On Midzuno's Inequality

By Hirojiro AOYAMA

In this series H. Midzuno has introduced the following inequality in regard to the confidence intervals for the mean, assuming that the distribution of population has 2λ-th order (λ=natural number) central moment:

$$P_r\{|\bar{x}-\bar{X}|\geqq k\sigma_{\bar{x}}\}\leqq\frac{1\cdot3\cdot5\cdots\cdots(2\lambda-1)}{k^{\lambda}} \tag{1}$$

where sample mean is $\bar{x}$, population mean, $\bar{X}$, standard deviation of sample mean, $\sigma_{\bar{x}}$.[(1)]

In his proof, however, some terms of higher order are neglected in regard to principal term. So here we will prove analytically that inequality not only for the mean but also for the original variate under the analogous assumptions.

Let the size of population be infinite and assume that

(i) $f(x)$ is the probability density function of random variable x in $(-\infty, \infty)$,

(ii) $f(x)$ is continuous and differentiable to the second derivative,

(iii) for any large integer λ

$$\lim_{x\to\pm\infty} x^{\lambda} f(x) = \lim_{x\to\pm\infty} x^{\lambda}\frac{df}{dx}=0,$$

(iv) $\int_{-\infty}^{\infty}\left|\frac{d^2f}{dx^2}\right|e^{\frac{x^2}{4}}dx$ exists,

(v) for any real number θ, exists

$$\int_{-\infty}^{\infty} e^{\theta x}|f(x)|dx.$$

Then the following equation holds uniformly

$$f(x)=\sum_{n=0}^{\infty}\alpha_n\Phi_n(x) \tag{2}$$

where $\Phi_n(x)$ denotes, by means of Hermite polynomial of n th order $H_n(x)$,

$$\Phi_n(x)=(-1)^n H_n(x)\Phi_0(x) \tag{3}$$

$$\Phi_0(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \tag{4}$$

and without loss of generality we assumed the 1st and 2nd central moment were 0 and 1 respectively.

Then we have

$$\left.\begin{aligned}\int_{-\infty}^{\infty} x^{\lambda}\Phi_n(x)\,dx&=0, && \text{for } n=2m+1\\ &=2^{m+1}\lambda(\lambda-1)\cdots\cdots(\lambda-m)\frac{(2\lambda)!}{2^{\lambda}\lambda!}, && \text{for } n=2m\end{aligned}\right\}\quad(5)$$

Therefore we can deduce from (2)

$$\mu_{\lambda}=\sum_{n=1}^{\infty}\int_{-\infty}^{\infty} x^{\lambda}\alpha_n\Phi_n(x)\,dx$$

$$=\frac{(2\lambda)!}{2^{\lambda}\lambda!}\{1+2^2\lambda(\lambda-1)\alpha_4+2^3\lambda(\lambda-1)(\lambda-2)\alpha_6+\cdots\cdots+2^{\lambda}\lambda!\alpha_{\lambda}\}\quad(6)$$

Evaluating α_r, we have( )

$$|\alpha_r|<\frac{KC}{\sqrt{(r+2)!}}\quad(7)$$

where

$$C=\int_{-\infty}^{\infty}\left|\frac{d^2f}{dx^2}\right|e^{\frac{x^2}{4}}dx\quad(8)$$

$$K^2=\frac{1}{2\pi}\int_{|t|=}\left|\frac{dt}{\sqrt{1-t^2}}\right|=1.18034.\quad(9)$$

So we have from (6)

$$\mu_{\lambda}<\frac{(2\lambda)!}{2^{\lambda}\lambda!}\left\{1+1.09\,C\left(\frac{(2^2\cdot 2!)^2}{5!}\binom{\lambda}{2}+\frac{(2^3\,3!)^2}{7!}\binom{\lambda}{3}+\cdots\cdots+\frac{(2^{\lambda}\lambda!)^2}{(2\lambda+1)!}\binom{\lambda}{\lambda}\right)\right\}\quad(10)$$

$$=1\cdot 3\cdot 5\cdots\cdots(2\lambda-1)\{1+g(\lambda,C)\}\quad(11)$$

When we are empirically permitted to take to α_4, we have

$$g^{*}(\lambda,C)=0.29067\,\lambda(\lambda-1)\,C\quad(12)$$

On the other hand we can easily have for any distribution

$$P_r(|x-E(x)|\geqq k\sigma)\leqq\frac{\mu_{\lambda}}{k^{\lambda}\sigma^{\lambda}},\quad(13)$$

so that we get from (11)

$$P_r(|x-E(x)|\geqq k\sigma)\leqq\frac{1\cdot 3\cdot 5\cdots\cdots(2\lambda-1)}{k^{\lambda}}(1+g(\lambda,C))\quad(14)$$

Putting $k=3$, we have following table and graphs for the right hand side of (14).

λ	the right hand side of (14)	$g(\lambda, C)$	$g^*(\lambda, C)$
1	0.111 11	0	0
2	0.037 04	0.5813 C	0.5813 C
3	0.020 58	2.2423 C	1.7440 C
4	0.016 00	5.9241 C	3.4880 C
5	0.016 00	13.4134 C	5.8134 C
10	0.187 77	383.6086 C	26.1603 C

Fig. 1.

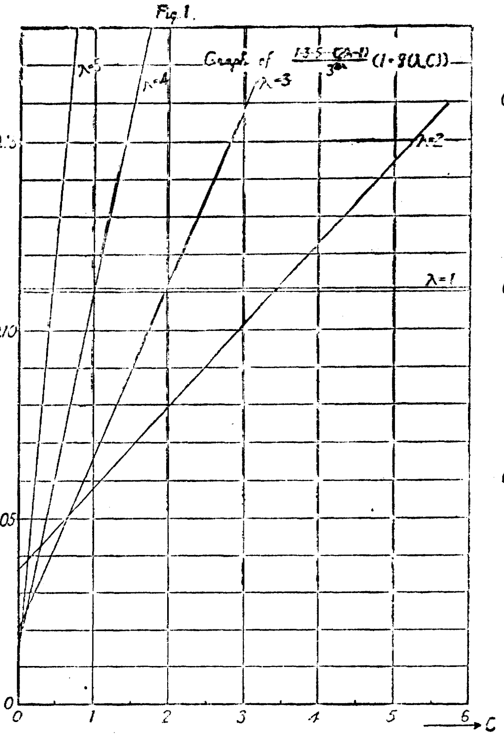
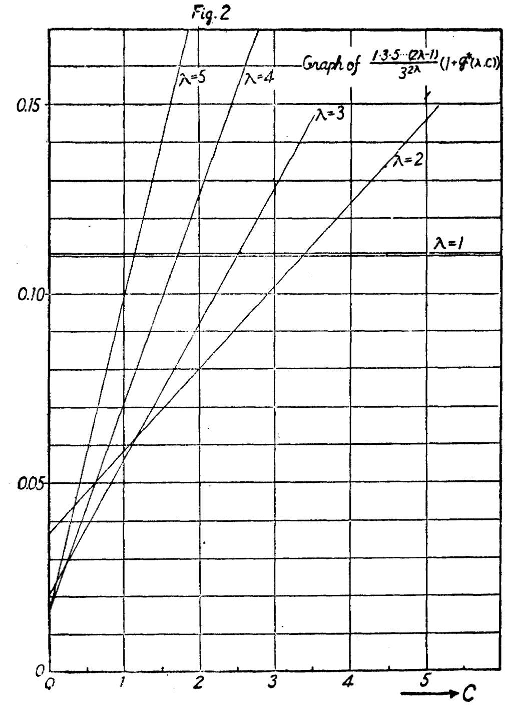


Fig. 2



From (14) we can get the maximum confidence interval taking adequate k and λ. For negligibly small C of uniform distribution, straight line distribution and distribution such as $|f''(x)|=O(e^{-\frac{x^2}{2}})$ etc., we get approximately

$$P_r(|x-E(x)|\geqq k\sigma)\leqq\frac{1\cdot3\cdot5\cdots\cdots(2\lambda-1)}{k^{\lambda}} \tag{15}$$

which is more effective in some points than Gauss' inequality and Cramér's inequality.[3]

When we should get the confidence interval for the mean, we have similarly,

$$P_r(|\bar{x}-E(\bar{x})|\geqq k\sigma_{\bar{x}})\leqq\frac{1\cdot3\cdot5\cdots\cdots(2\lambda-1)}{k^{\lambda}}(1+g(\lambda, C)) \tag{16}$$

Reference

(1) H. MIDZUNO: On Certain Groups of Inequalities, same series Vol. II, No. 1.
(2) J. SUETUNA: Theory of Probabilities, *Iwanami, Shoten Co.*, Japan 1941. H. CRAMÉR: Mathematical Methods of Statistics, *Princeton, Uuiv. Press* 1946 Part III.
(3) K. F. GAUSS: Theoria Combinationis Observationum. H. CRAMÉR: *loc. cit.* p. 256.

Institute of Statistical Mathematics.