On Optimal Tests for Separate Hypotheses and Conditional Probability Integral Transformations

by

C. P. Quesenberry and R. R. Starbuck

Institute of Statistics
Mimeo Series #985
Raleigh - 1975

ON OPTIMAL TESTS FOR SEPARATE HYPOTHESES AND CONDITIONAL PROBABILITY INTEGRAL TRANSFORMATIONS

C. P. Quesenberry and R. R. Starbuck
North Carolina State University

SUMMARY

Consider the problem of testing the composite null hypothesis that a random sample $X_1, \ldots, X_n$ is from a parent which is a member of a particular continuous parametric family of distributions against an alternative that it is from a separate family of distributions. It is shown here that in many cases a uniformly most powerful similar (UMPS) test exists for this problem, and, moreover, that this test is equivalent to a uniformly most powerful invariant (UMPI) test. It is also seen in the method of proof used that the UMPS test statistic is a function of the statistics $U_1, \ldots, U_{n-k}$ obtained by the conditional probability integral transformations (CPIT), and thus that no information is lost by these transformations. It is also shown that these optimal tests have power that is a monotone function of the null hypothesis class of distributions, so that, for example, if one additional parameter for the distribution is assumed known, then the power of the test can not decrease. Two readily established but important properties of CPIT transformations are given. It is first shown that the statistics given by these transformations are independent of the complete sufficient statistic, and that these statistics have important invariance properties. Some examples are given for particular families. These include testing the two-parameter uniform family against the two-parameter exponential family to illustrate the transformation approach given here for constructing UMPS tests. The problem of testing a scale parameter exponential family against a shape parameter lognormal family is considered for the shape parameter known, and for it unknown. Some empirical power results have been computed for the tests proposed here for these two problems, and these results are compared with those of other writers.

Some key words: Separate families, similar tests, invariant tests, Neyman-Pearson Lemma, conditional probability integral transformations, goodness-of-fit.

1. INTRODUCTION

Let $X_1, \ldots, X_n$ be a sample from a continuous parent distribution. We are interested in testing that this sample is from one family of distributions against the alternative that it is from another family where these families are separate as defined by Cox (1961). Cox (1961, 1962) has given a general method for deriving tests for these problems. His tests are based on the logarithm of the maximum likelihood ratio-MLR tests. Jackson (1968) and Atkinson (1970) have both considered the MLR method further and developed some particular tests.

Invariant tests are also important for the separate families testing problem. Lehmann (1959) gave the general theory of uniformly most powerful invariant (UMPI) tests. In an earlier paper Lehmann (1950) attributes the notion of invariance to R. A. Fisher, Hotelling, Pitman and others; and much of its development to Hunt and Stein in unpublished work. See, also, Hajek and Sidak (1967). Uthoff (1970, 1973) has considered UMPI tests for several particular problems. Dyer (1971, 1973, 1974) has considered several tests for separate hypotheses problems, including MLR and UMPI statistics. Antle, Dumonceaux, and Haas (1973) compared the power of MLR and UMPI tests for some location-scale parameter problems and recommended the MLR test over the UMPI test (when they differ), because it is more easily derived and has relatively good power. Dumonceaux and Antle (1973) give the MLR procedure for discriminating between lognormal and Weibull distributions.

Most goodness-of-fit tests can be used as tests of separate hypotheses. However, since these tests ignore the alternative hypothesis they would be expected to have less power than tests that utilize knowledge of the alternative family. There is no assurance that this loss of power will in fact occur, unless the test with which a goodness-of-fit test is being compared has maximum power among a class of tests to which both tests belong. Goodness-of-fit tests can have unexpected power properties, such as those reported by Dyer (1974), and independently by Stephens (1974). They found empirically that if the null hypothesis class is contracted by assuming a parameter to be known, that the power of some tests is decreased. This cannot happen for the most powerful tests considered here, and a statement and proof of this is given in Theorem 4.4. below.

The conditional probability integral transformations (CPIT) were introduced in O'Reilly and Quesenberry (1973). These transformations can be used to transform a sample $X_1, \ldots, X_n$ from a k-parameter continuous null hypothesis class to a set $U_1, \ldots, U_{n-k}$ of independently and identically distributed uniform random variables on the unit interval -- i.i.d. $U(0,1)$. Then many goodness-of-fit test statistics can be used to test the uniformity of the U's on the $(n-k)$ dimensional hypercube, and the test results applied to the original composite null hypothesis testing problem. The advantage of this approach is largely practical, for it allows the use of the same distributional results (tables, limiting distributions, etc.) to test a large number of different null hypothesis classes. One can, by this approach, obtain immediately an exact test for all sample sizes for a large number of composite hypotheses for which no test is presently available. In contrast to this, the large number of goodness-of-fit tests that have been proposed for composite null hypothesis problems have distributions that change when a different null hypothesis class is considered; and, in fact, from this perspective it should be observed that the tests for separate hypotheses have different distributions when either the null or the alternative family is changed. A question arises as to the power that can be achieved by this CPIT approach. We derive here a test that is most powerful among all tests based on the U's for testing separate families, and show that the test obtained is, in fact, a UMPS test for the problem. It is further shown that this UMPS test and a UMPI test are equivalent under conditions commonly met.

The basic approach adopted here of seeking a most powerful test among the class of similar tests is, we feel, obviously reasonable. If one is interested in testing that a sample is from a particular parametric family; such as $N(\mu,\sigma^2)$, say; against some alternative outside the family, then the test should treat all members of the family the same, i.e., all normal distributions are "equally" normal. Also, all of the test statistics, of which we are aware, that have been proposed for testing these types of composite null hypotheses have distributions that are the same for every member of the null hypothesis class. The tests therefore are all similar, and only similar tests seem to ever be used, anyway.

2. THE MODEL: TERMINOLOGY, NOTATION AND PRELIMINARY RESULTS

We shall frequently use terminology given in Lehmann (1959); Chapter 6, invariance, is especially relevant.

Let $\mathcal{X}$ denote a Borel set of real numbers, $\mathcal{A}$ the Borel subsets of $\mathcal{X}$, and $X = (X_1, \ldots, X_n)$ denote a vector of independently and identically distributed (i.i.d.) random variables, each distributed according to an absolutely continuous distribution P on the Borel space $(\mathcal{X},\mathcal{A})$; and, further, suppose that P is a member of a parametric class of distributions $\mathcal{P} = \{P_\theta;\ \theta \in \Omega\}$. The set Ω is assumed to be a k-dimensional Borel set with elements $\theta = (\theta_1, \ldots, \theta_k)$. It is also assumed that there exists a k-dimensional sufficient statistic $T = (T_1, \ldots, T_k)$ for $\mathcal{P}$ (or Ω), defined on the sample space $(\mathcal{X}^n,\mathcal{A}^n) = (\mathcal{X} \times \ldots \times \mathcal{X}, \mathcal{A} \times \ldots \times \mathcal{A})$. Also, put $\mathcal{P}^n = \{P^n;\ P^n = P \times \ldots \times P,\ P \in \mathcal{P}\}$, i.e., $\mathcal{P}^n$ is a class of product measures on $(\mathcal{X}^n,\mathcal{A}^n)$ corresponding to $\mathcal{P}$. The class $\mathcal{P}^n$ is also written as $\mathcal{P}^n = \{P^n_\theta;\ \theta \in \Omega\}$. If $(X_1, \ldots, X_{n-k})$ given T has an absolutely continuous distribution, we say $\mathcal{P}$ has absolute continuity rank n - k, i.e., a.c.r. $\mathcal{P}$ is n - k (cf. O'Reilly and Quesenberry (1973)).

Let g: $\mathcal{X} \to \mathcal{X}$ be a one-to-one transformation, and let g^n be the corresponding one-to-one transformation of $\mathcal{X}^n$ onto $\mathcal{X}^n$ defined by $g^n(x_1, \ldots, x_n) = (g(x_1), \ldots, g(x_n))$. For each g^n, suppose there exists a function $\bar{g}$: $\Omega \to \Omega$ (or $\mathcal{P}^n \to \mathcal{P}^n$) such that $P_{\bar{g}\theta}(X \in g^n A) = P_\theta(X \in A)$ for every $A \in \mathcal{A}^n$. Let G denote a transformation group on $\mathcal{X}$, g^n the corresponding transformation group on $\mathcal{X}^n$, and $\bar{G}$ the corresponding transformation group on Ω (that G^n is a transformation group is easily seen, that $\bar{G}$ is a group follows from Lehmann (1959), p. 214).

Denote by $\mathcal{B}_S$ the sub σ-algebra of $\mathcal{A}$ induced by a statistic S; by $h_1 \circ h_2$ the composition of a function h_1 with a function h_2, i.e., $h_1 \circ h_2(\cdot) = h_1(h_2(\cdot))$; by I_A the indicator function of a set A. With the usual abuse of notation the same symbol, e.g., g or g^{-1} will be used to denote a point function and the corresponding set function. The following lemma is a well-known result in probability theory that provides a convenient starting point for this work.

Lemma 2.1

For g: $\mathcal{X} \to \mathcal{X}$, one-to-one onto, and S a statistic on $(\mathcal{X}^n,\mathcal{A}^n)$,

$$P_\theta(A|S) = P_{\bar{g}\theta}(g^n A | g^n \mathcal{B}_S) \circ g^n \quad \text{a.s. } P_\theta \ \forall A \in \mathcal{Q}^n . \tag{2.1}$$

Proof. From the definition of an induced σ-algebra it follows easily that $g^n \mathcal{B}_S = \mathcal{B}_{S \circ g^{-n}}$. Thus $\forall B \in \mathcal{B}_S$,

$$\int_{g^n B} P_\theta(A|\mathcal{B}_S) \circ g^{-n} \, dP_{\bar{g}\theta} = \int_B P_\theta(A|\mathcal{B}_S) dP_\theta$$

$$= \int_B I_A dP_\theta = \int_{g^n B} I_{g^n A} dP_{\bar{g}\theta}, \text{ since } I_{g^n A} = I_A \circ g^{-n} ,$$

$$= \int_{g^n B} P(g^n A | g^n \mathcal{B}_S) dP_{\bar{g}\theta} .$$

Then by the Radon-Nikodym Theorem,

$$P_\theta(A|\mathcal{B}_S) \circ g^{-n} = P_{\bar{g}A}(g^n A | g^n \mathcal{B}_S) \quad \text{a.s. } P_{\bar{g}\theta} \ \forall A \in \mathcal{Q}^n ,$$

and (2.1) follows. This completes Lemma 2.1.

A transformation group on a space is said to be transitive if its maximal invariant is constant on the space. Transitive groups on parameter spaces will play an important role in this work.

Lemma 2.2. If

(a) G is a transformation group of increasing functions on $\mathcal{X}$ that induces a transitive group $\bar{G}$ on Ω; and

(b) T is a sufficient statistic for Ω, and is equivalent to $T \circ g^n$ for every $g \in G$;

then the distribution function of the conditional distribution of $(X_1, \ldots, X_n)$ for fixed T is invariant under G^n, i.e.,

$$F(x_1, \ldots, x_n \mid \mathcal{B}_T) = F(gx_1, \ldots, gx_n \mid \mathcal{B}_T) \circ g^n , \qquad (2.2)$$

a.s. $\mathcal{P}^n$ $\forall$ $g \in G$.

Proof. Let $\theta \in \Omega$ be fixed and θ' also be an element of Ω. Then there exists a $g \in G$ such that for the corresponding $\bar{g}$, $\theta' = \bar{g}\theta$. Let $J_x = \{(y_1, \ldots, y_n); y_i \leq x_i; i = 1, \ldots, n\}$ and $x = (x_1, \ldots, x_n)$. Then

$$F_\theta(x_1, \ldots, x_n|\mathcal{B}_T) = P_\theta(J_x|\mathcal{B}_T) \qquad \text{a.s.} \quad P_\theta,$$

$$= P_{\bar{g}\theta}(g^n J_x|\mathcal{B}_T) \circ g^n \qquad \text{a.s. } P_\theta \text{ , by Lemma 2.1 and (b),}$$

$$= P_\theta(g^n J_x|\mathcal{B}_T) \circ g^n \qquad \text{a.s. } P_\theta \text{ , by sufficiency of T,}$$

$$= F_\theta(gx_1, \ldots, gx_n|\mathcal{B}_T) \circ g^n \qquad \text{a.s. } P_\theta \; .$$

By the sufficiency of T, the subscript θ can be omitted on F, leaving

$$F(x_1, \ldots, x_n|\mathcal{B}_T) = F(gx_1, \ldots, gx_n|\mathcal{B}_T) \circ g^n \qquad \text{a.s. } \mathcal{P}^n ,$$

where the exceptional set may depend on g^n, and thus $F(x_1, \ldots, x_n|T)$ is almost invariant under G^n. However, since $\mathcal{P}^n$ is a dominated family on a Euclidean space, and both $\mathcal{X}^n$ and Ω are Euclidean sets, it follows by Lehmann (1959), Theorem 4 and discussion on p. 226, that the exceptional set does not depend on g^n. This completes Lemma 2.2.

It may appear that the conditions (a) and (b) are severely restrictive, but they are satisfied and readily verified for most of the separate families that have been considered in the literature. Important families that are easily verified include many location-scale families (for which $G = \{ax + b;\ a > 0, -\infty < b < \infty\}$), and shape parameter families such as the Pareto (for which $G = \{ x^a;\ a > 0\}$).

3. SOME PROPERTIES OF CONDITIONAL PROBABILITY TRANSFORMATIONS

Conditional probability integral transformations are introduced in O'Reilly and Quesenberry (1973). In this section we assume that a.c.r. $\mathcal{P}$ is n - k, i.e., that the conditional distribution of $(X_1, \ldots, X_{n-k})$ given T is absolutely continuous.

Put

$$u_1(x_1; T(x_1, \ldots, x_n)) = P(X_1 \le x_1 | T) ,$$

$$u_2(x_2; T(x_1, \ldots, x_n), x_1) = P(X_2 \le x_2 | T, x_1) , \qquad (3.1)$$

$$\vdots$$

$$u_{n-k}(x_{n-k}; T(x_1, \ldots, x_n), x_1, \ldots, x_{n-k-1}) = P(X_{n-k} \le x_{n-k} | T, x_1, \ldots, x_{n-k-1}) ,$$

and $u(x_1, \ldots, x_{n-k}, T) = (u_1, \ldots, u_{n-k})$. In O'Reilly and Quesenberry (1973) it is shown that $(U_1, \ldots, U_{n-k}) = (u_1(X_1, T) , \ldots, u_{n-k}(X_{n-k}; T, X_1, \ldots, X_{n-k-1}))$ are independently and identically distributed $U(0,1)$ random variables. From this and a result of Basu (1955), the next theorem is immediate.

Theorem 3.1. If $T = (T_1, \ldots, T_k)$ is a complete and sufficient statistic for Ω, then $(T_1, \ldots, T_k)$ and $(U_1, \ldots, U_{n-k})$ are independent vectors.

This theorem has important applications for constructing inference procedures that may be alternatives to nonparametric or robust procedures. The sufficient statistic T contains all the information for making inferences within the family $\mathcal{P}$ (or Ω), whereas the statistic $U = (U_1, \ldots, U_{n-k})$ contains information about the family $\mathcal{P}$. Thus U may be used to make inferences about the class $\mathcal{P}$, such as a goodness-of-fit test for the class $\mathcal{P}$, and T to make a parametric test within $\mathcal{P}$, and the independence exploited to assess overall error rates. Inferences based on U are considered in the following sections. In the next theorem we shall require that the statistic T be doubly transitive (cf. O'Reilly and Quesenberry (1973)).

Theorem 3.2. If the assumptions of Lemma 2.2 are satisfied, and T is doubly transitive, then u is equivalent to an invariant statistic, i.e.,

$$u(x_1, \ldots, x_n, T(x_1, \ldots, x_n)) = u(gx_1, \ldots, gx_n, T(gx_1, \ldots, gx_n))$$

a.s. $\mathcal{P}^n$ $\forall$ $g \in G$.

Proof. By Theorem 2.4 of O'Reilly and Quesenberry (1973),

$$u_j = E(F_j(x_j|T)|X_1, \ldots, X_{j-1}), \; j = k+1, \ldots, n \quad \text{a.s. } \mathcal{P} .$$

By Lemma 2.2 above, $F_j(x_j|T)$ is invariant under G, $j = k + 1, \ldots, n$. The result follows.

The following lemma is a consequence of the fact that the transforming functions of (3.1) are (conditional) continuous distribution functions.

Lemma 3.1. In the conditional space for fixed $T = t$, there is a.s. a one-to-one correspondence between $(u_1, \ldots, u_{n-k})$ and $(x_1, \ldots, x_n)$, i.e., $(U_1, \ldots, U_{n-k})$ and $(X_1, \ldots, X_n)$ are equivalent statistics, in this space.

4. MOST POWERFUL SIMILAR TESTS FOR SEPARATE FAMILIES

We consider using the sample $(X_1, \ldots, X_n)$ from a parent probability distribution P on $(\mathcal{X}, \mathcal{G})$ to test the null hypothesis

$$H: \; P \in \mathcal{P}_0 = \{P_\theta; \; \theta \in \Omega_0\} , \qquad (4.1)$$

against the composite alternative

$$K: \; P \in \mathcal{P}_1 = \{P_\tau; \; \tau \in \Omega_1\} , \qquad (4.2)$$

and $\mathcal{P}_0 \cap \mathcal{P}_1 = \emptyset$. It will sometimes be useful to consider a simple alternative

$$K': \; P = P_1 \, , \; P_1 \in \mathcal{P}_1 . \qquad (4.3)$$

We assume in the sequel that the a.c.r. of $\mathcal{P}_0$ is $n - k$.

Let $f_{(\cdot)}$ and $F_{(\cdot)}$ denote the density and distribution function, respectively, corresponding to $P_{(\cdot)}$. We assume that $\mathcal{P}_0$ and $\mathcal{P}_1$ are both identifiable classes, and that at every $x \in \mathcal{X}$ at least one distribution in each class has a positive density. Cox (1961) calls the classes $\mathcal{P}_0$ and $\mathcal{P}_1$ separate families of distributions if the density of an arbitrary member of either class cannot be obtained as the limit of a sequence of densities from the other class. We shall here be concerned entirely with such separate families $\mathcal{P}_0$ and $\mathcal{P}_1$.

Recall that a test is called similar-α for testing H vs K if it has constant probability α of rejection for every distribution in $\mathcal{P}_0$. More precisely, φ is a similar-α test function for H vs K if $E_{P_\theta}(\varphi) = \alpha$ for every $\theta \in \Omega_0$.

For $(u_1, \ldots, u_{n-k})$ as given by (3.1), let $h_1(u_1, \ldots, u_{n-k})$ denote the density of $(U_1, \ldots, U_{n-k})$ when $X_1, \ldots, X_n$ are i.i.d. from P_1 of K'. From the remark preceding Lemma 3.1, it follows that $h_1(u_1, \ldots, u_{n-k})$ is zero a.s. except in the unit hypercube. The next lemma is a direct application of the Neyman-Pearson Lemma.

Lemma 4.1. The most powerful level-α test of H versus K' based on u is

$$\psi(u) = \begin{cases} 1, & \text{if } h_1(u) > c \,, \\ 0, & \text{otherwise} \,, \end{cases} \tag{4.4}$$

where c is determined by $P\{h_1(U) > c\} = \alpha$, for $U = (U_1, \ldots, U_{n-k})$ i.i.d. $U(0,1)$ random variables.

Let $\varphi = \psi(u(x_1, \ldots, x_n))$, or, $\varphi = \psi \circ u$.

Theorem 4.1. If T is a boundedly complete sufficient statistic for $\mathcal{P}_0$ of (4.1), then the test $\varphi = \psi \circ u$ above is a MPS-α test for H versus K'.

Proof. By Lehmann (1959), Theorem 2, p. 134,

$$E_p\{\varphi(X_1, \ldots, X_n)\} = \alpha \qquad \forall\ P \in \mathcal{P}_o , \qquad \text{if and only if}$$

$$E_p\{\varphi(X_1, \ldots, X_n)|T\} = \alpha \text{ a.s. } \mathcal{P}^T .$$

Thus to find a most powerful test in the class of similar α-tests it is sufficient to find a most powerful conditional size-α test on the conditional space of $X_1, \ldots, X_n$ given T, i.e., to find the most powerful Neyman-structure test. But for $T = t$ fixed, $(X_1, \ldots, X_n)$ and $(U_1, \ldots, U_{n-k})$ are equivalent statistics by Lemma 3.1. Thus, the test φ is a MPS-α test. Theorem 4.1 is complete.

It will sometimes be the case that φ does not depend on P for $P \in \mathcal{P}_1$ of (4.2). Then, of course, φ is a uniformly most powerful similar-α (UMPS-α) test for H *versus* K . Conditions under which such tests exist are considered in the following theorem.

Theorem 4.2. If the conditions for both Theorem 3.2 and Theorem 4.1 are satisfied by $\mathcal{P}_o$, then a uniformly most powerful invariant level-α (UMPI-α) test exists for testing H *versus* K, provided the group $\overline{G}_1$ induced on $\mathcal{P}_1$ by G is also transitive. Moreover, this test is equivalent to the MPS-α test of Theorem 4.1, which is then UMPS-α.

Proof. If φ is invariant level-α, then since $\overline{G}$ is transitive it follows from Lehmann (1959), Theorem 3, p. 220, that

$$E_p(\varphi) = \alpha \qquad \forall\ P \in \mathcal{P}_o ,$$

i.e., φ is a similar-α test. Thus, if a test is MPS-α, it will be most powerful invariant level-α, provided it is invariant. But by Theorem 4.1, a MPS-α

test can a.s. be written as a function of $u_1, \ldots, u_{n-k}$ only, and is $\mathcal{B}_U$-measurable and invariant a.s. by Theorem 3.2. This completes Theorem 4.2.

If the null and alternative hypothesis classes are interchanged, i.e., if the problem of testing $\mathcal{P}_1$ vs $\mathcal{P}_o$ is considered, then its solution can be obtained directly from the solution for the problem of testing $\mathcal{P}_o$ vs $\mathcal{P}_1$ given in Theorem 4.2. We state this in Theorem 4.3.

Theorem 4.3. If the conditions for Theorem 4.2 are satisfied and if these conditions are satisfied with $\mathcal{P}_o$ and $\mathcal{P}_1$ interchanged, then a UMPS-α' test for $\mathcal{P}_1$ vs $\mathcal{P}_o$ is given by the test $\varphi' = 1-\psi(u(x_1, \ldots, x_n))$, where ψ is as given in Lemma 4.1 with $\alpha' = 1-\alpha$.

Proof. Let $\varphi = \psi(u(x_1, \ldots, x_n))$ as in Theorem 4.1. Then $\varphi' = 1-\varphi$, and if $P_o \in \mathcal{P}_o$ and $P_1 \in \mathcal{P}_1$, it follows that $E_{P_1}(\varphi') = \beta$, say, is a minimum for fixed $E_{P_o}(\varphi') = \alpha' = 1-\alpha$. Or, for β fixed α' is a maximum. This completes Theorem 4.3.

The next definition, and, particularly, Theorem 4.4 are motivated by results of Dyer (1974) and Stephens (1974). They found empirically that a number of well-known goodness-of-fit tests have the property that their power is less when the value of a parameter is assumed known than for the case when the same parameter is assumed unknown, under the null hypothesis. In the next definition and Theorem 4.4 conditions are given which assure that the power of the UMPS-α test for a smaller null hypothesis family is never less than that of the UMPS-α test for a larger family.

Two families of distributions on the same space $(\mathcal{X},\mathcal{A})$ are said to be conformable if there exists a group G of transformations on $\mathcal{X}$ and corresponding groups $\bar{G}_1$ and $\bar{G}_2$ on the parameter spaces that are transitive.

Consider two testing problems

$$H_1:\ \mathcal{P}_{1H}\ \underline{\text{versus}}\ K_1:\ \mathcal{P}_{1K}\ , \tag{4.5}$$

and

$$H_2:\ \mathcal{P}_{2H}\ \underline{\text{versus}}\ K_2:\ \mathcal{P}_{2K}\ , \tag{4.6}$$

where $\mathcal{P}_{1H} \subset \mathcal{P}_{2H}$, $\mathcal{P}_{1K} \subset \mathcal{P}_{2K}$, and $\mathcal{P}_{iH}$ and $\mathcal{P}_{iK}$ are conformable separate families of distributions, $i = 1, 2$.

Theorem 4.4. If φ_1 is UMPS-α for (4.5) and φ_2 is UMPS-α for (4.6), then for $P_1 \in \mathcal{P}_{1K}$ and $P_2 \in \mathcal{P}_{2K}$

$$E_{P_1}(\varphi_1) \geq E_{P_2}(\varphi_2)\ . \tag{4.7}$$

Proof. The class of tests that are similar-α for (4.6) is a subclass of the class of tests that are similar-α for (4.5). Thus the power of φ_i is constant on $\mathcal{P}_{ik}$; $i = 1, 2$; and the test with the maximum power in the superclass must have power not less than the test in the subclass. This completes Theorem 4.4.

In the next section we will consider some particular examples, and we conclude this section with some discussion of the results obtained. Theorem 4.1 shows that any MPS test can be written as a function of the CPIT U statistics, and thus the search for an MPS test can be made in the space of the U's. In this power sense, the CPIT U-transformations thus contain all the information in the sample.

The equivalence of UMPS and UMPI tests is of considerable theoretical interest, and in practice it offers the advantage of allowing the derivation of these tests by the approach that is easier for a particular case. In the derivation of the UMPS test by the CPIT transformation approach, the main task is to find the marginal density $h_1(u_1, \ldots, u_{n-k})$ of $(U_1, \ldots, U_{n-k})$ under the alternative hypothesis. In many problems, but not all, this is a difficult problem due to the rather complex nature of the u-transformations. In section 5 we consider some examples using the CPIT transformation approach.

The results of Theorem 4.3 are of some theoretical interest, we feel, in that it seems quite natural to use the same test statistic to test $\mathcal{P}_1$ vs $\mathcal{P}_0$ as for testing $\mathcal{P}_0$ vs $\mathcal{P}_1$. This property is not shared by the MLR tests of Cox (1961), in general. This result also provides another possible method for deriving the UMPS test, since the transformations and distributions involved are different and one derivation will sometimes be much easier than the other.

Finally, it should be mentioned that the results of this section may be very helpful for testing separate families for some cases when the assumptions required for the optimal tests discussed here do not hold. The most common problem where the assumptions do hold is probably that of testing two location-scale parameter families. However, if the null and alternative families are not conformable, then some alternative approach may be attractive. For example, by using Theorems 4.1 and 4.2 it is often possible to find a MPS test against one member of an alternative family, or, even against a subclass of an alternative family; but the test may depend upon further nuisance parameters. It may then be possible to obtain a nonoptimal but very good test by estimating the parameters in the forgoing optimal test. A problem of this type is considered in section 5 below, where we consider testing a scale parameter exponential family against a shape parameter lognormal family, and a very good test is obtained by estimating the lognormal parameter in the MPS test for testing a scale parameter exponential family against a particular lognormal distribution.

5. APPLICATIONS

In this section we consider two examples.

Example 5.1 Uniform vs exponential

Let $\mathcal{P}_0$ be the uniform distributions with densities

$$(\theta_2-\theta_1)^{-1} I_{(\theta_1,\theta_2)}(x); \ -\infty < \theta_1 < \theta_2 < +\infty \ ;$$

and $\mathcal{P}_1$ be the exponential family with densities

$$\lambda \exp\{-\lambda(x-\theta)\} I_{(\theta,\infty)}(x) \ , \ \lambda > 0 \ .$$

Both of these are location-scale parameter families, and the conditions for Theorem 4.2 are readily verified. Uthoff (1973) has given the UMPI test for testing $\mathcal{P}_0$ vs $\mathcal{P}_1$, which rejects for small values of $(\bar{x}-x_{(1)})/(x_{(n)}-x_{(1)})$. We give the CPIT derivation now. CPIT transformations for sample observations $x_1, \ldots, x_n$ for the uniform family are

$$u_i = (z_i-z_1)/(z_n-z_1) \ ; \ i = 2, \ldots, n-1 \ ; \qquad (5.1)$$

where $z_1 = x_{(1)}$, $z_n = x_{(n)}$, and the other z's are defined as follows. Suppose $z_1 = x_j$ and $z_n = x_k$, $j < k$. Then $z_i = x_{i-1}$, $i = 2, \ldots, j$; $z_i = x_i$, $i = j+1, \ldots, k-1$; and $z_i = x_{i+1}$, $i = k, \ldots, n-1$. It can be verified by direct Jacobian methods that these transformations give i.i.d. $U(0,1)$ random variables.

To find the UMPS test the joint distribution of $u_2, \ldots, u_{n-1}$ must be obtained under the assumption that $X_1, \ldots, X_n$ constitute a random sample from an exponential (θ,λ) distribution. Without loss of generality, let $(\theta,\lambda) = (0,1)$. The constant in the density h_1 is not needed and will be omitted in the following development. The joint density of $z_1, \ldots, z_n$ is

$$f(z_1, \ldots, z_n) \propto \exp\left(-\sum_{i=1}^{n} z_i\right) I_{(0,\infty)}(z_1) I_{(z_1,\infty)}(z_n) \prod_{i=2}^{n-1} I_{(z_1,z_n)}(z_i) \ . \quad (5.2)$$

The joint density of $u_1, \ldots, u_n$ is

$$g(u_1, \ldots, u_n) \propto u_n^{n-2} \exp\left\{-nu_1 - u_n\left(1 + \sum_{i=2}^{n-2} u_i\right)\right\} \cdot$$

$$I_{(0,\infty)}(u_1) I_{(0,\infty)}(u_n) \prod_{i=2}^{n-2} I_{(0,1)}(u_i) \ .$$

Integrating out u_1 and u_n gives for the joint density of $u_2, \ldots, u_{n-1}$,

$$h_1(u_2, \ldots, u_{n-1}) \propto \left(1 + \sum_{i=2}^{n-1} u_i\right)^{-(n-1)} \prod_{i=2}^{n-1} I_{(0,1)}(u_i) \ . \quad (5.3)$$

The rhs of (5.3) expressed in terms of $z_1, \ldots, z_n$ is

$$\left\{(z_n - z_1) / \sum_{i=1}^{n} (z_i - z_1)\right\}^{n-1} , \quad (5.4)$$

or, in terms of the original x's this is

$$\left\{(x_{(n)} - x_{(1)}) / \sum_{i=1}^{n} (x_i - x_{(1)})\right\}^{n-1} . \quad (5.5)$$

By Theorem 4.1 the hypothesis of uniformity is rejected in favor of the exponential alternatives if the quantity in (5.5) exceeds a constant, or, equivalently, if

$$T_{e,u} \equiv (\bar{x} - x_{(1)}) / (x_{(n)} - x_{(1)}) < c , \quad (5.6)$$

where $P(T_{e,u} < c | \mathcal{P}_o) = \alpha$.

Further, by Theorem 4.3 the UMPS-α test for testing $\mathcal{P}_1$ vs $\mathcal{P}_o$, i.e., exponential against uniform, is given by rejecting if

$$T_{e,u} > c \ , \tag{5.7}$$

where

$$P(T_{e,u} > c | \mathcal{P}_1) = \alpha \ .$$

Example 5.2 Exponential vs lognormal

The problem of deciding whether data is exponentially or lognormally distributed arises in the study of survival times of microorganisms which have been exposed to a disinfectant or poison (cf. Irwin (1942)). Cox (1961, 1962) developed an MLR test for this problem and gave the asymptotic distribution of the test statistic. For this problem $\mathcal{P}_0$ is the family of distributions with densities

$$\lambda \exp(-\lambda x) I_{(0,\infty)}(x), \ \lambda > 0 \ , \tag{5.8}$$

and $\mathcal{P}_1$ is the family with densities,

$$(x\sigma\sqrt{2\pi})^{-1} \exp\{-(\ln x)^2/2\sigma\} I_{(0,\infty)}(x) \ , \ \sigma > 0 \ . \tag{5.9}$$

We shall consider two cases. First, we consider testing $\mathcal{P}_0$ of (5.8) against a particular member of $\mathcal{P}_1$ of (5.9), i.e., the case for σ known. For the second case we consider testing $\mathcal{P}_0$ vs $\mathcal{P}_1$ with σ unknown. Since $\mathcal{P}_0$ and $\mathcal{P}_1$ are not conformable, the test obtained will not be UMPS or UMPI for this second case.

Case I, σ known

Srinivasan (1970) has studied the power of two Kolmogorov-Smirnov type goodness-of-fit statistics for this problem for a number of values of σ . See also, Schafer, Finkelstein, and Collins (1972) for corrections to Srinivasan's results.

By Theorem 4.2, the MPS and MPI tests are equivalent here, and we obtain the MPI test by applying a lemma of Hajek and Sidak (1967, p. 49), which says that the MPI test rejects for large values of the statistic

$$\int_0^\infty v^{n-1} f_1(vx_1, \ldots, vx_n)dv / \int_0^\infty v^{n-1} f_0(vx_1, \ldots, vx_n)dv \; , \tag{5.10}$$

where

$$f_o(y_1, \ldots, y_n) = \prod_{i=1}^{n} \exp(-y_i) I_{(0,\infty)}(y_i) \; ,$$

and

$$f_1(y_1, \ldots, y_n) = \left(\frac{1}{\sigma}\sqrt{\frac{2}{\pi}}\right)^n \prod_{i=1}^{n} y_i^{-1} \exp\{-(\ln y_i)^2/2\sigma\} I_{(0,\infty)}(y_i) \; .$$

This same formula can also be obtained from Lehmann (1959). The denominator of the ratio in (5.10) is

$$(n-1)!\left(\sum_{i=1}^{n} x_i\right)^{-n} . \tag{5.11}$$

The numerator of the ratio in (5.10) is proportional to

$$\left(\prod_{i=1}^{n} x_i\right)^{-1} \int_0^\infty v^{-1} \exp\{-\sum_{i=1}^{n} (\ln x_i + \ln v)^2/2\sigma^2\} dv \tag{5.12}$$

$$= \left(\prod_{i=1}^{n} x_i\right)^{-1} \int_{-\infty}^\infty \exp\{-\sum_{i=1}^{n} (t + \ln x_i)^2/2\sigma^2\} dt, \text{ letting } t = \ln v,$$

$$= (2\pi\sigma/n)^n \left(\prod_{i=1}^{n} x_i\right)^{-1} \exp\{-[\sum_{i=1}^{n} \ln^2 x_i - (\sum_{i=1}^{n} \ln x_i)^2/n]/2\sigma^2\} \; .$$

Therefore the test reduces to rejecting exponentiality if

$$(\sum_{i=1}^{n} x_i)^n (\prod_{i=1}^{n} x_i)^{-1} \exp \{-[\sum_{i=1}^{n} \ln^2 x_i - (\sum_{i=1}^{n} \ln x_i)^2/n]/2\sigma^2\} > c(\sigma) ,$$

or, equivalently, if

$$T_{e,L}(\sigma) \equiv \ln(\sum_{i=1}^{n} x_i) - (\sum_{i=1}^{n} \ln x_i)/n$$

$$- [\sum_{i=1}^{n} \ln^2 x_i - (\sum_{i=1}^{n} \ln x_i)^2/n]/2\,\sigma^2 > c(\sigma) , \qquad (5.13)$$

where $P(T_{e,L}(\sigma) > c(\sigma) | \mathcal{P}_o) = \alpha$. It is readily shown that this test is equivalent to the RML test of Cox (1961).

We have computed by Monte Carlo methods some significance points for the statistic $T_{e,L}(\sigma)$ for n = 10, 20, and for α = .01, .05, .10 . For these points we have further computed the power of the test for seven values of σ . These values are given in Table 5.1. All values were obtained by simulation using 5000 samples.

TABLE 5.1

Critical Values and Power of $T_{e,L}(\sigma)$ Test For Discriminating Between Exponential and Lognormal Distributions

H: Exponential $(0,\lambda)$ K: Lognormal $(0,\sigma^2)$, σ known

Reject H if $T_{e,L}(\sigma) > c(\sigma)$

		α=.01		α=.05		α=.10	
n	σ	$c(\sigma)$	Power	$c(\sigma)$	Power	$c(\sigma)$	Power
10	0.4	1.63	.93	1.16	1.00	.84	1.00
	0.6	2.07	.38	1.91	.80	1.80	.92
	0.8	2.24	.08	2.18	.36	2.14	.53
	1.0	2.38	.08	2.32	.21	2.30	.35
	1.4	2.75	.24	2.63	.40	2.58	.53
	2.0	3.07	.64	2.91	.78	2.84	.85
	2.4	3.20	.81	3.02	.90	2.93	.94
20	0.4	1.63	1.00	1.16	1.00	.83	1.00
	0.6	2.53	.93	2.34	1.00	2.23	1.00
	0.8	2.84	.43	2.78	.76	2.73	.89
	1.0	3.02	.23	2.99	.45	2.96	.63
	1.4	3.36	.48	3.28	.68	3.25	.77
	2.0	3.65	.91	3.54	.96	3.49	.98
	2.4	3.75	.98	3.63	.99	3.58	1.00

The second digit after the decimal is in some cases in doubt.

The purpose of this table is to allow comparisons of other tests with this best test. The same values of σ have been used as were used by Srinivasan (1970), and by Schafer, Finkelstein, and Collins (1972). The powers of Table 5.1 are a least upper bound for any test for these hypotheses. It should be pointed out, however, that it is not reasonable to expect a general goodness-of-fit test such as that of Srinivasan to compete well with the MPS test, because the goodness-of-fit test does not assume knowledge of the alternative, but is, presumably, effective against a broader range of alternatives.

Case II, σ unknown

The families of distributions with densities given by (5.8) and (5.9) are not conformable. The exponential family is a scale parameter family, i.e., the transfromations $\{g: \ g(x) = ax, \ a > 0\}$ induce the transitive group $\bar{G} = \{\bar{g}: \ \bar{g}(\sigma) = a\sigma, \ a > 0\}$ on the parameter space. However, the transformations $\{g: \ g(x) = x^a, \ a > 0\}$ induce this same transitive group $\bar{G}$ for the lognormal family. Therefore the forgoing theory gives no UMPS test, and we shall construct a (nonoptimal) test by replacing σ^2 in (5.13) by a sample estimator.

We shall use the estimator

$$\hat{\sigma}^2 = \{\sum_{i=1}^{n} (\ln x_i)^2 - (\sum_{i=1}^{n} \ln x_i)^2/n\}/(n-1) \ , \tag{5.15}$$

which has two important properties. It is invariant under scale transformations, and it is unbiased for σ^2 under the lognormal alternatives. Replacing σ^2 in (5.13) by $\hat{\sigma}^2$, a test for $\mathcal{P}_0$ of (5.8) against $\mathcal{P}_1$ of (5.9) is given by rejecting if

$$T^*_{e,L} \equiv \ln(\sum_{i=1}^{n} x_i) - (\sum_{i=1}^{n} \ln x_i)/n$$

$$-(n-1)\ln\{\sum_{i=1}^{n} (\ln x_i)^2 - (\sum_{i=1}^{n} \ln x_i)^2/n\}/2n > c \ , \tag{5.16}$$

where $P\{T^*_{e,L} > c|\mathcal{P}_0\} = \alpha$.

We give in Table 5.2 some significance points for the statistic $T^*_{e,L}$ for $n = 10$, 20, and $\alpha = .01, .05, .10$. The power of this test has also been computed for the same values of σ as were given in Table 5.1, and these values are also given in Table 5.2. The values in Table 5.2 are all based on 5,000 samples. In some cases the second digit after the decimal is doubtful.

TABLE 5.2

Critical Values and Power of $T^*_{e,L}$ for Discriminating

Between Exponential and Lognormal Distributions

H: Exponential $(0,\lambda)$ K: Lognormal $(0,\sigma^2)$, σ unknown

Reject H if $T^*_{e,L} > c$

n	c	σ	α=.01 Power	α=.05 Power	α=.10 Power
10	α = .01	0.4	.99	.99	1.00
	c = 2.02	0.6	.34	.69	.83
		0.8	.08	.26	.43
	α = .05	1.0	.04	.13	.26
	c = 1.90	1.4	.17	.29	.39
	α = .10	2.0	.55	.67	.75
	c = 1.84	2.4	.75	.83	.87
20	α = .01	0.4	1.00	1.00	1.00
	c = 2.17	0.6	.85	.98	.99
		0.8	.27	.63	.78
	α = .05	1.0	.12	.35	.52
	c = 2.10	1.4	.40	.55	.66
	α = .10	2.0	.86	.91	.93
	c = 2.07	2.4	.97	.98	.98

It is suggested that the power results given in Tables 5.1 and 5.2 be compared with those given by Schafer, et al (1972) for the statistics posed by Srinivasan and Lilliefors. Such comparison reveals that the power of the $T_{e,L}$ test is largest in all cases, as it must be, and that the power of the $T^*_{e,L}$ statistic is generally intermediate between that of $T_{e,L}$ and the best of the other two statistics. It is hoped that this case is representative of a larger class of situations, and that tests with high power can be obtained for other separate families problems which are not conformable, by substituting invariant estimators for nuisance parameters in pointwise optimal tests.

REFERENCES

Antle, C., R. Dumonceaux, and G. Haas. (1973). Likelihood ratio test for discrimination between two models with unknown location and scale parameters. Technometrics 15:19-28.

Atkinson, A. C. (1970). A method of discriminating between models. J. Royal Stat. Soc., Series B, 32:323-345.

Basu, D. (1955). On statistics independent of a complete sufficient statistic. Sankhya 15:337-380 and 20:223-226.

Chen, E. H. (1971). Random normal number generator for 32-bit-word computers. J. Amer. Stat. Assoc. 66:400-403.

Cox, D. R. (1961). Tests of separate families of hypotheses. Proceedings of the Fourth Berkely Symposium, Vol. 1, University of California Press, Berkeley, Calif. 105-23.

Cox, D. R. (1962). Further results on tests of separate families of hypotheses. J. Royal Stat. Soc., Ser. B, 24:406-424.

Dumonceaux, R. and C. Antle. (1973). Discrimination between the Log-Normal and the Weibull Distributions. Technometrics 15:923-926.

Dyer, A. R. (1971). A comparison of classification and hypothesis testing procedures for choosing between competing families of distributions, including a survey of the goodness of fit tests. Technical Memorandum No. 18, Aberdeen Research and Development Center, Aberdeen Proving Ground, Maryland.

Dyer, A. R. (1973). Discrimination procedures for separate families of hypotheses. J. Amer. Stat. Assoc. 68:970-974.

Dyer, A. R. (1974). Hypothesis testing procedures for separate families of hypotheses. J. Amer. Stat. Assoc. 69:140-145.

Geary, R. C. and E. S. Pearson. (1955). The ratio of the mean deviation to the standard deviation as a test of normality. Biometrika 27:310-335.

Hajek, J. and F. Sidak. (1967). Theory of Rank Tests. New York: Academic Press.

Irwin, J. O. (1942). The distribution of the logarithm of survival times when the true law is exponential. J. of Hygiene 42:328-333.

Jackson, O. A. Y. (1968). Some results on tests of separate families of hypotheses. Biometrika 55:355-363.

Lehmann, E. L. (1950). Some principles of the theory of testing hypotheses. Annals of Mathematical Statistics 21:1-26.

Lehmann, E. L. (1959). Testing Statistical Hypotheses. John Wiley and Sons, Inc., New York City, New York.

Moore, D. (1973). A note on Srinivasan's goodness fo fit test. Biometrika 60:209-211.

O'Reilly, F. J. and C. P. Quesenberry. (1973). The conditional probability integral transformation and applications to obtain composite chi-square goodness-of-fit tests. Annals of Statistics 1:74-83.

Schafer, R., J. Finkelstein, and J. Collins. (1972). On a goodness of fit test for the exponential distribution with mean unknown. Biometrika 59:222-224.

Srinivasan, R. (1970). An approach to testing the goodness of fit of incompletely specified distributions. Biometrika 57:605-611.

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. J. Amer. Stat. Assoc. 69:730-737.

Uthoff, V. A. (1970). An optimum test property of two well-known statistics. J. Amer. Stat. Assoc. 65:1597-1600.

Uthoff, V. A. (1973). The most powerful scale and location invariant test of the normal versus the double exponential. Annals of Statistics 1:170-174.