S.-T. Yau College Student Mathematics Contests 2011

Applied Math., Computational Math., Probability and Statistics

Individual

6:30–9:00 pm, July 9, 2011
(Please select 5 problems to solve)

1. Given a weight function $\rho(x) > 0$, let the inner-product corresponding to $\rho(x)$ be defined as follows:

$$(f, g) := \int_a^b \rho(x) f(x) g(x) \mathrm{d}x,$$

and let $\|f\| := (f, f)$.

(1) Define a sequence of polynomials as follows:

$$\begin{aligned} p_0(x) &= 1, \quad p_1(x) = x - a_1, \\ p_n(x) &= (x - a_n) p_{n-1}(x) - b_n p_{n-2}(x), \quad n = 2, 3, \cdots \end{aligned}$$

where

$$\begin{aligned} a_n &= \frac{(xp_{n-1}, p_{n-1})}{(p_{n-1}, p_{n-1})}, \quad n = 1, 2, \cdots \\ b_n &= \frac{(xp_{n-1}, p_{n-2})}{(p_{n-2}, p_{n-2})}, \quad n = 2, 3, \cdots . \end{aligned}$$

Show that $\{p_n(x)\}$ is an orthogonal sequence of monic polynomials.

(2) Let $\{q_n(x)\}$ be an orthogonal sequence of monic polynomials corresponding to the ρ inner product. (A polynomial is called *monic* if its leading coefficient is 1.) Show that $\{q_n(x)\}$ is unique and it minimizes $\|q_n\|$ amongst all monic polynomials of degree n.

(3) Hence or otherwise, show that if $\rho(x) = 1/\sqrt{1 - x^2}$ and $[a, b] = [-1, 1]$, then the corresponding orthogonal sequence is the Chebyshev polynomials:

$$T_n(x) = \cos(n \arccos x), \quad n = 0, 1, 2, \cdots .$$

and the following recurrent formula holds:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, \cdots .$$

(4) Find the best quadratic approximation to $f(x) = x^3$ on $[-1, 1]$ using $\rho(x) = 1/\sqrt{1 - x^2}$.

2. If two polynomials $p(x)$ and $q(x)$, both of fifth degree, satisfy

$$p(i) = q(i) = \frac{1}{i}, \qquad i = 2, 3, 4, 5, 6,$$

and

$$p(1) = 1, \qquad q(1) = 2,$$

find $p(0) - q(0)$.

3. Lay aside m black balls and n red balls in a jug. Supposes $1 \le r \le k \le n$. Each time one draws a ball from the jug at random.

1) If each time one draws a ball without return, what is the probability that in the k-th time of drawing one obtains exactly the r-th red ball?
2) If each time one draws a ball with return, what is the probability that in the first k times of drawings one obtained totally an odd number of red balls?

4. Let X and Y be independent and identically distributed random variables. Show that

$$E[|X + Y|] \ge E[|X|].$$

Hint: Consider separately two cases: $E[X^+] \ge E[X^-]$ and $E[X^+] < E[X^-]$.

5. Suppose that $X_1, \cdots, X_n$ are a random sample from the Bernoulli distribution with probability of success p_1 and $Y_1, \cdots, Y_n$ be an independent random sample from the Bernoulli distribution with probability of success p_2.

(a) Give a minimum sufficient statistic and the UMVU (uniformly minimum variance unbiased) estimator for $\theta = p_1 - p_2$.

(b) Give the Cramer-Rao bound for the variance of the unbiased estimators for $v(p_1) = p_1(1 - p_1)$ or the UMVU estimator for $v(p_1)$.

(c) Compute the asymptotic power of the test with critical region

$$|\sqrt{n}(\hat{p}_1 - \hat{p}_2)/\sqrt{2\hat{p}\hat{q}}| \ge z_{1-\alpha}$$

when $p_1 = p$ and $p_2 = p + n^{-1/2}\Delta$, where $\hat{p} = 0.5\hat{p}_1 + 0.5\hat{p}_2$.

6. Suppose that an experiment is conducted to measure a constant θ. Independent unbiased measurements y of θ can be made with either of two instruments, both of which measure with normal errors: for $i = 1, 2$, instrument i produces independent errors with a $N(0, \sigma_i^2)$ distribution. The two error variances σ_1^2 and σ_2^2 are known. When a measurement y is made, a record is kept of the instrument used so that after n measurements the data is $(a_1, y_1), \ldots, (a_n, y_n)$, where $a_m = i$ if y_m is obtained using instrument i. The choice between instruments is made independently for each observation in such a way that

$$P(a_m = 1) = P(a_m = 2) = 0.5, \quad 1 \le m \le n.$$

Let x denote the entire set of data available to the statistician, in this case $(a_1, y_1), \ldots, (a_n, y_n)$, and let $l_\theta(x)$ denote the corresponding log likelihood function for θ. Let $a = \sum_{m=1}^{n}(2 - a_m)$.

(a) Show that the maximum likelihood estimate of θ is given by

$$\hat{\theta} = \left(\sum_{m=1}^{n} 1/\sigma_{a_m}^2\right)^{-1} \left(\sum_{m=1}^{n} y_m/\sigma_{a_m}^2\right).$$

(b) Express the expected Fisher information I_θ and the observed Fisher information I_x in terms of n, σ_1^2, σ_2^2, and a. What happens to the quantity I_θ/I_x as $n \to \infty$?

(c) Show that a is an ancillary statistic, and that the conditional variance of $\hat{\theta}$ given a equals $1/I_x$. Of the two approximations

$$\hat{\theta} \,\dot{\sim}\, N(\theta, 1/I_\theta)$$

and

$$\hat{\theta} \,\dot{\sim}\, N(\theta, 1/I_x),$$

which (if either) would you use for the purposes of inference, and why?