# PRACTICE STATISTICS FINAL EXAM

1. The following are test scores on a biology exam: 64, 82, 87, 76, 93, 68, 71, 68, 75, 91, 84, 89. Calculate: mean, median, mode, range, standard deviation, Q1, Q3; make the Box plot.

2. The following speeds (in miles per hour) were measured by a radar device on a city street:
   36, 34, 35, 47, 59, 33, 40, 42, 31, 37, 30, 32, 25, 29, 28, 33, 52, 32, 42, 28, 30, 45, 30, 47.
   Construct a grouped frequency distribution using seven classes (begin with 25, use a class width of 5); construct the histogram; determine the measures of central tendency (mode, median, mean); describe the shape of the histogram (skewed left, skewed right or symmetric).

3. Classify each of the following as quantitative or qualitative (categorical): hair color; salary; GPA. Indicate the level of measurement for each.

4. Determine which measure of the center (mean, median, mode) which would best describe:
   a. The most popular boy's name in the U.S.
   b. The salary of employees at a doctor's office (one doctor, two nurses, one secretary and one bookkeeper)
   c. The average speed of cars on the New York Thruway.

5. In the standard normal distribution:
   a. Find the area to the right of z = -.58
   b. Find the area between z = -.36 and z =1.49
   c. Find the z value that has area to the left equal to .0307
   d. Find the z value that has area to the right equal to .0043

6. Scores on an IQ test for a group of high school students were approximately normally distributed with mean of 100 and standard deviation of 10.
   a. What percent of the students have IQ scores above 100?
   b. What percent of the students have IQ scores between 80 and 120?
   c. What is the probability that a student has an IQ score below 95?
   d. What is the probability that a student has an IQ score above 125?
   e. What is the probability that a student has an IQ score between 82 and 116?
   f. How many standard deviations from the mean is 113?
   g. What scores lie three standard deviations from the mean?
   h. What minimum IQ score would a student need to get to have an IQ in the top 10%?

7. A study was conducted to determine if the number of years of education a person completes after high school will affect their starting salary. The results, and corresponding summary statistics, were:

| # years of education (after high school) | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Mean 3.5 | St. dev. 2.45 |
|---|---|---|---|---|---|---|---|---|---|---|
| Starting salary | $11,250 | 11,900 | 14,800 | 15,500 | 26,000 | 28,350 | 29,400 | 36,700 | 21,737 | 9,549 |

   a. What is the explanatory variable? What is the response variable?
   b. Construct the scatterplot; describe the relationship.
   c. The correlation was determined to be .9674. Determine the slope, intercept and the equation of the least squares regression line. Add this line to the scatterplot. Interpret the slope.
   d. Predict the earnings of a person with four years of education (beyond high school). What would be the residual for this person?

8. If one fair die is thrown, what is the probability of getting: Three or more? Less than two? More than six?

9. A survey was conducted to determine how many high school juniors and seniors smoke. Results:

| | JUNIORS | SENIORS | TOTAL |
|---|---|---|---|
| Smoke | 88 | 142 | |
| Don't smoke | 107 | 143 | |
| TOTAL | | | |

Complete the table. Then if a student is chosen at random, what is the probability that the student:

a. Is a junior?
b. Smokes?
c. Is a junior and smokes?
d. Is a senior or does not smoke?
e. Is a senior given that they smoke?
f. Is a junior given that they do not smoke?

10. X is a normally distributed random variable with a mean of 70 and a standard deviation of 6. If a sample of size 36 is selected, what is the mean (expected value) of the sample mean? What is the standard deviation of the sample mean?

11. A can of soda has an average weight of 8 ounces with a standard deviation of 1.46 (assume a normal distribution).
    a. What is the probability you weigh one can and it contains more than 7.5 ounces?
    b. What is the probability you weigh one can and it is between 7.6 and 8.4 ounces?
    c. What is the probability you weigh 50 cans and the mean is less than 7.8 ounces?

12. The weights of full boxes of Jell-O are normally distributed with a standard deviation of 0.75 grams. A sample of 25 randomly selected boxes had a mean of 85 grams.
    a. What is the 95% margin of error?
    b. Find the 95% confidence interval for the true mean weight of a box of Jell-O.
    c. How many boxes must you sample to estimate the mean with 99% confidence if the margin of error is 0.2?

13. A company states that their customers wait an average of 4 minutes to be helped with a standard deviation of 1.2 minutes. You believe that the waiting time is longer. You randomly sample 36 customers and determine that the average wait time for this sample is 4.6 minutes. State the null hypothesis and the alternative hypothesis. Calculate the test statistic. Determine the P-value. At the 5% level of significance, state your conclusion.

14. Determine the binomial distribution when $n = 5$ and $p = 0.4$. Determine this distribution's mean and standard deviation. Describe the distribution's shape. What is the probability of at least one success?

15. Consider the binomial distribution with $n = 100$, $p = .7$. What is the probability that X is between 65 and 80?

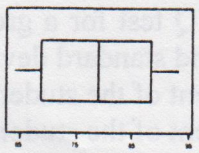
16. Review the vocabulary and the examples of observational studies and experiments described in Chapter 1 of the text. Be ready to comment on some specific situations (i.e., identify various sources of errors and possible bias issues).

# ANSWERS: PRACTICE STATISTICS FINAL EXAM

1. The following are test scores on a biology exam: 64, 82, 87, 76, 93, 68, 71, 68, 75, 91, 84, 89.
Calculate: mean, median, mode, range, standard deviation, Q1, Q3; make the Box plot.

| X | X-mean | SQ |
|---|---|---|
| 93 | 14 | 196 |
| 91 | 12 | 144 |
| 89 | 10 | 100 |
| 87 | 8 | 64 |
| 84 | 5 | 25 |
| 82 | 3 | 9 |
| 76 | -3 | 9 |
| 75 | -4 | 16 |
| 71 | -8 | 64 |
| 68 | -11 | 121 |
| 68 | -11 | 121 |
| 64 | -15 | 225 |
| Sum=948 | Sum=0 | Sum=1094 |

N=12
Mean=948/12=79
Median=(76+82)/2=79
Mode=68
Range=93-64=29
S=sqrt(1094/11)=9.97
Q1=(68+71)/2=69.5
Q3=(87+89)/2=88
IQR=88-69.5=18.5
5# summary:
64, 69.5, 79, 88, 93



(plot by hand; equal interval scale on horizontal)

2. The following speeds (in miles per hour) were measured by a radar device on a city street:
36, 34, 35, 47, 59, 33, 40, 42, 31, 37, 30, 32, 25, 29, 28, 33, 52, 32, 42, 28, 30, 45, 30, 47.
Construct a grouped frequency distribution using seven classes (begin with 25, use a class width of 5); construct the histogram; determine the measures of central tendency (mode, median, mean); describe the shape of the histogram (skewed left, skewed right or symmetric).

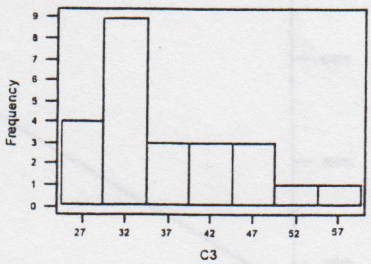
Note: if you are asked for measures after the grouping, then give grouped measures

| Speed | Frequency | Midpoint | Xf |
|---|---|---|---|
| 25-29 | 4 | 27 | 108 |
| 30-34 | 9 | 32 | 288 |
| 35-39 | 3 | 37 | 111 |
| 40-44 | 3 | 42 | 126 |
| 45-49 | 3 | 47 | 141 |
| 50-54 | 1 | 52 | 52 |
| 55-59 | 1 | 57 | 57 |
| N= | 24 | | Sum=883 |

**Grouped measures:**
Mean=883/24 =36.79
Median=32
Mode=32

**Ungrouped:**
Mean=36.54
Median=33.5
Mode=30

Shape: skewed right
Plot by hand using the midpoints:




3. Classify as quantitative or qualitative; identify level of measurement: hair color; salary; GPA.
Hair color is qualitative, nominal level; the other two are quantitative, ratio level.

4. Determine which measure of the center (mean, median, mode) which would best describe:
   a. The most popular boy's name in the U.S. Mode
   b. The salary of employees at a doctor's office (one doctor, two nurses, one secretary and one bookkeeper) Median
   c. The average speed of cars on the New York Thruway. Mean

Updated: 6/08

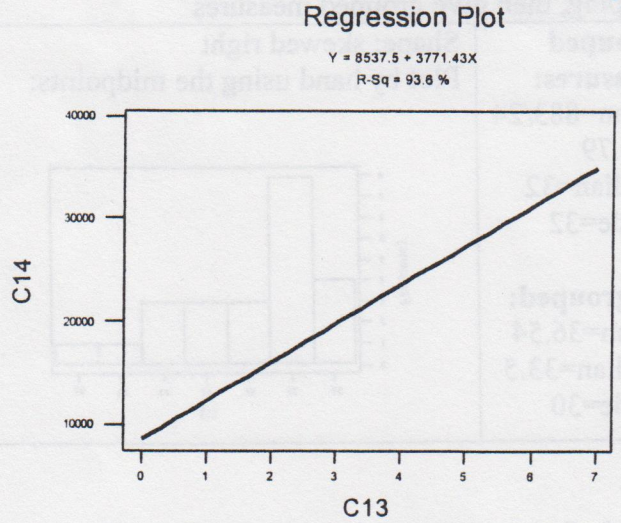
5. In the standard normal distribution:
   a. Find the area to the right of z = -.58. p = 1 - .2810 = .7190
   b. Find the area between z = -.36 and z =1.49 p = .9319 - .3594 = .5725
   c. Find the z value that has area to the left equal to .0307 z = -1.87
   d. Find the z value that has area to the right equal to .0043 z = 2.63

6. Scores on an IQ test for a group of high school students were approximately normally distributed with mean of 100 and standard deviation of 10.
   a. What percent of the students have IQ scores above 100? 50%
   b. What percent of the students have IQ scores between 80 and 120? 95%
   c. What is the probability that a student has an IQ score below 95? z = -.5, p = .3085
   d. What is the probability that a student has an IQ score above 125? z = 2.5, p = 1 - .9938 = .0062
   e. What is the probability that a student has an IQ score between 82 and 116? -1.8 < z < 1.6; p = .9452 - .0359
   f. How many standard deviations from the mean is 113? z = 1.3
   g. What scores lie three standard deviations from the mean? 100 ± 3(10) = 70 and 130
   h. What minimum IQ score would a student need to get to have an IQ in the top 10%? z = 1.28, X = 100 + 1.28(10) = 112.8 (would answer 113)

7. A study was conducted to determine if the number of years of education a person completes after high school will affect their starting salary. The results, and corresponding summary statistics, were:

| # years of education (after high school) | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Mean 3.5 | St. dev. 2.45 |
|---|---|---|---|---|---|---|---|---|---|---|
| Starting salary | $11,250 | 11,900 | 14,800 | 15,500 | 26,000 | 28,350 | 29,400 | 36,700 | 21,737 | 9,549 |

   a. What is the explanatory variable? X=# years What is the response variable? Y=salary $
   b. Construct the scatterplot; describe the relationship. Linear, positive, strong




   c. The correlation was determined to be .9674. Determine the slope, intercept and the equation of the least squares regression line. Add this line to the scatterplot. Interpret the slope.
   Slope=r sy/sx = 3771.43 (for every one year increase in schooling, salary increases $3771.43.)
   Intercept = Ybar-slope*Xbar = 21737 – 3771.43*3.5 = 8537.5 (predict salary of $8537.50 with 0 years of schooling beyond high school)
   Equation: Ypred=8537.50 + 3771.43 X
   d. Predict the earnings of a person with four years of education (beyond high school).
   @X=4, Ypred=8537.5+3771.43*4=$23,623 (rounded); error (residual) = $26,000-23,623 = $2377

8. If one fair die is thrown, what is the probability of getting: Three or more? 4/6 Less than two? 1/6 More than six? 0

9. A survey was conducted to determine how many high school juniors and seniors smoke. Results:

| | JUNIORS | SENIORS | TOTAL |
|---|---|---|---|
| Smoke | 88 | 142 | 230 |
| Don't smoke | 107 | 143 | 250 |
| TOTAL | 195 | 285 | 480 |

Complete the table. Then if a student is chosen at random, what is the probability that the student:

a. Is a junior? 195/480 = .40625 = 40.625%
b. Smokes? 230/480 = 47.92%
c. Is a junior and smokes? 88/480 =18.33%
d. Is a senior or does not smoke? 1 - 88/480 = 392/480 = (285+250-143)/480 = 81.67%
e. Is a senior given that they smoke? 142/230 = 61.74%
f. Is a junior given that they do not smoke? 107/250 = 42.8%

10. X is a normally distributed random variable with a mean of 70 and a standard deviation of 6. If a sample of size 36 is selected, what is the mean (expected value) of the sample mean? What is the standard deviation of the sample mean? Mean = 70, stdev of mean = $\sigma/\sqrt{n} = 6/\sqrt{(36)} = 6/6 = 1$

11. A can of soda has an average weight of 8 ounces with a standard deviation of 1.46 (assume a normal distribution).
    a. What is the probability you weigh one can and it contains more than 7.5 ounces? z = (7.5 - 8)/1.46 = -.34; p = 1 - .3669 = .6331
    b. What is the probability you weigh one can and it is between 7.6 and 8.4 ounces? z = (7.6 - 8)/1.46 = -.27; z = (8.4 - 8)/1.46 = +.27; area between = .6064 - .3936 = .2128
    c. What is the probability you weigh 50 cans and the **mean** of the cans is less than 7.8 ounces? now use z = (7.8 - 8)/[1.46/sqrt(50)] = -.2/.2065 = -.97; p = .1660

12. The weights of full boxes of Jell-O are normally distributed with a standard deviation of 0.75 grams. A sample of 25 randomly selected boxes had a mean of 85 grams.
    a. What is the 95% margin of error? Here, $z_c = 1.96$ (from table), $\sigma = .75$, and n = 25. $E = z_c * \sigma /\sqrt{(n)} = [1.96 * .75/\sqrt{(25)}] = .29$
    b. Find the 95% confidence interval for the true mean weight of a box of Jell-O. CI: xbar $\pm$ E = 85 $\pm$ .29; answer: 84.71 to 85.29
    c. How many boxes must you sample to estimate the mean with 99% confidence if the margin of error is 0.2? Here, $z_c = 2.576$ (from table), $\sigma = .75$, and the desired E = .2; use formula: $n = [z_c * \sigma /E]^2 = [2.576 * .75/.2]^2 = 9.66^2 = 93.32$; round up (always) so answer is n = 94

13. A company states that their customers wait an average of 4 minutes to be helped with a standard deviation of 1.2 minutes. You believe that the waiting time is longer. You randomly sample 36 customers and determine that the average wait time for this sample is 4.6 minutes. State the null hypothesis and the alternative hypothesis. Calculate the test statistic. Determine the P-value. At the 5% level of significance, state your conclusion.

**Hypotheses:** Ho: □ = 4 Ha: □ > 4. (upper tailed test); $\alpha = .05$
**Test statistic:** $z = (xbar - □)/[\sigma /\sqrt{(n)}] = (4.6 - 4)/[1.2/\sqrt{(36)}] = .6/[1.2/6] = .6/.2 = 3$
**P-value** = 1 - .9987 = .0013. (from table, area to the right of z = 3)
**Decision:** Since P-value is lower than $\alpha$ (.0013 is less than level of significance, $\alpha = .05$), reject Ho

**Conclusion**: There is sufficient evidence to conclude that the average wait time is longer than 4 minutes.

14. Determine the binomial distribution when n = 5 and p = 0.4. Determine this distribution's mean and standard deviation. Describe the distribution's shape. What is the probability of at least one success?

First, get q = 1- p = .6

Use the binomial formula to compute the probabilities: $P(r) = nCr * p^{r} * q^{(n-r)}$
(Usually round to three decimal places.)

| r | P(r) |
|---|---|
| 0 | 0.078 |
| 1 | 0.260 |
| 2 | 0.346 |
| 3 | 0.230 |
| 4 | 0.078 |
| 5 | 0.010 |

(Probabilities should add to 1; may have some round-off error, as shown here.)

Histogram is skewed to the right since p < .5. Mean = np = 5(.4) = 2; stdev = $\sqrt{[np(1 - p)]} = \sqrt{[5*.4*.6]} = \sqrt{1.2} = 1.095$.

$P(r \geq 1) = 1 - P(r = 0) = 1 - .078 = .922$ or $P(r \geq 1) = P(r = 1) + P(r = 2) + P(r = 3 + P(r = 4) + P(r = 5) = 0.260 + 0.346 + 0.230 + 0.077 + 0.010 = .924$ (either answer accepted; difference due to round-off error)

15. Consider the binomial distribution with n = 100, p = .7. What is the probability that X is between 65 and 80?

q = 1 – p = .3. First check the assumption that np = 70 and nq = 30 are both greater than 5.

Then can use the normal approx to the normal: mean = np = 70; stdev = $\sqrt{[npq]} = \sqrt{21} = 4.58$

Compute the two z values; remember to correct for continuity:
$z_1 = (80.5 - 70)/4.58 = 10.5/4.58 = 2.29$; $z_2 = (64.5 - 70)/4.58 = -5.5/4.58 = -1.20$

Look up both z values in the Normal Table and subtract the areas: p = .9890 - .1151 = .8739.

16. Review the vocabulary and the examples of observational studies and experiments described in Chapter 1 of the text. Be ready to comment on some specific situations (i.e., identify various sources of errors and possible bias issues). You should review the Chapter 1 vocabulary list and the text.