PhD QR exam 98.02 MathStat 第一卷

1. (25 points) Answer the following problems.

(a) (10 points) What does *variance stabilizing transformation* mean?

(b) (15 points) Let $X_1, ..., X_n$ be iid Bernoulli(θ). Find the variance stabilizing transformation h such that $h(0) = 0$, $h(1) = 1$, and $h'(t) \geq 0\ \forall t$.

2. (20 points) Suppose that given p, each of n trials are independent with success probability p. Assume that p has uniform prior on (0,1). Let X denote the number of successes in the n trials. Then,

(a) (10 points) Find $P(X = k)$ for $k = 0, 1, ..., n$.

(b) (10 points) Find $E(p|X)$.

3. (20 points) Consider the data set in the following table, where 12 observations are assumed to come from the bivariate normal distribution with $\mu_1 = \mu_2 = 0$, the correlation coefficient ρ, and variances σ_{11} and σ_{22}.

* Value not observed

1	1	-1	-1	2	2	-2	-2	*	*	*	*
1	-1	1	-1	*	*	*	*	2	2	-2	-2

In this data, 2 pairs have correlation 1, 2 pairs have correlation -1, and there are 8 missing values. Denote the covariance matrix as $\Sigma = [\sigma_{ij}]_{i,j=1,2}$. Assume the Jeffreys prior density on Σ; that is,

$$\xi(\Sigma) = \xi(\sigma_{11}, \sigma_{22}, \sigma_{12}) \propto |\Sigma|^{-(k+1)/2},$$

where $k = 2$ is the dimensional of the multivariate normal distribution.

(a) (10 points) Write down the joint posterior density of $(\sigma_{11}, \sigma_{22}, \rho)$.

(b) (10 points) Derive the marginal posterior density of ρ.

4. (15 points) The Kullback-Leibler information number between probability densities $f(x)$ and $g(x)$ is defined as

$$I(f, g) = E_f\Big(\log\frac{f(X)}{g(X)}\Big).$$

Show that $I(f, g) \geq 0$.

5. (20 points) Suppose that $X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j Z_j$ where $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ and Z_j are iid r.v.'s with mean 0 and variance σ^2. Let $\gamma(h) = E(X_t - \mu)(X_{t+h} - \mu)$, $-\infty < h < \infty$.

(a) (10 points) Show that $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$.

(b) (10 points) Show that $n\text{Var}(\bar{X}_n) \to \sum_{h=-\infty}^{\infty} \gamma(h)$.

Mathematical Statistics

第二卷 <1>

2/16/2009

1.(10%) Let X be a random variable with CDF F_X. Show that, if EX exists, then

$$EX=\int_0^{\infty}[1-F_X(x)]dx-\int_{-\infty}^{0}F_X(x)dx.$$

2.(16%) Let Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables having the uniform distribution $U[0, \theta]$. It is known that when $\theta \in \Theta=(0, \infty)$, the UMVUE of θ is $\frac{n+1}{n}X_{(n)}$. Now suppose that $\Theta=(1, \infty)$, show that the UMVUE for θ is $I_{[0,1]}(x_{(n)})+\frac{n+1}{n}x_{(n)}I_{(1,\infty)}(x_{(n)})$.

3.(12%) Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables having the exponential distribution $E(a, \theta)$, where $a \in R$, and $\theta > 0$.

a(6%) Show that $X_{(1)}$ and $\sum_{i=1}^{n}(X_i - X_{(1)})$ independent for any (a, θ).

b(6%) Show that $Z_i=(X_{(n)}-X_{(i)})/(X_{(n)}-X_{(n-1)})$, i=1, 2,..., $n-2$, are independent of $(X_{(1)}, \sum_{i=1}^{n}(X_i - X_{(1)}))$

4.(16%) Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables from the uniform distribution $U(\theta_1,\theta_2)$, $-\infty<\theta_1<\theta_2<\infty$.

a(8%) Show that the conditional distribution of $X_{(1)}$ given $X_{(n)}=x$ is the distribution of the minimum of a sample of size n-1 from the uniform distribution $U(\theta_1, x)$.

b(8%) Find a UMPU test of size α for testing $H_0:\theta_1 \le 0$ versus $H_1:\theta_1 > 0$.

5.(16%) Consider the bivariate discrete distribution with probability mass function

$p(x, y; \alpha,\beta,\gamma)=c\frac{\alpha^x\beta^y\gamma^{xy}}{x!y!}$, x = 0, 1, 2, ..., y = 0, 1, 2, ..., where $\alpha \ge 0$, $\beta \ge 0$, $0 \le \gamma \le 1$, and c is a normalizing constant. Consider a random sample of size n drawn from the bivariate distribution $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$. We want to construct a UMP test of size α of the null hypothesis that X and Y are independent against the alternative that X and Y are dependent.

a(6%) Determine the conditional distribution of Y given X, show that it's Poisson.

b(4%) Give the hypotheses of interest in terms of the parameters.

c(6%) Conduct the test with a sample of n=3, $(x_1, y_1)=(1, 3)$, $(x_2, y_2)=(2, 0)$ and $(x_3, y_3)=(1, 1)$.

6.(18%) Let X be a single observation from the discrete p.d.f.

$f_\theta(x)=[x!(1-e^{-\theta})]^{-1}\theta^x e^{-\theta} I_{\{1,2,\ldots\}}(x)$, where $\theta>0$ is unknown. Consider the estimation of $\theta/(1-e^{-\theta})$ under the squared error loss.

a(6%) Show that the estimator X is admissible.

b(6%) Show that X is not minimax unless $\sup_\theta R_T(\theta)=\infty$ for any estimator T=T(X).

c(6%) Find a loss function under which X is minimax and admissible.

7.(12%) Let $X_1, X_2, \ldots, X_n$ be i.i.d. from the discrete p.d.f. in the last problem where $\theta>0$ is unknown. Show that the likelihood equation has a unique root when the sample mean is greater than 1. Is this root an MLE of θ?

1. (40 pts; 10 pts for each part) Suppose that we have a sample $(X_1, \ldots, X_n)$ such that $X_1, \ldots, X_n$ are independent and identically distributed and $P(X_1 = 1) = p = 1 - P(X_1 = 0)$, where $0 < p < 1$. Suppose that $\alpha > 0$, $\beta > 0$, and $B(\alpha, \beta)$ is the Beta distribution with Lebesgue probability density function g, which is given by

$$g(y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1} I_{(0,1)}(y) \text{ for } y \in (-\infty, \infty),$$

where $I_{(0,1)}(\cdot)$ is the indicator function on $(0, 1)$ and $\Gamma(\cdot)$ is the Gamma function, which is defined by $\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt$ for $a > 0$.

   (a) Show that the Bayes estimator for p under square error loss with respect to the prior distribution $B(\alpha, \beta)$ is $(\alpha + \sum_{i=1}^n X_i)/(\alpha+\beta+n)$.

   (b) Consider estimating p using $\hat{p} = ((\sqrt{n}/2) + \sum_{i=1}^n X_i)/(\sqrt{n}+n)$. Show that $\hat{p}$ is minimax under square error loss.

   (c) Suppose that the loss of estimating p using a when $P(X_1 = 1) = p$ is $L(p, a) = (a-p)^2/p$. Determine whether the $\hat{p}$ in Part (b) is admissible under the loss L. Justify your answer.

   (d) Determine whether the $\hat{p}$ in Part (b) is consistent. Justify your answer.

2. (50 pts; 10 pts for each part) Suppose that we have a random sample $(X_1, \ldots, X_n)$ from a distribution with a Lebesgue probability density function f, which is given by

$$f(x) = \frac{\beta^\lambda}{\Gamma(\lambda)} \exp(\lambda x - \beta e^x) \text{ for all } x \in (-\infty, \infty),$$

where β and λ are positive parameters and $\Gamma(\cdot)$ is the Gamma function as in Problem 1. Suppose that α is a constant in $(0, 1)$.

   (a) Suppose that $\lambda = 1$. Find a statistic U that is complete and sufficient for β.

   (b) Suppose that $\lambda = 1$. Let $Y_i = e^{X_i}$ for $i = 1, \ldots, n$, $\bar{Y} = n^{-1}\sum_{i=1}^n Y_i$ and $V = \prod_{i=1}^n (Y_i/\bar{Y})$. Show that V is independent of the statistic U in part (a).

   (c) Suppose that T is a test for testing

$$H_0 : \lambda \le 1 \text{ versus } H_1 : \lambda > 1 \qquad (1)$$

   such that $\beta_T \le \alpha$ under H_0 and $\beta_T \ge \alpha$ under H_1, where β_T is the power function of T. Show that $E(T|U) = \alpha$ almost everywhere when $\lambda = 1$, where U is the statistic in part (a).

   (d) Find a UMPU level α test for the problem in (1). Also, express the rejection region of the UMPU test using the statistic V in part (b), if possible.

第三卷

(e) For every $c \in (-\infty, \infty)$, let g_c be the location transform such that $g_c(X_1, \ldots, X_n) = (X_1 + c, \ldots, X_n + c)$. Let $\mathcal{G} = \{g_c : c \in (-\infty, \infty)\}$. Show that $\mathcal{G}$ is a group of transformation and determine whether the testing problem in (1) is invariant under $\mathcal{G}$. Justify your answer.

3. (10 pts) Suppose that $X = (X_1, \ldots, X_n)$ is a random sample from $N(\theta, 1)$, where $\theta \in (-\infty, \infty)$. Consider two estimators of θ: $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$ and

$$T_n = \begin{cases} \bar{X} & \text{if } |\bar{X}| > n^{-1/4}; \\ \bar{X}/2 & \text{if } |\bar{X}| \le n^{-1/4}. \end{cases}$$

Determine whether T_n is asymptotically more efficient than $\bar{X}$. Justify your answer.