Homework #1 (Solutions)

STAT-UB.0003: Regression and Forecasting Models

For this assignment, you will need to use a z table and a t table. These are available in Appendix D of your textbook, as Tables 2 and 3.

Solutions adapted from N.S. Boudreau's *Instructor's Solution Manual* (2011).

1. McClave, Benson, & Sincich (MBS), Ex. 5.26

Solution:

(a)

$$E(\bar{X}) = \mu_{\bar{X}} = \mu = .10$$
$$\text{var}(\bar{X}) = \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} = \frac{(.10)^2}{50} = .0002$$

(b) Since $n > 30$, the Central Limit Theorem implies that $\bar{X}$ is approximately normal.

(c) First, compute $\sigma_{\bar{X}} = \sqrt{(.0002)} = .0141$. Now,

$$\begin{aligned} P(\bar{X} > .13) &= P\left(\frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} > \frac{.13 - (.10)}{(.0141)}\right) \\ &= P(Z > 2.13) \\ &= .0166 \end{aligned}$$

2. MBS, Ex. 5.32.

Solution:

(a) By the Central Limit Theorem, the sampling distribution of $\bar{X}$ is approximately normal with mean $\mu_{\bar{X}} = \mu$ and standard deviation $\sigma_{\bar{X}} = \sigma/\sqrt{n} = \sigma/\sqrt{100}$.

(b) The mean of the $\bar{X}$ distribution is equal to the mean of the distribution of the fleet, the fleet mean score.

(c) $\mu_{\bar{X}} = \mu = 30$ and $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 60/\sqrt{100} = 6$.

$$\begin{aligned} P(\bar{X} \geqslant 45) &= P(Z \geqslant \frac{45 - 30}{6}) \\ &= P(Z \geqslant 2.5) \\ &= .0062. \end{aligned}$$

(d) The sample mean of 45 tends to refute the claim. If the true fleet mean were as high as 30, observing a sample mean of 45 or higher would be extremely unlikely (probability .0062). Thus, we would infer that the true mean is actually not 30, but something higher. We would refute the company's claim that the mean "couldn't possibly be as large as 30."

3. MBS, Ex. 6.12.

Solution:

(a) $(66.350, 69.160)$

(b) Since n is large, we can use a z-based 90% confidence interval:

$$\begin{aligned}\bar{x} \pm 1.645\frac{s}{\sqrt{n}} &= 67.755 \pm 1.645\frac{26.871}{\sqrt{992}} \\ &= 67.755 \pm 1.403 \\ &= (66.355, 69.155)\end{aligned}$$

The discrepancy is due to using a z-based interval instead of a t-based interval. Basically, the intervals agree.

(c) We can be 90% confident that the true mean lies in this interval.

(d) This claim seems implausible given that 75 is very far outside the confidence interval (more than 8 standard errors away from the sample mean).

4. MBS, Ex. 6.27, parts (a–b). You can use Minitab, Excel, or a calculator to compute the mean and standard deviation.

Solution:

First, we compute $\bar{x} = 5$ and $s = \sqrt{5.2} = 2.2804$.

(a) For confidence level .90, $\alpha = .10$ and $\alpha/2 = .05$. With $n-1 = 6-1 = 5$ degrees of freedom, $t_{.05} = 2.015$. The 90% confidence interval is

$$\begin{aligned}\bar{x} \pm t_{.05}\frac{s}{\sqrt{n}} &= 5 \pm 2.015\frac{2.2804}{\sqrt{6}} \\ &= 5 \pm 1.88 \\ &= (3.12, 6.88)\end{aligned}$$

(b) The 95% confidence interval is

$$\begin{aligned}\bar{x} \pm t_{.025}\frac{s}{\sqrt{n}} &= 5 \pm 2.571\frac{2.2804}{\sqrt{6}} \\ &= 5 \pm 2.39 \\ &= (2.61, 7.39)\end{aligned}$$

5. Obtain the "NormTemp" dataset from the course website. This gives data on body temperatures for 130 randomly selected human subjects.

   (a) What is a reasonable population for this dataset?

   Solution: There are many possible valid answers to this question. One such answer is "the body temperatures of all people alive today."

   (b) Using Minitab, get a confidence interval for the population mean temperature. To do this, first read the data set into Minitab, and then use

   Stat ⇒ Basic Statistics ⇒ 1-Sample t

   The variable you need to use is *Temp*. Ask for a confidence interval with level 95.0. Copy and paste the Minitab output.

   Solution:
   The 95% confidence interval is $(98.122, 98.377)$.

   (c) What assumptions do you need for the confidence interval to be valid?

   Solution:
   We need the 130 randomly selected subjects to have been selected independently from the population; we need for the sample to be an unbiased random sample. (The central limit theorem is in force since the sample size is greater than 30.)

   (d) Are the results of the confidence interval surprising, in view of the fact that the population mean temperature is supposed to be 98.6 degrees?

   Solution:
   The confidence interval is surprising because it means that we are 95% confident that the population mean temperature is between 98.12 to 98.37 degrees whereas it is supposed to be 98.6 degrees.