Ma 494 — Theoretical Statistics

Test #2 — April 14, 2010

Prof. Sawyer — Washington University

Take-home examination. Open book and notes. Due at end of period on 04/14/2010. Six (6) problems on 3 pages. Not all parts of problems will be equally weighted.

1. (i) An experimenter wants to make a decision about whether a hypothesis H_1 is true for a null hypothesis H_0. Eight studies in the literature give conflicting results, but all reported P-values for their tests (which were based on independent datasets). The P-values, arranged in increasing order, are

$$0.014 \quad 0.04 \quad 0.065 \quad 0.073 \quad 0.21 \quad 0.31 \quad 0.57 \quad 0.87$$

Only two of the eight studies rejected H_0, but several other results are suggestive. Use Fisher's meta-analysis method to find an aggregate P-value for the eight studies. Do you accept or reject H_0?

(ii) Fisher's procedure is based on a UMP (uniformly most powerful) test of an hypothesis H_0 against an alternative hypothesis H_1. What are H_0 and H_1?

(*Hint*: This was covered in class, and is also covered in Section 8 of the Math 494 notes.)

2. Given an independent sample $X_1, X_2, \ldots, X_n$ that are uniformly distributed $U(0, \theta)$ (that is, for $0 \le X_j \le \theta$) for some θ, Suppose that we want to test

$$H_0 : \theta = 1 \quad \text{against} \quad H_1 : \theta < 1$$

It follows by arguing as in Section 6.5 of the text (see also Example 1 in Section 7 of the notes) that the generalized likelihood ratio test (GLRT) of H_0 against H_1 at level α has critical region $\mathcal{C}_\alpha = \{ X : X_{\max} \le \lambda_\alpha \}$ where $P(X_{\max} \le \lambda_\alpha \mid H_0) = \alpha$. (This test rejects H_0 if $X_{\max} \le \lambda_\alpha$.)

Suppose that $X_{\max} = 0.80$. What is the smallest value of n such that this test rejects H_0 at level of significance 0.05?

3. (Like Prob 7.3.4 page 479 in text.) Let $X_1, \ldots, X_n$ be independent normal $N(\mu, \sigma^2)$ and let $S^2 = (1/(n-1)) \sum_{j=1}^{n} (X_j - \overline{X})^2$. Prove that

$$\mathrm{Var}(S^2) \;=\; \frac{2\sigma^4}{n-1}$$

(*Hint*: Note that a χ^2_m distribution with m degrees of freedom has a gamma distribution gamma$(m/2, 1/2)$. You can use the formulas in Chapter 4 for the mean and variance of a gamma-distributed random variable.)

4. Let $X_1, \ldots, X_n$ be an independent sample, where each X_j has probability density $f(x, \theta) = (1/2)\theta x^{-3/2} \exp\bigl(-\theta/\sqrt{x}\bigr)$ for $x > 0$ and $\theta > 0$. Find the maximum likelihood estimator of θ. (*Hint*: Be careful!)

5. (Like Problem 7.3.6 in the text.)

(i) Prove that

$$\lim_{n\to\infty} P\left(\frac{\chi_n^2 - n}{\sqrt{2n}} \le y\right) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-(1/2)x^2}\, dx$$

for all y, where χ_n^2 represents a random variable with a chi-square distribution with n degrees of freedom. (*Hint*: Use the definition of a chi-square distribution on page 474 and the central limit theorem. You can use $Z^2 \approx \chi_1^2 \approx \text{gamma}(1/2, 1/2)$ and formulas in Chapter 4 to find the mean and variance of Z^2.)

(ii) Use part (i) to find $(\chi_{48}^2)_{0.90}$, the 90% quantile of a chi-square distribution with 48 degrees of freedom. (That is, $A = A_{0.90}$ where $P(\chi_{48}^2 \le A_{0.90}) = 0.90$.) How close is this approximate result to the exact value of this quantile in Table A.3 at the end of the textbook?

(iii) Use part (i) to find $(\chi_{150}^2)_{0.90}$.

6. Let $X_1, \ldots, X_n$ be independent $N(\mu, \sigma^2)$. Consider the test of the two composite hypotheses

$$H_0 : \mu = \mu_0 \qquad \text{against} \qquad H_1 : \mu \ne \mu_0$$

with no conditions on σ^2. The GLRT test statistic in this case is

$$\widehat{LR}_n(X) \;=\; \frac{\max_{\mu,\sigma^2} L(\mu, \sigma, X_1, \ldots, X_n)}{\max_{\sigma^2} L(\mu_0, \sigma, X_1, \ldots, X_n)} \;=\; \frac{L(\widehat{\theta}_{H_1}(X), X_1, \ldots, X_n)}{L(\widehat{\theta}_{H_0}(X), X_1, \ldots, X_n)} \tag{1}$$

where $\theta = (\mu, \sigma^2)$ in (1) and $\widehat{\theta}_{H_1}, \widehat{\theta}_{H_0}$ are the MLEs under the two hypotheses. Thus H_0 has $m_0 = 1$ free parameter (σ^2) and H_1 has $m_1 = 2$ free parameters (μ, σ^2), so that the difference between the numbers of parameters is $d = m_1 - m_0 = 1$. It then follows from Theorem 7.1 in the Math 494 notes (on the Math 494 Web site) that

$$\lim_{n\to\infty} P\bigl(2 \log \widehat{LR}_n(X) \le y\bigr) \;=\; P(\chi_1^2 \le y) \tag{2}$$

for all y. Prove (2) directly in this case.

(*Hints*: (i) By equation (7.10) in the Math 494 handout,

$$\widehat{LR}_n(X) \;=\; \left(1 \;+\; \frac{T(X)^2}{n-1}\right)^{n/2} \tag{3}$$

where $T(X)$ is the one-sample Student-t statistic $T(X)$. This is derived in Appendix 7.A.4 of the textbook except for a change in sign in the exponent, which follows from the fact that the textbook definition of the GLRT statistic (see Section 6.5) has the likelihood for H_0 in the numerator instead of the denominator and thus is one over the definition in (1) above.

(ii) When random variables converge pointwise to a constant, as in the law of large numbers, or a series of numbers converges to a limit, as in $P\big(g_n(Y_n) \le y\big) = P\big(Y_n \le g_n^{-1}(y)\big)$ where Y_n are random variables and $g_n^{-1}(y) \to z$, assume that you can interchange limits and probabilities.)