Stat 346/446: Theoretical Statistics II: Practice Exercises 5 Solutions

Textbook Exercises

8.29 (346 & 446)Let X be one observation from a Cauchy(θ) distribution.

(a) Show that this family does not have an MLR.
Let $\theta_2 > \theta_1$. Then

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} = \frac{1+(x-\theta_1)^2}{1+(x-\theta_2)^2} = \frac{1+(1+\theta_1)^2/x^2 - 2\theta_1/x}{1+(1+\theta_2)^2/x^2 - 2\theta_2/x}.$$

The limit of this ratio as $x \to \infty$ or as $x \to -\infty$ is 1. So the ratio cannot be monotone increasing (or decreasing) between $-\infty$ and ∞. Thus, the family does not have MLR.

(b) Show that the test

$$\phi(x) = \begin{cases} 1 & \text{if } 1 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

is most powerful of its size for testing $H_0 : \theta = 0$ versus $H_1 : \theta = 1$. Calculate the Type I and Type II Error probabilities.
By the Neyman–Pearson Lemma, a test will be UMP if it rejects when $f(x|1)/f(x|0) > k$, for some constant k. Examination of the derivative shows that $f(x|1)/f(x|0)$ is decreasing for $x \le (1-\sqrt{5})/2 \approx -0.618$, is increasing for $(1-\sqrt{5})/2 \le x \le (1+\sqrt{5})/2 \approx 1.618$, and is decreasing for $(1+\sqrt{5})/2 \le x$. Furthermore,

$$\frac{f(1|1)}{f(1|0)} = \frac{f(3|1)}{f(3|0)} = 2.$$

So rejecting if $f(x|1)/f(x|0) > 2$ is equivalent to rejecting if $1 < x < 3$. Thus, the given test is UMP of its size. The size of the test is

$$P(1 < X < 3 \mid \theta = 0) = \int_1^3 \frac{1}{\pi} \cdot \frac{1}{1+x^2}\,dx = \frac{1}{\pi}\left[\arctan x\right]_1^3 \approx 0.1476.$$

The Type II error probability is

$$1 - P(1 < X < 3 \mid \theta = 1) = 1 - \int_1^3 \frac{1}{\pi} \cdot \frac{1}{1+(x-1)^2}\,dx = 1 - \frac{1}{\pi}\left[\arctan(x-1)\right]_1^3 \approx 0.6476.$$

(c) Prove or disprove: The test in part (b) is UMP for testing $H_0 : \theta \le 0$ versus $H_1 : \theta > 0$. What can be said about UMP tests in general for the Cauchy location family?
We will not have $\frac{f(1|\theta)}{f(1|0)} = \frac{f(3|\theta)}{f(3|0)}$ for any other value of $\theta \ne 1$. Try $\theta = 2$, for example. So the rejection region $1 < x < 3$ will not be most powerful at any other value of θ. The test is not UMP for testing $H_0 : \theta \le 0$ versus $H_1 : \theta > 0$.

8.30 (446) Let $f(x \mid \theta)$ be the Cauchy *scale* pdf

$$f(x \mid \theta) = \frac{\theta}{\pi(\theta^2 + x^2)}, \quad -\infty < x < \infty, \quad \theta > 0.$$

(a) Show that this family does not have an MLR.
For $\theta_2 > \theta_1 > 0$, the likelihood ratio and its derivative are

$$\frac{f(x \mid \theta_2)}{f(x \mid \theta_1)} = \frac{\theta_2}{\theta_1} \cdot \frac{\theta_1^2 + x^2}{\theta_2^2 + x^2}, \quad \text{and} \quad \frac{d}{dx}\left(\frac{f(x \mid \theta_2)}{f(x \mid \theta_1)}\right) = \frac{\theta_2}{\theta_1} \cdot \frac{\theta_2^2 - \theta_1^2}{(\theta_2^2 + x^2)^2} \cdot x.$$

The sign of the derivative is the same as the sign of x (recall, $\theta_2^2 - \theta_1^2 > 0$), which changes sign. Hence the ratio is not monotone.

(b) If X is one observation from $f(x \mid \theta)$, show that $|X|$ is sufficient for θ and that the distribution of $|X|$ *does* have an MLR.
Because $f(x \mid \theta) = (\theta/\pi)(\theta^2 + |x|^2)^{-1}$, $Y = |X|$ is sufficient. Its pdf is

$$f(y \mid \theta) = \frac{2\theta}{\pi} \cdot \frac{1}{\theta^2 + y^2}, \quad y > 0.$$

Differentiating as above, the sign of the derivative is the same as the sign of y, which is positive. Hence the family has MLR.

8.33 (346 & 446)Let $X_1, \ldots, X_n$ be a random sample from the uniform$(\theta, \theta + 1)$ distribution. To test $H_0\colon \theta = 0$ versus $H_1\colon \theta > 0$, use the test

$$\text{reject } H_0 \text{ if } Y_n \geq 1 \text{ or } Y_1 \geq k,$$

where k is a constant, $Y_1 = \min\{X_1, \ldots, X_n\}$, $Y_n = \max\{X_1, \ldots, X_n\}$.

(a) Determine k so that the test will have size α.
From Theorems 5.4.4 and 5.4.6, the marginal pdf of Y_1 and the joint pdf of (Y_1, Y_n) are

$$f(y_1 \mid \theta) = n(1 - (y_1 - \theta))^{n-1}, \quad \theta < y_1 < \theta + 1,$$

$$f(y_1, y_n \mid \theta) = n(n-1)(y_n - y_1)^{n-2}, \quad \theta < y_1 < y_n < \theta + 1.$$

Under H_0, $P(Y_n \geq 1) = 0$. So

$$\alpha = P(Y_1 \geq k \mid 0) = \int_k^1 n(1 - y_1)^{n-1}\, dy_1 = (1-k)^n.$$

Thus, use $k = 1 - \alpha^{1/n}$ to have a size α test.

(b) Find an expression for the power function of the test in part (a).
For $\theta \leq k - 1$, $\beta(\theta) = 0$. For $k - 1 < \theta \leq 0$,

$$\beta(\theta) = \int_k^{\theta+1} n(1 - (y_1 - \theta))^{n-1}\, dy_1 = (1 - k + \theta)^n.$$

For $0 < \theta \leq k$,

$$\begin{aligned}\beta(\theta) &= \int_k^{\theta+1} n(1 - (y_1 - \theta))^{n-1}\, dy_1 + \int_\theta^k \int_1^{\theta+1} n(n-1)(y_n - y_1)^{n-2}\, dy_n\, dy_1 \\ &= \alpha + 1 - (1-\theta)^n.\end{aligned}$$

And for $k < \theta$, $\beta(\theta) = 1$.

(c) Prove that the test is UMP size α.
(Y_1, Y_n) are sufficient statistics. So we can attempt to find a UMP test using Corollary 8.3.13 and the joint pdf $f(y_1, y_n \mid \theta)$ in part (a). For $0 < \theta < 1$, the ratio of pdfs is

$$\frac{f(y_1, y_n \mid \theta)}{f(y_1, y_n \mid 0)} = \begin{cases} 0 & \text{if } 0 < y_1 \leq \theta,\ y_1 < y_n < 1, \\ 1 & \text{if } \theta < y_1 < y_n < 1, \\ \infty & \text{if } 1 \leq y_n < \theta + 1,\ \theta < y_1 < y_n. \end{cases}$$

For $1 \leq \theta$, the ratio of pdfs is

$$\frac{f(y_1, y_n \mid \theta)}{f(y_1, y_n \mid 0)} = \begin{cases} 0 & \text{if } y_1 < y_n < 1, \\ \infty & \text{if } \theta < y_1 < y_n < \theta + 1. \end{cases}$$

For $0 < \theta < k$, use $k' = 1$. The given test always rejects if $f(y_1, y_n \mid \theta)/f(y_1, y_n \mid 0) > 1$ and always accepts if $f(y_1, y_n \mid \theta)/f(y_1, y_n \mid 0) < 1$. For $\theta \geq k$, use $k' = 0$. The given test always rejects if $f(y_1, y_n \mid \theta)/f(y_1, y_n \mid 0) > 0$ and always accepts if $f(y_1, y_n \mid \theta)/f(y_1, y_n \mid 0) < 0$. Thus the given test is UMP by Corollary 8.3.13.

(d) Find values of n and k so that the UMP 0.10 level test will have power at least 0.8 if $\theta > 1$.
According to the power function in part (b), $\beta(\theta) = 1$ for all $\theta \geq k = 1 - \alpha^{1/n}$. So these conditions are satisfied for any n.

8.34 (346 & 446)In each of the following two situations, show that for any number c, if $\theta_1 \leq \theta_2$, then

$$P_{\theta_1}(T > c) \leq P_{\theta_2}(T > c).$$

(a) θ is a location parameter in the distribution of the random variable T.
Let $\theta_1 < \theta_2$, and suppose $T \sim f(t - \theta)$, i.e., θ is a location parameter. Let $X_1 \sim f(x - \theta_1)$, $X_2 \sim f(x - \theta_2)$, and let $Z \sim f(z)$ (i.e., the common density shifted by θ). Let $F(z)$ be the cumulative distribution function (CDF) corresponding to $f(z)$. Then for any real number x:

$$\begin{aligned} F(x \mid \theta_1) &= P(X_1 \leq x) = P(Z + \theta_1 \leq x) = P(Z \leq x - \theta_1) = F(x - \theta_1) \\ &\leq F(x - \theta_2) = P(Z \leq x - \theta_2) = P(Z + \theta_2 \leq x) = P(X_2 \leq x) = F(x \mid \theta_2). \end{aligned}$$

The inequality holds because $x - \theta_1 > x - \theta_2$, and F is non-decreasing. To get *strict inequality* for some x, choose an interval (a, b) of positive length $\theta_2 - \theta_1$, such that:

$$P(a < Z \leq b) = F(b) - F(a) > 0.$$

Now let $x = b + \theta_1$. Then:

$$\begin{aligned} F(x \mid \theta_1) &= F(b + \theta_1 - \theta_1) = F(b), \\ F(x \mid \theta_2) &= F(b + \theta_1 - \theta_2) = F(a), \\ \Rightarrow F(x \mid \theta_1) &= F(b) > F(a) = F(x \mid \theta_2). \end{aligned}$$

Hence,

$$F(x \mid \theta_1) < F(x \mid \theta_2), \quad \text{so} \quad P_{\theta_1}(T > c) > P_{\theta_2}(T > c).$$

(b) The family of pdfs of T, $\{g(t \mid \theta) : \theta \in \Theta\}$, has an MLR.
We can first prove the result for continuous distributions. (A similar argument can be adapted for discrete MLR families.) Let $F(t \mid \theta)$ denote the CDF of T under parameter θ. We aim to show that

$$F(t \mid \theta_1) \geq F(t \mid \theta_2), \quad \text{for all } t, \text{ if } \theta_1 < \theta_2,$$

which implies

$$P_{\theta_1}(T > c) = 1 - F(c \mid \theta_1) \leq 1 - F(c \mid \theta_2) = P_{\theta_2}(T > c).$$

Now define the function:

$$h(t) = F(t \mid \theta_1) - F(t \mid \theta_2).$$

Differentiate with respect to t:

$$\frac{d}{dt}h(t) = f(t \mid \theta_1) - f(t \mid \theta_2) = f(t \mid \theta_2)\left(\frac{f(t \mid \theta_1)}{f(t \mid \theta_2)} - 1\right).$$

Since the likelihood ratio $\frac{f(t|\theta_1)}{f(t|\theta_2)}$ is decreasing in t (because MLR is increasing in T as θ increases), the derivative of $h(t)$ changes sign at most once:

- it is negative before the curves cross,
- and positive after,

which implies that $h(t)$ has at most one minimum, and it is non-positive everywhere. Therefore, $h(t) \leq 0 \Rightarrow F(t \mid \theta_1) \leq F(t \mid \theta_2)$, which implies:

$$P_{\theta_1}(T > c) \leq P_{\theta_2}(T > c).$$

DS Section 9.3 Exercises

8 (346 & 446) Suppose that $X_1, \ldots, X_n$ form a random sample from the normal distribution with known mean 0 and unknown variance σ^2, and suppose that it is desired to test the following hypotheses:

$$H_0 \colon \sigma^2 \leq 2, \quad H_1 \colon \sigma^2 > 2.$$

Show that there exists a UMP test of these hypotheses at every level of significance α_0 $(0 < \alpha_0 < 1)$.
Since the mean is known and equal to 0, the sufficient statistic for σ^2 is

$$T = \sum_{i=1}^{n} X_i^2.$$

This statistic T is distributed as

$$T \sim \sigma^2 \cdot \chi_n^2,$$

because each $X_i \sim N(0, \sigma^2)$, so $X_i^2/\sigma^2 \sim \chi_1^2$, and the sum of n such independent variables gives χ_n^2. Now consider the likelihood ratio test: We compare the likelihood under H_1 $(\sigma^2 > 2)$ to that under H_0 $(\sigma^2 \leq 2)$. The most powerful test against $H_1 : \sigma^2 = \sigma_1^2 > 2$ versus $H_0 : \sigma^2 = 2$ by the Neyman–Pearson Lemma is based on the statistic T, and the rejection region takes the form:

$$\text{Reject } H_0 \text{ if } T = \sum_{i=1}^{n} X_i^2 \geq c$$

Because $T \sim \sigma^2 \cdot \chi_n^2$ and larger σ^2 leads to stochastically larger values of T, the test that rejects for large values of T is UMP for testing $H_0 : \sigma^2 \leq 2$ vs. $H_1 : \sigma^2 > 2$. To achieve a specified level of significance α_0, the constant c should be chosen so that

$$\Pr\left(\sum_{i=1}^{n} X_i^2 \geq c \,\middle|\, \sigma^2 = 2\right) = \alpha_0.$$

Since $\sum_{i=1}^{n} X_i^2$ has a continuous distribution and not a discrete distribution, there will be a value of c which satisfies this equation for any specified value of $\alpha_0 \in (0, 1)$.

9 (346 & 446) Show that the UMP test in Exercise 8 rejects H_0 when

$$\sum_{i=1}^{n} X_i^2 \geq c,$$

and determine the value of c when $n = 10$ and $\alpha_0 = 0.05$.
The first part of this exercise was answered in Exercise 8. When $n = 10$ and $\sigma^2 = 2$, the distribution of

$$Y = \sum_{i=1}^{n} \frac{X_i^2}{2}$$

will be the χ^2 distribution with 10 degrees of freedom, and it is found from a table of this distribution that $\Pr(Y \geq 18.31) = 0.05$. Also,

$$\Pr\left(\sum_{i=1}^{n} X_i^2 \geq c \,\middle|\, \sigma^2 = 2\right) = \Pr\left(Y \geq \frac{c}{2}\right).$$

Therefore, if this probability is to be equal to 0.05, then $c/2 = 18.31$ or $c = 36.62$.