STATISTICAL REPRODUCTION OF ORDERINGS AND TRANSLATION SUBFAMILIES

BY H. S. KONIJN

Tel Aviv University

0. Summary. For a given family of distributions let φ be a parameter which takes on different values on different subfamilies but is constant on each. The values of the parameter may be taken to define an ordering of the subfamilies. Various concepts of statistical reproduction of such an ordering are introduced. They generalize concepts such as unbiased estimators and tests.

In the latter part of the paper it is assumed that in each subfamily all distributions may be obtained from a single one by translation of the random variable. As an application we consider testing for decrease in variability when the individuals tested may have different means.

1. Introduction. The purpose of this paper is to introduce certain generalizations of concepts such as unbiased and consistent estimators and tests.

We shall consider an observable random variable X—usually it will have values in more than one dimension—with distribution function F which is known to belong to a certain family $\mathfrak{F}$ of distribution functions; and a specified parameter $\varphi(F)$ of F, more precisely a functional φ over $\mathfrak{F}$, which, for the sake of simplicity of presentation, we shall restrict to be unidimensional. On the basis of X it is required to make some "evaluation" $h(X)$ of $\varphi(F)$.

If φ can take only two values (which we may take to be 0 and 1 without loss of generality), the problem is one of testing. (The case in which φ can take on more than two values may also be of interest in testing problems; thus in testing the hypothesis $\varphi(F) = 0$ against $\varphi(F) > 0$, mean reproduction (defined below) refers to the behavior of the power function of the test.) Let $\mathfrak{F}_0$ be the subfamily of $\mathfrak{F}$ for which $\varphi(F) = 0$ and $\mathfrak{F}_1$ the subfamily for which $\varphi(F) = 1$, and let H_i be the hypothesis that F belongs to $\mathfrak{F}_i$ $(i = 0, 1)$. In the non-randomized case, to which we shall confine ourselves, one generally requires h to have the same two possible values as φ with certain maximum probabilities α and β of misevaluation (both strictly between 0 and 1):

$$P_F\{h(X) = 1\} \leqq \alpha \text{ when } \varphi(F) = 0, \quad P_F\{h(X) = 0\} \leqq \beta \text{ when } \varphi(F) = 1.$$

The test of H_0 consists in rejecting H_0 if $h(X) = 1$.

More generally, any φ partitions a family $\mathfrak{F}$ into a collection $\mathcal{P}$ of (mutually exclusive) subfamilies on each of which φ is constant and such that φ has a different value on each element (subfamily) of the collection. Moreover, this partition $\mathcal{P}$ of $\mathfrak{F}$ is linearly ordered by φ, and $\mathfrak{F}$ is partly ordered by φ. The problem may be to decide, on the basis of X, to which subfamily of the partition F belongs. It

Received 19 May 1965; revised 14 September 1965.

is not always practicable to require h to have the same range of values as φ, since such a restriction may lead to serious mathematical difficulties. Also, such a requirement may be inconsistent with other properties we may desire h to have.

In the testing problem we say that the test based on h is unbiased if $\alpha \leqq 1 - \beta$, and strictly unbiased if $\alpha < 1 - \beta$. In the more general problem we shall say that an "evaluation" h of φ is unbiased (or, more explicitly, mean unbiased) if $\mathcal{E}_F h(X) = \varphi(F)$ for all F in $\mathfrak{F}$, median unbiased if $P_F\{h(X) \geqq \varphi(F)\} \geqq \frac{1}{2}$ and $P_F\{h(X) \leqq \varphi(F)\} \geqq \frac{1}{2}$, i.e. if, under F, $\varphi(F)$ is a median of $h(X)$. In many problems such an estimator either does not exist or has much lower precision than slightly biased estimators. It is therefore worthwhile to consider requirements on h which are weaker than unbiasedness.

2. Mean or median point reproduction of a φ-ordering. In particular, we shall consider a requirement that h should, on the average, place each element of $\mathfrak{F}$ in the same relative position as φ does:

$$\mathcal{E}_F h(X) < \mathcal{E}_{F'} h(X) \quad \text{for} \quad \varphi(F) < \varphi(F') \tag{1}$$

or a similar requirement in terms of medians. (Thus, in place of (1) comes: any median under F of $h(X)$ is less than any median under F' of $h(X)$ if $\varphi(F)$ is less than $\varphi(F')$.) Note that the test previously discussed, if it exists, satisfies (1) if $\alpha < 1 - \beta$ since $P_F\{h(X) = 1\} \equiv \mathcal{E}_F h(X)$, and also satisfies the corresponding condition for medians if α and β are less than $\frac{1}{2}$ since then the median of $h(X)$ under F is 0 if $\varphi(F) = 0$ and is 1 if $\varphi(F) = 1$. (The referee suggests that it might be more appropriate to call a test satisfying $\alpha < 1 - \beta$ strongly unbiased, and one satisfying only (1) strictly unbiased. Note that in this paper we shall only consider an expectation to exist if it is finite.)

We shall call a mean or median reproduction[1] by h of the ordering (of $\mathcal{P}$ by φ), *proper* if h and φ have the same range of values (perhaps after deleting from the range of h, a set of values to which each element of $\mathfrak{F}$ assigns zero probability). As noted before, it may not be desirable to restrict consideration to proper reproduction. But then, since we observe only X, requirement (1) is not usually sufficient to permit us to associate F in some reasonable way with any definite element of $\mathcal{P}$, even if h should have complete precision in the sense that $h(X) \equiv \mathcal{E}_F h(X)$ with probability 1. In the case of complete precision, it would suffice to require in addition that the ranges of h and φ coincide, but whether this would suffice otherwise depends on the spacing of the φ values relative to the precision. An additional condition that would make it possible to associate F with a particular element of $\mathcal{P}$ in a reasonable manner is the existence of a functional φ_0 over $\mathfrak{F}$ which induces the same $\mathcal{P}$ (i.e., $\varphi_0(F) < \varphi_0(F')$ if and only if $\varphi(F) < \varphi(F')$) and which satisfies

$$\mathcal{E}_F h(X) < \varphi_0(F') \quad \text{and} \quad \varphi_0(F) < \mathcal{E}_{F'} h(X) \quad \text{for} \quad \varphi_0(F) < \varphi_0(F'). \tag{2}$$

[1] The general concepts "h reproduces the φ-ordering of $\mathfrak{F}$ in mean or in median" are defined at the end of this paragraph; equivalently, proper reproduction can be defined by (1) and the equality of the ranges of φ and h.

When (1) and (2) hold, we shall say that h *reproduces* the φ-ordering of $\mathfrak{F}$ in mean; when similar conditions hold in terms of medians, we shall say that h reproduces the φ-ordering in median.

If

$$(*) \qquad \varphi(F') < \varphi(F) < \varphi(F''),$$

and thus the same holds for φ_0, (2) implies that

$$\mathcal{E}_{F'}h(X) < \varphi_0(F) < \mathcal{E}_{F''}h(X) \quad \text{and} \quad \varphi_0(F') < \mathcal{E}_F h(X) < \varphi_0(F'').$$

Hence:

(A) If the range of φ is an interval, condition (2) implies condition (1), since in this case for every pair F', F'' with $\varphi(F') < \varphi(F'')$ there is an F satisfying $(*)$.

(B) If the range of φ is an open interval, condition (2) implies that h is an unbiased estimator of φ_0, since in this case for every F there is a pair F', F'' satisfying $(*)$ with $\varphi_0(F')$ and $\varphi_0(F'')$ arbitrarily close to $\varphi_0(F)$.

One could also define a concept of reproduction for the case in which h is an interval.

We briefly note that one obtains another concept of reproduction by replacing expectations in (1) and (2) by appropriate upper and lower fractiles.

3. Reproducibility. We shall say that a φ-ordering is mean reproducible if an h exists which reproduces the φ-ordering in mean. If this is still so when h is restricted to a given class $\mathfrak{H}$, we shall call the φ-ordering $\mathfrak{H}$-reproducible in mean. Similar concepts may be defined for medians and other fractiles.

4. Relation to certain other concepts in the literature. Consider again φ which is 0 over $\mathfrak{F}_0$ and 1 over $\mathfrak{F}_1$ and a procedure for deciding to which of these two subclasses which make up $\mathfrak{F}$ the distribution F of some X belongs if F is known to belong to $\mathfrak{F}$. If a procedure has a probability of erroneous decision which is never less than $\frac{1}{2}$, deciding by a flip of a coin would do at least as well.

Berger and Wald [3] consider a requirement that this probability be always less than $\frac{1}{2}$. They show that such a procedure exists if $\mathfrak{F}$ is a dominated family and if the set of probability mixtures of $\mathfrak{F}_0$ and the set of probability mixtures of $\mathfrak{F}_1$ have no common elements; and that under certain conditions the converse holds as well. (Actually [3] and [4] are in terms of the least upper bound of the actual errors involved. However, one may prove this result, and a parallel result in [4] which does not require domination, with just the condition on error probabilities stated in the text.) If the error probability is always less than $\frac{1}{2}$, the sum of the probabilities of the two kinds of error is always less than 1. The existence of a test with the latter property is called distinguishability in [4]; it is equivalent with the existence of a strictly unbiased test and proper mean reproducibility of the φ-ordering of $\mathfrak{F}$.

[3] also introduces the concept of distinct hypotheses: the hypothesis H_0 that F belongs to $\mathfrak{F}_0$ and the hypothesis H_1 that F belongs to $\mathfrak{F}_1$ are called distinct

if, given F_0 in $\mathfrak{F}_0$ and F_1 in $\mathfrak{F}_1$, there is a function h with possible values 0 and 1 such that $\mathcal{E}_{F_0}h(X) \neq \mathcal{E}_{F_1}h(X)$. Distinctness is implied by proper mean reproducibility of the φ-ordering. In [1] it is shown that, if $\mathfrak{F}$ is a denumerable collection of continous one dimensional distribution functions, the hypotheses are distinct.

[2] considers the case of infinite sequences $(X_1, X_2, \cdots, X_k, \cdots)$ of random variables with X_k taking values on a k-dimensional space, and speaks about distinguishability if the error probabilities can be made arbitrarily small: given any $\epsilon > 0$ there is an n and a function h_n over X_n with values 0 and 1, such that $\mathcal{E}h_n(X_n) < \epsilon$ in $\mathfrak{F}_0$ and $\mathcal{E}h_n(X_n) > 1 - \epsilon$ in $\mathfrak{F}_1$. This corresponds to asymptotic unbiasedness and consistency of tests. Hoeffding (who also considers sequential and randomized tests) calls this (finite) absolute distinguishability in [4] and distinguishability in [5].

5. Some propositions on mean reproducibility of an ordering of translation subfamilies. Let $\mathfrak{F}_c$, a subfamily of the partition $\mathcal{P}$, consist of the distributions $F_{c\theta}$ with θ ranging over a set A_c. If each $\mathfrak{F}_c$ contains a distribution F_c such that

$$F_{c\theta}(x) = F_c(x - \theta), \tag{3}$$

we shall say that $\mathcal{P}$ is a partition into *translation* subfamilies. It is possible to give a very simple criterion for the absence of mean reproducibility for this case.

We shall say that *condition A* is satisfied for two subfamilies $\mathfrak{F}_c$ and $\mathfrak{F}_{c'}$ if A_c and $A_{c'}$ are such that $F_{c'}$ assigns probability 1 to A_c and F_c assigns probability 1 to $A_{c'}$. For a simple illustration, let

$$F_c(x) = \textstyle\sum_{j \leq x} \binom{4}{j} p_c^{\,j}(1 - p_c)^{4-j}$$

and let A_c consist of the numbers 0, 1, 2, 3, 4 and 5 for each c. Each $\mathfrak{F}_c$ has 6 distributions in it; let $\varphi(F) = p_c$ for F in $\mathfrak{F}_c$ with p_c between 0 and 1. It follows from II below that on the basis of an observation from one of these populations one cannot find an estimate of $\varphi(F)$ whose mean will always be smaller for F in $\mathfrak{F}_c$ than for F in $\mathfrak{F}_{c'}$ if $p_c < p_{c'}$ even if we know p_c and $p_{c'}$.

We state these propositions:

I. *If the range of h is bounded from above or from below and condition A holds for two subfamilies of the partition, the φ-ordering is not mean reproducible.*

II. *If condition A holds for two subfamilies of the partition and the distributions in these two subfamilies assign probability* 1 *to a finite number of points, the φ-ordering is not mean reproducible.*

III. *If condition A holds for two subfamilies $\mathfrak{F}_c$ and $\mathfrak{F}_{c'}$ of the partition, if F_c and $F_{c'}$ have densities which tail off eventually*[2] *and are continuous except perhaps at a finite number of points, and if, for any h for which $\mathcal{E}_{F_c}h(X)$ and $\mathcal{E}_{F_{c'}}h(X)$ exist, $\mathcal{E}_{F_c}h(2X)$ and $\mathcal{E}_{F_{c'}}h(2X)$ also exist, the φ-ordering is not mean reproducible.*

Easy proofs of the first two propositions are given in the next section; let us,

---

[2] We shall say that a (possibly) multidimensional density f tails off eventually if, for some $k \geq 1$ and $a \geq 0$, $f(y) \leq kf(x)$ whenever $x \cdot y \geq x \cdot x \geq a$, where the dot indicates inner (or dot) product.

however, note two consequences. (Actually, what is shown is a stronger conclusion, viz. that (1) cannot be satisfied.) It follows from I that no strictly unbiased test exists of H_0 of Section 1 if $\mathfrak{F}_0$ and $\mathfrak{F}_1$ constitute a partition into translation families. From III it follows that, for most density functions that are met with in practice, sufficiently wide translation subfamilies have no unbiased estimators of $\varphi(F)$ if we impose the restraint that $\mathcal{E}_F h(2X)$ should exist whenever $\mathcal{E}_F h(X)$ does. Such a restraint is natural if φ is a *scale parameter*: $\varphi(F_c) = \tau \neq 0, \varphi(F_{c'}) = 1, F_c(x) = F_{c'}(\tau^{-1}x)$. It follows from I that no positive estimate h exists of $\varphi(F)$ with $\mathcal{E}_{F_{c\theta}} h(X) < \mathcal{E}_{F_{c'\theta'}} h(X)$ if $\tau < 1$, no matter what be the properties of the subfamilies. (In each of these cases we assumed condition A to hold.)

6. Proofs of Propositions I and II. To show II consider two subfamilies $\mathfrak{F}_c$ and $\mathfrak{F}_{c'}$ and let F_c assign positive probability p_q to a_q ($q = 1, \cdots, Q$) and $F_{c'}$ positive probability p_r' to b_r ($r = 1, \cdots, R$), with $\sum p_q = \sum p_r' = 1$; then

$$\mathcal{E}_{F_{c\theta}} h(X) = \textstyle\sum_q h(a_q + \theta)p_q, \tag{4a}$$

$$\mathcal{E}_{F_{c'\theta'}} h(X) = \textstyle\sum_r h(b_r + \theta')p_r'. \tag{4b}$$

Let φ be such that, e.g., $\varphi(F_c) < \varphi(F_{c'})$. We shall show that there is no h for which $\mathcal{E}_{F_{c\theta}} h(X) < \mathcal{E}_{F_{c'\theta'}} h(X)$ for all θ in A_c and all θ' in $A_{c'}$.

For suppose otherwise, then each of the numbers (4a) for θ in A_c is less than each of the numbers (4b) for θ' in $A_{c'}$. Since by condition A, $b_1, \cdots, b_R$ belong to A_c and $a_1, \cdots, a_Q$ belong to $A_{c'}$, each of the R numbers

$$\textstyle\sum_q h(a_q + b_r)p_q$$

is less than each of the Q numbers

$$\textstyle\sum_r h(b_r + a_q)p_r',$$

and so the (weighted) mean

$$\textstyle\sum_r \{\sum_q h(a_q + b_r)p_q\}p_r'$$

of the former numbers is less than the (weighted) mean

$$\textstyle\sum_q \{\sum_r h(b_r + a_q)p_r'\}p_q$$

of the latter numbers. Since the last two displayed expressions are equal, we have a contradiction.

In the proof of I we replace sums by integrals and identify two iterated integrals by Fubini's theorem for semibounded functions. The proof of III is more complicated and will be given elsewhere jointly with R. Sacksteder to whom I owe the idea of the proof of II.

7. An application. An example of a problem to which condition A applies is the following: Consider n patients and measure their blood pressure X once on each patient under standard conditions. (We assume that it is not possible or practicable to obtain several independent observations under standard conditions on each patient.) Let the n_1 first patients serve as controls, and let the other patients take a drug; we wish to study the possibility that the drug decreases the variability of the patients' blood pressure. Suppose that for the ith patient X has distribution $K_\sigma(x - \theta_i)$ if $i \leqq n_1$, and $K_\sigma(\tau^{-1}(x - \theta_i))$ if $i > n_1$. τ is unknown, except that it lies in the closed interval from b to 1, where b is a known positive number less than 1.

Condition A is fulfilled if it is known that the θ_i lie in an interval which includes a set to which $F_1^{(\sigma)}$ assigns probability 1, where $F_{\tau\theta}^{(\sigma)}$ is defined by

$$(5) \quad F_{\tau\theta}^{(\sigma)}(x) = K_\sigma(x_1 - \theta_1) \cdots K_\sigma(x_{n_1} - \theta_{n_1})K_\sigma(\tau^{-1}(x_{n_1+1} - \theta_{n_1+1})) \cdots K_\sigma(\tau^{-1}(x_n - \theta_n)),$$

when $x = (x_1, \cdots, x_n)$, $\theta = (\theta_1, \cdots, \theta_n)$. Usually σ is assumed known to lie in a certain interval and we have $K_\sigma(x) = K_1(x/\sigma)$.

If we do not wish to make an assumption about the range of σ, validity of condition A implies that we are not willing to put *a priori* bounds on the range of variation of the θ_i. If a variate with distribution K_1 can only take values between a_1 and a_2 ($> a_1$), and if σ is at most σ_0 when the θ_i can vary between $-c$ and $+c$, condition A is fulfilled if c exceeds the maximum $-\sigma_0 a_1$ and $\sigma_0 a_2$. Proposition I implies that under condition A we cannot estimate or test hypotheses about τ in any reasonable way when K_1 is known (and so also if K_1 is not fully known).

Consider a slightly different model in which it is supposed that the mean responses are also affected by the drug. Then,

$$(6) \quad F_{\tau\theta}^{(\sigma)\prime}(x) = K_\sigma(x_1 - \theta_1) \cdots K_\sigma(x_{n_1} - \theta_{n_1})K_\sigma(\tau^{-1}x_{n_1+1} - \theta_{n_1+1}) \cdots K_\sigma(\tau^{-1}x_n - \theta_n).$$

Note that (6) does not satisfy (3). Indeed, if, as is usual, the drugs are assigned to the patients by a random mechanism, the situation in this second model can be readily handled when $n_1 > 1$ and $n > 3$, and K_1 is the standard normal distribution. Writing

$$m(Y_1, \cdots, Y_l) = \textstyle\sum_{j=1}^{l} Y_j/l,$$
$$s^2(Y_1, \cdots, Y_l) = \textstyle\sum_{j=1}^{l} (Y_j - m(Y_1, \cdots, Y_l))^2/(l - 1),$$

we can use the F statistic $s^2(X_{n_1+1}, \cdots, X_n)/s^2(X_1, \cdots, X_{n_1})$ to make inferences about τ, if we approximate the situation by assuming that each X_i before treatment is the sum of two independent normal components, X_{i0} and θ_i, with means 0 and $\bar{\theta}$ and variances σ^2 and ω^2.

For then $(n - n_1 - 1)s^2(X_{n_1+1}, \cdots, X_n)/\{(\sigma^2 + \omega^2)\tau^2\}$ and $(n_1 - 1)s^2 \cdot (X_1, \cdots, X_{n_1})/(\sigma^2 + \omega^2)$ are independent chi-square variates with $n - n_1 - 1$ and $n_1 - 1$ degrees of freedom respectively, and we can use the lower $100\alpha\%$ point, F_α, of the corresponding F-distribution. E.g., to test whether $\tau = \tau_0$ against $\tau < \tau_0$, we reject when

$$\tau_0^{-2}s^2(X_{n_1+1}, \cdots, X_n)/s^2(X_1, \cdots, X_{n_1})$$

falls below F_α; the power of this test equals the integral of the F-density between 0 and $(\tau_0/\tau)^2 F_\alpha$.

8. Transformed translation subfamily problems. There are problems which can be transformed into a translation subfamily problem. Among the obvious examples are problems with positive variates and scale families, so that the logarithms of the variates form translation subfamilies.

In another class of problems a sufficient statistic for φ exists. Sometimes the families of distributions of a sufficient statistic corresponding to the families of parent distributions are translation families. The following case and its generalizations are discussed elsewhere ([6], [7]) in more detail.

Suppose $Y_1, \cdots, Y_n$ are joint normal with equal means μ, equal positive variances σ^2 and equal pairwise correlations ρ, all completely unknown and we wish to estimate the variance of $\bar{Y}$. It can be shown that, if an unbiased estimate of this parameter exists, it must depend on $\bar{Y}$ alone. But by III such an unbiased estimate does not exist.

REFERENCES

[1] Berger, A. (1950). On disjoint sets of distribution functions. *Proc. Amer. Math. Soc.* **1** 25–31.

[2] Berger, A. (1951). On uniformly consistent tests. *Ann. Math. Statist.* **22** 289–293.

[3] Berger, A. and Wald, A. (1949). On distinct hypotheses. *Ann. Math. Statist.* **20** 104–109.

[4] Hoeffding, W. (1956). The role of assumptions in statistical decisions. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **1** 105–114.

[5] Hoeffding, W. and Wolfowitz, J. (1958). Distinguishability of sets of distributions. *Ann. Math. Statist.* **29** 700–718.

[6] Konijn, H. S. (1964). On a theorem of Halmos concerning unbiased estimation of moments. *J. Austral. Math. Soc.* **4** 229–232.

[7] Konijn, H. S., Non-existence of consistent estimator sequences and unbiased estimators: a practical example. Submitted to *Sankhyā*.