eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea
CAMPUS OF INTERNATIONAL EXCELLENCE

Exercise E6

The Multiple Regression Model. Inference

Pilar González and Susan Orbe

Dpt. Applied Economics III (Econometrics and Statistics)

Contents

Contents

E6.1. Beach umbrella rental.

Part I. Simple Linear Regression Model.

Estimate a simple linear regression model to determine the number of rented umbrellas as a function of the temperature:

$$U_t = \alpha + \beta\, T_t + u_t \quad t = 1, \ldots, 22 \tag{1}$$

a. Estimate the coefficients of the model by OLS. Save the fitted values and the residuals.

b. Is temperature a statistically significant variable to explain the number of rented umbrellas?

c. If the temperature increases by two degrees Celsius, would it be possible an increase of 20 units in the number of rented umbrellas?

d. Estimate the number of rented umbrellas when the temperature is 42 degrees Celsius.

e. Assuming that the temperature is 42 degrees Celsius, estimate the upper limit for the number of rented umbrellas ($\alpha = 5\,\%$).

E6.1. Beach umbrella rental.

Part II. General Linear Regression Model.

Estimate a linear regression model to determine the number of rented umbrellas as a function of temperature, price and whether it is a windy week or not:

$$U_t = \gamma_1 + \gamma_2 T_t + \gamma_3 P_t + \gamma_4 WW_t + w_t \quad t = 1, \ldots, 22 \tag{2}$$

a. Estimate the model by OLS and write down the Sample Regression Function.

b. Are the explanatory variables individually significant? Are they jointly significant?

c. Do the results of the previous tests suggest the presence of a high degree of collinearity in the sample?

d. Assume that the temperature is 40 degrees Celsius and the average rental price €7.5. Estimate the upper limit for the number of rented umbrellas in a windy week and in a non-windy week ($\alpha = 5\,\%$).

e. Compare models (1) and (2). Which model is more appropriate to determine the number of rented umbrellas? Justify your answer in terms of the results of the tests performed and the properties of the OLS estimator in each model.

Contents

E6.2. Holiday cottages.

Model A.

Consider the linear regression model that specifies a linear relationship among the price of a room, the number of rooms and the price of breakfast as follows:

$$RP_i = \alpha_1 + \alpha_2\, NR_i + \alpha_3\, BP_i + u_i \tag{3}$$

a. Test the overall significance of the explanatory variables.

b. Is the variable number of rooms statistically significant?

c. Assume that the price of breakfast increases by €1 holding the number of rooms fixed. Estimate the lower and upper limits for the change in the price of the room ($\alpha = 5\,\%$).

d. Estimate the price of a room when the holiday cottage has 10 bedrooms and the breakfast price is €3.

e. Estimate the lower and upper limits for the price of a room in a holiday cottage that has 10 bedrooms and charges €3 for breakfast ($\alpha = 5\,\%$).

E6.2. Holiday cottages.

Model B.

Include two more explanatory variables in model (3): access to WiFi and location. Consider the regression model:

$$RP_i = \lambda_1 + \lambda_2 NR_i + \lambda_3 BP_i + \lambda_4 WIFIG_i + \lambda_5 WIFIS_i + \lambda_6 LOCC_i + u_i \qquad (4)$$

where $WIFIF$ takes the value 1 if the holiday cottage offers free WiFi access and 0 otherwise; $WIFIP$ takes the value 1 if the holiday cottage offers WiFi access for an additional fee and 0 otherwise; and, $LOCC$ takes the value 1 if the holiday cottage is in the town center and 0 otherwise.

a. Estimate the coefficients of model (4) by OLS and write down the Sample Regression Function.
b. Are the new variables included in the model jointly significant?
c. Are more expensive the holiday cottages in the town center?
d. Is the variable *access to WiFi* statistically significant?
e. It is believed that what is really important for business is to offer access to WiFi regardless of whether it is free or not. Is there any evidence in the sample to support this hypothesis?

E6.2. Holiday cottages.

Model C.

Consider the linear regression model:

$$\begin{aligned} RP_i &= \beta_1 + \beta_2\, NR_i + \beta_3\, BP_i + \beta_4\, WIFIF_i + \beta_5\, NPR_i + \\ &+ \beta_6\, BER_i + \beta_7\, LKR_i + u_i \end{aligned} \tag{5}$$

where NPR, BER and LKR take the value 1 if the holiday cottage is less than 1 km from a natural park, a beach or a lake, respectively; and 0 otherwise.

a. Estimate the coefficients of model (5) by OLS and write down the Sample Regression Function.

b. Is the variable *having free access to WiFi* statistically significant?

c. Are the variables *proximity to a natural park*, *proximity to a beach* and *proximity to a lake or reservoir* jointly significant?

E6.2. Holiday cottages.

d. Does the *proximity to a natural park* influence the average price of a room, holding the rest of the characteristics constant? And the *proximity to a beach*? And the *proximity to a lake or reservoir*?

e. Should the variable *location* be included in the specification of the regression model to determine the price of a room?

f. Given the result obtained in the previous item, what are the properties of the OLS estimators in model (5)?

g. Given all the results obtained so far, specify a regression model to determine room prices. Justify your answer.

Contents

E6.3. Soy milk.

Part I. General Linear Regression Model.

Consider the general linear regression model where sales (S, in thousands of euros) depend on prices per litter (P, in euro cents) and advertising expenditures (AE, in hundreds of euros euros) as follows:

$$S_t = \beta_1 + \beta_2 P_t + \beta_3 AE_t + \beta_4 AE_t^2 + u_t \quad t = 1990:1, \ldots, 2012:6. \tag{6}$$

a. Estimate the coefficients by OLS and write down the Sample Regression Function.

b. Are the variables price and advertising expenditures jointly significant? Are they individually significant?

c. Is the relationship between sales and advertising expenditures linear?

d. Is there any evidence in the sample that an increment of 50 cents in the price generates a decrease in sales of €750, holding advertising expenditures fixed?

e. If the price per litter is 75 cents and the firm spends €20000 on advertising, would it be possible for sales to reach the amount of 25 thousands of euros?

E6.3. Soy milk.

Part II. Trend and seasonality.

The previous model is augmented by including a time trend and seasonal dummy variables:

$$S_t = \beta_1 + \beta_2 P_t + \beta_3\, AE_t + \beta_4\, AE_t^2 + \beta_5\, time + \beta_6 dm1_t + \beta_7\, dm2_t + \ldots + \beta_{16} dm11_t + u_t \quad (7)$$

where the dummy variables dmj_t, $j = 1, 2, \ldots, 11$, take the value 1 if the observation t belongs to month j and 0 otherwise.

a. Estimate the coefficients of the model by OLS.

b. Estimate soy milk sales in the months of December and August.

c. Is the trend variable statistically significant?

d. Is there any evidence in the sample of the presence of seasonality in sales?

e. Add the regressor P^2 to model (7). Could you conclude that the relationship between sales and prices is quadratic?

f. Given the results obtained so far, specify a model to determine soy milk sales. Justify your answer.