Stat 8111 **Final Exam** December '17

Eight people took the exam and the scores ranged from 28 to 100. The average score was 49.25.

1. Let X be a real valued random variable where for each $\theta \in [0,1) = \Theta$ its probability function is given by

$$f_\theta(x) = \begin{cases} (1-\theta)\theta^x & \text{for} \quad x = 0, 1, \ldots \\ 0 & \text{otherwise} \end{cases}$$

That is X is geometric(θ).

i) For testing H:$\theta \le .5$ against K:$\theta > .5$ find the UMP level .125 test.

ii) Find the best unbiased estimator of $\gamma(\theta) = \theta/(1-\theta)$ when the loss function is squared error and the space of possible decisions is $\mathcal{D} = [0, \infty)$.

iii) If θ is known to belong to $\Theta^* = [\theta_1, \theta_2]$ where $0 < \theta_1 < \theta_2 < 1$ is the estimator in ii) admissible? Explain.

iv) If the parameter space is Θ^* and the decision space is taken to be $\mathcal{D} = \gamma(\Theta^*)$ discuss the unbiased estimation of $\gamma(\theta)$ with squared error loss.

v) Let the parameter space be Θ and consider estimating $\gamma_1(\theta) = \theta$ with $\mathcal{D} = [0,1]$ as the decision space and

$$L(\theta, d) = (\theta - d)^2/(1-\theta)$$

as the loss function. Let g is a continuous probability density function on Θ. Find an expression for the Bayes estimator of θ against g in terms of the moments of g.

Ans i) Easy to check that we have the MLR property in X so a UMP test rejects when $X > c$. Easy to see that $P_{\theta=1/2}(X \ge 3) = 1/8$.

ii) $E_\theta(X) = \theta/(1-\theta)$

iii) The estimator defined just below dominates X

$$\begin{aligned} \delta(x) &= \theta_1/(1-\theta_1) \quad \text{when} \quad x < \theta_1/(1-\theta_1) \\ &= x \quad \text{when} \quad \theta_1/(1-\theta_1) \le x \le \theta_2/(1-\theta_2) \\ &= \theta_2/(1-\theta_2) \quad \text{when} \quad x > \theta_2/(1-\theta_2) \end{aligned}$$

iv) By completeness X is the only unbiased "estimator" but it is no longer an estimator.

v) Note that the solution to the problem

$$\min_d \int (\theta - d)^2 h(\theta)\, d\theta$$

when h is not a probability density function is

$$d^* = \frac{\int \theta(h(\theta)\, d\theta}{\int h(\theta)\, d\theta}$$

So

$$\delta_g(x) = \frac{\int \theta\theta^x g(\theta)\, d\theta}{\int \theta^x g(\theta)\, d\theta} = m_g(x+1)/m_g(x)$$

2. A master of ceremonies (MC) announces that an urn contains a white balls and a blue balls where $a \geq 1$ is unknown. She will select n balls at random from the urn using one of two sampling plans. The first is simple random sampling with replacement (srswp) and the second is Polya sampling where after each draw the ball selected and another of the same color are returned to the urn before the next draw. Find the form of a UMP level α test for testing

H : The MC is using srswp. Against K : The MC is using Polya sampling.

Ans Let A_r be the event that we see exactly r white balls in some specific order. Then

$$P_{rs}(A_r) = (1/2)^n$$

and for a fixed a

$$P_{ps}(A_r) = \frac{(\prod_{i=0}^{r-1}(a+i))(\prod_{i=0}^{n-r-1}(a+i))}{\prod_{i=0}^{n-1}(2a+i)}$$

Hence

$$\frac{P_{ps}(A_r)}{P_{rs}(A_{rs})} \propto \{\prod_{i=0}^{r-1}(a+i)\}\{\prod_{i=0}^{n-r-1}(a+i)\}$$

which as a function of r is symmetric about $n/2$ and is strictly increasing in both directions as we move away from $n/2$. So the Ump level α test will reject when $|r - n/2| > c$ for some constant c.

3. Consider a no data decision problem with $\Theta = D = \mathcal{R}$. For $i = 1$ and 2 let f_i be a density function on $\mathcal{R}$ with mean m_i and variance $0 < \sigma_i^2 < \infty$. Find the decision d^* which solves the problem

$$\inf_{d \in \mathcal{R}} \int (\theta - d)^2 f_1(\theta)\, d\theta \quad \text{subject to} \quad \int (\theta - d)^2 f_2(\theta)\, d\theta \leq \sigma_2^2 + c$$

where $c > 0$ is some fixed constant.

Ans Since $E_2(\theta - d)^2 = \sigma_2^2 + (m_2 - d)^2$ we can restate the problem as finding the d^* which satisfies

$$\inf_{d \in \mathcal{R}} \int (\theta - d)^2 f_1(\theta)\, d\theta \quad \text{subject to} \quad (m_2 - d)^2 \leq c$$

Clearly the solution is given by

$$\begin{aligned} d^* &= m_1 && \text{when} \quad m_2 - \sqrt{c} \leq m_1 \leq m_2 + \sqrt{c} \\ &= m_2 - \sqrt{c} && \text{when} \quad m_1 < m_2 - \sqrt{c} \\ &= m_2 + \sqrt{c} && \text{when} \quad m_1 > m_2 + \sqrt{c} \end{aligned}$$

4. For a given positive integer θ Let X have the discrete uniform distribution over the set $\{1, 2, \ldots, \theta\}$. Suppose $\theta \in \{1, 2, \ldots, N\} = \Theta$ where $N > 2$ is a known positive integer. Suppose we wish to estimate $\gamma(\theta) = \theta$ with squared error loss and the decision space is $\mathcal{D} = [1, N]$.

i) Consider the prior distribution $\pi(\theta) \propto \theta$. Find its Bayes estimate of θ.

ii) Show that X is an admissible estimator for this problem. For what other loss functions is X admissible.

Ans i) Note $\pi(\theta) = 2\theta/(N(N+1))$ and so

$$p(x) = \sum_{\theta=x}^{N} (1/\theta)\pi(\theta) = \frac{N - x + 1}{N(N+1)}$$

and then

$$\begin{aligned} p(\theta|x) &= \frac{2\theta}{N(N+1)} \frac{1}{\theta} / \frac{N - x + 1}{N(N+1)} \\ &= \frac{1}{N - x + 1} \quad \text{for} \quad \theta = x, x+1, \ldots, N \end{aligned}$$

and we see that the Bayes estimate is $(x + N)/2$.

ii) Note X is unique stepwise Bayes against the sequence of priors (λ) where λ^i puts mass 1 on $\theta = i$ Note that this will work for any loss function is strictly increasing away from zero.

5. Consider a finite population of size N where the parameter space is $\mathcal{Y} = [0, \infty)^N$. Let λ denote a prior distribution over $\mathcal{Y}$ where the y_i's are independent and $E(y_i) = a_i$ and $V(y_i) = \sigma^2$ for $i = 1, \ldots, N$. Let π denote a sampling design with the property that $\pi(s) > 0$ if and only if $n(s) > 0$ where $n(s)$ is the number of units in the sample s.

i) Find δ_λ the Bayes estimator of the population total, T_y, when the loss function is squared error.

ii) Find $r(\delta_\lambda, \pi, \lambda)$, the Bayes risk for δ_λ when λ is the sampling design.

Ans i) Because the Bayes estimate does not depend on λ we have by independence of the y_i's that the Bayes estimate is

$$\delta_\lambda(y(s)) = \sum_{i \in s} y_i + \sum_{j \notin s} a_j$$

ii)

$$\begin{aligned} r(\delta_\lambda, \pi, \lambda) &= \int_y \sum_s (\delta_\lambda(y(s)) - T_y)^2 \pi(s)\, dy \\ &= \int_y (\sum_{j \notin s} (a_i] - y_j)^2 \pi(s))\ dy \\ &= \sum_s (\int_y (\sum_{j \notin s} (a_j - y_j)^2\ dy) \pi(s) \\ &= (N - n)\sigma^2 \end{aligned}$$