

2021-01-19

Ερω μιντεο

$$Y = b_0 + b_1 X + \varepsilon$$

Δειγμα

$$\begin{array}{c|c} X & Y \\ \hline x_1 & y_1 \\ \vdots & \vdots \\ x_n & y_n \end{array} \Rightarrow \hat{b}_0, \hat{b}_1 \quad \text{LSE}$$

$$\Rightarrow \hat{y}_j = \hat{b}_0 + \hat{b}_1 x_j$$

$$\text{κατανοιο-}j = e_j = y_j - \hat{y}_j$$

$$\text{jackknife residual } \tilde{r}_j = \frac{e_j}{S_{(-j)} \sqrt{n-j}}$$

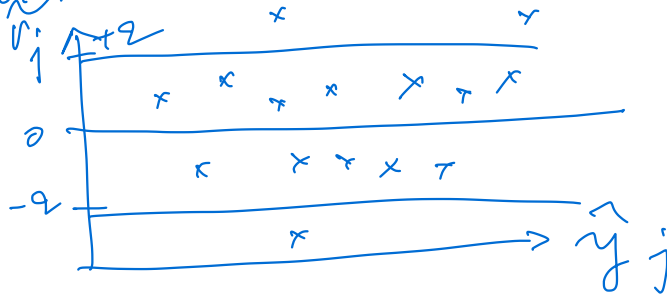
$$\text{Αν ισχύουν οι υποθέσεις } \begin{cases} \varepsilon_j \sim N(0, \sigma^2) \\ \varepsilon_j \text{ ανεξάρτητα} \\ \text{Var}(\varepsilon_j) = \sigma^2 \quad \forall j \end{cases}$$

$$\text{τότε } \tilde{r}_j \text{ αυτοσκέυσα } j=1, \dots, n$$

$$\text{Var}(\tilde{r}_j) = 1$$

$$\tilde{r}_j \sim t_{n-k-2} \quad (n \gg \Rightarrow \sim N(0,1))$$

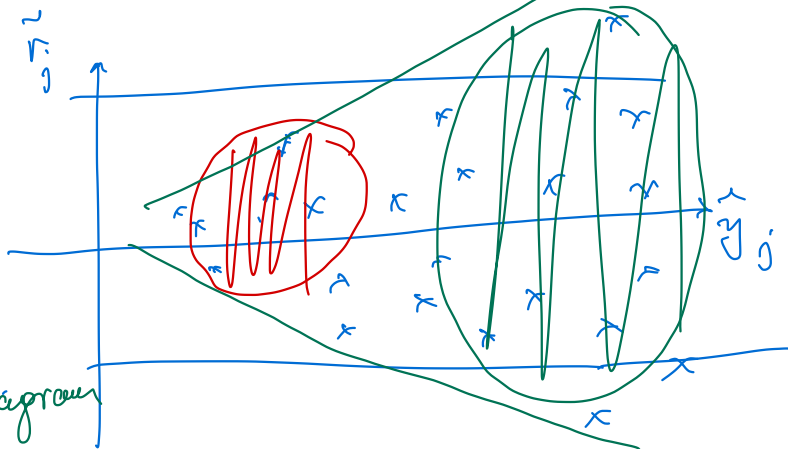
Διάγραμμα κατανομών



$$\text{Cov}(\tilde{r}_j, \hat{y}_j) = 0$$

Διαγνωστικοί Έλεγχοι Κατανομών

①



χοάνη

funnel diagram

ΕΞΕΡΘΟΚΕΔΩΣΗ

Διορθωτικά Μέτρα

Μετασχηματισμοί της Y

π.χ. $\tilde{Y} = \sqrt{Y}$

μεταβάλλει τις
ρετίες της Y

Χρειάζεται προσοχή

Εάν το μοντέλο $\tilde{Y} = \gamma_0 + \gamma_1 X + \varepsilon$

$$\Rightarrow E(\tilde{Y}) = \gamma_0 + \gamma_1 X$$

$$\Rightarrow E(\sqrt{Y}) = \gamma_0 + \gamma_1 X$$

γ_1 : μεταβολή της $E(\sqrt{Y})$ όταν η $X \rightarrow X+1$

π.χ. αν $E(\sqrt{Y})$ αυξάνει κατά $\gamma_1 = 1.2$

$$\Rightarrow E(Y) ?$$

ΔΕΝ ΙΣΧΥΕΙ

$$E(\sqrt{Y}) \neq \sqrt{E(Y)}$$

η 10 "ομαδοί" με/χωρί $\sqrt[3]{y}$, ή ακόμα $\log y$

Αν n y παίρνει f' αρμετικές κτλ

$$\begin{pmatrix} y_1 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\min(y) = y_1$$

$$\tilde{y}_j = y_j - y_1$$

Περίπτωση

Εάν n y_1

$$x_1 \quad y_1$$

$$x_2 \quad y_2$$

\vdots

$$x_n \quad y_n$$

Τροποποιημένη μέθοδος LSE

weighted LSE

$$\underbrace{\quad}_{n_1}$$

$$x_1 \quad x_1 \quad \dots \quad x_1$$

$$\underbrace{\quad}_{n_2}$$

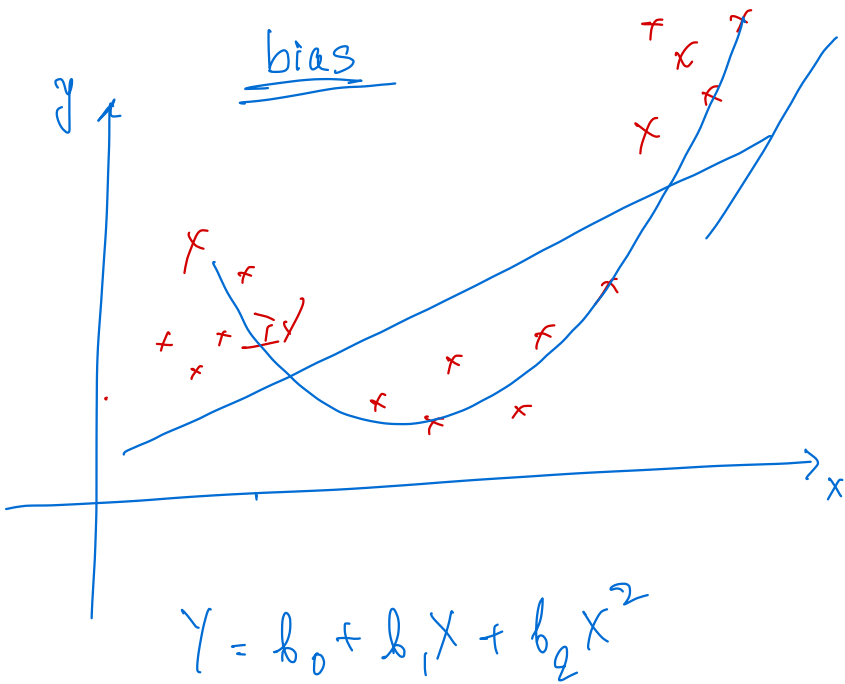
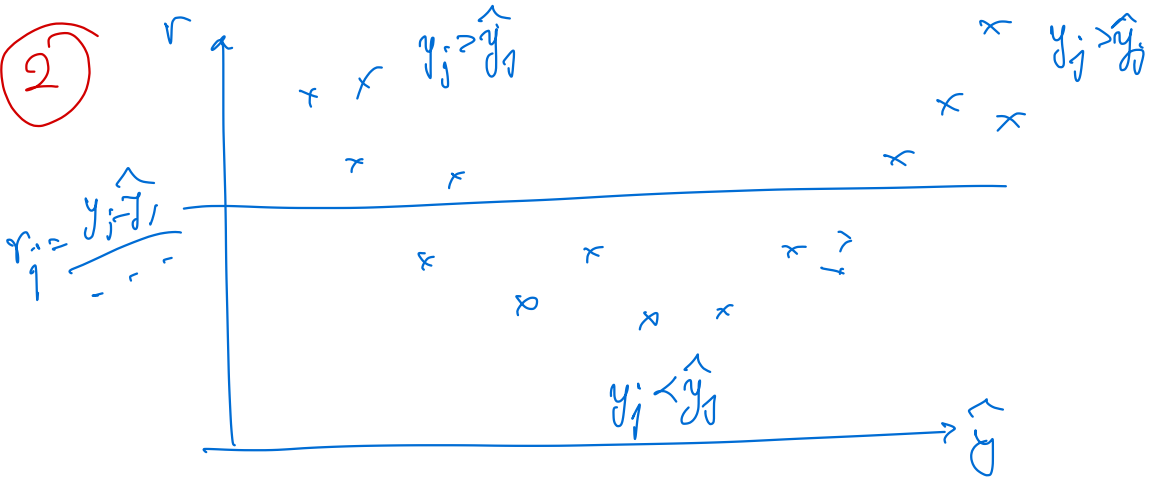
$$x_2 \quad x_2 \quad \dots \quad x_2$$

$$\underbrace{y_{11} \quad y_{12} \quad y_{1n_1}}_{y_1 = \bar{y}_{1n_1}}$$

$$y_1 = \bar{y}_{1n_1}$$

Εργασία Ορθοκεδαιμικότητας : Bartlett test

2

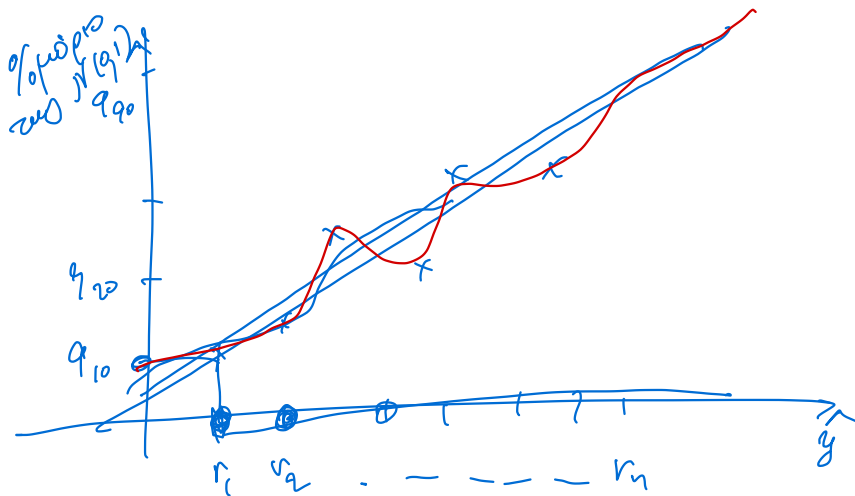


3

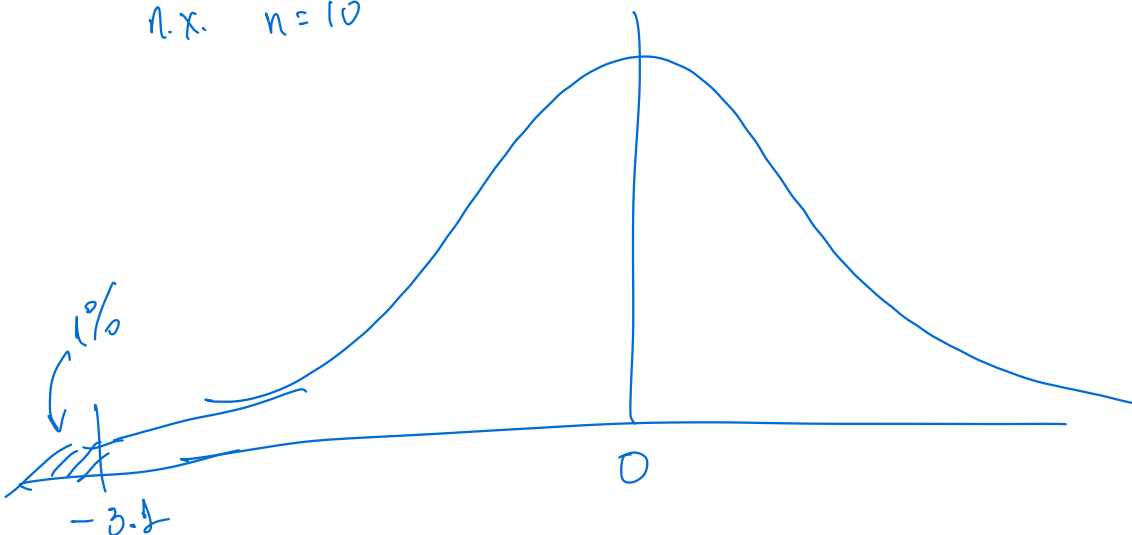
Κανονικότητα

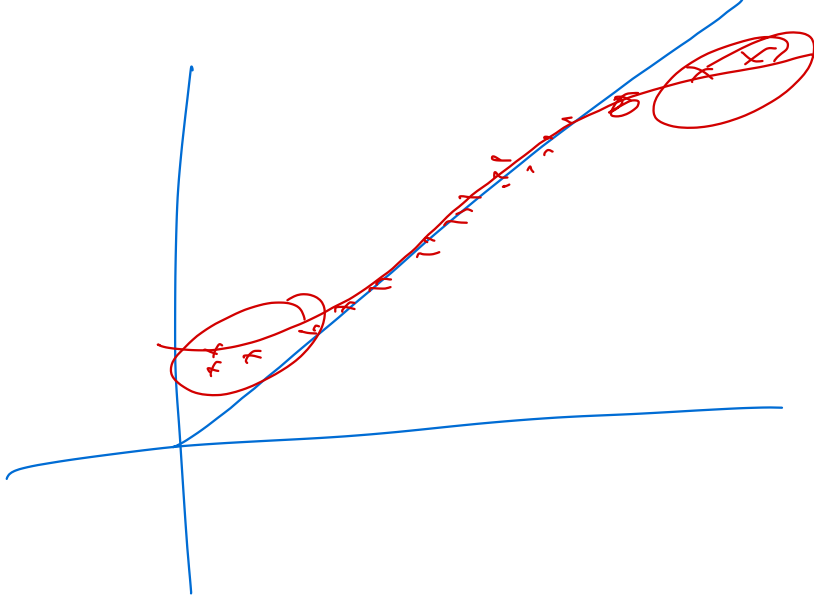
$$H_0: \tilde{r}_j \sim \mathcal{N}(0,1)$$

Γραφικός έλεγχος κανονικότητας qplot ?



π.χ. $n=10$





Ελεγχοί { Shapiro-Wilk (για μ και σ)
 KS (για K)

R : ks.test

Παράδειγμα

Κανονική

Κατανομή

$$r_j \sim \mathcal{N}(0,1)$$

Ερω

$$Y = b_0 + b_1 X + \varepsilon$$

$$\varepsilon \sim \mathcal{N}(0, \sigma^2)$$

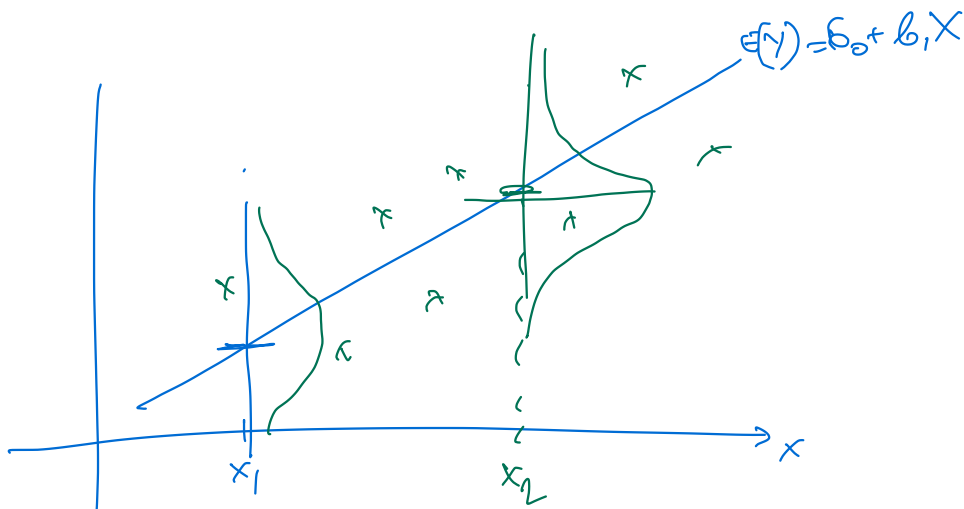
$$\Rightarrow Y_j = b_0 + b_1 x_j + \varepsilon$$

$$Y_j \sim \mathcal{N}(b_0 + b_1 x_j, \sigma^2)$$

Κάνω υποθέσεις για Y_j

και αν όλα γίνονται κανονικά $\Rightarrow X$

ΛΑΘΟΣ



4) Αυτοσυσκέυση Καταβολών

Υπόθεση: $\tilde{\gamma}_1, \tilde{\gamma}_2, \dots, \tilde{\gamma}_n$ ανεξάρτητα μεταβιβάσεις παρατηρήσεων

π.χ. παραβίαση

① Ερω π.χ. Y : αρ. ν. π.ε.μ X : ύψος
 $n = 100$ άτομα (20 οικογένειες)

Σε μια οικογένεια οι μετρήσεις Y υλοποιούνται μεταξί τους

②

x_1, x_2, \dots, x_n

: μετρήσεις σε διαδοχικές χρονικές στιγμές

y_1, y_2, \dots, y_n

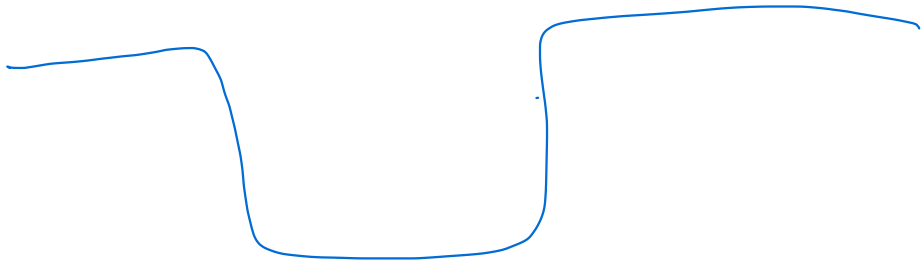
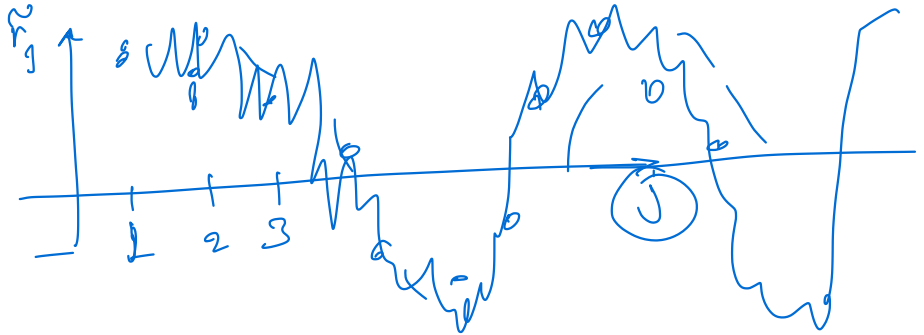
: μετρήσεις επί των

↓ ↓
 $\delta: 00\text{pm}$ $\delta: 10\text{pm}$

↳ μετρήσεις χρονολογικών

Αυτοσυσχέτιση συνδέεται μετρεσι διαδοχικών μετρήσεων σε μια χρονιά;

Έλεγχος αυτοσυσχέτισης



Έλεγχος

Durbin-Watson

d : Dubin-Watson Statistic

$$0 \leq d \leq 4$$

Όταν \mathbb{Z} αλυσούχ. $d = 2$

θετική αλυσούχ. $d \approx 4$

απλή $d \approx 0$

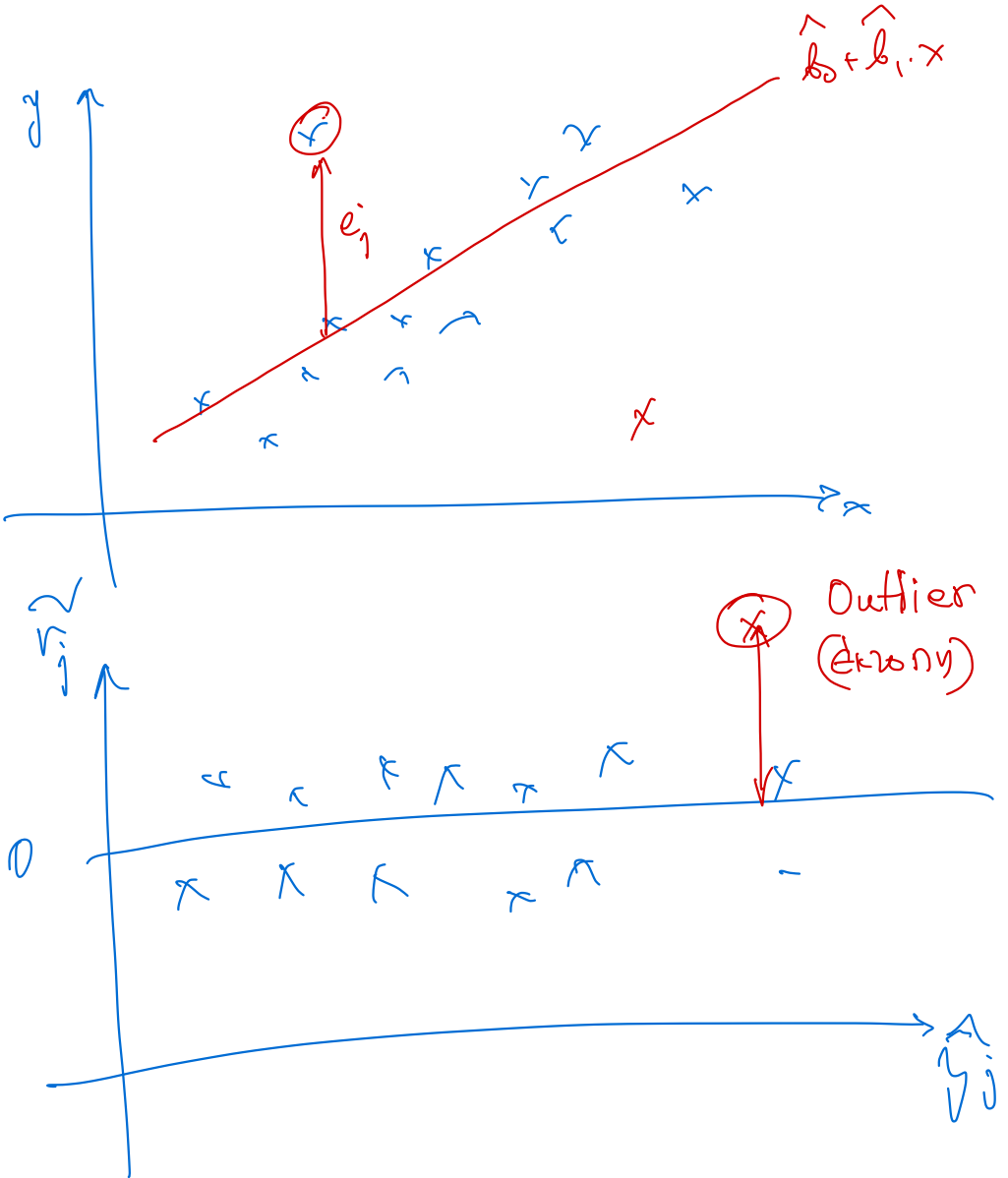
(p-value προσεγγιστικό)

Εκτροπές και Ενδραση της Παράλειψης

Outliers

Influential

Observation



Μπορεί να προέρχεται

- Άλλος καταγραφέας } Διάγραμμα
- " μέτρον }
- ?? σωστός καταγραφέας

Επίτηδες αν
"δρνίκε" οσ δείγμα

Διάγραμμα (κρίσιμα) για outliers

Προσεγγιστικά $|r| > 3$

Άκριβη κρίσιμα από νίματα

j : Outlier αν $|r_j| >$ κρίσιμα ζυγί
ωσ νίματα

Bonferroni Correction για πολλαπλούς ελέγχους

$$Y = b_0 + b_1 X_1 + \dots + b_{10} X_{10}$$

Stata

$\hat{\beta}$	$SE_{\hat{\beta}}$	$t = \frac{b_{10}}{SE_{b_{10}}}$	p-value
			< 0.05 *
			> 0.05

10 ελέγχοι

$$P(\text{κανένα λάθος σε 10 ελέγχους}) \approx (1 - 0.05)^{10}$$

$$P(\text{ζωηακίωσα ένα λάθος}) > \alpha$$

σε N ελέγχους καδίνας $\mu\epsilon P(\text{σφάλμα}) = \alpha$

$$P(\text{κανένα σφάλμα}) \approx (1 - \alpha)^N$$

Όταν $\alpha \approx 0$, $N \rightarrow \mu\epsilon\gamma\alpha\lambda\omega$

$$(1 - \alpha)^N \approx 1 - \alpha N + \frac{\alpha^2 N^2}{2}$$

$$\Rightarrow P(\text{ζωηακίωσα ένα λάθος}) \approx \underline{\underline{\alpha N}}$$

Αν θέσουμε $P(\text{ουλάχιστον ένα λάθος}) \leq \alpha$

τότε ο κάθε έλεγχος πρέπει να

δίνει σε α' : $N\alpha' = \alpha \Rightarrow \alpha' = \frac{\alpha}{N}$

Bonferroni
correctio

να τη συντηρήσει !!

Συν διαγραφή για outliers
υπερέχοντα ποσότητα έλεγχου

Επιδραστικές παραμετρώσεις

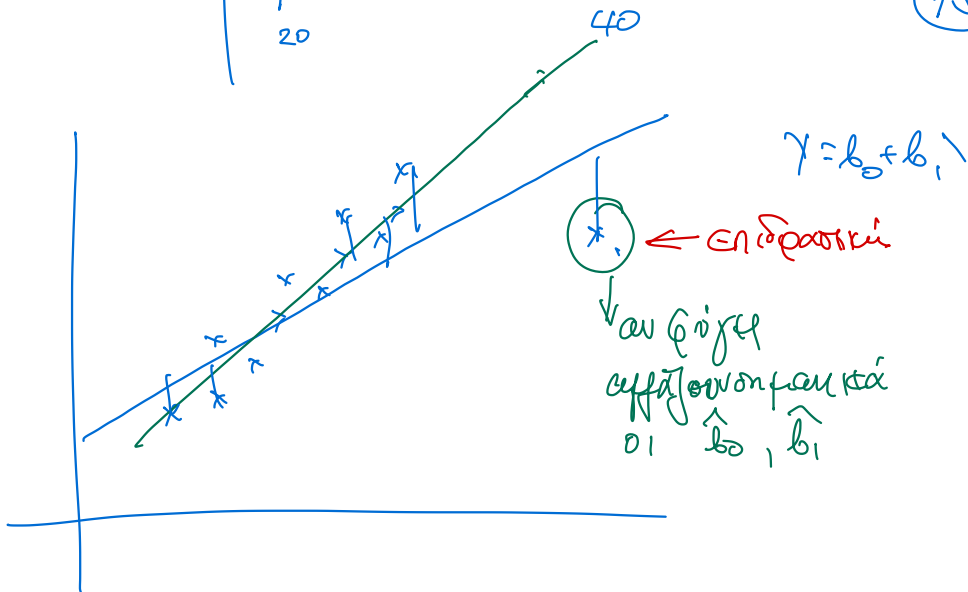
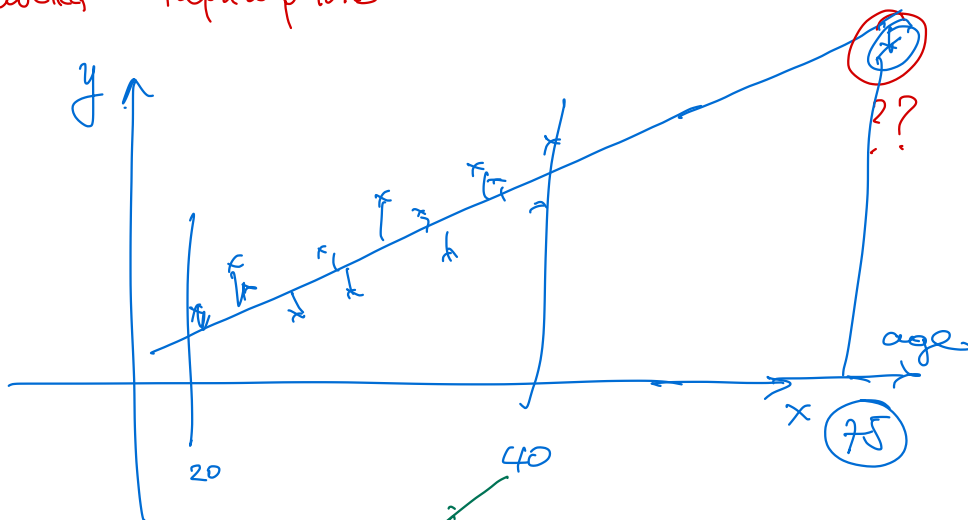


Table A-8A

Bonferroni corrected jackknife residual critical values

??

$\alpha = 0.1$

k	n=5	10	15	20	25	50	100	200	400	800
1	6.96	3.50	3.27	3.22	3.21	3.27	3.39	3.54	3.70	3.86
2	31.82	3.71	3.33	3.25	3.23	3.28	3.39	3.54	3.70	3.86
3		4.03	3.41	3.29	3.25	3.28	3.40	3.54	3.70	3.86
4		4.60	3.51	3.33	3.27	3.29	3.40	3.54	3.70	3.86
5		5.84	3.63	3.37	3.30	3.29	3.40	3.54	3.70	3.86
6		9.92	3.81	3.43	3.33	3.30	3.40	3.54	3.70	3.86
7		63.66	4.06	3.50	3.36	3.30	3.40	3.54	3.70	3.86
8			4.46	3.58	3.39	3.31	3.40	3.54	3.70	3.86
9			5.17	3.69	3.44	3.31	3.40	3.54	3.70	3.86
10			6.74	3.83	3.49	3.32	3.40	3.54	3.70	3.86
15				7.45	3.99	3.36	3.41	3.54	3.70	3.86
20					8.05	3.41	3.42	3.55	3.70	3.86
40						4.50	3.47	3.55	3.70	3.86
80							3.92	3.58	3.70	3.86

$\alpha = 0.05$

Max. Degrees (n)

k	5	10	15	20	25	50	100	200	400	800
1	9.92	4.03	3.65	3.54	3.50	3.51	3.60	3.73	3.87	4.02
2	63.66	4.32	3.73	3.58	3.53	3.51	3.60	3.73	3.87	4.02
3		4.77	3.83	3.62	3.55	3.52	3.60	3.73	3.87	4.02
4		5.60	3.95	3.67	3.58	3.53	3.61	3.73	3.87	4.02
5		7.45	4.12	3.73	3.61	3.53	3.61	3.73	3.87	4.02
6		14.09	4.36	3.81	3.65	3.54	3.61	3.73	3.88	4.02
7		127.32	4.70	3.89	3.69	3.54	3.61	3.73	3.88	4.02
8			5.25	4.00	3.73	3.55	3.61	3.73	3.88	4.02
9			6.25	4.15	3.79	3.56	3.61	3.73	3.88	4.02
10			8.58	4.33	3.85	3.57	3.61	3.73	3.88	4.02
15				9.46	4.50	3.61	3.62	3.74	3.88	4.03
20					10.21	3.67	3.63	3.74	3.88	4.03
40						5.04	3.69	3.75	3.88	4.03
80							4.23	3.78	3.88	4.03

$\alpha = 0.01$

	5	10	15	20	25	50	100	200	400	800
1	22.33	5.41	4.55	4.29	4.17	4.03	4.06	4.15	4.27	4.40
2	318.31	5.96	4.68	4.35	4.20	4.04	4.06	4.15	4.27	4.40
3		6.87	4.85	4.42	4.24	4.05	4.06	4.15	4.27	4.40
4		8.61	5.08	4.50	4.28	4.06	4.06	4.15	4.27	4.40
5		12.92	5.37	4.60	4.33	4.07	4.07	4.15	4.27	4.40
6		31.60	5.80	4.72	4.39	4.07	4.07	4.15	4.27	4.40
7		636.62	6.43	4.86	4.45	4.08	4.07	4.15	4.27	4.40
8			7.50	5.05	4.53	4.09	4.07	4.15	4.27	4.40
9			9.57	5.29	4.62	4.10	4.07	4.15	4.27	4.40
10			14.82	5.62	4.72	4.12	4.08	4.15	4.27	4.40
15				16.33	5.81	4.18	4.09	4.15	4.27	4.40
20					17.60	4.28	4.10	4.16	4.27	4.40
40						6.44	4.18	4.17	4.27	4.40
80							4.97	4.21	4.28	4.40

BA Alzaf
on jacob

Table A-9 Critical values for leverages, n = sample size, k = number of predictors

$\alpha = .10$

$k \backslash n$	1	2	3	4	5	6	7	8	9	10	15	20	40	80
10	0.626	0.759	0.847	0.911	0.956	0.984	0.997	1.000						
15	0.481	0.595	0.679	0.748	0.806	0.855	0.897	0.932	0.959	0.980				
20	0.394	0.491	0.565	0.627	0.682	0.731	0.775	0.815	0.851	0.883	0.988			
25	0.335	0.419	0.484	0.540	0.589	0.635	0.676	0.715	0.751	0.784	0.918	0.992		
30	0.293	0.366	0.424	0.474	0.519	0.560	0.599	0.635	0.669	0.701	0.837	0.937		
40	0.236	0.295	0.342	0.383	0.420	0.455	0.487	0.518	0.547	0.576	0.701	0.806		
60	0.172	0.214	0.248	0.279	0.306	0.332	0.356	0.380	0.402	0.424	0.524	0.612	0.888	
80	0.137	0.170	0.197	0.221	0.242	0.263	0.283	0.301	0.319	0.337	0.418	0.491	0.737	
100	0.114	0.141	0.164	0.183	0.201	0.219	0.235	0.250	0.266	0.280	0.348	0.410	0.625	0.941
200	0.064	0.079	0.091	0.102	0.111	0.121	0.130	0.138	0.146	0.155	0.192	0.227	0.353	0.568
400	0.036	0.043	0.050	0.055	0.060	0.065	0.070	0.075	0.079	0.083	0.104	0.122	0.190	0.311
800	0.020	0.024	0.027	0.030	0.032	0.035	0.037	0.040	0.042	0.044	0.055	0.065	0.100	0.164

$\alpha = .05$

$k \backslash n$	1	2	3	4	5	6	7	8	9	10	15	20	40	80
10	0.683	0.802	0.879	0.933	0.969	0.990	0.999	1.000						
15	0.531	0.639	0.719	0.782	0.835	0.880	0.916	0.946	0.969	0.986				
20	0.436	0.531	0.602	0.662	0.714	0.761	0.802	0.839	0.872	0.901	0.991			
25	0.372	0.454	0.518	0.573	0.621	0.665	0.705	0.742	0.776	0.807	0.931	0.994		
30	0.325	0.398	0.455	0.505	0.549	0.589	0.627	0.662	0.695	0.726	0.855	0.947		
40	0.261	0.321	0.368	0.409	0.446	0.480	0.512	0.543	0.572	0.600	0.722	0.823		
60	0.190	0.233	0.268	0.298	0.326	0.352	0.376	0.400	0.422	0.444	0.543	0.630	0.898	
80	0.151	0.185	0.212	0.236	0.258	0.279	0.299	0.318	0.336	0.353	0.435	0.508	0.751	
100	0.126	0.154	0.176	0.196	0.215	0.232	0.248	0.264	0.279	0.294	0.363	0.425	0.638	0.946
200	0.070	0.085	0.098	0.108	0.119	0.128	0.137	0.146	0.154	0.162	0.201	0.236	0.362	0.570
400	0.039	0.047	0.053	0.059	0.064	0.069	0.074	0.079	0.083	0.088	0.108	0.127	0.196	0.317
800	0.021	0.025	0.029	0.032	0.034	0.037	0.039	0.042	0.044	0.046	0.057	0.067	0.103	0.168

$\alpha = .01$

$k \backslash n$	1	2	3	4	5	6	7	8	9	10	15	20	40	80
10	0.785	0.875	0.930	0.965	0.986	0.997	1.000	1.000						
15	0.629	0.724	0.792	0.844	0.887	0.921	0.948	0.969	0.984	0.994				
20	0.524	0.612	0.677	0.731	0.777	0.817	0.852	0.883	0.910	0.933	0.996			
25	0.450	0.529	0.589	0.640	0.685	0.724	0.761	0.794	0.824	0.851	0.953	0.997		
30	0.394	0.466	0.521	0.568	0.610	0.648	0.683	0.716	0.746	0.774	0.889	0.964		
40	0.318	0.377	0.424	0.464	0.501	0.534	0.565	0.595	0.622	0.649	0.763	0.855		
60	0.231	0.275	0.310	0.341	0.369	0.395	0.420	0.443	0.465	0.487	0.584	0.668	0.917	
80	0.183	0.218	0.246	0.271	0.293	0.314	0.334	0.353	0.372	0.389	0.471	0.543	0.778	
100	0.152	0.181	0.205	0.225	0.244	0.262	0.279	0.295	0.310	0.325	0.394	0.456	0.666	0.956
200	0.085	0.100	0.113	0.124	0.135	0.145	0.154	0.163	0.172	0.180	0.219	0.255	0.383	0.598
400	0.046	0.054	0.061	0.067	0.073	0.078	0.083	0.088	0.092	0.097	0.118	0.138	0.208	0.330
800	0.025	0.029	0.033	0.036	0.039	0.041	0.044	0.046	0.049	0.051	0.062	0.073	0.110	0.175

Table A-10 Critical values for the maximum of N values of Cook's $d(i) \times (n - k - 1)$
(Bonferroni correction used) n observations and k predictors

$\alpha = 0.1$

k	n=5	10	15	20	25	50	100	200	400	800
1	14.96	11.13	11.84	12.68	13.46	16.39	19.97	23.94	28.70	33.80
2	40.53	12.21	12.09	12.63	13.22	15.65	18.64	22.09	25.96	30.12
3		13.30	12.09	12.35	12.79	14.84	17.48	20.52	23.86	27.50
4		15.21	12.18	12.14	12.45	14.23	16.62	19.36	22.30	25.97
5		19.33	12.44	12.03	12.21	13.76	15.95	18.49	21.39	24.51
6		31.06	12.94	12.01	12.04	13.39	15.43	17.81	20.36	23.51
7		96.01	13.79	12.08	11.94	13.10	15.02	17.27	19.75	22.42
8			15.26	12.26	11.90	12.85	14.70	16.83	19.20	21.73
9			18.00	12.55	11.91	12.66	14.40	16.52	18.62	21.45
10			23.93	13.02	11.97	12.50	14.16	16.16	18.43	20.55
15				27.66	13.60	12.01	13.39	15.16	17.00	19.34
20					30.94	11.83	12.92	14.53	16.31	18.35
40						15.95	12.26	13.56	15.10	16.83
80							13.49	13.05	14.39	15.85

$\alpha = 0.05$

k	n=5	10	15	20	25	50	100	200	400	800
1	24.97	15.24	15.55	16.37	17.18	20.41	24.31	28.83	33.88	40.15
2	82.06	16.56	15.63	16.01	16.56	19.08	22.33	26.05	30.20	33.96
3		18.16	15.50	15.49	15.85	17.93	20.72	24.14	27.57	32.06
4		21.28	15.59	15.14	15.33	17.06	19.63	22.49	25.83	29.31
5		28.40	15.94	14.95	14.96	16.41	18.70	21.39	24.42	28.24
6		50.22	16.70	14.91	14.70	15.91	17.97	20.54	23.48	26.68
7		192.90	17.99	15.00	14.55	15.50	17.49	20.00	22.35	25.67
8			20.32	15.25	14.48	15.19	17.05	19.31	22.06	24.44
9			24.78	15.69	14.49	14.92	16.69	18.85	21.34	24.29
10			34.72	16.38	14.58	14.70	16.38	18.42	20.49	23.33
15				39.98	16.94	14.03	15.36	17.16	19.39	21.75
20					44.63	13.79	14.81	16.52	18.46	20.32
40						19.50	13.92	15.22	16.83	18.76
80							15.55	14.58	15.99	17.52

$\alpha = 0.01$

p	n=5	10	15	20	25	50	100	200	400	800
1	77.29	28.72	26.88	27.24	27.92	31.46	36.10	41.22	49.42	68.39
2	415.27	30.97	26.13	25.65	25.81	28.12	32.61	37.34	44.99	57.70
3		35.12	25.66	24.22	24.33	26.17	29.15	34.23	37.55	52.58
4		44.09	25.82	23.58	23.20	24.56	27.31	31.26	35.28	40.60
5		66.83	26.66	23.20	22.49	23.39	25.84	29.44	34.14	36.91
6		150.47	28.48	23.12	22.00	22.55	24.35	28.42	31.04	36.91
7		964.09	31.80	23.34	21.71	21.79	24.19	26.87	31.04	33.55
8			37.84	23.93	21.59	21.26	23.28	25.83	29.31	33.55
9			50.10	24.93	21.64	20.76	22.23	25.62	28.21	30.50
10			80.67	26.54	21.83	20.37	22.11	24.53	28.21	30.50
15				92.09	27.02	19.16	20.22	22.40	25.64	27.73
20					102.32	18.82	19.18	21.32	23.31	25.21
40						29.95	18.04	19.32	21.17	22.91
80							20.67	18.57	20.12	22.90