Questions from: <http://www.math.wustl.edu/~jmding/math3200/chw/hw4.html>

1. *SAS code under “Problem 1” Heading. See Output pg. 1 for Partial SAS Output of the mean, standard deviation, z-interval, and t-interval for each sample of size 5; furthermore, the observation number denotes the number of a particular sample. See Output pg. 2 for a frequency output of the z-intervals and t-intervals which include the true mean, 7. (a) Z-Intervals including the true mean will have a value of XZ=1, and if it is not included the interval for that sample will have the value XZ=99. (b) Likewise, t-intervals containing the true mean will have a value of XT=1, and if it is not included the interval will have the value XT=99.*

Of the 1000 z confidence intervals, 88% (880 counts) contained the true mean, so the z-intervals had an estimated coverage probability of 88%. Of the 1000 t confidence intervals, 95.2% (952 counts) contained the true mean, so the t-intervals had an estimated coverage probability of 95.2%. I expected that the coverage probability for the t-interval would be higher because the bounds of the confidence intervals were calculated using sample statistics, and a true t-interval—based on the assumption that the true sigma is unknown—does use the sample standard deviation to approximate a data’s spread. A true z-interval should use the true population sigma to calculate the interval limits rather than the estimated (and often inaccurate) s for each sample. Thus, the z coverage probability is much lower than 95% while the t coverage probability is almost exactly 95%. Note: for a given confidence level, a t confidence interval will be wider than the z confidence interval to accommodate this sampling error. This is another reason that we expect the t intervals to have a higher coverage probability.

1. *SAS code under “Problem 2” heading. See Output pg. 3 for the lower 95% confidence limit on the mean octane rating. See output pg. 4 for the t-statistic and the bounds on its P-value.*

*(a)* The lower 95% confidence limit on the mean octane rating is 87.19489 where the sample mean is 87.395. Since the sample mean is above this confidence limit, we conclude with 95% confidence that the mean octane rating exceeds 87.

(b) H0: mu< 87. Ha: mu > 87. The t-statistic is 3.41 with 19 degrees of freedom. The corresponding p-value is 0.0029. Thus, the result is significant at the 0.005 (because p-value: 0.0029< 0.005, t-statistic: 3.41> 2.861) level but not at the 0.001 level (because p-value: 0.0029> 0.001, t-statistic: 3.41< 3.579).

1. *SAS code under “Problem 3” heading. See Output pg. 5 the normal probability plot. See Output pg. 6 for the upper 90% confidence interval for sigma.*

(a) The normal probability plot appears linear; thus, it is reasonable to assume that the data follow a normal distribution.

(b) H0: mu> 10. Ha: mu < 10

Thus, we reject the null hypothesis and conclude that since sigma < 10, the precision of the new device is better than the current monitor.

(c) The upper 90% confidence limit for sigma is 7.66607. Since this is lower than 10, we reject the null hypothesis and get the same results as in (b) with 90% confidence.

**PROBLEM 1:**

TITLE "Homework 4 Question 1";

**data** hw4q1;

do i=**1** to **1000**;

array z(**5**);

do j=**1** to **5**;

z[j]=rand('NORMAL',**7**,**5**);

zmean= mean(of z1-z5);

zstdev= std(of z1-z5);

zlower= zmean-**1.96**\*zstdev/sqrt(**5**);

zupper= zmean+**1.96**\*zstdev/sqrt(**5**);

\*Note: n=5 for each sample, so df=4 and alpha/2=0.05/2=0.025;

tlower= zmean-**2.776**\*zstdev/sqrt(**5**);

tupper= zmean+**2.776**\*zstdev/sqrt(**5**);

if zlower<**7** and zupper> **7** then XZ=**1**;

else XZ=**99**;

if tlower<**7** and tupper>**7** then XT=**1**;

else XT=**99**;

end;

output;

end;

drop i j;

**run**;

\*Only the first page of the SAS output has been included;

**proc** **print** data=hw4q1;

var zmean zstdev zlower zupper tlower tupper;

**proc** **freq** data=hw4q1;

tables XZ XT;

**run**;

**PROBLEM 2:**

TITLE "Homework 4 Question 2";

**data** hw4q2;

input rating @@;

datalines;

87.5 86.9 86.6 87.3 87.9 88 86.7 87.5 87.2 87

88.1 87.5 86.5 87.7 88 87.1 87 87.6 87.5 88.3

;

**run**;

**proc** **univariate** data=hw4q2 cibasic(type=lower alpha=**.05**);

var rating;

qqplot;

**run**;

TITLE2 "T-Test with H0: mu=87";

**proc** **ttest** data=hw4q2 h0=**87**;

var rating;

**run**;

**PROBLEM 3:**

TITLE "Homework 4 Question 3"

**data** hw4q3;

input results @@;

datalines;

125 123 117 123 115

112 128 118 124 111

116 109 125 120 113

123 112 118 121 118

122 115 105 118 131

;

**run**;

TITLE "HW 4, Q3: Testing Normality Assumption";

**proc** **CAPABILITY** data=hw4q3;

qqplot;

**run**;

TITLE "HW 4, Q3: Upper One-Sided 90% CI"

**proc** **univariate** data= hw4q3 alpha=**0.1** cibasic(type=upper alpha=**.1**);

**run**;