STAT 201 - Week-In-Review 9

Dr. Prasenjit Ghosh

Problem Solutions

1. Starbucks wants to know how much caffeine is in their coffee for a future marketing campaign, so the company takes a random sample of 64 coffees from their stores and takes them to a lab to test the amount of caffeine in each cup. From a previous study, they know the true mean of the caffeine in coffee is 190 mg with standard deviation 40 mg.

   (A) Can we apply the Central Limit Theorem to this problem? Why?
   - (a) Yes, because the question states that the population distribution is Normal.
   - (b) No, because the question does not state that the population distribution is Normal.
   - (c) *** Yes, because the random sample size is large enough.
   - (d) No, because the random sample size is not large enough.
   - (e) None of the above

   (B) What is the sampling distribution of $\bar{X}$?
   - (a) *** It is Normal with parameters $\mu_{\bar{X}} = 190$ mg and $\sigma_{\bar{X}} = \frac{40}{\sqrt{64}}$ mg
   - (b) It is Normal with parameters $\mu_{\bar{X}} = \frac{190}{\sqrt{64}}$ mg and $\sigma_{\bar{X}} = \frac{40}{\sqrt{64}}$ mg
   - (c) It is skewed with parameters $\mu_{\bar{X}} = \frac{190}{\sqrt{64}}$ mg and $\sigma_{\bar{X}} = \frac{40}{\sqrt{64}}$ mg
   - (d) It is skewed with parameters $\mu_{\bar{X}} = 190$ mg and $\sigma_{\bar{X}} = \frac{40}{\sqrt{64}}$ mg
   - (e) It is unknown

   (C) What is the probability that the sample mean is between 180 mg to 200 mg?
   - (a) 0.0455
   - (b) 0.1974
   - (c) 0.8026
   - (d) *** 0.9545 $\mathrm{P}(180 \le \bar{X} \le 200) = \mathrm{P}(Z \le 2) - \mathrm{P}(Z \le -2) = 0.97725 - 0.02275$
   - (e) 0.9750

   (D) What amount of caffeine corresponds to the 3rd quartile of the distribution of the sample mean? Report your answer to one decimal place.
   - (a) 186.6 mg
   - (b) 190 mg
   - (c) 195.2 mg
   - (d) 184.8 mg
   - (e) *** 193.4 mg $\bar{x}_{0.25} = \mu_{\bar{x}} + \sigma_{\bar{x}} z_{0.25} = 180 + \frac{40}{\sqrt{64}}(0.67449)$

2. It is known that the resistance of carbon resistors is approximately normally distributed with $\mu = 1200$ ohms and $\sigma = 120$ ohms. If 10 resistors are randomly selected from a shipment, what is the probability (up to four decimal places) that the average resistance will be more than 1250 ohms?

(a) 0.33742
(b) 0.66276
(c) 0.76528
(d) *** 0.09342 $P(\bar{X} > 1250) = 1 - P(Z \leq \frac{1250-\mu}{\sigma/\sqrt{n}}) \approx 1 - P(Z \leq 1.32) = 1 - 0.90658$
(e) 0.90658

3. The distribution of total body protein in healthy adult men is Normal with a mean of 12.3 Kg and a standard deviation of 0.4 Kg. If you take a random sample of 4 healthy adult men, what is the probability that their mean total body protein is between 12.2 Kg and 12.4 Kg?

(a) 0.19742
(b) *** 0.38292 $P(12.2 \leq \bar{X} \leq 12.4) = P(Z \leq 0.5) - P(Z \leq -0.5) = 0.69146 - 0.30854$
(c) 0.61708
(d) 0.80258
(e) We cannot calculate this probability because the sample size is too small

4. Let $\bar{X}$ be the mean of a random sample of size 16, drawn from a population with mean 224 and standard deviation 80.

(A) Find the mean and standard deviation of $\bar{X}$

(a) $\mu_{\bar{X}} = 2.24$, $\sigma_{\bar{X}} = 0.2$
(b) $\mu_{\bar{X}} = 224$, $\sigma_{\bar{X}} = 80$
(c) *** $\mu_{\bar{X}} = 224$, $\sigma_{\bar{X}} = 20$ $\mu_{\bar{X}} = \mu = 224$, and $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{80}{\sqrt{16}}$
(d) $\mu_{\bar{X}} = 22400$, $\sigma_{\bar{X}} = 2$
(e) $\mu_{\bar{X}} = 2240$, $\sigma_{\bar{X}} = 80$

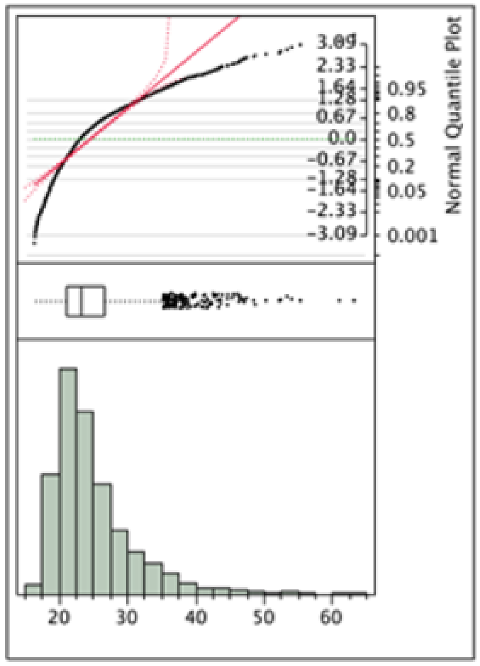
(B) Find the probability that $\bar{X}$ assumes a value smaller than 216.

(a) 0.16
(b) 0.84
(c) 0.95
(d) 0.05
(e) *** We don't know the distribution of $\bar{X}$

5. Random samples of size n were selected from a population with a known standard deviation. How is the standard deviation of the sampling distribution of the sample mean affected if the sample size is increased from 50 to 200?

(a) It remains the same.
(b) It is multiplied by four
(c) It is divided by four
(d) It is multiplied by two
(e) *** It is divided by two $\sigma_{\bar{X},1} = \frac{\sigma}{\sqrt{n}_1} = \frac{\sigma}{\sqrt{50}}$, and $\sigma_{\bar{X},2} = \frac{\sigma}{\sqrt{n}_2} = \frac{\sigma}{\sqrt{200}} \Rightarrow \sigma_{\bar{X},2} = \frac{\sigma_{\bar{X},1}}{2}$

6. A random sample of 2610 people aged 40 and above was selected from a general population of all Americans aged 40 and over. The body mass index (BMI) was measured on each individual. Below is a plot of this data set.



What can we say about the shape of the distribution of BMI for the population of all Americans aged 40 and over and the shape of the sampling distribution of $\bar{X}$ for samples of size 100 from this population?

(a) We expect the population distribution and the sampling distribution of $\bar{X}$ for samples of size 100 from this population to be right skewed.

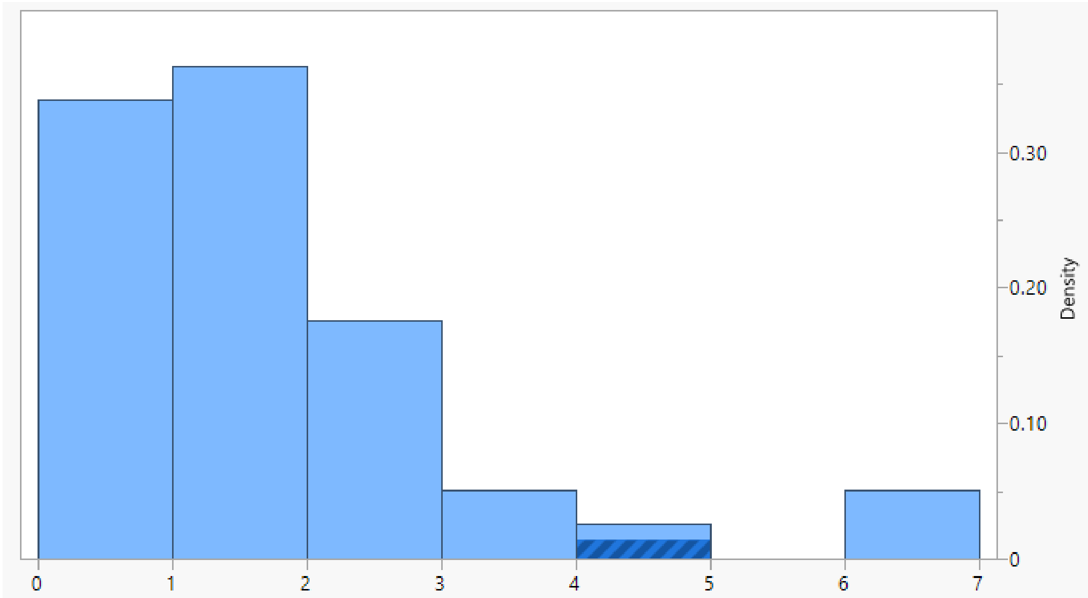
(b) We expect the population distribution and the sampling distribution of $\bar{X}$ for samples of size 100 from this population to be approximately normal.

(c) We expect the population distribution to be approximately normal and the sampling distribution of $\bar{X}$ for samples of size 100 from this population to be right skewed.

(d) *** We expect the population distribution to be right skewed and the sampling distribution of $\bar{X}$ for samples of size 100 from this population to be approximately normal.

(e) We don't know anything about the population distribution from a sample of only 2610 individuals randomly selected from the population and the sampling distribution of $\bar{X}$ for samples of size 100 from this population to be approximately normal.

7. Suppose, the following histogram is based on 80 observations of mass loss from a laboratory study of corrosion of steel embedded in a low-strength concrete.



Assume that the population distribution of mass loss (corrosion) has a mean of 1.765 and a standard deviation of 1.414 (both measured in appropriate units).

(A) What can we say about the distribution of the mass losses (corrosion) in the sample?

(a) The distribution of mass loss (corrosion) is heavily skewed towards the left with a mean of 1.765 and a standard deviation of 1.414.

(b) The distribution of mass loss (corrosion) is bell-shaped that is heavily skewed towards the right with mean 1.765 and standard deviation $\frac{1.414}{\sqrt{80}}$.

(c) *** The distribution of mass loss (corrosion) is heavily skewed towards the right with a mean of 1.765 and a standard deviation of 1.414.

(d) The distribution of mass loss (corrosion) is bell-shaped that is symmetric with respect to its mean of 1.765 and standard deviation of 1.414.

(e) The distribution of mass loss (corrosion) is bell-shaped that is symmetric with respect to its mean 1.765 and standard deviation $\frac{1.414}{\sqrt{80}}$.

(B) What is the probability that the mass loss (corrosion) lies between 1.6 and 1.9? Answer up to four decimal places. Select the nearest option.

(a) 0.08159
(b) 0.08359
(c) 0.08559
(d) 0.08759
(e) *** Can't be determined from the given information.

(C) What would be the distribution of the average mass loss (corrosion) in the sample?

(a) The average mass loss in the sample has a bell-shaped distribution that is heavily skewed towards the right with a mean of 1.765 and a standard deviation of 1.414.

(b) The average mass loss in the sample has a bell-shaped distribution that is heavily skewed towards the right with mean 1.765 and standard deviation $\frac{1.414}{\sqrt{80}}$.

(c) The average mass loss in the sample has a bell-shaped distribution that is heavily skewed towards the left with mean 1.765 and standard deviation $\frac{1.414}{\sqrt{80}}$.

(d) The average mass loss in the sample has a bell-shaped distribution that is symmetric with respect to its mean of 1.765 and standard deviation of 1.414.

(e) *** The average mass loss in the sample has a bell-shaped distribution that is symmetric with respect to its mean 1.765 and standard deviation $\frac{1.414}{\sqrt{80}}$.

(D) What is the probability that the average mass loss (corrosion) in the sample lies between 1.6 and 1.9? Answer up to four decimal places. Select the nearest option.

(a) 0.61317
(b) 0.63317
(c) *** 0.65317 $P(16 \leq \bar{X} \leq 1.9) = P(Z \leq 0.85) - P(Z \leq -1.04) = 0.80234 - 0.14917$
(d) 0.67317
(e) Can't be determined from the given information.

8. Suppose that the mean outstanding credit card balance for all young couples in the US is 650 USD with a standard deviation of 420 USD. The distribution is highly skewed to the right. Suppose now that one random sample of size 200 is selected from the population of young couples. The mean of this sample is 600 USD.

(A) Based on the above information, which of the following statements is true?

(a) The shape of the distribution of this one random sample is right-skewed with a mean of 650 USD and the shape of the sampling distribution of the sample mean $\overline{X}$ for samples of size 200 is approximately normal with a mean of 650 USD.

(b) *** The shape of the distribution of this one random sample is right-skewed with a mean of 600 USD and the shape of the sampling distribution of the sample mean $\overline{X}$ for samples of size 200 is approximately normal with a mean of 650 USD.

(c) The shape of the distribution of this one random sample is approximately normal with a mean of 600 USD and the shape of the sampling distribution of the sample mean $\overline{X}$ for samples of size 200 is approximately normal with a mean of 650 USD.

(d) The shape of the distribution of this one random sample is right-skewed with a mean of 600 USD and the shape of the sampling distribution of the sample mean $\overline{X}$ for samples of size 200 is approximately normal with a mean of 600 USD.

(e) The shape of the distribution of this one random sample is right-skewed with a mean of 600 USD and the shape of the sampling distribution of the sample mean $\overline{X}$ for samples of size 200 is also right-skewed but with a mean of 650 USD.

(B) Decide whether observing a future random sample of size 200 drawn from the given population distribution with the sample mean credit card balance less than or equal to its present realization is rare.

Solution: Let $\bar{X}$ be the sample mean credit card balance based a random sample of size n = 200. Then

$$\bar{X} \overset{a}{\sim} N\left(\mu, \frac{\sigma}{\sqrt{n}}\right), \text{ with } \mu = 650, \text{ and } \sigma = 420.$$

Hence,

$$\mathrm{P}(\bar{X} \leq 600) \approx \mathrm{P}\left(Z \leq \frac{600 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) \approx \mathrm{P}(Z \leq -1.68) = 0.04648,$$

which is small. Therefore, observing a future random sample of size 200 drawn from the given population distribution with the sample mean credit card balance less than or equal to its present realization of 600 USD is rare.

(C) What is the probability that the sample mean credit card balance of a random sample of size 200 is larger than 680 USD?

Solution: The required probability is:

$$\mathrm{P}(\bar{X} > 680) = \mathrm{P}\left(Z > \frac{680 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) \approx \mathrm{P}(Z > 1.01) = 1 - \mathrm{P}(Z \leq 1.01) = 1 - 0.84375 = 0.15625$$

9. The drying times for a certain type of cement are normally distributed with a standard deviation of 62 minutes. A researcher wishes to estimate the mean drying time for this type of cement. Find the least sample size n needed to assure with 90% confidence that the sample mean will not differ from the population mean by more than 5 minutes.

(a) 362
(b) 416
(c) *** 417
(d) 490
(e) 591

Explanation: Need to find the least value of n that satisfies $P(|\bar{X}-\mu| \leq 5) \geq 0.90$.

Observe that:

$$
\begin{aligned}
P(|\bar{X}-\mu| \leq 5) &= P\left(|Z| \leq \frac{5}{\sigma/\sqrt{n}}\right) = 2P\left(Z \leq \frac{5}{\sigma/\sqrt{n}}\right) - 1 \geq 0.90 \\
&\Rightarrow 2P\left(Z \leq \frac{5}{\sigma/\sqrt{n}}\right) \geq 1.90 \\
&\Rightarrow P\left(Z \leq \frac{5}{\sigma/\sqrt{n}}\right) \geq 0.95 = P(Z \leq 1.64485) \\
&\Rightarrow \frac{5}{\sigma/\sqrt{n}} \geq 1.64485 \\
&\Rightarrow n \geq \left(\frac{1.64485\sigma}{5}\right)^2 \approx 416.0025269 \text{ rounded up to } 417.
\end{aligned}
$$

10. Suppose that 65% of all college women have been on a diet within the last 6 months. A survey is planned to interview a simple random sample of 100 college women if they were on a diet within the last 6 months.

(A) What is the probability that 70% or more of the women in the sample have been on a diet in the last 6 months?

Solution: Let $\hat{p}$ be the sample proportion of women who have been on a diet in the last 6 months. Then

$$\hat{p} \overset{a}{\sim} N\left(0.65, \sqrt{\frac{0.65 * 0.35}{100}}\right).$$

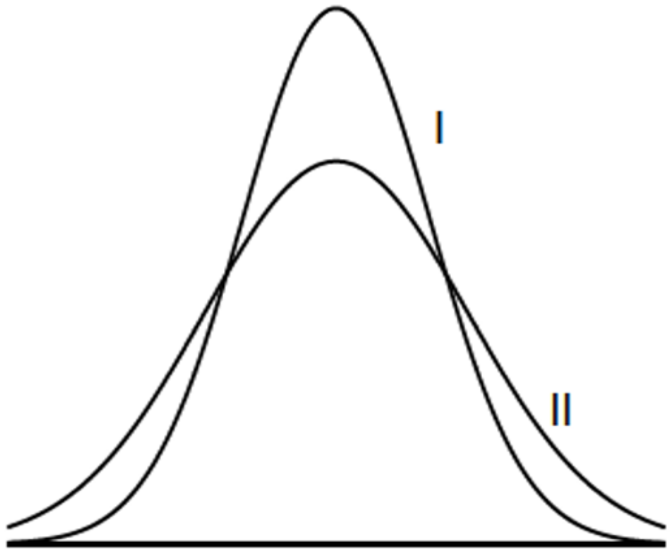
Hence, the required probability is:

$$
\begin{aligned}
P(\hat{p} \geq 0.70) \approx P\left(Z \geq \frac{0.70 - 0.65}{\sqrt{\frac{0.65 * 0.35}{100}}}\right) &= P(Z \geq 1.04828483672) \\
&\approx P(Z \geq 1.05) \\
&= 1 - P(Z < 1.05) \\
&= 1 - 0.85314 \\
&= 0.14686
\end{aligned}
$$

(B) What is the probability that less than 70% of the women in the sample have been on a diet in the last 6 months?

Solution: $P(\hat{p} < 0.70) = 1 - P(\hat{p} \geq 0.70) = 1 - 0.14686 = 0.85314$

11. The following graph shows the sampling distributions of two sample proportions based on two different sample sizes n_1 and n_2 drawn from the same population. What can we conclude about n_1 and n_2? Hint: Think in terms of the spread.



(a) $n_1 < n_2$.

(b) *** $n_1 > n_2$.

(c) $n_1 = n_2$.

(d) $n_1 \neq n_2$, but we do not have adequate information to conclude which one is larger.

(e) Something is wrong. If the population is the same in both cases, the sampling distributions should look the same.

12. Historically, 51% of voters in a certain state voted for a Republican candidate as state governor. A new governor election is coming up and a survey of randomly selected 100 voters from this state will be conducted. What is the probability that more than 55% will vote for the Republican candidate? Choose the closest answer.

(a) 0.14457

(b) *** 0.21186 $P(\hat{p} \geq 0.55) \approx P\left(Z \geq \frac{0.55-0.51}{\sqrt{\frac{0.51*0.49}{100}}}\right) \approx P(Z \geq 0.8) = 1 - 0.78814$

(c) 0.46182

(d) 0.53188

(e) 0.78814

13. Harley-Davidson motorcycles make up 16% of all the motorcycles registered in the United States. You plan to interview a simple random sample of 500 motorcycle owners. Without loss of generality, assume the selection of the owners to be independent of each other.

(A) Find the expected number of owners in your sample who own Harleys?

Solution: Let $\hat{p}$ be the sample proportion of owners in a sample of size $n = 500$ who own Harleys.

Then, $n\hat{p}$ is the number of owners in the sample who own Harleys.

According to the problem, the population proportion of owners who own Harleys is $p = 0.16$.

Hence the required expected number is:

$$\mathrm{E}(n\hat{p}) = n\mathrm{E}(\hat{p}) = np = (500)(0.16) = 80.$$

(B) Assume that, in your sample, you observed that 17% of the owners own Harleys. If you are asked to find the probability that the sample percentage of owners who own Harleys based on a future sample of 500 owners drawn from the same population is 17% or more, what sampling distribution would you use to compute the required probability?

(a) $\hat{p} \overset{a}{\sim} N\left(\mu_{\hat{p}} = 0.17,\ \sigma_{\hat{p}} = \sqrt{\frac{0.17 * 0.83}{500}}\right)$

(b) $\hat{p} \overset{a}{\sim} N\left(\mu_{\hat{p}} = 0.16,\ \sigma_{\hat{p}} = \sqrt{\frac{0.17 * 0.83}{500}}\right)$

(c) $\hat{p} \overset{a}{\sim} N\left(\mu_{\hat{p}} = 0.17,\ \sigma_{\hat{p}} = \sqrt{\frac{0.16 * 0.84}{500}}\right)$

(d) *** $\hat{p} \overset{a}{\sim} N\left(\mu_{\hat{p}} = 0.16,\ \sigma_{\hat{p}} = \sqrt{\frac{0.16 * 0.84}{500}}\right)$

(e) The sampling distribution of $\hat{p}$ is not approximately normal since the large sample size criteria are not satisfied.

(C) What would be the probability asked to evaluate in part (B) of this problem.

Solution: The required probability is:

$$\begin{aligned}
\mathrm{P}(\hat{p} \geq 0.17) &\approx \mathrm{P}\left(Z \geq \frac{0.17 - 0.16}{\sqrt{\frac{0.16 * 0.84}{500}}}\right) \\
&\approx \mathrm{P}(Z \geq 0.61) \\
&= 1 - \mathrm{P}(Z < 0.61) \\
&= 1 - 0.72907 \\
&= 0.27093
\end{aligned}$$