Topic: Conditions for inference with the SDSM

Question: A school finds that time spent studying for a test isn't normally distributed for their $2,800$ students. They would like to use the Central Limit Theorem to normalize the data. Which sample allows them to use the Central Limit Theorem?

Answer choices:

A The school samples 300 different students randomly

B The school samples 200 different students randomly

C The school samples 200 different students from an Honors classes

D The school samples 25 different students randomly

Solution: B

The random sample of 200 different students ("different" tells us that we're sampling without replacement) will allow the school to use the Central Limit Theorem. The sample is random, less than 10% of the population, and greater than 30, so it's large enough to normalize the data.

The other answer choices wouldn't allow the school to use the Central Limit Theorem. Answer choice A samples more than 10% of the population which doesn't maintain independence, answer choice C isn't random because selecting students in Honors classes will skew the data, and answer choice D doesn't have a big enough sample because a sample size smaller than 30 won't normalize the data.

Topic: Conditions for inference with the SDSM

Question: A company produces tires in a factory. Individual tires are filled to an approximate pressure of 36 PSI (pounds per square inch), with a standard deviation of 0.8 PSI. The pressure in the tires is normally distributed. The company randomly selects 125 tires to check their pressure. What is the probability that the mean pressure in the tires is within 0.1 PSI of the population mean?

Answer choices:

A $8.38\,\%$

B $91.62\,\%$

C $71.55\,\%$

D $83.84\,\%$

Solution: D

To verify normality, our sample must be random, no more than 10% of the population, and (if the population is not normal) the sample size must be greater than 30.

The sample was collected randomly. It's safe to assume that 125 tires is less than 10% of the total tires produced in the factory. The population is normal, so the sample size doesn't have to be greater than 30, but 125 is greater than 30 anyway. The sample space meets the conditions of normality.

Find the standard deviation of the sampling distribution.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$\sigma_{\bar{x}} = \frac{0.8}{\sqrt{125}}$$

$$\sigma_{\bar{x}} \approx 0.07155$$

We want to know the probability that the sample mean $\bar{x}$ is within 0.1 PSI of the population mean. We need to express 0.1 in terms of standard deviations.

$$\frac{0.1}{0.07155} \approx 1.39762 \approx 1.40$$

This means we want to know the probability $P(-1.40 < z < 1.40)$. Using a z-table, a z-value of -1.40 gives 0.0808 and a z-value of 1.40 gives 0.9192. The probability under the normal curve between these z-scores is

$$P(-1.40 < z < 1.40) = 0.9192 - 0.0808$$

$$P(-1.40 < z < 1.40) = 0.8384$$

$$P(-1.40 < z < 1.40) = 83.84\,\%$$

There's an $83.84\,\%$ chance that our sample mean will fall within 0.1 PSI of the population mean of 36 PSI.

Topic: Conditions for inference with the SDSM

Question: A large cookie company knows that the weight of their tins of Christmas cookies is normally distributed with a mean weight of 1 pound and a standard deviation of 0.2 pounds. If they take a random 50-tin sample, what is the probability that the sample mean $\bar{x}$ is within 0.05 pounds of the population mean?

Answer choices:

A $94.16\,\%$

B $92.32\,\%$

C $89.14\,\%$

D $84.98\,\%$

Solution: B

We were told that the sample was taken randomly. Our sample size is $n \geq 30$. We're sampling without replacement, but we can safely assume that this large cookie company makes more than 500 Christmas cookie tins. Therefore, we've met the random, normal, and independence conditions, respectively.

To answer the probability question, we'll start by finding the mean of the SDSM, but we know it'll be equal to the population mean, so $\mu = \mu_{\bar{x}} = 1$. The standard error will be

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$\sigma_{\bar{x}} = \frac{0.2}{\sqrt{50}}$$

$$\sigma_{\bar{x}} \approx 0.028$$

We want to know the probability that the sample mean $\bar{x}$ is within 0.05 pounds of the population mean, 1. A 0.05 interval around 1 gives us the interval 0.95 to 1.05, so

$$P(0.95 < \bar{x} < 1.05) \approx P\left(\frac{0.95 - 1}{0.028} < z < \frac{1.05 - 1}{0.028}\right)$$

$$P(0.95 < \bar{x} < 1.05) \approx P\left(-\frac{0.05}{0.028} < z < \frac{0.05}{0.028}\right)$$

$$P(0.95 < \bar{x} < 1.05) \approx P(-1.77 < z < 1.77)$$

Which means we want to know the probability of $P(-1.77 < z < 1.77)$. In a z-table, a z-value of 1.77 gives 0.9616,

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706

and a z-value of -1.77 gives 0.0384.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455

Which means the probability under the normal curve between these z-scores is

$$P(0.95 < \bar{x} < 1.05) \approx 0.9616 - 0.0384$$

$$P(0.95 < \bar{x} < 1.05) \approx 0.9232$$

$$P(0.95 < \bar{x} < 1.05) \approx 92.32\,\%$$

So there's an approximately $92.32\,\%$ chance that the mean $\bar{x}$ of the 50-tin sample the company takes will fall within 0.05 pounds of the population mean of $\mu = 1$ pound.