ECON 421
Term Project
Iqbal Bamrah
20682484

Task 1

(1) To interpret the estimators for $\beta 1$, $\beta 2$, and $\beta 3$, we would have to look at the independent variables within the LPM. Starting with $\beta 1$, holding all other variables constant, 1 more year of work experience increases the percentage points of labour force participation equalling 1 by $\beta 1$. For $\beta 2$, holding all other variables constant, 1 more year of education increases the percentage points of labour force participation equalling 1 by $\beta 2$. For $\beta 3$, holding all other variables constant, being 1 year older increase the percentage points of labour force participation equalling 1 by $\beta 3$

There would be problems with the interpretation of the estimated model because the relationship between the independent variables expr, educ, and age and the probability of lfprt equalling 1 is almost never linear. The second problem with the estimators is the fitted values of the probability of lfprt given the estimators can be less than 0 and greater than 1, which does not make any sense in regard to probability estimation using the estimators.

(2) LPM: $P(Y_i=1 \mid \{X_j\}_{j=1}^{n}) = X_i^T \beta_i^*$

Probit: $Y_i^* = X_i^T \beta_i^* + U_i$, $\quad U_i \sim N(0,1)$ $\qquad Y_i^* = \begin{cases} 1 & \text{if } Y_i^* > 0 \\ 0 & \text{otherwise} \end{cases}$

$$P(Y_i=1 \mid \{X_j\}_{j=1}^{n}) = G(X_i^T \beta_i^*)$$

$$P(Y_i=1 \mid \{X_j\}_{j=1}^{n}) = P(Y^* > 0) = P(X_i^T \beta_i^* + U_i > 0) = P(U_i < X_i^T \beta_i^*) = \Phi(X_i^T \beta_i^*)$$

$$G(X_i^T \beta_i^*) = \Phi(X_i^T \beta_i^*) \leftarrow \text{Probit regression model}$$

$i = 1, 2, 3$

(3) $\begin{cases} P(Y_i=1 \mid \{X_j\}_{j=1}^{n}) = G(X_i^T \beta_i^*) \\ P(Y_i=0 \mid \{X_j\}_{j=1}^{n}) = 1 - G(X_i^T \beta_i^*) \end{cases}$

$$f(Y_i \mid \{X_j\}_{j=1}^{n}; \beta_i^*) = G(X_i^T \beta_i^*)^{Y_i} [1 - G(X_i^T \beta_i^*)]^{1-Y_i}$$

$\hat{\beta}_{iMLE} := \operatorname{argmax} \bar{l}(\beta_i)$, where

$$\bar{l}(\beta_i) = \frac{1}{n} \sum_{i=1}^{n} \log(f(Y_i \mid \{X_j\}_{j=1}^{n}; \beta)) = \frac{1}{n} \sum_{i=1}^{n} Y_i \log(G(X_i^T \beta)) + \frac{1}{n} \sum_{i=1}^{n} (1-Y_i) \log(1 - G(X_i^T \beta))$$

F.O.C

$$0 = \frac{\partial \mathcal{L}(\beta_i)}{\partial \beta_i}\Big|_{\beta_i = \hat{\beta}_{iMLE}} = \frac{1}{n}\sum_{i=1}^{n} \frac{Y_i X_i G'(X_i^T \hat{\beta}_{iMLE})}{G(X_i^T \hat{\beta}_{iMLE})} + \frac{1}{n}\sum_{i=1}^{n} \frac{(1-Y_i) X_i G'(X_i^T \hat{\beta}_{iMLE})(-1)}{1 - G(X_i^T \hat{\beta}_{iMLE})}$$

where $\beta_i = \beta_1, \beta_2, \beta_3$

$X_i = X_1, X_2, X_3$

FOC of ML estimators

$$\sqrt{n}\,(\hat{\beta}_{iMLE} - \beta_i^*) \xrightarrow{d} N\left(0, [I(\beta_i^*)]^{-1}\right), \text{ where } I(\beta_i^*) = -E\left[\frac{\partial^2}{\partial \beta_i \partial \beta_i^T} \log\left(f(Y_i \mid \{X_j\}_{j=1}^n ; \beta_i^*)\right)\right]$$

asymptotic distribution

$$\frac{\partial P(Y_i = 1 \mid \{X_j\}_{j=1}^n)}{\partial X_i} = \Phi'(X_i^T \beta_i^*)\beta_i^*$$

Each unit increase in expr will increase the probability of lfprt=1 by $\Phi'(\text{expr}\,\hat{\beta}_{1,MLE})\hat{\beta}_{1,MLE}$ which explains the marginal effects of expr. The marginal effects of educ show that each unit increase in educ will increase the probability of lfprt=1 by $\Phi'(\text{educ}\,\hat{\beta}_{2MLE})\hat{\beta}_{2MLE}$. The marginal effects of age show that each unit increase in age will increase the probability of lfprt=1 by $\Phi'(\text{age}\,\hat{\beta}_{3MLE})\hat{\beta}_{3MLE}$. I expect the marginal effects of each independent variable to be positive because lfprt is positively effected by expr, educ, and age logically and based on that, I think the marginal effects of increase in a unit of experience, education, or age would positively effect the labour force participation.

Task 1

4) LPM

```
> LPM <- glm(lfprt ~ expr + educ + age, data = data)
> summary(LPM)

Call:
glm(formula = lfprt ~ expr + educ + age, data = data)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-0.49883  -0.26061  -0.02587   0.24996   0.83733

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.031691   0.047798   0.663    0.507
expr         0.078169   0.003163  24.718   <2e-16 ***
educ         0.050594   0.003256  15.538   <2e-16 ***
age         -0.022923   0.001063 -21.560   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.08929386)

    Null deviance: 202.039  on 999  degrees of freedom
Residual deviance:  88.937  on 996  degrees of freedom
AIC: 428.05

Number of Fisher Scoring iterations: 2
```

Here we can see the coefficients for the independent variables from the LPM model. The independent variables are statistically significant at the 5% p-value level. They all have a p-value far lower than 5% and so we can see that expr, educ, and age all independently have a relationship with the dependent variable lfprt based on the significance indicated from the p-values. Each variable is independently related to lfprt and the relationship is very significant based on how low the p-values are.

Probit Model

```
> Probit <- glm(formula = lfprt ~ expr + educ + age, data = data, family = binomial(link = 'probi
t'))
Warning messages:
1: glm.fit: algorithm did not converge
2: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

With the Probit model, due to too many iterations being created, the 1st error occurred and due to the probabilities being too close to 0 or 1, the regression did not give the proper output based on the model data. To adjust to the error, I created a Probit regression model without the independent variable, age, and the error disappeared and created the Probit model.

```
> Probit <- glm(formula = lfprt ~ expr + educ, data = data, family = binomial(link = 'probit'))
> summary(Probit)

Call:
glm(formula = lfprt ~ expr + educ, family = binomial(link = "probit"),
    data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3262  -0.5959  -0.2222   0.3719   2.3199

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -5.51549    0.32416  -17.01   <2e-16 ***
expr         0.37401    0.02495   14.99   <2e-16 ***
educ         0.23925    0.02113   11.32   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1187.79  on 999  degrees of freedom
Residual deviance:  737.12  on 997  degrees of freedom
AIC: 743.12

Number of Fisher Scoring iterations: 6
```

Here is the Probit regression model with the independent variables expr, and educ and we can see that both independent variables are statistically significant at the 5% p-value level as their p-values are far below 5%. This indicates that both expr, and educ are statistically significant and we can reject the null hypothesis that expr, and educ independently do not have a relationship with the dependent variable lfprt.

5)

```
> ProbitMar <- margins(Probit)
> summary(ProbitMar)
 factor    AME     SE       z      p  lower  upper
   educ 0.0488 0.0034 14.4123 0.0000 0.0422 0.0555
   expr 0.0764 0.0030 25.5167 0.0000 0.0705 0.0822
```

The marginal effects of expr, educ can be seen from AME. The marginal effects for educ is 0.0488 with a standard error of 0.0034 which is positive as I mentioned previously in (3) and this is due to the fact that educ has a positive relationship with lfprt and it can be seen from the Probit regression model. Also, we can see that educ has a p-value of 0.0000 which means it is statistically significant and we can reject the null hypothesis that the marginal effects of educ have no relationship with lfprt. Looking at expr, we can see the marginal effects are 0.0764 with a standard error of 0.0030 which indicate that expr also has a positive marginal effect on lfprt and this relationship between expr and lfprt is also statistically significant based on the p-value of 0.0000 we have for expr. This shows us that both expr, and educ are both statistically significant based on the marginal effects each independent variable has on lfprt or the labour force participation within our model.

6)
```
> fitted <- fitted.values(LPM)
> summary(fitted)
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-0.83733  0.04823  0.29046  0.28100  0.51597  1.28868
> fitted1 <- fitted.values(Probit)
> summary(fitted1)
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
0.0000026 0.0354838 0.1626904 0.2807827 0.5015170 0.9986317
```

From the fitted values of the LPM and the Probit model, fitted and fitted1 respectively, we can see that the values for the LPM extend to less than 0 and greater than 1, whereas the Probit model showcases the values between 0 and 1 which makes sense in this case as we are looking at the probability of labour force participation.

```
> stargazer(LPM, Probit, type = "text")

=================================================
                      Dependent variable:
                  -------------------------------
                              lfprt
                      normal          probit
                        (1)             (2)
-------------------------------------------------
expr                 0.078***        0.374***
                     (0.003)         (0.025)

educ                 0.051***        0.239***
                     (0.003)         (0.021)

age                 -0.023***
                     (0.001)

Constant              0.032         -5.515***
                     (0.048)         (0.324)

-------------------------------------------------
Observations          1,000           1,000
Log Likelihood      -210.023        -368.559
Akaike Inf. Crit.    428.047         743.118
=================================================
Note:               *p<0.1; **p<0.05; ***p<0.01
```

Looking at the AIC, we can see that based on this information the LPM is a better fit model for our data based on the AIC being lower. The AIC being lower indicates the model is better fit for the data and so based on this comparison of the two models, it is clear that based on this, the LPM is a better model for our data. But, the problem with this is that within our Probit model, the independent variable age could not be used and so I created a LPM without the independent variable age to match the Probit model and based on this, the AIC is higher for the LPM than the Probit model and based on that, I think the Probit model is a better indicator for prediction performance because with the same variables the Probit model has a lower AIC and this indicates the Probit model is better fit for our data.

```
> LPM1 <- glm(lfprt ~ expr + educ, data = data)
> stargazer(LPM1, Probit, type = "text")

=================================================
                      Dependent variable:
                  -------------------------------
                              lfprt
                      normal           probit
                        (1)              (2)
-------------------------------------------------
expr                 0.077***         0.374***
                     (0.004)          (0.025)

educ                 0.050***         0.239***
                     (0.004)          (0.021)

Constant            -0.658***        -5.515***
                     (0.043)          (0.324)

-------------------------------------------------
Observations          1,000            1,000
Log Likelihood       -401.537         -368.559
Akaike Inf. Crit.     809.074          743.118
=================================================
Note:                 *p<0.1; **p<0.05; ***p<0.01
```

As we can see from the stargazer, the AIC from the Probit model is lower than that of the LPM and so we can conclude that the Probit model is a better fit model for the data we have. Based on the information that the Probit model is better fit with the exclusion of age from both models, I think it is safe to assume that with age in both models, the Probit model would still be more fit for the data if the error did not occur.

Error within the Data

```
> summary(Probit1)

Call:
glm(formula = lfprt ~ expr + educ + age, family = binomial(link = "probit"),
    data = data, control = list(maxit = 50))

Deviance Residuals:
       Min          1Q      Median          3Q         Max
-3.722e-06  -2.110e-08  -2.110e-08   2.110e-08   3.105e-06

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)   -197.97  177096.46  -0.001    0.999
expr            40.95   36183.56   0.001    0.999
educ            27.31   24232.53   0.001    0.999
age            -13.65   12079.07  -0.001    0.999

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1.1878e+03  on 999  degrees of freedom
Residual deviance: 4.3136e-10  on 996  degrees of freedom
AIC: 8

Number of Fisher Scoring iterations: 33

> ProbitMar1 <- margins(Probit1)
> summary(ProbitMar1)
 factor     AME     SE       z      p   lower  upper
    age -0.0000 0.0000 -0.0000 1.0000 -0.0000 0.0000
   educ  0.0000 0.0000  0.0000 1.0000 -0.0001 0.0001
   expr  0.0000 0.0000  0.0000 1.0000 -0.0001 0.0001
```

From this data, this includes all the independent variables for the Probit model and from the error we can see that the data is very wrong as none of the variables are significant at any p-value. Also, form the marginal effects data, we can see that the marginal effects are non-existent for each variable and the p-value of each variable is 1 which indicates there is no relationship between these variables and the labour force participation which does not make any sense when looking at the other models created with the same data.

Task 2

$$Y_i^* = X_i^T\beta_0 + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2)$$

$$Y_i = \begin{cases} B_L & \text{if } Y_i^* \le B_L \\ Y_i^* & \text{if } B_L < Y_i^* < B_U \\ B_U & \text{if } Y_i^* \ge B_U \end{cases}$$

(1) Left censored
$$Y_i = \begin{cases} B_L & \text{if } Y_i^* \le B_L \\ Y_i^* & \text{if } Y_i^* > B_L \end{cases}$$

Right censored
$$Y_i = \begin{cases} B_U & \text{if } Y_i^* \ge B_U \\ Y_i^* & \text{if } Y_i^* < B_U \end{cases}$$

Interval censored
$$Y_i = \begin{cases} B_L & \text{if } Y_i^* \le B_L \\ Y_i^* & \text{if } B_L < Y_i^* < B_U \\ B_U & \text{if } Y_i^* \ge B_U \end{cases}$$

Left censored
$$Y_i = \begin{cases} B_L & P\left(\epsilon_i \le B_L - X_i^T\beta_0^* \,\middle|\, \{X_j\}_{j=1}^n\right) \ (*) \\ X_i^T\beta_0^* + \epsilon_i & \text{pdf}\left(Y_i \,\middle|\, \{X_j\}_{j=1}^n ; \beta_0^*, \sigma^{*2}\right) \ (\#) \end{cases}$$

$$(*) \quad P\left(\epsilon_i \le B_L - X_i^T\beta_0^* \,\middle|\, \{X_j\}_{j=1}^n\right) = P\left(\frac{\epsilon_i}{\sigma^*} \le \frac{B_L - X_i^T\beta_0^*}{\sigma^*} \,\middle|\, \{X_j\}_{j=1}^n\right) = \Phi\left(\frac{B_L - X_i^T\beta_0^*}{\sigma^*}\right) - 1$$

$$(\#) \quad \text{pdf}\left(Y_i \,\middle|\, \{X_j\}_{j=1}^n ; \beta_0^*, \sigma^{*2}\right) = \frac{1}{\sqrt{2\pi\sigma^{*2}}} e^{-\frac{1}{2}\left(\frac{Y_i - X_i^T\beta_0^*}{\sigma^*}\right)^2}$$
$$= \frac{1}{\sigma^*}\left[\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{Y_i - X_i^T\beta_0^*}{\sigma^*}\right)^2}\right.$$
$$= \frac{1}{\sigma^*}\phi\left(\frac{Y_i - X_i^T\beta_0^*}{\sigma^*}\right)$$

$$Y_i = \begin{cases} B_L & \Phi\left(\frac{B_L - X_i^T\beta_0^*}{\sigma^*}\right) - 1 \\ X_i^T\beta_0^* + \epsilon_i & \frac{1}{\sigma^*}\phi\left(\frac{Y_i - X_i^T\beta_0^*}{\sigma^*}\right) \end{cases}$$

$$\begin{bmatrix} \hat{\beta}_{0MLE} \\ \hat{\sigma}^2_{MLE} \end{bmatrix} := \underset{(\beta_0, \sigma^2)}{\operatorname{argmax}} \bar{l}(\beta_0, \sigma^2)$$

$$\bar{l}(\beta, \sigma^2) = \frac{1}{n}\sum_{i=1}^n \log\left(f\left(Y_i \,\middle|\, \{X_j\}_{j=1}^n ; \beta_0, \sigma^2\right)\right)$$
$$= \frac{1}{n}\sum_{i=1}^n \mathbb{1}(Y_i = B_L)\log\left(\Phi\left(\frac{B_L - X_i^T\beta_0}{\sigma}\right) - 1\right) + \frac{1}{n}\sum_{i=1}^n \mathbb{1}(Y_i \ne B_L)\log\left(\frac{1}{\sigma}\phi\left(\frac{Y_i - X_i^T\beta_0}{\sigma}\right)\right)$$

FOC
$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \bar{l}(\hat{\beta}_{0MLE}, \hat{\sigma}^2_{MLE})}{\partial \beta_0} \\ \frac{\partial \bar{l}(\hat{\beta}_{0MLE}, \hat{\sigma}^2_{MLE})}{\partial \sigma^2} \end{bmatrix} \longleftarrow \text{ML estimators for } \beta \text{ and } \sigma^2$$

Right censored
$$Y_i = \begin{cases} B_U & P\left(\epsilon_i \ge B_U - X_i^T\beta_0^* \,\middle|\, \{X_j\}_{j=1}^n\right) \ (\bullet) \\ X_i^T\beta_0^* + \epsilon_i & \text{pdf}\left(Y_i \,\middle|\, \{X_j\}_{j=1}^n ; \beta, \sigma^2\right) \ (\#) \end{cases}$$

$$(\bullet) \quad P\left(\epsilon_i \ge B_U - X_i^T\beta_0^* \,\middle|\, \{X_j\}_{j=1}^n\right) = P\left(\frac{\epsilon_i}{\sigma^*} \ge \frac{B_U - X_i^T\beta_0^*}{\sigma^*} \,\middle|\, \{X_j\}_{j=1}^n\right) = 1 - \Phi\left(\frac{B_U - X_i^T\beta_0^*}{\sigma^*}\right)$$

$$(\#) = \frac{1}{\sigma^*}\phi\left(\frac{Y_i - X_i^T\beta_0^*}{\sigma^*}\right)$$

$$y_i = \begin{cases} B_U & 1-\Phi\left(\frac{B_U - x_i^T\beta_0^*}{\sigma^*}\right) \\ x_i^T\beta_0^* + \epsilon_i & \frac{1}{\sigma^*}\phi\left(\frac{y_i - x_i^T\beta_0^*}{\sigma^*}\right) \end{cases}$$

$$\begin{bmatrix} \hat\beta_{MLE} \\ \hat\sigma^2_{MLE} \end{bmatrix} : \underset{(\beta,\sigma^2)}{\operatorname{argmax}} \bar{l}(\beta,\sigma^2) \qquad \bar{l}(\beta,\sigma^2) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}(y_i = B_U)\log\left(1-\Phi\left(\frac{B_U - x_i^T\beta_0}{\sigma}\right)\right) + \frac{1}{n}\sum_{i=1}^n \mathbb{1}(y_i \neq B_U)\log\left(\frac{1}{\sigma}\phi\left(\frac{y_i - x_i^T\beta_0}{\sigma}\right)\right)$$

FOC $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \bar{l}(\hat\beta_{MLE}, \hat\sigma^2_{MLE})}{\partial \beta_0} \\ \frac{\partial \bar{l}(\hat\beta_{MLE}, \hat\sigma^2_{MLE})}{\partial \sigma^2} \end{bmatrix}$ ← ML estimators for β and σ^2

Interval censored

$$y_i = \begin{cases} B_L & \text{if } y_i^* \le B_L \\ y_i & \text{if } B_L < y_i^* < B_U \\ B_U & \text{if } y_i^* \ge B_U \end{cases}$$

(*)
(∘)
(#)

$$\begin{bmatrix} \hat\beta_{MLE} \\ \hat\sigma^2_{MLE} \end{bmatrix} := \underset{(\beta,\sigma^2)}{\operatorname{argmax}} \bar{l}(\beta,\sigma^2)$$

$$\bar{l}(\beta,\sigma^2) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}(y_i = B_U)\log\left(1-\Phi\left(\frac{B_U - x_i^T\beta_0}{\sigma}\right)\right) + \frac{1}{n}\sum_{i=1}^n \mathbb{1}(y_i = B_L)\log\left(\Phi\left(\frac{B_L - x_i^T\beta_0}{\sigma}\right)\right) + \frac{1}{n}\sum_{i=1}^n \mathbb{1}(B_U \neq y_i \neq B_L)\log\left(\frac{1}{\sigma}\phi\left(\frac{y_i - x_i^T\beta_0}{\sigma}\right)\right)$$

FOC $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \bar{l}(\hat\beta_{MLE}, \hat\sigma^2_{MLE})}{\partial \beta} \\ \frac{\partial \bar{l}(\hat\beta_{MLE}, \hat\sigma^2_{MLE})}{\partial \sigma^2} \end{bmatrix}$ ← ML estimators for β and σ^2

(2) $\theta^* = \begin{bmatrix} \beta_0^* \\ \sigma^{*2} \end{bmatrix}$, $\hat\theta_{MLE} = \begin{bmatrix} \hat\beta_{MLE} \\ \hat\sigma^2_{MLE} \end{bmatrix}$, $\theta = \begin{bmatrix} \beta_0 \\ \sigma^2 \end{bmatrix}$

$\sqrt{n}\left(\hat\theta_{MLE} - \theta^*\right) \xrightarrow{d} N\left(0, [I(\theta^*)]^{-1}\right)$ ← asymptotic distribution

$$[I(\theta^*)] = -E\left[\frac{\partial^2}{\partial\theta\partial\theta^T}\log\left(f\left(y_i \mid \{x_j\}_{j=1}^n; \theta^*\right)\right)\right]$$

asymptotic variance, $[I(\theta^*)]^{-1}$ estimated through

$$\widehat{[I(\theta^*)]^{-1}} = \left(-\frac{\partial^2 \bar{l}(\theta)}{\partial\theta\partial\theta^T}\Big|_{\theta = \hat\theta_{MLE}}\right)^{-1}$$ ← asymptotic variance estimator for left, right, and interval censored response ML estimators.

Task 2

3)

Here we can see the summary of the 1,000 estimators as well as the asymptotic variance test of the 1,000 estimators.

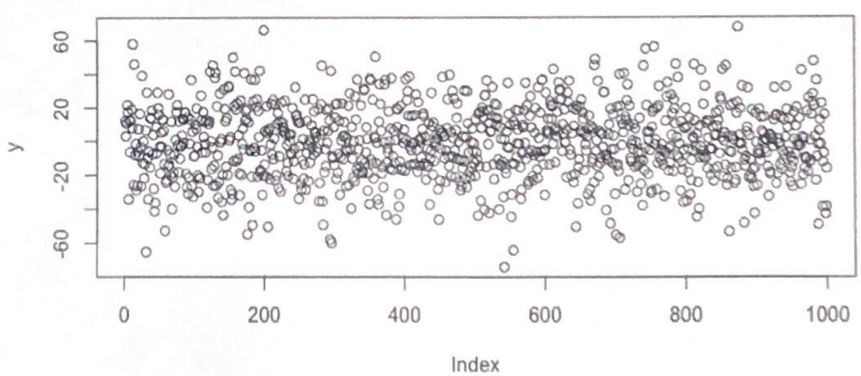
```
> summary(y)
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-73.8349 -14.5929  -1.1823  -0.7712  13.3218  67.9542
```

4)

```
> asymp.test(y, parameter = "var")

        One-sample asymptotic variance test

data:  y
statistic = 21.774, p-value < 2.2e-16
alternative hypothesis: true variance is not equal to 0
95 percent confidence interval:
 404.1087 484.0566
sample estimates:
variance
444.0827
```

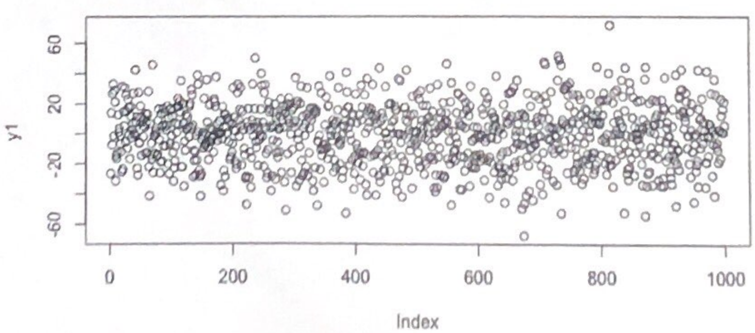


The model data is not consistent with the results from (2) because we can see that from the ML estimator created in (1) which was used to create the asymptotic distribution and asymptotic variance estimator in (2), the data from R shows us the asymptotic variance is 444.0827 which is the approximation of the variance within the data and this asymptotic variance created in R is much higher than what would be estimated using the estimator created in (2). We also have the plot of the 1,000 estimators we can see the variability within the data created for the simulations

5)
```
> cbind(asymp.test(y, parameter = "var"),asymp.test(y1, parameter = "var"))
            [,1]                                 [,2]
statistic   21.77382                             23.41772
p.value     0                                    0
conf.int    Numeric,2                            Numeric,2
estimate    444.0827                             402.9178
null.value  0                                    0
alternative "two.sided"                          "two.sided"
method      "One-sample asymptotic variance test" "One-sample asymptotic variance test"
data.name   "y"                                  "y1"
```

Here we have the comparison of the error term that follows a standard normal distribution "y", and the error term that follow a normal distribution "y1". From this data, we can see that the asymptotic variance estimate is lower when the error term follows a normal distribution to a standard normal distribution which makes sense as the variance in data from a normal distribution is lower than that of the standard normal distribution created. From this we can see that both variances are significant within our data based on the p-value of 0 and that the estimate indicates the normal distribution "y1" has a stronger indication of significance based on the asymptotic variance estimate being lower and the statistic being higher. The difference is not too big, but we can definitely see the difference within the change in the error term.

```
> summary(y1)
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-67.6071 -14.7897   0.2146  -0.5494  12.9398  71.9915
```



Here we have both the normal distribution summary statistics and the distribution of the estimators which show very similar data, with the small difference in the range of the data, which is all dependent on the distribution simulated within the data. With the two sets of data, I learned that the normal distribution shows slighlt more condensed data based on the asymptotic variance estimates above and how the standard error term displays a wider variability within the data in comparison to the normal distribution error term.