

BUS 204

Quiz 1

Scheduled for 9/27/20

Study Guide

as of 9/10/20

The following concepts will be on this quiz:

- 1. Statistical inference tools:** tools that allow a decision maker to reach a conclusion about a population of data based on a subset of data from population.
- 2. Experiment:** a process that produces a single outcome whose result cannot be predicted with certainty.
- 3. Sample:** a subset of the population.
- 4. Census:** an enumeration of the entire set of measurements taken from the whole population.
- 5. Non-statistical sampling techniques:** those methods of selecting samples using convenience, judgement, or other non-chance processes.
- 6. Quantitative data:** measurements whose values are inherently numerical.
- 7. Qualitative data:** data whose measurement scale is inherently categorical.
- 8. Parameter:** a measure computed from the entire population. As long as the population does not change, the value of the parameter will not change.
- 9. Statistic:** a measure computed from a sample that has been selected from a population. The value of the statistic will depend on which sample is selected.
- 10. Dependent variable:** a variable whose values are thought to be a function of, or dependent on, the values of another variable called the independent variable.

11. Independent variable: a variable whose values are thought to impact the values of the dependent variable. The independent variable, or explanatory variable, is often within the direct control of the decision maker.

12. Right-skewed data: a data distribution when the mean of the data is larger than the median.

13. Left-skewed data: a data distribution where the mean for the data is smaller than the median.

14. Population Variance: is the average of the squared distances of the data values from the mean. It is a measure of variability that utilizes all data. The variance is based on the deviation about the mean.

15. Standard deviation: is the positive square root of the variance. It is a distance from the mean on the x axis.

16. Standardization of data values (standard normal distribution): a normal distribution that has a mean = 0 and a standard deviation = 1,

Standardized data values are referred to as z-scores.

17. Coefficient of variation: the ratio of the standard deviation to the mean expressed as a percentage. The coefficient of variation is used to measure the relative variation in data.

18. Decision tree: a diagram that illustrates the correct ordering of actions and events in a decision-analysis problem. Each act or event is represented by a branch on the decision tree.

19. Axioms: givens and assumptions that are necessary in analytical work. In logic, an indemonstrable first principle, rule, or maxim, that has found general acceptance or is thought worthy of common acceptance whether by virtue of a claim to intrinsic merit or on the basis of an appeal to self-evidence. An example would be: "Nothing can both be and not be at the same time and in the same respect", or the "80-20 rule."

20. Consumer Research process (Burke video): is never fully reliable. Prospective consumers are irrational in their behavior.

21. Forensic research: a retrospective, descriptive research based on historical data. Typically, it is applied in GAPS analysis (explaining differences between stated goals and the actual results).

22. Predictive analysis - decisions making process about the future when the future is uncertain.

23. Empirical Rule: The empirical rule states that for a normal distribution, nearly all of the data will fall within three standard deviations of the mean. The empirical rule can be broken down into three parts:

- 68.28% of data falls within the first standard deviation from the mean.
- 95.44% fall within two standard deviations from the mean.
- 99.72% fall within three standard deviations from the mean..

24. Mean: a numerical measure of the center of a set of the quantitative measures computed by dividing the sum of the values by the number of values in the data.

25. Weighted mean: the mean value of data values that have been weighted according to their relative importance.

26. Median: is the center value that divides a data array into two halves.

27. Mode: is the value on a data set that occurs most frequently.

28. The concept of “self”: it is a product of all external influences on an individual and all internal characteristics of that individual.

29. Bias: is a disproportionate weight in favor of or against an idea or thing, usually in a way that is closed-minded, prejudicial, or unfair. Biases can be innate or learned. People may develop biases for or against an individual, a group, or a belief. In science and engineering, a bias is a systematic error. In statistics, it is the tendency of a predictive model to overestimate or underestimate the value of a continuous outcome.

30. Business Statistics: This is a form of mathematical analyses that are used to convert data into meaningful information. It uses quantified models, representations and synopses for a given set of experimental data or real-life studies. The discipline of Business Statistics studies methodologies to gather, review, analyze and draw conclusions from data.

31. Factorials: i.e. : $0! = 1$

32. Utility: In economics and finance, risk aversion (one's utility) is the behavior of humans (especially consumers and investors), who, when exposed to uncertainty, attempt to lower that uncertainty.

It is the hesitation of a person to agree to a situation with an unknown payoff rather than another situation with a more predictable payoff but possibly lower expected payoff.

For example, a risk-averse investor might choose to put their money into a bank account with a low but guaranteed interest rate, rather than into

a stock that may have high expected returns, but also involves a chance of losing value. It is a measure of the total worth or relative desirability of a particular outcome.

33. Normal distribution: is a bell-shaped distribution with the following properties. It is:

- a) unimodal: that is, the normal distribution peaks at a single value.
- b) symmetrical: this means that the two areas under the curve between the mean and any two points equidistant on either side of the mean are identical.
- c) is asymptotic to the x-axis (the normal curve approaches the x-axis as to moves toward the + or- infinity but never cross them.

also,

- d) one side of the distribution is the mirror image of the other.
- e) the mean, median, and mode are equal.
- f) the amount of variation in the random variable determines the width of the normal distribution.

34. Probability: the chance that a particular event will occur. The probability of an event will have a value in the range from 0 to 1.

35. Random variable: a variable that assigns a numerical value to each outcome of a random experiment or trial.

36. Discrete random variable: a random variable that can only assume a countable number of values.

37. Continuous random variable: random variables that can assume any value in an interval.

38. Outliers: observations with unusually large or unusually small values. Typically located above or below +/- two standard deviations from the mean. These extreme values are called outliers.

39. Frequency distribution: is a tabular summary of data showing the number (frequency) of data values in each of several categories.

40. Variation: captures differences in values of a variables of interest.

41. Cumulative relative frequency distribution: is a running total of the proportions of observations with a given array of data.

42. Uncertainty: no one can predict the future, yet business decisions about the future have to be made. Good decisions will create value and bad decisions will destroy it.