Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky

Katedra statistiky a pravděpodobnosti

Klasické a nově navržené popisné charakteristiky: porovnání výběrových vlastností na základě Monte Carlo simulace

doktorská dizertační práce

Doktorand:	Ing. Jiří Vohlídal
Školitel:	Prof. Ing. Václav Čermák, CSc.
Obor:	Statistika

Praha, duben, 2007

Prohlášení

Prohlašuji, že doktorskou práci na téma „*Klasické a nově navržené popisné charakteristiky: porovnání výběrových vlastností na základě Monte Carlo simulace*“ jsem vypracoval samostatně. Použitou literaturu a podkladové materiály uvádím v přiloženém seznamu literatury.

V Praze dne

……………………………….
podpis

Abstrakt

Na základě Monte Carlo simulace bylo provedeno porovnání estimátorů klasických a některých nově navržených měr variability, relativní variability, šikmosti a kurtozy. Z celkem 40 různých populací bylo pořízeno vždy 16 000 výběrů o rozsahu $n \in \{7; 15; 23; 31; 47; 63; 100; 200; 350; 500; 1000\}$. Z každého výběru byly vypočteny hodnoty estimátorů měr založených na momentech, kvantilech, L-momentech a jejich modifikacích a robustních měr založených na mediánu funkce lineární kombinace pořádkových statistik. Na základě experimentu bylo provedeno porovnání výběrových vlastností jednotlivých estimátorů z hlediska variability, vychýlení a rychlosti konvergence jejich výběrových rozdělení k normalitě. U estimátorů vybraných charakteristik byla dále na základě experimentálně odhadnuté průměrné síly testu posouzena vhodnost jejich použití jako testového kritéria při testu o rozdělení, z něhož výběr pochází. Zároveň byla porovnána síla závislosti estimátorů nově navržených charakteristik s estimátory momentovými s cílem posoudit, zda je možné danou charakteristiku skutečně považovat za vhodnou alternativu charakteristiky momentové.

Výsledky ukazují, že estimátory momentových měr jsou vyhovující pro popis souboru pouze při výběrech z populací nepříliš odlišných od normálního rozdělení. S rostoucí odlišností od normality rychle roste relativní variabilita i vychýlení jejich estimátorů a projevují se různé anomálie v jejich chování. Vhodnou alternativou k mírám založeným na klasických momentech i kvantilech by se mohly stát míry založené na L-momentech, jejichž estimátory vykazují ve většině případů nejlepší výběrové vlastnosti a zároveň vykazují vysokou míru závislosti s hodnotami estimátorů momentových charakteristik. Modifikace L-momentů, tzv. LQ- a TL-momenty, nepřinášejí oproti mírám založeným na L-momentech žádné zlepšení, v některých ohledech vykazují výrazně horší vlastnosti.

Klíčová slova

Metoda Monte Carlo, simulace, variabilita, relativní variabilita, šikmost, špičatost, kurtoza, L-moment, LQ-moment, TL-moment, kvantil, výběrové rozdělení, estimátor.

Abstract

A comparison of estimators of classical and some newly proposed measures of variability, relative variability, skewness and kurtosis has been done on the basis of the Monte Carlo simulation. In total 16,000 samples of the size $n \in \{7; 15; 23; 31; 47; 63; 100; 200; 350; 500; 1000\}$ were generated from 40 different populations. The estimators of moment-based, L-moment-based, LQ- and TL-moment-based, quantile-based and other measures were calculated for each sample. Sampling properties of the estimators were compared from the point of view of their variability, bias and convergence of their sampling distributions to normality. Possibility of a use of some estimators as a test statistic of the parent-distribution test was evaluated on the basis of the experimentally obtained test power. Because the newly proposed measures are being used as an alternative to the moment-based measures, existence of a statistically significant relation between the values of the newly proposed measures and the moment-based measures was also examined.

The obtained results have shown that the estimators of the moment-based measures are applicable only to populations that are close to the normal distribution. Their bias and variability increases very fast and some anomality appear in their behavior, if the parent population declines from the normal distribution. L-moment-based estimators were found to provide the best sampling properties in most cases. A strong relation between the moment-based and L-moment-based estimators was also observed. This makes L-moment-based measures a promising tool for characterisation of distribution properties, which can be used as an alternative to the classical moment-based and/or quantile-based characteristics. On the other hand, the measures based on modifications of the L-moments, i.e. LQ-moments and TL-moments do not provide any improvement in comparison with the L-moments. In some cases, their properties are even worse.

Keywords

Monte Carlo method; Simulation; Variability; Relative variability; Skewness; Kurtosis; L-moment; LQ-moment; TL-moment; Quantile; Sampling distribution; Estimator

Abstrakt

Aufgrund der Monte Carlo-Simulation wurde ein Vergleich der klassischen und einigen neu entworfenen Estimatoren der Variabilitätsmass, der relativen Variabilitätsmass, der Schiefe und Kurtosis durchgeführt. Von insgesamt 40 verschiedenen Populationen wurden immer 16 000 im Ausmass von $n \in \{7; 15; 23; 31; 47; 63; 100; 200; 350; 500; 1000\}$Stichproben getroffen. Von jeder Auswahl wurden Werte der an den auf Momenten basierten Estimatoren, Quantilen, L-Momenten und ihren Modifikationen und anderen Prinzipien berechnet. Aufgrund des Experimenten wurde ein Vergleich der Estimatoreneigenschaften der einzelnen Estimatoren durchgeführt, und zwar aus dem Gesichtspunkt der Variabilität, der Verzerrung und der Konvergenzgeschwindigkeit ihrer Stichprobenverteilung zur Normalverteilung. Bei Estimatoren gewählter Charakteristiken wurde weiter aufgrund der experimental geschätzten Durchschnittskraft des Testes die Verwendungstauglichkeit als Testkriterium bei dem Test auf Verteilung bewertet, von dem die Auswahl stammt.

Es wurde zugleich die Abhängigkeitskraft der Estimatoren der neu entworfenen Charakteristiken mit den Momentestimatoren verglichen, mit dem Ziel beurteilen zu können, ob es überhaupt möglich ist die angegebene Charakteristik in der Tat als geeignete Alternative zu der Momentcharakteristik zu betrachten. Die Ergebnisse zeigen, dass die Estimatoren der Momentmassen für die Komplexbeschreibung nur bei den nicht viel von der normalen Verteilung abweichenden Populationen geeignet sind. Mit der wachsenden Abweichung von der Normalität wächst die relative Variabilität und die Verzerrung ihrer Estimatoren und es erscheinen verschiedene Anomalien in deren Verhaltung. Geeignete Alternative zu den auf klassischen Momenten und Quantilen basierten Massen könnten die auf L-Momenten basierten Massen darstellen, deren Estimatoren in den meisten Fällen bessere Eigenschaften und zugleich einen hohen Abhängigkeitsmass mit den Werten den Estimatoren der Momentcharakteristiken nachweisen. Modifikationen der L-Momente, sog. LQ- und TL-Momente bringen im Vergleich zu den auf L-Momente basierten Werten keine Verbesserung, in einigen Hinsichten zeigen sie wesentlich schlechtere Eigenschaften auf.

Schlüsselwörter

Monte Carlo-Methode, Simulation, Variabilität, relative Variabilität, Schiefe, Kurtosis, L-Moment, LQ-Moment, TL-Moment, Quantil, Stichprobenverteilung, Estimator.

Obsah

ABSTRAKT	**3**
1. ÚVOD	**9**
2. ZKOUMANÉ POPISNÉ CHARAKTERISTIKY	**11**
2.1 Klasické a nově navržené popisné charakteristiky	**11**
2.1.1 Charakteristiky založené na L-momentech	12
2.1.2 Charakteristiky založené na LQ-momentech	15
2.1.3 Charakteristiky založené na TL-momentech	16
2.2 Charakteristiky variability	**18**
2.2.1 Směrodatná odchylka	18
2.2.2 Kvantilové míry	19
2.2.3 Charakteristika založená na L-momentech	20
2.2.4 Charakteristika založená na LQ-momentech	22
2.2.5 Charakteristika založená na TL-momentech	22
2.2.6 Průměrné absolutní odchylky	22
2.2.7 Charakteristika založená na mediánu funkce pořádkových statistik	24
2.3 Charakteristiky relativní variability	**26**
2.3.1 Variační koeficient	26
2.3.2 Standardizovaný variační koeficient	27
2.3.3 Poměrná kvartilová odchylka	28
2.3.4 Charakteristiky založené na L-momentech, LQ-momentech a TL-momentech	28
2.3.5 Poměrné absolutní odchylky	29
2.3.6 Charakteristika založená na mediánu funkce pořádkových statistik	31
2.4 Charakteristiky šikmosti	**31**
2.4.1 Momentová míra	31
2.4.2 Standardizovaný koeficient šikmosti	34
2.4.3 Kvantilová míra	35
2.4.4 Míry založené na L-momentech, LQ-momentech a TL-momentech	35
2.4.5 Charakteristika založená na mediánu funkce pořádkových statistik	36
2.4.6 Charakteristika založená na počtech hodnot	37
2.5 Charakteristiky kurtozy (špičatosti)	**39**
2.5.1 Momentová míra	40
2.5.2 Kvantilové míry	45
2.5.3 Míra založená na L–momentech	46
2.5.4 Míry založené na LQ-momentech a TL-momentech	46
3. VÝBĚROVÉ VLASTNOSTI A KRITÉRIA KVALITY ESTIMÁTORŮ	**48**

3.1 Absolutní charakteristiky polohy, vychýlení a variability 48

3.2 Relativní charakteristiky vychýlení a variability 49

3.3 Tvar výběrového rozdělení 52

3.4 Měří všechny charakteristiky stejnou vlastnost rozdělení? 53

3.5 Posouzení síly testu 53
3.5.1 Testy o normalitě rozdělení 53
3.5.2 Zobecnění testu o normalitě 54

4. POUŽITÁ DATA A PRAVDĚPODOBNOSTNÍ ROZDĚLENÍ 55

4.1 Teoretická pravděpodobnostní rozdělení 55

4.2 Konečné datové soubory 60

5. REALIZACE EXPERIMENTU 62

5.1 Popis realizace 62

5.2 Technika generování náhodných výběrů 63
5.2.1 Výběry z teoretických rozdělení 63
5.2.2 Výběry z konečných souborů 66

5.3 Metody určení kvantilů 67

6. VÝSLEDKY 68

6.1 Estimátory měr variability 68
6.1.1 Výběry z normálního rozdělení 68
6.1.2 Výběry z ostatních rozdělení 78

6.2 Estimátory měr relativní variability 87
6.2.1 Výběry z normálního rozdělení 87
6.2.2 Výběry z ostatních rozdělení 93

6.3 Estimátory měr šikmosti 99
6.3.1 Výběry z normálního rozdělení 100
6.3.2 Výběry z ostatních rozdělení 107

6.4 Estimátory měr kurtozy 122
6.4.1 Výběry z normálního rozdělení 122
6.4.2 Výběry z ostatních rozdělení 130

7. ZÁVĚRY 142

SEZNAM SYMBOLŮ 145

LITERATURA **147**

PŘÍLOHY **150**

1. Úvod

Pravděpodobnostní rozdělení nebo datový soubor lze charakterizovat z hlediska různých vlastností. Nejčastěji používanými a sledovanými vlastnostmi jsou poloha (úroveň), variabilita (případně relativní variabilita) a asymetrie (šikmost) rozdělení. Méně sledovanou vlastností rozdělení je jeho kurtoza (špičatost, někdy též označovaná jako exces). Pro každou z těchto vlastností existuje řada měr (charakteristik), pomocí nichž je lze kvantifikovat. Pokud dané pravděpodobnostní rozdělení neznáme, je možno jeho charakteristiky odhadnout z výběru pomocí jejich estimátorů. Každý estimátor má potom své pravděpodobnostní rozdělení (nazýváme jej výběrovým rozdělením estimátoru nebo zkráceně výběrovým rozdělením), které lze opět popsat pomocí těchto charakteristik.

Dlouhou dobu byly výše uvedené vlastnosti popisovány pouze pomocí momentových měr. Některé nedostatky těchto měr, které se projevují zejména u vyšších momentů, především vysoká variabilita jejich estimátorů a citlivost vůči odlehlým pozorováním, vedly k navržení alternativních měr, založených například na kvantilech nebo na L-momentech. Výběrové vlastnosti estimátorů některých měr lze za určitých předpokladů (např. pro výběry z normálního rozdělení) odvodit analytickou cestou. Pro většinu měr však výběrové vlastnosti jejich estimátorů analytickou cestou odvodit nelze a mnohdy nelze analyticky odvodit ani jejich limitní vlastnosti (tj. pro rozsah výběru $n \to \infty$). Pro jejich zkoumání a porovnání je tedy nutno použít metodu Monte Carlo.

Pro některé estimátory již byly pomocí metody Monte Carlo provedeny dílčí studie s cílem popsat jejich vlastnosti při výběrech z určitých rozdělení, zejména z rozdělení normálního. Tyto studie se zaměřovaly převážně na variabilitu a vychýlení, v některých případech také na tvar výběrového rozdělení. Souhrn dosavadních poznatků o výběrových vlastnostech jednotlivých estimátorů získaných analytickou i experimentální cestou uvádím v kapitole 2. Dosud však nebyla publikována ucelená studie, která by poskytla komplexní porovnání jednotlivých estimátorů pro výběry z různých typů populací.

Cílem této práce je pomocí simulace poskytnout doposud chybějící vzájemné porovnání výběrových vlastností estimátorů klasických a nově navržených měr

variability, relativní variability, šikmosti a kurtozy. Porovnání je provedeno především z hlediska vychýlení, variability a tvaru výběrového rozdělení, a to v závislosti na typu populace, z níž výběr pochází, a na rozsahu výběru. Zvláštní pozornost je věnována výběrům z normálního rozdělení.

Protože všechny nově zkonstruované popisné charakteristiky byly navrženy jako alternativa k charakteristikám momentovým, je žádoucí, aby mezi jejich estimátory a estimátory měr momentových existovala co nejsilnější přímá závislost. Jako další kritérium kvality jednotlivých charakteristik jsem tedy zvolil sílu této závislosti.

Velmi častou aplikací momentových měr šikmosti a kurtozy (špičatosti) je použití jejich estimátorů jako kritéria normality rozdělení. I pro tuto aplikaci se tedy naskýtá možnost využití nově navržených charakteristik. V práci proto dále provádím porovnání měr založených na momentech, kvantilech a L-momentech z hlediska vhodnosti použití jejich estimátorů pro testy hypotéz o tvaru rozdělení, z něhož výběr pochází. Porovnání je provedeno na základě odhadnuté síly testů založených na příslušné míře, která je vypočtena z experimentálně získaných distribučních funkcí výběrových rozdělení.

Pro některé nově navržené popisné charakteristiky jsou v této práci pořízeny též odhady význačných kvantilů a dalších popisných charakteristik výběrových rozdělení, které dosud nejsou k dispozici a které je možno tímto způsobem (alespoň pro výběry z normálního rozdělení) tabelovat. Zároveň je provedena verifikace některých již dříve publikovaných teoretických i experimentálně získaných výsledků týkajících se výběrových vlastností zkoumaných estimátorů.

2. Zkoumané popisné charakteristiky

2.1 Klasické a nově navržené popisné charakteristiky

Jak již bylo zmíněno v úvodu, v současné době nejčastěji používané jednorozměrné popisné charakteristiky jsou založeny na **momentech**. Obecně rozlišujeme několik typů momentů (klasifikace je převzata z ČERMÁK, 1993):

- *počátkové*[1] *momenty*, definované jako: $\mu'_r = E(X^r)$,

  kde $E(X^r)$ je střední hodnota náhodné veličiny X^r;

- *obecné momenty* okolo konstanty *a*: $\mu_r(a) = E(X-a)^r$;

- *centrální momenty*: $\mu_r = E(X-\mu'_1)^r = E[X-E(X)]^r$,

  což je speciální případ obecných momentů;

- *normované momenty*: $\mu_r(U) = \dfrac{E[X-E(X)]^r}{D(X)^{r/2}}$,

  kde $D(X)$ je rozptyl náhodné veličiny X;

- *počátkové absolutní momenty*: $\nu'_r = E(|X^r|)$;

- *obecné absolutní momenty*: $\nu_r(a) = E(|X-a|^r)$;

- *centrální absolutní momenty*: $\nu_r(a) = E(|X-\mu'_1|^r) = E(|X-E(X)|^r)$.

Podrobněji se lze o problematice momentů, způsobu jejich výpočtu a vztahů mezi nimi dočíst např. v ČERMÁK (1993), kde jsou také uvedeny obecné vztahy pro jejich výpočet u vybraných teoretických pravděpodobnostních rozdělení.

Nejstarší alternativou k momentovým mírám jsou míry založené na **kvantilech**. $100{\cdot}p\%$ kvantil náhodné veličiny X, který značíme X_p, je hodnota splňující současně nerovnosti:

$$P(X \le X_p) \ge p \quad \text{a} \quad P(X \ge X_p) \le (1-p),$$

[1] Někdy též označované jako obecné momenty, např. HEBÁK, KAHOUNOVÁ (1988).

kde P(A) značí pravděpodobnost náhodného jevu A.

Podrobněji se lze o způsobu určení kvantilů dočíst opět např. v ČERMÁK (1993), kde jsou uvedeny též kvantilové funkce Q(*u*) vybraných teoretických pravděpodobnostních rozdělení, pomocí nichž lze kvantily těchto rozdělení vypočítat. Způsob určení kvantilů z konečných souborů, který jsem použil ve své práci, je popsán v kapitole 5.

Další skupinou jsou charakteristiky založené na **L-momentech** a jejich analogiích **LQ-momentech** a **TL-momentech**, které popisuji v následujících podkapitolách, a charakteristiky založené **na mediánu funkce lineární kombinace pořádkových statistik**, jejichž princip výpočtu popisuji u definice charakteristik S_{HL} a *medcouple*.

2.1.1 Charakteristiky založené na L-momentech

L-momenty představují relativně nově navrženou sadu charakteristik založených na pořádkových statistikách. Popisují stejné vlastnosti rozdělení, jako klasické momenty, mají však oproti nim zejména tyto výhody:

- jejich výběrové protějšky (tzv. *l*-statistiky) se vyznačují jednak nižší variabilitou, jednak vyšší robustností vůči odlehlým pozorováním;
- jsou aplikovatelné na početnější třídu pravděpodobnostních rozdělení, neboť platí, že všechny L-momenty existují tehdy a jen tehdy, pokud má náhodná veličina X konečnou střední hodnotu E(X).

Naopak určitou nevýhodou L-momentů je nemožnost jejich určení analytickou cestou u mnohých typů rozdělení (mimo jiné u všech rozdělení, pro která nelze explicitně vyjádřit kvantilovou funkci). Drobným nedostatkem je také vyšší výpočetní náročnost jejich estimátorů.

Využití střední hodnoty lineárních kombinací pořádkových statistik pro charakterizování pravděpodobnostního rozdělení bylo navrženo v práci SILLITTO (1951). Teprve HOSKING (1991) zavedl pojem L-moment a navrhl použití podílů L-momentů jako alternativu k momentovým mírám relativní variability, šikmosti a špičatosti.

Obecně je r-tý L-moment definován vztahem:

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} \mathrm{E}(X_{r-k:r}),$$

kde $X_{i:n}$ je i-tá pořádková statistika při výběru o rozsahu n.

Pro první čtyři L-momenty potom po úpravách dostáváme:

$$\lambda_1 = \mathrm{E}(X) = \int_0^1 Q(u)du,$$

$$\lambda_2 = \frac{1}{2}\mathrm{E}(X_{2:2} - X_{1:2}) = \int_0^1 Q(u)(2u-1)du,$$

$$\lambda_3 = \frac{1}{3}\mathrm{E}(X_{3:3} - 2X_{2:3} + X_{1:3}) = \int_0^1 Q(u)(6u^2 - 6u + 1)du,$$

$$\lambda_4 = \frac{1}{4}\mathrm{E}(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) = \int_0^1 Q(u)(20u^3 - 30u^2 + 12u - 1)du,$$

kde $Q(u)$ je kvantilová funkce náhodné veličiny X.

Analogicky ke konstrukci měr založených na momentech jsou v práci HOSKING (1991) navrženy následující bezrozměrné charakteristiky (u každé uvádím též obor hodnot):

$V_L = \frac{L_2}{L_1} \in \langle 0;1 \rangle$ jako obdoba variačního koeficientu,

$\tau_3 \equiv \alpha_L = \frac{L_3}{L_2} \in \langle -1;1 \rangle$ jako obdoba momentového koeficientu šikmosti, tzv. *L-skewness* a

$\tau_4 \equiv \beta_L = \frac{L_4}{L_2} \in \langle -1;1 \rangle$ jako obdoba momentového koeficientu špičatosti, tzv. *L-kurtosis*.

L-momenty a charakteristiky na nich založené našly doposud širší uplatnění především v hydrologii a v klimatologii. V práci GUTTMAN (1994) je diskutována možnost použití estimátorů L-momentů pro odhad parametrů teoretických rozdělení používaných pro modelování průtoku řek a četností záplav. V práci

SANKARASUBRAMANIAN et al. (1999) byly pomocí simulace získány hodnoty směrodatných chyb výběrových rozdělení estimátorů měr založených na L-momentech pro vybraná teoretická rozdělení používaná v hydrologii a bylo provedeno porovnání jejich výběrových vlastností s estimátory měr momentových. V práci VOGEL et. al. (1993) je posouzena možnost nahrazení tzv. product moment diagramu - tj. dvojrozměrného grafu kombinací hodnot koeficientů šikmosti a špičatosti, na základě něhož je možné určit typ teoretického pravděpodobnostního rozdělení, viz např. JOHNSON et al. (1994) - jeho alternativou založenou na L-momentových mírách. Ve všech případech bylo chování měr založených na L-momentech shledáno jako uspokojivé.

ROYSTON (1992) vyzdvihuje lepší interpretovatelnost hodnot koeficientů šikmosti a špičatosti založených na L-momentech oproti mírám založeným na klasických momentech. Dále diskutuje možnost odvození transformace odstraňující šikmost ve smyslu L-momentové míry jako alternativu k transformaci Boxe a Coxe odstraňující šikmost momentovou. V práci ELAMIR, SEHEULT (2004) je dále odvozen vztah pro výčet kovarianční matice mezi estimátory L-momentů řádu $r \leq 4$.

2.1.1.1 Estimátory L-momentů

Pro odhad L-momentů byly navrženy dva druhy estimátorů. Jednak estimátory označované jako nestranné, jednak tzv. plotting-position estimátory. Práce HOSKING, WALLIS (1995) přináší srovnání obou typů estimátorů na základě výsledků Monte Carlo simulací a jednoznačně preferuje nestranné estimátory. Chováním plotting-position estimátorů jsem se proto dále nezabýval.

Estimátory L-momentů z výběru o rozsahu n, tzv. l-statistiky, získáme tak, že sestavíme všech $\binom{n}{r}$ r-prvkových kombinací a vypočteme průměr příslušné lineární kombinace pořádkových statistik získaných z těchto kombinací. Například pro $r = 4$ tak dostáváme:

$$l_4 = \frac{1}{4}\ \text{est}\,\mathrm{E}(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) = \frac{1}{4}\binom{n}{4}^{-1} \sum_{i}\sum_{i<j}\sum_{j<l}\sum_{l<m}\left(x_{i:n} - 3x_{j:n} + 3x_{l:n} - x_{m:n}\right)$$

Poznámka: Pro výpočet efektivnější je alternativní způsob vyjádření výběrové

l-statistiky: dle HOSKING, WALLIS (1995) lze l_r vyjádřit jako lineární kombinaci tzv. beta l-statistik b_k, kde

$$b_k = \frac{1}{n}\sum_{j=r+1}^{n}\frac{(j-1)(j-2)...(j-k)}{(n-1)(n-2)...(n-k)}x_{j:n}\,.$$

Například druhý až čtvrtý výběrový L-moment lze potom vypočítat podle vztahů:

$$l_2 = 2b_1 - b_0,$$

$$l_3 = 6b_2 - 6b_1 + b_0,$$

$$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0.$$

Každou r-tou l-statistiku tak můžeme jednoduše vyjádřit jako lineární kombinaci všech prvků setříděného datového souboru bez nutnosti vytvářet všechny možné r-prvkové kombinace.

2.1.2 Charakteristiky založené na LQ-momentech

S cílem dosáhnout ještě vyšší robustnosti než L-momenty byly navrženy dvě jejich analogie. Starší z nich, LQ-momenty, byla definována v práci MUDHOLKAR, HUSTON (1998). Hlavní výhodou LQ-momentů oproti L-momentům je skutečnost, že existují u všech pravděpodobnostních rozdělení.

Obecně je r-tý LQ-moment definován:

$$\xi_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k\binom{r-1}{k}\tau_{p,\alpha}(X_{r-k:r}),$$

kde $0 \le \alpha \le 1/2$, $0 \le p \le 1$, $r \in \{0, 1, 2,...\}$

a $\tau_{p,\alpha}(Y)$ je tzv. robustní charakteristika polohy veličiny Y definovaná:

$$\tau_{p,\alpha}(Y) = pQ_Y(\alpha) + (1-2p)Q_Y(1/2) + pQ_Y(1-\alpha)\,,$$

kde $Q_Y(\alpha)$ je hodnota kvantilové funkce veličiny Y v bodě α.

Charakteristika $\tau_{p,\alpha}(Y)$ je tedy vážený průměr mediánu, $100\cdot\alpha\%$ a $100\cdot(1 - \alpha)\%$ kvantilu. Pro vybraná α a p dostáváme robustní charakteristiky polohy známé pod názvy *medián*, *trimean* a *Gastwirth* (viz např. MUDHOLKAR, HUSTON 1998):

Medián: $\tau_{0,0} = Q_Y(0{,}50)$, tj. p = 0, α je libovolné;

Trimean: $\tau_{0{,}25;\,0{,}25} = 0{,}25Q_Y(0{,}25) + 0{,}5Q_Y(0{,}50) + 0{,}25Q_Y(0{,}75)$, tj. $p = \alpha = 0{,}25$;

Gastwirth: $\tau_{0{,}3;\,1/3} = 0{,}3Q_Y(1/3) + 0{,}4Q_Y(0{,}50) + 0{,}3Q_Y(2/3)$, tj. $p = 0{,}3$; $\alpha = 1/3$.

Jedinou odlišností konstrukce LQ-momentů od L-momentů je tedy nahrazení středních hodnot pořádkových statistik $E(X_{r-k\,:\,r})$ robustními charakteristikami polohy $\tau_{p,\alpha}(X_{r-k\,:\,r})$ těchto pořádkových statistik.

Analogicky ke konstrukci měr založených L-momentech jsou v práci MUDHOLKAR, HUSTON (1998) navrženy bezrozměrné charakteristiky šikmosti a kurtozy:

$\eta_3 = \frac{\xi_3}{\xi_2}$, obdoba L-momentového koeficientu šikmosti, tzv. *LQ-skewness*, a

$\eta_4 = \frac{\xi_4}{\xi_2}$, obdoba L-momentového koeficientu špičatosti, tzv. *LQ-kurtosis*.

Nevýhodou těchto charakteristik je skutečnost, že dosud nejsou prozkoumány z hlediska oboru hodnot. (U všech zkoumaných populací zahrnutých do experimentu se však jejich hodnoty nacházejí v intervalu <-1; 1>).

Estimátory LQ-momentů, tzv. výběrové LQ-momenty, jsou definovány analogicky jejich populačním protějškům s tím rozdílem, že kvantilové funkce pořádkových statistik jsou nahrazeny jejich odhady. Blíže viz MUDHOLKAR, HUSTON (1998).

2.1.3 Charakteristiky založené na TL-momentech

Novější analogií L-momentů jsou tzv. trimmed[2] L-momenty, zkráceně označované jako TL-momenty, jejichž konstrukce byla navržena v ELAMIR, SEHEULT (2004). TL-momenty existují u některých rozdělení, u nichž nejsou L-momenty definovány, nikoliv však u všech (narozdíl od LQ-momentů, které existují vždy). Obecně TL-momenty vycházejí z postupu aplikovaného např. při konstrukci

[2] Podle doslovného překladu ustřihnuté nebo zkrácené

robustní míry polohy *trimmed mean*, který spočívá v odstranění určitého počtu nejvyšších a nejnižších hodnot z datového souboru.

Definice *r*-tého TL-momentu je:

$$\lambda_r^{(t)} = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} \mathrm{E}(X_{r+t-k\,:\,r+2t}),$$

kde $X_{i\,:\,r+2t}$ je *i*-tá pořádková statistika při výběru o rozsahu $r + 2t$

a *t* je „délka oříznutí“ neboli počet nejvyšších a nejnižších prvků, které v každé kombinaci délky $r + 2t$ ignorujeme.

Odlišnost od L-momentů tedy spočívá v tom, že pro stanovení *r*-tého TL-momentu je třeba znát střední hodnoty prostředních *r* pořádkových statistik při výběrech o rozsahu $r + 2t$. Například pro první čtyři TL-momenty pro $t = 1$ dostáváme:

$$\lambda_1^{(1)} = \mathrm{E}(X_{2:3}) = 6\int_0^1 Q(u)u(1-u)du,$$

$$\lambda_2^{(1)} = \frac{1}{2}\mathrm{E}(X_{3:4} - X_{2:4}) = 6\int_0^1 Q(u)u(1-u)(2u-1)du,$$

$$\lambda_3^{(1)} = \frac{1}{3}\mathrm{E}(X_{4:5} - 2X_{3:5} + X_{2:5}) = \frac{20}{3}\int_0^1 Q(u)u(1-u)(5u^2 - 5u + 1)du,$$

$$\lambda_4^{(1)} = \frac{1}{4}\mathrm{E}(X_{5:6} - 3X_{4:6} + 3X_{3:6} - X_{2:6}) = \frac{15}{2}\int_0^1 Q(u)u(1-u)(14u^3 - 21u^2 + 9u - 1)du,$$

kde $Q(u)$ je kvantilová funkce náhodné veličiny *X*.

Analogicky ke konstrukci měr založených L-momentech jsou definovány charakteristiky šikmosti a kurtozy, které, stejně jako v případě LQ-momentů, zůstávají zatím neprozkoumány z hlediska přesných mezí oboru hodnot:

$\lambda_3^{(t)} = \dfrac{l_3^{(t)}}{l_2^{(t)}}$, obdoba L-momentového koeficientu šikmosti, tzv. *TL-skewness*, a

$\lambda_4^{(t)} = \dfrac{l_4^{(t)}}{l_2^{(t)}}$, obdoba L-momentového koeficientu špičatosti, *TL-kurtosis*.

Výběrové TL-momenty $l_r^{(t)}$ jsou konstruovány analogicky svým populačním protějškům. Estimátory TL-momentů lze, obdobně jako v případě L-momentů, vyjádřit jako lineární kombinaci všech prvků setříděného datového souboru a není třeba vytvářet všechny (r + 2t)-prvkové kombinace. Blíže viz ELAMIR, SEHEULT, 2004.

2.2 Charakteristiky variability

2.2.1 Směrodatná odchylka

Vymezení charakteristiky

Nejstarší a nejpoužívanější charakteristiku variability je zřejmě směrodatná odchylka, definovaná jako:

$$S_X = \sqrt{\mu_2}\ ,$$

kde μ_k je k – tý centrální moment.

Estimátor směrodatné odchylky je definován jako:

$$s_X = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}\ .$$

Dosavadní poznatky o výběrových vlastnostech estimátoru

V práci ČERMÁK, VRABEC (1999) je odvozen přibližný vztah pro směrodatnou chybu estimátoru s_X pro výběry s vracením z konečného základního souboru:

$$SE(s_X) \doteq \frac{1}{2} S_X \sqrt{\frac{1-f}{n}(\beta_M - 1)}\ ,$$

kde β_M je čtvrtý normovaný moment (tj. momentový koeficient špičatosti) a f je podíl vybíraných prvků ze základního souboru při výběru bez vracení. V případě výběru s vracením je $f = 0$ a po dosazení do předchozího vztahu tak dostáváme vztah pro výběr s vracením odvozený např. v JOHNSON et al. (1994).

Dále je v práci ČERMÁK, VRABEC (1999) odvozen přibližný vztah pro střední hodnotu estimátoru s_X pro výběry s vracením z konečného základního souboru:

$$E(s_X) \doteq S_X\left[1-\frac{1-f}{8n}(\beta_M-1)\right],$$

Z toho vztahu vyplývá, že se jedná o estimátor vychýlený, který vždy podhodnocuje, přičemž absolutní hodnota jeho vychýlení roste s mírou kurtozy v základním souboru:

$$B(s_X) \doteq -S_X\left[\frac{1-f}{8n}(\beta_M-1)\right].$$

Dále je např. v práci KONDO (1930) uveden vztah pro koeficient šikmosti a špičatosti výběrového rozdělení estimátoru s_X při výběrech z normálního rozdělení:

$$\alpha(s_X) = \frac{1}{2n},$$

$$\beta(s_X) = 3\left(1+\frac{1}{2n}\right).$$

Kromě toho je v KONDO (1930) uvedeno odvození směrodatné chyby, koeficientu šikmosti a špičatosti pro vybraná teoretická pravděpodobnostní rozdělení.

2.2.2 Kvantilové míry

Nejčastěji používané míry variability založené na kvantilech jsou kvartilová odchylka a dále ukazatel založený na tzv. kvazirozpětí. **Kvartilová odchylka**, která v literatuře častěji bývá označována δ, je definována jako

$$\delta \equiv S_Q = \frac{1}{2}(X_{0,75} - X_{0,25}),$$

kde X_P je p-procentní kvantil náhodné veličiny X. V případě normálního rozdělení nabývá hodnoty $0{,}6745\sigma$.

Charakteristika variability založená na **kvazirozpětí**, která bývá v literatuře označována jako ε, je definována

$$\varepsilon \equiv S_{QE} = \frac{1}{3{,}25}(X_{0{,}95} - X_{0{,}05}).$$

Normální rozdělení má hodnotu $S_{QE} = 0{,}4151\sigma$. Estimátory těchto měr s_Q a s_{QE} jsou konstruovány analogicky jejich populačním protějškům (v předchozích vzorcích se tedy kvantily veličiny X nahradí jejich odhady).

Dosavadní poznatky o výběrových vlastnostech estimátorů

V KENDAL, STUART (1963), str. 239 je odvozen vztah pro rozptyl odhadu kvartilového rozpětí. Pro jeho určení je třeba znát hodnotu hustoty pravděpodobnosti rozdělení, z něhož výběry provádíme, v bodech $X_{0,25}$ a $X_{0,75}$:

$$D(s_Q) = \frac{1}{64n}\left(\frac{3}{f(X_{0{,}25})^2} + \frac{3}{f(X_{0{,}75})^2} - \frac{2}{f(X_{0{,}25})f(X_{0{,}75})}\right).$$

Pro normální rozdělení po dosazení dostáváme

$$D(s_Q) = \frac{\sigma^2}{1{,}6158\,n},$$

resp.

$$SE(s_Q) = \frac{\sigma}{1{,}2711\sqrt{n}}.$$

2.2.3 Charakteristika založená na L-momentech

Vymezení charakteristiky

Pro kvantifikaci absolutní variability rozdělení lze použít již dříve zmíněný druhý L-moment, který je roven polovině tzv. Giniho střední diference G:

$$\lambda_2 \equiv S_L = \frac{1}{2}G = \frac{1}{2}\mathrm{E}(X_{2:2} - X_{1:2}),$$

jehož estimátorem je druhý výběrový L-moment $l_2 \equiv s_L$. V případě normálního rozdělení nabývá hodnoty $S_L = 0{,}5642\sigma$.

Dosavadní poznatky o výběrových vlastnostech estimátoru

V práci NAIR (1936) nebo ZENGA et al. (2004) je uveden obecný vztah pro rozptyl estimátoru Giniho střední diference *G*, z něhož lze jednoduše odvodit vztah pro rozptyl estimátoru s_L:

$$D(s_L) = \frac{1}{n(1-1)}\left[S_X^2 + (n-2)\Im - \frac{(2n-3)}{2}(2S_L)^2\right],$$

kde funkcionál $\Im$ lze vyjádřit pro spojitá pravděpodobnostní rozdělení například jako

$$\Im = \iiint |x-y||x-z| f(x) f(y) f(z)\, dx\, dy\, dz$$

nebo pro výběry s vracením z konečného datového souboru o rozsahu *N* jako

$$\Im = \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}|X_i - X_j||X_i - X_k|.$$

Pro případ výběrů z normálního rozdělení lze ze vztahu pro rozptyl estimátoru Giniho střední diference odvozeného např. v NAIR (1936) nebo ZENGA et al. (2004) získat vztah pro rozptyl estimátoru s_L:

$$D(s_L) = \frac{\sigma^2}{n(n-1)}\left[\frac{n+1}{3} + \frac{2\sqrt{3}(n-2) - 2(2n-3)}{\pi}\right].$$

V práci ZENGA et al. (2004) byl dále odvozen nevychýlený estimátor rozptylu estimátoru Giniho střední diference, který lze upravit na estimátor rozptylu estimátoru s_L:

$$est(D(s_L)) = \frac{1}{(n-2)(n-3)}\left[S_X^2 + (n-2)est(\Im) - \frac{(2n-3)}{2}(2s_L)^2\right],$$

kde

$$est(\Im) = \frac{1}{n(n-1)(n-2)}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}|x_i - x_j||x_i - x_k| - \frac{2s_X^2}{n-2}.$$

Z obecného vztahu pro výčet kovarianční matice mezi estimátory L-momentů řádu $r \leq 4$ uvedeného v práci ELAMIR, SEHEULT (2004) lze rozptyl estimátoru alternativně vyjádřit pomocí lineární kombinace středních hodnot funkcí pořádkových

statistik. V práci ZENGA et al. (2004) je dále navržen vzorec pro kriterium testu o hodnotě Giniho střední diference v populaci.

2.2.4 Charakteristika založená na LQ-momentech

Charakteristikou variability založenou na LQ momentech je výše zmíněný druhý LQ moment. Pro účely experimentu jsem bral v úvahu pouze variantu trimean, tj.

$$S_{LQ} \equiv \xi_2 = \frac{1}{2}\left[\tau_{0,25;\,0,25}(X_{2:2}) - \tau_{0,25;\,0,25}(X_{1:2})\right],$$

kde

$$\tau_{0,25;\,0,25}(X) = 0{,}25Q_X(0{,}25) + 0{,}5Q_X(0{,}5) + 0{,}25Q_X(0{,}75),$$

$Q_X(u)$ je kvantilová funkce náhodné veličiny *X*. Normálně rozdělená náhodná veličina má hodnotu $S_{LQ} = 0{,}5494\sigma$. Estimátorem této charakteristiky je druhý výběrový LQ-moment, který dále budu označovat s_{LQ}.

2.2.5 Charakteristika založená na TL-momentech

Charakteristikou variability založenou na TL-momentech je výše zmíněný druhý TL-moment $\lambda_2^{(t)}$, přičemž pro účely experimentu jsem bral v úvahu pouze variantu s délkou zkrácení $t = 1$:

$$\lambda_2^{(1)} \equiv S_{TL} = \frac{1}{2}E(X_{3:4} - X_{2:4}).$$

Normálně rozdělená náhodná veličina má hodnotu $S_{TL} = 0{,}2970\sigma$. Estimátor této míry, druhý výběrový TL-moment $l_2^{(1)}$, budu dále označovat s_{TL}. V práci ELAMIR et al. (2003) je uveden vztah pro výpočet kovarianční matice mezi výběrovým TL-momenty.

2.2.6 Průměrné absolutní odchylky

Do experimentu jsem zahrnul jednak průměrnou absolutní odchylku od průměru a dále průměrnou absolutní odchylku od mediánu. **Průměrná absolutní odchylka od průměru**, v literatuře označovaná např. jako *AAD* nebo ν_1, je definována jako:

$$S_{AAD} \equiv AAD = \int_E |x - \eta_1'| f(x)\, dx$$

nebo pro datový soubor o rozsahu *N* jako:

$$S_{AAD} \equiv AAD = \sum_{i=1}^{N} |x_i - \eta_1'| .$$

Průměrná absolutní odchylka od mediánu, v literatuře označovaná např. jako *AADM* nebo ν_1*, je analogicky definována jako:

$$S_{AADM} \equiv AADM = \int_E |x - Q(0{,}5)| f(x)\, dx$$

nebo pro datový soubor o rozsahu *N* jako:

$$S_{AADM} \equiv AADM = \sum_{i=1}^{N} |x_i - x_{0,5}| .$$

V případě normálního rozdělení nabývají obě charakteristiky hodnoty $S_{AAD} = S_{AADM} = 0{,}7979\sigma$.

Dosavadní poznatky o výběrových vlastnostech estimátorů

V KENDAL, STUART (1963), str. 240 je odvozen přibližný vztah pro rozptyl odhadu průměrné absolutní odchylky od konstanty a:

$$D(d_a) = \frac{1}{n}\left[\sigma^2 + (a - \mu)^2 - \delta_a^2\right],$$

kde μ resp. σ je střední hodnota resp. směrodatná odchylka v populaci a δ_a je skutečná hodnota průměrné absolutní odchylky v populaci. Pokud zanedbáme rozptyl odhadu konstanty (tj. výběrového průměru v případě estimátoru s_{AAD} resp. výběrového mediánu v případě s_{AADM}), získáme po dosazení přibližné vztahy pro rozptyl obou estimátorů:

$$D(s_{AAD}) \doteq \frac{1}{n}\left(\sigma^2 - S_{AAD}^2\right),$$

$$D(s_{AADM}) \doteq \frac{1}{n}\left[\sigma^2 + (X_{0,50} - \mu)^2 - S_{AAD}^2\right].$$

Pro výběry z normálního rozdělení byl dále odvozen přesný vztah pro rozptyl estimátoru s_{AAD}, opět viz KENDAL, STUART (1963). V GODWIN (1945) je dále odvozeno přesné výběrové rozdělení estimátoru s_{AAD} při výběrech z normálního

rozdělení, jehož význačné kvantily pro vybrané rozsahy výběru n jsou tabelovány v HARTLEY (1945).

2.2.7 Charakteristika založená na mediánu funkce pořádkových statistik

Z této kategorie charakteristik jsem do experimentu zahrnul charakteristiku zmíněnou např. v ROUSSEEUW, CROUX (1993), která nemá v literatuře standardně zavedené značení ani pojmenování:

$$S_{HL} = med\{|X_i - X_j|; i < j\}.$$

Tato charakteristika je analogií Hodges - Lehmannovy míry polohy

$$HL = med\left\{\frac{X_i + X_j}{2}; i < j\right\}.$$

Pro výpočet charakteristiky S_{HL} je tedy třeba vytvořit všechny možné dvouprvkové kombinace a z každé z nich vypočítat absolutní odchylku. Celkem tak vznikne $C_2(n)$ hodnot, z nichž se vypočte medián.

***Poznámka**: V DEVILLARD (1988) je uveden algoritmus pro výpočet mediánu bez nutnosti předchozího setřídění souboru, který jsem v tomto případě použil. Podrobněji viz kapitola 5.*

Normálně rozdělená náhodná veličina má hodnotu $S_{HL} = 0{,}9539\sigma$. ROUSSEEUW, CROUX (1993) proto navrhují dělení S_{HL} hodnotou 0,9539 z důvodu dosažení srovnatelnosti jeho hodnot s hodnotou směrodatné odchylky normálního rozdělení. Je však zřejmé, že u jiných rozdělení vztah $S_{HL} = 0{,}9539\sigma$ platit nemusí, proto navrženou korekci, stejně jako u ostatních estimátorů měr variability, ve této práci nepoužívám.

V práci ROUSSEEUW, CROUX (1993) jsou navrženy další robustní charakteristiky založené na hodnotě význačného kvantilu funkce pořádkových statistik. Primárním účelem těchto robustních charakteristik je však především poskytnutí počátečního odhadu pro různé iterativní algoritmy, spíše než popis variability datového souboru. Tyto míry jsem tedy do svého experimentu nezahrnoval.

Tabulka 2.1: Zkoumané estimátory měr variability a jejich teoretické střední hodnoty pro výběry z normovaného normálního rozdělení:

Míra variability	Výpočtový vzorec estimátoru	Střední hodnota pro N(0, 1)
Směrodatná odchylka	$s_X = \sqrt{\dfrac{\sum (x_i - \bar{x})^2}{n-1}}$	1
Kvartilová odchylka	$s_Q = \dfrac{1}{2}(x_{0,75} - x_{0,25})$	0,6745
Modifikované kvazirozpětí	$s_{QE} = \dfrac{1}{3{,}25}(x_{0,95} - x_{0,05})$	0,4151
Druhý L-moment	$s_L = \dfrac{1}{2}\, est\ E(X_{2:2} - X_{1:2})$	0,5642
Druhý LQ-moment (pro $p = \alpha = 0{,}25$)	$s_{LQ} = \dfrac{1}{2}\left[\tau_{0,25;\,0,25}(x_{2:2}) - \tau_{0,25;\,0,25}(x_{1:2})\right]$	0,5494
Druhý TL-moment (pro $t = 1$)	$s_{TL} = \dfrac{1}{2}\, est\ E\left(X_{3:4} - X_{2:4}\right)$	0,2970
Průměrná absolutní odchylka od průměru	$s_{AAD} = \sum_{i=1}^{n} \lvert x_i - \bar{x} \rvert$	0,7979
Průměrná absolutní odchylka od mediánu	$s_{AADM} = \sum_{i=1}^{n} \lvert x_i - x_{50} \rvert$	0,7979
Charakteristika založená na mediánu funkce pořádkových statistik	$s_{HL} = est\ med\left\{\lvert x_i - x_j \rvert;\ i < j\right\}$	0,9539

2.3 Charakteristiky relativní variability

2.3.1 Variační koeficient

Nejstarší a nejpoužívanější charakteristiku variability je zřejmě variační koeficient definovaný jako

$$V_X = \frac{\sqrt{\mu_2}}{\mu_1'},$$

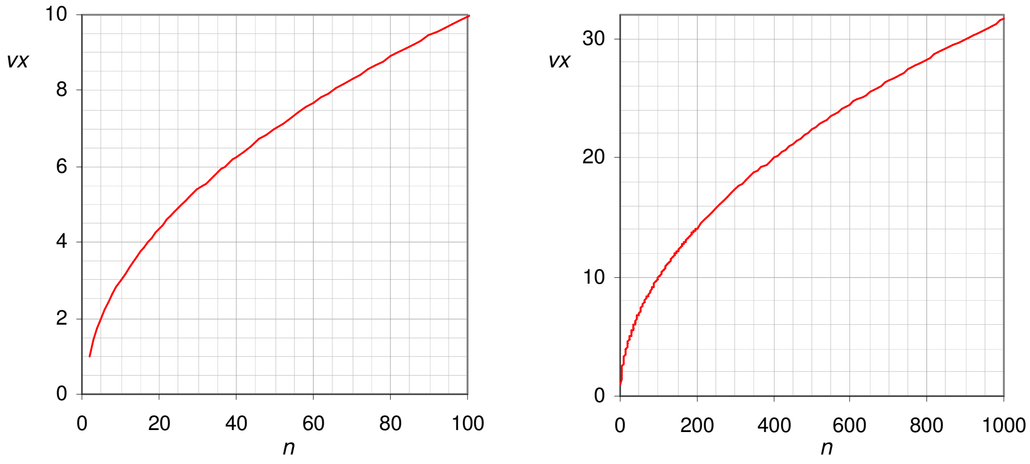
kde μ_k je k – tý centrální moment

a μ_k'je k – tý počátkový moment.

Jeho estimátor je definován analogicky jako

$$v_X = \frac{s_X}{\bar{x}}.$$

Obor hodnot estimátoru v_X je shora omezen, přičemž horní mez závisí na rozsahu výběru: Pro nezápornou náhodnou veličinu X může estimátor nabývat hodnot z intervalu $\left\langle 0; \sqrt{n-1} \right\rangle$, viz např. ČERMÁK (1985). Průběh horní meze oboru možných hodnot estimátoru v_X je znázorněn na obr. 2.1. Z existence této horní meze bez nutnosti dalšího zkoumání vyplývá, že pro populace s vysokou mírou relativní variability bude tento estimátor pro malé rozsahy výběru podhodnocovat.



Obr. 2.1: Horní mez oboru hodnot estimátoru v_X pro rozsahy výběru $n \leq 1000$ (vpravo pro všechny zkoumané rozsahy výběru, vlevo detail pro rozsahy $n \leq 100$).

Dosavadní poznatky o výběrových vlastnostech estimátoru

V práci ČERMÁK, VRABEC (1999) je odvozen přibližný vztah pro střední čtvercovou chybu estimátoru v_X pro výběry s vracením z konečného základního souboru, z něhož lze jednoduše odvodit vztah pro jeho směrodatnou chybu:

$$SE(v_X) \doteq \sqrt{\frac{1-f}{n} V_X^2 \left[\frac{1}{4}(\beta_M - 1) - \alpha_M V_X + V_X^2\right]},$$

kde a_M je třetí normovaný moment (tj. momentový koeficient šikmosti),
β_M je čtvrtý normovaný moment (tj. momentový koeficient špičatosti)
a f je podíl vybíraných prvků ze základního souboru při výběru bez vracení resp. $f = 0$ v případě výběru s vracením.

Pro výběry z normálního rozdělení potom dostáváme:

$$SE(v_X) \doteq \sqrt{\frac{1}{n} V_X^2 \left(\frac{1}{2} + V_X^2\right)}.$$

V McKAY (1931) je odvozen přibližný vztah pro kritérium testu o hodnotě variačního koeficientu, který lze použít přibližně pro $v_X \leq 0{,}33$. V práci IGLEWITZ et al. (1968) je navržen vztah pro interval spolehlivosti použitelný přibližně pro $v_X \leq 0{,}5$.

2.3.2 Standardizovaný variační koeficient

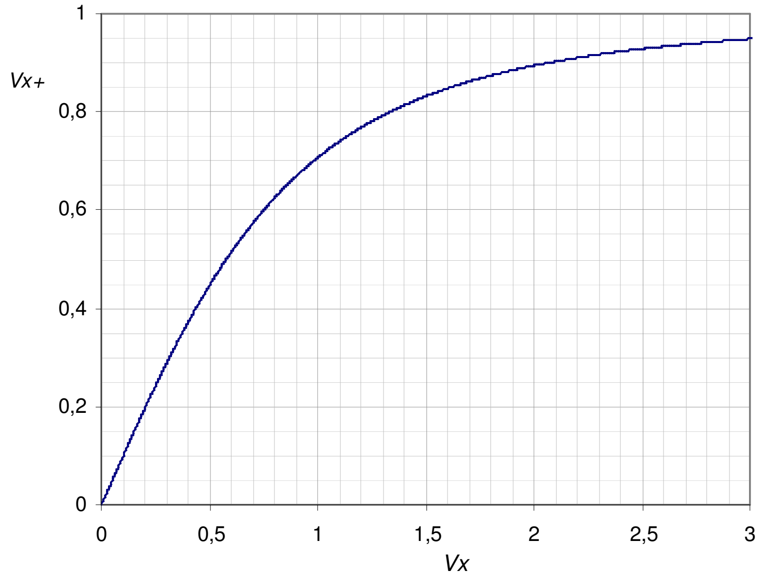
Nevýhody variačního koeficientu, spočívající zejména ve shora neomezeném oboru hodnot u teoretických rozděleních a dále v existenci výše zmíněné horní meze oboru hodnot závislé na rozsahu výběru n u konečných souborů, odstraňuje tzv. standardizovaný variační koeficient, navržený v ČERMÁK (1985):

$$V_X^+ = \sqrt{\frac{\mu_2}{\mu_2'}}$$

kde μ_2'je druhý počátkový moment. Modifikovaný variační koeficient nabývá hodnot z intervalu $\langle 0;1 \rangle$. Jednoduchou úpravou jej lze vyjádřit jako funkci variačního koeficientu V_X:

$$V_X^+ = \sqrt{\frac{V^2}{V^2+1}} \,.$$

Graf závislosti mezi hodnotami koeficientu V_X a V_X^+ je zobrazen na obr. 2.2. Je z něj zřejmé, že hodnoty obou koeficientů se výrazněji odlišují přibližně až pro hodnoty $V_X \geq 0{,}3$.



Obr. 2.2: Průběh závislosti hodnot koeficientů V_X (na horizontální ose) a V_X^+.

2.3.3 Poměrná kvartilová odchylka

Poměrná (někdy nazývaná jako relativní) kvartilová odchylka, v literatuře označovaná jako Q, je definována:

$$V_Q \equiv Q = \frac{X_{0,75} - X_{0,25}}{X_{0,75} + X_{0,25}} \,,$$

kde X_p je $100p$ procentní kvantil veličiny X.

2.3.4 Charakteristiky založené na L-momentech, LQ-momentech a TL-momentech

Charakteristika relativní variability **založená na L-momentech** je již zmíněný podíl druhého a prvního L-momentu, který dosud nemá ustálené pojmenování a

v literatuře bývá označován např. jako L-CV (zkratka z anglického názvu *L-Coefficient of Variation*):

$$V_L \equiv L\text{-}CV = \frac{L_2}{L_1},$$

kde L_r je r-tý L-moment.

Analogicky charakteristice V_L je konstruována charakteristika **založená na LQ-momentech**:

$$V_{LQ} = \frac{\xi_2}{\xi_1} = \frac{S_{LQ}}{\xi_1},$$

kde charakteristika polohy $\xi_1 = \tau_{0,25;\,0,25}(X)$ je první LQ-moment pro $p = \alpha = 0{,}25$.

Charakteristika relativní variability **založená na TL-momentech** je definována jako

$$V_{TL} = \frac{\lambda_2^{(1)}}{\lambda_1^{(1)}} = \frac{S_{TL}}{\lambda_1^{(1)}},$$

kde $\lambda_1^{(1)}$ je první TL-moment, $\lambda_1^{(1)} = E(X_{2:3})$.

2.3.5 Poměrné absolutní odchylky

Z průměrné absolutní odchylky od mediánu S_{AADM} je odvozena **poměrná absolutní odchylka od mediánu**, která bývá v literatuře označována jako R_1:

$$V_{AADM} \equiv R_1 = \frac{S_{AADM}}{\bar{X}}.$$

V práci Čermák (1985) je odvozeno, že obor hodnot této charakteristiky je <0; 1>.

Z průměrné absolutní odchylky od průměru S_{AAD} je odvozena **poměrná absolutní odchylka od průměru**, v literatuře označovaná jako R_2:

$$V_{AAD} \equiv R_2 = \frac{S_{AAD}}{\bar{X}}.$$

Obor hodnot této charakteristiky V_{AAD} je <0; 2>, přičemž pro konkrétní rozsah souboru n je to interval $\left\langle 0;\, 2\frac{n-1}{n} \right\rangle$.

Tabulka 2.2: Zkoumané estimátory měr relativní variability a jejich obory hodnot:

Míra variability	Výpočtový vzorec estimátoru	Obor hodnot
Variační koeficient	$v_X = \frac{s_X}{\bar{x}}$	$\left\langle 0;\, \sqrt{n-1} \right\rangle$; pro $n \to \infty$: $\left\langle 0;\, \infty \right\rangle$
Standardizovaný variační koeficient	$v_X^+ = \frac{s_X}{\sqrt{est\ E(X^2)}}$	<0; 1>
Poměrná kvartilová odchylka	$v_Q \equiv \frac{x_{0,75} - x_{0,25}}{x_{0,75} + x_{0,25}}$	<0; 1>
Charakteristika založená na L-momentech	$v_L = \frac{s_L}{\bar{x}}$	<0; 1>
Charakteristika založená na TL-momentech (pro $t = 1$)	$v_{TL} = \frac{l_2^{(1)}}{l_1^{(1)}} = \frac{s_{TL}}{l_1^{(1)}}$	<0; ?>
Charakteristika založená na LQ-momentech (pro $p = \alpha = 0,25$)	$v_{LQ} = \frac{est\ \xi_2}{est\ \xi_1} = \frac{s_{LQ}}{est\ \xi_1}$	<0; ?>
Poměrná absolutní odchylka od průměru	$v_{AAD} = \frac{s_{AAD}}{\bar{x}}$	$\left\langle 0;\, 2\frac{n-1}{n} \right\rangle$; pro $n \to \infty$: $\left\langle 0;\, 2 \right\rangle$
Poměrná absolutní odchylka od mediánu	$v_{AADM} = \frac{s_{AADM}}{\bar{x}}$	<0; 1>
Míra založená na mediánu funkce pořádkových statistik	$v_{HL} = \frac{s_{HL}}{\bar{x}}$	<0; ?>

2.3.6 Charakteristika založená na mediánu funkce pořádkových statistik

Na základě obecné definice pojmu relativní variabilita se nabízí možnost konstrukce míry relativní variability založené také na charakteristice S_{HL}. Definuji tak míru

$$V_{HL} = \frac{S_{HL}}{\overline{X}}.$$

Podotýkám, že se jedná o charakteristiku, která není běžně používaná a nebyla dosud prozkoumána ani z hlediska horní meze oboru možných hodnot.

2.4 Charakteristiky šikmosti

Pojmem šikmosti jako vlastnosti rozdělení se zabývá např. ČERMÁK (1975). Upozorňuje na častou chybu v učebnicích statistiky, ve kterých jsou pojmy šikmost a asymetrie používány jako synonyma. Ve skutečnosti je šikmost speciálním případem asymetrie. Dále poukazuje na často používanou chybnou implikaci, podle které z nulového koeficientu šikmosti (definovaného níže) vyplývá symetričnost pravděpodobnostního rozdělení, a pro ilustraci uvádí příklady asymetrických rozdělení, která mají nulovou hodnotu koeficientu šikmosti. Uvádí dále příklad datového souboru, ve kterém neplatí alternativně používaná definice pojmu šikmost, podle níž u kladně sešikmeného rozdělení má dolní polovina hodnot nižší variabilitu než horní polovina.

2.4.1 Momentová míra

Vymezení charakteristiky

Nejčastěji používanou mírou šikmosti je třetí normovaný moment, který bývá v literatuře označován většinou jako $\sqrt{\beta_1}$ případně α:

$$\alpha_M \equiv \sqrt{\beta_1} = \frac{\mu_3}{(\mu_2)^{3/2}},$$

kde μ_k je k-tý centrální moment.

Poznámka: *Označení* $\sqrt{\beta_1}$ *je použito podle charakteristiky* β_k, *viz např. KENDAL, STUART (1963) str. 85, která je definována zvlášť pro sudá a lichá k:*

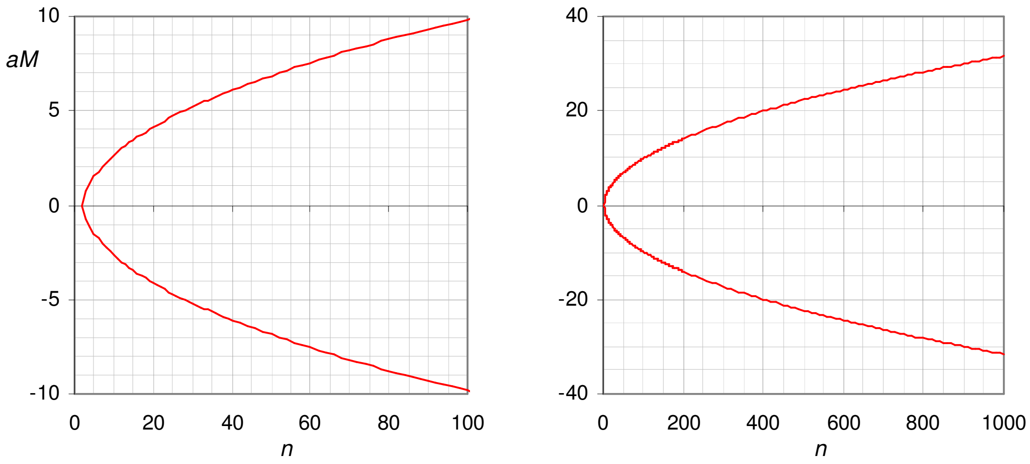
$$\beta_{2k+1} = \frac{\mu_3 \mu_{2k+3}}{\mu_2^{k+3}},$$

$$\beta_{2k} = \frac{\mu_{2k+2}}{\mu_2^{k+1}}.$$

Obor hodnot momentové míry není pro teoretické pravděpodobnostní rozdělení shora ani zdola omezen. Pro daný rozsah souboru n je však, jak bylo odvozeno např. ve WILKINS (1944) nebo ČERMÁK, HENZLER (1985), obor hodnot α_M omezen intervalem:

$$\left\langle -\frac{n-2}{\sqrt{n-1}}; +\frac{n-2}{\sqrt{n-1}} \right\rangle.$$

Jeho průběh v závislosti na rozsahu souboru n je znázorněn na obr. 2.3.



Obr. 2.3: Horní a dolní mez oboru hodnot míry α_M v závislosti na rozsahu souboru n (vpravo pro všechny zkoumané rozsahy výběru, vlevo detail pro rozsahy $n \leq 100$).

Estimátor momentové míry je definován analogicky ke konstrukci jeho populačního protějšku, populační momenty jsou tedy ve vztahu nahrazeny momenty výběrovými:

$$a_M = \frac{m_3}{m_2^{3/2}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{3/2}} ,$$

kde m_3 je třetí a m_2 druhý výběrový centrální moment a $\bar{x}$ je výběrový průměr.

Dosavadní poznatky o výběrových vlastnostech estimátoru

V učebnici KENDAL, STUART (1963) je odvozen limitní vztah pro rozptyl estimátoru a_M:

$$\lim_{n\to\infty} D(a_M) = \frac{1}{n}\left[\beta_1(4\beta_4 - 24\beta_2 + 36 + 9\beta_1\beta_2 - 12\beta_3 + 35\beta_1)\right] =$$

$$= \frac{1}{n}\left[\alpha^2\left(4\frac{\mu_6}{\mu_2^3} - 4\beta + 36 + 9\alpha^2\beta - 12\frac{\mu_3\mu_5}{\mu_2^4} + 35\alpha^2\right)\right],$$

kde μ_i je i-tý centrální moment, n rozsah výběru, α momentový koeficient šikmosti rozdělení a β momentový koeficient špičatosti rozdělení (definovaný v kapitole 2.5).

Pro výběry z normálního rozdělení po dosazení dostáváme

$$\lim_{n\to\infty} D(a_M) \doteq \frac{6}{n},$$

resp.

$$\lim_{n\to\infty} SE(a_M)\sqrt{n} \doteq 2{,}4495 .$$

V práci PEARSON (1930) byly odvozeny přibližné vztahy pro směrodatnou odchylku, koeficient šikmosti a špičatosti výběrového rozdělení estimátoru a_M pro výběry z normálního rozdělení:

$$\sigma(a_M) \doteq \sqrt{\frac{6}{n}}\left(1 - \frac{3}{n} + \frac{6}{n^2} - \frac{15}{n^3} + \ldots\right),$$

$$\alpha(a_M) = 0,$$

$$\beta(a_M) \doteq 3 + \frac{36}{n} - \frac{864}{n^2} + \frac{12{,}096}{n^3} - \dots .$$

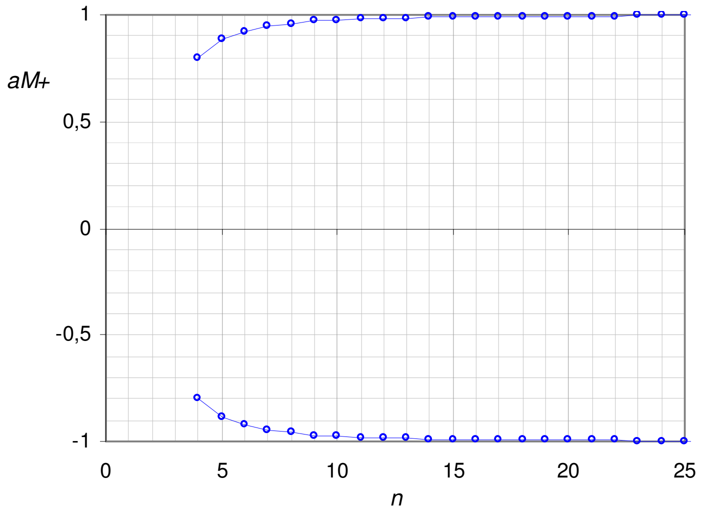
Pomocí simulace byly dále tabelovány hodnoty význačných kvantilů estimátoru α_M pro výběry z normálního rozdělení o rozsahu $20 \leq n \leq 100$, viz D'AGOSTINO, PEARSON (1973).

2.4.2 Standardizovaný koeficient šikmosti

Nedostatek momentové míry spočívající v teoreticky neomezeném oboru hodnot odstraňuje konstrukce, kterou navrhuje ČERMÁK (1985). V ní je výraz $\mu_2^{3/2}$ ve jmenovateli nahrazen třetím absolutním momentem:

$$\alpha_M^+ = \frac{E(X_i - \overline{X})^3}{E\left|X_i - \overline{X}\right|^3}.$$

Tento koeficient může u teoretických pravděpodobnostních rozdělení nabývat hodnot z intervalu $\langle -1; 1 \rangle$. Pro daný rozsah souboru je jeho obor hodnot omezen na interval $\left\langle -\frac{n^2 - 2n}{n^2 - 2n + 2}; +\frac{n^2 - 2n}{n^2 - 2n + 2} \right\rangle$, jak bylo odvozeno ve zmíněné práci. Průběh oboru hodnot v závislosti na rozsahu souboru n je znázorněn v grafu na obr. 2.4.



Obr. 2.4: Horní a dolní mez oboru hodnot estimátoru α_M^+ v závislosti na rozsahu souboru n.

2.4.3 Kvantilová míra

Kvantilový koeficient šikmosti, známý též jako Bowleyův koeficient, který bývá v literatuře označován τ, je definován jako:

$$\alpha_Q \equiv \tau = \frac{X_{0,75} - 2X_{0,5} + X_{0,25}}{X_{0,75} - X_{0,25}}$$

a nabývá hodnot z intervalu <-1; 1>.

Výběrové vlastnosti estimátoru

Pro estimátor a_Q byl ve WAGEMAKERS et al., 1995, odvozen limitní vztah pro rozptyl v případě výběru z normálního rozdělení:

$$\lim_{n\to\infty} \mathrm{D}(a_Q) = \frac{1{,}839}{n},$$

resp.

$$\lim_{n\to\infty} \mathrm{SE}(a_Q)\sqrt{n} = 1{,}356\,.$$

2.4.4 Míry založené na L-momentech, LQ-momentech a TL-momentech

Koeficient šikmosti **založený na L-momentech**, v literatuře nazývaný L-skewness a označovaný jako τ_3, je definován:

$$\alpha_L \equiv \tau_3 = \frac{L_3}{L_2}$$

a nabývá hodnot z intervalu <-1; 1>. Výběrové vlastnosti jeho estimátoru byly zkoumány v práci ČERMÁK, VRABEC (1998) metodou Monte Carlo pro výběry z vybraných teoretických rozdělení o rozsahu $n \leq 31$. Při porovnání s estimátory ostatních charakteristik zkoumaných v uvedené práci vykazovaly nejlepší vlastnosti estimátory a_L a a_M^+.

Analogicky je definován koeficient **založený na LQ momentech**

$$\alpha_{LQ} = \frac{\xi_3}{\xi_2} = \frac{\xi_3}{S_{LQ}}$$

a koeficient **založený na TL-momentech**, nazývaný jako TL-skewness

$$\alpha_{TL} = \frac{\lambda_3^{(1)}}{\lambda_2^{(1)}} = \frac{\lambda_3^{(1)}}{S_{TL}}.$$

2.4.5 Charakteristika založená na mediánu funkce pořádkových statistik

BRYS et al. (2004) navrhli třídu charakteristik šikmosti založenou na mediánu funkce pořádkových statistik. Jedná se o tzv. *medcouple* a dále z něj odvozený *medtriple*, *repeated medcouple* a *repeated medcouple*. Do zkoumání v rámci simulačního experimentu jsem zahrnul charakteristiku *medcouple*, v literatuře označovanou jako *MC*, která je definována:

$$\alpha_{MC} \equiv MC = med\ h(X_i, X_j) \qquad \text{pro } X_i \le \tilde{X}_{0,50} \le X_j,$$

kde funkce *h* je definována:

$$h(X_i, X_j) = \frac{(X_j - \tilde{X}_{0,50}) - (\tilde{X}_{0,50} - X_i)}{(X_j - X_i)} \qquad \text{pro } X_i \ne X_j,$$

$$h(X_i, X_j) = +1 \qquad \text{pro } X_i = X_j,\ i > j,$$

$$h(X_i, X_j) = 0 \qquad \text{pro } X_i = X_j,\ i = j,$$

$$h(X_i, X_j) = -1 \qquad \text{pro } X_i = X_j,\ i < j.$$

Pro výpočet charakteristiky medcouple je tedy třeba soubor setřídit a rozdělit na dvě překrývající se množiny tak, aby jednu množinu tvořily hodnoty $X_i \le \tilde{X}_{0,50}$ a druhou hodnoty $\tilde{X}_{0,50} \le X_j$. Z těchto množin se potom sestaví všechny možné kombinace a pro každou kombinaci se vypočte hodnota funkce *h*. Z hodnot funkce *h* se nakonec určí medián. Obor hodnot charakteristiky α_{MC} tvoří interval <-1; +1>.

2.4.6 Charakteristika založená na počtech hodnot

V učebnici CYHELSKÝ (1975) je navržena charakteristika šikmosti, založená na rozdílu počtu nadprůměrných a podprůměrných hodnot, označovaná též ω. Pro konečný datový soubor o rozsahu n je definována jako

$$\alpha_C \equiv \omega = \frac{n_D - n_H}{n},$$

kde n_D je počet „dolních" hodnot, tj. nižších než průměr a n_H je počet horních hodnot, tj. vyšších než průměr. Koeficient α_C lze alternativně vyjádřit (viz např. ČERMÁK, 1993) vztahem

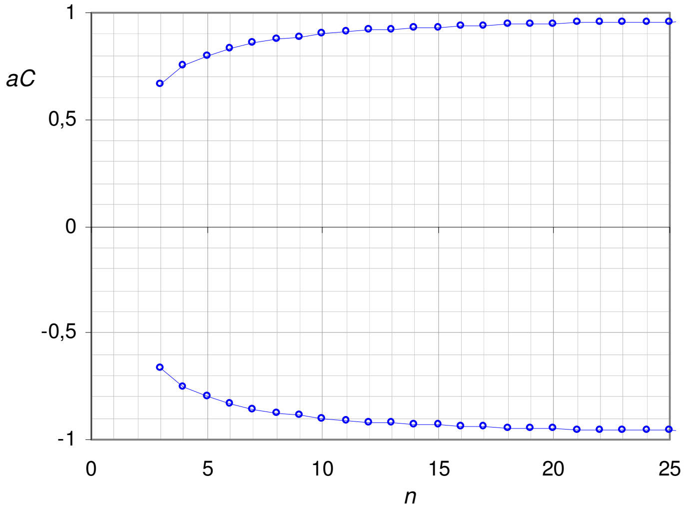
$$\alpha_C = \frac{(\overline{X}_H - \mu_1') - (\mu_1' - \overline{X}_D)}{\overline{X}_H - \overline{X}_D},$$

kde X_D je průměr dolních hodnot, X_H je průměr „horních" hodnot a μ_1' je celkový průměr.

Z alternativního vyjádření koeficientu α_C je zřejmá analogie s konstrukcí kvartilového koeficientu šikmosti: medián je nahrazen aritmetickým průměrem a kvartily jsou nahrazeny průměry „dolních" a „horních" hodnot. Jak ukázal ČERMÁK (1990), pro spojitou náhodnou veličinu lze koeficient α_C vyjádřit velmi jednoduše:

$$\alpha_C = 2F(\mu_1') - 1.$$

Pro teoretické pravděpodobnostní rozdělení může koeficient nabývat hodnot z intervalu (-1; 1), pro datový soubor o rozsahu n je to však interval $\left\langle -\frac{n-2}{n}; \frac{n-2}{n} \right\rangle$. Průběh horní a dolní meze oboru hodnot je znázorněn v grafu na obr. 2.5. Určitým nedostatkem tohoto koeficientu je nespojitost jeho oboru hodnot pro konečné datové soubory, který, jak vyplývá z prvního způsobu jeho vyjádření, je omezen na celočíselné násobky hodnoty $\frac{1}{n}$.



Obr. 2.5: Horní a dolní mez oboru hodnot estimátoru a_C v závislosti na rozsahu výběru n.

Tabulka 2.3: Zkoumané estimátory měr šikmosti a jejich obory hodnot:

Míra šikmosti	Výpočtový vzorec estimátoru	Obor hodnot
(Momentový) koeficient šikmosti	$a_M = \frac{m_3}{m_2^{3/2}}$	$\left\langle -\frac{n-2}{\sqrt{n-1}}; +\frac{n-2}{\sqrt{n-1}} \right\rangle$; pro $n \to \infty$: $(-\infty; \infty)$
Standardizovaný koeficient šikmosti	$a_M^+ = \frac{m_3}{est\ E\lvert X_i - \overline{X}\rvert^3}$	$\left\langle -\frac{n^2-2n}{n^2-2n+2}; +\frac{n^2-2n}{n^2-2n+2} \right\rangle$; pro $n \to \infty$: $\langle -1; 1 \rangle$
Bowleyův koeficient šikmosti	$a_Q = \frac{x_{0,75} - 2x_{0,5} + x_{0,25}}{x_{0,75} - x_{0,25}}$	<-1; 1>
Charakteristika založená na L-momentech	$a_L = \frac{l_3}{l_2}$	<-1; 1>
Charakteristika založená na TL-momentech (pro *t* = 1)	$a_{TL} = \frac{l_3^{(1)}}{l_2^{(1)}} = \frac{l_3^{(1)}}{s_{TL}}$	?
Charakteristika založená na LQ-momentech (pro $p = \alpha = 0,25$)	$a_{LQ} = \frac{est\ \xi_3}{est\ \xi_2} = \frac{est\ \xi_3}{s_{LQ}}$	?
Medcouple	$a_{MC} = med\ h(x_i, x_j)$	<-1; 1>
Koeficient založený na počtech hodnot	$\alpha_C = \frac{n_D - n_H}{n}$	$\left\langle -\frac{n-2}{n}; \frac{n-2}{n} \right\rangle$

2.5 Charakteristiky kurtozy (špičatosti)

Dlouhou dobu byla kurtoza jako vlastnost rozdělení charakterizována jedinou mírou, a to čtvrtým normovaným momentem. S touto mírou se lze setkat jak v popisné statistice, kde je uváděna jako jedna z elementárních charakteristik rozdělení, tak i v induktivní statistice: (i) na její velikosti závisí přesnost odhadu rozptylu a tedy i

směrodatných chyb (viz např. ČERMÁK, VRABEC 1999); (ii) její výběrová hodnota se používá v rámci explorativní analýzy dat jako jedno z kriterií normality rozdělení, a tedy i jako jedno z kritérií vhodnosti použití metody nejmenších čtverců (viz např. PEARSON 1930). Některé nedostatky čtvrtého normovaného momentu, především vysoká variabilita jeho estimátorů a citlivost vůči odlehlým pozorováním, později vedly k navržení alternativních měr kurtozy, založených jednak na kvantilových charakteristikách, jednak na L-momentech.

2.5.1 Momentová míra

Historicky nejstarší míra kurtozy, klasický koeficient špičatosti neboli čtvrtý normovaný moment proměnné X, je definován:

$$\beta_M = \mathrm{E}\left(\frac{X-\mu'_1}{\sigma}\right)^4 = \frac{\mathrm{E}(X-\mu'_1)^4}{\sigma^4} = \frac{\mu_4}{\mu_2^2},$$

kde μ'_1 je první obecný moment proměnné X (její střední hodnota), σ je směrodatná odchylka a $\mu_2 = \sigma^2$ je druhý a μ_4 je čtvrtý centrální moment proměnné X.

Oborem hodnot míry β_M je pro spojité náhodné veličiny interval $(1;\infty)$, pro diskrétní náhodné veličiny interval $<1;\infty)$. Pro daný rozsah výběru n je však jeho obor hodnot vymezen intervalem $\left\langle 1; \frac{n^2-3n+3}{n-1}\right\rangle$, jak bylo odvozeno např. v PICKARD (1951).

2.5.1.1 Interpretace momentové míry

Delší dobu byl čtvrtý normovaný moment jedinou používanou mírou této vlastnosti rozdělení. Vývoj takto reprezentovaného pojetí kurtozy prošel řadou etap, jak je popsáno v článku ČERMÁK, VODRÁŽKOVÁ (1991):

Nejstarší interpretací (z které vyplývá dodnes používané označení) je, že tato charakteristika je mírou špičatosti, resp. plochosti rozdělení. Za standard byla považována hodnota $\beta_M = 3$, tj. hodnota čtvrtého normovaného momentu u normálního

rozdělení. Na základě porovnání s touto hodnotou byla potom jednotlivá teoretická rozdělení klasifikována jako mezokurtická (středně špičatá) pro $\beta_M \cong 3$, leptokurtická (úzce špičatá) pro $\beta_M > 3$ resp. platykurtická (široce špičatá, plochá) pro $\beta_M < 3$. Předpokládalo se tedy, že velikost koeficientu špičatosti koresponduje s výškou rozdělení, tj. s hodnou hustoty pravděpodobnosti (v normovaném tvaru) v bodě E(X). Později se ukázalo, že tento vztah platí pouze u některých typů tzv. regulárních souměrných rozdělení, a že naopak zcela selhává u rozdělení nesouměrných. ČERMÁK, VODRÁŽKOVÁ (1991) dále uvádějí řadu příkladů dvojic rozdělení, kdy korespondence mezi výškou rozdělení a hodnotou čtvrtého normovaného momentu neplatí.

Další interpretace byla, že koeficient špičatosti je mírou váhy chvostu rozdělení, později potom, že je mírou dvouvrcholovosti (bimodality) rozdělení. V obou případech však byla demonstrována řada případů, kdy tyto interpretace selhávaly. V současné době uznávaná interpretace, kterou je možnou použít bez ohledu na typ rozdělení, vychází z následující úpravy vztahu pro koeficient β_M:

$$\beta_M = \mathrm{E}(U^4) = \mathrm{D}(U^2) + \mathrm{E}^2(U^2) = \mathrm{D}(U^2) + 1,$$

kde $U = (X - \mu'_1)/\sigma$ je normovaná náhodná proměnná. Z uvedeného rozkladu vyplývá, že tato míra závisí na variabilitě hodnot druhé mocniny normované proměnné U^2 okolo její střední hodnoty 1, tj. na variabilitě proměnné U okolo hodnot +1 a –1. Dosadíme-li za znormovanou veličinu U zpět veličinu X, dostáváme, že koeficient špičatosti β_M je mírou variability veličiny X okolo hodnot $\mu'_1 - \sigma$ a $\mu'_1 + \sigma$.

2.5.1.2 Estimátory momentové míry kurtozy

Pro odhad míry β_M se používá jednak estimátor založený na výběrových momentech, jednak estimátor založený na k-statistikách, tzv. Fisherův „téměř nestranný“ odhad.

a) Estimátor založený na výběrových momentech

Estimátor je definován jako poměr čtvrtého výběrového centrálního momentu a druhé mocniny druhého výběrového centrálního momentu, tj.

$$b_M = \frac{m_4}{m_2^2} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^4}{\left[\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2\right]^2},$$

kde m_4 je čtvrtý a m_2 druhý výběrový centrální moment a $\bar{x}$ je výběrový průměr. Jedná se o asymptoticky nestranný odhad.

b) Fisherův tzv. „téměř nestranný“ odhad

Estimátor, který budu dále označovat b_F, je definován analogicky jako b_M s tím rozdílem, že výběrové momenty m_4 a m_2 jsou nahrazeny tzv. *k*-statistikami (viz např. KENDAL, STUART, 1963):

$$b_F = \frac{k_4}{k_2^2} = \frac{n-1}{(n-2)(n-3)}\left[(n+1)\frac{m_4}{m_2^2} - 3(n-1)\right], \qquad n \geq 4,$$

kde: $k_2 = \frac{n}{n-1} m_2$ je druhá k-statistika,

$k_4 = \frac{n^2}{(n-1)(n-2)(n-3)}\left[(n+1)m_4 - 3(n-1)m_2^2\right]$ čtvrtá k-statistika,.

Estimátor může nabývat hodnot z intervalu $\langle -6; \infty)$, pro daný rozsah výběru n je však jeho obor hodnot vymezen intervalem $\left\langle \frac{(n-1)[(n+1)-3(n-1)]}{(n-2)(n-3)}; n \right\rangle$. Limita dolní meze tohoto intervalu pro $n \to \infty$ je rovna $\lim_{n\to\infty} b_M$ = -2. Pro výběry z normálního rozdělení je střední hodnota estimátoru nulová. Z výše uvedeného vyplývá, že estimátor b_F můžeme považovat za odhad hodnoty $\beta_M - 3$.

2.5.1.3 Výběrové vlastnosti estimátorů momentové míry

Pokud jde o **výběrové vlastnosti estimátoru b_M**, analytickou cestou lze odvodit limitní vztah pro jeho rozptyl (KENDAL, STUART, 1963, str. 243):

$$\lim_{n\to\infty} \mathrm{D}(\mathrm{b_M}) = \frac{1}{n}\left(\beta_6 - 4\beta_2\beta_4 + 4\beta_2^3 - \beta_2^2 + 16\beta_2\beta_1 - 8\beta_3 + 16\beta_1\right) =$$

$$= \frac{1}{n}\left(\frac{\mu_8}{\mu_2^4} - 4\beta\frac{\mu_6}{\mu_2^3} + 4\beta^3 - \beta^2 + 16\beta\alpha^2 - \frac{8\mu_5\mu_3}{\mu_2^4} + 16\alpha^2\right),$$

kde μ_i je i-tý centrální moment, n rozsah výběru, α koeficient šikmosti rozdělení a β koeficient špičatosti rozdělení. Pro výběry z normálního rozdělení po dosazení dostáváme

$$\lim_{n\to\infty} D(b_M) \doteq \frac{24}{n},$$

resp.

$$\lim_{n\to\infty} SE(b_M)\sqrt{n} \doteq 4{,}8990.$$

V PEARSON (1930) jsou dále odvozeny aproximace vztahů pro střední hodnotu, směrodatnou odchylku a pro třetí $\alpha(b_M)$ a čtvrtý $\beta(b_M)$ normovaný moment tohoto estimátoru při výběrech z normálního rozdělení:

$$E(b_M) \doteq \frac{3(n-1)}{n+1},$$

$$\sigma(b_M) \doteq \sqrt{\frac{24}{n}}\left(1 - \frac{15}{2n} + \frac{271}{8n^2} - \frac{2319}{16n^3} + \ldots\right),$$

$$\alpha(b_M) \doteq \frac{216}{n}\left(1 - \frac{29}{n} + \frac{519}{n^2} - \frac{7637}{n^3} + \ldots\right),$$

$$\beta(b_M) \doteq 3 + \frac{540}{n} - \frac{20{,}196}{n^2} + \frac{470{,}412}{n^3} - \ldots.$$

Z výše uvedeného vztahu pro střední hodnotu vyplývá, že estimátor b_M je v případě výběrů z normálního rozdělení vychýlený a velikost jeho vychýlení je:

$$B(b_M) \doteq -\frac{6}{n+1}.$$

Na základě simulace bylo později ověřeno (D'AGOSTINO et al., 1973), že uvedený vztah pro směrodatnou odchylku platí přibližně pro $n \geq 40$, zatímco vztahy pro třetí a čtvrtý normovaný moment výběrového rozdělení estimátoru platí přibližně až pro $n > 200$. Pomocí simulace byly dále tabelovány hodnoty význačných kvantilů

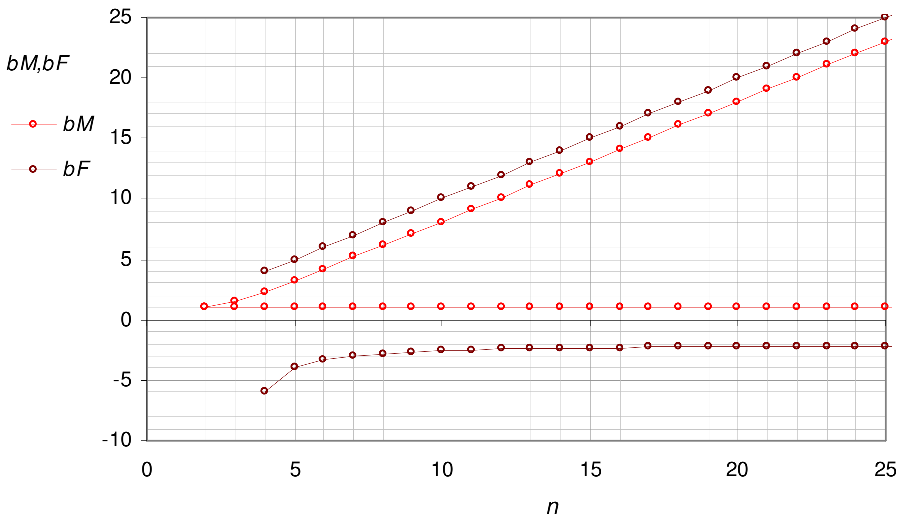
estimátoru b_M (opět pro výběry z normálního rozdělení), např. D'AGOSTINO et al. (1971) pro $n \le 50$ nebo D'AGOSTINO et al. (1973) pro $20 \le n \le 200$.

Pokud jde o **výběrové vlastnosti estimátoru b_F**, byl pro něj odvozen přibližný vztah pro rozptyl (KENDAL, STUART, 1963) pro případ výběru z normálního rozdělení:

$$\mathrm{D}(b_F) \doteq \frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}.$$

Pro libovolný rozsah výběru n potom při výběrech z normálního rozdělení platí, že $\mathrm{D}(b_F) < \mathrm{D}(b_M)$, odhad b_F je tedy při výběrech z normálního rozdělení vydatnější.

Závislost oborů hodnot estimátorů b_M a b_F na rozsahu výběru n je znázorněna na obr. 2.6. Ze skutečnosti, že obor hodnot obou estimátorů je shora omezen, již bez nutnosti použití simulace vyplývá, že pro vysoké hodnoty β_M a nízké rozsahy výběru n budou tyto estimátory poskytovat podhodnocené odhady. Otázkou potom je rychlost konvergence tohoto vychýlení k nule.



Obr. 2.6: Dolní a horní meze oborů hodnot estimátorů b_M a b_F v závislosti na rozsahu výběru n.

2.5.2 Kvantilové míry

Jako kvantilové míry kurtozy byly navrženy tzv. Moorsův koeficient a jeho standardizovaná modifikace (viz např. ČERMÁK, 1993).

Moorsův koeficient

Moorsův koeficient je definován jako funkce 1., 3., 5. a 7. oktilu a obou kvartilů rozdělení:

$$\beta_Q = \frac{x_{0,875} - x_{0,625} + x_{0,375} - x_{0,125}}{x_{0,75} - x_{0,25}},$$

kde x_p je p-procentní kvantil rozdělení. Oborem hodnot koeficientu je interval $<0;\infty)$. V případě normálního rozdělení je hodnota $\beta_Q = 1,2331$.

Standardizovaný Moorsův koeficient

Standardizovaná podoba Moorsova koeficientu je funkcí 1., 3., 5. a 7. oktilu rozdělení:

$$\beta_Q^+ = \frac{x_{0,875} - x_{0,625} + x_{0,375} - x_{0,125}}{x_{0,875} - x_{0,125}}.$$

Oborem hodnot je interval <0; 1>, v případě normálního rozdělení dostáváme $\beta_Q^+ = 0,7230$. Estimátory obou kvantilových měr jsou definovány analogicky jako populační koeficienty s tím rozdílem, že kvantily rozdělení jsou nahrazeny jejich výběrovými protějšky.

Výběrové vlastnosti estimátorů kvantilových měr

Pro estimátor Moorsova koeficientu byl ve WAGEMAKERS et al., 1995, odvozen limitní vztah pro rozptyl v případě výběrů z normálního rozdělení:

$$\lim_{n\to\infty} \mathrm{D}(b_Q) = \frac{3,153}{n},$$

resp.

$$\lim_{n\to\infty} \mathrm{SE}(b_Q)\sqrt{n} = 1{,}776 .$$

Ve zmíněné práci jsou dále uvedeny výsledky zkoumání výběrových vlastností β_Q pomocí simulace 200 výběrů z normovaného normálního rozdělení o rozsazích $n = \{50, 100, 200, 2000\}$. Vzhledem k velmi nízkému počtu výběrů však nelze na základě tohoto experimentu učinit žádné obecně platné závěry.

2.5.3 Míra založená na L–momentech

Míra kurtozy (špičatosti) založená na L-momentech, v literatuře označovaná τ_4 a nazývaná *L-kurtosis*, je definována jako:

$$\tau_4 \equiv \beta_{\mathrm{L}} = \frac{\lambda_4}{\lambda_2} .$$

Charakteristika nabývá hodnot z intervalu <-1, 1>. V případě normálního rozdělení je $\beta_{\mathrm{L}} = 0{,}1226$.

Estimátor a jeho výběrové vlastnosti

Estimátor je definován analogicky jako populační koeficient s tím rozdílem, že L-momenty jsou nahrazeny *l*-statistikami. Jeho výběrové vlastnosti byly, společně s estimátorem míry šikmosti a druhou *l*-statistikou, zkoumány na reálných datech (např. měsíční úhrn srážek, roční maximální průtok řek – viz GUTTANN, 1994). Ve všech případech se výběrové vlastnosti L-momentových měr jevily jako uspokojivé.

2.5.4 Míry založené na LQ-momentech a TL-momentech

Analogicky jako koeficient β_{L} je definován koeficient **založený na LQ-momentech**

$$\beta_{LQ} = \frac{\xi_4}{\xi_2} = \frac{\xi_4}{S_{LQ}} ,$$

který v případě normálního rozdělení nabývá hodnoty $\beta_{LQ} = 0{,}1179$, a koeficient **založený na TL-momentech**, nazývaný TL-kurtosis

$$\beta_{TL} = \frac{\lambda_4^{(1)}}{\lambda_2^{(1)}} = \frac{\lambda_4^{(1)}}{S_{TL}},$$

který v případě normálního rozdělení nabývá hodnoty β_{TL} = 0,0625.

Tabulka 2.4: Sledované estimátory měr kurtozy, jejich obory hodnot a limitní hodnoty pro výběry z normálního rozdělení:

Míra kurtozy	Výpočtový vzorec estiátoru	Obor hodnot	N(0, 1)
Momentový koeficient špičatosti	$b_M = \frac{m_4}{m_2^2}$	<0; ∞), resp. $\left\langle 1; \frac{n^2-3n+3}{n-1} \right\rangle$	3
Fisherův „téměř nestranný“ odhad	$b_F = \frac{k_4}{k_2^2}$	<-6; ∞), resp. $\left\langle \frac{(n-1)[(n+1)-3(n-1)]}{(n-2)(n-3)}; n \right\rangle$	0
Moorsův kvantilový koeficient	$b_Q = \frac{x_{0,875} - x_{0,625} + x_{0,375} - x_{0,125}}{x_{0,75} - x_{0,25}}$	<0; ∞)	1,2331
Normalizovaný Moorsův koeficient	$b_Q^+ = \frac{x_{0,875} - x_{0,625} + x_{0,375} - x_{0,125}}{x_{0,875} - x_{0,125}}$	<0; 1>	0,7230
Charakteristika založená na L-momentech	$b_L = \frac{l_4}{l_2}$	<-1; 1>	0,1226
Charakteristika založená na TL-momentech (pro t = 1)	$b_{TL} = \frac{l_4^{(1)}}{l_2^{(1)}} = \frac{l_4^{(1)}}{s_{TL}}$	?	0,0625
Charakteristika založená na LQ-momentech (pro $p = \alpha = 0{,}25$)	$b_{LQ} = \frac{est\ \xi_4}{est\ \xi_2} = \frac{est\ \xi_4}{s_{LQ}}$	?	0,1179

3. Výběrové vlastnosti a kritéria kvality estimátorů

Výběrová rozdělení estimátorů lze popsat z hlediska celé řady vlastností. Obecně se u nich sledují charakteristiky polohy (resp. vychýlení), variability a tvaru rozdělení, reprezentované mírou šikmosti, případně kurtozy. Dále se u výběrových rozdělení sleduje rychlost jejich konvergence k některému standardnímu rozdělení, zpravidla normálnímu, v závislosti na rozsahu výběru *n*. Pro potřeby vzájemného porovnání jednotlivých estimátorů je dále nutno charakteristiky vychýlení a variability vyjádřit tak, aby byly vzájemně srovnatelné. Sledují se tedy charakteristiky relativní variability a relativního vychýlení estimátoru. Zvláštní postavení pak mají hodnoty některých „význačných“ kvantilů výběrového rozdělení, které se používají pro induktivní úsudky o hodnotě zkoumané míry v populaci, tj. pro konstrukci intervalů spolehlivosti a testování hypotéz. Připomínám, že se vždy jedná o výběrové charakteristiky estimátoru pro konkrétní rozsah výběru a konkrétní typ a parametry rozdělní, resp. populace, z níž se výběry pořizují.

3.1 Absolutní charakteristiky polohy, vychýlení a variability

Charakteristiky polohy a vychýlení estimátoru

Jako míra polohy rozdělení estimátoru *t* se používá **střední hodnota**. Její odhad potom získáme jako

$$\text{est E}(t) = \frac{1}{m}\sum_{j=1}^{m} t_j \ ,$$

kde *m* je počet provedených výběrů a t_j je hodnota odhadu příslušné míry v *j*-tém výběru. Z odhadu střední hodnoty estimátoru se jednoduše získá odhad jeho **vychýlení**

$$\text{est B}(t) = est\ \text{E}(t) - \Theta ,$$

kde Θ je skutečná hodnota odhadované míry v populaci.

Charakteristiky variability estimátoru

Pro popis variability estimátoru se požívá (vedle střední čtvercové chyby - viz níže) **rozptyl**, jehož odhad se získá:

$$\text{est D}(t) = est\ \text{E}(t - est\ \text{E}(t))^2 = \frac{1}{m-1}\sum_{j=1}^{m}(t_j - est\ \text{E}(t))^2 ,$$

a **směrodatná odchylka**, jejíž odhad est S(t) se získá z odhadu rozptylu.

Charakteristika variability a vychýlení zároveň

Pro celkové posouzení kvality estimátoru jak z hlediska variability tak z hlediska vychýlení se používá **střední čtvercová chyba** (MSE, ze zkratky anglického označení *mean squared error*), resp. její odmocnina, **směrodatná chyba** (SE, dle anglického *standard error*). Jejích odhady se získají takto:

$$\text{est MSE} = est\ \text{E}(t - \Theta)^2 = \frac{1}{m-1}\sum_{j=1}^{m}(t_j - \Theta)^2 ,$$

případně

$$\text{est SE} = \sqrt{\text{est MSE}} .$$

Je zřejmé, že směrodatná chyba *SE* je z výše uvedených charakteristik nejdůležitější, protože v sobě zahrnuje současně vychýlení i variabilitu a nejlépe tak charakterizuje kolísání hodnot estimátoru t_j okolo skutečné hodnoty odhadované míry v populaci Θ.

3.2 Relativní charakteristiky vychýlení a variability

Pro vzájemné porovnání estimátorů jednotlivých měr z hlediska vychýlení a variability mají výše uvedené absolutní charakteristiky (tj. *B*, *D* a *SE*) ten nedostatek, že jsou závislé na měřítku jednotlivých měr. Je sice možné použít je pro vzájemné porovnání estimátorů, jejichž obor hodnot tvoří stejný, pokud možno uzavřený, interval, nicméně pro celkové porovnání všech estimátorů mezi sebou jsou nevhodné. Pro tyto účely je tedy třeba zvolit vhodné relativní charakteristiky.

Otázkou však je, jak zvolit vhodný základ, ke kterému budou absolutní míry vztaženy. Jednu z možností je zvolit princip odvozený od některé míry relativní variability, tj. vyjádřit absolutní vychýlení, směrodatnou odchylku nebo směrodatnou chybu v poměru buď ke skutečné hodnotě dané míry v populaci Θ, nebo ke střední hodnotě jejího estimátoru E(t). Tento postup lze bez problému použít pro porovnání estimátorů měr variability a relativní variability, neboť všechny mají dolní hranici oboru hodnot 0.

Tento postup však selhává při porovnání charakteristik šikmosti a kurtozy, z nichž některé mohou nabývat kladných i záporných hodnot. Důvod je ten, že většina používaných měr relativní variability, například variační koeficient, relativní kvartilová odchylka, ale také míra založená na L-momentech (viz např. ČERMÁK, 1993 nebo ČERMÁK, VRABEC, 1996) je použitelná pouze pro nezáporné náhodné veličiny (přesněji řečeno, pro veličiny, z nichž každá může nabývat buď pouze nezáporných nebo pouze nekladných hodnot), jinak selhávají. Jedinou mírou relativní variability částečně odstraňující tento nedostatek je standardizovaný variační koeficient zavedený v kapitole 2 (viz např. ČERMÁK, 1993), jehož jmenovatel je vždy nezáporný:

$$\mathrm{V}_{\mathrm{X}}^{+} = \frac{\mathrm{D}(X)}{\sqrt{\mathrm{E}\{X^2)}},$$

který nabývá hodnot z intervalu <0; 1>. Například pro relativní směrodatnou chybu estimátoru t bychom tak dostali:

$$\mathrm{SE}_{\mathrm{r}}^{V+}(t) = \frac{\mathrm{SE}(t)}{\sqrt{\mathrm{E}\{t^2)}}.$$

Tato míra však řeší zmíněný nedostatek pouze částečně: V případě, kdy $\mathrm{E}(X) \to 0$, nabývá V_X^+ hodnoty blízké k jedné. Pro účely vzájemného porovnání estimátorů měr šikmosti a kurtozy by tedy selhala v případech, kdy by daná míra nabývala hodnot blízkých nule, tj. pokud $\Theta \to 0$ resp. $\mathrm{E}(t) \to 0$.

Pro relativní vyjádření variability a vychýlení estimátorů měr šikmosti a kurtozy jsem využil následující skutečnosti: všechny zkoumané estimátory mají, pro daný rozsah výběru n, obor hodnot zdola omezen, přičemž dosažení spodní hranice oboru je sice možné, týká se však pouze teoretických případů (pro teoretické příklady rozdělení s minimální hodnotou kurtozy viz např. ČERMÁK, VODRÁŽKOVÁ, 1991).

Pro stanovení relativních charakteristik je tedy třeba absolutní hodnoty vychýlení, variability a směrodatné chyby vztáhnout k „posunuté“ míře polohy, tj. k rozdílu míry polohy Θ resp. E(t) a dolní hranice oboru hodnot daného estimátoru. Otázkou ještě zůstává, kterou charakteristiku zvolit jako míru polohy: Jednou z možností je střední hodnota estimátoru E(t), druhou možností je skutečná hodnota příslušné míry Θ v populaci. Zřejmě nelze jednoznačně určit, která z možností je lepší, ovšem vzhledem k tomu, že empirická střední hodnota estimátoru může být zatížena vychýlením, zvolil jsem první možnost, tj. skutečnou hodnotu v populaci Θ.

Pro posouzení relativní variability, relativního vychýlení a směrodatné chyby jsem tedy zavedl následující míry:

Relativní vychýlení:

$$\mathrm{B_r}(t) = \frac{\mathrm{B}(t)}{\Theta - t_{MIN}(n)},$$

kde Θ je skutečná hodnota příslušné míry v populaci

a $t_{MIN}(n)$ je dolní mez oboru hodnot estimátoru t (pro momentové estimátory měr šikmosti a kurtozy závislá na n).

Relativní směrodatná odchylka:

$$S_r(t) = \frac{\mathrm{S}(t)}{\Theta - t_{MIN}(n)}.$$

Relativní směrodatná chyba:

$$\mathrm{SE_r}(t) = \frac{\mathrm{SE}(t)}{\Theta - \mathrm{t_{MIN}}(n)}.$$

Další charakteristiky

Střední čtvercovou chybu lze rozložit na součet rozptylu estimátoru a čtverce jeho vychýlení (viz např. ČERMÁK, VRABEC, 1999), tedy $\mathrm{MSE} = \mathrm{D}(t) - \mathrm{B}^2(t)$. Pro měření relativního vychýlení jsem dále zavedl charakteristiku vyjadřující podíl čtverce vychýlení na střední čtvercové chybě definovanou jako:

$$\mathrm{B_r^{MSE}}(t) = \frac{\mathrm{B(t)}^2}{\mathrm{MSE}} \quad ,$$

která nabývá hodnot z intervalu <0; 1>. Dále jsem sledoval hodnoty výrazu $SE(t)\sqrt{n}$, neboť u některých estimátorů lze teoreticky odvodit konstantu, k níž tento výraz v případě výběrů z normálního rozdělení konverguje.

3.3 Tvar výběrového rozdělení

Pro popis tvaru výběrového rozdělení a také pro posouzení vhodnosti jeho aproximace rozdělením normálním se zpravidla sledují **momentové koeficienty šikmosti** a **špičatosti**, jejichž odhady se získají dle vztahů

$$est\ \alpha_M(t) = \frac{\mathrm{m}_3(t)}{\mathrm{m}_2^{3/2}(t)} = \frac{\sum_{i=1}^{m}\left(t_i - \mathrm{E}(t)\right)^3}{\left[\sum\left(t_i - \mathrm{E}(t)\right)^2\right]^{2/3}}$$

resp.

$$est\ \beta_M(t) = \frac{\mathrm{m}_4(t)}{\mathrm{m}_2^{2}(t)} = \frac{\sum_{i=1}^{m}\left(t_i - \mathrm{E}(t)\right)^4}{\left[\sum\left(t_i - \mathrm{E}(t)\right)^2\right]^{2}}.$$

U těchto charakteristik se sleduje, do jaké míry se jejich hodnoty liší od hodnot pro rozdělení normální, tj. 0 pro $\alpha_M(t)$ a 3 pro $\beta_M(t)$.

Pro každý sledovaný estimátor jsem dále provedl **Kolmogorovův - Smirnovův test o normalitě** jeho **výběrového rozdělení**. Test jsem prováděl na hladinách významnosti 1 %, 5 % a 10 % a pro každý rozsah výběru jsem sledoval, zda a pro jakou nejvyšší hladinu významnosti výběrové rozdělení testu vyhoví. Dále jsem pro jednotlivá výběrová rozdělení estimátoru získal odhad jejich hustoty pravděpodobnosti pomocí **normovaného histogramu rozdělení četností**.

3.4 Měří všechny charakteristiky stejnou vlastnost rozdělení?

Všechny nově zkonstruované jednorozměrné popisné charakteristiky byly navrženy jako alternativa k charakteristikám založených na klasických momentech. Je tedy žádoucí, aby mezi hodnotami estimátorů nově navržených a momentových charakteristik existovala statisticky významná závislost. Proto jsem se na základě výsledků simulace pokusil zhodnotit, zda tato závislost skutečně existuje a jak je silná. Pro posouzení této závislosti používám jednak Pearsonův korelační koeficient a jednak bodový graf hodnot estimátorů. Korelační matici a bodový graf jsem vytvořil pro každou populaci a rozsah výběru.

3.5 Posouzení síly testu

3.5.1 Testy o normalitě rozdělení

Jak již bylo zmíněno, významnou roli hrají momentové míry šikmosti a kurtozy při ověřování normality rozdělení. Zpravidla se postupuje tak, že se na základě hodnoty estimátoru získané z výběru testuje hypotéza H_0: $\Theta = \Theta_N$ proti alternativě H_1: $\Theta > \Theta_N$, případně H_2: $\Theta < \Theta_N$ nebo H_2: $\Theta \neq \Theta_N$, kde Θ je neznámá hodnota momentové míry kurtozy resp. šikmosti v populaci a Θ_N je hodnota kurtozy resp. šikmosti normálního rozdělení. Z tabelovaných hodnot kvantilů výběrového rozdělení estimátorů b_M resp. a_M se pro zvolenou hladinu významnosti stanoví kritický obor a pokud se hodnota estimátoru v tomto oboru nachází, testovanou hypotézu zamítáme. Zamítnutí testované hypotézy potom znamená také zamítnutí hypotézy o normalitě rozdělení, z něhož výběr pochází. Naskýtá se otázka, zda by k tomuto účelu nebylo vhodné používat také nově navržené míry kurtozy resp. šikmosti.

Efektivnost použití každého z estimátorů jako testového kriteria lze hodnotit podle síly testu, což je pravděpodobnost, že v případě platnosti alternativní hypotézy bude testovaná hypotéza zamítnuta. Tuto pravděpodobnost lze však jednoznačně určit pouze v případě jednoduché alternativní hypotézy, tj. hypotézy ve tvaru H_1: $\Theta = \Theta^*$. Při posouzení vhodnosti jednotlivých estimátorů šikmosti a kurtozy jako kritéria normality

jsem tedy postupoval následovně: pro zvolené *n* a hladinu významnosti $\alpha = 5\%$ jsem na základě empiricky získaných distribučních funkcí jednotlivých estimátorů při výběrech z normálního rozdělení stanovil pro každý z nich kritický obor. Ten jsem stanovil vždy jako oboustranný, tedy vymezený 2,5% a 97,5% kvantilem výběrového rozdělení příslušného estimátoru. Dále jsem pro všechny ostatní zkoumané populace na základě empirických výběrových distribučních funkcí estimátorů pro každý z nich vypočetl pravděpodobnost, že se jeho hodnota bude nacházet v kritickém oboru. Takto získanou pravděpodobnost lze považovat za odhad síly testu o normalitě rozdělení vůči příslušné alternativě.

3.5.2 Zobecnění testu o normalitě

Výše uvedený postup je možno zobecnit: můžeme si položit si otázku, nakolik je každý estimátor schopen rozlišit rozdělení s vysokou úrovní variability (resp. relativní variability, šikmosti nebo kurtozy) od rozdělení s nízkou úrovní této vlastnosti (a naopak). Pro posouzení schopnosti každého estimátoru rozlišit od sebe jednotlivá rozdělení jsem postup uvedený v předchozím odstavci provedl pro všechny možné variace zkoumaných populací, čímž jsem pro každý estimátor a rozsah výběru *n* získal odhad průměrné síly testu. Na základě odhadnuté průměrné síly testu jsem následně jednotlivé estimátory mezi sebou porovnal.

4. Použitá data a pravděpodobnostní rozdělení

Náhodné výběry byly prováděny jednak z teoretických pravděpodobnostních rozdělení a jednak z reálných datových souborů. Teoretická pravděpodobnostní rozdělení a datové soubory (dohromady označuji jako populace), na nichž byl experiment realizován, byly zvoleny tak, aby postihly co možná nejširší spektrum možných tvarů rozdělní, a to jak z hlediska kurtozy tak z hlediska zešikmení. Jsou zde zastoupeny populace s nízkou, střední i extrémně vysokou úrovní měr kurtozy, stejně tak jsou zastoupena rozdělení souměrná i extrémně zešikmená. V experimentu bylo prozkoumáno celkem 40 různých populací.

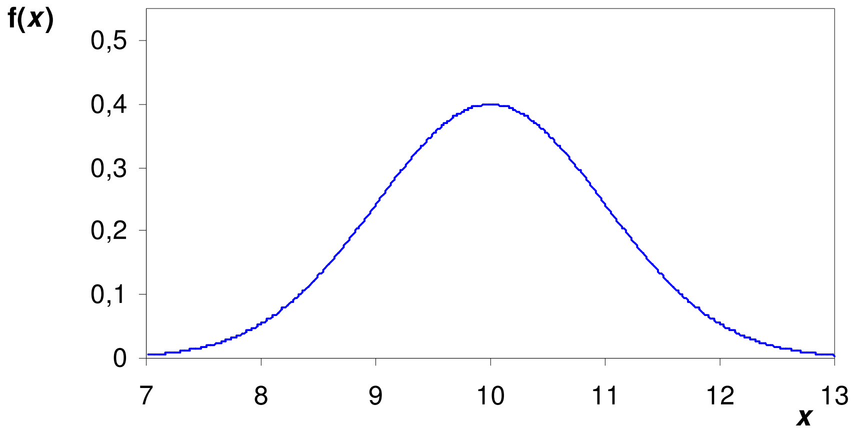
4.1 Teoretická pravděpodobnostní rozdělení

Označení parametrů jednotlivých teoretických rozdělení a někdy i jejich počty se v literatuře často liší. Jako standard jsem zvolil vždy základní tvar každého rozdělení dle příručky ČERMÁK (1993).

a) Normální rozdělení

Hustota pravděpodobnosti: $f_X(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$,

$-\infty < x < \infty,\ -\infty < \mu < \infty,\ 0 < \sigma < \infty$

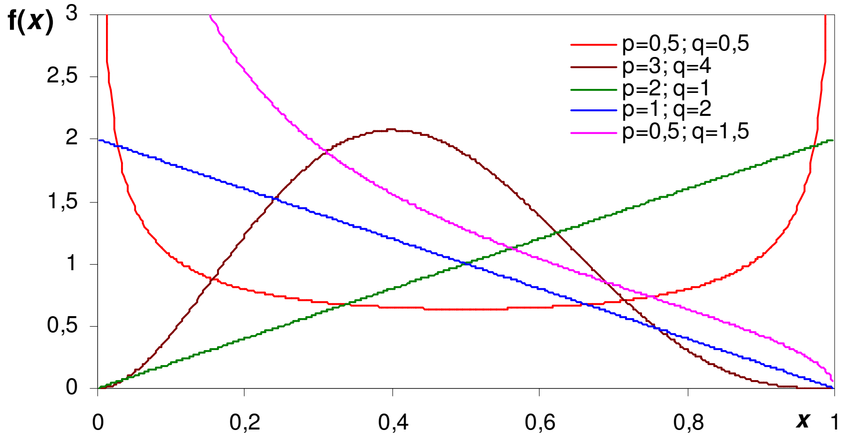


Obr. 4.1: Hustota pravděpodobnosti normálního rozdělení N(10; 1).

b) Rozdělení beta

Hustota pravděpodobnosti: $$f_X(x; p, q) = \frac{1}{B(p,q)} x^{p-1}(1-x)^{q-1},$$

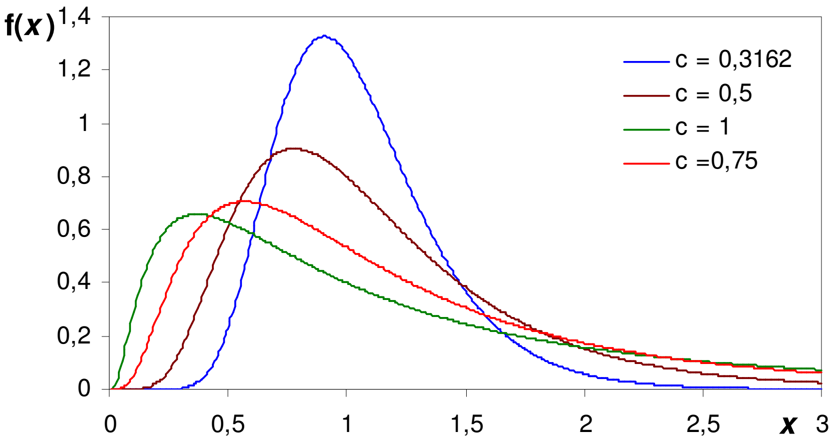
$0 < x < 1$, p, $q > 0$, B(p, q) je funkce beta



Obr. 4.2: Hustota pravděpodobnosti rozdělení beta pro (p; q) = {(0,5; 0,5), (3; 4), (1; 2), (2; 1), (0,5; 1,5)}.

c) Logaritmicko-normální rozdělení

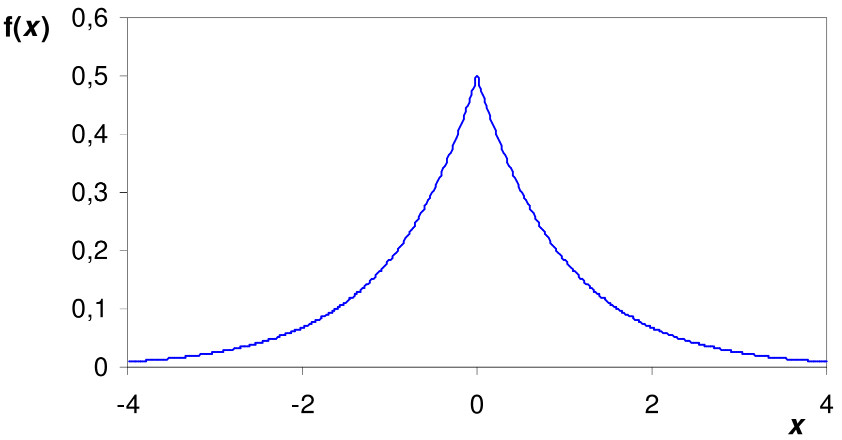
Hustota pravděpodobnosti: $f_X(x;c) = \frac{1}{c\sqrt{2\pi}}\frac{1}{x}\exp\left[-\frac{1}{2c^2}(\ln x)^2\right], \qquad x > 0,\ c > 0.$



Obr. 4.3: Hustota pravděpodobnosti logaritmicko-normálního rozdělení pro parametr $c = \{\sqrt{0{,}1};\ 0{,}5;\ 0{,}75;\ 1\}$

d) Rozdělení Laplaceovo

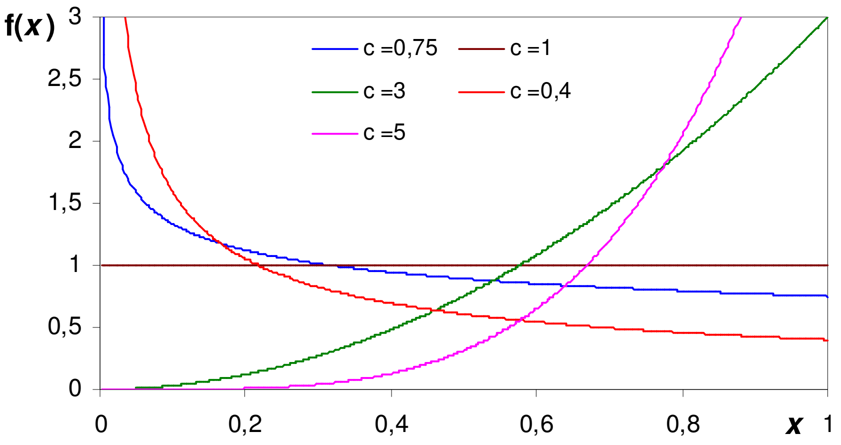
Hustota pravděpodobnosti: $f_X(x) = \frac{1}{2}e^{-|x|}, \qquad -\infty < x < \infty$



Obr. 4.4: Hustota pravděpodobnosti Laplaceova rozdělení.

e) Mocninové rozdělení

Hustota pravděpodobnosti: $f_X(x;c) = cx^c$, $0 < x < 1,\ c > 0$.

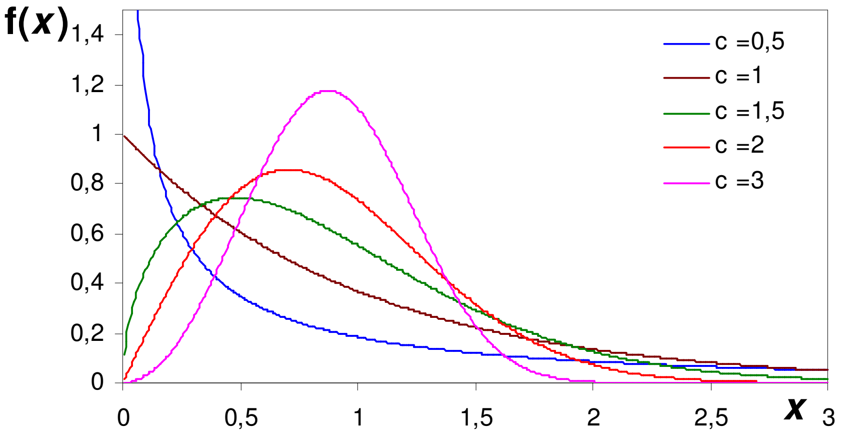


Obr. 4.5: Hustota pravděpodobnosti mocninového rozdělení pro parametry $c = \{0{,}4;\ 0{,}75;\ 1;\ 3;\ 5\}$.

Speciálně pro $c = 0$ dostáváme **rovnoměrné rozdělení** s parametry $a = 0$, $b = 1$.

f) Weibullovo rozdělení

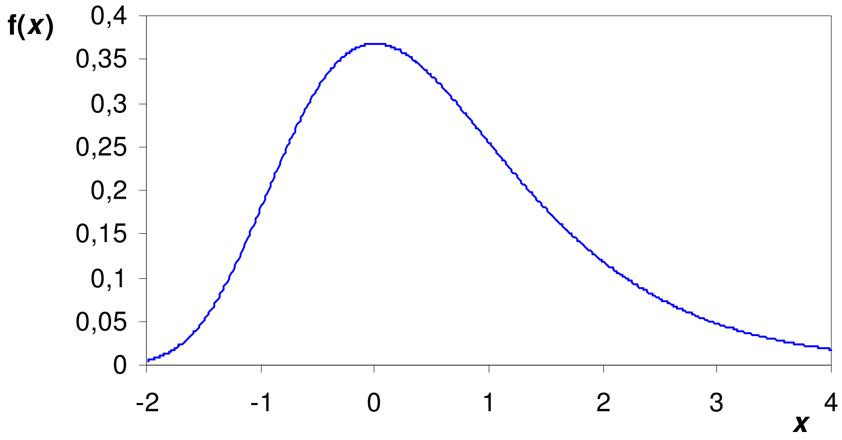
Hustota pravděpodobnosti: $f_X(x;c) = cx^{c-1}\exp(-x^c)$, $x > 0,\ c > 0$.



Obr. 4.6: Hustota pravděpodobnosti Weibullova rozdělení pro parametry $c = \{0{,}5;\ 1;\ 1{,}5;\ 2;\ 3\}$.

g) Gumbelovo rozdělení

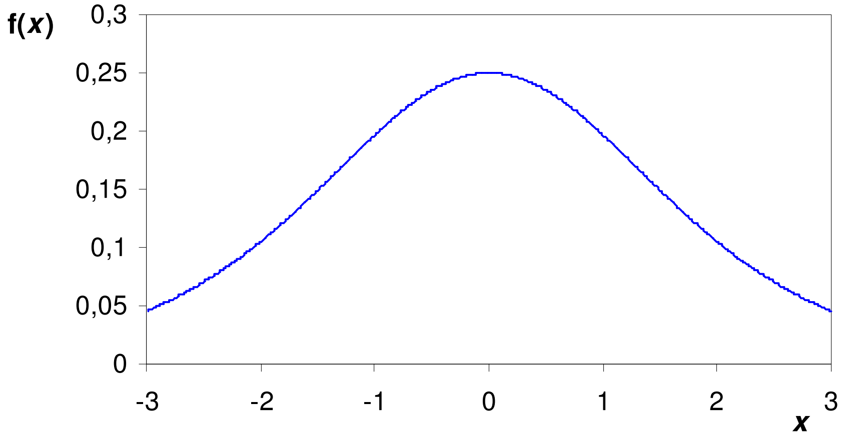
Hustota pravděpodobnosti: $f_X(x) = \exp(-x - \exp(-x))$, $x > 0$, $c > 0$



Obr. 4.7: Hustota pravděpodobnosti Gumbelova rozdělení.

h) Logistické rozdělení

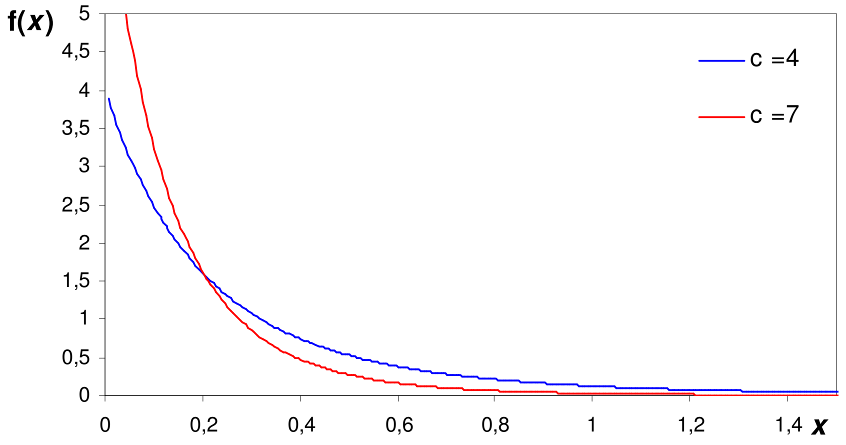
Hustota pravděpodobnosti: $f_X(x) = \dfrac{\exp(-x)}{(1+\exp(-x))^2}$, $x > 0$, $c > 0$



Obr. 4.8: Hustota pravděpodobnosti logistického rozdělení.

i) Paretovo rozdělení

Hustota pravděpodobnosti: $f_X(x;c) = \dfrac{c}{(x+1)^{c+1}}, \; x > 0, \, c > 0.$

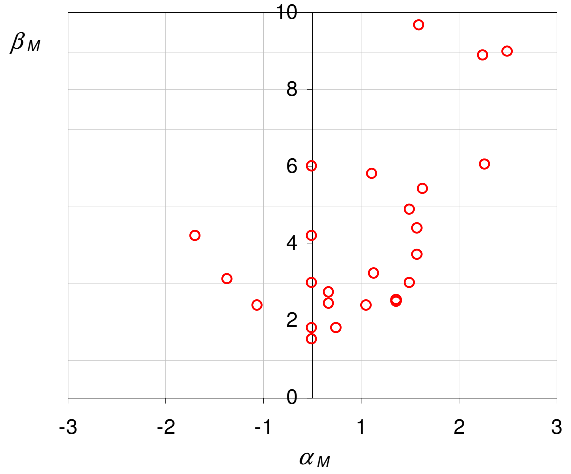


Obr. 4.9: Hustota pravděpodobnosti Paretova rozdělení s parametrem $c = \{4; 7\}$.

4.2 Konečné datové soubory

Kromě výše uvedených teoretických pravděpodobnostních rozdělení jsem do experimentu zahrnul reálné datové soubory. Jedná se o data z oblasti telekomunikací, služeb, zemědělství, financí aj., jejichž histogramy rozdělení četností a základní popisné charakteristiky uvádím v příloze.

Pro ilustraci je na obr. 4.10 znázorněn bodový graf hodnot momentových koeficientů šikmosti a špičatosti α_M a β_M pro ty zkoumané populace, které mají uvedené koeficienty z rozsahu $\alpha_M \in$ <-3; 3> a $\beta_M \in$ <1; 10>. (Při zahrnutí všech hodnot by byl graf nepřehledný, neboť oba koeficienty mají shora neomezený obor hodnot.)



Obr. 4.10: Bodový graf hodnot koeficientů šikmosti a špičatosti α_M a β_M ve zkoumaných populacích pro $\alpha_M \in$ <-3; 3> a $\beta_M \in$ <1; 10>.

5. Realizace experimentu

5.1 Popis realizace

Z konečných základních souborů a z teoretických pravděpodobnostních rozdělení byly vygenerovány náhodné výběry o následujících rozsazích: $n \in \{7; 15; 23; 31; 47; 63; 100; 200; 350; 500; 1000\}$. Pro každý rozsah výběru n bylo generováno celkem $m = 16\,000$ výběrů. Výběry generovány prostým náhodným výběrem, s vracením i bez vracení. Z každého vygenerovaného náhodného výběru byly pořízeny hodnoty všech sledovaných estimátorů.

***Poznámka**: Rozsahy výběrů pro $n \leq 31$ byly stanoveny z důvodu srovnatelnosti výsledků s dříve provedenými experimenty popsanými v ČERMÁK, LAUBER (1996) a ČERMÁK, VRABEC (1998).*

Po provedení všech $m = 16\,000$ výběrů o daném rozsahu n z jedné populace byly pro každý sledovaný estimátor vypočteny odhady všech charakteristik jeho výběrového rozdělení, které byly popsány v kapitole 3, tj. např. absolutní i relativní vychýlení, směrodatná chyba, koeficient šikmosti a špičatosti a byl proveden test Kolmogorova - Smirnova o normalitě jeho výběrového rozdělení. Všechny vypočtené charakteristiky výběrových rozdělení estimátorů byly uloženy do textového souboru pro další zpracování (grafické i tabelární). Dále byly pro každý estimátor uloženy jeho hodnoty získané ze všech $m = 16\,000$ výběrů pro další vyhodnocení z hlediska síly závislosti mezi jednotlivými estimátory.

Dosud popsaný postup byl realizován v programu, který jsem vytvořil v objektově orientovaném programovacím jazyku Microsoft Visual Basic 6.0. Následující zpracování probíhalo v tabulkovém procesoru Microsoft Excel 2000 s pomocí maker naprogramovaných v jazyku Visual Basic for Applications. Uložené textové soubory byly s pomocí maker naimportovány do tabulkového procesoru MS Excel. V něm byly vypočteny hodnoty korelačních koeficientů mezi jednotlivými estimátory a vytvořeny veškeré grafické a tabelární výstupy. Pomocí makra byly dále provedeny výpočty síly testů o tvaru rozdělení, z něhož výběr pochází.

5.2 Technika generování náhodných výběrů

5.2.1 Výběry z teoretických rozdělení

Obecný postup generování náhodných veličin je popsán např. v LAUBER, HUŠEK (1987). Všechny obecně používané metody generování náhodných veličin jsou založeny na principu transformace náhodných čísel na hodnoty veličiny X s požadovaným pravděpodobnostním rozdělením. Tato transformace se nejčastěji provádí *metodou inverzní transformace*, *zamítací metodou* nebo *kompoziční metodou.* Všechny tyto metody tedy vycházejí z předpokladu, že máme k dispozici ideální generátor náhodných čísel. Termínem *náhodná čísla* rozumíme nezávislé stejně rozdělené veličiny s rovnoměrným rozdělením na intervalu (0;1), tj. $R(0;1)$. Protože pro generování náhodných čísel se téměř výhradně používají metody založené na aritmetických algoritmech, označují se někdy takto získané veličiny jako *pseudonáhodná čísla*. Podrobnější popis metod generování náhodných čísel lze najít například v LAUBER, HUŠEK (1987) nebo HURT (1982).

5.2.1.1 Metoda inverzní transformace

Při generování hodnot z teoretických pravděpodobnostních rozdělení jsem používal výhradně metodu inverzní transformace, resp. její modifikaci. Obecně tato metoda vychází z následujícího poznatku (viz např. ANTOCH, VORLÍČKOVÁ, 1992): Nechť U je náhodná veličina s rozdělením $R(0;1)$ a nechť $\mathrm{Q}(u)$ je inverzní funkcí k distribuční funkci $F(x)$, neboli kvantilová funkce rozdělení s distribuční funkcí $F(x)$. Potom náhodná veličina $X = \mathrm{Q}(U)$ má rozdělení s distribuční funkcí $F(x)$.

Metoda inverzní transformace *pro spojité náhodné veličiny* spočívá v dosazení náhodných čísel do rovnice kvantilové funkce, tj. v aplikaci předpisu

$$x = \mathrm{Q}(u)$$

kde u je náhodné číslo
a x je hodnota náhodné veličiny s distribuční funkcí $F(x)$ a s kvantilovou funkcí $\mathrm{Q}(u)$.

Pokud známe kvantilovou funkci $\mathrm{Q}(u)$, je metoda inverzní transformace nejefektivnější z výše uvedených, neboť pro vygenerování n hodnot náhodné proměnné X stačí vygenerovat právě n náhodných čísel (viz např. LAUBER, HUŠEK, 1987).

V případech, kdy nejsme schopni kvantilovou funkci exaktně určit, lze využít *modifikace* této metody, kdy kvantilovou funkci nahradíme její vhodnou aproximací.

5.2.1.2 Metody generování hodnot ze zkoumaných rozdělení

Rozdělení s explicitně vyjádřenou kvantilovou funkcí

Při generování hodnot z jednotlivých teoretických rozdělení jsem v případě náhodných rozdělení, jejichž kvantilovou funkci lze explicitně vyjádřit, použil metodu inverzní transformace. Jedná se o následující rozdělení (vztahy pro výpočet kvantilové funkce viz také ČERMÁK, 1993):

- **Weibullovo rozdělení W(c)**

  $$x = [-\ln(u)]^{1/c}.$$

- **Mocninné rozdělení Power(c)**

  $$x = u^{1/c}$$

- **Laplaceovo rozdělení**

  $$x = \ln(2u) \qquad \text{pro } u \le 0{,}5;$$

  $$x = -2\ln(1-u) \qquad \text{pro } u \ge 0{,}5.$$

- **Rovnoměrné rozdělení R(a, b)**

  $$x = a + u(b-a).$$

- **Parettovo rozdělení Par(c)**

  $$x = \frac{1}{(u)^{1/c}} - 1.$$

- **Gumbelovo rozdělení**

  $$x = -\ln(-\ln u))$$

- **Logistické rozdělení**

  $$x = \ln\frac{u}{1-u}$$

Normální rozdělení

Pro generování výběrů z **normálního rozdělení**, jehož kvantilovou funkci nelze vyjádřit analyticky, jsem využil aproximaci kvantilové funkce normovaného normálního rozdělení navrženou v ACKLAM. Aproximace platí s relativní chybou $\pm 1{,}15 \cdot 10^{-9}$. Pseudokód jejího algoritmu je následující:

Definice koeficientů aproximace:

```
a(1) = -3.969683028665376e+01
a(2) =  2.209460984245205e+02
a(3) = -2.759285104469687e+02
a(4) =  1.383577518672690e+02
a(5) = -3.066479806614716e+01
a(6) =  2.506628277459239e+00

b(1) = -5.447609879822406e+01
b(2) =  1.615858368580409e+02
b(3) = -1.556989798598866e+02
b(4) =  6.680131188771972e+01
b(5) = -1.328068155288572e+01

c(1) = -7.784894002430293e-03
c(2) = -3.223964580411365e-01
c(3) = -2.400758277161838e+00
c(4) = -2.549732539343734e+00
c(5) =  4.374664141464968e+00
c(6) =  2.938163982698783e+00

d(1) =  7.784695709041462e-03
d(2) =  3.224671290700398e-01
d(3) =  2.445134137142996e+00
d(4) =  3.754408661907416e+00
```

Definice hranice horní a dolní oblasti:

```
p_low  = 0.02425
p_high = 1 - p_low
```

Aproximace pro dolní oblast:

```
If 0 < p < p_low Then
   q = sqrt(-2*log(p))
   x = (((((c(1)*q+c(2))*q+c(3))*q+c(4))*q+c(5))*q+c(6)) /
         ((((d(1)*q+d(2))*q+d(3))*q+d(4))*q+1)
End If
```

Aproximace pro centrální oblast:

```
If p_low <= p <= p_high Then
   q = p - 0.5
   r = q*q
   x = (((((a(1)*r+a(2))*r+a(3))*r+a(4))*r+a(5))*r+a(6))*q /
        (((((b(1)*r+b(2))*r+b(3))*r+b(4))*r+b(5))*r+1)
End If
```

```
Aproximace pro horní oblast:
If p_high < p < 1 Then
   q = sqrt(-2*log(1-p))
   x = -(((((c(1)*q+c(2))*q+c(3))*q+c(4))*q+c(5))*q+c(6)) /
          ((((d(1)*q+d(2))*q+d(3))*q+d(4))*q+1)
End If
```

V uvedeném zdroji jsou také k dispozici zdrojové kódy od různých autorů pro implementaci tohoto algoritmu ve standardních programovacích jazycích. Ve své práci jsem použil zdrojový kód pro Visual Basic 6 od autora HERRERO.

Logaritmicko-normální rozdělení

Pro generování výběrů z logaritmicko-normálního rozdělení jsem využil následujícího vztahu (např. LAUBER, HUŠEK, 1987): Má-li náhodná veličina X rozdělení LN(μ; σ), má náhodná veličina

$$Y = \ln X$$

rozdělení normální N(μ; σ). Jednoduchou úpravou dostaneme:

$$X = e^{Y}.$$

Generování hodnot proměnné X' z logaritmicko-normálního rozdělení LN(c) jsem tedy prováděl transformací vygenerovaných hodnot proměnné Y' z normálního rozdělení N(0; c) podle vztahu:

$$X' = e^{Y'}.$$

5.2.2 Výběry z konečných souborů

Pro generování prostého náhodného výběru o rozsahu n z konečného datového souboru o rozsahu N jsem použil analogii výběrové techniky založené na tabulkách náhodných čísel popsané v ČERMÁK, VRABEC (1999). Každému prvku jsem přiřadil jednoznačné pořadí. Pořadí prvku, který má být vybrán v daném kroku, jsem vygeneroval jako hodnotu diskrétního rovnoměrného rozdělení $R_D(1; N)$:

$$Z_i = int(N \cdot u) + 1,$$

kde Z_i je pořadí prvku vybraného v i-tém kroku, u je náhodné číslo a funkce $int(a)$ vrací celou část reálného čísla a.

5.3 Metody určení kvantilů

Pro určení kvantilů jsem použil *metodu váženého průměru*:

$$X_p = X_{(j+1)} + g \cdot (X_{(j+2)} - X_{(j+1)})$$

kde X_p je hledaný kvantil, j je celá část výrazu $(N-1) \cdot p$, g je desetinná část výrazu $(N - 1) \cdot p$ a $X_{(j)}$ je j-tý prvek v souboru v pořadí podle velikosti. Pro setřídění souboru, které je nutno provést pro určení pořádkových statistik, jsem používal algoritmus QuickSort.

Pro výpočet estimátorů charakteristik založených na mediánu funkce pořádkových statistik, tj. s_{HL} a a_{MC}, je však třeba nalézt medián hodnot $|x_i - x_j|$, resp. medián hodnot funkce h, jejichž počet se s rostoucím n rychle zvyšuje. Např. počet hodnot $|x_i - x_j|$ je roven počtu dvouprvkových kombinací z n hodnot, tj. $C_2(n)$. Pro tento výpočet jsem z důvodu optimalizace použil *metodu založenou na empirické distribuční funkci*:

$$X_p = X_{(j)} \qquad \text{pro } g = 0,$$

$$X_p = X_{(j+1)} \qquad \text{pro } g \neq 0.$$

Tato metoda vyžaduje znalost pouze jediné pořádkové statistiky a umožňuje tak dobu výpočtu zkrátit použitím algoritmu, který nevyžaduje předem setříděné pole. Pseudokód algoritmu nalezení j-té pořádkové statistiky bez nutnosti setřídění souboru je uveden např. v DEVILLARD (1998). Při aplikaci tohoto algoritmu v rámci provedeného experimentu se ukázalo, že uvedená metoda zkracuje čas výpočtu pořádkové statistiky oproti třídícímu algoritmu QuickSort při výpočtu charakteristiky s_{HL} pro rozsah výběru $n = 1000$, kdy vznikne celkem $C_2(1000) = 499\,500$ kombinací, přibližně o 40%. Z toho také vyplývá, že její použití je výhodné pouze v případě, že hledáme pouze jednu pořádkovou statistiku z daného souboru. Pro úplnost dodávám, že i přes použitou optimalizaci zabral výpočet hodnot estimátorů s_{HL} a a_{MC} přibližně 95% strojového času potřebného pro realizaci experimentu.

6. Výsledky

Při porovnání výběrových vlastností jednotlivých estimátorů nejsou patrné zásadní odlišnosti mezi výběrem s vracením a bez vrací. Níže uvedené výsledky tedy platí pro obě varianty prostého náhodného výběru. Veškeré grafy a tabulky prezentované v této práci byly získány na základě výběrů s vracením.

6.1 Estimátory měr variability

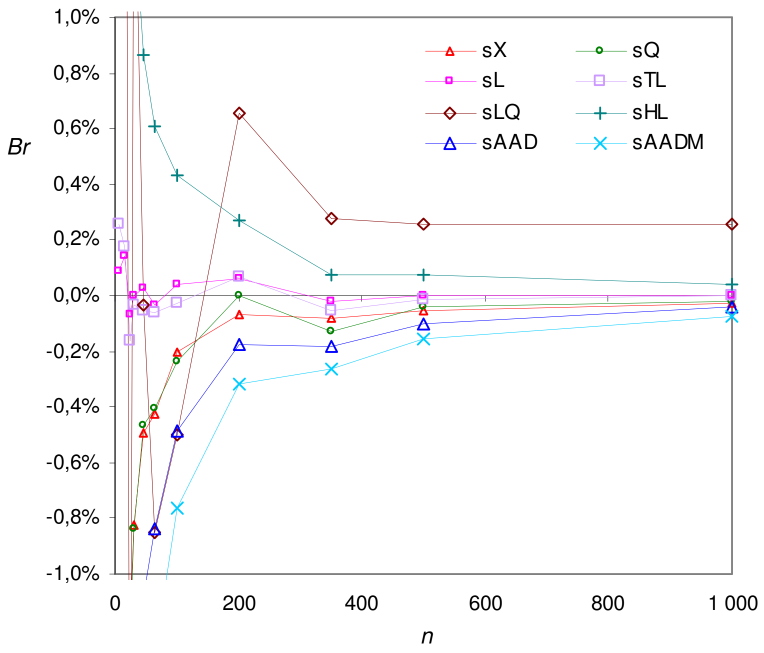
Při vzájemném porovnání výběrových vlastností estimátorů měr založených na kvantilech (s_Q a s_{QE}) docházíme ke stejným výsledkům. Zároveň docházíme k obdobným výsledkům při porovnání estimátorů měr založených na absolutní odchylce (s_{AAD} a s_{AADM}). Z důvodu přehlednosti proto v textu i ve většině grafů uvádím pouze s_Q resp. s_{AAD}, přičemž, není-li uvedeno jinak, stejné výsledky platí pro s_{QE}, resp. s_{AADM}.

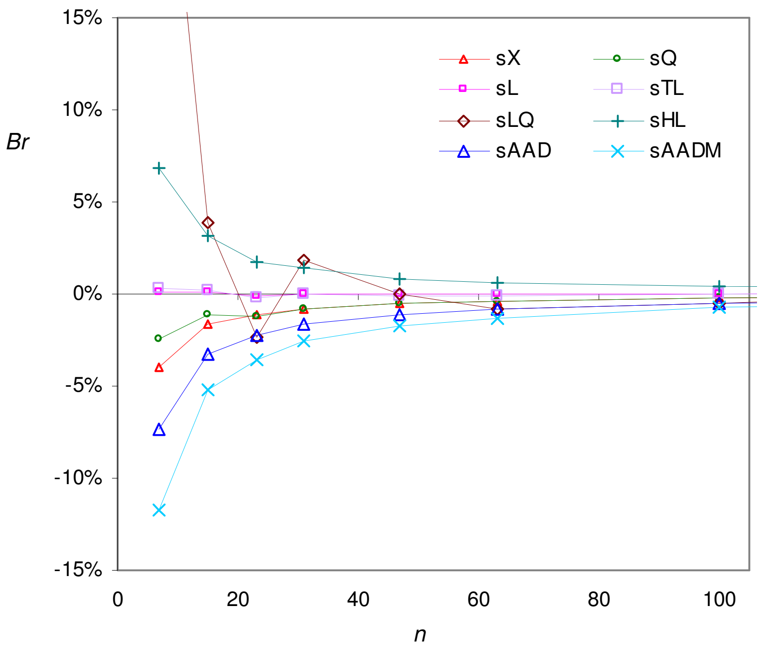
6.1.1 Výběry z normálního rozdělení

6.1.1.1 Variabilita a vychýlení

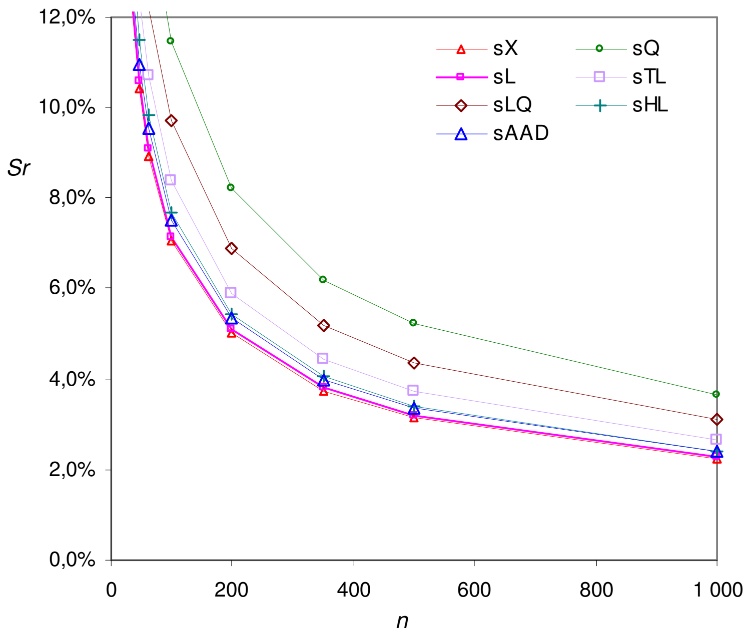
Porovnání z hlediska vychýlení

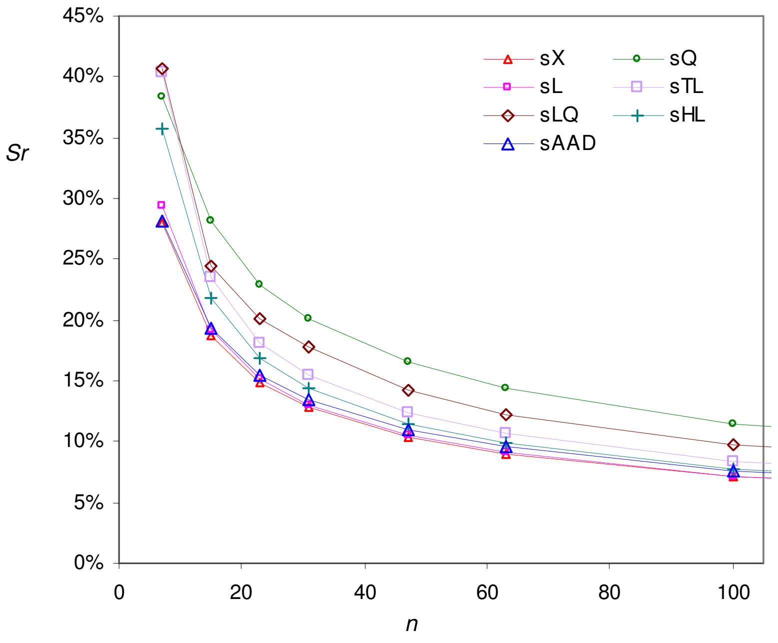
Při porovnání estimátorů z hlediska vychýlení při výběrech z normálního rozdělení dostáváme následující pořadí: nejnižší úroveň relativního vychýlení lze pozorovat shodně u s_L a s_{TL}, střední úroveň u s_Q, s_{ADD}, s_X, s_{HL} a s_{ADDM}. Nejvyšší úroveň vychýlení lze pozorovat u s_{LQ}, u něhož je konvergence vychýlení k nule nejpomalejší ze všech estimátorů. Za zmínku stojí nestandardní průběh vychýlení u estimátoru s_{LQ} pro nižší rozsahy výběru: tento estimátor pro některé rozsahy výběru podhodnocuje (n = 23, 47, 63, 100), pro jiné nadhodnocuje (n = 7, 15, 31, 200 a vyšší).



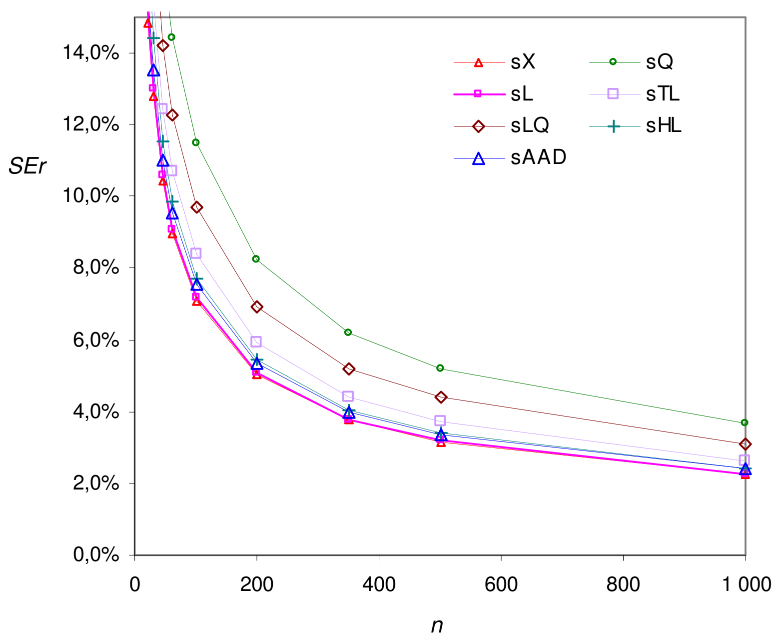


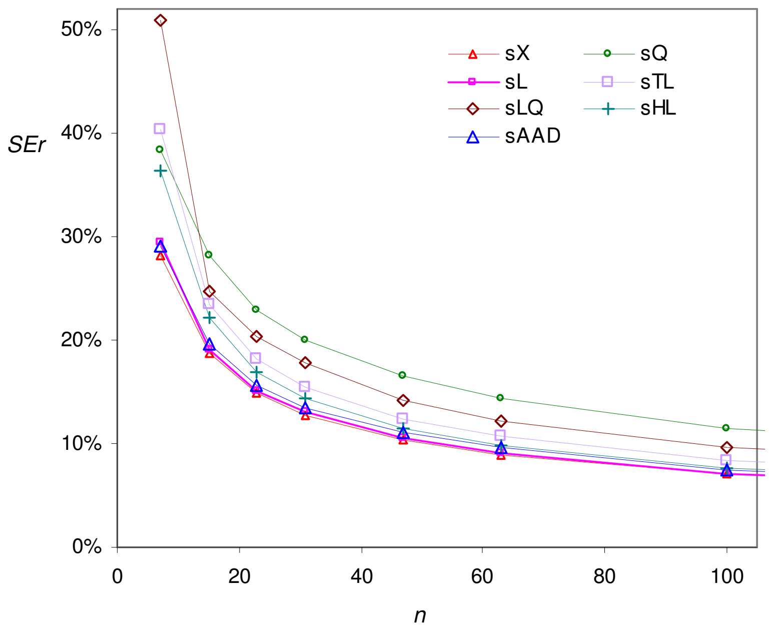
Obr. 6.1: Průběh odhadnutého relativního vychýlení estimátorů měr variability při výběrech z normálního rozdělení. (V horní části obrázku je zobrazen průběh pro všechny sledované rozsahy výběru *n*, ve spodní části je detail pro výběry o rozsahu $n \leq 100$).





Obr. 6.2: Průběh odhadnuté relativní směrodatné odchylky výběrových rozdělení estimátorů měr variability při výběrech z normálního rozdělení.





Obr. 6.3: Průběh odhadnuté relativní směrodatné chyby estimátorů měr variability při výběrech z normálního rozdělení.

Porovnání z hlediska variability

Při porovnání estimátorů z hlediska relativní variability při výběrech z normálního rozdělení dostáváme následující výsledek: nejnižší relativní směrodatnou odchylku lze pozorovat shodně u s_L a s_X. Střední úroveň variability lze pozorovat u s_{ADD}, s_{HL}, s_{TL} a s_{LQ} a nejvyšší u kvantilových měr.

Porovnání z hlediska vychýlení a variability současně

Při porovnání estimátorů z hlediska relativní směrodatné chyby při výběrech z normálního rozdělení dostáváme obdobný výsledek jako při jejich porovnání z hlediska relativní variability: nejnižší úroveň relativní směrodatné chyby lze pozorovat shodně u s_L a s_X, střední úroveň u s_{ADD}, s_{HL}, s_{TL} a s_{LQ} a nejvyšší u s_Q. Přestože estimátor s_{ADDM} vykazuje mírně vyšší vychýlení oproti estimátoru s_{ADD} (viz obr. 6.1), v průběhu relativní směrodatné chyby se tento rozdíl téměř neprojevuje a obě křivky by v grafu 6.3 splývaly.

6.1.1.2 Tvar výběrového rozdělení

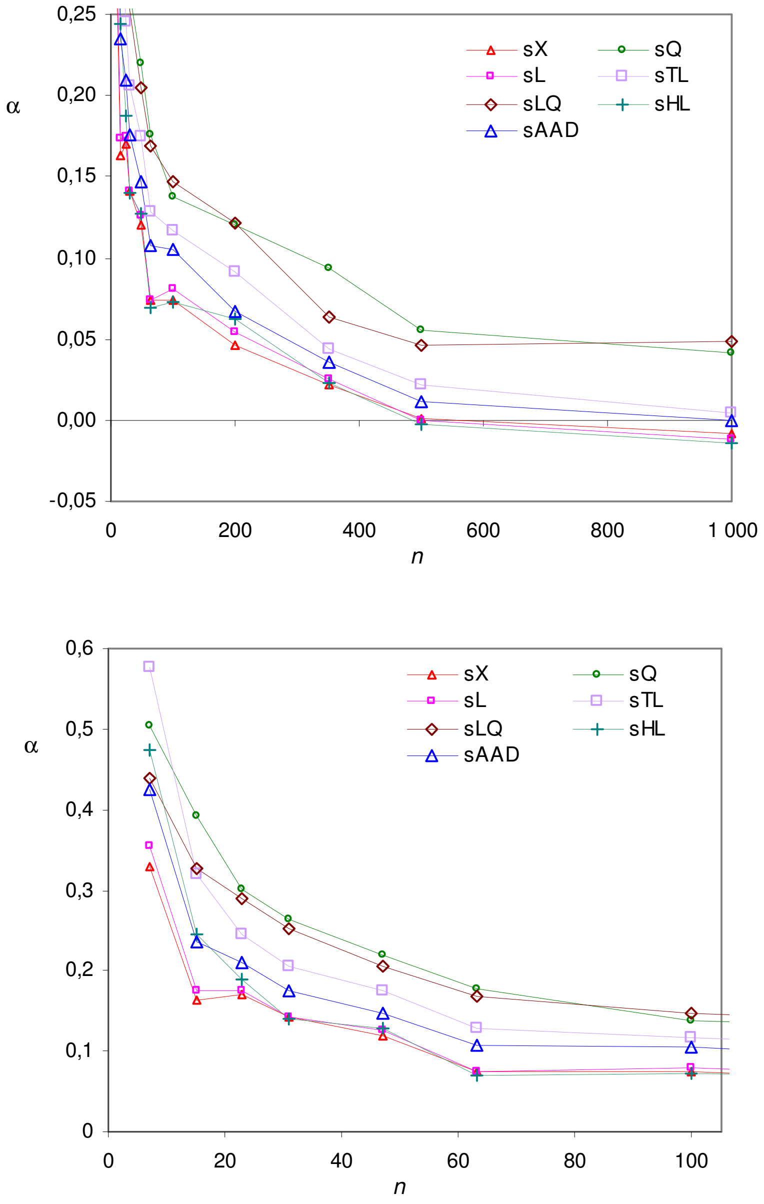
Porovnání z hlediska šikmosti a špičatosti výběrových rozdělení

Průběh odhadnutého koeficientu šikmosti výběrových rozdělení estimátorů v závislosti na rozsahu výběru je zobrazen v grafu na obr. 6.4. Je z něj patrné, že všechny estimátory mají výběrová rozdělení mírně kladně zešikmená, přičemž míra šikmosti s rozsahem výběru postupně klesá. Nejblíže k symetrickému rozdělení mají estimátory s_X a s_L, naopak nejvyšší zešikmení lze pozorovat u měr kvantilových a s_{LQ}. Z grafu na obrázku 6.5 je patrné, že hodnota odhadnutého koeficientu špičatosti v závislosti na rozsahu výběru konverguje u všech estimátorů velmi rychle k normální hodnotě 3.

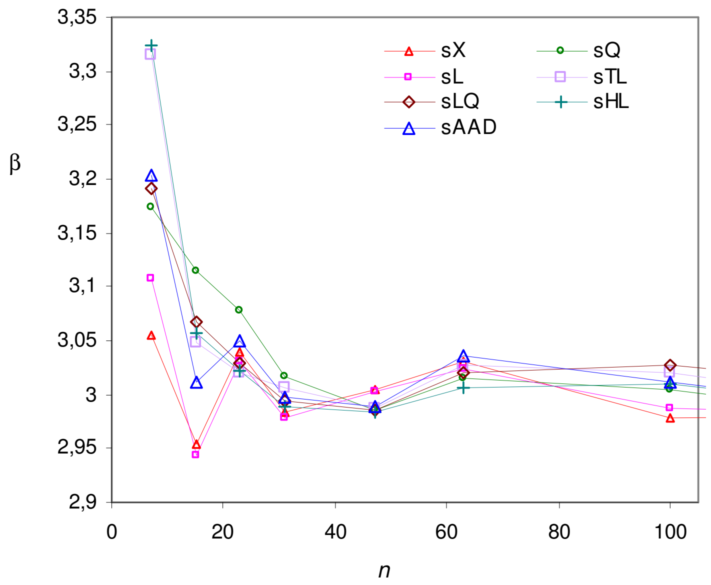
Otestování normality výběrových rozdělení

Výsledky testů Kolmogorova – Smirnova o normalitě výběrových rozdělení estimátorů jsou uvedeny v tabulce 6.1. Vyplývá z nich, že nejrychleji konvergují k normalitě

výběrová rozdělení estimátorů s_X, s_L a s_{HL}, nejpomaleji naopak výběrová rozdělení estimátorů s_Q a s_{LQ}.



Obr. 6.4: Průběh odhadnutého koeficientu šikmosti výběrových rozdělení estimátorů měr variability při výběrech z normálního rozdělení.



Obr. 6.5: Průběh odhadnutého koeficientu špičatosti výběrových rozdělení estimátorů měr variability při výběrech z normálního rozdělení pro rozsahy výběru $n \leq 100$.

Tabulka 6.1: Výsledky testů Kolmogorova – Smirnova o normalitě výběrových rozdělení estimátorů měr variability při výběrech z normálního rozdělení. Testy byly realizovány na hladinách významnosti 10 %, 5 % a 1 %, v tabulce je uvedena vždy nejvyšší hladina významnosti, na které výběrové rozdělení testu vyhovělo.

n	s_X	s_Q	s_L	s_{TL}	s_{LQ}	s_{HL}	s_{AAD}
- 23	-	-	-	-	-	-	-
31	1 %	-	1 %	-	-	1 %	-
47	1 %	-	1 %	-	-	1 %	-
63	5 %	-	1 %	-	-	1 %	1 %
100	5 %	-	1 %	1 %	-	1 %	1 %
200	10 %	1 %	10 %	10 %	-	10 %	10 %
350	10 %	5 %	10 %	10 %	10 %	10 %	10 %
500 +	10 %	10 %	10 %	10 %	10 %	10 %	10 %

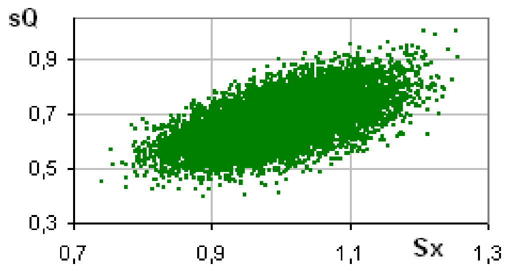
6.1.1.3 Měří všechny charakteristiky stejnou vlastnost rozdělení?

Hodnoty korelačních koeficientů vyjadřující sílu lineární závislost mezi směrodatnou odchylkou a ostatními estimátory pro výběry z normálního rozdělení o rozsahu 100 jsou uvedeny v tabulce 6.2. Stejné hodnoty korelačních koeficientů (s přesností +/- 0,03) byly získány také pro všechny ostatní rozsahy výběru pro $n \geq 23$. Na obrázku 6.6 až 6.9 jsou uvedeny bodové diagramy znázorňující průběh závislosti mezi hodnotami jednotlivých estimátorů, opět pro výběry o rozsahu $n = 100$ z normálního rozdělení, přičemž opět platí, že obdobný průběh závislosti můžeme pozorovat také pro všechny ostatní rozsahy výběru sledované v tomto experimentu.

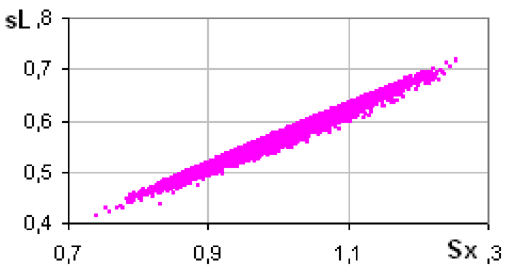
Z bodových grafů i z tabulky 6.2 je zřejmé, že nejsilnější závislost s hodnotami s_X je u estimátoru s_L. Vysoká míra závislosti s hodnotami s_X je dále patrná u estimátorů s_{AAD} a s_{HL}, nižší u s_{LQ} a nejnižší u s_Q.

Tabulka 6.2: Hodnoty korelačních koeficientů mezi estimátory s_X a s_Q, s_L, s_{TL}, s_{LQ}, s_{HL} a s_{AAD} při výběrech z normálního rozdělení o rozsahu $n = 100$.

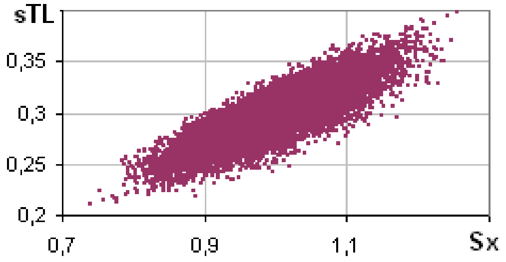
Estimátor	Korelační koeficient
s_Q	0,62
s_L	0,99
s_{TL}	0,85
s_{LQ}	0,73
s_{HL}	0,92
s_{AAD}, s_{AADM}	0,94



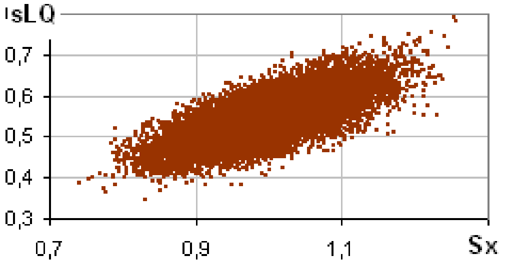
Obr. 6.6: Průběh závislosti hodnot estimátorů variability s_Q a s_X při výběrech z normálního rozdělení o rozsahu n = 100.



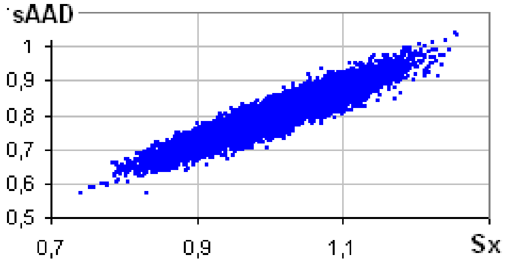
Obr. 6.7: Průběh závislosti hodnot estimátorů variability s_L a s_X při výběrech z normálního rozdělení o rozsahu n = 100.



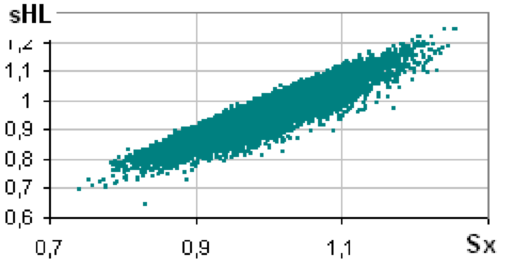
Obr. 6.8: Průběh závislosti hodnot estimátorů variability s_{TL} a s_X při výběrech z normálního rozdělení o rozsahu n = 100.



Obr. 6.9: Průběh závislosti hodnot estimátorů variability s_{LQ} a s_X při výběrech z normálního rozdělení o rozsahu n = 100.



Obr. 6.10: Průběh závislosti hodnot estimátorů variability s_{AAD} a s_X při výběrech z normálního rozdělení o rozsahu n = 100.



Obr. 6.11: Průběh závislosti hodnot estimátorů variability s_{HL} a s_X při výběrech z normálního rozdělení o rozsahu n = 100.

6.1.1.4 Verifikace dosavadních poznatků o chování estimátorů

Porovnání hodnot směrodatné chyby, vychýlení a koeficientů šikmosti a špičatosti výběrového rozdělení estimátoru s_X, získaných na základě experimentu, s teoretickými hodnotami vypočtenými podle vztahů uvedených v kapitole 2 je uvedeno v tabulce 6.3. Je z něj patrné, že hodnoty směrodatné chyby, vychýlení i koeficientu špičatosti získané pomocí experimentu se přibližně shodují s teoretickými hodnotami. Lze však pozorovat pomalejší konvergenci k normální hodnotě koeficientu šikmosti oproti analyticky odvozeným výsledkům.

Tabulka 6.3: Teoretické a empirické hodnoty směrodatné chyby, vychýlení a koeficientů šikmosti a špičatosti výběrového rozdělení estimátoru s_X při výběrech z normovaného normálního rozdělení.

n	**Teoretické hodnoty (viz kap. 2)**				**Empirické hodnoty**			
	$SE_T(s_X)$	**$B_T(s_X)$**	**$\alpha_T(s_X)$**	**$\beta_T(s_X)$**	**$SE(s_X)$**	**$B(s_X)$**	**$\alpha(s_X)$**	**$\beta(s_X)$**
7	0,2673	-0,0357	0,07	3,21	0,2824	-0,0399	0,33	3,06
15	0,1826	-0,0167	0,03	3,10	0,1872	-0,0165	0,16	2,95
23	0,1474	-0,0109	0,02	3,07	0,1485	-0,0117	0,17	3,04
31	0,1270	-0,0081	0,02	3,05	0,1279	-0,0083	0,14	2,98
47	0,1031	-0,0053	0,01	3,03	0,1043	-0,0049	0,12	3,01
63	0,0891	-0,0040	0,01	3,02	0,0896	-0,0042	0,07	3,03
100	0,0707	-0,0025	0,01	3,02	0,0707	-0,0020	0,07	2,98
200	0,0500	-0,0013	0,00	3,01	0,0503	-0,0007	0,05	2,98
350	0,0378	-0,0007	0,00	3,00	0,0376	-0,0008	0,02	3,02
500	0,0316	-0,0005	0,00	3,00	0,0314	-0,0005	0,00	3,01
1000	0,0224	-0,0003	0,00	3,00	0,0224	-0,0003	-0,01	3,04

Porovnání empiricky získaných hodnot směrodatné chyby výběrových rozdělení estimátorů s_X, s_L, s_Q, s_{AAD}, s_{AADM} s teoretickými hodnotami je uvedeno v tabulce 6.4. Z údajů vyplývá, že teoreticky odvozené vztahy platí u estimátorů s_L, s_{AAD} a s_{AADM} již pro $n = 7$, u estimátoru s_X pro $n \geq 15$ a u estimátoru s_Q přibližně až pro $n \geq 47$.

Tabulka 6.4: Rozdíl empiricky získané a teoretické hodnoty směrodatné chyby estimátorů měr variability vyjádřený jako procentní podíl z teoretické hodnoty pro výběry z normálního rozdělení:

n	$\frac{SE(s_X)}{SE_T(s_X)}-1$	$\frac{SE(s_Q)}{SE_T(s_Q)}-1$	$\frac{SE(s_L)}{SE_T(s_L)}-1$	$\frac{SE(s_{AAD})}{SE_T(s_{AAD})}-1$	$\frac{SE(s_{AADM})}{SE_T(s_{AADM})}-1$
7	6 %	-13 %	-1 %	2 %	3 %
15	3 %	-7 %	0 %	1 %	1 %
23	1 %	-6 %	-1 %	0 %	0 %
31	1 %	-4 %	-1 %	0 %	0 %
47	1 %	-3 %	0 %	0 %	0 %
63	1 %	-2 %	0 %	1 %	1 %
100	0 %	-2 %	0 %	0 %	0 %
200+	0 %	0 %	0 %	0 %	0 %

6.1.2 Výběry z ostatních rozdělení

6.1.2.1 Porovnání estimátorů z hlediska variability, vychýlení a tvaru rozdělení

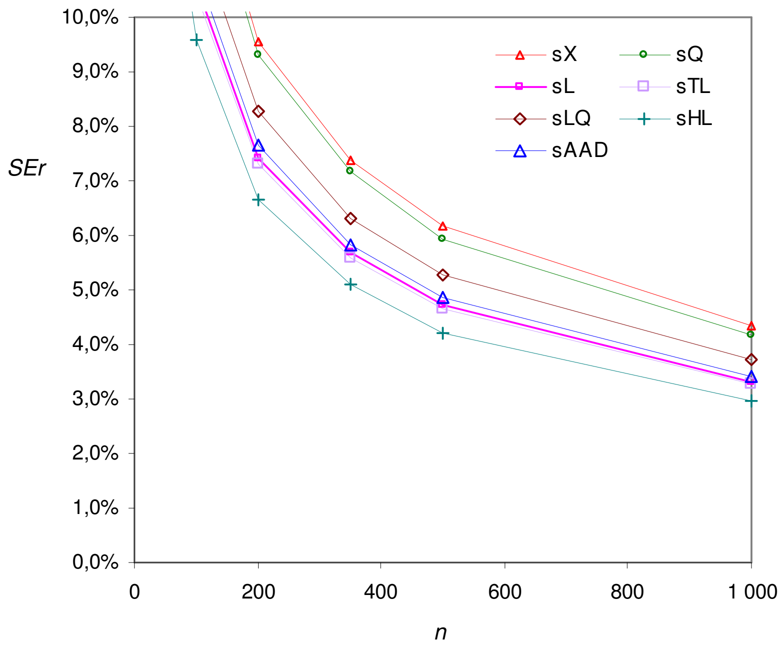
Při porovnání chování jednotlivých estimátorů v závislosti na typu rozdělení, z něhož výběr pochází, se ukazuje, že jejich výběrové vlastnosti závisejí na velikosti kurtozy v tomto rozdělení vzhledem k rozdělení normálnímu. Obdobné vlastnosti tedy mezi sebou vykazují vždy rozdělení s normální hodnotou kurtozy, s nízkou hodnotou kurtozy a vysokou hodnotou kurtozy.

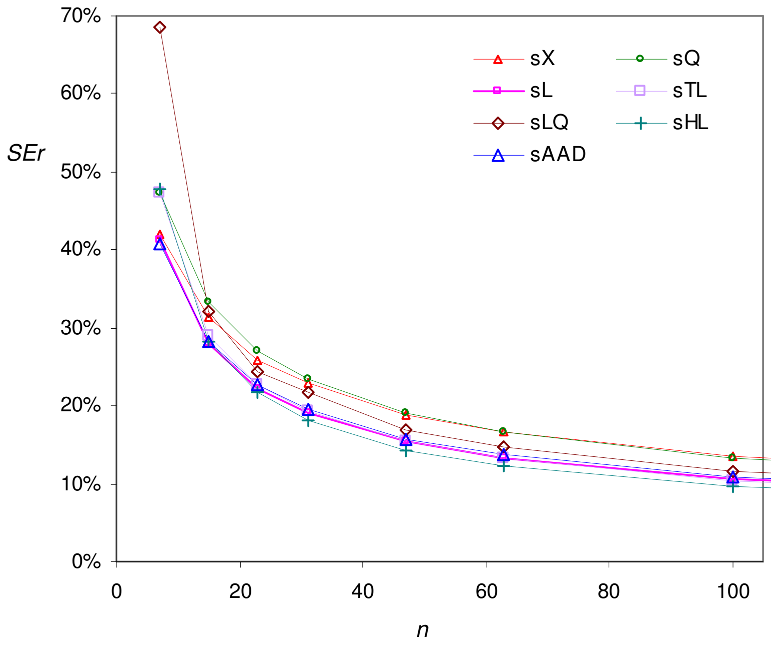
Pro **populace s přibližně normální nebo nižší než normální hodnotou kurtozy** jsou výsledky porovnání estimátorů z hlediska vychýlení, variability a konvergence k normalitě stejné, jako v případě výběrů z normálního rozdělení.

Při výběrech z **populací s vyšší než normální hodnotou kurtozy** byly zjištěny následující odlišnosti od výběrových vlastností estimátorů v případě výběrů z normálního rozdělení: **Směrodatná chyba** estimátoru s_X se se vrůstající úrovní kurtozy zvyšuje, zatímco pořadí ostatních estimátorů z hlediska relativní směrodatné chyby zůstává přibližně zachováno. Nejnižší relativní směrodatná chyba je tedy pozorována u estimátoru druhého L-momentu s_L. Výběrová rozdělení všech estimátorů se při

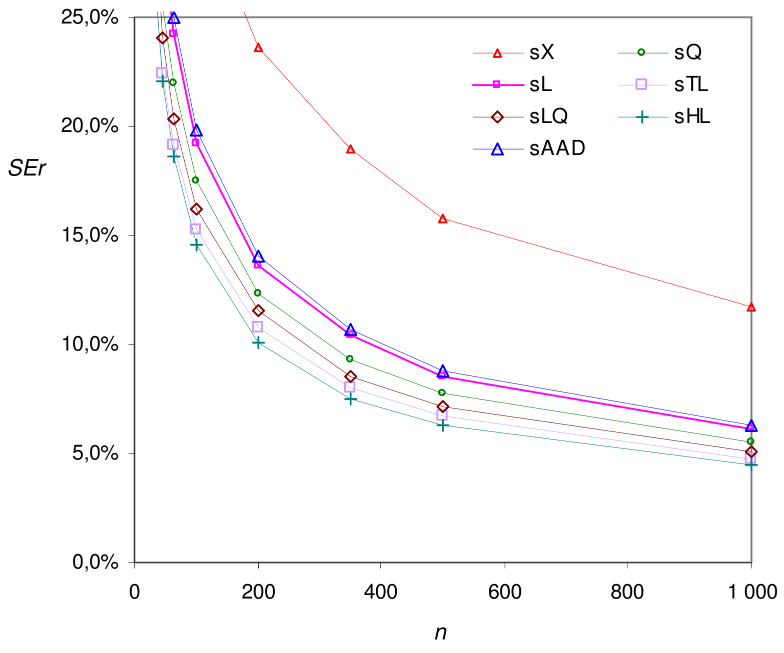
vzrůstající úrovni kurtozy kladně zešikmují a jejich konvergence k normalitě se zpomaluje. Např. při výběrech z Laplaceova rozdělení s hodnotou $\beta_M = 6$ vyhovují výběrová rozdělení estimátorů testu normality až pro výběry o rozsahu n od 200, v některých případech až od $n = 500$.

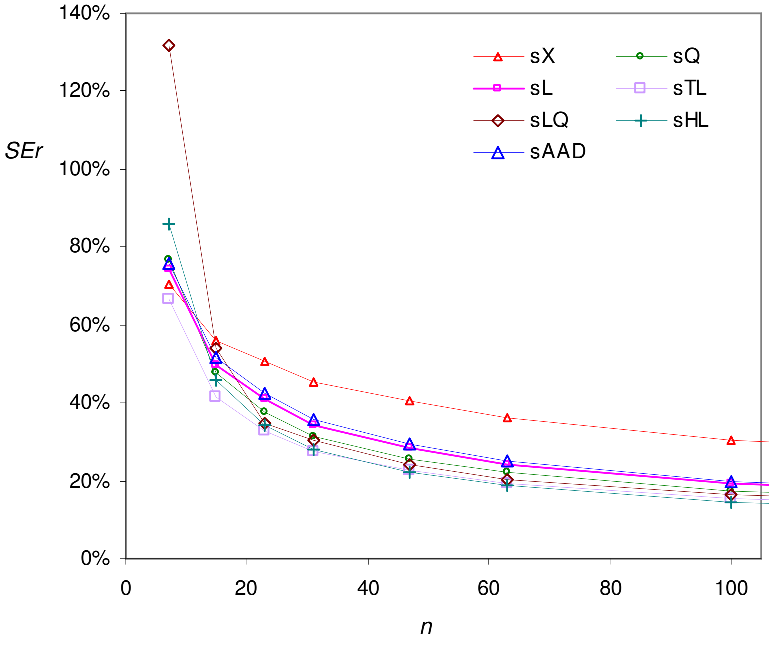
Pro **populace výrazně vyšší než normální hodnotou kurtozy** dosahuje nejvyššího relativního vychýlení i relativní směrodatné chyby estimátor s_X (estimátor podhodnocuje). Zároveň dochází k mírnému zvýšení relativní směrodatné chyby estimátoru s_L ve srovnání s ostatními estimátory. Nejlepšími estimátory z hlediska relativní směrodatné chyby jsou v tomto případě s_{TL}, s_{LQ} a s_{HL}; estimátor s_L společně s s_{AAD} vykazuje střední úroveň relativní směrodatné chyby. V těchto případech zároveň žádný z estimátorů nevyhovuje testu o normalitě výběrového rozdělení ani pro $n = 1000$. Tento výsledek zároveň potvrzuje obecně známou skutečnost, že na hodnotě koeficientu špičatosti v populaci závisí variabilita výběrového rozptylu a tedy i výběrové směrodatné odchylky (viz např. ČERMÁK, VRABEC 1999).





Obr. 6.12: Průběh odhadnuté relativní směrodatné chyby estimátorů měr variability při výběrech z logaritmicko-normálního rozdělení LN(c = 0,5) s hodnotou β_M = 8,89.





Obr. 6.13: Průběh odhadnuté relativní směrodatné chyby estimátorů měr variability při výběrech z logaritmicko-normálního rozdělení LN(c = 1) s hodnotou β_M = 113,9.

6.1.2.2 Měří všechny charakteristiky stejnou vlastnost rozdělení?

Síla závislosti mezi s_X a ostatními estimátory se mění s hodnotou špičatosti základního souboru následujícím způsobem:

Pro **populace s přibližně normální hodnotou kurtozy** jsou výsledky porovnání korelačních koeficientů shodné jako v případě výběrů z normálního rozdělení. Pro **populace s nižší než normální hodnotou kurtozy** lze vypozorovat, že s poklesem hodnoty koeficientu špičatosti se hodnoty korelačních koeficientů s s_X u všech estimátorů zvyšují. Jako příklad jsou v tabulce 6.5 uvedeny hodnoty korelačních koeficientů jednotlivých estimátorů s s_X pro výběry z rovnoměrného rozdělení (β_M = 1,8).

Tabulka 6.5: Hodnoty korelačních koeficientů mezi estimátorem s_X a estimátory s_Q, s_L, s_{TL}, s_{LQ}, s_{HL}, s_{AAD} a s_{AADM} při výběrech z rovnoměrného rozdělení o rozsahu n = 100 (β_M = 1,8).

Estimátor	**Korelační koeficient**
s_Q	0,84
s_L	1,00
s_{TL}	0,97
s_{LQ}	0,90
s_{HL}	0,97
s_{AAD}, s_{AADM}	0,97

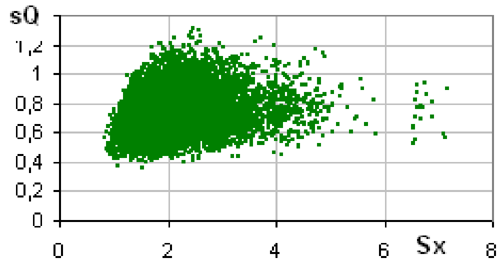
U **populací s vyšší než normální hodnotou kurtozy** lze pozorovat, že s rostoucí mírou kurtozy hodnoty korelačních koeficientů mezi s_X a ostatními estimátory klesají. Tento pokles přitom není u všech estimátorů stejný: k nejvyššímu poklesu dochází u estimátorů s_Q a s_{LQ}, dále u s_{HL} a s_{TL}. Naopak u estimátorů s_L, s_{AAD} a s_{AADM} je pokles korelačního koeficientu téměř zanedbatelný. V tabulce 6.6 jsou uvedeny hodnoty korelačních koeficientů pro výběry z Laplaceova rozdělení (β_M = 6). Je z nich patrné, že zatímco u estimátoru s_L poklesl korelační koeficient ve srovnání s výběry z normálního rozdělení přibližně o jednu třetinu, u estimátorů s_L resp. s_{AAD} poklesl pouze o 3 % resp. 4 %

Tabulka 6.6: Hodnoty korelačních koeficientů mezi estimátorem s_X a estimátory s_Q, s_L, s_{TL}, s_{LQ}, s_{HL}, s_{AAD} a s_{AADM} při výběrech z Laplaceova rozdělení o rozsahu n = 100 (β_M = 6).

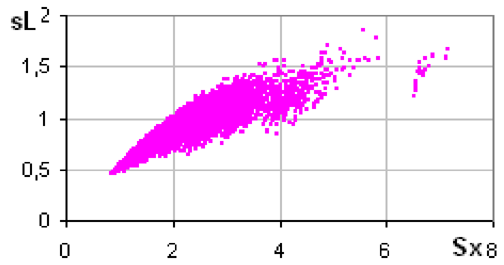
Estimátor	**Korelační koeficient**
s_Q	0,43
s_L	0,96
s_{TL}	0,71
s_{LQ}	0,56
s_{HL}	0,72
s_{AAD}, s_{AADM}	0,90

V případě populací **s extrémně vysokou hodnotu kurtozy** síla závislosti mezi s_X u většiny estimátorů dále klesá. Výjimku tvoří pouze estimátory s_L, s_{AAD} a s_{AADM}, u

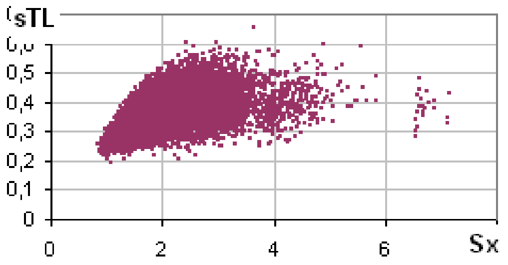
kterých je i v tomto případě patrná silná lineární závislost s hodnotami estimátoru s_X. Jako příklad uvádím hodnoty korelačních koeficientů a bodové grafy pro výběry z logaritmicko-normálního rozdělení LN(c = 1) s hodnotou β_M = 113,9.



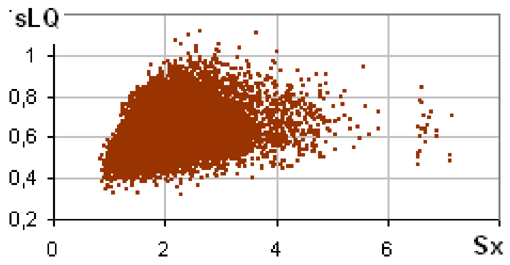
Obr. 6.14: Průběh závislosti hodnot estimátorů variability s_Q a s_X při výběrech z rozdělení o rozsahu LN(1) o rozsahu n = 100 (β_M = 113,9).



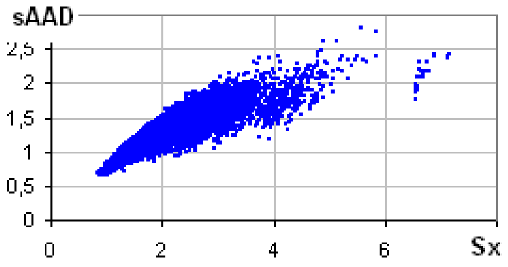
Obr. 6.15: Průběh závislosti hodnot estimátorů variability s_L a s_X při výběrech z rozdělení LN(1) o rozsahu n = 100.



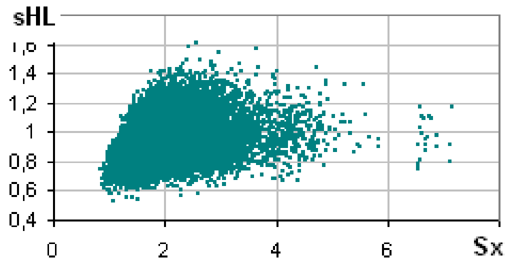
Obr. 6.16: Průběh závislosti hodnot estimátorů variability s_{TL} a s_X při výběrech z rozdělení LN(1) o rozsahu n = 100.



Obr. 6.17: Průběh závislosti hodnot estimátorů variability s_{LQ} a s_X při výběrech z rozdělení LN(1) o rozsahu n = 100.



Obr. 6.18: Průběh závislosti hodnot estimátorů variability s_{AAD} a s_X při výběrech z rozdělení LN(1) o rozsahu n = 100.



Obr. 6.19: Průběh závislosti hodnot estimátorů variability s_{HL} a s_X při výběrech z rozdělení LN(1) o rozsahu n = 100.

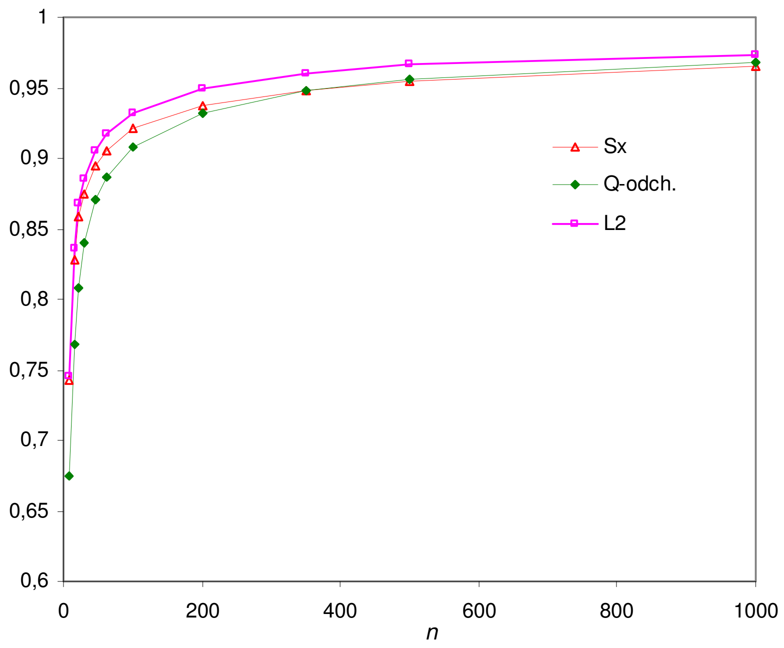
Tabulka 6.7: Hodnoty korelačních koeficientů mezi estimátorem s_X a estimátory s_Q, s_L, s_{TL}, s_{LQ}, s_{HL}, s_{AAD} a s_{AADM} při výběrech z logaritmicko-normálního rozdělení LN(c = 1) o rozsahu n = 100.

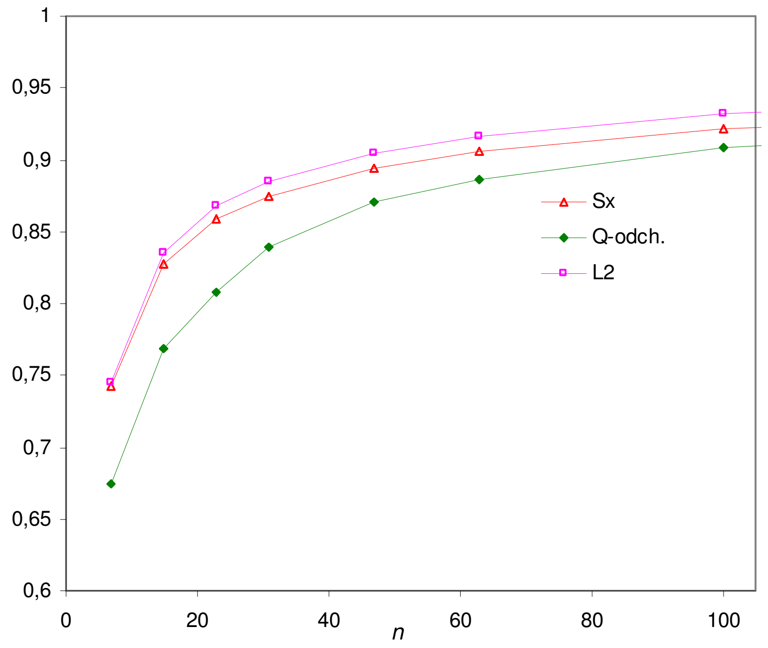
Estimátor	Korelační koeficient
s_Q	0,24
s_L	0,89
s_{TL}	0,46
s_{LQ}	0,32
s_{HL}	0,30
s_{AAD}	0,87
s_{AADM}	0,83

6.1.2.3 Síla testu

Hodnoty průměrné síly testů o rozdělení souboru, z něhož výběr pochází, jsou uvedeny na obr. 6.20. Je z něj patrné, že pro nejlepších výsledků dosahuje pro všechny sledované rozsahy výběru estimátor s_L. Naopak nejhorších výsledků dosahuje pro nižší rozsahy výběru estimátor s_Q, pro vyšší rozsahy výběru společně estimátory s_Q a s_X. Tento výsledek také můžeme interpretovat tak, že nejlepší schopnost od sebe odlišit soubory s vysokou a nízkou úrovní variability má estimátor s_L.

Je však třeba podotknout, že číselné hodnoty průměrné síly testů jsou výrazně ovlivněny výběrem populací zařazených do tohoto experimentu. Je například zřejmé, že pokud bychom do experimentu zařadili více populací podobných rozdělení normálnímu, získali bychom u všech estimátorů nižší hodnoty průměrné síly testu a naopak. Zde prezentované výsledky lze tedy posuzovat pouze **z hlediska pořadí**, nikoliv z hlediska absolutních hodnot.





Obr. 6.20: Průměrná síla testů o populaci, z níž výběr pochází, založených na estimátorech s_X, s_L, a s_Q. (V horní části obrázku je zobrazen průběh pro všechny sledované rozsahy výběru *n*, ve spodní části je detail pro výběry o rozsahu $n \leq 100$).

6.1.2.4 Verifikace dosavadních poznatků o chování estimátorů

Při porovnání teoretických hodnot směrodatných chyb, vypočtených dle vztahů z kapitoly 2, s hodnotami získanými na základě experimentu se ukazuje následující: v případě estimátoru s_X lze pozorovat, že u populací s vyšší než normální hodnotou kurtozy je empiricky zjištěná směrodatná chyba nižší než teoretická, u populací s nižší než normální hodnotu kurtozy je tomu naopak. V případě estimátoru s_{AAD} lze pozorovat kladný rozdíl mezi empiricky získanou a teoretickou hodnotou směrodatné chyby při výběrech z nesymetrických populací pro všechny sledovaná rozsahy výběru. Naopak u estimátoru s_{AADM} lze pozorovat shodu teoretických a empiricky získaných hodnot směrodatné chyby pro výběry ze všech zkoumaných populací o rozsahu $n \geq 15$. V případě estimátoru s_X je situace stejná, jako při výběrech z normálního rozdělení.

Pro estimátor s_L nebylo možné z důvodu výpočetní náročnosti funkcionálu $\Im$ (viz kapitola 2) v celé populaci spočítat jeho teoretickou hodnotu $D_{TEOR}(s_L)$. Nahradil jsem ji proto empiricky odhadnutou střední hodnotou jejího estimátoru $est\,D(s_L)$ dle ZENGA et al. (2004), kterou jsem z důvodu výpočetní náročnosti zjišťoval pouze pro rozsahy výběru $n \leq 100$. Výsledky získané při porovnání empiricky získané $est\,D(s_L)$ ukazují, že s růstem kurtozy v populaci se rozdíl mezi $E[est\,D(s_L)]$ a $D(s_L)$ zvyšuje.

Tabulka 6.8: Rozdíl empiricky získané a teoretické hodnoty směrodatné chyby estimátorů měr variability vyjádřený jako procentní podíl z teoretické hodnoty pro výběry z logaritmicko-normálního rozdělení s parametrem c = 0,5:

N	$\frac{SE(s_X)}{SE_T(s_X)}-1$	$\frac{SE(s_Q)}{SE_T(s_Q)}-1$	$\frac{SE(s_L)}{SE_T(s_L)}-1$	$\frac{SE(s_{AAD})}{SE_T(s_{AAD})}-1$	$\frac{SE(s_{AADM})}{SE_T(s_{AADM})}-1$
7	-20 %	-7 %	31 %	19 %	-2 %
15	-13 %	-4 %	17 %	20 %	0 %
23	-10 %	-3 %	10 %	20 %	0 %
31	-8 %	-3 %	9 %	20 %	0 %
47	-7 %	-3 %	6 %	19 %	0 %
63	-5 %	-2 %	5 %	20 %	0 %
100	-3 %	-2 %	2 %	20 %	0 %
200	-2 %	-2 %	.	19 %	0 %
350	0 %	0 %	.	20 %	1 %
500	-1 %	-1 %	.	20 %	1 %
1000	-1 %	-2 %	.	19 %	0 %

6.2 Estimátory měr relativní variability

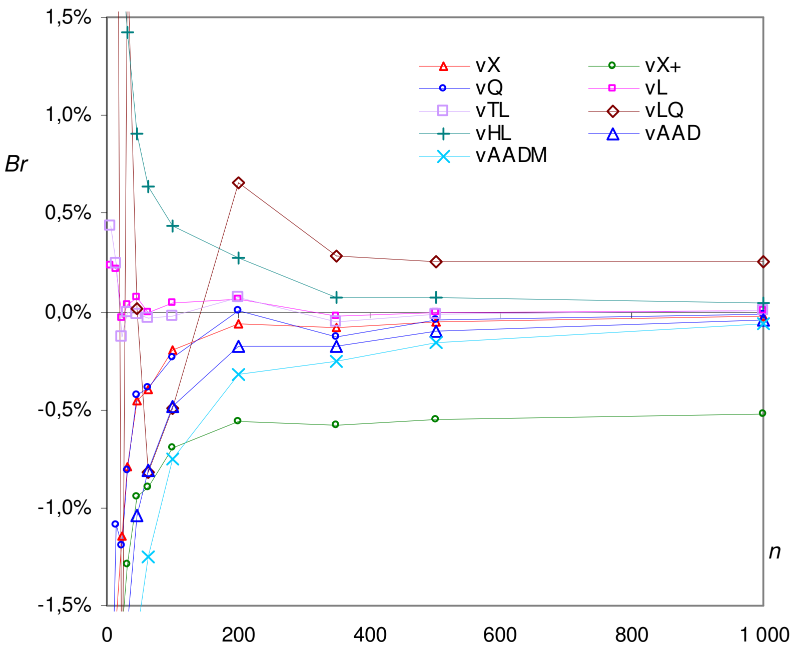
6.2.1 Výběry z normálního rozdělení

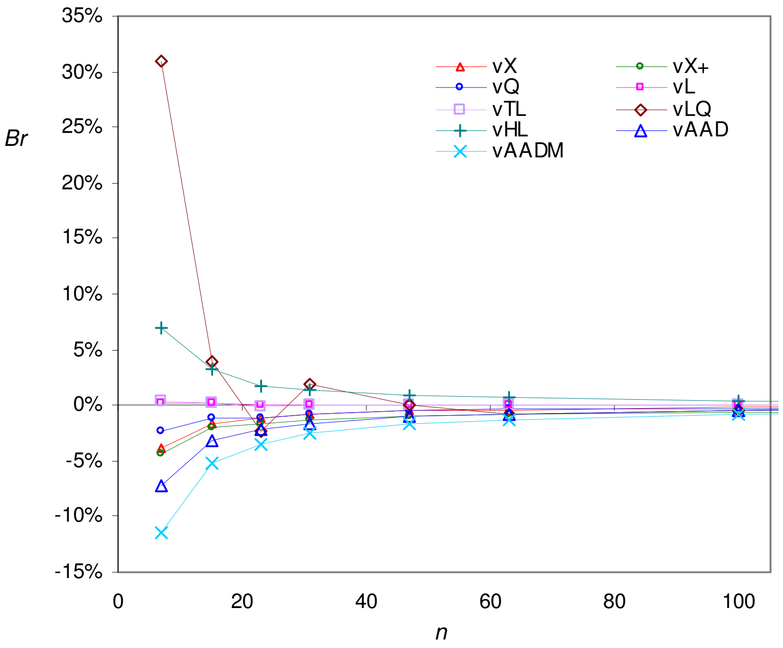
Při vzájemném porovnání estimátorů měr založených na průměrné absolutní odchylce od průměru (v_{AAD}) a od mediánu (v_{AADM}) docházíme v případě výběrů z normálního rozdělení k obdobným závěrům. Z důvodu přehlednosti proto v textu i ve většině grafů uvádím pouze první z nich.

6.2.1.1 Variabilita a vychýlení

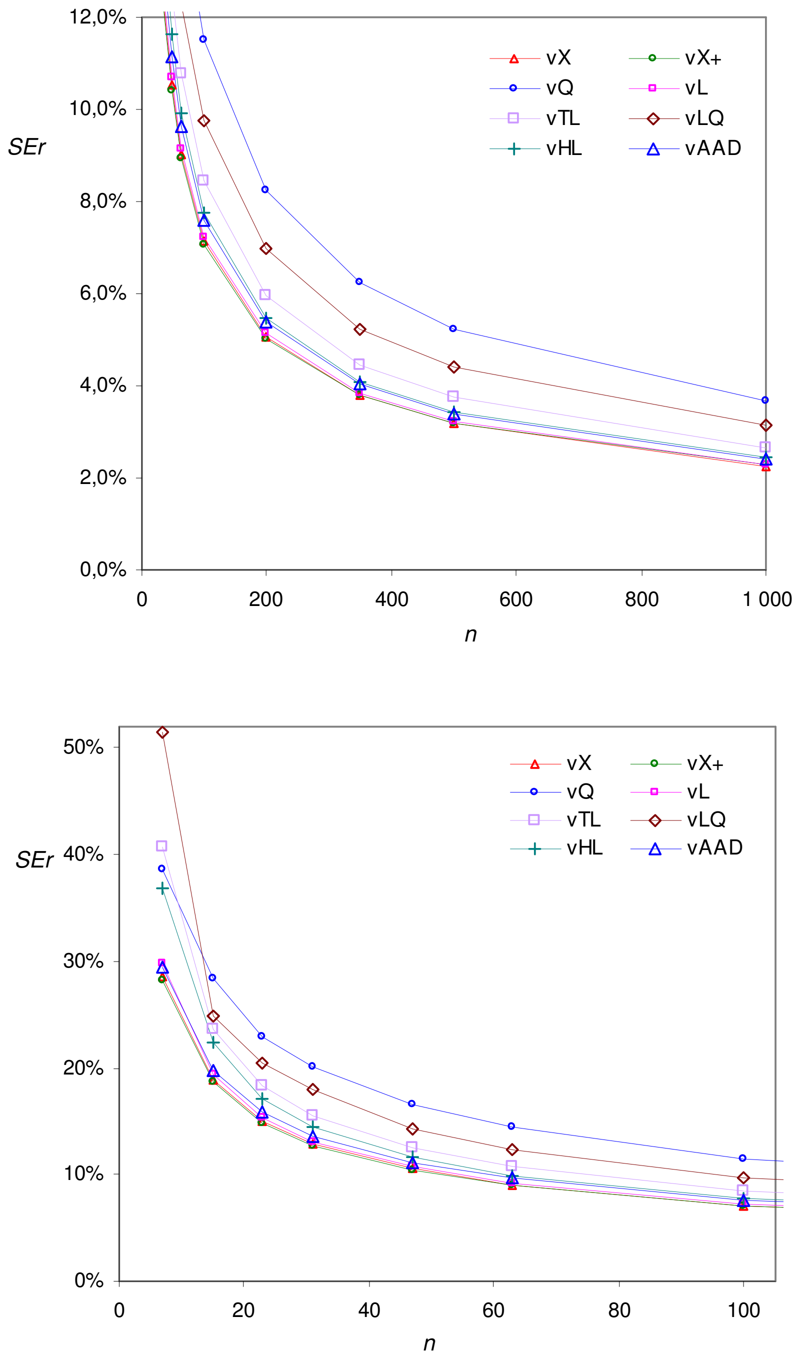
Při porovnání estimátorů měr relativní variability z hlediska vychýlení i variability při výběrech z normálního rozdělení dostáváme stejné pořadí, jaké jsme získali pro pořadí estimátorů měr absolutní variability, z nichž byly odvozeny. Nejnižší úroveň **relativního vychýlení** (obr. 6.21) lze tedy pozorovat shodně u v_L a v_{TL}, střední u v_Q, v_{ADD}, v_X a v_{HL}. Nejvyšší úroveň vychýlení lze pozorovat u v_{LQ} a u normovaného variačního koeficientu v_X^+.

Nejnižší úroveň **relativní směrodatné chyby** (obr. 6.22) lze pozorovat shodně u v_X, v_L a v_X^+. Střední úroveň lze pozorovat u v_{ADD}, v_{HL} následně u v_{TL} a v_{LQ} a nejvyšší u estimátoru kvantilové míry v_Q. Stejné pořadí dostáváme i při porovnání estimátorů z hlediska jejich relativní variability.





Obr. 6.21: Průběh odhadnutého relativního vychýlení estimátorů měr relativní variability při výběrech z normálního rozdělení.



Obr. 6.22: Průběh odhadnuté relativní směrodatné chyby estimátorů měr relativní variability při výběrech z normálního rozdělení. (Poznámka: Křivky pro estimátory v_X, v_X^+ a v_L v grafech splývají.)

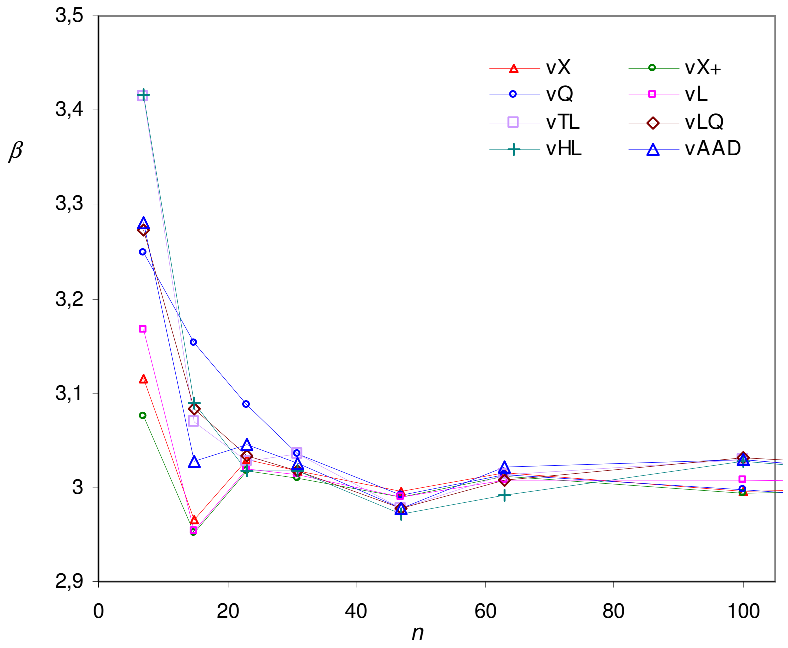
6.2.1.2 Tvar výběrového rozdělení

Porovnání z hlediska šikmosti a špičatosti výběrových rozdělení

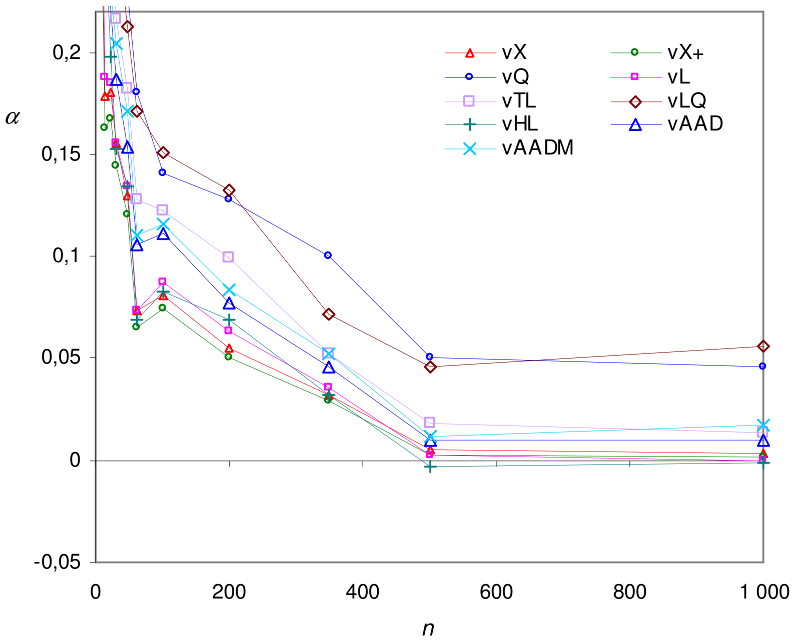
Průběh odhadnutého **koeficientu šikmosti** výběrových rozdělení estimátorů v závislosti na rozsahu výběru je zobrazen na obrázku 6.24. Je z něj patrné, že, obdobně jako v případě estimátorů měr absolutní variability, všechny estimátory mají výběrová rozdělení mírně kladně zešikmená, přičemž míra šikmosti s rozsahem výběru postupně klesá. Nejblíže k symetrickému rozdělení mají shodně estimátory v_X, v_L a v_X^+, naopak nejvyšší zešikmení lze pozorovat u měr kvantilových a v_{LQ}. Z grafu na obr. 6.23 je dále patrná rychlá konvergence **koeficientu špičatosti** výběrových rozdělení jednotlivých estimátorů v závislosti na rozsahu výběru k normální hodnotě 3.

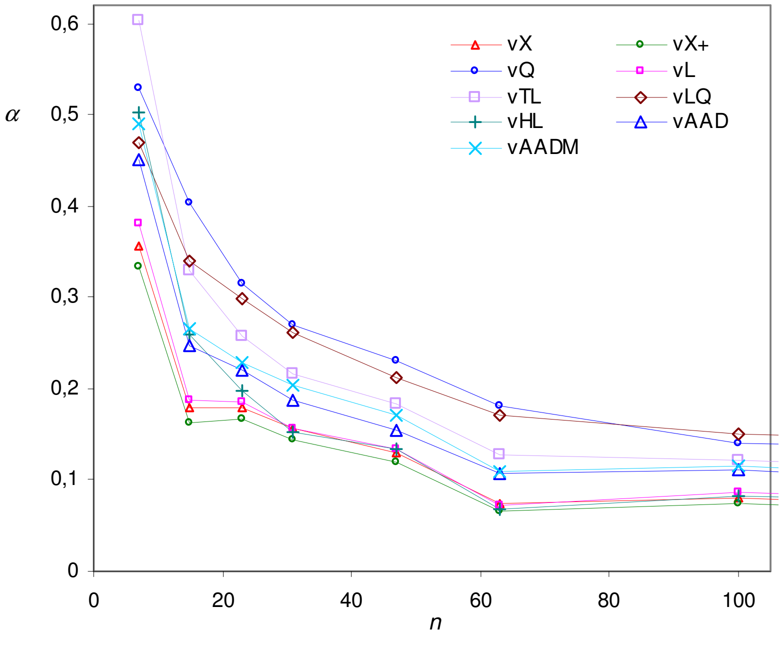
Otestování normality výběrových rozdělení

Výsledky testů Kolmogorova – Smirnova o normalitě výběrových rozdělení estimátorů jsou uvedeny v tabulce 6.9. Vyplývá z nich, že nejrychleji konvergují k normalitě výběrová rozdělení estimátorů v_X, v_L a v_X^+ a v_{HL}, nejpomaleji naopak výběrová rozdělení estimátorů v_Q a v_{LQ}.



Obr. 6.23: Průběh odhadnutého koeficientu špičatosti estimátorů měr relativní variability při výběrech z normálního rozdělení pro rozsahy výběru $n \leq 100$.





Obr. 6.24: Průběh odhadnutého koeficientu šikmosti estimátorů měr relativní variability při výběrech z normálního rozdělení.

Tabulka 6.9: Výsledky testů Kolmogorova – Smirnova o normalitě výběrových rozdělení estimátorů při výběrech z normálního rozdělení. (V tabulce je uvedena vždy nejvyšší hladina významnosti, na které výběrové rozdělení testu vyhovělo.)

n	v_X	v_X^+	v_Q	v_L	v_{TL}	v_{LQ}	v_{HL}	v_{AAD}
- 23	-	-	-	-	-	-	-	-
31	1 %	5 %	-	1 %	-	-	-	-
47	5 %	5 %	-	1 %	1 %	-	1 %	1 %
63	1 %	1 %	-	1 %	1 %	-	5 %	1 %
100	10 %	10 %	-	5 %	1 %	-	5 %	1 %
200	10 %	10 %	1 %	10 %	10 %	1 %	5 %	10 %
350 +	10 %	10 %	10 %	10 %	10 %	10 %	10 %	10 %

6.2.1.3 Měří všechny charakteristiky stejnou vlastnost rozdělení?

Hodnoty korelačních koeficientů i bodových grafů závislosti hodnotami estimátoru v_X a ostatních estimátorů měr relativní variability dávají vesměs stejné výsledky, jaké jsme dostali při posouzení závislosti mezi estimátory měr absolutní variability a s_X. Nejvyšší míra závislosti s hodnotami v_X je patrná u estimátoru v_L, nejnižší naopak u v_Q. Pro estimátor v_X^+ jsem neposuzoval sílu statistické závislosti s v_X, protože mezi hodnotami variačního koeficientu V_X a standardizovaného variačního koeficientu V_X^+ existuje závislost funkční (viz kapitola 2).

Tabulka 6.10: Hodnoty korelačních koeficientů mezi estimátory v_X a v_Q, v_L, v_{TL}, v_{LQ}, v_{HL} a v_{AAD} při výběrech z normálního rozdělení o rozsahu n = 100.

Estimátor	Korelační koeficient
v_Q	0,62
v_L	0,99
v_{TL}	0,85
v_{LQ}	0,73
v_{HL}	0,93
v_{AAD}	0,94

6.2.1.4 Verifikace dosavadních poznatků o chování estimátorů

Porovnání hodnot směrodatné chyby estimátoru v_X získaných na základě experimentu s teoretickými hodnotami vypočtenými podle vztahu uvedeného v kapitole 2 je uvedeno v tabulce 6.11. Je z ní patrné, že hodnoty směrodatné chyby získané pomocí experimentu se přibližně shodují s teoretickými hodnotami pro rozsahy výběru $n \geq 15$.

Tabulka 6.11: Rozdíl empiricky získané a teoretické hodnoty směrodatné chyby estimátoru v_X vyjádřený jako procentní podíl z teoretické hodnoty pro výběry z normálního rozdělení:

n	$\frac{SE(v_X)}{SE_T(v_X)} - 1$
7	6 %
15	2 %
23	1 %
31	1 %
47	1 %
63	0 %
100	0 %
200+	1 %

6.2.2 Výběry z ostatních rozdělení

6.2.2.1 Porovnání estimátorů z hlediska variability, vychýlení a tvaru rozdělení

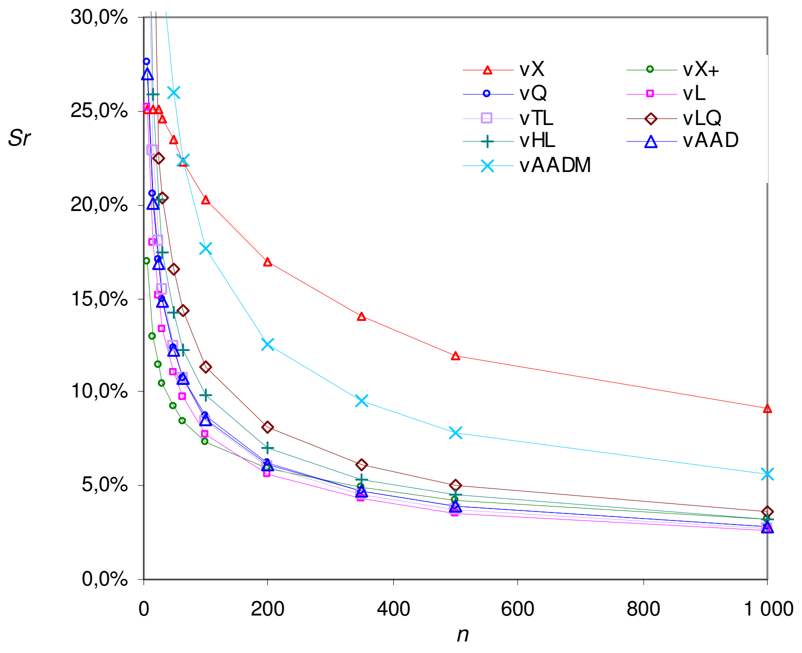
Při porovnání výběrových jednotlivých estimátorů se, stejně jako v případě estimátorů měr absolutní variability, ukazuje, že jejich výběrové vlastnosti závisejí na velikosti kurtozy v příslušné populaci vzhledem k rozdělení normálnímu.

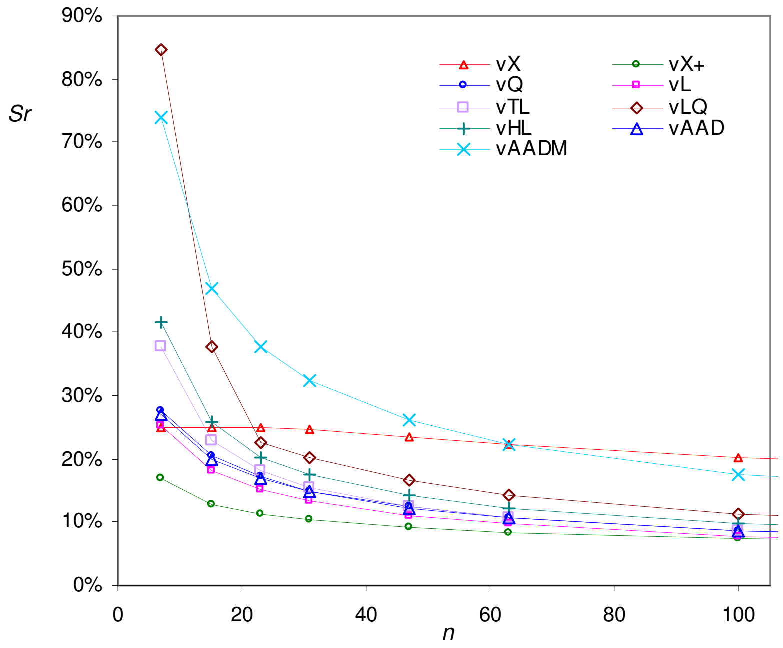
Pro **populace s přibližně normální nebo nižší než normální hodnotou kurtozy** jsou výsledky porovnání estimátorů z hlediska vychýlení, variability a konvergence k normalitě stejné, jako v případě výběrů z normálního rozdělení. Pro **populace s vyšší než normální hodnotou kurtozy** se ukazuje, že se vzrůstající úrovní kurtozy roste vychýlení i relativní směrodatná chyba estimátoru v_X (estimátor podhodnocuje). Nejlepší výsledky z hlediska relativní směrodatné chyby v těchto případech dává estimátor v_L, mírně

horší potom v_{AAD}, v_{TL} a v_X^+. Pokud jde o **konvergenci výběrových rozdělení k normalitě**, nejrychlejší je u výběrového rozdělení estimátoru v_L, které lze u většiny populací považovat za normální přibližně pro rozsah výběru $n \geq 100$. U dalších estimátorů nelze na základě experimentu učinit obecně platné závěry, protože rychlost konvergence jejich výběrových rozdělení k normalitě záleží na konkrétním základním souboru, ze kterého výběry provádíme.

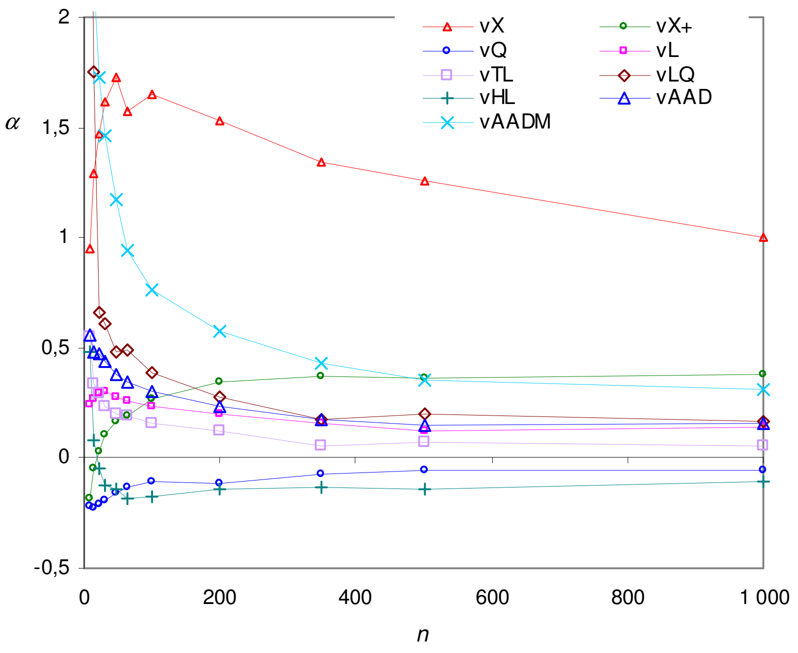
Pro populace **s extrémně vysokou hodnotou kurtozy** je třeba rozlišit situaci, kdy provádíme výběry nižšího rozsahu (přibližně pro $n \leq 63$) a vyššího rozsahu. Při výběrech nižšího rozsahu dosahuje nejvyšší relativní směrodatné odchylky i směrodatné chyby estimátor v_{AADM}, přičemž výrazný rozdíl jeho relativní směrodatné odchylky a chyby ve srovnání s estimátorem v_{AAD} lze pozorovat pro všechny sledované rozsahy výběru. Pro výběry vyššího rozsahu vykazuje nejhorší vlastnosti estimátor v_X, který dosahuje nejvyšších hodnot relativního vychýlení (podhodnocuje) i relativní směrodatné chyby. Velikost jeho relativní směrodatné chyby je pro rozsahy výběru $n \geq 100$ u těchto populací dvakrát až čtyřikrát vyšší než je její velikost ostatních estimátorů. Pokud jde o výběrové vlastnosti zbývajících estimátorů, nelze stanovit obecně platné pořadí z hlediska velikosti jejich relativní směrodatné chyby, ve většině případů však dosahují nejlepších výsledků společně estimátory v_L, v_{AAD}, v_{TL} a v_X^+. Z hlediska tvaru výběrových rozdělení ve většině případů platí, že jsou s výjimkou estimátoru v_X^+ kladně zešikmená a testu o normalitě buď nevyhovují vůbec nebo pouze pro rozsahy výběru $n = 1000$.

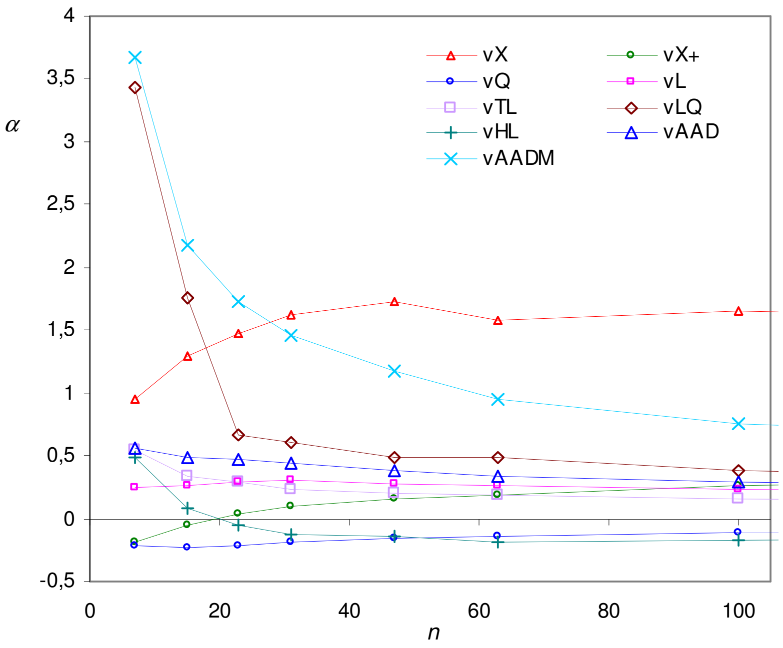
V chování estimátoru v_X lze při výběrech z populace s extrémně vysokou hodnotou kurtozy pozorovat následující anomálii: pro nižší rozsahy výběru (přibližně pro $n \leq 100$) dochází k nárůstu jeho směrodatné odchylky v závislosti na rozsahu výběru n a k nárůstu kladného zešikmení jeho výběrového rozdělení. Tato anomálie je zřejmě způsobena existencí horní hranice oboru možných hodnot tohoto estimátoru pro daný rozsah výběru $max_{v_X}(n) = \sqrt{n-1}$ pro nezápornou náhodnou proměnnou (viz kapitola 2).





Obr. 6.25: Průběh odhadnuté relativní směrodatné odchylky výběrových rozdělení estimátorů měr variability při výběrech z logaritmicko-normálního rozdělení LN(c = 1) s extrémně vysokou hodnotou kurtozy (β_M = 113,9)





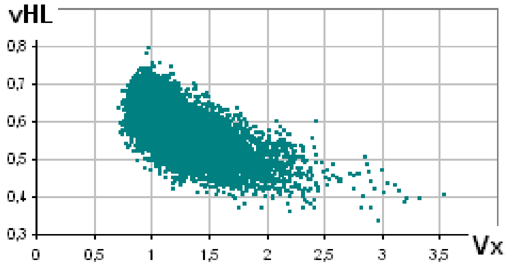
Obr. 6.26: Odhadnuté hodnoty koeficientu šikmosti estimátorů měr relativní variability při výběrech z logaritmicko-normálního rozdělení LN(c = 1) s hodnotou kurtozy β_M = 113,9.

6.2.2.2 Měří všechny charakteristiky stejnou vlastnost rozdělení?

Při porovnání síly závislosti mezi v_X a ostatními estimátory měr relativní variability lze pozorovat obdobný jev jako u estimátorů měr absolutní variability:

S rostoucí hodnotou míry kurtozy klesá síla závislosti mezi hodnotami příslušného estimátoru a v_X, přičemž pokles je pouze mírný u estimátorů měr v_L, v_{AAD} a v_{AADM}, zatímco u ostatních estimátorů je výrazně silnější (viz tab. 6.12).

V chování estimátoru v_{HL} byla shledána následující anomálie: jeho hodnoty jsou při výběrech z rozdělení s vyšší než normální hodnotou kurtozy negativně zkorelovány s hodnotami variačního koeficientu (obr. 6.27). Z tohoto důvodu jsem tento estimátor označil jako nevyhovující a vyřadil jsem jej z dalšího zkoumání.



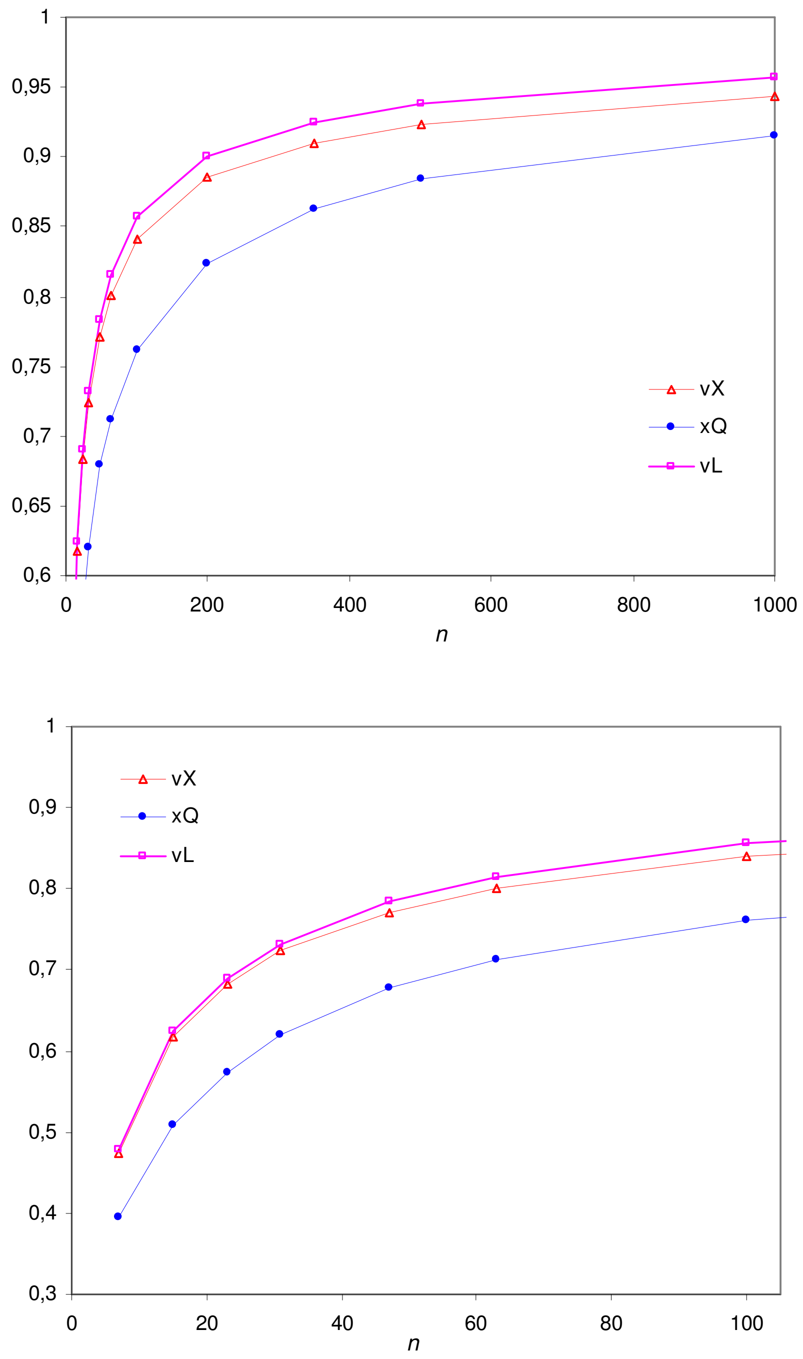
Obr. 6.27: Průběh závislosti hodnot estimátorů variability v_{HL} (na vertikální ose) a v_X při výběrech z logaritmicko-normálního rozdělení LN(c = 1) s hodnotou kurtozy β_M = 113,9.

Tabulka 6.12: Hodnoty korelačních koeficientů mezi estimátory v_X a v_Q, v_L, v_{TL}, v_{LQ}, v_{HL} a v_{AAD} při výběrech z logaritmicko-normálního rozdělení o rozsahu n = 100 (β_M = 113,9).

Estimátor	Korelační koeficient
v_Q	0,62
v_L	0,99
v_{TL}	0,85
v_{LQ}	0,73
v_{HL}	0,93
v_{AAD}	0,94

6.2.2.3 Síla testu

Hodnoty průměrné síly testů o rozdělení souboru, z něhož výběr pochází, jsou uvedeny na obr. 6.28. V grafu není uveden estimátor v_X^+, který lze vyjádřit jako transformaci v_X a dává tedy identické výsledky. Z grafu je patrné, že nejlepších výsledků dosahuje estimátor v_L, mírně horších v_X a nejhorších estimátor v_Q.



Obr. 6.28: Průměrná síla testů o populaci, z níž výběr pochází, založených na estimátorech v_X, v_L a v_Q v závislosti na rozsahu výběru n.

6.2.2.4 Verifikace dosavadních poznatků o chování estimátorů

Při porovnání empiricky získaných a teoreticky odvozených hodnot směrodatné chyby estimátoru v_X (viz kapitola 2) lze pozorovat obdobný jev jako v případě

estimátoru s_X: při výběrech z populací s vyšší než normální hodnotou kurtozy je empiricky získaná směrodatná chyba nižší než teoreticky odvozená, v případě výběrů z populací s nižší než normální hodnotou kurtozy je situace opačná.

Tabulka 6.13: Rozdíl empiricky získané a teoretické hodnoty směrodatné chyby estimátor v_X vyjádřený jako procentní podíl z teoretické hodnoty pro výběry z logaritmicko-normálního rozdělení:

n	$\frac{SE(v_X)}{SE_T(v_X)} - 1$
7	-27%
15	-19%
23	-16%
31	-12%
47	-10%
63	-7%
100	-5%
200	-3%
350	-2%
500	-1%
1000	-1%

6.3 Estimátory měr šikmosti

Pro porovnání estimátorů měr šikmosti jsem, stejně jako u estimátorů ostatních vlastností rozdělení, používal relativní charakteristiky vychýlení a variability vztažené k hodnotě $\Theta - t_{MIN}(n)$ definované v kapitole 3. Je však zřejmé, že tyto charakteristiky budou pro vyšší rozsahy výběru „stranit“ estimátoru a_M z důvodu, že s rostoucím rozsahem výběru n se dolní mez oboru hodnot tohoto estimátoru snižuje a hodnota výrazu $\Theta - t_{MIN}(n)$ tedy narůstá. Proto jsem pro porovnání průběhu výběrových vlastností estimátoru a_M s ostatními estimátory v závislosti na n prováděl paralelně také pomocí absolutních ukazatelů vychýlení a variability.

6.3.1 Výběry z normálního rozdělení

6.3.1.1 Variabilita a vychýlení

Průběh **relativního vychýlení** estimátorů je znázorněn na obr. 6.29. Je z něj patrné, že s výjimkou estimátoru a_{LQ} konverguje vychýlení velmi rychle k nule a ani pro malé rozsahy výběru nepřesahuje v absolutní hodnotě 0,06.

Při porovnání jednotlivých estimátorů z hlediska absolutní velikosti jejich **směrodatné odchylky** i **směrodatné chyby** získáváme následující pořadí: $a_{TL} \approx a_L < a_C < a_{LQ} < a_{MC} < a_Q < a_M^+ < a_M$. Při jejich porovnání z hlediska relativních ukazatelů, tj. směrodatné chyby (obr. 6.30) a odchylky vztažené k hodnotě $E(t)$-$t_{\mathrm{MIN}}(n)$, vykazuje estimátor a_M naopak nejnižší hodnoty, pořadí ostatních estimátorů se nemění.

6.3.1.2 Tvar výběrového rozdělení

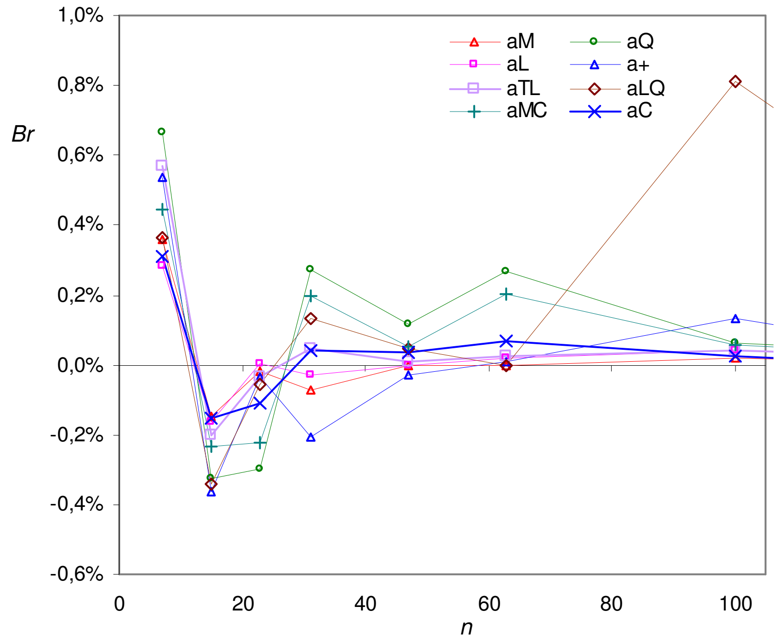
Z obrázků 6.31 a 6.32 je zřejmá rychlá konvergence koeficientů **šikmosti** i **špičatosti** výběrových rozdělení k normálním hodnotám 0 resp. 3. Za zmínku stojí divergence hodnoty koeficientu špičatosti od normální hodnoty v případě výběrového rozdělení estimátoru momentové míry a_M mezi rozsahy výběru $n = \{7, 15 \text{ a } 23\}$.

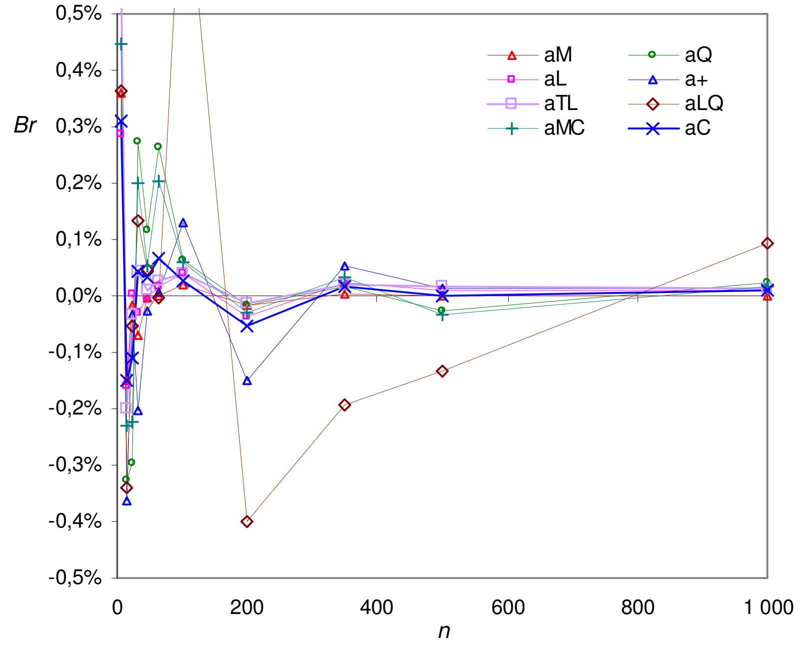
Otestování normality výběrových rozdělení

Na základě výsledku testů Kolmogorova – Smirnova o normalitě výběrových rozdělení estimátorů (tabulka 6.14) lze konstatovat, že výběrová rozdělení všech estimátorů s výjimkou a_C při výběrech z normálního rozdělení velmi rychle konvergují k rozdělení normálnímu. Rozdělení estimátorů a_L, a_{TL} vyhovují testu o normalitě pro všechny sledované rozsahy výběru, ostatní rozdělení lze považovat za normální počínaje rozsahem výběru $n = 23$ až 47. Výjimku tvoří estimátor a_C, jehož obor hodnot je, jak vyplývá z jeho definice, diskrétní. Histogram intervalového rozdělení četností hodnot estimátoru a_C je uveden na obr. 6.33.

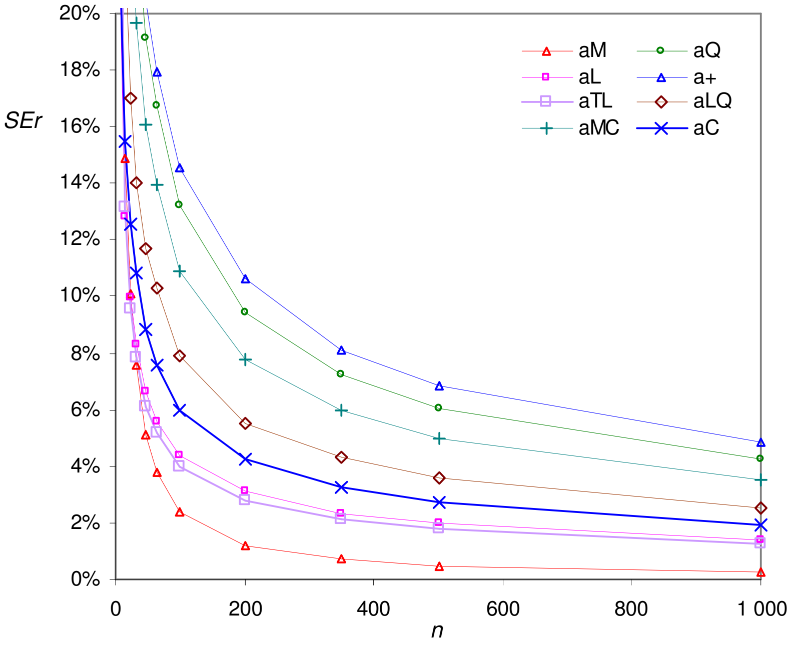
Skutečnost, že rozdělení estimátoru a_M testu o normalitě vyhovělo pro $n = 7$, ale nikoliv pro $n = \{15, 23\}$, lze vysvětlit existencí horní a dolní meze jeho oboru hodnot, které závisí na rozsahu výběru n. Tato hypotéza také vysvětluje nárůst hodnoty

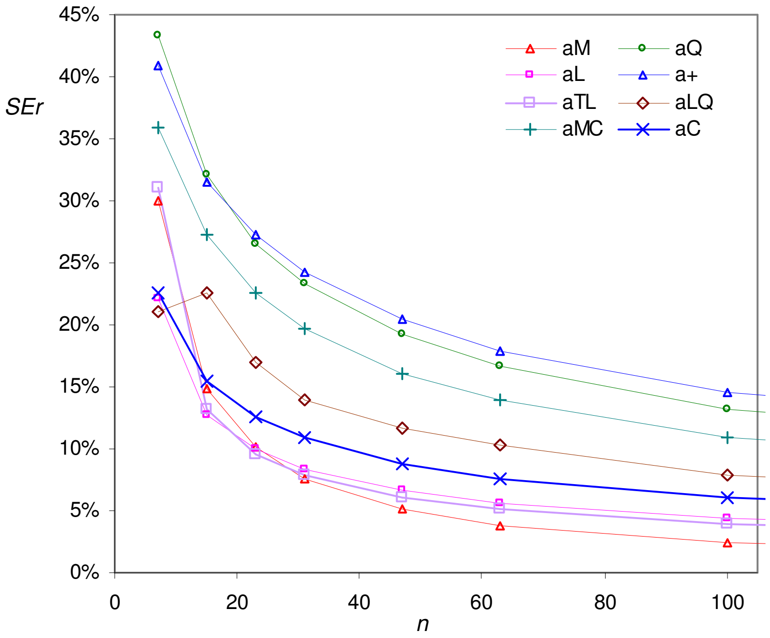
koeficientu špičatosti výběrového rozdělení tohoto estimátoru pro uvedené rozsahy výběru.



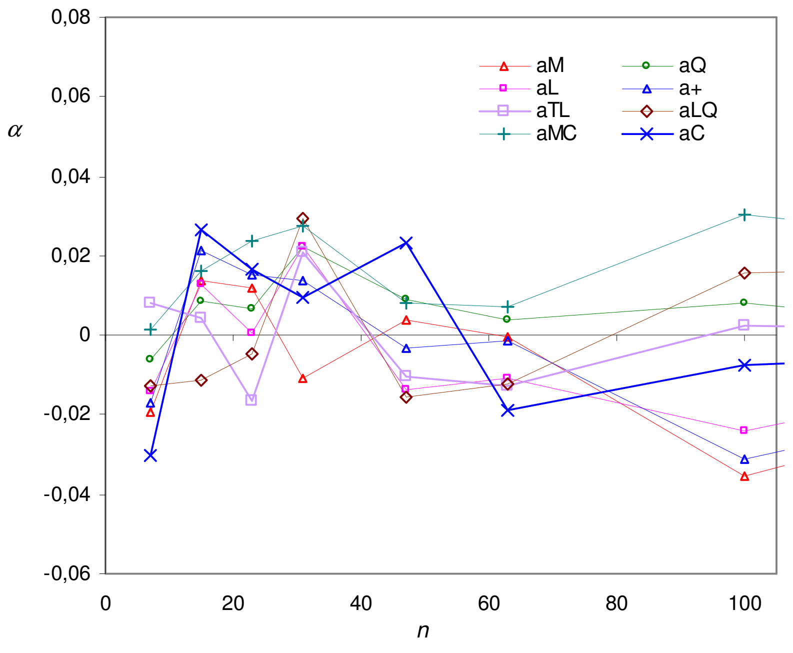


Obr. 6.29: Průběh odhadnutého relativního vychýlení estimátorů měr šikmosti při výběrech z normálního rozdělení.

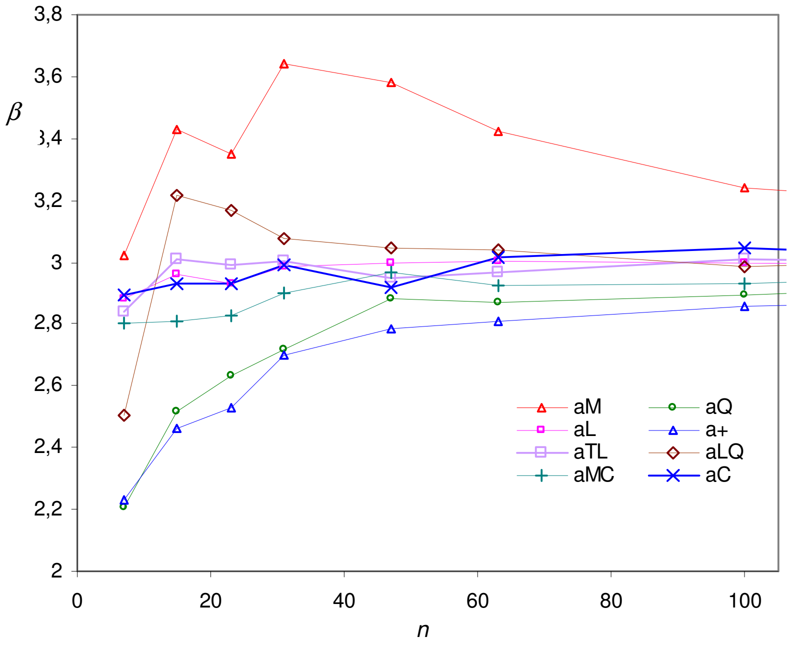




Obr. 6.30: Průběh odhadnuté relativní směrodatné chyby estimátorů měr šikmosti při výběrech z normálního rozdělení.



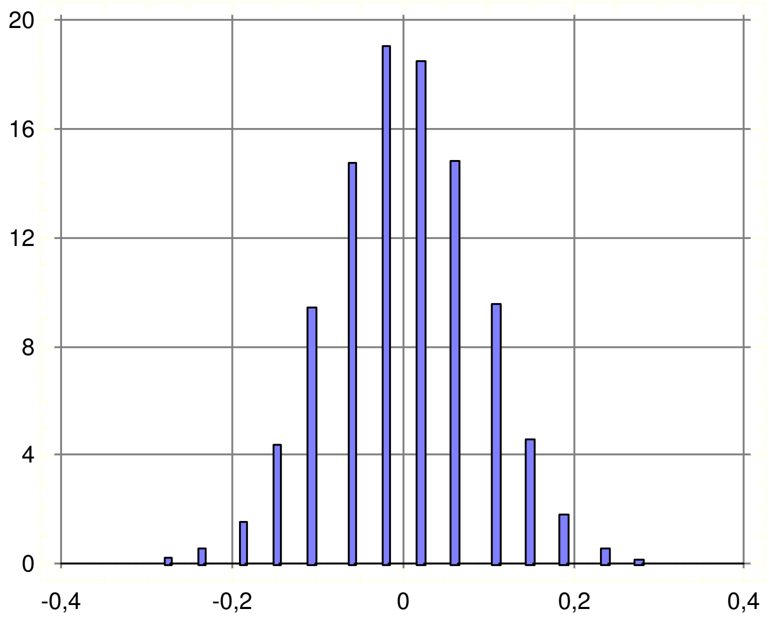
Obr. 6.31: Průběh odhadnutého koeficientu šikmosti estimátorů měr šikmosti při výběrech z normálního rozdělení pro rozsahy výběru $n \leq 100$.



Obr. 6.32: Průběh odhadnutého koeficientu špičatosti estimátorů měr šikmosti při výběrech z normálního rozdělení pro rozsahy výběru $n \leq 100$.

Tabulka 6.14: Výsledky testů Kolmogorova – Smirnova o normalitě výběrových rozdělení estimátorů měr šikmosti při výběrech z normálního rozdělení. (V tabulce je uvedena vždy nejvyšší hladina významnosti, na které výběrové rozdělení testu vyhovělo.)

n	a_M	a_Q	a_L	a_M^+	a_{TL}	a_{LQ}	a_{MC}	a_C
7	10 %	-	10 %	-	10 %	-	-	-
15	-	-	10 %	-	10 %	10 %	-	-
23	-	-	10 %	-	10 %	10 %	-	-
31	1 %	10 %	10 %	1 %	10 %	10 %	10 %	-
47	5 %	10 %	10 %	1 %	10 %	10 %	10 %	-
63+	10 %	10 %	10 %	10 %	10 %	10 %	10 %	-



Obr. 6.33: Histogram intervalového rozdělení četností empirického výběrového rozdělení estimátoru a_C pro výběry z normálního rozdělení o rozsahu n = 47.

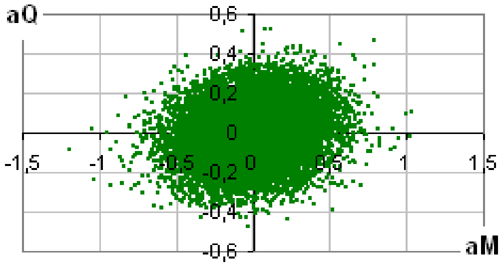
6.3.1.3 Měří všechny charakteristiky stejnou vlastnost rozdělení?

Jak je patrné z tabulky 6.15 a z obr. 6.34 až 6.40, nejvyšší sílu závislosti s hodnotami estimátoru momentové míry vykazuje estimátor normovaného koeficientu a_M^+ a estimátor míry založené na L-momentech a_L. Všechny ostatní míry vykazují sílu

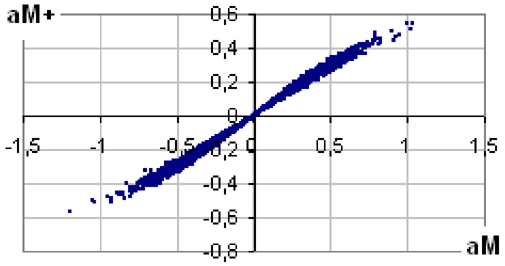
závislosti s a_M výrazně nižší, přičemž nejslabší je tato závislost u estimátoru míry založené na kvantilech a_Q. Z grafu na obr. 6.40 je také patrná nespojitost oboru hodnot estimátoru a_C.

Tabulka 6.15: Hodnoty korelačních koeficientů mezi estimátory a_M a a_Q, a_L, a_M^+, a_{TL}, a_{LQ}, a_{MC} a a_C při výběrech z normálního rozdělení o rozsahu n = 100.

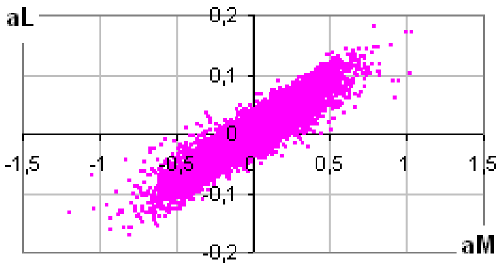
Estimátor	Korelační koeficient
a_Q	0,22
a_M^+	0,998
a_L	0,93
a_{TL}	0,64
a_{LQ}	0,46
a_{MC}	0,35
a_C	0,55



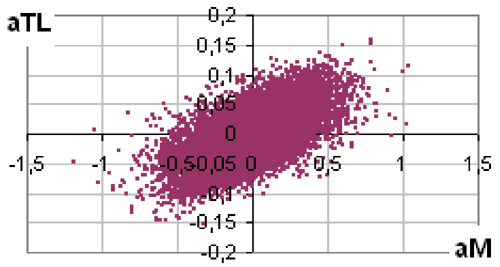
Obr. 6.34: Průběh závislosti hodnot estimátorů šikmosti a_Q a a_M při výběrech z normálního rozdělení o rozsahu n = 100.



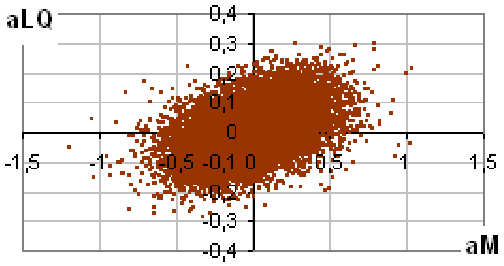
Obr. 6.35: Průběh závislosti hodnot estimátorů šikmosti a_M^+ a a_M při výběrech z normálního rozdělení o rozsahu n = 100.



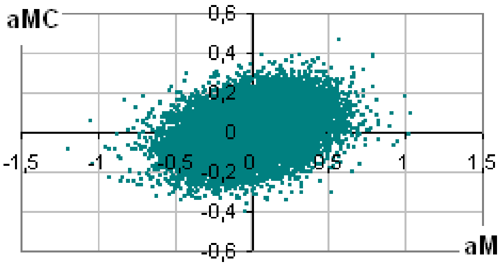
Obr. 6.36: Průběh závislosti hodnot estimátorů šikmosti a_L a a_M při výběrech z normálního rozdělení o rozsahu n = 100.



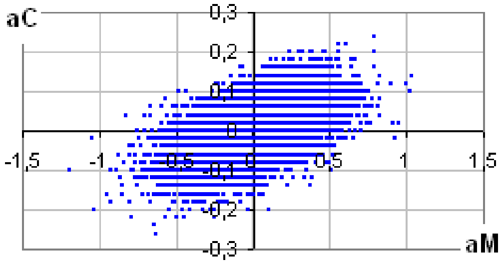
Obr. 6.37: Průběh závislosti hodnot estimátorů šikmosti a_{TL} a a_M při výběrech z normálního rozdělení o rozsahu n = 100.



Obr. 6.38: Průběh závislosti hodnot estimátorů šikmosti a_{LQ} a a_M při výběrech z normálního rozdělení o rozsahu n = 100.



Obr. 6.39: Průběh závislosti hodnot estimátorů šikmosti a_{MC} a a_M při výběrech z normálního rozdělení o rozsahu n = 100.



Obr. 6.40: Průběh závislosti hodnot estimátorů šikmosti a_C a a_M při výběrech z normálního rozdělení o rozsahu n = 100.

6.3.1.4 Verifikace dosavadních poznatků o chování estimátorů

Porovnání hodnot směrodatné chyby, vychýlení a koeficientů šikmosti a špičatosti výběrového rozdělení estimátoru a_M, získaných na základě experimentu, s aproximacemi teoretickými hodnot vypočtenými podle vztahů uvedených v kapitole 2 (dle PEARSON, 1930) je uvedeno v tabulce 6.16. Je z něj patrné, že aproximace směrodatné chyby platí pro všechny rozsahy výběru, zatímco aproximace koeficientu špičatosti výběrového rozdělení platí až pro $n \leq 47$. Zároveň v souladu s teoretickými výsledky platí, že při výběrech z normálního rozdělení lze estimátor považovat za nevychýlený a jeho výběrové rozdělení za symetrické pro všechny sledované rozsahy výběru.

Tabulka 6.16: Teoretické a empirické hodnoty směrodatné chyby, vychýlení a koeficientů šikmosti a špičatosti výběrového rozdělení estimátoru α_M při výběrech z normálního rozdělení. Poznámka: teoretické hodnoty $B_T(a_M) = \alpha_T(a_M) = 0$. U nevyplněných polí tabulky nedává aproximace smysluplné hodnoty.

n	Teoretické hodnoty (viz kap. 2)		Empirické hodnoty			
	$SE_T(a_M)$	$\beta_T(a_M)$	$SE(a_M)$	$B(a_M)$	$\alpha(a_M)$	$\beta(a_M)$
7	0,6019	-	0,6120	0,0073	-0,02	3,02
15	0,5200	-	0,5175	-0,0052	0,01	3,43
23	0,4493	2,9329	0,4517	-0,0007	0,01	3,35
31	0,3999	3,2626	0,4018	-0,0038	-0,01	3,64
47	0,3354	3,3749	0,3389	-0,0002	0,00	3,58
63	0,2944	3,3538	0,2949	-0,0001	0,00	3,42
100	0,2377	3,2736	0,2374	0,0021	-0,04	3,24
200	0,1706	3,1584	0,1711	-0,0024	0,01	3,10
350	0,1298	3,0958	0,1303	0,0009	0,03	3,07
500	0,1089	3,0685	0,1096	0,0002	0,00	3,07
1000	0,0772	3,0351	0,0771	0,0003	0,03	3,10

V tabulce v příloze č. 3 jsou dále uvedeny empiricky získané hodnoty výrazu $SE(\alpha_M)\sqrt{n}$. Je z nich patrné, že limitní vztah pro směrodatnou chybu estimátoru a_M uvedený v kapitole 2 (dle KENDAL, STUART, 1963) platí přibližně pro rozsahy výběru $n \geq 100$. Limitní vztah pro směrodatnou chybu estimátoru Bowleyova kvantilového koeficientu a_Q (WAGEMAKERS et al., 1995) platí přibližně pro $n \geq 47$.

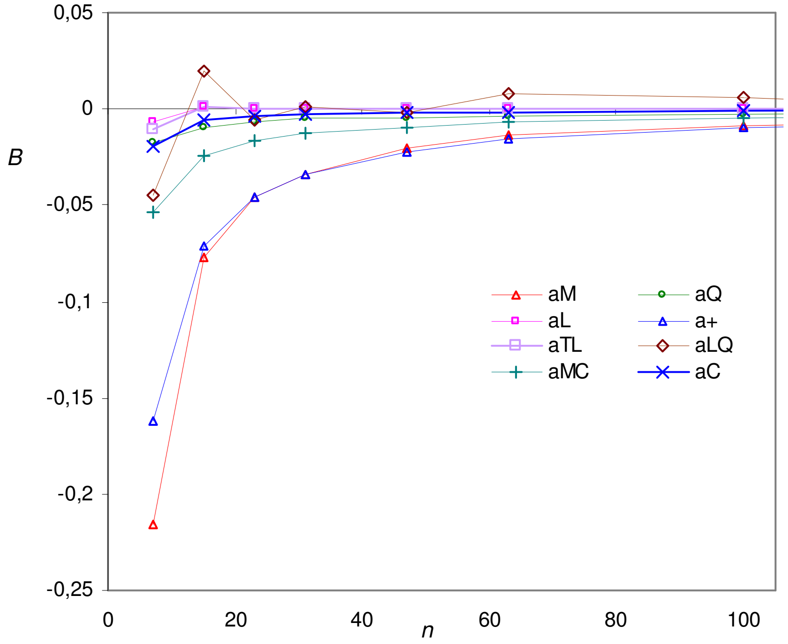
6.3.2 Výběry z ostatních rozdělení

6.3.2.1 Porovnání estimátorů z hlediska variability, vychýlení a tvaru rozdělení

Při porovnání estimátorů měr šikmosti při výběrech z ostatních rozdělení se ukazuje, že vlastnosti jejich výběrových rozdělení závisejí na absolutní hodnotě šikmosti v základním souboru. Pro tyto účely tedy rozlišuji situaci, kdy rozdělení populace je symetrické nebo mírně zešikmené a kdy je výrazně zešikmené.

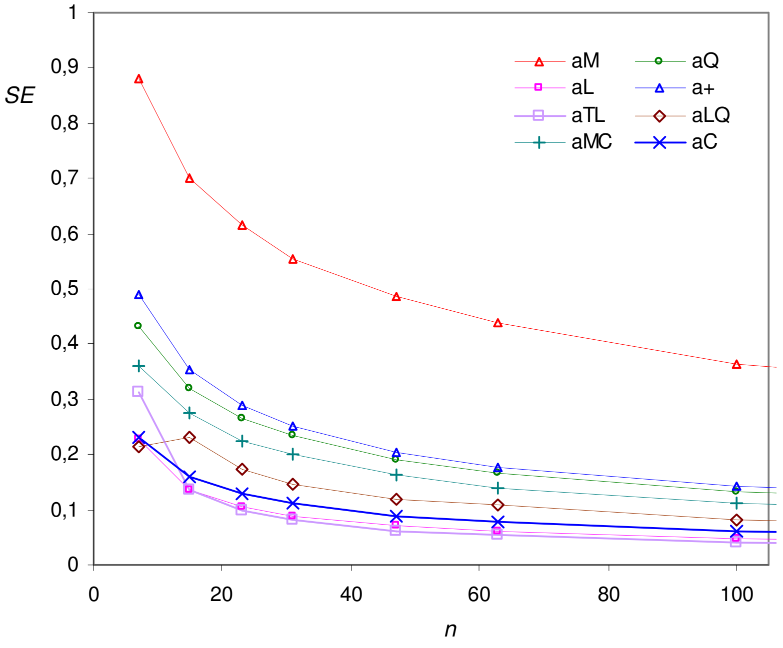
U rozdělení **symetrických** nebo jen **slabě zešikmených** jsou výsledky obdobné jako v případě výběrů z rozdělení normálního. V případě nenulové hodnoty koeficientu šikmosti lze pozorovat (obr. 6.41), že estimátory a_M a a_M^+ pro nižší hodnoty

n (přibližně pro $n \leq 100$) v absolutní hodnotě mírně podhodnocují (tzn. v případě záporného znaménka α_M nadhodnocují).



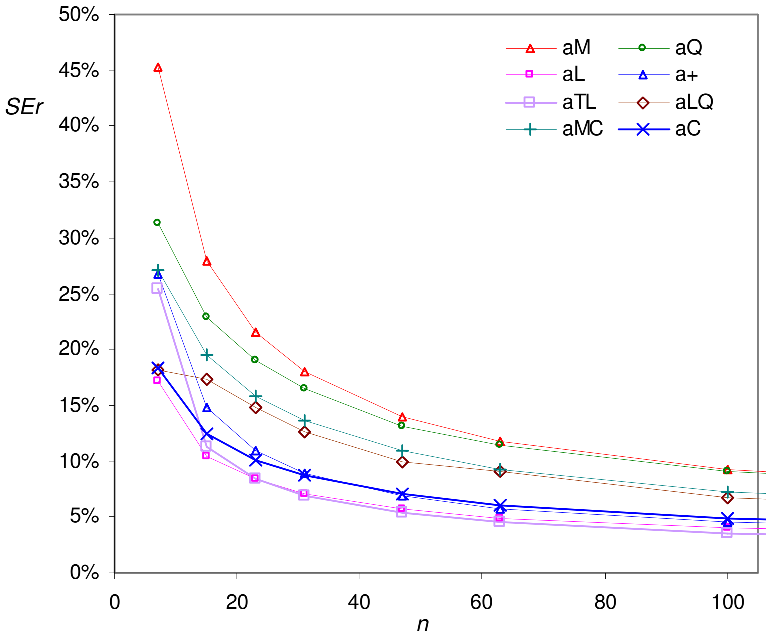
Obr. 6.41: Průběh vychýlení estimátorů měr šikmosti při výběrech z rozdělení beta ($p = 1$, $q = 2$) s hodnotou $\alpha_M = 0{,}57$ pro výběry o rozsahu $n \leq 100$.

Při výběrech z **výrazněji zešikmených populací** (přibližně s hodnotou $|\alpha_M| > 1$) můžeme pozorovat, že s rostoucí hodnotou α_M se **vychýlení** estimátorů a_M a a_M^+ zvyšuje (estimátory v absolutní hodnotě systematicky podhodnocují) a zejména u estimátoru a_M velmi pomalu konverguje k nule v závislosti na rozsahu výběru n. Spolu s vychýlením se s rostoucím zešikmením zvyšuje i **směrodatná chyba** těchto dvou estimátorů. Vychýlení ostatních estimátorů konverguje velmi rychle k nule, obdobně jako bylo pozorováno při výběrech ze symetrických rozdělení. Z hlediska hodnot směrodatné chyby zůstává zachováno stejné pořadí estimátorů jako v případě výběrů z normálního rozdělení, tj. nejnižších hodnot dosahuje u a_L a a_{TL}, nejvyšších naopak u estimátoru momentové míry a_M.



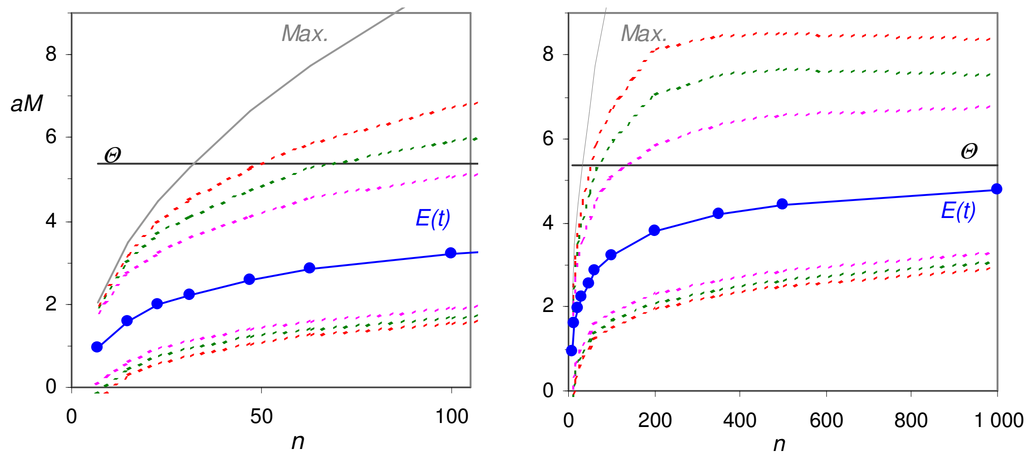
Obr. 6.42: Průběh směrodatné chyby estimátorů měr šikmosti při výběrech z logaritmicko-normálního rozdělení s parametrem c = 0,316 (α_M = 0,98) pro výběry o rozsahu $n \leq$ 100.

Při výběrech z rozdělení s vyšší hodnotou koeficientu šikmosti (přibližně $|\alpha_M| > 3$) estimátor a_M vychází pro nižší rozsahy výběru (přibližně pro $n \leq$ 100) jako nejhorší také při porovnání z hlediska **relativní směrodatné chyby**, a to i přes nedostatek tohoto kritéria zmíněný výše, kvůli kterému s rostoucím rozsahem výběru toto kritérium „straní“ estimátoru a_M (obr. 6.43).

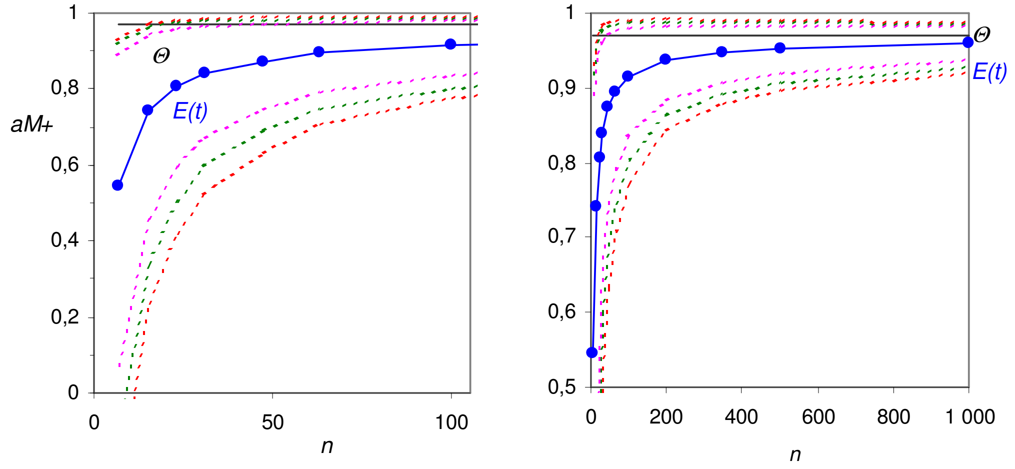


Obr. 6.43: Průběh relativní směrodatné chyby estimátorů měr šikmosti při výběrech o rozsahu $n \leq 100$ z populace s hodnotou $\alpha_M = 3{,}17$.

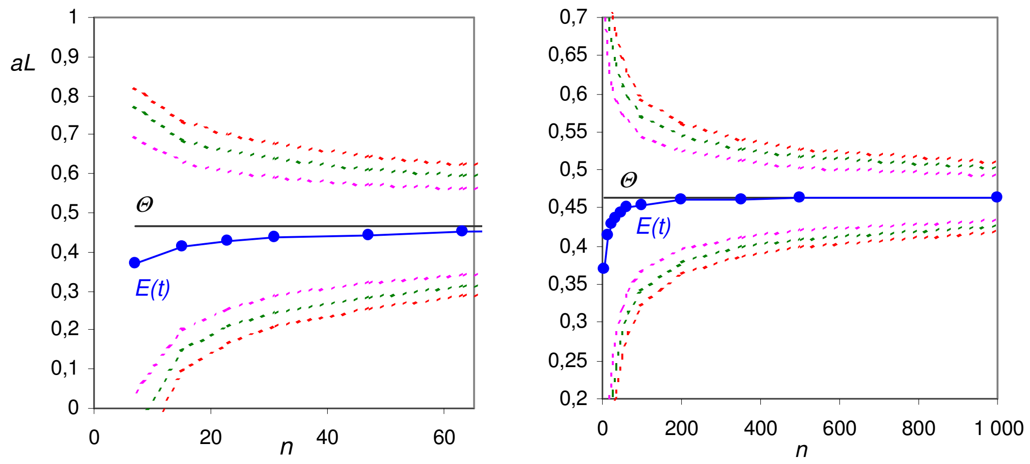
Průběh **vychýlení** estimátoru a_M při výběrech z výrazně zešikmeného rozdělení (s hodnotou koeficientu $\alpha_M = 5{,}4$) je patrný z obr. 6.44, který znázorňuje závislost odhadnuté střední hodnoty a význačných kvantilů výběrového rozdělení estimátoru a_M na rozsahu výběru n. Je z něj například patrná skutečnost, že při výběrech o rozsahu $n = 100$ estimátor ve více než 90% případů podhodnocuje. Tento jev je zřejmě způsoben existencí horní hranice oboru hodnot tohoto estimátoru, která je v grafu znázorněna šedou křivkou. Podobný jev, i když v menším rozsahu, lze pozorovat také u estimátoru a_M^+ (obr.6.45), jehož vychýlení však konverguje k nule výrazně rychleji. Pro srovnání dále uvádím průběh význačných kvantilů estimátoru a_L (obr. 6.46), u něhož je konvergence vychýlení k nule nejrychlejší.



Obr. 6.44: Průběhy empirické střední hodnoty E(*t*), význačných kvantilů a horní hranice oboru hodnot estimátoru a_M v porovnání se skutečnou hodnotou $\alpha_M = 5{,}4$ v populaci (v grafu jako Θ) v závislosti na rozsahu výběru při výběrech z logaritmicko-normálního rozdělení s parametrem $c = 1$. Přerušované křivky v pořadí shora: 97,5%, 95%, 90%, 10%, 5% a 2,5% kvantil.

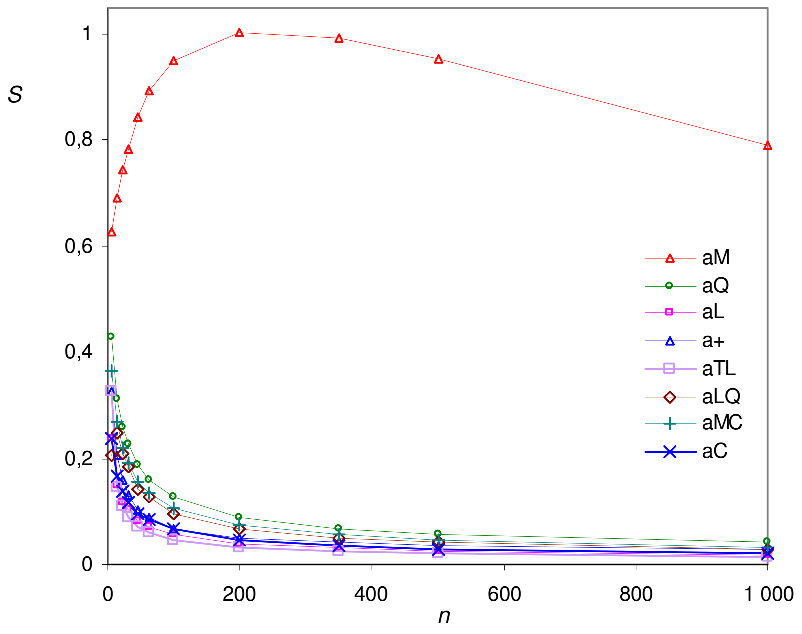


Obr. 6.45: Průběhy empirické střední hodnoty E(*t*) a význačných kvantilů estimátoru a_M^+ v porovnání se skutečnou hodnotou $\alpha_M^+ = 0{,}97$ v populaci (v grafu jako Θ) v závislosti na rozsahu výběru při výběrech z logaritmicko-normálního rozdělení s parametrem $c = 1$. Přerušované křivky v pořadí shora: 97,5%, 95%, 90%, 10%, 5% a 2,5% kvantil.



Obr. 6.46: Průběhy empirické střední hodnoty E(t) a význačných kvantilů estimátoru a_L v porovnání se skutečnou hodnotou $\alpha_L = 0{,}46$ v populaci (v grafu jako Θ) v závislosti na rozsahu výběru při výběrech z logaritmicko-normálního rozdělení s parametrem $c = 1$. Přerušované křivky v pořadí shora: 97,5%, 95%, 90%, 10%, 5% a 2,5% kvantil.

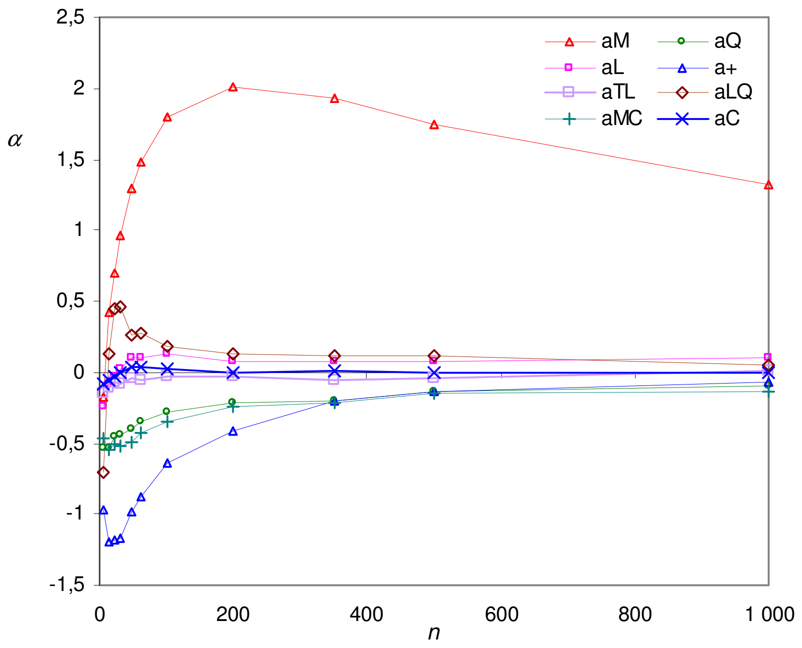
S existencí horní hranice oboru hodnot **estimátoru momentové míry** a_M zřejmě souvisí několik **anomálií** v jeho chování. Jednou z nich je **počáteční nárůst směrodatné odchylky** výběrového rozdělení estimátoru v závislosti na rozsahu výběru n pro populace s výraznějším zešikmením. Jedná se o obdobný jev, jaký byl pozorován u estimátoru variačního koeficientu v_X v případě populace s vysokou hodnotou kurtozy. V případě estimátoru a_M však tato anomálie projevuje v mnohem širším rozsahu (odhadnutá směrodatná odchylka estimátoru je v mnohých případech rostoucí funkcí n i pro rozsahy výběru $n = 500$ a její obrat v klesající funkci je ve srovnání v estimátorem v_X výrazně pomalejší).



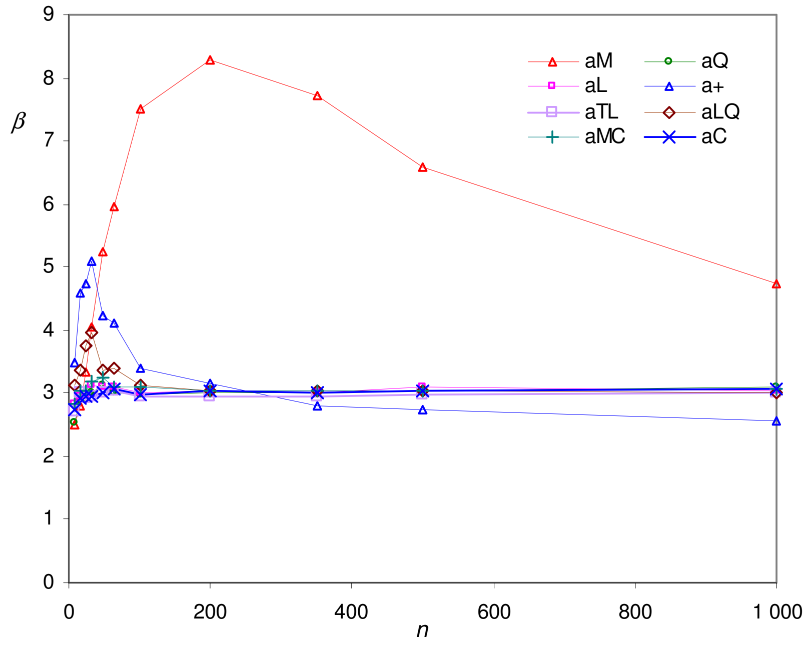
Obr. 6.47: Průběhy empirické směrodatné odchylky estimátorů měr šikmosti v závislosti na rozsahu výběru *n* při výběrech z populace hodnotou $\alpha_M = 3{,}24$.

Další anomálií v chování estimátoru momentové míry a_M, která se pro malé rozsahy výběru projevuje také u estimátoru a_M^+, je divergence od normálních hodnot **koeficientů šikmosti** a **špičatosti** výběrových rozdělení. Rozdělení estimátoru a_M je kladně zešikmené, přičemž míra jeho zešikmení s rozsahem výběru přibližně pro $n \leq 500$ roste. Rozdělení estimátoru a_M^+ je naopak záporně zešikmené a nárůst tohoto zešikmení se projevuje pouze pro nižší rozsahy výběru, přibližně pro $n \leq 31$ (obr. 6.48 a 6.49).

Při porovnání konvergence k normálním hodnotám šikmosti a špičatosti pro ostatní estimátory zjišťujeme, že rychlost této konvergence je pomalejší než v případě normálního rozdělení. Nejrychlejší konvergenci k normálním hodnotám vyvazují estimátory a_{TL}, a_C, a a_L. U všech ostatních estimátorů jsou však jejich odchylky od normálních hodnot výrazně menší než u a_M a a_M^+.



Obr. 6.48: Průběhy odhadnutého koeficientu šikmosti výběrových rozdělení estimátorů měr šikmosti v závislosti na rozsahu výběru *n* při výběrech z populace hodnotou $\alpha_M = 3{,}24$.

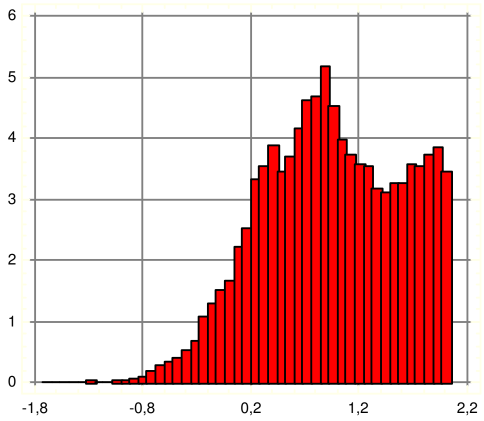


Obr. 6.49: Průběhy odhadnutého koeficientu špičatosti výběrových rozdělení estimátorů měr šikmosti v závislosti na rozsahu výběru *n* při výběrech z populace hodnotou $\alpha_M = 3{,}24$.

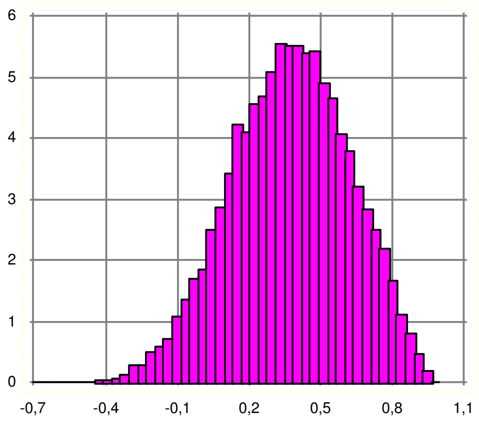
Konvergence k normalitě

Na základě **testování normality výběrových rozdělení** jednotlivých estimátorů při výběrech ze zešikmených populací lze konstatovat, že nejsilnější odchylky od normality vykazují výběrová rozdělení estimátorů a_M a a_M^+, která nevyhověla testu o normalitě pro žádnou ze zkoumaných populací, a to bez ohledu na rozsah výběru. Testům o normalitě nikdy nevyhovělo ani rozdělení estimátoru a_C z důvodu jeho nespojitosti. Z ostatních estimátorů vykazují z hlediska konvergence k normalitě nejlepší vlastnosti estimátory měr a_L a a_{TL}, jejichž výběrová rozdělení v některých případech vyhověla testu normality pro všechny zkoumané rozsahy výběru. Rychlost konvergence výběrových rozdělení k normalitě však závisí na konkrétní populaci a na základě získaných výsledků nelze stanovit žádné obecně platné pravidlo, které by říkalo, od jakého rozsahu je lze považovat za normální. (U některých výrazně zešikmených populací nevyhověly testu o normalitě ani rozdělení estimátorů a_L a a_{TL}, nebo vyhověly až pro vysoké rozsahy výběru.)

Další anomálie u estimátoru a_M byla zjištěna z hlediska tvaru jeho výběrového rozdělení. Při výběrech nižšího rozsahu (přibližně pro $n = \{7, 15, 23\}$) z populací s vysokým zešikmením (přibližně pro $\alpha_M > 5$) má jeho výběrové rozdělení tvar S-křivky s lokálním maximem při horní hranici jeho oboru hodnot (obr. 6.50). Pro srovnání uvádím také výběrové rozdělení estimátoru a_L pro stejnou populaci a rozsah výběru (obr. 6.51).



Obr. 6.50: Empirické výběrové rozdělení estimátoru a_M při výběrech z rozdělení LN(1) s hodnotu $a_M = 5{,}4$ při rozsahu výběru $n = 7$.



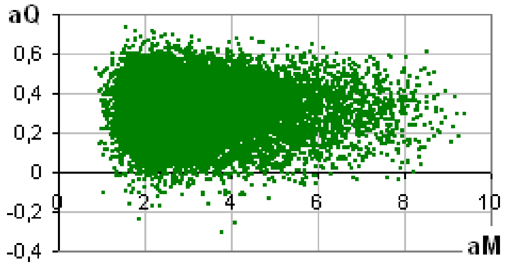
Obr. 6.51: Empirické výběrové rozdělení estimátoru a_L při výběrech z rozdělení LN(1) s hodnotu $a_L = 0{,}46$ při rozsahu výběru $n = 7$.

6.3.2.2 Měří všechny charakteristiky stejnou vlastnost rozdělení?

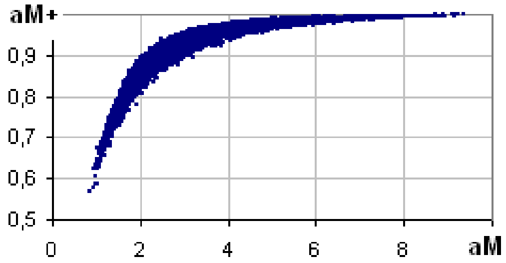
Při porovnání jednotlivých estimátorů z hlediska síly jejich závislosti s hodnotami estimátoru momentové míry docházíme k následujícím závěrům: V případě výběrů ze symetrických populací jsou síly závislostí přibližně stejné jako v případě výběrů z normální populace. S rostoucí mírou zešikmení síly těchto závislostí klesají; tento pokles je přitom pouze mírný v případě estimátorů a_L a a_M^+, zřetelnější je u estimátoru a_C. U ostatních estimátorů tento pokles s rostoucí úrovní zešikmění v populaci vede až k jejich faktické nezávislosti s hodnotami estimátoru momentové míry. Jako příklad jsou v tabulce 6.17 uvedeny hodnoty korelačních koeficientů a na obr. 6.52 až 6.58 bodové diagramy v případě výběrů z logaritmicko-normálního rozdělení s parametrem $c = 1$.

Tabulka 6.17: Hodnoty korelačních koeficientů mezi estimátory a_M a a_Q, a_L, a_M^+, a_{TL}, a_{LQ}, a_{MC} a a_C při výběrech z logaritmicko-normálního rozdělení s parametrem $c = 1$ o rozsahu $n = 100$ ($\alpha_M = 5{,}4$).

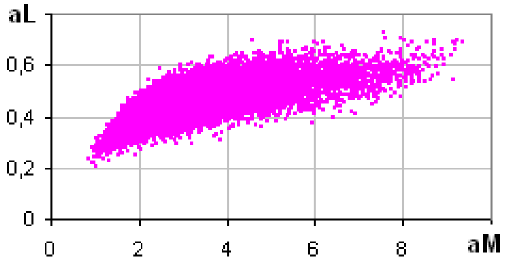
Estimátor	Korelační koeficient
a_Q	-0,04
a_M^+	0,82
a_L	0,72
a_{TL}	0,09
a_{LQ}	0,02
a_{MC}	-0,05
a_C	0,49



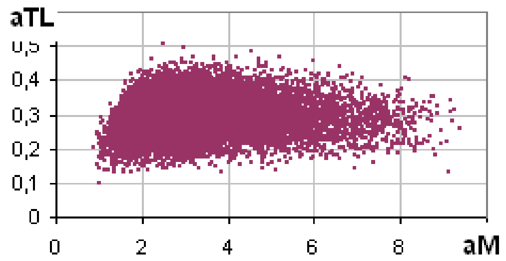
Obr. 6.52: Průběh závislosti hodnot estimátorů šikmosti a_Q a a_M při výběrech z logaritmicko-normálního rozdělení s parametrem $c = 1$ o rozsahu $n = 100$ ($\alpha_M = 5{,}4$).



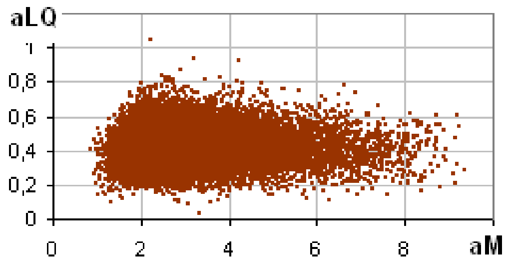
Obr. 6.53: Průběh závislosti hodnot estimátorů šikmosti a_M^+ a a_M při výběrech z rozdělení LN(1) o rozsahu $n = 100$.



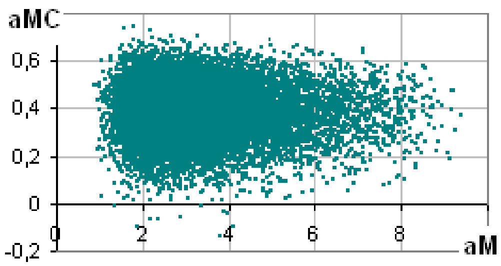
Obr. 6.54: Průběh závislosti hodnot estimátorů šikmosti a_L a a_M při výběrech z rozdělení LN(1) o rozsahu $n = 100$.



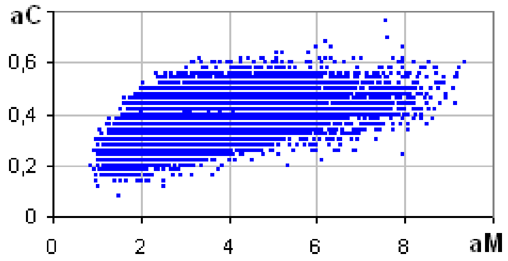
Obr. 6.55: Průběh závislosti hodnot estimátorů šikmosti a_{TL} a a_M při výběrech z rozdělení LN(1) o rozsahu $n = 100$.



Obr. 6.56: Průběh závislosti hodnot estimátorů šikmosti a_{LQ} a a_M při výběrech z rozdělení LN(1) o rozsahu n = 100.



Obr. 6.57: Průběh závislosti hodnot estimátorů šikmosti a_{MC} a a_M při výběrech z rozdělení LN(1) o rozsahu n = 100.



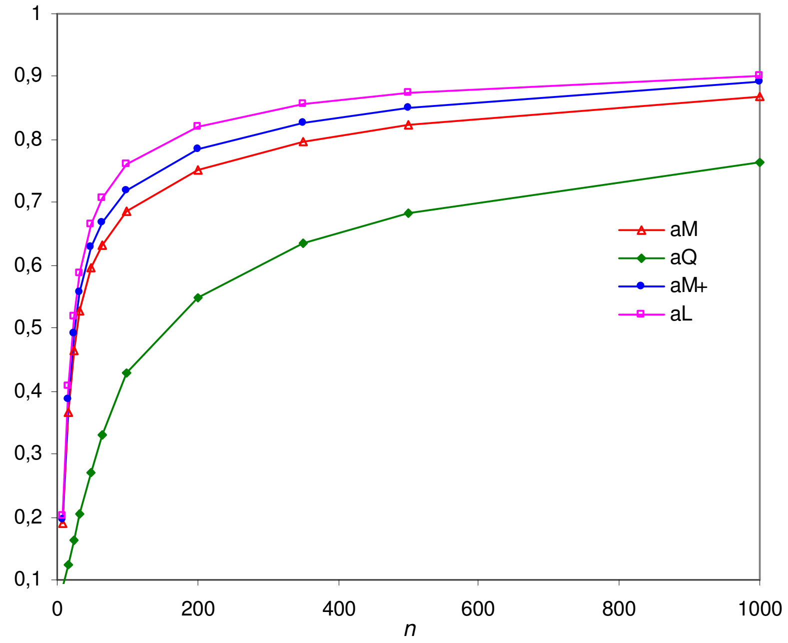
Obr. 6.58: Průběh závislosti hodnot estimátorů šikmosti a_C a a_M při výběrech z rozdělení LN(1) o rozsahu n = 100.

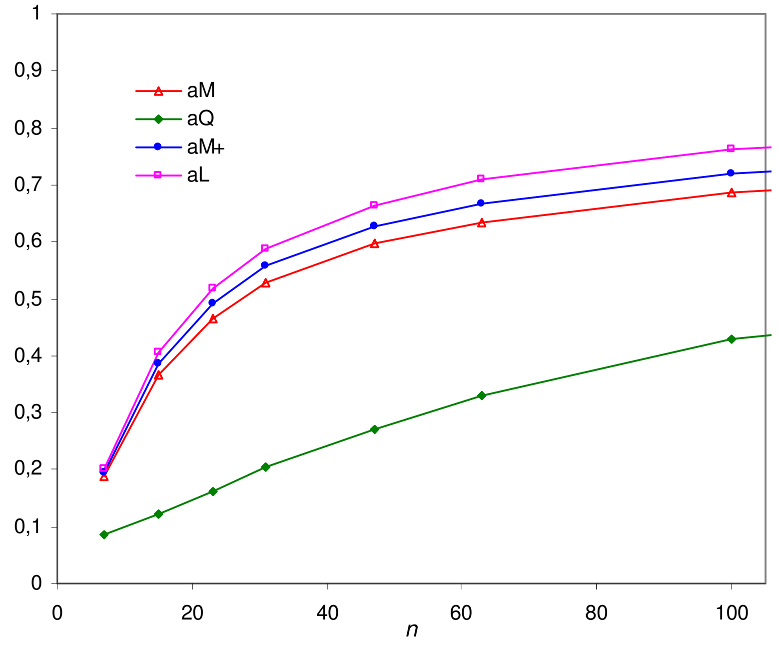
6.3.2.3 Síla testu

Hodnoty průměrné síly **testů o rozdělení** souboru, z něhož výběr pochází, jsou uvedeny na obr. 6.59. Z výsledků porovnání estimátorů měr šikmosti z hlediska průměrné síly testů o tvaru rozdělení základního souboru lze vyvodit obdobné závěry, jako v případě estimátorů měr relativní variability. Nejlepších výsledků dosahují estimátory míry založené na L-momentech a pouze nepatrně horších estimátor momentové míry a_M a standardizovaného koeficientu a_M^+. Naopak nejhorší výsledky vykazují estimátory měr založených na kvantilech.

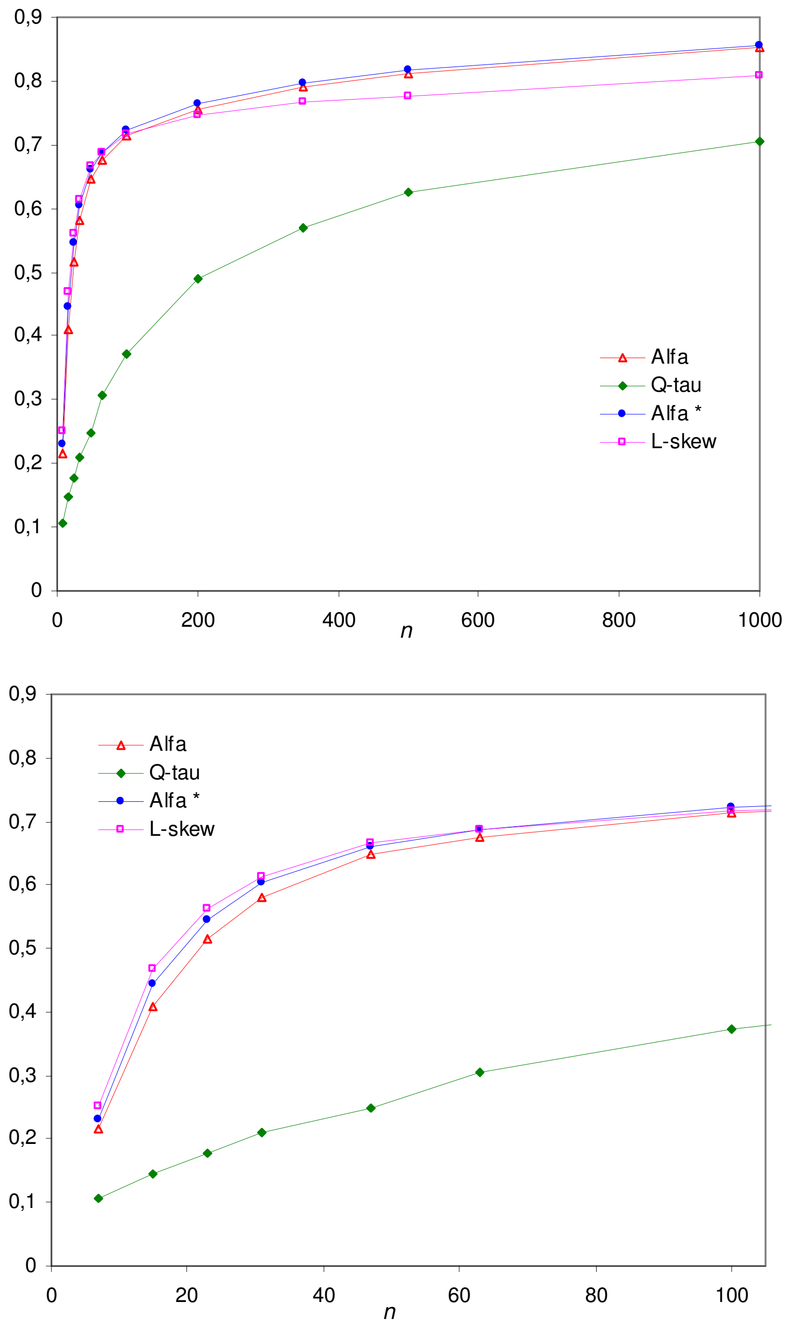
Pokud se zaměříme pouze na **testy o normalitě rozdělení základního souboru** (tj. nikoliv na průměr za všechny možné variace rozdělení specifikovaných v testované a alternativní hypotéze), je třeba rozlišit mezi výběry malého a velkého rozsahu. Při výběrech o rozsahu přibližně $n < 100$ jsou výsledky stejné jako při porovnání v rámci

všech zkoumaných populací navzájem. Pro vyšší rozsahy výběru dosahuje estimátor míry založené na L-momentech výsledků mírně horších oproti estimátorům měr založených na klasických momentech (obr. 6.60).





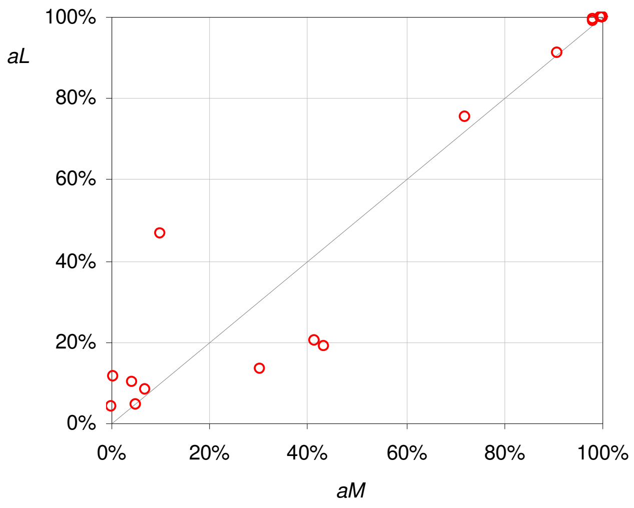
Obr. 6.59: Průměrná síla testů o populaci, z níž výběr pochází, založených na estimátorech a_M, a_M^+, a_L, a a_Q v závislosti na rozsahu výběru *n*.



Obr. 6.60: Průměrná síla testů o normalitě rozdělení, z něhož výběr pochází, založených na estimátorech a_M, a_M^+, a_Q a a_L, v závislosti na rozsahu výběru n.

Pro učinění zobecňujících závěrů o síle testů o normalitě však nemáme dostatek pozorování (tj. zkoumaných populací). Na rozdíl od případu, kdy jsme posuzovali

průměrnou sílu testu pro všechny variace populací, celkem tedy $V_2(K) = 1\,190$ porovnání, v případě testu o normalitě vycházíme pouze z $(K - 1) = 39$ porovnání. Zejména pro stanovení pořadí mezi estimátory a_M, a_M^+ a a_L by tedy bylo třeba provést rozsáhlejší zkoumání a vzít v úvahu nejen průměrnou sílu testu, ale také například porovnat průměrné rozdíly v síle testů založených na jednotlivých estimátorech pro konkrétní populace specifikované v alternativní hypotéze. Porovnání hodnot síly testů založených na estimátorech a_M a a_L je uvedeno v grafu na obr. 6.61, kde jsou proti sobě vyneseny konkrétní hodnoty síly testu založeného na a_M a a_L.



Obr. 6.61: Odhadnuté síly testů o normalitě rozdělení založených na estimátorech a_M (na horizontální ose) a a_L při výběrech o rozsahu $n = 100$. (Pozn.: velká část pozorování s odhadnutou sílou testu blízkou 100% splývá v pravém horním rohu grafu.)

6.4 Estimátory měr kurtozy

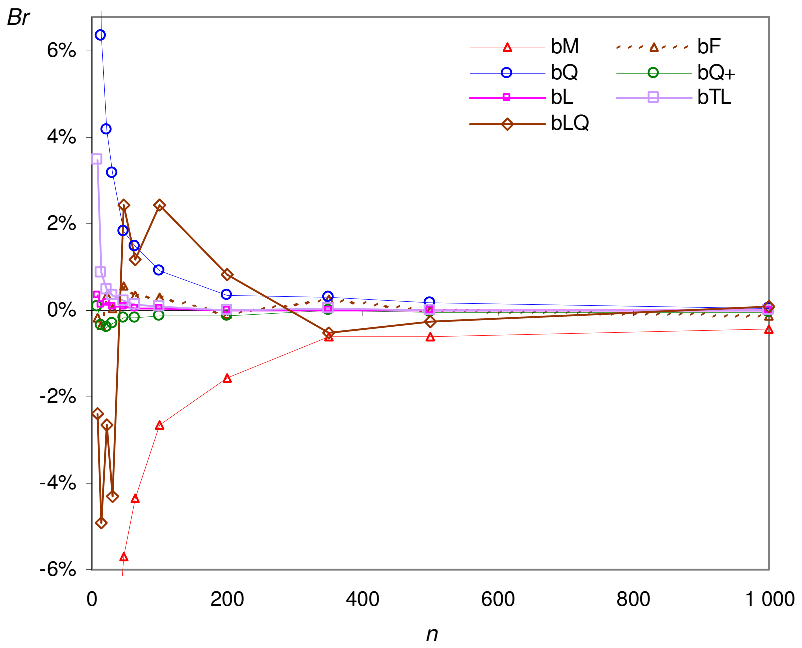
6.4.1 Výběry z normálního rozdělení

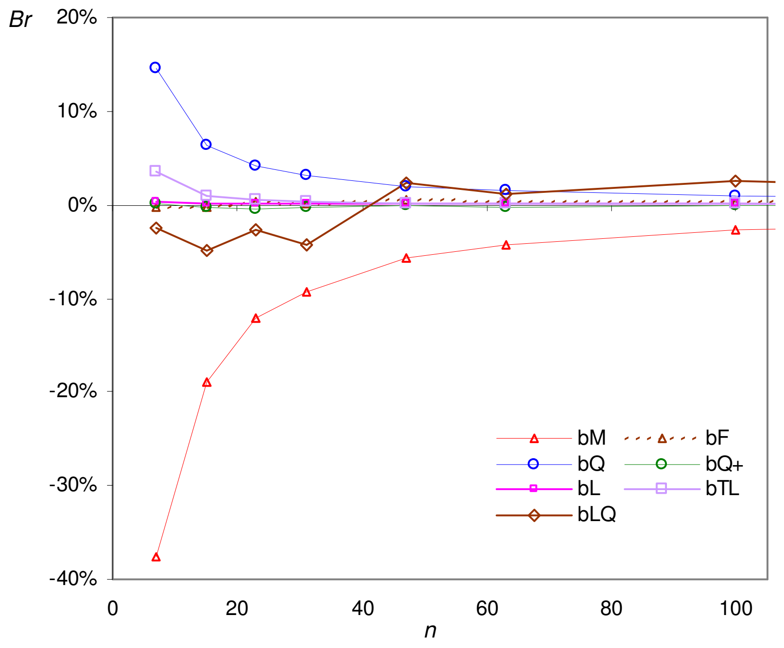
6.4.1.1 Variabilita a vychýlení

Při porovnání estimátorů b_M a b_F z hlediska **relativního vychýlení** (obr. 6.62) se potvrzuje obecně známá skutečnost, že b_M je v případě výběrů z normálního rozdělení estimátorem vychýleným (podhodnocuje), zatímco estimátor b_F lze považovat za téměř nevychýlený (viz např. KENDAL, STUART, 1963). Při celkovém porovnání lze za nevychýlené považovat estimátory b_F, b_L a b_Q^+ a pro rozsahy výběru $n \geq 63$ také b_{TL}. Nestandardní je průběh vychýlení estimátoru b_{LQ}, který pro některé rozsahy výběru výrazně podhodnocuje, pro jiné naopak nadhodnocuje.

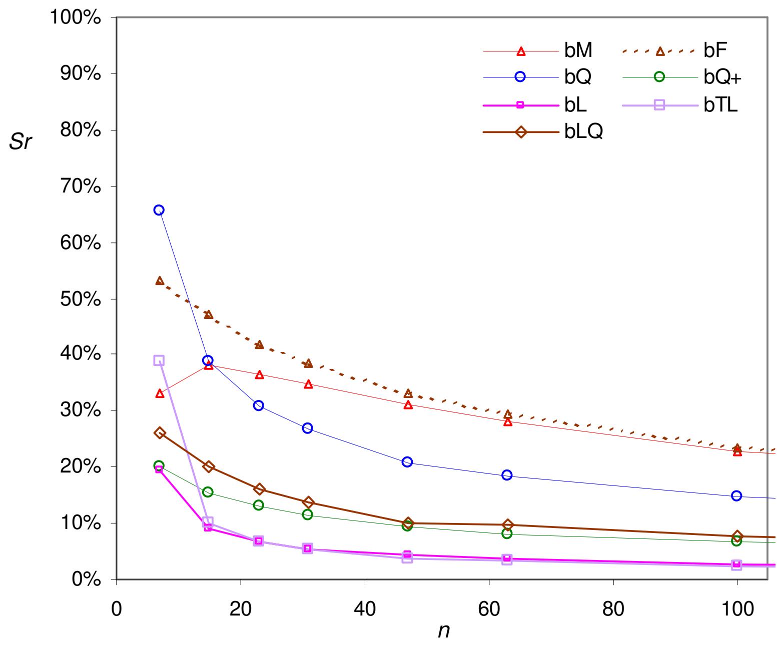
Porovnáním estimátorů b_M a b_F z hlediska **relativní variability** (obr. 6.63) se opět potvrzuje obecně známá skutečnost (např. KENDAL, STUART, 1963), že estimátor b_F se vyznačuje vyšší variabilitou než b_M. Rozdíl mezi rozptyly těchto dvou estimátorů přitom s rostoucím rozsahem výběru n klesá. Za zmínku stojí průběh směrodatné odchylky estimátoru b_M, která prochází maximem pro rozsah výběru $n = 15$. Jedná se zřejmě o obdobu anomálie, kterou bylo možné pozorovat také u estimátoru variačního koeficientu a momentové míry šikmosti v případě výběrů z určitých typů rozdělení. U estimátoru b_M se však tato anomálie projevuje i při výběrech z rozdělení normálního.

Při porovnání z hlediska **relativní směrodatné chyby** dostáváme (s výjimkou výběrů o rozsahu $n = 7$) následující pořadí estimátorů podle její velikosti: $b_L \approx b_{TL} < b_Q^+ < b_{LQ} < b_Q < b_M < b_F$. Pro rozsahy výběru $n \geq 200$ jsou přitom rozdíly mezi estimátory b_M a b_F téměř zanedbatelné. Při výběrech o rozsahu $n = 7$ vykazuje nejvyšší míru relativní směrodatné chyby estimátor b_Q, pořadí ostatních estimátorů se nemění (obr. 6.64).

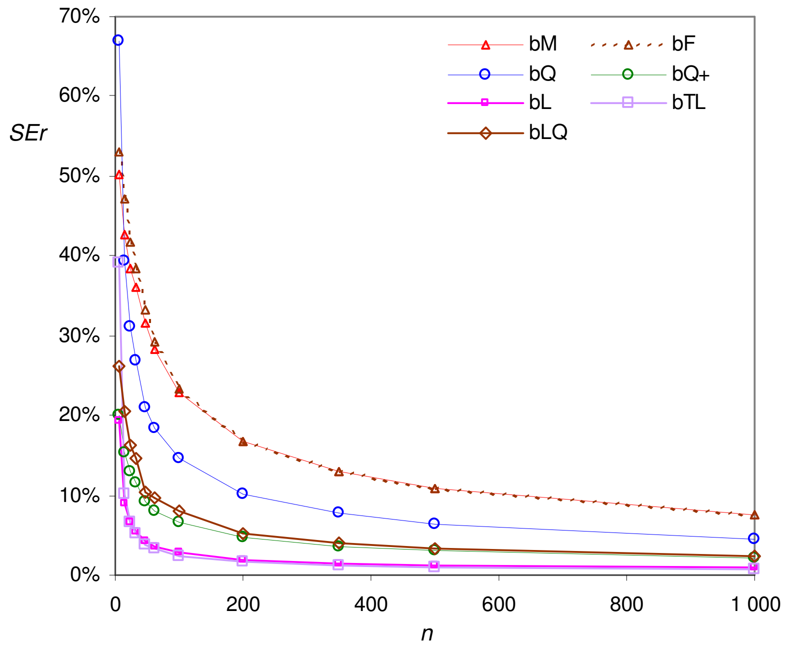




Obr. 6.62: Relativní vychýlení estimátorů měr kurtozoy při výběrech z normálního rozdělení v závislosti na rozsahu výběru.



Obr. 6.63: Relativní směrodatná odchylka estimátorů měr kurtozy při výběrech z normálního rozdělení pro rozsahy výběru $n \leq 100$.



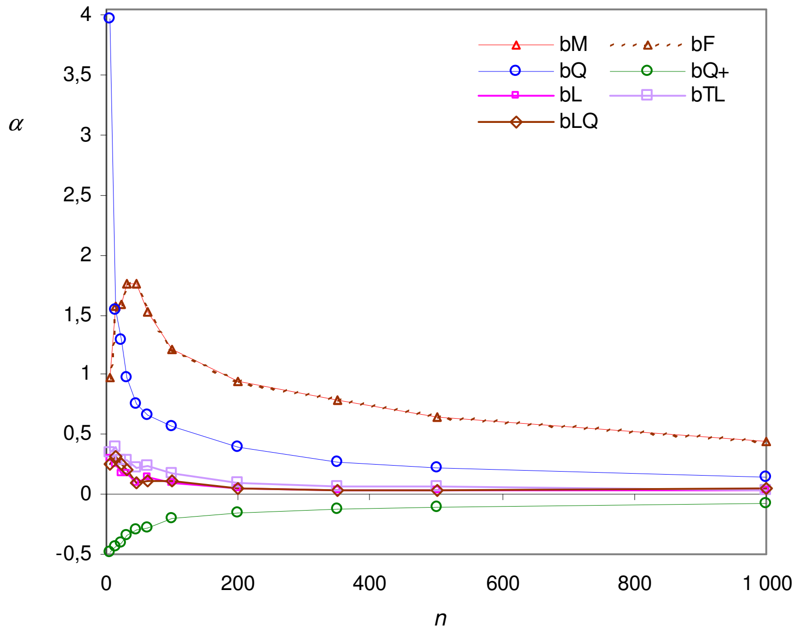
Obr. 6.64: Relativní směrodatná chyba estimátorů měr kurtozy při výběrech z normálního rozdělení v závislosti na rozsahu výběru.

6.4.1.2 Tvar výběrového rozdělení

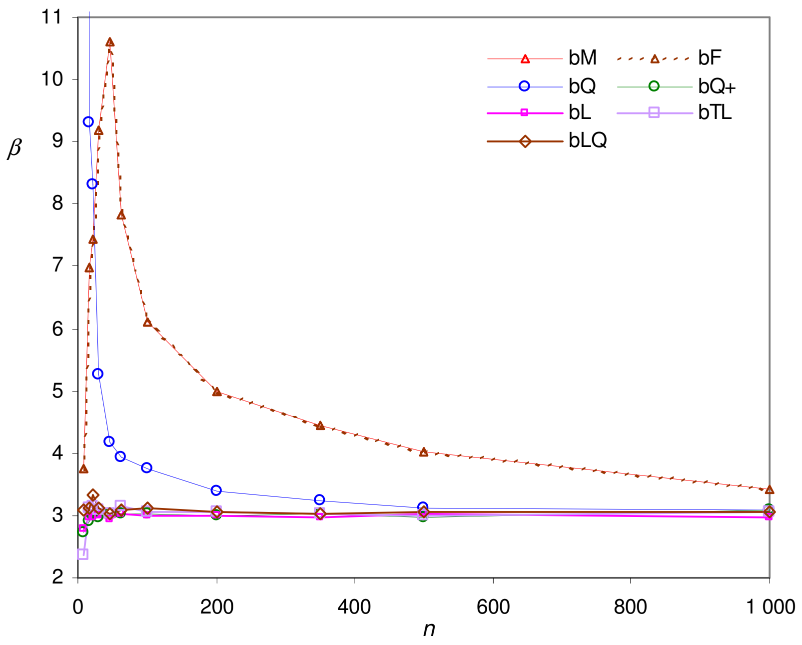
Při porovnání estimátorů b_M a b_F z hlediska tvaru jejich výběrových rozdělení docházíme ke zcela shodným výsledkům. Veškeré níže uvedené poznatky o tvaru rozdělení estimátoru b_M tedy platí také pro estimátor b_F.

Při porovnání estimátorů z hlediska **koeficientu šikmosti** jejich výběrových rozdělení lze konstatovat, že (s výjimkou výběrů o rozsahu $n = 7$) největší rozdíl oproti normální hodnotě 0 vykazuje estimátor b_M, u něhož lze pozorovat kladné zešikmení. Naopak nejrychlejší konvergenci k normální hodnotě vykazují estimátory b_L, b_{LQ} a b_{TL}. Za zmínku dále stojí počáteční nárůst zešikmení estimátoru b_M v závislosti na rozsahu výběru pro $n \leq 47$, což je zřejmě obdoba jevu pozorovaného u estimátoru a_M (obr. 6.65).

Z hlediska hodnot **koeficientu špičatosti** vykazuje (s výjimkou výběrů o rozsahu $n = 7$) největší rozdíly oproti normální hodnotě opět estimátor b_M a pro rozsahy výběru přibližně do $n \leq 350$ také estimátor b_Q. U ostatních estimátorů je zřejmá rychlá konvergence k normální hodnotě 3. Opět lze pozorovat anomálii u estimátorů b_M a b_F v podobě počáteční divergence od normální hodnoty pro $n \leq 47$ (obr. 6.66).



Obr. 6.65: Průběh odhadnutého koeficientu šikmosti estimátorů měr kurtozy v závislosti na rozsahu výběru.



Obr. 6.66: Průběh odhadnutého koeficientu špičatosti estimátorů měr kurtozy v závislosti na rozsahu výběru.

Otestování normality výběrových rozdělení

Testy Kolmogorova – Smirnova o normalitě výběrových rozdělení estimátorů měr kurtozy (tabulka 6.18) dávají obdobné výsledky jako hodnoty koeficientů šikmosti a špičatosti výběrových rozdělení: největší odlišnosti od normality vykazuje rozdělení estimátorů momentové míry, naopak nejrychlejší konvergenci k normalitě vykazují estimátory b_L, b_{LQ} a b_{TL}.

Tabulka 6.18: Výsledky testů Kolmogorova – Smirnova o normalitě výběrových rozdělení estimátorů měr kurtozy při výběrech z normálního rozdělení. (V tabulce je uvedena vždy nejvyšší hladina významnosti, na které výběrové rozdělení testu vyhovělo.)

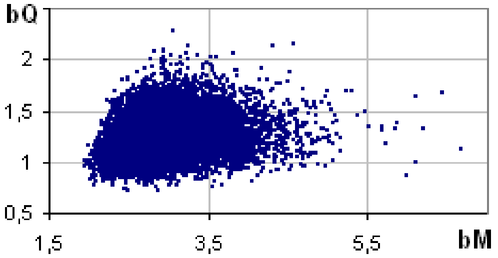
n	b_M	b_Q	b_Q^+	b_L	b_{TL}	b_{LQ}
- 31	-	-	-	-	-	-
47	-	-	-	1 %	-	10 %
63	-	-	-	10 %	-	1 %
100	-	-	-	5 %	-	10 %
200	-	-	1 %	10 %	10 %	10 %
350	-	-	5 %	10 %	10 %	10 %
500	-	-	1 %	10 %	10 %	10 %
1000	-	1 %	10 %	5 %	10 %	10 %

6.4.1.3 Měří všechny charakteristiky stejnou vlastnost rozdělení?

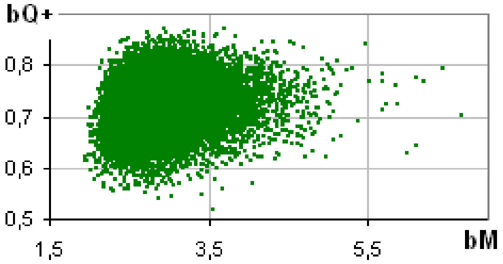
Nejvyšší sílu závislosti s hodnotami estimátoru momentové míry kurtozy vykazuje estimátor míry založené na L-momentech b_L. Všechny ostatní míry vykazují sílu závislosti s b_M výrazně nižší.

Tabulka 6.19: Hodnoty korelačních koeficientů mezi estimátory b_M a b_Q, b_Q^+, b_L, b_{TL} a b_{LQ} při výběrech z normálního rozdělení o rozsahu n = 100.

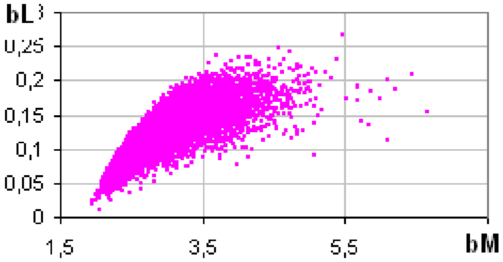
Estimátor	**Korelační koeficient**
b_Q	0,21
b_Q^+	0,15
b_L	0,79
b_{TL}	0,41
b_{LQ}	0,25



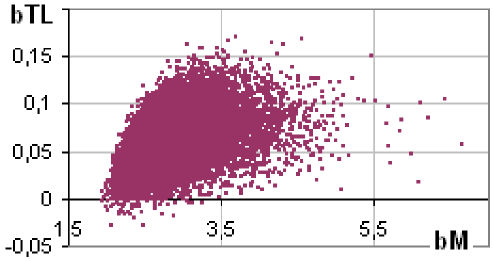
Obr. 6.67: Průběh závislosti hodnot estimátorů měr kurtozy b_M a b_Q při výběrech z normálního rozdělení o rozsahu n = 100



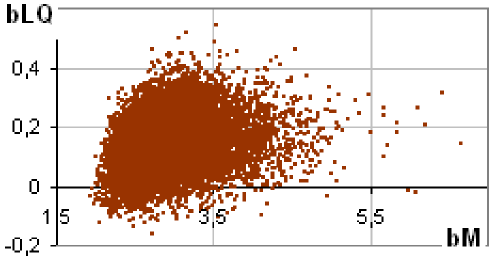
Obr. 6.68: Průběh závislosti hodnot estimátorů měr kurtozy b_M a b_Q^+ při výběrech z normálního rozdělení o rozsahu n = 100



Obr. 6.69: Průběh závislosti hodnot estimátorů měr kurtozy b_M a b_L při výběrech z normálního rozdělení o rozsahu n = 100



Obr. 6.70: Průběh závislosti hodnot estimátorů měr kurtozy b_M a b_{TL} při výběrech z normálního rozdělení o rozsahu n = 100



Obr. 6.71: Průběh závislosti hodnot estimátorů měr kurtozy b_M a b_{LQ} při výběrech z normálního rozdělení o rozsahu n = 100

6.4.1.4 Verifikace dosavadních poznatků o chování estimátorů

Porovnání hodnot směrodatné chyby, vychýlení a koeficientů šikmosti a špičatosti výběrového rozdělení estimátoru b_M, získaných na základě experimentu, s aproximací teoretických hodnot vypočtených podle vztahů uvedených v kapitole 2 (dle PEARSON, 1930) je uvedeno v tabulce 6.20. V tabulce není veden rozsah výběru $n = 7$, protože pro něj výše uvedené teoreticky odvozené aproximace nedávají smysluplné hodnoty. Z porovnání je patrné, že aproximace hodnot směrodatné chyby a vychýlení je vyhovující již pro $n = 47$, zatímco aproximace koeficientu špičatosti vyhovuje až pro $n \geq 350$ a aproximace koeficientu šikmosti vyhovuje pouze pro rozsahy výběru $n \doteq 200$.

Tabulka 6.20: Teoretické a empirické hodnoty směrodatné chyby, vychýlení a koeficietů šikmosti a špičatosti výběrového rozdělení estimátoru b_M při výběrech z normálního rozdělení. U nevyplněných polí tabulky nedává aproximace smysluplné hodnoty.

n	**Teoretické hodnoty (viz kap. 2)**				**Empirické hodnoty**			
	$SE_T(b_M)$	**$B_T(b_M)$**	**$\alpha_T(b_M)$**	**$\beta_T(b_M)$**	**$SE(b_M)$**	**$B(b_M)$**	**$\alpha(b_M)$**	**$\beta(b_M)$**
15	0,7686	-0,3750	-	-	0,8549	-0,3804	1,57	6,98
23	0,7417	-0,2500	-	-	0,7683	-0,2434	1,58	7,44
31	0,6937	-0,1875	2,43	-	0,7196	-0,1866	1,76	9,20
47	0,6105	-0,1250	2,50	14,48	0,6322	-0,1140	1,75	10,61
63	0,5486	-0,0938	2,19	11,57	0,5654	-0,0871	1,52	7,82
100	0,4547	-0,0594	1,63	8,40	0,4580	-0,0531	1,22	6,11
200	0,3337	-0,0299	0,94	5,70	0,3329	-0,0312	0,94	5,00
350	0,2563	-0,0171	0,57	4,54	0,2586	-0,0120	0,78	4,46
500	0,2158	-0,0120	0,41	4,08	0,2168	-0,0120	0,65	4,01
1000	0,1538	-0,0060	0,21	3,54	0,1515	-0,0083	0,44	3,42

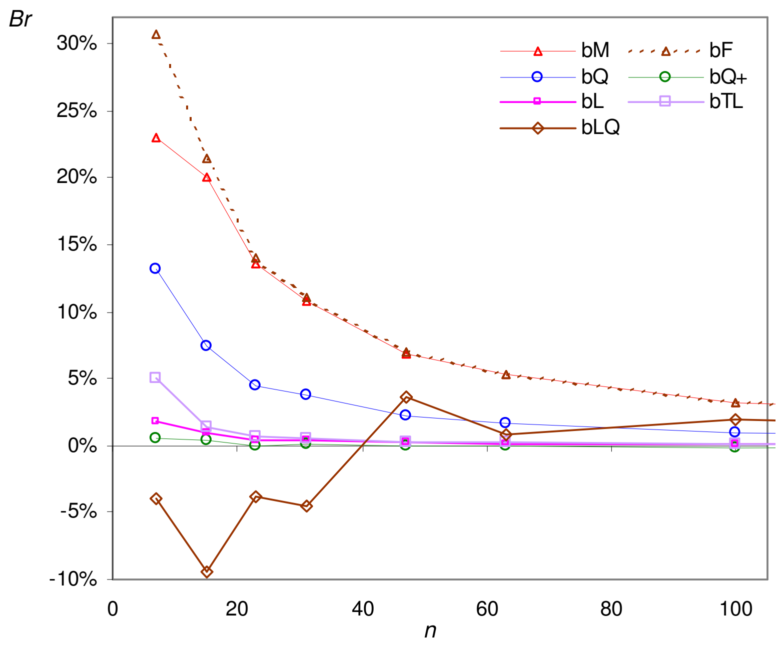
Z empiricky získaných hodnot výrazu $SE(\alpha_M)\sqrt{n}$ v tabulce v příloze č. 3 je vidět, že limitní vztah pro směrodatnou chybu estimátoru b_M uvedený v kapitole 2 (dle KENDAL, STUART, 1963) platí přibližně až pro rozsahy výběru $n \geq 350$. Oproti tomu limitní vztah pro směrodatnou chybu estimátoru Moorsova kvantilového koeficientu b_Q (WAGEMAKERS et al., 1995) platí již pro $n \geq 23$.

6.4.2 Výběry z ostatních rozdělení

6.4.2.1 Porovnání estimátorů z hlediska variability, vychýlení a tvaru rozdělení

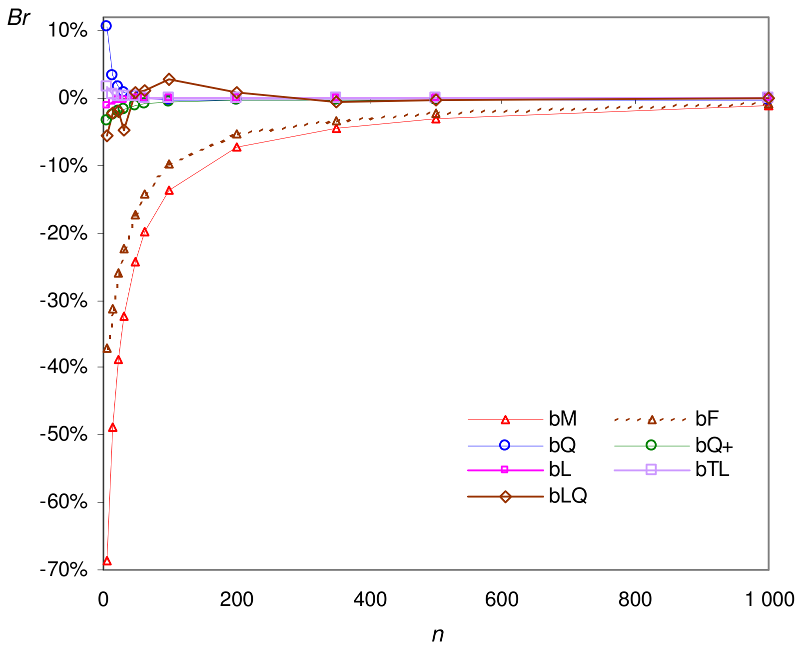
Při porovnání chování jednotlivých estimátorů v závislosti na typu rozdělení, z něhož výběr pochází, se ukazuje, že jejich výběrové vlastnosti závisejí na velikosti kurtozy v tomto rozdělení vzhledem k rozdělení normálnímu. Při **výběrech z populací s přibližně normální hodnotou kurtozy** dostáváme obdobné výsledky, jako v případě výběrů z normálního rozdělení. Jedinou odlišností je, že estimátor b_F při hodnotách $\beta_M > 3$ v populaci mírně podhodnocuje, naopak při $\beta_M < 3$ mírně nadhodnocuje. Estimátor b_M podhodnocuje v obou případech.

Pro **populace s nižší než normální hodnotou kurtozy** byly zjištěny odlišnosti od výběrových vlastností v případě výběrů z normálního rozdělení pouze u estimátorů momentové míry: S klesající úrovní kurtozy roste u estimátorů b_M i b_F jejich vychýlení (oba estimátory podhodnocují). Míra vychýlení i relativní směrodatné chyby je u estimátoru b_F vždy mírně vyšší než u b_M, přičemž rozdíly mezi oběma estimátory jsou pro rozsahy výběru $n \geq 63$ téměř zanedbatelné (obr. 6.72). Anomálie spojená s počátečním nárůstem směrodatné odchylky estimátoru b_M pro malé rozsahy výběru se neprojevuje. Projevuje se však anomálie spojená s divergencí koeficientů šikmosti a špičatosti výběrových rozdělení od normálních hodnot, ale pouze pro rozsahy výběru $n \leq 15$. Pořadí jednotlivých estimátorů z hlediska relativní směrodatné odchylky i relativní směrodatné chyby vychází stejně jako u normálního rozdělení.

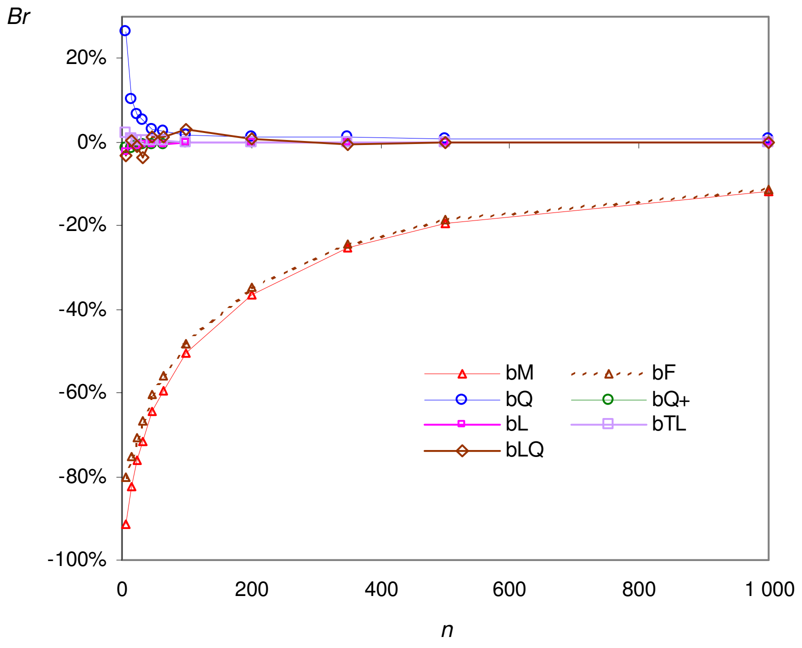


Obr. 6.72: Průběh relativního vychýlení estimátorů měr kurtozy při výběrech z rovnoměrného rozdělení (β_M = 1,8) pro výběry o rozsahu n ≤ 100

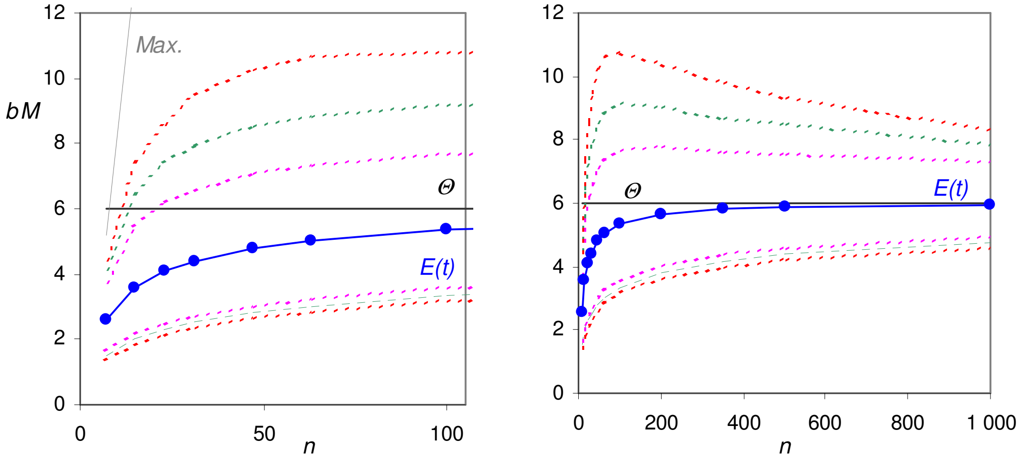
V případě výběrů z **populací s vyšší než normální mírou kurtozy** lze pozorovat výrazné odlišnosti ve srovnání s výběry z normálního rozdělení opět pouze u estimátorů momentové míry b_M a b_F. Při porovnání z hlediska **relativního vychýlení** zjišťujeme, že oba estimátory systematicky podhodnocují, přičemž velikost vychýlení se při stoupající míře kurtozy v populaci zvyšuje. Konvergence empirických středních hodnot estimátorů momentové míry k jejich limitní střední hodnotě je přitom velmi pomalá. Na obrázku 6.73 resp. 6.74 je uveden průběh závislosti vychýlení na rozsahu výběru n při výběrech z populací s hodnotou $\beta_M = 6{,}0$ resp. $\beta_M = 21{,}7$. Na obrázcích 6.75 a 6.74 je dále uveden průběh empirické střední hodnoty a význačných kvantilů výběrových rozdělení estimátoru b_M při výběrech z těchto populací s vyznačením horní meze oboru možných hodnot estimátoru b_M. Je z nich patrné, že např. při výběrech z populace s $\beta_M = 21{,}7$ o rozsahu $n = 100$ ve více nachází více než 95 % odhadů pod skutečnou hodnotou.



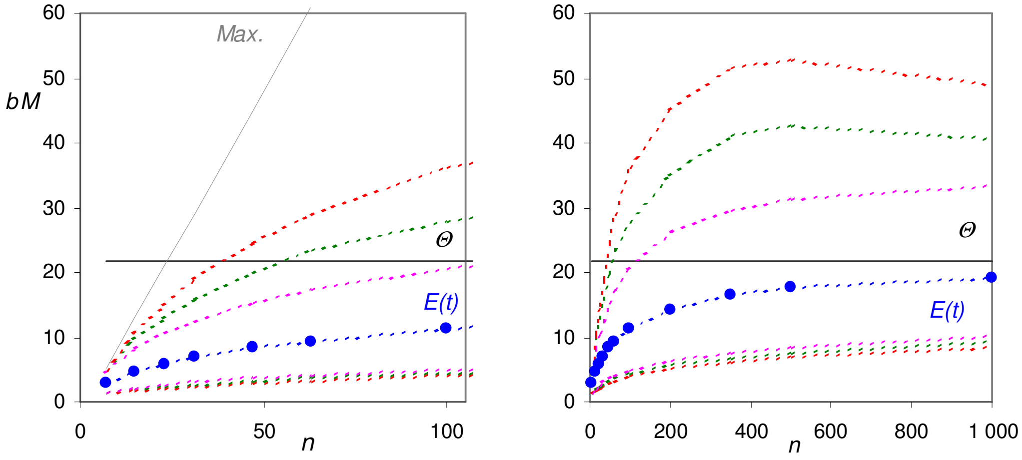
Obr. 6.73: Průběh relativního vychýlení estimátorů měr kurtozy při výběrech z Laplaceova rozdělení ($\beta_M = 6$)



Obr. 6.74: Průběh relativního vychýlení estimátorů měr kurtozy při výběrech ze souboru „Délka hovorů" s hodnotou $\beta_M = 21{,}7$



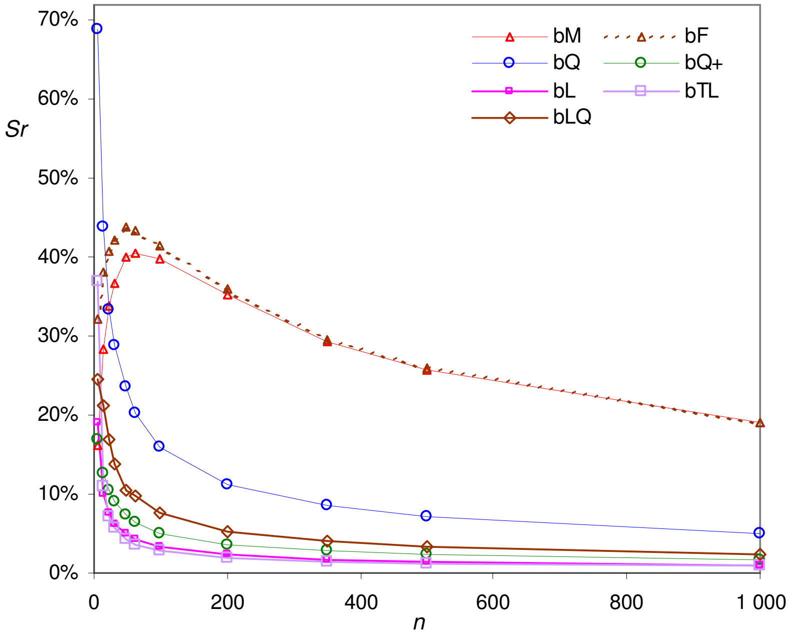
Obr. 6.75: Průběhy empirické střední hodnoty E(t) a význačných kvantilů estimátoru b_M v porovnání se skutečnou hodnotou $\beta_M = 6$ v populaci (v grafu jako Θ) v závislosti na rozsahu výběru při výběrech z Laplaceova rozdělení. Přerušované křivky v pořadí shora: 97,5%, 95%, 90%, 10%, 5% a 2,5% kvantil.



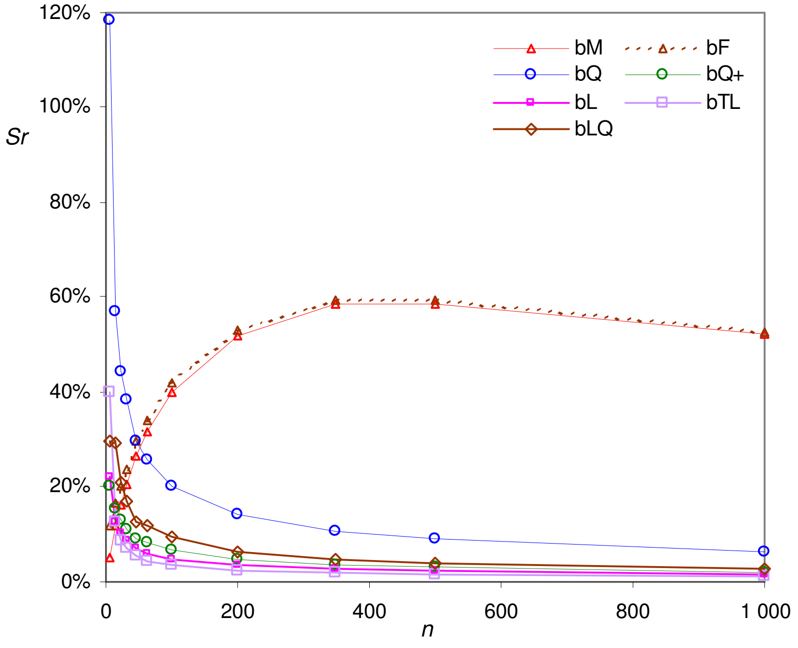
Obr. 6.76: Průběhy empirické střední hodnoty E(t) a význačných kvantilů estimátoru b_M v porovnání se skutečnou hodnotou β_M v populaci (v grafu jako Θ) v závislosti na rozsahu výběru při výběrech ze souboru „Délka hovorů“ s vysokou úrovní kurtozy ($\beta_M = 21{,}7$). Přerušované křivky v pořadí shora: 97,5%, 95%, 90%, 10%, 5% a 2,5% kvantil.

Obdobně jako v případech výběrů z normálního rozdělení je patrná anomálie v podobě počátečního nárůstu **směrodatné odchylky** výběrového rozdělení estimátoru b_M v závislosti na rozsahu výběru n, která se nyní projevuje i u výběrového rozdělení estimátoru b_F. S rostoucí úrovní kurtozy v populaci se přitom zvyšuje hodnota rozsahu

výběru n, do které směrodatná odchylka výběrových rozdělení roste. Například při výběrech z Laplaceova rozdělení s $\beta_M = 6$ směrodatná odchylka estimátoru b_M v závislosti na rozsahu výběru roste pro $n \leq 100$ (obr. 6.77), zatímco při výběrech z populace s $\beta_M = 21{,}7$ byl pozorován nárůst směrodatné odchylky až pro $n \leq 500$ (obr. 6.78). Při porovnání ostatních estimátorů z hlediska relativní směrodatné odchylky vychází nejlépe estimátory b_L a b_{TL}, nejhůře naopak estimátor b_Q.

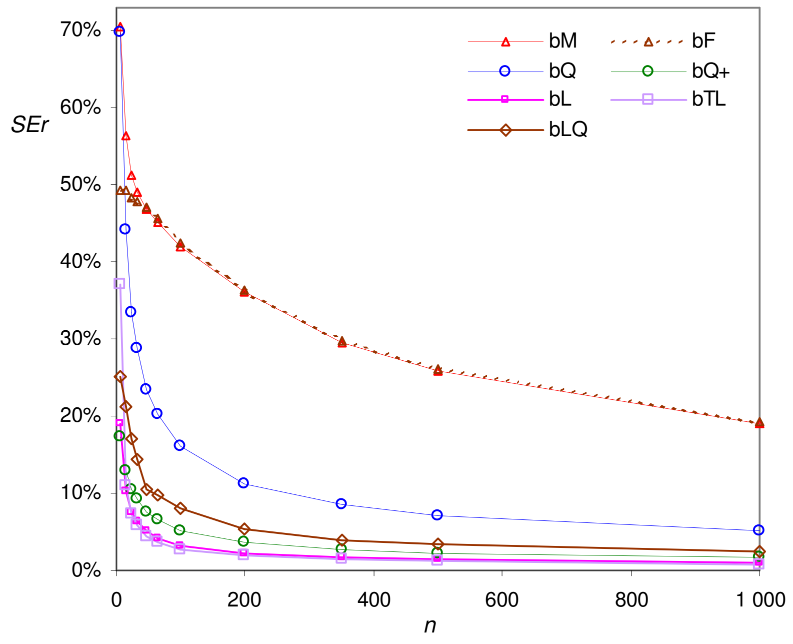


Obr. 6.77: Průběh relativní směrodatné odchylky estimátorů měr kurtozy při výběrech z Laplaceova rozdělení ($\beta_M = 6$)

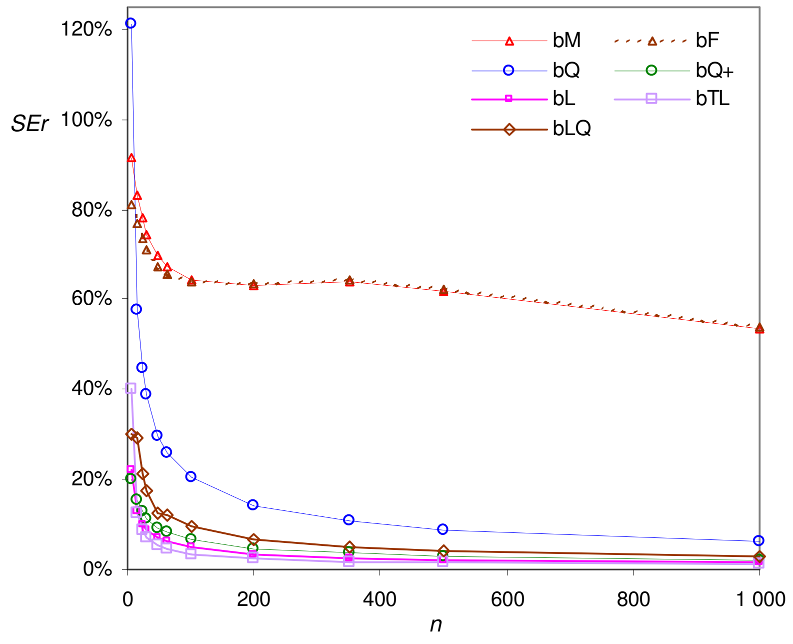


Obr. 6.78: Průběh relativní směrodatné odchylky estimátorů měr kurtozy při výběrech ze souboru „Délka hovorů“ s vysokou mírou kurtozy ($\beta_M = 21{,}7$)

Velikost vychýlení estimátorů b_M a b_F a anomálie v průběhu jejich směrodatné odchylky mají za následek, že jejich **relativní směrodatná chyba** konverguje k nule výrazně pomaleji, než v případě výběrů z normální populace (obr. 6.79). Rychlost konvergence směrodatné chyby k nule se s rostoucí mírou kurtozy v populaci snižuje. V některých případech (například při výběrech z populace s hodnotou $\beta_M = 21{,}7$, viz obr. 6.80) lze směrodatnou chybu pro vyšší rozsahy výběru považovat za **konstantní**, pro některé rozsahy výběru dokonce směrodatná chyba v závislosti na n mírně roste. Při porovnání ostatních estimátorů z hlediska velikosti relativní směrodatné chyby docházíme ke stejnému pořadí jako při jejich porovnání z hlediska relativní směrodatné odchylky. Nejlépe tedy vychází estimátory b_L a b_{TL}, nejhůře naopak estimátor b_Q.

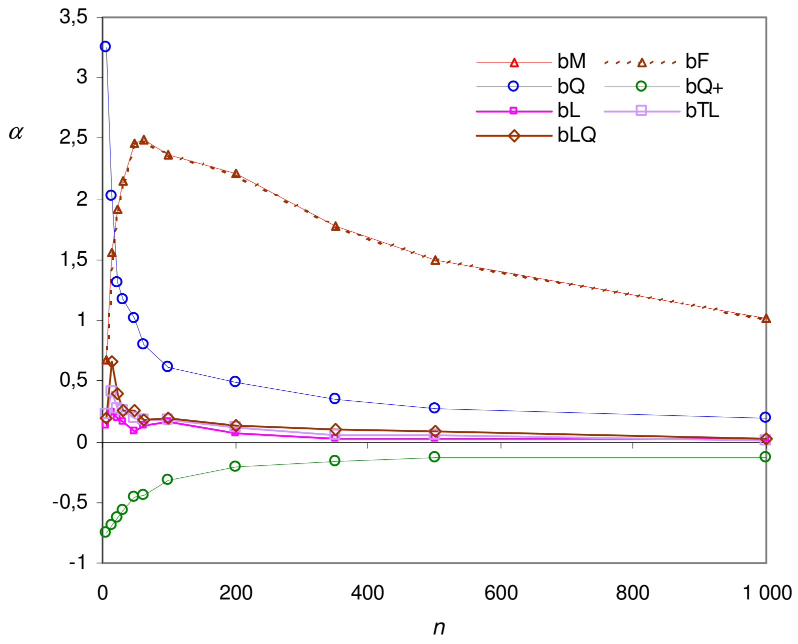


Obr. 6.79: Průběh relativní směrodatné chyby estimátorů měr kurtozy při výběrech z Laplaceova rozdělení ($\beta_M = 6$)



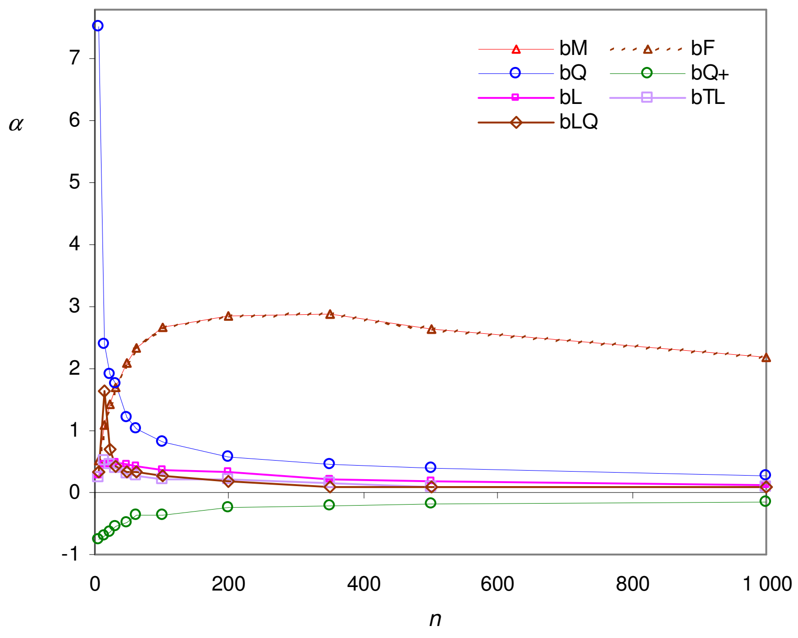
Obr. 6.80: Průběh relativní směrodatné chyby estimátorů měr kurtozy při výběrech ze souboru „Délka hovorů" s vysokou mírou kurtozy ($\beta_M = 21{,}7$)

Při zhodnocení estimátorů z hlediska **tvaru** jejich **výběrového rozdělení** se, stejně jako v případě výběrů z normálního rozdělení, projevuje anomálie v podobě divergence koeficientů šikmosti a špičatosti od normálních hodnot, která se zvyšuje v závislosti na hodnotě kurtozy v populaci. Na obrázku 6.81 a 6.82 je zobrazen průběh hodnot koeficientu šikmosti výběrových rozdělení jednotlivých estimátorů v případě výběrů z populace s hodnotou $\beta_M = 21,7$, u něhož se tato divergence od normálních hodnot projevuje přibližně pro $n \leq 350$. Při porovnání **ostatních estimátorů** měr kurtozy lze na základě výsledků testů Kolmogorova – Smirnova konstatovat, že rychlost konvergence jejich výběrových rozdělení k normalitě se s rostoucí mírou kurtozy zpomaluje. Platí však, že nejrychlejší konvergenci k normalitě lze pozorovat u estimátoru b_L.

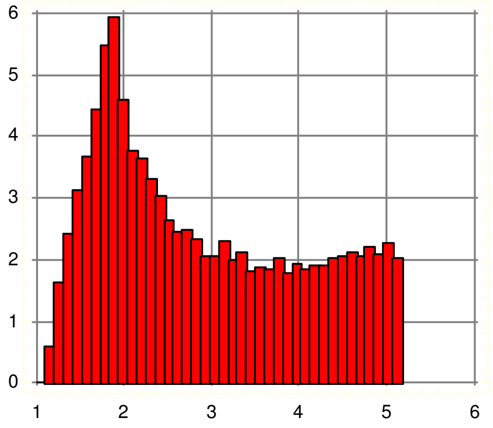


Obr. 6.81: Průběh odhadnutého koeficientu šikmosti estimátorů měr kurtozy při výběrech z Laplaceova rozdělení ($\beta_M = 6$)

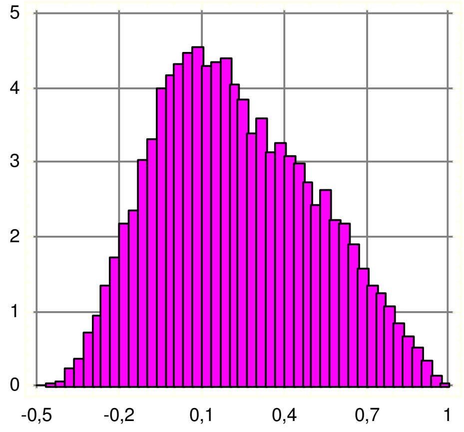
Při výběrech z rozdělení s extrémně vysokou hodnotou kurtozy lze pro malé rozsahy výběru pozorovat, že výběrové rozdělení estimátoru b_M je bimodální (obr. 6.83), což je obdobná anomálie, jakou je možné pozorovat v případě estimátoru a_M při výběrech z extrémně zešikmených populací. Pro srovnání uvádím též výběrové rozdělení estimátoru b_L pro stejnou populaci a rozsah výběru (obr. 6.84).



Obr. 6.82: Průběh odhadnutého koeficientu šikmosti estimátorů měr kurtozy při výběrech ze souboru „Délka hovorů“ s vysokou mírou kurtozy (β_M = 21,7)



Obr. 6.83: Empirické výběrové rozdělení estimátoru b_M při výběrech z rozdělení LN(1) s hodnotu β_M = 114 při rozsahu výběru n = 7.



Obr. 6.84: Empirické výběrové rozdělení estimátoru b_L při výběrech z rozdělení LN(1) s hodnotu β_M = 114 při rozsahu výběru n = 7.

6.4.2.2 Měří všechny charakteristiky stejnou vlastnost rozdělení?

Při porovnání estimátorů z hlediska síly jejich závislosti s estimátory momentové míry lze pozorovat, že s růstem úrovně kurtozy v populaci síla této závislosti rychle klesá a s poklesem kurtozy naopak roste. Při výběrech z populací s výrazně vyšší než normální hodnotou kurtozy lze konstatovat, že s výjimkou estimátoru b_L se jedná o faktickou nezávislost (tab. 6.21).

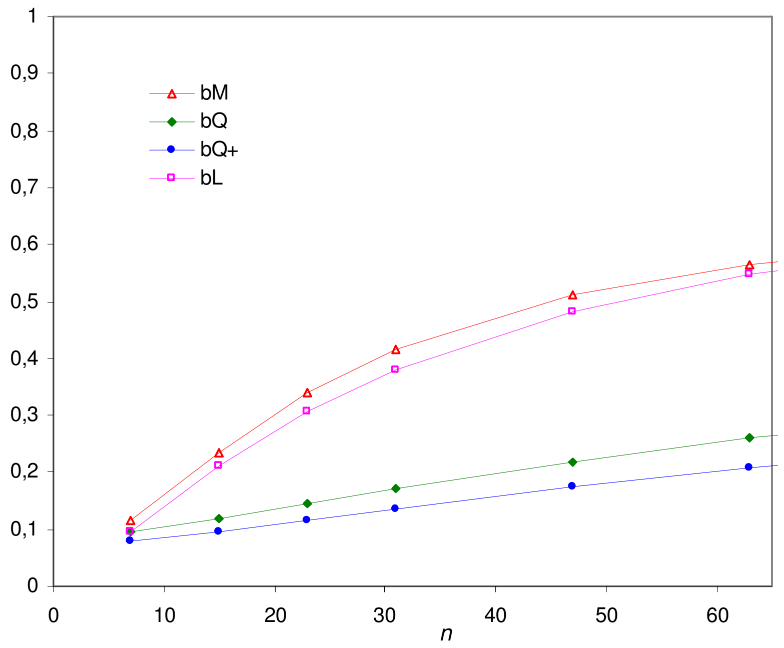
Tabulka 6.21: Hodnoty korelačních koeficientů mezi estimátory b_M a b_Q, b_Q^+, b_L, b_{TL} a b_{LQ} při výběrech z Laplaceova rozdělení (β_M = 6) o rozsahu n = 100.

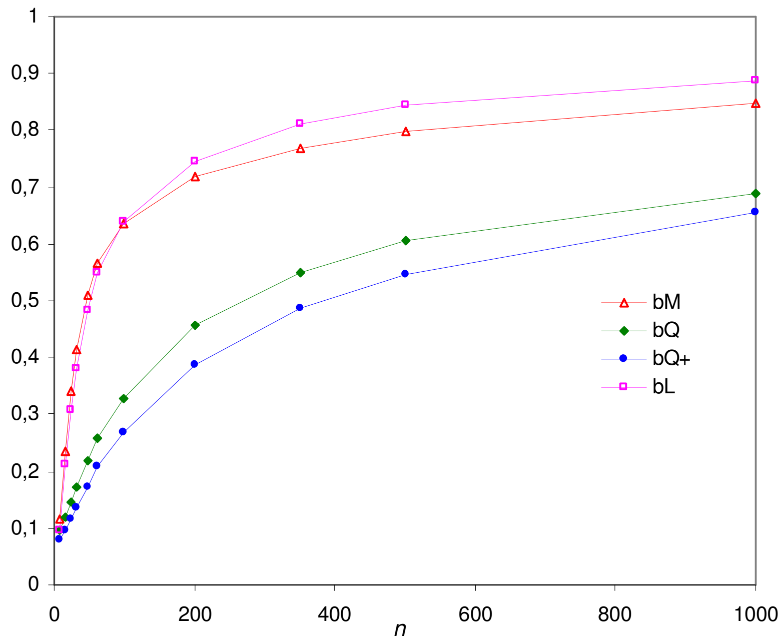
Estimátor	Korelační koeficient
b_Q	0,01
b_Q^+	-0,02
b_L	0,63
b_{TL}	0,14
b_{LQ}	0,05

6.4.2.3 Síla testu

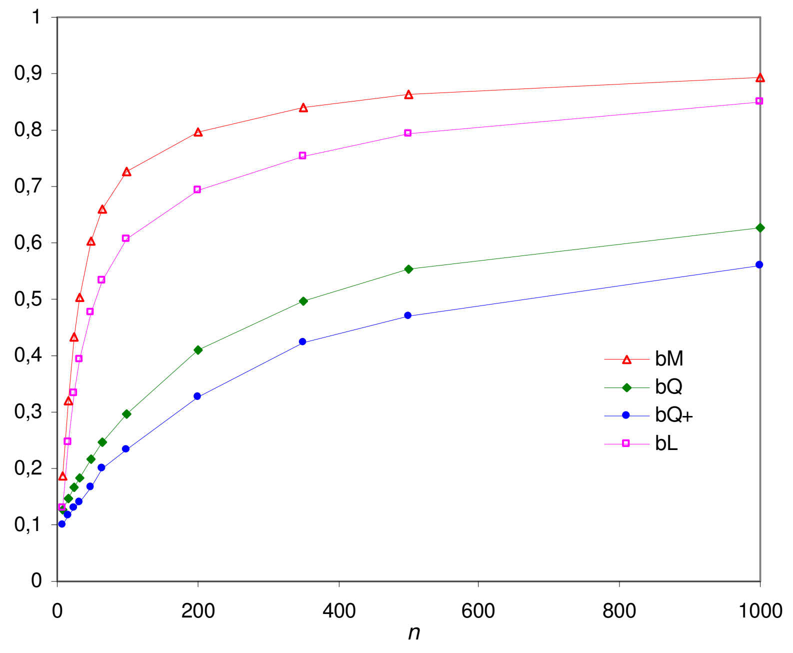
Při porovnání estimátorů měr kurtozy z hlediska průměrné síly jejich testů o tvaru rozdělení základního souboru lze vyvodit poněkud odlišné závěry než v případě estimátorů měr relativní variability a šikmosti. Pro výběry o rozsahu $n \leq 100$ dosahují nejlepších výsledků estimátory momentové míry a nepatrně horších estimátor míry založené na L-momentech. Při výběrech většího rozsahu naopak nejlepších výsledků dosahuje estimátor b_L. Nejhorších výsledků dosahuji estimátory měr založených na kvantilech, a to bez ohledu na rozsah výběru. Skutečnost, že pro nižší rozsahy výběru se jako nejefektivnější pro testování tvaru rozdělení základního souboru jeví estimátor momentové míry, je poněkud překvapivá, vezmeme-li v úvahu jeho výběrové vlastnosti zejména z hlediska vychýlení, které je nevyšší právě pro nižší rozsahy výběru.

Pokud se zaměříme pouze na **testy o normalitě rozdělení základního souboru**, vychází jako nejefektivnější estimátor b_M, pro všechny rozsahy výběru, přičemž b_M dává výsledky mírně horší (obr. 6.86). Lepší výsledky estimátoru b_M oproti b_L jsou patrné také z grafu na obr. 6.87, ve kterém jsou proti sobě vyneseny konkrétní hodnoty síly testu založeného na b_M a b_L. Jak jsem však uvedl v předchozí podkapitole, pro učinění zobecňujících závěrů o síle testů o normalitě by bylo třeba provést rozsáhlejší zkoumání.

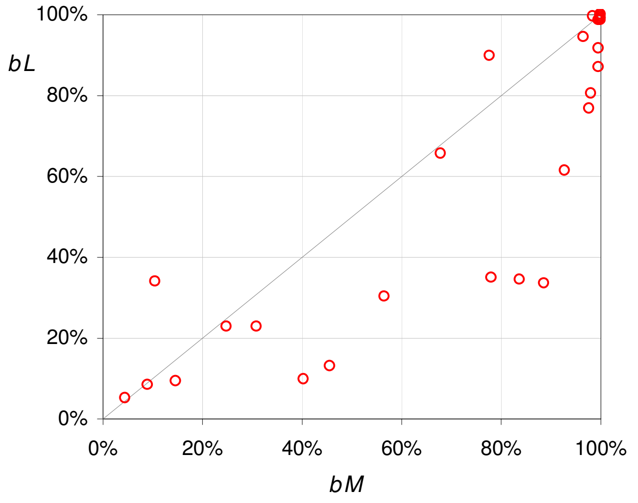




Obr. 6.85: Průměrná síla testů o populaci, z níž výběr pochází, založených na estimátorech b_M, b_Q, b_Q^+, a b_L v závislosti na rozsahu výběru n. (Poznámka: estimátor b_F, který není v grafu uveden, dává stejné výsledky jako b_M.)



Obr. 6.86: Průměrná síla testů o normalitě rozdělení, z něhož výběr pochází, založených na estimátorech b_M, b_Q, b_Q^+ a b_L,.



Obr. 6.87: Odhadnutá průměrná síla testů o normalitě rozdělení založených na estimátorech b_M (na horizontální ose) a b_L (na vertikální ose) při výběrech o rozsahu n = 100 (Poznámka: velká část pozorování s odhadnutou sílou testu blízkou 100% splývá v pravém horním rohu grafu)

7. Závěry

Na základě Monte Carlo simulace bylo provedeno porovnání estimátorů klasických a některých nově navržených měr variability, relativní variability, šikmosti a kurtozy. Z celkem 40 různých populací bylo pořízeno vždy 16 000 výběrů o rozsahu $n \in \{7; 15; 23; 31; 47; 63; 100; 200; 350; 500; 1000\}$. Z každého výběru byly vypočteny hodnoty estimátorů měr založených na momentech, kvantilech, L-momentech a jejich modifikacích a robustních měr založených na mediánu funkce lineární kombinace pořádkových statistik. Na základě experimentu bylo provedeno porovnání výběrových vlastností jednotlivých estimátorů z hlediska variability, vychýlení a rychlosti konvergence jejich výběrových rozdělení k normalitě. U estimátorů vybraných charakteristik byla dále na základě experimentálně odhadnuté průměrné síly testu posouzena vhodnost jejich použití jako testového kritéria při testu o rozdělení, z něhož výběr pochází. Zároveň byla porovnána síla závislosti estimátorů nově navržených charakteristik s estimátory momentovými s cílem posoudit, zda je možné danou charakteristiku skutečně považovat za vhodnou alternativu charakteristiky momentové. Na základě provedeného experimentu byly dále pořízeny hodnoty význačných kvantilů výběrových rozdělení jednotlivých estimátorů pro výběry z normálního rozdělení.

Získané výsledky ukazují, že estimátory momentových měr jsou vyhovující pro popis souboru pouze při výběrech z populací nepříliš odlišných od normálního rozdělení. S rostoucí odlišností od normality rychle roste relativní variabilita i vychýlení jejich estimátorů. U estimátorů koeficientu šikmosti a špičatosti a v některých případech také u variačního koeficientu se projevují anomálie např. v podobě počátečního nárůstu směrodatné odchylky výběrového rozdělení v závislosti na rozsahu výběru n a velmi pomalé konvergence vychýlení k nule, které jsou zřejmě způsobeny existencí horních mezí jejich oborů hodnot závislých na rozsahu výběru n.

Výsledky experimentu dále ukazují, že vhodnou alternativou k mírám založeným na klasických momentech i kvantilech by se mohly stát míry založené na L-momentech. Jejich estimátory vykazují ve většině případů nejlepší výběrové vlastnosti z hlediska relativní variability, vychýlení i konvergence výběrového rozdělení k normalitě. Také v případě výběrů z normálního rozdělení vykazují lepší nebo stejně

dobré výsledky, jako estimátory měr momentových. Zároveň vykazují vysokou míru závislosti s hodnotami estimátorů momentových charakteristik, k nimž byly navrženy jako alternativa.

Určité zlepšení oproti estimátoru klasického variačního koeficientu resp. koeficientu šikmosti přinášejí též estimátory jejich standardizovaných alternativ v_X^+ resp. a_M^+ a také estimátor a_C. Neprojevují se u nich výše zmíněné anomálie pozorované u estimátorů klasických momentových charakteristik, vykazují nižší relativní směrodatnou chybu a zejména u estimátorů v_X^+ a a_M^+ lze pozorovat silnou závislost s estimátory měr momentových. Ve srovnání s estimátory měr založených na L-momentech je však jejich relativní směrodatná chyba, zejména v případě výběrů z populací více odlišných od normálního rozdělení, vyšší. Nevýhodou estimátoru a_C je též nespojitost jeho oboru hodnot.

Dále se ukazuje, že estimátory měr založených na kvantilech, na mediánu funkce lineární kombinace pořádkových statistik a také na modifikacích L-momentů (tj. na LQ- a TL-momentech) nejsou vhodnou alternativou k mírám momentovým. Nedochází u nich sice k anomáliím pozorovaným u estimátorů momentových měr, jejich relativní směrodatné chyby jsou však vesměs (s výjimkou estimátorů založených na TL-momentech) vyšší než u charakteristik založených na L-momentech. V případě výběrů z populací nepříliš odlišných od normálního rozdělení často dávají dokonce horší výsledky než estimátory měr momentových. Především však vykazují velmi slabou závislost s estimátory měr založených na momentech, k nimž byly navrženy jako alternativa. V případě měr šikmosti a kurtozy se často jedná dokonce o nezávislost. Síla závislosti těchto estimátorů s estimátory měr momentových přitom rychle klesá s rostoucí odlišností[3] příslušné populace od normálního rozdělení. Proto nelze použití měr založených na TL-momentech doporučit ani při výběrech z populací výrazně odlišných od normálního rozdělení, přestože jejich estimátory v těchto případech vykazují často nižší relativní směrodatnou chybu než estimátory měr L-momentových. V této práci navržená míra relativní variability V_{HL} se též ukázala jako nevhodná, neboť u ní byla pro určité typy populací pozorována dokonce negativní závislost s estimátorem momentové míry.

[3] Odlišnost z hlediska hodnot koeficientů šikmosti a špičatosti.

Při porovnání měr založených na kvantilech, momentech a L-momentech z hlediska vhodnosti použití jejich estimátorů jako testového kriteria bylo zjištěno, že nejlepších výsledků dosahují míry založené na L-momentech. I přes výše popsané nedostatky však míry založené na momentech dosahují spolu se standardizovanými momentovými mírami V_X^+ a α_M^+ jen mírně horších výsledků než míry založené na L-momentech. V případě měr kurtozy při výběrech malého rozsahu dosahují dokonce výsledků mírně lepších.

Z uvedených výsledků vyplývají následující doporučení pro praxi: (i) Pro popis datového souboru jsou jednoznačně nejvhodnější charakteristiky založené na L-momentech, a to bez ohledu na typ populace, ze které výběr pochází. Pro účely popisu je tedy můžeme považovat za univerzální. (ii) Pro účely testování hypotéz o hodnotě určité popisné charakteristiky v populaci lze pro variabilitu, relativní variabilitu a šikmost doporučit estimátory měr založených na L-momentech. Pro kurtozu lze pro nižší rozsahy výběru (přibližně pro $n < 100$) doporučit běžně používaný estimátor momentové míry, pro výběry vyššího rozsahu vychází jako nejlepší opět estimátor míry založené na L-momentech.

Pro zavedení charakteristik založených na L-momentech do statisticko-ekonomické praxe bude třeba vyřešit několik dalších problémů: (i) Jednak bude třeba formulovat obecně platná doporučení, jak interpretovat konkrétní hodnotu vypočtené charakteristiky (např. obdobu pravidla 2σ nebo 6σ platného směrodatnou odchylku). (ii) Nutnou podmínkou pro jejich praktické využívání je zahrnutí procedur pro jejich výpočet ve standardních statistických paketech. (iii) Závěrem je třeba připomenout, že dosud nejsou, s výjimkou druhého L-momentu, navrženy vztahy pro intervalové odhady těchto charakteristik a nejsou navrženy ani estimátory těchto charakteristik pro složitější uspořádání výběru. Tento problém ovšem dosud přetrvává také u mnohých estimátorů charakteristik založených na klasických momentech nebo na kvantilech.

Seznam symbolů

α_C, a_C koeficient šikmosti založený počtech nadprůměrných a podprůměrných hodnot, resp. jeho estimátor

α_L, a_L koeficient šikmosti založený na L-momentech (L-skweness), resp. jeho estimátor

α_{LQ}, a_{LQ} koeficient šikmosti založený na LQ-momentech (LQ-skweness) pro $p = \alpha = 0{,}25$, resp. jeho estimátor

α_M, a_M (momentový) koeficient šikmosti, resp. jeho estimátor

α_M^+, a_M^+ standardizovaný koeficient šikmosti, resp. jeho estimátor

α_{MC}, a_{MC} medcouple (míra šikmosti založená na mediánu funkce lineární kombinace pořádkových statistik), resp. jeho estimátor

α_Q, a_Q Bowleyův (kvantilový) koeficient šikmosti, resp. jeho estimátor

α_{TL}, a_{TL} koeficient šikmosti založený na TL-momentech (TL-skweness) pro $t = 1$, resp. jeho estimátor

b_F Fisherův „téměř nestranný“ odhad hodnoty $\beta_M - 3$

β_L, b_L míra kurtozy založená na L-momentech (L-kurtosis), resp. její estimátor

β_{LQ}, b_{LQ} míra kurtozy založená na LQ-momentech (LQ- kurtosis) pro $p = \alpha = 0{,}25$, resp. její estimátor

β_M, b_M (momentový) koeficient špičatosti, resp. jeho estimátor

β_Q, b_Q Moorsův kvantilový koeficient špičatosti, resp. jeho estimátor

β_Q^+, b_Q^+ standardizovaný Moorsův kvantilový koeficient špičatosti, resp. jeho estimátor

$B(t)$ vychýlení estimátoru t

$B_r(t)$ relativní vychýlení estimátoru t

β_{TL}, b_{TL} míra kurtozy založená na TL-momentech (TL-kurtosis) pro $t = 1$, resp. její estimátor

$D(t)$ rozptyl estimátoru t

$MSE(t)$ střední čtvercová chyba estimátoru t

$S(t)$ směrodatná odchylka estimátoru t

S_{AAD}	průměrná absolutní odchylka od průměru
S_{AADM}	průměrná absolutní odchylka od mediánu
$SE(t)$	směrodatná chyba estimátoru t
$SE_r(t)$	relativní směrodatná chyba estimátoru t
S_{HL}	míra variability odvozená od Hodges – Lehmannovy míry polohy
S_L	druhý L-moment
S_{LQ}	druhý LQ-moment (pro $p = \alpha = 0{,}25$)
S_Q	kvartilová odchylka
S_{QE}	modifikované kvazirozpětí
$S_r(t)$	relativní směrodatná odchylka estimátoru t
S_{TL}	druhý TL-moment (pro $t = 0{,}25$)
S_X	směrodatná odchylka
V_{AAD}	poměrná absolutní odchylka od průměru
V_{AADM}	poměrná absolutní odchylka od mediánu
V_{HL}	míra relativní variability odvozená od Hodges – Lehmannovy míry polohy
V_L	míra relativní variability založená na L-momentech
V_{LQ}	míra relativní variability založená na LQ-momentech (pro $p = \alpha = 0{,}25$)
V_Q	poměrná kvartilová odchylka
V_{TL}	míra relativní variability založená na TL-momentech (pro $t = 1$)
V_X	variační koeficient
V_X^+	standardizovaný variační koeficient
Θ	skutečná hodnota určité popisné charakteristiky v populaci

Literatura

1. ACKLAM, P. J.: An algorithm for computing the inverse normal cumulative distribution function. http://home.online.no/~pjacklam/notes/invnorm/
2. ANTOCH J., VORLÍČKOVÁ, D.: *Vybrané metody statistické analýzy dat.* Praha, Academia 1992.
3. BICKEL, P.J., LEHMANN, E. L.: *Descriptive statistics for non-parametric models III: Dispersion.* The Annals of Statistics 4, 1139-1158. 1976.
4. BONETT, D. G.: *Confidence interval for a coefficient of quartile variation.* Computational Statistics & Data Analysis 50, č. 11, 2953-2957. 2006.
5. BONETT, D. G.: *Approximate confidence interval for standard deviation of nonnormal distributions.* Computational Statistics & Data Analysis 46, č. 3, 775-782. 2005.
6. BRYS, G., HUBERT, M., STRUYF, A.: *A Robust Measure of Skewness.* Journal of Computational and Graphical Statistics 13, č. 4, 996-1017. American Statistical Association, Institute of Mathematical Statistics, and Interface Foundation of North America 2004.
7. ČERMÁK, V.: *Diskrétní a spojitá rozdělení.* Praha, VŠE 1993.
8. ČERMÁK, V.: *Měří koeficient šikmosti souměrnost?* Statistická revue sv. 4, 211-220. Praha, Výzkumný ústav sociálně ekonomických informací při FSÚ 1975.
9. ČERMÁK, V.: *Srovnání některých relativních měr variability z hlediska jejich průběhu v rámci oboru možných hodnot.* Statistická revue sv. 9, 253-297. Praha, SEVT 1985.
10. ČERMÁK, V.: *Srovnání některých měr nesouměrnosti rozdělení z hlediska jejich chování v rámci intervalu možných hodnot.* Statistická revue sv. 10, 109-157. Praha, SEVT 1990.
11. ČERMÁK, V., HENZLER, J.: *Ještě jednou o koeficientu šikmosti.* Statistická revue 9, 229–312. Praha, SEVT 1985.
12. ČERMÁK, V., LAUBER, J.: *Sampling Distributions of Some Measures of Rel-variability: Results of a MC-simulation.* Acta Oeconomica Pragensia 4, č. 7. Praha, VŠE 1996.
13. ČERMÁK, V., VODRÁŽKOVÁ, I.: *Chování čtvrtého normovaného momentu u standardních rozdělení.* Ekonomicko–matematický obzor 27, č. 1. Praha, ČSAV 1991.
14. ČERMÁK, V., VRABEC, M.: *Sampling Distributions of Some Measures of Skewness: Results of a MC-Simulation.* Acta Oeconomica Pragensia 6, č. 2, VŠE Praha 1998.
15. ČERMÁK, V., VRABEC, M.: *Teorie výběrových šetření, část 1.* Praha, VŠE 1999.
16. CYHELSKÝ, L.: *Úvod do teorie popisné statistiky.* Praha, SNTL/Alfa 1974.
17. DEVILLARD, N: *Fast median search: an ANSI C implementation.* http://ndevilla.free.fr/median/median.pdf.1998.
18. D'AGOSTINO, R. B., PEARSON, E.S.: *Test for departure from normality. Empirical result for the distributions of* b_2 *and* $\sqrt{b_1}$. Biometrika 60, 613-622, 1973.
19. D'AGOSTINO, R. B., TIETJEN, G. L.: *Simulation probability points of* b_2 *for small samples.* Biometrika 58, 669-672, 1971.
20. D'AGOSTINO, R. B., TIETJEN, G. L.: *Approaches to the null distribution of* $\sqrt{b_1}$. Biometrika 60, 169-173, 1973.

21. DEVROYE, L.*: Non-Uniform Random Variate Generation.* New York, Springster-Verlag 1985.
22. ELAMIR, E. A. H., SEHEULT, A. H.*: Trimmed L-moments.* Computational Statistics & Data Analysis 43, 299-314. 2003.
23. ELAMIR, E. A. H., SEHEULT, A. H.*: Exact variance structure of sample L-moments.* Journal of Statistical Planning and Inference 124 (2), 337-359. 2004.
24. EVANS, W. D.*: On the Variance of Estimates of the Standard Deviation and Variance.* Journal of the American Statistical Association 46, č. 254, 220-224. 1951.
25. GODWIN, H. J.: *On the distribution of the estimate of mean deviation obtained from samples from a normal population.* Biometrika 33, 254, 1945.
26. GUTTMAN, N., B.: *On the sensitivity of sample L-moments to sample size*. Journal of Climate 6, 1026-1029. 1994.
27. HEBÁK, P. KAHOUNOVÁ, J.: *Počet pravděpodobnosti v příkladech.* Praha, SNTL 1988.
28. HARTLEY, H. O.: *Probability integral and percentage points of the mean deviation in normal samples.* Biometrika 33, 257, 1945.
29. HENDRICKS, W. A., ROBEY, K. W: *The Sampling Distribution of the Coefficient of Variation.* The Annals of Mathematical Statistics 7, č 3, 129-132. 1936.
30. HERRERO, J.: Microsoft Visual Basic implementation of algorithm for computing the inverse normal cumulative distribution function. http://home.online.no/~pjacklam/notes/invnorm/
31. HOSKING, J. R. M.: *L-moments: analysis and estimation of distributions using linear combinations of order statistics*. Journal of the Royal Statistical Society, Series B, 52, 105-124. 1990.
32. HOSKING, J. R. M., WALLIS, J. R.: *A comparison of unbiased and plotting-position estimators of L-moments*. Water Resources Research 31, 2019-2025. 1995.
33. HURT, J.: *Simulační metody*. Praha, SPN 1982.
34. IGLEWITZ, B., MYERS, R. H., HOWE, R. B.: *On the percentage points of the sample coefficient of variation.* Biometrika 55, č. 3, 580-581. 1968.
35. JOHNSON, N. L.: *The moment problem for unimodal distributions*. The Annals of Mathematical Statistics 22, 433-439.1951.
36. JOHNSON, N. L., KOTZ, S., BALAKRISHNAN, N.: *Continuous univariate distributions. Vol. 1*. New York: John Wiley & Sons. 1994.
37. KENDAL, M. G., STUART, A.: *Advanced Theory of Statistics. Vol. 1, Distribution Theory*. London, Charles Griffin & Company Limited 1963.
38. KONDO, T.: *A Theory of the Sampling Distribution of Standard Deviations*. Biometrika 2, 36-64. 1930.
39. KOTZ, S., JOHNSON, N. L.: *Breakthroughs in statistics. Vol. 2, Methodology and distribution*. New York; Berlin; Heidelberg: Springer. 1992.
40. LAUBER, J., HUŠEK, R.*: Simulační modely.* Praha, SNTL/ALFA 1987.
41. LOCKE, C., SPURRIER, J.: *Using U-statistics for testing Normality against nonsymmetric alternatives*. Biometrika 63, 143-147. 1976.
42. McKAY, A. T.: *The Distribution of the Estimated Coefficient of Variation*. Journal of the Royal Statistical Society 94, č 4, 564-567. 1931.

43. MOORS, J. J. A., WAGEMAKERS, R. TH. A., COENEN, V. M. J., HEUTS, R. M. J., JANSSENS M. J. T. B.: *Characterizing systems of distributions by quantile measures*. Statistica Neerlandica, 50, 417–430. 1996.

44. MUDHOLKAR, G. S., HUSTON, A. D.: *LQ-moments: Analogs of L-moments*. Journal of Statistical Planning and Inference 71, 191-208. 1998.

45. NAIR, U.: *The standard error of Gini's mean difference*. Biometrika 28, 428-436. 1936.

46. NORRIS, N.: *Some Efficient Measures of Relative Dispersion*. The Annals of Mathematical Statistics 9, č. 3, 214-220. 1938.

47. PEARSON, E. S.: *A Further Development of Test for Normality*, Biometrika 22, 239-249, 1930

48. PEARSON, E. S., D'AGOSTINO, R. B., BOWMAN, K.O.: *Test for departure from normality: comparison of power*. Biometrika 64, 231-246. 1977. (*)

49. PICKARD, H. C.: *A note on the Maximum Value of kurtosis.* The Annals of Mathematical Statistics 22, č 4, 480-482. 1951.

50. ROUSSEEUW, P. J., CROUX, C.: *Alternatives to the Median Absolute Deviation.* Journal of the American Statistical Association 88, č. 424, 1273-1283. 1993.(*)

51. ROYSTON, P.: *Which measures of skewness and kurtosis are the best?* Statistics in medicine 11, 333-343. 1992.

52. SANKARASUBRAMANIAN, A., SRINIVASAN, K.: *Investigation and comparison of sampling properties of L-moments and conventional moments*. Journal of Hydrology 218, 13-34. 1999.

53. SILLITTO, G. P.: *Interrelations between certain linear systematic statistics of samples from any continuous population,* Biometrika 38, 377-382. 1951.

54. VOGEL, R. M., FENNESSEY, N. M.: *L-moment diagram should replace product moment diagram.* Water Resource Research 29, 1745-1752. 1993.

55. WAGEMAKERS, R. TH. A., MOORS, J. J. A., JANSSENS M. J. T. B.: *Characterizing distributions by quantile measures*. Tilburg University 1992.

56. WILKINS, J. E.: *A note on skewness and kurtosis.* The Annals of Mathematical Statistics 15, č 2, 333-335. 1944.

57. ZENGA, M., POLISICCHIO, M., GRESELIN, F.: *The variance of Gini's mean difference and its estimators*. Statistica 64, 455-475. Università di Bologna 2004.

Přílohy

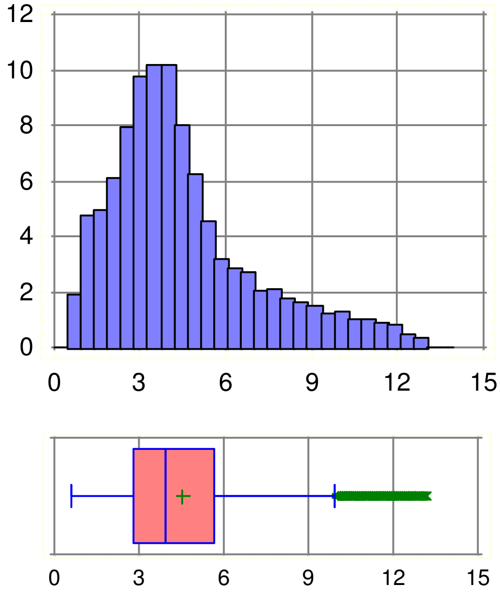
1. Základní charakteristiky proměnných z reálných datových souborů použitých pro realizaci experimentu
2. Základní charakteristiky výběrových rozdělení estimátorů měr variability pro výběry z normovaného normálního rozdělení
3. Základní charakteristiky výběrových rozdělení estimátorů měr šikmosti pro výběry z normálního rozdělení
4. Význačné kvantily estimátorů měr šikmosti pro výběry z normální rozdělení
5. Základní charakteristiky výběrových rozdělení estimátorů měr kurtozy pro výběry z normálního rozdělení
6. Význačné kvantily estimátorů měr kurtozy pro výběry z normální rozdělení
7. Základní popisné charakteristiky proměnných z teoretických pravděpodobnostních rozdělení

Příloha 1: Základní charakteristika proměnných z reálných datových souborů použitých pro realizaci experimentu

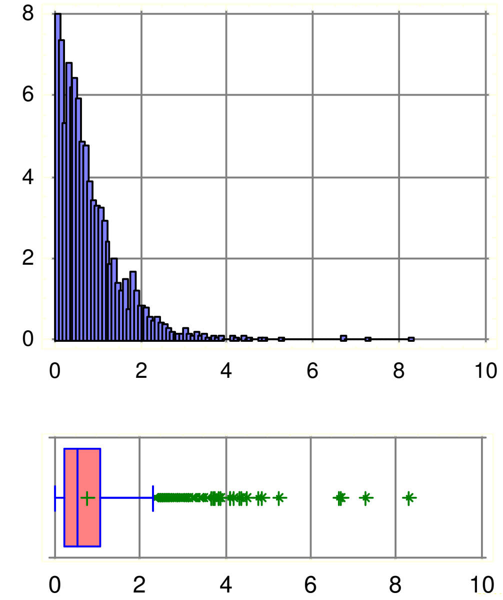
Tabulka P.1: Základní popisné charakteristiky proměnných z konečných datových souborů použitých v experimentu:

Soubor	N	$\overline{X}$	$\tilde{X}_{50}$	S_X	V_X	α_M	β_M
Katastry	13 022	4,54	3,91	2,60	0,573	1,07	3,72
Půda	2 258	758,30	555,00	771,43	1,017	2,43	14,67
Doba	50 000	9,61	7,67	8,01	0,833	1,77	6,04
Délka hovorů	50 000	124,39	77,00	137,25	1,103	3,17	21,69
SMS	50 000	100,88	48,00	182,35	1,808	39,46	4526,55
Cena ubytování	15 857	225,30	163,00	307,58	1,365	9,50	146,07
Zaměstnanci	15 848	5,31	2,00	17,34	3,269	16,18	380,38
Kapacita	15 857	31,64	13,00	78,62	2,485	19,08	624,83
Obvody	2 331	76,73	81,00	49,06	0,639	0,61	5,80
Počet hovorů	50 000	80,19	53,00	83,89	1,046	3,24	23,03
DJIA – změna	16 383	$3{,}3386 \cdot 10^{-4}$	$4{,}3101 \cdot 10^{-4}$	0,008864	-	-1,08	31,47
DJIA – rozpětí	16 383	0,0167903	0,0154406	0,007827	0,466	5,54	128,79
GE – změna	11 281	0,0001965	0	0,019360	-	-10,9	313,29
GE – rozpětí	11 281	0,018881	0,0162162	0,011232	0,595	3,25	32,90
USD	13 260	$2{,}456 \cdot 10^{-5}$	0	0,004714	-	0,26	8,11

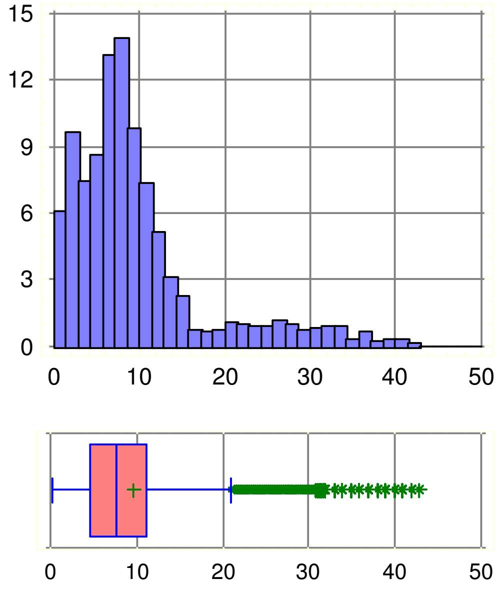
Označení souboru: **Katastry**
Statistická jednotka:
Katastrální území
Sledovaný znak:
Cena za 1 m² zemědělské půdy (2001)



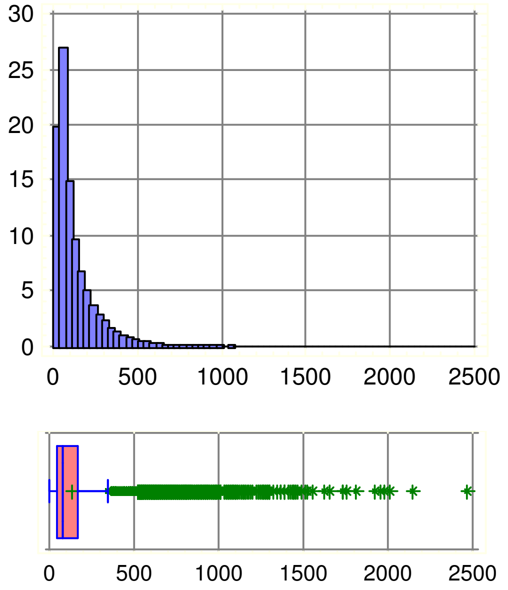
Označení souboru: **Půda**
Statistická jednotka:
Zemědělský podnik
Sledovaný znak: *Rozloha obdělávané zemědělské půdy (v tis. ha)*



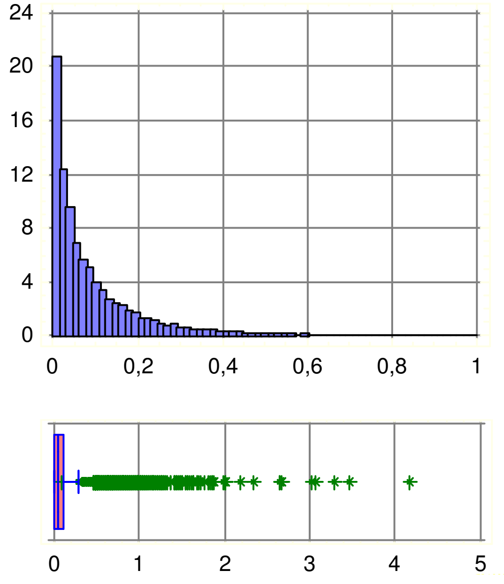
Označení souboru: **Doba**
Statistická jednotka:
Zákazník mobilního operátora
Sledovaný znak:
Doba trvání smlouvy (v měsících)



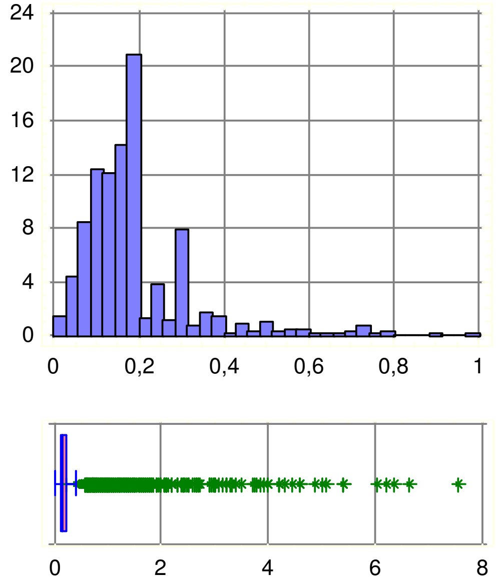
Označení souboru: **Délka hovorů**
Statistická jednotka:
Zákazník mobilního operátora
Sledovaný znak:
Průměrná délka hovorů za měsíc (v min)



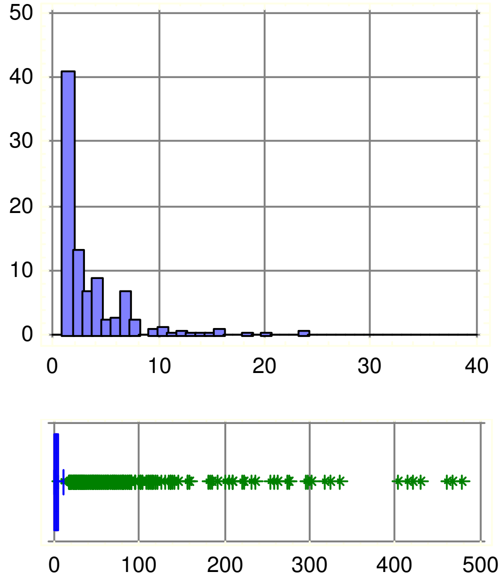
Označení souboru: **SMS**
Statistická jednotka:
Zákazník mobilního operátora
Sledovaný znak: *Průměrný počet odeslaných SMS za měsíc (v tis.)*



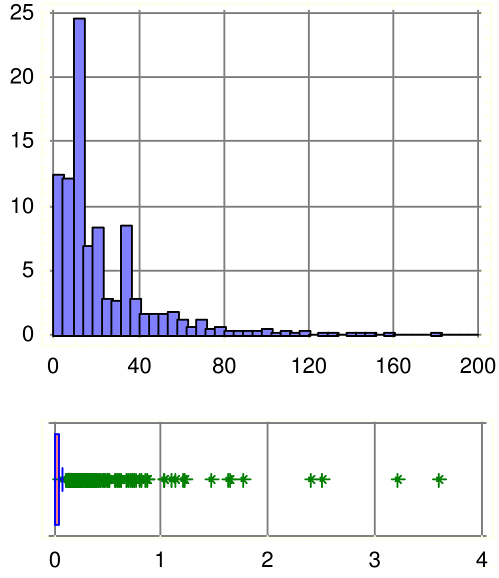
Označení souboru: **Cena ubytování**
Statistická jednotka:
Ubytovací zařízení v ČR (1997)
Sledovaný znak:
Cena za lůžko (v tis. Kč/den)



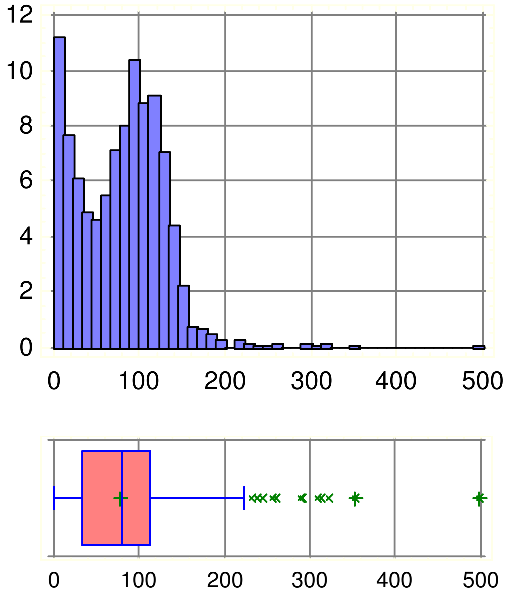
Označení souboru: **Zaměstnanci**
Statistická jednotka:
Ubytovací zařízení v ČR (1997)
Sledovaný znak:
Počet zaměstnanců



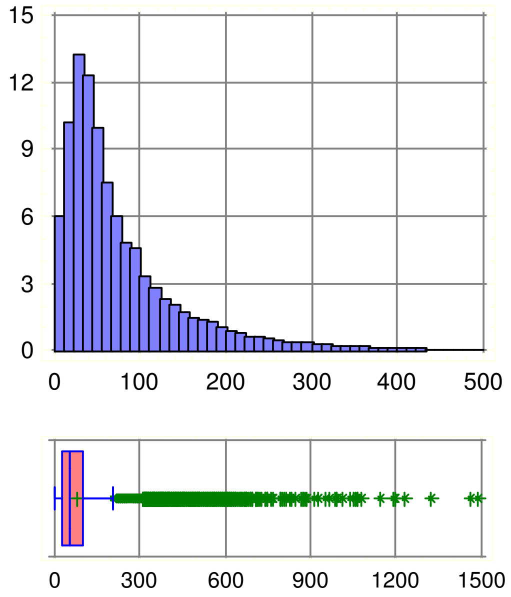
Označení souboru: **Kapacita**
Statistická jednotka:
Ubytovací zařízení v ČR (1997)
Sledovaný znak: *Počet lůžek*
Pozn.:spodní graf v tis.



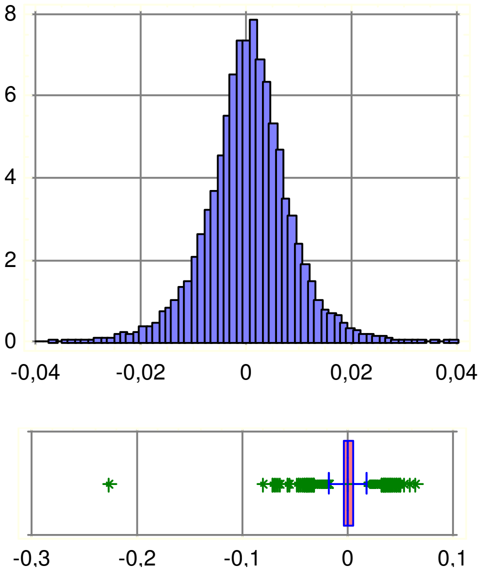
Označení souboru: **Obvody**
Statistická jednotka:
Sčítací obvod
Sledovaný znak:
Počet bytových jednotek v obvodu



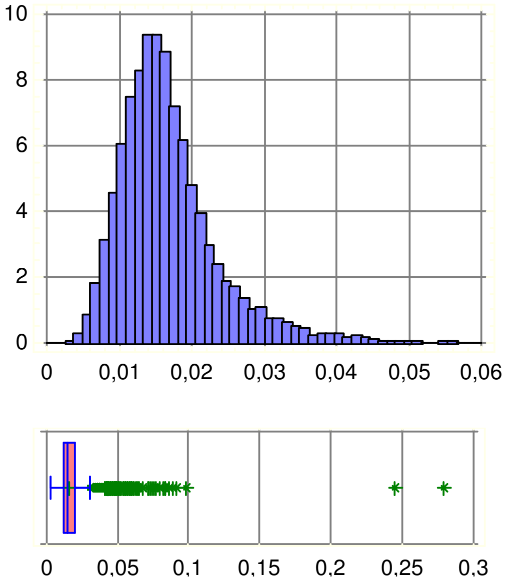
Označení souboru: **Počet hovorů**
Statistická jednotka:
Zákazník mobilního operátora
Sledovaný znak:
Průměrný počet hovorů za měsíc



Označení souboru: **DJIA – změna**
Sledovaný znak: *Relativní změna*[4] *indexu DJIA (Dow Jones Industrial Average Index) během jednoho dne od r. 1942 do r. 2007*

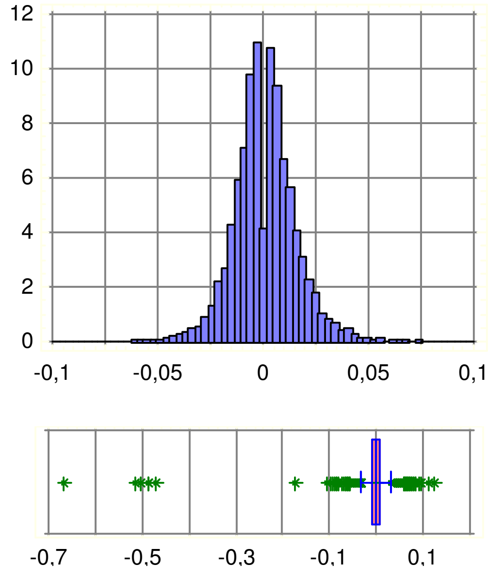


Označení souboru: **DJIA – rozpětí**
Sledovaný znak:
Relativní rozpětí[5] *hodnot indexu DJIA během jednoho dne od r. 1942 do r. 2007*

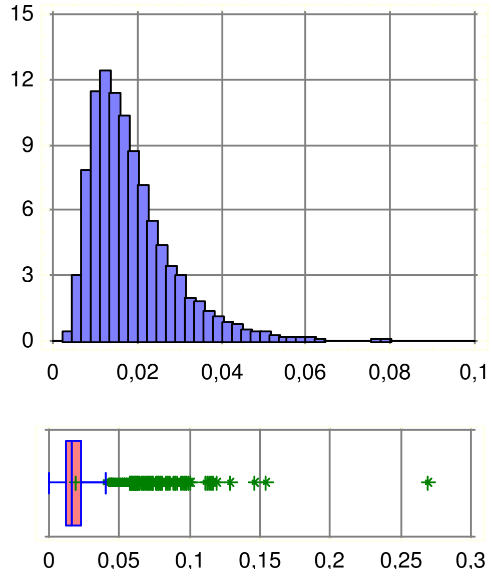


[4] (uzavírací hodnota v aktuální den - uzavírací hodnota předchozí den) / uzavírací hodnota předchozí den
[5] (maximální hodnota - minimální hodnota) / otevírací hodnota

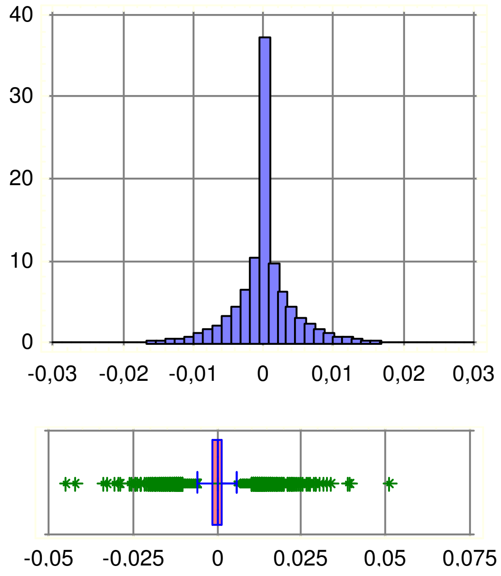
Označení souboru: **GE – změna**
Sledovaný znak: *Relativní změna kurzu akcie GE (General Electric) během jednoho dne od r. 1962 do r. 2006*



Označení souboru: **GE – rozpětí**
Sledovaný znak: *Relativní rozpětí kurzu akcie GE během jednoho dne od r. 1962 do r. 2006*



Označení souboru: **USD**
Sledovaný znak: *Relativní změna kurzu USD/GBP během jednoho dne od r. 1971 do r. 2006*



Příloha 2: Základní charakteristiky výběrových rozdělení estimátorů měr variability pro výběry z normovaného normálního rozdělení

	n	E(t)	B(t)	S(t)	SE(t)	SE(t)√n	$\alpha(t)$	$\beta(t)$
s_X	**7**	0,9601	-0,0399	0,2796	0,2824	0,7473	0,33	3,06
	15	0,9835	-0,0165	0,1865	0,1872	0,7250	0,16	2,95
	23	0,9883	-0,0117	0,1480	0,1485	0,7122	0,17	3,04
	31	0,9917	-0,0083	0,1276	0,1279	0,7122	0,14	2,98
	47	0,9951	-0,0049	0,1041	0,1043	0,7148	0,12	3,01
	63	0,9958	-0,0042	0,0895	0,0896	0,7109	0,07	3,03
	100	0,9980	-0,0020	0,0706	0,0707	0,7067	0,07	2,98
	200	0,9993	-0,0007	0,0503	0,0503	0,7116	0,05	2,98
	350	0,9992	-0,0008	0,0376	0,0376	0,7027	0,02	3,02
	500	0,9995	-0,0005	0,0314	0,0314	0,7025	0,00	3,01
	1 000	0,9997	-0,0003	0,0224	0,0224	0,7072	-0,01	3,04
	∞	**1,0000**	**0,0000**	**0,0000**			**0,00**	**3,00**
s_Q	**7**	0,6581	-0,0164	0,2582	0,2587	0,6845	0,50	3,17
	15	0,6666	-0,0079	0,1896	0,1898	0,7350	0,39	3,11
	23	0,6661	-0,0084	0,1539	0,1541	0,7392	0,30	3,08
	31	0,6688	-0,0057	0,1352	0,1353	0,7535	0,26	3,02
	47	0,6713	-0,0032	0,1115	0,1115	0,7644	0,22	2,98
	63	0,6717	-0,0028	0,0974	0,0974	0,7732	0,18	3,02
	100	0,6729	-0,0016	0,0774	0,0774	0,7744	0,14	3,00
	200	0,6745	0,0000	0,0554	0,0554	0,7834	0,12	2,94
	350	0,6736	-0,0009	0,0418	0,0418	0,7824	0,09	3,06
	500	0,6742	-0,0003	0,0352	0,0352	0,7867	0,05	3,01
	1 000	0,6744	-0,0001	0,0248	0,0248	0,7837	0,04	3,01
	∞	**0,6745**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**
s_{QE}	**7**	0,4050	-0,0101	0,1589	0,1592	0,4212	0,50	3,17
	15	0,4102	-0,0049	0,1167	0,1168	0,4523	0,39	3,11
	23	0,4099	-0,0051	0,0947	0,0949	0,4549	0,30	3,08
	31	0,4116	-0,0035	0,0832	0,0833	0,4637	0,26	3,02
	47	0,4131	-0,0019	0,0686	0,0686	0,4704	0,22	2,98
	63	0,4134	-0,0017	0,0599	0,0599	0,4758	0,18	3,02
	100	0,4141	-0,0010	0,0476	0,0477	0,4765	0,14	3,00
	200	0,4151	0,0000	0,0341	0,0341	0,4821	0,12	2,94
	350	0,4145	-0,0005	0,0257	0,0257	0,4815	0,09	3,06
	500	0,4149	-0,0002	0,0217	0,0217	0,4841	0,05	3,01
	1 000	0,4150	-0,0001	0,0153	0,0153	0,4823	0,04	3,01
	∞	**0,4151**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**
s_L	**7**	0,5647	0,0005	0,1662	0,1662	0,4397	0,36	3,11
	15	0,5650	0,0008	0,1082	0,1082	0,4191	0,17	2,94
	23	0,5638	-0,0004	0,0854	0,0854	0,4094	0,18	3,03
	31	0,5642	0,0000	0,0734	0,0734	0,4087	0,14	2,98
	47	0,5643	0,0001	0,0597	0,0597	0,4091	0,13	3,00
	63	0,5640	-0,0002	0,0512	0,0512	0,4068	0,07	3,02
	100	0,5644	0,0002	0,0404	0,0404	0,4040	0,08	2,99
	200	0,5645	0,0004	0,0287	0,0287	0,4060	0,05	2,97
	350	0,5641	-0,0001	0,0214	0,0214	0,4011	0,03	3,03
	500	0,5642	0,0000	0,0179	0,0179	0,4013	0,00	2,99
	1 000	0,5642	0,0000	0,0128	0,0128	0,4041	-0,01	3,05
	∞	**0,5642**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**

	n	E(t)	B(t)	S(t)	SE(t)	SE(t)√n	$\alpha(t)$	$\beta(t)$
s_{TL}	**7**	0,2978	0,0008	0,1199	0,1199	0,3173	0,58	3,32
	15	0,2976	0,0005	0,0698	0,0698	0,2703	0,32	3,05
	23	0,2966	-0,0005	0,0539	0,0539	0,2583	0,25	3,02
	31	0,2969	-0,0001	0,0459	0,0459	0,2553	0,21	3,01
	47	0,2969	-0,0002	0,0369	0,0369	0,2528	0,18	2,99
	63	0,2969	-0,0002	0,0318	0,0318	0,2525	0,13	3,03
	100	0,2970	-0,0001	0,0249	0,0249	0,2490	0,12	3,02
	200	0,2973	0,0002	0,0176	0,0176	0,2483	0,09	2,95
	350	0,2969	-0,0002	0,0131	0,0131	0,2460	0,04	3,04
	500	0,2970	0,0000	0,0111	0,0111	0,2472	0,02	2,97
	1 000	0,2970	0,0000	0,0078	0,0078	0,2480	0,00	3,03
	∞	**0,2970**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**
s_{LQ}	**7**	0,7179	0,1685	0,2234	0,2798	0,7404	0,44	3,19
	15	0,5708	0,0213	0,1340	0,1357	0,5256	0,33	3,07
	23	0,5366	-0,0129	0,1107	0,1115	0,5345	0,29	3,03
	31	0,5593	0,0098	0,0978	0,0983	0,5473	0,25	2,99
	47	0,5493	-0,0002	0,0782	0,0782	0,5359	0,21	2,99
	63	0,5447	-0,0047	0,0671	0,0673	0,5341	0,17	3,02
	100	0,5467	-0,0028	0,0534	0,0534	0,5342	0,15	3,03
	200	0,5530	0,0036	0,0379	0,0381	0,5383	0,12	2,98
	350	0,5509	0,0015	0,0285	0,0285	0,5338	0,06	3,01
	500	0,5508	0,0014	0,0240	0,0241	0,5381	0,05	2,95
	1 000	0,5508	0,0014	0,0170	0,0171	0,5403	0,05	3,00
	∞	**0,5494**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**
s_{HL}	**7**	1,0188	0,0648	0,3411	0,3472	0,9186	0,48	3,32
	15	0,9843	0,0304	0,2087	0,2109	0,8167	0,24	3,06
	23	0,9707	0,0168	0,1612	0,1620	0,7771	0,19	3,02
	31	0,9672	0,0132	0,1368	0,1374	0,7650	0,14	2,99
	47	0,9622	0,0082	0,1097	0,1100	0,7542	0,13	2,98
	63	0,9597	0,0058	0,0939	0,0941	0,7467	0,07	3,01
	100	0,9580	0,0041	0,0734	0,0735	0,7351	0,07	3,01
	200	0,9565	0,0026	0,0518	0,0519	0,7335	0,06	2,95
	350	0,9547	0,0007	0,0387	0,0387	0,7245	0,02	3,04
	500	0,9546	0,0007	0,0324	0,0324	0,7250	0,00	2,97
	1 000	0,9543	0,0004	0,0231	0,0231	0,7291	-0,01	3,06
	∞	**0,9539**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**

Příloha 3: Základní charakteristiky výběrových rozdělení estimátorů měr šikmosti pro výběry z normálního rozdělení

	n	E(t)	B(t)	S(t)	SE(t)	SE(t)√n	$\alpha(t)$	$\beta(t)$
a_M	**7**	0,0073	0,0073	0,6119	0,6120	1,6191	-0,02	3,02
	15	-0,0052	-0,0052	0,5174	0,5175	2,0041	0,01	3,43
	23	-0,0007	-0,0007	0,4517	0,4517	2,1664	0,01	3,35
	31	-0,0038	-0,0038	0,4017	0,4018	2,2369	-0,01	3,64
	47	-0,0002	-0,0002	0,3389	0,3389	2,3234	0,00	3,58
	63	-0,0001	-0,0001	0,2949	0,2949	2,3404	0,00	3,42
	100	0,0021	0,0021	0,2374	0,2374	2,3744	-0,04	3,24
	200	-0,0024	-0,0024	0,1711	0,1711	2,4196	0,01	3,10
	350	0,0009	0,0009	0,1303	0,1303	2,4382	0,03	3,07
	500	0,0002	0,0002	0,1096	0,1096	2,4504	0,00	3,07
	1 000	0,0003	0,0003	0,0771	0,0771	2,4388	0,03	3,10
	∞	**0,0000**	**0,0000**	**0,0000**		**2,4495**	**0,00**	**3,00**
a_Q	**7**	0,0067	0,0067	0,4337	0,4337	1,1475	-0,01	2,21
	15	-0,0033	-0,0033	0,3214	0,3214	1,2447	0,01	2,52
	23	-0,0030	-0,0030	0,2656	0,2657	1,2741	0,01	2,63
	31	0,0027	0,0027	0,2329	0,2329	1,2966	0,02	2,72
	47	0,0012	0,0012	0,1917	0,1917	1,3140	0,01	2,88
	63	0,0026	0,0026	0,1673	0,1673	1,3281	0,00	2,87
	100	0,0006	0,0006	0,1321	0,1321	1,3210	0,01	2,89
	200	-0,0002	-0,0002	0,0941	0,0941	1,3308	-0,01	2,98
	350	0,0002	0,0002	0,0725	0,0725	1,3556	-0,03	2,90
	500	-0,0003	-0,0003	0,0603	0,0603	1,3486	-0,02	2,96
	1 000	0,0002	0,0002	0,0424	0,0424	1,3412	0,01	3,08
	∞	**0,0000**	**0,0000**	**0,0000**	**0,0000**	**1,3560**	**0,00**	**3,00**
a_L	**7**	0,0029	0,0029	0,2208	0,2208	0,5843	-0,01	2,88
	15	-0,0016	-0,0016	0,1280	0,1280	0,4958	0,01	2,96
	23	0,0000	0,0000	0,0997	0,0997	0,4779	0,00	2,93
	31	-0,0003	-0,0003	0,0832	0,0832	0,4633	0,02	2,99
	47	0,0000	0,0000	0,0662	0,0662	0,4539	-0,01	3,00
	63	0,0002	0,0002	0,0561	0,0561	0,4452	-0,01	3,01
	100	0,0004	0,0004	0,0438	0,0438	0,4382	-0,02	3,00
	200	-0,0004	-0,0004	0,0309	0,0309	0,4375	0,01	2,98
	350	0,0002	0,0002	0,0234	0,0234	0,4369	0,04	3,00
	500	0,0001	0,0001	0,0197	0,0197	0,4395	0,01	2,97
	1 000	0,0001	0,0001	0,0138	0,0138	0,4366	0,03	3,07
	∞	**0,0000**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**
a_M^+	**7**	0,0051	0,0051	0,3875	0,3875	1,0254	-0,02	2,23
	15	-0,0036	-0,0036	0,3123	0,3123	1,2095	0,02	2,46
	23	-0,0003	-0,0003	0,2719	0,2719	1,3041	0,02	2,53
	31	-0,0020	-0,0020	0,2414	0,2414	1,3442	0,01	2,70
	47	-0,0003	-0,0003	0,2048	0,2048	1,4037	0,00	2,78
	63	0,0001	0,0001	0,1793	0,1793	1,4231	0,00	2,81
	100	0,0013	0,0013	0,1457	0,1457	1,4572	-0,03	2,86
	200	-0,0015	-0,0015	0,1060	0,1060	1,4996	0,00	2,91
	350	0,0005	0,0005	0,0811	0,0811	1,5172	0,03	2,96
	500	0,0001	0,0001	0,0683	0,0683	1,5280	0,00	2,98
	1 000	0,0001	0,0001	0,0482	0,0482	1,5249	0,03	3,06
	∞	**0,0000**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**

	n	E(t)	B(t)	S(t)	SE(t)	SE(t)√n	$\alpha(t)$	$\beta(t)$
a_{TL}	**7**	0,0057	0,0057	0,3112	0,3113	0,8236	0,01	2,84
	15	-0,0020	-0,0020	0,1314	0,1314	0,5090	0,00	3,01
	23	-0,0003	-0,0003	0,0959	0,0959	0,4601	-0,02	2,99
	31	0,0004	0,0004	0,0781	0,0781	0,4351	0,02	3,00
	47	0,0001	0,0001	0,0612	0,0612	0,4197	-0,01	2,95
	63	0,0003	0,0003	0,0516	0,0516	0,4095	-0,01	2,97
	100	0,0004	0,0004	0,0397	0,0398	0,3975	0,00	3,01
	200	-0,0001	-0,0001	0,0278	0,0278	0,3928	0,00	2,98
	350	0,0002	0,0002	0,0211	0,0211	0,3939	0,00	2,98
	500	0,0002	0,0002	0,0177	0,0177	0,3955	0,00	2,91
	1 000	0,0001	0,0001	0,0124	0,0124	0,3909	0,04	3,01
	∞	**0,0000**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**
a_{LQ}	**7**	0,0036	0,0036	0,2105	0,2106	0,5571	-0,01	2,50
	15	-0,0034	-0,0034	0,2255	0,2255	0,8733	-0,01	3,22
	23	-0,0005	-0,0005	0,1703	0,1703	0,8167	0,00	3,17
	31	0,0013	0,0013	0,1399	0,1399	0,7789	0,03	3,08
	47	0,0005	0,0005	0,1168	0,1168	0,8007	-0,02	3,04
	63	0,0000	0,0000	0,1033	0,1033	0,8196	-0,01	3,04
	100	0,0081	0,0081	0,0790	0,0794	0,7937	0,02	2,99
	200	-0,0040	-0,0040	0,0552	0,0554	0,7831	0,02	3,04
	350	-0,0019	-0,0019	0,0430	0,0430	0,8053	-0,03	2,96
	500	-0,0013	-0,0013	0,0355	0,0356	0,7954	-0,03	2,92
	1 000	0,0009	0,0009	0,0250	0,0250	0,7919	0,06	2,99
	∞	**0,0000**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**
a_{MC}	**7**	0,0045	0,0045	0,3593	0,3594	0,9508	0,00	2,80
	15	-0,0023	-0,0023	0,2721	0,2721	1,0538	0,02	2,81
	23	-0,0022	-0,0022	0,2258	0,2258	1,0828	0,02	2,83
	31	0,0020	0,0020	0,1964	0,1964	1,0933	0,03	2,90
	47	0,0005	0,0005	0,1610	0,1610	1,1036	0,01	2,97
	63	0,0020	0,0020	0,1396	0,1396	1,1079	0,01	2,92
	100	0,0006	0,0006	0,1088	0,1088	1,0884	0,03	2,93
	200	-0,0003	-0,0003	0,0779	0,0779	1,1014	0,01	3,00
	350	0,0003	0,0003	0,0596	0,0596	1,1146	-0,03	2,95
	500	-0,0003	-0,0003	0,0497	0,0497	1,1121	-0,02	2,93
	1 000	0,0002	0,0002	0,0350	0,0350	1,1080	0,02	3,05
	∞	**0,0000**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**
a_{C}	**7**	0,0031	0,0031	0,2253	0,2253	0,5961	-0,03	2,90
	15	-0,0015	-0,0015	0,1550	0,1550	0,6005	0,03	2,93
	23	-0,0011	-0,0011	0,1255	0,1255	0,6018	0,02	2,93
	31	0,0004	0,0004	0,1084	0,1084	0,6037	0,01	2,99
	47	0,0003	0,0003	0,0883	0,0883	0,6056	0,02	2,92
	63	0,0007	0,0007	0,0760	0,0760	0,6031	-0,02	3,02
	100	0,0003	0,0003	0,0599	0,0599	0,5990	-0,01	3,05
	200	-0,0005	-0,0005	0,0425	0,0425	0,6009	0,00	2,97
	350	0,0002	0,0002	0,0324	0,0324	0,6069	-0,05	2,98
	500	0,0000	0,0000	0,0273	0,0273	0,6098	0,02	2,94
	1 000	0,0001	0,0001	0,0190	0,0190	0,6008	0,03	3,05
	∞	**0,0000**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**

Příloha 4: Význačné kvantily výběrových rozdělení estimátorů měr šikmosti pro výběry z normálního rozdělení

	n	1%	2,5%	5%	10%	90%	95%	97,5%	99%
a_M	7	-1,4587	-1,2426	-1,0143	-0,7724	0,7858	1,0199	1,2320	1,4491
	15	-1,2649	-1,0347	-0,8515	-0,6525	0,6437	0,8509	1,0304	1,2833
	23	-1,0870	-0,8826	-0,7390	-0,5657	0,5713	0,7391	0,9018	1,1116
	31	-0,9983	-0,8036	-0,6476	-0,4972	0,4981	0,6491	0,8055	0,9736
	47	-0,8037	-0,6730	-0,5553	-0,4251	0,4214	0,5496	0,6666	0,8123
	63	-0,7091	-0,5748	-0,4779	-0,3710	0,3683	0,4795	0,5855	0,7048
	100	-0,5556	-0,4699	-0,3896	-0,3012	0,3044	0,3866	0,4658	0,5598
	200	-0,4008	-0,3423	-0,2831	-0,2217	0,2132	0,2788	0,3329	0,4053
	350	-0,3024	-0,2543	-0,2111	-0,1642	0,1678	0,2176	0,2616	0,3130
	500	-0,2580	-0,2151	-0,1795	-0,1392	0,1395	0,1819	0,2164	0,2554
	1 000	-0,1765	-0,1502	-0,1258	-0,0977	0,0998	0,1285	0,1507	0,1825
a_Q	7	-0,8679	-0,7887	-0,7022	-0,5850	0,5983	0,7167	0,7968	0,8713
	15	-0,6953	-0,6124	-0,5332	-0,4282	0,4226	0,5265	0,6108	0,7034
	23	-0,5895	-0,5101	-0,4382	-0,3519	0,3445	0,4329	0,5116	0,5891
	31	-0,5127	-0,4472	-0,3809	-0,3010	0,3059	0,3909	0,4562	0,5289
	47	-0,4391	-0,3756	-0,3143	-0,2454	0,2482	0,3188	0,3828	0,4453
	63	-0,3834	-0,3234	-0,2737	-0,2102	0,2187	0,2783	0,3263	0,3876
	100	-0,3018	-0,2591	-0,2196	-0,1695	0,1721	0,2175	0,2588	0,3041
	200	-0,2210	-0,1860	-0,1552	-0,1211	0,1203	0,1547	0,1829	0,2188
	350	-0,1684	-0,1432	-0,1199	-0,0939	0,0936	0,1179	0,1393	0,1658
	500	-0,1409	-0,1188	-0,1008	-0,0769	0,0772	0,0991	0,1191	0,1385
	1 000	-0,0992	-0,0827	-0,0687	-0,0536	0,0541	0,0690	0,0821	0,0984
a_L	7	-0,5072	-0,4343	-0,3644	-0,2810	0,2863	0,3650	0,4326	0,5137
	15	-0,2923	-0,2523	-0,2139	-0,1654	0,1629	0,2098	0,2487	0,2996
	23	-0,2279	-0,1947	-0,1635	-0,1282	0,1283	0,1647	0,1945	0,2306
	31	-0,1937	-0,1605	-0,1361	-0,1067	0,1070	0,1380	0,1622	0,1917
	47	-0,1536	-0,1306	-0,1101	-0,0845	0,0847	0,1086	0,1297	0,1534
	63	-0,1308	-0,1097	-0,0918	-0,0716	0,0716	0,0917	0,1095	0,1294
	100	-0,1020	-0,0858	-0,0716	-0,0560	0,0566	0,0715	0,0851	0,1024
	200	-0,0722	-0,0608	-0,0516	-0,0403	0,0393	0,0506	0,0598	0,0727
	350	-0,0540	-0,0456	-0,0381	-0,0297	0,0303	0,0393	0,0469	0,0556
	500	-0,0455	-0,0386	-0,0322	-0,0249	0,0253	0,0328	0,0386	0,0454
	1 000	-0,0318	-0,0267	-0,0224	-0,0174	0,0179	0,0229	0,0274	0,0325
a_M^+	7	-0,7861	-0,7184	-0,6346	-0,5170	0,5281	0,6352	0,7136	0,7843
	15	-0,6668	-0,5897	-0,5182	-0,4175	0,4151	0,5150	0,5916	0,6806
	23	-0,5890	-0,5146	-0,4480	-0,3582	0,3606	0,4465	0,5218	0,5968
	31	-0,5462	-0,4658	-0,3962	-0,3162	0,3155	0,3989	0,4675	0,5378
	47	-0,4652	-0,4008	-0,3384	-0,2668	0,2643	0,3354	0,3935	0,4634
	63	-0,4079	-0,3443	-0,2931	-0,2315	0,2312	0,2960	0,3484	0,4049
	100	-0,3315	-0,2848	-0,2408	-0,1886	0,1900	0,2377	0,2820	0,3304
	200	-0,2463	-0,2113	-0,1763	-0,1386	0,1343	0,1722	0,2048	0,2472
	350	-0,1858	-0,1574	-0,1313	-0,1031	0,1051	0,1361	0,1632	0,1922
	500	-0,1597	-0,1335	-0,1123	-0,0873	0,0877	0,1136	0,1353	0,1580
	1 000	-0,1099	-0,0934	-0,0791	-0,0614	0,0625	0,0802	0,0942	0,1137

	n	1%	2,5%	5%	10%	90%	95%	97,5%	99%
a_{TL}	7	-0,7138	-0,6084	-0,5023	-0,3908	0,4064	0,5174	0,6163	0,7302
	15	-0,3085	-0,2613	-0,2188	-0,1687	0,1667	0,2134	0,2553	0,3082
	23	-0,2262	-0,1879	-0,1581	-0,1243	0,1229	0,1565	0,1885	0,2199
	31	-0,1826	-0,1502	-0,1267	-0,0991	0,1025	0,1281	0,1542	0,1797
	47	-0,1411	-0,1205	-0,1004	-0,0788	0,0791	0,1002	0,1194	0,1423
	63	-0,1193	-0,1014	-0,0847	-0,0667	0,0665	0,0844	0,1002	0,1215
	100	-0,0916	-0,0777	-0,0648	-0,0503	0,0510	0,0661	0,0788	0,0944
	200	-0,0645	-0,0549	-0,0462	-0,0359	0,0354	0,0458	0,0540	0,0650
	350	-0,0488	-0,0407	-0,0346	-0,0272	0,0272	0,0349	0,0414	0,0492
	500	-0,0402	-0,0342	-0,0285	-0,0225	0,0230	0,0292	0,0343	0,0409
	1 000	-0,0280	-0,0238	-0,0201	-0,0158	0,0159	0,0205	0,0249	0,0292
a_{LQ}	7	-0,4602	-0,3984	-0,3442	-0,2770	0,2830	0,3516	0,3999	0,4652
	15	-0,5421	-0,4494	-0,3710	-0,2865	0,2830	0,3649	0,4430	0,5339
	23	-0,4101	-0,3371	-0,2775	-0,2163	0,2154	0,2804	0,3386	0,4002
	31	-0,3242	-0,2679	-0,2265	-0,1771	0,1817	0,2317	0,2802	0,3253
	47	-0,2708	-0,2295	-0,1918	-0,1485	0,1488	0,1910	0,2280	0,2748
	63	-0,2401	-0,2015	-0,1716	-0,1323	0,1310	0,1687	0,2011	0,2440
	100	-0,1728	-0,1465	-0,1217	-0,0925	0,1078	0,1387	0,1644	0,1925
	200	-0,1333	-0,1119	-0,0954	-0,0743	0,0669	0,0870	0,1057	0,1258
	350	-0,1021	-0,0871	-0,0737	-0,0575	0,0532	0,0684	0,0820	0,0975
	500	-0,0851	-0,0711	-0,0599	-0,0471	0,0446	0,0571	0,0674	0,0796
	1 000	-0,0559	-0,0472	-0,0397	-0,0311	0,0330	0,0425	0,0513	0,0603
a_{MC}	7	-0,8154	-0,7126	-0,6127	-0,4764	0,4915	0,6297	0,7155	0,8133
	15	-0,6164	-0,5301	-0,4556	-0,3610	0,3574	0,4516	0,5311	0,6246
	23	-0,5133	-0,4378	-0,3799	-0,2958	0,2900	0,3694	0,4398	0,5232
	31	-0,4438	-0,3824	-0,3238	-0,2528	0,2579	0,3313	0,3879	0,4584
	47	-0,3753	-0,3156	-0,2633	-0,2073	0,2064	0,2677	0,3227	0,3790
	63	-0,3182	-0,2696	-0,2280	-0,1781	0,1813	0,2328	0,2774	0,3232
	100	-0,2482	-0,2112	-0,1794	-0,1392	0,1407	0,1812	0,2151	0,2567
	200	-0,1814	-0,1535	-0,1275	-0,1005	0,1000	0,1275	0,1511	0,1808
	350	-0,1391	-0,1169	-0,0981	-0,0768	0,0770	0,0979	0,1154	0,1365
	500	-0,1159	-0,0980	-0,0822	-0,0645	0,0628	0,0805	0,0966	0,1140
	1 000	-0,0797	-0,0675	-0,0568	-0,0445	0,0447	0,0572	0,0681	0,0817
a_{C}	7	-0,4286	-0,4286	-0,4286	-0,1429	0,1429	0,4286	0,4286	0,4286
	15	-0,3333	-0,3333	-0,2000	-0,2000	0,2000	0,2000	0,3333	0,3333
	23	-0,3043	-0,2174	-0,2174	-0,1304	0,1304	0,2174	0,2174	0,3043
	31	-0,2258	-0,2258	-0,1613	-0,1613	0,1613	0,1613	0,2258	0,2258
	47	-0,1915	-0,1489	-0,1489	-0,1064	0,1064	0,1489	0,1915	0,1915
	63	-0,1746	-0,1429	-0,1111	-0,1111	0,1111	0,1111	0,1429	0,1746
	100	-0,1400	-0,1200	-0,1000	-0,0800	0,0800	0,1000	0,1200	0,1400
	200	-0,1000	-0,0800	-0,0700	-0,0500	0,0500	0,0700	0,0800	0,1000
	350	-0,0743	-0,0629	-0,0514	-0,0400	0,0400	0,0514	0,0629	0,0743
	500	-0,0640	-0,0520	-0,0440	-0,0360	0,0360	0,0440	0,0520	0,0640
	1 000	-0,0440	-0,0360	-0,0310	-0,0240	0,0240	0,0320	0,0380	0,0440

Příloha 5: Základní charakteristiky výběrových rozdělení estimátorů měr kurtozy pro výběry z normálního rozdělení

	n	E(t)	B(t)	S(t)	SE(t)	SE(t)√n	$\alpha(t)$	$\beta(t)$
b_M	7	2,2479	-0,7521	0,6630	1,0026	2,6526	0,98	3,76
	15	2,6196	-0,3804	0,7656	0,8549	3,3112	1,57	6,98
	23	2,7566	-0,2434	0,7287	0,7683	3,6845	1,58	7,44
	31	2,8134	-0,1866	0,6949	0,7196	4,0064	1,76	9,20
	47	2,8860	-0,1140	0,6218	0,6322	4,3342	1,75	10,61
	63	2,9129	-0,0871	0,5587	0,5654	4,4880	1,52	7,82
	100	2,9469	-0,0531	0,4550	0,4580	4,5805	1,22	6,11
	200	2,9688	-0,0312	0,3315	0,3329	4,7084	0,94	5,00
	350	2,9880	-0,0120	0,2583	0,2586	4,8383	0,78	4,46
	500	2,9880	-0,0120	0,2164	0,2168	4,8467	0,65	4,01
	1 000	2,9917	-0,0083	0,1513	0,1515	4,7916	0,44	3,42
	∞	**3,0000**	**0,0000**	**0,0000**		**4,8990**	**0,00**	**3,00**
b_F	7	-0,0050	-0,0050	1,5911	1,5911	4,2097	0,98	3,76
	15	-0,0078	-0,0078	1,0994	1,0994	4,2580	1,57	6,98
	23	0,0083	0,0083	0,9161	0,9161	4,3935	1,58	7,44
	31	0,0011	0,0011	0,8216	0,8216	4,5746	1,76	9,20
	47	0,0122	0,0122	0,6934	0,6936	4,7547	1,75	10,61
	63	0,0073	0,0073	0,6057	0,6058	4,8080	1,52	7,82
	100	0,0066	0,0066	0,4786	0,4786	4,7860	1,22	6,11
	200	-0,0014	-0,0014	0,3399	0,3399	4,8071	0,94	5,00
	350	0,0051	0,0051	0,2621	0,2621	4,9038	0,78	4,46
	500	0,0000	0,0000	0,2186	0,2186	4,8880	0,65	4,01
	1 000	-0,0023	-0,0023	0,1521	0,1521	4,8091	0,44	3,42
	∞	**0,0000**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**
b_Q	7	1,4138	0,1807	0,8063	0,8263	2,1862	3,98	62,76
	15	1,3118	0,0787	0,4802	0,4866	1,8845	1,55	8,31
	23	1,2845	0,0514	0,3790	0,3824	1,8341	1,29	9,31
	31	1,2721	0,0391	0,3289	0,3312	1,8441	0,97	5,25
	47	1,2559	0,0228	0,2568	0,2578	1,7674	0,76	4,17
	63	1,2516	0,0185	0,2265	0,2273	1,8039	0,65	3,94
	100	1,2443	0,0112	0,1812	0,1815	1,8151	0,57	3,74
	200	1,2374	0,0044	0,1257	0,1258	1,7794	0,39	3,39
	350	1,2372	0,0041	0,0954	0,0955	1,7861	0,28	3,23
	500	1,2355	0,0024	0,0798	0,0799	1,7859	0,22	3,12
	1 000	1,2339	0,0008	0,0559	0,0559	1,7687	0,15	3,10
	∞	**1,2331**	**0,0000**	**0,0000**	**0,0000**	**1,7760**	**0,00**	**3,00**
b_Q^+	7	0,7236	0,0006	0,1452	0,1452	0,3842	-0,49	2,71
	15	0,7205	-0,0025	0,1114	0,1114	0,4315	-0,44	2,90
	23	0,7201	-0,0029	0,0937	0,0937	0,4495	-0,41	2,99
	31	0,7208	-0,0022	0,0826	0,0827	0,4602	-0,35	2,97
	47	0,7218	-0,0012	0,0666	0,0666	0,4564	-0,29	3,02
	63	0,7217	-0,0013	0,0586	0,0586	0,4649	-0,29	3,04
	100	0,7223	-0,0007	0,0473	0,0473	0,4732	-0,21	3,02
	200	0,7223	-0,0007	0,0334	0,0334	0,4726	-0,16	2,99
	350	0,7231	0,0000	0,0253	0,0253	0,4735	-0,12	3,04
	500	0,7228	-0,0002	0,0215	0,0215	0,4809	-0,10	2,97
	1 000	0,7228	-0,0002	0,0150	0,0150	0,4733	-0,08	3,07
	∞	**0,7230**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**

	n	E(t)	B(t)	S(t)	SE(t)	SE(t)√n	$\alpha(t)$	$\beta(t)$
b_L	**7**	0,1266	0,0040	0,2159	0,2159	0,5713	0,31	2,79
	15	0,1244	0,0018	0,1013	0,1013	0,3923	0,26	2,96
	23	0,1247	0,0021	0,0739	0,0739	0,3545	0,18	2,96
	31	0,1237	0,0011	0,0610	0,0610	0,3396	0,18	3,00
	47	0,1238	0,0012	0,0472	0,0472	0,3234	0,10	2,92
	63	0,1233	0,0007	0,0401	0,0401	0,3183	0,14	3,03
	100	0,1233	0,0007	0,0310	0,0310	0,3104	0,10	3,00
	200	0,1226	0,0000	0,0214	0,0214	0,3020	0,06	3,01
	350	0,1229	0,0003	0,0160	0,0160	0,3002	0,04	2,97
	500	0,1226	0,0000	0,0134	0,0134	0,3004	0,03	3,01
	1 000	0,1226	0,0000	0,0094	0,0094	0,2968	0,03	2,97
	∞	**0,1226**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**
b_{TL}	**7**	0,0994	0,0369	0,4138	0,4154	1,0991	0,35	2,36
	15	0,0719	0,0095	0,1068	0,1072	0,4151	0,40	3,13
	23	0,0678	0,0054	0,0699	0,0701	0,3362	0,29	3,14
	31	0,0663	0,0038	0,0556	0,0557	0,3103	0,28	3,09
	47	0,0649	0,0024	0,0406	0,0407	0,2788	0,22	3,04
	63	0,0641	0,0016	0,0343	0,0344	0,2729	0,24	3,14
	100	0,0636	0,0011	0,0262	0,0262	0,2619	0,17	3,05
	200	0,0628	0,0003	0,0178	0,0178	0,2514	0,10	3,07
	350	0,0629	0,0004	0,0133	0,0133	0,2484	0,07	3,02
	500	0,0626	0,0002	0,0111	0,0111	0,2485	0,07	2,99
	1 000	0,0625	0,0000	0,0077	0,0077	0,2449	0,04	3,06
	∞	**0,0625**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**
b_{LQ}	**7**	0,0912	-0,0267	0,2901	0,2914	0,7709	0,25	3,08
	15	0,0629	-0,0550	0,2236	0,2302	0,8916	0,31	3,13
	23	0,0881	-0,0298	0,1790	0,1814	0,8701	0,26	3,32
	31	0,0697	-0,0481	0,1549	0,1622	0,9031	0,21	3,11
	47	0,1451	0,0273	0,1139	0,1172	0,8032	0,09	3,03
	63	0,1308	0,0130	0,1071	0,1079	0,8564	0,11	3,10
	100	0,1454	0,0275	0,0842	0,0886	0,8861	0,12	3,11
	200	0,1271	0,0092	0,0574	0,0581	0,8223	0,05	3,06
	350	0,1120	-0,0059	0,0438	0,0442	0,8274	0,03	3,01
	500	0,1149	-0,0030	0,0366	0,0367	0,8204	0,03	3,05
	1 000	0,1188	0,0009	0,0258	0,0258	0,8155	0,05	3,06
	∞	**0,1179**	**0,0000**	**0,0000**	**0,0000**		**0,00**	**3,00**

Příloha 6: Význačné kvantily výběrových rozdělení estimátorů měr kurtozy pro výběry z normálního rozdělení

	n	1%	2,5%	5%	10%	90%	95%	97,5%	99%
b_M	7	1,2561	1,3262	1,4124	1,5304	3,1918	3,5400	3,8749	4,2354
	15	1,5469	1,6311	1,7255	1,8467	3,6121	4,1044	4,5843	5,2543
	23	1,6945	1,7893	1,8862	2,0156	3,6771	4,1424	4,6069	5,2904
	31	1,8048	1,9093	1,9988	2,1120	3,6711	4,0923	4,5627	5,2101
	47	1,9471	2,0440	2,1295	2,2440	3,6458	4,0181	4,4254	4,9893
	63	2,0423	2,1312	2,2171	2,3207	3,6107	3,9350	4,2640	4,8048
	100	2,1932	2,2685	2,3475	2,4477	3,5241	3,7936	4,0556	4,4083
	200	2,3725	2,4430	2,5069	2,5929	3,4020	3,5740	3,7282	3,9461
	350	2,5017	2,5592	2,6160	2,6885	3,3203	3,4490	3,5740	3,7331
	500	2,5662	2,6236	2,6718	2,7290	3,2688	3,3694	3,4698	3,5960
	1 000	2,6789	2,7247	2,7635	2,8076	3,1872	3,2583	3,3203	3,4061
b_F	7	-2,3853	-2,2172	-2,0103	-1,7271	2,2602	3,0960	3,8997	4,7650
	15	-1,5481	-1,4272	-1,2915	-1,1176	1,4173	2,1242	2,8134	3,7755
	23	-1,3269	-1,2078	-1,0860	-0,9232	1,1655	1,7505	2,3344	3,1936
	31	-1,1914	-1,0679	-0,9620	-0,8282	1,0151	1,5131	2,0692	2,8346
	47	-1,0347	-0,9267	-0,8314	-0,7037	0,8595	1,2747	1,7289	2,3578
	63	-0,9367	-0,8403	-0,7471	-0,6348	0,7637	1,1153	1,4721	2,0583
	100	-0,7862	-0,7069	-0,6239	-0,5185	0,6138	0,8973	1,1728	1,5438
	200	-0,6129	-0,5405	-0,4751	-0,3869	0,4428	0,6192	0,7774	1,0008
	350	-0,4882	-0,4299	-0,3722	-0,2987	0,3422	0,4728	0,5997	0,7610
	500	-0,4261	-0,3681	-0,3194	-0,2616	0,2836	0,3852	0,4866	0,6141
	1 000	-0,3167	-0,2707	-0,2317	-0,1874	0,1942	0,2656	0,3279	0,4142
b_Q	7	0,4221	0,5047	0,5919	0,7109	2,2360	2,6984	3,2680	4,1683
	15	0,5451	0,6324	0,7107	0,8020	1,9071	2,1936	2,4780	2,9190
	23	0,6270	0,7046	0,7797	0,8663	1,7696	1,9836	2,1685	2,4553
	31	0,6842	0,7567	0,8190	0,8983	1,6986	1,8682	2,0435	2,2577
	47	0,7771	0,8373	0,8909	0,9560	1,5905	1,7168	1,8470	2,0105
	63	0,8172	0,8697	0,9217	0,9825	1,5460	1,6572	1,7644	1,8902
	100	0,8843	0,9337	0,9764	1,0267	1,4797	1,5685	1,6472	1,7482
	200	0,9746	1,0137	1,0460	1,0841	1,4002	1,4553	1,5030	1,5650
	350	1,0309	1,0647	1,0887	1,1176	1,3612	1,4005	1,4343	1,4738
	500	1,0621	1,0867	1,1094	1,1349	1,3386	1,3706	1,3986	1,4330
	1 000	1,1082	1,1278	1,1442	1,1642	1,3064	1,3281	1,3481	1,3701
b_Q^+	7	0,3565	0,4110	0,4581	0,5200	0,9037	0,9326	0,9534	0,9696
	15	0,4306	0,4794	0,5210	0,5693	0,8579	0,8860	0,9064	0,9287
	23	0,4759	0,5181	0,5559	0,5940	0,8357	0,8607	0,8795	0,9038
	31	0,5105	0,5480	0,5768	0,6108	0,8235	0,8477	0,8671	0,8863
	47	0,5533	0,5830	0,6073	0,6347	0,8050	0,8247	0,8418	0,8594
	63	0,5734	0,5992	0,6195	0,6446	0,7944	0,8136	0,8286	0,8444
	100	0,6042	0,6251	0,6418	0,6603	0,7817	0,7970	0,8103	0,8254
	200	0,6414	0,6541	0,6659	0,6791	0,7644	0,7756	0,7856	0,7953
	350	0,6618	0,6724	0,6803	0,6900	0,7550	0,7634	0,7707	0,7806
	500	0,6713	0,6793	0,6866	0,6947	0,7499	0,7569	0,7638	0,7709
	1 000	0,6869	0,6929	0,6979	0,7036	0,7418	0,7472	0,7516	0,7565

	n	1%	2,5%	5%	10%	90%	95%	97,5%	99%
b_L	7	-0,3021	-0,2552	-0,2054	-0,1449	0,4178	0,5036	0,5780	0,6573
	15	-0,0892	-0,0606	-0,0350	-0,0027	0,2583	0,2997	0,3352	0,3792
	23	-0,0358	-0,0144	0,0061	0,0317	0,2215	0,2499	0,2739	0,3057
	31	-0,0106	0,0108	0,0285	0,0474	0,2038	0,2273	0,2477	0,2710
	47	0,0191	0,0350	0,0479	0,0635	0,1861	0,2025	0,2183	0,2359
	63	0,0337	0,0471	0,0583	0,0722	0,1745	0,1906	0,2043	0,2218
	100	0,0536	0,0642	0,0733	0,0837	0,1637	0,1754	0,1861	0,1974
	200	0,0732	0,0805	0,0875	0,0954	0,1500	0,1575	0,1651	0,1733
	350	0,0858	0,0913	0,0967	0,1027	0,1437	0,1497	0,1547	0,1605
	500	0,0913	0,0966	0,1008	0,1055	0,1399	0,1448	0,1492	0,1543
	1 000	0,1007	0,1043	0,1073	0,1106	0,1347	0,1380	0,1410	0,1448
b_{TL}	7	-0,6151	-0,5662	-0,5093	-0,4227	0,6911	0,8453	0,9567	1,0558
	15	-0,1394	-0,1155	-0,0906	-0,0609	0,2138	0,2579	0,2980	0,3497
	23	-0,0806	-0,0585	-0,0410	-0,0197	0,1598	0,1876	0,2123	0,2417
	31	-0,0522	-0,0352	-0,0204	-0,0029	0,1397	0,1621	0,1833	0,2062
	47	-0,0237	-0,0103	0,0008	0,0144	0,1183	0,1337	0,1486	0,1668
	63	-0,0099	0,0006	0,0099	0,0209	0,1088	0,1229	0,1359	0,1516
	100	0,0057	0,0139	0,0221	0,0308	0,0976	0,1078	0,1173	0,1277
	200	0,0228	0,0282	0,0338	0,0401	0,0854	0,0924	0,0985	0,1053
	350	0,0325	0,0374	0,0412	0,0461	0,0800	0,0849	0,0892	0,0945
	500	0,0373	0,0410	0,0449	0,0486	0,0770	0,0812	0,0849	0,0889
	1 000	0,0447	0,0474	0,0499	0,0527	0,0724	0,0753	0,0778	0,0807
b_{LQ}	7	-0,5229	-0,4471	-0,3727	-0,2774	0,4663	0,5844	0,6973	0,8175
	15	-0,4011	-0,3446	-0,2882	-0,2177	0,3519	0,4477	0,5334	0,6355
	23	-0,3012	-0,2466	-0,1970	-0,1343	0,3200	0,3897	0,4552	0,5348
	31	-0,2694	-0,2193	-0,1758	-0,1246	0,2704	0,3338	0,3868	0,4551
	47	-0,1120	-0,0733	-0,0395	0,0004	0,2922	0,3359	0,3727	0,4201
	63	-0,1089	-0,0740	-0,0426	-0,0051	0,2698	0,3104	0,3479	0,3926
	100	-0,0455	-0,0164	0,0094	0,0388	0,2534	0,2863	0,3142	0,3474
	200	-0,0058	0,0145	0,0340	0,0546	0,2004	0,2232	0,2417	0,2641
	350	0,0111	0,0263	0,0399	0,0558	0,1681	0,1844	0,1983	0,2141
	500	0,0314	0,0429	0,0543	0,0679	0,1608	0,1743	0,1865	0,2027
	1 000	0,0591	0,0687	0,0769	0,0859	0,1520	0,1615	0,1698	0,1799

Příloha 7: Základní popisné charakteristiky proměnných z teoretických pravděpodobnostních rozdělení

Tabulka P.2: Základní popisné charakteristiky proměnných z teoretických pravděpodobnostních rozdělení použitých v experimentu:

Rozdělení	$\overline{X}$	$\tilde{X}_{50}$	S_X	V_X	α_M	β_M
LN(0,3162)	1,0513	1,00	0,34	0,32	1,01	4,86
LN(1)	1,6478	1,00	2,16	1,31	6,18	113,94
LN(0,5)	1,1331	1,00	0,60	0,53	1,75	8,90
B(0,5; 0,5)	0,5000	0,50	0,35	0,71	0,00	1,50
B(3; 4)	0,4286	0,42	0,17	0,41	0,18	2,44
B(1,5; 0,5)	0,2495	0,16	0,25	1,00	1,00	3,00
B(2; 1)	2/3	0,7	0,24	0,35	-0,57	2,40
B(1; 2)	1/3	0,3	0,24	0,71	0,57	2,40
Laplaceovo	0	0	2	-	0	6,00
N(10; 1)	10	10	1	0,10	0	3,00
Par(4)	0,3333	0,19	0,47	1,41	0,87	2,51
Par(7)	0,1667	0,10	0,20	1,18	3,25	22,67
Power(0,4)	0,2857	0,18	0,29	1,02	0,86	2,50
Power(0,75)	0,4286	0,40	0,30	0,70	0,25	1,81
Power(3)	0,7500	0,80	0,19	0,26	-0,86	3,09
Power(5)	0,8333	0,87	0,14	0,17	-1,18	4,20
R(0; 1)	0,5	0,5	0,29	0,58	0	1,80
Weibull(1,5)	0,8996	0,78	0,61	0,68	1,07	4,39
Weibull(1)	1	0,69	1,00	1	2	9
Weibull(2)	0,8862	0,83	0,46	0,52	0,63	3,24
Weibull(3)	0,8929	0,88	0,32	0,36	0,17	2,73
Gumbelovo	0,5772	0,37	1,28	2,22	1,14	5,40
Logistické	0	0	1,81	-	0	4,20