Practice Test Questions for Midterm: Review Ch 1 – 6

1. Download the HSB2 Dataset and import into a SAS session.

2. Test the claim that the mean writing score is equal to 50. Write all steps.

**Problem Statement**: Test the claim that the mean writing score is significantly different from 50.

**Assumptions**:

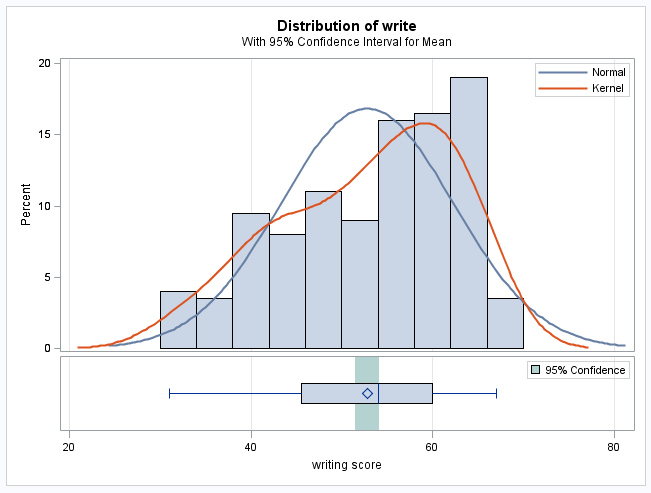
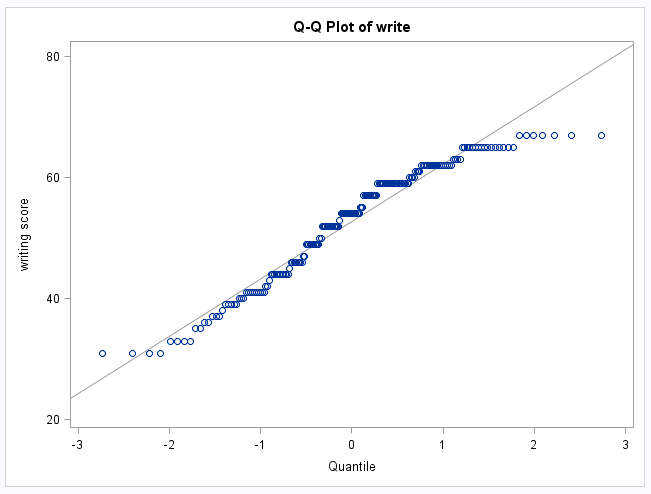
(Use proc ttest to get graphs to check assumptions).

\*To perform check assumptions and perform one-sample t-test;

proc ttest data = hsb H0 = 50;

var write;

run;



The histogram and q-q plot provide strong evidence of a left skew, although the sample size of 200 should ensure that the sample mean will be normally distributed (central limit theorem). We will assume the scores are independent of one another and proceed with a one-sample t-test.

\*A transformation is another option if it improves normality, but the inference will be on the median.

**Hypothesis test**:

1. Hypotheses:
2. Critical value:

\*To get critical value for one-sample t-test;

data quantile;

myquant = quantile('t', 0.975, 200-1);

run;

proc print data = quantile;

run;



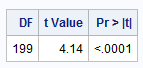
1. Test statistic: 4.14

\*To check assumptions and perform one-sample t-test;

proc ttest data = hsb H0 = 50;

var write;

run;





1. P-value: < 0.0001
2. Reject Ho at level alpha = 0.05
3. There is strong evidence to suggest at the alpha = .05 level of significance (p-value < .0001 from a one-sample t-test) that the mean writing score is different than 50 points. A 95% confidence interval for the true mean writing score is (51.5 points, 54.1 points). We can infer that the mean is not equal to 50 for the entire population of interest as the data was a random sample.

3. Test the claim that the mean Math score is different for males and females. Include all steps.

**Problem statement**: Test the claim that the mean writing score is different for males and females.

**Assumptions**:

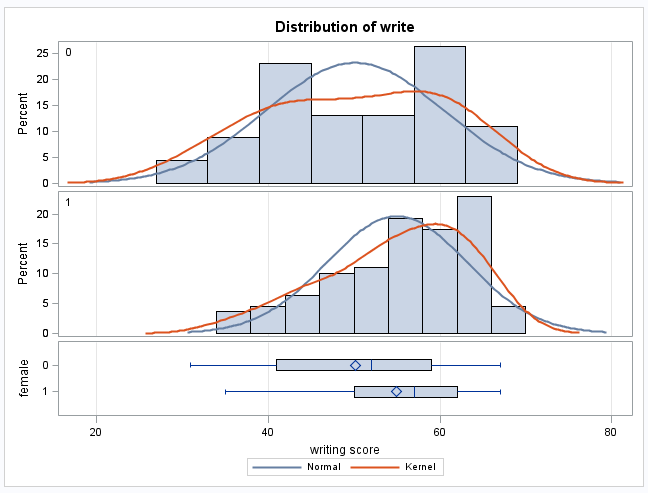
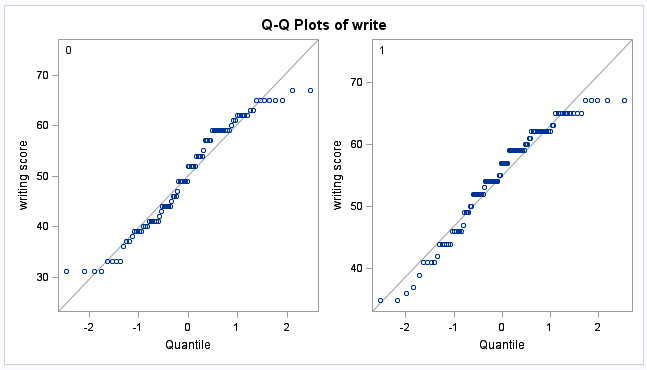
\*To check assumptions and then perform a two-sample t-test;

proc ttest data = hsb sides = 2;

class female;

var write;

run;



While there is not significant evidence that the male writing scores (female = 0) are not normally distributed, the histograms and q-q plots provide evidence of a slight left skewed distribution of writing scores for females (female = 1). However, since there are 91 in the male group and 109 in the female group, the Central Limit Theorem will ensure that sample means from these distributions are normally distributed, thus making the t-test robust to the normality assumption.

We will assume the scores are independent of one another both within and between groups.

The histograms are somewhat inconclusive on the question of equality of variances. For this reason, we seek secondary evidence in the form of a formal hypothesis test. Since there is some evidence that the writing scores are not normally distributed, the Brown-Forsythe test of equality of variance should be used instead of the F-Test.

\*To check equal variances assumption for the t-test;

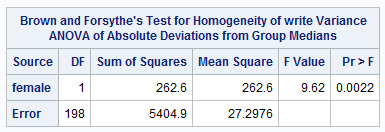
proc glm data = hsb;

class female;

model write = female;

means female/ hovtest = bf;

run;



There is some evidence that the standard deviations are different; therefore, since the Welch’s t-test is nearly as powerful as the Student t-test even when the standard deviations are the same, we will proceed with the Welch’s test of the difference of means.

\*A transformation is another option if it improves normality and makes variances more equal.

**Hypothesis test**:

1. Hypotheses:
2. Critical value: (run t-test to get Satterthwaite DF)

*±*

\*To get critical value for Welch's two-sample t-test;

data quantile;

thisquant = quantile('t', 0.975, 169.71);

run;

proc print data = quantile;

run;



1. Test statistic: -3.66

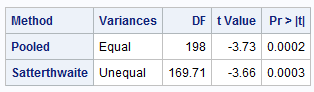
\*To check assumptions and then perform a two-sample t-test;

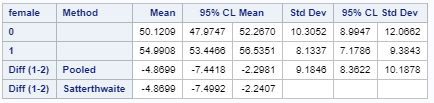
proc ttest data = hsb sides = 2;

class female;

var write;

run;





1. P-value: 0.0003
2. Reject Ho at level alpha = 0.05
3. There is strong evidence to suggest at the alpha = .05 level of significance (p-value = .0003) that the mean writing score of female high school students in the U.S. is different than the mean writing score of males. (Note the inference to the broader population.) This was an observational study; thus, no causal inference can be deduced. A 95% confidence interval for this difference is: (2.2 points, 7.5 points)-the positive difference in favor of females.

4. There are 4 unique race categories. Test to see if the 1st race has a different mean writing score than the 4th race.

**Problem Statement**: Test the claim that the writing scores of the first and fourth race significantly different. First, we will test the claim that any of the means/medians are different.

**Assumptions**:

Normality:

\*To check ANOVA assumptions for race groups;

proc glm data = hsb plots = all;

class race;

model write = race;

means race/ hovtest = bf;

run;

\*To address assumptions of the ANOVA for races with histograms and q-q plots;

proc univariate data = hsb;

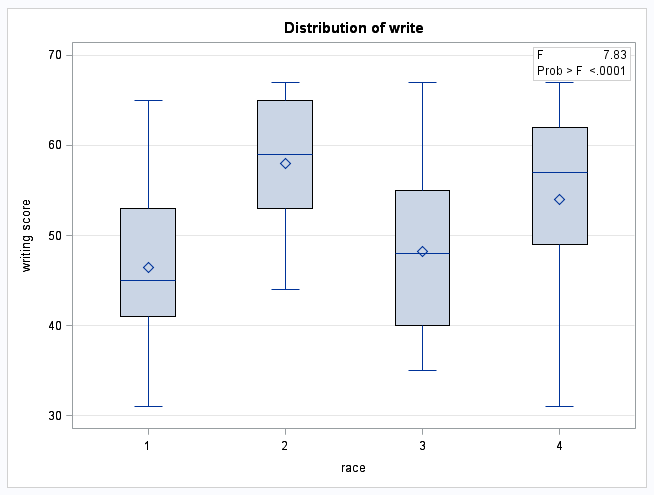
by race;

histogram write;

qqplot write;

run;

| Race = 1 | Race = 2 | Race = 3 | Race = 4 |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |



While the only histogram that provides strong evidence (n = 145 and a considerable left skew) against normality is race = 4, that is the group with the largest sample size. Remember that with small sample sizes, it is difficult to ascertain the true shape of the underlying distribution. There is no reason to believe the shape of the other races should be any different than the shape of the distribution of writing scores of race = 4. We will assume that the shapes (although maybe not the locations) of the distributions of the underlying populations of the first three groups are the same as that of the last (skewed) group. Some smaller sample sizes make it unclear that the CLT will kick in, calling the normality of some of the sample means into question.

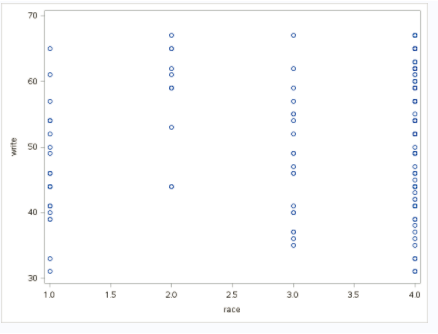
Equal standard deviations:

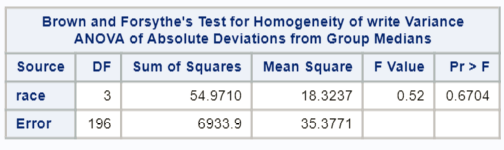
\*To address ANOVA assumptions for races with scatter plots;

proc sgplot data = hsb;

scatter x= race y = write;

run;





There is little evidence against equal standard deviations.

Independence: We will assume independence between and within groups.

Because normality is in question, we will conduct a Kruskal-Wallis Test and make our inference about the medians. (Note that we could have tried some transformations as well.)

**Hypothesis test**:

(1)

(2) Critical value: can skip

(3) Test statistic:

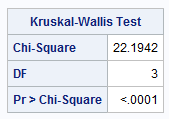
\*To perform nonparametric test for race groups;

proc npar1way data =hsb;

class race;

var write;

run;



(4) P-value: <0.0001

(5) Reject Ho

(6) There is sufficient evidence at the alpha = .05 level of significance (p-value < .0001) that at least one of the medians is different among the race groups. Next, we shall test to see if the 1st race has a different median than the 4th race. A multiple comparison adjustment is not needed here since our hypothesis was formulated before we looked at the data and we are only examining one comparison.

**Hypothesis test**:

(1)

(2) Critical value:

\*To get critical value for rank sum test;

data quantile;

aquant = quantile('NORMAL', 0.975);

run;

proc print data = quantile;

run;



(3) Test statistic: -3.72

\*To perform nonparametric comparison of two groups;

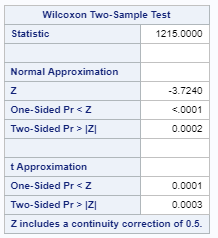
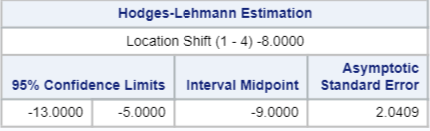
proc npar1way data = hsb Wilcoxon hl;

where race = 1 |race = 4;

class race;

var write;

run;

(4) P-value: <0.0001

(5) Reject Ho

(6) There is strong evidence to suggest at the alpha = .05 level of significance (p-value = .0002 from the Rank Sum Test) that the median writing score of U.S. high school students with race = 1 is different than the median writing score of those of race = 4. This was an observational study; thus, no causal inference can be deduced. The best estimate of the difference in locations is 8, in favor of race 4 scores, with a Hodges-Lehmann 95% confidence interval of the difference of (5 points, 13 points).

\*Note that the difference in medians is actually 12. That is because the HL confidence interval is not exactly for the medians when the data is not symmetric.

5. In the Midterm Review PowerPoint, we found that there was strong evidence that the 12 packs had a higher mean amount sold than the 18 or 30 packs. This may be due to a lot of factors. We would like to investigate if price is one of them. Maybe, on average, the 12 pack cans were priced lower (per can) than the 18 or 30 packs and this was the reason they sold more. Divide each price by the number of cans for each of the 52 weeks and create a column for “price per can.” Conduct a test to detect any difference in price per can for the three case amounts (12, 18, 30). The above analysis does not need to be long … in fact I would prefer it not be … but it should be thorough. Provide the plots and arguments necessary to address the assumptions and fully interpret each model and / or analysis. Be selective in which plots you choose and minimize them on your paper. Organize your presentation to make the most of your space. In short, think about what YOU would want to see if you asked for the information / analysis / report.

**Problem Statement**: Test the claim that the price per can of any type is significantly different. If an overall test shows significance, we will address which types are significantly different.

First, we will test the claim that any of the means/medians are different.

**Assumptions**:

Normality: There is strong evidence against normality, but there is likely sufficient data to invoke the Central Limit Theorem.

Equal Standard Deviations: There is strong evidence against equal standard deviations.

Independence: Independence between groups and within groups is likely violated, as prices are likely to be dependent on prior week pricing, and prices each week for each type are likely to be related. Although there are some methods to neutralize these effects, they will not be discussed here, and we will proceed with caution. A transformation could be considered, but we will use Welch’s ANOVA to see if any of the means are different.

\*To check equal variances assumption for ANOVA;

proc glm data = beer;

class type;

model pricepercan = type;

means type/ hovtest = bf;

run;

\*To address assumptions of the ANOVA with histograms and q-q plots;

proc univariate data = beer;

by type;

histogram pricepercan;

qqplot pricepercan;

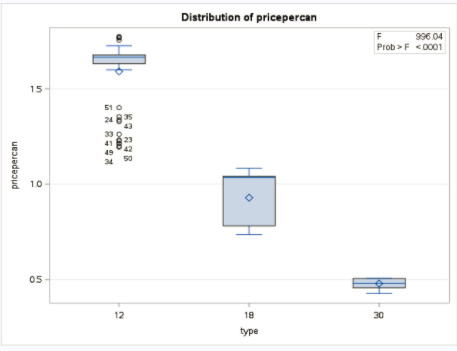
run;

\*To address ANOVA assumptions with scatter plots;

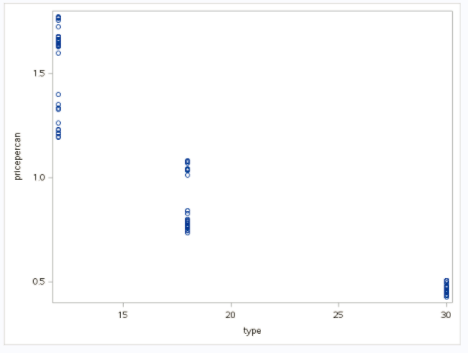
proc sgplot data = beer;

scatter x= type y = pricepercan;

run;



| Type = 12 | Type = 18 | Type = 30 |
| --- | --- | --- |
|  |  |  |
|  |  |  |



**Hypothesis test**:

(1)

(2) Critical value: can skip

(3) Test statistic: 27.62

\*To perform welch's ANOVA;

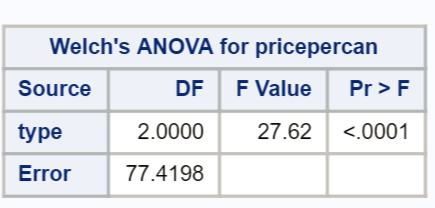
proc glm data = beer;

class type;

model pricepercan = type;

means type / welch;

run;



(4) P-value: <0.0001

(5) Reject Ho

(6) There is sufficient evidence at the alpha = .05 level of significance (p-value < .0001) that at least one of the mean prices per can is different among the three package groups. Next, we shall test to see which groups are different. A multiple comparison adjustment is needed here, since we are looking at all comparisons. Note that because standard deviations are not equal, we should NOT perform comparisons that involve pooled standard deviations.

However, we can perform three separate Welch’s t-tests using a Bonferroni adjustment.

If we set the family-wise alpha = 0.05, then the individual alpha for each test should be 0.05/3 = .016667.

**Hypothesis tests (three)**:

(1)Hypotheses:

| 12 vs. 18 | 12 vs. 30 | 18 vs. 30 |
| --- | --- | --- |
|  |  |  |
|  |  |  |

(2) Critical value:

\*To get critical value for all three tests (one value bc they have the same sample size);

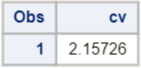
data critval;

cv = quantile('t', (1-(.05/3)), 52+52 -2);

run;

proc print data = critval;

run;



(3) Test statistic:

|  |  |  |
| --- | --- | --- |
| 12 vs. 18 | 12 vs. 30 | 18 vs. 30 |
| t-value: 21.72 | t-value: 45.52 | t-value: 23.74 |
| \*To perform 12 vs. 18 Welch's t-test with Bonferroni adjustment;  proc ttest data = beer alpha = 0.01666667;  where type = 12 |type = 18;  class type;  var pricepercan;  run; | \*To perform 12 vs. 30 Welch's t-test with Bonferroni adjustment;  proc ttest data = beer alpha = 0.01666667;  where type = 12 |type = 30;  class type;  var pricepercan;  run; | \*To perform 18 vs. 30 Welch's t-test with Bonferroni adjustment;  proc ttest data = beer alpha = 0.01666667;  where type = 18 |type = 30;  class type;  var pricepercan;  run; |
|  |  |  |

(4) P-value:

|  |  |  |
| --- | --- | --- |
| 12 vs. 18 | 12 vs. 30 | 18 vs. 30 |
| p-value: <0.0001 | p-value: <0.0001 | p-value: <0.0001 |

(5) Reject? (Compare p-values to 0.0166667.):

|  |  |  |
| --- | --- | --- |
| 12 vs. 18 | 12 vs. 30 | 18 vs. 30 |
| Reject Ho | Reject Ho | Reject Ho |

(6) There is sufficient evidence at the family-wise alpha = .05 level of significance (p-values < .0001) that all prices per can are significantly different. A 98.333% confidence interval for the difference in means of price per can of 12 and 18 packs is (58.73 cents, 73.57 cents), a 98.333% confidence interval for the difference in means of price per can of 12 and 30 packs is ($1.05, $1.17), and a 98.333% confidence interval for the difference in means of price per can of 18 and 30 packs is (40.31 cents, 49.66 cents) with the price per can decreasing as the cans per pack increases. We have very little information about the broader population from which this data came, so we can not generalize these results to a broader population. However, the assigning of package types and pricing is unclear, so we cannot draw strict causal inferences in a statistical sense. However, basic economic theory (economy of scale) tells us that companies often specifically price higher volumes at a cheaper unit price, and this case is no exception.