SOME BASIC THEOREMS OF DISTRIBUTION-FREE STATISTICS[1]

By C. B. Bell[2]

University of Washington and San Diego State College

1. Introduction and summary. The object of this article is the analysis of the concept of distribution-freeness. In Section 2 the common types of distribution-free (DF) statistics are defined and related. In Section 3 a generalization of the invariance principle is employed to construct DF statistics for classes of cpfs generated by groups of transformations of the sample space. It is shown in Section 4 that a function of several DF statistics is not necessarily DF, but that independence guarantees the desired result. The existence of a DF statistic is shown, in Section 5, to be equivalent to the existence of a suitable partition of the sample space. Several open problems are mentioned in Section 6, the final section.

2. Preliminaries. In general the notation of this paper is that of [2]. Ω will denote a class of cpfs on R_n ; $\Omega'(n)$ will denote the nth power class of the class Ω' of cpfs on R_1 ; and Ω_2 and Ω_2^* will denote, respectively, the class of all continuous cpfs on R_1 and the class of all strictly monotone continuous cpfs on R_1.

Definition 3.1. (i) A mapping $T(\cdot, \cdot)$ of $\Omega \times R_n$ into R_1 is called a statistic with respect to (wrt) the class Ω of cpfs on R_n if for each F in Ω, $T_F(\cdot) = T(F, \cdot)$ is $\mathcal{B}_n$-measurable. (ii) T_F is called the F-marginal statistic of T.

It is worthwhile to distinguish five types of distribution-free statistics relative to a given class Ω.

Definition 3.2. A statistic T is said to be of the given type wrt Ω if it satisfies the indicated DF conditions below:

(i) distribution-free (DF)	(A)
(ii) nonparametric distribution-free (NPDF)	(A), (B)
(iii) structure (d_n')	(C)
(iv) structure (d_n)	(C), (D)
(v) strongly distribution-free (SDF)	(E)

DF *Conditions.*

(A) There exists on R_1 one cpf $Q = Q_T$ such that for each F in Ω and real γ, $P_F\{x \,\varepsilon\, R_n : T(F, x) \leqq \gamma\} = Q_T(\gamma)$. [$Q_T$ is called the cpf of T.]

(B) $T_F = T_G$ for all F and G in Ω.

(C) $\Omega = \Omega'(n)$ for some class Ω' of cpfs on R_1 and

$$T_F(x_1, \cdots, x_n) = \Phi[F(x_1), \cdots, F(x_n)].$$

(D) The function Φ of condition (C) is symmetric.

Received 21 August 1963.

[1] The research on this paper was supported by the National Science Foundation through Grant NSF G 12973.

[2] Now at San Diego State College.

(E) $\Omega = \Omega'(n)$ for some class $\Omega' \subset \Omega_2^*$; and for each F and G in Ω' there exists on R_1 a cpf $Q_T(FG^{-1}; \cdot)$, depending only on FG^{-1}, such that $P_{F^{(n)}}\{x \varepsilon R_n : T(G^{(n)}, x) \leqq \gamma\} = Q_T(FG^{-1}; \gamma)$ for all real γ.

It is immediate from the definitions that

LEMMA 2.1.

(i) T *is* DF wrt Ω iff *for each* F *and* G *in* Ω *the measures* $P_F T_F^{-1} = P_G T_G^{-1}$ on $\mathfrak{B}_1$.

(ii) T *is* SDF wrt $\Omega'(n) \subset \Omega_2^*(n)$ iff $P_{F^{(n)}} T_{G^{(n)}}^{-1} = P_{H^{(n)}} T_{J^{(n)}}^{-1}$ *on* $\mathfrak{B}_1$ *whenever* F, G, H *and* J *are in* Ω' *and* $FG^{-1} = HJ^{-1}$.

(iii) *The following three conditions are equivalent*: (a) T *is* NPDF wrt Ω. (b) *Each* $T_F^{-1}(\mathfrak{B}_1)$ *is a σ-subalgebra of* $\mathfrak{B}_n$ *and is similar* wrt Ω; *and* $T_F^{-1}(A) = T_G^{-1}(A)$ *for all* F *and* G *in* Ω *and all* A *in* $\mathfrak{B}_1$. (c) *There exists a single measurable function* W *such that* $T_F = T_G = W$ *and* $P_F W^{-1} = P_G W^{-1}$ *on* $\mathfrak{B}_1$ *for all* F *and* G *in* Ω.

(iv) *Each statistic* SDF wrt $\Omega'(n)$ *is also* DF wrt $\Omega(n)$; *and each statistic* NPDF wrt Ω *is also* DF wrt Ω.

Further, Birnbaum, Rubin and Bell [2], [3] have essentially proved that the SDF and structure (d_n) statistics are equivalent in the following sense.

THEOREM 2.2. *If* $\Omega' \subset \Omega_2^*$, *then*

(i) *each statistic* T *of structure* (d_n') wrt $\Omega'(n)$ *is* SDF wrt $\Omega'(n)$;

(ii) *if* T *is* SDF wrt $\Omega'(n)$, *then there exists a statistic* S *of structure* (d_n') *such that* $P_F T_G^{-1} = P_F S_G^{-1}$ *for all* F *and* G *in* Ω; *and*

(iii) *if, further,* Ω *is symmetrically complete* (*of order* n), *then* T *is symmetric and* SDF wrt $\Omega'(n)$ iff T *is of structure* (d_n).

With these preliminaries, it is now possible to treat the generation, composition and existence of the various types of DF statistics.

3. Generation of DF statistics. A basic generation result is obtained by considering an application of the invariance principle related to the paper of Lukacs [6]; and illustrated by the following example.

EXAMPLE 3.1. If F_0 is an arbitrary cpf on R_1, let $F_{a,b}(x) = F_0(b^{-1}(x - a))$; $\Omega' = \{F_{a,b} : b > 0\}$; $T(F_{a,b}; x) = b^{-1}(x - a)$; and $S(F_{a,b}; x_1, x_2, x_3, x_4) = [x_1 - x_2][x_3 - x_4]^{-1}$. Clearly, then, T is DF wrt Ω' and $Q_T = F_0$. Further, S is NPDF wrt $\Omega'(4)$ if $P_{F_0}(x_3 = x_4) = 0$.

The immediate extension of the above example entails the replacing of the affine group by an arbitrary group of 1-1 transformations of R_n onto R_n.

DEFINITION 3.1. If F is an arbitrary cpf on R_n and $\mathcal{G} = [g]$ is an arbitrary group of 1-1 transformations of R_n onto R_n,

(a) F_g is the cpf on R_n induced by the probability measure $P_F g^{-1}$ on $\mathfrak{B}_n$; and

(b) $\Omega(F, \mathcal{G}) = [F_g : g \varepsilon \mathcal{G}]$.

It is now easy to construct a statistic DF wrt the generated class $\Omega(F, \mathcal{G})$.

THEOREM 3.1. *If* F *and* $\mathcal{G}$ *are as in Definition* 3.1, *then*

(i) *for each* $\mathfrak{B}_n$*-measurable real valued function* Ψ *on* R_n, T *is* DF wrt $\Omega(F, \mathcal{G})$ *if* $T(F_g, x) = \Psi g^{-1}(x)$ *for all* g *and* x; *and* $P_{Q_T} = P_F \Psi^{-1}$.

(ii) *If, further,* $\mathfrak{D}$ *is a sub-σ-algebra of* $\mathfrak{B}_n$ *and is invariant under* $\mathcal{G}$, *then each* $\mathfrak{D}$*-measurable real-valued function* T *is* NPDF wrt $\Omega(F, \mathcal{G})$.

PROOF.

(i) For each A in $\mathfrak{B}_1$ and $g \varepsilon \mathcal{G}$, one has for $H = F_g$, $P_H[\text{x}: T(H, x) \varepsilon A] = P_H T_H^{-1}(A) = P_H(g\Psi^{-1}(A)) = P_F(g^{-1}g\Psi^{-1}(A)) = P_F(\Psi^{-1}(A))$, which is the desired result.

(ii) Clearly, if $\mathfrak{D}$ is invariant under $\mathcal{G}$, $\mathfrak{D}$ is similar wrt $\Omega(F, \mathcal{G})$; and the result follows from Lemma 2.1.

It is now feasible to consider functions of DF statistics.

4. Functions of DF statistics. It is immediate from the definitions that a measurable function of a DF statistic is again a DF statistic of the same type. That this is not necessarily true for functions of more than one DF statistic is illustrated below.

EXAMPLE 4.1.

(i) Let a, b, c, and d be four different points of R_n such that $P_\alpha(\{x\}) = r - \alpha$ if $x = a$; $= r + \alpha$ if $x = b$; $= s + \alpha$ if $x = c$; $= s - \alpha$ if $x = d$, where $0 < r, s$ and $r + s = \frac{1}{2}$; and $\Omega = \{F_\alpha : 0 < \alpha < \min(r, s)\}$. For all F in Ω, let $V(F, x) = 1$ if $x = a$ or c; $=0$ otherwise; $U(F, x) = 1$ if $x = c$ or d; $= 0$ otherwise. Therefore, $P_F(V = 1) = r + s = \frac{1}{2}$ and $P_F(U = 1) = 2s$ for all F in Ω, and, hence, both U and V are NPDF wrt Ω. However, $P_\alpha(U + V = 1) = P_\alpha(\{a, d\}) = r + s - 2\alpha = \frac{1}{2} - 2\alpha$. Consequently, $U + V$ is not DF wrt Ω.

(ii) If one now changes the probability measures in (i) above in such a manner that $P_\alpha(\{x\}) = r - \alpha$ if $x = a$; $= r$ if $x = b$; $= s$ if $x = c$; $= s + \alpha$ if $x = d$, then neither U nor V is DF wrt Ω, but $U + V$ is NPDF wrt Ω.

In view of the pathological situation indicated by the preceding example one now seeks restrictions which guarantee that a function of a set of DF statistics be itself a DF statistic. For a variety of reasons, not the least of which are suggested by the k-sample statistics of Fisz [4] and Kiefer [5], and the well-known Studentization results, one is lead to the requirement of independence. Before investigating this restriction it is worthwhile to give a precise definition of the independence of a DF statistic wrt a class Ω.

DEFINITION 4.1. A family $\{T_r, r \varepsilon \mathfrak{L}\}$ of real-valued functions on $\Omega \times R_n$ is said to be mutually independent (MI) wrt the class Ω of cpfs on R_n if for each finite subset K of $\mathfrak{L}$, and each collection $\{D_r : r \varepsilon K\}$ of subsets of $\mathfrak{B}_1$

$$P_F(\bigcap_{r \varepsilon K}\{x: T_r(F, x) \varepsilon D_r\}) = \prod_{r \varepsilon K} P_F(\{x: T_r(F, x) \varepsilon D_r\}).$$

From the definitions, Lemma 2.1 and elementary product measure considerations, one can establish

LEMMA 4.1.

(a) *The following two conditions are equivalent*:

(i) $\{T_1, T_2, \cdots, T_k\}$ *is a family of statistics* MIDF wrt Ω.

(ii) *For each* F, G *in* Ω *and* $B \varepsilon \mathfrak{B}_k$

$$P_F\{x: [T_1(F, x), \cdots, T_k(F, x)] \varepsilon B\} = \lambda_G(B) = \lambda_F(B),$$

where λ_H is the product measure $\times_1^k P_H(T_i)_H^{-1}$ on $\mathfrak{B}_k$.

(b) Result (a) holds also for NPDF statistics.

One can similarly establish

LEMMA 4.2. *The following two conditions are equivalent*:

(i) $\{S_1, S_2, \cdots, S_k\}$ *is a family of statistics* MISDF wrt $\Omega'(n)$.

(ii) *Whenever* $B \varepsilon \mathcal{B}_k$; $F_1, G_1, F_2, G_2 \varepsilon \Omega'$ *and* $F_1G_1^{-1} = F_2G_2^{-1}$, *then* $P_{F_1}^{(n)} \{x: [S_1(G_1^{(n)}, x), \cdots, S_k(G_1^{(n)}, x)] \varepsilon B\} = \lambda_{F_1,G_1}(B) = \lambda_{F_2,G_2}(B)$, *where* $\lambda_{H,J}$ *is the product measure* $\times_1^k P_H(T_i)_J^{-1}$ *on* $\mathcal{B}_k$.

Employing these two lemmas and Theorem 2.2 one proves that functions of MIDF statistics are themselves DF, and that in the SDF case independence is not required if $\Omega(n)$ is symmetrically complete.

THEOREM 4.3.

(i) *If* $\{T_1, \cdots, T_k\}$ *is a family of statistics* MI DF wrt Ω *then for each real-valued measurable function* U *on* R_k, $V = U(T_1, \cdots, T_k)$ *is* DF wrt Ω. (*The result if valid if in* (i) "DF" *is replaced by* (ii) "NPDF" *or* (iii) "SDF.")

(iv) *If* Ω' *is symmetrically complete and* $T_1, \cdots, T_k$ *are each symmetric and* SDF wrt $\Omega'(n)$, *then for each measurable real-valued function* V, $V(T_1, \cdots, T_k)$ *is symmetric and* SDF wrt $\Omega'(n)$.

PROOF. (iv) follows immediately from Theorem 2.2.

(i) For each $F, G \varepsilon \Omega$ and $B \varepsilon \mathcal{B}_1$,

$$\begin{aligned} P_F V_F^{-1}(B) &= P_F\{x: V(F, x) \varepsilon B\} \\ &= P_F\{x: [T_1(F, x), \cdots, T_k(F, x)] \varepsilon U^{-1}(B)\} \\ &= \lambda_G(U^{-1}(B)) = \lambda_F(U^{-1}(B)) \end{aligned}$$

by Lemma 4.1. Therefore, $P_F V_F^{-1} = P_G V_G^{-1}$ and (i) follows from Lemma 2.1.

(ii) This follows immediately from (i) since in the NPDF case each $(T_i)_F = (T_i)_G = T_i$.

(iii) For the SDF case by definition Ω is of the form $\Omega'(n)$. Hence, whenever F, G, H and $J \varepsilon \Omega'$ and $FG^{-1} = HJ^{-1}$, then $P_F V_G^{-1}(B) = P_F\{x: [T_1(G, x), \cdots, T_k(G, x)] \varepsilon U^{-1}(B)\} = \lambda_{F,G}(U^{-1}(B)) = \lambda_{H,J}(U^{-1}(B)) = P_H V_J^{-1}(B)$. The result follows from Lemma 2.1.

One turns now to considerations of existence.

5. Existence of DF statistics. Up to this point it has been seen [Theorem 3.1] that there exist statistics DF wrt classes generated by groups of transformations. Necessary and sufficient conditions for the existence of statistics DF wrt an arbitrary class Ω follow readily from the two well known results below.

(α) If G is a cpf on R_1, then there exist (i) a continuous cpf F on R_1; (ii) Probabilities $\{p, p_1, \cdots\}$ with $\sum_i p_i = 1$; and (iii) real numbers $\{a_1, a_2, \cdots\}$ with the property that for all real γ, $G(\gamma) = (1 - p)F(\gamma) + p\sum_i p_i \epsilon(\gamma - a_i)$, where ϵ is the degenerate cpf with mass 1 at $x = 0$.

(β) If $(Z, \mathfrak{D}, P)$ is a NA (nonatomic) probability space, then for each cpf F on R_1, there exists a $\mathfrak{D}$-measurable function h on Z such that $Ph^{-1} = P_F$ on $\mathcal{B}_1$, i.e. h has cpf F.

THEOREM 5.1.

(i) *For each* $\Omega' \subset \Omega_2^*$ *and each* cpf J *on* R_1, *there exists a statistic* T *of structure* (d_n') *and, hence, SDF* wrt $\Omega'(n)$ *having* cpf $Q_T = J$.

(ii) *There exists a statistic* T DF wrt Ω *and having* cpf $Q_T = G$ *with continuous part* $(1 - p)F$ *and discrete part* $p \sum p_i \epsilon(\cdot - a_i)$ iff *for each* J *in* Ω *there exists a partition* $\{A_0(J), A_1(J), \cdots\}$ *of* R_n *such that*

(a) $P(A_i(J)) = pp_i$ *for* $i \geqq 1$; *and*

(b) P_J *is* NA *on* $A_0(J)$.

(iii) *There exists a statistic* T NPDF wrt Ω *and having* cpf $Q_T = G$ *of the type in* (ii) iff *there exists a* σ*-algebra* $\mathfrak{D}$ *contained in* $\mathcal{B}_n$ *and a similar partition* $\{A_0, A_1, \cdots\}$ *of* R_n *such that*

(a) $P_F(A_i) = pp_i$ *for* $i \geqq 1$ (*and all* $F \varepsilon \Omega$); *and*

(b) $A_0 \cap \mathfrak{D}$ *is a* NA σ*-ring of sets similar* wrt Ω.

PROOF.

(i) From the DF-ness of $G(x)$ and elementary product space considerations it follows that $S_G(x_1, \cdots, x_n) = [G(x_1), \cdots, G(x_n)]$ is a mapping such that $P_G^{(n)} S_G^{-1}$ is equal to Lebesgue measure λ_n on the Borel sets $\mathcal{B}_n(I)$ of the n-dimensional unit hypercube I_n. Result (β) then, guarantees the existence of a real-valued measurable function h on I_n such that $P_G^{(n)} S_G^{-1} h^{-1} = \lambda_n h^{-1} = P_J$ on $\mathcal{B}_1$.

Now one defines $T(G; x_1, \cdots, x_n) = hS_G(x_1, \cdots, x_n) = h[G(x_1), \cdots, G(x_n)]$. Since T has structure (d_n'), Theorem 2.2 guarantees the desired result.

(ii) (*Necessity*). If T is DF wrt Ω and $Q_T = G$, then for each H in Ω and $i \geqq 1$ let $A_i(H) = T_H^{-1}(\{a_i\})$, where a_i is a real number such that $P_H\{T_H = a_i\} > 0$, and let $A_0(H) = R_n - \bigcup_{i \geqq 1} A_i(H)$.

Clearly, $P_H(A_i(H)) = pp_i$ for $i \geqq 1$ and P_H is NA on $A_0(H)$.

(ii) (*Sufficiency*). For each J in Ω, one concludes, on applying result (B) to the conditional probability space induced by P_J on $A_0(J) \cap \mathcal{B}_n$, that there is a measurable and real-valued function S_J on $A_0(J)$ such that $P_J\{x: S_J(x) \leqq \gamma; x \varepsilon A_0(J)\} = (1 - p)F(\gamma)$ for all real γ.

Therefore, if $\{a_1, a_2, \cdots\}$ is a set of distinct real numbers and $T(H, x) = S_H(x)$ for $x \varepsilon A_0(H)$; $= a_i$ for $x \varepsilon A_i(H) (i \geqq 1)$, T is a statistic DF wrt Ω and having cpf G, since

$$P_H\{x: T(H, x) \leqq \gamma\} = P_H\{T_H(x) \leqq \gamma, x \varepsilon A_0\} + \sum_{i \geqq 1} P_H\{x \varepsilon A_i, T_H(x) \leqq \gamma\}$$

$$= (1 - p)F(\gamma) + p \sum_{i \geqq 1} p_i \epsilon(\gamma - a_i).$$

(iii) (*Necessity*). If T is NPDF wrt Ω and $Q_T = G$, then by Lemma 2.1 the desired sets and σ-algebra are

$$A_i = T^{-1}(\{a_i\}) \text{ for } i \geqq 1; \quad A_0 = R_n - \bigcup_i A_i; \text{ and } \mathfrak{D} = T^{-1}(\mathcal{B}_1).$$

(iii) (*Sufficiency*). This follows from the proof of (ii) in replacing "$\mathcal{B}_n$" by "$\mathfrak{D}$" and "$A_0(J)$" by "A_0." From this basic existence theorem follow several useful corollaries.

Since, if $T(F, x) = x$, T is DF wrt each Ω containing exactly one cpf; and since $\{A_0(F) = \phi, A_1(F) = R_n\}$ is a partition, one has

COROLLARY 5.2.

(i) *For each statistic T, there exists a class Ω such that T is* DF [NPDF, SDF] wrt Ω.

(ii) *The trivial (constant) statistic $T \equiv k$ is* DF *and* NPDF wrt *each* Ω, *and* SDF wrt *each* $\Omega \subset \Omega_2^*(n)$.

(iii) *If Ω contains one degenerate* cpf ϵ_1, *then the trivial statistic $T \equiv k$ is the only statistic* DF wrt Ω.

The next corollary answers the existence question for the nontrivial cases of DF and NPDF statistics. [Of course, Theorem 5.1 (i) summarizes the existence situation for SDF statistics.]

COROLLARY 5.3.

(i) *There exists a nontrivial,* i.e. *nonconstant, statistic* DF wrt Ω iff *there exists* $0 < p < 1$ *and a collection* $\{A(H): H \varepsilon \Omega\}$ *of sets of* $\mathcal{B}_n$ *such that* $P_H(A(H)) = p$ *for all H in Ω.*

(ii) *There exists a nontrivial statistic T* NPDF wrt Ω iff *there exists a set A similar* wrt Ω *and such that* $0 < P_H(A) < 1$ *for some H in Ω.*

From the preceding two lemmas and Example 4.1 one sees that the existence of a DF statistic is equivalent to the existence of a suitable partition of the sample space, and that for a degenerate cpf no nontrivial partition exists. Since each discrete cpf has a decomposition into degenerate cpfs one suspects that there will be some existence difficulties in this case. On the other hand, from (β) and from other considerations one sees that for continuous cpfs there is maximum flexibility and, in particular, there exists an infinitude of suitable partitions. More precisely, it is immediate that

COROLLARY 5.4.

(i) Ω *contains only continuous* cpfs iff *for each* cpf *H on R_1 there exists a statistic T* DF wrt Ω *and having* cpf $Q_T = H$.

(ii) *The following two conditions are equivalent:*

(a) *For each* cpf *J on R_1, there exists a statistic T* NPDF wrt Ω *with* $Q_T = J$; *and*

(b) $\mathcal{B}_n$ *contains a σ-algebra $\mathfrak{D}$ similar* wrt Ω *and such that P_F is* NA *on $\mathfrak{D}$ for some F in Ω.*

Having considered the generation, structure, composition and existence of real-valued DF statistics, it is worthwhile making some remarks concerning extensions, applications and open problems.

6. Concluding remarks, open problems. Before stating some open problems it should be re-emphasized that the DF property is relative rather than absolute; that DF concepts relative to the normals have been employed throughout classical statistics; that a DF statistic is not a single function but a family of marginal functions; and that the DF and NP properties are not equivalent.

Open Problems.

(1) How should one choose the Φ and Ψ of Theorems 2.2 and 3.1, respectively, in order that the resulting statistic T satisfies reasonable goodness criteria and has a readily computable cpf Q_T? [See Theorem 3.1 (i).]

(2) Which of Theorems 2.2, 3.1, 4.3 and 5.1 are valid for multivariate classes, vector-valued statistics and stochastic processes?

(3) How is Saunders' [7] "ample" estimate of a cpf related to the above DF statistics? [Essentially, a cpf $\hat{F}(\cdot; x_1, \cdots, x_n)$ is an ample estimate wrt Ω, if under F the stochastic process $\{\hat{F}F^{-1}(\xi): 0 \leqq \xi \leqq 1\}$ has a law $\tilde{Q}(\hat{F})$ which does not depend on F for F in Ω. It is easily established that the sample cpf F_n is an ample estimate for Ω_2 in the univariate random sampling case; and Saunders has proved that the cpf estimate based on maximum likelihood estimates is ample for the class of multivariate normals.]

(4) What are necessary and sufficient conditions that (a) the limit of a sequence of DF statistics is DF; and (b) each measurable function of a family of statistics is DF? (S. Pathak has shown that one need only consider linear combinations of DF statistics.)

REFERENCES

[1] Basu, D. (1958). On statistics independent of sufficient statistics. *Sankhyā* **20** 223–226.

[2] Bell, C. B. (1960). On the structure of distribution-free statistics. *Ann. Math. Statist.* **31** 703–709.

[3] Birnbaum, Z. W. and Rubin, H. (1954). On distribution-free statistics. *Ann. Math. Statist.* **25** 593–598.

[4] Fisz, M. (1960). Some non-parametric tests for the k-sample problem. *Colloq. Math.* **7** 289–296.

[5] Kiefer, J. (1959). K-sample analogues of the Kolmogorov-Smirnov and Cramér-v. Mises tests. *Ann. Math. Statist.* **30** 420–447.

[6] Lukacs, Eugene (1957). Certains tests independants de la distribution initiale. *Ann. Inst. H. Poincaré.* **15** 252–265.

[7] Saunders, Sam C. (1960). On the use of a distribution-free property in determining a transformation of one variate such that it will exceed another with a given probability. Math. Note No. 210, *Math. Res. Lab., Boeing Sci. Res. Lab.*

[8] Uzawa, H. (1960). Locally most powerful rank tests for two-sample problems. *Ann. Math. Statist.* **31** 685–703.

[9] Watson, G. S. (1957). Sufficient statistics, similar regions and distribution-free tests. *J. Roy. Statist. Soc. Ser. B* **19** 262–267.