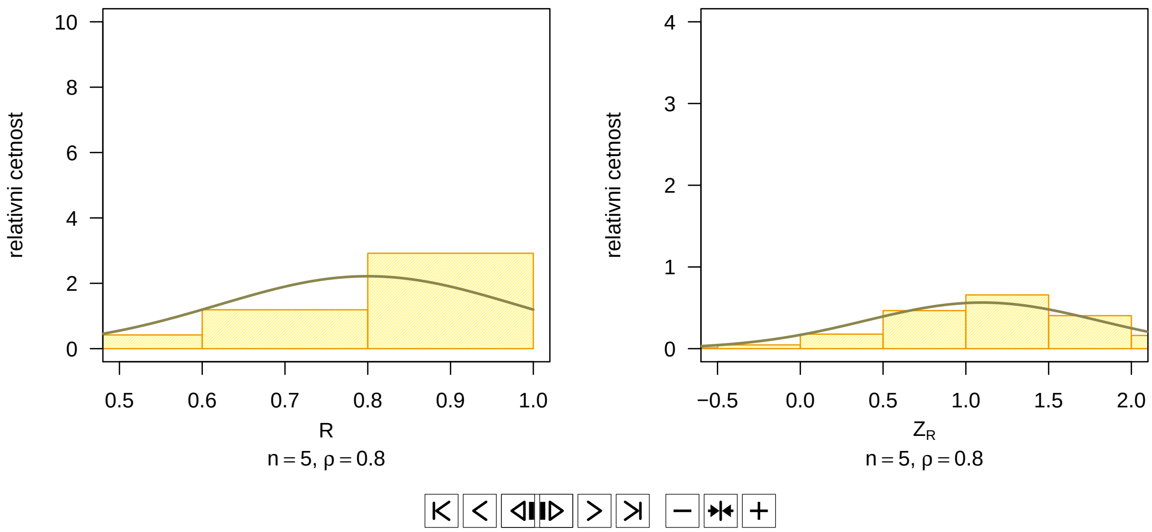
7 Test o korelačním koeficientu

Příklad 7.1. Konvergence ρ a ξ k normálnímu rozdělení pro $n \to \infty$
Proveďte simulaci pseudonáhodných čísel z $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 1$, $\sigma_2 = 1$, $M = 1\,000$, $\rho = 0.8$. Pro každé $m = 1, 2, \ldots, M$, vypočítejte realizaci výběrového korelačního koeficientu r_m a Fisherovy Z-proměnné $z_{R,m}$. Zobrazte histogramy simulovaných r_m a $z_{R,m}$ a superponujte je teoretickými hustotami příslušných normálních rozdělení. Vytvořte animaci zobrazující rozdělení výběrového korelačního koeficientu R a Fisherovy Z transformace pro různé rozsahy náhodného výběru $n \in \{5, 10, 15, \ldots, 65, 70\}$.

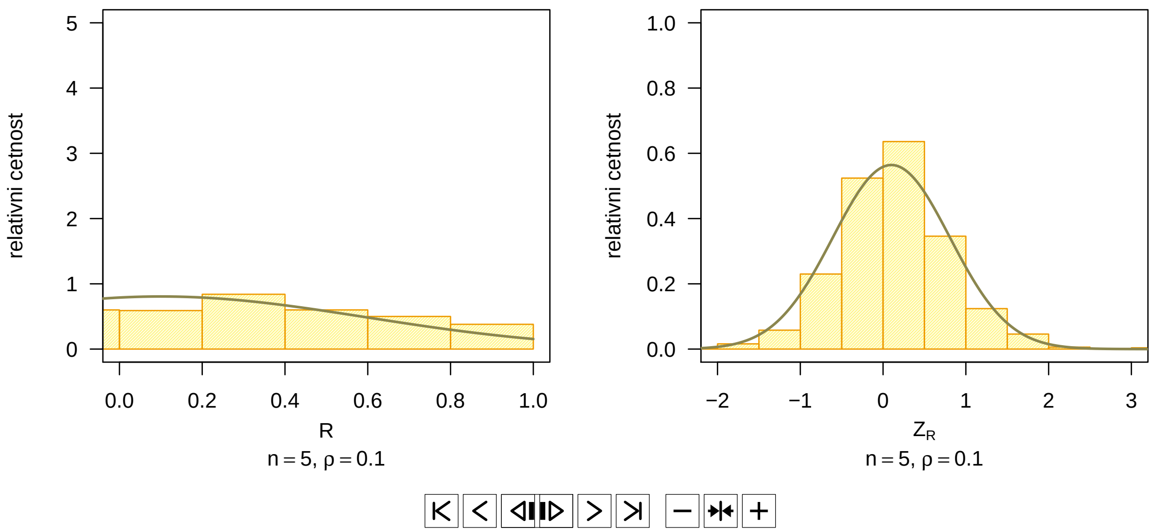
Do komentářů stručně uveďte porovnání kvality konvergence R a Z_R k normálnímu rozdělení pro $n \to \infty$ a závěr, pro jak velké n je již vhodné použít Pearsonův korelančí koeficient R, resp. Fisherovu Z-transformaci Z_R.



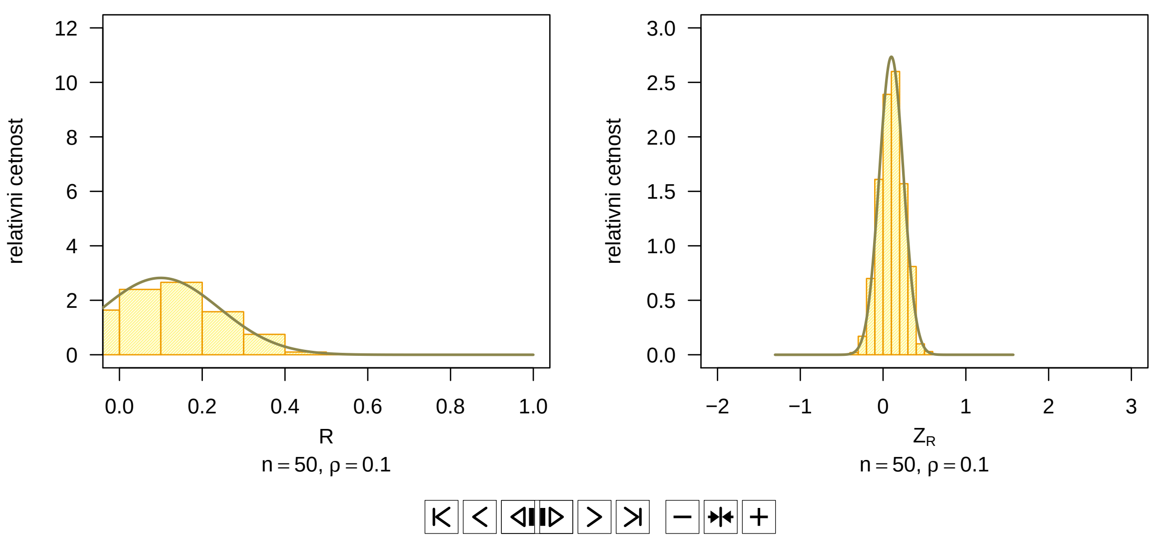
Obrázek 1: Rozdělení (a) výběrového korelačního koeficientu R, (b) Fisherovy Z-proměnné Z_r při měnícím se rozsahu náhodného výběru

Příklad 7.2. Konvergence ρ a ξ k normálnímu rozdělení pro $\rho \to 0.5$
Vytvořte animaci zobrazující konvergenci rozdělení výběrového korelačního koeficientu R a Fisherovy Z-transformace k normálnímu rozdělení pro $\rho \to 0.9$. Hodnoty koeficientu ρ volte $\rho \in \{0.1, 0.2, \ldots, 0.9\}$, rozsah náhodného výběru zvolte (a) $n = 5$, (b) $n = 50$.

Do komentářů stručně popište změnu kvality konvergence R a Z_R k normálnímu rozdělení pro $\rho \to 0.9$ (je-li nějaká). Také navzájem porovnejte situace pro $n = 5$ a $n = 50$.



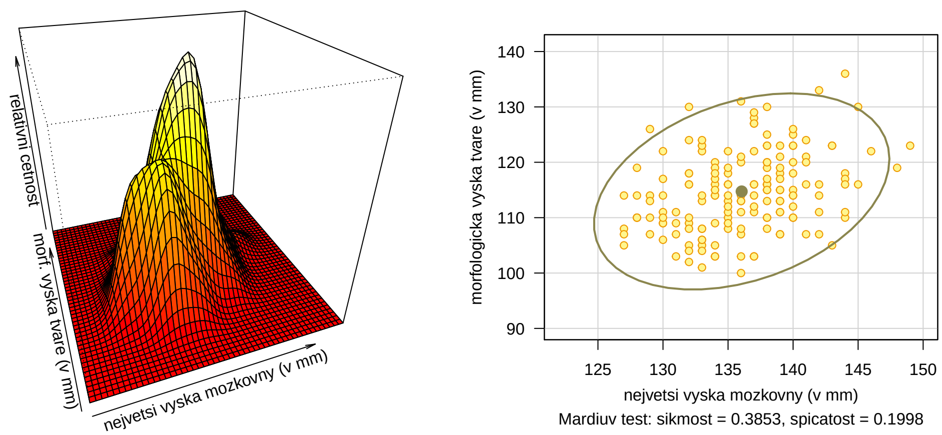
Obrázek 2: Konvergence rozdělení výběrového korelačního koeficientu R a Fisherovy Z-transformace Z_r k normálnímu rozdělení pro $\rho \to 0.9$ při pevně zvoleném rozsahu náhodného výběru $n = 5$



Obrázek 3: Konvergence rozdělení výběrového korelačního koeficientu R a Fisherovy Z-transformace Z_r k normálnímu rozdělení pro $\rho \to 0.9$ při pevně zvoleném rozsahu náhodného výběru $n = 50$

Příklad 7.3. Test o korelačním koeficientu ρ
Mějme datový soubor 05-one-sample-correlation-skull-mf.txt obshaující údaje o největší výšce mozkovny skull.pH (v mm) a morfologické výšce tváře face.H (v mm) starověké egyptské mužské a ženské populace.

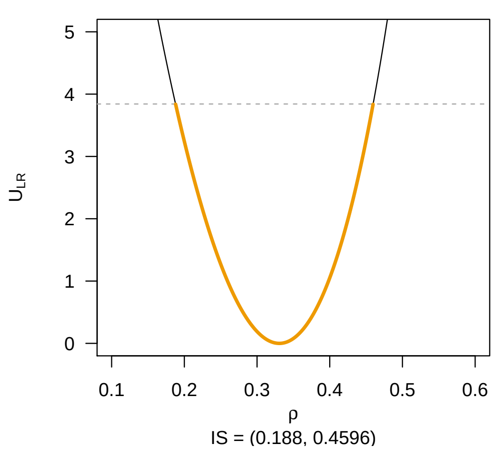
Na hladině významnosti $\alpha = 0.05$ testujte nulovou hypotézu o shodě korelačního koeficientu největší výšky mozkovny a morfologické výšky tváře u mužů s hodnotou 0.251. Testování proveďte pomocí (a) kritického oboru, (b) intervalu spolehlivosti, (c) p-hodnoty při použití (1) Waldovy testovací statistiky Z_W, (2) věrohodnostní testovací statistiky U_{LR}. Dále vykreslete graf zobrazující 95 % věrohodnostní interval spolehlivosti pro korelační koeficient ρ získaný na základě U_{LR} testovací statistiky.



Obrázek 4: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro největší výšku mozkovny a morfologickou výšku tváře mužů starověké egyptské populace (v mm)

Tabulka 1: Výsledky Waldova a věrohodnostního testu o korelačním koeficientu ρ

	$\widehat{\rho}$	statistika	$\mathcal{W}_{hh}$	$\mathcal{W}_{dh}$	IS$_{dh}$	IS$_{hh}$	p-hodnota
Waldův přistup	0.3306	1.1048	-1.9600	1.9600	0.1869	0.4606	0.2692
Věrohodnostní přístup	0.3306	1.2418		3.8415	0.1880	0.4596	0.2651



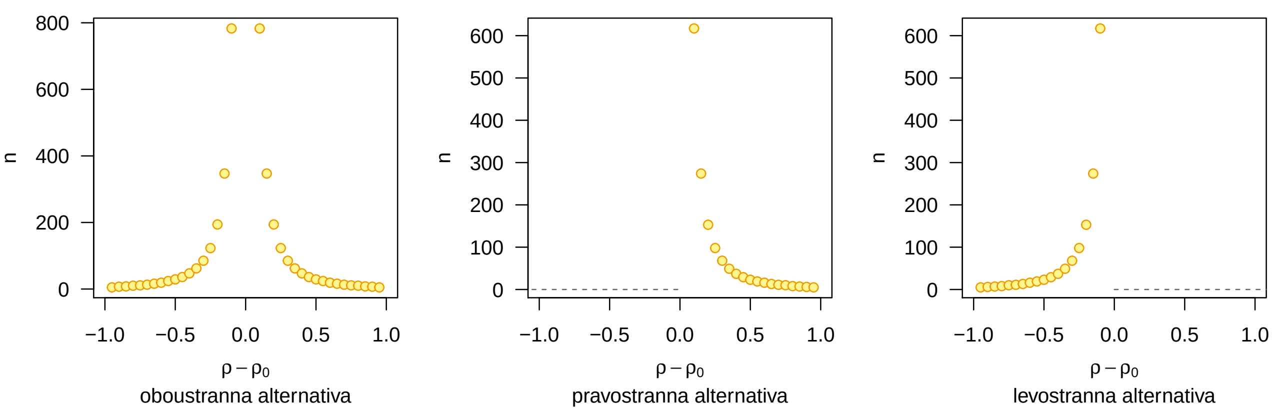
Obrázek 5: 95 % věrohodnostní interval spolehlivosti pro korelační koeficient ρ

Příklad 7.4. Minimální rozsah náhodného výběru
Předpokládejme, že $(X, Y)^T \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Testujeme všechny tři typy hypotéz:

1. $H_{01} : \rho = \rho_0$ oproti $H_{11} : \rho \neq \rho_0$ (oboustranná),
2. $H_{02} : \rho \leq \rho_0$ oproti $H_{12:}\ \rho > \rho_0$ (pravostranná),
3. $H_{03} : \rho \geq \rho_0$ oproti $H_{13} : \rho < \rho_0$ (levostranná).

Pro hypotézy (a), (b), (c) odvoďte vztah pro minimální rozsah náhodného výběru n při použití testovací statistiky pro rozdíl $\xi - \xi_0$ založené na Fisherově Z-proměnné. Naprogramujte funkci `min.n()`, která pro libovolnou z hypotéz (a), (b), (c) při předem stanovené hodnotě ρ, ρ_0, α a β^* vypočítá minimální potřebný rozsah n dvourozměrného náhodného výběru. Pomocí funkce `min.n()` vypočítejte minimální rozsah náhodného výběru pro hypotézy (a)–(c) při $\alpha = 0.05$ a $\beta^*(\xi) = 0.80$. Parametr $\rho_0 = 0$ a $\rho \in \{-0.95, -0.9, \ldots, -0.15, -0.1, 0.1, 0.15, \ldots, 0.9, 0.95\}$. Minimální rozsahy náhodných výběrů zakreslete jako bodový graf, kde na osu x vyneste rozdíl $\rho - \rho_0$, na osu y vypočítaný minimální rozsah n.

Určete (a) Jaký rozsah náhodného výběru potřebujeme na otestování nulové hypotézy $H_0 : \rho \leq \rho_0$, kde $\rho_0 = 0.85$, chceme-li, aby síla testu $\beta^*(\xi) = 0.85$ při stanovené hladině významnosti $\alpha = 0.1$, předpokládáme-li, že výběrový korelační koeficient bude nabývat hodnoty 0.60?; (b) Jaký rozsah náhodného výběru potřebujeme na otestování nulové hypotézy $H_0 : \rho = \rho_0$, kde $\rho_0 = 0.33$, chceme-li, aby síla testu $\beta^*(\xi) = 0.7$ při stanovené hladině významnosti $\alpha = 0.05$, předpokládáme-li, že výběrový korelační koeficient bude nabývat hodnoty 0.22?; (c) Jaký rozsah náhodného výběru potřebujeme na otestování nulové hypotézy $H_0 : \rho \geq \rho_0$, kde $\rho_0 = -0.74$, chceme-li, aby síla testu $\beta^*(\xi) = 0.9$ při stanovené hladině významnosti $\alpha = 0.01$, předpokládáme-li, že výběrový korelační koeficient bude nabývat hodnoty -0.5? Výsledné rozsahy pro přehlednost vložte spolu se zadanými parametry do přehledné tabulky.



Obrázek 6: Minimální rozsahy n náhodných výběrů při předem zvolených hodnotách ρ, ρ_0, α a β^*

Tabulka 2: Minimální rozsahy n náhodných výběrů při předem zvolených hodnotách ρ, ρ_0, α a β^*

	α	$\beta^*(\xi)$	$\beta(\xi)$	ρ_0	ρ	n
$H_0 : \rho \leq \rho_0$	0.10	0.85	0.15	0.85	0.60	20.00
$H_0 : \rho = \rho_0$	0.05	0.70	0.30	0.33	0.22	438.00
$H_0 : \rho \geq \rho_0$	0.01	0.90	0.10	-0.74	-0.50	84.00