# STAT 461: Homework 2

**Name:**Kyle Salitrik | ***ID#:*** *997543474* | ***PSU ID:*** *kps168*

September 5, 2018

## PROBLEM 1

The following is the assignments for the first experiment:

| | units | sample1 |
|---|---|---|
| 1 | 1 | high |
| 2 | 2 | high |
| 3 | 3 | none |
| 4 | 4 | low |
| 5 | 5 | medium |
| 6 | 6 | high |
| 7 | 7 | low |
| 8 | 8 | none |
| 9 | 9 | high |
| 10 | 10 | medium |
| 11 | 11 | none |
| 12 | 12 | none |
| 13 | 13 | low |
| 14 | 14 | none |
| 15 | 15 | low |
| 16 | 16 | medium |
| 17 | 17 | high |
| 18 | 18 | low |
| 19 | 19 | medium |
| 20 | 20 | medium |

## PROBLEM 2

The following is the assignments for the second experiment:

| | units | sample2 |
|---|---|---|
| 1 | 1 | none |
| 2 | 2 | high |
| 3 | 3 | medium |
| 4 | 4 | none |
| 5 | 5 | medium |
| 6 | 6 | low |
| 7 | 7 | high |
| 8 | 8 | low |
| 9 | 9 | low |
| 10 | 10 | none |
| 11 | 11 | medium |
| 12 | 12 | high |
| 13 | 13 | none |
| 14 | 14 | medium |
| 15 | 15 | medium |
| 16 | 16 | none |
| 17 | 17 | low |
| 18 | 18 | high |
| 19 | 19 | low |
| 20 | 20 | high |

## PROBLEM 3

The following is the assignments for the third experiment:

| | q3units | q3sample |
|---|---|---|
| 1 | 1 | r3 |
| 2 | 2 | r1 |
| 3 | 3 | r3 |
| 4 | 4 | r2 |
| 5 | 5 | r2 |
| 6 | 6 | r3 |

| | | | |
|---|---|---|---|
| 8 | 7 | 7 | r2 |
| 9 | 8 | 8 | r1 |
| 10 | 9 | 9 | r3 |
| 11 | 10 | 10 | r2 |
| 12 | 11 | 11 | r3 |
| 13 | 12 | 12 | r2 |
| 14 | 13 | 13 | r1 |

## PROBLEM 4

The first correlation was the lighthearted correlation of: US spending on science, space and technology correlated with Suicides by hanging, strangulation and suffocation. One thing interesting about this correlation is that with modern advancements in technology, it is easier to think of possible latent reasons for this correlation rather than counter-examples.

However, on face value, simply increasing the spending budgets for science, space and technology wouldn't induce more suicides. One important piece of information left out is whether or not this spending budget has been adjusted to account for inflation. Even if the dollar amount goes up, it doesn't necessarily mean that the percentage of relative value has increased from the last year.

In contrast, one factor that could increase suicide rates would be the amount of automation implemented by advances in technology, which drives people out of jobs. These forms of suicide are also fairly cheap to accomplish and the loss of income could push people who are already struggling in life over their breaking point.

## PROBLEM 5

### a)

$W \sim N(2-3+0,6+2+1)$

$W \sim N(-1,9)$

### b)

$Q = 2Y; Y \sim N(-3,2)$

$Q \sim N(2*-3,4*2)$

$Q \sim N(-6,8)$

### c)

$P = -2X+4; \quad 4 \sim N(4,0); \quad X \sim N(-2,6)$

$P \sim (-2*2,4*6)+(4,0)$

$P \sim (-4+4,24+0)$

$P \sim N(0,24)$

**d)**

$X \sim N(2,6)$

$M \sim aX + b : M \sim (0,1)$

$M \sim N(a * 2 + b, a^2 * 6)$

$a = \sqrt{\frac{1}{6}}$

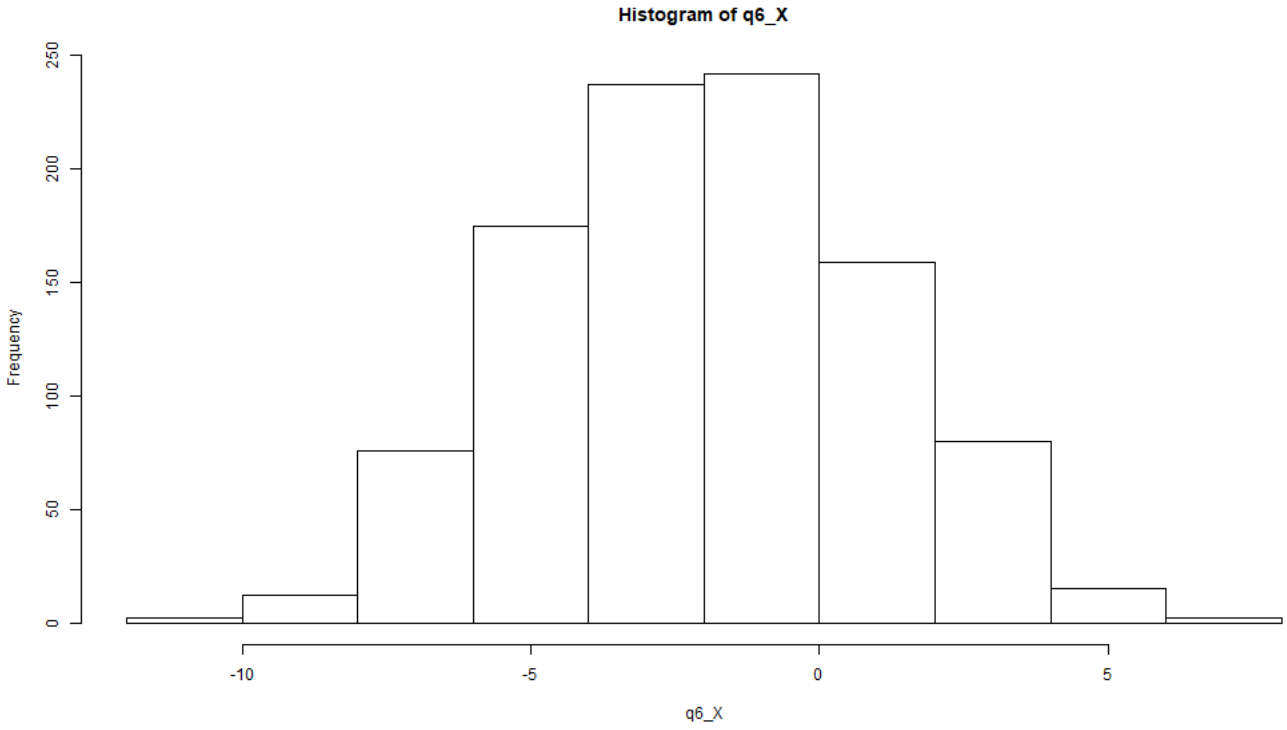
$b = -2 * a$

$M \sim N\left(\sqrt{\frac{1}{6}} * 2 - \sqrt{\frac{1}{6}} * 2, \left(\sqrt{\frac{1}{6}}\right)^2 * 6\right)$

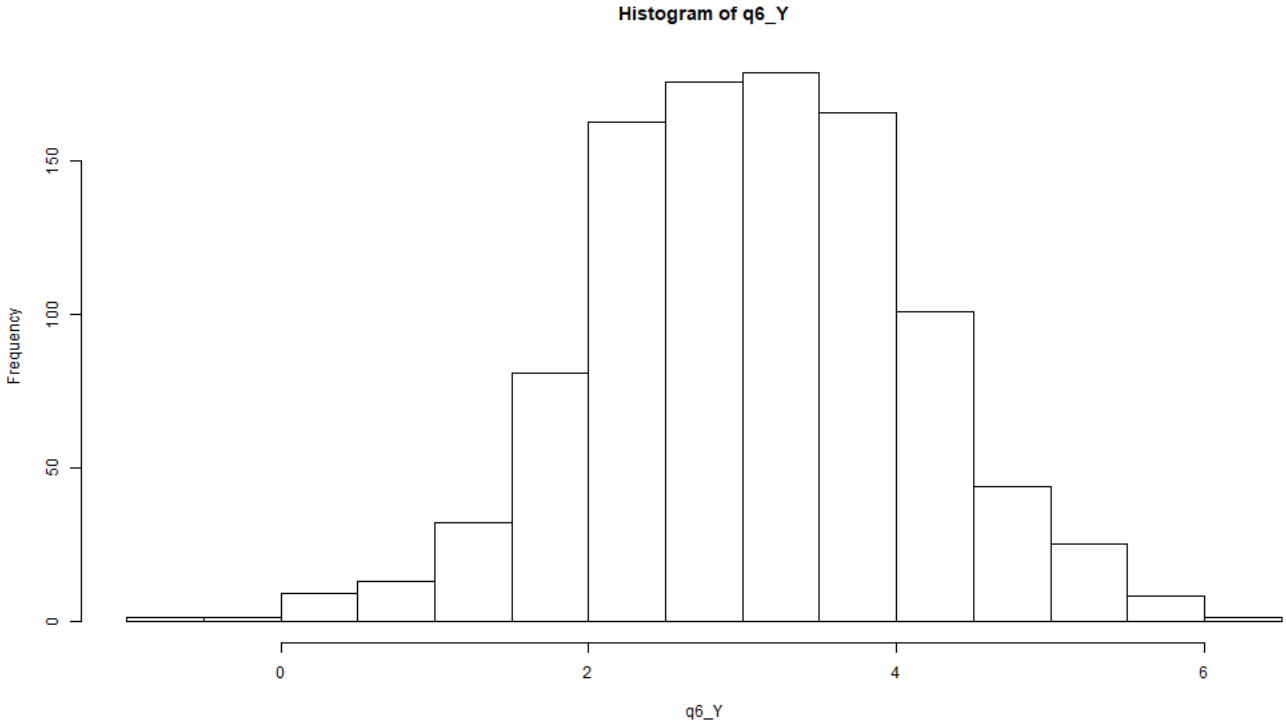
$M \sim N(0,1)$

## PROBLEM 6

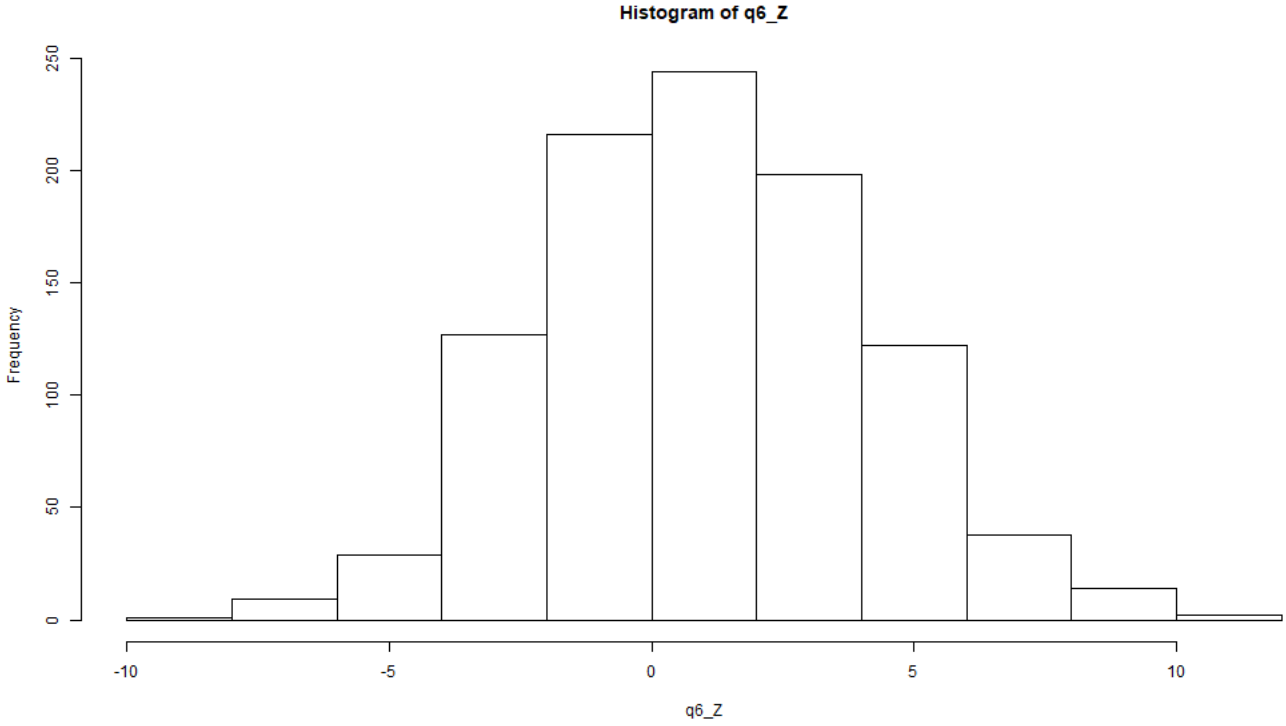
**a)**



**b)**


Histogram of q6_Y
Frequency
0
50
100
150
0
2
4
6
q6_Y

**c)**


Histogram of q6_Z
Frequency
0
50
100
150
200
250
-10
-5
0
5
10
q6_Z

## d)

Yes, although the distribution is a linear combination of the two independent distributions, the samples from Z are not dependent on samples from X or Y.

$P(Z|A,B) = P(Z)$

## e)

$Z = X + Y; \quad X \sim N(-2,3); \quad Y \sim N(3,1)$

$Z \sim (-2+3, 3+1)$

$Z \sim (1,4)$

```
> mean(q6_Z)
[1] 1.063592
> sd(q6_Z)
[1] 3.07499
```

# Code Appendix

```r
#####################################################################
#### Setup
#####################################################################
## Install and load libraries
# ipak function taken from: https://gist.github.com/stevenworthington/3178163
# ipak <- function(pkg) {
#   new.pkg <- pkg[!(pkg %in% installed.packages()[, "Package"])]
#   if (length(new.pkg))
#     install.packages(new.pkg, dependencies = TRUE)
#   sapply(pkg, require, character.only = TRUE)
# }
#
# packages <- c("ggplot2", "reshape2", "gridExtra", "TSA", "astsa", "orcutt",
#               "nlme", "fGarch", "vars")
# ipak(packages)

# Set up variables for first few questions
treatmentsSorted = c(rep("none",5), rep("low", 5), rep("medium", 5), rep("high", 5))
units = 1:length(treatmentsSorted)

#####################################################################
#### Problem 1
#####################################################################
sample1 = sample(treatmentsSorted)
experiment1 = data.frame(units, sample1)
experiment1

#####################################################################
#### Problem 2
#####################################################################
sample2 = sample(treatmentsSorted)
experiment2 = data.frame(units, sample2)
experiment2

#####################################################################
#### Problem 3
#####################################################################
q3treats = c(rep("r1", 3), rep("r2", 5), rep("r3", 5))
q3units = 1:length(q3treats)
q3sample = sample(q3treats)
q3experiment = data.frame(q3units, q3sample)
q3experiment


#####################################################################
#### Problem 6
#####################################################################
## Part A
q6_X = rnorm(1000, mean = -2, sd = 3)

png("./figures/p6_a.png", width = 1024, height = 576)
hist(q6_X)
dev.off()

## Part B
q6_Y = rnorm(1000, mean = 3, sd = 1)

png("./figures/p6_b.png", width = 1024, height = 576)
hist(q6_Y)
dev.off()

## Part C
q6_Z = q6_X + q6_Y

```

```r
png("./figures/p6_c.png", width = 1024, height = 576)
hist(q6_Z)
dev.off()


## Part E
# Z ~ N(1, 4)
mean(q6_Z)
sd(q6_Z)
```