Econ 312: Problem Set #2

Due:Thursday, March 11

Question.1 A manufacturer claims that a certain brand of VCR player has an average life expectancy of 5 years and 6 months with a standard deviation of 1 year and 6 months. Assume that the life expectancy is normally distributed.

(a) Selecting one VCR player from this brand at random, calculate the probability of its life expectancy exceeding 7 years.

Answer: $\Pr(Y > 7)$=$\Pr(Z > \frac{7-5.5}{1.5})$=$\Pr(Z > 1)$=0.1587

(b) The Critical Consumer magazine decides to test fifty VCRs of this brand. The average life in this sample is 6 years and the sample standard deviation is 2 years. Calculate a 99% confidence interval for the average life.

Answer: $6 \pm 2.58 \times \frac{2}{\sqrt{50}} = 6 \pm 0.73 = [5.27, 6, 73]$

(c)How many more VCRs would the magazine have to test in order to halve the width of the confidence interval?

Answer: $\frac{1}{2}\text{Width} = \frac{1}{2}(2.58 \times \frac{2}{\sqrt{50}}) = 2.58 \times \frac{2}{\sqrt{4\times50}}$, so, $n = 200$

Question.2 IQs of individuals are normally distributed with a mean of 100 and a standard deviation of 16. If you sampled students at your college and assumed, as the null hypothesis, that they had the same IQ as the population, then in a random sample of size

(a) n = 25, find $\Pr(\bar{Y} < 105)$.

Answer: $\Pr(\bar{Y} < 105) = \Pr(Z < \frac{105-100}{16/\sqrt{25}}) = \Phi(1.5625) = 0.9409$.

(b) n = 100, find $\Pr(\bar{Y} > 97)$.

Answer: $\Pr(\bar{Y} > 97) = \Pr(Z > \frac{97-100}{16/\sqrt{100}}) = \Pr(Z > -1.875) = \Pr(Z < 1.875) = \Phi(1.875) = 9696$.

(c) n = 144, find $\Pr(101 < \bar{Y} < 103)$.

$\Pr = \Pr(101 < \bar{Y} < 103) = \Pr(\frac{101-100}{16/\sqrt{144}} < Z < \frac{103-144}{16/\sqrt{144}}) = \Pr(0.75 < Z < 2.25) = \Phi(2.25) - \Phi(0.75) =$ .2144.

Question.3 Consider the following alternative estimator for the population mean: .

$$\tilde{Y} = \frac{1}{n}(\frac{1}{4}Y_1 + \frac{7}{4}Y_2 + \frac{1}{4}Y_3 + \frac{7}{4}Y_4 + \cdots + \frac{1}{4}Y_{n-1} + \frac{7}{4}Y_n)$$

Prove that $\tilde{Y}$ is unbiased and consistent, but not efficient when compared to $\bar{Y}$.

Answer:

$E(\tilde{Y}) = \frac{1}{n}(\frac{1}{4}E(Y_1) + \frac{7}{4}E(Y_2) + \frac{1}{4}E(Y_3) + \frac{7}{4}E(Y_4) + \cdots + \frac{1}{4}E(Y_{n-1}) + \frac{7}{4}E(Y_n)) = \frac{1}{n}\mu_Y(1/4 + 7/4 + 1/4 + 7/4 + \cdots + 1/4 + 7/4) = \frac{n}{n}\mu_Y = \mu_Y$. Hence, $\tilde{Y}$ is unbiased.

$Var(\tilde{Y}) = \frac{1}{n62}((\frac{1}{4})^2 Var(Y_1) + (\frac{7}{4})^2 Var(Y_2) + (\frac{1}{4})^2 Var(Y_3) + (\frac{7}{4})^2 Var(Y_4) + \cdots + (\frac{1}{4})^2 Var(Y_{n-1}) + (\frac{7}{4})^2 Var(Y_n)) = \frac{1}{n^2}\sigma_Y(1/16 + 49/16 + 1/16 + 49/16 + \cdots + 1/16 + 49/16) = \frac{1}{n^2}\frac{n}{2}(1/16 + 49/16)\sigma_Y = 1.5625\frac{\sigma_Y}{n}$.

Since $Var(\tilde{Y}) \to 0$ as $n \to \infty$, $\tilde{Y}$ is consistent. $\tilde{Y}$ has a larger variance than $\bar{Y}$ and is therefore not as efficient.

Question.4 You have obtained a sub-sample of 1744 individuals from the Current Population Survey (CPS) and are interested in the relationship between weekly earnings and age. The regression yielded the following result:

$$\widehat{Earn} = 239.16 + 5.20 \times Age, \quad R^2 = 0.05, SER = 287.21$$

where $Earn$ and Age are measured in dollars and years respectively.

(a) Interpret the results.

Answer: A person who is one year older increases her weekly earnings by $5.20. There is no meaning attached to the intercept. The regression explains 5 percent of the variation in earnings.

(b) Is the effect of age on earnings large?

Answer: Assuming that people worked 52 weeks a year, the effect of being one year older translates into an additional $270.40 a year. This does not seem particularly large in 2002 dollars, but may have been earlier.

(c) Why should age matter in the determination of earnings? Do the results suggest that there is a guarantee for earnings to rise for everyone as they become older? Do you think that the relationship between age and earnings is linear?

Answer: In general, age-earnings profiles take on an inverted U-shape. Hence it is not linear and the linear approximation may not be good at all. Age may be a proxy for “experience” which in itself can approximate “on the job training.” Hence the positive effect between age and earnings. The results do not suggest that there is a guarantee for earnings to rise for everyone as they become older since the regression R^2 does not equal 1. Instead the result holds “on average.”

Question.5 At a recent county fair, you observed that at one stand people’s weight was forecasted, and were surprised by the accuracy (within a range). Thinking about how the person could have predicted your weight fairly accurately (despite the fact that she did not know about your "heavy bones"), you think about how this could have been accomplished. You remember that medical charts for children contain 5%, 25%, 50%, 75% and 95% lines for a weight/height relationship and decide to conduct an experiment with 110 of your peers. You collect the data and calculate the following sums:

$$\sum_{i=1}^{110} Y_i = 17,375, \sum_{i=1}^{110} X_i = 7,665.5$$

$$\sum_{i=1}^{110}(Y_i - \bar{Y})^2 = 94,228.8, \sum_{i=1}^{110}(X_i - \bar{X})^2 = 1,248.9, \sum_{i=1}^{110}(X_i - \bar{X})(Y_i - \bar{Y}) = 7,625.9$$

(a) Calculate the slope and intercept of the regression and interpret these.

Answer:

$\hat{\beta}_1 = \frac{\sum_{i=1}^{110}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{110}(X_i - \bar{X})^2} = \frac{7,625.9}{1,248.9} = 6.1060934$

and

$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = \frac{17,375}{110} - 6.1060934 \times \frac{7,665.5}{110} = -267.5569$

(b) Find the regression R^2 and explain its meaning. What other factors can you think of that might have an influence on the weight of an individual?

$R^2 = ESS/TSS = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = \frac{\sum_{i=1}^{n}(\hat{\beta}_0 + \hat{\beta}_1 X_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = \frac{\sum_{i=1}^{n}(\bar{Y} - \hat{\beta}_1 \bar{X} + \hat{\beta}_1 X_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = \frac{\hat{\beta}_1^2 \sum_{i=1}^{n}(X_i - \bar{X})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = \frac{6.1060934^2 \times 1,248.9}{94,228.8} =$ 0.49416375

The regression R^2 indicates that about fifty percent of the variation in weight is explained by the model or height.