$$\frac{1}{B\left(\frac{1}{2}, \frac{\nu}{2}\right)} \int_{x_\alpha}^{1} p^{-\frac{1}{2}}(1 - p)^{(\nu-2)/2}\, dp = \alpha.$$

Using the approximation of ordinate over abscissa for the cumulative normal for extreme abscissa we find that z is the abscissa of a cumulative normal which is approximately equal to the power of the t-test for alternative δ. In a similar manner the normal approximation to the binomial yields $z = \delta\sqrt{r + 1}$ for the sign test. A fixed value of N and α determines r, α, x_α and we may solve for ν.

REFERENCES

[1] National Bureau of Standards, *Tables of the Binomial Probability Distribution*, U. S. Government Printing Office, Washington, D. C.

[2] K. Pearson, *Tables of the Incomplete Beta Function*, Cambridge University Press, 1934.

[3] Burton H. Camp, "Approximation to the point binomial," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 130–131.

[4] W. G. Cochran, "The efficiencies of the binomial test of significance of a mean and of a correlation coefficient," *J. Roy. Stat. Soc.*, Vol. 100 (1937), pp. 69–73.

[5] W. J. Dixon and A. M. Mood, "The statistical sign test," *J. Amer. Stat. Assn.*, Vol. 41 (1946), pp. 557–566.

[6] Murray F. Freeman and John W. Tukey, "Transformations related to the angular and the square root," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 607–611.

[7] T. A. Jeeves and Robert Richards, "A note on the power of the sign test," *Ann. Math. Stat.*, Vol. 21 (1950), p. 618.

[8] W. L. Nicholson, "A computing formula for the power of the analysis of variance test," submitted to *Ann. Math. Stat.*

[9] John E. Walsh, "On the power function of the sign test for slippage of means," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 358–362.

[10] John E. Walsh, "On the asymptotic power efficiency of the sign test for slippage of means," Douglas Aircraft Co., Inc. Report.

[11] John E. Walsh, "Some comments on the efficiency of statistical tests," *Human Biology*, Vol. 21 (1949), pp. 205–217.

[12] John E. Walsh, "Some significance tests for the median which are valid under very general conditions," *Ann. Math. Stat.*, Vol. 20, (1949), pp. 64–81.

[13] John E. Walsh, "Some bounded significance level properties of the equal-tail sign test," *Ann. Math. Stat.*, Vol 22 (1951), pp. 408–417.

THE ADMISSIBILITY OF CERTAIN INVARIANT STATISTICAL TESTS INVOLVING A TRANSLATION PARAMETER

By E. L. Lehmann[1] and C. M. Stein

University of California, Berkeley, and University of Chicago

1. Introduction. The notion of invariance (or symmetry) has such strong intuitive appeal that many current statistical procedures have the invariance property and are in fact the best invariant procedures although they were pro-

Received 2/10/53.

[1] This work is supported in part by the Office of Naval Research.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

posed long before a general discussion of invariance was available. Hotelling [1], [2] and Pitman [3], [4] emphasized the invariant nature of certain tests and estimates. A general definition of the notion for the problem of testing hypotheses was given by Hunt and Stein who showed that in this case under severe restrictions on the group of transformations an optimum invariant test is most stringent or more generally minimax with respect to an invariant loss function (see [5]). This result has been extended to more general decision problems and more general groups by Peisakoff [6]. However, these results do not prove admissibility of the procedures in question unless the group of transformations is compact.

The problem of admissibility in the case of point estimation of a location parameter was treated in the normal case by Blyth [7] and by Hodges and Lehmann [8] and for a general class of location parameter-problems by Blackwell [9]. In the latter paper the surprising fact was brought to light that even in the location parameter problem the best invariant estimate may, under certain circumstances, be inadmissible.

In the present note we prove under conditions which are presumably unnecessarily restrictive the admissibility of the most powerful invariant test for testing one location parameter family against another. As an example, consider the problem in which $Z_1, \cdots, Z_n$ are normally distributed with unknown mean ζ and variance σ^2. If we wish to test $H: \zeta \leq 0$ against the alternatives $K: \zeta > 0$, it was already pointed out in ([5], p. 15) that Student's t-test is admissible for this problem. This result is quite elementary and rests on the fact that unbiasedness in this case implies that the probability of rejection equals the level of significance for all points (ζ, σ) with $\zeta = 0$. However, this argument breaks down if we introduce an indifference zone and restrict our class of alternatives to $K': \zeta/\sigma \geqq \delta$ where δ is some specified positive number.

Consider now the general problem in which one observes a random point (X, Y) where X ranges over an arbitrary set, Y over the real line. There are two hypotheses H_i according to which the distribution of $(X, Y - \eta)$ is $F_i (= 1, 2)$ where η is unspecified. The problem discussed above is an example of this, if we take H_i to be $\zeta/\sigma = \delta_i$, $X = \sum Z_i/\sqrt{\sum Z_j^2}$, $Y = \log \sum Z_j^2$ and $\eta = \log \sigma$. As another example let $(Z_1 - \eta, Z_2 - \eta, \cdots, Z_n - \eta)$ have distribution F_i under H_i. Then we can take for X the set of differences $X = (Z_1 - Z_n, \cdots, Z_{n-1} - Z_n)$ and for Y the mean $\bar{Z}$ or the observation Z_n, or any of a number of other statistics.

2. The principal theorem. Let $\mathfrak{X}$ be a set (which for all practical purposes may be taken to be a Euclidean space), $\mathcal{A}$ a σ-algebra of subsets of $\mathfrak{X}$ (say, the ordinary Borel sets if $\mathfrak{X}$ is Euclidean), $\mathcal{R}$ the real line, $\mathcal{B}$ the set of all ordinary Borel subsets of $\mathcal{R}$, λ_1, λ_2 probability measures on $\mathcal{A}$ and for each x, let F_{1x}, F_{2x} be probability measures on $\mathcal{B}$ such that for each $B \varepsilon \mathcal{B}$, real k, and $i = 1, 2$ $\{x \mid F_{ix}(B) \leqq k\} \varepsilon \mathcal{A}$. We suppose that the distribution of the random point (X, Y) ranging over $\mathfrak{X} \times \mathcal{Y}$ is, for some real η, with $i = 1$ or 2

$$(1) \qquad P_{i\eta}((X, Y) \varepsilon C) = \int_C d\lambda_i(x) \int dF_{ix}(y - \eta).$$

A test for the hypothesis H_1 that it is $P_{1\eta}$ (with η unspecified) is a function φ on $\mathfrak{X} \times \mathfrak{R}$ to $[0, 1]$, $\mathfrak{A}\mathfrak{B}$ measurable. The test φ is said to be better than φ_0 if for all η

$$E_{1\eta}\varphi(X, Y) \leqq E_{1\eta}\varphi_0(X, Y)$$
$$E_{2\eta}\varphi(X, Y) \geqq E_{2\eta}\varphi_0(X, Y), \tag{2}$$

strictly better if (2) holds with strict inequality for some η. φ_0 is admissible if there exists no φ strictly better than φ_0.

THEOREM 1. *If* $E_{10} \mid Y \mid, E_{20} \mid Y \mid < \infty, 0 < c < 1,$

$$\lambda_1 \left\{ x \,\middle|\, \frac{d\lambda_2}{d(\lambda_1 + \lambda_2)}(x) = c \right\} = 0,$$

φ_0 is the test defined by

$$\varphi_0(x, y) = \begin{cases} 1 \text{ if } \dfrac{d\lambda_2}{d(\lambda_1 + \lambda_2)}(x) \geqq c \\ 0 \text{ if } \dfrac{d\lambda_2}{d(\lambda_1 + \lambda_2)}(x) < c, \end{cases} \tag{3}$$

and φ is better than φ_0, then $\varphi - \varphi_0 = 0$ a.e. $(\lambda_1 + \lambda_2)\mu$ where μ is ordinary Lebesgue measure on the real line.

COROLLARY. *If in addition all* F_{ix} *are absolutely continuous with respect to* μ, *then* φ_0 *is admissible.*

The corollary is an immediate consequence of Theorem 1.

PROOF OF THEOREM 1. Putting $\lambda = \lambda_1 + \lambda_2$, $f(x) = d\lambda_2/d(\lambda_1 + \lambda_2)(x)$ we can rewrite the condition (2) that φ be better than φ_0

$$\int_{f(x)<c} (1 - f(x))\, d\lambda(x) \int (\varphi - \varphi_0)(x, y)\, dF_{1x}(y - \eta) - \int_{f(x)\geqq c} (1 - f(x))\, d\lambda(x) \int (\varphi_0 - \varphi)(x, y)\, dF_{1x}(y - \eta) \leqq 0 \tag{4}$$

$$- \int_{f(x)<c} f(x)\, d\lambda(x) \int (\varphi - \varphi_0)(x, y)\, dF_{2x}(y - \eta) + \int_{f(x)\geqq c} f(x)\, d\lambda(x) \int (\varphi_0 - \varphi)(x, y)\, dF_{2x}(y - \eta) \leqq 0. \tag{5}$$

Multiplying (4) by c and (5) by $1 - c$ and adding we obtain

$$\begin{aligned} &\int_{f(x)<c} c(1 - f(x))\, d\lambda(x) \int (\varphi - \varphi_0)(x, y)\, dF_{1x}(y - \eta) \\ &\quad + \int_{f(x)\geqq c} (1 - c)f(x)\, d\lambda(x) \int (\varphi_0 - \varphi)(x, y)\, dF_{2x}(y - \eta) \\ &\leqq \int_{f(x)<c} (1 - c)f(x)\, d\lambda(x) \int (\varphi - \varphi_0)(x, y)\, dF_{2x}(y - \eta) \\ &\quad + \int_{f(x)\geqq c} c(1 - f(x))\, d\lambda(x) \int (\varphi_0 - \varphi)(x, y)\, dF_{1x}(y - \eta). \end{aligned} \tag{6}$$

In order to derive the conclusion of Theorem 1 from (6) we shall need the

LEMMA. *If $\mathfrak{X}$, $\mathfrak{A}$, $\mathfrak{R}$, $\mathfrak{B}$ are as before, ρ a probability measure on $\mathfrak{A}$, h_1, h_2 $\mathfrak{A}$-measurable functions on $\mathfrak{X}$ to* [0, 1] *with $h_1 - h_2 > 0$* a.e. (ρ), *ψ an $\mathfrak{A}\mathfrak{B}$-measurable function on $\mathfrak{X} \times \mathfrak{R}$ to* [0, 1], *and for each x, H_{1x}, H_{2x} probability measures on $\mathfrak{B}$ such that*

$$\text{(7)} \qquad \text{for each } B \in \mathfrak{B} \text{ real } k \text{ and } i = 1, 2, \{x \mid H_{ix}(B) \leqq k\} \in \mathfrak{A}$$

$$\text{(8)} \qquad \int h_i(x)\, d\rho(x) \int |y|\, dH_{ix}(y) < \infty.$$

$$\text{(9)} \qquad \begin{aligned} &\int h_1(x)\, d\rho(x) \int \psi(x, y)\, dH_{1x}(y - \eta) \\ &\qquad \leqq \int h_2(x)\, d\rho(x) \int \psi(x, y)\, dH_{2x}(y - \eta) \qquad \text{for all real } \eta \end{aligned}$$

then $\psi = 0$ a.e. ($\rho\mu$).

PROOF OF LEMMA. We can rewrite (9)

$$\text{(10)} \quad \int h_1(x)\, d\rho(x) \int \psi(x, y + \eta)\, dH_{1x}(y) \leqq \int h_2(x)\, d\rho(x) \int \psi(x, y + \eta)\, dH_{2x}(y).$$

Now

$$\text{(11)} \qquad \begin{aligned} &\int_{-n}^{n} d\eta \int \psi(x, y + \eta)\, dH_{2x}(y) - \int_{-n}^{n} \psi(x, \eta)\, d\eta \\ &\quad = \int dH_{2x}(y) \left[\int_{-n+y}^{n+y} \psi(x, \eta)\, d\eta - \int_{-n}^{n} \psi(x, \eta)\, d\eta \right] \\ &\quad \leqq \int_{y \geqq 0} dH_{2x}(y) \int_{n}^{n+y} \psi(x, \eta)\, d\eta + \int_{y \leqq 0} dH_{2x}(y) \int_{-n+y}^{-n} \psi(x, \eta)\, d\eta \\ &\quad \leqq \int |\, y\, |\, dH_{2x}(y) \end{aligned}$$

and

$$\text{(12)} \qquad \begin{aligned} &\int_{-n}^{n} d\eta \int \psi(x, y + \eta)\, dH_{1x}(y) - \int_{-n}^{n} \psi(x, \eta)\, d\eta \\ &\quad = \int dH_{1x}(y) \left[\int_{-n+y}^{n+y} \psi(x, \eta)\, d\eta - \int_{-n}^{n} \psi(x, \eta)\, d\eta \right] \\ &\quad \geqq -\int_{y \geqq 0} dH_{1x}(y) \int_{-n}^{-n+y} \psi(x, \eta)\, d\eta - \int_{y \leqq 0} dH_{1x}(y) \int_{n+y}^{n} \psi(x, \eta)\, d\eta \\ &\quad \geqq -\int |\, y\, |\, dH_{1x}(y). \end{aligned}$$

Integrating (10) with respect to η from $-n$ to n and using the final forms of (11), (12) we obtain

$$\begin{aligned}\int [h_1(x) - h_2(x)]\, d\rho(x) \int_{-n}^{n} \psi(x, \eta)\, d\eta \\ \leqq \int h_1(x)\, d\rho(x) \int |y|\, dH_{1x}(y) + \int h_2(x)\, d\rho(x) \int |y|\, dH_{2x}(y).\end{aligned} \tag{13}$$

Consequently,

$$\int [h_1(x) - h_2(x)]\, d\rho(x) \int_{-\infty}^{\infty} \psi(x, \eta)\, d\eta < \infty \tag{14}$$

and for every $\delta > 0$ there exists n such that

$$\int [h_1(x) - h_2(x)]\, d\rho(x) \int_{|\eta| \geqq n} \psi(x, \eta)\, d\eta \leqq \delta. \tag{15}$$

If instead of using the final forms of (11) and (12) for all x, we use them only in the range $h_1(x) - h_2(x) < \epsilon$ and use the next to final forms when $h_1(x) - h_2(x) \geqq \epsilon$ we obtain instead of (13).

$$\begin{aligned}&\int [h_1(x) - h_2(x)]\, d\rho(x) \int_{-n}^{n} \psi(x, \eta)\, d\eta \\ &\leqq \int_{h_1(x)-h_2(x)<\epsilon} d\rho(x) \left[h_1(x) \int |y|\, dH_{1x}(y) + h_2(x) \int |y|\, dH_{2x}(y) \right] \\ &+ \int_{h_1(x)-h_2(x)\geqq\epsilon} d\rho(x) \left[h_1(x) \left\{ \int_{y\geqq 0} dH_{1x}(y) \int_{-n}^{-n+y} \psi(x, \eta)\, d\eta \right.\right. \\ &\qquad \left. + \int_{y\leqq 0} dH_{1x}(y) \int_{n+y}^{n} \psi(x, \eta)\, d\eta \right\} \\ &+ h_2(x) \left.\left\{ \int_{y\geqq 0} dH_{2x}(y) \int_{n}^{n+y} \psi(x, \eta)\, d\eta + \int_{y\leqq 0} dH_{2x}(y) \int_{-n+y}^{-n} \psi(x, \eta)\, d\eta \right\}\right].\end{aligned} \tag{16}$$

The first term on the right-hand side can be made arbitrarily small by taking ϵ sufficiently small since $h_1(x) - h_2(x) > 0$ a.e. (ρ), $0 \leqq h_i(x) \leqq 1$ (using (8)). For, given $\epsilon > 0$, the second half of the last term can be made arbitrarily small by choosing $n \geqq n(\epsilon)$ sufficiently large since by (15)

$$\int_{h_1(x)-h_2(x)\geqq\epsilon} d\rho(x) \int_{|\eta|\geqq n} \psi(x, \eta)\, d\eta \leqq \delta/\epsilon.$$

Also

$$\begin{aligned}\int_{h_1(x)-h_2(x)\geqq\epsilon} d\rho(x) h_1(x) \int_{n/2\leqq y} dH_{1x}(y) \int_{-n}^{-n+y} \psi(x, \eta)\, d\eta \\ \leqq \int_{h_1(x)-h_2(x)\geqq\epsilon} d\rho(x)\, h_1(x) \int_{n/2\leqq y} y\, dH_{1x}(y).\end{aligned} \tag{17}$$

Again, for fixed ϵ this can be made arbitrarily small by choosing $n \geqq n(\epsilon)$ sufficiently large. Finally

$$\int_{h_1(x)-h_2(x)\geqq\epsilon} d\rho(x)\, h_1(x) \int_0^{n/2} dH_{1x}(y) \int_{-n}^{-n+y} \psi(x, \eta)\, d\eta$$
$$\leqq \int_{h_1(x)-h_2(x)\geqq\epsilon} d\rho(x)\, h_1(x) \int_{-\infty}^{-n/2} \psi(x, \eta)\, d\eta,$$

which is disposed of in the same way as the second half of the last term. The remaining integral with $y \leqq 0$ is analogous. Then, since the right-hand side of (16) is arbitrarily small for sufficiently large n, $\psi = 0$ a.e. $(\lambda_1 + \lambda_2)\mu$. This completes the proof of the Lemma.

To apply the Lemma to (6) we make the following identifications:

(i) If $f(x) < c$,

$$h_1(x) = c(1 - f(x)) \qquad h_2(x) = (1 - c)f(x)$$
$$H_{1x} = F_{1x} \qquad H_{2x} = F_{2x}.$$
$$\psi = \varphi - \varphi_0$$

(ii) If $f(x) \geqq c$

$$h_1(x) = (1 - c)f(x) \qquad h_2(x) = c(1 - f(x))$$
$$H_{1x} = F_{2x} \qquad H_{2x} = F_{1x}.$$
$$\psi = \varphi_0 - \varphi$$

In any case $\rho = \lambda/2$. The reader will readily verify that (7), (8), (9) are satisfied so that the theorem follows.

A moment's reflection shows that the origin of $\mathcal{R}$ for given x is arbitrary so that the hypotheses $E_{i0} \mid Y \mid < \infty$ could be replaced by: There exists an $\mathcal{A}$-measurable real-valued function τ on $\mathfrak{X}$ such that $E_{i0} \mid Y - \tau(X) \mid < \infty$.

It is seen that the admissibility of the noncentral t-test for testing $\zeta/\sigma = \delta_0$ against $\zeta/\sigma = \delta$, (central in case $\delta_0 = 0$) follows immediately from the theorem since

$$E \mid \log \sum Z_1^2 \mid < \infty$$

and $P(\sum Z_i = c\sqrt{\sum Z_i^2}) = 0$.

Another example is that of testing for the same random variables $\sigma = \sigma_0$ against $\sigma = \sigma_1$. Here we may take $X = \sum(Z_i - \bar{Z})^2$ and $Y = \sum Z_i$. Actually in this case the result can be proved quite easily by other means. Instead of taking for ζ the usual least favorable sequence of a priori distributions which in the limit is invariant, we may, if $\sigma_0 < \sigma_1$ take in H the a priori distribution $P(\zeta = a) = 1$ where a is any constant, and in K a normal distribution with mean a and variance $n(1/\sigma_1^2 - 1/\sigma_0^2)$. The Bayes solution is seen to be the F-test which is therefore admissible. (For details see [10]).

We can also consider the general linear hypothesis with no unknown means as nuisance parameters. For brevity we use the terminology of [5]. In the canonical form we have $U_1 \cdots U_m$, $V_1 \cdots V_n$ independently normally distributed with $EU_i = \psi_i$, $EV_j = 0$, $E(U_i - \psi_i)^2 = EV_j^2 = \sigma^2$ where σ^2, ψ_i are unknown and we want to test the hypothesis that all $\psi_i = 0$ say, against $\sum \psi_i^2 \geqq \gamma\sigma^2$. A sufficient statistic is $(U_1 \cdots U_m, \sum V_j^2)$. The problem is invariant under rotation of the vector $U_1, \cdots, U_m$ and multiplication of all U_i, V_j by the same constant c. Since the rotation group G_0 possesses a finite invariant measure, any test invariant under G_0 and admissible among all tests invariant under G_0 is admissible. Thus, in proving the usual F-test admissible we may restrict our attention to tests depending only on $(\sum U_i^2, \sum V_j^2)$. Under multiplication by c this goes into $(c^2 \sum U_i^2, c^2 \sum V_j^2)$. Taking $X = \sum U_i^2 / \sum V_j^2$ and $Y = \log \sum V_j^2$, applying Theorem 1 and the optimum property of the F-test among all those based only on $\sum U_i^2 / \sum V_j^2$, we obtain the admissibility of the usual test: Reject H_0 if $\sum U_i^2 \geqq k \sum V_j^2$. The same argument applies to the problem of testing $H_0 : \sum \psi_i^2 \leqq \gamma_1 \sigma^2$ against $H_1 : \sum \psi_i^2 \geqq \gamma_2 \sigma^2$ with $\gamma_2 > \gamma_1$.

REFERENCES

[1] H. HOTELLING, "Analysis of a complex of statistical variables into principal components," *J. Educational Psychology*, Vol. 24 (1933), pp. 417–441 and pp. 498–520.

[2] H. HOTELLING, "Relations between two sets of variates," *Biometrika*, Vol. 28 (1936), pp. 321–377.

[3] E. J. G. PITMAN, "The estimation of the location and scale parameters of a continuous population of any given form," *Biometrika*, Vol. 30 (1938), pp. 391–421.

[4] E. J. G. PITMAN, "Tests of hypotheses concerning location and scale parameters, *Biometrika*, Vol. 31 (1939), pp. 200–215.

[5] E. L. LEHMANN, "Some principles of the theory of testing hypotheses," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 1–26.

[6] M. PEISAKOFF, Unpublished Ph.D. thesis, Princeton University, 1950.

[7] C. BLYTH, "On minimax statistical decision procedures and their admissibility," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 22–42.

[8] J. L. HODGES, JR. AND E. L. LEHMANN, "Some applications of the Cramér-Rao inequality," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1951.

[9] D. BLACKWELL, "On the translation parameter problem for discrete variables," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 393–399.

[10] E. L. LEHMANN AND C. STEIN, "Most powerful tests of composite hypotheses, I. Normal distributions," *Ann. Math. Stat.*, Vol. 19 (1948), pp. 495–516.