Problem 1

The descriptive statistics for the three random vectors drawn from standard normal distribution are given in the table below.

	x (10)	y (10,000)	z (1,000,000)
Mean	-0.1329441	-0.0006762757	0.0004501281
Standard deviation	0.8136554	0.9997931	1.001349
Variance	0.6620352	0.9995862	1.0027

As the sample size increased, the sample mean approached the DGP mean of 0. Sample variance for larger samples is closer to 1. The variance of sample variance decreases as sample size increases, although we cannot conclude this from data.

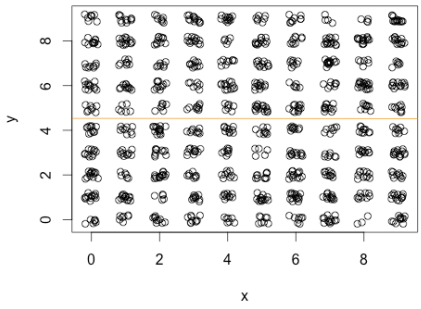
Problem 2

The raw generated series are given below together with a plot and a frequency table.

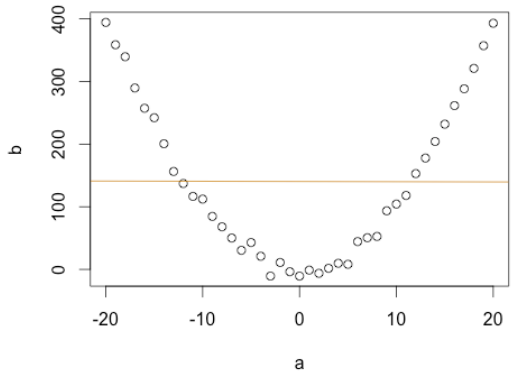
5,2,5,9,7,1,6,7,7,5,5,4,2,7,2,6,4,0,7,3,3,6,1,8,0,2,3,6,2,4,9,2,3,3,9,9,8,7,2,8,2,3,1,9,8,9,1,9,5,4,6,9,9,1,6,8,3,1,1,3,2,5,7,7,7,3,7,0,8,9,6,2,1,3,1,6,4,9,8,8,8,7,3,6,7,4,4,1,4,2,1,5,9,6,0,6,9,9,5,8,8,4,8,9,4,6,9,9,4,8,4,6,5,5,1,5,2,8,8,0,4,9,4,5,8,1,6,4,2,2,7,0,1,8,8,3,0,7,8,9,0,1,8,1,4,6,6,9,3,8,7,1,1,5,5,9,8,9,0,5,6,2,5,3,5,3,9,4,4,1,8,3,1,0,3,3,7,2,7,7,9,2,6,7,6,2,6,2,4,5,2,8,0,2,5,2,1,1,9,4,0,7,5,9,2,9,5,5,0,5,1,6,2,2,6,3,6,2,2,8,0,6,4,6,1,3,3,7,1,6,9,3,7,3,0,8,3,6,1,6,4,1,8,3,9,2,7,2,3,0,1,4,1,8,9,2,9,5,9,9,4,7,6,4,7,3,9,7,8,2,1,0,1,3,5,6,4,4,3,9,5,1,5,5,6,7,7,6,8,6,6,5,2,7,2,6,2,1,5,5,2,2,4,3,3,1,5,7,0,7,8,9,5,1,3,4,2,1,0,3,9,1,8,0,8,7,3,7,9,8,1,1,8,0,1,8,9,3,0,7,5,5,0,7,3,8,8,3,8,6,3,4,4,3,8,8,8,6,6,3,0,2,9,8,6,1,6,4,6,4,6,6,5,0,5,8,1,4,8,7,5,7,4,8,6,4,0,0,7,1,8,5,4,8,3,0,3,1,4,3,2,0,1,9,3,8,8,2,6,8,1,5,6,7,9,4,9,4,1,0,5,3,9,9,9,7,8,6,3,8,7,2,0,6,4,4,1,2,5,7,0,6,1,2,6,4,9,2,1,3,3,0,3,0,0,4,4,1,4,1,6,3,7,5,7,1,9,0,5,9,7,5,0,7,0,1,2,3,3,6,5,0,0,9,1,6,6,6,8,0,7,0,0,5,5,3,0,9,1,8,5,7,7,7,4,3,4,0,8,7,1,6,5,4,9,8,8,4,9,2,2,9,5,3,5,9,7,2,1,2,9,9,6,3,7,5,0,6,1,8,1,6,9,6,2,2,5,8,5,9,0,8,6,2,7,6,9,5,0,0,0,5,0,9,0,1,1,3,9,3,3,6,9,8,4,8,5,4,3,9,8,7,0,1,2,3,4,2,4,2,3,4,5,7,4,9,9,4,0,5,0,7,7,9,9,2,3,0,2,4,7,1,8,2,9,4,2,4,0,3,1,9,8,8,9,0,7,2,5,2,5,4,6,6,4,7,5,9,5,6,7,9,0,2,3,3,0,9,2,2,6,5,3,7,4,8,0,8,5,0,6,0,4,6,5,1,3,2,6,1,3,1,3,0,7,6,4,3,3,0,2,7,4,6,3,9,0,0,2,8,7,4,0,9,9,7,1,3,5,3,7,0,6,9,2,1,6,5,9,0,9,4,3,3,7,1,9,9,1,2,6,4,8,8,3,7,0,6,5,9,8,7,3,1,7,6,7,9,8,2,1,7,3,9,2,2,5,4,3,1,7,7,9,8,3,3,4,0,2,1,2,8,0,0,0,5,9,6,8,8,0,9,1,2,5,8,1,1,1,1,6,5,3,6,7,3,1,4,0,0,5,6,2,1,8,5,5,6,7,0,6,7,5,0,7,1,5,8,2,0,2,7,8,9,2,3,3,7,1,6,8,7,9,1,3,4,1,6,5,5,3,7,9,9,2,9,6,2,9,3,5,5,8,2,8,4,7,7,9,5,7,4,9,9,9,5,0,8,2,6,4,9,1,2,9,7,0,3,7,4,3,7,0,0,2,1,1,5,0,4,6,3,5,2,4,5,5,9,6,7,4,8,9,7,4,7,0,5,9,2,1,2,6,9,3,3,0,0,7,2,0,9,8,5,7,4,2,9,8,4,5,1,1,8,0,5,5,9,9,3,0,6,0,0,9,4,5,6,3,6,5,7,0,7,7,5,7,2,9,6,6,6,3,7,1,8,2,5,4,4,6,6,8,1,1,5,0,2,0,5,2,2,5,6,0,5,5,7,4,8,7,0,2,8,5,3,8,4,0,4,0,2,6,4,2,3,0,6,9,9

2,5,2,4,2,8,3,8,7,3,6,5,0,2,0,3,2,8,7,5,9,0,2,1,3,6,5,5,8,2,9,6,1,6,8,8,4,8,2,8,4,2,3,7,2,2,1,5,5,6,9,8,3,3,5,2,6,1,5,7,4,7,7,2,7,9,7,7,6,3,5,9,6,7,1,9,6,1,0,8,3,6,3,5,7,0,6,8,3,5,5,5,5,2,1,1,0,8,5,6,2,7,4,9,8,1,8,2,7,6,5,9,2,3,0,8,2,8,6,2,1,2,0,0,3,0,4,9,7,2,4,0,2,2,8,3,4,7,6,3,3,4,2,3,4,4,3,2,9,6,1,4,6,7,6,6,8,3,5,0,7,0,1,6,1,4,3,4,6,8,7,1,0,3,1,7,8,9,4,0,5,4,8,3,6,6,0,3,9,6,8,1,6,5,5,6,8,8,5,1,5,0,6,9,3,5,1,6,5,0,9,2,4,7,0,8,5,4,3,2,3,9,5,0,3,9,9,8,8,1,3,3,0,4,2,5,2,2,6,6,9,6,2,6,9,8,7,6,7,6,4,5,6,8,1,2,0,1,1,1,9,5,0,5,3,9,0,2,8,3,1,3,7,5,4,9,9,2,6,7,7,0,7,6,8,8,1,7,9,5,8,1,2,6,9,5,6,9,5,7,2,5,6,1,1,2,5,5,6,8,6,4,4,1,5,6,8,4,7,3,5,7,1,2,4,2,1,0,9,7,7,2,7,3,6,8,9,0,1,6,7,4,3,1,0,4,2,1,4,6,1,4,2,9,5,3,8,1,4,4,1,5,7,8,0,2,4,2,6,8,1,1,0,1,1,8,3,1,6,8,8,0,0,4,4,2,9,8,3,8,5,9,4,3,7,8,2,1,3,9,3,4,6,9,9,1,1,5,7,4,7,1,0,9,1,8,0,0,1,7,8,8,0,6,6,8,8,3,0,5,4,8,6,1,7,9,9,1,8,0,8,3,4,4,9,4,1,8,7,0,4,2,3,7,2,1,2,4,7,3,3,8,5,9,7,9,8,4,7,0,4,1,6,1,9,9,8,8,7,5,8,6,5,6,3,3,0,7,6,9,7,2,3,8,5,8,5,9,9,9,9,3,1,6,3,0,5,2,4,6,1,2,4,3,6,3,3,8,4,2,4,9,8,4,8,4,0,9,7,0,0,3,2,1,7,5,2,1,8,4,0,4,1,3,4,0,3,3,8,3,2,5,8,6,0,1,9,2,8,6,0,6,5,6,0,8,1,2,4,6,4,2,1,1,0,2,5,1,8,7,3,0,8,5,4,4,4,1,9,1,7,2,9,1,5,8,8,6,7,1,4,8,5,9,7,4,8,5,4,9,0,0,5,7,0,4,2,3,6,3,1,3,8,6,2,4,5,4,3,7,5,1,9,7,9,4,4,8,2,2,5,0,6,2,9,2,4,0,8,6,2,8,3,2,7,3,3,5,7,8,8,4,6,2,8,1,3,2,9,1,8,8,9,9,8,1,9,7,3,5,9,9,3,4,8,8,2,1,4,6,5,9,2,3,5,3,9,7,3,8,9,4,4,9,0,7,3,5,1,9,6,4,7,2,5,9,6,9,9,5,9,3,1,9,1,4,2,6,4,6,9,7,9,6,6,1,2,4,4,7,5,2,0,8,0,4,5,6,2,4,0,5,0,0,8,9,4,9,3,6,7,1,4,6,0,5,5,1,3,2,0,6,9,1,2,3,8,1,8,1,1,9,6,2,0,3,6,8,9,8,5,2,3,4,7,0,0,1,6,7,3,1,6,1,4,8,8,5,3,8,6,2,6,2,2,8,2,9,3,8,6,8,1,6,5,7,9,4,4,9,0,1,2,3,0,6,7,1,1,4,1,0,5,0,0,6,1,2,3,1,3,2,3,2,8,4,7,7,6,5,1,3,0,2,4,7,3,0,8,6,5,3,3,6,8,3,5,7,7,4,3,5,7,4,7,2,9,5,2,1,0,8,6,8,6,1,1,8,1,8,7,8,5,6,3,8,1,1,8,2,8,9,6,6,7,1,2,7,2,1,0,2,4,1,8,7,5,9,2,5,4,1,1,1,4,3,5,6,4,2,2,8,4,1,5,6,5,9,8,5,7,8,0,0,5,6,6,6,0,7,8,7,6,8,3,1,0,9,0,4,0,9,6,8,6,2,1,6,1,6,5,1,3,7,2,5,8,6,2,7,1,7,3,9,3,0,9,5,1,2

y \ x	0	1	2	3	4	5	6	7	8	9
0	7	7	10	10	7	8	8	13	3	11
1	13	14	5	10	10	13	15	12	9	13
2	13	6	14	8	9	10	13	10	11	10
3	11	10	10	11	8	5	12	7	11	14
4	11	9	15	9	10	9	9	7	10	8
5	8	6	7	9	7	16	12	12	9	9
6	12	11	11	7	13	12	6	8	17	13
7	8	8	7	10	9	11	8	14	6	6
8	11	14	11	13	5	13	9	15	15	13
9	8	11	8	13	12	7	9	4	4	15



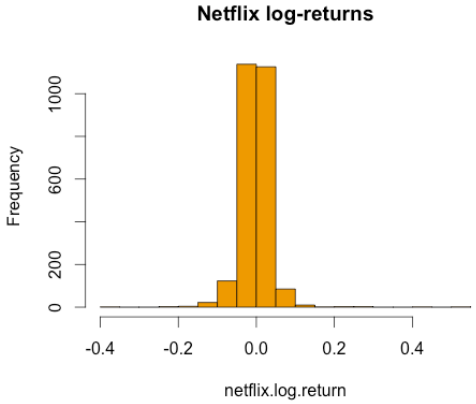
The regression coefficient in the linear model of y~x has t-statistic value -0.062, much less than 2 in absolute value. Based on the t-statistic, I would reject the hypothesis that there is a linear relationship between the two series. We cannot reject the hypothesis that there is any relationship between the two series. An example of two series that are related, but with a low t value for the regression coefficient is given in graph below.



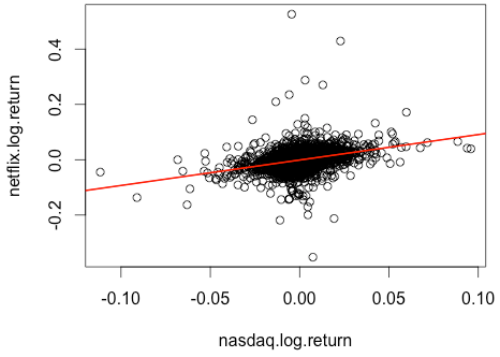
Series a and b show a strong relationship, but t value is -0.022

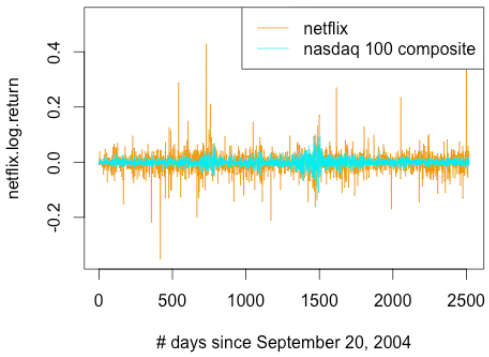
Problem 3

b.



c.





d.

The linear fit is

$$\widehat{\text{NFLX}} = -0.0009701 + 0.9208783 \times .\text{IXIS}$$

With 95% confidence intervals

$$\alpha \in (-0.002333502, 0.0003934015)$$
$$\beta \in (0.821543582, 1.0202130943)$$

so the 95% CIs for α and β include 0 and 1 respectively.

Problem 4

b.

Summary statistics:

```
       d                 x1
 Min.   :0.000   Min.   :0.0002863
 1st Qu.:0.000   1st Qu.:0.2493821
 Median :0.000   Median :0.4846455
 Mean   :0.477   Mean   :0.4873756
 3rd Qu.:1.000   3rd Qu.:0.7324941
 Max.   :1.000   Max.   :0.9990059
```

Standard deviation:

```
        d         x1
0.4997206 0.2833453
```

c.

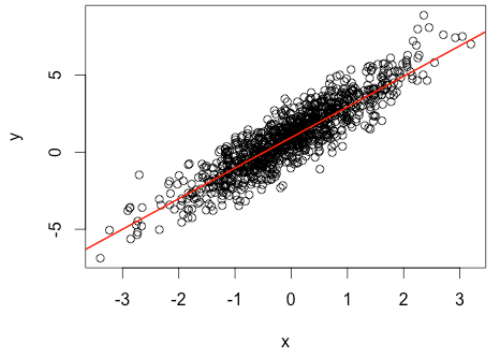
$$\hat{b} = -0.01795 + 1.01554\, x_1$$

d.

Predictions $\hat{b}(.65) = 0.642152$ and $\hat{b}(.99) = 0.9874363$, so with the threshold of 1, neither will be classified as spam.

Problem 5

b.



c.

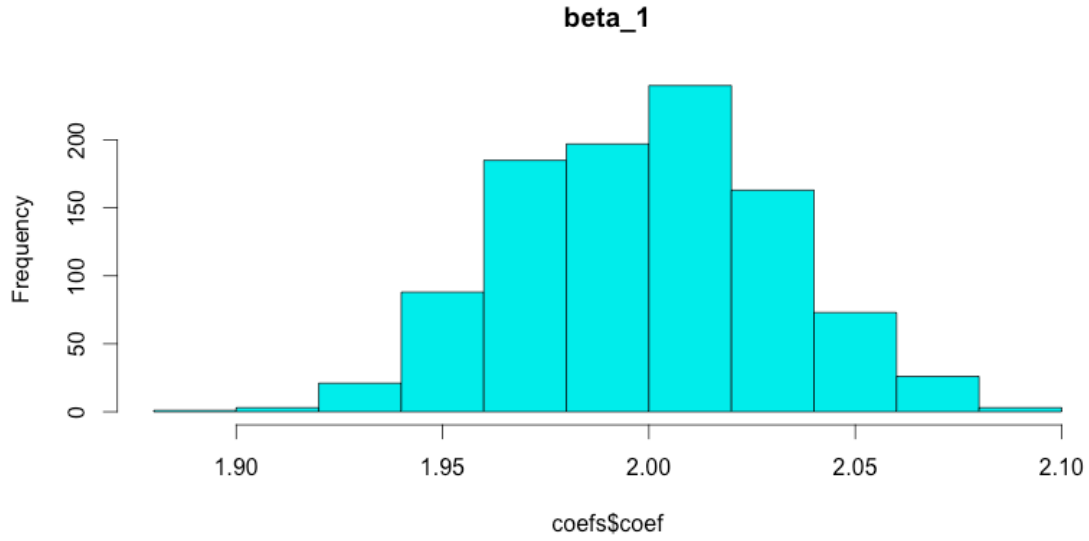
The estimated model is $\hat{y} = 0.95098 + 1.98407\, x$, with P-values for both coefficients $<< 0.01$.

d.

Generating the data and fitting the model five times yields the following values of $\hat{\beta}$: 2.055713, 1.991783, 1.992217, 1.962212, 2.006021.

e.

If the Monte Carlo simulation is performed 1000 times, the values of $\hat{\beta}$ should follow a bell-shaped distribution with centrality dependent on σ_ϵ (1) and the number of observations n simulated at each iteration of the MC simulation (1000). With known σ_ϵ, I think it may be a normal distribution with mean 2 and variance $\sigma_\epsilon/(n\ \mathrm{Var}(x))$.



f.

Histogram of $\exp(\beta_1)$ results of the MC simulation.