The Annals of Statistics
1973, Vol. 1, No. 1, 74-83

THE CONDITIONAL PROBABILITY INTEGRAL TRANSFORMATION AND APPLICATIONS TO OBTAIN COMPOSITE CHI-SQUARE GOODNESS-OF-FIT TESTS[1]

BY FEDERICO J. O'REILLY[2] AND C. P. QUESENBERRY

North Carolina State University

It is shown that certain conditional distributions, obtained by conditioning on a sufficient statistic, can be used to transform a set of random variables into a smaller set of random variables that are identically and independently distributed with uniform distributions on the interval from zero to one. This result is then used to construct distribution-free tests of fit for composite goodness-of-fit problems. In particular, distribution-free chi-square goodness-of-fit tests are obtained for univariate normal, exponential, and normal linear regression model families of distributions.

1. Introduction and summary. Let $X_1, \cdots, X_n$ denote a set of independent random variables that are identically distributed with distribution P, and corresponding distribution function F, which is a member of a class $\mathscr{P}$ of univariate distributions with absolutely continuous distribution functions. In Section 2 we establish results which show how, in some cases, certain conditional distributions obtained by conditioning on sufficient statistics can be used to transform the sample into a set of independently and identically distributed uniform random variables on the interval (0, 1)-i.i.d. $U(0, 1)$.

In Section 3 we use the foregoing results to construct chi-square type distribution-free tests of fit for composite problems. In Section 4 we consider some particular examples and set out the details of the tests. In particular, we consider tests for the two-parameter univariate normal family, the scale parameter exponential family, and the standard univariate normal regression model. In Section 5 we consider generalizations to multivariate distributions, and in Section 6 give a brief discussion of the foregoing results.

The major theorem of Section 2 relies upon the multivariate probability integral transformation of Rosenblatt [11]. Lilliefors [8], [9] has proposed a Kolmogorov–Smirnov type statistic for testing composite null hypothesis classes that can be characterized by location and scale parameters; that the statistic is distribution-free follows from results of David and Johnson [4]. He also simulated the distribution of this statistic for normal and exponential classes. Watson

Received November 17, 1971; revised June 20, 1972.

[1] This paper is adapted from the first author's doctoral dissertation written under the direction of Professor C. P. Quesenberry and approved by the Graduate Faculty of North Carolina State University in August 1971.

[2] Now at CIMASS, Universidad Nacional Autónoma de México.

AMS 1970 *subject classifications*. Primary 62; Secondary 71.

Key words and phrases. Conditional expectation, minimal sufficient statistic, absolute continuity, MVU function estimator, composite goodness-of-fit tests.

[14], [15] has shown how to construct a chi-square type statistic which is asymptotically distribution-free. There is a rather large literature on minimum variance unbiased estimation which is of general relevance here. We shall make specific reference to papers by Lieberman and Resnikoff [7], Ghurye and Olkin [5], and Sathe and Varde [12].

2. The conditional probability integral transformation (CPIT). Let $(R^n, \mathscr{B}^n, P^n)$ denote the probability space of the independent sample $X_1, \cdots, X_n$ from the Borel parent space $(R, \mathscr{B}, \mathscr{P})$. Let $X_i: R^n \to R$ denote the ith projection of R^n and denote by $\sigma(X_i)$ the sub σ-algebra of $\mathscr{B}^n$ induced by X_i; $i = 1, \cdots, n$.

Let $T_n: R^n \to R^{k_n} (k_n \in \{1, \cdots, n\})$ be a sufficient statistic for $\mathscr{P}$, and denote by $\tilde{F}_n(x_1, \cdots, x_\alpha)$ the Rao–Blackwell distribution function estimator:

$$E\{I_{[X_1 \leqq x_1, \cdots, X_\alpha \leqq x_\alpha]} \mid T_n\}, \qquad \alpha \in \{1, \cdots, n\},$$

where here I denotes the indicator function, i.e., I_A is one when $(X_1, \cdots, X_n) \in A$ for $A \in \mathscr{B}^n$, and is otherwise zero.

THEOREM 2.1. *If $\tilde{F}_n(x)$ is absolutely continuous* a.s., *then the* rv $\tilde{F}_n(X_i)$ *is uniformly distributed on the unit interval for $i = 1, \cdots, n$.*

PROOF. For a given $T_n = t$, $\tilde{F}_n(x_i)$ is the conditional distribution of X_i given T_n evaluated at t, which is absolutely continuous (except for a set of t values of zero probability). Therefore

$$P(\tilde{F}_n(X_i) \leqq y \mid T_n = t) = y \quad \text{a.s.}$$

by the probability integral transformation theorem. Observe that this conditional distribution does not depend on t a.s. The result follows.

In the sequel it will be assumed that T_n is symmetric in $X_1, \cdots, X_n$. Thus the ordering of the observations is immaterial in the next definition.

DEFINITION 2.1. The maximum value of $\alpha = 1, \cdots, n$ for which $\tilde{F}_n(x_1, \cdots, x_\alpha)$ is absolutely continuous a.e. $\mathscr{P}$ will be called the *absolute continuity rank of* $\mathscr{P}$Re T_n-a.c.r. $\mathscr{P}$Re T_n.

THEOREM 2.2. *The absolute continuity rank of* $\mathscr{P}$Re *the minimal sufficient statistic Z_n is not less than that given any sufficient statistic T_n.*

PROOF. Seheult and Quesenberry [13] have shown, essentially, that a necessary and sufficient condition for the existence of an unbiased $\sigma(T_n)$ measurable estimator $\tilde{f}(x_1, \cdots, x_\alpha)$ of the density function of $(X_1, \cdots, X_\alpha)$ is that $\tilde{F}_n(x_1, \cdots, x_\alpha)$ be absolutely continuous. But then the conditional expectation of $\tilde{f}(x_1, \cdots, x_\alpha)$ can be taken given Z_n and an unbiased $\sigma(Z_n)$ measurable estimator of the density results. Therefore, $\tilde{F}_n(x_1, \cdots, x_\alpha)$ obtained from Z_n is absolutely continuous.

In view of Theorem 2.2, we shall call the a.c.r. $\mathscr{P}$Re Z_n, the minimal sufficient statistic, simply the a.c.r. $\mathscr{P}$. The quantity a.c.r. $\mathscr{P}$ is, in general, a function of n. For many families $\mathscr{P}$ it is of the form $n - c$, for c a fixed integer. This is the case if Z_n is a vector with c components for an exponential class $\mathscr{P}$ of

continuous distributions. If the class $\mathscr{P}$ is not dominated by Lebesgue measure, then the a.c.r. $\mathscr{P}$ is zero for all sample sizes (making the convention that for $\alpha = 0$, $E\{I_{[X_1 \leqq x_1, \cdots, X_\alpha \leqq x_\alpha]} \mid T_n\}$ is an absolutely continuous df).

Next, let $\tilde{F}_n(x_j \mid x_1, \cdots, x_{j-1})$ be given by:

$$\tilde{F}_n(x_j \mid x_1, \cdots, x_{j-1}) = E(E(I_{[X_j \leqq x_j]} \mid T_n = t) \mid X_1 = x_1, \cdots, X_{j-1} = x_{j-1}) .$$

We shall use T in place of T_n in the next lemma and proof.

LEMMA 2.1. $\tilde{F}_n(x_j \mid x_1, \cdots, x_{j-1})$

$$= E(I_{[X_j \leqq x_j]} \mid T = t, X_1 = x_1, \cdots, X_{j-1} = x_{j-1}) \quad \text{a.s.}$$

PROOF. We will show only for $j = 2$. The general case is done similarly. Denote by $\tilde{F}(x_1)$ the marginal df of $\tilde{F}(x_1, x_2)$ computed for each $T = t$, i.e.,

$$\tilde{F}(x_1) = \lim_{x_2 \to \infty} \tilde{F}(x_1, x_2) = \lim_{x_2 \to \infty} E(I_{[X_1 \leqq x_1, X_2 \leqq x_2]} \mid T) .$$

By the monotone convergence theorem for conditional expectations the last term in the above expression is equal to $E(I_{[X_1 \leqq x_1]} \mid T)$ a.s.

Therefore, $\tilde{F}(x_1) = E(I_{[X_1 \leqq x_1]} \mid T)$ a.s. From its definition, for each value of T, $\tilde{F}(x_2 \mid x_1)$ obeys the relationship:

$$\tilde{F}(x_1, x_2) = \int_{-\infty}^{x_1} \tilde{F}(x_2 \mid u)\, d\tilde{F}(u) \quad \text{a.s.} \tag{2.1}$$

Now, $\tilde{F}(x_1, x_2)$, being a conditional expectation, must obey, for every $B_T \in \sigma(T)$, the relationship:

$$\int_{B_T} \tilde{F}(x_1, x_2)\, dP_T = P([X_1 \leqq x_1, X_2 \leqq x_2] B_T) . \tag{2.2}$$

Also, $E(I_{[X_2 \leqq x_2]} \mid T, X_1)$ obeys, for every $B_T \in \sigma(T)$,

$$\int_{B_T} \int_{-\infty}^{x_1} E(I_{[X_2 \leqq x_2]} \mid T, X_1)\, dP_{T, X_1} = P([X_1 \leqq x_1, X_2 \leqq x_2] B_T) . \tag{2.3}$$

From (2.1), (2.2), (2.3), and the fact that $dP_{T, X_1} = dP^T_{X_1}\, dP_T$, where $P^T_{X_1}(-\infty, u) = \tilde{F}(u)$; we have it that for every $B_T \in \sigma(T)$, and any $x_1 \in R$,

$$\int_{B_T} \int_{-\infty}^{x_1} \tilde{F}(x_2 \mid u)\, dP^T_{X_1}(u)\, dP_T = \int_{B_T} \int_{-\infty}^{x_1} E(I_{[X_2 \leqq x_2]} \mid T, X_1 = u)\, dP^T_{X_1}(u)\, dP_T ,$$

therefore

$$\tilde{F}(x_2 \mid x_1) = E(I_{[X_2 \leqq x_2]} \mid T, X_1 = x_1) \quad \text{a.s.}$$

THEOREM 2.3. *If the* a.c.r. $\mathscr{P}$ Re T_n' *is* $\alpha(> 0)$, *then*

$$\tilde{F}_n(X_1), \tilde{F}_n(X_2 \mid X_1), \cdots, \tilde{F}_n(X_\alpha \mid X_1, \cdots, X_{\alpha-1})$$

are i.i.d. $U(0, 1)$.

PROOF. By hypothesis, $\tilde{F}_n(x_1, \cdots, x_\alpha)$ is an absolutely continuous distribution function a.s. The result follows by applying the multivariate probability integral transformation given by Rosenblatt [12].

DEFINITION 2.2. The sequence of statistics $(T_n)_{n \geqq 1}$ is said to be *doubly transitive* if for each $n \geqq 1$; $\sigma(T_n, X_{n+1}) = \sigma(T_{n+1}, X_{n+1})$.

Berk [2] and Bahadur [1] have considered transitive sequences. When the

sequence $(T_n)_{n\geq 1}$ is doubly transitive the transformations in Theorem 2.3 can be simplified.

THEOREM 2.4. *If the* a.c.r. $\mathscr{P}$ Re T_n *is* $\alpha(>0)$ *and* $(T_n)_{n\geq 1}$ *is doubly transitive, then*

$$\begin{aligned}
&\tilde{F}_n(x_{n-1}\,|\,x_n) = \tilde{F}_{n-1}(x_{n-1}) \quad \text{a.s.}\\
&\tilde{F}_n(x_{n-2}\,|\,x_n, x_{n-1}) = \tilde{F}_{n-2}(x_{n-2}) \quad \text{a.s.}\\
&\qquad\vdots\\
&\tilde{F}_n(x_{n-\alpha}\,|\,x_n, \cdots, x_{n-\alpha+1}) = \tilde{F}_{n-\alpha}(x_{n-\alpha}) \quad \text{a.s.}
\end{aligned}$$

PROOF. By Lemma 2.1

$$\tilde{F}_j(x_j\,|\,x_n, \cdots, x_{j+1}) = E(I_{[X_j\leq x_j]}\,|\,T_n, X_n = x_n, \cdots, X_{j+1} = x_{j+1}) \quad \text{a.s.}$$

for $j = n - \alpha + 1, \cdots, n$. Since X_j is independent of $X_{j+1}, \cdots, X_n$ and double transitivity of $(T_n)_{n\geq 1}$ means that $\sigma(T_n, X_n) = \sigma(T_{n-1}, X_n)$, it follows that the r.h.s. of the last equation is equal to

$$E(I_{[X_j\leq x_j]}\,|\,T_j) = \tilde{F}_j(x_j) \quad \text{a.s.}$$

for $j = n - \alpha + 1, \cdots, n$.

COROLLARY 2.1. *If the* a.c.r. $\mathscr{P}$Re T_n *is* $\alpha(>0)$ *and* $(T_n)_{n\geq 1}$ *is doubly transitive then*

$$\tilde{F}_{n-\alpha+1}(X_{n-\alpha+1}), \cdots, \tilde{F}_n(X_n)$$

is a set of α i.i.d. $U(0, 1)$ rv's.

From Theorem 2.2 it follows that the maximum number of i.i.d. $U(0, 1)$ rv's is obtained in Corollary 2.1 when T_n is the minimal sufficient statistic.

The approach of this section can be used to transform a set $X_1, \cdots, X_n$ of rv's when the assumptions of identical distributions and independence of components is not fulfilled. The following theorem clearly holds.

THEOREM 2.5. *For any set* $X_1, \cdots, X_n$ *of real-valued* rv's, *if* $\tilde{F}_n(x_1, \cdots, x_\alpha)$ *is the* df *corresponding to a Rao–Blackwell estimating distribution* $\tilde{P}_n$ *that is dominated by* λ^α *(*α*-dimensional Lebesgue measure) for* $\alpha > 0$, *then the* α rv's

$$\tilde{F}_n(X_1), \tilde{F}_n(X_2\,|\,X_1), \cdots, \tilde{F}_n(X_\alpha\,|\,X, \cdots, X_{\alpha-1})$$

are i.i.d. $U(0, 1)$.

In many cases when some of the conditions of identical distributions, independence, transitivity, etc., are satisfied it is possible to write particularized versions of Theorem 2.5 which may be more useful or convenient. In Example 4.3 we shall use a version for which the identical distributions condition is not satisfied.

3. Distribution-free chi-square goodness-of-fit tests for composite hypotheses. In this section the results of Section 2 are applied to obtain tests for hypotheses of

the form H: $P \in \mathscr{P}$ vs. K: $P \notin \mathscr{P}$. The general strategy is to apply either Theorem 2.3 or Corollary 2.1 to obtain a set of α i.i.d. $U(0, 1)$ rv's under H. Under K this set of rv's will, in general, not be i.i.d. $U(0, 1)$. Any statistic which measures distance from uniformity in the transformed sample can be used as a test statistic. If a Kolmogorov–Smirnov statistic is used for the α transformed rv's, its distribution will be distribution-free (i.e., the same for any $P \in \mathscr{P}$); and, moreover, will be that given by Birnbaum [3].

If a chi-square type test statistic is considered, the statistic has an exact X^2 distribution (see (3.4) below), and, moreover, a significant additional advantage is that it turns out that the test statistic can be identified with a chi-square type statistic computed from the beginning sample, i.e., without actually performing the transformations. In the sequel it will be assumed that T_n is the minimal sufficient statistic. If any other sufficient statistic is used the result will sometimes be to waste some observations.

THEOREM 3.1. *Let the* a.c.r. $\mathscr{P}$ *be* $\alpha(> 0)$, *and let* $p_1, \cdots, p_k$ *be fixed positive numbers with* $p_1 + \cdots + p_k = 1$. *Then there exists* α *random partitions of* R:

$$\{(\tilde{C}_{(i-1)j}, \tilde{C}_{ij}]\}_{i=1}^{k}, \quad \textit{where} \tag{3.1}$$
$$-\infty = \tilde{C}_{0j} < \tilde{C}_{1j} < \cdots < \tilde{C}_{kj} = +\infty,$$

such that for $j = n - \alpha + 1, \cdots, n$ *the random vector* $(N_1, N_2, \cdots, N_k)$ *defined by*

$$N_i = \sum_{j=n-\alpha+1}^{n} I_{[\tilde{C}_{(i-1)j} < X_j \leqq \tilde{C}_{ij}]}, \qquad i = 1, \cdots, k$$

has a multinomial distribution with probability parameters $p_1, \cdots, p_k$; $\sum_{i=1}^{k} N_i = \alpha$.

PROOF. Apply Theorem 2.3, taking the transformations in reverse order. For each $j = n - \alpha + 1, \cdots, n$ let $Y_j = \tilde{F}_n(X_j \mid X_n, \cdots, X_{j+1})$, and define $\tilde{C}_{ij}$ by:

$$\tilde{F}_n(\tilde{C}_{ij} \mid X_n, \cdots, X_{j+1}) = \sum_{j=1}^{i} p_j; \qquad i = 1, \cdots, k-1. \tag{3.2}$$

Then

$$(\textstyle\sum_{j=1}^{i-1} p_j < Y_j \leqq \sum_{j=1}^{i} p_j) \quad \text{iff} \quad (\tilde{C}_{(i-1)j} < X_j \leqq \tilde{C}_{ij}). \tag{3.3}$$

The result follows since $Y_{n-\alpha+1}, \cdots, Y_n$ are i.i.d. $U(0, 1)$.

If we put

$$X^2 = \sum_{i=1}^{k} (N_i - \alpha p_i)^2 / \alpha p_i, \tag{3.4}$$

then this has the distribution of the X^2 statistic often used for testing simple hypotheses.

COROLLARY 3.1. *If the minimal sufficient statistic in Theorem* 3.1 *is doubly transitive, then the* X^2 *statistic of* (3.4) *converges in distribution to a* $\chi^2(k-1)$ rv *as* $n \to \infty$.

We advocate choosing $p_1 = \cdots = p_k = 1/k$ as has been proposed for the simple problem by Mann and Wald [10], and more recently by Good, *et al.* [6],

so that

$$X^2 = (k/\alpha) \sum_{i=1}^{k} N_i^2 - \alpha \,. \tag{3.5}$$

Tables that are useful here have been given by Good, *et al.* [6], and by Zahn and Roberts [16].

4. Examples. The results of the foregoing sections will here be applied to some problems of practical importance.

EXAMPLE 4.1. *Test of normality.* Let $\mathscr{P}$ be the class of univarite normal distributions with unknown parameters μ and σ^2. The minimal sufficient statistic is $T_n = (\bar{X}_n, S_n^2) = (\sum_1^n X_i/n, \sum_1^n (X_i - \bar{X})^2/n)$, and is clearly doubly transitive. By Theorem 2.4 we can use the functions $\tilde{F}_r(x_r)$ to construct the $\tilde{C}_{ij}$'s of Theorem 3.1. The estimator $\tilde{F}_r(z)$ is given (cf. Lieberman and Resnikoff [7]) for $r > 2$ by:

$$\begin{aligned}
\tilde{F}_r(z) &= 0 \,, \qquad \text{if} \quad z - \bar{X}_r < -(r-1)^{\frac{1}{2}} S_r \,, \\
&= 1 \,, \qquad \text{if} \quad (r-1)^{\frac{1}{2}} S_r < z - \bar{X}_r \,, \qquad\qquad \text{and} \\
&= G_{r-2}\{(r-2)^{\frac{1}{2}}(z - \bar{X}_r)/[(r-1)S_r^2 - (z - \bar{X}_r)^2]^{\frac{1}{2}}\} \,, \qquad \text{elsewhere,}
\end{aligned}$$

where G_{r-2} is the Student-t distribution with $r - 2$ df. For $r = 1, 2$, $\tilde{F}_r(z)$ is not continuous.

Since $\tilde{F}_r(z)$ is absolutely continuous for $r \geqq 3$, then for $j = n - \alpha + 1, \cdots, n$, using the double transitivity property, the transformation for x_j is given by $\tilde{F}_j(x_j)$. Thus, $\alpha = n - 2$.

Let $t(i/k, j-2)$ be the (i/k)th quantile of a Student-t distribution with $j - 2$ df; $j = 3, \cdots, n$. Obviously, this selection is for equiprobable cells.

The estimated quantiles $\tilde{C}_{ij}$, $i = 1, \cdots, k-1$, for the jth variable are given by the solutions to the equations:

$$(j-2)^{\frac{1}{2}}(\tilde{C}_{ij} - \bar{X}_j)/[(j-1)S_j^2 - (\tilde{C}_{ij} - \bar{X}_j)^2]^{\frac{1}{2}} = t(i/k, j-2) \,,$$

which gives:

$$\tilde{C}_{ij} = \frac{((j-1)/(j-2))^{\frac{1}{2}} t(i/k, j-2) S_j}{(1 + t^2(i/k, j-2)/(j-2))^{\frac{1}{2}}} + \bar{X}_j \,. \tag{4.1}$$

Note the sequential nature of the computational procedure, that $\bar{X}_3$, S_3^2 are computed first, then $\bar{X}_4$, S_4^2, etc.; and, also, if the cells are computed for a sample size n_0, and additional observations are then obtained, the formulas for the first cells are unaltered.

The case of testing normality when σ^2 is known can be obtained in a straightforward manner. The expression for $\tilde{F}_r(z)$ in this case is the df of a $N(\bar{X}_r, ((r-1)/r)\sigma^2)$ distribution.

EXAMPLE 4.2. *Test of exponentiality.* It is desired to test the composite null hypothesis that the parent has density $\lambda e^{-\lambda t}$, $\lambda > 0$. The sample mean $\bar{X}_n$ is the minimal sufficient statistic and is clearly doubly transitive. The Rao–Blackwell estimating distribution function is:

$$\tilde{F}_r(z) = 1 - [1 - zI_{(0, r\bar{X}_r)}(z)/(r\bar{X}_r)] + I_{[r\bar{X}_r, \infty)}(z) \,.$$

The absolute continuity rank is thus $n - 1$ and in this case no tables are needed to select the estimated quantiles, $\tilde{C}_{ij}$, $i = 1, \cdots, k - 1$; $j = 2, \cdots, n$.

The $\tilde{C}_{ij}$, are obtained by solving:

$$1 - [1 - \tilde{C}_{ij}/(j\bar{X}_j)]^{j-1} = i/k$$

which yields:

$$\tilde{C}_{ij} = \left[1 - \left(1 - \frac{i}{k}\right)^{1/(j-1)}\right] j\bar{X}_j \,. \tag{4.2}$$

Following the same steps as in the above examples, the random selection of cells can be easily obtained for the following cases given by Sathe and Varde [12]:

Incomplete Gamma; $\dfrac{\theta^{-p}}{\Gamma(p)} e^{-x/\theta} x^{p-1}$ for p known.

Weibull; $(p/\theta)x^{p-1}e^{-x^p/\theta}$, p known.

In both cases, the absolute continuity rank is $n - 1$, and in both cases, also, the corresponding minimal sufficient statistic is doubly transitive.

EXAMPLE 4.3. *Testing the fit of a normal regression model.* For $\mathbf{y}_n' = (y_1, \cdots, y_n)$ a vector rv consider testing the hypothesis:

$$H: \mathbf{y}_n \sim N(X_n\boldsymbol{\beta}, \sigma^2 I) \,, \tag{4.3}$$

where X_n is some $n \times p$ matrix ($n > p$) of full rank, and $(\boldsymbol{\beta}, \sigma^2)$ are $p + 1$ unknown parameters. This is the well-known univariate multiple regression model. In this section, a test is given for testing the hypothesis (4.3).

For the family $N(X_n\boldsymbol{\beta}, \sigma^2 I)$, clearly $(\mathbf{y}_n'\mathbf{y}_n, X_n'\mathbf{y}_n)$ is the minimal sufficient statistic, and is equivalent to the statistic

$$(X_n'\mathbf{y}_n, \mathbf{y}_n'(I - X_n(X_n'X_n)^{-1}X_n')\mathbf{y}_n) \,.$$

Since the family is defined with the knowledge of X_n, then for $\mathbf{x}_i'$ the ith row of X_n, the statistic $(\mathbf{y}_s'\mathbf{y}_s, X_s'\mathbf{y}_s, \mathbf{y}_{s+1})$ is known iff $(\mathbf{y}_{s+1}'\mathbf{y}_{s+1}, X_{s+1}'\mathbf{y}_{s+1}, y_{s+1})$ is known, which shows that the minimal sufficient statistic is doubly transitive.

Ghurye and Olkin ([5], page 1268, 4.2) give the MVUE of a related density which in the particular case considered here can be written as in the following. The conditional density of y_n given $\mathbf{t}_n = X_n'\mathbf{y}_n$ and $S_n^2 = \mathbf{y}_n'(I - X_n(X_n'X_n)^{-1}X_n')\mathbf{y}_n$, exists iff X_{n-1} is of full rank, and in this case it is given by the expression:

$$\frac{S_n^{-(n-p-2)}}{\beta(\frac{1}{2}, (n-p-1)/2)} (1 - \mathbf{x}_n'(X_n'X_n)^{-1}\mathbf{x}_n)^{-\frac{1}{2}} \times \left\{\Psi\left[S_n^2 - \frac{(y_n - \mathbf{x}_n'(X_n'X_n)^{-1}\mathbf{t}_n)^2}{(1 - \mathbf{x}_n'(X_n'X_n)^{-1}\mathbf{x}_n)}\right]\right\}^{(n-p-3)/2} ,$$

where $\Psi(z) = z$ if $z > 0$, and is otherwise zero. Therefore, for $j = p + 2, \cdots, n$, the conditional density of y_j given $\mathbf{t}_j = X_j'\mathbf{y}_j$, and $S_j^2 = \mathbf{y}_j'(I - X_j(X_j'X)^{-1}X_j')\mathbf{y}_j$ exists if the first $p + 1$ rows of the matrix X_n form a full rank matrix. It is assumed that this is the case. This result enables the writing of the conditional

distribution function of y_j given $\mathbf{t}_j$ and S_j^2 as a Student-t distribution function with $j - p - 1$ degrees freedom evaluated at:

$$U_j = \frac{(j-p-1)^{\frac{1}{2}}(y_j - \mathbf{x}_j'(X_j'X_j)^{-1}\mathbf{t}_j)}{\{(1 - \mathbf{x}_j'(X_j'X_j)^{-1}\mathbf{x}_j)S_j^2 - (y_j - \mathbf{x}_j'(X_j'X_j)^{-1}t_j)^2\}^{\frac{1}{2}}}.$$

Under the assumption that the ordering of the observations is given as above, then the conditional distribution of the jth observation ($j = p + 2, \cdots, n$) given S_j^2 and $\mathbf{t}_j$ is $G_{j-p-1}(U_j)$, for $U_j \in (-\infty, \infty)$, or, equivalently, for

$$|y_j - \mathbf{x}_j'(X_j'X_j)^{-1}\mathbf{t}_j| < \{(1 - \mathbf{x}_j'(X_j'X_j)^{-1}\mathbf{x}_j)S_j^2\}^{\frac{1}{2}}.$$

The formulas for the quantiles $\tilde{C}_{ij}$ (for k equiprobable cells); $i = 1, \cdots, k-1$; $j = p + 2, \cdots, n$; are given by

$$\tilde{C}_{ij} = \frac{\{S_j^2(1 - \mathbf{x}_j'(X_j'X_j)^{-1}\mathbf{x}_j)/(j-p-1)\}^{\frac{1}{2}}t(i/k, j-p-1)}{\{1 + t^2(i/k, j-p-1)/(j-p-1)\}^{\frac{1}{2}}} + \mathbf{x}_j'(X_j'X_j)^{-1}\mathbf{t}_j,$$

which clearly particularizes to the formulas given for the normal distribution of Example 1 by putting $p = 1$. See the remark at the end of Section 2.

5. Multivariate classes. The results obtained in earlier sections generalize to multivariate distributions in a natural manner. Indeed, this is a major appeal of the approach used in this paper.

Let $\mathbf{X}_1, \cdots, \mathbf{X}_n$ denote an independent sample of k-component vectors distributed according to $p \in \mathscr{P}$, and assume that the absolute continuity rank of $\mathscr{P}$ (here absolute continuity refers to domination by λ^k) is α (> 0). We can then obtain immediately multivariate analogues of Theorems 2.1, 2.3, 2.4, 2.5 and Corollary 2.1. The following generalizes Corollary 2.1 which is the most easily used in practice.

THEOREM 5.1. *If the above assumptions are met, and if the minimal sufficient statistic is doubly transitive, then the* $k \cdot \alpha$ rv's

$$\tilde{F}_j(X_{ij} \mid X_{1j}, \cdots, X_{(i-1)j}), \qquad i = 1, \cdots, k;\ j = n - \alpha + 1, \cdots, n;$$

are i.i.d. $U(0, 1)$, *where* $\tilde{F}_j(x_{ij} \mid x_{1j}, \cdots, x_{(i-1)j})$ *denotes the* df *of the conditional distribution of* X_{ij} *obtained from the distribution with* df $\tilde{F}_j(\mathbf{x}_j)$ *for fixed* $X_{1j} = x_{1j}, \cdots, X_{(i-1)j} = x_{(i-1)j}$.

PROOF. Consider $\tilde{F}_n(\mathbf{x}_{n-\alpha+1}, \cdots, \mathbf{x}_n)$, and note that for $j = n - \alpha + 1, \cdots, n$:

$$\tilde{F}_n(\mathbf{x}_j \mid \mathbf{x}_n, \cdots, \mathbf{x}_{j+1}) = \tilde{F}_j(x_j) \quad \text{a.s.},$$

which is the df of a distribution dominated by λ^k. Therefore each of these can be conditioned using the multivariate probability integral transformation to establish the theorem.

We give one example to illustrate this theorem.

EXAMPLE 5.1. Let $\mathbf{X}_1, \cdots, \mathbf{X}_n$ be a sample from the k-variate normal family $N(\boldsymbol{\mu}, V)$ where V *is known*. The minimal sufficient statistic is the vector of

sample means $\bar{\mathbf{X}}_n$, which is clearly doubly transitive. Ghurye and Olkin [5] give the conditional distribution of any one term in the sample given $\bar{\mathbf{X}}_n$, which distribution is $N(\bar{\mathbf{X}}_n, (1 - 1/n)V)$. The absolute continuity rank of the family is $n - 1$. For $j = 2, \cdots, n$ arbitrary, consider the sample mean $\bar{\mathbf{X}}_j$ computed from the first j observations, and let $\mathbf{x}_j' = (x_{1j}, \cdots, x_{kj})$ denote the jth observation vector.

From the conditional distribution $\tilde{F}_j(\mathbf{x}_j)$, which is the distribution $N(\bar{\mathbf{X}}_j, (1 - 1/j)V)$, evaluated at $\mathbf{x}_j$, the k conditional distributions

$$\tilde{F}_j(x_{ij}),\ F_j(x_{2j} \mid x_{1j}),\ F_j(x_{kj} \mid x_{1j}, \cdots, x_{(k-1)j})$$

are obtained in the following manner.

For any $s = 2, \cdots, k$, let σ_s^2 be the sth diagonal element of V, V_{s-1} be the sub-matrix of V obtained by considering the first $(s - 1)$ rows and columns, and let $\mathbf{v}_{s-1}$ be the vector for which the following relation holds:

$$V_s = \left[\begin{array}{c|c} V_{s-1} & \mathbf{v}_{s-1} \\ \hline \mathbf{v}_{s-1}' & \sigma_s^2 \end{array}\right].$$

Let $\bar{\mathbf{X}}_j' = (\bar{X}_{1j}, \cdots, \bar{X}_{kj})$; and then $\tilde{F}_j(x_{1j})$ is that of a $N(\bar{X}_{1j}, (1 - 1/j)\sigma_1^2)$ distribution evaluated at x_{ij}; and for $s = 2, \cdots, k$; $\tilde{F}_j(x_{sj} \mid x_{ij}, \cdots, x_{(s-1)j})$ is the df of a $N(\tilde{\mu}_{sj}, (1 - 1/j)\sigma^2_{s|1,\cdots,s-1})$ distribution, where

$$\tilde{\mu}_{sj} = \bar{X}_{sj} + \mathbf{v}_{s-1}' V_{s-1}^{-1}((x_{1j}, \cdots, x_{(s-1)j})' - (\bar{X}_{1j}, \cdots, \bar{X}_{(s-1)j})'),$$

and

$$\sigma^2_{s|1,\cdots,s-1} = \sigma_s^2 - \mathbf{v}_{s-1}' V_{s-1}^{-1} \mathbf{v}_{s-1}.$$

These are the usual conditional means and variances.

If sufficiently accurate tables are available, the transformations can be carried out, obtaining $(n - 1) \cdot k$ independent and uniformly distributed random variables in (0, 1). If a chi-square test statistic is to be used, the use of only some prescribed percentiles of the normal tables will be needed. In this latter case, suppose M cells are to be selected with equal probability. The expression for each of the $(n - 1) \cdot k \cdot (M - 1)$ percentile estimators is given by:

$$\tilde{C}_{ijs} = \{(1 - 1/j)\sigma^2_{s|1,\cdots,s-1}\}^{\frac{1}{2}} \cdot z_i + \tilde{\mu}_{sj};$$

where z_i is the (i/m)th percentile of a $N(0, 1)$ distribution, $i = 1, \cdots, M - 1$; $j = 2, \cdots, n$ and $s = 1, \cdots, k$.

6. Remark. In applying the tests of the foregoing sections caution must be exercised to insure that the observations are not systematically ordered; otherwise the distribution theory may be disturbed, in that, in general, a different set of values of the transformed observations will be obtained for different permutations of the observations; and, therefore, different values for the test statistic. This does not, however, in any way affect the validity of the tests set forth here.

Finally, we would like to point out that the general approach to composite

testing problems set out here constitutes an attractive alternative approach to presently used methods, particularly to likelihood ratio procedures.

REFERENCES

[1] BAHADUR, R. R. (1954). Sufficiency and statistical decision functions. *Ann. Math. Statist.* **25** 523–462.

[2] BERK, R. H. (1969). Strong consistency of certain sequential estimators. *Ann. Math. Statist.* **40** 1492–1495.

[3] BIRNBAUM, Z. W. (1953). Distribution free tests of fit for continuous distribution functions. *Ann. Math. Statist.* **24** 484–489.

[4] DAVID, F. N. and JOHNSON, N. L. (1948). The probability integral transformation when parameters are estimated from the sample. *Biometrika* **35** 182–192.

[5] GHURYE, S. G. and OLKIN, I. (1969). Unbiased estimation of some multivariate probability densities and related functions. *Ann. Math. Statist.* **40** 1261–1271.

[6] GOOD, I. J., GOVER, T. N., and MITCHELL, G. L. (1970). Exact distributions for X^2 and for the likelihood-ratio statistic for the equiprobable multinomial distribution. *J. Amer. Statist. Assoc.* **65** 267–283

[7] LIEBERMAN, G. J. and RESNIKOFF, G. J. (1955). Sampling plans for inspection by variables. *J. Amer. Statist. Assoc.* **50** 457–516.

[8] LILLIEFORS, H. W. (1967). On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *J. Amer. Statist. Assoc.* **62** 399–402.

[9] LILLIEFORS, H. W. (1969). On the Kolmogorov–Smirnov test for the exponential distribution with mean unknown. *J. Amer. Statist. Assoc.* **64** 387–389.

[10] MANN, H. B. and WALD, A. (1942). On the choice of the number of class intervals in the application of the chi-square test. *Ann. Math. Statist.* **13** 306–317.

[11] ROSENBLATT, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Statist.* **23** 470–472.

[12] SATHE, Y. S. and VARDE, S. D. (1969). On minimum variance unbiased estimation of reliability. *Ann. Math. Statist.* **40** 710–714.

[13] SEHEULT, A. H. and QUESENBERRY, C. P. (1971). On unbiased estimation of density functions. *Ann. Math. Statist.* **42** 1434–1438.

[14] WATSON, G. S. (1957). The χ^2 goodness-of-fit test for normal distributions. *Biometrika* **44** 336–348.

[15] WATSON, G. S. (1958). On chi-square goodness-of-fit tests for continuous distributions. *J. Roy. Statist. Soc. Ser. B* **20** 44–61.

[16] ZAHN, D. A. and ROBERTS, G. C. (1971). Exact χ^2 criterion tables with cell expectations one: An application to Coleman's measure of concensus. *J. Amer. Statist. Assoc.* **66** 145–148.

DEPARTMENT OF STATISTICS
NORTH CAROLINA STATE UNIVERSITY
RALEIGH, NORTH CAROLINA 27607