University of New Mexico
Homework III (Due Thursday, July 24)
Elements of Mathematical Statistics and Probability Theory
Stat 345 (Section 01), Summer 2014

Instructor: Fares Qeadan
Office: SMLC Room 319
Classroom: Dane Smith Hall 325
Course Website: http://www.mathalpha.com/teaching.html
Office Hours: F 8:00 -10:30 AM or by an appointment

E-mail: fqeadan@unm.edu

(1) _____ **statistics involves summarizing and describing important features of the data.**

    (a) Descriptive
    (b) Inferential

(2) **Techniques for using sample information to draw dependable conclusions about the population are gathered within the branch of statistics called _____ statistics.**

    (a) descriptive
    (b) inferential

(3) **Which of the following statements are false?**

    (a) A variable is any characteristic whose value may change from one object to another in the population.
    (b) All companies listed on the New York Stock Exchange is an example of a sample.
    (c) sampling frame is a listing of the individuals or objects to be sampled.
    (d) "All possible yields (in grams) from a certain chemical reaction carried out in a laboratory" is an example of conceptual or hypothetical population.

(4) **Which of the following statements regarding histograms are correct?**

    (a) A unimodal histogram is one that rises to a single peak and then declines, whereas a bimodal histogram is one that has two different peaks.
    (b) A unimodal histogram is positively skewed if the right or upper tail is stretched out compared to the left or lower tail.
    (c) A unimodal histogram is negatively skewed if the left or lower tail is stretched out compared to the right or upper tail.
    (d) A histogram is symmetric if the left half is a mirror image of the right half.
    (e) All of the above.

(5) **Which of the following statements are false?**

    (a) A physical interpretation of the sample mean $\bar{x}$ demonstrates how it measures the location (center) of a sample.
    (b) The sample median is the middle value when the observations are ordered from smallest to largest.
    (c) The sample median is very sensitive to extremely small or extremely large data values (outliers).

(d) The sample mean is very sensitive to extremely small or extremely large data values (outliers).

(6) **Boxplots have been used successfully to describe**

(a) center of a data set
(b) spread of a data set
(c) the extent and nature of any departure from symmetry
(d) identification of outliers
(e) All of the above

(7) **A sample of 24 offshore oil workers took part in a simulated escape exercise, resulting in the accompanying data on time (sec) to complete the escape:**

373 370 364 366 364 325 339 393
356 359 363 375 424 325 394 402
392 369 374 359 356 403 334 397

The sample mean and standard deviation are

(a) $\bar{x} = 369.833$ and $s_x = 20.234$
(b) $\bar{x} = 350.012$ and $s_x = 25.087$
(c) $\bar{x} = 369.833$ and $s_x = 25.087$

(8) **In a stem-and-leaf display, the _____ digit(s) is (are) the stem(s), whereas the _____ digit(s) is (are) the leaf (or leaves).**

(a) leading, trailing
(b) trailing, leading

(9) **The accompanying data specific gravity values for various wood types used in construction.**

.41 .41 .42 .42 .42 .42 .42 .43 .44
.54 .55 .58 .62 .66 .66 .67 .68 .75
.31 .35 .36 .36 .37 .38 .40 .40 .40
.45 .46 .46 .47 .48 .48 .48 .51 .54

From the stem-and-leaf display we can infer that the data in consideration is:

(a) symmetric
(b) positively skewed
(c) negatively skewed

(10) **For the following data we have constructed a histogram with 8 bins (classes) while the first class starting at 60:**

75 89 80 93 64 67 72 70 66 85
89 81 81 71 74 82 85 63 72 81

```
81 95 84 81 80 70 69 66 60 83
85 98 84 68 90 82 69 72 87 88
```

The highest frequency in the constructed histogram is:

(a) 15
(b) 12
(c) 13
(d) 9
(e) 10

(11) **A sample of n $= 10$ automobiles was selected, and each was subjected to a 5-mph crash test. Denoting a car with no visible damage by S (for success) and a car with such damage by F, results were as follows: S S S F F S S F S S. What is the value of the sample proportion of successes x/n?**

(a) 0.30
(b) 0.70
(c) 0.73
(d) 0.43
(e) 0.57

(12) **Suppose that the serum levels of total cholesterol y for a sample of adult males is bell-shaped with mean 180 and standard deviation 10. The distribution is sketched below.**
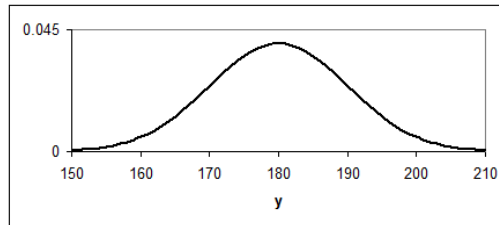
Approximately what percentage of sample males have y-values between 160 and 210?

(a) 50.0%
(b) 84.0%
(c) 68.0%
(d) 1.5%
(e) 98%

(13) **Mr. Smith is told that his cholesterol level y is located 2.5 standard deviations below the mean, that is, has a z-score $= $ -2.5. What is his cholesterol level? (you may need to use some information from the previous question)**

(a) 160
(b) 150
(c) 205
(d) 155

(e) None of the above

(14) sample of **20** glass bottles of a particular type was selected, and the internal pressure strength of each bottle was determined. Consider the following partial sample information:
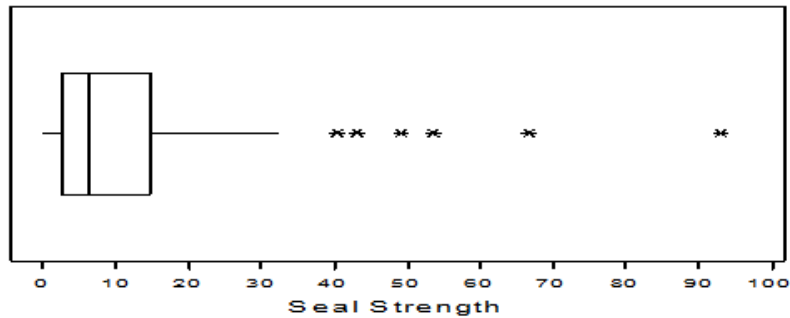
```
Q1 = 196.0
Median = 202.2
Q3 = 216.8
Three smallest observations 125.8 188.1 193.7
Three largest observations 221.3 230.5 250.2
```
Are there any outliers in the sample?

(a) Yes
(b) No
(c) There is insufficient information to answer the question

(15) **Below is a boxplot of the seal strengths of n $= 100$ sample potato chip bags.**

Seal Strength

Approximately, what percentage of bags had seal strengths below 14?

(a) 50%
(b) 40%
(c) 75%
(d) 25%
(e) 5%

(16) **Approximately what percentage of bags had seal strengths above 3?**

(a) 3%
(b) 25%
(c) 50%
(d) 75%
(e) 95%

(17) **The distribution of scores (from the boxplot in the previous question) is best described as**

    (a) symmetric about its mean
    (b) skewed to the right
    (c) skewed to the left

(18) **The Central Limit Theorem says the sampling distribution of the sample mean is approximately normal under certain conditions. What is a necessary condition for the Central Limit Theorem to be used?**

    (a) The population size must be large (e.g., at least 30).
    (b) The population from which we are sampling must not be normally distributed.
    (c) The population from which we are sampling must be normally distributed.
    (d) The sample size must be large (e.g., at least 30).

(19) **Which of the following statements about the sampling distribution of the sample mean is incorrect?**

    (a) The standard deviation of the sampling distribution is σ.
    (b) The mean of the sampling distribution is μ.
    (c) the sampling distribution is approximately normal whenever the sample size is sufficiently large ($n \geq 30$).
    (d) The sampling distribution is generated by repeatedly taking samples of size n and computing the sample means.

(20) **The compressive strength of concrete is normally distributed with $\mu = 2500$ psi and $\sigma = 50$ psi. Find the probability that a random sample of $n = 9$ specimens will have a sample mean diameter between 2499 psi to 2510 psi?**

    (a) 0.0872
    (b) 0.4128
    (c) 0.7503
    (d) 0.2497
    (e) 0.2503

(21) **A random sample of 64 observations produced a sample mean value of 82 and standard deviation of 5.5. The 90% confidence interval for the population mean μ is between**

    (a) 81.861 and 82.140
    (b) 80.654 and 83.351
    (c) 80.852 and 83.148
    (d) 81.312 and 82.691
    (e) None of the above answers are correct.

(22) **A random sample of 100 observations produced a sample proportion of .25. An approximate 90% confidence interval for the population proportion p is**

    (a) .248 and .252
    (b) .179 and .321

(c) .423 and .567
(d) .246 and .254
(e) None of the above answers are correct.

(23) **A 2001 survey found that the American family generates an average of 17.2 pounds of glass garbage each year. Assume the standard deviation of the distribution is 2.5 pounds, and that garbage quantity is approximately normally distributed. A researcher wishes to update this study for 2007. How large should his sample be to estimate the mean weight of garbage per family to within .1 pounds with 99% confidence?**

(a) 1692
(b) 2401
(c) 4161
(d) 5026
(e) Need more information.

(24) **The State News wants to know, within 10%, what proportion of the 48,000 UNM students approve of the new football coach. Their reporter will use a simple random sample and a 90% confidence interval. What is the minimum acceptable sample size?**

(a) 67
(b) 68
(c) 96
(d) 97
(e) 4,800

(25) **The objective of _____ is to select a single number such as $\bar{x}$ or s_x^2 , based on sample data, that represents a sensible value (good guess) for the true value of the population parameter, such as μ or σ^2 .**

(a) tree diagram.
(b) confidence interval.
(c) venn diagram.
(d) point estimation.
(e) contingency table.

(26) **If a confidence level of 90% is used to construct a confidence interval for the mean μ of a normal population when the value of the standard deviation σ is known, the z critical value is:**

(a) 1.645
(b) 1.96
(c) 2
(d) 2.58
(e) 2.33

(27) **Which of the following statements is true?**

(a) A confidence interval is always calculated by first selecting a confidence level, which is a measure of the degree of reliability of the interval.

(b) A confidence level of 95% implies that 95% of all samples would give an interval that includes the parameter being estimated, and only 5% of all samples would yield an erroneous interval.

(c) Information about the precision of an interval estimate is conveyed by the width of the interval.

(d) The higher the confidence level, the more strongly we believe that the value of the parameter being estimated lies within the interval.

(e) All of the above statements are true.

(28) **An airline has surveyed a simple random sample of air travelers to find out whether they would be interested in paying a higher fare in order to have access to email during their flight. Of the 400 travelers surveyed, 80 said email access would be worth a slight extra cost. Construct a 95% confidence interval for the population proportion of air travelers who are in favor of the airlines email idea.**

(a) $(14.85\%, 25.15\%)$
(b) $(16.71\%, 23.29\%)$
(c) $(17.12\%, 22.88\%)$
(d) $(16.08\%, 23.92\%)$
(e) None of the above.

(29) **Researchers have estimated that office workers in Germany receive an average of 15.0 fax messages per day. Assuming this finding to be based on a simple random sample of 80 German office workers, with a sample standard deviation of $s_x = 3.5$ messages, the constructed 90% confidence interval for the true population mean is:**

(a) $(14.35, 15.65)$
(b) $(14.49, 15.51)$
(c) $(14.22, 15.78)$
(d) $(13.97, 16.03)$
(e) None of the above.

(30) **A state politician would like to determine the average amount earned during summer employment bt state teenagers during the past summer's vacation period. She wants to have 95% confidence that the sample mean is within \$50 of the true population mean. Based on past studies, she has estimated the population standard deviation to be $\sigma = \$400$. What sample size is necessary to have 95% confidence that $\bar{x}$ will be within \$50 of the true population mean?**

(a) 245
(b) 246
(c) 240
(d) 200
(e) None of the above.

(31) **A tourist agency researcher would like to determine the proportion of U.S. adults who have ever vacationed in Mexico and wishes to be 99% confident that the sampling error will be no more than 0.03. Assuming the researcher has no idea regarding the true value of the population proportion, what sample size us necessary to have 99% confidence that the sample proportion**

will be within 0.03 of the actual population proportion?

(a) 950
(b) 1068
(c) 1849
(d) 1067
(e) None of the above.

(32) **A tourist agency researcher would like to determine the proportion of U.S. adults who have ever vacationed in Mexico and wishes to be 95% confident that the sampling error will be no more than 0.03. If the researcher believes the population proportion is no more than 0.3, and uses $\hat{p} = 0.3$ as the estimate, what sample size will be necessary?**

(a) 897
(b) 896
(c) 1067
(d) 1068
(e) None of the above.

(33) **In many situations, the _____ hypothesis is referred to as the research hypothesis, since it is the statement the researcher would really like to validate. Furthermore, in our treatment of hypothesis testing, the _____ hypothesis will always be stated as an equality claim.**

(a) null, alternative
(b) null, null
(c) alternative, null
(d) alternative, alternative

(34) **A _____ error involves not rejecting the null hypothesis H_0 when H_0 is false, while a _____ error consists of rejecting the null hypothesis H_0 when H_0 is true.**

(a) type I, type II
(b) type I, type I
(c) type II, type I
(d) type II, type II

(35) **The probabilities of type I and type II errors are traditionally denoted by the Greek letters _____ and _____, respectively. Furthermore, a _____ error is usually more serious than a _____ error.**

(a) α, $1 - \beta$, type I, type II
(b) α, β, type II, type I
(c) α, β, type I, type II
(d) β, α, type I, type II
(e) $1 - \beta$, α, type I, type II

(36) **If the P-value is smaller than or equal to the level of significance α , then the researcher should** _____ **at level α. Furthermore, If the P-value is larger than the level of significance α , then the researcher should** _____ **at level α.**

    (a) reject H_0, fail to reject H_0
    (b) fail to reject H_0, reject H_0
    (c) reject H_0, fail to reject H_a
    (d) fail to reject H_a, reject H_0
    (e) reject H_a, fail to reject H_a

(37) **Which of the following statements is not true?**

    (a) A statistical hypothesis is a claim or assertion either about the value of a single parameter, about the values of several parameters, or about the form of an entire probability distribution.
    (b) In any hypothesis-testing problem, there are two contradictory hypotheses under consideration.
    (c) A test of hypothesis is a method for using sample data to decide whether the null hypothesis should be rejected.
    (d) A type I error consists of not rejecting the null hypothesis H_0 when H_0 is false.

(38) **Which of the following statements is true?**

    (a) In general, the null hypothesis H_0 is a hypothesis which the researcher tries to disprove, reject or nullify.
    (b) The alternative hypothesis, denoted by H_a, is the assertion that is contradictory to the null hypothesis H_0.
    (c) The null hypothesis H_0 will be rejected in favor of the alternative hypothesis only if sample evidence suggests that H_0 is false.
    (d) If sample evidence does not strongly contradict the null hypothesis H_0 , we will continue to believe in the truth of H_0.
    (e) All of the above statements are true.

(39) **Which of the following statements is not correct?**

    (a) It is possible that the null hypothesis may be rejected when it is true.
    (b) It is impossible that the null hypothesis may be rejected when it is true.
    (c) It is possible that the null hypothesis may not be rejected when it is false.

(40) **In hypothesis-testing analysis, a type I error occurs only if**

    (a) the null hypothesis is rejected when it is true
    (b) the null hypothesis is rejected when it is false
    (c) the null hypothesis is not rejected when it is false
    (d) the null hypothesis is not rejected when it is true

(41) **In hypothesis-testing analysis, a type II error occurs only if**

    (a) the null hypothesis is rejected when it is true
    (b) the null hypothesis is rejected when it is false
    (c) the null hypothesis is not rejected when it is false

(d) the null hypothesis is not rejected when it is true

(42) **Which of the following statements is not true?**

    (a) A P-value conveys much information about the strength of evidence against the null hypothesis H_0 and allows an individual decision maker to draw a conclusion at any specified significance level α.

    (b) The P-value (or observed significance level) is the largest level of significance at which the null hypothesis H_0 would be rejected when a specified test procedure is used on a given data set.

    (c) If P-value $\leq \alpha$, we reject H_0 at level α.

    (d) If P-value $> \alpha$, we do not reject H_0 at level α.

    (e) All of the above statements are true.

(43) **Suppose that when data from an experiment was analyzed, the P-value for testing $H_0 :\mu \leq 50$ versus $H_a :\mu > 50$ was calculated as .0244. Which of the following statements is true?**

    (a) H_0 is rejected at .10 level.

    (b) H_0 is not rejected at .05 level.

    (c) H_0 is not rejected at .025 level.

    (d) H_0 is rejected at any level α.

    (e) All of the above statements are true.

(44) **Which of the following P-values will lead us to reject the null hypothesis at the .05 level?**

    (a) .10

    (b) .025

    (c) .075

    (d) .15

    (e) Any P-value greater than .05

(45) **Which of the following statements is not true?**

    (a) The reliability of hypothesis testing procedure in reaching a correct decision can be assessed by studying type I error probability.

    (b) The process of reaching a decision by using the methodology of classical hypothesis testing involves selecting a level of significance α and then rejecting or not rejecting the null hypothesis H_0 at that level α.

    (c) A small P-value would indicate statistical significance in that it would strongly suggest rejection H_0 of in favor of H_a.

    (d) A test with significance level α is one for which the type I error probability is controlled at the specified level.

    (e) A type I error is usually more serious than a type II error.

(46) **An economist believes that the mean debt carried by a new college graduate is at least $\$17,000$. Appropriate null and alternative hypotheses for testing this claim are:**

    (a) H_0: $\mu > \$17,000$, H_a: $\mu \leq \$17,000$

    (b) H_0: $\mu = \$17,000$, H_a: $\mu \leq \$17,000$

    (c) H_0: $\mu \leq \$17,000$, H_a: $\mu > \$17,000$

(d) H_0: $\mu \geq \$17,000$, H_a: $\mu < \$17,000$

(e) H_0: $\mu < \$17,000$, H_a: $\mu > \$17,000$

(47) **In carrying out a test of the claim in the previous question (at the .05 level) a researcher obtains a P-Value of .13. An appropriate formulation of her conclusion is**

(a) The data do not support the claim that the average debt carried by a recent college graduate is at least $\$17,000$.

(b) The data support the claim that the debt carried by a recent college graduate is at least $\$17,000$.

(c) The null hypothesis is rejected.

(d) We fail to reject the null hypothesis.

(48) **A physical therapist believes that the average weight of a backpack carried by an MSU sophomore on Monday mornings is at least 19 pounds. In carrying out her hypothesis test at the .05 level, the therapist obtains a p-value of .037. An appropriate formulation of her conclusion is**

(a) The data support the claim that the average weight of a backpack carried by an MSU sophomore on Monday mornings is at least 19 pounds.

(b) The data do not support the claim that the average weight of a backpack carried by an MSU sophomore on Monday mornings is at least 19 pounds.

(49) **A statistician carries out a hypothesis test at the 10% level for which he rejects the null hypothesis. He computes that the power of the test was 87%. What is the chance that he will have reached the wrong conclusion?**

(a) .05

(b) .10

(c) .13

(d) .87

(e) .90

(50) **A statistician carries out a hypothesis test at the 0.05 level. He computes that the power of the test was 87% against an alternative of interest. What is the chance that he will reach the right conclusion, if the alternative of interest is true?**

(a) 5%

(b) 95%

(c) 13%

(d) 87%

(e) 18%