Econometrics
Lab Hour – Session 4

Agustín Bénétrix
benetria@tcd.ie
Office hour:
Wednesday 4-5
Room 3021

Martin Schmitz
schmitzm@tcd.ie
Office hour:
Thursday 5-6
Room 3021

Question (i)

(i) The variable *favwin* is a binary variable if the team favored by the Las Vegas point spread wins. A linear probability model to estimate the probability that the favored team wins is

$$P(favwin = 1 | spread) = \beta_0 + \beta_1 spread.$$

Explain why, if the spread incorporates all relevant information, we expect $\beta_0 = .5$.

Answer (i)

- If *spread* is zero, there is no favorite, and the probability that the team we (arbitrarily) label the favorite should have a 50% chance of winning

Question (ii)

(ii) Estimate the model from part (i) by OLS. Test H_0: $\beta_0 = .5$ against a two-sided alternative. Use both the usual and heteroskedasticity-robust standard errors.

Question (ii)

```
                        Ordinary Least Squares Estimation
*******************************************************************************
 Dependent variable is FAVWIN
 553 observations used for estimation from    1 to   553
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 C                           .57695           .028235             20.4342[.000]
 SPREAD                     .019366           .0023386             8.2806[.000]
*******************************************************************************
 R-Squared                    .11067   R-Bar-Squared                    .10906
 S.E. of Regression           .40168   F-stat.    F(  1, 551)   68.5691[.000]
 Mean of Dependent Variable   .76311   S.D. of Dependent Variable       .42556
 Residual Sum of Squares     88.9038   Equation Log-likelihood       -279.2855
 Akaike Info. Criterion    -281.2855   Schwarz Bayesian Criterion    -285.6009
 DW-statistic                 2.1120
*******************************************************************************
```

```
                       Ordinary Least Squares Estimation
               Based on White's Heteroscedasticity adjusted S.E.'s
*******************************************************************************
 Dependent variable is FAVWIN
 553 observations used for estimation from    1 to   553
*******************************************************************************
 Regressor              Coefficient       Standard Error         T-Ratio[Prob]
 C                           .57695           .031657             18.2251[.000]
 SPREAD                     .019366           .0019218            10.0766[.000]
*******************************************************************************
```

Answer (ii)

(ii) The linear probability model estimated by OLS gives

$$\widehat{favwin} = \underset{\substack{(0.028)\\ [0.032]}}{0.577} + \underset{\substack{(0.0023)\\ [0.0019]}}{0.0194}\, spread$$

$n = 553, \quad R^2 = 0.111$

standard errors are in (·) and the heteroskedasticity-robust standard errors are in [·].

(ii) Answer

- Using the usual standard error, the ***t*** statistic for H_0: $\beta_0 = 0.5$ is (0.577 – 0.5)/0.028 = 2.75, which leads to rejecting H_0 against a two-sided alternative at the 1% level (critical value 2.58).

- Using the robust standard error reduces the significance but nevertheless leads to strong rejection of H_0 at the 2% level against a two-sided alternative: ***t*** = (0.577 – 0.5)/0.032= 2.41 (critical value 2.33).

Question (iii)

(iii) Is *spread* statistically significant? What is the estimated probability that the favored team wins when *spread* = 10?

Answer (iii)

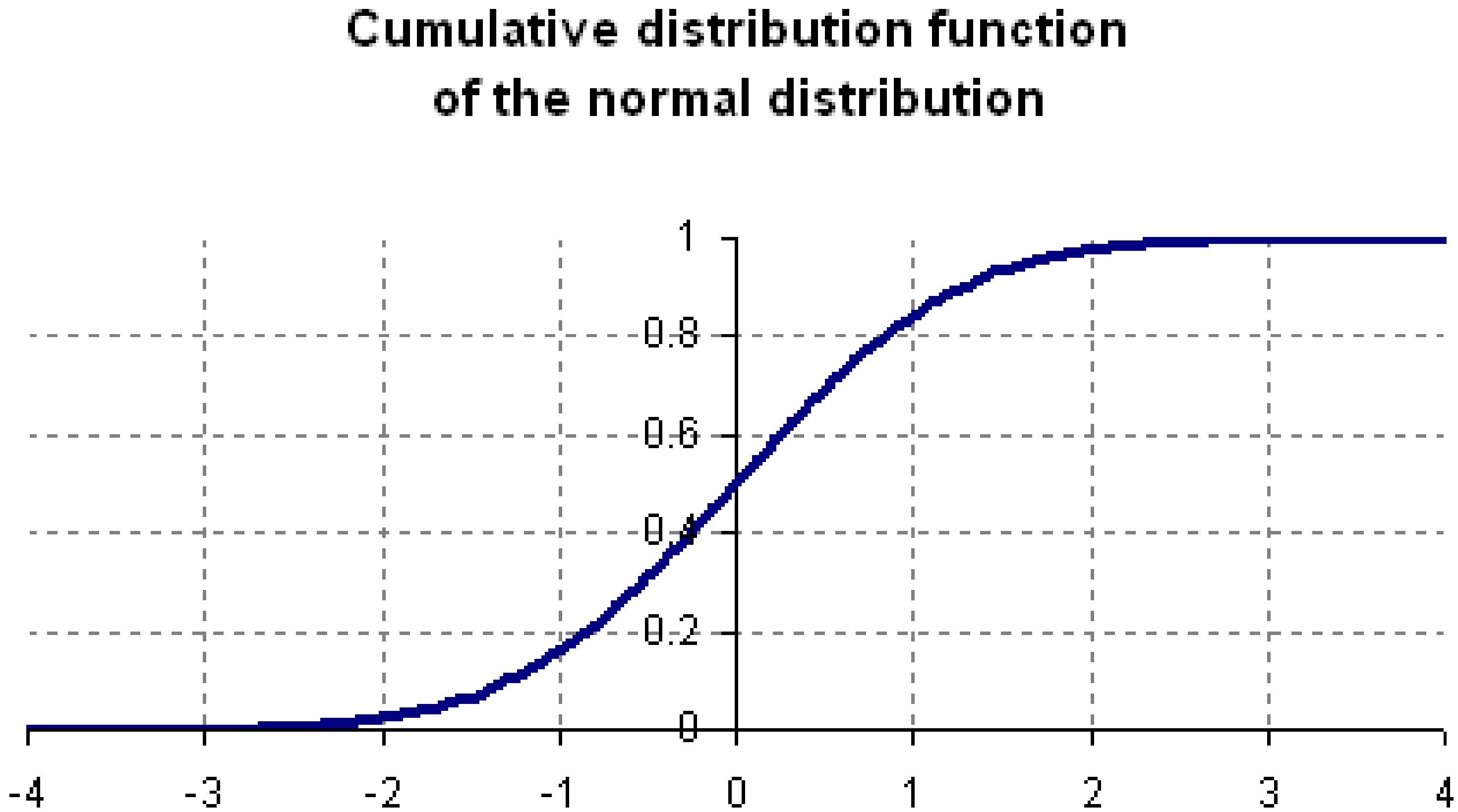
- As we expect, *spread* is very statistically significant using either standard error, with a ***t*** statistic greater than eight.

- If ***spread* = 10** the estimated probability that the favored team wins is 0.577 + 0.0194(10) = **0.771**

Question (iv)

(iv) Now, estimate a probit model for $P(favwin = 1|spread)$. Interpret and test the null hypothesis that the intercept is zero. [*Hint*: Remember that $\Phi(0) = .5$.]

Standard normal density function



Answer (iv)

The probit results are given in the following table:

Dependent Variable: *favwin*	
Independent Variable	Coefficient (Standard Error)
spread	0.0925 (0.0122)
constant	–0.0106 (0.1037)
Number of Observations	553
Log Likelihood Value	–263.56
Pseudo *R*-Squared	0.129

Answer (iv)

In the probit model $\mathbf{P}(\boldsymbol{favwin} = \mathbf{1}|\boldsymbol{spread}) = \Phi(\boldsymbol{\beta_0} + \boldsymbol{\beta_1 spread}),$

where $\Phi(\cdot)$ *denotes the standard normal cdf, if* $\beta_0 = 0$ *then*

$$P(favwin = 1|spread) = \Phi(\beta_1 spread)$$

and, in particular,

$$P(favwin = 1|spread = 0) = \Phi(0) = 0.5$$

This is the analog of testing whether the intercept is 0.5 in the LPM.

From the table, the ***t*** statistic for testing

H_0: $\beta_0 = 0$ *is only about -0.102, so we do not reject* H_0.

Question (v)

(v) Use the probit model to estimate the probability that the favored team wins when $spread = 10$. Compare this with the LPM estimate from part (iii).

Answer (v)

(v) When *spread* = 10 the predicted response probability from the estimated probit model is

$$\Phi[-0.0106 + 0.0925(10)] = \Phi(0.9144) \approx 0.820.$$

This is somewhat above the estimate for the LPM.

Question (vi)

(vi) Add the variables *favhome*, *fav25*, and *und25* to the probit model and test joint significance of these variables using the likelihood ratio test. (How many *df* are in the chi-square distribution?) Interpret this result, focusing on the question of whether the spread incorporates all observable information prior to a game.

Answer (vi)

- When ***favhome***, ***fav25***, and ***und25*** are added to the probit model, the value of the log-likelihood becomes –262.64.
- Therefore, the likelihood ratio statistic is
  2[–262.64 – (–263.56)] = 2(263.56 – 262.64) = 1.84.
- The *p*-value from the Chi^2_3 distribution is about 0.61, so ***favhome***, ***fav25***, and ***und25*** are jointly very insignificant.
- Once ***spread*** is controlled for, these other factors have no additional power for predicting the outcome.

- Project time….