pohlaví	předmět
chlapec	Tv
chlapec	Tv
chlapec	Tv
chlapec	Tv
chlapec	Tv
chlapec	Tv
chlapec	Tv
chlapec	Bi
chlapec	Bi
chlapec	Bi
chlapec	Bi
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	D
chlapec	D
chlapec	D
chlapec	D
chlapec	D
chlapec	D
chlapec	D
chlapec	D
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi

divka	M
divka	M
divka	M
divka	M
divka	M
divka	M
divka	M
divka	M
divka	M
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D

pohlaví	předmět
chlapec	Tv
chlapec	Tv
chlapec	Tv
chlapec	Tv
chlapec	Tv
chlapec	Tv
chlapec	Tv
chlapec	Bi
chlapec	Bi
chlapec	Bi
chlapec	Bi
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	M
chlapec	D
chlapec	D
chlapec	D
chlapec	D
chlapec	D
chlapec	D
chlapec	D
chlapec	D
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Tv
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi
divka	Bi

Počet z předmět	předmět			
pohlaví	Bi	D	M	Tv
divka	16	11	9	9
chlapec	4	8	11	7
Celkový součet	20	19	20	16

	Bi	D	M	Tv
dívka	16	11	9	9
očekávaný počet	12	11.4	12	9.6
chlapec	4	8	11	7
očekávaný počet	8	7.6	8	6.4
Celkový součet	20	19	20	16

divka	Bi
divka	M
divka	M
divka	M
divka	M
divka	M
divka	M
divka	M
divka	M
divka	M
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D
divka	D

Celkový součet
45
30
75

Celkový součet
45
45
30
30
75

čtyřpolní tabulky

	praváci	leváci
muži	41	9
ženy	46	4

$$\chi^2 = n\frac{(ad - b\ldots}{(a+b)(a+c)(c\ldots}$$

$$\frac{c)^{2}}{+ d)(b + d)}$$

čtyřpolní tabulky

	pravácí	leváci	celkem
muži	41	9	50
ženy	46	4	50
celkem	87	13	100

$$\chi^2 = n\frac{(ad - bc}{(a + b)(a + c)(c -}$$

chi2= **2.210433** < 3.841459 krit.

$$\frac{\ldots c)^2}{\ldots + d)(b + d)}$$

Příklad

Při přijímacím řízení se provádělo hodnocení komisí a hodnocení speciálním programem. Na za

Student	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Hodnocení komisí	***4***	***6***	***1***	***5***	***10***	***2***	***7***
Hodnocení programem	***1***	***3***	***5***	***7***	***8***	***4***	***6***
Diference pořadí							
Čtverec diference							

ákladě údajů o pořadí deseti studentů rozhodněte o tom, zda jsou obě hodnocení závislá.

H	I	J
3	*9*	*8*
2	*10*	*9*

$$rs = 1 - \frac{6\sum_{i=1}^{n} d_i^{\,2}}{n^3 - n}$$

Tabulka A.10: Kritické hodnoty Spearmanova korelačního koeficientu

n	*α*		*n*	*α*		*n*	*α*	
	0,05	0,01		0,05	0,01		0,05	0,01
			11	0,6091	0,7545	21	0,4351	0,5545
			12	0,5804	0,7273	22	0,4241	0,5426
			13	0,5549	0,6978	23	0,4150	0,5306
			14	0,5341	0,6747	24	0,4061	0,5200
5	0,9000	-	15	0,5179	0,6536	25	0,3977	0,5100
6	0,8286	0,9429	16	0,5000	0,6324	26	0,3894	0,5002
7	0,7450	0,8929	17	0,4853	0,6152	27	0,3822	0,4915
8	0,6905	0,8571	18	0,4716	0,5975	28	0,3749	0,4828
9	0,6833	0,8167	19	0,4579	0,5825	29	0,3685	0,4744
10	0,6364	0,7818	20	0,4451	0,5684	30	0,3620	0,4665

Příklad

Při přijímacím řízení se provádělo hodnocení komisí a hodnocení speciálním programem. Na za

Student	A	B	C	D	E	F	G
Hodnocení komisí	***4***	***6***	***1***	***5***	***10***	***2***	***7***
Hodnocení programem	***1***	***3***	***5***	***7***	***8***	***4***	***6***
Diference pořadí	3	3	-4	-2	2	-2	1
Čtverec diference	9	9	16	4	4	4	1

diskuse:

kdyby bylo právě opačné pořadí:

	1	2	3	4	5	6	7
	10	9	8	7	6	5	4
	9	7	5	3	1	-1	-3
1980	81	49	25	9	1	1	9
2							

rs= **-1**

ákladě údajů o pořadí deseti studentů rozhodněte o tom, zda jsou obě hodnocení závislá.

H	I	J
3	***9***	***8***
2	***10***	***9***
1	-1	-1
1	1	1

$$rs = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n}$$

rs= 0.69697 > 0,6364

8	9	10
3	2	1
-5	-7	-9
25	49	81

Tabulka A.10: Kritické hodnoty Spearmanova korelačního koeficientu

n	α		*n*	α		*n*	α	
	0,05	0,01		0,05	0,01		0,05	0,01
			11	0,6091	0,7545	21	0,4351	0,5545
			12	0,5804	0,7273	22	0,4241	0,5426
			13	0,5549	0,6978	23	0,4150	0,5306
			14	0,5341	0,6747	24	0,4061	0,5200
5	0,9000	-	15	0,5179	0,6536	25	0,3977	0,5100
6	0,8286	0,9429	16	0,5000	0,6324	26	0,3894	0,5002
7	0,7450	0,8929	17	0,4853	0,6152	27	0,3822	0,4915
8	0,6905	0,8571	18	0,4716	0,5975	28	0,3749	0,4828
9	0,6833	0,8167	19	0,4579	0,5825	29	0,3685	0,4744
10	0,6364	0,7818	20	0,4451	0,5684	30	0,3620	0,4665

Test1	Test2
80	65
50	60
36	35
58	39
72	48
60	44
56	48
68	61

Příklad Studenti

Máme k dispozici výsledky testů ze dvou předmětu zjištění u 8 n

Stanovte těsnost lineární závislosti těchto výsledků Spearmanovým a Pear

áhodně vybraných studentů.
sonovým koeficientem.

Test1	Test2
80	65
50	60
36	35
58	39
72	48
60	44
56	48
68	61

Příklad Studenti

Máme k dispozici výsledky testů ze dvou předmětu zjištění u 8 ná
Stanovte těsnost lineární závislosti těchto výsledků Spearmanovým a Pears

spearman		*d2*
8	8	*0*
2	6	*16*
1	1	*0*
4	2	*4*
7	4.5	*6.25*
5	3	*4*
3	4.5	*2.25*
6	7	*1*
		0.6012

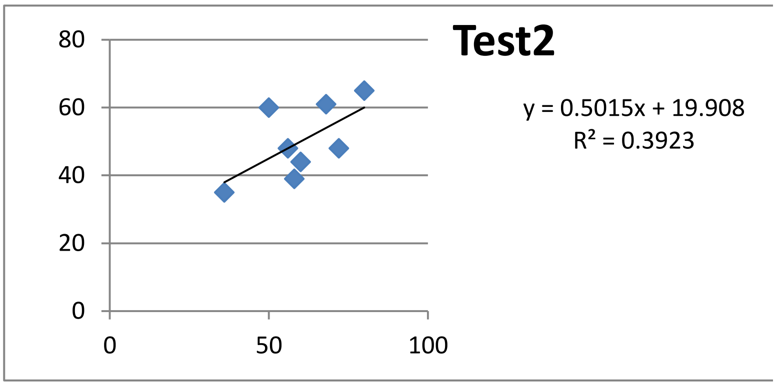
$$r_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n}$$

Test1	Test2	
X	Y	X-x
80	65	20
50	60	-10
36	35	-24
58	39	-2
72	48	12
60	44	0
56	48	-4
68	61	8
průměr x	*průměr y*	
60	**50**	

Pearsonův korelační koeficier

correl()
pearson()

ıhodně vybraných studentů.
onovým koeficientem.



Y-y	(X-x)*(Y-y			(X-x)^2	(Y-y)^2
15	300			400	225
10	-100			100	100
-15	360			576	225
-11	22			4	121
-2	-24			144	4
-6	0			0	36
-2	8			16	4
11	88			64	121
	suma		*suma*		suma
	654			1304	836
		násobek			1090144
		odmocnina			1044.1

nt r

r = 0.6264

0.6264

0.6264

stáří	cena
3	167
4	165
5	139
6	149
7	119
7	129
8	89
8	115
9	76
9	89

předpokládáme, že data jsou normálně rozložena

Jednoduchá regrese

Příklad 1:

Na základě údaje o stáří a ceně 10 ojetých aut značky Felicia Comb

1. zkonstruujte regresní model závislosti **ceny** auta na **stáří**,
2. posuďte jeho kvalitu a
3. použijte jej k odhadu střední hodnoty ceny aut starých 10 le

bi

et.

stáří	cena
3	167
4	165
5	139
6	149
7	119
7	129
8	89
8	115
9	76
9	89

předpokládáme, že data jsou normálně rozložena

Jednoduchá regrese

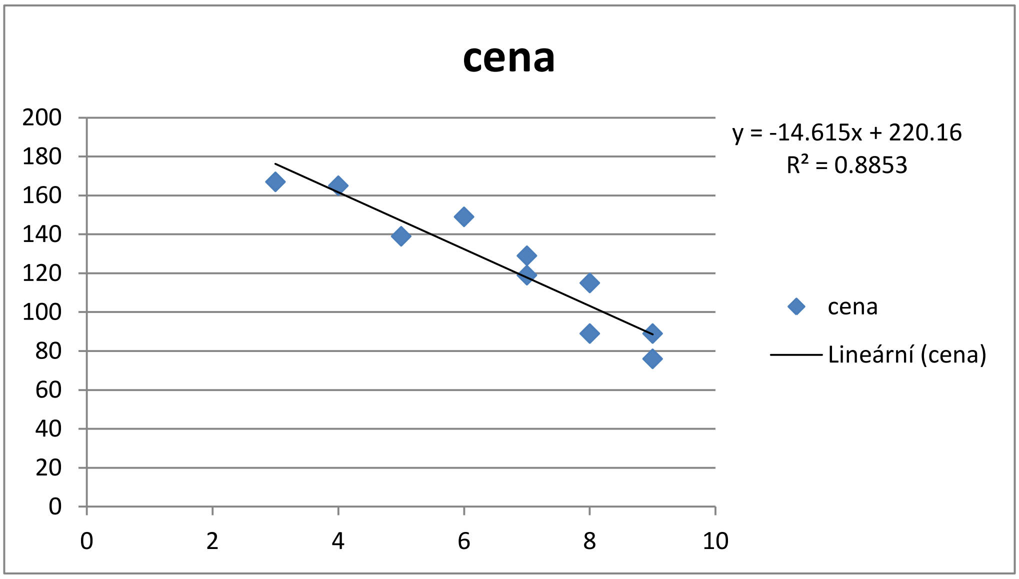
Příklad 1:

Na základě údaje o stáří a ceně 10 ojetých aut značky Felicia Coml

1. zkonstruujte regresní model závislosti **ceny** auta na **stáří**,
2. posuďte jeho kvalitu a
3. použijte jej k odhadu střední hodnoty ceny aut starých 10 l

-0.94092

0.940904



http://stp202.blog.cz/

bi

et.

VÝSLEDEK

Regresní statistika	
Násobné R	0.940915089
Hodnota spolehliv	0.885321204
Nastavená hodnc	0.870986355
Chyba stř. hodno	11.52386563
Pozorování	10

je to kladné????

ANOVA

	Rozdíl	*SS*	*MS*	*F*	*Významnost F*
Regrese	1	8201.704	8201.704	61.76006	4.96E-05
Rezidua	8	1062.396	132.7995		
Celkem	9	9264.1			

	Koeficienty	*'ba stř. hodı*	*t Stat*	*Hodnota P*	*Dolní 95%*	*Horní 95%*
Hranice	220.15625	12.80329	17.19528	1.33E-07	190.6318	249.6807
stáří	-14.61458333	1.859656	-7.85876	4.96E-05	-18.903	-10.3262

Dolní 95.0%	Horní 95.0%
190.6318	249.6807
-18.903	-10.3262

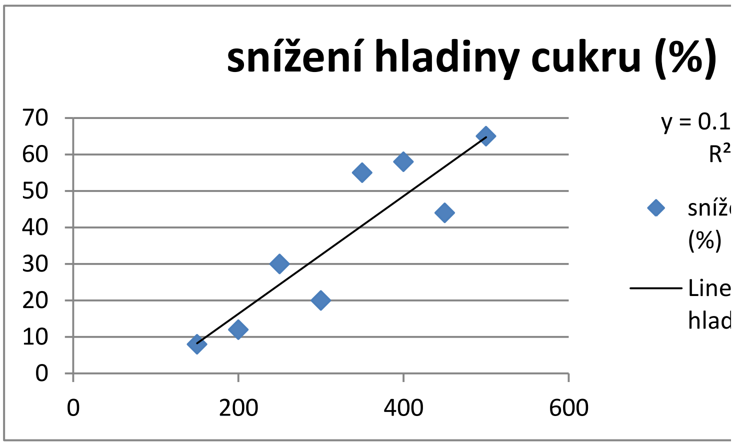
Byl vyvinut nový druh insulinu a zkoumá se závislost snížení hladiny cukru v krvi pacienta na množství nového insulinu určitou dobu před měřením.
Náhodně vybraným 8 pacientům byla naočkována různá množství insulinu a po určité době bylo těmto pacientům změřeno snížení cukru v krvi.
Výsledky měření: **prokažte silnou korelaci a zobrazte reziduály od regresní přímky**!

množství insulinu (ug)	150	200	250	300	350	400
snížení hladiny cukru (%)	8	12	30	20	55	58

!

	450	500
	44	65

Byl vyvinut nový druh insulinu a zkoumá se závislost snížení hladiny cukru v krvi pacienta na množství nového insulinu určitou dobu před měřením.
Náhodně vybraným 8 pacientům byla naočkována různá množství insulinu a po určité době bylo těmto pacientům změřeno snížení cukru v krvi.
Výsledky měření:



množství insulinu (ug)	150	200	250	300	350	400
snížení hladiny cukru (%)	8	12	30	20	55	58

	množství insulinu	*hladiny cukru (%)*	Analýza dat:korelace
množství insulinu (ug)	1		R2
snížení hladiny cukru (%)	0.897089	1	0.804769

Analýza dat:regrese

množství insulinu (ug)	snížení hladiny cukru (%)
150	8
200	12
250	30
300	20
350	55
400	58
450	44
500	65

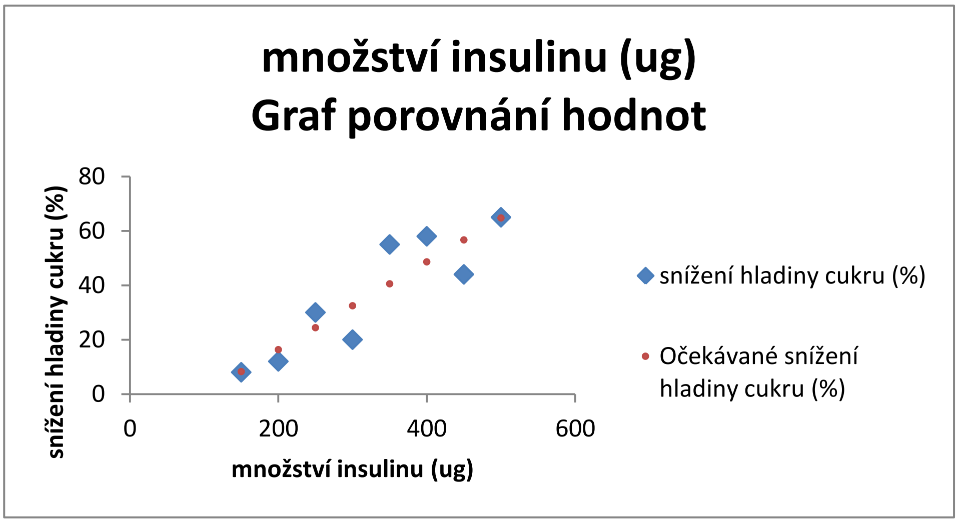
VÝSLEDEK

Regresní sta
Násobné R
Hodnota spolehl
Nastavená hodn
Chyba stř. hodn
Pozorování

ANOVA

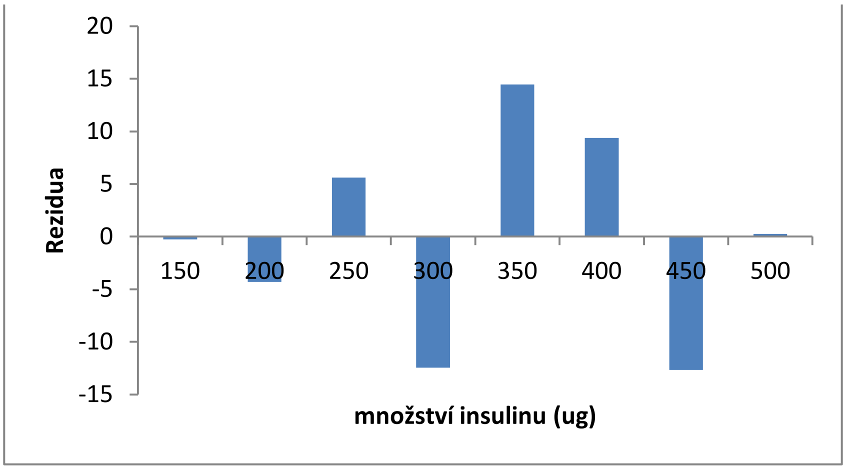
Regrese
Rezidua
Celkem

Hranice
množství insulin

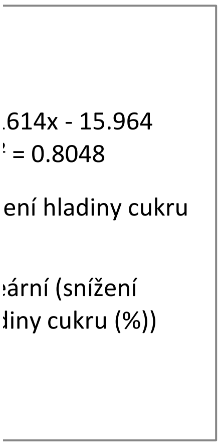


REZIDUA

Pozorování
1
2
3
4
5
6
7



20
15
10
5
0
-5
-10
-15
Rezidua
150
200
250
300
350
400
450
500
množství insulinu (ug)



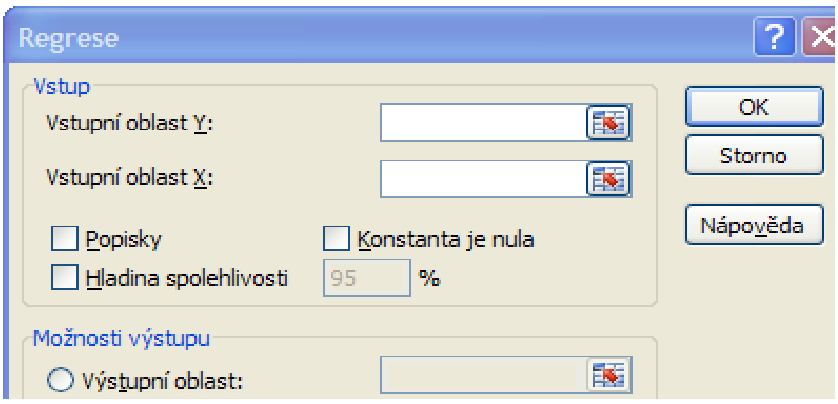
			linregrese()			klíč:	
		4.973209	0.16142857	-15.9642857	*1.427214*	a	b
			0.03245964	11.18563092		sa	sb
450	500		0.80476891	10.51812495		R2	syx
44	65		24.7328096	6	2.446912	*F*	stupvol
			2736.21429	663.7857143		*ssreg*	*ssres*

atistika
0.897089
0.804769
0.77223
10.51812
8

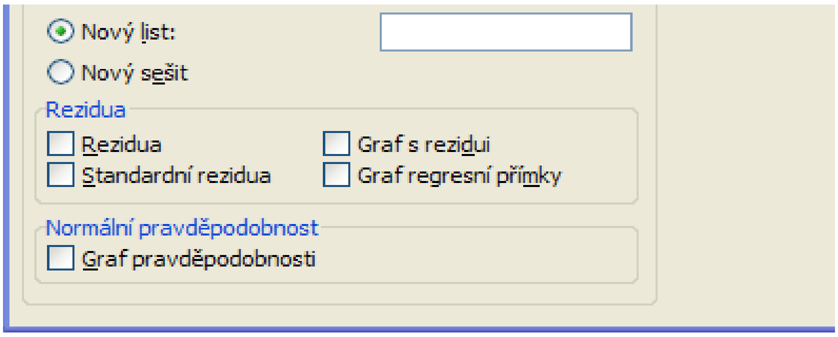
Rozdíl	*SS*	*MS*	*F*	*Významnost F*
1	2736.214	2736.214	24.7328096	0.002518757
6	663.7857	110.631		
7	3400			

Koeficienty	*ba stř. hodı*	*t Stat*	*Hodnota P*	*Dolní 95%*	*Horní 95%*	*Dolní 95.0%*	*Horní 95.0%*
-15.9643	11.18563	-1.42721	0.20342468	-43.3345385	11.40597	-43.3345	11.40597
0.161429	0.03246	4.973209	0.00251876	0.082002698	0.240854	0.082003	0.240854

nížení hlad	*Rezidua*
8.25	-0.25
16.32143	-4.32143
24.39286	5.607143
32.46429	-12.4643
40.53571	14.46429
48.60714	9.392857
56.67857	-12.6786



64.75	0.25


Nový list:
Nový sešit
Rezidua
Rezidua
Standardní rezidua
Graf s rezidui
Graf regresní přímky
Normální pravděpodobnost
Graf pravděpodobnosti

Testování, jehož smyslem je porovnání odhadů a, b se skutečnými hodnotami parametrů α, β, lze snadno provést za použití testu t: často budeme např. testovat, zda se posunutí neliší od nuly, tj. je-li vůbec statisticky významné, zda vlastně není $\alpha = 0$, event. zda je směrnice $\beta = 1$ apod. Testování koeficientu a pomocí Studentova testu t (viz odd. 2.6) se provede porovnáním vypočtené hodnoty t

$$t = \frac{|a - \alpha|}{s_a} \tag{3.2.12}$$

resp. při testování pro $\alpha = 0$

$$t = \frac{a}{s_a} \tag{3.2.13}$$

s tabelovanou kritickou hodnotou Studentova rozdělení (tab. 8), kterou hledáme pro počet stupňů volnosti $\nu = n - 2$ a pro zvolenou hladinu významnosti a zjistíme, zda je rozdíl statisticky významný.

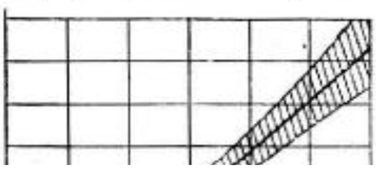
3.2.3 Interval spolehlivosti hodnoty vypočtené z regresní rovnice

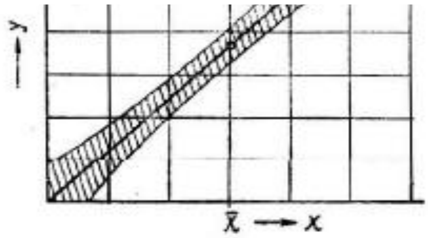
Pro hodnotu Y_i vypočítanou z lineární regresní rovnice můžeme jako pro každou náhodnou veličinu určit interval spolehlivosti (viz odd. 2.5). Dolní a horní mez intervalu spolehlivosti hodnoty Y_i, odpovídající zvolené hodnotě nezávisle proměnné x_i, lze za předpokladu normálního rozdělení vypočítat podle vztahu

$$L_{1,2} = Y_i \pm t_\alpha s_{y,x} \sqrt{\left[\frac{1}{n} + \frac{(X_i - \bar{x})^2}{\Sigma x_i^2 - (\Sigma x_i)^2/n}\right]} \tag{3.2.19}$$

V tomto vztahu je $s_{y,x}$ směrodatná odchylka, charakterizující rozptýlení kolem regresní přímky, určené podle vztahu (3.2.7) pro regresní rovnici (3.2.1), nebo podle vztahu (3.2.17) pro regresní rovnici (3.2.15); t_α je kritická hodnota Studentova rozdělení pro zvolenou hladinu významnosti α a pro počet stupňů volnosti $\nu = n - 2$ pro regresní přímku, vyjádřenou rovnicí (3.2.1) nebo pro $\nu = n - 1$ pro lineární regresi tvaru (3.2.15). Dále v ní vystupuje průměrná hodnota nezávisle proměnné $\bar{x}$ a hodnoty X_i a Y_i. Ze vztahu (3.2.19) je zřejmé, že horní a dolní hranice intervalu spolehlivosti závisejí na tom, jak se regresní hodnota nezávisle proměnné liší od průměru. To znamená, že hodnoty L_1 a L_2 vytvářejí okolo regresní přímky pás spolehlivosti, který je nejužší v bodě, kde $X_i = \bar{x}$, jak je znázorněno na obr. 12.

Praktický význam intervalu spolehlivosti viz odd. 2.5.



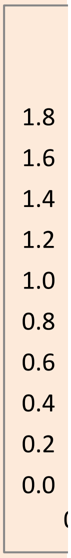


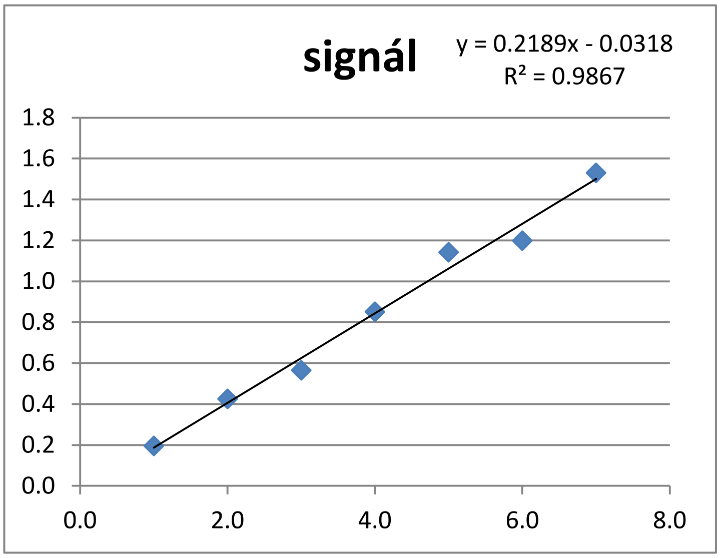
Obr. 12. **Pás spolehlivosti regresní lineární závislosti**

Koncentrace	signál
1	0.195
2	0.425
3	0.565
4	0.851
5	1.142
6	1.198
7	1.530

Koncentrace	signál			
1.0	0.195		**0.72**	
2.0	0.425		**3.4**	*3.4*
3.0	0.565			
4.0	0.851			
5.0	1.142			
6.0	1.198			
7.0	1.530			

0.218885714	-0.03177143	0.624046	2.570582	0.212531	0
0.011384271	0.05091201			0.004825	#N/A
0.986655231	0.0602399			0.996917	0.057093
369.6786448	5			1940.055	6
1.341506766	0.01814423			6.323745	0.019557





VÝSLEDEK

Regresní statistika	
Násobné R	0.99330521
Hodnota spolehlivo	0.98665523
Nastavená hodnot	0.98398628
Chyba stř. hodnoty	0.0602399
Pozorování	7

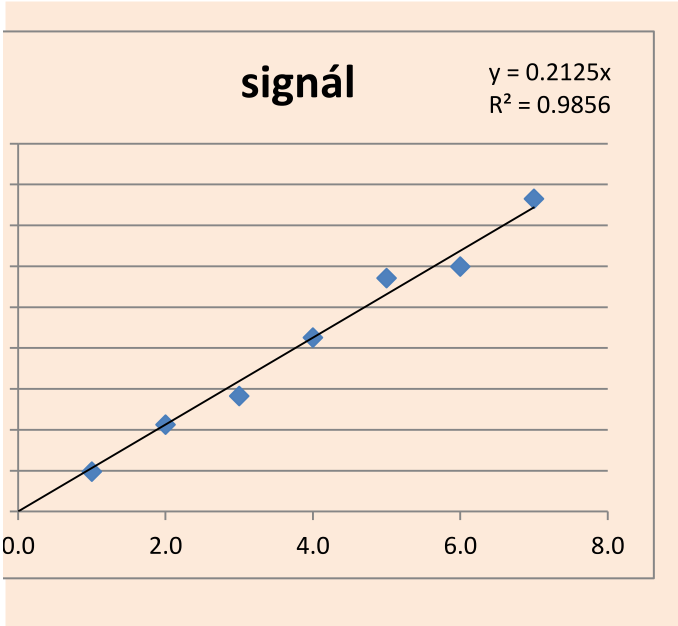
ANOVA

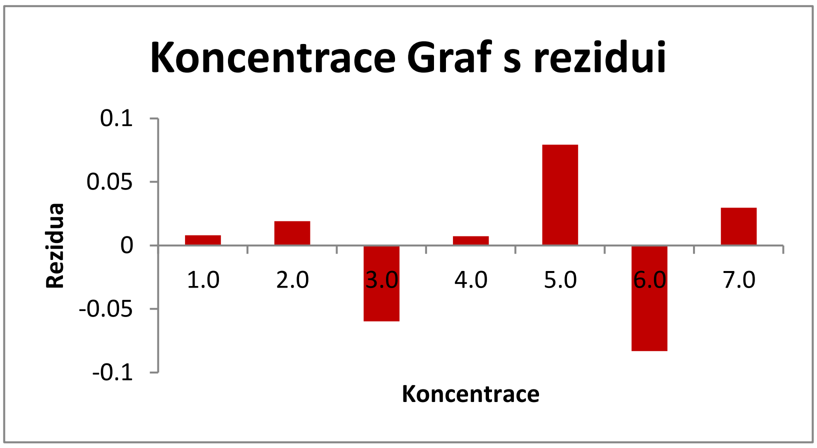
	Rozdíl	*SS*	*MS*	*F*	*ýznamnost F*
Regrese	1	1.341507	1.341507	369.6786	7.02E-06
Rezidua	5	0.018144	0.003629		
Celkem	6	1.359651			

	Koeficienty	*ba stř. hodn*	*t Stat*	*Hodnota P*	*Dolní 95%*	*Horní 95%*	*Dolní 95.0%*
Hranice	-0.03177143	0.050912	-0.62405	0.559962	-0.16264	0.099102	-0.16264
Koncentrace	0.21888571	0.011384	19.22703	7.02E-06	0.189622	0.24815	0.189622

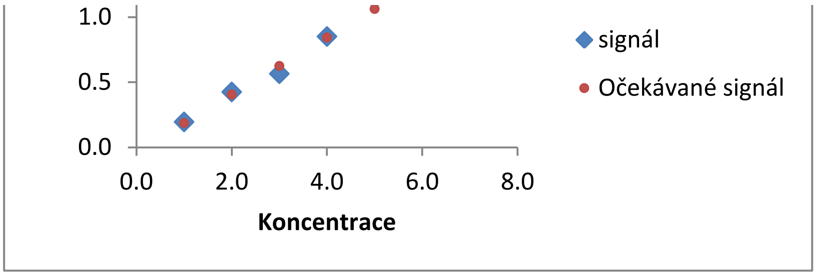
REZIDUA

Pozorování	čekávané sign	Rezidua
1	0.18711429	0.007886
2	0.406	0.019
3	0.62488571	-0.05989
4	0.84377143	0.007229
5	1.06265714	0.079343
6	1.28154286	-0.08314
7	1.50042857	0.029571





Horní 95.0%
0.099102
0.24815


1.0
0.5
0.0
0.0
2.0
4.0
6.0
8.0
Koncentrace
signál
Očekávané signál