Let $B_j$ ($j = 2, 3, \cdots q$) stand for the set of constants

$$\beta_{ru}, \qquad \begin{array}{l} r = l_{j-1} + 1, l_{j-1} + 2, \cdots l_j, \\ u = 1, 2, \cdots l_{j-1}, \end{array}$$

and let

(11) $$B_j = 0$$

imply the vanishing of all the constants of the set $B_j$. The $q$ sets of variates will be independent if (11) holds for all values of $j$ from 2 to $q$. Denote by $I(B_q, B_{q-1} \cdots, \cdots, B_2)$ the integral of (10) over the region $W - w$. Integrating first over the sub-region of $W - w$ for which

$$\prod_{j=2}^{q-1} \lambda_j$$

has a given value and using the result of section 4, it follows that

$$I(B_q, B_{q-1}, \cdots B_2) \leq I(0, B_{q-1}, \cdots B_2).$$

Also if $B_q = 0$, $\lambda_q$ is distributed independently of $\lambda_{q-1}$. Hence starting with $B_q = 0$ in (10) and integrating for $\lambda_{q-1}$ first, we obtain

$$I(0, B_{q-1}, B_{q-2}, \cdots B_2) \leq I(0, 0, B_{q-2}, \cdots B_2).$$

Thus finally, $I(B_q, B_{q-1}, \cdots B_2) \leq I(0, 0, \cdots 0)$, which proves the completely unbiassed character of the Wilks criterion.

## REFERENCES


[1] J. E. DALY, "On the unbiased character of likelihood ratio tests for independence in normal system," *Annals of Math. Stat.*, Vol. 11 (1940), p. 1.

[2] P. C. TANG, "The power function of analysis of variance tests with tables and illustrations of their use," *Stat. Res. Mem.*, Vol. 2 (1938), p. 126.

[3] R. D. NARAIN, "A new approach to sampling distribution of the multivariate normal theory. Part I," *Jour. Ind. Soc. Agr. Stat.*, Vol. 1 (1948), p. 59.

[4] R. D. NARAIN, "A new approach to sampling distributions of the multivariate normal theory. Part II," *Jour. Ind. Soc. Agr. Stat.*, Vol. 1 (1948), p. 137.


# ON THE DISTRIBUTION OF THE TWO CLOSEST AMONG A SET OF THREE OBSERVATIONS[1]


BY G. R. SETH

*Iowa State College*


**1. Introduction.** In this note we obtain the joint distribution of the two closest observation $x', x''$ ($x' < x''$) of the set $x_1, x_2, x_3$ ($x_1 \leq x_2 \leq x_3$) when the distribution of $x_1, x_2, x_3$ is given or can be obtained.[2] We will assume that in general the density function is given by $f(x_1, x_2, x_3)$ and that it is continuous in the

[1] The results in this paper were presented at a meeting of the Institute of Mathematical Statistics in Madison, Wisconsin, September 9, 1948.

[2] The author's attention was drawn to this problem while visiting the National Bureau of Standards in the Spring of 1948, by Mr. Julius Lieblein of the Statistical Engineering

variables involved. We also find the distributions of certain statistics depending on $x'$ and $x''$. We will denote the density and the cumulative distribution function of a normal variate with mean zero and unit variance by $\phi(x)$ and $G(x)$.

**2. Distribution of the two closest.** Let $x'$, $x''$ be the two closest among the set of $x_1, x_2, x_3$ $(x_1 \leq x_2 \leq x_3)$. Let $P(S_1, S_2, \cdots, S_k)$ denote the probability that the events $S_1, S_2, \cdots, S_k$ occur. Let us consider $P(x' < s, x'' < t)$, for $t < s$. For $s < t$, it reduces to $P(x'' < t)$ i.e. the marginal cumulative distribution of $x''$.

Now

$$
\begin{aligned}
P(x' < s, x'' < t) &= P(x_1 < s, x_2 < t, x_2 - x_1 < x_3 - x_2) \\
&\quad + P(x_2 < s, x_3 < t, x_2 - x_1 < x_3 - x_2).
\end{aligned}
\tag{1}
$$

The equalities, here as well as elsewhere, are omitted as the variables admit continuous distributions. Let the first and second terms on the right side in (1) be denoted by $P(A)$ and $P(B)$ respectively, where $A$, $B$ denote the events in the respective brackets. The event $B$ can be further split up into more elementary events whose probabilities can be easily found. $(B)$ can be seen to be equivalent to

$$
\begin{aligned}
&\left(x_1 < 2s - t, \quad x_1 < x_2 < \frac{x_1 + t}{2}, \quad x_2 < x_3 < 2x_2 - x_1\right) \\
&+ (2s - t < x_1 < s, \quad x_1 < x_2 < s, \quad x_2 < x_3 < 2x_2 - x_1) \\
&+ \left(x_1 < 2s - t, \quad \frac{x_1 + t}{2} < x_2 < s, \quad x_2 < x_3 < t\right).
\end{aligned}
$$

We may write (1) in the form of integrals and differentiating under the integral sign with respect to $t$ and $s$ we obtain

$$
\frac{\partial^2 P}{\partial s \partial t} = \int_{2t-s}^{\infty} f(s, t, x_3)\, dx_3 + \int_{-\infty}^{2s-t} f(x_1\, s, t)\, dx_1 . \tag{2}
$$

The right hand side of (2) gives the density function of $x'$, $x''$ at $x' = s$, $x'' = t$. Let $f_{ij}(x_i, x_j)$ be the density function of $x_i$ and $x_j$ $(i > j = 1, 2, 3)$. Then the density function $p(x', x'')$ of $x'$ and $x''$ can be put into the form

$$
\begin{aligned}
p(x', x'') &= f_{12}(x', x'')\,[1 - F_3(2x'' - x' \mid x_1 = x', x_2 = x'')] \\
&\quad + f_{23}(x', x'')\,[F_1(2x' - x'' \mid x_2 = x', x_3 = x'')],
\end{aligned}
\tag{3}
$$

where $F_i(x_i \mid x_j = l, x_k = m)$ represents the cumulative distribution function of the conditional density function of $x_i$ when $x_j$ and $x_k$ are fixed at the values $l$ and $m$ respectively. If, before ordering, the three observations are independent

Laboratory. He understands that Mr. Lieblein has in preparation for submission to the *Journal of Research of the National Bureau of Standards* a paper giving intensive consideration to the closest pair and other aspects of samples of three observations.

and from the same population having the density function $f(x)$, then (3) with the help of

$$f(x_1, x_2, x_3) = 6f(x_1)\,f(x_2)\,f(x_3)$$

reduces to

$$p(x', x'') = 6f(x')f(x'')[1 - F(2x'' - x') + F(2x' - x'')] \tag{4}$$

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

**3. Joint distribution of** $(x'' - x')$ **and** $(x'' - x')/(x_3 - x_1)$**.** Let $F_1(s, t)$ denote the cumulative distribution function of $u = x'' - x'$ and $w = \dfrac{x'' - x'}{x_3 - x_1}$. Then

$$F_1(s, t) = P\left[x'' - x' < s, \frac{x'' - x'}{x_3 - x_1} < t\right]. \tag{5}$$

The range for $u$ is $(0, \infty)$ and $w$ varies between 0 and $\frac{1}{2}$, and thus we limit ourselves to $s$ varying from 0 to $\infty$, and $t$ varying in $(0, \frac{1}{2})$.

After some manipulation of the probability statement and differentiating with respect to $s$ and $t$ under the integral sign, in a manner similar to that of the previous section, we obtain the joint density function of $u$ and $v$, given by

$$\begin{aligned} \frac{\partial^2 F_1(s, t)}{\partial s \partial t} &= \frac{s}{t^2}\left[\int_{-\infty}^{\infty} f\left(x_1, x_1 + s, x_1 + \frac{s}{t}\right) dx_1 \right. \\ &\qquad \left. + \int_{-\infty}^{\infty} f\left(x_1, x_1 + \frac{s(1 - t)}{t}, x_1 + \frac{s}{t}\right) dx_1\right] \\ &= f_1(s, t) \quad \text{(say)}. \end{aligned} \tag{6}$$

**4. Applications to normal distributions.** Let $f(x)$ in (4) be the density function of a normal distribution with mean $\theta$ and variance unity, then (6) reduces to

$$f_1(u, w) = \frac{2\sqrt{3}u}{\pi w^2} e^{-u^2 \frac{(1 - w + w^2)}{3w^2}}. \tag{7}$$

Further the marginal density of $u$ and $w$ will be given by

$$p(u) = 6\sqrt{2}\phi\left(\frac{u}{\sqrt{2}}\right)\left[1 - G\left(\frac{\sqrt{3}u}{\sqrt{2}}\right)\right], \tag{8}$$

$$p(w) = \frac{3\sqrt{3}}{\pi} \frac{1}{1 - w + w^2}, \qquad 0 < w < \tfrac{1}{2}, \quad \text{respectively.} \tag{9}$$

The distribution of $w$ has been obtained by J. Lieblein in an unpublished paper.

From (2) we can also obtain the joint density function of $u = x'' - x'$ and

$v = \frac{x' + x''}{2}$. When we integrate this joint density function with respect to $u$, we obtain the density function of $v = \frac{x' + x''}{2}$ as given by

$$
(16) \qquad p(v) = 6\sqrt{2}\phi[\sqrt{2}(v - \theta)]\left[1 + G\left(\frac{\sqrt{2}(v - \theta)}{\sqrt{11}}\right) - 2\int_0^\infty \phi(x)G\left(\frac{3x}{\sqrt{2}} + v - \theta\right) dx\right].
$$

The mean and the variance of the distribution of $v$ are given by $\theta$ and $\frac{1}{2} + \frac{\sqrt{3}}{4\pi}$ respectively.

It may be remarked that if there is a suspicion that one of the extreme observations in a sample of three does not belong to the normal population under consideration, then the median of the sample is a better estimate than the average of the two closest. The efficiency of the latter compared to that of the former is about 70%, for the variance of the median in this case is given by $1 + \frac{\sqrt{3}}{\pi}$ compared to $\frac{1}{2} + \frac{\sqrt{3}}{4\pi}$ of $v$, the average of the two closest. The efficiency is here defined as the ratio of the variances for the two estimates.

## ERRATA

By W. Feller

*Cornell University*

The author regrets the following inconsequential, but very disturbing, slips in his paper "On the Kolmogorov-Smirnov limit theorems for empirical distributions" (*Annals of Math. Stat.*, Vol. 19 (1948), pp. 177–189):

(1) In equation (1.4) on p. 178, the exponent $-\nu^2 z^2$ should be replaced by $-2\nu^2 z^2$. The same copying error occurs in the description of Smirnov's table on p. 279. The proof is correct as it stands.

(2) In the formulation of the *continuity-theorem* on p. 180 it is claimed that $u_k \to f(t)$ whereas in reality the continuity theorem permits only the conclusion that

$$
(*) \qquad \delta \sum_{r=1}^{k} u_r \to \int_0^t f(x)\, dx.
$$

This slip in formulation in no way affects the proofs since only (*) is used. (The assertion that the step functions $\{\xi_k\}$ converge pointwise is not based on a