April 5, 2010

Class # 17: Confidence Intervals for the Mean.

Reading: Section 7.1 of Devore.

Calculating a Confidence Interval: Assume that we randomly sample \( n \) subjects from a population that is normally distributed. That is, we collect observations \( X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2) \). If this is the case, it follows that

\[
\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \implies \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1).
\]

This last feature of the sample mean allows us to construct a \( 100(1 - \alpha)\% \) confidence interval for \( \mu \).

This \( 100(1 - \alpha)\% \) confidence interval is calculated as

\[
\bar{X} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}},
\]

where \( z_{\frac{\alpha}{2}} \) is the number such that \( P(Z \leq z_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2} \).

Interpretation of a Confidence Interval: Once the confidence interval has been calculated, no meaningful probabilistic statements can be made about whether \( \mu \) is in there or not. It either is or it isn’t.

Sample Size Considerations: The width of a confidence interval is \( w = 2z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \). The sample size necessary for a desired width is calculated as

\[
n = \left(2z_{\frac{\alpha}{2}} \frac{\sigma}{w}\right)^2.
\]

Homework: These problems are due at the beginning of class on 4/7/10.

1. Devore, page 218, #46.
3. Devore, page 218, #50.
4. Devore, page 218 #54.
5. Devore, page 218, #55.

Homework Solutions:

1. Recall that \( X \) and \( Y \) are independent if \( f(x, y) = f(x)f(y) \). In this case, \( f(x) = \int_{0}^{\infty} f(x, y)dy = \int_{0}^{\infty} 2xe^{-y}dy = 2x \int_{0}^{\infty} e^{-y}dy = 2x \) and \( f(y) = \int_{0}^{1} f(x, y)dx = e^{-y} \). So \( f(x, y) = f(x)f(y) \), and \( X \) and \( Y \) are independent.

2. (a) The joint pdf of \( X \) and \( Y \) is \( p(x, y) = p(x)p(y) \) where

\[
p(x) = \begin{cases} 
.1 & x = 0 \\
.2 & x = 1 \\
.4 & x = 2 \\
.3 & x = 3 
\end{cases}
\]

and

\[
p(y) = \begin{cases} 
.1 & y = 0 \\
.2 & y = 1 \\
.4 & y = 2 \\
.3 & y = 3 
\end{cases}
\]
(b) \[ P(X = Y) = p(0, 0) + p(1, 1) + p(2, 2) + p(3, 3) = p(0)p(0) + p(1)p(1) + p(2)p(2) + p(3)p(3) = \frac{1^2}{2} + \frac{.2^2}{.4} + \frac{.3^2}{.3} = .3 \]

(c) \[ P(X > Y) = p(0, 1) + p(0, 2) + p(0, 3) + p(1, 2) + p(1, 3) + p(2, 3) = .1 \times .2 + .1 \times .4 + .1 \times .3 + .2 \times .4 + .2 \times .3 + .4 \times .3 = .35 \]

3. (a) If \( X \) and \( Y \) are independent, then \((X, Y) \sim f(x, y)\) where \( f(x, y) = f(x)f(y) \). Since \( X \) and \( Y \) are uniformly distributed between 5 and 6, \( f(x) = 1 \) for \( 5 \leq x \leq 6 \) and \( f(y) = 1 \) for \( 5 \leq y \leq 6 \). Therefore, \( f(x, y) = 1 \) for \( 5 \leq x, y \leq 6 \).

(b) \[ P(5.25 \leq X \leq 5.75 \text{ and } 5.25 \leq Y \leq 5.75) = \int_{5.25}^{5.75} \int_{5.25}^{5.75} 1 \, dx \, dy = .25 \]

(c) This is very similar to the problem done in class. In this case, \( P(A) = 1 - P(A^c) \) where \( P(A^c) = \left[ \left( \frac{5}{6} \right)^2 / 2 \right] \times 2 = \frac{25}{36} \). Therefore, \( P(A) = \frac{11}{36} \).

4. (a) \[ E(X + Y) = \sum_{x,y} (x + y)p(x, y) = 0 \times .02 + 5 \times .06 + \cdots + 25 \times .01 = 14.1 \]

(b) \[ E(\max(X, Y)) = 0 \times .02 + 5 \times .06 + \cdots + 15 \times .01 = 9.6 \]

5. (a) \[ \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 44.25 - 5.55 \times 8.55. \] These numbers are calculated using \( E(XY) = \sum_{x,y} xy p(x, y) = 25 \times .15 + 150 \times .01 = 44.25 \), \( E(X) = \sum_{x} xp(x) = 0 \times .2 + 5 \times .49 + 10 \times .31 = 5.55 \), and \( E(Y) = \sum_{y} yp(y) = 0 \times .07 + \cdots + 15 \times .21 = 8.55 \)

(b) \[ \rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{-3.2025}{\sqrt{144.25 \times 144.75}} = -.207 \]