Practice Problems for the Midterm

1. Calculate the median for the following distribution:

```
. tab salarycat

    salarycat |      Freq.     Percent        Cum.
--------------+-----------------------------------
     0-15,000 |         32       44.44       44.44
15,001-30,000 |         36       50.00       94.44
30,001-45,000 |          4        5.56      100.00
--------------+-----------------------------------
        Total |         72      100.00
```

50% (i.e. the median) lies in the 15,001-30,000 interval, so…

$$L+\left[\frac{N(.5)-Cf}{f}\right]w = 15{,}001+\left[\frac{72(.5)-32}{36}\right]15{,}000$$ **= $16,667.67**

2. Calculate the mean for the following distribution of the maximum number of seats in a sample of sport utility vehicles:

6
6
9
9
6
5
5
5
7
5
5
5
8

#(X)	**Frequency(*f*)**	***f* * X**
5	**6**	**30**
6	**3**	**18**
7	**1**	**7**
8	**1**	**8**
9	**2**	**18**
å	**13**	**81**

$$\sum \frac{fx}{N} = \frac{81}{13}$$ **= 6.23**

3. Calculate the mean for the following distribution of the heights of females (in inches) of 16 female students in a physical education class.

Height (in inches)	Frequency		*f* * X	
61	3		183	
64	4		256	
67	7		469	
70	2		140	
	Σ = 16		Σ = 1048	

$$\sum \frac{fx}{N} = \frac{1048}{16}$$ **= 65.5″**

4. Calculate the Z-scores for the scores listing in the following distributions (assume normal distribution)

$Zscore = \frac{X - \overline{X}}{s}$ or $X = \overline{X} + Z(s)$

a) mean = 45 sd = 12

i. what percentage of scores are between 24 and 54 minutes?

Z = -1.75 and .75, so .4599 + .2734 = 73.33%

ii. what percentage of scores is greater than 39?

Z = -.5, area beyond = .3085, so 1-.3085 = 69.15%

b) mean = 505 sd = 111

i. what percentage of scores are less than 450?

Z = -.5, area beyond = .3085, so 30.85%

ii. what percentage of scores are between 200 and 450?

Z = -2.75 and -.5, so .4970 - .1915 = 30.55%

c) mean = 300 sd = 100

iii. what score must be achieved to be higher than 65.87% of the scores? **Find Z with an area beyond of 1 - .6587 = .3413, so Z = .41 which makes X = 341**

iv. what percentile corresponds to a score of 250? **Z = -.5, so the percentile is area beyond = 30.85%**

5. As a market researcher, you are using the number of sentences per advertisement as a measure of readability for magazine advertisements. You have a random sample of 54 advertisements with a mean of 12.4 sentences per advertisement. The standard deviation for your sample is 5. Construct a 95% confidence interval for the mean.

N = 54, $\overline{X}$ **= 12.4, Z = 1.96,** $S_{\overline{X}}$ **=** $\frac{5}{\sqrt{54}}$ **= .68, so**

C.I = 12.4 ± 1.96(.68) = 12.4 ± 1.33 = 11.07 to 13.73

6. In a survey of 1024 U.S. adults, 287 said that their favorite sport to watch is football. Find a 95% confidence interval for the population portion of U.S. who say their favorite sport to watch is football.

p= $\frac{287}{1024}$ **= .28,** S_P **=** $\sqrt{\frac{.28(.72)}{1024}}$ **= .014, so**

C.I. = .28 ± 1.96(.014) = .28 ± .027 = .253 to .307

7. A South Bend Mayor’s Office has issued a report of mortgage rates in the area stating that the current average mortgage rate is 6.7%. So, you randomly select 20 mortgage institutions and determine the current mortgage interest rate at each. Your sample mean rate is 6.93% with a standard deviation of 0.42%. Do you have reason to believe that the Mayor’s Office for the average mortgage rate for the area is significantly different than the one you found in your sample with a 99% confidence level?

N = 20, m **= 6.7%,** $\overline{X}$ **= 6.93%, s = .42,** $S_{\overline{X}}$ **=** $\frac{.42}{\sqrt{20}}$ **= .094, so**

C.I. = 6.93 ± 2.58(.094) = 6.93 ± .243 = 6.69% to 7.17%

No, the reported population mean fits just inside the 99% confidence interval, however, if we were to use a 95% level we would see that the sample

does significantly differ from the reported population because the population mean would no longer fall within the interval.

8. The average commute in a certain urban area is 30 minutes with a standard deviation 3.5 minutes. You gather information from with a random sample of 36 commute times and find a sample mean of 28.5 minutes. Does your sample underestimate the commute time for the area using a 99% confidence level?

N = 36, m **= 30,** s **= 3.5,** $\overline{X}$ **= 28.5,** $s_{\overline{X}} = \frac{3.5}{\sqrt{36}}$ **= .583, so**

C.I. = 30 ± 2.58(.583) = 30 ± 1.504 = 28.496 to 31.504

Given the above C.I., our sample mean of 28.5 fits just inside the confidence interval, so we'll assume it is not significantly different (i.e. not underestimating) from the true average…however, if we used a 95% level it would appear to underestimate the commute time as compared to the true (population) average.