NOTE ON A CONDITIONAL PROPERTY OF STUDENT'S t[1]

By R. J. Buehler and A. P. Feddersen

Iowa State University

1. Introduction. Fisher (1956) criticized the conditional behavior of Welch's solution of the Behrens-Fisher problem, showing an instance in which the conditional probability of rejecting the null hypothesis, given equality of sample variances, is for all parameters values bounded below by a number ($= 0.108$) greater than the nominal significance level ($= 0.100$). The associated 90% confidence intervals for the difference of population means would therefore cover the true difference less than $(90 - 0.8)\%$ of the time, for any parameter values, in a subset of outcomes in which the ratio of sample variances differs very slightly from unity.

Conditional properties of other procedures were subsequently studied under general formulations by Tukey (1958), Buehler (1959), Wallace (1959), and Stein (1961). Let z denote a sample point. Tukey introduced a general conditioning procedure by defining a "selection" as a function $k(z)$, $0 \leqq k \leqq 1$, which gives the probability that any outcome z is included in the conditioning subset. A "pure selection" is one for which $k(z)$ takes only the values 0 and 1; the set $\{z:k(z) = 1\}$ is then a conditioning set in the ordinary sense. For the usual confidence or fiducial intervals based on Student's t, Stein considered the selection $k(\bar{x}, s) = 1 - (2/\pi) \tan^{-1} (|\bar{x}|/2s)$ (where $n\bar{x} = \sum x_i$, $ns^2 = \sum(x_i - \bar{x})^2$) and showed that the selected intervals cover the true mean with probability greater than $\alpha + \epsilon$ where α is the nominal confidence level and $\epsilon > 0$ is independent of the parameters.

The two most evident respects by which Fisher's and Stein's examples differ are: (i) the direction of the inequality, and (ii) the use by Stein of a general (rather than a "pure") selection. The purpose of the present note is to show that at least in a special case Stein's selection can be replaced by a "pure" one having the same property. Thus Fisher's fiducial intervals based on Student's t could be criticized in very nearly the same way as Fisher himself has critized the Welch solution.

2. Calculations. We consider x normal, $Ex = \mu$, Var $x = \sigma^2$. Given two observatious, a 50% confidence or fiducial interval based on Student's t has the form $x_{\min} \leqq \mu \leqq x_{\max}$. Stein's example suggests conditioning on a subset of the form $|\bar{x}|/s \leqq$ constant, or equivalently $|x_1 - x_2| \geqq c|x_1 + x_2|$ (where $0 < c < 1$). We seek bounds for the conditional probability $P(A|B)$ where the sets A and B are defined by

$$A = \{x_1, x_2 : x_{\min} \leqq \mu \leqq x_{\max}\}, \qquad B = \{x_1, x_2 : |x_1 - x_2| \geqq c|x_1 + x_2|\}.$$

Received July 2, 1962; revised April 8, 1963.

[1] Supported by National Science Foundation Grant G14831.

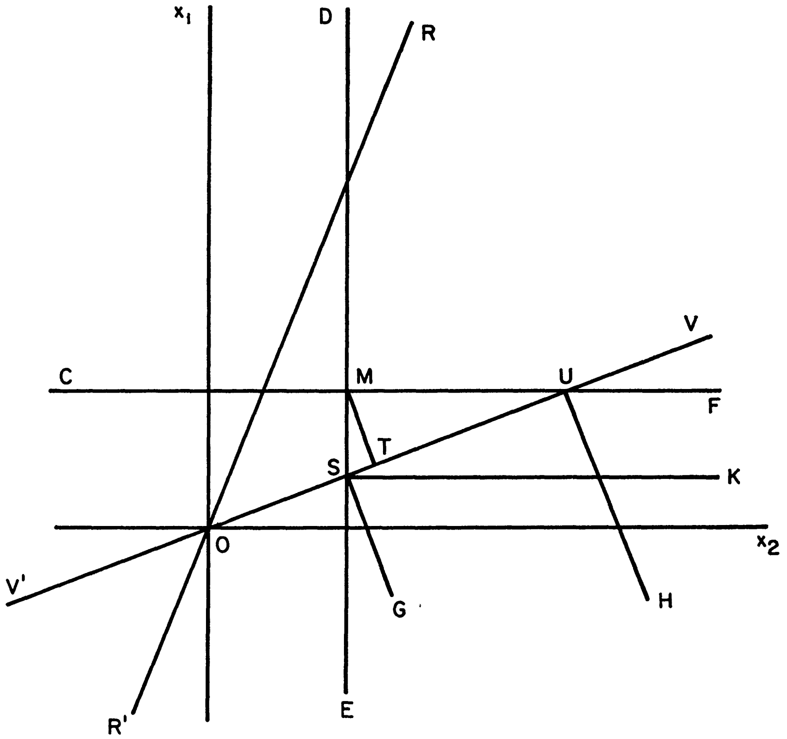

FIG. 1

In Figure 1, O is the origin and M is the mean (μ, μ) of the circular bivariate normal distribution ($\mu > 0$ is assumed without loss of generality). The set A comprises the quarter planes CMD and EMF, while B comprises the wedges $V'OR$ and $R'OV$. The lines OR and OV are easily found to have slopes b and $1/b$ where $b = (1 + c)/(1 - c) > 1$. Points S and U are determined by intersections of the boundaries of A and B as shown, and T is constructed on OV so that MT is perpendicular to OV. Lines SG and UH are constructed parallel to MT, and SK is parallel to the x_1 axis. The coordinates of S, T and U are found to be $(\mu, \mu/b)$, $(\mu b(b + 1)/(b^2 + 1), \mu(b + 1)/(b^2 + 1))$, and $(\mu b, \mu)$ respectively. Let the lengths of the lines MT, TU, ST, and MS be denoted by μa_1, μa_2, μa_3 and μa_4 respectively. Then $a_2 = b(b - 1)/(b^2 + 1)^{\frac{1}{2}}$, $a_3 = (b - 1)/b(b^2 + 1)^{\frac{1}{2}}$, $a_4 = (b - 1)/b$.

The intersection set AB comprises the region $ESUF$ plus its reflection in the diagonal $x_1 = x_2$. Since the region $GSUH$ is contained in $ESUF$, one has

$$P(AB) \geqq 2\{1 - F(\lambda a_1)\}\{F(\lambda a_2) - F(-\lambda a_3)\} \tag{1}$$

where $\lambda = \mu/\sigma$ and F is the standard normal c.d.f. Similarly since ESK is contained in $ESUF$,

$$(2) \qquad P(AB) \geqq 1 - F(\lambda a_4).$$

The quantity $2(1 - F(\lambda a_1))$ is seen to exceed $P(B)$ by twice the probability of the wedge $V'OR'$, and therefore

$$(3) \qquad P(B) \leqq 2\{1 - F(\lambda a_1)\}.$$

Also since the density at any point on the boundary OR is greater than the density at its image on OV' equidistant from O, one sees that if the mean of the distribution were moved from M to O, $P(B)$ would be increased. Thus

$$(4) \qquad P(B) \leqq \tfrac{1}{2} + (2/\pi) \tan^{-1} (1/b).$$

The quotient of (1) and (3) gives

$$(5) \qquad P(A \mid B) \geqq F(\lambda a_2) - F(-\lambda a_3).$$

Now put $b = 5$ (whence incidentally $c = \frac{2}{3}$). Then $a_2 = 3.922$, $a_3 = 0.1569$, $a_4 = 0.8$, and $(2/\pi) \tan^{-1} (1/b) = 0.1257$. Next put $\lambda = 0.57$ to give $\lambda a_2 = 2.236$, $\lambda a_3 = 0.08943$, $\lambda a_4 = 0.4560$, $F(\lambda a_2) = 0.9873$, $F(\lambda a_3) = 0.5536$, $F(\lambda a_4) = 0.6758$. From (5), $P(A \mid B) \geqq 0.5229$. From the quotient of (2) and (4), $P(A \mid B) \geqq 0.3242/0.6257 = 0.5181$. But the right hand side of (5) is an increasing function of λ so that our first bound gives $P(A \mid B) \geqq 0.5229$ for $\lambda \geqq 0.57$; and the quotient of (2) and (4) is a decreasing function of λ, whence $P(A \mid B) \geqq 0.5181$ for $\lambda \leqq 0.57$. This gives the desired result $P(A \mid B) \geqq 0.5 + 0.0181$ for all λ, and hence for all μ, σ.

Clearly the bound could be increased by sharpening some of the inequalities. It is not evident how much the "bias" of 0.0181 could be increased in this way. We are also unable to say what results would hold of other confidence levels α, other values of b, or other degrees of freedom. A few exploratory calculations have indicated, however, that the present method will not give any surprising increase in the bias when α and b are changed.

REFERENCES

BUEHLER, ROBERT J. (1959). Some validity criteria for statistical inferences. *Ann. Math. Statist.* **30** 845–863.

FISHER, SIR RONALD (1956). On a test of significance in Pearson's Biometrika Tables (No. 11). *J. Roy. Statist. Soc. Ser. B* **18** 56–60.

STEIN, CHARLES (1961). Estimation of many parameters. Inst. Math. Statist. Wald Lectures. Unpublished.

TUKEY, JOHN W. (1958). Fiducial inference. Inst. Math. Statist. Wald Lectures. Unpublished.

WALLACE, DAVID L. (1959). Conditional confidence level properties. *Ann. Math. Statist.* **30** 864–876.