Student name:\_\_\_\_\_\_\_\_\_\_

**1)** When forecasting with time-series data, it is highly recommended to test for the presence of a trend in the data. Testing for trend at the 10% level of significance

 A) requires use of the t distribution.
 B) can be accomplished by use of a standard 95% correlogram.
 C) can be accomplished by comparing the estimated autocorrelation coefficient with the number 2 divided by the square root of sample size.
 D) requires use of the standard normal probability distribution.

**2)** Type I error

 A) has a probability value equal to the significance level of any statistical test.
 B) is a measure of the uncertainty associated with rejecting any null hypothesis on the basis of sample data.
 C) is said to arise when we reject a true null hypothesis.
 D) All of the options are correct.
 E) both a. and b. are correct.

**3)** The correlation coefficient provides

 A) a measure of the strength of the linear association between a quantitative variable and a categorical variable.
 B) a measure of the extent to which changes in one variable cause changes in another variable.
 C) a measure of the strength of the linear association between two quantitative variables.
 D) a measure of the strength of the association (not necessarily linear) between two categorical variables.
 E) a measure of the strength of the linear association between two categorical variables.

**4)** Your authors present a guide to selecting an appropriate forecasting method based on

 A) quantitative background of the forecast user.
 B) quantity of historical data available.
 C) data patterns.
 D) All of the options are correct.
 E) forecast horizon.

**5)** A k-period plot of autocorrelations is called

 A) a simple regression line.
 B) a product diffusion curve.
 C) a scatter plot.
 D) an autocorrelation function.

**6)** A random sample of bolts is taken from inventory, and their lengths are measured. The average length in the sample is 5.3 inches with a standard deviation of .2 inches. The sample size was 50.A 95% confidence interval for the unknown population mean is

 A) None of the options are correct.
 B) 5.3 inches plus or minus .056.
 C) 4.9 to 5.7 inches.
 D) 5.3 inches.
 E) 4.784 to 5.816 inches.

**7)** When the correlation coefficient is negative, it means

 A) None of the above.
 B) X will not be a good predictor of Y.
 C) when X goes down, Y does too.
 D) there is a weak relationship.
 E) when X goes down, Y tends to go up.

**8)** The standard normal probability table

 A) shows a normal distribution with standard deviation equal to zero.
 B) None of the options are correct.
 C) is used to make inference for all normally distributed random variables.
 D) All of the options are correct.
 E) is equivalent to a t distribution if the sample size is less than 30.

**9)** Which probability distribution is appropriate for testing hypotheses concerning an unknown population mean when the sample variance is used to estimate the population variance?

 A) The standard normal distribution.
 B) The normal distribution with mean 0 and variance 1.
 C) None of the options are correct.
 D) The normal distribution with mean μ and variance σ2.
 E) The t distribution with n-1 degrees of freedom.

**10)** Autocorrelation refers to the correlation between a variable and

 A) another variable when the analysis is done on a computer.
 B) another very similar variable.
 C) itself when lagged one or more periods.
 D) itself.
 E) None of the options are correct.

**11)** Which frame of the correlation diagram (A through F) represents an imperfect negative linear correlation?

 A) D
 B) A
 C) E
 D) B
 E) F
 F) C

**12)** Time series data of a typical The GAP store should show which of the following data patterns?

 A) Trend
 B) Cyclical
 C) Random
 D) Seasonal
 E) All of the options are correct.

**13)** Which of the following is incorrect?

 A) The forecaster should be able to defend why a particular model or procedure has been chosen.
 B) Forecast errors should not be discussed since most people know that forecasting is an inexact science.
 C) None of the options are correct.
 D) You should tailor your presentation to the sophistication of the audience to maximize credibility in the forecast process.
 E) Forecast errors should be discussed in an objective manner to maximize management’s confidence in the forecast process.

**14)** A study examined the relationship between the sepal length and sepal width for two varieties of an exotic tropical plant. Varieties A and B are represented by x’s and o’s, respectively, in the following plot. Which of the following statements is FALSE?

 A) Considering each variety separately, there is a positive correlation between sepal length and sepal width.
 B) Considering both varieties together, the least squares regression line for predicting sepal length from sepal width has a positive slope.
 C) Considering variety A alone, there is a negative correlation between sepal length and sepal width.
 D) Considering variety B alone, the least squares regression line for predicting sepal length from sepal width has a negative slope.
 E) Considering both varieties together, there is a positive correlation between sepal length and sepal width.

**15)** Which measure of dispersion in a data set is the most intuitive and represents an average?

 A) Mean.
 B) Variance.
 C) Mode.
 D) Standard deviation.
 E) Range.

**16)** Of the following model selection criteria, which is often the most important in determining the appropriate forecast method?

 A) What is the forecast horizon?
 B) Technical background of the forecast user
 C) How much money is in the forecast budget?
 D) Patterns the data have exhibited in the past
 E) When is the forecast needed?

**17)** If the scatterplot of two variables has a circular pattern, this suggests the two variables have a population correlation coefficient of

 A) −.5.
 B) 0.
 C) +1.
 D) +.5.
 E) −1.

**18)** The median and mode may be more accurate than the sample mean in forecasting the populations mean when

 A) the sample has one large outlier.
 B) the population is assumed to be normally distributed.
 C) the sample size is small.
 D) the sample size is large.
 E) All of the options are correct.

**19)** A**utocorrelation Function for GDP** **Autocorrelation function for GDP with second differencing**Refer to the autocorrelation functions for Gross Domestic Product (GDP) presented above. The shape of these autocorrelation functions

 A) indicates that GDP is not autocorrelated.
 B) indicates that GDP is seasonal.
 C) indicates that GDP has a trend.
 D) indicates that GDP is stationary.

**20)** The correlation coefficient is also called

 A) Theil’s U.
 B) the variance.
 C) the Pearson product-moment correlation coefficient.
 D) the autocorrelation coefficient.

**21)** If the correlation between body weight and annual income were high and positive, we could conclude that

 A) high incomes cause people to gain weight.
 B) high income people tend to be heavier than low income people, on average.
 C) high income people tend to spend a greater proportion of their income on food than low income people, on average.
 D) low incomes cause people to eat less food.
 E) high incomes cause people to eat more food.

**22)** Which of the following is not a benefit of a scatter diagram?

 A) The strength of the relationship may be revealed.
 B) The nature of the X-Y relationship (linear of nonlinear) may be revealed.
 C) The sign of the correlation coefficient will be revealed.
 D) Displaying the population size.

**23)** Which of the following is not a measure of central tendency in a population?

 A) Mean.
 B) Mode.
 C) Median.
 D) Range.

**24)** Which of the following is not a foundation of classical statistics?

 A) Summary measures of probability distributions called descriptive statistics
 B) The use of sampling distributions, which describe the uncertainty in making inference about the population on the basis of a sample
 C) Probability distribution functions, which characterize all outcomes of a variable
 D) The concept of expected value
 E) None of the options are correct.

**25)** The correlation coefficient (ρ) is an extremely important descriptive statistic because

 A) It provides a unit-free measure of how two random variables move together.
 B) It provides a measure of the linear association between a pair of random variables.
 C) All of the options are correct.
 D) It provides the forecaster with a diagnostic tool of when regression analysis is appropriate for the business-forecasting problem.

**26)** For a collection of 15 X-Y data values, the sample correlation coefficient was estimated at -.63 from a sample of size 15. The calculated t value for a null of zero correlation is

 A) −2.92.
 B) 2.92.
 C) −1.92.
 D) None of the options are correct.
 E) 1.92.

**27)** The correlation coefficient (r) represents

 A) the degree of significance in the *t*-test between two variables.
 B) the linear association between two variables.
 C) the causal relationship between two variables.
 D) the positive association between two variables.

**28)** Which statistic is correctly interpreted as the "average" spread of data about the mean?

 A) Range.
 B) Standard deviation.
 C) Mean.
 D) Mode.
 E) Variance.

**29)** Correlation coefficients may range in value

 A) from zero to one.
 B) from zero to four.
 C) from zero to 100 percent.
 D) from −1 to +1.

**30)** Which of the following is not used to calculate the sample Pearson correlation coefficient for the variables X and Y?

 A) Sample mean of Y.
 B) Sample standard deviation of X.
 C) Sample mean of X.
 D) All of the options are used to calculate correlation coefficients.
 E) Sample covariance of X and Y.

**31)** The range of a data set

 A) is the difference between zero and the largest value in the data set.
 B) is the standard deviation divided by the number of observations.
 C) is the highest value in the data set.
 D) is the difference between the smallest value and the greatest value.

**32)** In order to conduct a correlation analysis, the collected data must be

 A) All of the options are correct.
 B) numerical
 C) constructed of categories
 D) related to the real world
 E) highly correlated

**33)** For which data frequency is seasonality not a problem?

 A) Annual.
 B) Quarterly.
 C) Weekly.
 D) Daily.
 E) Monthly.

**34)** The normal distribution

 A) is identical to the *t*-Distribution.
 B) is symmetrical around the mean.
 C) is fully defined by three characteristics: the median, the mode, and the mean.
 D) cannot be easily distinguished from a Chi-Square distribution.

**35)** Which statement is incorrect?

 A) The sample mean is the best estimator if sampling from a normal population.
 B) Confidence intervals provide no more information than point estimates.
 C) The sample mean is an unbiased estimator.
 D) Confidence intervals depend on sample size.
 E) The sample variance is an unbiased estimator.

**36)** Accuracy refers to

 A) how accurate the measures of central tendencywill be.
 B) the accuracy of the seasonal component of a forecasting model.
 C) how well the model works retrospectively.
 D) how well the model works in the forecast horizon.

**37)** Which of the following is a measure of central tendency?

 A) median
 B) autocorrelation coefficient
 C) *t*-statistic
 D) standard deviation

**38)** Based upon ten years of monthly data, the monthly rate of return of the DOW Jones 30 composite stock portfolio was normally distributed with mean .0084 and variance .0014. What is the probability, that in any given month, we observe a rate of return on the DOW above 10 percent?

 A) Not enough information is provided to answer the question.
 B) Three percent.
 C) Two percent.
 D) Less than one percent.

**39)** A medical researcher has just calculated a correlation coefficient of zero for two particular random variables. Which of the following statements is most accurate?

 A) There is no significant relationship between the two variables.
 B) There is no significant linear relationship between the two variables.
 C) There is no significant linear difference between the two variables.
 D) There is a significant linear relationship between the two variables.

**40)** Which of the following is not a descriptive statistic?

 A) Expected value.
 B) Range.
 C) Variance.
 D) Mean.
 E) None of the options are correct.

**41)** Last year's midterm results showed a mean of 51 points and a variance of 46. An approximate confidence interval is

 A) 5 to 97.
 B) 37.4 to 64.6.
 C) None of the options are correct.
 D) 44.2 to 57.8.

**42)** An unbiased model

 A) is one that does not consistently over-estimate or under-estimate the true value of a parameter.
 B) is one which contains no independent variable; it depends solely on time-series pattern recognition.
 C) is one made up by a team of forecasters.
 D) is one that consistently produces estimates with the smallest RMSE.

**43)** Suppose the sample Pearson correlation coefficient (r) is estimated to be .75 with a sample size of 35. The correct calculated value of the test statistic for a null of zero correlation is

 A) 44.1.
 B) None of the options are correct.
 C) 6.5.
 D) 2.5
 E) 8.6.

**44)** A sample of 100 are selected at random from a process with a mean of 500.52 and a standard deviation of 4.0. Estimate the probability that a sample of 100 would have a mean equal to or greater than 500.52 if the true population mean is really 500.0.

 A) about .4938.
 B) about .4032.
 C) about .9032.
 D) about .0968.

**45)** A trend in a time series

 A) is a long term change in the level of the data.
 B) occurs in a time series when there is a regular variation in the level of the data that repeats itself at the same time each year.
 C) is represented by wavelike upward and downward movements of the data around the long-term trend.
 D) contains the fluctuations that are not part of the other three components.

**46)** Which of the following is not a reason for testing if the population correlation coefficient is zero?

 A) To see if r and rho (ρ. are equal)
 B) To bring sample size into the analysis.
 C) All of the options are correct.
 D) To determine if a significant X-Y relationship exists.
 E) To make inference from sample to population.

**47)** Fit refers to

 A) how well the model works in the forecast horizon.
 B) the accuracy of the seasonal component of a forecasting model.
 C) how well the model works retrospectively.
 D) how accurate the measures of central tendency will be.

**48)** When running a hypothesis test, the process begins by setting up two hypotheses,

 A) the average hypothesis and the mean hypothesis.
 B) the null hypothesis and the alternative hypothesis.
 C) the theoretical hypothesis and the statistical hypothesis.
 D) Pearson’s hypothesis and the null hypothesis.

**49)** Which of the following is not a foundation of classical statistics?

 A) The knowledge of thousands and thousands of normal probability tables required for statistical inference of normally distributed random variables.
 B) The concept of expected value, which is the average value of a random variable taken over a large number of samples.
 C) Summary measures of probability distribution called descriptive statistics.
 D) Probability distribution function which characterizes all possible outcomes of a random variable.

**50)** The difference between seasonal and cyclical components is

 A) duration
 B) frequency
 C) predictability
 D) All of the options are correct.
 E) source

**51)** Quarterly time-series data with a trend can be applied to models that assume stationary data by

 A) taking the first difference of the original series.
 B) averaging the data over time.
 C) using a moving average.
 D) taking the fourth difference of the original series.

**52)** When we test the significance of a correlation coefficient, the null hypothesis is usually

 A) ignored.
 B) that the correlation coefficient is equal to one.
 C) that the correlation coefficient is equal to zero.
 D) tested with n-1 degrees of freedom.

**53)** Which of the following is not consistent with the presence of a trend in a time series?

 A) The autocorrelation function declines slowly towards zero as the lag increases.
 B) The autocorrelation function declines quickly to zero as the lag increases.
 C) The autocorrelation function of the first-differences quickly declines to zero.
 D) The autocorrelation function of the first-differences declines quickly to zero as the lag increases.

**54)** When evaluating a time series of data, it is useful to look at the correlation between successive observations over time. This measure of correlation is called

 A) a trial correlation.
 B) a *t*-distribution
 C) an autocorrelation.
 D) a paired correlation.

**55)** Consider the following equation:This equation represents

 A) a calculated *t*-statistic.
 B) the calculated Pearson product-moment correlation coefficient.
 C) the calculated Z-statistic.
 D) the calculated standard deviation.

**56)** Sampling distributions

 A) are used to make inference when the population of a variable is unobservable.
 B) All of the options are correct.
 C) exhibit important properties for the ranking of alternative estimators such as unbiasedness and efficiency.
 D) are the distributions of all possible values of a sample statistic based upon repeated sampling.

**57)** Which frame of the correlation diagram (A through F) represents a perfect inverse linear correlation?

 A) C
 B) A
 C) E
 D) B
 E) F
 F) D

**58)** In finance, an investor who ignores risk is termed "risk neutral." What descriptive statistic is our risk neutral investor ignoring when she generates stock portfolios?

 A) Mode.
 B) Median.
 C) None of the options are correct.
 D) Standard deviation.
 E) Mean.

**59)** Suppose you observe the entire population of a random variable and you wish to test some hypothesis about the mean. To perform your hypothesis test, you

 A) There is no answer to this question.
 B) apply the t distribution.
 C) apply a sampling distribution to the problem.
 D) obtain sample estimates of population parameters.
 E) simply find the population mean and compare it to the hypothesized value.

**60)** Which frame of the correlation diagram (A through F) represents a perfect positive linear correlation?

 A) F
 B) B
 C) D
 D) C
 E) A
 F) E

**61)** Which of the following correlation coefficients suggests the strongest relationship between the variables?

 A) 1.03
 B) .75
 C) −.02
 D) −.89

**62)** Which functions are not appropriate for use of the Pearson correlation coefficient to estimate the correlation between a pair of random variables?

 A) Cubic polynomials.
 B) Functions involving a variable raised to the one-half power.
 C) Reciprocal functions.
 D) Quadratic polynomials.
 E) Higher-order polynomials.

**63)** If we were to know the true population correlation, confidence intervals for the population correlation can be constructed using the \_\_\_\_\_\_\_ distribution.

 A) All of the options are correct.
 B) chi-square distribution
 C) F distribution
 D) standard normal distribution
 E) t distribution

**64)** Using a Student’s *t*-Distribution table and 85 degrees of freedom, what value of *t* would be associated with finding 97.5 percent of the area in the unshaded region (i.e., the area not in the tail)?

 A) 1.96
 B) 99%
 C) 0.050
 D) zero

**65)** Which of the following is not an attribute of a normal probability distribution?

 A) All of the options are correct.
 B) Most observations cluster around zero.
 C) Most observations cluster around the mean.
 D) The distribution is completely determined by the mean and variance.
 E) It is symmetrical about the mean.

**66)** A difference between the population standard deviation of the random variable X and the standard deviation of the sampling distribution of the sample mean is

 A) All of the options are correct.
 B) one is based upon the other.
 C) application to the t distribution.
 D) dependence on sample size.
 E) the possibility of sampling error.

**67)** The *t*-Distribution (also called the Student’s *t*-Distribution)

 A) resembles both a normal and a Gaussian distribution.
 B) resembles a Gaussian distribution.
 C) resembles Theil’s Distribution.
 D) resembles a Chi-Square distribution.
 E) resembles a normal distribution.

**68)** A large sample of X-Y data values are analyzed and reveal a correlation coefficient of −.88. Which statement is correct?

 A) If r had been +.88, the correlation would have been much stronger.
 B) The correlation is weak because r is less than -1.
 C) A fairly strong negative linear relationship exists.
 D) A weak negative relationship exists.

**69)** Which of the following is not part of the recommended nine-step forecast process?

 A) What role do forecasts play in the business decision process?
 B) Is there enough data?
 C) All of the options are correct.
 D) What exactly is to be forecast?
 E) How urgent is the forecast?

**70)** In a normal distribution, what percentage of the area under the curve is included between one standard deviation below the mean and one standard deviation above the mean?

 A) 99%
 B) 95%
 C) 68%
 D) There is no standard answer to the question.

**71)** Which frame of the correlation diagram (A through F) represents an imperfect positive linear correlation?

 A) B
 B) F
 C) D
 D) C
 E) A
 F) E

**72)** In calculating the sample variance, we subtract one from the sample size. This is because

 A) the sample mean is employed.
 B) the population mean is unknown.
 C) the sum of deviations about the sample mean is zero.
 D) of using the sample mean to estimate the population mean.
 E) All of the options are correct.

**73)** The arithmetic average of the relative frequency of the occurrence of some random variable is also called the \_\_\_\_\_\_\_.

 A) range
 B) None of the options are correct.
 C) mean
 D) variance
 E) standard deviation

**74)** Which time-series component is said to fluctuate around the long-term trend and is fairly irregular in appearance?

 A) None of the options are correct.
 B) Trend.
 C) Seasonal.
 D) Irregular.
 E) Cyclical.

**75)** Consider the following formula:This formula represents

 A) a calculated *t*-statistic.
 B) a calculated Z-statistic.
 C) a calculated standard deviation.
 D) a calculated correlation coefficient.

**76)** Why are forecasting textbooks full of applied statistics?

 A) Statistics is the study of uncertainty.
 B) Real-world business decisions involve risk and uncertainty.
 C) Forecasting attempts to generate certainty out of uncertain events.
 D) Forecasting ultimately deals with probability.
 E) All of the options are correct.

**77)** Which of the following statements about the probability of Type I and Type II error is not correct?

 A) Type I error cannot occur if the null hypothesis is false.
 B) If the null hypothesis is true, the results of the test will either be a correct conclusion or a Type I error.
 C) Type II error cannot occur if the null hypothesis is true.
 D) It is not possible to specify both the probabilities of Type I and II errors since only one of them can occur.

**78)** Suppose two random variables X and Y are related as follows: Y = 1/X2. The population Pearson correlation coefficient should be

 A) None of the options are correct.
 B) −1.
 C) .5.
 D) 0.
 E) +1.

**79)** If two large random samples are drawn from two populations, each having a mean of $100, the relevant sampling distribution of their difference has a mean of

 A) $200.
 B) the sum of the two sample means.
 C) the difference between the two sample means.
 D) 0.