Exercise 8.128

Suppose that we take a sample of size n_1 from a normally distributed population with mean and variance μ_1 and σ_1^2 and an independent of sample size n_2 from a normally distributed population with mean and variance μ_2 and σ_2^2. What can be done if we cannot assume that the unknown variances are equal but are fortunate enough to know that $\sigma_2^2 = k\sigma_1^2$ for some known constant $k \neq 1$? Suppose, as previously, that the sample means are given by $\bar{Y}_1$ and $\bar{Y}_2$ and the sample variances by S_1^2 and S_2^2, respectively.

a Show that $Z^\star$ given below has a standard normal distribution.

$$Z^\star = \frac{\left(\bar{Y}_1 - \bar{Y}_2\right) - (\mu_1 - \mu_2)}{\sigma_1 \sqrt{\frac{1}{n_1} + \frac{k}{n_2}}}$$

b Show that $W^\star$ given below has a χ^2 distribution with $n_1 + n_2 - 2$ df.

$$W^\star = \frac{(n_1 - 1)\, S_1^2 + (n_2 - 1)\, S_2^2/k}{\sigma_1^2}$$

c Notice that $Z^\star$ and $W^\star$ from parts (a) and (b) are independent. Finally, show that

$$T^\star = \frac{\left(\bar{Y}_1 - \bar{Y}_2\right) - (\mu_1 - \mu_2)}{S_p^\star \sqrt{\frac{1}{n_1} + \frac{k}{n_2}}}, \qquad \text{where } S_p^{2\star} = \frac{(n_1 - 1)\, S_1^2 + (n_2 - 1)\, S_2^2/k}{n_1 + n_2 - 2}$$

has a t distribution with $n_1 + n_2 - 2$ df.

d Use the result in part (c) to give a $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$, assuming that $\sigma_2^2 = k\sigma_1^2$.

e What happens if $k = 1$ in parts (a)-(d)?

Exercise 8.133

Suppose that two independent random samples of n_1 and n_2 observations are selected from normal populations. Further, assume that the populations possess a common variance σ^2. Let

$$S_i^2 = \frac{\sum_{j=1}^{n_i} \left(Y_{ij} - \bar{Y}_i\right)^2}{n_i - 1}, \quad i = 1, 2$$

a Show that S_p^2, the pooled estimator of σ^2 (which follows), is unbiased:

$$S_p^2 = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}$$

b Find $V\left(S_p^2\right)$.

Exercise 8.135

A confidence interval is unbiased if the expected value of the interval midpoint is equal to the estimated parameter. The expected value of the midpoint of the large-sample confidence interval (Section 8.6) is equal to the estimated parameter, and the same is true for the smallsample confidence intervals for μ and $(\mu_1 - \mu_2)$ (Section 8.8). For example, the midpoint of the interval $\bar{y} \pm ts/\sqrt{n}$ is $\bar{y}$, and $E(\bar{Y}) = \mu$. Now consider the confidence interval for σ^2. Show that the expected value of the midpoint of this confidence interval is not equal to σ^2.

Exercise 9.8

Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a probability density function $f(y)$, which has unknown parameter θ. If $\hat{\theta}$ is an unbiased estimator of θ, then under very general conditions

$$V(\hat{\theta}) \geq I(\theta), \qquad \text{where } I(\theta) = \left[nE\left(-\frac{\partial^2 \ln f(Y)}{\partial \theta^2} \right) \right]^{-1}$$

(This is known as the Cramer-Rao inequality.) If $V(\hat{\theta}) = I(\theta)$, the estimator $\hat{\theta}$ is said to be efficient.

a Suppose that $f(y)$ is the normal density with mean μ and variance σ^2. Show that $\bar{Y}$ is an efficient estimator of μ.

b This inequality also holds for discrete probability functions $p(y)$. Suppose that $p(y)$ is the Poisson probability function with mean λ. Show that $\bar{Y}$ is an efficient estimator of λ.

Exercise 8.128

a Following the notation of Section 8.8 and the assumptions given in the problem, we know that $\bar{Y}_1 - \bar{Y}_2$ is a normal variable with mean $\mu_1 - \mu_2$ and variance $\frac{\sigma_1^2}{n_1} + \frac{k\sigma_1^2}{n_2}$. Thus, the standardized variable Z^* as defined indeed has a standard normal distribution.

b The quantities $U_1 = \frac{(n_1-1)S_1^2}{\sigma_1^2}$ and $U_2 = \frac{(n_2-1)S_2^2}{k\sigma_1^2}$ have independent chi-square distributions with $n_1 - 1$ and $n_2 - 1$ degrees of freedom (respectively). So, $W^* = U_1 + U_2$ has a chi-square distribution with $n_1 + n_2 - 2$ degrees of freedom.

c By Definition 7.2 , the quantity $T^* = \frac{Z^*}{\sqrt{W^*/(n_1+n_2-2)}}$ follows a t-distribution with $n_1 + n_2 - 2$ degrees of freedom.

d A $100(1-\alpha)\%$ CI for $\mu_1 - \mu_2$ is given by $\bar{Y}_1 - \bar{Y}_2 \pm t_{\alpha/2} S_p^* \sqrt{\frac{1}{n_1} + \frac{k}{n_2}}$, where $t_{\alpha/2}$ is the upper-$\alpha/2$ critical value from the t-distribution with $n_1 + n_2 - 2$ degrees of freedom and S_p^* is defined in part (c).

e If $k = 1$, it is equivalent to the result for $\sigma_1 = \sigma_2$.

Exercise 8.133

We know that $E\left(S_i^2\right) = \sigma^2$ and $V\left(S_i^2\right) = \frac{2\sigma^4}{n_i - 1}$ for $i = 1, 2$.

a $E\left(S_p^2\right) = \frac{(n_1-1)E\left(S_1^2\right) + (n_2-1)E\left(S_2^2\right)}{n_1+n_2-2} = \sigma^2$

b $V\left(S_p^2\right) = \frac{(n_1-1)^2 V\left(S_1^2\right) + (n_2-1)^2 V\left(S_2^2\right)}{(n_1+n_2-2)^2} = \frac{2\sigma^4}{n_1+n_2-2}$.

Exercise 8.135

The midpoint of the CI is given by $M = \frac{1}{2}\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}} + \frac{(n-1)S^2}{\chi^2_{\alpha/2}}\right)$. Therefore, since $E\left(S^2\right) = \sigma^2$, we have

$$E(M) = \frac{1}{2}\left(\frac{(n-1)\sigma^2}{\chi^2_{1-\alpha/2}} + \frac{(n-1)\sigma^2}{\chi^2_{\alpha/2}}\right) = \frac{(n-1)\sigma^2}{2}\left(\frac{1}{\chi^2_{1-\alpha/2}} + \frac{1}{\chi^2_{\alpha/2}}\right) \neq \sigma^2$$

Exercise 9.8

a It is not difficult to show that $\frac{\partial^2 \ln f(y)}{\partial \mu^2} = -\frac{1}{\sigma^2}$, so $I(\mu) = \sigma^2/n$, Since $V(\bar{Y}) = \sigma^2/n, \bar{Y}$ is an efficient estimator of μ.

b Similarly, $\frac{\partial^2 \ln p(y)}{\partial \lambda^2} = -\frac{y}{\lambda^2}$ and $E\left(-Y/\lambda^2\right) = 1/\lambda$. Thus, $I(\lambda) = \lambda/n$. $\bar{Y}$ is an efficient estimator of λ.