AP Stats Ch 24,25 - Required Practice Name: Solutions Per: ____

8. Friday the 13th, II: The researchers in Exercise 7 also examined the number of people admitted to emergency rooms for vehicular accidents on 12 Friday evenings (6 each on the 6th and 13th).

Year	Month	6th	13th	D
1989	October	9	13	4
1990	July	6	12	6
1991	September	11	14	3
1991	December	11	10	-1
1992	March	3	4	1
1992	November	5	12	7

Based on these data, is there evidence that more people are admitted on average on Friday the 13th? Here are two possible analyses of the data:

Paired t-Test of mu1 = mu2 vs. mu1 < mu2
Mean of Paired Differences = 3.333
t-Statistic = 2.7116 w/5 df
P = 0.0211

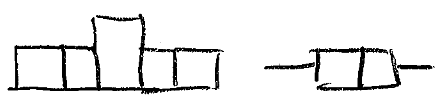
2-Sample t-Test of mu1 = mu2 vs. mu1 < mu2
Difference Between Means = 3.333
t-Statistic = 1.6644 w/9.940 df
P = 0.0636

a) Which of these tests is appropriate for these data? Explain.
b) Using the test you selected, state your conclusion.
c) Are the assumptions and conditions for inference met?

a) matched-pairs, matched by yr/month. Events within similar time period may be related.

b) With $\alpha = .05$, p-value $= .0211$ is low so reject H_0. We do have sufficient statistical evidence to conclude the mean number of ER admissions is higher on Friday 13th.

c) ✓ – SRS (assuming these dates are representative)
✓ – $6 < 10\%$ of all corresponding days
– Nearly Normal; Find differences:



yes, conditions for inference are met.

6. Rain. Simpson, Alsen, and Eden (*Technometrics* 1975) report the results of trials in which clouds were seeded and the amount of rainfall recorded. The authors report on 26 seeded and 26 unseeded clouds in order of the amount of rainfall, largest amount first. Here are two possible tests to study the question of whether cloud seeding works. Which test is appropriate for these data? Explain your choice. Using the test you select, state your conclusion.

Paired t-Test of $\mu(1 - 2)$
Mean of Paired Differences = −277.39615
t-Statistic = −3.641 w/25 df
p = 0.0012

2-Sample t-Test of $\mu1 - \mu2$
Difference Between Means = −277.4
t-Statistic = −1.998 w/33 df
p = 0.0538

a) Which of these tests is appropriate for these data? Explain.
b) Using the test you selected, state your conclusion.

independent (clouds are not matched)
the 2-Sample t-Test

With $\alpha = .05$, p-value $= .0538$ is high, so we fail to reject H_0. We do not have sufficient statistical evidence to conclude that cloud seeding works (the difference is mean rainfall is not statistically significant).

23. Braking. In a test of braking performance, a tire manufacturer measured the stopping distance for one of its tire models. On a test track, a car made repeated stops from 60 miles per hour. The test was run on both dry and wet pavement, with results as shown in the table. (Note that actual *braking distance*, which takes into account the driver's reaction time, is much longer, typically nearly 300 feet at 60 mph!)

a) Write a 95% confidence interval for the mean dry pavement stopping distance. Be sure to check the appropriate assumptions and conditions, and explain what your interval means.

b) Write a 95% confidence interval for the mean increase in stopping distance on wet pavement. Be sure to check the appropriate assumptions and conditions, and explain what your interval means.

Stopping Distance (ft)

Dry Pavement	Wet Pavement
145	211
152	191
141	220
143	207
131	198
148	208
126	206
140	177
135	183
133	223

Not matched

(a) this is a 1 Sample mean for dry only

conditions

✓ — SRS (assume these stops are representative)

✓ — 10 < 10% of all stops

✓ — Nearly normal:

NPP

Perform a T interval in TI-84 using dry data with C-level = .95 (133.61, 145.19)

We are 95% confident that the true average stopping distance on dry pavement is between 133.61 and 145.19 ft.

(b) this is 2 sample, independent test

conditions

✓ — SRSs (assume both are representative stops)

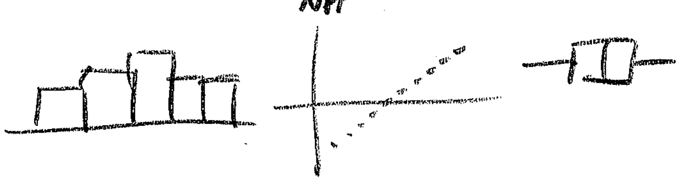
✓ — 10 < 10% of all stops

✓ — Nearly Normal:

✓ — groups independent

dry:

NPP



wet:

Perform a 2SampTInt in TI-84 dry data in L1, wet data in L2 w/ C-level = .95

(−74.62, −51.38) (dry − wet)

or

(51.38, 74.62) (wet − dry)

we are 95% confident that the mean increase in stopping distance (wet − dry) is between 51.38 and 74.62 ft.

25. Braking, test 2. For another test of the tires in Exercise 23, the company tried them on 10 different cars, recording the stopping distance for each car on both wet and dry pavement. Results are shown in the table.

a) Write a 95% confidence interval for the mean dry pavement stopping distance. Be sure to check the appropriate assumptions and conditions, and explain what your interval means.

b) Write a 95% confidence interval for the mean increase in stopping distance on wet pavement. Be sure to check the appropriate assumptions and conditions, and explain what your interval means.

Car #	Stopping Distance (ft) Dry Pavement	Wet Pavement
1	150	201
2	147	220
3	136	192
4	134	146
5	130	182
6	134	173
7	134	202
8	128	180
9	136	192
10	158	206

matched

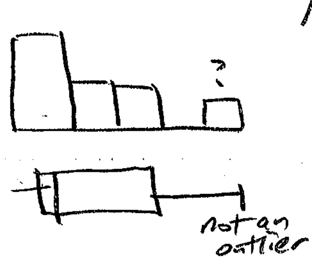
a) again, this just a 1 sample test on dry only

conditions

✓ – SRS (assume representative stops)

✓ – 10 < 10% of all stops

X – Nearly Normal:



this data look skewed right.

Perform a T-Interval in TI-84:
using dry data and C-level = .95

(131.79, 145.61)

we are 95% confident that the true average stopping distance on dry pavement is between 131.79 and 145.61 ft.
(But not sure if we can trust this result, data is not nearly normal).

b) this is a 2 sample, matched pair test. calculate differences in L3 = L2 − L1
wet − dry

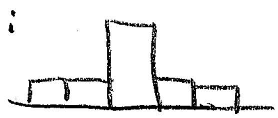
conditions

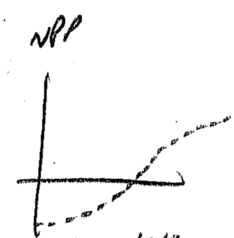
✓ – SRS (assume both samples are representative)

✓ – 10 < 10% of all stops

X – Nearly normal (differences):

✓ – groups matched by car





2 outliers

Perform a T-Interval in Ti-84
on the differences in L3
using C-level = .95

(38.782, 62.618)

we are 95% confident that the true mean increase in stopping distance (wet − dry) is between 38.782 and 62.618 ft.
(But we are not sure if we can trust this result, because the difference data contains outliers.)

19. Sex and violence. In June 2002, the *Journal of Applied Psychology* reported on a study that examined whether the content of TV shows influenced the ability of viewers to recall brand names of items featured in the commercials. The researchers randomly assigned volunteers to watch one of three programs, each containing the same nine commercials. One of the programs had violent content, another sexual content, and the third neutral content. After the shows ended, the subjects were asked to recall the brands of products that were advertised. Results are summarized below.

	Program Type		
	Violent	Sexual	Neutral
No. of subjects	108	108	108
Brands recalled			
Mean	2.08	1.71	3.17
SD	1.87	1.76	1.77

(a) Conduct a hypothesis test to determine whether there is evidence that viewer memory for ads may differ between programs with sexual content and programs with neutral content.

this is a 2sample (indep) test

Ho: $\mu_N = \mu_S$ mean number of brands recalled in ads is the same for sexual and neutral content.

HA: $\mu_N \neq \mu_S$ mean number of brands recalled in ads is different for sexual and neutral content.

conditions

✓ – SRSs (assume shows & ads are representative, and subjects)

✓ – 108 < 10% of all programs

✓ – groups are independent (randomly assigned)

✓ – Nearly Normal can assume both sexual and neutral are b/c $n = 108$.

Perform a 2SampTTest in Ti-84 using

$\bar{x}1 = 1.71$ $\bar{x}2 = 3.17$ $\mu_1 \neq \mu_2$

$Sx1 = 1.76$ $Sx2 = 1.77$ no-pooling

$n1 = 108$ $n2 = 108$

$t = -6.0785$

p-value $= 5.49 \times 10^{-9} \approx 0$

$df = 213.993$

with $\alpha = .05$, p-value ≈ 0 is low so we reject Ho.
We do have sufficient statistical evidence to conclude the mean number of brands recalled in ads is different for sexual and neutral program content.

(b) Find and interpret the confidence interval for the difference in brands recalled between sexual content programs and neutral content programs.

(conditions checked in a)

Perform a 2SampTInterval in Ti-84 using

$\bar{x}1 = 3.17$ $\bar{x}2 = 1.71$ C-level $= .95$

$Sx1 = 1.77$ $Sx2 = 1.76$ no pooling

$n1 = 108$ $n2 = 108$

(.98656, 1.9334) (neutral−sexual)

We are 95% confident that people will recall between 0.98656 and 1.9334 more brands in ads placed in neutral content compared to sexual content, on average.

19. Sleep. W. S. Gosset (Student) refers to data recording the number of hours of additional sleep gained by 10 patients from the use of *laevohyoscyamine hydrobromide*. We want to see if there is strong evidence that the herb can help people get more sleep.

a) State the null and alternative hypotheses clearly.

b) A t-test of the null hypothesis of no gain has a t-statistic of 3.680 with 9 degrees of freedom. Find the P-value.

c) Interpret this result by explaining the meaning of the P-value.

d) State your conclusion regarding the hypotheses.

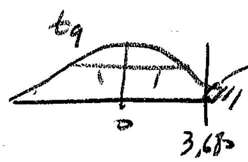
e) This conclusion, of course, may be incorrect. If so, which type of error was made?

gained implies before & after (matched pairs) and we are looking at the differences

a) Define D = difference in hours of sleep for each patient (w/herb − no herb)

Ho: $\mu_D = 0$ There is no difference in avg number of hours of sleep between herb and no herb.

HA: $\mu_D > 0$ Avg number of hours of sleep is higher for group taking the herb than no herb.

b) 

p-value = tcdf(3.680, 999, 9) = $\boxed{.0025}$

c) If there were actually no difference in sleep caused by the herb, we would see a difference in mean hours sleep as large as in this study (or larger) 0.25% of the time, just due to random chance.

d) With $\alpha = .05$, p-value = .0025 is low, so we reject Ho.
We do have sufficient statistical evidence to conclude that the average number of hours slept is higher for people taking the herb.

e)

	Ho T	Ho F
R	I	
NR		II

we rejected Ho, so if this was in error it would be a $\boxed{\text{Type I error}}$.

23. Lower scores? Newspaper headlines recently announced a decline in science scores among high school seniors. In 2000, 15,109 seniors tested by The National Assessment in Education Program (NAEP) scored a mean of 147 points. Four years earlier, 7537 seniors had averaged 150 points. The standard error of the difference in the mean scores for the two groups was 1.22.

a) Have the science scores declined significantly? Cite appropriate statistical evidence to support your conclusion.

b) The sample size in 2000 was almost double that in 1996. Does this make the results more convincing, or less? Explain.

a) different group of seniors (not matched) so this is a 2Sample (indp) test

H_0: $\mu_{2000} = \mu_{1996}$ Avg NAEP score has not changed.

H_A: $\mu_{2000} < \mu_{1996}$ Avg NAEP score has declined.

Conditions

✓ SRS (assume these are representative seniors)

✓ 15109 & 7537 are < 10% of all seniors in those years

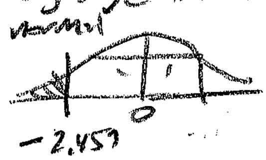
✓ groups are independent (different seniors)

✓ Nearly normal, n = 15109 & 7537 are ≥ 40 so yes

we can't do 2SampTTest because we don't have each year's standard error, but they gave us the standard error for the "difference in mean scores" so we can compute a test statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{SE_{\bar{x}_1 - \bar{x}_2}} = \frac{(147 - 150) - 0}{1.22} = -2.459$$

w/ n very large the t-distribution is approximately normal:



p-value = normalcdf(-999, -2.459, 0, 1)
= .00696

with α = .05, p-value = .00696 is low, so we reject H_0, we do have sufficient statistical evidence to conclude that average NAEP science scores have declined.

b) n = 15109 and 7537 are both very large and there is no reason n_1 must equal n_2 to conduct a 2 sample indp test, so this difference in sample size has no effect on the analysis or on how convincing it is.

Chapter 24 Practice Quiz
AP Statistics Quiz C – Chapter 24 Name________________

A total of 23 Gossett High School students were admitted to State University. Of those students, 7 were offered athletic scholarships. The school's guidance counselor looked at their composite ACT scores (shown in the table), wondering if State U. might admit people with lower scores if they also were athletes. Assuming that this group of students is representative of students throughout the state, what do you think?

Composite ACT Score		
Non-athletes		Athletes
25	21	22
22	27	21
19	29	24
25	26	27
24	30	19
25	27	23
24	26	17
23	23	

this is 2 sample (indep)

1. Test an appropriate hypothesis and state your conclusion.

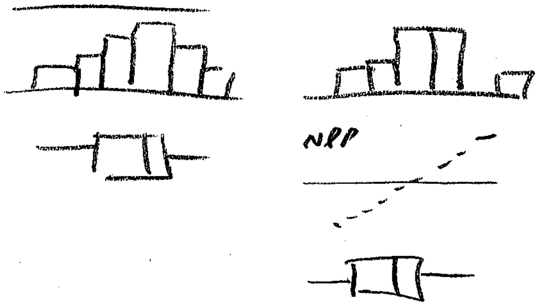
$H_0: \mu_{NA} = \mu_A$ mean ACT score for athletes and non-athletes are the same.

$H_A: \mu_{NA} > \mu_A$ mean ACT score for athletes is lower than non-athletes.

Conditions

✓ – SRS (says assume representative)
✓ – 23 < 10% of all students admitted
✓ – groups independent (likely)
✓ – Nearly normal:

non-athletes athletes



NPP

Perform a 2SampTTest in TI-84
using non-athlete data in L1, athlete in L2
$\mu_1 > \mu_2$, nonpooled
$t = 2.0212$
p-value = .0352
df = 10.1257

With $\alpha = .05$, p-value = .0352 is low so we reject H₀
we do have sufficient statistical evidence to conclude that the mean ACT score for athletes is lower than for non-athletes.

2. Create and interpret a 90% confidence interval.

(conditions checked in #1)
Perform a 2SampTInt in TI-84
using nonathlete data in L1, athlete in L2
C-level = .90, nonpooled

(.30209, 5.4836)

We are 90% confident that the athletes' mean ACT score is between 0.30209 and 5.4836 points lower than non-athletes' mean.

—or—

We are 90% confident that the difference in mean ACT score (non-athlete – athlete) is between 0.30209 and 5.4836 points.

Chapter 25 Practice Quiz

AP Statistics Quiz B – Chapter 25 **Name**____________________________

Most people are definitely dominant on one side of their body – either right or left. For some sports being able to use both sides is an advantage, such as batting in baseball or softball. In order to determine if there is a difference in strength between the dominant and non-dominant sides, a few switch-hitting members of some school baseball and softball teams were asked to hit from both sides of the plate during batting practice. The longest hit (in feet) from each side was recorded for each player. The data are shown in the table at the right. Does this sample indicate that there is a difference in the distance a ball is hit by batters who are switch-hitters?

Matched pairs (matched by batter)

1. Test an appropriate hypothesis and state your conclusion.

D = difference in hit lengths (dominant − non-dominant side)

$H_0: \mu_D = 0$ no difference in hit length dominant vs. nondominant side (on average)

$H_A: \mu_D \neq 0$ there is a difference in hit length for dominant and non-dominant side (on average)

c1 Dominant Side	c2 Non-dominant Side	c3 = c1−c2
142	119	23
144	118	26
153	126	27
148	119	
146	121	
149	125	
138	116	
145	120	
153	124	
160	138	
163	135	
170	144	
169	142	
151	128	
152	131	
167	141	
164	140	
165	140	
163	138	

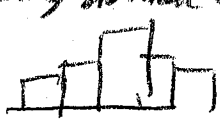
Conditions

✓ SRS – assume batters & hits are representative

✓ 19 < 10% of all batters

✓ matched pairs (by batter)

✓ Nearly normal differences



Perform a TTest on difference in L3 = (dom − non) using $\mu_0 = 0$, $\mu \neq \mu_0$

$t = 47.4$

p-value $= 2.3 \times 10^{-20} \approx 0$

(df = 18)

With $\alpha = .05$, p-value ≈ 0 is low, so we reject H_0. We do have evidence that there is a difference in average distance hit between dominant and non-dominant side.

2. Create and interpret a 95% confidence interval.

(Conditions checked above)

Perform a TInterval on differences using C-level = .95

(23.993, 26.217)

We are 95% confident that a batter's hit length is between 23.993 and 26.217 ft longer when batting on dominant side compared to non-dominant side, on average.