Statistics 101 Elementary Statistics
Reviewer for Final Examination S101-FE-004

I. True or False.

1. We use Descriptive Statistics to describe the population of interest using the data collected in a sample. (F)
2. In computing the sample standard deviation, the formula $s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}}$ is used. (F)
3. The sample mode is an unbiased estimator of the population mean. (F)
4. A statistics S is defined to be the difference between the 80th percentile and the 20th percentile (that is $S = P_{80} - P_{20}$). The statistics S is a measure of variability. (T)
5. In a class of 250 students, Harry's score of 89 was the 90th percentile. This implies that Harry belongs to the top 30 examinees in his class. (T)
6. The sample median is an unbiased estimator of the population mean. (F)
7. If we conduct a census then we can compute for the values of the parameters of interest using Descriptive Statistics. (T)
8. If X is a continuous random variable then P(X=0) is always equal to 0. (T)
9. Pearson's Product-Moment Correlation Coefficient is a widely used measure of relationship between two quantitative variables. A computed value of 0 for the Pearson coefficient implies that there is no relationship between the two variables. (F)
10. If A, B, and C are events then P(A∪B∪C) = P(A) + P(B) + P(C). (F)
11. Simple random sampling is a sampling design that involves dividing the population into nonoverlapping subpopulations, and then taking a sample in each subpopulation using probability sampling. (F)
12. The Central Limit Theorem requires that the parent population from which the random sample comes from is normally distributed for the distribution of the sample mean to approximate a normal distribution when the sample size is sufficient large. (F)
13. The t-test about the population mean is very sensitive to the assumption of normality. (F)
14. In a class of 10, the following are the weight of the students (in kg): 50, 45, 94, 53, 85, 70, 60, 35, 48, and 44. The range of students' weight is 60. (F)
15. If the ordinal level of measurement was used then the sum of all the data values will be meaningful. (F)
16. The median is more sensitive than the mean to the presence of outliers in a dataset. (F)
17. If the interval level of measurement was used then arranging the data values from smallest to largest value will be meaningful. (T)
18. The t-distribution approaches the standard normal distribution as its degrees of freedom approaches infinity. (T)
19. In hypothesis testing, if the p-value is 0.01 then Ho will be rejected at 0.05 level of significance. (T)
20. It is impossible to commit a Type II error when we decide to reject Ho. (T)
21. If ρ is equal to 1 then either X causes Y or Y causes X. (F)
22. The Pearson product moment correlation coefficient between X and Y, denoted by r, can be any real number. (F)
23. If the 90% confidence interval estimate of is (15.7, 18.1) then the chance that the value of μ is in this interval is 0.90. (F)

II. Identification.

1. The levels of measurement and their properties.
2. The main agencies under the Philippine Statistics Authority.
3. The conditions of a binomial experiment.

III. Fill in the blanks.

1. Statistics is defined as the science dealing with the _____, _____, _____, _____, and _____ of data.
2. In computing for the standard score of an observation, _____ is the formula used.
3. In a 4x3 table, the number of cells with expected frequencies less than 5 but greater than 1 must not be more than _____ for the chi-square test to be valid.
4. If we select a sample of size 25 from a population containing 125 elements using SRSWOR 2
5. from a population whose $\mu=400$ and $\sigma^2=100$, then the value of the standard error of the sample mean is _____.
6. The extraneous variations in the values of the response variable which tend to mask the true effects of the treatments due to the inherent variations among experimental units and lack of uniformity in the physical conduct of the experiment is called _____.
7. If X is a random variable and $F(\cdot)$ is its cumulative distribution function (CDF) then for any real number x, $F(x)$ = _____.
8. The field of statistics that is concerned with the development of the foundations of the methods used in applied statistics is called _____.
9. _____ is the error in the study that is attributed to the variation among the values of the statistic from the different possible samples consisting of n elements.
10. If there are 2500 elements in the population and a sample of size 10 will be selected using systematic sampling then the value of the sampling interval is _____.

IV. Multiple Choice.

1. If the mean, median, and mode of the distributions are 40, 45, and 70 respectively, then the distribution is:
   a) Symmetrical
   b) Positively skewed
   c) Negatively skewed
   d) There is insufficient information to answer the question.
2. Which among the following is not an assumption of the Central Limit Theorem?
   a) The random variables follow the same distribution.
   b) The random variables are independent.
   c) The random variables have finite variances.
   d) The random variables are normally distributed.
3. Suppose you want to show the trend in the Consumer Price Index (CPI) from 2001 to 2015, which of the following is the appropriate graphical tool to show this?
   a) Scatter plot
   b) Line chart
   c) Horizontal bar chart
   d) Histogram
4. Which of the following is not a property of a binomial experiment?
   a) The experiment consists of a sequence of *n* identical trials
   b) In each of the *n* trials, there are exactly two distinct events defined.
   c) The probability of an event in a trial changes from one trial to the next
   d) D. The sequence of trials in the experiment are independent
5. Which of the following statements is/are true?

   I. The standard error is a statistic.
   II. The standard error is not affected by the unit of measurement.
   III. The standard error is a measure of central tendency.

   a) I only

b) II only
c) III only
d) All of the above
e) None of the above

6. A type II error occurs when:
   a) The null hypothesis is incorrectly accepted when it is false
   b) The null hypothesis is incorrectly rejected when it is true
   c) The null hypothesis is correctly accepted when it is true
   d) The null hypothesis is correctly rejected when it is false

V. Problem Solving.

1. A manufacturer obtained data on breakdowns in two models of portable turntables designed to rotate food in microwave ovens. In a sample of 197 turntables of the first model, 33 broke down within two years of the date of purchase. In a sample of 290 units of the second model, there were 42 that broke down within two years. The manufacturer wishes to know if the proportions of break-down of the two models are different.
   a) State Ho and Ha. Define the parameters.
   b) Write the formula of the test statistic.
   c) Compute for the value of the test statistic.
   d) Compute for the p-value.
   e) Is there sufficient evidence at 0.05 level of significance to conclude that the proportions of break-down of the two models are different?
2. The temperature X (in oC) at noon on any day in March in Quezon City is known to follow a normal distribution with mean 36.5oC and variance of 4.0oC. Evaluate the probabilities of the following events for the weather forecaster who wishes to report them for his forecast in television:
   a) event that the temperature is more than 38.4 ^{o}C
   b) event that the temperature is between 35.5 oC and 38.5 ^{o}C
3. Company A believes that the proportion of households buying the leading brand of milk is more than 0.75. They wish to test this hypothesis at 0.1 level of significance. In a random sample of 300 households, there were 240 who bought the leading brand of milk in their last purchase.
   a) State Ho and Ha.
   b) Write the formula of the test statistic to be used.
   c) Compute for the value of the test statistic.
   d) Compute for the p-value.
   e) Is there sufficient evidence at 0.1 level of significance for Company A to conclude that proportion of households buying the leading brand of milk is more than 0.75?

References:

Almeda, Josefina V., Therese Ann G. Capistrano and Genelyn Ma. F. Sarte (20__). Elementary Statistics. University of the Philippines Press. Quezon City, Philippines.

Parel, Cristina P., et. al. (1966). Introduction to Statistical Methods (With Applications). Macaraig Publishing Company. Manila, Philippines.

Walpole, Ronald E. and Raymond H. Myers (1978). Probability and Statistics for Engineers and Scientistics. 2nd Ed. Macmillan Publishing Co., Inc. New York, USA.

Stat 114 and Stat 115 Sample Exams of Prof. Therese Ann G. Capistrano