Vybrané dvouvýběrové testy parametrických hypotéz

Martina Litschmannová

VŠB TECHNICKÁ UNIVERZITA OSTRAVA | FAKULTA ELEKTROTECHNIKY A INFORMATIKY | KATEDRA APLIKOVANÉ MATEMATIKY

Obsah

- Opakování: Testování hypotéz – základní pojmy
- Test o shodě rozptylů (F-test),
- Testy o shodě středních hodnot (t-test, Aspinové-Welchův test),
- Test o shodě mediánů (Mannův-Whitneyho test),
- Test o shodě parametrů dvou binomických rozdělení (test homogenity dvou binomických rozdělení),

- Párové testy (párový t-test, párový znaménkový test)

Jak postupovat při testování hypotéz?

Klasický přístup:

- Formulujeme **nulovou a alternativní hypotézu**.
- Stanovíme **hladinu významnosti** (p-st chyby I. druhu).
- Zvolíme tzv. **testovou statistiku**, tj. výběrovou charakteristiku, jejíž rozdělení závisí na testovaném parametru θ.

  Rozdělení testové statistiky za předpokladu platnosti H_0 nazýváme nulové rozdělení, x_p je p-kvantil nulového rozdělení.
- Ověříme **předpoklady testu**!
- Určíme **kritický obor** W^*:

Tvar alternativní hypotézy H_A	Kritický obor W^*
$\theta < \theta_0$	$(-\infty, x_\alpha)$
$\theta > \theta_0$	$(x_{1-\alpha}, \infty)$
$\theta \neq \theta_0$	$\left(-\infty, x_{\frac{\alpha}{2}}\right) \cup \left(x_{1-\frac{\alpha}{2}}, \infty\right)$

- Na základě konkrétní realizace výběru určíme **pozorovanou hodnotu x_{OBS} testové statistiky**.
- **Rozhodneme o výsledku testu:**

$x_{OBS} \in W^*$	Na hladině významnosti α zamítáme H_0 ve prospěch H_A.
$x_{OBS} \notin W^*$	Na hladině významnosti α nezamítáme H_0.

Jak postupovat při testování hypotéz?

Čistý test významnosti:

- Formulujeme **nulovou a alternativní hypotézu**.
- Stanovíme **hladinu významnosti** (p-st chyby I. druhu).
- Zvolíme tzv. **testovou statistiku**, tj. výběrovou charakteristiku, jejíž rozdělení závisí na testovaném parametru θ.
- Ověříme **předpoklady testu**!
- Na základě konkrétní realizace výběru určíme **pozorovanou hodnotu x_{OBS} testové statistiky**.
- Vypočteme ***p*-hodnotu** (angl. „*p*-value“ nebo „significance level“).

Tvar alternativní hypotézy H_A	$p\text{–}hodnota$
$\theta < \theta_0$	$p\text{–}hodnota = F_0(x_{OBS})$
$\theta > \theta_0$	$p\text{–}hodnota = 1 - F_0(x_{OBS})$
$\theta \neq \theta_0$	$p\text{–}hodnota = 2min\{F_0(x_{OBS}); 1 - F_0(x_{OBS})\}$

- **Rozhodneme** o výsledku testu.

$p\text{–}hodnota < \alpha$	Na hladině významnosti α zamítáme H_0 ve prospěch H_A.
$p\text{–}hodnota \geq \alpha$	Na hladině významnosti α nezamítáme H_0.

Chyba I. a II. druhu

- Testování hypotéz je rozhodovací proces, tj. mohou při něm vznikat chyby.

Rozhodnutí

Skutečnost

	Nezamítáme H_0	**Zamítáme H_0**
Platí H_0	Správné rozhodnutí Pravděpodobnost: 1 - α	Chyba I. druhu Pravděpodobnost: α
Platí H_A	Chyba II. druhu Pravděpodobnost: β	Správné rozhodnutí Pravděpodobnost: 1 - β

- **Chyba I. druhu**
  - ✓ zachytíme efekt, který ve skutečnosti neexistuje (falešně pozitivní výsledek)
  - ✓ α … hladina významnosti
- **Chyba II. druhu**
  - ✓ nezachytíme efekt, který ve skutečnosti existuje (falešně negativní výsledek)
  - ✓ $1 - \beta$ … síla testu (**power**), tj. p-st, že zachytíme efekt, který ve skutečnosti existuje

Statistická významnost vs. praktická významnost

- Při použití příliš **malého rozsahu výběru** bude i prakticky významný efekt označen za statisticky nevýznamný.
- Při použití příliš **velkého rozsahu výběru** bude i velmi malý efekt (prakticky nevýznamný) označen za statisticky významný.

Statistická významnost (dle výsledku testu)

Praktická významnost		**Ne (Nezamítáme H_0)**	**Ano (Zamítáme H_0)**
	Ne	OK	**Příliš velký rozsah výběru?**
	Ano	**Příliš malý rozsah výběru?**	OK

- Pomocí testování hypotéz **se snažíme detekovat** (tj. prokázat jako statisticky významný) takový **efekt, který je prakticky významný** (má praktickou důležitost) pro určitý řešený problém. (Power analýza)
- Tohoto cíle chceme dosáhnout **při co nejmenším možném rozsahu výběru** (počtu měřených hodnot).

Test o shodě rozptylů (F test)

Mějme dva **nezávislé** výběry $\boldsymbol{X_1}$ a $\boldsymbol{X_2}$ ze spojitého rozdělení.

$\forall j = 1, 2, \dots, n_1$, kde n_1 je rozsah prvního výběru: $E(X_{1j}) = \mu_1, D(X_{1j}) = \sigma_1^2$

$\forall j = 1, 2, \dots, n_2$, kde n_2 je rozsah druhého výběru: $E(X_{2j}) = \mu_2, D(X_{2j}) = \sigma_2^2$

a předpokládejme, že rozsahy výběrů nepřesahují 5 % velikosti populace, tj. $n_i \leq 0{,}05N_i$, neboli $N_i \geq 20n_i$ pro $i \in \{1,2\}$.

Název testu	**Nulová hypotéza** $(c \in \mathbb{R})$	**Další předpoklady testu**	**Testová statistika** $\mathbf{T(X)}$	**Nulové rozdělení**
Test o shodě rozptylů	$\frac{\sigma_1^2}{\sigma_2^2} = c$ většinou: $\frac{\sigma_1^2}{\sigma_2^2} = 1,$ tj. $\sigma_1^2 = \sigma_2^2$	normalita obou populací	$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}$	F_{n_1-1,n_2-1}

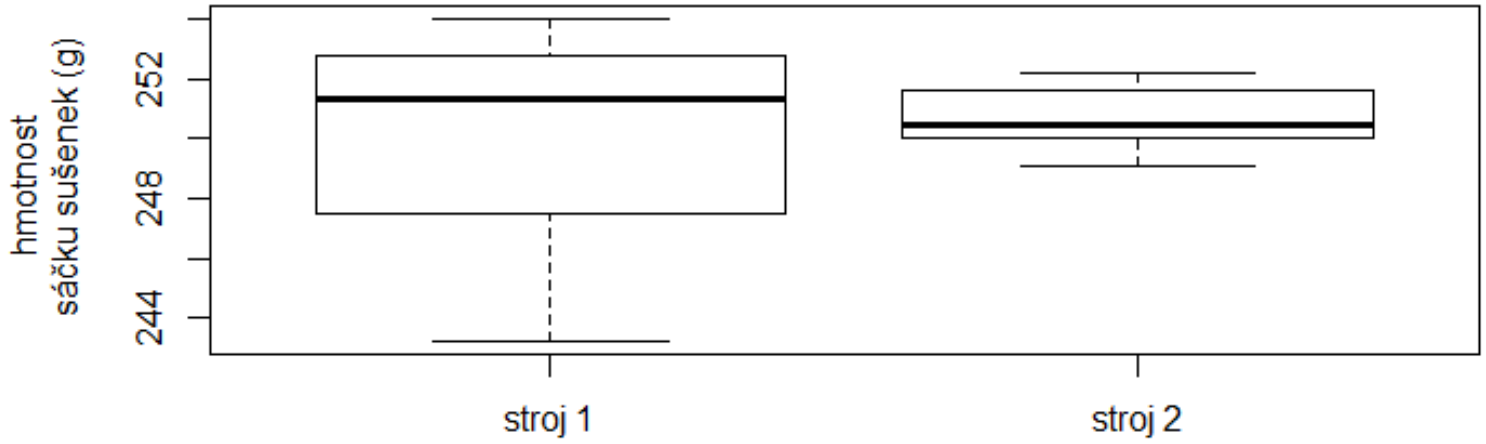
Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

a) Posuďte **na základě explorační analýzy** názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:



$$s_1^2 = 11{,}3$$
$$s_2^2 = 1{,}1$$

$$\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{s_1^2}{s_2^2} = 10{,}1 \gg 2$$

Volíme $s_1^2 > s_2^2$!

Předpokládáme, že tvrzení vedoucího střediska je opodstatněné. Nyní je vhodné toto tvrzení exaktně ověřit.

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

b) Posuďte názor vedoucího střediska **na základě intervalového odhadu** poměru rozptylů na 3% hladině významnosti

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

Bodový odhad: $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{s_1^2}{s_2^2} = 10{,}1 \;\; (s_1^2 = 11{,}3, \;\; s_2^2 = 1{,}1\,)$

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

b) Posuďte názor vedoucího střediska **na základě intervalového odhadu** poměru rozptylů na 3% hladině významnosti

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

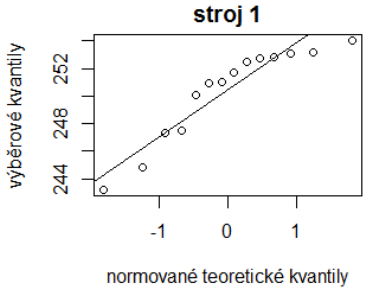
Řešení:

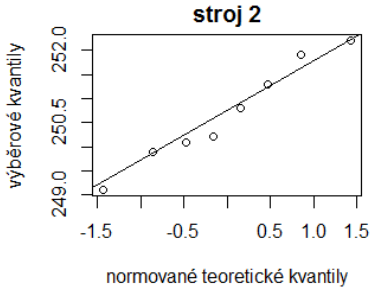
Bodový odhad: $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{s_1^2}{s_2^2} = 10{,}1 \;\; (s_1^2 = 11{,}3,\;\; s_2^2 = 1{,}1\,)$; 100(1 − α)% intervalový odhad $\frac{\sigma_1^2}{\sigma_2^2}$: $\left(\frac{1}{f_{1-\frac{\alpha}{2}}^{n_1-1,n_2-1}}\frac{s_1^2}{s_2^2}\,;\,\frac{1}{f_{\frac{\alpha}{2}}^{n_1-1,n_2-1}}\frac{s_1^2}{s_2^2}\right)$

Ověření předpokladů:

H_0: Data jsou výběrem z normálního rozdělení.

H_A: $\neg H_0$





skupina	Shapirův – Wilkův test (p-hodnota)
stroj 1	0,038
stroj 2	0,835

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

b) Posuďte názor vedoucího střediska **na základě intervalového odhadu** poměru rozptylů na 3% hladině významnosti

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

Bodový odhad: $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{s_1^2}{s_2^2} = 10{,}1 \;\; (s_1^2 = 11{,}3,\;\; s_2^2 = 1{,}1\,)$; 100(1 − α)% intervalový odhad $\frac{\sigma_1^2}{\sigma_2^2}$: $\left(\frac{1}{f_{1-\frac{\alpha}{2}}^{n_1-1,n_2-1}}\frac{s_1^2}{s_2^2}\,;\frac{1}{f_{\frac{\alpha}{2}}^{n_1-1,n_2-1}}\frac{s_1^2}{s_2^2}\right)$

Ověření předpokladů:

Předpoklad normality byl empiricky ověřen na základě Q-Q grafů (viz ….) a exaktně pomocí Shapirova – Wilkova testu (viz …….).

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

b) Posuďte názor vedoucího střediska **na základě intervalového odhadu** poměru rozptylů na 3% hladině významnosti

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

Bodový odhad: $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{s_1^2}{s_2^2} = 10{,}1 \;\; (s_1^2 = 11{,}3, \;\; s_2^2 = 1{,}1\,);$

97% intervalový odhad $\frac{\sigma_1^2}{\sigma_2^2}$: $\left(\frac{1}{f_{0{,}985}^{n_1-1,n_2,1}}\frac{s_1^2}{s_2^2}; \frac{1}{f_{0{,}015}^{n_1-1,n_2,1}}\frac{s_1^2}{s_2^2}\right)$

$$\left(\frac{1}{f_{0{,}985}^{13,7}} \cdot 10{,}1; \frac{1}{f_{0{,}015}^{13,7}} \cdot 10{,}1\right)$$

qf(0.985,13,7)=5,565; qf(0.015,13,7)=0,250

$$(2{,}2; 35{,}1)$$

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

b) Posuďte názor vedoucího střediska **na základě intervalového odhadu** poměru rozptylů na 3% hladině významnosti

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

Bodový odhad: $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{s_1^2}{s_2^2} = 10{,}1 \;\; (s_1^2 = 11{,}3, \;\; s_2^2 = 1{,}1\,);$

97% intervalový odhad $\frac{\sigma_1^2}{\sigma_2^2}$: $(2{,}2; 35{,}1)$

Rozptyl hmotnosti sáčků sušenek vyrobených na stroji 1 je cca 10,1 krát větší než rozptyl hmotnosti sáčků sušenek vyrobených na stroji 2. 97% intervalový odhad tohoto poměru je 2,2 až 35,1. Na hladině významnosti 3 % lze pozorovaný nesoulad ve variabilitě hmotností sáčků sušenek označit za statisticky významný.

(Kdyby byly rozptyly srovnatelné, museli bychom připustit, že lze očekávat poměr rozptylů roven 1, tj. 1 by musela ležet uvnitř nalezeného intervalového odhadu.)

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

c) **Ověřte klasickým testem** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}} \sim F_{n_1-1,n_2-1}$$

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

c) **Ověřte klasickým testem** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \sim F_{n_1-1,n_2-1} \qquad \Rightarrow \qquad \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů testu:

- nezávislé výběry (OK – každé pozorování se vztahuje k jiné statistické jednotce)
- normalita obou populací

skupina	Shapirův – Wilkův test (p-hodnota)
stroj 1	0,038
stroj 2	0,835

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

c) **Ověřte klasickým testem** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů testu:

- nezávislé výběry (OK – každé pozorování se vztahuje k jiné statistické jednotce)
- normalita obou populací (OK – předpoklad normality byl ověřen pomocí Shapirova – Wilkova testu (viz …….))

ad4) Kritický obor:

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

c) **Ověřte klasickým testem** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

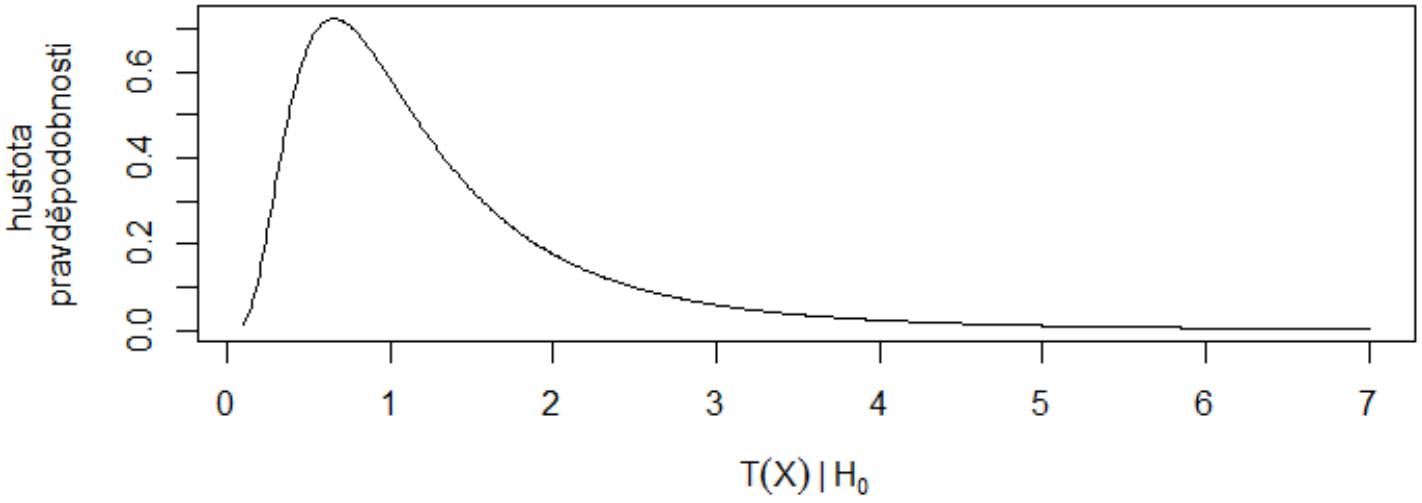
ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů: OK

ad4) Kritický obor:



Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

c) **Ověřte klasickým testem** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

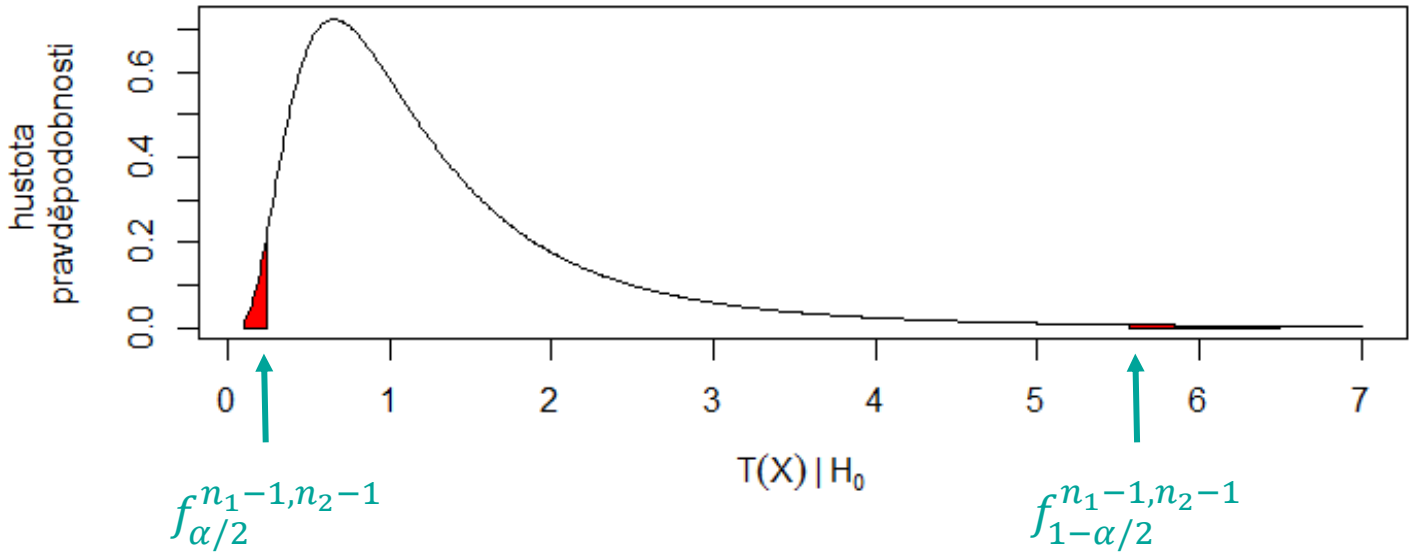
ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů: OK

ad4) Kritický obor:



Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

c) **Ověřte klasickým testem** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

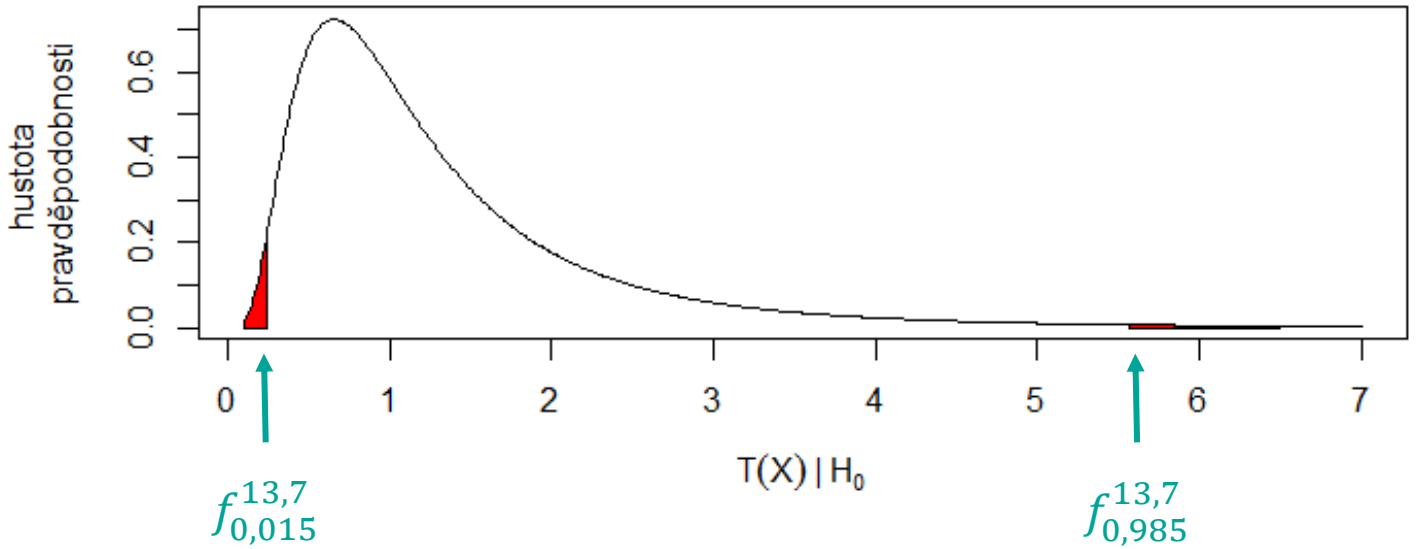
ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů: OK

ad4) Kritický obor:



Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

c) **Ověřte klasickým testem** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

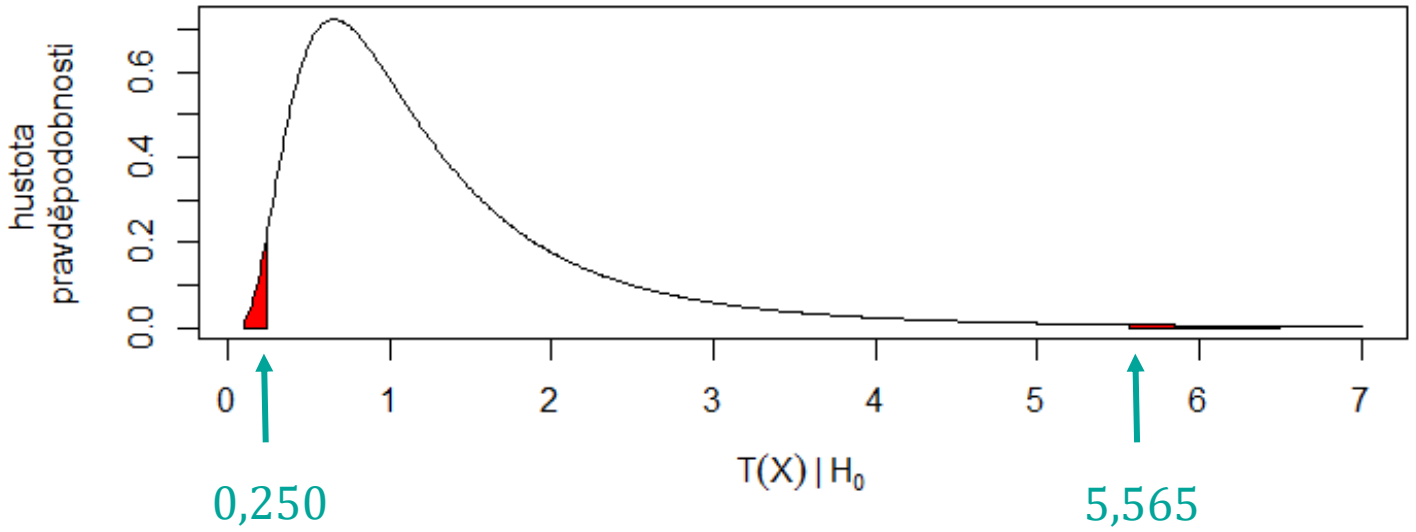
ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů: OK

ad4) Kritický obor:



Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

c) **Ověřte klasickým testem** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů: OK

ad4) Kritický obor: $W^* = (-\infty; 0{,}250) \cup (5{,}565; \infty)$

ad5) Pozorovaná hodnota: $x_{OBS} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} | H_0 = 10{,}1 \cdot 1 = 10{,}1$

ad6) $x_{OBS} \in W^* \quad \Rightarrow$ Na hladině významnosti 3 % zamítáme H_0 ve prospěch H_A.

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

c) **Ověřte klasickým testem** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů: OK

ad4) Kritický obor: $W^* = (-\infty; 0{,}250) \cup (5{,}565; \infty)$

ad5) Pozorovaná hodnota: $x_{OBS} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} | H_0 = 10{,}1 \cdot 1 = 10{,}1$

ad6) $x_{OBS} \in W^*$ $\Rightarrow$ Variabilita hmotností sáčků vyrobených na prvním a druhém stroji se statisticky významně liší.

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

d) **Ověřte čistým testem významnosti** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů: OK

ad4) Pozorovaná hodnota: $x_{OBS} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} | H_0 = 10{,}1 \cdot 1 = 10{,}1$

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

d) **Ověřte čistým testem významnosti** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů: OK

ad4) Pozorovaná hodnota: $x_{OBS} = 10{,}1$

ad5) $p - hodnota$:

$p - hodnota = 2 \cdot min\{F_0(x_{OBS}), 1 - F_0(x_{OBS})\}$,

kde $F_0(x)$ je distribuční funkce Fisherova – Snedecorova rozdělení s 13 stupni volnosti v čitateli a 7 stupni volnosti ve jmenovateli

$F_0(x_{OBS}) \cong 0{,}9974 \quad (pf(10.06,13,7))$, $1 - F_0(x_{OBS}) \cong 0{,}0026$

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

d) **Ověřte čistým testem významnosti** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů: OK

ad4) Pozorovaná hodnota: $x_{OBS} = 10{,}06$

ad5) $p - hodnota \cong 0{,}0052$

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

d) **Ověřte čistým testem významnosti** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů: OK

ad4) Pozorovaná hodnota: $x_{OBS} = 10{,}06$

ad5) $p - hodnota \cong 0{,}005 < 0{,}03$

ad6) Na hladině významnosti 3 % zamítáme H_0 ve prospěch H_A.

Příklad 1

Podnik Čoko používá dva typy strojů pro automatické balení sušenek. Vedoucí střediska je přesvědčen, že variabilita hmotností sáčků sušenek balených na těchto dvou strojích není stejná. Bylo tedy náhodně vybráno 14 balíčků sušenek balených na prvním stroji a 8 balených na druhém stroji.

d) **Ověřte čistým testem významnosti** na hladině významnosti 3 %, zda lze potvrdit názor vedoucího střediska.

STROJ 1 (g)	243,2	244,8	253,1	247,5	251,0	251,7	254,0	252,8	252,5	250,1	247,3	250,9	253,2	252,7
STROJ 2 (g)	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2						

Řešení:

ad1) $H_0\colon \sigma_1^2 = \sigma_2^2, \qquad H_A\colon \sigma_1^2 \neq \sigma_2^2$

ad2) Volba testové statistiky:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$

ad3) Ověření předpokladů: OK

ad4) Pozorovaná hodnota: $x_{OBS} = 10{,}06$

ad5) $p - hodnota \cong 0{,}005 < 0{,}03$

ad6) Variabilita hm. sáčků vyrobených na prvním a druhém stroji se statisticky významně liší.

Testy o shodě středních hodnot

Mějme dva **nezávislé** výběry $\boldsymbol{X_1}$ *a* $\boldsymbol{X_2}$ ze spojitého rozdělení.

$\forall j = 1, 2, \dots, n_1$, kde n_1 je rozsah prvního výběru: $E(X_{1j}) = \mu_1, D(X_{1j}) = \sigma_1^2$

$\forall j = 1, 2, \dots, n_2$, kde n_2 je rozsah druhého výběru: $E(X_{2j}) = \mu_2, D(X_{2j}) = \sigma_2^2$

a předpokládejme, že rozsahy výběrů nepřesahují 5 % velikosti populace, tj. $n_i \leq 0{,}05N_i$, neboli $N_i \geq 20n_i \; pro \; i \; \in \{1,2\}$.

Název testu	**Nulová hypotéza** $(\boldsymbol{c} \in \mathbb{R})$	**Další předpoklady testu**	**Testová statistika** $\boldsymbol{T(X)}$	**Nulové rozdělení**
Dvouvýběrový z test	$\mu_1 - \mu_2 = c$ většinou: $\mu_1 - \mu_2 = 0$, tj. $\mu_1 = \mu_2$	normalita obou populací nebo $n_1 > 30, n_2 > 30$, známe σ_1, σ_2	$\dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$	$N(0;1)$
Dvouvýběrový t test		normalita obou populací, neznáme σ_1, σ_2, $\sigma_1 = \sigma_2$	$\dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}}$	$t_{n_1+n_2-1}$
Aspinové – Welchův test		normalita obou populací, neznáme σ_1, σ_2, $\sigma_1 \neq \sigma_2$	$\dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$	t_ν kde, $\nu = \dfrac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{1}{n_1 - 1}\left(\dfrac{S_1^2}{n_1}\right)^2 + \dfrac{1}{n_2 - 1}\left(\dfrac{S_2^2}{n_2}\right)^2}$

Test o shodě mediánů

Mějme dva **nezávislé** výběry $\boldsymbol{X_1}$ a $\boldsymbol{X_2}$ ze spojitého rozdělení a předpokládejme, že rozsahy výběrů nepřesahují 5 % velikosti populace, tj. $n_i \leq 0{,}05N_i$, neboli $N_i > 20n_i$ pro $i \in \{1,2\}$.				
Název testu	**Nulová hypotéza** $(\boldsymbol{c} \in \mathbb{R})$	**Předpoklady testu**	**Testová statistika** $\boldsymbol{T(X)}$	**Kritický obor W**
Mannův – Whitneyho test	$x_{0,5}^1 = x_{0,5}^2$ (obecně: výběry pochází z populací se stejným rozdělením)	stejný tvar rozdělení	$\min(U_1, U_2)$, kde $U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - T_1$, $U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - T_2$ Poznámka: Pro výpočet nejprve seřadíme všechna pozorování od nejmenšího po největší tak, jako by byly z jednoho vzorku, a přiřadíme jednotlivým hodnotám jejich pořadí. Symbolem T_1 (resp. T_2) označíme součet pořadí hodnot příslušných první (resp. druhé) skupině.	$W = \{T: T \leq x_{krit}\}$ Kritické hodnoty x_{krit} jsou tabelovány (Tab. T7, volíme $n_1 > n_2$).

Test o shodě parametrů dvou binomických rozdělení

Mějme dva **nezávislé** výběry z alternativního rozdělení.
$\forall i = 1, 2, \dots, n_1$, kde n_1 je rozsah prvního výběru: $X_{1i} \sim A(\pi_1)$,
$\forall j = 1, 2, \dots, n_2$, kde n_2 je rozsah druhého výběru: $X_{2j} \sim A(\pi_2)$

a předpokládejme, že rozsahy výběrů nepřesahují 5 % velikosti populace, tj. $n_i \leq 0{,}05 N_i$, neboli $N_i \geq 20 n_i$ pro $i \in \{1,2\}$.

Název testu	**Nulová hypotéza** $(\boldsymbol{c} \in \mathbb{R})$	**Další předpoklady testu**	**Testová statistika** $\boldsymbol{T(X)}$	**Nulové rozdělení**
Test o shodě parametrů dvou bin. rozdělení	$\pi_1 - \pi_2 = c$ většinou: $\pi_1 - \pi_2 = 0$, tj. $\pi_1 = \pi_2$	$n_i > \frac{9}{p_i(1-p_i)}$ pro $i \in \{1,2\}$	$\frac{(p_1-p_2)-(\pi_1-\pi_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}}$	$N(0;1)$

Párové testy

Jsou-li výsledkem zjišťování dvojice náhodných veličin $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$, které tvoří páry závislých pozorování (jde o veličiny zjišťované na stejné statistické jednotce), musíme při ověřování shody polohy přistoupit k párovým testům.

Definujme soubor rozdílů (diferencí)

$$\boldsymbol{D} = (D_1, D_2, \ \dots, D_n), \quad \text{kde } D_i = X_i - Y_i.$$

Lze předpokládat, že náhodné veličiny $(D_1, D_2, \ \dots, D_n)$ jsou nezávislé a že mají stejné rozdělení se střední hodnotou μ_D. Test o shodě dvou středních hodnot prováděný na základě dvou závislých výběrů můžeme převést na jednovýběrový test o střední hodnotě aplikovaný na soubor diferencí (rozdílů) $\boldsymbol{D}$.

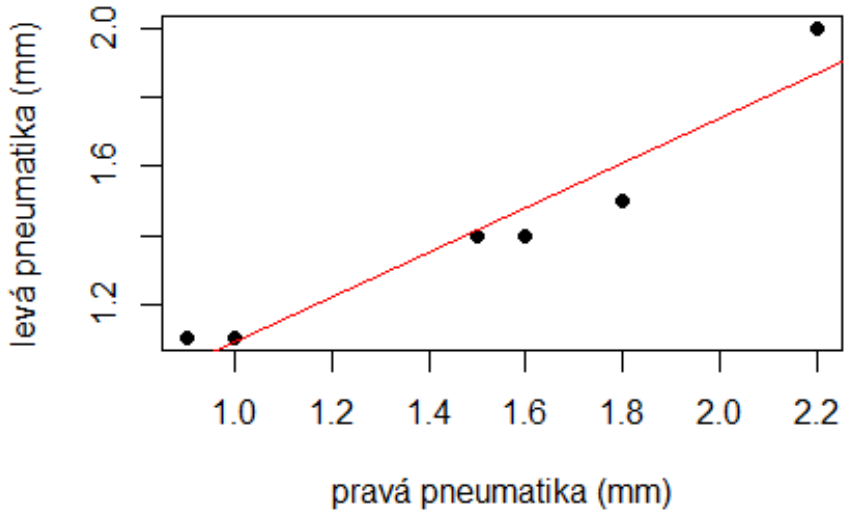
Názvy testů: párový t-test, párový Wilcoxonův test, …

Jak empiricky posoudit shodu párových dat?

U 6aut bylo zjištěno ojetí předních pneumatik (mm).

Ojetí předních pneumatik (mm)						
Pravá pneumatika	1,8	1,0	2,2	0,9	1,5	1,6
Levá pneumatika	1,5	1,1	2,0	1,1	1,4	1,4

Ojíždějí se pravá a levá pneumatika stejně?



$r = 0{,}962$ (výběrový korelační koeficient)

Velmi nevhodný postup!!!

Proč?

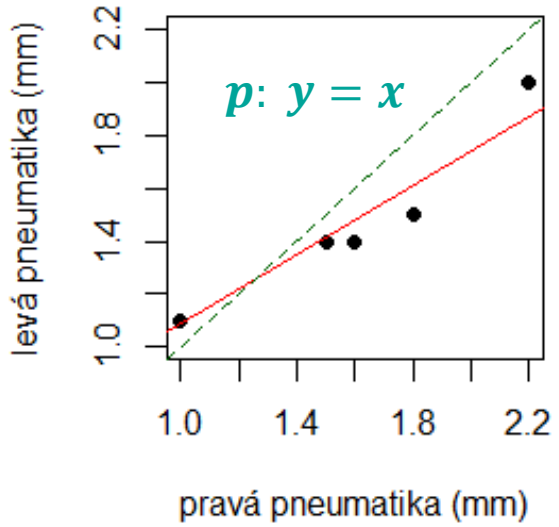
Ve skutečnosti **nechceme sledovat shodu dat s regresní přímkou**, ale shodu dat s osou prvního a třetího kvadrantu.

Jak empiricky posoudit shodu párových dat?

U 6aut bylo zjištěno ojetí předních pneumatik (mm).

Ojetí předních pneumatik (mm)						
Pravá pneumatika	1,8	1,0	2,2	0,9	1,5	1,6
Levá pneumatika	1,5	1,1	2,0	1,1	1,4	1,4

Ojíždějí se pravá a levá pneumatika stejně?



$r = 0{,}962$ (výběrový korelační koeficient)

Velmi nevhodný postup!!!

Proč?

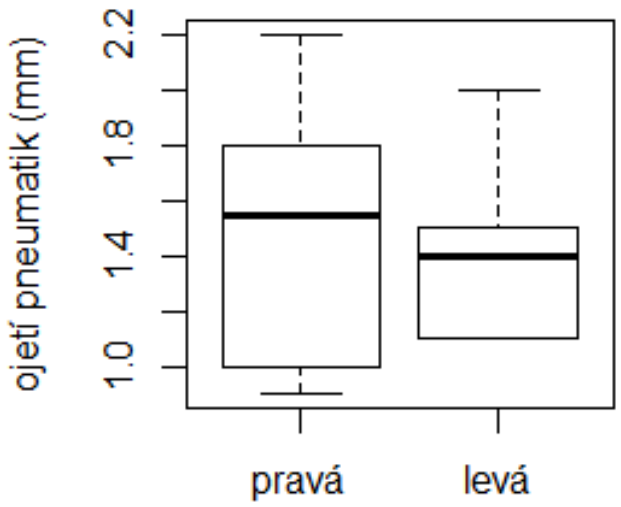
Ve skutečnosti **nechceme sledovat shodu dat s regresní přímkou**, ale shodu dat s osou prvního a třetího kvadrantu.

Jak empiricky posoudit shodu párových dat?

U 6aut bylo zjištěno ojetí předních pneumatik (mm).

Ojetí předních pneumatik (mm)						
Pravá pneumatika	1,8	1,0	2,2	0,9	1,5	1,6
Levá pneumatika	1,5	1,1	2,0	1,1	1,4	1,4

Ojíždějí se pravá a levá pneumatika stejně?



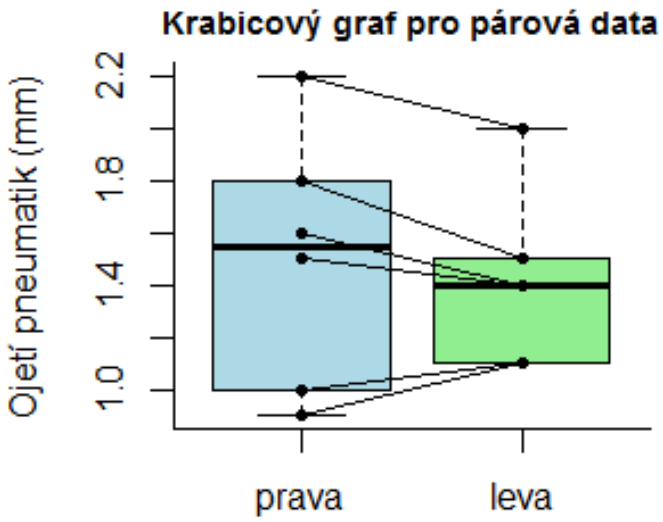
Na grafu nejsou vidět jednotlivé datové páry…

Jak empiricky posoudit shodu párových dat?

U 6aut bylo zjištěno ojetí předních pneumatik (mm).

Ojetí předních pneumatik (mm)						
Pravá pneumatika	1,8	1,0	2,2	0,9	1,5	1,6
Levá pneumatika	1,5	1,1	2,0	1,1	1,4	1,4

Ojíždějí se pravá a levá pneumatika stejně?

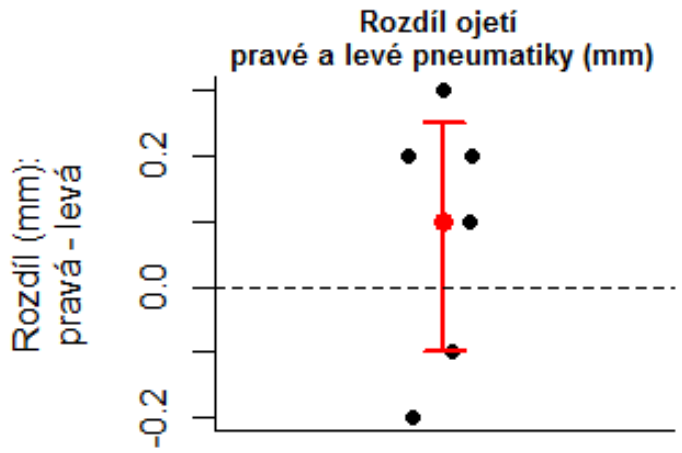


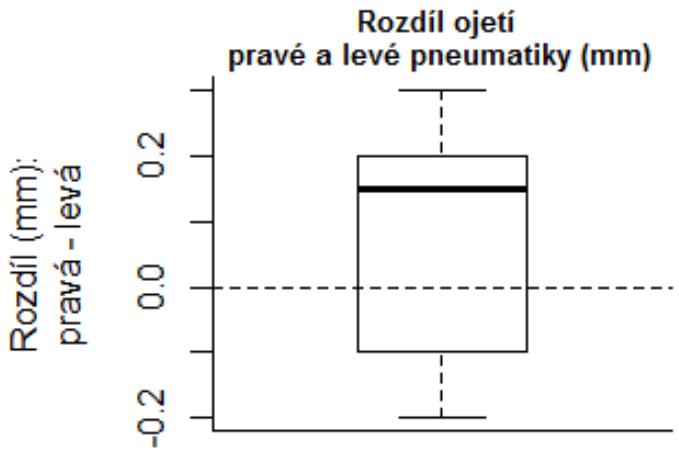
Jak empiricky posoudit shodu párových dat?

U 6aut bylo zjištěno ojetí předních pneumatik (mm).

Ojetí předních pneumatik (mm)						
Pravá pneumatika	1,8	1,0	2,2	0,9	1,5	1,6
Levá pneumatika	1,5	1,1	2,0	1,1	1,4	1,4

Ojíždějí se pravá a levá pneumatika stejně?





- Červená chybová úsečka označuje průměr a 95% IO stř. hodnoty
- Čerchovaná čára označuje nulový efekt rozdílů

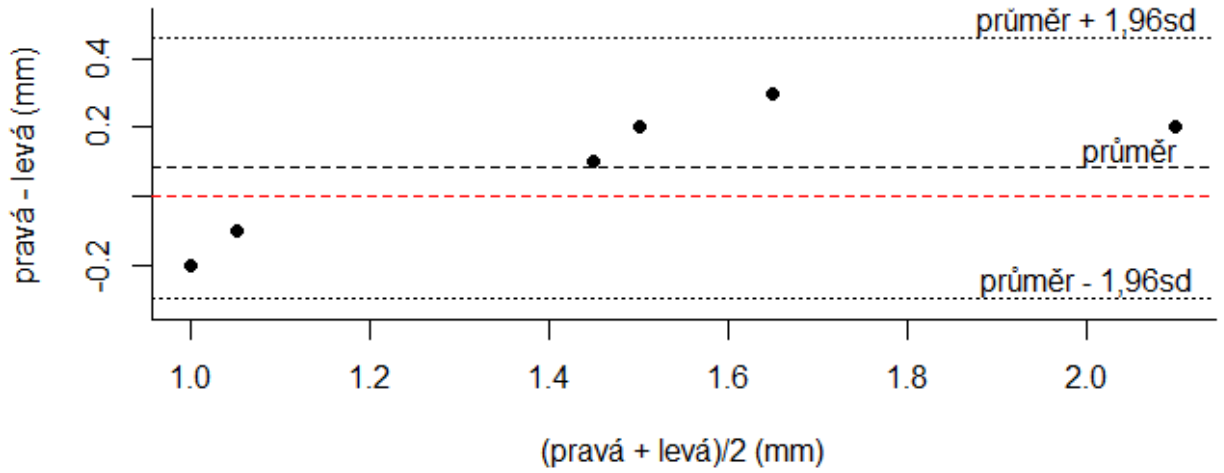
Jak empiricky posoudit shodu párových dat?

U 6aut bylo zjištěno ojetí předních pneumatik (mm).

Ojetí předních pneumatik (mm)						
Pravá pneumatika	1,8	1,0	2,2	0,9	1,5	1,6
Levá pneumatika	1,5	1,1	2,0	1,1	1,4	1,4

Ojíždějí se pravá a levá pneumatika stejně?

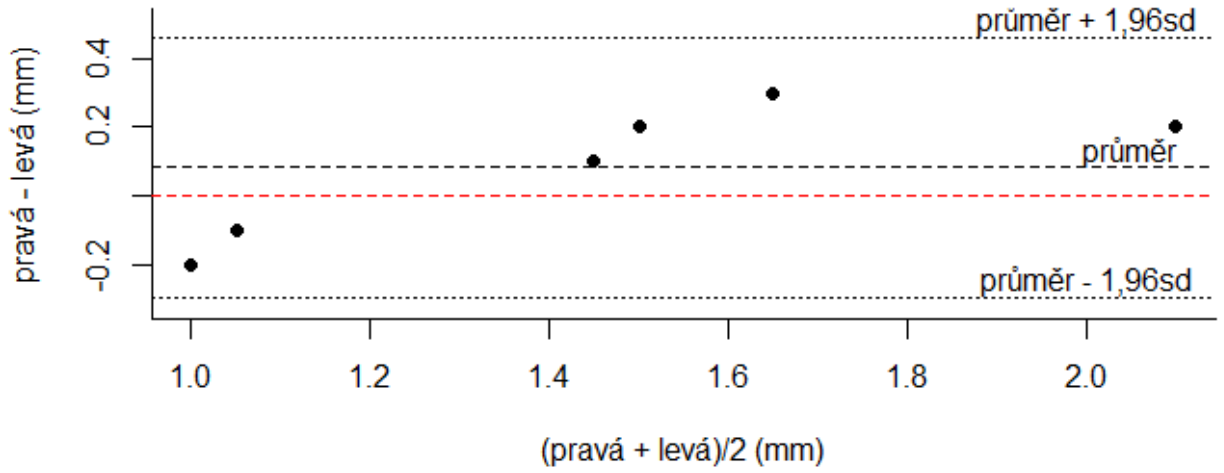
Blandův – Altmanův graf



Limity shody (limits of agreements, LoA) – vymezují interval, v němž lze očekávat naměřené hodnoty rozdílů párových měření v případě, že rozdíly mají normální rozdělení (nutno ověřit!!!)

Blandův – Altmanův graf

Graf pro posouzení shody párových dat



<u>Limity shody (limits of agreements, LoA)</u> – vymezují interval, v němž lze očekávat naměřené hodnoty rozdílů párových měření v případě, že rozdíly mají normální rozdělení (nutno ověřit!!!)

Odvození limit shody:

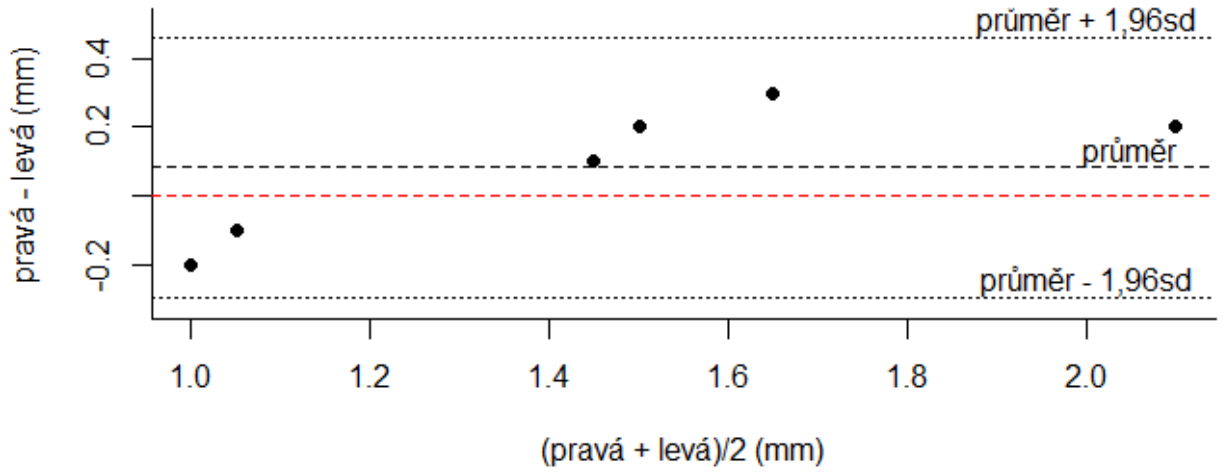
D … diference párových měření

$$D \sim N(\mu, \sigma^2) \quad \Rightarrow \quad \frac{D-\mu}{\sigma} \sim N(0,1)$$

$$P\left(z_{0,025} < \frac{D-\mu}{\sigma} < z_{0,975}\right) = 0,95$$

Blandův – Altmanův graf

Graf pro posouzení shody párových dat



Limity shody (limits of agreements, LoA) – vymezují interval, v němž lze očekávat naměřené hodnoty rozdílů párových měření v případě, že rozdíly mají normální rozdělení (nutno ověřit!!!)

Odvození limit shody:

D … diference párových měření

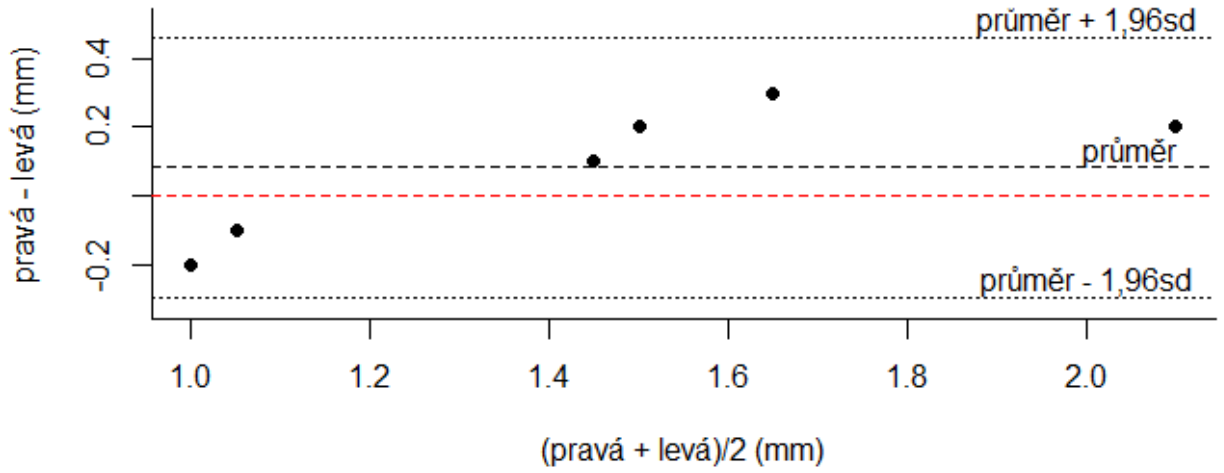
$$D \sim N(\mu, \sigma^2) \quad \Rightarrow \quad \frac{D-\mu}{\sigma} \sim N(0,1)$$

$$P\left(z_{0,025} < \frac{D-\mu}{\sigma} < z_{0,975}\right) = 0{,}95$$

$$P\left(-1{,}96 < \frac{D-\mu}{\sigma} < 1{,}96\right) = 0{,}95$$

Blandův – Altmanův graf

Graf pro posouzení shody párových dat



Limity shody (limits of agreements, LoA) – vymezují interval, v němž lze očekávat naměřené hodnoty rozdílů párových měření v případě, že rozdíly mají normální rozdělení (nutno ověřit!!!)

Odvození limit shody:

D … diference párových měření

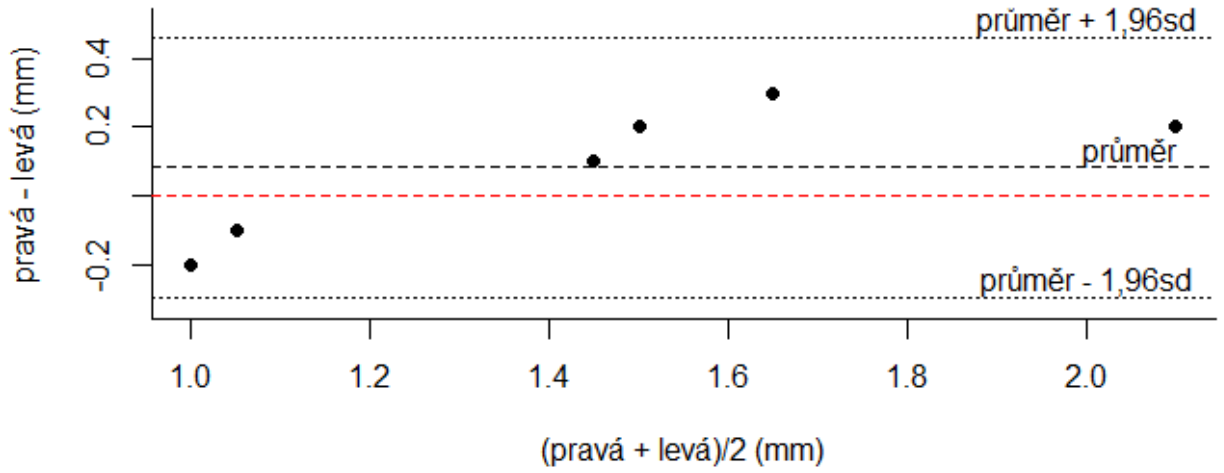
$$D \sim N(\mu, \sigma^2) \quad \Rightarrow \quad \frac{D-\mu}{\sigma} \sim N(0,1)$$

$$P\left(-1{,}96 < \frac{D-\mu}{\sigma} < 1{,}96\right) = 0{,}95$$

$$P(\mu - 1{,}96\sigma < D < \mu + 1{,}96\sigma) = 0{,}95$$

Blandův – Altmanův graf

Graf pro posouzení shody párových dat

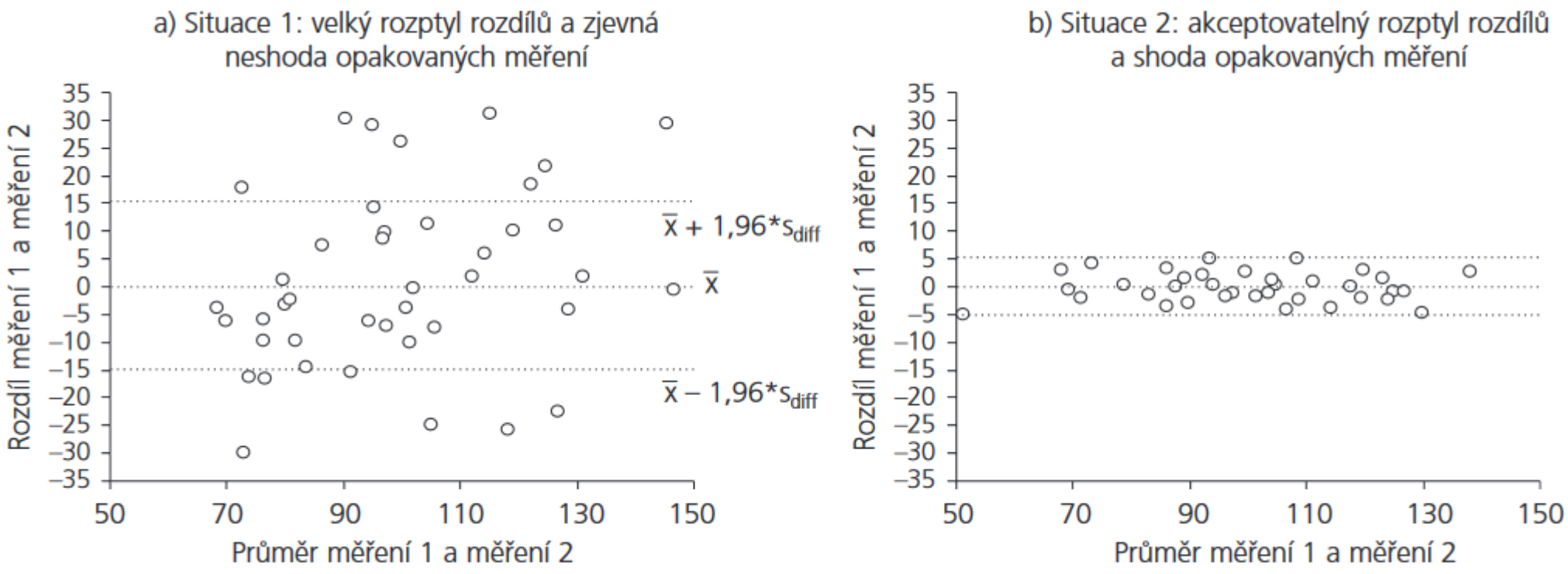


Limity shody (limits of agreements, LoA) – vymezují interval, v němž lze očekávat naměřené hodnoty rozdílů párových měření v případě, že rozdíly mají normální rozdělení (nutno ověřit!!!)

- Pokud jsou limity shody v praxi akceptovatelné jako hranice přijatelného rozdílu opakovaných měření, pak je lze využít jako míru reprodukovatelnosti nebo opakovatelnosti.
- Pokud diference opakovaných měření dané hranice překračují, nelze měření označit za reprodukovatelná (opakovatelná).
- Jako alternativní míra posouzení opakovatelnosti se používá koeficient opakovatelnosti (coefficient of repeatibility, CR).

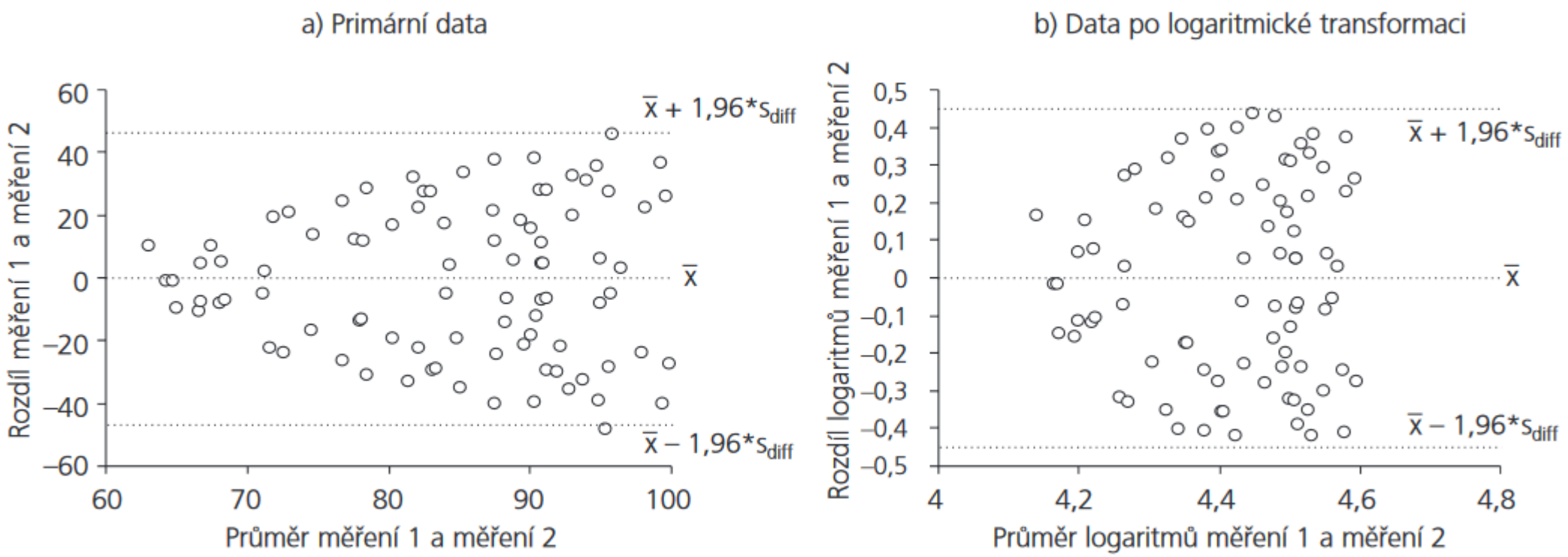
$$CR = 1{,}96s$$

Blandův – Altmanův graf



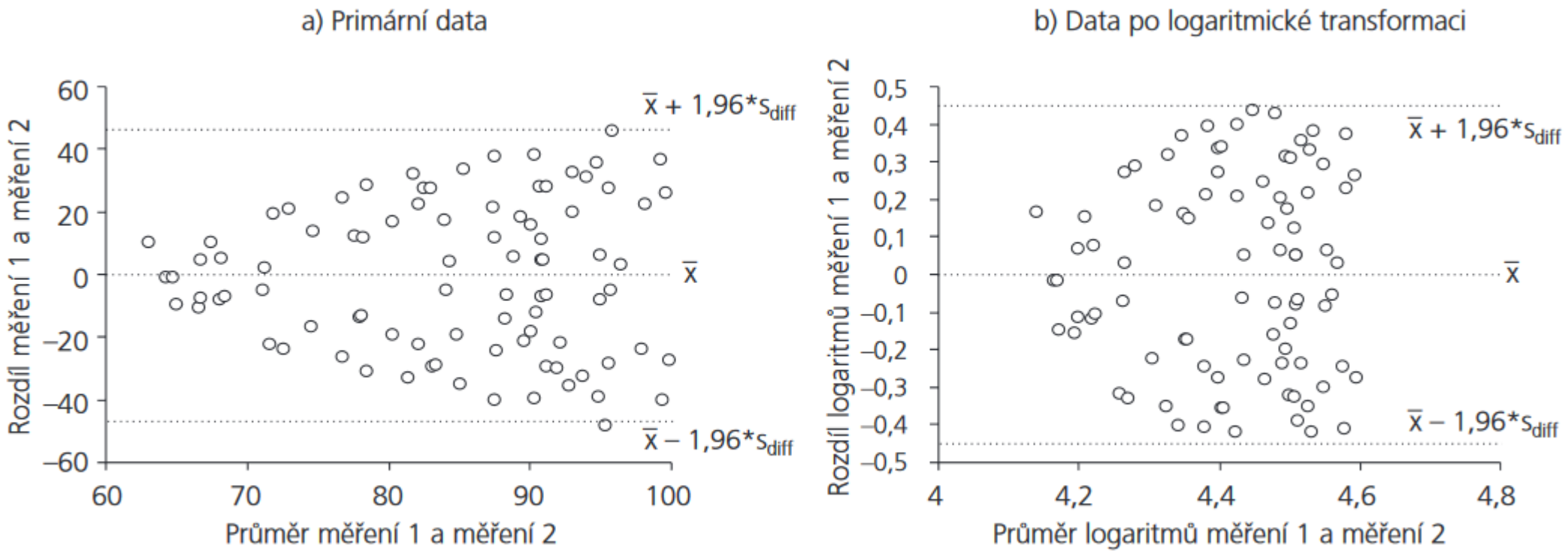
Zdroj: Dušek, L., Pavlík, T., Koptíková, J.: Analýza dat v neurologii VII. Reprodukovatelnost a opakovatelnost měření u spojitých dat. Česká a slovenská neurologie a neurochirurgie 2008; 71(1): 106-109.

Blandův – Altmanův graf



Zdroj: Dušek, L., Pavlík, T., Koptíková, J.: Analýza dat v neurologii VII. Reprodukovatelnost a opakovatelnost měření u spojitých dat. Česká a slovenská neurologie a neurochirurgie 2008; 71(1): 106-109.

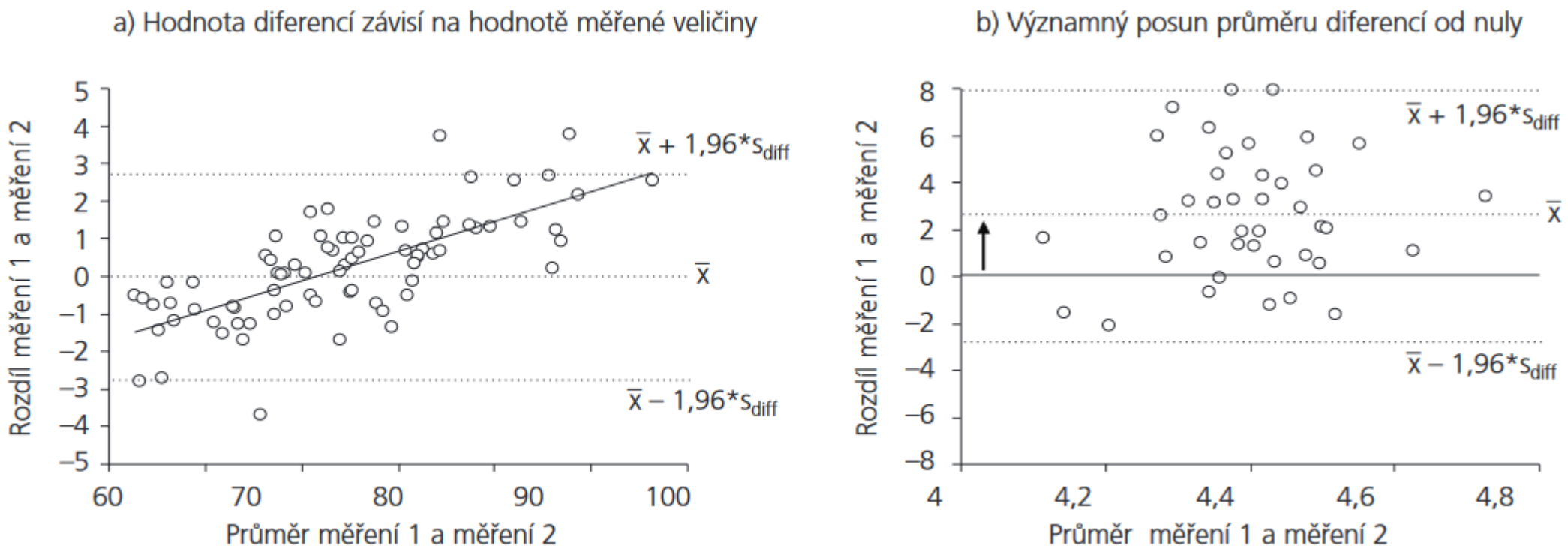
Blandův – Altmanův graf



Heteroskedasticita diferencí

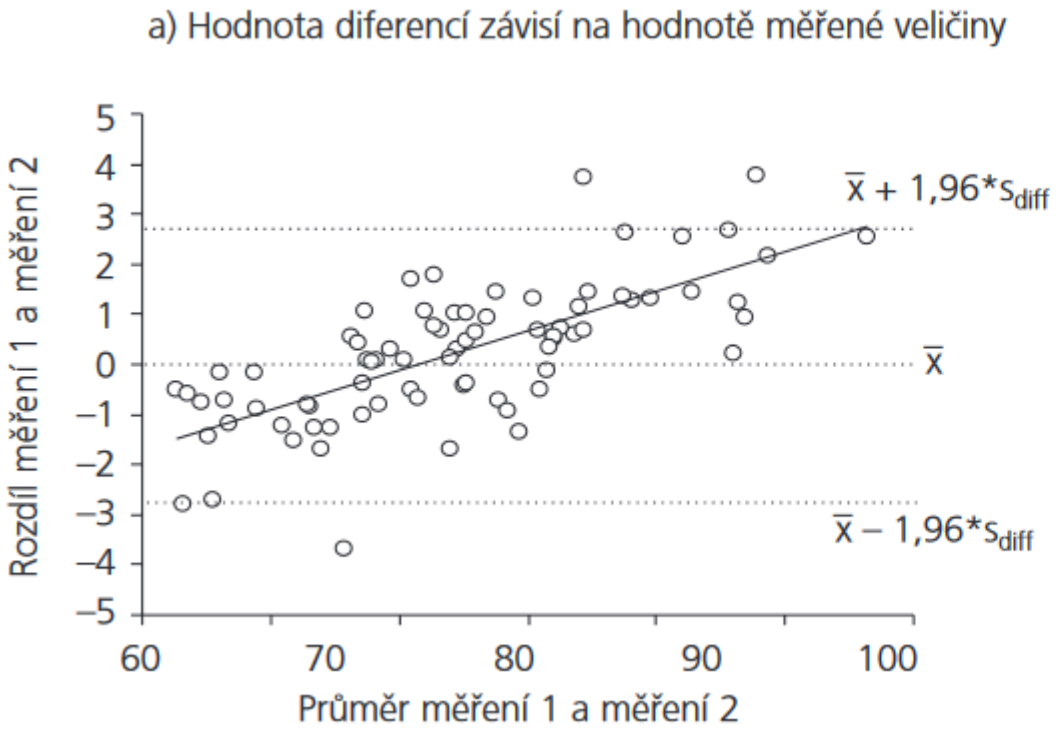
(rozptyl diferencí měření 1 a 2 se mění v závislosti na průměru měření, není odhalitelné pomocí t-testu)

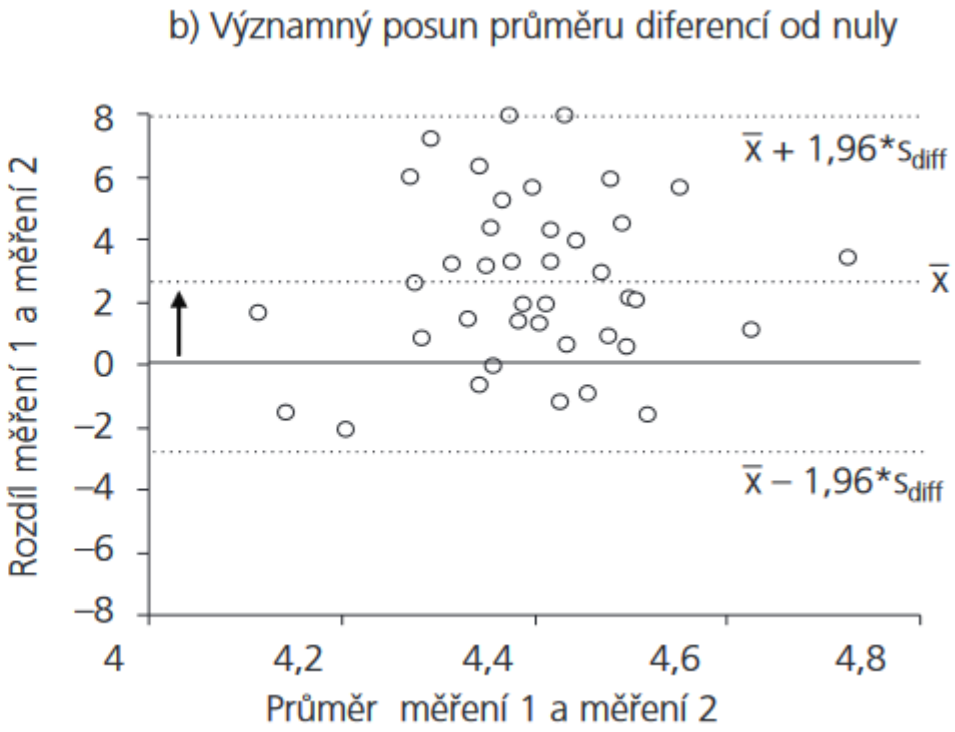
Blandův – Altmanův graf



Zdroj: Dušek, L., Pavlík, T., Koptíková, J.: Analýza dat v neurologii VII. Reprodukovatelnost a opakovatelnost měření u spojitých dat. Česká a slovenská neurologie a neurochirurgie 2008; 71(1): 106-109.

Blandův – Altmanův graf





Proporcionální chyba
(velikost odchylky měření 1 a 2 se mění v závislosti na průměru měření, není odhalitelná pomocí t-testu)

Vychýlení (angl. bias)
(detekovatelné t-testem)

Blandův – Altmanův graf

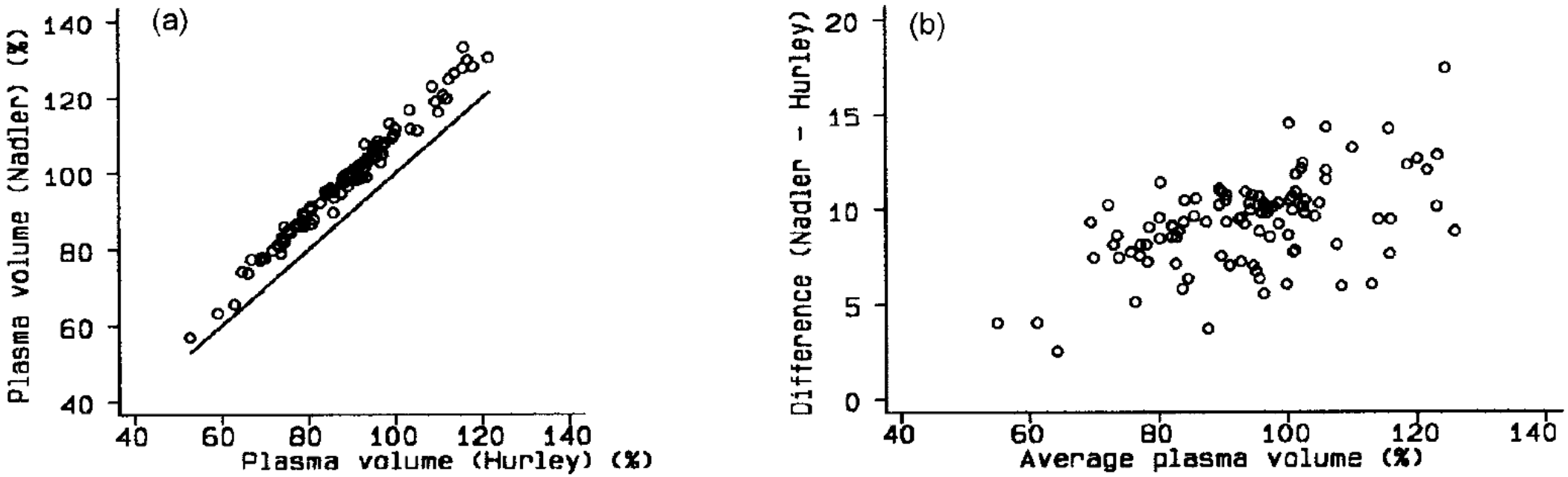


Figure 4 (a) Measurements of plasma volume as listed in Table 2; (b) plot of differences versus average with 95% limits of agreement

*Zdroj: BLAND, J Martin a Douglas G ALTMAN. Measuring agreement in method comparison studies. Statistical Methods in Medical Research [online]. 2016, **8**(2), 135-160 [cit. 2018-04-16]. DOI: 10.1177/096228029900800204. ISSN 0962-2802. Dostupné z: http://journals.sagepub.com/doi/10.1177/096228029900800204*

Blandův – Altmanův graf

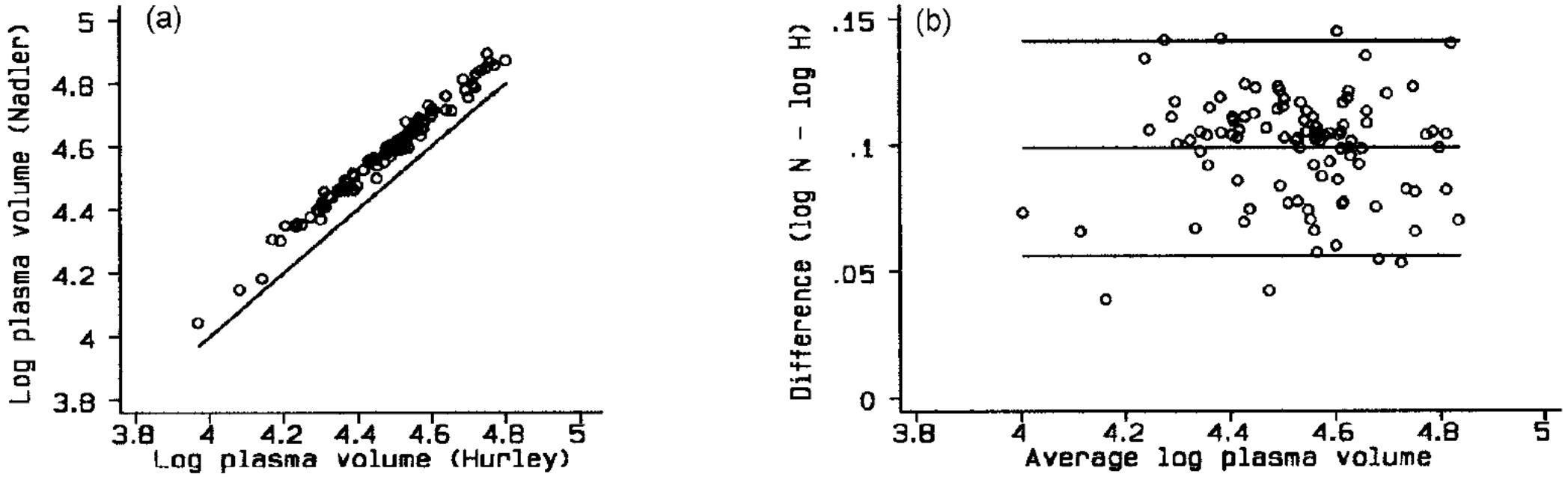


Figure 5 (a) Measurements of plasma volume after log transformation; (b) Difference between plasma volume measurements plotted against their average after $\log_e$ transformation with 95% limits of agreement

*Zdroj: BLAND, J Martin a Douglas G ALTMAN. Measuring agreement in method comparison studies. Statistical Methods in Medical Research [online]. 2016, **8**(2), 135-160 [cit. 2018-04-16]. DOI: 10.1177/096228029900800204. ISSN 0962-2802. Dostupné z: http://journals.sagepub.com/doi/10.1177/096228029900800204*

Blandův – Altmanův graf

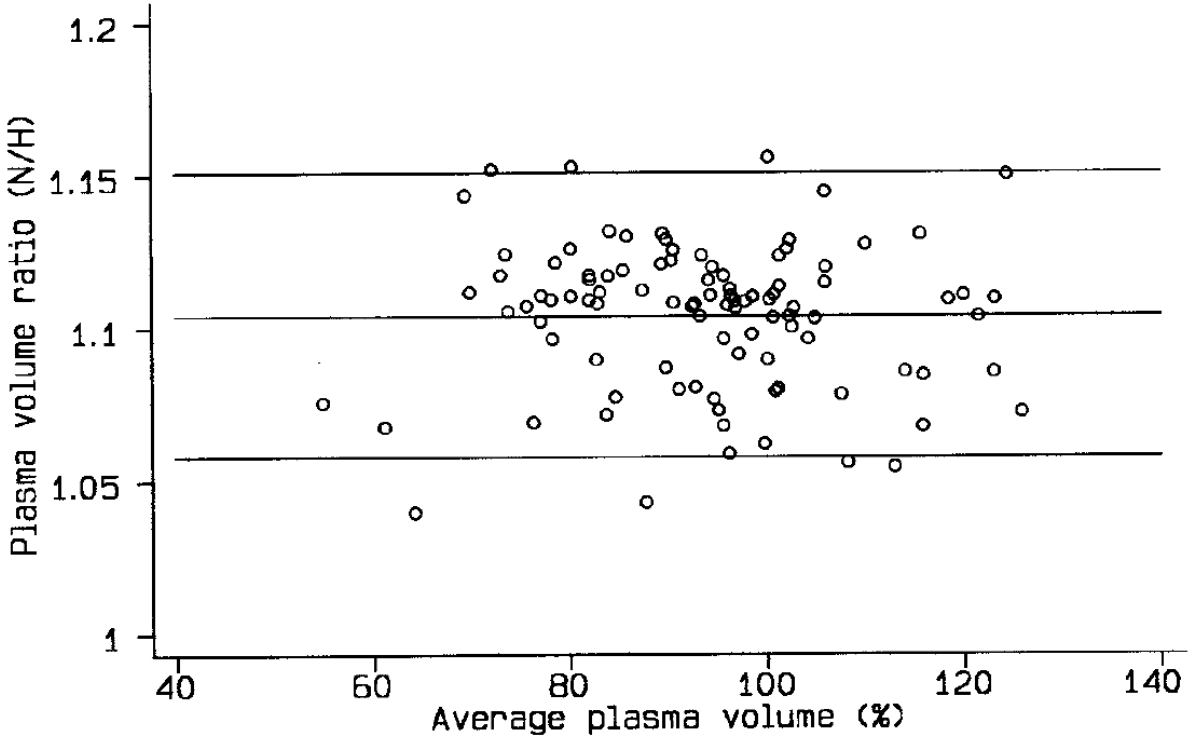


Figure 6 Ratio of plasma volume measurements plotted against their average

*Zdroj: BLAND, J Martin a Douglas G ALTMAN. Measuring agreement in method comparison studies. Statistical Methods in Medical Research [online]. 2016, **8**(2), 135-160 [cit. 2018-04-16]. DOI: 10.1177/096228029900800204. ISSN 0962-2802. Dostupné z: http://journals.sagepub.com/doi/10.1177/096228029900800204*

Blandův – Altmanův graf

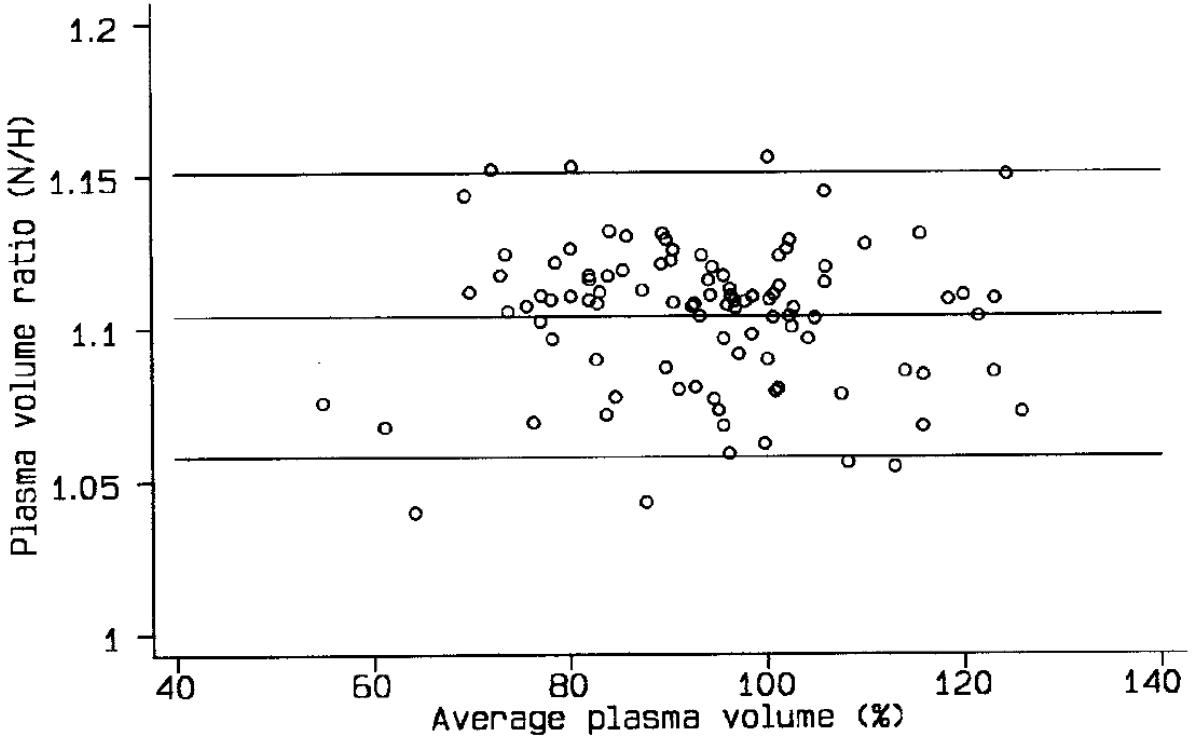


Figure 6 Ratio of plasma volume measurements plotted against their average

Jak najít limity shody pro poměr měření?

Blandův – Altmanův graf

Jak najít limity shody pro poměr měření?

Nechť jsou výsledkem zjišťování dvojice náhodných veličin $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$, které tvoří páry závislých pozorování (jde o veličiny zjišťované na stejné statistické jednotce).

Definujme soubor rozdílů (diferencí) přirozených logaritmů jednotlivých měření

$$\boldsymbol{D^{LN}} = \left(D_1^{LN}, D_2^{LN}, \dots, D_n^{LN}\right), \text{ kde } D_i^{LN} = \ln X_i - \ln Y_i.$$

Nechť $D_i^{LN} \sim N\left(\mu_D; \sigma_D^2\right)$, tj. $\frac{D_i^{LN} - \mu_D}{\sigma_D} \sim N(0; 1)$. Pak

$$P\left(z_{0,025} < \frac{D_i^{LN} - \mu_D}{\sigma_D} < z_{0,975}\right) = 0{,}95.$$

$$P\left(-1{,}96 < \frac{D_i^{LN} - \mu_D}{\sigma_D} < 1{,}96\right) = 0{,}95$$

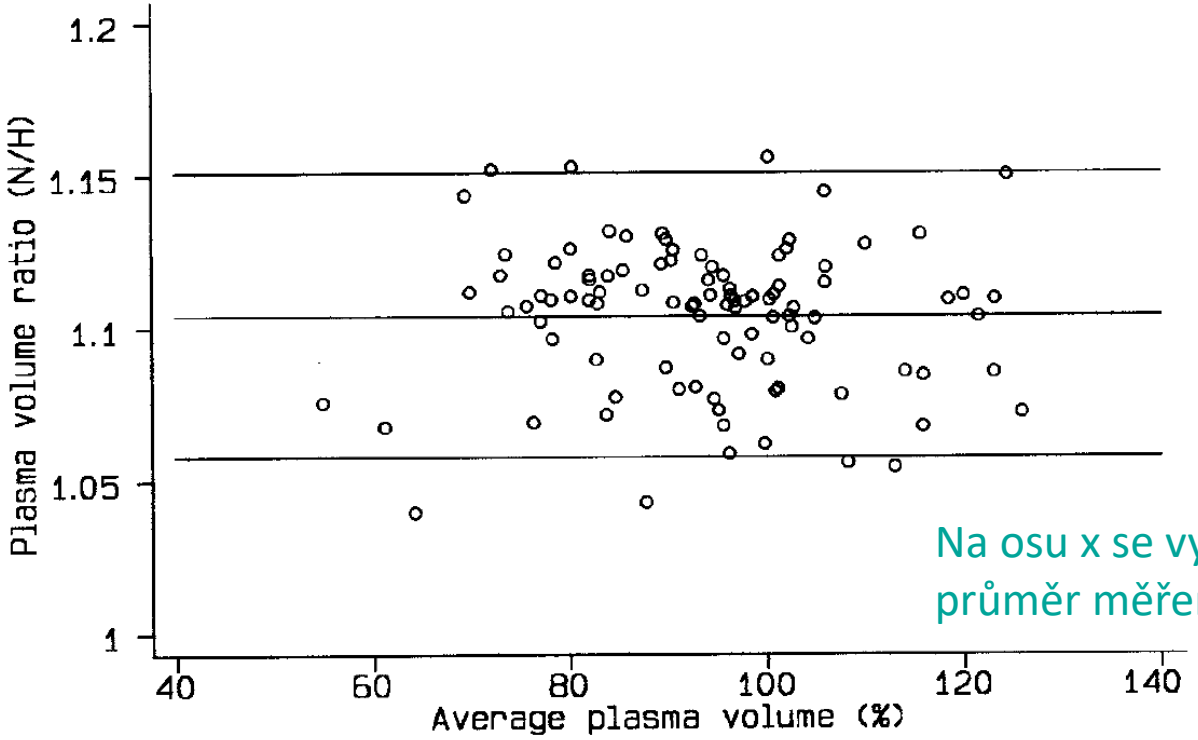
$$P\left(\mu_D - 1{,}96\sigma_D < D_i^{LN} < \mu_D + 1{,}96\sigma_D\right) = 0{,}95$$

$$P(\mu_D - 1{,}96\sigma_D < \ln X_i - \ln Y_i < \mu_D + 1{,}96\sigma_D) = 0{,}95$$

$$P\left(\mu_D - 1{,}96\sigma_D < \ln \frac{X_i}{Y_i} < \mu_D + 1{,}96\sigma_D\right) = 0{,}95$$

$$\boxed{P\left(e^{\mu_D - 1{,}96\sigma_D} < \frac{X_i}{Y_i} < e^{\mu_D + 1{,}96\sigma_D}\right) = 0{,}95}$$

Blandův – Altmanův graf



Na osu x se vynáší průměr měření.

Figure 6 Ratio of plasma volume measurements plotted against their average

Blandův – Altmanův graf

Poznámky:

- Vhodné používat společně s korelačním grafem posuzujícím shodu párových měření s přímkou $p\colon\ y = x$.
- Limity shody uvádět pouze v případě, že předpokládáme, že diference mají normální rozdělení.
- Vhodné ověřit, zda jsou limity shody v praxi akceptovatelné jako hranice přijatelného rozdílu (stanovuje expert v dané oblasti výzkumu).
- V případě, že se v párových datech projeví proporcionální chyba, pokusíme se ji ve vizualizaci eliminovat pomocí logaritmické transformace (a zároveň v původních datech hledáme její zdroj).
- Blandův – Altmannův graf poskytuje větší možnosti využití (např. doplnění intervalových odhadů pro střední hodnotu a limity shody, stanovení limit shody jako funkce průměru diferencí, …)

Zdroj:

*BLAND, J Martin a Douglas G ALTMAN. Measuring agreement in method comparison studies. Statistical Methods in Medical Research [online]. 2016, **8**(2), 135-160 [cit. 2018-04-16]. DOI: 10.1177/096228029900800204. ISSN 0962-2802. Dostupné z: http://journals.sagepub.com/doi/10.1177/096228029900800204*

Děkuji za pozornost!

martina.litschmannova@vsb.cz

VŠB TECHNICKÁ UNIVERZITA OSTRAVA | FAKULTA ELEKTROTECHNIKY A INFORMATIKY | KATEDRA APLIKOVANÉ MATEMATIKY