

2. kapitola

EDA: Transformace dat

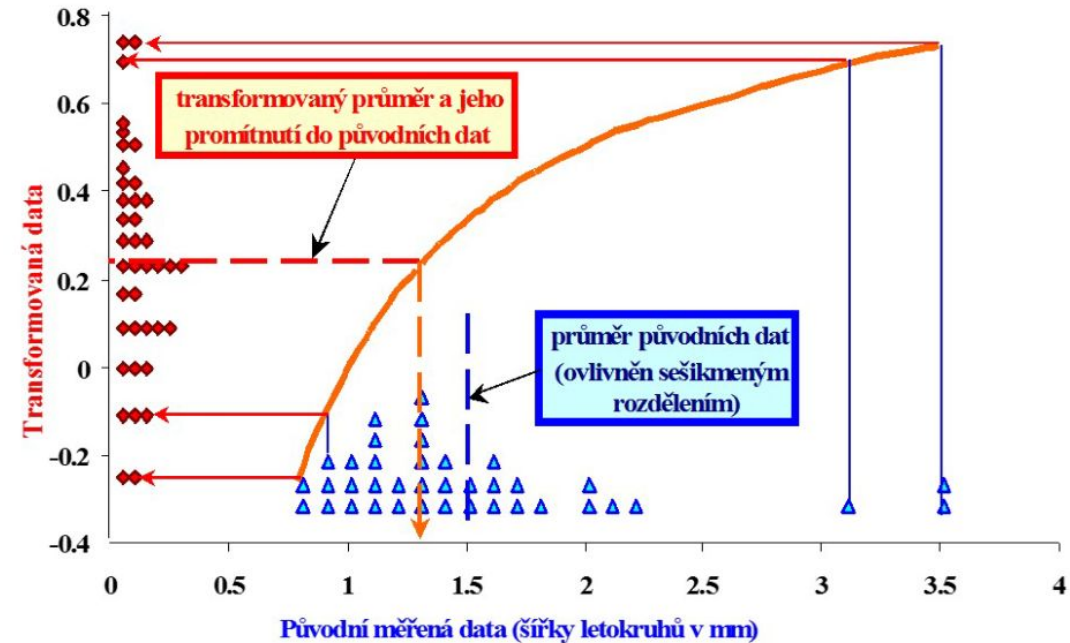
Symetrizující mocninná transformace

zesymetřičtění rozdělení výběru provede

$$y = g(x) = \begin{cases} x^\lambda & \lambda > 0 \\ \ln x & \lambda = 0 \\ -x^{-\lambda} & \lambda < 0 \end{cases} \quad \text{pro}$$

1. Nezachovává měřítko,
2. Není k λ všude spojitá,
3. Zachovává pořadí dat ve výběru,
4. Hodí se pouze pro kladná data.

TRANSFORMACE DAT - ukázka



Normalizační Box-Coxova transformace

přiblížení rozdělení výběru k normálnímu

$$y = g(x) = \begin{cases} \frac{x^\lambda - 1}{\lambda} & (\lambda \neq 0) \\ \ln x & (\lambda = 0) \end{cases}$$

Transformace

1. Je vzhledem k λ spojitá,
2. Je použitelná pouze pro kladná data, jinak $(x - x_0)$.

Schema mocinné a Box-Coxovy transformace

x	y (mocinná)	y (Box-Cox)
0.5	0.9117	-0.6621
0.9	0.9861	-1.0462
1.0	1.0000	0.0000
1.0	1.0000	0.0000
1.1	1.0128	0.0959
1.3	1.0356	0.2670
1.5	1.0556	0.4166
1.5	1.0556	0.4166
1.5	1.0556	0.4166
1.8	1.0815	0.6114
1.9	1.0893	0.6701
2.0	1.0968	0.7262
2.0	1.0968	0.7262
2.7	1.1416	1.0620
3.2	1.1678	1.2582
3.2	1.1678	1.2582
3.3	1.1726	1.2942
3.3	1.1726	1.2942
3.6	1.1862	1.3968
5.2	1.2459	1.8439
5.5	1.2552	1.9140
5.5	1.2552	1.9140
6.0	1.2699	2.0239
6.0	1.2699	2.0239
7.0	1.2962	2.2217
8.0	1.3195	2.3963
8.0	1.3195	2.3963
11.5	1.3849	2.8869

Zpětná transformace

$$\bar{y}, s^2(y), \bar{y} \pm t_{1-\alpha/2}(n-1) \cdot s(y)/\sqrt{n}$$

1. Nekorektní přístup:

$$\bar{x}_R = \bar{x}_\lambda = \left[\frac{\sum_{i=1}^n x_i^\lambda}{n} \right]^{1/\lambda}$$

$\bar{x}_R = \bar{x}_{-1}$ představuje *harmonický průměr*,
 $\bar{x}_R = \bar{x}_0$ představuje *geometrický průměr*,
 $\bar{x}_R = \bar{x}_1$ představuje *aritmetický průměr*,
 $\bar{x}_R = \bar{x}_2$ představuje *kvadratický průměr*.

Průměr

$$\bar{x} = 3.5714$$

$$\bar{y} = 1.1465$$

$$\bar{y} = 1.0987$$

Směrodatná odchylka

$$s = 2.7333$$

$$s = 0.1214$$

$$s = 0.9107$$

Retransform. průměr (mocnina)

$$\bar{x}_R = 2.788$$

$$s = 2.2147$$

Retransform. průměr (Box-Cox)

$$\bar{x}_R = 2.788$$

$$s = 2.2147$$

Správná hodnota je 2.79

2. **Korektní přístup:** z Taylorova rozvoje funkce $y = g(x)$ v okolí \bar{y} ,

$$\bar{x}_R \approx g^{-1} \left[\bar{y} - \frac{1}{2} \frac{d^2 g(x)}{dx^2} \left(\frac{dg(x)}{dx} \right)^{-2} s^2(y) \right]$$

Pro rozptyl vyjde $s^2(x_R) \approx \left(\frac{dg(x)}{dx} \right)^{-2} s^2(y)$.

(a) Pro speciální případ $\lambda = 0$, tzn. logaritmickou transformaci typu $g(x) = \ln x$, bude

$$\bar{x}_R \approx \exp [\bar{y} + 0.5 s^2(y)]$$

a rozptyl $s^2(x_R) \approx \bar{x}_R^{-2} s^2(y)$.

(b) Pro případ $\lambda \neq 0$:

$$\bar{x}_{R,1,2} = [0.5(1 + \lambda \bar{y}) \pm 0.5 \sqrt{1 + 2 \lambda (\bar{y} + s^2(y)) + \lambda^2 (\bar{y}^2 - 2 s^2(y))}]^{1/\lambda}$$

a rozptyl $s^2(x) = \bar{x}_R^{-2 \lambda + 2} s^2(y)$.

kde

$$I_D = g^{-1} \left(\bar{y} + G - t_{1-\alpha/2}(n-1) \frac{s(y)}{\sqrt{n}} \right)$$

$$I_H = g^{-1} \left(\bar{y} + G + t_{1-\alpha/2}(n-1) \frac{s(y)}{\sqrt{n}} \right)$$

$$G = -0.5 \frac{d^2 g(x)}{dx^2} \left(\frac{dg(x)}{dx} \right)^{-2} s^2(y)$$

kde $t_{1-\alpha/2}(n-1)$ je $100(1-\alpha/2)$ %ní kvantil Studentova rozdělení s $(n-1)$ stupni volnosti.

Rozptyl se vyčíslí

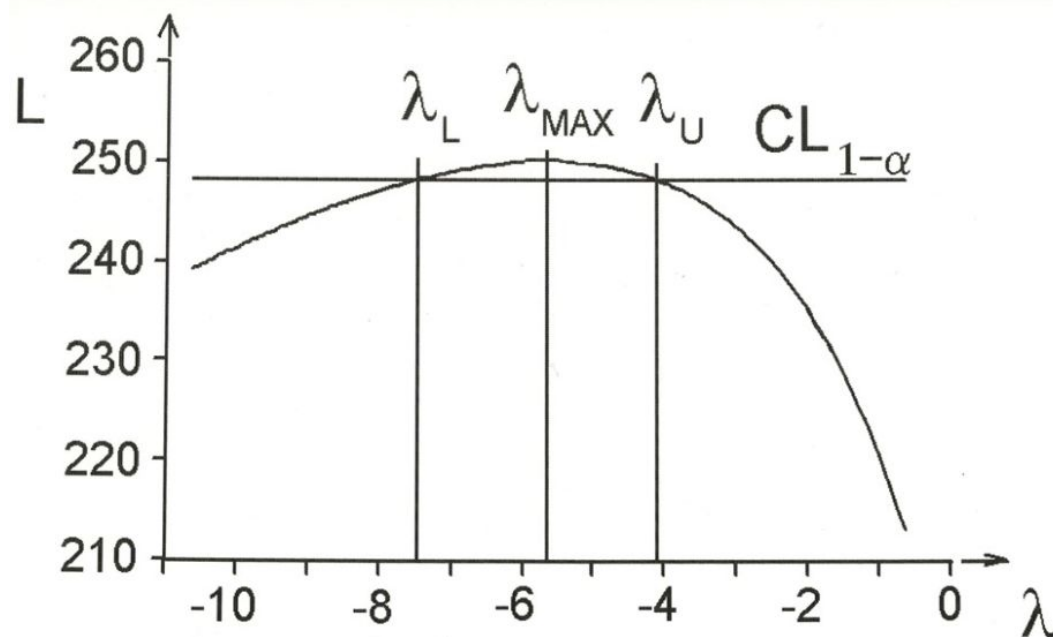
$$s^2(x_R) \approx \left(\frac{dg(x)}{dx} \right)^{-2} s^2(y)$$

kde derivace jsou vyčísleny v bodě $x = \bar{x}_R$.

100(1 - α)%ní interval spolehlivosti se vyčíslí dle

$$\bar{x}_R - I_D \leq \mu \leq \bar{x}_R + I_H$$

Graf maximální věrohodnosti



2. kapitola

EDA: Transformace dat

Řešení vzorového příkladu
z učebnice **Interaktivní statistická analýza dat**

P226

2. Interaktivní statistická analýza na PC

- Průzkumová analýza dat (EDA):**
Diagnostické grafy: stupeň symetrie rozdělení
lokální koncentrace
vybočující data
- Ověření předpokladů výběru dat:**
Diagnostiky: ověření normality,
ověření nezávislosti,
ověření homogenity,
ověření minimální četnosti.
- Transformace dat:** pro symetrii a pro heteroskedasticitu výběru
Box-Coxova transformace,
Mocninná transformace.
- Odhady parametrů polohy, rozptýlení a tvaru:**
Klasické odhady: aritmetický průměr, rozptyl
Robustní odhady: medián, uřezané průměry,
neparametrické odhady rozptylu.
Adaptivní odhady polohy a rozptýlení.

Zadání úlohy

Příklad 2.26 Atestace stopového Cu v kaolinu

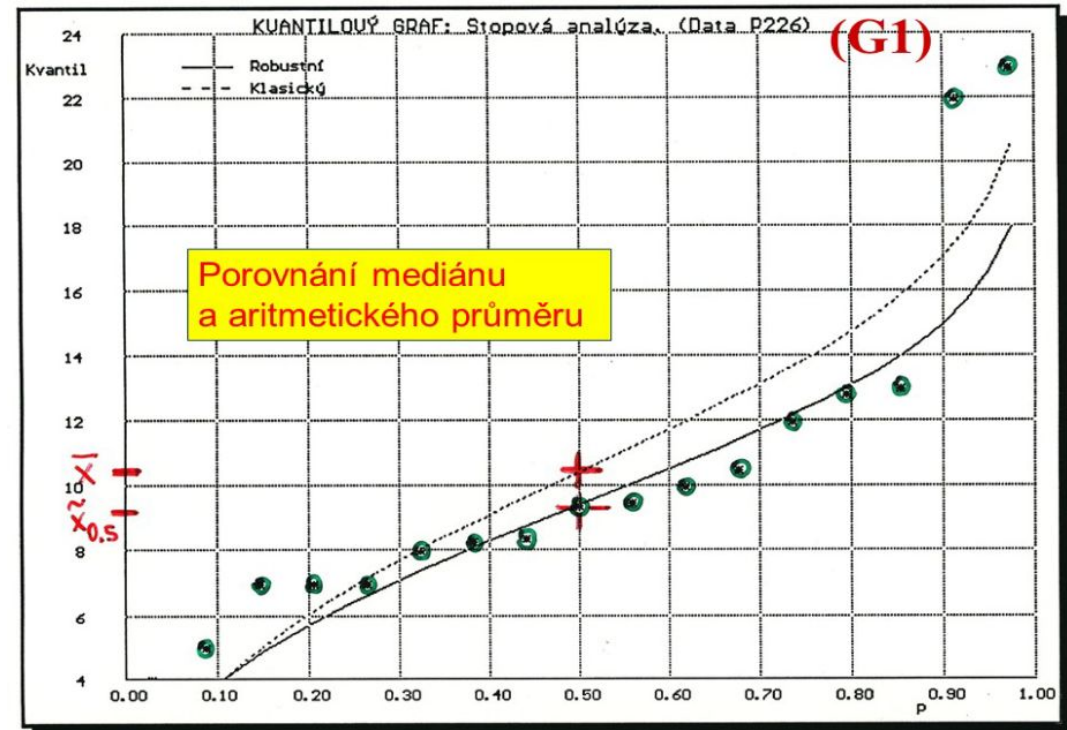
Ve vzorku kaolinu byl stanoven obsah Cu (ppm) a hodnoty ve vztupném pořadí jsou 4, 5, 7, 7, 7, 8, 8.3, 8.4, 9.4, 9.5, 10.0, 10.5, 12.0, 12.8, 13.0, 22.0, 23.0.

Stanovte míry polohy a rozptýlení uvedeného výběru.

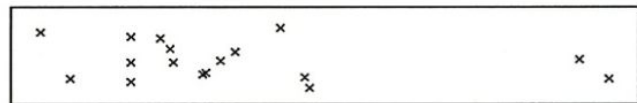
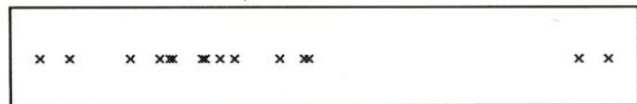
Řešení:

1. Výpočet na kalkulačce

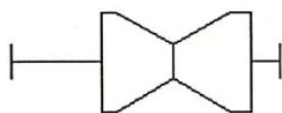
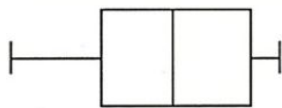
Aritmetický průměr	\bar{x}	=	10.406
Medián	$\tilde{x}_{0.5}$	=	9.400



BODOVÉ A KRABICOVÉ GRAFY: Stopová analýza, (Data P226)



Poloha mediánu

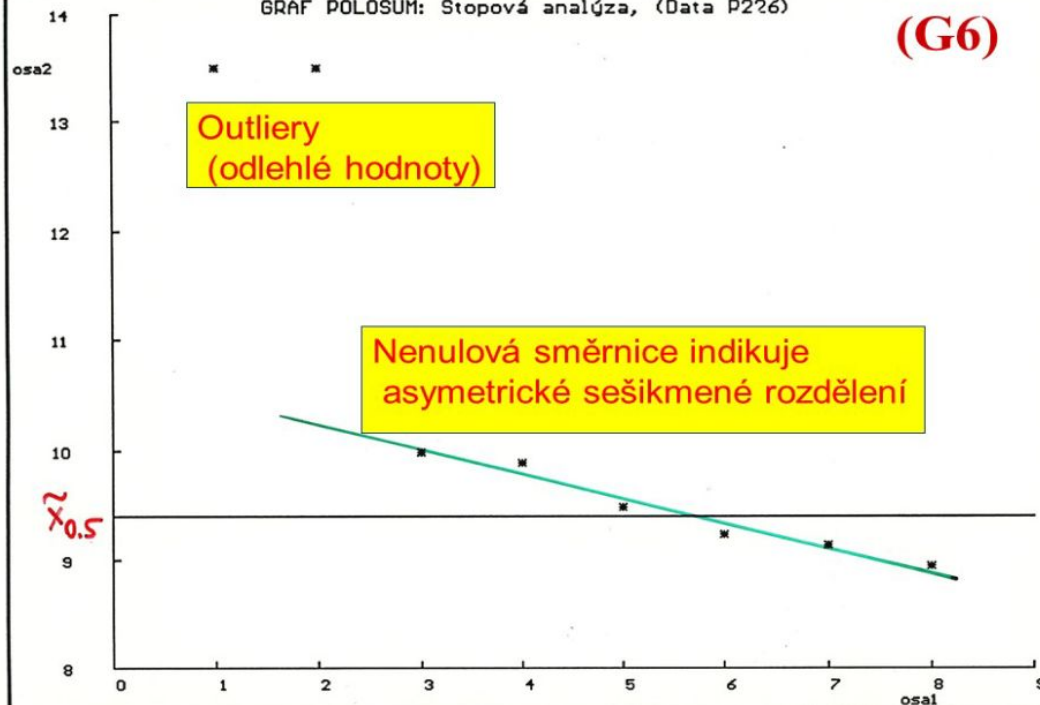


$\tilde{x}_{0.5}$



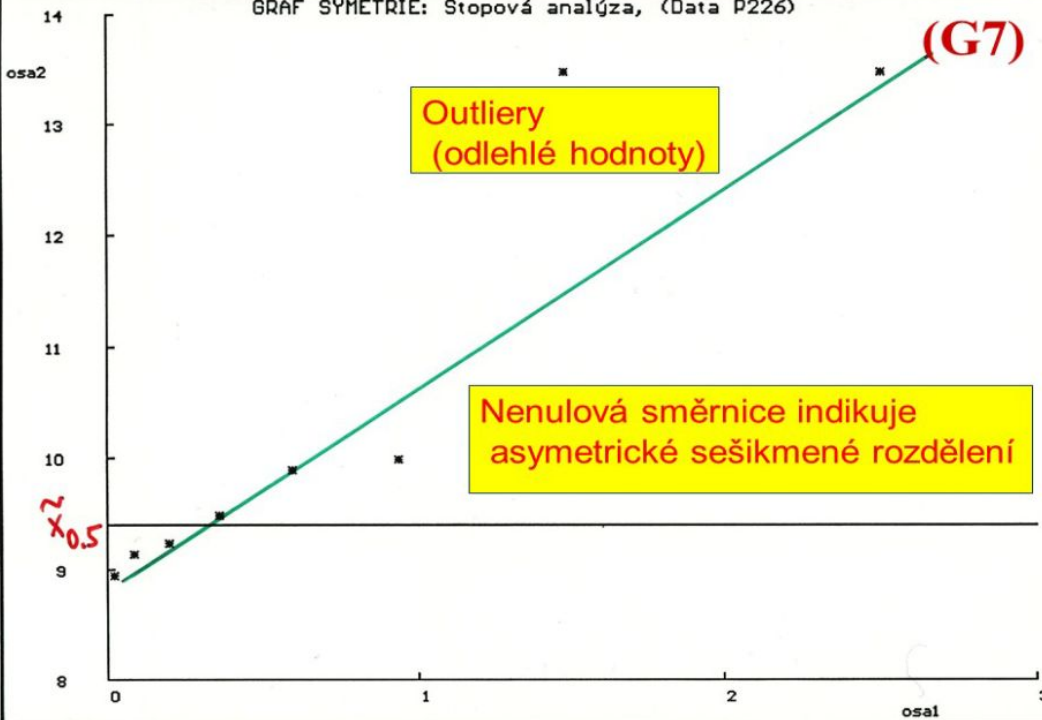
GRAF POLOSUM: Stopová analýza, (Data P226)

(G6)



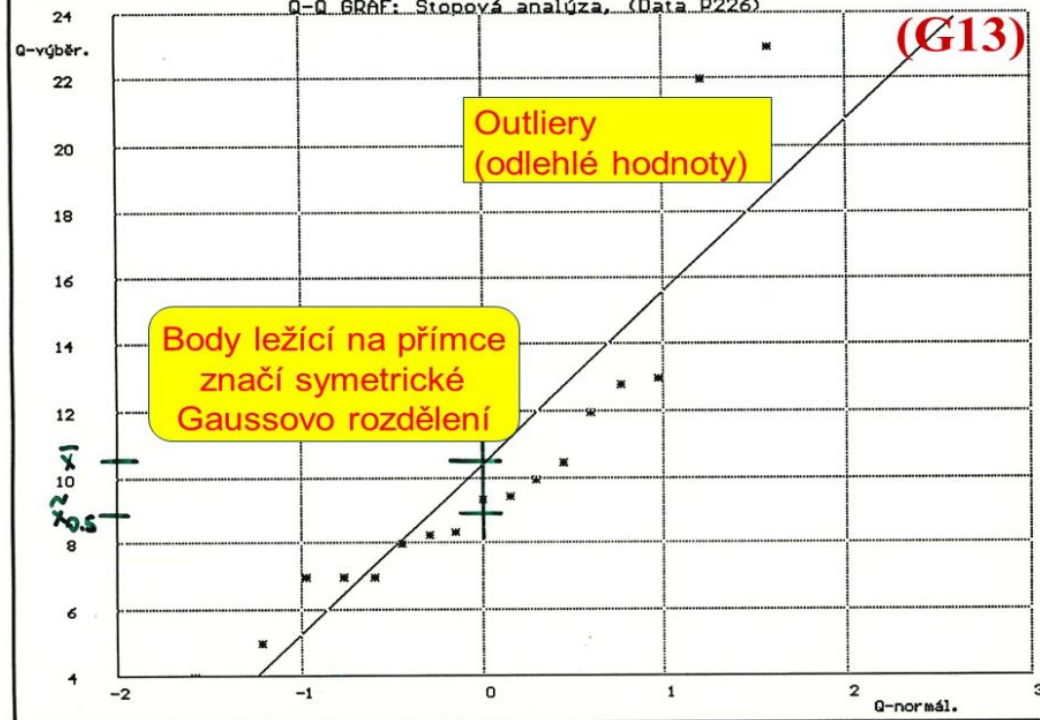
GRAF SYMETRIE: Stopová analýza, (Data P226)

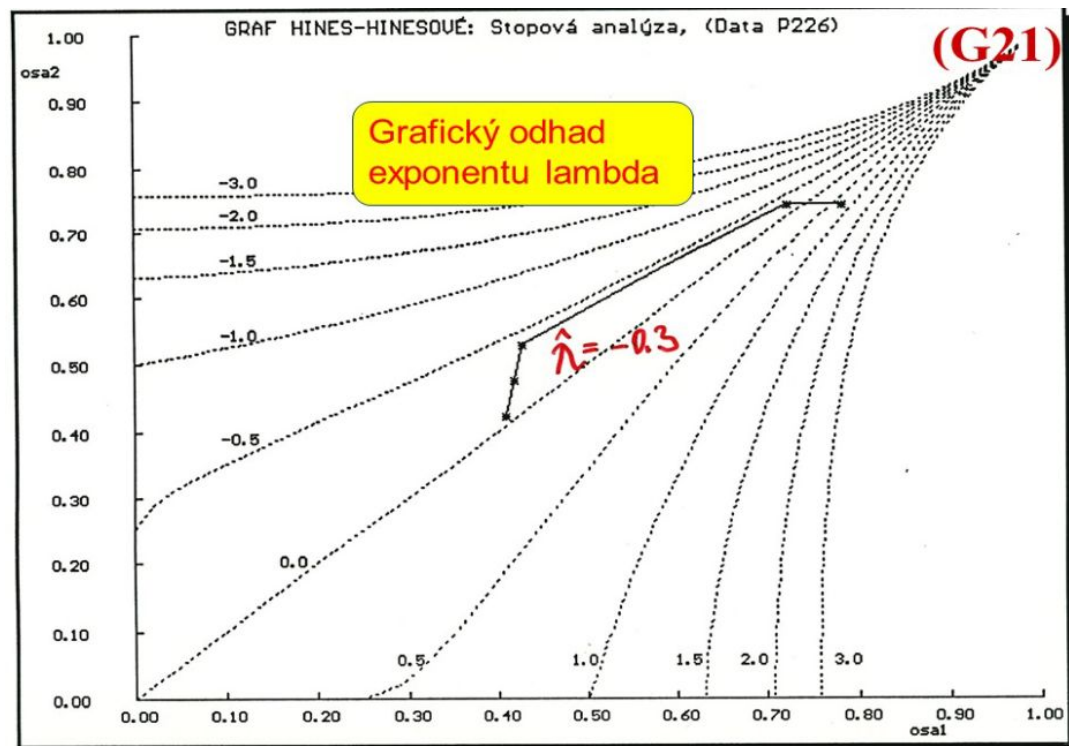
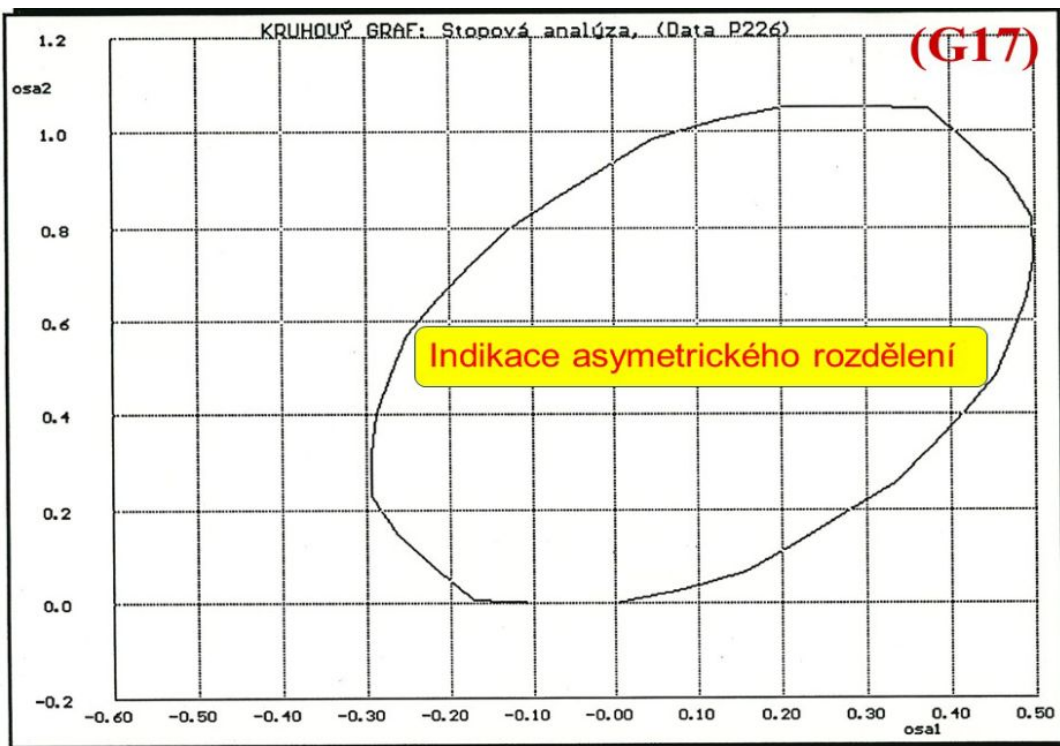
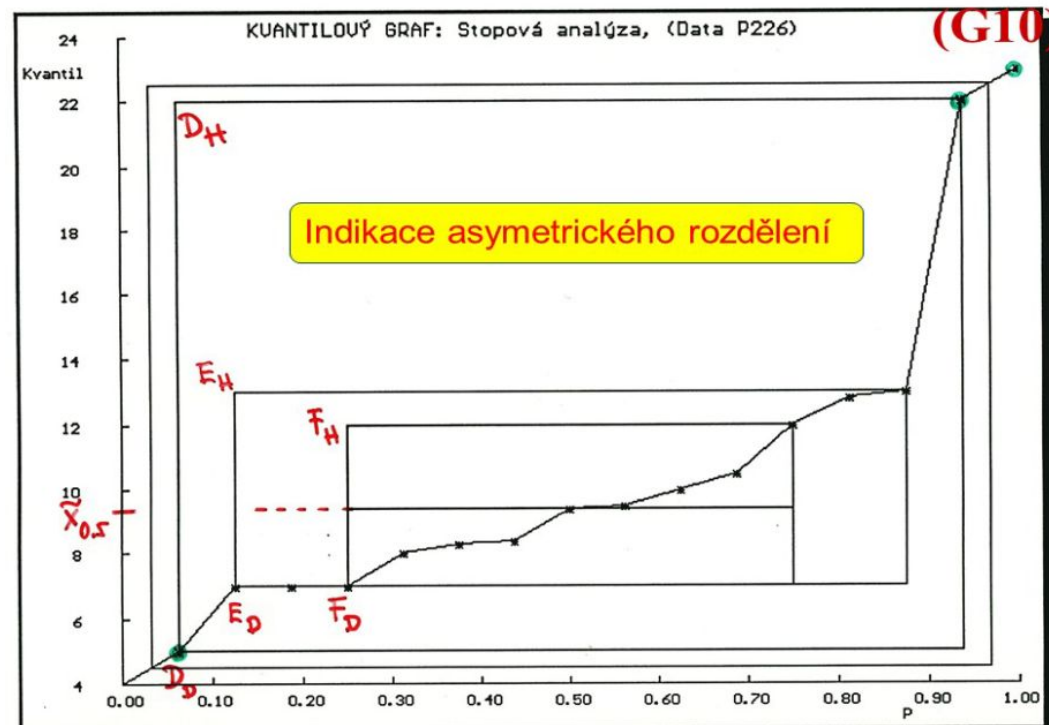
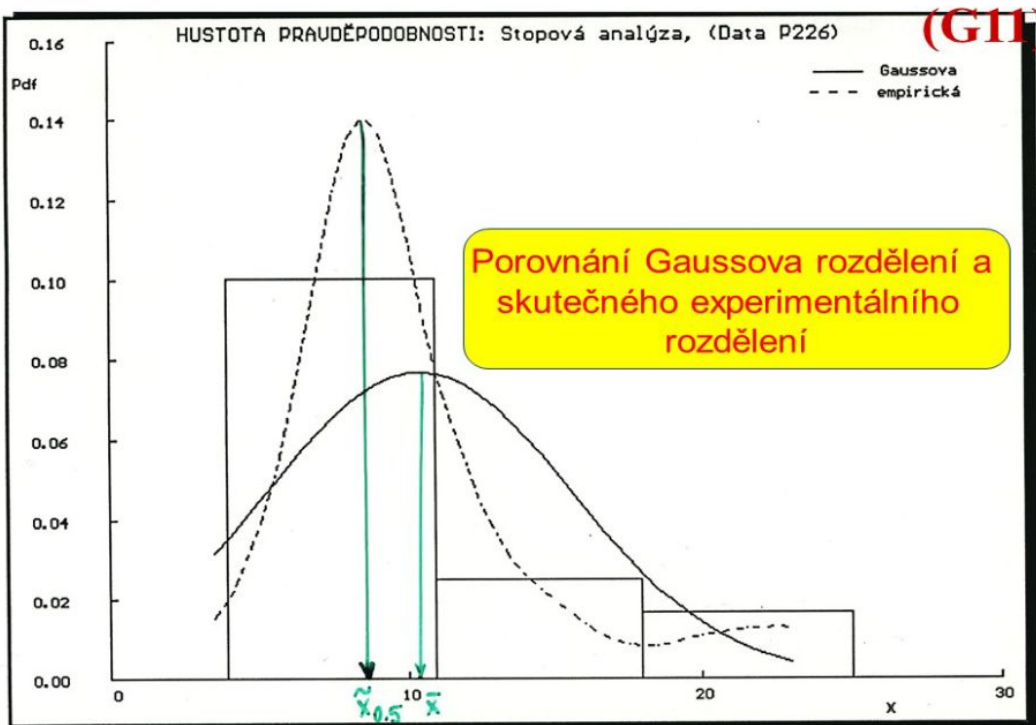
(G7)

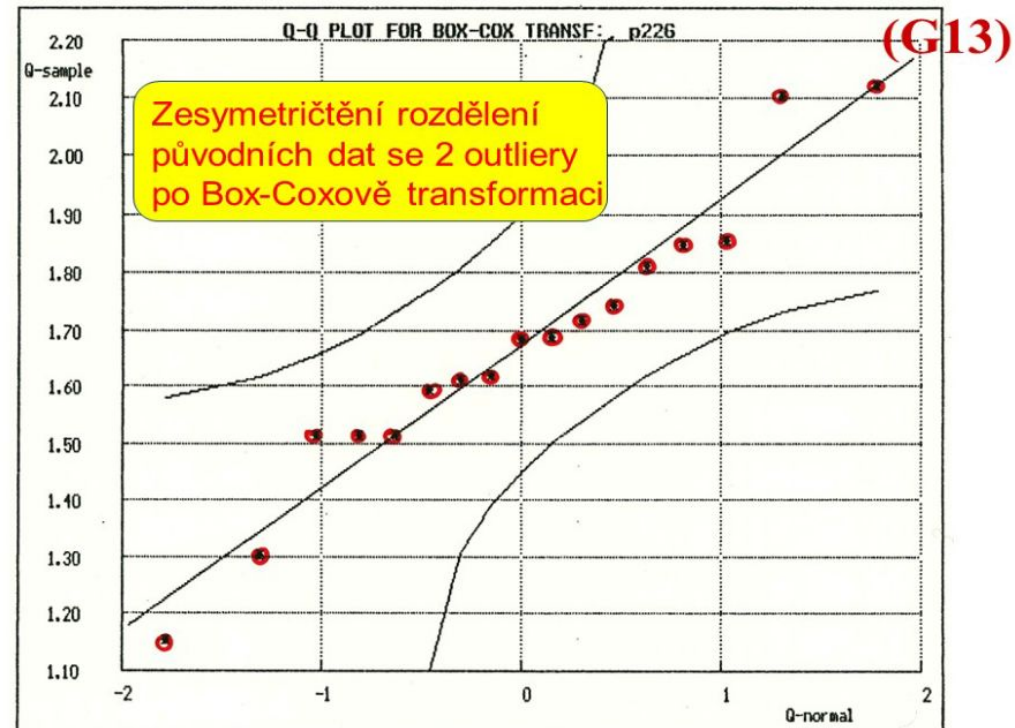
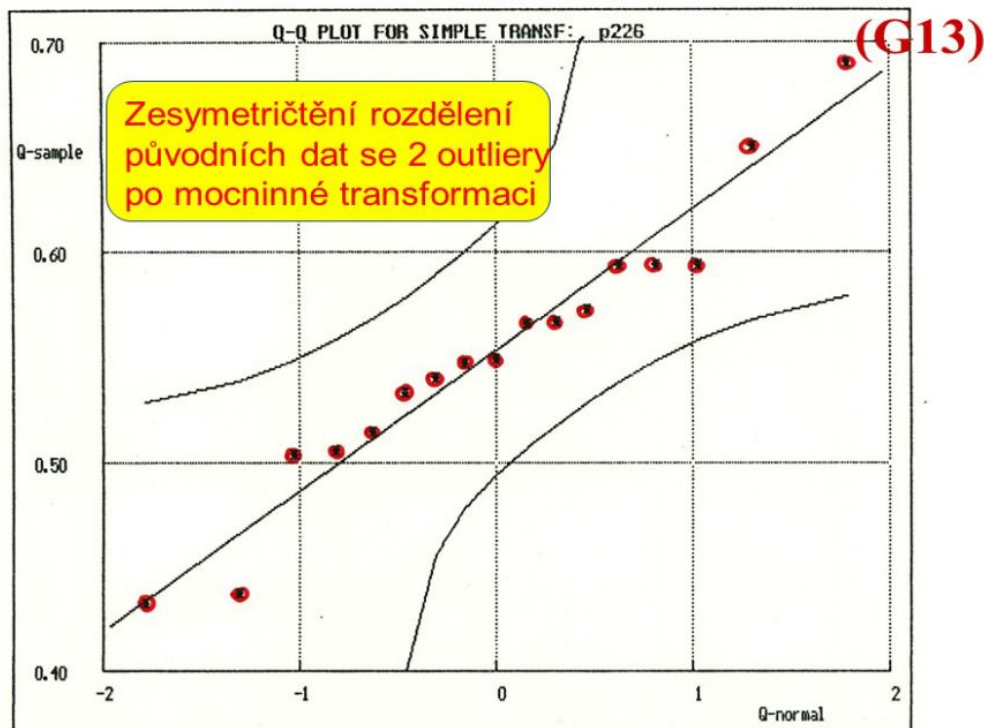
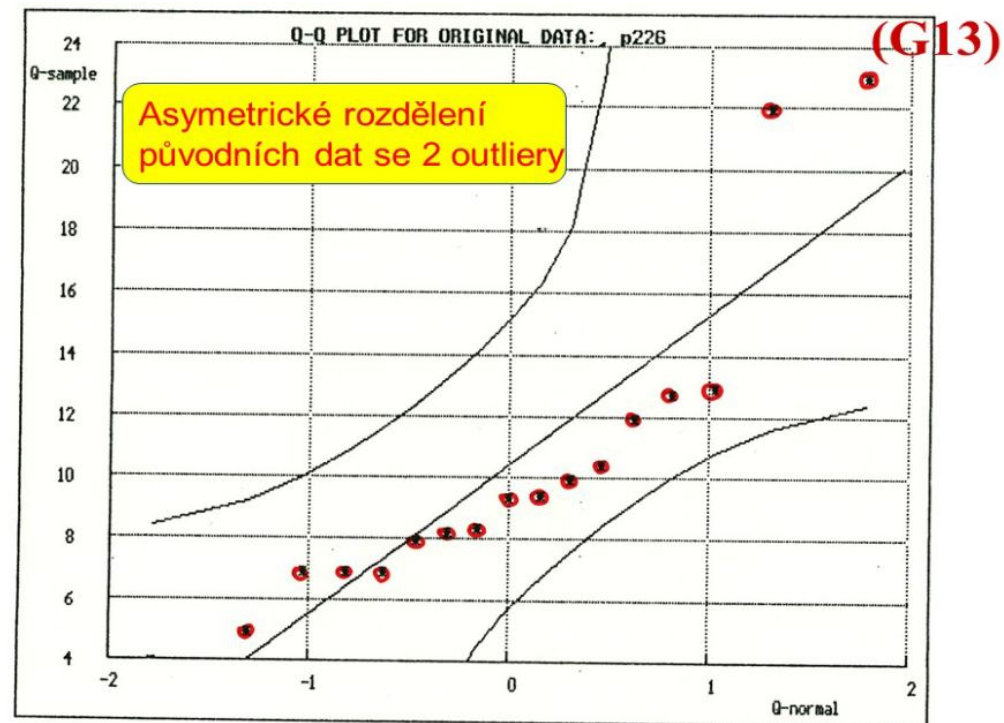
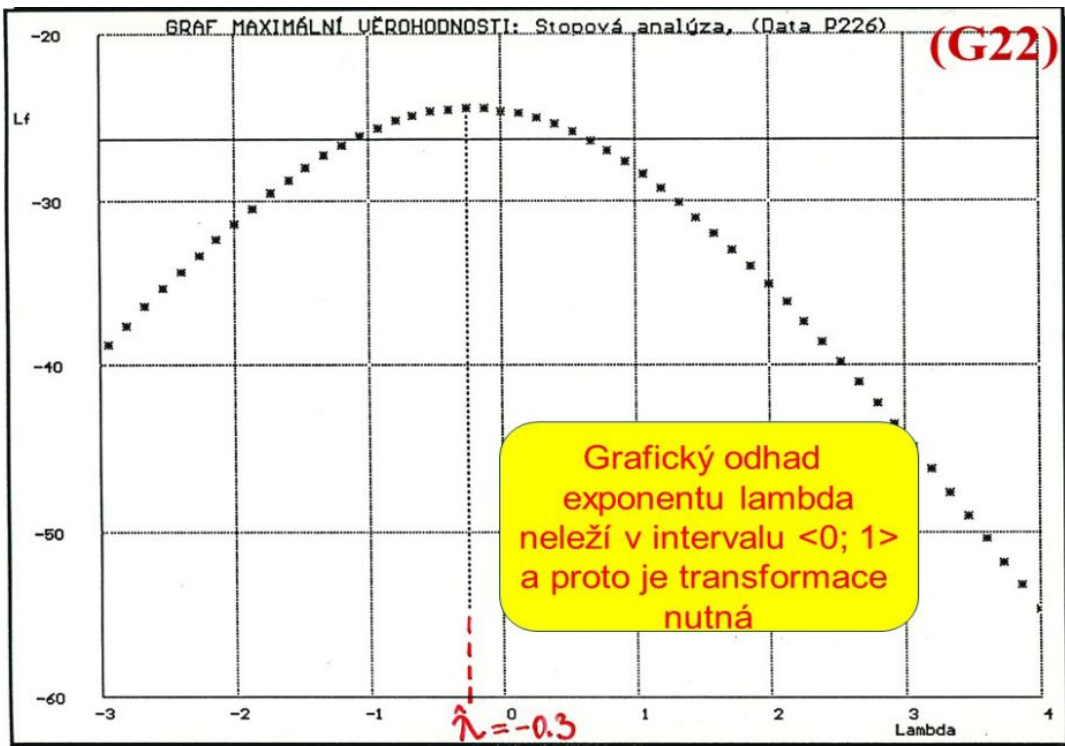


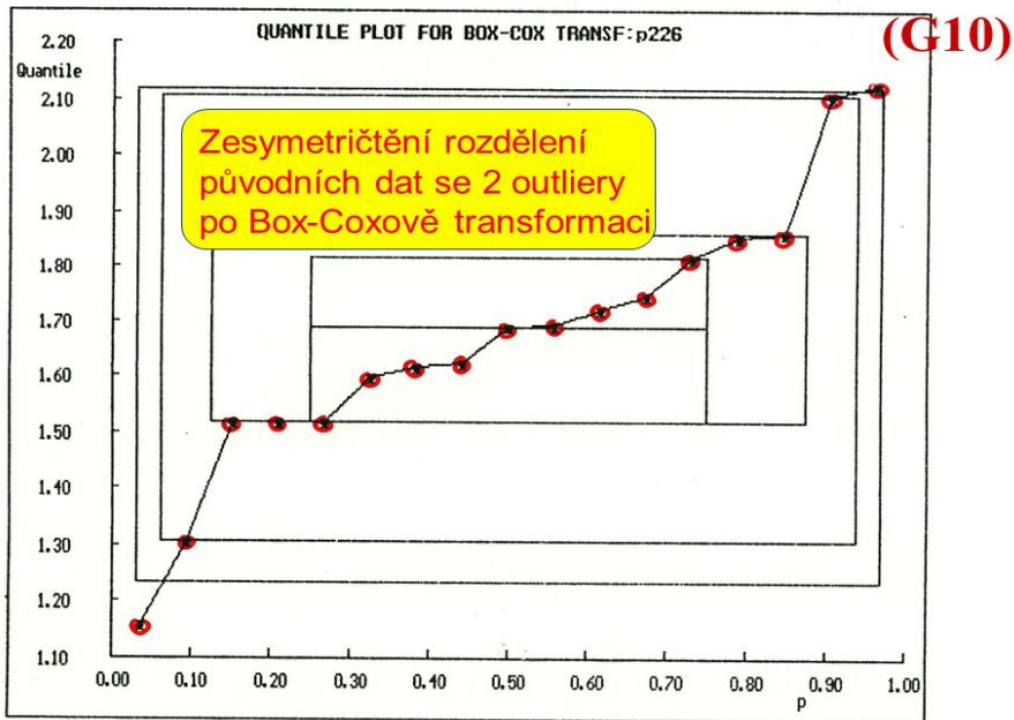
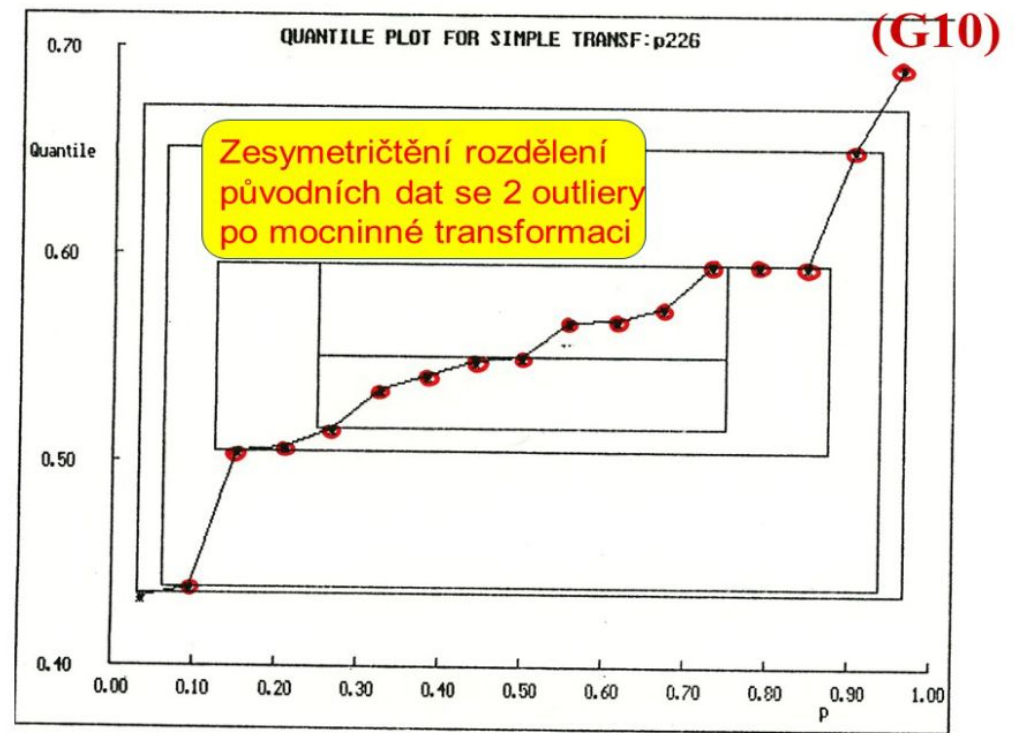
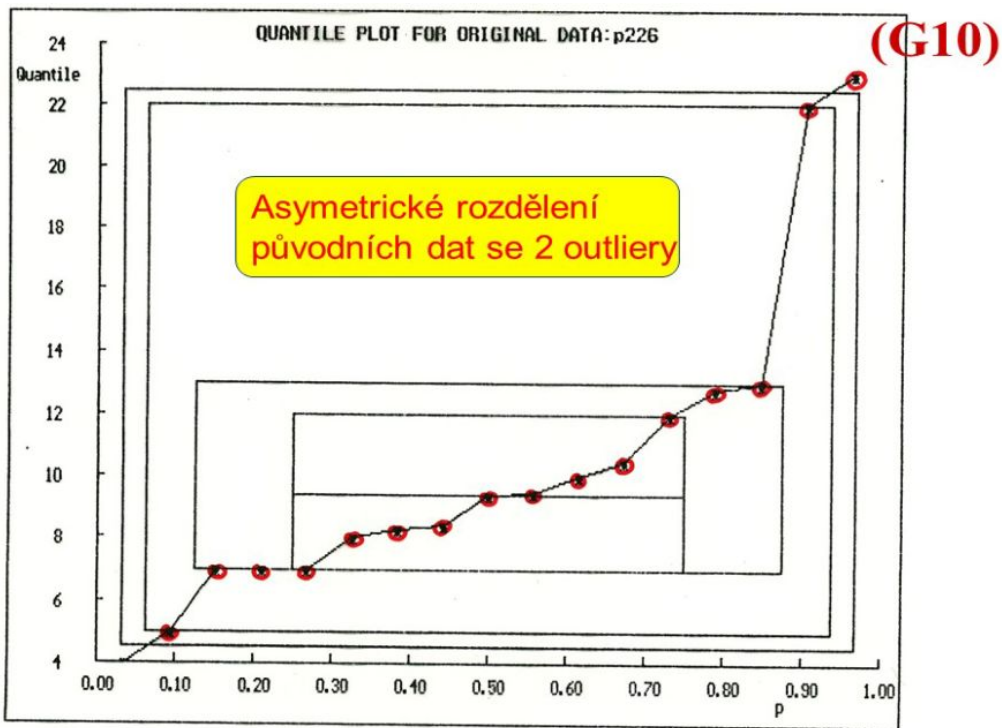
Q-Q GRAF: Stopová analýza, (Data P226)

(G13)

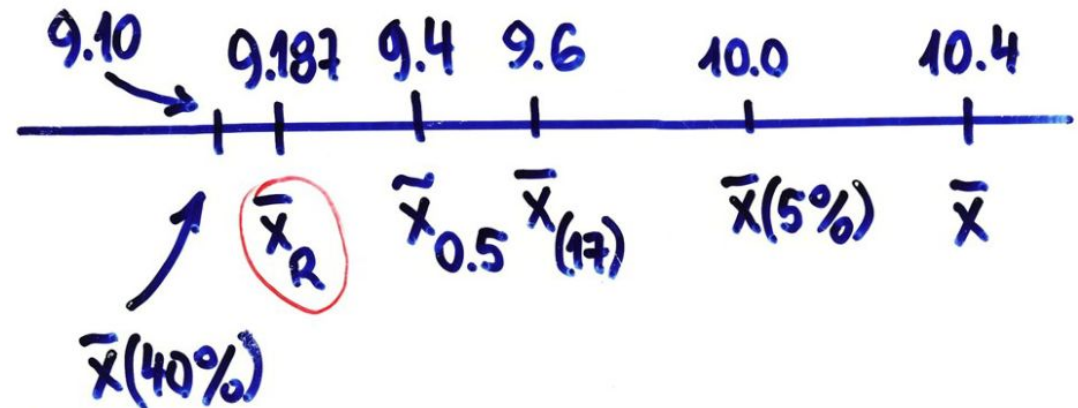








STŘEDNÍ HODNOTA



Přehled různých odhadů parametru polohy, když pravdivý odhad střední hodnoty je 9.19. Aritmetický průměr je zde nejhorší 10.40.

Title: p226

Základní předpoklady o výběru

INPUT

(1) CONDITIONS:

Sample size, n : 17
 Significance level, alpha: 0.050
 Output file name : P226A.TXT

(2) INPUT OF ORIGINAL DATA:

(1) 4.0000E+00 (6) 8.0000E+00 (11) 1.0000E+01 (16) 2.2000E+01
 (2) 7.0000E+00 (7) 8.3000E+00 (12) 1.0500E+01 (17) 2.3000E+01
 (3) 7.0000E+00 (8) 8.4000E+00 (13) 1.2000E+01
 (4) 7.0000E+00 (9) 9.4000E+00 (14) 1.2800E+01
 (5) 5.0000E+00 (10) 9.5000E+00 (15) 1.3000E+01

OUTPUT

(1) CLASSICAL ESTIMATORS OF LOCATION AND SHAPE:

Mean : 1.0406E+01 Variance : 2.6834E+01
 Standard deviation : 5.1802E+00 Skewness : 1.3991E+00
 Kurtosis : 4.2721E+00

(2) TEST FOR NORMALITY:

Quantile, Chi²(1-alpha,2) : 5.9915E+00
 Chi² statistics : 7.9084E+00
 Conclusion: Normality rejected
 Computed significance level : 1.9174E-02

Normalita zamítnuta.

ADSTAT (c) Trilobyte

UNIVARIATE DATA ANALYSIS

One Sample Analysis

Title: p226

Analýza jednoho výběru
 odhady parametrů polohy, rozptýlení
 a tvaru rozdělení

(1) CONDITIONS:

Sample size, n : 17
 Significance level, alpha: 0.050
 Output file name : P226.TXT

(2) INPUT OF ORIGINAL DATA:

(1) 4.0000E+00 (6) 8.0000E+00 (11) 1.0000E+01 (16) 2.2000E+01
 (2) 7.0000E+00 (7) 8.3000E+00 (12) 1.0500E+01 (17) 2.3000E+01
 (3) 7.0000E+00 (8) 8.4000E+00 (13) 1.2000E+01
 (4) 7.0000E+00 (9) 9.4000E+00 (14) 1.2800E+01
 (5) 5.0000E+00 (10) 9.5000E+00 (15) 1.3000E+01

OUTPUT

(1) ESTIMATORS OF DISTRIBUTION SHAPE:

Skewness : 1.3991E+00
 Kurtosis : 4.2721E+00

(3) TEST FOR INDEPENDENCE:

Quantile, t(1-alpha/2,n+1) : 2.1009E+00
 t statistics : 1.0381E+00
 Conclusion: Independence accepted
 Computed significance level : 1.5649E-01

Assumption of sample homogeneity:

Arithmetic mean : 1.0406E+01
 Variance : 2.6834E+01
 Standard deviation : 5.1802E+00

Internal bounds:

Lower bound: -3.1909E+00
 Upper bound: 2.2191E+01

(4) MINIMAL SAMPLE SIZES:

for 25% relative error of standard deviation: n = 14
 for 10% relative error of standard deviation: n = 83
 for 5% relative error of standard deviation: n = 328

(5) OUTLIERS DETECTION:

point number 17 (high) : 2.3000E+01
 Number of outliers : 1

Detekován jediný outlier,
 který ovšem nesmíme vyřadit.

Measures with excluded outliers

Mean : 9.6188E+00 Variance : 1.7388E+01
 Standard deviation : 4.1699E+00 Skewness : 1.6101E+00
 Kurtosis : 6.3399E+00

(2) CLASSICAL ESTIMATORS OF LOCATION AND SPREAD (Normality assumed):

Mean : 1.0406E+01
 Standard deviation : 5.1802E+00
 Variance : 2.6834E+01
 95.0 % confidence interval:
 Lower limit: 7.7425E+00 Upper limit: 1.3069E+01

(3) OTHER ESTIMATORS OF LOCATION:

Mode estimate : 7.5000E+00
 Mid-Sum estimate : 1.3500E+01

Klasické odhady jsou
 u asymetrického rozdělení
nepoužitelné!!!

(4) ROBUST STATISTICS OF LOCATION AND SPREAD (Normality not assumed):

Median measures:

Median : 9.4000E+00
 Standard deviation : 4.4596E+00
 Variance : 1.9888E+01
 Standard deviation : 2.0494E+00
 Variance (nonpar.) : 1.4316E+00
 Standard deviation : 1.0003E+00
 Variance (Marritz) : 1.0001E+00
 95.0 % confidence interval:

Lower limit: 7.2798E+00 Upper limit: 1.1520E+01

Robustní odhady jsou
 u asymetrického rozdělení
správně použitelné!

Trimming 5% (for P=0.05):

Mean : 1.0062E+01
 Standard deviation : 5.5644E+00
 Variance : 3.0963E+01
 Winsorized mean : 1.0406E+01
 Winsorized Standard deviation : 5.1209E+00
 Winsorized variance : 2.6223E+01
 95.0 % confidence interval:
 Lower limit: 7.1450E+00 Upper limit: 1.2979E+01

U uřezaných výběrů
 asymetrického rozdělení
 ztrácíme informaci,
 a proto jsou odhady
 nepoužitelné!!!

Trimming 10% (for P=0.10):

Mean : 9.6324E+00
 Standard deviation : 3.8390E+00
 Variance : 1.4738E+01
 Winsorized mean : 1.0406E+01
 Winsorized Standard deviation : 5.1209E+00
 Winsorized variance : 2.6223E+01
 95.0 % confidence interval:
 Lower limit: 6.7152E+00 Upper limit: 1.2549E+01

Trimming 40% (for P=0.40):

Mean : 9.1059E+00
 Standard deviation : 3.5294E+00
 Variance : 1.2457E+01
 Winsorized mean : 9.1412E+00
 Winsorized Standard deviation : 1.4739E+00
 Winsorized variance : 2.1722E+00
 95.0 % confidence interval:
 Lower limit: 7.2116E+00 Upper limit: 1.1000E+01

Biweight:

Mean : 8.8034E+00
 Standard deviation : 2.5962E+00
 Variance : 6.7403E+00
 Weight sqrt(w) : 3.7120E+00
 95.0 % confidence interval:
 Lower limit: 7.3207E+00 Upper limit: 1.0286E+01

(5) ADAPTIVE ESTIMATORS OF LOCATION AND SPREAD:

Hoog's estimates:

Relative tail length: 2.7797E+00
 Mean : 9.3353E+00
 Standard deviation : 3.7223E+00
 Variance : 1.3855E+01
 95.0 % confidence interval:
 Lower limit: 7.4215E+00 Upper limit: 1.1249E+01

Adaptivní odhady jsou
 u asymetrického rozdělení
 použitelné.

Úloha B2.16 Rozdělení výběru sedimentace červených krvinek

Z naměřených hodnot sedimentace červených krvinek u 40 lidí určete typ diskrétního rozdělení výběru a vhodný odhad střední hodnoty a míry variability. Jsou v datech vysoké odlehle hodnoty, indikující nemocného pacienta? O jaké rozdělení se jedná?

Data: Hodnoty sedimentace krve [mm/h]

4	5	10	6	6	7	5	28	5	7
11	9	6	5	9	9	5	7	8	5
11	6	8	6	7	12	9	6	27	8
12	7	10	9	6	5	9	9	7	7
13	8	9							

Box-Coxova transformace dat :

Název úlohy : B216
 Optimální parametr : -17,48291359
 Dolní mez parametru : -18,41256185
 Horní mez parametru : -13,23668249
 Věrohodnost bez transformace : 64,3775165
 Věrohodnost s transformací : 87,30301074
 Oprávněnost transformace : Ano
 Pravděpodobnost : 99,999999987257 %
 Zvolený parametr : -17,48291359
Oprávněný průměr : 7,354823842

Exponenciální transformace dat :

Název úlohy : B216
 Optimální parametr : 1,610244751
 Zvolený parametr : 1,610244751
 Oprávněnost transformace : Ano
Oprávněný průměr : 7,011200856
 Interval spolehlivosti :
Spodní : 6,387854325
Horní : 7,798715014

ADSTAT Box-Coxova transformace:

Oprávněný průměr: 7.0513

ADSTAT Mocninná transformace:

Oprávněný průměr: 7.0513

Klasické parametry :

Sloupce : B216
 Průměr : 8,3
 Spodní mez : 6,753108461
 Horní mez : 9,846891539
 Rozptyl : 23,39487179
 Směr. odchylka : 4,836824557
 Dolní mez : 3,975847146
 Horní mez : 6,181928518
 Robustní směr.odch. : 2,9652
 Šikmost : 3,175083132
 Odchylka od 0 : Významná
 Špičatost : 13,19172237
 Odchylka od 3 : Významná

Test normality: Momentový

Sloupce : B216
 Normalita : Zamítnuta
 Testové kritérium : 17,9334316
 Kritický kvantil chi2(22) : 5,991464547
 p-hodnota : 0,0001275865248

Vybočující body :

Sloupce : B216
 Homogenita : Zamítnuta
 Počet vybočujících bodů: 2
 Spodní mez : -0,48
 Horní mez : 15,48

Robustní parametry :

Sloupce : B216
 Medián : 7
 IS spodní : 3,387985583
 IS horní : 10,61201442
 Medianová směr. odchylka : 1,785747086
 Medianový rozptyl : 3,188892657
 10% Průměr : 7,444444444
 10% IS spodní : 6,694598938
 10% IS horní : 8,194289951
 10% Směr. odchylka : 1,74891908
 10% Rozptyl : 3,058717949
 20% Průměr : 7,34375
 20% IS spodní : 6,624492632
 20% IS horní : 8,063007368
 20% Směr. odchylka : 1,370476241
 20% Rozptyl : 1,878205128
 40% Průměr : 7,375
 40% IS spodní : 6,664146113
 40% IS horní : 8,085853887
 40% Směr. odchylka : 0,9013522605
 40% Rozptyl : 0,8124358974

Úloha B2.17 Rozdělení výběru koncentrace fibrinogenu v krevní plazmě

Nalezněte typ rozdělení koncentrace fibrinogenu v krevní plazmě u 50 pacientů a určete vysoké odlehle hodnoty nemocných pacientů, a to především těch, jež přesahují normu 3 až 5 g.l⁻¹. Vyčíslete k této průzkumové analýze kvantilové charakteristiky šikmosti a špičatosti: polosumu Z_L , rozpětí R_L , šikmosti S_L , pseudosigmu G_L a délky konců T_L pro kvartily a oktily a ukažte, jak charakterizují symetrii (tj. Z_L a S_L), rozptylení (tj. R_L) a špičatost (tj. G_L a T_L). Vyšetřete tvar rozdělení na základě grafu polosum, symetrie a špičatosti (symboly viz [19]). Které diagnostiky shodně indikují vybočující hodnoty?

Data: Úroveň fibrinogenu v krevní plazmě [g.l⁻¹]:

6.68	3.24	4.16	3.21	3.64	3.61	2.22	3.88	4.08	3.07
..
4.99	3.76	4.05	4.15	3.28	3.76	4.11	3.72	3.57	3.43

Box-Coxova transformace dat :

Název úlohy :	B217
Optimální parametr :	-7,857436752
Dolní mez parametru :	-12,85743675
Horní mez parametru :	-2,857436808
Věrohodnost bez transformace :	80,47189562
Věrohodnost s transformací :	86,29572015
Oprávněnost transformace :	Ano
Pravděpodobnost :	99,9357135055599 %
Zvolený parametr :	-7,857436752
Opravený průměr :	4,000741732

ADSTAT Box-Coxova transformace:

Opravený průměr: 4.0041

Exponenciální transformace dat :

Název úlohy :	B217
Optimální parametr :	0,3938140869
Zvolený parametr :	0,3938140869
Oprávněnost transformace :	Ano
Opravený průměr :	3,999281789
Interval spolehlivosti :	
Spodní :	3,759437754
Horní :	4,263636551

ADSTAT Mocnninná transformace:

Opravený průměr: 4.0041

Úloha C2.17 Podezřelá a odlehle hodnoty efektivního průměru pigmentů

Analýzou vzorku Colanyschwarz PR 100 na přístroji Particle Sizer BI-90 byly stanoveny hodnoty efektivního průměru dispergovaných částic pigmentu ve vodném prostředí. Úkolem je testovat normalitu, nezávislost, homogenitu rozdělení výběru a vyšetřit, zda jsou ve výběru přítomny odlehle hodnoty. Je vhodné odlehle hodnoty vypustit? Je rozdělení výběru symetrické? Kolik procent hodnot výběru leží nad 190 μm a kolik pod touto hodnotou?

Data: Efektivní průměr dispergované částice [μm]:

195	181	187	196	175	186
189	176	182	177	207	189
181	188	190	184	180	183
171	192	170	178	194	158
177	186	183	162	161	167
				137	

Box-Coxova transformace dat :

Název úlohy :	C217
Optimální parametr :	6,459022903
Dolní mez parametru :	1,459022903
Horní mez parametru :	11,45902285
Věrohodnost bez transformace :	64,3775165
Věrohodnost s transformací :	66,2374769
Oprávněnost transformace :	Ne
Pravděpodobnost :	94,6230756994289 %
Zvolený parametr :	6,459022903
Věrohodnost :	66,2374769
Opravený průměr :	183,1963019

Exponenciální transformace dat :

Název úlohy :	C217
Optimální parametr :	-0,230638504
Zvolený parametr :	-0,230638504
Oprávněnost transformace :	Ano
Opravený průměr :	183,1810576
Interval spolehlivosti :	
Spodní :	178,9922119
Horní :	187,0880463

ADSTAT Mocnninná transformace:

Opravený průměr: Nelze, soubor je delší než 99 prvků.

Klasické parametry :

Sloupce :	B217
Průměr :	4,1526
Spodní mez :	3,863218426
Horní mez :	4,441981574
Rozptyl :	1,036819633
Směr. odchylka :	1,018243405
Dolní mez :	0,853233979
Horní mez :	1,263725696
Robustní směr.odch.	0,704235
Detrendovaná směr.odch.(MR)	0,8618794326
Šikmost	1,470479424
Odchylka od 0 :	Významná
Špičatost :	6,818336949
Odchylka od 3 :	Významná

Test normality

Sloupce :	B217
Normalita :	Zamítnuta
Testové kritérium :	10,10458673
Kritický kvantil $\chi^2(22)$:	5,991464547
p-hodnota :	0,006394651332

Vybočující body :

Sloupce :	B217
Homogenita :	Zamítnuta
Počet vybočujících bodů :	1
Spodní mez :	1,18954
Horní mez :	6,92046

42

Klasické parametry :

Sloupce :	C217
Průměr :	181,85
Spodní mez :	177,5815618
Horní mez :	186,1184382
Rozptyl :	178,1307692
Směr. odchylka :	13,34656395
Dolní mez :	10,9708131
Horní mez :	17,05819662
Robustní směr.odch.	8,8956
Detrendovaná směr.odch.(MR)	10,7712766
Šikmost	-0,8730589259
Odchylka od 0 :	Významná
Špičatost :	4,867334943
Odchylka od 3 :	Významná

Test normality

Sloupce :	C217
Normalita :	Přijata
Testové kritérium :	4,967633653
Kritický kvantil $\chi^2(22)$:	5,991464547
p-hodnota :	0,08342420146

Vybočující body :

Sloupce :	C217
Homogenita :	Zamítnuta
Počet vybočujících bodů :	1
Spodní mez :	147,92
Horní mez :	217,08

44

Robustní parametry :

Sloupce :	B217
Medián :	3,945
IS spodní :	3,006839226
IS horní :	4,883160774
Medianová směr. odchylka :	0,4668453097
Medianový rozptyl :	0,2179445432
10% Průměr :	4,078409091
10% IS spodní :	3,841040691
10% IS horní :	4,315777491
10% Směr. odchylka :	0,6104265782
10% Rozptyl :	0,3726206073
20% Průměr :	4,06225
20% IS spodní :	3,8132099
20% IS horní :	4,3112901
20% Směr. odchylka :	0,5216640095
20% Rozptyl :	0,2721333388
40% Průměr :	4,008666667
40% IS spodní :	3,736714604
40% IS horní :	4,280618729
40% Směr. odchylka :	0,3152720626
40% Rozptyl :	0,09939647347

Robustní parametry :

Sloupce :	C217
Medián :	183
IS spodní :	176,2919732
IS horní :	189,7080268
Medianová směr. odchylka :	3,316387446
Medianový rozptyl :	10,99842569
10% Průměr :	182,3611111
10% IS spodní :	178,5006109
10% IS horní :	186,2216114
10% Směr. odchylka :	8,907375941
10% Rozptyl :	79,34134615
20% Průměr :	182,71875
20% IS spodní :	178,9174968
20% IS horní :	186,5200032
20% Směr. odchylka :	6,710821724
20% Rozptyl :	45,03512821
40% Průměr :	183,0416667
40% IS spodní :	179,8342918
40% IS horní :	186,2490415
40% Směr. odchylka :	3,501226891
40% Rozptyl :	12,25858974

Úloha E2.27 Odhad střední hodnoty stopové kontaminace půdy

Koncentrační pozadí detekované chemikálie v zemině bylo stanoveno na 10 vzorcích půdy u prokazatelně nekontaminovaného zemědělského pole. Zemina ze sousední oblasti byla pro koncentraci vyšší než horní mez 99% intervalu spolehlivosti nekontaminovaného pozadí, detekovaná proto jako znečištěná. Z nalezených dat vyčíslete bodový a intervalový odhad střední stopové koncentrace chemikálie ještě znečišťující půdu?

Data: Koncentrace pozadí c [mg/kg]: 1.4 0.6 1.2 1.6 0.5 0.7 0.3
0.8 0.2 0.9

Box-Coxova transformace dat :

Název úlohy :	E227
Optimální parametr :	-4,74437294
Dolní mez parametru :	-9,74437294
Horní mez parametru :	0,2556270039
Věrohodnost bez transformace :	16,09437912
Věrohodnost s transformací :	16,29669832
Oprávněnost transformace :	Ne
Pravděpodobnost :	47,5292590831894 %
Zvolený parametr :	-4,74437294
Věrohodnost :	16,29669832
Opravený průměr :	0,7737121682

Exponenciální transformace dat :

Název úlohy :	E227
Optimální parametr :	0,2982025146
Zvolený parametr :	0,2982025146
Oprávněnost transformace :	Ne
Opravený průměr :	0,7634940468
Interval spolehlivosti :	
Spodní :	0,4757731305
Horní :	1,115491942

ADSTAT Mocninná transformace:

Opravený průměr: 0,76125

Klasické parametry :

Sloupce :	E227
Průměr :	0,82
Spodní mez :	0,4899351539
Horní mez :	1,150064846
Rozptyl :	0,2128888889
Směr. odchylka :	0,4613988393
Dolní mez :	0,317366366
Horní mez :	0,8423344192
Robustní směr.odch.	0,51891
Detrendovaná směr.odch.(MR)	0,469858156
Šikmost	0,358852124
Odchylka od 0 :	Nevýznamná
Špičatost :	2,010656334
Odchylka od 3 :	Nevýznamná

Test normality:

Sloupce :	E227
Normalita :	Přijata
Testové kritérium :	0,5471251031
Kritický kvantil $\chi^2(22)$:	5,991464547
p-hodnota :	0,7606647541

Vybočující body :

Sloupce :	E227
Homogenita :	Přijata
Počet vybočujících bodů :	0
Spodní mez :	-0,834
Horní mez :	2,034

Robustní parametry :

Sloupce :	E227
Medián :	0,75
IS spodní :	-0,05792811249
IS horní :	1,557928112
Medianová směr. odchylka :	0,3571494173
Medianový rozptyl :	0,1275557063
10% Průměr :	0,8
10% IS spodní :	0,4092743983
10% IS horní :	1,190725602
10% Směr. odchylka :	0,319861081
10% Rozptyl :	0,1023111111
20% Průměr :	0,8
20% IS spodní :	0,4092743983
20% IS horní :	1,190725602
20% Směr. odchylka :	0,319861081
20% Rozptyl :	0,1023111111
40% Průměr :	0,7833333333
40% IS spodní :	0,3612044511
40% IS horní :	1,205462216
40% Směr. odchylka :	0,18636881
40% Rozptyl :	0,03473333333

Analýza malých výběrů

N :	10
Střední hodnota :	0,85
Spodní mez (5%) :	0,4839
Horní mez (95%) :	1,2161
Spodní mez (2.5%) :	0,3824
Horní mez (97.5%) :	1,3176
Pivotové rozpětí :	0,7