

Sampling and the Central Limit Theorem – Quiz

Probability & Statistics

Question 1

Why do we use samples instead of studying entire populations?

- A. Samples are always more accurate
- B. Populations are often too large or impractical to measure completely
- C. Statistics only works with samples
- D. Populations don't exist in reality

Question 1

Why do we use samples instead of studying entire populations?

- A. Samples are always more accurate
- B. Populations are often too large or impractical to measure completely
- C. Statistics only works with samples
- D. Populations don't exist in reality

Answer: B

Populations are often too large, expensive, or impractical to measure completely. We use samples to make inferences about the population.

Question 2

What does IID stand for?

- A. Independent and Identically Distributed
- B. Identical and Individual Data
- C. Integrated Information Distribution
- D. Independent Individual Datasets

Question 2

What does IID stand for?

- A. Independent and Identically Distributed
- B. Identical and Individual Data
- C. Integrated Information Distribution
- D. Independent Individual Datasets

Answer: A

IID means Independent and Identically Distributed - each observation is drawn independently from the same probability distribution.

Question 3

The sample mean \bar{x} is:

- A. A fixed population parameter
- B. A random variable that varies from sample to sample
- C. Always equal to the population mean
- D. The median of the sample

Question 3

The sample mean \bar{x} is:

- A. A fixed population parameter
- B. A random variable that varies from sample to sample
- C. Always equal to the population mean
- D. The median of the sample

Answer: B

The sample mean is a random variable - different samples give different values. It has its own distribution called the sampling distribution.

Question 4

The standard error of the sample mean is:

- A. σ / n
- B. $\sigma \times \sqrt{n}$
- C. σ / \sqrt{n}
- D. σ^2 / n

Question 4

The standard error of the sample mean is:

- A. σ / n
- B. $\sigma \times \sqrt{n}$
- C. σ / \sqrt{n}
- D. σ^2 / n

Answer: C

$SE(\bar{x}) = \sigma / \sqrt{n}$. As sample size n increases, the standard error decreases.

Question 5

If you quadruple the sample size, the standard error:

- A. Doubles
- B. Is cut in half
- C. Stays the same
- D. Is divided by 4

Question 5

If you quadruple the sample size, the standard error:

- A. Doubles
- B. Is cut in half
- C. Stays the same
- D. Is divided by 4

Answer: B

SE = σ / \sqrt{n} . If n becomes $4n$, SE becomes $\sigma / \sqrt{4n} = \sigma / (2 \sqrt{n})$, which is half the original.

Question 6

The Law of Large Numbers states that as n approaches infinity:

- A. Variance becomes infinite
- B. The sample mean converges to the population mean
- C. The distribution becomes uniform
- D. Probability becomes 0

Question 6

The Law of Large Numbers states that as n approaches infinity:

- A. Variance becomes infinite
- B. The sample mean converges to the population mean
- C. The distribution becomes uniform
- D. Probability becomes 0

Answer: B

The Law of Large Numbers: as sample size increases, the sample mean \bar{x} converges to the population mean μ .

Question 7

The Central Limit Theorem states that:

- A. All data follows a normal distribution
- B. Sample means are approximately normal for large n , regardless of the original distribution
- C. Variance decreases to zero
- D. Only normal data can be analyzed

Question 7

The Central Limit Theorem states that:

- A. All data follows a normal distribution
- B. Sample means are approximately normal for large n , regardless of the original distribution
- C. Variance decreases to zero
- D. Only normal data can be analyzed

Answer: B

The CLT says that sample means are approximately normally distributed for large n , no matter what shape the original population has.

Question 8

The common rule of thumb for CLT to apply is:

- A. $n \geq 5$
- B. $n \geq 10$
- C. $n \geq 30$
- D. $n \geq 100$

Question 8

The common rule of thumb for CLT to apply is:

- A. $n \geq 5$
- B. $n \geq 10$
- C. $n \geq 30$
- D. $n \geq 100$

Answer: C

$n \geq 30$ is the classic rule of thumb. However, for highly skewed or heavy-tailed distributions, larger n may be needed.

Question 9

For skewed populations, the CLT requires:

- A. Smaller sample sizes
- B. Larger sample sizes than for symmetric populations
- C. The same sample size as symmetric populations
- D. CLT doesn't apply to skewed data

Question 9

For skewed populations, the CLT requires:

- A. Smaller sample sizes
- B. Larger sample sizes than for symmetric populations
- C. The same sample size as symmetric populations
- D. CLT doesn't apply to skewed data

Answer: B

More skewed distributions require larger sample sizes for the normal approximation to be accurate.

Why is the CLT called the most important theorem in statistics?

- A. It proves all data is normal
- B. It allows us to use normal-based methods even when data isn't normal
- C. It eliminates the need for sampling
- D. It makes all variances equal

Why is the CLT called the most important theorem in statistics?

- A. It proves all data is normal
- B. It allows us to use normal-based methods even when data isn't normal
- C. It eliminates the need for sampling
- D. It makes all variances equal

Answer: B

The CLT justifies using normal distribution-based methods (like z-tests and confidence intervals) for sample means, even when the underlying data is not normal.

Question 11

In simple random sampling:

- A. Only certain members can be selected
- B. Every member has an equal chance of selection
- C. The largest values are always selected
- D. Samples are taken from specific strata

Question 11

In simple random sampling:

- A. Only certain members can be selected
- B. Every member has an equal chance of selection
- C. The largest values are always selected
- D. Samples are taken from specific strata

Answer: B

Simple random sampling gives every member of the population an equal probability of being selected, with selections being independent.

Stratified sampling involves:

- A. Selecting the first n individuals
- B. Dividing the population into groups and sampling from each
- C. Only sampling large values
- D. Random selection without any structure

Question 12

Stratified sampling involves:

- A. Selecting the first n individuals
- B. Dividing the population into groups and sampling from each
- C. Only sampling large values
- D. Random selection without any structure

Answer: B

Stratified sampling divides the population into homogeneous strata (groups) and then samples from each stratum, often proportionally.

The bootstrap method:

- A. Requires knowledge of the population distribution
- B. Resamples from the original sample to estimate sampling variability
- C. Only works for normal data
- D. Requires very small sample sizes

The bootstrap method:

- A. Requires knowledge of the population distribution
- B. Resamples from the original sample to estimate sampling variability
- C. Only works for normal data
- D. Requires very small sample sizes

Answer: B

Bootstrap resamples with replacement from the observed sample to approximate the sampling distribution without assuming any specific population distribution.

Question 14

As sample size n increases, the sampling distribution of \bar{x} :

- A. Becomes more spread out
- B. Becomes more concentrated around μ
- C. Becomes more skewed
- D. Stays exactly the same

Question 14

As sample size n increases, the sampling distribution of \bar{x} :

- A. Becomes more spread out
- B. Becomes more concentrated around μ
- C. Becomes more skewed
- D. Stays exactly the same

Answer: B

Larger n means smaller standard error, so the sampling distribution becomes more concentrated (narrower) around the population mean.

Question 15

The finite population correction is used when:

- A. The population is infinite
- B. A large fraction of the population is sampled
- C. The sample size is very small
- D. Data is not normally distributed

Question 15

The finite population correction is used when:

- A. The population is infinite
- B. A large fraction of the population is sampled
- C. The sample size is very small
- D. Data is not normally distributed

Answer: B

FPC applies when sampling without replacement from a finite population where the sample is a significant fraction (e.g., $\geq 10\%$) of the population.

Question 16

Biased sampling leads to:

- A. More accurate estimates
- B. Systematic over- or under-estimation of population parameters
- C. Smaller standard errors
- D. No effect on results

Question 16

Biased sampling leads to:

- A. More accurate estimates
- B. Systematic over- or under-estimation of population parameters
- C. Smaller standard errors
- D. No effect on results

Answer: B

Biased sampling systematically favors certain outcomes, leading to estimates that consistently miss the true population value in a particular direction.

Question 17

According to CLT, the sampling distribution of \bar{x} has mean:

- A. 0
- B. μ (population mean)
- C. σ
- D. n

Question 17

According to CLT, the sampling distribution of \bar{x} has mean:

- A. 0
- B. μ (population mean)
- C. σ
- D. n

Answer: B

$E[\bar{x}] = \mu$. The sampling distribution of the sample mean is centered at the population mean.

Question 18

Why do casino houses always win in the long run?

- A. They cheat
- B. Law of Large Numbers: the average converges to the expected value (which favors the house)
- C. They have more money
- D. Luck always favors the house

Why do casino houses always win in the long run?

- A. They cheat
- B. Law of Large Numbers: the average converges to the expected value (which favors the house)
- C. They have more money
- D. Luck always favors the house

Answer: B

The Law of Large Numbers ensures that over many games, the actual average outcome converges to the expected value, which is designed to favor the house.

Question 19

The sampling distribution of \bar{x} has variance equal to:

- A. σ^2
- B. σ^2 / n
- C. $\sigma^2 \times n$
- D. σ / n

Question 19

The sampling distribution of \bar{x} has variance equal to:

- A. σ^2
- B. σ^2 / n
- C. $\sigma^2 \times n$
- D. σ / n

Answer: B

$\text{Var}(\bar{x}) = \sigma^2 / n$. The variance of the sample mean decreases as sample size increases.

Question 20

If the original population is already normal, the sampling distribution of \bar{x} is:

- A. Approximately normal only for large n
- B. Exactly normal for any sample size n
- C. Uniform
- D. Exponential

Question 20

If the original population is already normal, the sampling distribution of \bar{x} is:

- A. Approximately normal only for large n
- B. Exactly normal for any sample size n
- C. Uniform
- D. Exponential

Answer: B

If the population is normal, \bar{x} is exactly normally distributed for any n . CLT is needed only when the population is non-normal.