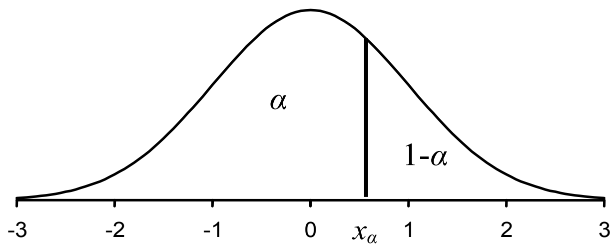
Z geometrického hľadiska rozdelí kolmica na os x v bode x_α obsah pod grafom funkcie hustoty na 2 časti, ľavá časť má obsah $1-\alpha$, vpravo ležiaca časť má obsah α . Je jasné, že α – kvantil je $(1-\alpha)$- kritická hodnota.



Obr. 3.8 Geometrický význam kvantilu

3.4 Niektoré rozdelenia pravdepodobnosti diskrétnej náhodnej premennej

3.4.1 Diskrétne rovnomerné rozdelenie R(*n*)

Náhodná premenná môže nadobúdať hodnoty 1, 2, 3, ... , *n*. Ak pravdepodobnosť výskytu ktorejkoľvek z nich je rovnaká t.j. $1/n$, hovoríme o rovnomernom rozdelení premennej. Číslo *n* je jediný parameter rozdelenia.

Napríklad, náhodná premenná *X* je počet padnutých bodov na hracej kocke. Výsledkom po hode kockou je jeden zo 6 javov, každý sa stane s pravdepodobnosťou $1/6$.

Tab. 3.3 Pravdepodobnostná tabuľka rovnomerného rozdelenia

x_i	1	2	3	4	...	n
p_i	$1/n$	$1/n$	$1/n$	$1/n$	...	$1/n$

Číselné charakteristiky: $E(X)=\frac{n+1}{2}$, $D(X)=\frac{n^2-1}{12}$

3.4.2 Alternatívne rozdelenie A(*p*)

Máme jeden pokus, v ktorom skúmaný jav A môže nastať alebo nenastať. Napríklad pohlavie vylosovanej osoby je mužské, vybraný výrobok je dobrý, športový klub vyhral zápas, študent urobil skúšku. Teda náhodná premenná X s pravdepodobnosťou p nadobúda hodnotu $x=1$ (ak jav nastal) a s pravdepodobnosťou $q=1-p$ hodnotu $x=0$ (ak jav nenastal). Číslo p je jediný parameter rozdelenia.

Tab. 3.4 Pravdepodobnostná tabuľka alternatívneho rozdelenia

x_i	0	1
p_i	1-p	p

Číselné charakteristiky:

$E(X)=p$, $D(X)=pq$, $S(X)=\frac{q-p}{\sqrt{pq}}$, $K(X)=\frac{1-6pq}{pq}$

3.4.3 Binomické rozdelenie Bi(*p*, *n*)

Súvisí s vetou o opakovaní nezávislých pokusov. Napríklad, zaujímame sa o počet chlapcov v trojdetnej rodine, alebo o počet dobrých výrobkov v kontrolnej vzorke. Ak ako náhodnú premennú X označíme počet výskytov javu A v sérii n nezávislých pokusov, a ak v každom pokuse nastane tento jav s pravdepodobnosťou p, potom náhodná premenná X má binomické rozdelenie pravdepodobnosti a hodnoty $x=0,1,2,\ldots,n$ nadobúda s pravdepodobnosťami:

$$p_x=P(X=x)=\binom{n}{x}p^x q^{n-x}, \qquad (3.21)$$

kde $p\in(0,1)$ a $q=1-p$. Čísla p, n sú parametre rozdelenia.

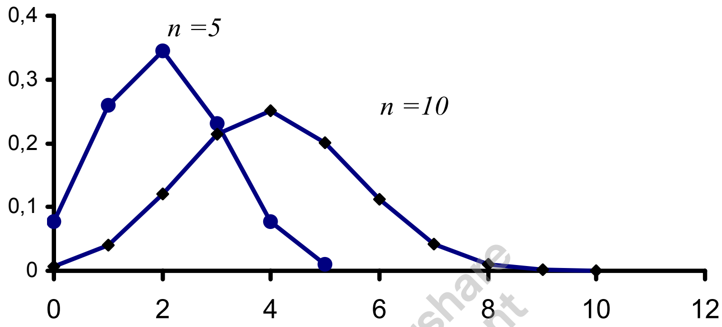
Číselné charakteristiky :

$E(X) = np$, $\quad D(X) = npq$, $\quad S(X) = \dfrac{q-p}{\sqrt{npq}}$, $\quad K(X) = \dfrac{1-6pq}{npq}$

Príklad 3.8

Znázornime pravdepodobnostné tabuľky a polygóny binomického rozdelenia pravdepodobnosti pre

a) $n = 4;\ \ p = 0{,}4;\ \ q = 0{,}6$ b) $n = 10;\ \ p = 0{,}4;\ \ q = 0{,}6$



Obr. 3.9 Polygóny rozdelenia pravdepodobnosti k Príkladu 3.8

Tab. 3.5 Pravdepodobnostné tabuľky k Príkladu 3.8

x_i	0	1	2	3	4	5	6	7	8	9	10
a) p_i	0,078	0,259	0,346	0,230	0,077	0,010	-	-	-	-	-
b) p_i	0,006	0,040	0,121	0,215	0,251	0,201	0,111	0,042	0,011	0,002	0,000

3.4.4 Poissonovo rozdelenie Po(λ)

Je špeciálne rozdelenie pre "maličké " pravdepodobnosti, jeho pravdepodobnostná funkcia sa používa na výpočet pravdepodobností zriedkavých udalostí.

Náhodná premenná X má Poissonovo rozdelenie s parametrom $\lambda > 0$, ak nadobúda hodnoty $x = 0, 1, 2, 3, \ldots$ s pravdepodobnosťami:

$$p_x = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \tag{3.22}$$

Číselné charakteristiky :

$E(X) = \lambda$, $D(X) = \lambda$, $S(X) = \frac{1}{\sqrt{\lambda}}$, $K(X) = \frac{1}{\lambda}$

Je to tiež rozdelenie počtu výskytov javu A za jednotku času. Ak priemerný počet objavení sa javu A za jednotku času je λ, potom pravdepodobnosť toho, že za jednotku času daný jav A nastane x - krát, má Poissonovo rozdelenie s parametrom λ. Príklady náhodných veličín s Poissonovým rozdelením:

Počet predmetov nájdených a odovzdaných denne v hypermarkete.

Počet telefonických hovorov v istom intervale.

Počet úrazov robotníkov za rok.

Poznámka 3.3

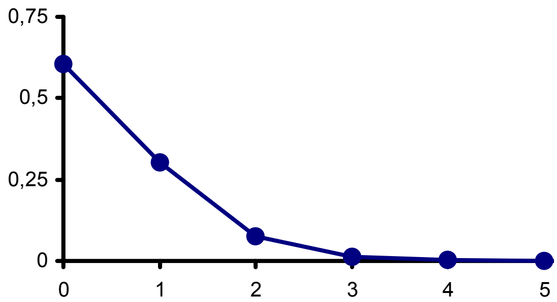
- Aj v binomickom rozdelení existujú zriedkavé javy, ale vždy sa doplňuje nastatie javu a nenastatie javu. Na mnohé javy sa však nedá aplikovať binomické rozdelenie. Napríklad, ak počas búrky bolo 10 bleskov, koľko bleskov nebolo? Ak sa od 10. hodiny do 12. hodiny narodilo 5 detí, koľko sa nenarodilo?
- Ak stredná hodnota a disperzia u neznámeho rozdelenia sú rovnaké, prichádza do úvahy práve Poissonovo rozdelenie.
- Poissonovo rozdelenie je rozdelenie „zriedkavých" javov.

Príklad 3.9

Pre $\lambda = 0{,}5$ uvádzame pravdepodobnostnú tabuľku a polygón rozdelenia pravdepodobnosti.

Tab. 3.6 Tabuľka rozdelenia pravdepodobnosti

x_i	0	1	2	3	4	5	...	...
p_i	0,607	0,303	0,076	0,013	0,002	1,6E-4	...	...



Obr. 3.10 Polygón rozdelenia pravdepodobnosti

Príklad 3.10

Povahu rozdelenia dobre ilustruje príklad z histórie. Nemecký štatistik Bartkiewicz zisťoval 10 rokov v 20 armádnych zboroch údaje o počte osôb zabitých úderom konského kopyta. Výsledky sú v tabuľke rozdelenia pravdepodobnosti, kde sme pravdepodobnosť nahradili relatívnou početnosťou:

Počet mŕtvych	**Počet prípadov**	**Relat. početnosť**	**Pravdepodobnosť**
0	109	109/200 = 0,545	0,543
1	65	65/200 = 0,325	0,331
2	22	22/200 = 0,11	0,101
3	3	3/200 = 0,015	0,021
4	1	1/200 = 0,005	0,003
	Spolu 200		

Stredná hodnota počtu mŕtvych za rok na jeden zbor je podľa tejto tabuľky 0,61. Ak sa tieto nehody dajú popísať Poissonovým rozdelením, potom λ = 0,61.

Z funkcie $p_x = P(X = x) = \dfrac{(0{,}61)^x \cdot e^{-0{,}61}}{x!}$ pre $x = 0, 1, 2, 3, 4$ vypočítame pravdepodobnosti $p_0 = 0{,}543$, $p_1 = 0{,}331$, $p_2 = 0{,}101$, $p_3 = 0{,}021$, $p_4 = 0{,}003$. Ak porovnáme posledné dva stĺpce tabuľky, vidíme, že toto rozdelenie dobre vystihlo rozdelenie počtu mŕtvych.

3.5 Niektoré rozdelenia pravdepodobnosti spojitej náhodnej premennej

3.5.1 Normálne rozdelenie (Gaussovo, Z – rozdelenie) N(μ,σ^2)

Je najdôležitejšie rozdelenie pravdepodobnosti z mnohých typov spojitých náhodných premenných, pretože mnohé náhodné premenné v technike, prírodných a spoločenských vedách majú toto rozdelenie, alebo dajú sa ním aproximovať iné spojité alebo diskrétne rozdelenia. Toto rozdelenie má aj náhodná premenná, ktorej hodnoty sú tvorené súčtom veľkého počtu malých a vzájomne nezávislých veličín. Normálne rozdelenie je závislé od dvoch parametrov μ, σ, označuje sa $\mathrm{N}(\mu, \sigma^2)$. Určujú graf funkcie hustoty rozdelenia (Obr. 3.11, Obr. 3.12).

Spojitá náhodná premenná X má **normálne rozdelenie pravdepodobnosti** s parametrami μ, σ^2, ak jej **funkcia hustoty** má tvar:

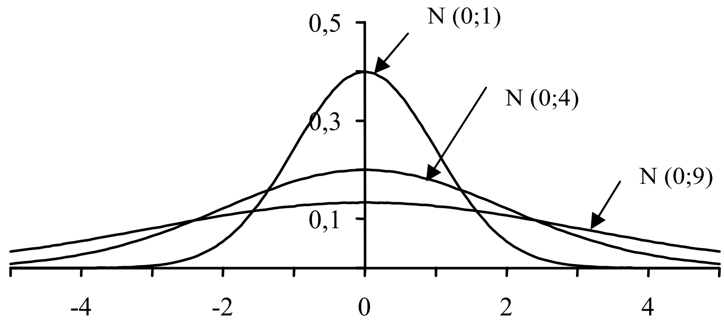
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \; x \in R \tag{3.23}$$

Distribučná funkcia rozdelenia N(μ, σ^2) má tvar

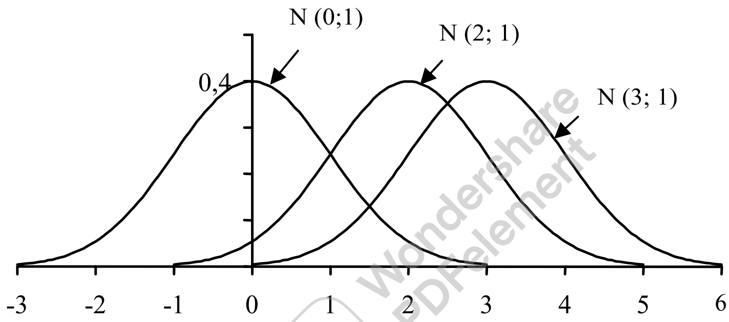
$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \tag{3.24}$$

Číselné charakteristiky :

$$E(X)=\mu, \quad D(X)=\sigma^2, \quad S(X)=0, \quad K(X)=0$$



Obr. 3.11 Graf funkcie hustoty normálneho rozdelenia pre rôzne σ

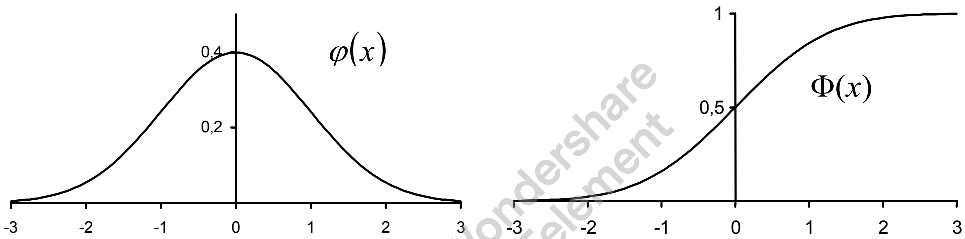


Obr. 3.12 Graf funkcie hustoty normálneho rozdelenia pre rôzne μ

Špeciálne pre $\mu=0$ a $\sigma^2=1$ dostaneme **normované (štandardizované) normálne rozdelenie,** ktoré sa označuje $N(0,1)$. Pre toto rozdelenie sa označujú funkcia hustoty $\varphi(x)$, distribučná funkcia $\Phi(x)$, graf funkcie hustoty sa volá Gaussova krivka (Obr. 3.13).

Vlastnosti normovaného normálneho rozdelenia:

- $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, $x \in R$, (3.25)
- $\varphi(x)$ je párna, t.j. $\varphi(-x)=\varphi(x)$, (3.26)
- v bode $x = 0$ má maximum $\varphi(0)=\frac{1}{\sqrt{2\pi}} \doteq 0{,}4$,
- inflexné body má v číslach $x=-1,\ x=1$,
- os x je asymptota grafu funkcie $\varphi(x)$,
- $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$, (3.27)
- $\Phi(-x)=1-\Phi(x)$. (3.28)



Obr. 3.13 Funkcia hustoty $\varphi(x)$ a distribučná funkcia $\Phi(x)$ rozdelenia N(0,1)

V štatistických tabuľkách sa nachádzajú funkcia hustoty $\varphi(x)$, distribučná funkcia $\Phi(x)$, kvantily normovaného normálneho rozdelenia, v programových štatistických balíkoch sú zaradené funkcia hustoty, distribučná funkcia, kvantily ľubovoľného normálneho rozdelenia. Vzhľadom na koeficienty asymetrie a špicatosti sa toto rozdelenie chápe ako základné, s ktorým sa porovnávajú iné rozdelenia (pozri časť číselné charakteristiky náhodnej premennej).

Vzťahy medzi ľubovoľným $N(\mu,\sigma^2)$ a $N(0,1)$:

- $f(x)=\frac{1}{\sigma}\,\varphi\left(\frac{x-\mu}{\sigma}\right)$ (3.29)
- $F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right)$ (3.30)
- $P(a\le X<b)=F(b)-F(a)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$ (3.31)

Platnosť týchto vzťahov dokážeme, ak vo funkcii hustoty rozdelenia $\mathrm{N}\left(\mu,\sigma^2\right)$ zvolíme substitúciu $z=\frac{x-\mu}{\sigma}$.

Príklad 3.11

Náhodná premenná má normálne rozdelenie N(1,1). Určite :

a) hodnoty $f(1)$, $F(1)$,

b) pravdepodobnosť, že premenná nadobudne hodnotu z intervalu $(1,2)$,

c) pravdepodobnosť, že absolútna hodnota náhodnej premennej nadobudne hodnoty väčšie ako 2.

Riešenie Využijeme štatistické tabuľky v [4], [15], [16], prípadne štatistické funkcie ponúkané EXCELom a predchádzajúce vlastnosti normálneho rozdelenia.

a) $f(1)=\varphi\left(\frac{1-1}{2}\right)=\varphi(0)=0{,}3989$ $\quad F(1)=\Phi\left(\frac{1-1}{2}\right)=\Phi(0)=0{,}5$

b) $P(1<X<2)=\Phi\left(\frac{2-1}{1}\right)-\Phi\left(\frac{1-1}{1}\right)=\Phi(1)-\Phi(0)=0{,}8413-0{,}5=0{,}3413$

c) $P(|X|>2)=1-P(|X|\le 2)=1-P(-2\le X\le 2)=1-\Phi\left(\frac{2-1}{1}\right)+\Phi\left(\frac{-2-1}{1}\right)=$

$=1-\Phi(1)+\Phi(-3)=1-\Phi(1)+1-\Phi(3)=1-0{,}8413+1-0{,}9987=0{,}16$

Príklad 3.12

Praktické informácie o $\mathrm{N}(m,\sigma^2)$ dostaneme odpoveďou na otázku, aká je $P(|X-m|<\varepsilon)$, kde ε je vopred dané, ľubovoľné kladné číslo.

$$P(|X-m|<\varepsilon)=P(-\varepsilon<X-m<\varepsilon)=P(m-\varepsilon<X<m+\varepsilon)=$$

$$\Phi\left(\frac{m+\varepsilon-m}{\sigma}\right)-\Phi\left(\frac{m-\varepsilon-m}{\sigma}\right)=\Phi\left(\frac{\varepsilon}{\sigma}\right)-\Phi\left(-\frac{\varepsilon}{\sigma}\right)=\Phi\left(\frac{\varepsilon}{\sigma}\right)-1+\Phi\left(\frac{\varepsilon}{\sigma}\right)=$$

$$=2\Phi\left(\frac{\varepsilon}{\sigma}\right)-1\,.$$

Zvoľme za ε postupne hodnoty $\sigma,\ 2\sigma,\ 3\sigma$. Postupne dostaneme nasledujúce tvrdenia:

$$P(|X-m|<\sigma)=2\Phi(1)-1=2.0{,}8413-1=0{,}6826$$

$$P(|X-m|<2\sigma)=2\Phi(2)-1=2.0{,}9972-1=0{,}9544$$

$$P(|X-m|<3\sigma)=2\Phi(3)-1=2.0{,}9987-1=0{,}9974$$

To znamená, že z celkovej plochy ohraničenej grafom funkcie $f(x)$ a osou x sa

- nad intervalom $(m-\sigma, m+\sigma)$ nachádza 68,26% plochy,
- nad intervalom $(m-2\sigma, m+2\sigma)$ nachádza 95,44% plochy,
- nad intervalom $(m-3\sigma, m+3\sigma)$ nachádza 99,74% plochy, t.j. takmer všetky hodnoty náhodnej premennej s $N(m,\sigma^2)$ ležia v tomto intervale. Táto posledná vlastnosť sa volá **„ pravidlo troch sigma“**. Využíva sa v praxi na prvý odhad štandardnej odchýlky pre náhodnú premennú, u ktorej predpokladáme normálne rozdelenie. Stačí, ak šírku intervalu, ktorý je určený najmenšou a najväčšou hodnotou premennej vydelíme šiestimi, výsledok je odhadom parametra σ.

EXCEL

Po voľbe ***Prilepiť funkciu/Štatistické*** sú pre normálne rozdelenie k dispozícii štatistické tabuľky, ich prehľad je v Tab. 3.6.

Tab. 3.6 Tabuľky normálneho rozdelenia v EXCELi

Označenie funkcie a zadávané parametre	**Výsledok**
NORMDIST (x, μ, σ, 1)	$F(x)$, hodnota distribučnej funkcie v čísle x pre N(μ, σ^2)
NORMDIST (x, μ, σ, 0)	$f(x)$, hodnota funkcie hustoty v čísle x pre N(μ, σ^2)
NORMINV (p, μ, σ)	$p\times$100%-ný kvantil pre N(μ, σ^2), p- pravdepodobnosť
NORMSDIST (z)	$\Phi(z)$, hodnota distribučnej funkcie v čísle z pre N(0, 1)
NORMSINV (p)	$p\times$100%-ný kvantil pre N(0, 1), p- pravdepodobnosť

Príklad 3.13

Vyriešime Príklad 3.11 s využitím EXCEL-u.

a) ***Prilepiť funkciu / Štatistické/*NORMDIST(1, 1, 1, 0)**.

Výsledok $F(1) = 0{,}3989$

***Prilepiť funkciu / Štatistické/*NORMDIST(1, 1, 1, 1** .

Výsledok $F(1) = 0{,}5$

b) ***Prilepiť funkciu / Štatistické/*NORMDIST(2, 1, 1, 1)**.

Výsledok $F(2) = 0{,}8413$

***Prilepiť funkciu / Štatistické/*NORMDIST(1, 1, 1, 1)**.

Výsledok $F(1) = 0{,}5$. $F(2) - F(1) = 0{,}3413$

c) ***Prilepiť funkciu / Štatistické/*NORMDIST(-2, 1, 1, 1)**.

Výsledok $F(-2) = 0{,}0013$

$$P(\,|X| > 2\,) = 1 - (F(2) - F(-2)) = 1 - F(2) + F(-2) = = 1 - 0{,}8413 + 0{,}0013 = 0{,}16$$

Príklad 3.14

Čas potrebný na vypracovanie kontrolnej práce má normálne rozdelenie s priemernou dobou vypracovania 110 minút a štandardnou odchýlkou 20 minút.

a) Koľko percent študentov dokončí prácu do 2 hodín?

b) Koľko času by bolo treba dať na prácu, aby ju dokončilo 90 % študentov?

Riešenie Náhodná premenná je čas potrebný na vypracovanie práce.

a) Zaujíma nás $P(X \le 120) =$NORMDIST(120;110;20;1)=0,6914, t.j. 69% študentov spraví prácu do 2 hodín.

b) V tomto prípade nás zaujíma ten čas t, pre ktorý platí $P(X \le t) = 0{,}9$, čo je 90%-ný kvantil tohto rozdelenia.

Preto $t =$ NORMINV(0,9;110;20)=135,6, t.j. študenti potrebujú na prácu 136 minút.

3.6 Aproximácia diskrétnych rozdelení normálnym

V tejto časti je stručne vysvetlený význam normálneho rozdelenia ako limitného rozdelenia pre mnohé náhodné premenné, pričom sú na ne kladené len dosť všeobecné predpoklady. Základnou vetou v štatistike je **centrálna limitná veta,** ktorá hovorí:

Ak náhodné premenné $X_1, X_2, X_3, \ldots, X_n$ sú nezávislé náhodné premenné s rovnakým rozdelením pravdepodobnosti t.j. $E(X_i) = \mu,\ D(X_i) = \sigma^2$, potom aritmetický priemer $\overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ je náhodná premenná, ktorá má pre $n \to \infty$ normálne rozdelenie s parametrami μ a σ^2 / n t.j. N($\mu, \sigma^2/n$).

Poznámka 3.4

- Ak pomocou $\overline{X}$ vytvoríme normovanú náhodnú premennú $Z = \dfrac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$, tak táto má rozdelenie $\mathrm{N}(0,1)$.

- Aproximácia rozdelenia premennej $\overline{X}$ je tým lepšia, čím je väčšie n.

Najznámejší a najpoužívanejší tvar centrálnej limitnej vety je integrálna **Moivrova - Laplaceova veta**, pomocou ktorej vieme **aproximovať binomické rozdelenie normálnym**. Táto veta hovorí:

Nech náhodná premenná X má binomické rozdelenie s parametrami n a p. Potom limitným rozdelením tohto rozdelenia pre $n \to \infty$ je normálne rozdelenie s parametrami $\mu = np$ a $\sigma^2 = npq$.

Poznámka 3.5

- Ak vytvoríme pomocou X normovanú náhodnú premennú $\dfrac{X-np}{\sqrt{npq}}$, môžeme tvrdenie vety vyjadriť v tvare $\lim\limits_{n\to\infty} P\left(\dfrac{X-np}{\sqrt{npq}} \le z\right) = \Phi(z)$, pre ľubovoľné $z \in R$.
- Pri dostatočne veľkom n môžeme vyjadriť pravdepodobnosť toho, že náhodná premenná s binomickým rozdelením pravdepodobnosti, nadobudne hodnoty z intervalu $\langle a,b)$ takto:

$$P(a \le X < b) \doteq F(b) - F(a) = \Phi\left(\frac{b-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np}{\sqrt{npq}}\right). \qquad (3.32)$$

- Táto aproximácia sa používa vtedy, ak $np > 5$ aj $nq > 5$. Používa sa aj kritérium $npq > 9$.

  Ak $p = 0{,}5$, sú nerovnosti splnené už pre $n > 10$.

  Ak $p = 0{,}8$ alebo $p = 0{,}2$, požadujeme $n > 25$.

  Ak $p = 0{,}05$ alebo $p = 0{,}95$, má byť $n > 100$.

- Ak sa pri tejto aproximácii diskrétneho rozdelenia spojitým zaujímame o pravdepodobnosť konkrétnej hodnoty x binomického rozdelenia, teda $P(X = x)$, použijeme funkciu hustoty normálneho rozdelenia:

$$P(X = x) \doteq f(x) = \frac{1}{\sigma}\varphi\left(\frac{x - np}{\sqrt{npq}}\right). \tag{3.33}$$

Pri malých hodnotách p, prípadne pri hodnotách p blízkych jednotke, táto aproximácia nie je dobrá a jej zlepšenie sa dosahuje len zvyšovaním počtu pokusov v Bernoulliho schéme. V tejto situácii dobrú aproximáciu binomického rozdelenia dáva Poissonovo rozdelenie. Platí nasledujúce tvrdenie.

Nech náhodná premenná X má binomické rozdelenie s parametrami n a p. Potom limitným rozdelením tohto rozdelenia pre $n \to \infty$ a súčasne $p \to 0$ je Poissonovo rozdelenie s parametrom $\lambda = np$.

Poznámka 3.6

Túto aproximáciu sa odporúča použiť ak $p < 0{,}1$ alebo $p > 0{,}9$.

Príklad 3.15

Poisťovňa poistila 1000 ľudí rovnakého veku. Pravdepodobnosť úmrtia v priebehu roka je pre každého z nich 0,008. Každý poistenec zaplatil 1200 Sk. V prípade jeho úmrtia dostanú jeho príbuzní 80 000 Sk. Aká je pravdepodobnosť, že poisťovňa utrpí stratu?

Riešenie Poisťovňa vybrala na poistnom 1 200 000 Sk. Z týchto peňazí je schopná vyplatiť $1200000/80000 = 15$ pozostalých. Teda poisťovňa utrpí stratu, ak v priebehu roka zomrie viac ako 15 poistencov. Preto nás zaujíma pravdepodobnosť tohto javu. Náhodná premenná X je počet zomretých poistencov za rok. Riadi sa Bi(0,008;1000).

Pretože $np = 1000 \cdot 0{,}008 = 8$ a $nq = 1000 \cdot 0{,}992 = 992$ môžeme binomické rozdelenie aproximovať normálnym s parametrami

$\mu = 8$; $\sigma^2 = npq = 1000 \cdot 0{,}008 \cdot 0{,}992 = 7{,}936$,   t.j. $N(8; 7{,}936)$.

$P(X > 15) = 1 - P(0 \leq X \leq 15) \doteq 1 - F(15) + F(0) = 1 - \text{NORMDIST}(15; 8; \sqrt{7{,}936}; 1)$

$+ \text{NORMDIST}(0; 8; \sqrt{7{,}936}; 1) = 0{,}0089$.

Pretože pravdepodobnosť $p = 0{,}008$ je malá, môžeme binomické rozdelenie aproximovať aj Poissonovým rozdelením s parametrom $\lambda = np = 8$, t.j. Po(8).

$$P(X > 15) = 1 - P(0 \leq X \leq 15) \doteq e^{-8} \cdot \left(1 + 8 + \frac{8^2}{2!} + \frac{8^3}{3!} + \cdots + \frac{8^{15}}{15!}\right) = 1 - 0{,}9944 = 0{,}0056 .$$

Pravdepodobnosť, že poisťovňa bude stratová je menšia ako 1% ako sme zistili oboma aproximáciami.

3.7 Rozdelenia funkcií náhodných premenných

Pri výberovom skúmaní v štatistike sa používajú ďalšie rozdelenia, ktoré vznikli ako funkcie iných rozdelení. Najčastejšie sa pri tomto skúmaní využíva

a) χ^2- rozdelenie (chí-kvadrát rozdelenie)

b) Studentovo rozdelenie (t-rozdelenie)

c) Fisherovo rozdelenie (F-rozdelenie).

3.7.1 χ^2 - rozdelenie

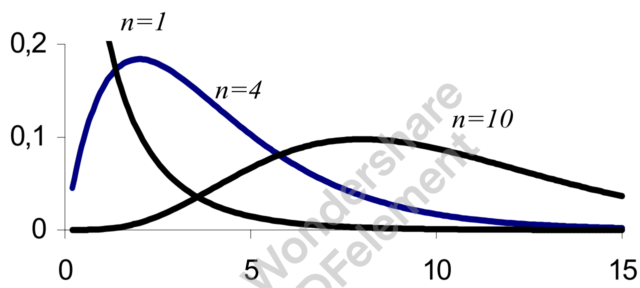
Pri konštrukcii tohto rozdelenia použijeme n nezávislých náhodných premenných $X_1, X_2, ..., X_n$, z ktorých každá má normované normálne rozdelenie $N(0,1)$. Súčet druhých mocnín (štvorcov) týchto náhodných premenných je náhodná premenná $\chi^2 = X_1^2 + X_2^2 + ... + X_n^2$. Jej rozdelenie sa nazýva **chí - kvadrát rozdelenie**.

Poznámka 3.7

Funkčný predpis funkcie hustoty rozdelenia nebudeme uvádzať.

Vlastnosti:

1. Počet sčítancov ***n*** určuje geometrický tvar rozdelenia. Číslo ***n*** sa volá **počet stupňov voľnosti** (d.f., degrees of freedom). Je to jediný parameter tohto rozdelenia.
2. $E(X)=n$
3. $D(X)=2n$
4. Koeficient asymetrie je $\sqrt{8/n}$, je to rozdelenie nesymetrické.
5. Koeficient špicatosti je $12/n$.
6. S rastúcim n sa oba koeficienty blížia k nule, pre $n \to \infty$ je χ^2- rozdelenie symetrické.
7. Pre $n > 30$ možno toto rozdelenie aproximovať normálnym rozdelením $\mathrm{N}(n, 2n)$.



Obr. 3.14 Funkcia hustoty χ^2- rozdelenia

3.7.2 Studentovo rozdelenie (t - rozdelenie)

Pomocou nezávislých náhodných premenných X a χ^2, kde X má normované normálne rozdelenie $\mathrm{N}(0,1)$ a χ^2 má χ^2- rozdelenie s n stupňami voľnosti vytvoríme novú náhodnú premennú T

$$T = \frac{X}{\sqrt{\dfrac{\chi^2}{n}}}$$

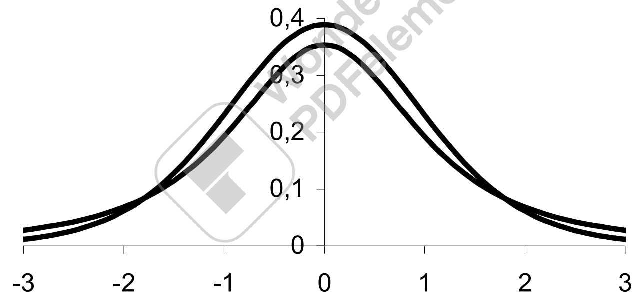
Jej rozdelenie sa nazýva **Studentovo rozdelenie s *n* stupňami voľnosti**.

Poznámka 3.8

Funkčný predpis funkcie hustoty rozdelenia nebudeme uvádzať.

Vlastnosti:

1. Číslo ***n*** – počet stupňov voľnosti je jediný parameter tohto rozdelenia, ktorý určuje aj geometrický tvar rozdelenia.
2. $E(T) = 0$
3. Krivka Studentovho rozdelenia je symetrická podľa $t = 0$ a veľmi sa podobá na krivku normovaného normálneho rozdelenia, je však plochšia. To znamená, že hodnoty vzdialenejšie od nuly majú väčšiu pravdepodobnosť nastatia ako pri normálnom rozdelení.
4. Pre $n > 30$ možno Studentovo rozdelenie aproximovať normovaným normálnym rozdelením.



Obr. 3.15 Funkcia hustoty t – rozdelenia

3.7.3 Fisherovo rozdelenie (F – rozdelenie)

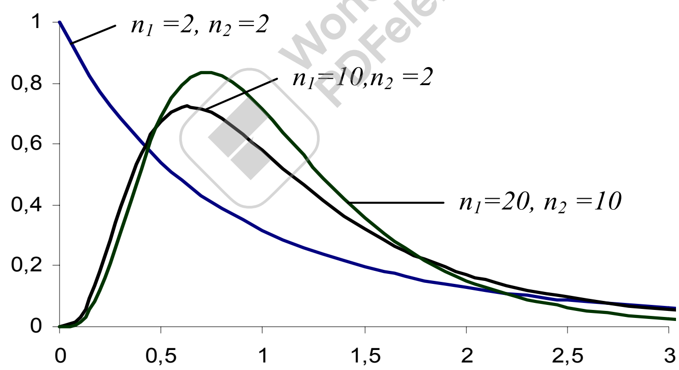
Je vytvorené pomocou dvoch náhodných premenných χ_1^2, χ_2^2, kde χ_1^2 má χ^2 - rozdelenie s n_1 stupňami voľnosti a χ_2^2 má χ^2 - rozdelenie s n_2 stupňami voľnosti. Ich podiel je náhodná premenná F,

$$F = \frac{\frac{\chi_1^2}{n_1}}{\frac{\chi_2^2}{n_2}} \quad ,$$

ktorej rozdelenie sa nazýva **Fisherovo rozdelenie s n_1 a n_2 stupňami voľnosti**.

Vlastnosti:

1. Parametrami rozdelenia sú dve čísla n_1 a n_2, ich poradie sa nemôže vymeniť.
2. Stupne voľnosti určujú geometrický tvar funkcie hustoty.



Obr. 3.16 Funkcia hustoty F – rozdelenia

EXCEL

Pri hľadaní kvantilov spomínaných rozdelení môžeme využiť ponuku štatistických funkcií v EXCELi. Treba podotknúť, že nie sú definované jednotne. K lepšej orientácii nám určite pomôže nasledujúca tabuľka. Okrem použitých označení štatistických funkcií, pomocou ktorých nájdeme potrebné kvantily, sú v nej zadané i hodnoty pravdepodobnosti, ktoré treba zadať pre príslušnú hodnotu hľadaného kvantilu.

Tab. 3.7 Hľadanie kvantilov v EXCELi

Funkcia	**Zadaná pravdepodobnosť**	**Nájdený kvantil**
NORMSINV	α	z_{α}
	1-α	$z_{1-\alpha}$
TINV	α	$t_{1-\alpha/2}$
	2α	$t_{1-\alpha}$
CHIINV	α	$\chi^2_{1-\alpha}$
	1-α	χ^2_{α}
FINV	α	$F_{1-\alpha}$
	1-α	F_{α}

V nasledujúcej tabuľke je urobený prehľad funkcií, ktoré budeme najčastejšie používať.

Tab. 3.8 Prehľad tabuliek ponúkaných EXCELom

Funkcia	**Parametre**	**Význam funkcie**
NORMSDIST	z	Distribučná funkcia rozdelenia N(0, 1) NORMSDIST(z) = Φ(z) = $P(X< z)$
NORMDIST	$x, \mu, \sigma, 1$	Distribučná funkcia rozdelenia $N(\mu, \sigma^2)$ NORMDIST $(x, \mu, \sigma, 1) = F(x) = P(X < x)$
NORMDIST	$x, \mu, \sigma, 0$	Funkcia hustoty rozdelenia $N(\mu, \sigma^2)$ NORMDIST $(x, \mu, \sigma, 0) = f(x)$
NORMSINV	p	Kvantily rozdelenia N(0, 1), je to inverzná funkcia k distrib. funkcii rozdelenia N(0, 1), určí číslo z_p na osi tak, aby $P(X<z_p) = p$ NORMSINV(p) = z_p $\Leftrightarrow$ $P(X<z_p) = p$
NORMINV	p, μ, σ	Kvantily rozdelenia $N(\mu, \sigma^2)$, je to inverzná funkcia k distrib. funkcii rozdelenia $N(\mu, \sigma^2)$, určí číslo u_p na osi tak, aby $P(X<u_p) = p$ NORMINV$(p, \mu, \sigma) = u_p$ $\Leftrightarrow$ $P(X<u_p) = p$
TDIST	$x, n, 1$ n – stupne voľnosti $x \geq 0$	Pre Studentove rozdelenie s n stupňami voľnosti určí pravdepodobnosť p = TDIST(x, n, 1) = $= P(X > x) = 1 - F(x)$
TDIST	$x, n, 2$ n – stupne voľnosti $x \geq 0$	Pre Studentove rozdelenie s n stupňami voľnosti určí pravdepodobnosť p = TDIST(x, n, 2) = $= P(\|X\| > x)$
TINV	p, n n – stupne voľnosti	Je to inverzná funkcia k funkcii TDIST(x, n, 2), určí číslo t na osi tak, aby $P(\|X\|>t) = p$ TINV$(p, n) = t \Leftrightarrow P(\|X\| >t) = p$
CHIDIST	x, n n – stupne voľnosti $x \geq 0$	Pre χ^2-rozdelenie s n stupňami voľnosti určí pravdepodobnosť p = CHIDIST$(x, n) = P(X > x) = 1 - F(x)$
CHIINV	p, n n – stupne voľnosti	Kritické hodnoty χ^2 –rozdelenia. Je to inverzná funkcia k funkcii CHIDIST, určí číslo χ^2 na osi tak, aby $P(X > \chi^2) = p$ CHIINV$(p, n) = \chi^2 \Leftrightarrow P(X > \chi^2) = p$
FDIST	x, n_1, n_2 n_1, n_2-stupne voľnosti $x \geq 0$	Pre Fisherove F-rozdelenie s n_1 a n_2 stupňami voľnosti určí pravdepodobnosť p = FDIST(x, n_1, n_2) = $P(X > x) = 1 - F(x)$
FINV	p, n_1, n_2 n_1, n_2-stupne voľnosti	Kritické hodnoty F –rozdelenia. Je to inverzná funkcia k funkcii FDIST, určí číslo f na osi tak, aby $P(X > f) = p$ FINV$(p, n_1, n_2) = f \Leftrightarrow P(X > f) = p$

Príklad 3.16 Nájdite

a) 95% kvantil normovaného normálneho rozdelenia (ozn. $z_{0,95}$),

b) 95% kvantil normálneho rozdelenia so strednou hodnotu μ =2 a disperziou σ^2 = 0,81 (ozn. $u_{0,95}$),

c) 90% kvantil Studentovho t - rozdelenia s 8-mi stupňami voľnosti (ozn. $t_{0,90;8}$),

d) 97,5% kvantil χ^2 - rozdelenia s 10-mi stupňami voľnosti (ozn. $\chi^2_{0,975;10}$),

e) 95% kvantil F - rozdelenia s n_1 = 8 a n_2 = 6 stupňami voľnosti (ozn. $F_{0,95;8;6}$),

f) 5% kvantil Studentovho t - rozdelenia s 6-mi stupňami voľnosti (ozn. $t_{0,05;6}$).

Riešenie

a) Na vstupnom paneli EXCELU vyberieme ponuku ***prilepiť funkciu*** f_x (alebo =), na ľavej strane z ponuky funkcií vyberieme ***viac funkcií,*** zvolíme ***štatistické,*** vyhľadáme **NORMSINV**, vpíšeme pravdepodobnosť 0,95. Vo výstupe máme $z_{0,95} = 1,644853$.

b) Postupne volíme ***=/ viac funkcií/ štatistické/*** **NORMINV** vpíšeme pravdepodobnosť 0,95, strednú hodnotu μ =2 a štandardnú odchýlku σ = 0,9. Vo výstupe máme $u_{0,95} = 3,48037$.

c) Funkcia **TINV** neponúka priamo kvantily rozdelenia, ale dá sa úpravou zadávanej pravdepodobnosti získať hľadaný kvantil. Ak potrebujeme *p*% kvantil, vpíšeme hodnotu pravdepodobnosti 2(1-*p*) (platí pre $p > 0,5$). Teda postupne volíme ***=/ viac funkcií/ štatistické/*** **TINV** vpíšeme pravdepodobnosť 2(1- 0,9) = 0,2 a 8 stupňov voľnosti. Vo výstupe je $t_{0,90;8} = 1,397$.

d) Funkcia **CHIINV** neponúka priamo kvantily rozdelenia, ale dá sa úpravou zadávanej pravdepodobnosti získať hľadaný kvantil. Ak potrebujeme *p*% kvantil, vpíšeme hodnotu pravdepodobnosti (1-p). Volíme ***=/ viac funkcií/ štatistické/*** **CHIINV** vpíšeme pravdepodobnosť 1 - 0,975 = 0,025 a 10 stupňov voľnosti. Vo výstupe je $\chi_{0,975;10} = 20,4832$.

e) Postup je rovnaký ako pri funkcii **CHIINV**. Na výpočet kvantilov tiež zmeníme zadávanú pravdepodobnosť na 1- *p*. Volíme ***=/ viac funkcií/ štatistické/***

FINV vpíšeme pravdepodobnosť 1- 0,95 = 0,05; 8 a 6 stupňov voľnosti. Vo výstupe je $F_{0,95;8;6} = 3{,}9715$.

f) Pre $p < 0{,}5$ nie je možné hneď použiť model c). Využijeme, že hustota pravdepodobnosti t-rozdelenia je symetrická podľa *t=0* a pre kvantily platí $t_p = -t_{1-p}$. Preto: $t_{0,05;6} = -t_{0,95;6} = -1{,}9431$.

Príklady na precvičenie

3.17 Pre náhodnú premennú X – doba čakania na električku z Príkladov 3.4 a 3.5 nájdite strednú hodnotu a rozptyl.

3.18 Strelec strieľa na cieľ, pričom má k dispozícii 4 náboje. Každý zásah znamená zisk 5 bodov, každé minutie stratu 2 bodov. Určite tabuľku rozdelenia pravdepodobnosti počtu získaných bodov, ak pravdepodobnosť zásahu pri každom výstrele je 0,8. Určite strednú hodnotu a strednú kvadratickú odchýlku počtu získaných bodov.

3.19 Podnik vyexpedoval zásielku s 5 výrobkami. Pravdepodobnosť, že sa jeden výrobok počas prepravy poškodí je $p = 0{,}2$. Aká je pravdepodobnosť, že sa počas prepravy nepoškodí ani jeden výrobok? Určite strednú hodnotu a rozptyl počtu poškodených výrobkov.

3.20 Distribučná funkcia spojitej náhodnej premennej má tvar:

$$F(x) = \begin{cases} 0 & x \le 0 \\ 2cx & 0 < x \le 2 \\ 1 & x > 2 \end{cases}$$

Určite c, $E(X)$, $D(X)$, $P(1 \le X < 3)$.

3.21 Každý z 300 uchádzačov o prácu robí test, pozostávajúci z 3 častí. Uchádzač bude prijatý, ak urobí všetky 3 časti testu. Zo skúseností s takýmito testami je známe, že prvú časť testu urobí 90% , druhú 95% a tretiu 85% uchádzačov. Aká je pravdepodobnosť, že test urobí

a) práve 200 uchádzačov

b) aspoň 200 uchádzačov

3.22 Vo veľkej skupine ľudí je 5% chorých na chrípku. Aká je pravdepodobnosť, že v skupine 400 ľudí je $5 \pm 0{,}25$ % chorých na chrípku?

3.23 Telefónna ústredňa obsluhuje 3000 účastníkov. Pravdepodobnosť, že nejaký účastník bude v priebehu hodiny telefonovať, je p=0,002. Vypočítajte pravdepodobnosť, že v priebehu hodiny budú telefonovať 4 účastníci.

3.24 Pravdepodobnosť, že výrobok neprejde kontrolou je 0,05. Určite pravdepodobnosť, že medzi 500 náhodne vybranými výrobkami bude

a) práve 20 výrobkov, ktoré neprejdú kontrolou

b) od 10 do 30 výrobkov, ktoré neprejdú kontrolou.

3.25 Pravdepodobnosť, že sa na gymnázium hlási žiak s vyznamenaním zo ZŠ je 0,7. Aká je pravdepodobnosť, že medzi 500 prihlásenými je viac ako 360 žiakov s vyznamenaním.