

ROZDELENIA NÁHODNEJ PREMENNEJ

Modely rozdelenia pravdepodobnosti náhodnej premennej

Konkrétne deje môžeme častokrát rozdeliť do určitých skupín, ktoré sú popísané určitým spoločným zákonom rozdelenia pravdepodobnosti.

Náhodné premenné zaradené do jednej skupiny a ich rozdelenia pravdepodobnosti popisujeme modelmi s určitými parametrami. *Voľbou parametrov je náhodná premenná jednoznačne určená.*

Pomocou vhodných metód vieme rozhodnúť, či zvolený model dostatočne presne popisuje zákonitosti daného deja.

Vybrané rozdelenia diskkrétnej náhodnej premennej

Rozdelenie diskkrétnej náhodnej premennej možno popísať pravdepodobnostnou funkciou $p(x)$, distribučnou funkciou $F(x) = P(X \leq x)$, resp. používame ich vybrané číselné charakteristiky

- **Rovnomerné diskkrétne rozdelenie (Uniform discrete distribution)**
- **Alternatívne rozdelenie**
- **Binomické rozdelenie (Binomial distribution)**
- **Poissonovo rozdelenie (Poisson distribution)**
- **Hypergeometrické rozdelenie (Hypergeometric distribution)**

Rovnomerné diskrétne rozdelenie (Uniform discrete distribution)

$$X \sim R(n)$$

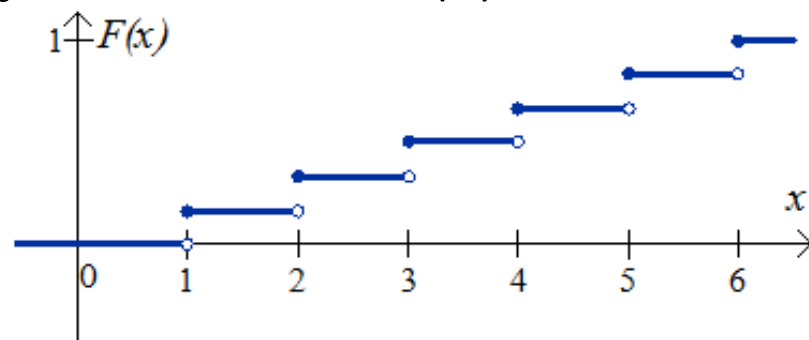
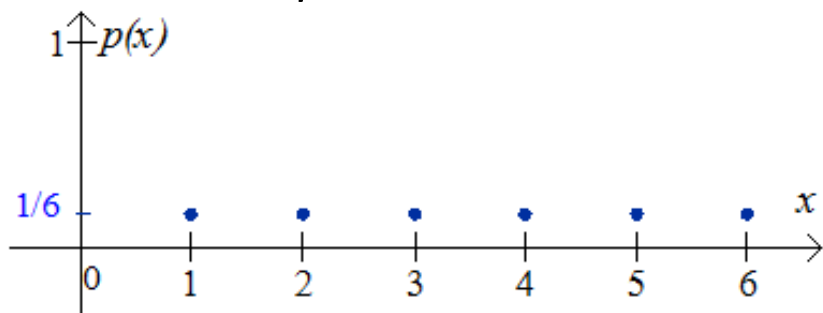
- náhodná premenná X môže nadobúdať hodnoty $1, 2, 3, \dots, n$, každú s rovnakou pravdepodobnosťou výskytu $\frac{1}{n}$;
- modelovanie úloh klasickej pravdepodobnosti

$X \sim R(n)$	
$p(x)$	$= \begin{cases} \frac{1}{n} & \text{pre } x_1, x_2, \dots, x_n; \\ 0 & \text{inak} \end{cases}$

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

$$D(X) = \frac{n^2-1}{12}$$

Pravdepodobnostná a distribučná funkcia rozdelenia $R(6)$



Alternatívne rozdelenie (Bernoulli distribution)

$$X \sim A(\pi)$$

- náhodná premenná X , udávajúca počet úspechov v jednom pokuse, parameter rozdelenia π vyjadruje pravdepodobnosť úspechu.

Náhodná premenná X priradí číslo 1, ak jav nastane a číslo 0, ak jav nenastane. Nadobúda hodnoty 0 alebo 1 $\rightarrow$ názov: nula – jednotkové rozdelenie.

$X \sim A(\pi)$	
$p(x)$	$= \begin{cases} \pi^x (1 - \pi)^{1-x} & \text{pre } x = 0, 1, \\ 0 & \text{inak} \end{cases}$

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

$$D(X) = \pi(1 - \pi)$$

Príklady alternatívneho rozdelenia pravdepodobnosti :

- počet nepodarkov pri náhodnom výbere jedného výrobku;
- počet zásahov pri jednom výstrele;
- počet spojení pri jednom vytočení telefónneho čísla;

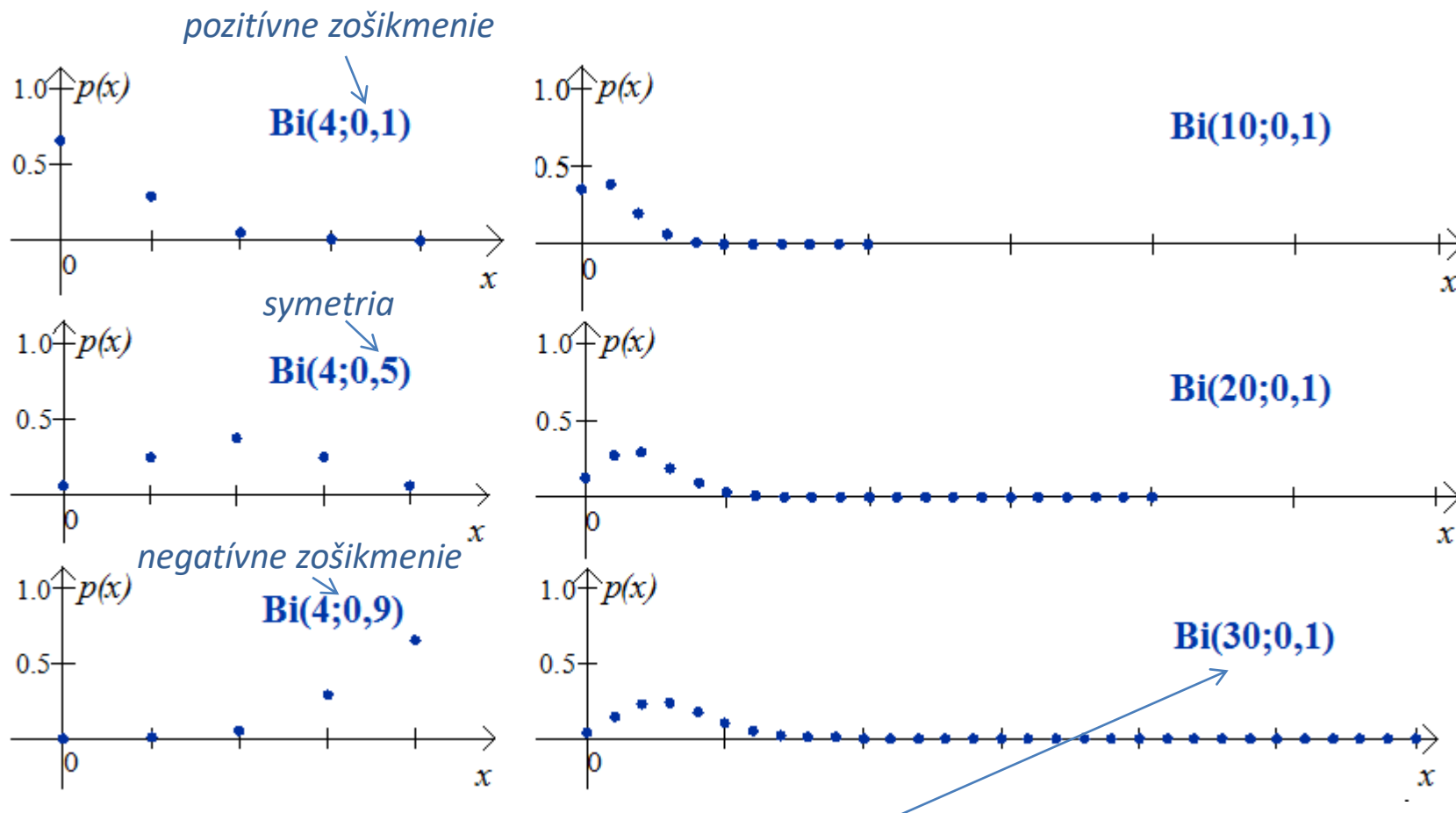
Binomické rozdelenie (Binomial distribution)

$$X \sim Bi(n, \pi)$$

- používa sa pri výbere s opakovaním (n **nezávislých opakovaných** pokusov). Pri každom z nich môže jav A nastať s rovnakou pravdepodobnosťou π a nenastať s pravdepodobnosťou $1 - \pi$ (tzv. Bernoulliho pokus) .
- X udáva koľkokrát jav A pri n -násobnom opakovaní pokusu nastal, môže nadobúdať niektorú hodnotu z množiny $\{0, 1, 2, \dots, n\}$, t.j. spolu $n + 1$ rôznych hodnôt.

$X \sim Bi(n, \pi)$	
$p(x) = \begin{cases} \binom{n}{x} \pi^x (1 - \pi)^{n-x} & \text{pre } x = 0, 1, \dots, n, \\ 0 & \text{inak} \end{cases}$	
Excel:	=BINOM.DIST(číslo_s; pokusy; pravdepodobnosť_s; kumulatívne) číslo_s je hodnota náhodnej premennej X , pokusy sú počet nezávislých pokusov (parameter n), pravdepodobnosť_s je pravdepodobnosť úspechu v jednom pokuse (parameter π), kumulatívne je logická premenná určujúca, či bude vypočítaná pravdepodobnostná funkcia (0) alebo distribučnú funkcia (1).

Tvar rozdelenia závisí od parametrov n a π .



S rastúcim n sa rozdelenie stáva stále viac symetrickým a konverguje k normálnemu rozdeleniu.

Náhodná premenná X má binomické rozdelenie, ak jej hodnoty predstavujú napr.:

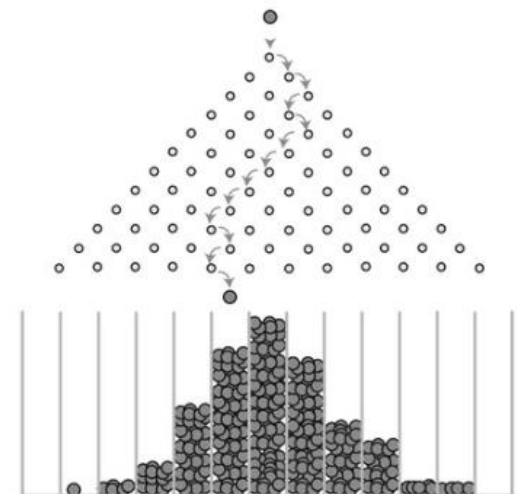
- počet chybných výrobkov zistených vo výbere n výrobkov, vyrábaných určitým strojom za stálych podmienok;
- počet chybných výrobkov zistených vo výbere n výrobkov z danej dávky, pri ktorom výrobok do nej vraciame;
- počet prípadov, v ktorých sa prejavila účinnosť podaného prípravku skúšaného na n objektoch;
- počet prípadov, v ktorých doba výrobnnej operácie prekročila určenú hodnotu, za predpokladu, že bola evidovaná doba n výrobných operácií;

Francis Galton – štatistický výskum v oblasti dedičnosti
– regresia ako proces návratu k priemeru

Galtonova doska

rozmiestnenie mincí približne podľa $Bi(n, 0,5)$

$$E(X) = n\pi, \quad D(X) = n\pi(1 - \pi)$$



Poissonovo rozdelenie (Poisson distribution)

$$X \sim Po(\lambda)$$

- Je to rozdelenie počtu výskytov javu A za jednotku času alebo oblasti, resp. špeciálne rozdelenie pre tzv. „maličké pravdepodobnosti“ (limitný prípad binomického rozdelenia).

$X \sim Po(\lambda)$	
$p(x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!} & \text{pre } x = 0, 1, \dots, \\ 0 & \text{inak} \end{cases}$	
Excel:	$=POISSON.DIST(x; \text{stred}; \text{kumulatívne})$ x je hodnota náhodnej premennej X ; stred je stredná hodnota λ ; kumulatívne je logická premenná – FALSE (0) určí, že počítame hodnotu pravdepodobnostnej funkcie, TRUE (1), ak počítame hodnotu distribučnej funkcie.

- pravdepodobnosť výskytu určitého javu v danom intervale (v čase alebo v priestore) je úmerná dĺžke tohto intervalu (napr. ak sa určitý jav vyskytne v priemere 12-krát za hodinu, tak za minútu sa vyskytne v priemere 0,2-krát);

Príklady Poissonovho rozdelenia pravdepodobnosti :

- počet ťažkých dopravných nehôd v určitom meste za deň;
- počet smrteľných pracovných úrazov v nejakom rezorte za rok;
- počet častíc vyslaných rádioaktívnou látkou počas určitého časového intervalu;
- počet chýb na zvare;
- počet bublín na ploche danej veľkosti na odliatku;
- počet mikrodefektov na zadanej vzorke materiálu;
- počet trhlín na 1km kábla;
- výskyt porúch zariadenia v čase t ;
- počet zákazníkov v obchode za časovú jednotku;
- počet červených krviniek v 1ml krvi;
- počet výskytov vzácného ochorenia na konkrétnom území, počet narodených dvojčiat, trojčiat v určitej populácii...

$$E(X) = \lambda, \quad D(X) = \lambda$$

Hypergeometrické rozdelenie (Hypergeometric distribution)

$$X \sim Hg(N, M, n)$$

- X predstavuje počet x prvkov so sledovanou vlastnosťou z n ($1 \leq n < N$) náhodne vybraných prvkov z konečnej N prvkovej množiny, medzi ktorými je M ($1 \leq M \leq N$) takých, ktoré majú sledovanú vlastnosť. Objekty vyberáme naraz alebo postupne bez vrátenia, realizujeme tzv. **závislé výbery**.

$$X \sim Hg(N, M, n)$$

$$p(x) = \begin{cases} \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} & \text{pre } \max\{0, n - N + M\} \leq x \leq \min\{n, M\}, x \in N_0, \\ 0 & \text{inak} \end{cases}$$

$N, M, n \in N: 1 \leq n < N, 1 \leq M < N$

Excel: =HYPGEOM.DIST(vzorka_s; veľkosť_vzorky; populácia_s; počet_pop; kumulatívne)
 vzorka_s je hodnota náhodnej premennej X ;
 veľkosť_vzorky určuje veľkosť výberu, t.j. n ;
 populácia_s udáva počet prvkov so sledovanou vlastnosťou – parameter M ;
 počet_pop je celkový počet prvkov – parameter N ;
 kumulatívne priradíme logickú hodnotu 0, ak počítame hodnotu $p(x)$ a logickú hodnotu 1, ak chceme vypočítať hodnotu distribučnej funkcie.

Vybrané rozdelenia spojitej náhodnej premennej

Rozdelenie spojitej náhodnej premennej možno popísať funkciou hustoty rozdelenia pravdepodobnosti $f(x)$, distribučnou funkciou $F(x) = P(X \leq x)$, resp. používame vybrané číselné charakteristiky.

- Rovnomerné (spojité) rozdelenie (Uniform distribution (continuous))
- Normálne (Gaussovo) rozdelenie (Normal (Gaussian) distribution)
- Exponenciálne rozdelenie (exponential distribution)
- Weibullovo rozdelenie (Weibull distribution)

Rovnomerné (spojité) rozdelenie (Uniform distribution (continuous))

$$X \sim R(\alpha, \beta)$$

- používa sa ako model rozdelenia vtedy, keď má spojitá náhodná premenná X na intervale (α, β) konštantnú hustotu pravdepodobnosti (geometrická definícia pravdepodobnosti);
- metóda Monte Carlo

$X \sim R(\alpha, \beta)$	
$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{pre } x \in (\alpha, \beta) \\ 0 & \text{inak} \end{cases}$	$F(x) = \begin{cases} 0 & \text{pre } x \in (-\infty, \alpha) \\ \frac{x - \alpha}{\beta - \alpha} & \text{pre } x \in (\alpha, \beta) \\ 1 & \text{pre } x \in \langle \beta, \infty \rangle \end{cases}$

Týmto rozdelením sa riadi napríklad:

- doba čakania na uskutočnenie javu, ktorý sa opakuje v pravidelných intervaloch, napr. doba čakania na dodávku tovaru, ktorá sa pravidelne opakuje, doba čakania na spoj, ktorý premáva v pravidelných intervaloch;
- veľkosť chyby pri zaokrúhľovaní čísel.

$$E(X) = \frac{\alpha + \beta}{2}, \quad D(X) = \frac{(\beta - \alpha)^2}{12}$$

Normálne (Gaussovo) rozdelenie (Normal (Gaussian) distribution)

$$X \sim N(\mu, \sigma^2)$$

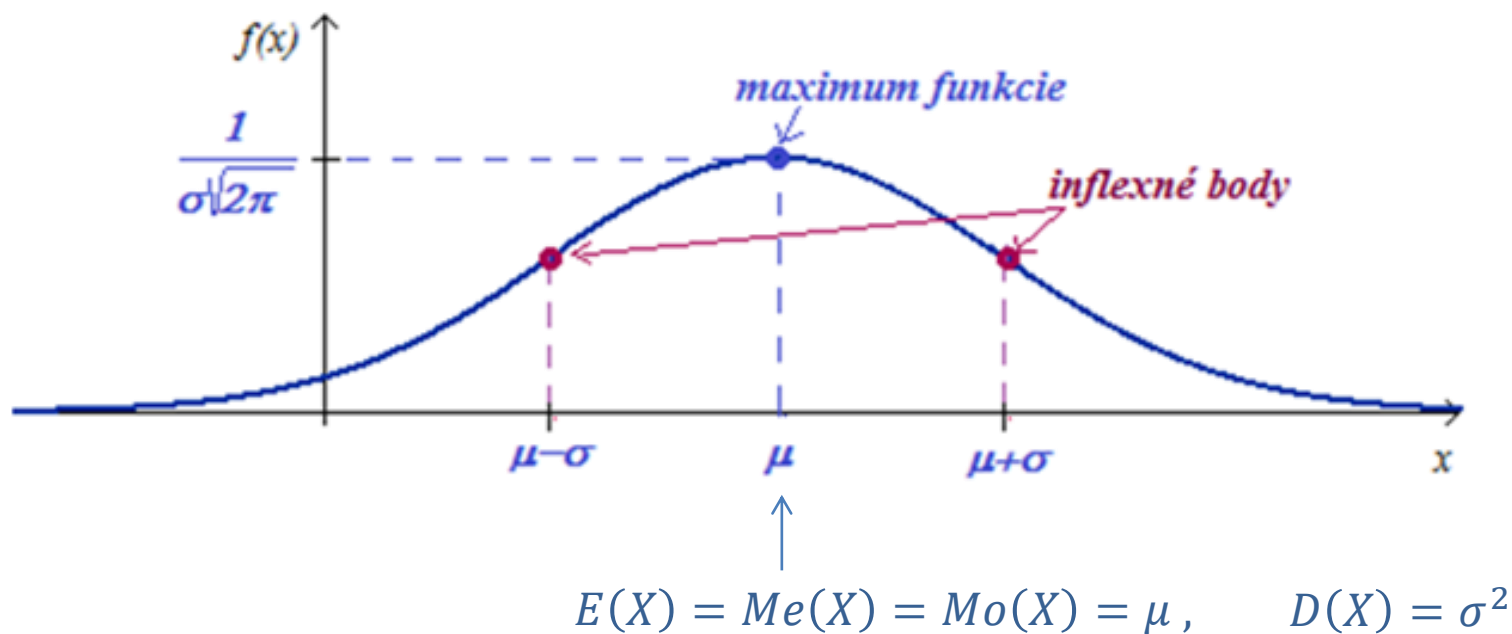
- najčastejšie a najdôležitejšie rozdelenie pravdepodobnosti v technike, prírodných, ekonomických a spoločenských vedách – náhodné výsledky sú vyjadrené ako aditívny výsledok mnohých nezávislých vplyvov;
- mimoriadny význam v teórii pravdepodobnosti a v matematickej štatistike;
- náhodné chyby – zákon chýb.

$X \sim N(\mu, \sigma^2)$	
$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}; x \in R$	$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt; x \in R$
Excel:	=NORM.DIST(x;stred;smerodajná_odch;kumulatívne) x je hodnota náhodnej premennej X; stred je stredná hodnota (priemer) μ daného rozdelenia; smerodajná_odch je smerodajná odchýlka σ; môžeme voľiť medzi frekvenčnou a distribučnou funkciou: kumulatívne zadávame 0 pre funkciu hustoty pravdepodobnosti $f(x)$ a 1 pre distribučnú funkciu $F(x)$.

Gaussova krivka má tvar zvona a je symetrická podľa osi $x = \mu$ (zo symetrie vyplýva $P(X > \mu) = P(X < \mu) = 0,5$).

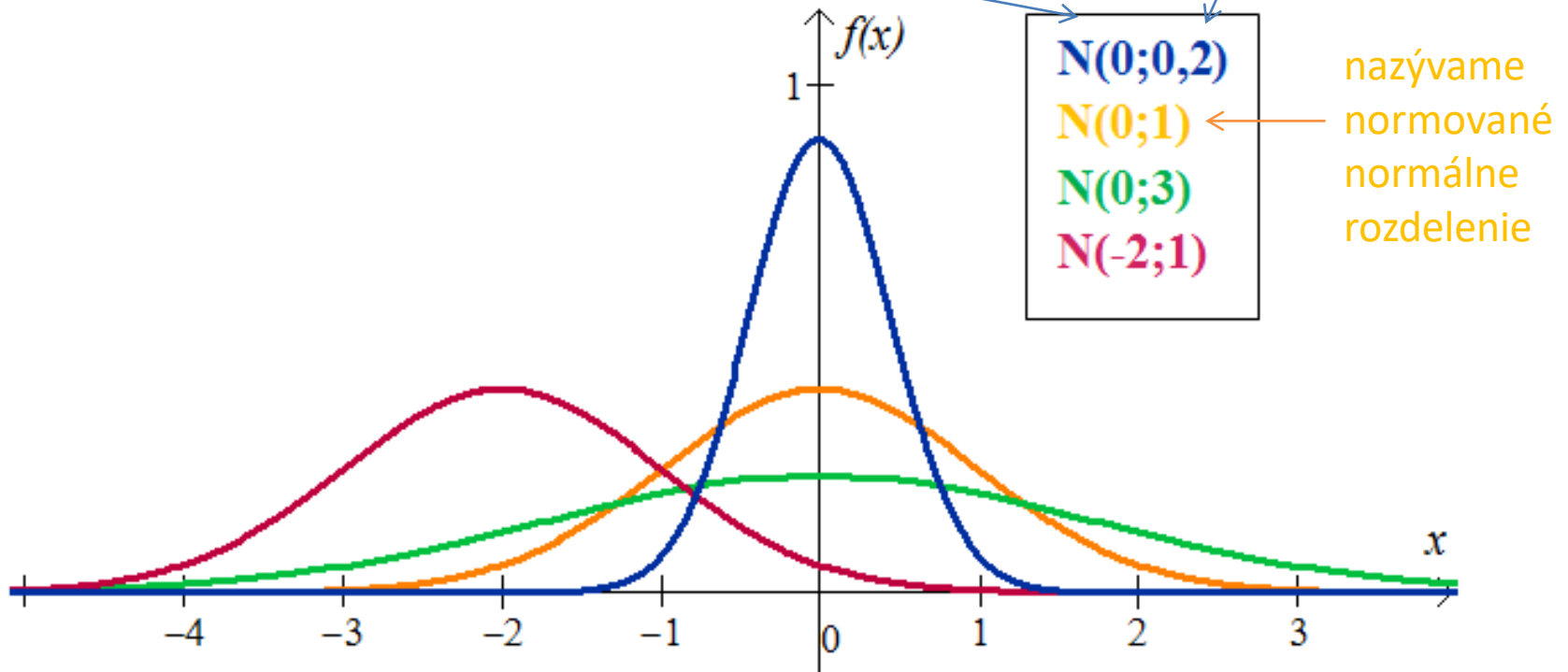
Parameter μ určuje, kde má krivka maximum s hodnotou $\frac{1}{\sigma\sqrt{2\pi}}$.

Parameter σ^2 je jediným parametrom, ktorý určuje tvar krivky, t.j. určuje ako sú od hodnoty μ vzdialené inflexné body, ktoré sú v $x = \mu + \sigma$ a v $x = \mu - \sigma$.



Umiestnenie Gaussovej krivky na x-ovej osi určuje stredná hodnota μ .

Tvar krivky ovplyvňuje disperzia σ^2 , resp. smerodajná odchýlka σ .



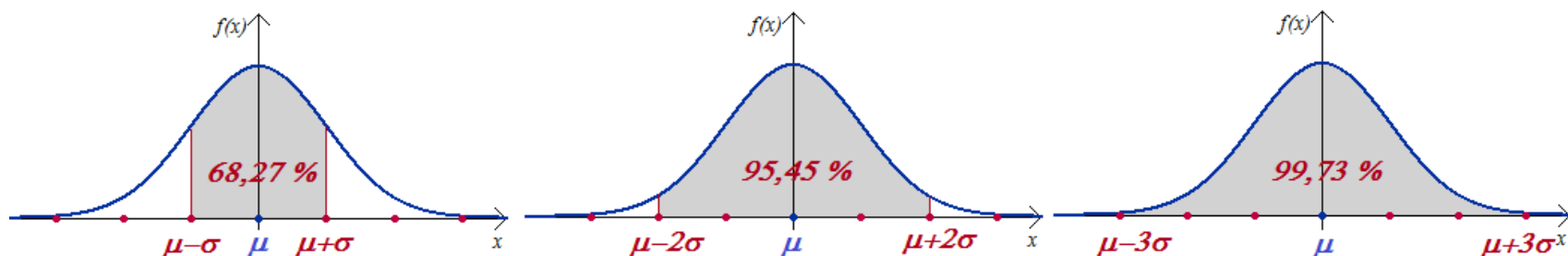
Pravidlo troch sigma

Pre náhodnú premennú $X \sim N(\mu, \sigma^2)$ platí:

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 0,6827$$

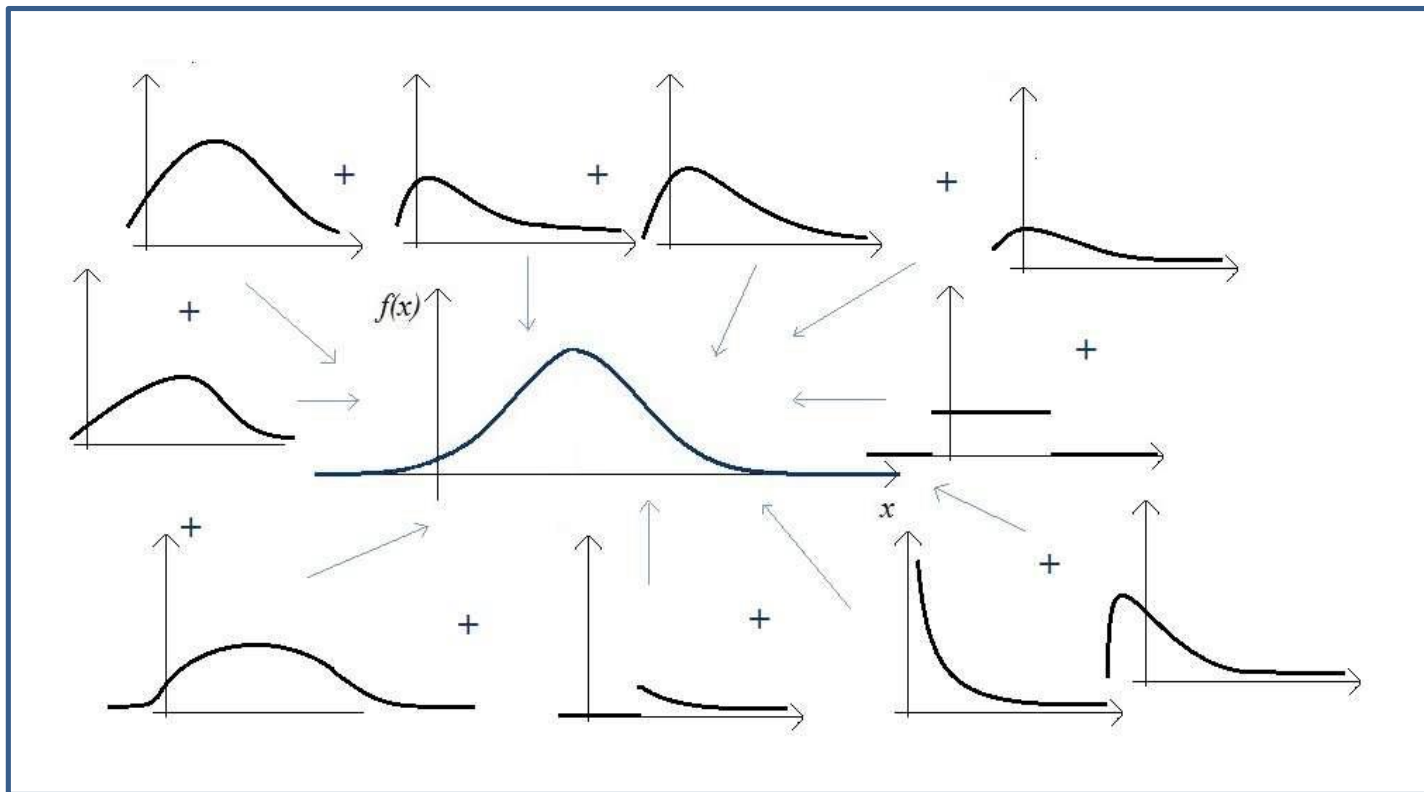
$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0,9545$$

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0,9973$$



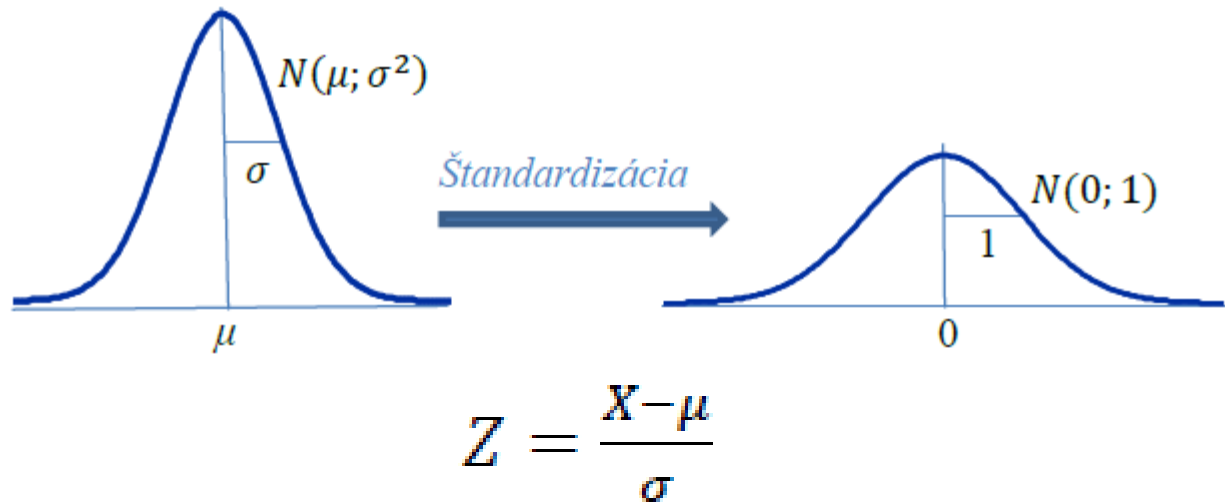
V praxi sa pravidlo troch sigma používa na prvý odhad štandardnej odchýlky pre náhodnú premennú, u ktorej predpokladáme normálne rozdelenie. Jej hodnotu dostaneme, ak rozpätie, t.j. rozdiel najväčšej a najmenšej hodnoty náhodnej premennej vydelíme šiestimi.

Centrálna limitná veta



Náhodná premenná, ktorá je určená pôsobením väčšieho počtu navzájom nezávislých vplyvov, má približne normálne rozdelenie, bez ohľadu na to, aké rozdelenia tieto nezávislé vplyvy majú.

Štandardizované (normované) normálne rozdelenie



$Z \sim N(0, 1)$	
$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}; z \in R$	$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt; z \in R$
Excel:	=NORM.S.DIST(z;kumulatívne) z je hodnota normovanej náhodnej premennej; kumulatívne môže nadobúdať hodnotu 0 pre $\varphi(z)$ a 1 pre hodnotu $\Phi(z)$.

História

- 1733 – De Moivre (hazardné hry)
- 18.-19. storočie – Gaussova -Laplaceova „krivka chýb“ (astronomické merania – pomocou MNŠ základe analýzy normálneho rozdelenia chýb predpovedal polohu planety Ceres)
- Adolph - Lambert Quételet – zakladateľ Kráľovskej štatistickej spoločnosti v Londýne (obvod prís škótskych vojakov) – príroda sa snaží vytvoriť „homme moyen“, avšak v rôznej miere chybuje
- Florence Nightingal , Francis Galton (Quincunx), ...
- Karl Pearson – normálna krivka → normálne rozdelenie → Gaussova krivka

Gaussova krivka



Gaussova krivka

Anglickí autori – the bell curve – zvonovitá krivka

Francúzski autori – krivka policajtného (Napoleonského) klobúka



Chvost normálneho rozdelenia je oblasť napravo od určitého bodu na pravej strane od strednej hodnoty alebo symetricky naľavo na ľavej strane krivky.

Pomenovanie „chvost“ má pôvod v obrázku W. Gossetta z r.1908.



Exponenciálne rozdelenie (Exponential distribution, Negative exponential distribution)

$$X \sim E(\lambda)$$

- toto rozdelenie má náhodná premenná X , ktorá vyjadruje dobu čakania medzi dvoma realizáciami náhodných javov, ktorých početnosť výskytu má Poissonovo rozdelenie.

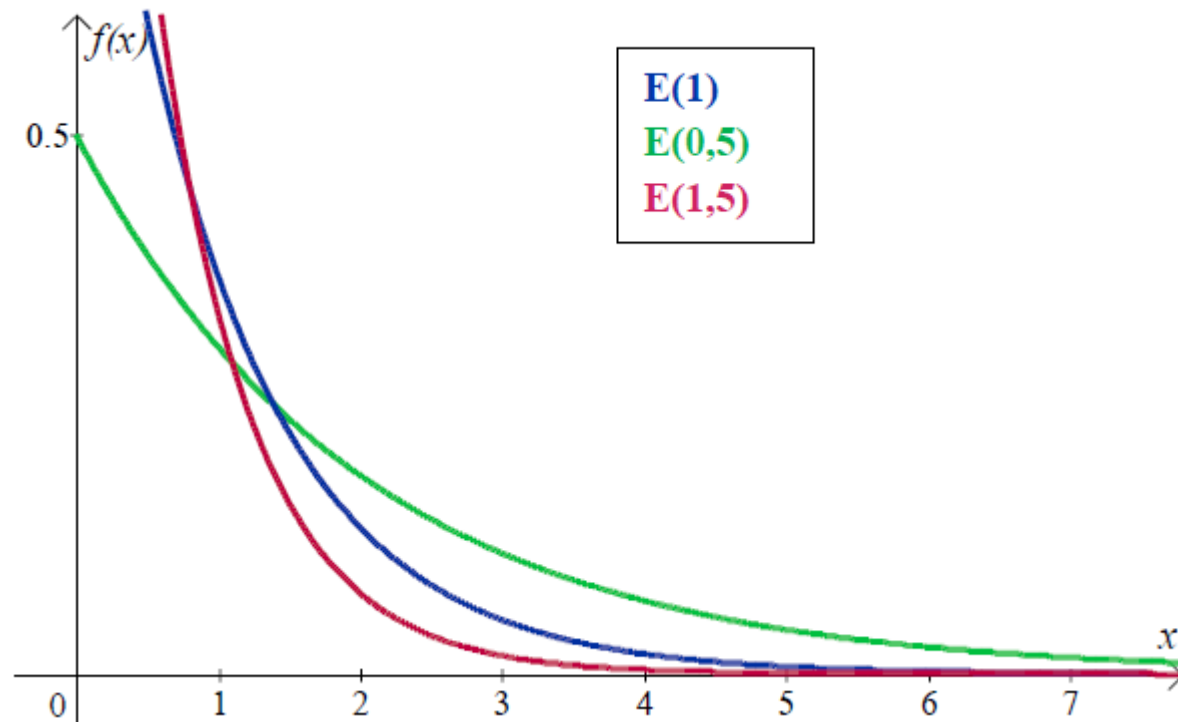
$X \sim E(\lambda)$		
$f(x)$	$= \begin{cases} \lambda e^{-\lambda x} & \text{pre } x \geq 0 \\ 0 & \text{pre } x < 0 \end{cases}$	$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{pre } x \geq 0 \\ 0 & \text{pre } x < 0 \end{cases}$
Excel:	=EXPON.DIST(x;lambda;kumulatívne) x je hodnota náhodnej premennej X ; lambda je strednou hodnotou rozdelenia; kumulatívne pre 0 vypočíta hodnotu $f(x)$, pre 1 hodnotu $F(x)$.	

Použitie tohto rozdelenia:

- dobře popisuje životnosť zariadení, u ktorých dochádza k poruche z náhodných príčin (životnosť je „čakanie na poruchu“);
- teória spoľahlivosti; teória hromadnej obsluhy; teória obnovy

Exponenciálne rozdelenie (Exponential distribution, Negative exponential distribution)

$$X \sim E(\lambda)$$



$$E(X) = \frac{1}{\lambda}, \quad D(X) = \frac{1}{\lambda^2}$$

Weibullovo rozdelenie (Weibull distribution)

$$X \sim W(\lambda, k)$$

- toto rozdelenie vyjadruje životnosť, bezporuchovosť technických zariadení, ktorým nevyhovuje exponenciálne rozdelenie, t.j. kde sa prejavuje mechanické opotrebovanie a únava materiálu .

$X \sim W(\lambda, k)$			
$f(x)$	$= \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} & \text{pre } x \geq 0 \\ 0 & \text{pre } x < 0 \end{cases}$	$F(x)$	$= \begin{cases} 1 - e^{-\left(\frac{x}{\lambda}\right)^k} & \text{pre } x \geq 0 \\ 0 & \text{pre } x < 0 \end{cases}$

Toto rozdelenie sa používa:

- v prípadoch, kedy bezporuchovosť závisí na veku, počte odpracovaných hodín, počte vykonaných cyklov, opotrebení (popisuje čas do poruchy systému, kedy porucha nastáva pri poruche najslabšieho komponentu systému).

Rozdelenia funkcií náhodných premenných (odvodené rozdelenia, výberové rozdelenia)

Majú mimoriadny význam v praxi pri analýze štatistických údajov získaných náhodným výberom, t.j. používajú sa pri výberovom skúmaní. Sú úzko spojené s normálnym rozdelením.

- **χ^2 rozdelenie** pri popise výberového rozptylu
- **Studentovo t-rozdelenie** najčastejšie k porovnaniu priemerov
- **Fisherovo F-rozdelenie** pri porovnaní rozptylov dvoch súborov alebo pri porovnaní závislých veličín v lineárnej regresii

χ^2 rozdelenie (chí-kvadrát, Helmertovo-Pearsonovo rozdelenie) (Chi-squared distribution)

$$\chi^2 \sim \chi^2(n)$$

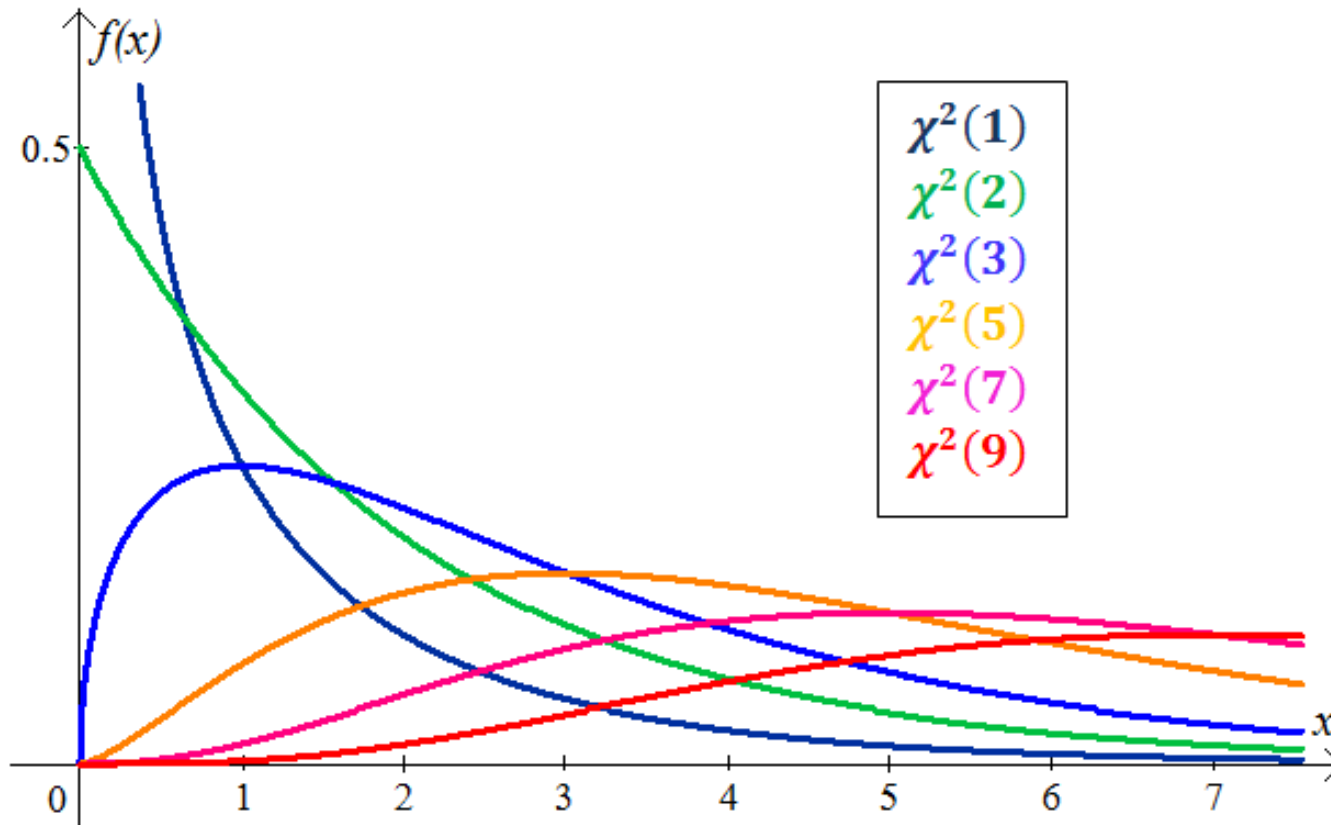
Nech $X_1, X_2, X_3, \dots, X_n$ je n nezávislých náhodných premenných, z ktorých každá má rozdelenie $N(0,1)$. Ich súčet štvorcov je náhodná premenná, ktorá má chí-kvadrát rozdelenie s n stupňami voľnosti.

$$\chi^2 = X_1^2 + X_2^2 + \dots + X_n^2 = \sum_{i=1}^n X_i^2$$

n – pod stupňami voľnosti rozumieme v štatistike počet nezávislých informačných jednotiek (počet meraní mínus počet odhadovaných parametrov)

$\chi^2 \sim \chi^2(n)$	
$f(x)$	$= \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \cdot x^{\frac{n}{2}-1} \cdot e^{-\frac{x}{2}}, x \in (0, \infty)$
Excel:	$=\text{CHISQ.DIST}(x;\text{stupeň voľnosti};\text{kumulatívne})$ x je hodnota náhodnej premennej χ^2 ; stupeň voľnosti je parameter určujúci počet stupňov voľnosti; kumulatívne je logická premenná určujúca, či bude vypočítaná funkcia hustoty pravdepodobnosti (parameter 0) alebo distribučnú funkcia (parameter 1).

$$E(X) = n, \quad D(X) = 2n$$



Počet stupňov voľnosti n je daný počtom nezávislých sčítancov a je jediným parametrom rozdelenia. S rastúcim počtom stupňov voľnosti sa chí-kvadrát rozdelenie „blíži“ k normálnemu rozdeleniu, t.j. pre $n \geq 30$ ho možno aproximovať (nahradiť) normálnym rozdelením $N(n; 2n)$.

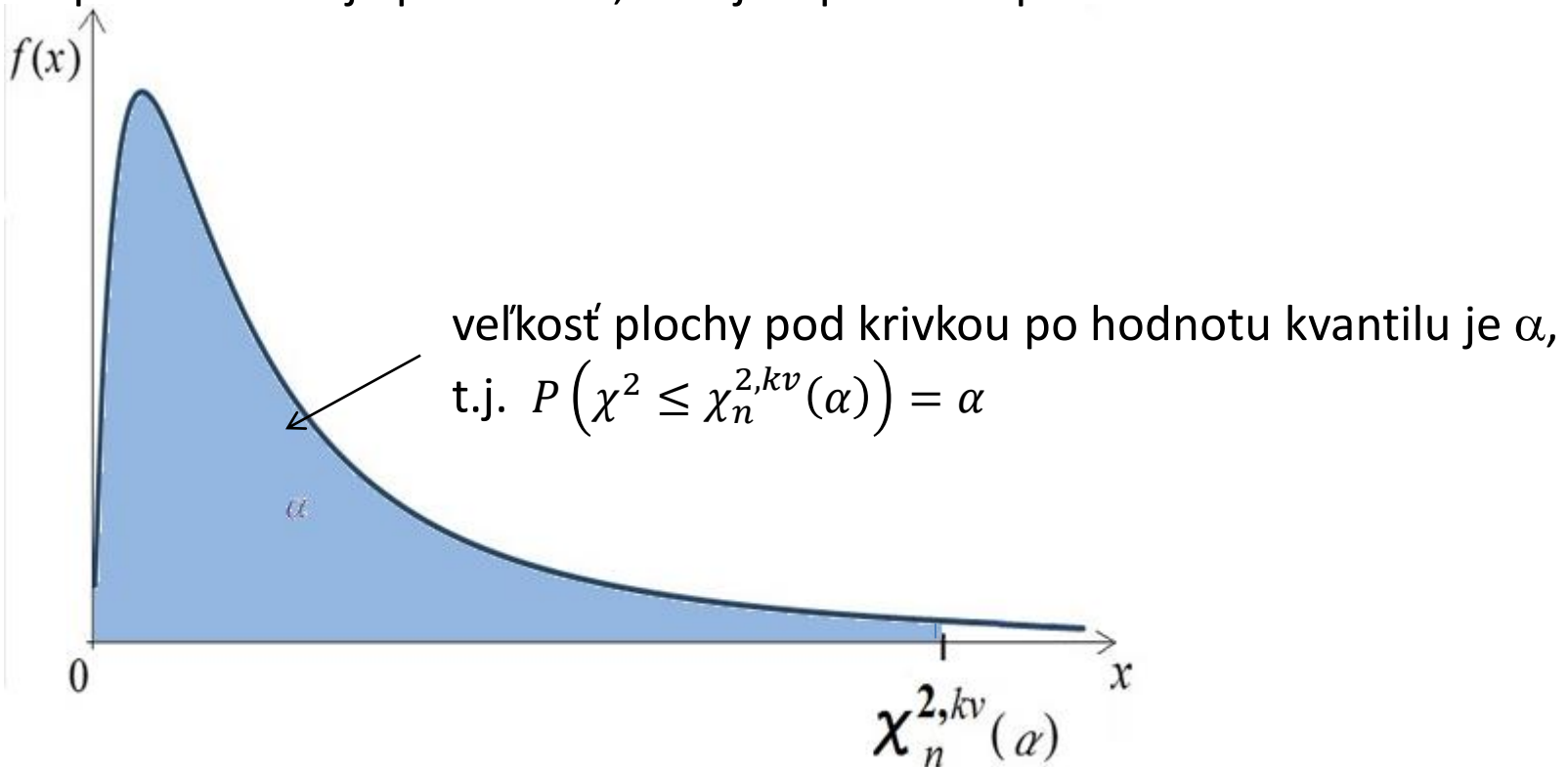
$$n \geq 30 \Rightarrow \chi^2(n) \sim N(n; 2n)$$

Kvantily chí-kvadrát rozdelenia

Kvantily chí-kvadrát rozdelenia budeme označovať $\chi_n^{2,kv}(\alpha)$, kde

$$P\left(\chi^2 \leq \chi_n^{2,kv}(\alpha)\right) = \alpha.$$

V Exceli: `=CHISQ.INV(pravdepodobnosť; stupeň voľnosti)`, kde pravdepodobnosť je pravdepodobnosť α , pre ktorú hľadáme kvantil $\chi_n^{2,kv}(\alpha)$; stupeň voľnosti je parameter, určujúci počet stupňov voľnosti.



Studentovo t – rozdelenie (Student's t – distribution)

$$t \sim t(n)$$

S týmto rozdelením sa stretávame v teórii odhadu, pri testoch štatistických hypotéz. Náhodná premenná t , ktorá má n stupňov voľnosti, je vytvorená z $n + 1$ nezávislých náhodných premenných $X, X_1, X_2, X_3, \dots, X_n$ s rozdelením $N(0,1)$ podľa predpisu

$$t = \frac{X}{\sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2}} = \frac{X}{\sqrt{\frac{1}{n} \chi^2}}$$

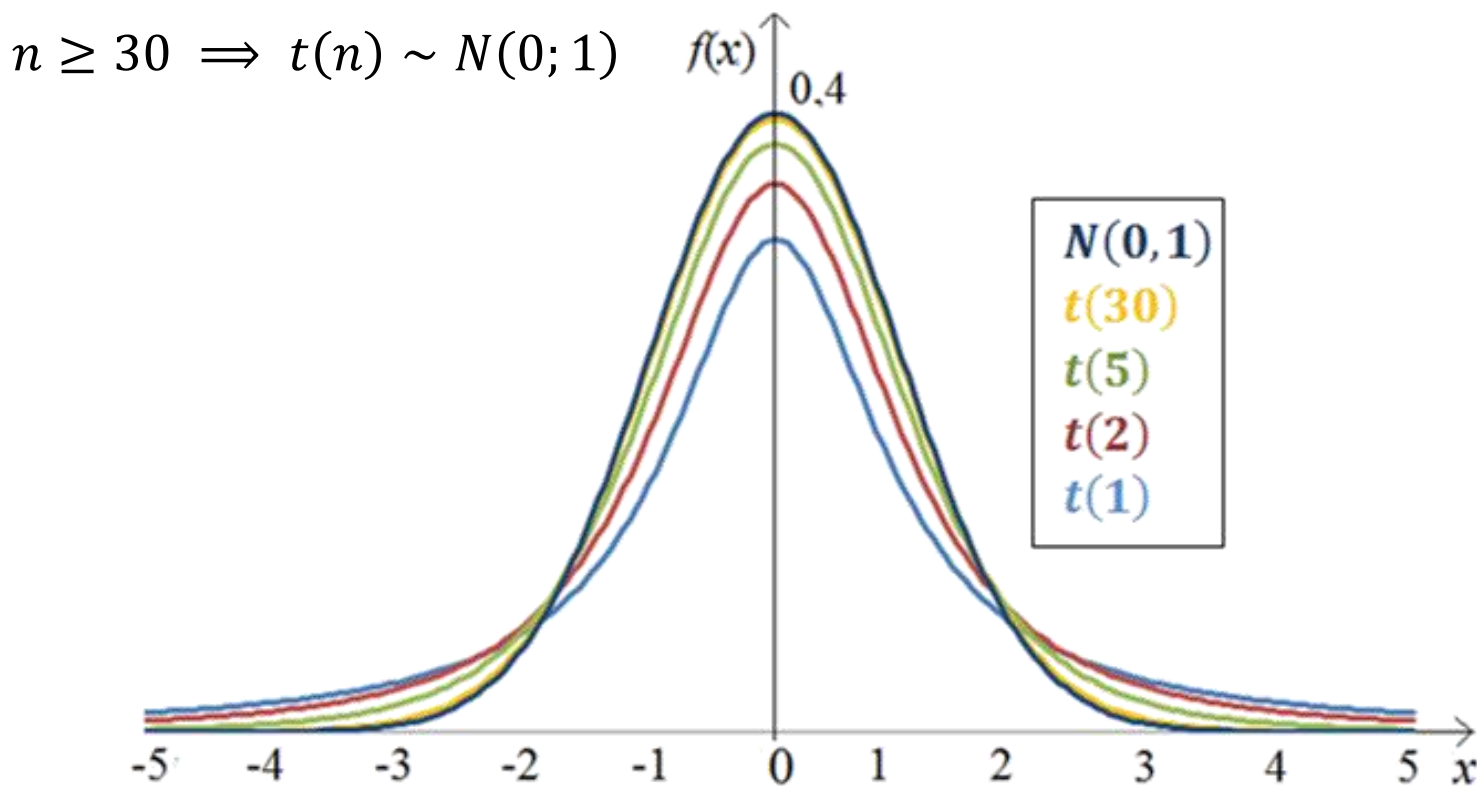
$t \sim t(n)$	
$f(x)$	$= \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2}) \cdot \sqrt{\pi n}} \cdot \left(1 + \frac{1}{x^2}\right)^{-\frac{n+1}{2}}, x \in (-\infty, \infty)$
Excel:	Na výpočet funkcie hustoty a distribučnej funkcie sa používa =T.DIST(x;stupeň voľnosti;kumulatívne) x je hodnota náhodnej premennej; stupeň voľnosti je parameter určujúci počet stupňov voľnosti; kumulatívne je logická premenná určujúca, či bude vypočítaná funkcia hustoty pravdepodobnosti (parameter 0) alebo distribučnú funkcia (parameter 1)

Původ názvu Studentovo má zajímavou historií. Irský statistik W. S. Gosset poprvé publikoval toto rozdělení anonymně pod pseudonymem "Student", protože jeho zaměstnavatel, pivovar Guinness v Dublinu, zakázal svým zaměstnancům publikovat pod svým vlastním jménem z obavy, že konkurence by odhalila tajemství jejich excelentního piva. Ve svém původním článku, Gosset použil označení "t" pro svoji statistiku. Odtud Studentovo t rozdělení pravděpodobnosti. Na práci Gosseta navázalo množství dalších statistiků; jmenujme alespoň R. A. Fishera, jehož jméno můžeme najít téměř ve všech směrech dalšího vývoje statistiky.



Zdroj: https://jv.wikipedia.org/wiki/William_Sealey_Gosset

S rastúcim počtom stupňov voľnosti sa t -rozdelenie „blíži“ k normovanému normálnemu rozdeleniu, pre $n \geq 30$ ho možno aproximovať normovaným normálnym rozdelením $N(0; 1)$

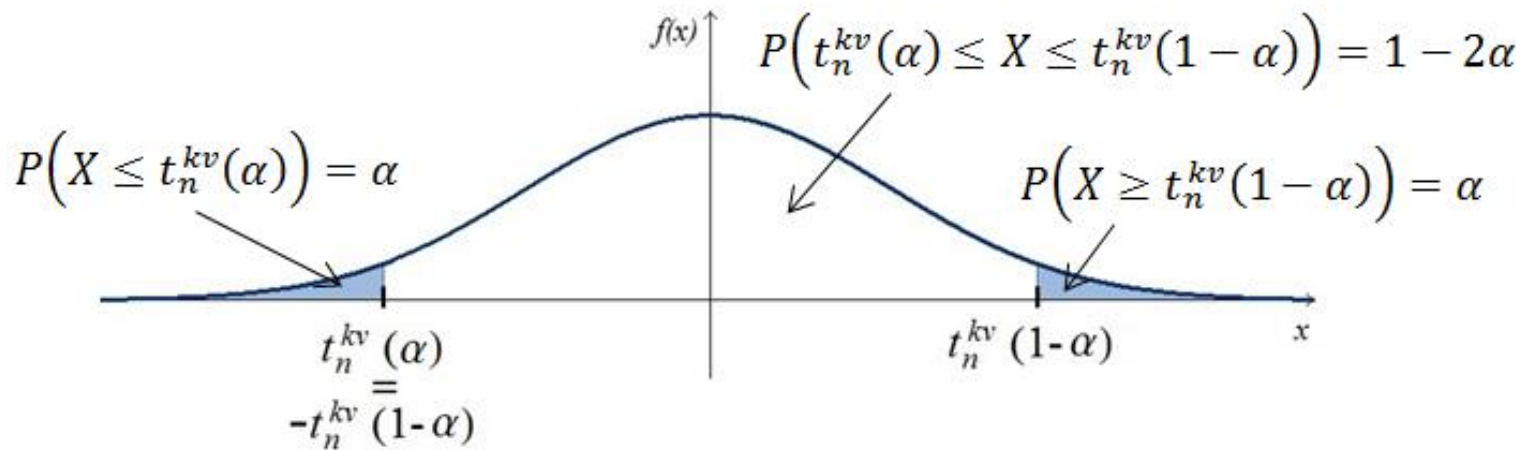


$E(X) = 0$ pre $k > 1$, pre $k = 1$ dostaneme Cauchyho rozdelenie

$D(X) = \frac{n}{n-2}$ pre $n > 2$

Kvantily t -rozdelenia

Kvantily t -rozdelenia budeme označovať $t_n^{kv}(\alpha)$, kde $P(t \leq t_n^{kv}(\alpha)) = \alpha$.
V Exceli: `=T.INV(pravdepodobnosť; stupeň voľnosti)`, kde pravdepodobnosť je pravdepodobnosť α , pre ktorú hľadáme kvantil $t_n^{kv}(\alpha)$; stupeň voľnosti je parameter, určujúci počet stupňov voľnosti.



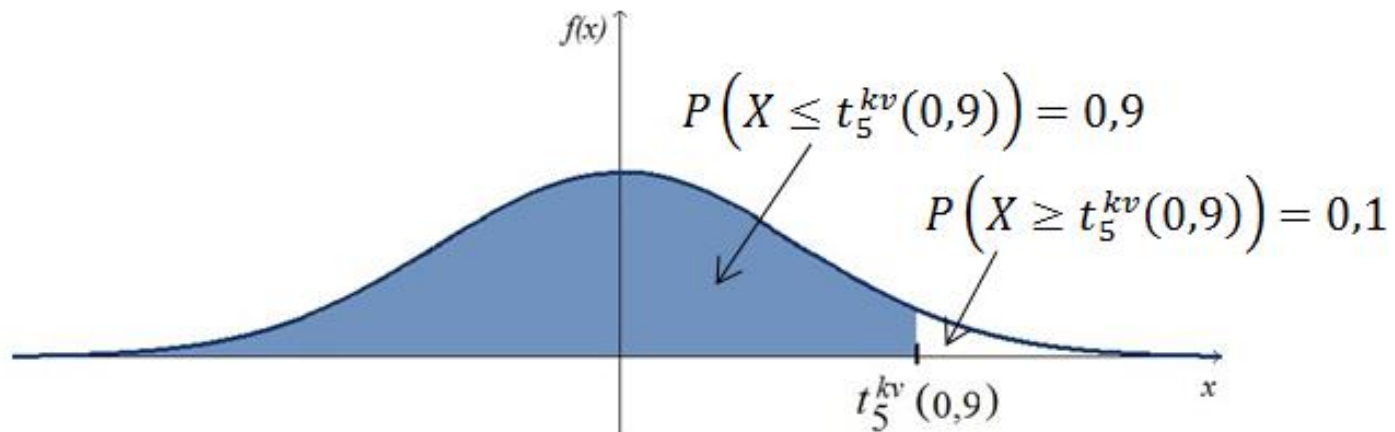
Zo symetrie Studentovho t -rozdelenia vyplýva, že $t_n^{kv}(\alpha) = -t_n^{kv}(1 - \alpha)$.

Príklad

Vypočítajte 90% kvantil Studentovho t -rozdelenia s 5 stupňami voľnosti. Aký je podiel hodnôt náhodnej premennej s daným rozdelením, ktoré presahujú číselnú hodnotu vypočítaného kvantilu?

Riešenie.

$$t_5^{kv}(0,9) = \text{T.INV}(0,9; 5) = 1,475884$$



***F*-rozdelenie alebo Fisherovo – Snedecorovo rozdelenie (Snedecor's *F*-distribution or Fisher-Snedecor distribution)**

$$F \sim F(n_1, n_2).$$

Nech $X_1, X_2, X_3, \dots, X_{n_1}$ a $Y_1, Y_2, Y_3, \dots, Y_{n_2}$ sú nezávislé náhodné premenné s normovaným normálnym rozdelením $N(0,1)$. Rozdelenie náhodnej premennej

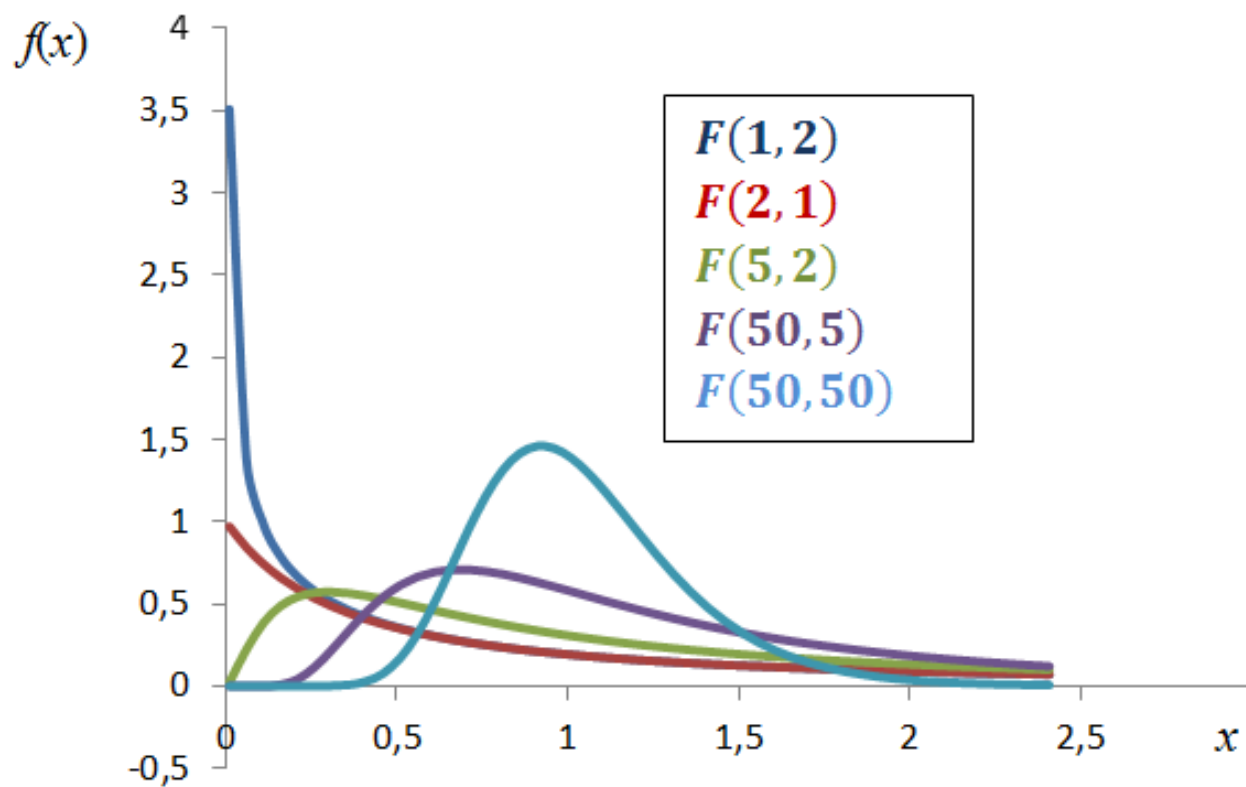
$$F = \frac{\frac{1}{n_1} \sum_{i=1}^{n_1} X_i^2}{\frac{1}{n_2} \sum_{i=1}^{n_2} Y_i^2} = \frac{\frac{1}{n_1} \chi_1^2}{\frac{1}{n_2} \chi_2^2}$$

sa nazýva Fisherovo – Snedecorovo rozdelenie s n_1 a n_2 stupňami voľnosti.

$$F \sim F(n_1, n_2)$$

Excel:	Na výpočet funkcie hustoty a distribučnej funkcie sa používa =F.DIST(x; stupeň voľnosti_1; stupeň voľnosti_2;kumulatívne) x je hodnota náhodnej premennej; stupeň voľnosti_1, stupeň voľnosti_2 sú parametre určujúce počet stupňov voľnosti n_1, n_2 ; kumulatívne je logická premenná určujúca, či bude vypočítaná funkcia hustoty pravdepodobnosti (parameter 0) alebo distribučná funkcia (parameter 1).
--------	--

Fisherovo rozdelenie má 2 parametre, stupne voľnosti n_1 a n_2 , je asymetrické.

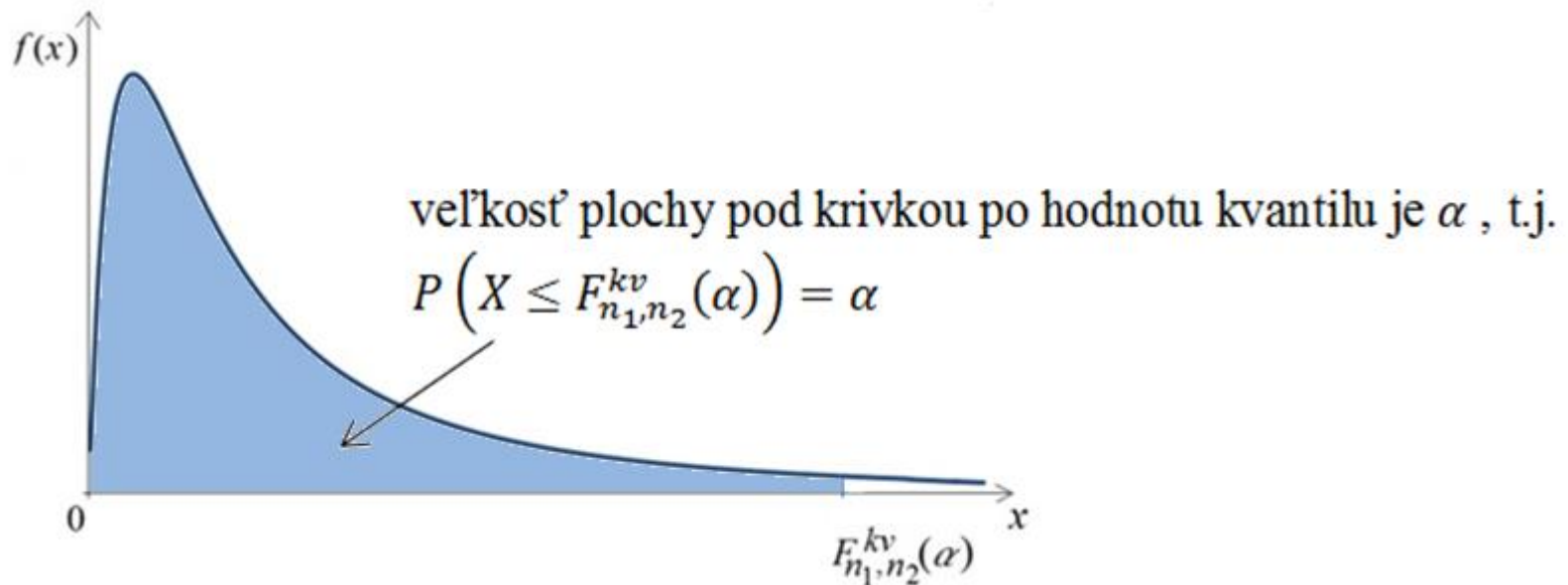


Kvantily F -rozdelenia

Kvantily F -kvadrát rozdelenia budeme označovať $F_{n_1, n_2}^{kv}(\alpha)$, kde

$$P\left(F \leq F_{n_1, n_2}^{kv}(\alpha)\right) = \alpha.$$

V Exceli: `=F.INV(pravdepodobnosť; stupeň voľnosti 1; stupeň voľnosti 2)`, kde pravdepodobnosť je pravdepodobnosť α , pre ktorú hľadáme kvantil $F_{n_1, n_2}^{kv}(\alpha)$; stupeň voľnosti 1 a 2 sú parametre, určujúce počet stupňov voľnosti.

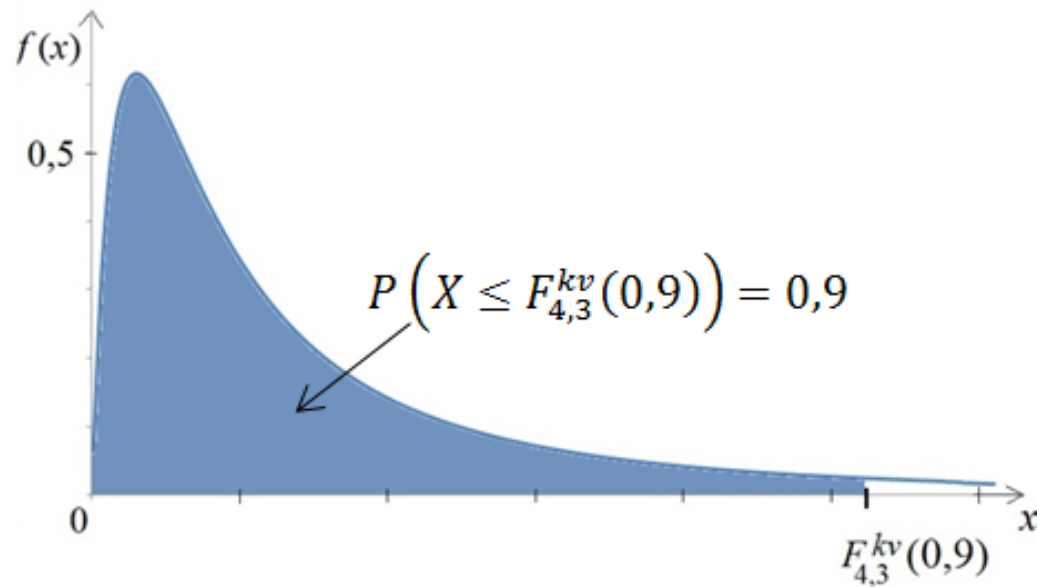


Príklad

Nájdite 90. percentil náhodnej premennej s rozdelením $F \sim F(4, 3)$.

Riešenie.

$$F_{4,3}^{kv}(0,9) = \text{F.INV}(0,9; 4; 3) = 5,342644$$



Príklady_rozdelenia

1. Poistovňa poistila 1000 ľudí rovnakého veku. Pravdepodobnosť úmrtia v priebehu roka je pre každého z nich 0,008. Každý poistenec zaplatil 39€. V prípade jeho úmrtia v priebehu roka dostanú pozostalí od poistovne 2600€. Aká je pravdepodobnosť, že poistovňa utrpí stratu?

→ poistovňa vyberá na poistnom $39 \cdot 1000 \text{ €} = 39\,000 \text{ €}$
je kuca schopná vyplatiť $\frac{39\,000 \text{ €}}{2\,600 \text{ €}} = 15 \text{ ľudí} \Rightarrow$
ak by bolo viac ako 15 ľudí

X - počet konšekcia prirodencov do roka

$$n = 1000$$

$$p = 0,008$$

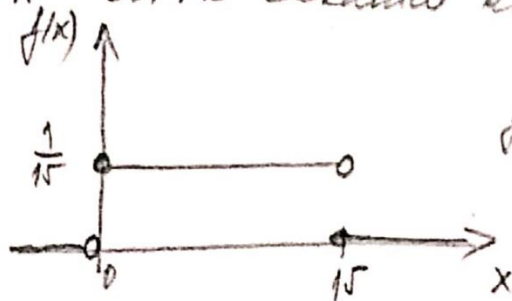
$$X \sim \text{Bi}(1000; 0,008)$$

$$P(X > 15) = 1 - \text{BINOM.DIST}(15; 1000; 0,008; 1) = 0,00798$$

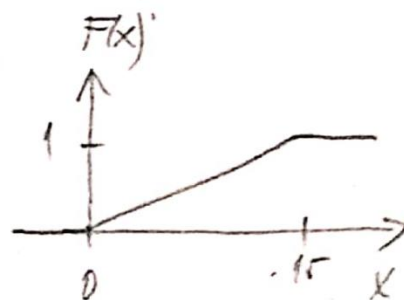
2. Autobus jazdí v pravidelných intervaloch po 15 minútach. Cestujúci príde na zastávku v ľubovoľnom okamihu. Sledujeme náhodnú premennú predstavujúcu čakanie na príchod autobusu.

- a) Popíšte túto náhodnú premennú pomocou funkcie hustoty pravdepodobnosti a distribučnej funkcie. Vyjadrite ich matematicky i graficky.
- b) Určte pravdepodobnosť, že cestujúci bude čakať na spoj najviac 5 minút, práve 10 minút, najmenej 3 minúty, 3 až 10 minút.

→ X - doba čakania na BUS



$$f(x) = \begin{cases} \frac{1}{15}; & x \in (0; 15) \\ 0; & \text{inak} \end{cases}$$



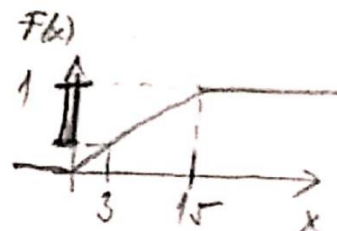
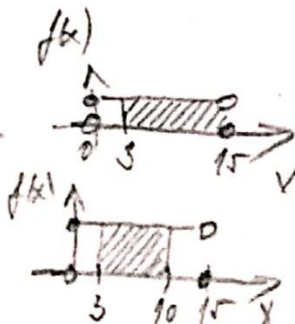
$$F(x) = \begin{cases} 0; & x \leq 0 \\ \int_0^x \frac{1}{15} dx = \frac{1}{15}x; & 0 < x < 15 \\ 1; & x \geq 15 \end{cases}$$

$$P(X \leq 5) = \frac{1}{15} \cdot 5 = \frac{5}{15} = \frac{1}{3}$$

$$P(X = 10) = 0$$

$$P(X \geq 3) = \frac{1}{15} \cdot (15 - 3) = \frac{12}{15} = \frac{4}{5}$$

$$P(3 \leq X \leq 10) = \frac{1}{15} (10 - 3) = \frac{7}{15}$$



$$P(X \geq 3) = F(15) - F(3) = \frac{1}{15} \cdot 15 - \frac{1}{15} \cdot 3 = \frac{12}{15}$$

Príklady_rozdelenia

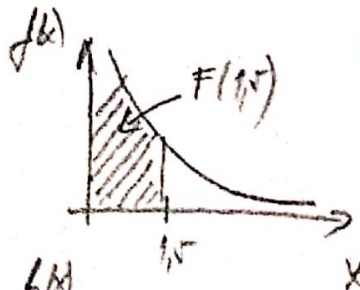
3. Stredná doba čakania zákazníka na obsluhu v určitej predajni je 2 minúty. Predpokladajme, že náhodná premenná X – doba čakania na obsluhu má exponenciálne rozdelenie. Aká je pravdepodobnosť, že náhodný zákazník bude obslúžený v dobe kratšej ako 1,5 minúty? Aký je podiel zákazníkov, ktorí budú čakať dlhšie ako 2 minúty?

→ X - doba čakania na obsluhu

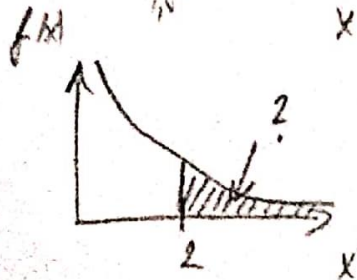
$$E(X) = 2 \text{ minúty} = \frac{1}{\lambda} \Rightarrow \lambda = \frac{1}{2}$$

$$X \sim \text{Exp}(\lambda)$$

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x}, & x \geq 0 \\ 0 & x < 0 \end{cases}$$



$$P(X < 1,5) = \text{EXPON.DIST}(1,5; 1/2; 1) \doteq 0,23618$$



podiel zákazníkov, kt. čakajú viac ako 2 minúty

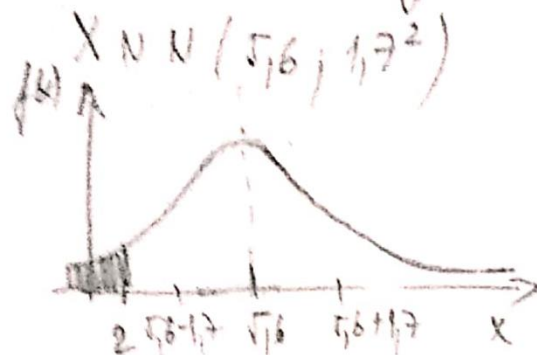
$$1 - \text{EXPON.DIST}(2; 1/2; 1) \doteq 0,367879 \Rightarrow 36,8\%$$

Príklady_rozdelenia

4. Firma poskytuje na svoje výrobky záručnú dobu 2 roky. Na každom výrobku, ktorý predá má firma zisk 520€. Ak zákazník vráti chybný výrobok v záručnej dobe, firma vymení výrobok za nový, čím stratí 1000€. Predpokladajme, že životnosť výrobku sa riadi normálnym rozdelením so strednou hodnotou 5,6 roku a smerodajnou odchýlkou 1,7 roku a viac ako 1 výmena neprichádza do úvahy. Vypočítajte pravdepodobnosť s akou sa výrobok pokazí počas záručnej doby a určte priemerný zisk za predaný kus. Určte, aká by mala byť záručná doba, aby priemerný zisk z predaného výrobku bol aspoň 508€.

Vypočítajte pravdepodobnosť s akou sa výrobok pokazí počas záručnej doby a určte priemerný zisk za predaný kus.
 Určte, aká by mala byť záručná doba, aby priemerný zisk z predaného výrobku bol aspoň 508€.

→ X - čas do poruchy výrobku v rokoch



$$P(X \leq 2) = \text{NORM.DIST}(2; 5.6; 1.7; 1) = 0.017102$$

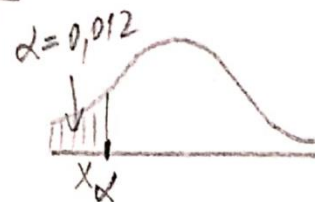
- priemerný zisk za predaný kus: $520 - 1000 \cdot 0.017 = 503 \text{ €}$
- aká má byť záručná doba, aby priemerný zisk bol aspoň 508 €

$$520 - 1000 \cdot p \geq 508$$

$$520 - 508 \geq 1000p$$

$$12 \geq 1000p$$

$$\frac{12}{1000} \geq p$$



$$x_\alpha = F^{-1}(\alpha)$$

$$x_\alpha = \text{NORM.INV}$$

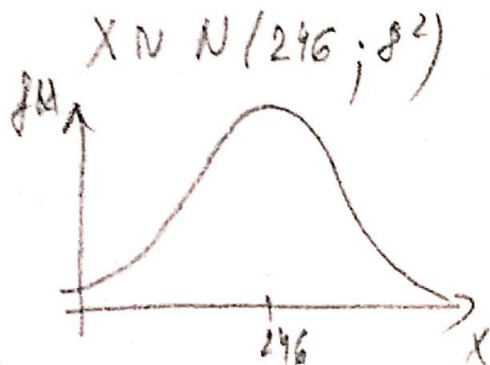
$$(0.012; 5.6; 1.7) = 1.76288 \Rightarrow$$

21,17 mes.

Príklady_rozdelenia

5. Maslo sa strojovo porcuje a balí automatom. Dlhodobým pozorovaním bolo zistené, že linka produkuje balíčky masla s priemernou hmotnosťou 246 gramov a smerodajnou odchýlkou 8 gramov. Predpokladajme, že hmotnosť masla je náhodná premenná s normálnym rozdelením. Aká je pravdepodobnosť, že náhodne vybrané balenie masla bude mať hmotnosť menšiu ako 250 gramov? Určte pravdepodobnosť, že náhodne vybrané balenie masla bude mať hmotnosť väčšiu ako 240 gramov. Aký je podiel balení, ktoré prejdú výstupnou kontrolou, ak je povolená tolerancia 250 gramov $\pm$ 10 gramov?

→ X - hmotnosť masla



$$P(X < 250) = \text{NORM.DIST}(250; 246; 8; 1)$$

$$P(X > 240) = 1 - \text{NORM.DIST}(240; 246; 8; 1)$$

$$P(240 \leq X \leq 260) = \text{NORM.DIST}(260; 246; 8; 1) - \text{NORM.DIST}(240; 246; 8; 1)$$

Príklady_rozdelenia

6. Určte medián náhodnej premennej s χ^2 rozdelením s 5 stupňami voľnosti.
7. Náhodná premenná sa riadi Studentovým t-rozdelením so 4 stupňami voľnosti. Aká je hodnota, ktorú presahuje 90% hodnôt danej premennej?
8. Aký je podiel hodnôt náhodnej premennej so Studentovým t-rozdelením so 4 stupňami voľnosti, ktoré presahujú číslo 1,7?

