Mohamed Khidher University

Faculty of economics, Commercial and Management Sciences

Department of commerce

Specialty: International Commerce & Finance

Module: Advanced Econometric

Academic year 2024/2025

<u>Exercise series N°03</u>

Exercise 1: choose the correct answer for the following questions

1. **What is the primary difference between a probit model and a logit model?**
   a) The probit model uses the logistic function, while the logit model uses the normal cumulative distribution function (CDF).
   b) The probit model assumes a normal distribution of errors, whereas the logit model assumes a logistic distribution of errors.
   c) The probit model is used for linear regression, while the logit model is used for classification.
   d) There is no fundamental difference between the two models.
2. **In a probit model, what type of function links the independent variables to the probability of an event occurring?**
   a) Linear function
   b) Logarithmic function
   c) Normal cumulative distribution function (CDF)
   d) Exponential function
3. **Which of the following is an assumption of the probit model?**
   a) The error terms follow a standard normal distribution.
   b) The dependent variable must be continuous.
   c) The independent variables must be normally distributed.
   d) The dependent variable must have more than two categories.
4. **If the estimated coefficient in a probit model is positive, what does it imply?**
   a) The probability of the event occurring decreases.
   b) The probability of the event occurring increases.
   c) The variable is not significant.
   d) The direction of the effect cannot be determined.

5. **If a probit model is estimated as follows:** $P(Y=1|X)=\Phi(0.5+1.5X)$

   What is the probability of Y=1 when X=1?
   a) 0.9772
   b) 0.5000
   c) 0.0228
   d) 0.7500

6. **Which of the following is a common issue when interpreting probit model coefficients?**
   a) The coefficients directly represent the change in probability.
   b) The coefficients do not have a direct probabilistic interpretation.
   c) The coefficients can be interpreted in the same way as those in a linear regression model.
   d) The coefficients are always positive.
7. **Which test is commonly used to assess the overall goodness-of-fit of a probit model?**
   a) Durbin-Watson test
   b) Wald test
   c) Hosmer-Lemeshow test
   d) McFadden's R^2
8. **Which of the following distributions does the probit model assume for the error term?**
   a) Standard normal distribution
   b) Exponential distribution
   c) Logistic distribution
   d) Poisson distribution
9. **You estimate a probit model and find the following output:** Intercept: -1.2, Coefficient on X: 2.4

What is the probability of Y=1 when X=1.0?
a) 0.8849
b) 0.1151
c) 0.5000
d) 0.7500

10. **A probit model estimates $P(Y=1|X)=\Phi(2.3X-1.1)$. If X=0.8, what is the probability of Y=1?**
    a) 0.7704
    b) 0.2296
    c) 0.5000

11. **In a probit model, a coefficient of** -0.7 **for an independent variable means:**
    a) The probability decreases, but the change depends on the standard normal distribution.
    b) The probability decreases by exactly 0.7 for every unit increase in the variable.
    c) The probability decreases by 70%.
    d) The probability is always negative.

12. **If $\beta 3=-0.4$ and X increases from 1 to 2, what happens to the probability of the event occurring?**
    a) It increases
    b) It decreases
    c) It stays the same
    d) It becomes exactly 0.5000

13. A researcher estimates a probit model with the following equation:

$Z=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3$

where: $\beta_0=-0.8$, $\beta_1=1.2$, $\beta_2=0.5$, $\beta_3=-0.4$

If $X_1=2$, $X_2=3$, and $X_3=1$, what is the predicted **Z-score** and the probability of event occurring?

a) z-score is 2.7, the probability id 0.9965

b) z-score is 1.8, the probability id 0.9641

c) z-score is 2.7, the probability id 0.9820

d) z-score is 1.8, the probability id 0.9965

Exercise 2

A bank wants to predict whether a loan applicant will default on their loan based on: **Credit Score (X1)** and **Loan Amount (X2)**

1. Write the Probit model equation to predict the probability of loan default ;

$$probability\ of\ default = \Phi(\beta_0 + \beta_1 Loan + \beta_2 Credit)$$

2. Estimate the parameters, interpret the results;

Dependent Variable: DEFAULT
Method: ML - Binary Probit (Newton-Raphson / Marquardt steps)
Date: 03/02/25 Time: 21:50
Sample: 1 30
Included observations: 30
Convergence achieved after 4 iterations
Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
CREDIT_SCORE	-0.017764	0.021377	-0.830975	0.4060
LOAN_AMOUNT	0.004789	0.151924	0.031524	0.9749
C	12.89640	17.68864	0.729078	0.4660

McFadden R-squared	0.281459	Mean dependent var	0.666667
S.D. dependent var	0.479463	S.E. of regression	0.419410
Akaike info criterion	1.114723	Sum squared resid	4.749425
Schwarz criterion	1.254843	Log likelihood	-13.72085
Hannan-Quinn criter.	1.159549	Deviance	27.44170
Restr. deviance	38.19085	Restr. log likelihood	-19.09543
LR statistic	10.74915	Avg. log likelihood	-0.457362
Prob(LR statistic)	0.004633		
Obs with Dep=0	10	Total obs	30
Obs with Dep=1	20		

$$probability\ of\ default = \Phi(12.89 + 0.0047 Loan - 0.0177 Credit)$$

Loan Coefficient (0.0047): A 1-unit increase in the loan amount increases the probability of default, since the coefficient is **positive**, higher loan amounts are associated with a **higher likelihood of default**.

Credit Coefficient (-0.0177): A 1-unit increase in credit score reduces the probability of default because the coefficient is **negative**. A **higher credit score is associated with lower default risk**, which aligns with intuition.

3. If a borrower has a **Credit Score of 650** and a **Loan Amount of $25,000$**, what is the predicted Z-score

$Z=12.89+(0.0047\times25000)-(0.0177\times650)$

$Z=118.885Z=118.885$

This **very high** Z-score suggests that the probability of default is **extremely close to 1** because the cumulative standard normal distribution Φ(Z) approaches 1 for large positive values.

4. Predict Default for a New Applicant, who has: **Credit Score =800, Loan Amount = $15,000$**

$Z=12.89+(0.0047\times15{,}000)-(0.0177\times800)$

$Z=12.89+70.5-14.16$, $Z=69.23$

where Φ(Z) is the **cumulative standard normal distribution (CDF)** evaluated at **Z = 69.23**.

Since **Z = 69.23** is extremely large, the normal CDF $\Phi(69.23)\approx1\Phi(69.23)\approx1$. This means the probability of default is **almost 100%**, which seems unrealistic.

5. If an applicant **increases** their loan amount from **$10,000$ to $20,000$**, and the coefficient for Loan what happens to the predicted Z-score?

Since only the **Loan** amount changes (Credit Score remains the same), the effect on **Z** is determined by the **Loan coefficient (0.0047)**:

$\Delta Z=0.0047\times(20{,}000-10{,}000)$

$\Delta Z=0.0047\times10{,}000$ $\Delta Z=47$

Z-score increases by 47 when the loan amount increases from $10,000$ to $20,000$.
Since the **Z-score represents the latent variable in a probit model**, this means that the **probability of default increases**. However, the exact change in probability depends on where the new Z-score falls on the **cumulative normal distribution Φ(Z))**.

If a **Credit Score = 750** and a **Loan Amount = $10,000$**, what is the predicted **Z-score** and probability of default? (Use Φ(-0.3) = **0.3821**) $Z=12.89+(0.0047\times10{,}000)-(0.0177\times750)$

$Z=46.615$ Since **Z = 46.615** is extremely high, it suggests a near-certain probability of default. However, this seems unrealistic for real-world cases, indicating a potential issue with the model coefficients or scaling.