Exercise 10.104

Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a uniform distribution over the interval $(0, \theta)$.

- **a** Find the most powerful α-level test for testing $H_0 : \theta = \theta_0$ against $H_a : \theta = \theta_\alpha$, where $\theta_a < \theta_0$.
- **b** Is the test in part **(a)** uniformly most powerful for testing $H_0 : \theta = \theta_0$ against $H_a : \theta < \theta_0$?
- **c** Find the most powerful α-level test for testing $H_0 : \theta = \theta_0$ against $H_a : \theta = \theta_a$, where $\theta_a > \theta_0$.
- **d** Is the test in part **(c)** uniformly most powerful for testing $H_0 : \theta = \theta_0$ against $H_a : \theta > \theta_0$?
- **e** Is the most powerful α-level test that you found in part **(c)** unique?

Exercise 10.113

Suppose that independent random samples of sizes n_1 and n_2 are to be selected from normal populations with means μ_1 and μ_2, respectively, and common variance σ^2.

a Show that in testing of $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 \neq \mu_2$ (σ^2 unknown) the likelihood ratio test reduces to the two-sample t test.

b Suppose that another independent random sample of size n_3 is selected from a third normal population with mean μ_3 and variance σ^2. Find the likelihood ratio test for testing $H_0 : \mu_1 = \mu_2 = \mu_3$ versus the alternative that there is at least one inequality. Show that this test is equivalent to an exact F test.

Exercise 10.104

Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a uniform distribution over the interval $(0, \theta)$.

a The likelihood function is $L(\theta) = \theta^{-n} I_{0,\theta}\left(y_{(n)}\right)$. To test $H_0 : \theta = \theta_0$ vs. $H_a : \theta = \theta_a$, where $\theta_a < \theta_0$, the best test is

$$\frac{L\left(\theta_0\right)}{L\left(\theta_a\right)} = \left(\frac{\theta_a}{\theta_0}\right)^n \frac{I_{0,\theta_0}\left(y_{(n)}\right)}{I_{0,\theta_a}\left(y_{(n)}\right)} < k.$$

So, the test only depends on the value of the largest order statistic $Y_{(n)}$, and the test rejects whenever $Y_{(n)}$ is small. The density function for $Y_{(n)}$ is $g_n(y) = ny^{n-1}\theta^{-n}$, for $0 \le y \le \theta$. For a size α test, select c such that

$$\alpha = P\left(Y_{(n)} < c \mid \theta = \theta_0\right) = \int_0^c ny^{n-1}\theta_0^{-n} dy = \frac{c^n}{\theta_0^n},$$

so $c = \theta_0 \alpha^{1/n}$. So, the RR is $\left\{Y_{(n)} < \theta_0 \alpha^{1/n}\right\}$.

b Since the RR does not depend on the specific value of $\theta_a < \theta_0$, it is UMP.

c The test is based on $Y_{(n)}$. In the case, the rejection region is of the form $\left\{Y_{(n)} > c\right\}$. For a size α test select c such that

$$\alpha = P\left(Y_{(n)} > c \mid \theta = \theta_0\right) = \int_c^{\theta_0} ny^{n-1}\theta_0^{-n} dy = 1 - \frac{c^n}{\theta_0^n},$$

so $c = \theta_0(1-\alpha)^{1/n}$.

d the test is UMP.

e It is not unique. Another interval for the RR can be selected so that it is of size α and the power is the same as in part a and independent of the interval. Example: choose the rejection region $C = (a, b) \cup (\theta_0, \infty)$, where $(a, b) \subset (0, \theta_0)$. Then,

$$\alpha = P\left(a < Y_{(n)} < b \mid \theta_0\right) = \frac{b^n - a^n}{\theta_0^n},$$

The power of this test is given by

$$P\left(a < Y_{(n)} < b \mid \theta_a\right) + P\left(Y_{(n)} > \theta_0 \mid \theta_a\right) = \frac{b^n - a^n}{\theta_a^n} + \frac{\theta_a^n - \theta_0^n}{\theta_a^n} = (\alpha - 1)\frac{\theta_0^n}{\theta_a^n} + 1,$$

which is independent of the interval (a, b) and has the same power as in part **(c)**.

Exercise 10.113

a Denote the samples as $X_1, \ldots, X_{n_1}$, and $Y_1, \ldots, Y_{n_2}$, where $n = n_1 + n_2$. Under H_a (unrestricted), the MLEs for the parameters are:

$$\hat{\mu}_1 = \bar{X}, \hat{\mu}_2 = \bar{Y}, \hat{\sigma}^2 = \frac{1}{n}\left(\sum_{i=1}^{n_1}\left(X_i - \bar{X}\right)^2 + \sum_{i=1}^{n_2}\left(Y_i - \bar{Y}\right)^2\right).$$

Under $H_0, \mu_1 = \mu_2 = \mu$ and the MLEs are

$$\hat{\mu} = \frac{n_1\bar{X} + n_2\bar{Y}}{n}, \hat{\sigma}_0^2 = \frac{1}{n}\left(\sum_{i=1}^{n_1}(X_i - \hat{\mu})^2 + \sum_{i=1}^{n_2}(Y_i - \hat{\mu})^2\right).$$

By defining the LRT, it is found to be equal to

$$\lambda = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2} \leq k, \text{ or equivalently reject if } \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right) \geq k'$$

Now, write

$$\sum_{i=1}^{n_1}(X_i - \hat{\mu})^2 = \sum_{i=1}^{n_1}\left(X_i - \bar{X} + \bar{X} - \hat{\mu}\right)^2 = \sum_{i=1}^{n_1}\left(X_i - \bar{X}\right)^2 + n_1(\bar{X} - \hat{\mu})^2,$$
$$\sum_{i=1}^{n_2}(Y_i - \hat{\mu})^2 = \sum_{i=1}^{n_2}\left(Y_i - \bar{Y} + \bar{Y} - \hat{\mu}\right)^2 = \sum_{i=1}^{n_2}\left(Y_i - \bar{Y}\right)^2 + n_2(\bar{Y} - \hat{\mu})^2,$$

and since $\hat{\mu} = \frac{n_1}{n}\bar{X} + \frac{n_2}{n}\bar{Y}$, and alternative expression for $\hat{\sigma}_0^2$ is

$$\sum_{i=1}^{n_1}\left(X_i - \bar{X}\right)^2 + \sum_{i=1}^{n_2}\left(Y_i - \bar{Y}\right)^2 + \frac{n_1 n_2}{n}(\bar{X} - \bar{Y})^2.$$

Thus, the LRT rejects for large values of

$$1 + \frac{n_1 n_2}{n}\left(\frac{(\bar{X} - \bar{Y})^2}{\sum_{i=1}^{n_1}\left(X_i - \bar{X}\right)^2 + \sum_{i=1}^{n_2}\left(Y_i - \bar{Y}\right)^2}\right).$$

Equivalently, the test rejects for large values of

$$\frac{|\bar{X} - \bar{Y}|}{\sqrt{\sum_{i=1}^{n_1}\left(X_i - \bar{X}\right)^2 + \sum_{i=1}^{n_2}\left(Y_i - \bar{Y}\right)^2}}.$$

This is equivalent to the two-sample t test statistic (σ^2 unknown) except for the constants that do not depend on the data.

b Using the sample notation $Y_{11}, \ldots, Y_{1n_1}, Y_{21}, \ldots, Y_{2n_2}, Y_{31}, \ldots, Y_{3n_3}$, with $n = n_1 + n_2 + n_3$, we have that under H_a (unrestricted hypothesis), the MLEs for the parameters are:

$$\hat{\mu}_1 = \bar{Y}_1, \hat{\mu}_2 = \bar{Y}_2, \hat{\mu}_3 = \bar{Y}_3, \hat{\sigma}^2 = \frac{1}{n}\left(\sum_{i=1}^{3}\sum_{j=1}^{n_i}\left(Y_{ij} - \bar{Y}_i\right)^2\right).$$

Under $H_0, \mu_1 = \mu_2 = \mu_3 = \mu$ so the MLEs are

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{3}\sum_{j=1}^{n_i} Y_{ij} = \frac{n_1\bar{Y}_1 + n_2\bar{Y}_2 + n_3\overline{Y_3}}{n}, \quad \hat{\sigma}_0^2 = \frac{1}{n}\left(\sum_{i=1}^{3}\sum_{j=1}^{n_i}(Y_{ij} - \hat{\mu})^2\right).$$

Similar to Ex. 10.112, ny defining the LRT, it is found to be equal to

$$\lambda = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2} \leq k, \text{ or equivalently reject if } \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right) \geq k'$$

In order to show that this test is equivalent to and exact F test, we refer to results and notation given in Section 13.3 of the text. In particular,

$$\begin{aligned} n\hat{\sigma}^2 &= \text{SSE} \\ n\hat{\sigma}_0^2 &= \text{TSS} = \text{SST} + \text{SSE} \end{aligned}$$

Then, we have that the LRT rejects when

$$\frac{\text{TSS}}{\text{SSE}} = \frac{\text{SSE} + \text{SST}}{\text{SSE}} = 1 + \frac{\text{SST}}{\text{SSE}} = 1 + \frac{\text{MST}}{\text{MSE}}\frac{2}{n-3} = 1 + F\frac{2}{n-3} \geq k',$$

where the statistic $F = \frac{\text{MST}}{\text{MSE}} = \frac{\text{SST}/2}{\text{SSE}/(n-3)}$ has an F-distribution with 2 numerator and $n-3$ denominator degrees of freedom under H_0. The LRT rejects when the statistic F is large and so the tests are equivalent,