Ma 494 — Theoretical Statistics

Problem Set #5 — Due March 24, 2010

Prof. Sawyer — Washington University

NOTE: 5 problems on 2 pages.

1. A statistician is studying observations $X_1, X_2, \ldots, X_n$ whose continuous or discrete density $f(x,\theta)$ depends on a parameter θ, and is interested in using an approximate 95% confidence interval for θ of the form $\big(A_n(X), B_n(X)\big)$.

The statistician would like to test the quality of the approximate confidence interval by simulation. To do this, she uses a computer to generate N independent random samples $X^k = (X_1^k, \ldots, X_n^k)$ where the X_j^k are independent random variables with density $f(x, \theta_0)$ for a value $\theta = \theta_0$. For each sample, the approximate 95% confidence interval $\big(A_n(X^k), B_n(X^k)\big)$ is constructed. Let M be the number of samples X^k $(1 \le k \le N)$ such that $\theta_0 \in \big(A_n(X^k), B_n(X^k)\big)$. Then M/N is an unbiased estimator of the true coverage probability p, and we can ask whether or not M/N is consistent with 0.95.

(i) Suppose $M/N = 0.94$ based on $N = 1000$ simulated samples. Is this consistent with $p = 0.95$? That is, does a symmetric 95% confidence interval for the true value of p based on $\widehat{p} = M/N$ contain $p_0 = 0.95$?

(ii) Suppose $M/N = 0.94$ based on $N = 10{,}000$ simulated samples. Is this consistent with $p = 0.95$?

(*Warning*: In calculating the width of the confidence interval for the true coverage probability p, use either $p = p_0 = 0.95$ or $p = \widehat{p}$. Do *not* use the conservative approximation for a binomial frequency ratio since p is most likely close to one. The conservative approximation can be much too conservative if p or $1-p$ are small.)

2. (Like Problem 5.8.6) Suppose that observations Y have a gamma distribution with parameters r and θ (see section 4.6 p329 in Chapter 4 of the text) and that we are interested in estimating θ. Taking a Bayesian viewpoint, we begin with a prior distribution $\pi_0(\theta)$ for θ, which we take to be also gamma distributed, now with parameters s and μ.

(i) Given one observation of Y, show that the posterior distribution $\pi_1(\theta \mid Y)$ is gamma with parameters $r+s$ and $Y+\mu$.

(ii) Find the Bayes estimator of θ based on one observation of Y and the square loss function.

3. Let $Y_1, Y_2, \ldots, Y_n$ be integer-valued random variables with the discrete geometric density function $f(y,p) = (1-p)^{y-1}p$ for $y = 1, 2, \ldots$ (see Section 4.4 p317). In particular, each $Y_i \ge 1$.

(i) (1/4) Find the maximum-likelihood estimator $\widehat{p}(Y)$ for $Y = (Y_1, \ldots, Y_n)$.

(ii) (3/4) Suppose that $\overline{Y} = 3.0$ and $n = 100$. Find the approximate asymptotic 95% confidence interval for p based on the likelihood. (*Hints*: See Section 5 in the main Math 494 handout on the Math 494 Web site. You can also use the formulas for $E(Y)$ and $\mathrm{Var}(Y)$ in Section 4.4 in the text.)

4. (Like Problem 6.2.2 page 438) A harmless concoction based on berries and roots is thought to increase the (Stanford-Binet) IQs of school-children with ADD (attention-deficit disorder). Past experience with ADD suggests that the IQ scores of children with ADD of this age are normally distributed with mean 95 and standard deviation 15. A random sample of 22 children with ADD are fed this concoction as a dietary supplement for two months. Let $\overline{Y}$ be the average of their IQ scores on a test given at the end of the period.

If the data is to be analyzed using $\alpha = 0.06$, what values of $\overline{Y}$ would cause the null hypothesis H_0 to be rejected, using the standard one-sided test? Assume that there is no effect on the standard deviation and that the scores are still normally distributed.

5. (Like Problem 6.4.14 page 461) A sample of size 1 (call it X) is taken from a distribution with density $f(x, \theta) = (\theta + 1)x^\theta$ for $0 \le x \le 1$. Let $H_0 : \theta = 1$ (so that the density is $2x$). Consider a test that rejects H_0 if $X \ge 0.90$.

(i) What is the level of significance α of this test?

(ii) If $H_1 : \theta = 10$, so that the density is $11x^{10}$ for $0 \le x \le 1$, what is the probability (β) of Type II error? What is the power of the test?