⑥ 林の Normality Test に就て

阪大理学部数学教室　　小川　潤次郎

Sample Size N が充分大きい場合に用い得る Normality Test として林知己夫氏わ次のような Test を提唱している。(1)
即ち大さ N の標本 $X_1, X_2, \cdots\cdots, X_N$ の順序が Randomize されているとき，これを2群に分って

第一群 ： $X_1, X_2, \cdots\cdots\cdots, X_{2n+1}$

第二群 ： $X_{2n+2}, \cdots\cdots\cdots, X_N$

として，第二群は改めて

$$Y_1, \quad Y_2, \cdots\cdots\cdots \quad Y_{f+1}$$

とかく。従って

$$(2n+1)+(f+1)=N$$

f わ小さく，n わ充分大と考える。
第一群の Median を $\widetilde{X}$ とし，第二群から統計量

$$S^2=\frac{1}{f}\sum_{i=1}^{f+1}(Y_i-\overline{Y}),\quad \overline{Y}=\frac{1}{\delta+1}\sum_{i=1}^{f+1}Y_i$$

を求める。 n が充分大ならば，母集団が $N(x;0;\sigma^2)$ のとき $\widetilde{X}$ の分布は漸近的に $N\left(\widetilde{x};0,\sigma^2\frac{\pi}{4n+2}\right)$ で，S^2 の分布は exact に自由度 f の χ^2 分布

$$p(s)ds = \frac{f^{\frac{f}{2}}}{2^{\frac{f-2}{2}}\Gamma\left(\frac{f}{2}\right)\sigma^{f}} s^{f-1} e^{-\frac{fs^2}{2\sigma^2}} ds$$

で，且つ $\tilde{X}$ と S とわ互に独立であることを用いて

$$T = \sqrt{\frac{4n+2}{\pi}} \cdot \frac{\tilde{X}}{S}$$

とおくと，T の漸近分布わ

$$p(t)dt = \frac{1}{\sqrt{\pi f}} \frac{\Gamma\left(\frac{f+1}{2}\right)}{\Gamma\left(\frac{f}{2}\right)} \left(1 + \frac{t^2}{f}\right)^{-\frac{f+1}{2}}$$

即ち，自由度 f の Student の t 分布となることを利用して，例えば $f = 10$，$n = 100$ ならば

$$t_{0.05} = 2.228$$

だから

$$\left|\frac{\tilde{x}}{s}\right| \geqq 2.228\sqrt{\frac{\pi}{4n+2}} \doteqdot 0.0174$$

のとき Normality の仮説を棄却しようと云ふのが林氏の Test である。

しかし，このとき対立仮説としてわ

$$N(x;\mu,\sigma^2) \qquad \mu \neq 0$$

を取り得ることを考えれば，統計量 T の値が

$$|T| \geqq t_\alpha \qquad (t：わ \alpha\% 点)$$

となったからと云って，Normality の仮説を棄却しないで $\mu = 0$ という仮説を棄却すると考えることも出来る訳で，むしろ，この方が自然

であると私は考える． Normality Test である為めにわ，考える統計量 T の標本分布わ，パラメーター μ，σ には無関係で Functunal Form のみに depend することが必要である。

この意味でわ Geary の b_1, b_2 を用いる Test わ

$$m_\nu = \frac{1}{N}\sum_{\alpha=1}^{N}(X_i-\bar{X})^\nu, \qquad \bar{X}=\frac{1}{N}\sum_\alpha X_\alpha$$

$$b_1 = \frac{m_3}{m_2^{3/2}}, \qquad b_2 = \frac{m_4}{m_2^2}$$

の分布は，母集団を $N(x;\mu,\sigma^2)$ とするとき，μ,σ^2 に無関係であるから合理的と云えるであろう。

だも，T の分布わ exact にわパラメーター μ,σ^2 に depend しても，practically に independent と考えられるならばよい。 林の Test の場合に，母集団を $N(x;\mu,\sigma^2)$とすると，$\tilde{x}$ と S の同時分布の Prob. element わ

$$\frac{\sqrt{2n+1}\; f^{\frac{f}{2}}}{2^{\frac{f-2}{2}}\pi\Gamma(\frac{f}{2})\sigma^{f+1}} s^{f-1} e^{-\frac{2n+1}{\pi\sigma^2}(\tilde{x}-\mu)^2-\frac{fs^2}{2\sigma}} d\tilde{x}ds$$

$$= \frac{\sqrt{2n+1}\; f^{\frac{f}{2}}}{2^{\frac{f-2}{2}}\cdot\pi\cdot\Gamma(\frac{f}{2})\sigma^{f+1}} e^{-\frac{2n+1}{\pi\sigma^2}\mu^2}$$

$$\times \sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{4n+2}{\pi\sigma^2}\mu\right)^r \tilde{x}^r s^{f-1} e^{-\frac{s^2}{2\sigma^2}\left(\frac{4n+2}{\pi}\frac{\tilde{x}^2}{s^2}+f\right)} d\tilde{x}ds$$

$$\sqrt{\frac{4n+2}{\pi}}\;\frac{\tilde{x}}{s} = t$$

とおくと，

$$p(t,s)\,dt\,ds=\frac{\sqrt{2n+1}\,f^{\frac{f}{2}}}{2^{\frac{f-2}{2}}\pi\cdot\Gamma(\frac{f}{2})\sigma^{f+1}}\,e^{-\frac{2n+1}{\pi\sigma^2}\mu^2}\sum_{r}\frac{1}{r!}\left(\frac{4n+2}{\pi}\right)^{\frac{r}{2}-1}\frac{\mu^r}{\sigma^{2r}}t^r$$

$$\times\, s^{f+r}e^{-\frac{s^2}{2\sigma}(t^2+f)}\,ds\,dt.$$

$$\delta^2\equiv\frac{4n+2}{\pi\sigma^2}\mu^2$$

とおいて，s を integrate out すると

$$p(t)\,dt=\frac{\left(\frac{f}{2}\right)^{\frac{f}{2}}e^{-\frac{\delta^2}{2}}}{\sqrt{2\pi}\,\Gamma\left(\frac{f}{2}\right)}\sum_{r=0}^{\infty}\frac{\Gamma\left(\frac{f+r+1}{2}\right)}{r!}(\delta t)^r\left(\frac{2}{f+t^2}\right)^{\frac{1}{2}(f+r+1)}dt. \qquad (2)$$

これは所謂 Non-central な t 分布である。$\delta=0$ なら $r=0$ の Term 丈のこつて

$$p_0(t)\,dt=\frac{\left(\frac{f}{2}\right)^{\frac{f}{2}}\Gamma\left(\frac{f+1}{2}\right)}{\sqrt{2\pi}\;\Gamma\left(\frac{f}{2}\right)}\left(\frac{2}{f+t^2}\right)^{\frac{1}{2}(f+1)}dt$$

となつて自由度 f の t 分布になる。 これでわ Non-central な t 分布の性質からして practically にも independent という訳にわ行かない。

多分林氏わ計算の簡便という点を無視して第一群の Median をとつたのであろうが，次のような Test を作れば，Geary より計算わ簡便で，一応の條件わ満たされる。 即ち X_1 ------ X_N を2群に分って

第一群　　$X_1,\ X_2,\ \cdots\cdots,\ X_{n_1}$

第二群　　$X_{n_1+1}=Y_1,\ X_{n_1+2}=Y_2,\ \cdots,\ X_N=Y_{n_2}$

$$n_1+n_2=N$$

として，これから

$$S_1^2 = \frac{1}{n_1-1} \sum_{\alpha=1}^{n_1} (X_\alpha - \bar{X}), \qquad \bar{X} = \frac{1}{n_1} \Sigma X_\alpha$$

$$S_2^2 = \frac{1}{n_2-1} \sum_{\alpha=1}^{n_2} (Y_\alpha - \bar{Y}), \qquad \bar{Y} = \frac{1}{n_2} \Sigma Y_\alpha$$

なる統計量を作ると

$$F = \frac{S_1^2}{S_2^2}$$

な，母集団を N（x；μ, σ^2）とすれば，μ, σ^2 に無関係につねに自由度（n_1-1　n_2-1）の F 分布に従う。

Student の有名な Data[3] についてこの Test を用いて見る。

$$X_1 = 1.9,\ X_2 = 0.8,\ X_3 = 1.1,\ X_4 = 0.1,\ X_5 = -0.1$$

$$Y = 4.4,\ Y_2 = 5.5,\ Y_3 = 1.6,\ Y_4 = 4.6,\ Y_5 = 3.4$$

$$\tilde{X} = 0.76 \qquad \bar{Y} = 3.90$$

$$4 S_1^2 = \Sigma X_\alpha^2 - 5 \times \bar{X}^2 = 2.5420$$

$$4 S_2^2 = \Sigma Y_\alpha^2 - 5 \times \bar{Y}^2 = 9.4400$$

$$F = \frac{2.5420}{9.4400} \doteqdot 0.27$$

$F_4^4(0.01) = 15.98$，$F_4^4(0.05) = 6.39$ だから Significant でない。即ち Normality を棄却する訳にわ行かない。

このような Normality Test の Efficiency については又別の機会にのべ度い。　（1950. 4. 10）

(1) C. Hayashi : Fragments of a New Test Formula of Normality, Annals of the Institute of Statistical Math. Vol. 1. No. 2 1950. pp. 125 − 130

(2) 例えば, C.C. Craig : Note on the Distribution of Non-Central t with An Applicatiin, Ann. of Math. Stat. Vol. 12. 1941.
N. L. Johnson and B. L. Welch : Applicatiins of the non-central t-distribution. Biimetrika Vol. 31. 1940. pp. 362 − 389.

(3) Student : The probable error of a mean : Brimetrika. Vol. 6. 1908. p. 1.

(受付 1950 9.14)