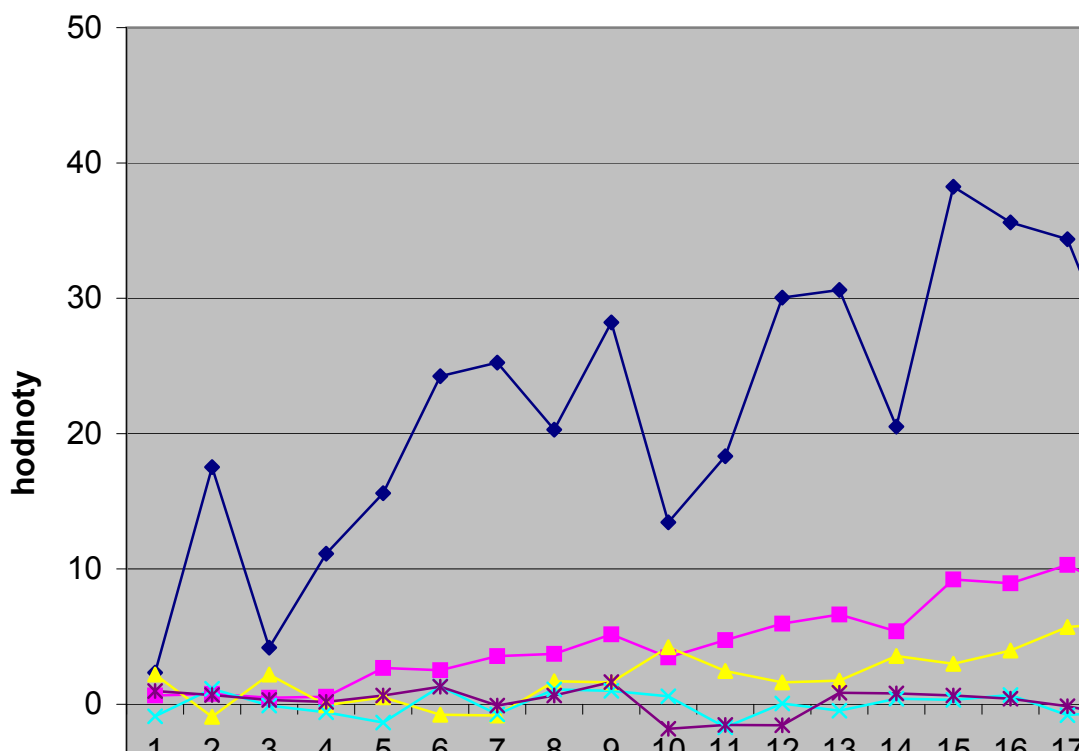


Příklad použití lineární regrese: data

logický čas t	vysvětlovaná: y	světlující veličiny: x ₁	x ₂	x ₃	x ₄
1	2,33	0,65	2,19	-0,89	0,97
2	17,52	0,76	-0,93	1,14	0,68
3	4,18	0,48	2,19	-0,1	0,31
4	11,12	0,55	-0,08	-0,6	0,17
5	15,6	2,68	0,51	-1,35	0,65
6	24,24	2,51	-0,78	1,32	1,3
7	25,24	3,56	-0,83	-0,75	-0,1
8	20,29	3,72	1,7	1,09	0,65
9	28,21	5,16	1,61	0,99	1,64
10	13,44	3,46	4,22	0,59	-1,82
11	18,33	4,74	2,46	-1,71	-1,53
12	30,05	5,95	1,61	0,05	-1,56
13	30,61	6,63	1,75	-0,48	0,84
14	20,52	5,39	3,57	0,41	0,8
15	38,24	9,22	2,99	0,32	0,64
16	35,61	8,94	3,97	0,68	0,41
17	34,36	10,29	5,71	-0,81	-0,15
18	23,87	8,16	6,04	-0,46	-0,97
19	28,98	9,64	5,37	-1,33	0,36
20	28,41	9,8	6,9	0,54	1,92

Modelované veličiny



Příklad použití lineární regrese: první model

Model

K vysvětlení se použijí pouze veličiny x_1 až x_3 .

Model má tvar: $y_t = b_0 + b_1 x_{t1} + b_2 x_{t2} + b_3 x_{t3} + e_t$.

pozorované hodnoty modelových proměnných				
t	y	x_1	x_2	x_3
1	2,33	0,65	2,19	-0,89
2	17,52	0,76	-0,93	1,14
3	4,18	0,48	2,19	-0,1
4	11,12	0,55	-0,08	-0,6
5	15,6	2,68	0,51	-1,35
6	24,24	2,51	-0,78	1,32
7	25,24	3,56	-0,83	-0,75
8	20,29	3,72	1,7	1,09
9	28,21	5,16	1,61	0,99
10	13,44	3,46	4,22	0,59
11	18,33	4,74	2,46	-1,71
12	30,05	5,95	1,61	0,05
13	30,61	6,63	1,75	-0,48
14	20,52	5,39	3,57	0,41
15	38,24	9,22	2,99	0,32
16	35,61	8,94	3,97	0,68
17	34,36	10,29	5,71	-0,81
18	23,87	8,16	6,04	-0,46
19	28,98	9,64	5,37	-1,33
20	28,41	9,8	6,9	0,54

parametry:	Odhady parametrů a dc	
	b_3	b_2
b_i	1,9023	-3,1069
sb_i	0,2434	0,1454
R^2, s	0,9921	0,9582
FR, n-k-1	669,1220	16,0000
	1842,8767	14,6889
$ b_i /sb_i$	7,8148	21,3667

Model je statisticky významný, protože FR (669)
Parametry jsou také statisticky významné, pro

Ověření splnění podmínek lineární regrese

t	yv	e	e^2	e^3	e^4	$e_t - e_{t-1}$
1	3,3836	-1,0536	1,1101	-1,1697	1,2324	1,1766
2	17,3970	0,1230	0,0151	0,0019	0,0002	-0,1212
3	4,1782	0,0018	0,0000	0,0000	0,0000	0,5468
4	10,5714	0,5486	0,3010	0,1651	0,0906	-1,1337
5	16,1850	-0,5850	0,3423	-0,2002	0,1172	0,2613
6	24,5638	-0,3238	0,1048	-0,0339	0,0110	0,4081
7	25,1557	0,0843	0,0071	0,0006	0,0001	-1,2562
8	21,4619	-1,1719	1,3734	-1,6095	1,8862	1,8317
9	27,5503	0,6597	0,4353	0,2872	0,1895	1,1820
10	11,5982	1,8418	3,3921	6,2475	11,5066	-1,5354
11	18,0236	0,3064	0,0939	0,0288	0,0088	0,6904
12	29,0532	0,9968	0,9936	0,9904	0,9872	-0,8297
13	30,4429	0,1671	0,0279	0,0047	0,0008	-0,9627
14	21,3156	-0,7956	0,6329	-0,5035	0,4006	0,1337
15	38,9019	-0,6619	0,4381	-0,2900	0,1919	0,8964
16	35,3755	0,2345	0,0550	0,0129	0,0030	1,3664

17	32,7591	1,6009	2,5629	4,1030	6,5686	-1,2571
18	23,5262	0,3438	0,1182	0,0407	0,0140	-1,4822
19	30,1184	-1,1384	1,2960	-1,4753	1,6795	-0,0402
20	29,5886	-1,1786	1,3891	-1,6372	1,9297	
součty:			14,6889	4,9633	26,8177	

Test normality reziduí

A3	0,3943	var A3	0,2236	norm test	test A3	0,8338
A4	-0,5142	var A4	0,5792	1,9600	test A4	0,3002

Test autokorelace reziduí Durbin-Watsonův

DW 1,39321

Test homoskedasticity reziduí Goldfeld-Quandtův

t	y	x ₁	x ₂	x ₃	e	e ²	
3	4,18	0,48	2,19	-0,1	0,0018	0,0000	SSE1
4	11,12	0,55	-0,08	-0,6	0,5486	0,3010	SSE2
1	2,33	0,65	2,19	-0,89	-1,0536	1,1101	SSE2/SSE1
2	17,52	0,76	-0,93	1,14	0,1230	0,0151	teoret.hodn
6	24,24	2,51	-0,78	1,32	-0,3238	0,1048	
5	15,6	2,68	0,51	-1,35	-0,5850	0,3423	
10	13,44	3,46	4,22	0,59	1,8418	3,3921	
7	25,24	3,56	-0,83	-0,75	0,0843	0,0071	
8	20,29	3,72	1,7	1,09	-1,1719	1,3734	Vynechávaná 4 prostřední pozo
11	18,33	4,74	2,46	-1,71	0,3064	0,0939	Vynechávaná 4 prostřední pozo
9	28,21	5,16	1,61	0,99	0,6597	0,4353	Vynechávaná 4 prostřední pozo
14	20,52	5,39	3,57	0,41	-0,7956	0,6329	Vynechávaná 4 prostřední pozo
12	30,05	5,95	1,61	0,05	0,9968	0,9936	F-rozdělení s $(T-T_2-2(k+1))/2$ a (
13	30,61	6,63	1,75	-0,48	0,1671	0,0279	
18	23,87	8,16	6,04	-0,46	0,3438	0,1182	
16	35,61	8,94	3,97	0,68	0,2345	0,0550	
15	38,24	9,22	2,99	0,32	-0,6619	0,4381	
19	28,98	9,64	5,37	-1,33	-1,1384	1,2960	
20	28,41	9,8	6,9	0,54	-1,1786	1,3891	
17	34,36	10,29	5,71	-0,81	1,6009	2,5629	

Razeno vzestupně podle hodnot proměnné x₁

Doprovodné statistiky:	
b_1	b_0
4,1659	9,1729
0,1007	0,3986
#N/A	#N/A
#N/A	#N/A
#N/A	#N/A
41,3735	23,0147

Vyhodnocení testů		
alfa	t-test	FR-test
0,05	2,1199	3,2389

n
20

0,122) je větší než FR-test (3,238867).
 protože $|b_1|/s_{b_1}$ je větší než hodnota t-testu (2,119905).

$(e_t - e_{t-1})^2$
1,3844
0,0147
0,2990
1,2852
0,0683
0,1666
1,5781
3,3550
1,3972
2,3573
0,4766
0,6883
0,9268
0,0179
0,8036
1,8670

1,5802
2,1970
0,0016
20,4647

Spočtená hodnota Durbin-Watsonova koeficientu autokorelace 1,393 spadá do pásma neurčitosti, kde je $d_L=1,00$, $d_U=1,68$

5,2726
6,8809
1,3050
6,3882

Protože podíl SSE_2/SSE_1 je menší než teoretická hodnota F-rozdělení o 4 a 4 stupních volnosti na hladině $\alpha = 0,05$; není proto důvod zamítnout hypotézu homoskedasticity

rování	T=	20
rování	T2=	4
rování	$2*(k+1)=$	8
rování	d.f.=	4
T-T ₂ -2(k+1))/2 stupni volnosti		

Příklad použití lineární regrese: rozš

Model

K vysvětlení se použijí všechny veličiny x_1 až x_4 .

Model má tvar: $y_t = b_0 + b_1 x_{t1} + b_2 x_{t2} + b_3 x_{t3} + b_4 x_{t4} + e_t$.

t	y	x_1	x_2	x_3	x_4	Odhady parametrů a dop	
1	2,33	0,65	2,19	-0,89	0,97	parametry	b_4 b_3
2	17,52	0,76	-0,93	1,14	0,68	b_i	-0,6242 2,1429
3	4,18	0,48	2,19	-0,1	0,31	sb_i	0,1753 0,1970
4	11,12	0,55	-0,08	-0,6	0,17	R^2, s	0,9957 0,7285
5	15,6	2,68	0,51	-1,35	0,65	FR, n-k-1	871,3371 15,0000
6	24,24	2,51	-0,78	1,32	1,3		1849,6054 7,9602
7	25,24	3,56	-0,83	-0,75	-0,1		
8	20,29	3,72	1,7	1,09	0,65	$ b_i /sb_i$	3,5608 10,8764
9	28,21	5,16	1,61	0,99	1,64		
10	13,44	3,46	4,22	0,59	-1,82		
11	18,33	4,74	2,46	-1,71	-1,53		Model je statisticky významný, proto
12	30,05	5,95	1,61	0,05	-1,56		Parametry jsou také statisticky významné
13	30,61	6,63	1,75	-0,48	0,84		
14	20,52	5,39	3,57	0,41	0,8		
15	38,24	9,22	2,99	0,32	0,64		
16	35,61	8,94	3,97	0,68	0,41		Test, zda přidání parametru
17	34,36	10,29	5,71	-0,81	-0,15		F12 F12-test
18	23,87	8,16	6,04	-0,46	-0,97		1,7300 2,3849
19	28,98	9,64	5,37	-1,33	0,36		
20	28,41	9,8	6,9	0,54	1,92		Neprokázalo set, že by přidání dalších

Vypovídací schopnost jednoc

směrodatné odchylky proměnných:

s_y s_{x1}

9,8877 3,3577

odhadnuté regresní koeficienty:

b_1

4,1996

absolutní korelace se závisle proměnnou:

$|\text{cor}(y, x_1)|$

1,4261

beta koeficienty:

beta1

58,50%

Proměnné x_1 až x_4 se na vysvětlení y podílejí významně.
Vliv proměnné x_4 je tedy celkem z

Ověření splnění podmínek lineární regrese								
t	yv	e	e ²	e ³	e ⁴	e _t - e _{t-1}		
1	2,6104	-0,2804	0,0786	-0,0221	0,0062	0,3288		
2	17,4716	0,0484	0,0023	0,0001	0,0000	0,1303		
3	4,0014	0,1786	0,0319	0,0057	0,0010	0,4504		
4	10,4910	0,6290	0,3957	0,2489	0,1566	-0,6922		
5	15,6632	-0,0632	0,0040	-0,0003	0,0000	-0,0420		
6	24,3452	-0,1052	0,0111	-0,0012	0,0001	-0,0057		
7	25,3509	-0,1109	0,0123	-0,0014	0,0002	-1,0947		
8	21,4957	-1,2057	1,4536	-1,7526	2,1131	2,4202		
9	26,9955	1,2145	1,4751	1,7915	2,1759	-0,6782		
10	12,9037	0,5363	0,2876	0,1543	0,0827	-0,9423		
11	18,7360	-0,4060	0,1649	-0,0669	0,0272	0,1598		
12	30,2962	-0,2462	0,0606	-0,0149	0,0037	0,7809		
13	30,0753	0,5347	0,2859	0,1529	0,0817	-1,0583		
14	21,0436	-0,5236	0,2742	-0,1436	0,0752	-0,1059		
15	38,8695	-0,6295	0,3962	-0,2494	0,1570	0,7305		
16	35,5090	0,1010	0,0102	0,0010	0,0001	1,4273		
17	32,8316	1,5284	2,3359	3,5700	5,4563	-1,7630		
18	24,1047	-0,2347	0,0551	-0,0129	0,0030	-0,5300		
19	29,7446	-0,7646	0,5847	-0,4470	0,3418	0,5637		
20	28,6109	-0,2009	0,0404	-0,0081	0,0016			
součty :			7,9602	3,2041	10,6833			
Test normality reziduí								
A3	0,6380	var A3	0,2236	norm test		test A3	1,3493	
A4	0,3720	var A4	0,5792	1,9600		test A4	0,8642	
Test autokorelace reziduí Durbin-Watsonův								
DW	2,1679							
Test homoskedasticity reziduí Goldfeld-Quandtův								
	y	x ₁	x ₂	x ₃	x ₄	e	e ²	
3	4,18	0,48	2,19	-0,1	0,31	0,1786	0,0319	
4	11,12	0,55	-0,08	-0,6	0,17	0,6290	0,3957	
1	2,33	0,65	2,19	-0,89	0,97	-0,2804	0,0786	
2	17,52	0,76	-0,93	1,14	0,68	0,0484	0,0023	
6	24,24	2,51	-0,78	1,32	1,3	-0,1052	0,0111	
5	15,6	2,68	0,51	-1,35	0,65	-0,0632	0,0040	
10	13,44	3,46	4,22	0,59	-1,82	0,5363	0,2876	
7	25,24	3,56	-0,83	-0,75	-0,1	-0,1109	0,0123	
8	20,29	3,72	1,7	1,09	0,65	-1,2057	1,4536	Vynechávaná 4
11	18,33	4,74	2,46	-1,71	-1,53	-0,4060	0,1649	Vynechávaná 4
9	28,21	5,16	1,61	0,99	1,64	1,2145	1,4751	Vynechávaná 4
14	20,52	5,39	3,57	0,41	0,8	-0,5236	0,2742	Vynechávaná 4
12	30,05	5,95	1,61	0,05	-1,56	-0,2462	0,0606	F-rozdělení s (T-

13	30,61	6,63	1,75	-0,48	0,84	0,5347	0,2859	
18	23,87	8,16	6,04	-0,46	-0,97	-0,2347	0,0551	
16	35,61	8,94	3,97	0,68	0,41	0,1010	0,0102	
15	38,24	9,22	2,99	0,32	0,64	-0,6295	0,3962	
19	28,98	9,64	5,37	-1,33	0,36	-0,7646	0,5847	
20	28,41	9,8	6,9	0,54	1,92	-0,2009	0,0404	
17	34,36	10,29	5,71	-0,81	-0,15	1,5284	2,3359	
		Řazeno vzestupně podle hodnot proměnné x_1						

Řízený model

Provodné statistiky:			Vyhodnocení testů :		
b ₂	b ₁	b ₀	alfa	t-test	FR-test
-3,1629	4,1996	9,3200	0,05	2,1314	3,2874
0,1117	0,0771	0,3058			
#N/A	#N/A	#N/A	n		
#N/A	#N/A	#N/A	20		
#N/A	#N/A	#N/A			
28,3246	54,4441	30,4743			

že FR (871,3371) je větší než FR-test (3,238867).
znamné, protože |b₁/s_{b1} je větší než hodnota t-testu (2,119905).

etru zlepšilo vypovídací schopnost modelu

ší proměnné statisticky významně zlepšilo shodu modelu s daty.

ktivých vysvětlujících proměnných - Beta koeficienty

s _{x2}	s _{x3}	s _{x4}	
2,3412	0,9128	1,0310	
b ₂	b ₃	b ₄	
-3,1629	2,1429	-0,6242	
cor(y,x ₂)	cor(y,x ₃)	cor(y,x ₄)	suma(cor(.))
0,7489	0,1978	0,0651	2,4379
beta2	beta3	beta4	suma(beta)
30,72%	8,11%	2,67%	1,0000

ní rozptylu veličiny y podílejí po řadě 58,5%, 30,72%, 8,11% a 2,67%.
anedbatelný.

Příklad použití lineární regrese: předp

Model

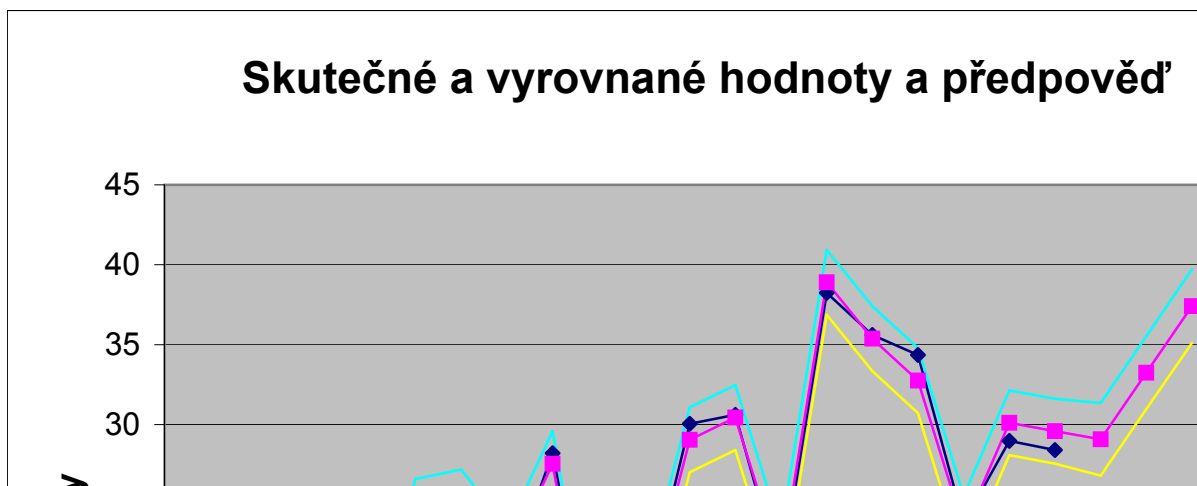
K vysvětlení se použijí pouze veličiny x_1 až x_3 .

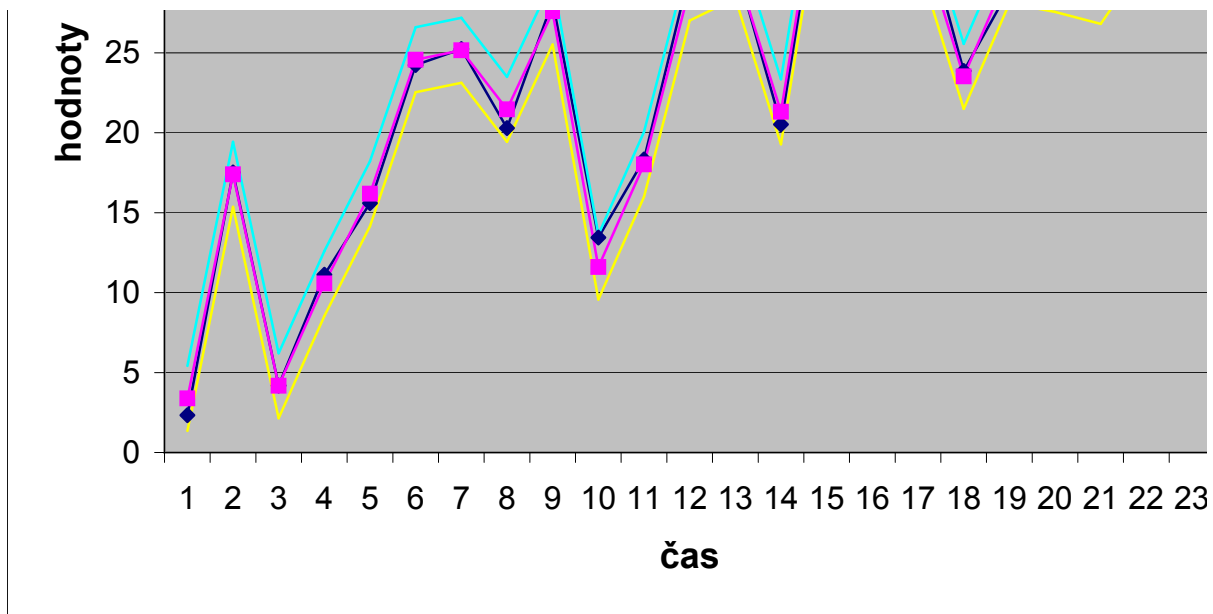
Model má tvar: $y_t = b_0 + b_1 x_{t1} + b_2 x_{t2} + b_3 x_{t3} + e_t$.

t	y	konst.	x_1	x_2	x_3	intervaly spolehlivosti poz	
						yv	yv _d
1	2,33	1	0,65	2,19	-0,89	3,3836	1,3524
2	17,52	1	0,76	-0,93	1,14	17,3970	15,3658
3	4,18	1	0,48	2,19	-0,1	4,1782	2,1470
4	11,12	1	0,55	-0,08	-0,6	10,5714	8,5402
5	15,6	1	2,68	0,51	-1,35	16,1850	14,1538
6	24,24	1	2,51	-0,78	1,32	24,5638	22,5326
7	25,24	1	3,56	-0,83	-0,75	25,1557	23,1245
8	20,29	1	3,72	1,7	1,09	21,4619	19,4307
9	28,21	1	5,16	1,61	0,99	27,5503	25,5191
10	13,44	1	3,46	4,22	0,59	11,5982	9,5670
11	18,33	1	4,74	2,46	-1,71	18,0236	15,9924
12	30,05	1	5,95	1,61	0,05	29,0532	27,0220
13	30,61	1	6,63	1,75	-0,48	30,4429	28,4117
14	20,52	1	5,39	3,57	0,41	21,3156	19,2844
15	38,24	1	9,22	2,99	0,32	38,9019	36,8707
16	35,61	1	8,94	3,97	0,68	35,3755	33,3443
17	34,36	1	10,29	5,71	-0,81	32,7591	30,7279
18	23,87	1	8,16	6,04	-0,46	23,5262	21,4950
19	28,98	1	9,64	5,37	-1,33	30,1184	28,0872
20	28,41	1	9,8	6,9	0,54	29,5886	27,5574
						střed	dolní mez

t_p	Předpovědi:					intervaly spolehlivosti před	
	konst.	x_{1p}	x_{2p}	x_{3p}	y_p	y_{p_d}	
21	1	10	7	0	29,0839	26,8130	
22	1	11	7	0	33,2498	30,9698	
23	1	12	7	0	37,4158	35,1067	
					střed	dolní mez	

Skutečné a vyrovnané hodnoty a předpověď





ověř' prvního modelu

or. hodnot

yv_n	f_t
5,4148	1,00
19,4282	1,00
6,2094	1,00
12,6026	1,00
18,2162	1,00
26,5950	1,00
27,1869	1,00
23,4931	1,00
29,5814	1,00
13,6294	1,00
20,0548	1,00
31,0844	1,00
32,4741	1,00
23,3468	1,00
40,9331	1,00
37,4067	1,00
34,7903	1,00
25,5574	1,00
32,1496	1,00
31,6198	1,00

horní mez

dp. hodnot

yp_n	
31,3547	1,1180
35,5299	1,1225
39,7248	1,1368

horní mez

s
0,9582
$t_{n-k-1}(\text{alfa})$
2,1199
alfa
0,05

ověř'

