

ΘΕΜΑ 1ο

α. Έχουμε αταξινόμητα δεδομένα και θα χρησιμοποιήσουμε τους τύπους του τυπολογίου.

Οι αντίστοιχοι αριθμητικοί υπολογισμοί περιέχονται στο επισυναπτόμενο αρχείο Excel, Φύλλο Askisi-1-erotima-a.

Ταξινομούμε αρχικά τα δεδομένα μας σε αύξουσα σειρά και έχουμε:

Αριθμητικός μέσος:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{2080}{48} \Rightarrow \bar{X} = 43,33$$

Επικρατούσα τιμή:

Παρατηρούμε ότι η τιμή η οποία εμφανίζεται πιο συχνά είναι το 59 με εμφάνιση 4 φορές, άρα:

$T_0 = 59$

Διάμεσος:

Η θέση της M είναι:

$$\frac{n+1}{2} = \frac{48+1}{2} = \frac{49}{2} = 24,50 = 24 + 0,50$$

άρα:

$$M = X_{24} + 0,50(X_{25} - X_{24}) = 44 + 0,50(45 - 44) \Rightarrow \mathbf{M = 44,5}$$

Εύρος:

$$R = X_{\max} - X_{\min} = X_{48} - X_1 = 64 - 21 \Rightarrow$$

$R=43$

Ενδοτεταρτημοριακό Εύρος:

Η θέση του Q_1 είναι:

$$\frac{n+1}{4} = \frac{48+1}{4} = \frac{49}{4} = 12,25 = 12 + 0,25$$

άρα:

$$Q_1 = X_{12} + 0,25(X_{13} - X_{12}) = 32 + 0,25(33 - 32) \Rightarrow \mathbf{Q_1 = 32,25}$$

Η θέση του Q_3 είναι:

$$\frac{3(n+1)}{4} = \frac{3(48+1)}{4} = \frac{144}{4} = 36,75 = 36 + 0,75$$

άρα:

$$Q_3 = X_{36} + 0,75(X_{37} - X_{36}) = 55 + 0,75(55 - 55) \Rightarrow \mathbf{Q_3 = 55,00}$$

Ενδοτεταρτημοριακό Εύρος $IR = Q_3 - Q_1 = 55,00 - 32,25 \Rightarrow IR = 22,75$

Διακύμανση

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{7406,67}{48-1} \Rightarrow S^2 = 157,59$$

Τυπική Απόκλιση

$$S = +\sqrt{S^2} = \sqrt{157,59} \Rightarrow S = 12,55$$

Συντελεστής ασυμμετρίας (β_3):

$$\beta_3 = \frac{\frac{\sum_{i=1}^n (X_i - \bar{X})^3}{n}}{S^3} = \frac{-15900,44}{(12,55)^3} \Rightarrow \beta_3 = -0,1674 < 0$$

άρα υπάρχει αρνητική ασυμμετρία, δηλ. σπανίζουν οι μικρές τιμές των ηλικιών των εργαζομένων.

β. Η κάθε τάξη θα είναι της μορφής $[a, \beta)$ δηλαδή ανοιχτό διάστημα από δεξιά, άρα δεν θα περιλαμβάνει την τιμή β .

Επομένως ο Πίνακας κατανομής συχνοτήτων θα έχει την ακόλουθη μορφή:

Τάξη	Κεντρική Τιμή Τάξης m_i	Συχνότητα f_i	Αθροιστική συχνότητα F_i
[20-30)	25	9	9
[30-40)	35	10	19
[40-50)	45	10	29
[50-60)	55	17	46
[60-70)	65	2	48
ΣΥΝΟΛΟ		48	

γ. Έχουμε ταξινομημένα δεδομένα και θα χρησιμοποιήσουμε τους τύπους του τυπολογίου.

Οι αντίστοιχοι αριθμητικοί υπολογισμοί περιέχονται στο επισυναπτόμενο αρχείο Excel, Φύλλο Askisi-1-erotima-g-d

Αριθμητικός μέσος

$$\bar{X} = \frac{\sum_{i=1}^5 f_i m_i}{\sum_{i=1}^5 f_i} = \frac{2090}{48} \Rightarrow \bar{X} = 43,5417$$

Επικρατούσα τιμή

Εντοπισμός της θέσης του T_0 :

Η τάξη με τη μεγαλύτερη συχνότητα (δηλ. το 17) είναι η 4^η . Άρα η επικρατούσα τιμή βρίσκεται στην τάξη αυτή, δηλ. στο διάστημα [50-60).

Υπολογισμός της τιμής του T_0 :

$$T_o = L_{T_o} + \delta \frac{\Delta_1}{\Delta_1 + \Delta_2} = 50 + 10 \frac{(17-10)}{(17-10) + (17-2)} = 50 + 10 \frac{7}{22} \Rightarrow$$

$$T_o = 53,18 \text{ €}$$

Πρώτο Τεταρτημόριο

Εντοπισμός της θέσης του Q_1

$$\frac{n}{4} = \frac{48}{4} = 12$$

δηλ. πριν από την τιμή 19 στην στήλη της αθροιστικής συχνότητας F_i , που αντιστοιχεί στην 2^η τάξη, άρα το Q_1 ανήκει στην 2^η τάξη, δηλ. στο διάστημα [30-40).

Υπολογισμός της τιμής του Q_1 :

$$Q_1 = L_{Q_1} + \frac{\delta}{f_{Q_1}} \left(\frac{n}{4} - F_{Q_1-1} \right) = 30 + \frac{10}{10} (12 - 9) \Rightarrow Q_1 = 33$$

Διακύμανση

$$S^2 = \frac{\sum_{i=1}^n f_i (m_i - \bar{X})^2}{\sum_{i=1}^n f_i - 1} = \frac{6997,92}{47} \Rightarrow S^2 = 148,89$$

Τυπική Απόκλιση

$$S = +\sqrt{S^2} = \sqrt{148,89} \Rightarrow S = 12,20$$

δ. Χρησιμοποιώντας τις ιδιότητες των στατιστικών μέτρων για την νέα μεταβλητή

$$Y = 10 X$$

καθώς επίσης και τα αποτελέσματα του ερωτήματος α. καταλήγουμε στα εξής:

Αριθμητικός Μέσος

$$\bar{Y} = 10\bar{X} = 10(43,33) = 433,3$$

Τυπική απόκλιση

$$S_y = 10S_x = 10(12,55) = 125,5$$

Εύρος

$$R_y = 10R_x = 10(43) = 430$$

Διακύμανση

$$S_y^2 = 10^2 S_x^2 = 100(157,59) = 15759 .$$

ΘΕΜΑ 2ο

α. Έστω τα ακόλουθα ενδεχόμενα:

E = το τεμάχιο δεν είναι ελαττωματικό

A₁ = το τεμάχιο προέρχεται από τη μηχανή Α, P(A₁)=0,25

A₂ = το τεμάχιο προέρχεται από τη μηχανή Β, P(A₂)=0,35

A₃ = το τεμάχιο προέρχεται από τη μηχανή Γ, P(A₃)=0,40

Επομένως σύμφωνα με τον τύπο της ολικής πιθανότητας θα έχουμε:

$$P(E) = P(A_1)P(E | A_1) + P(A_2)P(E | A_2) + P(A_3)P(E | A_3) \Rightarrow$$

$$P(E) = 0,25(1 - 0,05) + 0,35(1 - 0,04) + 0,40(1 - 0,02) \Rightarrow$$

$$P(E) = 0,25(0,95) + 0,35(0,96) + 0,40(0,98) \Rightarrow$$

$$\Rightarrow P(E) = 0,9655 = 96,55\%$$

β. Έστω τα ακόλουθα ενδεχόμενα:

E' = το τεμάχιο είναι ελαττωματικό

A₁ = το τεμάχιο προέρχεται από τη μηχανή Α, P(A₁)=0,25

A₂ = το τεμάχιο προέρχεται από τη μηχανή Β, P(A₂)=0,35

A₃ = το τεμάχιο προέρχεται από τη μηχανή Γ, P(A₃)=0,40

Επομένως σύμφωνα με τον τύπο της ολικής πιθανότητας θα έχουμε:

$$P(E') = P(A_1)P(E' | A_1) + P(A_2)P(E' | A_2) + P(A_3)P(E' | A_3) \Rightarrow$$

$$P(E') = 0,25(0,05) + 0,35(0,04) + 0,40(0,02) \Rightarrow$$

$$P(E') = 0,0345$$

Σύμφωνα με τον τύπο του Bayes θα έχουμε:

$$P(A_2 | E') = \frac{P(A_2)P(E' | A_2)}{P(E')} = \frac{0,35(0,04)}{0,0345} \Rightarrow \boxed{P(A_2 | E) = 0,4058 = 40,58\%}$$

ΘΕΜΑ 3ο

Έστω τα ακόλουθα ενδεχόμενα:

A= ένας πελάτης διαθέτει τρεχούμενο λογαριασμό

B= ένας πελάτης διαθέτει αποταμιευτικό λογαριασμό

Τότε από την εκφώνηση έχουμε ότι:

$$P(A)=0,55$$

$$P(B)=0,60$$

$$P(A \cap B) = 0,40$$

α. Επομένως έχουμε ότι:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0,55 + 0,60 - 0,40 \Rightarrow$$

$$P(A \cup B) = 0,75 = 75\%$$

β. Επίσης ισχύει:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{0,40}{0,60} \Rightarrow$$

$$P(A | B) = 0,6667 = 66,67\%$$

γ. Για οποιαδήποτε δύο ενδεχόμενα ισχύει:

$$P(A) = P(A \cap B) + P(A \cap B') \Rightarrow$$

$$P(A \cap B') = P(A) - P(A \cap B) = 0,55 - 0,40 \Rightarrow$$

$$P(A \cap B') = 0,15 = 15\%$$

ΘΕΜΑ 4ο

α. Έστω A_i , $i=1,2,3,4$ το ενδεχόμενο ο ερωτώμενος νέος να απάντησε ότι προτιμάει το i -στό αναψυκτικό. Από τη στιγμή που τα ανωτέρω ενδεχόμενα είναι ανεξάρτητα μεταξύ τους, θα ισχύει ότι:

$$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_1)P(A_2)P(A_3)P(A_4)$$

Όμως και τα αναψυκτικά είναι τέσσερα, άρα έχουμε ότι:

$$P(A_1) = P(A_2) = P(A_3) = P(A_4) = \frac{1}{4}$$

Συνεπώς τελικά θα ισχύει:

$$P(A_1 \cap A_2 \cap A_3 \cap A_4) = \left(\frac{1}{4}\right)^4 = \frac{1}{256} = 0,0039 = 0,39\%$$

β. Έστω το ενδεχόμενο:

Δ = ο ερωτώμενος απαντά ότι προτιμά τα δύο περισσότερα διαφημιζόμενα αναψυκτικά:

Τότε έχουμε ότι:

$$P(\Delta) = \frac{2}{4} = \frac{1}{2}$$

Πάλι όπως και στο α. Ερώτημα, εφόσον οι προτιμήσεις είναι ανεξάρτητες, θα έχουμε ότι:

$$P(\Delta \cap \Delta \cap \Delta \cap \Delta) = P(\Delta)P(\Delta)P(\Delta)P(\Delta) = \left(\frac{1}{2}\right)^4 = \frac{1}{16} = 0,0625 = 6,25\%$$

ΘΕΜΑ 5ο

Έστω η τυχαία μεταβλητή:

X = εβδομαδιαία έσοδα της επιχείρησης σε ευρώ

Τότε η X ακολουθεί την κανονική κατανομή:

$$X \sim N(\mu, \sigma^2) = N(1000, 125^2)$$

επομένως:

α.

i.

$$\begin{aligned} P(X < 800) &= P\left(\frac{X - \mu}{\sigma} < \frac{800 - 1000}{125}\right) = \\ P\left(Z < \frac{-8}{5}\right) &= P(Z < -1,6) = \Phi(-1,6) = 1 - \Phi(1,6) = \\ 1 - 0,9452 &= 0,0548 = 5,48\% \end{aligned}$$

ii.

$$\begin{aligned} P(X > 1250) &= 1 - P(X < 1250) = 1 - P\left(\frac{X - \mu}{\sigma} < \frac{1250 - 1000}{125}\right) = \\ &= 1 - P(Z < 2) = 1 - \Phi(2) = 1 - 0,9772 = 0,0228 = 2,28\% \end{aligned}$$

iii.

$$\begin{aligned} P(20 < X < 25) &= P\left(\frac{900 - 1000}{125} < \frac{X - \mu}{\sigma} < \frac{1100 - 1000}{125}\right) = \\ P(-0,8 < Z < 0,8) &= \Phi(0,8) - \Phi(-0,8) = \Phi(0,8) - (1 - \Phi(0,8)) = 2\Phi(0,8) - 1 = \\ 2\Phi(0,8) - 1 &= 2(0,7881) - 1 = 0,5762 = 57,62\% \end{aligned}$$

β. Η πιθανότητα τα έσοδα να μην ξεπερνούν το μέσο έσοδο είναι:

$$P(X < 1000) = P\left(\frac{X - \mu}{\sigma} < \frac{1000 - 1000}{125}\right) = P(Z < 0) = \Phi(0) = 0,5$$

Λόγω της ανεξαρτησίας των ενδεχομένων η πιθανότητα που ζητείται θα είναι το γινόμενο της ανωτέρω πιθανότητας με την πιθανότητα οι εβδομαδιαίες οφειλές να ξεπερνούν το μέσο έσοδο, δηλ. θα είναι ίση με:

$$(0,5)(0,02) = 0,01 = 1\%$$

ΘΕΜΑ 6ο

Ορίζουμε την τυχαία μεταβλητή:

X = αριθμός καταστημάτων που κλείνουν σε ένα χρόνο

τότε η τυχαία μεταβλητή X θα ακολουθεί την κατανομή Poisson:

$$X \sim P(\lambda) = P(9)$$

α. Ένα τετράμηνο αντιστοιχεί στο ένα τρίτο του έτους, επομένως τώρα η μέση τιμή θα είναι:

$\lambda = 9/3 = 3$ καταστήματα κλείνουν κατά μέσο όρο σε ένα τετράμηνο, άρα:

$$P(X=0) = \frac{e^{-3}(3)^0}{0!} = e^{-3} = 0,0498 = 4,98\%$$

β. Έχουμε για ένα χρόνο, οπότε $\lambda = 9$, ότι:

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) =$$

$$1 - \frac{e^{-9}(9)^0}{0!} = 1 - e^{-9} = 1 - 0,0001234 = 0,99988 = 99,988\%$$

ΘΕΜΑ 7ο

α. Η πιθανότητα επιτυχίας σε μία δοκιμή, δηλ. σε μία ερώτηση, είναι:

$$p = 1/4$$

Επομένως η τυχαία μεταβλητή X θα ακολουθεί τη διωνυμική κατανομή:

$$X \sim B(n=25, p=0.25)$$

Συνεπώς θα έχουμε ότι:

$$\mu = E(X) = np = 25(0,25) = 6,25 \text{ σωστές απαντήσεις}$$

$$\sigma^2 = V(X) = np(1-p) = 25(0,25)(0,75) = 4,6875 \Rightarrow$$

$$\sigma = \sqrt{V(X)} = \sqrt{4,6875} = 2,165$$

β. Πρέπει να υπολογίσουμε την πιθανότητα:

$$P(X \geq 13) = 1 - P(X < 13) = 1 - P(X \leq 12) =$$

$$1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) - P(X = 4) - P(X = 5) -$$

$$P(X = 6) - P(X = 7) - P(X = 8) - P(X = 9) - P(X = 10) - P(X = 11) - P(X = 12)$$

Επειδή όμως ο τύπος της διωνυμικής κατανομής θα απαιτήσει πάρα πολλούς αριθμητικούς υπολογισμούς, θα λύσουμε το πρόβλημα με την προσέγγιση της διωνυμικής από την κανονική κατανομή, αφού $n = 25$ (κοντά στο 30) και $np = 6,25$ (κοντά στο 5). Θεωρούμε λοιπόν ότι κατά προσέγγιση ισχύει:

$$X \sim N(np, \sigma^2) = N(6.25, 2.165^2)$$

Τώρα επιτυχών θεωρείται ο εξεταζόμενος όταν

$$X \geq 12,5$$

Οι φοιτητές θα κάνουν την δική τους εργασία σκεπτόμενοι πάνω στις ενδεικτικές απαντήσεις.

Σε περίπτωση αντιγραφής ή ύπαρξης παροραμάτων δεν φέρουμε καμία ευθύνη.

διότι κάνουμε προσέγγιση μιας διακριτής από μία συνεχή μεταβλητή. Τελικά θα έχουμε ότι:

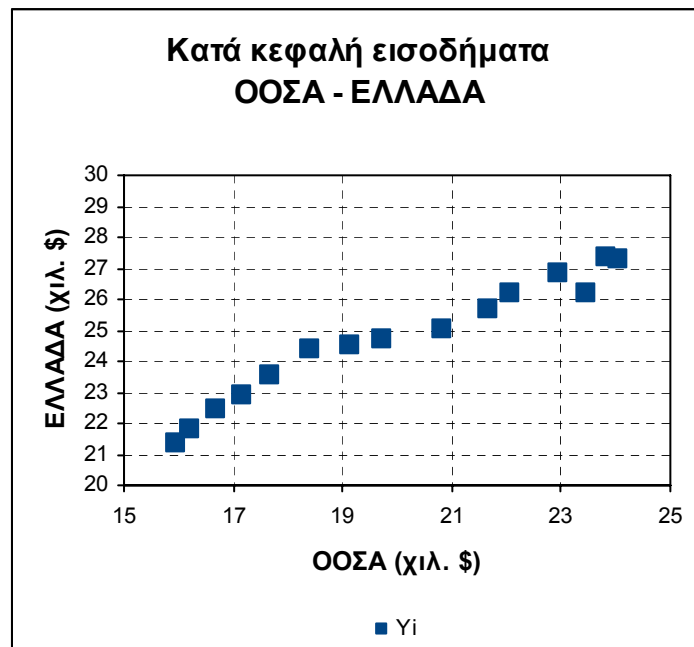
$$P(X \geq 12,5) = 1 - P(X < 12,5) = 1 - P\left(\frac{X - \mu}{\sigma} < \frac{12,5 - 6,25}{2,165}\right) = 1 - P(Z < 2.89) =$$

$$1 - \Phi(2.89) = 1 - 0,9981 = 0,0019 = 0,19\%$$

δηλ. η πιθανότητα δεν είναι ούτε καν μισό τοις εκατό.

ΘΕΜΑ 8ο

α. Θεωρώντας ως X το κατά κεφαλήν ΑΕΠ στις χώρες του ΟΟΣΑ και ως Y το κατά κεφαλήν ΑΕΠ στην Ελλάδα, παρατηρούμε ότι φαίνεται να υπάρχει μία θετική γραμμική συσχέτιση ανάμεσα στις δύο μεταβλητές:



β. Με βάση το τυπολόγιο και το φύλλο εργασίας askisi-8-erotima-b-c-d έχουμε:

Χρήση τύπων με αποκλίσεις από μέσες τιμές:

$$\beta_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\beta_0 = \bar{Y} - \beta_1 \bar{X}$$

Από τα στοιχεία μας μπορούμε να υπολογίσουμε τους αριθμητικούς μέσους των μεταβλητών X και

Υ οι οποίοι είναι:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{299,81}{15} = 19,9870$$

και

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n} = \frac{370,37}{15} = 24,6911$$

Για τη διευκόλυνση των πράξεων σχηματίζουμε τον παρακάτω πίνακα:

Πίνακας Θέματος 8

ΟΟΣΑ	ΕΛΛΑΔΑ	Υπολογισμοί Βασικών Αθροισμάτων							
X_i	Y_i	$X_i Y_i$	X_i^2	Y_i^2	$(X_i - \bar{X})$	$(Y_i - \bar{Y})$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$
15,95	21,33	340,19	254,49	454,76	-4,03	-3,37	13,58	16,28	11,33
16,21	21,83	353,95	262,92	476,48	-3,77	-2,86	10,8	14,23	8,19
16,70	22,47	375,17	278,88	504,7	-3,29	-2,23	7,32	10,81	4,95
17,17	22,93	393,6	294,76	525,59	-2,82	-1,77	4,98	7,94	3,12
17,68	23,56	416,44	312,5	554,94	-2,31	-1,13	2,62	5,33	1,29
18,41	24,36	448,48	338,94	593,43	-1,58	-0,33	0,52	2,49	0,11
19,13	24,50	468,58	365,8	600,23	-0,86	-0,19	0,17	0,74	0,04
19,72	24,74	487,74	388,72	611,97	-0,27	0,05	-0,01	0,07	0
20,82	25,05	521,57	433,47	627,58	0,83	0,36	0,3	0,69	0,13
21,65	25,69	556,28	468,9	659,94	1,67	1	1,66	2,78	1
22,06	26,21	578,31	486,81	687	2,08	1,52	3,16	4,31	2,31
22,96	26,85	616,55	527,23	720,99	2,97	2,16	6,43	8,85	4,67
23,84	27,38	652,9	568,52	749,81	3,86	2,69	10,38	14,87	7,24
24,05	27,29	656,26	578,28	744,75	4,06	2,6	10,55	16,49	6,75
23,45	26,19	614,17	549,86	685,99	3,46	1,5	5,19	11,99	2,25
299,81	370,37	7480,17	6110,09	9198,16			77,63	117,87	53,38

Κατά συνέπεια έχουμε ότι:

$$n=15$$

$$\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) = 77,63$$

$$\sum_{i=1}^n (X_i - \bar{X})^2 = 117,87$$

Επομένως η κλίση β_1 της ευθείας θα είναι:

$$\beta_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} = \frac{77,63}{117,87} = 0,658647$$

Για δε την εκτίμηση της σταθεράς β_0 έχουμε:

$$\beta_0 = \bar{Y} - \beta_1 \bar{X} = 24,6911 - (0,658647)(19,9870) = 11,526737$$

Άρα η εκτιμηθείσα εξίσωση παλινδρόμησης είναι:

$$\hat{Y} = 11,526737 - 0,658647X$$

Χρήση τύπων χωρίς αποκλίσεις από μέσες τιμές:

Ο εναλλακτικός τύπος υπολογισμού του συντελεστή β_1 που δεν απαιτεί να εκφρασθεί το Y σε αποκλίσεις από το μέσο του είναι:

$$\beta_1 = \frac{\sum_{i=1}^n X_i Y_i - n \bar{X} \bar{Y}}{\sum_{i=1}^n X_i^2 - n \bar{X}^2}$$

Από τον Πίνακα Θέματος 8 έχουμε ότι:

$$\beta_1 = \frac{7480,17 - 15 (19,9870) (24,6911)}{6110,09 - 15(19,9870)^2} = 0,658647$$

και επομένως:

$$\beta_0 = \bar{Y} - \beta_1 \bar{X} = 24,6911 - (0,658647)(19,9870) = 11,526737$$

Άρα η εκτιμηθείσα εξίσωση παλινδρόμησης είναι:

$$\hat{Y} = 11,526737 - 0,658647X$$

Παρατηρούμε ότι τα αποτελέσματα ταυτίζονται με εκείνα του πρώτου τρόπου.

Ερμηνεία συντελεστών.

- Η τιμή του συντελεστή $\beta_1=0,658647$ δηλώνει ότι μία αύξηση στο κατά κεφαλήν εισόδημα στις χώρες του ΟΟΣΑ κατά 1 χιλιάδα ευρώ (1000 ευρώ) θα οδηγήσει σε μία αύξηση στο κατά κεφαλήν εισόδημα στην Ελλάδα κατά 0,658647 χιλ. Ευρώ (658,647 ευρώ).
- Η τιμή του συντελεστή $\beta_0=11,526737$ δηλώνει ότι ακόμη κι αν δεν υπάρχει καθόλου κατά κεφαλήν εισόδημα στις χώρες του ΟΟΣΑ, στην Ελλάδα το κατά κεφαλήν εισόδημα θα είναι 11,526737 χιλ. Ευρώ (11526,737 ευρώ).

Συντελεστές με EXCEL

Οι φοιτητές θα κάνουν την δική τους εργασία σκεπτόμενοι πάνω στις ενδεικτικές απαντήσεις.

Σε περίπτωση αντιγραφής ή ύπαρξης παροραμάτων δεν φέρουμε καμία ευθύνη.

Χρησιμοποιώντας την Παλινδρόμηση από τα Εργαλεία Ανάλυση Δεδομένων του EXCEL έχουμε τα ίδια αποτελέσματα, όπως φαίνονται στο φύλλο Askisi-8-ipoerotima-b:

ΕΞΟΔΟΣ ΣΥΜΠΕΡΑΣΜΑΤΟΣ

Στατιστικά παλινδρόμησης	
Πολλαπλό R	0,978736577
R Τετράγωνο	0,957925287
Προσαρμοσμένο R Τετράγωνο	0,954688771
Τυπικό σφάλμα	0,415649904
Μέγεθος δείγματος	15

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ

	βαθμοί ελευθερίας	SS	MS	F	Σημαντικότητα F
Παλινδρόμηση	1	51,13393273	51,13393	295,9742	2,51671E-10
Υπόλοιπο	13	2,245942955	0,172765		
Σύνολο	14	53,37987568			

	Συντελεστής	Τυπικό σφάλμα	t	τιμή-P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95,0%	Υψηλότερο 95,0%
Τεταγμένη επί την αρχή	11,52673664	0,772687849	14,91771	1,48E-09	9,857446349	13,19602693	9,857446349	13,19602693
Μεταβλητή X 1	0,658647348	0,038284769	17,2039	2,52E-10	0,575938148	0,741356547	0,575938148	0,741356547

Εναλλακτικά μπορούμε να εργαστούμε χρησιμοποιώντας την εντολή LINEST() του EXCEL και να υπολογίσουμε τα ίδια αποτελέσματα με αυτά που ήδη έχουμε βρεί. Τέλος το ίδιο ακριβώς πετυχαίνουμε και με τις εντολές SLOPE() και INTERCEPT() του EXCEL. Τα εναλλακτικά αποτελέσματα παρουσιάζονται αναλυτικά στο φύλλο Askisi-8-ipoerotima-b-2 :

LINEST	
β_1	β_0
0,658647	11,526737
SLOPE	INTERCEPT
β_1	β_0
0,658647	11,526737

LINEST			
β_1	0,658647	11,526737	β_0
σ_1	0,04	0,77	σ_0
r^2	0,95793	0,42	σ_y
F	295,97	13	df
SS _{reg}	51,13	2,25	SS _{residue}

β. Με βάση το τυπολόγιο και το φύλλο εργασίας askisi-8-erotima-b-c-d υπολογίζουμε τον συντελεστή συσχέτισης:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} = \frac{77,63}{\sqrt{117,87 * 53,58}} = 0,97874$$

κι επειδή είναι αρκετά κοντά στο +1 συμπεραίνουμε ότι υπάρχει ισχυρή θετική γραμμική συσχέτιση ανάμεσα στο κατά κεφαλήν εισόδημα στην Ελλάδα (σε χιλ. ευρώ) Y και στο κατά κεφαλήν εισόδημα X στις χώρες του ΟΟΣΑ, δηλ. όταν αυξάνεται το εισόδημα στις χώρες του

Στο EXCEL χρησιμοποιούμε την συνάρτηση CORREL() και βρίσκουμε το ίδιο αποτέλεσμα.

δ. Γνωρίζουμε ότι ο συντελεστής προσδιορισμού R^2 ισούται με το τετράγωνο του συντελεστή συσχέτισης. Κατά συνέπεια, θα έχουμε ότι:

$$R^2 = (0,97874)^2 = \mathbf{0,95793}$$

Η τιμή αυτή δηλώνει ότι το 95,793% της μεταβλητότητας του κατά κεφαλήν εισοδήματος στην Ελλάδα (σε χιλ. ευρώ) Y ερμηνεύεται από το κατά κεφαλήν εισόδημα X των χωρών του ΟΟΣΑ.

Την ίδια τιμή βρίσκουμε και με την αντίστοιχη συνάρτηση RSQ() του EXCEL.

ε. Υπολογίζουμε πρώτα τις ποσοστιαίες μεταβολές κατ' έτος για τις χώρες του ΟΟΣΑ και για την Ελλάδα και κατόπιν χρησιμοποιούμε:

i)Είτε την Παλινδρόμηση από τα Εργαλεία Ανάλυση Δεδομένων του EXCEL και έχουμε τα

ΕΞΟΔΟΣ ΣΥΜΠΕΡΑΣΜΑΤΟΣ

Στατιστικά παλινδρόμησης	
Πολλαπλό R	0,748903918
R Τετράγωνο	0,560857078
Προσαρμοσμένο R Τετράγωνο	0,524261834
Τυπικό σφάλμα	1,295738909
Μέγεθος δείγματος	14

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ

	βαθμοί ελευθερίας	SS	MS	F	Σημαντικότητα F
Παλινδρόμηση	1	25,73134948	25,73135	15,32596	0,0020542
Υπόλοιπο	12	20,14727183	1,678939		
Σύνολο	13	45,87862131			

	Συντελεστής	Τυπικό σφάλμα	t	τιμή-P	Κατώτερο 95%	Υψηλότερο 95%	Κατώτερο 95,0%	Υψηλότερο 95,0%
Τεταγμένη επί την αρχή	-0,537997769	0,624296457	-0,861766	0,405708	-1,898222877	0,822227338	-1,898222877	0,822227338
Μεταβλητή X 1	0,724453486	0,185053253	3,914838	0,002054	0,32125709	1,127649881	0,32125709	1,127649881

αποτελέσματα που φαίνονται στο φύλλο Askisi-8-erotima-exodos:

ii) Είτε χρησιμοποιούμε την συνάρτηση LINEST() του EXCEL κι έχουμε τα ακόλουθα αποτελέσματα:

$$\Delta Y = -0,537998 + 0,724453 \Delta X$$

Έτος	ΟΟΣΑ	ΕΛΛΑΔΑ	ΟΟΣΑ	ΕΛΛΑΔΑ
	X_i	Y_i	$\Delta X_i (\%)$	$\Delta Y_i (\%)$
1995	15,95	21,33		
1996	16,21	21,83	1,64	2,36
1997	16,70	22,47	2,99	2,92
1998	17,17	22,93	2,81	2,05
1999	17,68	23,56	2,97	2,75
2000	18,41	24,36	4,14	3,41
2001	19,13	24,50	3,89	0,57
2002	19,72	24,74	3,09	0,97
2003	20,82	25,05	5,60	1,27
2004	21,65	25,69	4,01	2,55
2005	22,06	26,21	1,89	2,03
2006	22,96	26,85	4,07	2,44
2007	23,84	27,38	3,84	1,98
2008	24,05	27,29	0,85	-0,34
2009	23,45	26,19	-2,49	-4,03

LINEST			
β_1	0,724453	-0,537998	β_0
σ_1	0,19	0,62	σ_0
r^2	0,560857	1,3	σ_y
F	15,33	12	df
SS_{reg}	25,73	20,15	$SS_{residue}$

Ερμηνεία συντελεστών.

- Η τιμή του συντελεστή $\beta_1=0,724453$ δηλώνει ότι μία ποσοστιαία αύξηση στο κατά κεφαλήν εισόδημα στις χώρες του ΟΟΣΑ κατά 1 % θα οδηγήσει σε μία αύξηση στο κατά κεφαλήν εισόδημα στην Ελλάδα κατά 0,724453 %.
- Η τιμή του συντελεστή $\beta_0 = -0,537998$ δηλώνει ότι ακόμη κι αν δεν υπάρχει καθόλου ποσοστιαία μεταβολή στο κατά κεφαλήν εισόδημα στις χώρες του ΟΟΣΑ, στην Ελλάδα το κατά κεφαλήν εισόδημα θα μειωθεί κατά -0,537998 %.

Πρόβλεψη κατά κεφαλήν εισοδήματος στην Ελλάδα για τα έτη 2010 και 2011.

Από την σχέση που διέπει τις ποσοστιαίες μεταβολές:

$$\Delta Y = -0,537998 + 0,724453 \Delta X$$

Οι φοιτητές θα κάνουν την δική τους εργασία σκεπτόμενοι πάνω στις ενδεικτικές απαντήσεις.

Σε περίπτωση αντιγραφής ή ύπαρξης παροραμάτων δεν φέρουμε καμία ευθύνη.

βλέπουμε ότι για ποσοστιαία μεταβολή $\Delta X = -2,49$ (%) (έτος 2009 και 2010) του κατά κεφαλήν εισοδήματος στις χώρες του ΟΟΣΑ η αναμενόμενη αντίστοιχη ποσοστιαία μεταβολή για την Ελλάδα το 2010 και 2011 θα είναι:

$$\Delta Y = -0,537998 + 0,724453 (-2,49) = -2,34188597 \quad (\%)$$

Επομένως το εισόδημα το 2010 θα είναι:

$$Y_{2010} = [(100 + \Delta Y)/100] Y_{2009} = [(100 - 2,34188597)/100] (26,19) =$$

$$= 25,57666 \text{ χιλ. Ευρώ} \Rightarrow Y_{2010} = \mathbf{25576,66 \text{ Ευρώ}}$$

Με βάση αυτή την εκτίμηση το εισόδημα το 2011 θα είναι:

$$Y_{2011} = [(100 + \Delta Y)/100] Y_{2010} = [(100 - 2,34188597)/100] (25,57666) =$$

$$= 24,97768 \text{ χιλ. Ευρ.} \Rightarrow Y_{2011} = \mathbf{24977,68 \text{ Ευρώ.}}$$