Data Analysis In-Class Worksheet #06: Distribution & CI

Distributions

1. A normal distribution has _3_ SDs above the mean and _3__ SDs below the mean

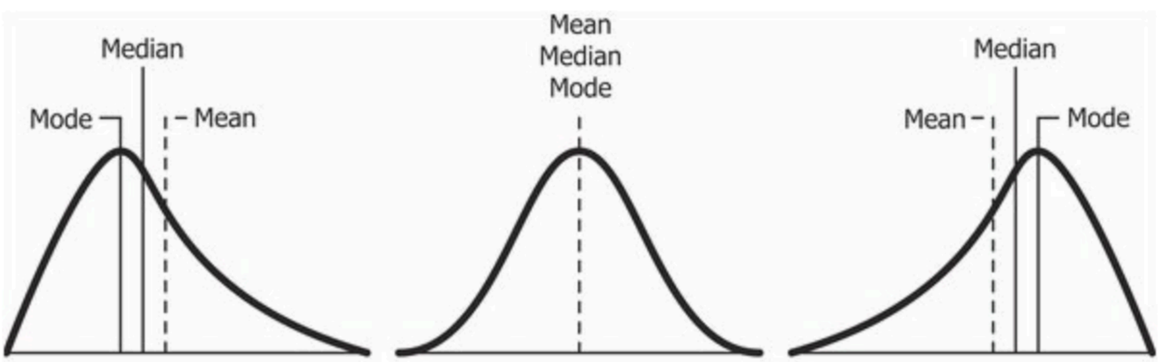
2. Which of the following is an outlier? ⬜ Z = -.3 ☑ Z = 5 ⬜ Z = 0

3. Which of these is a negative skew? ☑ -.3 ⬜ 1.2 ⬜ 0

4. Which of the following graphs shows a positive skew (Right Tail)?

☑ A ⬜ B ⬜ C

A B C



5. Our class dataset is a __sample____ of the _population__ of all Williamsburg properties

6. The _Shapiro-Wilk_ test determines if the population from which a sample is drawn is normally distributed or not.

7. If the test result has a p (probability) value of .04, that means

⬜ the population is normally distributed

☑ the population is NOT normally distributed (If $p < 0.05$, reject normality)

8. If the test result has a p (probability) value of .27, that means

☑ the population is normally distributed (If $p > 0.05$, accept normality)

⬜ the population is NOT normally distributed

9. If we want to convert a non-normal variable into a normal distribution, we can try to take __log or natural log__ of the variable.

10. The binomial distribution is the distribution of __Bernoulli__ trials, which depends on __p (probability)__, and _n (sample size) __

Confidence Interval

11. How does the point estimate differ from an interval estimate?

The point estimate is a single value (e.g., 17 years old), while the interval estimate gives you a range (e.g., 11–19 years old) with a confidence level

12. What is Standard Error? It's an estimate of the population's standard deviation

13. Which estimate indicates a higher level of confidence?

☑ I bet that Abhishek will get married between 28 and 30 years old (narrow range, higher precision)

⬜ I bet that Abhishek will get married between 18 and 95 years old

14. Take a sample of 16 stocks from a large population, with a sample mean return of 5.2%, and a sample standard deviation of 1.5%

Given n=16

μ_X =5.2%

σ_X =1.5%

$\mu = \mu_X \pm Z_{(1-\alpha)/2} * \sigma_X/\sqrt{n}$

μ is population mean

μ_X is sample mean

$Z_{(1-\alpha)/2}$ is Z-score corresponding to the confidence level

σ_X standard deviation

n sample size

$\sigma_X/\sqrt{n}$ is standard error (SE)

15. Calculate 95% confidence interval for the population mean

Z-score corresponding to 95% confidence interval is 1.96

$\mu = \mu_X \pm Z_{(1- \alpha)/2} * \sigma_X/\sqrt{n}$

$= 5.2\% \pm 1.96 * 1.5\% / \sqrt{16}$

$= (4.465\%, 5.935\%)$

16. How much is the margin of error?

Margin of Error (ME)= $Z_{(1-\alpha)/2} * \sigma_X / \sqrt{n}$

$1.96 * 1.5\% / \sqrt{16}$

0.735%

17. How much is the Standard Error?

$\frac{\sigma_X}{\sqrt{n}}$

$1.5\% / \sqrt{16}$

0.375%

Source: https://financetrain.com/confidence-interval-population-mean-known-population-variance/