

Οικονομικό Πανεπιστήμιο Αθηνών

Μάθημα: *“Εφαρμογές Στατιστικών Μεθόδων σε Επιχειρησιακά Προβλήματα”*

Συνάντηση: 3η

ΑΣΚΗΣΕΙΣ ΕΜΠΕΔΩΣΗΣ ΣΤΗΝ ΥΛΗ ΤΩΝ 2 ΠΡΩΤΩΝ ΣΥΝΑΝΤΗΣΕΩΝ

Ψ_i	X_i	$X_i - X_{\text{mean}}$	$\Psi_i - \Psi_{\text{mean}}$	$(X_i - X_{\text{m}}) * (\Psi_i - \Psi_{\text{m}})$
Ψ_1	X_1	$(X_1 - X_{\text{mean}})$	$(\Psi_1 - \Psi_{\text{mean}})$	$(X_1 - X_{\text{m}}) * (\Psi_1 - \Psi_{\text{m}})$
Ψ_2	X_2	$(X_2 - X_{\text{mean}})$	$(\Psi_2 - \Psi_{\text{mean}})$	$(X_2 - X_{\text{m}}) * (\Psi_2 - \Psi_{\text{m}})$
Ψ_3	X_3	$(X_3 - X_{\text{mean}})$	$(\Psi_3 - \Psi_{\text{mean}})$	$(X_3 - X_{\text{m}}) * (\Psi_3 - \Psi_{\text{m}})$
Ψ_4	X_4	$(X_4 - X_{\text{mean}})$	$(\Psi_4 - \Psi_{\text{mean}})$	$(X_4 - X_{\text{m}}) * (\Psi_4 - \Psi_{\text{m}})$
Ψ_5	X_5	$(X_5 - X_{\text{mean}})$	$(\Psi_5 - \Psi_{\text{mean}})$	$(X_5 - X_{\text{m}}) * (\Psi_5 - \Psi_{\text{m}})$
...	...	...	...	...
Ψ_n	X_n	$(X_n - X_{\text{mean}})$	$(\Psi_n - \Psi_{\text{mean}})$	$(X_n - X_{\text{m}}) * (\Psi_n - \Psi_{\text{m}})$
				$\Sigma(X_i - X_{\text{m}}) * (\Psi_i - \Psi_{\text{m}})$

Ψi	Xi	Xi - X_mean	Ψi - Ψ_mean	(Xi-X_m)*(Ψi-Ψ_m)	(Xi - X_m)^2	(Ψi - Ψ_m)^2
Ψ1	X1	(X1-X_mean)	(Ψ1-Ψ_mean)	(X1-X_m)*(Ψ1-Ψ_m)	(X1 - X_m)^2	(Ψ1 - Ψ_m)^2
Ψ2	X2	(X2-X_mean)	(Ψ2-Ψ_mean)	(X2-X_m)*(Ψ2-Ψ_m)	(X2 - X_m)^2	(Ψ2 - Ψ_m)^2
Ψ3	X3	(X3-X_mean)	(Ψ3-Ψ_mean)	(X3-X_m)*(Ψ3-Ψ_m)	(X3 - X_m)^2	(Ψ3 - Ψ_m)^2
Ψ4	X4	(X4-X_mean)	(Ψ4-Ψ_mean)	(X4-X_m)*(Ψ4-Ψ_m)	(X4 - X_m)^2	(Ψ4 - Ψ_m)^2
Ψ5	X5	(X5-X_mean)	(Ψ5-Ψ_mean)	(X5-X_m)*(Ψ5-Ψ_m)	(X5 - X_m)^2	(Ψ5 - Ψ_m)^2
...	...	...	...	...	...	...
Ψn	Xn	(Xn-X_mean)	(Ψn-Ψ_mean)	(Xn-X_m)*(Ψn-Ψ_m)	(Xn - X_m)^2	(Ψn - Ψ_m)^2
				Σ(Xi-X_m)*(Ψi-Ψ_m)	Σ(Xi - X_m)^2	Σ(Ψi - Ψ_m)^2

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Οικονομικό Πανεπιστήμιο Αθηνών

Μάθημα: “Εφαρμογές Στατιστικών Μεθόδων σε Επιχειρησιακά Προβλήματα”

Συνάντηση: 3η

Δίνονται τα πραγματικά στοιχεία που αφορούν τον αστικό πληθυσμό (%), το προσδόκιμο ζωής των γυναικών (σε έτη) και την παιδική θνησιμότητα (θάνατοι παιδιών ανά 1.000 γεννήσεις ζώντων) για ένα γκρουπ 15 ευρωπαϊκών χωρών κατά το έτος 1960.

Με μοναδικό κριτήριο την τιμή του συντελεστή γραμμικής συσχέτισης ποιο θα λέγατε πως είναι το πρόσημο της σχέσης που συνδέει:

α) τον αστικό πληθυσμό με το προσδόκιμο ζωής των γυναικών

β) τον αστικό πληθυσμό με την παιδική θνησιμότητα

Πως θα χαρακτηρίζατε την ένταση αυτών των σχέσεων;

Έτη	Χώρα	αστικός πληθυσμός (%)	προσδόκιμο ζωής γυναικών	παιδική θνητότητα (ανά 1.000 γενν.)
1960	Austria	65	72	43
1960	Belgium	92	73	34
1960	Bulgaria	37	71	54
1960	Belarus	32	73	45
1960	Switzerland	71	74	27
1960	Czechia	60	73	25
1960	Denmark	74	74	25
1960	Spain	57	72	55
1960	Finland	55	72	27
1960	France	62	73	29
1960	United Kingdom	78	74	27
1960	Greece	56	72	46
1960	Hungary	56	70	54
1960	Iceland	80	76	21
1960	Italy	59	72	52

1ο Βήμα: ΥΠΟΛΟΓΙΣΜΟΣ ΜΕΣΗΣ ΤΙΜΗΣ

αστικός πληθυσμός (μέση τιμή) = **62 %** Ψ_{mean}

προσδόκιμο ζωής των γυναικών (μέση τιμή) = **73 έτη** $X1_{\text{mean}}$

παιδική θνησιμότητα (μέση τιμή) = **37 θάνατοι** (ανά 1.000 γεννήσεις ζώντων) $X2_{\text{mean}}$

2ο Βήμα: Κατασκευή του αθροίσματος των **$(X_i - X_{\text{μέσο}}) * (\Psi_i - \Psi_{\text{μέσο}})$**

$$\Sigma(X1_i - X1_{\text{μέσο}}) * (\Psi_i - \Psi_{\text{μέσο}}) = \mathbf{172,02}$$

$$\Sigma(X2_i - X2_{\text{μέσο}}) * (\Psi_i - \Psi_{\text{μέσο}}) = \mathbf{-1.631,6}$$

3ο Βήμα: ΥΠΟΛΟΓΙΣΜΟΣ ΤΥΠΙΚΗΣ ΑΠΟΚΛΙΣΗΣ

αστικός πληθυσμός (τυπική απόκλιση) = 15,61 S_{ψ}

προσδόκιμο ζωής γυναικών (τυπική απόκλιση) = 1,39 $S_{\chi1}$

παιδική θνησιμότητα (τυπική απόκλιση) = 12,66 $S_{\chi2}$

4ο Βήμα: ΥΠΟΛΟΓΙΣΜΟΣ ΣΥΝΤΕΛΕΣΤΗ ΣΥΣΧΕΤΙΣΗΣ

$$\rho_{\chi1,\psi} = 172,02 / [14 * (15,61 * 1,39)] = + 0,56$$

$$\rho_{\chi2,\psi} = -1.631,6 / [14 * (15,61 * 12,66)] = - 0,59$$

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$r_{xy} \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y}$$

Οικονομικό Πανεπιστήμιο Αθηνών

Μάθημα: “Εφαρμογές Στατιστικών Μεθόδων σε Επιχειρησιακά Προβλήματα”

Συνάντηση: 3η

ΔΙΑΓΡΑΜΜΑ ΔΙΑΣΠΟΡΑΣ

Εργαστείτε στο ψηφιακό περιβάλλον της εφαρμογής

<https://www.statskingdom.com/index.html>

Από τα “βασικά στατιστικά” (**Basic stats**) επιλέξτε τη “συσχέτιση” (**correlation**) !!!

Εισάγετε τις τιμές των μεταβλητών σας στις 2 στήλες !!!

Όπου Ψ οι τιμές της **εξαρτημένης** μεταβλητής, X οι τιμές της **ανεξάρτητης** μεταβλητής

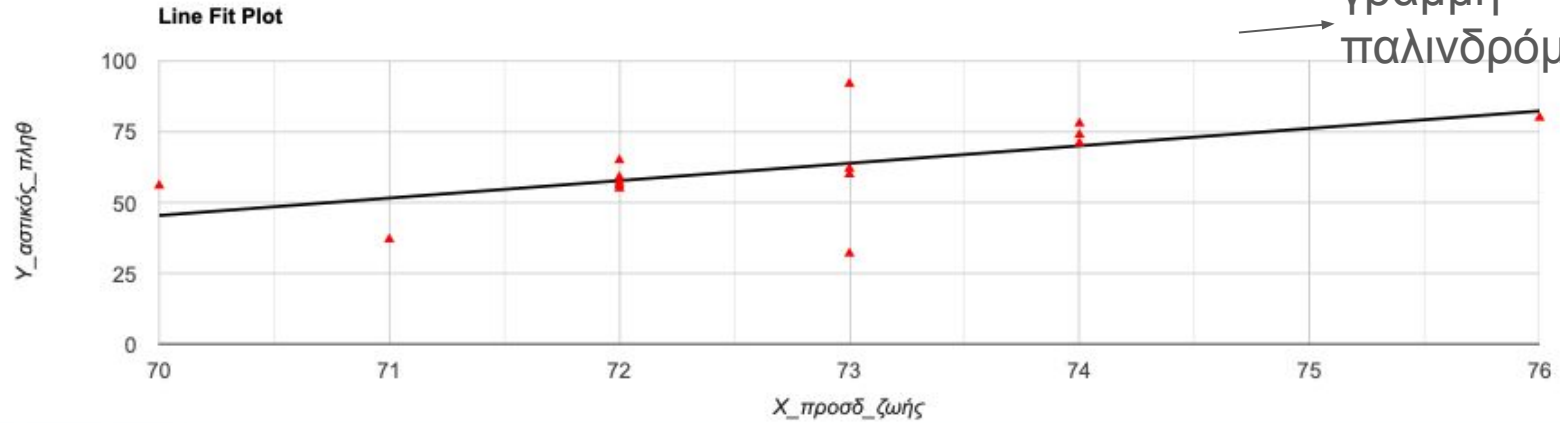
ΠΡΟΣΟΧΗ: Αν έχετε κόμματα (,) αντικαταστήστε με τελείες (.) ...



$r = 0.568$

τιμή συντελεστή
γραμμικής συσχέτισης

διάγραμμα
διασποράς και
γραμμή
παλινδρόμησης



Parameter	Value
Pearson correlation coefficient (r)	0.568
r^2	0.3226
P-value	0.02718
Covariance	12.719
Sample size (n)	15
Statistic	2.4882

→ τιμή συντελεστή συσχέτισης

→ τιμή συντελεστή προσδιορισμού
($R^2 = r^2$)

→ αριθμός παρατηρήσεων

$H_0: \rho = 0$ (ΔΕΝ υπάρχει γραμμική
συσχέτιση)

$H_1: \rho > 0$ ή $\rho < 0$
(υπάρχει γραμμική συσχέτιση)

Τρέχουμε το ίδιο παράδειγμα. Αυτή τη φορά αντι για χώρες της Ευρώπης χρησιμοποιούμε χώρες του λεγόμενου τρίτου κόσμου !!!

Δίνονται τα πραγματικά στοιχεία που αφορούν τον αστικό πληθυσμό (%), το προσδόκιμο ζωής των γυναικών (σε έτη) και την παιδική θνησιμότητα (θάνατοι παιδιών ανά 1.000 γεννήσεις ζώντων) για ένα γκρουπ 15 ευρωπαϊκών χωρών κατά το έτος 1960.

Με μοναδικό κριτήριο την τιμή του συντελεστή γραμμικής συσχέτισης ποιο θα λέγατε πως είναι το πρόσημο της σχέσης που συνδέει:

α) τον αστικό πληθυσμό με το προσδόκιμο ζωής των γυναικών

β) τον αστικό πληθυσμό με την παιδική θνησιμότητα

Πως θα χαρακτηρίζατε την ένταση αυτών των σχέσεων;

Έτη	Χώρα	αστικός πληθυσμός (%)	προσδόκιμο ζωής γυναικών	παιδική θνητότητα (ανά 1.000 γενν.)
1960	Afghanistan	8	33	355
1960	United Arab Emirates	74	52	201
1960	Burkina Faso	5	37	343
1960	Bangladesh	5	44	263
1960	Bahrain	82	51	198
1960	Bolivia	37	45	291
1960	Brazil	46	55	169
1960	Botswana	3	53	167
1960	Central African Republic	20	40	279
1960	Chile	68	59	153
1960	Cote d'Ivoire	18	40	319
1960	Cameroon	14	42	280
1960	Congo, Rep.	32	51	194
1960	Colombia	46	59	135
1960	Guyana	29	61	93

1ο Βήμα: ΥΠΟΛΟΓΙΣΜΟΣ ΜΕΣΗΣ ΤΙΜΗΣ

αστικός πληθυσμός (μέση τιμή) = **32 %** Ψ_{mean}

προσδόκιμο ζωής των γυναικών (μέση τιμή) = **48 έτη** $X1_{\text{mean}}$

παιδική θνησιμότητα (μέση τιμή) = **229 θάνατοι** (ανά 1.000 γεννήσεις ζώντων) $X2_{\text{mean}}$

2ο Βήμα: Κατασκευή του αθροίσματος των **$(X_i - X_{\text{μέσο}}) * (\Psi_i - \Psi_{\text{μέσο}})$**

$$\Sigma(X1_i - X1_{\text{μέσο}}) * (\Psi_i - \Psi_{\text{μέσο}}) = \mathbf{1.880}$$

$$\Sigma(X2_i - X2_{\text{μέσο}}) * (\Psi_i - \Psi_{\text{μέσο}}) = \mathbf{-14.934}$$

3ο Βήμα: ΥΠΟΛΟΓΙΣΜΟΣ ΤΥΠΙΚΗΣ ΑΠΟΚΛΙΣΗΣ

αστικός πληθυσμός (τυπική απόκλιση) = 26,06 S_{ψ}

προσδόκιμο ζωής γυναικών (τυπική απόκλιση) = 8,63 $S_{\chi1}$

παιδική θνησιμότητα (τυπική απόκλιση) = 80,45 $S_{\chi2}$

4ο Βήμα: ΥΠΟΛΟΓΙΣΜΟΣ ΣΥΝΤΕΛΕΣΤΗ ΣΥΣΧΕΤΙΣΗΣ

$$\rho_{\chi1,\psi} = 1.880 / [14 * (26,06 * 8,63)] = + 0,60$$

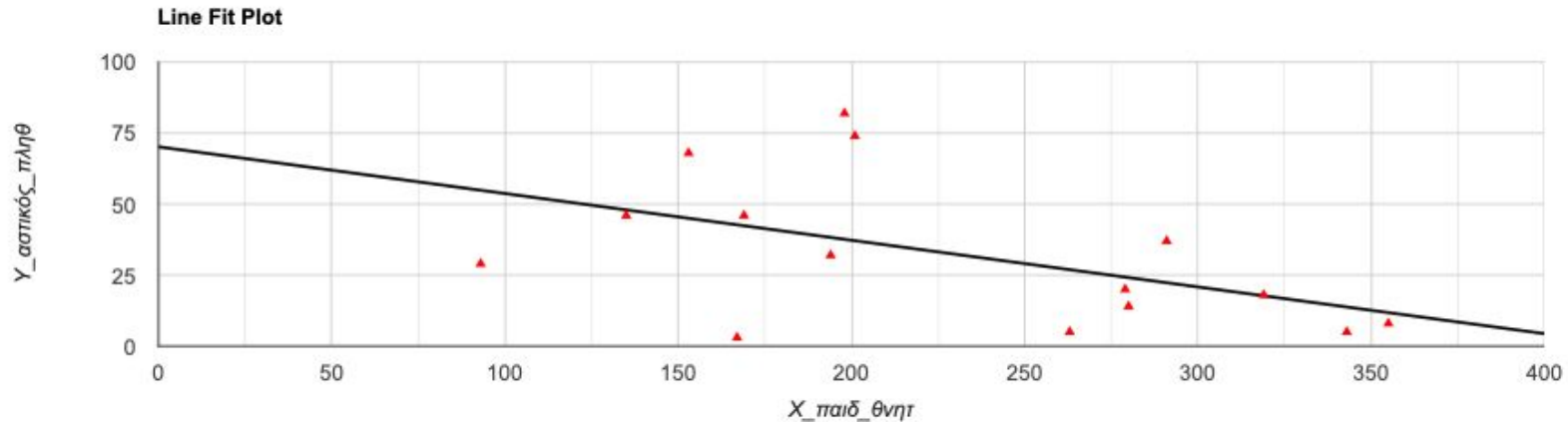
$$\rho_{\chi2,\psi} = -14.934 / [14 * (26,06 * 80,45)] = - 0,51$$

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$r_{xy} \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y}$$



↓ $r = -0.5074$



Parameter	Value
Pearson correlation coefficient (r)	-0.5074
r^2	0.2575
P-value	0.05352
Covariance	-1064.4524
Sample size (n)	15
Statistic	-2.123

$H_0: \rho = 0$ (ΔΕΝ υπάρχει γραμμική
συσχέτιση)

$H_1: \rho > 0$ ή $\rho < 0$
(υπάρχει γραμμική συσχέτιση)

Η τιμή του p-value είναι μεγαλύτερη από το 0,05 (έστω και οριακά).

Αυτό σημαίνει πως δε μπορούμε να απορρίψουμε την μηδενική υπόθεση.

Ωστόσο, $p_value < 0,1$ (10%) → απορρίπτουμε την H_0 στο Δ.Β του 90%