The Annals of Statistics
1987, Vol. 15, No. 2, 805–816

UNBIASEDNESS OF TESTS FOR HOMOGENEITY[1]

BY ARTHUR COHEN AND HAROLD B. SACKROWITZ

Rutgers University

Let X_i, $i = 1, 2, \ldots, k$, be independent random variables distributed according to a one-parameter exponential family with parameter θ_i. Assume also that the probability density function of X_i is a Pólya frequency function of order two (PF_2). Consider the null hypothesis H_0: $\theta_1 = \theta_2 = \cdots = \theta_k$ against the alternative K: not H_0. We show that any permutation invariant test of size α, whose conditional (on $T = \sum_{i=1}^k X_i$) acceptance sections are convex, is unbiased. A stronger result is that any size α test function φ, which is Schur-convex for fixed t, is unbiased. Previously, such a result was known only for the normal and Poisson cases.

1. Introduction and summary. Let $X_1, X_2, \ldots, X_k$ be independent random variables distributed according to a one-parameter exponential family with parameter θ_i, $i = 1, 2, \ldots, k$. That is, the joint density of the X_i is

$$f(\mathbf{x}, \boldsymbol{\theta}) = \left(\prod_{i=1}^{k} \beta(\theta_i)\right) \exp\left(\sum_{i=1}^{k} x_i \theta_i\right) \left(\prod_{i=1}^{k} h(x_i)\right), \tag{1.1}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_k)$, $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$. The dominating measure is Lebesgue for the continuous case and counting measure for the case where X_i are integer-valued. Assume $h(x_i)$ is a Pólya frequency function of order two (PF_2), that is, $h(x_i)$ is log concave. The problem is to test the null hypothesis H_0: $\theta_1 = \theta_2 = \cdots = \theta_k$ against the alternative K: not H_0. In this paper we study the issue of unbiasedness of permutation invariant (PI) tests. Note that any unbiased test for the model under discussion must be similar and therefore must have Neyman structure with respect to $T = \sum_{i=1}^k X_i = t$. See Lehmann (1959), Theorem 2, page 134. Hence, any unbiased test of size α must have conditional size α (given $T = t$). For $k = 2$, Lehmann (1959) exhibits a uniformly most powerful size α unbiased test for this model. See Lehmann (1959), Chapter 4, Sections 4 and 5. For X_i normal with unknown means or X_i Poisson, it is easily shown, using the results of Cohen and Sackrowitz (1975) and Marshall and Olkin (MO) (1979), pages 386 and 391, that any PI test, which has conditional acceptance sections that are convex, is unbiased. In fact, in these cases the stronger result of Schur-convex power functions follows.

For X_i gamma, $i = 1, 2, \ldots, k$, with equal shape parameter, but varying scale parameter, Cohen and Strawderman (1971) study the two-parameter family of permutation invariant test statistics,

$$R(\lambda, \eta) = \left(\sum_{i=1}^{k} x_i^{\lambda}\right)^{1/\lambda} \Bigg/ \left(\sum_{i=1}^{k} x_i^{\eta}\right)^{1/\eta}. \tag{1.2}$$

Received November 1985; revised September 1986.

[1]Research supported by NSF Grant DMS-8418416.

AMS 1980 *subject classification*. Primary 62F03.

Key words and phrases. Homogeneity, unbiasedness, similar test, Neyman structure, majorization, Schur convexity, stochastic ordering, Pólya frequency two.

For $\lambda \geq 0 \geq \eta$, tests defined by rejecting when $R(\lambda, \eta) > K$ were proven to be unbiased and have monotone power functions. See also MO (1979), page 387. However, the issue was unresolved for $R(\infty, 1)$, which is known as Cochran's test, and for $R(2,1)$, which is the locally most powerful locally unbiased test [see Cohen, Sackrowitz and Strawderman (1985)]. One application of the results of this paper is that all tests $R(\lambda, 1)$, $\lambda > 1$, are unbiased when the gamma density is PF_2.

In this paper we show that any PI size α test function which is Schur-convex for fixed t is unbiased. Such a result is applicable to all exponential family–PF_2 distributions that are continuous or integer-valued, including normal, Poisson, gamma (with shape parameter at least 1), binomial, geometric and others. Furthermore, since the homogeneity hypothesis remains invariant under a monotone transformation of the original random variables, the development applies to many other distributions, namely those which can be transformed monotonically into one of the distributions already mentioned. Thus, results apply to one-parameter beta distributions, Pareto distributions, log normal and others. Of course, the conditions required for unbiasedness would be expressed in terms of the transformed variates.

In the normal and Poisson cases the previous method used to prove unbiasedness used the Schur-convex preserving property of certain conditional densities. In fact, that property yielded the stronger result of monotonicity, in some sense, of power functions. However, for example, the relevant conditional densities may not have this property in the gamma case and does *not* have this property in the binomial case. (See Remark 5.6.) In this paper a different method of proof is given that works for the general class of distributions mentioned.

A useful theorem for our development has been established by Efron (1965). We give a slight modification of Efron's result in the next section along with some preliminaries. Section 3 contains a key stochastic ordering result. The main theorem is given in Section 4.

A complete proof will be given for the continuous case only. The main results are identical for the case of integer-valued random variables. However, because ties can occur in the latter case, more details would be required. Hence in Sections 2–4 continuity is assumed. In Section 5 we indicate how the integer-valued random variable case is treated.

2. Preliminaries. In this section we define majorization, Schur convexity and PF_2. We also note some basic results concerning these notions.

Let $X_{(1)} \geq X_{(2)} \geq \cdots \geq X_{(k)}$ denote the order statistics.

DEFINITION 2.1. For $\mathbf{u}, \mathbf{v} \in R^k$,

$$\mathbf{u} \preccurlyeq \mathbf{v}, \quad \text{if} \begin{cases} \sum_{i=1}^{m} u_{(i)} \leq \sum_{i=1}^{m} v_{(i)}, & m = 1, 2, \ldots, k-1, \\ \sum_{i=1}^{k} u_{(i)} = \sum_{i=1}^{k} v_{(i)}. & \end{cases} \tag{2.1}$$

When $\mathbf{u} \preccurlyeq \mathbf{v}$, $\mathbf{u}$ is said to be majorized by $\mathbf{v}$.

DEFINITION 2.2. A real-valued function U defined on a set $\mathscr{A} \subseteq R^k$ is said to be Schur-convex on $\mathscr{A}$ if $\mathbf{u} \preccurlyeq \mathbf{v}$ on $\mathscr{A}$ implies $U(\mathbf{u}) \le U(\mathbf{v})$.

Let g be the joint density of the order statistics given $\Sigma X_i = t$. Hence, the domain of g is $\{(x_{(1)}, x_{(2)}, \dots, x_{(k)}): x_{(1)} \ge x_{(2)} \ge \cdots \ge x_{(k)}, \Sigma_{i=1}^k x_{(i)} = t\}$.

THEOREM 2.3. *The ratio*

$$g(\mathbf{u}, \boldsymbol{\theta}|t)/g(\mathbf{u}, \bar{\boldsymbol{\theta}}|t), \tag{2.2}$$

where $\bar{\boldsymbol{\theta}} = (\bar{\theta}, \bar{\theta}, \dots, \bar{\theta})$, $\bar{\theta} = \Sigma_{i=1}^k \theta_i/k$, *is Schur-convex in* $\mathbf{u} = (x_{(1)}, x_{(2)}, \dots, x_{(k)})$.

PROOF. It follows from expression (5.12) in Cohen, Sackrowitz and Strawderman (1985) that (2.2) may be expressed as

$$C(t) \sum_{j=1}^{p!} e^{(\boldsymbol{\theta}_{(j)} - \bar{\boldsymbol{\theta}})\mathbf{u}'}, \tag{2.3}$$

which is the symmetrized version of $e^{\Sigma x_{(i)}(\theta_i - \bar{\theta})}$. Thus, (2.3) is symmetric and convex in $x_{(1)}, x_{(2)}, \dots, x_{(k)}$, and hence Schur-convex. See MO, page 67. □

When $\theta_1 = \theta_2 = \cdots = \theta_k$, the null hypothesis H_0 is true. We let $\boldsymbol{\theta}_0$ denote any vector which has all its coordinates equal. Note that conditional expectations given t under H_0 do not depend on the particular common coordinate value of the vector $\boldsymbol{\theta}_0$. Also note that in Theorem 2.3 that $\bar{\boldsymbol{\theta}}$ is a particular $\boldsymbol{\theta}_0$.

In the remainder of the paper dimensionality will be important so we let $\mathbf{X}^{(k)} = (X_1, X_2, \dots, X_k)$ and let $X_{(1)}^{(k)} \ge \cdots \ge X_{(k)}^{(k)}$ denote the order statistics for the random variables $X_1, \dots, X_k$. We also define upper (T's) tail sums of the order statistics by

$$T_i^{(k)} = X_{(1)}^{(k)} + \cdots + X_{(i)}^{(k)}, \qquad i = 1, \dots, k,$$

and $\mathbf{T}^{(k)} = (T_1^{(k)}, \dots, T_k^{(k)})$. We let $\mathscr{D}^{(k)}$ denote the range of $\mathbf{T}^{(k)}$, i.e.,

$$\mathscr{D}^{(k)} = \{(t_1, t_2, \dots, t_k): t_1 \ge t_2 - t_1 \ge \cdots \ge t_k - t_{k-1}\}$$

and define the family $\mathscr{H}_k$ of monotone functions on $\mathscr{D}^{(k)}$ as follows:

$\mathscr{H}_k = \{W(t_1, t_2, \dots, t_k)$: for fixed t_k, W is nondecreasing in t_i, $i = 1, 2, \dots, k-1$, on $\mathscr{D}^{(k)}\}$. The class $\mathscr{H}_k$ was considered by Nevius, Proschan and Sethuraman (1977).

The next theorem, which follows from results in MO (1979), suggests the view we take of test procedures.

THEOREM 2.4. *If ϕ is a permutation invariant test which is Schur-convex for each fixed $\Sigma_{i=1}^k X_i = t$, then ϕ, as a function of $T_1^{(k)}, \dots, T_k^{(k)}$, belongs to $\mathscr{H}_k$. In particular this holds if ϕ has convex acceptance regions for each fixed $\Sigma X_i = t$.*

PROOF. By arguments in MO (pages 67–68) if ϕ is permutation invariant and has convex acceptance sections, then ϕ is Schur-convex. The result now follows from MO (page 55) for the transformation to $\mathbf{T}^{(k)}$. □

REMARK 2.5. MO (page 55) is not limited to test functions. Hence if $h(\mathbf{x})$ is Schur-convex, then h as a function of $T_1^{(k)}, \dots, T_k^{(k)}$ belongs to $\mathscr{H}_k$.

DEFINITION 2.6. A probability density function on the real line $r(x)$ is said to be PF_2 if $x_2 \geq x_1$, $z_2 \geq z_1$ implies

$$r(x_1 - z_1)r(x_2 - z_2) - r(x_1 - z_2)r(x_2 - z_1) \geq 0. \tag{2.4}$$

Note that (2.4) is equivalent to the log concavity of r.

The following theorem is a slight modification of a result of Efron (1965).

THEOREM 2.7. *Let* $\mathbf{X} = (X_1, X_2, \dots, X_k)$, *where* X_i *satisfy* (1.1) *and* h *is* PF_2. *Let* A *be a rectangle set in* k-*space*: $A = \{\mathbf{x}: a_i \leq x_i \leq b_i,\ i = 1, 2, \dots, k\}$. *Fix* $s \leq s'$ *and let* $B = \{\mathbf{x}: s \leq \sum_{i=1}^k x_i \leq s'\}$. *Then*

$$E_{\boldsymbol{\theta}_0}\left\{ W(\mathbf{X}) | T_k^{(k)} = s, \mathbf{X} \in A \right\} \leq E_{\boldsymbol{\theta}_0}\left\{ W(\mathbf{X}) | T_k^{(k)} = s', \mathbf{X} \in A \right\},$$

for all W *defined and nondecreasing* (*coordinatewise*) *on* $A \cap B$.

PROOF. Let $M = \sup_{A \cap B} |W(\mathbf{x})|$. Define $\hat{W}(\mathbf{x}) = W(\mathbf{x})$ if $\mathbf{x} \in A \cap B$ but $\hat{W}(\mathbf{x}) = -M$ if $\sum x_i < s$ and $\hat{W}(\mathbf{x}) = M$ if $\sum x_i > s'$. Note that $\hat{W}$ is nondecreasing on all of A. Thus, using Corollary 2, page 277, of Efron (1965), along with the definition of $\hat{W}$, we obtain

$$\begin{aligned} E_{\boldsymbol{\theta}_0}\left\{ W(\mathbf{X}) | T_k^{(k)} = s, \mathbf{X} \in A \right\} &= E_{\boldsymbol{\theta}_0}\left\{ \hat{W}(\mathbf{X}) | T_k^{(k)} = s, \mathbf{X} \in A \right\} \\ &\leq E_{\boldsymbol{\theta}_0}\left\{ \hat{W}(\mathbf{X}) | T_k^{(k)} = s', \mathbf{X} \in A \right\} \\ &= E_{\boldsymbol{\theta}_0}\left\{ W(\mathbf{X}) | T_k^{(k)} = s', \mathbf{X} \in A \right\}. \qquad \square \end{aligned}$$

Note that the above result is equivalent to stochastic ordering of the conditional distribution of $\mathbf{X}$ (see MO, Chapter 17).

We remark that all results stated in the remainder of the paper concerning expectations hold only when the specified constants are such that the conditional distributions are defined. For example, in Theorem 2.8 below, x, x^* and t must be such that $x \geq (t - x^*)/k$.

One important consequence of Theorem 2.7 for functions in $\mathscr{H}_{k+1}$ is

THEOREM 2.8. *Let* $X_1, \dots, X_k$ *satisfy* (1.1) *with* h PF_2. *Let* $W \in \mathscr{H}_{k+1}$. *Then for* $x < x^*$

$$\begin{aligned} &E_{\boldsymbol{\theta}_0}\left\{ W\left(x^*, x^* + T_1^{(k)}, \dots, x^* + T_k^{(k)}\right) \middle| X_{(1)}^{(k)} \leq x,\ T_k^{(k)} = t - x^* \right\} \\ &\quad \geq E_{\boldsymbol{\theta}_0}\left\{ W\left(x, x + T_1^{(k)}, \dots, x + T_k^{(k)}\right) \middle| X_{(1)}^{(k)} \leq x,\ T_k^{(k)} = t - x \right\}. \end{aligned} \tag{2.5}$$

PROOF.

$$\begin{aligned} &E_{\boldsymbol{\theta}_0}\left\{ W\left(x^*, x^* + T_1^{(k)}, \dots, x^* + T_k^{(k)}\right) \middle| X_{(1)}^{(k)} \leq x,\ T_k^{(k)} = t - x^* \right\} \\ &\quad = E_{\boldsymbol{\theta}_0}\left\{ W\left(t - X_{(1)}^{(k)} - \cdots - X_{(k)}^{(k)}, \dots, \right.\right. \\ &\qquad\qquad \left.\left. t - X_{(k)}^{(k)}, t\right) \middle| X_{(1)}^{(k)} \leq x,\ T_k^{(k)} = t - x^* \right\}. \end{aligned} \tag{2.6}$$

We note that $W(t - X_{(1)}^{(k)} - \cdots - X_{(k)}^{(k)}, \dots, t - X_{(k)}^{(k)}, t)$ as a function of $\mathbf{X}^{(k)}$ is defined and nonincreasing in each X_i on

$$\{X_i \le x,\ i = 1, \dots, k\} \cap \left\{t - x^* \le \sum_{i=1}^{k} X_i \le t - x\right\}.$$

It follows from Theorem 2.7 that (2.6) is

$$\begin{aligned}&\ge E_{\theta_0}\left\{W\left(t - X_{(1)}^{(k)} - \cdots - X_{(k)}^{(k)}, \dots, t - X_{(k)}^{(k)}, t\right) \middle| X_{(1)}^{(k)} \le x,\ T_k^{(k)} = t - x\right\}\\ &= E_{\theta_0}\left\{W\left(x, x + T_1^{(k)}, \dots, x + T_k^{(k)}\right) \middle| X_{(1)}^{(k)} \le x,\ T_k^{(k)} = t - x\right\}. \qquad \square\end{aligned}$$

3. Monotonicity of expectations. This section contains the crucial property, as expressed in Theorem 3.2, of the tail sum random variables. Simply stated these theorems purport the notion that if the underlying distribution is PF_2 and $\Sigma_{i=1}^{k} X_i$ is fixed, then as the maximum $X_{(1)}^{(k)}$ increases, the upper tail sums increase (stochastically).

The purpose of the following lemma is to equate certain conditional distributions of high-dimensional order statistics with that of lower-dimensional order statistics.

LEMMA 3.1. *Assume H_0 is true. Consider the random variables $X_{(1)}^{(u+v)}, \dots, X_{(u)}^{(u+v)}, \dots, X_{(u+v)}^{(u+v)}$. Fix constants $r_1 > \cdots > r_u$ and let $R = \Sigma_{i=1}^{u} r_i$. The conditional distribution of $X_{(u+1)}^{(u+v)}, \dots, X_{(u+v)}^{(u+v)}$ given $X_{(i)}^{(u+v)} = r_i$, $i = 1, \dots, u$, $\Sigma_{i=1}^{u+v} X_{(i)}^{(u+v)} = t$, is the same as each of the following two conditional distributions*:

(i) *the distribution of $X_{(2)}^{(v+1)}, \dots, X_{(v+1)}^{(v+1)}$ given $X_{(1)}^{(v+1)} = r_u$, $\Sigma_{i=1}^{v+1} X_{(i)}^{(v+1)} = t - R + r_u$;*

(ii) *the distribution of $X_{(1)}^{(v)}, \dots, X_{(v)}^{(v)}$ given $\Sigma_{i=1}^{v} X_{(i)}^{(v)} = t - R$, $X_{(1)}^{(v)} \le r_u$.*

PROOF. The proof is immediate upon inspection of the relevant densities. Since all random variables are assumed to be continuous, it is easy to see that all the conditional density functions are the same. □

A Lemma 3.1-type result is the greatest obstacle in the path of an elegant exposition or the integer-valued case. Lemma 3.1 is not true for this case. The results in the remainder of this section and in Section 4 are true, as stated, even for the integer-valued case. The proofs for the integer-valued case require more detail and are discussed in Section 5.

THEOREM 3.2. *If $x < x^*$, then*

$$\begin{aligned}&E_{\theta_0}\{W(\mathbf{T}^{(k)}) | X_{(1)}^{(k)} = x,\ T_k^{(k)} = t\}\\ &\qquad \le E_{\theta_0}\{W(\mathbf{T}^{(k)}) | X_{(1)}^{(k)} = x^*,\ T_k^{(k)} = t\},\end{aligned} \tag{3.1}$$

for all $W \in \mathscr{H}_k$ and all $k = 2, 3, \dots$.

PROOF. The proof is by induction. First note that when $k = 2$, (3.1) reduces to $W(x, t) \leq W(x^*, t)$ which is trivially true as $W \in \mathscr{H}_2$. Next we show that if (3.1) is true for k, it will also be true for $k + 1$. By Lemma 3.1(ii) (with $u = 1$, $v = k$), and letting $W_{x^*}^*(\mathbf{T}^{(k)}) = W(x^*, x^* + T_1^{(k)}, \dots, x^* + T_k^{(k)})$

$$\begin{aligned}
&E_{\theta_0}\{W(\mathbf{T}^{(k+1)})|X_{(1)}^{(k+1)} = x^*, T_{k+1}^{(k+1)} = t\} \\
&\quad = E_{\theta_0}\{W_{x^*}^*(\mathbf{T}^{(k)})|X_{(1)}^{(k)} \leq x^*, T_k^{(k)} = t - x^*\} \\
&\quad = E_{\theta_0}\{W_{x^*}^*(\mathbf{T}^{(k)})|X_{(1)}^{(k)} \leq x, T_k^{(k)} = t - x^*\}P \\
&\qquad + E_{\theta_0}\{W_{x^*}^*(\mathbf{T}^{(k)})|x < X_{(1)}^{(k)} \leq x^*, T_k^{(k)} = t - x^*\}(1 - P),
\end{aligned} \tag{3.2}$$

where $P = P_{\theta_0}(X_{(1)}^{(k)} \leq x | X_{(1)}^{(k)} \leq x^*, T_k^{(k)} = t - x^*)$.

For $\mathbf{T}^{(k)} \in \mathscr{D}^{(k)}$ define the function $V^*(\mathbf{T}^{(k)}) = W_{x^*}^*(\mathbf{T}^{(k)})$ when $T_1^{(k)} \leq x^*$ and $V^*(\mathbf{T}^{(k)}) = \sup_{\mathbf{t} \in \mathscr{D}^{(k)}}\{W_{x^*}^*(\mathbf{t})\}$, otherwise. Thus V^* belongs to $\mathscr{H}_k$. By the induction hypothesis the integrals below are nondecreasing in u so that

$$\begin{aligned}
&E_{\theta_0}\{W_{x^*}^*(\mathbf{T}^{(k)})|x < X_{(1)}^{(k)} \leq x^*, T_k^{(k)} = t - x^*\} \\
&\quad = E_{\theta_0}\{V^*(\mathbf{T}^{(k)})|x < X_{(1)}^{(k)} \leq x^*, T_k^{(k)} = t - x^*\} \\
&\quad = \int_x^{x^*} E_{\theta_0}\{V^*(\mathbf{T}^{(k)})|X_{(1)}^{(k)} = u, T_k^{(k)} = t - x^*\} \\
&\qquad \times dP_{\theta_0}\big(X_{(1)}^{(k)} = u | x < X_{(1)}^{(k)} \leq x^*, T_k^{(k)} = t - x^*\big) \\
&\quad \geq E_{\theta_0}\{V^*(\mathbf{T}^{(k)})|X_{(1)}^{(k)} = x, T_k^{(k)} = t - x^*\} \\
&\quad \geq \int_{-\infty}^{x} E_{\theta_0}\{V^*(\mathbf{T}^{(k)})|X_{(1)}^{(k)} = u, T_k^{(k)} = t - x^*\} \\
&\qquad \times dP_{\theta_0}\big(X_{(1)}^{(k)} = u | X_{(1)}^{(k)} \leq x, T_k^{(k)} = t - x^*\big) \\
&\quad = E_{\theta_0}\{V^*(\mathbf{T}^{(k)})|X_{(1)}^{(k)} \leq x, T_k^{(k)} = t - x^*\} \\
&\quad = E_{\theta_0}\{W_{x^*}^*(\mathbf{T}^{(k)})|X_{(1)}^{(k)} \leq x, T_k^{(k)} = t - x^*\}.
\end{aligned}$$

Therefore, (3.2) is

$$\geq E_{\theta_0}\{W_{x^*}^*(\mathbf{T}^{(k)})|X_{(1)}^{(k)} \leq x, T_k^{(k)} = t - x^*\}. \tag{3.3}$$

It follows from Theorem 2.8 that (3.3) is

$$\geq E_{\theta_0}\{W_x^*(\mathbf{T}^{(k)})|X_{(1)}^{(k)} \leq x, T_k^{(k)} = t - x\}. \tag{3.4}$$

However, by Lemma 3.1(ii), (3.4) is equal to

$$E_{\theta_0}\{W(\mathbf{T}^{(k+1)})|X_{(1)}^{(k+1)} = x, T_{k+1}^{(k+1)} = t\},$$

which completes the proof. □

An important consequence of Theorem 3.2 is

COROLLARY 3.3. *For any* $W \in \mathscr{H}_k$, $1 \le m \le k-1$,

$$E_{\theta_0}\left\{ W(\mathbf{T}^{(k)}) | T_1^{(k)} = t_1, \dots, T_m^{(k)} = t_m, T_k^{(k)} = t_k \right\} \tag{3.5}$$

is a nondecreasing function of t_m.

PROOF. For each fixed $t_1, \dots, t_{m-1}, t_k$ write (3.5) as

$$E_{\theta_0}\left\{ W(\mathbf{T}^{(k)}) | X_{(1)}^{(k)} = t_1,\ X_{(2)}^{(k)} = t_2 - t_1, \dots, X_{(m)}^{(k)} = t_m - t_{m-1},\ T_k^{(k)} = t_k \right\},$$

which by Lemma 3.1(i) (with $u = m$, $v = k - m$) is equal to

$$E_{\theta_0}\left\{ W'(\mathbf{T}^{(k-m+1)}) | X_{(1)}^{(k-m+1)} = t_m - t_{m-1},\ T_{k-m+1}^{(k-m+1)} = t_k - t_{m-1} \right\},$$

where $W'(\mathbf{T}^{(k-m+1)}) = W(t_1, \dots, t_{m-1}, t_{m-1} + T_1^{(k-m+1)}, \dots, t_{m-1} + T_{k-m+1}^{(k-m+1)})$. Application of Theorem 3.2 completes the proof. □

4. Main result. In this section we prove that PI tests which are Schur-convex for fixed t are unbiased. The method of proof resembles that used by Perlman and Olkin (1980).

We will use Corollary 3.3 to prove

LEMMA 4.1. *Let* $W_1, W_2 \in \mathscr{H}_k$. *Then*

$$\begin{aligned} &E_{\theta_0}\left\{ W_1(\mathbf{T}^{(k)}) W_2(\mathbf{T}^{(k)}) | T_k^{(k)} = t \right\} \\ &\quad \ge E_{\theta_0}\left\{ W_1(\mathbf{T}^{(k)}) | T_k^{(k)} = t \right\} E_{\theta_0}\left\{ W_2(\mathbf{T}^{(k)}) | T_k^{(k)} = t \right\}. \end{aligned} \tag{4.1}$$

PROOF. In this proof we will suppress the superscript (k) notation as we will always operate in k-dimensions.

$$\begin{aligned} &E_{\theta_0}\{ W_1(\mathbf{T}) W_2(\mathbf{T}) | T_k = t \} \\ &\quad = E_{\theta_0}\Big\{ E_{\theta_0}\Big\{ E_{\theta_0}\Big\{ \cdots E_{\theta_0}\big\{ E_{\theta_0}\{ W_1(\mathbf{T}) W_2(\mathbf{T}) | T_1, \dots, T_{k-2}, T_k = t \} \\ &\qquad\qquad | T_1, \dots, T_{k-3}, T_k = t \big\} \cdots \Big\} | T_1, T_k = t \Big\} | T_k = t \Big\}. \end{aligned}$$

We evaluate the expression beginning with the innermost expectation. Since W_1, W_2 are nondecreasing in T_{k-1} for fixed $T_1, \dots, T_{k-2}, T_k$ they are, conditionally, positively correlated. That is,

$$= E_{\theta_0}\{ W_1(\mathbf{T}) W_2(\mathbf{T}) | T_1, \dots, T_{k-2}, T_k = t \} \ge W_1'(\mathbf{T}) W_2'(\mathbf{T}),$$

where $W_i'(\mathbf{T}) = E_{\theta_0}\{ W_i(\mathbf{T}) | T_1, \dots, T_{k-2}, T_k = t \}$. By Corollary 3.3 $W_i'(\mathbf{T})$ is nondecreasing in T_{k-2} for fixed $T_1, \dots, T_{k-3}, T_k$. Therefore, the correlation inequality can be applied to the second innermost expectation. That is,

$$E_{\theta_0}\{ W_1'(\mathbf{T}) W_2'(\mathbf{T}) | T_1, \dots, T_{k-3}, T_k = t \} \ge W_1''(\mathbf{T}) W_2''(\mathbf{T}),$$

where $W_i''(\mathbf{T}) = E_{\theta_0}\{ W_i(\mathbf{T}) | T_1, \dots, T_{k-3}, T_k = t \}$. Again by Corollary 3.3 $W_i''(\mathbf{T})$

is nondecreasing in T_{k-3} for fixed $T_1, \ldots, T_{k-4}, T_k$ so that once again the correlation inequality can be applied. Continuing in this fashion yields (4.1). □

THEOREM 4.2. *Let $X_i = 1, 2, \ldots, k$ have joint density* (1.1) *with $h(x_i)$ PF_2. Let $\phi(\mathbf{X})$ be a PI test of size α which is Schur-convex for fixed t. Then ϕ is unbiased. That is,*

$$E_{\theta}\phi(\mathbf{X}) \geq E_{\theta_0}\phi(\mathbf{X}). \tag{4.2}$$

PROOF. Write $\phi(\mathbf{X})$ as $\phi'(\mathbf{T}^{(k)})$ where ϕ' belongs to $\mathscr{H}_k$. (See Theorem 2.4.) Then

$$\begin{aligned} E_{\theta}\big(\phi(\mathbf{X})|T_k^{(k)}\big) &= E_{\theta}\big(\phi'(\mathbf{T}^{(k)})|T_k^{(k)}\big) \\ &= \int \phi'(\mathbf{t}^{(k)}) k_{\theta}(t_1, \ldots, t_{k-1}|t_k)\, dt_1\, dt_2, \ldots, dt_{k-1} \\ &= \int \phi'(\mathbf{t}^{(k)}) \frac{k_{\theta}(t_1, \ldots, t_{k-1}|t_k)}{k_{\bar{\theta}}(t_1, \ldots, t_{k-1}|t_k)} \\ &\qquad \times k_{\bar{\theta}}(t_1, \ldots, t_{k-1}|t_k)\, dt_1\, dt_2, \ldots, dt_{k-1}, \end{aligned} \tag{4.3}$$

where $k_{\theta}(t_1, \ldots, t_{k-1}|t_k)$ denotes the conditional density of $T_1, \ldots, T_{k-1}$ given $T_k = t_k$. Theorem 2.3 and Remark 2.5 imply that $\{k_{\theta}(t_1, \ldots, t_{k-1}|t_k)/k_{\bar{\theta}}(t_1, \ldots, t_{k-1}|t_k)\}$ belongs to $\mathscr{H}_k$. Now apply Lemma 4.1 in (4.3) to yield

$$E_{\theta}\phi(\mathbf{x}) \geq E_{\bar{\theta}}\phi(\mathbf{x}). \tag{4.4}$$

Since $E_{\bar{\theta}}\phi(\mathbf{x}) = E_{\theta_0}\phi(\mathbf{x})$, (4.2) follows. □

A consequence of Theorem 4.2 is

COROLLARY 4.3. *Under the condition of Theorem* 4.2, *let $\phi(\mathbf{x})$ be a size α-test with conditional convex acceptance sections. Then ϕ is unbiased.*

PROOF. Apply Theorems 2.4 and 4.2. □

5. The integer-valued case. In this section we will demonstrate the reasoning and additional detail needed to handle the integer-valued case by proving Theorem 3.2 for such random variables. The difficulty stems from the fact that, when ties are possible, Lemma 3.1 does not hold. However, by Remark 1 of Efron (1965), page 278, the results of that paper can be extended to cover integer-valued random variables. Hence, in this section we can assume that Theorems 2.7 and 2.8 hold.

In order to deal with ties, we define $L^{(v)}(\mathbf{X}^{(v)}, r)$ to be the number of $X_{(i)}^{(v)}$'s that are equal to r when $X_{(1)}^{(v)} \leq r$. That is,

$$L^{(v)}(\mathbf{X}^{(v)}, r) = \begin{cases} 0, & \text{if } X_{(1)}^{(v)} < r, \\ l, & \text{if } X_{(1)}^{(v)} = \cdots = X_{(l)}^{(v)} = r,\ X_{(l+1)}^{(v)} < r, \\ v, & \text{if } X_{(1)}^{(v)} = \cdots = X_{(v)}^{(v)} = r. \end{cases}$$

The following lemma, which we state without proof, concerns the relationship among various conditional distributions. Note that part (i) of Lemma 5.1 is analogous to Lemma 3.1(i).

LEMMA 5.1. *Assume H_0 is true. Consider the random variables $X_{(1)}^{(v)}, \dots, X_{(v)}^{(v)}$. Fix r.*

(i) *For $l = 0, 1, \dots, v-1$, the conditional distribution of $(X_{(l+1)}^{(v)}, \dots, X_{(v)}^{(v)})$ given $L^{(v)}(\mathbf{X}^{(v)}, r) = l, \sum_{i=1}^{v} X_{(i)}^{(v)} = t$ is the same as the conditional distribution of $(X_{(1)}^{(v-l)}, \dots, X_{(v-l)}^{(v-l)})$ given $X_{(1)}^{(v-l)} < r$ and $\sum_{i=1}^{v-l} X_{(i)}^{(v-l)} = t - lr$.*

(ii) *The conditional distribution of $L^{(v-1)}(\mathbf{X}^{(v-1)}, r) + 1$ given $X_{(1)}^{(v-1)} \le r$ and $\sum_{i=1}^{v-1} X_{(i)}^{(v-1)} = t - r$ is stochastically larger than that of $L^{(v)}(\mathbf{X}^{(v)}, r)$ given $X_{(1)}^{(v)} = r$, $\sum_{i=1}^{v} X_{(i)}^{(v)} = t$.*

The main departure from Section 3 is that we must now pay careful attention to the quantity

$$E\{W(\mathbf{T}^{(k)}) | L^{(k)}(\mathbf{X}^{(k)}, r) = l, T_k^{(k)} = t\}. \tag{5.1}$$

In this section all distributions will be assumed to belong to the null hypothesis space and so we will suppress the $\boldsymbol{\theta}_0$ notation as in (5.1) above. Note that (5.1) is meaningful only if l is a nonnegative integer between $t - k(r-1)$ and t/r.

It is easily seen that, for $k = 2$, (5.1) is nondecreasing in (the possible values of) $l = 0, 1, 2$. For $k = 2$, (3.1) is obviously true for integer-valued, as well as continuous random variables. The proof of Theorem 3.2 here will also be by induction. We assume:

(A.1) For each fixed $m = 2, \dots, k$, (3.1) holds with k replaced by m and $W \in \mathscr{H}_m$.

(A.2) For each fixed $m = 2, \dots, k$, (5.1) is nondecreasing (where defined) in $l = 1, \dots, m$ and $W \in \mathscr{H}_m$.

The proof of Theorem 3.2 will be complete once we show that (A.1) and (A.2) together imply that (3.1) is true with k replaced by $k+1$ *and* (5.1) with k replaced by $k+1$ is nondecreasing in $l = 1, \dots, k+1$. This will be done by Lemmas 5.4 and 5.3, respectively.

In the remainder of the paper we often use $a + \mathbf{t}$ to denote $(a + t_1, \dots, a + t_n)$ where a is a scalar and $\mathbf{t} = (t_1, \dots, t_n)$.

LEMMA 5.2. *If* (A.2) *holds, then for $W \in \mathscr{H}_k$*

$$\begin{aligned} &E\{W(\mathbf{T}^{(k)}) | X_{(1)}^{(k)} = r, T_k^{(k)} = t\} \\ &\quad \le E\{W(r, r + \mathbf{T}^{(k-1)}) | X_{(1)}^{(k-1)} \le r, T_{k-1}^{(k-1)} = t - r\}. \end{aligned} \tag{5.2}$$

PROOF. Write

$$E\{W(\mathbf{T}^{(k)})|X_{(1)}^{(k)} = r, T_k^{(k)} = t\}$$
$$(5.3) \qquad = \sum_{l=1}^{k} E\{W(\mathbf{T}^{(k)})|L^{(k)}(\mathbf{X}^{(k)}, r) = l, T_k^{(k)} = t\}$$
$$\times P\big(L^{(k)}(\mathbf{X}^{(k)}, r) = l|X_{(1)}^{(k)} = r, T_k^{(k)} = t\big),$$

which equals

$$(5.4) \qquad \sum_{l=1}^{k} E\{W(r, r + \mathbf{T}^{(k-1)})|L^{(k-1)}(\mathbf{X}^{(k-1)}, r) = l - 1, T_{k-1}^{(k-1)} = t - r\}$$
$$\times P\big(L^{(k)}(\mathbf{X}^{(k)}, r) = l|X_{(1)}^{(k)} = r, T_k^{(k)} = t\big),$$

since Lemma 5.1(i) implies that the two conditional distributions involved in the two expectations in (5.3) and (5.4) are the same. By monotonicity of (5.1) the expectations in (5.3), which are equal to those in (5.4), are nondecreasing in l. Thus, by Lemma 5.1(ii), (5.4) is

$$\leq \sum_{l=1}^{k} E\{W(r, r + \mathbf{T}^{(k-1)})|L^{(k-1)}(\mathbf{X}^{(k-1)}, r) = l - 1, T_{k-1}^{(k-1)} = t - r\}.$$

Also

$$P\big(L^{(k-1)}(\mathbf{X}^{(k-1)}, r) = l - 1|X_{(1)}^{(k-1)} \leq r, T_{k-1}^{(k-1)} = t - r\big)$$
$$= E\{W(r, r + \mathbf{T}^{(k-1)})|X_{(1)}^{(k-1)} \leq r, T_{k-1}^{(k-1)} = t - r\},$$

completing the proof. □

LEMMA 5.3. *If* (A.1) *and* (A.2) *hold and* $W \in \mathscr{H}_{k+1}$, *then* (5.1) *with* k *replaced by* $k + 1$ *is nondecreasing in* $l = 1, 2, \dots, k + 1$.

PROOF. Use Lemma 5.1(i) to write

$$E\{W(\mathbf{T}^{(k+1)})|L^{(k+1)}(\mathbf{X}^{(k+1)}, r) = l, T_{k+1}^{(k+1)} = t\}$$
$$= E\{W(r, 2r, \dots, lr, lr + \mathbf{T}^{(k-l+1)})$$
$$(5.5) \qquad |X_{(1)}^{(k-l+1)} \leq r - 1, T_{k-l+1}^{(k-l+1)} = t - lr\}$$
$$\leq E\{W(r, \dots, lr, lr + \mathbf{T}^{(k-l+1)})$$
$$|X_{(1)}^{(k-l+1)} = r - 1, T_{k-l+1}^{(k-l+1)} = t - lr\},$$

where this last inequality effectively follows from assumption (A.1). [This was actually shown in the proof of Theorem 3.2, as the portion of that proof which shows (3.2) to be greater than (3.3) does not depend on the continuity of the random variables.] By Lemma 5.2 it follows that (5.5) is

$$\leq E\{W(r,\ldots,lr,(l+1)r-1,(l+1)r-1+\mathbf{T}^{(k-l)}) \mid X_{(1)}^{(k-l)} \leq r-1,\ T_{k-l}^{(k-l)} = t-(l+1)r+1\},$$

which by an argument analogous to the proof of Theorem 2.8 is

$$\leq E\{W(r,\ldots,(l+1)r,(l+1)r+\mathbf{T}^{(k-l)}) \mid X_{(1)}^{(k-l)} \leq r-1,\ T_{k-l}^{(k-l)} = t-(l+1)r\},$$

which by Lemma 5.1(i) is

$$= E\{W(\mathbf{T}^{(k+1)}) \mid L^{(k+1)}(\mathbf{X}^{(k+1)}, r) = l+1,\ T_{k+1}^{(k+1)} = t\}.$$

This completes the proof. □

LEMMA 5.4. *If* (A.1) *and* (A.2) *hold and* $W \in \mathscr{H}_{k+1}$, *then* (3.1) *holds with* k *replaced by* $k+1$.

PROOF. Lemma 5.3 implies that

$$E\{W(\mathbf{T}^{(k+1)}) \mid X_{(1)}^{(k+1)} = x^*,\ T_{k+1}^{(k+1)} = t\} \geq E\{W(\mathbf{T}^{(k+1)}) \mid L^{(k+1)}(\mathbf{X}^{(k+1)}, x^*) = 1,\ T_{k+1}^{(k+1)} = t\},$$

which by Lemma 5.1(i) is

$$= E\{W(x^*, x^*+\mathbf{T}^{(k)}) \mid X_{(1)}^{(k)} \leq x^*-1,\ T_k^{(k)} = t-x^*\},$$

which by Theorem 2.8 is

$$\geq E\{W(x^*-1, x^*-1+\mathbf{T}^{(k)}) \mid X_{(1)}^{(k)} \leq x^*-1,\ T_k^{(k)} = t-x^*+1\}. \tag{5.6}$$

Since Lemma 5.3 essentially says that (A.1) and (A.2) imply that (A.2) holds with k replaced by $k+1$, it follows from Lemma 5.2 that (5.2) holds with k replaced by $k+1$. Thus, (5.6) is

$$\geq E\{W(\mathbf{T}^{(k+1)}) \mid X_{(1)}^{(k+1)} = x^*-1,\ T_{k+1}^{(k+1)} = t\},$$

which completes the proof. □

Thus, Theorem 3.2 holds for integer-valued random variables. To prove Corollary 3.3 in this case requires a slight generalization of Lemma 5.1 (to the case of $u+v$ variables where the first u are fixed) and arguments similar to those used above.

REMARK 5.5. In Cohen, Sackrowitz and Strawderman (1985), Example 5.4 offers a situation where $k=3$, f is an exponential family and a test with conditional acceptance regions that are convex is *not* unbiased. It is easily verified in this example that f is not PF_2.

REMARK 5.6. The following example illustrates that the method used to prove unbiasedness in MO (page 391) for the Poisson case, for example, cannot be used for the model in this paper. Let $k = 3$, $X_i \sim B(4, p_i)$, $t = 6$. The ordered possible outcomes are $(4,2,0)$, $(4,1,1)$, $(3,3,0)$, $(3,2,1)$ and $(2,2,2)$. The conditional test which rejects for outcomes $(4,2,0)$ and $(3,3,0)$ is Schur-convex yet we show that the conditional power function is *not* Schur-convex. Recognize that $(1, a, a)$ majorizes $((1 + a)/2, (1 + a)/2, a)$ in the parameter space and that for small a the power at $((1 + a)/2, (1 + a)/2, a) \to 1$ as $a \to 0$. This gives rise to many cases where $\mathbf{q}_1 \preccurlyeq \mathbf{q}_2$, $\mathbf{q}' = (p_1, p_2, p_3)$, yet $\beta(\mathbf{q}_1) > \beta(\mathbf{q}_2)$, where $\beta(\mathbf{q})$ is power. For instance, if $\mathbf{q}_2' = (0.88, 0.01, 0.01)$, $\beta(\mathbf{q}_2) = 0.4254$ and if $\mathbf{q}_1' = (0.45, 0.44, 0.01)$, $\beta(\mathbf{q}_1) = 0.9065$. If $\mathbf{q}_2' = (1.00, 0.01, 0.01)$, $\beta(\mathbf{q}_2) = 0.4286$ while $\mathbf{q}_1' = (0.5005, 0.5005, 0.01)$, $\beta(\mathbf{q}_1) = 0.9241$.

Acknowledgments. The method of proof used in Section 4 was supplied by Michael Perlman, who read an earlier version of the paper. Michael's elegant method replaced a proof that was lengthy and difficult. The authors are grateful to him for his contribution.

REFERENCES

COHEN, A. and SACKROWITZ, H. B. (1975). Unbiasedness of the chi-square, likelihood ratio, and other goodness of fit tests for the equal cell case. *Ann. Statist.* **3** 959–964.

COHEN, A., SACKROWITZ, H. B. and STRAWDERMAN, W. E. (1985). Multivariate locally most powerful unbiased tests. In *Multivariate Analysis VI* (P. R. Krishnaiah, ed.) 121–144. North-Holland, Amsterdam.

COHEN, A. and STRAWDERMAN, W. E. (1971). Unbiasedness of tests for homogeneity of variances. *Ann. Math. Statist.* **42** 355–360.

EFRON, B. (1965). Increasing properties of Pólya frequency functions. *Ann. Math. Statist.* **36** 272–279.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic, New York.

NEVIUS, S. E., PROSCHAN, F. and SETHURAMAN, J. (1977). Schur functions in statistics. II. Stochastic majorization. *Ann. Statist.* **5** 263–273.

PERLMAN, M. D. and OLKIN, I. (1980). Unbiasedness of invariant tests for MANOVA and other multivariate problems. *Ann. Statist.* **8** 1326–1341.

DEPARTMENT OF STATISTICS
HILL CENTER FOR THE
MATHEMATICAL SCIENCES
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903