ABSTRACTS OF PAPERS

(Abstract of a paper presented at the Western Regional meeting, Missoula, Montana, June 15–17, 1967. Additional abstracts appeared in earlier issues.)

18. Power approximations in multivariate exponential families. BRADLEY EFRON and DONALD TRUAX, Stanford University and University of Oregon.

Let $X_1, X_2, \cdots, X_n$ be independent and identically distributed random vectors from a d-dimensional exponential family, $dP_\theta(X) = e^{\theta' X - \psi(\theta)}\, d\mu(X)$, $\theta \varepsilon \boldsymbol{\theta}$. Another exponential family $d\tilde{P}_\theta(X) = e^{\theta' X - \tilde{\psi}(\theta)}\, d\tilde{\mu}(X)$ is said to *agree* with $dP_\theta(X)$ at $\theta = \theta_0$ (interior to $\boldsymbol{\theta}$) if $\theta_0 \varepsilon$ interior of, $\tilde{\theta}\, E_{\theta_0} X = \tilde{E}_{\theta_0} X$, and $\text{Cov}_{\theta_0}(X) = \widetilde{\text{Cov}}_{\theta_0}(X)$, where $\tilde{E}_{\theta_0}$ and $\widetilde{\text{Cov}}_{\theta_0}$ indicate the expectation and covariance of a vector from $d\tilde{P}_{\theta_0}$. Let C be a set close to $E_{\theta_0} X = \tilde{E}_{\theta_0} X$ in the sense that C is contained within a fixed sphere S_K: $\| X - E_{\theta_0} X \| \leqq K/n^{\frac{1}{2}}$ for some $K > 0$ not depending on n. It is shown that under certain regularity conditions,

$$P_\theta(\bar{X}_n \varepsilon C)/\tilde{P}_\theta(\bar{X}_n \varepsilon C) = e^{-n[\psi(\theta) - \tilde{\psi}(\theta)]}[1 + o_n(1)],$$

where $o_n(1)$ goes to zero uniformly for $\theta \varepsilon \boldsymbol{\theta} \cap \tilde{\boldsymbol{\theta}}$ and sets C as above. The usefulness of this approximation for both theoretical and numerical evaluation of power functions in hypothesis testing of exponential families is illustrated. The power of the likelihood ratio test of various null hypotheses is discussed, with particular reference to accurate "large deviation" results. The situation where θ_0 is on the boundary of $\boldsymbol{\theta}$ is also treated, and a theorem very similar to the above derived. (Received 9 June 1967.)

(Abstracts of papers to be presented at the Annual meeting, Washington, D. C., December 27–30, 1967. Additional abstracts appeared in the June and August issues and will appear in future issues)

5. Tables of quadratic forms in normal variables. N. L. JOHNSON and SAMUEL KOTZ, University of North Carolina, Chapel Hill, and Temple University.

Using formulae (18), (71), (74) and (76) derived in a paper published in the *Ann. Math. Statist.* **38** (1967) 823–837 detailed tables are calculated giving the probability density function and distribution function of $X = \sum_{j=1}^{n} \alpha_j U_j^2$ where the U_j's are independent unit normal variables and the α_j's are positive constants with $\sum_{j=1}^{n} \alpha_j = n$. The tables can be regarded as an extension to cases $n = 4, 5$, and 6 of G. Marsaglia's tables for $n = 2, 3$ (Report No. DI-82-0015-1, Math. Note 213, *Boeing Scientific Research Laboratories*, Seattle, Washington, August, 1960 and of H. Solomon's tables (Technical Report No. 45, Applied Math. and Statist. Laboratories, Stanford University, 1960). Some suggestions for interpolation with respect to α's are discussed. (Received 7 July 1967.)

6. Inequalities and tolerance limits for s-ordered distributions. MICHAEL J. LAWRENCE, Corning Glass Works. (Introduced by B. K. Shah.)

Define $F <_s G$ ($F <_r G$) if F and G have the same median, say the origin and $G^{-1}F(x)$ is concave-convex about the origin ($G^{-1}F(x)/x$ is increasing (decreasing) in x positive (negative)). Conservative tolerance limits are derived for distributions which are s-ordered with respect to the Laplace distribution. These are especially reasonable for mensuration data. In addition, many inequalities concerning combinations of order statistics are obtained. These results are useful in robustness studies of tolerance limits, estimates and statistical tests derived for specified distributions such as the normal distribution. Some examples are given. (Received 14 June 1967.)

7. Rank order tests for the homogeneity of dispersion matrices. PRANAB KUMAR SEN and MADAN LAL PURI, University of North Carolina, Chapel Hill, and Courant Institute of Mathematical Sciences, New York University.

In this paper we offer nonparametric tests for (i) the homogeneity of dispersion matrices assuming the identity of location parameters, (ii) the homogeneity of dispersion matrices without assuming the identity of location parameters, and (iii) the homogeneity of location parameters as well as the dispersion matrices, all against appropriate classes of alternatives. Some strictly distribution free permutation tests are developed and their asymptotic properties are studied with the aid of a generalization of the celebrated Chernoff-Savage Theorem (*Ann. Math. Statist.* **29** (1958) 972–994). Finally the asymptotic efficiencies of the proposed procedures relative to their likelihood ratio competitors (cf. T. W. Anderson, *Introduction to Multivariate Statistical Analysis* (1965), Chapter 10) are studied. The paper is in the spirit of Puri and Sen (1966), On a class of multivariate multiSample rank order tests, *Sankhyā, Ser. A* **28** 353–376. Sen and Puri (1967), On the theory of rank order tests for location in the multivariate one sample problem, *Ann. Math. Statist.*, **38**. Puri and Sen (1966), On a class of rank order tests for independence in multivariate distributions, Technical Report of the Courant Institute of Mathematical Sciences, July, 1966; Abstract *Ann. Math. Statist.* **37** 1863. (Received 22 May 1967.)

8. A decision theoretic approach to the problem of interval estimation. R. S. VALAND, E. R. Squibb and Sons. (Introduced by J. A. Hartigan.)

Problem of interval estimation has been studied within the framework of decision theory using the principle of invariance. The invariance principal is available only in decision problems which are invariant under certain transformations (see Blackwell and Girshick, *Theory of Games and Statistical Decisions* (1954)). Here I find a form of a best invariant interval estimate in the problem of estimating a location parameter with loss function as defined by $L(\theta, (a, b)) = h(a - \theta, b - \theta) + 1 - I_{(a, b)}(\theta)$, where $h(a, b)$ is defined on $\{(a, b); a \leqq b\}$. I find that if $E_0L(0, (X + b_1, X + b_2))$ exists for some (b_1, b_2), and if there exists a (b_1^0, b_2^0) such that

$$E_0L(0, (X + b_1^0, X + b_2^0)) = \inf_{(b_1,b_2)} E_0L(0, (X + b_1, X + b_2)),$$

where the infimum is taken over all (b_1, b_2) for which $E_0L(0, (X + b_1, X + b_2))$ exists, then $d_{b^0}(X) = (X + b_1^0, X + b_2^0)$ is a best invariant rule, and has a constant risk equal to $E_0L(0, (X + b_1^0, X + b_2^0))$. The condition under which a best invariant interval estimate of a location parameter is minimax is (A) if for every $\epsilon > 0$, there exists an N such that $\int_{-N}^{N} L(0, (X + b_1, X + b_2))\, dF(X/0) \geqq R0 + \epsilon$ for all b_1 and b_2 $(b_1 \leqq b_2)$ (see Kudô, H., Natural Science Report, *Ochnomizu Univ.* **6** (1955) 31–73). The condition (A) is satisfied if (1) $h(a, b)$ is bounded; (2) $h(a, b)$ and F are continuous; (3) $h(a, b) \to \infty$ as $b \to \infty$ or $a \to \infty$; and (4) $h(a, b)$ is non-decreasing in b for fixed a and is non-increasing in a for fixed b. It should be noted that for a loss function $L(\theta, (a, b)) = \max(\frac{1}{2}(b + a) - \theta, 0) + 1 - I(a, b)^{(\theta)}$, and for a distribution function $f(X, \theta) = [(X - \theta)(X - \theta - 1)]^{-1}$, $X = \theta + 1, \theta + 2, \cdots$, a best invariant interval estimate $d(X) = (X + C_a|X|, X - C_b|X|)$ exists but is not minimax $(\frac{1}{2}(C_a + C_b) > 1)$ (see Kiefer, *Ann. Math. Statist.* **28** (1957)

573–601; Blackwell Girshick, *Theory of Games and Statistical Decisions* (1954)). Finally, using proper transformations (see Lehmann, *Testing Statistical Hypothesis* (1959); Blackwell, *Ann. Math. Statist.* **22** (1951) 393–407) the corresponding results can be shown for the scale parameter of a distribution. (Received 5 July 1967.)

(Abstracts of papers not connected with any meeting of the Institute)

1. A distribution free version of the Smirnov p sample test. Peter J. Bickel, University of California, Berkeley.

Let $\mathbf{X}_i$, $i \geqq 1$, $\mathbf{Y}_j$, $j \geqq 1$, be two sequences of independent p vectors distributed according to cumulative distribution functions F and G respectively. Let F_m, G_n, H_N be the sample cumulative distributions of $\mathbf{X}_i$, $1 \leqq i \leqq m$, $\mathbf{Y}_j$, $1 \leqq j \leqq n$, and the pooled sample of $N = m + n$ observations, respectively. We show, Theorem. *If* $F = G$, $P_F\,[\mathfrak{L}(g(F_m) \mid H_N) \to \mathfrak{L}_F] = 1$ *where* $g(\cdot)$ *is a suitable functional defined on the space of bounded functions on* R^p, $\mathfrak{L}(\cdot \mid H_N)$ *is the (random) conditional law of* $g(F_m)$ *given* H_N *and* $\mathfrak{L}$ *is a limit law identified in terms of a given Gaussian process on* R^p. This result is applied to obtain a sequence of distribution free level α tests of the hypothesis $F = G$ consistent against all alternatives $F \neq G$. (Received 29 May 1967.)

2. Admissible Bayes procedures and classes of epsilon Bayes procedures for testing hypotheses in a multinomial distribution. Carl Morris, Stanford University.

The problem considered is the problem of testing a simple hypothesis $H_0 : p_i = p_i^0$, $i = 1, \cdots, k$, in a multinomial distribution with observations $N = (N_1, \cdots, N_k)$. It is shown for large k and near alternatives one only need consider quadratic forms in N. The constants for the quadratic form depend only on the first two moments of p_i under the prior distribution. A central limit theorem is proved which gives weak sufficient conditions on a_i and b_i for $\sum \{a_i N_i + b_i N_i^2\}$ to have an asymptotic normal distribution as $k \to \infty$. For large k it is proposed that to reject H_0 when $\sum N_i(N_i - 1)/p_i^0$ is large is minimax. Finally it is shown that the chi-square test, the likelihood ratio test and all quadratic forms are Bayes and admissible for fixed k and n. (Received 12 July 1967.)

3. A class of epsilon Bayes tests of a simple null hypotheses on many parameters in exponential families. Carl Morris, Stanford University.

$X_1, \cdots, X_k$ are observed as independent one-dimensional random variables, X_i being the sufficient statistic for a parameter θ_i in an exponential family. The problem of testing the hypothesis $H_0 : \theta_i = \theta_{0i}$ is considered when k is large. The usual hypothesis testing loss structure and near alternatives for moderate size and power are assumed. If $\epsilon > 0$ is given and k is large enough, it is shown that the Bayes test for a prior distribution π making the $\{\theta_i\}$ independent under H_1 has Bayes risk at most ϵ less than a test based on a non-negative definite quadratic form which rejects H_0 when $\sum \{a_i X_i + b_i X_i^2\} \geqq c$. The constants $\{a_i\}$ and $\{b_i\}$ are determined by the first two moments of $\{\theta_i\}$ under π, reducing the class of decision procedures under consideration. The approximate Bayes risk and location of the prior distribution under π is determined. The quadratic forms are shown to be proper Bayes for fixed k. For large k, ϵ-minimax tests are determined. Applications of the above results are made to the binomial, negative binomial, Poisson, normal and gamma distributions by introducing the concept of quadratic moment structure. Some large deviation theorems are proved for exponential families. (Received 12 July 1967.)

4. Sequential estimation of the mean of a distribution having a prescribed proportional closeness. R. C. SRIVASTAVA, Ohio State University.

Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed random variables each with mean μ and variance σ^2. Given $0 < \delta < 1$ and $0 < \alpha < 1$, the problem is to find an estimator $\hat{\mu}$ of μ such that

$$P_{\mu,\sigma^2}\{|\hat{\mu} - \mu| < \delta|\mu|\} \geqq \alpha \qquad (1)$$

for all μ, σ^2 $(\mu \neq 0)$. It is shown that there does not exist any fixed sample procedure which can satisfy (1). If the coefficient of variation $\lambda = \sigma/\mu$ of the given population is known, then the efficient fixed sample estimation procedure is to take n_0 observations and estimate μ by $\bar{x}_{n_0}$. Here n_0 is defined as: smallest integer $\geqq \lambda^2 a^2/\delta^2$ where a is defined by $(2\pi)^{-\frac{1}{2}} \int_{-a}^{a} \exp(-t^2/2)\, dt = \alpha$. However if λ is unknown, one can't attain the required property by a fixed sample procedure. In this paper, a sequential procedure is suggested and shown to possess these properties. This procedure is consistent and asymptotically efficient in the Chow-Robbins sense (*Ann. Math. Statist.* **36** (1965) 457–462). (Received 29 June 1967.)

5. An estimate of the density function. J. H. VENTER, Potchefstroom University.

Let $Y_1, \cdots, Y_n$ be an ordered independent sample from a density function f. Let $K_n(x)$ be the number of Y's $\leqq x$ and $\{r_n\}$ positive integers. The estimate $2r_n n^{-1}(Y_{K_n(x)+r_n} - Y_{K_n(x)-r_n+1})^{-1}$ of $f(x)$ is studied here. It is shown to be consistent for r_n of the form An^ν, $0 < \nu < 1$, if some mild conditions on f are satisfied. It is also shown that $(2r_n)^{\frac{1}{2}}(f_n(x)f(x)^{-1} - 1)$ has a $N(0, 1)$ limiting law under certain conditions on r_n and f. (Received 26 May 1967.)