Homework 9 solutions

Exercise 1

$X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma)$

$Y_1, \ldots, Y_n \overset{iid}{\sim} N(\mu, \sigma)$

$$\frac{L(x_1, \ldots, x_n; \mu, \sigma^2)}{L(y_1, \ldots, y_n; \mu, \sigma^2)} = \frac{(2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum (x_i-\mu)^2}}{(2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum (y_i-\mu)^2}}$$

$$= e^{\frac{1}{2\sigma^2}\left(\sum (y_i-\mu)^2 - \sum (x_i-\mu)^2\right)}$$

$$= e^{\frac{1}{2\sigma^2}\left(\sum y_i^2 - n\mu^2 - 2\mu \sum y_i - \sum x_i^2 - n\mu^2 + 2\mu \sum x_i\right)}$$

$$= e^{\frac{1}{2\sigma^2}\left[\sum y_i^2 - \sum x_i^2 - 2\mu\left(\sum y_i - \sum x_i\right)\right]}$$

THIS IS FREE OF μ AND σ^2 IF AND ONLY IF $\sum x_i^2 = \sum y_i^2$ AND $\sum x_i = \sum y_i$

THEREFORE, $\sum X_i$ AND $\sum X_i^2$ ARE JOINTLY MINIMAL SUFFICIENT STATISTICS FOR μ AND σ^2.

Exercise 2

The confidence interval for the ratio of two normal population variances $\frac{\sigma_1^2}{\sigma_2^2}$ is:

$$\frac{s_1^2}{s_2^2}\frac{1}{F_{1-\frac{\alpha}{2};n_1-1,n_2-1}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2}F_{1-\frac{\alpha}{2};n_2-1,n_1-1}.$$

In the above confidence interval, s_1^2 and s_2^2 are the sample variances based on two independent samples of size n_1, n_2 selected from two normal populations $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$. Here we use the result

$$\frac{\frac{s1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}} \sim F_{n_1-1,n_2-1}.$$

Exercise 3

We start by finding the distribution of $X_p - \bar{X}$.
$E(X_p - \bar{X}) = 0$ and $Var(X_p - \bar{X}) = \sigma^2(1 + \frac{1}{n})$. The distribution of $X_p - \bar{X}$ is:

$$X_p - \bar{X} \sim N(0, \sigma\sqrt{1 + \frac{1}{n}})$$

$$Z = \frac{X_p - \bar{X}}{\sigma\sqrt{1+\frac{1}{n}}}$$

$$t = \frac{\frac{X_p - \bar{X}}{\sigma\sqrt{1+\frac{1}{n}}}}{\sqrt{\frac{\frac{(n-1)s^2}{\sigma^2}}{n-1}}} \Rightarrow t = \frac{X_p - \bar{X}}{s\sqrt{1+\frac{1}{n}}}.$$

Since the above ratio follows the t distribution with $n-1$ degrees of freedom the $1-\alpha$ prediction interval for X_p is:

$$P(-t_{\alpha/2;n-1} \leq \frac{X_p - \bar{X}}{s\sqrt{1+\frac{1}{n}}} \leq t_{\alpha/2;n-1}) = 1 - \alpha$$

Or X_p will fall in $\bar{X} \pm t_{\alpha/2;n-1}s\sqrt{1+\frac{1}{n}}$.

Exercise 5

$$f(x) = \frac{2x}{\theta} e^{-\frac{x^2}{\theta}}, \quad x > 0$$

$$L = \frac{2^n \prod x_i}{\theta^n} e^{-\frac{1}{\theta}\sum x_i^2}$$

$$L = \frac{e^{-\frac{1}{\theta}\sum x_i^2}}{\theta^n} 2^n \prod x_i$$

$$L = g(u,\theta) \cdot h(x)$$

WHERE $g(u,\theta) = \dfrac{e^{-\frac{1}{\theta}\sum x_i^2}}{\theta^n}$

AND $h(x) = 2^n \prod x_i$

THEREFORE, $U = \sum x_i^2$ IS A SUFFICIENT STATISTIC FOR θ.

Exercise 5

$$f(x) = \frac{2x}{\theta} e^{-\frac{x^2}{\theta}}$$

$$\frac{L(x_1, \ldots, x_n; \theta)}{L(y_1, \ldots, y_n; \theta)} = \frac{x_1 \cdots x_n}{y_1 \cdots y_n} e^{-\frac{1}{\theta}\left[\sum x_i^2 - \sum y_i^2\right]}$$

THIS IS FREE OF θ.

THEREFORE $\sum x_i^2$ IS A MINIMAL SUFFICIENT STATISTIC FOR θ.

Exercise 6

$\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{9} + \frac{\sigma_2^2}{12}})$. Therefore we can write:

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{9} + \frac{\sigma_2^2}{12}}}.$$

And since $\sigma_1^2 = 3\sigma_2^2$ we get:

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma_2^2(\frac{3}{9} + \frac{1}{12})}}.$$

Now we need to define a χ^2 random variable. Because X and Y are independent we have:

$$\frac{(9-1)S_1^2}{\sigma_1^2} + \frac{(12-1)S_2^2}{\sigma_2^2} \sim \chi^2_{12+9-2} \sim \chi^2_{19}.$$

Using again $\sigma_1^2 = 3\sigma_2^2$ we get:

$$\frac{\frac{1}{3}8S_1^2 + 11S_2^2}{\sigma_2^2} \sim \chi^2_{19}.$$

Now we can define a variable that has a t distribution as follows:

$$t = \frac{\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma_2^2(\frac{3}{9} + \frac{1}{12})}}}{\sqrt{\frac{\frac{\frac{1}{3}8S_1^2 + 11S_2^2}{\sigma_2^2}}{19}}} \sim t_{19}$$

Finally we get:

$$t = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{3}8S_1^2 + 11S_2^2}} \frac{\sqrt{19}}{\sqrt{\frac{3}{9} + \frac{1}{12}}} = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{3}8S_1^2 + 11S_2^2}} \sqrt{\frac{228}{5}}.$$

We can use the above t_{19} random variable to construct a 95% confidence interval for $\mu_1 - \mu_2$. We want:

$$P(-t_{\frac{\alpha}{2};19} \leq \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{3}8S_1^2 + 11S_2^2}} \sqrt{\frac{228}{5}} \leq t_{\frac{\alpha}{2};19}) = 1 - \alpha.$$

After some manipulation we find that $\mu_1 - \mu_2$ will fall in the following interval with 95% confidence:

$$\bar{X} - \bar{Y} \pm t_{\frac{\alpha}{2};19} \sqrt{\frac{8}{3}S_1^2 + 11S_2^2} \sqrt{\frac{5}{228}}.$$