NOTES

This section is devoted to brief research and expository articles and other short items.

A NOTE ON THE POWER OF A NON-PARAMETRIC TEST

By F. J. Massey, Jr.

University of Oregon

1. Introduction. Let $x_1 < x_2 < \cdots < x_n$ be the ordered results of n independent observations of a random variable X which has a continuous cumulative distribution function $F(x)$. The following test for the hypothesis that $F(x)$ has some specified form, say $F_0(x)$, has been suggested by Wolfowitz [1].

Form the cumulative distribution of the sample and obtain the maximum deviation of this from $F_0(x)$. Thus if

$$\begin{aligned} S_n(x) &= 0 && \text{when } x < x_1, \\ &= \frac{k}{n} && \text{when } x_k \le x < x_{k+1}, \\ &= 1 && \text{when } x_n < x, \end{aligned}$$

the test statistic used would be

$$d = \max_x | F_0(x) - S_n(x) | \sqrt{n},$$

and the hypothesis would be rejected if d is large, say larger than d_α which is so chosen that the probability of a type I error is α. The limiting distribution (as $n \to \infty$) of d has been tabled [2], and a short table of the distribution of d for various small values of n ($n \le 80$) has been given [3].

The purpose of this note is as follows: 1. A lower bound for the power of the test is given. 2. This test is shown to be consistent against any continuous alternative $F(x) = F_1(x)$, where $F_1(x) \not\equiv F_0(x)$. 3. The test is shown to be biased for finite n. 4. An indication of similar results for a two sample test.

2. Lower bound for the power function. Let $\Delta = \max_x | F_0(x) - F_1(x) |$ and let x_0 be a value of x such that $\Delta = | F_0(x_0) - F_1(x_0) |$. The probability that $d > d_\alpha$ is certainly not less than $\Pr\{\sqrt{n} | F_0(x_0) - S_n(x_0) | > d_\alpha\}$. This is the same as

$$1 - \Pr\left\{F_0(x_0) - \frac{d_\alpha}{\sqrt{n}} < S_n(x_0) < F_0(x_0) + \frac{d_\alpha}{\sqrt{n}}\right\},$$

which, since $S_n(x_0)$ is the proportion of observations falling less or equal to x_0, is given by the binomial probability law.

If $F(x) = F_1(x)$ the probability of an observation being less than x_0 is $F_1(x_0)$. Since $F_0(x_0) = F_1(x_0) \pm \Delta$ the above probability can be written as follows:

$$1 - \Pr\{F_1(x_0) \pm \Delta - d_\alpha/\sqrt{n} < S_n(x_0) < F_1(x_0) \pm \Delta + d_\alpha/\sqrt{n}\}$$
$$= 1 - \Pr\{\pm\Delta - d_\alpha/\sqrt{n} < S_n(x_0) - F_1(x_0) < \pm \Delta + d_\alpha/\sqrt{n}\}$$
$$= 1 - \Pr\{(-d_\alpha \pm \Delta\sqrt{n})/\sqrt{F_1(x_0)(1 - F_1(x_0))} < (S_n(x_0) - F_1(x_0))\sqrt{n}/$$
$$\sqrt{F_1(x_0)(1 - F_1(x_0))} < (d_\alpha \pm \Delta\sqrt{n})/\sqrt{F_1(x_0)(1 - F_1(x_0))}\}.$$

Δ is fixed. It has been found [3] by observation for samples of size ≤ 80 that d_α actually decreases in size as n increases. For sufficiently large n both

$$-d_\alpha \pm \Delta\sqrt{n} \quad \text{and} \quad d_\alpha \pm \Delta\sqrt{n}$$

have the same sign and the law of large numbers indicates that the above probability approaches zero and the expression approaches unity.

The last expression above can also be used as a lower bound of the power of the test for finite n.

For large values of n this probability is given approximately by the normal distribution. Thus we can write for large n;

$$\text{power} > 1 - \int_{\lambda_1}^{\lambda_2} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt,$$

where

$$\lambda_1 = (-d_\alpha \pm \Delta\sqrt{n})/\sqrt{F_1(x_0)(1 - F_1(x_0))}$$

and

$$\lambda_2 = (d_\alpha \pm \Delta\sqrt{n})/\sqrt{F_1(x_0)(1 - F_1(x_0))}.$$

If n is so large that λ_1 and λ_2 are of the same sign and sufficiently different from zero we can replace $F_1(x_0)$ by $\frac{1}{2}$ and not decrease the value of the integral. In this case we might use as a working formula

$$\lambda_1 = 2(-d_\alpha \pm \Delta\sqrt{n}),$$
$$\lambda_2 = 2(d_\alpha \pm \Delta\sqrt{n}).$$

Since

$$1 - \frac{1}{\sqrt{2\pi}} \int_{\lambda_1}^{\lambda_2} e^{-t^2/2}\, dt$$

approaches one as n tends to infinity, the power, which is larger, must also approach one, and thus the test is consistent.

To demonstrate the biasedness of the test for fixed n consider the following picture.
The $F_0(x)$ is shown as a heavy line and an alternative $F_1(x)$ as a dash-dot line. $F_1(x)$ coincides with $F_0(x)$ except between the point $x = a$ and $x = b$. If $S_n(x)$ falls outside of the indicated band at any point we agree to reject the hypothesis $F(x) = F_0(x)$. If $F(x) = F_1(x)$ the $S_n(x)$ has no chance of being outside the band between $x = a$ and $x = c$, less chance between $x = c$ and $x = b$ than if

$F(x) = F_0(x)$, and the same chance for x larger than b. This indicates that the probability of rejecting $F(x) = F_0(x)$, if actually $F(x) = F_1(x)$, is greater than the probability of rejecting $F(x) = F_0(x)$ if this is actually true. Thus the test is biased.

3. Two sample test. Let $S_n(x)$ and $S'_m(x)$ be the cumulative distributions observed for samples of sizes n and m from two populations having continuous cumulative distribution functions $F(x)$ and $F'(x)$ respectively. Under the assumption that $F(x) = F'(x)$ the limiting distribution (as n and m tend to infinity) of $d' = (n^{-1} + m^{-1})^{-\frac{1}{2}} \max_x | S_n(x) - S'_m(x) |$ has been found and tabled [4], but the distribution of this statistic for small n and m is not known.

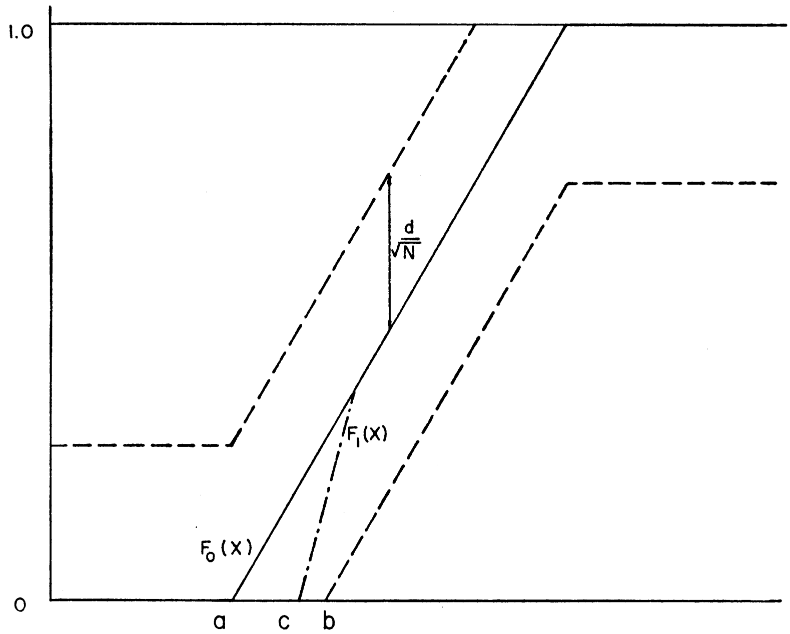

FIG. 1.

Suppose we wish to test the hypothesis that $F(x) = F'(x)$ at level of significance α and agree to reject this if d' is larger than d'_α, where d'_α is the value which would be exceeded a proportion α of the time if the hypothesis is true. The values of d'_α are not known for small samples but are for the limiting case [4].

The same argument as in Section 2 gives a limiting lower bound to the power of the test in terms of

$$\Delta = | F(x_0) - F'(x_0) |,$$

where x_0 is the value of x which maximizes $| F(x) - F'(x) |$, to be

$$1 - \int_{\lambda'_1}^{\lambda'_2} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt,$$

where

$$\lambda_1' = \left(-d_\alpha' \sqrt{\frac{1}{n} + \frac{1}{m}} \pm \Delta\right) \Big/ \sqrt{\frac{F(x_0)[1 - F(x_0)]}{n} + \frac{F'(x_0)[1 - F'(x_0)]}{m}}$$

and

$$\lambda_2' = \left(d_\alpha' \sqrt{\frac{1}{n} + \frac{1}{m}} \pm \Delta\right) \Big/ \sqrt{\frac{F(x_0)[1 - F(x_0)]}{n} + \frac{F'(x_0)[1 - F'(x_0)]}{m}}.$$

Since this lower bound approaches one as n and m approach infinity the power also approaches one and the test is consistent.

REFERENCES

[1] J. Wolfwitz, "Non-parametric statistical inference," *Proceedings of the Symposium on Mathematical Statistics and Probability*, University of California Press, 1949, pp. 93–113.
[2] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions," *Annals of Math. Stat.*, Vol. 19 (1948), pp. 279–281.
[3] F. Massey, "A note on the estimation of a distribution function by confidence limits," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 116–120.
[4] N. Smirnov, "On the estimation of the discrepancy between empirical curves of distribution for two independent samples," *Bull. Math. Univ. Moscou, Série Int.*, Vol. 2, fasc. 2 (1939).

ON OPTIMUM SELECTIONS FROM MULTINORMAL POPULATIONS[1]

By Z. W. Birnbaum and D. G. Chapman[2]
University of Washington

1. Introduction. Let $Y_1, Y_2, \cdots, Y_n$ be scores in n admission tests such as those used in educational institutions, personnel selection, or testing of materials, and let these scores be used as a basis for selecting a sub-population Π^* from an initial population Π. This selection is usually performed in such a manner that an achievement or performance score X has a distribution in Π^*, which shows some required improvement over the distribution of X in Π; such an improvement may for example consist in changing the expectation $E(X)$ of X in Π to a pre-assigned value $E^*(X)$ in Π^*. Among all selection procedures based on $Y_1, \cdots, Y_n$ and achieving the required improvement of the distribution of X, it appears desirable to find those which retain as large a portion of Π as possible. It will be shown that under certain assumptions the linear truncations studied in an earlier paper [1] are such optimal selections.

2. Selection, truncation, linear truncation. Let the frequency of individuals with the scores $(X, Y_1, \cdots, Y_n)$ be $F(X, Y_1, \cdots, Y_n)$ in Π and

[1] Presented at the New York meeting of the Institute of Mathematical Statistics on December 27, 1949.
[2] Research done under the sponsorship of the Office of Naval Research.