UNIVERZITA PALACKÉHO V OLOMOUCI
PŘÍRODOVĚDECKÁ FAKULTA
KATEDRA MATEMATICKÉ ANALÝZY A APLIKACÍ MATEMATIKY

Bakalářská práce

Typy jednorozměrných rozdělení pravděpodobnosti

Vedoucí bakalářské práce:
RNDr. et PhDr. Ivo Müller Ph.D.

Vypracovala:
Monika Langrová
MATEKO, IV. Ročník

Rok odevzdání: 2016

BIBLIOGRAFICKÁ IDENTIFIKACE

Autor: Monika Langrová

Název práce: Typy jednorozměrných rozdělení pravděpodobnosti

Typ práce: Bakalářská práce

Pracoviště: Katedra matematické analýzy a aplikací matematiky

Vedoucí práce: RNDr. et PhDr. Ivo Müller, Ph.D.

Rok obhajoby: 2016

Abstrakt:

Bakalářská práce se zabývá jednotlivými typy jednorozměrných rozdělení pravděpodobnosti. V první části jsou uvedeny základní pojmy související s tématem. V druhé a třetí části jsou stručně popsána jednotlivá rozdělení, a u vybraných typů ukázka vlivu parametrů na tvar rozdělení. V poslední části jsou porovnána vlastní experimentální data s teoretickým modelem.

Klíčová slova: pravděpodobnost, rozdělení, distribuční funkce, hustota, parametr, střední hodnota, rozptyl

Počet stran: 34

Počet příloh: 0

Jazyk: český

BIBLIOGRAPHICAL IDENTIFICATION

Author: Monika Langrová

Title: Types of one-dimensional probability distributions

Type of thesis: Bachelor‘s

Department: Department of Mathematical Analysis and Applications of Mathematics

Supervisor: RNDr. et PhDr. Ivo Müller, Ph.D.

The year of presentation: 2016

Abstract:

The bachelor thesis deals with individual types of one-dimensional probability distributions. The basic concepts related to the topic are stated in the first part. In the second and third part, the individual distributions and, concerning the selected types, the illustration of the parameters' influence of the form of distribution are briefly described. In the last section, own experimental data are compared with the theoretical model.

Key words: Probability, distribution, distributional function, density, parameter, expected value, variance

Number of pages: 34

Number of appendices: 0

Language: Czech

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracovala samostatně pod odborným vedením RNDr. et PhDr. Ivo Müllera, Ph.D., a že jsem v seznamu použité literatury uvedla všechny použité zdroje.

V Olomouci:

Poděkování

Ráda bych touto cestou poděkovala vedoucímu mé bakalářské práce RNDr. et PhDr. Ivo Müllerovi, Ph.D., za odborné vedení, poskytování rad a za čas, který mi věnoval při vypracování mé bakalářské práce.

Obsah

Úvod

Bakalářská práce se věnuje typům jednorozměrných rozdělení pravděpodobnosti. Rozdělení pravděpodobnosti náhodné veličiny získáme, jestliže každé hodnotě diskrétní náhodné veličiny, popř.: intervalu hodnot spojité náhodné veličiny, přiřadíme pravděpodobnost.

V první kapitole jsou uvedeny základní pojmy tykající se rozdělení pravděpodobnosti, seznámíme se zde se spojitou a diskrétní náhodnou veličinou, od kterých se dále odvíjí jednotlivá rozdělení.

Druhá kapitola se věnuje typům rozdělení pravděpodobnosti diskrétní náhodné veličiny. Stručně jsou zde popsány jednotlivé typy, uvedený příklad využití a u známějších typů je zde poukázáno na vliv parametrů na tvar rozdělení.

Třetí kapitola se věnuje typům rozdělení pravděpodobnosti spojité náhodné veličiny. Podobně jako ve druhé kapitole jsou stručně popsány jednotlivé typy, uvedené příklady využití a u známějších typů je zde poukázáno na vliv parametrů na tvar hustoty rozdělení.

V poslední kapitole jsou uvedeny vlastní data a výpočty pomocí nejvhodnějšího teoretického modelu.

Cílem bakalářské práce je seznámit se s jednotlivými typy jednorozměrných rozdělení pravděpodobnosti, u nejznámějších uvést číselné charakteristiky a poukázat na vliv parametrů na tvar rozdělení. V praktické části potom porovnat vlastní experimentální data s teoretickým modelem. Důraz byl kladen na zpracování, aby bakalářská práce byla srozumitelná a bylo by možné ji využít pro studenty jako učební text k porozumění daného tématu.

1. Základní pojmy

Tato kapitola je zpracována pomocí literatury [1],[2],[3],[4]

Označme si symbolem ω *elementární jev* jakož to výsledek náhodného pokusu. Množinu všech elementárních jevů označíme symbolem Ω jakož to *prostor elementárních jevů*. Nechť je na prostoru Ω dána nějaká σ-algebra A jeho podmnožin. Tyto podmnožiny nazveme *náhodnými jevy*. Trojici (Ω, A, P) nazveme *pravděpodobnostním prostorem*.

Definice 1.1 Nechť je dán pravděpodobnostní prostor (Ω, A, P). Reálnou funkci $X\colon \Omega \to R^1$ nazveme náhodnou veličinou, jestliže pro každé $x \in R^1$ platí

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathrm{A} \tag{1.1}$$

Množinu M $\subset R^1$ všech hodnot náhodné veličiny X nazýváme *obor hodnot náhodné veličiny X*.

Náhodná veličina je taková veličina, která mění své hodnoty v závislosti na náhodě. Rozdělujeme ji na 2 druhy- *diskrétní a spojitá*.

Diskrétní náhodná veličina může nabývat konečně nebo spočetně hodnot (hodnoty, které lze očíslovat). Často se dá vyjádřit jako počet, např.: počet ok při hodu kostkou, počet autonehod během týdne na území České Republiky.

Spojitá náhodná veličina může nabývat všech hodnot z určitého intervalu, např.: chyba v měření vzdálenosti 2 bodů.

1.1 Diskrétní náhodná veličina

Rozdělení pravděpodobnosti náhodné veličiny X nám určuje pravděpodobnostní chování náhodné veličiny. V případě diskrétní náhodné veličiny je rozdělení dáno

a.) výčtem hodnot N, kterých může veličina nabývat,

b.) pravděpodobnostmi, s jakými jednotlivé hodnoty x ∈ N nabývá, tj. P (X=x), x ∈ N.

Definice 1.2 Nechť X je náhodná veličina definována na pravděpodobnostním prostoru (Ω, A, P). Reálná funkce F_X definovaná na R^1 předpisem

$$F_X(x) = P(X \leq x), \quad x \in R_1, \tag{1.2}$$

se nazývá *distribuční funkce náhodné veličiny X.*

1.2 Charakteristiky diskrétní náhodné veličiny

Střední hodnota- charakterizuje nám polohu hodnot náhodné veličiny X, obvykle se značí E (X), v případě diskrétní náhodné veličiny je definována vztahem

$$E(X) = \sum_{x \in N} x\, P(X = x), \tag{1.3}$$

kde N je množina hodnot, jichž náhodná veličina X nabývá.

Rozptyl- jedná se o tzv.: druhý centrální moment náhodné veličiny a charakterizuje nám variabilitu pravděpodobnosti náhodné veličiny X. Značíme Var (X) a v případě diskrétní náhodné veličiny je definován vztahem

$$Var(X) = \sum_{x \in N} (x - E(X))^2 \cdot P(X = x) = \left(\sum_{x \in N} x^2 \cdot P(X = x)\right) - (EX)^2 \tag{1.4}$$

Směrodatná odchylka- značíme $\sigma(X)$ a je definována vztahem

$$\sigma(X) = \sqrt{Var(X)} \tag{1.5}$$

1.3 Spojitá náhodná veličina

U spojité náhodné veličiny obvykle pozorujeme, s jakou pravděpodobností se náhodná veličina realizuje uvnitř nějakého konečného nebo nekonečného intervalu (a, b).

Definice 1.3 Řekneme, že náhodná veličina X má rozdělení spojitého typu, existuje- li nezáporná reálná funkce *f (x) taková*, že pro libovolný interval (a, b) platí

$$P\big(X \in (a, b)\big) = \int_a^b f(x)dx \tag{1.6}$$

Funkci *f (x) se* říká *hustota pravděpodobnosti* a musí pro ni platit

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{1.7}$$

1.4 Charakteristiky spojité náhodné veličiny

Střední hodnota- Taktéž nám charakterizuje polohu hodnot náhodné veličiny, v případě spojité náhodné veličiny je definována vztahem

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)dx \tag{1.8}$$

Rozptyl- Taktéž nám určuje variabilitu náhodné veličiny, v případě spojité náhodné veličiny je definován vztahem

$$Var(X) = \int_{-\infty}^{\infty} \left(x - E(X)\right)^2 f(x)dx = (\int_{-\infty}^{\infty} x^2\, f(x)dx) - (EX)^2 \tag{1.9}$$

Směrodatná odchylka- Značíme $\sigma(X)$ a je definována vztahem

$$\sigma(X) = \sqrt{Var(X)} \tag{1.10}$$

2. Typy diskrétně rozdělených náhodných veličin

Tato kapitola je zpracována pomocí literatury [1],[2],[3],[4]

2.1 Binomické rozdělení

Náhodná veličina X má binomické rozdělení s parametry n a p, jestliže nabývá hodnot k= 0, 1, …, n s pravděpodobnostmi

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \tag{2.1}$$

Parametr *n* je přirozené číslo a $p \in (0, 1)$.

Říkáme, že *X* má *binomické rozdělení* a značí se $X \sim Bi(n, p)$. Lze interpretovat i stylem, že pokus má pouze dva výsledky- úspěch a neúspěch, přičemž pravděpodobnost úspěchu je rovna *p* a pravděpodobnost neúspěchu je *1- p*.

Náhodná veličina X s touto pravděpodobnostní funkcí se nazývá binomická náhodná veličina (má tzv. binomické rozdělení pravděpodobností).

Distribuční funkce binomické náhodné veličiny X má tvar

$$F_X(x) = \begin{cases} 0, & pro\ x < 0, \\ \sum_{k \le x} \binom{n}{k} p^k (1 - p)^{n-k}, & pro\ 0 \le x < n, \\ 1, & pro\ x \ge n. \end{cases} \tag{2.2}$$

V případě kdy n=*1* jedná se o tzv. *Alternativní rozdělení*.

Příklad č. 1: Student složí zkoušku, jestliže v testu odpoví správně alespoň na čtyři z pěti otázek. U každé otázky jsou čtyři možné odpovědi, z nichž jedna je správná. S jakou pravděpodobností student složí zkoušku, jestliže se vůbec nepřipravoval a odpovědi volí náhodně?

Řešení: Jelikož student volí zcela náhodně, p=1/4. Počet pokusů je n=5. Náhodná veličina X označující počet správných odpovědí má tedy binomické rozdělení Bi (n=5,p=1/4). Student složí zkoušku, jestliže správně zodpoví alespoň 4 otázky, tj. 4 nebo 5. Platí tedy

$$P(X \geq 4) = P(X = 4) + P(X = 5) = \binom{5}{4}\left(\frac{1}{4}\right)^4 \frac{3}{4} + \binom{5}{5}\left(\frac{1}{4}\right)^5 = 0{,}0156.$$

Bez přípravy složí student zkoušku s pravděpodobností 0,0156.

2.1.1 Charakteristiky binomického rozdělení

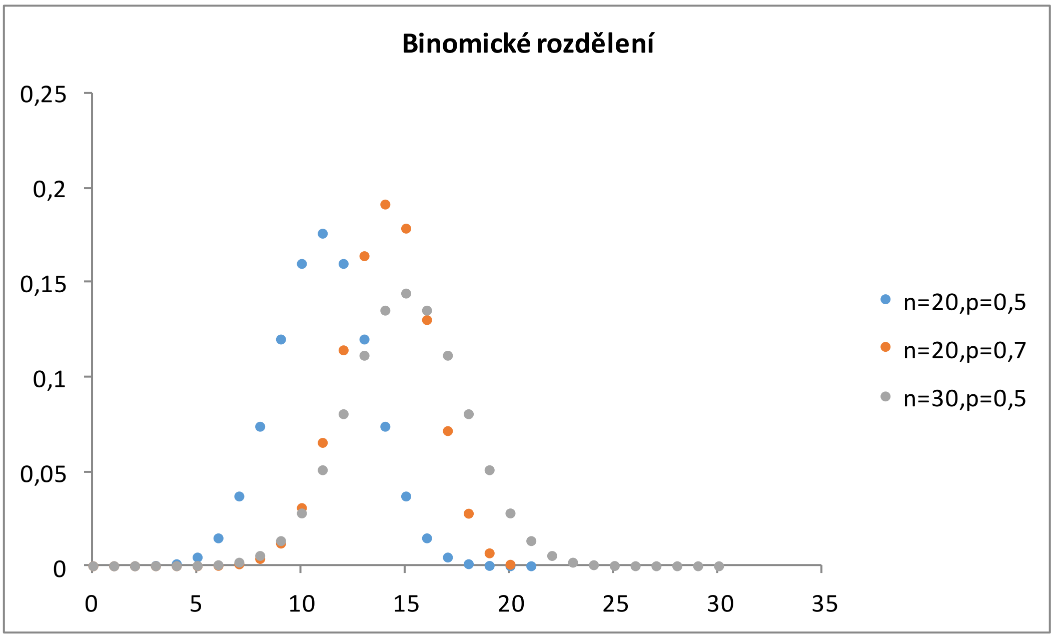
Střední hodnota:

$$E(X) = n \cdot p \tag{2.3}$$

Rozptyl:

$$Var(X) = n \cdot p \cdot (1 - p) \tag{2.4}$$

2.1.2 Vliv parametrů na tvar rozdělení



Obrázek č. 1

Na Obrázku č. 1 vidíme grafy *binomického rozdělení* závislých na parametrech *n* a *p*. Tvar binomického rozdělení se blíží tvaru poissonova rozdělení s parametrem λ, jestliže *n* je velké a *p* se blíží k nule.

2.2 Poissonovo rozdělení

Náhodná veličina X má *Poissonovo rozdělení* s parametrem λ >0, která nabývá hodnot k=0,1,…s pravděpodobnostmi

$$P(x = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \lambda > 0. \tag{2.5}$$

Označujeme X~Po (λ). Parametr λ je číslo a může být nazývaný intenzitou.

Distribuční funkce Poissonova rozdělení je rovna

$$F(x) = \begin{cases} 0, & pro\ x < 0, \\ \sum_{k \le x} \frac{\lambda^k}{k!} e^{-\lambda}, & pro\ x \ge 0. \end{cases} \tag{2.6}$$

Příklad č. 2: K holiči chodí „v průměru“ čtyři zákazníci za hodinu. S jakou pravděpodobností přijde během půl hodiny alespoň jeden zákazník?

Řešení: Za předpokladu, že chodí zákazníci k holiči zcela náhodně, řídí se náhodná veličina X označující počet zákazníku Poissonovým rozdělením. Za půl hodiny k holiči „průměrně“ přijdou dva zákazníci a tedy λ = 2. K holiči přijde alespoň jeden zákazník, přijde-li jeden nebo více a tedy

$$P(X \ge 1) = 1 - P(X = 0) = 1 - e^{-2} \doteq 1 - 0{,}135 = 0{,}865.$$

Během půlhodiny přijde k holiči alespoň jeden zákazník s pravděpodobností 0,865.

2.2.1 Charakteristiky Poissonova rozdělení

Střední hodnota:

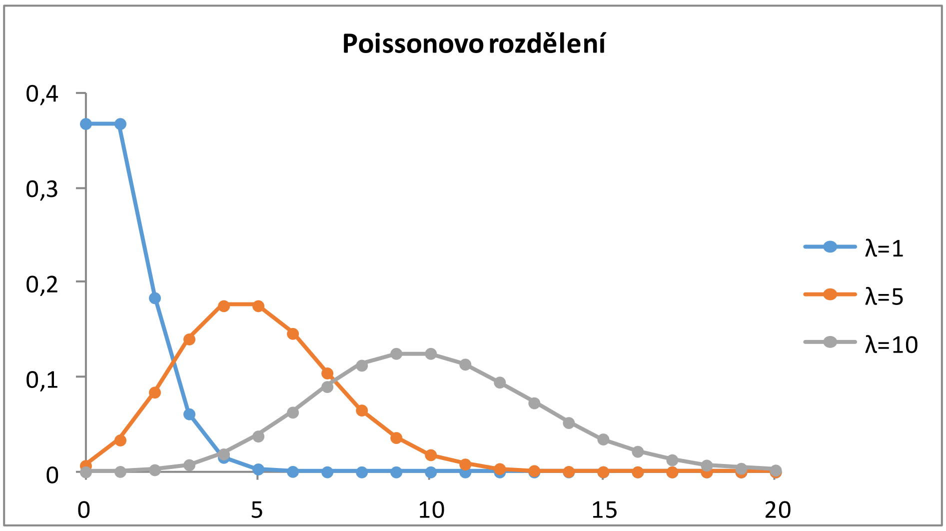
$$E(X) = \lambda \tag{2.7}$$

Rozptyl:

$$var(X) = \lambda \tag{2.8}$$

Střední hodnota a *Rozptyl* Poissonova rozdělení jsou stejné a rovnají se parametru λ.

2.2.2 Vliv parametru na tvar rozdělení



Obrázek č. 2

Na obrázku č. 2 vidíme 3 grafy *Poissonova rozdělení* v závislosti na parametru λ.

2.3 Hypergeometrické rozdělení

Náhodná veličina X má *hypergeometrické rozdělení* s parametry *N, A, n,* pro max (0, $A + n - N$) $\leq$ k $\leq$ min (A, n), jestliže

$$P(x = k) = \frac{\binom{A}{k}\binom{N-A}{n-k}}{\binom{N}{n}} \tag{2.9}$$

Parametry N, A a n jsou přirozená čísla, přičemž platí $A < N$, $n < N$.

Předpokládejme, že máme soubor N jednotek, kde A jednotek má sledovanou vlastnost. Z tohoto souboru vybereme zcela náhodně najednou nebo postupně bez vracení n jednotek. Náhodná veličina X označující počet vybraných jednotek vykazující sledovanou vlastnost se řídí hypergeometrickým rozdělením.

Příklad č. 3: Je známo, že mezi 10 součástkami jsou 2 vadné. Sestavíme-li přístroj z 2 náhodně vybraných součástek. S jakou pravděpodobností bude fungovat?

Řešení: Náhodná veličina X označující počet vadných součástek v přístroji se řídí hypergeometrickým rozdělením, přičemž N=10 (počet všech součástek), A=2 (počet vadných součástek) a n=2 (počet součástek použitých při výrobě přístroje). Přístroj bude fungovat právě tehdy, kdy nebude obsahovat žádnou vadnou součástku.

$$P(X=0)=\frac{\binom{2}{0}\binom{8}{2}}{\binom{10}{2}}=0{,}622$$

Přístroj bude fungovat s pravděpodobností 0,622.

2.3.1 Charakteristiky hypergeometrického rozdělení

Střední hodnota

$$E(X)=\frac{n\cdot A}{N} \tag{2.10}$$

Rozptyl

$$Var\,(X)=n\frac{A}{N}\left(1-\frac{A}{N}\right)\left(\frac{N-n}{N-1}\right). \tag{2.11}$$

2.4 Další diskrétní rozdělení

2.4.3 Alternativní rozdělení

Někdy taky nula-jedničkové rozdělení. Jedná se o speciální případ binomického rozdělení, kdy n=1. S pravděpodobností p nabývá hodnoty 1 a s pravděpodobností (*1*- p) nabývá hodnoty 0. Značíme Alt (p) a je dáno:

x	0	1
P (X=x)	1-p	p

Distribuční funkce je rovna

$$F(x)=\begin{cases} 0, & pro\ x<0,\\ 1-p, & pro\ 0\le x<1,\\ 1, & pro\ x\ge 1.\end{cases} \tag{2.12}$$

Střední hodnota

$$E(X)=p \tag{2.13}$$

*Rozpty*l

$$Var(X) = p \cdot (1-p) \tag{2.14}$$

2.4.2 Logaritmické rozdělení

Náhodná veličina X má *logaritmické rozdělení* s parametrem *α*, kde k >1, jestliže

$$P(X=k) = \frac{1}{-\ln(1-\alpha)} \frac{\alpha^k}{k} \tag{2.15}$$

Střední hodnota

$$E(X) = \frac{\alpha}{-(1-\alpha)\ln(1-\alpha)} \tag{2.16}$$

Rozptyl

$$Var(X) = \frac{\alpha}{-(1-\alpha)^2 \ln(1-\alpha)} \left[1 + \frac{\alpha}{\ln(1-\alpha)}\right] \tag{2.17}$$

2.4.3 Geometrické rozdělení

Náhodná veličina X má geometrické rozdělení s parametrem p, kde *k*= 0, 1, 2,…, a p ∈ (0,1), jestliže

$$P(X=k) = (1-p)^k p \tag{2.18}$$

Označujeme krátce $X \sim Ge(p)$.

Distribuční funkce je rovna

$$F(x) = \begin{cases} 0, & pro\ x < 0, \\ \sum_{k \le x} p(1-p)^k, & pro\ x \ge 0. \end{cases} \tag{2.19}$$

Střední hodnota

$$E(X) = \frac{1-p}{p} \tag{2.20}$$

Rozptyl

$$Var(X) = \frac{1-p}{p^2} \tag{2.21}$$

Geometrické rozdělení je zvláštním případem negativně binomického rozdělení.

2.4.4 Negativní binomické rozdělení

Nechť r > 0 a p ∊ (0,1). Nechť *X* nabývá pouze hodnot 0,1,…, a to s pravděpodobnostmi

$$P(X = k) = \binom{r+k-1}{k} p^r (1-p)^k, \quad \text{kde } k>0 \tag{2.22}$$

Pak říkáme, že *X* má *negativně binomické rozdělení* a píšeme $X \sim NBi(r, p)$.

Střední hodnota

$$E(X) = \frac{n(1-p)}{p} \tag{2.23}$$

Rozptyl

$$Var(X) = \frac{n(1-p)}{p^2} \tag{2.24}$$

3. Typy spojitě rozdělených náhodných veličin

Tato kapitola je zpracována pomocí literatury [1],[2],[3],[4].

3.1 Rovnoměrné rozdělení

Náhodná veličina X má *rovnoměrné rozdělení* na intervalu (a, b), jestliže pro hustotu $f(x)$ platí

$$f(x) = \begin{cases} \frac{1}{b-a}, & pro\ x \in (a,b) \\ 0, & pro\ x \notin (a,b). \end{cases} \tag{3.1}$$

Označujeme $X \sim Ro(a,b)$, kde a a b jsou parametry tohoto rozdělení. Pro distribuční funkci tohoto rozdělení platí

$$F(x) = \begin{cases} 0 & pro\ x \leq a \\ \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a} & pro\ a \leq x < b \\ 1 & pro\ x \geq b. \end{cases} \tag{3.2}$$

Příklad č. 4: Náhodná veličina X má Ro (0,2). Napište její hustotu a distribuční funkci a určete P[0< X< 0,5].

<u>Řešení:</u> Pro hustotu platí: $f(x) = \frac{1}{2-0} = \frac{1}{2}$ pro x ϵ (0,2).

Pro distribuční funkci platí: $\frac{x}{2}$

$P[0< X< 0{,}5]= \frac{0{,}5}{2} - \frac{0}{2} = \frac{1}{4}$

3.1.1 Číselné charakteristiky

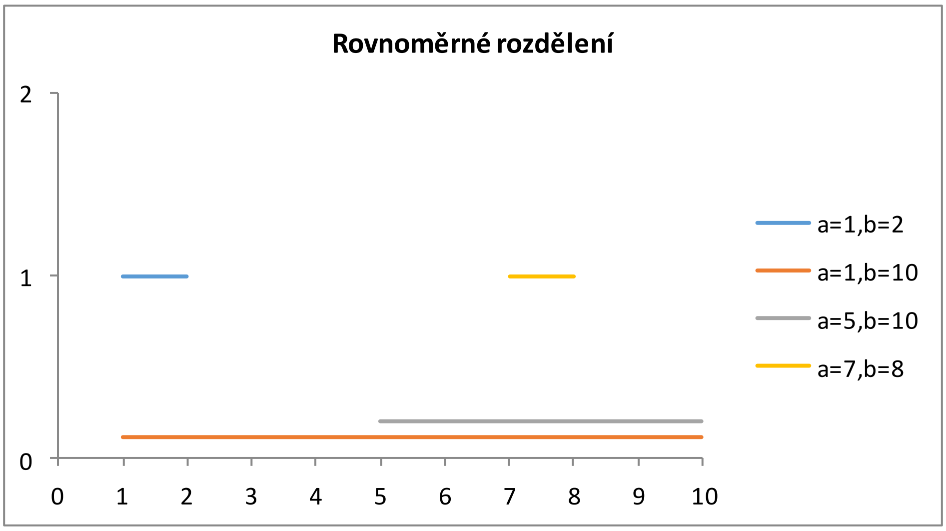
Střední hodnota

$$E(X) = \int_a^b x \frac{1}{b-a} dx = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2} \tag{3.3}$$

Rozptyl

$$Var(X) = \frac{(b-a)^2}{12} \tag{3.4}$$

3.1.2 Vliv parametrů na tvar rozdělení



Obrázek č. 3

Na obrázku č. 3 vidíme grafy hustot *rovnoměrného rozdělení* v závislost na parametrech a a b.

3.2 Normální rozdělení

Náhodná veličina X má *normální rozdělení*, jestliže $x \in R^1$ a pro hustotu ϕ(x) platí

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3.5}$$

Normální rozdělení označujeme $X \sim N(\mu, \sigma^2)$, kde μ, σ jsou parametry tohoto rozdělení a hustota je symetrická kolem parametru μ. Často se normální rozdělení nazývá taktéž Gaussovo rozdělení.

Nejznámější rozdělení mezi normálními rozdělení je tzv. Normální normované rozdělení s parametry (0, 1). Pro jeho hustotu a distribuční funkci platí

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \text{pro } \ x \in R^1 \tag{3.6}$$

$$\phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt\,, \quad \text{pro } \ x \in R^1. \tag{3.7}$$

Hodnoty distribuční funkce standardního normálního rozdělení najdeme ve statistických tabulkách.

Příklad č. 5: Měření vzdálenosti od objektu je spojeno se systematickými i náhodnými chybami. Systematická chyba zkracuje vzdálenost a je rovna 0,05 m. Náhodné chyby mají normální rozdělení se směrodatnou odchylkou σ= 0,1 m. Určete pravděpodobnost, že změřená vzdálenost nepřekročí vzdálenost skutečnou.

Řešení: Chyba měření X nám zde zastává rozdíl mezi naměřenou hodnotou a skutečností a má normální rozdělení. Vezmeme v úvahu, že přístroj má systematickou chybu, budou se tedy chyby měření pohybovat kolem -0,05m, a platí $X \sim N(\mu = -0{,}05, \sigma^2)$. Naměřená vzdálenost nepřekročí vzdálenost skutečnou, jestliže rozdíl mezi naměřenou hodnotou a skutečností bude záporný, neboli X<0.

$$P(X < 0) = F(0) = \phi\left(\frac{0 - (-0{,}05)}{0{,}1}\right) = \phi(0{,}5) = 0{,}6915.$$

Změřená vzdálenost nepřekročí skutečnou vzdálenost s pravděpodobností 0,6915.

3.2.1 Charakteristiky normálního rozdělení

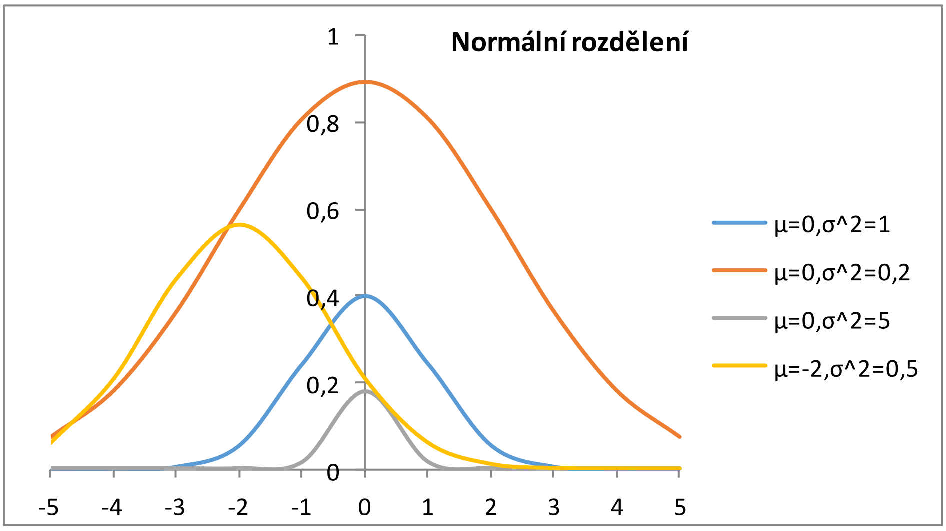
Střední hodnota

$$E(X) = \mu \tag{3.8}$$

Rozptyl

$$Var(X) = \sigma^2 \tag{3.9}$$

3.2.2 Vliv parametrů na tvar rozdělení



Obrázek č. 4

Na obrázku č. 4 vidíme grafy hustot *normálního rozdělení* s různými parametry μ a σ^2. Za nejznámější se považuje rozdělení N (0,1).

3.3 Exponenciální rozdělení

Náhodná veličina *X* má *exponenciální rozdělení*, jestliže pro hustotu platí

$$f(x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}}, & pro\ x\ > 0, \\ 0 & pro\ x\ \leq 0. \end{cases} \tag{3.10}$$

Označujeme $X \sim E(\lambda)$, kde $\lambda > 0$ je parametr exponenciálního rozdělení. Nejznámější využití exponenciálního rozdělení nalezneme v teorii spolehlivosti nebo v teorii obsluhy.

Příklad č. 6: Doba čekání hosta na pivo je v restauraci průměrně 5 minut. Určete pravděpodobnost, že budeme čekat na pivo déle než 12 minut. Pro distribuční funkci platí:

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{\frac{-1}{5}x} & x \geq 0 \end{cases}$$

<u>Řešení:</u> Hledaná pravděpodobnost je zde rovna:

$$P(X > 12) = P(12 < X < \infty) = F(\infty) - F(12) = 1 - \left(1 - e^{-\frac{1}{5}12}\right) = e^{-\frac{12}{5}} \cong 0{,}0907$$

3.3.1. Charakteristiky exponenciálního rozdělení

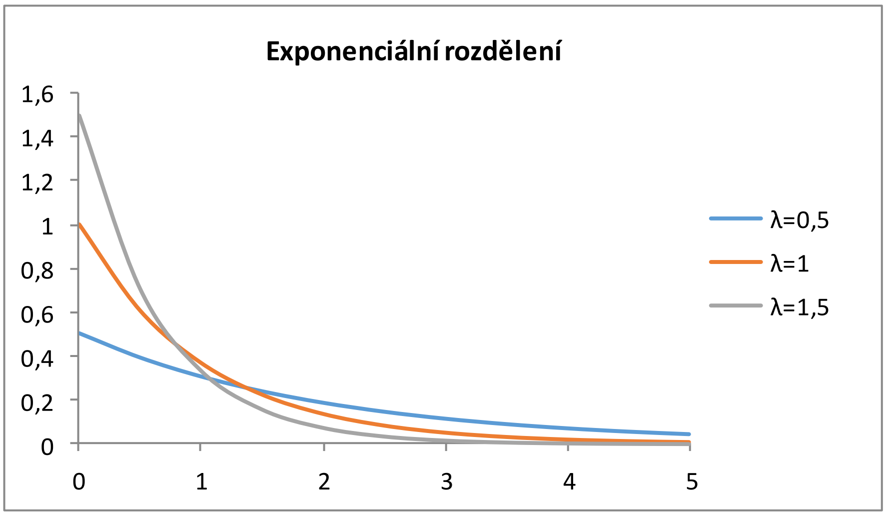
Střední hodnota

$$E(X) = \lambda \tag{3.11}$$

Rozptyl

$$Var(X) = \lambda^2 \tag{3.12}$$

3.3.2 Vliv parametru na tvar rozdělení



Obrázek č. 5

Na obrázku č. 5 vidíme grafy hustot *exponenciálního rozdělení* s parametrem λ.

3.4 Další spojitá rozdělení

3.4.1 Cauchyovo rozdělení

Náhodná veličina X má *Cauchyovo rozdělení*, jestliže pro hustotu platí

$$f(x) = \frac{1}{\pi}\frac{b}{b^2+(x-a)^2} \tag{3.13}$$

Označujeme $X{\sim}C(a, b)$, kde $a \in$ R a $b > 0$ jsou čísla a parametry rozdělení. Střední hodnota tohoto rozdělení neexistuje a nejčastěji se používá rozdělení C (0, 1).

3.4.2 Gama rozdělení

Náhodná veličina X má *gama rozdělení*, jestliže pro hustotu platí

$$f(x) = \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1}, \quad x > 0 \tag{3.14}$$

Gama rozdělení označujeme $X \sim Ga\,(a, p)$, kde $a > 0,\ p > 0$ jsou parametry rozdělení. Speciálním případem gama rozdělení je *exponenciální rozdělení*, které získáme při p = 1.

3.4.2 Beta rozdělení

Náhodná veličina X má *beta rozdělení*, jestliže pro hustotu platí

$$f(x) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}, \quad 0 < x < 1 \tag{3.15}$$

Beta rozdělní označujeme $X \sim B(a, b)$, kde $a > 0,\ b > 0$ jsou parametry tohoto rozdělení. Speciálním případem beta rozdělení je *rovnoměrné rozdělní*, které získáme při $a = b = 1$.

3.4.3 Rozdělení χ^2

Náhodná veličina X má *rozdělení* χ^2 *(chí kvadrát) o* n stupních volnosti, nechť $n \geq 1$ a pro hustotu rozdělení platí

$$f(x) = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{\frac{-x}{2}}, \quad x > 0 \tag{3.16}$$

Označujeme $X \sim \chi_n^2$. Parametr tohoto rozdělení je zde číslo n (stupeň volnosti), které podle podmínky nabývá pouze kladných hodnot. Jedná se o speciální případ gama rozdělení při $p = \frac{n}{2}$, $a = \frac{1}{2}$. Hustota rozdělení χ^2je kladná pouze pro kladné hodnoty argumentu. Využití rozdělení hlavně ve statistických aplikacích, kvantily tohoto rozdělení jsou uvedeny ve statistických tabulkách.

Střední hodnota tohoto rozdělení je rovna počtu stupňů volnosti a tedy platí

$$E(X) = n \tag{3.17}$$

Rozptyl je roven dvojnásobku počtu stupňů volnosti a tedy platí

$$Var(X) = 2n \tag{3.18}$$

3.4.4 Rozdělení *t*

Náhodná veličina X má *Rozdělení t (Studentovo rozdělení)* o n stupních volnosti, nechť $n \geq 0$ a pro hustotu rozdělení platí

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}}\left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} \tag{3.19}$$

Označujeme $X \sim t_n$, kde n *(stupeň volnosti)* je parametrem tohoto rozdělení a nabývá pouze hodnoty z množiny přirozených čísel. Hustota rozdělení je symetrická kolem nuly. Je-li $n > 1$, existuje střední hodnota a platí, že $E(X) = 0$. Pokud $n > 2$, existuje konečný rozptyl $Var(X) = \frac{n}{n-2}$. Pro $n = 1$ dostáváme *Cauchyovo rozdělení* C(0,1). Kvantily studentova rozdělení nalezneme taktéž ve statistických tabulkách.

4. Praktická část

V praktické části na vlastní sesbíraná data aplikuji nejvhodnější teoretický model a vypočítám charakteristiky tohoto rozdělení.

4.1 Praktický příklad č. 1

V tabulce č. 1 máme uvedené časové údaje jednotlivých příjezdů hostů do hotelu.

	Čas příjezdu		***Čas příjezdu***
Host č. 1	15:47	Host č. 12	16:50
Host č. 2	15:51	Host č. 13	15:54
Host č. 3	17:42	Host č. 14	13:27
Host č. 4	19:41	Host č. 15	15:52
Host č. 5	19:37	Host č. 16	19:36
Host č. 6	17:48	Host č. 17	18:39
Host č. 7	20:47	Host č. 18	18:36
Host č. 8	17:40	Host č. 29	15:53
Host č. 9	13:38	Host č. 20	19:50
Host č. 10	14:11	Host č. 21	17:50
Host č. 11	17:46	Host č. 22	20:55

Tabulka č. 1

Označme si X_i, i=1,2,…22 jako doby mezi příchody jednotlivých hostů. Příchod hosta je zcela náhodný. Řídíme se zde časovým intervalem, který začíná v 13:00 a končí ve 23:00. Během tohoto časového intervalu (10 hodin) se přijde ubytovat celkem 22 hostů, za hodinu se v průměru přijdou ubytovat 2 hosté. Po seřazení jednotlivých hodnot dostáváme hodnoty uvedené v tabulce č. 2.

X_i	Čas.int v min.	X_i	Čas.int v min.
$\mathbf{X_1}$	27	$\mathbf{X_{12}}$	4
$\mathbf{X_2}$	11	$\mathbf{X_{13}}$	2
$\mathbf{X_3}$	33	$\mathbf{X_{14}}$	2
$\mathbf{X_4}$	96	$\mathbf{X_{15}}$	46
$\mathbf{X_5}$	4	$\mathbf{X_{16}}$	3
$\mathbf{X_6}$	1	$\mathbf{X_{17}}$	57
$\mathbf{X_7}$	1	$\mathbf{X_{18}}$	1
$\mathbf{X_8}$	1	$\mathbf{X_{19}}$	4
$\mathbf{X_9}$	56	$\mathbf{X_{20}}$	9
$\mathbf{X_{10}}$	50	$\mathbf{X_{21}}$	57
$\mathbf{X_{11}}$	2	$\mathbf{X_{22}}$	8

Tabulka č. 2

Pro další výpočty využijeme rovnice pro výběrový průměr a výběrový rozptyl a následně porovnáme se střední hodnotou a rozptylem *exponenciálního rozdělení* s parametrem λ. (předpoklad).

Výběrový průměr:

$$\bar{X} = \frac{1}{n}\sum_{1}^{n} X_i = \frac{1}{22}\sum_{1}^{22} X_i = \frac{1}{22} \cdot 475 = 21{,}59091$$

Výběrový rozptyl:

$$S^2 = \frac{1}{n-1}\sum_{1}^{n} (X_i - \bar{X})^2 = \frac{1}{21} \cdot 15367{,}32 = 731{,}7771$$

Porovnání střední hodnoty a výběrového průměru:

$$E(X) = \bar{X}$$

$$\hat{\lambda} \doteq 21{,}6$$

Porovnání rozptylu a výběrového rozptylu:

$$Var(X) = S^2$$

$$\widehat{\lambda^2} = 731{,}7771$$

$$\hat{\lambda} \doteq 27{,}05$$

Z následujících výpočtů je zřejmé, že odhady parametru $\hat{\lambda}$ se nerovnají. V tomto případě data nejsou příliš vhodná pro exponenciální rozdělení.

4.2 Praktický příklad č. 2

V tabulce č. 3 máme uvedené jednotlivé časové údaje odjezdů hostů z hotelu.

	Čas odjezdu		*Čas odjezdu*
Host č. 1	*11:01*	Host č. 12	*7:53*
Host č. 2	*8:30*	Host č. 13	*8:43*
Host č. 3	*9:16*	Host č. 14	*8:45*
Host č. 4	*8:04*	Host č. 15	*11:09*
Host č. 5	*8:39*	Host č. 16	*9:13*
Host č. 6	*10:12*	Host č. 17	*9:07*
Host č. 7	*8:07*	Host č. 18	*8:37*
Host č. 8	*9:04*	Host č. 19	*8:35*
Host č. 9	*8:39*	Host č. 20	*8:20*
Host č. 10	*8:22*	Host č. 21	*7:01*
Host č. 11	*9:20*	Host č. 22	*9:16*

Tabulka č. 3

Označme si Y_i, i=1,2,...22 jako doby mezi odjezdy jednotlivých hostů. Odjezd hosta z hotelu je zde taktéž jako v předchozím příkladu zcela náhodný. Řídíme se zde časovým intervalem v rozmezí od 7:00 do 12:00, za hodinu se z hotelu odhlásí v průměru 4 hosté. V tabulce č. 4 máme seřazeny časové hodnoty mezi jednotlivými odjezdy hostu.

Y_i	Čas.int v min.	Y_i	Čas.int v min.
Y_1	1	Y_{12}	4
Y_2	52	Y_{13}	2
Y_3	11	Y_{14}	19
Y_4	3	Y_{15}	3
Y_5	13	Y_{16}	6
Y_6	2	Y_{17}	3
Y_7	8	Y_{18}	0
Y_8	5	Y_{19}	4
Y_9	2	Y_{20}	52
Y_{10}	2	Y_{21}	49
Y_{11}	0	Y_{22}	8

Tabulka č. 4

Stejně jako u prvního případu použijeme pro další výpočty stejné rovnice a následně porovnáme. Předpoklad zde zůstává stejný, a tedy jedná se o *exponenciální rozdělení* s parametrem λ.

Výběrový průměr:

$$\bar{X} = \frac{1}{n}\sum_{1}^{n} X_i = \frac{1}{22}\sum_{1}^{22} X_i = \frac{1}{22} \cdot 249 = 11{,}31818$$

Výběrový rozptyl:

$$S^2 = \frac{1}{n-1}\sum_{1}^{n}(X_i - \bar{X})^2 = \frac{1}{21} \cdot 5906{,}773 = 281{,}2749$$

Porovnání střední hodnoty a výběrového průměru:

$$E(X) = \bar{X}$$

$$\hat{\lambda} \doteq 11{,}32$$

Porovnání rozptylu a výběrového rozptylu:

$$Var(X) = S^2$$

$$\widehat{\lambda^2} = 281{,}2749$$

$$\hat{\lambda} \doteq 16{,}77$$

U následujících výpočtů je zřejmé, že rozdíl v odhadech parametrů je menší než v předchozím případu, avšak stále není roven. V tomto případu taktéž zamítám předpoklad o rovnosti exponenciálního rozdělení. Naměřená data nejsou příliš vhodná.

4.3 Praktický příklad č. 3

Uvažujeme situaci, ve které nastane jev A (vlak přijede včas) s pravděpodobností p (0,1). X je zde rovna počtu úspěchů, které nastanou v daném pokuse.

Ze stanice	Odjezd	Do stanice	Příjezd	Skutečný příjezd
Frenštát pod Radhoštěm	8:13	Valašské Meziříčí	8:45	8:49
Valašské Meziříčí	9:24	Frenštát p. R.	9:54	9:54
Frenštát p. R.	10:30	Val. Mez.	11:03	11:03
Val. Mez.	11:24	Frenštát p. R.	11:54	12:11
Frenštát p. R.	12:14	Val. Mez.	12:45	12:53
Val. Mez	13:24	Frenštát p. R.	13:56	13:56
Frenštát p. R.	13:34	Val. Mez.	14:07	14:09
Val. Mez.	14:30	Frenštát p. R.	15:08	15:10
Frenštát p. R.	14:43	Val. Mez.	15:17	15:17
Val. Mez.	15:30	Frenštát p. R.	16:08	16:17
Frenštát p. R.	15:43	Val. Mez.	16:17	16:19
Val. Mez.	17:23	Frenštát p. R.	17:53	17:54
Frenštát p. R.	16:44	Val. Mez.	17:17	17:18

Tabulka č. 5

V případě, kdy pouze sledujeme výskyt jevu A, zda vlak dojede včas či nikoliv, jedná se o Binomické rozdělení pravděpodobnosti. Vzhledem k tomu, že pravděpodobnosti, zda vlak přijede nebo nepřijede včas, jsou ovlivněny spoustou okolností, budeme tedy uvažovat nejjednodušší případ, kdy pravděpodobnosti příjezdu vlaku včas nejsou ničím ovlivněny a tedy p=0,5. Známe oba parametry rozdělení a můžeme tedy vyčíslit charakteristiky rozdělení a pravděpodobnost toho, že jev A nastane v pokusu právě 4x.

Střední hodnota:

$$E(X) = n \cdot p = 13 * 0{,}5 = 6{,}5$$

Rozptyl:

$$Var(X) = np \cdot (1 - p) = 6{,}5 * 0{,}5 = 3{,}25$$

Pravděpodobnost (X=4):

$$P(X = 4) = \binom{n}{x} \cdot p^x \cdot (1 - p)^{n-x} = \binom{13}{4} \cdot 0{,}5^4 \cdot 0{,}5^9 = 0{,}08728$$

4.4 Praktický příklad č. 4

Uvažujeme situaci, kdy máme misku s vruty s křížkovou drážkou. Při jisté dělnické práci doma budeme potřebovat 8 vrutů o velikosti 3×25 mm. Náhodná veličina X, která zde představuje počet vhodných vrutů s hledanou vlastností 3x25 mm, může mít rozdělení Hypergeometrické. Pro další výpočty využijeme hodnoty z tabulky č. 6. Velikosti jsou v tabulce uvedeny v mm.

šroub	velikost	šroub	velikost	šroub	velikost
1.	4×35	16.	4×35	31.	6×20
2.	3×25	17.	6×20	32.	4×50
3.	6×20	18.	3×25	33.	4×50
4.	5×20	19.	4×50	34.	3×25
5.	3×25	20.	4×35	35.	3×25
6.	4×50	21.	4×35	36.	4×50
7.	3×25	22.	4×35	37.	3×25
8.	5×20	23.	3×25	38.	3×50
9.	3×25	24.	4×50	39.	3×50
10.	3×25	25.	6×20	40.	3×25
11.	4×50	26.	6×20	41.	4×50
12.	4×50	27.	3×25	42.	4×50
13.	4×50	28.	4×50	43.	4×50
14.	5×20	29.	4×50		
15.	5×20	30.	3×25		

Tabulka č. 6

Dále si označme N=43, A=13 a n=8. Jelikož známe všechny parametry potřebné k určení charakteristik a pravděpodobností tohoto rozdělení, můžeme vypočítat:

Střední hodnotu

$$E(X) = \frac{8 \cdot 13}{43} \doteq 2{,}42$$

Pravděpodobnost toho, že si vytáhneme právě 8 vhodných vrutů:

$$P(X = 8) = \frac{\binom{13}{8} \cdot \binom{43-13}{8-8}}{\binom{43}{8}} = 0{,}00000021132$$

Pravděpodobnost toho, že si nevytáhneme ani jeden vhodný vrut:

$$P(X = 0) = \frac{\binom{13}{0} \cdot \binom{30}{8}}{\binom{43}{0}} = 0{,}04036263$$

Nyní pozměníme situaci, v případě, kdy budeme potřebovat 6 vrutů o velikosti 4x50 mm, si dále označíme N=43, A=14, a n=6.

<u>Střední hodnota</u>:

$$E(X) = \frac{6 \cdot 14}{43} = 1{,}95$$

<u>Pravděpodobnost</u>, že si vytáhneme právě 6 vhodných vrutů:

$$P(X = 6) = \frac{\binom{14}{6} \cdot \binom{43-14}{6-6}}{\binom{43}{6}} = 0{,}000493$$

<u>Pravděpodobnost</u>, že si nevytáhneme ani jeden vhodný vrut:

$$P(X = 0) = \frac{\binom{14}{0} \cdot \binom{43-14}{6}}{\binom{43}{6}} = 0{,}077917$$

Všechny data lze vypočítat díky znalosti pravděpodobnostní funkce, jednotlivých charakteristik a všech parametrů tohoto rozdělení. Z vypočítaných dat vidíme, že pravděpodobněji si vytáhneme 6 stejných vrutů o velikosti 4x50 mm než 8 stejných vrutů o velikosti 3x25 mm.

Závěr

Cílem bakalářské práce bylo stručně popsat jednotlivé typy jednorozměrných rozdělení pravděpodobnosti a u vybraných rozdělení poukázat na vliv parametrů na tvar rozdělení. V praktické části zpracovat vlastní data a porovnat s teoretickým modelem.

Typy rozdělení lze podle náhodné veličiny rozdělit na typy rozdělení diskrétní náhodné veličiny a typy rozdělení spojité náhodné veličiny.

Nejdůležitější je k popisu jednotlivých typů znát základní pojmy a charakteristiky. Na začátku práce jsem se tedy první věnovala popisu diskrétní a spojité náhodné veličiny, čím se tyto veličiny vyznačují a jak lze vypočítat nejznámější charakteristiky, jako jsou střední hodnota, rozptyl a směrodatná odchylka. Dále jsem se v teoretické části věnovala známým typům rozdělení, u nejznámějších poukázala na vliv parametrů na tvar rozdělení pomocí grafů zpracovaných v MS Excel 2007 a u známých typů uvedla příklad s výpočtem.

V praktické části jsem měla za úkol sesbírat vhodné data a pokusit se aplikovat je do nejvhodnějšího teoretického modelu.

Práce na toto téma pro mě byla velice přínosná a taktéž zajímavá. Je zpracována ve srozumitelném stylu s uvedenými příklady a výpočtem k porozumění využití jednotlivých typů v praxi.

Použitá literatura

[1] ANDĚL, Jiří. *Základy matematické statistiky.* 2. vyd. Praha: Matfyzpress, 2007. 358 s. ISBN 80-7378-001-1

[2] ZVÁRA, Karel a ŠTĚPÁN, Josef. *Pravděpodobnost a matematická statistika.* 5. vyd. Praha: Matfyzpress, 2012. 230 s. ISBN 978-80-7378-218-4

[3] KUNDEROVÁ, Pavla. *Základy pravděpodobnosti a matematické statistiky.* 1.vyd. Olomouc: Univerzita Palackého v Olomouci, 2004. 186 s. ISBN 80-244-0813-9

[4] JARUŠKOVÁ, Daniela. *Pravděpodobnost a matematická statistika.* 3. vyd. Praha: nakladatelství ČVUT, 2011. 138 s. ISBN 978-80-01-04829-0