大きさに無関係にきめられるということである。従来サンプリング法がややもすると、ロットの何分の 1 とるというように決められているのは、この点から考えて不合理なものが多い。

しかし技術的に考えると、たとえば、石炭のサンプルをとるときに、10,000 トンのロットでも、30 トンのロットでも、1 トンのサンプルでよいであろうか。この場合 3 つの問題を考えて、サンプルのとりかたを決めなければならない。

i) 母集団のバラツキが同じであれば、という条件を考えて見ると、10,000 トンのロットと 30 トンのロットでは、 σ_x が同じであるとは考えられず、ロットの大きさの大きくなる程、ことなつた条件の石炭が入ってきて σ_x が大きくなる場合が多い。完全な管理状態で生産し、出荷された製品なら、 σ_x はかわらぬが、わが国の鉱工業の現状では、このような状態は少ないであろう。しかし管理状態からの製品であれば、サンプルの量はロットに関係しないという利点がある。

ii) ランダムサンプリングであればという条件があるが、一般にロットが大きくなる程、ランダムサンプリングが困難になる。

iii) 経済的に考えて、10,000 トンのロットの平均値を知るための精度と、1,000 トンのロットの平均値を知るための精度と同じでは危険である。たとえば 1,000 トンにつき 1% 違つたと 10 万円購入入致がことなるときに、10,000 トンなら 100 万円の誤りをおかすことになる。

以上 i) ii) iii) の条件を考えると、統計的にサンプルの大きさ n は、ロットの大きさ N に殆んど関係ないからといって、ロットの大きさを無制限に大きくするよりは、技術的に考えて、ロットの大きさを適当に制限していくつかの副ロットに分けて、その各々からサンプリングした方がよい場合が多い。

以上 1, 2 の平均値の分布に関する性質は、もとの母集団の分布型にかかわらず成立する。

3. \bar{x} の分布型は、もとの母集団が正規分布の場合は勿論正規分布をなす。もし母集団の分布型が正規分布でない場合でも、サンプルの大きさ n が大きくなるにしたがって、正規分布に近づく。

通常われわれが工業で接するような母集団は、そう極端に歪んだ分布は少ないから、ランダムサンプリングしたときには、サンプルの大きさ $n = 4 \sim 5$ もとれば、正規分布をなすと見なしてよい。極端に歪んだ分布でも、 $n = 7 \sim 8$ くらいになれば正規分布と見なしうる。

以上 1, 2, 3 を総合すると、サンプルの大きさ n の平均値の分布は

$$N(\mu, \sigma^2/n)$$

である。

注意 1) 以上の平均値の分布に関する性質は、同じ母集団からのサンプルの平均値についての法則であって、別々の母集団からとったサンプルの平均値は、このような法則にはあてはまらない。勿論正規分布をなすということもこれだけではいえない。管理状態の工程からのいくつかの製品のロットからのサンプルであれば、正規分布に近い。

2) 以下の統計量の分布も、すべて同じ母集団からの統計量の分布であることを注意せよ。

分散の分布 ある正規無限母集団 $N(\mu, \sigma^2)$ からの n コのランダムサンプルの試料分散は、次の法則にしたがって分布する。

1. 試料分散 s^2 の平均値 μ_{s^2} , および s^2 の誤差分散 $V(s^2)$ は次の式であらわされる。

$$\mu_{s^2} = \frac{n-1}{n} \sigma^2 \quad (2.28)$$

$$V(s^2) = \frac{2(n-1)}{n^2} \sigma^4 \quad (2.29)$$

一般の母集団の場合には

$$V(s^2) = \frac{\beta_2 - 1}{n} \sigma^4 \quad (\text{ただし } n \gg 1)$$

この場合も n が大きくなれば $V(s^2)$ は急速に小さくなるが、 σ は 4 乗できいてくるので相当バラツキは大きい。

2. s^2 の分布型は複雑な型になるので省略するが、 n が小さいと相当歪んだ分布になる。

3. ns^2/σ^2 という統計量を考えると、この統計量 χ^2 の分布は

$$\chi^2 = \frac{ns^2}{\sigma^2}$$

$$\chi^2 \text{ の平均値 } \mu = n - 1 = f \quad (2.30)$$

$$\chi^2 \text{ の分散 } \sigma^2 = 2(n - 1) = 2f \quad (2.31)$$

という自由度 f のカイ二乗 (χ^2) 分布 (chi-square distribution) をなす。

注意 カイ二乗分布は自由度により、この場合にはサンプルの大きさにより変化する分布で、図 2.7 のような歪んだ分布型をもつ。この分布は後にのべる、母分散の信頼区間をもとめるときに用いる。

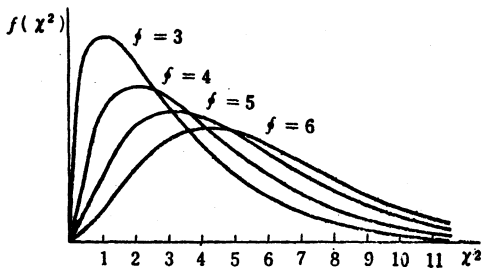


図 2.7 χ^2 分布と自由度

標準偏差の分布 ある正規無限母集団 $N(\mu, \sigma^2)$ からの n コのランダムサンプルの標準偏差 s の分布は次のようになる。

$$\text{平均値 } \mu_s = (1 - 3/4n - 7/32n^2 - \dots) \sigma = c_2 \sigma \quad (2.32)$$

$$\text{分散 } \sigma_s^2 = (1/2n - 2/8n^2 - 3/16n^3 \dots) \sigma^2 = \sigma^2/2n \quad (2.33)$$

1) この f を自由度という。

ただし c_2 はサンプルの大きさ n により変化する係数で、附表2 管理図の係数表よりもとめられる。

分布型は複雑な型になるので省略する。

範囲の分布 ある正規無限母集団 $N(\mu, \sigma^2)$ からの n コのランダムサンプルの範囲 R の分布は次のようになる。

$$\text{平均値} \quad \mu_R = d_2\sigma \quad (2.34)$$

$$\text{分散} \quad \sigma_R^2 = d_3^2\sigma^2 \quad (2.35)$$

$$\text{標準偏差} \quad \sigma_R = d_3\sigma \quad (2.36)$$

ただし、 d_2 および d_3 は n により変化する係数で、附表2 よりもとめられる。

分布型は複雑な型になるので省略する。

注意 以上の試料分散、試料標準偏差および範囲の分布は、正規型の無限母集団からのランダムサンプルであるという假定が入っているが平均値の分布のところの注意でものべたように、われわれが工場で接する程度の歪みかたをもった、抜取比が0.1より遙かに小さい母集団、あるいは無限母集団からのサンプルにおいては、近似的に大体そのまま当てはまると考えてよい場合が多い。

2.3.4 推定

これまでのべてきたことにより、母集団はある一定の母数、たとえば平均値や分散をもっているが、それからランダムサンプリングしたサンプルの測定値や統計量は常に一定の法則にしたがった分布をもってバラツいていることがわかった。ところがわれわれは、母数そのものを知ることはできないし、サンプリングしてデータをとる目的は、サンプルについてのデータを知るためではなくて、母集団についての知識を得たいためである。母集団を知るためには、母平均、母分散などの母数を**推定 (estimate)**すればよい。ところがサンプルからの統計量はバラツキをもっているので、この統計量より母数をもっともよく推定し、そのときの推定の精度や信頼度をはっきりさせて推定することが必要である。サンプリングを統計的に考えなければならぬ点はこのにある。

母数を推定するには、点推定と区間推定とがある。このほか他の特性値

を利用して推定する比推定¹⁾があるが問題の種類が一寸ことなるのでここではふれない。

点推定 (point estimation) とは、母数そのものをもっともよく一つの値で推定しようという方法であり、**区間推定 (interval estimation)** とは、ある区間をつくってその値の間に母数が入っていることを推定しようというのである。通常工場で行っているサンプリングの目的は、平均値の推定を所要経費である精度で行う場合が多い。

1. **点推定 (point estimation)** 母数を点推定により、サンプルより推定した値を**推定値 (estimate)** という。推定値は、母数をできるだけ間違いなく推定するような工夫が行われているが、あくまで母数そのものではなく、バラツキをもっているので母数とははっきり区別しなければならない。したがって本書においても、記号を区別して記すから、注意して読まれたい。たとえば分散についても、母分散と、試料分散と、母分散の推定値と3種類あるから注意されたい。

推定値の中、最もよい推定値を**不偏推定値 (unbiased estimate)** という。ある母数の不偏推定値としては、それに相当する統計量の分布の平均値を利用される。

母平均の点推定 母平均の推定には、サンプルの平均値を用いるのが、もっともよい²⁾。

いま母平均の不偏推定値を $\hat{\mu}^3$ とすると

$$\mu \leftarrow \hat{\mu} = \bar{x} \quad (2.37)$$

その誤差分散は、(2.25) 式と同じで、サンプル単位の母分散を σ_x^2 とすると

$$\sigma_{\bar{x}}^2 = \frac{N-n}{N-1} \frac{\sigma_x^2}{n} \quad (2.25)$$

1) §4・9参照。

2) 市場調査などの社会調査の場合には、これが簡単に成立しない。カテゴリーの入る場合があるが、工場におけるサンプリングの際にはサンプル単位の大きさが一定であるのでこのようなことは殆んどない。以下の推定についても同様。ただしサンプル単位の大きさにバラツキのある場合は注意を要する。

3) 推定値には原則として ^ 印をつけることにする。場合により ● という添字をつけることもある。

$$\Rightarrow \left(1 - \frac{n}{N}\right) \frac{\sigma_x^2}{n} \quad (\text{ただし } N \gg 1) \quad (2.26)$$

$$\Rightarrow \frac{\sigma_x^2}{n} \quad (\text{ただし } \frac{n}{N} \leq 0.1) \quad (2.27)$$

その標準誤差 $\sigma_{\bar{x}}$ も同様にして

$$\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \frac{\sigma_x}{\sqrt{n}} \quad (2.38)$$

$$\Rightarrow \left(1 - \frac{n}{2N}\right) \frac{\sigma_x}{\sqrt{n}} \quad (\text{ただし } N \gg 1) \quad (2.39)$$

$$\Rightarrow \frac{\sigma_x}{\sqrt{n}} \quad (\text{ただし } \frac{n}{N} \leq 0.1) \quad (2.40)$$

変動係数 $C_{\bar{x}}$ は、母集団の変動係数を C_x とすると

$$C_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \frac{C_x}{\sqrt{n}} \quad (2.41)$$

$$\Rightarrow \left(1 - \frac{n}{2N}\right) \frac{C_x}{\sqrt{n}} \quad (\text{ただし } N \gg 1) \quad (2.42)$$

$$\Rightarrow \frac{C_x}{\sqrt{n}} \quad (\text{ただし } \frac{n}{N} \leq 0.1) \quad (2.43)$$

母集団総計の点推定 たとえば、袋入の砂糖、綿花、肥料あるいは鉄板などのロットがあるとき、そのロット全量を n コのサンプルから推定したいときがこれに相当する。いま総計の母数を μ_X とすると、その不偏推定値 $\hat{\mu}_X$ は

$$\hat{\mu}_X = N\bar{x} = \frac{N}{n} \sum_{i=1}^n x_i \quad (2.44)$$

誤差分散は、(2.24) 式などを利用して

$$\sigma_{\bar{x}}^2 = N^2 \frac{N-n}{N-1} \frac{\sigma_x^2}{n} \quad (2.45)$$

$$\Rightarrow N^2 \frac{\sigma_x^2}{n} \quad (\text{ただし } \frac{n}{N} \leq 0.1) \quad (2.46)$$

標準誤差は同様にして

$$\sigma_{\bar{x}} = N \sqrt{\frac{N-n}{N-1}} \frac{\sigma_x}{\sqrt{n}} \quad (2.47)$$

$$\Rightarrow N \frac{\sigma_x}{\sqrt{n}} \quad (\text{ただし } \frac{n}{N} \leq 0.1) \quad (2.48)$$

変動係数 C_x は、平均値の場合と全く同じになる。

$$C_x = \sqrt{\frac{N-n}{N-1}} \frac{C_x}{\sqrt{n}} \quad (2.49)$$

$$\Leftrightarrow \frac{C_x}{\sqrt{n}} \quad (\text{ただし } \frac{n}{N} \leq 0.1) \quad (2.50)$$

備考 以上の平均値及び総計の点推定の精度については、後に各々の場合において論ずることとする。

分散の点推定 母分散の不偏推定量を $\hat{\sigma}^2$ とすると (σ^2 , V_{σ^2} ともあらわす) (これを不偏分散ともいう), (2.28) 式より

$$\sigma^2 \leftarrow \hat{\sigma}^2 \equiv \sigma^2_e = \frac{n}{n-1} s^2 \quad (2.51)$$

$$= \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{S.S.}{n-1} \quad (2.52)$$

この $n-1 = f$ を自由度 (degree of freedom) という。すなわち不偏分散は、偏差二乗和を自由度で割ったものである。

$$V(\hat{\sigma}^2) \Leftrightarrow \left(\frac{n}{n-1} \right)^2 V(s^2)$$

分散の不偏推定値はサンプリングにより相当大きなバラツキをもっている。推定の精度は平均値にくらべるとあまりよくない。

もし多くのサンプルから s^2 が多くもとめられているときには、 n が一定ならば平均値 \bar{s}^2 を用いた方が精度はよい。

不偏分散の比の分布 同じ正規無限母集団 $N(\mu, \sigma^2)$ から、二つのランダムサンプル (サンプルの大きさ n_1, n_2) をとったとき、その不偏分散の比の分布についてのべる。この分布は非常に重要な分布であり、応用範囲も非常に広い。いま、このような2つの不偏分散の比を F とすると

$$F \Leftrightarrow \frac{\frac{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2}{n_1 - 1}}{\frac{\sum_{j=1}^{n_2} (x_j - \bar{x}_2)^2}{n_2 - 1}} = \frac{V_{e_1}}{V_{e_2}} \quad (2.53)$$

1) この場合 $n-1$ で割ることに注意。試料分散 (2.11) 式では n で割っている。

この F は自由度($f_1=n_1-1$, $f_2=n_2-1$)の F 分布をなす。すなわち同じ母集団から、サンプルを二つとり、その不偏分散の比という統計量をもとめると、この比は一定の分布、 F 分布にしたがうことが、数学的に証明されている。

この F 分布の値は、一般に $V_{e_1} > V_{e_2}$ になるように、すなわち F の値が1より大きくなるように比をとった場合の、各自由度に相当する F の値が表としてあたえられている。多くの統計書、数値表にその値がでている。

注意 1) F 分布は、正規無限母集団からのランダムサンプルであるという假定があるが、われわれが工場で接するような母集団からのサンプルの場合には、 F 分布にそのままあてはまると考えてよい。

2) 自由度(1, f_2)の F の平方根は自由度 f_2 の t 分布に等しく、自由度(f_1, ∞)の F 分布は自由度 f_1 の χ^2/f_1 分布に等しい。

標準偏差の点推定 母標準偏差の不偏推定値 $\hat{\sigma}$ は、試料標準偏差からと範囲からと両方から推定する方法がある。

$$\sigma \leftarrow \hat{\sigma} = \frac{s}{c_2} \quad \text{または} \quad \frac{\bar{s}}{c_2} \quad (2.32 \text{ 式より}) \quad (2.54)$$

$$= \frac{R}{d_2} \quad \text{または} \quad \frac{\bar{R}}{d_2} \quad (2.34 \text{ 式より}) \quad (2.55)$$

上の2式ではいずれも \bar{s} , \bar{R} を用いた方が精度はよい。

変動係数の点推定 変動係数を推定するには、上にのべた平均値の不偏推定値と、標準偏差の不偏推定値を用いればよい。 \hat{C}_v は

$$\hat{C}_v = \frac{\hat{\sigma}}{\hat{\mu}} \quad (2.56)$$

注意 1) 平均値や総計の不偏推定値のパラツキをあらわす式(2.25~27)(2.38~43)(2.45~50)式では母分散、母標準偏差、母変動係数を用いているが、これが不明のときは、上にのべたこれらの不偏推定値を代入すればよい。

2) 分散の不偏推定値の平方根 σ_s も、標準偏差の推定値ではあるが、厳密にいうと標準偏差の不偏推定値ではない。同様に $(\hat{\sigma})^2$ (標準偏差の不偏推定値の二乗)は σ_s^2 (分散の不偏推定値)とはことなるし、また σ , $\hat{\sigma}$, σ_s , s はことなつた性格のものであり混同してはならない。しか

し工場におけるサンプリングの問題の場合のように、多くのデータがとれ、かつ工程が完全な管理状態にない場合には

$$\hat{\sigma} \approx \sigma_s$$

$$\hat{\sigma}^2 \approx \sigma_s^2$$

とみなしてよい場合が多い。

さらに非常に多くのデータが得られている場合には

$$\sigma \approx \hat{\sigma} \approx s$$

$$\sigma^2 \approx \sigma_s^2 \approx s^2$$

とみなしてよい。すなわち母分散そのものが既知の場合として、統計的に取り扱って差支えない。

2. 区間推定 信頼区間 母数がある区間に入っていることを推定する方法である。この推定の区間を信頼区間 (confidence interval) という。

母平均の区間推定 母平均がある区間に入っていることをサンプルから推定するには、母標準偏差、したがって母分散が既知の場合と、未知で不偏推定値あるいは試料標準偏差を用いた場合とではことなる。

a) 母標準偏差既知の場合

いまつぎのような統計量を考えると

$$u = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \quad (2.57)$$

\bar{x} の分布が $N(\mu, \frac{\sigma^2}{n})$ であり、 u という値は、これを規準化したものであるから、平均値0、分散1の正規分布 $N(0, 1)$ をなす。したがって附表1より、 u の絶対値が大きき値をとる確率は小さい。いま u を両側で5%の値を附表1よりもとめると、 $u=1.96$ となる。したがって u の絶対値が1.96より大きき値をとる確率は5%であり、1.96より小さい値をとる確率は95%である。すなわち平均値 μ 、分散 σ^2 という母集団から、 n コのサンプルをランダムに100回繰返しとった測定値の平均値 \bar{x} を用いて、 u の値を各々算出すると、その中95回は ± 1.96 以内にあることが期待される。いま (2.57) 式の右辺が、両側で $\alpha\%$ の u の値 u_α より小さい値をとると

すると

$$u_{\alpha} > \left| \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \right|$$

$$\therefore \bar{x} - u_{\alpha} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + u_{\alpha} \frac{\sigma}{\sqrt{n}} \quad (2.58)$$

が成立する確率は $(100 - \alpha)\%$ である ($\alpha = 5\%$ とすれば 95%)。この (2.58) 式は、 \bar{x} がこのような値をとる確率が 95% であることを示していると同時に、母平均 μ (未知) が、 95% の確率でこの区間に入ることを示している。この区間を信頼度 $(1 - \alpha)$ の母平均の信頼区間という。

信頼度 95% の母平均の信頼区間とは、したがって、つぎのような意味である。このようなサンプルを 100 回とり、その各々についてこのような母平均の信頼区間を推定すると、その中 95 回は母平均をこの区間内に含んでいると期待できることを意味している。

この限界値を信頼限界 (confidence limits) という。

b) 不偏分散または試料標準偏差を用いる場合—母標準偏差未知の場合
母標準偏差未知の場合には、とったサンプルからえられる不偏分散または試料標準偏差を用いる。いま母平均 μ をもつ正規母集団から、 n コのサンプルをとり、それより \bar{x} , σ_s または s をもとめて次式のような統計量 t を考えると

$$t = \frac{\bar{x} - \mu}{\sigma_s / \sqrt{n}} = \frac{\bar{x} - \mu}{s / \sqrt{n-1}} \quad (2.59)$$

t は、自由度 $f = n - 1$ の t 分布 (t -distribution) という分布をなすことが統計的にわかっている¹⁾。 t 分布は自由度 f (この場合は $n - 1$: サンプルの数より 1 少ない)、によってことなった分布曲線を示す。色々な自由度に対して、ある確率以下の値となる t の絶対値が附表 3 に示してある。

いま $n = 10$, すなわち自由度 $f = 10 - 1 = 9$ として、 $\alpha = 5\%$ の t の値 $t_{0.05}$ を附表 3 よりもとめると

$$t_{0.05} = 2.26 \quad (f = 9)$$

a) のときと同様に、自由度 9 のときに、 t の値が 2.26 より大きい値

1) この場合 σ の分布は、大体正規分布をなすとみとられるから、 μ との分布が正規分布でなくても (2.59) 式の t は t 分布をなすとしてよい。

をとる確率は5%より小さい。

いま一般に t_α (自由度 $f = n - 1$, $\alpha\%$) をもとめ、(ただし α は両側の確率、附表3にはこの値が示してある)

$$t_\alpha > \left| \frac{\bar{x} - \mu}{\sigma_s / \sqrt{n}} \right|$$

または

$$t_\alpha > \left| \frac{\bar{x} - \mu}{s / \sqrt{n-1}} \right|$$

$$\therefore \bar{x} - t_\alpha \frac{\sigma_s}{\sqrt{n}} < \mu < \bar{x} + t_\alpha \frac{\sigma_s}{\sqrt{n}} \quad (2.60)$$

または

$$\bar{x} - t_\alpha \frac{s}{\sqrt{n-1}} < \mu < \bar{x} + t_\alpha \frac{s}{\sqrt{n-1}} \quad (2.61)$$

a) と同様に (2.60) および (2.61) 式は、信頼度 $(100 - \alpha)\%$ の母平均の信頼区間を示す。この場合は不偏分散などサンプルからの値をつかったために、a) の u が正規分布をなすのが t 分布にか変わっただけで、信頼区間の意味などは全く a) と同様である。区間の中は母標準偏差既知の場合の方が遙かにせまい。

注意 1) t 分布は自由度によりやや尖りかたなどがかわってくるが、図 2.8 に示すように、左右対称の正規分布によく似た分布である。自由度 f が大きくなると、次第に正規分布に近づき、 f が 50 以上になれば、実用的には正規分布とみなして差支えない。理論的には、自由度無限大の場合に t 分布は正規分布になる。

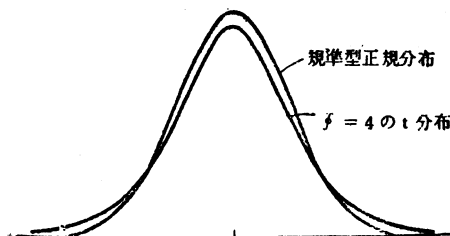


図 2.8 t 分布

2) 自由度 f の χ^2 の分布は、自由度 $(1, f)$ の F 分布をなす。

母分散の区間推定 母分散の信頼区間を推定するには前にのべた ns^2/σ^2 が自由度 $f = n-1$ のカイ二乗分布をなすという性質を利用する。

$$\chi^2 = \frac{ns^2}{\sigma^2}$$

カイ二乗分布は非対称な分布であるから、対称な分布である正規分布や t 分布を応用した平均値の区間推定のように行かない。したがって、たとえば信頼度 96% で母分散の信頼区間を推定したいときには、自由度 $f = n-1$ のカイ二乗分布の大きい方へ 2% の確率をとるような χ^2 の値 χ_1^2 と、小さい方へ 2% の確率をとるような χ^2 の値 χ_2^2 を附表 4 からもとめると、 ns^2/σ^2 という統計量は、信頼度 96% で χ_1^2 と χ_2^2 の間に入るということが期待できる。すなわち

$$\chi_2^2 < \frac{ns^2}{\sigma^2} < \chi_1^2 \quad (2.62)$$

ただし χ_1^2 : $f = n-1$ の確率 2% の χ^2 の値

χ_2^2 : $f = n-1$ の確率 98% の χ^2 の値

附表 4 のカイ二乗表は、ある χ^2 の値より大きい χ^2 の値の確率が表とになっているので、ある値より小さい χ^2 の値をとる確率は $(100-98)\% = 2\%$ となる。

(2.62) 式を書きなおすと

$$\frac{ns^2}{\chi_1^2} < \sigma^2 < \frac{ns^2}{\chi_2^2} \quad (2.63)$$

ただし、 χ_1^2 : 自由度 $f = n-1$ の $Pr = \alpha/2$ の χ^2 の値

χ_2^2 : " " " $= (1-\alpha/2)$ の χ^2 の値

すなわちこのような値を計算すると、母分散 σ^2 が、(2.63) 式で計算される区間に入るということが信頼度 $(1-\alpha)$ でいえる。母分散の信頼区間は、 n が大きくなると、 χ^2 の値が f により大きくなるので、 σ^2 が一定であれば、巾はせまくなる。

注意 1) 信頼度 $(100-\alpha)\%$ の区間推定をするとき、通常は大きい方の値をとる確率と小さい方の値をとる確率とを等しく、すなわち両者に $\alpha/2$ ずつとる。

2) 一般に区間推定の信頼度は通常, 90%, 95%, 96%, 98%, 99%, 99.9%などをとるが, この確率をいくらにとるかということは, 技術的に, その信頼区間をもとめる用途, 目的を考えてきめるべき問題である。たとえば, その区間に母数が含まれていなければ, 経済的に非常な損失を蒙るようなときは信頼度を大きく99%とか98%をとるべきであるし, 大体入っていればよいのであれば90%でもよい。

例1] ある薬品から10回サンプリングして, その融点を測定したら, 試料分散 $s^2=0.315$ となった。この薬品の融点の母分散の信頼度96%の信頼区間を推定せよ。

(2.63) 式において, $n=10 \therefore f=9$

$$Pr = (1-0.96) = 0.04 \quad \therefore \frac{Pr}{2} = 0.02$$

附表4 カイ二乗表より $\chi_1^2=19.68$ ($f=9$ $Pr=0.02$)

$$\chi_2^2 = 2.53 \quad (f=9 \quad Pr=(1-0.02)=0.98)$$

この値を (2.63) 式に入れると

$$\frac{10 \times 0.315}{19.68} < \sigma^2 < \frac{10 \times 0.315}{2.53}$$

$$0.160 < \sigma^2 < 1.245$$

この薬品の融点の母分散¹⁾は, 信頼度96%で0.160と1.245の間にある。

例2] 上の薬品の融点の測定を5ロットについて, 各6回ずつ測定したら, 試料分散 $s_1^2=0.285$, $s_2^2=0.354$, $s_3^2=0.328$, $s_4^2=0.296$, $s_5^2=0.315$ を得た。この会社の薬品の融点の母分散はどの位の範囲にあるか。

この場合には, カイ二乗分布の性質(加成性)を利用して, 次の一般式を利用する。

$$\frac{\sum_{i=1}^k n_i s_i^2}{\chi_1^2} < \sigma^2 < \frac{\sum_{i=1}^k n_i s_i^2}{\chi_2^2} \quad (2.64)$$

ただし χ_1^2 , χ_2^2 : それぞれ自由度 $f = \sum (n_i - 1)$ の $\alpha/2$ および

1) この分散は薬品自身の融点のパラツキと, 測定誤差によるパラツキとを含んでいる。

$(1-\alpha/2)$ の χ^2 の値

n_i : 各回の測定回数

k : 試料分散をもとめた組数

この例では $f = 5 \times 5 = 25$

$$\sum ns^2 = 6(0.285 + 0.354 + 0.328 + 0.296 + 0.315) = 9.468$$

この問題の題意から判断して、いま信頼度 96% にとると

$$\chi_1^2 = 41.57 \left(f = 25 \quad \frac{\alpha}{2} = 2\% \right)$$

$$\chi_2^2 = 12.70 \left(f = 25 \quad \left(1 - \frac{\alpha}{2} \right) = 98\% \right)$$

$$\therefore \frac{9.468}{41.57} < \sigma^2 < \frac{9.468}{12.70}$$

$$0.228 < \sigma^2 < 0.746$$

すなわち信頼度 96% で 0.227 から 0.745 の間に母分散はある。

2.3.5 サンプルの大きさ (sample size) と母集団についての情報

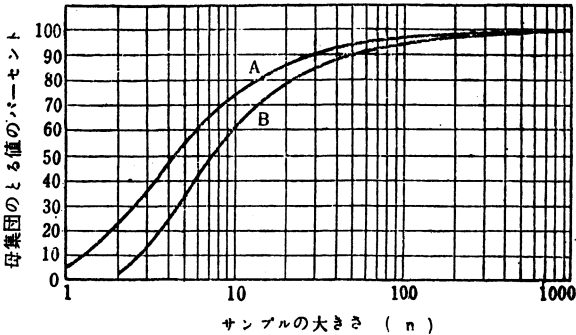
一般にサンプルがランダムにとられていれば、サンプルの大きさが多くなればなるほど、母集団についての情報はます。しかもサンプルの大きさが小さいときほど、1 コサンプル数が多くなると母集団についての情報は、急激にます。この間の関係を示す便利な図が図 2.9 である。

この図の A 曲線は母集団中のサンプル単位体の少なくとも何%が、95% の確率でサンプル中の最小値より大きい値をとるか、またはサンプル中の最大値よりも小さい値をとるかを示すものである。たとえば 1 コのサンプルをとり測定して a という値を得たとすると、母集団についての各測定値の少なくとも 5% は、95% の確率で a より大きい値をとり、また少なくとも 5% は a より小さい値をとることを示している。また 2 コのサンプルより、 a_1, a_2 ($a_1 \geq a_2$) という値を得たとすると、95% の確率で、母集団の少なくとも 22.5% は a_1 より小さい値をとり、22.5% は a_2 よりも大きい値をとることを示す。 $n=10$ ならば、74.1% が a_{\max} より小さく、 a_{\min} より大きい値をとる。すなわち n がますますしたがって、母集団のとるべき範囲が明確になってくる。一般に n コのサンプルがランダムにとられているときには、この母集団の値のうち、サンプルからの値の最大値より小さい

(または最小値より大きい) 値をとるパーセント P は

$$P = 100(0.05)^{\frac{1}{n}}\%$$

であらわされる。



- A 曲線 サンプルの大きさと、95%の確率でサンプルの最小値より大きい値 (または最大値より小さい値) を母集団の値がとるパーセントとの関係を示す。
- B 曲線 サンプルの大きさと、95%の確率でサンプルの最大値と最小値との間に母集団の値の入るパーセントとの関係を示す。

図 2.9 サンプルの大きさと母集団についての情報¹⁾

B曲線は、サンプルの最大値と最小値の間に95%の確率で母集団の値の何パーセントが入るかを示すものである。たとえば10コの測定値があるとき、その最大値と最小値の間に、母集団の値の60%が入るということが95%の確率でいえる。

1) ASTM: Symposium on Usefulness and Limitations of Samples p 13(1948)

第3章 サンプリング及び測定法の誤差

3・1 まえがき

われわれがデータをある目的でえようというときの手順¹⁾において、データをとるといふ実施段階について考えると、大きくわけて2つの段階にわけられる。

- 1) 母集団からサンプルをとる。場合によってはそのサンプルから更に測定用に少量のサンプルをとる。
- 2) そのサンプルを測定して、データを得る。

データをとる場合には必ずこの2つの段階を踏まなければならない。すなわちサンプルをとれば、必ず測定という問題が伴ってくる。ゆえにサンプリングを論ずるとき、特にサンプリングの誤差を論ずるときには、測定法の誤差をまず論じ、解決しておかなければならない。第3章ではこの2つの誤差のいろいろな関係につき、またサンプリングによる誤差及び測定法の誤差、各々について考えなければならぬことをのべる。

「工場で実際にデータをとるときに、あるいは仕様書、規格をきめるときに、測定誤差を考えずに、サンプリング誤差を論ずることは不可能であり、サンプリング誤差を考えずに、測定誤差のみを論ずることは、技術的、経済的に考えてノンセンスである。」

3・2 サンプリング誤差と測定誤差の関係

われわれがデータをとるために行う、母集団からサンプルをとるといふ過程、サンプルを測定するという過程、いずれの場合にも誤差を生じ、われわれが最後にえたデータには、この2つの誤差が加わってあらわれてくるのである。しかし測定を行わなければデータは出てこないのであるから、サンプリングを論ずるときにはまず

第1に測定法の誤差を検討しておかなければならない。

1) §1.4 参照。

この両者の誤差は互にからみ合ったままにしておいては、すなわちデータに含まれている2つの誤差—変動—を分離して論ずることができないので、

第2に、サンプルングの誤差と、測定法の誤差とを技術的に、または統計的に分離しなければならない。この分離を行ってはいじめて、両者の誤差を比較して論ずることができるのである。しかし場合によっては、特に破壊試験の場合などには、2つの誤差が交絡して、いろいろの技術的な工夫をしても分離を完全に行うことが困難な場合、あるいは不可能な場合もある。

第3に問題になることは、2つの誤差の分離ができた場合に、サンプルングの精度と、測定の精度の大きさの関係である。

いま測定値の精度を σ (母標準偏差)、サンプルングの精度を σ_s 、測定法の精度を σ_M とすると、サンプルングと測定法とは実用的には独立であるとみなせるのが普通であるから¹⁾、分散の加成性²⁾ から

$$\sigma^2 = \sigma_s^2 + \sigma_M^2 \quad (3.1)$$

いま $\sigma_s = 1.0\%$ 、 $\sigma_M = 0.2\%$ とすると

$$\sigma^2 = 1.00 + 0.04 = 1.04$$

$$\therefore \sigma \approx 1.02\%$$

すなわち σ_s に比し、 σ_M が $1/5$ くらいであると、測定法の誤差 σ_M はデータのバラツキ σ にはほとんど影響してこない。

2つの独立なバラツキの原因があり、それが加成されてえられるデータのバラツキとの関係を標準偏差であらわすと、表 3.1 のようになる。すなわち一方の原因を他方に比して標準偏差で $1/2$ か $1/3$ くらいの大きさに押えておけば実用的には十分である。たとえばある鉱石の成分のサンプルングの精度が 1% くらいあるときは、分析の精度は $0.3 \sim 0.5\%$ くらいならば十分で、これ以上小さくすることは実用的には、すなわち経済的にも技術的にもあまり意味がない。現在の測定法の精度が $1/10$ くらいであれば、むしろもう少し精度がわるくとも、迅速な、または安価な分析をやった方がよい場合がある。これと同様に現在行っている測定法の精度が 3%

1) 独立でない場合もあるから注意。

2) 3.2.1 分散の加成性参照。

くらいもあるときには、サンプルングの精度を1%以下に小さくすることも合理的でない。要するに、目的とするデータの精度 σ に応じて、サンプルング法及び測定法の精度を、統計的、技術的、経済的にバランスのとれた合理的な大きさに割り当てるべきである。

従来ややもすると分析の精度、測定法の精度を小さくすることのみを考える傾

向があるので、これに関連してもう少しのべておこう。それは工場で測定をする場合に、たとえば工程を管理しようとする場合に、たとえば工程のバラツキが $\sigma_p=3\%$ あるとすると、 $\sigma_M=0.1\%$ くらいの測定法（サンプルング誤差も含めて）を使用する必要はなく、 $\sigma_M=1\%$ 以下ならば十分である。すなわち

$$\sigma^2 = \sigma_p^2 + \sigma_M^2 = 3^2 + 1^2 = 10 \quad \therefore \sigma \approx 3.16 \%$$

σ_p の大きさの5%くらいしか大きくならぬから、たとえば管理図に書いた場合でも、管理限界の巾が僅かに大きくなるだけで、工程を管理するためにはあまり大きな影響を及ぼさない。この関係を頭において工程を管理するための測定法を決定すべきであろう。これは従来われわれの行っていた測定や試験は多くの場合、製品を保証するための試験、すなわち製品が消費者の要求に適するか否かの試験、保証試験 (suitability test) である。この場合でも勿論必要以上の精度は必要ないが、工程を管理して行くための試験、すなわち管理試験 (control test) と保証試験とは、はっきり区別すべき場合が多い。管理試験 (場合によってはサンプルング誤差も含めて) は工程の管理状態、すなわち σ_p に応じた精度を持てば十分であり、精度や正確さも保証試験に比して悪くてもよい場合が多く、相対的な値でもよいかから、迅速に安価に行える方がよい場合が多いのである。特に従来の化

表 3.1 二つの独立な原因とそれより得られるデータのバラツキ (標準偏差) との關係

原 因		結 果
σ_1	σ_2	σ
1.0	0	1.00
1.0	1/10	1.00
1.0	1/5	1.02
1.0	1/3	1.05
1.0	1/2	1.12
1.0	1	1.41
1.0	3/2	1.80

学分析のように時間と経費を要するものは、保証試験と管理試験とを区別した方がよく、最近欧米において非常に発達をとげている物理化学的な試験方法も、管理試験という面からの要求により発達したものが多いうのである¹⁾。

このように、われわれは各種バラツキとの相対的な大きさを比較して、もっとも合理的に、経済的な測定法を使用するように心掛けなければならない。

第4にサンプルング及び測定法の精度について考えなければならないことは、その測定値を用いる目的との関係である。すなわちえたデータの精度と、このデータを用いて行動 (action) を起すに必要な精度との関係である。たとえば、旋盤で丸棒を削ったときに、0.01 mm までの太さの差を知れば十分なときに、従来技術者は非常に精密な測定法、たとえば0.0001 mm 単位で測定できるメーターを使って安心であるとしている。しかしこれでは時間、経費その他不経済な場合があろう。また石炭の灰分を測定して、9.58 % と出たとしても、サンプルング及び測定法の精度から考えると、データの精度が±1 % くらいであるとすれば、このように細かい数字を記入してもあまり意味がないし、また工場の受入価格の決定にも、またボイラーや発生炉の原料に使用する場合でも、技術的に考えて見ると、小数点以下、特に2桁目はあまり意味がない場合が多いであろう。あるいは鉄鉱石が1万トン入ったときに、もし取引が水分を差引いた乾燥重量で取引するような場合には、重量の測定の精度が2 % くらいあるときには、水分を測定するための鉄鉱石のサンプルング並に測定法を含めた精度はトン数にもよるが、たとえば1 % 以下くらいにしておけば十分であろう。そして測定値も乾燥重量としては10トンあるいは1トン単位くらいにあらわしておけば十分であろう。このようにデータにあらわす必要な桁数は、そのデータの精度または使用目的を考えて、合理的な桁数に4捨5入するようにあらかじめ双方できめておけばよいのである。

以上要するに、従来われわれが多く行っていた、ただ数字的に多くの桁数のデータが得られたからといって、これを必要以上に並べておいたり、

1) Quality Control in Process Industries: J. G. Dobson: Ind. Eng. Chem. Vol. 43, No.12 p.2695 (1951 Dec.) 参照。

あるいは使用目的を考えずに、またサンプルング法や分析法の精度の向上を考えたりしていた態度は、かえって時間や経費の浪費である。また数字の桁数が多すぎるといことは計算間違いをしておかえって誤差を大きくしてしまう原因となったり、あるいは時間の浪費となる。

このようにいろいろのべて来たが、迅速に、精度よく、且つ安価に測定できる測定法があればこれを採用することは一向に差支えないのは勿論である。科学ならびに技術の進歩と共に、次第にそれに相応したサンプルング並に測定の精度の改良をはかるべきである。

3・2・1 分散の加成性

ある母集団 (μ, σ_p^2) があるときに、これから1つのサンプルをとり、測定をしてデータをうるのであるが、このときのデータ x はどのような構造をとるであろうか、次の式であらわされる。

$$x = \mu + x_s + x_M \quad (3.2)$$

ただし x_s : サンプルングによる μ からのへだたり

x_M : 測定誤差によるそのサンプルの真の値からのへだたり

μ は一定の値であり、 x_s, x_M はサンプルングごと、測定ごとに変化する値である。いまサンプルングがランダムに行われており、測定も管理されており、カタヨリもないものとすれば、 x_s の分散 σ_s^2 は σ_p^2 となり、 x_M の分散は σ_M^2 (測定の精度) となる。またもしサンプルが、 n コのサンプル単位を集めた混合サンプルであれば (たとえば石炭を1スコップを1インクリメントとして、 n 杯集めて1サンプルとした場合) $\sigma_s^2 = \sigma_p^2/n$ となる。

このような関係があるときに、1サンプルごとのデータのバラツキ、すなわち x の分散 σ^2 は次の式であらわされる。 μ は一定の値であるので

$$\sigma^2 = \sigma_s^2 + \sigma_M^2 \quad (3.1)$$

これには x_s と x_M とが互に独立 (independent) であるという条件が入っている。互に独立であるということは、 x_s が大きくなると、 x_M が大きく (あるいは小さく) なることがない、すなわち相関関係がないときには (3.1) 式が成立する。一般にサンプルングと測定の精度とは相関関係がないから、あるいはあっても非常に僅かであるから、この式が成立すると考えて差支えない場合が多い。この性質は分散が加算的に加えられることを

示しておりサンプルングや測定を論ずるときに非常に重要な性質であるので一般的な性質を次にのべる。

1. 和の誤差分散 n コの互に独立に変化する変数 x_i の和を X とすると

$$X = x_1 + x_2 + \cdots + x_n$$

X の分散 $V(X)$ は次式であらわされる。

$$\begin{aligned} V(X) &= V(x_1) + V(x_2) + \cdots + V(x_n) \\ &= \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2 = \sum \sigma_i^2 \end{aligned} \quad (3.3)$$

したがって同じ母集団からのサンプルであれば

$$V(X) = n\sigma^2 \quad (3.4)$$

2. 差の誤差分散 2つの変数 x_1, x_2 の差を X とすると

$$X = x_1 - x_2$$

$$V(X) = V(x_1) + V(x_2) = \sigma_1^2 + \sigma_2^2 \quad (3.5)$$

3. X がいくつかの変数の直線関係、すなわち 1 次式であらわされるときには、 k, a, b, \dots を常数とすると

$$X = k + ax_1 + bx_2 + \cdots$$

$$V(X) = a^2 V(x_1) + b^2 V(x_2) + \cdots \quad (3.6)$$

1) 3) をむすびつけると、同じ母集団 (μ, σ^2) からの n コのサンプルの平均値 \bar{x} の分散は、総和を X とすると

$$\bar{x} = \frac{X}{n} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

$$\therefore V(\bar{x}) = \frac{1}{n^2}(\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2)$$

しかるに $\sigma^2 = \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2$

$$\therefore V(\bar{x}) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n} = \frac{1}{n^2} V(X) \quad (2.27)$$

4. 以上の関係を一般式であらわすと、 X が互に独立な x_1, x_2, \dots という変数の関数であらわされる時、 X の誤差分散は

$$X = f(x_1, x_2, \dots, x_n)$$

$$V(X) \approx \left(\frac{\partial X}{\partial x_1}\right)^2 V(x_1) + \left(\frac{\partial X}{\partial x_2}\right)^2 V(x_2) + \cdots \quad (3.7)$$

標準誤差 σ_X は

$$\sigma_X \simeq \sqrt{V(X)} = \sqrt{\left(\frac{\partial X}{\partial x_1}\right)^2 V(x_1) + \left(\frac{\partial X}{\partial x_2}\right)^2 V(x_2) + \dots} \quad (3.8)$$

1) 2) 3) はこの特別な場合である。ただし (3.8) 式は、1)~3) のように線形の函数のときには、そのままあてはまるが、他の場合には近似式で、各 x_i の変化が各 x_i の真の平均値 μ_i の約 20% 以内のときにあてはまると考えてよい。

相関関係のある場合には、面倒な式になるので簡単な場合のみを示す。

x_1, x_2 という変数が母相関係数 ρ という相関関係をもつときには

$$X = x_1 \pm x_2$$

という X の分散は次式であらわされる。

$$V(X) = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2 \quad (3.9)$$

例1] 1,000トンの石炭が入荷した、平均湿分は約8%である。いま石炭の秤量の精度 $\sigma_W = 5$ トン、湿分のサンプリングと測定の精度 $\sigma_M = 0.5\%$ であるとする、乾燥炭の重量推定値 X の精度 σ_X はいくらになるか。

石炭の秤量を W トン、湿分を $M\%$ とすると

$$X = W \left(1 - \frac{M}{100}\right)$$

$$\begin{aligned} V(X) &\simeq \left(\frac{\partial X}{\partial W}\right)^2 V(W) + \left(\frac{\partial X}{\partial M}\right)^2 V(M) \\ &= \left(1 - \frac{M}{100}\right)^2 V(W) + W^2 V(M) = 0.92^2 \times 5^2 + 1000^2 \times 0.005^2 \\ &= 21.16 + 25.00 = 46.16 \end{aligned}$$

$\therefore \sigma_X = 6.79 \simeq 6.8$ トン

例2] ある化学薬品を乾燥重量で各 100g 瓶詰するのに、入れ目を平均 0.5g 程度におさえ、100g 以上入っていることを保証したい。水分は平均 5% で、その工程とサンプリングと測定を合せた精度は $\sigma_M = 0.3\%$ である。湿重量で入れ目の精度 σ_W をどのくらいにおさえたらよいか。

瓶詰重量は、湿重量で平均 $W = 100.5/0.95$

入れ目平均乾燥重量で 0.5g とし、100g 入を保証するためには $3\sigma_X = 0.5$ とすればよいであろう。

$$X = W \left(1 - \frac{M}{100} \right)$$

例1) と分散は同じ関係になるので、それぞれの値を代入すると

$$(0.5/3)^2 = (0.95)^2 V(W) + (100.5/0.95)^2 (0.003)^2$$

$$(0.95)^2 V(W) \rightleftharpoons 0.0278 - 0.1007 < 0$$

しかるに分散 $V(W)$ は負になることはないから、このようなことは不可能で、水分測定の精度を小さくしなければこの入れ目では 100g 入を保証することはできない。そこで平均 1.0g 程度の入れ目を許すならば同様に

$$(1/3)^2 = (0.95)^2 V(W) + (101/0.95)^2 (0.003)^2$$

$$V(W) \rightleftharpoons 0.0104$$

$$\sigma_W = 0.102\text{g}$$

すなわち湿重量で平均 $101\text{g}/0.95 \rightleftharpoons 106.32\text{g}$ として、精度 0.10g くらいに秤量しないと上の保証を行うことは困難である。

この場合はむしろ水分をよく管理し、またサンプルングや測定の精度を向上させて、 σ_M を小さくできれば、平均入れ目を小さくできよう。

3・3 信頼性、精度、正確さ

われわれがサンプルング、測定等の誤差を論ずるときには、従来はただ漠然と誤差(error)といていたが、これを3つにわけて信頼性(reliability)、精度 (precision) 及び正確さ (accuracy) をおのおの考えなければならぬ。

信頼性とは、サンプルングや測定が常に同じように、作業標準通りに行われているかどうか、したがってそのサンプルングや測定が信頼できるかどうかということである。いいかえると、これらの操作が管理状態にあるかどうか、操作に何か間違いや異状なこと (assignable cause) が時々おこっているかどうかということである。これらの操作が管理状態になれば、折角えたデータも、技術的に見て異状な値がでると、従来よく現場でいわれるように、「これはサンプルングがわるいのだ」「これは測定がわるいのだ」と簡単に片付けられてしまって、工程や製品に異状があっても気付かずに過ぎてしまい、何のためにデータをとっているのかわからなくな

つてしまう。したがって、サンプリングや測定を論ずる場合にも、これらの操作を管理状態に管理することがもっともたいせつなことである。このためには、サンプリングや測定が管理状態にあるか否かをチェックできるように実施しておいた方がよい。たとえば従来よく行われているように、化学分析を2回繰り返し行っているのは測定の管理状態をチェックしていることになる。しかしサンプリングについては、従来多くの場合チェックサンプリングが行われていないが、現場で実際にサンプリングを行う場合、作業標準通りに行われていない場合がよくあるので、これを管理するために経済的に事情が許すならば、チェックサンプリングを行っておいた方がよい。

この信頼性は、繰り返しの測定、あるいはチェックサンプル間のデータを群にとって、管理図を描き、特にその R 管理図を検討するのが、もっともよい。

サンプリングや測定に信頼性がなければ、すなわちサンプリングや測定という工程かが管理状態を示していなければ、その正確さや精度を推定したり、予想したりすることはできない。したがってサンプリングや測定の担当者は、まずその操作上の誤りを除去し、作業標準どおりに行い、あるいは行わせることに努力すべきで、しかる後正確さや精度を論ずることができるという段階になるのである。

正確さと精度は、従来は正確度、精度、誤差などという言葉で漠然と誤差のことを指して、従来大体同義語として区別せずに用いられているが、これは区別して用いなければならない。精度と、正確さ(正確度またはガタヨリ、バイアス)とは全く性質のことになったもので、これははっきり区別しなければならない。誤差 (error) とは、信頼性、精度、正確さを含めたものを、すなわち真の値からのデータのヘダタリ (以上3つの原因によるものを全部含めて) をいうのである。

精度 (precision) とは、同じサンプリング法で、1つの母集団から繰り返しサンプリングしたとき、あるいは同じまたは均一なサンプルを同じ測定法で繰り返し測定したときのデータの平均値からのバラツキかたである。母集団の平均値を推定するときにはサンプルの平均値のバラツキかたである。一般に完全なランダムサンプリングが行われているときのバラツ

キズなわち分布のひろがりの度合が、サンプルングの精度である。通常精度をあらわすには次の方法のいずれかを用いる。

- 1) 標準偏差 σ ; $\hat{\sigma}$ 2) 分散 σ^2 ; σ_s^2 .
- 3) 変動係数 $C = \frac{\sigma}{\mu}$; $c = \frac{s}{\bar{x}}$ または $c_s = \frac{\sigma_s}{\bar{x}}$
- 4) 範囲または平均範囲 R ; \bar{R} (ただし繰り返しの数を指定する)
- 5) 最大範囲 $D_1\bar{R}$ (信頼度をかえてもよい、繰り返し数を指定する)
- 6) 信頼区間 $\pm a_1\sigma_2$; $\pm a_1c_2$; $\pm A_2\bar{R}$ 等々 (ただし a_i 等の値は繰り返し数と必要とする信頼度によりきまる係数)

以上いずれを用いるかは、その場合、用途によりことなる。本書でも必要に応じて使いわけが、いずれの精度を用いているかは、識別して読んでいただきたい。また精度を指定、表示するときには、いずれのあらわしかたによるかを、示しておかなければならない。

特に平均値や総計などの精度を示すときには、信頼区間の考えかた¹⁾と似ている 6) の方法を多く用いる。サンプルングでは、平均値や総計を問題にすることが多いので少し説明を加える。たとえば入荷したボルトの不良数の総計、石炭の灰分や薬品の純分の平均値などを 99% の信頼度で $\pm 5\%$ 、 1% 、 0.5% 、の精度で知りたいというように用いられる。これはサンプルからもとめた平均値や総計の推定値の $\pm 5\%$ 、 1% 、 0.5% 内に、真の平均値や総計がある確率が 99% という意味である。たとえば信頼度 99% で $1\% = a_1\sigma_2\%$ となるように、 a_1 およびサンプルングのやりかたをきめればよい。すなわちこの精度と、信頼度をいかにとるかという要求によりサンプルング法が決定されるのである。たとえば、鉄鉱石の鉄分の平均値を $\pm 1\%$ の精度で知るようにサンプルングしたいときには、必ず信頼度²⁾ (信頼係数 confidence coefficient) (確率) を指定しなければ、サンプルング法は決定できないのである。信頼度としては、その経済的、技術的要求に応じて、通常 99%、98%、95%、90% などが用いられる。

例 1] あるロットより 10 コのランダムサンプルをとり、測定して平均値 10.5 mm ; 標準偏差 0.2 mm をえた。このサンプルング法による

1) § 2・3・4 参照。

2) 信頼性と混同しないように。

平均値の99%信頼度の精度はどのくらいになるか。

精度を $\beta\%$ とすると, t 分布より β は次式であらわされる。

$$\pm\beta = \pm t_{\alpha} \frac{s}{\sqrt{n-1}} = \pm t_{\alpha} \frac{\sigma_s}{\sqrt{n}} \quad (3.10)$$

ただし t_{α} : 危険率 $\alpha = (1 - \text{信頼度})$, 自由度 f の t 表の値. この場合は

$$f = n - 1$$

α : 1-信頼度

n : サンプルの大きさ

s : 試料標準偏差

σ_s : 不偏分散の平方根

(3.10) 式において, $\alpha = 100 - 99 = 1\%$, $s = 0.2 \text{ mm}$, $f = 10 - 1 = 9$
 t 表より

$$t_{0.01}(f=9) = 3.25$$

したがって精度は

$$\beta = 3.25 \times 0.2 \text{ mm} / \sqrt{9} \approx 0.22 \text{ mm}$$

すなわちこのロットの真の平均値 μ は, $10.5 \text{ mm} \pm 0.22 \text{ mm}$ 内にあるということが, 99%の信頼度でいえる。

例2] あるロットの各サンプル間の標準偏差 σ は 0.2 mm であることがわかっている (ただし測定誤差を含めて). このロットより10コランダムにサンプルをとって測定すると, その平均値の信頼度99%の精度はいくらか。

この場合は, 母標準偏差 σ 既知の場合であり, 正規分布を用いる。

$$\pm\beta = \pm u_{\alpha} \frac{\sigma}{\sqrt{n}} \quad (3.11)$$

ただし u_{α} : 危険率 $\alpha = (1 - \text{信頼度})$ の正規分布表 (両側) の値

(3.11) 式において, $\sigma = 0.2$; $\alpha = 1\%$, $n = 10$; $u_{0.01} = 2.58$

$$\therefore \beta = 2.58 \times 0.2 \text{ mm} / \sqrt{10} \approx 0.16 \text{ mm}$$

このようなサンプリング法を100回繰り返して行えば, 99回は平均値を $\pm 0.16 \text{ mm}$ 以内の誤差で推定できる。更に正規分布表より, 100回中95回は $\pm 0.12 \text{ mm}$, 90回は $\pm 0.10 \text{ mm}$, 80回は $\pm 0.08 \text{ mm}$, 50回は $\pm 0.04 \text{ mm}$ 以内の誤差で平均値を推定できることを示している。

注意 1) 例 1) は母標準偏差未知の場合、例 2) は既知の場合であるが、既知の場合の方が、平均値推定の精度ははるかによくなる。

2) 母標準偏差既知の場合とは、たとえば工程やサンプルングや測定が管理状態にあつて、いつも大体同じようなロットが出てくるときがこれに相当する。管理状態の工程では、このような利点がある。

例 3) 瓶詰の硝酸 500 本が入荷した。予備調査でその 5 本をランダムに抜きとり、その正味重量を測定したら、標準偏差 $s = 0.10$ kg であつた。

この硝酸ロットの一瓶の平均重量を ± 0.10 kg の精度で、その信頼度 99% で求めたい。サンプル何本をとればよいか。

この場合は予備調査の結果を利用して、所要精度をうるための本調査のサンプル数をきめる問題である。このときには次式を用いる。

$$\pm \beta = \pm t_{\alpha} \frac{\sigma_s}{\sqrt{k}} = \pm t_{\alpha} \frac{s}{\sqrt{k}} \sqrt{\frac{n}{n-1}} \quad (3.12)$$

ただし t_{α} : 信頼度 $(1-\alpha)$ の自由度 $f = n-1$ の t 表の値
 n : 予備調査のサンプルの大きさ
 k : 本調査のサンプルの大きさ
 σ_s, s : 予備調査での値

この例では $\beta = 0.10$, $s = 0.10$, $n = 5$, $t_{0.01} (f = 4) = 4.60$

$$\text{したがつて} \quad 0.10 = 4.60 \times 0.10 \times \sqrt{\frac{5}{4}} / \sqrt{k}$$

$$\therefore k = \frac{4.6^2 \times 0.1^2 \times 5}{0.10^2 \times 4} \approx 26.5$$

すなわち 27 本の正味重量を測定すればよい。

注意 1) この場合硝酸入荷量の総計を $\pm \beta$ の精度 (信頼度 $1-\alpha$) で知りたいときには、次式による。

$$\pm \beta = \pm t_{\alpha} \frac{N\sigma_s}{\sqrt{k}} = \pm t_{\alpha} \frac{Ns}{\sqrt{k}} \sqrt{\frac{n}{n-1}} \quad (3.13)$$

ただし N : ロットの大きさ

上例で総入荷量を 50 kg の精度 (信頼度 99%) で知りたいためには

$$50 = 4.60 \times 500 \times 0.10 \times \sqrt{\frac{5}{4}} / \sqrt{k}$$

$$\therefore k = 26.5$$

すなわち27本とれば十分である。

2) 予備調査の数が多い場合には、 t_{α} が小さくなるから、本調査の数は少なくなる。しかし n が15以上になると n を大きくした割に効果は少ない。 n が40以上になれば上 t_{α} のかわりに u_{α} 、 σ_s のかわりに s を用いて差支えない。

3) 混合試料の場合 化学分析のように、インクリメントなどでサンプルをとり、これを混合試料として分析する場合にも、予備調査で n コのインクリメントをランダムにとり、その σ_s, s をもとめておけば、信頼度 $(1-\alpha)$ 、精度 β の混合試料をえたいためには、上式より k コのインクリメントをとればよいことがわかる。しかしこの場合 §6 でのべるように、分析や縮分の精度を分離して別に考えなければならない。

以上の例でわかるように、ランダムサンプリングでは平均値の精度は標準偏差とサンプルを、何コとるかということにより支配されるが、 σ がきまるとサンプル数が多くなれば精度はよくなるが、一般に不経済になる。これをさけるために、管理状態にして σ を小さくしたり、統計的にサンプル数を小さくする工夫が行われる。

最大範囲を用いる方法は、繰り返し測定を行った場合などによく用いられる方法で、従来「2回の測定の差が1%以上であれば、測定をやり直す」という表現がよく行われているが、これは最大範囲を超えたならば——どの再現精度、繰返精度¹⁾をとるかに注意——その測定には、その操作中に何か異状があったのであるから、すなわち操作の管理状態が破れているのであるから、それらのデータは信頼性がなく、したがってそのデータは2つともすてよという意味になっていけば正しい。

変動係数は市場調査のような調査の場合によく用いられる方法で、工場でも調査関係には用いると便利な場合がある。

σ, σ^2, C, R などはいずれも母数について考える場合で、サンプリングの理論式を展開する場合には、便利なあらわしかたで、本書でもよく用いる。

正確さ(accuracy)とは、同じサンプリング法である母集団から繰り返し

1) §3-4-2 参照

サンプルングしたときに、その平均値の、母集団の真の平均値からのカタヨリ、あるいは同じサンプルを同じ測定法で繰り返し測定したときにえられるべき平均値の、サンプルの真の平均値からのカタヨリをいう。正確さは、**正確度**、**バイアス** (bias)、または**カタヨリ**などともいわれる。たとえば、良いものばかり選んでサンプルングしたり、悪いものばかりえらんでサンプルングしたり、粉塊の混合物があるときに塊を多くとるようなサンプルングをしたりすると、真の値からカタヨった値をうることになり、バイアスが入る。このようなサンプルングをバイアスのあるサンプルング、あるいは不正確なサンプルングという。また測定の場合にも甲という検査工は常に +0.01 mm 大きく目盛をよむ癖があるとか、ノギスの 0 点が -0.1 mm にあるのを知らずにそのまま用いたり、ピュレットやピペットを補正しないで用いたりするとバイアスが入り、不正確なデータを得ることになる。

この正確さ、バイアスは信頼性と精度を検討してから論ずべき問題である。すなわちたとえばある既知の成分の標準化合物を分析工に繰り返し分析させたときに、その分析が管理状態にあり、その精度が判明した後に、はじめて真の値と、測定の平均値が有意にことなっているかどうかを、統計的に (たとえば t -検定または管理図による有意差の検定などで) 検定できるのであり、この検定においてもし有意な差が認められなければ、カタヨリがある、不正確であるとはいいい切れないのである。これは假説検定の考えかたから明らかであろう。

次に図 3-1 に精度と正確さとの関係をいろいろの場合について、示した。この図はサンプルングについても、測定についても同じである。図 3-1 において μ は真の平均値を示し、山形の分布は、実際のデータから推定した分布型である。分布の巾の狭いのは精度のよいことを示し、巾の広いのは精度のわるいことを示している。

A 図は精度もよいし、この精度 (たとえば s) を用いて、 \bar{x} と μ と有意差の検定を行っても有意な差が認められないから、カタヨリがあるとはいえない良好な場合である。

B 図は精度はよいが、 \bar{x} と μ には有意差がはっきり認められる場合で、相当なカタヨリがあるから、カタヨリの原因をさがして修正する必要がある。

C 図は精度がよくないためか、 \bar{x} と μ は一見ことなっているように見え

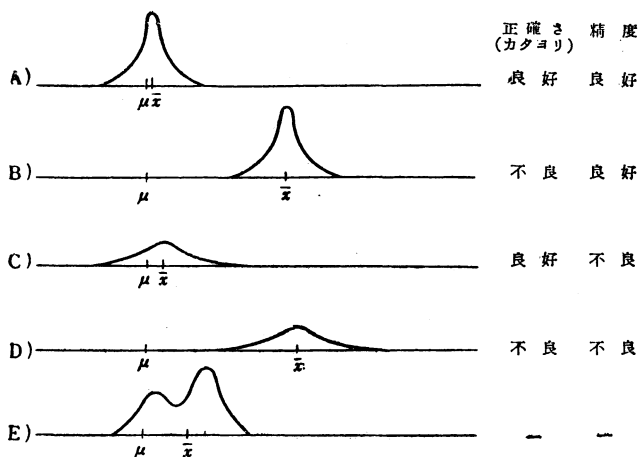


図 3.1 精度と正確さ (カタヨリ)

るが有意差はみとめられない。この場合は、現在の段階では、精度をよくするように努力すべきで、精度がよくなってからカタヨリを問題にすべきである。

D図は両方とも悪いことがはっきりわかる。精度、カタヨリ両方に対して処置をとる必要がある。以上のA、B、C、D図はいずれも信頼性のあるデータからの図である。しかし逆にこのような山型を示してもそのデータが信頼性があるとはいえない。

E図は、たとえばサンプリング法や精度が管理されていない状態で、信頼性のあまりない状態であるから、まずこれを管理する必要があることを示している。

以上でわかるように、サンプリングや測定の誤差を論ずるときには、信頼性、精度、正確さの順に統計的に検討して行かなければならない。特にサンプリングにおいては信頼性や正確さを解決するには、主として技術的に、精度をよくするには、相当統計的技術的に工夫を要する場合が多い。

3.4 測定誤差

測定誤差については、信頼性、精度、正確さ、及びサンプリング誤差と

測定誤差の分離について考えよう。

測定でバラツキを生ずる主な要因は a) 測定者 b) 測定器 c) 測定法(試薬等を含む)である。このほか勿論測定する材料、サンプルングも一緒になって、データのバラツキとなる。いろいろな原因を、統計的に異状な原因(assignable cause)をおさえて、測定法の信頼性を確立し、精度や正確さをきめて行く方法を論じよう。

3・4・1 測定の信頼性

測定についてもっともたいせつなことは§3・3にのべたように、その測定が信頼できるかどうかということである。測定が信頼できなければ、測定に未だ管理されない異状な原因が残っていれば、その測定の誤差は不定で、精度や正確さをもとめることはできない。

この信頼できるか否かは、統計的に検定するいろいろな方法があるが、ごく簡単で、しかも技術的にもっとも適当なのは管理図である。

管理図を用いる場合にもっとも注意しなければならないのは、何を群にとるかということ、一般に測定に繰り返しのあるときはこの繰り返しを群にとる。この場合同時に行った繰り返しを群にとるか、別のときに行った繰り返しを群にとるかにより群の意味が全くかわってくるから注意を要する¹⁾。化学分析の場合は多くの場合2回ずつ繰り返し分析を行っているからこれを群にとればよい。物理的測定の場合も同一試料につき繰り返しの行えるときにはその繰り返しを群にとる。破壊試験等で繰り返しの行えないときには、できるだけ同じと思われるサンプル、たとえば時間的または場所的に同時につくられたなるべく均一な試料について、いくつか測定を行った結果を群にとる。このように群のきめ方は、管理図により工程を管理するときと同様に、技術的、統計的な工夫が必要である。

このようにして書いたR管理図が管理状態であれば、その群の中には、異状な原因がなく、信頼できることを示している。すなわち、甲という測定者が、同時に同じ装置、同じ方法で同じ試料を繰り返し測定したデータを群にとっているときに、そのデータからかいたR管理図が管理されていれば、甲については、そのようにして同時に行った測定は信頼できることを意味しているのである。これだけでは甲が別の日に、または別の装

1) この点については3・4・2 繰返精度と再現精度のところ参照。

置、方法で測定したときの信頼性を保証することはできない。もしこのような信頼性を見たいときには、同じ試料を別の日に、別の装置で行った測定値を群にとればよい。あるいは同じ試料、または均一な標準試料を、同じ日に同じ条件で繰り返し測定を行い、また別の日にその試料を繰り返し測定し、これらのデータから、同じ日に行った繰り返しを群にとって $\bar{X}-R$ 管理図をかいて、 \bar{X} 管理図も R 管理図も管理されていけば、同時に行った測定は勿論のこと、別の日に行った測定も信頼できることを示している。このほか、どのような測定の信頼性を見たいかにより、いろいろな群のとり方や管理図のかきかたがあるから、その目的に応じて研究されるとよい。この場合、通常の3シグマ限界でもよいが、サンプリングや測定の管理ということを考えると、2シグマ限界を用いた方がよい場合もある。

この管理図をかくにあたっては、工程を管理するときとことなり、必ずしも $n=3\sim 6$ とする必要はない。繰り返しの回数によって、 $n=2$ でも3でもよい。また群の数は多いほどよいが10組くらいあれば一応管理図に書いて見るとよい。管理図を書いて見て R 管理図が一応管理状態にあれば、一定の繰り返しの数 n に対して、 $D_4\bar{R}$ 及び $D_3\bar{R}$ よりその測定者の管理限界を出し、これをその測定者に示しておいて、測定の繰り返しの結果の最大値と最小値との差が、この値以上のものは測定を全部やり直すことにしておく。この場合、繰り返しのデータの1つが、たしかにある原因によってバラツいたものであるということが技術的にはっきり判明した場合はそのデータだけすてることもよいが、はっきりしないときは全部やり直すべきである。

データの棄却についてもう1つ注意しておきたいことは、一般に実験や測定がすんでから、飛びはなれたデータがあると、それを勝手に、あるいは統計的に棄却することが行われている。勿論これも、そのデータの出る過程において、技術的に原因がはっきりした場合にはすてても差支えないが、このようなときは多くあとからこじつけることが多いから可及的やめ

- 1) この場合、従来よく行われていた、たとえば2回の繰り返しの差が1.0%以上あるときは、もう一回測定を行い、近い方の値2つの平均をとるという考え方はまちがっている。これによりカタヨリが入る。この1.0%が管理限界 $D_4\bar{R}-D_3\bar{R}$ に相当するならば、原則として2つともデータをすてて新たにはじめからやり直すようにしておいた方がよい。

の方がよい。またもしデータを出す過程において一寸異状があるなど思ったら、そこで測定を中止すべきである¹⁾。従来のようにそのまま最後までやってみて、よければ採用するし、わるければ棄却するという態度は正しくない。このようなことをすると測定にカタヨリが入りやすい。これはちょうど管理図法において、管理線を将来の工程管理のために延長する場合に限界外に点が出た場合でも、その原因が不明のときには、その点も含めて \bar{X} 、 \bar{R} 等を計算しているのと同じ考えかたである。

例1) ある会社の試験係において、職長、伍長、新入工員の3人について、同じ標準試料の強度を毎日2回ずつ繰り返して測定させた結果を人別に層別した管理図に書いて見ると図3・2のようになった。これを検討して見よう。

職長は \bar{R} は0.21で比較的よいが、 R 管理図に管理限界外の点が1点、 \bar{X} 管理図で上に2点出ている。すなわちこの管理図は管理状態になく、職長の測定だからといって、 \bar{R} は小さいが、信頼しうるデータではない。伍長は \bar{R} も小さく、 \bar{X} 、 R 管理図ともよく管理されている。このデータは十分信頼しうる。新入工員は、はじめの中は \bar{R} も大きく、 \bar{X} のパラッキも多く、共に限界外に出る点があつて、とても信用して使えなかつたが、最近では \bar{R} も小さくなり、限界外に出る点もなくなり、熟練して来たことを示し、まだ精度(\bar{R})はやや大きい、一応信頼して使うことができるようになった。

よく調べて見ると、職長は馴れているために、時々測定法を横着していることがわかり、また最近勤務状況もあまりよくなかつたので、他の重要でない勤務に配置転換した。

この場合、測定の精度は \bar{R} であらわされる。 \bar{R} の小さいほど精度はよい。正確さは、 \bar{X} 管理図の \bar{X} と、標準試料の既知の強度 $\mu = 17.60$ kg/cm²との差によって示されるが、 \bar{R} を用いて、管理図による有意差の検定を行つて、有意差がなければ、不正確、カタヨリがあるとはいえない。伍長の測定は、よく管理されているので、有意差の検定をやつて見よう。

1) はじめからこのデータはすべて練習のために最後までつづけるのなら勿論よい。難しい実験ではこのようなことが必要なこともよくある。

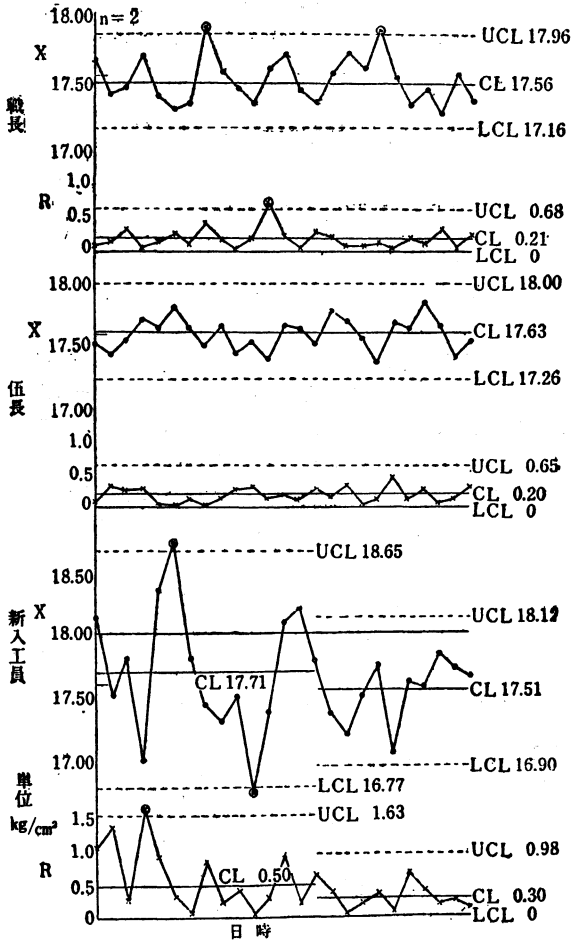


図 3.2 測定法の管理図

標準試料 (強度 17.60 kg/cm²) による測定法の管理

$$|\bar{X} - \mu| \geq A_2 \bar{R} / \sqrt{k} \quad (3.14)$$

ただし、 k は \bar{X} の点の数
ならば有意差ありといえる。

この例では

$$\begin{aligned} |\bar{X} - \mu| &= 17.63 - 17.60 = 0.03 \\ A_2 \bar{R} / \sqrt{k} &= 1.88 \times 0.20 / \sqrt{25} = 0.075 \\ \therefore |\bar{X} - \mu| &< A_2 \bar{R} / \sqrt{k} \end{aligned}$$

したがって、この伍長の測定は、標準値と有意な差があるとはいえない。
すなわち、この管理図からは不正確、カタヨリがありとはいえない。

注意 1) 管理図を使って平均値の有意差検定を行うときには、原則としてその管理図が管理状態を示していなければならない。しかし僅かに管理状態から外れている場合なら、有意差検定を行っても差支えないが、技術的には、まず管理状態にするように、次に精度 \bar{R} を小さくするように努力すべきである。一般に管理状態になるように努力し、その効果があがってくると、自ら \bar{R} も小さくなってくるものである。正確さは、常にこの \bar{R} とにらみ合せて、正しくなるように努力すべきである。

2) 伍長について、精度を標準偏差及び最大範囲の形でもとめると次のようになる。

$$\begin{aligned} \hat{\sigma} &= \bar{R} / d_2 = 0.20 / 1.13 = 0.177 \\ D_4 \bar{R} - D_3 \bar{R} &= 3.27 \times 0.20 - 0 = 0.654. \end{aligned}$$

3.4.2 測定の精度

測定の精度とは、同じサンプル、または十分に管理された工程でつくられた均一なサンプルについて、管理状態で繰り返し測定を行ってえた各データ間のバラツキの大きさを問題にする場合が多い¹⁾。測定値の平均値のバラツキを問題にすることもある。

測定の精度をあらわすには通常、標準偏差、分散、範囲(ただし繰り返しを指定して)または最大範囲 ($D_4 \bar{R}$) であらわす。従来よく行われていたように平均偏差 (mean deviation) とか確率誤差 (probable

1) サンプルングの精度は、平均値のバラツキを問題にする場合が多い。

error)¹⁾であらわすのは、精度を統計的にいろいろ検討するには不便である。

精度をもとめるには、いろいろな統計的方法があるが²⁾、多くの場合は管理図によるのが便利である。各測定値についての精度は繰り返しを群にとって、 R 管理図が管理されていれば

$$\hat{\sigma} = \bar{R}/d_2, \quad \text{または} \quad D_4\bar{R} - D_3\bar{R} = \text{最大範囲}$$

よりもとめられる。3・4・1の例を参照されたい。

この測定の精度について注意しなければならぬことが2つある。その一つは、破壊試験の場合の測定の精度とは何ぞやということ。他は繰返精度と再現精度の問題である。

第1の破壊試験の場合の測定の精度とは、真の測定の精度ではない。同じサンプルにつき繰り返し測定できないので、サンプリング誤差と測定精度とを分離できない。そこで均一なサンプルを各々標準試験法で、よく管理して測定し、その測定値の精度をもって、測定の精度と定義することになる。

たとえば抗張力の測定のような場合には、十分に管理された工程で同時につくられたできるだけ均一な材料よりサンプルを何箇所か作成して、これをよく管理した状態で、標準試験法で繰り返し測定したときのデータの精度をもって、測定の精度と定義する。すなわち破壊試験の場合の測定の精度とは、サンプルのバラツキ（小さい場合が多いが、大きい場合もある）を含めた測定値のバラツキをもって精度と定義するのである。化学分析の場合にも同様である。化学分析のときに、たとえば鉱石や、結晶の薬品を細かく粉碎し縮分して、よく混合し、これからサンプル1gをとって分析する場合もこの最後のサンプリングの精度は測定の精度に含まれてしまう。混合が十分ならばサンプリングの精度は非常に小さくなるが、このようにして求めた精度をやはり測定の精度と定義しているのである。

第2に、同じ測定者が同じ測定器を用いて、同じサンプルにつき同時に測定を繰り返したときの精度と、条件がいろいろかわった状態で測定したときの精度とは全く意味がちがうということである。これは精度をもとめるのに、管理図を用いるときその群のとりかたに注意しなければならぬ。

1) その値より大きい値の出る確率が50%の値。0.6745σにあたる。附表1参照。

2) 3・4・4並びに3・4・5参照。

精度の性格は群のとりかたによりかわるとのべたのと同じことである。たとえば同じ条件で、繰り返し測定したときの精度を繰返精度(repeatability)という。これに対し多くの条件で、たとえば同じサンプルを別の日に、別の機械で測定した測定値間の精度を再現精度(reproducibility)という。精度を論ずるときにこの両者は区別しなければならない。再現精度には、たとえば測定器をかえたとき、人をかえたとき、試薬をかえたとき、日時をかえたとき、あるいはこれらを組合せたときなどいろいろ考えられる。また研究所間でどのくらいちがうかという問題もあろう。いいかえると再現精度は繰返精度のパラツキのほかに、人、装置、試薬、日時などの測定に誤差をあたえる原因によるカタヨリを含んだものである。したがって再現精度の方が繰返精度よりも大きい。ある1つの測定値のパラツキを考えると

$$\sigma_{rpd}^2 = \sigma_{rpt}^2 + \sigma^2_K \quad \sigma^2_K > 0$$

$$\therefore \sigma_{rpd} > \sigma_{rpt}$$

ただし σ_{rpd} : 再現精度の分散

σ_{rpt} : 繰返精度の分散

σ^2_K : 測定に誤差をあたえる σ_{rpt} 以外の原因によるパラツキ

たとえば化学分析などでは他研究室間の再現精度は同一研究室内の繰返精度の2倍くらいになるのが普通である。

測定に誤差をあたえる要因をすべて発見したり、おさえてゆくことが、経済的にも、技術的にも困難な場合が多いので、規格とか、仕様書等で精度を論ずるときにはある再現精度を用いた方がよい。

精度が等しいか否かは、カイ二乗またはF分布を利用して検定する。

例 A, B の分析精度は各200回の分析においてよく管理されていて各2回の繰り返しの差が $\bar{R}_A = 0.35$, $\bar{R}_B = 0.32$ である。有意差ありといえるか。いま $n=2$ の $d_2=1.13$

$$\hat{\sigma}_A = \bar{R}_A/d_2 = 0.35/1.13 = 0.314 \quad \hat{\sigma}_B = 0.32/1.13 = 0.287$$

$$F_0 = (0.314)^2/(0.287)^2 = 1.19$$

$$\hat{\sigma}_A^2, \hat{\sigma}_B^2 \text{ の自由度は最大各々 } 200 \text{ である。 } F_{230}^{200}(0.05) = 1.26$$

したがって差があるとはいえない。

〔注〕この検定は厳密には正しくないが、自由度を大きく見ても有意差ありといえないのであるから、実用的には差支えない。

3・4・3 測定の正確さ

測定の正確さとは、まえにのべたように、測定が管理状態で行われているときに、真の値からのカタヨリをいい、カタヨリがあるというためには精度と比較して、有意差があるか否かを検定しなければならない。

測定の正確さを論ずる場合に、次の3つのことが問題になる。1) 真の値とは何ぞや 2) カタヨリの検定 3) カタヨリの推定と補正。

1. 真の値とは何ぞや われわれは厳密に言えば、測定により真の値を知り得ないのが通常である。理論的に真の値を推定できる場合もあろうが、このようなことは少ない。そこで、ある材料を、十分信頼できる測定法、測定者で測定した値を、あるいは熟練者による測定値を真の値と定義する。あるいは理論的にもとめた値を、真の値と定義することになる。したがって、標準試験とか、標準試薬、標準試料というものが、きめられ、これを真の値と定義する。たとえば米国においては、ASTM において、標準通り測定された値を真の値と定義しているものもある。非破壊試験で同じ試料を繰り返し測定できる場合には、標準の測定値と、同一試料を工場などで行った測定値とを比較すれば、測定の正確さをもとめることができる。

ところが、たとえば抗張力、爆発試験のような物理的な破壊試験、あるいは化学分析のような破壊試験の際には、真の値を推定することは不可能に近い。このような場合には、できるだけ均一なサンプルを作成し、あらかじめ決められた標準測定法により、十分管理して行った測定値の平均値を真の値と決めて、これと各工場と同種のサンプルについて行った測定の平均値の差を比較して正確さを決めることになる。この場合に注意しなければならぬことは、このデータには各サンプル間のバラツキと、測定法によるバラツキとが一緒に入っている(交絡している confound) ので、サンプルが異種のもので、サンプルにカタヨリがあったり、同種のサンプルでも、サンプリングがランダムに行われておらず、カタヨっていると、これらのサンプルのカタヨリを、測定法のカタヨリと誤認してしまうことがある。したがって、破壊試験の測定の正確さをチェックするときには、特にそのサンプルの準備や取り扱いに注意しなければならぬ。

2. カタヨリの検定法 カタヨリが確かにあるか否かは、有意差検定

による。これには通常、管理図による検定、正規または t 検定、分散分析法などが用いられる。

a. 管理図による検定 同じ標準試料などを繰り返し測定したデータがある場合には、1回の繰り返しを群にして、管理図をかく。このときの群のとりかたは精度の項でのべた通り。管理図が大体管理状態を示していれば、次の式により検定を行う。

イ) 真の値 (μ) 既知の場合。次の条件を \bar{x} が満足すれば、確かにカテゴリーあり。

$$|\bar{x} - \mu| \geq A_2 \bar{R} / \sqrt{k-1} \quad (3.14)$$

ロ) 標準法、熟練者によりもとめた値と比較する場合は次式による。

$$|\bar{x} - \bar{x}'| \geq A_2 \bar{R}_M \sqrt{\frac{1}{k} + \frac{1}{k'}} \quad (3.15)$$

ただし \bar{x}' : 標準法あるいは熟練者によりもとめた平均値

k' : " \bar{x} の点の数

k : 現在問題にしている測定法、測定者などによる \bar{x} の点の数

\bar{R}_M : 標準法及びその測定法などの \bar{R} の重みつき平均

$\bar{R}_M = (k\bar{R} + k'\bar{R}') / (k + k')$ ただし群の大きさは等しいものとする。

A_2 : 群の大きさにより決まる係数

注意 1) この式は精度に有意な差がないときに用いる。

2) 標準法とその測定法と群の大きさ n (1回の繰り返しの大きさ) がことなるときには、 \bar{R} , \bar{R}' は係数 d_2 を用いて、同じ n の値に補正しなければならない。

3) サンプルの品質特性がサンプル毎に変化するときには、真の値、あるいは標準法の値との差をとって、これで管理図をかき、管理図が管理状態を示したならば、その \bar{x} が $\mu' = 0$ に対して、有意差があるか否かを検定すればよい。

例 §3.4.1 の例において、職長と伍長の測定の正確さに有意差があるか否かを検定して見よう。

1) §3.4.1 の例参照。

$$\begin{aligned} k = k' = 25, \quad n = 2, \quad \therefore A_2 = 1.88, \quad \bar{x} = 17.63, \quad \bar{x}' = 17.56 \\ \bar{R} = 0.20, \quad \bar{R}' = 0.21 \quad \therefore \bar{R}_M = \frac{(25 \times 0.20 + 25 \times 0.21)}{(25 + 25)} = \frac{0.20 + 0.21}{2} \\ = 0.205 \end{aligned}$$

$$|\bar{x} - \bar{x}'| = 0.07 \quad A_2 \bar{R}_M \sqrt{\frac{1}{k} + \frac{1}{k'}} = 1.88 \times 0.205 \times \frac{\sqrt{2}}{5} \approx 0.109$$

$$\therefore |\bar{x} - \bar{x}'| < A_2 \bar{R}_M \sqrt{\frac{1}{k} + \frac{1}{k'}}$$

したがって、職長と伍長の測定の正確さに有意な差があるとはいえない。

b) 正規分布または t 分布を応用する方法 工場のように、比較的データを多くとれる場合には、管理図法を利用するのがよいか、数値計算により検定も行える。次式のいずれかによる。

イ) データの数 (n) の多いとき (40 以上, 20 以上でも大体よい)

$$\begin{aligned} u_0 &= \frac{\bar{x} - \mu}{s/\sqrt{n}} \approx \frac{\bar{x} - \mu}{\bar{R}/d_2 \cdot \sqrt{n}} = \frac{\bar{x} - \mu}{A_2 \bar{R}/3\sqrt{k}} \\ &= \frac{\bar{x} - \bar{x}'}{\sqrt{ns^2 + n's'^2}} \sqrt{nn'} \end{aligned} \quad (3.16)$$

ただし n' , s' はそれぞれ標準法によるデータの数及び標準偏差。この u_0 の出る確率を正規分布表よりもとめて検定する。

ロ) データの数の少ないとき

$$t_0 = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = \frac{\bar{x} - \mu}{\sigma_s/\sqrt{n}} \quad (\text{自由度 } f = n-1) \quad (3.17)$$

$$= \frac{\bar{x} - \bar{x}'}{\sqrt{ns^2 + n's'^2}} \sqrt{\frac{(n+n'-2)nn'}{n+n'}} \quad (3.18)$$

(自由度 $f = n+n'-2$)

この t_0 の値を、自由度を考慮して t 分布表と比較して検定する。

以上は、同じ試料について、標準法及びその測定法で行った場合であるが、各測定毎に、性質のことなつたサンプルで(ただし標準法及びその測定法では同じサンプルにつき各1回ずつ測定)、測定したデータがあるときは、次式による。

$$t = \frac{\bar{d}}{s_d/\sqrt{n-1}} \quad (\text{自由度 } f = n-1) \quad (3.19)$$

ただし n : サンプルの数 (大きさ) \bar{d} : 同じ試料につき (標準法の測定値) - (その測定法の測定値) = d としたとき, その d の平均値 s_d : d の標準偏差

これは分散分析法によっても検定できる.

例 新入工が入ったので, 熟練工と同時にガス分析を行わせ, その正確さを検定したい. 10日間にわたり, 毎日現場よりサンプルをとって来てこれを2人に分析させたデータが次の通りである. この場合は毎日サンプルがことなるので (3.19) 式を用いる. (3.18) 式を用いるのはまずい.

新入工	熟練工	$d \times 10$	$(d \times 10)^2$	
19.2	19.4	-2	4	$(\bar{d} \times 10) = 4.4$
18.5	18.0	5	25	$\sum (d - \bar{d})^2 \times 10^2 = \sum (10d)^2$
19.1	18.6	5	25	$-\frac{(\sum 10d)^2}{N}$
19.0	18.2	8	64	$= 340 - \frac{44^2}{10} = 146.4$
20.3	19.7	6	36	
21.4	20.6	8	64	$\therefore 10s_d = \sqrt{\frac{146.4}{10}} = 3.83$
18.2	18.5	-3	9	
19.5	19.1	4	16	$t_0 = \frac{4.4}{3.83/\sqrt{9}} \approx 3.45^{**1)}$
20.6	19.7	9	81	
21.4	21.0	4	16	
		44	340	分布表より自由度 $f = n - 1$
				$= 9$ の 1% 有意水準の値を見

ると $t_{0.01} = 3.25$ ($f = 9$)

したがって t_0 は非常に有意, すなわち新入工の分析値の正確さは熟練工に比したしかにちがう. データを見ると新入工は大きな分析値を出す. そのカタヨリは大体0.44%である. 原因をしらべたら, ゴム管のつなぎ方のわるいことがわかったので, この練習をさせたら, 後は有意差はなくなった.

c) 分散分析法を用いる方法 特にいろいろな要因の正確さに及ぼす影響をしらべるには, 分散分析法を用いるとよい. これについては § 3.4.4

1) 統計的検定では, 5%, 1%の危険率で有意な差があると認められたときには, それぞれ「*」「**」印であらわす.

を参照されたい。

検定について1つ注意しておきたいことは、同じ精度でも、測定の正確さは、その測定回数により有意に出てくることである。したがって技術的に第2種の過誤を考慮して、測定回数を決めなければならないことがある。

3. カタヨリの推定と補正 カタヨリの有無を検定して、有意差ありとみとめられた場合には、そのカタヨリを推定して、補正する必要がある。この推定には一般に点推定を用いる。たとえば標準法のときの平均値 \bar{x} 、問題としている測定法による平均値を \bar{y} とすると、カタヨリの推定値 a_e は

$$d_e = \bar{y} - \bar{x}$$

その推定の精度は、サンプリングの平均値の精度と同様にしてもとめられる。

また、相関分析を利用して、カタヨリを推定し、相関図により補正を行うことも実際によく行われている。

例 空試験 (blank test) を行うのは、カタヨリを補正するためである。これは滴定分析などで従来からよく行われている方法である。図3.3は、成分既知の材料を添加して、滴定し、その成分量を計算より推定した値と、既知の値との相関をとったもので、原点を通る45°の直線に対して、この

回帰線が有意差なしとみとめられれば、カタヨリがありとはいえず、実線の場合は実測値から理論値を推定する線が45°の線と平行線となり、相関係数が有意差のある場合で、既知材料を添加しないときに、実測値は3.0 mg となることを示してい

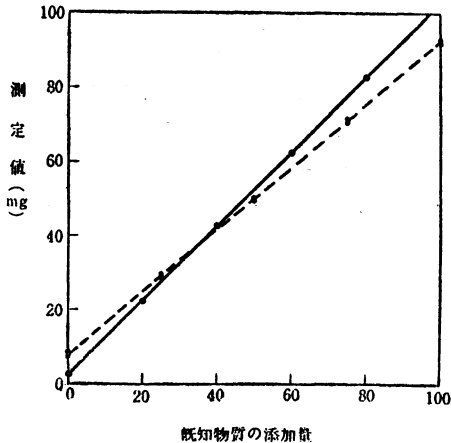


図 3.3

る。この値が空試験の値で、添加したときにも、同様に3.0 mg ずつ多く実測値が出ることを示しているからこれだけ補正しなければならぬ。破線は材料の添加量により、カタヨリの量がことなる場合で、線は45°の線とある角度で交わっている。この例では、約50 mg のところで交わっており、これより添加量少ないときには、実測値は大きい方へカタヨリ、多いときには、小さい方へカタヨっていることを示している。したがってこれだけ補正を加えてやらなければならぬ。点が線の両側にバラツキているのは、その精度を示している。

ピベット、ビュレット、メスフラスコ、分銅の補正なども、繰り返し実験を行って、有意差の検定を行ってから、補正を行うべきである。

測定法にカタヨリがある場合には、そのカタヨリの原因を除去することも必要であるが、これが不経済、または不可能で、かつそのカタヨリが管理状態で出てくるような場合には、あらかじめそのカタヨリを推定して、補正を加えてやれば十分な場合が多い。たとえば工程管理のために、少しバラツキは大きくても、簡易測定法を管理試験として利用することも多いが、この場合にも相関分析などを応用して標準測定法に対するカタヨリを推定して、補正を行い、その精度や信頼性をチェックしてから用いればよい。

3.4.4 測定誤差の解析

測定値のバラツキは、多くの原因によるカタヨリにより生ずるもので、われわれはその原因を追究して、カタヨリを除去し、誤差を小さくしなければならぬ。この誤差はいくつかの要因毎に分解できる。これらの要因は、各々独立な場合もあるし、相関関係をもつ場合、あるいは交互作用のある場合もある。これらの誤差の主な原因は

- 1) 測定者
- 2) 測定法及び測定器具
- 3) サンプル自身のバラツキ

である。これらの原因の中、どれが有意な原因であるかを解析するには、管理図法——特にいろいろに層別して——、分散分析法、相関分析法などが用いられる。

例 1) 分析誤差の大きい原因と技術的に考えられる、分析者、測定

器、標準液のいずれを問題とする必要のある原因かを調査したために、次のような実験を3元配置法で実験計画的に行った。6人の分析者に、No. 1 および No. 2 の測定器を使用させて、A標準液とB標準液を用いて、均一とみなせる同一試料を分析させた。ただし測定の種類は乱数表によりランダムに行った。

データは表3・2の通りであり、これを分散分析した結果が表3・3である。分散分析表を見ると、分析者間以下の3つの要因は、残差項に対して有意差が全く認められないので、これをプールして誤差項とみなす(この場合は、はじめに残差項のみをとり検定した場合より検出力はよくなる)。測定器間、標準液間には非常に大きな変動——カタヨリがある。標準液と測定器の間には、交互作用があるという結果にはなっ

表 3・2 分析データ

分析者	No.1 測定器		No.2 測定器	
	A標準液	B標準液	A標準液	B標準液
1	27.5	26.8	26.8	26.4
2	27.5	26.8	26.9	26.5
3	27.4	26.7	26.7	26.5
4	27.5	26.8	26.9	26.4
5	27.6	26.8	27.0	26.3
6	27.5	27.0	26.8	26.3

表 3・3 分散分析表

要因	自由度 f	V_e	F_0
測定器間	1	170.6	213**
標準液間	1	192.6	240**
標準液×測定器	1	8.3	10.4**
分析者間	5	0.6	0.8
分析者×標準液	5	1.1	
分析者×測定器	5	0.9	
残差	5	0.7	

いるが、これは技術的に考えられない。このような場合測定器と標準液に高度に有意差あるために、データだけから交互作用が有意と出て来ても、交互作用ありと判断することは危険である。統計的な結果は技術的に解釈しなければならない。

〔結論〕 1) 測定器、標準液間には高度に有意な差がある。

2) 分析者間には差はみとめられない。

〔処置〕 1) 測定器はよく調査したら、目盛が狂っていたので、補正して用いることとした。

2) 標準液は古くなると変化することがわかったので、1週間以上たったものは使用しないことにきめた。

例 2) A_1, A_2 2種類の材料が入ったので、 B_1, B_2 2人に試片各10枚を作成させて、 C_1, C_2 2人の検査工に同じ装置でその強度を試験させた。この場合技術的に交互作用があるとは考えられなかったので、 2×2 ラテン方格で繰り返しのある実験計画法を用い通常通り各条件で5個ずつ試験する計画をたてた。ただし試片の作成及び強度試験の順序はランダムに行わせた。そのデータ並びに解析結果は次の通りである。

表 3・4 試験データ

	A_1	A_2
B_1	C_1 9.6	C_2 11.4
	10.5	11.4
	10.2	10.6
	9.6	11.2
	9.7	11.5
B_2	C_2 10.7	C_1 11.1
	11.3	11.0
	11.4	10.3
	10.7	11.4
	10.9	11.4

表 3・5 補助表

(x)	A_1	A_2	和
B_1	C_1 -12	C_2 6	$\sum x_{B_1}$
	-3	6	
	-6	-2	-23
	-12	4	
	-11	7	
B_2	C_2 -1	C_1 3	$\sum x_{B_2}$
	5	2	
	6	-5	22
	-1	6	
	1	6	
和	$\sum x_{A_1}$ -34	$\sum x_{A_2}$ 33	-1

[手順] 1) 計算を簡単にするためにもとのデータから10.8を引き10倍する¹⁾。

2) 各行、欄の和及び総和をもとめる。

3) 修正項 CF をもとめる。

$$CF = \frac{(\sum x)^2}{n} = \frac{1^2}{20} = 0.05$$

4) 総変動 S_0 をもとめる。

$$S_0 = \sum x^2 - CF = 768.95$$

5) 材料間変動 S_A をもとめる。

$$\begin{aligned} S_A &= \frac{\sum x_{A1}^2 + (\sum x_{A2})^2}{n_A} - CF \\ &= \frac{(-31)^2 + (33)^2}{10} - 0.05 = 224.45 \end{aligned}$$

6) 試片作成者間変動 S_B をもとめる。

$$S_B = \frac{(\sum x_{B1})^2 + (\sum x_{B2})^2}{n_B} - CF = \frac{(-23)^2 + (22)^2}{10} - 0.05 = 101.25$$

7) 検査工間変動 S_C をもとめる。

$$S_C = \frac{(\sum x_{c1})^2 + (\sum x_{c2})^2}{n_c} - CF = \frac{(-32)^2 + (31)^2}{10} - 0.05 = 198.45$$

8) 誤差変動 S_E をもとめる。

$$S_E = S_0 - (S_A + S_B + S_C) = 244.80$$

9) 分散分析表を作成する。

要因	変動 (偏差2乗和)	f	V_e	F_0	$F_{1,16}$ $\alpha=0.05$
A	224.45	1	224.45	14.7**	4.49
B	101.25	1	101.25	6.6*	
C	198.45	1	198.45	13.0**	$\alpha=0.01$
E	244.80	16	15.3		8.53

[結論] 1) 材料間にはたしかに有意な差がある。

2) 検査工間にもたしかに有意な差がある。

1) 統計的な計算を行うときには、必ずこのように計算の簡便化をはかること。

3) 試片作成工間には有意な差があるようだ。

〔処置〕 1) A_1 の材料の方が強度が低いことがわかったので、 A_2 を採用することにした。

2) 検査工 B_2 は、試験器のネジの締めかたが不十分なことがわかったので、十分訓練した。後にもう一度チェックしたら有意差は認められなかった。

3) 試片作成者間の差の原因は未だわかっていないので、さらに検討して見る予定である。

備考 この場合誤差項の不偏分散 15.3 は、もとの単位になおして、0.153 とすると、これは強度測定の繰返精度であり、試片作成、試験工をも含めての再現精度 $\hat{\sigma}_{rpd}^2$ は

$$\hat{\sigma}_{rpd}^2 = \frac{101.25 + 198.45 + 244.80}{1 + 1 + 16} = 30.2$$

もとの単位になおして、0.302 である。

3.4.5 サンプルングの誤差と測定誤差の分離

まえにのべたように、われわれがうるデータのバラツキ σ は、サンプルングによるバラツキ σ_s と、測定によるバラツキ σ_M とが一緒になったものである。サンプルングを論ずるときには、これを分離しなければならぬ。いまあるデータのバラツキを考えると

$$\sigma^2 = \sigma_s^2 + \sigma_M^2 \quad \text{または} \quad V = V_s + V_M$$

σ^2 は実際にえたデータより推定できるので、 σ_M^2 がわかれば、 σ_s^2 はもとめられる。

1. σ_M が簡単にもとめられるとき 非破壊試験で同じサンプルを何回も測定できるとき、化学分析などの破壊試験でも流体で十分均一なサンプルをえられるとき、微粉体として均一と見なせるサンプルをえられるときは σ_M が容易にもとめられる。多くの場合管理図を用い \bar{R}_M より $\hat{\sigma}_M$ をもとめるか、あるいは分散分析法によりもとめられる。ただしこの場合次のことに注意しなければならない。

注意 1) 同じもの、均一なものというのが案外同じでない場合がある。たとえば丸棒の直径をはかるときを考えると、丸棒が場所により径にバラツキがあると、同じ太さを測定しているつもりでも、実はことなっ

た径を測定していることになり、この径のパラツキ（一種のサンプリング誤差である）は測定のパラツキの中に含まれてしまう。しかしこれを測定精度と定義することもあろう。流体や微粉体を分析する場合にも、よほどよく混合しないと均一と見なせない場合がある。たとえば石炭を60メッシュ以下に粉碎して、灰分を測定するときに、(A) 粉炭を全く混合せずにそのまま、(B) スプーンで攪拌してから、(C) 瓶に入れて十分に振盪してからサンプリングしたときのデータのパラツキの1例を示すと次の通りである。

この場合には、最後のサンプリング誤差を含めたものを、測定誤差と定義しているので、余程混合を十分に行わないと、測定精度をもとめても、意味がなくなってしまう。したがって測定精度を論ずる場合には混合法や測定法などをわかりやすく、具体的に示しておく必要がある。

2) もし σ_M に繰返精度をとると、 σ_s には、サンプリングの精度以外に、再現精度によるパラツキが入ってしまうことがあるから、このような試験では、時間的にあるいは人や装置をランダムにして再現精度をとるべきである。

3) もしサンプリングを何回か行い、それにカタヨリがあると、 V_s には、サンプリングの精度以外にカタヨリによるパラツキも入るから注意。

2. 測定誤差をもとめられぬとき 破壊試験などで、均一なサンプルの得られぬときである。

このときには一般に、前にのべたようにサンプリング誤差と測定誤差を分離することが困難であるので、できるだけ均一と見なしうる材料をうるように技術的に努力するより方法がない。たとえば短時間に管理状態でつくられた同じロットと見なせる材料から、ランダムにサンプルをとり、測定し、この測定値のパラツキを測定法の誤差 σ_M と定義する。

化学分析のように縮分を行って、分析をするときにも、最後の分析試料からのサンプリング誤差は測定誤差に含まれてしまう。たとえば表3.6のC法の場合でも、サンプリング誤差が全くないとはいえない。

表3.6 混合条件によるパラツキの変化

混合法	標準偏差%
A	1.10
B	0.51
C	0.21

もしこのような場合に、破壊試験などの特性値と、非破壊試験による特性値との間に、ほとんどバラツキのない、高度に有意な相関関係でもあれば、この特性値を用いて、サンプル間のバラツキを測定できれば、 σ_s をもとめることができる。しかしこの場合非破壊試験の測定法のバラツキが大きいときには、測定の繰返数を相当増加させなければならない。

参考 管理図により工程を管理するためのサンプルングと測定を考えると、完全な管理状態になった場合には、管理限界の巾は、 \bar{R} すなわち $\sigma^2 = \sigma_s^2 + \sigma_M^2$ によりきまってくる。したがって一応 σ_s 、 σ_M がどのくらいの大きさをチェックしておいた方がよい。

たとえばはじめに管理図をかくに当って、技術的に σ_s 、 σ_M が相当大きいことが考えられるときには、 σ 、 σ_s 、 σ_M 、などをあらかじめもめておいて、これらより管理限界の巾をきめて、管理図をかいて見るとよい。もしこの管理図が管理状態を示していれば、この工程を管理するには現在行っているサンプルング法や測定法ではバラツキが大きすぎることを示しているのだから、これを改良しなければならない。もし多くの \bar{X} の点が限界外に飛び出すようであれば、現在の方法で十分この工程が管理できることを示している。

3・5 サンプルングによる誤差

サンプルングによる誤差についても、信頼性、精度、正確さなどの考えかたや注意は、測定の誤差のときと同様である。またこれについては後章においていろいろのべるので、特に注意すべき点だけをのべよう。ただしここでは話を簡単にするために、測定の誤差は考えずに、サンプルングの誤差だけを考える。

工場におけるサンプルングにおいて誤差を生ずる主な原因は次の通りである。

- 1) 目的がはっきりしていなかったり、間違えている。
- 2) サンプルング法の標準が正しくない。
- 3) サンプルング係がサンプルング法標準やその目的をよく理解していない。または知っていても標準通りに正しくサンプルングを行わない。
- 4) サンプルングするものがランダムにとらず、よい、あるいはわるいも

のを意識的あるいは無意識的にとってしまう。このためには、あらかじめ TWI 方式などを活用して、十分その目的を認識させ、訓練を行い、その後も絶えざる管理や注意が必要である。

5) サンプルング法に対して、平均値、総計や分散などを推定する式が正しくない。

6) 母集団の変化。a) 母集団がサンプルをとった後に変化したり、b) 次々と来るロットが管理状態になく、相当性質のことになった母集団となっているのに気付かずに、同じサンプルング法をつづけている誤。

7) サンプルの変化。折角ランダムに合理的にとったサンプルも、その後の取り扱いかたが悪いとカタヨリを生ずる。

8) インクリメントの大きさが不適當であるためにカタヨリが入る。

以上各項目のうちこれまでのにべなかつた注意をのべよう。

目的については § 1 にいろいろのべたが、この目的はサンプルング法の設計者、サンプルングの実施者、測定、分析者にも徹底させておかなければならない。折角サンプルングがうまく行われていても、サンプル調製者や分析者がこれを知らぬために、縮分などの方法を間違えたり、完全にやらなかつたりしている例をよく見る。

サンプルング法標準の正しくない規格、仕様書は実に多い。

サンプルング実施者によるカタヨリは非常に多いので、是非ランダムサンプルング法の訓練やその後のチェックを行わなければならないが、心理的影響があるから注意しなければならない。たとえばあるロットについて、数名のサンプラーについて、実験計画的にサンプルをとらせ、これを分散分析している例を見るが、今日はチェックだということを知っていれば、いつも怠けてよい加減にとっているものも一生懸命にやるであろう。したがっていくら分散分析で検定して、有意差がなくとも、それはその実験についてだけの結論で、平生も正しくとっている証明にはならない。むしろ毎日合理的にチェックサンプルをとらせて、これを別々に各 1 回分析などして管理して行く方がよい。

推定式については、いろいろな場合について § 4 でのべる。

母集団の変化は、母集団がサンプルをとった以後に変化すると、サンプルから推定した母集団は、現在の材料、製品として存在する母集団とはこ

となったものとなり、前にとったサンプルからの推定では、カタヨリを生ずる。水分とか揮発性物質をもつ材料、あるいは化学反応などが進行中の母集団からサンプルをとると、たとえば屋外貯蔵の石炭、鉱石類の水分は雨が降ったり晴天つづきであると水分は変化している。また溶剤を含んだ綿菓やケーキのような場合も同様である。熟成、醗酵などが進行中の場合も同じ。

母集団について、もう一つ注意すべきことは、工場でわれわれがサンプルリングの対象としている母集団は、原料から製品にいたるまで、すべて工程により生産されたものである。たとえば原料として食塩が入荷した場合、これはやはり塩の生産者が生産したものである。したがって、その工程の管理状態により、あるいは別の供給者からの原料であれば、母集団の性質が変化していることがある。予備調査を行って、サンプルング法標準の決定を行ったのちに、工程の管理状態の変化などにより、入ってくる母集団の性質がかわると、このサンプルング法標準が不合理になることがある。したがってサンプルング法標準を決定するには、そのロットの管理状態をチェックして、その変化に対して合理的にきめておかねばならない。

サンプルの変化は、サンプルング後のサンプルの保存や準備などの取扱上の不注意によりおこる。たとえばサンプルを粉塊混合物の粉の部分だけこぼしたり、レットル等をはっきりさせておかなかつたために間違えたり、揮発性または吸湿性のサンプルを蓋をせずに放置したり、暑い所や湿度の高いところへおいたりするとサンプルは変化する。これらのことは主としてサンプルの保存の問題で、サンプル容器、レットル、保存場所など、当然技術的に払われるべき注意を嚴重にまもるように管理すればよい。分析試料を作成するために、サンプルの準備として、縮分とか、洗滌などを必要とするときに、その方法を十分統計的、技術的に検討して行わなければ、カタヨリが入ることがある。

インクリメントの大きさが適当でないと、サンプルをいくらランダムにとってもカタヨリが入る。従来行われているサンプルング法では、インクリメントの大きさが一般には大きすぎる例が多いが、また小さすぎる場合もある。たとえば、綿花、羊毛などのような繊維状のものをサンプルングするときには、インクリメントでサンプルを集めるが、この1インクリメ

ントには多数の長短の繊維が含まれてくる。この場合一般にインクリメントの大きさが小さいと、長い繊維を多くとりすぎることがある。また石炭のように粉塊状の混合物では、インクリメントの大きさが、その最大粒度に比して小さすぎると、小さい粒子のものを多くとりすぎたり、大きい粒子のものを多くとりすぎたりすることがある。このような場合には、一般にインクリメントの大きさを大きくすれば、これらのカタヨリをさけることができる。またもしカタヨリをさけるために、インクリメントの大きさが大きくなりすぎるようなときには、ロットを粉碎して、最大粒度を小さくしておいてからサンプリングするとか、繊維の方向を揃えておいてからサンプリングすれば、インクリメントの大きさをあまり大きくしないですむ。

以上のような、技術的なことを考慮にいれて、カタヨリや管理状態を見ながら、精度を統計的に検討して行く方法を § 4 以下にのべる。

第4章 サンプリングの種類と簡単な理論

4.1 まえがき

§2 にのべたように、サンプルはランダムにとらなければならない。しかし材料または製品などのロットから、完全にランダムにサンプルをとるといふ単純ランダムサンプリングは 1) 厄介または困難である、2) 精度もあまりよくない、3) 所要精度をうるためには、サンプルの大きさが大きくなり、不経済である、などの欠点をもっている。これを避けて、実用的な、経済的なサンプリングを行うためのいろいろの工夫が行われる。これには、そのロットに関する技術的、統計的な予備知識が必要で、ロットに関する予備知識が全くない場合には、これらの工夫を行うことが困難で、単純ランダムサンプリングを行うが、少しでも知識があれば、以下にのべるような工夫を行って、有利なサンプリング法を計画すべきである。

以下 §2 にのべた以外の主な型のサンプリング法について、そのやり方、簡単な推定式、その精度などについてのべる。サンプリング法を設計するときに注意しなければならぬことは、その目的についてであり、たとえば多くの場合ロットの平均値や総計を知りたい場合が多いから、それがいくつかの層に分けられるときに、ロット全体についての平均値を精度よく知りたいのか、各層毎に対する知識が重要であるのか、あるいは両者とも知りたいのかにより、サンプリング設計法が全くかわってくるので、その点をはじめによく考えておかなければならない。

本章ではやや玉石混濁的にいろいろなサンプリング法について、簡単にのべるが、これらは必ずしも各々独立に1つずつ用いるばかりでなく、これらを組合せて用いる場合も多い。サンプリングの目的と、そのロットに対する予備知識を利用して、いろいろなサンプリング法を縦横に組合せて、希望する精度や時間で母数をもっとも経済的に推定できるようなサンプリング法を決定しなければならない。したがってロットをどのように層別して層をいくつとるか、その各層からサンプルをどのくらいとるかとい

う、サンプルの最適の割当てという問題が入ってくる。これをサンプルの最適配分 (optimum allocation) という。これには、いろいろの経費や技術上の問題を考えに入れて、所要精度に対して、最低経費でサンプリングするための配分法、一定経費で最良精度になるようにサンプルを割当てる方法などが考えられる。

これらのいろいろな工夫は、所謂社会調査のための標本調査法において統計的に発達している。工場におけるサンプリングにおいては、社会調査の場合といろいろことになった点もあるので、まずその共通点、ことになった点について比較検討して見よう。

なお本章においては、分散の式は§4・9を除いては全部母分散、たとえば σ^2 , σ^2_b , σ^2_w , σ^2_o , σ^2_i などとしてのべている。実際にはこの推定値を入れることになる。その推定法については§5以下においてのべる。また計数の場合は本章の特別な場合と考えられるので§6・2にのべる。

4・2 工場におけるサンプリングと社会標本調査法の比較

この両者は非常に類似している点もあるが、こまかく考えて見ると、ことになっている点も多い。次にサンプリング法の設計から、サンプリングの実施、測定調査、その結果をいかにまとめて、利用するかという点まで、両者を比較して見よう。これらのことは、サンプリングを設計するときの基本的態度にも、大きな関係を持ち、また標本調査法の統計技術を工業に導入する場合にも考えなければならぬことである。

1. 対象物

本書がサンプリングの対象としている対象物は、主として品物 (physical materials) であって、意志をもっていないが、社会調査などの場合には、対象が多くは、人あるいはその集団で、意志、感情、思想をもっている。

2. 目的

PMS (physical materials sampling) は、その結果を見て、対象物、ロットに対して、あるいは工程に対して処置 (action) をとる。

SS (sampling survey) (market research を含む) では、その結果を見て、対象に対して、広告、啓蒙などの処置をとることもあるが、むしろ

ろ政策とか、経営方針とかに処置をとる。

3. 母集団の構造

SSの場合には母集団が非常に広範複雑で、層や集落の大きさも非常にことなる場合が多いが、PMSにおいては構造が比較的簡単な場合が多い。したがってサンプルング法も、PMSの方が簡単な式で十分な場合が多い。

4. 歴史

PMSはその性質上、昔からなんらかの形でサンプルングが行われていて、古い伝統があり、しかも不合理により加減に行われている。特に規格や仕様書において不合理なものが多い。しかし最近はサンプルングの合理化の必要性が次第に認識されつつある。SSは従来あまり行われておらず、有意的にえらんだサンプルによる調査、悉皆調査、あるいは勘による判断という問題もあり、またサンプルング以外に多くの問題がある。したがって、サンプルングによる調査の意味や必要性に問題や議論がある。

5. サンプルング法の設計、予備調査

両者とも統計的知識と、その調査事項や母集団についての技術的知識をもった技術者による設計が必要である。しかしSSでは、サンプルング実施回数も少なく、予備調査もある程度は必要であるが、PMSでは、毎日サンプルングを行う場合が多く、サンプルング法標準を作成しなければならず、その作成にあたっては、技術的、統計的に予備知識や調査などにより、十分検討しておかなければならない。したがってPMSにおけるサンプルングの研究は主として予備調査、予備実験を行うということになる。

SSでは、サンプルング法標準などの必要がなく、その目的に応じて設計すればよい場合が多い。

しかしSSではサンプルングの実施にあたって、予備調査の一つであるリストの作成という手間のかかる問題があるが、PMSでは、多くの場合、サンプル単位についてのリストなどを作成せずに現場でサンプルングする。

6. サンプルングの実施

SSにおいては、サンプルング法設計者がリストなどによりどのサンプル単位をとるかということまで、決定できる場合が多いが、PMSでは、

サンプリング法標準に従って、現場員がどのサンプルをとるかということ
を、対象物を目前に見ながら決定する場合が多いので、カタヨリも入りや
すく、またその技術的、良心的訓練が必要である。

7. サンプルの取扱

PMSではサンプル採取後、縮分したり、その他準備や保存上の問題が
あるが、SSでは、このような問題はないが、サンプリングすべく決定し
たサンプルの無い場合——無回答の1種——の問題がある。

8. 無回答

SSでは、サンプルのない場合、回答拒否など、かならず無回答の問題
があり、その取り扱い方が必ず問題になる。PMSでもサンプルの紛失や
測定、分析の失敗などにより、データの得られぬこともあるが、注意さえ
すれば、避けらるので、無回答の問題は少い。

9. サンプルの取り直し

SSではサンプルの取り直しは比較的問題はないが、PMSではサンプ
ルの取り直しが困難あるいは不可能な場合が多い。

10. 測定——調査

PMSでは多くの場合、測定法、分析法などの標準が出来ていて測定は
物理的、化学的に客観的に行える。もちろん測定者や測定器などによるカ
タヨリもあるが、これは技術的に注意などすれば少なくできる。SSで
は、調査員並に回答者の心理的な影響、環境や社会的な影響により、カ
タヨリやウソが入りやすい。しかし、分析者、測定者、調査員などを管理し
なければならぬことは同様である。

11. 測定法——調査法

PMSでは、測定法標準を作成することが必要であるが、SSでは、調
査毎に、質問表の作成という重要な問題がある。

12. サンプル単位

SSでは、その調査目的により、サンプル単位を、個人、世帯、部落な
どと比較的簡単にきめられる場合が多いが、PMSでは、たとえばインク
リメントの大きさ、試長の決定のように、サンプル単位のきめ方にいろ
んな問題がある。

SSでは、サンプル単位の大きさやその特性値に大きな差のある場合が

あるが、PMSではサンプル単位の大きさは、ほぼ一定で、特性値のパラツキにもあまり大きな変化がない場合が多い。

13. 母集団の分散

SSでは、母集団全体の分散は、母集団さえ決定すれば大体決ってくる。また分散に変化があっても、時間的に自然に少し変化するくらいである。PMSでは、その母集団がでてくる工程の管理状態や、そのロットの取扱い方法により、分散は大きく変化し、また分散を変化させることもできる。

14. 級間分散と級内分散

SSでは、調査区の決定方法により、級内分散 σ_w^2 と級間分散 σ_b^2 をいろいろかえられる。PMSでも場合によっては層別の仕方によりいろいろかえることはできるが、貨車積、込入、箱詰などの場合には、貨車、込入や箱内のバラツキと、箱間のバラツキは、工程などの管理状態により、 σ_w^2 と σ_b^2 のきまってしまう場合が多い。

15. 調査結果の統計的検討

SSでは、調査結果を統計的に検討して、結論をもとめると同時に次回調査のための貴重な資料とすることができるが、PMSでは、毎日行うサンプリグ結果から、統計的検討の資料をうるように行われていない場合が多い。たとえば、従来のサンプリグ法では、サンプル単位をいくつか集めて、1つの混合試料としてしまう場合には、サンプル単位間のバラツキの変化が大きい場合でも、これに対する知識がえられない場合が多い。しかし混合試料を作成するときには、はじめから、母集団についてのバラツキに関する知識を得られるようにサンプリグ法、混合試料の作成法を設計することができるからそのように努力すべきである。

16. 経費

SSでは、1回のサンプリグにおいても、サンプルの大きさ n が大きいので1回の調査に要する経費が相当大きい。PMSでは1回のサンプリグに要する経費は比較的安い。しかしPMSのサンプリグはこれを毎日何回も行うことが多いので、各1回の経費特に手間が大きな問題になる。

17. 推定式と集計経費

SSでは n が大きいので、調査後の集計整理に相当の時間と経費を要するが、PMSではこれが比較的簡単である。特に混合試料としてしまうとときには測定を数回やればよいので、集計よりも混合試料のつくり方が問題になる。

推定式もこれらのことを考えて決めなければならぬ。たとえばSSにおいても $\hat{\mu} = \frac{1}{n} \sum \sum x_{ij}$ という形式の式となるように設計すれば、集計も各サンプル単位の測定値を唯加え合せればよいので非常に容易である。PMSにおいても同様なことがいえるが、特に混合試料として測定する際には上のような推定式を用いるように設計しなければならない。たとえば $\hat{\mu} = \frac{1}{k} \sum_i \frac{1}{n_i} \sum_j^{n_i} x_{ij}$ という式になると、重みつき平均を考えなければならないので、簡単に混合試料とすることができず、各層毎に別々に混合試料をつくり、別々に測定しなければならない。すなわちいずれの場合も、集計の簡単にできるようなサンプリング法を設計すべきである。

4.3 層別サンプリング (stratified sampling)

単純ランダムサンプリングでは、母集団がいろいろ異質なものを含んでいて分散が大きいと、平均値の推定の精度は悪くなり、精度をあげるためには、サンプルの大きさを非常に大きくしなければならなくなる。もしこの母集団を実際に、あるいは頭の中でいくつかの層に分けることができる場合には、これを層別して、各層内からサンプルを各々ランダムにとる。このようにいくつかの層にわけて、その各層からサンプルをとるサンプリング法を層別サンプリングという。このサンプリング法によれば、各層内からランダムにサンプルをとればよいので、サンプリングすることも容易になる。また層別のやり方を、できるだけ各層内のバラツキ（級内分散 σ_w^2 ）を小さく均一にし、層間のバラツキ（級間分散 σ_b^2 ）を大きく、不均一になるように層別すれば (4.6) 式以下に示すように、平均値の精度には σ_b が入ってこないのと同じサンプル数の場合には、単純ランダムサンプリングの場合より、その総平均値の推定の精度は遙かによくなる。これは工場のサンプリングにおいても、非常に役立つ、よく用いられる方法である。

層別サンプリングしたときに、母集団全体についての推定の式や精度は次のようになる。いま N コよりなる母集団 (μ, σ^2) から、 n コのサンプルをとるのに、母集団を k コの層にわけて、その各層の大きさを N_i コとし、この各層より各 n_i コのサンプルをランダムにとると、

$$N = \sum_{i=1}^k N_i, \quad n = \sum_{i=1}^k n_i$$

母集団については、各層の母集団全体に対する割合を $p_i = N_i/N$ 、第 i 層の母平均 μ_i 、母分散 σ_i^2 とすると、

$$\mu = \sum_{i=1}^k p_i \mu_i \quad (4.1)$$

$$\mu_i = \frac{1}{N_i} \sum_{j=1}^{N_i} x_{ij} \quad (4.2)$$

$$\sigma^2 = \sum_{i=1}^k p_i \sigma_i^2 + \sum_{i=1}^k p_i (\mu_i - \mu)^2 = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^{N_i} (x_{ij} - \mu)^2 \quad (4.3)$$

$$\text{ただし} \quad \sigma_i^2 = \frac{1}{N_i} \sum_{j=1}^{N_i} (x_{ij} - \mu_i)^2$$

(4.3) 式の第1項を級内分散 σ_w^2 、第2項を級間分散 σ_b^2 という。

$$\sigma_w^2 = \sum_i p_i \sigma_i^2 \quad ; \quad \sigma_b^2 = \sum_i p_i (\mu_i - \mu)^2$$

したがって $\sigma^2 = \sigma_w^2 + \sigma_b^2$

変動係数 C についても同様に

$$C_w = \frac{\sigma_w}{\mu} \quad ; \quad C_b = \frac{\sigma_b}{\mu} \quad \text{とすると}$$

$$C^2 = C_w^2 + C_b^2$$

サンプルからの母平均の不偏推定は

$$\hat{\mu} \leftarrow \bar{x} = \sum_{i=1}^k p_i \bar{x}_i \quad (\text{ただし } \bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} ; \text{一般式}) \quad (4.4)$$

$$= \frac{1}{k} \sum_{i=1}^k \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \quad (N_i = \frac{N}{k} = \bar{N} \text{ 一定のとき}) \quad (4.5)$$

平均値の推定の分散(精度) $\sigma_{\bar{x}}^2$ は

$$\sigma_{\bar{x}}^2 = \sum_{i=1}^k \frac{N_i - n_i}{N_i - 1} \frac{\sigma_i^2}{n_i} p_i^2 \quad (\text{一般式}) \quad (4.6)$$

$$\Leftrightarrow \sum_i^k \left(\frac{1}{n_i} - \frac{1}{N_i} \right) p_i^2 \sigma_i^2 \quad (N_i \gg 1 \text{ のとき}) \quad (4.7)$$

$$\Leftrightarrow \sum_i^k \frac{p_i^2 \sigma_i^2}{n_i} = \frac{1}{N^2} \sum_i^k \frac{N_i^2 \sigma_i^2}{n_i} \quad (n_i/N_i < 0.1 \text{ のとき}) \quad (4.8)$$

$$\Leftrightarrow \frac{1}{k^2} \sum_i^k \frac{\sigma_i^2}{n_i} \quad (n_i/N_i < 0.1 \quad N_i = \frac{N}{k} = \text{一定}) \quad (4.9)$$

母集団の総計 μ_X を推定するには、 X は \bar{x}_i の各 N_i 倍、 σ^2_X は σ_i^2 の各 N_i^2 倍する。

$$\hat{p}_X \leftarrow X = \sum_{i=1}^k N_i \bar{x}_i = \sum_{i=1}^k \left(\frac{N_i}{n_i} \sum_{j=1}^{n_i} x_{ij} \right) \quad (4.10)$$

$$\sigma^2_X = \sum_{i=1}^k N_i^2 \frac{N_i - n_i}{N_i - 1} \frac{\sigma_i^2}{n_i} \quad (\text{一般式}) \quad (4.11)$$

$$\Leftrightarrow \sum_i^k N_i^2 \left(\frac{1}{n_i} - \frac{1}{N_i} \right) \sigma_i^2 \quad (N_i \gg 1) \quad (4.12)$$

$$\Leftrightarrow \sum_i^k N_i^2 \frac{\sigma_i^2}{n_i} \quad (n_i/N_i < 0.1) \quad (4.13)$$

これら分散の式を見るとわかるように、平均値や総計の精度は、級内の分散だけによりきまり、級間の分散が入ってこない。すなわち級間に行くら大きな差があっても、平均値の精度はかわらぬということは、層別サンプリングの非常に重要な特徴である。

上のべた式は、すべて母集団全体について推定を行うときの式である。各層毎についての推定の式は § 2.3 のランダムサンプリングの式を用いればよい。

上にはいろいろな場合のときの式をのべたが、工業においては、一般に $n_i/N_i < 0.1$ 、 N_i が大体一定という簡単な (4.9) 式の場合が多い。更に多いのは、次にのべる比例サンプリング、すなわち $n_i/N_i \equiv a$ 一定という場合である。

4.3.1 各層からのサンプルのとり方

層別サンプリングにおける層別のやり方は、技術的および統計的知識を利用して、意識的に層にわけが、各層からサンプルをとるには、やはり必ずランダムサンプリングをしなければならない。

ところが、層別サンプルングにおいては、各層からいかなる大きさの n_i のサンプルをとるべきかという問題がある。この問題は目的により大きく2つに分けられる。

そのサンプルングの目的が、(1) 母集団全体についての知識を得て、ロット全体に対して処置をとりたいのか、(2) あるいは各層についての知識を得て、各層に対して処置をとりたいのか、の2つに分けられる。(3) しかし各層の知識も得たいし、また全体についての知識も得たいという場合もある。これらの目的により、各層からとるサンプルの数を決める態度も相当異なってくる。

4.3.1.1 母集団全体についての知識のみを問題とするとき 母平均や、母集団の総計を推定するのになるべくサンプル数を少く、経費をやすく、かつ精度よく推定するために、いろいろの工夫が行われている。

1. 比例サンプルング (size proportionate sampling) 各層の大きさ N_i に比例して、一定の抜取比 a でサンプルングする方法で、従来われわれが多く行っていた方法である。工場においては、後にのべるように多くの場合比例サンプルングをすると精度がもっともよくなる。

$$\frac{n_1}{N_1} = \frac{n_2}{N_2} = \dots = \frac{n_i}{N_i} = \dots = \frac{n_k}{N_k} = \frac{n}{N} \equiv a$$

すなわちサンプルの全数を n とすると

$$n_i = \frac{N_i}{N} n = np_i = aN_i \quad (4.14)$$

したがって (4.4) (4.5) 式より

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij} \quad (4.15)$$

$$= \frac{1}{\bar{n}k} \sum_i \sum_j x_{ij} \quad (N_i = N/k = \bar{N} \text{ 一定}) \quad (4.16)$$

層別比例サンプルングでは、上式から判るように、平均値の推定は、 N_i が一定でなくとも、得たデータすべてを層に関係なく加えればよいので非常に簡単になる。同様にこのサンプルングならすべてのサンプルを一緒にして重量を測定して平均値をとったり、あるいは混合して、混合試料としてもよい。分散は

$$\sigma_a^2 = \sum_{i=1}^k \frac{N-n}{Np_i-1} \frac{\sigma_i^2}{n} p_i^2 \quad (4.17)$$

$$\begin{aligned} &\Leftrightarrow \left(\frac{1}{n} - \frac{1}{N} \right) \sum p_i \sigma_i^2 \\ &= \frac{(1-a)}{n} \sum p_i \sigma_i^2 \quad (Np_i = N_i \gg 1) \end{aligned} \quad (4.18)$$

$$\Leftrightarrow \frac{1}{n} \sum p_i \sigma_i^2 = \frac{\sigma_w^2}{n} \quad (n/N = a < 0.1) \quad (4.19)$$

$$\Leftrightarrow \frac{1}{nk} \sum \sigma_i^2 = (n/N < 0.1; N_i = \frac{N}{k} = \bar{N} \text{ 一定}) \quad (4.20)$$

ただし p_i 一定, すなわち N_i 一定ならば $\sigma_w^2 = \frac{\sum \sigma_i^2}{k}$

(4.20) 式のような条件が大体成立する場合が工場では多いが, このときには \bar{N} を正しく知る必要はない. ランダムサンプリングのときの平均値の精度 $\sigma_{\bar{x}(R)}^2$ は, (4.3) 式より

$$\begin{aligned} \sigma_{\bar{x}(R)}^2 &= \frac{\sigma_x^2}{n} = \frac{1}{n} \sum p_i \sigma_i^2 + \frac{1}{n} \sum p_i (\mu_i - u)^2 = \frac{\sigma_w^2 + \sigma_b^2}{n} \\ \sigma_{\bar{x}(R)}^2 - \sigma_{\bar{x}(P)}^2 &= \frac{\sigma_x^2}{n} - \frac{1}{n} \sum p_i \sigma_i^2 = \frac{1}{n} \sum p_i (\mu_i - \mu)^2 \\ &= \frac{\sigma_b^2}{n} \geq 0 \end{aligned}$$

すなわち実用的には, 層別のやり方がまずくて, $\mu_i = \mu$ すなわち $\sigma_b^2 = 0$ のときに, はじめてランダムサンプリングの精度に等しくなるが, 一般には, ランダムサンプリングの方が精度は悪い. 層別が非常にうまく行われ, 各層内において均一と見なせるときには, $\sigma_w^2 = 0$ すなわち $\sigma_{\bar{x}(P)}^2 = 0$ となり, サンプルの平均値がそのまま母平均となる. ある母集団については, σ^2 は一定であるから, σ_w^2 が小となるように, したがって層内をできるだけ均一になるように層別すれば, σ_b^2 が大となり, いいかえると層間の差をできるだけ大きくするように, サンプリング法を設計すれば, 全体の平均値の精度はよくなる. 従来のサンプリング法が無意識の中に層別サンプリングの形になっていることも多いから, 一応検討して見るとよ

例 1] 石炭が15トン積で10車入荷したので, 1貨車荷下しすること

に、ランダムにスコップ5杯ずつサンプルをとり、これを混合して、混合試料とすると、このサンプルから得られる灰分の値の精度はどのようになるか。

これは各貨車を層とし、10に層別し、各層の大きさは同じであるので、各層からサンプルをランダムに同数ずつとった——層別比例サンプリグ——場合である。この場合

$Np_i \gg 1$, $n/N < 0.1$, $N_i = \text{一定}$, $n_i = \text{一定}$
したがって (4.20) 式となる。

$$\sigma_{\bar{x}}^2 = \frac{1}{nk} \sum \sigma_i^2$$

もし貨車内のインクリメント間のバラツキ σ_i^2 がほぼ等しいとすれば、(このようなことが近似的に成立することは工場においてよくある)

$$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_{10}^2 = \sigma_{10}^2 = \sum p_i \sigma_i^2 \quad \text{とすると、}$$

$$\therefore \sum \sigma_i^2 = k\sigma_{10}^2$$

しかるに、(4.20) 式より

$$\sigma_{\bar{x}}^2 = \frac{\sigma_{10}^2}{\bar{n}k} = \frac{\sigma_{10}^2}{n}$$

この例では、 $\bar{n} = 5$, $k = 10$ $n = \bar{n}k = 50$

$$\therefore \sigma_{\bar{x}}^2 = \frac{\sigma_{10}^2}{50}$$

すなわち貨車間に相当の違いがあっても、混合試料の精度には影響なく、貨車内のインクリメント間のバラツキ σ_{10}^2 のみにより精度は決定される。

備考 これは次のようにしてももとめられる。第 i 層から重さ w_i のインクリメントをとり、その各層毎の混合試料の灰分—平均値—を \bar{x}_i とすると、全混合試料の灰分—総平均値— $\bar{\bar{x}}$ は、 W を全サンプル量とすると、比例サンプリグであれば、

$$\bar{\bar{x}} = (n_1 w_1 \bar{x}_1 + n_2 w_2 \bar{x}_2 + \dots + n_{10} w_{10} \bar{x}_{10}) / W$$

$$\begin{aligned} \therefore V(\bar{\bar{x}}) &= \left(\frac{n_1 w_1}{W} \right)^2 \sigma_{\bar{x}_1}^2 + \left(\frac{n_2 w_2}{W} \right)^2 \sigma_{\bar{x}_2}^2 + \dots + \left(\frac{n_{10} w_{10}}{W} \right)^2 \sigma_{\bar{x}_{10}}^2 \\ &= \sum \left(\frac{n_i w_i}{W} \right)^2 \sigma_{\bar{x}_i}^2 \end{aligned}$$

しかるに $n_i = \bar{n}$, $w_i = \bar{w}$ で一定とすると

$$V(\bar{x}) = \left(\frac{\bar{n}\bar{w}}{W} \right)^2 \sum_i^k \sigma_{\bar{x}_i}^2$$

しかるに $\sigma_{\bar{x}_i}^2 = \frac{\sigma_{x_i}^2}{n_i}$, いま $\sigma_{x_i}^2 = \sigma_w^2$ で一定とすると。

$$V(\bar{x}) = \left(\frac{\bar{n}\bar{w}}{W} \right)^2 \frac{k\sigma_w^2}{\bar{n}} = \frac{\sigma_w^2}{\bar{n}k} = \frac{\sigma_w^2}{n} \left(\because \frac{\bar{n}\bar{w}}{W} = \frac{1}{k} \right)$$

例 2] 一定時間に大体一定数ずつ生産される製品がある。この製品の1日の平均重量を知りたいので、サンプリング法をきめるためにこの工程から一定間隔で1日に100コのサンプルをとり——系統的サンプリング——その重量をおのおの測定した。全体のデータから分散をもとめランダムサンプリングの式を用いて、サンプル20の平均値の精度をもとめた。またこのデータを時間順に5コ毎に1つの層からのサンプルと考え、20の層から各1コずつランダムサンプルをとったと考えると¹⁾、たとえば時間順に5コ毎に1コずつ計20コのサンプルを5組つくりその各平均値のバラツキを実際にもとめて見た。この実験並びに計算の結果は表4.1の通りである。

全体のデータより各サンプル間の標準偏差をもとめると1.83gである。

したがって全体から20コのサンプルをランダムにとったものとし、その平均値をもとめれば、その平均値の標準偏差は、

$$s_{\bar{x}_{20}} = 1.83/\sqrt{20} \approx 0.409 \text{ g}$$

5コおきにとった各20コのサンプルの5組の各平均値の標準偏差 $s_{\bar{x}_i}$ は0.146gである。

$$s_{\bar{x}_{20}}/s_{\bar{x}_i} = 0.409/0.146 \approx 2.8$$

すなわち一定間隔で系統的に20コのサンプルをとると、その平均値のバラツキは、完全なランダムサンプリングのときよりもはるかに小さくなっている。

これは § 2.2.2 でのべたように系統的サンプリングが、一種の層別サンプリングの効果を発揮している場合である。たとえば、連続5コずつ

1) 各層からの5コは系統的サンプリングになっているが、これをランダムサンプルであるとす
る。

表 4・1 ある製品の重量 (単位 g) [データは測定順]

群番号	各サンプルのデータ					\bar{X}	R
1	104.9	104.2	103.8	101.6	104.3	104.36	1.1
2	102.3	103.0	102.6	102.6	102.9	102.68	0.7
3	104.0	105.3	103.9	103.9	104.0	104.22	1.4
4	100.2	103.1	101.7	101.7	103.0	101.94	2.9
5	102.6	102.9	101.9	105.5	103.0	103.18	3.6
6	101.6	100.9	103.3	101.2	99.7	101.34	3.6
7	102.1	101.9	102.4	101.7	103.7	102.36	2.0
8	100.0	100.8	98.9	99.9	103.4	100.60	4.5
9	104.8	103.2	102.9	103.6	102.2	103.34	2.6
10	103.0	102.6	100.4	103.0	103.4	102.48	3.0
11	102.8	101.6	101.6	101.9	102.5	102.08	1.2
12	101.5	103.7	103.9	101.3	103.6	102.80	2.6
13	102.4	103.3	104.0	103.9	103.8	103.48	1.6
14	106.7	105.6	104.5	105.5	105.3	105.52	2.2
15	106.9	106.2	106.3	105.7	105.6	106.14	1.3
16	102.7	102.8	102.5	100.0	100.7	101.74	2.8
17	102.5	102.9	103.2	103.8	104.8	103.44	2.3
18	106.3	103.8	104.0	104.0	104.0	104.42	2.5
19	103.9	103.7	103.7	104.6	103.5	103.88	1.1
20	105.5	101.2	102.8	103.4	101.9	102.96	4.3
計	2066.7	2062.7	2058.3	2061.8	2065.3	10314.8	47.3
平均	103.335	103.135	102.915	103.090	103.265	103.148	2.365
	\bar{x}_1	\bar{x}_2	\bar{x}_3	\bar{x}_4	\bar{x}_5	$\bar{\bar{x}}$	\bar{R}

$$s_{\bar{x}_i} \doteq 0.146$$

$$s_{\bar{x}_i} \doteq 1.83^{1)}$$

$$s_{\bar{x}_{20}} = \frac{1.83}{\sqrt{20}} = 0.409$$

$$\frac{s_{\bar{x}_{20}}}{s_{\bar{x}_i}} \doteq 2.8$$

$$R/d_2 = 1.02 = \hat{\sigma}$$

$$\text{cf} \left\{ \begin{array}{l} \hat{\sigma}/\sqrt{20} = 0.23 \\ \downarrow \\ s_{\bar{x}_i} = 0.15 \end{array} \right.$$

1) 全体のデータより度数表を作成し計算せるもの §2・3・1 表2・3と同様に。

をとって、この5コずつの母集団を一つの層と考えると、全母集団を20に層別して、その各層から各1コずつのサンプルをとりその平均値をもとめたような形になっている¹⁾。この場合各層の中からのサンプリングは一定間隔におこなっているので完全なランダムサンプリングとはいえない場合もあるので、完全な層別サンプリングの形にはならないこともあるが、一種の層別サンプリングの効果を発揮するのである。したがってこの各層間に大きな級間分散をもつようなとき、たとえば工程が管理状態になく、大波や傾向をもっているようなときには級間分散が大きくなるので、系統的サンプリングを行うと全くランダムなサンプリングを行ったときより平均値の精度はよくなる。この例でも、データを順に5コずつ群にとって管理図(図4・1)を書くと R 管理図は管理状態にあるが、 \bar{X} 管理図では限界外に出ている点が5点もあり、級間分散(群

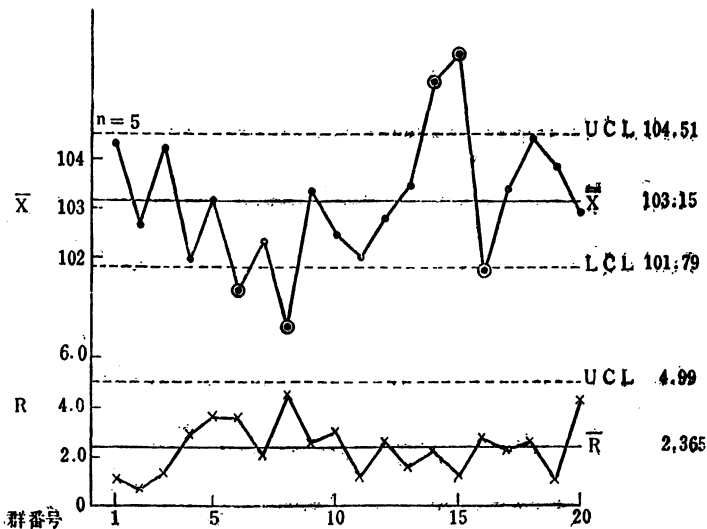


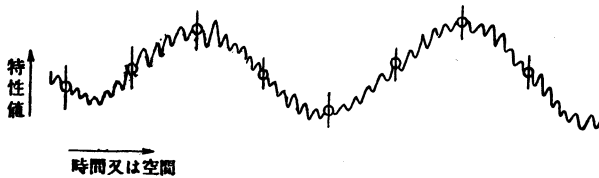
図 4・1 ある製品の重量の管理図

1) この例では実は各群毎に別の層である。

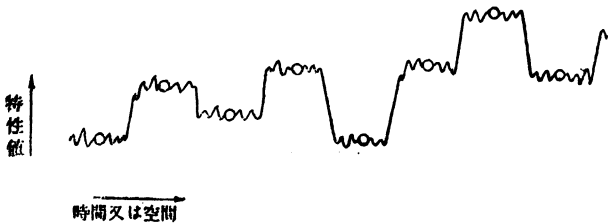
間分散)の大きいことを示している。級内分散は R 管理図は管理状態にあるから各層とも大体同じと認めてよいようである。 \bar{R} とり (4.20) 式を用いて $\hat{\sigma}_{\bar{x}_{20}} = 0.47$ となり $\sigma_{\bar{x}_{20}} = 0.41$ に近い。

注意 1) このようなことは、管理状態にない工程からの製品のロットの平均値を知るためのサンプリグのときによくおこる現象である。

2) 化学製品や、石炭、鉱石のようなものの製品の流れ、たとえば石炭や鉱石の輸送中に、流れの量に応じて一定間隔で1インクリメントずつサンプルをとり、これを混合して、混合試料として分析したときには、その平均値の精度は各インクリメント間のバラツキよりランダムサンプルであるとして、推定した平均値の精度よりも遙かによくなっていることが多い。したがって各インクリメントより標準偏差をもとめ、ランダムサンプルとして所要精度に必要なインクリメント数をもとめると、混



(i) 副次サンプリグの効果



(ii) 層別サンプリグの効果

図 4.2 系統的サンプリグ

合試料の精度は必要以上によくなっていることがある。いかえるとインクリメント数を取りすぎていることがあるから注意。

3) 系統的サンプリングは、以上のべたように、層別サンプリングの効果を発揮しているが、場合によっては § 4.5 にのべる副次サンプリングの形となっていることも多い。たとえば図 4.2 (i) のように、大波や傾向のあるときは、むしろ副次サンプリングに近いと考えた方がよい。

2. **ネイマン (Neymann) サンプリング** この方法は、各層へのサンプルの大きさ n_i の割当を、各層内の級内標準偏差 σ_i に比例して抜き取る方法である。すなわち k 層ある場合には、

$$n_1/N_1 \propto \sigma_1, \quad n_2/N_2 \propto \sigma_2, \quad \dots, \quad n_k/N_k \propto \sigma_k$$

または

$$n_i = n \frac{N_i \sigma_i^2}{\sum_i N_i \sigma_i^2} \quad (4.21)$$

これは $\sum n_i = n$ という条件で $\sigma_{\bar{x}}$ を最小にするには n_i をいかにとればよいかということからもとめたものである。このとき平均値の推定値は

$$\bar{x} = \sum_{i=1}^k p_i \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} = \sum_{i=1}^k p_i \bar{x}_i \quad (4.22)$$

$$\sigma_{\bar{x}}^2 = \frac{1}{n} (\sum p_i \sigma_i^2) - \frac{1}{N} \sum p_i \sigma_i^2 = \frac{1}{n} (\sum p_i \sigma_i^2) \quad (4.23)$$

このサンプリング法は、分散の大きい層からは採取比を大きくするのであるから、常識的にも考えられる結果である。 $\sigma_{\bar{x}}$ (ネイマン) は $\sigma_{\bar{x}}$ (比例) よりも小さくなり、平均値の精度はよいが、平均値をもとめるには、各層の重みつき平均値を用いなければならぬので、混合試料などをとるときには、全試料を混合するわけには行かず、各層毎に混合試料をつくり、これを分析して、その重みつき平均値をもとめなければならぬのである。

ネイマンサンプリングにおいて σ_i が一定のときは、通常の比例サンパ

1) σ_i は厳密にいうと $\sqrt{\frac{N_i}{N_i-1}} \sigma_i$ であるが、工業においては $N_i \gg 1$ が多いので

$$\sqrt{\frac{N_i}{N_i-1}} \approx 1 \text{ とする.}$$

リングとなり、工業においては大体 σ_i があまり違わぬ場合が多いので、比例サンプリグを行う場合が多い。

3. 経費を考慮に入れた割当法——デミング (Deming) サンプリグ
更に経費を考慮に入れて、たとえば、第 i 層での 1 測定当りの経費を c_i とすると、総経費 $C = \sum c_i n_i$ を一定として、 $\sigma_{\bar{x}}$ を最小にするためにはサンプルを次のように割当ればよい。

$$n_1/N_1 \propto \sigma_1/\sqrt{c_1}, \quad n_2/N_2 \propto \sigma_2/\sqrt{c_2} \quad \dots \quad n_k/N_k \propto \sigma_k/\sqrt{c_k}$$

または

$$n_i = \frac{C}{c_i} \frac{N_i \sigma_i \sqrt{c_i}}{\sum_j N_j \sigma_j \sqrt{c_j}} = \frac{C}{\sum_j N_j \sigma_j \sqrt{c_j}} \left(\frac{N_i \sigma_i}{\sqrt{c_i}} \right) \quad (4.24)$$

このときの平均値の推定およびその精度は、(4.4~5) および (4.6~9) 式であらわされる。これをデミングサンプリグという。

c_i が一定ならば、ネイマンサンプリグと同じになる。工場においては c_i はほとんど一定の場合が多いので、デミングサンプリグを用いることは少なく、 c_i 及び σ_i がほぼ一定の場合が多いので、そのとき $\sigma_{\bar{x}}$ を最小にする比例サンプリグ法が多く用いられる。

4.3.1.2 各層に関する知識も問題にするとき 各層に関する平均値をすべて同じ精度で知りたいときには、各層ごとに § 2 のランダムサンプリグでのべたように $\sigma_{\bar{x}}$ を級内分散とすると、 $\sigma_{\bar{x}} = \sigma_x/\sqrt{n}$ という関係より必要なサンプル数を決めるべきである。

各層の平均値も相当よい精度で得たいし、また全体の平均値の精度もよくしたいという場合には、前にのべた方法でもとめた、各層へのサンプルの割当数と、各層毎にランダムサンプリグの式でもとめた割当数とを考慮して、両者の中間的な値で妥協すべきであろう。たとえば層の大きさ N_i が小さいために、第 i 層からとるサンプル数 n_i が、比例サンプリグでは $n_i=3$ になり、層についてのバラツキから考えると $n_i=20$ と出たとすると、たとえば10くらいにとることもある。妥協していくらに決めるかは、そのサンプリグの目的と技術的経験的知識とからきめられるべきもので

一般にはいえぬ。たとえば $\sqrt{N_i}$, $\sqrt[3]{N_i}$ などに n_i をとることもある。

参考 またこの \sqrt{N} , $\sqrt[3]{N}$ に n をとるということは、単純ランダムサンプリングについてもよく行われている。 $\sigma_{\bar{x}}^2 = \sigma_x^2/n$ という関係から、 σ_x が一定ならば n は N には関係しないが、ロットが大きくなるとその結果が経済的に大きな影響をあたえることになるので、 n をロットの大きさ無関係という統計的結果と、経済的な問題との妥協で \sqrt{N} , $\sqrt[3]{N}$ と n を少し増加することがある。たとえば § 7. にのべるドラム罐やガロン罐のサンプリングのときには $\sqrt[3]{N}$ に n をとっている。

4.4 集落サンプリング (cluster sampling)

母集団をいくつかの層にわけ、その層の中からランダムにいくつかの層をサンプリングし、とった層はすべて調査する方法である。この層を集落 (cluster) という。社会調査において、層を町や村や調査区などとするので、集落という名前がついたのである。

層に分ける点は層別サンプリングと同じであるが、層別サンプリングではすべての層からサンプル単位をランダムにとるのに反して、集落サンプリングでは、いくつかの層をランダムにえらび、その層については全部調べるという点が丁度逆になっている。したがって、層をサンプル単位と考えた、単純ランダムサンプリングとも考えられるが、集落サンプリングでは集落をサンプル単位と考えずに、その中のサンプル単位についての品質特性を問題とし、従ってその集落の大きさと、その集落内の性質 (μ_i, σ_i^2) が、サンプル単位についての推定値の精度にどのような影響をもつかを考えている点が異なる。

集落サンプリングは、一般に各集落内を出来るだけ不均一に、各集落間のバラツキを出来るだけ小さくなるように層別する点が層別サンプリングと逆である。常識で考えてもわかるように、集落間分散 σ_b^2 が 0 であれば、一つの集落を十分に調査すれば、それで十分なことがわらう。

たとえば、工作機械が何台かあり、あるいは多頭の機械において、各機械をセットするやり方により、機械間に大きなバラツキが起りやすいか、一度セットしてしまえば、あとは各機械内には大したバラツキが起らぬことが技術的にわかっているようなときには、一定時間に各機械から出て来

る製品を集落と考えれば、1日の製品の平均品質を知るためには、一定時間毎に、すべての機械からでてくる製品を各1コずつ調査すればよい。これは一種の集落サンプリングである。また仄入、箱詰などの製品があるときに、仄あるいは箱をいくつかランダムにとり、この容器に入っている製品を全部調査するのも集落サンプリングである。集落サンプリングをうまく行う例としては、たとえば底の平らな壁の垂直なタンクに長く静置された液体が、水平の層には全く均一で、深さが深くなるに従って、たとえば比重が大きくなっているような場合には、タンクの上部からサンプリングパイプを入れて液を乱さないように、表面から底までの溶液を同じ割合でサンプリングしたときも一種の集落サンプリングである。サンプリングパイプ中の液を集落と考えれば垂直方向には液が不均一になっているので、集落内のバラツキは大きいですが、水平方向には均一、すなわち集落間のバラツキは小さい。したがって水平の層の均一性を保って、サンプリングが行われていれば、とったサンプルを十分に攪拌してその一部をとって測定すれば、そのデータがそのままタンク全体の液の平均値を示す。この場合集落間にバラツキのある恐れがあるときは、念のために数箇所からサンプルをとればよいであろうが、繰り返しサンプリングしたために、折角の集落を乱すようではかえって悪い影響をあたえるから、サンプリング技術に注意を要する。

また図4・3のように、管理状態の工程があるときには、一定間隔で数箇所ずつのサンプルをとったときにも、集落サンプリングと見なせる。すなわち管理状態の工程では、系統的サンプリングを相当に長い時間間隔で行っ

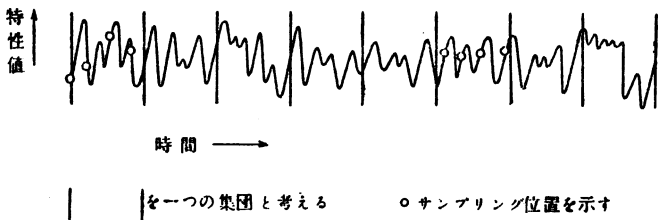


図 4・3 管理状態の工程

ても、その集落間のバラツキが少ないので全体の工程平均を精度よく推定できる¹⁾。

母集団については、次式のようになる。層別サンプリングのときと同様に

$$\mu = \frac{1}{N} \sum_i^M \sum_j^{N_i} x_{ij}$$

$$\sigma^2 = \sum_{i=1}^k p_i \frac{\sum_j^{N_i} (x_{ij} - \mu_i)^2}{N_i} + \sum_{i=1}^k p_i (\mu_i - \mu)^2$$

集落サンプリングのときの推定の式はランダムサンプリングのときと似ている。ただし集落サンプリングにおいて、推定したいのは、サンプル単位当りについてである。この場合もし集落当りの品質特性を問題にするのならば、これは単純ランダムサンプリングと全く同様になる。したがって、サンプリングの目的が、集落を構成しているサンプル単位の平均値であるか、集落当りの平均値であるかを注意しなければならない。

平均値の推定には

$$\hat{\mu} \leftarrow \bar{x} = \frac{M}{Nm} \sum_i^m X_i = \frac{M}{Nm} \sum_i^m \sum_j^{N_i} x_{ij} \quad (4.25)$$

$$= \frac{1}{\bar{N}m} \sum_i^m \sum_j^{N_i} x_{ij} = \frac{1}{n} \sum_i \sum_j x_{ij} \quad (N_i = \bar{N} \text{ のとき}) \quad (4.26)$$

$$\Rightarrow \frac{\sum_i^m \sum_j^{N_i} x_{ij}}{\sum_i^m N_i} = \frac{1}{n} \sum_i \sum_j x_{ij} \quad (N_i \neq \bar{N} \text{ のときは比推定}^2)) \quad (4.27)$$

ただし M : 全集落の数

m : サンプリングした集落の数

$N = \sum_i^M N_i$: 全サンプル単位数

N_i : 第 i 集落に含まれるサンプル単位数

$n = \sum_i^m N_i$: サンプリングした全サンプル単位数

$N_i = \bar{N}$, あるいは $N_i \neq \bar{N}$ とみなしうることは、工業において多い。

1) 図4.3に示した方法は、サンプリングした集落を全部調べていないので、実は次にのべる副次サンプリングに相当する。

2) §4.8 参照。

このときには (4.26) 式となり, N , N_i をもとめなくとも \bar{x} をもとめることができるので便利である.

(4.27) 式は比推定の式でありカタヨリが僅かにあるが, 実用的にはこれを用いて差支えない場合が多い. この場合には, 平均値の精度は比推定の式を用いるべきである.

総計の推定には

$$\hat{X} = \frac{M}{m} \sum_i^m \sum_j^{N_i} x_{ij} \quad (4.28)$$

$$= \frac{N}{n} \sum_i \sum_j x_{ij} \quad (N_i \approx \bar{N} \text{ のときは比推定}) \quad (4.29)$$

分散は, ランダムサンプルングのときの式において, 集落の大きさを考慮に入れて, σ^2 のかわりに σ_b^2 を, n のかわりに m を入れればよい. すなわち

$$V(\bar{x}) = \left(\frac{M}{N} \right)^2 \frac{M-m}{M-1} \frac{\sigma_b^2}{m} \quad (\text{一般式}) \quad (4.30)$$

$$= \left(\frac{M}{N} \right)^2 \frac{\sigma_b^2}{m} \quad (m/M < 0.1) \quad (4.31)$$

N_i が一定ならば $N_i = \bar{N}$

$$V(\bar{x}) = \frac{M-m}{M-1} \frac{\sigma_b'^2}{m} \quad (4.32)$$

$$= \frac{\sigma_b'^2}{m} \quad (m/M < 0.1) \quad (4.33)$$

ただし

$$\sigma_b^2 = \frac{1}{M} \sum_i^M (\sum_j^{N_i} x_{ij} - \bar{X})^2 \quad \bar{X} = \frac{1}{M} \sum_i^M X_i \quad X_i: \text{第 } i \text{ 層の総計}$$

$$\sigma_b'^2 = \frac{1}{M} \sum_i^M (\mu_i - \mu)^2 \quad \mu_i = \frac{1}{\bar{N}} \sum_j^{\bar{N}} x_{ij}: \text{第 } i \text{ 層の母平均}$$

総計についての推定値の精度は上式を N^2 倍すればよい.

つぎにランダムサンプルングと集落サンプルングの平均値の精度を比較して見よう. N_i が一定で \bar{N} のときには, 集落サンプルングでは

$$V_o(\bar{x}) = \frac{M-m}{M-1} \frac{\sigma_b'^2}{m} \approx \frac{M-m}{M} \frac{\sigma_b'^2}{m}$$

ランダムサンプリングでは同じサンプル数 $m\bar{N}$ をとれば

$$V_R(\bar{x}) = \frac{N-m\bar{N}}{N-1} \frac{\sigma^2}{m\bar{N}} \approx \frac{N-m\bar{N}}{N} \frac{\sigma^2}{m\bar{N}} = \frac{M-m}{M} \frac{\sigma^2}{m\bar{N}}$$

したがって α を次のようにとれば、 $\alpha \cong 1$ によりいずれが精度がよいかきまる。

$$\alpha = \frac{V_C(\bar{x})}{V_R(\bar{x})} = \bar{N} \left(\frac{\sigma'_b}{\sigma} \right)^2 = \frac{\bar{N}\sigma'_b{}^2}{\sigma^2} \quad (4.34)$$

$\bar{N}\sigma'_b{}^2 > \sigma^2$ ならば、あるいは $\sigma'_b{}^2 > \frac{\sigma^2}{\bar{N}}$ ならば、すなわち \bar{N} が一定ならば σ'_b が大きくなると、ランダムサンプリングの方が精度がよく、 σ'_b が小さくなり $\bar{N}\sigma'_b{}^2 < \sigma^2$ となれば、集落サンプリングの方がよい。すなわち集落間分散 $\sigma'_b{}^2$ が小さいか、あるいはこのように集落をえらべるときには、集落サンプリングがよい。

次に (4.34) 式からもう一つ考えなければならぬことは集落の大きさ \bar{N} である。 σ'_b が一定とすれば、 \bar{N} が大きい程、すなわち各集落の大きさが大きい程総平均値の精度は悪くなる。ところが実際には、集落の大きさが大きくなれば、 σ'_b が小さくなる場合が多く、 σ'_b は二乗できいてくることを考慮に入れて、集落化、層別を行わなければならない。また \bar{N} を小さくすると、 σ'_b が大きくなり、したがって所要精度をうるためには、とるべき層の数 m が増加してくるので、層を、たとえば箱や吠を抜き出し、運搬し、開ける費用が増加するから、この費用も考慮に入れていづれのサンプリング法をえらぶかきめるべきであろう。

集落のとり方 集落サンプリングにおいて、どの集落をサンプルとしてとるかというやり方は、各集落の大きさ N_i がほとんど一定のときには、集落を単位として、通常のランダムサンプリングと全く同様に、所要精度をうるに必要な集落数 m を乱数表などでランダムにサンプリングすればよい。

集落の大きさ N_i が相当にことなっているときには、集落を単位としてサンプリングすると、集落サンプリング法では集落内の2次単位の特性値の調査が目的であるので、2次単位当りの平均特性には、大きな集落は大きな影響をあたえ、小さな集落は小さな影響しか与えぬので、集落を同じ確率でとると結果がカタヨルおそれがある。そこで次にのべるような方法を

用いるとよい。

たとえば大工場において、平均賃金をもとめたい場合を考えよう¹⁾。各係を集落と考へて、調査するにはいかにサンプルングしたらよいか。係が300もあり、在籍員が2万名近くもあると集計にも時間や経費がかかるのでサンプルングにより調査しようというのである。まず表4.2のような表を作成し²⁾、各在籍員の累積和をつくる。いま平均値の精度の式より $m=18$ が必要であることがわかったとする。在籍員の総計が18,675名であるから、乱数表より00,001から18,675の間の数字18コをランダムにとる。

表 4.2 係と在籍員

係名	在籍員 N_i	累積和
1	38	38
2	144	182
3	92	274
4	25	299
⋮	⋮	⋮
M	38	18,675

かくして得られた乱数の一つがたとえば00235と出たとすると、この数字を含む係3をサンプルとする。かくして18の数字について18の係を抜き取る。そしてこの18の係について全員の給与をしらべる。この場合もし同じ係の人が2人あたれば、その係については、結果を2倍して計算すればよい。このようにすればカタヨラないサンプルングをすることが出来る。このサンプル

ング法を**確率比例サンプルング法** (size proportionate sampling; または sampling with probability proportionate to size) という。この方法は集落サンプルング以外にも用いられる。たとえば通常のランダムサンプルングにおいても、各サンプル単位の特性値の大きさが非常にことなるものがあるようなときには、その特性値の単位を考慮に入れて、過去のデータよりそれに比例させてサンプルングすればよいことがある³⁾。

またこの集落サンプルングは、工場に多くある工作機械、紡糸機、紡口

- 1) この場合集落サンプルングで行くべきか、層別サンプルング、副次サンプルングその他の方法で行くべきかという問題はあるが、それは場合によりことなるので、ここではただサンプルングのやり方を説明する意味でのべる。
- 2) この場合サンプル単位について、リストができていれば、このような表を作成する必要はなく、このリストからランダムサンプルングし、その単位の属している集落をしらべればよい。
- 3) この場合勿論層別サンプルングをした方がよい場合も多い。

などの機械装置,あるいは各係等の工具,消耗品,貯蔵品などの状況を調べるにも用いられる.たとえば工具などは全くランダムに各係内の各工具をしらべるのは大変な手間がかかるから,各係をえらび,その係のものを全部しらべるようにすれば,係毎にしらべればよいので楽である.集落サンプリングの集落が各地域的にきめられる場合に,これは1種の地域サンプリング法 (area sampling) である.

4.5 副次サンプリング (subsampling)・2段サンプリング (two-stage sampling)・多段サンプリング (multi-stage sampling)

集落サンプリングは,集落をサンプリングし,それを全部測定するのであるが,副次サンプリングでは,まず層別した層の中のいくつかをサンプリングし,次にサンプルした層の中から,いくつかの単位体をサンプリングする方法である.たとえば多くの箱詰の電球,ボルト,紐,束にした鋼材,缶詰や瓶詰の化学薬品,貨車積の石炭などがあるときに,まず箱,束,缶,貨車などをランダムにえらび,次にその箱などの中から数十本のボルト,数枚の鋼板,数十瓦の薬品,数缶の石炭などをサンプルとしてとる場合がこれに相当する.また特殊な場合としては石炭などの粉塊状物の大口試料から縮分して,分析用試料をつくるのもこれに相当する.

このサンプリング法では,まず第1段階として,箱,束,缶,貨車などを単位としてサンプリングする.これらのサンプリング単位を1次サンプリング単位 (primary sampling unit, 以後1次単位という) といい,これに対し,得た1次単位から第2段階としてとるサンプル単位を2次サンプリング単位 (secondary sampling unit, 以後2次単位という) という.このように2段にサンプリングするので,2段サンプリング法ともいう.

さらに,たとえば大箱入の電球,タバコ,石鹼などをサンプリングするのに,まず第1段階として,大箱をサンプリングし,さらにその大箱から小箱をサンプリングし,次に小箱中から数コの電球などをとるという3段サンプリング法も考えられる.さらに4段,5段サンプリングも考えられ,これら2段以上のサンプリング法を多段サンプリング法という.これに対しこれまでのべた,ランダムサンプリング法,層別サンプリング法,集落サ

ンプリング法はいずれも1段階サンプリング法 [one (or single)-stage sampling] である。

1段目の抜取比 a_1 , i 段目の抜取比を a_i とすると, k 段階サンプリングでは, 全抜取比 a は, 各段階毎に比例サンプリングすれば $a = a_1 \times a_2 \times \dots \times a_k$ となる。

副次サンプリング法は, 工業において多く用いられ, 経費の点からも, 精度の点からも極めて有利なサンプリング法となる場合が多い。

いま M コの1次単位よりなり, 全数 N コの2次単位をもつロットがある。その各1次単位が N_i コの2次単位から成立し, 各2次単位の特性値を x_{ij} とし, その母平均を μ_i とすると, 母集団については

$$\mu = \frac{1}{N} \sum_i^M \sum_j^{N_i} x_{ij} = \frac{1}{N} \sum_i^M N_i \mu_i = \sum_i^M p_i \mu_i \quad (N = \sum_i^M N_i) \quad (4.1)$$

$$\begin{aligned} \sigma^2 &= \frac{1}{N} \sum_i^M \sum_j^{N_i} (x_{ij} - \mu)^2 = \frac{1}{N} \sum_i^M \sum_j^{N_i} (x_{ij} - \mu_i)^2 + \frac{1}{N} \sum_i^M N_i (\mu_i - \mu)^2 \\ &= \sum_i^M p_i \sigma_i^2 + \sum_i^M p_i (\mu_i - \mu)^2 = \sigma_w^2 + \sigma_b^2 \end{aligned} \quad (4.3)$$

いまこのロットから m コの1次単位をとり, その各1次単位から n_i コのサンプルをとるとすると, サンプルからの総計及び母平均の推定式は

$$\hat{X} = \frac{M}{m} \sum_i^M \frac{N_i}{n_i} \sum_j^{n_i} x_{ij} \quad (4.35)$$

いま1次抜取比 $a_1 = \frac{m}{M}$, 2次抜取比 $a_{2i} = \frac{n_i}{N_i}$ とすると

$$\hat{X} = \frac{1}{a_1} \sum_i^M \frac{1}{a_{2i}} \sum_j^{n_i} x_{ij} \quad (4.36)$$

比例サンプリングで $a_{21} = a_{22} = \dots = a_{2i} = \dots$ 一定のときは

$$= \frac{1}{a_1 a_2} \sum_i^M \sum_j^{n_i} x_{ij} \quad (4.37)$$

$$= \frac{N}{n} \sum_i^M \sum_j^{n_i} x_{ij} \quad (4.38)$$

$$\hat{\mu} \leftarrow \bar{x} = \frac{M}{Nm} \sum_i^M \frac{N_i}{n_i} \sum_j^{n_i} x_{ij} \quad (4.39)$$

$$= \frac{1}{n} \sum_i \sum_j x_{ij} \quad (a_2 \text{一定}) \quad (4.40)$$

工業においては a_2 を大体一定にとる場合が多いので (4.40) 式を適用できる場合が多い。

これらの分散は

$$V(\bar{x}) = \frac{1}{N^2} \left\{ M^2 \frac{M-m}{M-1} \frac{\sigma^2_e}{m} + \frac{M}{m} \sum_i \left(N^2_i \frac{N_i - n_i}{N_i - 1} \frac{\sigma^2_i}{n_i} \right) \right\} \quad (4.41)$$

$$\Rightarrow \frac{1}{N^2} \left\{ \frac{1-a_1}{a_1^2} m \sigma^2_e + \frac{1-a_2}{a_1 a_2^2} \sum_i n_i \sigma^2_i \right\} \quad (4.42)$$

(さらに $M \gg 1$, $N_i \gg 1$, $n_i/N_i = a_2$ 一定のとき)

$$\Rightarrow (1-a_1) \frac{\sigma^2_b}{m} + (1-a_2) \frac{\sigma^2_w}{m\bar{n}} \quad (4.43)$$

(さらに $N_i = \bar{N}$, すなわち $n_i = \bar{n}$ 一定のとき)

$$\Rightarrow \frac{\sigma^2_b}{m} + \frac{\sigma^2_w}{m\bar{n}} = \frac{\sigma^2_b}{m} + \frac{\sigma^2_w}{n} \quad (4.44)$$

(さらに $m/M < 0.1$, $n_i/N_i = a_2 < 0.1$ のとき)

$$\text{ただし } \sigma^2_e = \frac{1}{M} \sum_i (X_i - \bar{X})^2 \quad \sigma^2_i = \frac{1}{N_i} \sum_j (x_{ij} - \mu_i)^2$$

N_i が一定のときには

$$\sigma^2_b = \frac{1}{M} \sum_i (\mu_i - \mu)^2 \quad \sigma^2_w = \frac{1}{MN} \sum_i \sum_j (x_{ij} - \mu_i)^2$$

工業においては, $M \gg 1$, a_2 が一定で 0.1 よりも小さく, N_i が一定で 1 より遙かに大きく, a_1 も 0.1 より小さいという場合が多いので, (4.38, 40, 44) 式などのいずれも, 最も簡単な式が近似的にあてはめられる場合が多い。

\hat{X} の分散は, これまでと同じように \bar{x} の分散を N^2 倍すればよい。

4.5.1 サンプルのとり方 サンプルの最適配分

- 1) $m=M$ のときは層別サンプリングとなる。また $n_i=N_i$ のときは集落サンプリングとなるすなわち層別サンプリング。集落サンプリングは副次サンプリングの特殊の場合とも考えられる。

副次サンプルング法により、サンプルングする際には、サンプルの割あて方とそのとり方に注意しなければならない。

われわれはもっとも精度よくサンプルの平均値より、母平均を推定しようというのであるが、これには(4.44)式と経費あるいは所要精度を考えて、何コの1次単位 m 、何コずつの2次単位 n_i をとったらよいかというサンプルの割あて方のことを考えなければならぬ。

1. サンプルング並びに測定、データの整理の費用が、2次単位のサンプル数 n によりきまる場合　すなわち1次単位のサンプルングについての費用が全体の費用に比して非常にやすいような場合には、総経費 T は次の式であらわされる。このような式を費用函数 (cost function) という。

$$T = k_2 n = k_2 m \bar{n} \quad (4.45)$$

ただし k_2 : 2次サンプル単位1コ当りの費用

このときに、 T が一定の場合には、サンプルの数 $n = m \bar{n}$ が決ってくる。したがって、通常の場合には、(4.44)式

$$V(\bar{x}) = \frac{\sigma_b^2}{m} + \frac{\sigma_w^2}{m \bar{n}}$$

において、 $V(\bar{x})$ を最小にするには、 σ_b 、 σ_w が一定で、 $m \bar{n}$ が一定ならば、 m が大きい方がよい。すなわち $\bar{n} = 1$ 、 $m = n$ にとればよい。たとえば、工程中を流れる製品をいくつかの層にわけて、副次サンプルングするような場合には、層をえらび、1次サンプルングする費用がほとんどかからぬから、なるべく多くの層 n コをとり、各層からは1コずつサンプルをとればよいことになる。このときには精度は

$$V(\bar{x}) = (\sigma_b^2 + \sigma_w^2)/n = (\sigma_b^2 + \sigma_w^2)k_2/T \quad (4.46)$$

$$m = T/k_2, \quad \bar{n} = 1 \quad \text{となる。}$$

また $V(\bar{x})$ を所要精度 B におさえて、経費を最小とするときにも同様に、 $\bar{n} = 1$ 、 $m = T/k_2$ となる。このときの m は次式であらわされる。

いま(3.10)式より $\beta^2 = t^2 V(\bar{x})$ とすると、(ただし t は信頼度と自由度によりきまる係数)

$$m \approx \frac{\sigma_b^2 + \sigma_w^2}{(\beta/t)^2} \quad (4.47)$$

$$\therefore T_{\min} = k_2 \frac{\sigma_b^2 + \sigma_w^2}{(\beta/t)^2} \quad (4.48)$$

2. 費用が1次単位のサンプル数 m によりきまる値 k_1 と、2次単位のサンプル数 n によりきまる値 k_2 とに分けて考えられる場合 この場合にはいろいろな費用函数が考えられるが、多くの場合に近似的に適用できる次式の場合について論じよう。

$$T = k_1 m + k_2 m \bar{n} \quad (4.49)$$

費用 T が一定で、平均値の誤差分散を最小とするには

$$m = \frac{T}{k_1 + k_2 \bar{n}} \quad (4.50)$$

この式を (4.44) 式のように簡単な場合を考えて、これに代入し、 $V(\bar{x})$ を \bar{n} の函数としてあらわし、 $V(\bar{x})$ が最小、すなわちこの函数を \bar{n} で微分して、それが0となるという条件でとくと

$$\bar{n} = \sqrt{\frac{k_1}{k_2}} \frac{\sigma_w}{\sigma_b} \quad (4.51)$$

$$\text{したがって} \quad m = \frac{T}{k_1 + \sqrt{k_1 k_2} (\sigma_w / \sigma_b)} \quad (4.52)$$

なるときに $V(\bar{x})$ が最小となる。変動係数 $C_{\bar{x}}$ を最小にするときも、(4.51, 4.52) 式に σ_w , σ_b のかわりに C_w , C_b を代入すればよい。

$V(\bar{x})$ を所要精度 β で押えて総経費 T を最小にするためには、前と同様にして、 $\beta^2 = t^2 V(\bar{x})$ なる故

$$\bar{n} = \sqrt{\frac{k_1}{k_2}} \frac{\sigma_w}{\sigma_b}, \quad m = \frac{\sigma_b^2 + \sigma_b \sigma_w \sqrt{\frac{k_2}{k_1}}}{(\beta/t)^2} \quad (4.53)$$

となる。

この2つの場合において、最小精度あるいは最小総経費をもとめるには、かくしてもとめた (4.51) (4.52) 式、あるいは (4.53) 式の値をそれぞれ (4.44) 式あるいは (4.49) 式に代入すればよい。

以上の結果から、 \bar{n} は所要精度、所要総経費には関係なく k_1/k_2 , σ_w/σ_b という相対的な値により決ってくることがわかる。所要精度が小さくなっ

1) 一般には、固定費 k があるから $T = k + k_1 m + k_2 m \bar{n}$ となる。

でも、所要総経費の多少にかかわらず、 \bar{n} はきまっでいて、変化するのは1次単位をとる数 m であるということは注目すべきことである。

(4.41~43) 式という条件、あるいは費用函数が他の形をとるときには、それぞれの値を上例と同様に代入して、解けばもとめられる。

例 1) 吸入の薬品が倉庫につんである。この吸入をとり出し、それからサンプルをとりたい。吸入をとり出す費用は、1吸入あたり k_1 で、それからサンプルをとり出す費用は1サンプル当り k_2 である。いま $k_1/k_2 \approx 24$ (すなわち倉庫に積みこんでしまった製品をとり出す費用は非常に大きいので、実際にこのようなことがよくある)、 $\sigma_w/\sigma_b = 1$ ($\sigma_w/\sigma_b \approx 1$ に近いことが多い) とすると、 \bar{n} , m をいくらにとったらよいか。

$$\bar{n} \approx \sqrt{24} \times 1 \approx 4.9 \approx 5$$

したがって所要精度 β 、あるいはきめられた総経費 T により m を決定すればよい。

例 2) 前例において、薬品吸入を倉庫に積みこむ途中においてサンプルリングするときには $k_1/k_2 \approx 1$ である。このときには \bar{n} , m , をいかにとったらよいか。

$$\bar{n} = \sqrt{1} \times 1 = 1$$

すなわちこの場合には、各吸入から1サンプルずつサンプルをとり、 $m = n$ というように、出来るだけ多くの1次単位をとった方がよい。この例において、もし $\bar{n} = 5$ ととれば、必要経費、あるいは精度は最小にならない。

副次サンプルングを実施する場合にもやはり、各2次単位が同じ確率でランダムにとられるように工夫しなければならない。1次単位の大きさ N_i がほぼ一定で、1次単位から2次単位を抜き取る比 a_2 も一定のときには、まず1次単位をランダムに、次に各1次単位から2次単位をランダムに抜き取ればよい。ところが N_i が一定でないのに、 n_i を一定にして、1次、2次サンプルングともランダムサンプルングすると

$$\frac{m}{M} \times \frac{\bar{n}}{N_1} \approx \frac{m}{M} \times \frac{\bar{n}}{N_2} \quad (\because N_1 \approx N_2)$$

したがって、2次単位のとられる確率がことになってくるので結果はカタヨる。この場合には1次単位のサンプルング法は、確率比例サンプルングで

N_i を考慮に入れて抜き取り，次の2次単位の抜き取りを n ずつランダムに行えばよい¹⁾。このときは

$$\frac{mN_1}{MN} \times \frac{n}{N_1} = \frac{mN_2}{MN} \times \frac{n}{N_2} = \frac{mn}{MN} \quad \text{一定}$$

となる。

1次単位についての情報を得たい場合は，このように N_i にかかわらず n を一定にとることがあるから，このときにはサンプリング法に注意しなければならない。しかし工場における資材のサンプリングにおいては，一般に N_i がほぼ一定の場合が多いので，1次，2次サンプリングともランダムに行ってよい場合が多い。

(4.35) 及び (4.39) 式で示されるように， N_i も変化し， n_i も変化するようにサンプリング法は工場で実際に用いられる場合は少ない。

注意 確率比例サンプリングは，異なった大きさの集落，層を単位としてサンプリングするときによく用いられる方法である。集落を単位とったときに，その集落中の各サンプル単位（2次単位）がカタヨらないように，同じ確率で抜き取られなければならない。これには，

1) 集落を抜き取るときに，その集落の大きさ N_i に比例した確率で抜き取り，次に抜き取られた各集落から，2次単位を同数 n ずつとる。

2) 集落をその大きさ N_i に関係なく同じ確率で抜き取り，次に2次単位は抜き取られた集落から N_i に比例して一定抜取比 α_2 で抜き取る。

3) 集落をその大きさ N_i に関係なく同じ確率で抜き取り，次に2次単位は抜き取られた集落から一定数 n ずつとる。

以上の3方法が考えられるが，本文でのべたように，1) はカタヨリがないが，3) は N_i が大きく異なっているときには，カタヨリが入る。これは大きい集落に対し重みがかかっていないからである。2) も精度があまりよくないが，それと同時に，はじめの設計において，全サンプル数 n が相当に変化して，所要精度がえられない場合もでてこよう。結局 1) がよい。

1) を実施するには，§4.4 でのべたような方法で1次単位をえらべば

1) §4.4 並びに本節注意参照。

よい。これを系統的に行うこともできる。§4.4の例において18,675/18=1037であるから、まず1~1037の中からランダムに1つの数字をえらび、たとえば0235と出れば、235, 235+1037, 235+2074, …というように18 \times を系統的にえらばよい。これは確率比例系統的サンプリングとなっている。このようにサンプリングの実際には、いろいろな方法を組合せて用いることが多い。

4.6 有意サンプリング (purposive sampling)

本書でははじめからのべて来たように、サンプリングはランダムに行われなければならないが、これとは反対に、選んでサンプリングする方法である。この方法は、そのロットに対する従来の知識や経験を用いて、全体の代表と思われる、あるいは平均値と思われるところを、各試料採取者の主観的な判断でサンプリングするので、どうしてもカタヨリが入りやすいから通常はさけた方がよい。特によいところをえらぶことになっていないのに、現場ではよいところをえらんでいたとなると、平均値の推定など非常にカタヨルから、このようなデータを統計的に処理しても意味がない。今後は極力有意サンプリングにならぬよう注意しなければならない。この方法によれば、時間や経費が非常にはぶける場合もあるので、従来よく用いられているが、一般には避けるべきである。

しかし特別の場合には、その試験の目的により、この方法を用いると時間や経費が非常にはぶけて、しかもサンプリングも容易になるので用いられることがある。たとえば、コークスの強度試験のように、炉の部分により強度のバラツキが非常に大きくかつ粒度に相当大小があるようなときには、その試験法の目的により、たとえばある粒度以下のものは熔鉱炉には使用しないので、途中で篩分してしまう場合には、このような粒度のサンプルは有意的に取らなくてもよい。また大きい粒度のものも全体からランダムにとるとバラツキが非常に大きくなり、平均値の精度がわるくてこまるようなときには、試験法をかえるべきであるが、もし試験法をかえることも困難であるような場合には、たとえばもし炉中心部のコークスには比較的バラツキも少なく、その平均値が炉全体の平均値からどのくらいカタヨっているかが前もって十分に統計的に検討してあり、それが管理状態に

あり、しかも炉中心部のコークスを間違いなくサンプリングできるようであれば、工程管理などのためにはこれをサンプルとしてもよい。ここへのべたことでわかるように、このように有意サンプリングを行ってよいためには、相当多くの条件がついているので、一般にはこのような条件が成立することは少なく、避けるべきである。

有意サンプリングを行ってよいためには次の条件を満足していなければならない。

1) ランダムサンプリング、その他のサンプリング法でいろいろ工夫をしても経費、時間、試験法、並びに精度の点から見てどうしても困難であること。2) 有意サンプリングが、主観的にえらぶのでなく、客観的に誰にでも容易に出来ること。しかも3) 有意サンプリングの結果が、真の母平均と有意差がないか、あるいは有意差があっても、そのカタヨリが§3・4・3の測定の正確さのところでのべたように、母平均と一定の関係をもっていること。カタヨリは管理状態になければならない。4) 有意サンプリングを行うと、経費、時間、試験法並びに精度の点から見て非常に有利な場合。

このような条件を満足すれば有意サンプリングを行った方がよい場合がある。

4・7 ある期間中の平均サンプリング (period average sampling)

ある期間中(時間的にも、位置的にも)サンプルを時々あるいは連続的にとって、たとえば工程中のガス体や、荷上中の鉱石類のようにこれを一緒にあつめて、混合試料として測定、分析などを行う方法である。これは流体、粉体、粉塊状体などの平均品質を知りたいときで、かつ個々の測定が化学分析などのように厄介、面倒なときによく用いられる方法である。これは系統的サンプリングを行って、それを混合試料として測定している場合に相当する。このサンプリング法で得られる情報は、その期間中の平均値のみで、その間の工程の変動などについて、したがってこの結果よりそのサンプリング法による平均値の精度をもとめることができない。これはサンプリングの目的が、ある期間中、あるいは1山の石炭の平均値のみ

を知りたいときには経済的なサンプリグ法であるが、あらかじめ予備実験でこのサンプリグの精度をもとめておかなければ、その平均値の精度は判らぬし、また工程を管理したいためのサンプリグとしては、その目的により、あまり期間を長くすることは不適当な場合が多い。

このサンプリグ法を行う場合にも、精度を推定できるようにサンプリグを行っておいた方がよい。たとえば、一定間隔にとったサンプルを交互に2つの試料容器に入れておいて、別々に分析するというような工夫が払われるべきである。

たとえば硫酸を4時間毎に5インクリメントずつ毎日6回サンプリグし、これを一つ容器に入れておいて、10日分まとめてから縮分して分析し窒素分を測定したときには、10日間の平均窒素分はわかっても、果して毎日でている製品の窒素分がどう変動しているかはわからず、10日平均が保証成分より高くても、これだけの情報では毎日の製品の平均品質を保証することもできない。これには1時間毎、1日毎の変動を知らなければならず、その目的に応じた間隔の平均サンプリグ、あるいは瞬間サンプリグをしなければならない。

一般に工程を管理するためには、ある一定間隔毎にいくつかのサンプルを一度にとった方がよい場合が多いが、このように短期間にとるサンプリグを平均サンプリグに対して、**瞬間サンプリグ**、**スナップサンプリング** (snap sampling) という。この場合に問題となるのは、測定法の問題で、§ 3・2にのべたように、管理試験により、簡単、迅速に測定できなければ、スナップサンプリングはできない。

工程管理の場合でも、工程が次第に管理状態となり、そのバラツキが保証成分以下にならぬことを示しているような場合でも、スナップサンプリングの間隔を長くした方が、長期間の平均サンプリグを行うよりはよい場合が多い。

あるロットの平均品質を知りたいために、平均サンプリグを行うことがよくあるが、このとき注意しなければならぬことが2つある。

1) ロットがコンベヤー上あるいはパイプ中などを流れているときには、その量に比例してサンプルをとること。単位体の場合には数を勘定しながら系統的サンプリグ法などでとればよいが、非常に多数の単位体、

あるいは集合体の場合には、技術的に流れの量を測定あるいは大体判断しながら、流れる量に比例してサンプリングしなければならない。これを確実に行うことが困難な場合がある。パイプ中を流れる流体の場合には、流量計と結びつけてサンプリングするとか、あるいは石炭、鉱石などをクレーンでコンベヤーまではこんでいるときには、クレーン何杯毎に1インクリメントとするようにすれば流量に応じてとることが出来る。

2) 人間がある期間中の平均サンプルをとろうとするときには、人手の問題とか、試料採取者の不注意などにより、ある期間のサンプリングを確実に行ってないかも知れぬ恐れがあるから、平均サンプルを得るためには、なるべく自動試料採取機 (automatic sampling machine) を用いた方がよい。最近この方面の研究が各所で行われ出したのは喜ばしいことである。この場合、自動採取機を流量、計量機と結びつけて、流量と比例してとるようにするか、あるいは自動採取機は一定間隔で動くようにして、流量が一定になるように管理するか、いずれかの注意が必要である。

4・8 反覆サンプリング (multiple sampling)

ランダムサンプリングその他のサンプリング法を、唯1回行うだけでなく、同じ母集団から2回以上繰り返しサンプリングする方法である。これは抜取検査 (sampling inspection) においても、2回抜取検査、逐次抜取検査として計数並びに計量の場合に行われているが、母集団の推定を行う場合にも用いられる。工場などにおいても、大きなロットについて、初めて、あるいは稀にサンプリングするときに用いられる。たとえば

1) はじめ小さいサンプルで母集団についての知識をうるために、予備調査を行い、次に大きなサンプルをとり本調査を行う場合。

2) はじめの調査結果を再検討するために行う場合。

3) はじめに小さいサンプルで精密な測定法と簡単な測定法との相関関係をもとめ、後に大きなサンプルは簡単な測定法で本調査を行う場合、たとえば石炭の発熱量を一々測定するのが困難な場合に、灰分と発熱量とを小さい同じサンプルについて各測定し、後に本調査では灰分分析により調査を行い、全体の発熱量を推定する場合。

4) 母集団に2つの品質特性があるときに、はじめの小さなサンプルで

2つの特性の比をもとめ、次に大きなサンプルで簡単に測定できる品質特性を測定して、その比を利用して推定を行う場合。

5) はじめのサンプルでは、推定の精度が不十分であるときに、更にサンプル数を増加し精度をよくするためにサンプリグする場合。

などいろいろ考えられる。いずれも本章にのべた他のサンプリグ法を結びつけて、技術的に、あるいは経済的にサンプリグの目的を達するために行われる。

4・9 比推定 (ratio estimate)

これはむしろ推定法の一つであるので、本章に入れるのは妥当ではないが、この推定によるときのサンプルの大きさのきめ方の問題もあるので本章に入れた。

サンプル単位に2つの特性 x , y があるときに、この x/y の比を用いて精度よく推定を行う方法で、一般に x と y の間の相関係数の大きいときに用いると有利である。これは一種の相関を利用した推定法である。

たとえば薬品、フィルム、ビール、石鹼、布地その他、貯蔵中に品質の低下する恐れのあるものを大量に貯蔵しているときに、貯蔵前の特性値と、貯蔵後、現在の特性値の比を用いて、現在の全体の品質をサンプルから推定する場合、ある品質特性を測定するのに、簡略法と精密法とあるときに、製品あるいはサンプル全部を精密法で測定するのは不経済であるので、その一部をサンプリグして、精密法と簡略法と両方で測定し、その比を利用して、残りを簡略法で測定して（一種の反覆サンプリグ）、推定する場合、あるいは出荷側が、各サンプル単位の重量や品質特性などを通知して来たときに、その一部を荷受側でチェック試験して、その比より入荷品の重量や品質を推定する場合、副ロットの大きさが相当ことなる製品中の不良率、不良数を推定する場合などにおいて2つの品質特性の間の相関係数の大きいときに、この推定法を用いると有利な場合がある。

4・9・1 ランダムサンプリグの比推定

サンプル単位 N 個の大きさの母集団において、各サンプル単位につき2種類の特性 x , y , が考えられる場合に、 X , Y , をその総和とすると

$$\varphi = \frac{X}{Y} = \frac{\sum_{i=1}^N x_i}{\sum_{i=1}^N y_i} \quad (4.54)$$

という比が考えられる。この φ は一種の母数である。これに対し、ランダムに n 個のサンプルをとり、その特性 x, y を各々測定すれば同様にし、統計量 f

$$f = \frac{x'}{y'} \quad (4.55)$$

がえられる。ただし x', y' は各々 x, y についての不偏推定とする。総和でも平均値でもよい。

この f を比推定という。 x', y' がランダムに変化する値であるから、 f も一定の値ではなく母数 φ のまわりにバラツク。しかし f は φ の不偏推定ではないが、そのカタヨリは僅かであるから、実用上は不偏推定とみなしてよい。したがって

$$X' = fY \quad (\text{ただし } Y \text{ は母集団についての総和})$$

という X' を推定したときに、これは総和 X の不偏推定値 \hat{X} とみなすことができる。この推定値の分散は次式によりあらわされる。

$$\sigma^2_{X'} \approx \varphi^2 \frac{N-n}{N-1} \frac{V}{n} \quad (4.56)$$

$$\begin{aligned} \text{ただし } V &= \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \varphi y_i}{\mu_x} \right)^2 \\ &= \varphi^2 \frac{N-n}{N-1} \frac{1}{n} \left(\frac{\sigma_x^2}{\mu_x^2} + \frac{\sigma_y^2}{\mu_y^2} - 2\rho C_x C_y \right) \end{aligned} \quad (4.57)$$

$$\begin{aligned} \text{ただし } \text{Cov}(x, y) &= \frac{1}{N} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \\ &= \varphi^2 \frac{N-n}{N-1} \frac{1}{n} (C_x^2 + C_y^2 + 2\rho C_x C_y) \end{aligned} \quad (4.58)$$

$$V(X') \approx X^2 \frac{N-n}{N-1} \frac{1}{n} \left(\frac{\sigma_x^2}{\mu_x^2} + \frac{\sigma_y^2}{\mu_y^2} - 2 \frac{\text{Cov}(x, y)}{\mu_x \mu_y} \right) \quad (4.59)$$

$$= X^2 \frac{N-n}{N-1} \frac{1}{n} (C_x^2 + C_y^2 - 2\rho C_x C_y) \quad (4.60)$$

$$(O_{X'})^2 = (O_f)^2 = V(f)/\varphi^2 = V(X')/X^2 \quad (4.61)$$

$$\hat{\sigma}_x^2 \approx \frac{1}{n} (C_x^2 + C_y^2 - 2\rho C_x C_y) \quad (n/N < 0.1) \quad (4.62)$$

$$\hat{\sigma}_x^2 \approx \frac{2C_x^2}{n} (1-\rho) \quad (n/N < 0.1, C_x \approx C_y) \quad (4.63)$$

ただし $\sigma_x^2, \sigma_y^2; \mu_x, \mu_y; C_x, C_y; \rho$ は母分散; 母平均; 母変動係数; および xy についての母相関係数である。

ところが x の総和 X の変動係数 $(C_X)^2$ は通常のランダムサンプルングのときには, (2.49) 式より

$$(C_X)^2 = \frac{N-n}{N-1} \frac{C_x^2}{n}$$

したがって

$$(C_{X'})^2 = (C_f)^2 \cong \frac{N-n}{N-1} \frac{C_x^2}{n} = (C_X)^2 \quad (4.64)$$

$$\text{ただし} \quad \rho \cong \frac{C_y}{2C_x} \quad (4.65)$$

すなわち $\rho > \frac{C_y}{2C_x}$ ならば, $(C_{X'})^2 < (C_X)^2$ となり比推定の方が精度がよくなる。

この関係は比推定を用いるべきか否かの判定基準となる重要な関係である。

以上の φ, C_x, C_y, ρ などは母数として取扱ったが, 実際にはこれらが不明であるので, n コのサンプルからこの値を推定しなければならない。

このときには, 上にのべた精度の各式の推定値は次のようになる。

$$\begin{aligned} \hat{\sigma}_f^2 &\approx f^2 \frac{N-n}{N} \frac{1}{n(n-1)} \sum \left(\frac{x_i - fy_i}{\bar{x}} \right)^2 \\ &= \frac{N-n}{N} \frac{1}{n(n-1)} \sum \left(\frac{x_i - fy_i}{\bar{y}} \right)^2 \end{aligned} \quad (4.66)$$

$$\approx \frac{1}{n(n-1)} \sum \left(\frac{x_i - fy_i}{\bar{y}} \right)^2 \quad (n/N < 0.1) \quad (4.67)$$

$$\approx f^2 \frac{N-n}{N} \frac{1}{n-1} \left(\frac{s_x^2}{\bar{x}^2} + \frac{s_y^2}{\bar{y}^2} - 2 \frac{\text{cov}(x, y)}{\bar{x}\bar{y}} \right) \quad (4.68)$$

1) $s_x^2, s_y^2, \text{cov}(x, y), \bar{y}$ を不偏分散の形で計算したときは, $n-1$ のかわりに n とすればよい。

$$\hat{\sigma}_f^2 = \frac{N-n}{N} \frac{1}{n-1} (c_x^2 + c_y^2 - 2rc_x c_y) \quad (4.69)$$

ただし $s_x^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$; $s_y^2 = \frac{1}{n} \sum (y_i - \bar{y})^2$

$\text{cov}(x, y) = \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})$: x と y の試料共分散 (covariance)

$$\begin{aligned} \hat{V}(x') &= \left(\frac{\hat{X}}{f} \right)^2 \hat{\sigma}_f^2 \\ &= Y^2 \hat{\sigma}_f^2 \end{aligned} \quad (4.70)$$

$$\begin{aligned} (\hat{O}_{x'})^2 &= (\hat{O}_f)^2 = \hat{\sigma}_f^2 / \bar{x}^2 = V(x') / (\hat{X})^2 \\ &\doteq \frac{N-n}{N} \frac{1}{n(n-1)} \sum \left(\frac{x_i - f y_i}{\bar{x}} \right)^2 \end{aligned} \quad (4.71)$$

$$\begin{aligned} &\doteq \frac{N-n}{N} \frac{1}{n(n-1)} \sum \left(\frac{x_i}{\bar{x}} - \frac{y_i}{\bar{y}} \right)^2 \\ &\doteq \frac{N-n}{N} \frac{1}{n-1} \left(\frac{s_x^2}{\bar{x}^2} + \frac{s_y^2}{\bar{y}^2} - 2 \frac{\text{cov}(x, y)}{\bar{x}\bar{y}} \right) \end{aligned} \quad (4.72)$$

$$\doteq \frac{N-n}{N} \frac{1}{n-1} (c_x^2 + c_y^2 - 2rc_x c_y) \quad (4.73)$$

$$\doteq \frac{1}{n-1} (c_x^2 + c_y^2 - 2rc_x c_y) \quad (n/N < 0.1) \quad (4.74)$$

ただし $r = \frac{\frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})}{s_x s_y}$

f , X' のある信頼度 α の精度 β をもとめるには, n の大きいときは

$$\beta = u_\alpha \hat{O}_{x'} \quad \text{または} \quad \beta = u_x \hat{\sigma}_f \doteq u_x \sqrt{\hat{V}(f)}$$

において, u_x の値を正規分布表よりもとめればよい. n の小さいときには自由度 $f = n-1$ の t 分布表よりもとめる.

4.9.2 層別サンプリングの比推定 ロットを k 層に層別し, その i 層から n_i コをサンプリングし, x, y , について測定したときの比推定は

1) $s_x^2, s_y^2, \text{Cov}(x, y), r$ を不偏分散の形で計算したときは, $n-1$ のかわりに n とすればよい.

a) 全体の比，総計をとるときは

$$f = \frac{\sum_i^k \sum_j^{n_i} \frac{x_{ij}}{a_i}}{\sum_i^k \sum_j^{n_i} \frac{y_{ij}}{a_i}} \quad \hat{X} = fY \quad (4.75)$$

$$\hat{\sigma}_f^2 \doteq f^2 \sum_i^k \frac{N_i - n_i}{N_i} \frac{1}{n_i - 1} \left(\frac{s_{ix}^2}{\bar{x}_i^2} + \frac{s_{iy}^2}{\bar{y}_i^2} - 2 \frac{\text{COV}_i(x, y)}{\bar{x}_i \bar{y}_i} \right) \quad (4.76)$$

$$\hat{V}(x) \doteq X^2 \sum_i^k \frac{N_i - n_i}{N_i} \frac{1}{n_i - 1} \left(\frac{s_{ix}^2}{\bar{x}_i^2} + \frac{s_{iy}^2}{\bar{y}_i^2} - 2 \frac{\text{COV}_i(x, y)}{\bar{x}_i \bar{y}_i} \right) \quad (4.77)$$

ただし $a_i = n_i/N_i$ ， s_{ix}^2 は i 層の x についての試料分散。他の記号は前と同じ。

b) 層毎に比を推定して，これを全層について加え合せたときは

$$f = \sum_i^k p_i f_i \quad (4.78)$$

$$\text{ただし } p_i = N_i/N, \quad f_i = \frac{\sum_j^{n_i} x_{ij}}{\sum_j^{n_i} y_{ij}}$$

$$\hat{\sigma}_f^2 = \sum_i^k p_i^2 \hat{\sigma}_{f_i}^2 \quad (4.79)$$

$$\hat{\sigma}_{f_i}^2 \doteq f_i^2 \frac{N_i - n_i}{N_i} \frac{1}{n_i - 1} \left(\frac{s_{ix}^2}{\bar{x}_i^2} + \frac{s_{iy}^2}{\bar{y}_i^2} - 2 \frac{\text{COV}_i(x, y)}{\bar{x}_i \bar{y}_i} \right) \quad (4.80)$$

c) 各層毎に比推定により総計を推定し，これを全層について加え合せたときは

$$\hat{X} = \sum_i^k \hat{X}_i = \sum_i^k f_i Y_i = \sum_i^k Y_i \frac{\sum_j^{n_i} x_{ij}}{\sum_j^{n_i} y_{ij}} \quad (4.81)$$

$$\hat{V}(\hat{X}) \doteq \sum_i^k \frac{N_i - n_i}{N_i} \frac{1}{n_i - 1} \hat{X}_i^2 \left(\frac{s_{ix}^2}{\bar{x}_i^2} + \frac{s_{iy}^2}{\bar{y}_i^2} - 2 \frac{\text{COV}_i(x, y)}{\bar{x}_i \bar{y}_i} \right) \quad (4.82)$$

比例サンプルングのときは， $n_i/N_i = a$ ，上式は次のように簡単になる。

a) 全体の比をとったとき

$$f = \frac{\sum_i^k \sum_j^{n_i} x_{ij}}{\sum_i \sum_j y_{ij}} \quad (4.83)$$

$$\hat{\sigma}_f^2 \doteq f^2(1-a) \sum_i^k \frac{1}{n_i-1} \left(\frac{s_{ix}^2}{\bar{x}_i^2} + \frac{s_{iy}^2}{\bar{y}_i^2} - 2 \frac{\text{COV}_i(x, y)}{\bar{x}_i \bar{y}_i} \right) \quad (4.84)$$

c) 層毎に各総計を比推定し、その和をもとめたとき
 \hat{X} は前と全く同じ。

$$\begin{aligned} \hat{V}(\hat{X}) &\doteq (1-a) \sum_i^k \frac{\hat{X}_i^2}{n_i-1} \left(\frac{s_{ix}^2}{\bar{x}_i^2} + \frac{s_{iy}^2}{\bar{y}_i^2} - 2 \frac{\text{COV}_i(x, y)}{\bar{x}_i \bar{y}_i} \right) \\ &= (1-a) \sum_i^k \hat{V}(\hat{X}_i) \end{aligned} \quad (4.85)$$

4.9.3 副次サンプリングの比推定 1次単位 m コをとり、1次単位から n_i コの2次単位をとったときは

$$f = \frac{\sum_i^m \sum_j^{n_i} x_{ij}}{\sum_i^m \sum_j^{n_i} y_{ij}} \quad \hat{X} = fY \quad (4.86)$$

$$\begin{aligned} \sigma_f^2 &\doteq \varphi^2 \left[(1-a_1) \frac{M^2}{m} \left(\frac{\sigma_{ex}^2}{X^2} + \frac{\sigma_{ey}^2}{Y^2} - 2 \frac{\text{COV}_e(x, y)}{XY} \right) \right. \\ &\quad \left. + \frac{1}{a_1} \sum_i (1-a_{2i}) \frac{N_i^2}{n_i} \left(\frac{\sigma_{ix}^2}{X^2} + \frac{\sigma_{iy}^2}{Y^2} - 2 \frac{\text{COV}_i(x, y)}{XY} \right) \right] \end{aligned} \quad (4.87)$$

ただし $a_1 = m/M$, $a_{2i} = n_i/N_i$

$$\sigma_{ex}^2 = \frac{1}{M} \sum_i^M (X_i - \bar{X})^2 \quad \sigma_{ix}^2 = \frac{1}{N_i} \sum_j^{N_i} (x_{ij} - \mu_{xi})^2$$

$$\text{COV}_e(x, y) = \frac{1}{M} \sum_i^M (X_i - \bar{X})(Y_i - \bar{Y}) \quad (4.88)$$

$$\text{COV}_i(x, y) = \frac{1}{N_i} \sum_j^{N_i} (x_{ij} - \mu_{xi})(y_{ij} - \mu_{yi}) \quad (4.89)$$

以上の σ_f^2 は母集団の値についての式で、推定値を用いるときは §5.3.4 にのべた考え方によりもとめた、 $\hat{\sigma}_{ex}^2$, $\hat{\sigma}_{ix}^2$, などを代入すればよい。

簡単な $n_i/N_i = a_2$ 一定で、 $N_i = \bar{N}$, $n_i = \bar{n}$ のときについての式を参考に次にかかげる。 f をもとめる式は一般の場合と同じ。

$$\hat{\sigma}_f^2 \doteq f^2 \left[(1-a_1) \frac{1}{m} \left(\frac{\hat{\sigma}_{bx}^2}{\bar{x}^2} + \frac{\hat{\sigma}_{by}^2}{\bar{y}^2} - 2 \frac{\widehat{Cov}_b(x, y)}{\bar{x}\bar{y}} \right) + (1-a_2) \frac{1}{m\bar{n}} \left(\frac{\hat{\sigma}_{wx}^2}{\bar{x}^2} + \frac{\hat{\sigma}_{wy}^2}{\bar{y}^2} - 2 \frac{\widehat{Cov}_w(x, y)}{\bar{x}\bar{y}} \right) \right] \quad (4.90)$$

ただし $\hat{\sigma}_{bx}^2$, $\hat{\sigma}_{wy}^2$, などは § 5.3.4 にのべると全く同様な方法で推定出来る。共分散は次のようにして推定する。

$$\widehat{Cov}_b(x, y) = \frac{1}{m-1} \sum_i^m (\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y}), \quad (4.91)$$

$$\widehat{Cov}_w(x, y) = \frac{1}{m(\bar{n}-1)} \sum_i^m \sum_j^{\bar{n}} (x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i)$$

とすると

$$\widehat{Cov}_w(x, y) \doteq \widehat{Cov}_w(x, y) \quad (4.92)$$

$$\widehat{Cov}_b(x, y) \doteq \widehat{Cov}_b(x, y) - \frac{1-a_2}{\bar{n}} \widehat{Cov}_w(x, y) \quad (4.93)$$

この値を上式に代入すればよい。この式の形は、 $\hat{\sigma}_{bx}^2$, $\hat{\sigma}_{wy}^2$ などとよく似ている。 \hat{X} の分散は、 $\hat{\sigma}_f^2$ の式の f^2 のかわりに \hat{X}^2 を代入すればよい。

例 1) 毎日生産されるパイライトシンダ－の残留硫黄分を精密測定法で測定することが困難なので、月に 2 日だけは 1 日の平均サンプルにつき精密法と簡易法で測定し、他の日は簡易測定法で測定した。シンダ－生産量は毎日パラツキは少しはあるが、ほとんど 30 トン前後であり、平均値は 30.0 トンである。いまこのパラツキは小さく無視しうる程度とする¹⁾。1 年 24 回の両測定法の比較より、精密法を x 、簡便法を y とすると

$$f = \frac{\sum x}{\sum y} = 1.11 \quad r = 0.93$$

となった。また精密法については、24 回の測定より $\bar{x} = 1.54\%$, $s_x = 0.20$, 簡便法については $\bar{y} = 1.69$, $s_y = 0.35$, $\sum_{i=1}^{365} y_i = 617\%$ となった。1 年間の残留硫黄分の総計は何トンか。ただし平均サンプルはうまくとり、真の平均サンプルをとっているものとする。

1) これはよほどよく管理された工場でないと思えないことも知れぬが、

$$r = 0.93 \quad \frac{c_y}{2c_x} = \frac{0.35/1.69}{2 \times 0.20/1.54} \doteq 0.8 \quad \therefore r > \frac{c_y}{2c_x}$$

したがって比推定の方が有利なことがわかる。比推定を計算するために $Y = \sum y_i \times \bar{w}$ をもとめる。ただし \bar{w} は1日のシンダー平均量。

$$Y = \bar{w} \sum y_i = 30 \times 6.17 = 185.1 \text{ トン}$$

$$\therefore \hat{X} = fY = 1.11 \times 185.1 = 205.5 \text{ トン}$$

精密法による残留硫黄分は 205.5 トンと推定される。

このときの精度は

$$\begin{aligned} \hat{V}(\hat{X}) &= \hat{X}^2 \frac{M-m}{M} \frac{1}{m-1} (c_x^2 + c_y^2 - 2r c_x c_y) \\ &\doteq 205.5^2 \times \frac{1}{24-1} \left[\left(\frac{0.20}{1.54} \right)^2 + \left(\frac{0.35}{1.69} \right)^2 \right. \\ &\quad \left. - 2 \times 0.93 \left(\frac{0.20}{1.54} \right) \left(\frac{0.35}{1.69} \right) \right] \\ &= 205.5^2 \times 0.00043 \end{aligned}$$

$$\hat{\sigma}_{\hat{X}} = 205.5 \times 0.0207 = 4.26$$

99%信頼度では $t_{0.01}(f=23) = 2.81$

$$\therefore \beta = 2.81 \times 4.26 \doteq 12.0$$

したがって、1年間にシンダー中に残った残留硫黄量は 205.5 トンでその99%信頼度の精度は ±12.0 トンである。

例 2) 工場にある材料5,000俵入荷した。各送状を合計すると305.5 トンとなっている。この中2,000俵はA社からで122.4トン、2,000俵はB社からで121.5トン、1,000俵はC社からで61.6トンであった。A、B社の俵から各20俵、C社の俵から10俵ランダムに抜き取り、各俵毎に秤量し、かつ送状の値と比較して次の値を得た。

	N_i	n_i	\bar{x}_i (秤量値)	s_{ix}	\bar{y}_i (送状)	r_{iy}	r	Y
A	2,000	20	60.06kg	0.20kg	60.90	0.15	0.93	122.4 トン
B	2,000	20	59.80	0.31	60.80	0.22	0.89	121.5
C	1,000	10	60.10	0.15	61.20	0.18	0.95	61.6

各社毎、並びに全俵の入荷時における重量はいくらか、またその精度はいくらか。

層別比例サンプルングのデータであるが、何れも $t > \frac{c_y}{2c_x}$ なる故に層別の比推定を行う。

まず各社別に総重量を推定する。

$$\text{A社} \quad f_A = \frac{\bar{x}_A}{\bar{y}_A} = \frac{60.06}{60.90} = 0.9861$$

$$\hat{X}_A = f_A Y_A = 0.986 \times 122.4 = 120.70 \text{ トン}$$

$$\begin{aligned} \hat{V}(\hat{X}_A) &= \hat{X}_A^2 \times \frac{1}{20-1} \left[\left(\frac{0.20}{60.06} \right)^2 + \left(\frac{0.15}{60.90} \right)^2 \right. \\ &\quad \left. - 2 \times 0.93 \left(\frac{0.20}{60.06} \right) \left(\frac{0.15}{60.90} \right) \right] = \hat{X}_A^2 \times 4.73 \times 10^{-8} \end{aligned}$$

$$\hat{\sigma}_{\hat{X}_A} = \hat{X}_A \times 2.17 \times 10^{-4} \doteq 0.03 \text{ トン}$$

$$\text{B社} \quad f_B = \frac{\bar{x}_B}{\bar{y}_B} = \frac{59.80}{60.80} = 0.9836$$

$$\hat{X}_B = f_B Y_B = 0.9836 \times 121.5 = 119.51 \text{ トン}$$

$$\begin{aligned} \hat{V}(\hat{X}_B) &= \hat{X}_B^2 \times \frac{1}{20-1} \left[\left(\frac{0.31}{59.80} \right)^2 + \left(\frac{0.22}{60.80} \right)^2 \right. \\ &\quad \left. - 2 \times 0.89 \left(\frac{0.31}{59.80} \right) \left(\frac{0.22}{60.80} \right) \right] \\ &= \hat{X}_B^2 \times 0.34 \times 10^{-6} \end{aligned}$$

$$\therefore \hat{\sigma}_{\hat{X}_B} = \hat{X}_B \times 0.58 \times 10^{-3} \doteq 0.07 \text{ トン}$$

$$\text{C社} \quad f_C = \frac{\bar{x}_C}{\bar{y}_C} = \frac{60.10}{61.20} = 0.9820$$

$$\hat{X}_C = f_C Y_C = 0.9820 \times 61.6 = 60.49 \text{ トン}$$

$$\begin{aligned} \hat{V}(\hat{X}_C) &= \hat{X}_C^2 \times \frac{1}{9-1} \left[\left(\frac{0.15}{60.10} \right)^2 + \left(\frac{0.18}{61.20} \right)^2 \right. \\ &\quad \left. - 2 \times 0.95 \left(\frac{0.15}{60.10} \right) \left(\frac{0.18}{61.20} \right) \right] = \hat{X}_C^2 \times 0.103 \times 10^{-6} \end{aligned}$$

$$\hat{\sigma}_{\hat{X}_C} = \hat{X}_C \times 0.32 \times 10^{-3} \doteq 0.02 \text{ トン}$$

全体については

$$\hat{X} = \sum_{i=1}^3 \hat{X}_i = 120.70 + 119.51 + 60.49 = 300.70$$

$$\hat{\sigma}_X^2 = \sum_{i=1}^3 \hat{\sigma}_{X_i}^2 \doteq (0.03)^2 + (0.07)^2 + (0.02)^2 \doteq 62 \times 10^{-4}$$

$$\therefore \hat{\sigma}_X = 0.08 \text{ トン}$$

第5章 サンプルングと予備知識

5・1 予備知識と予備調査

§ 4 にのべたサンプルング法は、すべて母集団の性質、たとえばその分布型、母分散 σ^2 、母級間分散 σ^2_b 、母級内分散 σ^2_w などが予備知識としてわかっている場合に、ロットの平均値などを推定する方法、ならびにそのときの精度についてのべた。したがってわれわれは合理的なサンプルングと測定を行ってこれら母集団の平均値を欲する精度、信頼度で経済的に推定するためには、予備知識として、対象となるロットの性質をあらかじめ知っておかなければならない。ロットについての統計的技術的な予備知識があってはじめて合理的なサンプルング法を設計し実施することができるのである。また一度実施したサンプルングの結果も § 1 にのべたように、欲する精度となっているか否かをチェックし、更にこの結果を、次のサンプルング、あるいは新しいサンプルング法の予備知識として利用できるようにしておくことが望ましい。

予備知識として特に大切なのは

- 1) 測定法について
- 2) ロットの歴史と取扱
- 3) 工程の管理状態
- 4) 各種分散の推定

である。前3者は工場におけるサンプルングに独特のものであり、4は標本調査法の統計的技術に負うところが多い。

工場におけるサンプルングは、サンプルング法標準、あるいは仕様書として合理的なものを決定しなければならないし、決定後も必要に応じ改訂しなければならないが、多くの場合一度決定するとこれを改訂することも仲々困難なので、予備知識は技術的、統計的に十分に検討して正しいものを得ておかなければならない。この点標本調査法の場合よりも予備調査が重要である。

予備知識としては、従来からもっている技術的、経済的知識、管理図による知識などは非常に役立つものであるから、サンプルング法立案者は、関

係者、特に現場の人々、経営的にそのサンプルングの目的、精度、経費などを決定する立場にある人々と十分打ち合わせる必要がある。工場には昔からの多くのデータや知識があるので、わざわざ分布型を知るための予備実験などを行う必要のない場合も多いことを銘記すべきである。工場のデータをよく調べて見ると、案外役に立つデータがころがっている。これより得られた知識（たとえば \bar{x} , s^2 ）を用いて、サンプルング法を設計して見るとこれだけの知識で一応十分な場合が多い。

従来のデータをよく調べて見て、得られる予備知識が不十分なときに、はじめて予備調査を行うべきである。また稀に入る材料のサンプルング、ある特別なロットの調査など、市場調査のサンプルング、毎日の作業として行わない特別のサンプルングのときには、予備調査にあまり金をかけることも不経済であるので、標本調査法の場合のように、従来の知識から推定される平均値や分散を、あまり信頼のおけない場合でも、大体の目安として利用することがある。

予備調査を計画する場合でも、従来からもっている予備知識を十分に活用しなければならぬことは勿論である。分布状況を知るためには、多くの場合に、個々のサンプル単位を、あるいは場合によってはよく考えて作成した混合試料を、一つずつ測定しなければならないので、相当の経費と時間と関係者の協力と努力とが必要である。したがって測定の能力などを考えて、少ない予備調査より、最大の情報を得るように計画しなければならぬ。また経済や生産と非常に密接な関係のあるサンプルング法、たとえば仕様書や標準規格などを合理的にきめるには、相当大がかりな予備調査が必要であるから、この予備調査、サンプルング法の研究は関係者間の密接な協力と分担と理解により、統計的な計画と、実施と、解析を伴ったものでなければならない。サンプルング法の研究とは、大部分がこの予備調査である。

予備知識として、用いる測定法について、信頼性、精度、および正確さがどの位であるかを知っておく必要がある。しかしこれについては既に§3でのべたので、本章では、ロットの歴史、管理状態、各種分散の推定法、予備調査のやり方の2～3の例についてのべる。

5・2 ロットの歴史と工程の管理状態

ロットの歴史とは、そのロットがどのような原料をつかって、どのような工程を通り、どのような取り扱いをうけて出て来ているかということと、その工程や取り扱いの管理状態が、すなわち従来のロットがどのような管理状態のものであったかということである。この両者はサンプリング法と密接な関係があり、これに関する予備知識なしでは、合理的、経済的なサンプリング法を決定することは困難である。工程の管理状態というのは広義に言えばロットの歴史に含まれるが、ロットの歴史は現実はそのロットの通って来た過程であり、工程の管理状態とは将来のロットの性質の推定の問題も含み非常に重要なことであるので、§5・1 では特に2項目に分けたのである。

工程管理のときと同様に、サンプリングという立場から考えても、ロットの歴史はハッキリさせておかなければならない。ロットの歴史をハッキリさせるためには、ロットを技術的に考えて必要な程度に層別して流すか、あるいは完全に均一化し、よく混合して流すかいずれかである。ロットの分散を推定するにも、次々と全く異なった歴史のよくわからぬロットが流れて来たのでは、合理的なサンプリング法を決定することは困難である。製品については、自社でロットを流すのに、その製品ロットの歴史がわかるようにすればよいので、たとえば層別して流すとか、ロットと共にカードをつけて流すとか品質管理的に工場を運営することによって比較的容易に行われる。この場合層別のやり方などは管理図法による品質管理のときと同様である。すなわち、合理的なサンプリングを行い、製品の品位を経済的に精度よく推定するためには、**品質管理の確実な実施**ということが有利になる。

原材料の受入れのように、他産業からの受入れの場合においても、その原材料を作成する会社、鉱業所などにおいて、どのようにそのロットを生産しているか、その出荷はどのように行っているか、あるいはどのように運搬されて来たかなどが問題になる。たとえば、炭坑において、採炭をどのように行っているか、層別採炭を行っているか、原炭の処理はロットや層別がはっきりしているか、あるいは洗炭工程の管理状態はどうであるか、

層別選炭を行っているか、精炭ポケットにどのように入れているか、出荷に際してはどのように貨車積みしているか、ある一日に入荷してくるいくつかの貨車について精炭ポケット、積み込みなどが一定しているか、輸送中に粉塊の分離などが起っているか、そのときの天気状況などはどうであったか、等々入荷したロットのいろいろな歴史がはっきりしていれば、これは非常に重要な予備知識になる。したがって受入資材については、もし調査が容易に行えるならば、出荷側とよく打ち合せて調査しておくことよい。このほかロットのパラツキがどんな原因により、何故起るかということを知っておくことも、予備知識としては重要なことである。

しかしロットの歴史において更に大切なことは、そのロットの経て来た**工程が管理状態にあるか否か**ということである。工場におけるサンプルングは、毎日、毎時出るロットをサンプルング法標準によりサンプルングするのである。しかるにサンプルング法標準は毎日出てくるロットの分布型（特にいろいろな母分散）によりきめられるので、工程が管理されていないと、次々と出てくるロットの母分散や母平均が異状な変動を示しており、特に母分散が推定できない状態にあると、あるサンプルング法標準をきめても、その精度が果して期待した通りになるか否かは全く不明である。したがって管理状態にない工程からロットについては、所要精度 β を得るためには、母分散を十分に大きく仮定しておいて、必要以上にサンプルの大きさ n をきめておかなければならないことになる。またサンプルングする間隔も短くして、ロットの品質をよほど注意してチェックしておかなければならないことになる。

管理状態にある工程からのロットであれば、母分散も大体正しく推定できるから、サンプルング法標準も合理的にきめることができ、サンプルの大きさも経済的にきめられ、しかもサンプルングの間隔を管理状態に応じて少なくして行くことが出来るので、サンプルングにとって非常に有利である。

以上のことからわかるように、工場におけるサンプルングでは、ロットの分布型が、次々とくるロットにより、いかに変化するかということが非常に重要で、予備調査を行う場合でも、ある日に入荷した1つのロット、ある日に生産された1つのロットについてのみ詳細な分散などの調査を行

っただけでは不十分な場合が多い。一つあるいは数コのロットによる知識のみによってサンプリング法標準を作成するのは、むしろ危険な場合がある。それよりも、予備調査、予備実験としては、毎日のサンプル数は少なくとも、なるべく長期間にわたって、ロットの分布型、特に分散の管理状態や大きさを調べるの方が大切である。この調査は、相当の時間と経費を要する大きな問題である。たとえば英国で石炭のサンプリングの規格を決定するにも、十年近く統計的な研究実験を行っている。したがって、関係者の協力と分担とが是非必要なのである。予備調査としては、ある期間にわたる調査をできるだけ少ない、経済的な実験から、いかに多くの情報をうるかということに、統計的な考え方や手法をいかに活用して行くかということが問題になる。

しかし実際には、これらの膨大な信頼できる予備調査を行うまで待つわけにも行かぬし、工程が管理状態になるのを待つわけにも行かない場合もあるので、少しは不合理でも、従来のサンプリング法よりは合理的にきめられれば、僅かな予備知識から一応暫定的なサンプリング法標準を作成した方が有利な場合も多いと思われる。この暫定案はよりよい予備調査ができたなら工程の管理状態や、ロットの取り扱い方により、次々と改訂して行くべきである。

ロットの歴史として技術的にも、工程管理上からももう一つ重要なことは、**ロットの取り扱い方**である。たとえば多種類の鉄鉱石が入荷したときに、これを如何に取り扱うかということである。工程管理上から考えても、原材料の平均品質とロット内のバラツキは、製成品質に大きな影響をあたえる場合が多い。しかるにこの原材料の取り扱い方や、あるいは原材料を取り扱う設備、施設が従来あまり考えられていない場合が多い。この取扱いかたが、サンプリングの目的や方法と密接な関係をもつ。この合理的な取り扱い方としては、ロット別、品種別に層別するか、均一化を行うかという2つの方法がある。

従来よく行われている方法は層別法であり、各鉱種毎に別々のところに積んでおくという方法である。原料の品種が非常に多く、しかも各品種について、その品質が管理状態になく、量的にも不安定な場合には、この方法も効果がある。しかし層別した場合には、サンプリングを各層について多数回行わなければならない非常に厄介で、この場合は全層の総平均値を知

るようなサンプルングをしても意味がない。原料が量的にも質的にもある程度管理されてくれば均一化の方が、工程管理上から見ても有利である。この点について最近、石炭、鉱石類において行われている方法は層積法 (layering または bedding) である。層積法とはたとえば、石炭類が入荷したときに、これをただ山積みにするのではなく、貯蔵場にうすく拡げてその上へ上へと順に薄層を積み上げて行く。別種のもが入ればまたこれをその上に同様に薄層として積む。品質標準に合うように割合を考えながら、次々と上へこのように層として積み上げ、使用するときは、これを端から垂直に切りとって、簡単な混合を行いつつ使用して行く。このように積み上げたり、切り取って行く装置も考案されている。これは次々と生産される硫酸や過磷酸石灰などの肥料類、本船積みで多量に入った石炭や鉄鉱石なども薄層にして、つぎつぎに積んで、端から垂直に使用して行けば、石炭や肥料の品質は比較的均一になるので、原料としても、製品としても均一なものを容易に保証できるのみならず、サンプルングも総平均値だけを知ればよいのであるから、非常に容易になり、精度もよくなる。これは丁度集落サンプルングを行っているような効果がでているのである。

ここには一例を示したにすぎないが、工程管理やサンプルング法を考えながら材料などの取り扱い方を常に考えておかなければならない。材料などのこれらの取り扱い、まことにサンプルングの原理のところでのべたように、「材料などの輸送中に行え」すなわち材料が入荷中、ある場所から他へ、あるいは工程中などで毎日通常に行っている輸送の際に行うべきことを付け加えておこう。

5・3 各種分散の推定法

母集団の分布型という、平均値、分散、歪度、尖度ということになるが、サンプルング法の設計や精度の決定に最も問題になるのは分散であるので、ここではその推定法についてのべよう。本章で推定した分散を § 4 でのべた式の母分散のかわりに代入して、サンプルング法を設計することになる。この推定については、§ 5・2 にのべたように、管理状態という思想をできるだけ導入しなればならない。

計数値の場合も考え方はここでのべることと同じであるが、推定式などについては § 6・2 にゆずる。

5・3・1 ランダムサンプリングの分散

ある一つの母集団から n コのサンプルをランダムにとったときの分散の不偏推定値は §2・3・4 にのべたように次式となる。これはすべての場合の基本式である。

$$V_s(x) = \hat{\sigma}^2 = \frac{N-1}{N} \frac{\sum_i^n (x_i - \bar{x})^2}{n-1} \approx \frac{\sum_i^n (x_i - \bar{x})^2}{n-1} = \frac{n}{n-1} s^2 \quad (5.1)$$

この場合の精度は §2・3・4 にのべたように、カイ二乗分布からもとめることができる。 n が大きくなれば推定の精度はよくなる。これよりも便利なのは、次にのべる試料分散の分布を利用することであろう。(5.1)式よりわかるように $\hat{\sigma}^2$ の精度は、 n が大きくなれば s^2 の精度に殆んど等しい。

試料分散の分散及び変動係数は次式によりあらわされる。一般の分布で n が大きいときには近似的に

$$V(s^2) \approx \sigma^4 \frac{\beta_2 - 1}{n} \quad (5.2)$$

$$C(s^2) = \frac{V(s^2)}{s^2} \approx \sqrt{\frac{\beta_2 - 1}{n}} \quad (5.3)$$

$$C(s) = \frac{\sigma_s}{s} \approx \frac{1}{2} \sqrt{\frac{\beta_2 - 1}{n}} \quad (5.4)$$

ただし β_2 : 4 次のモーメント

正規分布の場合には (正規分布では $\beta_2 = 3$ である)

$$V(s^2) = \frac{2\sigma^4}{n} \left(1 - \frac{1}{n}\right) \quad (5.5)$$

$$C(s^2) = \frac{V(s^2)}{s^2} = \sqrt{\frac{2}{n-1}} \quad (5.6)$$

$$C(s) = \frac{\sigma_s}{s} \approx \frac{1}{\sqrt{2(n-1)}} \quad (5.7)$$

(5.4) 及び (5.7) 式でわかるように、試料標準偏差 s の変動係数は一般には β_2 とサンプルの大きさ n により、正規分布では n のみにより決定される。正規分布よりはなれるにしたがって β_2 は 3 より小さくまたは大きくなる。通常われわれが接するような一つ山の連続分布をもつデータでは、 β_2 は 2~8 くらいである。はなれ島のようにデータの飛びはなれた

もののあるときには、 β_2 が8より遙かに大きくなるが、このようなときには層別サンプルングすべきであるから、通常8以下と考えてよい。このときサンプルの大きさと、その精度との関係を上式よりもとめて見ると、表5・1のようになる。

表 5・1 サンプルの大きさの分散の精度

n	$\beta_2=8$ の分布		正規分布	
	$C(s^2)$ %	$C(s)$ %	$C(s^2)$ %	$C(s)$ %
200	19	9	10	5
100	26	13	14	7
70	32	16	17	8.5
50	37	19	20	10
20	—	—	32	16

上表よりわかるように、標準偏差として10~20%くらいの精度で知るには、正規分布から相当はなれているときでも $n=50\sim 100$ 、正規分布に近いときには $n=20\sim 50$ もあれば十分なことがわかる。

また平均値を所要精度 B でおさえるためには、§3にのべたように

$$\beta = t_{\alpha} \frac{\sigma_e}{\sqrt{k}} \quad (3\cdot 12)$$

ただし σ_e : n コの予備調査のデータよりもとめた不偏分散の平方根

t_{α} : 自由度 $f = n - 1$ の危険率 α の t 表の値

k : 本調査に必要なサンプルの大きさ

k は20あるいは30以上になると、あまり変らぬので、この面から見ると80も必要でないことになる。したがってロットの分散が工程によりバラツクことを考えると、普通の分布ならば $n=20\sim 50$ くらいで十分な場合が多いのではあるまいか。特に化学分析などのように時間と経費を要するものでは、その能力や経済性を考えなくてはならない。予備調査において、このようなデータを得たら、データの数が少ない場合でも、一度ヒストグラムをかいて、その分布型を眺めて見ることも役に立つ。

歪度や尖度は通常サンプルング法の設計には、あまり問題にならない場合が多いが、これをもとめるには、サンプルングによるバラツキを考える

と少なくとも 100 以上、できれば 250 位のデータが必要である。

以上は一つの母集団についてのべたのであるが、工程の管理状態ということ considering と、ある一つのロットについて 80 コという多くの予備調査をするよりもむしろ次のような方法をえらんだ方がよい。

たとえば毎日原料が入荷したり、製品がロットとしてでて行くようなときには、それらの各ロットから毎日 $n=3\sim 6$ くらいのサンプルをランダムにとり、それを各々別々に測定する。このようなデータが 100 コくらい、すなわち 20~25 日分くらいデータがたまったらば、各ロットを群にとり、 $n=3\sim 6$ で $\bar{X}-R$ 管理図をかく。この管理図が管理状態にあれば、そのロットの出てくる工程は管理状態にあるのであるから、将来のロットも大体同じ \bar{R} をもっていることが期待できる。そこで $\hat{\sigma}=\bar{R}/d_2$ という関係を用いて

$$\beta = u_{\alpha} \frac{\hat{\sigma}}{\sqrt{k}} = u_{\alpha} \frac{\bar{R}}{d_2 \sqrt{k}} \quad (5.8)$$

u_{α} : 危険率 α の正規分布表の値

より所要精度をうるに必要なサンプルの大きさ k をきめることができる。この場合サンプリングの精度という立場から考えると、 \bar{X} 管理図は管理されていないともよいが、一もし管理されていけばサンプリングすべきロットの数を 1/5 に、あるいはサンプルの大きさを 1/5 くらいに減らしてもよいといわれている。一 R 管理図は管理されていないとはならない。すなわち 25 点中 0 点、35 点中 1 点以内、100 点中 2 点以内が管理限界外にとび出しているようなときは、大体管理状態とみなして、将来のロットについて \bar{R} を適用することができるが、これ以上限界外に点が飛びだしているときには、この \bar{R} をそのまま将来のサンプリング法設計に用いることは危険である。この場合には、限界外の点についてはすべて、できれば管理状態とみなせる場合の限界外の点も、その異状を起した原因が判明し、その原因が工程、サンプリング法あるいは測定法にあり、かつ今後それが再発しないような処置 (action) がとれれば、そのデータを除いて \bar{R} をもとめ、この \bar{R} による管理図が管理状態にあればその \bar{R} を用いることができる。もしその原因が今後完全に予防できぬときには、この \bar{R} からもとめたサンプルの大きさでは、所要精度を常に保証することはできない。このような







場合には、 R の大きい値に対してサンプルの大きさをきめておけば、所要精度は大體得られるのでこれを用いることもある。しかしこれでは通常の場合には精度が良すぎて不経済であるし、また何時、より大きなバラツキ R をもったロットがでるかもわからぬので、不安心である。

もしロット内のバラツキが大きくなったか、小さくなったかを工程などから予知できる場合には、サンプリング法標準を厳重（サンプルの大きさを大きく）、普通、緩和（小さく）などと何段階かにわけてきめておくとよい。また前のようにしていた R 管理図が R が大きく、あるいは R が小さくなったことを示したならば、 \bar{R} ならびに限界線を計算しなおして、サンプルの大きさを計算しなおすべきであろう。

注意 以下にのべる管理図を用いて分散を推定するときには、以上のべたようなことを常に考えなければならない。

1. 分散、標準偏差の簡易推定法 §5・3・1 では、なるべく正しく分散や標準偏差を推定し、管理状態と結びつけて、できるだけ合理的なサン

表 5・2 分布型と分散の推定

データの数 n	$n > 300$		$100 > n > 20$
	分散	標準偏差	標準偏差
	$h^2/12$	$0.29h$	$0.44h$
	$h^2/24$	$0.20h$	$0.30h$
	$h^2/18$	$0.24h$	$0.36h$
	$h^2/8$	$0.35h$	$0.53h$
	$h^2/16$	$0.25h$	$0.38h$
	$h^2/36$	$0.17h$	$0.26h$

プリング法を設計するための予備調査の考え方についてのべたが、場合によっては、予備知識を使って簡単に標準偏差を推定して、サンプリング法を立案することもある。

たとえばロットの大体の分布の形がわかり、理論的、技術的、あるいは経験的にサンプル単位体のとりうる最大値と最小値が推定できるときには、その差 h を利用して、表5・2左側に示すように簡単に分散や標準偏差を推定できる。この場合、最大値と最小値に関するデータや知識が少ないときには、表の右側にあるくらいの値にとっておく方が安全な場合もある。

また4〜5コの大きさのサンプルによる知識しかないときには

$$\hat{\sigma} = R/d_2$$

という関係から σ を大体推定することもできる。しかしこの R は相当にバラツクから、推定した $\hat{\sigma}$ より大きい値を採用しておいた方が安全である。

2. 集合体の場合 単位体の場合には、これまでのべたことが、そのままあてはまる。集合体の場合にもこれまでのべた考え方はそのままあてはまるが、これ以外に§1・5・2にのべたように、サンプル単位の決定という問題を考えなければならない。§7, §8にものべるので、ここではその基本的な考え方だけをのべる。

集合体の場合には、サンプル単位——インクリメントの大きさ、試長——をいかにきめるかということが、予備調査の一つの大きな問題となる。これはたとえばインクリメントの大きさにより、インクリメント間の分散、すなわち母分散がいろいろに変化するからである。これは試長についても全く同様である。

ある大きさのインクリメントで予備調査を行った結果、インクリメント間の標準偏差が大きすぎると、 $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ という関係があるから、一定精度をうるためには、 n を大きくしなければならぬ。一つの均一なロットにおいて、インクリメントの大きさ(w)を変化してインクリメント間の標準偏差(σ_i)との関係をプロットすると、図5・1のような関係となる。インクリメントの大きさが小さい中は、少し w が大きくなっても σ_i が急に小さくなるが、ある程度まで大きくなるとその σ_i を小さくする効果は少なくなってくる。したがって σ_i の大きすぎるときには、 w を大きくし

て、必要なサンプルの大きさ (n) を減らした方がよい場合もあるが、ある程度以上大きくなると、その割に n は減らなくなる。したがって、サンプリング、縮分、測定などの経費を考えると、インクリメントの大きさと、所要経費との関係は図 5・2 のように、経費が最小値となるインクリメントの大きさがある筈である。したがって、 w と n と経費を考えて、所要精度をうるのに最低に近い経費となるような w を予備調査よりもとめなければならない。

一般に w が小さすぎると、 σ_i が大きくなるばかりでなく、カタヨリが入ることがあり、 w が大きすぎてもカタヨリが入ったり、不経済になったりするからこの点も考慮に入れてインクリメントの大きさを決定する。

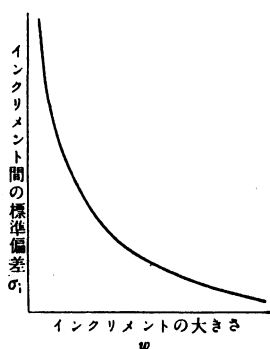


図 5・1 インクリメントの大きさとインクリメント間の標準偏差の関係

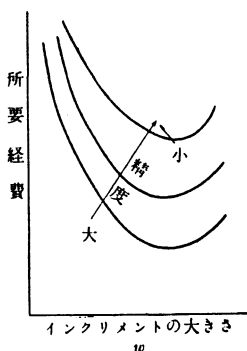


図 5・2 インクリメントの大きさと所要経費の関係

5.3.2 層別サンプリングの分散

層別サンプリングでは、各層の級内分散 σ_i^2 、 σ_w^2 を推定することが必要である。級内分散を推定するには、各層毎に、§5.3.1 でのベタランダムサンプリングの場合をそのまま適用して推定すればよい。この場合各層毎に、できれば各層すべて同じように管理されていることが必要である。一般には (4.6) 式の σ_i^2 は次のように (5.1) 式と同様に推定される。

$$\hat{\sigma}_i^2 = \frac{N_i - 1}{N_i} \cdot \frac{\sum_j^{n_i} (x_{ij} - \bar{x}_i)^2}{n_i - 1}$$

$$\Leftrightarrow \frac{\sum (x_{ij} - \bar{x}_i)^2}{n_i - 1} = \frac{n_i}{n_i - 1} s_i^2 \cdots \cdots (N_i \gg 1) \cdots \cdots \quad (5.9)$$

比例サンプリング ($n_i/N_i = a = \text{一定}$) で予備調査を行ったときには $n/N \leq 0.1$ ならば, (4.19) 式において

$$\sum \sigma_i^2 p_i = \hat{\sigma}_{w'}^2 \Leftrightarrow \frac{1}{n} \sum_i^k n_i \frac{\sum_j^{n_i} (x_{ij} - \bar{x}_i)^2}{n_i - 1} = \frac{1}{n} \sum_i^k \frac{n_i}{n_i - 1} s_i^2 \quad (5.10)$$

さらに $N_i = N/k = \text{一定}$ のときには, $n_i = \bar{n}$ なる故, (4.20) 式では

$$\frac{1}{k} \sum \sigma_i^2 = \hat{\sigma}_{w'}^2 = \frac{1}{k(\bar{n} - 1)} \sum_i^k \sum_j^{\bar{n}} (x_{ij} - \bar{x}_i)^2 \quad (5.11)$$

もし層の数が非常に多いときには, k 層の中 k' 層をランダムにとり, その各層から \bar{n} コずつをランダムにとれば, 副次サンプリングであり, (4.20) 式のときには

$$\hat{\sigma}_{w'}^2 = \frac{1}{k'(\bar{n} - 1)} \sum_i^{k'} \sum_j^{\bar{n}} (x_{ij} - \bar{x}_i)^2 \quad (5.12)$$

予備調査を比例サンプリングで行えば, 上式のようになるが, 各層の σ_i^2 を大体同じ精度で推定するためには, N_i が異なっているときでも, 工業においては一般に n_i を一定に, $n_i = \bar{n}$ としてよい。

平均値の推定値及び精度をもとめるには, 一般には各層の大きさ N_i を知っておく必要がある。しかし層別比例サンプリングで $n/N < 0.1$, N_i も大体一定という場合には N_i を知る必要はない。また貯蔵品などの場合には大体 N_i のわかっている場合が多い。

1. 管理図による簡便法 σ_w は, 管理図法を用いても大体の値を簡単に推定できる。たとえばある製品, 原料などの流れがあるときに, その流れから一定間隔で系統的サンプリングを行ったときに, その特性値たとえば $n=5$ ずつに順に区切って k コの群をつくり, $\bar{X}-R$ 管理図にあらわすと, これは5 コずつをとったもとの流れを各層とする, 層別サンプリングとなるこれは k 層の各層から系統的に $n=5$ のサンプルをとった層

ロットの平均値は管理されていることになる。R管理図が管理されていればこの \bar{R} は、通常のサンプルの大きさ N の1/2(A, Bへ半分ずつ分けて入れたから)、 $N/2$ のインクリメントのときの平均値の精度をあらわす。したがって N インクリメントのときには、 $\bar{R}/\sqrt{2}$ となり、通常のサンプリングあるいはAB試料の平均値の精度は、信頼度99.7%で大体 $D_4\bar{R}/\sqrt{2}$ となる。

この方法はサンプリング法やロットのパラツキをチェックしつつ精度がもとめられ、しかもあまり手数のかからぬ容易に行える方法であるから、できれば平常のサンプリングをこのように切りかえることをおすすめしたい。

注意 この場合後にのべるように、縮分、分析によるパラツキを考えに入れなければならない。

5.3.3 集落サンプリングの分散

集落サンプリングしたときの平均値の精度は級間分散 σ_b^2 または $\sigma_b'^2$ だけにより支配される(4.30~33)式集落サンプリングは§4.4にのべたように、集落を単位と考えたときのランダムサンプリングと同様であるから、次の式により級間分散を推定できる。予備調査として M 集落から m 集落をとれば

$$\hat{\sigma}_b^2 = \frac{1}{m-1} \sum_i^m \left(\sum_j^{N_j} x_{ij} - \bar{X}' \right)^2 \quad (M \gg 1) \quad (5.13)$$

$$\hat{\sigma}_b'^2 = \frac{1}{m-1} \sum_i^m (\bar{x}_i - \bar{\bar{x}})^2 \quad (N_i = \bar{N} \text{ 一定}, M \gg 1) \quad (5.14)$$

$$\text{ただし } \bar{X}' = \frac{1}{m} \sum_i^m \sum_j^{N_i} x_{ij}; \quad \bar{x}_i = \frac{1}{N} \sum_j^{N} x_{ij}$$

$$\bar{\bar{x}} = \frac{1}{m} \sum_i^m \bar{x}_i = \frac{1}{m\bar{N}} \sum_i^m \sum_j^{N} x_{ij}$$

工程を母集団と考えているから、分散の推定のときは母集団修正はいらない。また予備調査において、サンプリングした集落を全部調査するのは経費がかかるから、§5.3.4にのべる副次サンプリング法の推定法を利用して、級間分散をもとめた方が便利な場合も多い。

5.3.4 副次サンプリングの分散

副次サンプリングにおいて、所要精度を最小経費でサンプリングするよ

うに設計するためには、級間分散と級内分散とを推定しなければならない。この結果により、層別サンプルングあるいは集落サンプルングにすべきであるかもきまるので、重要である。予備調査は主として級間分散、級内分散の推定にある。

いま M コの1次単位から m コサンプルングし、 N_i コから成り立つ i 番目の1次単位から n_i コの2次単位をとり測定した場合を考える。一般の場合 (4.41) 式のときには、 σ_e 、 σ_i 、は次の式により推定される。ただし $n_i/N_i = \alpha_2$ 一定とする。

$$\hat{\sigma}_e^2 = \frac{M-1}{M} \left(s_e^2 - \frac{1-\alpha_2}{\alpha_2} s_w^2 \right) \quad (5.15)$$

$$\hat{\sigma}_i^2 = \frac{1}{n_i-1} \sum_j^{n_i} (x_{ij} - \bar{x}_i)^2 \quad (5.16)$$

$$\text{ただし } s_e^2 = \frac{1}{m-1} \sum_i^m (X_i - \bar{X})^2; X_i = \frac{N_i}{n_i} \sum_j^{n_i} x_{ij}; \bar{X} = \frac{1}{m} \sum_i^m X_i$$

$$s_w^2 = \frac{1}{m} \sum_i^m \frac{N_i}{n_i-1} \sum_j^{n_i} (x_{ij} - \bar{x}_i)^2$$

われわれが通常工場で接するように、各1次単位に含まれる2次単位の数 N^* がほぼ一定 \bar{N} で、2次単位の抜取比 $\bar{n}/\bar{N} = \alpha_2$ 一定、 $m/M < 0.1$ 、 $\bar{n}/\bar{N} \leq 0.1$ 、というとき、(4.43) 式には次の位を代入すればよい。

$$\hat{\sigma}_w^2 = \frac{\bar{N}-1}{\bar{N}} \cdot \frac{1}{m(\bar{n}-1)} \sum_i^m \sum_j^{\bar{n}} (x_{ij} - \bar{x}_i)^2 \quad (5.17)$$

$$\approx \frac{1}{m(\bar{n}-1)} \sum_i^m \sum_j^{\bar{n}} (x_{ij} - \bar{x}_i)^2 = s_w^2 \quad (N \gg 1) \quad (5.18)$$

$$\hat{\sigma}_b^2 = \frac{M-1}{M} \left\{ s_b^2 - \frac{\bar{N}-\bar{n}}{\bar{N}} \frac{s_w^2}{\bar{n}} \right\} \quad (5.19)$$

$$= s_b^2 - \frac{s_w^2}{\bar{n}} \quad (M \gg 1, \quad \alpha_2 < 0.1) \quad (5.20)$$

$$\text{ただし } s_w^2 = \frac{1}{m(\bar{n}-1)} \sum_i^m \sum_j^{\bar{n}} (x_{ij} - \bar{x}_i)^2 \quad (5.21)$$

$$s_b^2 = \frac{1}{m-1} \sum_i^m (\bar{x}_i - \bar{\bar{x}})^2 \quad (5.22)$$

注意 毎日入るロットに就ての予備調査の時には、工程を母集団と考

隔を短くした方が都合がよい場合もあろう。工程の総平均値の推定の精度をよくするという立場から考え、 σ_b を小さくし、 σ_w を大きくするようにすれば一定のサンプルの大きさ \bar{n} , m に対しては都合がよい¹⁾から、 \bar{n} コをとる間隔を延ばして σ_w を大きくするようにすればよい。極端な場合には $m=n$, $\therefore \bar{n}=1$ となり、前のべた系統的サンプルングにすればよいことがわかる。すなわち平均値の精度をよくするためには、時間をおいて \bar{n} コずつ $\bar{n} \times m = n$ をとるよりも、同じサンプルの大きさ n ならば、 n を一定間隔でとった方がよいことになる。もし σ_b が小さく、 σ_w が大きいときに、 \bar{X} 管理図は管理状態を示し、平均値の推定の精度は m , \bar{n} によってあまりかわらぬから、一定間隔で \bar{n} コずつとるサンプルング法で差支えない。このようにして推定した平均値は、製品を層積法により山積みしておけば、そのロットの平均品質の保証にもそのまま用いることができる。

2. 級内分散、級間分散の推定の精度 級内分散の推定値 $\hat{\sigma}_w^2$, 級間分散の推定値 $\hat{\sigma}_b^2$ の分散は次式により示される。 N 一定、 \bar{n} 一定、 $m/M \leq 0.1$, $\bar{n}/N \leq 0.1$ なるときは、正規母集団では

$$V(\hat{\sigma}_w^2) = \frac{2\sigma_w^4}{m(\bar{n}-1)} \quad (5.21)$$

$$V(\hat{\sigma}_b^2) = \frac{2}{\bar{n}^2} \left\{ \frac{\sigma_w^4}{m(\bar{n}-1)} + \frac{(\sigma_w^2 + \bar{n}\sigma_b^2)^2}{m-1} \right\} \quad (5.25)$$

註 $m(\bar{n}-1)$, $m-1$ はそれぞれ級内分散、級間分散の自由度

表 5.3 サンプルの大きさ n 一定するとき、1次単位よりとる2次単位の大きさ \bar{n} のきめ方²⁾
($\hat{\sigma}_b^2$ の推定精度を最適にする方法)

$n = m\bar{n}$	σ_b^2/σ_w^2 の 値		
	$\bar{n}=2$	$\bar{n}=3$	$\bar{n}=4$
12	>0.48	0.21~0.48	0.10~0.21
24	>0.50	0.31~0.50	0.19~0.31
36	>0.63	0.34~0.63	0.22~0.34
48	>0.65	0.36~0.65	0.21~0.36
60	>0.66	0.37~0.66	0.25~0.37

1) (4.43) 式参照.

2) ASTM ; Symposium on Bulk Sampling p. 10

この関係式より、 σ_b^2/σ_w^2 の値により、 $m\bar{n}=n$ —一定のときに、 σ_b^2 をもつともよい精度で推定するためには、 m 、 \bar{n} をいかにわりつけたらよいかもとめられる。 σ_w^2 の推定の精度は、一般に σ_b^2 の推定の精度よりもよいから、 σ_b^2 の推定の精度を最良になるようにえらんだ方がよい。通常あらわれる σ_b^2/σ_w^2 の範囲において、便利な表を表 5.3 に示す。一般に工程が管理されていけば $\sigma_b^2/\sigma_w^2 < 1$ であるが、たとえば $\sigma_b^2/\sigma_w^2 = 0.5$ くらいときには、 $m\bar{n} = 36$ ならば、表より $\bar{n} = 3$ とすればよいことがわかる。

5.4 級内分散、級間分散の推定の例 (特に集合体の場合)

すでに §5.3 において、集合体の場合のインクリメントなどサンプル単位間の分散をもとめる方法をのべたが、集合体では多くの場合、化学分析などの面倒な測定を行い、しかも混合試料をつくるという性質に対応して、2, 3 の推定法の工夫をのべてみよう。

この場合になるべく多くのロットを用いて、できるだけ分析回数を少なくしてこれを管理しながら推定する方法と、またやむをえずあるロットについて級間、級内分散の推定を行う場合でも、できるだけ経済的に行う方法をのべて見よう。しかしこれらの例はほんの 1 つの方法である。統計的方法や考え方を縦横に使いこなして、いろいろな方式を考案されたい。

呷入の硫酸、瓶詰の薬品類、貨車や船のいくつかの船倉につんで来た石炭や鉍石類、あるいは工程中を流れる製品、コンベヤー上を輸送中の原材料などのように、ロットが実際に、あるいは頭のなかでいくつかの副ロット (sublot) —— 呷、瓶、貨車、船倉、一定時間、一定量 —— に分けられるときに、これらの副ロットから更にサンプル単位を (集合体の場合にはインクリメント) とる場合で、工場によく用いられる方法である。このときは層別、あるいは副次サンプリングとなり、場合によっては、3 段、4 段などの多段サンプリングも行われる。

このようなときに、いずれのサンプリング法をとり、サンプルをいかに割当てるとかということを決定するためには、副ロット内分散、級内分散 (σ_w^2) と副ロット間分散、級間分散 (σ_b^2) を知る必要がある。すなわち貨車内、呷内分散、錠剤間分散と、貨車間、呷間、瓶間分散がわからなければ、合理的にサンプルを割当てることができなければ、もとめた平均値の

精度を知ることもできない。

ところが工鉱業における製品、半製品、原料などはいずれも工程から生産されるので、その管理状態により分散 σ_w^2 , σ_b^2 が変化する。したがってこれらを常にチェックできるようにサンプルング法、分析、測定法をきめておかなければならない。この点は単位体の場合でも、集合体の場合でも常に心掛けなければならないことである。ところが集合体で、化学分析などするときには、通常1つの混合試料にしてしまうので、これらの分散に関する情報はすべて失われてしまう。この場合単位体の通常の物理的測定のとおりと同様に、サンプル単位を各々測定する方法も考えられるが、これは化学分析のときには、非常な時間と経費がかかるので、管理試験を工夫するか、あるいは混合試料のつくりかたに工夫を行うとよい。

5・4・1 毎日サンプルングを行いつつ、サンプルング法をチェックし、分散の推定を行う方法

1. 管理図を用いる方法 管理図を用いる方法については、既に§5・3にいくつかの方法をのべたが、毎日 M コの副ロットよりなる原料ロットが入って来る場合を考えよう。あるいは毎日2,000トンずつの石炭がコンベヤーを流れる場合でも、これを100トンずつの副ロットと考えれば $M=20$ となる。

M の小さいときには、その各副ロット、各車、各貨車などから、きめられた大きさのインクリメントをランダムに2つずつ（一般には \bar{n} コずつ、ただし \bar{n} は偶数）とる。 M の大きいときあるいは M が変化するときには、副ロット m コをとることになる。試料容器は A_1, A_2 , 2箱準備する。

この2つずつのインクリメントを、たとえば奇数番目のロットについては、同一副ロットの2つのインクリメントを別々に分けて、1つずつランダムに A_1, A_2 の箱に別々に入れる。すべての副ロットについて、1つずつわけて別々に2つの箱に入れる。かくして A_1, A_2 中には各 M コ（一般には $M\bar{n}/2$ コ）のインクリメントよりなる混合試料ができる。この2つの混合試料を各々別々に縮分し、別々に各1回分析する。この場合もし各2回ずつ分析すれば、分析の誤差によるバラツキももとめられる。これをAサンプルング法とし、この2つの混合試料よりの分析値の差を R_A とする。この各混合試料は層の数 M , 各層からのサンプルの大きさ $n_i=1$

なる層別サンプリングのサンプルである。もし M が大きく、あるいは一定でないために、その中から m コをランダムにとったとするとこれは m コの1次単位をとり、各1コを2次単位をとった副次サンプリングとなる。

偶数番目に入ってくるロットについては、同様に B_1, B_2 の試料容器を準備して、各副ロットからの2つのインクリメントを、今度は奇数番目の副ロットについては一緒にして B_1 の試料箱へ、偶数番目の副ロットについては、2つとも B_2 の箱へ入れる。奇数番目の副ロットから集めた混合試料 B_1 と、偶数番目から集めた混合試料 B_2 の2つを、各々別々に縮分し、分析する。このサンプリング法を B とし、この2つの分析値の差を R_B とする。このサンプリング法は、1次単位の大きさ $M/2$ または $m/2$ 、2次単位の大きさ2つづの層別または副次サンプリングである。いまこれを図示すると図5.3のようになる。

2 L コのロットが入った後に、 A, B サンプリング法によるデータを別別にまとめて、 \bar{R}_A, \bar{R}_B 及び $\bar{x}_A, \bar{x}_B, \bar{x}$ をもとめる。

\bar{R}_A, \bar{R}_B より $n=2$ の d_2 の値を用いて、 $\hat{\sigma}_A, \hat{\sigma}_B$ をもとめる。

$$\hat{\sigma}_A = \bar{R}_A/d_2 \quad \hat{\sigma}_B = \bar{R}_B/d_2$$

このデータより、 $n=2$ で $\bar{X}_A - \bar{R}_A$ 管理図と、 $\bar{X}_B - \bar{R}_B$ 管理図をかいて見る。これを一緒に重ねたのが図5.4である。この管理図において、 R 管理図が両者とも、各々管理状態にあるならば、かくしてもとめた $\hat{\sigma}_A, \hat{\sigma}_B$ より次のようにして、 $\hat{\sigma}_w, \hat{\sigma}_b$ を推定できる。

いま縮分や分析の誤差はないものとする。(通常は必ずこれらの誤差があるから、注意1)にのべるように、その補正をしなければならない)。

A サンプリング法は層別サンプリングであるから、一般に各層から \bar{n} コ (\bar{n} は偶数) のサンプルをとったときには、データの数が多いため、 $(\hat{\sigma}_A)^2 \approx \hat{\sigma}_A^2, (\hat{\sigma}_B)^2 \approx \hat{\sigma}_B^2$ と見なすと

$$(\bar{R}_A/d_2)^2 \approx \hat{\sigma}_A^2 = \frac{\hat{\sigma}_w^2}{M(\bar{n}/2)} = \frac{2\hat{\sigma}_w^2}{M\bar{n}} \quad (5.26)$$

したがって

$$\hat{\sigma}_w^2 = \frac{M\bar{n}}{2} \hat{\sigma}_A^2 \quad (5.27)$$

次に B サンプリング法は副次サンプリング法であるから、(4.42) 式より M が一定で全部の副ロット M から2次単位をとるときには、

$$\begin{aligned}
 (\bar{R}_H/d_2)^2 = \hat{\sigma}_B^2 &= \frac{\hat{\sigma}_b^2}{M/2} + \frac{\hat{\sigma}_w^2}{\bar{n}M/2} \\
 &= \frac{2\hat{\sigma}_b^2}{M} + \frac{2\bar{n}\hat{\sigma}_w^2}{M\bar{n}}
 \end{aligned}
 \tag{5.28}$$

注意 この場合工程を母集団と考えているから母集団修正はいらない。

$$\therefore \hat{\sigma}_b^2 = \frac{M}{2}(\hat{\sigma}_B^2 - \hat{\sigma}_A^2) \tag{5.29}$$

副ロット M が大きいとき，あるいは M が一定しないときには， M から， m の副ロット --1次単位--をとるので

$$\hat{\sigma}_w^2 = \hat{\sigma}_A^2 m\bar{n}/2 \tag{5.30}$$

$$\hat{\sigma}_B^2 = \frac{\hat{\sigma}_b^2}{m/2} + \frac{2\hat{\sigma}_w^2}{m\bar{n}} = 2\left(\frac{\hat{\sigma}_b^2}{m} + \frac{\hat{\sigma}_w^2}{m\bar{n}}\right)$$

$$\therefore \hat{\sigma}_b^2 = \frac{m}{2}(\hat{\sigma}_B^2 - \hat{\sigma}_A^2) \tag{5.31}$$

かくしてサンプリングのときに一寸注意をし，毎日縮分回数が1回増加し，分析回数は2回平常通り行うだけで，あまり手間をかけずに副ロット内と副ロット間の分散 $\hat{\sigma}_w^2$ ， $\hat{\sigma}_b^2$ を容易に推定できる。しかも2.にのべる1回のロットだけから推定するのとちがって，長い工程の管理状態を考えての推定値であるし，またこれより管理状態を大体判定できるので遙かに実用的である。

管理図は図5.4のように， R 管理図を2つ， \bar{X} 管理図には \bar{R}_A と \bar{R}_B による2組の限界線を引く。

点は各ロット毎に，たとえば A サンプルングのロットについては， R_A 管理図と \bar{X} 管理図に青色，「 \cdot 」印などでプロットし， B サンプルングのロットは R_B 管理図と \bar{X} 管理図に赤色，「 \times 」印などことなつたマークでプロットする。点はロット番号1, 2, 3, ……順にうつ。 \bar{X} 管理図は各ロット毎に， R 管理図には各々1点おきにプロットすることになる。このとき \bar{X} 管理図の限界線も \bar{x}_A ， \bar{x}_B によりわかりやすいように色別して記入しておけば，各々について管理状態が判定できる。また \bar{x}_A と \bar{x}_B とは意味がちがうが，ロットの平均値の大体の傾向を見ることができ。

層別サンプリングになっているときには R_A 管理図により級内分散，各ロット内のバラツキの管理状態が判定できる。 R_B 管理図は級内分散と級

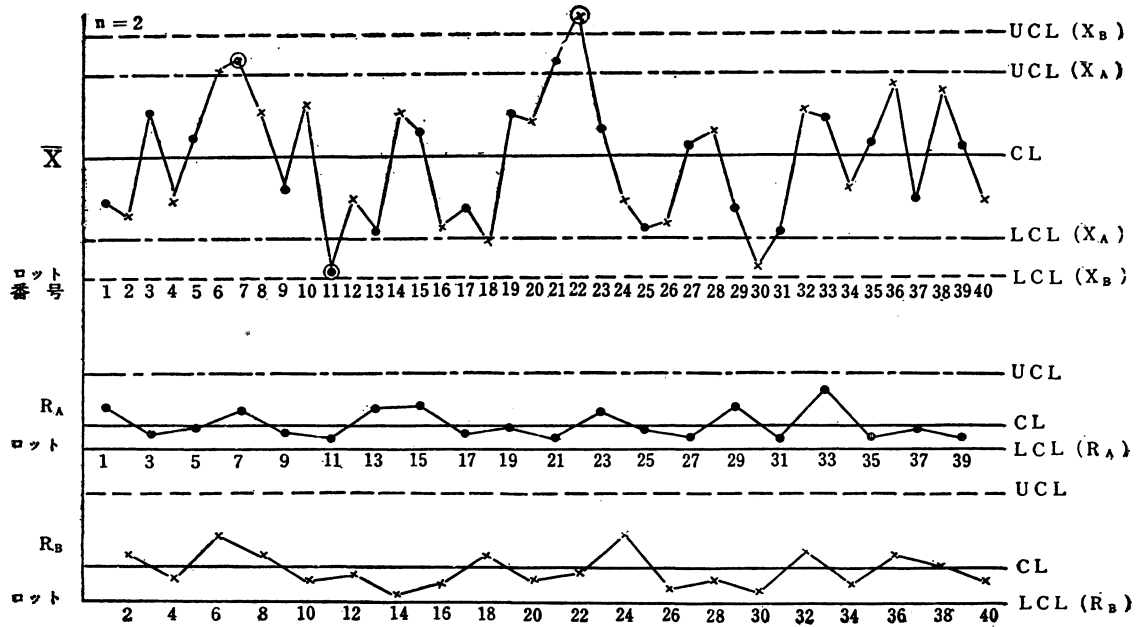


図 5.4 $\bar{X}-R_A-R_B$ 管理図

間分散をふくむが, 副ロットからとるサンプルの大きさ \bar{n} を大きくすれば (5.28) (5.31) 式からわかるように, σ_w の影響を小さくすることができるから, σ_b の管理状態を大体見ることができる. もし管理されていない R 管理図のあるときには, その状態により大きい R_A, R_B を $D_1\bar{R}$ に含むような \bar{R} をえらんで上にのべた式に代入すれば, 大体満足できる精度がえられる.

このとき \bar{X} 管理図がいずれも管理状態にあるときには, 平均値の精度が十分ならば, ロットは十分管理されているのであるから, サンプル数を減少させればよいし, また精度が不十分ならば, さらにサンプル数を増加させるとかサンプリング法を検討しなければならない.

注意 1) 以上の式においては, 縮分誤差, 分析誤差がないものとしているが, 実際にはこの2つの分散も相当に大きくなることが多いから, 予備実験においてこの大きさをもとめておく必要がある. いま縮分による分散を V_R , 分析による分散を V_M とすると, 上にのべた式は次のようになる.

$$\hat{\sigma}_w^2 = \frac{m\bar{n}}{2} (\hat{\sigma}_A^2 - V_R - V_M) \quad (5.32)$$

$$\hat{\sigma}_b^2 = \frac{m}{2} (\hat{\sigma}_B^2 - \hat{\sigma}_A^2) \quad (5.33)$$

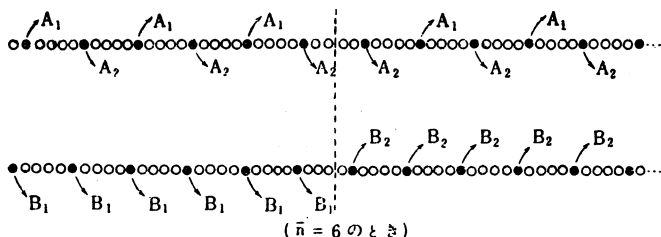
2) 工業におけるサンプリングでは毎日入ってくるロットは工程からの製品であるから, 副ロットは工程からのサンプリングと考えられるので, 分散の推定の際には, $m/M > 0.1$ でも母集団修正はいらない. いかえると変量模型で解析することになる. しかしあるロットの平均値を推定する精度をもとめるときには, そのロットを問題にしているので, 母集団修正が必要である. このことはすべての場合に注意しなければならぬことである. 何を母集団と考えているか注意せよ.

3) R 管理図が管理されていれば, \bar{R}_A, \bar{R}_B , または $A_2\bar{R}_A, A_2\bar{R}_B$ は, それぞれ \bar{x}_A, \bar{x}_B の精度をあらわす.

4) ここにのべた方法には, いろいろな類似法が考えられる. たとえば, 副ロットをランダムにとり, その副ロット内をさらに層別して, 層間, 層内のバラツキを考えて, σ_b^2, σ_s^2 (層間分散), σ_w^2 (層内分散) をもとめるために A, B, C, 3方法で行い, この3者を推定すること, あるいはまず層別してから, 箱, 吠などの2次単位をランダムにとり,

さらにその中から3次単位，サンプル単位体をランダムにとるという方法なども考えられる。

5) 工程中，コンベヤーで輸送中に一定間隔でとったサンプルも，時間により，あるいは日により，次図のようにA，B法を繰り返せば，この例と全く同様に解析できる。



このとき間隔をいくりにするか，副ロットをどこで区分するかは，上述の方法でいろいろやって見て，R管理図が管理されている状態をもとめて，これより $\hat{\sigma}_b$ ， $\hat{\sigma}_w$ を推定して， m ， \bar{n} をもとめる精度になるように決定すればよい。

2. 分散分析を用いる方法¹⁾ 1.にのべた例で， M が一定でないときに副ロットを M コずつ全部サンプルングしたときには， M が日により変化するので管理図による方法は適用できぬ。次のような分散の解析により推定する。

記号は1と全く同じとする。全部の副ロットからサンプル \bar{n} をとったとする。ただし副ロットの大きさ \bar{N} は等しく $\bar{n}/\bar{N} \leq 0.1$ とすれば，副ロットの数が増えるので，重みをつけて考えると，(4.20) (4.43) 式より次の式が導かれる。

$$\hat{\sigma}_w^2 = \frac{\sum_i^L M_i R^2 A_i}{4L} \times \bar{n} \quad (5.34)$$

$$\begin{aligned} \hat{\sigma}_b^2 &= \frac{\sum_j^L M_j R^2 B_j}{4L} - \frac{\hat{\sigma}_w^2}{\bar{n}} \\ &= \frac{1}{4L} \left(\sum_j^L M_j R^2 B_j - \sum_i^L M_i R^2 A_i \right) \end{aligned} \quad (5.35)$$

1) ASTM. Symposium on Bulk Sampling p.8.

またもし副ロットを全部とらず m_i, m_j だけ副ロットをとったときには, 上式において M_i, M_j のかわりに m_i, m_j を入れればよい. m が一定なら, 1. と同じになり, また通常の一元配置法で解析できる.

以上の式において $\hat{\sigma}_w^2, \hat{\sigma}_b^2$ の分散がもとめられる. 簡単な $\bar{n}=2$ のときの式を参考にあげておく.

$$V(\hat{\sigma}_w^2) = \frac{2\sigma_w^4}{L} \quad (5.36)$$

$$V(\hat{\sigma}_b^2) = \frac{2}{L} \left\{ \left(\sigma_b^2 + \frac{\sigma_w^2}{2} \right)^2 + \frac{\sigma_w^4}{4} \right\} \quad (5.37)$$

注意 上式はすべて, 縮分誤差, 分析誤差を考慮していない. 実際には必ずこれを考慮に入れて, 計算しなければならない. たとえば (5.34), (5.35) 式は次のようになる. V_M を縮分誤差を含めた測定誤差の分散の推定値とすると

$$\hat{\sigma}_w^2 = \frac{\sum M_i R^2_{Ai}}{4L} \times \bar{n} - \frac{\bar{n} V_M - \sum M_i}{2L} \quad (5.38)$$

$$\Rightarrow \frac{\sum M_i R^2_{Ai}}{4L} \times \bar{n} - \frac{\bar{n} \bar{M}_i}{2} V_M \quad (5.39)$$

$$\begin{aligned} \hat{\sigma}_b^2 &= \frac{1}{4L} \left(\sum M_j R^2_{Bj} - \sum M_i R^2_{Ai} \right) \\ &\quad - \frac{V_M}{2L} \left(\sum M_j - \frac{\bar{n}}{2} \sum M_i \right) \end{aligned} \quad (5.40)$$

$$\begin{aligned} \Rightarrow &\frac{1}{4L} \left(\sum M_j R^2_{Bj} - \sum M_i R^2_{Ai} \right) \\ &\quad - \frac{1}{2} \left(\bar{M}_j - \frac{\bar{n}}{2} \bar{M}_i \right) V_M \end{aligned} \quad (5.41)$$

すなわち級内分散を相当に大きく推定してしまう. 級間分散の方は大してかわりない. この式は V_M が混合試料の量により殆んどかわらず一定とみなせる場合である.

もし混合試料の量により V_M が相当かわるときには, V_M のかわりにそのサンプル量に応じた V_{M_i} を考えて

$$\frac{1}{L} \sum_i^L V_{M_i} M_i, \quad \frac{1}{L} \sum_j^L V_{M_j} M_j$$

を代入すればよい。

また M_i, M_j のかわりに m_i, m_j を代入しても全く同じになる。

例 ある吠入の工業薬品が、多数の吠の貨車積でしばしば入荷する。いまその入荷毎に30吠ずつ抜きとり、それから各2インクリメントをとり、A、B サンプルング法により各2つずつの混合試料をとった。その成分を60ロットにつき測定した結果は表5・4のようになった。この吠間、及びインクリメントを単位とする吠内分散をもとめよ。ただし縮分を含めた測定分散は0.01である。

表 5・4 薬品の成分

ロット番号	サンプルした吠の数 m_i	測定値		差 R_A %	$m_i R^2_A$
		A_1 %	A_2 %		
1	30	61.2	63.8	0.4	4.8
3	"	64.5	64.2	0.3	2.7
5	"	64.5	64.0	0.5	7.5
⋮					
57	"	63.8	63.7	0.1	0.3
59	"	63.7	64.0	0.3	2.1
$L=30$	900				118.6
		B_1	B_2	R_B	$m_j R^2_B$
2	30	64.0	63.5	0.5	7.5
4	"	64.4	64.7	0.3	2.7
6	"	64.0	64.4	0.4	4.8
⋮					
58	"	63.4	63.5	0.1	0.3
60	"	64.0	63.4	0.6	10.8
$L=30$	900				183.2

(5.39), (5.41) 式を用いて

$$\begin{aligned}\hat{\sigma}_w^2 &= \frac{\bar{n}}{4L} \sum m_i R^2_{Ai} - \frac{\bar{n}V_M}{2L} \sum m_i \\ &= \frac{2 \times 118.6}{4 \times 30} - \frac{2 \times 0.01 \times 900}{2 \times 30} = 1.677 \\ \hat{\sigma}_b^2 &= \frac{1}{4L} \left(\sum m_j R^2_{Bj} - \sum m_i R^2_{Ai} \right) - 0 \\ &= \frac{1}{4 \times 30} \left(183.2 - 118.6 \right) = 0.538\end{aligned}$$

即ち級内分散が非常に大きい。この場合もし測定誤差考えないと

$$\hat{\sigma}_w^2 = \frac{2 \times 118.6}{4 \times 30} = 1.977 \quad \hat{\sigma}_b^2 = 0.538$$

となり、級内分散は相当大きくなり、一定精度をうるためのサンプルの大きさを非常に多くとりすぎるようになる。

註 この例では m_j が一定でインクリメントの数各 2 なるゆえ、 $\hat{\sigma}_b$ には測定誤差が影響してこないが、 m_j が変化するか、 $\bar{n} > 2$ となると少し影響をあたえるようになる。

5.4.2 1 ロットについて $\hat{\sigma}_b$, $\hat{\sigma}_w$ をもとめる方法

§ 5.4.1 にのべた方法は、相当数のロットがつづけて入る場合には用いることができ、縮分や分析に殆んど特別の費用がかからないのでいろいろ有利な点をもっているから、できるだけこのような方法を用いるべきである。しかし入るロットの数が非常に少ない場合、あるいは、はじめに大体 $\hat{\sigma}_b$, $\hat{\sigma}_w$ を推定しておきたいときに、次のような方法を用いるとよい。しかし 1 つのロットについて、いくら精密に $\hat{\sigma}_b$, $\hat{\sigma}_w$ をもとめても、それはあくまでもそのロットに関する知識であって、次々と入るロットについては、これをそのまま適用できない。したがって本章でのべる方法でもとめた結果より、サンプリング法をきめたならば、あと § 5.4.1, § 5.3 などでべた方法により、管理状態をチェックして行くことが望ましい。

この場合には主として分散分析法が用いられるが、サンプリングの予備調査のときには、その構造が変量模型となることが多いから、注意を要する。これについては附録を参照されたい。

例1) M コの副ロットからなるロットがあるときに、この中から m コの副ロットをランダムに抜き取り、この副ロットからインクリメントを \bar{n} コずつとり、その各々を測定し、データをもとめ、一元配置法で分散分析を行う。

副ロット	1	2	3	……	m
インクリメント ごとのデータ	x_{11}	x_{21}	x_{31}	……	x_{m1}
	x_{12}	x_{22}	\vdots		\vdots
\bar{n}	\vdots	\vdots			\vdots
	$x_{1\bar{n}}$	\vdots			$x_{m\bar{n}}$
平均	\bar{x}_1	\bar{x}_2	\bar{x}_3	……	\bar{x}_m

表 5・5 分散分析表

二乗和	自由度	分散	$E(v^2)$
副ロット間	$m-1$	\hat{V}_b	$\sigma_w^2 + \sigma_M^2 + \bar{n}\sigma_b^2$
残り	$m(\bar{n}-1)$	\hat{V}_K	$\sigma_w^2 + \sigma_M^2$
計	$m\bar{n}-1$		

表 5・5 からわかるように、残りの不偏分散 \hat{V}_K から、測定誤差を引けば、 $\hat{\sigma}_w$ がもとめられ、 $\hat{\sigma}_b$ は $(\hat{V}_b - \hat{V}_K)/\bar{n}$ よりもとめられる。

ある袋入の材料が入ったときに、その中10袋をランダムにえらび、それからランダムに各2コのインクリメントをとり分析した結果次のようになった。分析誤差(縮分誤差を含めて) $\hat{\sigma}_M^2 = 0.04$ ということを予備実験からもとめてある。

袋	1	2	3	4	5	6	7	8	9	10	
インクリ メント	1	6.37	7.50	5.75	7.09	7.31	5.85	6.46	4.62	6.41	5.44
	2	6.46	6.36	6.05	7.38	6.78	5.75	5.44	6.16	6.26	6.46
平均		6.42	6.93	5.90	7.24	7.04	5.80	5.95	5.39	6.34	5.95

表 5.6 分散分析表

	由 自 度	不 偏 分 散
袋 間	9	0.74
袋 内	10	0.31

$$\therefore \hat{\sigma}_{w}^2 = 0.31 - 0.04 = 0.27$$

$$\hat{\sigma}_{b}^2 = \frac{0.74 - 0.31}{2} = 0.215$$

この場合, もし測定誤差が不明のときは, 2元配置の繰り返しのある場合の変量模型の型 (§ A1.4d) で解析すればよい. すなわち各インクリメントを十分粉碎して縮分するときに, 最初の縮分でえた2つの部分を各々別々に縮分して, 2つの分析試料をえてこれを各々分析してえた2つのデータを繰り返しにとる. ただし各インクリメントについての縮分と分析の順序はランダムに行う. このデータは次表のようになる.

袋	1	2	3	m
インクリメント 1	— } r	—	—	x_{ijk}	—
↓ n	—	—	—	—	—

表 5.7 分散分析表

要 因	二 乗 和	自 由 度	不 偏 分 散 の 期 待 値
袋	$\bar{n}r \sum (\bar{x}_i - \bar{x})^2$	$m - 1$	$\sigma_M^2 + r\sigma_w^2 + \bar{n}r\sigma_b^2$
インクリメント	$r \sum \sum (\bar{x}_{ij} - \bar{x}_i)^2$	$m(\bar{n} - 1)$	$\sigma_M^2 + r\sigma_w^2$
繰り返し	$\sum \sum \sum (\bar{x}_{ijk} - \bar{x}_{ij})^2$	$m\bar{n}(r - 1)$	σ_M^2

表 5.7 にしたがって解析を行えば, $\hat{\sigma}_M^2$, $\hat{\sigma}_w^2$, $\hat{\sigma}_b^2$ が同時に推定できる.

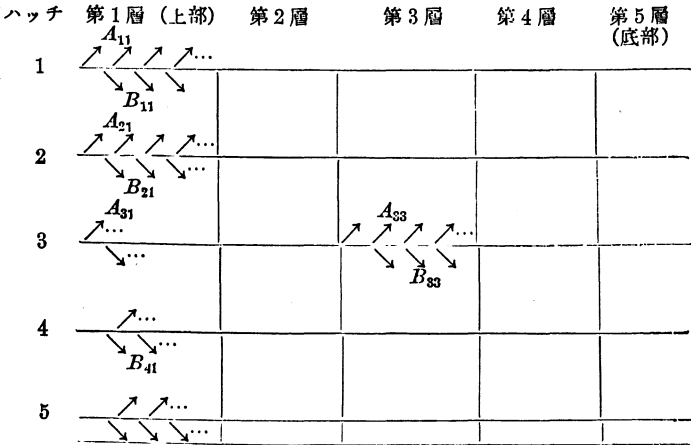
例 2) 船の各ハッチより荷下しの中のコンベヤーなどからのサンプリングの場合の一例. コンベヤー上を流れる製品などのバラツキの推定法

については、既にいろいろの方法をのべたが、ここではハッチ間、ハッチ内のバラツキと、現在実施しているサンプルングの精度を同時にともめる方法をのべよう。目的はハッチ間のバラツキを見たいのであるから、荷上げ、コンベヤーの流し方、あるいはサンプルング場所を工夫して、各ハッチ毎に層別してサンプルングしなければならない。この場合すべてのハッチについてサンプルングすれば、層別サンプルングであり、分散分析するときには、その船のその積荷のハッチと考えれば母数模型となる。またハッチをつぎつぎと入る船の1つのサンプルを考えるか、あるいは多くのハッチの中から、ランダムにいくつかとった場合には、副次サンプルングであり、ハッチ間については変量模型となる。ハッチ内のバラツキを見るために、荷下した順にコンベヤーなどでサンプルングするのであるから、その目的により、荷上係と荷上の順序などをよく打ち合せておかなければならない。たとえばハッチ内を各ハッチについて順にいくつかの層、たとえば5層に層別すれば、荷上中をその荷上量に応じて5層にわけてサンプルングすることになり、この層間については場合により変量模型あるいは母数模型となる。

たとえば従来ある一定の大きさのインクリメントで2,000トンにつき100とっていたとすると、次のようにサンプルングを行う。いま5つのハッチに各2,000トンずつ入って来たとする。これを5層に、すなわち400トンずつに層を考えると、各層につき20ずつのインクリメントを一定間隔でとることになる。このとき試料容器を各副ロットの各層につきAB2つずつ、この場合には $5 \times 5 \times 2 = 50$ 準備する。各層についてA、B箱にインクリメントを交互に、あるいはジックザグに、AB BA AB BA…と別々に入れる。かくして各層毎にA、B2種類の10インクリメントずつの混合試料をつくり、これを各別々に縮分し各1個ずつ分析する。これは次図のようになる。

A_{ij} , B_{ij} を各々繰り返しと見て、この場合は母数模型と考えて分散分析すればハッチ間 σ_b^2 , ハッチ内 σ_w^2 , 各混合試料間 σ_i^2 , 交互作用 σ_I^2 (各ハッチによるハッチ内の積荷の状況の差) などがもとめられる。 σ_i^2 をもとめるには、測定誤差 σ_M^2 を差引かなければならない。

各層毎に x_{11} , x_{12} , x_{13} , x_{14} , x_{15} などとプロットして見ると、各ハッ



このデータを表にまとめると次のようになる。

表 5.8 サンプルングのデータ

ハッチ番号 層番号	1	2	3	4	5	平均	平均
1	A_{11} B_{11}) x_{11}	A_{21} B_{21}) x_{21}	—	—	A_{51} B_{51}) x_{51}	$\bar{A}_{.1}$ $\bar{B}_{.1}$	$\bar{x}_{.1}$
2	A_{12} B_{12}) x_{12}	—	—	—	—	$\bar{A}_{.2}$ $\bar{B}_{.2}$	$\bar{x}_{.2}$
3		—	—	—	—	$\bar{A}_{.3}$ $\bar{B}_{.3}$	$\bar{x}_{.3}$
4	—	—	—	—	—	$\bar{A}_{.4}$ $\bar{B}_{.4}$	$\bar{x}_{.4}$
5	A_{15} B_{15}) x_{15}	—	—	—	A_{55} B_{55}) x_{55}	$\bar{A}_{.5}$ $\bar{B}_{.5}$	$\bar{x}_{.5}$
平均	$\bar{A}_{.1}$ $\bar{B}_{.1}$	$\bar{A}_{.2}$ $\bar{B}_{.2}$	$\bar{A}_{.3}$ $\bar{B}_{.3}$	$\bar{A}_{.4}$ $\bar{B}_{.4}$	$\bar{A}_{.5}$ $\bar{B}_{.5}$	\bar{A} \bar{B}	
平均	$\bar{x}_{.1}$	$\bar{x}_{.2}$	$\bar{x}_{.3}$	$\bar{x}_{.4}$	$\bar{x}_{.5}$		$\bar{\bar{x}}$

ただし x_{ij} は A_{ij}, B_{ij} の平均値
 $\bar{A}_{.j}, \bar{B}_{.j}, \bar{x}_{.j}$ は A_{ij}, B_{ij}, x_{ij} の各行の平均値
 $\bar{A}_{.i}, \bar{B}_{.i}, \bar{x}_{.i}$ は各列の平均値

表 5・9 分散分析表

要 因	自 由 度	不偏分散の期望値
ハッチ間 (A)	$a-1$	$\sigma_i^2 + \sigma_M^2 + b r \sigma_b^2$
ハッチ内 (B)	$b-1$	$\sigma_i^2 + \sigma_M^2 + a r \sigma_w^2$
交互作用 (A × B)	$(a-1)(b-1)$	$\sigma_i^2 + \sigma_M^2 + r \sigma_I^2$
混合試料間	$ab(r-1)$	$\sigma_i^2 + \sigma_M^2$
この場合	$a=5 \quad b=5 \quad r=2$	

チ内の積荷の傾向がわかり、 $\bar{x}_1, \dots, \bar{x}_j, \dots$ などをプロットして全体としての傾向をつかむこともできる。

また \bar{A}_i と \bar{B}_i より各 2,000 トン毎の \bar{x}_i の精度をもとめることができる。平常は 100 インクリメントを混合試料とし縮分し分析するのであるから、このときの一つデータの精度 σ^2 は次のようになる。

$$\sigma^2 = \sigma_s^2 + \sigma_R^2 + \sigma_M^2$$

σ_s : 100 インクリメントでサンプルングすることによるバラツキ

σ_R, σ_M : それぞれ縮分及び分析のバラツキ

しかるに \bar{A}_i と \bar{B}_i のバラツキ $V(AB)$ は、各 50 のインクリメントよりなり、これを 5 回に縮分分析しているので、このときはサンプルの量が少ないから縮分の誤差を σ_R' とすれば、

$$V(AB) = 2\sigma_s^2 + \frac{\sigma_R'^2}{5} + \frac{\sigma_M^2}{5}$$

となる。 $\sigma_R, \sigma_R', \sigma_M$ がわかっていれば、 $V(AB)$ より $\hat{\sigma}_s$ をもとめこれを上式に代入すれば $\hat{\sigma}_s$ がえられる (100 インクリメントと 50 インクリメントでは $\sigma_R^2 \approx \sigma_R'^2$ という場合が多い)。

注意 1) 100 インクリメントの混合試料にすると、さらに層別サンプルングの効果を發揮して、50 インクリメントのときの $\hat{\sigma}_s$ の $\frac{1}{2}$ 以下になり、実際の σ^2 は上式でもとめたものより小さくなる。

2) この場合インクリメントの大きさをいろいろにかえたときの条件を見たければ、各層毎にインクリメントの大きさをかえて、たとえば次のようなラテン方格を組んで、同様にサンプルングすればよい。

表 5・10 ラテン方格

ハッチ番号 層番号	1	2	3	4	5
1	w_2	w_3	w_4	w_5	w_1
2	w_3	w_4	w_5	w_1	w_2
3	w_5	w_1	w_2	w_3	w_4
4	w_4	w_5	w_1	w_2	w_3
5	w_1	w_2	w_3	w_4	w_5

w_i は各異なったインクリメントの大きさを示し、 w_i 間のバラツキはインクリメントによるカタヨリをあらわす。ただしこの場合は現行法の精度などはもとめられない。

例 3) インクリメントあるいは混合試料を 2 つにわけて縮分、分析回数を減小させる方法。

例 1), 2) にのべたような各ロットについて詳しい情報を得ようという方法は、非常に多くのサンプルを縮分、分析しなければならないので、いくつかのロットについてこのようなことを行うことは、不経済、あるいは不可能な場合が多い。したがってこれを減小させる工夫が必要である。

たとえば次のような方法が考えられる。サンプルのとり方は前にのべた方法と全く同様である。いま m 輻の各炭車から n コずつインクリメントをランダムにとつたとする。この nm の分析をするのは大変である。そこでこの各インクリメントを十分に粉碎し、混合しこれを各 2 分する。この 2 分はリッフルサンプラーを正しく用いた程度で十分である。この 2 分したサンプルの一方は各炭車毎に n コあつめて、1 つの試料箱に入れる。これが m コできる。他方の 1/2 は、各炭車より 1 つずつあつめて 1 つの試料箱に入れる。これが n コできる。このとり方を図示すると図 5・5 のようになる。この $n+m$ コを各々別々に縮分し、分析して、そのデータの m コ、及び n コより各々不偏分散をもとめると、それは次のような構造になっている。

各炭車毎の m コのデータからの不偏分散 V_m

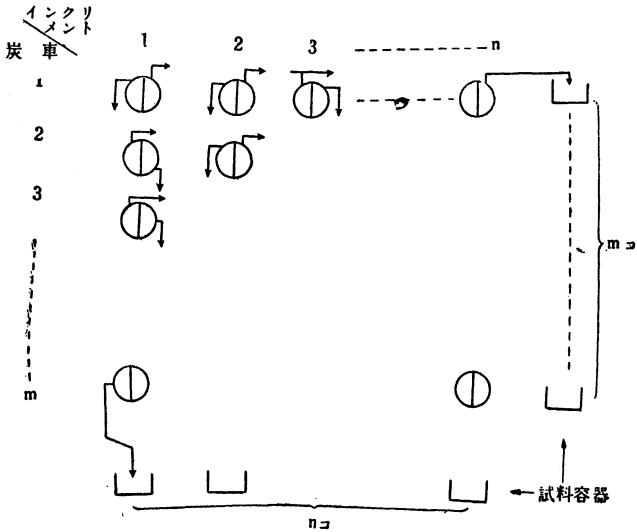


図 5.5 インクリメントの集め方

$$V_m = \frac{\sum_i^m (x_i - \bar{x})^2}{m-1} = \sigma_b^2 + \frac{\sigma_w^2}{n} + \frac{\sigma_1^2}{n} + \sigma_2^2 \quad (5.42)$$

各炭車からあつめた混合試料 n コのデータからの不偏分散 V_n^2

$$V_n = \frac{\sum_j^n (x_j - \bar{x})^2}{n-1} = \frac{\sigma_w^2}{m} + \frac{\sigma_1^2}{m} + \sigma_2^2 \quad (5.43)$$

ただし σ_1^2 : インクリメントを粉碎混合して2分するための誤差分散

σ_2^2 : 混合試料を縮分, 分析するための誤差分散

V_m, V_n をもとめ, σ_1, σ_2 に各々既知の知識よりその不偏推定値をいれれば, $\hat{\sigma}_b, \hat{\sigma}_w$ がもとめられる。

この方法によれば $m \times n$ コの縮分, 分析が, $m+n$ コとなり, 相当時間と経費の節約ができる。

注意 この方法は水分のように, 粉碎混合することにより揮発性成分に変化を来すようなものには, 適用困難である。

第6章 単位体のサンプリング

§ 5までは主として、一般論についてのべてきたが、本章以降において少し具体的に論じて見よう。

単位体とは1コ1コが殆んど同じような材料や製品から成り立っている場合で、1コ毎にサンプリングや測定のできる場合である。この単位体を、あるいは何か一緒にサンプリングすることもある。この場合には、後にのべる集合体の場合より、サンプリングを行う場合もランダムにとりやすく、各単位体について測定を行える場合が多く、その結果より、サンプリングの精度や級間分散、級内分散などをすぐにもとめられるので、いろいろな点において集合体のサンプリングより簡単であり、予備調査を特別に行う必要のない場合が多い。

単位体の場合には、その特性の見かたにより、計量値を推定する場合と、計数値たとえば不良品の全数を推定する場合とあるのでこの両者にわけて考えよう。計数値の場合については、単位体の場合のみになるので、本章にまとめる。

6・1 計量値のサンプリング

単位体の計量値についてのサンプリング法並びにその理論式は、各々の単位体について測定する場合が多いので、殆んど§ 4、§ 5にのべたことがそのままあてはまる。したがって本節では主として応用例について話をすすめる。

比推定については、§ 4・9に例をのべたので省略する。

6・1・1 ランダムサンプリング (計量値の場合)

単位体が多くあるときに、その中から n コをランダムサンプリングすることは、集合体の場合にくらべて容易である。しかし知ろうとするロット全体からランダムにとらなければならない。たとえば箱の上方にあるもの

1) 厳密な意味において同じものではないので、ここでは殆んど同じようなという表現を用いた。

のみとったり、倉庫の入口にあるものだけとったり、あるいは工程を流れる製品を1時間目毎に流れて来たものだけからランダムにとっても、一般には箱全体、倉庫全体、1時間にできた全製品を代表するサンプルにはならない。もしこれらの一部からとったサンプルについての測定値であれば、そのデータをあとでいくら統計的に正しい式を用いて推定したり、精度をもとめたりしても、あまり意味がない。したがって前にものべたように、どのようなロットの品質特性を知りたいのか、目的をよく考えて、ロットの定義をはっきりきめておかねばならない。

注意 しかし、製品がランダムに運搬され、倉庫などにつままれたときには、その表面の製品の品質特性をランダムとみなしうることもある。しかし時間とともに、その環境により変化するような品質特性の場合には、倉庫に入れてから時間がたつとランダムとはみなせなくなる。

ランダムサンプリングのときの平均値、総計の推定式は§2・3・4にのべ

た通り、
$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$
 その精度 β は、信頼度 $(1-\alpha)$ で

$$\pm\beta = \pm t_{\alpha} \frac{s}{\sqrt{n-1}} \quad (6.1)$$

ただし s : 試料標準偏差 $s = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$

t_{α} : 自由度 $f = n-1$ の t 表の危険率 α の値

もし次々と来るロットが管理状態にあり、過去の多くのデータから \bar{x} , \bar{R} などが既知のときには、今後の n コのサンプルの平均値の精度は

$$\pm\beta = \pm u_{\alpha} \frac{\hat{\sigma}}{\sqrt{n}} = \pm u_{\alpha} \frac{1}{\sqrt{n}} \frac{\bar{s}}{c_2} = \pm u_{\alpha} \frac{1}{\sqrt{n}} \frac{\bar{R}}{d_2} \quad (6.2)$$

ただし c_2 , d_2 はサンプルの大きさ n によりきまる係数 (附表2参照)

ところがここで推定した $\hat{\sigma}$ は各サンプル間のバラツキばかりでなく、測定誤差を含んでいる。われわれはこの両者を一緒にして精度を考えているのであるから、一応上式のまま用いてもよいわけである。しかしさらに精度を小さくするためには、サンプル間のバラツキと、測定法によるバラツキとはっきり区別して考えないと、無駄に多くのサンプルをとってしまうことがある。そこで測定法の精度も考えながら、全体の平均値の精度

をよくする方法を考えて見よう。

単位体を1コとり、それを1回測定したときには、その測定値の構造は次のようになる。

$$x_i' = \mu + x_i + z_i \quad (6.3)$$

ただし μ : そのロットの母平均

x_i : 各サンプルの母平均からのカタヨリ

z_i : 測定によるバラツキ (測定にカタヨリはないものとする)

いま σ_s^2 : 母集団の各単位体の分散=単位体のサンプリングによる分散

σ_M^2 : 測定によるバラツキ

とすると、 x_i , z_i は平均値0, 分散は各々 σ_s^2 , σ_M^2 という分布をしていると考えられる。いま測定法の値 x_i' と、サンプルの真の値 $\mu + x_i$ とに相関がなければ、各データのバラツキ $\sigma_{x'}$ は次のようになる。

$$\sigma_{x'}^2 = \sigma_s^2 + \sigma_M^2 \quad (6.4)$$

1. n コのサンプル単位体を各々1回ずつ測定する場合

サンプルの平均値の分散 $V(\bar{x})$ は

$$V(\bar{x}) = \frac{\sigma_{x'}^2}{n} = \frac{1}{n} (\sigma_s^2 + \sigma_M^2) \quad (6.5)$$

2. サンプルを n コとり、各サンプルを各 k 回ずつ繰り返し測定する場合

$$V(\bar{x}) = \frac{1}{n} \left(\sigma_s^2 + \frac{\sigma_M^2}{k} \right) \quad (6.6)$$

これは測定の精度が悪いときに、たとえば化学分析などで、分析精度の悪いときに用いると、この n コのサンプルよりのデータ $n \times k$ コの平均値の精度は1) の場合よりもよくなる。したがって σ_s に比し、 σ_M の大きいときには、サンプルを多くとるよりも、測定を繰り返し行った方がよい。

3. サンプル n コを全部一緒にして測定する場合

たとえば、a) 鉄板の重量の平均値を知りたいときに、1枚ずつはからずに n 枚のサンプルを一緒にして測定する場合、b) あるいは錠剤などを n コランダムサンプリングし、これを一緒にして分析してその平均値を知ろうとする場合がこれに相当する。このときにはえられるデータは次のような構造もっている。

a) 全量 X , 平均値 \bar{x} は, 測定を 1 回行う場合には

$$X = n\mu + \sum_i^n x_i + z_i \quad (6.7)$$

$$\bar{x} = \frac{X}{n} = \mu + \frac{\sum_i^n x_i}{n} + \frac{z_i}{n} \quad (6.8)$$

したがって平均値 \bar{x} の分散は, 多量のものを測定するときの精度を $\sigma_{M'}$ とすると

$$V_1(\bar{x}) = \frac{n\sigma_s^2}{n^2} + \frac{\sigma^2_{z'}}{n^2} = \frac{\sigma_s^2}{n} + \frac{\sigma^2_{M'}}{n^2} = \frac{1}{n} \left(\sigma_s^2 + \frac{\sigma^2_{M'}}{n} \right) \quad (6.9)$$

(6.5) 式と (6.9) 式を比較していずれが精度がよいか比較して見よう。

$\sigma^2_M = \sigma^2_{M'}$, すなわち測定の精度が, その量によりかわらないときには,

$$V(\bar{x}) = \frac{\sigma_s^2 + \sigma^2_M}{n} > \frac{1}{n} \left(\sigma_s^2 + \frac{\sigma^2_M}{n} \right) = \frac{1}{n} \left(\sigma_s^2 + \frac{\sigma^2_{M'}}{n} \right) = V_1(\bar{x})$$

となり, 全部一緒にして測定した方が, 平均値の精度はよい。

$n\sigma^2_M = \sigma^2_{M'}$, すなわち精度が分散として, その測定の絶対量に比例して大きく, 悪くなるようなときには

$$V(\bar{x}) = \frac{1}{n} (\sigma_s^2 + \sigma^2_M) = \frac{1}{n} \left(\sigma_s^2 + \frac{\sigma^2_{M'}}{n} \right) = V_1(\bar{x})$$

となり, 精度は同じである。しかし仕事量は測定が 1 回ですむから, 3)の方が有利であるが, 情報は 1)の方が遙かに多い。

$n\sigma^2_M < \sigma^2_{M'}$, すなわち量が多くなると, 精度が著しく悪くなるようなときには, $V(\bar{x}) < V_1(\bar{x})$ となり, はじめて 1 枚ずつ測定した方が精度はよい。

$n\sigma_M < \sigma_{M'}$ 程度でも $V(\bar{x}) > V_1(\bar{x})$ である。

重量物の秤量するときなど, 秤の精度—感度—を考えて測定法をきめるとき, 一般に $n\sigma^2_M < \sigma^2_{M'}$ になることは少ないから, 一緒にした方が精度はよくなる。しかし重量物に対する秤量器はカタヨリが入りやすく, そのチェックも困難な場合が多いから注意を要する。

b) 化学分析の場合のように, たとえば錠剤を n コとり, これを一緒に粉碎混合して, 混合試料を作成し, 縮分して測定するときには, この測定値が平均値となる。この測定誤差 σ_M は, 縮分誤差と分析誤差の和である。

このときは $\bar{x} = \mu + \frac{\sum_{i=1}^n x_i}{n} + z_i$ という構造となる。

$$V_1(\bar{x}) = \frac{\sigma_s^2}{n} + \sigma_M^2 \quad (6.10)$$

測定を k 回、現在よく行われている方法は 2 回 ($k=2$)、繰り返せば、その平均値の精度は

$$V_2(\bar{x}) = \frac{\sigma_s^2}{n} + \frac{\sigma_M^2}{k} \quad (6.11)$$

ところがもしサンプルを n コずつ 2 回別々にとり、別々に縮分し、各 1 回分析を行えば、その平均値は

$$V_3(\bar{x}) = \frac{V_1(\bar{x})}{2} = \frac{\sigma_s^2}{2n} + \frac{\sigma_M^2}{2} \quad (6.12)$$

したがって $V_3(\bar{x}) < V_2(\bar{x})$

精度も後者の方がよくなり、またもしあらかじめ σ_M がわかっているときには、2つの測定値 x_1, x_2 から $V_3(\bar{x})$ をもとめることができるから、このサンプリング法の精度、錠剤間のバラツキ σ_x に関する情報もうることが出来るので有利である。

この場合各サンプルを $n/2$ コずつ 2 回別々にとりこれを別々に縮分、分析した平均値は、サンプルの合計の大きさは同じであり、 $V_2(\bar{x})$ と精度も

$$V_4(\bar{x}) = \frac{1}{2} \left(\frac{\sigma_s^2}{n/2} + \sigma_M^2 \right) = \frac{\sigma_s^2}{n} + \frac{\sigma_M^2}{2} \quad (6.13)$$

同じであり、しかも 2つのデータのバラツキから、平均値の精度やサンプルについての情報もえられるので、この方がよい場合が多い。

したがって、サンプルの縮分、調整が経済的に行われ、かつ分析者の技術に信頼性のある場合には、この方法を採用した方がよいと思う。この方法によれば、2つのデータ x_1, x_2 から得られる知識を利用して、 x_1, x_2 の分散 $V_5(x) \approx 2V_4(\bar{x})$ より平均値の精度を毎回もとめることができるし、これを利用して、常に合理的なサンプリング法を再設計できるし、あるいはこの2つのデータを群にとって、工程を管理することが、できる場合もあろう。

この場合、縮分、分析を2つのサンプルについて同じ人がやったときには、 σ_M^2 としては当然その縮分、分析の繰返精度を、各々他の人がやっ

たときには、再現精度を用いなければならない。

これらのサンプリング誤差と測定誤差との関係は、他のサンプリング法
のときにも、同様に考えておかなければならない。

6.1.2 層別サンプリング (計量値の場合)

層別サンプリングは、§4.3 にのべたように、ロットをいくつかの層
に、分けられるときに、その各層からサンプリングする方法であるが、す
べての層からサンプルをとると、その平均値のパラツキは、級内分散 σ_w^2
だけによるという特長をもっている。しかし級間分散 σ_b^2 と級内分散 σ_w^2
を知れば、副次サンプリングを行った方が経済的に有利な場合もある。従
来の全数検査的な考え方で行われている多くのサンプリングは、層別サン
プリングになっているものが多く、また必要以上の精度となっている場合
も多いのではないかと思われる。

たとえば瓶詰になった錠剤を各瓶からサンプルをとったり、臥詰の肥料
を各臥詰の組毎に層別して、各組から数臥のサンプルをとって正味重量を
測定したり、工程を流れてくる電球などの単位体を一定間隔で系統的サン
プリングした場合、いずれも層別サンプリングの効果を発揮して、その平
均値の精度は相当よくなる。

単位体で、各単位毎に測定できる場合には、そのデータより、§5にの
べたように、 $\hat{\sigma}_b$ 、 $\hat{\sigma}_w$ をもとめて、平均値の精度をもとめたり、各層に関
する情報がえられ、各層毎に層別した管理図がかけて、品質の保証と品質
の管理が同時に行えるので有利であるから、経済的、技術的に事情が許せ
ば、単位体毎に測定を行うべきである。化学分析のように混合試料とする
場合でも、そのサンプリングの目的が、総平均値だけを知ればよいのか、
各瓶毎の平均値や、各瓶間の分散 σ_b^2 、各錠剤間の分散 σ_w^2 を知りたいの
かを考えないと、層別サンプリングは精度がよいということだけを考え
て、全部を一緒に混合して分析し、ロットの総平均値だけを知っても役に
立たないことがあるから注意しなければならない。工程管理ということか
ら考えると、サンプリング法を工夫して、層間、副ロット間のパラツキや
傾向を知ることが非常に重要であることを忘れてはならない。われわれは
ロット全体に対してと同時に、各層に対して処置、行動をとらなければな
らぬことが多い。したがって、面倒でも、このような知識をえられるよう

に錠剤の混合法を工夫して、何回か分析を行わなければならないことがある。この場合、縮分や分析法が管理されていれば、従来のように同じサンプルを2回ずつ分析することをやめて、その手間をサンプル数を増加することに向けるべきである。この点からいっても、縮分、分析という工程を管理状態におくことが有利であることが判ろう。

例 1] 10本の錠剤 500錠入の瓶があるときに、そのある成分につき、各瓶内の錠剤間のバラツキが管理されていて、いずれも $\sigma_w = 1.0\%$ であることがわかっている。この総平均値を精度 (信頼度 99%) を 0.3% 以下で知るためには、各瓶より何錠ずつのサンプルをとればよいか。ただし縮分、分析の精度 $\sigma_M = 0.03\%$ であるとする。

$$\beta = 0.3 = u_\alpha \sigma_{\bar{x}}$$

ただし $u_\alpha = 2.57$ (正規分布表より $\alpha = 1\%$ の値)

いま分析を2回繰り返すとし、サンプリングによる平均値のバラツキを σ_s とすると

$$\sigma_{\bar{x}}^2 = \sigma_s^2 + \frac{\sigma_M^2}{2} \quad \text{しかるに} \quad \sigma_s^2 = \frac{\sigma_w^2}{n} = \frac{\sigma_w^2}{nk} \quad \therefore \bar{n} = \frac{\sigma_w^2}{k\sigma_s^2}$$

これらに各数値を入れて計算すると

$$\sigma_{\bar{x}} = 0.3/2.57 = 0.117$$

$$\sigma_s^2 = \sigma_{\bar{x}}^2 - \frac{\sigma_M^2}{2} = (0.117)^2 - \frac{(0.03)^2}{2} = 0.0137 - 0.0005 = 0.0132$$

$$\bar{n} = \frac{\sigma_w^2}{k\sigma_s^2} = \frac{1.00}{10 \times 0.0132} \approx 7.6$$

すなわち各瓶から、ランダムに8錠ずつとり、混合試料として2回分析すれば目的の精度が得られる。

この場合次のような仮定が入っていることに注意。

- 1) 各錠剤の目方は大体等しいとしている。
- 2) したがって各錠剤の成分 (%) の変動係数は、各錠剤の成分重量の変動係数と等しいとみなしている。

この仮定が成立しないときには、成分重量、たとえば mg で測定して、これについて分散をもとめ、上のように計算をすればよい。

6・1・3 副次サンプリング (計量値の場合)

§4・5, §5 にのべた通りの式の考え方でよく, 別に特記すべきことはない。応用例によって説明しよう。

例 ある電気部品を多量に生産している工場において, 100 コを1ロットとして工程を流し1時間に大体6ロット生産している。その電気抵抗を管理するために, 1時間毎に巡回検査し, その1時間に生産した6ロットから1ロットをランダムに抜き取り, 更にそのロット中から5コをランダムに抜き取り, その電気抵抗を測定して管理図を作成し工程を管理している。A組についてある1日について測定したデータは表6・1の通りである。ただし測定の精度 σ_M は1オームである。

- 1) 1日の電気抵抗の総平均値並びにその精度はいくらか。
- 2) 1日の総平均値をもっとも経済的に精度よく推定するためには, サンプルング法を如何にすべきか。

表6・1 ある電気部品の電気抵抗 (単位オーム)

時 間	9	10	11	12	14	15	16	17	
測 定 値	515	506	509	510	508	507	513	506	
	506	506	509	514	505	510	506	505	
	515	508	504	506	511	506	509	508	
	508	509	507	514	511	512	508	505	
	513	506	507	505	508	510	507	508	
\bar{x}_i	511.4	507.0	507.2	509.8	508.6	509.0	508.6	506.4	$\bar{\bar{x}}=508.5$
R_i	9	3	5	9	6	6	7	3	$\bar{R}=6.0$

$$D_4 \bar{R} = 2.11 \times 6.0 = 12.66 > R_i$$

したがって級内分散は大体管理されているとみなせる。

1) この例は系統的副次サンプリングの例である。もし1時間に生産された製品全体から単位体をランダムサンプリングすれば, これは層別サンプリングであり, また1時間毎を層と考え, その中からロットをランダムにとり, 更にそのロットから5コずつランダムにとると考えれば層別副次サンプリング法とも考えられる。この場合は副次サンプリングと考えて, といて見よう。1次単位は1時間毎のロットからランダムに

とっているので系統的サンプリングにおいて問題になる周期性の問題はない。

全ロット数 $M=6 \times 8=48$

サンプリングせるロット数 $m=8$

単位体のサンプル総数 $n=m\bar{n}=40$

総平均値 \bar{x} は (4.40) 式より

$$\bar{x} = \frac{1}{n} \sum_i^m \sum_j^{\bar{n}} x_{ij} = \frac{1}{40} \sum_i^8 \sum_j^5 x_{ij} = 508.5$$

平均値の分散は (4.43) 式において $a_1=1/6$, $a_2=5/100 < 0.1$

$$\therefore V(\bar{x}) = \frac{5}{6} \cdot \frac{\sigma_b^2}{8} + \frac{\sigma_w^2}{40}$$

σ_b , σ_w をデータから推定しなければならない。しかるに (5.18) 及び (5.19) 式より

$$\begin{aligned} \hat{\omega}_w^2 &\Leftrightarrow \frac{1}{m(\bar{n}-1)} \sum_i^m \sum_j^{\bar{n}} (x_{ij} - \bar{x}_i)^2 \quad (\bar{n}/N < 0.1 \text{ なるゆえ}) \\ \hat{\sigma}_b^2 &= \frac{M-1}{M} \left\{ \frac{1}{m-1} \sum_i^m (\bar{x}_i - \bar{x})^2 - \frac{N-\bar{n}}{N} \frac{\hat{\sigma}_w^2}{\bar{n}} \right\} \\ &\Leftrightarrow \frac{1}{m-1} \sum_i^m (\bar{x}_i - \bar{x})^2 - \frac{\hat{\sigma}_w^2}{\bar{n}} \end{aligned}$$

この場合測定誤差があるとして、その精度を σ_M とすると

$$\begin{aligned} \hat{\sigma}_w^2 &\Leftrightarrow \frac{1}{m(\bar{n}-1)} \sum_i^m \sum_j^{\bar{n}} (x_{ij} - \bar{x}_i)^2 - \sigma_M^2 \\ \hat{\sigma}_b^2 &\Leftrightarrow \frac{1}{m-1} \sum_i^m (\bar{x}_i - \bar{x})^2 - \frac{\hat{\sigma}_w^2}{\bar{n}} - \frac{\sigma_M^2}{\bar{n}} \\ &= \frac{1}{m-1} \sum_i^m (\bar{x}_i - \bar{x})^2 - \frac{1}{\bar{n}} (\hat{\sigma}_w^2 + \sigma_M^2) \end{aligned}$$

しかるに上のデータを一元配置法 (変量模型) で分散分析すると、次のような結果になり、その分散の期待値は表 6.2 の $E(V)$ 欄のようになる。したがって次のようにして各分散がもとめられる。

残差項の分散 σ_F^2 は、副ロット内の単位体間の分散 σ_w^2 と測定誤差による分散 σ_M^2 の和である。

$$\sigma_F^2 = \sigma_w^2 + \sigma_M^2$$

表 6.2 分散分析表

	二乗和	自由度	平均分散	$E(V)$	
ロット間	$S_b = \frac{1}{\bar{n}} \sum_i^m (X_i - \bar{X})^2$ $= \bar{n} \sum_i (\bar{x}_i - \bar{\bar{x}})^2$	93.6	$m-1=7$	13.4	$\sigma_K^2 + \bar{n}\sigma_b^2$
残差	$S_w = \sum \sum (x_{ij} - \bar{x}_i)^2$	254.4	$m(\bar{n}-1)=32$	8.0	σ_K^2
計	$\sum \sum (x_{ij} - \bar{\bar{x}})^2$	348.0	$m\bar{n}-1=39$		

したがって上表より

$$\hat{\sigma}_w^2 = \sigma_K^2 - \sigma_M^2 = \frac{S_w}{m(\bar{n}-1)} - \sigma_M^2 = 8.0 - 1.0 = 7.0 \quad \therefore \hat{\sigma}_w \approx 2.65$$

$$\hat{\sigma}_b^2 = \frac{S_b}{\bar{n}(m-1)} - \frac{\sigma_K^2}{\bar{n}} = \frac{S_b}{\bar{n}(m-1)} - \frac{1}{\bar{n}} (\hat{\sigma}_w^2 + \sigma_M^2)$$

$$= (13.4 - 8) / 5 \approx 1.1$$

となる。

測定誤差を考えにいれると、平均値の精度は

$$V(\bar{\bar{x}}) = \frac{5}{6} \frac{\hat{\sigma}_b^2}{8} + \frac{\hat{\sigma}_w^2}{40} + \frac{\sigma_M^2}{40} = \frac{5}{6} \frac{1.1}{8} + \frac{1}{40} (7.0 + 1.0) = 0.31$$

$$\sqrt{V(\bar{\bar{x}})} = 0.56$$

すなわち1日の製品の総平均値 $\bar{\bar{x}} = 508.5$ オームで、精度(標準偏差)は 0.56 オームである。

2) $m\bar{n}$ をいかに割当てたら、精度がもつともよくなるか。ただし $m\bar{n}$ は現在通り40とする。経費を考えると、1回まわったときに、6ロットの中から何ロットとつてもかわりないから m は巡回数に關係してこない。

したがって経費 $T = k_2 m\bar{n}$ と検査個数にのみよりきまると考えられる。

§4.5.1 にのべたように、 $\bar{n}=1$, $m=40$ にとつた方がよい。すなわち各巡回毎に各6ロット中から5ロットをランダムにえらび、各ロットより1コずつサンプルをとつた方が精度はよくなる。すなわち

$$V(\bar{\bar{x}}) = \left(1 - \frac{40}{48}\right) \frac{1.1}{40} + \frac{1}{40} (7.0 + 1.0) \approx 0.20$$

平均値の精度を考えるとこの方がよいが、工程管理という面から見るとその管理状態により、いずれをえらぶべきかを決定すべきである。

注意 1) このサンプリング法は、層別副次サンプリングともみなせるが、この場合には各層（1時間）から1ロットしかとっていないので、層内のロット間分散がもとめられぬので、層別副次サンプリングとしての平均値の精度はもとめられぬ。1時間に2ロット以上とってあればもとめられる。

2) このようなデータがB組についてもあるときに、AB組の平均値間の有意差はAB組を母数模型、ロット間などを変量模型とみて、二元配置法によって検定できる。

3) 工場において単位体が非常に多数あると各1次単位内の2次単位の数 N_i の不明な場合がある。このような場合は N_i が非常に多く $N_i \Rightarrow \bar{N}$ とみなせる場合が多い。この \bar{N} を推定するには

a) 1次単位をいくつかランダムサンプリングし N_i をかぞえ \bar{N} , σ_N を推定する。

b) 2次単位のいくつかの目方の平均値をもとめ、1次単位の重量を測定して、計算により \bar{N} をもとめる。

いずれにしても \bar{N} が相当大きいので少しバラツキがあってもその変動係数が小さい場合には、推定値や精度の式は上の公式通りに用いて差支えない。

6・2 計数値のサンプリング

単位体の場合には、たとえばその良、不良を判定する計数値の場合がある。あるロット中に不良品が何%あるか、何コあるかという、二項分布、多項分布、ポアソン分布に関することを推定する場合である。また計数値のサンプリングについては、その不良率が何%の製品は採択するか、棄却するかという、抜取検査の問題がある。後者については他書に譲ってここでは、主として推定の場合をのべる。

計数値をサンプリングにより推定する問題は、所謂標本調査法で、yes-or-noの答をうるような調査のときによく出てくるので、その方の書籍には相当詳細に出ているから、それを参照されたい。

6・2・1 ランダムサンプリング (計数値の場合)

ランダムサンプリングに関する注意は全く前にのべた通りである。推定式などについては、原則は§4にのべたと同じである。この場合は二項分布、多項分布あるいはポアソン分布となるので、その点がことなる。二項分布の性質から大きさ N の母集団の不良率を P とすると、良品率 Q は $Q=1-P$ 、母集団については

$$\mu = P = \frac{R}{N} \quad (\text{ただし } R: \text{ロット中の不良品数}) \quad (6.14)$$

$$\sigma = \sqrt{PQ} = \sqrt{P(1-P)} \quad (6.15)$$

$$\sigma^2 = PQ = P(1-P) \quad (6.16)$$

n コのサンプルをとり、その中に r コ不良品が入るとき、そのサンプル不良率 p は

$$p = \frac{r}{n} \quad (6.17)$$

このサンプル不良率がそのまま P の推定値 ($\hat{\mu}_p$) になる。

$$\hat{\mu}_p = \hat{\mu}(r/n) = p = \frac{r}{n} \quad (6.18)$$

これはちょうど計量値の場合、 n コの平均値に相当するものである。

$$\sigma_p^2 = \sigma_{\bar{x}}^2 = \frac{N-n}{N-1} \frac{\sigma^2}{n} = \frac{N-n}{N-1} \frac{PQ}{n} \approx \left(1 - \frac{n}{N}\right) \frac{PQ}{n} \quad (6.19)$$

$$\approx \frac{PQ}{n} \quad (n/N < 0.1) \quad (6.20)$$

$$\approx \frac{P}{n} \quad (n/N < 0.1, \quad P \leq 0.1) \quad (6.21)$$

$$\sigma_p = \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{PQ}{n}} \approx \left(1 - \frac{n}{2N}\right) \sqrt{\frac{PQ}{n}} \approx \sqrt{\frac{PQ}{n}} \approx \sqrt{\frac{p}{n}} \quad (6.22)$$

$$C_p = \frac{\sigma_p}{\mu_p} = \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{Q}{nP}} \approx \left(1 - \frac{n}{2N}\right) \sqrt{\frac{1}{n} \left(\frac{1}{P} - 1\right)} \quad (6.23)$$

$$\approx \sqrt{\frac{1-P}{nP}} = \sqrt{\frac{1/P-1}{n}} = \sqrt{\frac{1-P}{r}} \quad \left(\frac{n}{N} < 0.1\right) \quad (6.24)$$

$$\approx \frac{1}{\sqrt{nP}} \quad (n/N < 0.1, \quad P \leq 0.1) \quad (6.25)$$

ロット中の不良品数の推定値は

$$\hat{X} = \frac{N}{n} r = \frac{r}{a} = NP \quad (\text{ただし } n/N = a) \quad (6.26)$$

その精度は

$$\sigma_x^2 = N^2 \sigma_p^2 = N^2 \frac{N-n}{N-1} \frac{PQ}{n} \approx N(N-n) \frac{PQ}{n} \quad (6.27)$$

$$\approx N^2 \frac{PQ}{n} \quad (n/N < 0.1) \quad (6.28)$$

$$\approx N^2 \frac{P}{n} \quad (n/N < 0.1, P \leq 0.1) \quad (6.29)$$

$$C_x = \frac{\sigma_x}{\hat{X}} = \frac{N\sigma_p}{NP} = C_p \quad (6.30)$$

精度の式において P, Q は母集団についての値、すなわち母数であるが、実際にはサンプルからの推定値をいれる。すなわち σ^2 の不偏推定は

$$\hat{\sigma}^2 = \frac{n}{n-1} pq = \frac{n}{n-1} p(1-p) \quad (6.31)$$

$$\approx pq \quad (n \gg 1)$$

この値を上記の精度の各式に代入すればよい。

これらの精度の式を用いて、ある信頼度における不良率推定の精度をもとめるには、次のようにする。

1) $p \leq 0.5$ で $\bar{r} \geq 5$ 、すなわち不良率が 50% 以下でサンプル中に入る不良品数が 5 以上のときは、2 項分布はほとんど正規分布とみなしてもよいので、計量値の場合と全く同様に精度がもとめられる。 $\bar{r} \geq 3$ でも実用上は正規分布と近似的にみなして取り扱ってもよい場合が多い。

2) $\bar{r} < 3$ のとき、すなわち $np < 3$ であるから、サンプルの大きさが小さいか、あるいは不良率の非常に小さく、サンプル中の平均不良数が 2 以下になるようなときには、

$p \leq 0.05$ ならばポアソン分布表¹⁾により、

$0.10 > p > 0.05$ ならばポアソン分布、2 項分布いずれでもよい。

$p > 0.10$ ならば 2 項分布表¹⁾よりもとめられる。

1) 統計の数値表参照。

注意 1) 上の精度の式より次のような面白いことがわかる。

a) $n/N < 0.1$ ならば、推定の精度はロットの大きさ N には関係なく、 n によってのみきまる。これは前にものべたことであるが、従来のロットの大きさに比例して、サンプリングしているのは正しくない。統計的に精度のみから考えれば、一定精度をうるためには、 n はロットの大きさによりほとんどかわらぬ。したがって工程や製品が管理状態にあり、なるべく大きな等質のロットが得られるようになれば、同じ精度で小さいサンプル採取比 n/N ですむので、サンプリング経費の面から見ても有利である。しかし大きい N から小さい n をランダムにサンプリングする技術的困難さ、大きなロットについて間違った判定や処置をとると経済的に大きな影響をあたえるということを考えると、あまり N を大きくすることも考えもので、たとえば 1,000 とか 10,000 とかある程度の大きさのいくつかの副ロットに切って、サンプリングした方がよい場合も多いから注意。

b) 精度 σ_p は同じ n でも、 $p \leq 0.5$ のときは不良率 p が大きくなれば悪くなる。したがって精度の絶対値 σ_p を一定で推定するためには、一般に不良率の大きいもの程サンプルの大きさ n を大きくしなければならない。

c) 精度 C_p または C_x は、不良率が小さくなると、同じ n でも大きくなる。したがって相対的な精度 C_p 、 C_x を一定として推定するためには、 p の小さいほど、多くのサンプルをとらなければならない。

2) 良品、不良品を調査する際にも誤差はある。良品を不良品に間違える率を P_M' 、不良品を良品に間違える率を P_M とすると、調査の結果不良品としてあらわれる見かけ上の不良品数 R' は、

$$R' = R + (N - R)P_M' - RP_M \quad (6-32)$$

$$P' = \frac{R'}{N} = P + QP_M' - PP_M = P(1 - P_M - P_M') + P_M' \quad (6-33)$$

$$= P(1 - 2P_M) + P_M \quad (\text{ただし } P_M = P_M' \text{ とする})$$

$$\therefore V(P') = (1 - 2P_M)^2 V(p) + (1 - 2P)^2 V(P_M) \quad (6-34)$$

ただし $n/N < 0.1$ 、 $P_M = P_M'$ とし、 P_M' は P によりかわらぬも

のとする。

すなわち不良率の推定値には、カタヨリが入り、分散も厄介な型になる。

いま実際にあらわれた不良率を p' とし、近似的に

$$V(\hat{p}') = \frac{p'q'}{n-1} \quad \text{とすると}$$

$$V(\hat{p}) = \frac{1}{(1-2P_M)^2} \left[\frac{p'q'}{n-1} - (1-2p)^2 \frac{P_M(1-P_M)}{n} \right] \quad (6.35)$$

$$\text{ただし} \quad p = \frac{p' - P_M}{1 - 2P_M}$$

例 1] 1日に約 10,000 コ作成する製品中からランダムに 100 コサンプリングして、不良数 12 コであった。1日の不良率及びその推定の 95%信頼度の精度はいくらか。

$$p = \frac{12}{100} = 0.12$$

$$\hat{\sigma}_p = \sqrt{\frac{0.12 \times 0.88}{100}} = 0.032$$

($n > 1$ なる故 99 でなく 100 を用いる)

$r = 12$, $p = 0.12$ なる故正規分布とみなすと

$$\beta = u_{\alpha} \hat{\sigma}_p = 1.96 \times 0.032 \approx 0.063$$

不良率の推定値は 12.0% であり、その 95% 信頼度の精度は 6.3% である。

例 2] 不良率 P の母集団から、 n コのサンプルをとったとき、その中にある平均不良箇數 \bar{r} が変化するとそのサンプル不良率 p の変動係数 C_p はいかに変化するか。

$$C_p \approx \sqrt{\frac{1-P}{nP}} = \sqrt{\frac{1-P}{\bar{r}}}$$

p をいろいろにかえて、 \bar{r} により C_p がいかに変るかを表にまとめると表 6.3 のようになる。ただし小数点 3 位を 4 捨 5 入

i) 表からわかるように、 C_p の値はサンプル中にある不良品数によって大きく支配され、不良率にはあまり大きく支配されない。 C_p を一定にするには、 \bar{r} が大体同じになるようにサンプルの大きさをきめればよい。

したがって n は P に反比例して大きくなる。

表 6.3 P と \bar{r} の変化による C_p の変化

C_p の値		\bar{r}	4	9	16	25	36	49	64	100
P	$1-P$	$\sqrt{\bar{r}}$	2	3	4	5	6	7	8	10
		$\sqrt{1-P}$								
0.01	0.99	0.99	0.50	0.33	0.25	0.20	0.16	0.14	0.12	0.10
0.02	0.98	0.99	0.50	0.33	0.25	0.20	0.16	0.14	0.12	0.10
0.05	0.95	0.97	0.48	0.32	0.24	0.19	0.16	0.14	0.12	0.10
0.10	0.90	0.95	0.48	0.32	0.24	0.19	0.16	0.14	0.12	0.10
0.20	0.80	0.89	0.44	0.30	0.22	0.18	0.15	0.13	0.11	0.09
0.30	0.70	0.84	0.42	0.28	0.21	0.17	0.14	0.12	0.10	0.08

ii) \bar{r} が 35 以上になると P の値に関係なく C_p は 15% 以下になる。

iii) この表より $P < 0.1$ ならば $1-P \approx 1$ とみなしてよいことがわかる。

6.2.2 層別サンプリング (計数値の場合)

たとえばボルトが 5 台の機械により生産されたときに、各機械毎の製品を層とみて、各層からサンプルをとり、全体の平均不良率や、不良個数を推定する問題である。

N : ロットの全単位体数

N_i : 第 i 層の全単位体数 $N = \sum_i^k N_i$

$b_i = N_i/N$ (§ 4 では p_i としたが不良率 p_i との混同をさけるためにここでは b_i とした)

k : 層の数

n : 全サンプル数

n_i : 第 i 層からとるサンプル数 $n = \sum_i^k n_i$ $a_i = n_i/N_i$

P : ロット不良率

\hat{P} : その不偏推定量

P_i : 第 i 層の不良率

\hat{P}_i : //

X : 全不良品数

\hat{X} : //

X_i : 第 i 層の不良品数

\hat{X}_i : //

r_i : サンプル中の全不良数

r_i : 第 i 層からのサンプル中の不良数

p : 全サンプル不良率

p_i : 第 i 層のサンプル不良率

推定値は次のようになる.

$$\hat{X}_i = \frac{N_i}{n_i} r_i = N_i p_i \quad (6.36)$$

$$\hat{X} = \sum_i^k \hat{X}_i = \sum_i^k \frac{N_i}{n_i} r_i = \sum_i^k N_i p_i \quad (6.37)$$

$$= \frac{1}{a} \sum_i^k r_i \quad (\text{比例サンプリング } n_i/N_i = n/N = a) \quad (6.38)$$

$$\hat{P}_i = \frac{\hat{X}_i}{N_i} = p_i = \frac{r_i}{n_i} \quad (6.39)$$

$$\begin{aligned} \hat{P} &= \frac{\hat{X}}{N} = \frac{1}{N} \sum_i^k \frac{N_i}{n_i} r_i \\ &= \frac{1}{N} \sum_i^k N_i p_i = \sum b_i p_i \quad (6.40) \end{aligned}$$

$$= \frac{1}{n} \sum_i^k r_i = \frac{r}{n} = p \quad (\text{比例サンプリング}) \quad (6.41)$$

$$= \frac{1}{k} \sum_i^k p_i \quad (N_i = \bar{N} \text{ 一定}) \quad (6.42)$$

これらの精度は次のように §4・3 にのべた精度の式の σ_i のかわりに $P_i Q_i = P_i(1-P_i)$ とおいた形になる. \hat{X}_i , \hat{P}_i の精度は §6・2・1 と同じ.

$$V(\hat{P}) = \sum_i^k \frac{N_i - n_i}{N_i - 1} \frac{P_i(1-P_i)}{n_i} b_i^2 \quad (6.43)$$

$$\Leftrightarrow \sum_i^k \frac{P_i(1-P_i)}{n_i} b_i^2 \quad (n_i/N_i < 0.1) \quad (6.44)$$

$$\Leftrightarrow \frac{1}{k^2} \sum_i^k \frac{P_i(1-P_i)}{n_i} \quad (6.45)$$

$$\Leftrightarrow \frac{1}{n^2} \sum_i^k n_i P_i(1-P_i) \quad (n_i/N_i < 0.1, \text{ 比例サンプリング}) \quad (6.46)$$

$$\hat{\sigma}^2 = \frac{1}{nk} \sum_i^k P_i(1-P_i) \quad (6.47)$$

($n_i/N_i < 0.1$, $N_i = \bar{N} = \text{一定}$, 比例サンプリング)

$$V(\hat{X}) = N^2 V(\hat{P}) \quad (6.48)$$

サンプルの値から分散を推定するときには、上式の $\sigma^2_i = P_i(1-P_i)$ のかわりに

$$\hat{\sigma}^2_i = \frac{n_i}{n_i-1} p_i(1-p_i) \quad (6.49)$$

(ただし $n \gg 1$ のときは係数は不要)

を代入すればよい。

このほか §4.3 にのべたことが大部分あてはまる。

例 ある電気部品 2,000 コずつのロットが 7 ロットある。いま各ロットより各々ランダムに 50 コずつサンプルをとり、試験したところ、その不良数は 7, 2, 3, 6, 5, 9, 4 コであった。7 ロットの平均不良率、およびその推定の精度はいくらか。

表 6.4 計 算 表

ロット番号	n_i	r_i	p_i	$1-p_i$	$p_i(1-p_i)$
1	50	7	0.14	0.86	0.120
2	50	2	0.04	0.96	0.038
3	50	3	0.06	0.94	0.056
4	50	6	0.12	0.88	0.106
5	50	5	0.10	0.90	0.090
6	50	9	0.18	0.82	0.148
7	50	4	0.08	0.92	0.074
計	350	36	—	—	0.632

$$\hat{P} = \frac{r}{n} = \frac{36}{7 \times 50} = 0.1029 = 10.29\%$$

$$\begin{aligned} \hat{V}(\hat{P}) &= \frac{1}{nk} \sum_i^7 P_i(1-P_i) \approx \frac{1}{nk} \sum_i^7 p_i(1-p_i) \\ &= \frac{1}{350 \times 7} \times 0.632 = 2.58 \times 10^{-4} \end{aligned}$$

$$\begin{aligned}\therefore \hat{\sigma}_p &= 1.61 \times 10^{-2} \\ &= 1.61\%\end{aligned}$$

すなわち平均不良率 10.29% で、その精度は、標準偏差として 1.61% である。

6.2.3 副次サンプリング (計数値の場合)

たとえば薄鉄板が 100 枚束で何十束も入ったときに、その中から何束かをランダムにサンプリングし、次にその束の 100 枚中から何枚かを各々サンプリングして、その良不良を判定し、全ロットの不良数、不良率を推定する問題である。

この場合も §4.5 にのべた式に、2 項分布の場合の値をそれぞれ代入すればよい。記号は §6.2.2 と殆んど同じである。

M : 副ロット (1 次単位) の数

m : サンプリングする副ロット (1 次単位) の数

不良品全数および平均不良率の推定は

$$\hat{X} = \frac{M}{m} \sum_i^m \frac{N_i}{n_i} r_i = \frac{M}{m} \sum_i^m N_i p_i \quad (6.50)$$

$$\begin{aligned}&= \frac{1}{a_1} \sum_i^m N_i p_i \quad (m/M = a_1) \\ &= \frac{\sum_i r_i}{a_1 a_2} \quad (\text{比例サンプリング } n_i/N_i = a_2 \text{ 一定}) \quad (6.51)\end{aligned}$$

$$\hat{P} = \frac{\hat{X}}{N} = \frac{M}{Nm} \sum_i^m \frac{N_i}{n_i} r_i = \frac{M}{m} \sum_i^m b_i p_i \quad (6.52)$$

$$= \frac{M}{m} \frac{1}{Na_2} \sum_i^m r_i = \frac{r}{Na_1 a_2} \quad (\text{比例サンプリング}) \quad (6.53)$$

$$= \frac{r}{n} = p \quad (\text{比例サンプリング, } N_i = \bar{N}) \quad (6.54)$$

これらの精度は §4.5 の (4.41~4.44) 式において、 σ_x^2 , σ_i^2 , σ_b^2 , σ_w^2 , に次の値を代入すればよい。

$$\sigma_x^2 = \frac{1}{M} \sum_i^m \left(N_i P_i - \frac{NP}{M} \right)^2 \quad \sigma_i^2 = P_i Q_i \quad (6.55)$$

N_i が一定のときには

$$\sigma_b^2 = \frac{1}{M} \sum_i^M (P_i - P)^2 \quad \sigma_w^2 = \frac{1}{M} \sum_i^M P_i Q_i \quad (6.56)$$

サンプルからこれらの値を推定するには、次のように行う。

前にのべたように一般に母集団修正不要の際には、サンプルの級間分散の期待値は

$$E(s_b^2) = \sigma_b^2 + \frac{\sigma_w^2}{\bar{n}} \quad (n_i = \bar{n} \text{ 一定})$$

したがって、(4.44) 式より

$$V(\hat{P}) \approx \frac{\sigma_b^2}{m} + \frac{\sigma_w^2}{m\bar{n}} = \frac{1}{m} E(s_b^2) \quad (6.57)$$

($N_i = \bar{N}$, $n_i = \bar{n}$, $m/M < 0.1$, $n_i/N_i < 0.1$ のとき)

しかるに

$$s_b^2 = \frac{1}{m-1} \sum_i^m (p_i - p)^2 = \frac{1}{m-1} \sum_i^m \frac{1}{\bar{n}^2} (r_i - \bar{r})^2 \quad (6.58)$$

$$\therefore V(\hat{P}) \approx \frac{1}{m(m-1)} \sum_i^m (p_i - p)^2 = \frac{1}{m\bar{n}^2(m-1)} \sum_i^m (r_i - \bar{r})^2 \quad (6.59)$$

$V(\hat{X})$ は $V(\hat{P})$ を N^2 倍すればよい。一般の場合は計量の場合を参照されたい。

経費と結びつけて、サンプルの最適配分を知りたいときには、 σ_b^2 , σ_w^2 を各々推定しなければならない。このときには

$$\hat{\sigma}_w^2 = \frac{\bar{n}}{m(\bar{n}-1)} \sum_i^m p_i q_i \quad (6.60)$$

$$\approx \frac{1}{m} \sum_i^m p_i q_i \quad (\bar{n} \gg 1) \quad (6.61)$$

$$\hat{\sigma}_b^2 = E(s_b^2) - \frac{\sigma_w^2}{\bar{n}} = \frac{1}{m-1} \sum_i^m (p_i - p)^2 - \frac{1}{m(\bar{n}-1)} \sum_i^m p_i q_i \quad (6.62)$$

よりもとめればよい。

例 1 箱ボルト 500 コ入のものが 100 箱入った。この中から 5 箱サンプリングし、各箱から 50 コのボルトをサンプリングし、調査したところ不良品の数が、5, 10, 7, 3, 4 であった。不良ボルトの全数は何本

か. その精度はいくらか. また今後どのようにサンプリングしたら, もっとも推定の精度がよくなるか.

$$a_1 = 5/100 = 0.05 \quad a_2 = 50/500 = 0.10$$

$$\therefore \hat{X} = \frac{100}{5} \times \frac{500}{50} \times (5+10+7+3+4) = 5800 \text{ 本}$$

$$\begin{aligned} V(\hat{X}) &= N^2 V(\hat{p}) = \frac{N^2}{m\bar{n}^2(m-1)} \sum_{i=1}^m (r_i - \bar{r})^2 \\ &= \frac{(500 \times 100)^2}{(50)^2 5 \times 4} [(-0.8)^2 + (4.2)^2 + (1.2)^2 + (2.8)^2 + (-1.8)^2] \\ &= 1000^2 \frac{30.8}{20} = 1.54 \times 1000^2 \end{aligned}$$

$$\therefore \hat{\sigma}_X = 1240 \text{ 本}$$

したがって 2.5 シグマをとると

$$2.5 \hat{\sigma}_X = 3100 \text{ 本}$$

最適配分をもとめるために $\hat{\sigma}_w^2$, $\hat{\sigma}_b^2$ をもとめると

$$\begin{aligned} \hat{\sigma}_w^2 &\Rightarrow \frac{1}{5} (10 \times 90 + 20 \times 80 + 14 \times 86 + 6 \times 94 + 8 \times 92) \times 10^{-4} \\ &= 1000.8 \times 10^{-4} \end{aligned}$$

$$\hat{\sigma}_b^2 = \frac{1}{4} \frac{30.8}{50^2} - \frac{\hat{\sigma}_w^2}{50} = 10.8 \times 10^{-4}$$

いま経費 (T) 一定で精度よく推定することを考える. $T = 3,000$ 円
 $k_1 : k_2 = 200 \text{円} : 10 \text{円}$ とすると

$$T = k_1 m + k_2 m \bar{n}$$

$$\bar{n} = \sqrt{\frac{k_1}{k_2} \frac{\sigma_w^2}{\sigma_b^2}} = \sqrt{\frac{200}{10} \times \frac{1001.8}{10.8}} \Rightarrow 43 \text{ 本}$$

$$m = \frac{T}{k_1 + k_2 \bar{n}} = \frac{3,000}{200 + 430} \Rightarrow 5 \text{ 箱}$$

したがって 5 箱とり, 各 40 本ずつとれば $T = 3,000$ 円 となり, 精度がもっともよい.

備考 この計算には, 次のような表を用いるのも便利である,

表 6・5 計 算 表

副ロット 番号	N_i	n_i	r_i	p_i	q_i	$p_i q_i$	$p_i - \bar{p}$	$(p_i - \bar{p})^2$
15	500	50	5	0.10	0.90	0.0900	-0.016	256×10^{-6}
38	500	50	10	0.20	0.80	0.1600	0.084	$7056 \times "$
51	500	50	7	0.14	0.86	0.1204	0.024	$576 \times "$
73	500	50	3	0.06	0.94	0.0564	-0.056	$3136 \times "$
88	500	50	4	0.08	0.92	0.0736	-0.036	$1296 \times "$
計	2500	250	29	—	—	0.5004	—	1.232×10^{-2}

$$p = \frac{29}{250} = 0.116$$

6・2・4 比推定 (計数の場合)

§ 4 では計量の場合だけのべたのでここでは、計数値の場合の比推定について、ランダムサンプリングの場合をのべる。

たとえば、1日に生産された鋼材 N 本から n 本をランダムにとり、検査する場合を考えると、このときにたとえば、1級品、2級品、不良品と分けられるときに、

i) 1級品の数と2級品の数の比を推定する場合

ii) 不良品中原因 A による不良品の割合 (比) を推定する場合
にわけられる。i) は1本の鋼材が1級品ときまれば、これは2級品には入らない。すなわち1級品と2級品とは相反した特性である。排反的な特性である。これに対し ii) は原因 A による不良品は、不良品であって (特性 B をもって)、しかもその原因が A によるものであるというのであるから、原因 A による不良品は不良品という特性に含まれている。このように2つの特性が、排反的であるか、一方が他に含まれているかにより、統計的な取り扱いかたがことになっている。

比の推定は、両者とも同様に次の式であらわされる。

$$f = \frac{p_A}{p_B} = \frac{r_A}{r_B} \quad (6 \cdot 63)$$

ただし $p_A = r_A/n$ n : サンプルの大きさ
 $p_B = r_B/n$

		i) の場合	ii) の場合
$r_A : A$ という特性をもつサ ンプルの数		2 級品数	原因 A による不良品数
$r_B : B$	〃	1 級品数	不良品数

この推定の分散は i) と ii) ではことなる。

i) 排反的な場合

$$\hat{V}(f) = f^2 \frac{N-n}{N-1} \frac{1}{n-1} \left[\frac{1}{p_A} + \frac{1}{p_B} \right] \quad (6.64)$$

$$\hat{C}_f^2 = \frac{N-n}{N-1} \frac{1}{n-1} \left[\frac{1}{p_A} + \frac{1}{p_B} \right] \quad (6.65)$$

ii) 特性 B が、特性 A を含んでいる場合

$$\hat{V}(f) = f^2 \frac{N-n}{N-1} \frac{1}{n-1} \left[\frac{1}{p_A} - \frac{1}{p_B} \right] \quad (6.66)$$

$$\hat{C}_f^2 = \frac{N-n}{N-1} \frac{1}{n-1} \left[\frac{1}{p_A} - \frac{1}{p_B} \right] \quad (6.67)$$

例] ある合成樹脂製品 5,000 コのロットから 100 コをランダムサンプリングし、検査したら、1 級品 65 コ、2 級品 17 コ、不良品数 18 コ、不良品中表面不良によるもの 10 コあった。

i) このロットの 2 級品、1 級品の比はいくらか。またその精度はいくらか。

ii) ロットの不良品中表面不良によるものの割合、およびその精度はいくらか。

この問題は i) は排反的な現象であり、ii) は一方の特性が他に含まれている。

$$i) \quad p_A = \frac{17}{100}, \quad p_B = \frac{65}{100}$$

$$\therefore f = \frac{p_A}{p_B} = \frac{17}{65} = 0.262$$

(6.65) 式より

$$\hat{C}_f^2 = \frac{N-n}{N-1} \frac{1}{n-1} \left(\frac{1}{p_A} + \frac{1}{p_B} \right) = \frac{1}{100} \left[\frac{100}{17} + \frac{100}{65} \right] = 0.074$$

$$\therefore \hat{C}_f = \sqrt{0.074} = 0.27 \quad f \text{ の変動係数は } 27\%$$

$$\hat{s}_f \approx 0.071$$

$$\text{ii) } P_A = \frac{10}{100}, P_B = \frac{18}{100}$$

AはBに含まれている。したがって、分散は(6-66), (6-67)式を用いる。

$$f = \frac{10}{18} = 0.556$$

$$\hat{\sigma}_f^2 \approx \frac{1}{100} \left[\frac{100}{10} - \frac{100}{18} \right] = 0.044 \quad \therefore \hat{\sigma}_f = 0.21$$

$$\hat{s}_f = 0.117$$

第7章 集合体のサンプリング

7・1 まえがき

1コ1コ単位体としてサンプリングするのではなく、集合した材料などの中から、その一部分を単位としてサンプリングする場合で、この場合には計数的な測定を行うことは少なく、計量的な場合が大部分である。

集合体のサンプリングが工場においてもっとも問題になる点が多い。これには統計的な手法よりも、統計的な考え方を応用して、技術的に考えて行かなければならぬ問題が多い。§4、§5にのべたいろいろな式も勿論縦横に使われるが、これと同時に技術的な知識の活用が特にたいせつである。

集合体のサンプリングにおいては、特に次のようなことが問題になる。

- i) サンプル単位——インクリメントの大きさ、試長——をきめること
- ii) サンプル単位をカタヨリのないようにとる技術上の問題
- iii) ロットの状況によりランダムサンプリングが経済的、技術的に困難あるいは不可能な場合が多い。
- iv) 混合サンプルの縮分法が統計的、技術的に問題になる。

以下の各節ではこれらのことを考えながら、統計的、技術的に話をすすめて行こう。

集合体を、対象物の状況により次のようにわけて考える。

1) 流体のサンプリング（ガス体および液体）

2) ケーキ状、泥状体のサンプリング

たとえばフィルターケーキ、湿潤パルプ、その他有機合成反応にある反応中または中間体など

3) 連続体のサンプリング

たとえば、糸、電線、紙、帯鋼など長い連続的な材料（流体を除く）

4) 同じ大きさの副ロットにわけられている集合体（固体）のサンプリング

たとえば、貨車積の鉱石類、臥詰の肥料、袋入、瓶、樽詰などの工業薬品

5) 粉塊混合物のサンプリング

本章においては、1), 2), 3), 4) についてのべ、5) はいろいろな問題があるので章をわけて、§ 8 でのべる。

7・2 流体のサンプリング

流体にも非常に混合しやすい水素ガスのようなものから、タール、アスファルト、濃いコロイド溶液類のように、温度その他の条件を変化させなければ流体とみなせないようなものまでいろいろあるが、一般に流体のサンプリングは、技術的な注意をほらえば、固体のサンプリングより容易な場合が多い。まず共通点、特殊な点について考えて行こう。

流体が完全に均一に混合している場合にはサンプル単位の量や数に関係なく、極く少量のサンプルで十分であり、たとえば1滴の液、1 cc のガスでも、サンプルとしてもそのロットの性質を十分に示してくれる。不均一な流体でも、容易に十分混合できる場合には、わざわざ面倒なサンプリング理論を考えるよりも、均一に混合することを技術的に考えた方が経済的で、よい場合が多い。しかしある混合法、攪拌法によって、均一に混合できるか否かは、あらかじめいくつかの部分から何 cc ずつかのサンプルをとって、その混合状況を統計的にチェックしておく必要がある。完全に混合していると思って安心していても、混合が不十分で、それからの1つのサンプルにより判定していると、カタヨリや大きなバラツキが入っていることがあるから注意を要する。

このチェックには、たとえばよく攪拌した後に、ロットの、技術的に考えてもっとも混合されにくい部分や、よく混合されていると思われる部分から、いくつかのサンプルをとり、そのサンプルの各々を別々に更によく混合して、いくつかに分けて繰り返し測定し、これを一元配置法により解析してサンプル間に有意差が出なければ、まず混合は十分とみなしてよい。またもし従来のデータから測定誤差の分散 σ^2_M がわかっており、よく管理された測定者、測定法により上にとったサンプルを測定させる場合には、各サンプルを1回ずつ測定させて、その測定値の不偏分散と σ^2_M と比較して、 σ^2_M が大きければ問題はないが、 σ^2_M が小さくとも、 σ^2_M を

分母として比をとり，自由度 $f_1 = n - 1$ (ただし n はサンプルの数)， $f_2 = \infty$ として， F 検定により有意差の検定をすればよい。

場合によっては，不均一なものをわざわざ混合しなくとも，簡単に層別，集落，あるいは系統的サンプリング法などにより，よい精度でサンプリングを行うこともできる。

一般に流体は，特にパイプ中，装置中などを連続的に流れている流体は，プロセス的な変化をなし，次のような特徴をもっている。

1. 流体は混合が比較的良好に行われる場合が多いので，その特性の変化は次第に緩慢になる。特に太いパイプや，途中にタンクやガスホルダーなどのあるときは，特性のバラツキは小さくなるのが普通である。

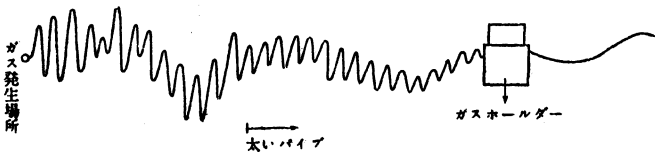


図 7.1 流体のバラツキの変化

2. したがって流体の特性は長いパイプやガスホルダーを出た後には一般に比較的大きな周期や傾向で，平均値のまわりにバラツク，次の工程へガスを引き渡すのが，ホルダー後であるような場合には，発生直後の小波，たとえば通常ガス発生炉によくあらわれる小波よりは，むしろホルダー後の大波を管理すべき場合が多い。したがって，タンク在庫量や，ホルダーの高さを管理すると，平均サンプルをとりやすくなることがある。

われわれがあるガスのロット，たとえば1時間，10分，1分にガスの流れを切って考えて，その平均値を知るためにサンプリングするには，一般に連続的にサンプルをとり，ある期間の平均サンプルとする。この場合，流量や圧力の変化があると，流量に比例した量ずつのサンプルをとることが技術的に困難になる。あるいは一定間隔で系統的サンプリングをして，混合試料としてもよい。このときには，流量に比例させるということと，サンプリング間隔と特性の周期的変化について，§ 2.2.2 にのべたよう

な注意が必要である。

流量と、流体の特性が互に独立に、相関関係がなくランダムに変化する時に連続的に或いは一定間隔で一定量ずつのサンプルをとっているならば、混合試料の品質特性はカタヨリはないが、その精度は流量が一定の時、流量に比例してサンプリングする時よりも悪くなる。流量と特性の変化が相関があるときには、カタヨリが入る。これは各サンプルを別々に測定して、その平均をとるときも同様で、現在よく工場で行われているように、ただ一定時間に一定量のサンプルをとり、その測定値を平均して、1日の平均成分として推定しているときにも、カタヨリが入る。しかし流量や圧力の変動係数が、品質特性の変動係数に比し小さいときは、あまり問題にならない。

3. 流体は流動論にしたがって運動する。したがって

a) その流体が層流をなして流れているか、混流(乱流)をなして流れているか、レイノルズ数などから計算してチェックしておく必要がある。混流のときには、比較的均一になっているから、パイプ中のサンプリング場所をあまり注意する必要はないが、管壁附近では層流をなし、不均一になっている場合が多いから、管壁からのみサンプリングするのは正しくない場合が多い。層流のときには、管内の流速分布を考えて、サンプリングパイプを配置しなければならない。

b) 一般に管壁、炉壁、曲り部分などの近くでは、運動、混合が不十分な場合があり、その他死角にある流体は、正しくサンプリング時現在の特性を示さぬことがあるから、サンプリングパイプ取付場所に注意しなければならない。

c) 温度が不均一、比重が不均一な場合には場所による不均一や対流などがあるから、サンプリング法、場所に注意しなければならない。

d) 粘度の大きい流体は、不均一になりやすいから、もしできれば、加熱して粘度を小としてサンプリングするか、または固化して、固体としてのサンプリング法を適用した方がよい場合もある。

4. 流体の場合には、大口試料を均一に混合することは容易であるから、縮分の誤差はほとんど考えなくてよい場合が多い。しかし混合は十分に行わなくてはならない。特に半固溶体のときには加熱したり、また溶剤

などでうすめることもある。

注意 1) 一般に温度、圧力、流量などを一つの測定値であって、品質特性と考えない傾向があるが、これも流体の一つの重要な品質特性である。特に流体の場合には、これらの特性値が工程に大きな影響をあたえている場合が多い。したがって、温度、圧力、流量の測定時間も一つのサンプリング上の問題である。しかしこの場合には、実際にサンプルを取り出すという問題はないが、測定時間は、やはり一つの工程からのサンプリングと考えて、本書の各所で述べた、統計的な考え方や管理状態を考えて決定すべきである。

2) 一般に流体の場合には、自動計器が発達している。したがってサンプリング法の検討と同時に、技術的にこれらの計器を研究することが必要である。またこれを何所に設置すべきかということにも統計的な検討が必要である。

7.2.1 気体のサンプリング

気体の場合には特に流動が激しいから、流動に関する知識、あるいはガスの燃焼とか、反応とかについての技術的な知識を用いて、知りたい特性を代表するようにサンプリングしなければならない。一般に気体では混合がよく行われるので、採取場所、採取法などを技術的に十分考えて、サンプリング法をきめれば、サンプリングによる誤差はおこらず、ただちにロットの特性や変化を知ることができる。したがって統計的な面のみでなく、技術的にサンプリング法を十分に検討すべきである。気体のサンプリングは容易であるというので、以下本章にのべるような注意事項がまもられていない場合が多いから、一度工場の現状をふりかえっていただきたい。

気体を状態により分類すると

- i) 移動中のガス (パイプ中、工程中)
- ii) 静止中のガス (ガスホルダー、ポンペ)

この各々について、サンプリング上注意すべきことをのべる。

1. 移動中の気体 気体は静止中のものよりも移動中の場合が多い。移動中の気体は、一般に時間的経過により、質的並びに量的に変化するので、サンプリングの際には、一つの工程 (プロセス) からのサンプリングと考えなければならない。特に反応工程中、あるいは反応直後の

ガスは、一般にガス体は反応速度が早いために、その変動の周期も極めて短い場合がある。このようなときには、スナップサンプリングして、短時間をロットと考えて、工程を管理することができぬ、あるいは管理しても意味の少ない場合も多いから、その変動を輸送中のパイプ、あるいはガスタンクなどで緩和して、大波あるいは傾向的な変動にかえた後にサンプリングすべき場合もある。小波は途中で相当の空間をおけば、小さくなる場合が多いが、もしこのような小波の除去が不可能な場合には、サンプリングして分析するよりも、自動記録計などをつけて測定した方がよい。またプロセス的な傾向の強いガスでは、自動記録計あるいは自動調節装置が必要な場合も多い。しかしこれらの装置をつけるときにも、そのサンプリングパイプ、サンプリング法については、以下にのべるような注意が必要である。

一般に気体サンプリングに際し注意すべきことは、工程の変動（流量、組成、圧力、温度）、採取場所、採取法、採取器などについてである。

a) 流動中の気体のサンプリング法としては、サンプリング時間の非常に短いスナップサンプリングとある期間中の平均サンプリング法の2種がある（両者の差は採取時間の長短のみであるが）。工程を管理するためには、スナップサンプリングの方が多くの情報が得られるが、前にのべたようにスナップサンプリングでは、工程管理の目的からいって意味のない場合があり、サンプリングの目的を考えて、スナップの時間を検討する必要がある。

平均サンプリングも、ガスの長期間、たとえば1日の平均品質を保証するために行われることもあるが、一般に品質のバラツキを知る必要がある場合が多いので、あまり長時間の平均サンプリングでは意味のない場合が多い。平均サンプリングでは、ガス量に比例してサンプリングする装置が必要である。

短い週期の変動をもつ気体では、スナップサンプリングを系統的サンプリング法で行っても、ランダムサンプリングとみなしうる。

b) 一般にサンプリングパイプは、ガスパイプの中央部まで挿入すること。層流のときにはパイプの流速などを考えて、直径方向にいろいろの深さに数本のサンプリングパイプを付けなければならない場合もある。ボイ

ラー等において燃焼状況を見るときには、その目的とするガス組成、変動状況をよく示すようなところにパイプを挿入しなければならない。管壁、炉壁など死角になるところからサンプルをとっている場合が案外多い。

c) サンプリングパイプは、特にゴミ、炭塵などの固形物、昇華物、ミストなどを含んでいるガスの際には、真直ぐな、比較的大きな口をもったものがよい。小孔を多くあけたものは穴がつまりやすく、スリット型の口は特別の場合以外はあまりよくない。また掃除や手入の容易な型にしておくこと。

d) ガス採取に際しては、サンプリングパイプ中のガスが完全に目的とするロットと同じ組成のガスと置換してから採取せねばならない。したがってサンプリングパイプは可及的短い方がよい場合が多い。自動指示計、自動記録計、自動制御装置などにおいても、サンプリングパイプはなるべく短い方がよく、もしメーター室などまで相当長くサンプリングパイプを配管しなければならぬときには、その時間の遅れをあらかじめ調査しておく必要がある。特に高圧パイプのときには、時間の遅れが甚しいから、高圧部のパイプを短くするなど、注意を要する。

e) サンプリングパイプの材質は、強度、耐熱、耐蝕性を考慮して決定しなければならない。これは特に、高温、高圧、あるいは腐蝕性ガスのときに考慮しなければならぬ問題で、サンプリングパイプの僅かな節約をして、工程などの管理が不能になり、「百文惜しみの銭失い」にならぬよう留意したいものである。

f) 気体は反応速度が早く、特にボイラー、炉など高温の場合には、サンプリングパイプが触媒作用をなし、ガスの成分が変化して、サンプルがロットの組成を示さなくなることがあるから、注意しなければならない。ガラス、磁器、シリカチューブなどはこの点よいか、耐熱性、強度などに留意して保護管をつける必要がある。ガス管、不銹鋼管などが通常よく用いられるが、平炉やガラス炉などのように高温ガスを採取するときには、水冷式にすべきである。これは腐蝕を防ぐと同時に、採取したガスが直ちに冷却されるので、触媒作用によるガス成分の変化を防ぐことができる。サンプルの組成が変化しては、ロットの推定にカタヨリが入る。しかし工程管理のためには、常に同じようにカタヨリが入り、そのカタヨリが管理

されていれば、十分役立つことに注意。

g) 負圧のところからのサンプリングは、器具に洩れがないように十分点検し、注意しなければならぬ。採取場所は目的の場所にできるだけ近いところを選ばねばならない。たとえばボイラー等では、ダンパーの前でとるべきであり、ダンパー後の煙突附近からとったのでは、ダンパー、煙道その他における空気洩れにより、われわれの知りたい真の燃焼を知ることができない。サンプル採取器には特別な装置をつけて、たとえばガス成分のとけぬような液体を用い、サンプル採取後圧力を常圧、できれば正圧にして空気の混入を防ぐとよい。

h) 正圧のところからサンプリングしたときには、採取後もサンプルをある程度正圧に保っておいた方がよい。

2. 静止中の気体 静止中の気体としては、容器、ポンプなどに入ってから完全に静止中のガスと、ガスホルダーのように一方からガスが流入し、他方から流出しつつあるような半静止状態のものとする。

一般に気体のサンプリングは静止中の方が、正しい代表的なサンプルをとることが技術的に容易である。

容器などに入っているガス体は、通常十分に混合されているとみなせるから、どこから採取してもよいか、一応混合状態、特に比重差による不均一をチェックしておく必要がある。

ポンプ入りなどの液化ガスのサンプリングは一般に液化せる部分からサンプリングすべきである。ガス化せる部分のみをとると、これは沸点の高い気化しやすい部分であるから気相、液相の平衡関係により、カタヨったサンプルをとることになる。

一般に気相部分の重量は、液相部分の重量に比し極めて僅かであるから、液相部分をとれば十分である。

ホルダーなどのように半静止中のものも大体均一に混合しているとみなしうるが、ホルダー中のガスの平均品質を知るためには、ガス流入口、流出口附近からのサンプリングはさけて、均一に混合しているとみなしうる場所からサンプリングすべきである。

7.2.2 液体のサンプリング

液体の場合も大体 §7.2, §7.2.1 にのべたことが当てはまる。この場

台も技術的な問題が多い。

液体の性質、状態、容器などによって、サンプリング法がわけられる。

性質としては主として粘度——したがって均一性——と揮発性並びに異物混入に留意しなければならぬ。

状態としては、移動中と静止中にわけられるが、気体が移動中のサンプリングの問題が多いのに比し、液体の場合には静止中のサンプリングの問題の方が多い¹⁾。静止中にもっとも問題になるのは、その容器中、容器間のバラツキの問題である。しかし固体の場合に比し、液体は必ず容器に入っており、容器内のバラツキは技術的に解決できる場合が多いので、固体のサンプリングよりも容易である。

タンクや容器に入った液体からサンプルをとり平均品質を推定することは、固体よりは容易であるが、その不均一性、バラツキを考えると、やはり移動中すなわちタンクに入れるとき、あるいは出すときにサンプリングした方が問題は少ない。

移動中の液体は、気体の場合より流量が一定の場合が多く、圧力により体積も変化せぬので、連続サンプリングなど技術的に容易に行える。

一般に粘度や濃度、比重が大になるに従って不均一になり、また結晶、不純物、異物などの混入や析出も多くなってくるので、ガス体の場合よりサンプリングに対する技術的工夫、並びにバラツキの大きさやその変化状況を統計的に検討する必要が生じてくる。

液体のサンプリングについては、石油製品は、非常に粘度の低い、揮発性のものから、非常に粘度の高い半固溶体のものまであり、液体の種類を大体含んでいるので、これについて話をすすめて行く。他の液体についてはこれを参考として、サンプリング法を研究されたい。ASTM²⁾も1950年に特にサンプリング法を決定しており、英国³⁾においても大体これと類似しているので、これらを中心に注意事項をのべて行く。

- 1) 更に固体になると益々静止中のものが多くなる。一般に静止中はランダムサンプリングが困難になるので、移動のチャンスを利用してサンプリングを心掛けることになる。
- 2) ASTM Manual on Measurement and Sampling of Petroleum and Petroleum Products. 1950.
- 3) Standard Methods for Testing Petroleum and its Product. Institute of Petroleum.

1. サンプリング法一覽表 液体のサンプリングにおいても、1つずつのサンプルを問題にするとときと、いくつかのサンプルを一緒にして混合試料の平均品質を問題にするとときとあるが、一般にサンプリング法とその適用範囲をまとめて表7・1に示す。この表では、サンプリング法、すなわち採取器、採取法を、液体の揮発性、粘度、状態、容器などによりいくつかに分類している。

表7・1 サンプリング法とその適用範囲

方 法	液 体 の 性 質	容 器 , 状 態
瓶、ピッカーによるサンプリング	蒸気圧 16 lb Rvp.* 以下の液体。加熱により液体となる半液体	タンク車、タンクトラック、船、バージタンク、貯蔵タンク
タップサンプリング	26 lb Rvp. 以下の液体	貯蔵タンク
連続サンプリング	2 lb Rvp. 以下の液体	移動中 (パイプなど)
柄杓サンプリング	2 lb Rvp. 以下の液体、半液体	開放又は流出口における流れ
チューブサンプリング	同 上	ドラム缶、樽、ガロン缶
シーフサンプリング	同上の底部のサンプルをとるとき	タンク車、貯蔵タンク
ボーリングサンプリング	ワックス及び軟固体	樽、箱、袋詰、ケーキ状
グラブサンプリング	塊状固体	貨車、コンベヤー、バンカー、瓶、袋、樽、箱
グリースのサンプリング	グリース	釜、ドラム缶、ガロン缶
工業用芳香族炭化水素のサンプリング	ベンゼン、トルエン、キシレン、粗並びに精製ソルベンメナフサ	タンク車、タンクトラック、船及びバージタンク、貯蔵タンク、ドラム缶、ガロン缶

* Rvp.=Reid Vapour Pressure リードの蒸気圧。石油製品の蒸気圧は一般にリード法で測定される。ガソリンのように揮発性の液体は蒸気圧が高い。通常の自動車用ガソリンは最大13.5 (冬) 11.5 (春、秋) 9.5 (夏) lb 以下である。16~26 lb 以上高いものは特に低沸点溜分のロスに注意しなければならない。

以上の分類は、次のようなことに留意して分類している。

i) 揮発性、揮発性物質のロスによるサンプルのカタヨリをさけるため

に、サンプルの採取法、容器、保存法に十分注意する。

ii) 粘度が大になると不均一になり、サンプリング法に工夫が必要である。

iii) 不純物、スラッジ（結晶）、などの沈澱する場合には、これを層別して別にサンプリングする。

iv) 液体が2層に分離する場合があるから、それに留意する。

v) 容器や流体の運動状態。

以上の点を技術的によく考えて、サンプリング法を設計し、サンプリングを実施しなければならない。

サンプリングに際しては、容器や手などの清掃、揮発性可燃物のときには、発火、爆発に対する注意など技術的常識を実行することも必要である。

2. サンプルの種類 いろいろな方法によりえられるサンプルの種類をまず定義しておこう。

a) 平均サンプル (average sample) 平均サンプルとは、容器中の各部分から、その液体の量に比例してとった混合サンプルである。しかし一般にはこのようなことは困難な場合が多いので、以下にのべるようなサンプルをとる。

容器に入っている液体は、はげしく攪拌して、できるだけ均一に混合し、未だ攪拌のとまらぬ中にサンプリングする（攪拌サンプル）。パイプ中を流れる場合には、垂直管にサンプリングパイプをつけて、一定量ずつサンプリングする（連続サンプル）。通常の連続的なポンプ輸送の場合、流量がほぼ一定とみなしうるときには、厳密に流量に比例したサンプリングでなくとも十分である。水平管は、粘度の大きい、あるいは異物などを含む液体のときには、特に流速ののろい場合には、不適當である。

b) 全層サンプル (all-levels sample) サンプル採取器を容器の底部まで入れ、蓋をあけて、これを次第に引き上げて得たサンプルである。この場合引き上げた後にサンプル採取器が一杯になってはまずい。全層サンプルは平均サンプルでない場合が多い。たとえば容器の横断面積が等しくなかったり、あるいは引き上げる速度をうまく調節できないことが多いからである。採取器中へ流体の入る速度は、大体深さの平方根に比例するので、引き上げ速度を流入速度が一定になるように調節しなければな

らず、これが困難、厄介である。後にのべるチューブサンプリング法か、全層サンプルをとるのに用いられる。

c) 上部サンプル (upper sample) タンクなどに入っている液体の上部3分の1の大体中央部からとったサンプル (図7.2 参照)。英国では

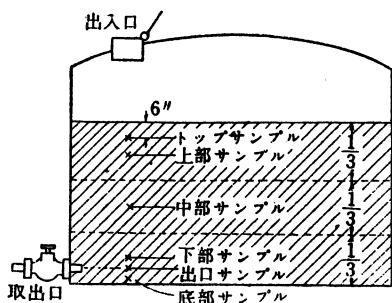


図7.2 タンクからのサンプリングの深さ (ASTM)

上から10%のところからとったサンプルとしている。

d) 中部サンプル (middle sample) タンクなどに入っている液体の深さの丁度中央のところからとったサンプル (図7.2 参照)

e) 下部サンプル (lower sample) タンク中の液体の下部3分の1の中央からとったサンプル (図7.2

参照)。英国では下から10%のところからとったサンプル。

表7.2 水平円筒型タンクからのサンプリング法及びサンプルの混合割合 (ASTM)

液体の入っている割合 %	サンプルをとる底からの高さ %			混合割合		
	上部	中部	下部	上部	中部	下部
100	80	50	20	3	4	3
90	75	50	20	3	4	3
80	70	50	20	2	5	3
70	—	50	20	—	6	4
60	—	50	20	—	5	5
50	—	40	20	—	4	6
40	—	—	20	—	—	10
30	—	—	15	—	—	10
20	—	—	10	—	—	10
10	—	—	5	—	—	10

f) 混合サンプル (1つのタンクの場合) (single tank composite sample) 1つのタンクの上部, 中部, 下部サンプルを混合したもの。

タンクの型が図7.2のように横断面積の等しい場合には, 上中下部サンプルを各々同量ずつまぜればよい。英国の場合は上下部サンプルが少しちがうので, 1:3:1の割合にとっている。

横断面積の等しくない場合には, 各部の容積に比例した割合で混合しなければならない。円筒型の横型タンクの場合には, 表7.2または表7.3のような割合でとる。この場合タンクが満量していない場合には, 表7.2のように各部サンプルをとる位置や量が変わってくることに注意。これらの混合割合は, 石油製品についての表であるから, 他の場合には, これできいかチェックして, 表を作成しておくとい。

表 7.3 水平円筒型タンク
よりの混合サンプル混合
割合(英国)(満量の場合)

種 類	混合割合
上部サンプル	1
中部サンプル	8
下部サンプル	1

g) 混合サンプル (多くのタンクの場合) (multiple tank composite sample)

船で多くのタンクをもつとき, パージ取する場合に多く用いられる。各タンク的全層サンプル, あるいは混合サンプルをよく混合した後, 各タンクの液量に比例して全部一緒に混合したサンプルである。

h) トップサンプル (top sample) 液面から6吋の深さからとったサンプル (図7.2参照)。

液上層を層別して, 何か特殊な性質をもっていないかをしらべる。

i) 出口サンプル (outlet sample) タンク取出口からとったサンプル。取出口はタンク下方からのみ取り出すようになったものも, またスウィングパイプにより上部から取り出せるようになっているものもある。このスウィングパイプは他の液体の出荷のときにも, 特に水分や異物が底にたまりやすいときに用いたらよい。

j) 連続サンプル (continuous sample) 輸送中のパイプから, 輸送期間中連続に流量に比例してとったサンプル。

k) 柄杓サンプル (dipper sample) 柄杓その他の汲取用の道具を用いて流れている流体を, その全期間中一定間隔でその流れの各部分より一

定量ずつ集めたサンプル。

l) 攪拌サンプル (mixed sample) 液体を容器中で十分に混合、または激しく攪拌した後、必要量を注出し、または取り出したサンプル。

m) チューブサンプル (tube sample) またはシーフサンプル (thief sample) チューブあるいはシーフ (図 7・6 参照) を用いて、容器中の特別な場所からとったサンプル。これは平均サンプルのみでなく、底とか特別に異状のありそうなところからのサンプリングにも用いられる。

n) ドレインサンプル (drain sample) 水抜きバルブよりとったサンプル。たとえばタンク式では次にのべる底部サンプルと同じものになることがある。

o) 底部サンプル (bottom sample) タンク、容器などの底やパイプなどのもっとも低いところからとったサンプルで、n) o) は多くのスラッジ、ゴミ、その他の固形物や水分等特別な調査をするのに用いる。

p) ボーリングサンプル (boring sample) ワックス、固型油脂類などのように、常温で軟い固体状のものを樽、箱、袋などの中からサンプリングするとき、上から下まで穴をあけてとるサンプル。

q) グラブサンプル (grab sample) 粉塊混合物のサンプリングで、あるインクリメントで混合物を一定量ずつとったサンプルである。p) q) は主として固体のサンプリングの考え方に属する。

3. サンプル容器 通常透明または褐色ガラス瓶、缶等が用いられる。

透明ガラス瓶は中がよく見えて、透明度、水分、固体不純物などをそのままで見ることができるので便利である。褐色瓶は光線により、変化をおこすような液体を入れるときに用いる。これらガラス容器は破損しやすい。しかし耐蝕性がある。

缶類は外側から見えないという欠点がある。ハンダ付は外側からしたもののがよい。またハンダ付に用いるフラックスは洗滌容易なものでなければならぬ。このフラックスがサンプルに影響をあたえることがあるから注意。また缶の材料は液により腐蝕されないものでなければならぬ。

栓や蓋が必ずついたものでなければならぬ。これは揮発性物質のロスと他からの夾雑物の混入を防ぐために必要である。特に揮発性物質の際に

は密栓できるもの。また液やその蒸気により腐蝕や変化しないものでなければならぬ。場合によっては栓に耐蝕性の薄片のカバー（たとえば金属薄片、合成樹脂のシート）をしてもよい。

容器は使用前に十分清掃、乾燥しておかなければならない。特に缶の場合には、外部から見えぬから十分に行わなければならない。水溶液を入れるときは、使用前にサンプルと同じ液でよく洗滌すれば乾燥の不要のこともある。

掃除、乾燥がすんだら直ちに栓をしておくこと。

サンプルをとったら直ちに所要事項を書いたレッテルをはること。レッテルがはがれたり、汚損したり、消えたり、腐蝕されたりしないように十分に注意すること。

サンプルをサンプル容器にうつすときに、特に揮発性のサンプルは、揮発分のロスのないように注意して迅速に行わなければならない。揮発性の大きいサンプルは、試験室で栓をあける前に冷却しておくこと。

サンプルは膨脹、収縮があるから、サンプル容器に一杯に入れてはならない。

サンプルを保存したり、輸送したりするときには、密栓をして合成樹脂製のキャップをしてしっかりしぼるか、しっかりした栓をして、紙などでつつみ、パラフィンなどをつけること。栓に直接にパラフィンやシールワックスをつけてはならない。なおこれをあけるとときには、サンプル中にこれらのものが入らぬように留意すること。

4. サンプリングを行う時期および個数

a) 製品 製品を出荷あるいは荷下しするときには、出す方と入る方と両方あるいはいずれか一方のタンクからサンプルをとる。必要ならばパイプラインからとる。このときパイプライン中に前の製品がのこっていると、製品が汚れたり、不純になったりするから、パイピングに対する注意、あるいははじめにパイプから出た製品は他へ除くとかしなければならぬことがある。

b) 船またはバージのタンク、タンク車 積込み直後または積下し直前にとる。

c) ドラム缶、ガロン缶、樽、箱などに入っている場合 ASTM で

はロットにつき表7・4にあるサンプル数以上の容器をランダムにとるよ
うに、きめられている。これはロットの大きさを N とすると、 \sqrt{N} (た
だし端数を切り上げ) にサンプルの大きさ n をとっている。これは前にも
のべたように $\sigma_{\bar{x}} = \sigma_x / \sqrt{n}$ で n は N には統計的には無関係であるが、 σ_x
の変化、経済性を考えると必ずしも N に無関係に n をきめてはならない
とのべた (§6・2・1) ことを示している一例である。表7・4 はアメリカに
おける1ロットについての表で、多くの生産者から集めたようなロット、
あるいはアメリカの大製油所と管理状態のことなった製品のロットにおい
て、そのままあてはまるか否かは疑問であり、一応ドラム缶間などのバラ
ツキや管理状態をチェックしてからでなければ、この数字を直ちに一般の
場合に採用することは危険であろう。

表7・4 容器入の出荷ロットからのサンプリング必要最小個数

ロットの 大きさ	サンプル の大きさ	ロットの大きさ	サンプル の大きさ	ロットの大きさ	サンプル の大きさ
1	1	513~729	9	4,097~4,913	17
2~8	2	730~1,000	10	4,914~5,832	18
9~27	3	1,001~1,331	11	5,833~6,859	19
28~64	4	1,332~1,728	12	6,860~8,000	20
65~125	5	1,729~2,197	13	8,001~9,261	21
126~216	6	2,198~2,744	14	9,262~10,648	22
217~343	7	2,745~3,375	15	10,649~12,167	23
344~512	8	3,376~4,096	16	12,168~13,824	24

5. サンプリング法 以下各サンプリング法について注意すべき点を
簡単にのべる (表7・1 参照)。

a) 容器によるサンプリング いろいろの容器があるが (図7・3)、
いずれも栓のある、しかも任意のところで栓を外すことができ、重りをつ
けて液体中に沈むものでなければならぬ。容器の口の径は通常 15~40
mm くらいのものが用いられるが、液体の粘度大なるものは大きい口径の
ものを用いる。これでとるサンプルは、全層、上、中、下部、トップ、出
口サンプルなどである。またこれらを混合して混合サンプルとすることも
ある。この混合は試験室で行わなければならない。一般のほとんどすべて

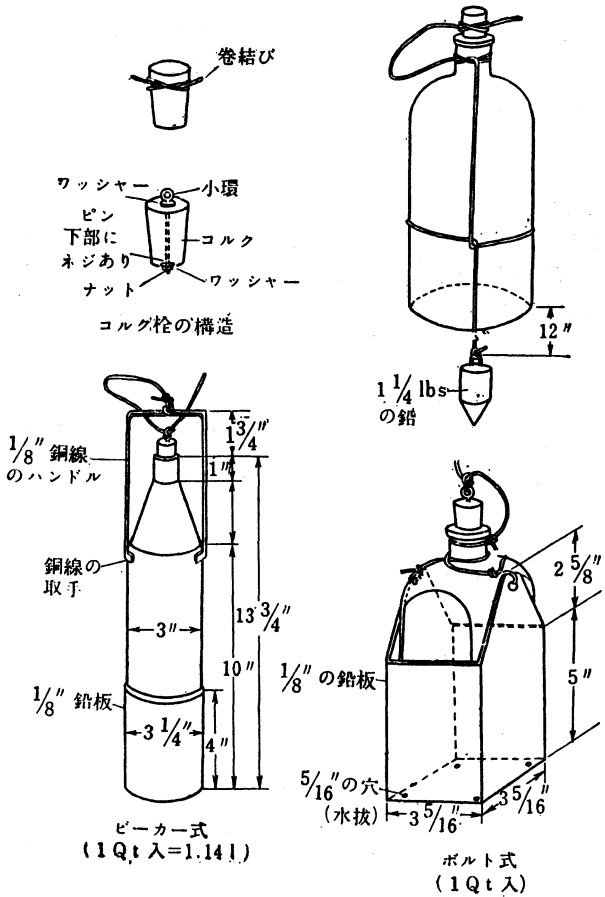


図 7.8 サンプリング容器

の液体（製品および原油）に用いられる。

b) タップサンプリング 揮発性の液体を揮発分のロスなくサンプリ

ングする方法である。タップはタンクの内側 1 m くらいまでパイプを入れ、いろいろな高さのところに少なくとも 3 本は必要である。サンプリングパイプはサンプル容器の底まで入るようにしておき、サンプリングに際してはパイプを液中に浸して、液をとること。揮発性の大きい場合には、瓶を金物で補強しておくこと。

c) 連続サンプリング パイプ中を流れている、蒸気圧の低い (2 lb 以下) 液体、半液体をサンプリングするのに用いる。パイプの中に、図 7・4 のように、管壁より管の半径の $1/10$, $3/10$, $6/10$ のところに、各 3 本のパイプを入れて、この各々から同量ずつのサンプルが流れ出すようにサンプルをとる。採取量は全期間中流れる液量の 0.1%, しかし 150 l 以上はならない。図のようにパイプ垂直部で行うこと。

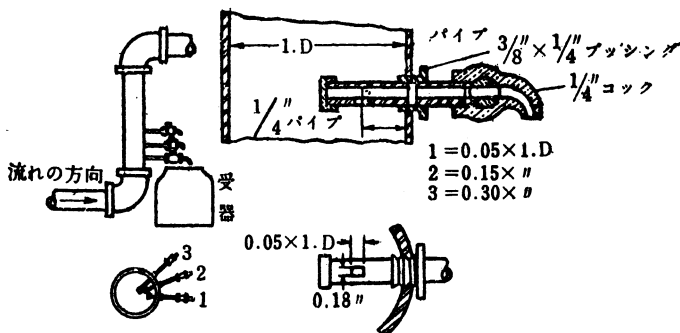


図 7・4 連続サンプリングのサンプリングパイプ

d) 柄杓サンプリング 蒸気圧の低い (2 lb Rvp. 以下) 場合に用いる。柄杓でサンプリングするのであるから、2 吋以下の細いパイプからの出口などでサンプリングするのに用いられる。たとえば缶詰の際、柄杓の大きさはできれば、流出口の流れの全断面積より大きいものを用い、全断面積をカバーするようにサンプリングした方がよい。通常は一定間隔で、系統的サンプリングを行う。大口試料の量はロットの 0.1%, しかし 150 l 以上はならない。

e) チューブサンプリング ドラム、樽、ガロン缶入などで、蒸気圧

の低い (2 lb Rvp. 以下) 液体のサンプリングに用いる。ガラスまたは金属性の管で 500 cc~1 l 入り、かつ容器の底から 5 mm のところまでとどくものを用いる。管の上端は親指で押えやすい。また下方の口も比較的小さい形にしておく (図 7.5 参照)。管の上端を親指で押えて、適当の深さまで管を入れて、親指をはなせば、そのサンプルがとれる。全層サンプルをとるときには、指で押えずに上から底までチューブを徐々に入れて、一杯になったら指で押えて引き出せばよい。水分とかゴミなどを層別して見たいときは、指で押えて底までチューブを入れてサンプルをとる。

ガロン缶以上のものはこのサンプリング法でサンプルをとるが、ガロン缶より小さい缶の場合には、缶全量をサンプルとする。

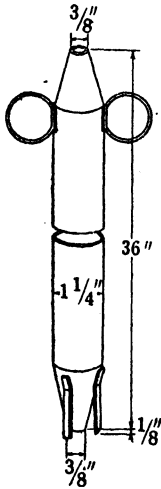


図 7.5 サンプリング用チューブ

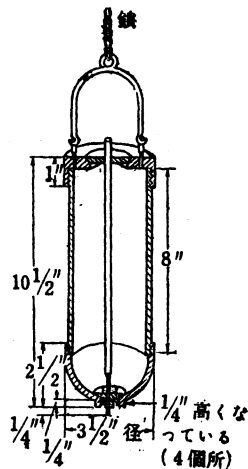


図 7.6 サンプリング用シーフ

f) シーフサンプリング タンク車やタンクの底部サンプルをとるのに用いる。シーフは図 7.6 のように、シーフがタンク底部につくと自然に蓋があき、引き上げると自然に蓋がしまるようになっている。

g) ボーリングサンプリング ワックスや軟い固体をとかしてサンプリング

リングするのが困難なときに用いる。樽、箱、袋、塊などをまず表7・4により1次サンプリングし、次に各1次単位より2次単位として中心部、周辺部、その中間から3本のボーリングをとり、これをサンプルとする。別々に試験室まで持参し、試験室でよく見て外観上3つの2次単位が異なっていれば別々に分析する。外観上差がないときは一緒に混合してもよい。もしサンプルを縮分する必要があるときは、冷却し粉碎して縮分する。表面の異物は、サンプルに入れるか、入れぬかは目的によりきめよ。

h) グラブサンプリング 塊状固体のサンプリングに用いる。これに用いるシャベルは最大粒度のものが十分入るものでなければならない。このサンプリングは、後に粉塊混合物のサンプリングでのべるが、他の粉塊混合物よりバラツキの小さい場合が多いから、サンプリングも容易であるが、荷下し、積荷中、コンベヤーで輸送中のように動いているときにサンプリングした方がよい。貨車からとるときには少なくとも12個所からとること。8個所は貨車の四隅の各上下から、4個所は中央附近の上下各2個所。

大口試料はロットの0.1%、ただし50lb以上、1,000lb以下のこと。縮分法は§8参照。

i) グリースのサンプリング グリースは、その品種、製造条件に相当変化がある。また製造工程中と、これを缶詰、樽詰などしたあとでは、性質も、分離状況も相当バラツキがあるので次のような考え方が必要である。

イ) 製品の品質を調査するためには釜、冷却槽などの工程中からサンプリングしてはいけない。必ず容器に入れてから、12時間以上冷却後サンプルをとること。

ロ) 各ロット毎に変化のあるものであるから、各ロット毎にサンプルをとること。

ハ) 一缶の中でも表面と内部と異なることがあるから、あらかじめ外観上両者の差異をしらべること。また攪拌してサンプルをとってはならぬ。他のグリース状のものでも、このようなことがよくある。

グリースについては、表面と表面から15cm位下部とか見かけ上不均一のときには、両方から別々に約500gのサンプルをとる。均一と見られ

るときには、表面から 7~8 cm 下からサンプルをとる。1つのロットで容器毎のバラツキの大きいときには、サンプルは缶毎に別々にとり、別々に分析すること。

サンプルはグリースの固さによりスプーンまたはナイフなどでとる。

j) 工業用芳香族炭化水素のサンプリング 試験用に最小1ポンドのサンプルが必要である。

他は大体その容器により、これまでのべたサンプリング法を応用すればよい。底部サンプルは通常、水分、異物混入による濁を見るためにとるが、底部サンプルが濁っていたり、異物が入っていたりしたときには、さらに1フィート上から別にサンプルをとる、これも汚れているときには2フィート上からサンプルをとる。

注意 1) 以上の数字は主として ASTM についてのべたのであるが、数字よりもむしろ考え方を理解していればよい。しかし考え方も不十分などところがあるから、本書の他の部分にのべた考え方を参照されたい。またこの数字はアメリカにおける管理状態に対して作成されたもので、そのまま日本の現状に適用できるかは別問題である。

日本におけるその液体の管理状態を、統計的に検討しておかなければ数字の合理性は証明できない。

2) 以上のべた中には、インクリメント間のバラツキや、インクリメントの大きさ、平均値の精度などについては、殆んどふれていないが、以上のようなサンプリングを行うと、比較的よく混合されるときにはインクリメント間の分散が非常に小さく、あるいは0になる。しかしグリースやワックスなどからのサンプリングのときには、インクリメント間のバラツキに注目しなければならない。

3) 水分、ゴミなどは、通常容器の底にあるので、これを層別してサンプリングする。サンプリングの目的によっては、全層にわたって比例サンプリングすることもある。

4) グリース、ワックスのように半固溶体のサンプリングには、後にのべる §7.3 ケーキ状のサンプリング、§7.5 副ロットにわけたときの固体のサンプリングの考え方を利用すべきである。

5) 熔融体、たとえば熔銑などのような金属類の液体、あるいはカー

パイドのように炉からタップするときには液状のもの、あるいは熔融塩類のようなものも、大体流体についてのべた考え方で行けばよい。出鉄中、出鋼中の金属、タップ中のカーバイドは流動中の液体であるから、その1出鉄、1タップの平均品質を知るためには、その均一性、あるいはそのプロセス的な傾向や周期性を調査して、サンプリング法をきめなければならない。多くの場合一定間隔のスナップサンプリングにより、系統的サンプリングを行えば、相当よい精度で平均品質を知ることができよう。しかしこの場合には、比重の関係で不純物が表面に浮き、あるいは下方に沈んでいることが多いから、サンプリングの目的を考えて、それをとるか否かを決めなければならない。

容器に入ったこれらの熔融体は、完全に均一とみなせるものと、比重などの関係で上下層で品質のことになっている場合とある。均一のときには、表面に近いところ（表面はよくない）から小さなインクリメントをとればよいのであるから技術的にも容易であるが、不均一ときには、これらの熔融体は高温の場合が多いので全層から、あるいは層別してサンプリングすることが困難であるので、できるだけ移動中、すなわち注入中、注出中にサンプリングした方がよい。

すぐに凝固はじめるような場合には、凝固に際して成分の分離のおこる場合が多いから、凝固中にサンプリングするのはよくない。むしろ完全に凝固してしまってから、固体として、§7・5あるいは§8などにのべるような方法でサンプリングすべきである。

またこれらの熔融体は高温の場合が多いので、空気水分などによる酸化などにより、サンプルの組成がロットと異なってくる場合、あるいはロットが、時間と共に変化してサンプルと異なってしまう場合が多いから、その目的を考えて、サンプルの保存法、時間的経過による組成の変化などを検討する必要がある。またサンプリング装置の技術的検討も必要であろう。

たとえば出鋼のときに、ある品質特性が時間的にある一定の傾向や周期をもっていて、その状況が管理されていれば、適当な時期に1サンプルとれば、工程管理のためには差支えない場合もあろうが、その時間的経過の管理状態が保証されていないかぎり、1サンプルでは不十分である。

7.3 ケーキ状、泥状体のサンプリング

化学系工業に半製品としてよくある状況で、たとえば固体中に水分、薬品類、溶剤、油などを多量に含んだ状態、あるいは液体中に多量の固体を含んだ状態のものを対象とするサンプリングである。このような対象物でも、たとえば苛性ソーダを含んだ1束のパルプ、フィルタープレスから出した1つ1つのケーキの目方を対象とするときは、単位体のサンプリングとして取り扱えばよい。ここではこのような対象物の、液体含有量、純度などを目的とする場合を考えよう。

この場合の特徴としては、次のような点である。

1) 液体あるいは固体の含有量により、品質特性のパラツキが大きく変化する。たとえば、フィルターケーキ中の液体量——平均値——の少ない場合には、パラツキも比較的小さく、分布が少し歪んでいる。平均値が大きくなるにしたがいパラツキも大きくなり、ある程度以上ますとパラツキはまた小さくなり、液体量が非常に増加し、泥状体になると更にパラツキが小さくなり、最後には殆んど均一状態となる。これは粉塊混合物の場合 (§8.3.2, 図8.3 参照) とよく似ている。

2) 液体の含有量が、場所的、あるいは時間的に傾向や大波をもつ場合が多く、それが工程の管理状態により変化する。

3) サンプル採取器に特別な技術的工夫を要する場合が多い。

4) サンプリングに際し、液体をしぼり出すようなカタヨリが入りやすい。たとえば湿潤パルプからポーリングしてサンプルをとると、液体分を他へ圧出してしまうことがある。多くの場合この条件により、インクリメントの大体の大きさがきまってくる。

5) 縮分に大きな問題がある。

以上のような点があるので、この分類に属するサンプリングは仲々厄介である。しかし統計的に考えると、§7.5, §8 とよく似ている。たとえば、フィルターケーキが各々別個にあるときには、ある大きさのインクリメントでとったときに、ケーキ内の分散とケーキ間の分散を考えれば、副次、層別、集落サンプリングのいずれかを行えばよいことになり、これらの分散の大きさや管理状態を推定することが問題になってくる。またもし

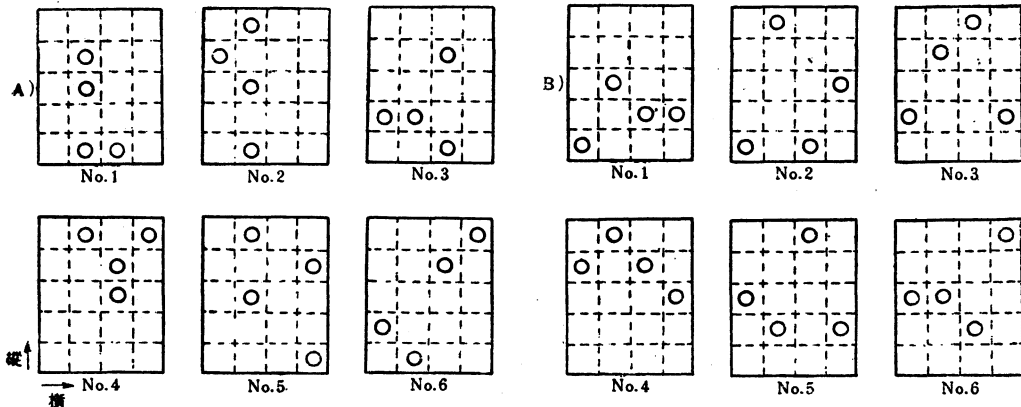
ケーキなどがバラバラになって混合状態にあるときには、粉塊混合物のサンプリングと同様に考えて行けばよい。

a) インクリメントの大きさと数 インクリメントの大きさの決定は、i) 液体分の圧出作用によるカタヨリのない程度大きくしなければならない。ii) カタヨリのないインクリメントについて、§ 8にのべるように、インクリメント間の分散と所要精度をうるのに必要なインクリメントの数との関係を考えて、もっとも経済的に、しかも実行しうる程度にインクリメントの大きさを決定することになる。この点については§ 8を参照されたい。

b) サンプリング法 ケーキやパルプのように、目の前に性状を見ながらインクリメントをとることになると、経験者程液体分の多少が直感的にわかるので、サンプリングにカタヨリが入りやすいから、サンプル採取者には、ランダムにとるという教育を行い、それを実行させることが現実には大きな問題になろう。これにはいろいろな工夫が考えられるが、たとえばパルプのように形の一定のものであれば、乱数表のかわりに図7・7のようなサンプリングカードのようなものを100枚程準備しておき、このカードをランダムにぬいて、カードに指定された箇所をインクリメントでとるようにするのも1つの考案であろう。このほかいろいろのカード類の工夫が考えられるから、各現場の実情に応じて考案されたい。

c) 縮分 この種のサンプリングにおいて多く問題になるのは、液体分含有量である。したがって以上のような工夫をして集めた各インクリメントを一緒にして混合試料とし測定する場合が多く、この混合試料を縮分するという問題がある。ところが§ 8にのべるように縮分の誤差を少なくするためには、なるべく粉碎して行わなければならないが、液体含有量を目的とする場合、あるいは純度を目的とするときにも液体含有量が大きな影響をあたえるので、粉碎、縮分をしていたのでは、液体部分の揮発、流出等によりサンプルがカタヨってしまう。液体含有量が非常に少なく、揮発によるカタヨリの考えられぬ場合はよいが、しからざるときは普通の縮分法は適用できない。

この対策も技術的にいろいろ考えられるが、たとえば i) 最も簡単な方法は混合試料全量をそのまま乾燥して、湿分を測定してから他の特性を分



作成法 A. カードをたとえ20の区割に仕切り、1から20まで番号をうち、乱数表を引いて4区割をランダムにとる。これはランダムサンプリングである。

B. もしパルプが横方向にバラツキが大きいということが、予備知識としてわかっていれば、インクリメントを4コとするには、各列に層別して、各列からランダムにサンプリングした方が、層別サンプリングの効果により精度がよくなる。そこで各列毎に1から5までの数より1コを乱数表よりきめて、サンプリングすべき箇所にマークを打つ。

備考] 1) もし対象物が大きく、インクリメントの大きさが小さく、ある方向へのバラツキが大きいときに、とるべきインクリメント

の数が3というときには、層間分散が大きくなるように、3つの層にわけ、その中に多くの区割をつくつて、その中からランダムに1区割をあるいは層の大きさに比例して区割をえらぶ。——層別サンプリング

あるいは層間分散が大きくなる方向に多くの列で区切り、まず列を3コランダムにえらび、さらにその列内で1区割をランダムにえらぶようにすれば、副次サンプリングの形になる。層間分散が小さくなるように列で区切り、列をランダムにえらび、えらんだ列を全部とれば集落サンプリングとなる。

2) 区割の大きさは、区割内で有意選出が行われぬように、インクリメントの大きさをくわく小さくしておいた方がよい。

図 7.7 サンプルカード

析する。ii) 湿分測定用のサンプルを別にとる。iii) § 8・5・3 にのべるスコップ法で混合試料中から分析用サンプルをとる。iv) 混合試料の重量を測定しておいて、これを全量、液体で、あるいは水などで稀釈、あるいは溶解して、均一な一定量の液状として、その中から一定量の平均サンプルをとり、測定する。等々の技術的な工夫が必要である。

注意 1) この種のサンプリングは、多くの場合同一工場の中間工程においておこるので、たとえば原単位の計算には各ロットごとのサンプリングの精度を要求しなくともよく、1月の平均値の精度がよければ役立つことも多いのではないと思われる。1日の精度がわるくとも、1月の精度はよくなることに注意。ただしカタヨリはこまる。

2) 有機化学工業のバッチ式の工程において、その品質特性、収率などの管理図において \bar{R} の大きい場合には、中間体のケーキ類のサンプリングの正確さや精度がよくない場合があるから、 \bar{R} が大きいときには一度サンプリング法をチェックしてみるとよい。

7・4 連続体のサンプリング

糸、紙、電線、フィルムのような長く連続した帯状、線状の固体の一部をとって、その品質を推定するためのサンプリングである。

このサンプリングにおいては、次のようなことが問題となる。

1) サンプリングするために、その連続体の一部を切りとらねばならない。製品となると1巻になっている。この1巻からランダムにサンプリングするということは、多くの場合その部分を切りとらねばならぬので、その製品の価値が低下し、また場合によってはそのようなサンプリングをすることが不可能な場合がある。

2) サンプリングするのにサンプル単位としてサンプルの長さ、試長をどの位にとったらよいかということの決定、これは後にのべるようにサンプリングの目的と、その製品や工程の大波、小波、傾向などのバラツキ状況によってきめられる。

3) 工程中の連続体は、たとえば紡糸中の糸、抄紙中の紙、圧延中の帯鋼のようなものから切り取ってサンプリングすることは機械の運転を休止したり、阻害する。またこの場合、非常な高速で移動しているので、その

まま測定できる場合でも、サンプリングや測定が技術的に困難な場合が多い。ところが製品になると1巻になってしまう。

4) これらの連続体は、長さの方向に、その品質特性に大波や傾向のある場合が多い。巾の方向に傾向をもつことも多い。したがってその変化状況、並びにその管理状態を予備知識として知る必要がある。

5) これらのサンプリングにおいて、品質特性として強度、巾並びに厚さ、長径と短径のように2次元的な長さ、あるいは外觀上の不良や欠陥を調査することを目的とする場合が多い（勿論成分水分などを問題とする場合もあるが）。

これらのことを考えながら、連続体のサンプリング法並に測定法の考え方を論じて行こう。

a) 測定法 まず測定法であるが、1巻に巻きとってしまったものはサンプリングが困難になるので、サンプリングの原則である移動中に、すなわち巻きとっている間にサンプリングして、切斷せずに工程中に測定する方法があれば一番簡単である。物理的に、物理化学的に、あるいは電気的に迅速に、しかも連続的に測定できる方法が考えられれば、平均品質を知るサンプリングの面から見ても、あるいは工程管理という面から見ても非常に有利で、技術的にこの方向への研究が必要である。特に強度、分析試験のような破壊試験の場合には、それと密接な相関関係のある測定容易な品質特性を発見することが有利である。たとえば、米国では、抄紙機で紙の厚さをはかるのに、放射性物質を用いて、一方より照射し、その透過量を測定して、紙の厚さを測定しているということである。あるいはX線、電気抵抗、高周波などによる迅速測定法も考えられよう。このように連続体のサンプリングは、測定法により非常に容易になってくるので、測定法の研究が一つの重要な問題である。

b) 1巻内のバラツキ 連続体は1巻内のバラツキに大波や傾向のある場合が多いから、予備知識として、どうしてもその調査が必要である。いろいろな試長に切って、系統的に、特性値の変化を調査する必要がある。この調査も、工業におけるサンプリングの予備調査の原則として、1巻をしらべるのみでなく、1ロット内での各1巻内の変化、ロットの異なるときのこれらの変化、並びにその管理状態を知るように、サンプリング

間隔は大きくてもよいから、多くの巻についてまず大体のバラツキの傾向をつかみ、次に必要ならば詳細な調査に移るべきである。

特に外周の近くの特性値が、1巻の平均値やバラツキとどのような関係にあるかをチェックしておく。たとえば外側は厚いとか、太いとか、細いとか、あるいは帯状のものならば、その巾方向に対する厚さがどのような傾向をもっているかなどは是非調査しておかなければならない。これは連続体の場合には、外周ならば切り取ることも容易であるし、経済的であるので、外周だけをサンプリングしてよいか否かというチェックになる。たとえば1巻内の変化がランダムであれば、あるいは外周の品質特性と平均値などとの間に管理されたカテゴリや相関関係があれば、外周だけをサンプリングしてよいことになる。また連続体の成分の変化は原料における成分のバラツキが引伸ばされてあらわれるのが通常であるから、原料のバラツキを調査管理すればよい場合が多い。一般にある試長のサンプルを1巻からいくつかとるときには、系統的サンプリングを行った方が平均値の精度はよくなる。

c) 1巻間のバラツキ 連続体の場合には、原料別、装置別、交替別などに層別すると、相続く1巻は比較的似ている場合が多い。したがって十分層別した合理的なロットを考えれば、巻間のバラツキ（級間分散）は級内分散に比し小さい。このときには副次サンプリングとしてロットから m 巻を1次単位としてランダムにえらび、その中から試長によるサンプルを n コとるときには、1巻を無駄にするという経費を考えると m を小さくし、 n を大きくあるいは集落として全部をサンプルとした方が平均値の精度がよくなる。

また工程によっては、相連続する2巻において、前巻の末端と、後巻の始端とは特性値が連続的に変化している場合がある。

合理的なロットや層別が行えず、各巻の平均値は大波、あるいは傾向をもって変化し、巻間分散が大きいときには、各巻をサンプリングするときに、ランダムサンプリングするよりは、系統的サンプリングした方が精度はよくなる。

d) ロット間のバラツキ c) にのべたように、合理的にとったロットについては、ロット内の各巻間のバラツキは小さいが、ロット間のバラ

ツキはこれに比し大きい。したがって、ロットの取り方が悪く、層別の仕方がわるいと、あるいは層別できぬときには、各巻間のバラツキは c) にのべたように小さくはならず、大きくなる。したがって連続体の場合には、特にロットの概念をハッキリさせて、製品を流しておかぬと、サンプリングも厄介になり、折角サンプリングしてもとめたデータにより action をとろうとしても、action が困難になる。

e) 試長並びにその数の決定 試長の決定はまずサンプリングの目的が平均値だけ知ればよいのか、あるいは平均品質のみでなく、ある長さについての特性値を保証することも目的としているかにより変わってくる。平均品質だけを精度よく推定したいためには § 1.5.2, § 8 にのべるように、試長間の分散とサンプルの数、その経済性を考えて決定すればよい。一般に連続体は波があるので、試長が大波の周期に近くなると、各試長間の分散は急速に小さくなる。したがって各種試長のサンプルをランダムにとり、その分散を検討するとその最小になるところの試長が大体波の周期に等しいという推定をすることもできる。また試長間の分散が、試長 l に比例して小さくなるようであれば、その連続体のバラツキは大体ランダムになっているとみなせる。

ある長さの製品についての品質あるいは平均品質を保証するという立場をとる場合には、試長をその長さにとらなければならない。そして平均品質を知るための試長の数は、そのバラツキにより、大波のあるときには系統的サンプリングの精度の向上ということを考えて、所要精度より決定される。

f) これらの品質特性を管理図により工程管理せんとするときには、サンプリング法と群のつくりかたをよく研究しなければならない。これについては § 9.1.1 参照。

また糸の強度などのように最小値により品質特性の決定されるものにおいて、強度を計量的に測定する場合は、上にのべたような考えかたでサンプルの数を決定すればよい。この場合 1 本 1 本別々に試験する場合には、その分布型より、その強度が管理状態にあれば比較的少ない数で最低品質を保証できる。一種の集合試料として数本一緒にして試験するような場合には、各糸に均等に力がかかる状態で試験しなければならぬという、試験

法自身の技術的問題もあるが、さらにたとえばはじめ各糸に均等に力がかかっても、もっとも弱い1本の糸から切れるという最小値の分布に関するような問題が起ってくる。この場合に1本の糸の最小値をもとめたりする統計学について研究が目下盛んに行われているが、ここでは省略する。しかし何本かの糸の強さというものを単位として保証するというのであれば、この測定値をこれまでのべたようにサンプル単位として統計的に取り扱えばよい。

ある試長についてある強度以下の部分が何%以下ということを保証する場合は、計数値の場合で、一定試長についての良、不良を、抜取検査法より考えてサンプルの数をきめればよいが、計量的に考えた方がサンプルの数は少なくてすむ。その不良率の推定は、§ 6・2にのべた単位体の計数値の場合になる。

7・5 同じ大きさの副ロットにわけられている集合体のサンプリング

吸入の確安、瓶詰の薬品類、貨車積の石炭、鉄鋼材などのように、ロットがいくつかの同じ大きさの副ロットにわけられている場合で、まず副ロットをとり、次にその副ロット内からインクリメントでサンプルをとるという、典型的な副次サンプリングの場合である。したがって既に本書の各所、§ 4・5、§ 5、§ 6・1・3などで詳細にのべているのでその章を参照していただければよい。

サンプリングを実施する際の技術的問題について気付いた点をのべておく。

1. インクリメントの決定 (§ 1・5・2, § 8 参照) この場合に1呎、1瓶の平均値を保証したり、推定したりするという要求と、すなわち副ロットについての知識と、ロット全体についての知識とを得たい場合がある。このときには必ずしも § 4・5にのべた m , \bar{n} のきめ方に従わず、全体の平均値の精度を少しおとしても、 \bar{n} を大きくしなければならぬことがある。たとえば $\bar{n}=1, 2$ と求められても、各副ロットの品質を推定する目的からいうと、計算より $\bar{n}=10$ となった場合には、たとえば $\bar{n}=4, \bar{n}=5$ というような妥協的な値をとることになる。またたとえば副ロットが N_f

このインクリメントから成り立っているときに、 $\sqrt{N_i}$ 、 $\sqrt{N_i}$ に n_i をとることもある。この場合は勿論各インクリメントを全部一緒にして混合試料にしてはまずいので、できれば各副ロット毎、あるいはいくつかの副ロットからのサンプルをよく考えて混合して、数コの混合試料、少なくとも2コの混合試料にわけて、副ロット間の分散を求められるようにしておかなければならない。そしてこれを群にとって、管理図にえがき、管理状態を検討しながら、規格値などと比較して行けばよい。全体の推定値があまり問題にならないようなときには、副ロットの数に関係なく、 $m=3\sim6$ くらいの副ロットをとり、各副ロット毎の混合試料を別々に分析して、これを管理図にかき、管理状態になっており、しかも規格値に合格していれば、これで十分である。

2. たとえば鉄板、鋼材などのように、**segregation** を起しやすいようなものでは、1枚のある成分の平均値でなく、すなわち副ロットの平均値でなく、インクリメント毎の特性値が問題となる場合がある。このような場合は、鉄板のどの辺に **segregation** が多いかということが技術的、あるいは統計的に判明する場合が多く、しかもその場所が大体一定しているような場合には、むしろ平均値を知るためにサンプリングするよりは、最悪値を推定するように、鉄板内を技術的に層別して、悪い層内からランダムにインクリメントをとって、各一枚につき別々に、あるいは混合試料をつくり、分析した方が合理的な場合もあろう。またこのように分離をおこなっている製品、別の例でいえば缶の内外、瓶の上下で品質特性が相当大きくバラツいているときに、平均値を精度よく推定したい場合には、副ロット内で層別サンプリングした方が有利である。

3. 合理的なロットをとること。いろいろな工程や商社からのロットが混合していると、副ロット内、副ロット間の分散が異なってくるので、工程などが管理されていないときは、十分層別したロットを対象としてサンプリングするように計画すべきである。工程が管理されてくれば、大きいロットに対して同数のサンプルで同じ精度で推定できるので有利である。

副ロットに分けられている場合は、多くの場合どこかの工程でこの作業が行われているのであるから、§5・2 にのべたように、その工程において、副ロット間分散、副ロット内分散が小さくなるように、集合体の取り

扱いに注意しておけばサンプリングは非常に容易になる。副ロットの品質を保証するという立場にたっても、工程管理と結びついて、はじめてサンプリングも容易になり、その意味もでてくる場合が多いことをここで再び強調しておきたい。

4. 従来多くの副ロットから成り立つロットを調査するときに、非常に労力をかけて、しかもよい加減の全数調査をやっている場合が多い。ロットさえ合理的にとっておけば、よい加減の全数調査よりも、管理図と結びつけてにらみながら精密なサンプリング調査を行った方が、遙かに信頼性があり、精度もよく経費も安く調査できることを銘記すべきである。

5. 鉄板、紙、フィルム、シートのような板状体からランダムサンプリングするためには、図7.7のようなサンプリングカードの活用がのぞましい。

第8章 粉塊混合物のサンプリング

8・1 まえがき

粉塊混合物のサンプリングは、粉塊混合物であるがための特質——粉と塊との分離、インクリメントによるサンプリング、縮分を必要とする——を除いては、これまでのべて来た考え方、統計的な手法すべて同じである。しかしこの特質のために、また工業原料として多量に取り扱われ、しかも噸当りの単価が比較的安いために（したがってサンプリングに噸当りあまり経費をかけることは意味がない）、サンプリングに最も多くの技術上の問題があり、またこれを経済的に行うという点に厄介な問題がある。これらの合理的なサンプリングは世界的に目下研究されている問題で、未だ決定的な解決はされていない。英国においては1930年代より、石炭について研究がすすめられ、多くの実験結果が報告されている¹⁾が、その他の粉塊混合物についてのサンプリング研究は、ほとんどない。従来からの研究と、私の考えを、技術的、統計的にのべてみよう。

粉塊混合物としては、石炭、コークスをはじめとし、鉄鉱石、硫化鉄などの鉄石類、石灰石、粘土、岩塩、カーバイド、肥料その他多くの工業原料、製品があるが、問題となる品質特性が異なるだけで、技術的、統計的な考え方はほとんどすべて同じであるので、ここでは主として石炭、コークスについてのべる。

これまでのべて来たように、ロット全体の平均値を知りたいためのサンプリングについてのべるが、各副ロットの平均値、あるいは工程中を洗っている場合の短期間の平均値をスナップサンプリングする場合でも、考え方は全く同じで、ロットの非常に小さい場合として考えて行けばよい。

粉塊混合物の分類

粉塊混合物でも、微粉状のものから、大塊までいろいろの状態があるので、これをサンプリングの面から見て分類すると次のようになる。

1) BS 403 (1930); 768 (1937); 1017 (1942); 735 (1944) など

1. 大きさが殆んど揃っていて、かつ1コずつサンプリングできる場合
 このようなことは、実際には殆んどなく、各塊により目方の大小がある。この目方のバラツキが特性値、たとえば灰分のバラツキに比し相当小さければ、各塊を単位体と考えて、通常の単位体のサンプリングと考えて行けばよい。また大きさが大体均一であれば、粉と塊とが分離を起すということもないから、サンプリングは非常に容易になる。50mm以上の塊コークスのみをサンプリングする、従来のコークス潰裂強度測定のためのサンプリングはこれに近い。

2. 各種粒度のものが混合している場合 ほとんどすべての場合がこれに属する。この場合には、粉と塊との分離という問題が起るのでサンプリングが最も厄介なもので、本章ではこれについてのべる。この混合状態もいろいろの場合があり、大塊中に少量の微粉のある場合、連続的な粒度分布をもつ場合、粉中に塊の点在する場合、しかも塊が300mm以上、100mm以上もあるような場合、粉炭のように最大粒度が10~15mm、あるいは3~5mmくらいの場合など、いろいろの場合がある。

3. 微粉の場合 最大粒の小さい場合で、2)と連続していて、どこから3)に属するかを最大粒度によりきめることはできず、むしろ最大粒度と最小粒度が分離した状態にあるか、否かにより、分離していない状態ならば3)として取り扱ってよい。この場合は流体に近い性質をもっているので、§7にのべた考え方と、§8にのべる考え方を一緒にしてサンプリング法を設計することになるが、一般に最大粒の小さい程サンプリングは容易である。微粉の場合、 $10^{-3}\text{cm}=0.01\text{mm}$ 以下で、水分が少なくサラサラしているときは、流体とみなしてよい。しかし水分が多くなると、水分の多い部分だけ固まって、粒状、塊状になり、2)に比較的似た性状を示すが、はじめから大きい粒子の場合より粒子間のバラツキは小さいのが普通である。

1) 2) 3) と分類はしたが、3)の流体とみなせる特別な場合を除いては、通常は以下にのべる考え方で共通にすんでよい。

8・2 バラツキをあたえる原因

われわれが粉塊混合物からサンプリングして、これを縮分し、分析し、

データをうるまでのバラツキを大きくわけて考えると、

- 1) ロット自身のバラツキ
- 2) 大口試料 (1次試料) サンプリングによるバラツキ
- 3) 分析用試料 (2次試料) のサンプリング, すなわち縮分によるバラツキ
- 4) 分析によるバラツキ (分析用試料から分析試料採取のバラツキを含む)

これらのバラツキが一緒になって、データのバラツキとなる。これらのバラツキ, すなわち誤差を考えると、常に考えねばならぬことは、§ 3 にのべた、信頼性、カタヨリ (正確さ)、精度である。

信頼性は主として、技術を統計的に管理する問題である。

カタヨリは、粉塊混合物のサンプリングにおいて、粉塊の分離と水分等の蒸発によるもので、大口試料採取や、縮分において問題が多い。このカタヨリは、統計的な検討を要することは勿論であるが、主として技術的に解決して行かなければならぬことが多い。

精度は、分散の加成性より次のようにあらわされる。

$$\sigma_x^2 = \sigma_p^2 + \sigma_i^2 + \sigma_R^2 + \sigma_M^2 + \sigma_e^2 \quad (8.1)$$

ただし σ_x : 分析して得た1つのデータの精度 σ_p : ロット内の精度

σ_i : 均一な層からのインクリメントサンプリングの精度

σ_R : 縮分による分析用試料サンプリングの精度 σ_M : 分析の精度

σ_e : ロット, 大口試料サンプリング, 分析用試料サンプリング分析などが互に独立でない, すなわち相関関係をもつときの補正のための項。ただし $\sigma_e^2 \geq 0$

通常は相関がないか, あるいはあっても非常に小さい場合が多いので, $\sigma_e \approx 0$ とみなせる。

σ_p は、出荷先の管理状態、その後の取り扱いによって変化するサンプリング法設計に大きな影響をあたえるバラツキであり、これと σ_i とが一緒になって大口試料採取の精度 σ_e となる。したがってわれわれの得た1つのデータは、そのロットの母平均に対して

1) たとえば 100 g の 60 メッシュ以下の試料をつくり、これから分析するために、1 g の試料をとるといのが普通である。ここでは前者を分析用試料 (2次試料)、後者を分析試料と区別した。

$$\sigma_x^2 = \sigma_s^2 + \sigma_R^2 + \sigma_M^2 \quad (8.2)$$

この σ_x は、大口試料をとり、それを縮分して1つの分析用試料をとり、それを1回分析したときに得た1つのデータの精度である。従来われわれが行っていたように、分析を同じ分析用試料について2回行っているときは、その平均値のバラツキ $\sigma_{\bar{x}}$ は

$$\sigma_{\bar{x}}^2 = \sigma_s^2 + \sigma_R^2 + \frac{\sigma_M^2}{2} \quad (8.3)$$

粉塊混合物のサンプリングにおいては、一般に σ_s が一番大きく、 σ_R がこれにつき、 σ_M が最小となっている場合が多い。しかし σ_R が大きい場合も相当あり、分析法の確立していない、あるいは管理されていない特殊な測定では、 σ_M が大きくなっていることもある。

§3 にのべたように¹⁾、われわれがえたデータの精度を向上させるためには、これら3つの精度の大きさを相対的に決定して行かなければならない。

すなわち $\sigma_{\bar{x}}$ は、これら3つの中、ほとんど最大のバラツキをもつものにより支配されるから、他はその分散として1/10～1/20、標準偏差として1/3～1/5 くらいに押えておけば十分である。もし現行法がそれ以下に小さくなっているときには、たとえば分析の精度が大口試料の σ_s の1/10 という場合には、もっと精度を悪くして、簡易分析を行い、時間を節約した方が有利である。

合理的なサンプリングを達成するには、えたデータの精度をある信頼度で目的とする巾以内におさえて、これらのバラツキを管理状態とし、カタヨリのない状態で、経済的にデータが得られるように設計し実行することである。このために、これらのバラツキをあたえる原因を考えてみよう。

- 1) ロットの生産管理状況
- 2) ロットの輸送中荷下し法、貯蔵法、使用法その他の取扱
- 3) 各粒間の特性値の分布状況
 - i) 粒度による特性値のカタヨリ
 - ii) 同一粒度の粒子間における特性値のバラツキ
- 4) ロット内の粒度分布のカタヨリ

1) §3.2 参照.

- 5) ロットをいくつかの副ロットに層別したときに、その層間および層内のバラツキ、あるいはプロセスとして見たときの特性値の変化状況
- 6) ロット内の水分分布のカタヨリ
- 7) 大口試料採取に際しては
 - i) ロットの量 (重量)
 - ii) 大口試料の量 (重量) $W = nw_i$
 - a) インクリメントの大きさ (重量) w_i
 - b) インクリメントの数 n
 - iii) インクリメントの取り方
- 8) 縮分に際しては
 - i) 大口試料の量 (重量)
 - ii) 粉砕法
 - iii) 粉砕粒度
 - iv) 縮分法並びに回数
 - v) 2次試料 (分析用試料) の量 (重量)
- 9) 分析に際しては
 - i) 2次試料から分析試料の取り方
 - ii) 分析法, 分析者, 分析条件, 分析回数

以上 1)~6) ロットに関する問題, 7) 大口試料採取に関する問題, 8) 二次試料採取に関する問題, 9) 分析に関する問題とわけて話をすすめて行こう。

8.3 サンプルングの面から見た粉塊混合物

粉塊混合物のロットをサンプルングの対象と考えた場合にもっている特徴を十分に検討しておくことは、後に大口試料、2次試料採取法を考える基礎となるので、まずロットについて考えよう。

8.3.1 ロットの管理状態

われわれの対象とするロットは § 5 にのべたように、すべてある工程からの生産品である。たとえば採炭、選炭という工程を通った製品である。さらに工程という言葉を広義に解釈すれば、出荷状況、港頭における取扱、輸送中の状況、工場などへ到着後の荷下し法、運搬法、貯蔵法、使用法なども1つの工程である。われわれはこれらの工程を通して来た製品である石炭をサンプルングするのである¹⁾。したがってこれらの工程の管理状態に

1) これらの取り扱い方がサンプルングをする目的と関連のあることに注意。

より、ロットの性質——ロット内のバラツキとロット間のバラツキ——は大きく変化する。ロットをいくつかの副ロットにわけて考えたときに、たとえば貨車を副ロットと考え、数車が入荷したときに、貨車間のバラツキと貨車内のバラツキは、選炭状況、積荷状況、輸送状況により大きく変化する。これらのロット内のバラツキは、粒度分布のカタヨリと、同じ粒度間のバラツキにより生ずる。船積の場合でも、同様である。この副ロット間 (σ_0)、副ロット内のバラツキ (σ_w)、あるいは次々と船から荷下しされるときの連続的なプロセスとしての特性値の変化は、大口試料採取の際にも、またこれを原料として使用する工程の管理、品質管理の面にも、非常に大きな影響をあたえる。これらの原材料の品質の推定あるいは保証のためのサンプリングも、工程においてこのロットをいかに利用するかということを考えなければならぬので、工程管理のためのサンプリングと兼用している場合が多い。ロットが管理状態でバラツキが小さく均一になっているということは、サンプリングを容易にし、工程管理を容易にする。このことは、不均一になりやすい、多量に原料として取り扱われる粉塊混合物の場合には、特に重要な問題である。

したがって、粉塊混合物を購入の契約書には、その平均値のみならず、できればバラツキ、たとえば貨車間のバラツキ、管理状態も明記しておくべきであろう。また入荷後のロットの取扱も、折角ロットの平均値を精度よく推定するのであるから、そのデータがそのまま利用できるように、§5・2にのべたような層積法 (bedding) または層別貯炭などにより管理すべきである。最近某製鉄所におけるサンプリング法研究の大きな副産物として明らかになったことは、同じハッチから荷下した輸入鉄鉱石に、場所により鉄分50%弱の部分から、70%弱の部分まであり、平均鉄分60%というバラツキの大きいロットがあった。従来焙鉄炉ではこれをすべて60%とみなして操業していた矛盾が判明した例がある。輸入石炭についても同様の例がある。最近建設される某製鉄所においても、層積法の工夫が払われ、また硫安工場でも硫安の堆積に層積法を用いはじめた、ということで、誠によい傾向であると思う。

以上のようにある1つのロットが管理されていることが必要であるが、これと共に次々と入ってくるロットが管理されていないと、§5にのべ

た通り合理的なサンプルング法的设计が困難になる。

8.3.2 各粒間の特性値の分布

粉塊混合物は、いろいろの粒度の多くの粒子の混合体である。この各粒子の間にバラツキがあるので、ロット内にバラツキを生じ、サンプルングは誤差を伴うのである。いまこのバラツキを 1) 粒度による特性値のカタヨリ、2) 同一粒度の各粒子間の特性値のバラツキ、の2つにわけて比較的管理された均一な層について、すなわち σ_x について話をすすめる。

1. 粒度による特性値のカタヨリ 粉塊混合物を篩分けすると、いくつかの粒度のものにわけられるが、この各粒度ごとの特性の平均値が各々異なっている。すなわちカタヨリがある。一般に粉塊混合物では、このカタヨリは必ずあるもので、図 8.1 の A、B 曲線に示すように、傾向として最大粒度、および最小粒

度においてカタヨリが著しい。たとえば、石炭の灰分においても、例外はあるが、一般に最大粒度および最小粒度、あるいはいずれか一方の、灰分が多い。また鉄鉱石においては、一般に最小粒度の鉄分の少ない場合が多い。中間粒度のものは、カタヨリはたいして大きくないのが普通であるが例外的に C 曲線のようなものもあるかも知れない。

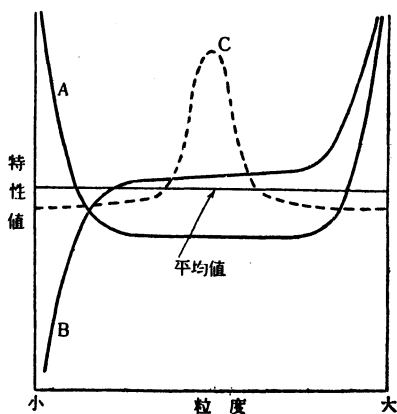


図 8.1 粒度とその平均品質の関係

このようにどんな曲線をとるにしても、粒度によりカタヨリがあるので、サンプル中の粒度分布が、ロットの粒度分布とほぼ等しくなるように、少なくとも最大粒度と最小粒度はロットと大体同じ割合にサンプルングすることが必要である。ところが特に最大粒度と最小粒度のものは分離を起してロット内でカタヨリを生じやすい。もしこの粒度の平均品質のバラツキがなく、いずれの粒度もほぼ同じ平均品質間をもつならば、サンプルング

は非常に容易になる。

2. 同一粒度の各粒子間における特性値のバラツキ 同一粒度の石炭でも、その1粒ずつの特性値の間にはバラツキをもっている。このバラツキは、1粒内の純物質と不純物質の量により支配される。たとえば石炭の1粒は、固有灰分をもった純炭分と外来灰分である岩石類との混合体である。したがって純炭分が均一であっても、外来灰分の混入量により、各粒子間の灰分分布は大体ポアソン分布、2項分布に近い形をとる。

この分布は、a) 粒度と、b) 平均品位すなわち分布の平均値によりかわってくる。

a) 粒度とそのバラツキ 同一粒度の各粒子間のバラツキは、粒度により変化する。たとえば大きい粒度のものは、重量も大きいので、2項分布 $\sigma^2 = \bar{p}\bar{q}/n$ (\bar{p} 純炭分の平均含有率、 \bar{q} ボタの平均含有率) において、 n が大きい場合の平均値をとる形になり、分散は小さくなり、次第に正規分布に近い分布をとる。 n が小さく、すなわち粒度が小さくなると分散は大きくなり、次第に歪んだ分布になる。極端に小さな粒子になると、純炭分と外来灰分とが粉碎により分離して、純炭分に近いところと、外来灰分に近いところに2つの山のある2峰分布になる。しかし実際には粒度による選炭の難易により、分布型が少しかわってくる。たとえば微粉のみを上手に浮游選炭すれば、粒度微小の図8・2の右側の山はほとんどなくなり、選炭がわるければ、塊でも二峰分布になるであろう。

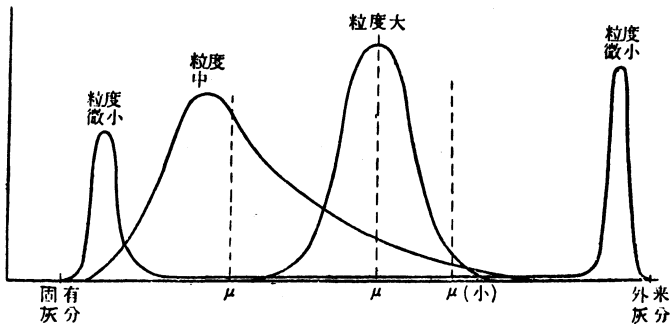


図8・2 各粒度の1粒間のバラツキ (一例)

b) 平均品位とバラツキ 粒度により n がかわりバラツキが変化したが、2項分布は p の変化によっても分布型がかわる。この p に相当するものが平均品位である。すなわち平均灰分含有量、平均鉄分、平均窒素分、平均硫黄分などにより分布型が変化してくる。この関係を図示すると大体図8.3のような形をとる。

(A) 平均灰分の少ない場合には、灰分の多い方へ、(F) ポタのように平均灰分の多い場合には、灰分の少ない方へスソを引いた分布になり、(B~E) その中間では正規分布に近い形となる。しかも精炭では平均灰分の少ない程、(C→A) 分布のバラツキは小さくなり、ポタでは平均灰分の多い程 (F→D) バラツキは小さくなる。この場合も $p < 0.1$ となればポアソン分布と見なしてよい。

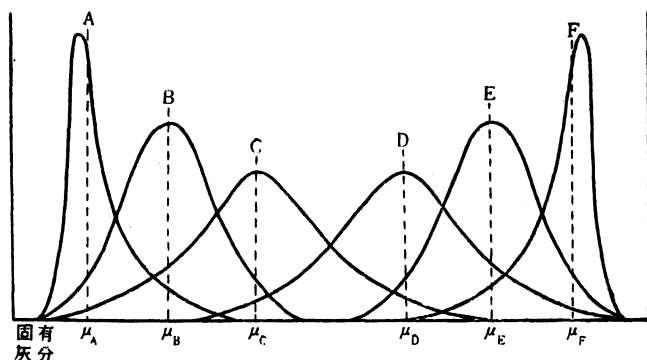


図 8.3 平均灰分と分布型、バラツキの大きさとの関係

実際には純炭分が必ずしも均一でないため、また選炭方式により分布型がやや変化することがある。インクリメントでサンプリングしたときにも各インクリメントの分布をとると、大体上にのべたような傾向となる。

また鉄鉱石、硫化鉄、硫安などについても同様で、たとえば鉄鉱石では鉄分の分布は、鉄の酸化物の理論量を最高値として、平均鉄分のよい鉱石ではFのような分布をし、平均品位が下り、粘土、岩石などの混入が多くなると、E、D、の型となる。硫安の場合も硫酸アンモニアの理論窒素量

よりも平均窒素分が下り、水分などが多くなればやはり $F \rightarrow E \rightarrow D$ という傾向をたどる。これは純物質と不純物との混合体と見られる場合にすべてあてはまり、一般に平均品位の悪くなる程、サンプル数を増加しなければならぬという重要な性質を示している。近似的に

$$\frac{\text{標準偏差}}{|\text{純物質品位}-\text{実際の品位}|} \approx \text{一定} \quad (8.4)$$

といわれている。石炭の場合には〔標準偏差/(全灰分-固有灰分)〕、硫化鉄の場合には〔標準偏差/(理論硫黄分-全硫黄分)〕、硫安の場合には〔標準偏差/(理論窒素分-全窒素分)〕という関係が、実際のわれわれの対象とする範囲(図 8.3 における $A \rightarrow B \rightarrow C$, または $F \rightarrow E \rightarrow D$ の範囲)のロットについて大体成立する。また〔標準偏差/ $\sqrt{\text{全灰分}-\text{固有灰分}}$ 〕 \approx 一定となる場合があるともいわれている(図 8.7 参照)。

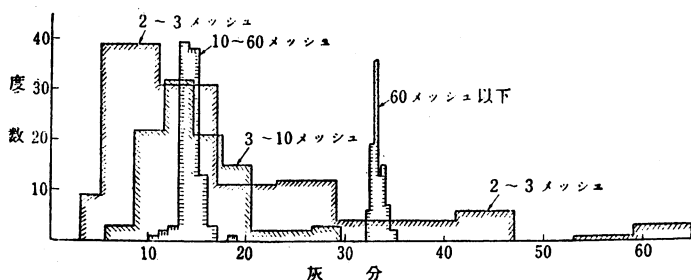
c) 粒度とそのサンプルに及ぼす影響 粒度とその平均灰分量により、各粒度間のバラツキや分布型の変化する状況は a) b) に示した通りであるが、実際にサンプル中にこれらの粒子が入ったときの影響を考えなければならぬ。

一定量のサンプルをとったときに、その中に入ってくる各粒の数 n_i は、大塊はきわめて少なく、微粉はきわめて多い。したがって図 8.2 に示したように各粒毎のバラツキは微粉の方が大きい。サンプル中に入るその数がきわめて多いので、その平均値の分布はほとんどバラツキのない正規分布となる。これに対し大塊の方は、各塊間のバラツキは粉よりも小さいことが多いが、数が少ないために、その平均値のバラツキは大きい。

$$\left(\because \sigma_{\bar{x}_i}^2 = \frac{\sigma_{n_i}^2}{n_i} \right)$$

この関係を実験的にもとめた図 8.4 を見れば明らかであろう。この図はコークス原料炭の A 洗粉炭である。これを篩分けして、各粒度毎に大きさ約 1g のインクリメントでランダムサンプリングしてえたデータのヒストグラムである。

どの粒度の 1g がサンプル中にとられるかにより、そのカタヨリも、バラツキの影響もよくわかる。すなわち同じ 1g がサンプル中にランダムにとられても、大きい粒度ほどサンプルにバラツキをあたえる影響が大きい



メッシュ	n	\bar{x}	s	最大値	最小値	備 考
~2	20	23.11	15.3	61.8	6.4	各 1 コ
2~3	120	17.35	13.3	64.7	3.0	〃
3~10	100	14.56	4.4	28.4	5.9	各 1g インクリメント
10~60	100	14.13	1.0	18.3	10.2	〃
60~	100	33.05	0.61	35.0	32.0	〃

図 8・4 粒度と 1g インクリメントのパラツキとカタヨリ (A 洗粉炭)

ことがはっきりしている。したがって一般に粉塊混合物のサンプルングでは、そのロット中の最大粒度に着目しなくてはならないのである。

またこの図より、粒度間のカタヨリの状況もよくわかるであろう。

8・3・3 粒度分布

各粒度間および各粒度内の品質がバラついても、これらの粒子がロット内でランダムに分布していればサンプルングも比較的容易であるが、粉塊混合物では、その粒度分布がロット内においてランダムになっていない、必ずカタヨリをもっている。このために粉塊混合物のサンプルングは非常に厄介となる。

いくら均一に混合しても、その各部分はロットの平均粒度分布からのランダムサンプルとは見なせない。粒度による分離 (segregation) カタヨリが起っている場合が多い。われわれはこのような部分からサンプルングして、ロットの粒度分布に近いサンプルをとろうというのである。

たとえば山積にした石炭について考えれば、粒度の細かいものは塊の間を

通って下方へころがり落ちやすい。また船積、貨車積、バケットコンベヤー中のものは、輸送中の振動や衝撃などにより、粉や小粒が下方に落ちたり或いは積込の際に転落などして、塊が下方に落ちて多くなることもある。一般に粒度差の大きくなる程、特に最大粒の大きくなる程、また乾燥している程この分離ははげしくなる。したがって最大粒が大きくなる程、乾燥炭程、サンプリング法に技術的工夫や注意が必要となる。山積、船積、貨車積のように静止状態にあるものは、粉塊の分布に必ずカタヨリがあり、しかも全体からランダムサンプリングすることが不可能となるから、「粉塊混合物の合理的なサンプリングは静止中に行うことは困難である。原則として移動中、すなわちロットの積みかえ中など動いているときにサンプリングすること」という原則となる。着荷したロット、生産したロットは必ず他へ移動する筈だから、そこでサンプリングするように工夫するのが最もよい。しかし粉炭のように、比較的粒度の小さい場合には、そのロットの取り扱い方がうまく行われていれば、静止中のロット、たとえば貨車積のロットからでもカタヨリのないサンプリングを行うことができる。

粒度分布のカタヨリという点からいうと、ロットを篩分けして、各粒度ごとに層別して、その中から層別比例サンプリングを行えばよいが、これは実際的でない。しかし1個で5 kg も 10 kg 以上もあるような大塊の入っているときには、その大塊は篩分けして、層別サンプリング、あるいはこれを粉砕してもとのロットにもどしてサンプリングした方がよい場合がある。この場合も管理という面からいうと、このような大塊のつこないように、購買部、生産者、契約書などに action すべきであろう。

従来は、ロットの粒度分布を目で見て、大体その割合にサンプリングすることにより、粒度分布のカタヨリを避けようとしていた場合が多いが、目で見た粉塊の割合は、実際に人により異なっており、カタヨリが入りやすい。いろいろのロットにつき、目で見た判定と実際に篩分けして、目の訓練をするということも考えられるが、これは非常に困難であろう。むしろ粒度分布について、サンプル採取者が注意なくとも、ロットに近い品質をサンプリングによってもとめられるように、サンプリング法をきめるべきである。

8・3・4 水分分布

ロットについて、もう1つ非常に厄介な問題は水分分布のカタヨリである。粉塊混合物は一般に屋外に貯蔵、無蓋で輸送される場合が多く、天候により水分が大きく支配される。水分が粉塊混合物の固有水分以下の場合に、輸送中の乾燥や降雪、降雨の影響がなければ、大体各粒度に特有な水分分布（たとえば粉になる程水分量が多い、非常に乾燥している場合は粉の方が少ない等々）であるから、水分をそう気にする必要はないが、このような保証のえられることは少ない。

したがって、殆んどすべての場合、下方に、中心部に、上方に水分が多いというカタヨリをもっている。

しかるに水分は多くの場合、原材料である混合物の受渡の量に関係してくる。すなわちある含水重量、あるいは乾燥重量を、この測定水分より求めて、代金の受払や、原単位の計算の基礎とする重要な量である。従来水分換算をやっていない契約書や検査規格がよくあるが、これは技術的に考えて見てもおかしい。

水分も1つの品質特性であるから、その分布のバラツキやカタヨリは、§8・3・2 にのべたように、水分の非常に少ないとき、あるいは極端に多く全体が完全に濡れているようなときには、比較的小さい。しかし非常に濡れていて、秤量後に水分が多量流出するような場合は、秤量直後にサンプリングしなければ、水分用サンプリングの目的からいっても不合理である。

水分分布に大きなバラツキやカタヨリのある場合には、静止中のロットから水分測定用のサンプルを合理的にとることはできないという原則となり、移動中にとらなければならない。また浸水などにより、船底などが特に濡れているときには、その部分を層別してサンプリングすることも考えられる。この場合秤量が水を含んだまま行われていれば、最後の水もサンプリングする必要がある。

また水分のサンプリングについては、後にものべるが、サンプルの水分が測定するまでに蒸発などして変化することが多いので、サンプルの保存法や処理法に十分な技術的注意が必要である。

8・4 大口試料（1次試料）の集め方

大口試料をサンプリングする際に考えなければならぬことは、

- 1) 目的とする品質特性と所要精度
- 2) ロットの状況
- 3) ロットの量
- 4) 大口試料の量=(インクリメントの大きさ)×(インクリメントの数)
- 5) インクリメントの取り方

である。

目的とする品質特性によりサンプリング法が全くかわってくることがある。たとえば、水分測定用のサンプリングと、灰分、純度などを目的とするサンプリング法のことなる場合がある。またコークスのロットを熔鉱炉の原料としてのサンプリング法と、コークス炉管理用のサンプリング法とでは当然異なってくる。また石炭においても灰分の推定のためのサンプリングと、揮発分測定用のサンプリング法とは異なる。しかし石炭の工業分析のように、同じサンプルから2つ以上の品質特性を測定するようときには、一般にバラツキの大きい方あるいは所要精度の高い方の特性が所要精度以内になるようにサンプリングすれば、他の特性についての精度は十分である。石炭では、一般に灰分のバラツキがもっとも大きいから灰分を対象としてサンプリング法の研究を行う。コークスの場合には水分のバラツキが最も大きいのが普通である。しかし最近の製鉄コークスのように、平均水分が低く2～3%になっているときには、灰分などのバラツキが最も大きいこともある。

(8・3)式であらわしたような所要精度 σ_x を達成するためには、 σ_s 、 σ_R 、 σ_M を管理する必要がある。 σ_R 、 σ_M については後にのべるが、一般に σ_R 、 σ_M は縮分法、測定法がきまり、管理されてくれば大体その大きさを推定できる。したがって

$$\sigma_s^2 = \sigma_x^2 - (\hat{\sigma}_R^2 + \hat{\sigma}_M^2)$$

より、 σ_s をどの程度に押えたらよいか推定することができる。この σ_s になるようなサンプリング法は、後にのべるようにインクリメントを決定し

て、§4, §5, §6 でのべたようないろいろの統計的な手法や考え方を利用して、インクリメントをどこから、どのようにとるかを決定すればよい。この決定が大口試料の集め方になる。

・ σ_p すなわち σ_b , σ_w の決定にはロットの状況についての予備知識が是非必要であるが、これについては §5, §8・3 にのべたのでここでは主として σ_i についてのべる。

8・4・1 インクリメントの大きささと数

粉塊混合物は1粒ずつサンプリングすることは不可能に近いので、通常スコップなどでサンプリングしている。まえにものべたように¹⁾、1インクリメントとは「スコップなどで、1回に、すなわち1回の運動で1杯とったサンプル」のことである。小さいスコップで何回か集めて、大型スコップに1杯とったときは、その大型スコップ1杯が1インクリメントではなく、小さいスコップ1杯が1インクリメントであり、何回かとってスコップに1杯、以に1杯集めたものは1インクリメントとはいえない。粉塊混合物ではこのインクリメントをサンプル単位とする。

このようにしてインクリメントをいくつか集めた混合試料を大口試料とするが、これには、1) インクリメントの大きさ（1杯にとる重量、あるいはスコップの大きさ）、2) インクリメントの数、3) インクリメントの取り方、の3者を決定すればよい。次にインクリメントの大きささと数をいかにして決めるかをのべるが、まず管理されたロット、すなわち $\sigma_p \approx 0$ のロットについてインクリメント向のバラッキ σ_i について話をすすめよう。

1. ロットの平均品質とインクリメントの分布 一般に平均品質の変化により、ランダムサンプリングされたインクリメントの分布の分散、標準偏差は §8・3・2 でのべた図8・3と同じ傾向で変化する。すなわちロットの平均品質が悪くなるとインクリメント間のバラッキは大きくなる。勿論インクリメントでサンプリングしているので、1インクリメントの品質は、その中に含まれる多くの粒子の平均品質をとることになるから、そのバラッキは小さくなり、純度のよい場合でも歪みかたは少なく、正規分布に近づく。

1) §1-5-2 参照。

たとえば全灰分A10%，B15%の2種の石炭があるときに，固有灰分が3%とすると， $15-3 > 10-3$ すなわちAの方がインクリメント間の分散は小さくなる。

硫化鉄であれば，硫黄の理論含有量よりも少なくなれば，コークスは水分が多くなれば，ある値までは，インクリメント間のバラツキが大きくなる。また粒度に比しインクリメントの大きさが大きいときには，その分布型は正規分布に近づく。

したがって σ_s を一定にするためには，ロットがきまれば σ_p は一定であるから，

$$\sigma_s^2 \approx \frac{\sigma_p^2 + \sigma_i^2}{n}$$

ただし σ_i^2 : 均一な層でのインクリメント間の分散 n : インクリメントの数

ロットがきまれば σ_p は一定としても一般に平均品質が悪くなると， n を大きくしなければならない。すなわち大口試料の量が多くなる。

2. 粒度とインクリメントの大きさ 最大粒度が大きくなると，インクリメントはカタヨリを生じやすくなり，またバラツキも大きくなる。

インクリメントをいくらランダムにとっても，インクリメントの大きさが小さくなり，最大粒子がインクリメント中に入りにくくなると，インクリメントを集めた混合試料中の粒度分布は，ロットの粒度分布に比し大きい粒子の割合が少ないというカタヨリが入る。従ってインクリメントの大きさは，最大粒度が十分ランダムに入りうる程度に大きくしなければならない。一般にインクリメントのとり方により変化するが，インクリメントの大きさ（容量）を最大粒子の100倍以上，少なくとも10倍以上にしないとカタヨリが入りやすい。

同じロットについて，すなわち平均品質も，最大粒度も同じ場合には，§5・3・1にのべたように，インクリメント間の精度 σ_i は，インクリメントの大きさ w_i が大きくなると，小さくなる。しかし直線的に小さくなるわけではなく 図8・5のような漸近線になるからあまり w_i を大きくしても不経済である。

この曲線において，モード（山の極大となっているところ）よりも左側で， σ_i が小さくなっているのは，最大粒子がインクリメントに入らなく

なったために、バラツキが小さくなったので、これ以下では明らかにカタヨリが入っていることを示す。

σ_i と w_i とはモードより右側の部分では $\log\text{-}\log$ をとると直線関係になるともいわれている。しかし粉塊混合物の性質や状態により変化するものであるから、一応目的とするいろいろの混合物について、実験的に求めてみる必要がある。

図 8.5 には数本の曲線が写っているが、これは最大粒度が変化した場合の曲線の傾向を示したもので、前述のように、最大粒度が大きくなると曲線は σ_i の大きくなる右方へずれて行く。

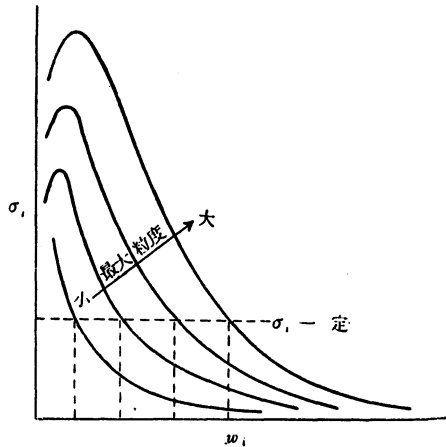


図 8.5 インクリメントの大きさ (w_i) と インクリメント間のバラツキ (σ_i)

いま粒度に関係なく、 σ_s 、したがって σ_i および n を一定にするためには、この曲線群を図のように横軸に水平な線で切り、各粒度に応じた w_i をとるようにすればよい。粒度に応じて w_i をきめておけば、すなわちスコップをきめておけば一定の大口試料の精度 σ_s をうるためには、 n を一定にきめることができる。石炭のサンプリングにおいて ASTM も BS もこのように最小のインクリメントの大きさ w_i を決定している。表 8.1 にその数値を示す。

注意 1) インクリメントの大きさを一定にしてとるということが実際上は問題になる。各インクリメント毎に正味 500 g, 1 kg あるか否かを一々チェックしながらサンプリングすることは不可能である。したがってインクリメントの大きさ (重量) が実際作業として一定になるような工夫が必要である。これには、インクリメントをとる道具の工夫や研

表 8・1 最大粒度とインクリメントの大きさ
A) ASTM (D 492-48)

粒度 (in)	1)	1)	1)	2)	3)	3)	4)	5)
	5/8以下	5/8~1 ¹ / ₄	1 ¹ / ₄ ~2	2~6	2~4	4~6	A	B
最小のインクリメントの大きさ (lb)	2	4	6	10	10	15	30	20

注) 1) 最大粒度以下の石炭および表の数値の間に篩分けされた石炭 2) 2吋から6吋までに篩分けされた石炭 3) 最大粒度2~4吋, 4~6吋以下の石炭 4) 最大粒度6吋以上, 最小粒度5/8吋以上の石炭, 塊炭 5) 最大粒度6吋以上の石炭, 脈炭もこれに属す. 最大粒度6吋以上で最小粒度5/8吋以下で篩分けした石炭, 塊炭

B) BS (1017)

最大粒度 ¹⁾ (in)	塊 ²⁾	3	2	1 ¹ / ₂	1	1/2
最小のインクリメントの大きさ (lb)	10	6	4	3	2	1

1) ある粒度の篩で篩分けしたときに残量5%以下の粒度をいう. 2) 10吋度近くもあるような大きい塊炭は, 層別して別にサンプリングすること.

究が必要であり, 一般に各大きさのインクリメントをとるに適した特別のスコップを作成すると便利である. たとえば 図 8・6 のような底の平

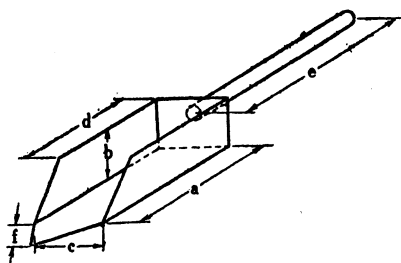


図 8・6 インクリメント採取用スコップの例

すい形の研究が必要である.

らな, 両壁が比較的高く, 垂直に立った, 必要に応じて先端のとがった¹⁾スコップがよい. 両壁の高さや, 角度, 先端の形, その他 a, b, c, d, e の長さなど, ロットの状態, 特に粒度, 水分, 採取場所により, バイアスのないようなインクリメントがとれる, 使いや

1) 混合物中にまじこむためには, 先をある程度とがらしておいた方がよい.

また自動試料採取器を設計する場合も、バイアスの入らぬ、一定量ずつとれるという条件を満足するように考えた方がよい。

2) 注意 1) にのべたような工夫をしないで、通常のスcoopなどでサンプリングすると、インクリメント間の重量のバラツキが大きくなり、これが特性値のバラツキに比して相対的に無視できなくなる場合がある。また大粒が落ちやすくなり、カタヨリが入る。インクリメントが混合試料にあたる影響 y_i は

$$y_i = w_i x_i \quad (8.5)$$

ただし x_i : 1 インクリメント内の特性値の平均値。 w_i : 1 インクリメントの重さ

多くの場合 w_i と x_i とは相関関係なく変化するから

$$V_{y_i} = \left(\frac{\partial y_i}{\partial w_i} \right)^2 V_{w_i} + \left(\frac{\partial y_i}{\partial x_i} \right)^2 V_{x_i} = x_i^2 V_{w_i} + w_i^2 V_{x_i} \quad (8.6)$$

たとえば、インクリメントの大きさの平均 $\mu_{w_i} = 1 \text{ kg}$ 、標準偏差 $\sigma_{w_i} = 50 \text{ g}$ 、ロットの平均特性 $\mu_{x_i} = 8\%$ 、インクリメントの特性 x_i 間の標準偏差 $\sigma_{x_i} = 2\%$ とすると

$$V_{y_i} = (8)^2 \times (50)^2 + (1,000)^2 \times (2)^2 = 1.6 \times 10^6 + 4 \times 10^6$$

となり、 w_i のバラツキによる影響は無視できるが、

$$\mu_{w_i} = 1 \text{ kg}, \sigma_{w_i} = 100 \text{ g}, \mu_{x_i} = 20\%, \sigma_{x_i} = 3\%$$

というときには

$$V_{y_i} = (20)^2 \times (100)^2 + (1,000)^2 \times 3^2 = 4 \times 10^6 + 9 \times 10^6$$

$$\text{または } \sigma_{y_i}^2 = \sigma_{w_i}^2 + \sigma_{x_i}^2 = (0.1)^2 + (0.15)^2 \quad (8.7)$$

となり無視出来ない。一般に、インクリメントの特性値の変動係数に対し、インクリメントの大きさの変動係数が $1/3 \sim 1/5$ 位になるように、インクリメント採取器の設計ならびに採取法の管理をしておけばよい。しかし実際にはこれも困難な場合には、混合試料としての実際の精度をもとめた方がよい。

3. インクリメントの大きさと数の決定 インクリメントの大きさと数を決定するには、ロットの状況 σ_p 、ロットの大きさ(量)、サンプリング法も考えなければならぬが、ロットの状況がランダムに変化しているとき

には、所要精度 σ_s とその信頼度 $(1-\alpha)$ 、最大粒度、平均品質特性によりインクリメントの大きさ w_i と数 n_i は決定される。§ 8.4.1 の 1, 2 にのべたように次のような手順で決定すると実用的に便利である。

- 1) 図 8.5 に示したように、粒度により w_i を決定する。
- 2) 図 8.7 に示すように、同じ w_i では平均品質特性により σ_i が変化するのでその σ_i を用いて次式より n を決定する。

$$\sigma_s = u_\alpha \sqrt{\frac{\sigma_p^2 + \sigma_i^2}{n}} \quad (8.8)$$

ただし u_α : 信頼度 $(1-\alpha)$ によりきまる係数

実際には σ_p が相当大きいので § 4 にのべたように層別サンプリングその他を行つて、 σ_p を小さくするための工夫を行う。均一な層においては $\sigma_p = 0$

注意 σ_i は w_i により変化するが、 σ_p は w_i によりかわらぬ。

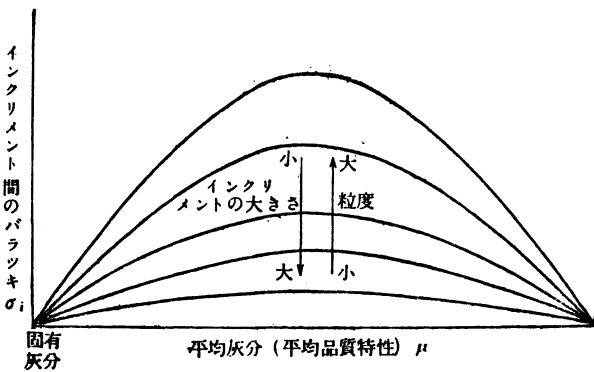


図 8.7 平均品質とインクリメント間のバラツキ

通常われわれが接するロットは、左右の端の部分で、大体直線関係がある。したがって $\sigma_i / (\text{平均灰分} - \text{固有灰分})$ または $\sigma_i / \sqrt{\text{平均灰分} - \text{固有灰分}}$ が一定という関係が成立する場合が多い。

BS および ASTM の石炭のインクリメントの大きさや数の決定もこの方式にしたがっている。次に参考にその表をかかげる。

表 8.2 インクリメントの数 n の最小値
(中塊炭以下の場合) (英国規格)

灰分 精度 x'	0~6	6~10	10~15	15~20
±0.25	167	376	667	1043
±0.50	42	94	167	261
±0.75	19	42	74	116
±1.00	11	24	42	65

注 1) 精度の信頼度は99%. 2) インクリメントの大きさ w_i は表 8.1 B により決定. 3) 大口試料最低採取量は $w_i \times n$ よりもとめられる.
4) 固有灰分 4% の場合の表である (比重 1.6 より軽い部分¹⁾).
5) 本表よりもとめた n コ以上のインクリメントをとれば, 表の精度よりよい精度でロットの平均値を推定できる. 6) 貨車からのサンプリングの際は表 8.5 及び 6 による.

表 8.3 塊炭の大口試料最小量 (英国規格)
インクリメントの大きさ $w_i = 10 \text{ lb}$ 精度 ±1.0% (信頼度 99%)

全灰分 % x'	インクリメントの数 n			
	最小粒度			篩分けされていない場合 原炭
	3 in	1½ in	1/2 in	
4% 以下	10	10	10	—
4~5%	15	18	18	22
5~6%	24	30	30	36
6~7%	33	41	41	50
7~8%	43	53	53	64
8% 以上	50	63	63	75

注 1) 表 8.2 及び 3 は固有灰分 4% の炭種に対する表である. したがって, 固有灰分 6% の炭種ならば, 灰分割の値は全灰分 $x\%$ のときは $x' = x - 6 + 4 = x - 2$ でもとめられる x' を見ればよい. たとえば $x = 9 \sim 10$ は, $x' = 7 \sim 8$ に相当する. 4 吋以下の切込炭で, 全灰分 = 10% の炭種ならば表 8.3 より灰分 $10 - 2 = 8\%$ 灰分のところ, すなわち 10 lb のインクリメントを $n = 64$ とればよい.
2) 最小粒度が小さくなると, 分離が起りやすいので, n を大きくしてある.

1) 本邦炭では炭種によりもつと軽い比重, たとえば 1.3~1.4 をとるべきである.

表 8・4 インクリメントの数 n の最小値 (米国規格)

粒 度 in	5/8以下	5/8~1 ¹ / ₄	1 ¹ / ₄ ~2	2~6	2~4	4~6	A	B
インクリメントの大きさ (w_i) lb	2	4	6	10	10	15	30	20
灰 分 %	インクリメントの数 (n)							
<8	15	15	15	15	15	15	35	35
8.0~9.9	20	20	20	20	20	20	35	35
10.0~14.9	35	35	35	35	35	35	35	50
15.0≤	35	35	35	35	35	35	35	50

注 1) 粒度の意味については表 8・1 A の注参照。 2) 精度は平均灰分の ± 10%、(すなわち変動係数の形であらわしてある) 信頼度 95%。 3) 精度を平均灰分の ± 5% にするには、 n を 4 倍、± 3.33% にするには 9 倍すればよい (∵ 精度は n の平方根に反比例するから)。 4) 大口試料の量 $W = w_i \times n$ 5) 以上の 3 表は本邦炭すべてにそのままあてはまるわけではない。

8・4・2 インクリメントの数の決定に及ぼす他の要因

§ 8・4・1 にのべたことは、粉塊混合物の特性を考えて、ランダムサンプリングとみなせる場合である。しかし実際にインクリメントの数を決定するにあたっては、1) ロットの大きさ、2) ロット内の傾向、大波、3) サンプリングコストなどを考えなければならない。

1. **ロットの大きさ** 従来のサンプリング法はロットの 1/1,000, 1/2,000 というように決めているがこれはおかしい。

通常の工業におけるサンプリングでは、ロットの大きさがサンプルの大きさに比し大きいので、 $\sigma_p = 0$ で σ_i が一定なら (2・27) 式に示したようにロットの大きさに関係なく、サンプルの大きさ、この場合はインクリメントの数 n が決定される。しかしこれは § 2・3・3 の 1) 注意にのべたように、あくまで数学理論によるのであって、いろいろな誤差をとめないやすい、われわれの行うサンプリング、縮分、測定では、もし大きな誤差が入ると、また小さな精度でも、たとえば 1 万トンの契約となると、その結果にあたる経済的な影響が大きい。また 1 万トンの大きなロットの平均値を知っても層積法による貯炭などを行わなければ、使用目的からいって技

術的に意味のない場合もあるので、あまり大きなロットについて混合試料をつくらずにいくつかの副ロットにわけて、その各副ロットについて、所要精度をうるための n を決定し、各副ロット毎に別々にインクリメントを集めて混合試料とし、縮分、分析する。たとえば石炭、鉱石類のように単価の比較的やすいものでは、1,000トン、2,000トンくらいの副ロットに、金銀鉱のように、高価なものでは、100トン、20トンなどの小さい副ロットにわけるのがよい。

このようにすれば全ロットについての平均値の精度は、副ロットすべてのデータを平均するので、層別サンプリングの効果で、層内のインクリメント間の分散を σ_w^2 とすると

$$\sigma_s^2 = \frac{\sigma_w^2}{kn} \quad (8.9)$$

ただし k : 副ロットの数 n : 各副ロットからとるインクリメントの数、精度はよくなり、縮分や分析の誤差も大体 $1/k$ になる。したがってもし大口試料採取、縮分、分析に誤りがあってもその影響は小さくなり、経済的に安全である。

また一般にロットが大きくなると、工程が管理されていなければ副ロット間のバラツキにより σ_p が大きくなるので、 n は大きくなる。工程が管理されていて、 σ_p が小さく、経済性ということも考えなければという前提のもとに、数学的にサンプルの大きさはロットの大きさに無関係といえるのであるが、目的と経済性を考えるといくらか関係がある。

2. ロット内の傾向、大波のバラツキ σ_p 船積にて多量に入荷した鉱石類は、ハッチ毎に、同じハッチ内でも上下前後により、品質にカタヨリがある。また貨車で入荷した中塊炭でも、各貨車間、貨車内にはカタヨリがある。したがって、ロット内に傾向や大波の変化がある。管理された工程からの製品で、粉塊の分離が行われていない場合には $\sigma_p = 0$ でインクリメントはランダムに変化するということも考えられるが、現在の日本ではこのようなことは少なく、多少の傾向や大波がある。サンプリングの際に、これらの変化を2つの面から考慮しなければならない。

a) サンプリングの目的として、このような大波や傾向を知る必要はないか。この点については§ 8.3.1にのべた通り、各副ロット毎の知識が必

要な場合が多い。これは層積法などのロットの取扱法、使用法とも関係してくるが、サンプリングの目的を技術的に考えて、副ロットにわけてサンプリングする必要がある。このとき、全ロットについては経済性を考えて、各副ロットについては技術的目的を考えて、 k , n をきめればよい。

b) 大波や傾向を知る必要なく、平均品質を推定すればよいときには、 σ_p を小さくするため、§4にのべた n の割合など統計的な工夫を行う。すなわち w_i の決定には σ_p はほとんど影響せぬが、 n は σ_i のみならず σ_p によつても大きく支配される。したがつて前記英米の n の表が現在の日本の管理状態に適用できるか否かは疑問である。またたとえば、一定間隔でインクリメントをとつたときに、全体でのインクリメント間のバラツキ $\sigma_{i'}^2 (= \sigma_p^2 + \sigma_i^2)$ より $\sigma_i^2 = \sigma_{i'}/n$ とはならない。これは σ_p が層別サンプリングなどの効果により小さくなるからで、 n は上式よりもとめられる値よりも遙こ小さくなる。しかしいろいろなロットに対するサンプリング法標準を決めなければならぬ場合には、安全な側へ、すなわちランダムサンプリングのときの上式を用いて、 $\sigma_{i'}$ より n をきめておけば安全である。しかし一般に、大波や傾向のあるロット—— σ_p が大きくなる——から所要精度をうるための n は、管理されたロット—— σ_p は小さい——からの必要なインクリメントの数よりも、あまり大きくならないのが通常である。

3. サンプリングコストとインクリメントの数 粉塊混合物のサンプリングにおいては、サンプリングの経費は多くの場合人件費であり、サンプリングする時間に比例してきまってくる。したがってコンベヤーを流れるロットを一定間隔でサンプリングするときには、コストは n には殆んど関係しないで、むしろロットの流れる時間、たとえば入荷、出荷、輸送時間によりきまる。したがって一定時間にとる n 、すなわちあるロットからとるインクリメントの数は多くても少なくとも、大差ない場合が多い。これに比し粉炭を貨車でサンプリングするようなときには、サンプリング時間はインクリメントの数により決定されるから、サンプリング法を決定する際には費用函数を考えなければならない。

このほか、大口試料の量と縮分誤差の関係など、いろいろの要因を研究、実験してはじめてインクリメントの大きさと数が決定される。

8.4.3 インクリメントの取りかた

インクリメントの取り方は、大口試料の集めかたである。

大口試料はとりなおすことができぬし、インクリメントの取り方によりカタヨリが入る場合が多いから、あらかじめ十分に計画をたて、十分注意して集めなければならぬ。

§ 8.4.1, 8.4.2 できめた所要インクリメントの大きさ以上で所要精度をうるに必要な数をとる取り方の問題で、如何にロットの粒度分布、成分分布と同じようにカタヨリなくとるかという技術の問題である。

1. **インクリメントは原則として移動中にとること。** 粉塊混合物は、積んだままでは、あるいは積んでしまった状態では、完全なランダムサンプリングを行うことは不可能な場合が多い。特に粒度差のあるとき、乾燥しているときは、不可能である。洗粉炭のように、粒度差小さく、最大粒度が小さく、湿っているときには、貨車積のままサンプリングしても、あまりカタヨリがないかも知れぬが、なるべく避けた方がよい。積荷中、荷下し中にとるのがよい。

1米以上山積になった石炭などからは、合理的なサンプリングはできない。米英でも3呎以上積んだ場合のサンプルは、特に“top sample”という特別の名称をあたえ区別している。これも貯炭場への運搬時にサンプリングすべきである。

船積で入荷したときも同様で、荷下し期間中にとらなければならない。石川の実験によれば、置いてある、静止した1kg乾燥粉炭より100gのインクリメントでサンプリングしても、カタヨリを生じやすい結果がでている。

2. **インクリメントはロット全体について、一定間隔でサンプリングすればよい。** 一定間隔で系統的サンプリングすることになるから、周期性にさえ注意すればよい。たとえばクレーンでコンベヤーへ落すときに、常にその先に洗れて来たものばかりとらぬようにする。一部分から何杯もつづけてとるのはよくない。

3. **インクリメントは、時間的、場所的に目の前にある部分よりランダムにとること。** ややもすると、粉の方が成分の悪いことを知っているとし生産者は粉の少ない方を、消費者は多い方を有意的にえらび勝ちであるから、§ 7.4 にのべたようなサンプリングカードの工夫もよいであろう。従来多くの場合ロットと同じような粒度分布になるようにとれと記されている

が、これは実際問題として仲々困難であり、かえってカタヨリが入りやすい。粒度分布のことは念頭においておくことは必要であるが、むしろサンプリング法標準を設計するときに必要なのであって、標準通りにサンプリングすれば、採取者が粒度の判定を主観的に考えなくとも、サンプリングできるようにしておくべきである。粒度分布は混合試料としたときに全体としてロットと大体同じ分布をもっていればよい。

4. ロットが大きいときは、経済性を考えて、一定量ずつの、たとえば1,000 トンの副ロットについて大口試料1つをとり、各大口試料を別々に縮分、分析すべきである。各炭種が混載されているときには、炭種毎に副ロットとするのもよい。

5. 特に大きな塊の入っている場合には篩分けして層別し、大塊だけ別にサンプリングするか、あるいは粉碎してサンプリングするとよい。しかし大塊の量が微量か、その平均特性があまりカタヨっていないときには、無視してもよい。

6. 混合試料は水分にカタヨリを生ぜしめないように、サンプル容器、保存法、保存場所に留意しなければならぬ。なるべくならば、金属性の蓋付の容器がよい。

7. 湿分測定用に必要なサンプル量を、通常必要な大口試料の量に加えて大口試料をそれだけ増加して一緒にとる。

8. インクリメントはたとえば船から荷上げ中連続的にとらなければならぬので、人員、サンプリングの確実な実行という面から考えても、一定間隔でとる自動サンプリングマシンの設置が望ましい。最近いろいろな研究が行われている。

a) コンベヤーからのサンプリング 粉塊混合物はコンベヤーでサンプリングするのが一番容易である。しかもコンベヤーの落ち口でとるのがよい。

コンベヤー上、あるいは落ち口でとる場合には、その流れ全体から——巾および厚さの方向に——一定間隔でインクリメントをとる。もし全体が一度にとれぬときは、左右、中央等各部から交互にとる。コンベヤー上でとるときには、スコップをコンベヤーの底まで入れぬと粉炭をとりそこなう。そのコンベヤーに前のコンベヤー、シュートからどのように入ってく

るかを考えずに、片側からだけサンプリングすると、巾方向へのカタヨリが入りやすい。

一定間隔は、あらかじめロット全量をインクリメントの数で割ってもとめ、一定量通過毎にインクリメントをとる。たとえばクレーンで何杯目毎に、コンベヤスケールである一定量毎に、とる。しかし荷上作業量が管理されていけば、一定時間毎にとってよい。

バケットコンベヤでは、粉塊の分離が起りやすいから、注意を要する。たとえばバケットからでるときに、上方および下方から交互にとる。

コンベヤに最初に流れてくる部分、側の方を転がるものはとらぬこと。

b) シュートからのサンプリング 一定間隔でシュート中よりとる。スコップでシュートの前後左右より交互にとること。他はAと同じ。

c) 人力運搬の場合のサンプリング バイスケ等で積荷、荷下しするときには、運搬者をランダムにえらび、その一方の籠をあけさせて、その中から1インクリメントをランダムにとる。

d) 山積からのサンプリング 1米以上になると技術的に殆んど不可能である。これは船積、貨車積の際にもいえることで、山の表面やスノの方からだけサンプリングしたのではカタヨリが入りやすい。

e) 本船からのサンプリング 積んだままでサンプリングできない。荷下し中、できれば、A、B、Cのようなところととる。コンベヤ、シュートでとるのがもっともよいか、本船から直ちに山積するようなときには、一定間隔で、たとえばクレーンで20回つかむ毎に、20回目のクレーンでつかんだ場所から、あるいはこれからつかむ所からランダムにとる。機帆船あるいはハシケ取りの場合も同様である。さらに級間、級内分散を考えて、属別、副次サンプリングなどすべきである。

f) 貨車からのサンプリング できれば積荷中、荷下し中にサンプリングすること。貨車の表面からとってはいけない。

貨車間、貨車内などのバラツキ σ_b 、 σ_w などが管理されていて、わかっているときには、副次、層別、多段サンプリングにおけるサンプルの割当てを考えて、貨車をランダムに n 車えらび、次にその中から n コずつインクリメントをランダムにえらぶ。貨車内においても、上下、端と中央、前後などにカタヨリのあるときは、貨車内で層別比例サンプリングをすれば

よい。通常よく管理された同一炭坑からきた石炭であれば、貨車間と貨車内のバラツキは大体同じくらいとなる場合が多い。

もしこれらの予備知識のないときには、1インクリメントずつでもよいから、全車両からとること。

スコップで荷下しするときには、全車両荷下し予定時間中、一定期間毎にスコップを止めさせて、そのスコップよりインクリメントをとる。

粉炭、特に湿っているときには、貨車上でたとえば1尺くらい掘って穴をあけ、穴の壁を下から上へ次には底から1インクリメントずつサンプリングすれば大したカタヨリはでない。貨車上の位置はランダムに、あるいは対角線上隅に近い方（たとえば隅から1/6のところ）、あるいは中央などからランダムにとるか、ランダムカードを用いてとるとよい。とる場所は貨車毎にかえるとよい。湿分試料は、特に全層からランダムにとる必要があるが、炭層の大体中央の部分からとれば大体よいこともある。しかし表面からだけとつてカタヨリがないか、チェックしておく必要がある。

塊炭、または塊炭と中塊炭以下が混合しているときには、なるべくなら貨車上でサンプリングはさける。止むをえぬときは貨車上に粉炭の場合よりやや深く穴を掘って、塊炭、中塊炭を別々にサンプリングし、これを縮分後その重量比に混合するように英国ではきめている（この場合湿分用サンプルは別にとること）。塊炭及び中塊炭以下の各インクリメントの大きさ及び数は表8・5、8・6の通りである。粉炭の量の少ないときは、それを無視してよい。 w_4 を小さく、 n を大きくしている点に注意

表 8・5 貨車からのサンプリング (BS1017)
塊炭のインクリメントの大きさと数 (精度 $\pm 1\%$; 信頼度99%)

灰分(%)	インクリメントの 最小の大きさ(lb)	インクリメント 最 小 数	大口試料の最小量 (lb)
4 以下	2	50	100
4~5	〃	73	145
5~6	〃	120	240
6~7	〃	165	330
7~8	〃	213	425
8 以上	〃	250	500

表 8・6 貨車からのサンプリング (BS1017)

中塊炭以下のインクリメントの大きさと数 (精度 $\pm 1\%$; 信頼度99%)

灰分 (%)	最小インクリメントの大きさ (lb)	最小インクリメントの数		
		1 吋以下	2 吋以下	3 吋以下
6	1	20	40	50
7	1	25	53	65
8	1	33	65	80
9	1	40	78	98
10	1	45	90	113
11	1	53	105	133
12	1	60	120	150
13	1	68	135	170
14	1	75	150	190
15	1	85	170	213
16	1	95	185	235
17	1	103	205	255
18	1	110	220	280
19	1	120	240	300
20	1	130	260	325
21	1	140	280	350
22	1	150	300	375
23	1	160	320	405
24	1	173	345	433
25	1	185	370	463

注 3 吋以下の中塊炭が 20% 以上入っているときには、インクリメントの大きさを 2 ポンドとする。たとえば坑口から出たばかりの原炭ではこのようなことがある。

底開きの炭車，カーダンパーで荷下しする場合などは，特別な装置を考案して，荷下し中にサンプリングできるように工夫すべき場合もあろう。

8・5 縮分および分析誤差

縮分および分析の工程が、合理的に、管理状態で行われていければ、その誤差は、大口試料採取の誤差よりも遙かに小さくなるのが通常である。

すなわち分散として各々大口試料採取の誤差の $1/10$ 以下、少なくとも $1/5$ 以下になる。しかし従来の方法を見ると、たとえば縮分が、不合理に、乱暴に行われているために、あるいは大口試料の取り方が不合理で、1トン2トンという多量とっているために——インクリメントの大きさの大きすぎる場合が多い——縮分の誤差が意外に大きくなっている場合がある。また縮分を非常に面倒な、不経済な工程にしている場合も多い。したがって大口試料の量に比し分析用試料の量の小さい粉塊混合物のサンプリングにおいては、縮分の合理化は経済的にも、誤差という問題から見ても重要な問題である。

縮分誤差もやはり、信頼性、精度、正確度とわけて論ぜられるが、一般に正確度は、たとえば縮分中の水分の蒸発、超微粉の散逸、不純物（たとえば岩石分）のカタヨリなどで、縮分法と操作について技術的工夫をすれば避けうる。信頼性は縮分という工程管理の問題であるが、一般に円錐四分法のように、主として人手による混合作業は、慎重に行われぬとあまり信頼できず、しかもバラツキが大きい。むしろ縮分工程の管理が容易な縮分法、縮分装置を用いるべきである。精度はここでも統計の問題となるので、後に論ずることとする。

分析誤差については、既に § 3 に詳細にのべたので省略するが、§ 3 にのべたように、分析用試料から分析試料採取のときの誤差も、混合を十分に行われぬと案外大きくなる——この誤差は分析誤差としてあらわれる——から注意を要する。縮分、分析誤差を論ずるときは、分析誤差は繰り返し精度でなく、人や日がかかったときの再現精度を用いるべきである。また従来精密測定法を用いているために、大口試料採取の誤差に比し測定誤差が小さすぎる場合も多く、もっと分析の簡易化をはかり、迅速分析とすべき場合が多いことを再び強調しておく。

8・5・1 縮分に関する原則

縮分も大口試料から分析用試料をとる1つのサンプリングであり、大口試料の母平均を誤差を小さく分析用試料より推定しようというのである。これは一種の集落サンプリングであり、集落をとるときにカタヨリの入ら

ぬように、集落間のバラツキ——精度——をなるべく小さくして、信頼性のあるサンプリングを経済的に行いたいのである。このような縮分を行うための原則は、大体大口試料採取の考え方と似ているが、次の通りである。

1. 縮分は必ず細い粒度で行うこと。すなわち縮分前に必ず粉砕すること。

粉塊混合物のサンプリングの精度は最大粒度により大きく影響されるといふ原則がここでもあてはまり、はじめから経済的に可能な範囲に粉砕しておく。縮分の精度をよくするためには、できれば2回——縮分前の粉砕と分析用試料の粉砕——通常3回——さらに縮分中に1回粉砕する——ぐらいにとどめる。すなわち、初めに5 m/m以下、できれば3 m/m以下少なくとも10 m/m以下に粉砕しておくといふ。初めの粉砕を十分にしておけば、縮分は非常に楽になる。しかしこのためには粉砕機の能力と経済性の問題を伴う。大口試料の量を無駄に多く取りすぎていると、かえって粗砕で縮分することになり、誤差が大きくなりやすい。合理的な量の大口試料をとれば縮分も合理的に行えるようになる。しかし塊炭、塊鉱のようなときには、精度をよくしようとするより大口試料の量が相当多量になるので、用途、目的によっては、粉砕後サンプリングするように工夫するとか、縮分前の粉砕を数段に行わなければならぬこともある。

この面から見ると、粉砕法と粉砕設備の研究が重要な問題である。

2. 粉砕後十分混合するか、あるいは混合不十分でも集落サンプリングできるような縮分法を用いる。いいかえると円錐四分法はなるべく避けて、試料分割器 (sample divider, たとえば riffle sample divider, cascade sample divider) を用いた方がよい。

円錐四分法はその操作を見てもわかるように、混合が十分に行われており、円錐を4分した各部分間が均質であるときに、はじめて誤差が小さくなる。ところが1トンもある混合物を均一化することは非常に困難であり、これを日常作業として行くと、どうしても混合をよい加減に行い勝ちである。これに比し試料分割器は、自動的にランダムに分れるようにできているので、混合が十分である方がよいか、十分でなくとも縮分の誤差は小さい。

1) 使用目的からいうとあまり精度を必要としない場合でも、日本人は精度をよくしたがる癖がある。

円錐四分法を乱暴に行ったとき、丁寧にいったとき、分割器を乱暴に用いたときについての石川の実験によると、分割器の精度がもっともよく、丁寧に行っても円錐四分法の方が悪くなっている。

円錐四分法のやり方は周知のことであるので省略するが、

- a) 円錐をつくる前に十分に混合すること。
- b) あまり細かく粉砕しすぎると混合が不十分になることがある。適当な粒度がある。

リップルサンプラー (riffle sample divider) とは、図 8・8 のような金属製の器具で、a の部分は一定間隔に平行の間隙があり、この間隙が左右に交互に開いていて、入ったサンプルはこれを通して2分される。間隙の寸法は粒度により、10m/m~30m/m くらいまであり、適当なものを用いる。大体最大粒径の2~3倍¹⁾ くらいのものを用いるとよい。サンプルが間隙につまらなければよい。この間隙の数は通常12~20 くらいである。サンプル受器は微粉の飛散を防ぐために、図のように分割器出口にピッタリ合うものがよい。サンプル装入器は大体分割器の長さと同じくらいのものでよい。これを使用する際には、装入器を左右へ僅かにふりながら、サンプルが均一に間隙に入るようにする。サンプルは前によく混合しておけばよいが、混合が比較的不十分でも、左右の間隙に比較的ランダムに装入されるので、精度はよい。この装置は最も簡単で、故障も少なく便利であるが、あまり湿ったものは間隙をつめてしまう恐れがある。しかし湿分用の試料は別項のように別にとるので、通常は大気または熱乾燥した後、縮分器にかける。

カスケード・カンプレアー (cascade sample divider) にもいろいろな形式がある。図 8・9 はその一例である。図のAは固定されており、Bは40~60 r. p. m. で回転するようになっている。図ではABの間隔が大部あいているが、実際には微粉のロスのないようにできるだけ密接させ、あるいはBの上端の中にAが入りこむようにしておけばよい。Bの横断面は右図のようにいくつかの部分にわけてある。図では、1~4 をとれば1/2に、2~5をとれば1/6に、3~6をとれば1/3に縮分される。試料はHの口か

1) 間隔がせますぎても、広すぎてもよくない。

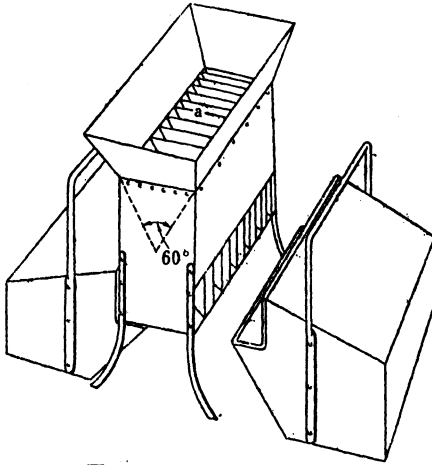


図 8・8 リッフル・サンプラー

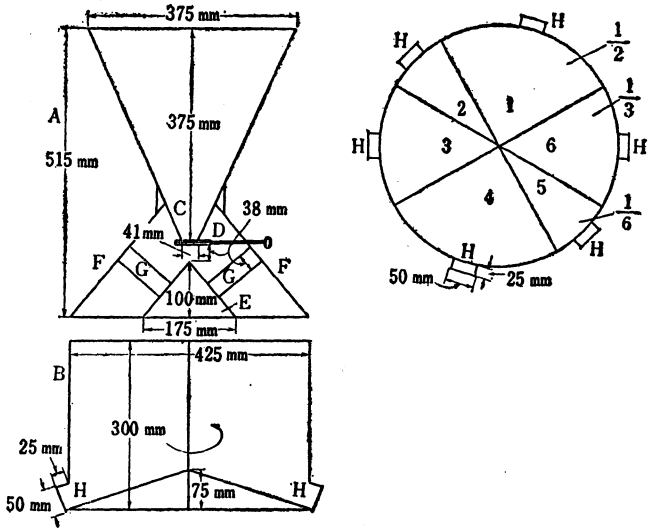


図 8・9 カスケード・サンプラー

らとり出す。カスケード・サンプラーでは、回転しながら分割するので、装入試料が均一に混合されておらず周期性をもって出てくると、誤差が大きくなりやすいので、Bの回転速度を比較的早く、かつ1セクションだけでなく、2-5をとるようにすれば誤差は小さくなる。

この外回転縮分器 (rotating sample divider) もある。図8・10にその例を示す。Dは12個の三角形容器を組み合わせてつくっており、各々取り外せる。

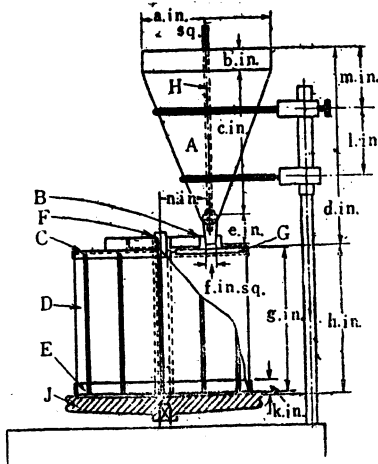


図8・10 回転縮分器

これが60 r.p.m で回転している。対称の2コをとれば1/6, 4コ毎にとれば1/3, 3コ毎にとれば1/4などに縮分できる。大型は6 m/m以下の粒度, 容量約10 kg, 小型は1.2 m/m以下の粒度, 容量2.5 kgのものがある。

3. 縮分誤差は、縮分を合理的に行えば、大口試料の量には殆んど影響されず、縮分後のサンプル量により大体決定される。

したがって、1) 3) を一緒にすると、「縮分誤差は、縮

分するときの粒度、縮分後のサンプル量および縮分法により大体決定される」。

実際には、縮分工程が、粉砕、混合が不十分であったりする場合があるので、必ずしも、1トンからの縮分も、10 kgからの縮分も一致するとはいえないが、しかし合理的に粉砕されており、縮分が合理的に行われれば、1トンのものを2分したときには、集落化がうまく行われ、集落間のバラツキ σ_b 、すなわち縮分の誤差が無視しうる程に小さくなる。この誤差は、最大粒度に比し最後のサンプル量が少なくなるに従って大きくなるので、主として最大粒度と、最後のサンプル量により決定されるのである。これは次のようなことを示している。いま第 n 段の縮分誤差を σ_n と

すると、全縮分誤差 σ_R は

$$\sigma_R^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \dots + \sigma_n^2$$

$$\sigma_R^2 \approx \sigma_n^2$$

すなわち $\sigma_1^2 \ll \sigma_2^2 \ll \dots \ll \sigma_n^2$ となることを示している。しかし縮分工程の管理状態により必ずしもこのようにならないことに注意。管理状態ならば大体この仮定は成立する。この間の関係については、E. G. Bailey¹⁾の古い研究がある。これについては § 8・5・4 にのべる。

4. 大口試料採取のときと同様に、縮分誤差は平均品質——したがってバラツキ——によっても影響され、一般に品質が悪くあるいは良くなるにしたがい縮分誤差の大きくなる傾向がある。したがってたとえば灰分の多い石炭程、粉砕粒度を小さくするか、縮分後のサンプル量を多くしなければならぬ。

5. 縮分の誤差 σ_R は、大口試料採取の誤差 σ_1 の 1/10~1/20 にしておけば十分である。経済的には、同程度、あるいは 1/5 くらいにすることもある。

注意 1) 採取した大口試料は縮分前、並びに縮分中に変化のおこらぬよう保存法、取扱法に留意しなければならない。粉塊混合物のサンプルは比較的少量になるので、取扱が乱暴になりがちであるから注意を要する。特に水分の変化——発散したり、雨にかかったり——に留意。また異物の混入にも注意。サンプル容器、その置場、などをよく管理しなければならぬ。湿分測定用サンプルの縮分については § 8・5・3 参照。

2) 縮分の際の粉砕は、磨砕 (grinding) よりも破砕 (crushing) の方がよい。磨砕では発熱によるサンプルの変質、鉄粉などの混入のおそれがある。鉄分の混入を防ぐには、特に硬いサンプル——たとえばコークス——のときは、粉砕機の材質はマンガン鋼など特殊鋼を用いるとよい。

3) 円錐四分法はなるべくさけた方がよいが、もしこれを行うときには、完全な混合と、正確な 4 分が大切である。このために、鉄板などを十字形にくみ、正確に直角にわけられるような道具をつくっておくとよ

1) E. G. Bailey, J of Ind. & Eng. Chem. Vol. 1, 1909, p. 161.

い。これを円錐を平にひろげた上にあてて、4分するのである。

以上の考え方により縮分法が決定される。

その一例を英米において、決められた石炭の例をのべよう。

8.5.2 石炭の縮分法

英国では、はじめに 1/4~3/16 吋以下に、米国では 100%が 3/8 吋以下で、95%以上が 4.76mm 以下に、粉碎することになっている。はじめにこのくらいまで粉碎しないと縮分の誤差が大きくなる。

英国では¹⁾、大口試料採取の際の精度により次の3方法にわけている。

第1法 大口試料採取の精度が ± 0.25 , $\pm 0.5\%$ (信頼度99%) としたときに用いる。縮分の精度を 0.25% 以下になるようにきめている。

1) 粉碎した試料を試料分割器により、表 8.7 の量まで縮分する。

2) この試料を BS 14メッシュ (1.2 mm) 以下に粉碎し、縮分して2封度とする。灰分 6% 以下のものは 1/2 封度でよい。

3) この試料をさらに BS 72メッシュ (0.2 mm) 以下に粉碎して分析試料とする。分析は2回行う。

注意 もし 1) で縮分したはじめの半分のサンプルを、別に縮分して分析すれば、精度は十分に保証される。

表 8.7 縮分第1法

粉碎後の石炭の粒径 (吋)	灰分 (%)	縮分後のサンプル量 (lb)	最大誤差 ($P_r=0.99$)
1/4 以下	{ 10	30	0.25
	{ 6	10	0.22
3/16 以下	{ 15	20	0.25
	{ 10	10	0.25
	{ 6	3	0.24

第2法 大口試料採取の精度が ± 0.75 , $\pm 1.0\%$ (信頼度99%) の場合に用いる。縮分の精度を 0.45% (信頼度99%) 以下になるようにきめている。

1) BS 1017, 1942; BS 735, 1944.

- 1) 粉碎した試料を試料分割器で表 8・8 の量まで縮分する。
- 2) これを BS 7 メッシュ (2.4 mm) 以下に粉碎し、試料分割器で 2 封度とする。灰分 6% 以下のときには 1 封度でよい。
- 3) これを BS 36 メッシュ (0.4 mm) 以下に粉碎し、水分および灰分の定量に用いる。他の分析はサンプル 2 オンス以上をとり、BS 72 メッシュ (0.2 mm) 以下に粉碎して用いる。

注意 本法ははじめから BS 36 メッシュ以下に粉碎できれば、その方が精度はよい。

表 8・8 縮分第 2 法

粉碎後の石炭の粒径 (吋)	灰分 (%)	縮分後のサンプル量 (lb)	最大誤差 ($P_r=0.99$)
1/4 以下	15	20	0.41
	10	10	0.39
	6	5	0.30
3/16 以下	15	16	0.42
	10	3	0.42
	6	1½	0.31

第 3 法 大口試料採取の精度が $\pm 1.0\%$ (信頼度 99%) のときに用いる。ただし縮分の誤差が $\pm 0.5\%$ (信頼度 99%) 以上にならぬように縮分法を考えなければならぬ——分散として 1/4 になる——また表 8・9 は、一般に最大粒度と縮分後のサンプル量と精度との関係をあらわしている。

1) 粉碎した試料を試料分割器で、表 8・9 にしたがって表の量以上まで縮分する。この場合縮分の誤差が $\pm 0.5\%$ (信頼度 99%) 以上にならぬよう、粉碎、縮分の組み合わせを工夫しなければならない。

2) これを BS 7 メッシュ (2.4 mm) 以下に粉碎し、試料分割器で 2 封度とする。ただし灰分 6% 以下のときは 1 封度でよい。

3) 第 2 法に同じ。

注意 以上の表 8・7, 8, 9 はいずれも表 8・2 と同様に固有灰分 4% としたときの表である。したがって本邦炭の様に固有灰分の異つた石炭ではそのままはあてはまらぬのは、§ 8・3 にのべたとおりである。

表 8・9 縮分第3法並びに一般的関係¹⁾

灰分(%)	6					10					15					20					最大誤差	
	2	1½	1	1/2	1/4	2	1½	1	1/2	1/4	2	1½	1	1/2	1/4	2	1½	1	1/2	1/4	$P_r = 0.90$	$P_r = 0.99$
縮分後の サンプル 量(封度)	1500	680	290	35	5	5600	2500	1100	130	17	12600	5700	2400	300	39	21600	9800	4200	500	66	0.2	0.31
	900	410	180	21	3	3300	1500	630	77	10	7200	3300	1400	170	22	12600	5700	2400	290	39	0.25	0.38
	600	280	120	14	2	2100	930	400	48	6	4300	2000	830	100	13	7600	3400	1500	180	23	0.3	0.46
	330	150	64	8	1	1100	480	210	25	3	2200	970	420	50	7	3400	1600	660	80	11	0.4	0.62
	200	92	39	5	—	660	300	130	15	2	1300	580	250	30	4	2000	910	390	47	6	0.5	0.78
	130	60	25	3	—	460	210	88	11	1	860	390	170	20	3	1300	590	250	31	4	0.6	0.94
	86	39	17	2	—	320	150	62	8	—	620	280	120	15	2	930	420	180	22	3	0.7	1.1

1) BS 735, 1944.

固有灰分の量に従って、[平均灰分 - (固有灰分 - 4)]%のところの表を見れば大体よい。各種本邦炭についてこのような表を実験的に作成するか、あるいはチェックしておけば安心である。

表 8・9 の使い方

i) 縮分時に 1½ 吋の最大粒度をもつ灰分 15% (固有灰分 4%) の石炭試料 1 トンあるときに、これを縮分して約 280 封度のサンプルとすると、最大誤差は表より ±1.1% ($P_r=0.99$) となり、相当大きな誤差となる。これを 1/4 吋以下に粉碎しておけば、39 封度まで縮分しても、誤差は ±0.31% ($P_r=0.99$) となり遙かに誤差が小さくなる。

ii) 固有灰分 8% の炭種で、平均灰分 17% くらいの炭種ならば

$$17\% - (8\% - 4\%) = 13\%$$

したがって 15% の欄で縮分すれば十分である。

iii) 最大粒度 50 m/m, 平均灰分 18% (固有灰分 4%) の大口試料 2 トンあるときに、これをそのまま 1 トン (2,240 封度) まで縮分し、次にこれを 25 mm 以下に粉碎してこれを 2 回縮分して 250 kg (約 550 封度) とし、次にこれを 6 mm 以下に粉碎して 5 kg (約 11 封度) としたときの縮分の誤差 ($P_r=0.99$) はいくらになるか。

a) 2 吋で 2,240 封度までの縮分誤差 ($P_r=0.99$) $s_1 \approx 0.78$

b) 25 m/m で 250 kg までの縮分誤差 ($P_r=0.99$) $s_2 \approx 0.62$

c) 6 m/m で 5 kg までの縮分誤差 ($P_r=0.99$) $s_3 = 0.62$

したがって分散の加成性より全体の縮分の誤差 s ($P_r=0.99$) は

$$s = \sqrt{s_1^2 + s_2^2 + s_3^2} = \sqrt{0.78^2 + 0.62^2 + 0.62^2} \approx \sqrt{1.38} \approx 1.17\%$$

すなわち大口試料の平均値は、えられたデータを 18.1% とすると、分析誤差を考えなければ、18.1% ± 1.2% すなわち 19.3% から 16.9% の間にある確率が 99% という非常に精度のわるいことになる。

もしはじめから、全量を 6 m/m 以下に粉碎し、10 kg (約 22 封度) まで縮分すれば、精度は 0.5% となり、はじめから十分粉碎すれば非常に精度がよくなることを示している。

米国では¹⁾、英国におけるデータを参照して、石炭の縮分法を次のようにきめている。

1) ASTM Standards on Coal & Coke. 1949.

- 1) 全量 3/8吋 (約 9.5 m/m) 以下, 95% 以上 4.76 m/m 以下に粉碎し, 60封度以上まで縮分, リップルサンプラーの間隙 (3/4吋~1吋) 使用.
- 2) 次に全量 4.76 m/m 以下に粉碎し, 1/2~3/4 吋のリップルサンプラーで, 15封度まで縮分.
- 3) 次に 2.38 m/m 以下95%以上になるように粉碎し, 1/2 吋以下のリップルサンプラーで, 1/4 封度以上まで縮分する. この際, 見かけ上ぬれているときには, 空気乾燥してから粉碎する.

微粉碎により, 1) 2) の段階を省略してもよいことにしている.

米国では, 変動係数で精度をだしているもので, 平均灰分の大小にあまり関係なく, 一定の精度となるが, 縮分の絶対誤差を相当大きく, 0.5~0.8 ($P_r=0.99$) くらいにとっているようである.

8.5.3 湿分測定用サンプルの縮分

湿分測定用サンプルの縮分は, 水分の発散によるカタヨリをさけるという技術的条件を考慮して行わなければならない. これは他の粉塊混合物も同様で, いずれも選鉱, 輸送, 貯蔵中に多くの水分をもっており, しかも契約量の算出に密接な関係があるので, 経済的にも大切である.

湿分の多少により, 次のような特別の方法を用い, 一般の縮分法は用いない.

測定の際にも, 縮分の関係から 1 kg~10 kg という多量を用いるのが通常である. したがって湿分測定用試料として必要量を, 大口試料採取のときに, 灰分等の所要精度よりもとめたインクリメント数に加えて, 多くとっておく必要がある.

湿分測定用サンプルの縮分の原則は,

- 1) サンプルの保存に留意し, 縮分などの操作は手早く行うこと.
- 2) 粉碎や円錐四分法, リップルサンプラーなどの使用はできるだけさける. 特に湿分の多いときの取り扱いには注意を要する.

第1法 (一般法) 特に目で見えてぬれていることがわかるときにはこの方法によること.

採取した大口試料を混合などせず手早く円錐状に積み, これを平に, 各部分同じ厚さになるようにひろげる. この厚さは, 次にのべるインクリメントをとるときに, スコップが底まで十分とどき, サンプリングできる

程度に薄くする。もし湿分の少ないときは、円錐を3回位積みかえるとよい。

サンプルのとり方は、粒度により次のように行う。

- a) 最大粒度 1吋まで 1封度のインクリメント 10 (計 10封度)
- b) " 2吋まで 2封度 " 10 (計 20封度)
- c) " 3吋まで 3封度 " 10 (計 30封度)

しかし2吋以上のときは、これを2吋以下に手早く破碎してインクリメントサンプリングすることもある。これによりカタヨリが入るか否かチェックしておくとうい。10のインクリメントはランダムにとるが、なるべく層間分散が大きくなるように、同じ大きさの10区劃にわけ、その各区劃からランダムに1インクリメントずつ層別比例サンプリングした方がよい。

第2法 目で見て乾燥している場合には、大口試料を混合などせず、手早く1/2吋以下に破碎し、円錐状に積み、平にして、10コ以上のインクリメントで10封度のサンプルをとる。

以上のようにしてとったサンプルが水分測定用として多量すぎるときには、更にこれを縮分してサンプルをとらなければならない。このときには次のようにする。

目で見て乾燥しているときには、1/2吋以下に手早く破碎して、10コ以上のインクリメントで測定に必要なサンプルをとる。

ぬれているときには、このまま1/2吋以下に粉碎するとカタヨリが入るので、一応うすくひろげて風乾(6~8時間)、あるいは強制通風、温風(50°C以下1.5~3時間)送風により大体恒量になるまでの湿分を測定して、後に粉碎して前記と同様にサンプリングし、湿分を測定し、両者を加えればよい。

尙精度をよくしようとして、インクリメントの数を増加させると、カタヨリが入るといわれている¹⁾くらいであるから、このサンプルの取り扱いには余程手早く行い、注意しなければならない。

以上の間、サンプルは密閉蓋付の缶に入れておかなければならぬのは勿論である。

1) BS 735, 1944, p. 63.

8・5・4 粒度—重量比の定理 (size-weight-ratio theory)

粉塊混合物のサンプリングの精度は、最大粒度とサンプルの量により大きく支配されるということに着目して、E. G. Bailey¹⁾が求めた定理であるが、これはその後 E. S. Grumell 等の研究²⁾により、実験的にそのままではあてはまらぬということが証明されている。この定理はプロセス的な変動 σ_p の多い大口試料採取にはあまり参考にならぬが、縮分の大体の目安をきめるには一応役立つこともあるのでここに紹介しておく。定理は次の通りである。

「粉塊混合物のサンプリングにおいては、そのサンプリングの精度 σ_s^2 は、最大粒の平均重量 w_i と、サンプル採取量 W の比の百分率を粒度—重量比 (Size-weight ratio, S. W. R.) とすると、SWR に比例する。

$$\sigma_s^2 = k \frac{w_i \times 100}{W} \quad (8 \cdot 10)$$

$$= k(\text{SWR})$$

この定理は、次のような欠点をもっている。

- 1) インクリメントによるサンプリングの概念が入っていない。
- 2) ロット内はランダムであると仮定していることになるが、実際は級間分散、級内分散、あるいはプロセス的な変化がある。
- 3) サンプリングのパラツキは最大粒度の目方 w_i によってのみ支配されるという考え方であるが、これは最大粒に灰分が全部集まっていると考えている。
 - a) 最大粒が何%あるかにより、サンプル中に含まれる割合や数がことなる。
 - b) 最大粒が大きくなるにしたがって、必ずしも各粒子間のパラツキが大きくなり、小さくなることもある。
 - c) 最大粒以外の粒子もサンプリングの精度に影響をあたえる。
 - d) w_i は炭種により相当異っており、必ずしも粒度により一定ではない。

したがって、多くの場合この定理からもとめたサンプル量は実際よりも

1) 前掲 (p.277)

2) BS 403, 1930, BS 763, 1937.

多すぎ、あるいは所要粉砕粒度は細かすぎる傾向にある。しかし前述のように大体的見当をつけるには便利である。

Bailey はある種の米炭について

$$r = 0.522\sqrt{ap} - 0.353p \quad (8.11)$$

ただし記号は原著のまま

r : サンプル中の灰分の確率誤差 = $0.6745\sigma_r$

a : 外来灰分パーセント

p : SWR

という式を出している。 p は通常 10^{-2} 以下の小さい値であるから、第2項はほとんど無視できる。すると (8.10) 式の k を平均外来灰分パーセント μ の一次関数とすればよいことになる。

いま $k = c\mu$ とすると、ただし c は一定数

$$\sigma_s^2 \approx c\mu \times (\text{SWR}) \quad (8.12)$$

w_i は炭種により大体きまるので、各 c をもとめておけばよい。いま

$$w_i = \rho d^3 \quad (8.13)$$

ただし ρ : 外来灰分の比重、その他の場合には不純分の比重

d : 最大粒径¹⁾

とする。すなわち近似的に最大粒は全部不純物であるとし、粒形を正方形であると仮定する。すると

$$\text{SWR} = \frac{\rho d^3}{W} \times 100 \quad (8.14)$$

$$\therefore \sigma_s^2 = \mu \times \left(\frac{c\rho d^3}{W} \times 100 \right) \quad (8.15)$$

実際にサンプリング実験を行つて、えた精度 σ_s 、 μ 、 ρ 、 d 、 W より c をもとめて、いろいろの場合の c についての知識をえておけば、それより、所要精度をうるためのサンプリング法、すなわち d 、あるいは W をうるることができる。

c は前述の仮定についての補正係数に相当するもので、一般に1より小さく、炭種、粒度、最大粒の数などによりいろいろな値をとりうる。ある

1) 最大粒としては、その粒度より大きいものが5%くらいある粒径をとつておけばよいであろう。

炭種について大体の値をもとめておけばよい。

一般にこれを求めるのは大変であるから、まず近似的に

$$c = 1 \quad (8.16)$$

とすると

$$\sigma_s^2 = \mu \times \left(\frac{\rho d^3}{W} \times 100 \right) = \mu (\text{SWR}) \quad (8.17)$$

となり、近似的にサンプリング法の見当をつけることができる。しかし c は1より小さいので、このようにしてもとめたサンプル採取量 W は相当大きくなりすぎ、所要粉碎粒度 d は立方根できいてくるので、少し小さすぎる値となる。

例1) ある金属鉱石において、その純硫化物の比重が2.5、その平均含有量は大体30%である。この鉱石の最大粒径20 m/mであるときに、その平均成分を標準偏差で0.3%位の精度でもとめるには、サンプルをどのくらいとったらよいか（この場合は金属化合物の方が少ないので、これを不純物と見ればよい）。

$$V(x) = \sigma_s^2 = 0.09 \quad \mu = 30\% \quad \rho = 2.5 \quad d = 2 \text{ cm}$$

$$\therefore \text{SWR} \approx \sigma_s^2 / \mu = 0.09 / 30 = 0.003$$

$$W = \frac{\rho d^3}{\text{SWR}} \times 100 = \frac{2.5 \times 2^3 \times 100}{0.003} \text{ g} \approx 6.7 \times 10^5 \text{ g} = 670 \text{ kg}$$

すなわち670 kgのサンプルをとれば十分である。

注意 1) 実際には $c < 1$ であるから、これ程のサンプルはいらない。

2) この値は、サンプル採取量を示してくれるだけで、インクリメントの大きさや数は示していない。

例2) 例1)においてもとめたサンプルを粉碎縮分して10 gの分析試料をうるには、どの程度の粒度まで粉碎すればよいか。

通常縮分の誤差 σ_R^2 は σ_s^2 の1/10くらいにとれば十分であるから

$$\sigma_R^2 = \frac{\sigma_s^2}{10} \quad \therefore \text{SWR} = \frac{\sigma_R^2}{\mu} = \frac{\sigma_s^2}{\mu \times 10} = 0.0003$$

$$\therefore d^3 = \frac{W(\text{SWR})}{100\rho} = \frac{10 \times 0.0003}{100 \times 2.5} \approx 12 \times 10^{-6} \text{ cm}^3$$

$$\therefore d \approx 2.3 \times 10^{-2} \text{ cm} = 0.23 \text{ mm}$$

すなわち Tyler 60 メッシュまで粉碎すればよい。この粒径は実際必

要な粒度よりも過小となっている。

注意 実際実験では $\sigma_s^2=0.09$ とするためには、 $SWR=0.01$ となった。すなわち 例 1) では (8・12) 式より

$$SWR = \frac{\sigma_s^2}{c\mu} \quad \therefore 0.01 = \frac{0.09}{30c} \quad \therefore c = \frac{3}{10} = 0.3$$

$$W = \frac{100r d^3}{SWR} = \frac{100 \times 2.5 \times 2^3}{0.01} = 2 \times 10^5 \text{ g} = 200 \text{ kg}$$

サンプルの量は実際には $3/10$ の 200 kg でよかった。

$$\text{例 2) では, } SWR = \frac{1}{10} \times 0.01 = 0.001$$

したがって

$$d^3 = \frac{10 \times 0.001}{100 \times 2.5} = 40 \times 10^{-6} \text{ cm}^3 \quad \therefore d = 3.4 \times 10^{-2} \text{ cm} = 0.34 \text{ mm}$$

0.34 mm, すなわち Tyler の 40 メッシュ程度まで粉碎すればよい。

この例で見ればわかるように、この定理によれば安全な側にデータがでるので、目的とする精度は十分にえられるが、経済性を考えると必ずしも合理的ではないことがわかる。

8・6 コークスのサンプリング

コークスは、水分のバラツキと粒度が比較的大きいものが多いという特徴をもっているので、これを例として簡単にのべる。

コークスのサンプリングにおいて注意すべきことは次のとおりである。これらはコークス以外の類似の粉塊混合物においても注意しなければならぬことである。

1. サンプリングの目的が、工業分析にあるか、強度試験にあるか、篩分試験にあるか。強度試験の場合には、たとえば高炉用として、粉を除去したものの強度を目的としているのか、あるいはコークス炉の運転を管理するためかにより、大分サンプリング法が異なってくる。

2. 強度試験の際には、同じ1本の乾溜室内のコークスでも強度に相当カタヨリやバラツキがあり、これをランダムサンプリングしたのでは、バラツキが大きすぎて、特に強度の弱いときには品質の推定の精度が悪くな

ってしまう。これは現在の潰裂強度試験そのものにも問題があるが、サンプリングとしては、炉内の上下のカタヨリに対して、技術的に目で見えてはつきり層別できるのであれば、層別比例サンプリングの形でサンプリングすれば精度はよくなる。あるいは中央部の形のととのったサンプルだけとるのであれば、平均品質と中央部の平均品質の相関関係をあらかじめもとめておけばよい。これらの有為サンプリングを行うことは、特に目で見ただけで主観的に選んでとるので精度はよくなるが、カタヨリが入りやすくあまりすすめられない。これについては特殊問題となるのでここでは省略するが、この場合も1コずつとるより、インクリメントでサンプリングすべきである。この方がカタヨリが少ない。

また強度としてある期間の平均値を推定するためには、その期間中から適当な時間だけにサンプリングするのはよくない。そのロットすべての部分についてあるいはその期間中から一定間隔で系統的サンプリングしなければならない場合が多い。

3. 工業分析としては、バラツキのもっとも大きいのは湿分で、これを目標にサンプリング法を決定しておけば、灰分などの精度は十分にえられる。湿分が小さく、2%位になると、大体灰分などのバラツキと同じ程度になる。

4. 湿分のバラツキは平均8~9%になるまで増加する。それ以上は通常はあまり大きくなるしない。

5. 湿分は粒度の小さいもの程多いのが通常である。

6. 湿分を目的とするときはコークスは温い、あるいは熱い間にサンプリングしてはならぬ。ロットとサンプルのその後の冷却、乾燥状態がことなるのでカタヨリが入る。

7. コークスは輸送、貯蔵中の気象状況により、雨、雪、天気、風、また輸送中に表面が乾燥するなどの変化が大きいから、表面からだけサンプリングすることは特に避けるべきである。

8. 多くの場合最大粒度はあまり変化せぬ。

9. 乾燥したコークスは吸湿性が強い。

以上のことを考えて、すなわち、コークスの粒度、湿分含有量、所要精度によりインクリメントの数と大きさを決定しカタヨリのないように技術

的に考えて、サンプリング法を決定しなければならない。

英国では¹⁾次のように決めている。

粒度により2種に分ける。

1) 消火したときに1吋以下のものが7.5%以下で、湿分5%以下のとき。

2) その他のコークス、およびブリーズ。

インクリメントの大きさは最小2.5封度で一定にしているが、これはコークスにより粒度に大した変化がないからであろう。ブリーズなど細かい場合には、もっとインクリメントの大きさを小さくしてもよいように思われる。インクリメントの数は表8・10にしたがっている。インクリメントの大きさが小さく、数が大きいのは、コークスが部分によりかなり不均一だからである。

表 8・10 インクリメントの最小数

種 類	1		2	
	インクリメントの数	大口試料最小量 (lb)	インクリメントの数	大口試料最小量 (lb)
2 以下	20	50	35	88
2~3	44	110	70	175
3~4	74	185	120	300
4~5	95	238	170	425
5~6	112	280	235	588
6~7	118	295	280	700
7~8	126	315	300	750
8 より大	134	335	305	763

表8・10の精度は信頼度99%で±1%の場合の表である。実際にこれ程の精度を必要としないときには、インクリメントの数の平方根に逆比例して精度は大きくなる。たとえばインクリメントの数を1/2にすれば、精度は大体±1.4%になる。

1) BS 1017, 1942.

この他更に粒度により分類し、所要精度 ± 0.25 , ± 0.5 , ± 0.75 などの表もあたえているが、省略する。

縮分法は大体石炭湿分サンプリングのときと同様である。

手早く 1/2 吋以下に粉碎し——この際篩分けなどしてはならぬ——約 500 g のインクリメントで 50 コとる。大口試料の非常に多いときには、粉碎機の出口から一定間隔で、あまり多くないときは、これを円錐につきみ、平にひろげて、底の粉まで残さぬようにランダムに 50 コとる。次にこれを手早く円錐につきみ、平にして、必要試料——通常約 1 kg、ブリーズのときは 2 kg 以上——を 20 インクリメントでとり、サンプルとし、湿分を測定し、以後の分析用に用いる。その後の縮分も、手早くインクリメントでサンプリングする。

注意 1) 石炭及びコークスについてのべた数字は英米のものについてであり、また疑問の点もあるので、わが国の石炭にそのままあてはまらぬものもあると思われる。たとえば一番問題になる σ_p すなわちロットの管理状態がことなるであろう。また固有灰分、外來灰分の関係もことなるであろう。したがって、以上のべた粉塊混合物のサンプリングの考え方にしたがって、本邦炭、輸入炭などについて、§5 にのべたような方法を工夫して予備実験を行い、わが国に適した、その工場に適した、もっとも経済的、合理的なサンプリング法を作成し、わが国産業の合理化をはかっていただきたい。

2) 粉塊混合物のサンプリングの場合には、こゝではロットの管理状態 σ_p と、比較均一な、すなわちごく近いところからとつたインクリメント間のバラッキ σ_s とにわけて話をすゝめてきたが、予備実験としては、これを一緒にした σ_s をもとめるように行つてもよい。

第9章 綜 括

9・1 品質管理とサンプリング

統計的品質管理の普及とともに、工場に従来からあるデータを利用して管理図をかくときに、まず問題になったのはデータの歴史である。特にデータの歴史に大きな関係をもつサンプリングにいろいろな問題がおこってきた。

いまその主な問題点をあげて見ると

- 1) 目的の不明確なサンプリングが多い。
- 2) カタヨリのあるサンプリングが多い。
- 3) サンプリングの精度が殆んどすべて不明である。

精度という面から見るとサンプルが多量すぎるか、少量すぎるか、特に多量すぎるために時間的、経費的に不経済となり、确实に行われていないか、あるいは多量すぎるためにかえって精度をわるくしている。

- 4) サンプリング、測定という工程の管理が行われていない。

- 5) 管理図とサンプリングの関係

6) 相関分析、分散分析など統計的方法を用いる場合のサンプリング以上のうち 1)～4) はすでに各章にのべてあるので、ここでは 5)、6) について、簡単にのべる。しかしサンプリングの問題は管理という考えかたとともに、統計的品質管理の基礎となるので、これまでのべてきたことを十分考えて研究する必要がある。

9・1・1 管理図とサンプリング

管理図において、サンプリングが問題になるのは、これまでのべてきた一般的な問題と、管理図特有の群分けのやりかたとサンプリング法との関係の問題である。

一般的な問題の主な点をのべると、

- 1) 工程管理の場合には、常に工程を母集団と考えているので、一般に無限母集団として取り扱う。

2) 管理の目的により、群の作りかたとサンプリングの方法がきまってくるから、その管理図をいかなる目的でかくかということをはっきりさせることである。従来サンプリングを考えない、目的のはっきりしない管理図がわが国の多くの工場に見られる。

3) カタヨリのあるサンプル、現場で有意的にえらんだサンプルは工程管理のための管理図用のデータにはならず、サンプリングのカタヨリを管理している場合がある。しかし誰でもが機械的に同じようにカタヨリの入ったサンプルをとれる場合には、そのバラツキは小さくなるから、有利な場合があるが、これには試料採取者間のカタヨリに差のないことを十分チェックして、行わなければ危険である。

4) あるロットの平均値推定のサンプリングの精度が、工程の変動(標準偏差として)の1/2以上になると、管理限界の巾が広くなり、工程の変動に対する検定力がわるくなるから、サンプリングの精度をよくしなければならぬ。サンプリングの精度が工程の変動の1/5~1/10以下のときには、サンプリングの精度をおとして、簡易化した方がよい。このことは群の分けかたと関係があるので、後にものべる。

5) サンプリング間隔は、可避原因(assignable cause)による製品のバラツキを逃さぬ程度にしておけば理想的である。しかし可避原因の周期とサンプリングの周期が一致していなければ、いつかは必ずつかまえることができる。そのときに可避原因を完全に除去して今後起らぬように処置(action)をとるのが品質管理の原則であるから、そう神経質になることはない。一般に管理状態が長くつづき、工程が安心できれば、サンプリング間隔を長くし、管理状態がやぶれば間隔を短くするのが常道である。しかしこれも経費とにらみ合せて決定すべきである。たとえば、連続25点以上限界外に点がとびださず、連(run)や傾向(trend)もなければ、サンプリング間隔を倍にする。あるいは p 管理図のときにはサンプルの大きさを1/5にへらすなど。限界外に点が出たらまたもとにもどす。

6) \bar{x} 管理図では、一般に R 管理図からもとめた群内変動より管理限界を決定している。

次の(9・1)及び(9・2)式はランダムサンプリング、またはそれに近いときは、そのまま用いてもよいが、特別な場合(たとえば異質の層が1つ

の群に入っている)には不適當なことがある。

$$A_2R = 3 \sqrt{\frac{\sigma_x}{n}} \quad (9.1)$$

\bar{x} のバラツキは、 σ_μ を群間のバラツキ、 σ_x を群内のバラツキとすると

$$\sigma_{\bar{x}}^2 = \sigma_\mu^2 + \frac{\sigma_x^2}{n} \quad (9.2)$$

となる。また σ_μ 、 σ_x は、サンプリングのやり方、群への分け方により変化する。

σ_μ が σ_x^2/n に比し、相対的に大きくなると限界外にとびだす点が増加してくる。工程を管理して行く管理図を作成する場合には、 σ_μ があまり大きくなりすぎても使用できないし、また $\sigma_{\bar{x}}^2 < \sigma_x^2/n$ という関係となり一見管理状態を示す場合)でも、製品のバラツキが規格などに対して不満足のとときは、やはり使用困難で、いずれの場合も群分け、サンプリング法を工夫する必要がある。特に集合体の場合には、各単位のサンプル²⁾のとり方により、 σ_x が変化するから、サンプル単位間のバラツキがあまり大きくならぬよう工夫することが必要である。

7) R 管理図は、 σ_x より管理限界がきまるので、群内に異質の層のサンプルが入っていなければ、その検定力は、工程自身の変動、群の大きさ n 、サンプル単位のサンプリングの精度により支配される。工程の変動を鋭敏に検定するためには、 n のとり方をかえるか、一定間隔における n を大きくするか (これは不適當な場合がある)、サンプル単位のサンプリングの精度を小さくするように工夫すればよい。群内に異質の層のサンプルが入っているときには、管理図の原則に従って群の分けかたをかえるか、層別する。

サンプル単位のサンプリング精度をよくすることは、これまでのべてきた方法で、容易に実施しうるから、ここでは主として群へのわけかたと、 σ_μ と σ_x との関係について論ずる。

群分けとサンプリング 管理図におけるサンプリングは、層別できぬ

- 1) サンプリングの構造がランダムサンプリングになつておらず、いくつかの層からサンプルをとり層別サンプリングの構造になつている場合など、 \bar{x} の点は中心線の近くにあつまる。
- 2) 一般に管理図法では、サンプルを群の意味に用いるが、ここでは群とよび、各単位サンプルをサンプルとよんで区別することにする。

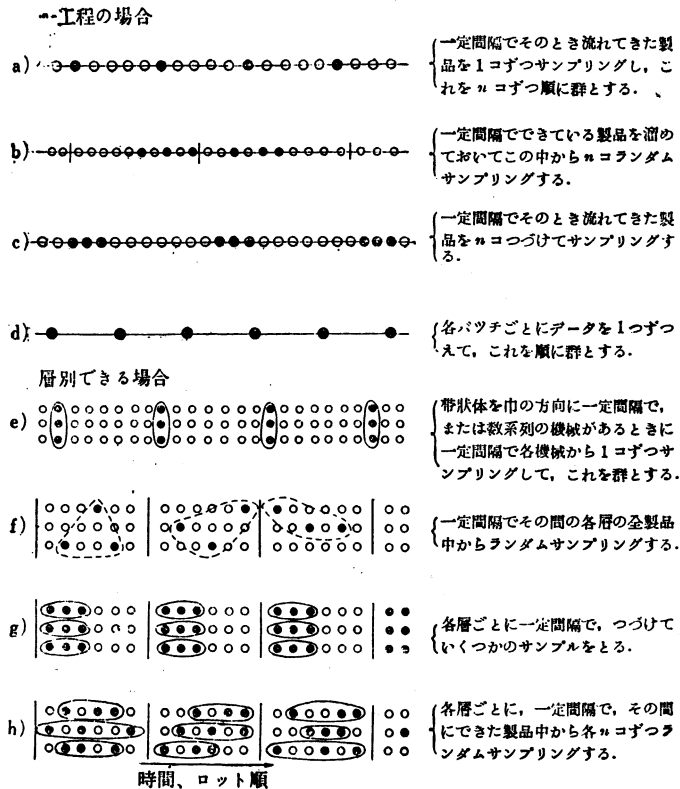


図 9・1 管理図におけるサンプリング

一工程の場合と、層別できる場合とあり 図 9・1 の様な数種の型に入る。

通常の工程管理のためのサンプリングは、一定間隔で、すなわち系統的サンプリング¹⁾を行うのが通常である。したがって図 2・3 にのべたような工程の周期などと、サンプリング間隔に対する注意が必要である。管理

1) §2・2・2 参照。

図の場合には、各サンプル単位の周期と、群を単位とした周期とを考えなければならぬ。サンプル単位としての周期性は § 2・2・2 にのべたので、ここでは群についての周期性を考えよう。しかし何れの場合でも、周期性によるカテゴリーさえさけられるならば、

- 1) 単純ランダムサンプリング
- 2) 層別サンプリング
- 3) 副次サンプリング

いずれかに似たサンプリングになつている (図 2・3 参照)。ランダムサンプリングのときは別に問題ないが、層別及び副次サンプリングとなつているときには、群分けの問題が級内分散 σ^2_{α} (群内分散) と級間分散 σ^2_{μ} (群間分散) とに大きく関係してくる。特に \bar{X} 管理図の検定力に大きな関係をもつ。

\bar{X} 管理図は、完全な管理状態の工程なら、図 9・1 のどのサンプリング法をとっても、ランダムサンプリングとなるから、どのように群分けしても大差ない。しかし管理という面から見れば層別しておく方が有利である。ところが実際には、管理状態の工程は少なく、また管理図が一見管理状態を示していても、完全な管理状態ではない場合が多いので、工程の周期性や傾向、層間のバラツキによりどのサンプリング法をとり、どのように群分けすると、どのような検定力をもつかということが問題になる。そこで工程の周期、層間のバラツキなどとむすびつけて考えなければならぬ。

層別できぬ場合 層別できぬ工程の場合には 1), n をサンプリングする群の周期が、工程の周期よりも長ければ、大体ランダムサンプリングとなり、 σ_{μ} が小さくなるから、a) b) d) では \bar{X} 管理図は大体管理状態を示す。しかしそれでも工程が不満足ならば、 n を小さくするかサンプリング間隔を短くして、群の周期を短くするか、あるいは c) のようなサンプリング法にかえる。工程の周期が非常に短く、管理図は管理状態を示すが、規格など (合理的にきめられているものとする) に対して不満足なときは、自動記録計、あるいは自動制御装置をつけて、連続的にサンプリングし作業標準のように取り扱って管理することが必要な場合もある。工程の周期の方が群の周期より長くなれば、 σ_{μ} は次第に大きくなり、管理限界外にとびだす点は増加してくる。

1) 巾のない線材からのサンプリングもこれと同様に考えられる。

この場合群はその群の期間の工程からのサンプルであるから、いまその期間の工程を母集団と考えれば、各サンプル単位は a) は母集団からの系統的サンプリングであり、b) はランダムサンプリングであり、c) は母集団をさらにいくつかの副ロットにわけ、その中から1副ロットとりその中からサンプリングした形になる。したがって一般にはサンプリング法をかえただけであるから

$$\sigma^2 = \sigma_{x_a}^2 + \sigma_{\mu_a}^2 = \sigma_{x_b}^2 + \sigma_{\mu_b}^2 = \sigma_{x_c}^2 + \sigma_{\mu_c}^2 \quad (9.3)$$

$$\sigma_{x_b}^2 \geq \sigma_{x_a}^2 \geq \sigma_{x_c}^2 \quad (9.4)$$

$$\sigma_{\mu_b}^2 \leq \sigma_{\mu_a}^2 \leq \sigma_{\mu_c}^2 \quad (9.5)$$

ただし添字 a, b, c, は前図サンプリング法をしめす。また全工程を母集団としてを群を考えれば、a) では層別サンプリングに近く、b) は完全な層別サンプリング、c) は副次サンプリングに近い形式となっている。

d) は各バッチのすべてのデータを用いているので、十分層別するか、工程の周期を考えて、大波や傾向をもつときは n を小さくして σ_{μ} を小さくするか¹⁾、あるいはバッチ内からいくつかのデータをとって群とし、群内変動に相当するものをもとめるかすれば、検定力はよくなる。たとえば、もし可能ならば、バッチ内から n コのサンプルをとり、これより別々にデータをもとめ、これを群とする。すなわちバッチ内分散を σ_{μ} として用い、バッチ間分散を σ_{μ}^2 とすることも考えられる。

層別できる場合 いくつかの機械、工程からの製品の場合、あるいは鋼板、紙のように巾のある帯状製品の場合には、技術的に機械別、工程別、あるいは巾、長さの方向別に層別できる。この場合はサンプリングの面から見るといろいろ面白い問題がある。

機械別、工程別に層別できる場合には、層別サンプリングをして、各層ごとに管理図をかくのが原則である。すなわち e) のようなサンプリングを行わずに、g) h) のようなサンプリングを行うのが原則である。各層ごとに層別すれば、e) は a) と同じになり、g) は c) と、h) は b) と同じである。しかし層間に有意差がなければ、あるいは特別の場合には、e)

1) 一般に \bar{x} 管理図では、同じ σ_{μ} ならば、群の大きさ n を小さくすると検定力がわるくなるが、 \bar{y} のバラツキに対しては σ_{μ}/\sqrt{n} で影響するので、 σ_{μ} を小さくした方が効果がある。

のようにサンプリングして、時間ごとに、すなわち各層を1つの群にまとめることもある。f) のようなサンプリングは、品質管理では原則として行わない。

いま各時間ごとの層内のバラツキを σ_w 、層間のバラツキを σ_h 、時間ごとのバラツキを σ_t とすると¹⁾、

f) の場合は R から求めた \bar{x} の管理限界、すなわち群内のバラツキ σ_x は

$$\sigma_x^2 = \sigma_h^2 + \sigma_w^2 \quad (9.6)$$

大きさ n コの群の平均値 \bar{x} のバラツキ $\sigma_{\bar{x}}^2$ は

$$\sigma_{\bar{x}}^2 = \sigma_t^2 + \frac{\sigma_h^2 + \sigma_w^2}{n} \quad (9.7)$$

したがって、 \bar{X} 管理図では主として時間ごと、一定時間ごとのバラツキ σ_t を検定し、管理することになる。

e) の場合は、各層から1コずつサンプルをとっている。すなわち層別サンプリングである。したがって \bar{x} のバラツキは、層の数を k とすれば、 $\sigma_{\bar{x}}$ には σ_h は入ってこないから、

$$\sigma_{\bar{x}}^2 = \sigma_t^2 + \frac{\sigma_w^2}{k \times 1} \quad (9.8)$$

しかるに R から管理限界を計算すると

$$\sigma_{\bar{x}}^2 = \sigma_h^2 + \sigma_w^2 \quad (9.9)$$

\bar{x} のバラツキには σ_h が入ってこないにもかかわらず、限界の巾にはこれが入ってくるので巾が広くなりすぎ、 \bar{x} の点は中心線の近くにおち不適当な管理図となる。このときは、各層内の σ_w をもとめて、——たとえば各層からとるサンプル数を $n(n \geq 2)$ にして——これより管理限界線を計算しなければならぬ。このとき管理限界線は $\bar{x} \pm 3 \frac{\sigma_w}{\sqrt{nk}}$ となる。

この管理図は、各層に共通な原因によるバラツキを見るのに役立つ。

また e) において、層ごとに層別して、時間順に n コのデータを群とし

1) 帯状体の場合には、一定の長さ内の長さ方向のバラツキ σ_w 、巾の方向のバラツキが σ_h 、一定長さ間のバラツキが σ_t に相当する。

2) 実際には e) のときは系統的サンプリングが副次サンプリングの効果を發揮して (9.6) 式の σ_x よりこの方が小さくはなる。特に σ_w が小さくなる。

て、管理図をかくことにすれば、各層別した管理図の限界の巾 σ_x は、 σ と同じで

$$\sigma_x^2 = \sigma^2_{t_1} + \sigma^2_{w'} \quad (9.10)$$

ただし $\sigma^2_{w'}$: 副次サンプリングの効果により f) の場合の群内のラツキ σ^2_w より小さくなるので記号をかえた。

$\sigma^2_{t_1}$: 群内の時間によるバラツキ

\bar{x} のバラツキは

$$\sigma_{\bar{x}}^2 = \sigma^2_{t_2} + \frac{\sigma^2_{t_1} + \sigma^2_{w'}}{n} \quad (9.11)$$

ただし σ_{t_2} : 群間の (時間による) バラツキ

g) のように、層別して、一定時間に n コずつづけてとる場合には

$$\sigma_x^2 = \sigma^2_{w'} \quad (9.12)$$

$$\sigma_{\bar{x}}^2 = \sigma^2_{t_1} + \sigma^2_{t_2} + \frac{\sigma^2_{w'}}{n} \quad (9.13)$$

となり、 \bar{x} 管理図の検定力はもつともよくなる。

g) の場合、 k コの各層から n コずつのデータを取り、これを一まとりに群にしたときは層別サンプリングになるので、 \bar{x} のバラツキは

$$\sigma_{\bar{x}}^2 = \sigma^2_{t_1} + \sigma^2_{t_2} + \frac{\sigma^2_{w'} k}{nk} \quad (9.14)$$

となる。このデータの限界を nk コのデータの R からもとめると

$$\sigma_x^2 = \sigma^2_b + \sigma^2_{w'} \quad \therefore 3 \frac{\sigma_x}{\sqrt{nk}} = 3 \sqrt{\frac{\sigma^2_b + \sigma^2_{w'}}{nk}} \quad (9.15)$$

となりやはり限界の巾が広すぎ、 \bar{x} の点は中心線の近くにおちる様になる

さらに一般に M コの層があるときに、その中から m 層を取り、各層から n コずつランダムにとって、群とした場合には、副次サンプリングのようになるから

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2_b}{m} + \frac{\sigma^2_{w'} k}{mn} \quad (9.16)$$

したがって、これを管理限界に用いるべきであるが、実際にはこのようことはほとんどあるまい。

1) 工程を母集団と考えているから、母集団修正はつかない。次にのべる帯状体の場合も同様にらない。

以上の式は何れも測定誤差を考えていない式であるが、実際には測定誤差を考えて計算しなければならない。

帯状体の場合も、e) f) g) h) などとはほぼ同様に考えられるのでこれは例題によって説明しよう。紙、ビニルシート、帯鋼、鋼板、テープ類など巾をもって、巾の方向と長さの方向にバラツキの考えられるものが（鋼板のように切れていてもよい）これである。

いまロールにある種のシートをまきとっている例を考えよう。シートは一定の長さごとに切って、1ロールとしている。切られる時期は工程と無関係に一定の長さで切るので、その切口附近の引張強さのバラツキは、各ロール内バラツキがランダムにあらわれている。一定時間ごとにサンプルをとって測定するが、1ロールを途中で切ってサンプリングすることは、実際上不可能であるので、その時間に巻きとっている1ロールが終わったら、その最も外側に相当する3mのところから n コのサンプルをとることにした。したがって、1ロールをサンプルの1次単位と考えると、ロールは大体系統的サンプリングされることになる。次にその1ロールから3mという2次単位を、さらにそれから n コの3次サンプル単位をとることになるので、結局系統的3段サンプリングをしていることになる。いま過去のデータ、および実験計画的に分散分析法を用いて、調査した結果

二箇所附近におけるサンプリングと測定をふくめた誤差 $\sigma^2_M = 8$

巾の方向のバラツキ $\sigma^2_b = 6$

3m内の長さの方向のバラツキ $\sigma^2_w = 14$

1ロール内各3m間のバラツキ $\sigma^2_t = 4$

ロール間のバラツキ $\sigma^2_R = 10$

となり、巾方向のバラツキは、たとえば右側が強いか、中央が強いとかいうカタヨリや傾向はなく、ランダムに変わっていた。

この場合いろいろなサンプリング法が考えられ

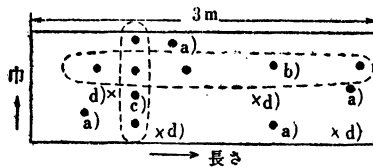


図 9・2

るが、4 コのサンプルをとり群とすると、いくつかの場合に R から求めた \bar{X} 管理図の管理限界すなわち群内の σ_x と \bar{x} のバラツキ $\sigma_{\bar{x}}$ がいかになるか考えて見よう。測定は各サンプルごとに1回行うものとする。

a) ロールの端から3 m切ったものより、サンプリングカード¹⁾によりランダムに4 コのサンプルをとる(図9・1 f)と同じ。

b) 巾方向にランダムに、或る一定場所をきめて、長さの方向に4等分して4 コのサンプルをとる(図9・1 h)の1層をとった場合と似ている。

c) 長さの方向にランダムに一定場所をきめて、巾の方向に4等分して4 コのサンプルをとる(図9・1 e)と同じ。

d) 巾の方向にランダムに2箇所きめて、長さの方向にランダムに各2コずつ計4 コのサンプルをとる(図9・1 h)の2層をとった場合と似ている。

いま \bar{X} の管理限界は、4 コのデータの \bar{R} よりもとめるものとする、 $(\bar{R}/t_2)^2 = \sigma_x^2$ 、群内のバラツキ σ_x^2 がいかになるかを計算すればよい。分散の加成性より

$$a) \sigma_x^2 = 14 + 6 + 8 = 28 \quad \therefore A_2\bar{R} = \frac{3}{\sqrt{n}}(\sigma_x) = \frac{3}{\sqrt{4}}\sqrt{28} \approx 7.9$$

$$\sigma_{\bar{x}}^2 = 10 + 4 + \frac{14 + 6 + 8}{4} = 21 \quad \therefore 3\sigma_{\bar{x}} \approx 13.7$$

$$b) \sigma_x^2 = 14 + 8 = 22 \quad \therefore A_2\bar{R} = \frac{3}{\sqrt{4}}\sqrt{22} \approx 7.0$$

$$\sigma_x^2 = 10 + 4 + 6 + \frac{14 + 8}{4} = 25.5 \quad \therefore 3\sigma_x = 15.3$$

$$c) \sigma_x^2 = 6 + 8 = 14 \quad \therefore A_2\bar{R} = \frac{3}{\sqrt{4}}\sqrt{14} \approx 5.6$$

$$\sigma_x^2 = 10 + 4 + 14 + \frac{6 + 8}{4} = 31.5 \quad \therefore 3\sigma_x = 16.8$$

$$d) \sigma_x^2 = 14 + 6 + 8 = 28 \quad \therefore A_2\bar{R}^2 = \frac{3}{\sqrt{4}}\sqrt{28} \approx 7.9$$

$$\sigma_{\bar{x}}^2 = 10 + 4 + \frac{6^3}{2} + \frac{14 + 8}{4} = 22.5 \quad \therefore 3\sigma_{\bar{x}} = 14.2$$

1) §7・4 参照。

2) このようにサンプリングしたときは、唯 $A_2\bar{R}$ から管理限界をもとめるのは計測の構造からは正しくないが、ここでは機械的に計算した。

3) この場合、巾の方向のバラツキもやはり工程を母集団に考えているので母集団修正はいらない。長さ、ロール内、ロール間いずれもいらない。

この場合いずれの方法をとっても、通常のように \bar{R} から管理限界をもとめたのでは、限界外にとび出す点が相当多いが、いずれも何を管理しているのかわからぬ状態になる。すなわち a) では σ^2_R と σ^2_ϵ により \bar{X} 管理図の点が限界外に、b) では σ^2_R , σ^2_ϵ , σ^2_b により、c) では σ^2_R , σ^2_ϵ , σ^2_w により、d) では σ^2_R , σ^2_ϵ , σ^2_b により、主としてとびだす。ロール間のバラツキを管理するためには σ^2_ϵ が比較的小さいから、a) なら使えるかも知れぬが、a) の方法を用いる場合も σ^2_ϵ もふくめて、 \bar{X} 管理図の限界線として

$$3\sigma_x = 3 \times \sqrt{4 + \frac{14+6+8}{4}} = 3 \times \sqrt{11} = 10.0$$

とすれば数学的には合理的であろう。

しかしこの場合は σ^2_w が一番大きく、長さ方向のバラツキを加えると

$$14+4+10 = 28$$

となり、長さ方向に非常に大きなバラツキがあり、巾のバラツキはこれに比し遙かに小さいから、シートの端だけでよいから、1ロール全体から系統的にサンプリングして、1ロール毎に群として管理した方が適当かも知れない。

もし a) の場合、測定誤差 σ^2_M が相当大きいからというので、1箇所からとったサンプルを2回ずつ測定し、その平均値を各々データとして4コの平均値を群にとると

$$\sigma^2_x = 14+6+\frac{8}{2} = 24$$

となる。強度試験などのバラツキも大きく、工程にもいろいろなバラツキがあるようなときは、経済的に事情が許せば、同じ近くのサンプルでの測定の繰り返し数を増加して、その平均値の精度をあげれば、ある目的に対しての管理図の検定力はよくなる。

このようにいろいろなバラツキを調査して、これをよく考えて、その目的に応じて、工夫すれば、比較的合理的なサンプリング法を設計することができる。しかしこのバラツキは工程の管理状態により、次第に変化するものであるから、常にチェックできるようにしておくとうい。

管理図の書き方、群の分け方ということサンプリングという面から見ると

- 1) 工程のいろいろな分散の調査
- 2) 系統的サンプリングのサンプル単位の周期、群の周期と、工程の周期、傾向との関係。
- 3) サンプル単位のきめ方

などを管理の目的と技術とをむすびつけて考え、研究する必要がある。

以上は主として検定力とその合理性についてのべたが、最後に一言ふれておきたいことは、管理図の検定力をよくすることが管理図の目的ではない。合理的な限界を計算することは、もちろん非常に参考にはなるが、このような計算をしなくとも、いろいろ層別し、群分けして見て、技術的に、管理責任者がもっとも管理しやすいような管理図をつくりあげることができる。むしろこれを組織、作業標準、品質標準とむすびつけて管理状態の管理図をつくり、管理の道具としていかに使って行くかということの方が品質管理としては、遙かに重要なことであり、管理図の主目的はこの点にあるということである。

9・1・2 その他の統計的方法とサンプリング

統計的方法を用いて、いろいろの推定や検定を行うときにもサンプリングは基本的な問題である。たとえば実験順序を時間的、空間的にランダムにするということは、非常に重要な原則であるが、現実には物を対象にしてのサンプリングの問題は少ないので、特に気づいた点だけを簡単にのべる。

1. サンプリングの精度をよくすること

一般に検定や推定の際には、技術的に小さな原因と考えられるものは、ランダム化して誤差分散に入れてしまう。サンプリングの誤差も誤差分散に入る。誤差分散をなるべく小さくして、検定などの効率をよくするためには、サンプリング誤差を小さくした方が有利である。たとえば第2種の過誤を考えたときに、実験回数が非常に小さくてすむ。したがって、サンプルの大きさを大きくするとか、サンプル単位の大きさや、サンプリング法を考えて、サンプルのパラッキを小さくするという工夫が必要である。

2. もしロットが層別できるときは、それを層別してサンプリングし、この層別を要因にとるか、あるいは一緒にして層別サンプリングによる精

度の向上をはかると有利である。

3. サンプルングの研究に、特にいろいろなバラツキの推定に分散分析がよく用いられる。この予備調査に用いられる分散分析は、通常工程管理のために多く用いられる母数模型ではなく、変量模型のことが多いから注意されたい。変量模型のことを詳細にのべることは、本書の範囲をこえるので、附録 A・1 に代表的な変量模型のことを簡単にのべておいた。

9・2 サンプルング問題の解決手順

サンプルングに問題が起って来たときに、どのように考えて行けばよいか。これはいろいろの経済情勢、技術、設備の状況により千差万別ではあるが、一応これまでのべてきたことをまとめて見よう。

- 1) 目的の明確化
- 2) ロットの歴史、管理状態、その後の取り扱いに対する研究と行動 (action)
- 3) 測定法の検討
- 4) 各種分散の推定並びに周期性、傾向の検討
 - a) 過去のデータの調査
 - b) 予備調査実験の計画、実施、解析
- 5) サンプルング法標準の設定
- 6) サンプルング並びに測定管理
- 7) データの活用、行動

基礎 1. サンプルングの重要性の啓蒙、普及 2. 品質管理の必要性、考え方の啓蒙、普及、実施 3. 統計的な考え方の啓蒙、普及 4. 組織の合理化

1. 目的の明確化

何のためにサンプルングするかはっきりさせることである。必ずデータの活用、行動を考えよ。

これについては既に各所でのべて来たがもっとも重要なことで、2) の特にそのロットをいかに取り扱うかということと密接な関係がある。一般に工業において日常作業として行うサンプルングでは、工程管理のためにはバラツキを知るように、原材料、出荷製品の場合にも、バラツキを知るためのサンプルングが必要であるが、そのロットを技術的に均一化できる

ような場合、例えば層積法 (bedding) の活用できる場合には、平均値だけ知るサンプリングでよい場合も多い。特に化学分析のように、測定に時間や経費のかかる場合には、平均値だけを目的とすれば、非常に経済的になる。バラツキを知るためのサンプリングは、一般にサンプル数、測定数が増えて厄介である。しかし技術的に均一化の困難な場合には、平均値だけ知るサンプリングをしてもいみのない場合が多いから注意、また保証のためのサンプリングか、工程管理のためのサンプリングかもはっきり区別しなければならぬ。

2. ロットの歴史、管理状態、取り扱いに対する研究と行動

工程やロットが管理状態にあり、層別化あるいは均一化が行われていれば、サンプリングは非常に容易になる。バラツキの大きいロットについて、困難なサンプリング法で、技術的、統計的に苦勞するよりも、工程やロットが管理状態になるように行動をとる方が経済的である。これには品質管理の実施、契約書の合理化、経営者、購買、資材関係者の教育、原材料供給者への行動、などが必要であろう。

ロットをその後いかに取り扱って、どのように使用するかということの合理化も大切である。取り扱いには層別と均一化とあるが、均一化は一寸した技術的工夫で行えることもあるし、サンプリングも全体の平均値を知ればよいだけであるから非常に容易になる。層別、層積法 (bedding) など、原料についても、製品についても、今後大いに研究すべきであろう。

3. 測定法の検討

これについては§3にのべた通りであるが、測定の管理、生のデータを大切にすること、データの修正や棄却などの注意が必要である。

4. 各種分散の推定並びに周期性、傾向の検討

工程やロットの管理状態は、変化するものであるから、長期間にわたる、多種類の工程やロットについての検討と、その変化状況をチェックすることが必要であることを特に強調しておきたい。

5. サンプリング法標準の設定

サンプリング法標準は必ず作成しておかなければならないが、これについて注意すべき点は次の通りである。

- a) 必ず書いたものにしておくこと、

- b) 具体的に、誰にでもわかるようにサンプルング法標準書をつくる。
- c) 必要に応じて、定期的な改訂できるようにしておくこと。
- d) 国家あるいは同業団体の規格として、あるいは会社間、会社内の仕様書として、公認されたものにしておくこと。
- e) 統計的に合理化せんとし、あまり複雑なものにならぬよう、実際に容易に行えるものであること。これにはサンプルングの場所、装置、道具、量などを技術的に検討しなければならぬ。
- f) サンプルングの管理責任者をきめておくこと。
- g) 信頼性をえやすいサンプルング法であること。
- h) カタヨリの入りにくいサンプルング法であること。ランダムサンプルングしやすい形にしておくこと。
- i) 精度がわかっているサンプルング法であること。信頼度は目的に応じ、95、90、80%くらいが多くの場合実用的であろう。99%という信頼度における精度を向上させることは、サンプルの大きさが大きくなりすぎて、実用的でなくなる場合がある。
- j) サンプルングによりえられる情報の価値が、サンプルングに必要な経費よりも大になるようなサンプルング法であること。たとえば精度やカタヨリをあまり小さくしようとしても意味のない場合がある。
- k) 経済的に事情が許せば、サンプルング法や測定法の、信頼性、精度、カタヨリをチェックできるようにしておくこと。たとえば出荷側と受入側の連絡。

要するに、サンプルング法標準は、理想的、科学的なものではなく、実用的、技術的、経済的なものでなければならぬ。

6. サンプルング並びに測定の管理

サンプルング法標準ができたなら、これを普及徹底させなければならぬ。

また実際に実行できるように教育しなければならぬ。これにはTWI方式など大いに活用すべきである。

その後、実際その通り実行しているか否かを管理しなければならぬ。

これにはサンプルング法標準と並行的に、測定法標準を作成しておくべきである。従来からある測定法標準には不合理な、技術的でないものが多いから、たとえ従来法があっても再検討して、5)にのべたと同様にしてい

理化した測定法標準書を作成しておくべきである。

7. データの活用と行動

サンプリングをし、データをとる目的からいって、えたデータを活用し、これにより行動をとらなければ全く意味がない。データが途中で修正されていないか。活用して、はたして合理的な行動を実際にとっているか否かをチェックしなければならぬ。さらにその行動の結果までチェックして、はじめて目的を達したといえよう。

基礎1～4 啓蒙、普及、実施、合理化

サンプリングの重要性、品質管理の必要性、考え方、統計的な考え方、組織の合理化などの啓蒙、普及、実施は、合理的なサンプリング法を確立するための基礎になるもので、実際にはサンプリング法の研究より以上に大切な問題である。したがって、数人の品質管理技術者が、統計的にいろいろひねくりまわし、難しい実験をし、サンプリング法標準をつくっても、この基礎がなければ何にもならない。この基礎が、現場、工場、会社、さらに広くは日本人全体にまで普及し、実行されるようになれば、合理的なサンプリングは自ら行われるようになろう。

特に為政者、経営者、工場幹部の考え方の啓蒙、合理化、決断がもっとも重要な基礎条件である。

9.3 むすび

品質管理の実施とともにその不合理が明らかになってきたわが国の従来から行われているサンプリング法の合理化は、わが国産業の合理化のためにも、日本産業の世界市場への発展のためにも、是非必要なことで、これには関係者全員の協力が必要である。サンプリングの研究は非常に時間、労力、経費、努力の必要なもので、関係者が協力してそれを分担してはじめて、目標に一步步近づけることができる。サンプリングの研究は非常に地味な、経費のかかるものであるが、これが合理化されれば社会生活に貢献するところは誠に大きいと思う。工業を通じての社会生活の向上に日夜努力しておられる読者に、また産業の合理化に、このサンプリングに関する拙著が、何らか益するところがあれば筆者の望外の喜びである。

附 録

A 1. 分散分析の構造

サンプリングの予備実験をする場合に、大体の分散の推定をするのに実験計画的にデータを取り、これを分散分析法により解析すると便利である。§5では予備調査のやり方をいろいろのべたが、更にいろいろな手法を使うやりかたである。しかしこの方法は分散の大体の大きさを推定するにはよいが、工業製品の場合には、工程によりバラツキが変化するので、なるべくなら§5・1にのべたように、長期間における簡単な調査が望ましい。通常の工程管理のための実験計画が大部分母数模型であり、また技術的知識を用いて、母数模型の形でやらなければ、いろいろの面から、特に処置(action)をとるという意味において損である。これに対し、サンプリングの予備調査の場合には、バラツキについては、常にその製品の生産工程を母集団と考えており、またロット、あるいは副ロットからランダムにサンプリングすることになるので、いわゆる変量模型になる場合が多い。母数模型と変量模型では、構造がことなるので、検定法も、推定法もことなってくる。これを次に簡単に解説する。

A 1・1 母数模型と変量模型¹⁾

実験を行うときに、技術的に重要だと考えられる原因を因子としてとりあげ、たとえば温度 T_1, T_2, T_3 , 触媒 C_1, C_2, C_3 という各3水準を考えているときには、技術的に温度 T_1 では結果はいかになるかというように、各水準について母平均がありと考え、この母平均の間のバラツキを問題にしている。このように技術的にある条件を指定できる場合は各水準について母平均という母数を対象に考えているので、これを母数模型という。従来通常に行われている分散分析は多くこの型である。母数模型のときには、 $T_i C_j$ という条件すなわち温度 T_i , 触媒 C_j という条件で行ってえ

1) 増山元三郎；工学技術者のための実験計画法の話（検定編）日本化学会講習会テキスト 1951. 参照。

たデータ x_{ij} の構造を考えると

$$x_{ij} = \mu + a_i + b_j + z_{ij} \quad (\text{A1}\cdot 1)$$

ただし μ : この実験計画で行った全体の条件からえられる総平均

a_i : T_i で行ったために, μ よりのカタヨリ. すなわち $\mu + a_i$ は T_i という条件の際の母平均. ただし $\sum a_i = 0$

b_j : C_j で行ったためのカタヨリ. ただし $\sum b_j = 0$

z_{ij} : $T_i C_j$ で行ったときの実験誤差. すなわちランダム化された原因によるデータのバラツキ. これはランダムにかわるので, 同じ実験を何回も繰り返したときに, z_{ij} の母集団の平均値は 0 になる. すなわち z_{ij} の期待値 $E(z_{ij})$ は

$$E(z_{ij}) = 0$$

(A1·1) 式で μ , a_i , b_j は一定の値をもち, z_{ij} だけがランダムにかわる値である.

これに対し, たとえば入荷してきた石炭について貨車 3 台をランダムにえらび, その中からインクリメントを各 5 コヅつとったときには, 貨車間と, 貨車内のインクリメントごとのバラツキを知りたいのである. 貨車ごとのバラツキも, 3 台とった貨車間のバラツキではなく, 貨車ごとの平均値の全体としてのバラツキを知りたいのである. ここにえらばれた 3 台の貨車の平均値は, 母数模型のように一般に指定できず, ランダムにかわるものとしている. 貨車内のインクリメントについても同様である. したがってこのときには

$$x_{ijk} = \mu + a_i + b_{ij} + z_{ijk}$$

ただし μ : この実験での総平均

a_i : i 番目にランダムサンプリングされた貨車の平均値の μ よりのカタヨリ. これはランダムにかわる変量であるので, 多くの貨車, 貨車に炭坑でつみこむ工程, という母集団を考えると, その母集団全体での平均値即ち期待値は 0 になる.

$$E(a_i) = 0 \quad (\text{母数模型のときは } \sum a_i = 0)$$

b_{ij} : i 番目の貨車内で j 番目にサンプリングされたインクリメントの $\mu + a_i$ よりのカタヨリ. すなわち貨車内でインクリメントをとったためのカタヨリ. これもランダムにかわる変量であるので, 多くのインクリメントという母集団を考

えると、その平均値は0となる。

$$\therefore E(b_{ij}) = 0$$

$$z_{ijk} : \text{縮分, 分析誤差} \quad E(z_{ijk}) = 0$$

以上の例でわかるように、実験を計画するときに、母数模型では a_i , b_j を一定の値、母数と考えており、変量模型では、 a_i , b_{ij} 何れも変量と考えている。サンプリングの予備調査ではランダムサンプリングの場合が多いので、変量模型の場合が多く、また母数模型と変量模型では分散分析表における、不偏分散の構造——不偏分散の期待値——がことなっているから、検定法も、分散の推定法も異なる場合が多いので、ここに附録としてのべるのである。ただしここでは不偏分散の期待値をもとめる計算の手順をのべるだけの紙数がないので、主として結果のみを記すこととした。

ところが変量模型にも2種類ある。対応のある場合と対応のない場合である。たとえば前の貨車のサンプリングのデータが第A1.1表のようにえられたとする。この表において1番目の貨車の1番目のインクリメントと、2番目の貨車の1番目のインクリメントでは、1番目という名称は似ているが、ともにランダムにとられているので1番目という意味は技術的に全くない。すなわちインクリメント間のバラツキという因子は貨車ごとに全く対応がないのである。これに対し貨車という因子の方は、1輛目の貨車ということは、1列目のデータにつ

表 A 1.1 二元配量のデータ

		貨 車 間		
		1	2	3
イン ク リ メ ン ト 間	1	=	=	=
	2	=	=	=
	3	=	=	=
	4	=	=	=
	5	=	=	=

いては、技術的にも意味のあることで、各列ということでも列ごとに対応がある。この例では貨車間に対応のある変量模型であり、インクリメント間に対応のない変量模型である。

サンプリングの予備調査では、母数模型、対応のある変量模型、対応のない変量模型、その他特別な模型のものがよくあらわれるので、どの模型と考えるべきかをその構造を技術的によく考えて、決定しなければならない。構造模型をまちがえると、分散の推定や、検定法をまちがえる。

次に例によってこれを説明する。

法 1) 一般にサンプリングの予備調査としては、工程からの製品である工業材料を対象とするので、母数模型の場合は少なく、対応のある変数、対応のない変数模型の場合が多い。

2) 分散分析表で母数模型のときと変数模型のときと同じ型となる場合でも検定の際の第 1 種の過誤はかわりないが、第 2 種の過誤が一般にかわってくる。

A 1.2 一元配置法

一元配置法では、母数模型も変数模型も同じ形になる。一般に一元配置法のデータは、次のようにあらわされる。

母数模型 $x_{ij} = \mu + a_i + z_{ij}$

ただし $\sum a_i = 0$ $E(z_{ij}) = 0$ $A: k$ 水準 繰り返しの数: r

変数模型 $x_{ij} = \mu + a_i + z_{ij}$

ただし $E(a_i) = 0$ $E(z_{ij}) = 0$ $A: k$ 水準 繰り返しの数: r

例 母数模型 ある標準試料を経験年数 5 年以上の甲、5~1 年の乙、新入工丙の 3 種の工具に、各 5 回ずつ分析させた。船の上層、中層、下層にわけて、その中からランダムに r コずつのサンプルをとる（技術的に上層、中層、下層ということが、品質に差がありと思われる場合）。

変数模型 ある標準試料を、多くの分析工の中から、ランダムに 3 人えらび、各 5 回ずつ分析させた。貨車をランダムに k 台とり、その中から r コずつのインクリメントをランダムにとり各 1 回分析する。

不偏分散の期待値 $E(V_e)$ をもとめる手順は次の通りである¹⁾。以下の各構造のものも同様にもとめられるので、特別な場合を除いて省略する。

B. 母数模型の場合

$$\bar{x}_{i.} = \mu + a_i + \bar{z}_{i.}$$

$$\bar{\bar{x}} = \mu + \bar{\bar{z}}$$

ただし $\bar{x}_{i.}$: 第 i 層の r コのデータの平均値, $\bar{\bar{x}}$: データ全体の平均値

$$E(S_A) = rE \sum_{i=1}^k (\bar{x}_{i.} - \bar{\bar{x}})^2 = rE \sum_{i=1}^k \{(\mu + a_i + \bar{z}_{i.}) - (\mu + \bar{\bar{z}})\}^2$$

1) この証明は興味のある方だけよんでいただければよい。

$$\begin{aligned}
 &= rE\{\sum^k (a_i - 0)^2 + (\bar{z}_{i.} - \bar{z}_{..})^2\} \\
 &= r(k-1)\sigma^2_A + r(k-1)\frac{\sigma^2_K}{r} = r(k-1)\sigma^2_A + (k-1)\sigma^2_K
 \end{aligned}$$

ただし σ^2_A : 因子A間の分散

$$\sigma^2_A = E \sum^k (a_i - 0)^2 / f_A = E \sum^k a_i^2 / (k-1)$$

σ^2_K : 誤差分散

$$\sigma^2_K = E \sum^k \sum^r (z_{ij} - \bar{z}_{i.})^2 / k(r-1) = rE \sum^k (\bar{z}_{i.} - \bar{z}_{..})^2 / (k-1)$$

$$\therefore E(V_A) = \frac{E(S_A)^2}{k-1} = \sigma^2_K + r\sigma^2_A$$

$$\begin{aligned}
 E(S_K) &= E \sum^k \sum^r (x_{ij} - \bar{x}_{i.})^2 \\
 &= E \sum^k \sum^r \{(\mu + a_i + z_{ij}) - (\mu + a_i + \bar{z}_{i.})\}^2 \\
 &= E \sum^k \sum^r (z_{ij} - \bar{z}_{i.})^2 = k(r-1)\sigma^2_K
 \end{aligned}$$

$$\therefore E(V_K) = \frac{1}{k(r-1)} \times k(r-1)\sigma^2_K = \sigma^2_K$$

したがって分散分析表の構造は次のようになる。

表 A 1・2 一元配置法分散分析表 (A : 母数または変量)

要因	S.S.	f	$E(V_e)$
A	$S_A = r \sum^k (\bar{x}_{i.} - \bar{x})^2$	$k-1$	$\sigma^2_K + r\sigma^2_A$
誤差	$S_K = \sum^k \sum^r (x_{ij} - \bar{x}_{i.})^2$	$k(r-1)$	σ^2_K
計	$\sum^k \sum^r (x_{ij} - \bar{x})^2$	$kr-1$	

注 S.S. の項を実際に計算するには § 2・3 にのべたように簡略計算を行う。

- 1) 母数模型では分散はこの型になる。
- 2) 不偏分散 = 偏差 2 乗和 / 自由度

b. 変量模型の場合 (因子 A: 対応のある場合)

$$x_{ij} = \mu + a_i + z_{ij} \quad \text{ただし } \bar{a} : a_i \text{ の試料平均}$$

$$\bar{x}_{i.} = \mu + a_i + \bar{z}_{i.}$$

$$\bar{\bar{x}} = \mu + \bar{a} + \bar{z}_{..}$$

$$\begin{aligned} E(S_A) &= rE \sum^k (\bar{x}_{i.} - \bar{\bar{x}})^2 = rE \{ \sum^k (a_i - \bar{a}) + \sum^k (\bar{z}_{i.} - \bar{z}_{..}) \}^2 \\ &= r(k-1)\sigma_a^2 + (k-1)\sigma^2_R \end{aligned}$$

ただし $\sigma^2_A = E \sum^k (a_i - \bar{a})^2 / f_A$ 1) σ^2_R : まえと同じ

$$E(S_R) = E \sum^k \sum^r (x_{ij} - \bar{x}_{i.})^2 = E \sum^k \sum^r (z_{ij} - \bar{z}_{i.})^2 = k(r-1)\sigma^2_R$$

したがって一元配置法変量模型の場合の分散分析表は母数模型の場合と全く同じ。

これらの分散分析表の $E(V_e)$ の構造を考えて、 σ^2_R 、 σ^2_A を推定することになる。たとえば貨車間の分散 σ^2_A 、貨車内のインクリメント間の分散 σ^2_R を推定するには

$$\sigma^2_R = E(V_R)$$

$$\sigma^2_A = \frac{E(V_A) - E(V_R)}{r}$$

注 1) この場合 σ^2_A が σ^2_R に比して小さい場合には、 V_A, V_R はそれからのランダムサンプルとしてもとめているので、 $E(V_A) < E(V_R)$ となり、一見 σ^2_A が「-」となることがある。このときは $\sigma^2_A \hat{=} 0$ とし取り扱って、実用上差支えない。

2) このような推定を行うためには、 σ^2_R 、すなわちたとえば貨車内のインクリメント間のバラツキが管理状態にあることが望ましい。管理状態にないと、推定の結果が信頼しにくくなり、極めて大ザッパな推定を行つていることになる。したがって繰り返しのデータを用いて、R管理図を書いて一応チェックしておくといふ。このことはあとのすべての型の場合と同じ。

3) 分散分析で推定した分散は、自由度が小さい場合が多いので推定の精度はあまりよくない。実験全体を繰り返すとよい。

1) 変量模型では分散はこの型になる。誤差分散は対応のない変量模型である。

4) この一元配置法による解析の例では、 σ^2_R に、縮分の誤差 $\sigma^2_{R'}$ 、分析の誤差 σ^2_M などが入っているから、インクリメント間の真のパラツキを推定するためには、 σ^2_R 、 σ^2_M を引かなければならない。

5) すべての因子が対応のない変量となる場合は、技術的には考えられない。

A 1・3 二元配置法 (繰り返しのない場合)

因子として、すなわちパラツキの要因を2つ考えた場合である。

例 a) 母数模型 原料生産者、 A_1, A_2, \dots, A_k について、ある月の上旬 B_1 、中旬 B_2 、下旬 B_3 についての製品を比較する場合。両因子とも母数模型。

b) 変量模型 A 原料を生産者が多数ある中、ランダムに A_1, A_2, \dots, A_k をえらび、 B 原料も同様に、生産者をランダムに B_1, B_2, \dots, B_l をえらんで実験した場合。両因子とも対応のある変量模型。

c) 混合型 A_1 工程と A_2 工程との比較を、原料が各種あるうちランダムに B_1, B_2, \dots, B_l について比較する場合。 A は母数、 B は対応のある変量模型である。

d) 変量模型 b) において、 A 原料生産者 A_1, \dots, A_k 各々について、ランダムにドラム缶 l 本ずつとって各1回ずつ試験する場合。 A は対応のある変量、 B は対応のない変量。これはドラム缶ごとのパラツキを要因にとっているように一見思われるが、一元配置法と全く同じ型になる。

a. 母数模型の場合

$$x_{ij} = \mu + a_i + b_j + z_{ij}$$

$$\text{ただし } \sum a_i = 0; \sum b_j = 0; E(z_{ij}) = 0$$

$$\bar{x}_{i.} = \mu + a_i + \bar{z}_{i.} \quad \sigma^2_A = \sum a_i^2 / f_A = \sum a_i^2 / k - 1$$

$$\sigma^2_B = \sum b_j^2 / f_B = \sum b_j^2 / l - 1$$

$$\bar{x}_{.j} = \mu + b_j + \bar{z}_{.j}$$

$$\bar{\bar{x}} = \mu + \bar{z}_{..}$$

b. 変量模型の場合 (2因子ともに対応のある場合)

$$x_{ij} = \mu + a_i + b_j + z_{ij}$$

$$\text{ただし } E(a_i) = 0; E(b_j) = 0; E(z_{ij}) = 0$$

表 A 1・3 二元配置法分散分析表 (A: 母数, B: 母数)

要因	S.S.	f	E(V)
A	$S_A = l \sum (\bar{x}_{i.} - \bar{\bar{x}})^2$	$k-1$	$\sigma^2_E + l\sigma^2_A$
B	$S_B = k \sum (\bar{x}_{.j} - \bar{\bar{x}})^2$	$l-1$	$\sigma^2_E + k\sigma^2_B$
E	$S_E = S_0 - S_A - S_B$	$(k-1)(l-1)$	σ^2_E
計	$S_0 = \sum \sum (x_{ij} - \bar{\bar{x}})^2$	$kl-1$	

注) この σ^2_E には交互作用を含んでいる。

$$\bar{x}_{i.} = \mu + a_i + \bar{b} + \bar{z}_{i.} \quad \sigma^2_A = E \sum (a_i - \bar{a})^2 / (k-1)$$

$$\sigma^2_B = E \sum (b_{ij} - \bar{b})^2 / (l-1)$$

$$\bar{x}_{.j} = \mu + \bar{a} + b_j + \bar{z}_{.j}$$

$$\bar{\bar{x}} = \mu + \bar{a} + \bar{b} + \bar{z}_{..}$$

分散分析表は a) と同じになる。

c. 混合型 (A は母数, B は変量 (対応あり))

$$x_{ij} = \mu + a_i + b_j + z_{ij}$$

$$\text{ただし } \sum a_i = 0; E(b_i) = 0; E(z_{ij}) = 0$$

$$\bar{x}_{i.} = \mu + a_i + \bar{b} + \bar{z}_{i.}$$

$$\bar{x}_{.j} = \mu + b_j + \bar{z}_{.j}$$

$$\bar{\bar{x}} = \mu + \bar{b} + \bar{z}_{..}$$

分散分析表は a) と同じになる。

d. 変量模型 (A は対応のある変量, B は対応のない変量)

$$x_{ij} = \mu + a_i + b_{ij} + z_{ij}$$

$$\text{ただし } E(a_i) = 0; E(b_{ij}) = 0; E(z_{ij}) = 0$$

$$\sigma^2_A = E \sum (a_i - \bar{a})^2 / k - 1$$

$$\sigma^2_B = E \sum \sum (b_{ij} - \bar{b}_{i.})^2 / k(l-1) = lE \sum (\bar{b}_{i.} - \bar{b}_{..})^2 / (k-1)$$

$$\bar{x}_{i.} = \mu + a_i + \bar{b}_{i.} + \bar{z}_{i.}$$

$$\bar{\bar{x}} = \mu + \bar{a} + \bar{b}_{..} + \bar{z}_{..}$$

$\bar{x}_{.j}$ は対応がないからとって意味がない。

$$E(S_A) = lE \sum (\bar{x}_{i.} - \bar{x})^2 = l(k-1)\sigma_A^2 + (k-1)\sigma_B^2 + (k-1)\sigma_K^2$$

$$E(S_{B(A)}) = k(l-1)(\sigma_B^2 + \sigma_K^2)$$

対応がないから、 S_B はとって意味がない。

分散分析表は次の通りになる。

表 A 1・4 二元配置法分散分析表

[A: 変量 (対応あり) 又は母数, B: 変量 (対応なし)]

要因	S.S.	f	E(V)
A	$S_A = l \sum (x_{i.} - \bar{x})^2$	$k-1$	$\sigma_K^2 + \sigma_B^2 + l\sigma_A^2$
B(A)	$S_{B(A)} = \sum \sum (x_{ij} - \bar{x}_{i.})^2$	$k(l-1)$	$\sigma_K^2 + \sigma_B^2$
計	$S_0 = \sum \sum (x_{ij} - \bar{x})^2$	$kl-1$	

この表は l を r , $\sigma_K^2 + \sigma_B^2$ を σ^2_K とすれば一元配置法の場合表 A1・1 と全く同じ型になる。すなわち変量模型の対応のない場合は、単なる繰り返しも見られる。これは一元配置法と同じ型で解析すればよい。この場合には、たとえば A をランダムにとった貨車、 B を各貨車からのインクリメントとすれば、 $B(A)$ の項の $\sigma_K^2 + \sigma_B^2$ は、インクリメント間のバラツキ σ_B^2 と、縮分と測定の誤差 σ_K^2 とは分離できない。

注 A が母数模型の場合でも全く同じ型になる。

A 1・4 二元配置法 (繰り返しのある場合)

この場合も、母数模型、変量模型で対応のある場合、対応のない場合のいろいろの組合せがある。

a. 2 因子とも母数模型の場合

サンプリングの調査としては、母数模型と考えて行う場合は少なく、変量模型として考えなければならぬ場合が多いが、いまここに倉庫 1 号、2 号、... k 号 (因子 A) があり、これに大量のロットが入っている。この南側、中央、北側 (因子 B 、一般には l 水準) と層別できるときに、このきまったロットを調査するための予備調査として、各倉庫の各層からランダムに r コずつのサンプルをとり測定した場合がこの例である。この場合各倉庫

ごとの母平均間のバラツキ，南側，北側などと位置もきめて考えられるから，母数として取り扱ってよい¹⁾。

$$x_{ijk} = \mu + a_i + b_j + I_{ij} + z_{ijk}$$

ただし I_{ij} : A と B との交互作用 $\sum_i I_{ij} = 0$; $\sum_j I_{ij} = 0$

x_{ijk} : i 番目の倉庫の j 層の k 番目のランダムサンプルの測定値

$$\bar{x}_{ij.} = \mu + a_i + b_j + I_{ij} + \bar{z}_{ij.}$$

$$\bar{x}_{i..} = \mu + a_i + \bar{z}_{i..}$$

$$\bar{x}_{.j.} = \mu + b_j + \bar{z}_{.j.}$$

$$\bar{x} = \mu + \bar{z} \dots$$

この分散分析表は次の通りである。

表 A 1・5 繰り返しのある二元配置法分散分析表 (A, B : 母数)

要因	S.S.	f	E(V)
A	$S_A = rlE \sum_i (\bar{x}_{i..} - \bar{x})^2$	$k-1$	$\sigma_K^2 + rl\sigma_A^2$
B	$S_B = rkE \sum_j (\bar{x}_{.j.} - \bar{x})^2$	$l-1$	$\sigma_K^2 + rk\sigma_B^2$
$A \times B$	$S_{A \times B} = S_{AB} - S_A - S_B$	$(k-1)(l-1)$	$\sigma_K^2 + r\sigma_{AB}^2$
AB	$S_{AB} = rE \sum_i \sum_j (\bar{x}_{ij.} - \bar{x})^2$	$kl-1$	σ_K^2
E	$S_K = E \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{ij.})^2$	$kl(r-1)$	
計	$S_0 = E \sum_i \sum_j \sum_k (x_{ijk} - \bar{x})^2$	$klr-1$	

$$\begin{aligned} \text{ただし } \sigma_{AB}^2 &= \sum_{ij} (I_{ij} - \bar{I}_{i.} - \bar{I}_{.j} + \bar{I})^2 / f_{A \times B} \\ &= \sum \sum I_{ij}^2 / [(k-1)(l-1)] \end{aligned}$$

注 1) 検定は E の不偏分散により行えばよい。

2) 各分散の推定はデータよりもめた不偏分散を代入して， σ_K^2 ， σ_A^2 ， σ_B^2 ， σ_{AB}^2 などをもとめればよい。

b. A : 母数， B : 変量 (対応のある場合)

1) 一般のサンプリングの予備調査の場合は，このような特定のロットについてのみ問題とする場合は少ない。

ある工程に k 台の機械がある。この機械はおのおのあるきまった特長をもっているで、機械 (A) は母数と考えられる。この機械からの製品を一定時間ごとに l 回、各 r コずつランダムサンプルをとり測定した。一定時間ごとに k 台の各機械からサンプルをとるので、工程は共通的な原因で時間によりランダムにかかわるとすれば、時間 (B) については対応のある変量模型と考えられる。あるいは k 台の機械に、工員中からランダムにえらんだ l 人に運転させ、 r コずつサンプルをとった場合もこの型となる。この場合工員を技術的に考えて l 層にわければ、工員間 (B) は母数模型となる。

$$x_{ijk} = \mu + a_i + b_j + I_{ij} + z_{ijk}$$

$$\text{ただし } \sum a_i = 0 ; E(b_j) = 0 ; E(I_{ij}) = 0^{1)} ; E(z_{ijk}) = 0$$

$$\sigma^2_A = \sum a_i^2 / (k-1) \quad \sigma^2_B = E \sum (b_j - \bar{b})^2 / (l-1)$$

$$\begin{aligned} \sigma^2_{AB} &= E \sum \sum (I_{ij} - \bar{I}_{..})^2 / (kl-1) \\ &= lE \sum (\bar{I}_{i.} - \bar{I}_{..})^2 / (k-1) = kE \sum (\bar{I}_{.j} - \bar{I}_{..})^2 / (l-1) \end{aligned}$$

$$x_{ij.} = \mu + a_i + b_j + I_{ij} + \bar{z}_{ij.}$$

$$x_{i..} = \mu + a_i + \bar{b} + \bar{I}_{i.} + \bar{z}_{i..}$$

表 A 1・6 繰り返しのある二元配置法分散分析表

〔A: 母数 B: 変量 (対応のある場合)〕

要因	S.S.	f	E(V)
A	$S_A = r l \sum (\bar{x}_{i..} - \bar{\bar{x}})^2$	$k-1$	$\sigma^2_K + r\sigma^2_{AB} + rl\sigma^2_A$
B	$S_B = r k \sum (\bar{x}_{.j} - \bar{\bar{x}})^2$	$l-1$	$\sigma^2_K + r\sigma^2_{AB} + rk\sigma^2_B$
A × B	$S_{A \times B} = S_{AB} - S_A - S_B$	$(k-1)(l-1)$	$\sigma^2_K + r\sigma^2_{AB}$
AB	$S_{AB} = r \sum \sum (\bar{x}_{ij.} - \bar{\bar{x}})^2$	$kl-1$	
E	$S_K = S_0 - S_{AB}$ $= \sum \sum \sum (x_{ijk} - \bar{x}_{ij.})^2$	$kl(r-1)$	σ^2_K
計	$S_0 = \sum \sum \sum (x_{ijk} - \bar{\bar{x}})^2$	$klr-1$	

1) 変量の因子に関係ある交互作用はすべて変量模型となる。

$$\bar{x}_{.j} = \mu + b_j + \bar{I}_{.j} + \bar{e}_{.j}$$

$$\bar{\bar{x}} = \mu + \bar{b} + \bar{I}_{..} + \bar{e}...$$

注 1) A, B の不偏分散には σ^2_{AB} という項が入っているので、 σ^2_E で検定するのは間違い。この場合交互作用が技術的に全くないと考えられるときには、 $\sigma^2_{AB}=0$ 、したがって $[S_{A \times B} + S_E] / [\phi_{A \times B} + \phi_E]$ という項により検定すべきである。一般には A, B は $A \times B$ の項を誤差項として検定しなければならない。 $A \times B$ の項は、 E を誤差項として検定すればよい。

2) 分散推定の場合も上の表にしたがって行えばよい。交互作用による分散をサンプリング法決定の際にいかに取り扱うべきかは技術的によく意味を考える必要がある。たとえばこの例では、 B は時間により、 A 、すなわち各機械に共通な影響をあたえる原因によるバラツキを示す。したがって $A \times B$ は時間によって各機械に特有なバラツキをあらわす。そこで各機械から何時間ごとにサンプルをとったらよいかというサンプルの適正配分をするときには、この例では、機械によって時間的にことなつた変動をもっているのだから、 B を対応のない変量模型の型として取り扱うか、あるいは $\sigma^2_{AB} + \sigma^2_B$ を時間によるバラツキと考えて、サンプルを割当てればよいであろう。この場合機械により層内（機械内）の一定時間の間のバラツキがことなつたものがあることを示している。

c. A : 母数 B : 変量 (対応のない場合)

A_1, A_2, \dots, A_k ハッチのある船で石炭が入ったときに、各ハッチごとにランダムに l コのサンプルをとり、これを縮分して両端の 2 つをとり 2 つのサンプル¹⁾をつくり、これを各々 1 回分析した場合がこの例である。その船のハッチ間の差というものを調査対象に考えれば²⁾、ハッチ間は母数に考えられ、各ハッチのサンプル間は互に対応がない変量となる。

$$x_{ijk} = \mu + a_i + b_{ij} + z_{ijk}$$

$$\text{ただし } \sum a_i = 0 ; E(b_{ij}) = 0 ; E(z_{ijk}) = 0$$

$$\sigma^2_A = \sum a_i^2 / (k-1)$$

$$\sigma^2_B = E \sum_{i=1}^k \sum_{j=1}^l (b_{ij} - \bar{b}_{.i})^2 / (kl-1) = lE \sum_{i=1}^k (\bar{b}_{.i} - \bar{b}_{..})^2 / (k-1)$$

1) §A 1-5 f) 図 A 1-3 のように、図 A 1-3 では分析は各 2 回であるが、各 1 回の場合。

2) 次の船のハッチ間というように、一般にあるところから送られてくるというときには、ハッチ間のバラツキを考えると、変量模型となり、d) の型となる。

$$\bar{x}_{ij.} = \mu + a_i + b_{ij} + \bar{z}_{ij.} \quad ; \quad \bar{x}_{i..} = \mu + a_i + \bar{b}_{i.} + \bar{z}_{i..}$$

$$\bar{x} = \mu + \bar{b}_{..} + \bar{z}_{..}$$

注 1) このように対応のない場合は、一般にこの要因と他の要因の交互作用というものは考えても対応のない値であるから全く意味がなく、これは b_{ij} に含まれてしまう。これを $B(A)$ とあらわす。

表 A 1・7 繰り返しのある二元配置法分散分析表
[A: 母数 B: 変量 (対応のない場合)]

要因	S.S.	f	E(V)
A (ハッチ間) B(A) (サンプル間)	$S_A = r l \sum (\bar{x}_{i..} - \bar{x})^2$ $S_{B(A)} = S_{AB} - S_A$	$k-1$ $k(l-1)$	$\sigma_K^2 + r\sigma_{B(A)}^2 + rl\sigma_A^2$ $\sigma_K^2 + r\sigma_{B(A)}^2$
AB (全サンプル間) E	$S_{AB} = r \sum \sum (\bar{x}_{ij.} - \bar{x})^2$ $S_K = S_0 - S_{AB}$ $= \sum \sum \sum (x_{ijk} - \bar{x}_{ij.})^2$	$kl-1$ $kl(r-1)$	σ_K^2
計	$S_0 = \sum \sum \sum (x_{ijk} - \bar{x})^2$	$klr-1$	

2) 検定は、ハッチ間を検定するには $B(A)$ により、サンプル間 $B(A)$ を検定するには、 E により行う。

3) 推定も上表にしたがって行う。この場合 $B(A)$ はサンプル間のバラツキを示し、交互作用を考えていないから、問題はない。

4) この例では σ_K^2 は縮分誤差と測定誤差を加えたものである。

d. A: 変量 (対応のある場合) B: 変量 (対応のない場合)

たとえば、入荷する鉱石を、貨車を k 台ランダムにとり、それから各々 l コのサンプルをとり、各別々に縮分して r コずつのサンプルをつくり、これを各 1 回測定した場合、倉庫にある大箱入の製品を、大箱を k 個ランダムにとり、その中から l 個ずつ小箱をランダムにとり、小箱中から r 個の製品を各ランダムにとりだし 1 回ずつ測定した場合、などがこれである。この型はサンプリングの調査のときに多い。

$$x_{ijk} = \mu + a_i + b_{ij} + z_{ijk} \quad \text{ただし } E(a_i) = 0 \text{ 以外は c) と同じ}$$

$$\bar{x}_{4j} = \mu + a_4 + b_{4j} + \bar{z}_{4j}$$

$$\bar{x}_{4..} = \mu + a_4 + \bar{b}_{4.} + \bar{z}_{4..}$$

$$\bar{\bar{x}} = \mu + \bar{a} + \bar{b}_{..} + \bar{z}_{...}$$

結果は全く c) と同じ型になり、検定法も推定も同じ。

注 以上の例でわかるように、対応のない変量の因子を考えるという事は、繰り返しと同じような意味になる。したがって二元配置法で繰り返しのある場合は、二元配置法にさらに因子 c を考え、 c が対応のない変量模型と考えれば、すべて三元配置法の繰り返しのない場合と同じになる。

e. A, B : 変量 (対応のある場合)

たとえば、例 1) 紙や鉄板など帯状のものから、図 A 1・1 のようにランダムに縦方向に巾 10cm に k 箇所、次にその中から横方向に巾 10cm に

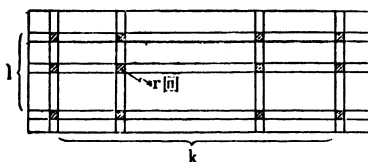


図 A 1・1

各同じ場所を l 箇所サンプリングした場合に、その 10cm 角のものを r 回測定した場合は $\S A1 \cdot 3 b)$ である。これを 1 回測定した場合は $\S A1 \cdot 3 b)$ である。

例 2) または k 枚をランダムサンプリングし、その各々か

ら縦方向に l 箇所から同じ場所をサンプリング、さらにその紐状体からランダムに r 箇所からサンプルをとった場合、例 3) あるいは糸などのポピンが沢山あるときに、ポピンをランダムにとり、各ポピンより、端から各同じ長さのところを l 箇所とり、これを r 回試験した場合もこれである。これらの場合において 2 つ目の因子 B 、すなわち例 1) 横方向、例 2) 縦方向、あるいは例 3) ポピンの端から l 箇所とるといふ因子は、考えかたによっては対応のない変量模型ともいえる。これはその材料に対する技術的知識をいかにもっているかにより、あるいは調査の目的によりきまる問題で、統計学の問題ではない。すなわち B について技術的にランダムに変化していないで、例 1) では横方向に例 2) では縦方向に、例 3) ではポピンの内外により同じような周期や傾向で変化していると技術的に想像されるときは、対応のある変量とみなし、これらがランダムにかわっ

ているとみなせるときは対応のない変量として、解析すべきである。

因子 B が対応ある変量と考えられる場合の分散分析表は b) と全く同じ型になる。

注 実際には、このような型を考えることは少なく、d) として考える場合が多い。

A 1.5 多元配置法

多元配置法は全く二元配置法の延長であるから、因子の構造に関する考え方をそのまま延長すればよい。

この場合は交互作用の取り扱い方に注意を要する。

次にいくつかの型の結果のみを示す。

a. 三元配置法 (繰り返しのない場合) ($A \cdot B \cdot C$: 母数)

表 A 1.8 分散分析表 (A, B, C : 母数)

要因	$S.S.$	f	$E(V)$
A	$S_A = ml \sum (\bar{x}_{i..} - \bar{\bar{x}})^2$	$k-1$	$\sigma_K^2 + ml\sigma^2_A$
B	$S_B = km \sum (\bar{x}_{.j.} - \bar{\bar{x}})^2$	$l-1$	$\sigma_K^2 + km\sigma^2_B$
C	$S_C = kl \sum (\bar{x}_{..k} - \bar{\bar{x}})^2$	$m-1$	$\sigma_K^2 + kl\sigma^2_C$
$A \times B$	$S_{A \times B} = S_{AB} - S_A - S_B$	$(k-1)(l-1)$	$\sigma_K^2 + m\sigma^2_{AB}$
$A \times C$	$S_{A \times C} = S_{AC} - S_A - S_C$	$(k-1)(m-1)$	$\sigma_K^2 + l\sigma^2_{AC}$
$B \times C$	$S_{B \times C} = S_{BC} - S_B - S_C$	$(l-1)(m-1)$	$\sigma_K^2 + k\sigma^2_{BC}$
$A \times B \times C$	$S_{A \times B \times C} = S_0 - S_A - S_B - S_C - S_{A \times B} - S_{B \times C} - S_{C \times A}$	$(k-1)(l-1)(m-1)$	σ_K^2
計	$S_0 = \sum \sum \sum (x_{ijk} - \bar{\bar{x}})^2$	$klm-1$	

b. 三元配置法 (繰り返しのない場合) (C : 対応のない変量)

この場合は、 A, B がいかなる構造と考えるかによつて、すべて § A1.4 の A, B のその構造に相当するものと全く同じとなる。たとえば A が対応のある変量、 B が対応のない変量ならば、§ A 1.4 d) と同じ。

c. 三元配置法 (繰り返しのない場合) (A, B, C : 全部対応のある変量模型)

表 A 1・9 分散分析表 (A, B, C: 対応のある変量)

要因	S.S.	f	E(V)
A	$S_A = ml \sum (\bar{x}_{i..} - \bar{\bar{x}})^2$	k-1	$\sigma^2_K + m\sigma^2_{AB} + l\sigma^2_{AC} + lm\sigma^2_A$
B	$S_B = km \sum (\bar{x}_{.j.} - \bar{\bar{x}})^2$	l-1	$\sigma^2_K + k\sigma^2_{BC} + m\sigma^2_{AB} + km\sigma^2_B$
C	$S_C = kl \sum (\bar{x}_{..k} - \bar{\bar{x}})^2$	m-1	$\sigma^2_K + l\sigma^2_{AC} + k\sigma^2_{BC} + kl\sigma^2_C$
A × B	$S_{A \times B} = S_{AB} - S_A - S_B$	(k-1)(l-1)	$\sigma^2_K + m\sigma^2_{AB}$
A × C	$S_{A \times C} = S_{AC} - S_A - S_C$	(k-1)(m-1)	$\sigma^2_K + l\sigma^2_{AC}$
B × C	$S_{B \times C} = S_{BC} - S_B - S_C$	(l-1)(m-1)	$\sigma^2_K + k\sigma^2_{BC}$
A × B × C	$S_{A \times B \times C} = S_0 - S_A - S_B - S_C - S_{A \times B} - S_{B \times C} - S_{A \times C}$	(k-1)(l-1)(m-1)	σ^2_K
計	$S_0 = \sum \sum \sum (x_{ijk} - \bar{\bar{x}})^2$	klm-1	

注] この場合には厳密な検定法はない。1)

d. 四元配置法 [A: 変量 (対応あり), B, C, D: 変量 (対応なし)]

たとえば、いくつかの炭坑から石炭が入るときに、全体に共通サンプリング法をきめようという場合である。まず炭坑 (A) を k 坑ランダムにえらび、次に入ってくる列車 (B) をおのおの l 列車ずつランダムにえらび、さらに各列車から貨車 (C) を m 車とり、その中からインクリメント (D) を r コとり各 1 回分析する場合はこれである。これは上と同じ条件で A, B, C の 3 因子をとつた繰り返しのある三元配置法と全く同じである。各インクリメントからのデータ x_{ijkl} は

$$x_{ijkl} = \mu + a_i + b_{ij} + c_{ijk} + d_{ijkl} + \varepsilon_{ijkl}$$

ただし a_i : 炭坑によるバラツキ $E(a_i) = 0$

b_{ij} : 各炭坑内での列車によるバラツキ $E(b_{ij}) = 0$

c_{ijk} : 各列車内での貨車によるバラツキ $E(c_{ijk}) = 0$

d_{ijkl} : 各貨車内でのインクリメントによるバラツキ

$$E(d_{ijkl}) = 0$$

1) 前記増山: 日化プリント参照.

x_{ijkl} : 縮分, 分析によるバラツキ $E(x_{ijkl}) = 0$

$$\bar{x}_{ijk.} = \mu + a_i + b_{ij} + c_{ijk} + \bar{d}_{ijk.} + \bar{z}_{ijk.}$$

$$\bar{x}_{ij..} = \mu + a_i + b_{ij} + \bar{c}_{ij.} + \bar{d}_{ij..} + \bar{z}_{ij..}$$

$$\bar{x}_{i...} = \mu + a_i + \bar{b}_{i.} + \bar{c}_{i...} + \bar{d}_{i...} + \bar{z}_{i...}$$

$$\bar{x} = \mu + \bar{a} + \bar{b} + \bar{c} + \bar{d} + \bar{z}$$

$$E(S_A) = lmn E \sum_k (\bar{x}_{i...} - \bar{x})^2$$

$$E(S_{AB}) = mn E \sum_k \sum_l (\bar{x}_{ij..} - \bar{x})^2 \quad E(S_{B(A)}) = E(S_{AB} - S_A)$$

などよりもとめると, 分散分析表は次の通りになる.

表 A 1・10 分散分析表

[A: 変量 (対応あり), B, C, D: 変量 (対応なし)]

要因	S.S.	f	E(V)
A	$S_A = lmn \sum_k (\bar{x}_{i...} - \bar{x})^2$	$k-1$	$\sigma_K^2 + \sigma_D^2 + r\sigma_C^2 + nr\sigma_B^2 + lmr\sigma_A^2$
B(A)	$S_{B(A)} = S_{AB} - S_A$	$k(l-1)$	$\sigma_K^2 + \sigma_D^2 + r\sigma_C^2 + nr\sigma_B^2$
C(AB)	$S_{C(AB)} = S_{ABC} - S_{AB}$	$kl(m-1)$	$\sigma_K^2 + \sigma_D^2 + r\sigma_C^2$
D(ABC)	$S_{D(ABC)} = S_0 - S_{ABC}$	$klm(r-1)$	$\sigma_K^2 + \sigma_D^2$
計	$S_0 = E \sum_k \sum_l \sum_m \sum_r (x_{ijkl} - \bar{x})^2$	$klmr-1$	

§ A 1・4 c) d) を一因子増した型となる. その項の注を参照. この例では $V_{D(ABC)}$ は, インクリメント間, 縮分, 分析の誤差を含んだものである.

e. 縮分の場合 1. (各因子とも条件つき対応のない変量の多元配置法)

縮分は通常何段階にも行つて, 1 トンの鉾石から分析試料 1g をとつている. この各段階の縮分工程を各々因子とし, そのバラツキ (誤差) をもとめ, あるいは検定するときは一種の特別な変量模型の多元配置法となる.

いま図 A 1・2 のように円錐四分法あるいはリップルサンプラーなどで

何段階か縮分したものをすべてとつておき、各々縮分していつて、分析用試料をつくり、これからの分析試料各 1g をとり分析する場合を考える。

ここでは 4 回縮分して、16 コの分析用試料をえて、これを各 2 回ずつ分析した場合を考えよう。もちろん縮分の順序はランダムにし、32 回の分析は、分析者にランダムに割りあて、時間的にもランダムに行わなければならない。分析データを、 x_{ijklm} であらわすと

$$x_{ijklm} = \mu + a_i + b_{ij} + c_{ijk} + d_{ijkl} + z_{ijklm}$$

ただし a_i : 第 1 段の縮分 (A) によるバラツキ $\sum a_i = 0$ 2 水準

$$b_{ij} : \text{第 2 段の縮分 (B) によるバラツキ } \sum b_{ij} = 0 \quad "$$

$$c_{ijk} : \text{第 3 段の縮分 (C) によるバラツキ } \sum c_{ijk} = 0 \quad "$$

$$d_{ijkl} : \text{第 4 段の縮分 (D) によるバラツキ } \sum d_{ijkl} = 0 \quad "$$

$$z_{ijklm} : \text{分析によるバラツキ } E(z_{ijklm}) = 0 \quad "$$

縮分という工程は 2 つにわかれる、わかれ方は全くランダムで変量である。しかしそのわかれたことによるバラツキは、たとえば一方に灰分が a_i 増えれば、他方では必ず a_i へるという特別な条件をもっている。すなわち母集団全体をランダムに 2 つにわけける工程である。したがつてどの段階においても、 $\sum a_i = 0$ というような条件が成立する。これはちやうど母数模型における条件と同じであるので、分散分析表も母数模型とよく似ているが、対応のない変量であるから交互作用はない。

$$\bar{x}_{ijkl.} = \mu + a_i + b_{ij} + c_{ijk} + d_{ijkl} + \bar{z}_{ijkl.}$$

$$\bar{x}_{ijk..} = \mu + a_i + b_{ij} + c_{ijk} + \bar{z}_{ijk..}$$

$$\bar{x}_{ji...} = \mu + a_i + b_{ij} + \bar{z}_{ji...}$$

$$\bar{x}_i.... = \mu + a_i + \bar{z}_i....$$

$$\bar{\bar{x}} = \mu + \bar{z}....$$

注 このような構造を考えていることは、縮分した各部分の重量のバラツキを無視している。実際にはこの影響はほとんど問題にならない。

表 A 1・11 縮分の分散分析表 (各部を全部分析した場合)

要因	S.S.	水準	f	E(V)
1 段目縮分 A	$S_A = 2^4 \sum (\bar{x}_i \dots - \bar{x})^2$	2	1	$\sigma^2_E + 32\sigma^2_A$
2 " B(A)	$S_{B(A)} = S_{AB} - S_A$	2	2	$\sigma^2_E + 16\sigma^2_B$
3 " C(AB)	$S_{C(AB)} = S_{ABC} - S_{AB}$	2	4	$\sigma^2_E + 8\sigma^2_C$
4 " D(ABC)	$S_{D(ABC)} = S_{ABCD} - S_{ABC}$	2	8	$\sigma^2_E + 4\sigma^2_D$
分析誤差 E	$S_{E(ABCD)} = S_0 - S_{ABCD}$	2	16	σ^2_E
計	$S_0 = \sum \sum \sum \sum (x_{ijklm} - \bar{x})^2$		31	

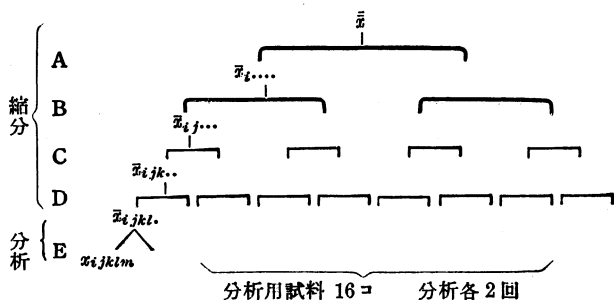


図 A 1・2 縮分法

注 1) このようにすればどの段階の縮分誤差が大きいか、検定もできるし、各縮分誤差を推定もできる。全縮分誤差 σ^2_R は

$$\sigma^2_R = \sigma^2_A + \sigma^2_B + \sigma^2_C + \sigma^2_D$$

2) 全縮分誤差をもとめるには、次のようにしてもよい。

$$E(S_{ABCD}) = 15\sigma^2_E + 32(\sigma^2_A + \sigma^2_B + \sigma^2_C + \sigma^2_D) = 15\sigma^2_E + 32\sigma^2_R$$

したがって $\hat{\sigma}^2_R = (S_{ABCD} - 15\hat{\sigma}^2_E) / 16$

縮分 3 回目までの誤差 $\sigma^2_{R_3}$ ならば、同様にして

$$E(S_{ABC}) = 7\sigma^2_E + 32\sigma^2_{R_3}$$

3) この例では σ^2_E は 1 サンプルを、人と時間をランダムにして測定しているので、§ 3・4・2 にのべた、同一研究所内における再現精度

である。この場合、1 サンプルの2 回の分析を同一人に平行にやらせた場合は σ^2_R は1 人が同時に行った繰返精度になり、他の人、他の時に行ったときのカタヨリは、 σ^2_D などに入ってしまう σ^2_R を大きく推定することになる。したがって同一人が同時に全部行える場合は、 σ^2_R には測定の影響が入らぬからよいが、そうでないときには、各サンプルに測定のカタヨリが入らぬように、人、場所、時間など一定にできない条件をランダム化しておかなければならない。実験や測定順序のランダム化は、サンプリング調査のときにもたいせつである。

4) この場合 σ^2_R に対して $A, B(A), C(AB), D(ABU)$ を検定して、有意差があれば、その段階を検討すべきである。また注 2) に記したように $S_{ABCD}/15$ ($f=15$) に対して検定すれば、縮分全体の誤差の有意差が検定できる。

5) 縮分法が $1/2$ ずつにわけず、一回で $1/10$ とし、その各縮分段階で2 ずつサンプルをとるときは、A 1・5 d) の型となる。

例) 4 メッシュ以下の石炭 1 kg を、間隔 10 mm のリップルサンブラーで、4 回縮分して、約 62.5 g ずつのサンプル16 コをえて、これを各 2 回ずつ灰分を分析して表 A 1・12 のようなデータをえた。

計算手順 1) 全体のデータから 15.2 を引き、100 倍して補助表をつくる(計算を便にするため、表を上下逆にしてある)。

$$2) CF = (151)^2/32 = 712.5$$

$$3) S_0 = 6^2 + (-30)^2 + 22^2 + \dots + (-69)^2 + (-56)^2 - CF \\ = 37,766.5$$

$$4) S_A = \frac{1}{16} \{(-13)^2 + (-138)^2\} - CF = 488.3$$

$$5) S_{AB} = \frac{1}{8} \{(-43)^2 + 30^2 + (-132)^2 + (-6)^2\} - CF = 1,813.6$$

$$6) S_{B(A)} = S_{AB} - S_A = 1,325.3$$

$$7) S_{ABC} = \frac{1}{4} \{(-27)^2 + (-16)^2 + \dots + 84^2 + (-90)^2\} - CF \\ = 8,138.3$$

$$8) S_{C(AB)} = S_{ABC} - S_{AB} = 6,324.7$$

表 A 1・12 縮分実験データ (石炭の灰分, 単位%)

A ₁								A ₂							
B ₁₁				B ₁₂				B ₂₁				B ₂₂			
C ₁₁₁		C ₁₁₂		C ₁₂₁		C ₁₂₂		C ₂₁₁		C ₂₁₂		C ₂₂₁		C ₂₂₂	
D ₁₁₁₁	D ₁₁₁₂	D ₁₁₂₁	D ₁₁₂₂	D ₁₂₁₁	D ₁₂₁₂	D ₁₂₂₁	D ₁₂₂₂	D ₂₁₁₁	D ₂₁₁₂	D ₂₁₂₁	D ₂₁₂₂	D ₂₂₁₁	D ₂₂₁₂	D ₂₂₂₁	D ₂₂₂₂
15.26	15.42	15.45	15.00	15.69	15.39	14.84	15.12	14.56	15.08	14.67	15.62	14.93	15.76	15.20	14.51
14.90	14.95	15.09	15.10	15.34	15.24	14.98	15.30	15.65	14.88	15.20	14.62	15.31	15.64	15.55	14.64

表 A 1・13 補 助 表

6	22	25	-20	49	19	-36	-8	-64	-12	-53	42	-27	56	0	-69
-30	-25	-11	-10	14	4	-22	10	45	-32	0	-58	11	44	35	-56
-24	-3	14	-30	63	23	-58	2	-19	-44	-53	-16	-16	100	35	-125
-27		-16		86		-56		-63		-69		84		-90	
-43				30				-132				-6			
-13								-138							
-151															

- 9) $S_{ABCD} = \frac{1}{2}\{(-24)^2 + (-3)^2 + \dots + 35^2 + (-125)^2\} - CF$
 $= 20,295.0$
- 10) $S_{D(AEC)} = S_{ABCD} - S_{ABC} = 12,156.7$
- 11) $S_R = S_0 - S_{ABCD} = 17,471.5$
- 12) 分散分析表をつくる.

表 A 1・14 分散分析表

要 因	S. S.	f	V
A	$S_A = 488.3$	1	488.3
B(A)	$S_{B(A)} = 1,325.3$	2	662.6
C(AB)	$S_{C(AB)} = 6,324.7$	4	1,581.2
D(ABC)	$S_{D(ABC)} = 12,156.7$	8	1,519.6
E	$S_R = 17,471.5$	16	1,092.0
計	$S_0 = 37,766.5$	31	

- 13) 検定 各要因とも有意差は認められない.
- 14) 全体の縮分誤差に対して検定する.

$$S_{ABCD}/f_{ABCD} = \sigma^2_R + \frac{16}{15}(\sigma^2_A + \sigma^2_B + \sigma^2_C + \sigma^2_D)$$

$\sigma^2_R = \sigma^2_A + \sigma^2_B + \sigma^2_C + \sigma^2_D = 0$ という仮説をたてれば

$$\therefore F_0 = \frac{S_{ABCD}/f_{ABCD}}{S_R/f_R} \doteq 1.24 \quad (f_1 = 15, f_2 = 16)$$

しかるに $F_{15,16}(0.05) = 2.35$ F_0 はこれより小さいから、仮説 $\sigma^2_R = 0$ は棄却されない.

15) 推定 縮分誤差が有意に大きいとは認められなかったが、その大きさを推定して見る.

$$\left. \begin{aligned} \hat{\sigma}^2_A &= (V_A - V_E)/32 \doteq 0 \\ \hat{\sigma}^2_B &= (V_{B(A)} - V_E)/16 \doteq 0 \end{aligned} \right\} 0 \text{ より小さいから } 0 \text{ とみなす.}$$

$$\hat{\sigma}^2_C = (V_{C(AB)} - V_E)/8 \doteq 56 \quad \text{もとの単位になおして } \hat{\sigma}^2_C = 0.0056$$

$$\hat{\sigma}^2_D = (V_{D(ABC)} - V_E)/2 \doteq 107 \quad \hat{\sigma}^2_D = 0.0107$$

$$\hat{\sigma}^2_R = V_R = 1092$$

$$\hat{\sigma}^2_R = 0.109$$

$$\hat{\sigma}^2_R = (S_{ABCD} / f_{ABCD} - V_R) \times \frac{15}{32} = 123$$

$$\hat{\sigma}^2_R = 0.0123$$

16) 結論 i) この場合の誤差分散は再現精度であるが、これがもつとも大きいようである。

ii) この実験では縮分誤差は大して問題にする必要ない程小さい。しかし、サンプルの量がへってくると、縮分誤差が大きくなるようである。

注] この実験は縮分と測定の順序はランダムに行つた。

f. 縮分の場合 2

e) のように、縮分をすべてやり、分析するという事は、非常に時間と経費を要することであるから、もつと簡単に一部をすてた場合を考えよう。

図 A 1.3 に示したように 1 段目の縮分のみは両側をとり、次段階からは、縮分の方の部分をランダムにすてることにする。図では説明上外側のみとつてあるが、何れの部分をとるかはランダムにきめる。そして最後に 2 コの分析用試料をえて、これを各 2 回ずつ計 4 回の分析を、人や時間をランダムにかえて行う。(σ_R は再現精度となる)

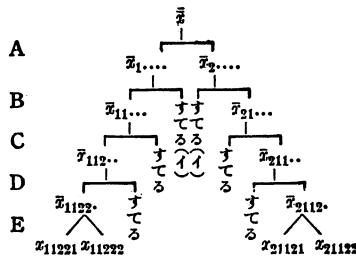


図 A 1.3 縮分法 (一部すてる場合)

$$x_{ijklm} = \mu + a_i + b_{ij} + c_{ijk} + d_{ijkl} + z_{ijklm}$$

ただし $a_i, b_{ij}, c_{ijk}, d_{ijkl}, z_{ijklm}$ は e) の場合と同じパラツキをあらわす。

$$\sum a_i = 0 \quad E(b_{ij}) = 0 \quad E(c_{ijk}) = 0 \quad E(d_{ijkl}) = 0 \quad E(z_{ijkl}) = 0$$

すなわち $b_{ij}, c_{ijk}, d_{ijkl}$ は片方をランダムにとつているので、期待値が 0 という、対応のない変量模型の普通の型となる。

$$\left. \begin{aligned} \bar{x}_{ijkl} &= \mu + a_i + b_{ij} + c_{ijk} + d_{ijkl} + \bar{z}_{ijkl} \\ \bar{x}_{ijk..} &= \mu + a_i + b_{ij} + c_{ijk} + d_{ijkl} + \bar{z}_{ijkl} \end{aligned} \right\}$$

$$\left. \begin{aligned} \bar{x}_{ijk\dots} &= \mu + a_i + b_{ij} + c_{ijk} + d_{ijkl} + \bar{x}_{ijkl} \\ \bar{x}_{i\dots} &= \mu + a_i + b_{ij} + c_{ijk} + d_{ijkl} + \bar{x}_{ijkl} \end{aligned} \right\} \text{全部等しい}$$

$$\bar{\bar{x}} = \mu + \bar{b}_{..} + \bar{c}_{...} + \bar{d}_{....} + \bar{\bar{x}}\dots$$

すなわち $\bar{x}_{ijkl\dots} = \bar{x}_{ijk\dots} = \bar{x}_{ij\dots} = \bar{x}_{i\dots}$

$$E(S_A) = 2E \sum (\bar{x}_{i\dots} - \bar{\bar{x}})^2 = \sigma^2_R + 2\sigma^2_D + 2\sigma^2_C + 2\sigma^2_B + 4\sigma^2_A$$

これ以外に縮分のバラツキはもとめられない。

したがって分散分析表は、次の通りになる。

表 A 1・15 分散分析表

要因	S.S.	f	E(V)
縮分 分析	$S_A = 2 \sum (\bar{x}_{ijkl\dots} - \bar{\bar{x}})^2$	1	$\sigma^2_R + 2(\sigma^2_D + \sigma^2_C + \sigma^2_B + 2\sigma^2_A)$
	$S_R = S_0 - S_A$	2	σ^2_R
計	$S_0 = \sum \sum (x_{ijklm\dots} - \bar{\bar{x}})^2$	3	

注 1) このような実験をすれば、比較的簡単に測定の変差と、縮分の変差 σ^2_R とを検定、および分離して推定できる。

$$\sigma^2_R \approx 2\sigma^2_A + \sigma^2_B + \sigma^2_C + \sigma^2_D$$

しかし、縮分の変差を各縮分工程ごとに分離することはできない。

2) この例において、図 A 1・3 において、イ) のところをすてずに B の片側からもう 1 つの部分を取り、これを同様に縮分してもう 1 つずつのサンプルをつくり、これを各 2 回ずつ、全体で 8 回の分析をランダムに行った場合には、 $E(b_{ij}) = 0$ という関係が $\sum b_{ij} = 0$ となり、次のようなかわった分散分析表となる。

検定は σ^2_R を誤差分散にとって行うべきであるが、何れが有意と出ても、どの段階の縮分の変差が大きいのかはつきりわからぬ。

表 A 1・16 分散分析表

要因	S.S.	f	E(V)
A	$S_A = 4 \sum_i (\bar{x}_{i..} - \bar{x})^2$	1	$\sigma^2_E + 2\sigma^2_D + 2\sigma^2_C + 8\sigma^2_A$
B(A)	$S_{B(A)} = S_{ABCD} - S_A$	2	$\sigma^2_E + \sigma^2_D + \sigma^2_C + 4\sigma^2_B$
ABCD	$S_{ABCD} = 2 \sum_i \sum_j (\bar{x}_{ij..} - \bar{x})^2$	3	$\sigma^2_E + 2(\sigma^2_D + \sigma^2_C) + \frac{8}{3}(\sigma^2_B + \sigma^2_A)$
E	$S_E = \sum_i \sum_j \sum_k \sum_l \sum_m (x_{ijklm} - \bar{x}_{ijkl.})^2$	4	σ^2_E
計	$S_0 = \sum_i \sum_j \sum_k \sum_l \sum_m (x_{ijklm} - \bar{x})^2$	7	

A 1・6 ラテン方格法

ラテン方格法は、A, B, C, 3種の因子をとりあげることができる。繰り返しを行うときには、通常は別の型の方格を用い、Dという因子を入れて実験を計画するが、測定や分析を2回繰り返したときは、同じ型のままのデータが入ることがある。

a. 繰り返しのない場合

A, B, C, いずれも母数模型と考えられるときは、1辺pの、すなわち各p水準のラテン方格を用いれば、分散分析表は次のとおりになる。

表 A 1・17 ラテン方格分散分析表 (繰り返しのない場合)

要因	S.S.	f	E(V)
A	$S_A = p \sum_i (\bar{x}_{i..} - \bar{x})^2$	p-1	$\sigma^2_E + p\sigma^2_A$
B	$S_B = p \sum_j (\bar{x}_{.j.} - \bar{x})^2$	p-1	$\sigma^2_E + p\sigma^2_B$
C	$S_C = p \sum_k (\bar{x}_{..k} - \bar{x})^2$	p-1	$\sigma^2_E + p\sigma^2_C$
E	$S_E = S_0 - S_A - S_B - S_C$	(p-1)(p-2)	σ^2_E
計	$S_0 = \sum (x_{ijk} - \bar{x})^2$	p ² -1	

注 1) いずれかの因子が対応のない変量模型のときには、二元配置

法の繰り返しのない場合と同じになる。このときはラテン方格にはならぬ。

2) A, B, C のいずれかの因子, あるいは全部の因子が対応のある変量模型の場合にも, 分散分析表はいずれも母数模型の場合と同じになる。

b. 同じ型で繰り返しのある場合

たとえばラテン方格でえたサンプルを各 r 回ずつ分析した場合は次のようになる。

表 A 1・18 同じ型で繰り返しのあるラテン方格

要 因	$S.S.$	f	$E(V)$
A	$S_A = rp \sum (\bar{x}_{i...} - \bar{\bar{x}})^2$	$p-1$	$\sigma^2_M + r\sigma^2_K + rp\sigma^2_A$
B	$S_B = rp \sum (\bar{x}_{.j..} - \bar{\bar{x}})^2$	$p-1$	$\sigma^2_M + r\sigma^2_K + rp\sigma^2_B$
C	$S_C = rp \sum (\bar{x}_{..k.} - \bar{\bar{x}})^2$	$p-1$	$\sigma^2_M + r\sigma^2_K + rp\sigma^2_C$
E	$S_E = S_{ABC} - S_A - S_B - S_C$	$(p-1)(p-2)$	$\sigma^2_M + r\sigma^2_K$
ABC	$S_{ABC} = r \sum \sum (\bar{x}_{ijk.} - \bar{\bar{x}})^2$	p^3-1	
M	$S_M = \sum \sum \sum (x_{ijkl} - \bar{x}_{ijk.})^2$	$p^3(r-1)$	σ^2_M
計	$S_0 = \sum \sum \sum (x_{ijkl} - \bar{\bar{x}})^2$	p^3r-1	

注 1) A, B, C を検定するには E を誤差項として, E を検定するには測定誤差を誤差項として検定する。

2) 新たに因子 D を入れるときには, 通常同じ型を用いるとカタヨリが入ったりするので, 別の型のラテン方格により実験を行う。

c. 別の型で繰り返しのある場合

D という因子を入れ, D_1, D_2, \dots, D_r という条件でランダムにえらんだ r コのラテン方格を用いて実験した場合は次のようになる。

表 A 1・19 A, B, C, D (いずれも母数模型の場合)

要因	S.S.	f	E(V)
A	$S_A = rp \sum (\bar{x}_i \dots - \bar{\bar{x}})^2$	$p-1$	$\sigma^2_E + rp\sigma^2_A$
B	$S_B = rp \sum (\bar{x}_{.j.} - \bar{\bar{x}})^2$	$p-1$	$\sigma^2_E + rp\sigma^2_B$
C	$S_C = rp \sum (\bar{x}_{..k} - \bar{\bar{x}})^2$	$p-1$	$\sigma^2_E + rp\sigma^2_C$
D	$S_D = p^2 \sum (\bar{x}_{...m} - \bar{\bar{x}})^2$	$r-1$	$\sigma^2_E + p^2\sigma^2_D$
A × D	$S_{A \times D} = S_{AB} - S_A - S_D$	$(p-1)(r-1)$	$\sigma^2_E + p\sigma^2_{A \times D}$
B × D	$S_{B \times D} = S_{BD} - S_B - S_D$	$(p-1)(r-1)$	$\sigma^2_E + p\sigma^2_{B \times D}$
C × D	$S_{C \times D} = S_{CD} - S_C - S_D$	$(p-1)(r-1)$	$\sigma^2_E + p\sigma^2_{C \times D}$
E	$S_E = S_0 - S_A - S_B - S_C - S_{A \times D} - S_{B \times D} - S_{C \times D}$	$r(p-1)(p-2)$	σ_{K^2}
計	$S_0 = \sum \sum \sum (x_{ijkl} - \bar{\bar{x}})^2$	p^2r-1	

表 A 1・20 A: 変量 (対応あり); B, C, D: 母数模型の場合

要因	S.S.	f	E(V)
A	(A 1・19 表と同じ)	$p-1$	$\sigma^2_E + p\sigma^2_{A \times D} + pr\sigma^2_A$
B		$p-1$	$\sigma^2_E + pr\sigma^2_B$
C		$p-1$	$\sigma^2_E + pr\sigma^2_C$
D		$r-1$	$\sigma^2_E + p\sigma^2_{A \times D} + pr\sigma^2_D$
A × D		$(p-1)(r-1)$	$\sigma^2_E + p\sigma^2_{A \times D}$
B × D	$(p-1)(r-1)$	$\sigma^2_E + p\sigma^2_{B \times D}$	
C × D	$(p-1)(r-1)$	$\sigma^2_E + p\sigma^2_{C \times D}$	
E		$r(p-1)(p-2)$	σ^2_E
計		p^2r-1	

附 表

附表 1. 正規分布の性質

“	μ ± 1σ 内に 入る確率 1-2P	μ ± 1σ の両側 外にでる確率 2P	μ ± 1σ の片側 外にでる確率 P	確率密度高さ
$\left(= \frac{x-\mu}{\sigma} \right)$				
	0	0.00	1.00	0.50
0.50	.33	.62	.32	0.352
0.6745	.50	.50	.25	—
1.00	.68	.32	.16	0.242
1.28	.80	.20	.10	—
1.50	.87	.13	.07	0.130
1.645	.90	.10	.05	—
1.96	.95	.05	.025	—
2.00	.954	.046	.023	0.054
2.33	.98	.02	.01	—
2.50	.988	.012	.006	0.0175
2.58	.990	.010	.005	—
2.81	.995	.005	.0025	—
3.00*	.9973	.0027*	.00135	0.0044
3.09	.998	.002	.001	—
3.29	.999	.001	.0005	—
4.00	.9999	.0001	.0000	0.0001

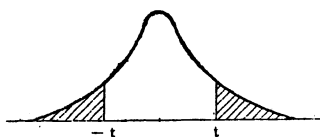
備考 * ±3σ 外にでる確率は実用的には 0 と見て差支えない。

附表 2. 管理関係数表 (3 シグマ)

群の大きさ	X̄ 管理図		R 管理図				s と σ̂ の関係		R と σ̂ の関係		σ と σ _R の関係
	A	A ₂	D ₁	D ₂	D ₃	D ₄	e ₂	1/e ₂	d ₂	1/d ₂	d ₃
2	2.12	1.88	0	3.69	0	3.27	0.564	1.773	1.13	0.896	0.853
3	1.73	1.02	0	4.35	0	2.58	0.724	1.382	1.69	0.591	0.889
4	1.50	0.73	0	4.70	0	2.28	0.798	1.253	2.06	0.486	0.880
5	1.34	0.58	0	4.92	0	2.12	0.841	1.189	2.33	0.430	0.864
6	1.23	0.48	0	5.08	0	2.00	0.869	1.151	2.53	0.395	0.848
7	1.13	0.42	0.21	5.20	0.08	1.92	0.898	1.126	2.70	0.370	0.833
8	1.06	0.37	0.39	5.31	0.14	1.86	0.903	1.108	2.85	0.351	0.820
9	1.00	0.34	0.55	5.39	0.18	1.82	0.914	1.094	2.97	0.337	0.803
10	0.95	0.31	0.69	5.47	0.22	1.78	0.923	1.084	3.08	0.325	0.797

公式集 \bar{X} 管理図: $\bar{x} \pm A\sigma$; $\bar{x} \pm A_2\bar{R}$
 R 管理図: $D_2\sigma$, $D_1\sigma$; $D_4\bar{R}$, $D_3\bar{R}$
 X 管理図: $\bar{x} \pm \sqrt{n}A\sigma$; $\bar{x} \pm \sqrt{n}A_2\bar{R}$
 s , \bar{s} と $\hat{\sigma}$ との関係: $\hat{\sigma} = \bar{s}/c_2$; $\hat{\sigma} = s/c_2$
 R , \bar{R} と $\hat{\sigma}$ との関係: $\hat{\sigma} = \bar{R}/d_2$; $\hat{\sigma} = R/d_2$
 σ と σ_R との関係: $\sigma_R = d_3\sigma$
 係数の関係: $A = 3/\sqrt{n}$; $A_2 = 3/d_2\sqrt{n}$; $D_1 = d_2 - 3d_3$,
 $D_2 = d_2 + 3d_3$, $D_3 = 1 - 3d_3/d_2$, $D_4 = 1 + 3d_3/d_2$

附表 3. t 分布表



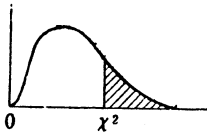
$Pr.$ は t の絶対値が表の値以上になる確率(両側)

\mathcal{F} は自由度

例. $\mathcal{F} = 10$ のとき t が 2.23 以上になる確率は 0.05

$\mathcal{F} \backslash Pr.$	0.2	0.1	0.05	0.02	0.01	0.001
1	3.08	6.31	12.71	31.82	63.66	636.62
2	1.89	2.92	4.30	6.97	9.93	31.60
3	1.64	2.35	3.18	4.54	5.84	12.94
4	1.53	2.13	2.78	3.75	4.60	8.61
5	1.48	2.02	2.57	3.37	4.03	6.86
6	1.44	1.94	2.45	3.14	3.71	5.96
7	1.42	1.89	2.37	3.00	3.50	5.41
8	1.40	1.86	2.31	2.90	3.36	5.04
9	1.38	1.83	2.26	2.82	3.25	4.78
10	1.37	1.81	2.23	2.76	3.17	4.59
11	1.36	1.80	2.20	2.72	3.11	4.44
12	1.36	1.78	2.18	2.68	3.06	4.32
13	1.35	1.77	2.16	2.65	3.01	4.22
14	1.35	1.76	2.15	2.62	2.98	4.14
15	1.34	1.75	2.13	2.60	2.95	4.07
16	1.34	1.75	2.12	2.58	2.92	4.02
17	1.33	1.74	2.11	2.57	2.90	3.97
18	1.33	1.73	2.10	2.55	2.88	3.92
19	1.33	1.73	2.09	2.54	2.86	3.88
20	1.33	1.73	2.09	2.53	2.85	3.85
25	1.32	1.71	2.06	2.49	2.79	3.73
30	1.31	1.70	2.04	2.46	2.75	3.65
40	1.30	1.68	2.02	2.42	2.70	3.55
60	1.30	1.67	2.00	2.39	2.66	3.46
120	1.29	1.66	1.98	2.36	2.62	3.37
∞	1.28	1.65	1.96	2.33	2.58	3.29

附表 4. カイ二乗分布表

Pr. は χ^2 が表の値以上になる確率 (片側)

\$ は自由度

Pr. \$	0.99	0.95	0.90	0.20	0.10	0.05	0.01	0.001
1	0.00	0.00	0.02	1.64	2.71	3.84	6.64	10.83
2	0.02	0.10	0.21	3.22	4.61	5.99	9.21	13.82
3	0.11	0.35	0.58	4.64	6.25	7.82	11.35	16.27
4	0.30	0.71	1.06	5.99	7.78	9.49	13.28	18.47
5	0.55	1.14	1.61	7.29	9.24	11.07	15.09	20.52
6	0.87	1.63	2.20	8.56	10.65	12.59	16.81	22.45
7	1.24	2.17	2.83	9.80	12.02	14.07	18.43	24.32
8	1.65	2.73	3.49	11.03	13.35	15.51	20.09	26.13
9	2.09	3.32	4.17	12.24	14.68	16.92	21.67	27.88
10	2.56	3.94	4.87	13.44	15.99	18.31	23.21	29.59
11	3.05	4.57	5.58	14.63	17.28	19.68	24.73	31.26
12	3.57	5.23	6.30	15.81	18.55	21.03	26.22	32.91
13	4.11	5.89	7.04	16.99	19.81	22.36	27.69	34.53
14	4.66	6.57	7.79	18.15	21.05	23.69	29.14	36.12
15	5.23	7.26	8.55	19.31	22.31	25.00	30.58	37.70
16	5.81	7.96	9.31	20.47	23.54	26.30	32.00	39.25
17	6.41	8.67	10.08	21.62	24.77	27.59	33.41	40.79
18	7.01	9.39	10.86	22.76	25.99	28.87	34.81	42.31
19	7.63	10.12	11.65	23.90	27.20	30.14	36.19	43.82
20	8.26	10.85	12.44	25.04	28.41	31.41	37.57	45.32
21	8.90	11.59	13.24	26.17	29.62	32.67	38.93	46.80
22	9.54	12.34	14.04	27.30	30.81	33.92	40.29	48.27
23	10.20	13.09	14.85	28.43	32.01	35.17	41.64	49.73
24	10.86	13.85	15.66	29.55	33.20	36.42	42.99	51.18
25	11.52	14.61	16.47	30.68	34.38	37.65	44.31	52.62
26	12.20	15.38	17.29	31.80	35.56	38.89	45.64	54.05
27	12.88	16.15	18.11	32.91	36.74	40.11	46.95	55.48
28	13.56	16.93	18.94	34.03	37.92	41.34	48.28	56.89
29	14.26	17.71	19.77	35.14	39.09	42.56	49.59	58.30
30	14.95	18.49	20.60	36.25	40.26	43.77	50.89	59.70

備考 F表は略した。

附表 5. 乱 数 表 (1)

81 28 37 85 17	69 35 24 77 61	15 69 25 80 89	13 83 01 12 51
41 66 57 44 43	65 41 05 37 37	05 39 63 70 58	22 29 39 27 75
32 02 74 03 07	25 19 48 03 46	35 43 18 03 51	42 38 87 71 72
38 72 04 51 62	37 09 79 40 77	53 43 63 36 07	14 31 11 06 41
04 59 74 55 09	03 77 79 22 25	71 29 13 04 29	07 07 47 68 57
48 30 73 23 77	96 41 09 82 35	45 21 89 83 21	56 31 48 92 62
07 02 18 42 73	95 99 99 51 95	90 37 69 49 34	29 19 29 52 48
89 66 17 41 68	60 59 25 47 02	84 08 96 85 33	28 17 07 82 86
61 96 95 01 66	65 51 77 21 40	17 23 10 36 37	50 77 30 61 88
56 64 99 30 82	87 11 57 44 53	74 08 05 20 03	49 10 92 01 63
21 25 19 39 83	53 47 62 57 38	93 48 76 73 69	14 69 17 66 20
47 79 02 70 43	38 82 07 92 85	07 17 01 61 87	01 05 35 29 73
39 81 85 08 35	47 36 65 10 12	97 20 96 60 87	86 18 64 59 61
23 93 45 51 06	59 53 68 98 09	33 71 72 50 49	23 88 39 64 64
52 35 31 50 76	28 27 89 32 88	70 65 76 13 06	51 99 03 33 22
50 66 31 38 84	38 18 23 04 18	80 65 96 60 58	72 86 39 89 19
51 81 42 52 76	49 20 45 29 36	15 29 61 45 43	38 95 08 29 22
37 12 65 58 54	33 37 08 22 61	31 50 57 07 41	99 78 42 74 34
07 85 32 90 20	62 65 51 16 49	23 54 23 34 09	66 03 81 85 45
84 16 33 34 55	99 71 95 88 57	79 57 13 36 20	48 38 46 91 62
45 10 31 67 95	54 73 16 86 58	06 88 64 37 68	53 82 04 33 21
83 41 19 11 63	92 14 77 37 15	52 60 08 01 78	52 05 49 43 46
88 63 80 59 48	03 89 89 87 23	28 68 34 21 31	25 30 49 85 49
27 37 68 45 51	52 96 98 47 73	32 47 60 91 09	71 58 09 90 23
30 18 16 55 43	84 95 27 65 93	84 62 87 30 88	09 55 00 55 14
70 74 65 24 77	57 21 24 63 03	06 37 34 60 38	17 66 74 82 30
22 73 72 42 62	43 43 14 81 73	62 98 53 95 69	94 53 73 20 84
10 27 16 27 91	82 97 41 22 16	78 46 55 08 65	77 16 78 89 10
76 03 75 59 41	22 68 55 72 19	22 97 10 02 07	85 89 47 59 31
24 99 64 32 45	01 02 24 53 22	99 76 59 95 83	97 56 56 45 74
04 80 27 26 84	05 64 07 61 13	51 15 86 75 08	88 16 84 41 80
47 44 40 54 95	05 83 95 23 33	44 81 13 61 42	08 70 72 43 72
77 67 98 97 09	82 44 54 36 27	86 21 43 95 53	48 70 24 71 71
25 98 64 20 63	19 30 76 13 67	47 41 66 94 72	18 07 58 28 29
90 62 54 36 03	03 24 59 79 22	49 27 26 39 86	23 37 73 38 05
23 09 86 07 93	18 95 48 70 47	66 41 63 45 51	80 93 62 27 14
80 73 72 65 82	85 79 07 38 10	11 97 16 47 00	72 64 47 13 05
27 09 21 95 11	57 61 47 11 98	12 43 20 52 24	24 76 71 60 99
60 51 05 61 53	01 44 08 70 87	09 76 99 93 77	61 54 51 62 66
53 14 90 84 44	11 37 00 52 02	63 83 82 87 48	26 71 77 97 79
17 77 88 00 31	30 10 62 96 15	90 94 80 88 69	27 35 16 13 67
60 77 05 63 44	41 38 69 45 21	65 88 61 69 05	21 84 22 39 80
50 27 94 60 11	72 53 90 03 85	51 81 21 33 76	59 13 99 11 04
02 70 89 14 60	26 86 45 70 61	70 43 08 70 78	82 95 47 99 81
35 73 19 96 82	63 46 45 40 04	78 90 17 45 66	94 93 97 74 24
97 28 64 86 90	88 53 17 97 69	51 40 34 81 91	30 88 12 84 32
07 32 49 55 27	37 06 52 33 96	97 72 06 40 97	58 78 67 66 03
42 22 82 06 10	21 16 56 80 95	79 06 29 69 51	19 34 34 25 47
98 96 10 47 68	68 70 65 29 21	69 94 39 68 00	11 10 85 48 83
89 58 00 03 07	80 90 47 53 93	60 47 19 62 92	61 15 70 94 91

亂 數 表 (2)

19 28 60 40 00	55 16 76 86 38	61 55 43 06 76	19 03 28 41 59
87 73 76 11 94	42 52 02 43 24	08 83 96 51 33	15 69 46 65 24
63 49 16 97 22	12 37 28 48 31	07 15 20 96 43	28 33 77 0 17
56 23 40 15 42	15 95 43 15 40	87 29 86 73 72	67 19 39 66 84
61 36 00 68 06	77 67 73 70 45	97 42 92 05 85	37 95 75 08 27
35 61 76 38 36	06 91 65 14 95	13 54 31 44 33	60 38 06 41 42
70 78 22 99 57	08 05 88 71 81	68 34 06 33 58	73 03 14 82 41
78 66 59 94 71	55 56 53 24 14	29 15 93 32 46	24 89 14 49 95
42 40 39 97 27	86 44 34 07 09	95 45 74 09 45	66 37 54 13 54
11 85 40 07 46	68 29 94 53 25	83 45 69 62 25	44 63 35 66 33
30 15 59 35 54	91 96 05 80 58	45 14 89 49 27	31 23 37 50 30
62 42 06 86 88	41 99 65 74 16	43 32 18 37 70	76 38 30 80 70
80 02 44 93 67	91 17 75 43 72	62 95 83 55 92	17 66 92 78 26
47 88 32 10 07	02 99 37 56 99	88 94 08 29 01	78 83 06 42 54
26 71 13 55 96	74 62 44 34 17	61 73 90 74 52	84 91 45 30 64
67 33 44 08 58	30 07 16 95 47	57 42 85 27 83	85 69 59 30 26
92 97 07 99 44	29 94 82 38 87	29 98 80 27 98	08 45 66 25 71
95 51 13 44 77	10 80 59 24 89	10 14 62 42 33	45 27 19 29 02
50 19 24 83 11	54 04 64 68 94	86 34 78 83 54	95 46 06 44 59
23 64 90 84 03	26 89 71 41 61	87 48 13 33 73	92 89 73 78 65
28 69 74 23 33	49 11 44 09 12	22 22 17 39 56	85 64 37 11 66
55 32 08 88 87	49 51 09 24 48	79 95 33 92 51	54 93 69 85 87
96 90 72 88 45	00 97 55 70 96	98 69 77 97 95	07 61 79 68 96
84 97 02 24 34	47 26 43 24 01	03 37 43 16 17	55 22 45 33 79
90 30 94 60 87	74 83 94 67 75	05 35 28 17 08	63 22 00 32 68
27 89 33 90 77	88 26 28 53 33	99 36 22 47 95	25 98 45 27 10
42 14 27 16 68	38 92 24 93 12	85 64 76 64 02	43 03 43 92 21
59 95 13 35 65	08 39 00 48 36	21 43 68 95 41	33 12 84 43 94
63 50 57 53 20	24 66 46 97 23	82 89 39 24 30	33 36 95 86 89
63 20 06 45 10	62 02 01 87 30	55 31 71 54 35	74 73 33 95 04
15 50 87 82 11	62 23 79 39 91	17 76 35 54 94	59 06 65 89 82
49 26 21 18 74	33 93 84 62 02	21 55 50 96 31	31 39 95 04 47
93 77 12 97 26	75 85 09 42 25	13 88 96 50 89	40 18 66 16 86
20 07 15 25 01	67 60 75 24 48	24 53 93 81 48	84 95 88 01 03
68 16 90 02 49	73 85 97 73 29	84 64 87 96 37	33 85 91 95 21
81 27 06 15 95	90 74 85 29 59	27 70 56 00 26	01 98 80 22 34
76 56 20 94 05	15 77 86 87 28	27 18 72 95 07	31 20 64 48 78
79 55 32 26 75	53 77 80 55 45	15 11 96 82 55	29 73 88 67 97
58 06 80 58 14	32 27 49 16 89	18 78 94 81 12	30 64 42 97 12
84 43 54 52 93	62 68 87 16 47	95 40 64 80 18	93 44 26 42 64
28 82 17 02 15	36 07 68 69 59	14 02 44 58 35	69 51 82 73 68
13 68 70 47 99	49 54 63 57 16	16 92 76 99 76	17 82 69 17 18
36 26 04 49 47	83 60 25 45 73	65 28 17 04 85	60 15 32 07 44
18 14 12 07 80	28 04 70 45 72	12 62 17 52 51	83 70 24 71 75
10 47 52 57 22	82 48 83 58 09	72 33 76 41 52	95 53 14 87 56
77 03 04 27 50	09 02 81 68 91	44 96 29 59 76	66 92 85 32 42
98 75 13 48 33	76 45 89 79 12	15 60 80 37 14	71 43 14 07 06
39 97 96 02 69	88 75 72 66 57	14 46 40 87 98	91 53 28 26 18
54 53 10 38 68	39 51 29 80 66	61 35 56 98 22	86 56 94 30 26
30 33 34 02 41	26 55 77 99 71	99 76 65 35 42	81 95 07 07 42

亂 數 表 (3)

33 49 89 82 75	03 23 47 55 54	43 94 08 69 70	26 68 16 46 60
04 02 13 09 27	75 93 74 83 62	40 28 60 54 16	82 92 11 75 34
50 35 71 38 86	30 87 56 98 42	92 40 09 00 71	62 97 93 74 66
60 77 19 48 51	07 90 21 39 93	49 38 85 76 55	52 57 84 95 03
93 11 56 26 18	10 32 10 86 58	58 07 95 75 54	29 44 82 12 13
52 56 39 03 13	29 63 13 04 33	39 86 79 32 82	33 70 81 13 87
55 54 73 28 08	84 70 33 18 02	76 12 09 44 46	40 56 03 35 53
44 67 81 40 77	66 93 70 68 42	19 63 75 50 96	91 80 45 77 07
46 49 05 96 75	71 52 06 57 82	95 17 73 90 69	25 37 01 72 88
30 91 62 83 51	38 23 03 54 83	18 77 06 44 29	67 35 12 43 03
89 13 31 83 95	16 63 30 00 85	43 56 42 50 23	51 43 40 70 77
62 68 97 41 56	03 34 67 45 95	64 85 64 24 24	64 78 34 76 13
17 54 32 30 63	25 39 06 49 35	64 62 69 77 12	12 98 27 53 58
19 86 81 08 78	97 83 77 80 12	24 24 83 57 19	79 84 20 61 50
93 97 45 79 47	88 12 02 41 12	00 43 11 05 02	08 23 22 24 96
17 33 11 26 36	03 72 25 61 98	67 23 52 15 43	36 17 70 59 32
42 47 14 39 59	57 31 47 47 55	54 80 58 24 63	37 33 18 06 14
35 34 40 44 24	41 69 53 63 51	52 13 69 33 14	17 82 61 20 58
26 58 44 34 85	23 84 58 30 88	57 04 58 11 13	55 91 90 38 07
98 78 41 58 17	75 87 40 45 52	37 57 22 11 11	35 83 22 47 80
56 58 07 23 55	72 86 29 33 00	22 38 69 46 57	93 22 95 14 69
31 33 64 69 17	84 99 85 67 52	09 50 50 87 07	73 49 00 11 70
78 69 80 55 25	63 52 52 44 61	42 52 47 66 63	07 39 05 00 05
98 73 27 24 16	56 10 61 63 66	26 59 50 90 93	34 52 56 90 46
95 47 54 28 37	59 83 63 25 62	03 77 58 46 57	33 47 54 79 04
03 22 02 69 26	05 04 39 12 78	41 52 17 38 34	33 64 48 60 29
72 90 83 40 93	30 32 65 26 13	24 61 64 50 87	34 19 76 23 48
70 95 82 26 89	59 74 06 25 39	75 86 86 00 18	72 13 14 85 89
85 67 38 32 41	96 84 00 76 33	20 10 28 62 18	81 59 96 53 36
32 88 01 97 63	59 47 72 98 73	02 93 92 06 25	64 48 25 75 36
08 66 21 03 72	65 04 55 47 33	78 58 47 40 97	00 88 55 71 01
97 96 32 25 88	05 99 76 74 95	45 78 46 04 93	83 43 19 19 85
80 42 04 33 36	66 83 05 25 53	44 51 72 79 77	63 05 14 99 63
26 46 38 96 18	03 81 13 21 13	41 84 59 02 52	22 20 08 69 16
03 07 48 59 18	58 74 27 47 61	20 40 00 24 03	48 99 61 90 06
98 48 28 02 55	79 99 83 69 88	82 25 87 65 06	12 83 77 37 89
64 61 30 69 41	67 77 91 09 14	59 16 39 03 03	16 42 35 43 49
98 83 95 70 58	56 08 82 80 30	14 10 21 82 20	00 78 13 43 43
76 25 01 26 11	29 07 95 06 31	62 94 04 97 75	48 31 18 14 22
97 13 10 26 73	43 23 72 02 74	01 82 30 28 67	85 20 73 00 37
53 21 05 47 91	92 14 68 81 37	55 98 33 43 70	80 39 80 86 48
14 92 04 64 57	33 01 03 03 77	94 68 50 76 44	09 48 35 62 91
32 50 05 22 94	96 50 07 62 21	03 37 33 33 96	59 33 05 66 56
24 68 84 55 79	31 49 63 03 28	12 25 78 66 06	17 48 54 26 30
30 32 28 04 43	13 84 47 10 53	89 30 46 37 78	19 83 49 64 29
04 61 89 59 07	74 92 51 75 46	49 93 14 69 21	84 02 40 39 57
42 71 73 94 75	97 11 00 15 31	04 15 22 68 17	30 60 02 31 04
83 34 10 86 77	45 61 75 12 28	67 57 27 55 67	11 88 87 59 42
24 92 91 31 95	02 06 38 12 26	04 18 11 23 22	38 41 61 29 24
82 75 03 07 39	52 49 34 18 03	29 17 84 25 64	47 10 90 44 09

参考数値表

合理的なサンプリング研究並びに実施に数値表が必要である。本書では2〜3の数表をつけたにすぎぬが、下記の数値表を用いると便利である。

品質管理用数値表：日科技連リサーチグループ編，1952，日本科学技術連盟

品質管理のためにも，サンプリング研究のためにも便利なもので安価ではあるし，技術者必携の書。(100円)

新編統計数値表：統計科学研究会（増山，北川）編，1952，河出書房
非常に多くの数値表とその解説があり，本格的に勉強される方はもたれるとよい。(2,300円)

Statistical Tables：Fisher and Yates 1938, 1949 Oliver and Boyd
統計数値表と同様に多くの表と解説があり，よい本である。(15\$)

索引

ア 行

一元配置法	310
1次サンプリング単位 (1次単位)	131
1段サンプリング法	132
インクリメント	15, 234, 240
————間のバラッキ	258
————の大きさ	15, 106, 161, 258
————の取りかた	266
英国規格	263
液体のサンプリング	218
受入検査	17
——仕様	18, 20
M^2 分布	63
円錐四分法	273
大口試料(1次試料)の集め方	256
帯状体	299

カ 行

カイニ乗(χ^2)分布	58, 63, 67, 336
回転縮分器	276
ガウスの誤差曲線	51
攪拌サンプリング	224
確 率	52
——誤差	90

確率比例サンプリング	130, 137
カスケードサンプラー	274
カタヨリ	13, 84, 93, 97
下部サンプル	222
空試験	97
假りの平均	45
管理試験	73
——図	86, 91, 94, 159, 163, 167, 170, 291
——係数表	334
規 格	5, 17
規準化	49
期待値	54, 310
気体のサンプリング	215
級	41
級間分散	112, 114, 295
級内分散	112, 114, 295
級の巾	41
供給源	11
均一化	155
区間推定	60, 64
グラフサンプリング	230
グラフサンプル	224
繰返精度	92, 326
群	11, 293
群 分	293
計数值	51

系統的サンプリング 34, 37, 119,
122, 126, 164, 168,
213, 232, 294

契約書 3, 5

計量値 51

ケーキ状, 泥状体の
サンプリング 233

誤差 78, 79

——分散 54

工程の管理状態 153

交絡 93

合理的なサンプリング 7, 151

混合試料 (混合サンプル) 223

サ行

再現精度 92, 325

最大範囲 80, 90

最多値 46

最適配分 109, 133

三元配置法 321

算術平均 44

サンプリング 9

——カード 234

——間隔 (抜取間隔)

34, 292

——誤差 71, 104

——の手順 7

——の原則 25

——分類 21

——パイプ 217, 228

サンプリング法標準 (書) 8, 151,
305

サンプル 9

——の大きさ 14, 45, 69

——の単位 15

——のもつべき条件 12

仕切=ロット 10

自動試料採取機 141

シーフサンプリング 229

試長 15, 161, 236

出荷検査 19

——仕様 20

濕分測定用サンプルの縮分 282

縮分 234, 271, 323

集合体 21, 22

——のサンプリング 211

修正項 47

自由度 58, 62

集落 125, 129

——サンプリング 125, 165, 166

仕様書 5, 17

上部サンプリング 222

試料=サンプル 9

——抜取=サンプリング 9

——標準偏差 49

——分割器 273

——分散 47

——平均 45

眞の値 93

信頼度 64

	索 引
信頼係数	80
——限界	65
——性	13,78
——度	13,80
推 定	59
——値	60,109
ス ョ ー	50
スナップサンプリング	140,216
正確さ	78,83
——度	84
正規分布	51,57,334
精 度	78,79
石炭の縮分法	278
設計仕様	18,20
全数検査, 全数調査	4,10
全層サンプル	221
相関分析法	98
測定誤差	71,85
——の信頼性	86
——正確さ	93
——精度	90
——法標準書	306
測 度	44
総 計	61
層積法 (bedding)	156,248
相対度数	42
——分布	42
層 別	155,293
——サンプリング	113,122, 145,162,166,192,295

	索 引
層別比例サンプリング	116
夕 行	
対応のある変量模型	309
——ない——	309
大母集団	11
多元配置法	321
多段サンプリング	131
トップサンプリング	227
単位体	14,21
——のサンプリング	187
単純ランダムサンプリング	27,295
地域サンプリング法	131
チェックサンプリング	79
チグザグサンプリング	39
チップ実験	14
中央値	46
中心的傾向	44
柱状図	42
チューブサンプリング	228
——サンプル	224
中部サンプル	222
底部——	224
t 分布	63
出口サンプル	223
データの棄却	87
——歴史	5
デミング (Deming) サンプリ ン グ	124
点推定	60

とがり	44
特別な変量模型	323
独立	75
統計量	40,44
——の分布	53
度数多角形	42
——分布	42
——曲線	43
——表	41
トップサンプル	223
ドレインサンプル	224

ナ 行

二元配置法	313,315
2次サンプリング単位	
(2次単位)	131
2段サンプリング	131
抜取比	34
ネイマン (Neymann) サンプル	
リング	123

ハ 行

バイアス	84
破壊試験	4,72,91,103
バッチ	10
バラツキ	44
——の測定	46
範囲	49,90
——の分布	59
半固溶体のサンプリング	231

販売仕様	19,20
反覆サンプリング	141
比推定	142,208
柄杓サンプリング	228
——サンプル	223
微粉	244
費用函数	134,135
標準誤差	54
——試験	93
——試薬	93
——試料	93
——偏差	47,90
——の分布	58
標本=サンプル	9
——調査法	6,109
平らたさ	44
比例サンプリング	116
変量模型	303,307
副次サンプリング	122,131,147,
	165,194,205,240,295
不偏推定値	60
不偏分散	62
——の期待値	310
——の比	62
粉塊混合物	22,121,156,
	164,169,181,243
分散	46,90
——の加成性	72,75
——の簡易推定法	160
——の推定	156

分散の分布	57
——分析	96, 98, 176, 179
——の構造	307
分析試料	245
——用試料	245
平均サンプル	139, 321
——値	44
——の分布	54
——偏差	50, 90
平方和	46
米国規格	264
bedding	156, 248
偏 差	46
——二乗和	46
変動係数	49
保証試験	73
母集団	3, 9
——の大きさ	14, 45
——修正	175
母 数	40, 44
——模型	307
母標準偏差	47
母分散	47, 109
母平均	45
ボーリングサンプリング	229
——サンプル	224

マ 行

無限母集団	14, 55, 167, 291
無作為化	14

メジアン	46
モード	46

ヤ 行

有意差検定	93
——サンプリング	26, 138
ゆがみ	44
有限母集団	14, 55
——修正 (有限修正)	55
四元配置法	322
予備知識	8, 151
——調査	8, 108, 110, 151
熔融体	231

ラ 行

ラテン方格法	331
乱 数	30
——器	33
——サイ	33
——表	30, 337
ランダム	13
——サンプリング	14, 24, 142, 157, 187, 198
離散的分布	51
リップルサンプラー	274
流体のサンプリング	212
粒 度	249
——・重量比の定理	284
——分布	253
累積度数曲線	44

累積度数多角形	44
——分布	43
layering	156
レンジ	49
連続体	22
——のサンプリング	236
——サンプリング	228
——サンプル	223

連続分布	51
ロット	10
——の大きさ	14
——取り扱い方	155
——歴史	153

ワ 行

歪度	44
----	----

著者の略歴

東京大学助教授、
昭和14年東京大学工学部応用化学科卒

工場における ¥ 350
サンプリング (地方売価 370)

昭和27年7月20日印刷・昭和27年7月25日発行
昭和28年4月1日第2版印刷・昭和28年4月5日第2版発行
昭和28年9月10日第3版印刷
昭和28年9月15日第3版発行



著 作 者 石 川 馨 かたし
発 行 者 丸 善 株 式 会 社
代 表 者 司 忠
東 京 都 中 央 区 日 本 橋 通 2 丁 目 6 番 地
印 刷 者 林 清 市
東 京 都 中 央 区 京 橋 3 丁 目 2 番 地

発行所

東京都中央区日本橋通2丁目6番地
丸 善 株 式 会 社

印刷・中央印刷株式会社
製本・株式会社三水舎