DEPARTMENT OF MATHEMATICAL STATISTICS

Qualifying Examination

Statistics

Friday November 16, 1979

1 o'clock

Do Any 4 Questions

Answer each question in a separate book

1. Recall Helmert's transformation for $X_1, \ldots X_n$ iid $N(\mu, \sigma^2)$ r. v. 's:

$$U_1 = \bar{X}_n$$
$$U_i = X_i - \bar{X}_i, \quad i = 2, \ldots n$$

where $\bar{X}_i = (X_1 + \ldots + X_i)/i$. The U_i are independent normal variates.

Let now $W_1 = [n/(n-1)] \cdot (X_n - \bar{X}_n)^2$, $Q = \sum_i (X_i - \bar{X}_n)^2$ and $Z = W_1/Q$.

a) Express the statistic Z in terms of the standardized U_i.

b) From that representation and elementary probability methods show that the conditional distribution of Z given Q is Beta $(1/2, (n-2)/2)$.

Hint: Consider first the independent variates W_1 and $W_2 = Q - W_1$. From that the joint density of (Z, Q) is easily derived.

$x = 1, 2, \cdots \in N$

2. Let X be an observation from P_θ population, where P_θ is defined by $P_\theta[X = i] = \frac{1}{\theta}$ for $i = 1, \ldots, \theta$.
$= 0$ otherwise.

$E_\theta g(x) = 0$ for all θ

$\theta = 1 \Rightarrow g(1) = 0$

$\theta = 2 \Rightarrow g(1) + g(2) = 0 \Rightarrow g(2)$

$\cdots$

$g(x) = 0$ and sure x is

(i) Show that $\mathcal{P} = \{P_\theta, \theta = 1, 2 \ldots\}$ is complete and find the U.M.V.U. (uniformly minimum variance unbiased) estimate of θ. $EX = \frac{1+\theta}{2}$

$T = 2X - 1$ is UMV

(ii) For any integer K, let $\mathcal{P}_K = \{P_\theta, \theta = 1, 2 \ldots, \theta \neq K\}$. Show that $\mathcal{P}_K$ is not complete and your estimate in (i) is not U.M.V.U. in this case.

the UMVU is $2X - 1 + I_{[X=K]} - I_{[X=K+1]}$

$E g(x) = 0 \Rightarrow g(K) + g(K+1) = 0$ is ok. $\to$ not complete

$g(i) = 0$ for $i = 1, 2, \ldots, K-1$. $g(K+2) = 0$ $g(K+3) = 0$ $g(K+\ell) = 0$, $\ell \geq 2$

3. An individual belongs to one of two populations Π_0 or Π_1 with prior probability p_θ that a randomly selected individual

belongs to Π_θ ($\theta = 0, 1$). The individuals in Π_θ have a distribution of scores X which is $N(\mu_\theta, 1)$ where μ_0 and μ_1 may be assumed known. The loss function $L(\theta, a)$ for deciding that the individual belongs to population Π_a when in fact he belongs to Π_θ is given by

设 分成 $\mu_0=\mu_1$, $\mu_0<\mu_1$, $\mu_0>\mu_1$ 5 讨论之.

$$L(\theta, a) = c_1(1-a)\theta + c_0 a(1-\theta) \qquad (a = 0, 1;\ \theta = 0, 1)$$

where c_0 and c_1 are known positive constants.
What is the Bayes decision rule $\delta(X)$ against the prior (p_0, p_1) for deciding population membership based on X?

4. Let $y_t = t\beta + \epsilon_t$, $t = 1, 2, \ldots n$, with ϵ_t i.i.d $N(0, \sigma^2)$.
Let $\delta_t = y_{t+1} - y_t$, $t = 1, \ldots, n-1$

a) Find the mean and variance of δ_t and of

$$\bar{\delta} = \frac{1}{n-1} \sum_{1}^{n-1} \delta_t.$$

$E\delta_t = Ey_{t+1} - Ey_t = \beta$
$Var\,\delta_t = Var(\epsilon_{t+1}) + Var(\epsilon_t) = 2\sigma^2$
$E\bar\delta = \beta$, $Var\,\bar\delta = Var(\frac{1}{n-1}(y_n - y_1)) = \frac{2\sigma^2}{(n-1)^2}$

b) Is $\underline{\delta} = (\delta 1, \ldots, \delta_{n-1})'$ sufficient for β? 不是

c) Give the best linear estimate you can for β based on $\underline{\delta}$, and compare it with the least squares estimate based on $y_1, \ldots y_t$.

5. Consider the regression model, $y_i = x_{i1}\beta_1 + \ldots + x_{ip}\beta_p + \epsilon_i$, $i = 1, 2..$ where ϵ_i are i.i.d. random variables with mean zero and variance $\sigma^2 > 0$. Let X_n denote the design matrix (x_{ij}) $1 \le i \le n$; $1 \le j \le p$. Assume that $(X_n' X_n)^{-1}$ exists for $n \ge p$. Let b_n be the least squares estimate of β_1. $\because rk(X_n) = p \Rightarrow \beta_1$ is estimable

(i) Show that $var(b_n)$ is a decreasing function of $n (n \ge p)$.

除非 $x_{n+1,j} \equiv 0$ $\forall j$ 否则

$b_n = c_1 y_1 \cdots + c_n y_n$
$b_{n+1} = c_1 y_1 + \cdots + c_n y_n + c_{n+1} y_{n+1}$

无偏线性类中最小者.
若取 $c_{n+1} = 0$ 则为 b_n, 可知

(ii) If $p = 2$, $E|\epsilon_i|^3 < \infty$, $\operatorname{var}(b_n) \to 0$ and all design points (x_{i1}, x_{i2}) are taken from a compact subset of $\{(x_1, x_2) | x_2 = 1\}$ then $(b_n - \beta_1)/\{\operatorname{var}(b_n)\}^{\frac{1}{2}} \xrightarrow{\mathcal{L}} N(0, 1)$.

$$\frac{\Sigma(x_i-\bar{x})\varepsilon_i}{\sqrt{\sigma^2\Sigma(x_i-\bar{x})^2}} = \frac{\Sigma(x_i-\bar{x})(y_i-Ey)}{\sqrt{\sigma^2\Sigma(x_i-\bar{x})^2}} \xrightarrow{?} N(0,1)$$

6. Let p_o, p_1, p_2 be unknown probabilities, with $\Sigma p_i = 1$, and let A_i be corresponding events, with $\cup A_i = S$, the certain event. We may interpret these as indiference, preference for treatment I, and preference for treatment II, resp., on a single trial of an experiment. Consider now n independent such trials with constant $(p_i, i = 0, 1, 2)$, and let X_i be the number of times A_i occurs.

a) Consider first the conditional statistical model that describes the assumed behaviour of (X_1, X_2) given that X_o equals a fixed given constant. Five the UMP level α test for $p_1 = p_2$ vs. $p_1 > p_2$ under such conditional model. Specifically: give the test statistic, the distribution of the test statistic, and the form of the critical function.

b) Consider now such conditional test as a test depending on X_o, X_1 and X_2, for $p_1 = p_2$ vs. $p_1 > p_2$, under the original unconditional model. Show that this test is UMP among all tests ϕ such that $E(\phi | X_o) \equiv \alpha$ under the null hypothesis (i.e., whenever $p_1 = p_2$).

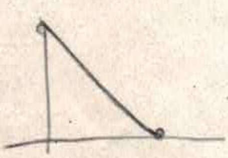
ϕ $E(\phi) \equiv \alpha$

a) given $x_0 = k$. $\Rightarrow$ $x_1 + x_2 = n-k$.

$$\boxed{P(X_1 = l \mid x_0 = k) = \binom{n-k}{l}\left(\frac{p_1}{p_1+p_2}\right)^l \left(\frac{p_2}{p_1+p_2}\right)^m}$$

分子 = 1项

$l + m = n - k$



在这个线上的分布

for such l, $\dfrac{P_{p_1>p_2}(X_1 = l | X_0 = k)}{P_{p_1=p_2}(X_1 = l | X_0 = k)} > \lambda$ reject H_0

$\Rightarrow l > k$, reject.

size. $P_{p_1=p_2}(X_1 = l | X_0 = k) = \binom{n-k}{l}\left(\frac{1}{2}\right)^{n-k}$ for $X_0 = 0, 1, 2,$

b) $E(\phi | X_0) \equiv \alpha$ $E(\phi | X_0) >$ [illegible] $\tilde{\phi}$ s.t. $E(\tilde{\phi} | X_0) \equiv \alpha$