Journal of Science of the Gakugei Faculty,
Tokushima University
Vol. 1, 1950

UNBIASED ESTIMATE OF THE MEAN ABSOLUTE DEVIATION

By

Yoshikatsu WATANABE

(Received Sept. 15, 1950)

If the observed values of a quantity z be $z_1, \dots, z_n$ and the mean value $\bar{z}=\sum z_j/n$, then the deviations are $z_j-\bar{z}=x_j$, $j=1, \dots, n$. The unbiased estimate of the mean square deviation is given by the well-known Bessel's formula

$$\hat{\sigma}^2 = \sum x_j^2/(n-1). \tag{1}$$

However its demonstrations are found hardly legitimate in classical books on least squares, except some fews, e. g. A. F. Craig's elegant proof given in Bulletin of the American Math. Soc., 1936, vol. 42. He pointed out that (1) means nothing but the expectation of the sum of squares

$$\sum x_j^2 = V, \quad \text{i.e.} \quad E(V) = (n-1)\sigma^2, \tag{2}$$

and proved (2) under the assumption that x distributes normally. In the present note a similar process is applied to generalize Peters' formula in regard to the mean absolute deviation ϑ

$$\hat{\vartheta} = \sum |x_j| / \sqrt{n(n-1)}.$$

1° Characteristic. As well known, the distribution function (density of probability) $f(x)$ as well as its characteristic $g(t)$ are defined as follows[1)]

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ixt} g(t)dt, \quad g(t) = \int_{-\infty}^{\infty} e^{ixt} f(x)dx. \tag{3}$$

Or, in case of many variables,

$$\left.\begin{aligned} f(x_1, \dots, x_m) &= \frac{1}{(2\pi)^m}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-L} g(t_1, \dots, t_m)dt_1 \dots dt_m, \\ g(t_1, \dots, t_m) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{L} f(x_1, \dots, x_m)dx_1 \dots dx_m, \\ \text{where} \qquad L &= ix_1t_1 + \cdots + ix_mt_m. \end{aligned}\right\} \tag{4}$$

1) See the annexed References.

Now, for a single valued continuous function $u=U(x_1, \dots, x_n)$, where $x_1, \dots, x_n$ are assumed to be independent, the distribution function $F(u)$ shall be given by

$$F(u)du=\int_D \cdots \int f(x_1, \dots, x_n)dx_1 \dots dx_n \ ,$$

where $x_1, \dots, x_n$ should be taken over the domain D that satisfies the inequality $u \leqq U(x_1, \dots, x_n) \leqq u+du$. To avoid this inconvenience, let us multiply, after Cauchy's devise, both sides by the function[2]

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dt \int_u^{u+du} e^{it(U-\xi)} d\xi \ ,$$

which becomes $=1$ in $\langle u, u+du \rangle$, and otherwise $=0$. Then the domain of integration can be extended to the whole n-dimensional space R_n, so that

$$F(u)du=\frac{1}{2\pi}\int_{-\infty}^{\infty} dt \int_u^{u+du} e^{-it\xi} d\xi \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{itU} f(x_1, \dots, x_n)dx_1 \dots dx_n \ .$$

Or putting the inner integral

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{itU} f(x_1, \dots, x_n)dx_1 \dots dx_n = G(t) \ , \tag{5}$$

we get

$$F(u)du \simeq \frac{1}{2\pi}\int_{-\infty}^{\infty} \left[e^{-itu} du \right] G(t) dt \ ,$$

hence

$$F(u)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itu} G(t) dt \ . \tag{6}$$

Thus $G(t)$ gives the characteristic of $F(u)$. Furthermore letting

$$u_k=U_k(x_1, \dots, x_n) \ , \quad k=1, 2, \dots, m, \quad (m<n) \ ,$$

the distribution function $F(u_1, \dots, u_m)$ shall be defined by

$$F du_1 \dots du_m = \int_D \cdots \int f(x_1, \dots, x_n) dx_1 \dots dx_n \ ,$$

where D denotes such a domain that $u_k \leqq U_k \leqq u_k+du_k$, $k=1, \dots, m$. Here again repeating Cauchy's devises m times, and putting as the characteristic

$$G(t_1, \dots, t_m)=\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\vartheta} f(x_1, \dots, x_n)dx_1 \dots dx_n \ , \tag{7}$$

---

2) Cf. the annexed References (*).

where $\theta = it_1U_1 + it_2U_2 + \cdots + it_mU_m$, we obtain

$$F(u_1, \ldots, u_m) = \frac{1}{(2\pi)^m}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-\theta}G(t_1, \ldots, t_m)dt_1 \ldots dt_m\,. \tag{8}$$

2° In our case, assuming that x distributes normally

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{x^2}{2\sigma^2}\right\} \tag{9}$$

under the conditions that

$$u_k = \sum_{\nu=1}^{n} a_{k\nu}x^{\nu}\,, \quad k = 1, \ldots, m, \text{ and } u_{m+1}(=u) = \sum_{\nu=1}^{n}|x_\nu|\,, \tag{10}$$

where $1 \leq m \leq n$, and all the coefficients are real, the characteristic in accordance with (7) is given by

$$G(t_1, \ldots, t_{m+1}) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left\{it_1u_1 + \cdots + it_{m+1}u_{m+1}\right\}f(x_1, \ldots, x_n)dx_1 \ldots dx_n\,,$$

where f is determined from (9) to be

$$f(x_1, \ldots, x_n) = \frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left\{-V/2\sigma^2\right\}, \quad V = \sum_{\nu=1}^{n} x_\nu^2\,.$$

Hence we obtain

$$\left.\begin{aligned} G(t_1, \ldots, t_{m+1}) &= \frac{1}{(\sqrt{2\pi}\sigma)^n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{\Theta}dx_1 \ldots dx_n\,, \\ \text{where} \quad \Theta &= i\sum_{l=1}^{m+1} t_l u_l - V/2\sigma^2\,. \end{aligned}\right\} \tag{11}$$

In order to perform the integration, we write

$$\begin{aligned} \Theta = &-\frac{1}{2\sigma^2}\sum_{x_\nu>0}{}' \left\{\left[x_\nu - i\sigma^2(s_\nu + t_{m+1})\right]^2 + \sigma^4(s_\nu + t_{m+1})^2\right\} \\ &-\frac{1}{2\sigma^2}\sum_{x_\nu<0}{}'' \left\{\left[x_\nu - i\sigma^2(s_\nu - t_{m+1})\right]^2 + \sigma^4(s_\nu - t_{m+1})^2\right\} \end{aligned} \tag{12}$$

where

$$s_\nu = \sum_{l=1}^{m} a_{l\nu}t_\nu\,. \tag{13}$$

So that (11) becomes

$$G(t_1, \ldots, t_{m+1}) = \prod_{\nu=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\left[I_\nu\exp\left\{-\frac{\sigma^2}{2}(s_\nu + t_{m+1})^2\right\} + J_\nu\exp\left\{-\frac{\sigma^2}{2}(s_\nu - t_{m+1})^2\right\}\right], \tag{14}$$

where

$$\begin{matrix} I_\nu \\ J_\nu \end{matrix} = \int_0^{\infty}\exp\left\{-\frac{1}{2\sigma^2}\left[x_\nu - i\sigma^2(s_\nu \pm t_{m+1})\right]^2\right\}dx_\nu\,. \tag{15}$$

The integrals (15) can be found by utilizing Cauchy's integral theorem to be

$$\sqrt{\frac{\pi}{2}}\sigma \pm i\sigma\int_0^{\sigma(s_\nu \pm t_{m+1})} e^{t^2/2}dt ,$$

so that (14) may be transformed into

$$G(t_1, \dots, t_{m+1}) = \prod_{\nu=1}^{n} \exp\left\{-\frac{\sigma^2}{2}(s_\nu^2 + t_{m+1})\right\}\Big[\cosh(\sigma^2 s_\nu t_{m+1}) + \frac{i}{\sqrt{2\pi}}\Big(\exp\left(-\frac{1}{2}\sigma^2 s_\nu t_{m+1}\right)\int_0^{\sigma(s_\nu + t_{m+1})} e^{t^2/2}dt - \exp\left(-\frac{1}{2}\sigma^2 s_\nu t_{m+1}\right)\int_0^{\sigma(s_\nu - t_{m+1})} e^{t^2/2}dt\Big)\Big]. \quad (16)$$

In particular letting $t_{m+1}=0$, we obtain as the characteristic of the combined distributions function $\Phi(u_1, \dots, u_m)$,

$$G(t_1, \dots, t_m, 0) = \prod_{\nu=1}^{n} \exp\left\{-\frac{\sigma^2}{2}s_\nu^2\right\} \equiv \exp\left\{-\frac{\sigma^2}{2}Q\right\} . \quad (17)$$

Since the numbers considered are all real

$$Q = \sum_{\nu=1}^{n} s_\nu^2 = \sum_{k,l=1}^{m} \sum_{\nu=1}^{n} a_{k\nu}a_{l\nu}t_k t_l = \sum_{k,l} b_{kl}t_k t_l$$

is a positive definite Hermite form (Bt, t), so that the matrix $B=(b_{kl})$ can be transformed into a diagonal one Λ by taking an adequate orthogonal matrix $C=(c_{kl})$ so as $C'BC=\Lambda$: or in other words, by the orthogonal transformation $t=Cz$: $t_l = \sum_k c_{lk}z_k$, we obtain

$$Q = \sum_{l=1}^{m} \lambda_l z_l^2 \ (>0 \text{ except when all } z_l = 0) ,$$

where the coefficients are all >0, and the jacobian $J=\dfrac{\partial(t_1, \dots, t_m)}{\partial(z_1, \dots, z_m)}=1$. Thus we get

$$G(t_1, \dots, t_m, 0) = \exp\left\{-\frac{\sigma^2}{2}\sum_{l=1}^{m}\lambda_l z_l^2\right\} , \quad (18)$$

and by (8) the corresponding distribution function becomes

$$\Phi(u_1, \dots, u_m) = \frac{1}{(2\pi)^m}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left\{-L-\frac{\sigma^2}{2}Q\right\}dt_1 \dots dt_m$$
$$= \frac{1}{(2\pi)^m}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left\{-L-\frac{\sigma^2}{2}Q\right\}dz_1 \dots dz_m , \quad (19)$$

where $L = i\sum_{l=1}^{m} t_l u_l = i\sum_l u_l \sum_k c_{lk}z_k = i\sum_k v_k z_k$, and $v_k = \sum_l c_{lk}u_l$.

Now that the multiple integral in (19) might be decomposed into a product of the form

$$\prod_{l=1}^{m}\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left\{-\frac{\sigma^2}{2}\left(\lambda_l z_l^2 + \frac{2i}{\sigma^2}v_l z_l\right)\right\}dz_l , \quad (20)$$

the integration could be performed by availing Cauchy's integral theorem, and we get finally

$$\Phi(u_1, \dots, u_m) = \frac{1}{\sqrt{\lambda_1 \dots \lambda_m}(\sqrt{2\pi}\sigma)^m} \exp\left\{-\frac{1}{2\sigma^2}\sum_{l=1}^{m}\left(\sum_{k=1}^{m} c_{lk}u_k\right)^2 / \lambda_l\right\}. \tag{21}$$

In particular, if $u_1 = \dots = u_m = 0$, we have

$$\Phi(0, \dots, 0) = 1/\sqrt{\lambda_1 \dots \lambda_m}\,(\sqrt{2\pi}\sigma)^m. \tag{22}$$

3° The expectation of $\sum|x_\nu| = u$. Let $F(u_1, \dots, u_m, u)$ be the combined distribution function of $u = \sum|x_\nu|$, and $u_k (k=1, \dots, m)$ given in (10), and let us find the expectation of u, when $u_1, \dots, u_m$ are assumed to be fest. Here, since the compound probability for $u_1, \dots, u_m$ is $\Phi(u_1, \dots, u_m)\, du_1 \dots du_m$, while the compound probability for $u_1, \dots, u_m$ and u is $F(u_1, \dots, u_m, u)\, du_1 \dots du_m du$, the relative probability becomes F/Φ, and accordingly the required expectation can be given by

$$\bar{u} = \int u\frac{F}{\Phi}du, \tag{23}$$

where the integration should be extended over all the possible values of $u(\geqq 0)$ as far as $u_1, \dots, u_m$ preserve the given fest values. But in virtue of (4), the characteristic $G(t_1, \dots, t_{m+1})$ of $F(u_1, \dots, u_m, u)$ is given by

$$G(t_1, \dots, t_m, t_{m+1}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp\left\{L + iut_{m+1}\right\} F du_1 \dots du_m du,$$

where $L = i\sum_{l=1}^{m} t_l u_l$. Whence we get

$$\left(\frac{\partial G}{\partial t_{m+1}}\right)_{t_{m+1}=0} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} iue^L \frac{F}{\Phi} \Phi du du_1 \dots du_m,$$

which becomes after substitution of (23)

$$\left(\frac{\partial G}{\partial t_{m+1}}\right)_0 = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} i\bar{u}e^L \Phi du_1 \dots du_m.$$

Therefore inversely we get by (4)

$$i\bar{u}\Phi = \frac{1}{(2\pi)^m}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-L}\left(\frac{\partial G}{\partial t_{m+1}}\right)_0 dt_1 \dots dt_m. \tag{24}$$

On the other hand we have by (16)

$$G = \prod_{\nu=1}^{n} g_\nu(t_{m+1}),$$

where $g_\nu(t_{m+1})$ denotes the ν-th factor in (16), and consequently

$$g_\nu(0)=\exp\left\{-\frac{\sigma^2}{2}s_\nu^2\right\}, \qquad s_\nu=\sum_{l=1}^{m}a_{l\nu}t_l ,$$

and
$$g_\nu'(0)=i\sigma\sqrt{\frac{2}{\pi}}\left[1-\sigma s_\nu\exp\left(-\frac{\sigma^2}{2}s_\nu^2\right)\int_0^{\sigma s_\nu}e^{t^2/2}dt\right] .$$

$$\therefore\quad \left(\frac{\partial G}{\partial t_{m+1}}\right)_0=\sum_{\nu=1}^{n}g_\nu'(0)\exp\left(-\frac{\sigma^2}{2}\sum_{\mu\neq\nu}^{n}s_\mu^2\right) .$$

This value being substituted in (24), we ought to integrate it, which is somewhat troublesome. However if the first m conditions of (10) be linear homogeneous, that is, if $u_1=\cdots=u_m=0$, then we can perform further integrations. Really, form the result just obtained together with (22), we get

$$\bar{u}=\frac{\sqrt{\lambda_1\ldots\lambda_m}\,\sigma^{m+1}}{\sqrt{2\pi}^{m}}\sqrt{\frac{2}{\pi}}\sum_{\nu=1}^{n}\left\{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left[-\frac{\sigma^2}{2}\left(Q-s_\nu^2\right)\right]dt_1\ldots dt_m\right.$$
$$\left.-\sigma\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left(-\frac{\sigma^2}{2}Q\right)s_\nu\int_0^{\sigma s_\nu}e^{t^2/2}dt_1\ldots dt_m\right\} , \tag{25}$$

Let us evaluate the multiple integrals in the above large bracket, (i) and (ii) say. In order to facilitate the integral (i), we must adopt another orthogonal transformation $t=C_\nu\zeta$, so that the expression in the exponent, which is also a positive definite Hermite form, becomes

$$Q_\nu\equiv Q-s_\nu^2=\sum_{\mu\neq\nu}s_\mu^2=\sum_{k,l}^{m}\left(\sum_{\mu\neq\nu}^{n}a_{k\mu}a_{l\mu}\right)t_kt_l=\sum_l\lambda_{l\nu}\zeta_l^2 ,$$

and after integrations we get

$$(\mathrm{i})=\left(\frac{\sqrt{2\pi}}{\sigma}\right)^m\Big/\sqrt{\lambda_{1\nu}\ldots\lambda_{m\nu}} . \tag{26}$$

In regard to (ii), we utilize at first the before mentioned transformation $t=Cz$, so that $Q=\sum_l\lambda_l z_l^2$, and $s_\nu=\sum_l a_{l\nu}t_l=\sum_l d_{\nu l}z_l$, where $d_{\nu l}$ denotes the νl-element of the matrix $d=a'c$. Integrating by parts, we obtain

$$(\mathrm{ii})=\sigma\sum_{l=1}^{m}\frac{d_{\nu l}^2}{\lambda_l}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left\{-\frac{\sigma^2}{2}\left[\sum_{k=1}^{m}\lambda_k z_k^2-\left(\sum_{k=1}^{m}d_{\nu k}z_k\right)^2\right]\right\}dz_1\ldots dz_m .$$

Here the expression in the exponent becomes

$$Q_\nu^*\equiv\sum_{k,l}^{m}(\delta_{kl}\sqrt{\lambda_k\lambda_l}-d_{\nu k}d_{\nu l})z_kz_l \quad (\delta_{kl}=\text{Kronecker's delta})$$
$$=\sum_{l=1}^{m}\lambda_{l\nu}^*\zeta_l^{*2}$$

by a third orthogonal transformation $z=C_\nu^*\zeta^*$, and after integration, we get

$$(\text{ii})=\frac{\sqrt{2\pi}^m}{\sigma^{m+1}}\sum_{l=1}^{m}\frac{d_{\nu l}^2}{\lambda_l\sqrt{\lambda_{1\nu}^*\dots\lambda_{m\nu}^*}}\;. \tag{27}$$

Substituting (26) and (27) in (25), we get, as the generalized Peters' formula,

$$\bar{u}=\sqrt{\lambda_1\dots\lambda_m}\sqrt{\frac{2}{\pi}}\,\sigma\sum_{\nu=1}^{n}\left[\frac{1}{\sqrt{\lambda_{1\nu}\dots\lambda_{m\nu}}}-\sum_{l=1}^{m}\frac{d_{\nu l}^2}{\lambda_l\sqrt{\lambda_{1\nu}^*\dots\lambda_{m\nu}^*}}\right]. \tag{28}$$

Specially if $m=1$, we obtain (in omitting the suffix 1)

$$\lambda=\sum_{\nu=1}^{n}a_\nu^2\,,\qquad \lambda_\nu=\sum_{\mu\neq\nu}^{n}a_\mu^2=\lambda-a_\nu^2\,,\qquad \lambda_\nu^*=\lambda_\nu\quad(\because\ d_\nu=a_\nu)\,,$$

and hence

$$\bar{u}=\sqrt{\frac{2}{\pi}}\,\sigma\sum_{\nu=1}^{n}\sqrt{\frac{\lambda-a_\nu^2}{\lambda}}=\sqrt{\frac{2}{\pi}}\,\sigma\sum_{\nu=1}^{n}\sqrt{\sum_{\mu\neq\nu}^{n}a_\mu^2\Big/\sum_{\nu=1}^{n}a_\nu^2}\;. \tag{29}$$

More specially in case that all $a_\nu=1$, i.e. $u_1\equiv\sum x_\nu=0$, as in the case of the residual sum of least squares, we have $\lambda=n$, and

$$\bar{u}=\sqrt{\frac{2}{\pi}}\,\sigma\sqrt{n(n-1)}\;.$$

But the absolute mean is

$$\vartheta=\frac{2}{\sqrt{2\pi}\,\sigma}\int_0^\infty x\exp\left\{-\frac{x^2}{2\sigma^2}\right\}dx=\sqrt{\frac{2}{\pi}}\,\sigma\;,$$

so that we get $\bar{u}=\vartheta\sqrt{n(n-1)}$, which means that

$$\hat{\vartheta}=\frac{u}{\sqrt{n(n-1)}}=\frac{\sum|x_\nu|}{\sqrt{n(n-1)}}\;, \tag{30}$$

the so-called Peters' formula.

Usually the unbiased estimate of σ conveniently calculated from Bessel's formula $\hat{\sigma^2}=\sum x_\nu^2/(n-1)$, to be $\hat{\sigma}=\sqrt{\sum x_\nu^2/(n-1)}$. However this is not correct, because $\sqrt{\hat{\sigma^2}}\neq\hat{\sigma}$. It will be rather reasonable to avail the above obtained result $\bar{u}=\sqrt{\frac{2}{\pi}}\,\sigma\sqrt{n(n-1)}$, and to put

$$\hat{\sigma}=\sqrt{\frac{\pi}{2}}\,\frac{\sum|x_\nu|}{\sqrt{n(n-1)}}\;. \tag{31}$$

Gakugei Faculty, Tokushima University.

References

According to Fourier's integral-theorem, we have

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}dt\int_{-\infty}^{\infty}e^{it(\xi-x)}f(\xi)d\xi\ , \qquad i=\sqrt{-1}\ .$$

This follows readily from Fourier's expansion

$$f(x)=\frac{1}{l}\sum_{n=0}^{\infty}\int_{-l}^{l}f(\xi)\cos\frac{n\pi}{l}(\xi-x)d\xi$$

by writing $\frac{n\pi}{l}=t$, $\frac{\pi}{l}=dt$, and making $n, l\to\infty$,

$$f(x)=\frac{1}{\pi}\int_{0}^{\infty}dt\int_{-\infty}^{\infty}f(\xi)\cos t(\xi-x)d\xi\ .$$

(cf. e.g. Prof. Takagi's Treatise on Analysis, p. 336).
Or

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}dt\int_{-\infty}^{\infty}f(\xi)e^{it(\xi-x)}d\xi=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-ixt}dt\int_{-\infty}^{\infty}e^{it\xi}f(\xi)d\xi\ ,$$

whence the relations (3) immediately follow.

Specially, if in $a\leqq x\leqq b$, $f(x)=1$, and otherwise $f(x)=0$, we obtain

$$\left.\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty}dt\int_{-\infty}^{\infty}e^{it(x-\xi)}d\xi&=1\ , &&\text{in } a\leqq x\leqq b\ ,\\ &=0\ , &&\text{otherwise.}\end{aligned}\right\} \qquad (*)$$