evropský sociální fond v ČR
EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

TESTOVÁNÍ NORMALITY ZE ZAOKROUHLENÝCH DAT

Michal Friesl

Katedra matematiky
Fakulta aplikovaných věd
Západočeská univerzita v Plzni

Zaokrouhlená data

Model

- náhodný výběr $X_1, \ldots, X_n$ z F
- pozorujeme zaokrouhlené hodnoty $X_1^d, \ldots, X_n^d$

Úkol

- testovat $H_0\colon F \in \{\mathrm{N}(\mu, \sigma^2), \mu \in \mathbf{R}, \sigma > 0\}$
- tj. nespecifikované normální rozdělení

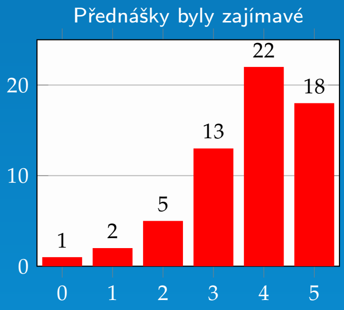
Přičemž nejspíš

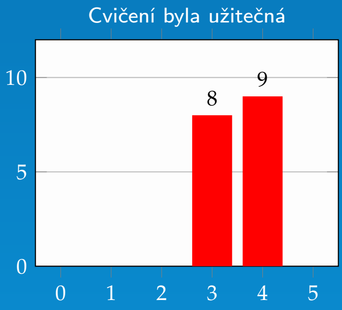
- šířka intervalů není zanedbatelná vzhledem k rozptylu
- rozsah výběru n malý

Příklad: Hodnocení kvality výuky

Studentské hodnocení předmětů

- k předmětu 3 až 5 tvrzení (Přednáška byla srozumitelná,. . .)
- pozorování: 0,1,. . . ,5 (míra souhlasu studenta)



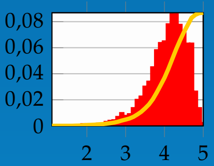


Příklad: Hodnocení kvality výuky

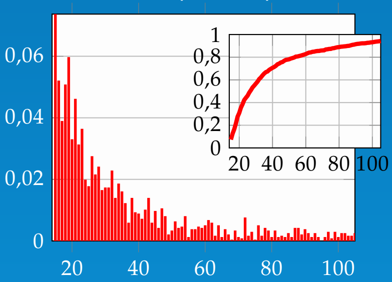
Alespoň 15 odpovídajících u 2361 tvrzení, průměrně

- 41 odpovídajících (v 50 % méně než 27)
- průměrné hodnocení 4,05
- směrodatná odchylka 0,95 (v 50 % menší než 0,85)

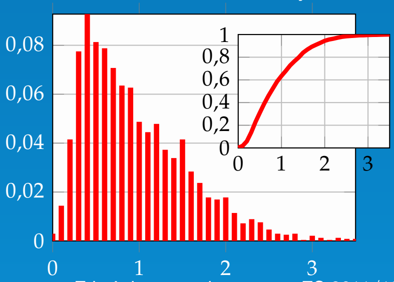
Rozdělení průměrů



Rozdělení počtů odpovědí



Rozdělení směrod. odchylek

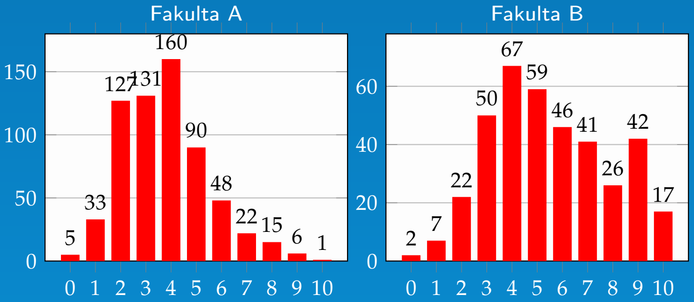


Zdroj dat: portal.zcu.cz — ZS 2011/12

Příklad: Vstupní test

Vstupní test znalostí studentů 1. ročníku z matematiky

- 10 příkladů
- pozorování: 0 1 . . . 10 (počet správně zodpovězených)



- směrodatné odchylky kolem 1,9, průměr kolem 4 až 6 bodů

Zdroj dat: stat.kma.zcu.cz/testik11/

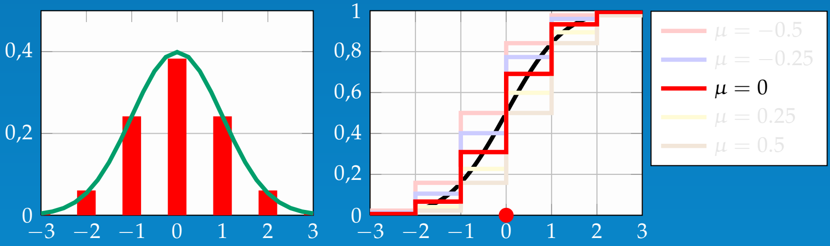
Následuje

1. **Diskretizované normální rozdělení**
   - odhady parametrů
2. **Připomenutí χ^2 testu**
   - asymptotický při $n \to \infty$, ale při malých n?
3. **Kolmogorovův-Smirnovův test a jeho varianty**
   - pro spojitá data, ale použít pro seskupená a složenou hypotézu?
4. **Porovnání testů**
5. **Cenzorovaná data**
   - výsledky příkladů

Diskretizované normální rozdělení

Tvar rozdělení X_d závisí i na μ (hlavně při menších σ)

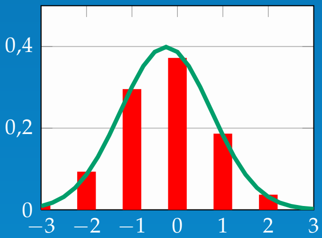
- pravděpodobnosti a distribuční funkce $\mathrm{N}(\mu, 1)$ pro různá μ

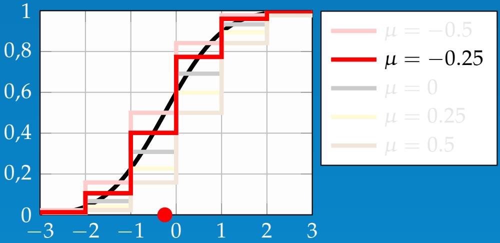


Diskretizované normální rozdělení

Tvar rozdělení X_d závisí i na μ (hlavně při menších σ)

- pravděpodobnosti a distribuční funkce $\mathrm{N}(\mu, 1)$ pro různá μ

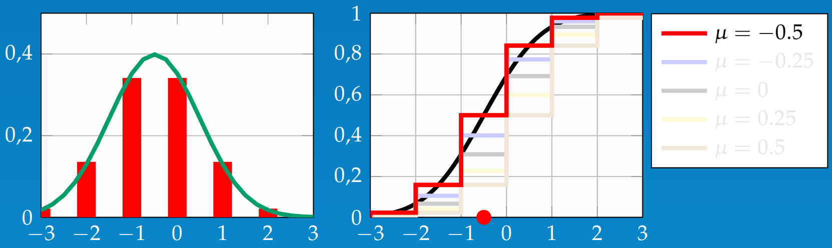




Diskretizované normální rozdělení

Tvar rozdělení X_d závisí i na μ (hlavně při menších σ)

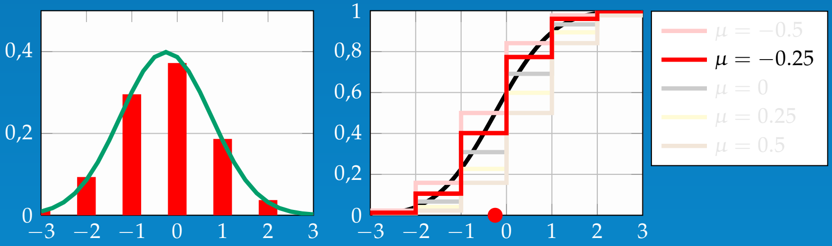
- pravděpodobnosti a distribuční funkce $\mathrm{N}(\mu, 1)$ pro různá μ



Diskretizované normální rozdělení

Tvar rozdělení X_d závisí i na μ (hlavně při menších σ)

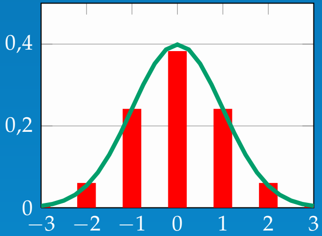
- pravděpodobnosti a distribuční funkce $\mathrm{N}(\mu, 1)$ pro různá μ

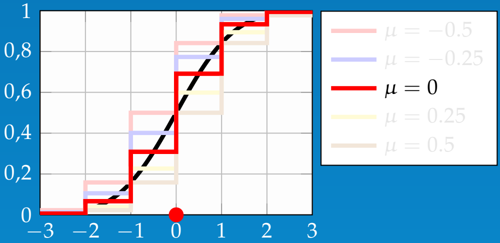


Diskretizované normální rozdělení

Tvar rozdělení X_d závisí i na μ (hlavně při menších σ)

- pravděpodobnosti a distribuční funkce $\mathrm{N}(\mu, 1)$ pro různá μ

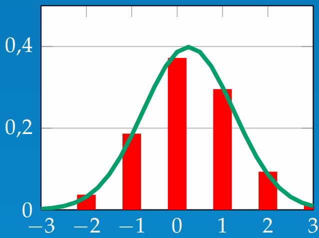


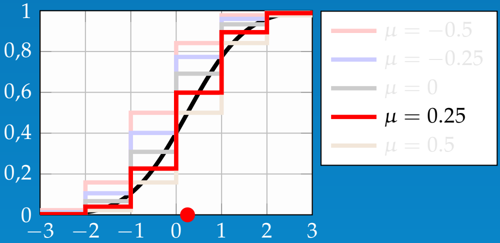


Diskretizované normální rozdělení

Tvar rozdělení X_d závisí i na μ (hlavně při menších σ)

- pravděpodobnosti a distribuční funkce $\mathrm{N}(\mu, 1)$ pro různá μ

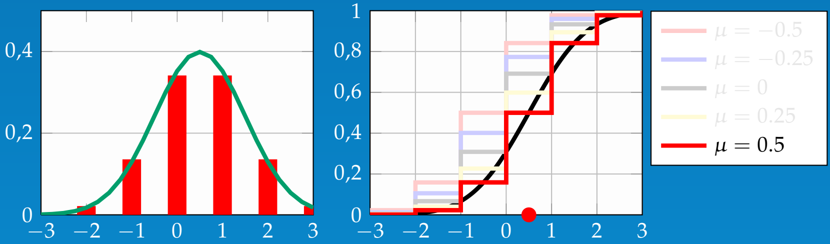




Diskretizované normální rozdělení

Tvar rozdělení X_d závisí i na μ (hlavně při menších σ)

- pravděpodobnosti a distribuční funkce $\mathrm{N}(\mu, 1)$ pro různá μ



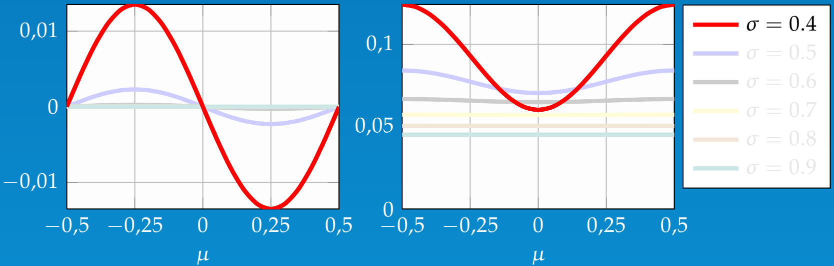
Momenty a parametry

Zaokrouhlením se změní momenty…

- není už $\mathrm{E}\,X_d = \mu$, $\sqrt{\operatorname{var} X_d} = \sigma$

Změna střední hodnoty a směrodatnmé odchylky při různých σ

- u μ malá, u σ větší



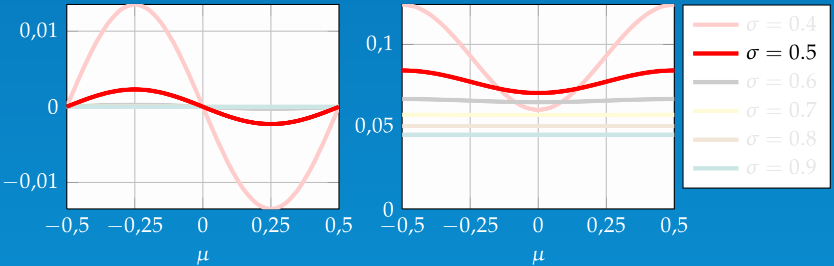
Momenty a parametry

Zaokrouhlením se změní momenty…

- není už $\operatorname{E} X_d = \mu$, $\sqrt{\operatorname{var} X_d} = \sigma$

Změna střední hodnoty a směrodatnmé odchylky při různých σ

- u μ malá, u σ větší



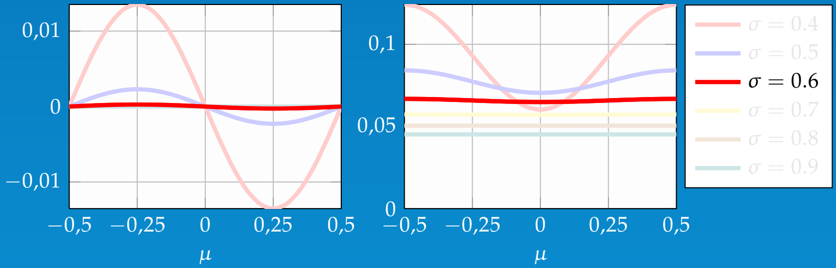
Momenty a parametry

Zaokrouhlením se změní momenty...

- není už $\mathrm{E}\,X_d = \mu$, $\sqrt{\operatorname{var} X_d} = \sigma$

Změna střední hodnoty a směrodatnmé odchylky při různých σ

- u μ malá, u σ větší



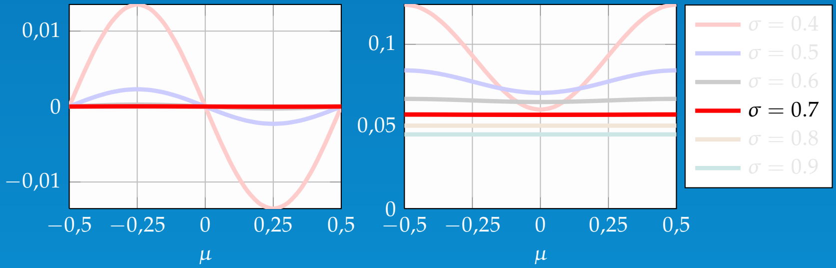
Momenty a parametry

Zaokrouhlením se změní momenty…

- není už $\operatorname{E} X_d = \mu$, $\sqrt{\operatorname{var} X_d} = \sigma$

Změna střední hodnoty a směrodatnmé odchylky při různých σ

- u μ malá, u σ větší



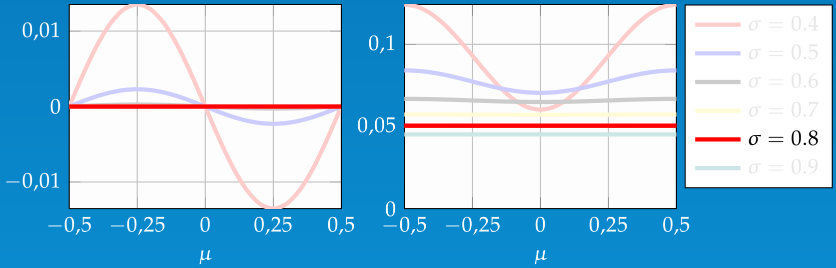
Momenty a parametry

Zaokrouhlením se změní momenty...

- není už $\mathrm{E}\,X_d = \mu$, $\sqrt{\mathrm{var}\,X_d} = \sigma$

Změna střední hodnoty a směrodatnmé odchylky při různých σ

- u μ malá, u σ větší



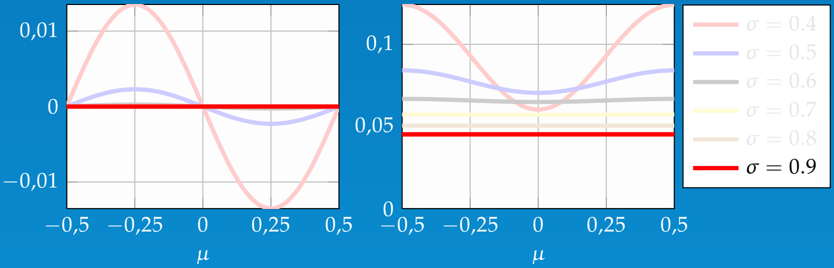
Momenty a parametry

Zaokrouhlením se změní momenty…

- není už $\mathrm{E}\, X_d = \mu$, $\sqrt{\mathrm{var}\, X_d} = \sigma$

Změna střední hodnoty a směrodatnmé odchylky při různých σ

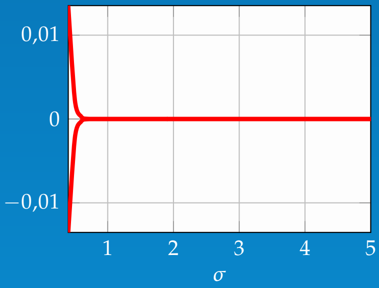
- u μ malá, u σ větší



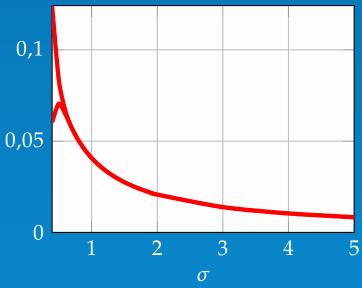
Změna momentů celkově

Největší a nejmenší změna při daném σ

Střední hodnota

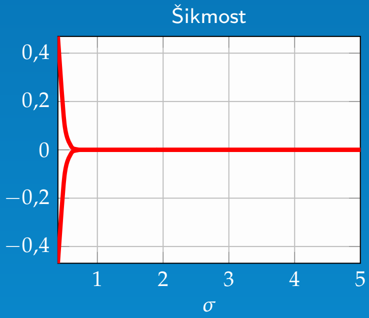


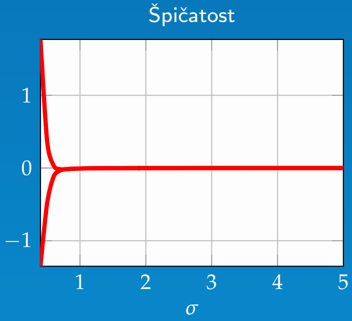
Směrodatná odchylka



Změna momentů celkově

Největší a nejmenší změna při daném σ





Odhady parametů

Výběrové momenty

- ze spojitých dat (nedostupné)
- ze zaokrouhlených dat — odhadují momenty různé od parametrů

Obecné přístupy

- metodou maximální věrohodosti
- metodou minimálního χ^2, resp. modifikovanou
  — nutno řešit numericky

Jednoduché

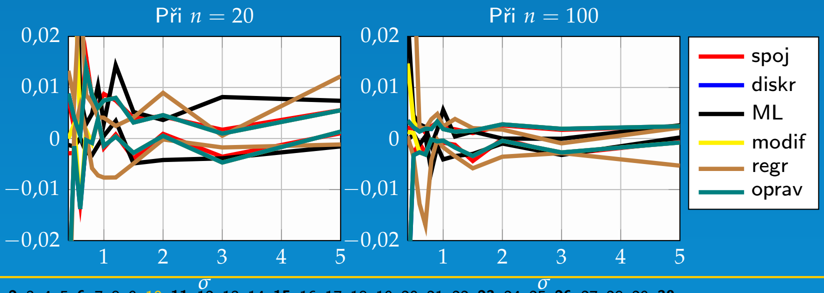
- regresně z QQ grafu
- výběrové momenty s opravou u s_d

Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \ldots, 0{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Rozmezí výchylek odhadů parametru μ v závislosti na σ:

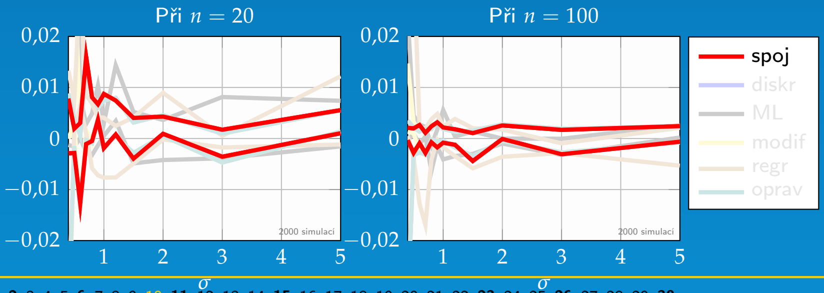


Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \ldots, 0{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Rozmezí výchylek odhadů parametru μ v závislosti na σ:

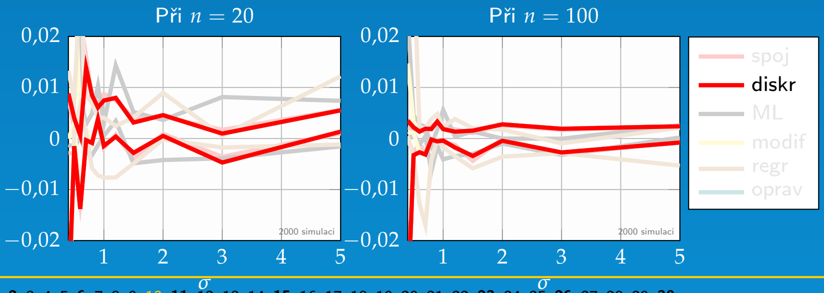


Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \ldots, 0{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Rozmezí výchylek odhadů parametru μ v závislosti na σ:

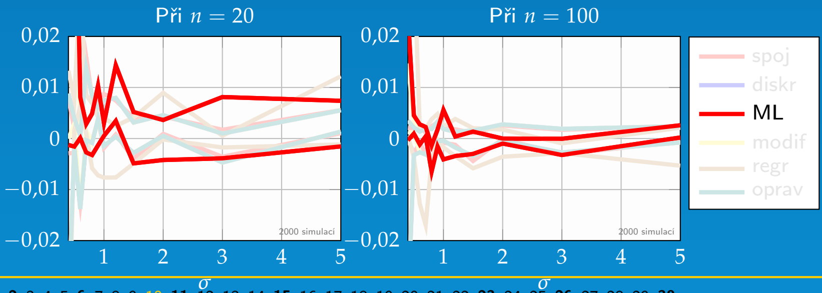


Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \ldots, 0{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Rozmezí výchylek odhadů parametru μ v závislosti na σ:

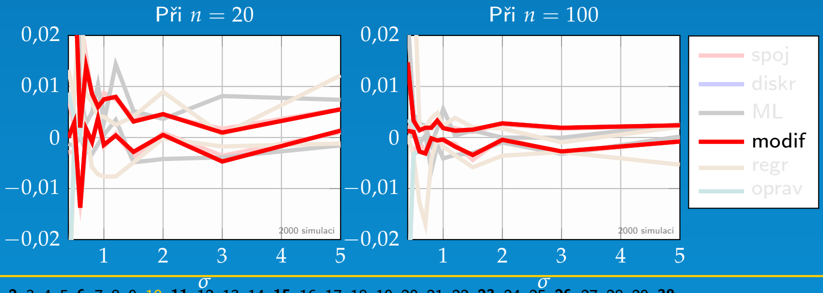


Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \ldots, 0{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Rozmezí výchylek odhadů parametru μ v závislosti na σ:

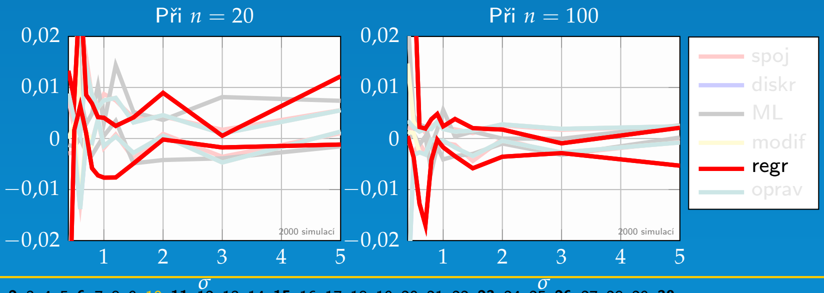


Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \dots, 0{,}5$
- $\sigma = 0{,}4, \dots, 5$

Rozmezí výchylek odhadů parametru μ v závislosti na σ:

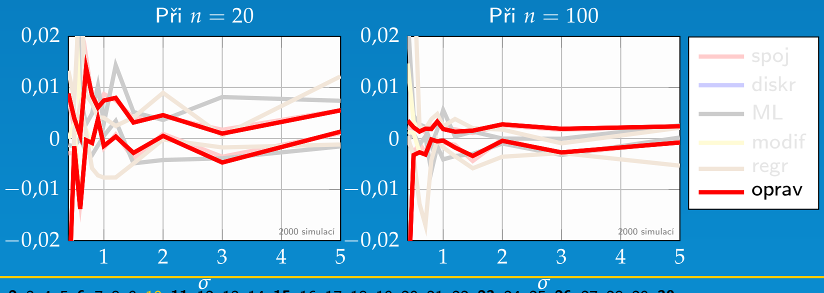


Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \ldots, 0{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Rozmezí výchylek odhadů parametru μ v závislosti na σ:

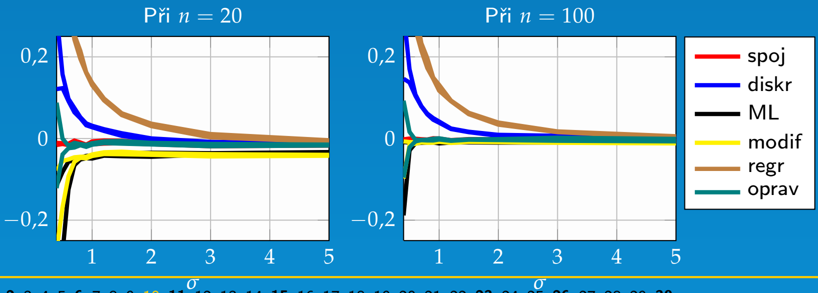


Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \dots, 0{,}5$
- $\sigma = 0{,}4, \dots, 5$

Rozmezí výchylek odhadů parametru σ v závislosti na σ:



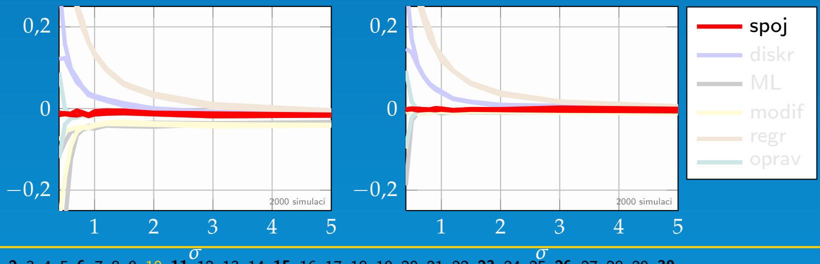
Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \dots, 0{,}5$
- $\sigma = 0{,}4, \dots, 5$

Rozmezí výchylek odhadů parametru σ v závislosti na σ:

Při $n = 20$ Při $n = 100$

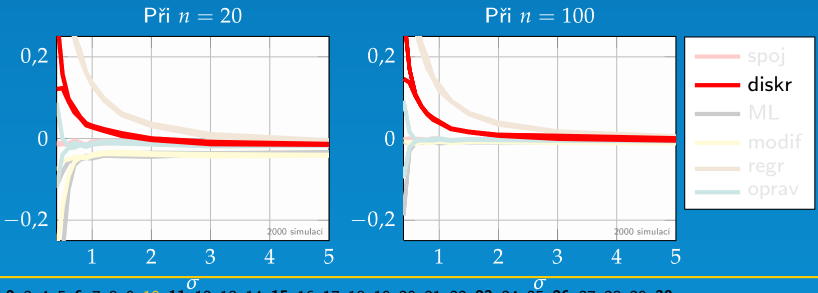


Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \ldots, 0{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Rozmezí výchylek odhadů parametru σ v závislosti na σ:

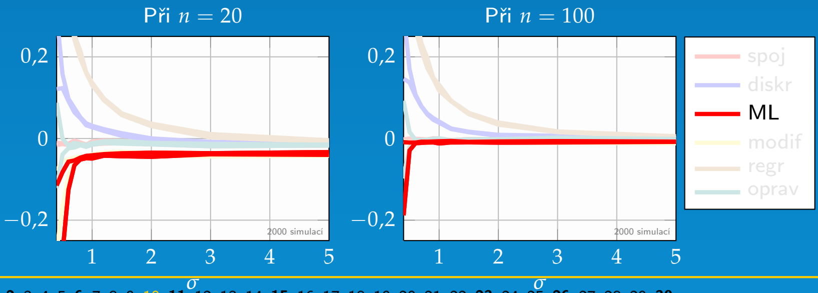


Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \ldots, 0{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Rozmezí výchylek odhadů parametru σ v závislosti na σ:

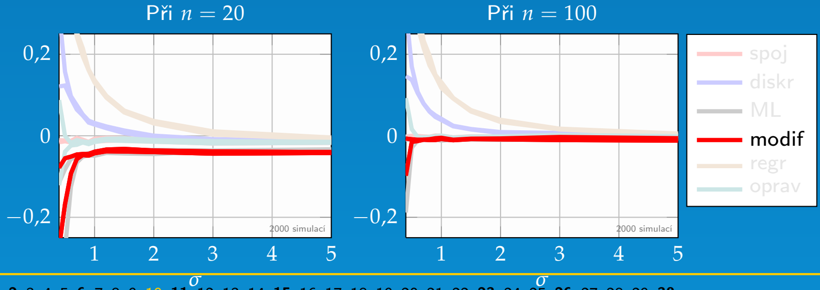


Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \ldots, 0{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Rozmezí výchylek odhadů parametru σ v závislosti na σ:

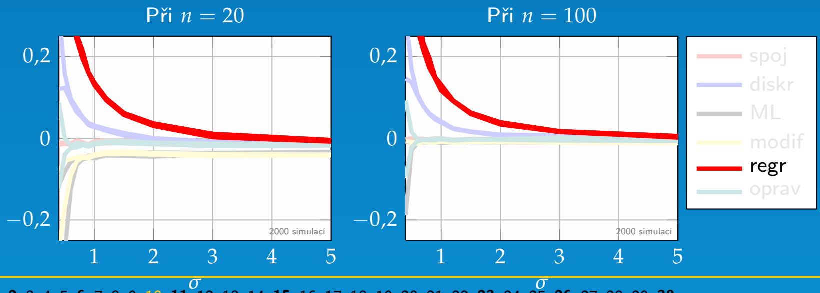


Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \ldots, 0{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Rozmezí výchylek odhadů parametru σ v závislosti na σ:



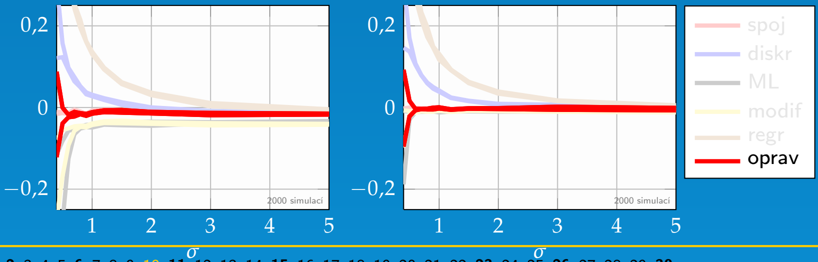
Porovnání odhadů

Nasimulované odhady z $n = 20$ a $n = 100$ diskrétních pozorování při

- $\mu = -0{,}5, \ldots, 0{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Rozmezí výchylek odhadů parametru σ v závislosti na σ:

Při $n = 20$ Při $n = 100$



Chí-kvadrát test

Pearsonův chí-kvadrát test dobré shody

- pro data rozdělená do k tříd
- rozdíly mezi pozorovanými (n_i) a očekávanými (o_i) četnostmi

$$\chi^2 = \sum \frac{(n_i - o_i)^2}{o_i}$$

Kritická hodnota

- asymptoticky při $n \to \infty$ kvantil χ^2_{k-1} rozdělení
- nezáleží na pořádí tříd

V případě neznámých parametrů

- při výpočtu o_i dosazení vhodných odhadů neznámých parametrů
- a snížení počtu stupňů volnosti

Použít i při malém rozsahu?

Přijatelnost asymptotiky

- všechna $o_i > 5$ (nebo mírnější)
- v rovnoměrném případě stačí $n > 2k$
- při malých o_i se třídy sdružují
- alternativně u malých čísel procházení konfigurací

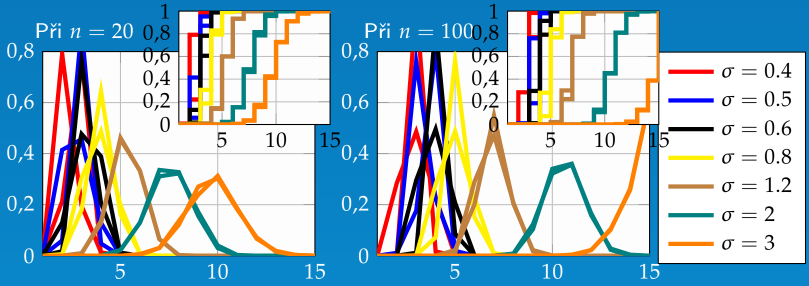
Algoritmus sdružování

- zde iteračně: od krajů třída s menší o_i se sdruží se sousední, až má $o_i > 5$
- po splnění u krajní třídy se přechází postupně dovnitř
- potřebujeme ale počet tříd $k \geqq 4$.

Počty tříd

V případě normálního rozdělení

- rozdělení počtu neprázdných tříd (před sdružováním)

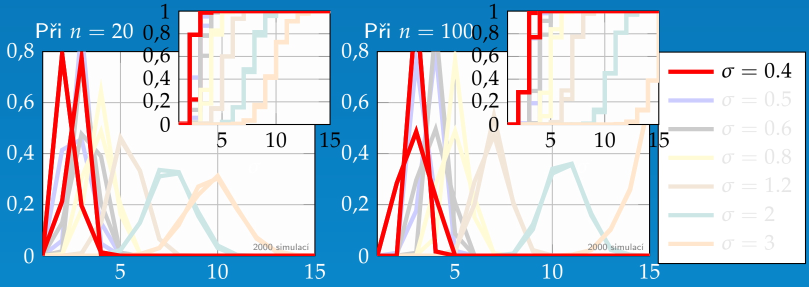


Např. při $\sigma = 0{,}6$ s 60–80 % pravděpodobností třídy méně než 4, při $\sigma = 0{,}8$ s 20–30 %

Počty tříd

V případě normálního rozdělení

- rozdělení počtu neprázdných tříd (před sdružováním)

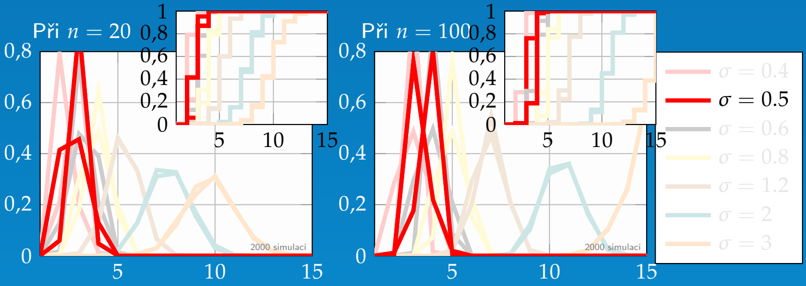


Např. při $\sigma = 0{,}6$ s 60–80 % pravděpodobností třídy méně než 4,
při $\sigma = 0{,}8$ s 20–30 %

Počty tříd

V případě normálního rozdělení

- rozdělení počtu neprázdných tříd (před sdružováním)

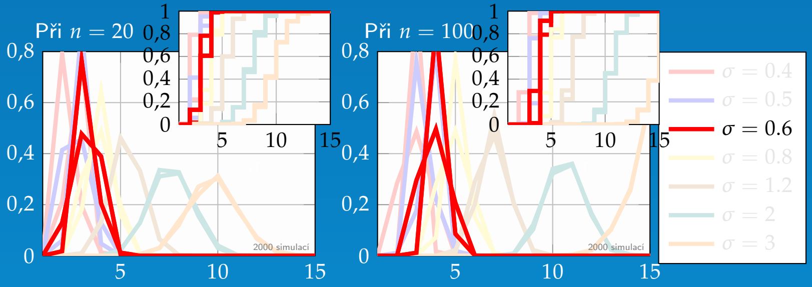


Např. při $\sigma = 0{,}6$ s 60–80 % pravděpodobností třídy méně než 4, při $\sigma = 0{,}8$ s 20–30 %

Počty tříd

V případě normálního rozdělení

- rozdělení počtu neprázdných tříd (před sdružováním)

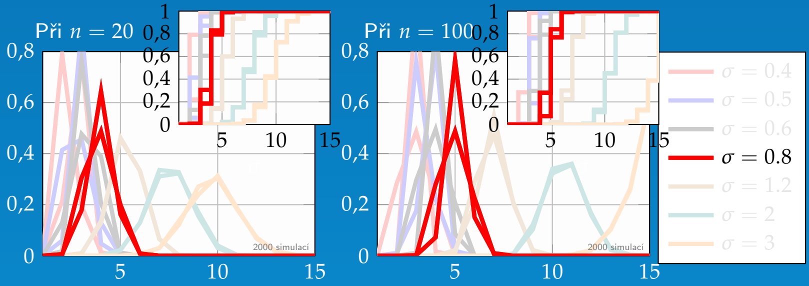


Např. při $\sigma = 0{,}6$ s 60–80 % pravděpodobností třídy méně než 4,
při $\sigma = 0{,}8$ s 20–30 %

Počty tříd

V případě normálního rozdělení

- rozdělení počtu neprázdných tříd (před sdružováním)

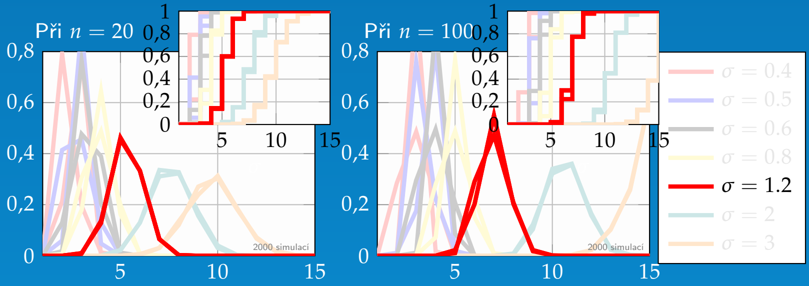


Např. při $\sigma = 0{,}6$ s 60–80 % pravděpodobností třídy méně než 4, při $\sigma = 0{,}8$ s 20–30 %

Počty tříd

V případě normálního rozdělení

- rozdělení počtu neprázdných tříd (před sdružováním)

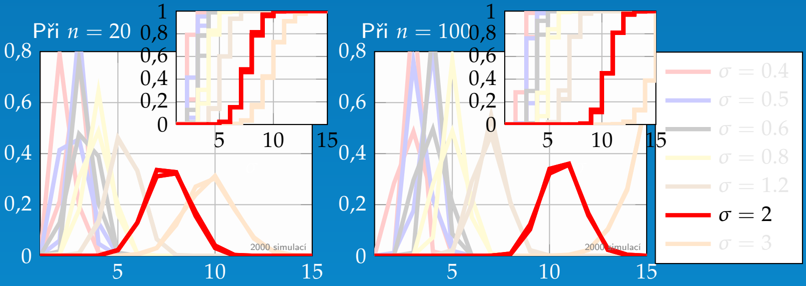


Např. při $\sigma = 0{,}6$ s 60–80 % pravděpodobností třídy méně než 4, při $\sigma = 0{,}8$ s 20–30 %

Počty tříd

V případě normálního rozdělení

- rozdělení počtu neprázdných tříd (před sdružováním)

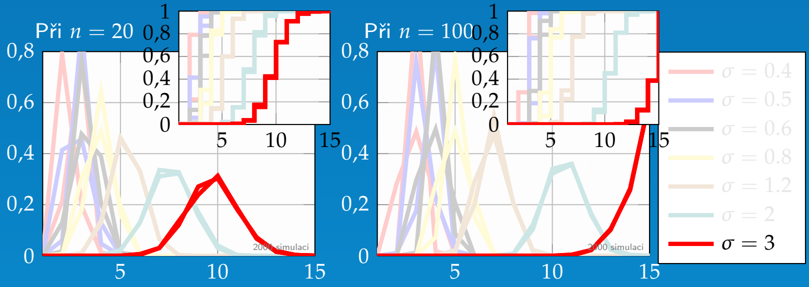


Např. při $\sigma = 0{,}6$ s 60–80 % pravděpodobností třídy méně než 4, při $\sigma = 0{,}8$ s 20–30 %

Počty tříd

V případě normálního rozdělení

- rozdělení počtu neprázdných tříd (před sdružováním)

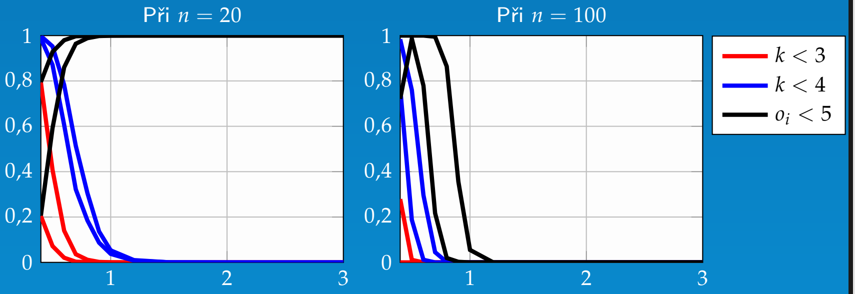


Např. při $\sigma = 0{,}6$ s 60–80 % pravděpodobností třídy méně než 4, při $\sigma = 0{,}8$ s 20–30 %

Počty tříd

V případě normálního rozdělení

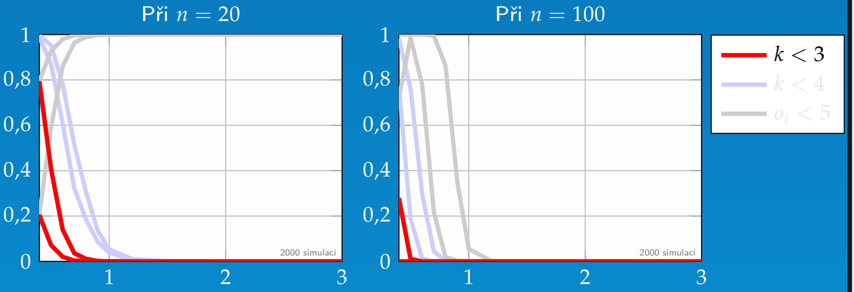
- procento výskytu malého počtu tříd
- resp. nízkých očekávaných četností po sdružení



Počty tříd

V případě normálního rozdělení

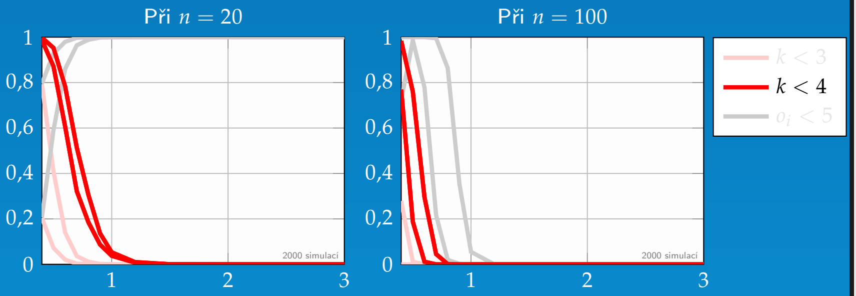
- procento výskytu malého počtu tříd
- resp. nízkých očekávaných četností po sdružení



Počty tříd

V případě normálního rozdělení

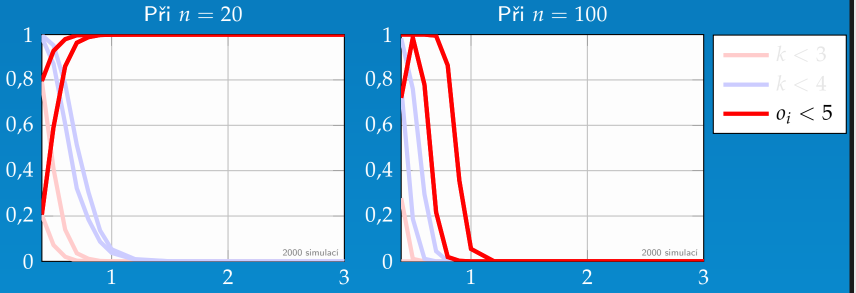
- procento výskytu malého počtu tříd
- resp. nízkých očekávaných četností po sdružení



Počty tříd

V případě normálního rozdělení

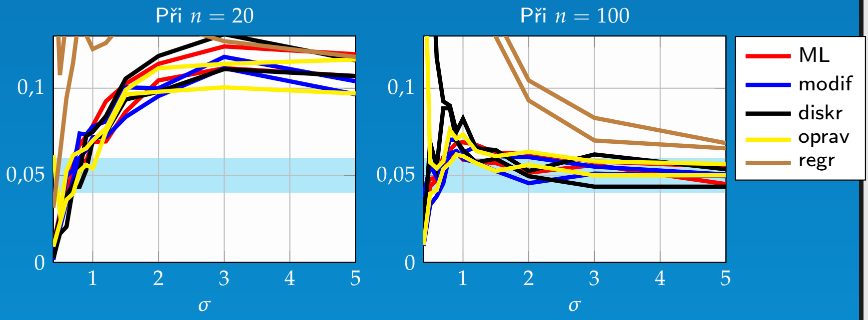
- procento výskytu malého počtu tříd
- resp. nízkých očekávaných četností po sdružení



Dosažená hladina významnosti

Pravděpodobnost zamítnutí při H_0

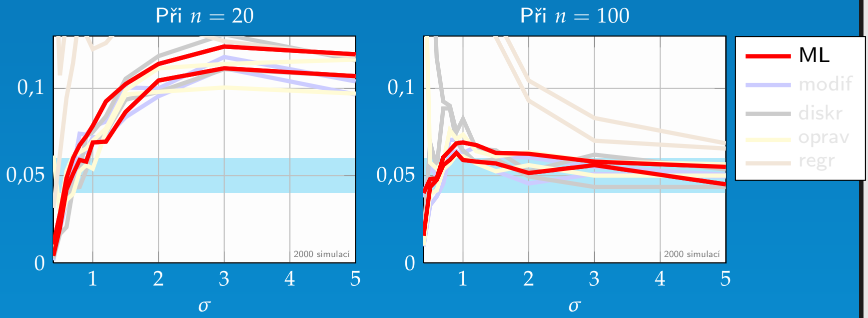
- nominální 5 %
- skutečná (v závislosti na σ):



Dosažená hladina významnosti

Pravděpodobnost zamítnutí při H_0

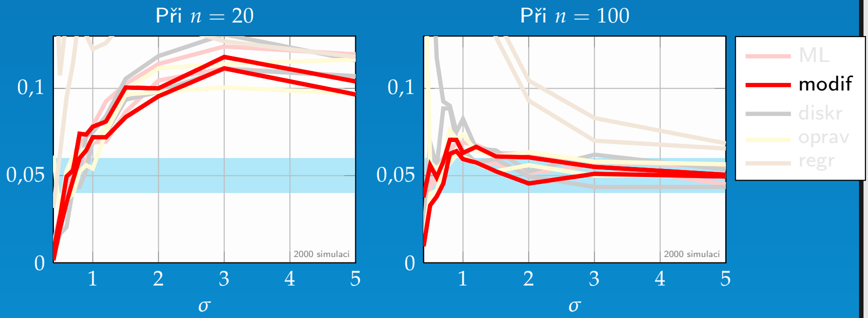
- nominální 5 %
- skutečná (v závislosti na σ):



Dosažená hladina významnosti

Pravděpodobnost zamítnutí při H_0

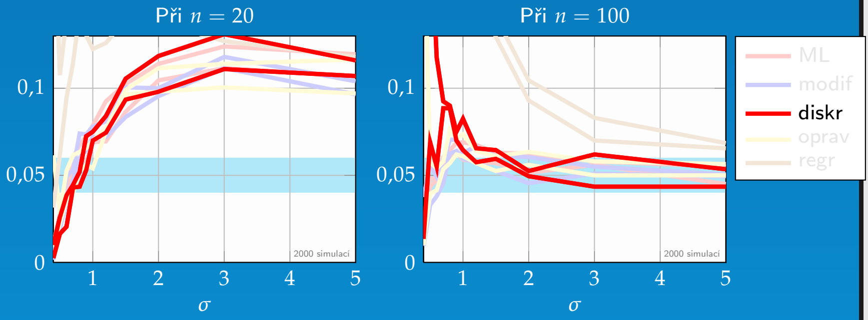
- nominální 5 %
- skutečná (v závislosti na σ):



Dosažená hladina významnosti

Pravděpodobnost zamítnutí při H_0

- nominální 5 %
- skutečná (v závislosti na σ):

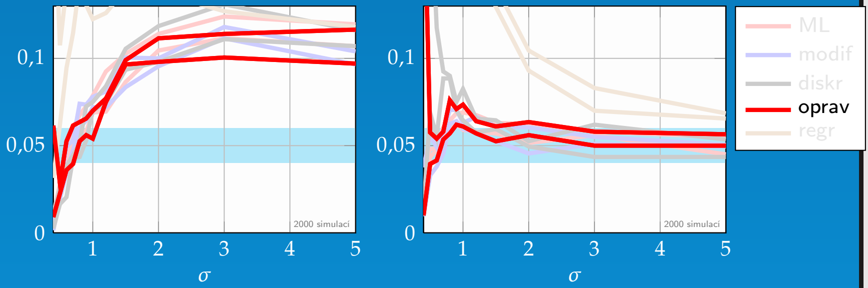


Dosažená hladina významnosti

Pravděpodobnost zamítnutí při H_0

- nominální 5 %
- skutečná (v závislosti na σ):

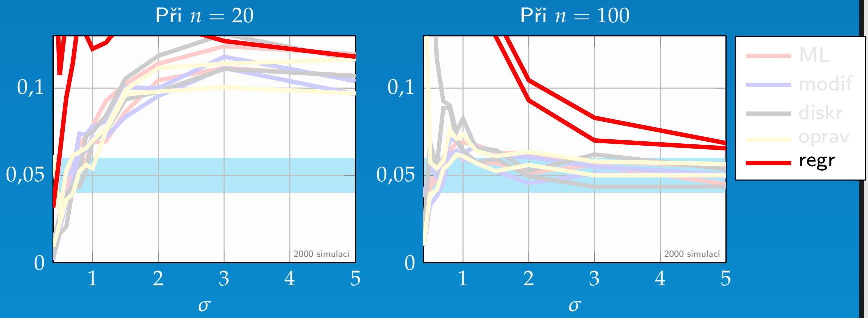
Při $n = 20$ Při $n = 100$



Dosažená hladina významnosti

Pravděpodobnost zamítnutí při H_0

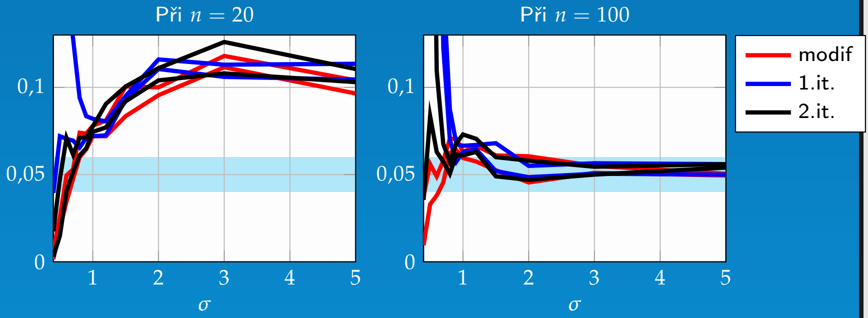
- nominální 5 %
- skutečná (v závislosti na σ):



Dosažená hladina významnosti

Vliv zjednodušených odhadů

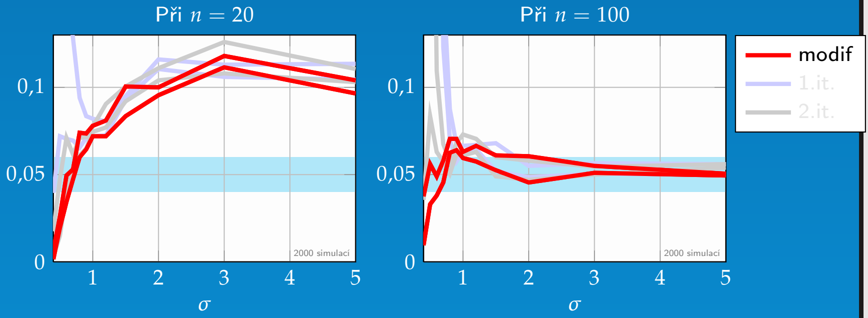
- při použití jen 1. resp. 2. iterace



Dosažená hladina významnosti

Vliv zjednodušených odhadů

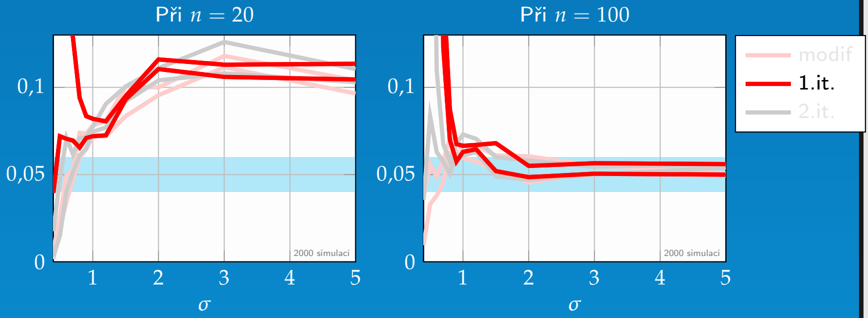
- při použití jen 1. resp. 2. iterace



Dosažená hladina významnosti

Vliv zjednodušených odhadů

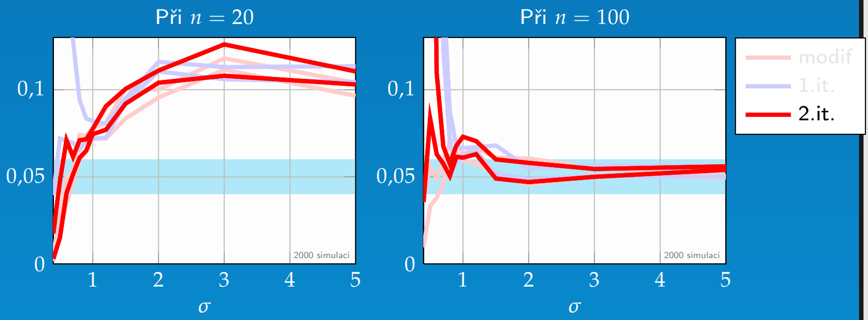
- při použití jen 1. resp. 2. iterace



Dosažená hladina významnosti

Vliv zjednodušených odhadů

- při použití jen 1. resp. 2. iterace

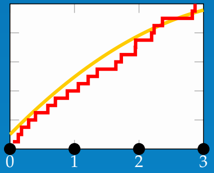


Kolmogorovův-Smirnovův test

Kolmogorovův-Smirnovův test $H_0 : F = F_0$ pro spojitá data

- rozdíl mezi empirickou (F_n) a teoretickou (F_0) distribuční funkcí
- $KS = \sup_{-\infty<x<\infty} \left| F_0(x) - F_n(x) \right|$

$$= \max_{1\leq i\leq n} \left| F_0(X_{(i)}) - \underbrace{F_n(X_{(i)}\pm)}_{i/n \text{ a } (i-1)/n} \right|$$



Kritická hodnota

- pro konečná n tabulky, $\sqrt{n}KS$ asymptoticky kvantil Brownova mostu
- stejná pro všechna rozdělení F

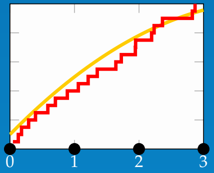
Podobně test Cramérův-von Misesův, Andersenův-Darlingův,...

$$\int (F_n - F_0)^2 \,\mathrm{d}F_0, \qquad \int \frac{(F_n - F_0)^2}{F_0(1-F_0)} \,\mathrm{d}F_0$$

Kolmogorovův-Smirnovův test

Kolmogorovův-Smirnovův test $H_0 : F = F_0$ pro spojitá data

- rozdíl mezi empirickou (F_n) a teoretickou (F_0) distribuční funkcí
- $$KS = \sup_{-\infty<x<\infty} \left| F_0(x) - F_n(x) \right|$$
  $$= \max_{1\leqq i\leqq n} \left| F_0(X_{(i)}) - \underbrace{F_n(X_{(i)}\pm)}_{i/n \text{ a } (i-1)/n} \right|$$



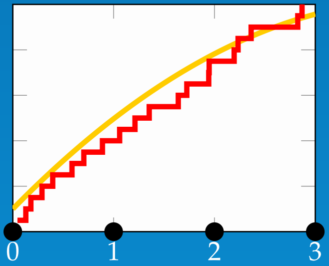
V případě neznámých parametrů měřítka či polohy

- do F dosazení odhadů invariantních ke změně měřítka a polohy
- pro každý typ rozdělení/odhadů jiné kritické hodnoty, ale nezávislé na parametrech F (Lillieforsův test)

Použít pro diskrétní data?

Místo nedostupné F_n použít F_n^d spočtenou z diskrétních dat

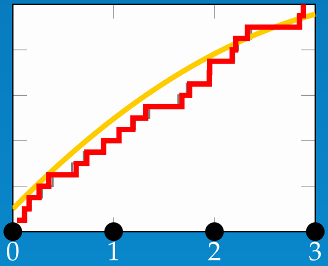
$$KS^d = F_n^d - F_0$$



Použít pro diskrétní data?

Místo nedostupné F_n použít F_n^d spočtenou z diskrétních dat

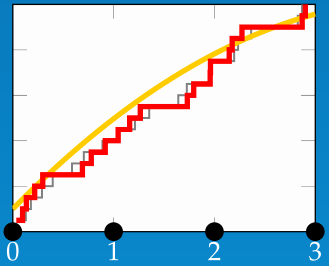
$$KS^d = F_n^d - F_0$$



Použít pro diskrétní data?

Místo nedostupné F_n použít F_n^d spočtenou z diskrétních dat

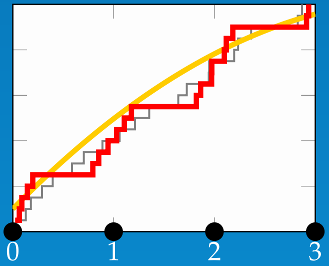
$$KS^d = F_n^d - F_0$$



Použít pro diskrétní data?

Místo nedostupné F_n použít F_n^d spočtenou z diskrétních dat

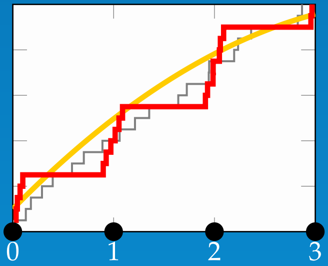
$$KS^d = F_n^d - F_0$$



Použít pro diskrétní data?

Místo nedostupné F_n použít F_n^d spočtenou z diskrétních dat

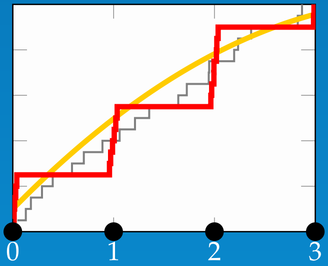
$$KS^d = F_n^d - F_0$$



Použít pro diskrétní data?

Místo nedostupné F_n použít F_n^d spočtenou z diskrétních dat

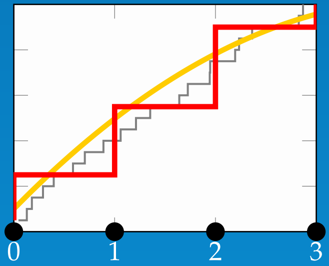
$$KS^d = F_n^d - F_0$$



Použít pro diskrétní data?

Místo nedostupné F_n použít F_n^d spočtenou z diskrétních dat

$$KS^d = F_n^d - F_0$$



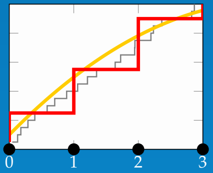
Použít pro diskrétní data?

Místo nedostupné F_n použít F_n^d spočtenou z diskrétních dat

$$KS^d = F_n^d - F_0 = \underbrace{F_n^d - F_n}_{D} + \underbrace{F_n - F_0}_{S} \approx \frac{\pm \mathrm{N}(p\sqrt{n}, p(1-p))}{\sqrt{n}} + \frac{BR}{\sqrt{n}}$$

kde

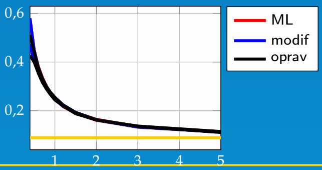
$$D \sim \pm \frac{1}{n} \mathrm{Bi}(n, \underbrace{|F_0^d(x) - F_0(x)|}_{p})$$



Dle závislosti šířky intervalu na počtu pozorování n

- $\lessapprox c/\sqrt{n}$: rozhodne S
- $\approx c/\sqrt{n}$: nějaké limitní rozdělení KS^d
- $\gtrapprox c/\sqrt{n}$ (náš případ): rozhodne D

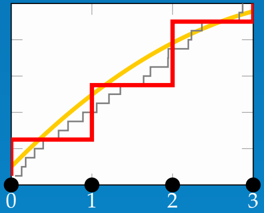
Při $n = 100$



Kritické hodnoty při $\alpha = 5\,\%$

Další varianty

Další možnosti porovnání F_n a F_0:

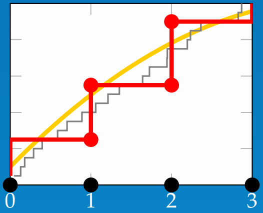


Zprůměrování (KSH)

- místo „vzdálených“ limit zleva/zprava se vezme jejich průměr

Další varianty

Další možnosti porovnání F_n a F_0:

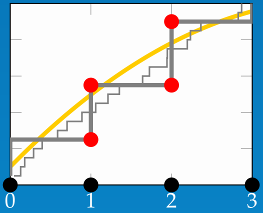


Zprůměrování (KSH)

- místo „vzdálených“ limit zleva/zprava se vezme jejich průměr

Další varianty

Další možnosti porovnání F_n a F_0:

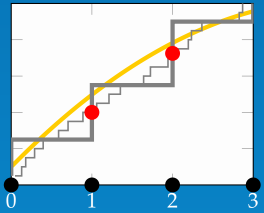


Zprůměrování (KSH)

- místo „vzdálených“ limit zleva/zprava se vezme jejich průměr

Další varianty

Další možnosti porovnání F_n a F_0:

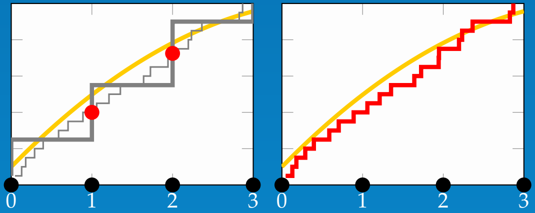


Zprůměrování (KSH)

- místo „vzdálených“ limit zleva/zprava se vezme jejich průměr

Další varianty

Další možnosti porovnání F_n a F_0:

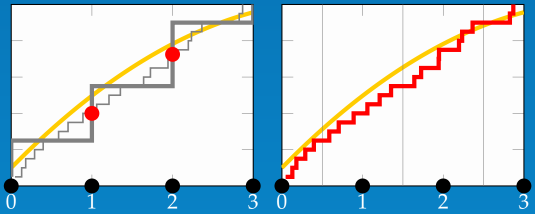


Kolmogorovův-Smirnovův test pro diskrétní rozdělení (KSD)

- ve skutečnosti známe přírůstky F_n přes intervaly
- porovnají se F_n a F_0 v krajních bodech intervalů
- tj. kumulované relativní četnosti a pravděpodobnosti tříd

Další varianty

Další možnosti porovnání F_n a F_0:

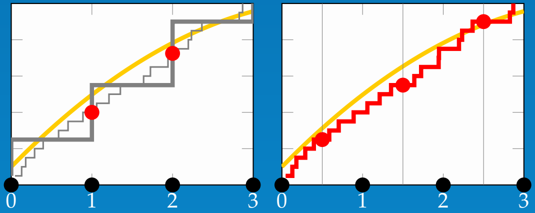


Kolmogorovův-Smirnovův test pro diskrétní rozdělení (KSD)

- ve skutečnosti známe přírůstky F_n přes intervaly
- porovnají se F_n a F_0 v krajních bodech intervalů
- tj. kumulované relativní četnosti a pravděpodobnosti tříd

Další varianty

Další možnosti porovnání F_n a F_0:

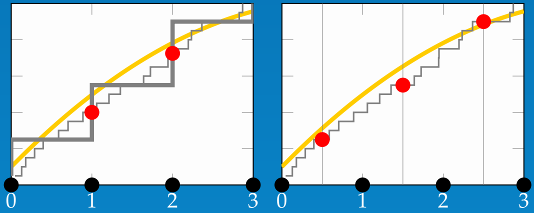


Kolmogorovův-Smirnovův test pro diskrétní rozdělení (KSD)

- ve skutečnosti známe přírůstky F_n přes intervaly
- porovnají se F_n a F_0 v krajních bodech intervalů
- tj. kumulované relativní četnosti a pravděpodobnosti tříd

Další varianty

Další možnosti porovnání F_n a F_0:

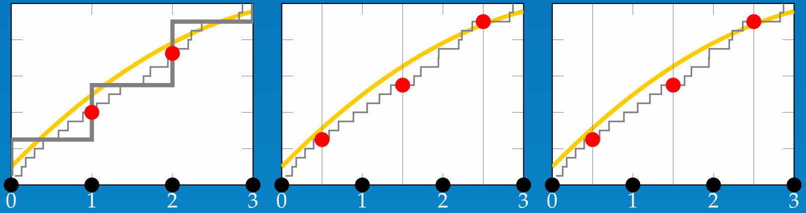


Kolmogorovův-Smirnovův test pro diskrétní rozdělení (KSD)

- ve skutečnosti známe přírůstky F_n přes intervaly
- porovnají se F_n a F_0 v krajních bodech intervalů
- tj. kumulované relativní četnosti a pravděpodobnosti tříd

Další varianty

Další možnosti porovnání F_n a F_0:

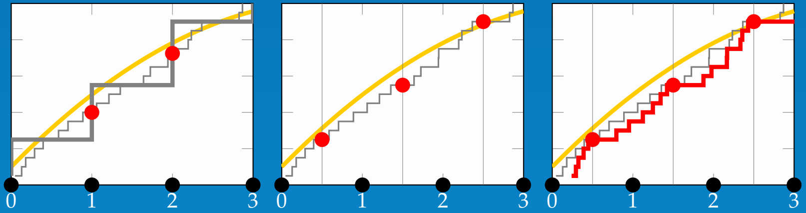


Randomizovaná verze (KSR)

- povede na „dogenerování" neznámého průběhu F_n mezi krajními body intervalů

Další varianty

Další možnosti porovnání F_n a F_0:



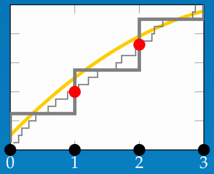
Randomizovaná verze (KSR)

- povede na „dogenerování" neznámého průběhu F_n mezi krajními body intervalů

KSH - zprůměrování

Zprůměrování limity F_n zprava a zleva

- $$KSH = \sup_i \left| F_0(X_{(i)}) - \frac{F_n(X_{(i)}+) + F_n(X_i-)}{2} \right|$$
  $$= \sup_i \left| F_0(X_{(i)}) - \frac{(i-1)/n + i/n}{2} \right|$$



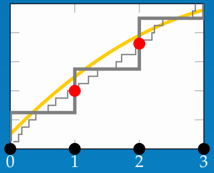
Oproti KSd

- sice bližší shoda s F_0
- není ale ani nyní $KSH \to 0$ při $n \to \infty$

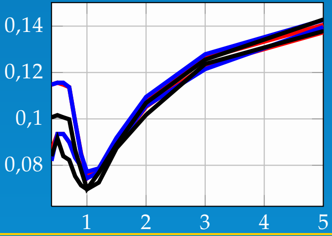
KSH - zprůměrování

Zprůměrování limity F_n zprava a zleva

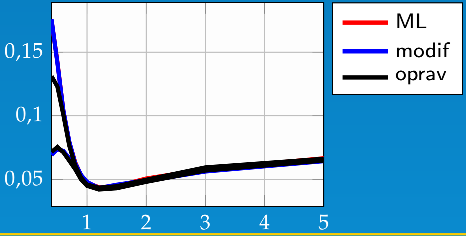
- $$KSH = \sup_i \left| F_0(X_{(i)}) - \frac{F_n(X_{(i)}+) + F_n(X_i-)}{2} \right|$$
  $$= \sup_i \left| F_0(X_{(i)}) - \frac{(i-1)/n + i/n}{2} \right|$$



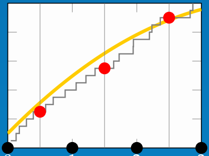
Při $n = 20$



Při $n = 100$



KSD - diskrétní verze KS



Sledují se krajní body intervalů

- $$KSD = \sup_j \left| F_n(d_{(j)}) - F_0(d_{(j)}) \right|$$
  $$= \max_i \left| F_n\left(X_{(i)} \pm \frac{1}{2}\right) - F_0\left(X_{(i)} \pm \frac{1}{2}\right) \right|$$

  když $-\infty = d_0 < d_1 < \cdots < d_{k-1} < d_k = \infty$ jsou dělící body

Oproti verzi pro spojitá data

- porovnání probíhá v méně bodech, proto jiné kritické hodnoty
- pro každé dělení a pro každé rozdělení (a jeho parametry) jiné

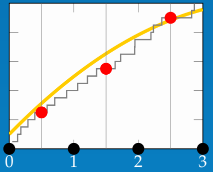
Při neznámých parametrech

- použít naivně kritické hodnoty KSD počítané s odhady pro $\sigma = \widehat{\sigma}_d$
- skutečná hladina významnosti může být jiná

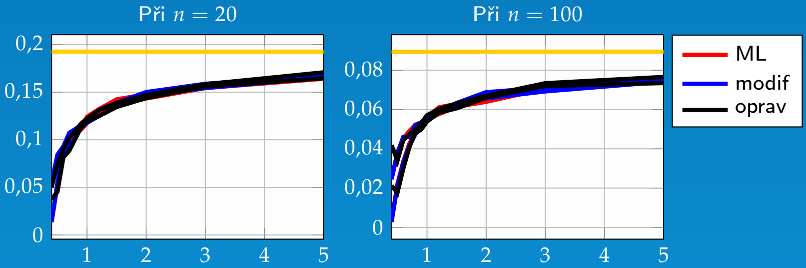
KSD - diskrétní verze KS

Sledují se krajní body intervalů

- $$KSD = \sup_j \left| F_n(d_{(j)}) - F_0(d_{(j)}) \right| = \max_i \left| F_n\left(X_{(i)} \pm \frac{1}{2}\right) - F_0\left(X_{(i)} \pm \frac{1}{2}\right) \right|$$



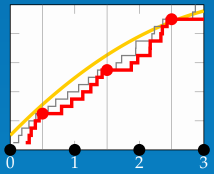
Kritické hodnoty při $\alpha = 5\,\%$



KSR - randomizovaný test

Zamítnout $H_0 : F = F_0$ s pravděpodododobností

- $\mathrm{P}\big(KS(X_1, \dots, X_n) > KS_{\mathsf{krit}} \mid X_1^d, \dots, X_n^d\big)$



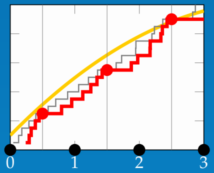
Prakticky

- pravděpodobnost sice neznáme
- umíme ale rozhodnutí snadno vygenerovat
- přesná hladina významnosti

Postup

- známe počty pozorování v intervalech
- dle $\mathrm{N}(\mu, \sigma^2)$ přiřadit konkrétní hodnoty z intervalu
- na dogenerovaná pozorování standardní spojitý KS test

KSR - randomizovaný test



Zamítnout $H_0 : F = F_0$ s pravděpododobností

- $\mathrm{P}\big(KS(X_1, \ldots, X_n) > KS_{\mathsf{krit}} \mid X_1^d, \ldots, X_n^d\big)$

Při neznámých parametrech

- použít odhady a Lillieforsovy kritické hodnoty
- neumíme ale dogenerovat dle $\mathrm{N}(\mu, \sigma^2)$
- nezbývá než generovat dle $\mathrm{N}(\widetilde{\mu}, \widetilde{\sigma}^2)$

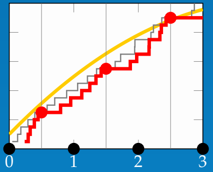
Tj.

- místo $F_{\widehat{\mu}, \widehat{\sigma}^2}(X_{(i)}) = \Phi\left(\dfrac{\mu + \sigma u_{R_{(i)}} - \widehat{\mu}}{\widehat{\sigma}}\right) = \Phi\left(\dfrac{\mu - \widehat{\mu}}{\widehat{\sigma}} + \dfrac{\sigma}{\widehat{\sigma}} u_{R_{(i)}}\right)$
  v KS něco trochu jiného
- změní se skutečná hladina významnosti, snad málo

KSR - randomizovaný test

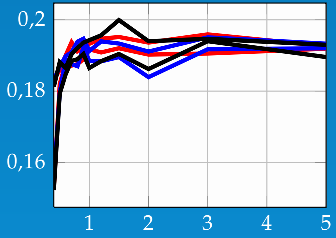
Zamítnout $H_0 : F = F_0$ s pravděpododobností

- $\mathrm{P}\big(KS(X_1, \ldots, X_n) > KS_{\text{krit}} \mid X_1^d, \ldots, X_n^d\big)$

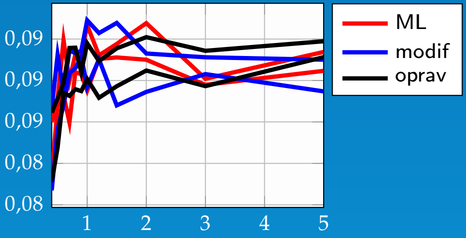


Kritické hodnoty při $\alpha = 5\,\%$:

Při $n = 20$



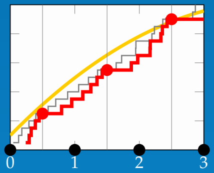
Při $n = 100$



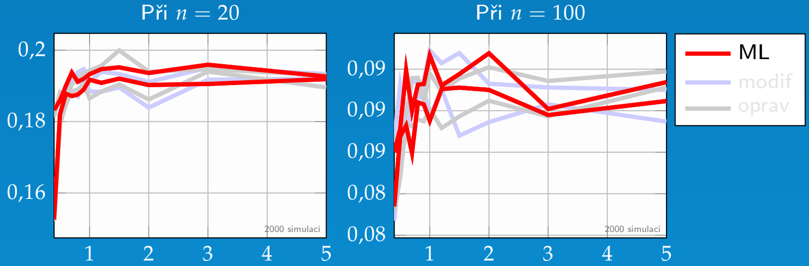
KSR - randomizovaný test

Zamítnout $H_0 : F = F_0$ s pravděpododobností

- $\mathrm{P}\big(KS(X_1, \ldots, X_n) > KS_{\mathsf{krit}} \mid X_1^d, \ldots, X_n^d\big)$



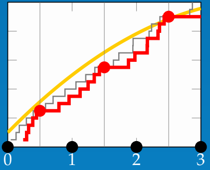
Kritické hodnoty při $\alpha = 5\,\%$:



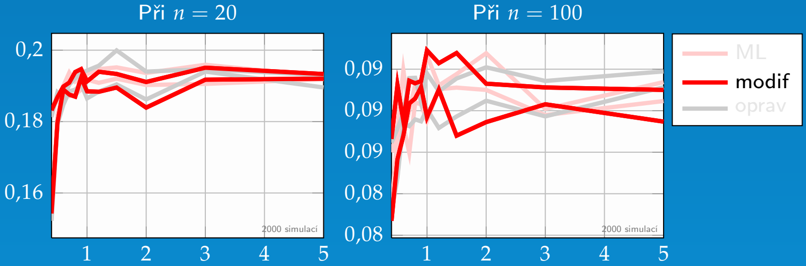
KSR - randomizovaný test

Zamítnout $H_0 : F = F_0$ s pravděpododobností

- $\mathrm{P}\big(KS(X_1, \ldots, X_n) > KS_{\text{krit}} \mid X_1^d, \ldots, X_n^d\big)$



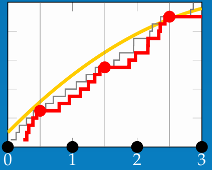
Kritické hodnoty při $\alpha = 5\,\%$:



KSR - randomizovaný test

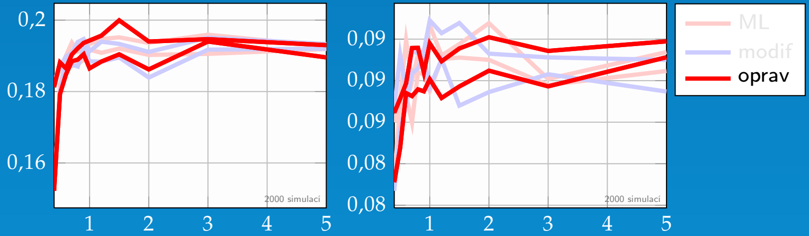
Zamítnout $H_0 : F = F_0$ s pravděpododobností

- $\mathrm{P}\big(KS(X_1, \ldots, X_n) > KS_{\text{krit}} \mid X_1^d, \ldots, X_n^d\big)$



Kritické hodnoty při $\alpha = 5\,\%$:

Při $n = 20$ Při $n = 100$

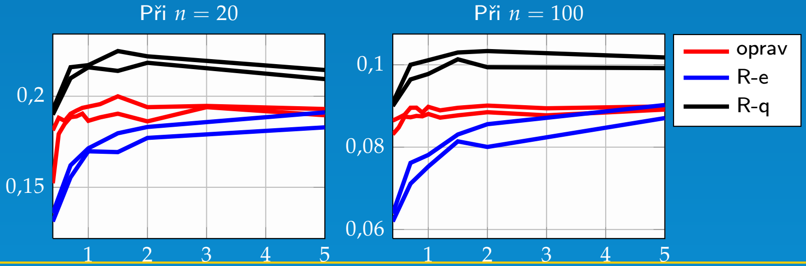


Nerandomizovaná verze

Zamítnout, když

- podmíněná střední hodnota $\mathrm{E}(KS \mid X_1^d, \ldots, X_n^d) > \mathrm{E}_{\mathsf{krit}}$
- podmíněný kvantil $> q_{\mathsf{krit}}$
- vše nutno zjistit simulacemi a závisí na μ, σ

Kritické hodnoty při $\alpha = 5\,\%$

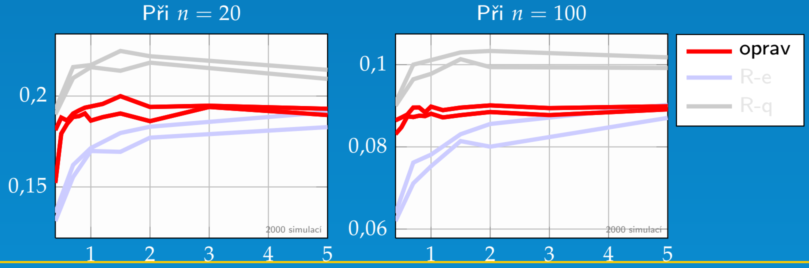


Nerandomizovaná verze

Zamítnout, když

- podmíněná střední hodnota $\mathrm{E}(KS \mid X_1^d, \ldots, X_n^d) > \mathrm{E}_{\text{krit}}$
- podmíněný kvantil $> q_{\text{krit}}$
- vše nutno zjistit simulacemi a závisí na μ, σ

Kritické hodnoty při $\alpha = 5\,\%$

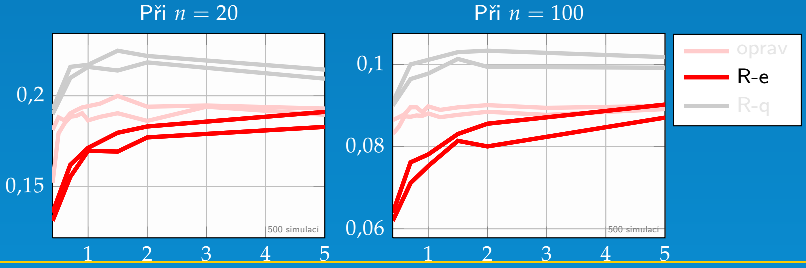


Nerandomizovaná verze

Zamítnout, když

- podmíněná střední hodnota $\mathrm{E}(KS \mid X_1^d, \ldots, X_n^d) > \mathrm{E}_{\text{krit}}$
- podmíněný kvantil $> q_{\text{krit}}$
- vše nutno zjistit simulacemi a závisí na μ, σ

Kritické hodnoty při $\alpha = 5\,\%$

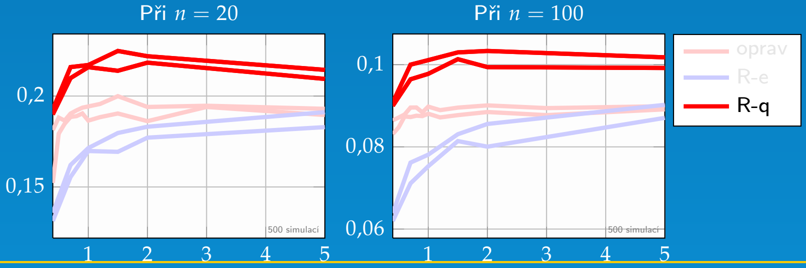


Nerandomizovaná verze

Zamítnout, když

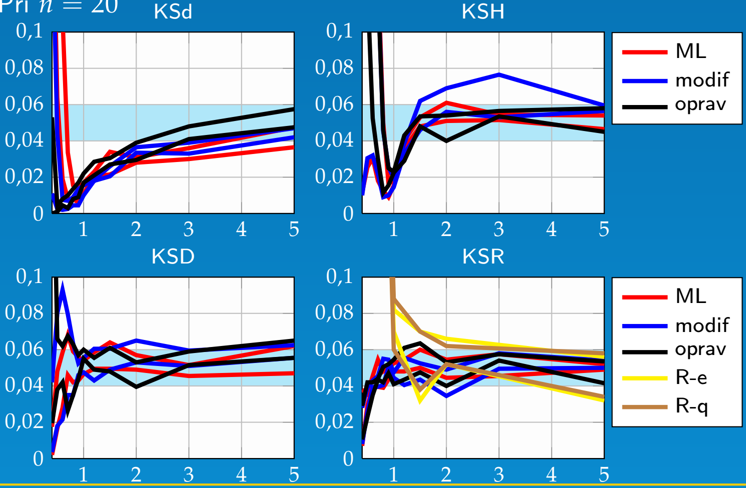
- podmíněná střední hodnota $\mathrm{E}(KS \mid X_1^d, \dots, X_n^d) > \mathrm{E}_{\text{krit}}$
- podmíněný kvantil $> q_{\text{krit}}$
- vše nutno zjistit simulacemi a závisí na μ, σ

Kritické hodnoty při $\alpha = 5\,\%$



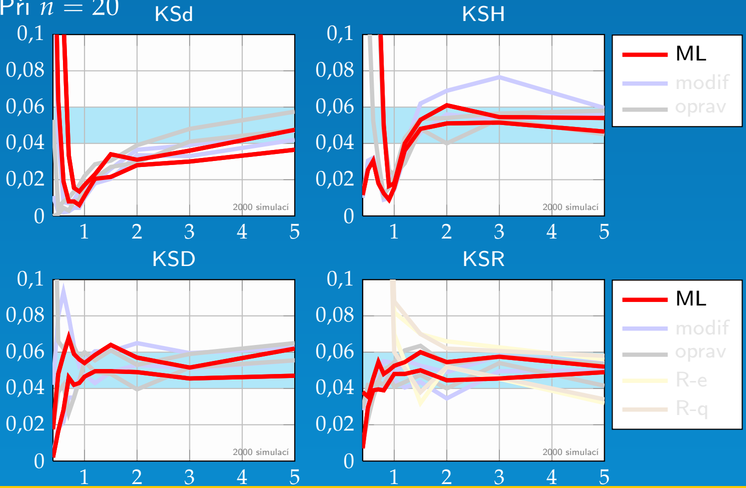
Dosažená hladina významnosti

Při $n = 20$



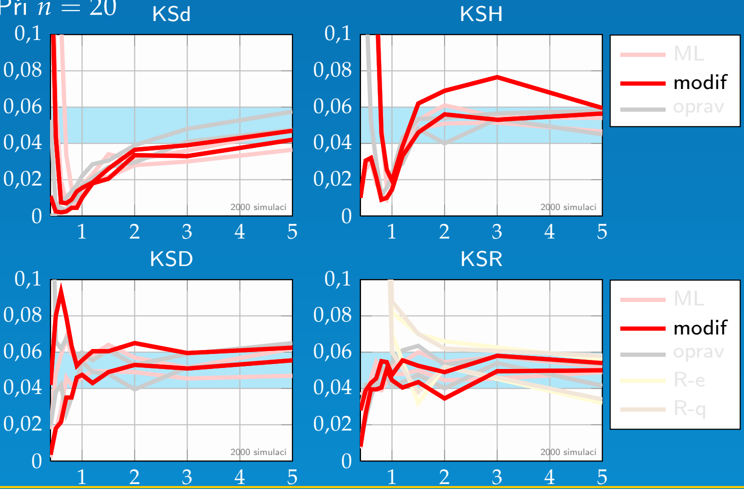
Dosažená hladina významnosti

Při $n = 20$



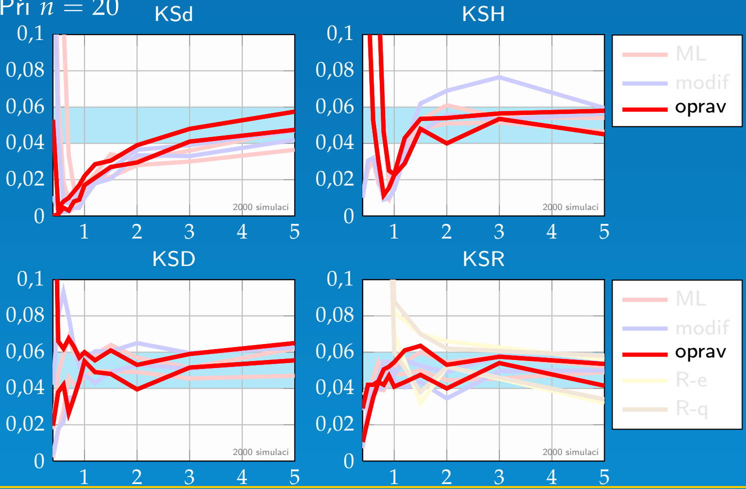
Dosažená hladina významnosti

Při $n = 20$



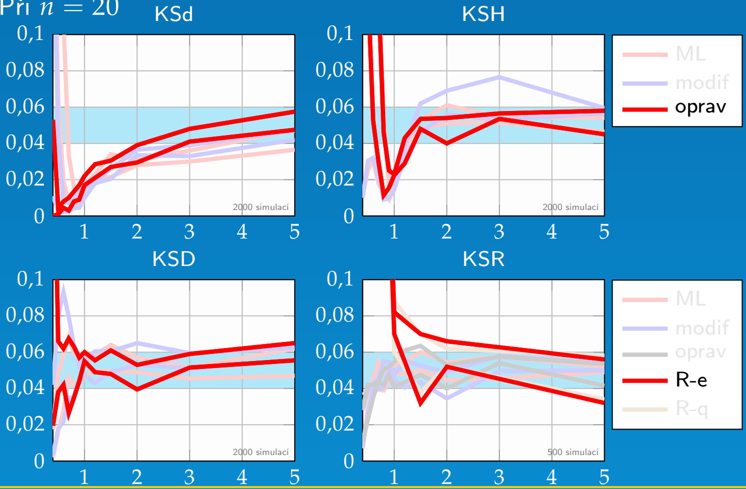
Dosažená hladina významnosti

Při $n = 20$



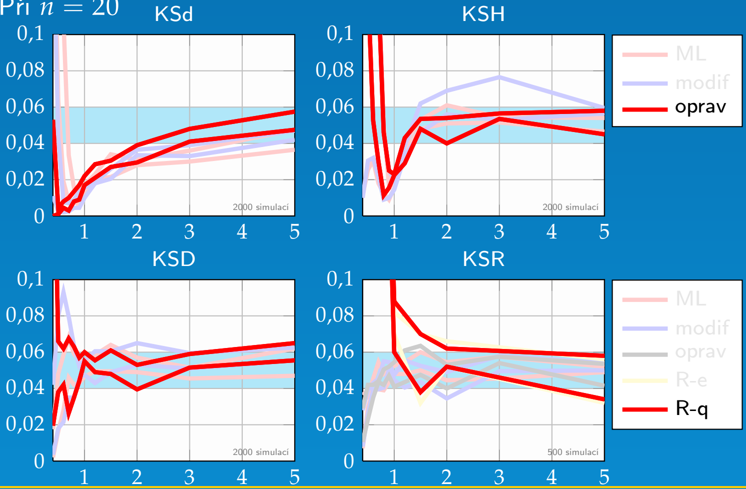
Dosažená hladina významnosti

Při $n = 20$



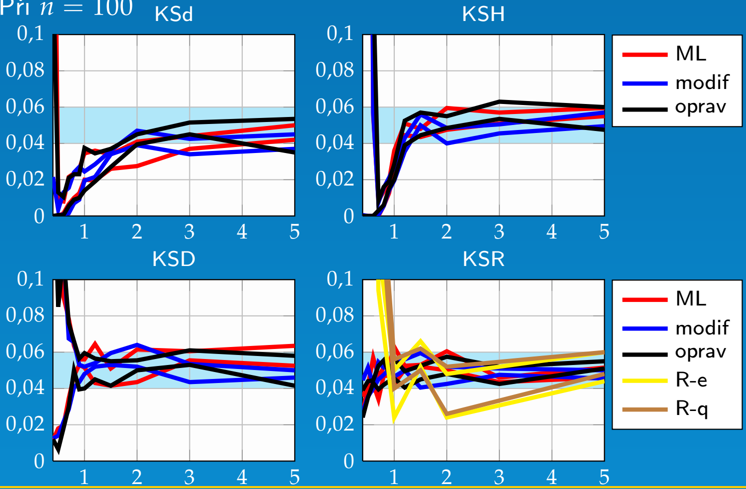
Dosažená hladina významnosti

Při $n = 20$



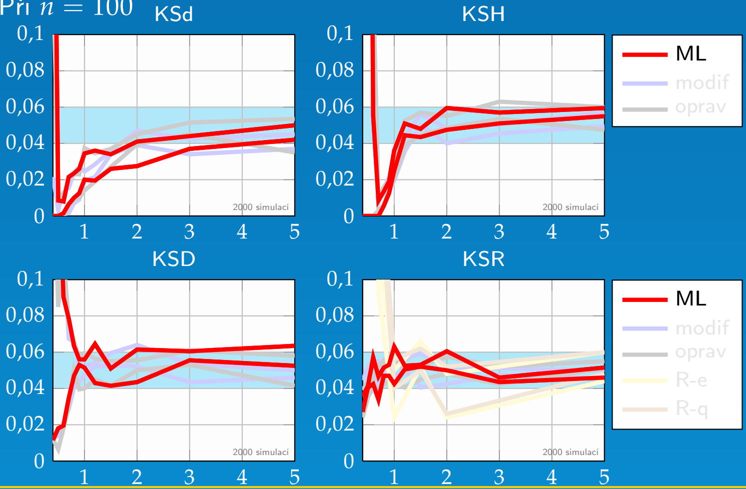
Dosažená hladina významnosti

Při $n = 100$



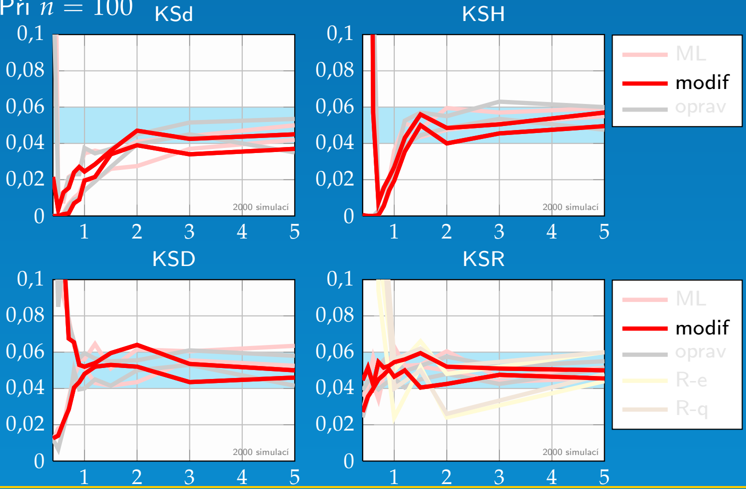
Dosažená hladina významnosti

Při $n = 100$



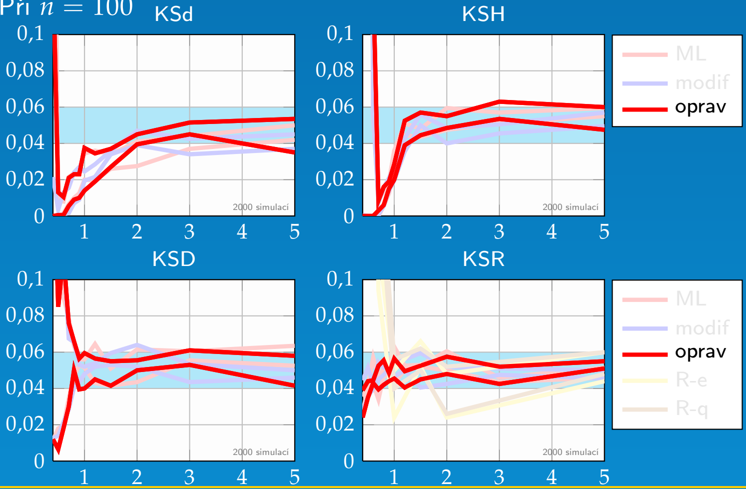
Dosažená hladina významnosti

Při $n = 100$



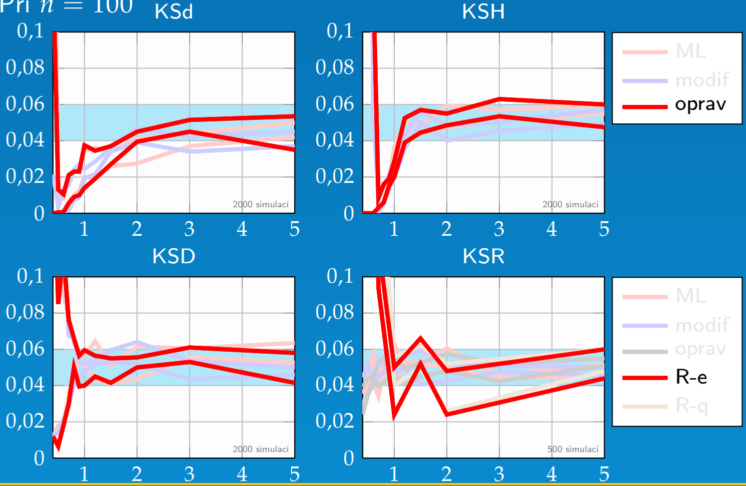
Dosažená hladina významnosti

Při $n = 100$



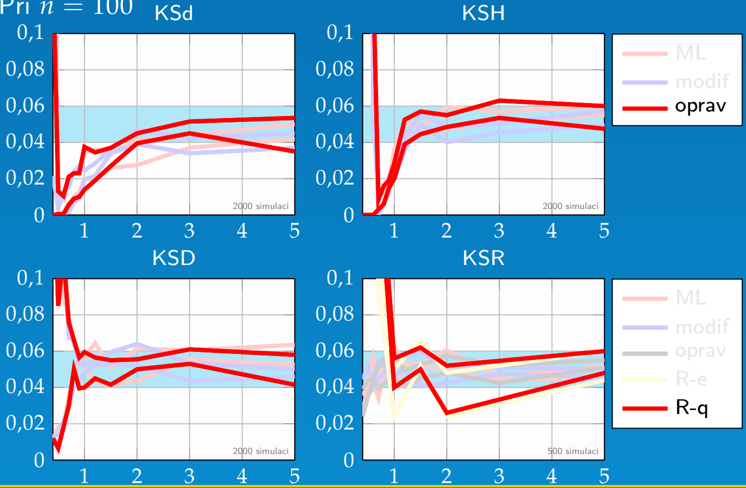
Dosažená hladina významnosti

Při $n = 100$



Dosažená hladina významnosti

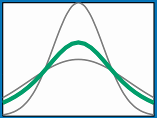
Při $n = 100$



Porovnání testů

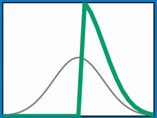
Skutečná hladina významnosti

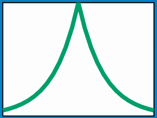
- může se lišit od nominální (aproximace, odhady...
- simulace pro $\mu \in \langle 0; 0{,}5 \rangle$ a $\sigma \in \langle 0{,}4; 5 \rangle$

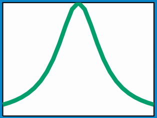


Síla proti alternativám

- směs normálních rozdělení
  $30\,\% \cdot \mathrm{N}(\mu, 1) + 70\,\% \cdot \mathrm{N}(\mu, 4)$
- useknuté normální
  $\mathrm{N}(\mu, \sigma^2)$ na (μ, ∞)
- Laplaceovo (dvojitě exponenciální) rozdělení
  $\mu \pm \mathrm{Exp}(\theta)$, $0{,}4 \leqq \theta \leqq 5$
- Studentovo t-rozdělení
  $\mu + t_m$, $1 \leqq m \leqq 10$







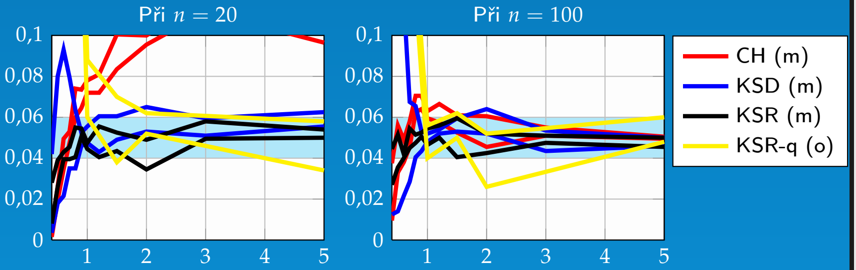
Skutečné hladiny významnosti

Chí-kvadrát test

- při $n = 20$ hladina významnosti mnohem větší (tudíž asi i síla)

Kolmogorovův-Smirnovův test

- varianty (kromě KSR) spíš při $\sigma > 0,8$



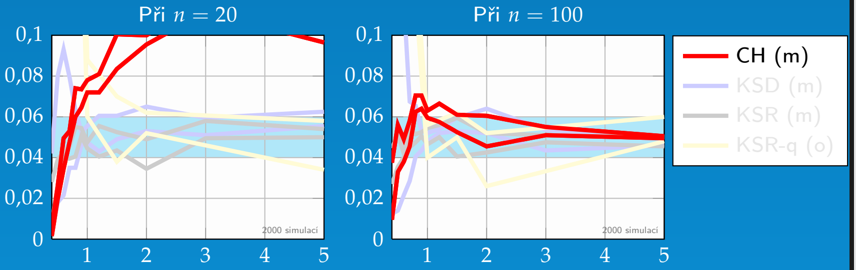
Skutečné hladiny významnosti

Chí-kvadrát test

- při $n = 20$ hladina významnosti mnohem větší (tudíž asi i síla)

Kolmogorovův-Smirnovův test

- varianty (kromě KSR) spíš při $\sigma > 0,8$



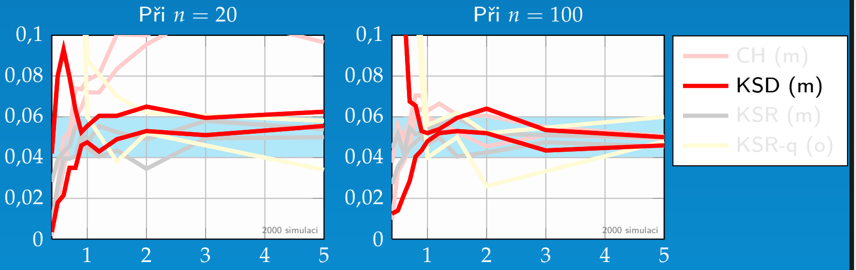
Skutečné hladiny významnosti

Chí-kvadrát test

- při $n = 20$ hladina významnosti mnohem větší (tudíž asi i síla)

Kolmogorovův-Smirnovův test

- varianty (kromě KSR) spíš při $\sigma > 0,8$



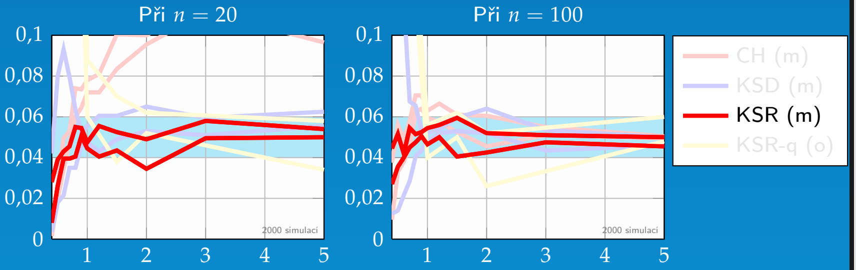
Skutečné hladiny významnosti

Chí-kvadrát test

- při $n = 20$ hladina významnosti mnohem větší (tudíž asi i síla)

Kolmogorovův-Smirnovův test

- varianty (kromě KSR) spíš při $\sigma > 0,8$



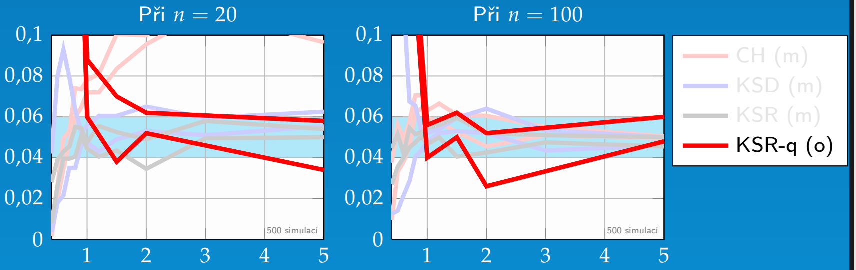
Skutečné hladiny významnosti

Chí-kvadrát test

- při $n = 20$ hladina významnosti mnohem větší (tudíž asi i síla)

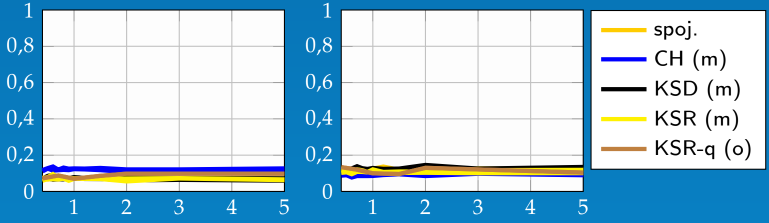
Kolmogorovův-Smirnovův test

- varianty (kromě KSR) spíš při $\sigma > 0,8$

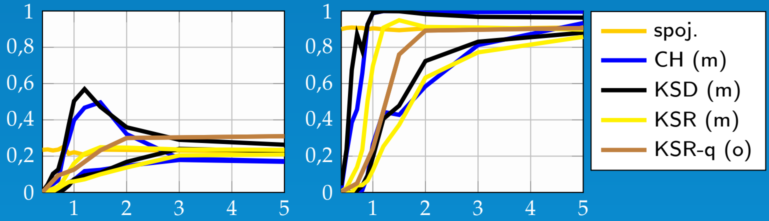


Síla testů

Směs norm. Při $n = 20$ Při $n = 100$

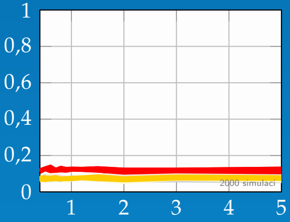


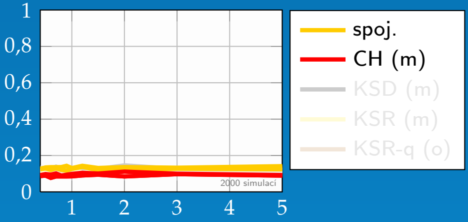
Useknuté normální



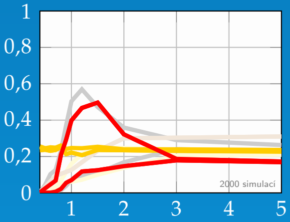
Síla testů

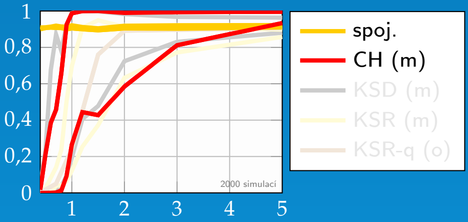
Směs norm. Při $n = 20$ Při $n = 100$





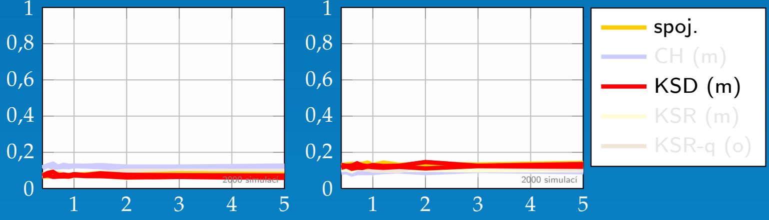
Useknuté normální



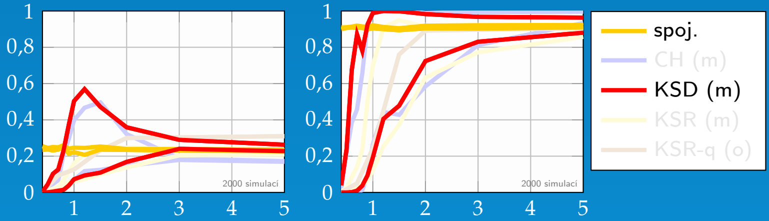


Síla testů

Směs norm. Při $n = 20$ Při $n = 100$

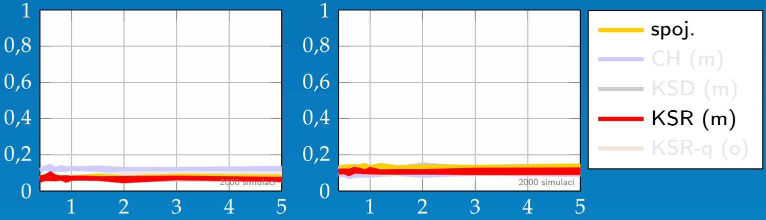


Useknuté normální

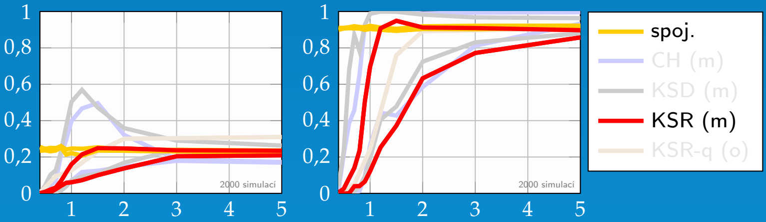


Síla testů

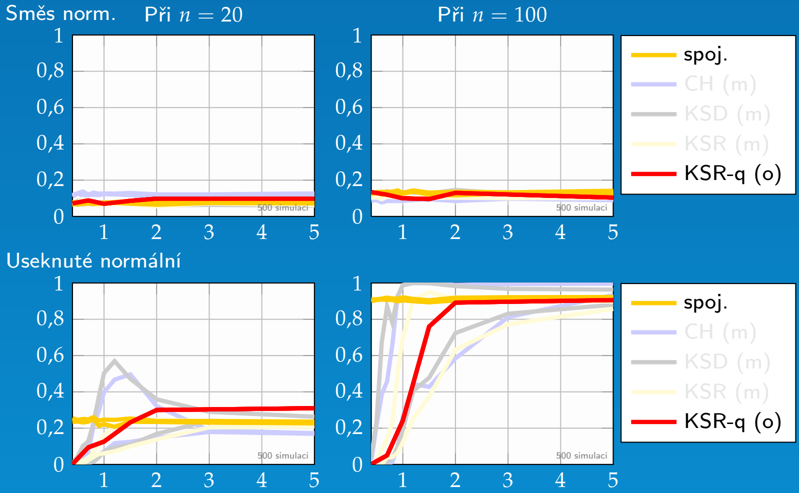
Směs norm. Při $n = 20$ Při $n = 100$



Useknuté normální

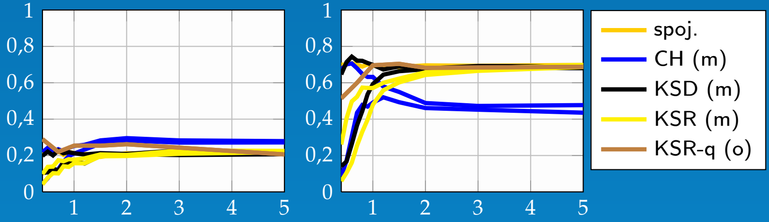


Síla testů

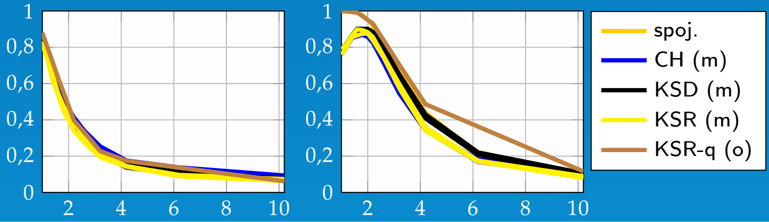

Směs norm.
Při $n = 20$
Při $n = 100$
1
0,8
0,6
0,4
0,2
0
1 2 3 4 5
500 simulací
spoj.
CH (m)
KSD (m)
KSR (m)
KSR-q (o)
Useknuté normální

Síla testů

Laplaceovo Při $n = 20$ Při $n = 100$



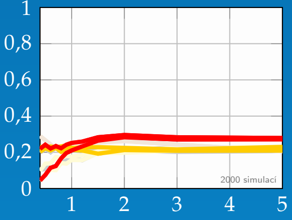
Studentovo



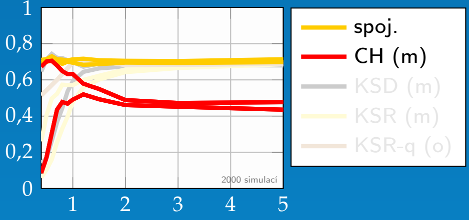
Síla testů

Laplaceovo

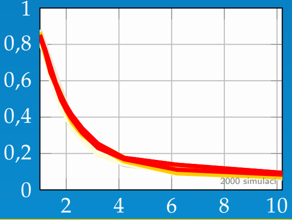
Při $n = 20$

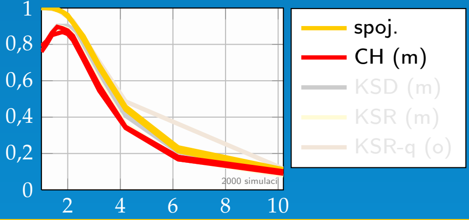


Při $n = 100$

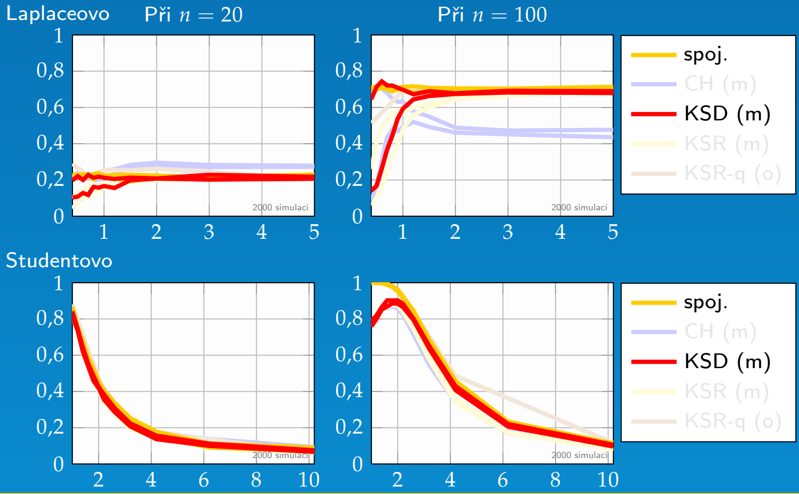


Studentovo

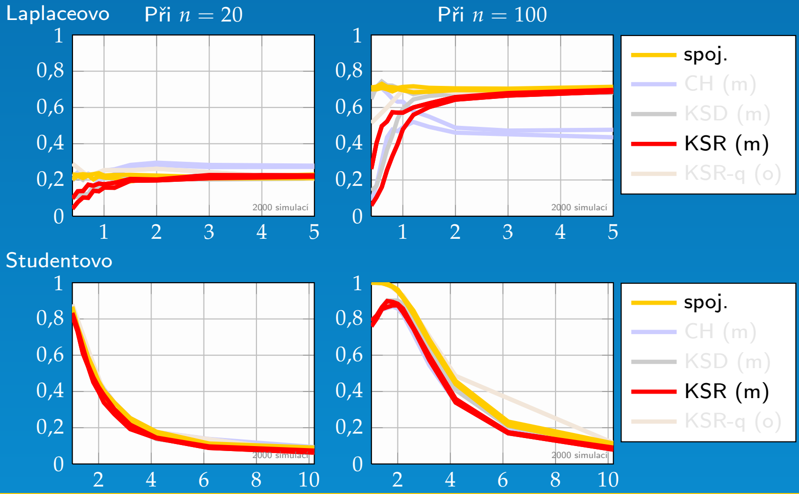




Síla testů

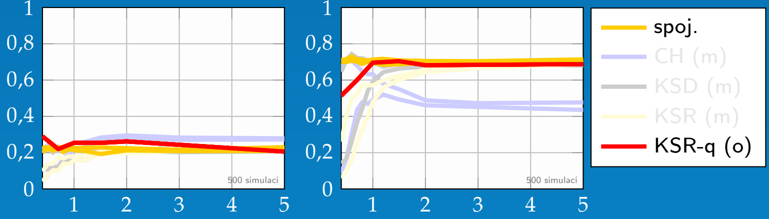

Laplaceovo
Při $n = 20$
Při $n = 100$
Studentovo
1
0,8
0,6
0,4
0,2
0
1 2 3 4 5
2 4 6 8 10
2000 simulací
spoj.
CH (m)
KSD (m)
KSR (m)
KSR-q (o)

Síla testů

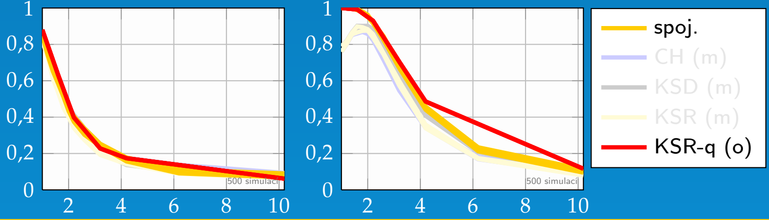

Laplaceovo
Při $n = 20$
Při $n = 100$
Studentovo
2000 simulací
spoj.
CH (m)
KSD (m)
KSR (m)
KSR-q (o)

Síla testů

Laplaceovo Při $n = 20$ Při $n = 100$



Studentovo



Cenzorovaná data

V nastíněných příkladech

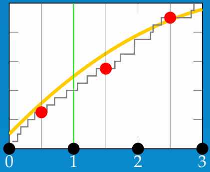
- obor pozorovaných hodnot omezený
- předp. v krajních třídách zahrnuta i pozorování menší/větší
- cenzorující hodnota nenáhodná

Chí-kvadrát test

- „nevadí", cenzorovaná data přispějí do krajních tříd

Kolmogorovův-Smirnovův test a varianty

- porovnáme jen část distribuční funkce
- kritická hodnota závisí na procentu cenzorování
- u KSd/KSH/KSD závislost i na prametrech



Cenzorovaná data

V nastíněných příkladech

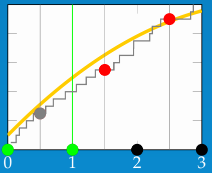
- obor pozorovaných hodnot omezený
- předp. v krajních třídách zahrnuta i pozorování menší/větší
- cenzorující hodnota nenáhodná

Chí-kvadrát test

- „nevadí“, cenzorovaná data přispějí do krajních tříd

Kolmogorovův-Smirnovův test a varianty

- porovnáme jen část distribuční funkce
- kritická hodnota závisí na procentu cenzorování
- u KSd/KSH/KSD závislost i na prametrech



Cenzorovaná data

V nastíněných příkladech

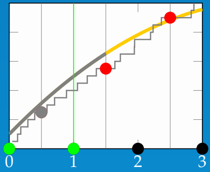
- obor pozorovaných hodnot omezený
- předp. v krajních třídách zahrnuta i pozorování menší/větší
- cenzorující hodnota nenáhodná

Chí-kvadrát test

- „nevadí“, cenzorovaná data přispějí do krajních tříd

Kolmogorovův-Smirnovův test a varianty

- porovnáme jen část distribuční funkce
- kritická hodnota závisí na procentu cenzorování
- u KSd/KSH/KSD závislost i na prametrech



Cenzorovaná data

V nastíněných příkladech

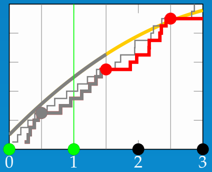
- obor pozorovaných hodnot omezený
- předp. v krajních třídách zahrnuta i pozorování menší/větší
- cenzorující hodnota nenáhodná

Chí-kvadrát test

- „nevadí“, cenzorovaná data přispějí do krajních tříd

Kolmogorovův-Smirnovův test a varianty

- porovnáme jen část distribuční funkce
- kritická hodnota závisí na procentu cenzorování
- u KSd/KSH/KSD závislost i na prametrech

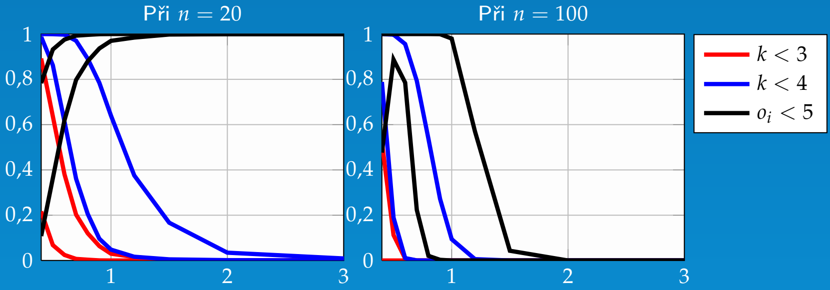


Chování testů

Uvažujeme výběry z $\mathrm{N}(\mu, \sigma^2)$, data cenzorovaná zleva 0 a zprava 5

- $\mu = 3{,}5, \dots, 4{,}5$
- $\sigma = 0{,}4, \dots, 5$

Problémy s počty tříd

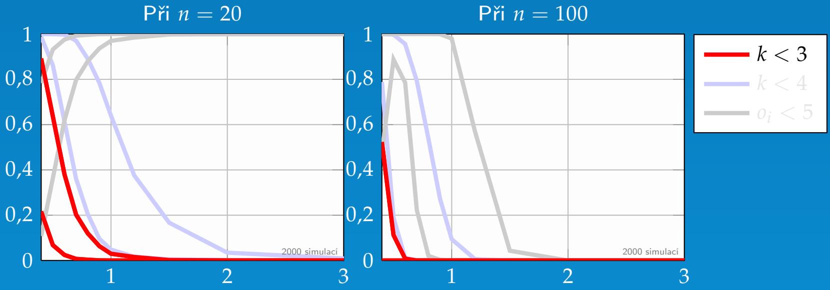


Chování testů

Uvažujeme výběry z $\mathrm{N}(\mu, \sigma^2)$, data cenzorovaná zleva 0 a zprava 5

- $\mu = 3{,}5, \ldots, 4{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Problémy s počty tříd



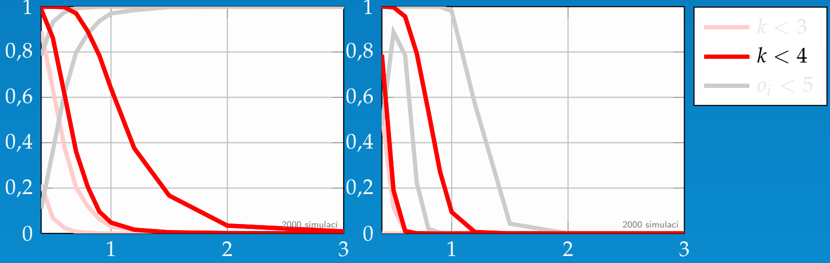
Chování testů

Uvažujeme výběry z $\mathrm{N}(\mu, \sigma^2)$, data cenzorovaná zleva 0 a zprava 5

- $\mu = 3{,}5, \ldots, 4{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Problémy s počty tříd

Při $n = 20$ Při $n = 100$

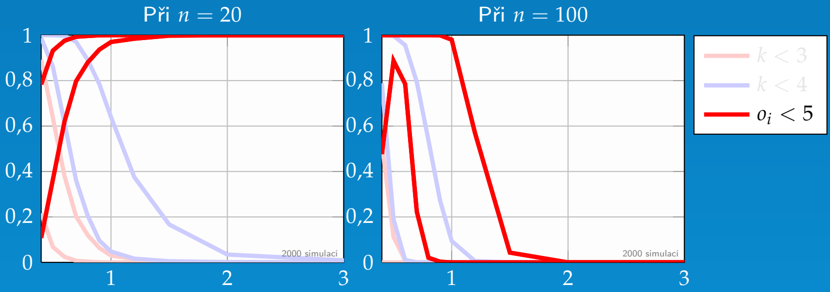


Chování testů

Uvažujeme výběry z $\mathrm{N}(\mu, \sigma^2)$, data cenzorovaná zleva 0 a zprava 5

- $\mu = 3{,}5, \ldots, 4{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Problémy s počty tříd

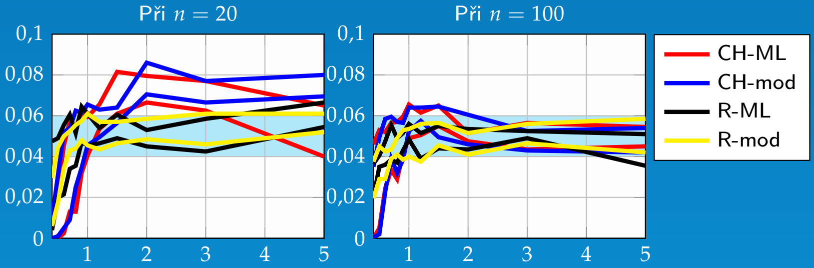


Chování testů

Uvažujeme výběry z $\mathrm{N}(\mu, \sigma^2)$, data cenzorovaná zleva 0 a zprava 5

- $\mu = 3{,}5, \ldots, 4{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Dosažené hladiny významnosti

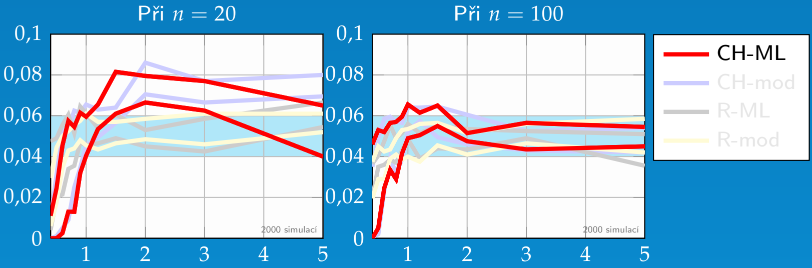


Chování testů

Uvažujeme výběry z $\mathrm{N}(\mu, \sigma^2)$, data cenzorovaná zleva 0 a zprava 5

- $\mu = 3{,}5, \ldots, 4{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Dosažené hladiny významnosti

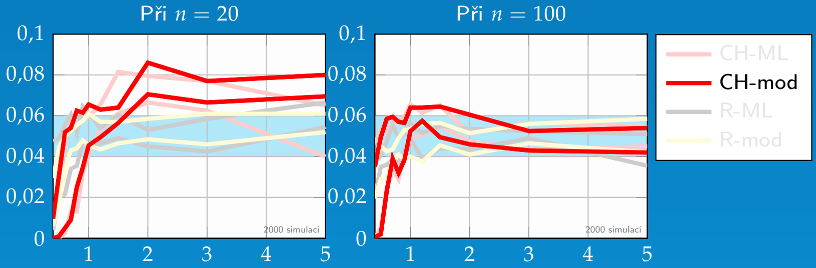


Chování testů

Uvažujeme výběry z $\mathrm{N}(\mu, \sigma^2)$, data cenzorovaná zleva 0 a zprava 5

- $\mu = 3{,}5, \ldots, 4{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Dosažené hladiny významnosti

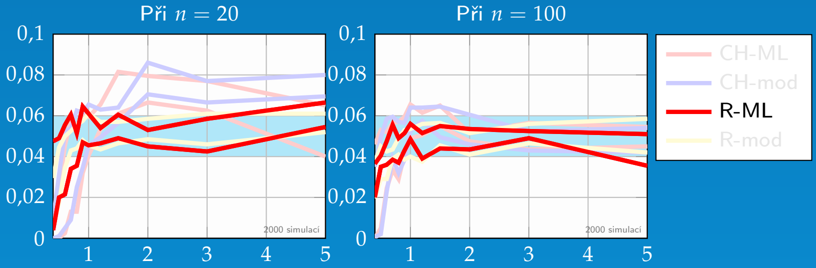


Chování testů

Uvažujeme výběry z $\mathrm{N}(\mu, \sigma^2)$, data cenzorovaná zleva 0 a zprava 5

- $\mu = 3{,}5, \ldots, 4{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Dosažené hladiny významnosti

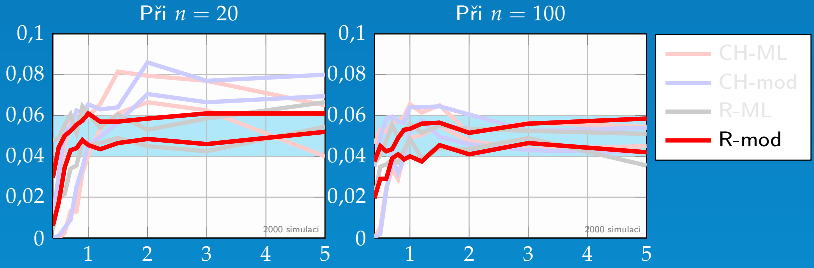


Chování testů

Uvažujeme výběry z $\mathrm{N}(\mu, \sigma^2)$, data cenzorovaná zleva 0 a zprava 5

- $\mu = 3{,}5, \ldots, 4{,}5$
- $\sigma = 0{,}4, \ldots, 5$

Dosažené hladiny významnosti

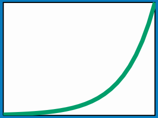


Chování testů

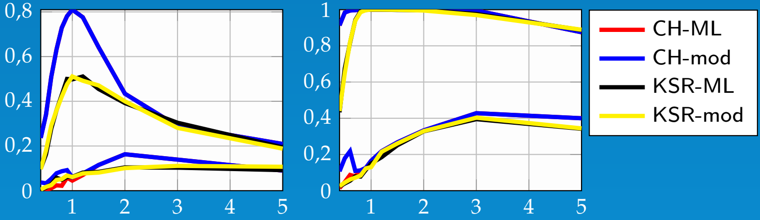
Síla proti alternativám

Exponenciální

- $\mu - \mathrm{Exp}(\theta)$



Při $n = 20$ Při $n = 100$

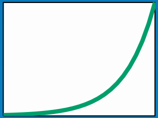


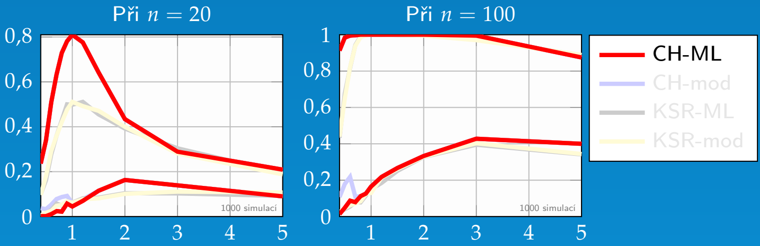
Chování testů

Síla proti alternativám

Exponenciální

- $\mu - \mathrm{Exp}(\theta)$



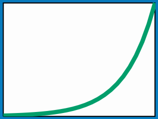


Chování testů

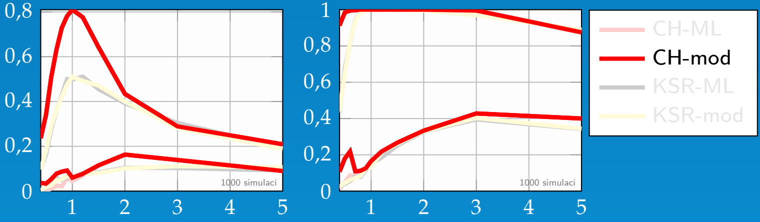
Síla proti alternativám

Exponenciální

- $\mu - \mathrm{Exp}(\theta)$



Při $n = 20$ Při $n = 100$

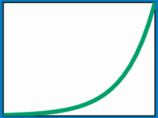


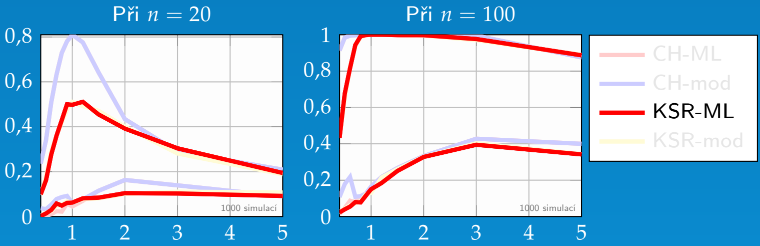
Chování testů

Síla proti alternativám

Exponenciální

- $\mu - \mathrm{Exp}(\theta)$



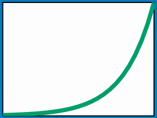


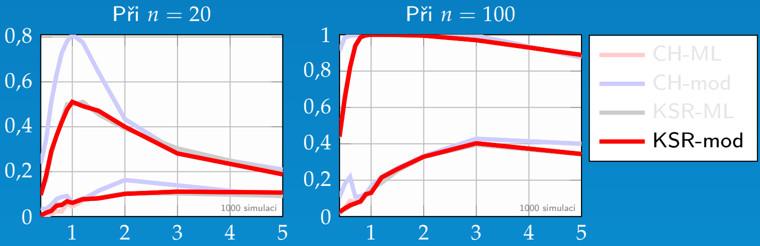
Chování testů

Síla proti alternativám

Exponenciální

- $\mu - \mathrm{Exp}(\theta)$



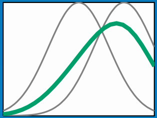


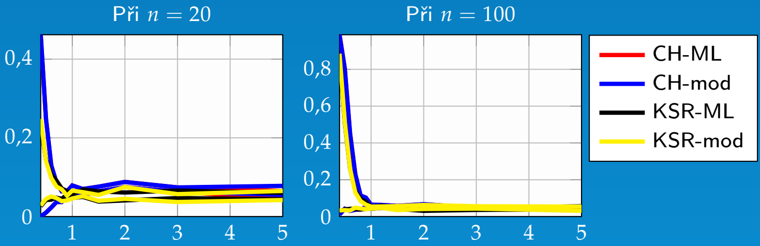
Chování testů

Síla proti alternativám

Směs normálních

- $30\,\% \cdot \mathrm{N}(\mu, \sigma^2) + 70\,\%\,\mathrm{N}(4, \sigma^2)$



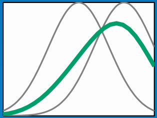


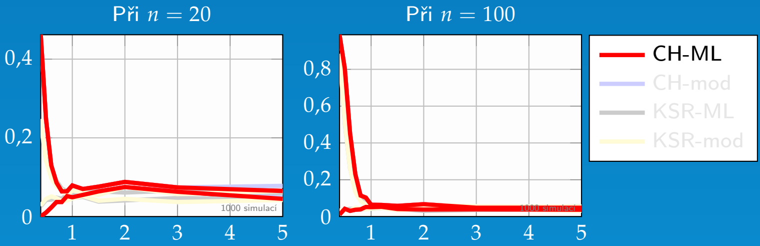
Chování testů

Síla proti alternativám

Směs normálních

- $30\,\% \cdot \mathrm{N}(\mu, \sigma^2) + 70\,\% \, \mathrm{N}(4, \sigma^2)$



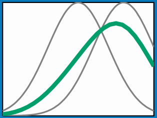


Chování testů

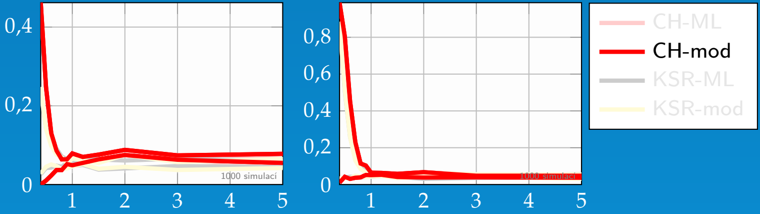
Síla proti alternativám

Směs normálních

- $30\,\% \cdot \mathrm{N}(\mu, \sigma^2) + 70\,\%\, \mathrm{N}(4, \sigma^2)$



Při $n = 20$ Při $n = 100$

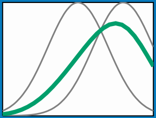


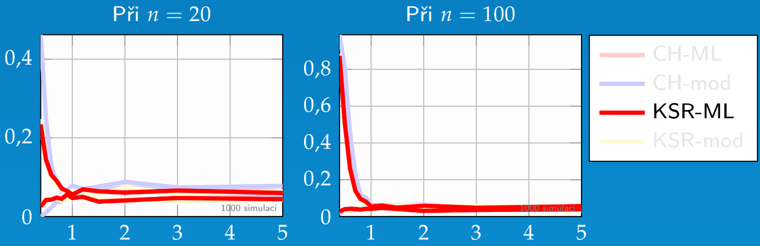
Chování testů

Síla proti alternativám

Směs normálních

- $30\,\% \cdot \mathrm{N}(\mu, \sigma^2) + 70\,\% \, \mathrm{N}(4, \sigma^2)$



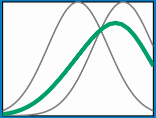


Chování testů

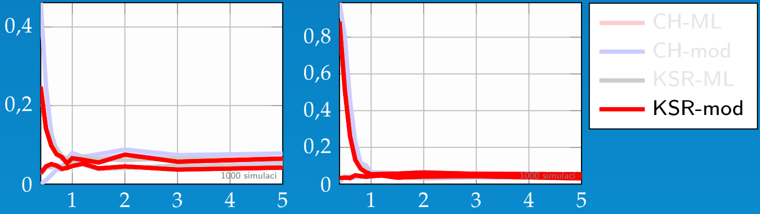
Síla proti alternativám

Směs normálních

- $30\,\% \cdot \mathrm{N}(\mu, \sigma^2) + 70\,\% \, \mathrm{N}(4, \sigma^2)$



Při $n = 20$ Při $n = 100$



Příklad: Hodnocení kvality výuky

Vyhodnocena normalita u 241 otázek o předmětech jedné fakulty

- z toho

Test	# zam. H_0	Procentně
CH-ML	45	18,6 %
CH-modif	45	18,6 %
KSR-ML	21	8,7 %
KSR-modif	17	7,1 %

Neshoda mezi testy u 40 otázek, z toho

- ve 20 případech rozhodnutí těsné (blízko 5 % hranice)
- resp. ve 21 shoda tří testů ze čtyř
- ve 31 případech aspoň jedno z toho

Příklad: Vstupní test

Fakulta	Písemek	Výsledek
A	386	zam.
B	647	zam.
C	350	zam.
D	60	nezam.
E	113	nezam.
F	172	zam.
G	23	nezam.

Po fakultách

- všechny testy shodné rozhodnutí
- normalita zamítnuta u větších fakult

Po studijních oborech

- testováno 24 oborů, z nich:

Test	# zam. H_0	Procentně
CH-ML	6	25,0 %
CH-modif	7	29,2 %
KSR-ML	6	25,0 %
KSR-modif	8	33,3 %

- neshoda mezi testy u 7 oborů
  (z nich u 3 těsně a u dalších 2 odlišný 1 test)

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

TESTOVÁNÍ NORMALITY ZE ZAOKROUHLENÝCH DAT

Michal Friesl

Katedra matematiky
Fakulta aplikovaných věd
Západočeská univerzita v Plzni

- http://home.zcu.cz/~friesl/Rob12.html