Exercise 1

Given $X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ and $Y_1, \ldots, Y_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2),$ let

$$L \left( \frac{1}{\sigma^2} \sum (\bar{X} - \mu)^2 - \frac{1}{\sigma^2} \sum (\bar{Y} - \mu)^2 \right) = \frac{e^{\frac{1}{2}\sigma^2 \left[ \sum (\bar{Y} - \mu)^2 - \sum (\bar{X} - \mu)^2 \right]}}{(2\sigma^2)^{-\frac{n}{2}} \cdot e^{\frac{1}{2}\sigma^2 \left[ \sum (\bar{X} - \mu)^2 \right]}}$$

Thus, this is FREE of $\mu$ and $\sigma^2$ if and only if $\sum \bar{X} = \sum \bar{Y}$ and $\sum \bar{X} = \sum \bar{Y}.$

Therefore, $\Sigma X_i$ and $\Sigma X_i$ are jointly minimal sufficient statistics for $\mu$ and $\sigma^2.$
Exercise 2

The confidence interval for the ratio of two normal population variances $\frac{s_1^2}{s_2^2}$ is:

$$\frac{s_1^2}{s_2^2} \sim \frac{1}{\chi^2_{n_1-1,n_2-1}} \leq \frac{s_1^2}{s_2^2} \leq \frac{s_1^2}{s_2^2} F_{1-\frac{\alpha}{2},n_2-1,n_1-1}.$$

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

$$\frac{s_1^2}{s_2^2} \sim F_{n_1-1,n_2-1}.$$

Exercise 3

We start by finding the distribution of $X_p - \bar{X}$.

$E(X_p - \bar{X}) = 0$ and $V(X_p - \bar{X}) = \sigma^2(1 + \frac{1}{n})$. The distribution of $X_p - \bar{X}$ is:

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

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

$$t = \frac{X_p - \bar{X}}{s \sqrt{1 + \frac{1}{n}}} \Rightarrow t = \frac{X_p - \bar{X}}{s \sqrt{1 + \frac{1}{n}}}.$$

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

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

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

\[ L = \frac{e}{\theta^n} 2^n \prod x_i \]

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

WHERE \[ g(u, \theta) = \frac{e}{\theta^n} \]

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

Therefore, \[ U = \sum x_i \]

is a sufficient statistic for \( \theta \).
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 \( \theta \).

Therefore \( \sum x_i^2 \) is a minimal sufficient statistic for \( \theta \).
Exercise 2

\[ X - \bar{Y} \sim N(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}). \] Therefore we can write:

\[ Z = \frac{X - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}. \]

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

\[ Z = \frac{X - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2(n_1 + n_2)}}. \]

Now we need to define a \( \chi^2 \) random variable. Because \( X \) and \( Y \) are independent we have:

\[ \frac{(9 - 1)S_1^2}{\sigma_1^2} + \frac{(12 - 1)S_2^2}{\sigma_2^2} \sim \chi_{12 + 9 - 2}^2 \sim \chi_{19}^2. \]

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

\[ \frac{\frac{4}{3}8S_1^2 + 11S_2^2}{\sigma_2^2} \sim \chi_{19}^2. \]

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

\[ t = \frac{\frac{X - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\frac{4}{3}8S_1^2 + 11S_2^2}{\sigma_2^2}}}}{\sqrt{\sigma_2^2(n_1 + n_2)}} \sim t_{19}. \]

Finally we get:

\[ t = \frac{X - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\frac{4}{3}8S_1^2 + 11S_2^2}{\sigma_2^2} + \frac{11S_2^2}{12}}} = \frac{X - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{228}{5}}}. \]

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

\[ P(-t_{0.025;19} \leq \frac{X - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\frac{4}{3}8S_1^2 + 11S_2^2}{\sigma_2^2} + \frac{11S_2^2}{12}}} \leq t_{0.025;19}) = 1 - \alpha. \]

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

\[ X - \bar{Y} \pm t_{0.025;19} \sqrt{\frac{\frac{4}{3}8S_1^2 + 11S_2^2}{\sigma_2^2} + \frac{11S_2^2}{12}} \sqrt{\frac{228}{5}}. \]