1. Suppose $X_1, \ldots, X_n$ are i.i.d Poisson($\lambda$) r.v.’s.

a) The likelihood function for $\lambda$ is

$$L(\lambda) = P(X_1 = x_1, \ldots, X_n = x_n|\lambda) = \prod_{i=1}^{n} \left( \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \right) = \frac{\lambda^y e^{-n\lambda}}{\prod x_i!}, \text{ where } y = \sum_{i=1}^{n} x_i.$$  

The only function of the $X_i$’s that can’t be separated from $\lambda$ in the likelihood function is $Y = \sum X_i$, so this is a sufficient statistic for $\lambda$. In the factorization theorem, $g(y, \lambda) = \lambda^y e^{-n\lambda}$ and $h(x) = (\prod x_i!)^{-1}$.

b) To show $Y$ is sufficient using the definition of sufficiency, we need to show that $P(X_1 = x_1, \ldots, X_n = x_n|Y = y)$ does not involve $\lambda$.

$$P(X = x|Y = y) = \frac{P(X_1 = x_1, \ldots, X_n = x_n, Y = y)}{P(Y = y)} = \frac{\left( \frac{\lambda^y e^{-n\lambda}}{\prod x_i!} \right) \left( (n\lambda)^y e^{-n\lambda} \right)^{-1}}{y! \prod x_i! \left( \frac{1}{n} \right)^y, \text{ for } \sum_{i=1}^{n} x_i = y.}$$

This expression does not involve $\lambda$, which proves that $Y = \sum X_i$ is sufficient for $\lambda$. Furthermore, this is a Multinomial probability function, with each of the $y$ events is equally likely (probability $1/n$) to be associated with any of the $n X_i$’s.

2. The strategy for constructing a confidence interval is to find an interval with probability $1 - \alpha$ of containing some function of the data and the parameter of interest (a “pivot”). We then manipulate this interval to find bounds for the parameter that depend only on the data. These bounds form a random interval with probability $1 - \alpha$ of containing the parameter. Once the data have been observed, it is no longer appropriate to make a probability statement, because the parameter is not considered random, and the data have been observed and are hence no longer random. We are “confident” that the observed interval contains the parameter, because there was a small probability ($\alpha$) of getting an interval that fails to contain the parameter.

a) We know that $\bar{X} \sim N(\mu_x, \sigma_x^2/n)$, so $Z = \sqrt{n}(\bar{X} - \mu_x)/\sigma_x \sim N(0, 1)$. With $\sigma_x$ known, we can use $Z$ as a pivot to find a CI for $\mu_x$. Choosing $z^*$ such that $P(Z < z^*) = 1 - \alpha/2$ (e.g., $z^* = 1.96$ corresponds to $\alpha = 0.05$) we have $P(-z^* < Z < z^*) = 1 - \alpha$. This implies

$$P(\bar{X} - z^*\sigma_x/\sqrt{n} < \mu_x < \bar{X} + z^*\sigma_x/\sqrt{n}) = 1 - \alpha.$$ 

Therefore, having observed $\bar{X} = \bar{x}$, $\bar{x} \pm z^*\sigma_x/\sqrt{n}$ is a $1 - \alpha$ CI for $\mu_x$. 
b) We know that \( \bar{X} - \bar{Y} \sim N(\mu_x - \mu_y, \sigma_x^2/n + \sigma_y^2/m) \). So with \( \sigma_x \) and \( \sigma_y \) known,
we can use \( Z = (\bar{X} - \bar{Y} - (\mu_x - \mu_y)) / \sqrt{\sigma_x^2/n + \sigma_y^2/m} \sim N(0,1) \) as a pivot to find a CI for \( \mu_x - \mu_y \):

\[
P \left( \bar{X} - \bar{Y} - z^* \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}} < \mu_x - \mu_y < (\bar{X} - \bar{Y}) + z^* \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}} \right) = 1 - \alpha,
\]

so \( \bar{x} - \bar{y} \pm z^* \sqrt{\sigma_x^2/n + \sigma_y^2/m} \) is a \( 1 - \alpha \) CI for \( \mu_x - \mu_y \).

c) We can use the expression in 6d as a pivot to find CI’s for \( \sigma_x^2 \) or \( \sigma_y^2 \). Let \( x_l \) and \( x_u \) be the \( \alpha/2 \) and \( 1 - \alpha/2 \) quantiles of the \( \chi^2_{(n-1)} \) distribution. Then

\[
P \left( x_l < \frac{(n-1)s_x^2}{\sigma_x^2} < x_u \right) = P \left( \frac{(n-1)s_x^2}{x_u} < \sigma_x^2 < \frac{(n-1)s_x^2}{x_l} \right) = P \left( s_x \sqrt{\frac{n-1}{x_u}} < \sigma_x < s_x \sqrt{\frac{n-1}{x_l}} \right) = 1 - \alpha.
\]

For \( n = 20 \) \((n-1 = 19)\) and \( \alpha = .05 \), we have \( x_l \approx 8.91 \) and \( x_u \approx 32.85 \). With \( s_x = 10 \), a 95% CI for \( \sigma_x \) is

\[
(10 \sqrt{\frac{19}{32.85}}, 10 \sqrt{\frac{19}{8.91}}) \approx (7.6, 14.6).
\]

d) From the definition of the \( F \) distribution, we know that

\[
\frac{s_y^2/\sigma_y^2}{s_x^2/\sigma_x^2} = \frac{s_y^2}{s_x^2} \frac{\sigma_x^2}{\sigma_y^2} \sim F(m-1, n-1).
\]

Let \( x_l \) and \( x_u \) be the \( \alpha/2 \) and \( 1 - \alpha/2 \) quantiles of the \( F(m-1, n-1) \) distribution. Then

\[
P \left( x_l < \frac{s_y^2/\sigma_y^2}{s_x^2/\sigma_x^2} < x_u \right) = P \left( x_l \frac{s_y^2}{s_x^2} < \frac{\sigma_y^2}{\sigma_x^2} < x_u \frac{s_y^2}{s_x^2} \right) = P \left( \sqrt{x_l} \frac{s_y}{s_x} < \frac{\sigma_y}{\sigma_x} < \sqrt{x_u} \frac{s_y}{s_x} \right) = 1 - \alpha.
\]

With \( m = 25, n = 20 \) and \( \alpha = 0.05 \), we have \( x_l \approx 0.43 \) and \( x_u \approx 2.45 \). For \( s_x = 10 \) and \( s_y = 15 \), a 95% CI for \( \sigma_x/\sigma_y \) is

\[
\left( \sqrt{0.43} \frac{10}{15}, \sqrt{2.45} \frac{10}{15} \right) \approx (0.44, 1.04).
\]

This interval contains the value 1.0, so it is plausible that \( \sigma_x = \sigma_y \).