Faculty of economics, Commercial   *Specialty: International Commerce & Finance*
and Management Sciences   *Module: Advanced Econometric*
Department of commerce   *Academic year 2024/2025*

Exercise series N°2

Exercise1: Choose the correct answer for the following questions

1. **What is a qualitative variable in statistical modeling?**
   a) A variable that represents quantities and can take any numerical value.
   b) A variable that represents categories or characteristics, often non-numerical.
   c) A variable that is always continuous and measurable.
   d) A variable that only includes time-related data.
2. **Which of the following is an example of a qualitative variable?**
   a) Age of an individual (in years).
   b) Monthly income (in dollars).
   c) Employment status (Employed, Unemployed).
   d) Temperature (in Celsius).
3. **The Logit Model is used to:**
   a) Model a continuous dependent variable based on multiple predictors.
   b) Predict probabilities of outcomes for a binary dependent variable.
   c) Determine the linear relationship between independent and dependent variables.
   d) Identify time-series patterns in data.
4. **The Logit Model formula can be expressed as:**
   a) $P(Y=1)=\beta_0+\beta_1 X$.
   b) $P(Y=1)=1/(1+e^{-(\beta_0+\beta_1 X)})$
   c) $Y=\ln(\beta_0+\beta_1 X)$
   d) $Y=e^{\beta_0+\beta_1 X}$
5. **The Logit Model is best suited for:**
   a) Predicting stock prices.
   b) Modeling binary outcomes such as "success vs. failure."
   c) Determining relationships between continuous variables.
   d) Identifying seasonal trends in data.
6. **In a Logit Model, the coefficients (β) represent:**
   a) The change in the log-odds of the dependent variable for a one-unit change in the predictor.
   b) The probability of success.
   c) The likelihood of multicollinearity.
   d) The mean value of the independent variable.
7. **Which of the following is an application of the Logit Model?**
   a) Predicting the likelihood of a customer purchasing a product (Yes/No).
   b) Estimating the annual revenue of a company.
   c) Modeling temperature changes over time.

d) Determining the variance of residuals in linear regression.

8. The dependent variable in a Logit Model:

a) Can take on any real number.

b) Must be a binary variable (e.g., 0 or 1).

c) Must be a categorical variable with more than two levels.

d) Represents the residual error.

9. If a logistic regression model produces a coefficient of 0.7 for a variable, what does the corresponding odds ratio indicate?

a) The odds increase by **70%** for a one-unit increase in the independent variable

b) The odds increase by **100%** for a one-unit increase in the independent variable

c) The odds increase by $e^{0.7} \approx$ **2.01 or 101%** for a one-unit increase in the independent variable

d) The probability of the event occurring is exactly 0.7

10. In a dataset for a Logit Model, which of the following is most appropriate for the dependent variable?

a) The height of individuals in centimeters.

b) The income of individuals in dollars.

c) Whether a student passes or fails an exam (Pass = 1, Fail = 0).

d) The number of hours worked per week.

11. Which of the following metrics is commonly used to assess the goodness-of-fit of a logistic regression model?

a) Adjusted R^2

b) Pseudo R^2 (e.g., McFadden's R^2)

c) Mean Squared Error (MSE)

d) Variance Inflation Factor (VIF)

Exercise 2

The goal of this study is to **predict whether a customer will purchase a product (Y=1) or not (Y=0)** based on four key independent variables: **Income** – The financial capability of the customer, which may affect purchasing power. **Age** – The age of the customer, which could influence spending behavior. **Advertisement Exposure** – The number of ads seen by the customer, which may impact product awareness and decision-making. **Product Price** – The cost of the product, which plays a critical role in affordability and purchase likelihood.

1. Write the logit model, and the logit regression equation
2. Using the given data:Perform logistic regression (using EViews). Obtain the estimated values of β_0, β_1, β_2 and β_3.
3. Explain the effect of independent variables on the likelihood of purchase based on the sign and magnitude of β_1 and β_2, β_3.
4. Predict the probability of purchase for a customer with income 70, age 40, Ads seen 7 and price of 67
5. Which factor has the **strongest influence** on customer purchasing decisions?

Solution

1. Write the logit model, and the logit regression equation

$$P(Y = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 Inc + \beta_2 Age + \beta_3 Adv + \beta_4 Price)}}$$

Logi regression : $ln\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \beta_0 + \beta_1 Inc + \beta_2 Age + \beta_3 Adv + \beta_4 Price$

Dependent Variable: PURCHASE_DECISION
Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)
Date: 02/17/25   Time: 09:05
Sample: 1 40
Included observations: 40
Convergence achieved after 9 iterations
Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PRODUCT_PRICE	0.107490	0.119576	0.898928	0.3687
INCOME	-0.078677	0.188334	-0.417756	0.6761
AD_EXPOSURE	1.111428	1.073290	1.035533	0.3004
AGE	-0.229501	0.222031	-1.033643	0.3013
C	-15.41308	8.295061	-1.858103	0.0632

McFadden R-squared	0.800026	Mean dependent var		0.575000
S.D. dependent var	0.500641	S.E. of regression		0.225012
Akaike info criterion	0.522706	Sum squared resid		1.772057
Schwarz criterion	0.733816	Log likelihood		-5.454125
Hannan-Quinn criter.	0.599037	Deviance		10.90825
Restr. deviance	54.54837	Restr. log likelihood		-27.27418
LR statistic	43.64012	Avg. log likelihood		-0.136353
Prob(LR statistic)	0.000000			

Obs with Dep=0	17	Total obs		40
Obs with Dep=1	23			

$$logit\ (purchase) = -15.41 - 0.078Inc - 0.229Age + 1.11Adv + 0.107Price$$

-The logit regression shows that the income has a coefficient of -0.078, indicating a weak negative relationship with the probability of purchase.

A one-unit increase in income decreases the odds of purchase by approximately 7.5%. however, since the **p-value is 0.67 (> 0.05)**, we **cannot conclude that this relationship is statistically significant**.

-The logit regression shows that age has a coefficient of -0.22, indicating a medium negative relationship with the probability of purchase.

A one-unit increase in Age decreases the odds of purchase by 2.26%. However, since the **p-value is 0.30 (> 0.05)**, we **cannot conclude that this relationship is statistically significant**.

-The logit regression shows that price has a coefficient of 0.107, indicating a medium positive relationship with the probability of purchase.

A one-unit increase in price increase the odds of purchase by 11.29%. However, since the **p-value is 0.36 (> 0.05)**, we **cannot conclude that this relationship is statistically significant**.

--The logit regression shows that Ads seen has a coefficient of 1.111, indicating a strong positive relationship with the probability of purchase.

A one-unit increase in Ads seen increases the odds of purchase by 203.7%. However, since the **p-value is 0.30 (> 0.05)**, we **cannot conclude that this relationship is statistically significant**.

Predict the probability of purchase for a customer with income 70, age 40, Ads seen 7 and price of 67

$$\widehat{P} = \frac{e^{(-15.41-0.078*70+1.111*7+0.107*67-0.22*40)}}{1+e^{(-15.41-0.078*70+1.111*7+0.107*67-0.22*40)}} = \frac{e^{-21,89}}{1+e^{-21,89}} = 3{,}10449E\text{-}10$$

the **strongest influence** on customer purchasing decisions is the Ads seen as one unit increase in advertising exposure increase the probability of purchase by 203,7%