

ProbCalc: Manuál

Adam Karaba

Účel

Program ProbCalc původně vznikl, jako nástroj k testování různých generátorů náhodných čísel v průběhu vývoje metody určené k optimalizaci parametrů kinetického modelu pyrolýzy uhlovodíků kdy bylo třeba jednak navrhnout vhodnou optimalizační metodu založenou na stochastických postupech, jednak vybrat vhodný generátor pseudo-náhodných čísel a implementovat transformační metody pro generování pseudo-náhodných čísel do různých rozdělení pravděpodobnosti [1].

Při této vývojové činnosti bylo zjištěno, že v podstatě nejsou dostupné implementace pravděpodobnostních rozdělení ve formě zdrojových kódů nebo knihoven pod volnou licenci, ba naopak, že se vždy jedná o součást softwaru, které je komerční povahy. Jelikož implementace výpočtů hodnot distribučních funkcí není triviální a implementace inverzního postupu je ještě podstatně složitější, rozhodli jsme se tento software poskytnout v případě nekomerčního použití pod licencí Creative Commons (viz oddíl Licence).

Výpočet kritických hodnot nebo hodnot pravděpodobnosti implementovaných rozdělení je v tomto programu umožněn. Stejně tak možné generovat pseudo-náhodné posloupnosti s vybraným rozdělením pravděpodobnosti a tyto posloupnosti exportovat. Je umožněno snadno provést χ^2 test dobré shody s vygenerovanými nebo importovanými posloupnostmi čísel.

Licence

Program ProbCalc i tento manuál byly vytvořeny Adamem Karabou a podléhají licenci [Creative Commons Uved'te autora-Neužívejte dílo komerčně-Nezasahujte do díla 3.0 Česká republika](#).

Používáním díla vyjadřujete svůj souhlas s podmínkami licence.

Obsah

1	Průvodce pro rychlé seznámení.....	9
1.1	Práce s pravděpodobnostním rozdělením.....	9
1.1.1	Přímý výpočet: hustota a pravděpodobnost	9
1.1.2	Inverzní výpočet: kritická hodnota rozdělení.....	10
1.2	Generování pseudo-náhodných čísel s požadovaným rozdělením.....	11
1.3	Práce s daty a formátování čísel	12
1.4	Prokládání rozdělení a test dobré shody	13
2	Technické aspekty vestavěného generátoru pseudonáhodných čísel	15
3	Transformace pseudonáhodných čísel na zadané rozdělení.....	15
3.1	Inverzní metoda	15
3.2	Vylepšená inverzní metoda	16
3.3	Empirické postupy	17
4	Implementovaná rozdělení pravděpodobnosti.....	17
4.1	Standardní rovnoměrné rozdělení (Uniform std).....	17
4.2	Rovnoměrné rozdělení (Uniform)	18
4.3	Normované normální rozdělení (Normal std)	18
4.4	Normální rozdělení (Normal)	18
4.5	χ -kvadrát rozdělení (Chi2).....	19
4.6	t-rozdělení (Student)	20
4.7	Logaritmicko-normální rozdělení (LogNormal)	22
4.8	Exponenciální rozdělení (Exponential)	23
4.9	Fisher-Snedecorovo rozdělení (F-distribution)	24
4.10	Cauchyho-Lorenzovo rozdělení (Cauchy).....	25
4.11	Logistické rozdělení (Logistic).....	25
4.12	Beta-rozdělení (Beta).....	25
4.13	Rayleighovo rozdělení (Rayleigh).....	28
4.14	Weibullovo rozdělení (Weibull).....	29
4.15	Paterovo rozdělení (Pareto)	30
4.16	Rozdělení extrémní hodnoty (ExtrValue).....	31
5	Literatura	33

1 Průvodce pro rychlé seznámení

Tato kapitola je stručným průvodcem pro používání vyvinutého software ProbCalc (dále jen program), zahrnuje několik ilustrací pro snadnější pochopení postupu práce s programem.

1.1 Práce s pravděpodobnostním rozdělením

Program umožňuje vyčíslit hustotu pravděpodobnosti a hodnotu distribuční funkce pro zadané hodnoty náhodné veličiny na základě zadaného tvaru rozdělení spolu s jeho parametry. Umožňuje také inverzní postup: tj. vyčíslení kritické hodnoty pro zadanou pravděpodobnost. Postup provedené těchto výpočtů bude nyní ilustrován na dvou příkladech.

1.1.1 Přímý výpočet: hustota a pravděpodobnost

Mějme náhodnou veličinu X , která se řídí normálním rozdělením s parametry $\mu = 3,5$; $\sigma = 1,25$. Následuje ukázka, jakým způsobem vypočítat hodnotu distribuční funkce (tj. pravděpodobnost, že veličina nabude dané a/nebo nižší hodnoty) pro hodnoty náhodné veličiny $X_i = \{-1 ; 3 ; 5,5\}$:

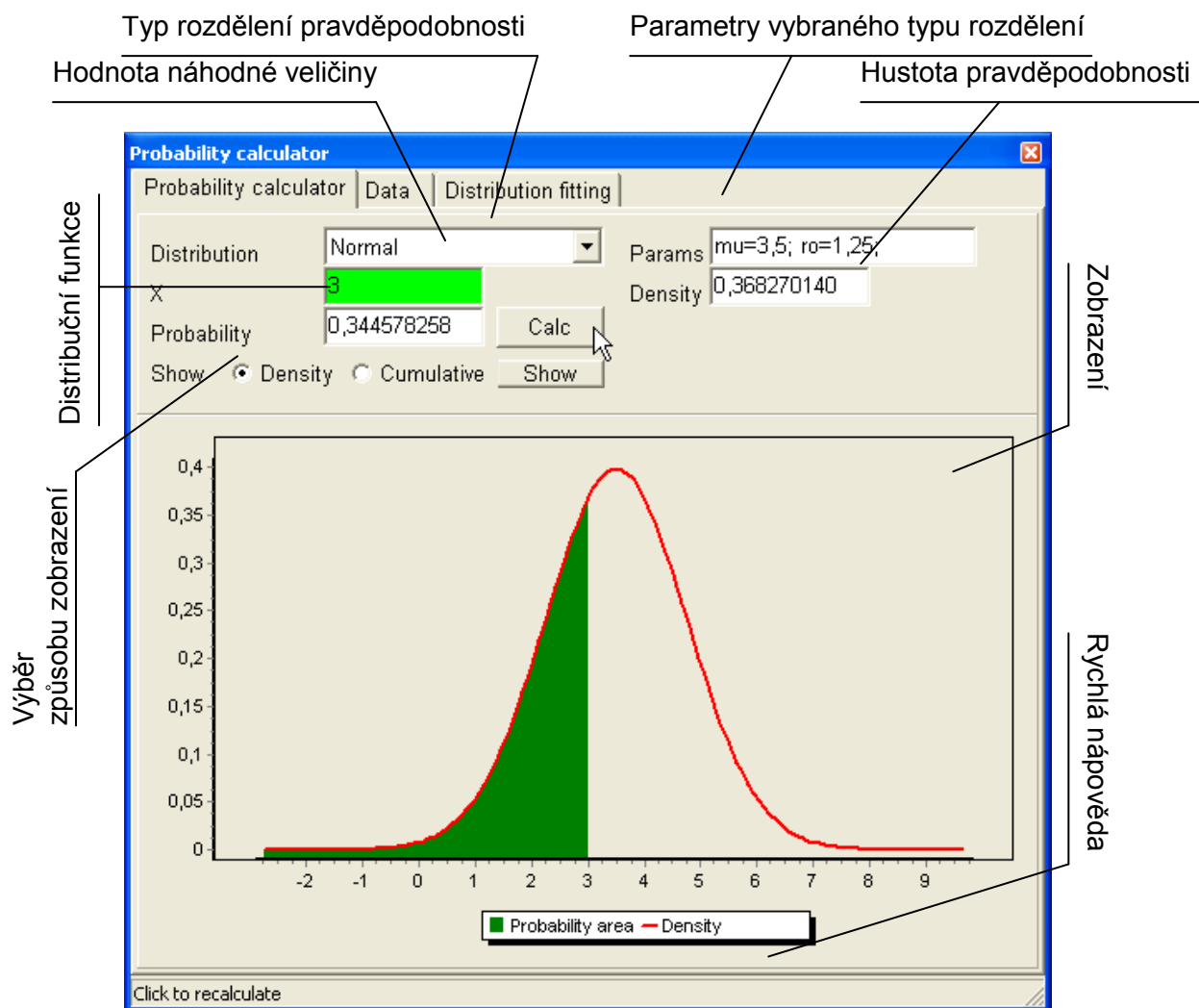
- Po spuštění programu přejděte na kartu **Probability calculator**
- V comboboxu **Distribution** zvolte druh rozdělení (v příkladu Normální rozdělení **Normal**). Pokud je to požadováno, budete dotázáni na hodnoty parametrů rozdělení, pro příklad:
 - Střední hodnota μ (**mu**) je rovna 3,5
 - Hodnota směrodatné odchylky σ (**ro**) je rovna 1,25

Zadané parametry se zobrazí v editboxu **Params** vedle zvoleného rozdělení¹².

- Do pole hodnoty náhodné veličiny **X** zadejte požadovanou hodnotu, pro kterou má být spočtena hodnota distribuční funkce (v příkladu 3) a stiskněte tlačítko **Calc** nebo klávesu ENTER. Program provede výpočet a zobrazí výsledky. Pokud je některý ze vstupů neplatný, je obarven červeně. Pokud je výpočet v pořádku, obarví se rozhodující vstup zeleně.
- Hodnota distribuční funkce je zobrazena v poli **Probability** (0,344578258), hodnota hustoty pravděpodobnosti v poli **Density** (0,368270140).
- Graf rozdělení je zobrazen ve spodní části okna. Uživatel může pomocí přepínačů (**Density** a **Cumulative**) vybrat jeden ze dvou možných způsobů zobrazení daného rozdělení: hustotu pravděpodobnosti nebo distribuční funkci v závislosti na hodnotě náhodné veličiny X .

¹ Zobrazené parametry rozdělení (po jejich zadání) není možné v editboxu interaktivně měnit. Pokud je zapotřebí parametry změnit, vyberte znovu typ rozdělení a budete dotázáni na hodnoty parametrů.

² Pokud je při zadávání parametrů rozdělení uživatelem předán neplatný vstup (např. písmeno namísto čísla), je tato skutečnost ohlášena a program nepřevezme zadané rozdělení. Projeví se to tím, že daný tvar rozdělení není vybrán. Proveďte znovu výběr rozdělení.

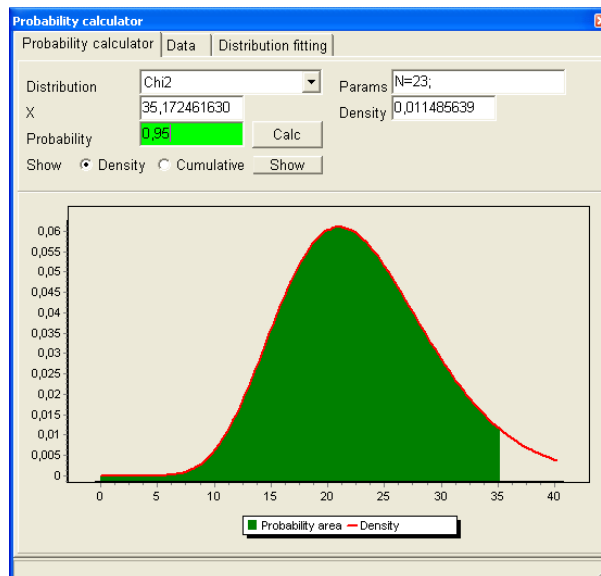


Obrázek 1: Okno programu při výpočtu hodnot distribuční funkce zadaného rozdělení

1.1.2 Inverzní výpočet: kritická hodnota rozdělení

Inverzní postup zahrnuje výpočet takové hodnoty náhodné veličiny se zadaným rozdělením (vč. jeho parametrů), pro niž daná distribuční funkce vrací předem zadanou pravděpodobnost. Mějme veličinu X s rozdělením χ^2 se 23 stupni volnosti. Nalezneme takovou hodnotu, které (a nebo nižší) náhodná veličina X nabude s pravděpodobností 95 %. Z pohledu testování hypotéz se jedná o výpočet kritické hodnoty rozdělení χ^2 se 23 stupni volnosti pro hladinu významnosti 0,05. Poustup je následující (o ovládání jednotlivých prvků programu, které byly diskutovány v odst. 1.1.1 viz tamtéž):

- Po spuštění aplikace přejděte na kartu **Probability calculator**
- V comboboxu **Distribution** vyberte typ rozdělení **chi2**, budete dotázáni na parametry rozdělení: počet stupňů volnosti **N** zadejte 23.
- Do editboxu **probability** zadejte požadovanou hodnotu pravděpodobnosti 0,95 a stiskněte tlačítko **Calc** nebo klávesu ENTER. Program provede výpočet a zobrazí výsledky. Případné nesprávně zadané vstupy budou označeny červeně. V případě úspěšného výpočtu bude rozhodující vstup označen zeleně.
- Editbox veličiny **X** obsahuje hledanou kritickou hodnotu (35,17246163).



Obrázek 2: Okno programu při výpočtu kritické hodnoty rozdělení

1.2 Generování pseudo-náhodných čísel s požadovaným rozdělením

Při některých výpočtech je zapotřebí generovat pseudonáhodná čísla s předem zadaným tvarem rozdělení a parametry. Postup generování bude nyní předveden na příkladu generování 1000 hodnot veličiny mající t-rozdělení (Studentovo) s 11 stupni volnosti:

- Po spuštění programu přejděte na kartu **Data**
- Stiskněte možnost generovat náhodná čísla **Generace RND**, bude zobrazeno modální okno **Random generation form**, proto vyplňte požadované vstupní údaje:

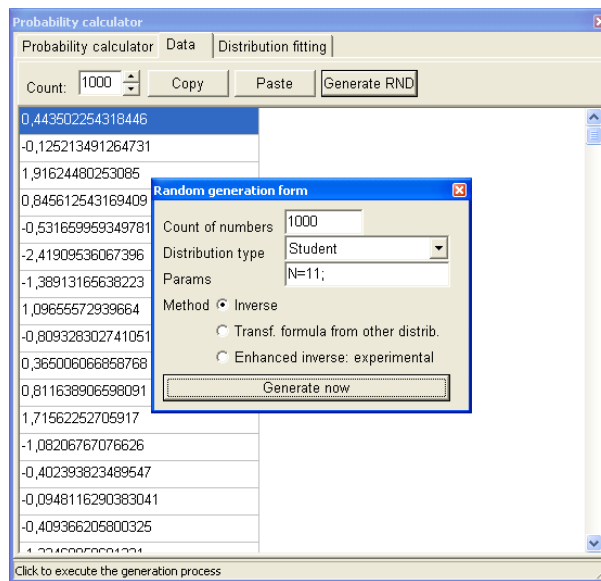
- **Count of Numbers**: Požadovaný počet vygenerovaných hodnot (1000)
- **Distribution**: Tvar rozdělení (podle příkladu **Student**). Po jeho výběru budete dotázáni na parametry rozdělení (pokud jsou zapotřebí). V případě tohoto příkladu:

- Počet stupňů volnosti **N** = 11

Pokud dojde při zadávání parametrů k chybě, bude tato skutečnost oznámena a výběr tvaru rozdělení bude zrušen. V takovém případě a nebo pokud chcete parametry rozdělení změnit, zvolte znovu tvar rozdělení a budete dotázáni na jeho parametry.

- **Method**: Metoda generování náhodných čísel (v případě některých rozdělení není možné měnit):
 - **Inverse**: Transformace čísel inverzní metodou (doporučeno!)
 - **Transf. formula from other distrib.**: Jiný transformační postup (podrobnosti jsou uvedeny v dokumentaci u jednotlivých typů rozdělení). Může výrazně zpomalit proces!
 - **Enhanced inverse**: experimentální je experimentální metoda taktéž popsána dále v dokumentaci. Může výrazně zpomalit proces!

- Stiskněte tlačítko **Generace now**, program provede generování čísel.



Obrázek 3: Generování náhodných čísel se zadaným rozdělením pravděpodobnosti

Po provedení této operace doporučujeme vyzkoušet příklad prokládání rozdělení v odstavci 1.4.

1.3 Práce s daty a formátování čísel

Karta **Data** slouží k práci s načtenými daty a k výměně dat s ostatními aplikacemi. Manipulací s počtem načtených hodnot, dosáhnete prodloužení nebo ořezání velikosti bufferu. Hodnoty lze v zobrazené tabulce přímo editovat. Po opuštění tabulky bude uživatel dotázán zda chce provedené změny aplikovat. K zápisu hodnot používejte standardní formát zápisu reálného čísla v dekadické soustavě. Desetinný oddělovač je načítán z jazykového nastavení systému (aktuálně správný je ten, který je automaticky zapsán z numerické klávesnice). Čísla lze zapisovat i v plovoucí řádové čárce, pozice oddělovače je uvedena znakem „e“ resp. „E“ za nímž následuje směr posunu a počet dekadických míst. Formát zápisu čísla platí pro veškeré vstupy tohoto programu! Uvedme několik příkladů s českým nastavením zápisu čísla:

- 1,25
- -35,78
- 12
- -12,38E-13
- 1,2E16

Pokud si uživatel přeje data exportovat z programu, stiskne tlačítko **Copy**. Data jsou pak přenesena jako sloupcový vektor do schránky ve std. formátu a lze je vložit do kompatibilního software.

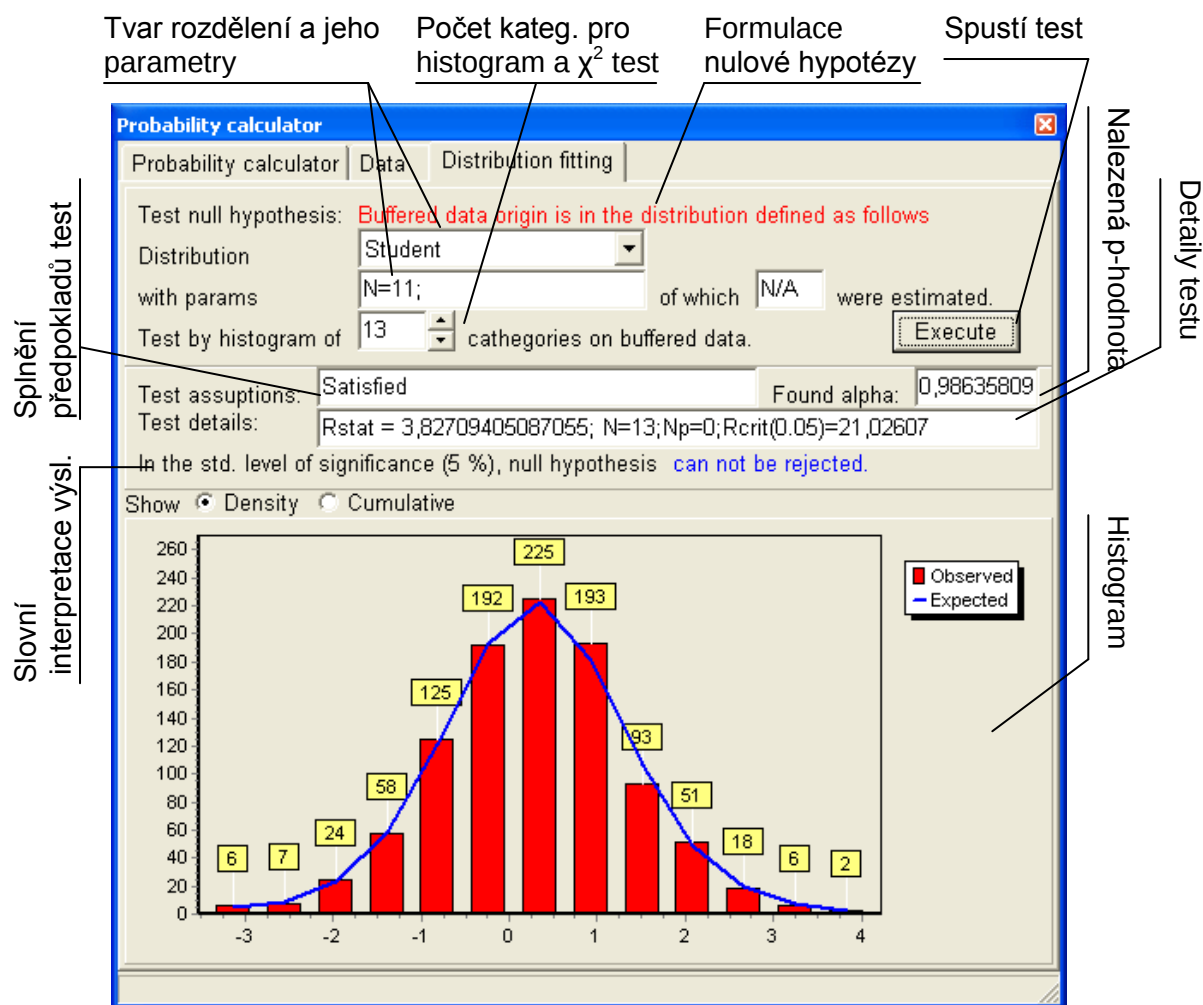
Pokud si uživatel přeje importovat data z jiné aplikace, doporučujeme tato data načíst do schránky a vložit je do programu stiskem tlačítka **Paste** taktéž na kartě **Data**. Běžné tabulkové procesory jako je Matlab, Excel, Statistica apod. jsou plně kompatibilní. Pokud

dojde k pokusu o vložení řádkového vektoru namísto sloupcového, bude uživatel o této skutečnosti informován. Pokud je na vstupu rozpoznána dvojrozměrná tabulka a nebo není struktura rozpoznána vůbec, je operace vkládání zastavena a uživatel je o této skutečnosti informován chybovým hlášením programu.

1.4 Prokládání rozdělení a test dobré shody

Uživatel může s použitím programu otestovat zda předložená data pochází ze základního souboru s definovaným rozdělením a parametry tohoto rozdělení může buď zadat nebo nechat automatizovaně odhadnout. Před provedením následujícího příkladu doporučujeme vyzkoušet si příklad s generováním hodnot z t-rozdělením s 11 stupni volnosti (viz odst. 1.2) neboť tento příklad předpokládá, že taková data máme k dispozici. Tato data jsou vygenerována, ručně zadána nebo importována na kartě **Data** ještě před provedením následujícího postupu. K provedení testu shody je v programu implementován známý χ^2 test dobré shody, který lze realizovat následovně:

- Přejděte na kartu **Distribution fitting**
- Vyberte tvar rozdělení (**distribution**) proti kterému chcete náhodný výběr otestovat (podle příkladu rozdělení **Student**)
 - Bude zobrazeno modální okno, které umožňuje zadat parametry rozdělení. Pokud tyto parametry zadáte, nebudou již dále odhadovány. Pokud necháte políčka **Auto** zaškrtnuta, budou takto označené parametry odhadnuty výpočetně. Zadání potvrďte stiskem tlačítka **OK**.
- Nastavte počet kategorií, do kterých budou data rozdělena pro provedení χ^2 testu dobré shody
- Stiskněte tlačítko **Exekute** k provedení testu.
- Pole **Test assumption** obsahuje hodnocení, za jsou předpoklady χ^2 testu dobré shody splněny z cela, z části nebo hrubě překročeny. Výsledek testu interpretujte s ohledem na tyto okolnosti.
- Pole **found alpha** obsahuje vypočtenou p-hodnotu
- Pole **Test details** obsahuje další podrobnosti
- Je připojeno slovní rozhodnutí o zamítnutí nulové hypotézy.
- Spodní část okna uvádí porovnání očekávaných a pozorovaných četností v jednotlivých kategoriích.



Obrázek 4: Provedení χ^2 testu dobré shody

Parameters estimation

N = 11 ☐ Auto

OK

Obrázek 5: Zadávání parametrů hypotetického rozdělení

2 Technické aspekty vestavěného generátoru pseudonáhodných čísel

Při vývoji aplikace a při vývoji původní optimalizační metody, pro kterou se tento software stal testovacím nástrojem, bylo otestováno velmi mnoho generátorů pseudonáhodných čísel pomocí komerčních balíčků testů na náhodnost.

V souladu s výsledky výzkumníků v této oblasti se pro simulace a výpočty jako nejvýhodnější generátor jeví tzv. Mersennův twister. Implementovaná rutina je varianta známá jako MT19937 [2], dále jen vestavěný generátor.

Při generování reálných čísel dochází k přeskupení 32-bitových slov z výstupu vestavěného generátoru do 52-bitových sekvencí, kterým je upraveno měřítko tak, aby byl výstup tvořen reálnými čísly v rozsahu 0 až 1 s rovnoměrným rozdělením. Tato čísla jsou tedy generována s plnou přesností 52 bitů.

Na takto generovaná čísla jsou následně aplikovány transformační metody podle požadavků uživatele programu (viz odst. 1.2).

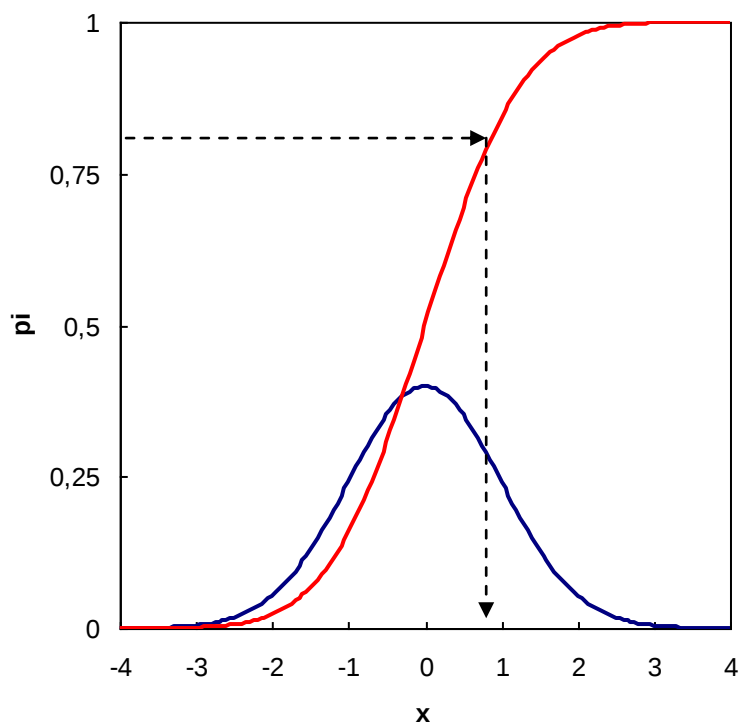
Interní generátor čísel je iniciován z aktuálních hodnot systémového času, počtu úderů časovače od spuštění operačního systému a dalších systémových proměnných. Z takto sestaveného semínka je jednocestnými funkcemi vybudován Seed pro vestavěný generátor. Jedná se o jednocestné funkce doporučené pro tento generátor literárními zdroji [3,4]. Následně je interní generátor zahříván tak dlouho, dokud není dosaženo dostatečné entropie uvnitř jeho Seedu. Generátor není umožněno úmyslně znovu nastavit do stejného stavu jako při posledním spuštění. Pokud si uživatel přeje uchovat posledně vygenerovanou posloupnost čísel, doporučuje se mu, aby si ji uložil.

3 Transformace pseudonáhodných čísel na zadané rozdělení

Tento odstavec dokumentuje transformační postupy použité pro generování pseudonáhodných čísel ze zadaného rozdělení s využitím vestavěného generátoru (postup práce je popsán v odst. 1.2). Vnitřní generátor pseudonáhodných čísel, popsáný v kap. 2, poskytuje čísla se standardním rovnoměrným rozdělením. Pokud je žádána soubor hodnot náhodné veličiny o jiném rozdělení, je použita jedna z transformačních metod (dle výběru uživatele), které jsou popsány v této kapitole.

3.1 Inverzní metoda

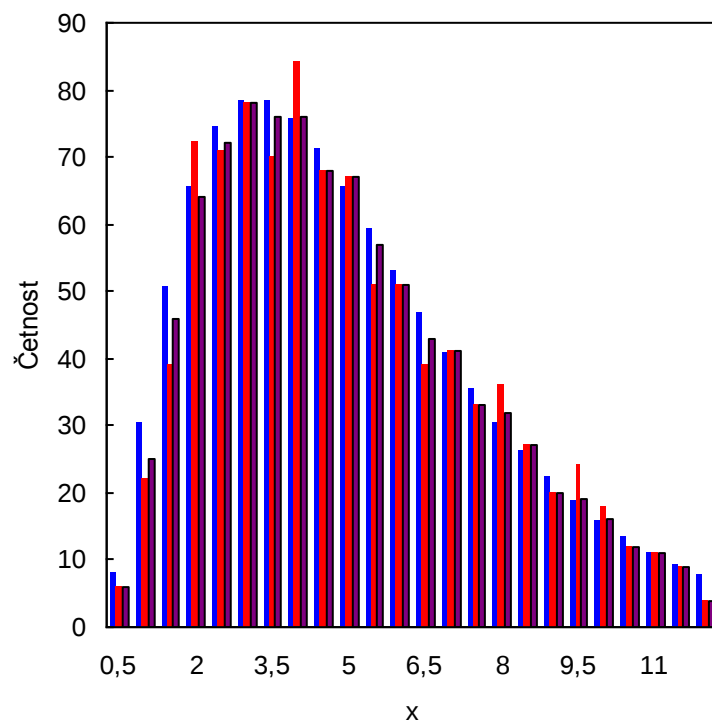
Inverzní transformační metoda (ozn. **inverse**) používá inverzní funkci k distribuční funkci zadaného rozdělení k nalezení hodnot náhodné veličiny. Vstupem je tedy číslo se standardním rovnoměrným rozdělením jehož hodnota vypovídá o pravděpodobnosti, že náhodná veličina se zadaným (požadovaným) rozdělením nabude hledané (nebo nižší) hodnoty. Výstupem je kritická hodnota daného rozdělení o daných parametrech. Postup je ilustrován na obrázku.



Obrázek 6: Ilustrace inverzního postupu při generování čísel s normovaným normálním rozdělením (hustota — a distribuční funkce — daného rozdělení, - ➔ postup transformace)

3.2 Vylepšená inverzní metoda

Vylepšená inverzní metoda (ozn. **Enhanced inverse: experimental**) využívá inverzní metodu, popsanou v odst. 3.1, ke generování jednotlivých posloupností. Na vygenerovanou posloupnost se následně aplikuje postup, který rozčlení data do kategorií a porovná četnost jednotlivých kategorií s očekávanou hodnotou četnosti zjištěnou teoreticky s použitím distribuční funkce požadovaného rozdělení. Následně jsou pomocí inverzní metody přepsány ty hodnoty, které nevyhovují zadanému rozdělení přičemž je pravděpodobnost přepisu přímo úměrná přebytku četnosti v kategorii, do které daná hodnota patří. Tento postup je experimentální a výsledné posloupnosti splňují velmi dobře požadavek na tvar a parametry rozdělení, jeho nevýhodou je značná výpočetní náročnost, resp. delší výpočetní čas.



Obrázek 7: Demonstrace vylepšené inverzní metoda při aplikaci na rozdělení χ^2 se 7 stupni volnosti: histogram četnosti v generované posloupnosti - ■ očekávaná četnost (z hypotetického = požadovaného rozdělení), ■ pozorovaná četnost při prvním průchodu algoritmu, ■ pozorovaná četnost při druhém průchodu algoritmu (tj. po první opravě)

3.3 Empirické postupy

Empirické postupy (ozn. transf formula from other distrib.) jsou založeny na empirických vztazích mezi jednotlivými rozděleními a budou vysvětleny pro jednotlivá rozdělení pravděpodobnosti.

4 Implementovaná rozdělení pravděpodobnosti

Tato kapitola popisuje rozdělení pravděpodobnosti, která byla do softwaru implementována, stručně popisuje význam a označené parametry.

4.1 Standardní rovnoměrné rozdělení (Uniform std)

Má definovanu hustotu pravděpodobnosti (1):

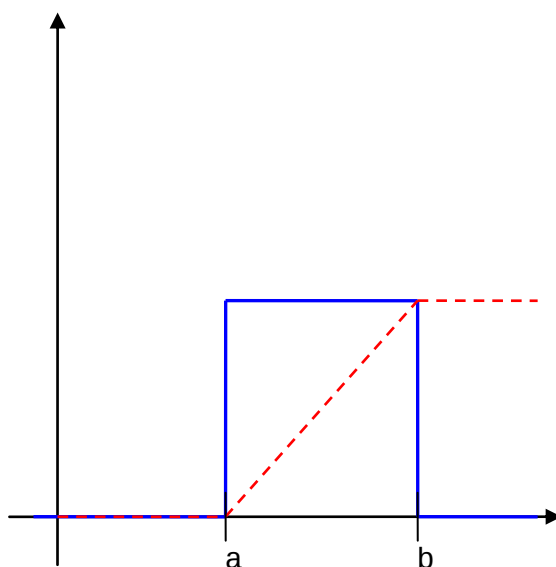
$$f(x) = \begin{cases} 1 & \text{pro } x \in (0;1) \\ 0 & \text{pro } x \notin (0;1) \end{cases} \quad (1)$$

a jedná se tedy o speciální případ rovnoměrného rozdělení pravděpodobnosti (viz odst. 4.2), kde budou uvedeny další podrobnosti.

4.2 Rovnoměrné rozdělení (Uniform)

Je definováno hustotou (2) s parametry $-\infty < a < b < \infty$, které jsou v software pro intuitivnost označovány jako minimum a maximum. Hustota a distribuční funkce jsou zobrazeny na obrázku Obrázek 8.

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{pro } x \in (a; b) \\ 0 & \text{pro } x \notin (a; b) \end{cases} \quad (2)$$



Obrázek 8: Hustota pravděpodobnosti — a distribuční funkce - - - rovnoměrného rozdělení

Jeho střední hodnota a rozptyl jsou dány relacemi (3):

$$E(X) = \frac{a+b}{2} \quad (3)$$
$$D(X) = \frac{(b-a)^2}{12}$$

4.3 Normované normální rozdělení (Normal std)

Normované normální rozdělení je Normální rozdělení (viz odst. 4.4) s parametry $\mu = 0$; $\sigma = 1$.

4.4 Normální rozdělení (Normal)

Veličina mající normální rozdělení je podle centrální limitní věty (odkazy) vytvořena součtem mnoha veličin o rovnoměrném rozdělení pravděpodobnosti (viz odst. 4.2) a jde tedy o přirozené rozdělení, ze kterého pochází většina fyzikálně měřitelných veličin. Toto

rozdělení s parametry $-\infty < \mu < \infty$ (střední hodnota ozn. **mu**) a $0 < \sigma < \infty$ (směrodatná odchylka ozn. **ro**) je dáno hustotou pravděpodobnosti:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4)$$

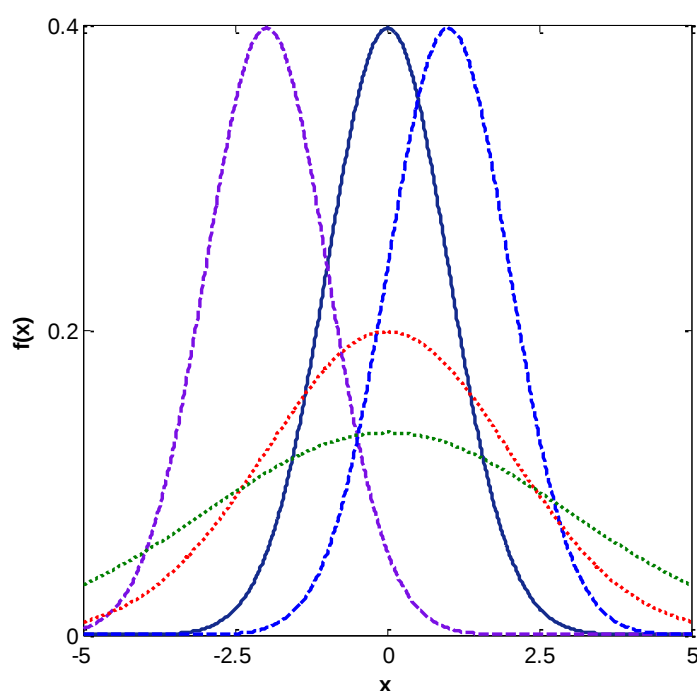
Jejíž tvar pro různé volby parametrů já znázorněn na obrázku. Distribuční funkci

$$F(x) = \int_{-\infty}^x f(x) dx \quad (5)$$

nelze zapsat s použitím elementárních funkcí. Střední hodnota a rozptyl náhodné veličiny s tímto rozdělením jsou dány relacemi:

$$E(X) = \mu \quad (6)$$

$$D(X) = \sigma^2$$



Obrázek 9: Hustota pravděpodobnosti normálního rozdělení pro hodnoty parametrů: — $\mu = 0; \sigma = 1$ (normované), $\mu = 0; \sigma = 2$, $\mu = 0; \sigma = 3$, - - - $\mu = 1; \sigma = 1$; - - - $\mu = -2; \sigma = 1$

4.5 χ -kvadrát rozdělení (*Chi2*)

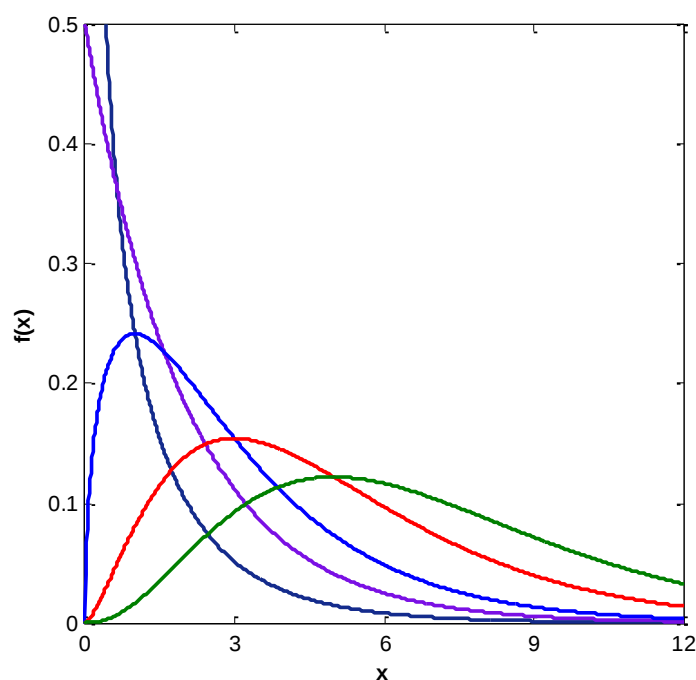
Rozdělení χ^2 o $N > 0$ stupni volnosti (ozn. **N**) je dáno hustotou:

$$f(x) = \begin{cases} 0 & \text{pro } x \leq 0 \\ \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} & \text{pro } x > 0 \end{cases} \quad (7)$$

jejíž příklady jsou zobrazeny na obrázku. Toto rozdělení se používá v souvislosti s jeho definicí kdy veličina Y pochází z χ^2 rozdělení s N stupni volnosti pokud platí:

$$Y = \sum_{i=1}^N X_i^2 \quad (8)$$

kde X_i je N vzájemně nezávislých veličin veličiny pocházejících z normovaného normálního rozdělení $N(0; 1)$, podrobnosti viz odst. 4.1. Tuto definici lze rovněž využít jako empirický postup pro transformaci hodnot pocházejících z $N(0; 1)$ do $\chi^2(N)$.



Obrázek 10: Hustota pravděpodobnosti rozdělení χ^2 pro stupně volnosti: — $N = 1$, — $N = 2$, — $N = 3$, — $N = 5$, — $N = 7$

4.6 t -rozdělení (Student)

Studentovo rozdělení nebo též t -rozdělení $t(N)$ s $N > 0$ stupni volnosti (ozn. $\underline{N}$) je definováno hustotou pravděpodobnosti:

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \left(\frac{x^2}{n} + 1\right)^{-\frac{n+1}{2}} \quad (9)$$

Mějme dvě nezávislé veličiny U a V mající normované normální rozdělení (viz odst. 4.3) a rozdělení $\chi^2(N)$ s N stupni volnosti, pak rozdělení veličiny

$$X = \frac{U}{\sqrt{\frac{V}{n}}} \quad (10)$$

nazýváme Studentovým rozdělením (příp. t-rozdělením) s N stupni volnosti $t(N)$. Tento vztah je hojně využíván ve statistice a také k transformacím souborů hodnot mezi těmito rozděleními. Kromě uvedeného postupu lze rovněž provést méně přesnou transformaci hodnot s použitím relace:

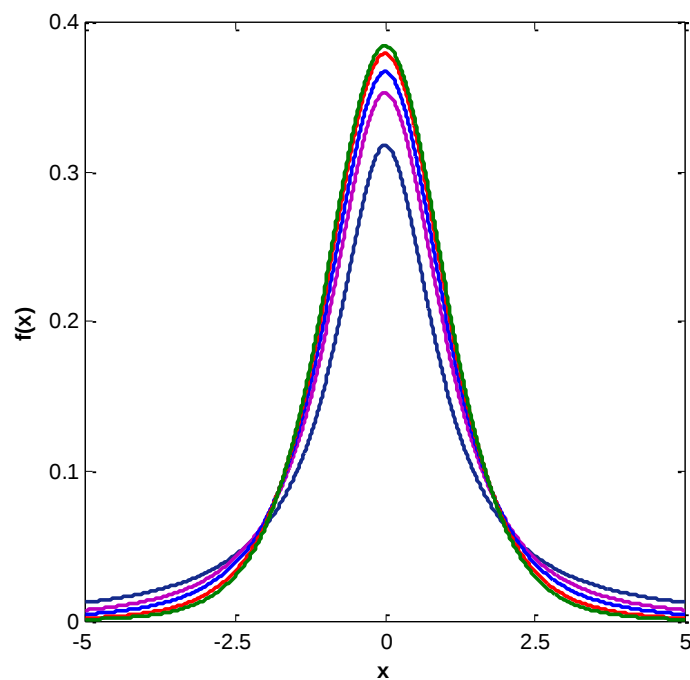
$$t = z\sqrt{N-1} \quad (11)$$

kde veličina z pochází ze souboru s normovaným normálním rozdělením a veličina t pak vykazuje znaky t-rozdělení s N stupni volnosti.

Střední hodnota a rozptyl základního souboru majícího t-rozdělení je dáno:

$$\begin{aligned} E(X) &= 0 \\ D(X) &= \frac{N}{N-2} \end{aligned} \quad (12)$$

pro $N > 2$.

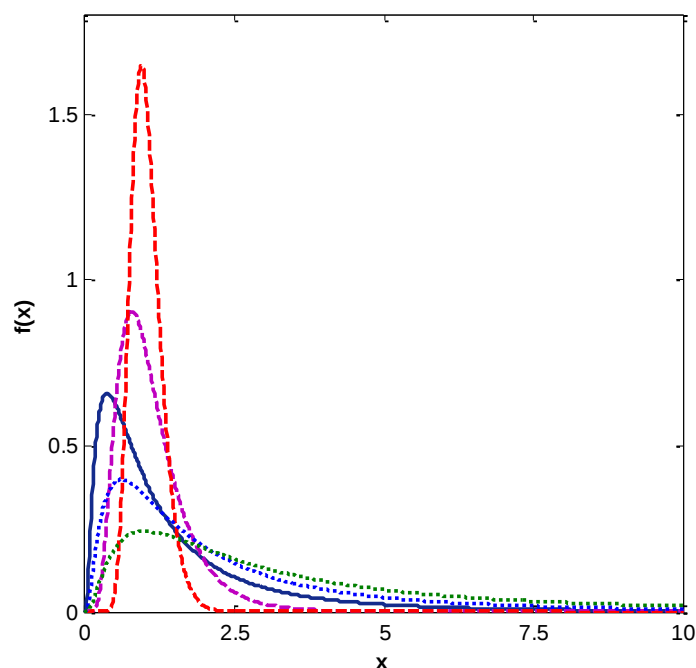


Obrázek 11: Hustota pravděpodobnosti Studentova rozdělení pro stupně volnosti: — $N = 1$, — $N = 2$,
— $N = 3$, — $N = 5$, — $N = 7$

4.7 Logaritmicko-normální rozdělení (LogNormal)

Logaritmicko-normální rozdělení $LN(\mu; \sigma)$ s parametry μ (ozn. **mu**) a σ (ozn. **ro**) je rozdělení takové veličiny X , že veličina $Y = \ln(X)$ má normální rozdělení $N(\mu; \sigma)$ s parametry střední hodnoty μ a směrodatné odchylky σ . Jeho hustota je dán předpisem:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} \text{ pro } x > 0 \quad (13)$$



Obrázek 12: Hustota pravděpodobnosti logaritmicko normálního rozdělení $LN(\mu; \sigma)$ pro hodnoty parametru: — $LN(0; 1)$, - - $LN(0; 0,5)$, - - $LN(0; 0,25)$, $LN(0,5; 1)$, $LN(1; 1)$

Empirická transformační formule pro veličinu Y pocházející z normálního rozdělení s parametry μ a σ ve tvaru

$$X = e^Y \quad (14)$$

zavádí veličinu X s logaritmicko-normálním rozdělením s parametry μ a σ , příp. lze použít alternativní transformaci:

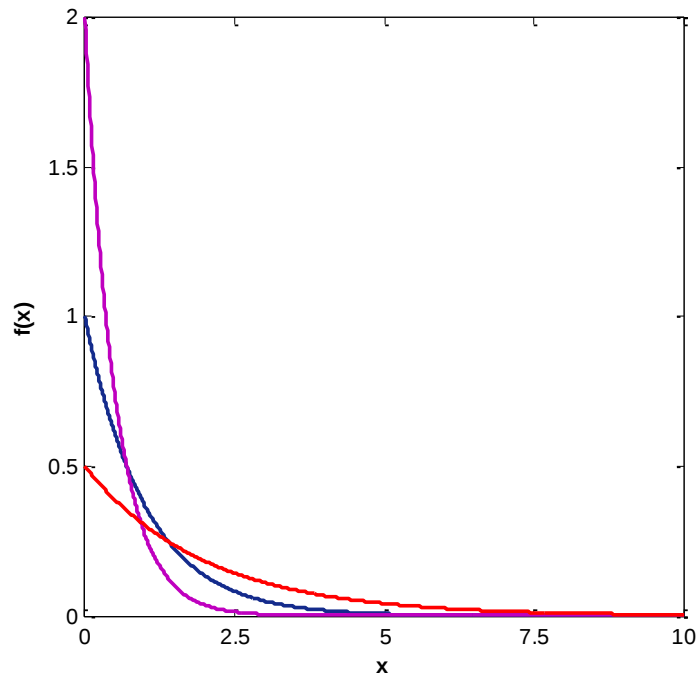
$$X = e^{\mu + \sigma Y} \quad (15)$$

pro Y pocházejícího ze souboru s normovaným normálním rozdělením.

4.8 Exponenciální rozdělení (Exponential)

Náhodná veličina X pochází z exponenciálního rozdělení s parametrem $\lambda > 0$ (ozn. **lambda**) právě tehdy když má její hustota pravděpodobnosti tvar:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{pro } x \geq 0 \\ 0 & \text{pro } x < 0 \end{cases} \quad (16)$$



Obrázek 13: Hustota pravděpodobnosti exponenciálního rozdělení pro hodnoty parametru: — $\lambda = 1$,
— $\lambda = 2$; — $\lambda = 0,5$

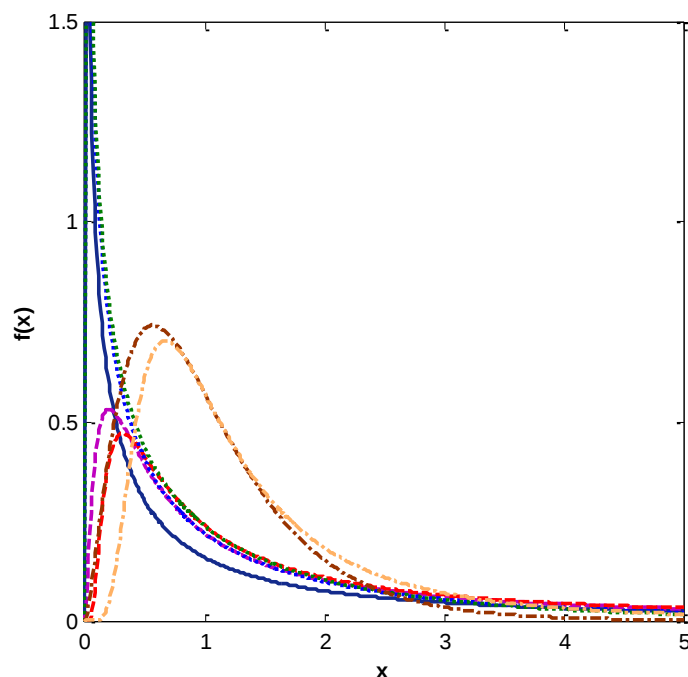
4.9 Fisher-Snedecorovo rozdělení (F-distribution)

Mějme vzájemně nezávislé náhodné veličiny U_1 a U_2 , které pocházejí ze základních souborů o χ^2 rozdělení s počty stupňů volnosti N_1 a N_2 , pak rozdělení veličiny X vzniklé relací

$$X = \frac{U_1/N_1}{U_2/N_2} \quad (17)$$

nazveme Fisher-Snedecorovým rozdělením (zkr. F-rozdělením) o parametrech N_1, N_2 (ozn. $\mathbf{N}, \mathbf{M}$), jehož hustota pravděpodobnosti je dána vztahem:

$$f(x) = \frac{\Gamma\left(\frac{N_1 + N_2}{2}\right) N_1^{\frac{N_1}{2}} N_2^{\frac{N_2}{2}}}{\Gamma\left(\frac{N_1}{2}\right) \Gamma\left(\frac{N_2}{2}\right)} \frac{x^{\frac{N_1}{2}-1}}{(N_1 + N_2 x)^{\frac{N_1 + N_2}{2}}} \quad \text{pro } x \geq 0 \quad (18)$$



Obrázek 14: Tvar hustoty pravděpodobnosti F-rozdělení $F(N_1, N_2)$ pro kombinace parametrů:
— $F(1; 1)$, - - $F(5; 1)$, - . $F(35; 1)$, $F(1; 5)$, $F(1; 35)$, - . $F(5; 35)$, - . $F(35; 5)$

4.10 Cauchyho-Lorenzovo rozdělení (Cauchy)

Cauchyho (též Cauchyho-Lorenzovo) rozdělení o parametrech umístění a (ozn. a) a měřítko λ (ozn. lambda) má hustotu danu vztahem:

$$f(x) = \frac{1}{\pi\lambda \left[1 + \left(\frac{x-a}{\lambda} \right)^2 \right]} \quad (19)$$

4.11 Logistické rozdělení (Logistic)

Logistické rozdělení o parametrech μ (ozn. mu) a $s > 0$ (ozn. s) je dáno hustotou:

$$f(x) = \frac{e^{-\frac{x-\mu}{s}}}{s \left(1 + e^{-\frac{x-\mu}{s}} \right)^2} \quad (20)$$

4.12 Beta-rozdělení (Beta)

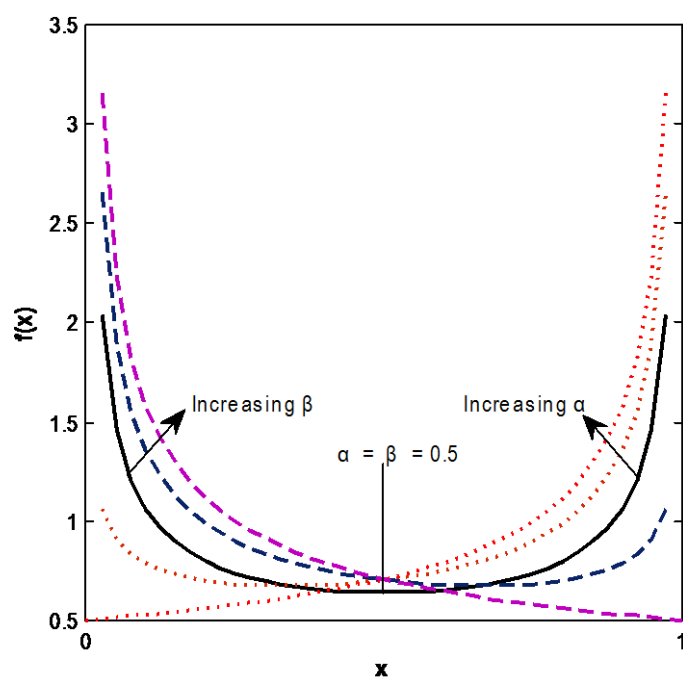
Beta-rozdělení o parametrech $\alpha, \beta > 0$ (ozn. alpha, beta) je dáno hustotou ve tvaru:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{C(\alpha, \beta)} \text{ pro } 0 \leq x \leq 1 \quad (21)$$

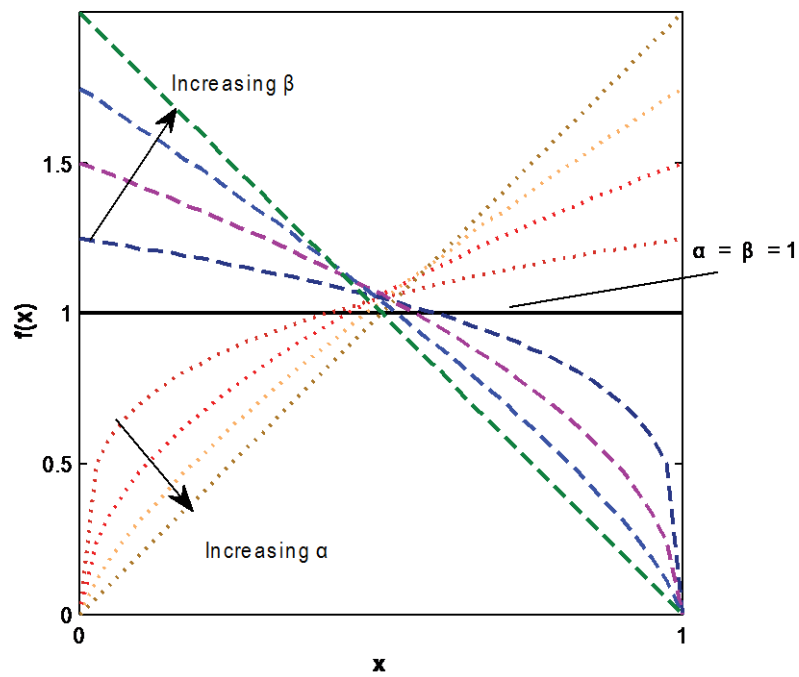
kde je normovací konstanta C dána výrazem:

$$C(\alpha, \beta) = \int_0^1 w^{\alpha-1}(1-w)^{\beta-1} dw = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \quad (22)$$

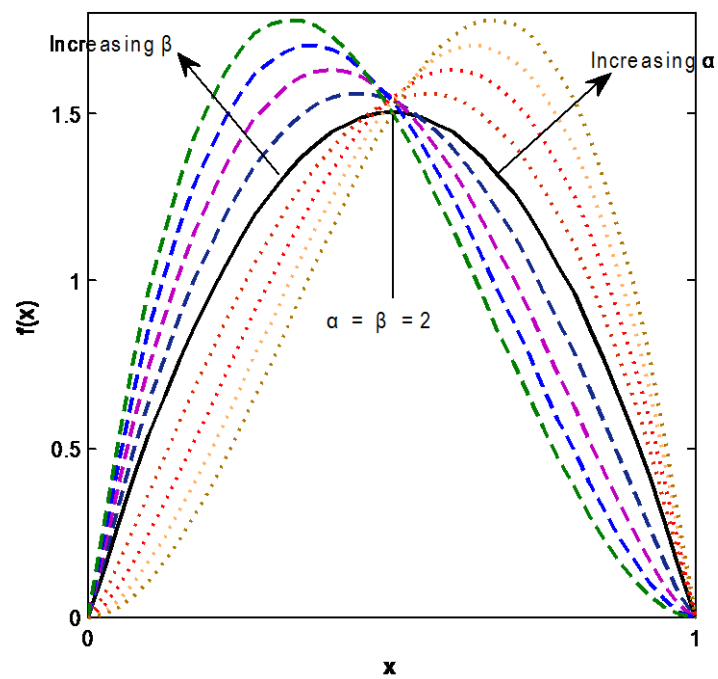
Tvar tohoto rozdělení je velmi proměnný a složitý jak ilustrují následující grafy.



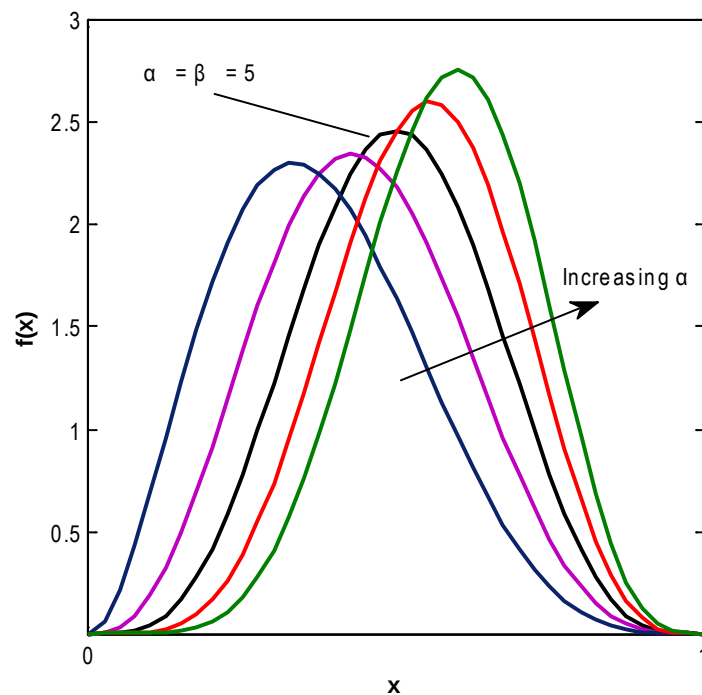
Obrázek 15: Tvar hustoty beta rozdělení s proměnnými parametry, pro fixovaný parametr = 0,5



Obrázek 16: Tvar hustoty beta rozdělení s proměnnými parametry, pro fixovaný parametr = 1



Obrázek 17: Tvar beta rozdělení s proměnnými parametry, pro fixovaný parametr = 2



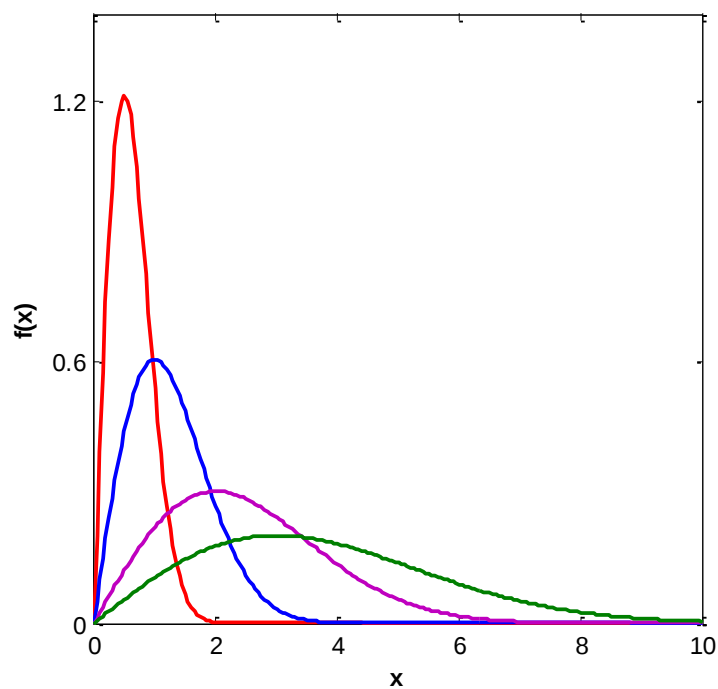
Obrázek 18: Tvar beta rozdělení s fixním $N_2 = 5$ a proměnným $N_1 = 3 \dots 7$

4.13 Rayleighovo rozdělení (Rayleigh)

Rayleighovo rozdělení s parametrem σ (ozn. **ro**) má hustotu pravděpodobnosti danu předpisem:

$$f(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \quad \text{pro } x \geq 0 \quad (23)$$

Hustota je pro několik hodnot parametrů znázorněna na obrázku.



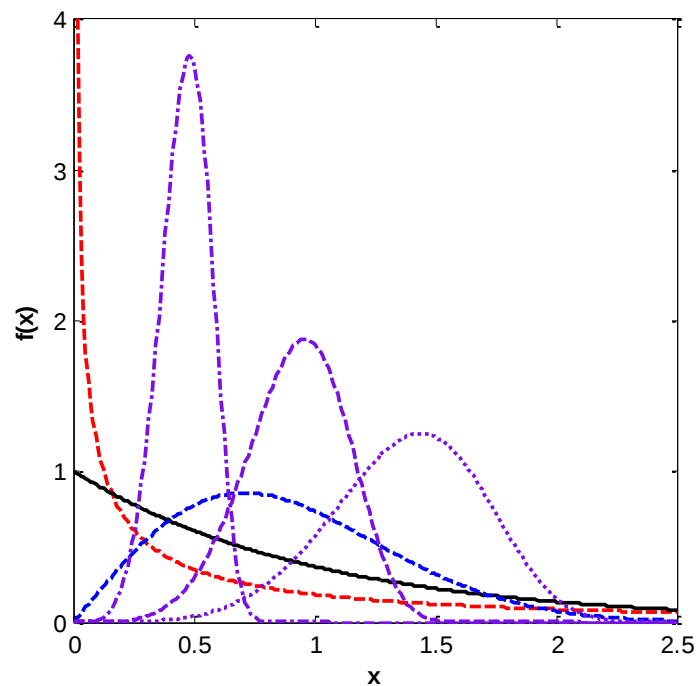
Obrázek 19: Hustota Rayleighova rozdělení pro parametry měřítka: — $\sigma = 0,5$; — $\sigma = 1$; — $\sigma = 2$; — $\sigma = 3$

4.14 Weibullovo rozdělení (Weibull)

Weibullovo rozdělení s parametry $k > 0$ (tvar ozn. **k**) a $\lambda > 0$ (měřítko, ozn. **lambda**) má následující tvar hustoty:

$$f(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda} \right)^{k-1} e^{-\left(\frac{x}{\lambda} \right)^k} \quad \text{pro } x \geq 0 \quad (24)$$

Tato hustota je pro některé vybrané hodnoty parametrů předvedena na obrázku.”

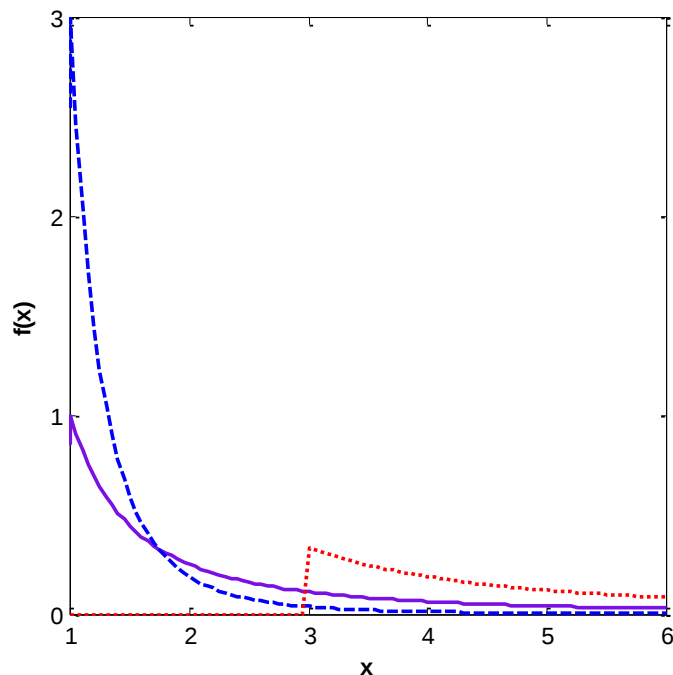


Obrázek 20: Weibullovo rozdělení pro hodnoty parametrů: — $\lambda = 1$; $k = 1$; - - $\lambda = 1$; $k = 0,5$; - - $\lambda = 1$; $k = 2$; - - $\lambda = 1$; $k = 5$; — $\lambda = 0,5$; $k = 5$; $\lambda = 1,5$; $k = 5$

4.15 Paterovo rozdělení (Pareto)

Paretovo rozdělení s parametry $x_m > 0$ (měřítko, ozn. **xm**) a α (tvar, ozn. **alpha**) má hustotu danu předpisem:

$$f(x) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}} \quad \text{pro } x \geq x_m \quad (25)$$



Obrázek 21: Hustota Paretova rozdělení pro parametry: — $x_m = 1; \alpha = 1$; - - $x_m = 1; \alpha = 3$; $x_m = 3; \alpha = 1$

4.16 Rozdělení extrémní hodnoty (ExtrValue)

Rozdělení extrémní hodnoty s parametry polohy, měřítka a tvaru $-\infty < \mu < \infty$, $\sigma > 0$ a $-\infty < \xi < \infty$ (ozn. **mu**, **ro** a **xi**) je dáno hustotou pravděpodobnosti ve tvaru:

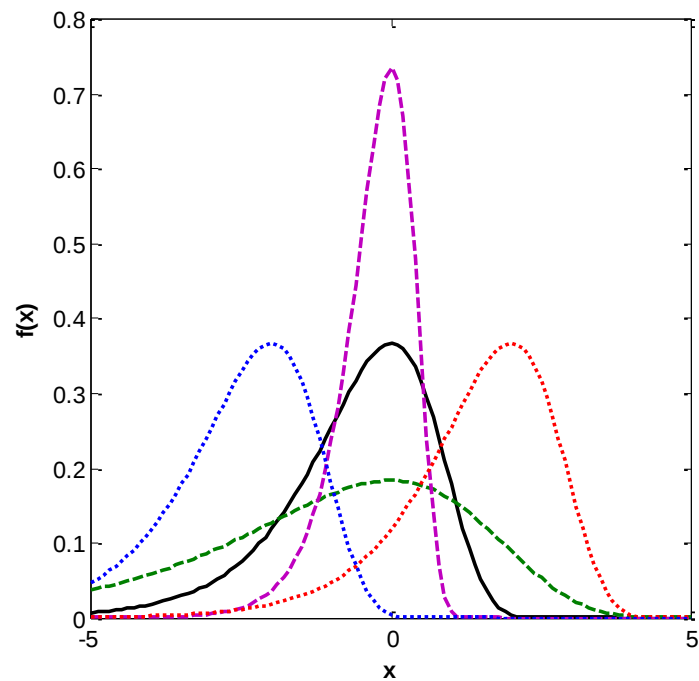
$$f(x) = \frac{1}{\sigma} t(x)^{\xi+1} e^{-t(x)} \quad (26)$$

kde $t(x)$ je dáno restrikcí:

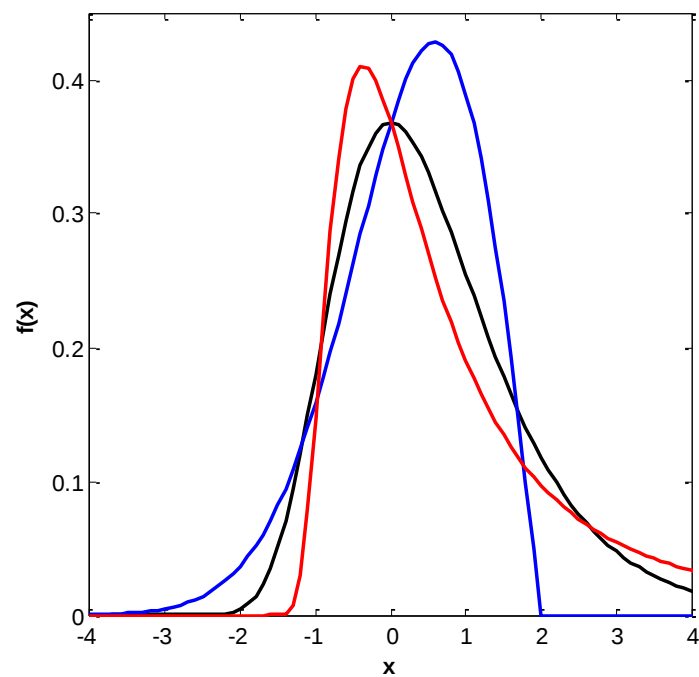
$$t(x) = \begin{cases} \left(1 + \left(\frac{x - \mu}{\sigma}\right)^\xi\right)^{-1/\xi} & \text{pro } \xi \neq 0 \\ e^{-\frac{x - \mu}{\sigma}} & \text{pro } \xi = 0 \end{cases} \quad (27)$$

přitom pro veličinu x jsou přípustné hodnoty:

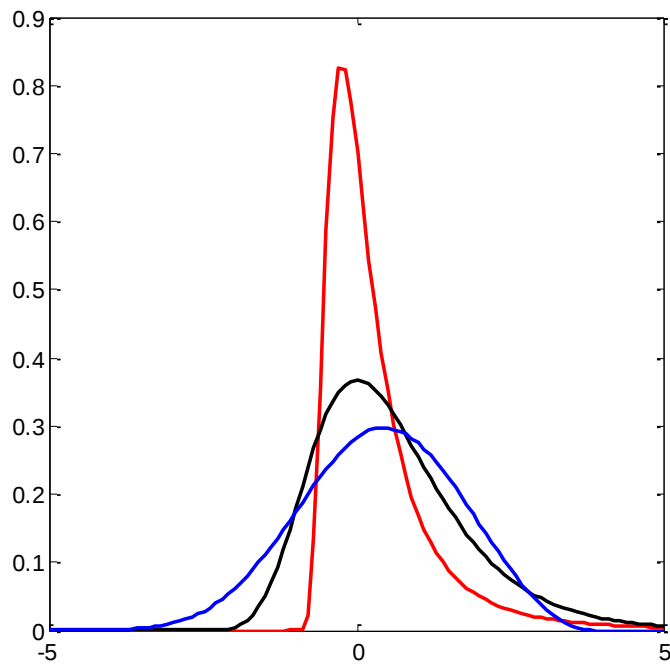
$$\begin{aligned} x &\in \left[\mu - \frac{\sigma}{\xi}; +\infty\right) & \text{pro } \xi > 0 \\ x &\in (-\infty; +\infty) & \text{pro } \xi = 0 \\ x &\in \left(-\infty; \mu + \frac{\sigma}{\xi}\right] & \text{pro } \xi < 0 \end{aligned}$$



Obrázek 22: Hustota pravděpodobnostního rozdělení extrémní hodnoty pro $\xi = 0$, ostatní parametry:
 — $\mu = 0; \sigma = 1$; — $\mu = 0; \sigma = 0,5$; — $\mu = 0; \sigma = 2$; $\mu = -2; \sigma = 1$; $\mu = 2; \sigma = 1$



Obrázek 23: Hustota pravděpodobnostního rozdělení extrémní hodnoty pro $\mu = 0; \sigma = 1$, ostatní parametry: — $\xi = 0$; — $\xi = -0,5$; — $\xi = 0,5$



Obrázek 24: Hustota pravděpodobnostního rozdělení extrémní hodnoty pro — $\xi = 0$; $\mu = 0$; $\sigma = 1$ a dále se zachováním střední hodnoty a rozptylu pro volbu — $\xi = -1/3$; — $\xi = 1/3$

5 Literatura

- [1] Karaba A., Zamostny P., Lederer J., Belohlav Z.: *Ind. Eng. Chem. Res.*, 52, 15407 (2013).
- [2] Matsumoto M., Nishimura T.: *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 8, 3 (1998).
- [3] Matsumoto M., Nishimura T.: v knize: *A nonempirical test on the weight of pseudorandom number generators*. Springer 2002.
- [4] Knuth D.E.: *The art of computer programming, volume 2 (3rd ed.): seminumerical algorithms*. Addison-Wesley Longman Publishing Co., Inc. 1997.