STAT 2290 Homework 5

Hand in on April 8 or 9 with your quiz.

Read these instructions:

- Show your work. Answers without proper justifications receive no credit.

Sample proportions

Problem 1.
The Transportation Security Administration (TSA) is responsible for airport safety. On some flights, TSA officers randomly select passengers for an extra security check before boarding. One such flight had 76 passengers—12 in first class and 64 in coach class. TSA officers selected an SRS of 10 passengers for screening. Let $\hat{p}$ be the proportion of first-class passengers in the sample.
(a) Is the 10% condition met in this case? Justify your answer.
(b) Is the Large Counts condition met in this case? Justify your answer.

Answer
(a) No: here the sample size is $n = 10$ out of a population size of 76, but $76 \ngeq 10 \times 10 = 100$.
(b) Since $\hat{p}$ and $1 - \hat{p}$ are proportions, they are numbers between 0 (=0%) and 1 (=100%). The Large Counts condition forces $n\hat{p} = 10\hat{p} \geq 10 \Rightarrow \hat{p} \geq 1 \Rightarrow \hat{p} = 1$ and also $n(1 - \hat{p}) = 10(1 - \hat{p}) \geq 10 \Rightarrow 1 - \hat{p} \geq 1 \Rightarrow 1 - \hat{p} = 1 \Rightarrow \hat{p} = 0$. These two are contradictory to each other, so both parts of the Large Counts condition cannot be satisfied at the same time.

Problem 2.
Explain why you cannot use the methods of the lesson on "Sampling Distributions" to find the desired probability in the following situation:

> A factory employs 3000 unionized workers, of whom 30% are Hispanic. The 15-member union executive committee contains 3 Hispanics. What would be the probability of 3 or fewer Hispanics if the executive committee were chosen at random from all the workers?

Answer:
This problem is about sampling $n = 15$ people to understanding the sampling distribution of the sample proportion $\hat{p}$ of the proportion of Hispanics in the 15-person sample. To apply the methods of this lesson, the Large Counts condition must hold. But $np = 15 \cdot 0.3 < 10$ so we cannot apply the methods of this lesson. (Since we cannot ensure the sampling distribution of $\hat{p}$ is normal.)

Problem 3.
Do you go to church? The Gallup Poll asked a random sample of 1785 adults whether they attended church during the past week. Let $\hat{p}$ be the proportion of people in the sample who attended church. A newspaper report claims that 40% of all U.S. adults went to church last week. Suppose this claim is true.
(a) What is the mean of the sampling distribution of $\hat{p}$? Why?
(b) Find the standard deviation of the sampling distribution of $\hat{p}$. Check to see if the 10% condition is met.
(c) Is the sampling distribution of $\hat{p}$ approximately normal? Check to see if the Large Counts condition is met.
(d) Of the poll respondents, 44% said they did attend church last week. Find the probability of obtaining a sample of 1785 adults in which 44% or more say they attended church last week if the newspaper report's claim is true. Does this poll give convincing evidence against the claim? Explain.
(e) What sample size would be required to reduce the standard deviation of the sampling distribution to one-third the value you found in Part (b)? Justify your answer.

Answer:
(a) The mean of the sampling distribution of $\hat{p}$ is always equal to the population proportion p of 0.4.

(b) $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.4 \cdot 0.6}{1785}} = 0.0115954207.$

(c) Since $n\hat{p} = 1785 \times 0.4 = 714 \geq 10$ and $n(1-\hat{p}) = 1785 \times 0.6 = 1071 \geq 10$, the Large Counts condition holds, so the sampling distribution of $\hat{p}$ approximately normal.

(d) We compute the z-score of how far 0.44 deviates from the mean of 0.4 in our normal distribution with $\sigma = 0.0116$: $z = \frac{0.44 - 0.40}{0.0116} = 3.448$ and `pnorm(3.448,lower.tail=FALSE)=0.00028`. Therefore there is a 0.028% chance that the poll gives a result as extreme as or more extreme than $\geq 44\%$. This poll gives **convincing evidence** against the original claim.

(e) If the original sampling distribution had standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$, then $\sqrt{\frac{p(1-p)}{9n}} = \frac{1}{3} \cdot \sqrt{\frac{p(1-p)}{n}}$ has one-third that value. So we need 9 times the original sample size of 1785.

Problem 4. (On-time shipping)
A mail-order company advertises that it ships 90% of its orders within three working days. You select an SRS of 100 of the 5000 orders received in the past week for an audit. The audit reveals that 86 of these orders were shipped on time.
(a) If the company really ships 90% of its orders on time, what is the probability that the proportion in an SRS of 100 orders is 0.86 or less? Show your work.
(b) A critic says, “Aha! You claim 90%, but in your sample the on-time percentage is lower than that. So the 90% claim is wrong.” Explain in simple language why your probability calculation in (a) shows that the result of the sample does not refute the 90% claim.

Answer:
(a) If the company really ships 90% of its orders on time, then the pop. proportion of % of orders shipped on time is known to be 0.9. The standard deviation of the sampling distribution for $n = 100$ is
$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.9 \cdot 0.1}{100}} = 0.03.$

Since $N = 5000 > 10n = 1000$, $np = 100 \cdot 0.9 = 90 \geq 10$, and $n(1-p) = 100 \cdot 0.1 = 10 \geq 10$, both the 10% Condition and the Large Counts Condition are met. So we can further assume the sampling distribution is normal.

We compute $P(\hat{p} < 0.86)$ via `pnorm(0.86, mean=0.9, sd=0.03)` or by `pnorm((0.86-0.9)/0.03)=0.0912`.

(b) The probability we found is that there is roughly a 1-in-10 chance that an SRS of 100 orders have average shipping time that is considered "on-time". A 1-in-10 chance is large for a large business, but still on the lower side. An audit with a larger sample size is needed to determine whether the company's claim is true or not.

Sample mean (σ known)

Problem 5. (Dead battery?)
A car company has found that the lifetime of its batteries varies from car to car according to a Normal distribution with mean μ = 48 months and standard deviation σ = 8.2 months. The company installs a new brand of battery on an SRS of 8 cars.
(a) If the new brand has the same lifetime distribution as the previous type of battery, describe the sampling distribution of the mean lifetime $\bar{x}$.
(b) The average life of the batteries on these 8 cars turns out to be $\bar{x}$ = 42.2 months. Find the probability that the

sample mean lifetime is 42.2 months or less if the lifetime distribution is unchanged. What conclusion would you draw?

Answer:
(a) The sampling distribution is normal (because the pop. distribution is normal) and has mean 48 months and standard deviation $\sigma/\sqrt{8} = 8.2/\sqrt{8} \approx 2.899$ months.

(b) Using (a): $P(\bar{x} < 42.2)$ equals `pnorm(42.2, mean=48, sd=2.899)=0.0227129096429935` so there is a 2.27% chance that we obtain a sample mean of 42.2 months or less from an SRS of 8 cars if the true population mean is 48 months. Conclusion: there is strong evidence that the new brand of batteries has a shorter average lifetime compared to the old brand of batteries.

Problem 6. (Airline passengers get heavier)
In response to the increasing weight of airline passengers, the Federal Aviation Administration (FAA) told airlines to assume that passengers average 190 pounds in the summer, including clothes and carry-on baggage. But passengers vary, and the FAA did not specify a standard deviation. A reasonable standard deviation is 35 pounds. Weights are not Normally distributed, especially when the population includes both men and women, but they are not very non-Normal. A commuter plane carries 30 passengers.
(a) Explain why you cannot calculate the probability that a randomly selected passenger weighs more than 200 pounds from the above information alone.
(b) Find the probability that the total weight of 30 randomly selected passengers exceeds 6000 pounds. Show your work. (Hint: To apply the central limit theorem, restate the problem in terms of the mean weight.)

Answer:
(a) If we draw a sample of size $n = 1$, then the sampling distribution is equal to the population distribution which is not normal, so we cannot do normal distribution calculations to calculate the probability that a randomly selected passenger weighs more than 200 pounds from the above information alone.

(b) The sampling distribution of sample means of $n = 30$ passenger weights is 190 lbs.

The 10% condition of $N \geq 10n = 10 \cdot 30$ is met since there are many passengers at an airport. So we can calculate the standard deviation of the sampling distribution: $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 35/\sqrt{30} \approx 6.39$ lbs.

Since $n = 30 \geq 30$, the sampling distribution is approximately normal.

Our sample mean is $6000/30 = 200$ lbs.

Therefore $P(\bar{x} \geq 200)$ is `pnorm(200, mean=190, sd=6.39, lower.tail=FALSE) = pnorm(10/6.39, lower.tail=FALSE) = 0.059`.

There is 5.9% probability that a random sample of 30 passengers average 200 lbs. or more.