Pravděpodobnostní metody ve strojírenství

12. Testy o kvalitativních výsledcích

FAKULTA STROJNÍ ČVUT V PRAZE

prof. RNDr. Gejza Dohnal, CSc.

ak. rok 2023/2024

12. Testy o kvalitativních výsledcích

Klíčové pojmy:

- poměry (proporce)
- Kontingenční tabulka, čtyřpolní tabulka

Klíčové vztahy:

- Test chí-kvadrát
- Fisherův test

Test poměrů

Příklady:

- Statistický odhad pravděpodobnosti
- Odhad volebních preferencí
- Poměr neshodných výrobků
- Poměr žen v určité funkci
- … atd.

odhadujeme pomocí relativní četnosti $\hat{p} = \frac{n_A}{N}$

Jednovýběrový test:

$$H_0 : \quad p = p_0$$
$$H_A : \quad p \neq p_0$$

$$T = \frac{\hat{p} - Np_0}{\sqrt{Np_0(1-p_0)}} \sim \mathrm{N}(0,1)$$

```
> prop.test(36, 100, 0.3)

	1-sample proportions test with continuity correction

data:  36 out of 100, null probability 0.3
X-squared = 1.4405, df = 1, p-value = 0.2301
alternative hypothesis: true p is not equal to 0.3
95 percent confidence interval:
 0.2681721 0.4627255
sample estimates:        p   0.36
```

Test poměrů

Dvouvýběrový test:

$$H_0: \quad p_1 = p_2 = p$$
$$H_A: \quad p_1 \neq p_2$$

$$T = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{p_0(1-p_0)(1/N_1 + 1/N_2)}} \sim \mathrm{N}(0,1)$$

Příklad:

Bylo dotazováno 800 absolventů ČVUT a UK. Z výzkumu plyne, že 322 dotazovaných jsou absolventi ČVUT kteří pracují dále ve svém oboru, 234 dotázaných absovovalo UK a pracují také ve svém oboru. Ve 158 případech odpovědí bylo zjištěno, že absolventi ČVUT ve svém oboru nepracují a zbývajících 86 jsou absolventi UK, kteří také nepracují ve svém oboru.

Liší se statisticky významně poměr absolventů pracujících v oboru pro tyto dvě univerzity?

$$\hat{p}_1 = 322/480 = 0,67, \quad \hat{p}_2 = 234/320 = 0,73$$

	ČVUT	UK	celkem
v oboru	322	234	556
mimo obor	158	86	244
celkem	480	320	800

Test poměrů

Dvouvýběrový test:

$$H_0: \quad p_1 = p_2 = p$$
$$H_A: \quad p_1 \neq p_2$$

$$T = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{p_0(1-p_0)(1/N_1 + 1/N_2)}} \sim \mathrm{N}(0,1)$$

Příklad:

```
> prop.test(c(322,234),c(480,320))

   2-sample test for equality of proportions with continuity correction

data:  c(322, 234) out of c(480, 320)
X-squared = 3.0273, df = 1, p-value = 0.08187
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.127257708  0.006424375
sample estimates:
   prop 1    prop 2
0.6708333 0.7312500
```

	ČVUT	UK	celkem
v oboru	322	234	556
mimo obor	158	86	244
celkem	480	320	800

Test nezávislosti v kontingenční tabulce

Příklad:

Bylo dotazováno 800 absolventů ČVUT a UK. Z výzkumu plyne, že 322 dotazovaných jsou absolventi ČVUT kteří pracují dále ve svém oboru, 234 dotázaných absovovalo UK a pracují také ve svém oboru. Ve 158 případech odpovědí bylo zjištěno, že absolventi ČVUT ve svém oboru nepracují a zbývajících 86 jsou absolventi UK, kteří také nepracují ve svém oboru.

Lze považovat uplatnění se v oboru studia stochasticky závislé na absolvované univerzitě?

	ČVUT	UK	celkem
v oboru	0,4025	0,2925	0,695
mimo obor	0,1975	0,1075	0,305
celkem	0,6000	0,4000	1,000

0,695x0,6 = 0,417 ≠ 0,4025

Test nezávislosti v kontingenční tabulce

Příklad:

Bylo dotazováno 800 absolventů ČVUT a UK. Z výzkumu plyne, že 322 dotazovaných jsou absolventi ČVUT kteří pracují dále ve svém oboru, 234 dotázaných absovovalo UK a pracují také ve svém oboru. Ve 158 případech odpovědí bylo zjištěno, že absolventi ČVUT ve svém oboru nepracují a zbývajících 86 jsou absolventi UK, kteří také nepracují ve svém oboru.

Lze považovat uplatnění se v oboru studia stochasticky závislé na absolvované univerzitě?

Test chí-kvadrát:

$$H_0 : n_{ij} = \frac{n_{i.} n_{.j}}{n}$$ pro všechny dvojice i,j

$$H_A : n_{ij} \neq \frac{n_{i.} n_{.j}}{n}$$ pro nějakou dvojici i,j

kde $n_{i.} = \sum_{j=1}^{n} n_{ij}, \; n_{.j} = \sum_{i=1}^{n} n_{ij}.$

	ČVUT	UK	celkem
v oboru	0,4025	0,2925	0,695
mimo obor	0,1975	0,1075	0,305
celkem	0,6000	0,4000	1,000

0,695x0,6 = 0,417 ≠ 0,4025

Test nezávislosti v kontingenční tabulce

Příklad:

Bylo dotazováno 800 absolventů ČVUT a UK. Z výzkumu plyne, že 322 dotazovaných jsou absolventi ČVUT kteří pracují dále ve svém oboru, 234 dotázaných absovovalo UK a pracují také ve svém oboru. Ve 158 případech odpovědí bylo zjištěno, že absolventi ČVUT ve svém oboru nepracují a zbývajících 86 jsou absolventi UK, kteří také nepracují ve svém oboru.

Lze považovat uplatnění se v oboru studia stochasticky závislé na absolvované univerzitě?

	ČVUT	UK	celkem
v oboru	n_{11}	n_{12}	$n_{1.}$
mimo obor	n_{21}	n_{22}	$n_{2.}$
celkem	$n_{.1}$	$n_{.2}$	n

Test chí-kvadrát:

$$H_0 : n_{ij} = \frac{n_{i.}n_{.j}}{n}$$ pro všechny dvojice i,j

$$H_A : n_{ij} \neq \frac{n_{i.}n_{.j}}{n}$$ pro nějakou dvojici i,j

kde $n_{i.} = \sum_{j=1}^{n} n_{ij}, \; n_{.j} = \sum_{i=1}^{n} n_{ij}.$

Test nezávislosti v kontingenční tabulce

Test chí-kvadrát:

Označme $n_{ij}^* = \dfrac{n_{i.}n_{.j}}{n}$ tzv. teoretické (očekávané) četnosti. Potom testová statistika

bude mít tvar: $$X^2 = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{ij}-n_{ij}^*)^2}{n_{ij}^*} \sim \chi^2(r-1)(s-1)$$

Hypotézu H_0 o nezávislosti znaků zamítáme na hladině významnosti α, pokud hodnota testové statistiky X^2 překročí $(1-\alpha)$-kvantil rozdělení chí-kvadrát o $(r-1)(s-1)$ stupních volnosti.

Podmínky pro použití testu nezávislosti v kontingenční tabulce:

- žádná teoretická četnost nesmí být menší než 1
- nejvíce 20 % teoretických četností může být menších než 5
- pro tabulku 2x2 (čtyřpolní) musí navíc platit, že:
    - n > 40, jinak
    - pro 20 < n < 40, je nutná úprava test. kriteria pomocí tzv. Yatesovy korekce
    - pro n<20 je třeba použít tzv. přesný Fisherův test.

Test nezávislosti v kontingenční tabulce

Příklad:

Bylo dotazováno 800 absolventů ČVUT a UK. Z výzkumu plyne, že 322 dotazovaných jsou absolventi ČVUT kteří pracují dále ve svém oboru, 234 dotázaných absovovalo UK a pracují také ve svém oboru. Ve 158 případech odpovědí bylo zjištěno, že absolventi ČVUT ve svém oboru nepracují a zbývajících 86 jsou absolventi UK, kteří také nepracují ve svém oboru.

Lze považovat uplatnění se v oboru studia stochasticky závislé na absolvované univerzitě?

	ČVUT	UK	celkem
v oboru	322	234	556
mimo obor	158	86	244
celkem	480	320	800

```
> pruzkum <- matrix(c(322, 234, 158, 86),2)
> colnames(pruzkum)= c("CVUT","UK")
> rownames(pruzkum)= c("v oboru", "mimo obor")
> pruzkum
          CVUT  UK
v oboru    322 158
mimo obor  234  86

> chisq.test(pruzkum)
  Pearson's Chi-squared test with Yates' continuity correction

 data: pruzkum, X-squared = 3.0273, df = 1, p-value = 0.08187
```

Test nezávislosti v kontingenční tabulce

Fisherův test je založen na tzv. poměru šancí (Odds ratio) výskytu určité události v závislosti na události druhé (v našem příkladu „CVUT“ a „v oboru“). Kvantifikuje tedy sílu vztahu mezi těmito dvěma veličinami.

$$\mathrm{OR} = \frac{P(\text{jev nastane})}{P(\text{jev nenastane})} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

OR = 1 : žádná závislost mezi zkoumanými událostmi – pravděpodobnost nastání jedné události je stejná nezávisle na přítomnosti či nepřítomnosti druhé
OR > 1 : přítomnost jedné události navyšuje šanci na nastání události druhé
OR < 1 : přítomnost jedné události snižuje šanci na nastání události druhé

	ČVUT ano	ČVUT ne	celkem
v oboru ano	322	234	556
v oboru ne	158	86	244
celkem	480	320	800

```
> fisher.test(pruzkum)

    Fisher's Exact Test for Count Data

data:  pruzkum,  p-value = 0.07187
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval: 0.5407144 1.0343715
sample estimates: odds ratio 0.7492708
```

Pravděpodobnostní metody ve strojírenství

13. Síla testu a rozsah výběru

FAKULTA STROJNÍ ČVUT V PRAZE

prof. RNDr. Gejza Dohnal, CSc.

ak. rok 2023/2024

13. Síla testu a rozsah výběru

Klíčové pojmy:

- Síla testu
- Rozsah výběru

Klíčové vztahy:

- Test chí-kvadrát
- Fisherův test

Síla testu

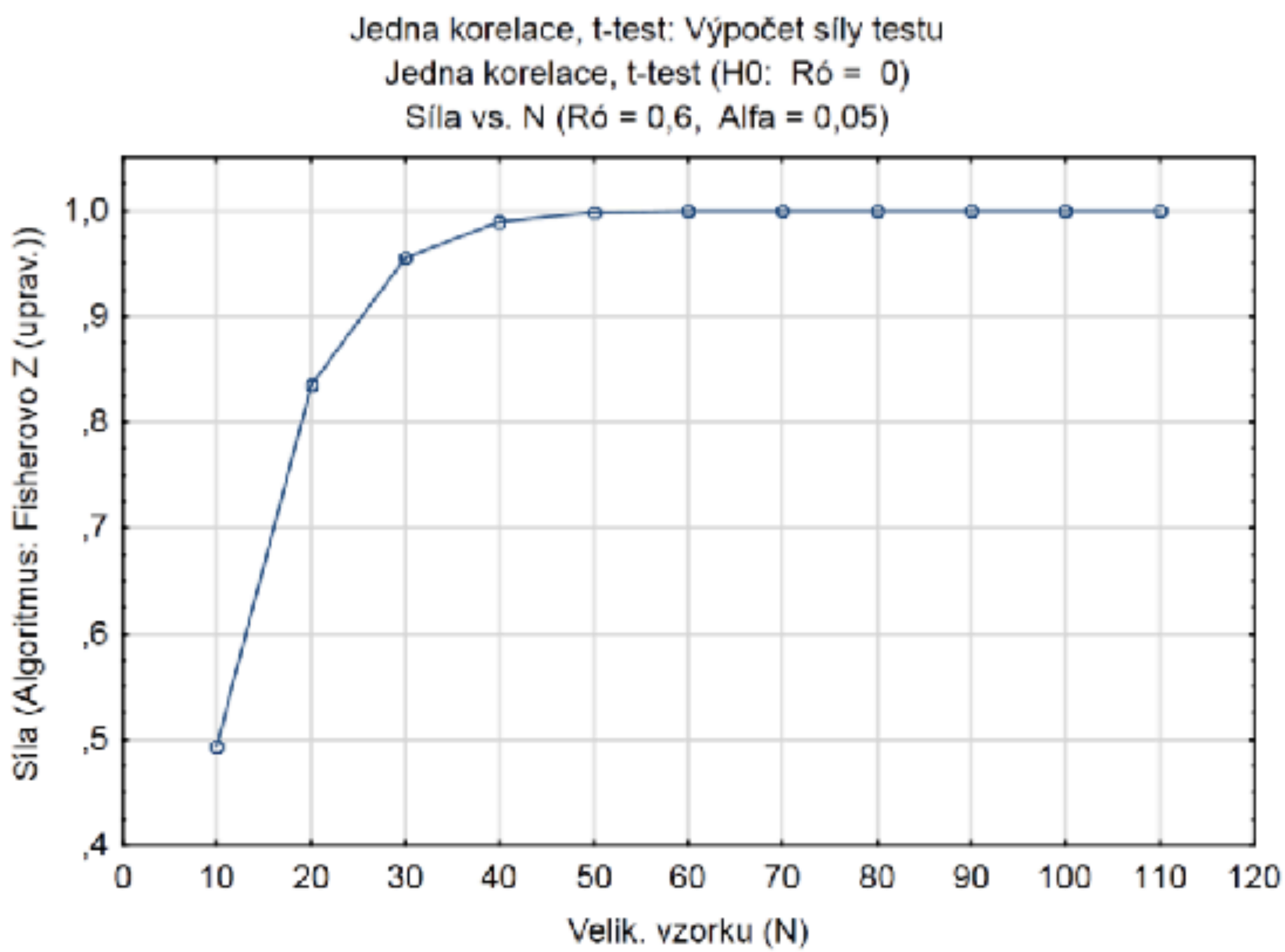
Síla testu je pravděpodobnost toho, že test zamítne hypotézu když neplatí, tedy je rovna (1- β), kde β je pravděpodobnost chyby II. druhu (= nezamítáme nulovou hypotézu když nepatí)

Síla testu je vlastnost testu, tedy závisí především na:

- testové statistice
- alternativní hypotéze
- na velikosti výběru
- no a samozřejmě na datech

V případě t-testu závisí na

- rozptylu dat σ^2
- předpokládané diferenci δ
- na velikosti výběru n

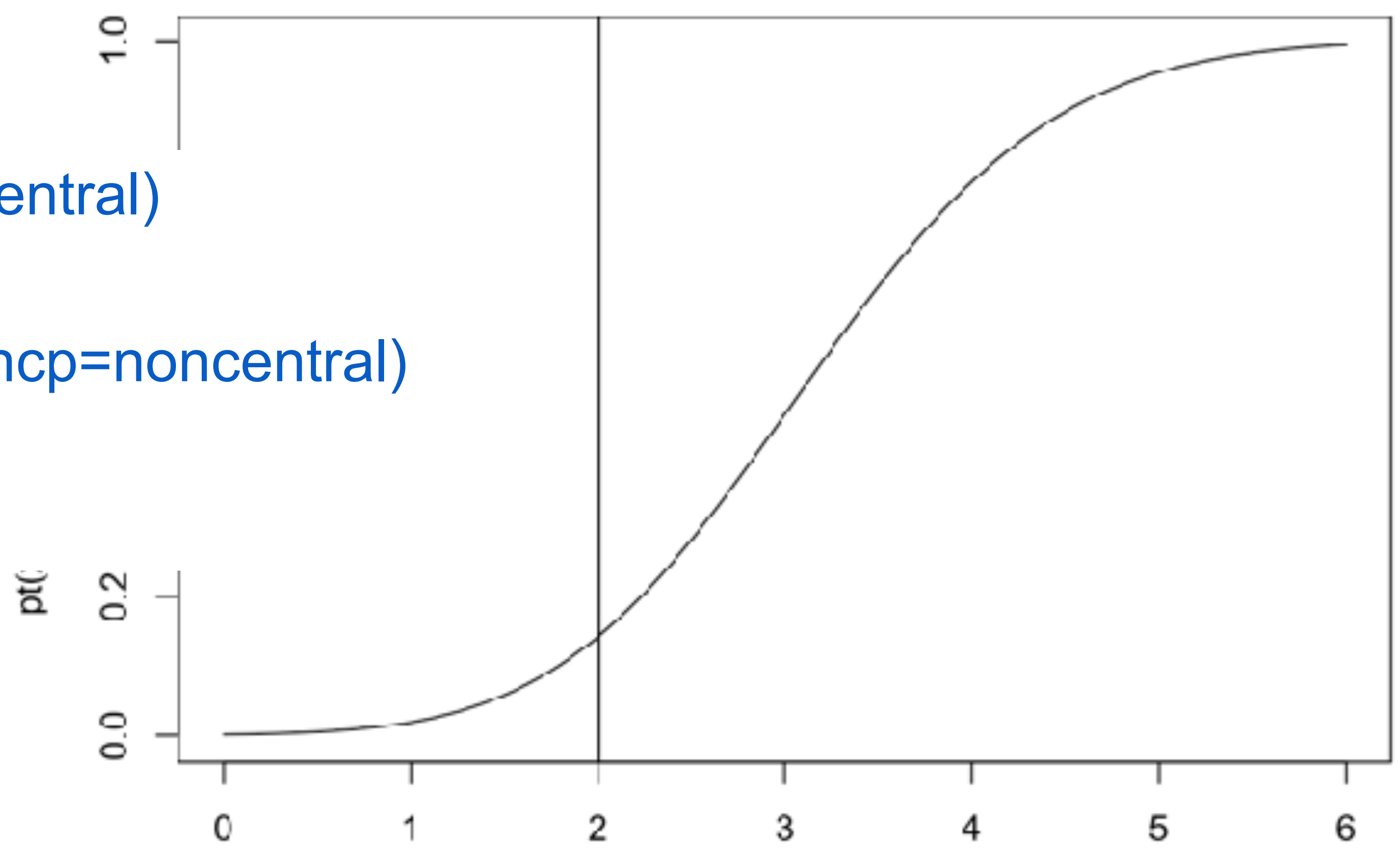


Síla testu

Jednovýběrový t-test (nebo párový t-test): Za předpokladu alternativní hypotézy H_A: $\mu = m_0 + \delta$ má testová statistika tzv. necentrální t-rozdělení s parametrem necentrality $\nu = \frac{\delta}{\sigma}\sqrt{n}$

```
> davky <- data.matrix(read.table("davkovac.txt"))
> n = 50;  delta = 0.2
> noncentral = delta*sqrt(n)/sd(davky)
> curve(pt(x,n-1,ncp=noncentral), from=0, to=10)
> abline(v=qt(0.975, n-1))

> pt(qt(0.975,n-1), n-1,ncp=noncentral)
[1] 0.1438221
> power=1-pt(qt(0.975,n-1), n-1,ncp=noncentral)
> power
[1] 0.8561779
```



Síla testu

Jednovýběrový t-test (nebo párový t-test):

```
> power.t.test(delta=0.2, sd=sd(davky), sig.level=0.05, power=0.9,
    type="one.sample")

     One-sample t test power calculation

              n = 57.18447
          delta = 0.2
             sd = 0.458528
      sig.level = 0.05
          power = 0.9
    alternative = two.sided
```

Při ručním počítání obvykle předpokládáme znalost směrodatné odchylky σ a namísto t-rozdělení použijeme standardní normální rozdělení (a jeho kvantily u_p). Požadovaná síla testu je zde označena β. Potom lze odvodit přibližný vztah:

$$n = \left[\left(\frac{u_{\alpha/2} + u_\beta}{\delta/\sigma} \right)^2 \right] + 1$$

Síla testu

Dvouvýběrový t-test: Za předpokladu alternativní hypotézy H$_A$: $\mu_X - \mu_Y = \delta$ má testová statistika necentrální t-rozdělení s parametrem necentrality

$$\nu = \frac{\delta}{\sigma\sqrt{1/n_X + 1/n_Y}}$$

```
> x <- data.matrix(read.table("dodavatelX.txt"))
> y <- data.matrix(read.table(„dodavatelY.txt"))
> z=c(x,y)

> power.t.test(delta=2, sd=sd(z), sig.level=0.05, power=0.9)

     Two-sample t test power calculation

              n = 43.59226
          delta = 2
             sd = 2.847857
      sig.level = 0.05
          power = 0.9
    alternative = two.sided

NOTE: n is number in *each* group
```

Síla testu

Dvouvýběrový t-test: Za předpokladu alternativní hypotézy H_A: $\mu_X - \mu_Y = \delta$ má testová statistika necentrální t-rozdělení s parametrem necentrality

$$\nu = \frac{\delta}{\sigma\sqrt{1/n_X + 1/n_Y}}$$

```
> x <- data.matrix(read.table("dodavatelX.txt"))
> y <- data.matrix(read.table(„dodavatelY.txt"))
> z=c(x,y)

> power.t.test(n=100, delta=2, sd=sd(z), sig.level=0.05)

     Two-sample t test power calculation

              n = 100
          delta = 2
             sd = 2.847857
      sig.level = 0.05
          power = 0.9985668
    alternative = two.sided

NOTE: n is number in *each* group
```

Síla testu

Dvouvýběrový t-test: Za předpokladu alternativní hypotézy H_A: $\mu_X - \mu_Y = \delta$ má testová statistika necentrální t-rozdělení s parametrem necentrality

$$\nu = \frac{\delta}{\sigma\sqrt{1/n_X + 1/n_Y}}$$

```
> x <- data.matrix(read.table("dodavatelX.txt"))
> y <- data.matrix(read.table(„dodavatelY.txt"))
> z=c(x,y)

> power.t.test(delta=2, sd=sd(z), sig.level=0.05, power=0.9,
  alt="one.sided")

     Two-sample t test power calculation

              n = 35.42511
          delta = 2
             sd = 2.847857
      sig.level = 0.05
          power = 0.9
    alternative = one.sided

NOTE: n is number in *each* group
```

Síla testu

Dvouvýběrový t-test: Za předpokladu alternativní hypotézy H_A: $\mu_X - \mu_Y = \delta$ má testová statistika necentrální t-rozdělení s parametrem necentrality

$$\nu = \frac{\delta}{\sigma\sqrt{1/n_X + 1/n_Y}}$$

```
> x <- data.matrix(read.table("dodavatelX.txt"))
> y <- data.matrix(read.table(„dodavatelY.txt"))
> z=c(x,y)

> power.t.test(delta=2, sd=5.0, sig.level=0.05, power=0.9,
  alt="one.sided")

     Two-sample t test power calculation

              n = 107.7313
          delta = 2
             sd = 5
      sig.level = 0.05
          power = 0.9
    alternative = one.sided

NOTE: n is number in *each* group
```

Síla testu

Dvouvýběrový t-test: Za předpokladu alternativní hypotézy H_A: $\mu_X - \mu_Y = \delta$ má testová statistika necentrální t-rozdělení s parametrem necentrality

$$\nu = \frac{\delta}{\sigma\sqrt{1/n_X + 1/n_Y}}$$

Při ručním počítání obvykle předpokládáme znalost směrodatné odchylky σ a namísto t-rozdělení použijeme standardní normální rozdělení (a jeho kvantily u_p). Požadovaná síla testu je zde označena β. Potom lze odvodit přibližný vztah:

$$n = 2 \cdot \left[\left(\frac{u_{\alpha/2} + u_\beta}{\delta/\sigma} \right)^2 \right] + 1$$

(Toto n je pro každou skupinu.)

Síla testu

Porovnání dvou poměrů:

```
> power.prop.test(n=400, p1=0.67, p2=0.73)

     Two-sample comparison of proportions power calculation

              n = 400
             p1 = 0.67
             p2 = 0.73
      sig.level = 0.05
          power = 0.4567771
    alternative = two.sided

NOTE: n is number in *each* group
```

Síla testu

Porovnání dvou poměrů:

```
> power.prop.test(power=0.9, p1=0.67, p2=0.73)

    Two-sample comparison of proportions power calculation

              n = 1223.788
             p1 = 0.67
             p2 = 0.73
      sig.level = 0.05
          power = 0.9
    alternative = two.sided

NOTE: n is number in *each* group
```

Síla testu

Porovnání dvou poměrů:

```
> power.prop.test(power=0.85, p1=0.67, p2=0.73)

     Two-sample comparison of proportions power calculation

              n = 1045.926
             p1 = 0.67
             p2 = 0.73
      sig.level = 0.05
          power = 0.85
    alternative = two.sided

NOTE: n is number in *each* group
```

Pravděpodobnostní metody ve strojírenství

12. Testy dobré shody

FAKULTA STROJNÍ ČVUT V PRAZE

prof. RNDr. Gejza Dohnal, CSc.

ak. rok 2023/2024

12. Testy dobré shody

Klíčové pojmy:

- Nulová a alternativní hypotéza
- Hladina významnosti testu
- Chyby 1. a 2. druhu
- p-hodnota

Klíčové vztahy:

- Test ANOVA, podmínky
- Kruskal-Wallis test

Pravděpodobnostní modely

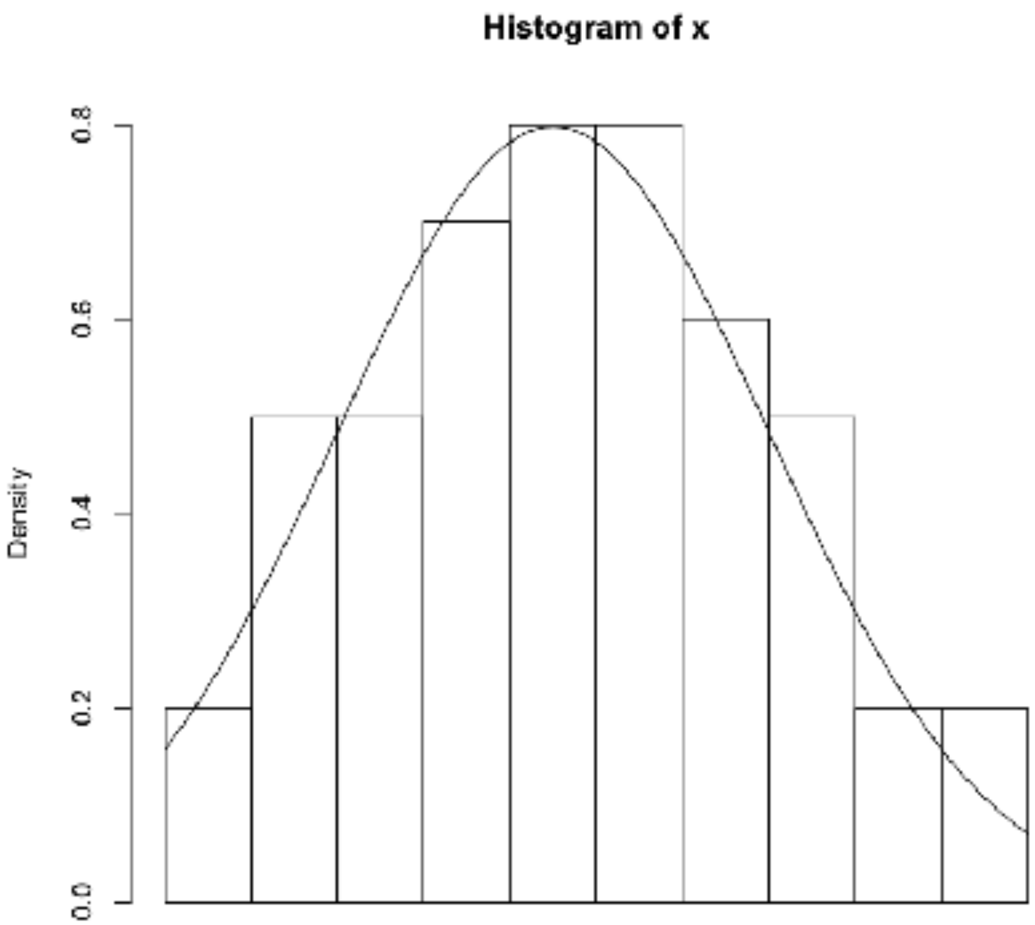
1) Diskrétní:

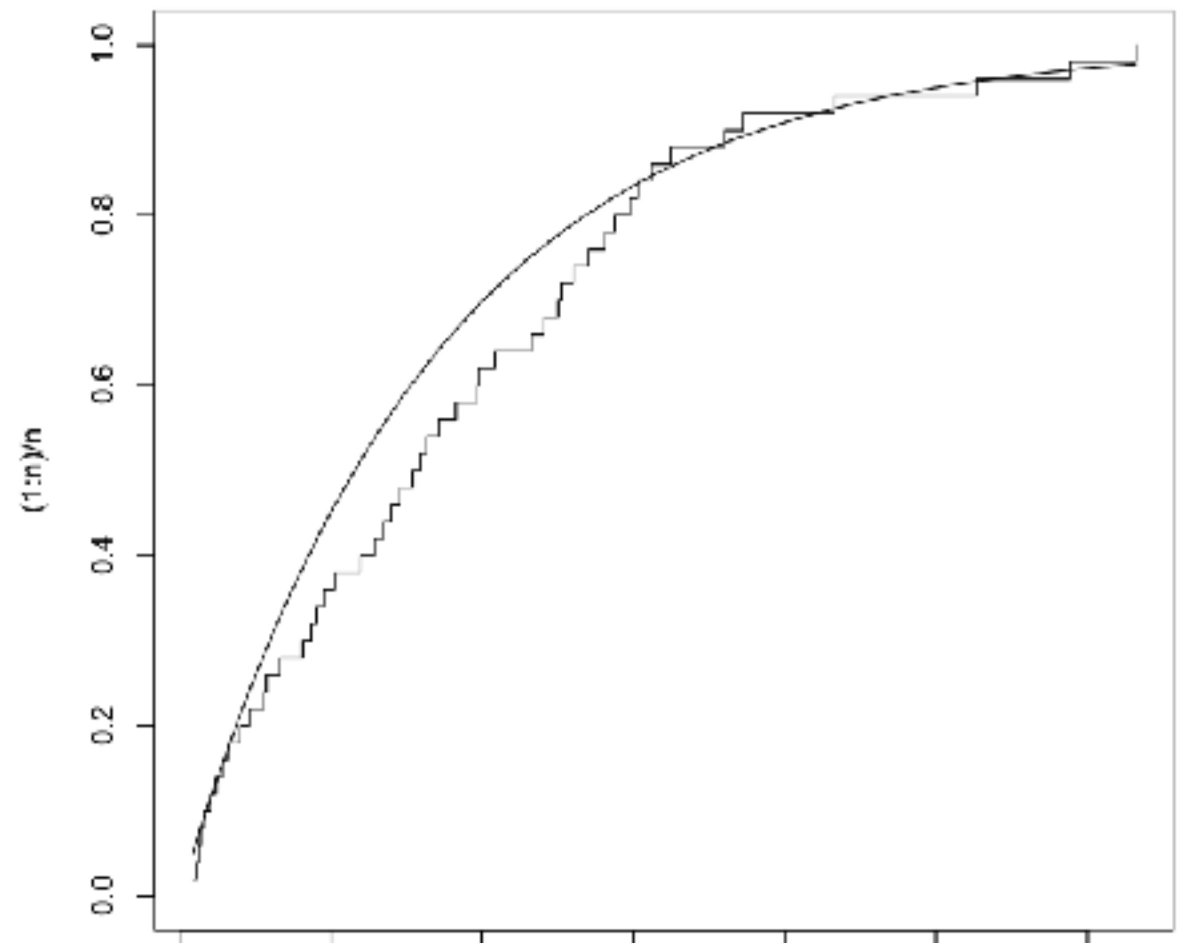
- Rovnoměrný $\Omega = \{1, 2, \dots, N\}$
- Alternativní $\Omega = \{0, 1\}$
- Binomický $\Omega = \{0, 1, \dots, n\}$
- Hypergeometrický $\Omega = \{\max(0, n + M - N), \dots, \min(n, M)\}$
- Geometrický $\Omega = \{0, 1, 2, \dots\}$
- Poissonův $\Omega = \{0, 1, 2, \dots\}$

2) Spojité:

- Rovnoměrný $\Omega = \langle a, b \rangle$
- Normální $\Omega = (-\infty, \infty)$
- Exponenciální $\Omega = \langle 0, \infty)$
- Weibullův $\Omega = \langle 0, \infty)$
- Logaritmicko-normální $\Omega = \langle 0, \infty)$

Testy dobré shody



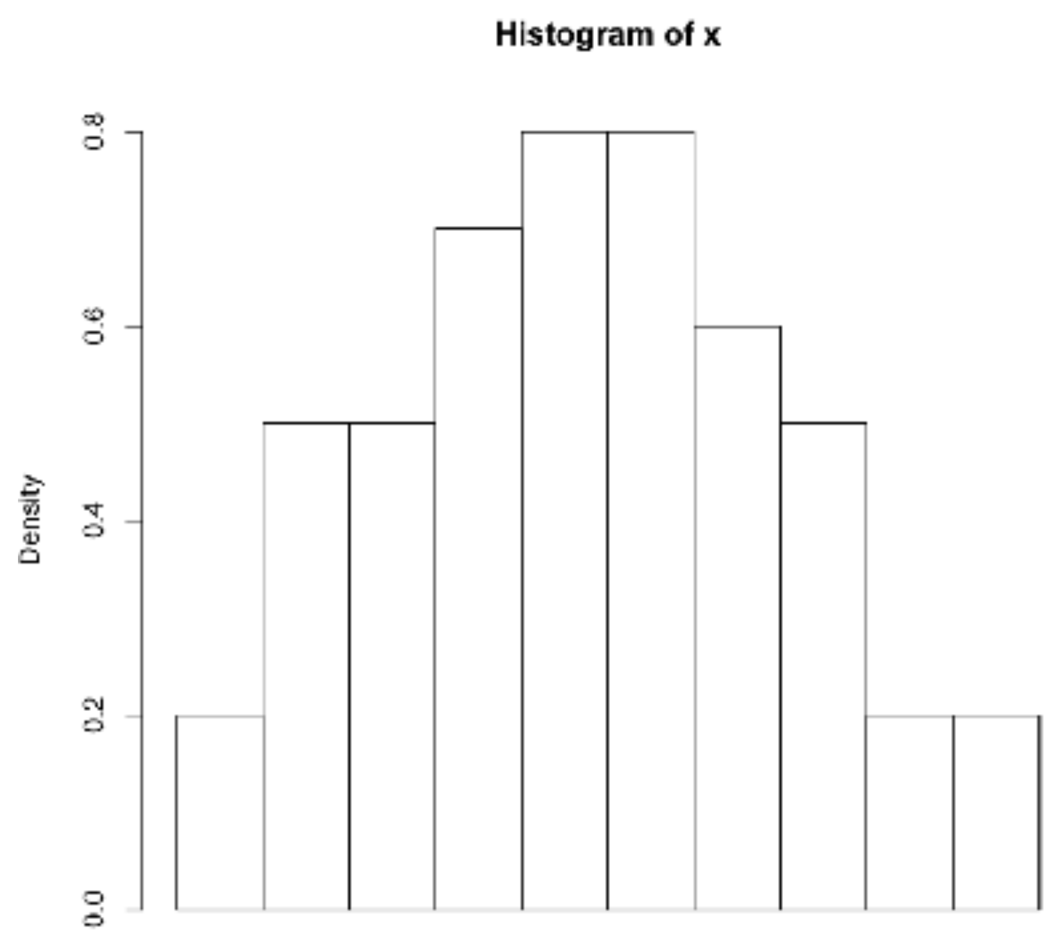


Jaká je shoda pozorovaného experimentu s teoretickým modelem?

Testy dobré shody

Co máme k dispozici?

1) Pozorování výsledků experimentu (měření) = data

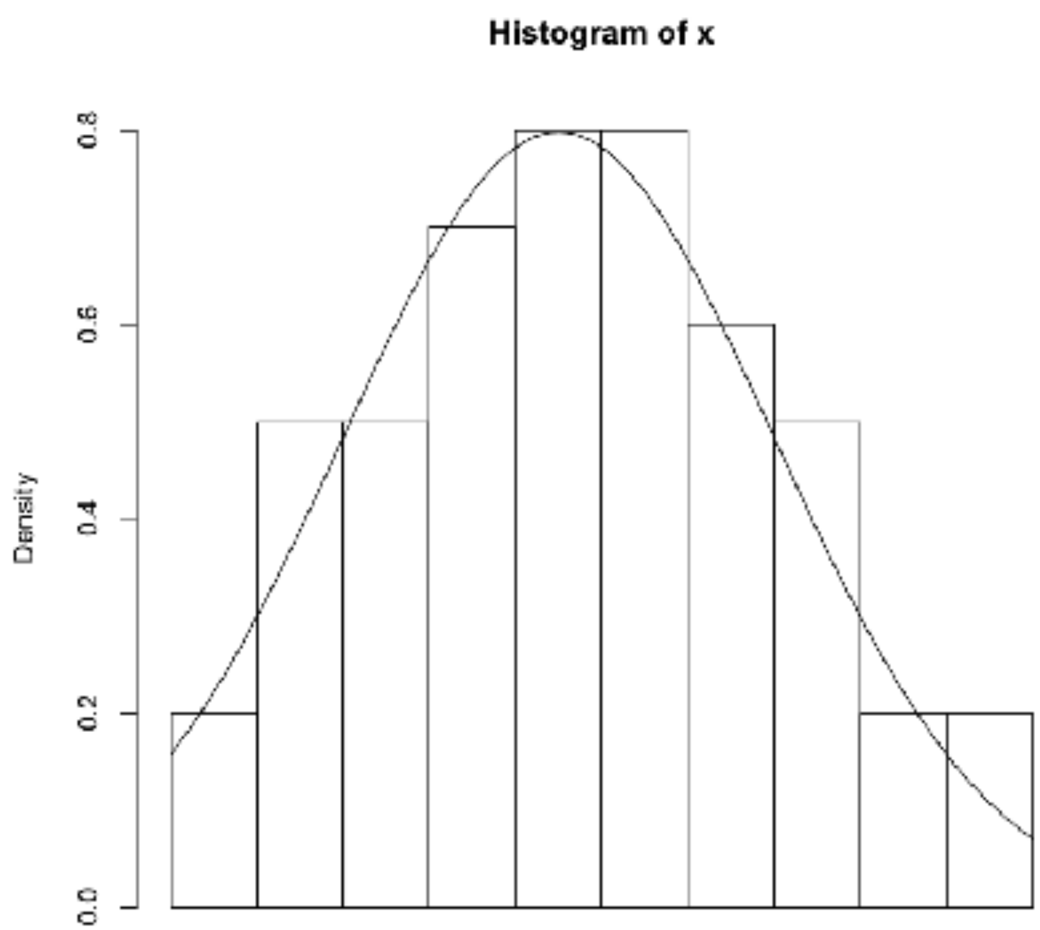


Testy dobré shody

Co máme k dispozici?

1) Pozorování výsledků experimentu (měření) = data

2) Představu o hypotetickém (teoretickém) rozdělení pozorované veličiny



Testy dobré shody

Co s tím?

1) Základní informaci o rozdělení nám poskytují výběrové momenty:

1. výběrový moment = aritmetický průměr: .... $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
   (bodový odhad střední hodnoty)

2. výběrový centrální moment = výběrový rozptyl .... $m_2 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2$

3. výběrový centrální moment .... $m_3 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^3$
   (bodový odhad
   koeficientu šikmosti: $S_{kew} = \frac{m_3}{\sqrt{m_2^3}}$

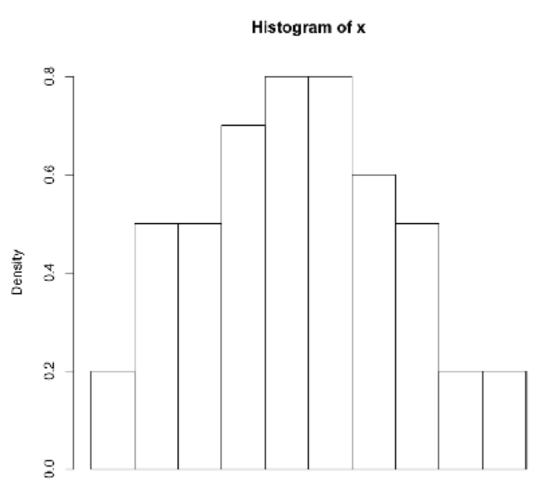
4. výběrový centrální moment .... $m_4 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^4$
   (bodový odhad
   koeficientu špičatosti: $K_{urt} = \frac{m_4}{m_2^2}$ -3

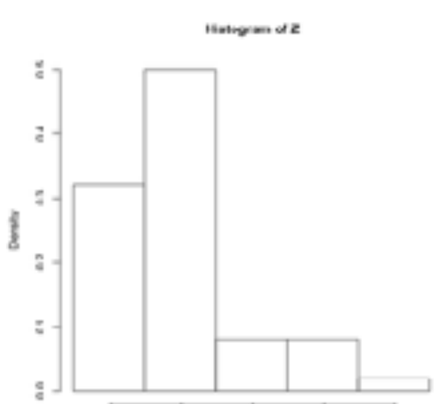
Testy dobré shody

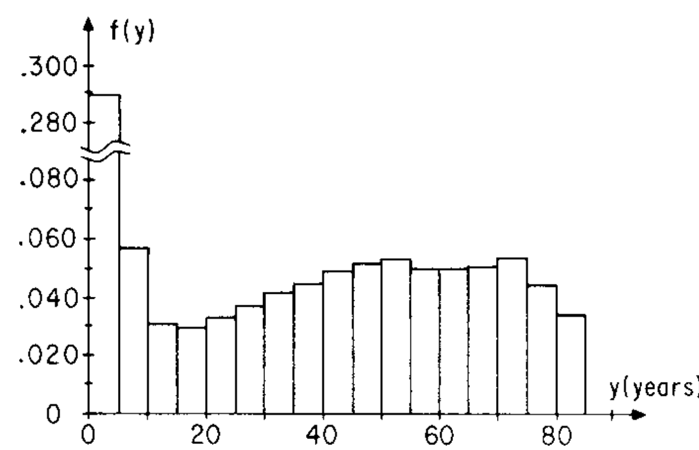
Co s tím?

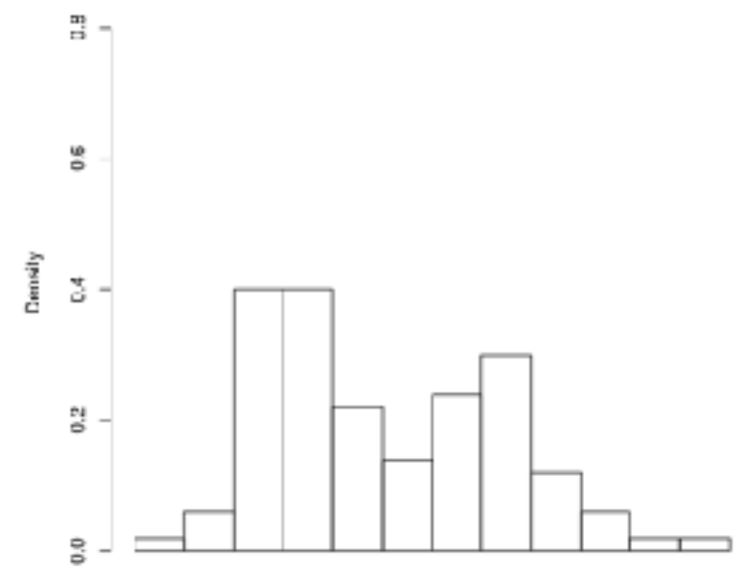
2) Grafická analýza

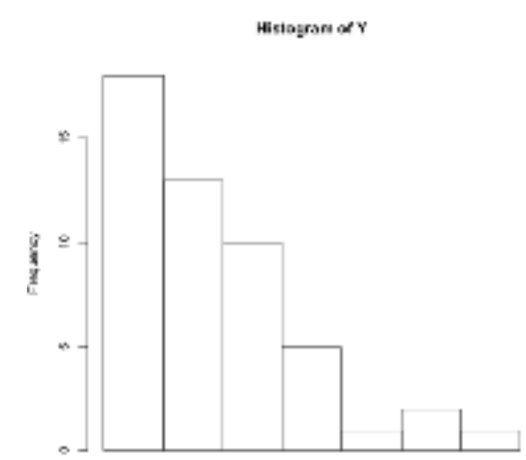
- Histogram   - poskytuje předběžnou představu o tvaru hustoty











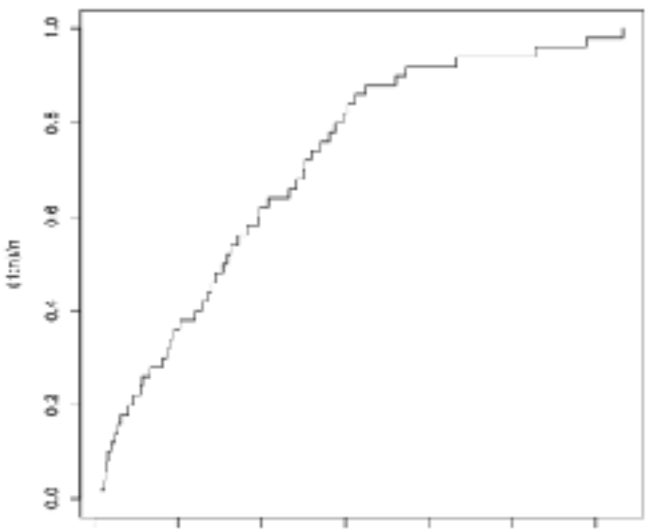
Lze použít například Sturgessovo pravidlo pro volbu počtu tříd: $k = 1 + \frac{10}{3} \log_{10} N$

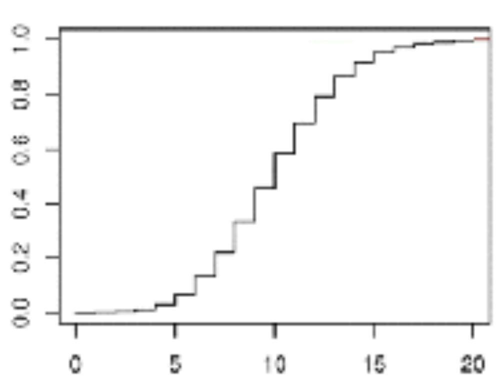
Testy dobré shody

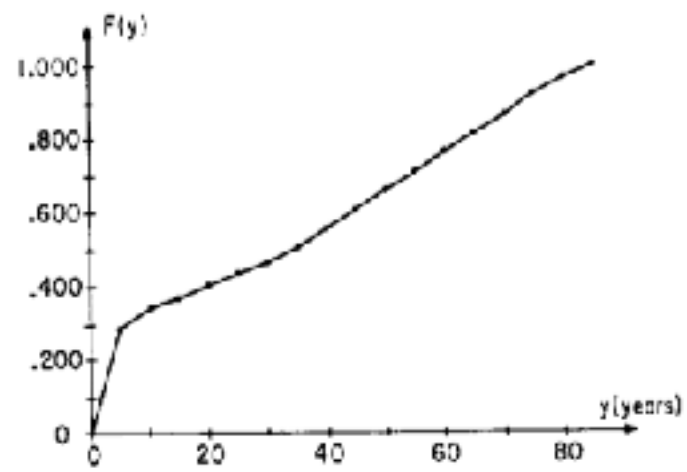
Co s tím?

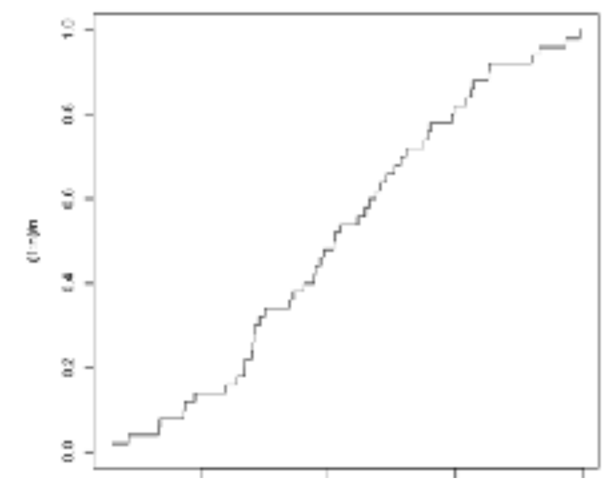
2) Grafická analýza

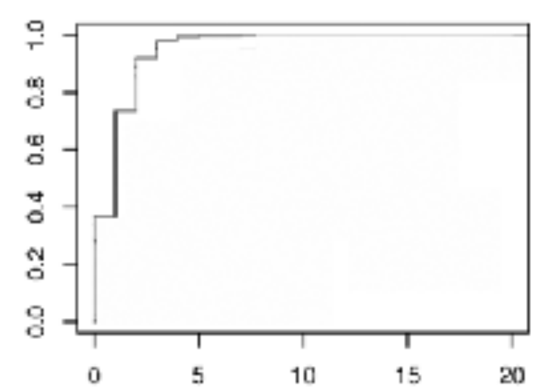
- Histogram - poskytuje předběžnou představu o tvaru hustoty
- Empirická distribuční funkce - poskytuje předběžnou představu o tvaru distribuční funkce











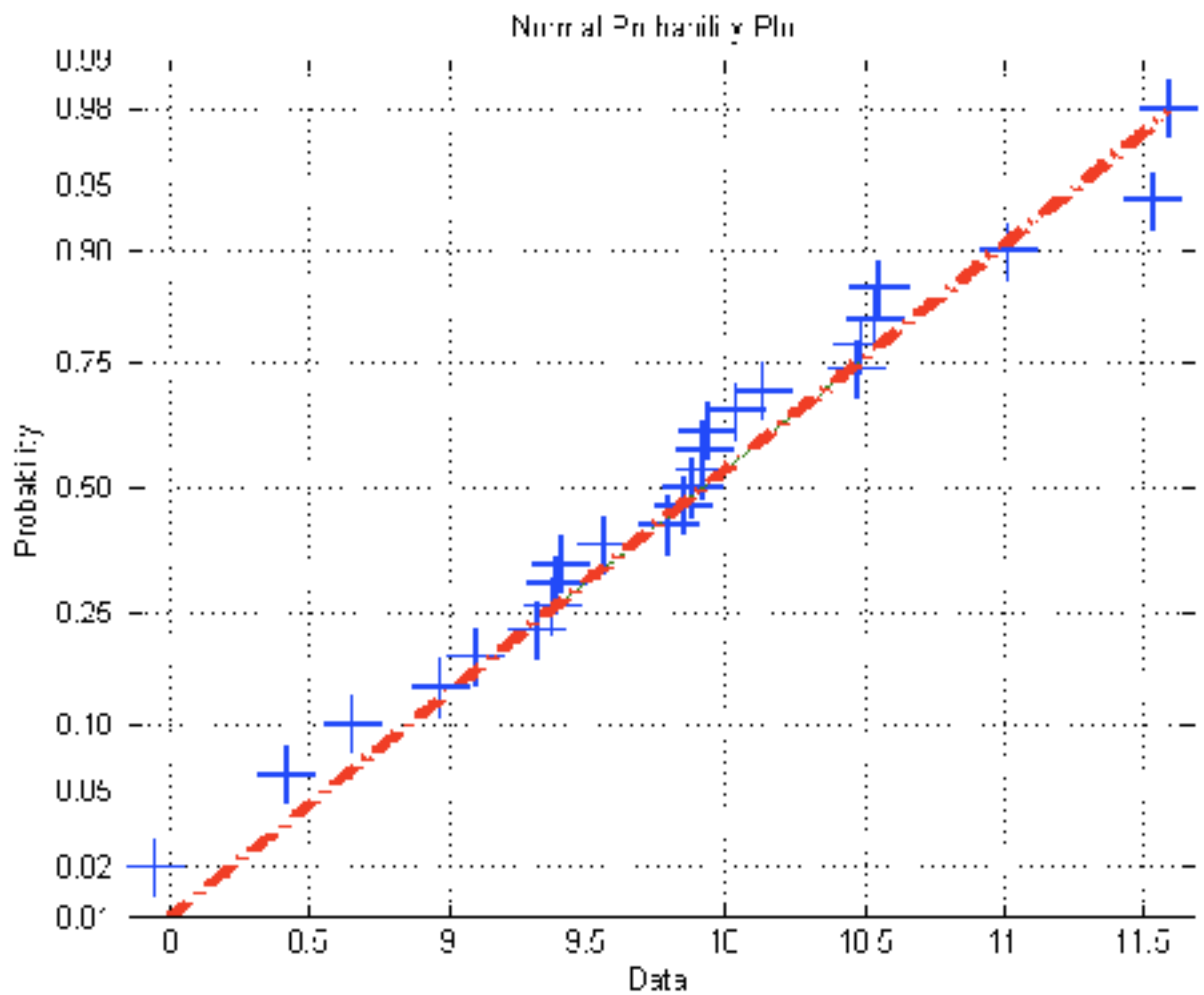
Testy dobré shody

Co s tím?

2) Grafická analýza

- Histogram - poskytuje předběžnou představu o tvaru hustoty
- Empirická distribuční funkce - poskytuje předběžnou představu o tvaru distribuční funkce
- Pravděpodobnostní papír
  osa x: lineární
  osa Y: transformované
  "pravděpodobnostní" měřítko

  Zakreslujeme dvojice $(x_{(i)}, i/n)$



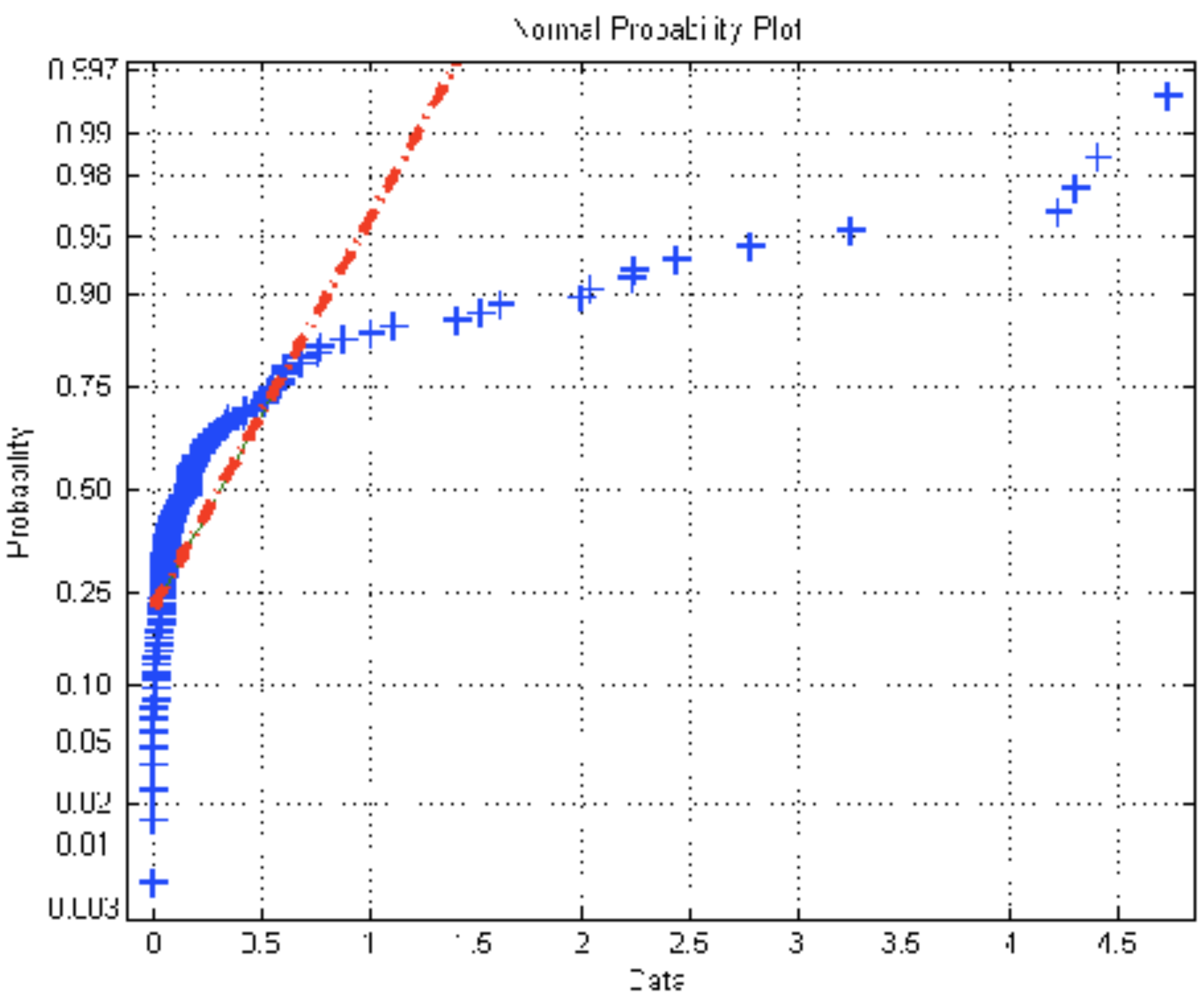
Testy dobré shody

Co s tím?

2) Grafická analýza

- Histogram - poskytuje předběžnou představu o tvaru hustoty
- Empirická distribuční funkce - poskytuje předběžnou představu o tvaru distribuční funkce
- Pravděpodobnostní papír
  osa x: lineární
  osa Y: transformované „pravděpodobnostní" měřítko

  Zakreslujeme dvojice $(x_{(i)}, i/n)$



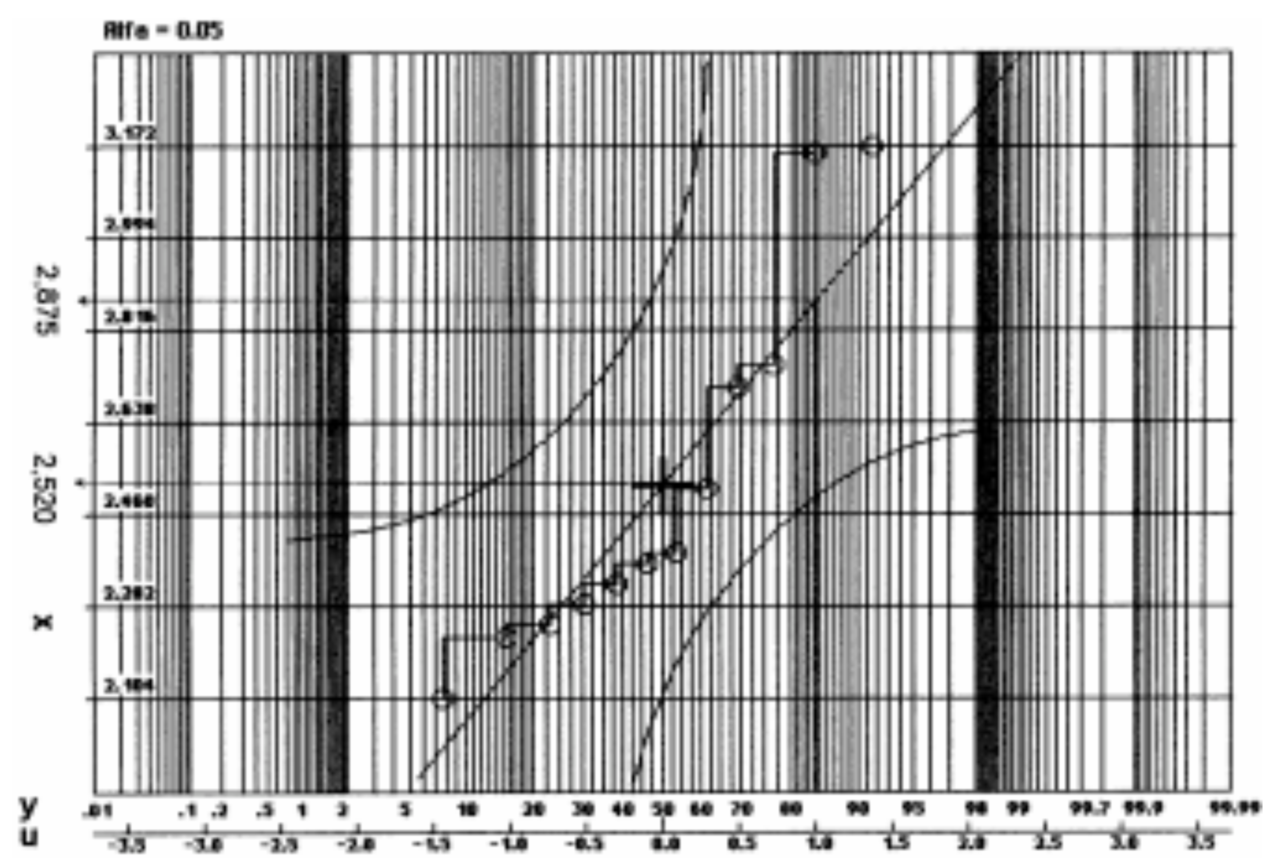
Testy dobré shody

Co s tím?

2) Grafická analýza

- Histogram - poskytuje předběžnou představu o tvaru hustoty
- Empirická distribuční funkce - poskytuje předběžnou představu o tvaru distribuční funkce
- Pravděpodobnostní papír
  osa x: lineární
  osa Y: transformované
  "pravděpodobnostní" měřítko

  Zakreslujeme dvojice $(x_{(i)}, i/n)$



Testy dobré shody

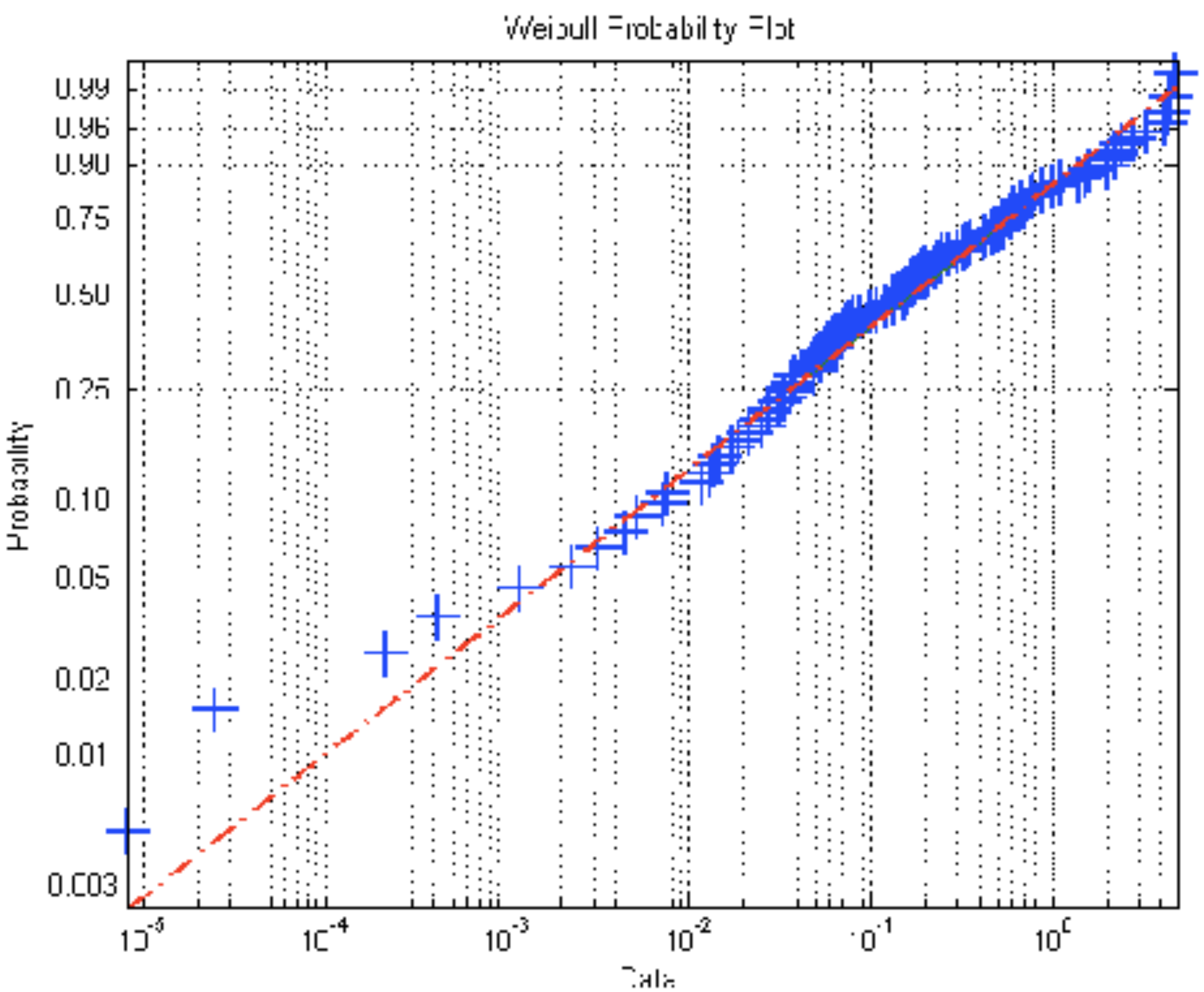
Co s tím?

2) Grafická analýza

- Histogram - poskytuje předběžnou představu o tvaru hustoty
- Empirická distribuční funkce - poskytuje předběžnou představu o tvaru distribuční funkce

- Pravděpodobnostní papír
  osa x: lineární
  osa Y: transformované
  "pravděpodobnostní" měřítko

  Zakreslujeme dvojice $(x_{(i)}, i/n)$



Testy dobré shody

Co s tím?

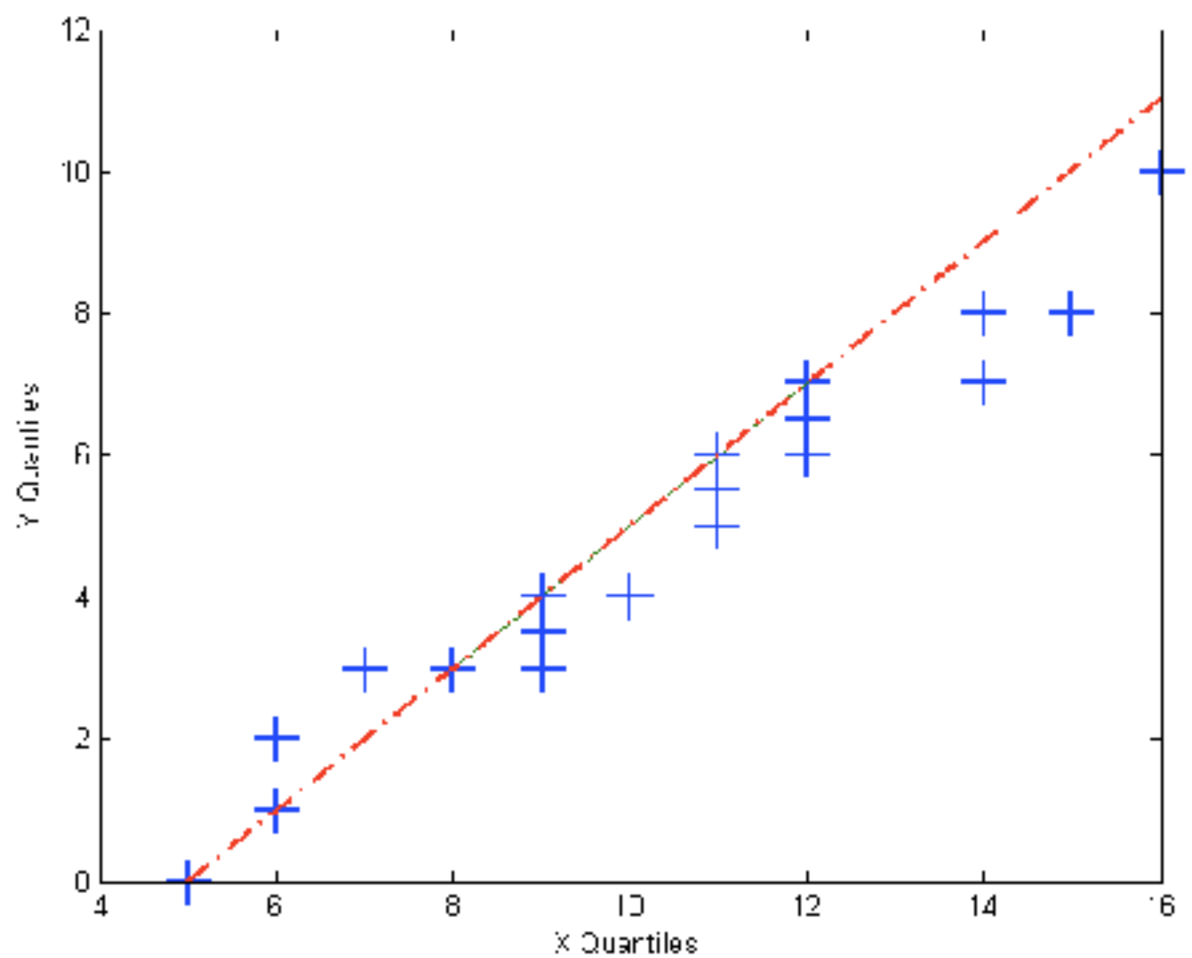
2) Grafická analýza

- Histogram - poskytuje předběžnou představu o tvaru hustoty
- Empirická distribuční funkce - poskytuje předběžnou představu o tvaru distribuční funkce
- Pravděpodobnostní papír
  osa x: lineární
  osa Y: transformované "pravděpodobnostní" měřítko

  Zakreslujeme dvojice $(x_{(i)}, i/n)$

- Q-Q graf
  osa x: měření
  osa y: kvantily hypotetické d.f.

  Zakreslujeme dvojice $(x_{(i)}, F^{-1}(i/n))$



Testy dobré shody

Co s tím?

Pomocí grafické analýzy můžeme metodou srovnání se standardními modely pouze odhadnout typ rozdělení

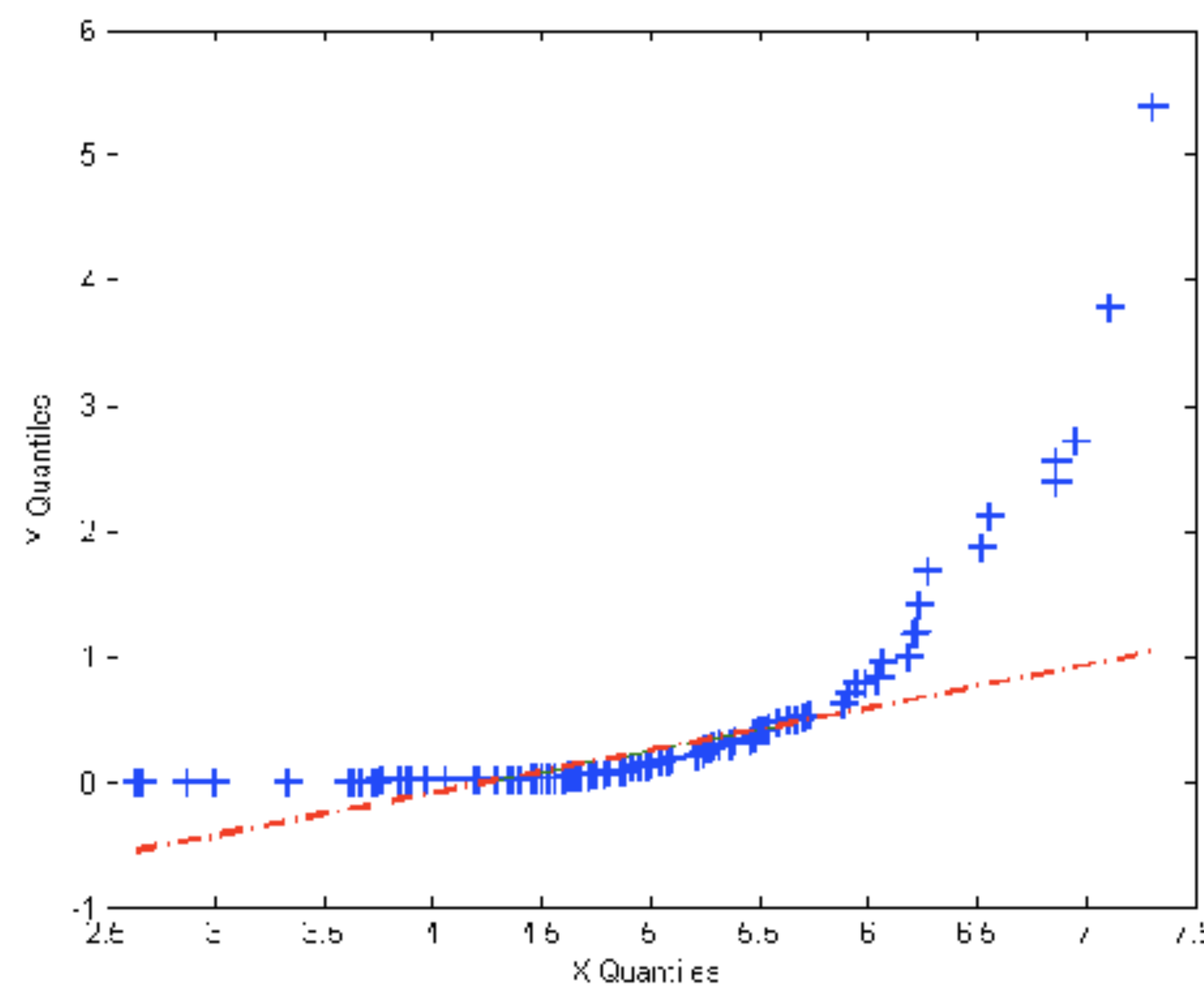
2) Grafická analýza

- Histogram - poskytuje předběžnou představu o tvaru hustoty
- Empirická distribuční funkce - poskytuje předběžnou představu o tvaru distribuční funkce
- Pravděpodobnostní papír
  osa x: lineární
  osa Y: transformované "pravděpodobnostní" měřítko

  Zakreslujeme dvojice $(x_{(i)}, i/n)$

- Q-Q graf
  osa x: měření
  osa y: kvantily hypotetické d.f.

  Zakreslujeme dvojice $(x_{(i)}, F^{-1}(i/n))$



Testy dobré shody

Co můžeme dělat dál?

3) Kvantitativní testy hypotéz o daném typu rozdělení

- poskytnou nám objektivní míru shody dat s teoretickým modelem

nulová hypotéza : $H_0 : \quad F(x) = F_0(x)$

alternativní hypotéza: $H_A : \quad F(x) \neq F_0(x)$

testová statistika : $T(X_1, X_2, \ldots, X_n)$

hladina významnosti: α

chyba 1. druhu: zamítneme hypotézu, která platí

chyba 2. druhu: nezamítneme hypotézu, která neplatí

hladina významnosti testu: pravděpodobnost chyby 1. druhu

síla testu: pravděpodobnost zamítnutí hypotézy, když neplatí

p-hodnota: nejmenší hladina významnosti, při které bychom ještě zamítli nulovou hypotézu.

Testy dobré shody

Co můžeme dělat dál?

3) Kvantitativní testy hypotéz o daném typu rozdělení

- poskytnou nám objektivní míru shody dat s teoretickým modelem

nulová hypotéza : $H_0 : \quad F(x) = F_0(x)$

alternativní hypotéza: $H_A : \quad F(x) \neq F_0(x)$

testová statistika : $T(X_1, X_2, \ldots, X_n)$

hladina významnosti: α

- Chí-kvadrát test dobré shody
- Kolmogorov-Smirnovův test
- Testy normality (Shapiro-Wilkův test, testy na základě šikmosti a špičatosti, Lilieforsův, Anderson-Darlingův test)

Chí-kvadrát test dobré shody

Test srovnává empirické a teoretické četnosti při zadaném třídění:

i) provedeme roztřídění naměřených hodnot do k tříd

ii) napočítáme empirické četnosti $n_1, n_2, ..., n_k$

iii) napočítáme pravděpodobnosti tříd $p_1, p_2, ..., p_k$ při hypotetickém rozdělení (kde $p_j = F(x_{j+1})$- $F(x_j)$)

iv) napočítáme teoretické četnosti $np_1, np_2, ..., np_k$

v) pokud pro všechna $j = 1, 2, ..., k$ platí $np_j > 5$, spočítáme hodnotu testové statistiky

$$\chi^2 = \sum_{j-1}^{k} \frac{(n_j - np_j)^2}{np_j}$$

vi) neplatí-li podmínka v bodě (v), provedeme úpravu třídních intervalů (nemusejí být stejně velké)

vii) známe-li parametry hypotetického rozdělení předem, bude mít testová statistika rozdělení $\chi^2(k-1)$ a nulovou hypotézu zamítneme, pokud $\chi^2 \geq \chi^2_{1-\alpha}(k-1)$, kde $\chi^2_{1-\alpha}(k-1)$ je $(1-\alpha)$-kvantil chí-kvadrát rozdělení o $(k$-$1)$ stupních volnosti.

viii) pokud neznámé parametry hypotetického rozdělení odhadujeme z naměřených dat, bude mít testová statistika chí-kvadrát rozdělení o $(k$-r-$1)$ stupních volnosti, kde r je počet odhadovaných parametrů. Nulovou hypotézu v tomto případě zamítneme, pokud bude $\chi^2 \geq \chi^2_{1-\alpha}(k-r-1)$.

Chí-kvadrát test dobré shody

24.52586	24.17119	24.54486	24.44240	23.93455	24.20389	24.19974	24.34851	23.94024	24.21022
24.87474	25.06155	25.48924	25.32572	23.71721	24.61622	25.06676	24.90055	24.36213	24.98580
24.80591	24.20853	24.72623	24.64437	24.70405	23.97645	25.29837	24.46910	24.99453	25.42994
24.66147	24.75773	25.03970	24.44901	25.13285	24.40205	24.78721	23.83656	24.17186	23.65390
24.48244	24.68550	24.22988	23.83956	24.09777	24.52098	24.89240	24.25332	24.14259	25.12906

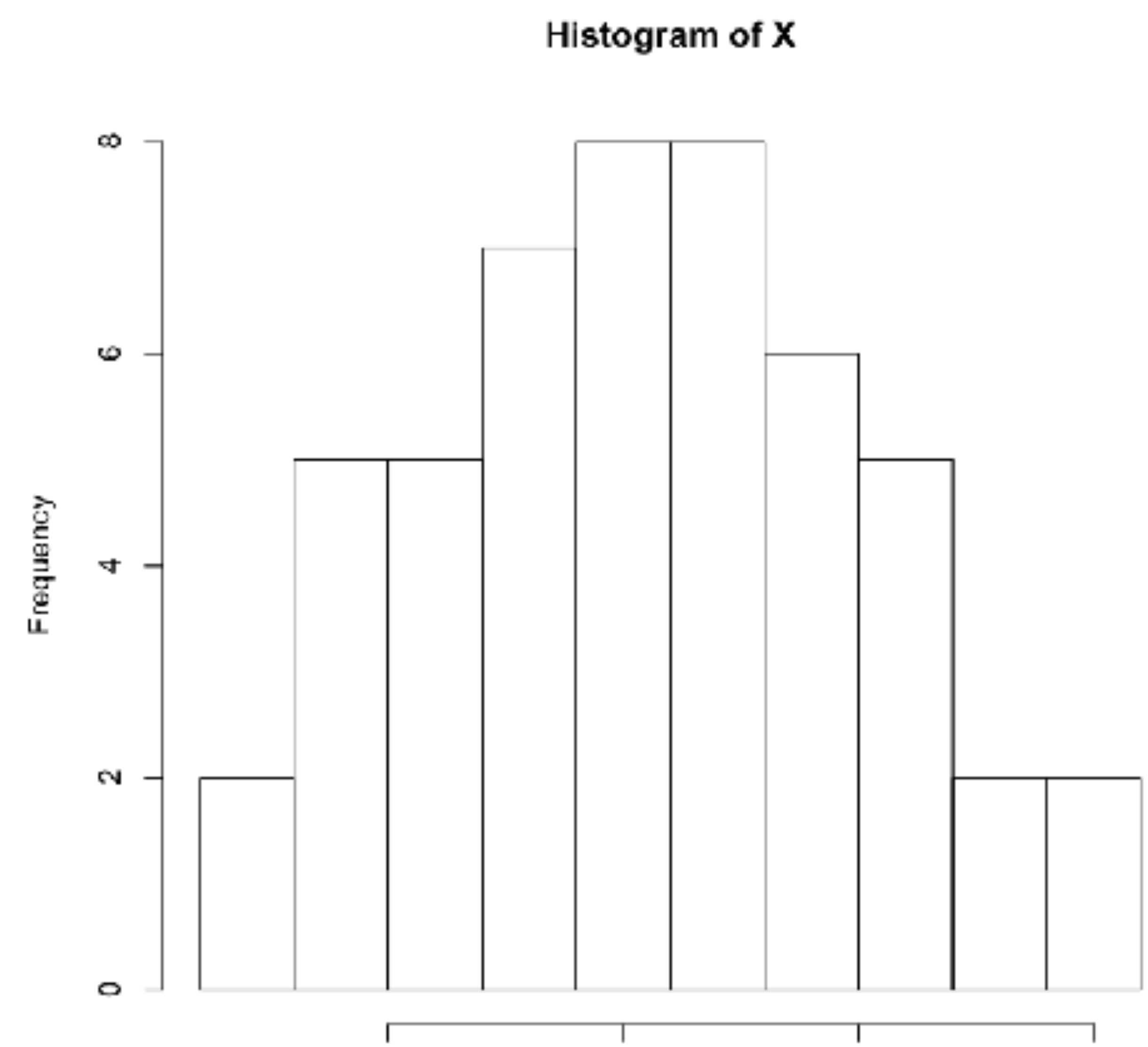
$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i = 24.54689$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{2}(X_i - \bar{X})^2 = 0,2102477$$

$$s = \sqrt{0,2102477} = 0,4585$$

$$N(24,55;0,2102)$$

$$H_0 : F(x) = F_{N(24,55;0,21024)}(x)$$

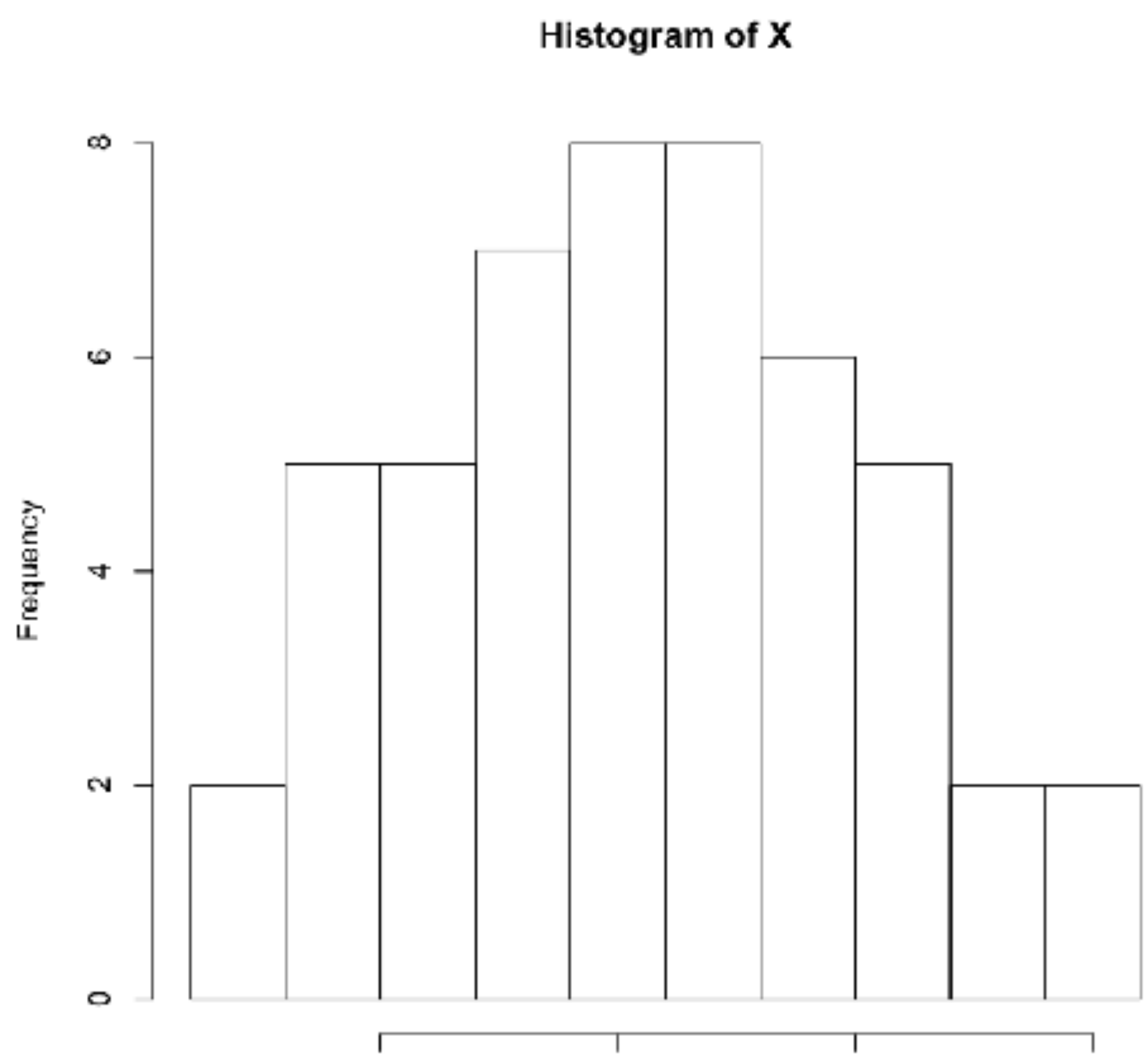


Chí-kvadrát test dobré shody

24.52586	24.17119	24.54486	24.44240	23.93455	24.20389	24.19974	24.34851	23.94024	24.21022
24.87474	25.06155	25.48924	25.32572	23.71721	24.61622	25.06676	24.90055	24.36213	24.98580
24.80591	24.20853	24.72623	24.64437	24.70405	23.97645	25.29837	24.46910	24.99453	25.42994
24.66147	24.75773	25.03970	24.44901	25.13285	24.40205	24.78721	23.83656	24.17186	23.65390
24.48244	24.68550	24.22988	23.83956	24.09777	24.52098	24.89240	24.25332	24.14259	25.12906

	i	ni	pi	npi	(ni-npi)2/npi
23,6	23,8	2	0,0634	3,17	0,4323
23,8	24	5	0,0743	3,72	0,4433
24	24,2	4	0,1187	5,94	0,6312
24,2	24,4	8	0,1572	7,86	0,0025
24,4	24,6	8	0,1727	8,63	0,0463
24,6	24,8	8	0,1572	7,86	0,0025
24,8	25	6	0,1187	5,94	0,0007
25	25,2	5	0,0743	3,72	0,4433
25,2	25,4	2	0,0386	1,93	0,0026
25,4	26	2	0,0248	1,24	0,4636
	suma	50	1,0000	50,00	2,4684

$$H_0 : F(x) = F_{N(24,55;0,21024)}(x)$$



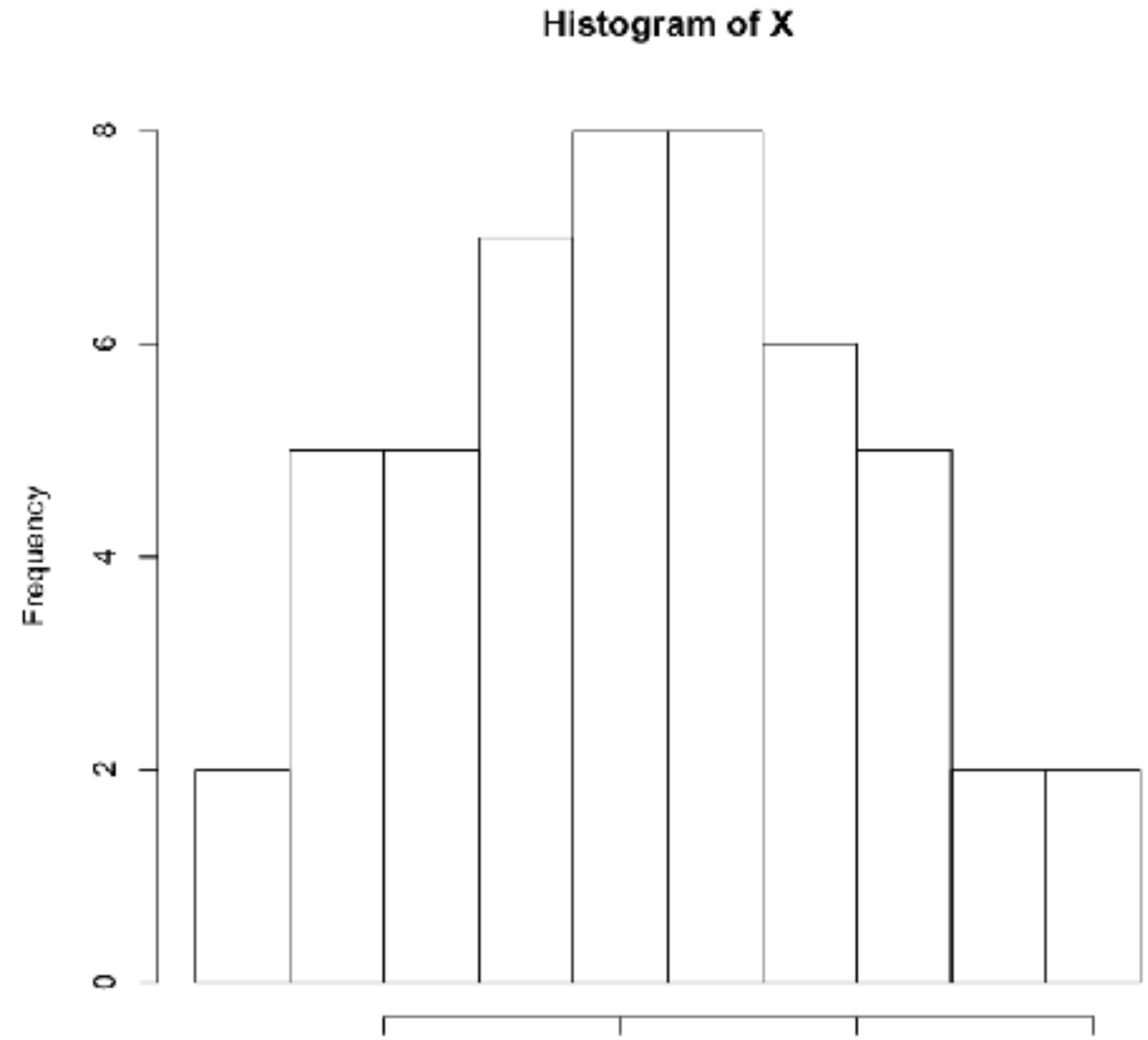
Chí-kvadrát test dobré shody

24.52586	24.17119	24.54486	24.44240	23.93455	24.20389	24.19974	24.34851	23.94024	24.21022
24.87474	25.06155	25.48924	25.32572	23.71721	24.61622	25.06676	24.90055	24.36213	24.98580
24.80591	24.20853	24.72623	24.64437	24.70405	23.97645	25.29837	24.46910	24.99453	25.42994
24.66147	24.75773	25.03970	24.44901	25.13285	24.40205	24.78721	23.83656	24.17186	23.65390
24.48244	24.68550	24.22988	23.83956	24.09777	24.52098	24.89240	24.25332	24.14259	25.12906

	i	ni	pi	npi	(ni-npi)2/npi
23,6	24	7	0,1377	6,89	0,0018
24	24,2	4	0,1187	5,94	0,6312
24,2	24,4	8	0,1572	7,86	0,0025
24,4	24,6	8	0,1727	8,63	0,0463
24,6	24,8	8	0,1572	7,86	0,0025
24,8	25	6	0,1187	5,94	0,0007
25	26	9	0,1377	6,89	0,6482
	suma	50	1,0000	50,00	1,3332

$$\chi^2 = 1,3332 \leq \chi_{0.95}(5) = 11,0705$$

$$H_0 : F(x) = F_{N(24,55;0,21024)}(x)$$



Nulovou hypotézu nezamítáme, a tedy
data lze považovat za normálně rozdělená na hladině významnosti 5 %.

Kolmogorov-Smirnovův test dobré shody

Test srovnává empirickou a teoretickou distribuční funkci pomocí maximálního rozdílu hodnot.

i) seřadíme n naměřených hodnot podle velikosti od nejmenší do největší

ii) pro každou hodnotu $x_{(i)}$ spočteme rozdíly $\left|F_0(x_{(i)}) - \frac{i}{n}\right|, \quad \left|F_0(x_{(i)}) - \frac{i-1}{n}\right|$

iii) největší z těchto rozdílů je hodnota testové statistiky $D(n)$

iv) pokud je hypotetické rozdělení známé včetně parametrů, použijeme krok (v). Jinak musíe použít některou z modifikací K-S testu (Liliefors, Anderson-Darling)

v) pro malá n tuto hodnotu porovnáme s tabulkovou kritickou hodnotou $d_{1-\alpha}(n)$ pro K-S-test. Pro velká n můžeme použít aproximaci

$$d_{1-\alpha}(n) = \sqrt{(1/2n)\ln(2/\alpha)}$$

vi) Pokud je $D(n) \geq d_{1-\alpha}(n)$, nulovou hypotézu zamítáme.

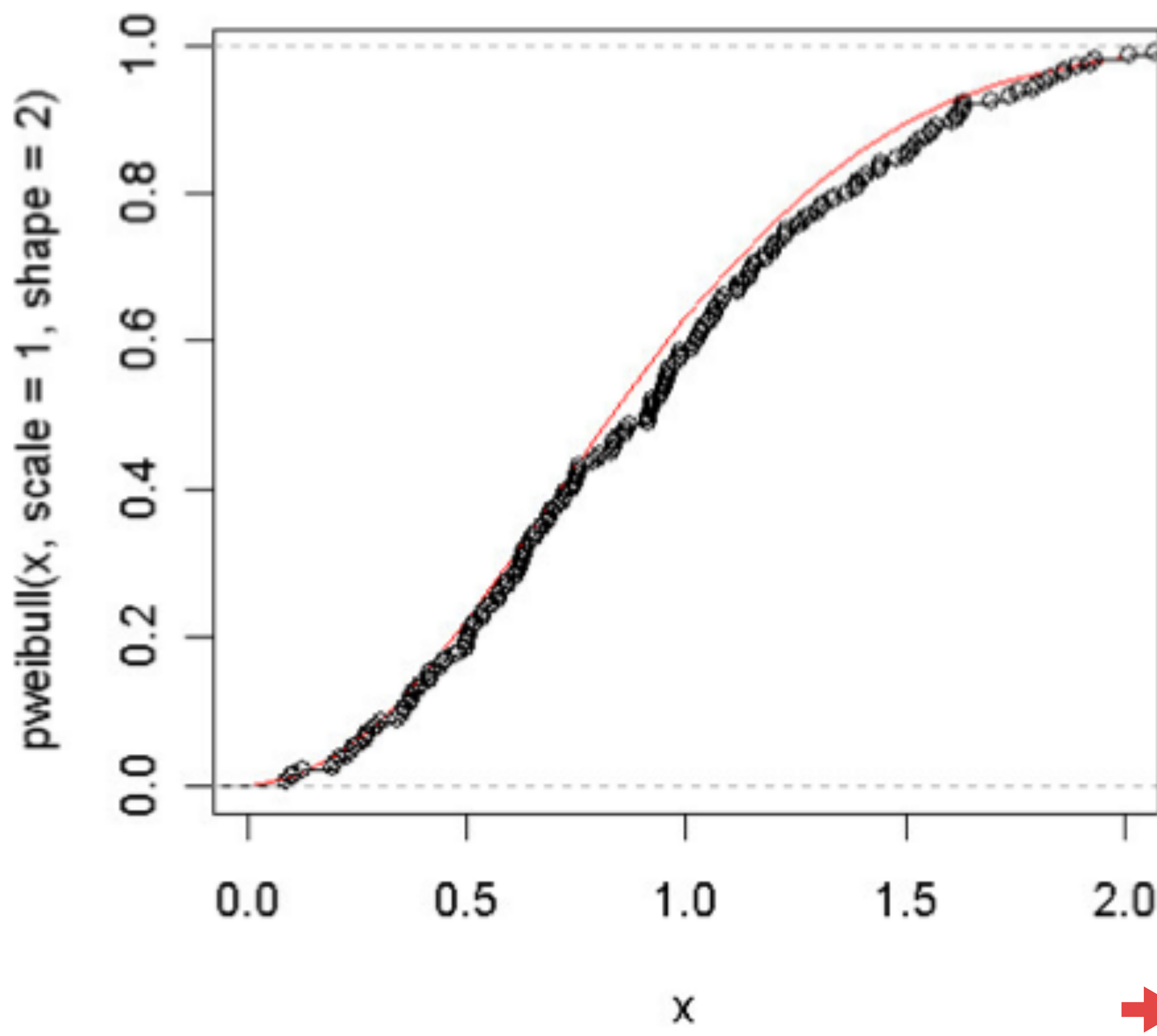
Kolmogorov-Smirnovův test dobré shody

```
> ks.test(x.wei,"pweibull", shape=2,scale=1)

        One-sample Kolmogorov-Smirnov test

data: x.wei D = 0.0623, p-value = 0.4198
alternative hypothesis: two.sided
```

```
> x<-seq(0,2,0.1)
> plot(x,pweibull(x,scale=1,
  shape=2),type="l",col="red")
> plot(ecdf(x.wei),add=TRUE)
```



Testy normality

Testy na základě šikmosti a špičatosti

- Za předpokladu, že výběr pochází z normálního rozdělení, platí pro index šikmosti: $E(S_{kew}^{norm}) = 0$ a $Var(S_{kew}^{norm}) = \dfrac{6(n-2)}{(n+1)(n+3)}$.
- Pro index špičatosti platí vztahy: $E(K_{urt}^{norm}) = -\dfrac{6}{n+1}$ a $Var(K_{urt}^{norm}) = \dfrac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$.
- Máme-li dostatečný počet pozorování (řádově stovky), mají statistiky

$$T_3 = \frac{S_{kew}^{norm}}{\sqrt{Var(S_{kew}^{norm})}} \qquad T_4 = \frac{K_{urt}^{norm} - E(K_{urt}^{norm})}{\sqrt{Var(K_{urt}^{norm})}}$$

přibližně standardní normální rozdělení pravděpodobnosti.

- Tedy hypotézu o normalitě na základě šikmosti zamítáme, pokud bude platit $|T_3| \geq u_\alpha$, nebo pokud bude $p \leq \alpha$, kde $p = 2\min\{\Phi(T_3), 1 - \Phi(T_3)\}$.
- Hypotézu o normalitě na základě špičatosti zamítáme, pokud bude platit $|T_4| \geq u_\alpha$, nebo pokud bude $p \leq \alpha$, kde $p = 2\min\{\Phi(T_4), 1 - \Phi(T_4)\}$.
- Oba testy by se měly používat současně, proto se často používá kombinovaný test s testovou statistikou $T_{34} = T_3^2 + T_4^2$, která má χ^2-rozdělení o 2 stupních volnosti. Hypotézu o normalitě potom zamítáme, když $T_{34} \geq \chi_\alpha^2(2)$.

Testy normality

Shapirův-Wilkův test

Jeden z nejsilnějších testů normality

$$SW = \frac{\left[\sum_{i=1}^{n} a_{(i)} x_{(i)}\right]^2}{\sum_{i=1}^{n} a_{(i)}^2 \sum_{i=1}^{n} (x_{(i)} - \bar{x})^2}$$

kde $a_{(i)} = \Phi^{-1}\left(\frac{8i-3}{8n+2}\right)$ a kritické hodnoty jsou tabelovány.

```
>  shapiro.test(x.norm)

         Shapiro-Wilk normality test

data: x.norm W = 0.9938, p-value = 0.5659
```

Lilieforsův test

Testová statistika je totožná s Kolmogorov-Smirnovovým testem, parametry hypotetického rozdělení odhadujeme z dat a kritické hodnoty hledáme v tabulkách

Testy normality

Anderson-Darlingův test

Test je modifikací Kolmogorovova-Smirnovova testu (používá empirickou distribuční funkce a uspořádaný výběr) s testovou statistikou

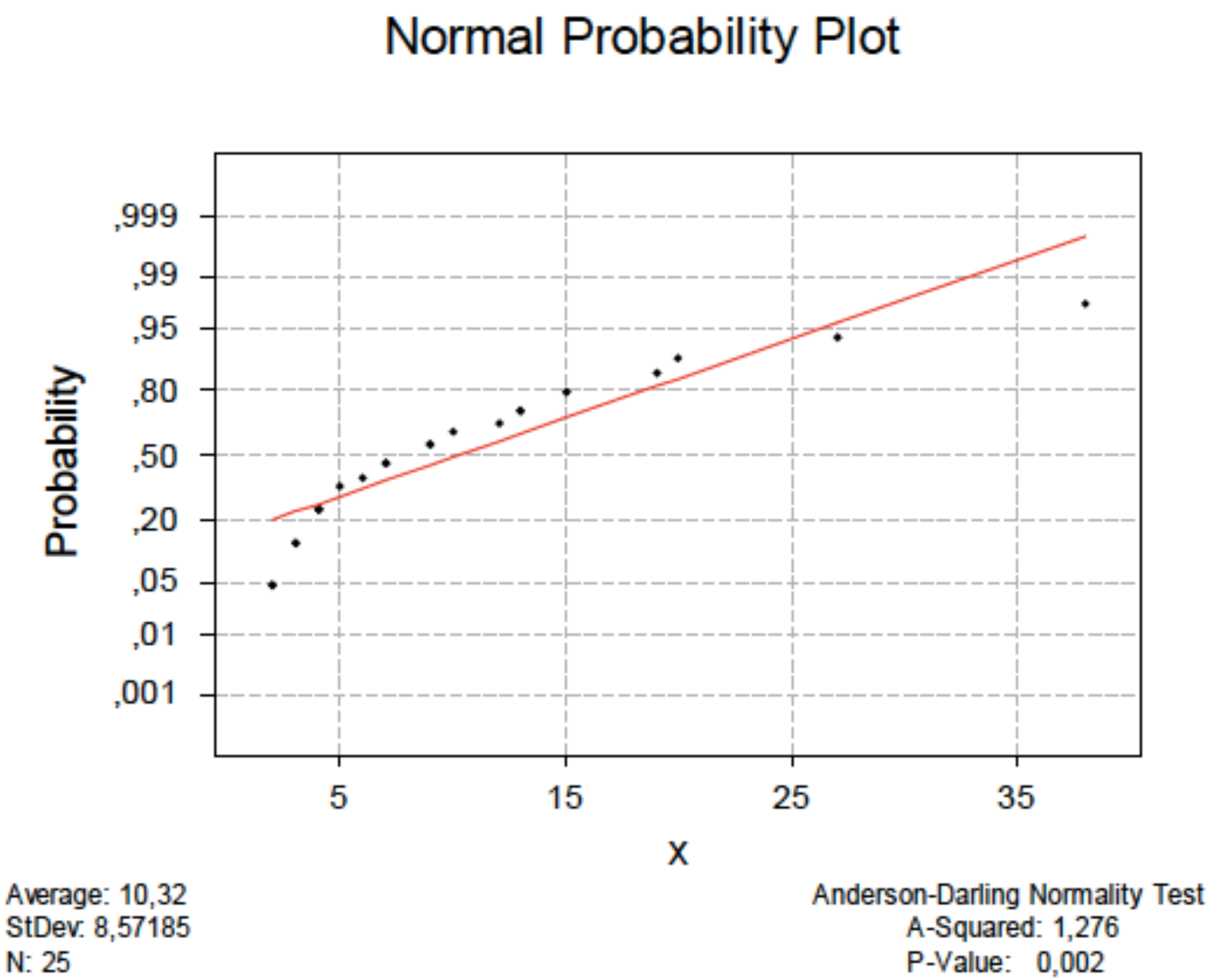
$$AD = -\frac{\sum_{i=1}^{n}(2i-1)\big(\ln F_0(x_{(i)}) + \ln(1-F_0(x_{(n-i+1)}))\big)}{n} - n$$

Kritické hodnoty jsou pro malá n tabelovány, pro velká n lze použít aproximaci:

$$ad_{0,95} = 1,0348(1 - 1,013/n - 0,93/n^2)$$

Testy normality

Anderson-Darlingův test



Testy normality

Anderson-Darlingův test