

## 5 Test o střední hodnotě při neznámém rozptylu

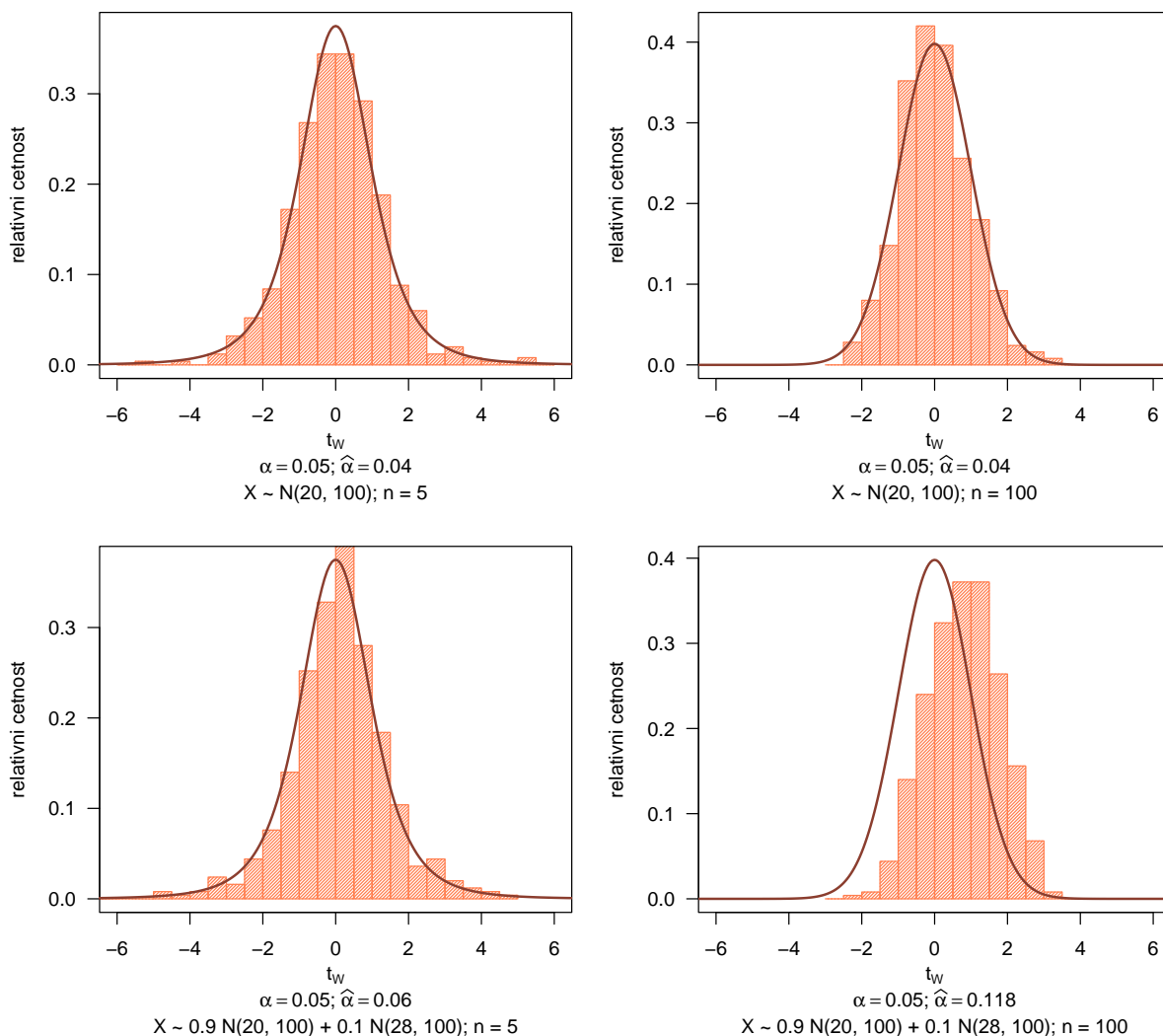
### Příklad 5.1. MC experiment: Aktuální vs. nominální hladina významnosti - Testovací statistika

Nechť

(a)  $X \sim N(20, 100)$ ;

(b)  $X \sim pN(20, 100) + (1 - p)N(28, 100)$ , kde  $p = 0.9$ , tedy jde o směs dvou normálních rozdělení  $X \sim N(20, 100)$  a  $X \sim N(28, 100)$  v poměru 9 : 1.

Pro obě části (a) i (b) vygenerujte  $M = 500$  náhodných výběrů s rozsahem  $n = 5$ , resp.  $n = 100$  a vypočítejte hodnotu testovací statistiky  $T_W$  pro test nulové hypotézy  $H_0 : \mu = \mu_0 = 20$  oproti  $H_1 : \mu \neq 20$ , když  $\sigma^2$  neznáme. Hodnoty testovacích statistik  $T_W$  zanešte do histogramu. Vždy spočítejte, kolik testovacích statistik  $T_W$  spadá do kritického oboru  $W$ . Toto číslo podělené hodnotou  $M$  představuje aktuální hladinu významnosti  $\hat{\alpha}$ . Porovnejte tuto hodnotu s nominální hladinou významnosti  $\alpha$  a v každé ze čtyř situací určete, zda je test konzervativní či liberální.



Obrázek 1: Rozdělení Waldovy testovací statistiky  $U_W$  pro test o střední hodnotě  $\mu$  při neznámém rozptylu  $\sigma^2$ ; porovnání aktuální a nominální hladiny významnosti  $\alpha$

**Příklad 5.2. Rozdělení testovacích statistik  $t_W$  a  $t_W^2$** 

Nechť náhodný výběr  $X$  pochází z normálního rozdělení, t.j.  $X \sim N(\mu, \sigma^2)$ , kde  $\mu = 150$  a  $\sigma^2 = 30^2$ . Rozsah náhodného výběru zvolte (a)  $n = 10$ ; (b)  $n = 100$ . Pomocí simulační studie ( $M = 1000$ ) ověřte, že pro test o střední hodnotě  $\mu$  když  $\sigma^2$  neznáme, platí za platnosti  $H_0$  následující:

$$1. T_W = \frac{(\bar{x} - \mu_0)}{s/\sqrt{n}} \sim t_{n-1};$$

$$2. T_W^2 = \frac{(\bar{x} - \mu_0)^2}{s^2/n} \sim F_{1, n-1}.$$

Vygenerujte  $M = 1000$  pseudonáhodných výběrů o rozsahu  $n = 5$ . Pro každý náhodný výběr vypočítejte hodnoty realizací testovacích statistik  $T_W$  a  $T_W^2$ . Vytvořte histogramy těchto testovacích statistik a superponujte je křivkami příslušného rozdělení. Dále vytvořte animaci, pomocí které vizualizujete přesnost rozdělení testovacích statistik  $T_W$ , resp.  $T_W^2$  při různých rozsazích náhodného výběru. Animaci spravte pro měnící se  $n \in \{5, 10, 15, 20, 25, 30, 35, 40, 50, 100, 250, 500\}$

Obrázek 2: Studentovo rozdělení testovací statistiky  $T_W$  a Fisherovo rozdělení testovací statistiky  $T_W^2$  pro test o střední hodnotě  $\mu$  při neznámém rozptylu

**Příklad 5.3. Rozdělení testovacích statistik  $U_W$ ,  $U_S$  a  $U_{LR}$** 

Nechť náhodný výběr  $X$  pochází z normálního rozdělení, t.j.  $X \sim N(\mu, \sigma^2)$ , kde  $\mu = 150$  a  $\sigma^2 = 30^2$ . Rozsah náhodného výběru zvolte (a)  $n = 10$ ; (b)  $n = 100$ . Pomocí simulační studie ( $M = 1000$ ) ověřte, že pro test o střední hodnotě  $\mu$  když  $\sigma^2$  neznáme platí za platnosti  $H_0$  následující:

1.  $U_W = t_W^2 \frac{n}{n-1} \sim \chi_1^2$ ;

2.  $U_S = \frac{\frac{nt_W^2}{n-1}}{1 + \frac{t_W^2}{n-1}} \sim \chi_1^2$ ;

3.  $U_{LR} = n \ln \left( 1 + \frac{t_W^2}{n-1} \right) \sim \chi_1^2$ .

Vygenerujte  $M = 1000$  pseudonáhodných výběrů o rozsahu  $n = 5$ . Pro každý náhodný výběr vypočítejte hodnoty realizací testovacích statistik  $U_W$ ,  $U_S$  a  $U_{LR}$ . Vytvořte histogramy těchto testovacích statistik a superponujte je křivkami  $\chi^2$  rozdělení. Dále vytvořte animaci, pomocí které vizualizujete přesnost  $\chi^2$  rozdělení testovacích statistik  $U_W$ ,  $U_S$  a  $U_{LR}$  při různých rozsazích náhodného výběru. Animaci spravte pro měnící se  $n \in \{5, 10, 15, 20, 25, 30, 35, 40, 50, 100, 250, 500\}$ .

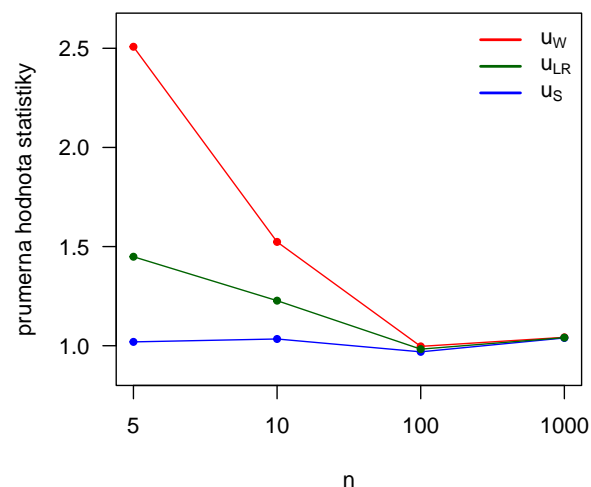
Obrázek 3:  $\chi^2$  rozdělení testovacích statistik  $U_W$ ,  $U_S$  a  $U_{LR}$  pro test o střední hodnotě  $\mu$  při neznámém rozptylu

**Příklad 5.4. Porovnání testovacích statistik  $U_W$ ,  $U_S$  a  $U_{LR}$** 

Pomocí simulační studie porovnejte tvary tří testovacích statistik (1)  $U_W = \frac{n}{n-1}t_W^2$ , (2)  $U_S = \frac{nt_W^2}{n-1+t_W^2}$  a (3)  $U_{LR} = n \ln \left(1 + \frac{t_W^2}{n-1}\right)$ , kde  $t_W^2 = \frac{(\bar{x} - \mu_0)^2}{S_{n-1}^2/n}$  a ukažte, že platí vztah  $U_S < U_{LR} < U_W$ .

Vygenerujte  $M = 1000$  náhodných výběrů  $X \sim N(0, 3)$  a pro každý výběr vypočítejte hodnoty testovacích statistik  $U_W$ ,  $U_S$  a  $U_{LR}$  pro test nulové hypotézy  $H_0 : \mu = \mu_0 = 0$  oproti alternativní hypotéze  $H_1 : \mu \neq \mu_0$ . Pro každou testovací statistiku následně najdete její jádrový odhad a křivky jádrového odhadu pro pevně zvolené  $n$  vykreslete do jednoho grafu. Vytvořte animaci zobrazující tvary křivek jádrových odhadů testovacích statistik při rostoucím rozsahu náhodných výběrů  $n$ . Rozsah výběrů  $n$  volte postupně  $n = 2, 3, \dots, 9, 10, 15, 20, 25, 30, 35, 40, 50, 100, 250, 500$ . Dále vypočítejte průměrné hodnoty testovacích statistik  $U_W$ ,  $U_S$  a  $U_{LR}$  pro  $n = 5, 10, 100$  a  $1000$  a zanešte je do jednoho grafu.

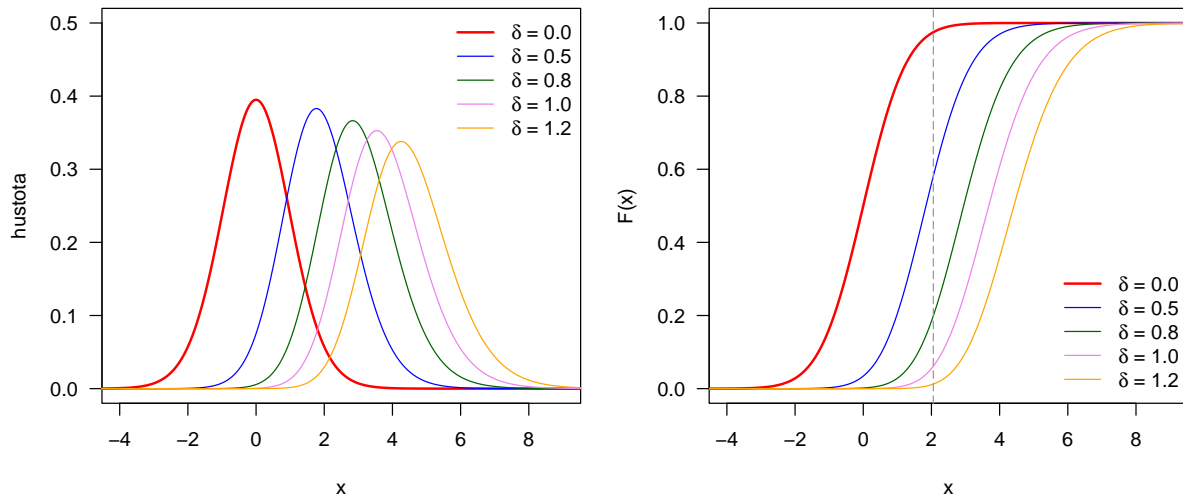
Obrázek 4: Porovnání tvarů testovacích statistik  $U_W$ ,  $U_S$  a  $U_{LR}$  testu o střední hodnotě při neznámém rozptylu  $\sigma^2$



Obrázek 5: Porovnání průměrů testovacích statistik  $U_W$ ,  $U_S$  a  $U_{LR}$  pro  $n = 5, 10, 100, 1000$

**Příklad 5.5. Hustota a distribuční funkce centrálního a necentrálního  $t$ -rozdělení**

Nakreslete (a) hustotu; (b) distribuční funkci jednoho centrálního a čtyř necentrálních  $t$ -rozdělení  $t_{n-1,\lambda}$  ( $\delta = \mu - \mu_0$  a  $\lambda = \delta/(\sigma/\sqrt{n})$ ) do jednoho obrázku tak, aby byly odlišitelné barvou nebo typem čáry. Zvolte  $\mu_0 = 0$ ,  $\sigma = 1.4$ ,  $n = 26$ ,  $\delta = 0, 0.5, 0.8, 1$  a  $1.2$ .



Obrázek 6: Graf (a) hustoty; (b) distribuční funkce (i) centrálního  $t$ -rozdělení; (ii) čtyř necentrálních  $t$ -rozdělení s parametrem necentrality  $\delta$

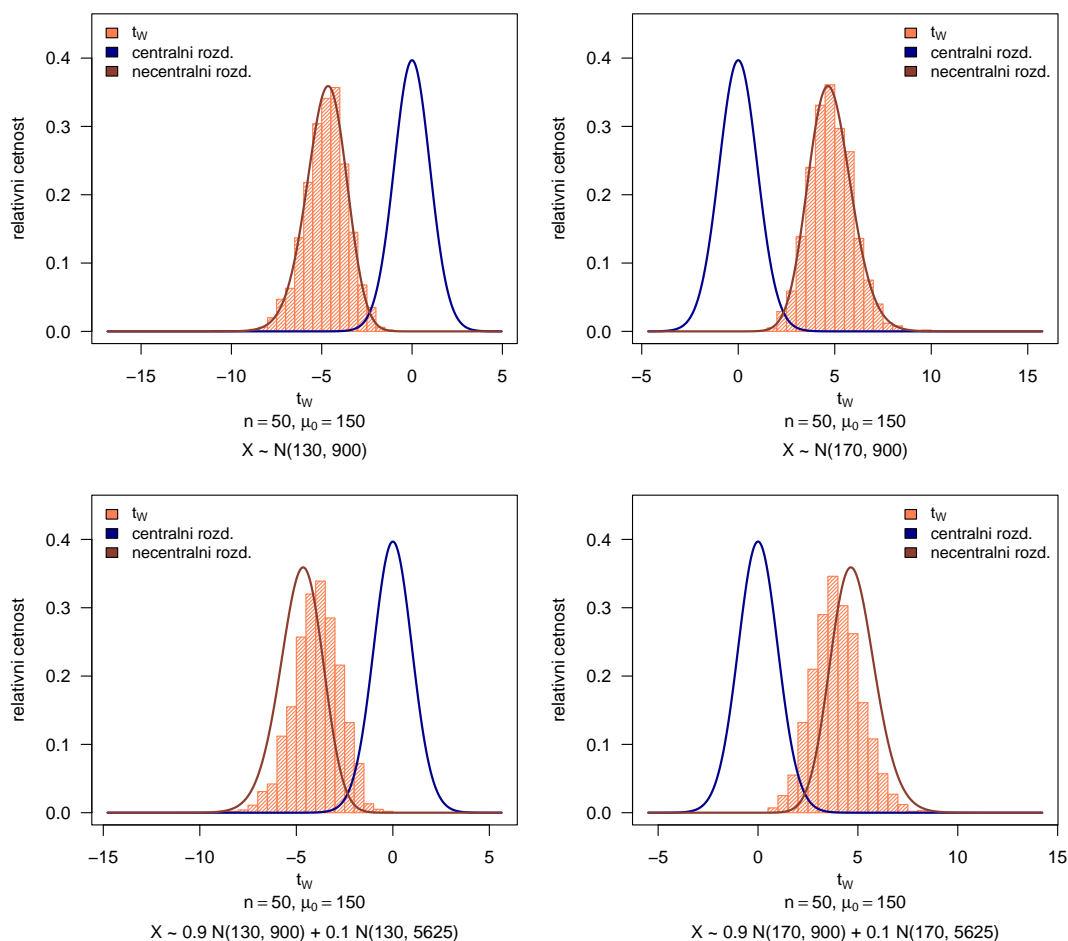
**Příklad 5.6. Rozdělení testovací statistiky pro test o střední hodnotě  $\mu$ , když  $\sigma^2$  neznáme**

Nechť náhodný výběr  $X$  pochází z normálního rozdělení, t.j.  $X \sim N(\mu, \sigma^2)$ . Pomocí simulační studie porovnejte rozdělení testovací statistiky  $T_W$  pro test nulové hypotézy  $H_0: \mu = 150$  (alternativní hypotéza  $H_1: \mu \neq 150$ ), když rozptyl  $\sigma^2$  neznáme, s rozdělením testovací statistiky stanovené na základě náhodného výběru se střední hodnotou  $\mu$ . Parametry zvolte (a)  $\mu = 130$ ,  $\sigma^2 = 30^2$ ,  $n = 50$ ; (b)  $\mu = 170$ ,  $\sigma^2 = 30^2$ ,  $n = 50$ .

Nechť dále  $X$  pochází ze směsi dvou normálních rozdělení, t.j.  $X \sim [pN(\mu, 30^2) + (1-p)N(\mu, 75^2)]$ , kde  $p = 0.9$  a (c)  $\mu = 130$ ; (d)  $\mu = 170$ . Proveďte simulační studii popsanou výše také pro tento náhodný výběr.

Nasimulujte  $M$  pseudonáhodných výběrů,  $M = 1, \dots, 2000$  a pro každý vypočítejte realizaci testovací statistiky  $t_{W,\lambda}^{(m)} = \frac{\bar{x}_m - \mu_0}{s} \sqrt{n}$  pro nulovou hypotézu  $H_0: \mu = 150$  oproti  $H_1: \mu \neq 150$ . Vykreslete histogram testovacích statistik  $T_W$  a superponujte jej jednak křivkou hustoty Studentova rozdělení  $t_{n-1,\lambda}$  s parametrem nacentrality  $\lambda$  ( $\lambda = \frac{\mu - \mu_0}{\sigma} \sqrt{n}$ , kde  $\mu$  je skutečná střední hodnota (relevantní za platnosti  $H_1$ ) a  $\sigma^2$  je skutečný rozptyl) a jednak křivkou hustoty centrálního Studentova rozdělení  $t_{n-1}$ . Obě křivky potom vzájemně okometricky porovnejte.

*Poznámka:* U směsi křivka hustoty nacentrálního rozdělení nesuperponuje histogram statistik  $t_W$  dostatečně. Zamyslete se nad tím, proč, a co z toho pro nás do praxe vyplývá. Odpověď uveďte formou komentáře.



Obrázek 7: Porovnání centrálního a nacentrálního Studentova  $t$ -tozdělení s rozdělením testovací statistiky  $T_W$  testu o střední hodnotě  $\mu$  při neznámém rozptylu

**Příklad 5.7. Silofunkce testu o střední hodnotě  $\mu$  když  $\sigma^2$  neznáme**

Předpokládejme, že  $X \sim N(\mu, \sigma^2)$ , kde  $\sigma^2$  neznáme, ale předpokládáme, že  $\sigma^2 = 10^2$ , a  $n = 100$ . Necht'  $\theta = \mu$ . Na hladině významnosti  $\alpha = 0.05$  testujeme hypotézu  $H_{01} : \mu = \mu_0$  oproti  $H_{11} : \mu \neq \mu_0$  (oboustranná), kde  $\mu_0 = 150$ .

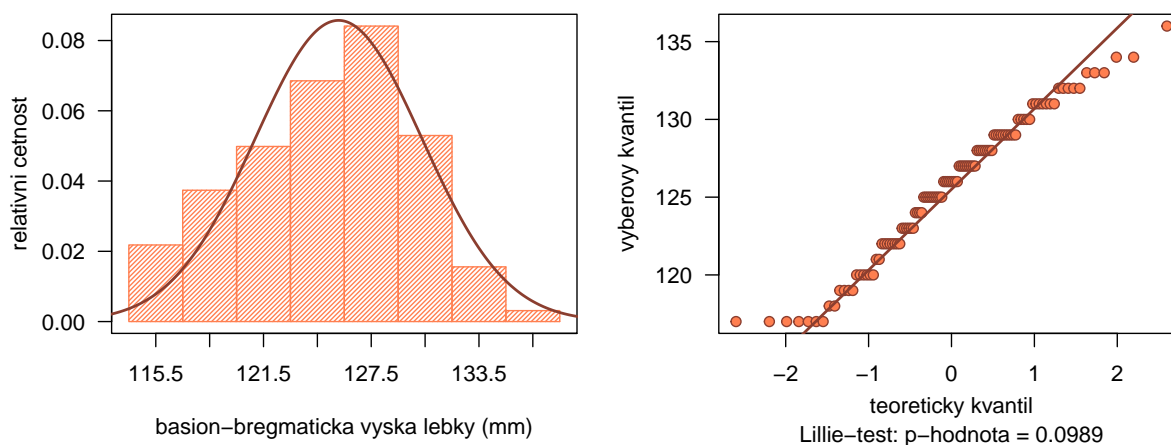
Vytvořte animaci zobrazující změnu polohy necentrálního rozdělení vzhledem k hodnotě centrálního rozdělení testovací statistiky testu o  $\mu$  když  $\sigma^2$  neznáme, spolu s barevně odlišenou oblastí kritického oboru a exaktní síly; změnu hodnoty exaktní silofunkce; při měnící se střední hodnotě náhodného výběru  $\mu = 140, 141, \dots, 146, 146.5, \dots, 153.5, 154, 155, \dots, 160$ .

Obrázek 8: Průběh síly testu o střední hodnotě  $\mu$  když  $\sigma^2$  neznáme – oboustranná alternativa

**Příklad 5.8. Test o střední hodnotě  $\mu$  když  $\sigma^2$  neznáme; praktický příklad**

Z archivních materiálů (Schmidt, 1888) máme k dispozici původní kranio-metrické údaje 215 dospělých mužů a 107 dospělých žen ze starověké egyptské populace o basion–bregmatické výšce lebky. Údaje jsou k dispozici v souboru 11-two-samples-means-skull.txt. Současně máme k dispozici hodnoty basion–bregmatické výšky ( $\bar{x}_m = 133.977$  mm;  $\bar{x}_f = 126.942$  mm) a hodnoty směrodatné odchylky ( $s_m = 5.171$  mm;  $s_f = 4.430$  mm) mužů a žen novověké egyptské populace.

Na hladině významnosti  $\alpha = 0.05$  otestujte nulovou hypotézu  $H_0 : \mu_f = 126.942$  oproti alternativní hypotéze  $H_1 : \mu_f \neq 126.942$ . Před testováním ověřte předpoklad normality. Testování proveďte pomocí (a) kritického oboru; (b) intervalu spolehlivosti; (c) p-hodnoty při použití testovacích statistik (1)  $T_W$ , (2)  $U_W$ , (3)  $U_S$ , (4)  $U_{LR}$ . Dále vykreslete grafy zobrazující věrohodnostní intervaly spolehlivosti pro test střední hodnotě získané na základě testovacích statistik  $U_W$ ,  $U_S$  a  $U_{LR}$ .

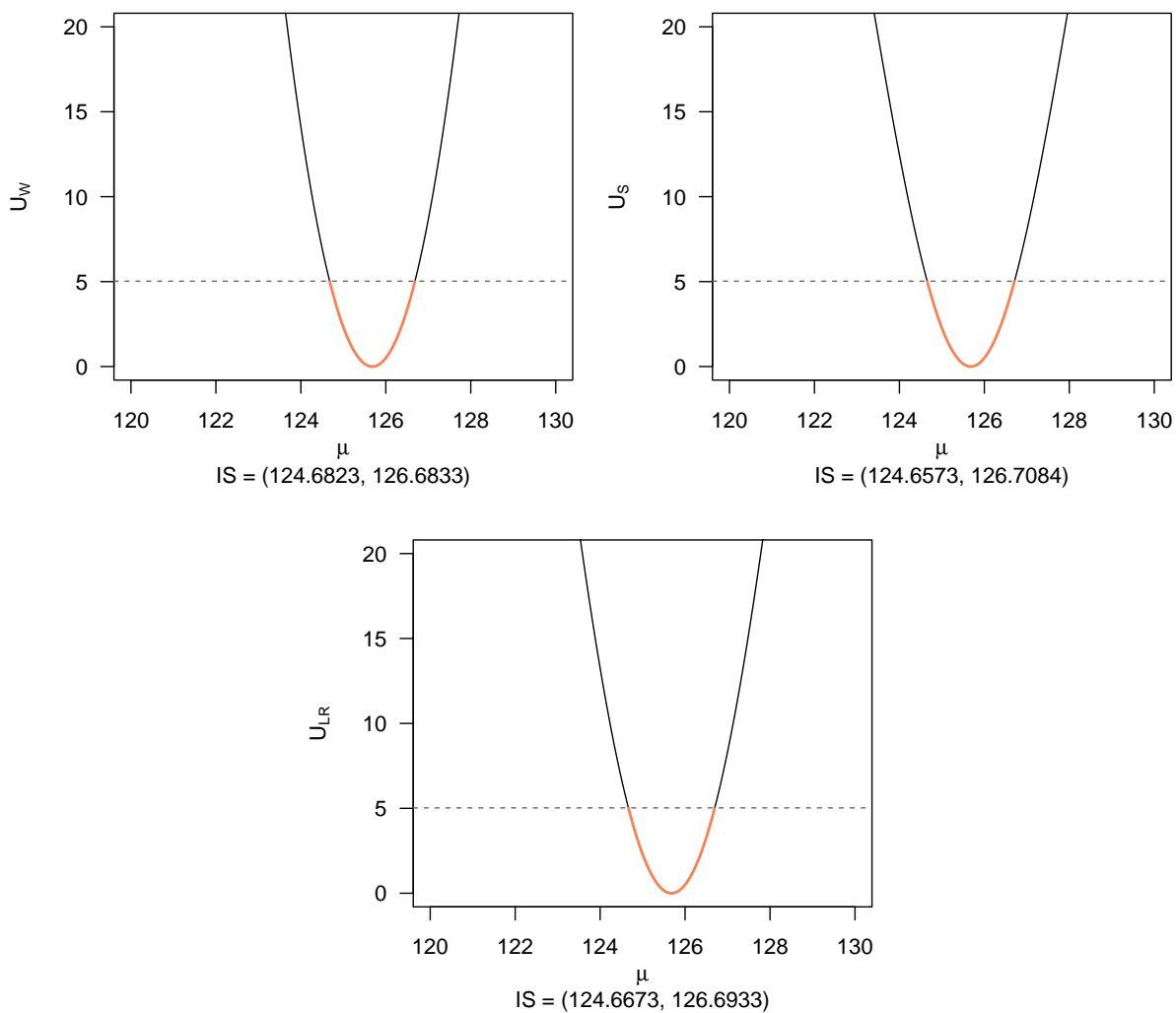


Obrázek 9: Vizualizace normality basion-bregmatické výšky lebky u žen pomocí (a) histogramu, (b) Q-Q grafu

Tabulka 1: Výsledky testu o střední hodnotě  $\mu$  při neznámém rozptylu  $\sigma^2$  při použití testovacích statistik  $T_W$ ,  $U_W$ ,  $U_S$  a  $U_{LR}$

Statistika	$\hat{\mu}$	statistika	$W_{hh}$	$W_{dh}$	$IS_{dh}$	$IS_{hh}$	p-hodnota
$T_W$	125.6822	-2.8004	-2.2737	2.2737	124.6594	126.7051	0.0061
$U_W$	125.6822	7.9163		5.0239	124.6823	126.6833	0.0049
$U_S$	125.6822	7.3710		5.0239	124.6573	126.7084	0.0066
$U_{LR}$	125.6822	7.6371		5.0239	124.6673	126.6933	0.0057





Obrázek 10: 95 % Věrohodnostní empirické intervaly spolehlivosti pro test o střední hodnotě  $\mu$  získané na základě testovacích statistik  $U_W$ ,  $U_S$  a  $U_{LR}$