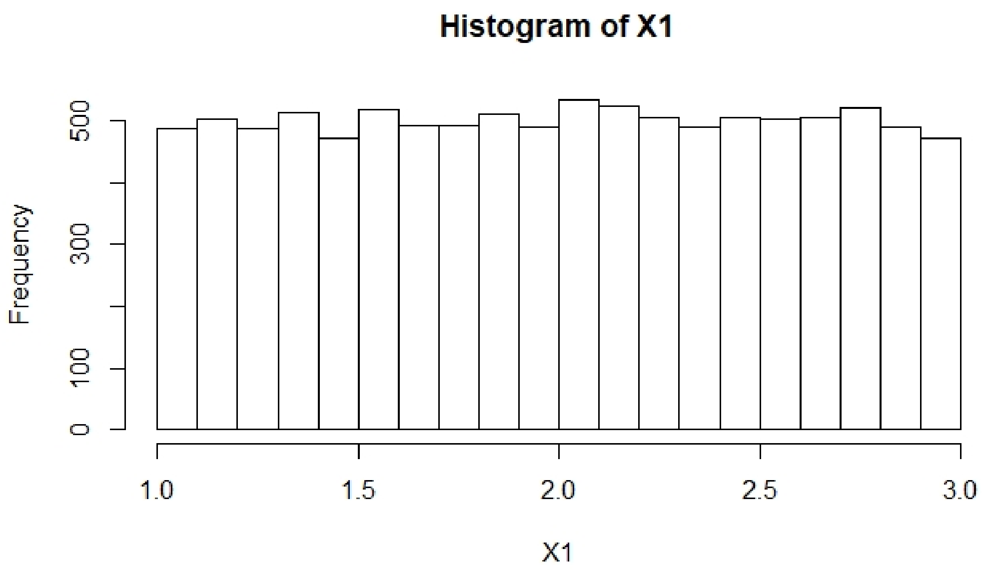
ECON 613 Applied Econometrics Homework2

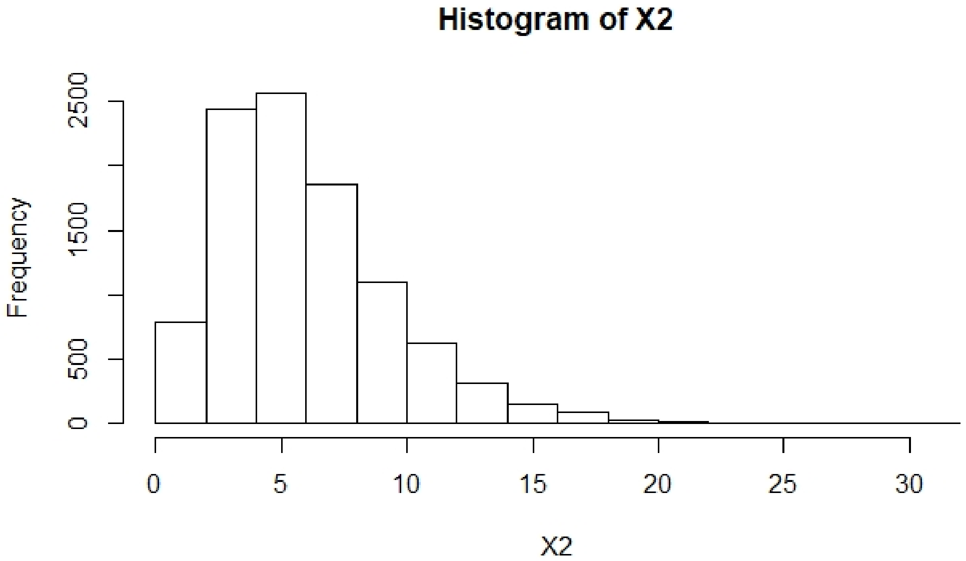
Promvarat Pradit

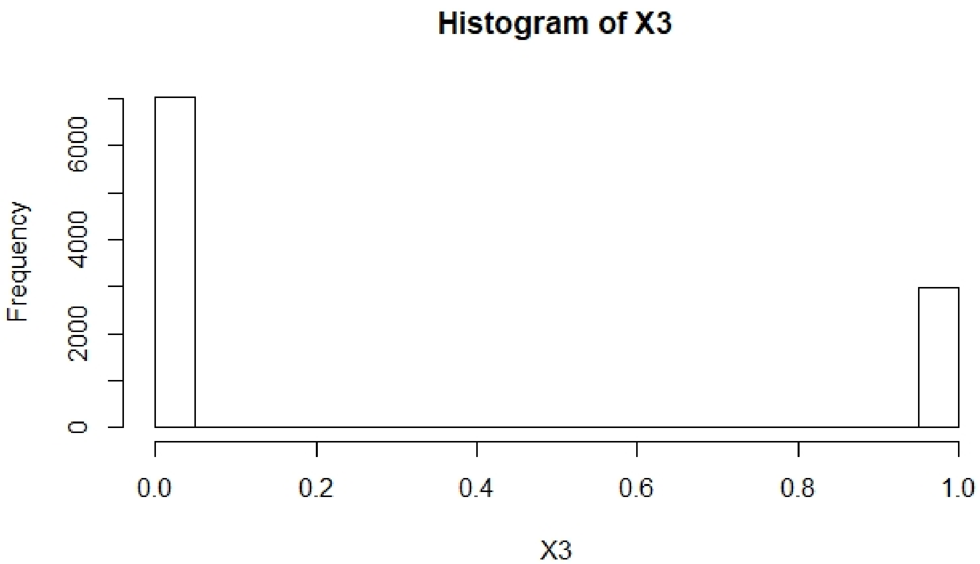
Febuary 12, 2019

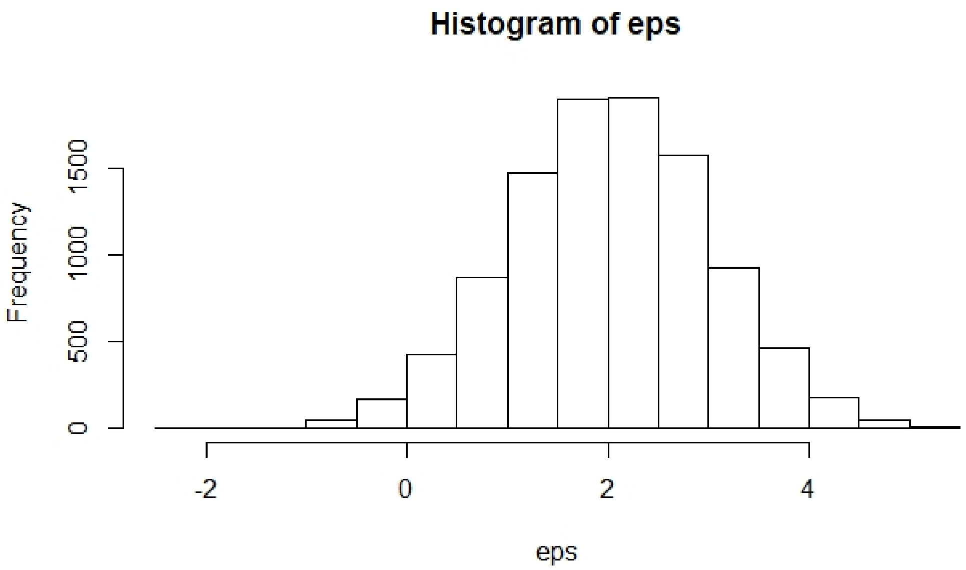
Exercise 1 Data Creation

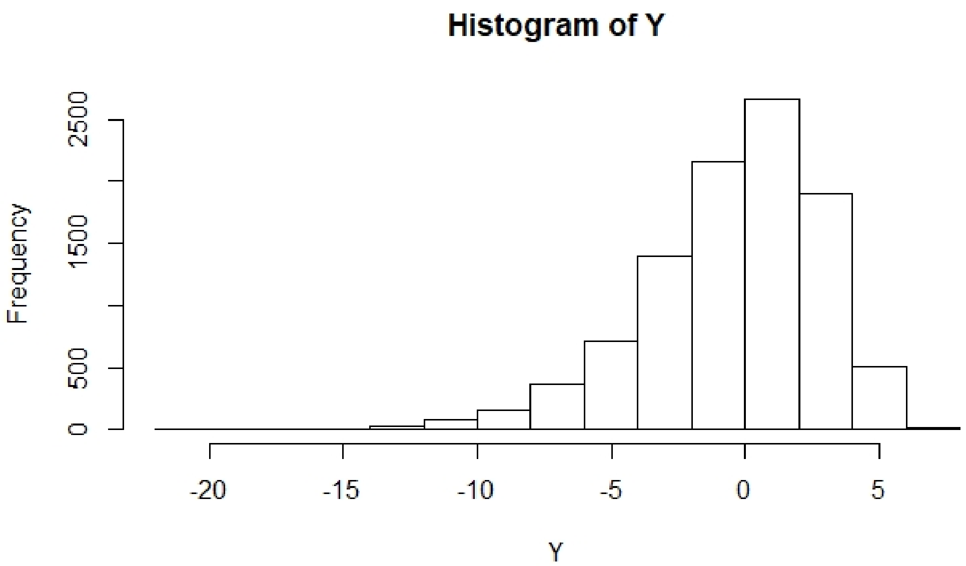
Histogram of the simulated variables will be reported here

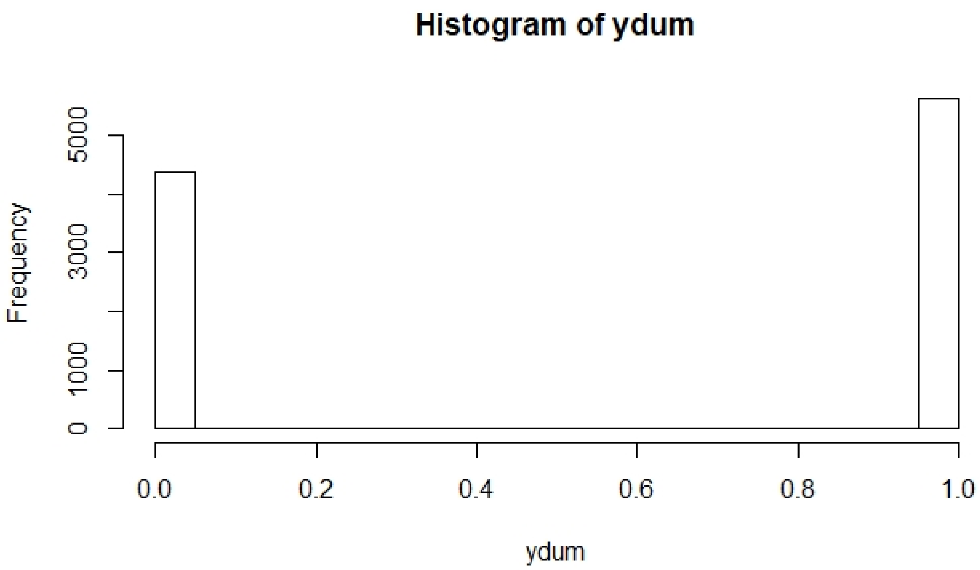



Histogram of X2
Frequency
0
500
1500
2500
0
5
10
15
20
25
30
X2


Histogram of X3
Frequency
0
2000
4000
6000
0.0
0.2
0.4
0.6
0.8
1.0
X3


Histogram of eps
Frequency
0
500
1000
1500
-2
0
2
4
eps


Histogram of Y
Frequency
0
500
1500
2500
-20
-15
-10
-5
0
5
Y


Histogram of ydum
Frequency
0
1000
3000
5000
0.0
0.2
0.4
0.6
0.8
1.0
ydum

Exercise 2 OLS

- Correlation between Y and X1

Correlation betwwen Y and X1 is approximately 0.2018 which is way less than the actual value of 1.2. This is because correlation reflects only how much Y and X1 move in accordance with each other, not the marginal effect of X1. From exercise 1, we know that $Y = 1.2X1 - 0.9X2 + 0.1X3 + eps$), X1 explains only some portion of Y, so the correlation of 0.2018 is appropriate. In addition, the value of correlation is between -1 to 1.

- OLS Coefficients

Using formula $\hat{\beta} = (X'X)^{-1}X'Y$, estimation result is shown in column 1 of Figure 1. The second column of the figure is the result obtained from R built-in package. Both of the estimates are very similar, and they are also close to the actual value.

Figure 1: OLS Coefficient

	Standard formular	R OLS package
c	2.4866583	2.48667726
X1	1.2181458	1.21809470
X2	-0.8984368	-0.89860058
X3	0.0810096	0.08064399

- OLS Standard error

Standard error from each method is reported in Figure 2. Overall, all methods provide similar results. R built-in package gives the closest SE to the Standard formula $(SE(\hat{\beta}) = \sigma^2(X'X)^{-1})$. For bootstrap, the larger the number of replications is the closer the value of SE to the one from Standard formula.

Figure 2: OLS Standard error

	Standard formular	Bootstrap 49	Bootstrap 499	R OLS package
c	0.040600003	0.039812673	0.03919052	0.040606137
X1	0.017439030	0.016871321	0.01750368	0.017441654
X2	0.002894556	0.003140135	0.00302345	0.002895197
X3	0.021873127	0.019786333	0.02143240	0.021876482

Exercise 3 Numerical Optimization

The function that return probit's likelihood and implementation of steepest ascent algorithm is written in the associated R script file. The result of the optimization is reported in Figure 3. The optimization result in a coefficients that have correct sign. However, the relative magnitudes differ from the actual parameter, which is probably caused by my algorithm having large tolerance level so the steepest ascent stop searching before the actual global maximum is found.

Figure 3: Probit Coefficient from steepest ascent optimization

	Steepest ascent	Actual value
intercept	0.6532120	0.5
X1	1.1385014	1.2
X2	-0.5035009	-0.9
X3	0.1627279	0.1

Figure 4: Model comparison

	Dependent variable:		
	ydum		
	Probit	*Logistic*	*Linear Probability*
	(1)	(2)	(3)
X1	1.128***	2.037***	0.142***
	(0.042)	(0.079)	(0.006)
X2	-0.882***	-1.591***	-0.104***
	(0.018)	(0.036)	(0.001)
X3	0.093**	0.167**	0.020***
	(0.047)	(0.084)	(0.007)
Constant	2.985***	5.381***	0.895***
	(0.099)	(0.186)	(0.013)
Observations	10,000	10,000	10,000
R^2			0.551
Adjusted R^2			0.551
Log Likelihood	-2,215.375	-2,217.146	
Akaike Inf. Crit.	4,438.750	4,442.293	
Residual Std. Error			0.333 (df = 9996)
F Statistic			4,088.631*** (df = 3; 9996)
Note:			*p<0.1; **p<0.05; ***p<0.01

Exercise 4 Discrete Choice

Estimation results from probit, logit and linear probability model is shown in Figure 4. Overall, coefficient for each variable have same sign in all 3 models. X1 and X2 have positive effect on probability of ydum = 1; if X1 or X2 increase, ceteris paribus, ydum will have **more** chance of being 1. On the contrary, X3 has negative coefficient; if X3 increases, ydum will have **less** chance of being 1. Moreover, all of these coefficients are significant at confidence level of 95%. These results are consistent with the actual parameter value, exercise 1.

Magnitude wise, we can only intepret the magnitude of linear probabiliy model because its coefficient reflects marginal effect while probit and logit coefficient reflect the effect on latent variable. From linear probability model, an increase of 1 unit of X1 **increases** the chance that ydum will be 1 by 0.142%, and an increase in X3 by 1 unit **increases** the chance that ydum will be 1 by 0.02%. In contrast, a unit increase in X2 is associated with the **reduction** in the chance ydum will be 1 by 0.104%.

Exercise 5 Marginal Effects

Compute marginal effect

Using average marginal effects (ME) in sample method, ME can be compute by this process.

- Compute $\hat{ydum_i} = X_i\beta$ for each observatons.
- Compute pdf, $F'(X_i\beta)$, which is the first derivative of $F(X\beta)$ for each value of $X_i\beta$. Probit = normal distribution pdf, and Logit = logistic distribution pdf.
- Compute average value of all pdf $\frac{1}{n}\sum(F'(X_i\beta))$
- Compute marginal effect by $ME = \frac{1}{n}\sum(F'(X_i\beta))\beta$

The result is demonstrated in Figure 5

Figure 5: Marginal Effect of Probit and Logit model

	Probit ME	Logit ME
(Intercept)	0.36780949	0.36639804
X1	0.13892065	0.13869164
X2	-0.10866015	-0.10832145
X3	0.01149834	0.01138724

Compute standard deviation

The SD for marginal effect calculated by both delta method and bootstrap is reported in Figure 6. Overall, both practices give similar result. However, delta method give smaller SD, which is expected because delta method calculates asymptotic SD which assume n approaches infinity. The steps taken to obtain result from delta method is outlined below.

- Delta method
  1. Compute $\hat{ydum_i} = X_i\beta$ for each observatons.
  2. Compute pdf, $F'(X_i\beta)$, which is the first derivative of $F(X\beta)$ for each value of $X_i\beta$. Probit = normal distribution pdf, and Logit = logistic distribution pdf.
  3. Compute Jacobian matrix, $J = \frac{1}{n}F'(X\beta)X$
  4. Compute coefficient var-cov matrix from Probit and Logit model, V
  5. Compute Var$(\epsilon), \sigma^2 = JVJ^t$
  6. Compute SD of each coefficient β_i ME by $SD(\beta_i) = \sqrt{\frac{\sigma^2}{Var(x_i)}}$

Figure 6: SD of Marginal Effect

	Probit SD Delta	Probit SD Bootstrap	Logit SD Delta	Logit SD Bootstrap
X1	0.0045611300	0.004784985	0.0042715956	0.004783979
X2	0.0007570389	0.001323131	0.0007089831	0.001322472
X3	0.0057204904	0.005873441	0.0053573614	0.005888638