

UNIVERZITA PARDUBICE

**Fakulta chemicko-technologická
Katedra analytické chemie**

Licenční studium chemometrie na téma

Statistické zpracování dat

Semestrální práce ze 6. soustředění

Předmět: 3.3 Tvorba nelineárních regresních modelů v analýze dat

Vedoucí licenčního studia: Prof. RNDr. Milan Meloun, DrSc.

Vypracoval: Ing. Jiří Souček, Ph.D.

Licenční studium Statistické zpracování experimentálních dat.

Předmět: 3.3 Tvorba nelineárních regresních modelů v analýze dat

Přednášející: Prof. RNDr. Milan Meloun, DrSc., Prof. Ing. Jiří Militký, CSc.

Zadání: Vypracovat alespoň dvě úlohy, nejlépe z dat vlastního pracoviště.

Výpočet výškových křivek stromů na pokusné ploše. Pro konstrukci a výpočet výškových křivek byly použity různé funkce s rozdílným počtem parametrů:

Logaritmická funkce $h=A1+A2*\ln(d)$

Naeslund $h=((d^2/A1+A2*d)^2) + 1,3$

Prodan $h=d^2/(1+A2*d+A3*d^2)$

Která křivka prokládá data nejlépe? Jaký má rozdílný počet parametrů vstupujících do výpočtu vliv na přesnost?

Data: vysky-mochov.vts

Odhadování parametrů

Logaritmická křivka

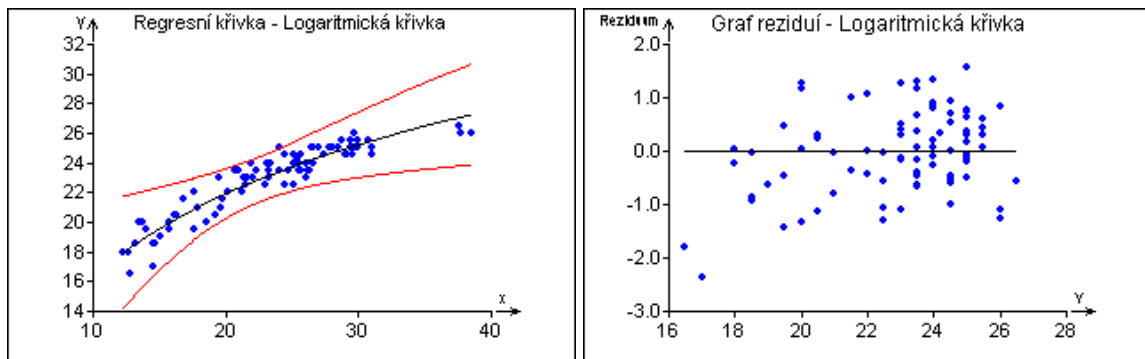
Odhady	Parametr	Směr. odchylka	Dolní mez	Horní mez
P1	-2,504	0,992	-4,476	-0,532
P2	8,157	0,317	7,527	8,787

Naeslund

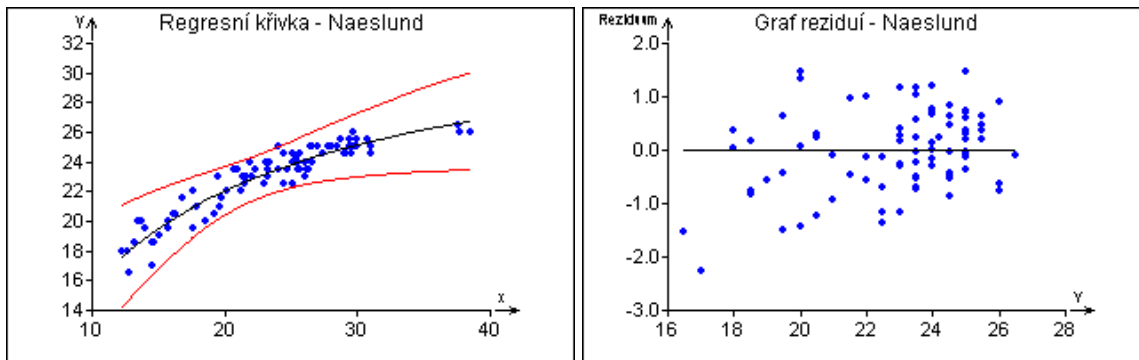
Odhady	Parametr	Směr. odchylka	Dolní mez	Horní mez
P1	0,888	0,0364	0,8167	0,9590
P2	0,175	0,0016	0,1720	0,1783

Prodan

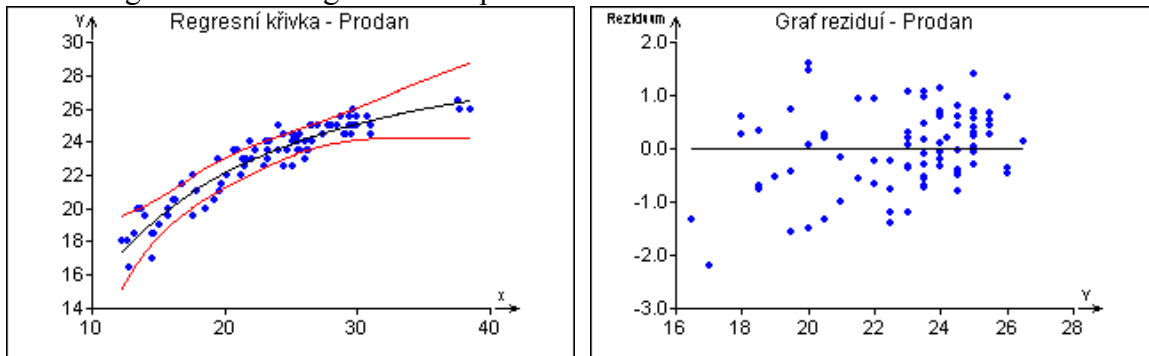
Odhady	Parametr	Směr. odchylka	Dolní mez	Horní mez
P1	1,544	0,915	-0,275	3,363
P2	0,191	0,091	0,009	0,373
P3	0,032	0,002	0,028	0,036



Obr. 1 Regresní křivka a graf reziduí pro logaritmickou funkci



Obr. 2 Regresní křivka a graf reziduí pro funkci Naeslunda



Obr. 3 Regresní křivka a graf reziduí pro funkci Prodana

Jednotlivé regresní křivky poměrně dobře prokládají data, s nárůstem parametrů se zúžují pásy spolehlivosti ve středních hodnotách. Rezidua tvoří náhodné mraky, to naznačuje vhodnost modelu.

Korelační matice parametrů

Logaritmická:	P1	P2
P1	1	-0,9964
P2	-0,9964	1

Naeslund:	P1	P2
P1	1	-0,9674
P2	-0,9674	1

Prodán:	P1	P2	P3
P1	1	-0,9877	0,9588
P2	-0,9877	1	-0,9901
P3	0,9588	-0,9901	1

Jednotlivé korelační matice ukazují na těsnou korelaci mezi jednotlivými parametry

Základní statistické charakteristiky

	Regresní rabat	AIK	Se	Šikmost	Špičatost
Logaritmická	88,747	-39,706	0,785	0,134	2,973
Naeslund	89,661	-46,997	0,752	0,130	3,026
Prodán	89,8	-46,235	0,751	0,128	3,053

Závěr:

Regresní rabat pro jednotlivé funkce srovnatelné procento bodů odpovídajících navrženým modelům. Odhady šikmosti i špičatosti odpovídají normálnímu rozdělení. Hodnota AIK dosáhla minimální hodnoty u Prodanovy funkce, rozdíl mezi funkcí Prodana (3 parametry) a Naeslunda (2 parametry) je u většiny sledovaných charakteristik minimální.

Úloha 2.

Výpočet výškových křivek stromů (výška stromů, výška nasazení zelené koruny) na pokusné ploše. Pro konstrukci a výpočet výškových křivek byly použity různé funkce s rozdílným počtem parametrů:

Logaritmická funkce $h=A1+A2*\ln(d)$

Naeslund $h=((d^2/A1+A2*d)^2) + 1,3$

Data: pok_hora.vts

Výpočet výškové křivky

Odhadování parametrů

Naeslundova funkce

Odhady parametrů	Parametr	Směr. odchylka	Dolní mez	Horní mez
P1	1,599	0,135	1,328	1,869
P2	0,142	0,005	0,133	0,152

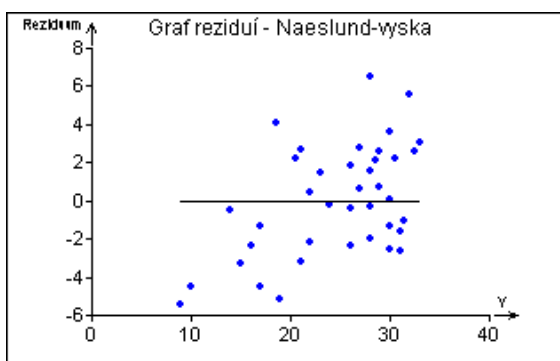
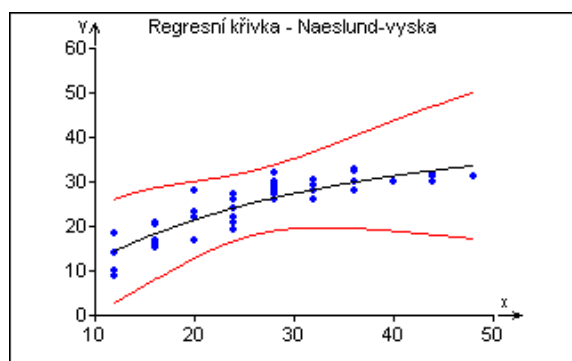
Logaritmická funkce

Odhady parametrů	Parametr	Směr. odchylka	Dolní mez	Horní mez
P1	-21,974	3,197	-28,409	-15,538
P2	14,507	0,998	12,497	16,516

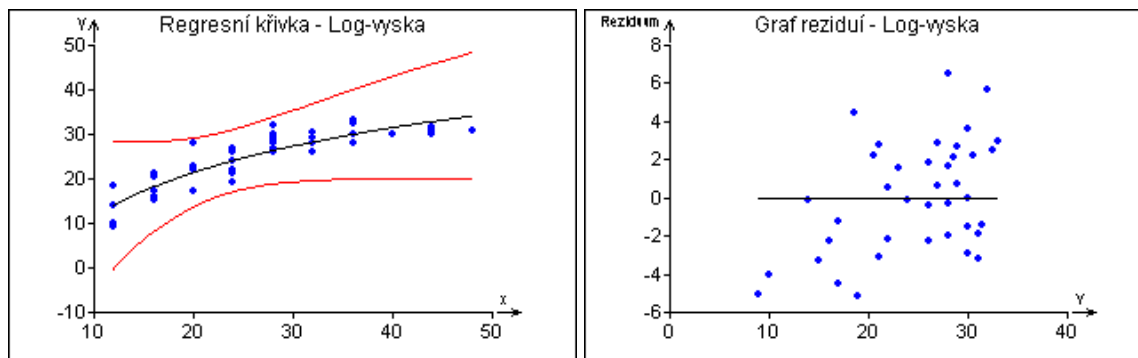
Korelační matice parametrů

	Naeslundova funkce	
	P1	P2
P1	1	-0,945
P2	-0,945	1

	Logaritmická funkce	
	P1	P2
P1	1	-0,993
P2	-0,993	1



Obr. 1 Regresní křivka a graf reziduí pro Naeslundovu funkci



Obr. 2 Regresní křivka a graf reziduí pro Logaritmicou funkci

Základní statistické charakteristiky

	Regresní rabat	AIK	Se	Šikmost	Špičatost
Logaritmická	82,1	97,216	2,697	0,023	2,706
Naeslund	82,3	96,717	2,683	0,004	2,807

Navržené výškové křivky prokládají data obdobně, základní statistické charakteristiky obou funkcí jsou srovnatelné. Odhady šikmosti i špičatosti reziduí odpovídají normálnímu rozdělení.

	Naeslundova funkce	Logaritmická funkce
Cook-Weisbergův test heteroskedasticity		
Hodnota kritéria CW :	1,807	1,139
Kvantil $\chi^2(1-\alpha,1)$:	3,842	3,842
Pravděpodobnost :	0,179	0,286
Závěr :	Homoskedasticita reziduí.	Homoskedasticita reziduí

Jarque-Berrův test normality		
Hodnota kritéria JB :	0,109	0,357
Kvantil $\chi^2(1-\alpha,2)$:	5,992	5,992
Pravděpodobnost :	0,947	0,837
Závěr :	Normální rozdělení reziduí	Normální rozdělení reziduí

Waldův test autokorelace		
Hodnota kritéria WA :	1,137	1,207
Kvantil $\chi^2(1-\alpha,1)$:	3,842	3,842
Pravděpodobnost :	0,179	0,286
Závěr :	Nevýznamná	Nevýznamná

Znaménkový test reziduí		
Hodnota kritéria Sg :	2,180	2,743
Kvantil $N(1-\alpha/2)$:	1,960	1,960
Pravděpodobnost :	0,029	0,006
Závěr :	V reziduích je trend!.	V reziduích je trend!

Výpočet křivky nasazení koruny

Odhady parametrů

Naeslundova funkce

Odhady parametrů	Parametr	Směr. odchylka	Dolní mez	Horní mez
P1	1,689	0,279	1,128	2,249
P2	0,212	0,011	0,190	0,233

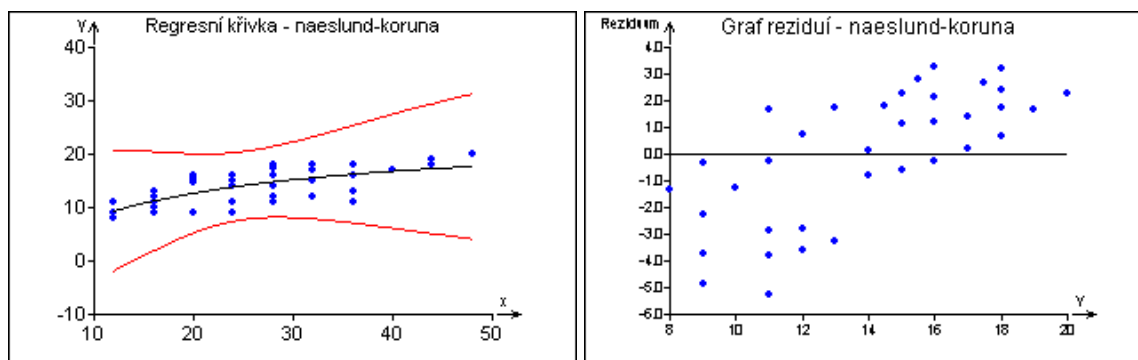
Logaritmická funkce

Odhady parametrů	Parametr	Směr. odchylka	Dolní mez	Horní mez
P1	-5,975	2,820	-11,652	-0,298
P2	6,214	0,881	4,442	7,987

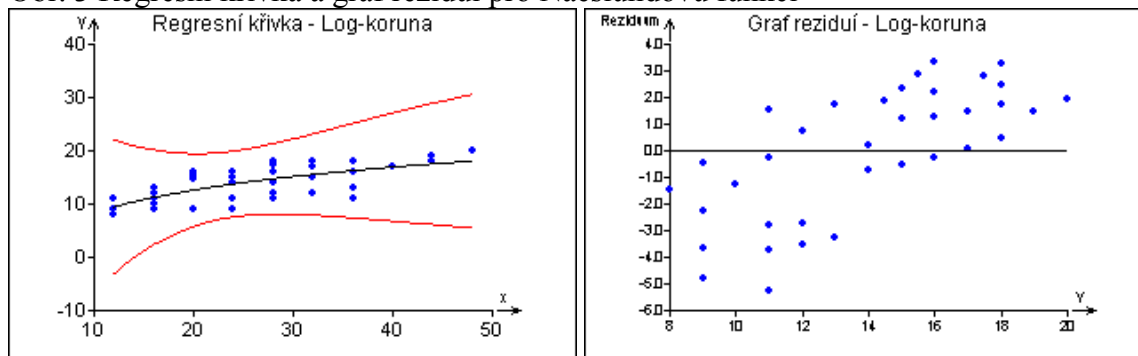
Korelační matice parametrů

	Naeslundova funkce	
	P1	P2
P1	1	-0,939
P2	-0,939	1

	Logaritmická funkce	
	P1	P2
P1	1	-0,993
P2	-0,993	1



Obr. 3 Regresní křivka a graf reziduí pro Naeslundovu funkci



Obr. 4 Regresní křivka a graf reziduí pro logaritmickou funkci

Základní statistické charakteristiky

	Regresní rabat	AIK	Se	Šikmost	Špičatost
Logaritmická	52,0	85,179	2,380	0,172	2,235
Naeslund	51,8	85,416	2,385	0,211	2,230

	Naeslundova funkce	Logaritmická funkce
Cook-Weisbergův test heteroskedasticity		
Hodnota kritéria CW :	1,807	1,139
Kvantil $\chi^2(1-\alpha,1)$:	3,842	3,842
Pravděpodobnost :	0,179	0,286
Závěr :	Homoskedasticita reziduí.	Homoskedasticita reziduí

Jarque-Berrův test normality		
Hodnota kritéria JB :	0,109	0,357
Kvantil $\chi^2(1-\alpha,2)$:	5,992	5,992
Pravděpodobnost :	0,947	0,837
Závěr :	Normální rozdělení reziduí	Normální rozdělení reziduí

Waldův test autokorelace		
Hodnota kritéria WA :	1,137	1,207
Kvantil $\chi^2(1-\alpha,1)$:	3,842	3,842
Pravděpodobnost :	0,179	0,286
Závěr :	Nevýznamná	Nevýznamná

Znaménkový test reziduí		
Hodnota kritéria Sg :	2,180	2,743
Kvantil $N(1-\alpha/2)$:	1,960	1,960
Pravděpodobnost :	0,029	0,006
Závěr :	V reziduích je trend!.	V reziduích je trend!

	Naeslundova funkce	Logaritmická funkce
Cook-Weisbergův test heteroskedasticity		
Hodnota kritéria CW :	1,786	1,350
Kvantil $\chi^2(1-\alpha,1)$:	3,842	3,842
Pravděpodobnost :	0,182	0,245
Závěr :	Homoskedasticita reziduí.	Homoskedasticita reziduí

Jarque-Berrův test normality		
Hodnota kritéria JB :	2,873	2,545
Kvantil $\chi^2(1-\alpha,2)$:	5,992	5,992
Pravděpodobnost :	0,238	0,280
Závěr :	Normální rozdělení reziduí	Normální rozdělení reziduí

Waldův test autokorelace		
Hodnota kritéria WA :	0,721	0,602
Kvantil $\chi^2(1-\alpha,1)$:	3,842	3,842
Pravděpodobnost :	0,182	0,245
Závěr :	Nevýznamná	Nevýznamná

Znaménkový test reziduí		
Hodnota kritéria Sg :	1,035	1,035
Kvantil $N(1-\alpha/2)$:	1,960	1,960
Pravděpodobnost :	0,301	0,301

Závěr :	V reziduích není trend	V reziduích není trend.
---------	------------------------	-------------------------

Závěr:

Regresní rabat pro výpočet výšky stromu je výrazně vyšší než rabat pro výpočet nasazení koruny. Použité modely pro výškový růst mají srovnatelné výsledky.