

2 PRAVDĚPODOBNOST

2.1 Základní pojmy

2.1.1 Náhodný pokus

Při výuce na základní či střední škole jsme se všichni setkali s pokusy, prováděnými v hodinách chemie či fyziky. Tyto pokusy měly za cíl demonstrovat určité procesy, a jejich výsledek byl dopředu znám. Ohlášený výsledek však záležel na dodržení podmínek pokusu a pokud tyto byly dodrženy, očekávaný výsledek se dostavil. Výsledek pokusu byl tedy jednoznačně předurčen podmínkami, za kterých pokus probíhal.

Na druhé straně se setkáváme s velkým množstvím pokusů (obecně činností), jichž výsledek může být různý, i když se snažíme dodržet veškeré podmínky pokusu. Důležitá jsou zřejmě slova snažíme dodržet. Většinou je totiž, na rozdíl od laboratoře, těchto podmínek příliš mnoho a dodržet je bývá velmi obtížné. Znamená to tedy, že v těchto případech podmínky, které předurčují výsledek pokusu, jsou jen obtížně kontrolovatelné.

Pokud např. cestujeme každé ráno do práce stejnou tramvají po stejné trase, ve stejný čas dle jízdního rádu, nikdy nejsme dopředu schopni určit, jak přesně dlouho pojedeme. Budeme-li volat každé ráno v 8:00 hod. na určité telefonní číslo, ze stejného telefonu, u stejného operátora, nikdy dopředu nevíme, zda se dovoláme. Pokud ano, tak zda volaný přijme hovor po prvním, druhém či např. až pátém zazvonění. Stejně tak, začneme-li používat nový počítač, nevíme, jak dlouho vydrží pracovat do první poruchy – zda to bude méně než 1000 hodin, či více než 2000 hodin apod.

Podobných činností by se dala vyjmenovat celá řada. Ve všech těchto případech sice nejsme schopni určit dopředu přesný výsledek, ale na základě zkušeností z minulosti jsme většinou schopni určit, zda určitý výsledek nastává často, velmi často, velmi zřídka či naprostě ojediněle.

Náhodným pokusem budeme rozumět takovou činnost, jejíž výsledek není jednoznačně předurčen podmínkami, za kterých pokus probíhá, a tento pokus je (alespoň teoreticky) mnohonásobně opakovatelný ze stejných podmínek.

Jedním z cílů teorie pravděpodobnosti je nahradit neurčitá tvrzení o výsledcích náhodných pokusů mírou četnosti výskytů těchto výsledků a dále zavést pravidla pro počítání s těmito mírami.

Protože lidstvo od pradávna zajímalo, jaké jsou šance na vítězství v některých poměrně jednoduchých hrách, ať už se jednalo o házení mincí, losování z osudí, hru v kostky, karetní hry apod., jsou počátky počtu pravděpodobnosti spjaty právě s těmito jednoduchými hazardními hrami. Jedná se o hry, jejichž výsledky jsou výsledkem reálnice náhodných pokusů. Navíc se jedná o hry poměrně jednoduché, u kterých jsou dobře známy podmínky, za kterých probíhají, a možných výsledků nebývá mnoho.

2.1.2 Náhodný jev

Vysvětlení, co budeme rozumět termínem náhodný jev, podávají autoři Likeš, Machek (1991) zhruba takto:

Náhodným jevem je jakékoliv tvrzení o výsledku náhodného pokusu, o kterém lze po uskutečnění pokusu jednoznačně rozhodnout, zda je při dané realizaci pokusu pravdivé či nepravdivé. Např. při hodu klasickou kostkou je tvrzení „padne číslo 6“ náhodným jevem. Stejně tak je náhodným jevem tvrzení „ze čtyř po sobě narozených dětí budou právě 2 chlapci“, tvrzení „politická strana xyz získá v parlamentních volbách nejméně 10 % hlasů“ či tvrzení „na území hlavního města Prahy nedojde zátra k více než 20 dopravním nehodám“.

2.1.3 Operace s náhodnými jevy

Jevy budeme značit velkými písmeny ze začátku latinské abecedy, tedy budeme hovořit o jevu A , B , C , ... atd. Protože jevy jsou vlastně tvrzení o výsledku náhodného pokusu, můžeme tato tvrzení různě skládat, vymezovat jejich postavení vůči sobě na vzájem, vytvářet jejich negaci, tvořit z nich výroky apod. Budeme tedy provádět tzv. operace s náhodnými jevy. Tyto operace a zacházení s jevy samotnými, jak dále uvidíme, mají analogii s teorií množin, a to včetně terminologie.

Jev nemožný je jev, který nikdy nemůže nastat (nenastává za žádných okolností). Jev nemožný budeme značit \emptyset .

Jev jistý je jev, který nastává vždy (nastává za všech okolností). Budeme jej značit S (nyní porušíme dohodu, že jevy budeme značit velkými písmeny ze začátku abecedy – nicméně toto označení pro jev jistý se vžilo a v literatuře se běžně používá).

Jevem doplňkovým (opačným, komplementárním) k jevu A je takový jev, který nastává vždy, když nenastává A . Budeme ho značit symbolem \bar{A} .

Příklad 2.1

Uvažujme náhodný pokus, kterým je hod kostkou. Označme jako jev A „padne sudé číslo“, jako jev B „padne číslo dělitelné 6“, jako jev C „padne číslo menší než 5“. Pak

- \bar{A} je jev „padne liché číslo“,
- \bar{B} je jev „nepadne číslo 6“,
- \bar{C} je jev „padne číslo 5 nebo 6“.

Sjednocením jevů A a B rozumíme takový jev $A \cup B$, který nastává právě tehdy, pokud nastane alespoň jeden z jevů A a B .

Sjednocení lze rozšířit na konečný počet n jevů A_1, A_2, \dots, A_n (resp. nekonečný počet jevů A_1, A_2, \dots). Sjednocením jevů $A_1 \cup \dots \cup A_n = \bigcup_{i=1}^n A_i$ (resp. $A_1 \cup A_2 \cup \dots = \bigcup_{i=1}^{\infty} A_i$) rozumíme jev, který nastává právě tehdy, nastane-li alespoň jeden z jevů A_1, A_2, \dots, A_n (resp. A_1, A_2, \dots).

Příklad 2.2

Uvažujme náhodný pokus, kterým je hod kostkou. Označme jako jev A_i „padne číslo i “ pro $i = 1, 2, \dots, 6$. Pak sjednocením

- $A_3 \cup A_6$ je jev „padne číslo 3 nebo číslo 6“,
- $A_1 \cup A_2 \cup A_3$ je jev „padne číslo menší než čtyři“,
- $A_2 \cup A_4 \cup A_6$ je jev „padne sudé číslo“.

Průnikem jevů A a B rozumíme takový jev $A \cap B$, který nastává právě tehdy, nastanou-li oba jevy A a B současně.

Průnik jevů lze stejně jako u sjednocení rozšířit na konečný počet n jevů A_1, A_2, \dots, A_n (resp. na nekonečný počet jevů A_1, A_2, \dots). Průnikem jevů $A_1 \cap \dots \cap A_n = \bigcap_{i=1}^n A_i$ (resp. $A_1 \cap A_2 \cap \dots = \bigcap_{i=1}^{\infty} A_i$) rozumíme jev, který nastává právě tehdy, nastanou-li všechny jevy A_1, A_2, \dots, A_n (resp. A_1, A_2, \dots) současně.

Příklad 2.3

Uvažujme náhodný pokus, kterým je hod kostkou. Označme jako jev A „padne sudé číslo“, jako jev B „padne číslo dělitelné 3“ a jako jev C „padne číslo větší než 2“. Pak

- $A \cap B$ je jev „padne číslo 6“,
- $A \cap C$ je jev „padne číslo 4 nebo 6“,
- $B \cap C$ je jev „padne číslo 3 nebo 6“,
- $A \cap B \cap C$ je jev „padne číslo 6“.

Jevy A a B nazýváme **disjunktní** (neslučitelné), jestliže jejich průnikem je jev nemožný. Tedy pokud platí $A \cap B = \emptyset$. Pojem disjunktních jevů můžeme zobecnit pro libovolný konečný či nekonečný systém jevů. Takový systém se nazývá disjunktní, pokud jsou disjunktní kterékoliv dva různé jevy z tohoto systému.

Příklad 2.4

Uvažujme náhodný pokus, kterým je hod hrací kostkou. Označme jako jev A „padne sudé číslo“, jako jev B „padne liché číslo“. Pak $A \cap B = \emptyset$ a tudíž se jedná o disjunktní jevy.

Rozdílem jevů A a B rozumíme takový jev $A \setminus B$, který nastává právě tehdy, pokud nastane jev A , ale nikoliv jev B .

Rozdíl jevů můžeme vyjádřit i pomocí průniku jevů A a \bar{B} jako $A \setminus B = A \cap \bar{B}$.

Příklad 2.5

Uvažujme náhodný pokus, kterým je hod kostkou. Označme jako jev A „padne sudé číslo“, jako jev B „padne číslo dělitelné 3“, jako jev C „padne číslo menší než 4“. Pak rozdílem jevů

$$\begin{aligned} A \setminus B & \text{ je jev } „\text{padne číslo 2 nebo číslo 4}\", \\ A \setminus C & \text{ je jev } „\text{padne číslo 4 nebo číslo 6}\\", \\ B \setminus C & \text{ je jev } „\text{padne číslo 6}\\", \\ C \setminus B & \text{ je jev } „\text{padne číslo 1 nebo číslo 2}\\". \end{aligned}$$

■ Rekneme, že jev A implikuje jev B (jev A má za následek jev B), jestliže jev B nastane vždy, když nastane jev A .

Fakt, že jev A implikuje jev B , zapisujeme jako $A \subset B$. Pokud platí, že jev A implikuje jev B a zároveň jev B implikuje jev A , pak jsou si jevy A a B rovny.

Příklad 2.6

Uvažujme náhodný pokus, kterým je hod hrací kostkou. Označme jako jev A „padne číslo 2“, jako jev B „padne sudé číslo“. Pak jev A implikuje jev B , neboť pokud padne na kostce číslo 2, platí, že padlo sudé číslo, čili nastal jev B .

2.1.4 Elementární jev a prostor elementárních jevů

Elementárním jevem budeme rozumět „nejjednodušší výsledek“ náhodného pokusu, který už dále nejde rozložit. Jedná se o jev, který nelze vyjádřit jako sjednocení dalších různých jevů.

Jev A nazveme **elementárním jevem**, pokud neexistují jevy B a C různé od A takové, že $A = B \cup C$.

Pro libovolný náhodný pokus bude velmi důležité určit všechny elementární jevy, spjaté s tímto pokusem. Tyto jevy budou pak tvořit prostor elementárních jevů.

Prostorem elementárních jevů S budeme rozumět množinu všech elementárních jevů, které mohou nastat jako výsledek daného náhodného pokusu.

Jakýkoliv další náhodný jev pak bude podmnožinou prostoru elementárních jevů, jakýkoliv elementární jev je prvkem tohoto prostoru. Všem operacím prováděným s náhodnými jevy, jako je sjednocení, průnik a rozdíl, odpovídá sjednocení, průnik a rozdíl příslušných podmnožin prostoru elementárních jevů. Operace s náhodnými jevy pak v podstatě znamená operaci s příslušnými množinami a jak již bylo v předchozím textu uvedeno, řídí se stejnými pravidly.

Poznamenejme ještě, že jevu nemožnému odpovídá prázdná množina a jevu jistému odpovídá celý prostor elementárních jevů.

Příklad 2.7

Uvažujme náhodný pokus, kterým je hod hrací kostkou. Označme jako jev E_i „padne číslo i “ pro $i = 1, 2, \dots, 6$. Pak prostorem elementárních jevů pro náhodný pokus „hod kostkou“ je množina obsahující všechny elementární jevy

$$S = \{E_1, E_2, E_3, E_4, E_5, E_6\}.$$

Jakýkoliv další jev můžeme vyjádřit jako sjednocení těchto elementárních jevů. Položme za jev B „padne sudé číslo“, za jev C „padne číslo dělitelné třemi“. Pak $B = E_2 \cup E_4 \cup E_6$ a $C = E_3 \cup E_6$.

Jevu B odpovídá podmnožina $E_B = \{E_2, E_4, E_6\}$ množiny E , jevu C odpovídá podmnožina $E_C = \{E_3, E_6\}$ množiny S .

Příklad 2.8

Uvažujme náhodný pokus, kterým je hod 2 hracími kostkami. Označme jako jev E_{ij} „na 1. kostce padne číslo i a zároveň na 2. kostce padne číslo j “ pro $i, j = 1, 2, \dots, 6$. Pak prostorem elementárních jevů pro náhodný pokus „hod 2 kostkami“ je množina obsahující všech 36 elementárních jevů $S = \{E_{11}, E_{12}, \dots, E_{65}, E_{66}\}$.

Jakýkoliv další jev opět můžeme vyjádřit jako sjednocení těchto elementárních jevů. Položme jako jev B „padne sudé číslo na 1. kostce a zároveň padne sudé číslo na 2. kostce“, jako jev C „na 1. kostce padne číslo dělitelné 3 a zároveň na 2. kostce padne sudé číslo“. Pak

$$B = E_{22} \cup E_{24} \cup E_{26} \cup E_{42} \cup E_{44} \cup E_{46} \cup E_{62} \cup E_{64} \cup E_{66}.$$

Jevu B odpovídá podmnožina $E_B = \{E_{ij}\}$ množiny E o 9 prvcích (elementárních jevech) pro $i, j = 2, 4, 6$. Dále

$$C = E_{32} \cup E_{34} \cup E_{36} \cup E_{62} \cup E_{64} \cup E_{66},$$

přičemž jevu C odpovídá podmnožina $E_C = \{E_{32}, E_{34}, E_{36}, E_{62}, E_{64}, E_{66}\}$ množiny S o 6 prvcích (elementárních jevech).

2.1.5 Stabilita relativních četností

Budeme-li mnohokrát opakovat stejný náhodný pokus a budeme-li sledovat výskyt určitého jevu v každé realizaci tohoto pokusu, obdržíme nejrůznější výsledky, které se budou nahodile, a tudíž zcela nepravidelně střídat. V jednotlivém pokusu tak nejsme schopni dopředu určit, jaký výsledek máme očekávat. Hodnotíme-li však všechny dosažené výsledky najednou (za předpokladu velkého množství realizací pokusu), objevují se u mnoha náhodných pokusů jisté pravidelnosti či zákonitosti, které lze dobře interpretovat. Tuto situaci vysvětlíme pomocí tzv. **stability relativních četností**.

Stabilitou relativních četností rozumíme fakt, že relativní četnost výskytu sledovaného jevu při velkém počtu opakování náhodného pokusu kolísá kolem určitého čísla,

přičemž velikost kolísání se zmenšuje při zvyšujícím se počtu pokusů. Při velkém počtu pokusů se tedy relativní četnost ustaluje na určitém čísle. Poznamenejme ještě závěrem, že stabilita relativních četností je empirickým základem (myšleno odvozeným z pozorování) pravděpodobnosti.

Příklad 2.9

V tabulce 2.1 jsou čísla tažená v loterii Sportka za období 21. 4. 1957 – 1. 2. 2012.

Protože všechna čísla mají stejnou šanci být vylosována (49 čísel, tedy šance je 1:49), není důvod se domnívat, že by se jednotlivé četnosti výrazně odlišovaly. Je samozřejmé, že absolutní četnost každého čísla (tedy počet případů, kdy bylo číslo vylosováno) je jiná. Podíváme-li se však na relativní četnosti, je vidět, že se příliš neodlišují od 1/49, tedy od čísla 0,020 4. Odchylky jsou vesměs řádově až na 4. desetinném místě, tedy v desítitisících. Potvrzuje se nám tedy, že při velkém množství realizací pokusu relativní četnost kolísá kolem čísla 0,020 4.

Tab. 2.1 Čísla tažená ve Sportce

četnost											
číslo	absolutní	relativní	číslo	absolutní	relativní	číslo	absolutní	relativní	číslo	absolutní	relativní
1	957	0,0200	8	983	0,0206	15	998	0,0209	22	989	0,0207
2	988	0,0207	9	963	0,0202	16	930	0,0195	23	960	0,0201
3	936	0,0196	10	942	0,0197	17	965	0,0202	24	988	0,0207
4	923	0,0193	11	992	0,0208	18	1000	0,0209	25	965	0,0202
5	987	0,0207	12	1009	0,0211	19	1003	0,0210	26	976	0,0204
6	1039	0,0218	13	1009	0,0211	20	1018	0,0213	27	934	0,0196
7	962	0,0201	14	966	0,0202	21	960	0,0201	28	975	0,0204
číslo	absolutní	relativní	číslo	absolutní	relativní	číslo	absolutní	relativní	číslo	absolutní	relativní
29	999	0,0209	36	1008	0,0211	43	965	0,0202			
30	987	0,0207	37	979	0,0205	44	997	0,0209			
31	987	0,0207	38	958	0,0201	45	913	0,0191			
32	994	0,0208	39	993	0,0208	46	994	0,0208			
33	933	0,0195	40	1017	0,0213	47	1008	0,0211			
34	958	0,0201	41	915	0,0192	48	983	0,0206			
35	976	0,0204	42	949	0,0199	49	916	0,0192			

Zdroj: www.sazka.cz

2.1.6 Klasická definice pravděpodobnosti

Klasickou definici pravděpodobnosti můžeme použít pro **konečný prostor** elementárních ievů. Zavedl ji statistik Laplace a budeme ji interpretovat v upraveném tvaru.

Pokud má prostor elementárních jevů n prvků, které pokládáme za stejně možné, a jestliže jev A je sjednocením m elementárních jevů, pak je pravděpodobnost jevu A definována jako

$$P(A) = \frac{m}{n}. \quad (2.1)$$

Uvedený vzorec lze interpretovat i takto: pokud nazveme elementární jevy, při jejichž výskytu nastává jev A , jako **výsledky příznivé jevu A** , pak klasická definice pravděpodobnosti říká, že pravděpodobnost jevu A je rovna **podílu výsledků příznivých jevu A a počtu všech možných výsledků**.

V této definici je třeba zdůraznit stejnou možnost výskytu jednotlivých elementárních jevů. Při výpočtu pravděpodobnosti na základě klasické definice se uplatňuje zejména kombinatorika, tedy operace jako kombinace, variace a permutace, a to jak s opakováním, tak bez opakování.

2.1.7 Geometrická interpretace

Modifikací klasické definice pravděpodobnosti získáme tzv. geometrickou definici pravděpodobnosti. Označme jako S prostor stejně možných elementárních jevů. Pokud je tento prostor geometrickou oblastí (libovolné dimenze), jejíž velikost (délka, obsah, objem, ...) je konečná a je rovna $V(S)$, a část této oblasti představující jev A má velikost $V(A)$, pak je pravděpodobnost jevu A rovna

$$P(A) = \frac{V(A)}{V(S)}. \quad (2.2)$$

Příklad 2.1

Hodiny, které nebyly včas nataženy, se po určité době zastaví. Jaká je pravděpodobnost, že se velká ručička zastaví mezi číslicemi 4 a 6 (jev A)?

Řešení

Hledaná pravděpodobnost je rovna podílu délky oblouku mezi číslicí 4 a 6 a délky obvodu celého číselníku. V našem případě je tedy

$$P(A) = \frac{V(A)}{V(S)} = \frac{2}{12} = \frac{1}{6}.$$

2.1.8 Statistická definice pravděpodobnosti

Klasická definice pravděpodobnosti je omezena pouze na konečný prostor elementárních jevů a má i některá další omezení, jako je stejná možnost nastoupení elementárních jevů. V případě, že tyto předpoklady nejsou splněny, již s klasickou definicí nevystačíme. Tyto potíže dále vedly k definici tzv. **statistické (empirické) definici pravděpodobnosti**, navržené R. von Misesem. Tato definice pravděpodobnosti je založena na relativních četnostech a na stabilitě relativních četností.

Připomeňme, že stabilitou relativních četností rozumíme fakt, že relativní četnost sledovaného jevu A při velkém počtu opakování náhodného pokusu kolísá kolem určitého čísla, přičemž velikost kolísání se zmenšuje při zvyšujícím se počtu pokusů. Při velkém počtu pokusů se tedy relativní četnost ustavuje na určitém číslu. Je tudíž

přirozené chápat toto číslo jako míru početnosti výskytů příslušného jevu a nazvat jej pravděpodobností tohoto jevu.

Pravděpodobnost jevu A je takové číslo $P(A)$ přiřazené jevu A , které má tu vlastnost, že relativní četnost $p(A)$ jevu A se s rostoucím počtem realizací pokusu blíží číslu $P(A)$ (Likeš, Machek, 1991).

2.1.9 Axiomy pravděpodobnosti

Nyní uvedeme tzv. **axiomy pravděpodobnosti**. Tyto axiomy jsou založeny na vlastnostech relativních četností – uvědomme si, že relativní četnosti jsou empirickým protějškem pravděpodobnosti.

1. Pravděpodobnost náhodného jevu A je číslo mezi 0 a 1

$$0 \leq P(A) \leq 1. \quad (2.3)$$

2. Nechť A_1, A_2, \dots je konečný nebo spočetný systém disjunktních náhodných jevů, pak pravděpodobnost sjednocení jevů je rovna součtu pravděpodobností jednotlivých jevů

$$A_i \cap A_j = \emptyset, \quad i \neq j = 1, 2, \dots \Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i). \quad (2.4)$$

3. Pravděpodobnost jevu jistého je rovna 1

$$S \text{ je jistý} \Rightarrow P(S) = 1.$$

2.1.10 Vlastnosti pravděpodobnosti

Všechny dále uvedené vlastnosti pravděpodobnosti jsou důsledkem výše uvedených axiomů a dají se tedy z těchto axiomů odvodit.

1. Součet pravděpodobnosti jevu A a pravděpodobnosti doplňkového jevu \bar{A} je roven 1

$$P(A) + P(\bar{A}) = 1. \quad (2.5)$$

2. Pravděpodobnost jevu nemožného je rovna nule

$$P(\emptyset) = 0. \quad (2.6)$$

3. Pravděpodobnost rozdílu $A \setminus B$ jevu A a B se rovná

$$P(A \setminus B) = P(A) - P(A \cap B). \quad (2.7)$$

4. Pravděpodobnost sjednocení dvou jevů je rovna součtu pravděpodobností prvního a druhého jevu ménus pravděpodobnost jejich průniku

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \quad (2.8)$$

5. Pokud jev A implikuje jev B , pak pravděpodobnost jevu A je menší nebo rovna pravděpodobnosti jevu B

$$A \subset B \Rightarrow P(A) \leq P(B). \quad (2.9)$$

2.1.11 Nezávislost náhodných jevů

Klíčovou úlohu v celé pravděpodobnosti sehrává pojem **nezávislost**. Budeme jí pro náhodné jevy chápát následovně.

Řekneme, že jevy A a B jsou vzájemně nezávislé, jestliže

$$P(A \cap B) = P(A)P(B), \quad (2.10)$$

tedy jestliže pravděpodobnost jejich průniku je rovna součinu jejich pravděpodobností. Slovo **vzájemně** se často vymezuje a hovoříme pouze o nezávislosti jevů.

Pokud jsou jevy A a B nezávislé, jsou nezávislé i jevy A a \bar{B} , dále jevy \bar{A} a B a také jevy \bar{A} a \bar{B} . Pracujeme-li s více než dvěma jevy, definujeme jejich nezávislost obdobně. Předchozí definice pak musí platit pro každou dvojici, trojici, čtverici atd.

2.1.12 Podmíněná pravděpodobnost a nezávislost

Statistická definice pravděpodobnosti byla založena na stabilitě relativních četností. Na základě stejného principu budeme chápát i podmíněnou pravděpodobnost.

Pravděpodobnost jevu A podmíněnou jevem B definujeme jako

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (2.11)$$

Pro $P(B) > 0$. Uvedený vzorec lze přepsat do tvaru

$$P(A \cap B) = P(A)P(A|B), \quad (2.12)$$

který se používá v případech, kdy známe podmíněnou pravděpodobnost a pravděpodobnost jevu podmínějícího, a chceme určit pravděpodobnost průniku těchto jevů.

Pro **nezávislé** jevy A a B , pro které $P(B) > 0$ resp. $P(A) > 0$, platí

$$P(A|B) = P(A), \quad (2.13)$$

resp.

$$P(B|A) = P(B). \quad (2.14)$$

Platí totiž

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A) \quad (2.15)$$

a analogicky i pro $P(B|A)$. Uvedené vzorce se někdy používají jako definice nezávislosti jevů.

2.1.13 Úplná pravděpodobnost

Uvažujme dva disjunktní jevy B_1 a B_2 , jejichž pravděpodobnosti $P(B_1) > 0$ a $P(B_2) > 0$ jsou známé. Dále uvažujme libovolný jev A s neznámou pravděpodobností $P(A)$. Jev A můžeme zapsat jako sjednocení dvou disjunktních jevů $A \cap B_1$ a $A \cap B_2$. Pak

$$P(A) = P(A \cap B_1) + P(A \cap B_2). \quad (2.16)$$

Jelikož

$$P(A \cap B_1) = P(B_1)P(A|B_1) \text{ a } P(A \cap B_2) = P(B_2)P(A|B_2),$$

můžeme psát

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2). \quad (2.17)$$

Pokud tedy známe podmíněné pravděpodobnosti $P(A|B_1)$ a $P(A|B_2)$, můžeme určit neznámou pravděpodobnost $P(A)$. V praxi se totiž mnohdy stává, že není těžké určit podmíněné pravděpodobnosti pro určitý jev, je však dosti obtížné určit jeho pravděpodobnost přímo. Nyní uvedené úvahy zobecníme pro více jevů B_1, B_2, \dots, B_n .

Věta o úplné pravděpodobnosti

Nechť B_1, B_2, \dots, B_n jsou navzájem disjunktní jevy, pro které platí

$$P(B_j) > 0, \quad (2.18)$$

$$\bigcup_{j=1}^n B_j = S, \quad (2.19)$$

kde S je jev jistý.

Nechť A je libovolný jev s nenulovou pravděpodobností, pro který jsou podmíněny pravděpodobnosti $P(A|B_j)$ známé pro $j = 1, 2, \dots, n$. Pak pravděpodobnost jevu A je

$$P(A) = \sum_{j=1}^n P(A|B_j)P(B_j). \quad (2.20)$$

Vše uvedené vlastnosti podmíněné pravděpodobnosti nám umožňují definovat tzv. Bayesovu větu, jejíž znění dále uvedeme.

Příklad 2.11

Velkoobchod odebírá počítače od dvou dodavatelů. První dodavatel pokrývá odběr velkoobchodu z 80 %, přičemž 75 % dodávky tvoří počítače osazené procesorem Intel. Druhý dodavatel pokrývá odběr velkoobchodu ze zbývajících 20 %, přičemž 60 % dodávky tvoří počítače osazené procesorem Intel. Pokud by si zákazník vybral počítač zcela náhodně (bez ohledu na procesor), jaká je pravděpodobnost, že tento počítač bude osazen procesorem Intel?

Řešení

Označme jevy

A – náhodně vybraný počítač je osazen procesorem Intel,

B_1 – náhodně vybraný počítač je od prvního dodavatele,

B_2 – náhodně vybraný počítač je od druhého dodavatele.

Ze zadání víme, že

$$\begin{aligned} P(B_1) &= 0,80, & P(A|B_1) &= 0,75, \\ P(B_2) &= 0,20, & P(A|B_2) &= 0,60. \end{aligned}$$

Podle předchozích vzorců můžeme psát

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) = 0,80 \cdot 0,75 + 0,20 \cdot 0,60 = 0,72.$$

Bayesova věta

Nechť B_1, B_2, \dots, B_n jsou navzájem disjunktní jevy, pro které platí

$$P(B_j) > 0, \quad j = 1, 2, \dots, n, \quad (2.21)$$

$$P\left(\bigcup_{j=1}^n B_j\right) = \sum_{j=1}^n P(B_j) = 1, \quad j = 1, 2, \dots, n. \quad (2.22)$$

Nechť A je libovolný jev s nenulovou pravděpodobností, pro který jsou podmíněny pravděpodobnosti $P(A|B_j)$ známé pro všechna $j = 1, 2, \dots, n$. Pak pravděpodobnost jevu B_j podmíněná jevem A je rovna

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{j=1}^n P(A|B_j)P(B_j)}, \quad j = 1, 2, \dots, n. \quad (2.23)$$

2.2 Rozdělení pravděpodobnosti

Rozdělení pravděpodobnosti sehrává zásadní úlohu jak v pravděpodobnosti, tak v celé matematické statistice (viz další kapitola). V každé teoretické úloze, při každé aplikaci nás bude zajímat, s jakou pravděpodobností nastávají sledované jevy. Klíčovou rolí přitom sehrává náhodná veličina a její chování.

2.2.1 Náhodná veličina

Většina náhodných pokusů má výsledek vyjádřený reálným číslem. I v případě, že výsledek náhodného pokusu je kvalitativní povahy (barva očí, stupeň dosaženého vzdělání, pohlaví atd.), stejně většinou výsledky převádíme na čísla – určujeme počet, relativní četnost atd.

V těchto případech tedy pracujeme s náhodným pokusem, jehož prostor elementárních jevů S je množina reálných čísel R nebo některá její podmnožina. Potom hodnotou náhodné veličiny X budeme rozumět výsledek náhodného pokusu, vyjádřený reálným číslem. Tím jsme naznačili, co je hodnotou náhodné veličiny, nedefinovali jsme však náhodnou veličinu jako takovou. **Náhodnou veličinu** v tomto textu budeme chápat jako funkci definovanou na prostoru elementárních jevů, která každému elementárnímu jevu přiřadí číslo na reálné ose.

Na rozdíl od náhodných jevů, kdy jsme používali velká písmena ze začátku abecedy, náhodné veličiny budeme značit velkými písmeny z konce abecedy X, Y, Z, \dots

Pokud budeme schopni jakékoli podmnožině reálné osy přiřadit pravděpodobnost, že náhodná veličina nabývá hodnoty právě z této podmnožiny, hovoříme o tom, že známe rozdělení pravděpodobnosti této náhodné veličiny. Většinou je onou zmiňovanou podmnožinou nějaký interval. V celém textu budeme používat symboliku známou z teorie množin.

62 Pravděpodobnost

můžeme psát

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2). \quad (2.17)$$

Pokud tedy známe podmíněné pravděpodobnosti $P(A|B_1)$ a $P(A|B_2)$, můžeme určit neznámou pravděpodobnost $P(A)$. V praxi se totiž mnohdy stává, že není těžké určit podmíněné pravděpodobnosti pro určitý jev, je však dosti obtížné určit jeho pravděpodobnost přímo. Nyní uvedené úvahy zobecníme pro více jevů B_1, B_2, \dots, B_n .

Věta o úplné pravděpodobnosti

Nechť B_1, B_2, \dots, B_n jsou navzájem disjunktní jevy, pro které platí

$$P(B_j) > 0, \quad j = 1, 2, \dots, n, \quad (2.18)$$

$$\bigcup_{j=1}^n B_j = S, \quad (2.19)$$

kde S je jev jistý.

Nechť A je libovolný jev s nenulovou pravděpodobností, pro který jsou podmíněny pravděpodobnosti $P(A|B_j)$ známé pro $j = 1, 2, \dots, n$. Pak pravděpodobnost jevu A je

$$P(A) = \sum_{j=1}^n P(A|B_j)P(B_j). \quad (2.20)$$

Výše uvedené vlastnosti podmíněné pravděpodobnosti nám umožňují definovat tzv. Bayesovu větu, jejíž znění dále uvedeme.

Příklad 2.11

Velkoobchod odebírá počítače od dvou dodavatelů. První dodavatel pokrývá odběr velkoobchodu z 80 %, přičemž 75 % dodávky tvoří počítače osazené procesorem Intel. Druhý dodavatel pokrývá odběr velkoobchodu ze zbyvajících 20 %, přičemž 60 % dodávky tvoří počítače osazené procesorem Intel. Pokud by si zákazník vybral počítač buď zcela náhodně (bez ohledu na procesor), jaká je pravděpodobnost, že tento počítač bude osazen procesorem Intel?

Řešení

Označme jevy

A – náhodně vybraný počítač je osazen procesorem Intel,

B_1 – náhodně vybraný počítač je od prvního dodavatele,

B_2 – náhodně vybraný počítač je od druhého dodavatele.

Ze zadání víme, že

$$\begin{aligned} P(A|B_1) &= 0,75, \\ P(B_1) &= 0,80, \\ P(B_2) &= 0,20, \end{aligned}$$

$$\begin{aligned} P(A|B_2) &= 0,60. \\ P(A) &= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) = 0,80 \cdot 0,75 + 0,20 \cdot 0,60 = 0,72. \end{aligned}$$

Bayesova věta

Nechť B_1, B_2, \dots, B_n jsou navzájem disjunktní jevy, pro které platí

$$P(B_j) > 0, \quad j = 1, 2, \dots, n, \quad (2.21)$$

$$P\left(\bigcup_{j=1}^n B_j\right) = \sum_{j=1}^n P(B_j) = 1, \quad j = 1, 2, \dots, n. \quad (2.22)$$

Nechť A je libovolný jev s nenulovou pravděpodobností, pro který jsou podmíněny pravděpodobnosti $P(A|B_j)$ známé pro všechna $j = 1, 2, \dots, n$. Pak pravděpodobnost jevu B_j podmíněná jevem A je rovna

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{j=1}^n P(A|B_j)P(B_j)}, \quad j = 1, 2, \dots, n. \quad (2.23)$$

2.2 Rozdělení pravděpodobnosti

Rozdělení pravděpodobnosti sehrává zásadní úlohu jak v pravděpodobnosti, tak v celé matematické statistice (viz další kapitola). V každé teoretické úloze, při každé aplikaci nás bude zajímat, s jakou pravděpodobností nastávají sledované jevy. Klíčovou rolí přitom sehrává náhodná veličina a její chování.

2.2.1 Náhodná veličina

Většina náhodných pokusů má výsledek vyjádřený reálným číslem. I v případě, že výsledek náhodného pokusu je kvalitativní povahy (barva očí, stupeň dosaženého vzdělání, pohlaví atd.), stejně většinou výsledky převádíme na čísla – určujeme počet, relativní četnost atd.

V těchto případech tedy pracujeme s náhodným pokusem, jehož prostor elementárních jevů S je množina reálných čísel R nebo některá její podmnožina. Potom hodnotou náhodné veličiny X budeme rozumět výsledek náhodného pokusu, vyjádřený reálným číslem. Tím jsme naznačili, co je hodnotou náhodné veličiny, nedefinovali jsme však náhodnou veličinu jako takovou. **Náhodnou veličinu** v tomto textu budeme chápat jako funkci definovanou na prostoru elementárních jevů, která každému elementárnímu jevu přiřadí číslo na reálné ose.

Na rozdíl od náhodných jevů, kdy jsme používali velká písmena ze začátku abecedy, náhodné veličiny budeme značit velkými písmeny z konce abecedy X, Y, Z, \dots

Pokud budeme schopni jakékoli podmnožině reálné osy přiřadit pravděpodobnost, že náhodná veličina nabývá hodnoty právě z této podmnožiny, hovoříme o tom, že známe rozdělení pravděpodobnosti této náhodné veličiny. Většinou je onou zmiňovanou podmnožinou nějaký interval. V celém textu budeme používat symboliku známou z teorie množin.

2.2.2 Distribuční funkce

Univerzální formou popisu pravděpodobnostního chování náhodné veličiny je **distribuční funkce**. Distribuční funkci náhodné veličiny X budeme rozumět reálnou funkcí $F(x)$, definovanou jako

$$F(x) = P(X \leq x), \quad \forall x \in R. \quad (2.24)$$

Distribuční funkce $F(x)$ v bodě x tedy vyjadřuje pravděpodobnost, že náhodná veličina X nabude hodnoty menší nebo rovné x . V tomto zápisu je důležité si uvědomit, že velkým písmenem značíme náhodnou veličinu, zatímco malým písmenem značíme libovolné číslo na reálné ose. Nerovnost ve výrazu (2.24) je neostrá, což se projevuje v bodech nespojitosti, jak ostatně dále uvidíme.

Vlastnosti distribuční funkce

Distribuční funkce má vlastnosti, které vyplývají především z faktu, že se jedná o pravděpodobnost:

1. $0 \leq F(x) \leq 1, \quad \forall x \in R.$

Tato vlastnost je zřejmá, pracujeme s pravděpodobností.

2. $\forall x_1, x_2 \in R; x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2).$

Distribuční funkce je neklesající.

3. $\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow +\infty} F(x) = 1.$

Uvedené limity budeme symbolicky zapisovat $F(-\infty) = 0$ a $F(\infty) = 1$. V případě, že definiční obor náhodné veličiny je interval (a, b) , platí

$$F(a) = 0, \quad F(b) = 1. \quad (2.25)$$

4. Distribuční funkce je zprava spojita a má nejvýš konečně nebo spočetně mnoho bodů nespojitosti.

Je třeba dodat, že distribuční funkce jednoznačně popisuje pravděpodobnostní chování náhodné veličiny. To znamená, že pokud známe distribuční funkci náhodné veličiny, máme kompletní informaci o jejím chování.

Z definice distribuční funkce vyplývá, že platí

$$P(x_1 < X \leq x_2) = F(x_2) - F(x_1), \quad -\infty < x_1 < x_2 < \infty. \quad (2.26)$$

2.2.3 Rozdělení diskrétního a spojitého typu

V praxi se většinou setkáváme s náhodnými veličinami dvojího typu. Nejlépe budeme celou situaci ilustrovat na příkladech. Hodnotou náhodné veličiny může být např.

- počet sourozenců náhodně vybrané osoby,
- počet dopravních nehod na území hlavního města Prahy za jeden pracovní den,
- počet zákazníků ve frontě u pokladny v supermarketu,
- počet cestujících, které přepraví metro za 1 hodinu atd.

Ve všech těchto případech pracujeme s náhodnou veličinou X , která nabývá hodnot z nějaké konečné nebo spočetné množiny $\{x_1, x_2, \dots\}$. Často se přitom jedná o celočíselné hodnoty. V takovém případě hovoříme o náhodné veličině, která má pravděpodobnostní rozdělení diskrétního typu, resp. o diskrétní (nespojitě) náhodné veličině.

Na druhé straně může být hodnotou náhodné veličiny například

- doba životnosti určitého zařízení,
- hmotnost či výška náhodné vybrané osoby,
- délka života právě narozeného jedince,
- chyba vzniklá vážením, měřením či zaokrouhlováním atd.

V těchto případech pracujeme s náhodnou veličinou X , která nabývá libovolných hodnot z nějakého intervalu. Pak hovoříme o náhodné veličině, která má pravděpodobnostní rozdělení spojitého typu, zkráceně o spojité náhodné veličině.

Diskrétní náhodná veličina

Definujme nyní diskrétní (nespojitu) náhodnou veličinu. Řekneme, že náhodná veličina X má rozdělení diskrétního typu, existuje-li konečná nebo spočetná množina reálných čísel $\{x_1, x_2, \dots\}$ taková, že pro každé x_j z této množiny je pravděpodobnost $P(X = x_j) > 0$ a součet všech těchto pravděpodobností přes všechna x_j z této množiny je roven jedné

$$\sum_{x_j} P(X = x_j) = 1. \quad (2.27)$$

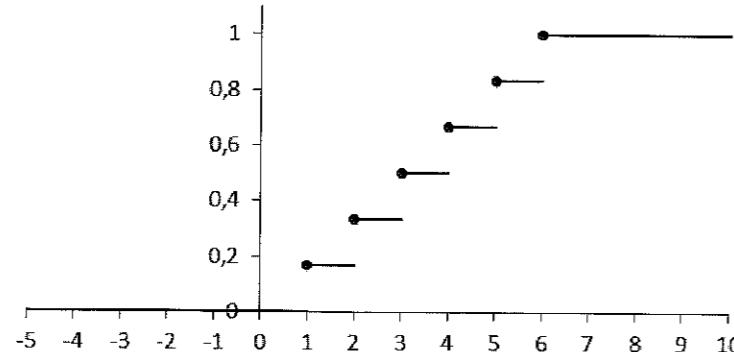
Pravděpodobnosti pro všechny hodnoty z množiny $\{x_1, x_2, \dots\}$ můžeme znázornit tabulkou, grafem či vzorcem. Poslední možnost se jeví jako zdaleka nevhodnější. Funkce $P(X = x)$ se nazývá pravděpodobnostní funkci náhodné veličiny X a většinou se udává právě tato funkce.

Známe-li hodnoty pravděpodobnostní funkce diskrétní náhodné veličiny X , snadno určíme její distribuční funkci jako

$$F(x) = \sum_{x_j \leq x} P(X = x_j), \quad x \in R. \quad (2.28)$$

Distribuční funkci náhodné veličiny X v bodě x tedy získáme jako součet pravděpodobností pro všechny hodnoty x_j , $j = 1, 2, \dots$, menší nebo rovné x . Z uvedeného vztahu vyplývá, že i pravděpodobnostní funkce jednoznačně popisuje rozdělení náhodné veličiny. Je tedy jedno, zda známe distribuční funkci či pravděpodobnostní funkci, obě funkce nám poskytují úplnou informaci o rozdělení náhodné veličiny. Existuje tudíž jednoznačný vztah mezi distribuční funkci a pravděpodobnostní funkci. K úplnému popisu pravděpodobnostního chování nám tedy postačuje i znalost pravděpodobnostní funkce.

Jak může vypadat graf distribuční funkce diskrétní náhodné veličiny, je vidět z obrázku 2.1.



Obr. 2.1 Distribuční funkce diskrétní náhodné veličiny

Jde se o distribuční funkci náhodné veličiny X , která má pravděpodobnostní funkci

$$\begin{aligned} P(x) &= \frac{1}{6}, & x = 1, 2, \dots, 6, \\ &= 0, & \text{jinak.} \end{aligned} \quad (2.29)$$

Takovouto náhodnou veličinou může být třeba výsledek hodu hrací kostkou. Distribuční funkce potom nabývá hodnot

$$\begin{aligned} F(x) &= 0, & x < 1, \\ &= 1/6, & 1 \leq x < 2, \\ &= 2/6, & 2 \leq x < 3, \\ &= 3/6, & 3 \leq x < 4, \\ &= 4/6, & 4 \leq x < 5, \\ &= 5/6, & 5 \leq x < 6, \\ &= 1, & x \geq 6. \end{aligned} \quad (2.30)$$

Uvažovaná náhodná veličina se řídí diskrétním rovnoměrným rozdelením. U tohoto rozdelení nabývá náhodná veličina X hodnot $j = 1, 2, \dots, n$, každé z nich se stejnou pravděpodobností $1/n$. V předchozím příkladu je $n = 6$. Pravděpodobnostní funkce náhodné veličiny X pak tvar

$$\begin{aligned} P(x) &= \frac{1}{n}, & x = 1, 2, \dots, n, \\ &= 0, & \text{jinak,} \end{aligned} \quad (2.31)$$

a distribuční funkce má tvar

$$\begin{aligned} F(x) &= 0, & x < 1, \\ &= \frac{k}{n}, & k \leq x < k+1, \quad k = 1, 2, \dots, n-1, \\ &= 1, & x \geq n. \end{aligned} \quad (2.32)$$

Graf na obrázku 2.1 dobře ilustruje tvar distribuční funkce diskrétní náhodné veličiny – jedná se o „schody“ (v tomto případě stejně „vysoké a dlouhé“, což není obecně pravidlem). Navíc jsou z tohoto grafu jasné patrné i vlastnosti distribuční funkce.

Spojitá náhodná veličina

Náhodná veličina X má rozdelení spojitého typu, existuje-li nezáporná reálná funkce $f(x)$ taková, že pro všechna reálná x lze distribuční funkci $F(x)$ vyjádřit ve tvaru

$$F(x) = \int_{-\infty}^x f(t) dt, \quad -\infty < x < \infty. \quad (2.33)$$

Funkce $f(x)$ se nazývá hustota pravděpodobnosti náhodné veličiny X . Většinou mluvíme zkráceně pouze o hustotě náhodné veličiny. Na rozdíl od diskrétní náhodné veličiny je distribuční funkce spojitá pro všechna reálná x . Z definice vyplývá, že platí

$$f(x) = \frac{dF(x)}{dx} \quad (2.34)$$

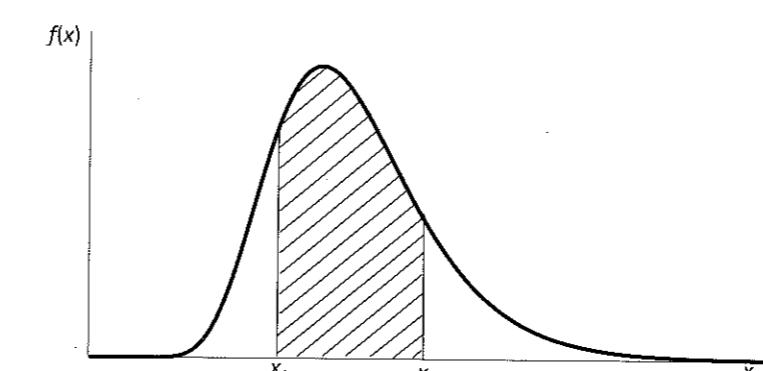
ve všech bodech, kde derivace existuje.

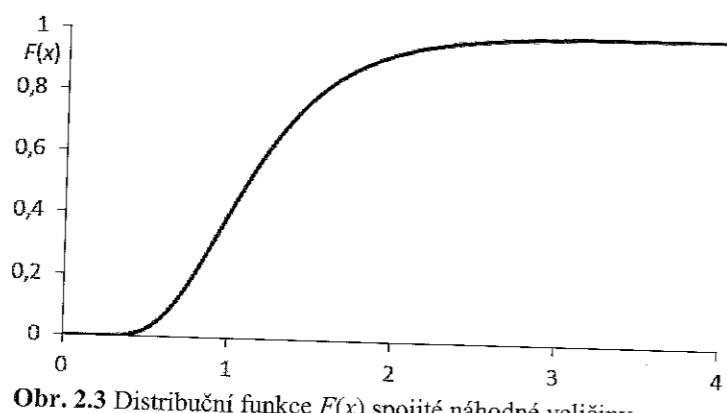
Podobně jako u diskrétní veličiny je jedno, zda známe distribuční funkci či hustotu. Obě funkce plně charakterizují pravděpodobnostní rozdelení náhodné veličiny a díky výše uvedeným vztahům dokážeme z jedné funkce určit druhou.

Z definice distribuční funkce a hustoty pravděpodobnosti plyne

$$P(x_1 < X \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx \quad (2.35)$$

pro všechna reálná $-\infty < x_1 < x_2 < \infty$. Uvažovaná pravděpodobnost je tedy rovna velikosti vyšrafované plochy pod křivkou hustoty, vymezenou body x_1 a x_2 na obrázku 2.2. Na obrázku 2.3 je znázorněn průběh distribuční funkce stejné náhodné veličiny.

Obr. 2.2 Pravděpodobnost $P(x_1 < X < x_2)$

Obr. 2.3 Distrifuční funkce $F(x)$ spojité náhodné veličiny

Z definice distrifuční funkce a jejích vlastností vyplývá, že

$$\int_{-\infty}^{\infty} f(x) dx = 1. \quad (2.36)$$

Pokud náhodná veličina X nabývá hodnot z intervalu (a, b) , a je tedy $f(x) = 0$ pro $x \leq a$ a $x \geq b$, lze předchozí vztah přepsat do podoby

$$\int_a^b f(x) dx = 1. \quad (2.37)$$

Dále pro každé reálné x platí

$$P(X = x) = 0. \quad (2.38)$$

Pak je již zřejmé, že platí i vztah

$$\begin{aligned} P(x_1 \leq X \leq x_2) &= P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) = \\ &= P(x_1 < X < x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx. \end{aligned} \quad (2.39)$$

2.2.4 Číselné charakteristiky náhodných veličin

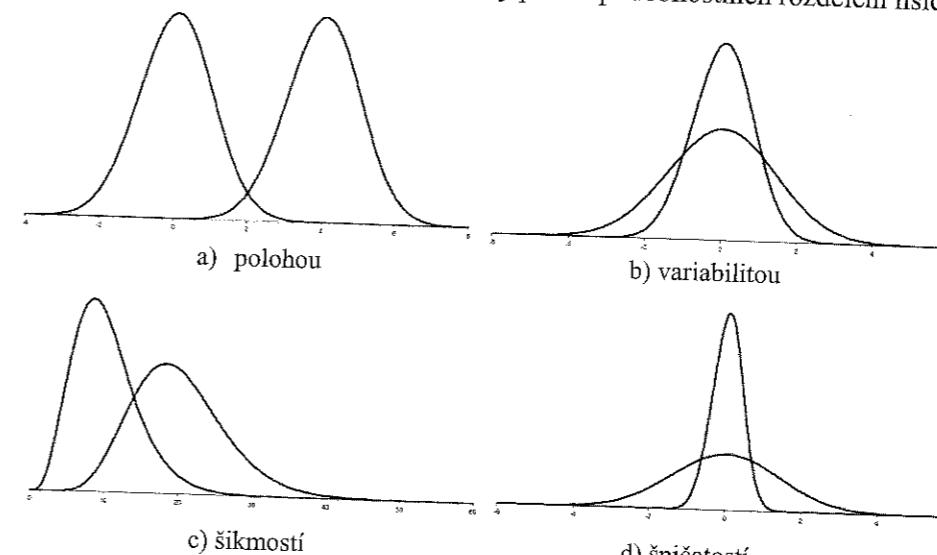
Z předchozího textu již víme, že pravděpodobnostní chování (pravděpodobnostní rozdelení) náhodné veličiny X je obecně určeno tvarem distrifuční funkce $F(x)$, a to jak pro spojitou, tak pro diskrétní náhodnou veličinu. V případě diskrétní náhodné veličiny můžeme její chování plně charakterizovat pravděpodobnostní funkcí $P(x)$, v případě spojité náhodné veličiny hustotou pravděpodobnosti $f(x)$. Ve všech případech nám znalost příslušné funkce podává úplnou informaci o náhodné veličině X .

Mnohdy se však jako nanejvýš praktické ukazuje shrnutí důležitých informací do několika málo čísel, která nám ilustrují vlastnosti náhodné veličiny. Tato čísla označujeme jako číselné charakteristiky náhodné veličiny.

Číselné charakteristiky náhodných veličin lze rozdělit do čtyř skupin, a to na charakteristiky

- a) polohy,
- b) variability,
- c) šíkmosti,
- d) špičatosti.

Na obrázcích 2.4 jsou znázorněny hustoty pravděpodobnostních rozdelení lišící se



Obr. 2.4 Různé tvary rozdelení pravděpodobnosti náhodné veličiny

Střední hodnota

Střední hodnota náhodné veličiny X je číslo, kolem kterého kolísá výskyt náhodné veličiny. Střední hodnota, někdy též označovaná jako očekávaná hodnota, náhodné veličiny X se značí $E(X)$. Je to nejdůležitější charakteristika polohy.

Pro nespojitou náhodnou veličinu X je střední hodnota definována jako

$$E(X) = \sum_{x_j} x_j P(X = x_j), \quad (2.40)$$

kde $x_j, j = 1, 2, \dots$, jsou hodnoty, kterých náhodná veličina X nabývá s pravděpodobností $P(X = x_j) > 0$.

Pro spojitolou náhodnou veličinu X je střední hodnota definována jako

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx, \quad (2.41)$$

kde $f(x)$ je hustota pravděpodobnosti náhodné veličiny X .

Podívejme se, jak vypadá střední hodnota pro náhodnou veličinu X s diskrétním rovnoměrným rozdělením

$$E(X) = \sum_{x=1}^n x P(X=x) = \sum_{x=1}^n x \frac{1}{n} = \frac{1}{n} \cdot \frac{n}{2} (1+n) = \frac{n+1}{2}. \quad (2.42)$$

Vlastnosti střední hodnoty

Při výpočtu střední hodnoty se často používají její vlastnosti, které přímo vyplývají z definice a z obecných vlastností sumy a integrálu:

1. $E(a) = a.$

Střední hodnota konstanty a je rovna konstantě a .

2. $E(bX) = bE(X).$

Střední hodnota součinu bX , kde b je konstanta a X náhodná veličina, je rovna součinu b a střední hodnoty X .

3. $E(a + bX) = a + bE(X).$

Tuto rovnost dostáváme spojením první a druhé vlastnosti.

4. $E(X + Y) = E(X) + E(Y).$

Střední hodnota součtu dvou náhodných veličin je rovna součtu středních hodnot těchto veličin.

5. $E(a + bX + cY) = a + bE(X) + cE(Y).$

Tuto vlastnost dostáváme spojením všech předchozích vlastností. Podotkněme ještě, že c je konstanta.

6. $E(XY) \neq E(X)E(Y).$

Obecně tedy střední hodnota součinu dvou náhodných veličin není rovna součinu středních hodnot těchto veličin. Rovnost v tomto vztahu platí pouze za předpokladu nezávislosti obou náhodných veličin.

Rozptyl

Jestliže střední hodnota náhodné veličiny X je číslo, kolem kterého kolísá výskyt náhodné veličiny, rozptyl určuje velikost tohoto kolísání. Jedná se tedy o míru variability.

Rozptyl náhodné veličiny X se označuje jako $D(X)$ či $\text{var}(X)$ a je definován jako

$$D(X) = E\{[X - E(X)]^2\}. \quad (2.43)$$

Tento vzorec lze snadnou úpravou převést do tvaru

$$D(X) = E(X^2) - [E(X)]^2. \quad (2.44)$$

Často se pro měření variability místo rozptylu používá směrodatná odchylka, která se počítá jako odmocnina z rozptylu.

Vlastnosti rozptylu

Při výpočtu rozptylu se stejně jako u střední hodnoty často používají vlastnosti, které přímo vyplývají z definice a z obecných vlastností sumy a integrálu:

1. $D(X) \geq 0.$

Rozptyl náhodné veličiny X je nezáporný, přičemž roven nule může být pouze v případě, kdy X je degenerovanou náhodnou veličinou, tedy konstantou – viz dále.

2. $D(a) = 0.$

Rozptyl konstanty je nula.

3. $D(bX) = b^2 D(X).$

Rozptyl součinu bX , kde b je konstanta a X náhodná veličina, je roven součinu čtverce konstanty b a rozptylu veličiny X .

4. $D(a + bX) = b^2 D(X).$

Tuto rovnost dostáváme spojením první a druhé vlastnosti.

5. $D(X + Y) = D(X) + \text{cov}(X, Y) + D(Y).$

Rozptyl součtu dvou náhodných veličin je roven součtu rozptylů těchto veličin a dvojnásobku jejich kovariance. Kovariance je přitom definována

$$\text{cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}, \quad (2.45)$$

což lze snadno upravit do tvaru

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y). \quad (2.46)$$

Kovariance je nástroj pro měření lineární závislosti náhodných veličin, podrobněji bude tento pojem vysvětlen v části věnované náhodnému vektoru.

6. $D(aX + bY + c) = a^2 D(X) + 2ab \text{cov}(X, Y) + b^2 D(Y),$

kde a, b, c jsou konstanty. Tato vlastnost je spojením všech předchozích vlastností.

Příklad 2.12

Nechť náhodné veličiny X a Y mají následující střední hodnotu a rozptyl: $E(X) = -1$, $E(Y) = 1$, $D(X) = 18$, $D(Y) = 2$. Kovariance těchto veličin je rovna $\text{cov}(X, Y) = -3$. Určete střední hodnotu, rozptyl a kovarianci náhodných veličin Z a Q , které jsou definovány jako $Z = X - Y + 1$ a $Q = X + 3Y + 2$.

Řešení

Při výpočtu budeme vycházet z výše uvedených vlastností rozptylu. Nejprve spočítáme střední hodnoty a rozptyly veličin Z a Q

$$E(Z) = E(X) - E(Y) + 1 = -1 - 1 + 1 = -1,$$

$$E(Q) = E(X) + 3E(Y) + 2 = -1 + 3 + 2 = 4,$$

$$D(Z) = D(X - Y + 1) = D(X) + D(Y) - 2\text{cov}(X, Y) = 18 + 2 - 2 \cdot (-3) = 26,$$

$$D(Q) = D(X + 3Y + 2) = D(X) + 9D(Y) + 2 \cdot 3\text{cov}(X, Y) = 18 + 9 \cdot 2 + 6 \cdot (-3) = 18.$$

Dále určíme kovarianci veličin Z a Q . Nejprve spočítáme střední hodnotu součinu náhodných veličin Z a Q

$$\begin{aligned} E(ZQ) &= E[(X - Y + 1) \cdot (X + 3Y + 2)] = \\ &= E(X^2 + 3XY + 2X - XY - 3Y^2 - 2Y + X + 3Y + 2) = \\ &= E(X^2 + 2XY + 3X + Y - 3Y^2 + 2) = \\ &= E(X^2) + 2E(XY) + 3E(X) + E(Y) - 3E(Y^2) + 2. \end{aligned}$$

Musíme tedy vypočítat $E(X^2)$, $E(XY)$ a $E(Y^2)$

$$\begin{aligned} E(X^2) &= D(X) + E(X)^2 = 18 + 1^2 = 19, \\ E(Y^2) &= D(Y) + E(Y)^2 = 2 + 1^2 = 3, \\ E(XY) &= \text{cov}(X, Y) + E(X) \cdot E(Y) = -3 + (-1) \cdot 1 = -4. \end{aligned}$$

Konečně můžeme dosadit a vypočítat

$$E(ZQ) = 19 + 2 \cdot (-4) + 3 \cdot (-1) + 1 - 3 \cdot 3 + 2 = 2.$$

Potom tedy

$$\text{cov}(Z, Q) = 2 - (-1) \cdot 4 = 6.$$

Podívejme se na další charakteristiky náhodné veličiny. Mezi charakteristiky polohy dále patří momenty náhodné veličiny a kvantily.

Momenty náhodné veličiny

Obecné momenty

Definujme r -tý obecný moment jako

$$\mu'_r(X) = E(X^r), \quad r = 0, 1, 2, \dots \quad (2.47)$$

Nezajímavá je hodnota pro $r = 0$, neboť $\mu'_0(X) = E(X^0) = 1$. Dosadíme-li za $r = 1$, máme $\mu'_1(X) = E(X)$. To znamená, že střední hodnota je prvním obecným momentem náhodné veličiny X .

Centrální momenty

Předpokládejme, že $E(X)$ je konečná střední hodnota náhodné veličiny X . Potom. r -tý centrální moment má tvar

$$\mu_r(X) = E\{[X - E(X)]^r\}, \quad r = 0, 1, 2, \dots \quad (2.48)$$

Nezajímavá je opět hodnota pro $r = 0$ či $r = 1$, neboť $\mu_0(X) = 1$ a $\mu_1(X) = 0$. Pro $r = 2$ obdržíme

$$\mu_2(X) = E\{[X - E(X)]^2\} = \mu'_2(X) - [E(X)]^2, \quad r = 0, 1, 2, \dots \quad (2.49)$$

To znamená, že druhým centrálním momentem náhodné veličiny X je její rozptyl. Ten můžeme tedy zapsat jako

$$\mu_2(X) = D(X) = E(X) - [E(X)]^2. \quad (2.50)$$

Často se uvádějí i centrální momenty vyšších řádů. Lze je nalézt v Marek (2012).

Kvantily

Definujme $100P\%$ kvantil pravděpodobnostního rozdělení (často hovoříme i o kvantili náhodné veličiny X – rozumíme tím veličinu, mající příslušné rozdělení) jako číslo x_p takové, že pro dané P , $0 < P < 1$, platí

$$F(x_p) \leq P, \quad F(x_p + 0) \geq P, \quad (2.51)$$

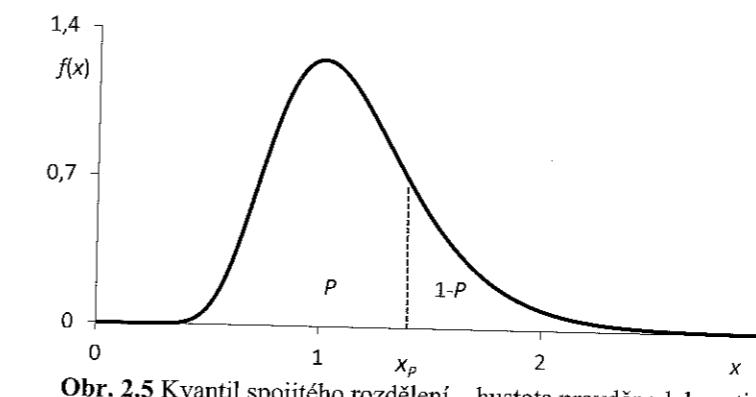
kde $F(x_p + 0)$ značí limitu funkce F v bodě x_p zprava. Z definice je zřejmé, že hodnota kvantilu není jednoznačně určena, neboť může existovat více než jedna hodnota (dokonce nekonečně mnoho hodnot), splňující výše uvedené nerovnosti.

Pokud má náhodná veličina X spojité rozdělení, pak platí

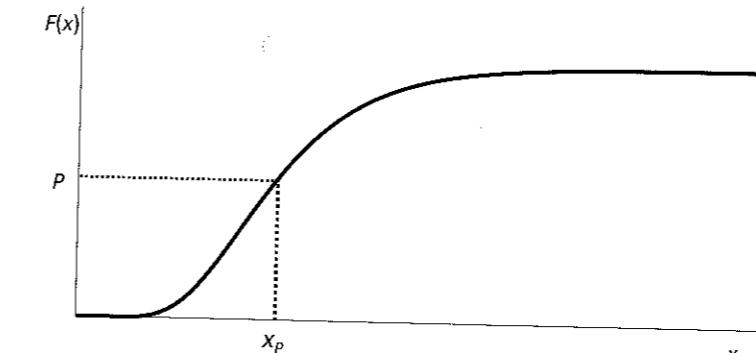
$$F(x_p) = P. \quad (2.52)$$

Je-li distribuční funkce náhodné veličiny X spojitá a navíc rostoucí ve všech bodech, je kvantil určen jednoznačně. S kvantily spojité náhodné veličiny se budeme setkávat zejména při testování hypotéz a konstrukci intervalů spolehlivosti.

Z definice kvantilu je zřejmé, že kvantil x_p je číslo, které dělí plochu pod křivkou hustoty v poměru P (vlevo od hodnoty kvantilu) a $1 - P$ (vpravo od hodnoty kvantilu). Celá situace je znázorněna na obrázku 2.5.



Obr. 2.5 Kvantil spojitého rozdělení – hustota pravděpodobnosti



Obr. 2.6 Kvantil spojitého rozdělení – distribuční funkce

Na obrázku 2.6 je zobrazena distribuční funkce a kvantil x_p . Je vidět, že funkční hodnota distribuční funkce v bodě kvantilu x_p je rovna P .

Některé kvantily mají speciální názvy

hodnota P	kvantil	název
0,5	$x_{0,5}$	medián
0,25	$x_{0,25}$	dolní kvartil
0,75	$x_{0,75}$	horní kvartil
0,1; ...; 0,9	$x_{0,1}, \dots, x_{0,9}$	decily
0,01; ...; 0,99	$x_{0,01}, \dots, x_{0,99}$	percentily

Podívejme se, jaká je hodnota mediánu pro vybrané nespojitě a spojité rovnoměrné rozdělení. Nejprve uvažujme rozdělení náhodné veličiny, představující výsledek hodu hrací kostkou. Její pravděpodobnostní funkce je dána vzorcem (2.29) a distribuční funkce je dána vzorcem (2.30). Je zřejmé, že definici mediánu vyhovuje libovolné číslo $x \in (3,4)$, neboť pro takovouto hodnotu platí

$$F(x+0) \geq 0,5, \quad F(x) \leq 0,5. \quad (2.53)$$

Nyní se podívejme na medián spojitého rovnoměrného rozdělení. Distribuční funkce tohoto rozdělení na intervalu (α, β) má tvar

$$\begin{aligned} F(x) &= 0, & x \leq \alpha, \\ &= \frac{x-\alpha}{\beta-\alpha}, & \alpha < x < \beta, \\ &= 1, & x \geq \beta. \end{aligned} \quad (2.54)$$

Vzhledem k tomu, že platí $F(x_{0,5}) = 0,5$, dostaneme medián z rovnosti

$$\frac{x_{0,5}-\alpha}{\beta-\alpha} = 0,5, \quad (2.55)$$

a tudíž

$$x_{0,5} = \frac{\alpha+\beta}{2}. \quad (2.56)$$

Hodnoty kvantilů pro spojitu náhodnou veličinu můžeme snadno určit – stačí si uvědomit, že k distribuční funkci existuje funkce inverzní. Je tedy

$$x_p = F^{-1}(P), \quad 0 < P < 1. \quad (2.57)$$

Modus

Jako charakteristika polohy se používá i tzv. **modus**. Značí se \hat{x} a je to taková hodnota, pro kterou platí

a) pro diskrétní náhodnou veličinu

$$P(X=\hat{x}) \geq P(X=x_j), \quad j=1, 2, \dots \quad (2.58)$$

b) pro spojitu náhodnou veličinu

$$f(\hat{x}) \geq f(x), \quad -\infty < x < \infty. \quad (2.59)$$

Jde tedy o hodnotu, ve které nabývá u diskrétní náhodné veličiny pravděpodobnostní funkce maxima nebo u spojité náhodné veličiny nabývá maxima hustota pravděpodobnosti. Z definice je zřejmé, že podobně jako kvantil, ani modus není určen jednoznačně.

Charakteristiky založené na kvantilech

Prestože kvantily jsou používány jako charakteristiky polohy, je možné pomocí nich sestrojit i charakteristiky variability. Nejčastěji se používá

mezikvartilové rozpětí

$$x_{0,75} - x_{0,25},$$

mezidecilové rozpětí

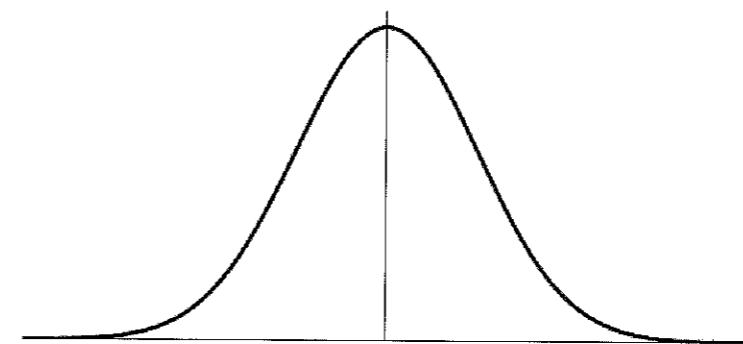
$$x_{0,90} - x_{0,10},$$

mezipercentilové rozpětí

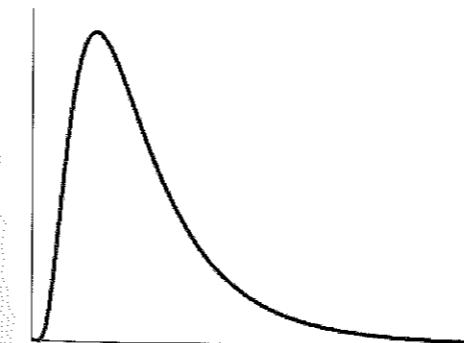
$$x_{0,99} - x_{0,01}.$$

Šíkmost

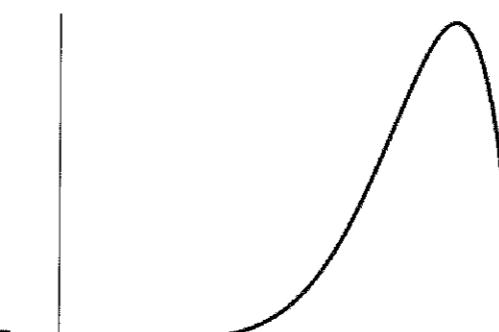
Charakteristiky šíkmosti popisují symetrii či asymetrii pravděpodobnostního rozdělení náhodné veličiny X . Různě sešikmená rozdělení jsou znázorněna na obrázcích 2.7-2.9.



Obr. 2.7 Nezešikmené (symetrické) rozdělení



Obr. 2.8 Kladně zešikmené rozdělení



Obr. 2.9 Záporně zešikmené rozdělení

Jako charakteristika šíkmosti se nejčastěji používá **koeficient šíkmosti** ve tvaru

$$\alpha_3(X) = \frac{\mu_3(X)}{[D(X)]^{3/2}}. \quad (2.60)$$

Špičatost

Jako charakteristika špičatosti se používá tzv. **koeficient špičatosti**

$$\alpha_4(X) = \frac{\mu_4(X)}{[D(X)]^2} - 3 = E\left[\frac{(X - E(X))^4}{\sqrt{D(X)}}\right] - 3. \quad (2.61)$$

Hodnota -3 není ve vzorci náhodou. Špičatost rovnou 0 (měřeno tímto vzorcem) má totiž **normální rozdělení** – viz dále v přehledu rozdělení. To znamená, že hodnota uvedeného zlomku ve výrazu je rovna 3 . Pokud tedy od zlomku číslo 3 odečteme, do staneme nulovou špičatost pro normální rozdělení, které tedy sehrává jakousi roli normy ohledně špičatosti – normy proto, že se jedná o nejznámější rozdělení, které zná každý, kdo se pravděpodobností zabývá. Pokud tedy má náhodná veličina X symetrické rozdělení a $\alpha_4(X) > 0$ (resp. $\alpha_4(X) < 0$), znamená to, že na koncích rozdělení je pravděpodobnostní funkce $P(X=x)$ či hustota pravděpodobnosti $f(x)$ větší (resp. menší), než hustota pravděpodobnosti normálního rozdělení se stejnou střední hodnotou a stejným rozptylem.

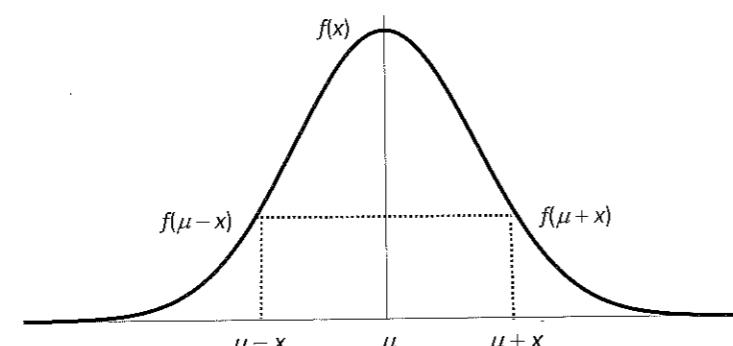
2.2.5 Symetrické rozdělení

V předchozím textu jsme hovořili o symetrii či asymetrii pravděpodobnostního rozdělení náhodné veličiny X , ale přesně jsme tyto pojmy zatím nevymezili. Učiníme tak v této kapitole. Budeme vycházet z práce autorů Likeš, Machek (1991).

Řekneme, že náhodná veličina X má **spojité rozdělení symetrické podle bodu μ** , jestliže pro její hustotu pravděpodobnosti $f(x)$ platí

$$f(\mu - x) = f(\mu + x), \quad -\infty < x < \infty. \quad (2.62)$$

Ukázka symetrického rozdělení je na obrázku 2.10.



Obr. 2.10 Hustota pravděpodobnosti symetrického rozdělení

Není těžké spočítat, že pro distribuční funkci platí

$$F(x) = 1 - F(2\mu - x), \quad -\infty < x < \infty. \quad (2.63)$$

Pro $100P\%$ kvantil x_p , $0 < P < 1$, tak můžeme psát

$$F(x_p) = P = 1 - F(2\mu - x_p), \quad 0 < P < 1, \quad (2.64)$$

a odtud

$$x_{1-P} = 2\mu - x_p. \quad (2.65)$$

Položíme-li $P = 0,5$, dostaneme $x_{0,5} = \mu$. Tedy platí, že medián je roven bodu symetrie μ . Totéž platí i pro střední hodnotu

$$E(X) = \mu. \quad (2.66)$$

To znamená, že pro symetrické rozdělení je střední hodnota rovna bodu symetrie μ .

2.3 Náhodný vektor

Doposud jsme pracovali s jednorozměrnou náhodnou veličinou – výsledkem náhodného pokusu, který jsme chápali jako hodnotu náhodné veličiny, bylo jedno reálné číslo. V mnoha situacích (tj. náhodných pokusech) je však výsledkem náhodného pokusu celá p -tice čísel. Jako příklad uvedeme cenu akcie za jeden den obchodování na burze – můžeme sledovat cenu při otevření burzy, cenu při uzavření, nejvyšší cenu během dne, nejnižší cenu během dne apod. U pacienta, který podstoupí vyšetření u lékaře, může být sledována hmotnost, výška, teplota a krevní tlak.

Ve všech těchto situacích nepracujeme s jednou náhodnou veličinou, ale s p náhodnými veličinami X_1, \dots, X_p . Tuto p -tici označíme jako $X = (X_1, \dots, X_p)'$ a budeme ji nazývat **náhodným vektorem** nebo **p -rozměrnou náhodnou veličinou**.

Obdobně jako v jednorozměrném případě budeme pracovat s **pravděpodobnostním rozdělením náhodného vektora** – viz Likeš, Machek (1991). Pokud je tedy dán vztah mezi podmnožinami p -rozměrného euklidovského prostoru R^p a jejich pravděpodobnostmi, pracujeme s tzv. **sdruženým rozdělením pravděpodobnosti** náhodných veličin X_1, \dots, X_p .

2.3.1 Distribuční funkce

Distribuční funkci náhodného vektoru $X = (X_1, \dots, X_p)'$ budeme rozumět reálnou funkci $F(x_1, \dots, x_p)$, definovanou jako

$$F(x_1, \dots, x_p) = P(X_1 \leq x_1, \dots, X_p \leq x_p), \quad \forall x_1, \dots, x_p \in R. \quad (2.67)$$

Distribuční funkce $F(x_1, \dots, x_p)$ v bodech x_1, \dots, x_p vyjadřuje pravděpodobnost, že náhodná veličina X_1 nabude hodnoty menší nebo rovné x_1 a zároveň náhodná veličina X_2 nabude hodnoty menší nebo rovné x_2 , až náhodná veličina X_p nabude hodnoty menší nebo rovné x_p . Distribuční funkci $F(x_1, \dots, x_p)$ náhodného vektoru $X = (X_1, \dots, X_p)'$ označujeme jako **sdruženou distribuční funkci náhodných veličin** X_1, \dots, X_p . Analogicky jako v jednorozměrném případě má sdružená distribuční funkce určité vlastnosti, které jsou zobrazením vlastností distribuční funkce jednorozměrné náhodné veličiny. Položíme-li totiž $p = 1$, pracujeme s jednorozměrnou náhodnou veličinou.

2.3.2 Rozdělení diskrétního a spojitého typu

Podobně jako v jednorozměrném případě budeme i pro náhodný vektor rozlišovat vícerozměrné rozdělení diskrétního a spojitého typu. Nejprve se budeme zabývat rozdělením diskrétního typu. V následujících definicích vycházíme z Likeš, Machek (1991).

Rozdělení diskrétního typu

Náhodný vektor $X = (X_1, \dots, X_p)'$ má **rozdělení diskrétního typu**, existuje-li konečná nebo spočetná množina p -tic reálných čísel x_1, \dots, x_p taková, že pro každý prvek této množiny je pravděpodobnost $P(X_1 = x_1, \dots, X_p = x_p) > 0$ a součet pravděpodobností pro všechny prvky této množiny je roven jedné,

$$\sum_{x_1} \dots \sum_{x_p} P(X_1 = x_1, \dots, X_p = x_p) = 1. \quad (2.68)$$

Funkce $P(X_1 = x_1, \dots, X_p = x_p)$ se nazývá pravděpodobnostní funkce náhodného vektora X nebo o ní hovoříme jako o sdružené pravděpodobnostní funkci náhodných veličin X_1, \dots, X_p . Zkráceně ji budeme značit $P(\mathbf{x}) = P(x_1, \dots, x_p)$, kde $\mathbf{x} = (x_1, \dots, x_p)'$.

Stejně jako pro jednu diskrétní náhodnou veličinu i pro diskrétní náhodný vektor $X = (X_1, \dots, X_p)'$ existuje vzájemně jednoznačný vztah mezi distribuční funkcí a pravděpodobnostní funkcí, který lze zapsat ve tvaru

$$F(x_1, \dots, x_p) = \sum_{t_1 \leq x_1} \dots \sum_{t_p \leq x_p} P(X_1 = t_1, \dots, X_p = t_p). \quad (2.69)$$

Pro $p = 2$ pak můžeme psát

$$P(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2) = \sum_{a_1 < t_1 \leq b_1} \sum_{a_2 < t_2 \leq b_2} P(X_1 = t_1, X_2 = t_2), \quad (2.70)$$

pro všechny $P(X_1 = t_1, X_2 = t_2) > 0$. Distribuční funkci p -rozměrného náhodného vektora budeme zkráceně zapisovat jako $F(\mathbf{x})$, kde $F(\mathbf{x}) = F(x_1, \dots, x_p)$.

Příklad 2.13

V tabulce 2.2 jsou uvedeny hodnoty, kterých nabývá dvourozměrný náhodný vektor $X = (X_1, X_2)'$ a sdružené pravděpodobnosti $P(X_1 = x_1, X_2 = x_2)$. Na těchto údajích budeme ilustrovat výše uvedenou teorii.

Tab. 2.2 Sdružené pravděpodobnosti

$x_1 \backslash x_2$	1	2	3	$P(X_1 = x_1)$
0	0,28	0,21	0,09	0,58
1	0,18	0,07	0,17	0,42
$P(X_2 = x_2)$	0,46	0,28	0,26	1

Hodnoty pravděpodobnostní a distribuční funkce ve vybraných bodech (x_1, x_2)

$$\begin{aligned} P(1,1) &= P(X_1 = 1, X_2 = 1) = 0,18, & F(1,1) &= P(X_1 \leq 1, X_2 \leq 1) = 0,46, \\ P(2,1) &= P(X_1 = 2, X_2 = 1) = 0, & F(2,1) &= P(X_1 \leq 2, X_2 \leq 1) = 0,46, \\ P(0,3) &= P(X_1 = 0, X_2 = 3) = 0,09, & F(0,3) &= P(X_1 \leq 3, X_2 \leq 3) = 0,58, \\ P(4,4) &= P(X_1 = 4, X_2 = 4) = 0, & F(4,4) &= P(X_1 \leq 4, X_2 \leq 4) = 1, \\ P(-1,2) &= P(X_1 = -1, X_2 = 2) = 0, & F(-1,2) &= P(X_1 \leq -1, X_2 \leq 2) = 0, \\ P(X_1 < 1, 1 \leq X_2 \leq 2) &= 0,49, & F(2,5; 1) &= P(X_1 \leq 2,5, X_2 \leq 1) = 0,46, \\ P(X_1 < 1, 1 < X_2 \leq 3) &= 0,24. & & \end{aligned}$$

Rozdělení spojitého typu

Náhodný vektor $X = (X_1, \dots, X_p)'$ má **rozdělení spojitého typu**, existuje-li nezáporná reálná funkce $f(x_1, \dots, x_p)$ taková, že pro všechna reálná x_1, \dots, x_p lze distribuční funkci $F(x_1, \dots, x_p)$ vyjádřit ve tvaru

$$F(x_1, \dots, x_p) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_p} f(t_1, \dots, t_p) dt_1 \dots dt_p. \quad (2.71)$$

Funkce $f(x_1, \dots, x_p)$ se nazývá hustota pravděpodobnosti náhodného vektora X nebo též sdružená hustota pravděpodobnosti náhodných veličin X_1, \dots, X_p . Většinou mluvíme zkráceně pouze o sdružené hustotě.

Ve všech bodech, kde derivace existuje, platí následující vztah mezi hustotou a distribuční funkcí

$$f(x_1, \dots, x_p) = \frac{\partial^p F(x_1, \dots, x_p)}{\partial x_1 \dots \partial x_p}. \quad (2.72)$$

To znamená, že ve všech bodech, kde derivace existuje, získáme hustotu jako parciální derivaci distribuční funkce podle všech proměnných. Je zřejmé, že nezáleží na pořadí proměnných, podle kterých distribuční funkci derivujeme.

Z definice distribuční funkce dostáváme

$$P(a_1 < X_1 \leq b_1, \dots, a_p < X_p \leq b_p) = \int_{a_1}^{b_1} \dots \int_{a_p}^{b_p} f(x_1, \dots, x_p) dx_1 \dots dx_p \quad (2.73)$$

pro každé $-\infty < a_j \leq b_j < \infty, j = 1, 2, \dots, p$.

Dále platí

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_p) dx_1 \dots dx_p = 1. \quad (2.74)$$

Příklad 2.14

Mějme dvourozměrný náhodný vektor $\mathbf{X} = (X_1, X_2)'$ s distribuční funkcí

$$\begin{aligned} F(x_1, x_2) &= (1 - e^{-x_1})(1 - e^{-x_2}), \quad x_1 > 0, x_2 > 0, \\ &= 0, \quad \text{jinak.} \end{aligned}$$

Určete hustotu pravděpodobnosti $f(x_1, x_2)$.

Řešení

Najdeme parciální derivace distribuční funkci $F(x_1, x_2)$ podle obou proměnných a obdržíme

$$f(x_1, x_2) = \frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2} = \frac{\partial^2 (1 - e^{-x_1})(1 - e^{-x_2})}{\partial x_1 \partial x_2} = \frac{\partial e^{-x_1}(1 - e^{-x_2})}{\partial x_2} = e^{-x_1} e^{-x_2}.$$

Poté po úpravě

$$\begin{aligned} f(x_1, x_2) &= e^{-x_1 - x_2}, \quad x_1 > 0, x_2 > 0, \\ &= 0, \quad \text{jinak.} \end{aligned}$$

2.3.3 Marginální rozdělení

Marginální rozdělení budeme nejprve definovat pro případ dvourozměrného náhodného vektoru $\mathbf{X} = (X_1, X_2)'$ a poté definici zobecníme pro p -rozměrný případ. Pro náhodný vektor $\mathbf{X} = (X_1, X_2)'$ jsme definovali sdruženou distribuční funkci, která popisovala pravděpodobnostní chování celého vektoru (v tomto případě obou náhodných veličin X_1 a X_2 najednou). Často nás však zajímá pravděpodobnostní chování jednotlivých složek náhodného vektoru. To v našem případě znamená chování náhodné veličiny X_1 a X_2 . Pokud tedy pracujeme s pravděpodobnostním rozdělením jednotlivých složek náhodného vektoru, hovoříme o **marginálních rozděleních**.

Nechť náhodný vektor $\mathbf{X} = (X_1, X_2)'$ má sdruženou distribuční funkci $F(x_1, x_2)$. Pak **marginální distribuční funkce** $F_1(x_1)$ náhodné veličiny X_1 je rovna

$$F_1(x_1) = P(X_1 \leq x_1) = \lim_{x_2 \rightarrow \infty} P(X_1 \leq x_1, X_2 \leq x_2) = \lim_{x_2 \rightarrow \infty} F(x_1, x_2), \quad -\infty < x_1 < \infty. \quad (2.75)$$

Zcela analogicky obdržíme marginální distribuční funkci $F_2(x_2)$.

Obdobný vztah jako pro marginální distribuční funkci obdržíme pro pravděpodobnostní funkci v případě diskrétního náhodného vektoru a pro hustotu pravděpodobnosti v případě spojitého náhodného vektoru.

Nechť tedy náhodný vektor $\mathbf{X} = (X_1, X_2)'$ má rozdělení diskrétního typu a nechť $P(X_1 = x_1, X_2 = x_2)$ je sdružená pravděpodobnostní funkce veličin X_1 a X_2 . Pak marginální pravděpodobnostní funkce náhodné veličiny X_1 má tvar

$$P(X_1 = x_1) = \sum_{x_2} P(X_1 = x_1, X_2 = x_2). \quad (2.76)$$

Pravděpodobnost $P(X_1 = x_1)$ zapisujeme zkráceně jako $P_1(x_1)$. To znamená, že marginální pravděpodobnostní funkci náhodné veličiny X_1 získáme tak, že sečteme sdruženou pravděpodobnostní funkci přes všechny hodnoty x_2 (tedy té náhodné veličiny, kterou „nechceme“).

Obdobně marginální pravděpodobnostní funkce náhodné veličiny X_2 má tvar

$$P(X_2 = x_2) = \sum_{x_1} P(X_1 = x_1, X_2 = x_2). \quad (2.77)$$

Pro náhodný vektor $\mathbf{X} = (X_1, X_2)'$, který má rozdělení spojitého typu se sdruženou hustotou $f(x_1, x_2)$, je marginální hustota pravděpodobnosti náhodné veličiny X_1 rovna

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_2, \quad -\infty < x_1 < \infty. \quad (2.78)$$

Marginální hustotu pravděpodobnosti náhodné veličiny X_1 tedy získáme tak, že sdruženou hustotu integrujeme přes hodnoty náhodné veličiny X_2 (tedy veličiny, kterou „nechceme“). Obdobně

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_1, \quad -\infty < x_2 < \infty. \quad (2.79)$$

Není těžké provést zobecnění pro p -rozměrný případ (viz Marek, 2012).

2.3.4 Charakteristiky náhodného vektoru

Stejně jako tomu bylo u jedné náhodné veličiny X , budou nás i u náhodného vektoru $\mathbf{X} = (X_1, \dots, X_p)'$ zajímat číselné charakteristiky. Ty, jak již víme, shrnují v několika málo číslech informace o pravděpodobnostním chování náhodné veličiny a tedy i náhodného vektoru. Nejdůležitějšími charakteristikami náhodné veličiny X byly střední hodnota $E(X)$ a rozptyl $D(X)$. Obdobně tomu bude i u náhodného vektoru. Pro náhodný vektor $\mathbf{X} = (X_1, \dots, X_p)'$ budeme definovat jeho střední hodnotu jako vektor středních hodnot $[E(X_1), \dots, E(X_p)]$. U rozptylu je situace složitější. Pro náhodný vektor \mathbf{X} používáme jako charakteristiku **kovarianční matici**. Ta má tvar

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \dots & \dots & \dots & \dots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{pmatrix}, \quad (2.80)$$

kde

$$\sigma_{jk} = \text{cov}(X_j, X_k) = E\{[X_j - E(X_j)][X_k - E(X_k)]\}, \quad j, k = 1, 2, \dots, p. \quad (2.81)$$

Matice Σ je čtvercová matice typu $p \times p$, je semidefinitní a symetrická podle hlavní diagonály, neboť

$$\text{cov}(X_j, X_k) = \text{cov}(X_k, X_j), \quad j, k = 1, 2, \dots, p, \quad (2.82)$$

a tedy $\sigma_{jk} = \sigma_{kj}$.

Není těžké upravit kovarianci do podoby, kterou používáme k výpočtům

$$\text{cov}(X_j, X_k) = E(X_j X_k) - E(X_j)E(X_k), \quad j, k = 1, 2, \dots, p. \quad (2.83)$$

Položíme-li $j = k$, obdržíme

$$\text{cov}(X_j, X_j) = D(X_j), \quad j = 1, 2, \dots, p. \quad (2.84)$$

Tzn., že na hlavní diagonále kovarianční matic se nachází rozptyly $D(X_1), \dots, D(X_p)$ náhodných veličin X_1, \dots, X_p .

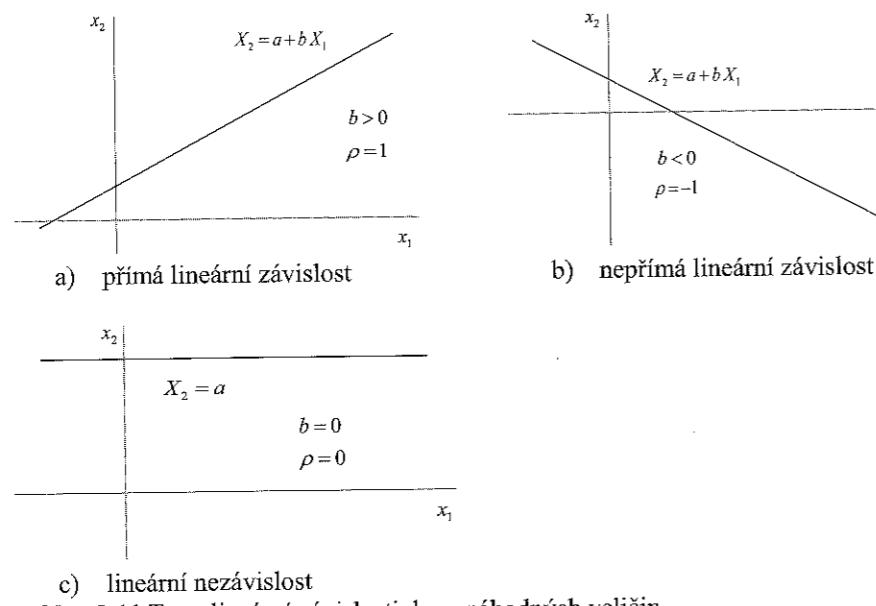
Kovarianci lze upravit do tvaru korelačního koeficientu náhodných veličin X_j a X_k

$$\rho(X_j, X_k) = \frac{\text{cov}(X_j, X_k)}{\sqrt{D(X_j)D(X_k)}}, \quad j, k = 1, 2, \dots, p, \quad (2.85)$$

pro $D(X_j) > 0, D(X_k) > 0$. Korelační koeficient se používá jako míra lineární závislosti náhodných veličin. Je to jedna z nejpoužívanějších charakteristik v nejrůznějších statistických metodách.

Korelační koeficient nabývá hodnot v intervalu $(-1; 1)$. Je-li kovariance v čitateli vzorce korelačního koeficientu rovna nule, je roven nule i korelační koeficient a hovoříme o tom, že náhodné veličiny jsou **nekorelované** (tj. lineárně nezávislé).

Již bylo řečeno, korelační koeficient se používá jako míra lineární závislosti. Je-li konstanta $b > 0$, hovoříme o **přímé lineární závislosti** („rostе jedna veličina, pak roste i druhá veličina“), v případě $b < 0$, jde o **nepřímou lineární závislost** („rostе jedna veličina, pak druhá veličina klesá“). Celá situace je ukázána na obrázcích 2.11a)- c).



Obr. 2.11 Typy lineární závislosti dvou náhodných veličin

Důležitou charakteristikou náhodného vektoru je **korelační matic**

$$\rho(\mathbf{X}) = \begin{pmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1p} \\ \dots & \dots & \dots & \dots \\ \rho_{p1} & \rho_{p2} & \dots & \rho_{pp} \end{pmatrix}. \quad (2.86)$$

Jedná se o čtvercovou matici typu $p \times p$, kde se na místě (i, j) nachází korelační koeficient $\rho(X_i, X_j)$, $i, j = 1, 2, \dots, p$. Matice je symetrická podle hlavní diagonály. Na ní se nachází korelační koeficienty

$$\rho(X_j, X_j) = 1, \quad j = 1, 2, \dots, p. \quad (2.87)$$

To znamená, že korelační matice má na hlavní diagonále samé jedničky

$$\rho(\mathbf{X}) = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1p} \\ \dots & \dots & \dots & \dots \\ \rho_{p1} & \rho_{p2} & \dots & 1 \end{pmatrix}. \quad (2.88)$$

Jsou-li všechny kovariance nulové (veličiny X_1, \dots, X_p jsou nekorelované), platí

$$\rho(X_j, X_k) = 0, \quad j \neq k = 1, 2, \dots, p. \quad (2.89)$$

Kovarianční matice je pak **jednotkovou maticí**

$$\rho(\mathbf{X}) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}. \quad (2.90)$$

2.3.5 Nezávislost náhodných veličin

Nejprve budeme definovat nezávislost pro dvě náhodné veličiny. Uvažujme dvouzdrobný náhodný vektor $\mathbf{X} = (X_1, X_2)'$ se sdruženou distribuční funkcí $F(x_1, x_2)$. **Náhodné veličiny X_1 a X_2 jsou nezávislé** právě tehdy, když

$$F(x_1, x_2) = F_1(x_1) F_2(x_2) \quad (2.91)$$

pro všechna reálná x_1 a x_2 . Jinými slovy, náhodné veličiny X_1 a X_2 jsou nezávislé právě tehdy, když jejich sdružená distribuční funkce je rovna součinu jejich marginálních distribučních funkcí.

Dále pro nezávislé náhodné veličiny platí

$$P(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2) = P(a_1 < X_1 \leq b_1) P(a_2 < X_2 \leq b_2). \quad (2.92)$$

Pro náhodný vektor $\mathbf{X} = (X_1, X_2)'$, který se řídí rozdělením diskrétního typu, platí, že náhodné veličiny X_1 a X_2 jsou nezávislé právě tehdy, když

$$P(X_1 = x_1, X_2 = x_2) = P_1(X_1 = x_1) P_2(X_2 = x_2) \quad (2.93)$$

pro všechna reálná x_1 a x_2 . Platí tedy, že diskrétní náhodné veličiny X_1 a X_2 jsou nezávislé právě tehdy, když jejich sdružená pravděpodobnostní funkce je rovna součinu jejich marginálních pravděpodobnostních funkcí.

Podobně pro náhodný vektor $\mathbf{X} = (X_1, X_2)'$, který se řídí rozdělením spojitého typu, platí, že náhodné veličiny X_1 a X_2 jsou nezávislé právě tehdy, když

$$f(x_1, x_2) = f_1(x_1) f_2(x_2) \quad (2.94)$$

pro všechna reálná x_1 a x_2 . Platí tedy, že náhodné veličiny X_1 a X_2 jsou nezávislé právě tehdy, když jejich sdružená hustota pravděpodobnosti je rovna součinu jejich marginálních hustot.

Z nezávislosti náhodných veličin X_1 a X_2 vyplývají některé další vztahy, které nyní uvedeme. Předpokládejme tedy nezávislost náhodných veličin X_1 a X_2 . Pak platí

a) $E(X_1^{r_1} X_2^{r_2}) = E(X_1^{r_1}) E(X_2^{r_2})$,

pokud uvedené střední hodnoty existují (r_1 a r_2 jsou přirozená čísla). Z uvedeného vztahu plyne i rovnost $E(X_1 X_2) = E(X_1) E(X_2)$.

b) $\text{cov}(X_1, X_2) = 0$.

Nezávislé náhodné veličiny jsou tedy nekorelované.

c) $D(X_1 + X_2) = D(X_1) + D(X_2)$.

Výše uvedené úvahy lze zobecnit pro náhodný vektor $\mathbf{X} = (X_1, \dots, X_p)'$, kde $p \geq 3$.

2.3.6 Podmíněné rozdělení

Nejprve budeme definovat podmíněné rozdělení pro dvě náhodné veličiny. Začneme případem diskrétního rozdělení.

Uvažujme dvourozměrný náhodný vektor $\mathbf{X} = (X_1, X_2)'$, řídící se rozdělením diskrétního typu se sdruženou pravděpodobnostní funkcí $P(X_1 = x_1, X_2 = x_2)$. **Podmíněnou pravděpodobnostní funkci náhodné veličiny X_1 za podmínky, že veličina X_2 nabyla hodnoty x_2** , definujeme jako

$$P(X_1 = x_1 | X_2 = x_2) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_2 = x_2)}, \quad P(X_2 = x_2) > 0. \quad (2.95)$$

Jinými slovy, podmíněná pravděpodobnostní funkce náhodné veličiny X_1 za podmínky, že veličina X_2 nabyla hodnoty x_2 , je rovna podílu sdružené pravděpodobnostní funkce veličin X_1 a X_2 a marginální pravděpodobnostní funkce veličiny X_2 .

Zcela analogicky můžeme vyjádřit podmíněnou pravděpodobnostní funkci náhodné veličiny X_2 za podmínky, že veličina X_1 nabyla hodnoty x_1 , jako

$$P(X_2 = x_2 | X_1 = x_1) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_1 = x_1)}, \quad P(X_1 = x_1) > 0. \quad (2.96)$$

Opět není těžké zobecnit uvedená tvrzení pro $p > 2$ – viz Marek (2012).

2.4 Vybraná diskrétní rozdělení

2.4.1 Alternativní rozdělení

Řekneme, že náhodná veličina X má alternativní rozdělení s parametrem π , jestliže její pravděpodobnostní funkce má tvar

$$\begin{aligned} P(x) &= \pi^x (1-\pi)^{1-x}, & x = 0, 1, \\ &= 0, & \text{jinak,} \end{aligned} \quad (2.97)$$

kde $0 < \pi < 1$. Náhodná veličina X tedy nabývá pouze dvou hodnot 0 a 1. Parametr π pak představuje pravděpodobnost nastoupení určitého sledovaného jevu (pro $x = 1$, pravda, úspěch), hodnota $1 - \pi$ pak představuje pravděpodobnost, že k nastoupení sledovaného jevu nedojde (pro $x = 0$, nepravda, neúspěch). To, že se náhodná veličina řídí alternativním rozdělením, budeme symbolicky zapisovat $X \sim A(\pi)$.

Momenty

Střední hodnota náhodné veličiny X s rozdělením $A(\pi)$ je

$$E(X) = \pi, \quad (2.98)$$

rozptyl je roven

$$D(X) = \pi(1 - \pi). \quad (2.99)$$

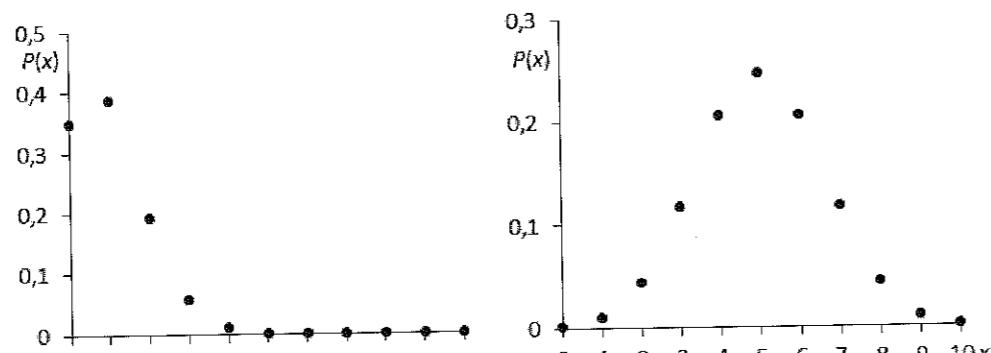
2.4.2 Binomické rozdělení

Náhodná veličina X má binomické rozdělení s parametry n a π , jestliže její pravděpodobnostní funkce má tvar

$$\begin{aligned} P(x) &= \binom{n}{x} \pi^x (1-\pi)^{n-x}, & x = 0, 1, \dots, n, \\ &= 0, & \text{jinak,} \end{aligned} \quad (2.100)$$

kde $0 < \pi < 1$ a $n \in \mathbb{N}$. Parametr π (stejně jako u alternativního rozdělení) představuje pravděpodobnost nastoupení určitého sledovaného jevu v jednom pokusu, náhodná veličina X pak představuje počet nastoupení sledovaného jevu v n nezávislých pokusech. Tímto rozdělením se tedy řídí náhodná veličina X , představující počet výskytů sledovaného jevu v n nezávislých pokusech, jestliže pravděpodobnost výskytu tohoto jevu je v každém pokusu stejná a je rovna témuž číslu π (výběr s vrácením v případě opakování pokusů). To, že se náhodná veličina řídí binomickým rozdělením, budeme symbolicky zapisovat $X \sim Bi(n, \pi)$. Název rozdělení pochází z jeho vlastnosti – pravděpodobnost $P(x)$ je totiž obecným členem binomického rozvoje $[(1-\pi) + \pi]^n$.

Podívejme se nyní na obrázky 2.12 a 2.13, které zobrazují průběh pravděpodobnostní funkce pro různé parametry n a π .

Obr. 2.12 Pravděp. funkce rozdělení $Bi(10; 0,1)$ Obr. 2.13 Pravděp. funkce rozdělení $Bi(10; 0,5)$

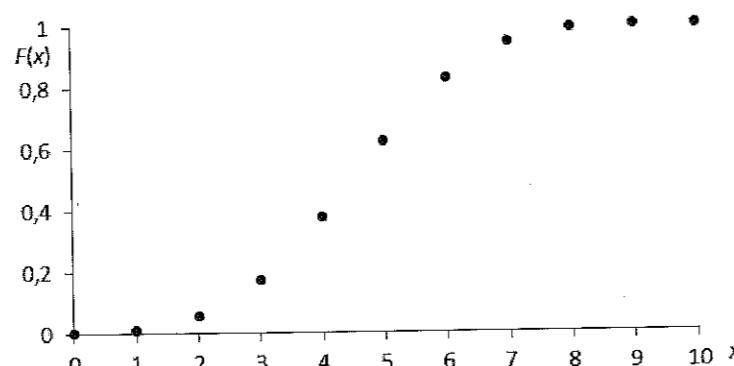
Distribuční funkce

Distribuční funkce $F(x) = P(X \leq x)$, $x \in R$ binomického rozdělení má tvar

$$\begin{aligned} F(x) &= 0, & x < 0, \\ &= \sum_{i=0}^{\lfloor x \rfloor} \binom{n}{i} \pi^i (1-\pi)^{n-i}, & 0 \leq x < n, \\ &= 1, & x \geq n, \end{aligned} \quad (2.101)$$

kde $\lfloor x \rfloor$ je celá část čísla x . Hodnoty distribuční funkce, případně hodnoty pravděpodobnostní funkce, lze snadno získat např. v MS Excel, který obsahuje celou řadu funkcí pro nespojitá i spojité rozdělení.

Např. pro $n = 10$ a $\pi = 0,5$ má distribuční funkce tvar uvedený na obrázku 2.14.

Obr. 2.14 Distribuční funkce rozdělení $Bi(10; 0,5)$

Momenty

Střední hodnota náhodné veličiny s rozdělením $Bi(n, \pi)$ je rovna

$$E(X) = n\pi, \quad (2.102)$$

rozptyl je roven

$$D(X) = n\pi(1 - \pi). \quad (2.103)$$

Vzhledem k tomu, že $0 < \pi < 1$, je $E(X) > (1 - \pi)E(X) = D(X)$, platí

$$E(X) > D(X). \quad (2.104)$$

2.4.3 Poissonovo rozdělení

Náhodná veličina X má Poissonovo rozdělení s parametrem λ , jestliže její pravděpodobnostní funkce má tvar

$$\begin{aligned} P(x) &= \frac{\lambda^x}{x!} e^{-\lambda}, & x = 0, 1, \dots, \\ &= 0, & \text{jinak,} \end{aligned} \quad (2.105)$$

kde $\lambda > 0$. Fakt, že se náhodná veličina řídí Poissonovým rozdělením, budeme symbolicky zapisovat $X \sim Po(\lambda)$.

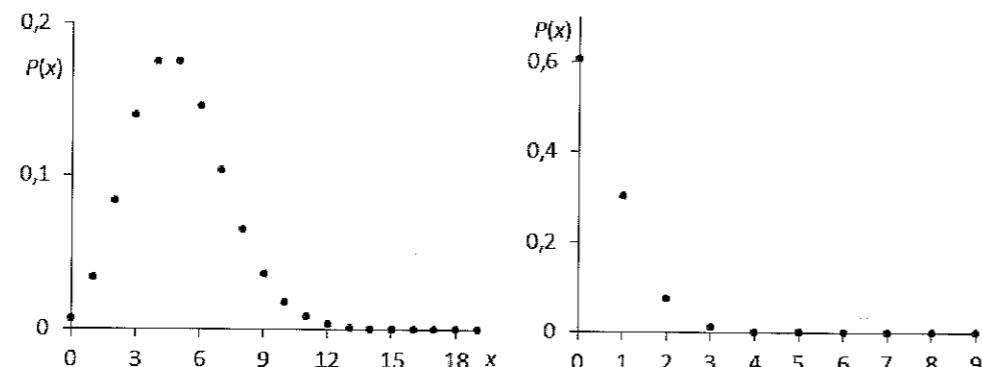
Tímto rozdělením se často řídí náhodná veličina představující počet výskytů sledovaného jevu za časovou nebo prostorovou jednotku, pokud tyto jevy nastávají náhodně v čase a nezávisle na sobě. Protože parametr λ je střední hodnotou, představuje hodnota λ střední počet výskytu sledovaného jevu.

Pokud náhodnou veličinou je počet výskytů sledovaného jevu v určitém časovém intervalu délky t , pak platí tyto skutečnosti – viz Hebká, Kahounová (1983):

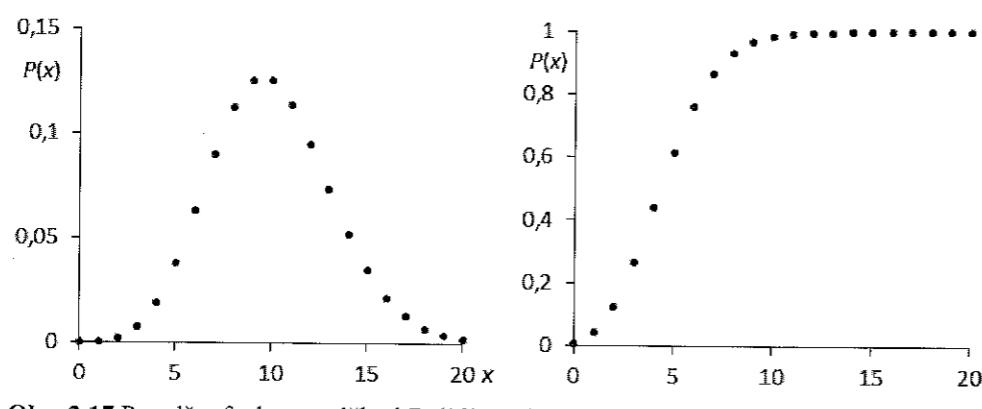
- Jev může nastat v kterémkoliv časovém okamžiku.
- Počet výskytu jevu během časového intervalu závisí pouze na délce tohoto intervalu, nikoliv na jeho počátku. Přitom nezávisí na tom, kolikrát sledovaný jev nastal před jeho počátkem.
- Pravděpodobnost, že jev nastane více než jednou v intervalu délky t , konverguje k nule rychleji než t .
- Je-li λ střední hodnota počtu výskytů sledovaného jevu za časovou jednotku, pak uvedená veličina má rozdělení $X \sim Po(t\lambda)$.

Tímto rozdělením se tedy řídí např. počet zákazníků ve frontě na pokladnu, počet odbavených cestujících za jednotku času, počet dopravních nehod na určité komunikaci za jeden pracovní den, počet signálů na telefonní ústředně za určitý časový interval, počet operací zpracovaných za jednu sekundu procesorem osobního počítače, počet pojistných událostí za jeden rok pro určitý druh pojistění, počet překlepů na jednu stranu textu atd. Již teď je zřejmé, že toto rozdělení bude sehrávat klíčovou roli v teorii hromadné obsluhy (společně s exponenciálním rozdělením tvoří základ pro tzv. Poissonův proces) a používá se tam, kde vznikají nároky na určitou službu náhodně v čase.

Na obrázcích 2.15 – 2.18 je vidět tvar pravděpodobnostní a distribuční funkce Poissonova rozdělení pro různé hodnoty parametru λ a jsou z nich na první pohled zřejmé i výše popsané skutečnosti.



Obr. 2.15 Pravděp. funkce rozdělení $Po(0,5)$ Obr. 2.16 Pravděp. funkce rozdělení $Po(5)$



Obr. 2.16 Pravděp. funkce rozdělení $Po(5)$ Obr. 2.17 Pravděp. funkce rozdělení $Po(10)$



Obr. 2.17 Pravděp. funkce rozdělení $Po(10)$ Obr. 2.18 Distrib. funkce rozdělení $Po(5)$



Obr. 2.18 Distrib. funkce rozdělení $Po(5)$



Obr. 2.19 Pravděp. funkce rozdělení $R(10)$ Obr. 2.20 Distribuční funkce rozdělení $R(10)$



Obr. 2.20 Distribuční funkce rozdělení $R(10)$



Obr. 2.19 Pravděp. funkce rozdělení $R(10)$ Obr. 2.20 Distribuční funkce rozdělení $R(10)$



Obr. 2.20 Distribuční funkce rozdělení $R(10)$

Momenty

Střední hodnota i rozptyl náhodné veličiny X s rozdělením $Po(\lambda)$ jsou rovny parametru λ

$$E(X) = D(X) = \lambda. \quad (2.106)$$

Aproximace binomického rozdělení rozdělením Poissonovým

V praxi se většinou používá approximace rozdělení $Bi(n, \pi)$ rozdělením $Po(n\pi)$, pokud je $\pi < 0,1$ a $n > 30$. Zároveň je třeba podotknout, že z výpočetního hlediska ztrácí v dnešní době approximace částečně smysl, neboť přesné hodnoty binomického rozdělení můžeme většinou bez problémů spočítat i bez specializovaných statistických pro-

gramů – snadno je lze spočítat např. v MS Excel. Nicméně teoretický význam aproximace přetravává a často se používá při nejrůznějších teoretických odvozeních a formálních důkazech.

2.4.4 Rovnoměrné rozdělení

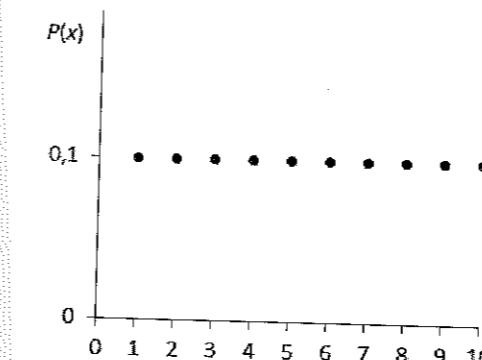
Náhodná veličina X má rovnoměrné diskrétní rozdělení s parametrem n , jestliže její pravděpodobnostní funkce má tvar

$$\begin{aligned} P(x) &= \frac{1}{n}, & x = 1, 2, \dots, n, \\ &= 0, & \text{jinak.} \end{aligned} \quad (2.107)$$

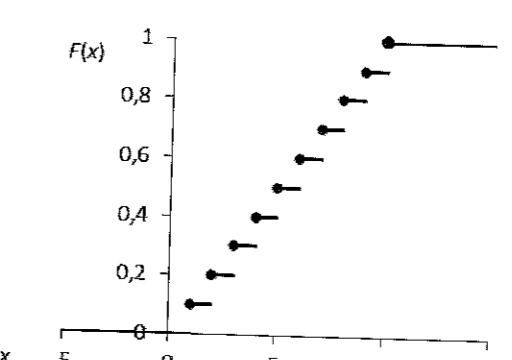
Fakt, že se náhodná veličina řídí rovnoměrným diskrétním rozdělením, budeme symbolicky zapisovat $X \sim R(n)$.

Jedná se o velmi jednoduché rozdělení. Jde vlastně o výběr jedné jednotky z konečné množiny jednotek, kdy každá jednotka má stejnou pravděpodobnost vybrání. Nejlépe ho můžeme ilustrovat na následujícím příkladu. Mějme osudí s n koulemi, z nichž každá má jinou barvu. Náhodně vybereme jednu kouli pravděpodobnost, že bude jedna určitá koule vybrána (např. koule červené barvy) je rovna $p = 1/n$.

Na obrázcích 2.19 a 2.20 vidíme pravděpodobnostní a distribuční funkci rovnoměrného rozdělení pro $n = 10$.



Obr. 2.19 Pravděp. funkce rozdělení $R(10)$



Obr. 2.20 Distribuční funkce rozdělení $R(10)$

Momenty

Střední hodnota náhodné veličiny s rozdělením $R(n)$ je

$$E(X) = \frac{n+1}{2}, \quad (2.108)$$

rozptyl je roven

$$D(X) = \frac{1}{n} \sum_{i=1}^n i^2 - \left(\frac{n+1}{2} \right)^2. \quad (2.109)$$

2.4.5 Hypergeometrické rozdělení

Náhodná veličina X má hypergeometrické rozdělení s parametry N, M a n , jestliže její pravděpodobnostní funkce má tvar

$$P(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}, \quad x = \max(0, M-N+n), \dots, \min(M, n). \quad (2.110)$$

Přitom N, M a n jsou přirozená čísla, $1 \leq n < N$, $1 \leq M < N$. Fakt, že se náhodná veličina řídí hypergeometrickým rozdělením, budeme symbolicky zapisovat $X \sim Hg(M, N, n)$.

Pro lepší pochopení si představme následující model. Mějme osudí s N koulemi, z toho jich M má určitou sledovanou vlastnost (např. jsou červené). Provedeme náhodný výběr bez vracení o velikosti n a budeme zkoumat, s jakou pravděpodobností se ve výběru vyskytuje x koulí se sledovanou vlastností (x červených koulí).

Momenty

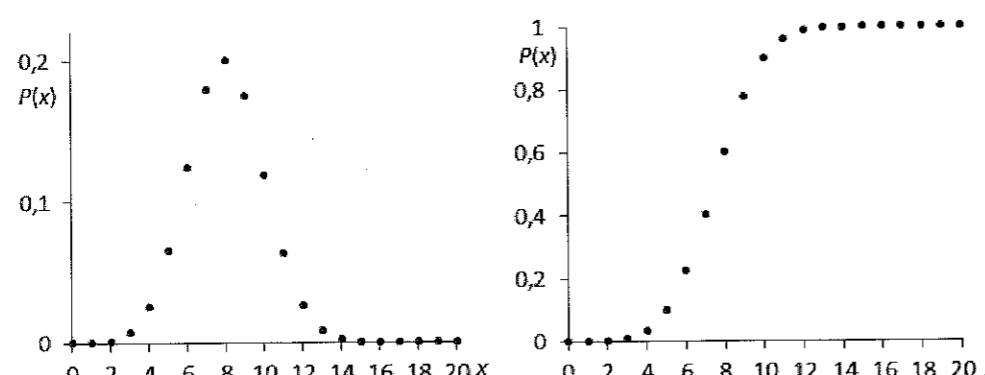
Střední hodnota náhodné veličiny X s rozdělením $Hg(M, N, n)$ je

$$E(X) = n \frac{M}{N}, \quad (2.111)$$

rozptyl je roven

$$D(X) = n \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N-n}{N-1}. \quad (2.112)$$

Pro $\pi = M/N$ je střední hodnota hypergeometrického rozdělení stejná jako střední hodnota rozdělení $Bi(n, \pi)$. Jaký tvar má pravděpodobnostní a distribuční funkce tohoto rozdělení je vidět na obrázcích 2.21 a 2.22.



Obr. 2.21 Pravděp. funkce rozdělení $Hg(20,100,40)$ Obr. 2.22 Distrib. funkce rozdělení $Hg(20,100,40)$

2.5 Vybraná spojitá rozdělení

Se spojitymi rozděleními se budeme velice často setkávat v matematické statistice. Těchto rozdělení je velké množství, my se omezíme na přehled těch nejpoužívanějších.

2.5.1 Rovnoměrné spojité rozdělení

Řekneme, že náhodná veličina X má rovnoměrné spojité rozdělení s parametry α a β jestliže její hustota pravděpodobnosti má tvar

$$\begin{aligned} f(x) &= \frac{1}{\beta - \alpha}, & \alpha < x < \beta, \\ &= 0, & \text{jinak,} \end{aligned} \quad (2.113)$$

kde α a β jsou reálná čísla. Fakt, že se náhodná veličina řídí rovnoměrným rozdělením, budeme symbolicky zapisovat $X \sim R(\alpha, \beta)$.

Speciálním případem je rovnoměrné rozdělení na intervalu $(0,1)$, které má hustotu pravděpodobnosti

$$\begin{aligned} f(x) &= 1, & 0 < x < 1, \\ &= 0, & \text{jinak.} \end{aligned} \quad (2.114)$$

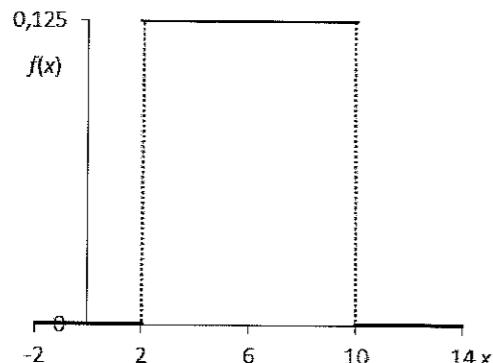
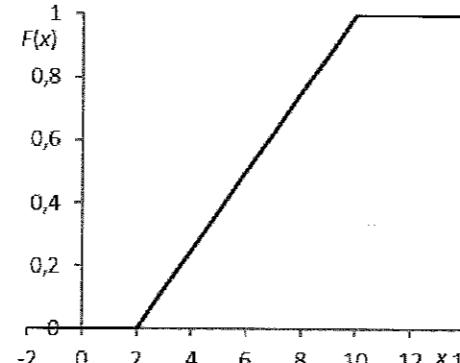
Distribuční funkce

Distribuční funkce $F(x)$ rovnoměrného rozdělení má tvar

$$\begin{aligned} F(x) &= 0, & x \leq \alpha, \\ &= \frac{x-\alpha}{\beta-\alpha}, & \alpha < x < \beta, \\ &= 1, & x \geq \beta. \end{aligned} \quad (2.115)$$

Rovnoměrné rozdělení se používá pro modelování doby čekání na události, které nastávají pravidelně v čase (doba čekání na příjezd metra, když náhodně přijde na stanici, doba čekání na zprávy v každou celou hodinu, pokud náhodně zapneme rádio atd.). Hodnoty náhodné (resp. pseudonáhodné) veličiny s rozdělením $R(0, 1)$ je většinou možné vygenerovat v různých programech (např. MS Excel) a tyto pak slouží jako základ pro generování náhodných veličin z dalších rozdělení.

Na obrázcích 2.23 a 2.24 je ukázka hustoty pravděpodobnosti a distribuční funkce rovnoměrného rozdělení na intervalu $(2, 10)$. Je vidět, že hustota tvoří vlastně obdélník, jehož plocha je rovna jedné. Pravděpodobnost, že náhodná veličina s rozdělením $R(2, 10)$ nabude hodnoty z určitého intervalu, je potom dáná plochou obdélníku, vymezeného krajními body tohoto intervalu, osou x a hustotou pravděpodobnosti. To samozřejmě velmi usnadňuje výpočet.

Obr. 2.23 Hustota pravděp. rozd. $R(2, 10)$ Obr. 2.24 Distribuční funkce rozd. $R(2, 10)$

Momenty

Střední hodnota náhodné veličiny s rozdelením $R(\alpha, \beta)$ je

$$E(X) = \frac{\alpha + \beta}{2}, \quad (2.116)$$

rozptyl je roven

$$D(X) = \frac{(\beta - \alpha)^2}{12}. \quad (2.117)$$

2.5.2 Normální rozdělení

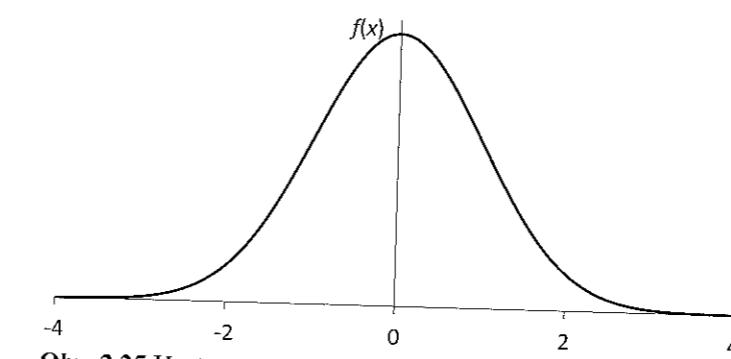
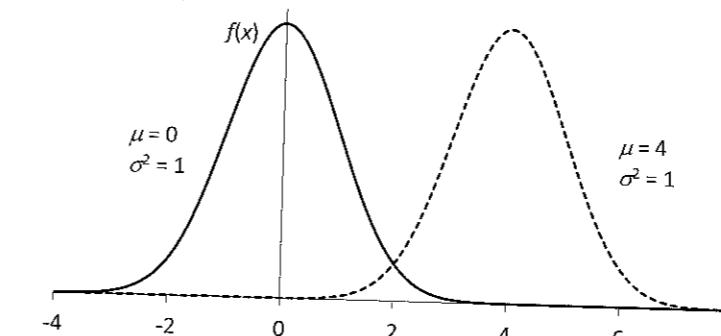
Normální rozdělení je bezesporu nejdůležitějším rozdělením ze všech pravděpodobnostních rozdělení. Hraje zcela zásadní roli jak v pravděpodobnosti, tak v matematické statistice, a to jak v oblasti teoretické, tak v oblasti praktické.

Náhodná veličina X má normální rozdělení s parametry μ a σ^2 , jestliže její hustota pravděpodobnosti má tvar

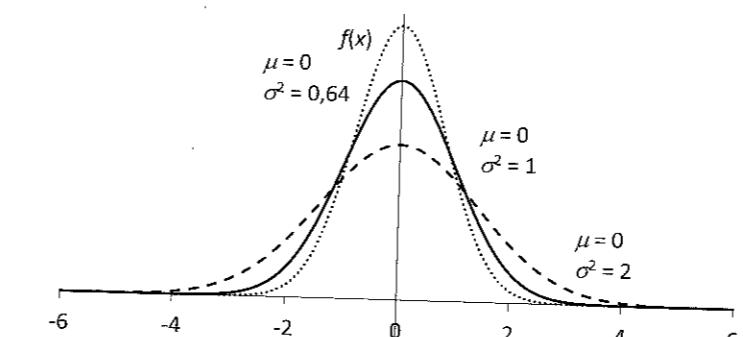
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty, \quad (2.118)$$

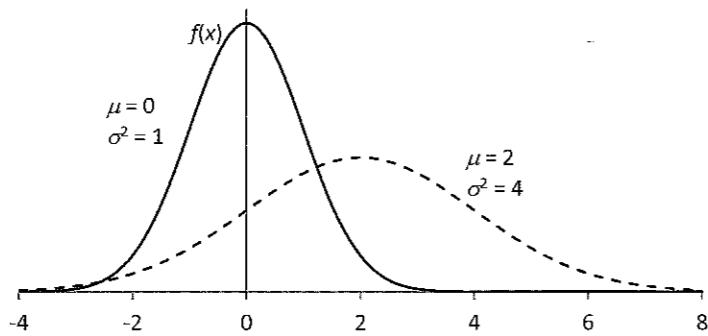
kde $-\infty < \mu < +\infty$, $\sigma^2 > 0$. Fakt, že se náhodná veličina řídí normálním rozdělením, budeme symbolicky zapisovat $X \sim N(\mu, \sigma^2)$. Normální rozdělení je symetrické, jeho hustota pravděpodobnosti je symetrická kolem bodu μ . Někdy je toto rozdělení označováno jako rozdělení Gaussovo či rozdělení Laplaceovo-Gaussovo.

Jak hodnoty parametrů ovlivňují konkrétní tvar hustoty pravděpodobnosti, vidíme na obrázcích 2.25 – 2.28. Zatímco parametr μ určuje polohu rozdělení na reálné ose, parametr σ^2 určuje konkrétní tvar hustoty (variabilitu). Na obrázku 2.25 je zobrazena hustota pravděpodobnosti normálního rozdělení s parametry 0 a 1.

Obr. 2.25 Hustota pravděpodobnosti rozdělení $N(0, 1)$ Obr. 2.26 Hustota pravděpodobnosti rozdělení $N(0, 1)$ a $N(4, 1)$

Na obrázku 2.26 jsou hustoty normálního rozdělení, lišícího se polohou – tj. hodnotou parametru μ . Hodnota parametru σ^2 je v obou případech stejná. Následuje obrázek 2.27 tří hustot normálního rozdělení se stejnou polohou (tj. se stejnou hodnotou parametru μ) a různou variabilitou (tj. různou hodnotou parametru σ^2). Na obrázku 2.28 jsou dvě hustoty normálního rozdělení, lišící se polohou (tj. s různou hodnotou parametru μ) i variabilitou (tj. s různou hodnotou parametru σ^2).

Obr. 2.27 Hustota pravděpodobnosti rozdělení $N(0; 0,64)$, $N(0, 1)$ a $N(0, 2)$

Obr. 2.28 Hustota pravděpodobnosti rozdělení $N(0, 1)$ a $N(2, 4)$

Normované normální rozdělení

Položíme-li hodnoty parametrů $\mu = 0$ a $\sigma^2 = 1$, obdržíme tzv. normované normální rozdělení $N(0, 1)$. Pokud má náhodná veličina U normované normální rozdělení, má její hustota pravděpodobnosti tvar

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}, \quad -\infty < u < +\infty. \quad (2.119)$$

Toto rozdělení je symetrické kolem bodu $u = 0$, a proto můžeme psát

$$\varphi(-u) = \varphi(u), \quad -\infty < u < +\infty. \quad (2.120)$$

Protože hodnoty parametrů 0 a 1 jsou spíše výjimečné, je třeba najít způsob, jak od obecného normálního rozdělení (s libovolnými parametry) přejít k rozdělení normovanému. Postup, který tento přechod umožňuje, je normování. Pokud má náhodná veličina X obecné normální rozdělení s parametry μ a σ^2 , pak náhodná veličina

$$U = \frac{X - \mu}{\sigma} \quad (2.121)$$

má normované normální rozdělení $N(0, 1)$. Je zřejmé, že z tohoto vztahu můžeme vyjádřit náhodnou veličinu X ve tvaru

$$X = \sigma U + \mu, \quad (2.122)$$

což nám umožní přejít od rozdělení $N(0, 1)$ k obecnému rozdělení $N(\mu, \sigma^2)$.

Distribuční funkce

Distribuční funkci normovaného normálního rozdělení budeme značit symbolem Φ . Pro toto rozdělení – díky symetrii kolem bodu $\mu = 0$ – obdržíme

$$\Phi(u) = 1 - \Phi(-u), \quad -\infty < u < +\infty, \quad (2.123)$$

a dále

$$F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right), \quad -\infty < x < +\infty. \quad (2.124)$$

Tento vztah budeme často využívat při výpočtu konkrétních hodnot pravděpodobností. Jelikož uvedené integrály nejdou explicitně vyčíslit, jsou hodnoty distribuční funkce normovaného normálního rozdělení tabelované, a lze je nalézt ve statistických tabulkách či učebnicích statistiky. Stačí přitom tabelovat hodnoty distribuční funkce pro $u > 0$, hodnoty pro $u < 0$ se získají díky symetrii rozdělení.

Momenty

Střední hodnota náhodné veličiny s rozdělením $N(\mu, \sigma^2)$ je

$$E(X) = \mu, \quad (2.125)$$

rozptyl je roven

$$D(X) = \sigma^2. \quad (2.126)$$

Střední hodnota a rozptyl jsou tedy přímo parametry rozdělení. Jak již bylo vidět z předchozích obrázků, parametr μ určuje polohu a parametr σ^2 určuje variabilitu.

Kvantily

Z definice kvantilu lze pro $100P\%$ kvantil u_P normovaného normálního rozdělení psát

$$\Phi(u_P) = P, \quad 0 < P < 1. \quad (2.127)$$

Ze symetrie rozdělení $N(0, 1)$ dále vyplývá, že

$$u_{1-P} = -u_P, \quad 0 < P < 1. \quad (2.128)$$

Přejdeme-li nyní k obecnému normálnímu rozdělení $N(\mu, \sigma^2)$, platí

$$P = F(x_P) = \Phi\left(\frac{x_P - \mu}{\sigma}\right). \quad (2.129)$$

Odtud je již zřejmé, že platí

$$u_P = \frac{x_P - \mu}{\sigma}, \quad 0 < P < 1, \quad (2.130)$$

a tedy

$$x_P = \mu + \sigma u_P, \quad 0 < P < 1. \quad (2.131)$$

Hodnoty kvantilů normovaného normálního rozdělení lze opět nalézt v tabulkách, přičemž (z důvodu symetrie rozdělení) se publikují pouze hodnoty kvantilů pro $P > 0,5$.

Lineární kombinace nezávislých náhodných veličin s normálním rozdělením
Často nás zajímá nejen rozdělení náhodných veličin jako takových, ale také rozdělení jejich součtu případně průměru. Pro náhodné veličiny s normálním rozdělením platí, že jejich lineární kombinace má opět normální rozdělení. Speciálně tedy, pokud X_1, \dots, X_n jsou nezávislé náhodné veličiny z rozdělení $N(\mu, \sigma^2)$, má

- náhodná veličina $Y = \sum_{i=1}^n X_i$ rozdělení $N(n\mu, n\sigma^2)$,
- náhodná veličina $Y = \sum_{i=1}^n X_i / n$ rozdělení $N(\mu, \sigma^2/n)$.

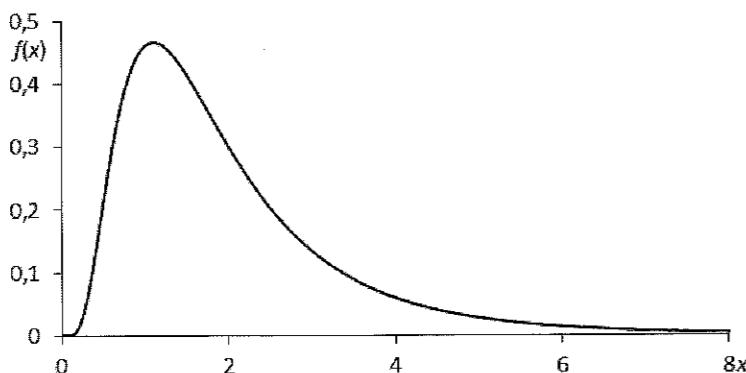
2.5.3 Logaritmicko-normální rozdelení

Řekneme, že náhodná veličina X má logaritmicko-normální rozdelení s parametry μ a σ^2 , $-\infty < \mu < +\infty$, $\sigma^2 > 0$, jestliže její hustota pravděpodobnosti má tvar

$$\begin{aligned} f(x) &= \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, & x > 0, \\ &= 0, & x \leq 0. \end{aligned} \quad (2.132)$$

Fakt, že se náhodná veličina řídí logaritmicko-normálním rozdelením, budeme symbolicky zapisovat $X \sim LN(\mu, \sigma^2)$.

Toto rozdelení se používá velmi často pro modelování mezd či obecně příjmů obyvatelstva, uplatňuje se i v oblasti pojišťovnictví pro modelování výše škod. Již z jeho názvu vyplývá příbuznost s normálním rozdelením. Uvažujeme-li totiž náhodnou veličinu $Y = \ln X$, kde X má rozdelení $LN(\mu, \sigma^2)$, má náhodná veličina Y rozdelení $N(\mu, \sigma^2)$. Naopak, pokud má náhodná veličina Y rozdelení $N(\mu, \sigma^2)$, má veličina $X = \exp(Y)$ rozdelení $LN(\mu, \sigma^2)$. Na rozdíl od normálního rozdelení se jedná o rozdelení asymetrické, jak je ostatně patrné i z obrázku 2.29.



Obr. 2.29 Hustota pravděpodobnosti rozdelení $LN(0,5; 0,4)$.

Momenty

Střední hodnota náhodné veličiny s rozdelením $LN(\mu, \sigma^2)$ je

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right), \quad (2.133)$$

rozptyl je roven

$$D(X) = \exp(2\mu + \sigma^2) [\exp(\sigma^2) - 1]. \quad (2.134)$$

Střední hodnota a rozptyl v tomto případě tedy nejsou přímo parametry rozdelení, jsou pouze funkcí těchto parametrů. Nicméně není těžké určit, že

$$\mu = \ln E(X)^2 - \frac{\ln E(X^2)}{2}, \quad (2.135)$$

a

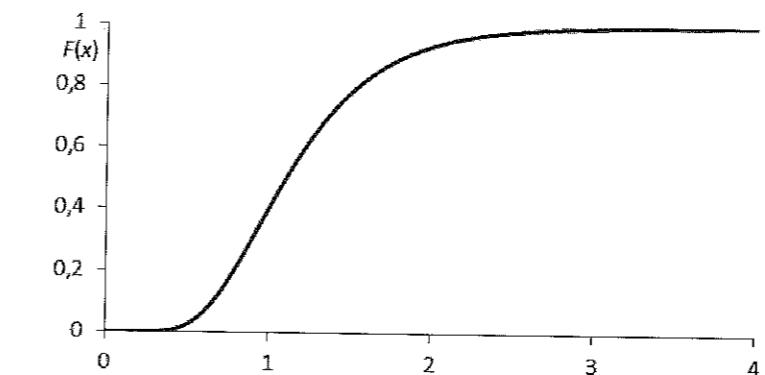
$$\sigma^2 = \ln\left(\frac{E(X^2)}{E(X)^2}\right). \quad (2.136)$$

Pro kvantily lze snadno odvodit, že platí

$$x_P = \exp(\mu + \sigma u_P), \quad 0 < P < 1, \quad (2.137)$$

kde u_P je $100P\%$ kvantil rozdelení $N(0, 1)$.

Tvar distribuční funkce vidíme na obrázku 2.30.



Obr. 2.30 Distribuční funkce rozdelení $LN(0,1; 0,16)$

2.5.4 Exponenciální rozdelení

Řekneme, že náhodná veličina X má exponenciální rozdelení s parametry A a δ , $-\infty < A < +\infty$, $\delta > 0$, jestliže její hustota pravděpodobnosti má tvar

$$\begin{aligned} f(x) &= \frac{1}{\delta} e^{-\frac{(x-A)}{\delta}}, & x > A, \\ &= 0, & x \leq A. \end{aligned} \quad (2.138)$$

Fakt, že se náhodná veličina řídí exponenciálním rozdelením, budeme symbolicky zapisovat $X \sim E(A, \delta)$.

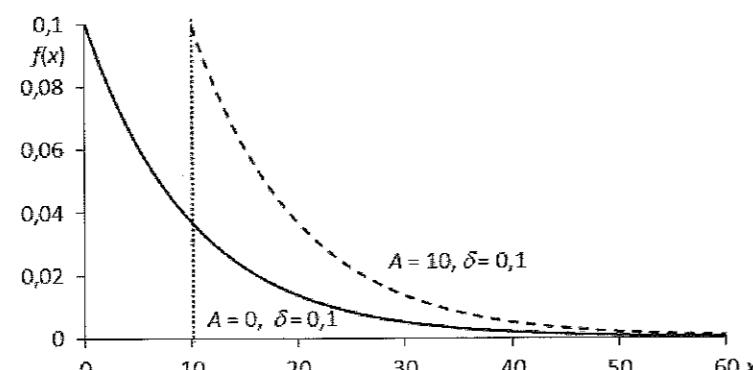
Toto rozdelení se používá velmi často v teorii spolehlivosti a životnosti, v teorii hromadné obsluhy (spolu s Poissonovým rozdelením tvoří tzv. Poissonův proces), v teorii obnovy, v oblasti modelování výše škod ve neživotním pojistění apod.

O tomto rozdelení se říká, že tzv. „nemá paměť“. Pokud jako náhodnou veličinu X označíme životnost (dobu života) nějakého zařízení, pak pravděpodobnost, že zařízení,

které pracovalo bez poruchy a hodin, bude pracovat bez poruchy ještě alespoň dalších x hodin, je rovna pravděpodobnosti, že zařízení, které ještě nebylo v provozu, bude pracovat alespoň x hodin. To znamená, že zařízení jako by zapomnělo již odpracovanou dobu. Takováto zařízení se v praxi vyskytuje všude tam, kde k poruchám dochází ze zcela náhodných (často vnějších) příčin a nikoliv v důsledku opotřebení (mechanické opotřebení, únavy materiálu apod.). Jedná se často o dobu životnosti elektronických zařízení. Tuto vlastnost mají např. počítačové paměti, procesory apod. Pokud dochází k opotřebení, je možné životnost modelovat některými dalšími pravděpodobnostními rozděleními.

Obecně platí, že toto rozdělení je vhodné k modelování doby výskytu mezi jednotlivými událostmi, které nastávají náhodně v čase. Často se jedná o požadavky na určitou službu – požadavky na přistávací dráhu na letištích, požadavky na obsluhu automatů (turnikety, telefonní ústředny) atd.

Podívejme se nyní na tvar hustoty pravděpodobnosti pro různé hodnoty parametrů A a δ na obrázku 2.31.



Obr. 2.31 Hustota pravděpodobnosti rozdělení $E(0; 0,1)$ a $E(10; 0,1)$

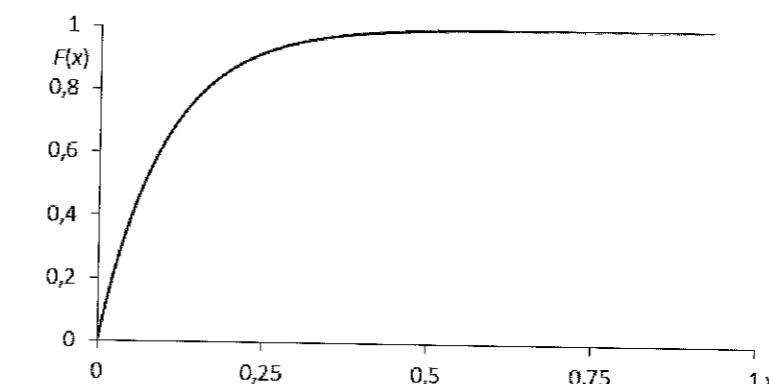
Je zřejmé, že parametr A je pouze posunutí hustoty, jeho změna tedy ovlivňuje polohu, ale nikoliv samotný tvar hustoty. Pokud je $A = 0$, bývá často exponenciální rozdělení označováno jako $E(\delta)$ a mlčky se předpokládá nulovost parametru A .

Distribuční funkce

Distribuční funkce exponenciálního rozdělení má tvar

$$\begin{aligned} F(x) &= \int_A^x f(t) dt = 1 - e^{-\frac{(x-A)}{\delta}}, & x > A, \\ &= 0, & x \leq A. \end{aligned} \quad (2.139)$$

Její tvar pro $A = 0$ a $\delta = 10$ je znázorněn na obrázku 2.32.



Obr. 2.32 Distribuční funkce rozdělení $E(0, 10)$

Kvantily

Pro $100P\%$ kvantil x_P exponenciálního rozdělení $E(A, \delta)$ můžeme psát

$$x_P = A - \delta \ln(1 - P), \quad 0 < P < 1. \quad (2.140)$$

Přitom

$$-\ln(1 - P), \quad 0 < P < 1, \quad (2.141)$$

je $100P\%$ kvantil rozdělení $E(0,1)$. To znamená, že pro medián platí

$$x_{0,5} = A + \delta \ln(2) = A + 0,69315\delta. \quad (2.142)$$

Momenty

Střední hodnota náhodné veličiny s rozdělením $E(A, \delta)$ je

$$E(X) = A + \delta, \quad (2.143)$$

rozptyl

$$D(X) = \delta^2. \quad (2.144)$$

Rozptyl tedy nijak nezávisí na parametru A .

2.5.5 Rozdělení chí-kvadrát

Náhodná veličina X má rozdělení chí-kvadrát s parametrem v , jestliže její hustota pravděpodobnosti má tvar

$$\begin{aligned} f(x) &= \frac{1}{2^{v/2} \Gamma\left(\frac{v}{2}\right)} x^{v/2-1} \exp\left(-\frac{x}{2}\right), & x > 0, \\ &= 0, & x \leq 0. \end{aligned} \quad (2.145)$$

Rozdělení je tedy jednoparametrické, parametr v se označuje jako stupeň volnosti. Toto rozdělení budeme značit $\chi^2(v)$.

Kvantity

Hodnoty $100P\%$ kvantilu x_p chí-kvadrát rozdělení $\chi^2(\nu)$ není snadné spočítat, a proto bývají tabelovány nebo se určují numericky.

Momenty

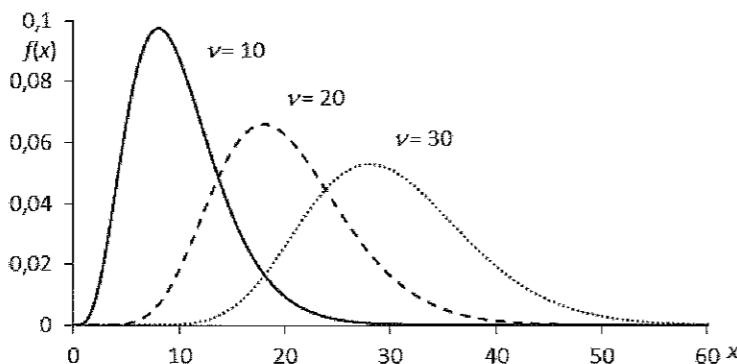
Střední hodnota náhodné veličiny s rozdělením $\chi^2(\nu)$ je pak

$$E(X) = \nu, \quad (2.146)$$

a rozptyl je

$$D(X) = 2\nu. \quad (2.147)$$

Jak hodnota parametru ν ovlivňuje tvar rozdělení, je vidět na obrázku 2.33.



Obr. 2.33 Hustota pravděpodobnosti rozdělení $\chi^2(10)$, $\chi^2(20)$ a $\chi^2(30)$

Dále uvedeme několik tvrzení, která jsou užitečná v matematické statistice. Využijeme je zejména při testování hypotéz. Tvrzení uvedeme ve formě vlastnosti, přesnou formulaci lze nalézt v Marek (2012).

Vlastnosti

- Pro nezávislé náhodné veličiny X_1, \dots, X_n s rozdělením $N(0, 1)$ platí, že náhodná veličina $Y = X_1^2 + \dots + X_n^2$ má rozdělení χ^2 o n stupních volnosti.
- Pro nezávislé náhodné veličiny X a Z s rozdělením $\chi^2(n)$ a $\chi^2(m)$ má náhodná veličina $Y = X + Z$ rozdělení χ^2 o $n + m$ stupních volnosti.
- Nechť X_1, \dots, X_n jsou nezávislé náhodné veličiny s rozdělením $N(\mu, \sigma^2)$. Označme $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / n - 1$, kde $\bar{X} = \sum_{i=1}^n X_i / n$. Pak platí
 - $(n - 1)S^2/\sigma^2$ má pro $n > 1$ a $\sigma^2 > 0$ rozdělení $\chi^2(n - 1)$,
 - pro $n > 1$ jsou veličiny \bar{X} a S^2 nezávislé.

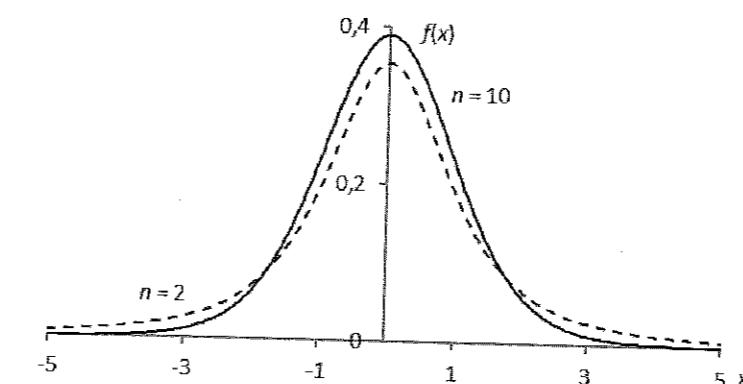
2.5.6 Studentovo rozdělení (t-rozdělení)

Náhodná veličina X má Studentovo rozdělení (t rozdělení) o n stupních volnosti, jestliže její hustota pravděpodobnosti má tvar

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < x < \infty, \quad (2.148)$$

kde n je přirozené číslo. Toto rozdělení budeme značit $t(n)$.

Jak počet stupňů volnosti ovlivňuje tvar rozdělení, je vidět na obrázku 2.34.



Obr. 2.34 Hustota pravděpodobnosti rozdělení $t(2)$ a $t(10)$

Studentovo rozdělení poprvé použil W. S. Gosset, který vystupoval pod pseudonymem Student (odtud název rozdělení). Toto rozdělení se používá zejména při testování hypotéz – velká část testových statistik má právě toto rozdělení (viz kapitola 3).

Na obrázku 2.34 je vidět, že se jedná o symetrické rozdělení kolem bodu $x = 0$. Z tohoto důvodu platí pro hustotu vztah $f(x) = f(-x)$, $-\infty < x < \infty$. Jeho tvar nápadně připomíná hustotu pravděpodobnosti normovaného normálního rozdělení $N(0, 1)$. Při určení hodnot $100P\%$ kvantilu t_p t-rozdělení využíváme symetrii kolem bodu 0. Hodnoty kvantilů jsou tabelované, přičemž platí

$$-t_p(n) = t_{1-p}(n), \quad 0 < P < 1. \quad (2.149)$$

Momenty

Střední hodnota existuje pro $n > 1$ a je rovna

$$E(X) = 0, \quad (2.150)$$

rozptyl existuje pro $n > 2$ a je roven

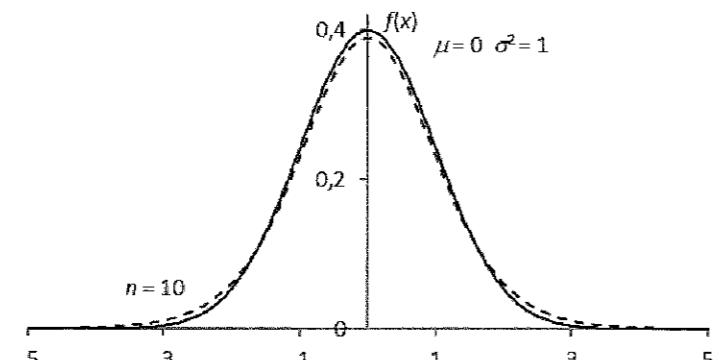
$$D(X) = \frac{n}{n-2}. \quad (2.151)$$

Vzhledem k tomu, že se jedná o jednovrcholové symetrické rozdělení kolem bodu 0, jsou si rovny střední hodnota, medián i modus a můžeme tedy psát

$$E(X) = t_{0,5}(n) = \hat{x} = 0. \quad (2.152)$$

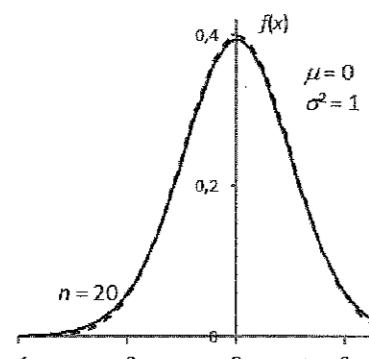
Podobnost s normálním rozdělením

Na obrázku 2.35 je vidět, že mezi hustotou t rozdělení s 10 stupni volnosti (čárkovaná čára) a hustotou normovaného normálního rozdělení (plná čára) není velký rozdíl.

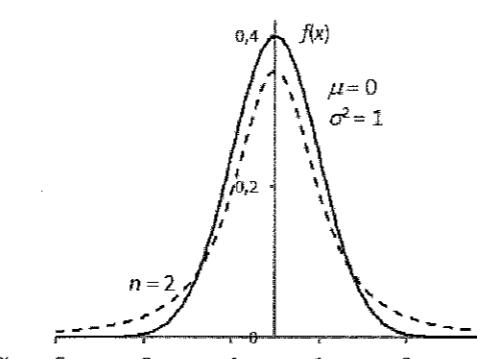


Obr. 2.35 Hustota pravděpodobnosti rozdělení $t(10)$ a $N(0, 1)$

Konec t rozdělení však klesají k nule pomaleji, než je tomu u normovaného normálního rozdělení. Toho se využívá zejména při testování hypotéz, když se pracuje s t rozdělením testové statistiky. Rozdíl mezi hustotami obou rozdělení bude menší a menší při rostoucím n . Budeme-li pracovat s velkým počtem stupňů volnosti (např. $n = 20$), je rozdíl zanedbatelný. Naproti tomu, při malém počtu stupňů volnosti, je rozdíl mezi oběma hustotami mnohem výraznější. To dokládají obrázky 2.36 a 2.37.



Obr. 2.36 Hustota pravděp. rozd. $t(20)$ a $N(0, 1)$



Obr. 2.37 Hustota pravděp. rozd. $t(2)$ a $N(0, 1)$

I pro toto rozdělení uvedeme dvě tvrzení, která jsou užitečná v matematické statistice. Tvrzení uvedeme ve formě vlastnosti, přesná formulace je v Marek (2012).

Vlastnosti

- Pro nezávislé náhodné veličiny X a Z takové, že $X \sim N(0, 1)$ a $Z \sim \chi^2(k)$, má náhodná veličina T Studentovo rozdělení s k stupni volnosti

$$T = \frac{X}{\sqrt{Z/k}}.$$

- Pro nezávislé náhodné veličiny X_1, \dots, X_n z rozdělení $N(\mu, \sigma^2)$, kde $n \geq 2$ a $\sigma^2 > 0$, má náhodná veličina T Studentovo rozdělení s $n-1$ stupni volnosti

$$T = \frac{\bar{X} - \mu}{S} \sqrt{n}.$$

2.5.7 Fisherovo-Snedecorovo rozdělení (F rozdělení)

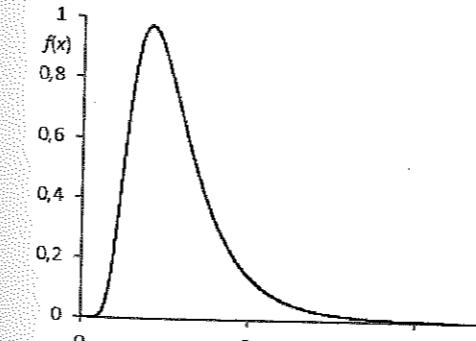
Náhodná veličina X má Fisherovo-Snedecorovo rozdělení (F rozdělení) o m a n stupních volnosti, jestliže její hustota pravděpodobnosti má tvar

$$f(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right)^{m/2} x^{m/2-1} \left(1 + \frac{m}{n}x\right)^{\frac{m+n}{2}}, \quad x > 0, \\ = 0, \quad x \leq 0,$$
(2.153)

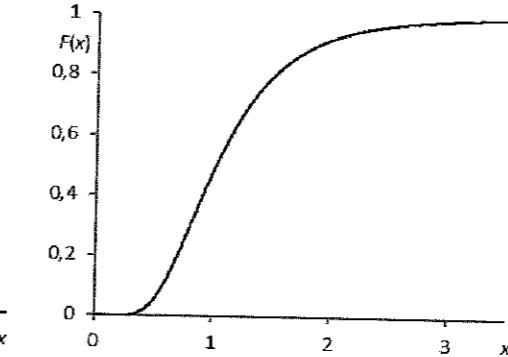
kde m a n jsou přirozená čísla. Toto rozdělení budeme značit $F(m, n)$.

Toto rozdělení se používá zejména při konstrukci intervalů spolehlivosti a při testování hypotéz – některé testové statistiky mají právě toto rozdělení. Jedná se zejména o testy týkající se rozptylu, některé testy v regresní analýze apod.

F rozdělení je kladně zešikmené jednovrcholové rozdělení se dvěma parametry, které nazýváme stupně volnosti. Jejich hodnoty výrazně ovlivňují tvar rozdělení, jak je vidět na obrázku 2.38. Tvar distribuční funkce tohoto rozdělení je na obrázku 2.39.



Obr. 2.38 Hustota pravd. rozdelení $F(30, 15)$



Obr. 2.39 Distribuční funkce rozděl. $F(30, 15)$

Náhodnou veličinu s F -rozdelením obdržíme, pokud dělíme dvě nezávislé náhodné veličiny s χ^2 -kvadrát rozdelením, každou z nich přitom počtem stupňů volnosti.

Vlastnosti

- Pro nezávislé náhodné veličiny X a Z takové, že $X \sim \chi^2(m)$ a $Z \sim \chi^2(n)$, má náhodná veličina $F = (X/m)/(Y/n)$ rozdelení F o m a n stupních volnosti.
- Nechť X_1, \dots, X_n jsou nezávislé náhodné veličiny z rozdelení $N(\mu_1, \sigma_1^2)$ a Y_1, \dots, Y_m jsou nezávislé náhodné veličiny z rozdelení $N(\mu_2, \sigma_2^2)$, kde $n \geq 2$, $m \geq 2$, $\sigma_1^2 > 0$ a $\sigma_2^2 > 0$. Předpokládejme, že oba výběry jsou nezávislé. Pak náhodná veličina $Z = S_X^2/S_Y^2$ má rozdelení $F(n-1, m-1)$.

Kvantily

Hodnoty kvantilů jsou tabelovány. Při určení hodnot $100P\%$ kvantilu f_P rozdelení F využíváme následující vlastnost. Pro kvantily F rozdelení platí vztah

$$f_P(m, n) = \frac{1}{f_{1-P}(n, m)}, \quad 0 < P < 1. \quad (2.154)$$

Momenty

Střední hodnota existuje, pokud $n > 2$ a je rovna

$$E(X) = \frac{n}{n-2}, \quad (2.155)$$

rozptyl existuje pro $n > 4$ a je roven

$$D(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}. \quad (2.156)$$

2.6 Limitní věty

V teorii pravděpodobnosti sehrávají klíčovou úlohu tzv. limitní věty. Jedná se o **zákon velkých čísel** (ZVČ) a **centrální limitní větu** (CLV). Obě věty existují v nejrůznějších podobách. Název těchto vět je často spojován s jejich autory a jednotlivá znění říkají zhruba totéž, liší se však poměrně výrazně silou předpokladů. Také důkazy těchto vět bývají různě složité – opět jsou poplatné použitým předpokladům. Pokud by se čtenář o této části teorie pravděpodobnosti chtěl dozvědět více, doporučujeme publikace Rao (1978), Rényi (1972). My uvedeme pouze několik základních znění těchto vět a soustředíme se spíše na věty s jednoduššími předpokladami.

Až dosud jsme pracovali s jednotlivými náhodnými veličinami či s náhodnými vektory. V následujících úvahách budeme pracovat s **posloupnostmi náhodných veličin**, případně s **posloupnostmi jejich pravděpodobnostních rozdělení**.

Obecně nás tedy bude zajímat, zda posloupnosti náhodných veličin nevykazují při rostoucím n určité obecné vlastnosti, které nebudou závislé na konkrétním pravděpodobnostním rozdělení původních náhodných veličin.

Připomeňme v této souvislosti pojmenování **stabilita relativních četností**. Tuto stabilitu lze zhruba interpretovat tak, že se zvyšujícím se počtem realizací náhodného pokusu se pozorovaná relativní četnost (empirická veličina) ustalovala na určité hodnotě (teoretické veličině), kterou jsme nakonec využili při definici pravděpodobnosti. Podobně i zákon velkých čísel říká, že za jistých podmínek se empirické charakteristiky budou blížit svým teoretickým protějškům.

Centrální limitní věta hovoří o tom, jak se při velkém n chová úhrn či průměr nezávislých náhodných veličin.

K tomu, abychom obě věty definovali přesněji, potřebujeme zavést nové typy konvergence. V případě ZVČ se jedná o **konvergenci podle pravděpodobnosti** a v případě CLV o **konvergenci v distribuci**.

V celé části věnované limitním větám předpokládejme, že n je přirozené číslo, tedy lze psát $n \in N$. Tento předpoklad již nebudeme v jednotlivých větách znova opakovat.

Konvergence podle pravděpodobnosti

Nechť $\{X_n\}$ je posloupnost náhodných veličin. Řekneme, že posloupnost $\{X_n\}$ konverguje k a podle pravděpodobnosti, jestliže pro každé $\varepsilon > 0$ platí

$$\lim_{n \rightarrow \infty} P(|X_n - a| \geq \varepsilon) = 0. \quad (2.157)$$

Uvedená limita říká, že pravděpodobnost absolutní odchyly od a konverguje k nule pro n rostoucí nad všechny meze. Je třeba si uvědomit i to, že se nejedná o klasickou konvergenci, ale o konvergenci pravděpodobnostní. Tuto limitu lze zapsat také jako

$$\lim_{n \rightarrow \infty} P(|X_n - a| < \varepsilon) = 1. \quad (2.158)$$

2.6.1 Slabý zákon velkých čísel

Nechť $\{X_n\}$ je posloupnost náhodných veličin a $\{X_n - E(X_n)\}$ je posloupnost odchylek veličin X_n od svých středních hodnot $E(X_n)$. Položme

$$Z_n = \frac{1}{n} \sum_{j=1}^n [X_j - E(X_j)]. \quad (2.159)$$

Řekneme, že pro posloupnost $\{X_n\}$ platí **slabý zákon velkých čísel**, jestliže posloupnost $\{Z_n\}$ konverguje k nule podle pravděpodobnosti. Tzn., že pro každé $\varepsilon > 0$ platí

$$\lim_{n \rightarrow \infty} P \left[\left| \frac{1}{n} \sum_{j=1}^n [X_j - E(X_j)] \right| \geq \varepsilon \right] = 0. \quad (2.160)$$

Je třeba si uvědomit, co tento výraz znamená. Uvedenou limitu lze totiž přepsat do tvaru

$$\lim_{n \rightarrow \infty} P[|Z_n| \geq \varepsilon] = \lim_{n \rightarrow \infty} P\left[\left|\frac{1}{n} \sum_{j=1}^n X_j - \frac{1}{n} \sum_{j=1}^n E(X_j)\right| \geq \varepsilon\right] = 0. \quad (2.161)$$

Jinými slovy, aritmetický průměr $\bar{X}_n = \sum_{j=1}^n X_j / n$ konverguje podle pravděpodobnosti k aritmetickému průměru ze středních hodnot $\sum_{j=1}^n E(X_j) / n$.

Hovoříme o tzv. **slabém** zákonu velkých čísel. Existuje i **silný** zákon velkých čísel. Obě verze ZVČ se odlišují především typem konvergence. Zatímco u slabého zákonu hovoříme o konvergenci **podle pravděpodobnosti**, u silného zákonu bychom pracovali s konvergencí ve smyslu **skoro jistě**. Budeme se nyní věnovat podmínkám, za kterých ZVČ platí.

Čebyševova věta

Nechť $\{X_n\}$ je posloupnost náhodných veličin s konečným rozptylem a nulovou kovariancí, tedy pro ně platí

$$D(X_i) \leq K \quad \text{a} \quad \text{cov}(X_i, X_j) = 0, \quad i \neq j, i, j = 1, 2, \dots,$$

kde $K < \infty$. Pak pro posloupnost $\{X_n\}$ platí slabý zákon velkých čísel.

Bernoulliova věta

Nechť $\{X_n\}$ je posloupnost vzájemně nezávislých náhodných veličin s alternativním rozdelením se stejným parametrem π , $0 < \pi < 1$. Pak pro každé $\varepsilon > 0$ platí

$$\lim_{n \rightarrow \infty} P\left[\left|\frac{1}{n} \sum_{j=1}^n X_j - \pi\right| \geq \varepsilon\right] = 0. \quad (2.162)$$

Bernoulliova věta je tedy speciálním případem Čebyševovy věty. Uvědomme si, že veličina $\sum_{j=1}^n X_j$ představuje v případě alternativního rozdelení počet úspěchů (četnost výskytu určitého jevu) v n nezávislých opakování téhož pokusu, přičemž pravděpodobnost úspěchu je rovna π . Veličina $\sum_{j=1}^n X_j / n$ pak představuje relativní četnost výskytu tohoto jevu. Bernoulliova věta tak vlastně tvrdí, že relativní četnost sledovaného jevu stochasticky konverguje k pravděpodobnosti tohoto jevu.

Důsledek Čebyševovy věty

Nechť $\{X_n\}$ je posloupnost nekorelovaných náhodných veličin se stejnou střední hodnotou μ a stejným konečným rozptylem σ^2 . Platí tedy

$$E(X_i) = \mu, \quad D(X_i) = \sigma^2 \quad \text{a} \quad \text{cov}(X_i, X_j) = 0, \quad i \neq j, i, j = 1, 2, \dots,$$

kde $K < \infty$. Pak pro posloupnost $\{X_n\}$ platí slabý zákon velkých čísel. Můžeme tedy psát

$$\lim_{n \rightarrow \infty} P\left[\left|\frac{1}{n} \sum_{j=1}^n X_j - \mu\right| \geq \varepsilon\right] = 0. \quad (2.163)$$

Uvedené tvrzení plyne přímo z Čebyševovy věty a má i důležitý praktický důsledek. Za uvedených předpokladů totiž můžeme střední hodnotu μ při dostatečně velkém množství pozorování odhadnout aritmetickým průměrem.

Konvergence v distribuci

Nechť $\{X_n\}$ je posloupnost náhodných veličin, jimž přísluší posloupnost distribučních funkcí $\{F_n(x)\}$. Dále nechť X je náhodná veličina s distribuční funkcí $F(x)$. Řekneme, že posloupnost $\{X_n\}$ konverguje v distribuci k X , jestliže

$$\lim_{n \rightarrow \infty} F_n(x) = F(x) \quad (2.164)$$

ve všech bodech spojitosti funkce $F(x)$.

Jestliže funkci $F(x)$ je distribuční funkce normálního rozdělení, pak říkáme, že posloupnost $\{X_n\}$ konverguje v distribuci k rozdělení $N(\mu, \sigma^2)$, jestliže

$$\lim_{n \rightarrow \infty} F_n(x) = \Phi\left(\frac{x - \mu}{\sigma}\right), \quad -\infty < x < \infty, \quad (2.165)$$

kde $\Phi(u)$ je distribuční funkce normovaného normálního rozdělení.

2.6.2 Centrální limitní věta

Centrální limitní věta hovoří o tom, jak se chová součet velkého množství nezávislých náhodných veličin. Ukazuje se totiž, že za poměrně obecných předpokladů lze pravděpodobnostní rozdělení takového součtu approximovat normálním rozdělením. O veličinách, jejichž limitním rozdělením je rozdělení normální, říkáme, že mají asymptoticky normální rozdělení. Je tedy zřejmé, že normální rozdělení sehrává v teorii limitních vět klíčovou úlohu.

Centrální limitní věta existuje v různých verzích, které se liší sítou svých předpokladů. My zde uvedeme tzv. Lindebergovu-Lévyho větu a Moivreovu-Laplaceovu větu. Existuje ale řada dalších, jejichž znění a důkazy lze nalézt např. v Rao (1978), Rényi (1972).

Lindebergova-Lévyho věta

Nechť $\{X_n\}$ je posloupnost vzájemně nezávislých náhodných veličin se stejným pravděpodobnostním rozdělením se střední hodnotou μ a konečným rozptylem σ^2 . Pak posloupnost $\{Y_n\}$, kde

$$Y_n = \frac{1}{\sigma\sqrt{n}} \left(\sum_{j=1}^n X_j - n\mu \right), \quad (2.166)$$

konverguje v distribuci k rozdělení $N(0, 1)$.

Z této věty vyplývá, že pro velká n lze rozdelení pravděpodobnosti náhodné veličiny $\sum_{j=1}^n X_j, j = 1, 2, \dots$, approximovat rozdelením $N(n\mu, n\sigma^2)$ a rozdelení náhodné veličiny $\sum_{j=1}^n X_j/n$ lze approximovat rozdelením $N(\mu, \sigma^2/n)$. Důležité je, že tak lze učinit za poměrně obecných předpokladů a bez znalosti původního rozdelení veličin X_j .

Moivreova-Laplaceova věta

Speciálním případem Lindebergovy-Lévyho věty je následující věta: Nechť $\{X_n\}$ je posloupnost vzájemně nezávislých náhodných veličin s alternativním rozdelením se stejným parametrem π , $0 < \pi < 1$. Pak posloupnost $\{Y_n\}$, kde

$$Y_n = \frac{1}{[n\pi(1-\pi)]^{1/2}} \left(\sum_{j=1}^n X_j - n\pi \right), \quad (2.167)$$

konverguje v distribuci k rozdelení $N(0, 1)$.

Příklad 2.15

Na základě minulých pozorování víme, že průměrná hmotnost balíku zasílaného poštou je 5 kg. Hmotnost zasílaného balíku můžeme považovat za náhodnou veličinu se střední hodnotou 5 kg a se směrodatnou odchylkou 1,5 kg (ta byla rovněž spočítána na základě minulých pozorování). Jaká je pravděpodobnost, že 100 zasílaných balíků, které se nakládají do poštovního auta, převýší svou hmotností užitečné zatížení nákladního prostoru tohoto automobilu, které činí 550 kg?

Řešení

Hmotnost i -tého balíku zasílaného poštou můžeme považovat za náhodnou veličinu X_i pro $i = 1, 2, \dots, 100$ se střední hodnotou $\mu = 5$ kg a směrodatnou odchylkou $\sigma = 1,5$ kg. Součet hmotnosti všech 100 balíků označme Y , tedy $Y = \sum_{i=1}^{100} X_i$. Hmotnosti jednotlivých balíků jsou nezávislé. Dále platí, že všechny veličiny X_i mají stejnou střední hodnotu a rozptyl. Jsou tedy splněny podmínky použití centrální limitní věty a lze psát

$$Y = \sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2).$$

Po dosazení obdržíme

$$Y = \sum_{i=1}^{100} X_i \sim N(100 \cdot 5, 100 \cdot 2,25), \text{ tedy } n = 100.$$

Hledaná pravděpodobnost je pak rovna

$$\begin{aligned} P(Y > 550) &= 1 - P(Y \leq 550) = 1 - P\left(\frac{Y - 500}{15} < \frac{550 - 500}{15}\right) = \\ &= 1 - P(U \leq 3,333) = 1 - F(3,333) = 1 - 0,99957 = 0,0004. \end{aligned}$$

3 ANALÝZA DAT Z VÝBĚROVÝCH ŠETŘENÍ

V této části použijeme znalosti z předchozích dvou kapitol věnovaných popisné statistice a teorii pravděpodobnosti k tomu, abychom se pokusili usuzovat na základě omezené informace obsažené v nějakém vzorku na celou populaci, která je cílem zkoumání. Takovými postupy se zabývá matematická statistika, která používá metody teorie pravděpodobnosti při analýze dat. Metody a postupy, které umožňují zobecňovat výsledky z části na celek, se nazývají statistickou indukcí.

Nejprve se budeme věnovat problémům spojeným s pořizováním dat, ze kterých by bylo možné činit spolehlivé závěry, a dále představíme základní metody zobecňování, kterými jsou bodový a intervalový odhad a testování statistických hypotéz.

3.1 Výběrová šetření

Žádná statistická metoda nemůže poskytnout hodnotné výsledky, jsou-li špatná data, se kterými pracujeme. O kvalitě dat rozhoduje již fáze statistického šetření, v rámci kterého je získáváme. První podstatnou okolností při jeho přípravě je rozsah **statistikého souboru**, který má být zkoumán (používáme také název **základní soubor**, **základní populace** nebo jen **populace**). Může to být soubor všech domácností v České republice, soubor všech voličů, kteří se mohou zúčastnit určitých voleb, nebo soubor všech příjemců nějaké dávky.

Ve vyčerpávajícím **zjišťování** se prosetřují veškeré jednotky základního souboru. Uvažujme rozsáhlý soubor – řekněme, že ho tvoří například všichni obyvatelé nějakého území. Realizace takového statistického šetření bude vyžadovat náročnou organizaci, množství pracovníků, značné finanční i časové náklady. I zpracování jeho výsledků může zabrat relativně dlouhé období. Na druhé straně dobře provedené vyčerpávající zjišťování může poskytnout detailní informace o každé jednotce v základním souboru a přesné charakteristiky sledovaných veličin. Tyto jeho přednosti však zmíněnou náročnost vyčerpávajícího šetření můžou využít, a tak nebývá realizováno často. Jako příklad uvedeme sčítání lidu, domů a bytů, které náš stát a jeho statistické orgány realizují ve zhruba desetiletých intervalech.

Provedení vyčerpávajícího zjišťování však mohou bránit i jiné okolnosti. V průmyslové výrobě mohou určité zkoušky výrobků při ověřování jejich vlastností – například jejich odolnosti při určité zátěži – vést až k jejich znehodnocení, a tím k vyřazení z dalšího použití. Při dotazování zákazníků obchodu jejich kompletní soubor jednoduše ani nemáme k dispozici.

V takovém případě lze přistoupit k nevyčerpávajícímu, **výběrovému zjišťování**. To znamená, že prosetřeny jsou pouze některé jednotky, které byly ze základního souboru vybrány. V ekonomické praxi je sledován cenový vývoj nebo vývoj životních nákladů. Rozsáhlá komoditní skladba trhu a liberální tvorba cen vytvářejí situace, kdy