

Εργαστηριακές Ασκήσεις στη Ανάλυση Δεδομένων

Γραμμική παλινδρόμηση Ανάλυση Διακύμανσης

Δεδομένα Χι	Δεδομένα Υι	Χι ^2	Χι*Υι	Υεκτιμ	(Υι-Υμεσο)^2	(Υι-Υεκτιμ)^2	(Υεκτιμ-Υμεσο)^2			
1,1	1	1,21	1,1	1,251710655	2,3104	0,063358254	1,608557863			
1,7	2	2,89	3,4	1,907722385	0,2704	0,008515158	0,374883878			
3,8	4	14,44	15,2	4,203763441	2,1904	0,04151954	2,835059325			
2,5	3,1	6,25	7,75	2,782404692	0,3364	0,10086678	0,068856222			
2,2	2,5	4,84	5,5	2,454398827	0,0004	0,002079467	0,004303514			
5	2,52				5,108	0,216339198	4,891660802			
Πλήθος μετρήσεων ΣΥι = 2,52					Σ(Υι-Υμεσο)^2=0,216339198		Σ(Υεκτιμ-Υμεσο)^2=4,89166			
							67,83321058			
ΕΞΟΔΟΣ ΣΥΜΠΕΡΑΣΜΑΤΟΣ					F = 67,83321058					
Στατιστικά παλινδρόμησης										
Πολλαπλό R	0,978594393									
R Τετράγωνο	0,957646985									
Προσαρμοσμένο R	0,943529314									
Τυπικό σφάλμα	0,268538761									
Μέγεθος δείγματος	5									
ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ										
	<i>βαθμοί ελευθερί</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Σημαντικότητα F</i>					
Παλινδρόμηση	1	4,891661	4,891661	67,83321058	0,003747363					
Υπόλοιπο	3	0,216339	0,072113							
Σύνολο	4	5,108								
	<i>Συντελεστής</i>	<i>πικό σφάλ</i>	<i>t</i>	<i>τιμή-P</i>	<i>Κατώτερο 95%</i>	<i>Υψηλότερο 95%</i>	<i>τώτερο 95, λότερο 95,0%</i>			
Τεταγμένη επί την	0,049022483	0,323162	0,151696	0,889053518	-0,97942253	1,077467496	-0,97942	1,077467		
Μεταβλητή Χ 1	1,093352884	0,132751	8,236092	0,003747363	0,670878624	1,515827143	0,670879	1,515827		
	$Y' = 0,049 + 1,093352 * X$									

Στον παραπάνω πρώτο πίνακα βρίσκουμε το F της ANOVA (Ανάλυσης Διακύμανσης της Παλινδρόμησης) με τη δημιουργία του πρώτου πίνακα και στη συνέχεια με τη βοήθεια του Excel.

1. Περιγράψτε πως δημιουργούμε τον πρώτο πίνακα
2. Τι είναι το Πολλαπλό R; Είναι αξιόπιστο το μοντέλλο;
3. Στον πίνακα ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ που μας δίνει το Excel πως βρίσκουμε τα SS πως τα MS και πως το F. Το F είναι σημαντικό; Οι μεταβλητές x και y είναι συσχετισμένες; Γιατί;
4. Δώστε ένα διάστημα εμπιστοσύνης για τους συντελεστές του γραμμικού μοντέλου. Ποια τιμή παίρνει ο συντελεστής της μεταβλητής x και ποια ο σταθερός όρος;

Διαστήματα εμπιστοσύνης και στατιστικός έλεγχος

Με χρήση του λογισμικού PSPP να εφαρμοστεί ένα μονοδειγματικό T-test στην παρακάτω μεταβλητή για διάφορες τιμές της εκτιμώμενης μέσης τιμής. Να αναλυθούν τα εξαγόμενα αποτελέσματα

3,00000
3,10000
3,00000
2,50000
3,20000
2,20000
3,20000
2,60000
3,50000
3,30000
4,20000
3,10000
3,20000
2,80000
3,70000
3,00000
3,10000
2,90000
3,70000
3,30000

Αποτελέσματα με μέση τιμή 3.0

Εξαγωγή — PSPPIRE Απεικ. Εξαγωγής

Archivio Διόρθωση Παράθυρα Βοήθεια

T-TEST

T-TEST /TESTVAL=3.0
 /VARIABLES= Μετ0001 /MISSING=ANALYSIS
 /CRITERIA=CI(0.95).

Στατιστικά Μονο-Δειγματικού Τεστ

	N	Μέσος Όρος	Τυπ. Απόκλιση	T.Σ. ΜΟ
Μετ0001	20	3,13	,44	,10

Μονο-Δειγματικό Τεστ

	Τιμή Τεστ = 3,000000					
	t	ΒΕ	Στ. Σημ. (2-κατ/νσης)	Διαφορές Μέσου	95% Διάστημα Εμπιστοσύνης για Διαφορά	
					Κατώτερο	Ανώτερο
Μετ0001	1,31	19	,207	,13	-,08	,34

Εκθετική Παλινδρόμηση Παράδειγμα του Baxter, σελ. 37-38 στις Σημειώσεις

Σύμφωνα με την παρακάτω εικόνα,

1. Μεταξύ ποιων μεταβλητών πραγματοποιείται παλινδρόμηση;
2. Είναι αξιόπιστο το μοντέλο και γιατί
3. Πως βρίσκουμε το predict(ln) και πως το predict(build_ind)

Αυτόματη αποθήκευση Baxter 03122020.ods Αναζήτηση (Alt+X) Konstantinos Perakis

Αρχείο Κεντρική Εισαγωγή Διάταξη σελίδας Τύποι Δεδομένα Αναθεώρηση Προβολή Βοήθεια Σχόλια Κοινή χρήση

Αναίρεση Επικόλληση Προχέρο Γραμματοσειρά Στοιχισμός Αριθμός Στυλ Κελιά Επεξεργασία

A1 km_cc

km_cc	build_ind	ln(build_ind)	predict(ln)	predict(build_ind)	
0,5	0,79	-0,235722334	-0,33307	0,716642554	ΕΞΟΔΟΣ ΣΥΜΠΕΡΑΣΜΑΤΟΣ
1	0,505	-0,68319685	-0,59114	0,553635449	
1,5	0,39	-0,94160854	-0,84921	0,427705847	Στατιστικά παλινδρόμησης
2	0,32	-1,139434283	-1,10728	0,33042012	Πολλαπλά 0,97438
2,5	0,26	-1,347073648	-1,36535	0,255262948	R Τετράγωνο 0,949417
3	0,157	-1,851509474	-1,62342	0,197200984	Προσαρμογή 0,945804
3,5	0,156	-1,857899272	-1,88149	0,152345761	Τυπικό σφ 0,293555
4	0,098	-2,3227878	-2,13956	0,117693282	Μέγεθος 16
4,5	0,08	-2,525728644	-2,39763	0,090922835	
5	0,076	-2,577021939	-2,6557	0,070241579	ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ
5,5	0,082	-2,501036032	-2,91377	0,05426447	βαθμοί ελευθερίας SS MS F ιαντικότητα F
6	0,074	-2,603690186	-3,17184	0,041921505	Παλινδρόμο 1 22,64417 22,64417 262,7713 1,82E-10
6,5	0,05	-2,995732274	-3,42991	0,032386063	Υπόλοιπο 14 1,206442 0,086174
7	0,025	-3,688879454	-3,68798	0,025019548	Σύνολο 15 23,85062
7,5	0,012	-4,422848629	-3,94605	0,019328616	
8	0,01	-4,605170186	-4,20412	0,01493214	Συντελεστές σφάλματος t τιμή-P
					ιπώτερο 95% ιχίτερο 95% ιχίτερο 95,0%
					Τεταγμένη -0,07511 0,153941 -0,4879 0,633176 -0,40528 0,255064 -0,40528 0,255064
					Μεταβλητή -0,51614 0,03184 -16,2102 1,82E-10 -0,58443 -0,44785 -0,58443 -0,44785
		0,927643905	είναι το e ^λ		

Εφαρμογή της Εκθετικής Παλινδρόμησης

Ενεργοποιήστε τα Windows Μετάβαση στις ρυθμίσεις για ενεργοποίηση των Windows

Φύλλο1

Ετοιμο Προσβασιμότητα: Διερεύνηση

3:51 μμ 23/1/2022

Cluster Analysis (k-means)

σελ. 67-71 στις Σημειώσεις με χρήση του λογισμικού SPSS