UC Berkeley
Department of Statistics

STAT 210A: Introduction to Mathematical Statistics

Problem Set 9
Fall 2014

Issued: Friday, November 7 **Due:** Thursday, November 13 (beginning of class)

Problem 9.1
Suppose $X \sim N_p(\theta, I)$ and consider testing $H_0 : \theta \in \Omega_0$ versus $H_1 : \theta \notin \Omega_0$.

(a) Show that the likelihood ratio test statistic λ is equivalent to the distance D between X and Ω_0, defined as

$$D = \inf\{\|X - \theta\| \,:\, \theta \in \Omega_0\}.$$

(b) Using part (a), a generalized likelihood ratio test will reject H_0 if $D > c$. What is the significance level α for this test if $p = 2$ and $\Omega_0 = \{\theta : \theta_1 \leq 0, \theta_2 \leq 0\}$?

Problem 9.2
Suppose that $Y_1, \ldots, Y_n$ are i.i.d. samples from a normal $N(\mu, \sigma^2)$ distribution where $\sigma^2 > 0$ is known. We wish to test the hypothesis $H_0 : \mu = \mu_0$ versus the alternative $H_1 : \mu = \mu_1 > \mu_0$ using the sample mean $\bar{Y}_n = \frac{1}{n}\sum_{j=1}^n Y_j$.

(a) Suppose that we reject the null hypothesis if $\bar{Y}_n \geq t_\alpha$. Specify the choice of t_α that yields a test of level α.

(b) Show that the power of this test can be expressed as

$$\beta(\mu_1) \quad = \quad \Phi(z_\alpha + \delta_n)$$

where $z_\alpha = \sqrt{n}\,(\mu_0 - t_\alpha)/\sigma$, and $\delta_n = \sqrt{n}\,(\mu_1 - \mu_0)/\sigma$. Explain what happens to the power as $(\mu_1 - \mu_0)$ increases/decreases, σ increases/decreases, or n increases. Why is this reasonable?

(c) Now suppose that we instead observe the thresholded quantities $Z_i = \mathbb{I}[Y_i \geq \mu_0]$. (This might happen when data communication rates are limited, so that only a single bit can be transmitted.) Using the normal approximation to the binomial distribution, show that the power function for a level-α test using the mean $\bar{Z}_n$ can be approximated as

$$\beta_Z(\mu_1) \quad \approx \quad \Phi\left\{\frac{n\Phi(\delta_n/\sqrt{n}) - n/2 + \sqrt{n}z_\alpha/2}{\left[n\Phi(\delta_n/\sqrt{n})\left\{1 - \Phi(\delta_n/\sqrt{n})\right\}\right]^{1/2}}\right\}.$$

(d) Using the approximation $\Phi(\delta_n/\sqrt{n}) \approx \frac{1}{2} + (\delta_n/(\sqrt{2\pi n})$, show that the power from (c) further simplifies to

$$\beta_Z(\mu_1) \quad \approx \quad \Phi\left(z_\alpha + \sqrt{\frac{2}{\pi}}\delta_n\right).$$

Note: Comparing this power to that from part (b) provides an indication of what is lost by the thresholding operation.

Problem 9.3

(A nonparametric hypothesis test) A set of i.i.d. samples $Y_1, \ldots, Y_n$ is drawn from an unknown distribution F. The null hypothesis H_0 asserts that F has median $\mu = \mu_0$, while the alternative H_1 asserts that $\mu > \mu_0$. (Note that both hypotheses are composite, but neither is based on a parametric model, since only μ_0 is given while the form of F is unknown.)

(a) Consider the statistic $S = \sum_{i=1}^{n} \mathbb{I}[Y_i > \mu_0]$. Compute the exact distribution of S under the null hypothesis. How could this be useful in performing a hypothesis test?

(b) Consider the test $\delta_s(Y)$ that rejects H_0 when S exceeds some threshold s. Use asymptotic theory to approximate the level $\alpha(s) = \mathbb{E}_0[\delta_s(Y)]$ of this test under H_0 as a function of s. (This test is known as a nonparametric *sign test.*)

Problem 9.4

Suppose that we use a prior $\lambda = [\lambda_0 \quad (1-\lambda_0)]$ for a binary hypothesis test $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$ where $\lambda_0 = \mathbb{P}[\theta = \theta_0]$.

(a) Show that the Bayes risk under $0-1$ loss can be written as

$$r(\lambda, \delta) \quad = \quad \lambda_0 \mathbb{E}_0[\delta(X)] + (1-\lambda_0)\left\{1 - \mathbb{E}_1[\delta(X)]\right\}.$$

(b) Show that any test minimizing the Bayes risk can be expressed as a likelihood ratio test where the threshold t is a function of λ_0.

(c) Suppose that we are given i.i.d. samples $X_1, \ldots, X_n$ from $\mathbb{P}_\theta$ (where θ is either equal to θ_0 or θ_1). For any fixed $\lambda_0 \in (0,1)$, let $\delta_{n,\lambda}$ denote the Bayes test derived in part (b). Prove that $\lim_{n \to +\infty} r(\lambda, \delta_n) = 0$.

Problem 9.5

For each of the following problems, compute the generalized likelihood ratio test, and compute its asymptotic distribution under the specified hypothesis H_0.

(a) Let $X_1, \ldots, X_n$ be an i.i.d. sample of $N(\mu_x, \sigma_x^2)$ variates, and let $Y_1, \ldots Y_n$ be an i.i.d. sample of $N(\mu_y, \sigma_y^2)$ variates. Consider testing $H_0 : \mu_x = \mu_y$ and $\sigma_x^2 = \sigma_y^2$.

(b) For $i = 1, \ldots, k$, let $X_{i1}, \ldots, X_{in}$ be independent samples from Poisson distributions with means θ_i, respectively. Consider testing $H_0 : \theta_1 = \theta_2 = \ldots \theta_k$.

(c) Let $X_1, \ldots, X_n$ be an i.i.d. sample from the exponential distribution with parameter θ, and $Y_1, \ldots Y_n$ an i.i.d. sample from the exponential distribution with parameter μ. Consider testing $H_0 : \mu = 2\theta$.