HW2

Jiamin Xuan

[Jx624@nyu.edu](mailto:Jx624@nyu.edu)

1.

a) **10 draws:**

> mean(x1) : [1] -0.3514216

> var(x1):[1] 1.497763

> sd(x1):[1] 1.223831

b)**10,000 draws:**

> mean(x2):[1] -0.0138902

> var(x2):[1] 0.9787875

> sd(x2):[1] 0.9893369

c)**1,000,000 draws:**

> mean(x3):[1] 0.001993352

> var(x3):[1] 1.003208

> sd(x3):[1] 1.001603

d) According to the law of large numbers, the mean is getting close to 0, and variance, standard deviation is getting close to 1.

e) Please see attached code.

2.

a) Please see the attached text files

b) y=0.05594x+4.15239

Call:lm(formula = y ~ x)

Residuals:

Min 1Q Median 3Q Max

-4.6558 -2.4321 -0.2083 2.5679 4.8476

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.15239 0.16083 25.818 <2e-16 \*\*\*

x 0.05594 0.03080 1.816 0.0697 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.832 on 998 degrees of freedom

Multiple R-squared: 0.003294, Adjusted R-squared: 0.002295

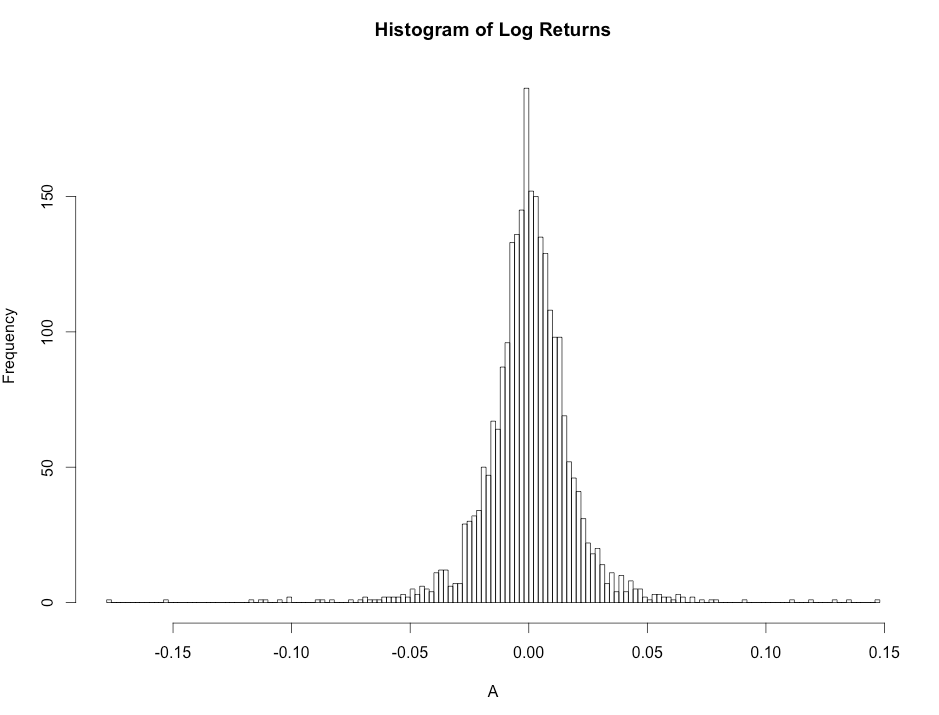
F-statistic: 3.298 on 1 and 998 DF, p-value: 0.06966

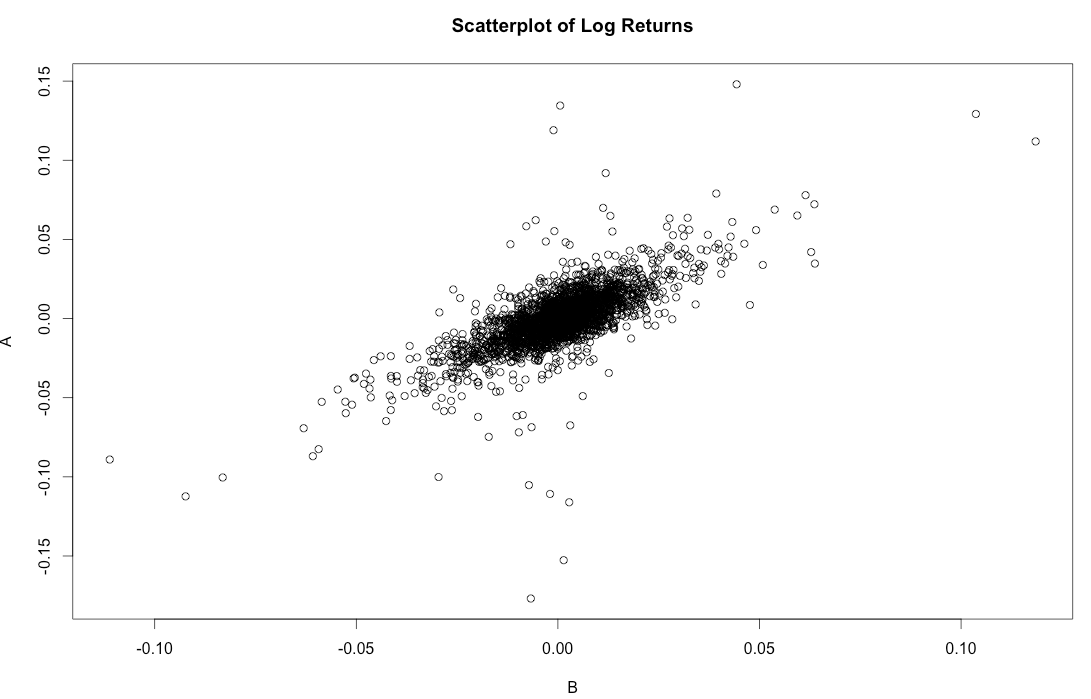
c) T-statistic is 1.816, the absolute value of which is no greater than 2. So we fail to reject the null hypothesis (no relationship between two series)

d) Please see attached code and files

3.

a) please see attached csv files and code

b)

c)

d) Estimated “alpha” is -0.0003168 and the estimated “beta” is 1.0092986

2.5 % 97.5 %

(Intercept) -0.0008424302 0.0002089022

B 0.9713263866 1.0472708429

Alpha: (-0.0008424302, 0.0002089022)

So Alpha is not significantly different than zero at a 95% level

Beta: (0.9713263866, 1.0472708429)

Yes, 95% confidence level for “beta” include one

e) please see attached csv files and code

4.

a) please see attached dta files and code

b)

> summary(train)

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

> table(train[,1])

0 1

523 477

c)

Call:

lm(formula = d ~ x1, data = train)

Coefficients:

(Intercept) x1

-0.01795 1.01554

The estimated coefficient is 1.01554

d) predicted value is 0.642152, would not classify it as spam.

e) predicted value is 0.9874363, would not classify it as spam.

f) please see attached dta files and code

5.

c)Call:

lm(formula = y ~ x)

Residuals:

Min 1Q Median 3Q Max

-3.1102 -0.6469 0.0183 0.7168 2.9349

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.99838 0.03217 31.04 <2e-16 \*\*\*

x 2.01242 0.03214 62.62 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.017 on 998 degrees of freedom

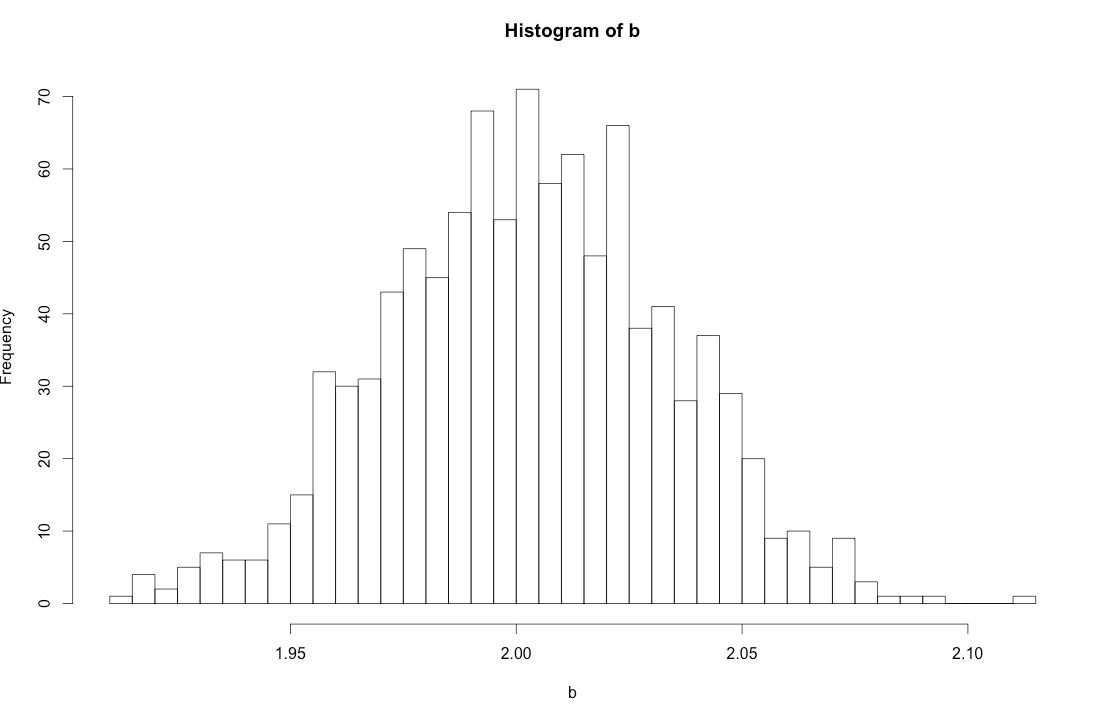
Multiple R-squared: 0.7971, Adjusted R-squared: 0.7969

F-statistic: 3921 on 1 and 998 DF, p-value: < 2.2e-16

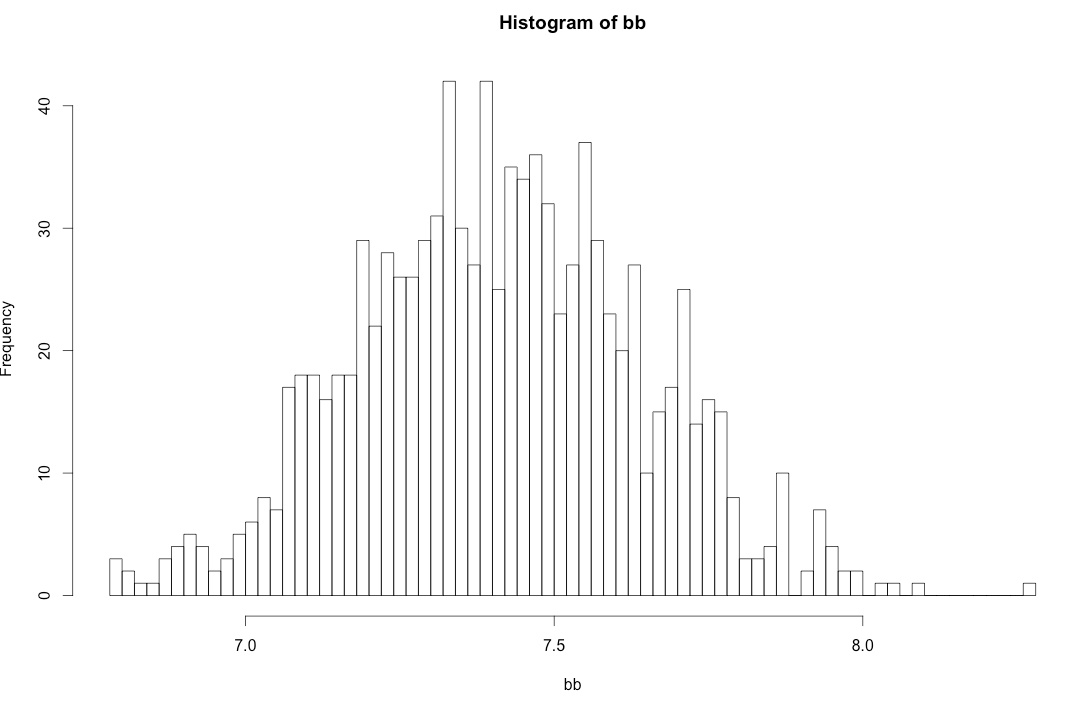
The estimate intercept and slope is 0.998 and 2.012 which close to 1 and 2

d) it’s close to result in c)

e)



f)transform b into exp(b)



g) see attached file and code