AP Statistics Name: Key

Chapter 13 Review Hour: _____

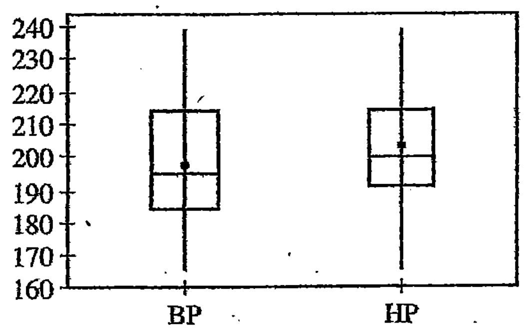
1. A 2001 Harris Poll of 1011 American adults found that 88% said they would trust teachers to tell the truth. This represented an increase of 2% from a similar poll taken in 1998, when 86% said they would trust teachers to tell the truth. Using a 95% confidence interval for the difference between the 2001 and 1998 results, decide whether there is convincing evidence of an increase in the proportion of adults who trust teachers to tell the truth. Assume that the sample sizes were the same in both years and that the samples were random and independent.

(−.0085, .05008)

(A) Yes, because the 95% confidence interval for the difference contains 2%.
(B) Yes, because the 95% confidence interval for the difference contains 0%.
(C) No, because the 95% confidence interval for the difference contains 2%.
(D) No, because the 95% confidence interval for the difference contains 0%.
(E) No, because the 95% confidence interval for the difference does not contain 2%.

37. In a study of the weights of college athletes, player weights for a random sample of baseball players (BP) and for an independent random sample of hockey players (HP) were compared. The computer output shown below gives the results of a test of $H_0: \mu_{BP} = \mu_{HP}$ *versus* $H_a: \mu_{BP} < \mu_{HP}$.

BOXPLOTS OF BP AND HP
(means are indicated by solid circles)



Two Sample T-Test and Confidence Interval

Two sample T for BP vs HP

	N	Mean	StDev	SE Mean
BP	37	197.8	19.4	3.2
HP	34	204.2	17.7	3.0

95% CI for μ BP − μ HP: (−15.1, 2.4)
T-Test μ BP = μ HP (vs <): T = −1.45 $\boxed{P = 0.076}$ DF = 69

Which of the following is the best conclusion that can be drawn from the analysis?

(A) The mean weight of baseball players is not significantly lower than the mean weight of hockey players at the 0.05 level.
(B) The mean weight of baseball players is not significantly lower than the mean weight of hockey players at the 0.10 level.
(C) The mean weight of baseball players is significantly higher than the mean weight of hockey players at the 0.10 level.
(D) The mean weight of baseball players is significantly lower than the mean weight of hockey players at the 0.05 level.
(E) The mean weight of baseball players is significantly different from the mean weight of hockey players at the 0.05 level.

.076 > .05 ; Fail to reject H_0 .076 < .1 ; reject H_0

3. A randomized experiment was performed to determine whether two fertilizers, A and B, give different yields of tomatoes. A total of 33 tomato plants were grown; 16 using fertilizer A, and 17 using fertilizer B. The distributions of the data did not show marked skewness and there were no outliers in either data set. The results of the experiment are shown below.

	Fertilizer A	Fertilizer B
Average number of tomatoes per plant	19.54	23.39
Standard deviation	3.68	4.93
Number of plants	16	17

Which of the following statements best describes the conclusion that can be drawn from this experiment?

(A) There is no statistical evidence of difference in the yields between fertilizer A and fertilizer B ($p > 0.15$).

(B) There is a borderline statistically significant difference in the yields between fertilizer A and fertilizer B ($0.10 < p < 0.15$).

(C) There is evidence of a statistically significant difference in the yields between fertilizer A and fertilizer B ($0.05 < p < 0.10$).

(D) There is evidence of a statistically significant difference in the yields between fertilizer A and fertilizer B ($0.01 < p < 0.05$).

(E) There is evidence of a statistically significant difference in the yields between fertilizer A and fertilizer B ($p < 0.01$).

$\mu_A - \mu_B \neq 0$

2 Samp T- test

p-value .016

✓ (2010 #5)

4. A large pet store buys the identical species of adult tropical fish from two different suppliers—Buy-Rite Pets and Fish Friends. Several of the managers at the pet store suspect that the lengths of the fish from Fish Friends are consistently greater than the lengths of the fish from Buy-Rite Pets. Random samples of 8 adult fish of the species from Buy-Rite Pets and 10 adult fish of the same species from Fish Friends were selected and the lengths of the fish, in inches, were recorded, as shown in the table below.

		Length of Fish	Mean	Standard Deviation
B	Buy-Rite Pets ($n_B = 8$)	3.4 2.7 3.3 4.1 3.5 3.4 3.0 3.8	3.40	0.434
F	Fish Friends ($n_F = 10$)	3.3 2.9 4.2 3.1 4.2 4.0 3.4 3.2 3.7 2.6	3.46	0.550

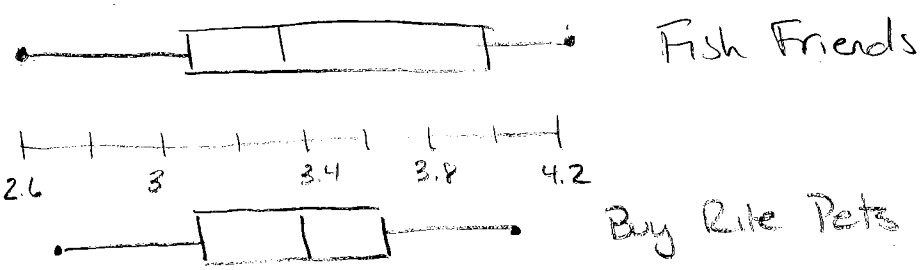
Do the data provide convincing evidence that the mean length of the adult fish of the species from Fish Friends (F) is greater than the mean length of the adult fish of the same species from Buy-Rite Pets (B)?

F > B

① $H_0: \mu_F - \mu_B = 0$ $\quad H_a: \mu_F - \mu_B > 0$

② Two sample t-test for $\mu_F - \mu_B$ where $\mu_F - \mu_B$ is the difference in the mean length of all adult fish of this species from Fish Friends (μ_F) and Buy-Rite Pets (μ_B)

③ It is stated that these are random samples of fish lengths from the two suppliers, so they are random and independent. Graphs of the sample data are both ~N so we can assume the sampling distributions are also ~N.



④ $$t = \frac{3.46 - 3.4}{\sqrt{\frac{.550^2}{10} + \frac{.434^2}{8}}} = .259 \qquad df = 16 \qquad \text{P-value} = P(t > .259) = .400$$

(2 Samp T-test) (tcdf)

⑤ Since the p-value $.400 > \alpha = .05$, we fail to reject H_0. There is not convincing evidence that the mean length of all adult fish of this species from Fish Friends is greater than the mean length of all adult fish of this species from Buy-Rite Pets.

(2006B #2)

5. A large company has two shifts – a day shift and a night shift. Parts produced by the two shifts must meet the same specifications. The manager of the company believes that there is a difference in the proportions of parts produced within specifications by the two shifts. To investigate his belief, random samples of parts that were produced on each of these shifts were selected. For the day shift, 188 of its 200 selected parts met specifications. For the night shift, 180 of its 200 selected parts met specifications.

(a) Use a 96 percent confidence interval to estimate the difference in the proportions of parts produced within specifications by the two shifts.

① Two sample z Confidence interval for $p_D - p_N$ where $p_D - p_N$ is the difference in the proportion of all parts produced during the day shift (p_D) and the night shift (p_N) that meet specifications.

② It is stated that these are independent, random samples from the two shifts. $200\left(\frac{188}{200}\right)=188$ $\quad 200\left(\frac{200-188}{200}\right)=12$ $\quad 200\left(\frac{180}{200}\right)=180$ $\quad 200\left(\frac{20}{200}\right)=20$
Since all three are greater than 10, the sampling distribution is ~N.

③ $\left(\frac{188}{200}-\frac{180}{200}\right) \pm 2.054\left(\frac{\frac{188}{200}\left(\frac{12}{200}\right)}{200}+\frac{\left(\frac{180}{200}\right)\left(\frac{20}{200}\right)}{200}\right) = (-.0156, .0952)$

④ Based on these samples, I am 96% confident that the difference in the proportion of all parts produced within specifications by the two shifts (day − night) is between −.0156 and .0952.

Calculator Check: 2 Prop z Int: (−.0156, .09557)

(b) Based only on this confidence interval, do you think that the difference in the proportions of parts produced within specifications by the two shifts is significantly different from 0? Justify your answer.

Since 0 is in the 96% Confidence interval, then it is likely that there is not a difference in the proportions of all parts produced within specifications by the two shifts.

(2012 #4)

6. A survey organization conducted telephone interviews in December 2008 in which 1,009 randomly selected adults in the United States responded to the following question.

At the present time, do you think television commercials are an effective way to promote a new product?

Of the 1,009 adults surveyed, 676 responded "yes." In December 2007, 622 of 1,020 randomly selected adults in the United States had responded "yes" to the same question Do the data provide convincing evidence that the proportion of adults in the United States who would respond "yes" to the question changed from December 2007 to December 2008?

① $H_0: p_8 - p_7 = 0$ $H_a: p_8 - p_7 \neq 0$

② Two sample z test for $p_8 - p_7$ where $p_8 - p_7$ is the difference in the proportion of all adults who would have responded "yes" to the question in 2008 (p_8) and Dec 2007 (p_7).

③ It is stated these are random, independent samples taken in 2007 and 2008. $1009\left(\frac{676}{1009}\right) = 676$ $1009\left(\frac{333}{1009}\right) = 333$ $1020\left(\frac{622}{1020}\right) = 622$ $1020\left(\frac{398}{1020}\right) = 398$. Since all are ≥ 10, the sampling distribution is $\sim N$.

④ $\hat{p}_c = \frac{676+622}{1009+1020} = \frac{1298}{2029}$

$$z = \frac{\frac{676}{1009} - \frac{622}{1020}}{\sqrt{\frac{\frac{1298}{2029}\left(1 - \frac{1298}{2029}\right)}{1009} + \frac{\frac{1298}{2029}\left(1 - \frac{1298}{2029}\right)}{1020}}} = 2.823$$

P-value $= 2P(z > 2.823) = .005$ $\left(2(.00236) = .00472\right)$

⑤ Since the p-value $= .005 < \alpha = .05$, we reject the H_0. There is convincing evidence that the proportion of all U.S. adults who would respond "yes" to the question has changed from Dec 2007 to Dec. 2008.

(2006 #4)

7. Patients with heart-attack symptoms arrive at an emergency room either by ambulance or by self-transportation provided by themselves, family, or friends. When a patient arrives at the emergency room, the time of arrival is recorded. The time when the patient's diagnostic treatment begins is also recorded.

An administrator of a large hospital wanted to determine whether the mean wait time (time between arrival and diagnostic treatment) for patients with heart-attack symptoms differs according to the mode of transportation. A random sample of 150 patients with heart-attack symptoms who had reported to the emergency room was selected. For each patient, the mode of transportation and wait time were recorded. Summary statistics for each mode of transportation are shown in the table below.

	Mode of Transportation	Sample Size	Mean Wait Time (in minutes)	Standard Deviation of Wait Times (in minutes)
A	Ambulance	77	6.04	4.30
S	Self	73	8.30	5.16

(a) Use a 99 percent confidence interval to estimate the difference between the mean wait times for ambulance-transported patients and self-transported patients at this emergency room.

① Two sample t confidence interval for $\mu_A - \mu_S$ where $\mu_A - \mu_S$ is the difference in the mean wait time for all ambulance-transported patients (μ_A) and all self-transported patients (μ_B) at this ER.

② The sample of patients (wait times) is stated as random and each patient may chose mode of transportation so they are also seen as independent. $n_A = 77 \geq 30$ $n_B = 73 \geq 30$ Since both are ≥ 30 the CLT states the sampling distribution will be $\sim N$.

③ $(6.04 - 8.30) \pm 2.611 \sqrt{\frac{4.3^2}{77} + \frac{5.16^2}{73}} = (-4.291, -.2291)$ df = 140.372
(2 Samp T-test)

④ Based on these samples, I am 99% confident that the difference in mean wait times for all ambulance-driven and self-driven patients at this ER is between −4.291 and −.2291 minutes.

(b) Based only on this confidence interval, do you think the difference in the mean wait times is statistically significant? Justify your answer.

Since 0 is not in the 99% confidence interval, the difference in mean wait times for the two types of transportation is statistically significant.