Table I gives approximate 90 per cent probability limits for estimates of π based on 101 falls by the three methods considered. For estimates based on the average number of intersections the limits are given by $\pi \pm 1.645\, \sigma_{\hat{\pi}}$. For the estimate based on variation in the number of intersections, the limits are the estimates corresponding to $V = (1/1.28)(1 + 2/\pi - 16/\pi^2) = .0121$ and $V = 1.24(1 + 2/\pi - 16/\pi^2) = .0192$ where 1/1.28 and 1.24 are the 5th and 95th percentiles of the distribution of $F_{100,\infty}$ respectively.

TABLE I

Estimates of π based on 101 falls; 90% probability limits

Buffon needle case	
$L = 1$	2.75 to 3.53
Cartesian grid system	
Mean number of intersections	3.09 to 3.19
Variation in number of intersections	3.138 to 3.146

The estimate based on variation in the number of intersections is relatively insensitive to counting and measurement errors. Thus a 10 per cent error in measuring L will produce only $1/10$ of 1 per cent error in the estimate of π. A similar error in measuring L will produce a 10 per cent error in the estimate of π by the other methods. It should be remarked that the situation here is unusual in that the sample variance provides a much better estimate of the true mean number of intersections than does the sample mean. This is in contrast with the case of the Poisson distribution for which the sample mean provides the best estimate of the population variance.

REFERENCES

[1] E. Kasner and J. Newman, *Mathematics and the Imagination*, Simon and Schuster, 1940, pp. 246–247.

[2] J. V. Uspensky, *Introduction to Mathematical Probability*, McGraw-Hill Book Co., 1937, pp. 112–115, 251–257.

A HIGHER ORDER COMPLETE CLASS THEOREM[1]

By Lionel Weiss

University of Virginia and Cornell University

1. Introduction. The purpose of this note is to show that one can prove complete class theorems in which the risk for each possible distribution is not only a scalar, as is usual in the Wald theory, but actually a vector with as many com-

Received 5/7/53.

[1] Research under contract with the Office of Naval Research.

ponents as desired. The proof is an almost trivial reformulation of the proof of Theorem 1 in [1]. Other known results may be extended in a similar manner. An example of the application of the theorem will be found in Section 3. This is a complete class theorem which requires no assignment of a loss (weight) function, but only the classification of decisions into two classes, favorable and unfavorable, for each distribution function. This result may be satisfactory to those who maintain that in some or many situations the assignment of a loss function is difficult.

2. Complete class theorems when the risk is a vector. Let x be the generic point of a Euclidean space Z (the extension of the results of this paper to general abstract spaces is trivial). $F_1(x), F_2(x), \cdots, F_m(x)$ are m (>1) given cumulative probability distributions on Z. The statistician is presented with an observation on the chance variable X which is distributed in Z according to an unknown one of $F_1, \cdots, F_m$. On the basis of this observation he has to make one of L decisions, say $d_1, \cdots, d_L$. Let s be a positive integer and $W_{ijk}(x)$ $(i = 1, \cdots, m;$ $j = 1, \cdots, L; k = 1, \cdots, s)$ be measurable functions of x such that

$$\int_Z | W_{ijk}(x) | \, dF_i < \infty .$$

A randomized decision function, hereafter called "test" for short, and generically designated by $\eta(x)$, is defined as follows: $\eta(x) = [\eta_1(x), \cdots, \eta_L(x)]$, where

(a) $\eta(x)$ is defined for all x,

(b) $0 \leqq \eta_j(x)$ for $j = 1, \cdots, L$,

(c) $\sum_{j=1}^{L} \eta_j(x) = 1$ identically in x,

(d) $\eta_j(x)$ is measurable for $j = 1, \cdots, L$.

Let $r_{ik} = \int_Z (\sum_{j=1}^{L} \eta_j(x) W_{ijk}(x)) \, dF_i$ and $r^s = (r_{ik})$, $(i = 1, \cdots, m; k = 1, \cdots, s)$. Thus to each test $\eta(x)$ there corresponds the sth order risk point r^s. The test T with sth order risk point r^s will be said to be uniformly better (s) than the test T' with sth order risk point $r'^s = (r'_{ik})$ if $r_{ik} \leqq r'_{ik}$ for every i and k, with the inequality sign holding for at least one pair (i, k). A test T will be called admissible (s) if there exists no test uniformly better (s) than T. A class C of tests will be called complete (s) if, for any test T' not in C, there exists a test T in C which is uniformly better (s) than T'. A complete (s) class will be called minimal if no proper subclass of it is complete (s).

Wald's proof given in [1] obviously holds and we may state: The class of all admissible (s) tests is a minimal complete (s) class.

Any set $\xi = (\xi_{ik})$, $(i = 1, \cdots, m; k = 1, \cdots, s)$ of nonnegative numbers which add to unity (a convenient normalization) will be called an a priori distribution (s). A Bayes solution (s) with respect to ξ is a test T^* which minimizes

$$\sum_{i,k} \xi_{ik} r_{ik}(T) \tag{1}$$

with respect to all tests T.

THEOREM. *Every admissible (s) test is a Bayes solution (s) with respect to some a priori distribution (s). Hence the class of Bayes solutions (s) is complete (s).*

PROOF. Let $\{G_{ik}(x)\}$ $(i = 1, \cdots, m; k = 1, \cdots, s)$ be $m \cdot s$ distribution functions on Z, and let $G_{ik}(x) = F_i(x)$ for every i and k. Suppose the statistician has to make one of a set of L decisions which we may call $d_1, \cdots, d_L$. Let $W_{ijk}(x)$ be the loss incurred when x is the observed point, $G_{ik}(x)$ is the distribution function of X, and the decision d_j is made. Let $r'^*(\eta)$ be the first order risk point of this problem for the decision function $\eta(x)$. Then $r^s(\eta) = r'^*(\eta)$. The desired result now follows from Theorem 1 of [1], for the requirement made there that the F_i are distinct is never used.

Let f_i be the density function of F_i with respect to a measure μ with respect to which all F_i are absolutely continuous. There is always such a measure. To construct an sth order Bayes solution with respect to $\xi = (\xi_{ik})$ $(i = 1, \cdots, m; k = 1, \cdots, s)$ one may proceed as follows: $\eta_j(x) = 0$ for all j $(j = 1, \cdots, L)$ for which $\sum_{i=1}^{m} \sum_{k=1}^{s} \xi_{ik} f_i(x) W_{ijk}(x)$ is not a minimum with respect to j; $\eta_j(x)$ is defined arbitrarily between zero and one, inclusive, for all other j, provided only that every component of the resulting $\eta(x)$ is measurable and the sum is always one.

Other results found in [1] and elsewhere may be extended in a manner similar to that of the present theorem.

It is also obvious that one may prove similar results with the inequality signs reversed, by using anti-Bayes solutions (s), that is, tests which maximize (1) instead of minimizing it.

Using the ideas of [2] the above results may easily be extended to the case where the number of distributions and/or decisions is infinite and where the observations are taken sequentially, to obtain ϵ-complete (s) theorems.

3. Application to controlling probabilities of making the various decisions. $P(i, j \mid T)$ will denote the probability of making decision d_j when F_i is actually the distribution and the test T is employed. In other respects the notation of Section 2 is used. For each i, we suppose that there are certain decisions which are favorable (i.e., we prefer to make them when F_i is the distribution), and the others are unfavorable (we prefer not to make them when F_i is the distribution). We assume that for each i there is at least one favorable and one unfavorable decision. For our present purpose, s is equal to L, the number of decisions. For given i and k, we define $W_{ijk}(x)$ as follows. If d_k is favorable relative to F_i, $W_{ijk}(x) = 0$ if $j = k$, $W_{ijk}(x) = 1$ if $j \neq k$. If d_k is unfavorable relative to F_i, $W_{ijk}(x) = 1$ if $j = k$, $W_{ijk}(x) = 0$ if $j \neq k$. Then we have the following result. Let T be any test which is not a Bayes solution (s). There is a Bayes solution (s), T', such that for any i $(i = 1, \cdots, m)$ and any j such that d_j is unfavorable relative to F_i, we have $P(i, j \mid T') \leqq P(i, j \mid T)$; while for any i and any j' such that $d_{j'}$, is favorable relative to F_i, we have $P(i, j' \mid T') \geqq P(i, j' \mid T)$. The inequality sign holds in at least two of these relations, one in each set.

4. Other applications. We mention two other applications of the results. If there are s individuals with possibly different loss functions, $W_{ijk}(x)$ can denote the loss suffered by individual k when d_j is made and F_i is true and x is observed. Or different true situations may lead to the same distribution of the observable chance variable, so that $W_{ijk}(x)$ is the loss incurred under the kth true situation leading to the distribution F_i. The range of k may depend upon i, and all the results hold.

REFERENCES

[1] A. WALD AND J. WOLFOWITZ, "Characterization of the minimal complete class of decision functions when the number of distributions and decisions is finite," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1951.

[2] J. WOLFOWITZ, "On ϵ-complete classes of decision functions," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 461–465.

CORRECTION OF A PROOF*

BY J. KIEFER

Cornell University

In the proof of Theorem 3 of "On Wald's Complete Class Theorems" (*Ann. Math. Stat.*, Vol. 24 (1953), pp. 70–75), the inequality appearing in the definition of $r_{2,m}(\xi)$ should be altered to read $r(\xi, \delta^m) \geqq r(\xi, \delta_2) - \epsilon/2$; the remainder of the proof is then easily altered to give the desired result. Without the $\epsilon/2$, one would still have to prove that the space $\mathfrak{D}$ is large enough to give $\lim_{m\to\infty} r_{2,m}(\xi) < \infty$. The author is indebted to Mr. Jerome Sacks for pointing out this fact.

* Received 7/11/53.

ABSTRACTS OF PAPERS

(Abstracts of papers presented at the Stanford meeting of the Institute, June 19–20, 1953)

1. **On the Probability Function of the Quotient of Sample Ranges from a Rectangular Distribution.** LEO A. AROIAN, Hughes Aircraft and Development Laboratories, Culver City.

In a recent paper Paul R. Rider (*J. Amer. Stat. Assn.*, Vol. 46 (1951), pp. 502–507) has derived the probability function of $u = R_1/R_2$, the quotient of the sample ranges of two independent random samples from $f(x) = 1/x_0$ for $0 \leqq x \leqq x_0$, $f(x) = 0$ elsewhere, where R_1 is the sample range in a sample of m and R_2 is the sample range in a sample of n from $f(x)$. The power function of the test is derived, the tables are extended for the 5 per cent, $2\frac{1}{2}$ per cent, 1 per cent, and $\frac{1}{2}$ per cent levels of significance. In case m and n large a Cornish-Fisher expansion for the levels of significance is derived. The transformation $w = \frac{1}{2} \log_e u$ is found convenient and use is made of the moment generating function of w to find the