AP Statistics – Test #13 review   Name ____________________________

1. Suppose you rolled a die 60 times and observed 14 ones, 8 twos, 7 threes, 16 fours, 7 fives, and 8 sixes. Is this a fair die, or do you think it's loaded?

2. Five different vehicle models sales rate for Ivegotwheels Company are being compared to determine if the public prefers each vehicle type equally or not. The calculated chi-square statistics is 10.12. Is there sufficient evidence to say that the proportion of preferences is different or not?

3. A study of accident records at a large engineering company in England (***The Lancet***, October 22, 1994, page 1137) reported the following number in injuries on each shift for 1 year:

Shift:	Morning	Afternoon	Night
# of Injuries	1372	1578	1686

   Is there sufficient evidence to say that the number of accidents on the three shifts is not the same? Test at the .05, .01, and .001 levels.

4. A highway superintendent states that five bridges into a city are used in the ratio 2:3:3:4:6 during the morning rush hour. A highway study of a simple random sample of 6000 cars indicates that 720, 970, 1013, 1380, and 1917 cars use the five bridges, respectively. Can the superintendent's claim be rejected at the 2.5% or 5% level of significance?

5. Which of the following is NOT true of the Chi-Square probability density function?
   a) For small degrees of freedom, the curve displays right-skewness.
   b) As the degrees of freedom increase, the curve slowly approaches a normal curve.
   c) Chi-square is defined only for positive values of the variable.
   d) The area under a chi-square distribution is 1.
   e) All of these are true about the Chi-square probability density function.

6. An experiment was done by 15 students in a statistics class at the University of California at Davis to see if manual dexterity was better for the dominant hand compared to the nondominant hand (left or right). Each student measured the number of beans they could place in a cup in 15 seconds, once with the dominant hand and once with the nondominant hand. The order in which the two hands were measured was randomized for each student.

Student	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Dominant hand	22	19	18	17	15	16	16	20	17	15	17	17	14	20	26
Nondominant hand	18	15	13	16	17	16	14	16	20	15	17	17	16	18	25

a) Compute a 90% confidence interval for the mean difference in the number of beans that can be placed into a cup in 15 seconds by the dominant and nondominant hands.
b) Write a sentence or two using the interval to address the question of whether manual dexterity is better, on average, for the dominant hand.

7. In a study of memory recall, eight students from a large psychology class were selected at random and given 10 minutes to memorize a list of 20 nonsense words. Each was asked to list as many of the words as he or she could remember both 1 hour and 24 hours later. Is there evidence to suggest that the mean number of words recalled after 1 hour exceeds the mean recall after 24 hours **by more than 3 words**? Use .01 significance level.

Student	1	2	3	4	5	6	7	8
1 hr later	14	12	18	7	11	9	16	15
24 hours later	10	4	14	6	9	6	12	12

8. Over the ten year period from 1990 to 2000, the unemployment rates from Australia and the UK were reported as follows:

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Australia	7.3	9.6	10.7	10.9	9.7	8.5	9.6	8.6	8.0	7.2	6.3
UK	6.9	8.8	10.1	10.5	9.7	8.7	8.2	7.0	6.4	6.0	5.5

Find a 90% confidence interval for the mean difference in unemployment rates for Australia and the UK.

9. Tell what kind of test you would perform in the following situations:

   a. You want to see whether freshmen do better on their second calculus test of the semester than the first. You randomly select a group of freshmen, and record the scores of their first two tests.

   b. You want to see whether males or females in college drink more alcohol on the weekends. You randomly select 20 males and 20 females and gather data about the amount of alcohol they consume on 2 randomly selected weekends.

   c. You want to see whether more alcohol is consumed on homecoming weekend than the weekend before finals. You randomly select 20 males and 20 females and gather data about the amount of alcohol they consume on these two weekends.

AP Statistics – Test #13 review Answers!

1. If this is a fair die, then I would expect to get 10 of each type of roll. So that gives me:

Roll value:	1	2	3	4	5	6
Observed:	14	8	7	16	7	8
Expected:	10	10	10	10	10	10
$\frac{(\text{Observed - Expected})^2}{\text{Expected}}$	1.6	0.4	0.9	3.6	0.9	0.4

I'm going to use a χ^2 test here, with 5 degrees of freedom (# of categories minus 1).

I need to check my conditions for a χ^2 test:
1. I'm using counts.
2. Each roll is definitely random and independent of all others.
3. We expect to see at least 5 in every count.

The sum of the last row is 7.8, which is the χ^2 test statistic. We now calculate our *p*-value = $P(\chi_5^2 > 7.8) = 16.76\%$.

With a p-value of 16.76%, this means that there's a 16.76% chance of results this far off from the model happening just by chance, which is not that unusual. I would have a hard time saying that the die isn't fair.

2. H_0 in this case would be that all the model preferences are the same, and H_A would be that they are different. We've got 5 categories here, so we've got 4 degrees of freedom. $P(\chi_4^2 > 10.12) = 3.8\%$, which is a pretty low *p*-value. I would say that there is sufficient evidence to say that the proportion of preferences is different.

3. If the number of accidents on each shift were the same, then I would expect 1545.33 accidents per shift, and I would have the following table:

Shift:	Morning	Afternoon	Night
Observed:	1372	1578	1686
Expected:	1545.3333	1545.3333	1545.3333
$\frac{(\text{Observed - Expected})^2}{\text{Expected}}$	19.44	0.69	12.80

I'm using a χ^2 test here, with 2 degrees of freedom.

Conditions for a χ^2 test:
1. We have counts, rather than other types of measures.

2. We have to *assume* that the data was randomly collected.
3. We expect to see at least 5 in every count.

The sum of the last row is 32.93, which is the χ^2 test statistic.
$P(\chi_2^2 > 32.92) = 0.0000076\%$, which is REALLY unlikely.

With a *p*-value that low, there's an almost zero probability that results this far off the model would happen just by chance. We would reject the null hypothesis at just about any level, and state with confidence that there is strong evidence that the number of accidents on the three shifts is not the same. (Assuming that the data was randomly collected, of course)

4. Using the ratios and data given, we get the following table:

Observed:	720	970	1013	1380	1917
Expected:	666.67	1000	1000	1333.33	2000
$\frac{(\text{Observed - Expected})^2}{\text{Expected}}$:	4.27	0.9	0.169	1.63	3.44

I have a χ^2 test, with 4 degrees of freedom.

Conditions for the χ^2 test:
1. We have counts, rather than other types of measures.
2. The data was collected using an SRS.
3. We expect to see at least 5 in every count.

The sum of the last row is 10.4135, which is the χ^2 test statistic.
$P(\chi_4^2 > 10.4135) = 3.4\%$.

With a *p*-value of 3.4%, we can reject the superintendent's claim at the 5% level of significance, but not the 2.5% level of significance.

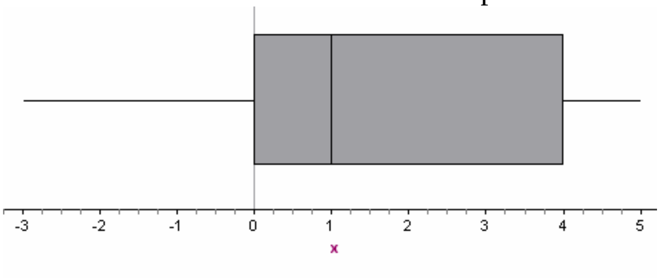
5. a) This is true for all χ^2 distributions, but especially true for those with small degrees of freedom.
b) True, although the distribution is always skewed somewhat to the right.
c) True.
d) True for *any* distribution, not just χ^2.
e) This is the correct answer.

6. We want to use a matched-pair *t*-interval here, since the two samples are not independent of one another, and each hand in the "dominant" sample can be matched to the same person's hand in the "nondominant" sample. So we have the following:

Student	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Dominant hand	22	19	18	17	15	16	16	20	17	15	17	17	14	20	26
Nondominant hand	18	15	13	16	17	16	14	16	20	15	17	17	16	18	25
Difference	4	4	5	1	-2	0	2	4	-3	0	0	0	-2	2	1

We need to check the conditions/assumptions:

1. Randomization. We don’t know how well the students involved represent the population at large, so we will assume that they are representative. We don’t have any reason to believe they aren’t. The order in which the two hands were used was randomized, in case people have a tendency to get either better or worse at the task the second time around.
2. Normal model. If we look at a boxplot of the differences, we see:



This isn’t real symmetric, but it’s not horribly asymmetric either, and there aren’t any outliers. It does look like it might be bimodal though, which could be a concern.
3. The sample size of 15 is far below 10% of the total population.

The 90% confidence interval is [-.0402, 2.1735], which is 1.067±1.107.

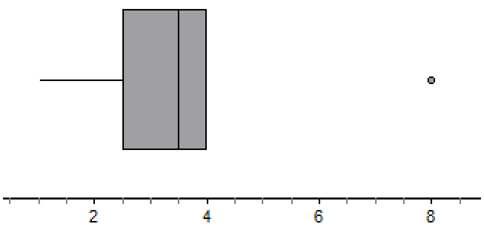
Because my confidence interval contains the value 0, I cannot say that dexterity is better, on average, for the dominant hand. My estimate is that the average difference in the number of beans that can be placed into the cup is about 1.067, but my margin of error is slightly larger than my estimate.

7. Using a matched pair *t*-test, we have:

Student	1	2	3	4	5	6	7	8
1 hr later	14	12	18	7	11	9	16	15
24 hours later	10	4	14	6	9	6	12	12
Difference	4	8	4	1	2	3	4	3

Conditions/assumptions:

1. Randomization. Again, we don’t know how well the students involved represent the population at large, so we will assume that they are representative. And again, we don’t have any reason to believe they aren’t. They were selected at random, but it's not necessarily true that psychology students are representative of the entire population.
2. Normal model. If we look at a boxplot of the differences, we see:



This is unfortunately not the boxplot we were looking for. We should probably do the test twice, once with the outlier and once without. Unfortunately, when the distribution has the outlier removed, it looks very skewed to the left. So we're not in great shape either way.
3. Independence: here we're in good shape, because there is no reason to think that any student would effect another student, and the number of students is far less than 10% of the population.

So we do our *t*-test, where
H_0: $\mu = 3$
H_A: $\mu > 3$

With the outlier, the results of our test are:

$t = .855$ (7 degrees of freedom)

$\bar{x} = 3.625$

$p\text{-value} = 0.21$

Without the outlier, the results of the test are:

$t = 0$ (6 degrees of freedom)

$\bar{x} = 3$

$p\text{-value} = 0.5$

Either way, the probability of our getting a sample mean this great or greater is way more than .01, assuming the null hypothesis is true. So I fail to reject the null hypothesis, and find that there is not sufficient evidence to say that the mean number of words recalled after one hour exceeds the mean recall of words after 24 hours by more than 3 words. I do, however, have to mention that the data appear to be non-normal, so these results may not be valid.

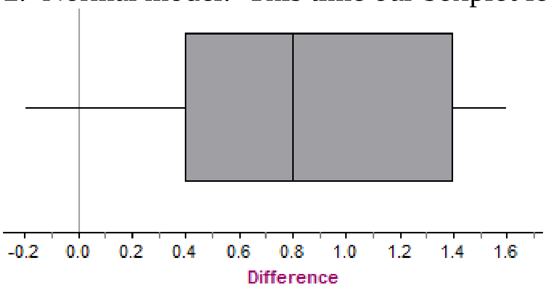
8. Again, use a matched pair *t*-interval:

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Australia	7.3	9.6	10.7	10.9	9.7	8.5	9.6	8.6	8.0	7.2	6.3
UK	6.9	8.8	10.1	10.5	9.7	8.7	8.2	7.0	6.4	6.0	5.5

Conditions/assumptions:

1. Randomization. The years were not selected randomly; they are 11 consecutive years from the recent past. The results of the confidence interval are dependent on these years being representative of all years.

2. Normal model: This time our boxplot looks good:



The data look reasonably symmetric, with no outliers.

3. Independence: Since these years are consecutive, one might also argue that they are not independent of one another, further throwing doubt on the conclusions of the study.

The 90% confidence interval for the difference between Australian and UK unemployment rates is [.4451, 1.1185], which is 0.7818 ± 0.3367. I am 90% confident that the difference in unemployment rates lies within that interval, assuming that the years in the study are representative of all years, and that all years are relatively independent.

9. a) Matched pair *t*-test, since each freshman is taking two tests.
b) We are comparing males and females here. The males and females are independent of one another, so there is no reason to match a particular male with a particular female. So we would use a 2-sample *t*-test.
c) This time we are comparing one weekend with another weekend, and each time we look at the same population. So we would do a matched pair *t*-test and measure the difference in alcohol consumed on each weekend for each specific person.