Homework 7

Aleck Zhao

November 4, 2016

1. Suppose an iid vector of data $\underline{X} = (X_1, \cdots, X_n)$ can belong to one of two classes Y, where $Y = 0$ or $Y = 1$. A *decision rule* or *classifier* g is a function $g : \mathbb{R}^n \to \mathbb{R}$ that assigns to any n-tuple of data a value 0 or 1. Suppose that the so-called *class conditional* densities f_0 and f_1 of $\underline{X}$ are given,

$$f_j(x_1, \cdots, x_n) = f(x_1, \cdots, x_n \mid Y = j), \quad j = 0, 1$$

are given. Define $L^0(g)$ and $L^1(g)$ as follows:

$$L^0(g) = P(g(\underline{X}) = 1 \mid Y = 0), \quad L^1(g) = P(g(\underline{X}) = 0 \mid Y = 1)$$

For $c > 0$, define the decision rule

$$g_c(x_1, \cdots, x_n) = \begin{cases} 1 \text{ if } cf_1(x_1, \cdots, x_n) > f_0(x_1, \cdots, x_n) \\ 0 \text{ otherwise} \end{cases}$$

Prove that for any classifier g, if $L^0(g) < L^0(g_c)$, then $L^1(g) > L^1(g_c)$. In other words, if L^0 is required to be kept under a certain level, then the decision rule minimizing L^1 has the form g_c for some c.

2. Suppose we consider a Bayesian framework for hypothesis testing, in which we consider testing a simple null vs a simple alternative:

$$H_0 : \mu = \mu_0, \quad H_a : \mu = \mu_a$$

Suppose we have the probabilities $P(\mu = \mu_0)$ and $P(\mu = \mu_a)$ as the *prior probabilities* that the null or alternative are true. Suppose we are given a distribution of the data under both null and alternative, so that

$$f(x_1, \cdots, x_n \mid \mu = \mu_0), \quad f(x_1, \cdots, x_n \mid \mu = \mu_a)$$

are given. How would you use the prior and likelihood to construct a test of hypothesis. Be completely specific about your test statistic and how it is computed.

11. Suppose $X_1, \cdots, X_n$ are iid standard normal data. Show that the vector of random variables given by $(X_1 - \bar{X}, \cdots, X_n - \bar{X})$ is independent of $\bar{X}$ for the case $n = 2$. Use this independence for more general n to show that for any sample size n, the scaled sample variance $(n-1)s^2$ is the sum of squares of independent standard normal random variables.

Chapter 9: Hypothesis Testing and Assessing Goodness of Fit

10. Suppose that $X_1, \cdots, X_n$ form a random sample from a density function, $f(x \mid \theta)$, for which T is a sufficient statistic for θ. Show that the likelihood ratio test of $H_0 : \theta = \theta_0$ vs $H_A : \theta = \theta_1$ is a function of T. Explain how, if the distribution of T is known under H_0, the rejection region of the test may be chosen so that the test has the level α.

12. Let $X_1, \cdots, X_n$ be a random sample from an exponential distribution with the density function $f(x \mid \theta) = \theta \exp(-\theta x)$. derive a likelihood ratio test of $H_0 : \theta = \theta_0$ vs $H_A : \theta \neq \theta_0$, and show that the rejection region is of the form $\left\{ \bar{X} \exp(-\theta_0 \bar{X}) \le c \right\}$.

13. Suppose, to be specific, that in Problem 12, $\theta_0 = 1, n = 10$, and that $\alpha = 0.05$. In order to use the test, we must find the appropriate value of c.

    a. Show that the rejection region is of the form $\{\bar{X} \leq x_0\} \cup \{\bar{X} \geq x_1\}$, where x_0 and x_1 are determined by c.

    b. Explain why c should be chosen so that $P(\bar{X}\exp(-\bar{X}) \leq c) = 0.05$ when $\theta_0 = 1$.

    c. Explain why $\sum_{i=1}^{10} X_i$ and hence $\bar{X}$ follow gamma distributions when $\theta_0 = 1$. How could this knowledge be used to choose c?

    d. Suppose that you hadn't thought of the preceding fact. Explain how you could determine a good approximation to c by generating random numbers on a computer.

14. Suppose that under H_0, a measurement X is $N(0, \sigma^2)$, and that under H_1, X is $N(1, \sigma^2)$ and that the prior probability $P(H_0) = 2P(H_1)$. The hypothesis H_0 will be chosen if $P(H_0 \mid x) > P(H_1 \mid x)$. For $\sigma^2 = 0.1, 0.5, 1.0, 5.0$:

    a. For what values of X will H_0 be chosen?

    b. In the long run, what proportion of the time will H_0 be chosen if H_0 is true $2/3$ of the time?

18. Let $X_1, \cdots, X_n$ be iid random variables from a double exponential distribution with density

$$f(x) = \frac{1}{2}\lambda\exp(-\lambda\,|x|).$$

Derive a likelihood ratio test of the hypothesis $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda = \lambda_1$ where λ_0 and $\lambda_1 > \lambda_0$ are specified numbers. Is the test uniformly most powerful against the alternative $H_1 : \lambda > \lambda_0$?

20. Consider two PDFs on [0, 1]: $f_0(x) = 1$ and $f_1(x) = 2x$. Among all tests of the null hypothesis $H_0 : X \sim f_0(x)$ versus the alternative $X \sim f_1(x)$, with significance level $\alpha = 0.10$, how large can the power possibly be?

24. Let X be a binomial random variable with n trials and probability p of success.

    a. What is the GLR for testing $H_0 : p = 0.5$ vs $H_A : p \neq 0.5$?

    b. Show that the test rejects for large values of $|X - n/2|$.

    c. Using the null distribution of X, show how the significance level corresponding to a rejection region $|X - n/2| > k$ can be determined.

    d. If $n = 10$ and $k = 2$, what is the significance level of the test?

    e. Use the normal approximation to the binomial distribution to find the significance level if $n = 100$ and $k = 10$.

26. True or false:

    a. The generalize likelihood ratio statistic Λ is always less than or equal to 1.

    b. If the p-value is 0.03, the corresponding test will reject at the significance level 0.02.

    c. If a test rejects at a significance level 0.06, then the p-value is less than or equal to 0.06.

    d. The p-value of a test is the probability that the null hypothesis is correct.

    e. In testing a simple versus simple hypothesis via the likelihood ratio, the p-value equals the likelihood ratio.

    f. If a chi-square test statistic with 4 degrees of freedom has a value of 8.5, the p-value is less than 0.05.

30. Suppose that the null hypothesis is true, that the distribution of the test statistic, T say, is continuous with CDF F and that the test rejects for large values of T. Let V denote the p-value of the test.

    a. Show that $V = 1 - F(T)$.
    b. Conclude that the null distribution of V is uniform. (Hint: Prop C Section 2.3)
    c. If the null hypothesis is true, what is the probability that the p-value is greater than 0.1?
    d. Show that the test that rejects if $V < \alpha$ has significance level α.