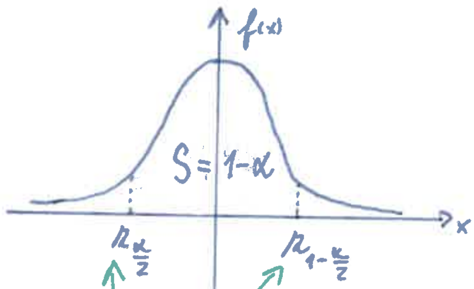
Poznámka: Pokud $X \to N(\mu, \sigma^2)$ pak $Z = \frac{\bar{X}-\mu}{\sigma}\sqrt{n} \longrightarrow N(0,1)$

Pokud n je velké $(n>30)$ pak $Z = \frac{\bar{X}-\mu}{\sigma}\sqrt{n} \xrightarrow{\cdot} N(0,1)$
(viz centrální limitní věta)

⇒ S pravděpodobností $1-\alpha$ nastane situace:



$$u_{\frac{\alpha}{2}} < Z < u_{1-\frac{\alpha}{2}}$$

$$-u_{1-\frac{\alpha}{2}} = u_{\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{\sigma}\sqrt{n} < u_{1-\frac{\alpha}{2}} \quad / \cdot \frac{\sigma}{\sqrt{n}} > 0$$

kvantily rozdělení $N(0,1)$

$u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$

$$-\frac{u_{1-\frac{\alpha}{2}} \cdot \sigma}{\sqrt{n}} < \bar{X} - \mu < \frac{u_{1-\frac{\alpha}{2}} \cdot \sigma}{\sqrt{n}} \quad / - \bar{X}$$

$$-\bar{X} - \frac{u_{1-\frac{\alpha}{2}} \cdot \sigma}{\sqrt{n}} < -\mu < -\bar{X} + \frac{u_{1-\frac{\alpha}{2}} \cdot \sigma}{\sqrt{n}} \quad / \cdot (-1)$$

$$\bar{X} + \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}} > \mu > \bar{X} - \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$

$$\Rightarrow P\left(\bar{X} - \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}} < \mu < \bar{X} + \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right) = 1-\alpha$$

⇒ Pravděpodobnost, že měřením zjistíme průměrnou hodnotu $\bar{x}$ takovou, že

$\mu \in \left(\bar{x} - \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}}, \bar{x} + \frac{\sigma \cdot u_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right)$ = Intervalový odhad μ $(1-\alpha)100\%$-ní

je rovna $1-\alpha$. Tzn. budeme-li určovat průměry, a vypočítávat tyto intervaly, pak asi v cca $(1-\alpha)\cdot 100\%$ případů bude hodnota μ ležet uvnitř tohoto intervalu (spočítaného pro zrovna naměřenou hodnotu $\bar{x}$).

⇒ Využijeme v případě, že $X \to N(\mu, \sigma^2)$, nebo $n>30$ a σ známe.

Př.: Předpokládejme, že hmotnost kaprů je náhodná veličina s normálním rozdělením pravděpodobnosti se směrodatnou odchylkou $\sigma = 0{,}5$ kg.

a) Měřením hmotnosti 30 kaprů jsme zjistili jejich průměrnou hmotnost 2,3 kg. Určete 95% interval spolehlivosti pro střední hodnotu jejich hmotnosti.

Odhadovaný parametr: $\mu = EX$ (bodovým odhadem μ je $\bar{X} = 2{,}3$ kg)

Ověření předpokladů: $\boxed{X \to N(\mu, \sigma^2) \ldots \text{víme ze zadání}}$

95% intervalový odhad μ:

$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}} ; \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}\right)$$

kvantil rozdělení $N(0,1)$

, kde $\alpha = 0{,}05$ ← hladina významnosti

$\bar{X} = 2{,}3$ kg

$\sigma = 0{,}5$ kg

$n = 30$

$$\left(\underbrace{2{,}3 - \frac{0{,}5}{\sqrt{30}} 1{,}96}_{=\, 2{,}1210\ldots \text{ zaokrouhlit dolů}} ; \underbrace{2{,}3 + \frac{0{,}5}{\sqrt{30}} 1{,}96}_{=\, 2{,}4789\ldots \text{ zaokrouhlit nahoru}}\right)$$

qnorm(0.975, 0, 1) $= z_{0.975} \doteq 1{,}96$

$\Rightarrow$ 95% IO : $(2{,}12 ; 2{,}48)$

b) Kolik kaprů musíme zvážit, aby 95% IO $\mu = (\bar{X} - \Delta, \bar{X} + \Delta)$, kde $0 \le \Delta \le 0{,}1$ kg?

$$0{,}1 \geq \Delta = \frac{\sigma}{\sqrt{n}} \cdot z_{0{,}975} \quad \Big/ \cdot \frac{\sqrt{n}}{0{,}1} > 0$$

$$\sqrt{n} \geq \frac{\sigma}{0{,}1} z_{0{,}975}$$

$$n \geq \left(\frac{\sigma}{0{,}1} z_{0{,}975}\right)^2 = \left(\frac{0{,}5}{0{,}1} \cdot 1{,}96\right)^2 \doteq 96{,}0347 \quad \Rightarrow$$

Měli bychom zvážit alespoň 97 kaprů.

Příklad: Jaký musí být počet pozorování, jestliže chceme najít 95% interval spolehlivosti pro střední hodnotu hemoglobinu s chybou nejvýše $1\, g\cdot l^{-1}$? Hodnotu populačního rozptylu σ^2 odhadujeme na $46\, g^2/l^2$.

Odhadujeme: rozsah výběru = n

Předpoklady: „bez splnění předpokladu normality nelze řešit."

$\Rightarrow$ Hledáme interval $\left(\bar{X} - \underbrace{\frac{\sigma \cdot z_{0{,}975}}{\sqrt{n}}}_{\Delta}, \bar{X} + \underbrace{\frac{\sigma \cdot z_{0{,}975}}{\sqrt{n}}}_{\Delta} \right)$ takový, že:

$$\Delta = \frac{\sigma \cdot z_{0{,}975}}{\sqrt{n}} \leq 1 = \max \Delta \quad / \cdot \sqrt{n}$$

$$\sigma \cdot z_{0{,}975} \leq \sqrt{n}$$

$$(\sigma \cdot z_{0{,}975})^2 \leq n \qquad / \ \sigma = \sqrt{46}, \ z_{0{,}975} \doteq 1{,}96$$

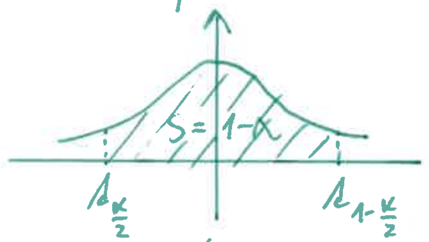
$$(\sqrt{46} \cdot 1{,}96)^2 \leq n$$

$$176{,}7 \doteq 46 \cdot 1{,}96^2 \leq n$$

Počet pozorování musí být alespoň 177.

Poznámka: Pokud $X \to N(\mu, \sigma^2)$, pak

a) $\dfrac{\bar{X}-\mu}{S}\sqrt{n} \longrightarrow t_{n-1}$ ... studentovo rozdělení s $n-1$ stupni volnosti



$$\underbrace{t_{1-\frac{\alpha}{2}} > \frac{\bar{X}-\mu}{S}\sqrt{n} > t_{\frac{\alpha}{2}}}_{\text{nastane s pravděpodobností } 1-\alpha} = -t_{1-\frac{\alpha}{2}} \quad \Big/ \cdot \frac{S}{\sqrt{n}}$$

je symetrické $\Rightarrow t_{\frac{\alpha}{2}} = -t_{1-\frac{\alpha}{2}}$

$$\frac{S\, t_{1-\frac{\alpha}{2}}}{\sqrt{n}} > \bar{X}-\mu > -\frac{S \cdot t_{1-\frac{\alpha}{2}}}{\sqrt{n}} \quad \Big/ -\bar{X}$$

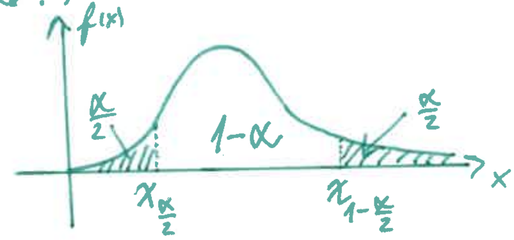
$$-\bar{X}+\frac{S\, t_{1-\frac{\alpha}{2}}}{\sqrt{n}} > -\mu > -\bar{X}-\frac{S \cdot t_{1-\frac{\alpha}{2}}}{\sqrt{n}} \quad \Big/ \cdot(-1)$$

$$\bar{X}-\frac{S \cdot t_{1-\frac{\alpha}{2}}}{\sqrt{n}} < \mu < \bar{X}+\frac{S \cdot t_{1-\frac{\alpha}{2}}}{\sqrt{n}} \Rightarrow$$

$$\boxed{(1-\alpha).100\%\ \text{IO pro } \mu : \left(\bar{X}-\frac{S \cdot t_{1-\frac{\alpha}{2}}}{\sqrt{n}},\ \bar{X}+\frac{S \cdot t_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right)}$$

výběrová směrodatná odchylka

b) $\dfrac{(n-1)S^2}{\sigma^2} \longrightarrow \chi^2_{n-1}$ .... Chi kvadrát rozdělení s $n-1$ stupni volnosti (není symetrické!)



$$\chi_{\frac{\alpha}{2}} < \frac{(n-1)S^2}{\sigma^2} < \chi_{1-\frac{\alpha}{2}}$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\sigma^2 < \frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}} \wedge \sigma^2 > \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}} \quad \Big/ \sigma > 0$$

$$\sigma < \sqrt{\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}}} \wedge \sigma > \sqrt{\frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}}} \Rightarrow$$

$$\boxed{(1-\alpha).100\%\ \text{IO pro } \mu : \left(\sqrt{\frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}}} ; \sqrt{\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}}}\right)}$$

$\Rightarrow$ Využijeme v případě, že $X \to N(\mu, \sigma^2)$ a σ neznáme.

Př. Předpokládejme, že délka trubky je náhodná veličina X normálním rozdělením pravděpodobnosti (tj. $X \to N(\mu, \sigma^2)$). Změřením 18-ti trubek jsme zjistili jejich průměrnou délku $\bar{x} = 300$ cm s výběrovou směrodatnou odchylkou $s = 2$ cm. Určete 90% intervalový odhad pro μ a σ.

a) 90% IO μ : $(\Rightarrow \alpha = 0{,}1)$

Předpoklad: $X \to N(\mu, \sigma^2)$... víme ze zadání

$\bar{X} = 300$ cm , $s = 2$ cm , $n = 18$, $t_{1-\frac{\alpha}{2}} = t_{0{,}95} \doteq 1{,}74$ ... kvantil t_{n-1} $\Rightarrow$

$= qt(0.95, 17)$

obecně: $\left(\bar{X} - \frac{s \cdot t_{0{,}95}}{\sqrt{n}} \; ; \; \bar{X} + \frac{s \cdot t_{0{,}95}}{\sqrt{n}} \right)$

90% IO μ : $(299{,}1 \; ; \; 300{,}9)$ cm

b) 90% IO σ :

$s = 2$ cm , $n = 18$

$\chi^2_{17,\, 0{,}95} = 27{,}58711 = qchisq(0.95, 17)$

$\chi^2_{17,\, 0{,}05} = 8{,}67176 = qchisq(0.05, 17)$

$$\left(\sqrt{\frac{(n-1)\, s^2}{\chi^2_{n-1,\, 1-\frac{\alpha}{2}}}} \; ; \; \sqrt{\frac{(n-1)\, s^2}{\chi^2_{n-1,\, \frac{\alpha}{2}}}} \right)$$

$\doteq 1{,}570006$; $\doteq 2{,}800276$

90% IO σ : $(1{,}5 \; ; \; 2{,}9)$

3. Úkolem je určit průměrnou hladinu cholesterolu v séru v určité populaci mužů. V náhodném výběru (pocházejícím z normálního rozdělení) 25 mužů je výběrový průměr 6,3 mmol/l a výběrová směrodatná odchylka 1,3 mmol/l.

Odhadovaný parametr: střední hodnota hladiny cholesterolu- μ

Ověření předpokladů: hladina cholesterolu má podle zadání normální rozdělení pravděpodobnosti

Intervalový odhad parametru:

$$\overline{x} = 6,3 \text{ mmol/l} \qquad s = 1,3 \text{ mmol/l} \qquad 1-\alpha = 0,95 \qquad P(5,8 < \mu < 6,8) = 0,95$$

95% intervalový odhad střední hladiny cholesterolu je $(5,8; 6,8)$ mmol/l.

```
#---------------------------------------------------------------------------------
# a) Určete 95% intervalový odhad (= interval spolehlivosti) střední hladiny cholesterolu v séru

# Předpokládáme normalitu dat (dle zadání)

n = 25       # rozsah souboru
x.bar = 6.3    # mmol/l .... průměr (bodový odhad střední hodnoty)
s = 1.3    # mmol/l .... výběrová směrodatná odchylka (bodový odhad sm. odchylky)
alpha = 0.05 # hladina významnosti (spolehlivost 1-alpha = 0.95)

## Oboustranný intervalový odhad střední hodnoty

x.bar-qt(1-alpha/2,n-1)*s/sqrt(n)   # dolní mez IO = 5,763386
x.bar+qt(1-alpha/2,n-1)*s/sqrt(n)   # horní mez IO = 6,836614


# 95% intervalový odhad střední hladiny cholesterolu je (5,8; 6,8) mmol/l.
# (výsledek ze skript)                                      ==================
```

```
#---------------------------------------------------------------------------------
# b) Určete 95% intervalový odhad (= interval spolehlivosti) směrodatné odchylky σ.

# Oboustranný intervalový odhad směrodatné odchylky:

n = 25         # rozsah souboru
s = 1.3           # mmol/l.... výběrová směrodatná odchylka (bodový odhad sm. odchylky)
alfa = 0.05             # hladina významnosti (spolehlivost 1-alpha = 0.95)

sqrt( ((n-1)*s^2) / qchisq(1-alfa/2,n-1) )    # dolní mez IO = 1,015077
sqrt( ((n-1)*s^2) / qchisq(alfa/2,n-1) )      # horní mez IO = 1,808498

# 90% intervalový odhad směrodatné odchylky délky trubky je (1,0; 1,9) mmol/l.
#                                                      =====================
```

Poznámka: Pokud $n > \frac{9}{p(1-p)}$, pak (CLV ⇒)

$$\sqrt{n}\,\frac{(P-\pi)}{\sqrt{\pi(1-\pi)}} \stackrel{\cdot}{\longrightarrow} N(0,1)$$

kde P… výběrová relativní četnost (výskytu nějakého jevu)
π… pravděpodobnost (výskytu nějakého jevu)

$$u_{\frac{\alpha}{2}} < \sqrt{n}\,\frac{(P-\pi)}{\sqrt{\pi(1-\pi)}} < u_{1-\frac{\alpha}{2}} \qquad \Big/\ u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$$

$$-u_{1-\frac{\alpha}{2}} < \sqrt{n}\,\frac{(P-\pi)}{\sqrt{\pi(1-\pi)}} < u_{1-\frac{\alpha}{2}} \qquad \Big/ \cdot \sqrt{\frac{\pi(1-\pi)}{n}}$$

$$-u_{1-\frac{\alpha}{2}}\sqrt{\frac{\pi(1-\pi)}{n}} < P-\pi < u_{1-\frac{\alpha}{2}}\cdot\sqrt{\frac{\pi(1-\pi)}{n}} \qquad \Big/ -P$$

$$-P-u_{1-\frac{\alpha}{2}}\sqrt{\frac{\pi(1-\pi)}{n}} < -\pi < -P+u_{1-\frac{\alpha}{2}}\cdot\sqrt{\frac{\pi(1-\pi)}{n}} \qquad \Big/ \cdot(-1)$$

$$P+u_{1-\frac{\alpha}{2}}\sqrt{\frac{\pi\cdot(1-\pi)}{n}} > \pi > P-u_{1-\frac{\alpha}{2}}\sqrt{\frac{\pi\cdot(1-\pi)}{n}} \qquad \Rightarrow$$

odhadneme π pomocí p.

$$\boxed{(1-\alpha)\cdot 100\%\ \text{IO pro } \pi:\ \left(P-u_{1-\frac{\alpha}{2}}\sqrt{\frac{p\cdot(1-p)}{n}}\ ;\ P+u_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right)}$$

Waldův (asymptotický) interval spolehlivosti

$\boxed{\text{Tzn. využijeme pokud } n > \frac{9}{p(1-p)}}$ $\qquad p$ … zjištěná hodnota P.

Př.: Předpokládejme, že v náhodném výběru 200 mladých mužů má 120 z nich zvýšenou hladinu cholesterolu. Určete 95% intervalový odhad pravděpodobnosti, že mladý muž má zvýšenou hladinu cholesterolu (π).

$n = 200$

$p = \frac{120}{200} = 0{,}6$ ... relativní četnost = bodový odhad π.

$\alpha = 0{,}05$

Předpoklady: $n = 200 > \frac{9}{0{,}6(1-0{,}6)} = 37{,}5$

$n < 0{,}05N \Rightarrow N > 20n = 4000$

$z_{1-\frac{\alpha}{2}} \doteq 1{,}96 = \text{qnorm}(1-0.05/2)$

$(1-\alpha)\cdot 100\%$ IO pro π: $\underbrace{\left(p - z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\ ,\ p + z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right)}_{\text{Waldův (asymptotický) interval spolehlivosti}}$

95% IO pro π: $\Big(\underbrace{0{,}6 - 1{,}96\cdot\sqrt{\frac{0{,}6\cdot 0{,}4}{200}}}_{\doteq\ 0{,}5321049}\ ;\ \underbrace{0{,}6 + 1{,}96\cdot\sqrt{\frac{0{,}6\cdot 0{,}4}{200}}}_{\doteq\ 0{,}6678951}\Big)$

$n \in (30, 2000\rangle \Rightarrow$ zaokr. na 3 platné cifry.
Nejvyšší platné místo: první nenulová číslice zleva. Nejnižší = poslední číslice (včetně 0)

$\Rightarrow$ Waldův 95% IO π: $(0{,}532\ ;\ 0{,}668)$

4. Předpokládejme, že v náhodném výběru 200 mladých mužů má 120 z nich vyšší než doporučenou hladinu cholesterolu v séru. Určete 95% interval spolehlivosti pro procento mladých mužů s vyšší hladinou cholesterolu v populaci.

Odhadovaný parametr: relativní četnost mladých mužů s vyšší hladinou cholesterolu v populaci - π

Ověření předpokladů:

předpoklad 1: $n > 30 \Rightarrow n = 200 \Rightarrow$OK
předpoklad 2: $n/N < 0,05 \Rightarrow$ výběr menší než 5 % populace považujeme za splněný
předpoklad 3: $n \cdot p \cdot (1-p) > 9 \Rightarrow 200 \cdot \frac{120}{200} \cdot (1 - \frac{120}{200}) = 40 > 9 \Rightarrow$OK

Intervalový odhad parametru (Waldův – dle CLV):

$$p = \frac{120}{200} \qquad 1 - \alpha = 0,95$$

$$P\left(p - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p(1-p)}{n}} < \pi < p + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p(1-p)}{n}}\right) = 1 - \alpha$$

$$P(0,532 < \pi < 0,668) = 0,95$$

95% Waldův intervalový odhad pravděpodobností, že muž v dané věkové kategorii bude mít vyšší než doporučenou hladinu cholesterolu v séru je $(53,2\ \%; 66,8\ \%)$.

Intervalový odhad parametru (Clopperův - Pearsonův – Statgraphics):

$P(0,528 < \pi < 0,668) = 0,95$

95% Cloperův-Pearsonův intervalový odhad pravděpodobností, že muž v dané věkové kategorii bude mít vyšší než doporučenou hladinu cholesterolu v séru je $(52,8\ \%; 66,8\ \%)$.

Mějme realizaci náhodného výběru $\boldsymbol{x}$ ze spojitého rozdělení, tj. $\boldsymbol{x} = (x_1, \ \ldots, \ x_n)$ a předpokládejme, že rozsah výběru nepřesahuje 5 % velikosti populace ($n \leq 0{,}05N$, neboli $N \geq 20n$).

Odhadovaný parametr		**Předpoklady**	**Meze oboustranného intervalového odhadu**		**Dolní mez levostranného intervalového odhadu**	**Horní mez pravostranného intervalového odhadu**
			$\boldsymbol{M_D}$	$\boldsymbol{M_H}$	$\boldsymbol{M_D^*}$	$\boldsymbol{M_H^*}$
Míra polohy	$\boldsymbol{\mu}$	normalita nebo $n > 30$, známe σ	$\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}$	$\bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}$	$\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}$	$\bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha}$
		normalita, neznáme σ	$\bar{x} - \frac{s}{\sqrt{n}} t^{n-1}_{1-\frac{\alpha}{2}}$	$\bar{x} + \frac{s}{\sqrt{n}} t^{n-1}_{1-\frac{\alpha}{2}}$	$\bar{x} - \frac{s}{\sqrt{n}} t^{n-1}_{1-\alpha}$	$\bar{x} + \frac{s}{\sqrt{n}} t^{n-1}_{1-\alpha}$
Míra variability	$\boldsymbol{\sigma^2}$	normalita	$\frac{(n-1)s^2}{\chi^{n-1}_{1-\frac{\alpha}{2}}}$	$\frac{(n-1)s^2}{\chi^{n-1}_{\frac{\alpha}{2}}}$	$\frac{(n-1)s^2}{\chi^{n-1}_{1-\alpha}}$	$\frac{(n-1)s^2}{\chi^{n-1}_{\alpha}}$

Mějme realizaci náhodného výběru $\boldsymbol{x}$ z alternativního rozdělení, tj. $\boldsymbol{x} = (x_1, \ \ldots, \ x_n)$ a předpokládejme, že rozsah výběru nepřesahuje 5 % velikosti populace ($n \leq 0{,}05N$, neboli $N \geq 20n$).

Odhadovaný parametr		**Předpoklady**	**Meze oboustranného intervalového odhadu**		**Dolní mez levostranného intervalového odhadu**	**Horní mez pravostranného intervalového odhadu**
			$\boldsymbol{M_D}$	$\boldsymbol{M_H}$	$\boldsymbol{M_D^*}$	$\boldsymbol{M_H^*}$
parametr bin. rozdělení	$\boldsymbol{\pi}$	$n > \frac{9}{p(1-p)}$	$p - z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}$	$p + z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}$	$p - z_{1-\alpha} \sqrt{\frac{p(1-p)}{n}}$	$p + z_{1-\alpha} \sqrt{\frac{p(1-p)}{n}}$

Poznámka: Funkce, které lze použít pro intervalový odhad mediánu v software R, včetně předpokladů pro jejich použití, najdete v R-taháku.

5. V rámci výzkumné studie pracujeme s náhodným výběrem 70 žen z české populace. U každé z žen byl změřen hemoglobin s přesností 0,1 g/100 ml. Naměřené hodnoty jsou v uvedeny v souboru . Nalezněte 95% intervalové odhady směrodatné odchylky a střední hodnoty hemoglobinu v populaci českých žen. (Normalitu ověřte na základě exploračních grafů.)

Odhadované parametry: střední hodnota hladiny hemoglobinu - μ,
směrodatná odchylka hladiny hemoglobinu - σ

Ověření předpokladů: pomocí exploračních grafů $\Rightarrow$ OK

Intervalový odhad parametru 1:

```
# a) Určíme 95% oboustranný intervalový odhad střední hodnoty
t.test(hem$hodnoty,alternative = "two.sided",conf.level=0.95)
```

$$P(11,65 < \mu < 12,32) = 0,95$$

95% intervalový odhad střední hodnoty obsahu hemoglobinu v krvi českých žen je $(11,7; 12,3)$ g/100 ml.

Intervalový odhad parametru 2:

```
varTest(hem$hodnoty,conf.level = 0.95) #je 95% intervalový odhad rozptylu
```

$$P(1,2 < \sigma < 1,7) = 0,95$$

95% intervalový odhad směrodatné odchylky obsahu hemoglobinu v krvi českých žen je $(1,2; 1,7)$ g/100 ml.

```
## Načtení dat z xlsx souboru (pomoci balíčku readxl)
## adresu v následujícím příkazu je nutno upravit dle vlastního nastavení
hem = read_excel("intervalove_odhady.xlsx",
                 sheet = "Hemoglobin")
colnames(hem)="hodnoty"


#----------------------------------------------------
## Explorační analýza - zjistit, zda není třeba odstranit odlehlá pozorování
boxplot(hem$hodnoty)

# Data neobsahují odlehlá pozorování.
length(hem$hodnoty)
sd(hem$hodnoty)
summary(hem$hodnoty)


#----------------------------------------------------
# Ověření normality
qqnorm(hem$hodnoty)
qqline(hem$hodnoty)
hist(hem$hodnoty)

skewness(hem$hodnoty)           # koef šikmosti = 0 u normálního rozdělení; \in [-2,2] => OK
moments::kurtosis(hem$hodnoty)-3  # standardizovaný koef špičatosti = 0 u normálního rozdělení; \in [-2
# Šikmost i špičatost odpovídá norm. rozdělení. Pro konečné rozhodnutí o normalitě dat použijeme
# test normality.

# Známe-li testování hypotéz, ověříme předpoklad normality Shapirovým . Wilkovým testem.
shapiro.test(hem$hodnoty)
# Na hl. významnosti 0.05 nelze předpoklad normality zamítnout (p-hodnota=0.522, Shapirův-Wilkův test).
```

7. V průběhu experimentu sledujeme vliv chlazení (skupina 1 - žádné, skupina 2 - chlazení vodou) okolních struktur na největší rozměr poškození tkáně slinivky břišní . Kvantifikujte efekt vlivu chlazení a určete jeho 95% intervalový odhad.

Odhadované parametry: střední hodnota poškození tkáně skupina 1 - μ_1,
střední hodnota poškození tkáně skupina 2 - μ_2

Ověření předpokladů: normalita rozdělení poškození tkáně - skupina 1 $\Rightarrow$ OK
normalita rozdělení poškození tkáně - skupina 2 $\Rightarrow$ OK
shoda rozptylů zamítnuta pro $\alpha = 0,05$

Intervalový odhad středního poškození tkáně pro skupinu 1:

$$P(24,83 < \mu_1 < 25,1) = 0,95$$

Intervalový odhad středního poškození tkáně pro skupinu 2:

$$P(21,29 < \mu_2 < 22,39) = 0,95$$

Intervalový odhad rozdílu středních poškození tkáně pro skupinu 1 a skupinu 2:

$$P(2,59 < \mu_1 - \mu_2 < 3,75) = 0,95$$

Intervalový odhad rozdílu středních poškození tkáně neobsahuje nulu, tzn. μ_1 je statistick výzmanně vyšší než μ_2 $\Rightarrow$ dochází k výrazně vyššímu poškození tkáně, jestliže není prováděno chlazení okolních struktur.

$[(24,83; 25,19),\ (21,29; 22,39),\ (2,59; 3,75)]$