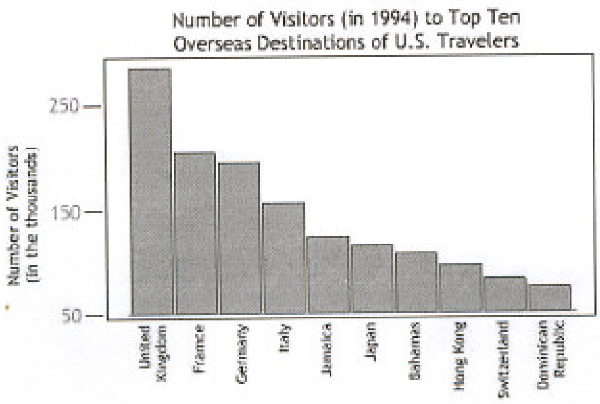
360 Possible Points.

Elements of Statistics (Math 106) - Final Exam **Name** Brad Hartlaub - Solutions
Fall 2002

Part I. True or False. Each problem is worth 5 points.

T F The bar chart shows the numbers of 1994 visitors to the top ten overseas destinations of U.S. travelers. There were just as many travelers to the *top two* destinations as there were to the next eight.



275 + 200 = 475 < 190 + 180 + 150 + 125 + ...

T F 50% of the data in a boxplot lie in the interquartile range.

T F Set A appears to be bimodal. (See graph below.)

Set A		Set B
5	1	2
1 0	2	3 4
9 5 4	3	0 0 4 5 8 8
8 7 7 3	4	0 6
	5	7 7 9 9 9 9
8	6	5

$\frac{38+40}{2} = 39$

$\tilde{x} = 39$ for both sets.

T F Set A and set B have the same median. (See graph above)

T F Simpson's paradox occurs only in those tables that when combined exhibit a different ordering of the conditional probabilities.

T F An investigator wishes to select a sample of students at a particular high school. He decides to choose the first 50 students who enter the cafeteria during the late lunch period. The resulting sample can be considered a *simple random sample* from the population of students at this school.

T F In stratified random sampling, individuals are selected at random from only a few randomly selected strata.

T F The reason that stratified random sampling is preferred over simple random sampling is because it reduces the risk of selection bias.

T F A doctor wishes to study the relationship between exercise and cholesterol level. Thirty subjects available for the study are randomly divided into two groups. One group will not exercise while the second group will exercise three times a week. Blood cholesterol levels will be measured at the beginning and at the end of the 3-month study. Such a study is called an *experiment*.

T F Randomization is important in experiments because it reduces the chance that a lurking variable will confound the results.

T F To test the effectiveness of a pre-wash stain treatment, 20 pieces of white cotton cloth are stained with various products such as ketchup, mustard, oil, and dirt. Half of each cloth is treated with a pre-wash stain treatment, and the other half is left untreated. Which half gets the treatment is determined by a flip of a coin. The 20 pieces of cloth are washed, and the difference in whiteness between the treated and untreated halves is evaluated. This is an example of a completely randomized design.

Block design

T F A mechanical engineer is studying the force developed by a drill press. She suspects that the rate at which material is fed to the press may be an important factor. She has 4 feed rates. There are 8 operators who normally use the drill press. Each operator uses the drill press at each of the 4 feed rates. The order of the feed rates is randomized for each operator. This is an example of a randomized complete block design.

T F If a discrete random variable, X, has the following probability distribution, then $P(X < 3) = 0.4$.

X	0	1	2	3	4
Prob	0.25	0.10	0.05	0.10	0.50

$.25 + .10 + .05 = .4$

T F If a random variables Z is the difference of 2 independent random variables, P and Q; Z = P - Q and P and Q are independent, then the standard deviation of Z is the sum of the standard deviations of P and Q.

T F The total score obtained in the SAT test is a random variable described by adding the total of the 2 components of the test, the verbal and the math. The variance of the total can be found by adding the variances of each of the two components.

M & V scores are not indep.

T F If the random variable X follows a $N(\mu, \sigma)$ distribution, then for any specific value x the area to the right of $\mu + x$ is the same as the area to the left of $\mu - x$ because of symmetry.

T F If a statistic is a biased estimator of a population parameter of interest, increasing the sample size results in less bias.

Part II. Multiple Choice. Each problem is worth 5 points.

1. Here are the body and backpack weights of the eight members of Summit Squad, a group of backpackers from Walker High.

Predictor　　　Response

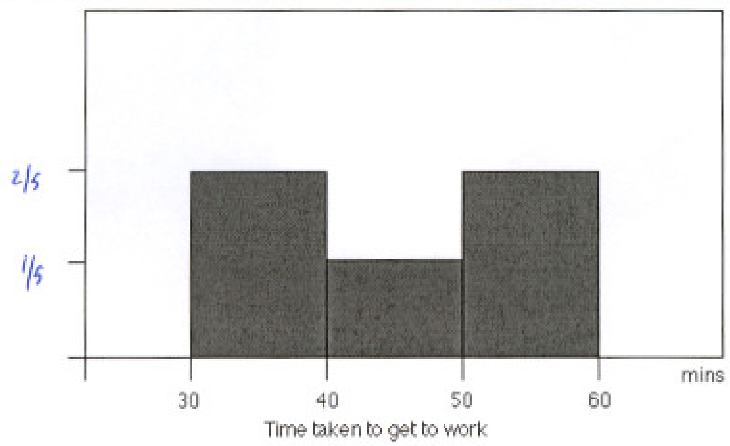
Name	Body Weight (lb.)	Backpack Weight (lb.)
Amble	120	26
Plodder	187	30
Slog	109	26
Cliff	103	24
Saunter	131	29
Thumper	165	35
Belay	158	31
Rocky	116	28

Perform a least squares regression analysis. Which student has the largest residual?

a. Amble
b. Plodder
c. Cliff
d. Saunter
e. Thumper

e

2. The time it takes to get to work is distributed according to the following distribution.



$2x + x + 2x = 1$
$5x = 1$
$x = \frac{1}{5}$

What is the probability that it takes less than 40 minutes to get to work?

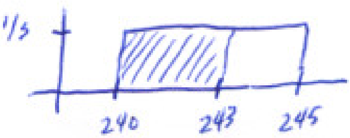
a. 0
b. 0.333
c. 0.4
d. 0.5
e. 0.6

c

3. Application temperatures for a manufacturing process that seals bags vary uniformly from 240 to 245°° F. What is the probability that the application temperature for a randomly selected bag will be at most 243°° F?

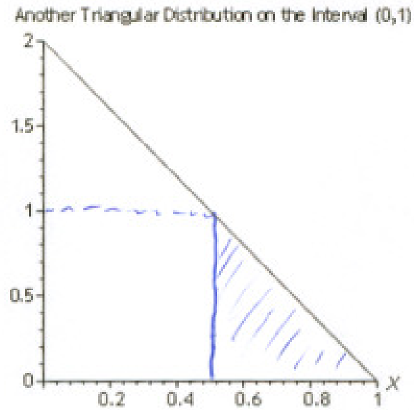
a. 0.2
b. 0.3
c. 0.4
d. 0.6
e. 1.0

d



$3 \times \frac{1}{5} = 3/5$

4. Suppose that a random-number generator follows the triangular distribution shown in the graph.



1 − Area of shaded triangle
$= 1 - \frac{1}{2}\left(\frac{1}{2}\right)(1) = 1 - \frac{1}{4}$
$= 3/4$

What is the chance that a randomly selected number will be at most 0.5?

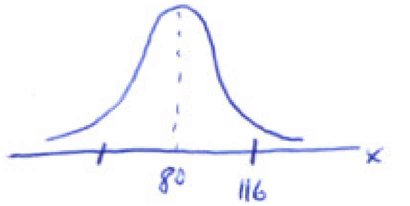
a. 0.0
b. 0.25
c. 0.5
d. 0.75
e. 1.0

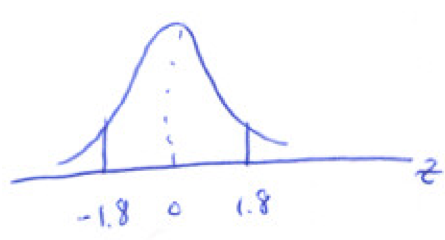
d

5. Suppose that X follows a N(80, 20) distribution and Z is a standard normal random variable. All but one of the probabilities below are equal. Which one is different?

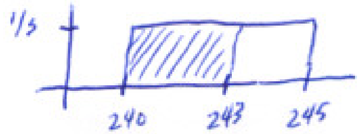
a. $P(Z \leq 1.8)$
b. $P(X \leq 121.8)$
c. $P(X \leq 116)$ $\frac{116-80}{20} = 1.8$
d. $P(X \geq 44)$ $\frac{44-80}{20} = -1.8$
e. $P(Z \geq -1.8)$

b





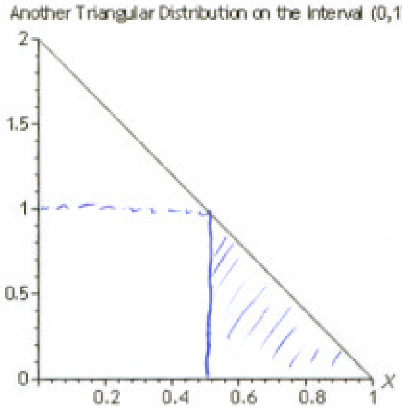
3. Application temperatures for a manufacturing process that seals bags vary uniformly from 240 to 245°° F. What is the probability that the application temperature for a randomly selected bag will be at most 243°° F?



$3 \times \frac{1}{5} = 3/5$

a. 0.2
b. 0.3
c. 0.4
d d. 0.6
e. 1.0

4. Suppose that a random-number generator follows the triangular distribution shown in the graph.



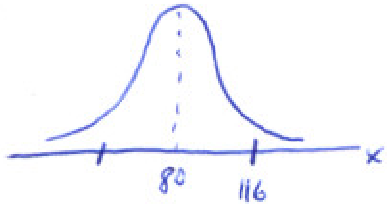
$1 -$ Area of shaded triangle
$= 1 - \frac{1}{2}\left(\frac{1}{2}\right)(1) = 1 - \frac{1}{4}$
$= 3/4$

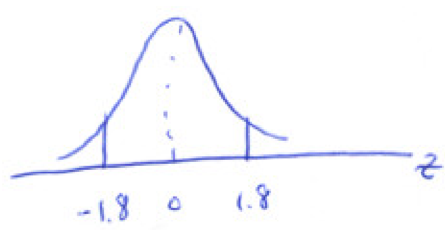
What is the chance that a randomly selected number will be at most 0.5?

a. 0.0
b. 0.25
c. 0.5
d d. 0.75
e. 1.0

5. Suppose that X follows a N(80, 20) distribution and Z is a standard normal random variable. All but one of the probabilities below are equal. Which one is different?

a. $P(Z \leq 1.8)$
b b. $P(X \leq 121.8)$
c. $P(X \leq 116)$ $\frac{116-80}{20} = 1.8$
d. $P(X \geq 44)$ $\frac{44-80}{20} = -1.8$
e. $P(Z \geq -1.8)$



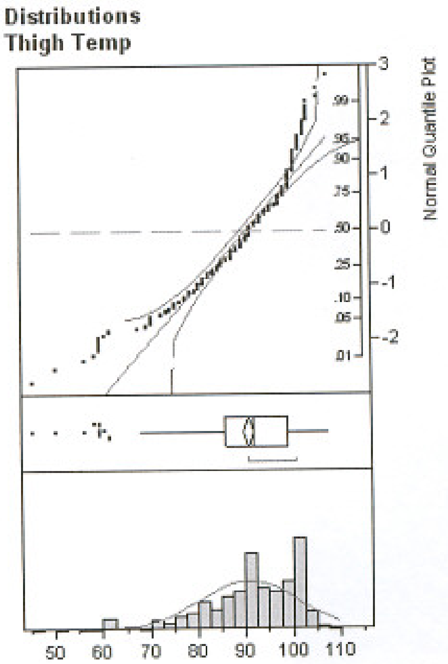


6. As part of an experiment to evaluate student achievement in math, a pretest and a posttest were given to each student. The investigators were testing the hypothesis of zero change in test scores. Which of the following must have an approximately normal distribution for the hypothesis test to be a valid procedure?

a. The pretest scores.
b. The posttest scores.
c. The change in scores, (posttest - pretest).
d. More than one of the above.

c

7. The following data displays are thigh temperatures for a sample of deer. Which statement below best describes your assessment of the plausibility of modeling these data with a normal curve?



a. The boxplot indicates some skew, but the points are basically linear in the normal probability plot.

b. The $\frac{IQR}{\sigma}$ ratio of 1.290 indicates significantly lighter-than-normal tails. This is clear evidence that the data are not normal.

c. All the graphical displays show that there is significant skew in these data, clearly indicating nonnormality.

c

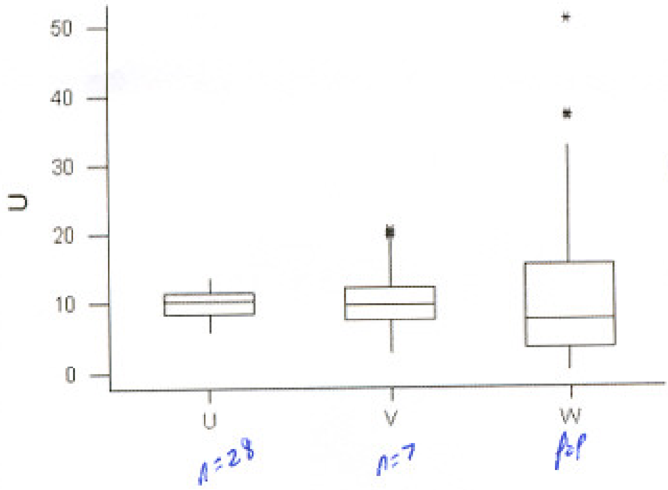
d. The boxplot does show some outliers at the low end, but with this number of observations, outliers are inevitable and do not suggest a lack of normality.

8. Which of the following confidence level/sample size combinations will result in the narrowest confidence interval for a population mean?

a. 90% confidence level, n=50
b. 90% confidence level, n=100
c. 95% confidence level, n=100
d. 99% confidence level, n=50
e. 99% confidence level, n=100

b

9. The boxplots below show the simulated sampling distributions for a population distribution with center at 10 and sample means of size 7 and 28 from this population. Which of the following statements lists the correct order for the densities for the population, the mean for sample size 7, and the mean for sample size 28?



e

a. U, V, W
b. U, W, V
c. V, W, U
d. W, U, V
e. W, V, U

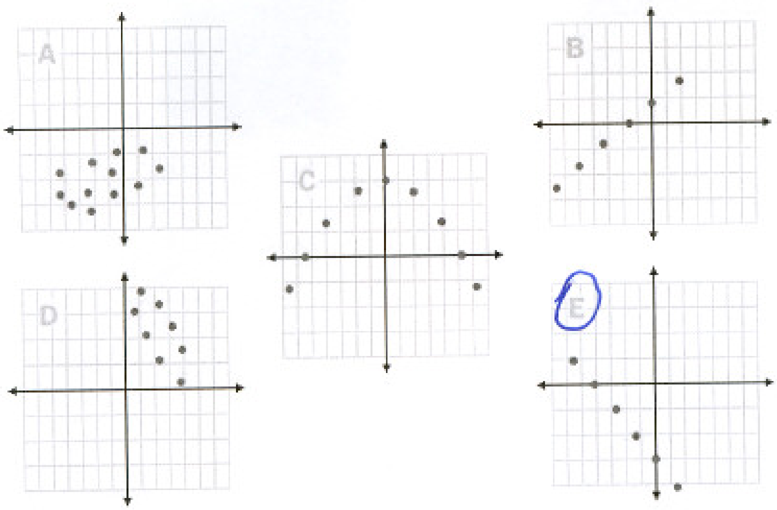
10. Which of the following is *not* a characteristic of a normal probability distribution?

c

a. The distribution is symmetrical.
b. The mean, median, and mode are all equal.
c. The standard deviation must be equal to 1.
d. The mean can be negative, zero, or positive.

11. Which of the scatter plots displays a correlation closest to -1?

e



http://agilemind.thinkfive.com/Kms/review/thinkfive

Part III. Problems. The point values for each part are in parentheses. To receive maximum credit, please show your work.

1. When modern PCs with statistical analysis software operate, data are frequently swapped back and forth between the computer and the hard drive. Because of this, the choice of disk drives is an important factor when large data sets are analyzed. The Acme Statistics Group installed 7 hard drives of each of two brands on 14 randomly selected computers. (Each hard drive was "rated" at the same data transfer speed.) They measured the times (in seconds) to perform a specified set of statistical calculations. The data for the two hard drives are given in the table below. Your task is to estimate the difference in means for the computation times for the two brands of hard drive. [33] AP #2.

Time to Complete Statistical Analysis

Brand A	Brand B
9.2	8.1
9.4	7.9
8.4	7.9
9.2	8
8.6	8.25
8.6	7.8
8.9	8.4

a. Based on the information above, which procedure would you tentatively decide to utilize: the two-sample independent t or the paired t? Justify your response with an appropriate statistical argument. (5)

-2 rationale
-5 paired t.

Since the computers were randomly selected, and thus the assignment to treatments was independent, the two-sample independent t should be used.

b. The procedures mentioned in the previous part assume sampling from normal populations. Assess the credibility of this assumption graphically. (5)

-3 graphs with no comments
-3. no linear assoc.
-2, residual plot

Although there appears to be a slight skew in the distribution for Brand B (refer to histograms, boxplots, or normal prob. plot) the data are consistent with the normal assumption. Normal prob. plots show roughly linear trend.

c. Construct a 95% confidence interval for the difference between the mean times to complete the statistical analysis. (5)

-1 incorrect C.V.
-1, incorrect SE.

Stat > Basic Stat > 2-sample t.

95% CI for $\mu_{Brand A} - \mu_{Brand B}$ is (.478, 1.22)

95% CI for $\mu_{Brand B} - \mu_{Brand A}$ is (-1.22, -.478)

-1, Pooled t (.492, 1.208)
-3, Paired t (.453, 1.247)

-5 incorrect C.I w/ no work.
-1, pooled t.

d. Based on your confidence interval for the difference between the mean times, is there sufficient evidence that the means are different for the two brands of hard drives? Provide appropriate statistical justification for your response. (10)

Correct conclusion, but
-3 do not check if 0 is in the C.I.
-5, No with reasonable rationale.
★ Many students did not use C.I.

Yes, there is sufficient evidence at the .05 level that the means differ for the two brands. 0 is not in the 95% C.I. ⇒ $H_0: \mu_A = \mu_B$ would be rejected in favor of $H_a: \mu_A \neq \mu_B$ at $\alpha = 0.05$.

2. One problem with owning a fishing-boat fleet is that barnacles attach themselves to the bottom of the boats. Two new barnacle-resistant paints have come on the market, each claiming to resist barnacle infestation. One fleet owner tested the two paints by painting both sides of his boats, each side with one brand, randomly assigned. (Yes, they were the same color!) At the end of the fishing season, the boats were raised and the barnacles scraped off. The amount of infestation, in pounds, is presented in the table below. You have been asked to perform the data analysis and report to the fleet owner. [34] AP #1

Boat Number	Paint 1	Paint 2	Boat Number	Paint 1	Paint 2
1	32.5	33.0	5	25.5	29.5
2	34.7	34.5	6	32.5	32.5
3	25.0	29.9	7	33.0	34.2
4	32.7	28.9	8	3.9	25.0

a. Assess the credibility of the assumption(s) necessary to test the hypothesis of equal barnacle resistance for these two paints. (5)

-2 do not mention boat 8
-3 do not check normality after mention paired data.
-1, fine until they assume indep of paint 1 & paint 2.

Paired Data => Form Difference (Paint 2 - Paint 1)
A normal probability plot for the paired differences clearly shows 1 very unusual difference (for boat 8).
I would be very cautious in using any procedures based on the normal assumption for these data.

-3 use 2-sample t

b. Is there sufficient evidence that the paints differ in their resistance to barnacle infestation? Clearly state your hypotheses, test statistic, p-value, and conclusion. (10)

Use sign or signed rank test.

$H_0: \eta_{diff} = 0$ vs $H_1: \eta_{diff} \neq 0$ 3

2 Below 0 1 equal 0 $\boxed{5 \text{ above } 0}$ 2

-1, just say don't reject Ho.

p-value = .4531 2

We do not have significant evidence 3 to reject the claims that the paints are the same.

Paired t

$H_0: \mu_{diff} = 0$ vs $H_1: \mu_{diff} \neq 0$

$t = 1.29$

p-value = .239

Same conclusion.

Signed Rank - W = 23, p-value = .151 => same conclusion

c. Write a short summary report for the fishing-fleet owner, explaining your analysis in "everyday" language. (He does not know any statistics.) (10)

-2, just say there's not stat evidence to conclude paints are diff.
-5, Because of the wacky result for boat #8 we can't say which paint is better.

We evaluated the two paints by painting one side of each boat with one of the paints, and the other with the other paint. When both sides were evaluated for infestation the ~~means and medians~~ amounts of barnacles were very similar (as measured by means & medians). Thus, it appears that the paints are about equal in their ability to inhibit infestation.

3. In a June 2000 CNN/Time Magazine poll, 30% of American adults answered "disturbing trend" to the question "Do you think that reality-based television shows are a disturbing trend for society or just harmless entertainment?" A student interested in statistics wants to know if high school seniors feel the same way at her school. She decides to take a random sample of 90 seniors and finds that 35 answer "disturbing trend." Perform a test of hypotheses to answer the student's question at the 5% significance level and clearly state the conclusion. (20) [31] AP #2.

3 $H_0: p = .3$ vs $H_a: p \neq .30$

3 $\hat{p} = \frac{35}{90} = .3889$

2 $np_0 = 90(.3) = 27$ and $n(1-p_0) = 90(.7) = 63$ are both greater than 10

Stat > Basic Stat > 1-Prop.

4 $z = \dfrac{.3889 - .3}{\sqrt{\frac{.3(.7)}{90}}} = 1.84$ ✓ Exact p-value is .084 (using MTB)

3 p-value $= 2 \times .0329 = .0658$

Since $.0658 > .05$ we cannot reject H_0.

5 There is not sufficient evidence to conclude that the high school seniors think reality TV programs are a disturbing trend for society at a different proportion than adults in America.

4. One of the perils of traveling by air is that people must share a common armrest. When one person uses the armrest, this can be bothersome to the other person. A recent study randomly sampled male/female pairs of individuals not traveling together but sitting next to each other and sharing an armrest. They were asked whether they were bothered when their seatmate used the common armrest. Here is a contingency table of their responses: [37] AP #1

	Bothered	Not Bothered	Total
Female	19 25.4	26 19.6	45
Male	38 31.6	18 24.4	56
Total	57	44	101

Expected cell counts.

a. Based on these data, what is the expected number of females who were not bothered by their seatmates' use of the armrest? (5)

-4 give $\frac{26}{45} = .578$ or 57.8%

$E[\text{females not bothered}] = \frac{45(44)}{101} = 19.604$

b. Is there any association between gender and propensity to be bothered by a seatmate's use of the common armrest? Be sure to show all your steps. (10)

-8, do not perform a significance test. -5, ~~an~~ compare appropriate cond. probabilities.

3 H_0: gender + bother-state are indep.
H_a: gender + bother-state are dep.

4 $\chi^2 = 6.669$
p-value $= P(\chi^2_1 > 6.669) = .0098$

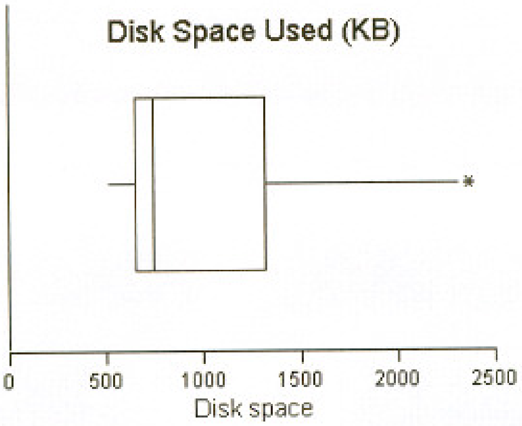
3 Since $.0098 < .01$, reject H_0 and conclude that gender and bother-state appear to be related.

c. In a few sentences, how would you interpret your results for someone not versed in statistics? (10)

-5, I cannot conclude there is a relationship.

By looking at the observed + expected counts, it appears that fewer females and more males are bothered than would be the case if gender and bother state were indep. [Sociological explanations are not appropriate.]

5. Management of shared disk space on a computer network is an important problem. Below are graphical and numerical summaries of the amount of disk space used for a simple random sample of 80 users of a large computer network.



$$n = 80; \bar{y} = 973.6\,kilobytes\,(KB); s = 483.1\,KB$$

-1, say skewed to the left

-1, we know sigma

a. Describe the shape of the boxplot. What conditions necessary for inference may be violated? Explain briefly. (5)

The boxplot has a long whisker on the right side, indicating that the distribution of disk space is skewed to the right. A condition necessary for inference is that the data come from a normal population. The boxplot clearly indicates that our sample data come from a population that is not normally distributed.

b. What does the central limit theorem say about the distribution of the sample mean? (10)

The central limit theorem says that when sampling at random from a nonnormally distributed population, as the sample size increases the sampling distribution of the sample mean becomes more and more symmetric & bell-shaped. (i.e., like a normal dist.)

-5, outliers impact the dist of $\bar{x}$.

-3, say dist of $\bar{x}$ gets normal if pop. is normal

-6, dist of $\bar{x}$ will remain the same

-2 don't mention $\bar{x}$ approx normal

c. Given the central limit theorem, test the hypothesis that the population mean amount of disk space used is 1000 KB against an alternative that less than 1000 KB is used, on average. (10)

3 $H_0: \mu = 1000$ KB vs $H_a: \mu < 1000$ KB

Test Stat: $t = \dfrac{973.6 - 1000}{483.1/\sqrt{80}} = -.49$

4

p-value $= P(t_{79} < -.49) = .3127$

-2 no conclusion in context.

3 Since the p-value > .05, we cannot reject H_0. The average amount of disk space used is not significantly smaller than 1000 KB.

6. 20-ounce bottles of Brown 'n Sweet soda are labeled with a sticker that says "1 in 6 wins a cool prize." Suzie Sucrose purchases a bottle of Brown 'n Sweet and twists off the cap hopefully. Her hopes are dashed when she sees the words "Tough luck, try again," on the underside of the cap. Suzie starts sniveling, then whines to her mother, "Pleeeeze can I buy five more bottles? I want to win a prize!" Her mother explains that buying five more bottles won't guarantee that she wins the cool prize. Suzie is bewildered. If her mother gives in to Suzie's whining and lets her buy five more Brown 'n Sweet sodas, what is the probability that Suzie wins

$X \sim$ Binomial $[n=5,\ p=1/6]$ Calc > Prob Dist > Binomial [19] STQ1

a. exactly one cool prize; (5) $P(X=1) = 5\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^4 = .4019$

-1, use n=6 to get .7368

b. at most one cool prize; (5) $P(X \leq 1) = .8038$

c. at least one cool prize; (5) $P(X \geq 1) = 1 - P(X<1) = 1 - P(X=0) = 1 - .4019 = .598122$

-3, 1-(b)

d. less than one cool prize; (5) $P(X<1) = P(X=0) = .4019$

e. more than one cool prize? (5) $P(X>1) = 1 - P(X \leq 1) = 1 - .8038 = .196245$

7. The Associated Press (Dec. 16, 1991) reported that in a random sample of 507 adult U.S. citizens, only 142 could correctly identify the Bill of Rights as the first ten amendments to the U.S. Constitution.

[27] AP #1

a. Identify the parameter of interest and provide the Wilson estimate and the traditional estimate for this parameter. (10)

-1 Number
-1, in the sample.
-3, provide C.I.'s.

p = proportion of adult U.S. citizens who can identify the Bill of Rights

traditional estimate 3 $\hat{p} = \frac{X}{n} = \frac{142}{507} = .280$

Wilson estimate 3 $\tilde{p} = \frac{X+2}{n+4} = \frac{144}{511} = .2818$

b. Calculate and interpret a 90% confidence interval for the proportion of adult U.S. citizens who could correctly identify the Bill of Rights. (15)

Use C.I based on Wilson estimate: 90% CI for p is

$$\tilde{p} \pm z^* \sqrt{\frac{\tilde{p}(1-\tilde{p})}{n+4}} \qquad .2818 \pm 1.6449\sqrt{\frac{.2818(1-.2818)}{511}}$$

$.2818 \pm .0327$

$(.2419, .3145)$

12

-2, incorrect C.V.

-2, do not use $\tilde{p}$.

Tradition CI. $(.24099, .319165)$

-2, use sample mean in interpretation

3 We can be 90% confident that the proportion of U.S. adults that can correctly identify the Bill of Rights is between .2419 and .3145.

8. The airline industry is interested in estimating the average weight of luggage (both checked and carry-on) brought by an individual passenger for a cross-country flight. Experience has shown that the length of time a passenger will be away from home affects how much luggage she will carry on a flight. A random sample of 500 passengers who were traveling from their home airports on a cross-country, round-trip ticket was selected. Each passenger's checked and carry-on luggage was weighed. Here are some summary statistics for this sample:

	Duration of Trip Predictor (x)	Weight of Luggage Response (Y)
Mean	5 days	65.2 pounds
Standard Deviation	3 days	14.3 pounds

Major Trouble. [6] AP #2

-3, say yes but then mention trip ≠ 0 days. Revisit

a. Could the distribution of trip durations be approximately normally distributed? Justify your answer. (5) It is very unlikely that trip durations are approximately normal. The shortest possible trip is still greater than 0 days. It seems more likely that the distribution is skewed to the right. Also, if trip durations are measured only in whole numbers of days, the distribution is discrete rather than continuous.

b. A scatter plot of the data reveals a correlation of 0.63 between trip duration and weight of luggage. Suppose that each passenger is given a two-pound deli snack to carry onto the plane. Now each passenger's total luggage weight has increased by two pounds. What effect will this have on the correlation coefficient? Explain. (5)
Adding two pounds to the luggage weight simply shifts all the points in the scatterplot by two units in the same direction. This will have no effect on the correlation coefficient. It will still be .63.

-3, simulated data set up correctly but use incorrect
-5, interchange x & y
-8, MTB > Reg > Reg

c. Find the equation of the least squares regression line for predicting luggage weight from trip duration. Show your method. (10)

$b = r\frac{s_y}{s_x} = \frac{.63(14.3)}{3} = 3.003$

$a = \bar{y} - b\bar{x} = 65.2 - 3.003(5) = 50.185$

Thus, the least squares regression line for predicting luggage weight is

$\hat{y}_{wt} = 50.185 + 3.003\,x$ ← duration of trip.

d. Interpret the slope and y-intercept of your least squares regression line from part c. in the context of this problem. (5)

The slope of 3.003 means that our model predicts an increase of 3 pounds in luggage weight for each additional day of trip duration.

The y-intercept means that our model predicts a luggage weight of 50.185 pounds for a trip that lasts 0 days. It will take some clever argument to give meaning to this. (Intercept has no practical interpretation in this setting.)

9. Suppose the average number of people in attendance at a showing of the new release of a movie in a local theater is 120 with a standard of 16. There are five showings of the movie each day. [24] AP #2.

a. Approximate the probability that at least 4250 people will see the new movie in this theater in the next week (7 days) by using the sampling distribution of the sample mean. What assumptions are you making in your calculations? (10)

-7 conduct a hypothesis test.

Let T represent the total number of people who see the new movie in the next 7 days. Properties of means and variances are used to show $E[T] = 7 \times 5 \times 120 = 4200$ and $Var[T] = 7 \times 5 \times 16^2 = 8960$. $SD[T] = 94.6573$ The normal approximation can be used to find $P(T > 4250) \cong 1 - .7013 = .2987$. Thus, there is about a 30% chance that at least 4250 people will see the movie in the next week.

-2, no comment on assumptions.

Major Assumptions - the # of people attending the 35 different showings can be represented by independent and identically distributed R.V.'s. with $\mu = 120$ $\sigma = 16$ AND ∴ the normal approx. can be used. Reasonable arguments supporting or questioning the indep assumption should be accepted.

b. Past records indicate that the average attendance increases by 25 at each showing on Saturday and Sunday if the weather is bad and decreases by 25 if the weather is good. How would your approximate probability change if weather forecasters indicate that there will be a stormy weekend? Explain. (10)

-5, correct E[T] but don't say anything about prob.

For a stormy weekend the expected total will increase by 250.

$E[T] = 5 \times 5 \times 120 + 2 \times 5 \times 145 = 4450$.

Thus, the prob of getting at least 4250 people to see the new movie will increase substantially to approximately 98%.

[NOTE: we must assume the variability stays the same on weekends, even though this may seem unlikely, since no additional information is given.

Note Also applies to part ⓒ

c. Would your approximate probability in part (a) change if the forecast calls for Saturday to be a sunny day and rain on Sunday? Explain. (5)

Since the expected weekly total will not change, the approximate probability will not change if the forecast calls for one sunny day and one rainy day.

$E[T] = 5 \times 5 \times 120 + 5 \times 145 + 5 \times 95 = 4200$.