LPM Probit Logit

Taken from `https://www.econometrics-with-r.org/`; created via `rawr::rawr(...)`

Eyayaw Beze

July 14, 2020

Linear Probability Model (LPM)

```
# inspect the data
head(HMDA)
```

deny	pirat	hirat	lvrat	chist	mhist	phist	unemp	selfemp	insurance	condomin	afam	single	hschool
no	0.221	0.221	0.80000	5	2	no	3.9	no	no	no	no	no	yes
no	0.265	0.265	0.92188	2	2	no	3.2	no	no	no	no	yes	yes
no	0.372	0.248	0.92040	1	2	no	3.2	no	no	no	no	no	yes
no	0.320	0.250	0.86047	1	2	no	4.3	no	no	no	no	no	yes
no	0.360	0.350	0.60000	1	1	no	3.2	no	no	no	no	no	yes
no	0.240	0.170	0.51053	1	1	no	3.9	no	no	no	no	no	yes

```
summary(HMDA)
```

```
##   deny          pirat            hirat            lvrat          chist    mhist
##  no :2095   Min.   :0.000   Min.   :0.000   Min.   :0.020   1:1353   1: 747
##  yes: 285   1st Qu.:0.280   1st Qu.:0.214   1st Qu.:0.653   2: 441   2:1571
##             Median :0.330   Median :0.260   Median :0.780   3: 126   3:  41
##  phist          unemp       selfemp    insurance  condomin    afam
##  no :2205   Min.   : 1.80   no :2103   no :2332   no :1694   no :2041
##  yes: 175   1st Qu.: 3.10   yes: 277   yes:  48   yes: 686   yes: 339
##             Median : 3.20
##  single     hschool
##  no :1444   no :  39
##  yes: 936   yes:2341
##
##  [ reached getOption("max.print") -- omitted 3 rows ]
```

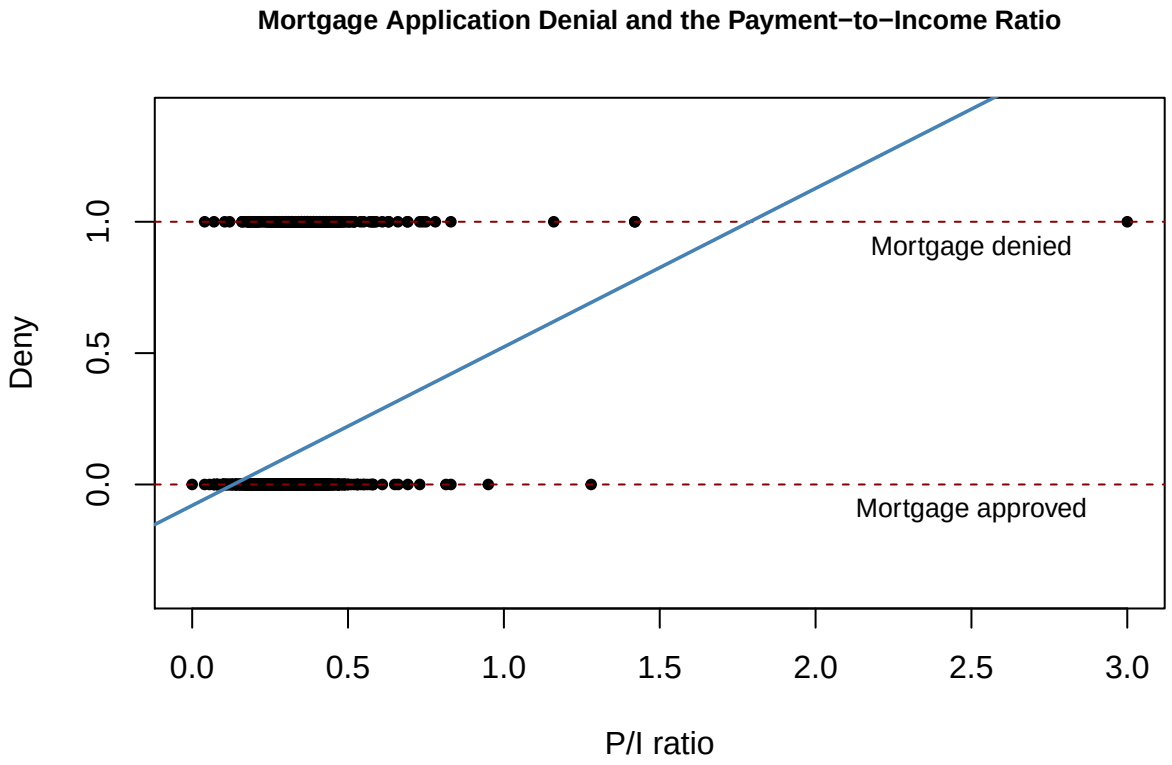
```
as.numeric(HMDA$deny)
```

```
##  [1] 1 1 1 1 1 1 1 1 2 1 1 1 2 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## [39] 1 1 1 1 2 2 1 1 2 1 2 1
##  [ reached getOption("max.print") -- omitted 2330 entries ]
```

```r
# convert 'deny' to numeric
HMDA$deny <- as.numeric(HMDA$deny) - 1

# estimate a simple linear probability model
denymod1 <- lm(deny ~ pirat, data = HMDA)
denymod1
```

```
##
## Call:
## lm(formula = deny ~ pirat, data = HMDA)
##
## Coefficients:
## (Intercept)        pirat
##     -0.0799       0.6035
```

```r
# plot the data
plot(x = HMDA$pirat,
     y = HMDA$deny,
     main = "Mortgage Application Denial and the Payment-to-Income Ratio",
     xlab = "P/I ratio",
     ylab = "Deny",
     pch = 20,
     ylim = c(-0.4, 1.4),
     cex.main = 0.8)

# add horizontal dashed lines and text
abline(h = 1, lty = 2, col = "darkred")
abline(h = 0, lty = 2, col = "darkred")
text(2.5, 0.9, cex = 0.8, "Mortgage denied")
text(2.5, -0.1, cex= 0.8, "Mortgage approved")

# add the estimated regression line
abline(denymod1,
       lwd = 1.8,
       col = "steelblue")
```




```r
# print robust coefficient summary
coeftest(denymod1, vcov. = vcovHC, type = "HC1")
```

```
##
## t test of coefficients:
##
##             Estimate Std. Error t value    Pr(>|t|)
## (Intercept)  -0.0799     0.0320   -2.50       0.012 *
## pirat         0.6035     0.0985    6.13 0.000000001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# rename the variable 'afam' for consistency
colnames(HMDA)[colnames(HMDA) == "afam"] <- "black"

# estimate the model
denymod2 <- lm(deny ~ pirat + black, data = HMDA)
coeftest(denymod2, vcov. = vcovHC)
```

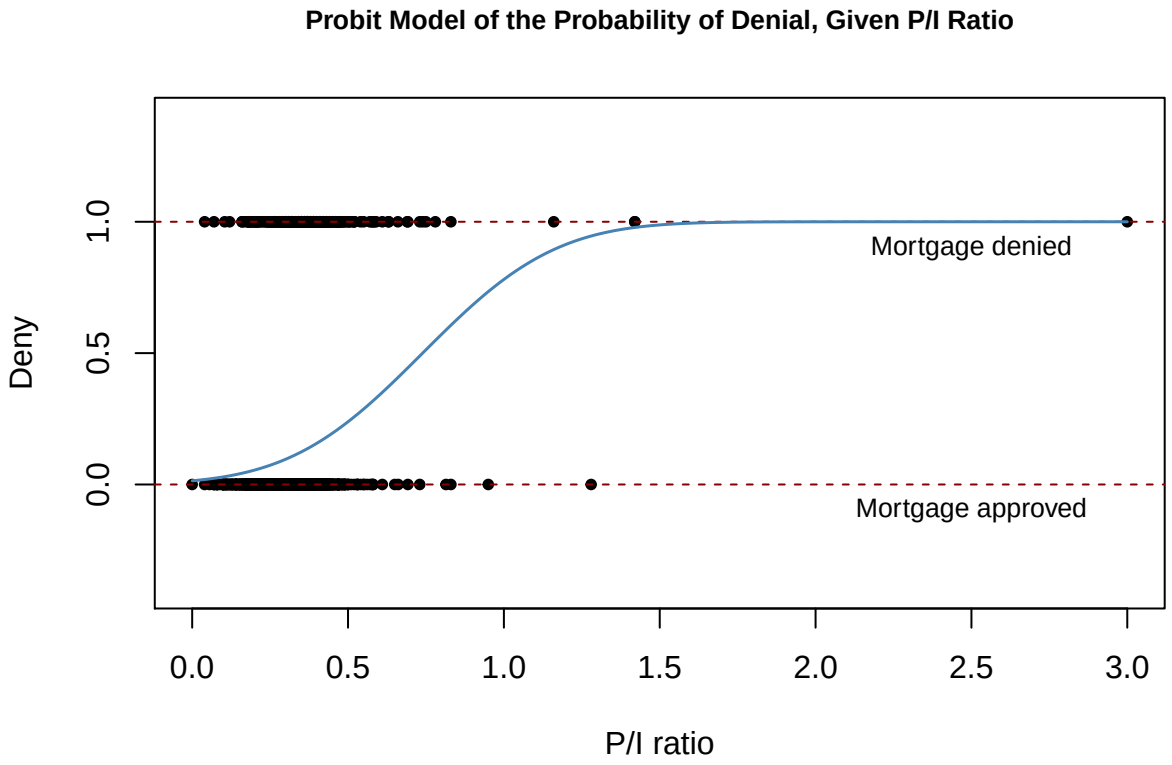
```
##
## t test of coefficients:
##
##             Estimate Std. Error t value        Pr(>|t|)
## (Intercept)  -0.0905     0.0334   -2.71          0.0068 **
## pirat         0.5592     0.1037    5.39 0.0000000757501 ***
## blackyes      0.1774     0.0251    7.08 0.0000000000019 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Probit Model

```r
# estimate the simple probit model
denyprobit <- glm(deny ~ pirat,
                  family = binomial(link = "probit"),
                  data = HMDA)

coeftest(denyprobit, vcov. = vcovHC, type = "HC1")
```

```
##
## z test of coefficients:
##
##             Estimate Std. Error z value    Pr(>|z|)
## (Intercept)   -2.194      0.189  -11.61     < 2e-16 ***
## pirat          2.968      0.537    5.53 0.000000033 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# plot data
plot(x = HMDA$pirat,
     y = HMDA$deny,
     main = "Probit Model of the Probability of Denial, Given P/I Ratio",
     xlab = "P/I ratio",
     ylab = "Deny",
     pch = 20,
     ylim = c(-0.4, 1.4),
     cex.main = 0.85)

# add horizontal dashed lines and text
abline(h = 1, lty = 2, col = "darkred")
abline(h = 0, lty = 2, col = "darkred")
text(2.5, 0.9, cex = 0.8, "Mortgage denied")
text(2.5, -0.1, cex= 0.8, "Mortgage approved")

# add estimated regression line
x <- seq(0, 3, 0.01)
y <- predict(denyprobit, list(pirat = x), type = "response")

lines(x, y, lwd = 1.5, col = "steelblue")
```




```
# 1. compute predictions for P/I ratio = 0.3, 0.4
predictions <- predict(denyprobit,
                       newdata = data.frame("pirat" = c(0.3, 0.4)),
                       type = "response")

# 2. Compute difference in probabilities
diff(predictions)
```

```
##        2
## 0.060814
```

```
denyprobit2 <- glm(deny ~ pirat + black,
                   family = binomial(link = "probit"),
                   data = HMDA)

coeftest(denyprobit2, vcov. = vcovHC, type = "HC1")
```

```
##
## z test of coefficients:
##
##             Estimate Std. Error z value    Pr(>|z|)
## (Intercept)  -2.2588     0.1766  -12.79     < 2e-16 ***
## pirat         2.7418     0.4977    5.51 0.000000036 ***
## blackyes      0.7082     0.0831    8.52     < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# 1. compute predictions for P/I ratio = 0.3
predictions <- predict(denyprobit2,
                       newdata = data.frame("black" = c("no", "yes"),
                                            "pirat" = c(0.3, 0.3)),
                       type = "response")

# 2. compute difference in probabilities
diff(predictions)
```

```
##       2
## 0.15781
```

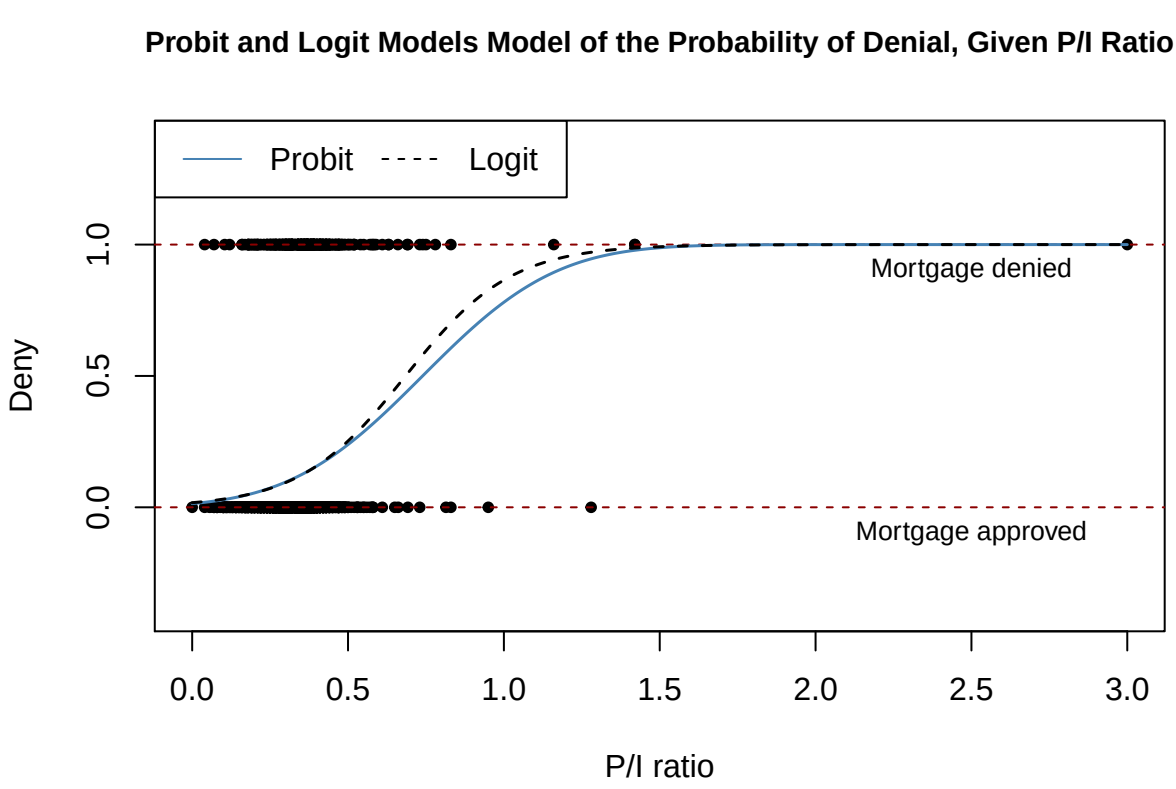
Logit Model

```r
denylogit <- glm(deny ~ pirat,
                 family = binomial(link = "logit"),
                 data = HMDA)

coeftest(denylogit, vcov. = vcovHC, type = "HC1")
```

```
##
## z test of coefficients:
##
##             Estimate Std. Error z value    Pr(>|z|)
## (Intercept)   -4.028      0.359  -11.22     < 2e-16 ***
## pirat          5.884      1.000    5.88 0.000000004 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# plot data
plot(x = HMDA$pirat,
     y = HMDA$deny,
     main = "Probit and Logit Models Model of the Probability of Denial, Given P/I Ratio",
     xlab = "P/I ratio",
     ylab = "Deny",
     pch = 20,
     ylim = c(-0.4, 1.4),
     cex.main = 0.9)

# add horizontal dashed lines and text
abline(h = 1, lty = 2, col = "darkred")
abline(h = 0, lty = 2, col = "darkred")
text(2.5, 0.9, cex = 0.8, "Mortgage denied")
text(2.5, -0.1, cex= 0.8, "Mortgage approved")

# add estimated regression line of Probit and Logit models
x <- seq(0, 3, 0.01)
y_probit <- predict(denyprobit, list(pirat = x), type = "response")
y_logit <- predict(denylogit, list(pirat = x), type = "response")
```

```
lines(x, y_probit, lwd = 1.5, col = "steelblue")
lines(x, y_logit, lwd = 1.5, col = "black", lty = 2)

# add a legend
legend("topleft",
       horiz = TRUE,
       legend = c("Probit", "Logit"),
       col = c("steelblue", "black"),
       lty = c(1, 2))
```




```
# estimate a Logit regression with multiple regressors
denylogit2 <- glm(deny ~ pirat + black,
                  family = binomial(link = "logit"),
                  data = HMDA)

coeftest(denylogit2, vcov. = vcovHC, type = "HC1")
```

```
##
## z test of coefficients:
##
##             Estimate Std. Error z value    Pr(>|z|)
## (Intercept)   -4.126      0.346  -11.92     < 2e-16 ***
## pirat          5.370      0.964    5.57 0.000000025 ***
## blackyes       1.273      0.146    8.71     < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# 1. compute predictions for P/I ratio = 0.3
predictions <- predict(denylogit2,
                       newdata = data.frame("black" = c("no", "yes"),
                                            "pirat" = c(0.3, 0.3)),
                       type = "response")

predictions
```

```
##        1        2
## 0.074851 0.224146
```

```r
# 2. Compute difference in probabilities
diff(predictions)
```

```
##       2
## 0.14929
```