# 6ο φροντιστήριο

1. Στον παρακάτω πίνακα σας δίνονται οι παρατηρήσεις αναφορικά με τη ζήτηση του αγαθού Α

| Y=ποσότητα σε κιλά | X=τιμή σε ευρώ |
|---|---|
| 100 | 13 |
| 82 | 13.5 |
| 86 | 13.2 |
| 90 | 13 |
| 60 | 14 |
| 70 | 13.7 |

a. Αν η συνάρτηση ζήτησης είναι της μορφής $Y_i = \beta_0 + \beta_1 X_i + u_i$, $i = 1, \dots, 6$, να εκτιμηθούν οι παράμετροι $\beta_0$ και $\beta_1$ με τη μέθοδο ελαχίστων τετραγώνων.

b. Να υπολογιστεί η ελαστικότητα ζήτησης στο σημείο $(\bar{X}, \bar{Y})$.

c. Να σχολιάσετε τα αποτελέσματα από οικονομική άποψη (αναφερθείτε στα πρόσημα των εκτιμήσεων καθώς και στο μέγεθός τους).

d. Να αξιολογηθούν τα αποτελέσματα με βάση στατιστικά κριτήρια. Σημείωση: Υπολογίστε τον συντελεστή προσδιορισμού, τα τυπικά σφάλματα των συντελεστών και κατόπιν να ελέγξετε αν οι συντελεστές είναι ίσοι με το μηδέν, δηλ. $\beta_0 = 0$ και $\beta_1 = 0$. Τις εναλλακτικές υποθέσεις θα τις διατυπώσετε σύμφωνα με την οικονομική θεωρία.

e. Να κατασκευάσετε ένα 95% διάστημα εμπιστοσύνης για τον συντελεστή $\beta_1$.

f. Να κατασκευάσετε ένα 99% διάστημα εμπιστοσύνης για τη ζητούμενη ποσότητα όταν η τιμή είναι 13.6 ευρώ.

2)

g. Ποιες υποθέσεις πρέπει να ισχύουν και γιατί για να απαντήσουμε σε κάθε μια από τις παραπάνω ερωτήσεις?

Λύση:

a. Αν $y_i = Y_i - \bar{Y}, x_i = X_i - \bar{X}$, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{6}80.40 = 13.4$, $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{6}488 = 81.33$, $\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X}) = -28$ και $\sum_{i=1}^{n}(X_i - \bar{X})^2 = 0.82$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{-28}{0.82} = -34.146$$

$$\hat{\beta}_0 = \bar{\mathrm{Y}} - \hat{\beta}_1 \bar{\mathrm{X}} = 81.33 - (-34.146) * 13.4 = 538.88$$

Συνεπώς η παλινδρόμησης στο δείγμα είναι $\hat{Y}_i = 538.88 - 34.146X_i$

b. Η μέση ελαστικότητα ζήτησης δίνεται ως εξής

$$\bar{\varepsilon} = \hat{\beta}_1 \frac{\bar{\mathrm{X}}}{\bar{\bar{\mathrm{Y}}}} = -34.146\frac{13.4}{81.33} = -5.63$$

c. Τα πρόσημα των συντελεστών είναι αυτά που αναμένονται από την οικονομική θεωρία, δηλ. θετικό πρόσημο για το σταθερό όρο και αρνητικό για το συντελεστή $\beta_1$. Η ζήτηση του αγαθού είναι ελαστική αφού η ελαστικότητα στο σημείο των μέσων κατ' απόλυτη τιμή είναι μεγαλύτερη από τη μονάδα.

d. Ο συντελεστής προσδιορισμού μπορεί να υπολογιστεί από τη σχέση $R^2 = 1 - \frac{\sum_{i=1}^{n}\hat{u}_i^2}{\sum_{i=1}^{n}(Y_i - \bar{\mathrm{Y}})^2} = 1 - \frac{73.24}{1029.33} = 0.93$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\hat{u}_i^2}{n-k} = \frac{73.25}{6-2} = 18.31$$

$$var(\hat{\beta}_0) = \sigma^2 \frac{\sum_{i=1}^{n}X_i^2}{n\sum_{i=1}^{n}(X_i - \bar{X})^2} = 18.31\frac{1078.18}{6*0.82} = 4011.4,$$

$$se(\hat{\beta}_0) = \sqrt{4011.4} = 63.33$$

$$var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=}^{n}(X_i - \bar{X})^2} = \frac{18.31}{0.82} = 22.32, \; se(\hat{\beta}_1) = \sqrt{22.32} = 4.72$$

Έλεγχοι υποθέσεων

$H_0: \beta_0 = 0$

$H_1: \beta_0 > 0$

$$t = \frac{\hat{\beta}_0 - \beta_0}{se(\hat{\beta}_0)} \sim t_{(6-2)}$$

$t = \frac{\hat{\beta}_0 - \beta_0}{se(\hat{\beta}_0)} = \frac{538.88 - 0}{63.33} = 8.5$ και $t_{4,0.05} = 2.132$, $t > t_{4,0.05}$ επομένως η $H_0$ απορρίπτεται.

$H_0: \beta_1 = 0$

$H_1: \beta_1 < 0$

$t = \frac{\hat{\beta}_1 - \beta_1}{se(\hat{\beta}_1)} = \frac{-34.146 - 0}{4.72} = -7.23$ και $t_{4,0.05} = 2.132$, $t < t_{4,0.05}$ επομένως η $H_0$ απορρίπτεται.

e. Το διάστημα εμπιστοσύνης δίνεται ως εξής

$$\Pr\left(\hat{\beta}_1 - t_c \cdot se(\hat{\beta}_1) \leq \beta_1 \leq \hat{\beta}_1 + t_c \cdot se(\hat{\beta}_1)\right) = 0.95$$

$t_{4,0.025} = 2.78$

$$\hat{\beta}_1 - t_c \cdot se(\hat{\beta}_1) \leq \beta_1 \leq \hat{\beta}_1 + t_c \cdot se(\hat{\beta}_1)$$

$$-34.146 - 2.78 * 4.72 \leq \beta_1 \leq -34.146 + 2.78 * 4.72$$

$$-47.248 \leq \beta_1 \leq -21.043$$

Αν πάρουμε έναν άπειρο αριθμό δειγμάτων μεγέθους 6 παρατηρήσεων από τον πληθυσμό και για κάθε δείγμα υπολογίσουμε ένα διάστημα εμπιστοσύνης , 95% των

4/

διαστημάτων αυτών θα περιλαμβάνει την παράμετρο του πληθυσμού $\beta_1$ (το διάστημα που υπολογίσαμε θεωρούμε ότι είναι ένα από αυτά)

f. Για $X_f = 13.6$, $\hat{Y}_f = 538.88 - 34.146X_f = 74.49$

$$var(\hat{Y}_f) = \sigma^2\left[\frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} + 1\right] = 4.717$$

$$\hat{Y}_f - t_c \cdot \sqrt{var(\hat{Y}_f)} \leq Y_f \leq \hat{Y}_f + t_c \cdot \sqrt{var(\hat{Y}_f)}$$

$$t_{4,0.005} = 4.60$$

$$74.49 - 4.60 * \sqrt{4.717} \leq Y_f \leq 74.49 + 4.60 * \sqrt{4.717}$$

$$52.77 \leq Y_f \leq 96.21$$

σε επίπεδο σημαντικότητας 99% η ποσότητα που προβλέπουμε ότι θα ζητηθεί όταν η τιμή είναι 13.6 ευρώ βρίσκεται στο διάστημα $74.49 \pm 21.717$

g. Ισχύουν οι κλασικές υποθέσεις

9. Έστω ότι σας δίνεται ένα δείγμα 12 παρατηρήσεων για την τιμή ($X_i$) και την ποσότητα ($Y_i$) μιας μάρκας καφέ.

| i | Ποσότητα | Τιμή |
|---|---|---|
| 1 | 89 | 1 |
| 2 | 86 | 1 |
| 3 | 74 | 1 |
| 4 | 79 | 1 |
| 5 | 68 | 1 |
| 6 | 84 | 1 |
| 7 | 139 | 0.95 |
| 8 | 122 | 0.95 |
| 9 | 102 | 0.95 |
| 10 | 186 | 0.85 |
| 11 | 179 | 0.85 |
| 12 | 187 | 0.85 |

Η γραμμική σχέση που τις συνδέει τις μεταβλητές δίνεται από το παρακάτω γραμμικό υπόδειγμα $Y_i = \beta_0 + \beta_1 X_i + u_i$

i. Υπολογίστε τους εκτιμητές ΕΤ $\hat{\beta}_0$ και $\hat{\beta}_1$.

ii. Υπολογίστε το $TSS, ESS$ και $RSS$ και τον συντελεστή προσδιορισμού $R^2$.

iii. Υπολογίστε τη διακύμανση του διαταρακτικού όρου.

iv. Υπολογίστε το τυπικό σφάλμα του $\hat{\beta}_1$ και προβείτε σε έλεγχο στατιστικής σημαντικότητας για τον εκτιμητή της κλίσης ($a = 5\%$). Σας δίνεται ότι η κριτική τιμή είναι $t_{0.025,10} = 2.23$

v. Υπολογίστε το 95% διάστημα εμπιστοσύνης του συντελεστή κλίσης $\beta_1$

**Υπόδειξη**: Χρησιμοποιήστε αποκλίσεις από τους μέσους

*Λύση:*

| $i$ | $Y_i$ | $X_i$ | $Y_i-\bar{Y}$ | $X_i-\bar{X}$ | $(Y_i-\bar{Y})(X_i-\bar{X})$ | $(X_i-\bar{X})^2$ | $(Y_i-\bar{Y})^2$ | $\hat{Y}_i$ | $\hat{u}_i$ | $\hat{u}_i^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 89 | 1 | -27.250 | 0.05 | -1.3625 | 0.0025 | 742.5625 | 81.5833 | 7.4167 | 55.0069 |
| 2 | 86 | 1 | -30.250 | 0.05 | -1.5125 | 0.0025 | 915.0625 | 81.5833 | 4.4167 | 19.5069 |
| 3 | 74 | 1 | -42.250 | 0.05 | -2.1125 | 0.0025 | 1785.0625 | 81.5833 | -7.5833 | 57.5069 |
| 4 | 79 | 1 | -37.250 | 0.05 | -1.8625 | 0.0025 | 1387.5625 | 81.5833 | -2.5833 | 6.6736 |
| 5 | 68 | 1 | -48.250 | 0.05 | -2.4125 | 0.0025 | 2328.0625 | 81.5833 | -13.5833 | 184.5069 |
| 6 | 84 | 1 | -32.250 | 0.05 | -1.6125 | 0.0025 | 1040.0625 | 81.5833 | 2.4167 | 5.8403 |
| 7 | 139 | 0.95 | 22.750 | 0 | 0.0000 | 0 | 517.5625 | 116.2500 | 22.7500 | 517.5625 |
| 8 | 122 | 0.95 | 5.750 | 0 | 0.0000 | 0 | 33.0625 | 116.2500 | 5.7500 | 33.0625 |
| 9 | 102 | 0.95 | -14.250 | 0 | 0.0000 | 0 | 203.0625 | 116.2500 | -14.2500 | 203.0625 |
| 10 | 186 | 0.85 | 69.750 | -0.1 | -6.9750 | 0.01 | 4865.0625 | 185.5833 | 0.4167 | 0.1736 |
| 11 | 179 | 0.85 | 62.750 | -0.1 | -6.2750 | 0.01 | 3937.5625 | 185.5833 | -6.5833 | 43.3403 |
| 12 | 187 | 0.85 | 70.750 | -0.1 | -7.0750 | 0.01 | 5005.5625 | 185.5833 | 1.4167 | 2.0069 |
| sum | 1395 | 11.40 | 0.00 | 0.00 | -31.2000 | 0.045 | 22760.25 | 1395.0 | 0.00 | 1128.25 |

i. Εκτιμητές ΕΤ: $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$ και $\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(Y_i-\bar{Y})(X_i-\bar{X})}{\sum_{i=1}^{n}(X_i-\bar{X})^2}$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(Y_i-\bar{Y})(X_i-\bar{X})}{\sum_{i=1}^{n}(X_i-\bar{X})^2} = \frac{-31.2}{0.045} = -693.33$$

$$\overline{Y} = \frac{1395}{12} = 116.25 \text{ και } \overline{X} = \frac{11.40}{12} = 0.95$$

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{X} = 116.25 - (-693.33) * 0.95 = 794.91$$

Το εκτιμημένο υπόδειγμα είναι $\hat{Y}_i = 794.91 - 693.33X_i$

ii. $TSS = \sum_{i=1}^{n}(Y_i - \overline{Y})^2 = 22760.25$

$$RSS = \sum_{i=1}^{n} \hat{u}^2 = 1128.25$$

$$TSS = ESS + RSS \Rightarrow ESS = TSS - RSS = 22760.25 - 1128.25 = 21632$$

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{1128.25}{22760.25} = 0.95$$

iii. $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-k} = \frac{1128.25}{12-2} = 112.825$

iv. $se(\hat{\beta}_1) = \sqrt{\mathrm{Var}(\hat{\beta}_1)} = \sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}} = \sqrt{\frac{112.825}{0.045}} = 50.072$

$H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

$$t = \frac{\hat{\beta}_1 - \beta_1}{SE(\hat{\beta}_1)} = \frac{-693.33}{50.072} = -13.847$$

Για τον δικατάληκτο έλεγχο η $H_0$ απορρίπτεται αν η απόλυτη τιμή του στατιστικού κριτηρίου είναι μεγαλύτερη ή ίση της κριτικής τιμής της κατανομής του $|t| > t_c$, καθώς $|-13{,}847| < 2.23$ η $H_0$ απορρίπτεται

v. Το διάστημα εμπιστοσύνης δίνεται ως εξής

$\hat{\beta}_1 \pm t_c \cdot SE(\hat{\beta}_1)$

$\hat{\beta}_1^L = \hat{\beta}_1 - t_c \cdot SE(\hat{\beta}_1)$ και $\hat{\beta}_1^U = \hat{\beta}_1 + t_c \cdot SE(\hat{\beta}_1)$ αποτελούν αντίστοιχα το κάτω και το άνω όριο του διαστήματος εμπιστοσύνης $\beta_1$

Η κριτική τιμή δίνεται ως εξής $t_c = t_{a/2,n-k} = t_{0.025,10} = 2.23$

Άρα $\hat{\beta}_1^L = \hat{\beta}_1 - t_c \cdot SE(\hat{\beta}_1) = -693.33 - 2.23 * 50.072 = -804.994$

και $\hat{\beta}_1^U = \hat{\beta}_1 + t_c \cdot SE(\hat{\beta}_1) = -693.33 + 2.23 * 50.072 = -581.672$