ECON 7320 Advanced Microecnometrics

Assignment III

Student name: Chi-Sheng Lo
Student ID number: 44046918

Q1. Hypothesis Testing

Given:

y = β_0 + xβ_1 + ε

ε ~ N(0, σ^2) in which x is independent of ε

(i) Write down the conditional distribution of y given x

y|x ~ N(β_0 + xβ_1, σ^2)

(ii) Write down the likelihood and log likelihood functions

Likelihood function:

$$L(\beta, \sigma^2) = \prod_{i=1}^{N} f(\varepsilon_i) = \prod_{i=1}^{N} f(y_i - \beta_0 - x_i\beta_1)$$

$$= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\boldsymbol{\sigma}^2}} \exp\left[- \frac{(y_i - \beta_0 - x_i\beta_1)^2}{2\sigma^2}\right] = \left(\frac{1}{2\pi}\right)^{N/2} (\sigma^2)^{-N/2} \exp\left[- \frac{(y_i - \beta_0 - x_i\beta_1)^2}{2\sigma^2}\right]$$

Log likelihood function:

$$\ln L(\beta, \sigma^2) = -\frac{N}{2}\ln 2\pi - \frac{N}{2}\ln\sigma^2 - \frac{1}{2\boldsymbol{\sigma}^2}\sum_{i=1}^{N}(y_i - \beta_0 - x_i\beta_1)^2 = -\frac{N}{2}\ln(2\pi\sigma^2) - \frac{\sum_{i=1}^{N}(y_i - \beta_0 - x_i\beta_1)^2}{2\sigma^2}$$

(iii) Find the maximum likelihood estimators $\boldsymbol{\beta_0}$ and $\boldsymbol{\beta_1}$

$$\frac{\partial \ln L}{\partial \beta_0} = -\frac{2\sum_{i=1}^{N}(-1)(y_i - \beta_0 - x_i\beta_1)}{2\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{N}(y_i - \beta_0 - x_i\beta_1) = 0$$

$$\frac{\partial \ln L}{\partial \beta_1} = -\frac{2\sum_{i=1}^{N}(-x_i)(y_i - \beta_0 - x_i\beta_1)}{2\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{N} x_i(y_i - \beta_0 - x_i\beta_1) = 0$$

$\beta_0 = \bar{y} - \beta_1\bar{x}$

$\sum_{i=1}^{N} x_i(y_i - \beta_0 - x_i\beta_1) = 0$

$\sum_{i=1}^{N} x_i y_i - x_i\beta_0 - x_i^2\beta_1 = 0$

$\sum_{i=1}^{N} x_i y_i - x_i(\bar{y} - \beta_1\bar{x}) - x_i^2\beta_1 = 0$

$\sum_{i=1}^{N} x_i y_i - x_i\,\bar{y} + x_i\beta_1\bar{x} - x_i^2\beta_1 = 0$

$\sum_{i=1}^{N}(y_i - \bar{y} + \beta_1\bar{x} - x_i\beta_1)\,x_i = 0$

$\sum_{i=1}^{N} y_i - \bar{y} + \beta_1(\bar{x} - x_i) = 0$

$\sum_{i=1}^{N}(y_i - \bar{y}) = -\beta_1\sum_{i=1}^{N}(\bar{x} - x_i)$

Finally:

$\widehat{\beta_1} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N}(x_i - \bar{x})^2}$ and $\widehat{\beta_0} = \bar{y} - \bar{x}\widehat{\beta_1}$

(iv) Test H_0: β_1 = 0 against a two-sided alternative

LR = − 2 (−143.361 + 145.878) = 5.034

Since the LR is 5.034 which is larger than the critical value of χ^2 at 3.84, we reject the H_0: β_1 = 0 at five percent significance level.

(v) Use θ = (β_0, β_1, σ^2) ^{T} and output from part (iv) to test H_0: β_1 = 0 against a two-sided alternative

t = $\frac{\widehat{\beta_1}}{\widehat{SE}[\widehat{\beta_1}]} = \frac{0.2560}{\sqrt{0.0091}} = \frac{0.2560}{0.0954}$ = 2.683

Since the t-statistics is 2.683 which is larger than the critical value at 1.984, We reject the H_0: β_1 = 0 at five percent significance level.

Q2. Panel Data

Given:

$y_{it} = \alpha_i + \beta_i x_{it} + \varepsilon_{it}$

Where β_i ~ N(β, σ_β^2), α_i ~ N(α, σ_α^2), and ε_{it} ~ N(0, σ_ε^2) are independent of one another and of x_{it}.

(i) Write down the likelihood function and the corresponding log likelihood function

Likelihood function:

$$L = \prod_{i=1}\prod_{t=1}(2\pi\sigma_\varepsilon^2)^{-1/2}\exp\left[-\frac{(y_{it}-\alpha_i-\beta_i x_{it})^2}{2\sigma_\varepsilon^2}\right]$$

Log likelihood function:

$$\ln L = \sum_{i=1}\sum_{t=1}\ln\left[(2\pi\sigma_\varepsilon^2)^{-1/2}\right] + \left[-\frac{(y_{it}-\alpha_i-\beta_i x_{it})^2}{2\sigma_\varepsilon^2}\right]$$

$$= -\frac{NT}{2}\ln 2\pi\sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2}\sum_{i,t}(y_{it}-\alpha_i-\beta_i x_{it})^2$$

(ii) Comment on the feasibility of maximum likelihood estimation of the parameters

Since β_i ~ N(β, σ_β^2) where: $\beta_i = \beta + \sigma_\beta\eta_i$ and η_i ~ N(0,1)

Then:

Likelihood function:

$$L = \prod_{i=1}\prod_{t=1}(2\pi\sigma_\varepsilon^2)^{-1/2}\exp\left[-\frac{(y_{it}-\alpha_i-\beta_i x_{it}-\sigma_\beta\eta_i x_{it})^2}{2\sigma_\varepsilon^2}\right]$$

Log likelihood function:

lnL = $-\frac{NT}{2}\ln 2\pi\sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2}\sum_{i=1}\sum_{t=1}(y_{it} - \alpha_i - \beta_i x_{it} - \sigma_\beta \eta_i x_{it})^2$

Comment

Since η_i is unobserved, then the likelihood cannot be maximized directly. Therefore, it is not feasible to conduct maximum likelihood estimation of the parameters for this case. See part (iii) for alternative solution.

(iii) Describe how to obtain a simulated maximum likelihood estimator of the parameters

Step 1: Let $\eta_i^{(s)}$ be a draw from N(0,1):

L($y_{it}|x_{it}$, η_i = $\eta_i^{(s)}$) = $\prod_{i=1}\prod_{t=1}(2\pi\sigma_\varepsilon^2)^{-1/2}exp\left[-\frac{(y_{it}-\alpha_i-\beta_i x_{it}-\sigma_\alpha \eta_i^{(s)})^2}{2\sigma_\varepsilon^2}\right]$

This step makes unobservables become observables.

Step 2: Converts $\eta_i^{(s)}$ by averaging out the effect of $\eta_i^{(s)}$:

L($y_{it}|x_{it}$) = E_η[L($y_{it}|x_{it}, \eta_i$)] = $\int$ L($y_{it}|x_{it}$)f(η_i)dη_i $\cong \frac{1}{S}\sum_{s=1}$ L($y_{it}|x_{it}, \eta_i \ = \ \eta_i^{(s)}$) for large S

This step ensures that L($y_{it}|x_{it}, \eta_i \ = \ \eta_i^{(s)}$) is converted to L($y_{it}|x_{it}$).

Step 3: Sets maximum simulated likelihood estimator that maximizes (with respect to θ):

$\frac{1}{S}\sum_{s=1}$ L($y_{it}|x_{it}, \eta_i \ = \ \eta_i^{(s)}$)

Q3. Quantile Regression

(i) Write down the corresponding quantile regression (QR) model

Since the dependent variable ltotexp = ln(totexp), the equivariance property of QR comes into play. Given Q_q(ln(y)|x, d) = $x^T \beta_q$, we have Q_q(y|x, d) = exp[Q_q(ln(y)|x, d)] = exp($x^T \beta_q$). More precisely, the marginal effect is expressed as $\frac{\partial Q_q(\text{y|x,d})}{\partial x_j}$ = exp($x^T \beta_q$) β_j.

At 0.1 quantile:

$$\{\widehat{lntotexp}\} = \underset{(0.4022)}{3.8670} + \underset{(0.0260)}{0.5392}totchr_1 + \underset{(0.0053)}{0.0193}age_i - \underset{(0.0704)}{0.0127}female_1 + \underset{(0.1951)}{0.0734}white_1 + \underset{(0.0712)}{0.3957}suppins_1$$

At 0.9 quantile:

$$\{\widehat{lntotexp}\} = \underset{(0.5086)}{8.3226} + \underset{(0.0337)}{0.3580}totchr_1 + \underset{(0.0067)}{0.0059}age_i - \underset{(0.0968)}{0.1576}female_1 + \underset{(0.2350)}{0.3052}white_1 - \underset{(0.0853)}{0.0143}suppins_1$$

(ii) Interpret the regression results, focusing on the suppins

People with supplementary insurance (suppins) has a total expenditure about 39.57% higher than those with no suppins, holding other variables constant, at the 0.10 conditional quantile. However, people with suppins has a total expenditure about 1.43% lower than those without, holding other variables constant, at the 0.90 conditional quantile.

Apparently, suppins has higher influence at the 0.10 conditional quantile than at the 0.90 conditional quantile of expenditure. Further, the standard error of 0.10 conditional quantile of expenditure is also smaller suggesting better precision at the left of distribution. Besides, the health-status variable (totchr) also has higher impact at the 0.10 conditional quantile. Among the three sociodemographic variables, both age and female have higher impacts at the lower conditional quantile and white has lower impact at the 0.10 conditional quantile than at the 0.90 conditional quantile.

(iii) Use the STATA output to obtain quantile treatment effect for a 65-year old white female with no chronic conditions and q = 0.9

At 0.9 quantile:

$x^T\widehat{\beta_q}$ = constant + $0.3580totchr_1$ + $0.0059age_i$ + $-0.1576female_1$ + $0.3052white_1$

= 8.3226 + (0 × 0.3580) + (65 × 0.0059) + (1 × − 0.1576) + (1 × 0.3052) = 8.85

Treatment effect at 0.9 quantile:

T(x, q) = Q_q(y|x, 1) − Q_q(y|x, 0) = $\left[\exp\left[\ln\left[\frac{Q_q(\text{y}|\text{x},1)}{Q_q(\text{y}|\text{x},0)}\right]\right] - 1\right] \times \exp\left[x^T\widehat{\beta_q}\right]$

= [exp[− 0.0143] − 1] × exp[8.85] = − 0.0142 × 6974.39= − 99.04

(iv) Explain how to construct a 0.95 confidence interval for the quantile treatment effect in part (iii)

To construct the confidence interval (CI) for QTE, I need to acquire the standard error of QR estimator which comes from the general condition in the following:

$\widehat{\beta_q} \xrightarrow{a}$ N(β_q, A^{-1}BA^{-1}) where A = $\sum_i q(a-1)x_i x_i^T$ and B = $\sum_i f_{u_q}(0|x_i)\, x_i x_i^T$.

$f_{u_q}(0|x_i)$ is the conditional density of the error term u_q = y − $x^T\beta_q$ evaluated at u_q = 0.

During this process, I compute the variance specifically for the quantile regression with bootstrap estimator. At first, I draw a bootstrap sample: ($y_i^b, x_i^{T\,b}$) in which b = 1,…, B, i = 1,…, N, from the empirical distribution of (y_i, x_i^T). Then we estimate the conditional quantile function $x_i^{T\,b}\widehat{\beta_q^b}$ where $\widehat{\beta_q^b}$ is the bootstrap estimate of β_q. Finally, the bootstrap estimate of variance is described as: $\widehat{VAR}[\overline{\beta_q}]$ = $\frac{N}{B}\sum(\widehat{\beta_q^b} - \overline{\beta_q^b})\,(\widehat{\beta_q^b} - \overline{\beta_q^b})^T$ where $\overline{\beta_1^b}$ = $B^{-1}\sum\widehat{\beta_q^b}$. Thereafter, the standard error is described as $\widehat{SE}[\overline{\beta_q}]$ = $\sqrt{\widehat{VAR}[\overline{\beta_q}]}$. Recall that the treatment effect at 0.9 quantile from part (iii) is: T(x, q) = Q_q(y|x, 1) − Q_q(y|x, 0) = $\left[\exp\left[\ln\left[\frac{Q_q(\text{y}|\text{x},1)}{Q_q(\text{y}|\text{x},0)}\right]\right] - 1\right] \times \exp\left[x^T\widehat{\beta_q}\right]$ = − 99.04. Therefore, I can construct a two-sided 95% CI for the quantile treatment effect in part (iii): $\left(-99.04 - 1.96\widehat{SE}\left[\overline{\beta_q}\right], -99.04 + 1.96\widehat{SE}\left[\overline{\beta_q}\right]\right)$.

Q4. Data Analysis

Given:

$y^* = \beta x - \varepsilon$

$$y = \begin{cases} 1 \text{ if } y^* \geq 0 \\ 0 \text{ otherwise} \end{cases}$$

where y^* is a latent variable and ε and x are independently distributed.

Requirements for simulation:

1. N = 100 and β = 1

2. x follows a normal distribution with mean 0 and variance 1

(i) Show that above is the probit model when ε follows the standard normal distribution

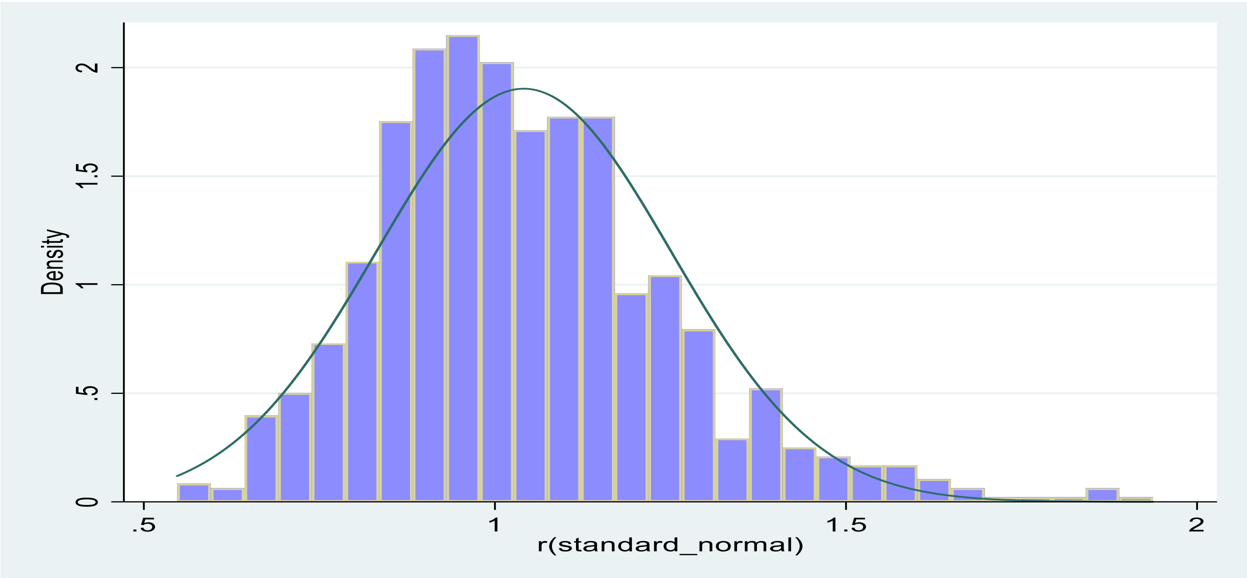
ε ~ NID(0,1)

$P(y = 1|x) = P(y^* \geq 0|x) = P(\beta x - \varepsilon \geq 0|x) = P(\varepsilon \leq \beta x) = \Phi(\beta x)$

(ii) Conduct a Monte Carlo experiment:

(a) The properties of the probit estimator of β when ε follows the standard normal distribution

Observations	Mean	Standard Deviation	Min	Max
1000	1.0411	0.2096	0.5475	1.9371



Simulation setting: Seed at 703 with 1000 repetitions

Skewness/Kurtosis tests for Normality

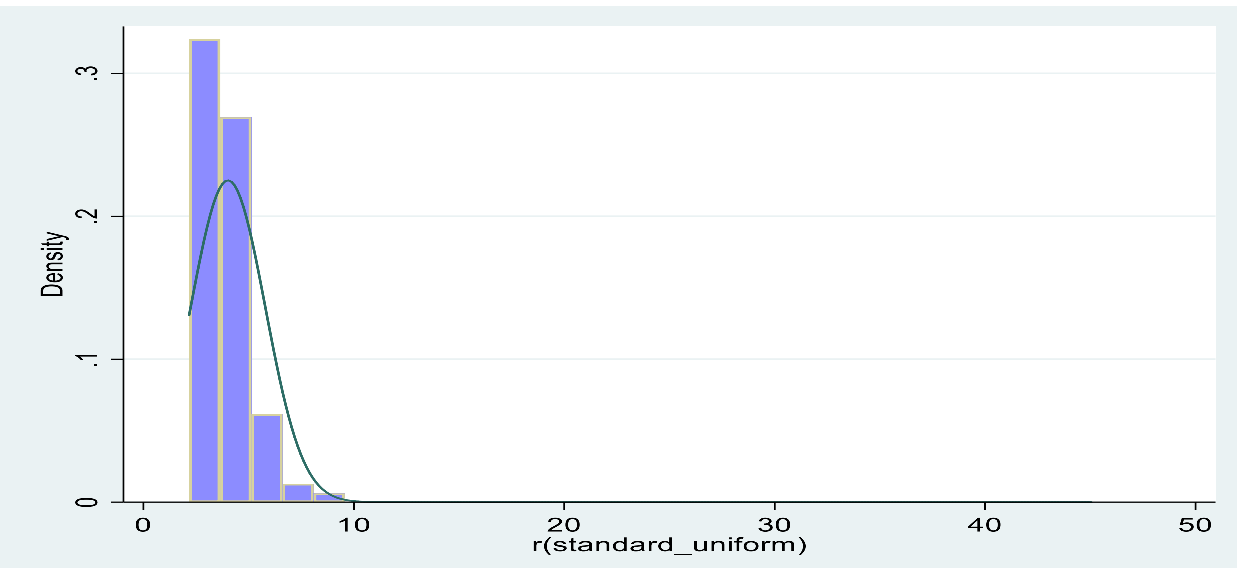
Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2) (joint)	Prob>chi2 (joint)
standard_n~l	1,000	0.0000	0.0000	.	0.0000

Comment

The histogram of the probit estimator of β when ε follows the standard normal distribution is quite symmetrical with slight skewed to the right. The skewness and kurtosis tests for normality can also confirm that the histogram is not normally distributed. Having said that, the estimator can still be unbiased with correct pre-requisite which is the standard normal distribution for the probit model.

(b) The properties of the probit estimator of β when ε follows the standard uniform distribution

Observations	Mean	Standard Deviation	Min	Max
1000	4.0195	1.7720	2.1733	45.0451



Simulation setting: Seed at 703 with 1000 repetitions

Skewness/Kurtosis tests for Normality

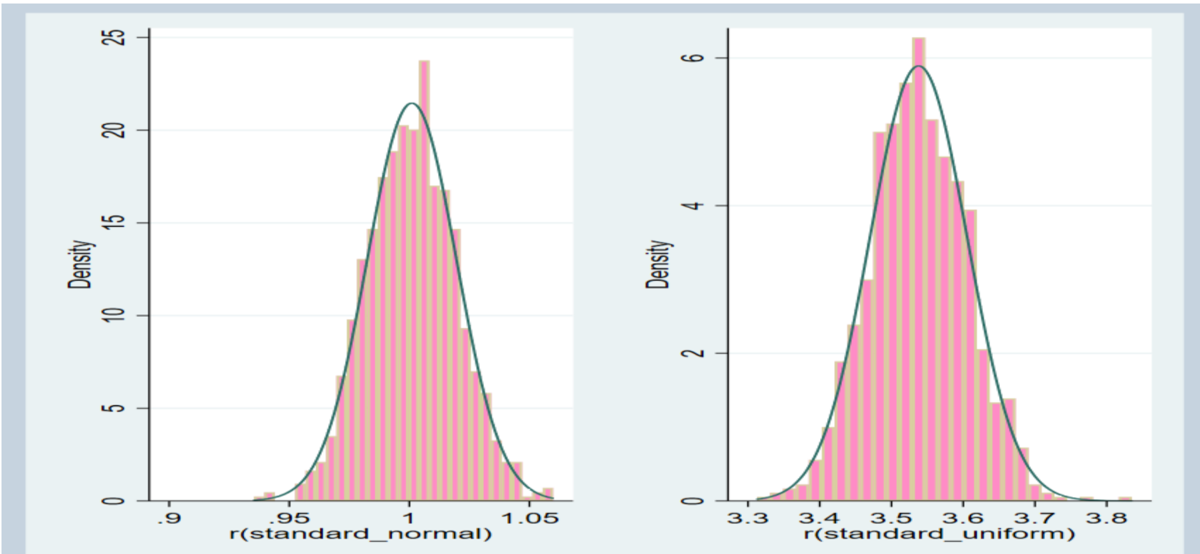
Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2) (joint)	Prob>chi2 (joint)
standard_u~m	1,000	0.0000	0.0000	.	.

Comment

The histogram of the probit estimator of β when ε follows the standard uniform distribution is not symmetrical (with long tail to the right or positive skewness) at all. The skewness and kurtosis tests for normality also confirm that the histogram is not normally distributed. This estimator will be biased with incorrect pre-requisite for the probit model.

(iii) Repeat part (ii) with N = 10,000

	Observations	Mean	Standard Deviation	Min	Max
Standard Normal	1000	1.0010	0.0186	0.9353	1.0598
Standard Uniform	1000	3.5378	0.0677	3.3129	3.8349



(LHS): Probit estimator of β when ε follows the standard normal distribution

(RHS): Probit estimator of β when ε follows the standard uniform distribution

Simulation setting: Seed at 703 with 1000 repetitions

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	joint adj chi2(2)	joint Prob>chi2
standard_n~l	1,000	0.6329	0.8393	0.27	0.8740
standard_u~m	1,000	0.1922	0.1105	4.24	0.1202

Comment

At the first glance, both distributions are much more symmetrical when N = 10000 than when N = 100. As N gets much larger, the distribution of the probit estimator of β when ε follows the standard normal distribution becomes much more symmetric with no apparent skewness and no excess kurtosis. After conducting the skewness and kurtosis tests, we can confirm that it is normally distributed when N gets to 10000. The parameters have converged to true values in which this phenomenon implies consistency in the estimator as the sample size becomes large.

To the contrast, as N gets much larger, the distribution of the probit estimator of β when ε follows the standard uniform distribution is also quite symmetric based on the skewness and kurtosis tests for normality. However, the parameters cannot converged to their true values for incorrect specification; thus, we cannot conclude that the estimator is consistent when it comes to asymptotic properties.

Appendix: STATA code for Q4

```
clear all
capture log close
set more off
log using ECON7320A3.log, replace

program Q4N100, rclass
//standard normal//
	drop _all
	quietly set obs 100
	gen e=rnormal(0,1)
	gen x=rnormal(0,1)
	gen ystar=1*x-e
	gen y=(ystar>=0)
	quietly probit y x
	return scalar standard_normal=_b[x]
//standard uniform//
	drop _all
	quietly set obs 100
	gen e=runiform(0,1)
	gen x=rnormal(0,1)
	gen ystar=1*x-e
	gen y=(ystar>=0)
	quietly probit y x
	return scalar standard_uniform=_b[x]
end
```

```
Q4N100

return list


//Simulation//

simulate standard_normal=r(standard_normal) standard_uniform=r(standard_uniform), seed(703) reps(1000):
Q4N100

//Summary of statistical property//

sum

//Non-normal test//

sktest standard_normal

sktest standard_uniform

//graphing//

histogram standard_normal, normal fcolor(blue*.45)

graph save SND, replace

histogram standard_uniform, normal fcolor(blue*.45)

graph save UND, replace

graph combine SND.gph UND.gph, scheme(economist)


clear all


program Q4N10000, rclass

//standard normal//

        drop _all

        quietly set obs 10000

        gen e=rnormal(0,1)

        gen x=rnormal(0,1)

        gen ystar=1*x-e

        gen y=(ystar>=0)

        quietly probit y x

        return scalar standard_normal=_b[x]

//standard uniform//
```

```
        drop _all
        quietly set obs 10000
        gen e=runiform(0,1)
        gen x=rnormal(0,1)
        gen ystar=1*x-e
        gen y=(ystar>=0)
        quietly probit y x
        return scalar standard_uniform=_b[x]
end

Q4N10000
return list

//Simulation//
simulate standard_normal=r(standard_normal) standard_uniform=r(standard_uniform), seed(703) reps(1000): Q4N10000
//Summary of statistical property//
sum
//Non-normal test//
sktest standard_normal standard_uniform
//Graphing//
histogram standard_normal, normal fcolor(pink*.45)
graph save SND, replace
histogram standard_uniform, normal fcolor(pink*.45)
graph save UND, replace
graph combine UND.gph UND.gph, scheme(economist)
log close
```