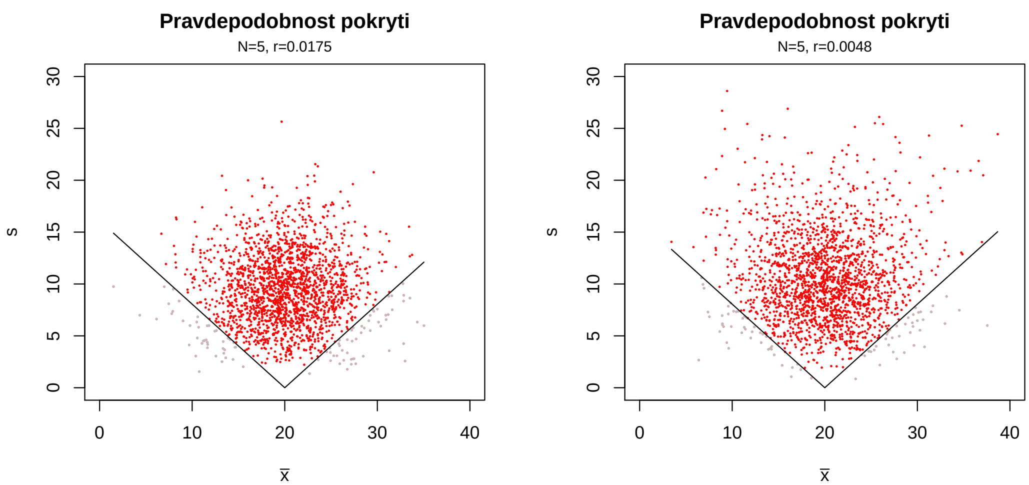
Zadání příkladů – Statistická inference II – 2017

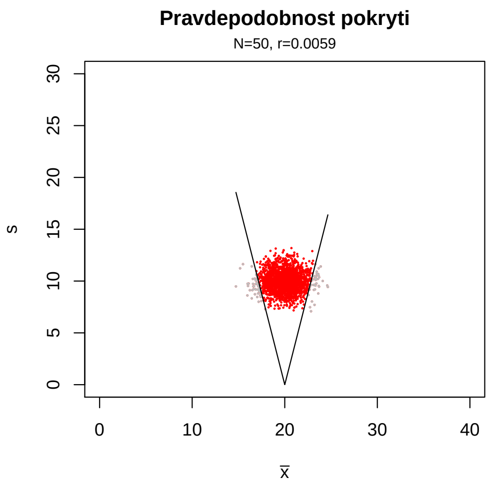
Příklad 12. nezávislost μ a σ^2; pravděpodobnost pokrytí Nechť $X \sim N(\mu, \sigma^2)$, kde $\mu = 20$ a $\sigma^2 = 100$. Pomocí simulační studie vypočítejte Pearsonův korelační koeficient $r_{\bar{X},S}$. Nakreslete šedou barvou rozptylový graf $(\bar{x}_m, s_m)$, kde $m = 1, 2, \ldots, M$, přičemž $M = 5\,000$. Černou barvou vyznačte v grafu takové body $(\bar{x}_m, s_m)$, pro které platí $t_{W,m} = \left|\frac{\bar{x}_m - \mu}{s_m}\sqrt{n}\right| < t_{n-1}(\alpha/2)$. Dále vykreslete hranice, které jsou definovány body $(\bar{x}_m, s_m)$, jež splňují vztah $t_{W,m} = t_{n-1}(\alpha/2)$. Vypočítejte pravděpodobnost pokrytí 95 % DIS pro μ jako podíl $\sum_m I(t_{W,m} < t_{n-1}(\alpha/2))/M$. Zvolte (a) $n = 5$, (b) $n = 50$ a (c) $n = 100$.

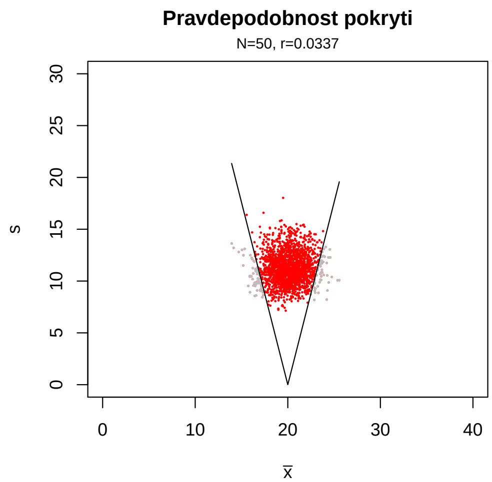
Simulaci proveďte také za předpokladu, že data pochází ze smíšeného rozdělení $X \sim [pN(\mu, \sigma_1^2) + (1-p)N(\mu, \sigma_2^2)]$, kde $p = 0.9$, $\mu = 20$, $\sigma_1^2 = 100$ a $\sigma_2^2 = 400$.

```
##                                         n
## aktualni pst.pokryti                0.952
## nominalni pst.pokryti (spolehlivost) 0.950
##                                         n
## aktualni pst.pokryti                0.958
## nominalni pst.pokryti (spolehlivost) 0.950
```

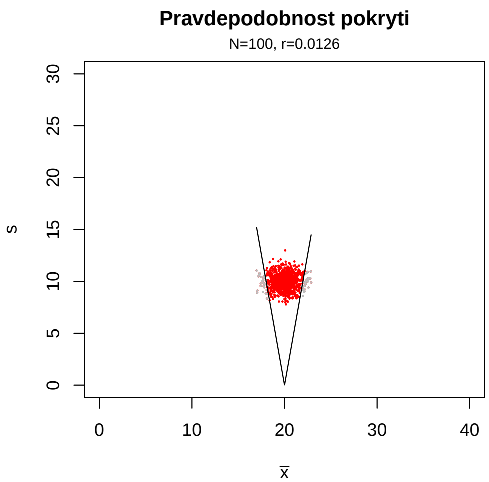


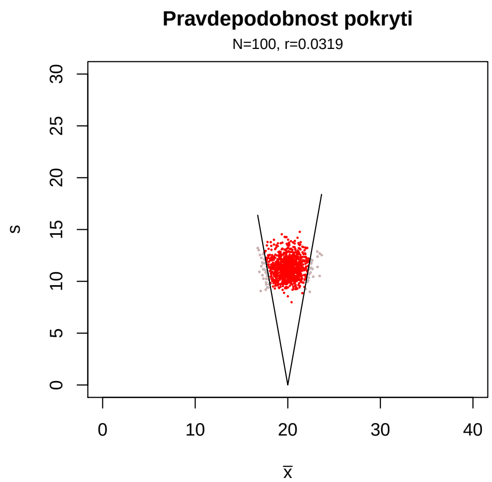

```
##                                         n
## aktualni pst.pokryti                0.958
## nominalni pst.pokryti (spolehlivost) 0.950
##                                          n
## aktualni pst.pokryti                0.9425
## nominalni pst.pokryti (spolehlivost) 0.9500
```






```
##                                          n
## aktualni pst.pokryti                 0.946
## nominalni pst.pokryti (spolehlivost) 0.950
##                                          n
## aktualni pst.pokryti                 0.942
## nominalni pst.pokryti (spolehlivost) 0.950
```

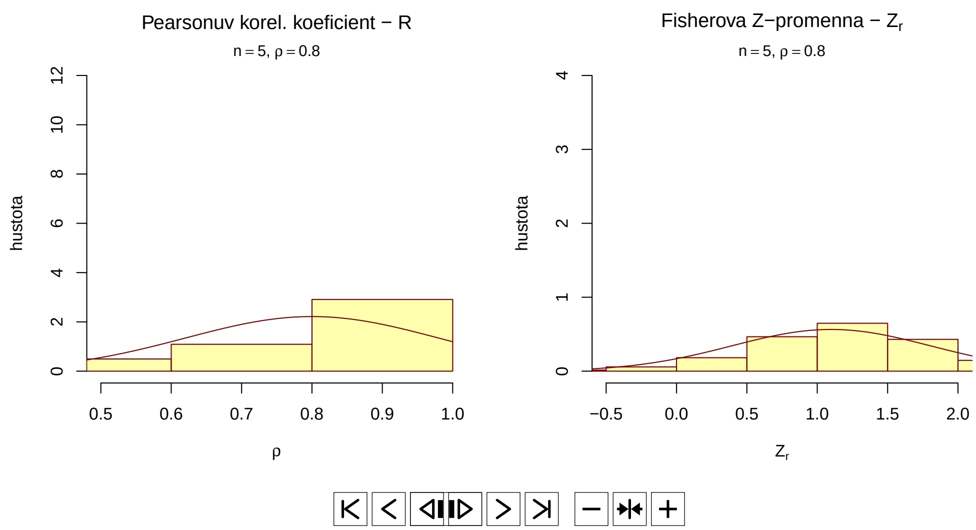




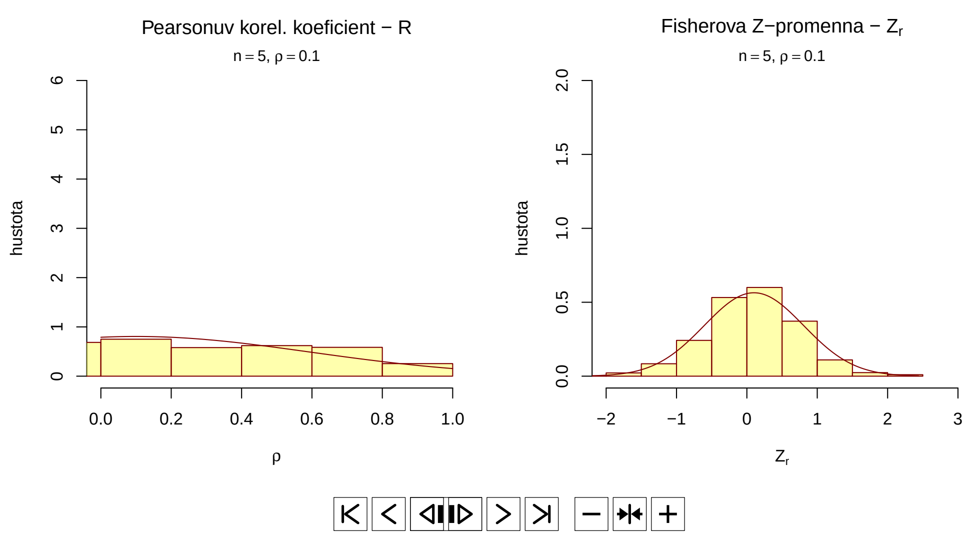
Příklad 13. konvergence ρ a ξ k normálnímu rozdělení

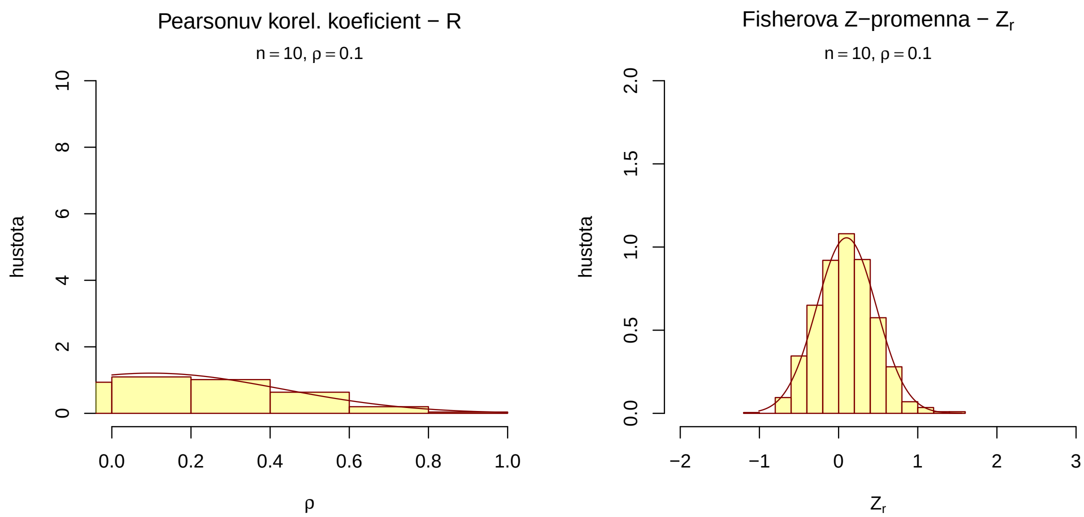
1. Proveďte simulaci pseudonáhodných čísel z $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, kde $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 1$, $\sigma_2 = 1$, $M = 10\,000$, $rho = 0.8$.

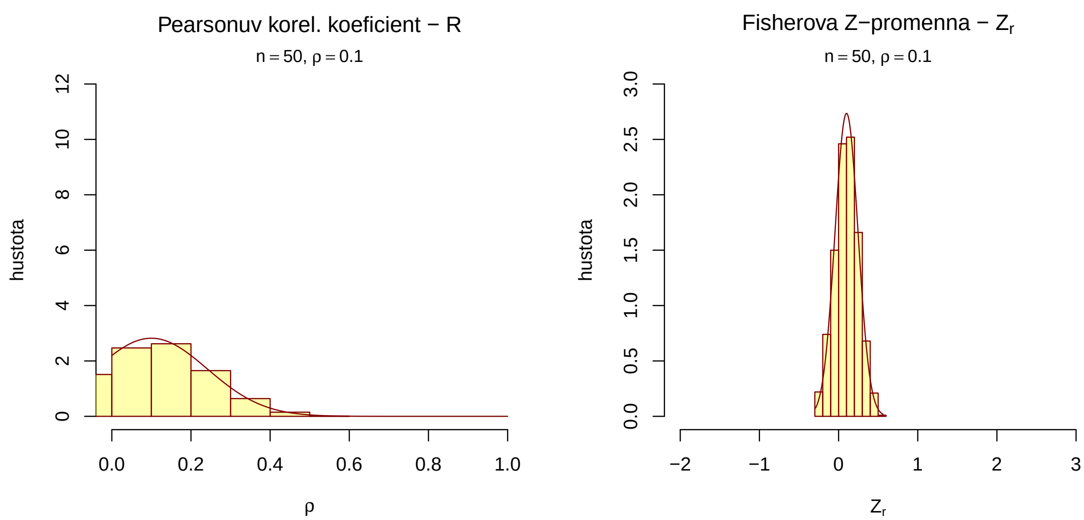
   Pro každé $m = 1, 2, \ldots, M$, vypočítejte realizaci výběrového korelačního koeficientu r_m a Fisherovy Z-proměnné $z_{R,m}$. Zobrazte histogramy simulovaných r_m a $z_{R,m}$ a superponujte je teoretickými hustotami příslušných normálních rozdělení. Vytvořte animaci zobrazující rozdělení výběrového korelačního koeficientu R a Fisherovy Z transformace pro různé rozsahy náhodného výběru $n \in \{5, 10, \ldots, 70\}$.



2. Vytvořte animaci zobrazující rozdělení výběrového korelačního koeficientu $R \in \{0.1, 0.2, \ldots, 0.9\}$ a Fisherovy Z transformace pro tyto různé korelační koeficienty. Animace vytvořte pro (a) $n = 5$, (b) $n = 10$, (c) $n = 100$.




Pearsonuv korel. koeficient − R
n = 10, ρ = 0.1
hustota
ρ
Fisherova Z−promenna − Z_r
n = 10, ρ = 0.1
hustota
Z_r


Pearsonuv korel. koeficient − R
n = 50, ρ = 0.1
hustota
ρ
Fisherova Z−promenna − Z_r
n = 50, ρ = 0.1
hustota
Z_r

Příklad 14. Test o korelačním koeficientu ρ: Mějme data one-sample-correlation-skull-mf.txt a proměnné největší výška mozkovny skull.pH a morfologická výška tváře face.H (obě v mm) starověké egyptské mužské populace, o kterých předpokládáme, že mají dvourozměrné normální rozdělení $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

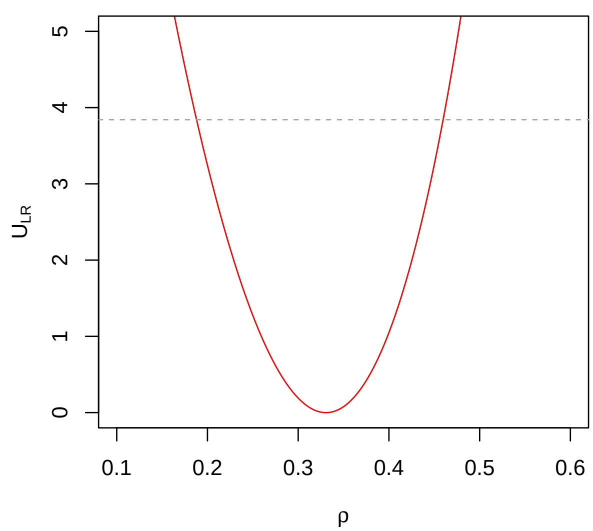
1. Otestujte hypotézu o tom, že korelační koeficient největší výšky mozkovny a morfologické výšky tváře je rovný 0.251.

2. Vypočítejte $100 \times (1-\alpha)\%$ empirický DIS pro korelační koeficient, kde koeficient spolehlivosti $1-\alpha = 0.95$.

Použijte

(a) Waldovu testovací statistiku Z_W;

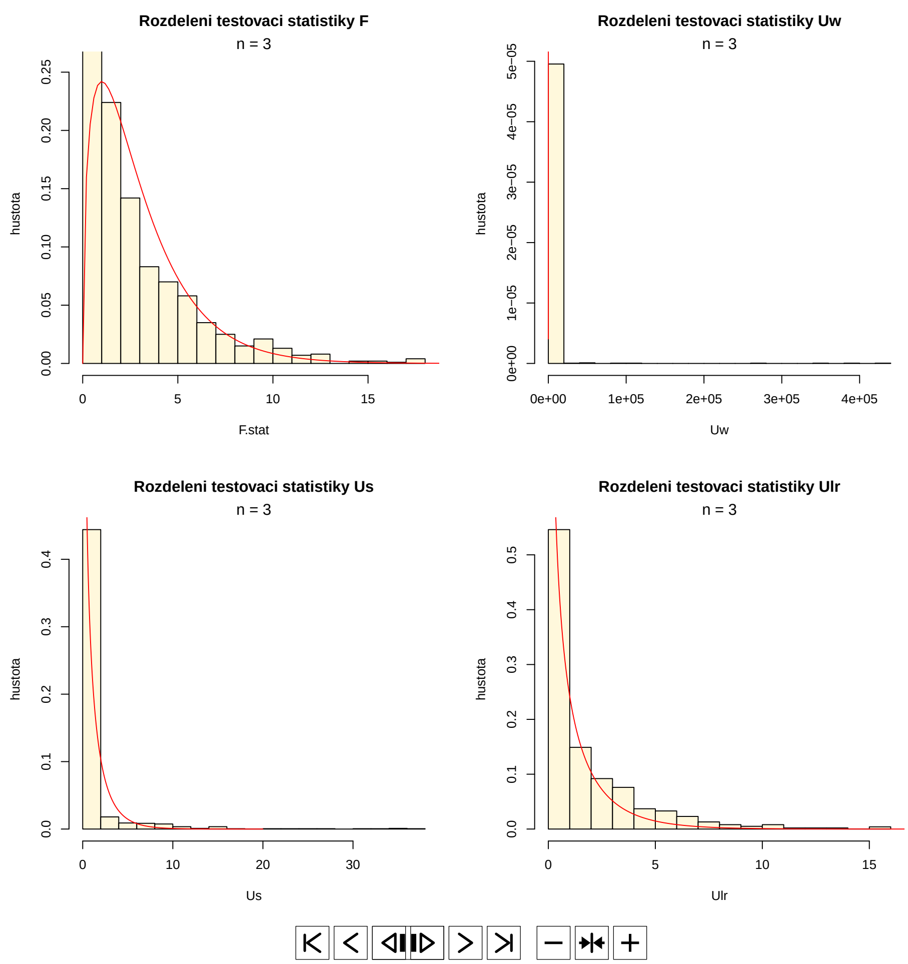
(b) testovací statistiku poměrem věrohodnosti U_{LR}.



	test.statistika	W	kritický obor	p-hodnota
Waldův přístup	1.104799	$(-\infty, -1.959964\rangle \cup \langle 1.959964; \infty)$	$(0.1868617, 0.4605561)$	0.2692467
Věrohodnostní přístup	1.241757	$\langle 3.841459; \infty)$	$(0.1880626, 0.4590998)$	0.2651328

Příklad 15. Rozdělení testovacích statistik F, U_W, U_S a U_{LR} jednovýběrového testu o rozptylu σ^2
Pomocí simulační studie ověřte, že

- statistika $F = \dfrac{nS_n^2}{\sigma_0^2} \sim \chi_n^2$
- statistika $U_w = \dfrac{n}{2}\left(\dfrac{n}{F_{obs}} - 1\right)^2 \sim \chi^2(1)$
- statistika $U_S = \dfrac{n}{2}\left(\dfrac{F_{obs}}{n} - 1\right)^2 \sim \chi^2(1)$
- statistika $U_{LR} = F_{obs} - n\left(1 + \ln\left(\dfrac{F_{obs}}{n}\right)\right) \sim \chi^2(1)$
- Vygenerujte $M = 1000$ pseudonáhodných výběrů o rozsahu $n = 5$. Pro každý náhodný výběr vypočítejte hodnoty realizací testovacích statistik F, U_W, U_S a U_{LR}. Vytvořte histogramy těchto testovacích statistik a superponujte je křivkami příslušného rozdělení.
- Vytvořte animaci, pomocí které bude zřejmé, jak se s rostoucím n rozdělení testovacích statistik F, U_W, U_S a U_{LR} asymptoticky blíží k příslušnému rozdělení. Animaci spravte pro měnící se $n \in \{3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000\}$



Příklad 16. Pokračování příkladu 15

- Vygenerujte $M = 1000$ pseudonáhodných výběrů o rozsahu $n = 5$. Pro každý náhodný výběr stanovte hodnoty realizací testovacích statistik F, U_W, U_S a U_{LR}. Pomocí funkce `density()` vypočítejte jádrové odhady rozdělení testovacích statistik U_W, U_S a U_{LR}. Tyto jádrové odhady zobrazte do jednoho grafu a barevně je odlište. Do grafu dokreslete křivku hustoty χ^2 rozdělení o jednom stupni volnosti a vertikální referenční čáru v hodnotě 0.

- Vytvořte animaci, pomocí které bude zřejmé, jak se s rostoucím n rozdělení testovacích statistik U_W, U_S a U_{LR} asymptoticky blíží k $\chi^2(1)$ rozdělení. Animaci spravte pro měnící se $n \in \{3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000\}$