MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

BAKALÁRSKA PRÁCA

Tomáš Krupa

Štatistické testy normality

Katedra pravděpodobnosti a matematické statistiky

Vedúci bakalárskej práce: RNDr. Matúš Maciak, Ph.D.
Študijný program: Obecná matematika

Praha 2023

Vyhlasujem, že som túto bakalársku prácu vypracoval samostatne a výhradne s použitím citovaných prameňov, literatúry a ďalších odborných zdrojov. Táto práca nebola použitá k získaniu iného alebo rovnakého titulu.

Beriem na vedomie, že sa na moju prácu vzťahujú práva a povinnosti vyplývajúce zo zákona č. 121/2000 Sb., autorského zákona v platnom znení, najmä skutočnosť, že Univerzita Karlova má právo na uzavretie licenčnej zmluvy o použití tejto práce ako školského diela podľa §60 odst. 1 autorského zákona.

V .... dňa ....
Podpis autora

Na tomto mieste sa chcem poďakovať RNDr. Matúšovi Maciakovi, Ph.D., vedúcemu mojej bakalárskej práce, za cenné rady, rýchle odpovede a veľké množstvo trpezlivosti pri tvorbe tejto práce.

Názov práce: Štatistické testy normality

Autor: Tomáš Krupa

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedúci bakalárskej práce: RNDr. Matúš Maciak, Ph.D., Katedra pravdepodobnosti a matematické statistiky

Abstrakt: Cieľom tejto práce je predstaviť známe, v praxi používané testy normality a porovnať ich. Prvá kapitola pozostáva zo základných pojmov a vlastnosti normálneho rozdelenia. V druhej kapitole je spracovaných 6 testov normality, konkréte Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk, Anderson-Darling, D'Agostino-Pearson a Jarque-Bera. Pre každý test je okrem iného uvedená testová štatistika a tvar kriticého oboru. Tretia kapitola s empirickou štúdiou obsahuje dve časti. V prvej časti je stručne vysvetlený charakter štúdie a empiricky skontrolovaná deklarovaná hladina testov. V druhej časti je empiricky porovnaná sila testov proti rôznym alternatívam a diskusia výsledkov.

Kľúčové slová: štatistický test, normálne rozdelenie, Gaussovo rozdelenie

Title: Statistical tests of normality

Author: Tomáš Krupa

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Matúš Maciak, Ph.D., department

Abstract: The aim of this paper is to present the well-known normality tests used in practice and to compare them. The first chapter consists of the basic concepts and properties of the normal distribution. In the second chapter 6 normality tests are treated, namely Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk, Anderson-Darling, D'Agostino-Pearson and Jarque-Bera. For each test, test statistic and shape of critical region are given, among others. The third chapter, with empirical study, contains two parts. In the first part, nature of the study is briefly explained and level of significance declared by tests is empirically-checked. In the second part, power of tests is empirically compared against various alternatives and the results are discussed.

Keywords: statistical test, normal distribution, Gaussian distribution

Obsah

Úvod

Predpoklad normality je nezanedbateľný pre veľa štatistických procedúr. Jedná sa o analýzu rozptylu, niektoré verzie t-testu, lineárnu regresiu atď. Dá sa povedať, že vlastnosť normality sa nejakým spôsobom vyskytuje v každom odbore čerpajúcom zo štatistiky (napr. niektoré vlastnosti zdravej populácie v biológii majú spravidla normálne rozdelenie). Začiatkom 20. storočia dokonca panoval názor, že to je jediné rozdelenie, ktoré sa v praxi testuje. Z týchto dôvodov je dôležité mať (a správne používať) spoľahlivé metódy overenia normality. Grafickými metódami sa v tejto práci zaoberať nebudeme, venovať sa budeme presnejšej metóde, ktorou sú štatistické testy.

Za posledné storočie bolo predstavených veľa rôznych testov normality, líšia sa predpokladmi, charakteristickou vlastnosťou normality, ktorú testujú a pre nás najpodstatnejšou vlastnosťou - silou testu. Táto práca poskytne stručný popis a simulačnú štúdiu 6 testov normality. Jedná sa o testy Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk, Anderson-Darling, D'Agostino-Pearson a Jarque-Bera.

V prvej kapitole definujeme normálne rozdelenie a dokážeme jednu z jeho základných vlastností. Potom vysvetlíme pojmy hladina testu a silofunkcia a zavedieme použité značenie základných popisných charakteristík náhodného výberu.

V druhej kapitole sa bližšie zoznámime so spomínanými 6 testami normality. Vyjadríme sa k testovým štatistikám, kritickým oborom, implementácii a predpokladom týchto testov. Predpoklady majú formu modelu, ak pri príslušnom teste model nie je, znamená to, že žiadne zvláštne predpoklady nemá.

Tretia kapitola pozostáva z overenia deklarovanej hladiny testov a empirického porovnania sily proti rôznym alternatívam. Na konci obsahuje diskusiu, v rámci ktorej interpretujeme výsledky empirickej štúdie.

1. Základné pojmy

V prvej kapitole uvádzame definície (a jedno základné tvrdenie), ktoré sa v rámci témy testov normality skutočne nedajú opomenúť. Budeme čerpať hlavne z poznámok k prednáške z matematickej štatistiky (Kulich a Omelka, 2022). Pojmy nulová hypotéza H_0, alternatíva H_1, testová štatistikaT_n, kritický obor $\mathcal{C}_n$ a model $\mathcal{F}$ považujeme za fundamentálne a dovolíme si ich používať bez toho aby sme ich definovali.

1.1 Definícia normality

Definícia 1 (Náhodná veličina a jej rozdelenie)**.** *Nech $(S,\mathcal{S})$ je merateľný priestor a $(\Omega,\mathcal{A},P)$ je pravdepodobnostný priestor. Potom náhodná veličina (n.v.) je merateľné zobrazenie $X : (\Omega,\mathcal{A}) \mapsto (S,\mathcal{S})$. Rozdelenie n.v. X je pravdepodobnostná miera na $\mathcal{S}$ daná predpisom: $P_X(B) = P(X \in B),\ B \in \mathcal{S}$.*

V tejto práci budeme pracovať iba s reálnymi náhodnými veličinami, teda $(S,\mathcal{S}) = (\mathbb{R},\mathcal{B})$, kde $\mathcal{B}$ je Borelovská $\sigma-$algebra.

Normálne, resp. Gaussovské, rozdelenie náhodnej veličiny je spojité (voči Lebesgueovej miere) pravdepodobnostné rozdelenie, ktoré závisí na parametroch $\mu \in \mathbb{R}$ a $\sigma^2 > 0$, kde μ je stredná hodnota a σ^2 je rozptyl.

Definícia 2 (Normálne rozdelenie)**.** *Nech $\mu \in \mathbb{R},\sigma^2 > 0$ a n.v. X má rozdelenie $X \sim N(\mu,\sigma^2)$. Potom hustota n.v. X je daná vzťahom:*

$$f_X(x) = \varphi_{(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}$$

Pre distribučnú funkciu n.v. X platí:

$$F_X(x) = \Phi_{(\mu,\sigma^2)}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt, x \in \mathbb{R}$$

Keďže hustota $\varphi_{(\mu,\sigma^2)}(x)$ nemá primitívnu funkciu, počítajú sa hodnoty $\Phi_{(\mu,\sigma^2)}(x)$ numericky, resp. sú tabelované. V prípade, že $\mu = 0$ a $\sigma^2 = 1$, sa jedná o normované normálne rozdelenie, ktorého hustotu značíme $\varphi(x)$ s ditribučnou funkciou $\Phi(x), x \in \mathbb{R}$.

Tvrdenie 1 (Vlastnosť normálneho rozdelenia)**.** *Nech $X \sim N(0,1), b \in \mathbb{R}$, $a \in (0,\infty)$ a $Y = aX + b$, potom $Y \sim N(b,a^2)$.*

Dôkaz.

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(aX + b \leq y) = \\ &= P\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right) = \Phi\left(\frac{y-b}{a}\right) \end{aligned}$$

Potom pre hustotu n.v. Y platí:

$$\begin{aligned} f_Y(y) = (F_Y(y))' &= \left(\Phi\left(\frac{y-b}{a}\right)\right)' = \varphi\left(\frac{y-b}{a}\right)\left(\frac{y-b}{a}\right)' = \\ &= \varphi\left(\frac{y-b}{a}\right)\frac{1}{a} = \frac{1}{\sqrt{2\pi}}e^{-\frac{(x-b)^2}{2a^2}}\frac{1}{a} = \frac{1}{\sqrt{2\pi a^2}}e^{-\frac{(x-b)^2}{2a^2}} \end{aligned}$$

Čo je z definície hustota rozdelenia N(b,a^2).

$\square$

1.2 Hladina a sila testu

Definícia 3 (Náhodný výber)**.** *Nech* $\{X_i, i = 1, \ldots, n\}, n \in \mathbb{N}$ *sú nezávislé náhodné veličiny s rovnakým rozdelením. Potom* $\{X_i, i = 1, \ldots, n\}$ *nazveme náhodný výber a označíme* $\mathbb{X}_n$*, n nazveme rozsah (alebo veľkosť) výberu. Usporiadaním prvkov náhodného výberu vzostupne získame usporiadaný náhodný výber* $\{X_{(i)}, i = 1, \ldots, n\}$ *kde platí* $X_{(i)} \leq X_{(i+1)}, \forall i \in \{1, \ldots, n-1\}$.

Definícia 4 (Štatistický test)**.** *Definujme štatistický test ako funkciu náhodného výberu* $\Theta : \mathbb{R}^n \mapsto \{0,1\}$*, kde zamietame* $H_0 \Leftrightarrow \Theta(\mathbb{X}_n) = 1$.

Štatistické testy v tejto práci budú stále tvaru $\Theta(\mathbb{X}_n) = \mathbf{1}_{[T_n \in \mathcal{C}_n]}$, kde $n \in \mathbb{N}$ je veľkosť náhodného výberu, T_n je testová štatistika a $\mathcal{C}_n$ je kritický obor.

Teda výsledkom použitia štatistického testu na nejaký náhodný výber je zamietnutie, alebo nezamietnutie nulovej hypotézy. Preto musí testovanie vyústiť do jednej z týchto štyroch možností:

- Nulová hypotéza platí a my sme ju nezamietli. Rozhodli sme správne.
- Nulová hypotéza platí a my sme ju zamietli. Nastala chyba prvého druhu.
- Nulová hypotéza neplatí a my sme ju zamietli. Rozhodli sme správne.
- Nulová hypotéza neplatí a my sme ju nezamietli. Nastala chyba druhého druhu.

Pravdepodobnosť chyby prvého druhu je pre nás dôležitejšia, preto ju berieme do úvahy už pri konštrukcii testu. Nazývame ju hladinou testu (značíme α). V tejto práci budú mať všetky testy štandardnú hladninu $\alpha = 0.05$.

Predmetom tejto práce je porovnanie štatistických testov (konkrétne testov normality), preto dôležité vysvetliť, čo vlastne od dobrého testu požadujeme. Prirodzene chceme, aby test dodržiaval určenú hladinu (alebo bol aspoň mierne konzervatívny). Ďalej by sme chceli, aby test minimalizoval chybu druhého druhu, ktorá ale závisí na konkrétnom prvku z alternatívy. Objekt, ktorý charakterizuje ako spoľahlivo zamieta test daný prvok z alternatívy sa nazýva silofunkcia. Uvedieme riadnu definíciu:

Definícia 5 (Hladina, silofunkcia a konzistencia testu)**.** *Nech máme daný test hypotézy* H_0 *proti alternatíve* H_1 *s testovou štatistikou* T_n *a kritickým oborom* $\mathcal{C}_n, n \in \mathbb{N}$*. Označme* $\mathcal{F}_0 \subset \mathcal{F}$ *množinu rozdelení spĺňajúcich* H_0*. Nech pre* $\alpha \in (0,1)$ *platí:*

$$\sup_{F \in \mathcal{F}_0} P_F(T_n \in \mathcal{C}_n) = \alpha, \ \forall n \in \mathbb{N}$$

Potom α *je hladina daného testu.*

Označme $\mathcal{F}_1 \subset \mathcal{F}$ *množinu rozdelení, ktoré nespĺňajú* H_0*, potom pre silofunkciu daného testu platí:*

$$\beta_n(F) = P_F(T_n \in \mathcal{C}_n), \ \forall F \in \mathcal{F}_1, n \in \mathbb{N}$$

Nech platí:

$$\lim_{n\to\infty} \beta_n(F) = 1, \ \forall F \in \mathcal{F}_1$$

Potom je daný test konzistentný.

1.3 Základné popisné štatistiky

Drvivá väčšina štatistík, ktoré uvedieme v 2. kapitole bude obsahovať vo svojom vzorci aspoň jednu základnú popisnú štatistiku. Preto pripomenieme ich definíciu a zavedieme značenie používané v tejto práci:

Definícia 6 (Základné popisné štatistiky)**.** *Nech* $\{X_i, i = 1, \dots, n\}, n \in \mathbb{N}$ *je náhodný výber, potom:*
Výberový priemer je tvaru:

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Výberový rozptyl je tvaru:

$$S_n{}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

Empirická distribučná funkcia je tvaru:

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{[X_i \leq x]}, x \in \mathbb{R}$$

Veta 2 (O výberovom priemere a výberovom rozptyle)**.** *Nech* $\{X_i, i = 1, \dots, n\}$, $n \in \mathbb{N}$ *je náhodný výber z rozdelenia* P_X, *potom sú výberový priemer* $\overline{X}_n$ *a výberový rozptyl* $S_n{}^2$ *nestrannými, konzistentnými odhadmi strednej hodnoty* $E(X)$, *rozptylu* $var(X)$ *rozdelenia náhodného výberu.*

Dôkaz. Veta je dokázaná v práci Kulich a Omelka (2022, str.19 a 20)

□

Veta 3 (O empirickej distribučnej funkcii)**.** *Nech* $\{X_i, i = 1, \dots, n\}, n \in \mathbb{N}$ *je náhodný výber z rozdelenia* P_X *s distribučnou funkciou* $F_X(x), x \in \mathbb{R}$, *potom je empirická distribučná funkcia* $\widehat{F}_n(x)$ *nestranný a stejnoměrne konzistentný odhad distribučnej funkcie* $F_X(x)$, $\forall x \in \mathbb{R}$.

Dôkaz. Vetu sme prebrali zo zdroja Kulich a Omelka (2022, str.57).

□

2. Vybrané testy normality

V tejto kapitole predstavíme šesť testov normality, ktoré sa chystáme porovnať. Menovite sa jedná o Kolmogorov-Smirnov (skratkou K-S) sekcia(2.1), Lilliefors sekcia(2.2), Shapiro-Wilkov (skratkou S-W) sekcia(2.3), Anderson-Darlingov (skratkou A-D) sekcia(2.4), D'Agostino-Pearsonov (Omnibus) sekcia(2.5) a Jarque-Bera (skratkou J-B) sekcia(2.6) test. Keďže testujeme normalitu, bude sa jednať o testy neparametrické, tzv. "testy dobrej zhody". Všetky zmienené testy pracujú s nejakou variantou tejto nulovej hypotézy a alternatívy:

Nulová hypotéza a alternatíva:

$$H_0 : \exists\mu \in \mathbb{R}, \exists\sigma \in \mathbb{R}, \sigma \neq 0 : X \sim N(\mu,\sigma^2)$$
$$H_1 : \forall\mu \in \mathbb{R}, \forall\sigma \in \mathbb{R}, \sigma \neq 0 : (X \nsim N(\mu,\sigma^2))$$

2.1 Kolmogorov-Smirnov test normality

Tento test je založený na jednovýberovom K-S teste Kolmogorov (1933), ktorý poznáme z Kulich a Omelka (2022, str.94). Testujeme celú distribučnú funkciu skutočného rozdelenia náhodného výberu, resp. jej dobrý odhad (empirickú distribučnú funkciu). Model bude rovnaký ako pri jednovýberovom K-S teste, z rovnakého dôvodu ako pre jednovýberový K-S test Kulich a Omelka (2022, str.97). Respektíve preto, že testy používajúce vlastnosti empirickej distribučnej funkcie majú problém s náhodnými výbermi, ktoré obsahujú viac rovnakých hodnôt. Spojitosť skutočného rozdelenia zaručí, že táto situácia nenastane skoro iste.

Model:

$$\mathcal{F} = \{\text{všetky spojité rozdelenia}\}$$

Testová štatistika:

$$K_n = \underset{x\in\mathbb{R}}{sup}|\widehat{F_n}(x) - F(x)|$$

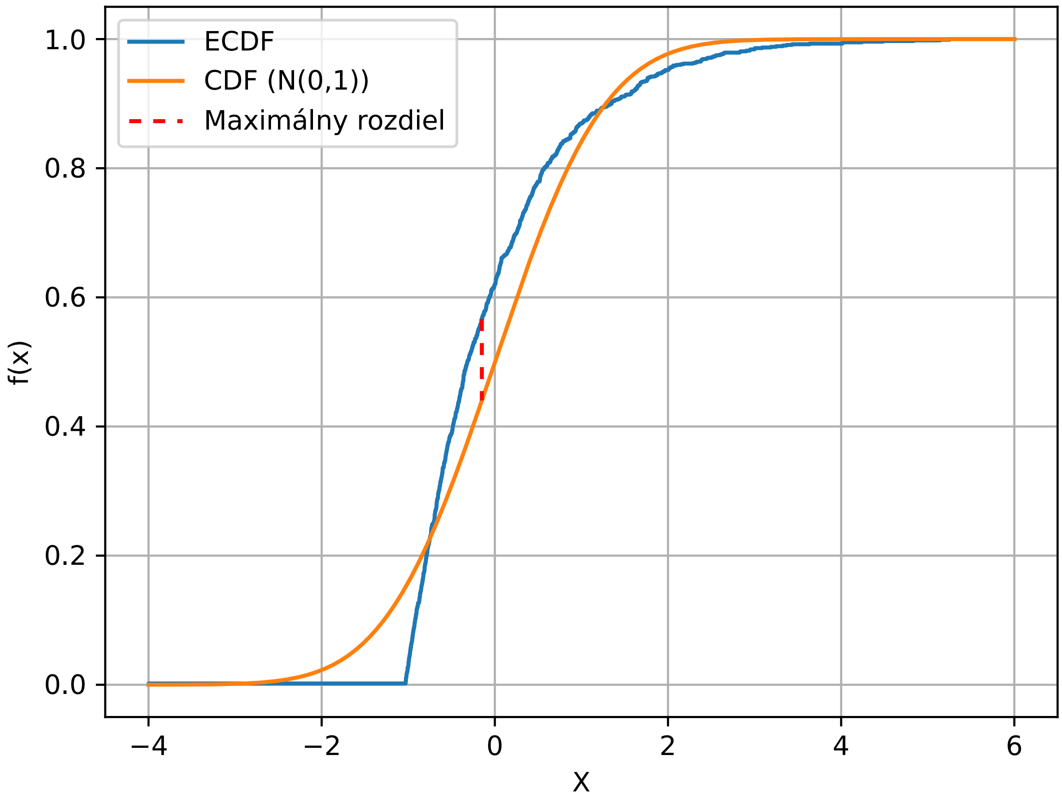
Kde $\widehat{F_n}(x)$ je empirická distribučná funkcia testovaného náhodného výberu a $F(x)$ je distribučná funkcia normálneho rozdelenia so strednou hodnotou $\mu = \overline{X}_n$ a rozptylom $\sigma^2 = {S_n}^2$. Pre predstavu testovej štatistiky viď. obr.2.1.

Testová štatistika použitá v implemetácii z kapitoly 3:

$$\widetilde{K_n} = \underset{x\in\mathbb{R}}{sup}|\widetilde{F}_n(x) - \Phi(x)|$$

Kde empirická distribučná funkcia $\widetilde{F}_n(x), n \in \mathbb{N}, x \in \mathbb{R}$ je daná transformovaným náhodným výberom $\{Y_i, i = 1, \ldots ,n\}$, pre ktorého prvky platí:

$$Y_i = \frac{X_i - \overline{X}_n}{\sqrt{{S_n}^2}},\ \forall i \in \{1, \ldots ,n\}$$



Obrázek 2.1: *Znázornenie K-S štatistkiky*
Modrou je znázornená empirická distribučná funkcia transformovaného náhodného výberu so skutočným rozdelením $Exp(1)$. Oranžová funkcia je $\Phi(x)$. Červenou farbou je označená hodnota testovej štatistiky $\widetilde{K}_n$ v tomto prípade.

Veta 4. *Nech $n \in \mathbb{N}$, potom $K_n \overset{s.j.}{=} \widetilde{K}_n$.*

Dôkaz. Zafixujme hodnoty náhodného výberu, ďalej nech $Y \sim N(0{,}1)$, potom s použitím zavedeného značenia platí:

$$K_n = \sup_{x\in\mathbb{R}}\left|F(x) - \widehat{F}_n(x)\right| = \sup_{x\in\mathbb{R}}\left|P\left(x \geq Y \times {S_n}^2 + \overline{X}_n\right) - \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{[X_i \leq x]}\right| =$$
$$= \sup_{x\in\mathbb{R}}\left|P\left(\frac{x-\overline{X}_n}{{S_n}^2} \geq Y\right) - \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{\left[\frac{X_i-\overline{X}_n}{{S_n}^2} \leq \frac{x-\overline{X}_n}{{S_n}^2}\right]}\right| =$$
$$= \sup_{x\in\mathbb{R}}\left|\Phi\left(\frac{x-\overline{X}_n}{{S_n}^2}\right) - \widetilde{F}_n\left(\frac{x-\overline{X}_n}{{S_n}^2}\right)\right| = \sup_{x\in\mathbb{R}}\left|\Phi(x) - \widetilde{F}_n(x)\right| = \widetilde{K}_n$$

Kde druhá rovnosť platí kvôli vlastnostiam normálneho rozdelenia Tvrdenie (1) a piata rovnosť platí, lebo pre dané $\overline{X}_n, {S_n}^2 \in \mathbb{R}$ je transformácia $y = \frac{x-\overline{X}_n}{\sqrt{{S_n}^2}}$ bijekciou z $\mathbb{R}$ na $\mathbb{R}$.

□

Rovnako ako pri jednovýberovom K-S teste zamietame v prípade, že je K_n príliš veľká, teda kritický obor bude tvaru (c_n,∞), $c_n \in (0,\infty)$.

Pretože sme v testovej štatistike použili výberový priemer a výberový rozptyl, nepoznáme rozdelenie testovej štatistiky K_n. Pri klasickom K-S teste by sme poznali aspoň asymptotické rozdelenie $\sqrt{n}K_n$. K-S test normality používa (chybne) rovnaký kritický obor ako dvojvýberový K-S test, čo spôsobuje silnú konzervativitu testu.

Ďalej uvedieme niekoľko užitočných vlastností K-S testu normality.

Veta 5. *Za platnosti H_0 nezávisí rozdelenie testovej štatistiky K_n na parametroch $\mu \in \mathbb{R}$ a $\sigma \in \mathbb{R}, \sigma \neq 0$.*

Dôkaz. Jednotlivé kroky sú podrobne rozpísané v práci Macoun (2021, str. 14, Veta 13).

□

Veta 6. *K-S test normality je konzistentný.*

Dôkaz. Tvrdenie je dôsledkom Vety 1 a práce Macoun (2021, str.10, Veta 9).

□

2.2 Lillieforsov test normality

Korekciou kritického oboru K-S testu normality vznikol nový, tzv. Lillieforsov, výrazne menej konzervatívny test normality. Tento test je v ostatných veciach s K-S testom identický, má rovnaký model, štatistiku, je konzistentný z rovnakých dôvodov. Prvá korekcia prebehla v článku Lilliefors (1967), kde boli metódou Monte-Carlo získané vhodnejšie medze kritických oborov (viď. tabuľka na str.400). Dnes sa používajú (a v našej simulačnej štúdii sú použité) skoro rovnaké kritické obory, rozdiel je väčšinou až na treťom desatinnom mieste. Tabuľka používaná v našej štúdii je dohľadateľná v oficiálnej dokumentácii modulu zmieneného v tretej kapitole.

2.3 Shapiro-Wilkov test normality

S-W test normality Shapiro a Wilk (1965) je založený na princípe pomeru dvoch štatistík odhadujúcich rovnaký parameter, v tomto prípade sa jedná o rozptyl.

Model:

$$\mathcal{F} = \{\text{všetky spojité rozdelenia}\}$$

V nasledujúcich riadkoch zavedieme značenie premenných, ktoré sa vyskytujú v testovej štatistike S-W testu:

Nech $\{Y_{(i)}, i = 1, \dots, n\}, n \in \mathbb{N}$ je usporiadaný náhodný výber z normovaného normálneho rozdelenia. Nazerajme naň ako na náhodný vektor $\mathbf{Y} = (y_1,...,y_n)^\top$, kde $y_i = Y_{(i)} \;\; \forall i \in \{1,...,n\}$. Nech $\mathbf{m} = E(\mathbf{Y})$ je stredná hodnota a $\mathbf{V} = var(\mathbf{Y})$ rozptylová matica náhodného vektora $\mathbf{Y}$, potom definujeme vektor $\mathbf{a} = (a_1, \dots, a_n)^\top \in \mathbb{R}^n$:

$$\mathbf{a}^\top = \frac{\mathbf{m}^\top \mathbf{V}^{-1}}{(\mathbf{m}^\top \mathbf{V}^{-1} \mathbf{V}^{-1} \mathbf{m})^{1/2}}$$

Testová štatistika:

$$W_n = \frac{(\sum_{i=1}^{n} a_i X_{(i)})^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

Kde $\{X_i, i = 1, \ldots, n\}$ je testovaný náhodný výber a $\{X_{(i)}, i = 1, \ldots, n\}$ je usporiadaný testovaný náhodný výber.

V menovateli testovej štatistiky si môžeme všimnúť výberový rozptyl (člen $\frac{1}{n-1}$ sa skrátil), ktorý nezávisí na normalite dát. Odhad rozptylu v čitateli naopak na rozdelení náhodného výberu závisí. Prehľadné odvodenie testovej štatistiky metódou lineárnej regresie nájdeme v práci Malíková (2014, str.6-8).

Ďalej zmienime dve podstatné vlastnosti Testovej štatistiky:

Veta 7. *Testová štatistika Shapiro-Wilkovho testu je invariantná vzhľadom na strednú hodnotu a rozptyl skutočného rozdelenia náhodného výberu.*

Dôkaz. Vyplýva z tvrdenia dokázaného v práci Shapiro a Wilk (1965, str.593). □

Veta 8. *Maximálna hodnota testovej štatistiky Shapiro-Wilkovho testu je 1.*

Dôkaz. Tvrdenie je detailne dokázané v práci Malíková (2014, str.11). □

Keďže sú čitateľ a menovateľ v testovej štatistike Shapiro-Wilkovho testu nutne kladné, musí byť kladná aj daná testová štatistika. Teda hodnoty testovej štatistiky sa musia vyskytovať v intervale $W_n \in (0,1]$. Pretože je tento test postavený na princípe pomeru dvoch odhadov rovnakého parametra, mala by testová štatistika dosahovať za platnosti H_0 hodnoty okolo 1, preto bude kritický obor tvaru $(c_n,1], c_n \in \mathbb{R}$. Presné rozdelenie testovej štatistiky Shapiro-Wilkovho testu nepoznáme, preto je kritický obor určený metódou Monte-Carlo. Tiež treba zmieniť, že výpočet koeficientov $a_i, i \in \{1,...,n\}$ (relatívne) výpočetne náročný, preto sa pri implementácii tohoto testu používajú pre menšie n tabelované konštanty a pre väčšie n aproximácie. Hľadaním dobrých aproximácii sa zaoberal najmä Royston (1995). Implementácia, ktorú použijeme v tretej kapitole je založená hlavne na jeho práci.

2.4 Anderson-Darlingov test normality

A-D test je pomerne univerzálny test dobrej zhody, ktorý podobne ako jednovýberový K-S test pouíva vlastnosti empirickej distribučnej funkcie. Ako by sme od testu pracujúceho s empirickou distribučnou funkciou očakávali, A-D test predpokladá spojitosť skutočného rozdelenia náhodného výberu.

Model:

$$\mathcal{F} = \{\text{všetky spojité rozdelenia}\}$$

Obmedzme zatiaľ tento test na nulovú hypotézu a alternatívu:

$$H_0 : X \sim N(0,1)$$
$$H_1 : X \nsim N(0,1)$$

Ako uvidíme, od jednovýberového K-S testu sa A-D test (s našou H_0) líši najmä ”metrikou”, ktorou určuje veľkosť rozdielu medzi empirickou distribučnou funkciou a $\Phi(x)$.Testová štatistika totiž predstavuje istým spôsobom (ako je uvedené nižšie) váženú ”integrálnu vzdialenosť” empirickej distribučnej funkcie $\widehat{F_n}(x)$ a $\Phi(x), x \in \mathbb{R}$.

Testová štatistika:

$$A^2{}_n = n \int_{-\infty}^{\infty} \frac{(\widehat{F_n}(x) - \Phi(x))^2}{\Phi(x)(1 - \Phi(x))} \, d\Phi(x) = n \int_{-\infty}^{\infty} (\widehat{F_n}(x) - \Phi(x))^2 w(x) \, d\Phi(x)$$

Kde $w(x), x \in \mathbb{R}$ je vyššie zmienená kladná váhová funkcia.

Je zjavné, že testová štatistika je nezáporná (čitateľ je nezáporný, menovateľ je kladný, a funkcia je $\Phi(x)$ neklesajúca). Ako by sme očakávali, $A^2{}_n$ bude za platnosti H_0 dosahovať menšie hodnoty, ako za platnosti H_1-rovnako ako pri Lillieforsovom teste. Teda kritický obor je tvaru $[0,c_n), c_n \in \mathbb{R}$.

Testová štatistika používaná pri výpočte:

$$D^2{}_n = -n - \frac{1}{n} \sum_{i=1}^{n} (2i - 1)[log(\Phi(X_{(i)}) + log(1 - \Phi(X_{(n-i+1)}))]$$

Zdôrazňujeme, že podľa zavedeného značenia sa v testovej štatistike $D^2{}_n$ vyskytujú prvky usporiadaného náhodného výberu.

Tvrdenie 9. *Nech $n \in \mathbb{N}$, potom $A^2{}_n = D^2{}_n$.*

Dôkaz. Tvrdenie je detailne dokázané v práci Anderson a Darling (1954, str.768) □

Môžeme si všimnúť, že za platnosti H_0 je váhová funkcia $w(x)$ strednou hodnotou náhodnej veličiny $n(\widehat{F_n}(x) - \Phi(x))^2 \; \forall x \in \mathbb{R}$. Tento test sa teda snaží ”vyrovnať” výberovú chybu, čo sa mu darí na celom $\mathbb{R}$. Súčasne pozorujeme, že váhová funkcia rastie smerom do chvostov, čo spôsobí lepšiu citlivosť tohoto testu na špecifický typ alternatív.

Tvrdenie 10. *Nech $n \in \mathbb{N}$ a platí nulová hypotéza, potom:*

$$nE(\widehat{F_n}(x) - \Phi(x))^2 = \frac{1}{w(x)}, \forall x \in \mathbb{R}$$

Dôkaz. Z definície empirickej distribučnej funkcie vyplýva, že za platnosti H_0 bude mať $n\widehat{F_n}(x)$ pre zafixované $x \in \mathbb{R}$ rozdelenie $n\widehat{F_n} \sim Bi(\Phi(x),n)$. Potom zrejme platí:

$$\begin{aligned} nE(\widehat{F_n}(x) - \Phi(x))^2 &= nE(\widehat{F_n}(x))^2 - 2E(n\widehat{F_n}(x))\Phi(x) + n(\Phi(x))^2 = \\ &= nE(\widehat{F_n}(x))^2 - 2n(\Phi(x))^2 + n(\Phi(x))^2 = \\ &= nE(\widehat{F_n}(x))^2 - n(\Phi(x))^2 \end{aligned}$$

V druhej rovnosti sme využili, že poznáme strednú hodnotu binomického rozdelenia, ďalej použijeme, že poznáme jeho rozptyl:

$$n\Phi(x)(1-\Phi(x)) = var(n\widehat{F_n}(x) = E(n\widehat{F_n}(x))^2 - (E(n\widehat{F_n}(x)))^2 =$$
$$= n^2 E(\widehat{F_n}(x))^2 - (n\Phi(x))^2$$

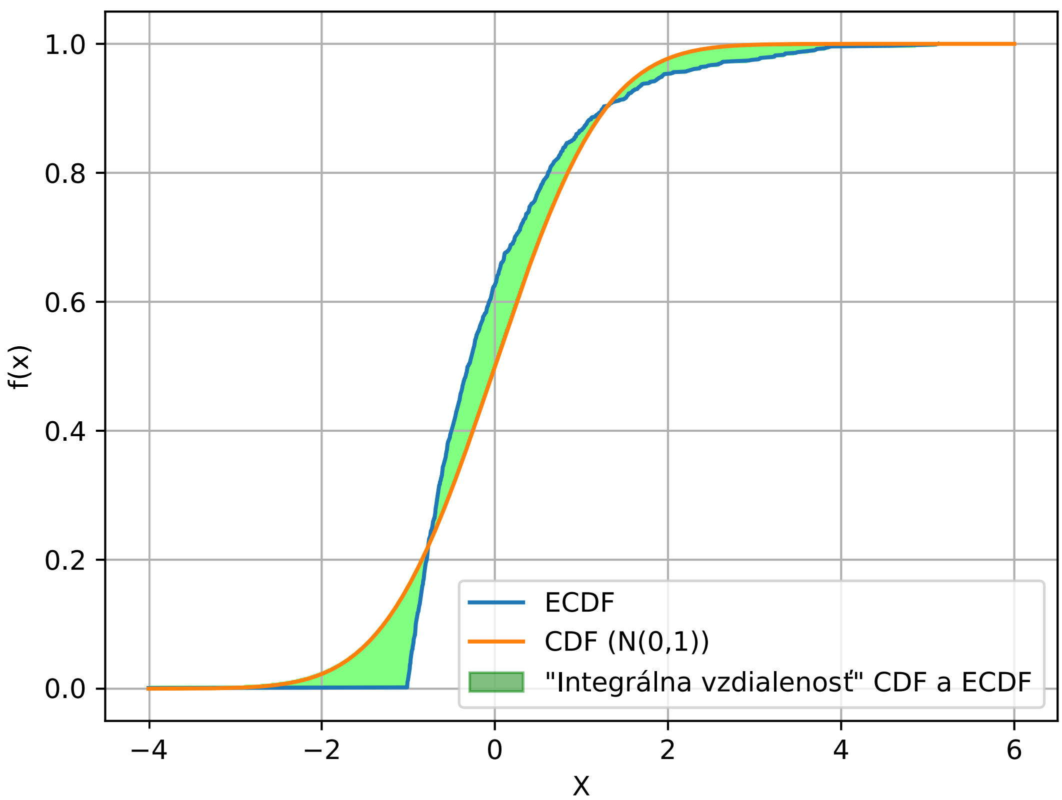
Teda:

$$E(\widehat{F_n}(x))^2 = \frac{\Phi(x)(1-\Phi(x))}{n} + (\Phi(x))^2$$

Dosadíme:

$$E(\widehat{F_n}(x) - \Phi(x))^2 = n\left(\frac{\Phi(x)(1-\Phi(x))}{n} + (\Phi(x))^2\right) - n(\Phi(x))^2 =$$
$$= \Phi(x)(1-\Phi(x))$$

$\square$



Obrázek 2.2: *Znázornenie A-D štatistkiky*
Modrou je znázornená empirická distribučná funkcia transformovaného náhodného výberu so skutočným rozdelením $Exp(1)$. Oranžová funkcia je $\Phi(x)$. Zelenou farbou je označená plocha medzi funkciami.

Dôležitou vlastnosťou testovej štatistiky je, že poznáme jej asymptotické rozdelenie. Je síce dosť škaredé a spravidla sa počíta numerickými metódami, ale už v práci Anderson a Darling (1954, str.768) je uvedený jednoduchší postup ako získať jeho aproximáciu na 4 platné desatinné miesta.

Pretože potrebujeme testovať nulovú hypotézu a alternatívu;

$$H_0 : \exists\mu \in \mathbb{R}, \exists\sigma \in \mathbb{R}, \sigma \neq 0 : X \sim N(\mu,\sigma^2)$$

$$H_1 : \forall\mu \in \mathbb{R}, \forall\sigma \in \mathbb{R}, \sigma \neq 0 : (X \nsim N(\mu,\sigma^2))$$

musíme A-D test upraviť. Jediný rozdiel v testovej štatistike oproti jednoduchšej H_0 je, že úplne na začiatku náhodný výber "znormujeme" takýmto spôsobom:

$$Y_i = \frac{X_i - \overline{X}_n}{\sqrt{{S_n}^2}},\ \forall i \in \{1,..,n\}$$

Kde $\{Y_i\}, i \in \{1,...,n\}$ je transformovaný náhodný výber, s ktorým bude test ďalej pracovať. Pre predstavu testovej štatistiky viď. obr.2.2. Podobne ako v prípade K-S testu normality sa kvôli použitiu výberového priemeru a výberového rozptylu v testovej štatistike drasticky zmenia vlastnosti testu. Stredná hodnota a rozptyl testovej štatistiky totiž závisia na tom, ktoré parametre normálneho rozdelenia zafixujeme v rámci H_0. Asymptotické rozdelenie testovej štatistiky poznáme, je overené metódou Monte-Carlo, ale vôbec nie je pekné (viď. Stephens (1976)).

2.5 D'Agostino-Pearsonov test normality

Tento test zamieta na základe šikmosti a špicatosti rozdelenia náhodného výberu. Vlastne sa jedná o kombináciu 2 rôznych testov normality, preto sa s ním môžeme v literatúre stretnúť pod názvom "Omnibus normality test". Prvý sa týmto prístupom k testovaniu normality zaoberal E.S.Pearson okolo roku 1930.

Definícia 7 (Šikmosť rozdelenia)**.** *Nech X je náhodná veličina, $var(X) = \phi^2$, $\phi \in (0,\infty)$ a $E(X) = \mu$, $\mu \in \mathbb{R}$, potom šikmosťou rozdelenia X rozumieme:*

$$\gamma_1(X) = E\left(\frac{X-\mu}{\phi}\right)^3$$

Pre predstavu viď. obr.2.3.

Definícia 8 (Špicatosť rozdelenia)**.** *Nech X je náhodná veličina, $var(X) = \phi^2$, $\phi \in (0,\infty)$ a $E(X) = \mu$, $\mu \in \mathbb{R}$, potom šikmosťou rozdelenia X rozumieme:*

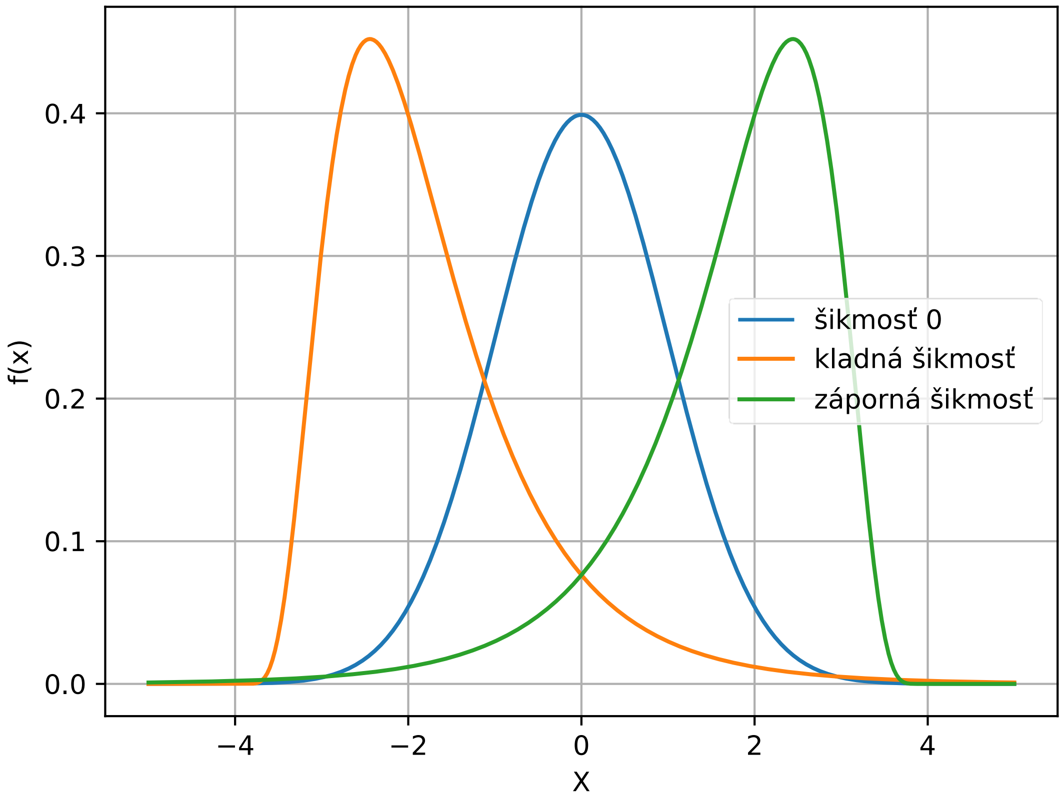
$$\gamma_2(X) = E\left(\frac{X-\mu}{\phi}\right)^4$$

Pre predstavu viď. obr.2.4.

Z vlastnosti normálneho rozdelenia Tvrdenie(1) plynie, že šikmosť a špicatosť nedegenerovaného normálneho rozdelenia nezávisia na jeho strednej hodnote ani rozptyle. Pripomenieme vzorec na výpočet momentov normovaného normálneho rozdelenia:

$$E(X)^k = \begin{cases} 0, & \text{pre nepárne } k \in \mathbb{N} \\ (k-1)!!, & \text{pre párne } k \in \mathbb{N} \end{cases}$$

Teda každé normálne rozdelenie má šikmosť $\gamma_1 = 0$ a špicatosť $\gamma_2 = 3$.



Obrázek 2.3: *Znázornenie Šikmosti*
Hustota normovaného normálneho rozdelenia má modrú farbu. Oranžovou je hustota rozdelenia Lognorm(-4,2), zelenou je hustota zrkadlového lognormálneho rozdelenia s rovnakými koeficientmi.

Prirodzené odhady šikmosti a špicatosti náhodného výberu získané momentovou metódou sú:

Výberová šikmosť:

$$\hat{\gamma}_1 = \frac{m_3}{{m_2}^{3/2}}$$

Výberová špicatosť:

$$\hat{\gamma}_2 = \frac{m_4}{{m_2}^2}$$

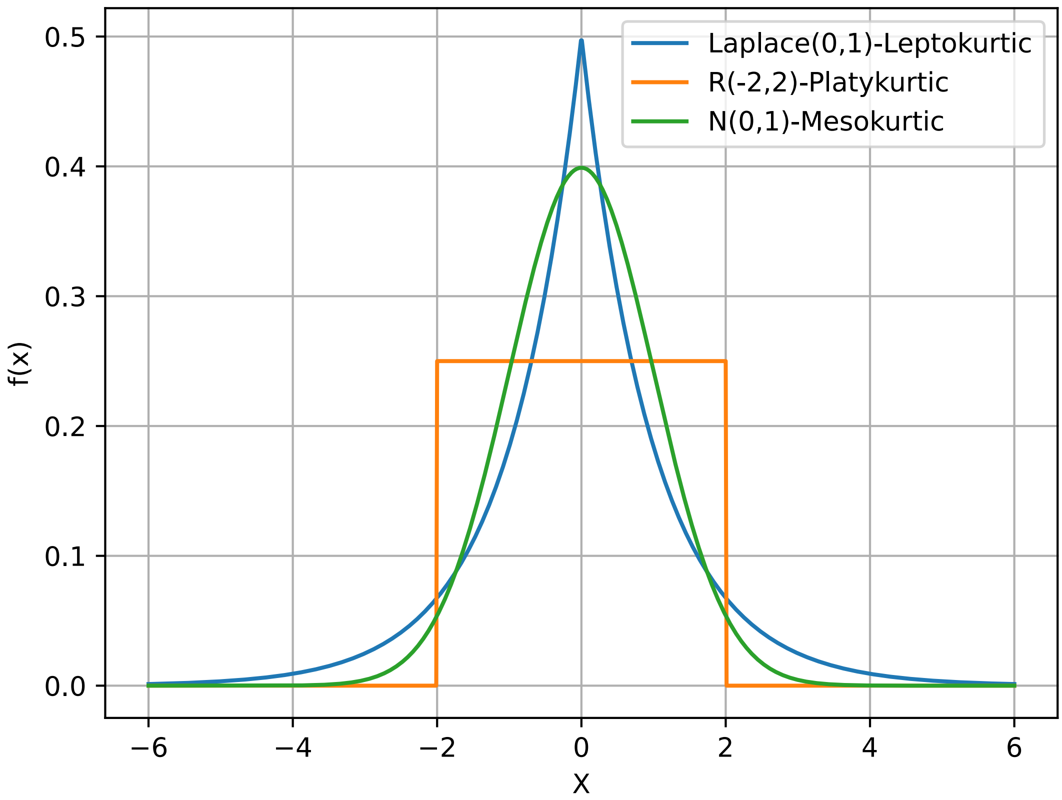
Kde $m_k = \sum_{i=1}^{n}(X_i - \overline{X}_n)^k$.

Testová štatistika:

$$O_n = (Z_1(\hat{\gamma}_1))^2 + (Z_2(\hat{\gamma}_2))^2$$

Detailne rozpísané transformácie $Z_1(\hat{\gamma}_1)$ a $Z_2(\hat{\gamma}_2)$ sú uvedené v prílohách (A.1). Komentár k tvaru transformácií sa nachádza v práci D'Agostino a Belanger (1990).

Testové štatistiky $Z_1(\hat{\gamma}_1)$ a $Z_2(\hat{\gamma}_2)$ majú obe, za platnosti nulovej hypotézy, približne normované normálne rozdelenie. Teda testová štatistika má za platnosti nulovej hypotézy asymptotické rozdelenie $O_n \overset{\text{as}}{\approx} \chi^2_2$ (viď. D'Agostino a Pearson (1973)). Zmienená práca je primárny zdroj k tomuto testu. Kritický obor má tvar $[0, c_n), c_n \in \mathbb{R}$, kde c_n sú tabelované (pre menšie náhodné výbery), získané metódou Monte-Carlo. Pre väčšie náhodné výbery sa miesto c_n použije $Q(0.95)$, teda kvantil χ^2_2 rozdelenia. To znamená, že zamietame pre veľké hodnoty testovej štatistiky.



Obrázek 2.4: *Znázornenie Špicatosti*
Na obrázku vidíme hustoty troch rozdelení. Laplaceovo rozdelenie je zrejme špicatejšie (resp. má ťažšie chvosty, angl. Leptokurtic) ako normálne, rovnomerné rozdelenie je špicaté menej (a chvosty vôbec nemá, angl. Platykurtic). O rozdeleniach so špicatosťou menšou ako normálne rozdelenie budeme vravieť, že sú placaté.

2.6 Jarque-Berrov test normality

Rovnako ako D'Agostino-Pearsonov využíva aj tento test charakteristickú šikmosť a špicatosť normálneho rozdelenia.

Testová štatistika:

$$J_n = \frac{n}{6}\left(\hat{\gamma}_1{}^2 + \frac{(\hat{\gamma}_2 - 3)^2}{4}\right)$$

Vo vyjadrení testovej štatistiky sme použili značenie zavedené pre D'Agostino-Pearsonov test.

Zaujímavosťou je, že tento test bol prvý krát spomenutý v práci Bowman a Shenton (1975). Meno Jarque-Bera nesie, pretože v článku Jarque a Bera (1987) autori odvodili testovú štatistiku metódou Lagrangeovho multiplikátora.

Na základe tejto skutočnosti sú testu pripisované podobné asymptotické vlastnosti ako má test pomerom vierohodností. Kritický obor má rovnaký tvar ako pri D'Agostino-Pearsonovom teste (tabelované konštanty sú samozrejme iné).

3. Simulačná štúdia

V tejto kapitole uvedieme výsledky Monte-Carlo simulácie 11 rôznych scenárov. Výsledky 9 scenárov budú grafickej formy s menším počtom opakovaní ($o = 1000$), 2 výsledky s väčším počtom opakovaní ($o = 10000$, pre kontrolu) budú v tabuľkách. Začneme empirickou kontrolou deklarovanej hladiny $\alpha = 0.05$. Potom sa budeme zaoberať pozorovanou silou testov normality z kapitoly 2 proti alternatívam:

1. beta rozdelenie
2. Cauchyho rozdelenie
3. exponenciálne rozdelenie
4. gamma rozdelenie
5. Laplaceovo rozdelenie
6. lognormálne rozdelenie
7. Paretovo rozdelenie
8. Studentovo rozdelenie

Viac, ako sily testov proti daným alternatívam pre konkrétne veľkosti náhodného výberu, nás bude zaujímať relatívny výkon testov (budeme testy porovnávať medzi sebou). Pretože budeme porovnávať na základe grafov, málo významné rozdiely medzi testami budú zanedbané. Ku každej alternatíve uvedieme jej vlastnosti. Výsledkom simulačnej štúdie bude priradenie testov k rodinám rozdelení, voči ktorým sú silnejšie. V prípade K-S testu sa overí nutnosť Lillieforsovej korekcie.

Všeobecne prevláda názor, že na náhodný výber menšieho rozsahu je najvhodnejšie použiť S-W test normality (zhodnú sa na tom skoro všetky zdroje). Preto sa budeme zaoberať testovaním náhodných výberov veľkosti $n \in \{20{,}300\}$.

Pracovali sme v programe Python a použili štandardnú implementáciu testov normality z modulu "stats" z knižnice "scipy"a
z modulu "statsmodels.stats.diagnostic".

3.1 Kontrola deklarovanej hladiny

Odhad skutočnej hladiny testu α_n (resp. chyby I. druhu), pre náhodný výber dĺžky n, budeme pre každý test počítať týmto spôsobom:

$$\hat{\alpha}_n = \frac{Z}{o}$$

Kde n je veľkosť náhodného výberu, $o = 1000$ (resp. $o = 10000$) je počet opakovaní sceára a $Z \leq o$ je množstvo náhodných výberov, pre ktoré príslušný test zamietol H_0.

Pre každé $n \in \mathbb{N}$ sa jedná o odhad nestranný a konzistentný, pretože náhodná veličina Z má rozdelenie $Z \sim Binom(\alpha_n, o)$.

Rovnakú metódu výpočtu použijeme aj pre odhad síl testov proti alternatívam. V tom prípade bude mať náhodná veličina Z rozdelenie $Z \sim Binom(\beta_n, o)$, kde β_n je skutočná sila konkrétneho testu pre veľkosť náhodného výberu n proti danej alternatíve.

Podľa grafu 3.1 sa dá povedať, že významnému porušeniu hladiny pre náhodné výbery veľkostí $n \in \{20,300\}$ sa vyhli všetky testy, a K-S test má problém s konzervativitou. Túto domnienku potvrdzujú výsledky v tabuľke 3.1.Teda má zmysel zaoberať sa ďalej silou týchto testov pre skúmanú veľkosť náhodného výberu.

3.2 Pozorovanie sily testov

Cauchyho rozdelenie:

Toto rozdelenie síce nemá dobre definovanú šikmosť a špicatosť, ale je známe svojimi ťažkými chvostami a symetriou preto ho pre naše účely prehlásime za špicaté so šikmosťou $\gamma_1 = 0$.

Na grafe 3.2 vidíme odhad sily testov proti Cauchyho rozdeleniu. Všetky testy si počínali celkom dobre, aj keď K-S test výrazne zaostával.

Exponenciálne rozdelenie:

Exponenciálne rozdelenie je asymetrické s kladnou šikmosťou (má pravý chvost). Ďalej je špicaté a nenulovú hustotu má iba na kladných číslach.

Na grafe 3.3 vidíme odhad sily testov proti exponenciálnemu rozdeleniu. Testy zoradíme od najsilnejších po najslabšie (malé rozdiely v sile zanedbáme):

1. D'Agostino-Pearson a Shapiro-Wilk
2. Anderson-Darling, Jarque-Bera a Lilliefors
3. Kolmogorov-Smirnov

Beta rozdelenie:

Špicatosť a šikmosť beta rozdelenia závisí na jeho parametroch. V prípade $\alpha = \beta = 2$ má toto rozdelenie šikmosť $\gamma_1 = 0$ a je placaté. Bez ohľadu na parametre má beta rozdelenie nenulovú hustotu iba na intervale $[0,1]$.

Na grafe 3.4 vidíme odhad sily testov proti beta rozdeleniu. V tomto prípade sa sila testov dosť líši. Od najlepšieho po najhorší zoradíme testy (s ľahkou ujmou na obecnosti) takto:

1. D'Agostino-Pearson
2. Shapiro-Wilk
3. Anderson-Darling
4. Jarque-Bera
5. Lilliefors

6. Kolmogorov-Smirnov

Paretovo rozdelenie:

Kvôli hodnote parametra $\alpha = 1$ nie je špicatosť tohoto rozdelenia dobre definovaná. Pretože má ťažké chvosty, pre naše účely ho prehlásime za špicaté. Paretovo rozdelenie je kladne šikmé, teda asymetrické, a nenulovú hustotu má iba na intervale $[x_m,\infty)$.

Na grafe 3.5 vidíme odhad sily testov proti Paretovmu rozdeleniu. Všetky testy si počínali celkom dobre, aj keď K-S test zaostával.

Gamma rozdelenie:

Toto rozdelenie je bez ohľadu na parametre kladne šikmé (teda asymetrické s pravým chcvostom) a špicaté. Nenulovú hustotu má iba na kladných číslach.

Na grafe 3.6 vidíme odhad sily testov proti gamma rozdeleniu. Sila testov je pomerne rôznorodá, preto ich môžeme zmysluplne zoradiť od najsilnejších po najslabšie (nie významné rozdiely opäť zanedbáme):

1. Shapiro-Wilk
2. D'Agostino-Pearson, Anderson-Darling a Jarque-Bera
3. Lilliefors
4. Kolmogorov-Smirnov

Lognormálne rozdelenie:

Lognormálne rozdelenie je bez ohľadu na parametre kladne šikmé a špicaté. Nenulovú hustotu má iba na kladných číslach.

Na grafe 3.7 je odhad sily testov proti lognormálnemu rozdeleniu. Sily testov nie sú dramaticky odlišné, ale vidíme, že zoradené od najsilnejších po najslabšie by mali vyzerať asi takto:

1. Shapiro-Wilk
2. Anderson-Darling
3. D'Agostino-Pearson a Jarque-Bera
4. Lilliefors
5. Kolmogorov-Smirnov

Laplaceovo rozdelenie:

Toto rozdelenie má šikmosť $\gamma_1 = 0$ a je špicaté. Zmienené vlstnosti na parametroch nezávisia.

Na grafe 3.8 vidíme odhad sily testov proti Laplaceovmu rozdeleniu. Rozdiely síl testov sú v tomto prípade ťažko rozoznateľné, ale môžeme sa pokúsiť ich zoradiť:

1. Anderson-Darling
2. Shapiro-Wilk, D'Agostino-Pearson a Jarque-Bera
3. Lilliefors
4. Kolmogorov-Smirnov

Studentovo rozdelenie:

Pretože v našom prípade počet stupňov voľnosti je 7, sú šikmosť $\gamma_1 = 0$ a špicatosť (rozdelenie je špicaté) dobre definované.

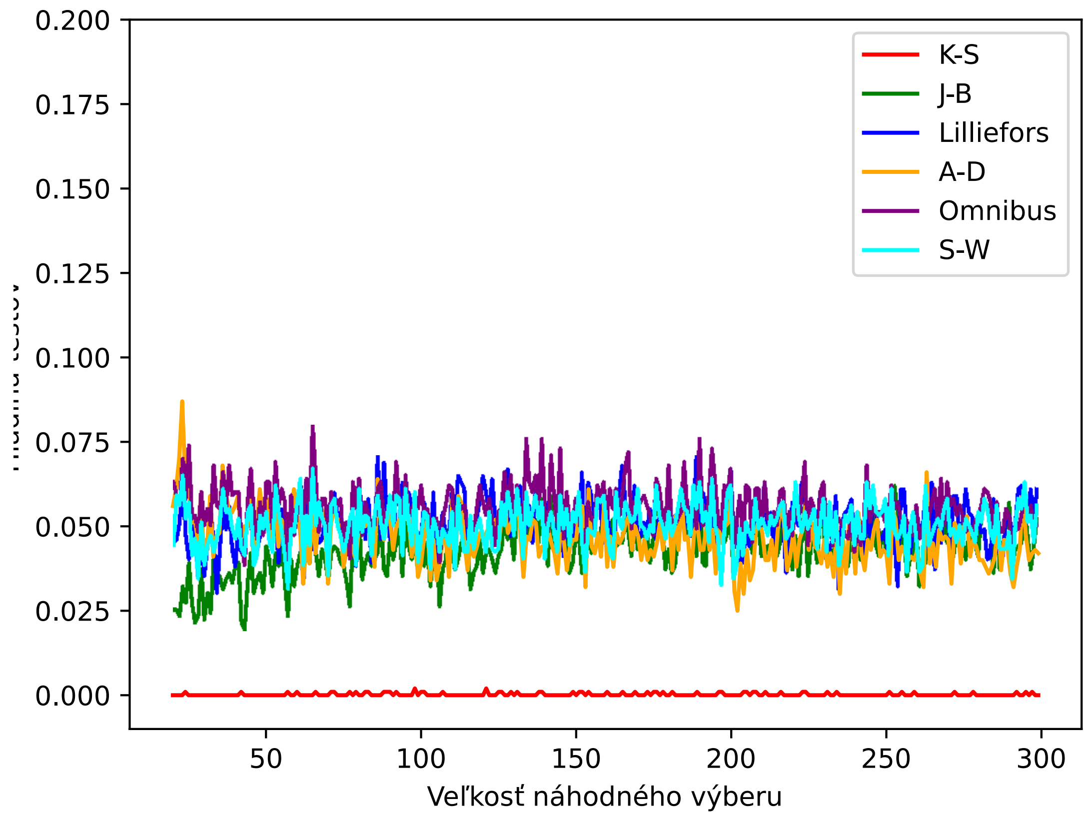
Na grafe 3.9 vidíme odhad sily testov proti Studentovmu rozdeleniu. Rozdiely sily testov síce nie sú veľké, ale v tomto prípade je ich poradie relatívne jednoznačné:

1. Jarque-Bera
2. D'Agostino-Pearson
3. Shapiro-Wilk
4. Anderson-Darling
5. Lilliefors
6. Kolmogorov-Smirnov

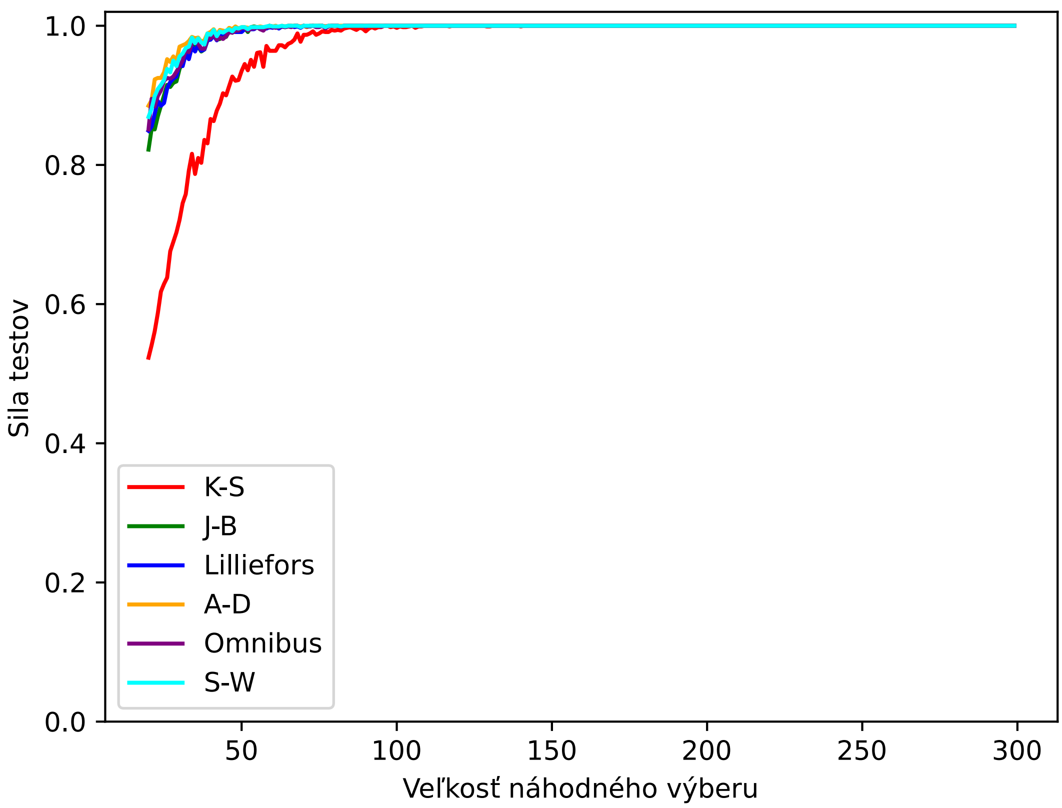
3.3 Diskusia

V rámci rodiny testov používajúcich empirickú distribučnú funkciu sú výsledky našich simulácií celkom jednoznačné (a potvrdené výsledkami s vyššou presnosťou v tabuľkách 3.2,3.3). Vo všetkých prípadoch bol A-D test striktne lepší ako test Lillieforsov, ktorý bol striktne lepší ako K-S test. Overili sme, že Lillieforsova korekcia K-S testu je skutočne potrebná. Klasický K-S test je totiž silne konzervatívny, najlepšie to vidno pri testovaní rozdelení beta a Student. Čo sa týka štvorice testov A-D, S-W, Omnibus a J-B, nemôžeme ani o jednom z nich povedať, že je striktne lepší ako ostatné. Toto tvrdíme, lebo sme pre každý z týchto testov našli rozdelenie, voči ktorému bol najsilnejší. Ďalej si všimneme, že S-W test fungoval pozoruhodne dobre na všetky rozdelenia, ktoré majú nenulovú hustotu iba na kladných číslach (aj na beta rozdelenie). Tento test bol nadpriemerne silný aj proti všetkým šikmým rozdeleniam (túto vlastnosť by sme očakávali skôr od Omnibus testu). Netvrdíme, že S-W test je najlepší proti každému rozdeleniu, ktoré je šikmé, alebo má nenulovú hustotu iba na kladných číslach. Chceme len zdôrazniť, že túto hypotézu výsledky simulácií nevylučujú.

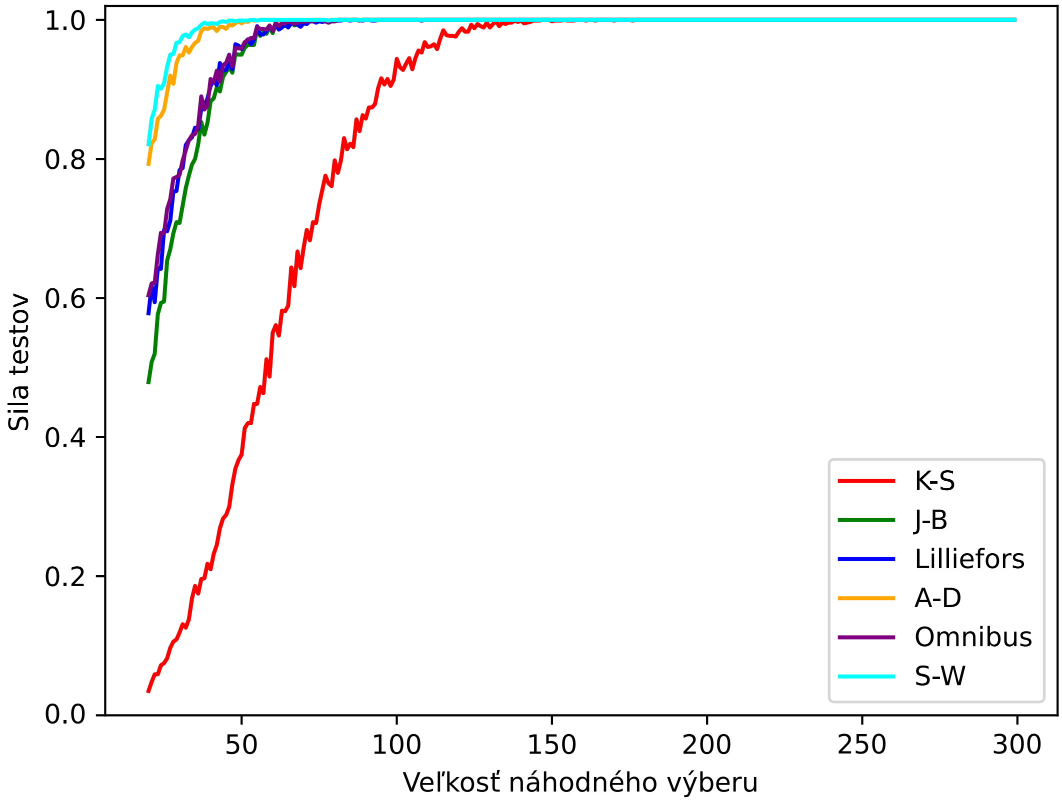
Nakoniec chceme pripomenúť, že naše simulácie prebehli iba pre jednu hladinu $\alpha = 0.05$, relatívne úzky interval veľkostí náhodných výberov $n \in \{20,...,300\}$ a 9 rozdelení (z toho iba 6 malo pre nás nejakú výpovednú hodnotu). Tým chceme povedať, že všeobecné a bezpečné závery sa na základe takej malej vzorky vyvodiť nedajú.



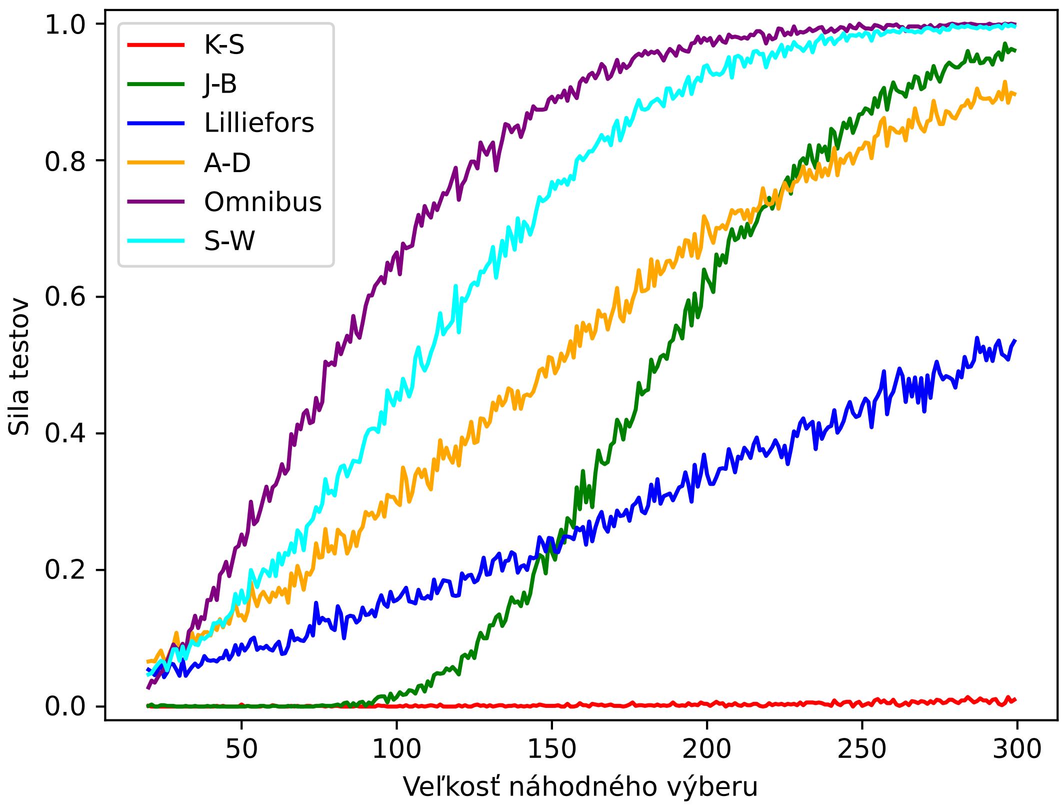
Obrázek 3.1: *Graf odhadu chyby I. druhu*
Náhohodné výbery boli z N(0,10) rozdelenia.



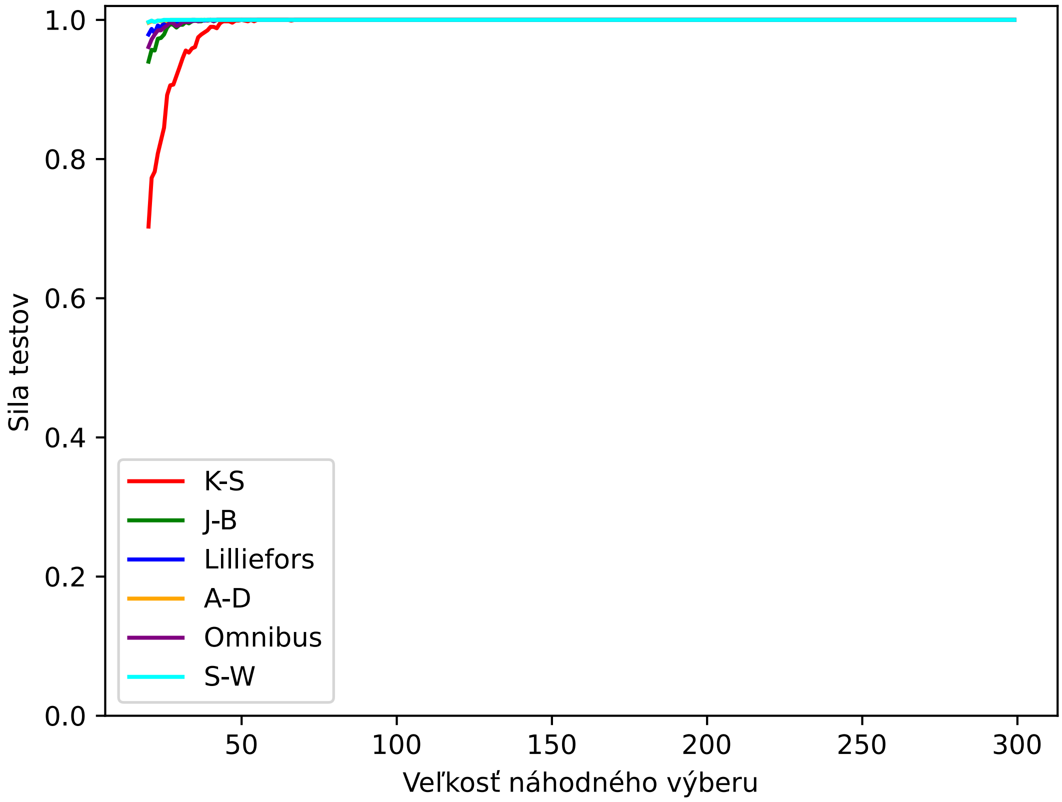
Obrázek 3.2: Na grafe je odhad sily testov normality proti Cauchyho rozdeleniu.



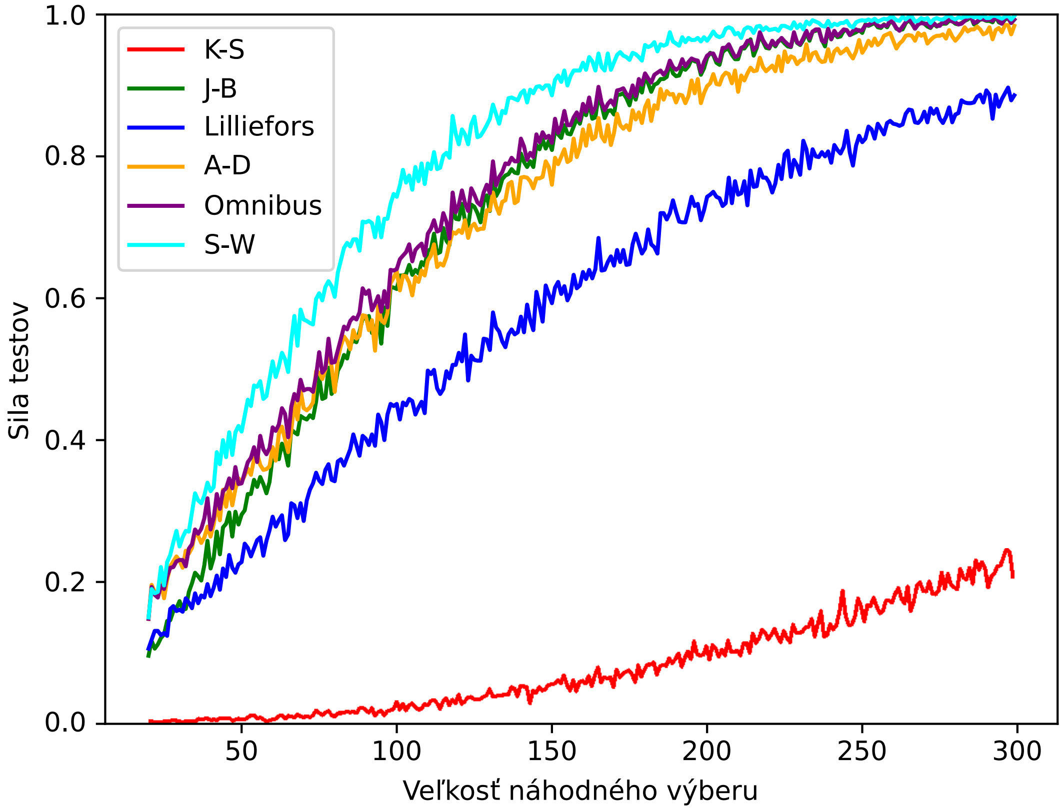
Obrázek 3.3: Na grafe je odhad sily testov normality proti exponenciálnemu rozdeleniu s parametrom $\lambda = 1$.



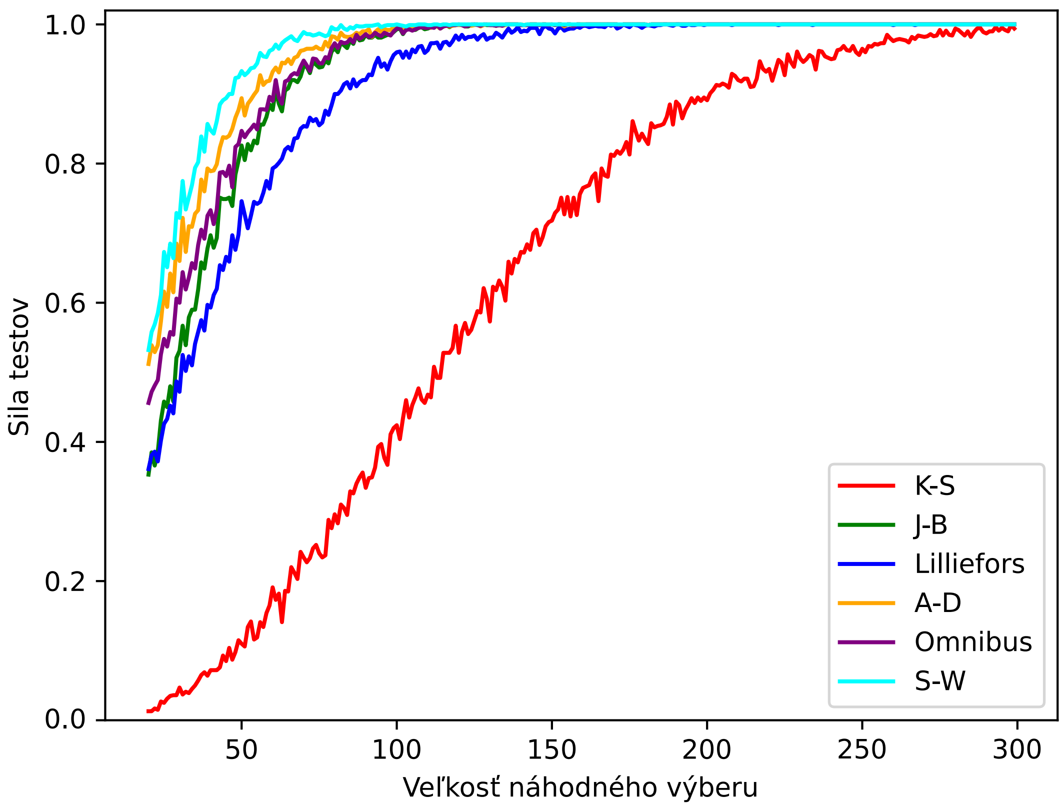
Obrázek 3.4: Na grafe je odhad sily testov normality proti beta rozdeleniu s parametrom $\alpha = \beta = 2$.



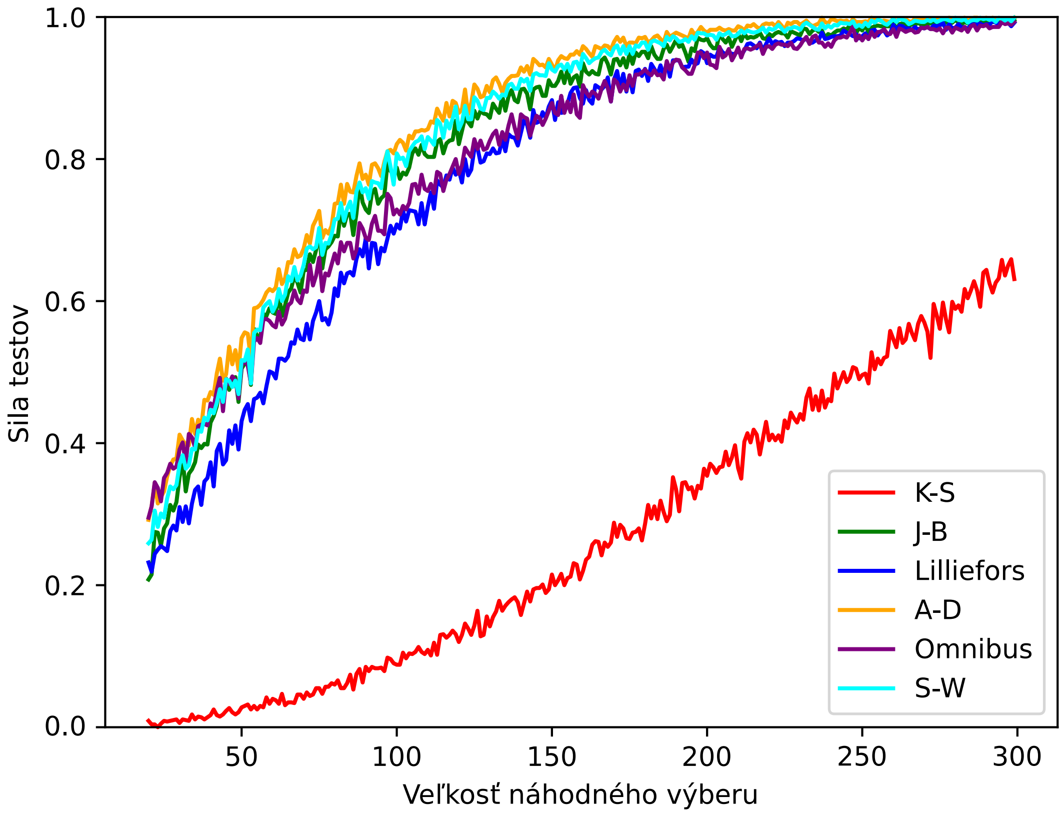
Obrázek 3.5: Na grafe je odhad sily testov normality proti Paretovmu rozdeleniu s parametrami $\alpha = 1, x_m = 0$.



Obrázek 3.6: Na grafe je odhad sily testov normality proti gamma rozdeleniu s parametrami $\alpha = 7.5, \beta = 1$.



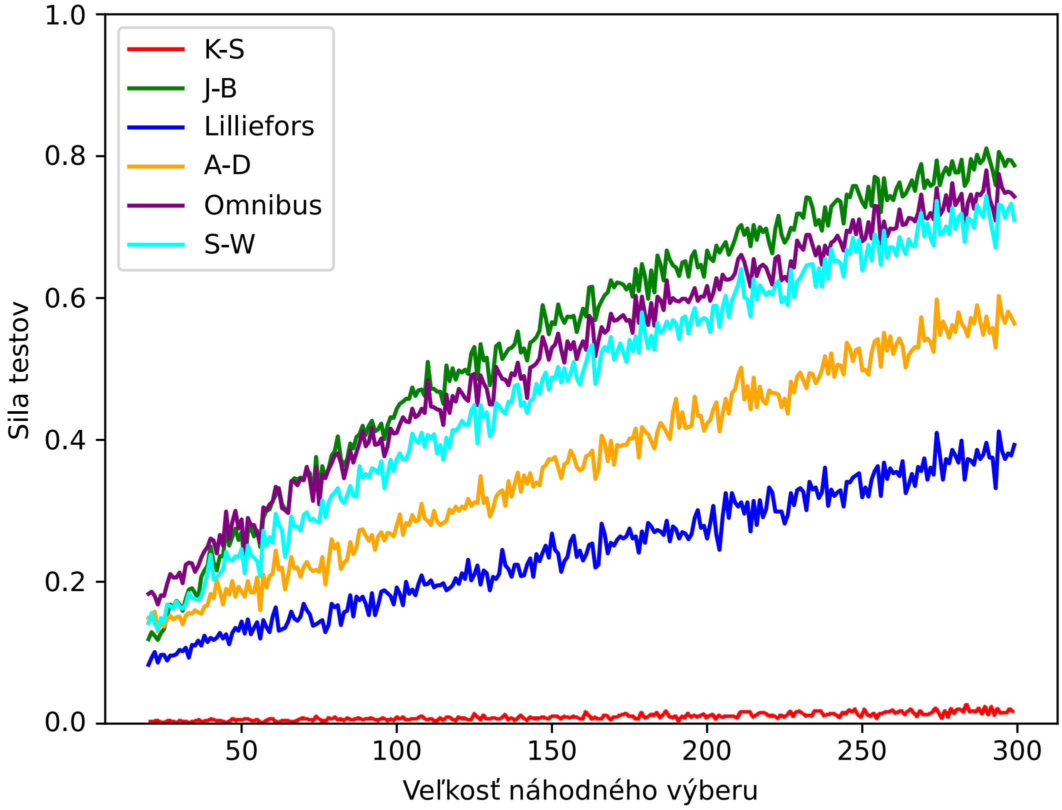
Obrázek 3.7: Na grafe je odhad sily testov normality proti lognormálnemu rozdeleniu s parametrami $\mu = 0.5, \sigma = 0.2$.



Obrázek 3.8: Na grafe je odhad sily testov normality proti Laplaceovmu rozdeleniu s parametrami $\mu = 0, b = 4$.

Testy	**Veľkosti náhodného výberu**				
—	n = 20	n = 50	n = 100	n = 200	n = 300
K-S	0.0001	0.0001	0.0003	0.0002	0.0003
Lilliefors	0.0479	0.0483	0.0490	0.0506	0.0502
S-W	0.0471	0.0518	0.0506	0.0515	0.0484
A-D	0.0580	0.0517	0.0486	0.0472	0.0413
Omnibus	0.0534	0.0556	0.0544	0.0572	0.0511
J-B	0.0211	0.0390	0.0408	0.0475	0.0452

Tabulka 3.1: Odhad hladiny testov za platnosti nulovej hypotézy.
Značenie: Kolmogorov-Smirnov (K-S), Shapiro-Wilk (S-W),
Anderson-Darling (A-D), D'Agostino-Pearson (Omnibus), Jarque-Bera (J-B)



Obrázek 3.9: Na grafe je odhad sily testov normality proti Studentovmu rozdeleniu so 7 stupňami voľnosti (resp. $\nu = 7$).

Testy	Veľkosti náhodného výberu				
—	n = 20	n = 50	n = 100	n = 200	n = 300
Empirická sila testu proti alternatíve beta(2,2)					
K-S	0.0001	0.0002	0.0002	0.0022	0.0084
Lilliefors	0.0483	0.0810	0.1516	0.3345	0.5412
S-W	0.0547	0.1532	0.4508	0.9259	0.9974
A-D	0.0724	0.1344	0.3008	0.6816	0.8991
Omnibus	0.0386	0.2414	0.6615	0.9721	0.9986
J-B	0.0015	0.0003	0.0123	0.6104	0.9616
Empirická sila testu proti alternatíve Cauchy					
S-W	0.8622	0.9958	0.9999	0.9999	0.9999
Omnibus	0.8578	0.9924	0.9999	0.9999	0.9999
A-D	0.8860	0.9966	0.9999	0.9999	0.9999
Lilliefors	0.8385	0.9937	0.9999	0.9999	0.9999
J-B	0.8160	0.9929	0.9999	0.9999	0.9999
K-S	0.4957	0.9297	0.9990	0.9999	0.9999
Empirická sila testu proti alternatíve exp(1)					
K-S	0.0446	0.3730	0.9310	0.9999	0.9999
Lilliefors	0.5762	0.9635	0.9998	0.9999	0.9999
S-W	0.8351	0.9995	0.9999	0.9999	0.9999
A-D	0.7960	0.9972	0.9999	0.9999	0.9999
Omnibus	0.6068	0.9677	0.9999	0.9999	0.9999
J-B	0.4846	0.9592	0.9998	0.9999	0.9999
Empirická sila testu proti alternatíve gamma(7.5,1)					
K-S	0.0006	0.0041	0.0208	0.0979	0.2362
Lilliefors	0.1113	0.2340	0.4427	0.7343	0.8946
S-W	0.1715	0.4220	0.7398	0.9685	0.9982
A-D	0.1702	0.3379	0.6104	0.8956	0.9834
Omnibus	0.1704	0.3537	0.6359	0.9368	0.9953
J-B	0.1000	0.3010	0.6082	0.9317	0.9952

Tabulka 3.2: Kontrola odhadu sily testov prvá časť.
Značenie: Kolmogorov-Smirnov (K-S), Shapiro-Wilk (S-W),
Anderson-Darling (A-D), D'Agostino-Pearson (Omnibus), Jarque-Bera (J-B)

Testy	**Veľkosti náhodného výberu**				
—	n = 20	n = 50	n = 100	n = 200	n = 300
Empirická sila testu proti alternatíve Laplace(0,4)					
K-S	0.0045	0.0253	0.0966	0.3540	0.6488
Lilliefors	0.2170	0.4273	0.7038	0.9435	0.9926
S-W	0.2624	0.5215	0.7961	0.9746	0.9980
A-D	0.2986	0.5503	0.8215	0.9803	0.9989
Omnibus	0.3055	0.5065	0.7390	0.9411	0.9897
J-B	0.2164	0.5097	0.7828	0.9648	0.9959
Empirická sila testu proti alternatíve lognorm(0.5,0.2)					
K-S	0.0162	0.1121	0.4157	0.8945	0.9907
Lilliefors	0.3415	0.7037	0.9499	0.9992	0.9999
S-W	0.5228	0.9242	0.9988	0.9999	0.9999
A-D	0.4998	0.8748	0.9933	0.9998	0.9999
Omnibus	0.4423	0.8253	0.9913	0.9999	0.9999
J-B	0.3370	0.7995	0.9899	0.9999	0.9999
Empirická sila testu proti alternatíve Pareto(1,0)					
K-S	0.7053	0.9988	0.9999	0.9999	0.9999
Lilliefors	0.9792	0.9999	0.9999	0.9999	0.9999
S-W	0.9968	0.9999	0.9999	0.9999	0.9999
A-D	0.9959	0.9999	0.9999	0.9999	0.9999
Omnibus	0.9638	0.9999	0.9999	0.9999	0.9999
J-B	0.9420	0.9999	0.9999	0.9999	0.9999
Empirická sila testu proti alternatíve Student(7)					
K-S	0.0009	0.0032	0.0056	0.0117	0.0201
Lilliefors	0.0928	0.1218	0.1831	0.2821	0.3968
S-W	0.1329	0.2339	0.3618	0.5807	0.7302
A-D	0.1328	0.1857	0.2681	0.4300	0.5807
Omnibus	0.1709	0.2793	0.4032	0.6166	0.7566
J-B	0.1137	0.2660	0.4295	0.6636	0.8006

Tabulka 3.3: Kontrola odhadu sily testov druhá časť.
Značenie: Kolmogorov-Smirnov (K-S), Shapiro-Wilk (S-W),
Anderson-Darling (A-D), D'Agostino-Pearson (Omnibus), Jarque-Bera (J-B)

Záver

V rámci prvej časti tejto práce sme stručne popísali niekoľko známych testov normality. Konkrétne sa jedná o testy Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk, Anderson-Darling, D'Agostino-Pearson a Jarque-Bera. Venovali sme sa rôznym prístupom k testovaniu normality. Niektoré vychádzajú z empirickej distribučnej funkcie a merajú jej odlišnosť od distribučnej funkcie nejakého normálneho rozdelenia. Iné používajú odhady šikmosti a špicatosti, ktoré sú charakteristické pre normálne rozdelenie, a zaoberajú sa vlastnosťami ich transformácií (a nepotrebujú predpoklad spojitosti skutočného rozdelenia náhodného výberu). Dôraz bol kladený na to, aby si vedel nejakú predstavu o princípoch jednotlivých testov vytvoriť aj čitateľ, ktorý ich možno vidí prvý krát.

Ďalej sme skúmali skutočnú silu predstavených testov. Druhá časť obsahuje výsledky simulačnej štúdie prevedenej metódou Monte-Carlo v programe Python. Táto štúdia nám síce na konci neodhalila jeden, jednoznačne najlepší, test normality, ale nejaké závery sa z nej iste vyvodiť dajú. Kolmogorov-Smirnov test bez korekcie by sa na testovanie normality používať nemal, a ani s korekciou väčšinou nie je taký efektívny ako ostatné zmienené testy. Jarque-Berov test má pre výbery malej veľkosti konzervatívne sklony, ale pre väčšie náhodné výbery sa v sile vyrovná zvyšným testom. Prečo je ťažké tieto testy porovnať sme už zmienili v diskusii. Na základe tejto štúdie odporúčame používať Shapiro-Wilkov test, keď to z nejakého dôvodu čitateľovi nevyhovuje, Anderon-Darlingov a D'Agostino-Pearsonov test sú tiež dobrá voľba.

Zoznam použitej literatúry

Anderson, T. a Darling, D. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, **49**, 765–769.

Bowman, K. a Shenton, L. (1975). Omnibus contours for departures from normality based on $\sqrt{b_1}$ and b_2. *Biometrika*, **62**, 243–250.

D'Agostino, R. a Belanger, A. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, **44**, 316–321.

D'Agostino, R. a Pearson, E. (1973). Tests for departure from normality. empirical results for distributions of b_2 and $\sqrt{b_1}$. *Biometrika*, **60**, 613–622.

Jarque, C. a Bera, A. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, **55**, 163–172.

Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione. *Inst. Ital. Attuari*, **4**, 1–11.

Kulich, M. a Omelka, M. (2022). NMSA331 matematická statistika 1 - poznámky k přednášce. `https://www2.karlin.mff.cuni.cz/~komarek/vyuka/2022_23/nmsa331/ms1.pdf`. Navštívené: 2023-03-23.

Lilliefors, H. (1967). On the kolmogorov-smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, **62**, 399–402.

Macoun, J. (2021). Bakalářská práce - Lillieforsův test normality. `https://dspace.cuni.cz/bitstream/handle/20.500.11956/127887/130305186.pdf?sequence=1&isAllowed=y`. Navštívené: 2023-03-23.

Malíková, K. (2014). Bakalářská práce - Shapirův-Wilkův test normality. `https://dspace.cuni.cz/bitstream/handle/20.500.11956/73041/BPTX_2013_1_11320_0_317024_0_140283.pdf?sequence=1&isAllowed=y`. Navštívené: 2023-03-23.

Royston, P. (1995). A remark on algorithm as 181: The w-test for normality. *The Journal of the Royal Statistical Society*, **44**, 547–551.

Shapiro, S. a Wilk, M. (1965). An analysis of variance tests for normality (complete samples). *Biometrika*, **52**, 591–611.

Stephens, M. (1976). Asymptotic results for goodness-of-fit statistics with unknown parameters. *The Annals of Statistics*, **4**, 357–369.

A. Prílohy

A.1 Transformácie testovej štatistiky Omnibus testu

$$
\begin{aligned}
A &= \hat{\gamma}_1 \left(\frac{(n+1)(n+3)}{6(n-2)} \right)^{1/2} \\
B &= \frac{3(n^2+27n-70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)} \\
C &= (2B-1)^{1/2} - 1 \\
D &= \sqrt{\frac{2}{ln(C)}} \\
E &= \sqrt{\frac{2}{C-1}} \\
Z_1(\hat{\gamma}_1) &= D * ln \left(\frac{A}{E} + \left[\left(\frac{A}{E} \right)^2 + 1 \right]^{1/2} \right) \\
F &= \frac{3(n-1)}{n+1} \\
G &= \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)} \\
H &= \frac{\hat{\gamma}_2 - F}{\sqrt{G}} \\
I &= \frac{6(n^2-5n+2)}{(n+7)(n+9)} \left(\frac{6(n+3)(n+5)}{n(n-2)(n-3)} \right)^{1/2} \\
J &= 6 + \frac{8}{I} \left(\frac{2}{I} + \left[1 + \frac{4}{I^2} \right]^{1/2} \right) \\
Z_2(\hat{\gamma}_2) &= \left(1 - \frac{2}{9J} - \left[\frac{1-2/J}{1+H\sqrt{2/(J-4)}} \right]^{1/3} \right) / \sqrt{2/(9J)}
\end{aligned}
$$