Indian Statistical Institute

Bangalore Centre

B.Math Third Year 2017-2018

Second Semester

Semestral Examination Date : 9.5.18

Statistics IV

Answer as much as you can. The maximum you can score is 120.

The notation used have their usual meaning unless stated otherwise.

Time :- 3 hours

1. Consider a study with three categorical variables X, Y and Z in which X and Y takes two values while Z takes K values.

   (a) Define conditional odds ratio θ_k^{XY}, $1 \leq k \leq K$.

   (b) Show that each of the following is a sufficient condition for θ_k^{XY} being independent of k.

   (i) X is conditionally independent of Z, given Y.

   (ii) Y is conditionally independent of Z, given X. [3 + 4 x 2 =11]

2. Suppose $X = (X_1, \cdots X_k)'$ follows multinomial distribution with parameters $(n, \pi_1, \cdots \pi_k)$. Let

$$V = (V_1, \cdots V_k)', \; V_i = (X_i - n\pi_i)/\sqrt{n\pi_i}.$$

   (a) Consider a $t \times k$ matrix B, $t \leq k$. Find the asymptotic distribution of $B'V$, as $n \to \infty$.

   (b) Suppose each π_i is a function of $\theta_1, \cdots \theta_q$, $q < k$. Let $M = ((M_{ij}))$, $m_{ij} = (\sqrt{\pi_i})^{-1}\frac{\partial \pi_i}{\partial \theta_j}$.

   Show that the information matrix of the θ vector is $= nM'M$.

   [Recall that the (s,t)th element of the information matrix is

$$E[\frac{\partial log(f(x,\theta))}{\partial \theta_s}\frac{\partial log(f(x,\theta))}{\partial \theta_t}]$$

   [8 + 8 = 16]

3. Suppose $X_1, \cdots X_n$ is a random sample from a continuous distribution with median θ. Let R_i^a be the rank of $|X_i|$ and $T^+ = \sum_{X_i > 0} R_i^a$. Let $W_{ij} = (1/2)(X_i + X_j), 1 \leq i \leq j \leq n$.

   (a) Show that if $\theta = 0$, the distribution of T^+ is symmetric about $n(n+1)/4$.

   (b) Suppose $\theta = 0$. Let W^+ denote the number of positive W_{ij}'s. Show that $W^+ = T^+$, provided (i)$X_i \neq 0$ and (ii) $|X_i| \neq |X_j|, j \neq i, i, j = 1, \cdots n$.

   (c) Show how you can find a confidence interval for θ using W_{ij}'s.

   (d) Show that the median of the W_{ij}'s is a reasonable point estimator of θ. [4 + 7 + 8 + 8 = 27]

4. (a) When is a decision rule said to be admissible ?

(b) Define a minimal complete class. Show that if a minimal complete class exists, then it consists exactly of the admissible rules. [3 + (3 + 6) = 12]

5. Consider a subset S of a k-dimensional Euclidean space.

(a) When is S said to be bounded from below ? What is the lower boundary of S ? When is S said to be closed from below ?

(b) State a condition on S which implies that the lower boundary of S is not empty ; proof is not required. [(3 + 4 + 2) + 3 = 12]

6. (a) Define (i) a minimax decision rule and (ii) a Bayes' decision rule w.r.t. a prior distribution π.

(b) In the following statement fill in the blank with one of the given words. Justify your answer.

"If the loss function $L(\theta, a)$ is $(\theta - a)^2$, then an unbiased estimator is - the Bayes' rule."

(i) always, (ii) often, (iii) sometimes, (iv) rarely, (v) never.

[Hint : State and prove a result and provide an example supporting your answer. State clearly the properties that you use, without proof.] [(3 x 2) + (2 + 9 + 3) = 20]

7. Consider a decision problem. Let D denote the class of all randomized decision rules and Π the class of all prior distributions of $\theta \in \Theta$.

(a) Define a least favorable prior distribution for the problem.

(b) Show that for every decision rule $\delta \in D$ the following holds.

$$\sup_{\pi \in \Pi} r(\pi, \delta) = \sup_{\theta \in \Theta} R(\theta, \delta).$$

(c) Suppose Θ is finite and the risk set is bounded from below.

(i) Show that the following relation holds and there exists a least favorable distribution π_0.

$$\inf_{\delta \in D} \sup_{\pi \in \Pi} r(\pi, \delta) = \sup_{\pi \in \Pi} \inf_{\delta \in D} r(\pi, \delta) = V. \tag{1}$$

(ii) Suppose further the risk set is closed from below. Then there exists a minimax rule δ_0, which is Bayes'w.r.t. π_0.

(d) (i) Define a lower semicontinuous function.

(ii) Suppose Θ is infinite and D satisfies the following conditions.

(Ai) D has a compact subset C, which is essentially complete.

(Aii) $R(\theta, \delta)$ is lower semicontinuous in $\delta \in C$, for all $\theta \in \Theta$.

Then show that the relation in (1) holds and a minimax rule exists.

[State clearly the properties of lower semicontinuous functions and all other results of analysis that you use]. [3 + 5 + (10 + 10) + (3 + 7) = 38]