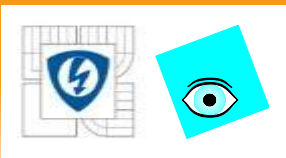


Statistická analýza jednorozměrných dat

Prof. RNDr. Milan Meloun, DrSc.
Univerzita Pardubice, Pardubice

31.ledna 2011

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.



Kapitola 2.4

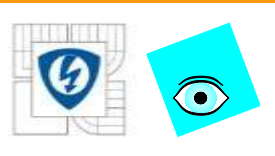
Řešení praktických úloh z Kompendia

EXPLORATORNÍ ANALÝZA DAT

31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





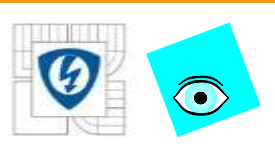
Postup analýzy dat

1. Průzkumová analýza dat:

Diagnostické grafy: stupeň symetrie a rozdělení
 lokální koncentrace dat
 vybočující data

2. Ověření předpokladů výběru dat:

Diagnostiky, testy: ověření normality
 ověření nezávislosti
 ověření homogenity
 určení minimální četnosti

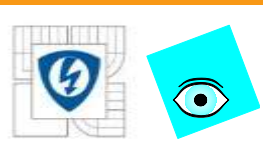


3. Transformace dat:

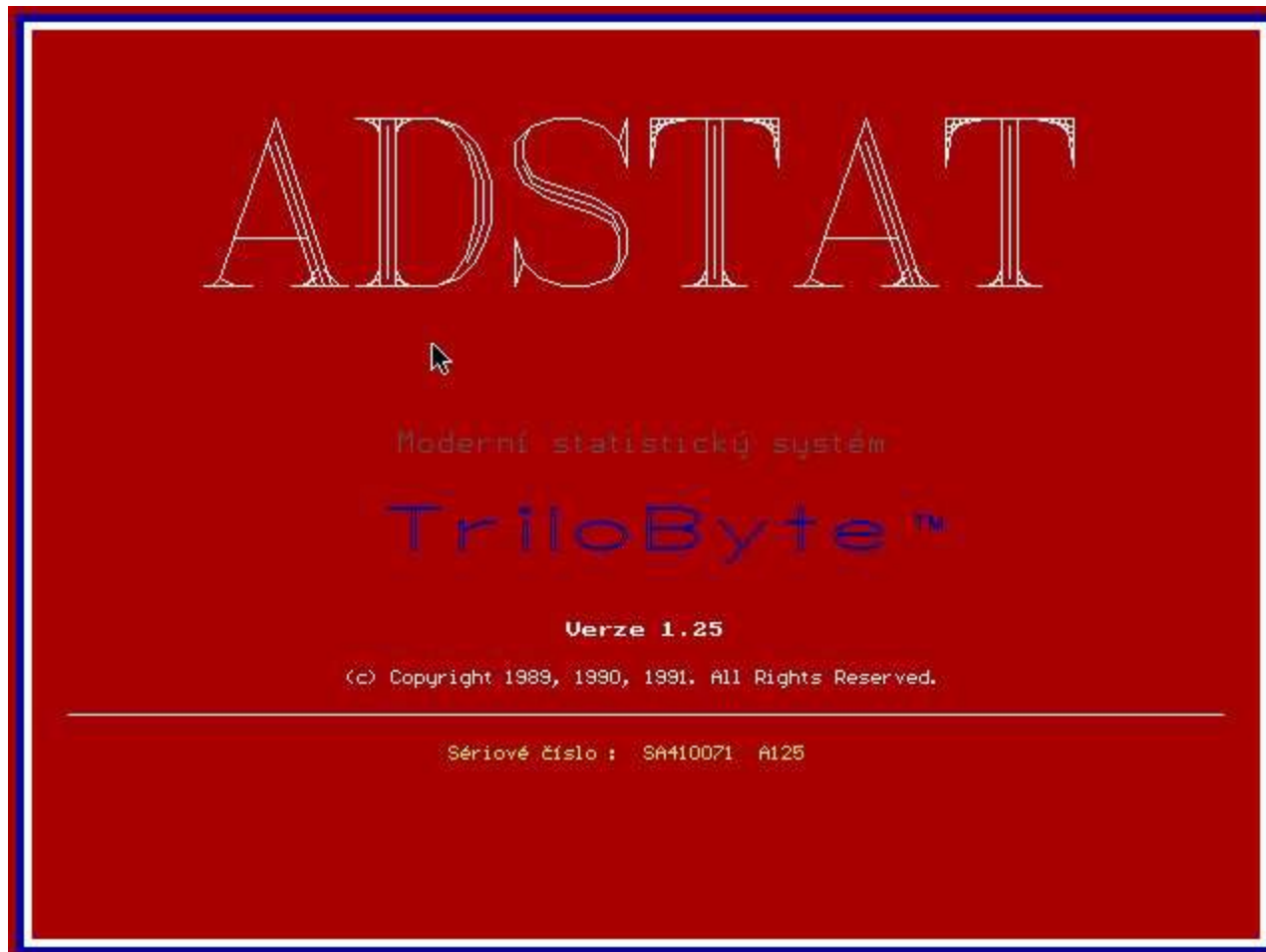
Analýza dat: originální data
 data po mocninné transformaci
 data po Box-Coxově transformaci

4. Parametry polohy, rozptýlení a tvaru:

Analýza 1 výběru: klasické odhady – průměr, rozptyl
 robustní odhady – medián, uřezané
 průměry, winsorizovaný
 rozptyl, interkvantilové rozpětí
 adaptivní odhady



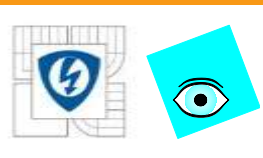
ADSTAT



31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





ADSTAT





ADSTAT

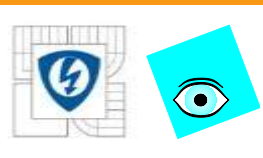
F1=nápověda F2=uložení F3=čtení F4=numerický mód F5=zoom ESC=konec
Počet řádků: 7 Počet sloupců: 2 Uodorovně demo.dat

Numerický editor

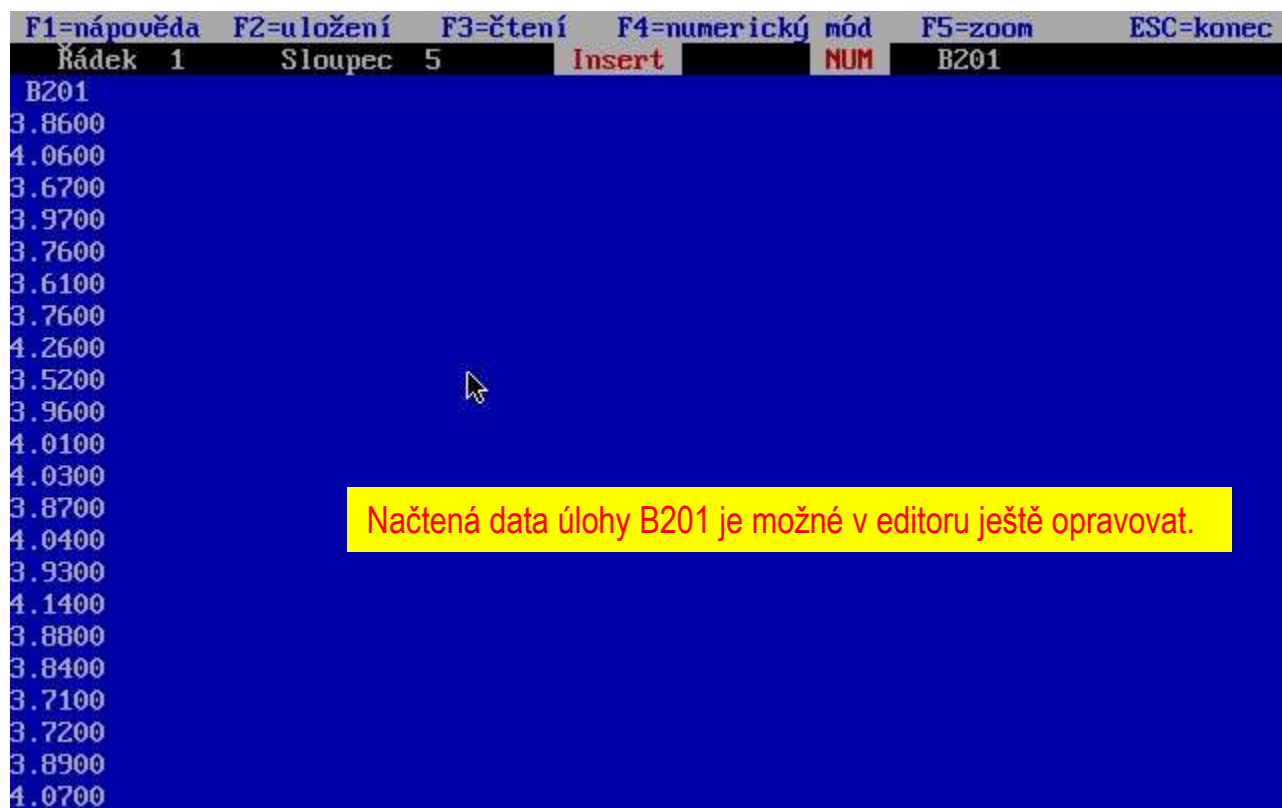
	1	2
1	6.000000	6.000000
2	Čtení ze souboru: B201	
3		
4		
5		
6	7.000000	8.000000
7	9.000000	1.000000 E+01
8	1.500000 E+01	1.200000 E+01

Načtení dat po F3.

značení bloku: Z=začátek K=konec B=buňka X=zrušení/obnovení
S=sloupec R=řádek N=nová poloha MEZERA=změna směru
blokové operace: F6=kopírování F7=zaplnění F8=transformace
ostatní operace: F9=změna velikosti matice F10=transpozice



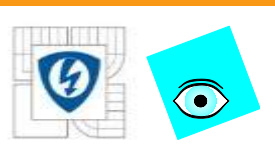
ADSTAT



31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

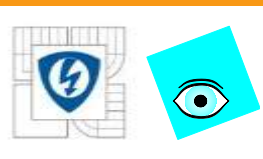




Průzkumová analýza dat

Diagnostické grafy: stupeň symetrie rozdělení
lokální koncentrace dat
vybočující data





Průzkumová analýza dat

F1-nápověda
Řádek 1
B201
3.8600
4.0600
3.6700
3.9700
3.7600
3.6100
3.7600
4.2600
3.5200
3.9600
4.0100
4.0300
3.8700
4.0400
3.9300
4.1400
3.8800
3.8400
3.7100
3.7200
3.8900
4.0700

Základní Statistika: Analýza jednoho výběru

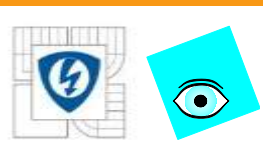
Data	Metoda	Zadání	Výpočet	Výsledky	Graf	Konec
Analýza 1 výběru Porovnání 2 výběrů Porovnání rozdělení Základní předpoklady Mocninná transformace Exploratorní analýza Šíření chyb Spojité rozdělení Diskrétní rozdělení						

Volba exploratorní analýzy pro spojité rozdělení.

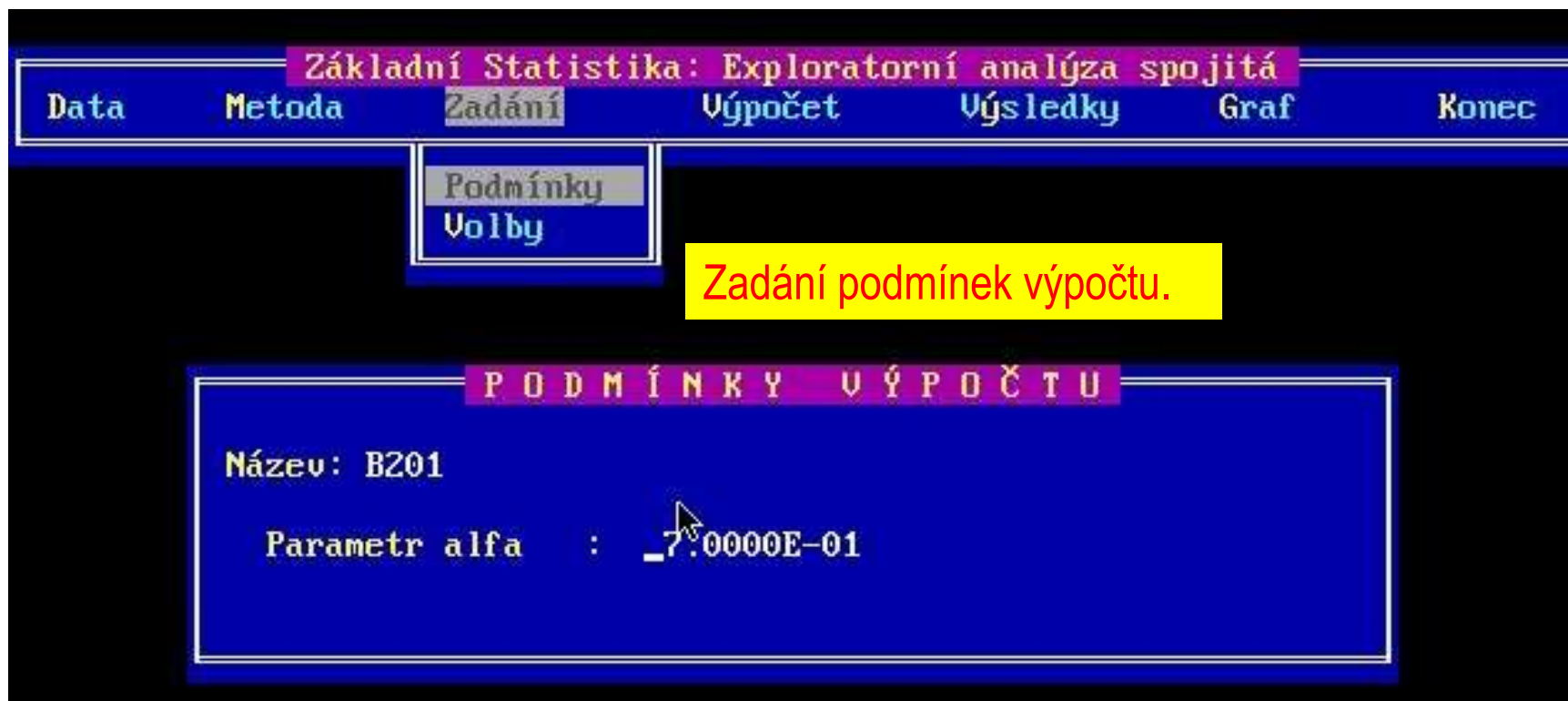
31.1.2011

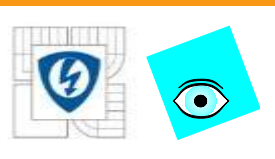
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





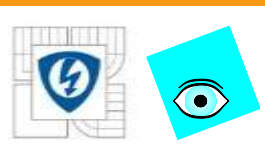
Průzkumová analýza dat



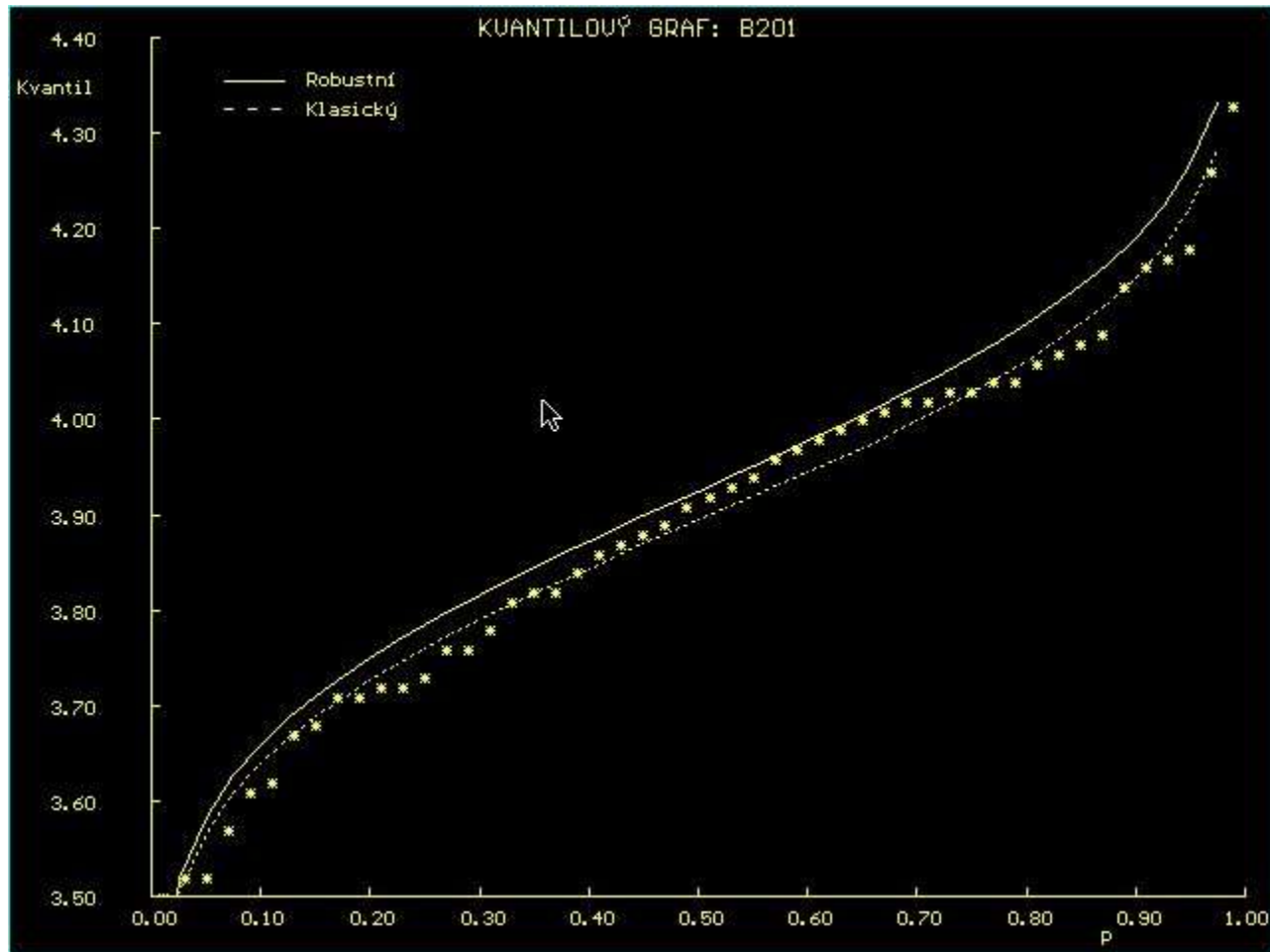


Průzkumová analýza dat





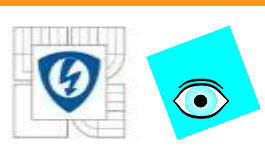
Kvantilový graf ukazuje na symetrii a odlehlé body.



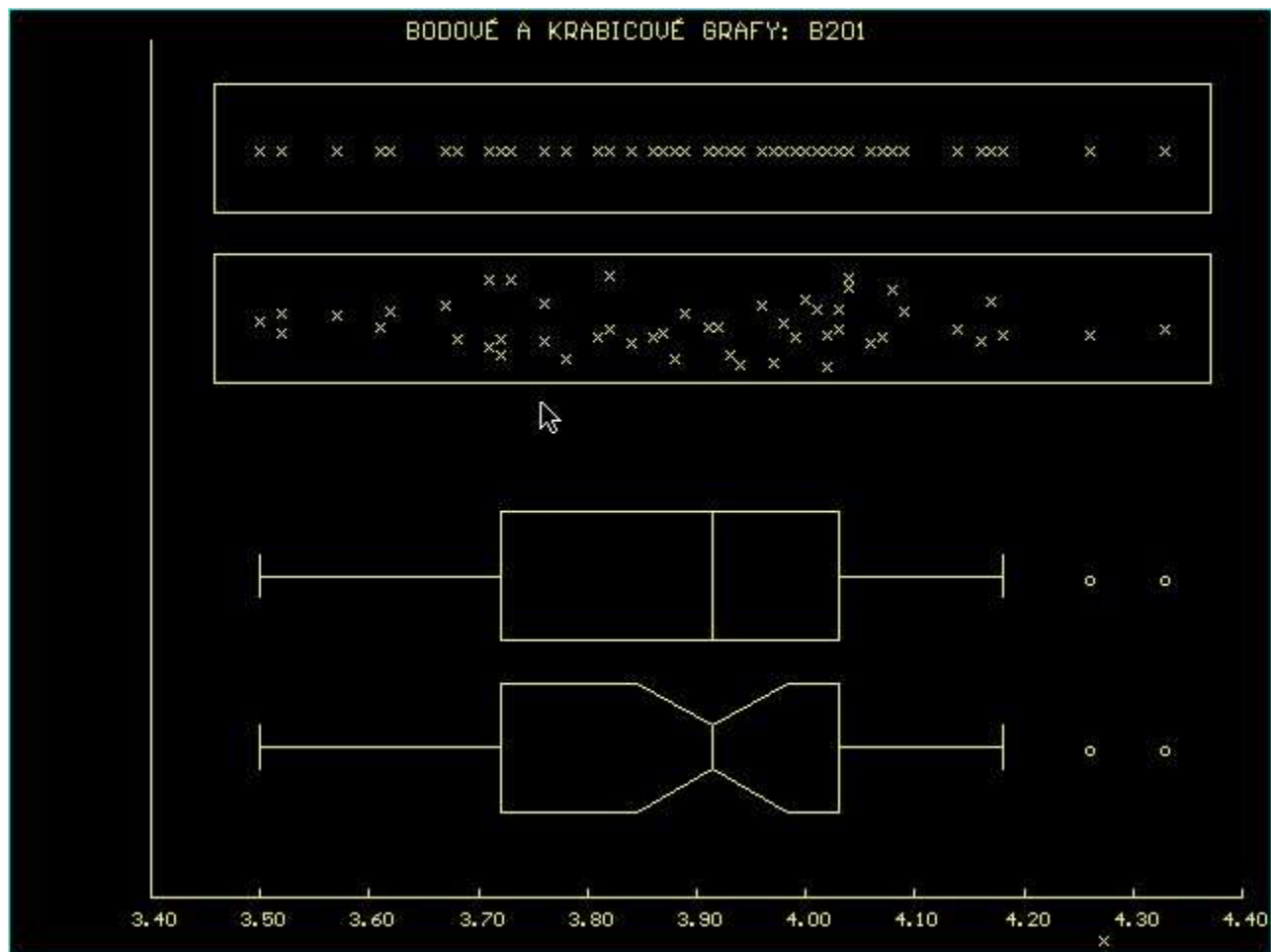
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





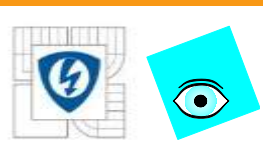
Kvantilový graf ukazuje na symetrii a žádné odlehlé body.



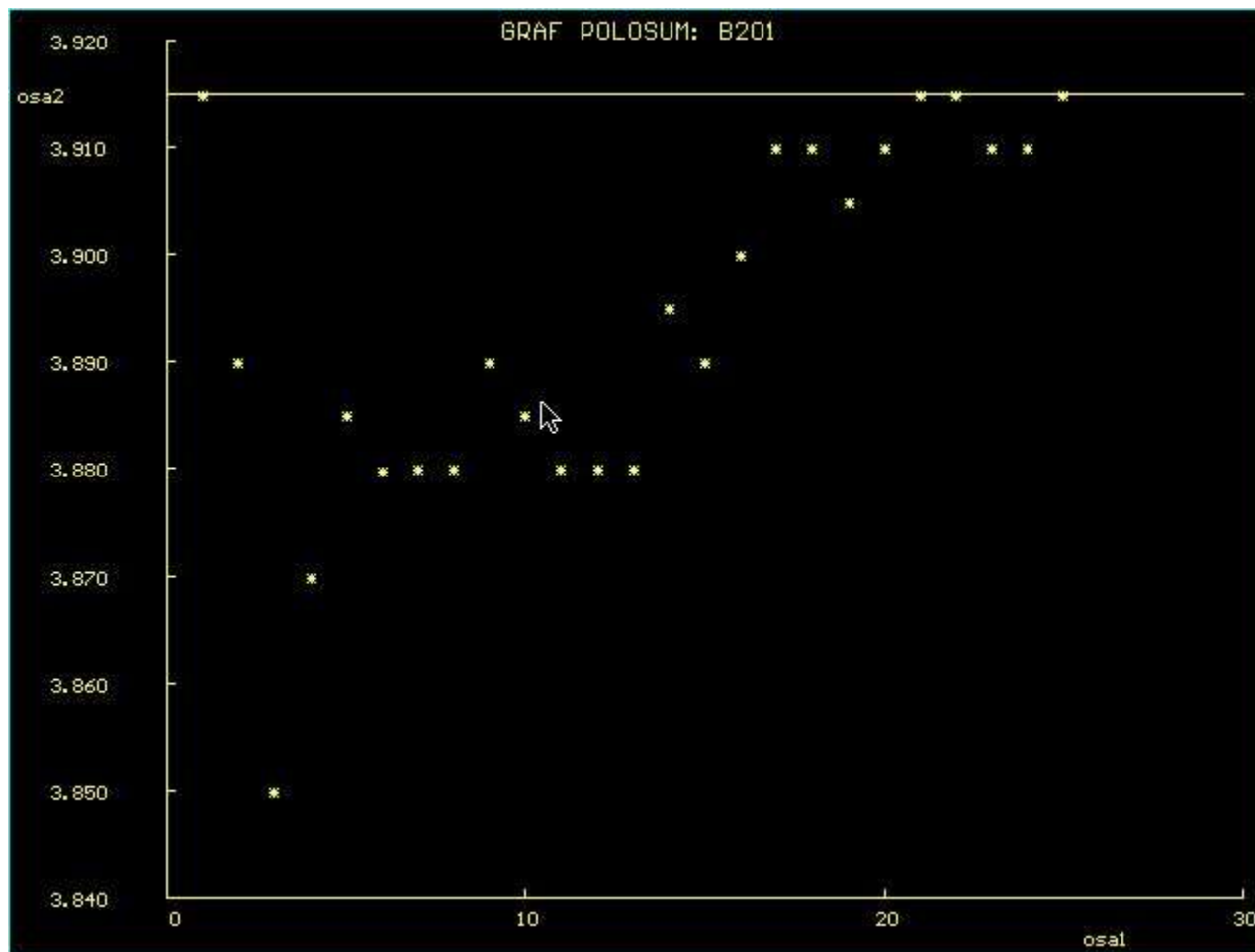
31.1.2011

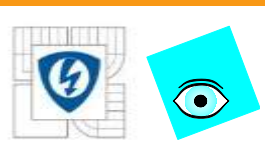
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



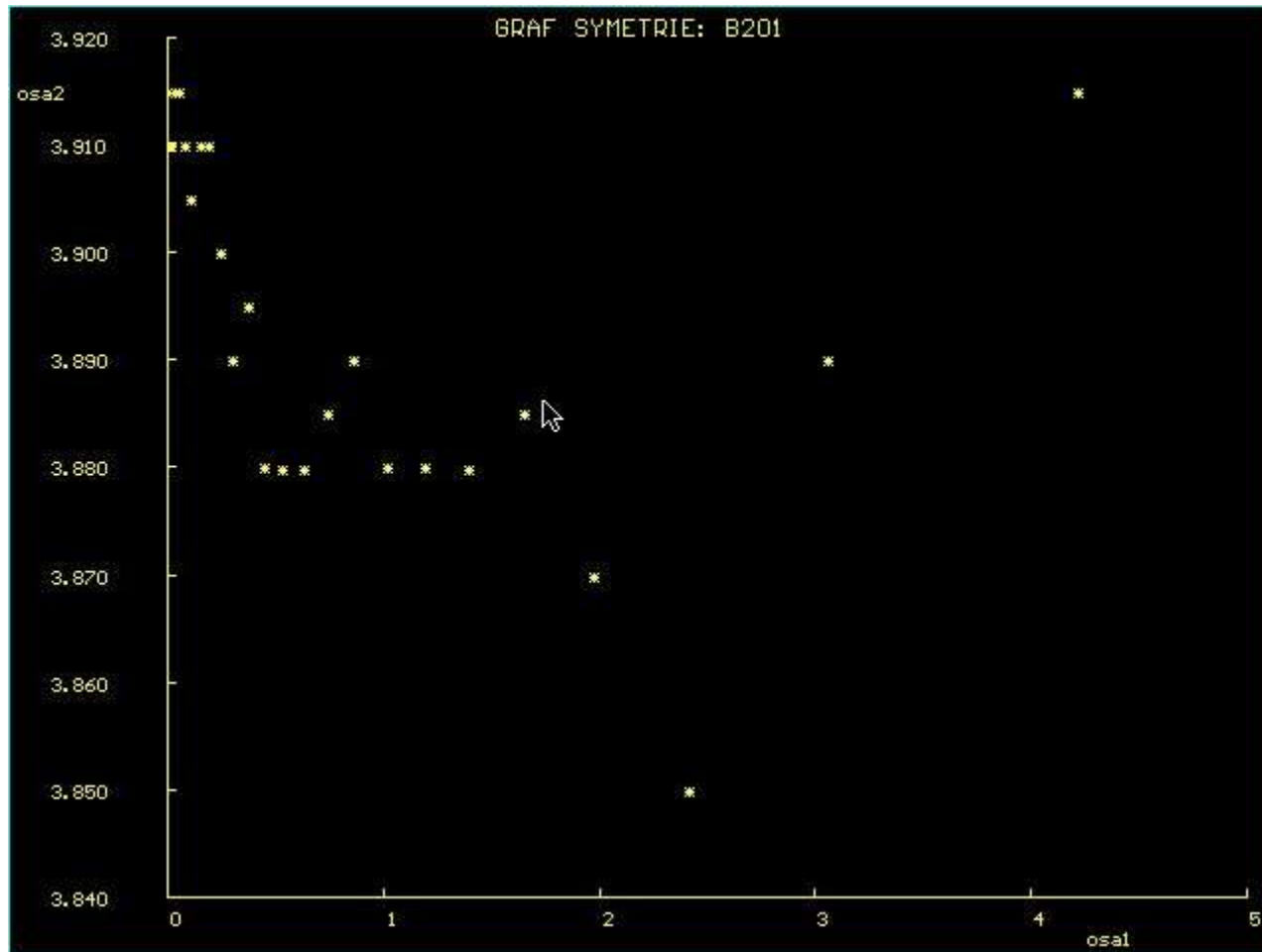


Velký zoom bodů u obou grafů soustředí body blízko rovnoběžné přímky a dokazuje tím symetrii rozdělení.





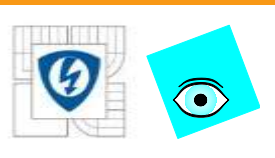
Velký zoom bodů u obou grafů soustředí body blízko rovnoběžné přímky a dokazuje tím symetrii rozdělení.



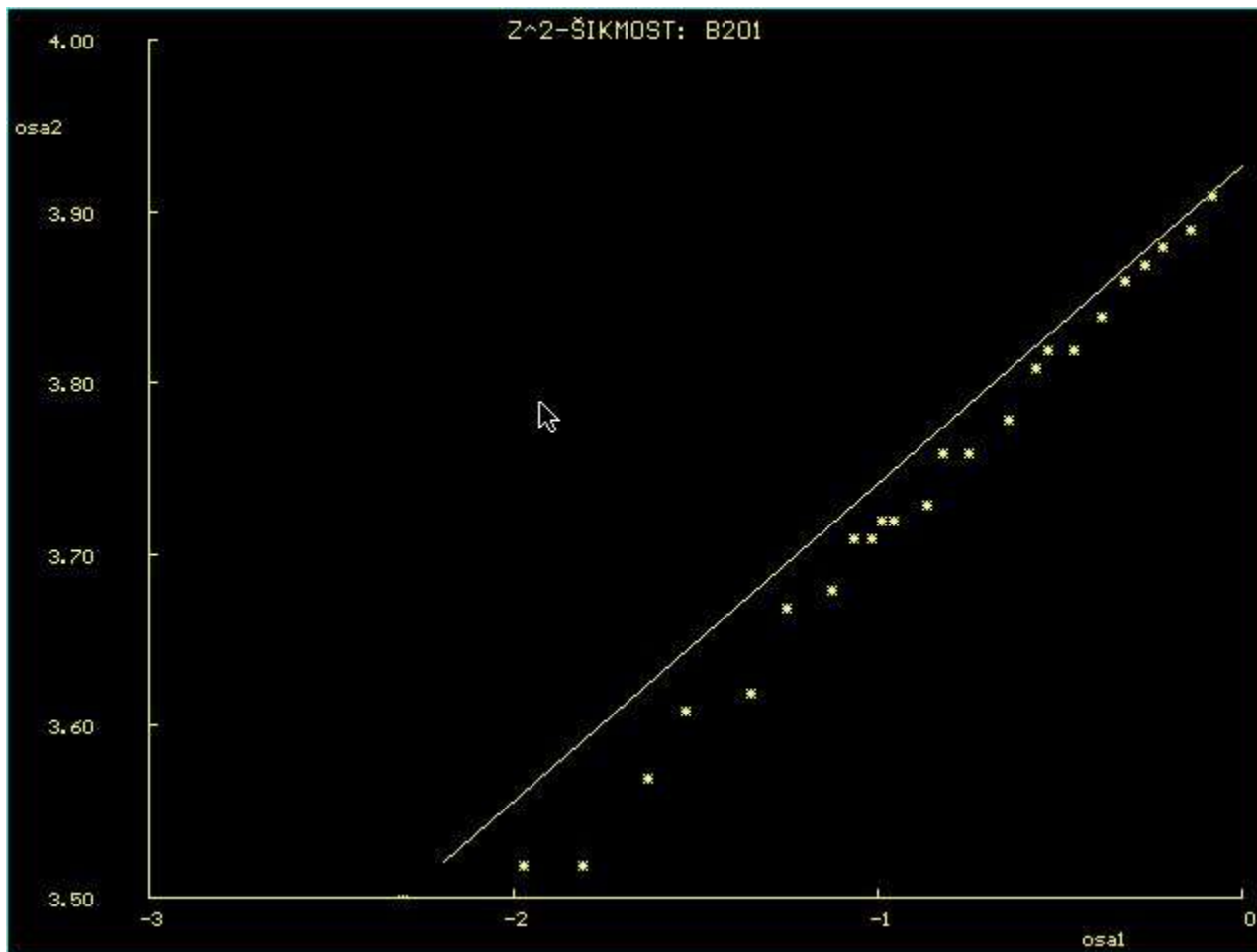
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





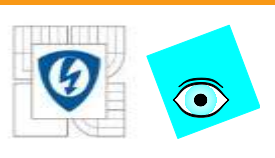
Graf šikmosti ukazuje na symetrii, protože body leží téměř na přímce vytyčené pro Gaussovo rozdělení.



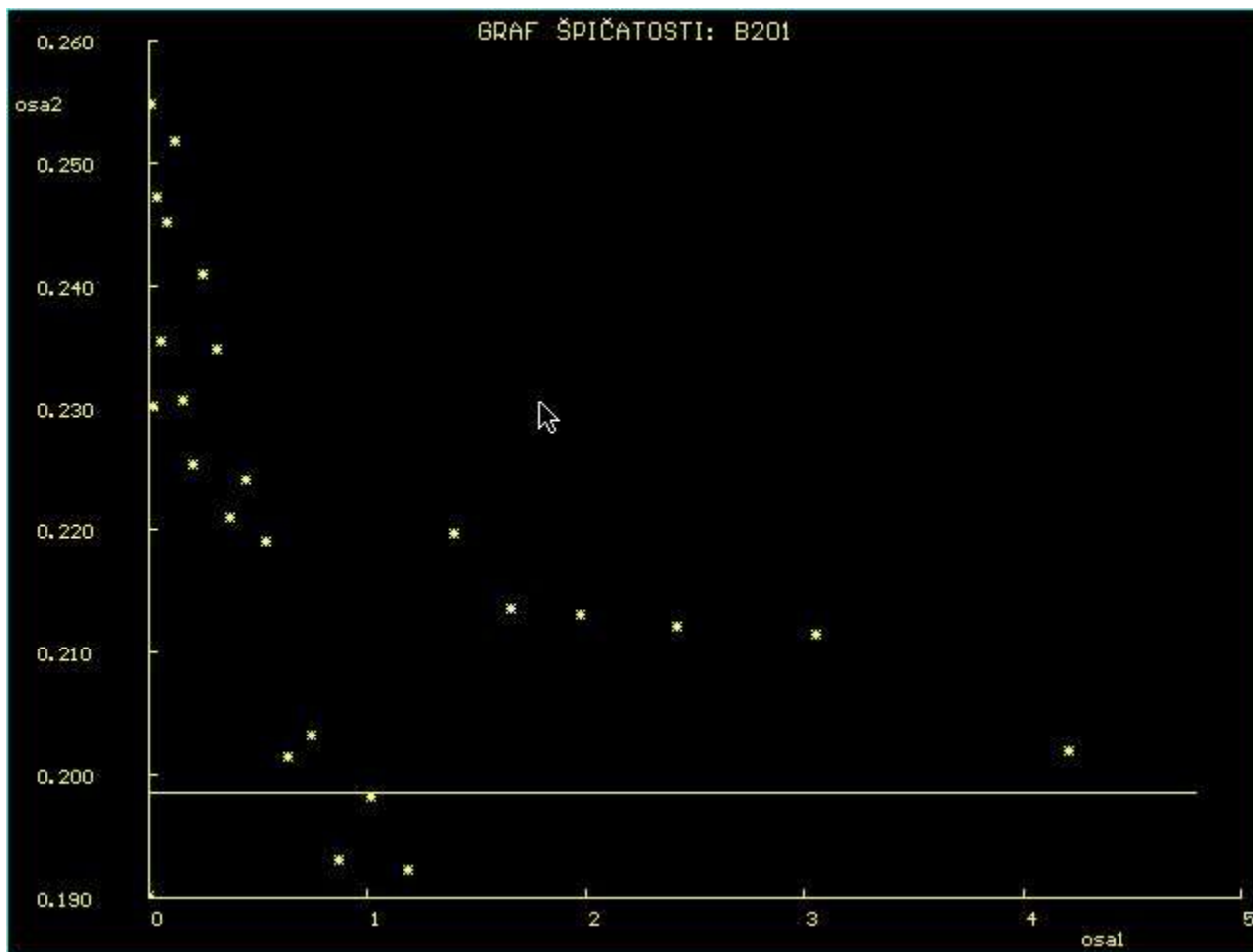
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





Graf špičatosti ukazuje na hodnotu špičatosti 0.2.



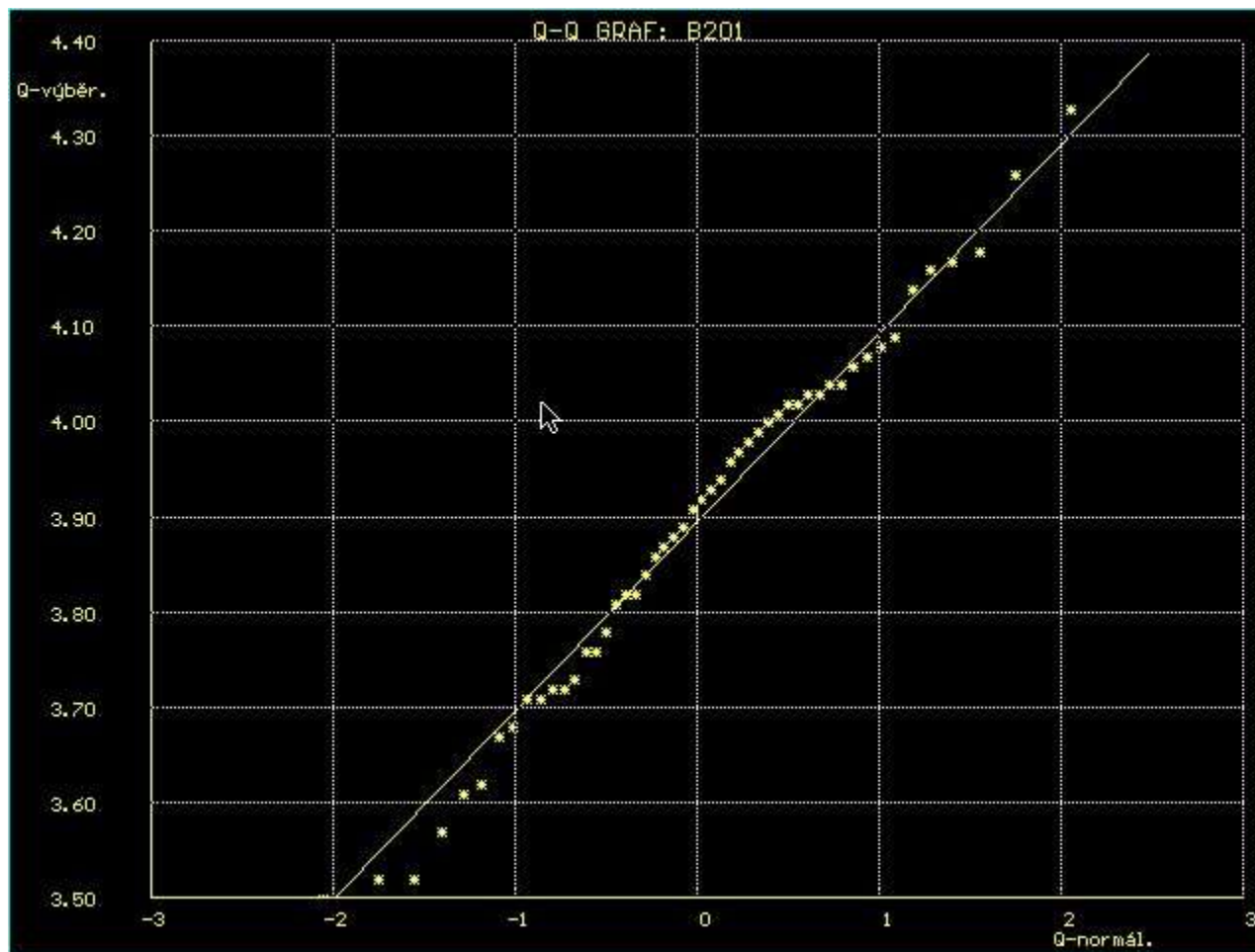
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





Oba Q-Q grafy ukazují na symetrii a nahoře 2 a dole 5 odlehlých bodů. Průsečík s kolmicí v bodě $x=0$ značí aritm. průměr.



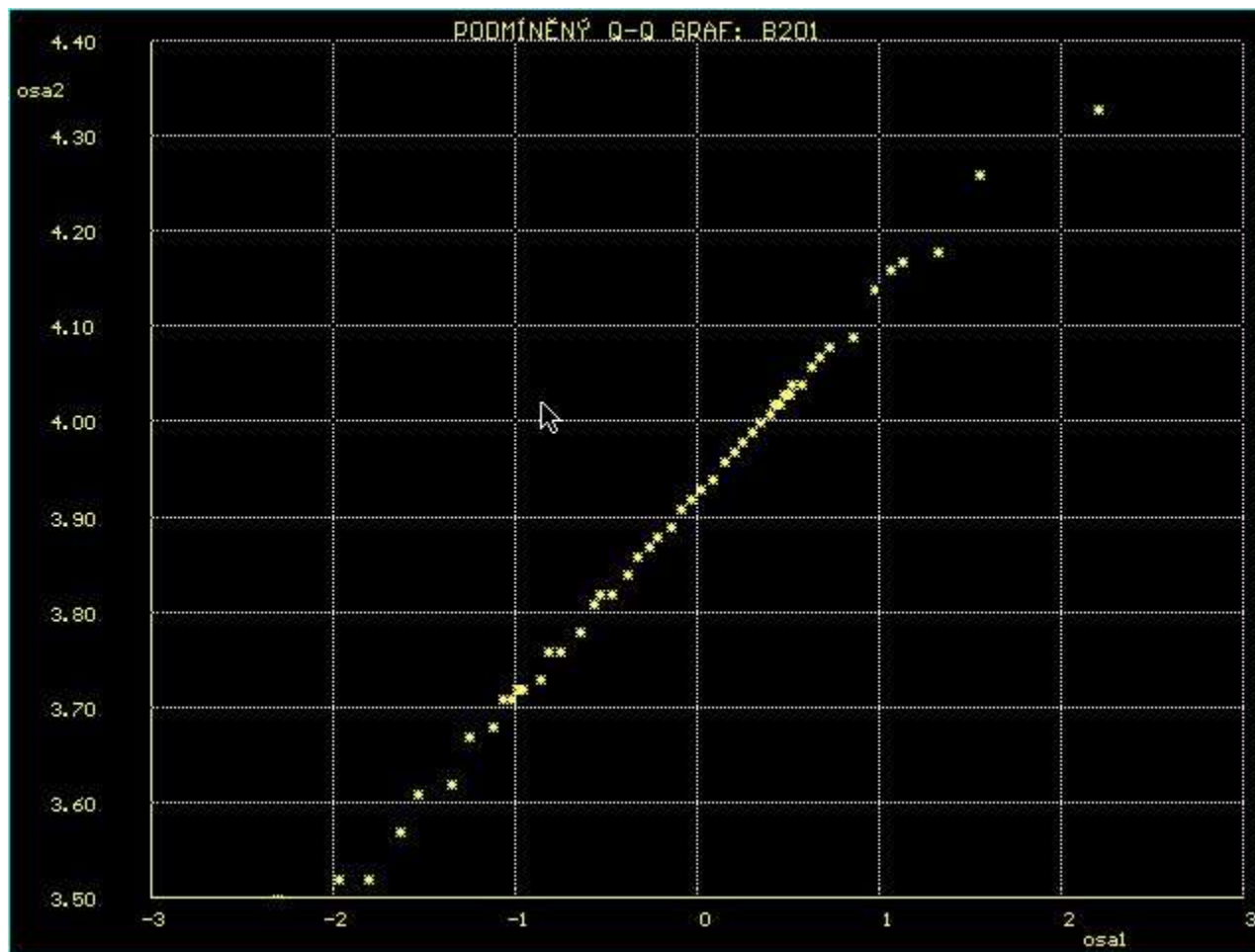
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





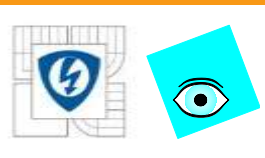
Oba Q-Q grafy ukazují na symetrii a nahoře 2 a dole 5 odlehlých bodů. Průsečík s kolmicí v bodě $x=0$ značí aritm. průměr.



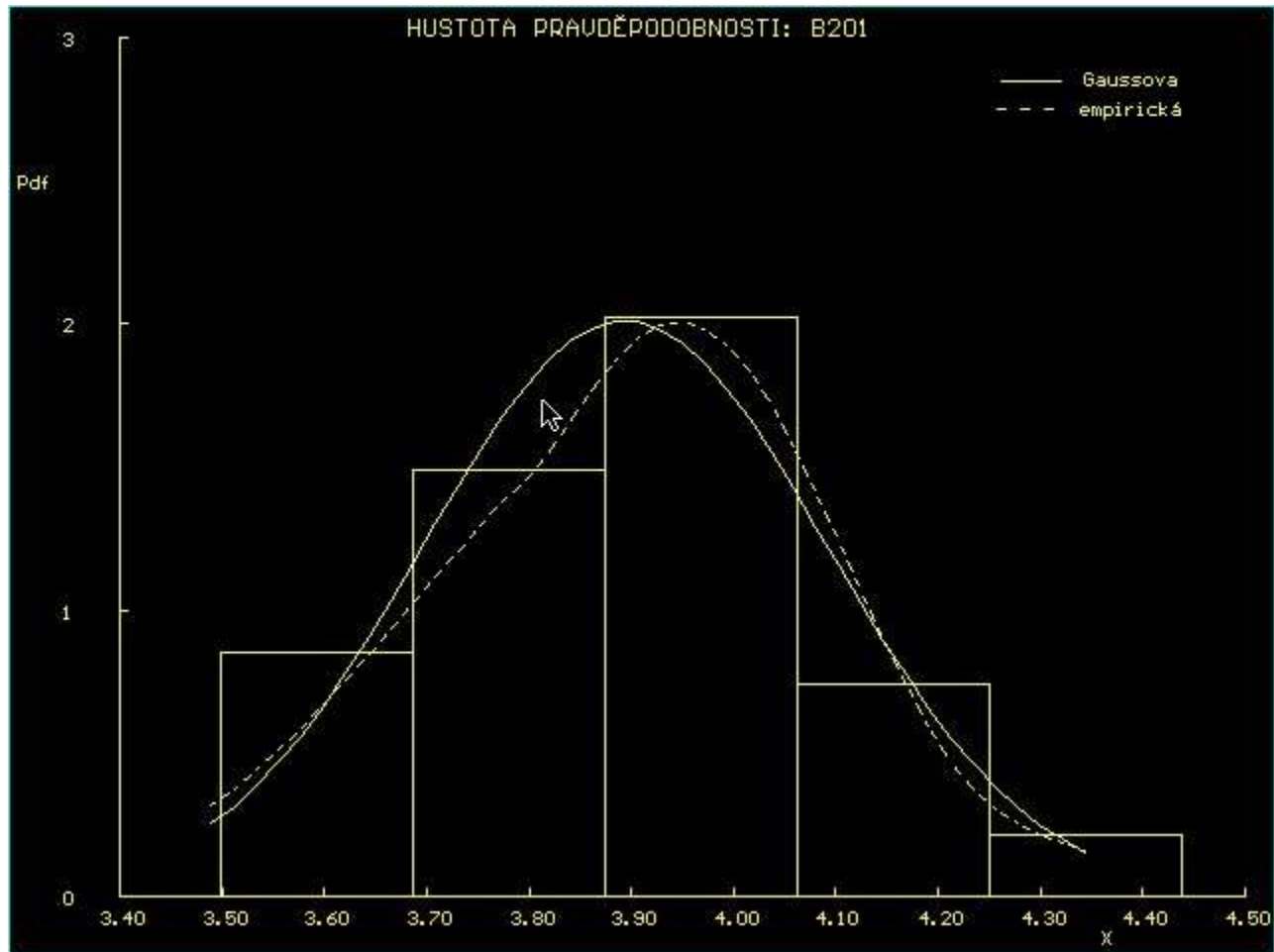
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





Graf hustoty pravděpodobnosti ukazuje na symetrii. Kolmice z vrcholu empirické je medián a z vrcholu Gaussovy je průměr.



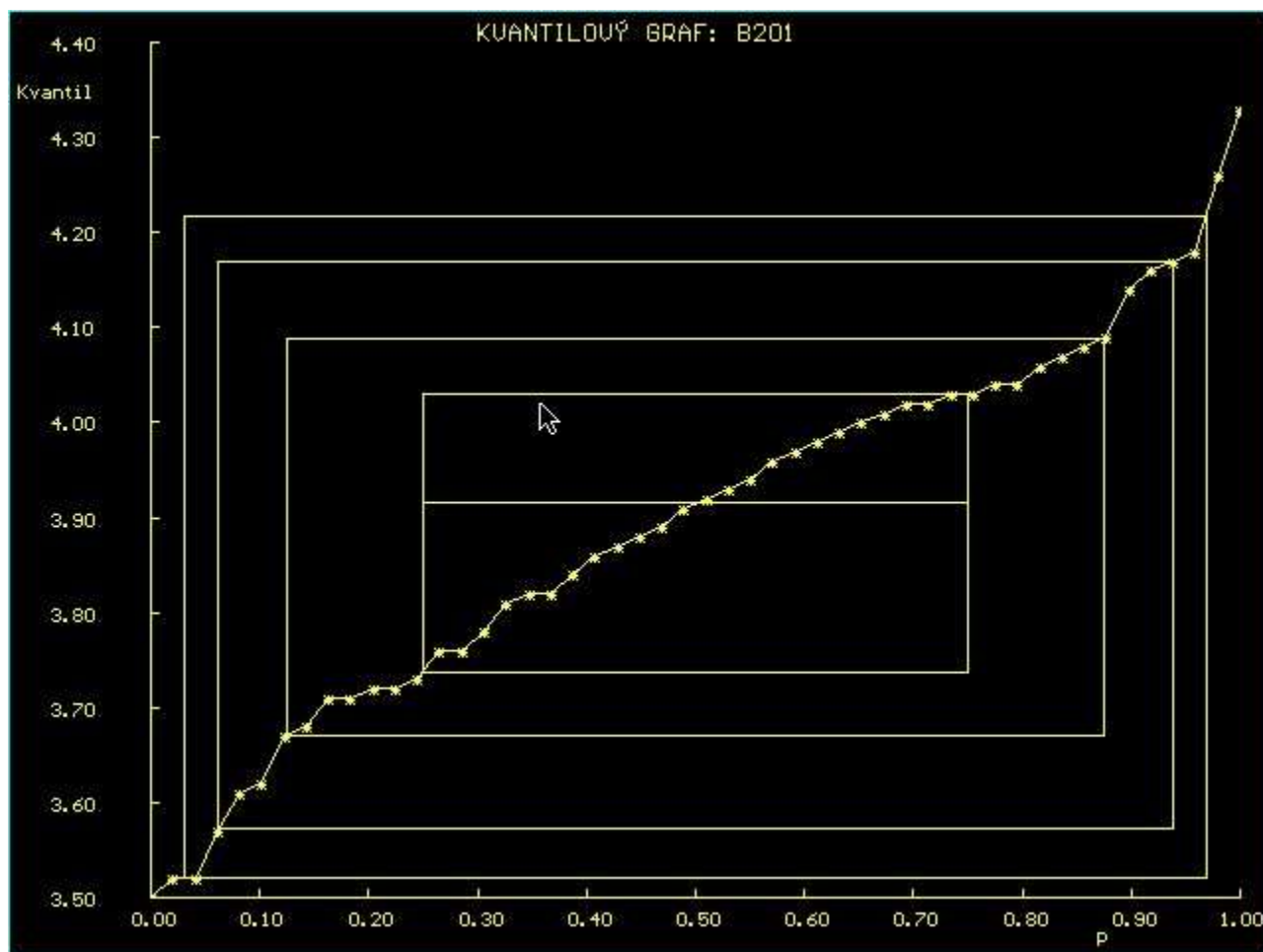
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





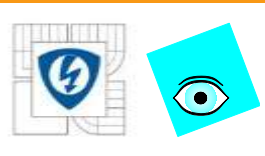
Graf rozptýlení s kvantily vykazuje symetricky umístěné obdélníky a tím dokazuje symetrii rozdělení..



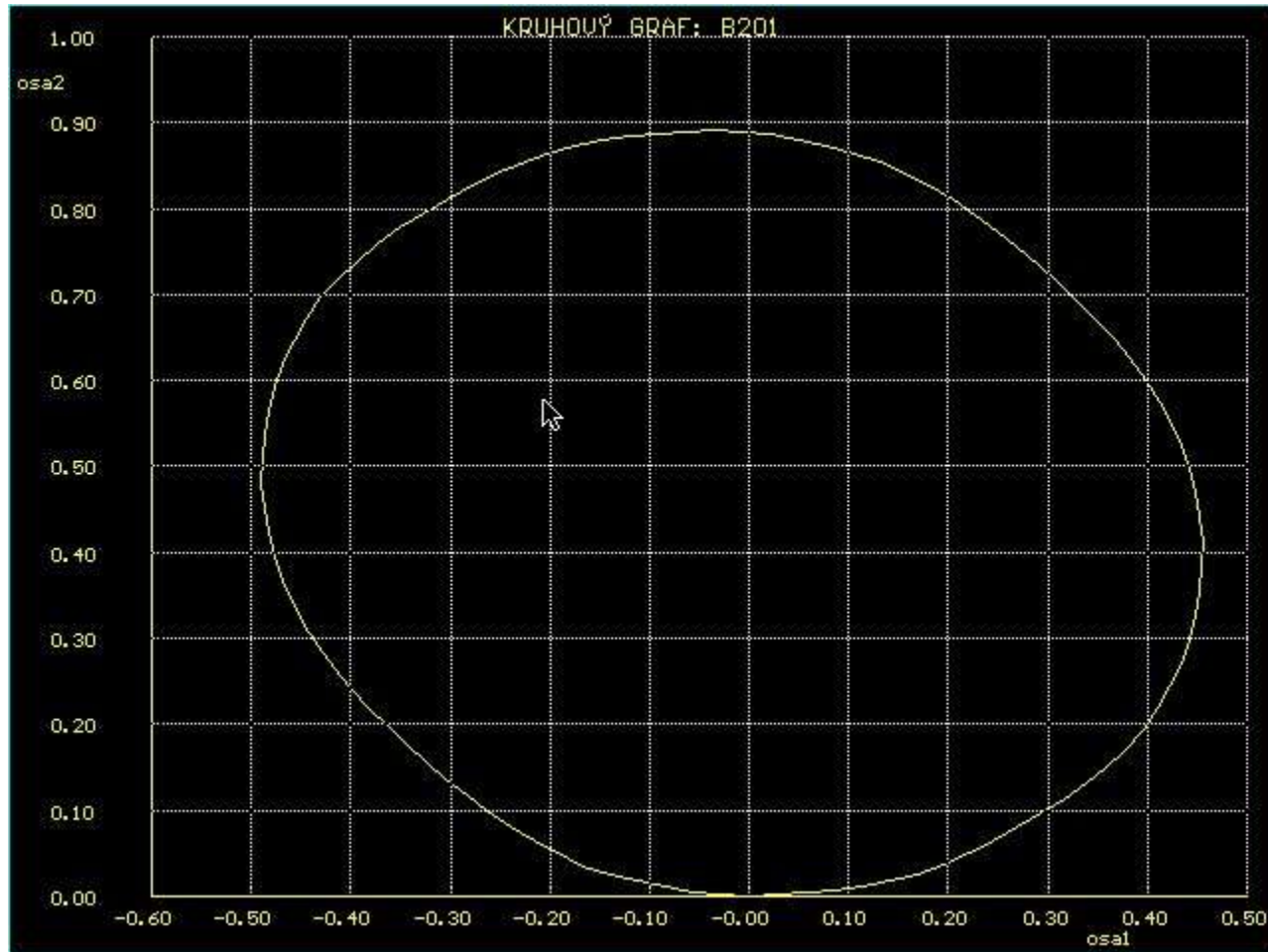
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





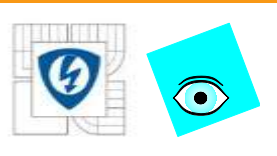
Kruhový grafem je elipsa s x-osou rovnoběžná, a proto je rozdělení symetrické, rovnoměrné až Gaussovo.



31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



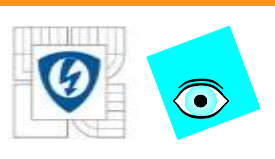


Závěr EDA

1. Rozdělení je symetrické (mezi rovnoměrným a Gaussovým).
2. Rozdělení obsahuje nahoře 2 a dole cca 5 odlehlých bodů – outlierů.
3. Není asi nutné užít transformaci dat.



31.1.2011



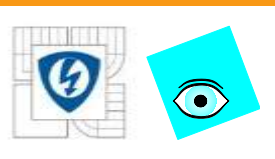
Načtení druhé kapitoly.



31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

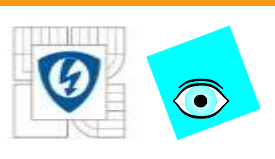




Volba výpočtů Základní statistika.

	A1	3.86							
		B201	B202	B203					
1		3.86	0.6544	49.6					
2		4.06	0.6121	47.9					
3		3.67	0.6438	50.0					
4		3.97	0.6510	50.1					
5		3.76	0.6592	50.1					
6		3.61	0.6525	50.1					
7		3.76	0.6515	50.2					
8		4.26	0.6545	50.2					
9		3.52	0.6519	50.2					
10		3.96	0.6504	50.2					
11		4.01	0.6416	50.3					
12		4.03	0.6656	50.3					
13		3.87	0.6626	55.3					
14		4.04	0.6546	50.3					
15		3.93	0.6342	50.3					
16		4.14	0.6413	50.4					
17		3.88	0.6588	50.4					
18		3.84	0.6531	50.4					
19		3.71	0.6461	50.5					
20		3.72	0.6196	50.5					
21		3.89		50.5					
22		4.07		50.72	0.0690	6	4.24	20.12	79.99
23		3.82		50.92	0.6800	0	3.30	20.09	79.68
24		4.33		50.40	3.2100	4	5.22	19.82	80.16
25		4.00			2.5500	2	5.61	19.99	79.82
26		3.99			0.2100	1	5.57	19.96	79.64
27		4.02			0.4100	1	4.49	20.00	80.17
28		3.82			0.0800	4	5.47	19.98	80.30
29		3.62			1.5100	0	3.96	20.16	79.72
30		3.68			0.1300	1	5.05	19.94	80.08

Základní statistika...	
Porovnání dvou výběrů...	
Pravděpodobnostní modely...	
Transformace...	B207b
Pravděpodobnostní kalkulátor...	80.18
Testování	80.07
Simulace	79.56
ANOVA	80.28
Plánované experimenty	80.02
Responsní povrch...	79.85
	80.15
Vicerozměrné metody	79.76
Regrese	79.96
Kalibrace...	80.03
Partial Least Squares...	79.89
	80.32
Shewhartovy diagramy	79.92
Rozšířené diagramy	79.96
Způsobilost...	80.25
Paretův diagram...	80.12
Přejímka	80.01
	79.81
Grafy...	80.29
Rychlé zobrazení dat...	80.20
	80.08



Zadání výpočtu této úlohy.

Název úlohy : B201

Řád trendu 4

Testuj hodnotu 0

Vyhlazení hustoty 0.5

Řád autokorelace 4

Hladina významnosti 0.05

Sloupce

- B201
- B202
- B203
- B204
- B205
- B206
- B207a

K výpočtu použij

- ☐ Všechna data
- ☒ Sloupce
- ☐ Průměry podskupin

Data

- ☒ Všechna
- ☐ Označená
- ☐ Neoznačená
- ☐ Podle filtru

Popis [Žádný]

Grafy

- Autokorelace
- Trendy vyhlazení
- Kvantilový graf
- PP graf
- Graf rozptýlení s kvantily
- Graf polosum
- Graf symetrie
- Graf spícatosti
- Kruhový graf

Protokoly

- Klasické parametry
- Robustní parametry
- Test normality
- Vybočující body
- Autokorelace
- Významnost trendu
- Vyhlazené hodnoty
- Rezidua

Časová osa

Časová osa

Vyber vše

Doporučené

Nápověda

Použít

Zpět

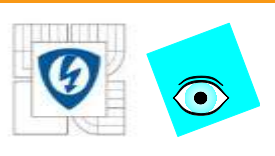
OK

Mění

Standardní

Všechny grafy

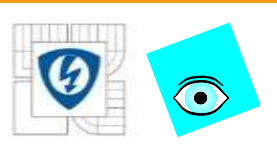
Všechny protokoly



Vysvětlení

Data a parametry: Data jsou organizována do sloupců. V prvním řádku jsou názvy sloupce. Jsou-li ve vstupním dialogovém panelu vybrány “Všechna data” nebo “Sloupce”, mohou být délky jednotlivých sloupců různé. Má-li se výpočet provést pro “Průměry podskupin, měly by všechny sloupce mít stejnou délku. Minimální počet sloupců je 1. Minimální počet dat je 2.

Řád trendu určuje z kolika po sobě jdoucích dat budou počítány klouzavé průměry a klouzavé mediány. Hodnota by měla být menší než polovina počtu dat.

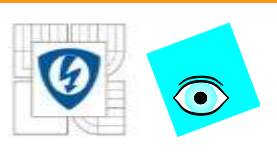


Vysvětlení

Testuj hodnotu zadává se hodnota pro T-test. Program testuje na zadané hladině významnosti, zda tato hodnota může být shodná se střední hodnotou dat.

Vyhlazení hustoty udává šířku jádra pro jádrové vyhlazení pro graf hustoty pravděpodobnosti. Čím je větší, tím bude křivka hustoty pravděpodobnosti hladší a naopak. Hodnota musí být větší než nula.

Řád autokorelace udává do kterého řádu se budou počítat autokorelační koeficienty. Hodnota musí být alespoň o 2 menší než počet platných dat.



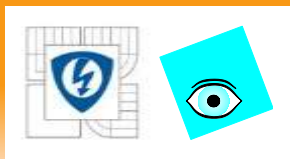
Vysvětlení

Hladina významnosti udává spolehlivost pro intervaly spolehlivosti a statistické testy. Musí být větší než nula a menší než 0.5. Vynásobena 100 udává hodnotu v procentech. Obvyklá hodnota je 0.05 (tedy 5%).

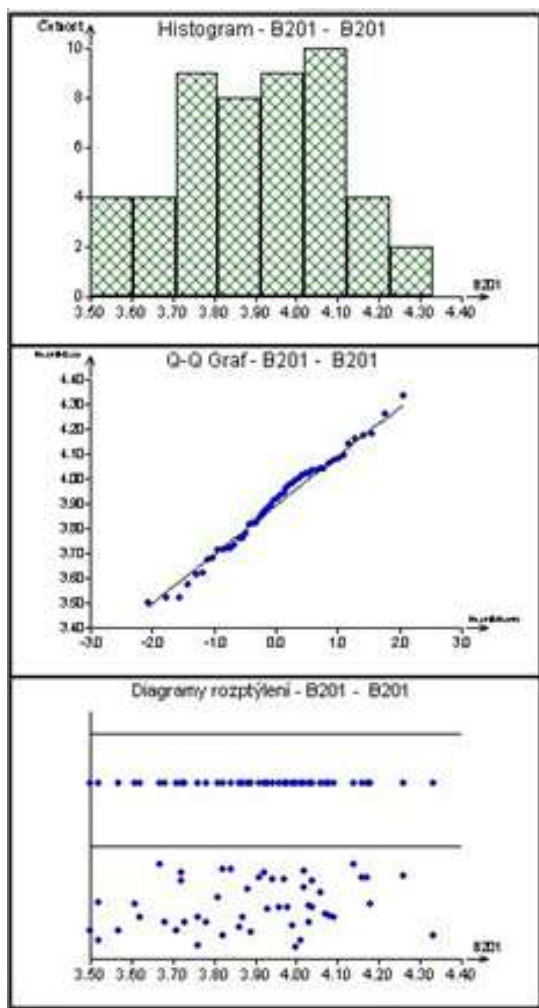
Všechna data všechny vybrané sloupce budou brány jako jediný sloupec.

Sloupce Výpočet se provede pro každý z vybraných sloupců zvlášť.

Průměry podskupin Výpočet se provede pro řádkové průměry z vybraných sloupců. Pokud nejsou sloupce stejně dlouhé, nebo obsahují chybějící data, provede se výpočet jen pro úplné řádky. Tento výpočet má význam především pro diagnostiku dat pro regulační diagramy typu x-průměr.



Grafy výstupu Exploratorní analýzy a jejich vysvětlení.

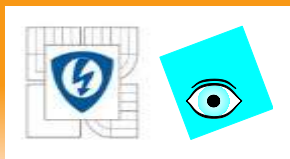


Histogram: nejstarší diagram rozdělení ukazuje na symetrii rozdělení. Histogram četností dat v jednotlivých třídách s konstantní šířkou; optimální počet tříd je stanovován automaticky s ohledem na počet dat.

Q-Q graf: kvantil-kvantilový graf vykazuje většinu bodů na přímce.

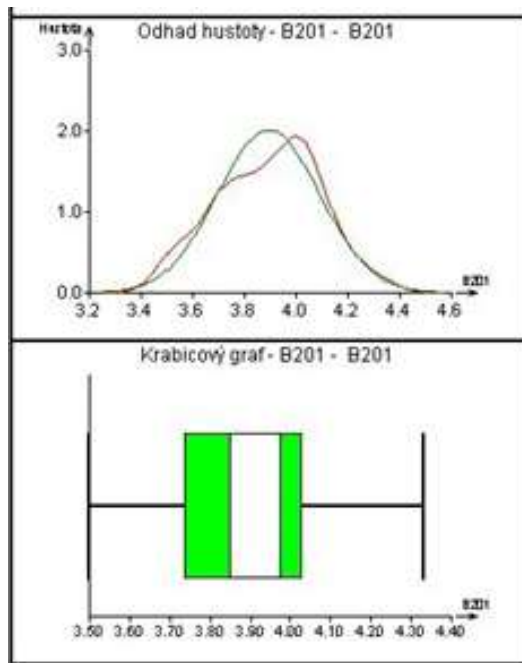
4 body dole a 2 body nahoře jsou podezřelé, že jsou odlehlé. Graf pro diagnostiku normality a odlehlých měření; pro normální data bez odlehlých měření má tvar přímky; pro normální data s odlehlými měřeními má tvar přímky s koncovými body ležícími mimo tuto přímku; pro systematicky sešikmená data s kladnou šikmostí (např. rozdělení lognormální, exponenciální) má nelineární konvexní tvar. Pro systematicky sešikmená data se zápornou šikmostí má nelineární konkávní tvar.

Diagramy rozptýlení: horní diagram rozptýlení a dolní rozmítnutý diagram rozptýlení ukazují na podezřelé body z odlehlosti. Zobrazuje všechna data ve skutečném měřítku na ose X. Popis osy Y nemá význam. Aby nedošlo ke splývání shodných nebo blízkých dat, jsou ve spodní polovině grafu zobrazena též data, ale náhodně rozmítnutá ("rozházená") na ose Y.

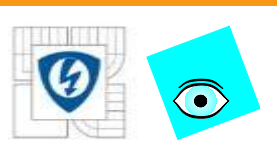


Grafy výstupu Exploratorní analýzy a jejich vysvětlení.

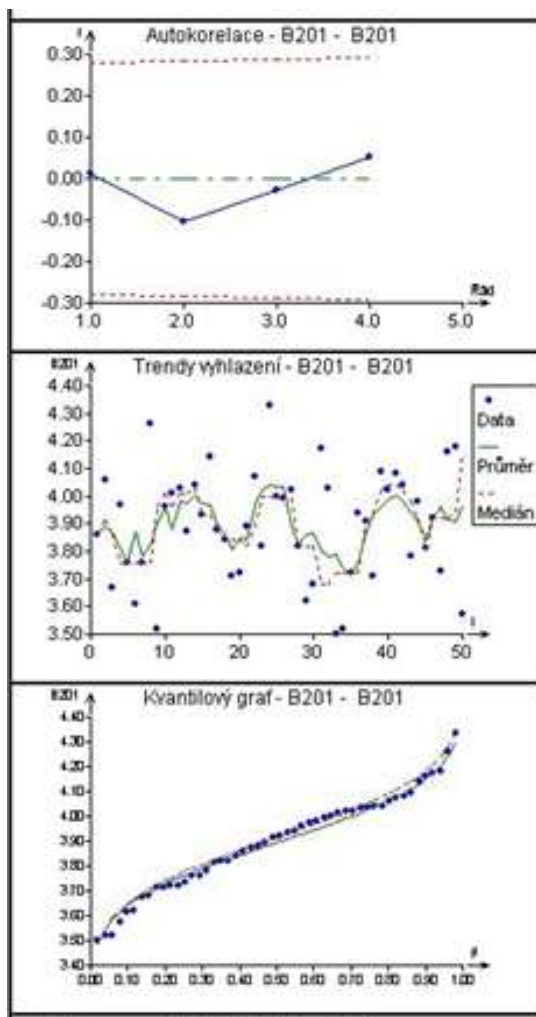
Jádrový odhad hustoty pravděpodobnosti: zelená křivka značí proložení teoretickým Gaussovým normálním rozdělením na bázi aritmetického průměru a směrodatné odchylky, zatímco červená křivka značí robustní proložení skutečným experimentálním rozdělením na bázi mediánu a kvartilového rozpětí. Porovnání průběhu hustoty pravděpodobnosti normálního rozdělení (plná čára) s jádrovým odhadem hustoty vypočítaným na základě dat (přerušovaná čára). Jádrový odhad používá Gaussovské jádro. Hladkost křivky jádrového odhadu je dána parametrem "Vyhlazení hustoty" v dialogovém panelu, čím je parametr menší, tím podrobnější je průběh. V případě normality a většího množství dat jsou si obě křivky blízké.



Krabicový graf: ukazuje na symetrické rozdělení bez vybočujících bodů (outlierů). Větší obdélník ohraničuje vnitřních 50% dat, horní okraj zeleného (vyšrafovaného) obdélníku odpovídá 75% kvantilu, spodní okraj zeleného obdélníku odpovídá 25% kvantilu, střed bílého pruhu v zeleném obdélníku odpovídá mediánu, šířka pruhu odpovídá intervalu spolehlivosti mediánu, dva černé proužky jsou tzv. vnitřní hradby. Data mimo vnitřní hradby jsou označena červeným proužkem a lze je považovat za vybočující měření.



Grafy výstupu Exploratorní analýzy a jejich vysvětlení.



Graf autokorelace: data jsou nezávislá, protože leží mezi červenými intervaly.

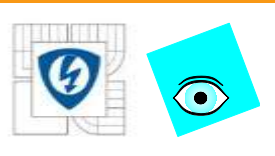
Graf autokorelačních koeficientů až do řádu autokorelace zadaného v dialogovém panelu. Červené meze ohraničují interval, v němž jsou koeficienty statisticky nevýznamné na zadané hladině významnosti. Překročí-li některý autokorelační koeficient tyto meze, je třeba považovat za závislá.

Graf trendů vyhlazení: v datech není trend.

Grafické znázornění trendu v datech pomocí klouzavého průměru (plná křivka) a klouzavého mediánu (přerušovaná křivka) na základě "Řádu trendu" zadaného v dialogovém panelu. Čím je řád trendu větší, tím jsou křivky hladší, méně citlivé na lokální poruchy a zachycují spíše globální dlouhodobý trend. Menší řád zachycuje spíše lokální chování dat. Klouzavý medián je méně citlivý (robustní) na lokální poruchy v datech a jednotlivá odlehlá měření a tedy vhodnější pro tyto případy.

Kvantilový graf: ukazuje na Gaussovo rozdělení.

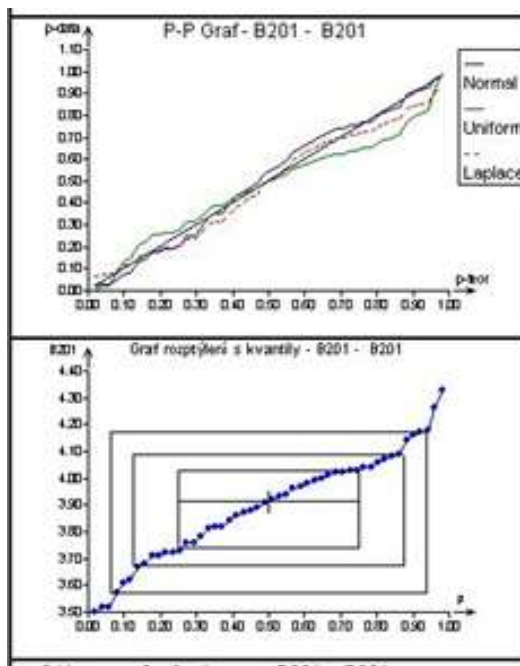
Zobrazuje empirické kvantily dat proložené kvantilovou funkcí normálního rozdělení. Zelená křivka odpovídá funkci s klasickým průměrem a rozptylem (nerobustní), červená křivka odpovídá mediánu a mediánové odchylce (robustní). Podle toho, která z křivek lépe prokládá data, je vhodné zvolit jako odhad střední hodnoty průměr nebo medián.



Grafy výstupu Exploratorní analýzy a jejich vysvětlení.

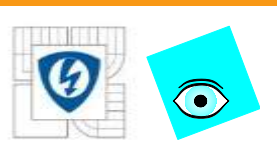
P-P graf: podobnost exp. křivky s normální ukazuje na Gaussovo rozdělení.

Porovnává data s normálním (modrá křivka plná), Laplaceovým (zelená křivka) a rovnoměrným rozdělením pomocí teoretické a empirické distribuční funkce. Která křivka leží nejbližce černé přímce $y = x$, to rozdělení odpovídá experimentálním datům. Graf slouží pro rozlišení symetrických rozdělení podle špičatosti. Podobnost rovnoměrnému rozdělení ukazuje na možné vyloučení vysokých a nízkých hodnot, podobnost s Laplaceovým rozdělením ukazuje na možnou nekonstantnost rozptylu dat.



Graf rozptýlení s kvantily: symetrie obdélníků ukazuje na Gaussovo rozdělení.

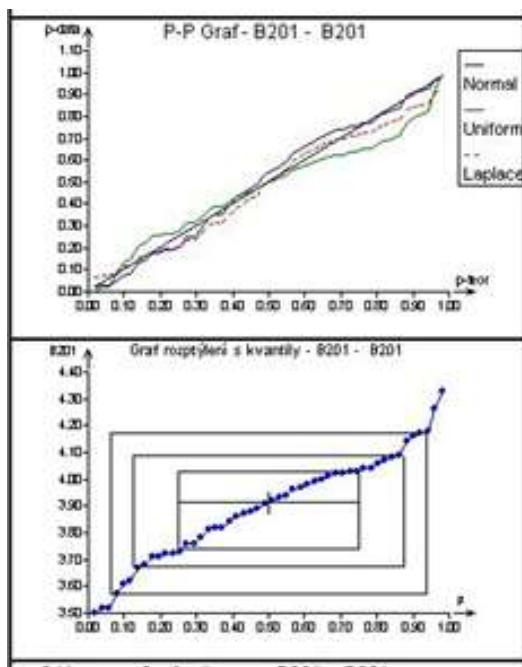
Body grafu jsou vizuálně i významově shodné s kvantilovým grafem. Vzájemná poloha obdélníků odpovídá symetrii, resp. asymetrii rozdělení. Vodorovná příčka uprostřed nejmenšího obdélníku označuje medián, svislá úsečka na příčce odpovídá intervalu spolehlivosti mediánu.



Grafy výstupu Exploratorní analýzy a jejich vysvětlení.

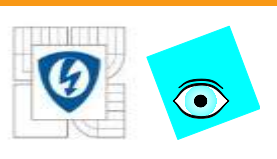
P-P graf: podobnost exp. křivky s normální ukazuje na Gaussovo rozdělení.

Porovnává data s normálním (modrá křivka plná), Laplaceovým (zelená křivka) a rovnoměrným rozdělením pomocí teoretické a empirické distribuční funkce. Která křivka leží nejbližce černé přímce $y = x$, to rozdělení odpovídá experimentálním datům. Graf slouží pro rozlišení symetrických rozdělení podle špičatosti. Podobnost rovnoměrnému rozdělení ukazuje na možné vyloučení vysokých a nízkých hodnot, podobnost s Laplaceovým rozdělením ukazuje na možnou nekonstantnost rozptylu dat.

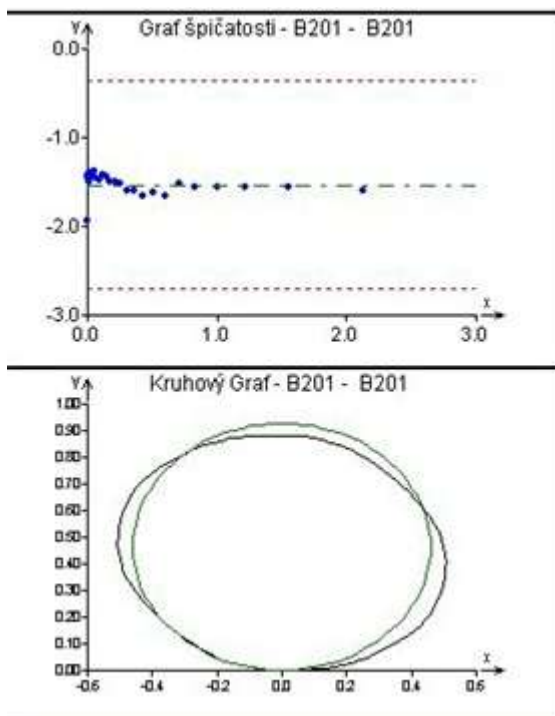


Graf rozptýlení s kvantily: symetrie obdélníků ukazuje na Gaussovo rozdělení.

Body grafu jsou vizuálně i významově shodné s kvantilovým grafem. Vzájemná poloha obdélníků odpovídá symetrii, resp. asymetrii rozdělení. Vodorovná příčka uprostřed nejmenšího obdélníku označuje medián, svislá úsečka na příčce odpovídá intervalu spolehlivosti mediánu.



Grafy výstupu Exploratorní analýzy a jejich vysvětlení.

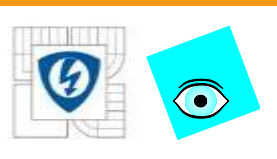


Graf špičatosti: většina bodů je v mezích a proto jde o symetrické rozdělení.

Má podobný význam jako dva předchozí grafy. Směrnice případného trendu je úměrná odchylky špičatosti od 3. V případě výrazně nenormální špičatosti rozdělení vykazují body výrazný trend. Body jsou konstruovány ze dvojic dat (první, poslední; druhý, předposlední; atd.), proto označením bodu jsou označena dvě příslušná data v tabulce.

Kruhový graf: exp. elipsa se značně blíží teoretické pro normální rozdělení.

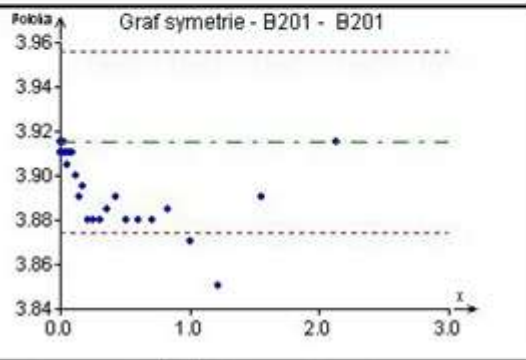
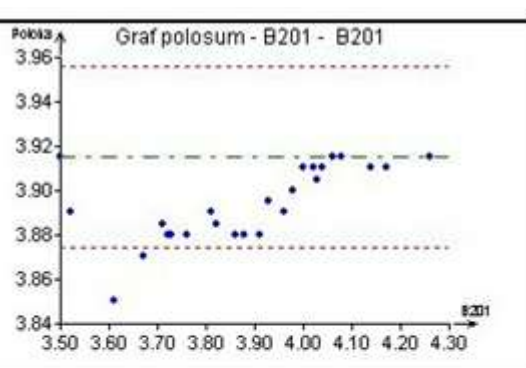
Slouží ke komplexnímu vizuálnímu posouzení normality na základě kombinace šikmosti a špičatosti. Zelený kruh (elipsa) je optimální tvar pro normální rozdělení, černý "kruh" představuje data. V případě normálních dat se obě křivky téměř kryjí.



Grafy výstupu Exploratorní analýzy a jejich vysvětlení.

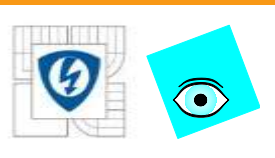
Graf polosum: většina bodů je v mezích a proto jde o symetrické rozdělení.

Citlivý indikátor asymetrie rozdělení. V ideálním případě leží body na horizontální přímce. Horizontální přímka, na níž leží poslední bod, představuje medián a červené přerušované meze jeho interval spolehlivosti. V případě asymetrického rozdělení vykazují body výrazný trend (rostoucí pro zápornou šikmost, nebo klesající, pro kladnou šikmost) výrazně překračující přerušované meze. Body jsou konstruovány ze dvojic dat (první, poslední; druhý, předposlední; atd.), proto označením bodu jsou označena dvě příslušná data v tabulce.



Graf symetrie: většina bodů je v mezích a proto jde o symetrické rozdělení.

Má podobný význam jako předchozí graf polosum. Směrnice případného trendu je úměrná šikmosti. V případě asymetrického rozdělení vykazují body výrazný trend (rostoucí pro zápornou šikmost, nebo klesající, pro kladnou šikmost) výrazně překračující přerušované meze. Body jsou konstruovány ze dvojic dat (první, poslední; druhý, předposlední; atd.), proto označením bodu jsou označena dvě příslušná data v tabulce.



Popis výsledků

Základní analýza dat

Název úlohy :	B201
Data:	Všechna
Řád trendu :	4
Testovaná hodnota :	0
Vyhlazení hustoty :	0.5
Hladina významnosti :	0.05
Název sloupce :	B201
Počet platných dat :	50

Klasické parametry :

Název sloupce :	B201
Průměr :	3.894
Spodní mez :	3.837588535
Horní mez :	3.950411465
Rozptyl :	0.0394
Směr. odchylka :	0.1984943324
Šikmost :	-0.119116936
Odchylka od 0 :	Nevýznamná
Špičatost :	2.40159769
Odchylka od 3 :	Nevýznamná
Polosuma :	3.915
Modus :	3.955352941

Sloupec

Počet řádků

v datech.

Počet platných dat

v datech.

Počet chybějících hodnot

v datech.

Název sloupce.

Celkový počet řádků

Celkový počet platných dat

Počet prázdných buněk

Klasické parametry

Aritmetický průměr

Odhad střední hodnoty pro normálně rozdělená data.

Spodní mez

Spodní mez intervalu spolehlivosti aritmetického průměru na zadané hladině významnosti.

Horní mez

Horní mez intervalu spolehlivosti aritmetického průměru na zadané hladině významnosti.

Rozptyl

Odhad rozptylu.

Směrodatná odchylka

Druhá odmocnina z rozptylu.

Šikmost

Odhad třetího statistického momentu, šikmosti.

Rozdíl od 0

Normální a každé symetrické rozdělení má šikmost nulovou. Je-li hodnota šikmosti statisticky významně odlišná od 0, nelze data považovat za symetrická. Spolehlivější je však test normality.

Špičatost

Odhad čtvrtého statistického momentu, špičatosti.

Rozdíl od 3

Normální rozdělení má špičatost 3. Je-li hodnota špičatosti statisticky významně odlišná od 3, lze předpokládat, že data neodpovídají normálnímu rozdělení. Spolehlivější je však test normality.

Polosuma

Odhad polosumy, tedy středu nejmenší a největší hodnoty.

Modus

Odhad modu rozdělení, tedy maxima na křivce hustoty pravděpodobnosti.

Robustní parametry :

Název sloupce :	B201
Medián :	3.915
IS spodní :	3.709937536
IS horní :	4.120062464
Medianová směr. odchylka :	0.1020426907
Medianový rozptyl :	0.01041271072
10% Průměr :	3.895227273
10% IS spodní :	3.835978432
10% IS horní :	3.954476113
10% Směr. odchylka :	0.149154143
10% Rozptyl :	0.02224695837
20% Průměr :	3.897
20% IS spodní :	3.835649773
20% IS horní :	3.958350227
20% Směr. odchylka :	0.1243187477
20% Rozptyl :	0.01545515102
40% Průměr :	3.901666667
40% IS spodní :	3.838030654
40% IS horní :	3.96530268
40% Směr. odchylka :	0.08167506551
40% Rozptyl :	0.006670816327

Robustní parametry:

Medián

Odhad mediánu, tedy 50% kvantilu. Tento odhad střední hodnoty je spolehlivější než aritmetický průměr v případě porušení normality dat nebo přítomnosti vybočujících bodů.

IS spodní

Spodní mez intervalu spolehlivosti mediánu na zadané hladině významnosti.

IS horní

Horní mez intervalu spolehlivosti mediánu na zadané hladině významnosti.

Mediánová směr.odchylka

Odhad směrodatné odchylky na základě mediánu.

Mediánový rozptyl

Odhad rozptylu na základě mediánu.

10% uřezaný průměr

Aritmetický průměr pro symetrickém uřezání 10% dat, tedy 5% nejmenších a 5% největších hodnot. Tento robustní odhad střední hodnoty se doporučuje v případě podezření na vybočující body.

IS spodní

Spodní mez intervalu spolehlivosti uřezaného průměru na zadané hladině významnosti.

IS horní

Horní mez intervalu spolehlivosti uřezaného průměru na zadané hladině významnosti.

Rozptyl

Odhad rozptylu na základě mediánu.

Směr. odchylka

Odhad směrodatné odchylky na základě mediánu.

40% uřezaný průměr

Aritmetický průměr pro symetrickém uřezání 40% dat, tedy 20% nejmenších a 20% největších hodnot. Tento robustní odhad střední hodnoty se doporučuje v případě podezření na velký počet vybočujících bodů.

IS spodní

Spodní mez intervalu spolehlivosti uřezaného průměru na zadané hladině významnosti.

IS horní

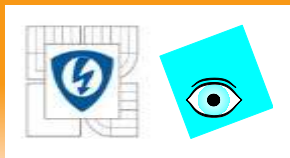
Horní mez intervalu spolehlivosti uřezaného průměru na zadané hladině významnosti.

Rozptyl

Odhad rozptylu na základě mediánu.

Směr. odchylka

Odhad směrodatné odchylky na základě mediánu.



Popis výsledků

Znaménkový test :

Závěr :

Data jsou nezávislá

Znaménkový test

Závěr

Neparametrický test nezávislosti dat, závislost je indikována, obsahují-li data shluky po sobě jdoucích sekvencí se shodným znaménkem odchylky od průměru. Slovní závěr testu závislosti dat.

Test normality

Kombinovaný test normality založený na shodě šikmosti a špičatosti s normálním rozdělením.

Normalita

Slovní závěr testu na zadané hladině významnosti.

Vypočtený

Vypočtená testovací statistika.

Teoretický

Příslušný kvantil t-rozdělení.

Pravděpodobnost

Pravděpodobnost odpovídající vypočtené statistice.

Vybočující body

Robustní test na přítomnost vybočujících měření založený na kvantilovém odhadu vnitřních mezí dat.

Homogenita

Slovní závěr testu, nejsou-li v datech vybočující měření, je předpoklad homogenity přijat.

Počet vybočujících bodů Počet případných měření přesahujících přípustné meze, které je možno považovat za vybočující.

Dolní hranice

Dolní hranice, pod níž je možno data považovat za vybočující.

Horní hranice

Horní hranice, nad níž je možno data považovat za vybočující.

Test normality :	
Název sloupce :	B201
Průměr :	3.894
Rozptyl :	0.0394
Šikmost :	-0.119116936
Špičatost :	2.40159769
Normalita :	Přijata
Vypočtený :	0.2378781079
Teoretický :	5.991464547
Pravděpodobnost :	0.8878619108
Vybočující body :	
Název sloupce :	B201
Homogenita :	Přijata
Počet vybočujících bodů :	0
Spodní mez :	3.04482
Horní mez :	4.70518

Autokorelace :
 Řád autokorelace : 4
 Název sloupce : B201
 Počet : 0.0521889568

Řád autokorelace 1
 Korelační koeficient : 0.01257847301
 Pravděpodobnost : 0.4658208006
 Závěr : Nevýznamný

Řád autokorelace 2
 Korelační koeficient : -0.1020884699
 Pravděpodobnost : 0.2449546948
 Závěr : Nevýznamný

Řád autokorelace 3
 Korelační koeficient : -0.02901067026
 Pravděpodobnost : 0.4232554223
 Závěr : Nevýznamný

Řád autokorelace 4
 Korelační koeficient : 0.0521889568
 Pravděpodobnost : 0.3652515549
 Závěr : Nevýznamný

Test významnosti trendu :
 Název sloupce : B201
 Směrnice : 0.0004225690276
 Významnost : Nevýznamný
 Pravděpodobnost : 0.5847032011

Autokorelace

Odhady autokorelačních koeficientů a jejich významnost na zadané hladině významnosti.

Řád autokorelace Koeficient

Řád autokorelace.

Hodnota autokorelačního koeficientu, formálně odpovídá párovému korelačnímu koeficientu a má stejné vlastnosti.

Pravděpodobnost

Pravděpodobnost nevýznamnosti autokorelačního koeficientu; je-li menší než zvolená hladina významnosti, je autokorelace významná.

R0Krit

Kritická hodnota autokorelačního koeficientu, nad níž se korelace považuje za významnou.

Vysledek

Slovní vyjádření významnosti autokorelace.

Vyhlazené hodnoty

Hodnoty klouzavých průměrů a mediánů. Při vhodné volbě řádu trendu lze získat po odečtení těchto hodnot od vstupních dat detrendovaná data, která lze použít pro konstrukci regulačních diagramů v případě, že je v datech příliš silný trend.

Rezidua

Odchylky (rozdíly) naměřených a vyhlazených hodnot, (nameřená - vyhlazená).

Test významnosti trendu

Test významnosti lineárního trendu v datech.

Vypočtený

Testovací statistika pro směrnici přímky vypočtená z dat.

Teoretický

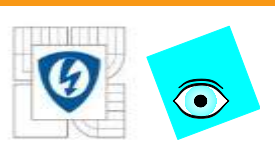
Příslušný kvantil t-rozdělení.

Trend

Hodnota směrnice přímky proložené daty

Pravděpodobnost

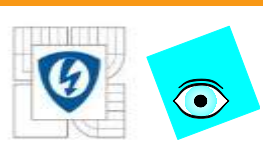
Pravděpodobnost toho, že je lineární trend nevýznamný; vyjde-li menší než zadaná hladina významnosti (tedy obvykle 0.05), považuje se trend za významný.



Ověření předpokladu výběru dat

Diagnostiky, testy: ověření normality, nezávislosti, homogenity, určení minima četností





Data a podmínky výpočtu.

V Ý S L E D K Y

ZÁKLADNÍ STATISTIKA

Základní předpoklady

Název : B201

V S T U P

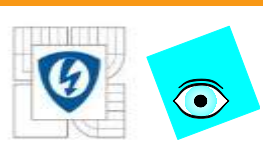
(1) DATA A PODMÍNKY:

Počet dat : 50
Hladina významnosti alfa : 0.050
Název výstupního souboru : RESULTS.TXT

(2) VSTUPNÍ DATA:

(1)	3.8600E+00	(14)	4.0400E+00	(27)	4.0200E+00	(40)	4.0200E+00
(2)	4.0600E+00	(15)	3.9300E+00	(28)	3.8200E+00	(41)	4.0800E+00
(3)	3.6700E+00	(16)	4.1400E+00	(29)	3.6200E+00	(42)	4.0400E+00

Napověda-F1 Řádek: 1 - 23 Celkem: 81 Délka: 2739



Výsledky

V Ý S L E D K Y

(1) KLASICKÉ ODHADY PARAMETRŮ:

Průměr	: 3.8940E+00	Rozptyl	: 3.9398E-02
Směrodatná odchylka	: 1.9849E-01	Šikmost	: -1.1890E-01
Špičatost	: 2.4015E+00		

(2) TEST NORMALITY:

Data jsou normálního rozdělení.

Tabulkový kvantil $\chi^2(1-\alpha, 2)$	: 5.9915E+00
χ^2 -statistika	: 7.8013E-01
Závěr: Předpoklad normality přijat	
Vypočtená hladina významnosti	: 6.7701E-01

(3) TEST NEZÁVISLOSTI:

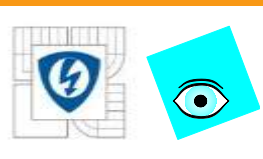
Data jsou nezávislá.

Tabulkový kvantil $t(1-\alpha/2, n+1)$	: 2.0076E+00
Test autokorelace	: 2.9905E-01
Závěr: Předpoklad nezávislosti přijat	
Vypočtená hladina významnosti	: 3.8306E-01

Předpoklad homogenity výběru:

Data jsou homogenní.

Napověda-F1 Řádek: 37 - 59 Celkem: 81 Délka: 2739



V datech nejsou outliers.

V Ý S L E D K Y

Aritmetický průměr : 3.8940E+00
Rozptyl : 3.9398E-02
Směrodatná odchylka: 1.9849E-01

Unitřní meze:
Spodní mez : 3.0764E+00
Horní mez : 4.6837E+00

(4) MINIMÁLNÍ VELIKOST VÝBĚRU:

pro 25% relativní chybu směrodatné odchylky : n = 7
pro 10% relativní chybu směrodatné odchylky : n = 36
pro 5% relativní chybu směrodatné odchylky : n = 141

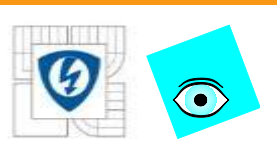
(5) DETEKCE ODLEHLÝCH BODŮ:

Ve výběru nejsou odlehlé body

Parametry s vymezenými odlehlými hodnotami:

Průměr	: 3.8940E+00	Rozptyl	: 3.9398E-02
Směrodatná odchylka	: 1.9849E-01	Šikmost	: -1.1915E-01
Špičatost	: 2.4016E+00		

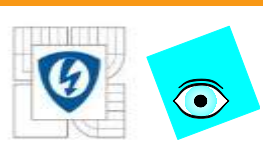
Napověda-F1 Řádek: 60 - 81 Celkem: 81 Délka: 2739



Závěr základních předpokladů

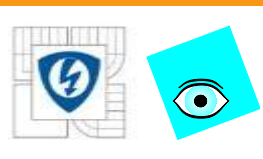
Data vykazují Gaussovo normální rozdělení a není proto třeba použít jakoukoliv transformaci dat.

Test prokázal, že data jsou nezávislá a bez vybočujících hodnot.



Určení typu rozdělení výběru.





Data a podmínky výpočtu.

Č E K Á M ...

Základní Statistika: Porovnání rozdělení

Data	Metoda	Zadání	Výpočet	Výsledky	Graf	Konec
------	--------	---------------	---------	----------	------	-------

Podmínky

Volby

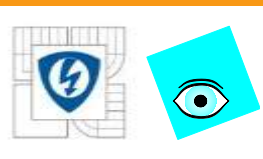
PODMÍNKY VÝPOČTU

Název: B201

Počáteční podmínky:

Parametr pro Weibullovo rozd.: 1.0000E+00
Parametr for Paretovo rozdělení: 1.0000E+00
Parametr pro Gamma rozdělení: 1.0000E+00

Zadej podmínky pro výpočet



Data a podmínky výpočtu.

Základní Statistika: Porovnání rozdělení

Data	Metoda	Zadání	Výpočet	Výsledky	Graf	Konec
------	--------	---------------	---------	----------	------	-------

Podmínky

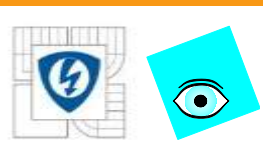
Volby

V O L B Y

Kreslit Q-Q graf : Kreslit TDF graf : Ano
Hladina významnosti: 0.050

Název výst. souboru: RESULTS .TXT

Zadej počáteční volby



Data a podmínky výpočtu.

V Ý S L E D K Y

ZÁKLADNÍ STATISTIKA

Porovnání rozdělení

Název : B201

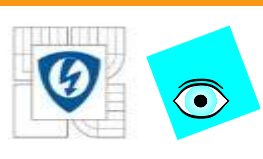
V S T U P

(1) DATA A PODMÍNKY:
Počet dat : 50
Hladina významnosti alfa : 0.050
Název výstupního souboru : RESULTS.TXT

(2) VSTUPNÍ DATA:

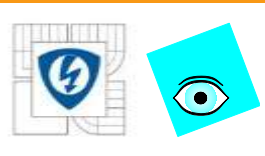
(1)	3.8600E+00	(14)	4.0400E+00	(27)	4.0200E+00	(40)	4.0200E+00
(2)	4.0600E+00	(15)	3.9300E+00	(28)	3.8200E+00	(41)	4.0800E+00
(3)	3.6700E+00	(16)	4.1400E+00	(29)	3.6200E+00	(42)	4.0400E+00
(4)	3.9700E+00	(17)	3.8800E+00	(30)	3.6800E+00	(43)	3.7800E+00

Napověda-F1 Řádek: 1 - 23 Celkem: 158 Délka: 8464

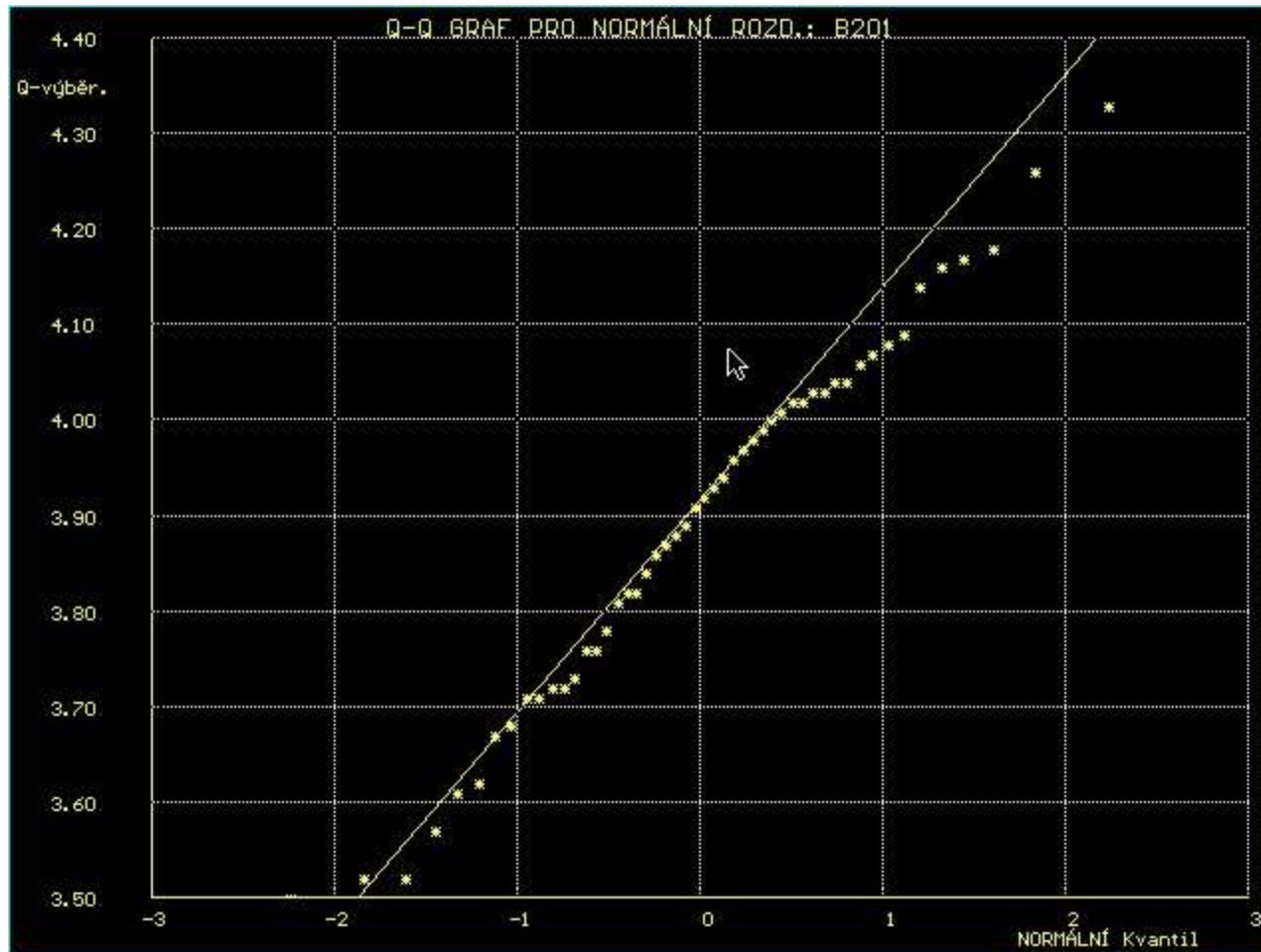


Sleduje se nejvyšší hodnota korelačního koeficientu blízka 1, a to pro prvních 5 rozdělení.

V Ý S L E D K Y				
(1) LINEARITA V GRAFU KVANTIL-KVANTIL (Q-Q):				
Čís.	Rozdělení	Směrnice	Úsek	Korelační koeficient
0	Laplaceovo	1.4496E-01	3.8940E+00	9.7403E-01
1	Normální	2.0124E-01	3.8940E+00	9.9432E-01
2	Exponenciální	1.8826E-01	3.7087E+00	8.9901E-01
3	Rovnoměrné	6.7467E-01	3.5567E+00	9.8601E-01
4	Lognormální	9.4522E-02	3.7446E+00	8.2336E-01
5	Gumbelovo	1.5654E-01	3.9824E+00	9.7451E-01
6	Weibullovo(0.5)	4.0831E-02	3.8185E+00	7.1528E-01
6	Weibullovo(1.0)	1.8826E-01	3.7087E+00	8.9901E-01
6	Weibullovo(1.5)	3.1787E-01	3.6089E+00	9.5789E-01
6	Weibullovo(2.0)	4.2672E-01	3.5171E+00	9.8026E-01
6	Weibullovo(2.5)	5.2410E-01	3.4298E+00	9.8961E-01
6	Weibullovo(3.0)	6.1507E-01	3.3452E+00	9.9358E-01
6	Weibullovo(3.5)	7.0227E-01	3.2623E+00	9.9511E-01
6	Weibullovo(4.0)	7.8709E-01	3.1806E+00	9.9546E-01
6	Weibullovo(4.5)	8.7035E-01	3.0996E+00	9.9523E-01
6	Weibullovo(5.0)	9.5252E-01	3.0191E+00	9.9470E-01
6	Weibullovo(5.5)	1.0339E+00	2.9390E+00	9.9404E-01
6	Weibullovo(6.0)	1.1147E+00	2.8593E+00	9.9332E-01
6	Weibullovo(6.5)	1.1951E+00	2.7798E+00	9.9259E-01
Napověda-F1		Řádek: 36 - 58	Celkem: 158 Délka: 8464	



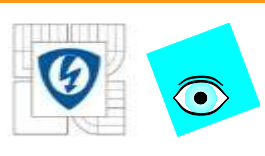
Nejlépe leží body na přímce pro normální rozdělení.



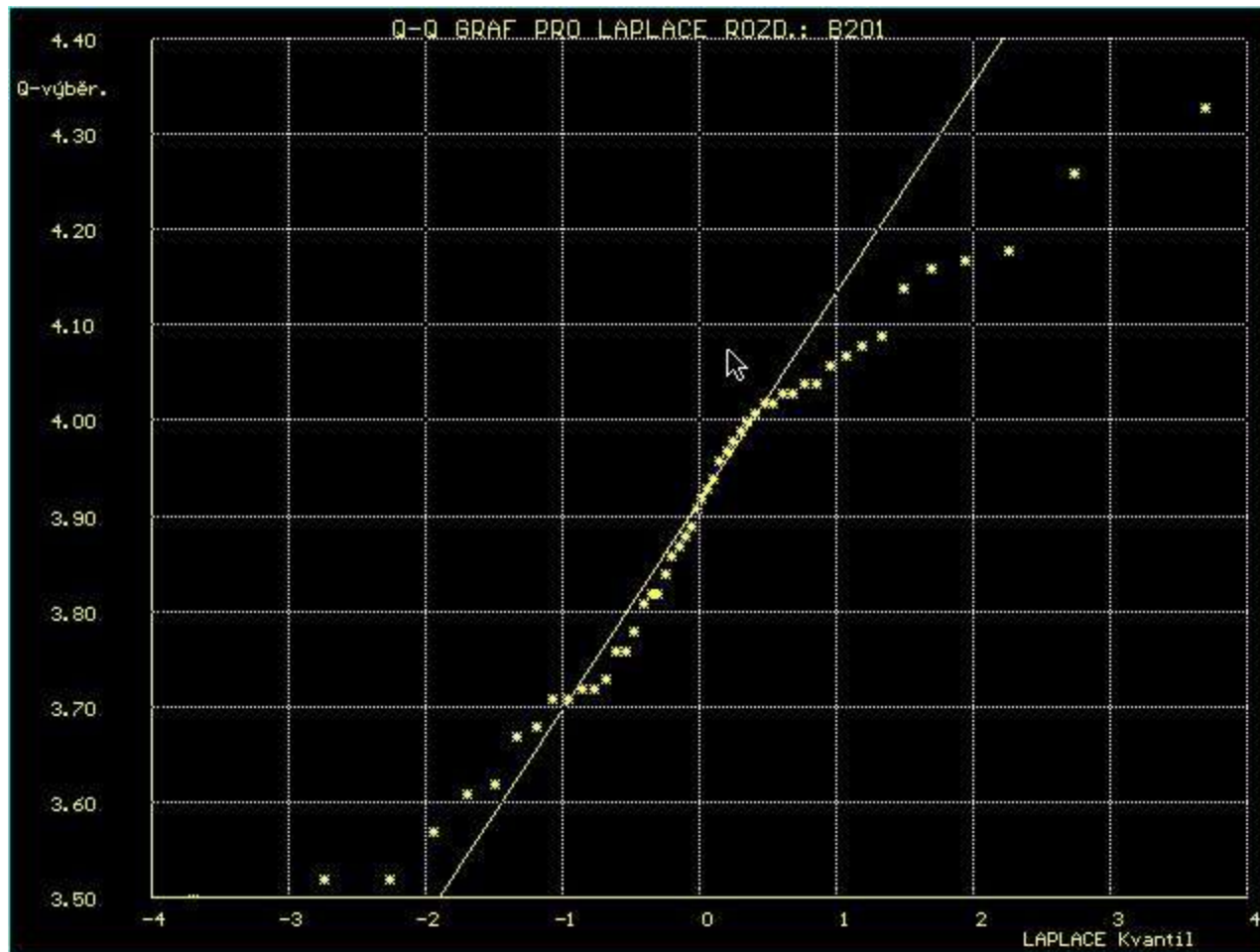
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





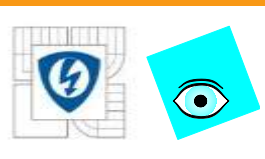
Nejlépe leží body na přímce pro normální rozdělení.



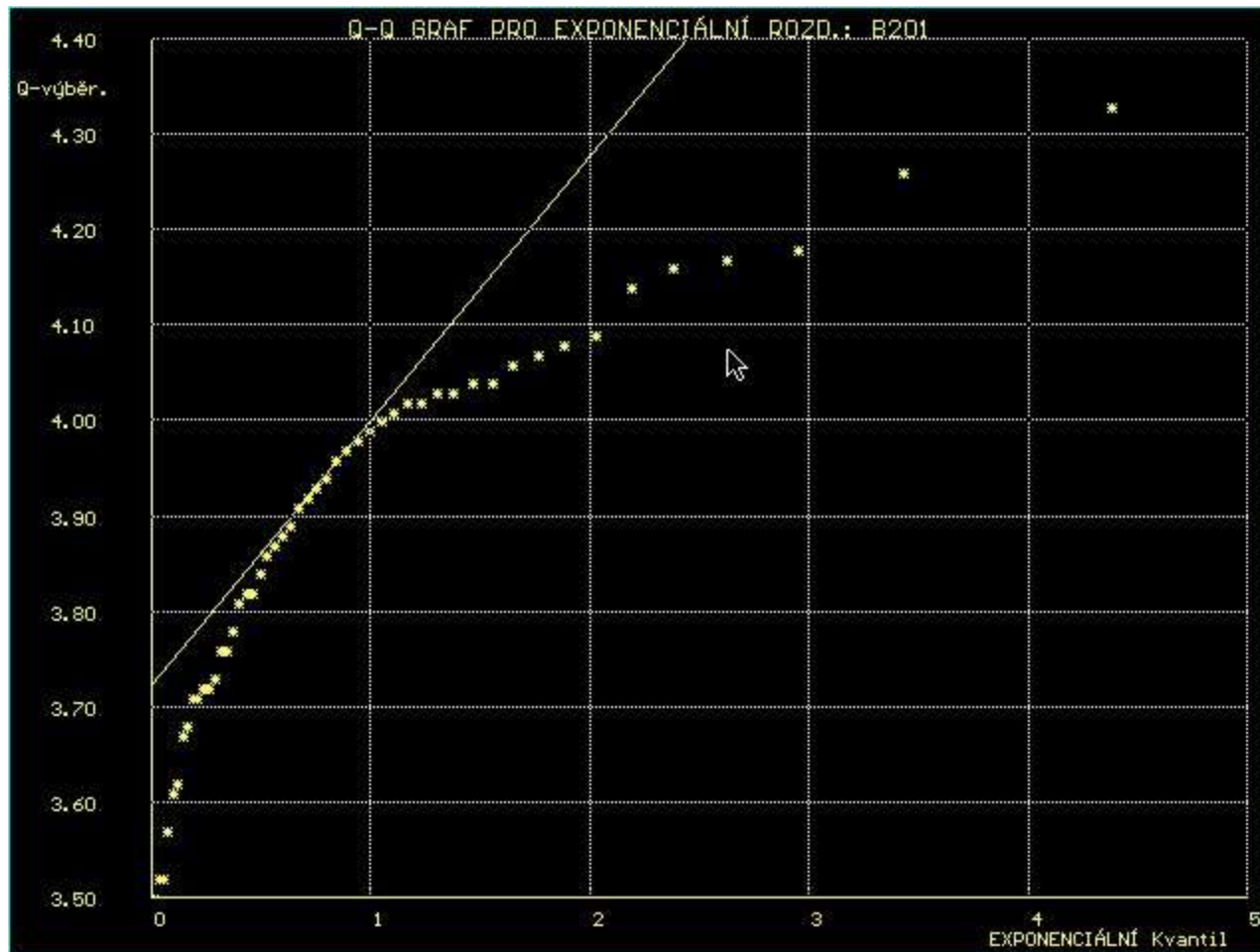
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





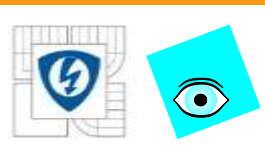
Nejlépe leží body na přímce pro normální rozdělení.



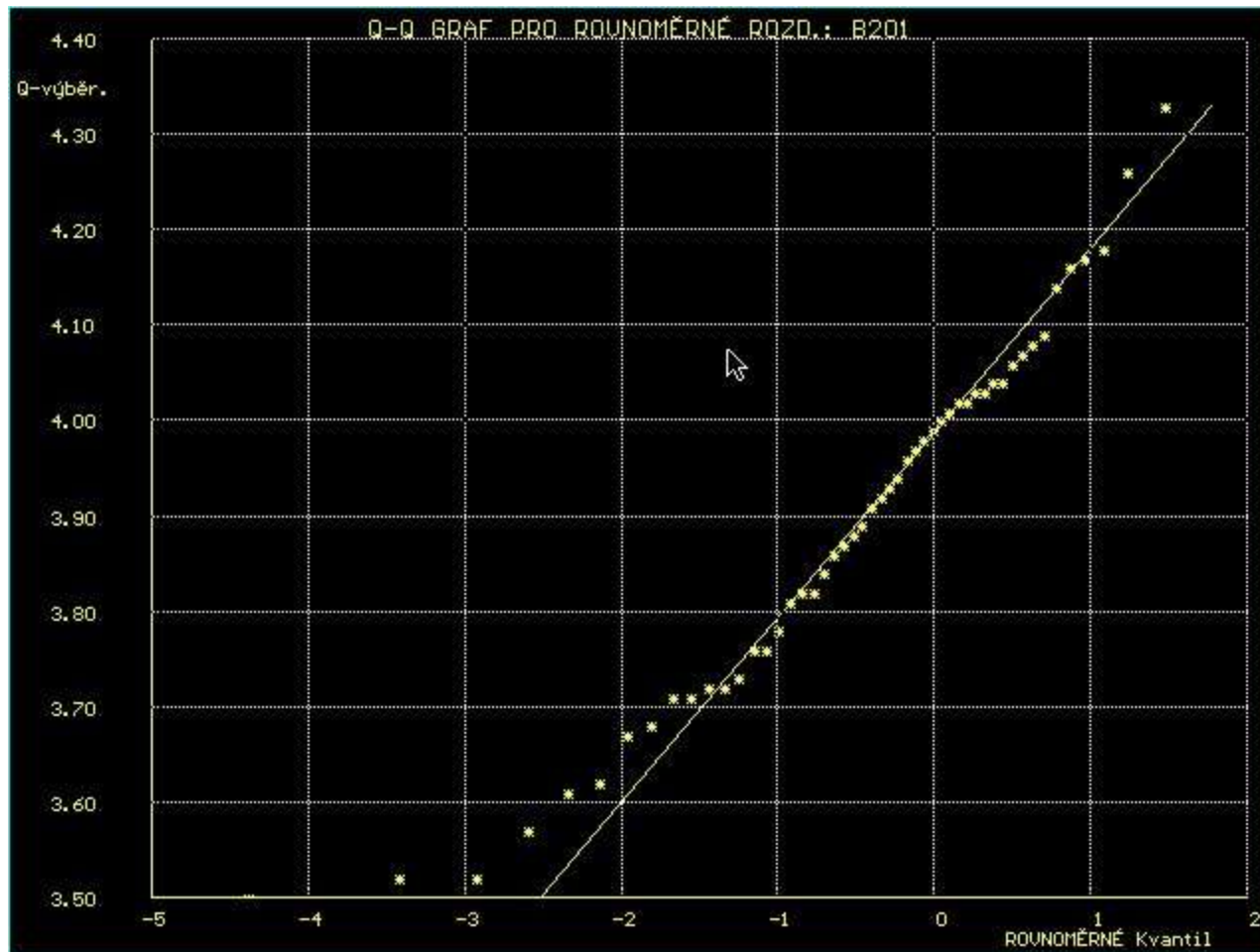
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





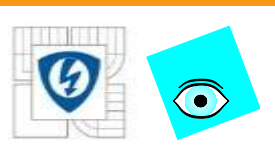
Nejlépe leží body na přímce pro normální rozdělení.



31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

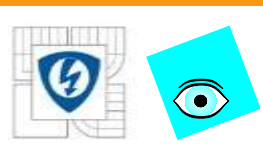




Transformace dat

Analýza dat: originální data, data po mocninné transformaci, data po Box-Coxově transformaci





Transformace dat

Č E K Á M ...

Základní Statistika: Mocnná transformace

Data	Metoda	Zadání	Výpočet	Výsledky	Graf	Konec
------	--------	--------	---------	----------	------	-------

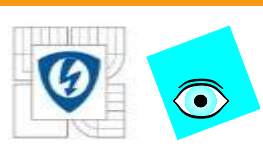
Podmínky
Volby

PODMÍNKY VÝPOČTU

Název: B201

Prostá mocnina : 5.0000E-01 Box – Coxova mocnina : 5.0000E-01

Zadej podmínky pro výpočet



Transformace dat

V Ý S L E D K Y

(2) PROSTÁ MOCNINNÁ TRANSFORMACE:

(A) Optimální hodnoty mocniny pro vybraná kritéria:

Optimální mocnina:	2.1333E+00	pro šikmost	: 1.1508E-03
Optimální mocnina:	-4.0000E+00	pro špičatost	: 3.3114E+00
Optimální mocnina:	4.0000E+00	pro asymetrii	: 7.4458E-03
Optimální mocnina:	4.0000E+00	pro asymetrii, rob.	: 1.7341E-01
Optimální mocnina:	-4.0000E+00	pro Hinkley-asymetrii	: 7.1361E-04

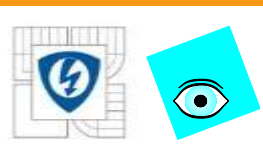
Zvolená mocnina : 2.13

Průměr	: 1.8232E+01
Rozptyl	: 3.8861E+00
Směrodatná odchylka	: 1.9713E+00
Šikmost	: 1.1508E-03
Špičatost	: 2.4333E+00
Opravený průměr	: 3.8996E+00

(B) Kvantilové míry:

Kvantil	P	Spodní mez	Horní mez	Polorozptyl
---------	---	------------	-----------	-------------

Napověda-F1 Řádek: 63 - 85 Celkem: 185 Délka: 7564



Transformace dat

```
U Ý S L E D K Y
(3) BOX-COXOVA TRANSFORMACE:

(A) Optimální hodnoty mocniny pro vybraná kritéria:

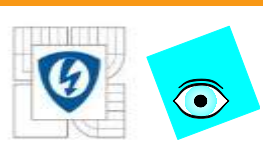
Optimální mocnina: 2.1333E+00 pro šikmost           : 1.1508E-03
Optimální mocnina: -4.0000E+00 pro špičatost        : 3.3114E+00
Optimální mocnina: 4.0000E+00 pro asymetrii         : 7.4458E-03
Optimální mocnina: 4.0000E+00 pro asymetrii, rob.   : 1.7341E-01
Optimální mocnina: -4.0000E+00 pro Hinkley-asymetrii: 1.7840E-04
Optimální mocnina: 2.0000E+00 pro věrohodnost      : 8.0914E+01

Zvolená mocnina : 2.13

Průměr           : 8.0777E+00
Rozptyl          : 8.5387E-01
Směrodatná odchylka : 9.2405E-01
Šikmost          : 1.1508E-03
Špičatost        : 2.4333E+00
Opravený průměr   : 3.8996E+00

(B) Kvantilové míry:

Napověda-F1      Řádek: 96 - 118      Celkem: 185      Délka: 7564
```



Transformace dat

```
U V S L E D K Y

ZÁKLADNÍ STATISTIKA
MOCNINNÁ TRANSFORMACE

Název : B201

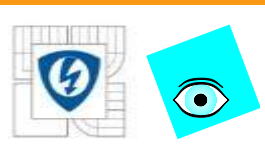
U S T U P

(1) DATA A PODMÍNKY:
    Počet dat                : 50
    Hladina významnosti alfa : 0.050
    Název výstupního souboru  : RESULTS.TXT

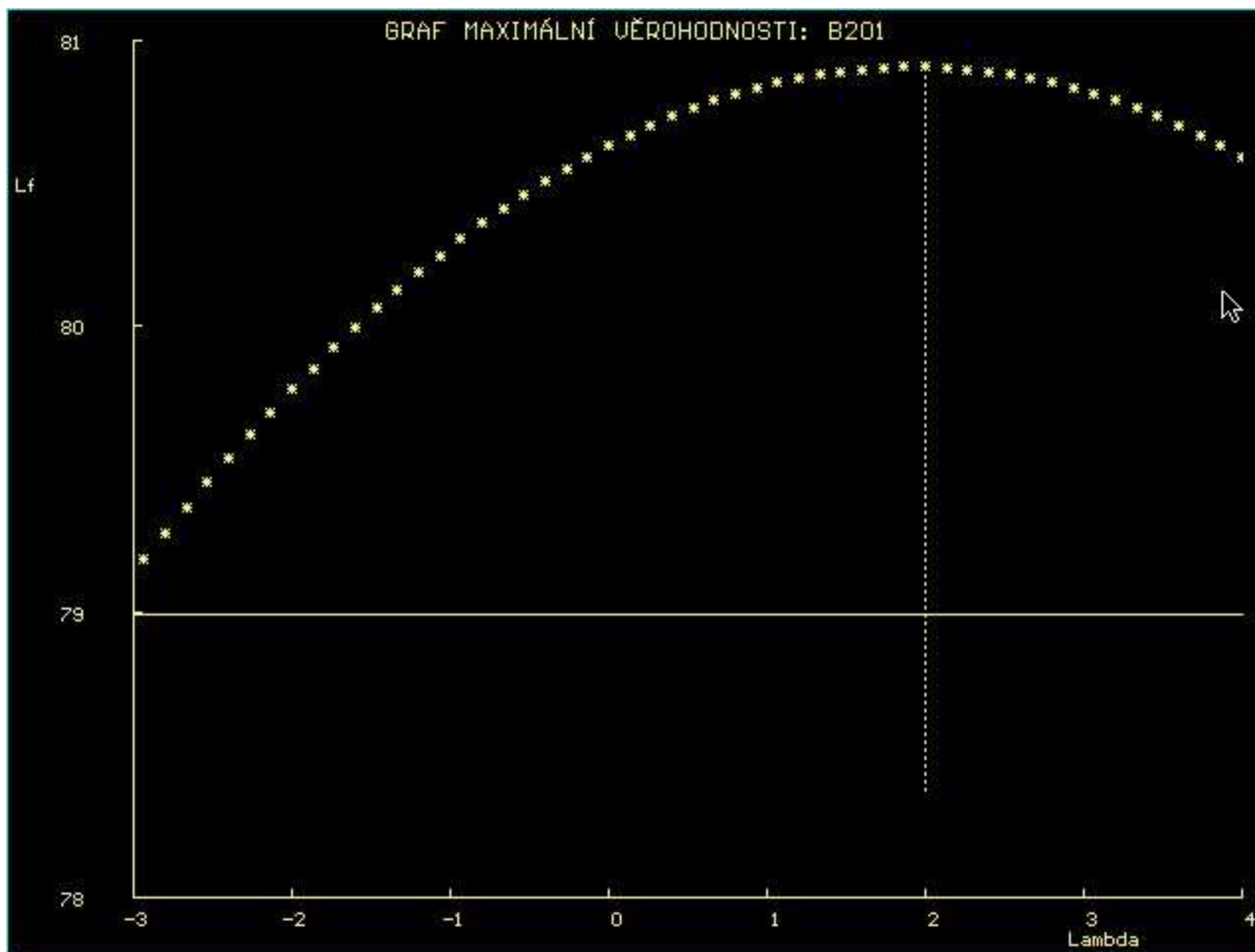
(2) VSTUPNÍ DATA:

( 1) 3.8600E+00 ( 14) 4.0400E+00 ( 27) 4.0200E+00 ( 40) 4.0200E+00
( 2) 4.0600E+00 ( 15) 3.9300E+00 ( 28) 3.8200E+00 ( 41) 4.0800E+00
( 3) 3.6700E+00 ( 16) 4.1400E+00 ( 29) 3.6200E+00 ( 42) 4.0400E+00
( 4) 3.9700E+00 ( 17) 3.8800E+00 ( 30) 3.6800E+00 ( 43) 3.7800E+00

Napověda-F1   Řádek: 1 - 23                Celkem: 185   Délka: 7564
```



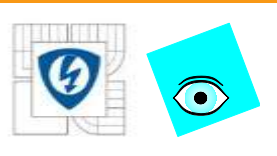
Transformace dat



31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

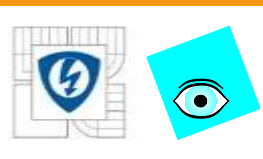




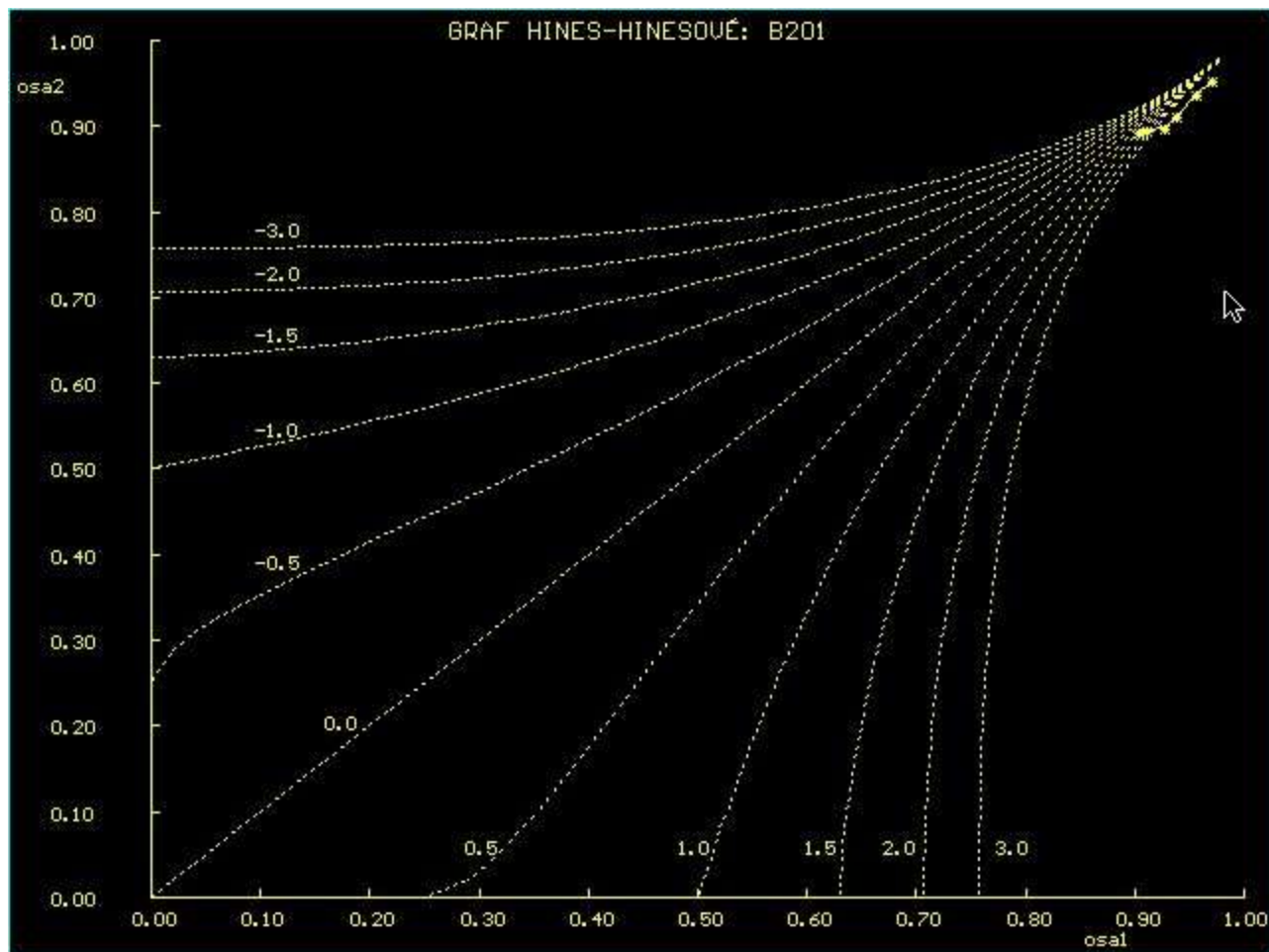
Rozhodčí kritérium použití transformace dat:

Nachází-li se pod segmentem na x-ové ose číslo 1 , pak je transformace zbytečná a lze využít pro odhad střední hodnoty aritmetický průměr.

Závěr: Jelikož zde je číslo 1 pod segmentem, je ze statistického hlediska retransformovaný průměr shodný s aritmetickým.



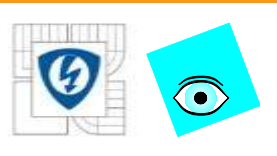
Transformace dat



31.1.2011

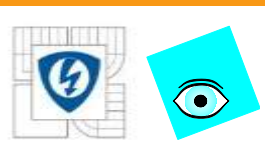
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



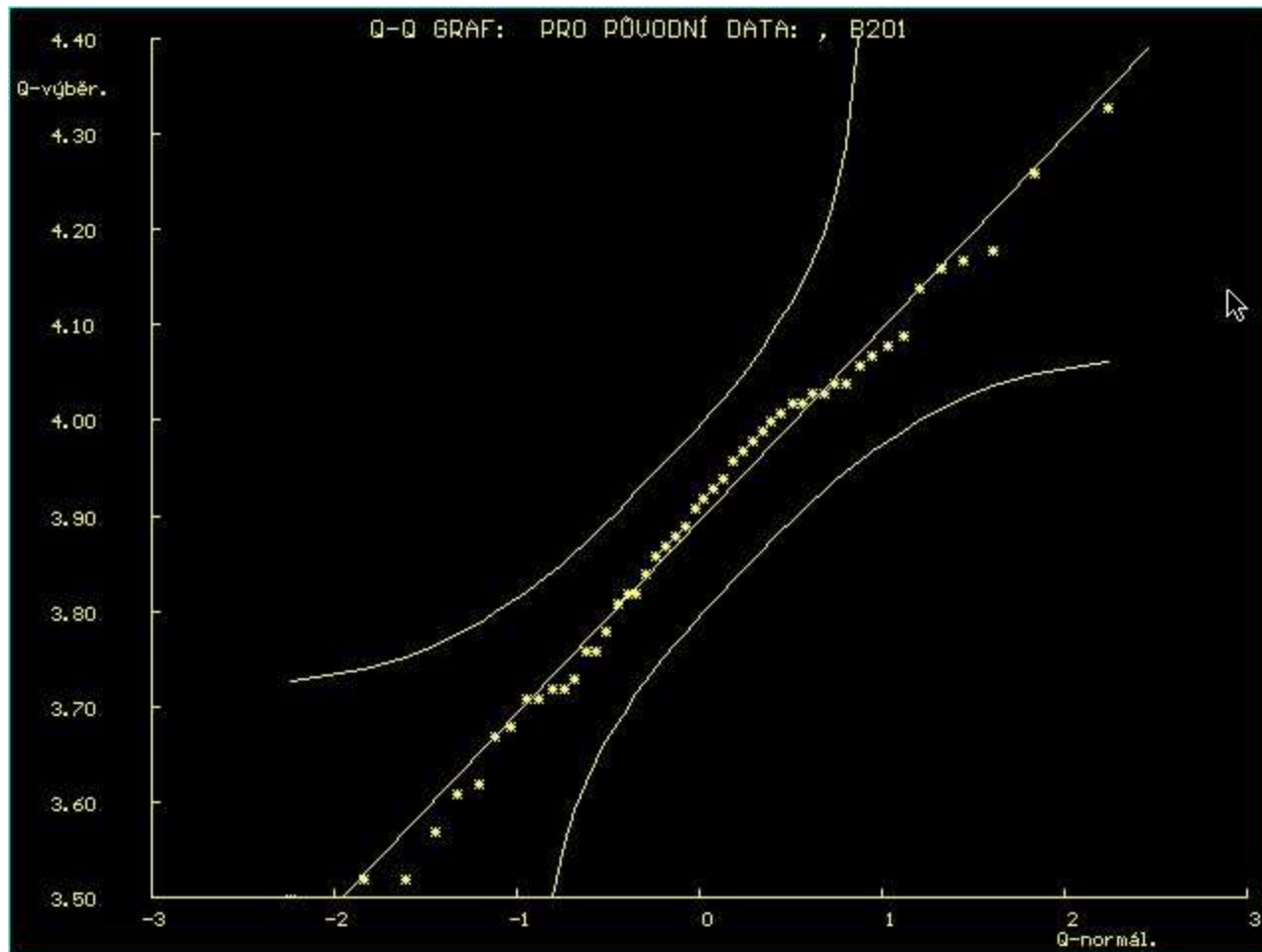


Určení nejlepšího odhadu exponentu λ Grafickou metodou grafem Hinese a Hinesové:

V nomogramu Hines-Hines se odhadne poloha křivky s body, která se přibližuje některému průvodiči v nomogramu.



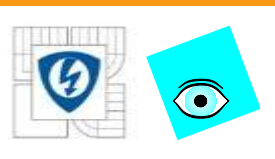
Ukázka postupného zlepšení symetrie rozdělení po transformaci dat



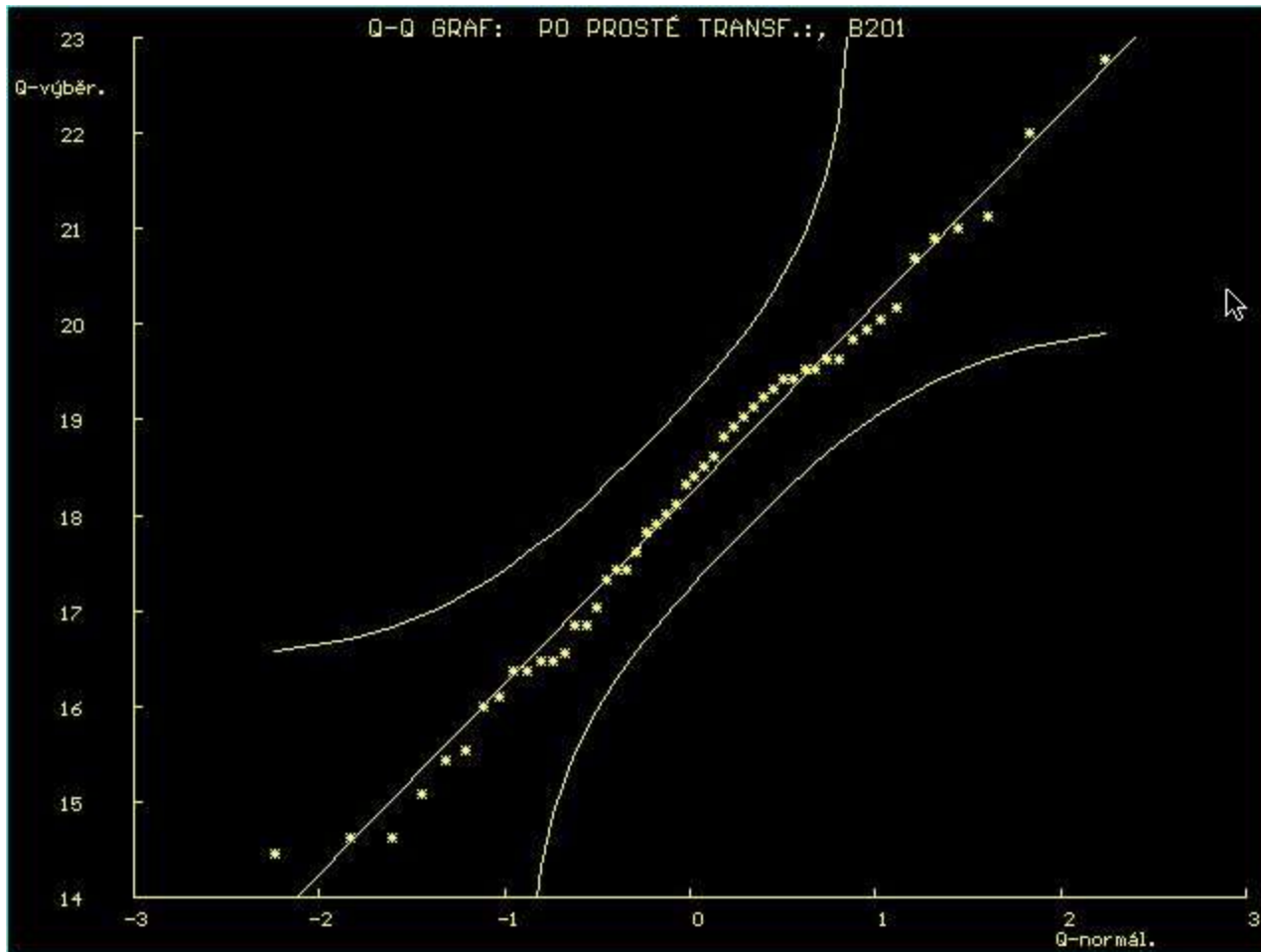
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





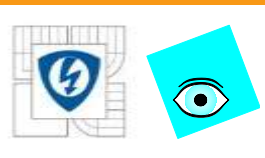
Ukázka postupného zlepšení symetrie rozdělení po transformaci dat



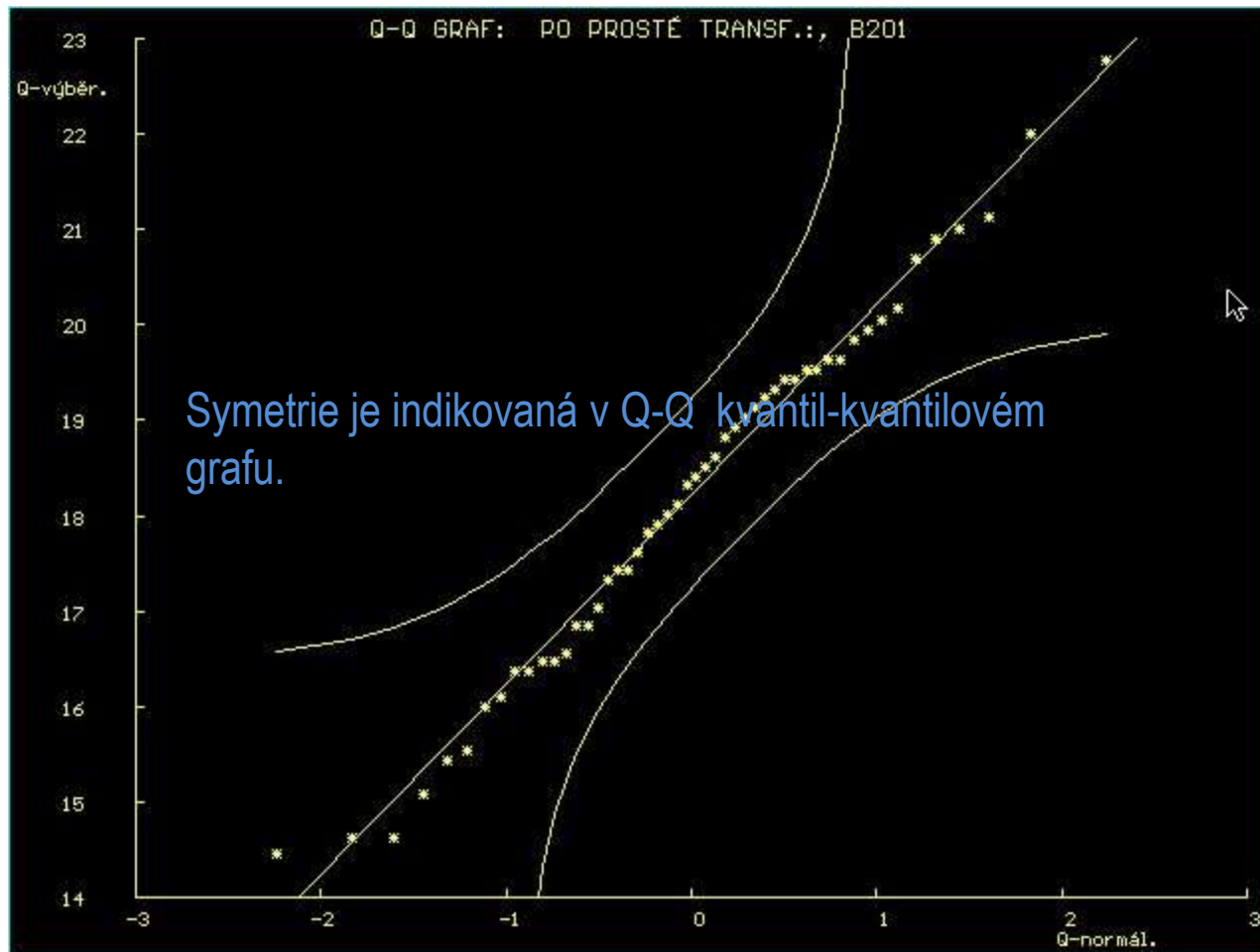
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





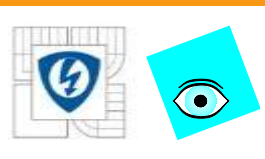
Ukázka postupného zlepšení symetrie rozdělení po transformaci dat



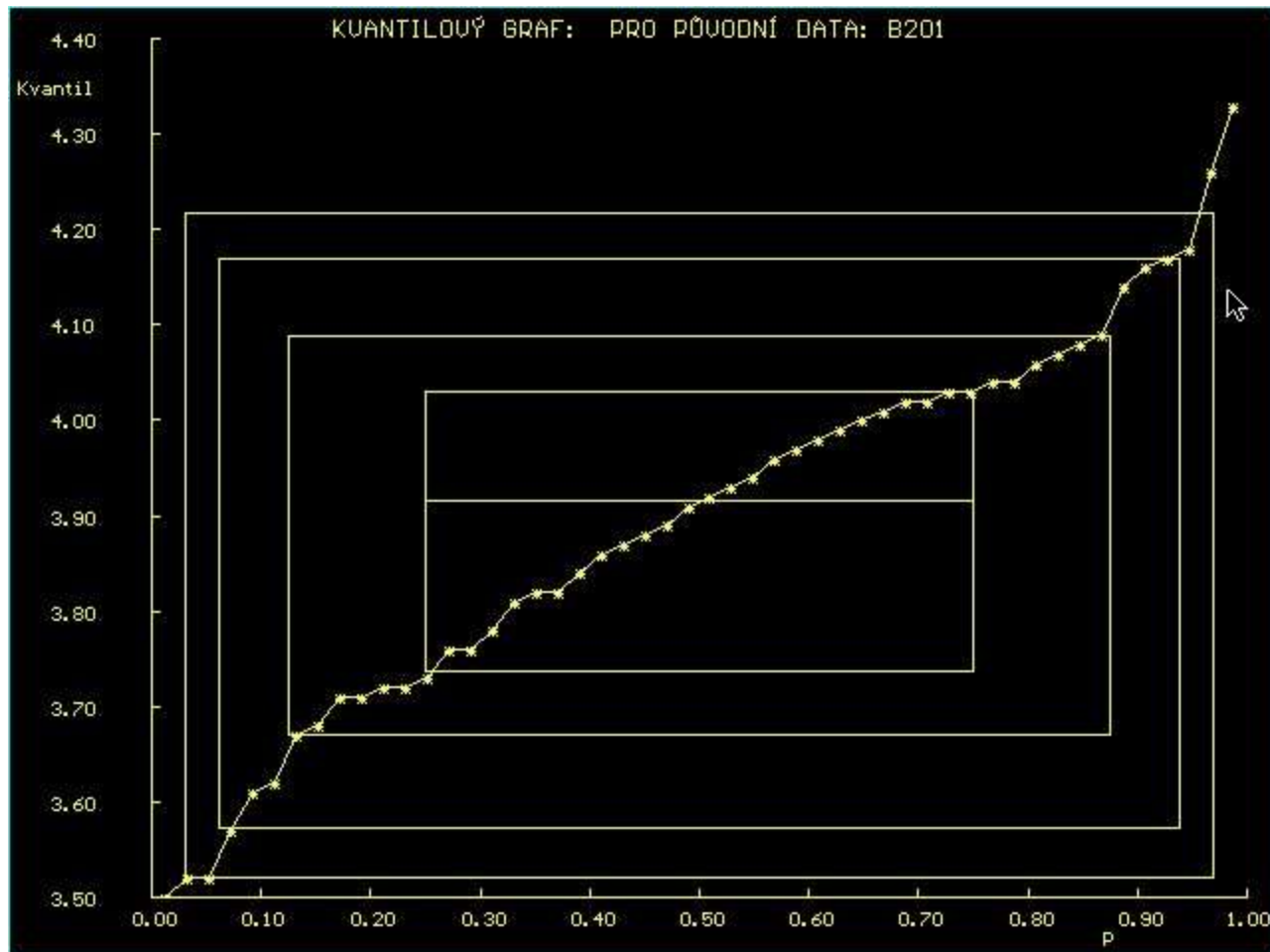
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





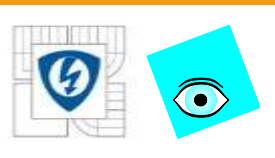
Ukázka postupného zlepšení symetrie rozdělení po transformaci dat.



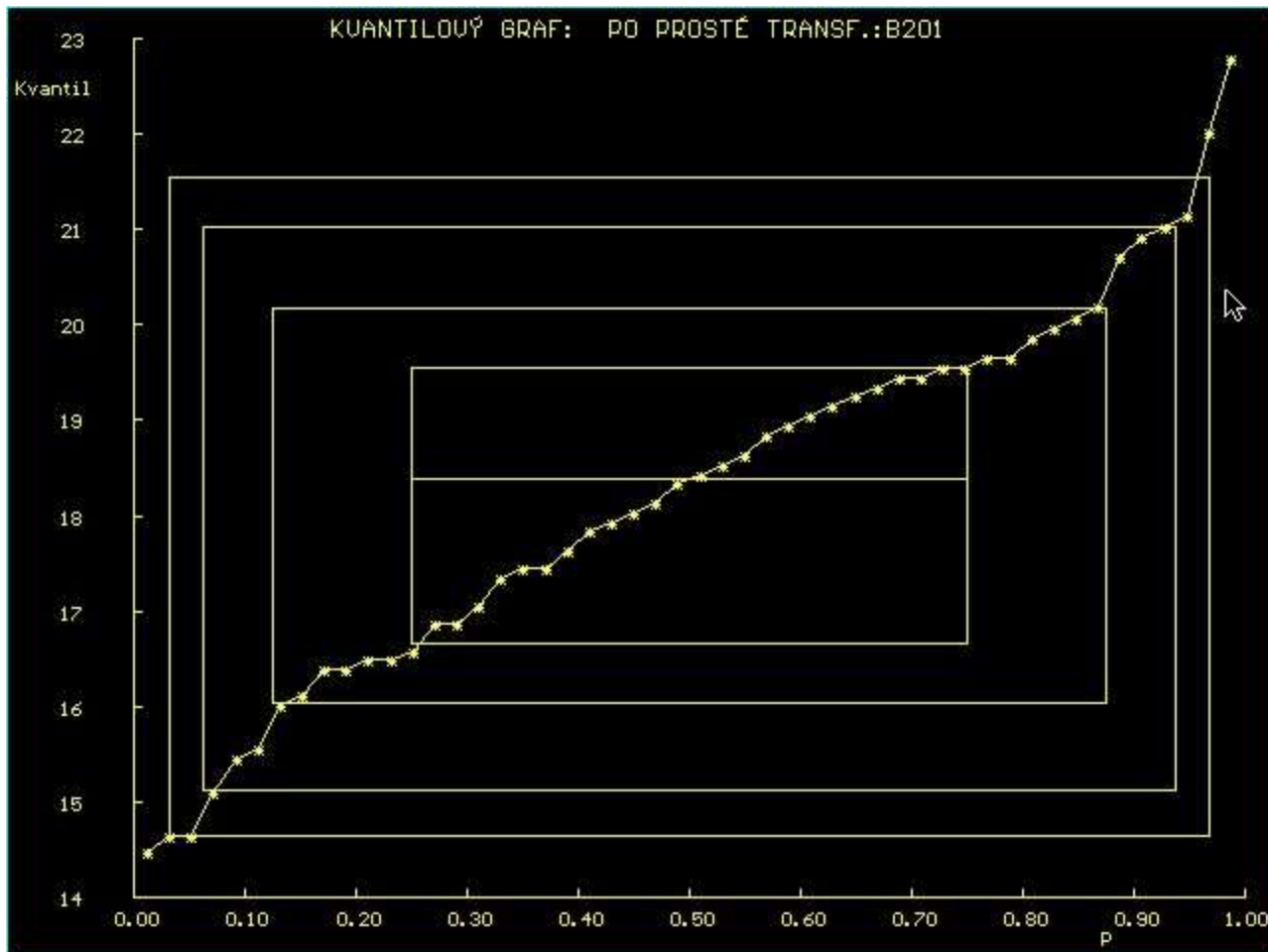
31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



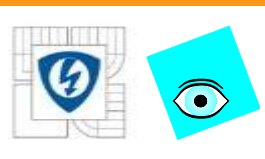


Ukázka postupného zlepšení symetrie rozdělení po transformaci dat.

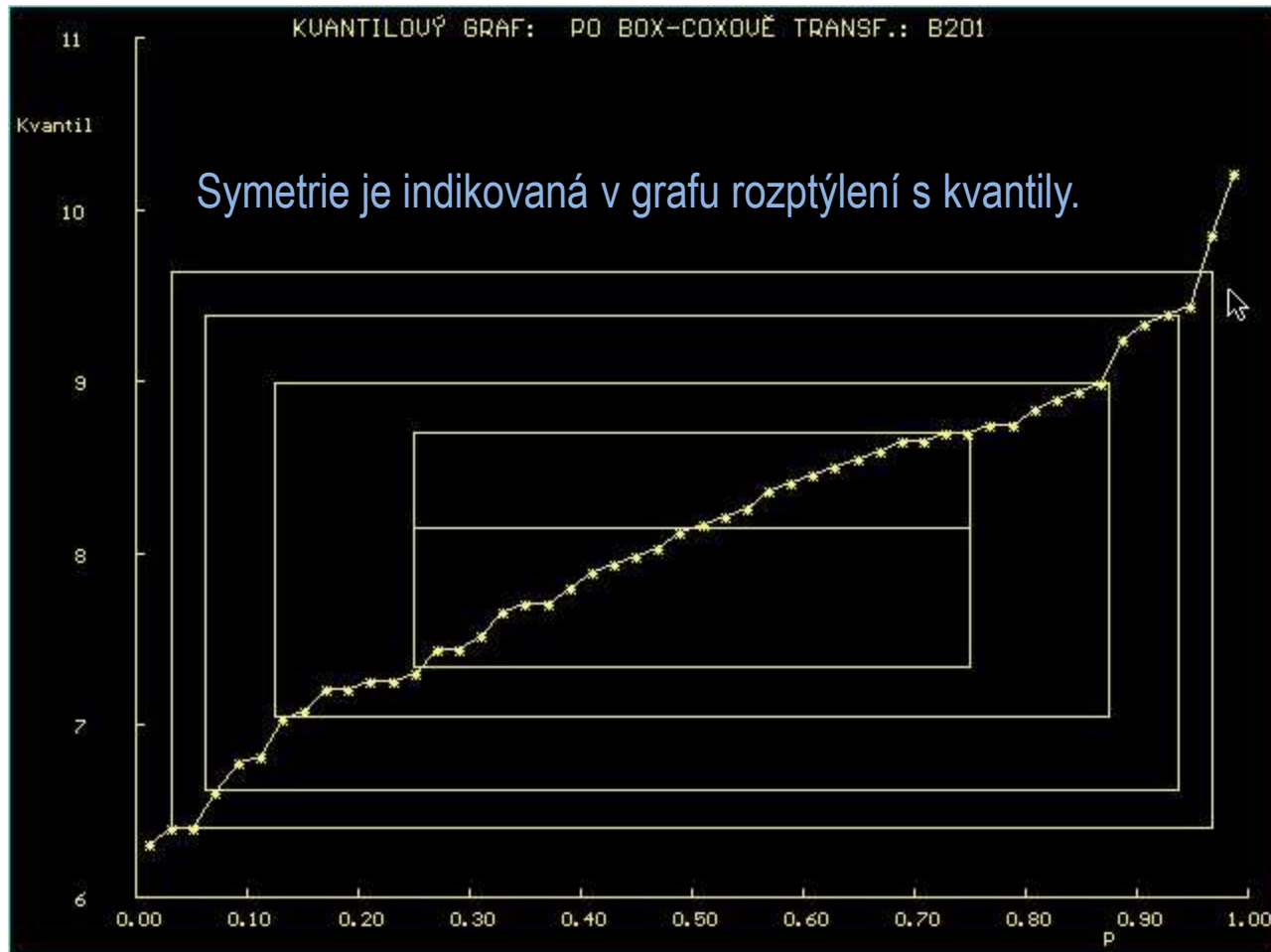


31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



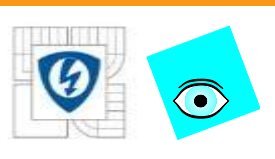
Ukázka postupného zlepšení symetrie rozdělení po transformaci dat.



31.1.2011

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





Načtení úlohy

Název úlohy : B201

Transformace :
☒ Exponenciální ☒ Box-Coxova
☐ Transformovaná data

Parametr ?
Parametr

Sloupce
B201
B202
B203
B204
B205
B206
B207a
B207b
B208

☒ Vyber vše

Data
☒ Všechna
☐ Označená
☐ Neoznačená
☐ Podle filtru

Popis [Žádný]

Název úlohy : B201

Transformace :
☒ Exponenciální ☒ Box-Coxova
☐ Transformovaná data

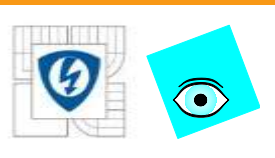
Parametr -0.05795288085937 ?
Parametr 2.22048034667969 ?

Sloupce
B201
B202
B203
B204
B205
B206
B207a
B207b
B208

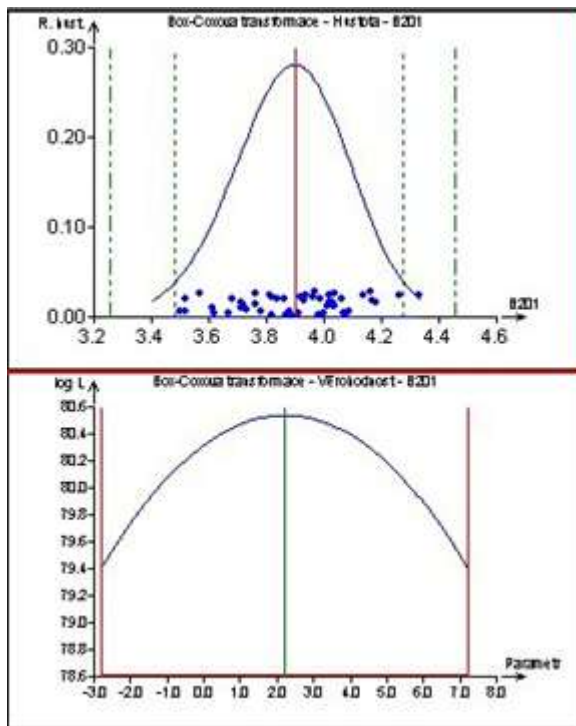
☒ Vyber vše

Data
☒ Všechna
☐ Označená
☐ Neoznačená
☐ Podle filtru

Popis [Žádný]

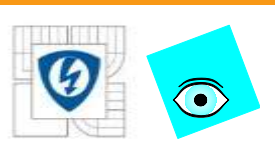


Výklad



Graf hustoty představuje tvar rozdělení, který nejlépe vystihuje data prostředky Box-Coxovy transformace. Svislé čáry představují kvantily (hodnoty) odpovídající (od středu) mediánu (50% kvantil), kvartilu (25% kvantily ohraničující 50% dat), $\pm 2s$ (zhruba 2.5% kvantily ohraničující interval 95% dat), 0.5% kvantily ohraničující 99% dat a $\pm 3s$ (ohraničující 99.73% dat).

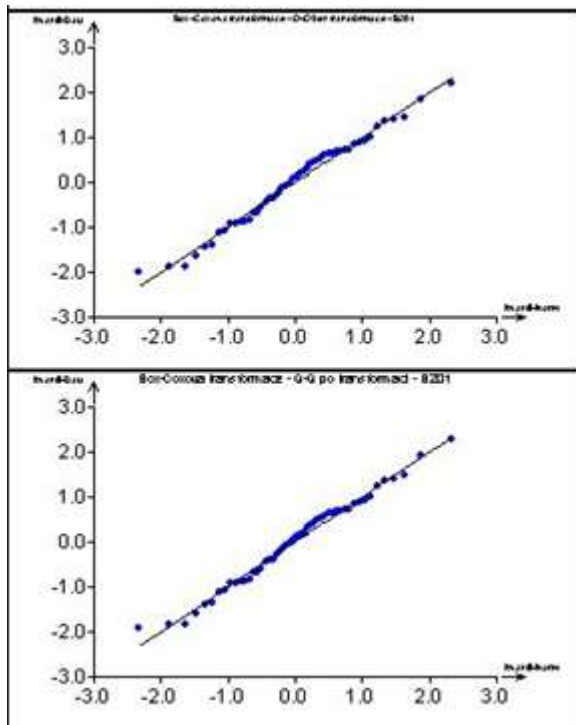
Graf logaritmu závislosti věrohodnostní funkce (osa y) na exponentu lambda. Maximu odpovídá optimální hodnota lambda. Vodorovná přímka odpovídá spodní mezi 95% intervalu spolehlivosti maxima věrohodnosti a svislé přímky odpovídají intervalu spolehlivosti odhadu lambda. Obsahuje-li tento interval 1, není nutné transformovat a je možné použít odhady v Základní statistice, případně transformaci s $\lambda = 1$.

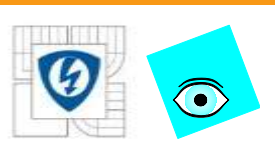


Výklad

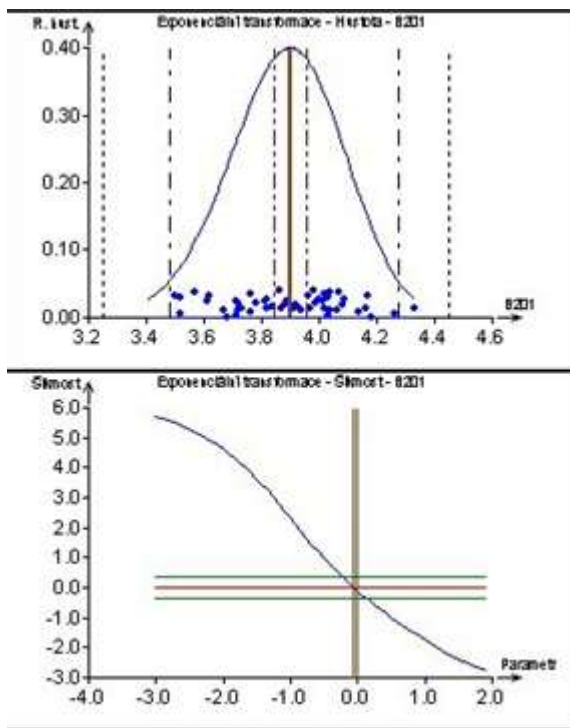
QQ-graf původních dat, shodný s QQ-grafem v Základní statistice. Metoda transformace bývá zpravidla užitečný jen pro systematicky prohnutý tvar bodů v QQ-grafu, viz vlevo. Proti statistikám má QQ-graf výhodu v možnosti vizuálně posoudit, zda je nelinearita (tedy odchylka od normality) způsobena jen několika body, nebo všemi daty.

QQ-graf dat po transformaci: Je-li tvar bodů blíže přímce než na předešlém grafu, je transformace úspěšná. Ke kvantitativnímu posouzení je však třeba použít statistik uvedených v protokolu.



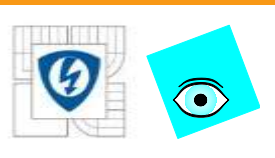


Výklad

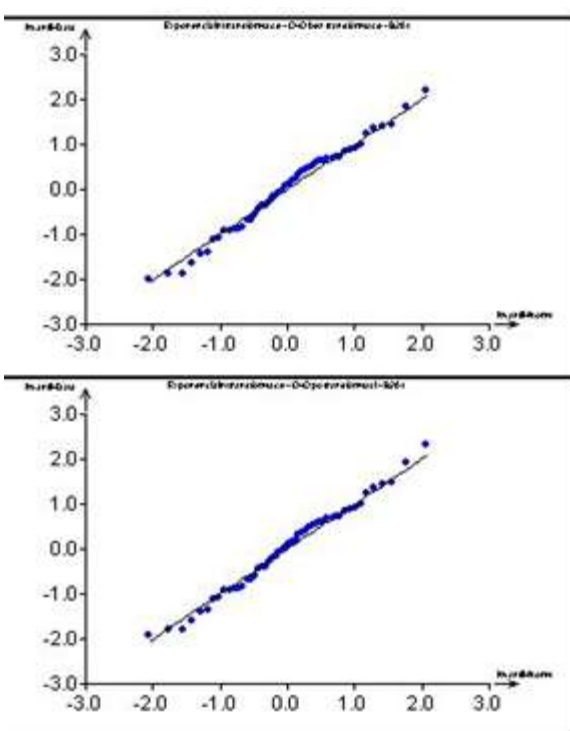


Graf hustoty představuje tvar rozdělení, který nejlépe vystihuje data prostředky Box-Coxovy transformace. Svislé čáry představují kvantily (hodnoty) odpovídající (od středu) mediánu (50% kvantil), kvartilu (25% kvantily ohraničující 50% dat), $\pm 2s$ (zhruba 2.5% kvantily ohraničující interval 95% dat), 0.5% kvantily ohraničující 99% dat a $\pm 3s$ (ohraničující 99.73% dat).

Závislost šikmosti transformovaných dat na parametru transformace: Nulová šikmost odpovídá optimálnímu parametru. Význam tohoto grafu je podobný jako u předchozího grafu věrohodnosti, slouží k nalezení parametru transformace a určení statistické významnosti transformace. Leží-li průsečík svislé zelené přímky s křivkou mimo interval spolehlivosti šikmosti (vodorovné zelené přímky), je transformace opodstatněná.

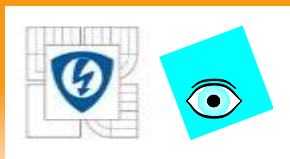


Výklad



QQ-graf původních dat, shodný s QQ-grafem v Základní statistice. Metoda transformace bývá zpravidla užitečný jen pro systematicky prohnutý tvar bodů v QQ-grafu, viz vlevo. Proti statistikám má QQ-graf výhodu v možnosti vizuálně posoudit, zda je nelinearita (tedy odchylka od normality) způsobena jen několika body, nebo všemi daty.

QQ-graf dat po transformaci: Je-li tvar bodů blíže přímce než na předešlém grafu, je transformace úspěšná. Ke kvantitativnímu posouzení je však třeba použít statistik uvedených v protokolu.



Optimální parametr

Nejlepší doporučená hodnota parametru r , při níž je dosaženo nejlepší shody s normálním rozdělením na základě maximální věrohodnosti.

Dolní a horní mez parametru

Interval spolehlivosti optimální hodnoty r . Hodnoty uvnitř tohoto intervalu poskytnou podobný efekt jako hodnota optimální. Interval se obvykle zužuje s rostoucím počtem dat. Je-li uvnitř tohoto intervalu jednička, není účelné data transformovat buď z toho důvodu, že data již normální jsou, nebo je dat příliš málo, interval spolehlivosti r je příliš široký a není možné najít jednoznačnou transformaci. Je-li uvnitř intervalu nula, lze rozdělení dat považovat za lognormální.

Věrohodnost bez transformace

Hodnota logaritmu věrohodnosti normality netransformovaných dat vzhledem k normálnímu rozdělení. Protože se jedná o logaritmy, odpovídá rozdílu o 1 řádovému rozdílu.

Věrohodnost transformací

Hodnota logaritmu věrohodnosti po transformaci s optimálním parametrem, tedy maximální dosažitelná shoda s normálním rozdělením. Protože se jedná o **logaritmy, odpovídá rozdílu o 1 řádovému rozdílu**.

Oprávněnost transformace

Slovní vyjádření opodstatnění transformace. NE znamená, že transformace neposkytuje významný přínos. ANO znamená doporučení transformace. Transformace se doporučuje, je-li účinnost transformace vyšší než 95%.

Účinnost transformace

Statistická významnost transformace. Je v podstatě kvantitativním vyjádřením oprávněnosti transformace (předchozí položka). Je-li účinnost větší než 95%, považuje se transformace za oprávněnou a doporučenou, jinak se transformace nedoporučuje. Hranice 95% však není striktní. Je-li účinnost blízká 95%, lze rozhodnout i opačně.

Zvolený parametr

Hodnota uživatelem zadaného parametru v dialogovém panelu. Tato hodnota může být různá od doporučené optimální hodnoty.

Věrohodnost

Logaritmus věrohodnosti odpovídající zvolenému parametru.

Opravený průměr

Aritmetický průměr vypočítaný metodou Box-Coxovy transformace. V případě asymetrických dat odpovídá střední hodnotě lépe než prostý průměr počítaný v Základní statistice.

LCL

Doporučená hodnota spodní kontrolní meze pro případ konstrukce Shewhartova regulačního diagramu typu X-průměr. Počet sloupců je chápán jako velikost podskupiny.

UCL

Doporučená hodnota spodní kontrolní meze pro případ konstrukce Shewhartova regulačního diagramu typu X-průměr. Počet sloupců je chápán jako velikost podskupiny.

LWL

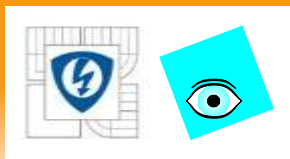
Doporučená hodnota spodní varovné meze.

UWL

Doporučená hodnota horní varovné meze.

Box-Coxova transformace dat :

Název úlohy :	B201
Data:	Všechna
Vybrané sloupce :	B201
Optimální parametr :	2.220480347
Dolní mez parametru :	-2.779519653
Horní mez parametru :	7.220480291
Věrohodnost bez transformace :	80.47189562
Věrohodnost s transformací :	80.53909592
Oprávněnost transformace :	Ne
Pravděpodobnost :	28.6029656815284 %
Zvolený parametr :	2.220480347
Věrohodnost :	80.53909592
Opravený průměr :	3.898743595
LCL :	3.254893143
UCL :	4.453982689
LWL :	3.695897375
UWL :	4.091950865



Optimální parametr

Nejlepší doporučená hodnota parametru r , při níž je dosaženo nejlepší shody s normálním rozdělením na základě maximální věrohodnosti.

Dolní a horní mez parametru

Interval spolehlivosti optimální hodnoty r . Hodnoty uvnitř tohoto intervalu poskytnou podobný efekt jako hodnota optimální. Interval se obvykle zužuje s rostoucím počtem dat. Je-li uvnitř tohoto intervalu jednička, není účelné data transformovat buď z toho důvodu, že data již normální jsou, nebo je dat příliš málo, interval spolehlivosti r je příliš široký a není možné najít jednoznačnou transformaci. Je-li uvnitř intervalu nula, lze rozdělení dat považovat za lognormální.

Věrohodnost bez transformace

Hodnota logaritmu věrohodnosti normality netransformovaných dat vzhledem k normálnímu rozdělení. Protože se jedná o logaritmy, odpovídá rozdílu o 1 řádovému rozdílu.

Věrohodnost transformací

Hodnota logaritmu věrohodnosti po transformaci s optimálním parametrem, tedy maximální dosažitelná shoda s normálním rozdělením. Protože se jedná o **logaritmy, odpovídá rozdílu o 1 řádovému rozdílu**.

Oprávněnost transformace

Slovní vyjádření opodstatnění transformace. NE znamená, že transformace neposkytuje významný přínos. ANO znamená doporučení transformace. Transformace se doporučuje, je-li účinnost transformace vyšší než 95%.

Účinnost transformace

Statistická významnost transformace. Je v podstatě kvantitativním vyjádřením oprávněnosti transformace (předchozí položka). Je-li účinnost větší než 95%, považuje se transformace za oprávněnou a doporučenou, jinak se transformace nedoporučuje. Hranice 95% však není striktní. Je-li účinnost blízká 95%, lze rozhodnout i opačně.

Zvolený parametr

Hodnota uživatelem zadaného parametru v dialogovém panelu. Tato hodnota může být různá od doporučené optimální hodnoty.

Věrohodnost

Logaritmus věrohodnosti odpovídající zvolenému parametru.

Opravený průměr

Aritmetický průměr vypočítaný metodou Box-Coxovy transformace. V případě asymetrických dat odpovídá střední hodnotě lépe než prostý průměr počítaný v Základní statistice.

LCL

Doporučená hodnota spodní kontrolní meze pro případ konstrukce Shewhartova regulačního diagramu typu X-průměr. Počet sloupců je chápán jako velikost podskupiny.

UCL

Doporučená hodnota spodní kontrolní meze pro případ konstrukce Shewhartova regulačního diagramu typu X-průměr. Počet sloupců je chápán jako velikost podskupiny.

LWL

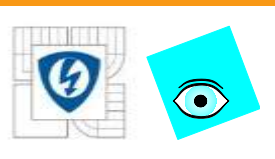
Doporučená hodnota spodní varovné meze.

UWL

Doporučená hodnota horní varovné meze.

Exponenciální transformace dat :

Název úlohy :	B201
Data:	Všechna
Vybrané sloupce :	B201
Optimální parametr :	-0.05795288086
Zvolený parametr :	-0.05795288086
Oprávněnost transformace :	Ne
Opravený průměr :	3.899547054
Interval spolehlivosti :	
Spodní :	3.842812816
Horní :	3.95535866
LCL :	3.250862643
UCL :	4.447294615
LWL :	3.479724116
UWL :	4.274265575



Parametry polohy, rozptýlení a tvaru

Analýza 1 výběru: klasické odhady (průměr, rozptyl), robustní odhady (medián, uřezané průměry, winsorizovaný rozptyl, interkvantilové rozpětí), adaptivní odhady

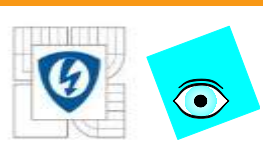
```
U Ý S L E D K Y
ZÁKLADNÍ STATISTIKA
Analýza jednorozměrného výběru

Název : B201

U S T U P

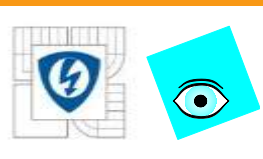
(1) DATA A PODMÍNKY:
    Počet dat : 50
    Hladina významnosti alfa : 0.050
    Název výstupního souboru : RESULTS.TXT

(2) VSTUPNÍ DATA:
    ( 1) 3.8600E+00 ( 14) 4.0400E+00 ( 27) 4.0200E+00 ( 40) 4.0200E+00
    ( 2) 4.0600E+00 ( 15) 3.9300E+00 ( 28) 3.8200E+00 ( 41) 4.0800E+00
    ( 3) 3.6700E+00 ( 16) 4.1400E+00 ( 29) 3.6200E+00 ( 42) 4.0400E+00
    ( 4) 3.9700E+00 ( 17) 3.8800E+00 ( 30) 3.6800E+00 ( 43) 3.7800E+00
    ( 5) 3.7600E+00 ( 18) 3.8400E+00 ( 31) 4.1700E+00 ( 44) 3.9800E+00
Napověďa-F1 Řádek: 1 - 23 Celkem: 112 Délka: 3926
```



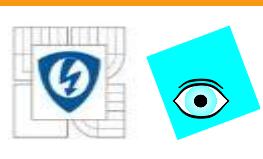
Parametry polohy, rozptýlení a tvaru

V Ý S L E D K Y			
(1) PARAMETRY TVARU:			
	Šikmost	:	-1.1912E-01
	Špičatost	:	2.4016E+00
(2) KLASICKÉ ODHADY PARAMETRŮ :			
	Průměr	:	3.8940E+00
	Směr. odchylka	:	1.9849E-01
	Rozptyl	:	3.9400E-02
	95.0% spolehlivost:		
	Spodní mez:	3.8376E+00	Horní mez: 3.9504E+00
(3) OSTATNÍ ODHADY POLOHY:			
	Odhad modu	:	4.0300E+00
	Odhad polosumy	:	3.9150E+00
(4) ROBUSTNÍ ODHADY PARAMETRŮ :			
	Medián	:	3.9150E+00
	Směr. odchylka mediánu:		2.6014E-01
	Rozptyl mediánu	:	6.7675E-02
	Rozptyl (nepar.)	:	2.0250E-03
Napověda-F1		Řádek: 35 - 57	Celkem: 112 Délka: 3926



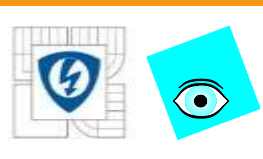
Parametry polohy, rozptýlení a tvaru

V Ý S L E D K Y			
(4) ROBUSTNÍ ODHADY PARAMETRŮ :			
Medián	:	3.9150E+00	
Směr. odchylka mediánu:		2.6014E-01	
Rozptyl mediánu	:	6.7675E-02	
Rozptyl (nepar.)	:	2.0250E-03	
Směr. odchylka mediánu:		4.5000E-02	
Rozptyl (Marritz)	:	1.9250E-03	
Směr. odchylka mediánu:		4.3875E-02	
95.0% spolehlivost:			
Spodní mez:	3.8268E+00	Horní mez:	4.0032E+00
Uřezání 5% (pro P=0.05):			
Průměr	:	3.8942E+00	
Směr. odchylka	:	2.0668E-01	
Rozptyl	:	4.2717E-02	
Průměr, winsor.	:	3.8898E+00	
St.odch. winsor.	:	1.9526E-01	
Rozptyl, winsor.	:	3.8128E-02	
95.0% spolehlivost:			
Spodní mez:	3.8356E+00	Horní mez:	3.9528E+00
Uřezání 10% (pro P=0.10):			
Napověda-F1		Řádek: 53 - 75	Celkem: 112 Délka: 3926



Parametry polohy, rozptýlení a tvaru

V Ý S L E D K Y			
Uřezání 10% (pro $P=0.10$):			
Průměr	:	3.8970E+00	
Směr. odchylka	:	2.1447E-01	
Rozptyl	:	4.5998E-02	
Průměr, winsor.	:	3.8936E+00	
St.odch. winsor.	:	1.8942E-01	
Rozptyl, winsor.	:	3.5879E-02	
95.0% spolehlivost:			
Spodní mez:	3.8357E+00	Horní mez:	3.9583E+00
Uřezání 40% (pro $P=0.40$):			
Průměr	:	3.9130E+00	
Směr. odchylka	:	2.7268E-01	
Rozptyl	:	7.4357E-02	
Průměr, winsor.	:	3.9146E+00	
St.odch. winsor.	:	1.1569E-01	
Rozptyl, winsor.	:	1.3384E-02	
95.0% spolehlivost:			
Spodní mez:	3.8271E+00	Horní mez:	3.9989E+00
Biweight:			
Průměr	:	3.8966E+00	
Směr. odchylka	:	1.9826E-01	
Napověda-F1		Řádek: 75 - 97	Celkem: 112 Délka: 3926



Parametry polohy, rozptýlení a tvaru

V Ý S L E D K Y			
Biweight:			
Průměr	:	3.8966E+00	
Směr. odchylka	:	1.9826E-01	
Rozptyl	:	3.9306E-02	
Váhy sqrt(w)	:	6.7659E+00	
95.0% spolehlivost:			
Spodní mez:	3.8377E+00	Horní mez:	3.9555E+00
(5) ADAPTIVNÍ ODHADY PARAMETRŮ:			
Hoggovy odhady:			
Relativní délka konců	:	2.3753E+00	
Průměr	:	3.8940E+00	
Směr. odchylka	:	1.9849E-01	
Rozptyl	:	3.9400E-02	
95.0% spolehlivost:			
Spodní mez:	3.8376E+00	Horní mez:	3.9504E+00
Napověda-F1		Řádek: 94 - 112	Celkem: 112 Délka: 3926