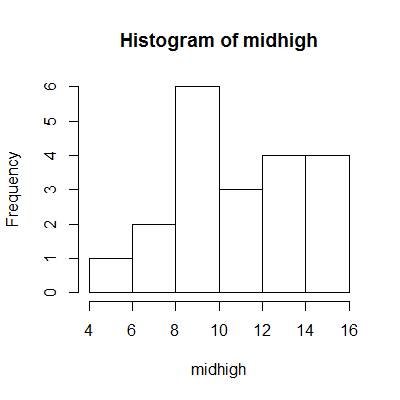
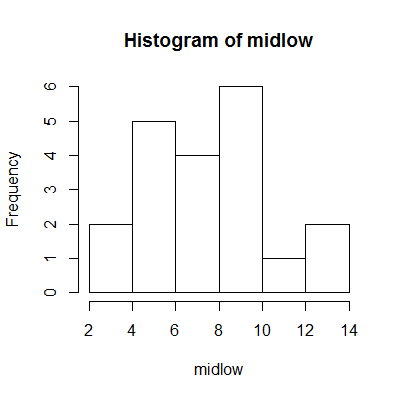
# MA4413 R Assignment

# *for*

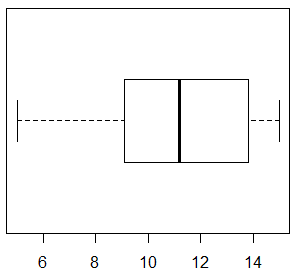
# Philip Waldron (Seed: 14173026)

**1) Analysis of Midterm Scores**

**ii) Graphical and Numerical Summaries**

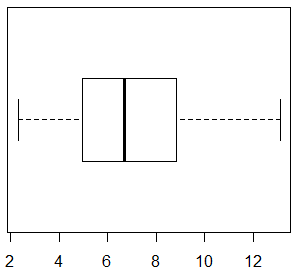
Scores

Scores



**Boxplot of midlow**

**Boxplot of midhigh**

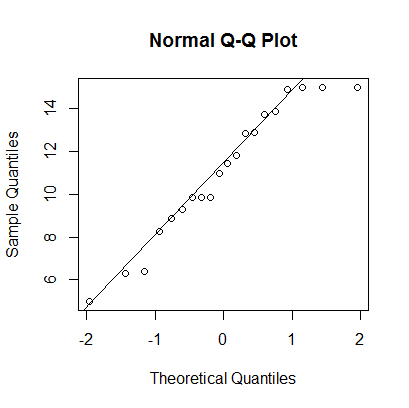
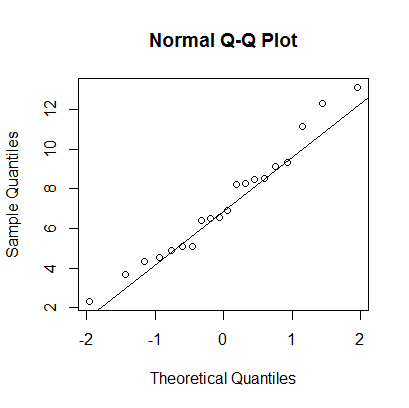


|  |  |  |
| --- | --- | --- |
|  | **midhigh** | **midlow** |
| **Mean** | 11.06 | 7.23 |
| **Standard Deviation** | 3.12 | 2.89 |
| **1st Quartile** | 9.20 | 5.02 |
| **3rd Quartile** | 13.77 | 8.68 |
| **Inter Quartile Range** | 4.57 | 3.66 |
| **Minimum** | 4.99 | 2.31 |
| **Maximum** | 15.00 | 13.12 |

**Shape:** midhigh has a negative skew while midlow is symmetric.

**Centre:** In this case is the mean, which is 11.06 for midhigh and 7.23 for midlow. On average midhigh has higher score than midlow.

**Spread:** In this case is the standarddeviation, with both midhigh and midlow having similar spreads.

**iii) Check for Normality of Data**

**midlow**

**midhigh**

From the above Q-Q plots we see that midlow is approximately normally distributed as it generates a straight line from corner to corner (as the histogram is symmetric).

The Q-Q plot for midhigh has a slight downward curve in points (as the histogram has a negative skew) producing a line that doesn’t go from corner to corner, hence midhigh is not normally distributed.

**iv) Confidence Intervals and Hypothesis Testing**

*>t.test(c(midhigh, midlow), mu=7.5)*

One Sample t-test

data: c(midhigh, midlow)

t = 2.9329, df = 39, p-value = 0.005596

alternative hypothesis: true mean is not equal to 7.5

95 percent confidence interval:

8.009979 10.276521

sample estimates:

mean of x

9.14325

Null hypothesis: **Ho: μ = 7.5**

Alternative hypothesis: **Ha: μ ≠ 7.5**

We reject the null hypothesis as the p-value is very small (0.005596 < 0.05) given the confidence interval.

*>var.test(midhigh, midlow)*

F test to compare two variances

data: midhigh and midlow

F = 1.1674, num df = 19, denom df = 19, p-value = 0.7393

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

0.4620629 2.9493261

sample estimates:

ratio of variances

1.167379

Null hypothesis: **Ho: V1:V2 = 1**

Alternative hypothesis: **Ha: V1:V2 ≠ 1**

We Accept the null hypothesis as the p-value is very large (0.7393 > 0.05) strongly supporting the chance of the ratio of variances being equal to 1, given the 95 percent confidence interval 0.4620629 – 2.9493261.

*> t.test(midhigh, midlow, var.equal=TRUE)*

Two Sample t-test

data: midhigh and midlow

t = 4.0244, df = 38, p-value = 0.0002626

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

1.90118 5.74982

sample estimates:

mean of x mean of y

11.0560 7.2305

Null hypothesis: **Ho: μ1: μ2 = 1**

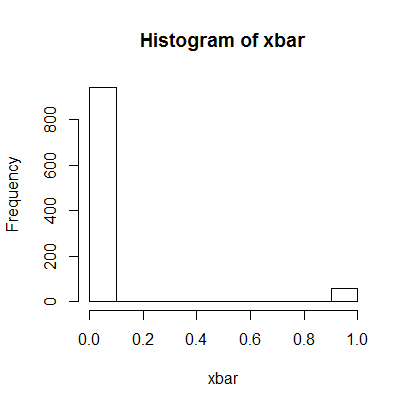
Alternative hypothesis: **Ha: μ1: μ2 ≠ 1**

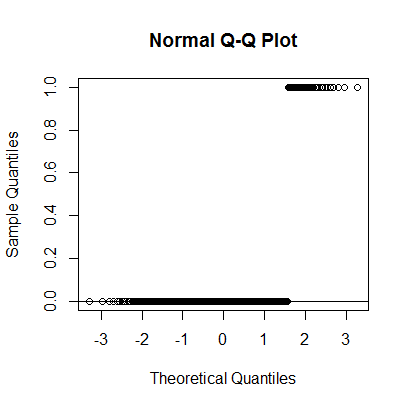
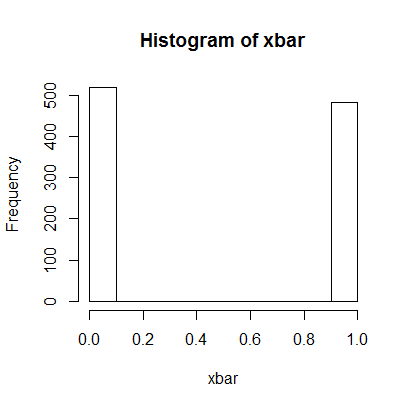
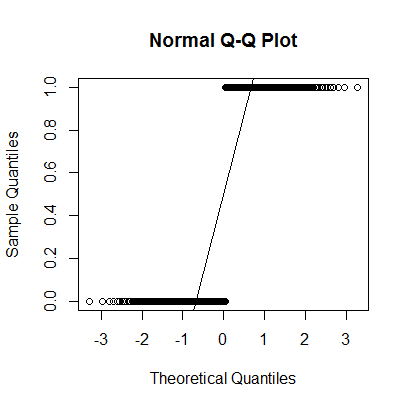
We reject the null hypothesis as the p-value is very small (0.0002626 < 0.05) given the 95 percent confidence interval 1.90118-5.74982.

**v) Brief Summary of Analysis**

The main results of my analysis show that the average between midterm results taking into account Sulis activity from samples of two groups, high and low, differ significantly, with higher Sulis activity showing higher scores on average. The low Sulis activity group shows a normal distribution while high Sulis activity is negatively skewed, favouring higher scores more often.

When both sample groups are taken into account, on average students get a grade higher than the halfway mark. The variance and standard deviation for both groups are also very similar, while the mean values are set apart.

****2**) Simulation Study (Central Limit Theorem)**

****

**3) Probability Questions**

i) Pr(X ≥ 6) where X ~ Binomial(n = 10, p = 0.65) = 0.7515

ii) Pr(X < 30) where X ~ Binomial(n = 100, p = 0.2) = 0.9888

iii) Pr(15 ≤ X ≤ 30) where X ~ Binomial(n = 50, p = 0.32) = 0.6698

iv) Pr(X = 8) where X ~ Poisson(λ = 6) = 0.1033

v) Pr(X > 35) where X ~ Poisson(λ = 41) = 0.8031

vi) Pr(2 ≤ X ≤ 5) where X ~ Poisson(λ = 1) = 0.2636

vii) Pr(X > 12) where X ~ N(μ = 7, σ = 2.5) = 0.02275013

viii) Pr(X > 9.8) where X ~ N(μ = 10, σ = 1) = 0.5792597

ix) Pr(X < 38) where X ~ N(μ = 50, σ = 5) = 0.009875511

x) Pr(4 < X < 8) where X ~ N(μ = 5, σ = 3.6) = 1.820151