⑬ 標準偏差の推定値について

東北大学

淡中　忠郎

日本応用力学会編「応用統計学」8．26に石田保士氏が正規分布をする母集団の標準偏差について一つの推定法を紹介している。以下にその方法の精度について考えて見度い。

石田氏の他の方法（$n \leqq 10$ の場合）に関しては北川氏が研究していられる由である。

今 x_1，x_2，-----，x_n を大きさ n の試料とし，母集団のS.D. σ を

$$\sigma_e = \frac{1}{nc}\left(|x_1 - x_2| + |x_2 - x_3| + \cdots + |x_{n-1} - x_n| + |x_n - x_1|\right)$$

によって推定する（$c = \frac{2}{\sqrt{\pi}} = 1.128 \cdots$）．問題はこの推定値の分散で，最初はその正確な値を求めることが不可能でないかと思われたので数値計算で概数を求め，後になって実際の値を出すことが出来た。勿論出来て了えば簡単なことであるが，上の推定値を実際問題に応用される際多少の参考になることゝ思い，こゝに発表することゝした次第である。便宜上母集団平均値を0として

$$\phi(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x^2/2\sigma^2}$$

とおく，されば

$$E|x-a| = \int_{-\infty}^{\infty} |x-a| \phi(x) dx$$

$$= \int_{-\infty}^{a} (a-x)\phi(x) dx + \int_{a}^{\infty} (x-a)\phi(x) dx$$

$$= a\left(\int_{-\infty}^{a} \phi(x) dx - \int_{a}^{\infty} \phi(x) dx\right) + \left(-\int_{-\infty}^{a} x\phi(x) dx + \int_{a}^{\infty} x\phi(x) dx\right)$$

この右辺の第一項を少し変形し，第二項は部分積分法で容易にその値が求められることに注意して

$$E|x-a| = 2|a|\int_{0}^{|a|} \phi(t) dt + 2\sigma^2\phi(a)$$

が得られる．　之を用いれば

$$d = E(|x_1 - x_2||x_2 - x_3|)$$

$$= \iiint |x_1 - x_2||x_2 - x_3| \phi(x_1)\phi(x_2)\phi(x_3) dx_1 dx_2 dx_3$$

$$= \int_{-\infty}^{\infty} \left(2x\int_{0}^{x} \phi(t) dt + 2\sigma^2\phi(x)\right)^2 \phi(x) dx$$

（x_1，x_3 で積分をする際に前の公式を用い，残つた x_2 を x と書けばよい。）次に

$$d' = \int_{0}^{\infty} \left(x\int_{0}^{x} \phi(t) dt + \phi(x)\right)^2 \phi(x) dx$$

$$= I_1 + I_2 + I_3 \qquad (d = 8d')$$

と置く。　右辺の三項は被積分項の平方を展開して得られる三つの積分に対応するものである。之を逐次に求める。

$$I_3 = \int_0^\infty \phi(x)^3 dx = \left(\frac{1}{\sqrt{2\pi}}\right)^2 \frac{1}{\sqrt{3}} \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-3x^2/2} d(\sqrt{3}x)$$

$$= \frac{1}{4\sqrt{3}\pi} \quad \left(\int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{2} \text{ に注意} \right)$$

又　$\phi'(x) = -x\phi(x)$;　$\frac{d}{dx}\left(\phi^2(x)\right) = -2x\phi'(x)$　を用いて I_2 を部分積分法により計算すれば

$$I_2 = \int_0^\infty \left(2x\phi^2(x) \int_0^x \phi(t)\,dt \right) dx$$

$$= \left[\int_0^x \phi(t)\,dt \left(-\phi^2(x)\right) \right]_0^\infty + \int_0^\infty \phi(x)^3 dx$$

$$= 0 + I_3 = \frac{1}{4\sqrt{3}\pi}$$

次に　$\phi'(x) = -x\phi(x)$,　$\phi''(x) = -\phi(x) + x^2\phi(x)$,

$x^2\phi(x) = \phi(x) + \phi''(x)$　であるから

$$I_1 = \int_0^\infty \left(\int_0^x \phi(t)\,dt \right)^2 \phi(x)\,dx + \int_0^\infty \left(\int_0^x \phi(t)\,dt \right)^2 \phi''(x)\,dx$$

$$= I_4 + I_5$$

ここに　I_4 は部分積分法により

$$I_4 = \left[\left(\int_0^x \phi(t)dt \right)^2 \int_0^x \phi(t)dt \right]_0^\infty$$

$$- \int_0^\infty 2 \left(\int_0^x \phi(t)dt \ \phi(x) \int_0^x \phi(t)dt \right) dx$$

$$= \left(\frac{1}{2} \right)^3 - 2 I_4 \qquad \therefore I_4 = \frac{1}{24}$$

最後に

$$I_5 = \int_0^\infty \left(\int_0^x \phi(t)dt \right)^2 \phi''(x)\, dx$$

$$= \left[\left(\int_0^x \phi(t)dt \right)^2 \phi'(x) \right]_0^\infty - \int_0^\infty 2 \int_0^x \phi(t)dt\, \phi(x) \phi'(x)\, dx$$

$$= 0 + 2 \int_0^\infty \left(x \phi^2(x) \int_0^x \phi(t)dt \right) dx = I_2 = \frac{1}{4\sqrt{3}\pi}$$

以上をまとめると

$$d = 8d' = \frac{6}{\sqrt{3}\pi} + \frac{1}{3}$$

次に目的の σ_e の分散は

$$Var\, \sigma_e = E(\sigma_e - \sigma)^2 = E\sigma_e^2 - \sigma^2$$

$$= \frac{1}{n^2 c^2} - \left(|x_1 - x_2| + \cdots + |x_{n-1} - x_n| + |x_n - x_1| \right)^2 - \sigma^2$$

であるが，平方を展開した n^2 個の項の中で

1) $E|x_1 - x_2|^2$ の形の項が n 項，

2）　$E|x_1-x_2||x_2-x_3|$　の形の項が $2n$ 項,

3）　$E|x_1-x_2||x_3-x_4|$　の形の項は残りの n^2-n-2n

$= n(n-3)$ 項

あることは容易に分る。　ところが

$E|x_1-x_2|^2 = x_1-x_2$　の分散 $=2\sigma^2$.

$E|x_1-x_2||x_2-x_3| = d$ は既にもとめた通り．

$E|x_1-x_2||x_3-x_4| = E|x_1-x_2|E|x_3-x_4| = (c\sigma)^2 = \frac{4}{\pi}\sigma^2$

であるから之等を代入して

$$\mathrm{Var}\,\sigma_e = \frac{1}{n^2c^2}\left\{2n+n(n-3)\frac{4}{\pi}+2nd\right\}\sigma^2-\sigma^2$$

$$= \frac{1}{n}\left(\sqrt{3}+\frac{2}{3}\pi-3\right)\sigma,$$

$$\underline{\sigma_e \text{ の S.D.} = 0.9091\sigma/\sqrt{n}}$$

となつて目的が達せられた．

（前にも述べた様に最初は d の計算が不可能と思われたので数値積分によることにした。　そのためには被積分項を評価して $\int_a^\infty$ の部分が省略出来る程度に a をもとめ $\int_0^a$ を計算すればよいと思つたのであるが，山下千歳氏（東北大教養部）の次の考えにより著しく計算を簡単にすることが出来た。　現在の問題では最早その必要はないのであるが，他の問題に応用出来る可能性を考えて附記して置く．

即ち d の被積分項の平方内は x が大きい時略々 x に等しいから

$$2x\int_0^x \phi(t)dt + 2\phi(x) = x(1+2\delta(x))$$

と置いて見る（ $\sigma=1$ と仮定した），さすれば

$$\delta(x) = \frac{\phi(x)}{x} - \int_x^\infty \phi(t)\,dt.$$

之を微分して　$\phi'(x) = -x\,\phi(x)$　なることから

$$\delta'(x) = \frac{\phi'(x)}{x} - \frac{\phi(x)}{x^2} + \phi(x) = -\frac{\phi(x)}{x^2} < 0$$

従つて

$$d = 2\int_0^\infty = 2\int_0^a + 2\int_a^\infty$$

で，

$$2\int_a^\infty x^2\phi(x)dx < 2\int_a^\infty < 2(1+2\delta(a))\int_a^\infty x^2\phi(x)dx$$

$$\left(\int_a^\infty x^2\phi(x)dx = a\,\phi(a) + \int_a^\infty \phi(x)dx\right)$$

となり，a を適当に大きくとれば $2\int_a^\infty$ が数表によつてもとめることが出来る．　$2\int_0^a$ は例えば Simpson の方法でもとめられば充分精密にもとめることが出来る）

終りに，石田氏の推定値 σ_e と，標本標準偏差 s の精度を比較して見る．

n が充分大きい時 σ_e はその形から見て大体正規分布をなすものと考えてよいであらう．

従って $\sigma_e \leqq C_n \sigma$ （危険率 0.05 ）の様な C_n がもとめられる。

S の場合にはその求め方はよく知られているから両者を比較して見れば大体の見当がつくであろう。

次にその概略の値を掲げる。

n	σ_e の場合	S の場合
100	1.1495	1.11
200	1.1058	1.08
300	1.0863	1.066
400	1.0748	1.058
500	1.0669	1.051

以　上