STATISTICS PROBLEMS

I. Richard Savage

Special Report No. 6

University of Minnesota
Minneapolis, Minnesota
TABLE OF CONTENTS

Introduction ..............................................

Mathematical Problems .................................

Problems from References ..............................

Applied Problems .........................................

(i)
Introduction

Characteristically, the problems in this collection are stated out of context. Thus, it is necessary to have a fair amount of maturity to even understand the statement of the problems. Having understood the meaning of the parts of a problem, the reader must then locate the appropriate body of theory which can help in the solution of the problem. In most problem collections these preliminaries are handed to the reader in advance; most problem collections are organized in terms of subject matter, and in textbooks the relevant theory usually precedes each "set" of problems. The problems in this collection are on the average quite difficult; some of them can serve as the basis for serious research. Solutions of the problems can be quite involved and can be at several levels of generality.

This collection of problems has grown over a period of ten years. The problems arose in lectures, in written preliminary and oral examinations, in the literature, and in the course of research. No attempt has been made to make the collection comprehensive or well balanced. Ability to solve these problems will not necessarily make one a better statistician, but will suggest that one can handle much of the classical analysis required in research in mathematical statistics.

A reader should not expect to be able to go through and do the problems in order. Rather the reader should attempt to work a problem and if his work does not seem fruitful he should pass on to another problem. The reader will note that the problems have not been arranged in order of difficulty.

The "mathematical problems" are self contained statements which require a proof of some assertion. The well informed reader should be able to solve these problems with no further assistance. In practice, however, their
Introduction

solution will often require reading in relevant textbooks or research literature. The "problems from references", instruct the reader to look at some specific point of the literature and to verify a statement made at that point; often the statement originally contained an "obvious" or "it can easily be shown". Solutions to these problems should be easier than the "mathematical problems" since the reader is led directly to the most relevant reference. These problems were included to give the reader some experience in working out details of reading other people's research. The "applied problems" usually involve some data and its numerical analysis. More important, the solution of these problems requires the reader to make some reasonable assumptions about the relevant distributions from which the data were obtained. Presumably, the experience required on the part of the reader, is either actual work in applying statistics or else fairly extensive reading in the applications of statistical theory. Since these problems do not have "correct" solutions, as the others usually do, the reader can never feel well satisfied with his handling of them. The collection is strongest in its representation of "mathematical problems", and weakest in "applied problems".

The author will appreciate receiving comments about this collection. In particular, errors if discovered and communicated, would help in improving future editions. Additional problems would be desirable for expanding this collection.

Finally, the reader must be on guard for errors in statements of problems, i.e., he should be looking for counter examples.
Mathematical Problems

** * * * * *

Suppose \( k \overset{d}{=} p(k; \lambda) \) and \( x \overset{d}{=} N(x; 0,9) \), \( k \) and \( x \) independent. Let \( y = \alpha k + x \).

This model might arise in a situation in which the Poisson distributed number of events \( (k) \) cannot be counted, but a measurement proportional to \( k \) (i.e., \( \alpha k \)), can be obtained with measurement error, \( x \). Suggest estimators for \( \alpha \) and \( \lambda \), based on \( n \) independent observations on \( y, y_1, y_2, \ldots, y_n \). If the variance of \( x \) were not known, how would you estimate the three unknown parameters?

** * * * * *

You can place fifty white and fifty black balls in two urns in any manner desired. Your opponent then selects one of the urns at random and selects one of the balls that it contains at random. How can you minimize [maximize] the probability that he will select a white ball?

** * * * * *

Let \( \{X_n\} \) be independent R. V.'s and let \( \sigma_n^2 \) be a sequence of positive constants with \( \sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} = \infty \). Define

\[
P(X_n = n) = P(X_n = -n) = \frac{\sigma_n^2}{2n^2}; P(X_n = 0) = 1 - \frac{\sigma_n^2}{n^2} \quad \text{if } \sigma_n \leq n;
\]

\[
P(X_n = \sigma_n) = P(X_n = -\sigma_n) = \frac{1}{2} \quad \text{if } \sigma_n > n.
\]

Show that \( \{X_n\} \) does not obey the SLLN. (To this extent, the Kolmogoroff condition is "necessary").

** * * * * *

Consider a Poisson process such that the probability of no occurrences in an interval of length \( t \) is \( e^{-mt} \). Let \( T \) be the time required to observe the first occurrence, and let \( N[K/T] \) be the number of failures in the next \( K/T \) units of time. Find the first two moments of \( N[K/T] T/K \).
Mathematical Problems

** * * * * *

Let $\mathbf{X} = (X_1, \ldots, X_{2n})$ be a random row-vector with $\mathbb{E}\mathbf{X} = \theta \mathbf{A}$, where $\theta = (\theta_1, \ldots, \theta_k)$, $k \leq 2n$, is unknown and $\mathbf{A}$ is a $k \times 2n$ known matrix of rank $k$. The covariance matrix of $\mathbf{X}$ consists of $n$ $2 \times 2$ diagonal blocks, each of which is $\left( \begin{smallmatrix} 3 & 3 \\ 3 & 2 \end{smallmatrix} \right)$, all the remaining elements being 0. Obtain the least-squares estimate $\hat{\theta}$ of $\theta$.

Under what further assumptions on $\mathbf{X}$ is it true that for any given $k$-dimensional vector $\ell$, the unique u.m.v. unbiased linear estimate of $\ell \theta'$ is $\ell \hat{\theta}'$? Sketch a proof of your statement.

** * * * * *

Let $(X_n, n = 1, 2, \ldots)$ be a sequence of independent and identically distributed random variables, and let $S_n = X_1 + \ldots + X_n$. State the most general form that you know of sufficient conditions on the $X$-distribution in order that $S_n$, suitably normalized, should have a normal limiting distribution. Give a precise statement of the result, and indicate its proof.

** * * * * *

Let $X_1, \ldots, X_m$ be a sample from $F[X]$ and $Y_1, \ldots, Y_n$ be a sample from $G[X]$. Assume that $F$ and $G$ are continuous c.d.f.'s, and that $F = G^\Delta$ where $\Delta > 0$.

Find:

a. $P$ [exactly $i$ of the $X$'s are larger than all of the $Y$'s]

b. $P$ [exactly $i$ of the $X$'s are less than all of the $Y$'s]

c. Consider $i$ of part a, as a random variable and find its moments.

d. Discuss limits of the quantities found in a and b as the sample sizes tend to infinity.

** * * * * *

If $P(X < Y) = 1$, then $\mathbb{E}(XY) \geq \mathbb{E}X \mathbb{E}Y$.

-2-
Mathematical Problems

Let \([X_1]\) be a sequence of constants defined on \([\Omega, \mathbb{A}, \mathbb{P}]\) and \(c, c_1, c_2, \ldots\) sequence of constants. Prove or counter:

(i) \(X^{(i)}_{n} \xrightarrow{a.c.} X^{(i)}, \ i=1, \ldots, K, \) and \(g(X^{(1)}, \ldots, X^{(K)})\) continuous imply \(g(X^{(1)}, \ldots, X^{(K)}) \xrightarrow{a.c.} g(X^{(1)}, \ldots, X^{(K)})\)

(ii) \(X_n + c_n \xrightarrow{P} X, \ X_n \xrightarrow{P} X - c\) imply \(c_n \rightarrow c\)

(iii) \(X_n - c_n \xrightarrow{L} 0, \ X_n \xrightarrow{P} c\) imply \(c_n \rightarrow c\)

Let the following model be given:

\[ Y_i = \alpha x_i + u_i \]

where the \(x's\) and \(y's\) are observed, the \(x's\) are fixed variates, the \(u's\) are jointly normally distributed with means zero and a covariance matrix which is known up to a constant of proportionality (and is independent of the \(x's\)).

What is the maximum likelihood estimator of \(\alpha\)? Is the least squares estimator of \(\alpha\) unbiased? (Justify your answers.)

Define precisely what is meant by a balanced incomplete block design.

State a standard fixed-effects model, indicate the least squares estimates of the parameters and give the analysis of variance.

Prove that if the convolution of two \(k\)-dimensional c.d.f.'s is the multivariate normal c.d.f., the components (if non-degenerate) are likewise. Hint: Reduce matters to the univariate case and apply the result for \(k=1\).

Covariances of order statistics are \(> 0\).
Mathematical Problems

* * * * *

Consider a set of $k$ given populations with densities, all on the interval $[0,1]$, given by

$$
\frac{k!}{(i-1)!(k-i)!} x^{k-i} (1-x)^{i-1} \quad (i=1,2,\ldots,k).
$$

It is known that each population corresponds to one of these densities and no two have the same density, but no a priori information is given about the exact correspondence between the $k$ populations and the $k$ values of $i$. A fixed sample of size $n$ is taken from each population. On the basis of this, we wish to select the population with the highest value of $i$. If $n=1$, what rule would you use? Write an expression for the probability of a correct selection for $n=1$. If we take $n > 1$ observations, is there a sufficient statistic for each population? What rule would you use for general $n$ ($n \geq 1$) to select the best population?

* * * * *

Let $[X_i]$ be a sequence of R.V.'s defined on $(\Omega, A, P)$ and $(c, c_1, c_2)$ sequence of constants. Construct a sequence of R.V.'s (i) converging in probability (but not a.c.) to a non-constant R.V., (ii) converging in r-mean (all $r > 0$) and almost certainly. The theorem of Section 20.6 of Cramer's "Mathematical Methods of Statistics" is an immediate consequence of what previously examined result?

* * * * *

The following experiment is performed. An observation $x$ is made of the Poisson random variable $X$ with parameter $\lambda$. Then a binomial event with probability $p$ of success is repeated $x$ times and $\sigma$ successes are observed. What is the distribution of the random variable $R$ associated with $\sigma$?
Suppose that $X_1, X_2, \ldots$ is a sequence of random variables converging in distribution to the random variable $X$. Suppose further that $X$ and the $X_i$'s are all independent. In a recent statistical article it was asserted that, given the above, $X_i$ converges to $X$ in probability.

Show, by a simple counter-example, that this assertion cannot in general be true.

The article mentioned above gives the following "proof" for its assertion. Here $\phi_Y(t)$ denotes the characteristic function of $Y$.

\[
\phi_{X_i - X}(t) = \phi_{X_i}(t)/\phi_X(t) \quad \text{(by independence)}
\]

\[
\lim_{i \to \infty} \phi_{X_i}(t) = \phi_X(t) \quad \text{(by hypothesis)}
\]

\[
\therefore \lim \phi_{X_i - X}(t) = 1
\]

\[
\therefore X_i - X \text{ converges in distribution to that distribution}
\]

\[
\text{with all its mass at zero.}
\]

\[
\therefore X_i - X \to 0 \text{ in probability.}
\]

\[
\therefore X_i \to X \text{ in probability.}
\]

Point out exactly wherein this proof is faulty. Under what additional conditions on the distribution of $X$ (if any) does the assertion hold? Give a proof of your conclusion.

If $X_i$, and $X_j$, are order statistics with $i < j$, then find parent distributions such that

\[
E(X_j/X_i = E X_i) = E X_j
\]

Find uncorrelated, dependent random variables with normal marginal distributions.
**Mathematical Problems**

* * * * *

Let \( \{X_n\} \) be independent R.V.'s.

i. \( P(X_n = b_n) = P_n = 1 - P(X_n = -b_n) \), Characterize the set of sequences \( (b_n, P_n) \) for which \( \sum_{i=1}^{n} X_i \) converges a.e. Specialize for \( P_n = \frac{1}{2} \); for \( b_n = b \); for \( b_n = n^{-\alpha} \).

ii. \( P(X_n = n + n^{-2}) = P_n = 1 - P(X_n = n - n^{-2}) \). For what values of \( P_n \) does \( \sum_{i=1}^{n} X_i \) conv. a.e.? conv. a.e. when centered? Select a sequence of centering constants.

iii. \( X_n \) is \( N(\mu_n, \sigma_n^2) \). Characterize the parameter set for which \( \sum_{i=1}^{n} X_i \) conv. a.e. when centered.

iv. \( X_n \) is uniformly distributed in \((-b_n, b_n)\). For what sequences \( b_n \) does \( \sum_{i=1}^{n} X_i \) conv. a.e.?

* * * * *

Prove the Levy-Polya inversion formula for the multivariate case: If \( I \) is a continuity interval, defined by \( a_i < b_i, \ i=1,\ldots,k \)

\[
P(I) = \lim_{C \to \infty} \frac{1}{(2C)^k} \int_C \cdots \int_C \pi e^{-it_j a_j - it_j b_j} \varphi(t_1, \ldots, t_k) dt_1, \ldots, dt_k
\]

* * * * *

Suppose \( X_1, \ldots, X_n \) are the ordered values of independent and identically distributed exponential R.V.'s. Determine the joint distribution of the successive differences between the \( X_i \)'s, i.e., \( \{X_2 - X_1, X_3 - X_2, \ldots, X_n - X_{n-1}\} \).
Mathematical Problems

* * * * *

Prove the "Generalized Tchebycheff Inequality"

\[ P(\{|Y| \geq C\} < \frac{E[g(Y)]}{g(C)} \]

where \( g(Y) \) is non-negative, non-decreasing (for \( Y > 0 \)) and an even function, while \( C \) is an arbitrary positive constant. Choose \( Y = X - EX \) and \( g(Y) = |Y|^k \), \( k > 0 \) to obtain more conventional types of Tchebycheff Inequalities.

Demonstrate that if \( P(\{|Y| \leq M\} = 1 \), for every \( C > 0 \)

\[ P(\{|Y| \geq C\} > \frac{E[g(Y)]-g(C)}{g(M)} \]

Show that even if \( Y \) is not a bounded R.V. that the above remains true with \( g(M) \) replaced by the (almost everywhere) sup \( g(X) \). Also prove if \( M \) is median of \( X \), \( \sigma^2 \) its variance, then \( EX - \sigma \sqrt{2} \leq M \leq EX + \sigma \sqrt{2} \).

* * * * *

Let \( \{X_i\} \) be a sequence of R.V.'s defined on \((\Omega, A, P)\) and \( c, c_1, c_2, \ldots \) sequence of constants. Prove that if \( g(x) \) is non-negative, even and monotonically increasing for \( x > 0 \) that each of (i) and (ii) imply \( X_n \xrightarrow{P} X \).

(i) \[ \text{Eg}(X_{n+p} - X) < \epsilon \text{ for all } \epsilon > 0, \ p > 0 \text{ and } n > N. \]

(ii) \[ \lim_{n \to \infty} \text{Eg}(X_n - X) = 0. \]

Also, if \( g(x) \) is continuous and bounded with \( g(0) = 0 \), (i) and (ii) are necessary that \( X_n \xrightarrow{P} X \).

* * * * *

Prove that if \( F(x) \) is a c.d.f., \( \overline{F}(x) = 1 - F(-x+0) = 1 - \lim_{y \to -x} F(y) \) is also a c.d.f. Show that \( G = F \ast \overline{F} \) is a symmetric c.d.f. If \( M \) is the median of \( F \), what is the median of \( \overline{F} \)?
Mathematical Problems

Let \( m(i) \) be a non increasing function, \( N \) is a non negative integer and
\[
0 < p = 1-q < 1 \text{ then }
\]
\[
\sum_{i=0}^{N} \binom{N}{i} p^i q^{N-i} m(i) > \sum_{i=0}^{N+1} \binom{N+1}{i} p^i q^{N+1-i} m(i)
\]

A NSC that \( F(x) \) belongs to the domain of attraction of a stable law with parameters \( c_1, c_2, \alpha \in (0, 2) \) is that
\[
\lim_{x \to \infty} \frac{F(-x)}{1-F(x)} = \frac{c_1}{c_2} \quad \text{and} \quad \lim_{x \to \infty} \frac{1-F(x)+F(-x)}{1-F(\lambda x)+F(-\lambda x)} = \lambda^\alpha, \text{ all } \lambda > 0.
\]

\( X \) and \( Y \) are said to be mutually symmetric if \( F_X(-t) = 1-F_Y(t) \). Is this an equivalence relationship? If \( X \) and \( Y \) are independent and mutually symmetric, what can you say about the distributions of \( X+Y \) and \( X-Y \)?

Explain the distinction between a ranking problem and a test of a "null" hypothesis. Under what circumstances is a ranking formulation appropriate? When is it more appropriate than a test of a null hypothesis? Illustrate your answers by formulating a problem in both ways. Then give a critical comparison of the two formulations.

Find a sequence of random variables such that no moments exist for any of the random variables in the sequence and yet the sequence converges in probability to a normal random variable.
Mathematical Problems

** ** ** **

Show
\[ \sum_{i=0}^{c} \binom{m}{i} p^i q^{m-i} \geq \sum_{i=0}^{c} \binom{m+n}{i} p^i q^{m+n-i} \]

where \( c, m, \) and \( n \) are integers and \( 0 \leq p, 1-q \leq 1 \). When does the equality hold?

** ** ** **

If \( F_n \rightarrow \) two c.d.f.'s \( F \) and \( G \) at their continuity points, show that \( F(x) \equiv G(x) \). Further if \( m \) and \( m_n \) are the unique medians of the c.d.f.'s \( F \) and \( F_n \), \( n=1,2,\ldots \) and \( F_n \rightarrow F \), then \( m_n \rightarrow m \). Can an analogous statement be made if the medians are not unique?

** ** ** **

Let \( \{X_n\} \) be independent. Does \( \{X_n\} \) obey the SLLN for

i. \( P(X_n = t 2^n) = \frac{1}{2} \)

ii. \( P(X_n = t 2^n) = 2^{-(2n+1)} \)

\( P(X_n = 0) = 1 - 2^{-2n} \)

iii. \( P(X_n = t n) = \frac{1}{2} n^{-\frac{1}{2}} \)

iv. \( P(X_n = t 1) = \frac{1}{2} (1 - 2^{-n}) \)

\( P(X_n = 0) = 1 - n^{-\frac{1}{2}} \)

\( P(X_n = 2^n) = 2^{-(n+1)} \)

** ** ** **

Consider a recurrent event with independent and identically distributed positive integer valued waiting times between recurrences. At time \( n \) let \( Y_n = \) elapsed time since last recurrence. Show that the sequence \( \{Y_n\} \) constitutes a Markov chain. Determine the transition probabilities. Give a necessary and sufficient condition that there exists an initial stationary distribution, and determine it when it exists.

** ** ** **

Present a singular distribution.

-9-
Mathematical Problems

** * * * *

Using an example from any field you wish, explain to a reasonably intelligent (but non-statistically trained) individual the meaning of the following multiple regression concepts:

a. Partial regression coefficient.
b. Intercorrelation.
c. Auto-correlation of residuals.
d. Standard error of regression.

** * * * *

In the analysis of variance dealing with "linear hypotheses" a Pythagorean n-dimensional triangle plays an important role in the geometric interpretation of the analysis. Explain how this comes about; what the various parts of this triangle signify; and how the geometry helps us to obtain the appropriate F-statistic. Illustrate your answer.

** * * * *

If \( X_1, X_2, \ldots, X_n \) are independent observations on a binomial distribution with \( P(X_i = 1) = p \), find the Bayes estimate of \( p \) if the a priori distribution of \( p \) is

a. \( C(a,b) p^{a-1}(1-p)^{b-1} \) \( a,b \) known

b. \( P(p = \frac{i}{n}) = \frac{1}{n+1} \) \( i=0,1,\ldots,n \)

c. \( P(p = \frac{1}{4}) = P(p = \frac{3}{4}) = \frac{1}{2} \)

Use squared error as your loss function.

** * * * *

Write a short essay on tests of structure \( S(\alpha) \). Proofs are not asked for. Proper arrangement and careful statement of all relevant results are important.
Mathematical Problems

*****

What is a latin square design? What is meant by a pair of orthogonal latin squares? Illustrate your answers. What interesting results can you mention about $6 \times 6$ latin squares and about latin squares of side length $4t+2$, $t$ an integer? Show the breakdown in degrees of freedom and the analysis of a $6 \times 6$ latin square if there are two observations in each cell.

*****

What are retrospective studies? What are their advantages and what are their disadvantages? What would be the opposite kind of study? What would be its relative merits?

*****

Show that if $F_n(x) \rightarrow F(x)$ for all $x$, $F_n(x)-F_m(x) \rightarrow 0$, all $x$ and that if $F_n(x)-F_m(x) \rightarrow 0$ uniformly in $x$ that $F_n \rightarrow F$, a c.d.f.. Is the latter true if $F_n(x)-F_m(x) \rightarrow 0$ for all $x$?

*****

Prove the simpler Helly-Bray Theorem that $\lim\limits_{n \rightarrow \infty} \int_a^b g(x) dF_n = \int_a^b g(x) dF$ where $F_n \rightarrow F$, a c.d.f., $a$ and $b$ are continuity points of $F(x)$ and $g(x)$ is any continuous function.

*****

Let $[X_i],[Y_i]$ be sequences of R.V.'s defined on $[\Omega,A,P]$. Demonstrate that $X_n \xrightarrow{c} 0 \Rightarrow X_n \xrightarrow{a.c.} 0$ and that the converse is true if the $X_n$ are independent. Can independence be eliminated?

*****

If $F$ is a step function with $n$ points of increase and $F = F_1 * F_2 * \ldots * F_m$ where $F_i$ are any non-degenerate (proper) c.d.f.'s, what are the possible (integral) values of $m \geq 2$?
Mathematical Problems

***

Some of the more common distributions in statistics have interesting relationships. Can you point out any relationship between the following pairs?

a. The binomial distribution and the incomplete beta distribution.
b. The Gamma distribution and the chi-square distribution.
c. The Poisson distribution and the Gamma distribution.
d. The exponential distribution (with density \( \frac{1}{\theta} e^{-x/\theta} \) \( x > 0, \theta > 0 \)) and the chi-square distribution.
e. The F-distribution and the t-distribution.
f. The t-distribution and the Cauchy distribution.
g. The normal distribution and any other with finite second moment.

***

The nonparametric analogue to quadratic regression could be convexity or concavity of the regression function. How could one test nonparametrically this type of hypothesis?

***

Let \( X_n(\omega) \) be a sequence of R.V.'s defined on \( (\Omega, \mathcal{A}, \mathbb{P}) \) with \( P_n[B'] = P \left( \frac{X_n}{n} \right) \)

\( B' \in \mathcal{B} \) (Borel sets of \( R' \)) the induced probability measures and \( F_n(x) = \frac{X_n}{n} \)

the corresponding c.d.f.'s. Prove that a NSC that \( F \to F, \) a c.d.f., is that \( P_n[B'] \to P[B'], \) a probability measure, for every "continuity interval", \( B' \in \mathcal{B}. \)

Note: An interval \( (a, b) \) is a "continuity (or continuity-bordered) interval" of a measure \( P \) if \( P(a) = P(b) = 0. \) How is \( P \) related to \( F? \)

***

Suggest several statistical procedures that encourage one to look at the data in order to find new hypotheses.
Mathematical Problems

** * * * *

1. Define the term 'confidence interval'.

2. Suppose you have drawn a sample of four observations: 3, 5, 2, 10.

3. Assuming that the population is known to be normal (Gaussian), construct the confidence interval with a given (by you) confidence level.

4. Perform the same task as under No. 3 above, but with the additional assumption that the population variance is known to equal 10.

5. Do you know about difficulties involved in constructing confidence intervals due to the presence of 'nuisance parameters'? If so, explain.

6. Under what circumstances is the confidence interval a useful tool of statistical inference?

7. Do you know of controversies related to confidence intervals?:
   a. as against 'fiducial limits';
   b. as to their usefulness (the views of L. J. Savage).

** * * * *

Research problem.

To get an estimator with a sample of N, find the best systematic estimator with a sample of size n < N and then average this over all samples of size n. The estimator that is to be used comes out as a simple linear function in terms of the original N ordered random variables. The problem of finding the variance of this estimator is not trivial. Since the proposed statistic is a U statistic it is normal for large samples and probably is a good statistic to use. It is not clear that it is much better than simply averaging over disjoint sets of observations.

** * * * *

A NSC that \( \text{EX}^2 < \infty \) is \( \sum_{j=1}^{\infty} 2^{2j} P(|X| > 2^j) \). How can this be generalized?

-13-
Mathematical Problems

** ** ** **

In least squares theory suppose

\[ \begin{align*}
EX_1 &= \theta_1 + \theta_2 \\
EX_2 &= \theta_1 + \theta_3 \quad \text{and} \quad \sigma(X_i X_j) = \sigma^2 \delta_{ij} \quad \text{(Kronecker delta)} \\
EX_3 &= \theta_2 - \theta_3
\end{align*} \]

Let \( A \) be defined by writing \( EX = A\theta \). What is the rank of \( A \)? Can one obtain a linear unbiased estimate of any of the \( \theta_i \)? Characterize the functions of \( \theta \) that can be so estimated. Is this related to the concept of estimability? Define this concept. Does one need the assumption of normality above? What does the assumption of normality enable one to do?

** ** ** **

Find the covariance between order statistics from the exponential distribution. (Hint: See relevant table in Sarhan and Greenberg's Order Statistics.

** ** ** **

One observation is taken on a geometrical distribution

\[ P(X = x) = q^{x-1} p \quad (x=1,2,...) \]

where \( p = 1-q \) is such that \( 0 \leq p \leq 1 \). It is desired to estimate \( p \). It is known that an unbiased estimate is \( \delta \) given by

\[ \delta(X) = 1 \text{ if } X=1 \]
\[ = 0 \text{ if } X \neq 2 \]

Show it is the best unbiased estimate. Show it is the only unbiased estimate. Does this estimate throw away "information"? Discuss this question.

** ** ** **

What sources are there for the history of statistics and probability theory?
Mathematical Problems

* * * * *

1. Define a class of distributions (the wider the better) \( F(x; \theta) \), with \( \theta \) denoting the population mean, such that the sample mean is a consistent estimator of \( \theta \). (Define the term 'consistent' in a rigorous manner, but also give a 'popular' verbal interpretation.) Give precise reference to the theorem(s) on which your answer is based.

2. Give an example where the sample mean is not a consistent estimator of \( \theta \). Sketch a proof of lack of consistency.

3. If you were teaching an elementary course in statistical theory (say at Mood level), how would you explain the desirability of having consistency among the properties of an estimator?

4. Could you imagine a situation where you would prefer to use an estimator lacking consistency rather than one having this property? If not, explain your reasoning. If yes, give an example.

* * * * *

When is the two sample Wilcoxon test the locally most powerful rank test, i.e., characterize all of the situations.

* * * * *

Prove

\[
S_n = \sum_{i=1}^{n} (X_i - \bar{x})^2
\]

where \( \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \) is an increasing function of \( n \).

* * * * *

Prove that if \( \sum_{i=1}^{n} (X_i - C_i) \) conv. a.e. and \( \sum d_i < \infty \) and \( C_i' = C_i + d_i \) that

\[
\sum_{i=1}^{n} (X_i - C_i') \text{ conv. a.e.}
\]
Mathematical Problems

Let $X_1, \ldots, X_n$ be a sample from a population with density function $f$,

$$f(x) = \begin{cases} 
2(\pi \sigma^2)^{-\frac{1}{2}} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} & \text{if } x > \mu, \\
0 & \text{otherwise},
\end{cases}$$

$\mu$ and $\pi$ being unknown. Find a minimal sufficient statistic for $(\mu, \sigma)$, and prove normality. Is the minimal sufficient statistic complete? Justify your answer.

For some combinations of parameters a good approximation to the binomial is given by the normal distribution, and for some other combinations good approximations are given by the Poisson. Do there exist combinations such that the normal and Poisson are both good approximations?

Is it true that if $F_1(x)$ is a step function c.d.f. with $n_1$ points of increase, that $F = F_1 * F_2$ has at least $n_1 + n_2 - 1$ points of increase? Consider also the corresponding statement with the c.d.f.'s not necessarily step functions.

If $X_1, \ldots, X_n$ are independently and identically distributed from a Cauchy distribution with a location parameter, find the minimal sufficient statistic.

Let $\{X_n\}$ be independent identically distributed with finite variances.

Take $A_n \to 0$ and $S_n = \sum_{i=1}^{n} X_i$. Show that $\{A_i, S_i\}$ obeys the WLLN but not $\{S_i\}$

i.e. $\frac{1}{n} \sum_{i=1}^{n} (S_i - ES_i) \overset{p}{\to} 0.$
Discuss fully the asymptotic properties of Wilcoxon's one sample test (Fisher's test using ranks). There are \( n \) given values of a random variable \( X \). The null hypothesis is that the distribution of \( X \) is symmetrical about 0. Show that the two-sided test is consistent against

\[
2p_1p_2p + p_1^2 + \frac{1}{2},
\]

where \( p_1 = P(X > 0) \), \( p_2 = P(X < 0) \), \( p = P(X_1 + X_2 > 0) \),

where \( X_1 \) is a random positive value of \( X \), and \( X_2 \) is a random negative value of \( X \).

If the frequency function of \( X \) is \( f(x-\theta) \), where \( f(x) \) is even in \( x \), calculate the resolving power of the test at \( \theta = 0 \). Justify any assumptions of asymptotic normality.

Find the distribution of the waiting time required to get at least one stone in every box when each toss is independent of all of the others and the probability that the stone goes into box \( i \) is \( p_i \) where \( i = 1, \ldots, I \). Prove that the waiting time is finite with probability one when all of the \( p_i \)'s are positive.

Let \( F, f \) be respectively the c.d.f. and the density function of the standard normal distribution. Prove the following relations with an arbitrary:

\[
\begin{align*}
(a) \quad & \int_{-\infty}^{\infty} F(x+a) f(x) dx = F(a/\sqrt{2}), \\
(b) \quad & \int_{-\infty}^{\infty} F(ax) f(x) dx = \frac{1}{a}, \\
(c) \quad & \int_{0}^{\infty} F(ax) f(x) dx = \frac{1}{2} + \left(1/\pi\right) \arctan a.
\end{align*}
\]
Consider a random sample of \( n \) from a normal distribution with zero mean. Under the null hypothesis its variance is 1 and under the alternatives its variance exceeds 1. Take as given the fact that a uniformly most powerful test of given size for the null hypothesis is obtained by rejecting when the sum of squares of the observations is too large. Further any other critical region of the same size, with power equal to that of the test just mentioned at a single alternative, differs from it on a set of zero Lebesgue measure.

Show that for each alternative distribution the power of the above test is strictly increasing in \( n \).

Suppose \( \theta \) has cumulative distribution function \( 1 - \exp(-\theta) \) for \( \theta \geq 0 \) and 0 for \( \theta < 0 \), and \( X \) has a Poisson distribution with parameter \( \theta \). Given \( X = x \), find the posterior distribution of \( \theta \). If \( d(X) \) is an estimate of \( \theta \), and losses are of the form \((d(x)-\theta)^2\), choose \( d(x) \) to minimize \( E(d(x)-\theta)^2 \), with expectations over \( x \) and \( \theta \).

Let \( S_n = \sum_{i=1}^{n} X_i \) where \( \{X_i\} \) are independent and identically distributed nondegenerate R.V.'s. Prove that for any \( C > 0 \), \( \lim_{n \to \infty} P(|S_n| > C) = 1 \).

Let \( \{X_n\} \) be independent of \( X_j \) for all \( n \neq j+1, j-1 \) and \( \sigma_n^2 < cn^\alpha \), \( 0 \leq \alpha < 1 \). Then \( \{X_n\} \) obeys the WLLN.

Show that the variance covariance matrix of the order statistics from the normal distribution (with variance one) is doubly stochastic.
Mathematical Problems

** ** ** **

Consider the linear hypothesis model, \( EX = Ab \), where \( X = \{x_1, \ldots, x_n\} \) is a column vector of uncorrelated random variables with common, finite 2nd moment. \( A \) is a matrix of known constants, and \( b = \{b_1, \ldots, b_r\} \) is a column vector of unknown constants. The variance is not known.

1. When is it possible to find linear estimators of the form \( \sum c_i b_i \)? (\( c_i \)'s are known constants.)

2. When \( \sum c_i b_i \) is estimable by linear estimators, what is the minimum variance unbiased estimator in the class of linear estimators?

3. Impose the additional assumption that the \( X_i \)'s are normally distributed. Show that the answer to 2 is a minimum variance estimator in the class of all unbiased estimators (not just linear ones.)

** ** ** **

Suppose \( k_i \) has a Poisson distribution with parameter \( \lambda \), for \( i = 1, 2, \ldots, n \). Suppose the \( k_i \) are independent. Let \( y = (1/n)(k_1 + k_2 + \ldots + k_n) \). What are the mean and variance of \( y \)? If \( n = 10 \) and \( \lambda = 2 \), what is the exact probability that \( y = 2 \)? (You needn't evaluate; just give a numerical expression.)

** ** ** **

\( F(x,y) \) is the joint cumulative distribution function of the random variables \( X \) and \( Y \), and \( F(x,y) \) is continuous. May we conclude

(a) \( X \) and \( Y \) possess a joint density function?

(b) \( P(X = 0) = 0? \)

(c) \( P(X = Y) = 0? \)

For each part, if you answer yes, sketch a proof; if no, give a counter-example.

** ** ** **

Prove: truncation reduces variance.
Mathematical Problems

* * * * *

If X and Y have a joint normal distribution how do you find the correlation between f[X] and g[X] where f and g are nice functions? There is a problem in Kendall's problem book giving higher moments for bivariate normals which might be useful in this pursuit. The presumption being that under some reasonable conditions these correlations will be about the same as the originals. In Morris's "Varieties of Values" he suggests that this is not really to be expected.

* * * * *

Let \([X_i]\) be a sequence of constants defined on \((\Omega, A, P)\) and \(c, c_1, c_2, \ldots\) sequence of constants.

If \(C = \{w | \sum_{1}^{\infty} X_i(w) \text{ converges}\}\) and \(D_{m,n,N} = \{w | \max_{1 \leq j \leq K \leq N} \sum_{i=j}^{K} X_i(w) < \frac{1}{m}\}\),
prove that \(P(C) = \lim_{m \to \infty} \lim_{n \to \infty} \lim_{N \to \infty} P(D_{m,n,N})\).

* * * * *

Let \((X_i, Y_i)\) for \(i=1, \ldots, n\) be a sample from a bivariate normal distribution. Assume that all you are given are the order statistics for the X's and Y's. What information, if any, do you have about the population correlation?

* * * * *

A and B are playing a two person zero sum game with A attempting to maximize. No matter what strategy B chooses A will know about it. Under these circumstances when is it most favorable for A to also pay for spying to learn precisely which move B is going to use?

* * * * *

In the one-sample rank order problem compute exactly the probabilities of the rank orders when the observations are taken from \(F(x) = H^\Delta(x)\) where \(\Delta > 1\) and \(H(-x)H(x) = 1\).
Mathematical Problems

Let \([X_i]\) be a sequence of R.V.'s defined on \((\Omega, \mathcal{A}, P)\) and \(c, c_1, c_2, \ldots\) sequence of constants.

Prove or counter: If \(\{X_n\}\) are independent, \(X_n \xrightarrow{P} X\) implies \(P\{X=c\} = 1\).

Locate the false link in the following chain of reasoning and prove it is erroneous: \(X_n \xrightarrow{L} X \xrightarrow{L} X_n - X \xrightarrow{L} 0 \xrightarrow{P} X_n - X \xrightarrow{P} 0 \xrightarrow{P} X_n \xrightarrow{P} X\).

Does the C.L.T. hold for independent r.v.'s \(\{X_n\}\) defined by

1. \(P\{X_j = \pm 2^j\} = \frac{1}{2}\)
2. \(P\{X_j = \pm 2^j\} = 2^{-2(2^j+1)}, \quad P\{X_j = 0\} = 1 - 2^{-2^j}\)
3. \(P\{X_j = \pm 1\} = \frac{1}{2}(1 - 2^{-j}), \quad P\{X_j = 2^j\} = 2^{-(j+1)}\)
4. \(P\{X_j = \pm j\} = \frac{(j-1)}{2^j}, \quad P\{X_j = 0\} = \frac{1}{2} - \frac{(j-1)}{2^j}\).

Consider a random walk on the place starting from \(b\) with absorbing barriers at \(a\) and \(c\) where \(a < b < c\). Let \(y\) satisfy \(a \leq y \leq c\). Find the expected number of visits to \(y\) before absorption. Do in general, and for the special case of equi-probability.

If \(X\) has a binomial distribution compute the probability that \(X\) will take on an even value. When is this probability greater than one half? Do the same thing for a Poisson distribution. Consider the probability that \(X=0\). [mod 4].

If \(X_1, \ldots, X_n\) is a sample from a uniform distribution with unknown range and location, find the maximum likelihood estimates, and the sufficient statistics. Are the maximum likelihood estimates unbiased, consistent, sufficient, or normal?
Mathematical Problems

* * * * *

Characterize the parameter sequences for which the C. L. T. holds if 

(i) \( X_n \) is uniformly distributed in \((-b_n, b_n)\)

(ii) \( X_n \) has a gamma distribution 
\[ f_n(x) = \frac{1}{\Gamma(\lambda_n)} x^{\lambda_n-1} e^{-x}, \quad x > 0 \]

(iii) \( X_n \) has a Poisson distribution with parameter \( \lambda_n \).

Failing this, state necessary and the best sufficient conditions you can find.

* * * * *

From a statistical standpoint what are the losses in information that arise from coding data? If you wish to estimate the mean of a normal distribution, how bad is it to group the data before estimating? This type of problem arises in coding for IBM equipment. Is there a general theory or should each case be treated separately?

* * * * *

Let \([X_i]\) be a sequence of R.V.'s defined on \((\Omega, \mathcal{A}, P)\).

Demonstrate that if \( X_n \xrightarrow{P} X \), a NSC that \( X_n \xrightarrow{r-m.} X \) is the uniform integrability of the sequence of integrals \( \{E|X_n - X|^r\} \). Also, if \( X_n \xrightarrow{r-m.} X \) (non-constant) \( r \geq 2 \), \( \rho(X_n, X) \rightarrow 1 \) where \( \rho \) is the correlation coefficient.

* * * * *

If \( X = (X_1, X_2, \ldots, X_k) \) and \( x \) are \( k \)-dimensional vector R.V.'s such that the distribution of \( X_n \) converged to that of \( x \) (at every continuity point of the latter), prove that the (univariate) c.d.f. of \( g(X_1^{(n)}, \ldots, X_k^{(n)}) \) converges to that of \( g(X_1, \ldots, X_k) \) at every continuity point of the latter if \( g \) is continuous.

* * * * *

An urn contains \( N \) balls, labelled 1,2,\ldots,\( N \). \( N \) is not known. \( n(n \leq N) \) balls are drawn without replacement. Determine a minimum variance unbiased estimator of \( N \).
Mathematical Problems

***

If \(a, b, c\) are uniform \((0,1)\) independent random variables, what is the probability that \(ax^2 + bx + c\) has real roots? What can you say about the distribution of the roots?

***

Research Problem

Consider a finite Markoff chain with states \(1, \ldots, N\) and let \(a_t\) be the state of the chain at time \(t\). Instead of observing \(a_t\) we observe \(x_t\) a random variable whose distribution depends on \(a_t\). From observations on the \(X\) process compute the a posteriori probability of being in state \(k\) at time \(t\) given you know all of the parameters of the chains and the conditional distributions. Also consider the problem of estimating some of these things. See Girschick and Rubin on Bayes quality control.

***

Prove

\[
\sum_{i=1}^{n} M_{i,n} \left( \frac{i-1}{d-1} \right) \left( \frac{n-i}{s-d} \right) < M_{d,s} \left( \frac{n}{s} \right)
\]

where \(1 \leq i, s \leq n, 1 \leq d \leq s\) and \(M_{a,b}\) is the expected value of the \(a^{th}\) order statistic in a sample of size \(b\).

***

Let \([X_i]\) be a sequence of R.V.'s defined on \((\Omega, A, P)\).

Prove if

\[
\sum_{n=1}^{\infty} P\left[ \left| X_{n+1} - X_n \right| > \epsilon_n \right] < \infty \quad \text{where} \quad \epsilon_n > 0 \quad \text{and} \quad \sum_{n=1}^{\infty} \epsilon_n < \infty
\]

that \(X_n \xrightarrow{a.s.} X\).

***

If \(X\) and \(Y\) have a joint bivariate normal distribution, when are \(S = X+Y\) and \(T = X-Y\) independently distributed?
Mathematical Problems

** * * * * *

For real $\lambda$, let $F(x;\lambda)$ denote a family of c.d.f.'s (in the variable $x$) none of whose $\lambda$ discontinuities is also a discontinuity of the c.d.f. $G(\lambda)$.

Prove that $H(x) = \int_{-\infty}^{\infty} F(x;\lambda) dG(\lambda)$ is a c.d.f.

** * * * * *

Let $[X_i]$ be a sequence of R.V.'s defined on $(\Omega,A,P)$ and $(c,c_1,c_2)$ sequence of constants.

Prove or disprove: If $X_n \xrightarrow{p} X$, and $\{X_n\}$ are uniformly bounded R.V.'s, i.e., $P(|X_n| \leq c) = 1$, $X_n \xrightarrow{r.m.} X$, $r > 0$.

** * * * * *

What is probit analysis?

** * * * * *

Prove that if $\{X_n\}$ are R.V.'s such that (i) $\sigma^2 < C$ and (ii) $cov(X_n,X_m) < 0$, all $n,m = 1,2,\ldots$ or (ii') $cov(X_n,X_m) \to 0$ uniformly as $|n-m| \to \infty$, then $\{X_n\}$ obeys the WLLN.

** * * * * *

If $\phi(t)$ is a c.f. show that $e^{\lambda[\phi(t)-1]}$, where $\lambda > 0$ is a c.f.. Prove that $R(\phi(t))$ is a c.f.. Is $\exp\{\sin t\}$ a c.f.? Validate your reply.

** * * * * *

Let $[X_i]$ be a sequence of R.V.'s defined on $(\Omega,A,P)$.

Show that $X_n^{(i)} \xrightarrow{p} X^{(i)}$, $i=1,2,\ldots,K$, $g(X_1,X_2,\ldots,X_K)$ continuous imply $g(X_n^{(1)},\ldots,X_n^{(K)}) \xrightarrow{p} g(X^{(1)},\ldots,X^{(K)})$. Is it true for $K-1$?

** * * * * *

What are some of the major contributions to statistics and probability of Laplace, K. Pearson, R. A. Fisher, Todhunter, A. Wald, and J. Neyman?

-24-
Mathematical Problems

** * * * *

If \( f, F \) are standard normal density and c.d.f. respectively and \( c \) is any constant show that

\[
\int_{-\infty}^{\infty} F(x+c) f(x) dx = F\left(\frac{c}{\sqrt{2}}\right).
\]

Hint: Use a probability argument.

** * * * *

Let \( S_n \) be a monotone sequence of R.V.'s. Is it true that

\[
S_n \xrightarrow{a.e.} S \iff S_n \xrightarrow{L} S?
\]

** * * * *

Prove the (Raikov-Ottaviani) theorem that if \( F_1 F_2 = F(x) = \sum_{j < x} \frac{\lambda^j e^{-\lambda}}{j!} \) then \( F_1 \) and \( F_2 \) are Poisson c.d.f.'s. Hint: Use prob. generating functions rather than c.d.f.'s.

** * * * *

Assume \( X \) is normal with mean \( \theta \) and variance 1. If we know a priori that \( \theta \) is less than one in absolute value, what is the maximum likelihood estimate of \( \theta \)?

** * * * *

Construct a sequence of R.V.'s

(i) Obeying the WLLN but not the SLLN.

(ii) Obeying the WLLN but having infinite expectations.

** * * * *

If \( X_1, \ldots, X_n \) is a sample from the uniform \((0,1)\) distribution find the distribution of

\[
P = \prod_{i=1}^{n} X_i.
\]
Mathematical Problems

***

Let \([X_1], [Y_1]\) be sequences of R.V.'s defined on \((\Omega, \mathcal{A}, P)\).

Prove that \(X_n \overset{L}{\to} X\) and \(X_n - Y_n \overset{P}{\to} 0\) imply \(Y_n \overset{L}{\to} X\).

***

Prove that if \(F\) is an arbitrary c.d.f., \(F(x) = F_d(x) + F_c(x)\) where \(F_d\) is a step function, \(F_c\) is continuous and both are "distribution functions" (though not generally c.d.f.'s). Hint: Define \(F_d\) via the discontinuities of \(F\), then take \(F_c = F - F_d\) and show \(F_d, F_c\) to have the asserted properties.

***

Demonstrate that if \(\varphi(t)\) is a c.f., \(\varphi'(0)\) finite does not preclude
\[
\int_{-\infty}^{\infty} |X| dF(x) = \infty. \quad \text{Hint: take } \varphi(t) = c \sum_{j=2}^{\infty} \frac{\cos jt}{j \log j}. \quad \text{Find the corresponding c.d.f.}.
\]

***

In the interpretation of the meaning of a simple correlation coefficient, do you prefer to think in terms of \(r^2\) rather than \(r\)? Why?

***

Is the Helly-Bray Theorem valid for multivariate c.d.f.'s, i.e., if \(g(x_1, \ldots, x_k)\) is continuous and bounded and \(F_n \to F\), a c.d.f., does
\[
\lim_{n \to \infty} \int_{\mathbb{R}^k} g(x_1, \ldots, x_k) dF_n(x_1, \ldots, x_k) = \int_{\mathbb{R}^k} g(x_1, \ldots, x_k) dF(x_1, \ldots, x_k)\]

***

Discuss the practical value in application of good experimental designs. Give an example of how error control is obtained by replication or stratification.

***

If \(N\) has \(k\) significant figures, how many significant figures are in \(N^\frac{1}{k}\)?

-26-
Mathematical Problems

Let $X_1, X_2, X_3, \ldots$ be a sequence of R.V.'s such that $F_{X_n}(x) \to F_X(x)$ at every continuity point of the latter. Prove if $g(x)$ is a continuous function that $F_{g(X_n)}(x) \to F_{g(X)}(x)$ at every continuity point of the latter.

In random sampling from a bivariate normal distribution, how could you make up an unbiased estimator of the population product moment correlation? How many bivariate observations do you need?

If $F_m(x)$ and $G_n(x)$ are sample c.d.f.'s find the relationship between
\[ T = \int_{-\infty}^{\infty} [F_m(x) - G_n(x)]d[F_m(x) + G_n(x)] \]
and the Wilcoxon statistic.

Is there anything like the Tukey-Scheffe multiple confidence intervals for the chi-square goodness of fit tests? Thus if we reject the null hypothesis, can we make legitimate statements about where the trouble is?

Let $\{X_n\}$ be independent with $P(X_n = +cn^\alpha) = \frac{1}{2}n^{-\beta}$, $P(X_n = 0) = 1 - n^{-\beta}$ where $c \neq 0$, $\beta \geq 0$. Characterize the region in the $(\alpha, \beta)$ half-plane where the classical (generalized) WLLN holds; where does the classical SLLN hold?

Let $X_1, \ldots, X_n, Y_1, \ldots$ be mutually independent random variables, each of which is $N(0,1)$. Obtain the density function of $R = (\Sigma X_i Y_i)(\Sigma X_i^2 Y_i^2)^{-\frac{1}{2}}$.

Prove that a bounded R.V. cannot be infinitely divisible.

-27-
Mathematical Problems

Let $X_n$ be independent with $P(X_n = \pm 2^n) = \frac{1}{2}$. Prove that the distributions of the partial sums $S_n$ converge to a uniform distribution. Interpret, give an appropriate prob. space and prove that if $\sum a_n^2 < \infty$, $P(\sum_1^n \pm a_n)$ converges = 1.

What is the relationship between the Wilcoxin and Mann-Whitney statistics?

Let $X_j$ be independent, uniformly distributed R.V.’s in $(-\frac{1}{j}, \frac{1}{j})$, $j=1,2,\ldots$. Show that $F_n(x) = P(\sum_1^n X_j < x) \rightarrow F(x)$, a c.d.f. (F need not be specified)

Prove the Theorem (of Pólya) that if a sequence of c.d.f.’s converges to a continuous c.d.f. that the convergence is uniform in $(-\infty, \infty)$. Can a sequence of continuous c.d.f.’s converge to a discontinuous c.d.f.?

If $P(X < 0) = 0$ then $EX \geq E[X|X \leq A]$. Can this be improved?

Demonstrate that the least-squares estimators of the coefficients in $y = \alpha + \beta x$ also are maximum likelihood estimators. Be certain to state explicitly all assumptions made.

Assume $y(t) = x_1(t) + x_2(t)$ where $x_1(.)$ and $x_2(.)$ are Poisson processes with parameters $A$ and $B$ respectively. Find $P(X_1(t) = k|y(T))$ for $0 \leq T \leq t$.

Then formulate and solve the analogous problem for Weiner processes.

Clearly distinguish between the statistical models for correlation and regression.
Mathematical Problems

*****

(a) Prove that \( \lim_{h \to \infty} \frac{1}{h} \int_{-h}^{h} F(u) \, du = F(\infty) \) and \( \lim_{h \to \infty} \frac{1}{b} \int_{-b}^{b} F(u) \, du = F(-\infty) \).

State a closely related proposition for sequences of numbers.

(b) If \( \beta_k = E|X|^k < \infty \), log \( \varphi(t) = \sum_{j=1}^{k} \frac{\gamma_j}{j!} (it)^j + o(t^k) \).

*****

Prove that a Poisson random variable with large parameter has a distribution which can be approximated (well) by a normal distribution.

*****

Show that if \( \sum_{i=1}^{n} a_i = 1, a_i \geq 0 \) that \( \sum_{i=1}^{n} a_i \varphi_i(t) \) is a c.f. if \( \varphi_i(t) \) is a c.f., \( i=1,2,\ldots,n \). Under what conditions, if any, is this true for \( n = \infty \)?

*****

Find a distribution such that \( E|X|^k \) fails to exist if \( k \geq 0 \).

*****

Show that \( F \), the family of all c.d.f.'s is a semi-group under the convolution operation with identity element

\[
I(x) = \begin{cases} 
1, & x > 0 \\
0, & x \leq 0. 
\end{cases}
\]

Demonstrate that \( F \) is not a group.

*****

Show that a NSC that \( F_n(x_1, \ldots, x_k) \) converge to a c.d.f. \( F(x_1, \ldots, x_k) \) is that for the corresponding c.f.'s \( \lim_{n} \varphi_n(t_1, \ldots, t_k) = g(t_1, \ldots, t_k) \) for any fixed vector \( t = (t_1, \ldots, t_k) \) with \( g(t_1, \ldots, t_k) \) continuous as the origin \( t = 0 \).

*****

If \( EX = EY = 0 \) and \( f(x) = EY|X \) is non decreasing, then \( EXY \geq 0 \).
Mathematical Problems

* * * * *

From the definition of stable laws prove, without invoking the class L or resorting to the form of a stable c.f., that the Poisson distribution cannot be stable. Under the same proviso prove that if \( X \) is stable and has a finite 2nd moment that \( X \) is Gaussian. Construct a distribution belonging to the domain of attraction of no proper stable law.

* * * * *

What is the nature of the controversy between the 'Bayesians' and 'non-Bayesians'?

* * * * *

If \( X = \sum_{i=1}^{N} X_i \), then \( \sigma_X \leq \sum_{i=1}^{N} \sigma_{X_i} \).

* * * * *

Formulate two different models justifying the use of least squares in estimating regression coefficients.

* * * * *

Under what normalization and for what positive values of \( \alpha \) (if any) does the C.L.T. hold for a sequence of independent identically distributed random variables with c.d.f.

\[
F_\alpha(x) = \begin{cases} 
\frac{1}{2} & \text{for } |x| \leq 1, \\
\frac{1}{2|x|^{\alpha}} & \text{for } x < -1, \\
1 - \frac{1}{2|x|^{\alpha}} & \text{for } x > 1.
\end{cases}
\]

* * * * *

Construct an example in which you can find a strategy which you can show to be admissible.

* * * * *

Let \( X_{n,i} \) be \( i^{th} \) order statistic in sample of \( n \). Then

\[
E(X_{n,j}|(X_{n,i}=a)) \geq EX_{n-i,j-i} \text{ for } j > i.
\]
Mathematical Problems

***

Let X be Poisson with parameter \( \lambda \). If \( X = \sigma \) let the conditional number observed be binomial \((p, \sigma)\). What is the distribution of the number observed?

***

Truncation from the left increases the value of the mean [provided it exists].

***

\( X \) is asymptotically normal \((a_n, b_n)\). Investigate the asymptotic normality of \( X^2 \).

***

Use contour integration to find the c.f. of \( X^2 \) where \( X \) is normally distributed with \( EX=0, EX^2=1 \).

***

If \( \{X_n\} \) are independent and uniformly bounded random variables, prove that the C.L.T. (if \( s_n \to \infty \)) and the SLLN hold.

***

Give examples of distributions which might occur in practice which have the property that they are neither purely discrete nor purely continuous.

***

If \( X \) is a random variable and \( P[Y < 0]=0 \) then \( P[X \leq t] \geq P[X+Y \leq t] \).

***

If \( X \) and \( Y \) are independent Poisson random variables, what can be said about the conditional distribution of \( X \) given \( X+Y \)?

***

Prove that the Levy distance \( \rho[F, G] \leq \rho_1[F, G] = \sup_x |F(x) - G(x)| \).

Is \( \rho_1[F_n, F] \to 0 \) a NSC that \( F_n \to F \)? Prove your assertion.
Mathematical Problems

* * * * *

Show that $F_n(x)/F(x)$ is bounded with probability one. $F_n(x)$ is the sample c.d.f.

* * * * *
Problems from References

* * * * *

Read the review of the article by Cibisov (Math. Rev., No. A1776, page 327, Vol. 24, Number 3A, September 1962.). Explain the difference between the author and the reviewer. Defend one of the positions or else form a new position.

* * * * *

If X and Y are independent Poisson random variables what is the conditional distribution of X given X+Y? 


* * * * *

Compare the results of


Apply the techniques of both papers to several examples.

* * * * *


* * * * *

Explain the uses of the results given in abstract 1, page 767, AMS, Vol. 26, (1955) by Walsh.

* * * * *

In R.A. Fisher's Statistical Methods and Scientific Inference, interpret the paragraph on page 14 beginning "Bayes' definition is ...". On page 47 what value of \( \theta \) corresponds to the null hypothesis and derive equations 21 on page 48.

* * * * *

State precisely what are the various methods for obtaining confidence intervals for proportions described by Crow, Biometrika, 1956, 423-435. In practice does it make much difference what you do? Can you supply some other reasonable procedures than those described?
Problems from References


Give an elementary proof of the main result in the abstract of Shapiro on page 364 of the Bull. Amer. Math. Soc. (1956)


Prove the result of the second paragraph of section 5 on page 162 of the Brown and Mood article in the Second Berkeley Symposium. If you have trouble, see Hodges, page 261 of Biometrika, 1955.

What affect does truncation, i.e., elimination of part of the sample space have on such properties as sufficiency, completeness, admissibility, and unbiasedness? Give examples and theories for the various concepts as well as a formal definition of "truncation". You might look at AMS 20 (1949) pp. 309-311, article by Tukey.

Describe the term you would use from the NBS Tables of the Binomial Distribution to approximate the Poisson term, p (3; 4.9).
Problems from References

* * * * *

If a distribution has finite variance then the median and mean can differ by at most square root of two times the standard deviation. When you can prove the above result look in the AMS 3, page 141 (1932), and Bull. Amer. Math. Soc., September 1955, abstract 629.

* * * * *

See AMS 26, page 540 (1955), abstract 14, and derive some of the results.

* * * * *

Show that the variance covariance matrix of the order statistics from the normal distribution (with variance one) is doubly stochastic.

See AMS 17, page 71 (1946).
AMS 19, page 270 (1948).
AMS 26, page 395 (1955), (Lemma 4.1.1).

* * * * *

On page 451 of Cramer's Mathematical Methods of Statistics verify the alternative definition of \( \omega^2 \) and the formulas for the moments of \( \omega^2 \).

* * * * *

Read Chapter 3 of Language as Choice and Chance by G. Herdan. Explain the fundamental probability set-up made by the author. Do you think the paradox is a paradox? Does \( v_m \) depend on the sample size very much? Is it clear that \( v_m \) will converge?

* * * * *

In the journal Indagationes Mathematicae, page 601 (1954) derive the equations between [13] and [14]. The article is by Stoker. You might go ahead and try some of the other equations.
Problems from References

* * * * *

Discuss examples and supply the details of the article on operations research in *Industrial Quality Control*, October 1955 (Lothrop).

* * * * *

Examine the article by Saucer and Deabler in the *Journal of Consulting Psychology* 20, (1956) pp. 385-389. Do the authors ask a reasonable question? Do they answer their own question? Can you answer their question? What happened to the missing observations? Is the measurement scale too coarse? Should there be some sort of interpolation between the frequencies before and after the threshold?

* * * * *

Verify the footnote on page 4 of *Industrial Quality Control*, volume XI, No. 5, February 1955.

* * * * *

Monica Lawlor "Cultural Influence on Preference for Designs" *J. of Ab. and Soc. Psychology* 51 (1955) pp. 690-692. Is the chi-square analysis in this article correct? In any case does this seem like a particularly good mode for making an analysis of this data?

* * * * *

Facing page 44 in the June 1956 *Industrial Quality Control*, what is the meaning of ± as used in the advertisement?

* * * * *

Discuss the model and verify some of the results contained in the abstract by Tweedie on page 525 of the *JASA* (1956).

* * * * *

Derive (5.8), (5.9), (5.10), and (5.11) on pages 595-596 of *AMS* 27 (1956).
Problems from References

Read the article by K. Purcell in the Journal of Consulting Psychology 20 (1956), pages 361-364. What is the meaning of reliability as used here? What model does the author have in mind and what would you expect to be a reasonable model for the problem? What holds the two parts of this study together? Was the analysis as performed done correctly? What form would your analysis of the data take?

Derive the results of the abstract by Edgett on page 508 of the JASA (1956).


a. Explain the model behind the analysis of variance.

b. Discuss confidence intervals of page 276.

c. How would you handle the data?


Prove the results stated in the following review of the Bhattacharyya article in Math. Reviews, Vol. 16 (1955), page 727:

Let \( x \) be the number of successes in \( n \) binomial trials, \( p \) being the success probability. If \( \tau(p) \) is a polynomial in \( p \) of degree \( \leq n \), it is shown how to construct a function \( t(x) \) such that \( E(t) = \tau \); \( t \) is a polynomial in \( x \) of degree \( \leq n \) and is a minimum-variance estimator, since any function \( t(x) \) is a minimum variance estimator of its expectation in binomial estimation. If \( \tau \) is not of the above form, no minimum variance estimator exists; a number of estimates are given which minimize the bias according to some criterion.
Problems from References

** * * * *

Are the results of the abstract by Jaspen in JASA (1956) on page 513 correct? Are they useful?

** * * * *

In the American Scientist, January 1957, consider the last paragraph begun on page 21 regarding "super nova". How would you justify the probability conclusions made by the author? Is there other astronomical data and theory required and/or available?

** * * * *

Verify the statement in the review of the article by Finkel'stein in Math Reviews (1957) page 1217.

** * * * *

Prove 1-6 on pages 21-22 and 1-9 on pages 55-58 of Shannon and Weaver, on information theory.

** * * * *


** * * * *

Explain the meaning of, and give the proof of, the theorem by Huzurbazar in Math. Rev. (1956) page 120.

** * * * *

In Sankhya 17 (1956-1957) page 217, is the last statement of the first paragraph exact or a limiting result? In the same paper can you do anything with the last paragraph on page 219?

** * * * *

Page 166 of AMS, (1951), first line of proof of Theorem 3.1.
Problems from References

* * * * *

Compare (verbally and numerically) the paper by Chu, pp. 780-789 in AMS (1956) and Chung, pp. 447-465 AMS (1946).

* * * * *


* * * * *

Derive results of Section 2 in the paper by Rosenblatt, AMS (1956), pp. 832-837.

* * * * *

Fraser's Non Parametric Methods, page 293. Ex 4(a). Note: $\xi$ should be replaced by $\xi \sigma$ where $\sigma$ is the standard deviation of $Z_i$.

* * * * *

In Proc. Camb. Phil. Soc. 51 (1955) page 761, is equation 3 correct? Explain your answer?

* * * * *


If $X_1, X_2, X_3, X_4$ are i.i.d. with the density $f(x)$, compute the probability that

$$X_1 \leq \frac{(w-1)X_2 + X_3}{w}, \quad X_4 \leq (1-w)X_3 + wX_2$$

for $0 \leq w \leq 1$. Are the probabilities distribution free?

* * * * *
Applied Problems

* * * * *

Assuming the usual assumptions underlying the analysis of variance are met, consider the following:

Genotypically identical grass-plants were inoculated with four spore concentrations of a fungus. The spores were allowed to germinate in five different dew chambers. Duplicate pots of each concentration appeared in each dew chamber. The data represent counts of lesions appearing on the plants.

<table>
<thead>
<tr>
<th>Dew Chamber</th>
<th>50,000</th>
<th>100,000</th>
<th>150,000</th>
<th>200,000</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>106</td>
<td>154</td>
<td>185</td>
<td>91</td>
<td>923</td>
</tr>
<tr>
<td>B</td>
<td>96</td>
<td>57</td>
<td>109</td>
<td>125</td>
<td>984</td>
</tr>
<tr>
<td>C</td>
<td>135</td>
<td>126</td>
<td>173</td>
<td>122</td>
<td>867</td>
</tr>
<tr>
<td>D</td>
<td>98</td>
<td>96</td>
<td>85</td>
<td>149</td>
<td>91</td>
</tr>
<tr>
<td>E</td>
<td>68</td>
<td>109</td>
<td>111</td>
<td>44</td>
<td>804</td>
</tr>
<tr>
<td></td>
<td>92</td>
<td>91</td>
<td>144</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>89</td>
<td>69</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td></td>
<td>130</td>
<td>83</td>
<td>128</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td></td>
<td>123</td>
<td>100</td>
<td>118</td>
<td>122</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1079</td>
<td>1005</td>
<td>1331</td>
<td>1088</td>
<td>4503</td>
</tr>
</tbody>
</table>

\[(106)^2 + (96)^2 + \ldots + (121)^2 = 548,713\]
\[\{(202)^2 + (233)^2 + \ldots + (243)^2\}/2 = 523,708\]
\[\{(1079)^2 + (1005)^2 + \ldots + (1088)^2\}/10 = 512,957\]
\[\{(923)^2 + (984)^2 + \ldots + (925)^2\}/8 = 509,239\]
\[(4503)^2/40 = 506,925\]

a. Complete the analysis of variance, including tests of significance.

b. Partition the spore concentration sum of squares into orthogonal linear, quadratic, etc. components and test for significance.

* * * * *

Given that \(k\) has the distribution, \(p(k;100)\), give numerical values for the following: \(E(k)\), standard deviation of \(k\), third moment about the mean, fourth moment about the mean.
Applied Problems

** * * * * **

4 biochemistry laboratories were selected within the north central region. Within each lab a analysts were appointed to analyze the oil content of soybeans of m different methods. d duplicate determinations were made for each method. Assume all anova assumptions are met. Write in the appropriate d.f. in the blanks provided. Then determine the M.S. expectations under three conditions:

(a) all effects random (b) all effects fixed and (c) labs and methods fixed--all other effects random. State the appropriate F tests under all three models.

<table>
<thead>
<tr>
<th>S.F.</th>
<th>d.f.</th>
<th>M.S</th>
<th>M.S. Expectation</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laboratories</td>
<td></td>
<td>M₆</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysts in labs</td>
<td></td>
<td>M₅</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td></td>
<td>M₄</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M X L</td>
<td></td>
<td>M₃</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M X A in L</td>
<td></td>
<td>M₂</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duplicate Determinations</td>
<td></td>
<td>M₁</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* * * * * *

Explain the following quotation which the chairman of the board of United States Steel made in a speech of 15 January 1957, "...Fifteen years ago only 5 percent of the family units in America had incomes of $5,000 or more. Today 44 percent of an increased number of family units are in this bracket...A few years ago we took a survey among our stockholders which yielded some surprising information. It showed that more than half of our shareowners had incomes that were less than the average wages we were then paying to our steelworkers. Nearly three-fifths of these stockholders had incomes of less than $5,000 a year; and about a third of them were in the brackets below the $3,000-a-year level."
Applied Problems

***

An experiment was made to yield one observation on each of the random variables $y_{ijk}$, $i=1,2,\ldots,12$, the assumption being that $y_{ijk} = u + a_i + b_i + c_{ijk}$, where $u$ is a constant and the $a_i$, $b_i$'s and $c_i$'s are independent, normally distributed random variables with zero means and respective variances $\sigma_a^2$, $\sigma_b^2$, $\sigma_c^2$. Obtain a 95 percent confidence interval for $\sigma_b^2$ from the following table:

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>mean square</th>
</tr>
</thead>
<tbody>
<tr>
<td>a classes</td>
<td>11</td>
<td>3.5629</td>
</tr>
<tr>
<td>b classes</td>
<td>48</td>
<td>1.2055</td>
</tr>
<tr>
<td>c classes</td>
<td>120</td>
<td>.6113</td>
</tr>
</tbody>
</table>

***

Suppose it is hypothesized ($H_0$) that $k \sim p(k;2)$. Suppose the alternative ($H_1$) is that $k$ has a binomial distribution with $n=10$ and $\pi=.20$. Let $k_1, k_2, \ldots, k_{100}$ be 100 independent observations on $k$. If these observations are classified into 6 classes, $k=0$, $k=1$, $k=2$, $k=3$, $k=4$, and $k=5$, and a Pearson $\chi^2$ Goodness-of-Fit Test is done, what is the probability of accepting $H_1$ if it is true?

***

A factory has 10,000 light bulbs of a given size used to light the plant. Present an argument for expecting that the number of bulbs burning out per day will resemble a Poisson distribution, when recorded over a long period of time.

***

Suppose that larvae of a particular species are randomly distributed over a field with density per square foot denoted by $\delta$. If $\delta \geq 5$, how many plots of size one square foot should be randomly selected in order to obtain an estimate of $\delta$, based on the number of larvae counted in the plots, with coefficient of variation no larger than .30?
Applied Problems

Let $\lambda_a$ and $\lambda_b$ denote the rate of counted emissions/min. for two separate radioactive materials. Suppose each material is counted until n counts are accumulated, and the waiting times, $t_{na}$ and $t_{nb}$, are recorded. How many events (n) should be counted to yield an interval estimate of $\lambda_a \lambda_b$ with endpoints differing by four-fold? (C.C. = .95)

Epileptic seizures occur in some patients in bursts or clusters in time. Suppose the clusters occur randomly for a specific patient at the rate of one a week, and that a cluster will usually consist of about ten seizures spread over two days. On the average how many seizures per day may be expected? What will be the variance of seizures per day? How many days would the patient have to be observed to give high assurance of rejecting the hypothesis of Poissoness, if the Index of Dispersion Test for Poissoness is to be used?

Question for following data is on the top of next page.

"Split-split-plot" diallel analysis when the $F_1$'s are blocked according to line (whole plots) and reciprocals (sub-plots). Selfs not included.

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>MEAN SQUARE EXPECTATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>$r-1$</td>
<td>$S_R, M_R$</td>
<td>1 0 0 4 0 8(n-1) 0 2n(n-1)</td>
<td></td>
</tr>
<tr>
<td>General combining ability (Gca)</td>
<td>$n-1$</td>
<td>$S_G, M_G$</td>
<td>1 0 0 4 4r 4(n-2) 4r(n-2)</td>
<td></td>
</tr>
<tr>
<td>$R \times Gca$</td>
<td>$(r-1)(n-1)$</td>
<td>$S_{RG}, M_{RG}$</td>
<td>1 0 0 4 0 4(n-2)</td>
<td></td>
</tr>
<tr>
<td>Specific combining ability (Sca)</td>
<td>$n(n-3)/2$</td>
<td>$S_s, M_s$</td>
<td>1 0 0 4 4r</td>
<td></td>
</tr>
<tr>
<td>$R \times Sca$</td>
<td>$(r-1)n(n-3)/2$</td>
<td>$S_{Rs}, M_{Rs}$</td>
<td>1 0 0 4</td>
<td></td>
</tr>
<tr>
<td>Duplicate sub-plots in replications</td>
<td>$rn(n-1)/2$</td>
<td>$S_d, M_d$</td>
<td>1 0 4</td>
<td></td>
</tr>
<tr>
<td>Reciprocal crosses</td>
<td>$n(n-1)/2$</td>
<td>$S_r, M_r$</td>
<td>1 4r</td>
<td></td>
</tr>
<tr>
<td>Duplicate reciprocal crosses</td>
<td>$(2r-1)n(n-1)/2$</td>
<td>$S_e, M_e$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>$2rn(n-1)-1$</td>
<td>$S_t$</td>
<td>-43-</td>
<td></td>
</tr>
</tbody>
</table>
Applied Problems

* * * * *

The diallel analysis can be used to estimate genetic parameters of a conceptual population from which a set of inbred lines were randomly selected. Two such parameters are additive genetic ($\sigma_g^2$) and dominance ($\sigma_d^2$) variances. These quantities are commonly used to estimate heritability, $H$, as follows;

$$ H = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_d^2 + \sigma_p^2} $$

The data on bottom of previous page gives the analysis of variance "key-out", including mean square expectations, for a particular type of a diallel mating system. For the fruit weight of 25 berries (in grams) the following mean squares were determined.

\[ M_r = 180.51, \quad M_g = 1741.28, \quad M_{rg} = 138.88, \quad M_s = 81.44, \quad M_{rs} = 32.51 \]
\[ M_d = 75.54, \quad M_r = 36.47, \quad M_e = 18.37. \]

From these data, estimate $H$ given that $r=2$, $n=9$, $\sigma_g^2=4\sigma_a^2$, $\sigma_d^2=4\sigma_y^2$, $\sigma_p^2=\epsilon^2+\sigma_a^2+\sigma_d^2+\sigma_y^2$. Also construct a mean square for testing the significance of $Qca$.

* * * * *