The Annals of Mathematical Statistics
1969, Vol. 40, No. 2, 720–727

ABSTRACTS OF PAPERS

(An abstract of a paper presented at the Annual meeting, Madison, Wisconsin, August 26–30, 1968. Additional abstracts appeared in earlier issues.)

92. The distribution of the sample correlation coefficient with one variable fixed. DAVID HOGBEN, National Bureau of Standards.

For the usual straight-line model, in which the independent variable takes on a fixed, known set of values, it is shown that the sample correlation coefficient is distributed as Q with $(n - 2)$ degrees of freedom and noncentrality $\theta = (\beta/\alpha)/\overline{\sum (x_i - x)^2}$. The Q variate has been defined and studied previously by Hogben et al. (*Ann. Math. Statist.* **35** 298–314 and 315–318). It is noted that the square of the correlation coefficient is distributed as a non-central beta variable.

(Abstracts of papers presented at the Central Regional meeting, Iowa City, Iowa, April 23–25, 1969. Additional abstracts have appeared in earlier issues and will appear in future issues.)

6. Jackknifing U-statistics. JAMES N. ARVESEN, Purdue University.

Previous work of Hoeffding on U-statistics (*Ann. Math. Statist.* **19** (1948) 293–325), and Miller on the jackknife (*Ann. Math. Statist.* **35** (1964) 1594–1605) is combined to obtain the following result. Let $X_1, \cdots, X_n$ be independent and identically distributed random variables, and $U(X_1, \cdots, X_n)$ be the U-statistic based on these random variables. Assume $U(X_1, \cdots, X_n)$ is an unbiased estimate of η, and Var $(U(X_1, \cdots, X_n)) \rightarrow \sigma^2 < \infty$ as $n \rightarrow \infty$. Let f denote a real-valued function defined on the real line, which in a neighborhood of η has a bounded second derivative. Let $\hat{\theta}$ denote the jackknife estimate of $\theta = f(\eta)$, and s_f^2 denote the sum of squares as defined in Miller. Then as $n \rightarrow \infty$, $n^{1/2}(\hat{\theta} - \theta)/s_f \rightarrow_{\mathfrak{L}} N(0, 1)$. The result is then extended to functions of q U-statistics, that is real-valued functions defined on R^q. Finally an extension is presented to the case where $X_1, \cdots, X_n$ are independent (not necessarily identically distributed). Applications are then presented to obtain both asymptotic tests and confidence intervals for variance, components in Model II ANOVA. The unbalanced case is also treated. (Received 27 January 1969.)

7. Bayesian prediction and population size assumptions. T. L. BRATCHER and W. R. SCHUCANY, Southern Methodist University.

This paper is concerned with the distribution of the number successes in a random sample given the results of a previous sample from the same population. Assuming uniform weights (i.e., uniform prior) on the proportion of successes in the original population, Bayes rule is utilized to obtain the desired distribution. If the size of the population is finite, say N, then the hypergeometric density gives the probabilities for the number of successes in a random sample. On the other hand, if N is infinite, the binomial gives the probabilities. Somewhat surprisingly, the resulting distribution is independent of the population size N and is the same for both the finite and infinite cases. (Received 9 January 1969.)

8. Useful bounds in packing problem. BODH RAJ GULATI and E. G. KOUNIAS, Eastern Connecticut State College and McGill University.

Let $m_t(r, s)$ denote the maximum number of points in finite projective geometry $PG(r - 1, s)$ of $(r - 1)$-dimensions based on Galois field $GF(s)$, where s is a prime or power of a prime, so that no t of the chosen points are conjoint (a set of t points are conjoint if they lie on a flat space of dimensions not greater than $t - 2$). Bose has shown (*Sankhyā* **8**) that $m_t(r, s)$ also symbolises the maximum number of factors that can be accommodated in a symmetrical factorial design in which each factor is at $s = p^n$ levels, blocks are of size s^r and no t-factor or lower order interaction is confounded. These bounds, which have been lately useful in error correcting codes and information theory, were introduced by Bose, Barlotti, Seiden, Segre, Qvist and many others. Their investigations are restricted to $t = 3$ but no general methods are available for $t \geqq 4$. Our results for $m_t(t, s) \leqq s + t - 1$, $s = 2^n$ and $m_t(t, s) \leqq s + t - 2$, $s = p^n$ (p odd) agree with those of Bush (*Ann. Math. Statist.* **23**). We have further established that $m_t(t + 1, s) \leqq s^2 + t - 2$ for s odd. Example of 10 points for $s = t = 3$ was given by Bose (*Sankhyā* **8**) and of 11 points for $s = 3$, $t = 4$ by G. Tallini (*Acta Arithemetica* **7** (1961)). (Received 11 December 1968.)

9. Distribution of Wilks' likelihood ratio criterion in the complex case. A. K. Gupta, The University of Arizona.

The null distributions of Wilks' likelihood ratio criterion Λ, in the complex case, are derived, and it is shown that the density and the distribution function of Λ have exact closed form representations for all p, the number of variates, and for all f_2, the hypothesis degrees of freedom. Earlier Gupta (to appear) and Pillai and Gupta (*Biometrika* **56** (1969)) using convolution techniques, have obtained the density and the distribution functions of Λ in the real case. Based on a result of Khatri (*Ann. Math. Statist.* **36** (1965)), the techniques of the author's earlier papers have been used to derive the results of the present paper. (Received 4 February 1969.)

10. Inter-relations among estimators based on different definitions of efficiency (preliminary report). A. R. Padmanabhan, Ohio State University.

Among all unbiased estimators of a given estimable function, the uniformly minimum variance estimator (UMV), if exists, is usually taken to be the best. This paper gives a new justification for considering variance as the criterion instead of any higher order absolute central moment. It is shown that if a bounded estimator is the best in terms of variance, (i.e. if it is UMV), then it is automatically the best in terms of any higher absolute central moment. Under some restrictions, this holds for even unbounded UMV's. If an estimator T is locally the best in terms of the pth absolute central moment, ($p > 1$) and takes only two values then, for any $r > p$, T is the best (locally) in terms of the rth absolute central moment as well. However, if T takes even three values, this result may be false. Finally, the non-existence of the UMV's in the case of the rectangular distribution $R(\theta - \frac{1}{2}, \theta + \frac{1}{2})$ is generalized with variance replaced by an arbitrary absolute central moment of order > 1. (Received 6 January 1969.)

11. Multivariate analysis of covariance based on general rank scores. P. K. Sen and M. L. Puri, University of North Carolina and Indiana University.

The purpose of the paper is two-fold: (i) to develop the asymptotic distribution theory of the normal theory likelihood ratio test statistic for the general (multivariate) linear hypothesis problem when the underlying distribution is not necessarily normal and (ii) to extend the results of the authors' earlier paper [*Ann. Math. Statist.* **40** (1969), 000] on the univariate analysis of covariance to the multivariate case. It is shown that the normal theory likelihood ratio statistic has asymptotically a chi-square distribution (with appropriate degrees of freedom) when the underlying distribution has finite moments up to the second order. Further, using the results of Puri and Sen [*Sankhyā Ser. A.* **28** (1966) 353–376], the permutation as well as unconditional distribution theory of the rank order tests statistics for the MANOCA problem is studied and the allied efficiency results are considered. (Received 20 January 1969.)

12. Estimation of effects in a fixed effect model. Jagbir Signh, Ohio State University.

A linear model $E(y_i) = \sum_{j=1}^{p} x_{ji}\beta_j$ is considered where $\{x_i\}$ are known inputs and unknown effects $\{\beta_j\}$ are more or less idealized formulations of some properties of interest in the phenomena underlying the observations. Following an idea of Katti (*Biometrics* **18** (1962) 139–147) a sampling scheme is proposed to estimate $\{\beta_j\}$ when some *a priori* knowledge about them is available in the form of guess $\beta_0' = (\beta_{10}, \cdots, \beta_{p0})$. Choose two integers $n \geqq p$ and $r > 1$. Corresponding to $(x_{1i}, \cdots, x_{pi})$ determine one y_i at random; $i = 1, 2, \cdots, n$. The vector $Y_1' = (y_1, \cdots, y_n)$ may be considered as constituting one replication of the experiment with design matrix $X = (x_{ji})$. If the experiment is replicated r times, then $Y = (Y_1, \cdots, Y_r) = (x'\beta, \cdots, x'\beta)' = X'_{nr\times p\times 1}\beta + \epsilon$. Let r_1 and r_2 be two integers to be determined so that $r_1 + r_2 = r$. Define $b_1 = (r_1S)^{-1}x(Y_1 + \cdots + Y_{r1})$ and $b_2 = (r_2S)^{-1} x(Y_{r_1+1} + \cdots + Y_r)$, where $S = xx'$. Estimate the vector β by b as follows: $b = b_1$ if $b_1 \varepsilon R$ and by $(r_1b_1 + r_2b_2)/r$ otherwise, where R is chosen in some optimum way. For instance if we choose R so that $|E(b - \beta_0)(b - \beta_0)'|$ is minimum then it is determined an ellipsoid centred at β_0. Other reasonable criterion are also considered for specifying R. Satistical inference aspect about $\{\beta_j\}$ is investigated. (Received 7 February 1969.)

13. On Bartlett's test and Lehmann's test for homogeneity of variances. Nariaki Sugiura and Hisao Nagao, University of North Carolina and Hiroshima University.

The purpose of this paper is to compare Bartlett's test (modified likelihood ratio test) and Lehmann's test (asymptotically UMP invariant test) for homogeneity of variances of k normal populations. Bartlett's test is known to be unbiased, whereas Lehmann's test is shown to be biased. These two test statistics are known to have asymptotically the same χ^2 distribution with $k - 1$ degrees of freedom under the null hypothesis. We have investigated the limiting distributions under the sequence of alternatives with arbitrary rate of convergence to the null hypothesis, as sample sizes tend to infinity. Depending on the rate of convergence, they are given by χ^2, noncentral χ^2, and normal distributions. Asymptotic expansions of the nonnull distributions of the two test criteria under fixed alternatives, as well as of the null distribution of Lehmann's test, are obtained, by which some numerical examples for the approximate powers are computed. (Received 27 January 1969.)

14. On some statistical inferences in a probabilistic pseudo-metric space (preliminary report). Chia Kuei Tsao, Wayne State University.

Suppose $\{(S, \mathfrak{A}, P_\theta), \theta \varepsilon \Omega\}$ is a family of probability spaces, where $P = \{P_\theta, \theta \varepsilon \Omega\}$ is dominated by some σ-finite measure μ over $(S, \mathfrak{A})$. Let $F = \{f(x; \theta), \theta \varepsilon \Omega\}$ be the family of probability densities with respect to μ. Let $X = (X_1, \cdots, X_N)$ be a random vector on S having pdf in F. We assume: (1) $f(x; \theta_1)/f(x; \theta_2) > 0$ a.e. μ, $\theta_1, \theta_2 \varepsilon \Omega$, (2) $e(\theta_1, \theta_2) = E_{\theta_1}(\ln [f(x; \theta_1)/f(x; \theta_2)]) \geqq 0$, (3) $d_f(\theta_1, \theta_2) = (e(\theta_1, \theta_2) + e(\theta_2, \theta_1))^{1/2}$ satisfies triangle inequality, (4) for some fixed closed subset w of Ω, there exists a continuous function $t(\theta)$ from Ω to w such that for any θ in Ω, the distance $D(w, \theta)$ from θ to w is given by $D(w, \theta) = d(t(\theta), \theta)$, and (5) there exists an unbiased (or consistent) estimator $\hat{\theta} = u(X)$ of θ having a covariance matrix $(1/N^\epsilon)\Sigma$, $\epsilon > 0$ and Σ being a bounded covariance matrix. Under conditions (1), (2) and (3), the pair (Ω, d_f) is a probabilistic pseudo-metric space and under the additional conditions (4) and (5), three random variables $L(X, \theta) =$

ln $[f(X;\theta)/f(X;t(\theta))]$, $L(X;u(X))$ and $d(t(u(X)), u(X))$ may be used in decision functions (such as test statistics or estimators of distances between θ and w). Some properties of these criteria are studied and certain asymptotic optimum properties are discussed. A few well-known statistics (both parametric and non-parametric) may be shown to be special cases of these random variables, e.g., under suitable conditions, the statistics $L(X, u(X))$ and/or $d(t(u(X)), u(X))$ lead to the likelihood ratio criteria and/or asymptotically locally optimum non-parametric statistics. Some new statistics can also be obtained through $L(X, u(X))$ or $d(t(u(X)), u(X))$ by using various suitable $u(X)$. (Received 7 January 1969.)

(Abstracts of papers to be presented at the Western Regional meeting, Monterey, California, May 7–9, 1969. Additional abstracts will appear in future issues.)

1. Some useful bounds in symmetrical factorial designs. BODH RAJ GULATI, Eastern Connecticut State College. (By title)

Let $m_t(r, s)$ denote the maximum number of points that can be chosen in the finite projective geometry PG$(r - 1, s)$ of $r - 1$ dimensions based on Galois field GF(s), where s is a prime or power of a prime, so that any set of t points are linearly independent. It is well known that $m_t(r, s)$ also symbolises the maximum number of factors that can be accommodated in a symmetrical factorial design in which each factor is at s levels, blocks are of size s^r, and no t-factor or lower order interaction is confounded. In a technical report RM-117 of Michigan State University, 1964, E. Seiden studied the maximum number of points in PG$(r - 1, 2)$, no *four* coplanar. We have generalised the following results for any arbitrary $t \geqq 4$.

(i) $m_t(t, 2) = t + 1$, (ii) $m_t(t + 1, 2) = t + 2$, (iii) $m_t(t + 2, 2) = t + 4$ for $t = 4, 5$ and $= t + 3$ for $t > 5$, (iv) $t + 4 \leqq m_t(t + 3, 2) \leqq t + 7$, (v) $t + 5 \leqq m_t(t + 4, 2) \leqq t + 13$. Examples are given demonstrating that for $t = 5$, the upper bounds are achieved in (iv) and (v). (Received 4 February 1969.)

2. A two sample test of equality of coefficients of variation. RONALD K. LOHRDING, University of California Los Alamos Scientific Laboratory.

A likelihood ratio testing procedure, which assumes a normal distribution of the samples, is developed in this paper for testing the equality of two coefficients of variation. The asymptotic distribution of the test statistic is found and the distribution is tabled for small samples. The rapidity of convergence of the small sample distribution to the asymptotic distribution is demonstrated graphically. The power of this test statistic is investigated by simulation techniques. (Received 10 February 1969.)

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

1. Optimum best linear unbiased estimates of the location and scale parameters based on selected order statistics from finite cesored samples (preliminary report). LAI K. CHAN, University of Western Ontario.

Let $X_{(r_1)} < X_{(r_1+1)} < \cdots < X_{(n-r_2)}$, $r_1, r_2, \geqq 0$, be a censored sample corresponding to an ordered random sample of size n from an absolute continuous distribution whose cdf is of the form $F((x - \mu)/\sigma)$, where μ and σ are called the location and scale parameters, respectively. For an integer k, $0 < k \leqq n - r_2 - r_1 + 1$, let $s(n_1, \cdots, n_k)$ be the set of order statistics with ranks $n_1, \cdots, n_k$ taken from the censored sample. An optimum best linear unbiased estimate (BLUE) of μ, σ or (μ, σ) based on $s(n_1, \cdots, n_k)$ is defined to be the BLUE (obtained by the Gauss-Markov theorem. cf. Sarhan & Greenberg, *Contributions to Order Statistics*, (1962), Wiley, Chapter 3) such that its variance (or generalized variance when (μ, σ) is estimated) is the minimum among the variances of the BLUE's based on the choices of $\binom{n-r_1-r_2+1}{k}s(n_1, \cdots, n_k)$. Tables up to 6 decimal places of the ranks, coefficients, variances (covariances), efficiencies (relative to the BLUE's based on the corresponding complete samples) of the optimum BLUE's based on $k = 1, 2, 3, 4$ order statistics from all possible censored samples from the following distributions have been completed: (1) normal distribution, $n = k(1)20$, (2) Cauch distribution, $n = 5(1)20$, (3) double exponential distribution, $n = k(1)20$, (4) logistic distribution, $n = k(1)10, 15, 20$. (Received 18 November 1969.)

2. Marginal homogeneity of multidimensional contingency tables. S. KULLBACK, The George Washington University.

Tests of marginal homogeneity in a two-way contingency table given by Bhapkar (1966), Caussinus (1966), and Stuart (1955) do not seem to lend themselves to extension to the question of m-way marginal homogeneity in an N-way $r \times r \times \cdots \times r$ contingency table, $m < N$. The principle of minimum discrimination information estimation and the associated minimum discrimination information statistic applied by Ireland, Ku and Kullback to the problem of marginal homogeneity in an $r \times r$ contingency table can be easily extended to the case of a multidimensional contingency table. Estimates of the cell entries under the hypothesis of marginal homogeneity are given. Relationships among the tests of homogeneity for m-way, $m = 1, 2, \cdots, N - 1$, marginals are given by an analysis of information. Numerical results are given for two sample $3 \times 3 \times 3$ tables. (Received 8 January 1969.)

3. A bound for the variation of Gaussian densities. S. KULLBACK, The George Washington University.

Schwartz and Root used Mehler's identity to obtain a bound for the integral of the absolute difference between the bivariate gaussian density function and the product of its corresponding marginal densities. The result was also extended to the case of two dependent gaussian vectors. The bounds were given in terms of the correlation coefficient in the bivariate case and canonical correlations in the two vector case. In this note an information-theoretic inequality is applied to derive a better bound than reached above and to extend the result to the case of $m > 2$ dependent gaussian vectors. No series expansion is required as in Schwartz and Root. (Received 10 January 1969.)

4. Locally asymptotically most powerful tests about the effects of K treatments. S. R. KULKARNI, Karnatak University.

In this paper we obtain optimal asymptotic test of the general linear hypothesis in the effects, $\xi_1, \xi_2, \cdots, \xi_k$, of k treatments $T_1, T_2, \cdots, T_k$, when the underlying density involves several parameters which are functions of $\xi_1, \xi_2, \cdots, \xi_k$ and a random variable. More specifically let $t = m$ with probability π_m, $m = 1, 2, \cdots k!$, $\sum \pi_m = 1$. t can be associated with the random choice of one of the $k!$ arrangements of the k treatments, for experimentation. Let Y_{ij} be the response on the jth member of the ith set and $\mathbf{Y}_i = (Y_{i1}, \cdots Y_{ik})$, $i = 1, 2, \cdots, N$. Under a mild assumption the joint density of $(t_i, \mathbf{Y}_i)$ will be $g(t, \boldsymbol{\pi})p(\mathrm{y} \mid \boldsymbol{\theta}_1(\xi_1 g_{11}^{(t)} + \cdots + \xi_k g_{1k}^{(t)}), \cdots, \boldsymbol{\omega}_k(\xi_1 g_{k1}^{(t)} + \cdots + \xi_k g_{kk}^{(t)}), \boldsymbol{\omega})$, where $\boldsymbol{\pi} = (\pi_1, \pi_2, \cdots, \pi_{k!})$, $g(t, \boldsymbol{\pi}) = \pi_m$ if $t = m$, $\boldsymbol{\omega}_j(\xi)$ is a vector of l_j functions and is of the form $\boldsymbol{\omega}_j(\xi) = \boldsymbol{\omega}_j + \xi\boldsymbol{\omega}_j' + \xi^2(2!)^{-1}\boldsymbol{\omega}_j'' + o(\xi^2)$, where $\boldsymbol{\theta}_j$ is a row vector of l_j components and $\boldsymbol{\theta}_j'$ and $\boldsymbol{\theta}_j''$ have obvious meaning, $\boldsymbol{\omega}$, is a vector of s components, and $g_{1j}(t) = 0$ or 1 and is defined to satisfy (i) $\sum_{j=1}^{k} g_{lj}(t) = 1$, $l = 1, 2, \cdots, k$, and (ii) $g_{lj}(t) = 1$ if for the tth arrangement the jth treatment occurs on the lth plot. On the basis of N observed vectors

$(t_i, \mathbf{Y}_i)$, we obtain locally asymptotically most powerful test of general linear hypothesis in ξ's. The linear and multiplicative parametric functions are treated as special cases. (Received 17 January 1969.)

5. Small sample distributions of score product-moment statistics (preliminary report). PETER A. W. LEWIS, IBM Research Center, Yorktown Heights.

The small sample distributions of score product-moment statistics of orders one, two and three have been obtained by synthetic sampling for normal scores, exponential scores, double-exponential scores, uniform scores, half-gamma scores and half-Weibull scores. The rate of convergence to the asymptotic normal distribution depends critically on the skewness of the parent populations, especially in the case of positive random variables. For normal scores the convergence is complete for series of length $n = 40$ or greater, but for exponential scores convergence is complete only for series of length $n = 10{,}000$ or greater. The convergence is even slower for half-gamma and half-Weibull scores. This work is part of a continuing large-scale computational investigation of tests of randomness in time series. (Received 7 January 1969.)

6. Applications of order statistics to the multivariate exponential distribution. RUSSELL MAIK, Southern Colorado State College.

The concept of an order statistic for a univariate distribution is well known. The question naturally arises as to how this concept could be extended to a p-variate random vector. Among the various possibilities are to order the vectors by their minimum or maximum coordinate. For example, the random vector $\mathbf{X}$ is said to be less than or equal to the random vector $\mathbf{Y}$ if $\min_{1 \leq i \leq p} x_i \leq \min_{1 \leq i \leq p} y_i$. These two definitions are examined in detail for both continuous and discrete distributions. The method of ordering the vectors would depend on the type and purpose of the experiment and on the underlying distribution. The results are then applied to the multivariate exponential (as defined by Marshall and Olkin in the *J. Amer. Statist. Assoc.* March, (1967)) when the vectors are ordered by their minimum coordinate. Among the numerous results we obtain are the joint asymptotic distribution of the ith order statistic from the bivariate exponential. The estimation of the parameters of the multivariate exponential using this definition is left for a latter paper. (Received 19 December 1968.)

7. A characterization based on the absolute differences of two independent copies of a random variable. PREM S. PURI, Purdue University.

A discrete random variable X is said to have x as a possible value if $\Pr(X = x) > 0$. Obviously such an X cannot have more than a denumerable number of possible values. Let X_1 and X_2 be two independent copies of a nonnegative discrete random variable X. The following theorem characterizes all the nonnegative discrete distributions with the property that the distribution of the absolute difference $|X_1 - X_2|$ is the same as that of X. The reader may find a different characterization based on $|X_1 - X_2|$ in author's earlier paper (P. S. Puri, *Proc. Nat. Acad. Sci.* **56** (1966) 1059–1061). THEOREM. *Let X_1 and X_2 be two independent copies of a nonnegative discrete random variable X. Then X and the absolute difference $|X_1 - X_2|$ has the same distribution, if and only if the distribution of X is given for some positive constant a, by* $\Pr(X = 0) = p_0$; $\Pr(X = ka) = 2p_0(1 - p_0)(1 - 2p_0)^{k-1}$; $k = 1, 2, 3, \cdots$, *where either $p_0 = 1$ or $0 < p_0 \leq \frac{1}{2}$.* Note that the case with $p_0 = 1$ is that of a degenerate random variable X with $\Pr(X = 0) = 1$, while the case with $p_0 = \frac{1}{2}$, corresponds to the one with $\Pr(X = 0) = \Pr(X = a) = \frac{1}{2}$. (Received 20 January 1969.)

8. Interval estimation of the largest variance of k normal populations (preliminary report). K. M. LAL SAXENA, University of Kansas.

Let $\pi_1, \cdots, \pi_k$ be k normal populations with unknown means μ_i and unknown variance σ_i^2. Let $\sigma^{2*} = \max_{1 \leqq i \leqq k} \sigma_i^2$. There is no *a priori* knowledge about $\mathfrak{d}^2 = (\sigma_1^2, \cdots, \sigma_k^2)$. For preassigned $a(< 1)$ and $\gamma \varepsilon (0, 1)$, the following is the single stage procedure for obtaining a confidence interval I such that $P(\sigma^{2*} \varepsilon I \mid \mathfrak{d}^2) \geqq \gamma$, based on the largest of the sample variances. Take n observations from each population and assert that $I = (aT_n^*, b_nT_n^*)$, where $T_n^* = \max_{1 \leqq i \leqq k} T_{in}$ and $T_{in} = \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2/(n - 1)$. It is proved, under certain conditions, that $\inf_{\mathfrak{d}_2} P(\sigma^{2*} \varepsilon I \mid \mathfrak{d}^2) = P(\sigma^{2*} \varepsilon I \mid \mathfrak{d}_0^2)$ where $\mathfrak{d}_0^2$ has $\sigma_1^2 = \cdots = \sigma_k^2$. The sample size required is the smallest integer n (say n^*) such that $G_n^k(1/a) - (1 - G_n(1/a))^k \geqq \gamma$, where $G_n(t/\sigma_i^2)$ is the distribution function of T_{in}. Then $b_{n*} = \max (1, \inf (b: ae^{1/a} \leqq be^{1/b}), 1/G_{n*}^{-1}(1 - G_{n*}(1/a)))$. Two other types of confidence intervals are considered, namely $I_1 = (aT_{n*}, \infty)$, $a < 1$; $I_2 = (0, bT_{n*})$, $b > 1$. For I_1 the sample size required is the smallest n such that $G_n^k(1/a) \geqq \gamma$. For I_2 the sample size required is the smallest n such that $1 - G_n(1/b) \geqq \gamma$. These results easily extend to the case when the populations belong to a scale parameter family. (Received 11 December 1969.)

9. Estimation of location parameter on the basis of pooling data. J. SINGH, University of California, Berkeley.

Suppose that x_i $(i = 1, \cdots, n_1)$ and y_j $(j = 1, \cdots, n_2)$ represent two independent samples from the distributions $F(x - \theta)$ and $F(y - \theta - \Delta)$ where F is assumed to be symmetric and continuous. If x and y represent measurements on the same biological substance obtained at two different laboratories, a common procedure will consist in deciding on the basis of a preliminary test of the hypothesis $\Delta = 0$ whether or not the samples may be pooled in finding an estimate for θ. Let $X^{(1)} < X^{(2)} < \cdots < X^{(n_1)}$ and $Y^{(1)} < Y^{(2)} < \cdots < Y^{(n_2)}$ be the ordered samples. Let $Z^{(1)} < Z^{(2)} < \cdots < Z^{(m)}$ where $m = (n_1 + n_2)$ be the ordered sample when X's and Y's are combined together. We use Wilcoxon's test for the hypothesis $\Delta = 0$. The following procedure is suggested for estimating θ: (i) If the hypothesis $\Delta = 0$ is accepted, use $T_1 = \text{Med}\, \{\frac{1}{2}(Z^{(i)} + Z^{(j)})\}$ $(1 \leqq i \leqq j \leqq m)$ as an estimate of θ. (ii) If the hypothesis $\Delta = 0$ is rejected, use $T_2 = \text{Med}\, \{\frac{1}{2}(x^{(i)} + x^{(j)})\}$ $(1 \leqq i \leqq j \leqq n_1)$ as an estimate of θ. We denote this estimate by T^*. The distribution of T^* and its other asymptotic properties are investigated. We also plan to compare this estimate with the Hodges-Lehmann (1963) estimate, here denoted by T_2. Next let us consider the problem of interval estimation for θ. Let $D^{(1)} < \cdots < D^{n_1(n_1+1)/2}$ be the ordered set of $n_1(n_1 + 1)/2$ averages $(x_i + x_i)/2$ $(i \leqq j)$ and $E^{(1)} < \cdots < E^{m(m+1)/2}$ be the ordered set of $m(m + 1)/2$ averages $(Z_i + Z_j)/2$ $(i \leqq j)$. If V is the Wilcoxon one sample statistic based on n_1 observations we can get two numbers c_1 and c_2 such that $P[c_1 < V \leqq c_2] = 1 - \alpha$. c_1' and c_2' can be similarly obtained for a sample of size m. We then give the following procedure for determining a $100(1 - \alpha)$ percent confidence interval for θ: (i) If the hypothesis $\Delta = 0$ is accepted, use the limits $E^{(c_1'+1)} \leqq \theta \leqq E^{(c_2'+1)}$. (ii) If the hypothesis $\Delta = 0$ is rejected, use the limits $D^{(c_1+1)} \leqq \theta \leqq D^{(c_2+1)}$. (Received 3 February 1969.)

10. Λ-minimax estimates (preliminary report). DANIEL L. SOLOMON, Cornell University.

Statistical decision problems are considered in which the decision maker is assumed to have prior information but cannot completely specify a prior distribution. His prior knowledge is reflected in his willingness to specify a subset, Λ, (called an incompleteness specification), of the class of all prior distributions. He is then recommended to select a decision rule to minimize the maximum over distributions in Λ of the Bayes risk. Such a rule is called Λ-minimax after Blum and Rosenblatt (*Ann. Math. Statist.* **38** (1967) 1671–78).

Estimation of the mean of a p-variate random variable with known covariance matrix is treated. Here Λ is the class of all prior distributions with given covariance matrix and mean in a specified compact convex set. For quadratic loss, the linear (of the form $\delta(x) = Bx + C$) Λ-minimax rule is obtained. Also obtained are Λ-minimax estimates of a (univariate) normal scale parameter when the mean is (1) known and (2) unknown. Problem (2) is reduced to problem (1). Finally, a cost is postulated of obtaining an incompleteness specification. The design problem consists of selecting a sample size and an incompleteness specification to minimize the total expected loss. Examples are given. (Received 3 January 1969.)

11. Equivalence of "test deficiency" and general deficiency for dichotomies.

Erik Nikolai Torgersen, University of California, Berkeley.

In this paper we consider dichotomies, i.e. experiments where the parameter space contains two points. It is shown that to any dichotomy F and any ordered pair (e_1, e_2) of non negative numbers there correspond a dichotomy G on the same sample space which represents the minimum of all dichotomies (regardless of sample space) which are (e_1, e_2) deficient (as defined by LeCam in 1964) relative to F. The construction implies an extension of the "errors of the first and the second kind" criterion for comparison of dichotomies given by Blackwell in 1953. In particular it is shown that the Lévy diagonal distance between the curves representing the relation between errors of the first and the second kind is a natural distance for dichotomies. It is shown that this distance is, but for a trivial modification, the same as that defined by LeCam in 1964. The criterion implies for dichotomies a result proved by LeCam, that this distance is equivalent to the Lévy diagonal distance between laws of likelihood ratios. (Received 13 January 1969.)