ABSTRACTS OF PAPERS

(Abstracts of papers presented at the Western Regional meeting, Berkeley, California, July 19–21, 1965. Additional abstracts appeared in earlier issues and others will appear in the October issue.)

5. Selection From Multivariate Normal Populations. KHURSHEED ALAM and M. HASAEEB RIZVI, Indian University and Ohio State University. (By title)

Two problems of selection are considered; Problem I concerns selection of t largest (that is, with largest parameter) of k populations and Problem II concerns selection of a subset containing the t largest. Procedures R_1 and R_2 respectively for Problems I and II are given. R_1 selects the t largest populations corresponding to t largest x_i's, x_i being a statistic form ith population π_i. R_2 selects a subset such that π_i is retained in the subset iff $d(x_i, x_{[k-t+1]}) \leqq \epsilon$, where $\epsilon > 0$ and $x_{[i]}$ denotes the ith smallest x_i and d is a metric; two metrics $d_1(y, z) = z - y$ and $d_2(y, z) = z/y$ are considered. The probability of a correct selection is required to be no less than P^*, $1/\binom{k}{t} < P^* < 1$, for both the problems. This P^* condition determines the common sample size for R_1 and the constant ϵ for R_2. Some operating characteristics of these procedures for a "monotone class" of populations are shown. Application of R_1 and R_2 to multivariate normal populations $\pi_i:N(\mu_i, \sum_i)$, $i = 1, \cdots, k$, is given when populations are ranked according to $\theta_i = \mathbf{\mu}_i' \sum_i^{-1} \mathbf{\mu}_i$; both cases of known or unknown $\sum_i$ are treated. Parametric subspaces where the P^* condition is satisfied are exhibited. Upper bounds for expected size of the selected subset are obtained when $t = 1$.

6. A New Proof of Some Results of Rényi. MIKLÓS CSÖRGŐ, Princeton University.

In his paper, "On the theory of order statistics," *Acta Math. Acad. Sci. Hungar.* **4** (1953), 191–231, Rényi divides the usual Kolmogorov-Šmirnov statistics by $F(x)$, the continuous distribution function of the population from which one assumes having a random sample, and derives their limit distributions by reducing the respective Markov processes to additive Markov processes. In this paper, using the ideas of Doob and Donsker, it is shown that, in the limit, Rényi's random variables can be replaced by a specific Brownian movement process, and this way his original theorems are proved.

7. Characterization Theorems of the Weibull and the Weibull-Gamma Distributions. SATYA D. DUBEY, Procter and Gamble Co. (By title)

Let $X_1, X_2, \cdots, X_n$ be n independent identically distributed random variables and let $Y_n = \min(X_1, X_2, \cdots, X_n)$. Then we obtain the following characterization theorems. THEOREM 1. *If each X_i has the* Weibull *distribution with the parameters α, β and γ, then Y_n obeys the* Weibull *law with the parameters α, $n^{-1}\beta$ and γ. Conversely, if Y_n has the* Weibull *distribution with the parameters μ, σ and λ, then each X_i obeys the* Weibull *law with the parameters μ, $n\sigma$ and λ.* COROLLARY 1. *Theorem* 1 *is valid for the exponential distribution also.* THEOREM 2. *If each X_i has the* Weibull-Gamma (W-G) *distribution with the parameters α, γ, δ and η, then Y_n obeys the* W-G *law with the parameters α, γ, $n\delta$ and η. Conversely, if Y_n has the* W-G *distribution with the parameters μ, λ, θ and σ, then each X_i obeys the* W-G *law with the parameters μ, λ, $n^{-1}\theta$ and σ.* COROLLARY 2. *Theorem* 2 *is also valid for the* Lomax *(Exponential-Gamma) and the* Burr *distributions.*

8. An Analysis for Shift in Mean of a Nonstationary Time Series. M. J. Hinich and J. U. Farley, Carnegie Institute of Technology.

Consider a consumer who is switching between Brand X and other brands of some product. Let p be the unknown probability of his choosing Brand X instead of some other. Due to the effects of advertisement it is hypothesized that p is not constant over time, but instead shifts in a stepwise manner from time to time. Let r_i be the proportion of choice of Brand X in m choices made by a consumer during the ith observation period. Let $X_i = 2$ arc sin $r_i^{\frac{1}{2}}$, which is approximately normal $N(2 \text{ arc sin } p^{\frac{1}{2}}, 1/m)$. Observing over n periods we have $X_1, \cdots, X_n$ where it is assumed that the $X_i - 2$ arc sin $p_i^{\frac{1}{2}}$ are jointly normal with a known covariance matrix. We assume that there is at most one shift over the n periods and that the shift has probability γ of occurring in any of the periods. If the shift occurred in the period j,

$$p_i = \theta_1 \text{ for } i < j$$
$$= \theta_2 \text{ for } i \geqq j$$

where the θ_i are unknown. This has probability γ and the probability of a shift sometime over the whole period is $n\gamma < 1$. Efficient estimators are given for θ_1 and θ_2 for the case where the shift $|\theta_2 - \theta_1| \ll 1/m$. The estimators are linear functions of the statistics $\sum X_i$ and $\sum iX_i$. An approximation to the likelihood-ratio test for a shift over the n periods is developed.

9. On the Property (W) of the Class of Statistical Decision Functions. Hirokichi Kudō, University of California, Berkeley.

The property (W) is introduced by LeCam [*Ann. Math. Statist.* **26** (1955)] as an extension of Wald's concept of weak compactness in the intrinsic sense and is an essential part of the assumptions for the general complete class theorems. Several sufficient conditions for the property (W) and their applications will be discussed. For this purpose, we shall introduce a geometrical notion of half-closedness of a family of non-negative extended functions. Roughly speaking, the class of all decision functions has the property (W) if the family $\{L(\cdot, a): a\varepsilon$ the action space $A\}$ is half-closed, where $L(\theta, a)$ is the loss function. However this condition is not enough for the property (W) of the closed subset of the class of all decision functions. For such a subset the following condition will be needed: for any positive integer n and any θ there is a compact $C_{n,\theta} \subset A$ such that $n \leqq \inf_{a \notin C_{n,\theta}} L(\theta, a)$.

10. Sequential Procedures for Selecting the Best One of Several Binomial Populations. Edward Paulson, Queens College and Courant Institute of Mathematical Sciences, New York University.

Single-stage procedures for selecting the binomial population with the greatest probability of a success were given by Sobel and Huyett [*Bell System Tech. J.* **36** (1957), 537–576]. In this paper corresponding sequential procedures are developed based on taking a random number of measurements with each population not yet eliminated at every stage of the experiment. Let p_i be the probability for population $\prod_i$, denote the ordered probabilities by $p_{[1]} \geqq p_{[2]} \geqq \cdots \geqq p_{[k]}$, and let $\prod_{[i]}$ be the population with probability $p_{[i]}$ $(i = 1, 2, \cdots, k)$. Let N_{ir} be a double sequence of independent random variables each having a Poisson distribution with mean $= J$, and let S_{ir} and F_{ir} be the number of successes and failures when N_{ir} measurements are taken from $\prod_i$ at the rth stage of the experiment. Let $\alpha = (1 - P)/(k - 1)$, $\lambda = (1 + .75d)/(1 - .75d)$, $A = \log \alpha / \log \lambda$, $B = J[d(\lambda^2 - 1) - (\lambda - 1)^2]/(\lambda \log \lambda)$, where d and P are specified in advance. A sequential procedure satisfying the requirement $P[\prod_{[1]}$ is selected $| \, p_{[1]} - p_{[2]} \geqq d] \geqq P$ is obtained as follows: at the rth stage ($r = 1, 2, \cdots$) of the experiment we eliminate any population $\prod_i$ left after the first $r - 1$ stages for which $\sum_{\beta=1}^{\beta=r} (S_{i\beta} - F_{i\beta}) \leqq \max_j [\sum_{\beta=1}^{\beta=r} (S_{j\beta} - F_{j\beta})] + A + rB$, where the maximum is taken over all populations left after the $(r - 1)$ stage. The experiment terminates when only one population is left, which is selected as the best one. A similar sequential procedure has been obtained satisfying the requirement $P[\prod_{[1]}$ is selected $| \, p_{[1]}/p_{[2]} \geqq c] \geqq P$.

11. Confidence Limits for the Product of N Binomial Parameters. MELVIN D. SPRINGER and WILLIAM E. THOMPSON, GM Defense Research Laboratories, General Motors Corporation.

The fiducial probability density function of the product of N binomial parameters is derived in closed form, from which exact confidence limits are obtained. Examples are given to illustrate the procedure for determining explicit confidence limits. The procedure permits the tabulation of the fiducial distribution of the product of an arbitrary number of binomial parameters, and has an immediate application to problems of reliability of serial systems.

12. An Admissible Test with Monotone Power Function for the Equality of Two Covariance Matrices. M. S. SRIVASTAVA, University of Toronto. (By title)

Samples of sizes $N_1(> p + 1)$ and $N_2(> p + 1)$ are drawn from $N(\mu_1, \Sigma_1)$ and $N(\mu_2, \Sigma_2)$, respectively, where $N(\mu_j, \Sigma_j)$ denotes a p-variate non singular normal distribution with mean vector μ_j and covariance matrix Σ_j. On the basis of these data we wish to test the null hypothesis $H_0: \Sigma_1^{-1} = \Sigma_2^{-1}$ against the one-sided alternative $H_1: \Sigma_1^{\frac{1}{2}}\Sigma_2^{-1}\Sigma_1^{\frac{1}{2}} = I + \eta\eta'$, where $\eta\eta'$ is a $p \times p$ symmetric and positive semidefinite matrix of rank $\geqq 1$ (η may be taken to be a matrix of order $p \times r$) and the square root $\Sigma_1^{\frac{1}{2}}$ is any factorization of Σ_1. It has been shown that the test with critical region $\{(S_1, S_2): \det (S_1 + S_2)/\det S_2 > R\}$ is admissible. Also, the power of this test depends on the characteristic roots of $\Sigma_1\Sigma_2^{-1}$, and is a monotone increasing function of each of the ordered root of $\Sigma_1\Sigma_2^{-1}$.

13. Uniform Convergence of Losses in the Fixed and Sequential $m \times n$ Compound Decision Problem. J. VAN RYZIN, Argonne National Laboratory. (By title)

The notation is that of the author's abstracts (*Ann. Math. Statist.* **35**, 1847) for the fixed case and (*Ann. Math. Statist.* **36**, 362–363) for the sequential case. Let $L_N(T^*, \theta) = N^{-1} \cdot \sum_{k=1}^{N} L(T^*, \theta_k)$ where $L(T^*, \theta_k)$ is the loss of the procedure $T^* = \{t(x_k, \xi_N), k = 1, \cdots, N\}$ for the fixed case and $T^* = \{t(x_k, \xi_{k-1}), k = 2, 3, \cdots\}$ or $T^* = \{t(x_k, \xi_k), k = 1, 2, \cdots\}$ for the sequential case. Then, for every $\epsilon > 0$, $P[(\log N)^{-1}N^{\frac{1}{2}} |L_N(T^*, \theta) - \phi_N(\theta)| \geqq \epsilon] \to 0$ uniformly in the sequence θ for suitably chosen estimates ξ_k, $k = 1, 2, \cdots$. The estimates $\xi_k = (\xi_{k1}, \cdots, \xi_{km})$ are taken as $\xi_{ki} = \bar{g}_{i,k}$ in Equation (14), p. 18 of Robbins (*Ann. Math. Statist.*, **35**, 1–20) satisfying suitable integrability conditions in both the fixed and sequential cases, plus certain non-degeneracy assumptions in the sequential case. In particular, if $m = n = 2$, the sequential case results improve on Theorem 1 of Samuel (*Ann. Math. Statist.*, **35**, 1606–1621) by yielding uniformity in θ, a convergence rate of $(\log N)^{-1} N^{\frac{1}{2}}$, and extension to arbitrary pairs of distributions (P_1, P_2), $P_1 \neq P_2$, while requiring a weakening from convergence with probability one to convergence in probability. Under this same weakening, if $m = n = 2$, the fixed case result yields a convergence rate improvement for the class of estimates given above on a two-sided version of Theorem 3 of Hannan and Robbins (*Ann. Math. Statist.*, **26**, 37–51).

14. Theory of Modal Unbiased Estimation. M. T. WASAN, Queen's University, Kingston, Ontario. (By title)

A modal unbiased estimate of a parameter of a density function is defined and for some important density functions modual unbiased estimates are constructed. It is shown that a modal unbiased estimate may not be unique. A method of construction of it is given and conditions are singled out to show when it does exist. Furthermore it is proved that maximum likelihood estimate and BAN estimate are asymptotically a modal unbiased estimate. The relative efficiency of a modal unbiased, a mean unbiased and a median unbiased estimate are discussed. For a given loss function it is proved that for exponential and Weibull density functions a modal unbiased estimate has uniformly smaller risk than that of a mean unbiased estimate, for other given convex loss function they are equivalent. It is shown also that a modal unbiased estimate has mean square error uniformly smaller than that of median unbiased estimate. The methods are given to find what type of loss function can be appropriate for an estimate of given function of parameter of a density function. A method of estimation by confidence set is discussed. It is proved that for Pareto and exponential density function confidence interval based on the modal unbiased estimate is shorter in length than that of uniformly minimum variance unbiased estimate of a parameter for a given probability of confidence.

(*Abstracts of papers to be presented at the Annual Meeting, Philadelphia, Pennsylvania, September 8–11, 1965. Additional abstracts appeared in the August issue and others will appear in future issues.*)

4. Characterization Theorems for Beta and "FMEL" Distributions. SATYA D. DUBEY, Procter and Gamble Co. (By title)

Let $X_1, X_2, \cdots, X_n$ be n independent identically distributed random variables and let $Y_n = \min(X_1, X_2, \cdots, X_n)$. Then we obtain the following characterization theorems. THEOREM 1. *If each X_i has a* beta *distribution with the parameters α and* 1, *then Y_n^α obeys the* beta *law with the parameters* 1 *and n. Conversely, if Y_n has a* beta *distribution with the parameters β and* 1, *then each X_i^β obeys the* beta *law with the parameters* 1 *and* n^{-1}. COROLLARY 1. *Theorem* 1 *is valid for the* uniform *distribution in a weak sense when* $\alpha = \beta = 1$. THEOREM 2. *If each X_i has a* flexible modified exponential-model Lomax (FMEL) *distribution with the parameters a, b, c and μ then Y_n obeys the* FMEL *law with the parameters na, b, nc and μ. Conversely, if Y_n has a* FMEL *distribution with the parameters α, β, γ and δ then each X_i obeys the* FMEL *law with the parameters $n^{-1}\alpha$, β, $n^{-1}\gamma$ and δ.* COROLLARY 2. *Theorem* 2 *is valid for the* modified exponential-model Lomax *distribution also.*

5. Characterization Theorems for Extreme Value and Logistic Distributions. SATYA D. DUBEY, Procter and Gamble Co. (By title)

Let $X_1, X_2, \cdots, X_n$ be n independent identically distributed random variables and let $Y_n = \min(X_1, X_2, \cdots, X_n)$. Then we obtain the following characterization theorems. THEOREM 1. *If each X_i has a* generalized extreme value (GEV) *distribution with the parameters, α, μ and σ, then Y_n obeys the* GEV *law with the parameters $n\alpha$, μ and σ. Conversely, if Y_n has a* GEV *distribution with the parameters β, γ and δ, then each X_i obeys the* GEV *law with the parameters $n^{-1}\beta$, γ and δ.* COROLLARY 1.1. *Theorem* 1 *is valid for the* truncated GEV *distribution also.* COROLLARY 1.2. *Theorem* 1 *is valid for the* extreme value *distribution in a* *weak sense.* THEOREM 2. *If each X_i has a* GEV-Gamma (GEV-G) *distribution with the parameters p, q, μ and σ, then Y_n obeys the* GEV-G *law with the parameters np, q, μ and σ. Conversely, if Y_n has a* GEV-G *distribution with the parameters β, γ, δ and ν, then each X_i obeys the* GEV-G *law with the parameters $n^{-1}\beta, \gamma, \delta$ and ν.* COROLLARY 2.1. *Theorem 2 is also valid for the truncated* GEV-G *distribution which includes the standard logistic distribution as a special case.* COROLLARY 2.2. *Theorem 2 is valid for the logistic distribution of Gupta and Shah only in a weak sense.*

6. The Moments of a Doubly Non-Central t-Distribution. MARAKATHA KRISHNAN, State University of New York at Buffalo and Indian Statistical Institute, Madras.

If X is normally distributed with mean δ and variance 1, and Y^2 is independently distributed as non-central χ^2 with ν degrees of freedom and non-central parameter λ, then the ratio $X\nu^{\frac{1}{2}}/Y$ follows a doubly non-central t-distribution with ν degrees of freedom and non-central parameters δ and λ. Robbins [*Ann. Math. Statist.* **19** (1948) 406–410] and Patnaik [*Sankhyā* **15** (1955) 343–372] have considered this distribution in certain situations, although tables of its probability integral do not exist. Approximations to the doubly non-central t-distribution require its moments. This paper gives analytic expressions for the moments and recurrence relations for the first four raw moments of the doubly non-central t-distribution.

7. A New Method of Estimating Treatment Effects or Treatment Differences in Balanced Incomplete Block Designs. PAUL S. LEVY, Harvard Medical School and Peter Bent Brigham Hospital. (By title)

Graybill and Deal [*Biometrics* **15** (1959) 543–550] have introduced an estimator $\hat{t}_{iw}$ $(= (\hat{w}_1 t_i + \hat{w}_2 t_{bi})/(\hat{w}_1 + \hat{w}_2))$ of a treatment effect, τ_i, in a BIBD. This estimator is, in effect, a weighted combination of the intra-block estimator, t_i, and the inter-block estimator, t_{bi}, with weights $\hat{w}_1$ and $\hat{w}_2$ chosen in inverse proportion to the estimated variances of t_i and t_{bi}. Because of the mutual independence of t_i, t_{bi}, $\hat{w}_1$ and $\hat{w}_2$, and because $\hat{w}_1$ and $\hat{w}_2$ involve mean squares, the method of Meier [*Biometrics* **9** (1953) 59–73] can be used to obtain an approximate expression for the variance of $\hat{t}_{iw}$. If block effects were assumed to be nonexistent, and the design analysed as a one way ANOVA model, the estimator, $\bar{t}_i$, of τ_i so obtained can be expressed as $(f_1 t_i + f_2 t_{bi})/(f_1 + f_2)$ where f_1 and f_2 are constants. A method introduced by this author [*Ann. Math. Statist.* **35** (1964) 1394–1395] is proposed to construct an estimator $\hat{t}_{i0}$ of τ_i which is a linear combination of $\hat{t}_{iw}$ and $\bar{t}_i$. This estimator has the property of being close to $\hat{t}_{iw}$ when block effects are large and to $\bar{t}_i$ when block effects are small. Treatment differences, $\tau_i - \tau_j$ can be handled analogously.

8. Bayesian Comparison of Means of a Mixed Model with Application to Regression Analysis. GEORGE C. TIAO, University of Wisconsin.

In this paper, a Bayesian approach is adopted to analyse the two way mixed model, $y_{ij} = \mu_i + a_j + e_{ij}$, $i = 1, \cdots, m$, $j = 1, \cdots, k$, where μ_i's are location parameters, a_j's the random effects and e_{ij}'s the errors. Under the usual assumption of normality, independence and a "non-informative" prior distribution, the joint posterior distribution of the parameters in the model is obtained. Special attention is then given to the problem of comparing two means. It is shown that the criterion for comparison follows a modified Student-t distribution. This allows the investigator to extract more "information" from the data than what is usually possible in the sampling theory framework. A procedure with similar properties for an overall comparison of more than two means is also derived. The corresponding criterion here follows a modified F distribution. An asymptotic method which yields excellent approximations to the modified t and F distributions is developed. Finally the analysis is extended to regression models with variance-components appearing in the error term.

Abstracts of papers not connected with any meeting of the Institute.

1. Stochastic Approximation and Nonlinear Regression–I. A. Albert and L. A. Gardner, Jr., ARCON Inc. and Lincoln Laboratory, M.I.T.

Let $y_1, y_2, \cdots$ be an observable time series with uniformly bounded variances and $\mathcal{E}y_n = F_n(\theta)$. The unknown parameter θ is an interior point of a prescribed interval Ω (finite or infinite) on which each F_n is a given differentiable function with sgn $F_n'(x) = s_n$ for all $x \varepsilon \Omega$. For each n let $a_n(x_1, \cdots, x_n)$ be defined and positive over the n-fold product space Ω_n of Ω. For points in Ω_n set $T_n(x_1, \cdots, x_n) = x_n + s_n a_n(x_1, \cdots, x_n)\cdot[y_n - F_n(x_n)]$, and let $[T]_\Omega$ denote the truncation of T in the interval Ω. We consider rapidly computable iterative schemes (a must in real-time estimation problems) of the form $t_{n+1} = [T_n(t_1, \cdots, t_n)]_\Omega$ arbitrarily initialized by $t_1 \varepsilon \Omega$. The variety of choices available for $\{a_n\}$ which yield probability one convergence of $\{t_n\}$ to θ is evident from the following conditions, formulated in terms of $\gamma_n = \inf_\Omega |F_n'(x)|$, $\lambda_n = \sup_\Omega |F_n'(x)|$, $\underline{a}_n = \inf_{\Omega_n} a_n(x_1, \cdots, x_n)$ and $\bar{a}_n = \sup_{\Omega_n} a_n(x_1, \cdots, x_n)$: (1) $\limsup \lambda_n \bar{a}_n < 1$, (2) $\sum \gamma_n \underline{a}_n = \infty$, and (3) $\sum \bar{a}_n < \infty$ or, if $\{y_n\}$ is an independent process, (3′) $\sum \bar{a}_n^2 < \infty$. We apply our general result to gain functions such that $\underline{a}_n \geqq a\gamma_n/\Gamma_n^2$ and $\bar{a}_n \leqq b\gamma_n/\Gamma_n^2$, for some $0 < a \leqq b < \infty$, where $\Gamma_n^2 = \gamma_1^2 + \cdots + \gamma_n^2$. Then (1) holds if $\limsup \lambda_n/\gamma_n < \infty$ and $\gamma_n^2/\Gamma_n^2 \to 0$, and (2) and (3′) if $\Gamma_n^2 \to \infty$. (3) is satisfied whenever γ_n increases as some power of n (e.g. polynomial regression). When γ_n increases so rapidly that γ_n^2/Γ_n^2 does not approach 0, strong consistency is retained if we replace a_n by a_n/n. In practice we use gains which are computed recursively. The two most important ones are deterministic weights γ_n/Γ_n^2 and random weights $|F_n'(t_n)|/\sum_1^n F_k'^2(t_k)$.

2. Stochastic Approximation and Nonlinear Regression–II. A. Albert and L. A. Gardner, Jr., ARCON Inc. and Lincoln Laboratory, M.I.T.

Let $y_1, y_2, \cdots$ be independent random variables with $\mathcal{E}y_n = F_n(\theta)$, where θ is an unknown interior point of an interval Ω (finite or infinite) on which each F_n is a given differentiable function with sgn $F_n'(x) = s_n$ for all $x \varepsilon \Omega$. Let $\lambda_n = \sup_\Omega |F_n'(x)|$, $\gamma_n = \inf_\Omega |F_n'(x)|$ and $\Gamma_n^2 = \gamma_1^2 + \cdots + \gamma_n^2$. Let $a_n(x_1, \cdots, x_n)$ be defined at every point in the n-fold product space Ω_n of Ω, and suppose $\inf_{\Omega_n} a_n(x_1, \cdots, x_n) \geqq a\gamma_n/\Gamma_n^2$ and $\sup_{\Omega_n} a_n(x_1, \cdots, x_n) \leqq b\gamma_n/\Gamma_n^2$ for some $0 < a \leqq b < \infty$. For points in Ω_n define $T_n(x_1, \cdots, x_n) = x_n + s_n a_n(x_1, \cdots, x_n)[y_n - F_n(x_n)]$ and let $[T]_\Omega$ be the truncation of T in the interval Ω. Let t_1 be arbitrary in Ω, and recursively define $t_{n+1} = [T_n(t_1, \cdots, t_n)]_\Omega$ for all $n \geqq 1$. The following conditions ensure $e_n = \mathcal{E}(t_n - \theta)^{2p} \to 0$ (p an integer): (1) $\limsup \lambda_n/\gamma_n < \infty$, (2) $\Gamma_n^2 \to \infty$, (3) $\gamma_n^2/\Gamma_n^2 \to 0$ and (4) the $2p$th moment of $y_n - F_n(\theta)$ is uniformly bounded. (The theorem remains true when (3) fails provided a_n is replaced by a_n/n.) Under the same conditions $e_n = O(1/\Gamma_n^{2p})$ when $a > \frac{1}{2}$. If (3) is strengthened to $\sum \gamma_n^4/\Gamma_n^4 < \infty$, then e_n is at most the order of $(\log \Gamma_n^2/\Gamma_n^2)^p$ when $a = \frac{1}{2}$, and $1/\Gamma_n^{4ap}$ when $a < \frac{1}{2}$. In applications the regression functions are usually defined only over a finite or perhaps semi-infinite interval J. The procedure with $\Omega = J$ is called "truncated." The procedure with F_n linearly extended from J to $\Omega = (-\infty, \infty)$ is called "untruncated."

3. Stochastic Approximation and Nonlinear Regression–III. A. ALBERT and L. A. GARDNER, JR., ARCON Inc. and Lincoln Laboratory, M.I.T. (Preliminary Report)

Let $y_1, y_2, \cdots$ be independent random variables with common variance σ^2 and $\mathcal{E}y_n = F_n(\theta)$, where θ is an unknown interior point of an interval $J = [A, B]$ (finite or infinite) on which each F_n is a given differentiable function with sgn $F_n'(x) = s_n$ for all $x \varepsilon J$. Let $\lambda_n = \sup_J |F_n'(x)|$, $\gamma_n = \inf_J |F_n'(x)|$ and $\Gamma_n^2 = \gamma_1^2 + \cdots + \gamma_n^2$. Assume (1) lim sup $\lambda_n/\gamma_n < \infty$, (2) $\Gamma_n^2 \to \infty$, (3) $\sum \gamma_n^4/\Gamma_n^4 < \infty$, and (4) the functions $h_n(x) = F_n'(x)/\gamma_n$ $(n = 1, 2, \cdots)$ have a common modulus of continuity on J, and converge as $n \to \infty$ to $h(x)$ at each $x \varepsilon J$. Extend F_n by linearity from J to the entire real line, and define $R(x) = \int_A^x |h(\xi)|\, d\xi$ for $-\infty < x < \infty$. Let $M = R^{-1}$ be the inverse function and define $\bar{F}_n(y) = F_n(M(y))$ for $-\infty < y < \infty$. Let $\bar{t}_1$ be arbitrary and $\bar{t}_{n+1} = \bar{t}_n + s_n a_n [y_n - \bar{F}_n(\bar{t}_n)]$ where a_n is either the deterministic gain γ_n/Γ_n^2 or the random gain $|\bar{F}_n(\bar{t}_n)|/\sum_1^n \bar{F}_k'^2(\bar{t}_k)$. The $t_n = M(\bar{t}_n)$'s are called "untruncated transformed" θ-estimates. If (5) $y_n - F_n(\theta)$ has a uniformly bounded moment of order exceeding 2, then $[\sum_1^n F_k'^2(\theta)]^{\frac{1}{2}}(t_n - \theta)$ is asymptotically $N(0, \sigma^2)$. When at least one of A, B is finite, the estimate t_n obtained by first truncating the right side of the equation for $\bar{t}_{n+1}$ from above by $R(B)$ and below by 0 (called "truncated transformed") has the same limiting distribution provided (5′) $\Gamma_n^{\nu}\gamma_n \to \infty$ for some $0 \leqq \nu < \infty$ and $y_n - F_n(\theta)$ has a uniformly bounded moment of integer order $\geqq 4 + 2\nu$. In either case, we can replace θ by t_k in the norming sequence. These iterations performed in the transformed parameter space are asymptotically efficient when and only when each y_n is $N(F_n(\theta), \sigma^2)$. If the deterministic gains are used in either the "untruncated" or "truncated" iteration but without transforming, i.e. $M(y) \equiv y$, then the variance of the limiting distribution is increased by the factor $h^2(\theta)/(2\,|h(\theta)| - 1)$. With the random gains the variance is unaltered provided $\sup_J |h(x)|^2 < 2 \inf_J |h(x)|$.

4. A Method of Sample Selection with Unequal Probabilities without Replacement. J. DURBIN, London School of Economics and Political Sciences.

Choose the first sample unit with selection probabilities $p_1, \cdots, p_N$ where $\sum p_i = 1$. Suppose the ith unit is chosen. Choose the second unit with selection probabilities $p_j' = \lambda p_j\, \{(1 - 2p_i)^{-1} + (1 - 2p_j)^{-1}\}$ $(j \neq i)$, where λ is such that $\sum_{j \neq i} p_j' = 1$. Since $\sum_{j \neq i} \cdot p_j\{(1 - 2p_i)^{-1} + (1 - 2p_j)^{-1}\} = 1 + \sum_{k=1}^N p_k(1 - 2p_k)^{-1}$ it follows that λ is independent of i. The probability p_ip_j' of getting the ith unit first and the jth unit second is consequently equal to the probability of getting the jth unit first and the ith unit second. The total probability of selecting the ith and jth units is therefore $2\lambda p_ip_j\{(1 - 2p_i)^{-1} + (1 - 2p_j)^{-1}\}$ and the total probability of selecting the ith unit is $2p_i$. These results are useful for estimating variances in stratified multi-stage samples in the manner described by Yates and Grundy and by Durbin, *J. Roy. Statist. Soc. Ser. B.* **15** (1953) 253–261 and 262–269. For samples of size three take the third set of selection probabilities proportional to $p_k'\{(1 - 2p_j')^{-1} + (1 - 2p_k')^{-1}\}$ $(k \neq i, j)$ and so on for larger sample sizes.

5. Some Aspects of Simultaneous Equations Estimation Theory with Complex Observations. D. G. KABE, Northern Michigan University.

Let $By_t + \Gamma x_t = u_t$ be a simultaneous equations model, where B is a $G \times G$ Hermitian positive definite matrix, Γ a $G \times K$ matrix of rank $G(<K)$, y_t, x_t, and u_t are complex column vectors of G, K, and G elements respectively. The error vector u_t has mean zero and a Hermitian positive definite matrix as its covariance matrix or u_t has a G variate complex normal distribution. Then several results in the real simultaneous equations estimation theory may be generalized with minor changes to the complex case represented by the

above model. Assuming a sample of N observations on y_t and x_t is available, (i.e., $t = 1, 2, \cdots, N$) we obtain the estimates of the parameters of the above model, under certain conditions, using the usual methods such as indirect least squares etc. We also consider the estimation of the parameters of the first equation of the above model by using two stage and three stage least squares and generalize our results to k class estimators.

6. A Continued-Fraction Algorithm for the Automatic Computation of Markov Transition Functions. David G. Kendall, Churchill College, Cambridge.

Continued-fraction algorithms for the exponential function are adapted to permit the computation of the transition functions $p_{ij}(t)$ for a Markov process with bounded infinitesimal generator Ω by the iterative evaluation of the convergents to a continued fraction involving Ω. Analytical error estimates lead one to anticipate high accuracy with a small number of iterations, and preliminary numerical studies have now confirmed this expectation. The program will be made generally available after further testing.

7. Efficiency of Des Raj's Estimator. Pramod K. Pathak, Indian Statistical Institute.

Consider a population $\Pi = (U_1, \cdots, U_j, \cdots, U_N)$ of N element. Let Y_j and P_j respectively be the value of a Y-characteristic and the probability of selection associated with the jth population unit U_j. It is shown that under sampling with unequal probabilities without replacement the variance of Des Raj's estimator [*J. Amer. Statist. Assoc.* **51** (1956) 269–289] of the population total $Y = \sum Y_j$ is given by $(\frac{1}{2})n^{-2} \sum_{j,j'=1}^{N} (Y_jP_{j'} - Y_{j'}P_j)^2 \cdot P_j^{-1}P_{j'}^{-1}\{1 + \sum_{k=2}^{n} Q_{jj'}(k)\}$, where n is the sample size and $Q_{jj'}(k)$ denotes the probability of non-inclusion of U_j and $U_{j'}$ in the first $(k-1)$ sample units. A simple upper bound to the above variance is obtained on observing that $Q_{jj'}(k) \leqq (1 - P_j - P_{j'})^{k-1}$. Des Raj's estimator is then compared with other estimators. It is found that in many situations commonly met in practice Des Raj's estimator has smaller variance than the estimators considered by Sampford [*Biometrika* **49** (1962) 27–40], J. Rao, Hartley and Cochran [*J. Roy. Statist. Soc. Ser. B* **24** (1962) 482–491], Hartley and J. Rao [*Ann. Math. Statist.* **33** (1962) 350–374], Stevens [*J. R. Statist. Soc. Ser. B* **20** (1958) 394–397] and Hájek [*Ann. Math. Statist.* **35** (1964) 1491–1523].

8. Limit Theorems Involving Capacities. Sidney C. Port, RAND Corporation.

In an irreducible, transient Markov chain let $V_B(T_B)$ be the time of first (last) visit to a finite nonempty set B. If the chain has an invariant measure $\pi(x)$, then let $E_B(n) = \sum_x \pi(x)P_x(V_B \leqq n)$. We show $[E_B(n) - E_B(n-1) - C(B)] \downarrow 0$ and $E_B(n) - nC(B) \uparrow$, where $C(B) = \sum_{x \varepsilon B} \pi(x)P_x(V_B = \infty)$. Also if $R_n(x, y) = \sum_{j=n+1}^{\infty} P_x(X_j = y)$, then if $\sum_n R_n(0, 0) < \infty$ we have $E_x(V_B \mid V_B < \infty) P_x(V_B < \infty) < \infty$, $E_xT_B < \infty$ and $\lim_{n\to\infty} [E_B(n) - nC(B)] < \infty$, while if $\sum_n R_n(0, 0) = \infty$ we have $\sum_{x \varepsilon B} E_x(V_B \mid V_B < \infty) P_x \cdot (V_B < \infty) = \infty$, $E_xT_B = \infty$ and $\lim_{n\to\infty} [E_B(n) - nC(B)] = \infty$. If the chain satisfies condition (A): $\lim_{n\to\infty} \sum_{j=0}^{n} R_j(x, y)/\sum_{j=0}^{n} R_j(0, 0) = A(x, y)$ exists for all x, y, then (1) $\sum_{j=0}^{n} P_x(j < V_B < \infty) \sim Q(x; B) \sum_{j=0}^{n} R_j(0, 0)$, (2) $\sum_{j=0}^{n} P_x(T_B > j) \sim D(x; B) \sum_{j=0}^{n} R_j(0, 0)$ and (3) $E_B(n) - nC(B) \sim L(B) \sum_{j=0}^{n-1} R_j(0, 0)$, and the constant Q, D, and L are explicitly determined. $E_B(n)$ is shown to be the mean, $EN_n(B)$, of the number of distinct particles which enter the set B by time n in certain system of infinitely many particles. We prove that $P(N_n(B)/n \to C(B)) = 1$. When $\sum R_n(0, 0) = \infty$, Condition A is shown to hold for all irreducible, transient (a) *Random Walk* on the integer lattice points in E^d, (here $A(x, y) = 1$) (b) *Discrete Time Birth and Death Processes* (here $A(x, y) = \pi(y)$), and (c) *Discrete Time Renewal Processes*. For certain Random Walks stronger results (like $P_x(n < V_B < \infty) \sim Q(x; B)R_n(0, 0)$) are shown to be true.

9. Asymptotic Efficiency of Some Nonparametric Tests of Dispersion. HANS K. URY, California State Department of Public Health, Berkeley.

In a recent paper (*Ann. Math. Statist.* **34** 973–983), Moses showed that no rank test (i.e., a test invariant under strictly increasing transformation of the scale) can hope to be a satisfactory test against dispersion alternatives unless some sort of strong restriction (such as known median difference) is placed on the class of admissible distribution pairs. He also considered some "rank-like" tests which consist of applying a rank test to some index of dispersion computed within small subgroups; these will be of exact size and should prove to be robust. For a particular case, Wilcoxon's test applied to variances-within-triads, the ARE against normal alternatives was shown to be 0.5. In the present paper, the ARE against normal alternatives is computed for different subgroup sizes when Wilcoxon's test is applied (a) to subgroup ranges and (b) to subgroup variances, and (c) when the Fisher-Yates-Hoeffding-Terry test is applied to subgroup variances. For (a), the maximum of 0.69 is reached for subgroups of size 8; for (b) and (c), the ARE tends to $3/\pi$ and 1, respectively, with increasing subgroup size. The question of decreasing "efficiency" because of an insufficient number of subgroups is considered in a numerical example. Finally, some analogous sign tests are briefly investigated.

10. The Voting Paradox—A Summary of Related Research. ZALMAN USISKIN, Niles Township High School West, Skokie, Illinois.

At the time of the publication of my paper (Max-min probabilites in the voting paradox, *Ann. Math. Statist.* **35** 857–862), I was aware of very little related research. I wish to thank Harry H. Ku, H. A. David, and Thomas Cover for calling my attention to some of the references listed here. Steinhaus and Trybula [On a paradox in applied probabilities, *Bull. Acad Polon. Sci.* **7** (1959) 67–69] seem to have been the first to consider the problem of independence, though the paradox itself dates back at least to Condorcet [*Essai sur l'Application de l'Analyse a la Probabilite des Decisions Rendues a la Pluralite des Voix*, De l'Imprimerie royale, (1785)] and Black [*The Theory of Committees and Elections*, Cambridge Univ. Press, (1958)] discusses the dependent case $n = 3$. Trybula [On the paradox of three random variables, *Zastos. Mat.* **5** (1961) 321–332] obtains the bound for the independent case $n = 3$. David discusses both cases [*The Method of Paired Comparisons*, C. Griffin & Co., Ltd., (1963)]. Chang, translated by Ku, [The maximum and minimum probabilities of cyclic stochastic inequalities, *Chinese Math.* **2** (1963) 279–283] states but does not prove Theorem 2 of my paper, proves Theorem 3 and proves one of the bounds given in Section 4 of the same paper. My paper goes beyond Chang's in obtaining the bounds explicitly, showing that they are achievable and by what distributions, and in giving sharper lower bounds for the bounds.