ON THE COEFFICIENT OF MEAN DIFFERENCE OF CONTINUOUS DISTRIBUTIONS

BY S.N. NATH

State Statistical Bureau, Govt. of West Bengal, Calcutta

1. INTRODUCTION

The coefficient of mean difference which is due to Gini (1912) may be defined by :

$$\Delta_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} | x - y | \, dF(x) \, dF(y)$$

where $F(x)$ is a continuous distribution function of x $(-\infty \leqslant x \leqslant \infty)$. The appearance of absolute values in the definition makes it extremely difficult to calculate the integral. However, we can get rid of the absolute values in the integrand by using a simpler result due to Kendall (1952).

$$\Delta_1 = 2 \int_{-\infty}^{\infty} F(x) \, \{1 - F(x)\} dx$$

where

$$F(x) = \int_{-\infty}^{x} f(x) dx$$

is the distribution function. But, it is hardly possible to calculate the above integral, except of course, when $F(x)$ is of a simpler form.

Until now no attempt has been made to express Δ_1 in terms of parameters of the distribution function $F(x)$. In this paper we shall integrate this integral and express Δ_1 in the form of an infinite series whose terms depend on $F(x)$ and its derivatives at $x=0$. The expression is checked against well known results for Δ_1.

2. Expression of $\triangle_1$ in an Infinite Series

We shall prove the followi.·g theorem in this section.

Theorem 1. If $F(x)$ be a single valued function of x, continuous in the range $-\infty \leqslant x \leqslant \infty$ and monotone increasing, and if first movement exists, then $\triangle_1$ can be expressed as an infinite series :

$$\triangle_1 = 2(2F_o - 1)\,\mu_1' + 4\sum_{i=1}^{\infty} (b_i/i+2)\,[(1-F_o)^{i+2} - (-F_o)^{i+2}]$$

where

$$F_o = \int_{-\infty}^{0} dF(x),$$

$$\mu_1' = \int_{-\infty}^{\infty} xd\,F(x)$$

and

$$b_i = \left(\frac{1}{i!}\right)(d^{i-1}/dx^{i-1})\left[\{F'(0) + \frac{x}{2!}\,F''(0) + \frac{x^2}{3!}\,F'''(0) + \ldots\}^{-i}\right]_{x=0}$$

Proof.

Under the usual conditions, Maclaurin's Expansion gives :

$$F(x) = F(0) + xF'(0) + \frac{x^2}{2!}F''(0) + \ldots + \frac{x^n}{n!}F^{(n)}(0) + \ldots$$

or,

$$G(x) = x\,F'(0) + \frac{x^2}{2!}F''(0) + \frac{x^3}{3!}F'''(0) + \ldots + \frac{x^n}{n!}F^{(n)}(0) + \ldots$$

where

$$G(x) = F(x) - F(0).$$

By reversing the above series we get,

$$x = \sum_{i=1}^{\infty} b_i\,\{G(x)\}^i \qquad \ldots(2\cdot1)$$

where

$$b_i=(1/i!)(d^{i-1}/dx^{i-1})\left[\{F'(0)+\frac{x}{2!}F''(0)+\frac{x^2}{3!}F'''(0)+\ldots\}^{-i}\right]_{x=0}$$

(Bromwitch, 1947).

Substituting this value of x we have.

$$\mu_1'=\int_{-\infty}^{\infty} xdF$$

$$=\int_{-\infty}^{\infty}[\Sigma b_i\, G^i(x)]\, dG(x)$$

$$=\sum_{i=1}^{\infty}(b_i/i+1)\,[\,(1-F_o)^{i+1}-(-F_o)^{i+1}] \quad \ldots(2\ 2)$$

$$\triangle_1=2\int_{-\infty}^{\infty}F(x)\,\{1-F(x)\}\,dx$$

$$=2\int_{-F_o}^{1-F_o}\left[\sum_{i=1}^{\infty}(G+F_o)\,(1-F_o-G)ib_i\,G^{i-1}\right]dG$$

$$=2\sum_{i=1}^{\infty}ib_i\int_{-F_o}^{1-F_o}\left[[F_o(1-F_o)+(1-2F_o)G-G^2]\,G^{i-1}\,dG\right.$$

$$=2\sum_i b_iF_o(1-F_o)\left[(1-F_o)^i-(-F_o)^i\right]+$$

$$2\sum_i(1-1/i+1)b_i(1-2F_o)\left[(1-F_o)^{i+1}-(-F_o)^{i+1}\right]$$

$$-2\sum_i(1-2/i+2)b_i\left[(1-F_o)^{i+2}-(-F_o)^{i+2}\right]$$

$$=2(2F_o-1)\sum_{i=1}^{\infty}(b_i/i+1)\left[(1-F_o)^{i+1}-(-F_o)^{i+1}\right]$$

$$+4 \sum_{i=1}^{\infty} (b_i/i+2) \left[(1-F_o)^{i+2}-(-F_o)^{i+2} \right].$$

Therefore,

$$\triangle_1=2(2F_o-1)\,\mu_1'+4 \sum_{i=1}^{\infty} (b_i/i+2) \left[(1-F_{10})^{i+2}-(-F_o)^{i+2} \right],$$

[follows from (2·2)]

3. EXPRESSION OF $\triangle_1$ IN SPECIAL CASES

(*a*) Distributions symmetrical about $x=0$.

For a distribution symmetrical about $x=0$ we have $F_o=1/2$. Hence by Theorem 1, we get,

$$\triangle_1=4 \sum_{i=1}^{\infty} (b_i/i+2) \left[(1/2)^{i+2}-(-\tfrac{1}{2})^{i+2}\right]$$

$$= \sum_{i=1}^{\infty}(b_{2i-1}/_{2i+1})/2^{2i-2}. \qquad \ldots(3{\cdot}1)$$

(*b*) Distributions for which $0 \leqslant x \leqslant \infty$.

For such distributions

$F_o=0$ and $\triangle_1$ becomes

$$\triangle_1=-2\mu_1'+4 \sum_{i=1}^{\infty} (b_i/i+2). \qquad \ldots(3{\cdot}2)$$

4. ON THE b COEFFICIENTS.

The b's may be calculated by the formula (2·1) for b_i.

But, it is convenient to calculate the b's by the method of undetermined coefficients.

$$G(x)=xF'(0)+x^2F''(0)/2!+\ldots+x^nF^{(n)}(0)/n!+\ldots$$

and

$$x= \sum_{i=1}^{\infty} b_i\, G^i(x).$$

Substituting this value of x in the infinite series for $G(x)$.

$$G=[b_1F'(0)]G+b_2F'(0)+b_1^2F''(0)/2]G^2+[b_3F'(0)+b_1b_2F''(0)+b_1^3F'''(0)/6]G^3$$
$$+[b_4F'(0)+b_2^2F''(0)/2+b_1b_3F''(0)+b_1^2b_2F'''(0)/2+b_1^4F''''(0)/24]G^4+\ldots \quad \ldots(4\cdot1)$$

and equating coefficients of G, G^2, G^3, G^4...on both sides, we get:

$$\left.\begin{aligned}
b_1&=1/F'(0)=1/f_o\\
b_2&=-b_1^2F''(0)/2F'(0)=-f'(0)/2f_o^3\\
b_3&=-b_1^3\,F'''(0)/6\,F'(0)-b_1b_2\,F''(0)/F'(0)\\
&=\{f'(0)\}^2/2f_o^5-f''(0)/6\,f_o^4\\
b_4&=(1/F'(0))\,[-b_2^2\,F''(0)/2-b_1b_3\,F''(0)-b_1^2b_2\,F'''(0)/2-b_1^4\,F''''(0)/24]\\
&=-5\{f'(0)\}^3/8f_o^7+5f''(0)f'(0)/12f_o^6-f'''(0)/24f_o^5
\end{aligned}\right\}\quad\ldots(4\cdot2)$$

5. IN THIS SECTION THE RESULTS OBTAINED WILL BE CHECKED BY THE KNOWN VALUE OF $\triangle_1$ IN THE CASES WHERE IT CAN BE DIRECTLY OBTAINED.

(*i*) Rectangular Distribution

$$dF=dx/k, \qquad 0\leqslant x\leqslant k$$
$$\mu_1'=k/2, \qquad F(x)=x/k.$$

So,

$$\triangle_1=2\int_0^k (x/k)(1-x/k)dx=k/3. \quad \ldots(5\cdot1)$$

From result (3·2) in this case

$$\triangle_1=-k+4\sum_{i=1}^{\infty}(b_i/i+2).$$

The b's may be calculated from (4·2) we have

$$b_1=k, \qquad b_i=0, \qquad i>1.$$

Hence, $\triangle_1=k/3$, which is the same as obtained in (5·1).

(*ii*) Exponential Distribution.

$$dF=(1/\sigma)e^{-x/\sigma}dx, \qquad 0\leqslant x\leqslant \infty$$

$$F(x)=1-e^{-x/\sigma},$$

$$\mu_1'=\int_0^\infty x dF=\sigma\Gamma(2)=\sigma$$

$$\triangle_1=2\int_0^\infty (1-e^{-x/\sigma})(e^{-x/\sigma})dx=\sigma. \qquad \ldots(5{\cdot}2)$$

In this case b's are directly calculated as follows. We have,

$$F(x)=1-e^{-x/\sigma},\ F(0)=0,$$

and

$$F^{(n)}(0)=(-1)^{n+1}/\sigma^n.$$

And from (4·1)

$$\begin{aligned}G(x)&=(1/\sigma)[b_1G(x)+b_2G^2(x)+\ldots]-(1/\sigma^2)[b_1G(x)+b_2G^2(x)+\ldots]^2/2!\\&\quad+(1/\sigma^3)[b_1G(x)+b_2G^2(x)+\ldots]^3/3!+\ldots\\&=1-\exp.[-\{b_1G(x)+b_2G^2(x)+\ldots\}/\sigma]\end{aligned}$$

or,

$$\begin{aligned}-(b_1/\sigma)G(x)-(b_2/\sigma)G^2(x)-\ldots&=\log(1-G(x))\\&=-G(x)-\frac{G^2(x)}{2}-\frac{G^3(x)}{3}-\ldots\end{aligned}$$

[as $G(x) < 1$].

Equating coefficients of G, G^2, G^3,...etc. on both sides.

$$b_1=\sigma,\quad b_2=\sigma/2.\quad b_3=\sigma/3,\ldots,\quad b_i=\sigma/i,\ldots.$$

Substituting these values of b's in the expression for $\triangle_1$ obtained from (3·2).

$$\triangle_1=-2\sigma+4\sigma\left[\frac{1}{1\cdot3}+\frac{1}{2\cdot4}+\frac{1}{3\cdot5}+\ldots+\frac{1}{n(n+2)}+\cdots\right].$$

Since

$$\sum_{n=1}^{\infty}\frac{1}{n(n+2)}=\sum_{i=1}^{\infty}\frac{1}{2}\left[\frac{1}{n}-\frac{1}{n+2}\right]=3/4$$

$\triangle_1=\sigma$, which is the same as the result obtained in (5·2).

6. Here We Shall Calculate Δ_1 in the Case of Normal Distribution.

$$dF = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} dx , \quad -\infty \leqslant x \leqslant \infty$$

$$F'(x) = (1/\sigma\sqrt{2\pi})\, e^{-x^2/2\sigma^2}, \; F'(0) = 1/\sigma\sqrt{2\pi}, \; F''(0) = 0$$

and using Hermite polynomials, generally

$$F^{(2m)}(0) = 0, \quad F^{(2m+1)}(0) = -(2m-1)\, F^{(2m-1)}(0).$$

From (4·2) or by following (4·1) we have $b_{2i} = 0$

and

$$b_1 = \sigma\sqrt{2\pi}, \quad b_3 = \sigma(\sqrt{2\pi})^3/6, \quad b_5 = \sigma.7(\sqrt{2\pi})^5/120$$

$$b_7 = \sigma \cdot 127(\sqrt{2\pi})^7/5040, \text{ etc.}$$

Substituting these values of b-coefficients in the expression

$$\Delta_1 = \sum_{i=1}^{\infty} (b_{2i-1})/(2i+1)2^{2i-2},$$

we get

$$\Delta_1 = \sigma[\sqrt{2\pi}/3 + (\sqrt{2\pi})^3/120 + (\sqrt{2\pi})^5/1920 + 127\,(\sqrt{2\pi})^7/2903040 + \ldots]$$

$$= \sigma[0{\cdot}835 + 0{\cdot}131 + 0{\cdot}051 + 0{\cdot}027 + \ldots]$$

$$= \sigma.[1{\cdot}044 + \ldots]$$

$$= (1{\cdot}044)\sigma, \text{ approximately.}$$

REFERENCE

1. Bromwitch T.J.I.A. & Macrobert, T.M. (1917) : An Introduction to the Theory of Infinite Series. (Macmillan & Co., London) pp. 158-159.

2. Gini, C. (1912) : "Variabititae Mutabitita" Studi Economico—Giuridici della R. Umiversita, Cagliari, anne 3, Part 2, p. 80.

3. Kendall, M.G. (1952) : The Advanced Theory of Statistics (Charles Griffin & Company Ltd., London) pp, 42-48.