The Annals of Mathematical Statistics
1970, Vol. 41, No. 1, 328–332

ABSTRACTS OF PAPERS

(*Abstracts of papers to be presented at the Central Regional meeting, Dallas, Texas, April* 8–10, 1970. *Additional abstracts will appear in future issues.*)

124-2. A class of conjugate prior distributions, and optimal allocation: a multivariate extension. RONNIE L. MORGAN, Oregon State University.

An extension to a paper entitled "A class of conjugate prior distributions, and optimal allocation" contributed by the author at the 1969 region Spring Meeting of the Joint Statistical Conference at Iowa City, Iowa is given here. A univariate one-parameter family of sampling distributions of the form $f(x \mid \tau) = \exp[x\tau - \theta(\tau)]f(x)$ was considered in the earlier paper. The multivariate family defined for a p-dimensional variate $\mathbf{x} = (x_1, x_2, \cdots, x_p)$ by $f(\mathbf{x} \mid \tau) = \exp[\sum_{i=1}^{m} \varphi_i(\mathbf{x})\tau_i - \theta(\tau)]f(\mathbf{x})$ is considered here. A class of conjugate priors of the form $g(\tau) \propto \exp[\mathbf{a}^t\tau - b\theta(\tau)]$ is assumed. Let $\theta_i(\tau) = \partial\theta(\tau)/\partial\tau_i$, and $\theta_{ij}(\tau) = \partial^2\theta(\tau)/\partial\tau_i\,\partial\tau_j$, then $\theta_i(\tau) = E[\varphi_i(\tilde{\mathbf{x}})]$ and $\theta_{ij}(\tau) = \text{Cov}[\varphi_i(\mathbf{x}), \varphi_j(\mathbf{x})]$. The following theorems are proved, under certain regularity conditions. THEOREM 1. $E[\theta_i(\tilde{\tau})] = a_i/b$. THEOREM 2. $E[\theta_{ij}(\tilde{\tau})] = b\,\text{Cov}[\theta_i(\tilde{\tau}), \theta_j(\tilde{\tau})]$. THEOREM 3. $E[\text{Cov}\{\theta_i(\tilde{\tau}), \theta_j(\tilde{\tau})\}] = b\,\text{Cov}[\theta_i(\tilde{\tau}), \theta_j(\tilde{\tau})]/(n+b)$. The expectation in Theorem 3 is with respect to the marginal distribution of $\mathbf{x}$ as determined by the assumed prior and Cov denotes the posterior covariance. The results of this paper are applied to several optimal allocation problems. In particular when application is made to optimal stratified sampling, the resulting solution is shown to be a generalization of the Neyman allocation formula. (Received October 14, 1969.)

124-3. Bulk queues with phase input. ASHA S. KAPADIA, Arthur D. Little, Inc.

In this paper the expected number of idle servers in the steady state has been obtained for the following three queues. (i) $P_{\{q_i\}}/P/1$, (ii) $P_{(K)}/P_{(N)}/1$, (iii) $P_{\{q_i\}}/M/k$. Here P represents the phase distribution (defined as the distribution of a random sum of exponential random variables).

$A_{\{q_i\}}/B/k$ refers to a queuing system which has interarrival distribution A, service time distribution B, number of servers k, batch arrival distribution $\{q_i\}\cdot(q_i =$ probability an arrival batch has i customers in it). $A_{(L)}/B_{(N)}/k$ represents batch arrivals of fixed size L and batch service of fixed size N. (Received October 30, 1969.)

(*Abstracts of papers to be presented at the Eastern Regional meeting, Chapel Hill, North Carolina, May* 5–7, 1970. *Additional abstracts will appear in future issues.*)

125-3. Queuing with general and special service. ASHA S. KAPADIA, Arthur D. Little, Inc.

Consider a system with k general and m special servers. Customers entering the system require either special or general service and they queue up in front of the special and general servers respectively. Whenever a general server is idle, he takes in a customer from the special queue, provided there are more than m customers in it. However, when a special server is idle, he cannot take in a general customer for service because special servers can perform special service only. The interarrival times of general and special customers have phase distribution (defined as the distribution of a random sum of exponential random variables) and the service time distribution for each customer is assumed exponential. It is interesting to observe that from the point of view of expected number of idle servers in the steady state, the system behaves as though the servers have no interaction where as in fact they sometimes do. The result is generalized to the case of

$(N-1)$ different types of special service. Also to the case of phase service time distribution. (Received October 30, 1969.)

(Abstracts of papers contributed by title)

70T-1. Linear sufficiency in survey-sampling. M. K. Ramakrishnan, Indian Statistical Institute.

Godambe [*J. Roy. Statist. Soc. Ser. B.* **28** (1966) 310–319] defined "linear sufficiency" for survey-sampling and has shown that, for any design p with fixed sample size n for which $p(s) > 0$ for all the $\binom{N}{n}$ samples, the unique unbiased, linear sufficient estimator for the population total which also satisfies the principle of censoring is given by $t = [\binom{N-1}{n-1}p(s)]^{-1}\sum_{i\in s} y_i$. In this paper we show that Godambe's definition of linear sufficiency is different from the original definition of linear sufficiency due to Barnard [*J. Roy. Statist. Soc. Ser. B.* **25** (1963) 124–127] which is also applicable to survey-sampling. It is shown that the Horvitz–Thompson estimator is the unique unbiased linear sufficient estimator in the class, L, of all linear estimators of the population total. Under an alternative definition, nonexistence of a linear sufficient estimator in L is established. Consequently it is noted that there does not exist a uniformly minimum mean square estimator in L. Extensions to wider classes of estimators are also considered. (Received September 25, 1969.)

70T-2. On the number of averages deviating from the mean. D. W. Müller, University of California at Berkeley.

Let $X_1, X_2, \cdots$ be a sequence of i.i.d. random variables having $EX_1 = 0$, $EX_1^2 = 1$. For given $\varepsilon > 0$ denote by N_ε the (random) number of n such that $|n^{-1}(X_1 + \cdots + X_n)| > \varepsilon$. For $\varepsilon \to 0$, $\varepsilon^2 N_\varepsilon$ has a limit distribution with density $f(t) = f^+(t) - \sum_{n>0}(2n+1)!!^2(2n+1)^2 [f^+(t(2n+1)^2)+2]^* \prod_{j=0}^{n}{}^* [f^+(t(2j+1)^2) - \delta]$; here $f^+(t) = (2/\pi t)^{\frac{1}{2}} e^{-\frac{1}{2}t} - \operatorname{erfc}(t^{\frac{1}{2}})$ (= the limit density of the corresponding one-sided problem), $\operatorname{erfc}(t) = (2/\pi)^{\frac{1}{2}} \int_t^\infty \exp(-\frac{1}{2}u^2)\,du$, δ is the Dirac delta "function". $f(t)$ represents the density of the random variable $\int_0^\infty 1_{[|B(t)|>t]}\,dt$ depending on a Wiener process B (under $P[B(0) = 0] = 1$). The essential tool used in the proof is the identity $L(a, \lambda) + \lambda \int_0^\infty E^a 1_{[|B(s)|>B(0)+s]} \{e^{-\lambda(T(s)-s)} L(T(s)+a, \lambda)\}\,ds = 1$, satisfied by $L(a, \lambda) = E^a \exp(-\lambda \int_0^\infty 1_{[|B(u)|>B(0)+u]}\,du)$ $(a, \lambda > 0)$ and $T(s) = \inf\{u \geqq s: |B(u)| = B(0)+u\}$ $(\leqq +\infty)$. This identity extends to a wider class of diffusion processes. (Received October 7, 1969).

70T-3. Asymptotic relative efficiencies (ARE) of certain asymptotically most powerful rank-order tests and contiguity. Konrad Behnen, Institut für Mathematische Statistik.

Hajek [*Ann. Math. Statist.* **33** (1962) 1124–1147] obtained asymptotically most powerful rank-order tests under certain regularity conditions with respect to parametric classes of alternatives, which are contiguous to the hypothesis. In this paper similar methods are used to derive for any real Lebesgue—a.e. continuous measurable function b defined on $(0, 1)$ with $0 < \int_{(0,1)} b^2(x)\,dx < \infty$, $\int_{(0,1)} b(x)\,dx = 0$, $\int_{(0,t)} b(x)\,dx \leqq 0$, $t \in (0, 1)$ a class of non-parametric alternatives and a two-sample rank-order test, which is asymptotically most powerful for these alternatives. No regularity conditions on the underlying distribution functions are necessary. Analogous results are given for the problem of symmetry and for the correlation problem. In particular, the Wilcoxon, the Fisher–Yates, the van der Waerden-X and the signed rank tests can be derived in this way for those three problems.

The ARE of such tests under general alternatives (contiguous to the hypothesis) can be established by the following method: It is shown that contiguity of the sequence $\{v_n^{(n)}\}$ of product probability measures $v_n^{(n)} = n_n x \cdots x v_n$ (n-times) to the sequence $\{u_n^{(n)}\}$ of product probability measures $u_n^{(n)} = u_n x \cdots x u_n$ (n-times) implies $\lim_{n\to\infty} \sup(n^{\frac{1}{2}}\|u_n - v_n\|) < \infty$. This fact together with an approximation of the underlying rank statistics by sums of independent random variables implies the asymptotic normality of these rank statistics under general alternatives (contiguous to the hypothesis). (This leads to the ARE of tests in the usual way.)

Moreover, results are derived on the existence of bounds for the ARE of two tests under general alternatives (contiguous to the hypothesis). In particular, the results of Hodges and Lehmann [*Proc. Fourth Berkeley Symp. Math. Statist. Prob.* 1 (1961) 307–317] on the lower bound $\pi/6$ for the ARE of the (two-sample) Fisher–Yates test to the Wilcoxon test are generalized: Here it can be shown that in the two-sample problem and also in the problem of symmetry $\pi/6$ and in the correlation problem $(\pi/6)^2$ is a general lower bound. (Received October 9, 1969.)

70T-4. On weak convergence of stochastic processes with multi-dimensional time parameter. GEORG NEUHAUS, Institut für Mathematische Statistik.

A space D_k is introduced consisting of functions on the k-dimensional unit-cube which generalizes the well-known space $D[0, 1]$ (cf. Billingsley, *Convergence of Probability Measures*, Wiley, New York (1968)). It is possible to define a Skorohod-metric on D_k, such that D_k becomes a separable complete metric space. Compact sets in D_k are characterized by a certain "modulus" and a criterion for weak sequential compactness of a sequence of measures on the Borel σ-algebra on D_k is given. As an application of these results it is shown that two "empirical processes" which generalize the one-dimensional empirical process converge weakly to certain Gaussian processes. Especially the results imply the convergence in distribution of the Kolmogorov test statistic for independence. (Received October 9, 1969.)

70T-5. Sequential estimation of an integer mean—Poisson case (preliminary report). GEORGE P. MCCABE, JR., Columbia University.

Let $X_1, X_2, \cdots$ be a sequence of i.i.d. Poisson random variables with mean λ. It is assumed that the true value of the parameter λ lies in the set $\{1, 2, 3, \cdots\}$. From observations on the sequence it is desired to guess the true value of the parameter with a uniformly (for all λ) small probability of error. There is no fixed sample size procedure which can accomplish this. A sequential plan, based on a likelihood ratio criterion, following Robbins, *Sequential estimation of an integer mean* (to be published in the Herman Wold *Festschrift*), is set forth. The procedure, which depends on a parameter $\alpha > 1$, is such that (1) P_λ (error) $< 2/(\alpha - 1)$ for all λ, and (2) E_λ(sample size) $\sim k_\lambda \log \alpha$, as $\alpha \to \infty$, where $k_\lambda = (1 - \lambda \log(1 + 1/\lambda))^{-1}$. The procedure is asymptotically optimal as $\alpha \to \infty$. The results can be extended to more general families of distributions. A computer study of the properties of the procedure for different values of α is in progress. (Received October 14, 1969.)

70T-6. Generalization of Thompson's distribution v. ANDRE G. LAURENT, Wayne State University.

Let X be a univariate random variable with cdf $F(x; m, \theta) = G[x - m)/\theta]$. Let $X_1, \cdots, X_n$ be a sample of n observations of X. Let $m^*(X_1, \cdots, X_n)$ and $\theta^*(X_1, \cdots, X_n)$ be "origin and scale statistics." It is straightforward that $(m^* - m)/\theta$, θ^*/θ, $(m^* - m)/\theta^*$, $(X_i - m^*)/\theta^*$ are parameter free distributed (the latter is a maximal ancillary statistic). A special case is when m^* and θ^* are the maximum likelihood estimates $\hat{m}$ and $\hat{\theta}$ of m and θ. If there exist a complete sufficient statistic T, let $H(z)$ be the cdf of $(X_i-\hat{m})/\hat{\theta}$; then, $H[(x-u)/v]$ is the conditional distribution of X_i given $\hat{m}=u$, $\hat{\theta}=v$, and $H[(x-\hat{m})/\hat{\theta}]$ is the minimum variance unbiased estimator of $F(x; m, \theta)$. (Received October 23, 1969.)

70T-7. Generalization of Thompson's distribution VI. ANDRE G. LAURENT, Wayne State University.

Let X be a univariate random variable with continuous cdf $F(x, \theta)$. Let $X_1, \cdots, X_n$ be a sample of n observations of X, and $T(X_1, \cdots, X_n)$ be a complete sufficient statistic. (i) There exists a maximal ancillary statistic $U(X_i, T)$ that is a monotone function of X_i and whose cdf will be denoted $H(u)$. (ii) The cdf of X_i given $T=t$ is $G(x;t)=H(u)$, where u denotes $U(x, t)$. (iii) If $V(X_i)$ is an unbiased estimator of $g(\theta)$ to be estimated, then $V(X_i)=W(U, T)$. Let $E[W(U, t)=w(t)$, where E denotes an expected value; then $w(T)$ is the m.v.u. estimator of $g(\theta)$. This holds true if g, W, w are indexed by x. Either U is easily identifiable and its distribution is known or both are obtained by deriving the conditional cdf of X_i given $T=t$. Both methods have been used in previous papers by the writer. The results generalise to the case of a statistic $U(Y, T)$, where Y is a subsample. (Received October 23, 1969.)

70T-8. One-sided confidence interval based on a censored sample for an unknown distribution function. ANDRE G. LAURENT, Wayne State University.

Let $F(x)$ be the cdf of a continuous unidimensional random variable X, and $F_{r,n}(x)$ be the corresponding censored cumulative frequency function based on the r first order statistics X_j, $j=1$ to $r<n$, of a sample of n independent observations of X. Let $0<t<1$. Then for all x's: $\Pr[tF_{r,n}(x) \leqq F(x)] = (1-t)[1-I_{rt/n}(r, n-r)]$, where $I_u(p, q)$ denotes the incomplete beta function of argument u. Proof: one considers the case of X rectangular without loss of generality, and one obtains $P[\bigcap_j X_j \geqq jt/n]$ for $j=1$ to r. This extends a result by Robbins (*Ann. Math. Statist.* **25** 409). (Received October 23, 1969.)

70T-9. Simple distribution-free confidence intervals for location difference. P. VAN DER LAAN, Philips' Industries, I.S.A.-Research.

In this paper the construction of simply applicable confidence intervals for location difference of two populations with the same shape for their distributions is considered. The simplest confidence bounds are based on the difference of the ith order statistic of the first sample of size m and the jth order statistic of the second sample of size n. Generalisations of these confidence bounds are also considered. With a certain selection procedure "optimal" confidence bounds are selected for sample sizes $m, n = 1, 2, \cdots, 15$. This selection procedure is based on the power function of the tests to which these confidence bounds correspond. Normal and Homogeneous shift alternatives, as well as Lehmann alternatives have been taken into consideration. The selected tests are compared with Student's two-sample test in case of Normal shift alternatives and with Wilcoxon's two-sample test in case of Lehmann alternatives. The asymptotic relative efficiency of an asymptotic form of the test relative to Student's and Wilcoxon's test is determined in various cases. An extensive analytical and numerical study is made about the behavior of moments and densities of $F^k(x)$ for various classes of distribution functions $F(x)$. (Received October 23, 1969.)

70T-10. Probability limits on order statistics from truncated distributions. SATYA D. DUBEY, New York University.

In this paper necessary explicit formulas for computing the exact desired probability limits on any order statistic from *truncated* Burr, exponential, Gumbel, logistic, Lomax, Rayleigh, Weibull, Weibull-gamma and from other useful distributions have been derived. Formulas for computing exact desired probability limits on any order statistic from uniform distribution, defined over any finite interval, have been obtained as special cases of three different distributions. Several examples have been worked out to illustrate the usefulness of the results. Some properties of these probability limits have been investigated. Suitable transformations have been derived to reduce the above-mentioned distributions to the standard exponential distribution. The effect on probability limits of dependent data has been demonstrated. Some sharp distribution-free probability inequalities have been established. Finally, a method has been discussed to compute exact desired probability limits on any order statistic from mixed distributions. (Received October 31, 1969.)

70T-11. The probability that the empirical function lies between two non-decreasing functions. G. P. STECK, Sandia Laboratories.

Let $X_1, X_2, \cdots, X_m$ be independent observations with distribution function F and let F_m denote the empirical distribution function. Let $P_m(\alpha, \beta \mid F) = P(\alpha(x) \leqq F_m(x) \leqq \beta(x) \mid F)$ where α and β are both continuous distribution functions with the same support as F. It is easy to see that $P_m(\alpha, \beta \mid F) = P_m(\alpha F^{-1}, \beta F^{-1} \mid x)$. By passing to the limit in a result due to Steck [*Ann. Math. Statist.* **40** (1969) 1449–1466] concerning the joint distribution of ranks it is possible to express $P_m(\alpha, \beta \mid F)$ as a determinant. In particular, if αF^{-1} and βF^{-1} are linear, this result generalizes Durbin [*Ann. Math. Statist.* **39** (1968) 398–411] to the case of two lines not necessarily parallel. Also, taking $\alpha = \alpha(F)$ and $\beta = \beta(F)$ makes it possible to construct confidence regions for F of more general shape than shifted versions of F_m. (Received November 4, 1968.)