Statistics Qualifying Examination, November 4, 1977

Instructions: Do any four problems. Use a separate book for each answer.

1. Let $\underset{\sim}{p}=(p_1,\ldots,p_k)$ and $\underset{\sim}{q}=(q_1,\ldots,q_k)$ be probability vectors. Let $\underset{\sim}{m}=(m_1,\ldots,m_k)$ be multinomial with parameters M and $\underset{\sim}{p}$, and let $\underset{\sim}{n}=(n_1,\ldots,n_k)$ be independent of $\underset{\sim}{m}$ with the multinomial distribution with parameters N and $\underset{\sim}{q}$, so that $\sum_{i=1}^{k} m_i=M$ and $\sum_{i=1}^{k} n_i=N$. Give an unbiased estimate of $\sum_{i=1}^{k}(p_i-q_i)^2$ based on $\underset{\sim}{m}$ and $\underset{\sim}{n}$.

define $X_i^1 = \begin{cases} 1 & \text{ith trial turns out to be the first event} \\ 0 & \text{otherwise.} \end{cases}$ $\Rightarrow \sum_1^M X_i^1 = m_1$

$(\sum X_i^1)^2 - \sum X_i^{1^2} = \sum X_i^1 X_j^1$ take E we see $E\,\frac{m_1^2-m_1}{M(M-1)} = p_1^2$. similarly for q_1^2

$\Rightarrow (p_1-q_1)^2 = p_1^2+q_1^2-2q_1p_1$, $\frac{m_1^2-m_1}{M(M-1)} + \frac{n_1^2-n_1}{N(N-1)} - 2\frac{m_1}{M}\frac{n_1}{N}$ should be an unbiased est. of $(p_1-q_1)^2$

Sum up we get what we need.

2. Let $X_1,\ldots,X_n$ be independent random variables each having the Poisson distribution $P(X_i=x) = \frac{\lambda^x e^{-\lambda}}{x!}$, $x=0,1,2,\ldots$

a) Find a lower bound for the variance of an unbiased estimate of the parameter $\theta=\lambda^2$. $\geq \frac{4\lambda^3}{n}$

b) Find the minimum variance unbiased estimate of θ. $\frac{\sum X_i'}{n}$

c) Is the variance of the estimator in (b) equal to the Cramér-Rao lower bound? Explain your answer. Yes.

3. Smith and Jones play the following game repeatedly. Each one picks a number from {1,2,3}. If the numbers differ by 2, the game is replayed. If the numbers differ by 1, the player with the smaller number wins \$1.00 from the player with the larger number. If the numbers are equal, both players must

con't.)

pay the house \$2.00.

a) What is the minimax strategy for Smith?

b) Suppose that Smith knows that Jones randomizes his guesses so as to average 2. What, if any, adjustments should Smith make in his strategy to reduce his losses?

4. Consider two simple linear regression models

$$y_i=\alpha+\beta u_i+\varepsilon_i \qquad (i=1,\ldots,m)$$

$$z_j=\alpha'+\beta' v_i+\varepsilon'_j \qquad (j=1,\ldots,n) \; ,$$

where the errors $\varepsilon_1,\ldots,\varepsilon_m,\varepsilon'_1,\ldots,\varepsilon'_n$ are i.i.d. $N(0,\sigma^2)$ random variables.

a) How would you test the null hypothesis that the two regression lines are in fact parallel? $\beta=\beta'$

b) Suppose it is known that $\alpha=\alpha'$, and call the common value α^*. What are the least squares estimates of α^*, β, and β'? joint

5. Let $\{X_i\}$ be independent $N(\mu_i,1)$ random variables for $i=1,2$.

a) Characterize the most powerful test of size α of H_0: $\mu_i=0$ $(i=1,2)$ vs. H_1: $\mu_1=1,\mu_2=-2$.

b) Find the largest set of alternatives (μ_1,μ_2) for which the test in (a) is uniformly most powerful.

6. Let (X_i,Y_i) $i=1,\ldots,n$ be paired observations in which the $\{X_i\}$ are known constants, and the $\{Y_i\}$ are independent binary variables with $P[Y_i=1|X_i]=\theta_i(X_i)=1-P[Y_i=0|X_i]$.

Consider the linear logistic model

$$\log\left(\frac{\theta_i}{1-\theta_i}\right)=\alpha+\beta X_i$$

(6 con't)

a) Find the log-likelihood function $l(\alpha,\beta|X_1,\ldots,X_n)$. What are the sufficient statistics for α and β?

b) Let $\left\{\begin{array}{l} X_1=\ldots=X_5=-1 \\ X_6=\ldots=X_{10}=0 \\ X_{11}=\ldots X_{15}=+1 \end{array}\right\}$ and suppose the number of $Y_j=1$ is equal to $\begin{cases} 1 & \text{for } j=1,\ldots 5 \\ 0 & \text{for } j=6,\ldots,10 \\ 5 & \text{for } j=11,\ldots,15. \end{cases}$

Find a UMP one-sided test of $H_0:\beta=0$ vs $H_1:\beta>0$, and test the hypothesis at the $\alpha=0.2$ level.

7. Let (X,Y) have a bivariate normal distribution with zero means, unit variances, and correlation ρ. Let $P=P[X>0 \text{ and } Y>0]$.

a) ρ and P satisfy a relation of the form $\rho=f(P)$. Find the function f.

$\rho = \cos(2\pi P)$

b) In a sample of 144 pairs of observations $(X_i,Y_i,)$ $i=1,\ldots,144$, 24 pairs fell in the first quadrant of the X-Y plane. Give an approximate 95% confidence interval for ρ.

Binomial case

$P\left\{P(X>0\ Y>0) \in \left(\frac{24}{144}-\varepsilon,\ \frac{24}{144}+\delta\right)\right\} \approx .95$

find ε and δ.

Confidence Interval $\Rightarrow$ $\left(\rho_L = \cos\left[2\pi\left(\frac{24}{144}-\varepsilon\right)\right],\ \rho_U = \cos\left[2\pi\left(\frac{24}{144}+\delta\right)\right]\right)$

is a confidence Interval for ρ