Rovnoměrné rozdělení

- Nejjednodušší pravděpodobnostní rozdělení pro diskrétní náhodnou veličinu.
- V literatuře se také nazývá jako **klasické** rozdělení pravděpodobnosti.
- Náhodná veličina může nabývat *n* hodnot se stejnou pravděpodobností 1/n.
- Náhodná veličina X má rovnoměrné rozdělení s parametrem *n* pokud

  $p(a) = P(X=a) = 1/n$, pro $a = 1, 2, 3, \ldots, n$

  $p(a) = P(X=a) = 0$, pro ostatní a.
- Př.: házení kostkou, n=6, všechny možné výsledky jsou stejně pravděpodobné

Bernoulliho rozdělení

- Také nazývané jako **alternativní** rozdělení.
- Popisuje experiment, jehož možnými výsledky jsou jen dvě možné hodnoty. Např. (úspěch, neúspěch) nebo (1, 0).

DEFINITION. A discrete random variable X has a *Bernoulli distribution* with parameter p, where $0 \leq p \leq 1$, if its probability mass function is given by

$$p_X(1) = \mathrm{P}(X = 1) = p \quad \text{and} \quad p_X(0) = \mathrm{P}(X = 0) = 1 - p.$$

We denote this distribution by $Ber(p)$.

- Pravděpodobnostní funkce je daná vztahem:

  $p(a) = P(X=a) = p^a(1-p)^{1-a}$, pro a = 0, 1

  $p(a) = 0$, pro ostatní *a*

Binomické rozdělení

- Popisuje mi, jak je rozložena pravděpodobnost, že nastane náhodný jev *k* krát, pokud budu provádět *n* nezávislých pokusů. Každý náhodný jev má stejnou pravděpodobnost, že nastane.

DEFINITION. A discrete random variable X has a *binomial distribution* with parameters n and p, where $n = 1, 2, \ldots$ and $0 \le p \le 1$, if its probability mass function is given by

$$p_X(k) = \mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for } k = 0, 1, \ldots, n.$$

We denote this distribution by $Bin(n, p)$.

- Tedy $p_X(k)$ nám říká, jaká je pravděpodobnost, že jev nastane právě *k*-krát. Neboli s jakou pravděpodobností náhodná veličina X bude rovna hodnotě *k*.

Binomické rozdělení

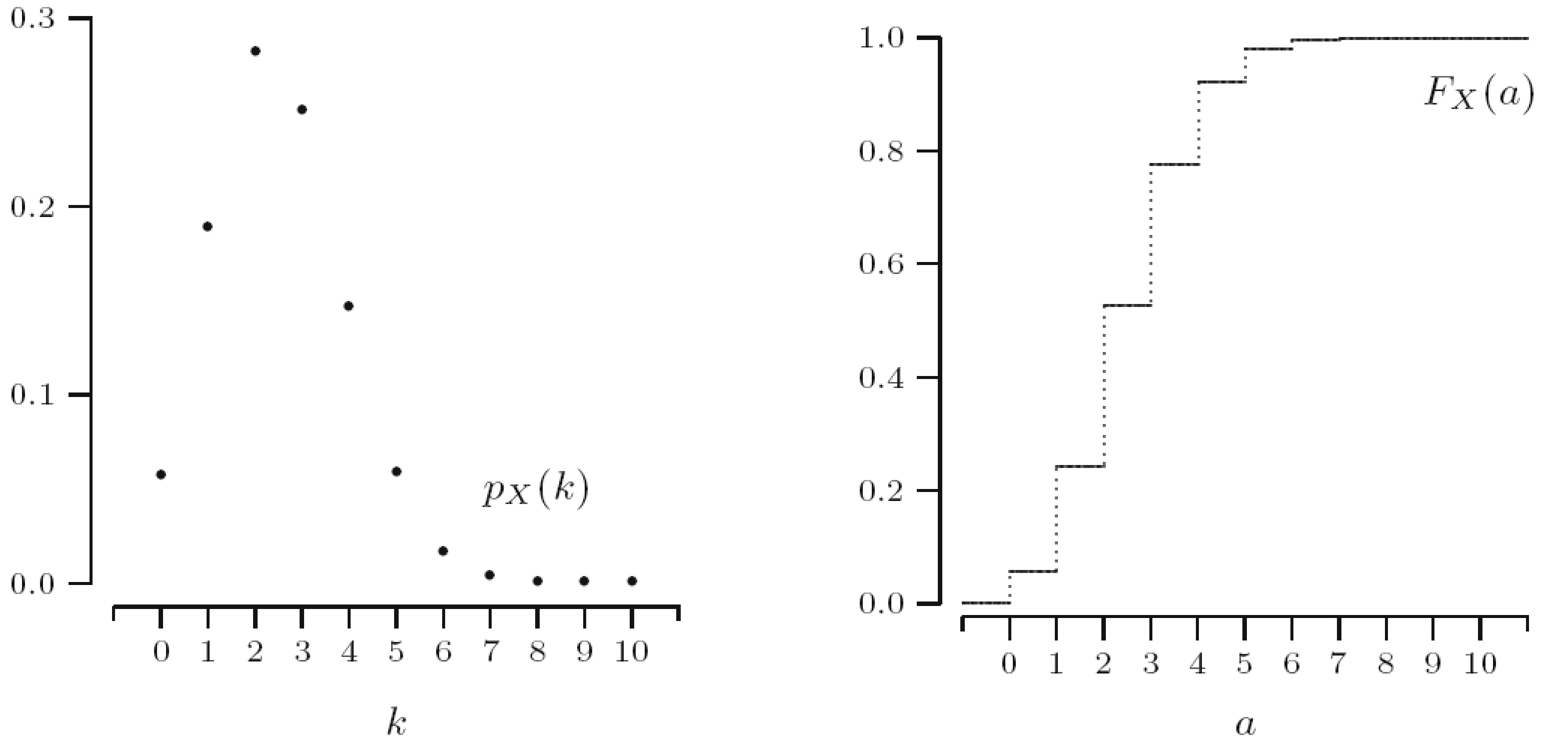
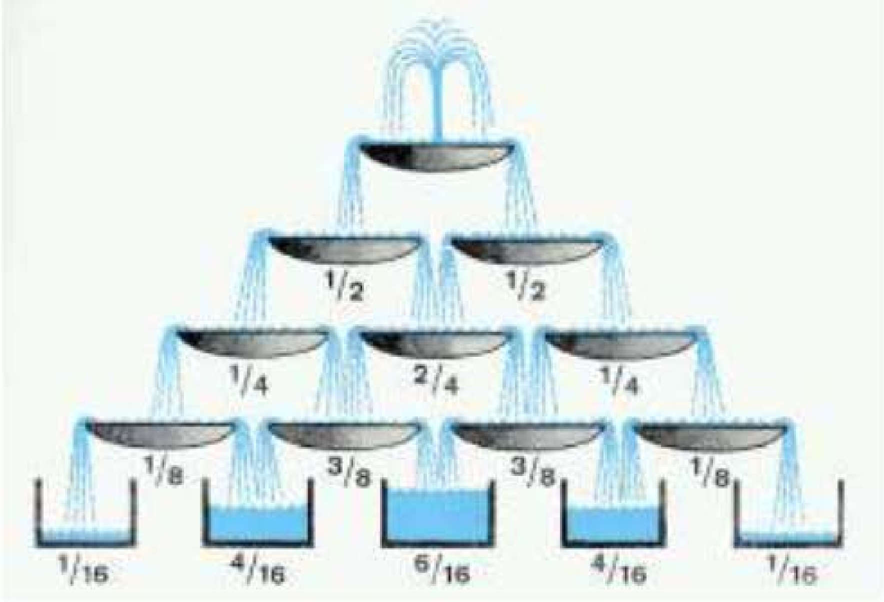


Fig. 4.2. Probability mass function and distribution function of the $Bin(10, \frac{1}{4})$ distribution.

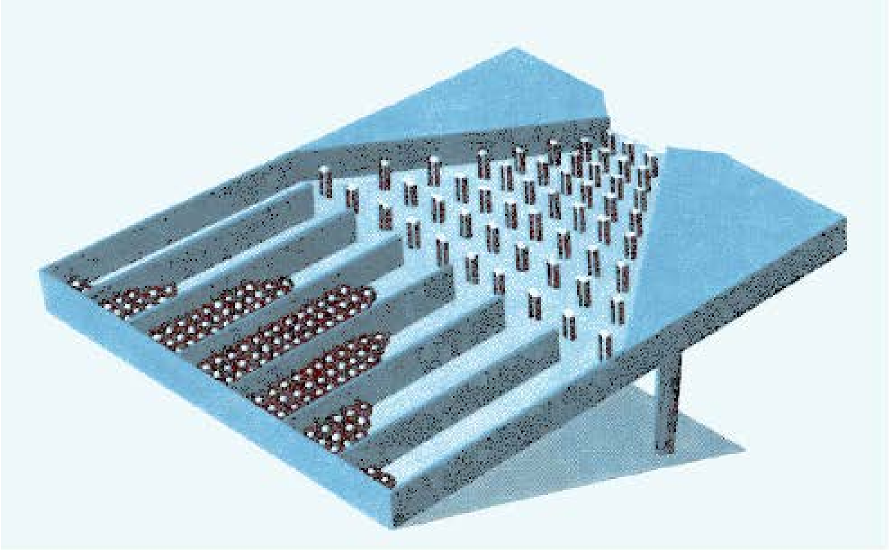
- Pravidelný čtyřstěn; 10 hodů; $p_X(k)$ nám říká, jaká je pravděpodobnost, že při 10 hodech padne třeba číslo 2 právě *k*-krát.

Binomické rozdělení

- Př.: model římské kašny nebo Galtonovo prkno (princip hracích automatů)



Binomické rozdělení zobrazené pomocí modelu římské kašny – nádržky se naplní podle Pascalova trojúhelníku: 1:4:6:4:1



Galtonovo prkno s kolíčky – kuličky se náhodným mechanismem rozdělí do přihrádek podle Pascalova trojúhelníku: 1:6:15:20:15:6:1

Binomické rozdělení

- Speciální případy
  - Pokud n=1 a tedy k={0, 1}, pak dostaneme Bernoulliho rozdělení.
  - Pokud $n \rightarrow \infty$ a $p \rightarrow 0$ pak dostaneme Poissonovo rozdělení
  - Pokud $p \rightarrow 1/2$, lze pro *n* v řádu několika desítek dobře aproximovat Gaussovým rozdělením.

Poissonovo rozdělení

- Je limitním případem binomického rozdělení pro jevy, které jsou velmi řídké (málo pravděpodobné), tedy jejich pravděpodobnost p limitně jde k nule. Ale musím provést velmi mnoho opakování pokusu.
- V takovém případě, lze binomické rozdělení dobře aproximovat Poissonovým rozdělením.
- Jednotlivé jevy musí být **nezávislé**.
- Musím znát střední četnost výskytu jevů. Je to zásadní parametr.
- Nemá smysl se ptát po počtu událostí, které nenastaly.
- Podle tohoto rozdělení se řídí četnosti jevů, které jsou málo časté, např.: počet poruch výrobního zařízení v časové jednotce, počet signálů v telefonní ústředně za jednotku času, počet částic za jednotku času registrovanou Geiger-Müllerovým počítačem.

Poissonovo rozdělení

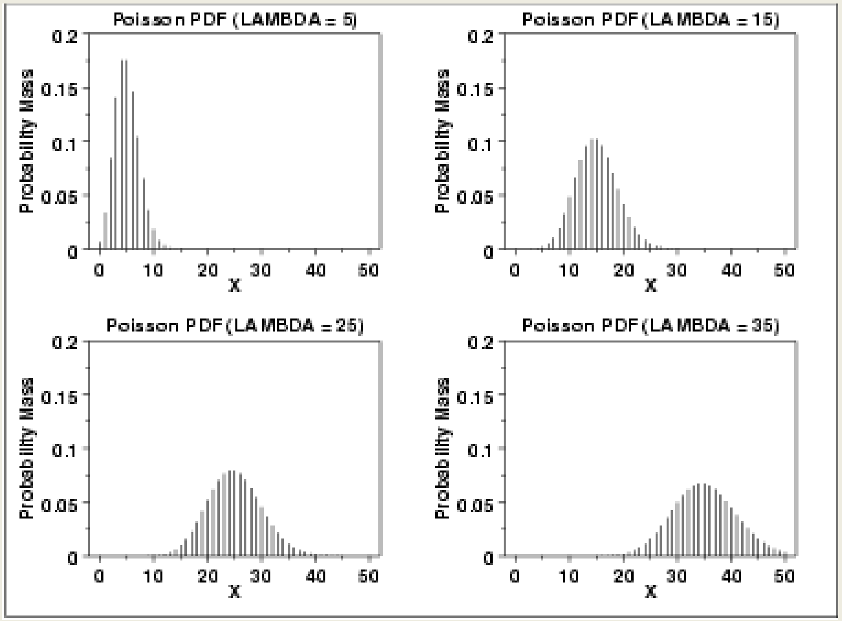
- Typicky pro n > 30 a p < 0,1 lze binomické rozdělení velmi dobře aproximovat Poissonovým rozdělením.
- Pravděpodobnostní funkce:

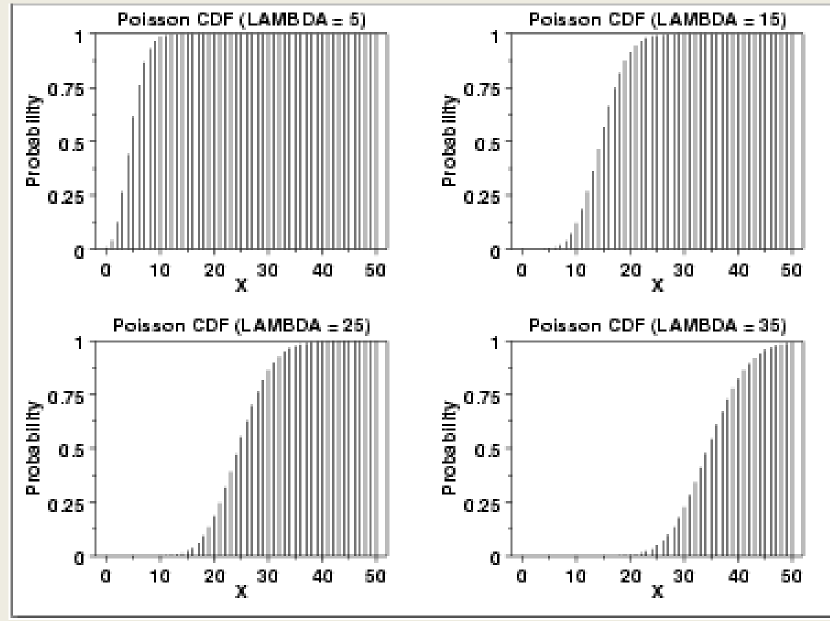
  $p_X(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$, kde λ = n·p,
  pro k=0,1,2… Pro ostatní *k* je $p_X(k)$ = 0.

- Značíme ho **Poiss(λ)**. Konstantu λ nazýváme „očekávaná hodnota" a *k* je počet hledaných jevů.
- λ je v podstatě průměrná (střední) hodnota (četnost), že daný náhodný jev nastane. Je to jediný parametr, který odlišuje jednotlivá Poissonova rozdělení.
- Pokud pro jiný náhodný experiment bude λ třeba 10x větší, pak nelze mechanicky vynásobit 10 celou pravděpodobnostní funkci.
- Distribuční funkce:

$$F(a) = \sum_{k=0}^{a} \frac{\lambda^k}{k!} e^{-\lambda},$$

Poissonovo rozdělení





- Všimněte si, že pro velká λ je Poissonovo rozdělení podobné normálnímu rozdělení. Ale podstata náhodných jevů, které obě rozdělení popisují je zcela odlišná!!!

Geometrické rozdělení

- Nezávislé opakování náhodného experimentu – házení mincí.
- Možnými výsledky jsou jen dvě možné hodnoty. Např. (úspěch, neúspěch) nebo (1, 0). Pravděpodobnost, že jev nastane je *p*.

DEFINITION. A discrete random variable X has a *geometric distribution* with parameter p, where $0 < p \leq 1$, if its probability mass function is given by

$$p_X(k) = \mathrm{P}(X = k) = (1-p)^{k-1}p \quad \text{for } k = 1, 2, \ldots .$$

We denote this distribution by $Geo(p)$.

- Náhodná veličina X udávající počet neúspěšných opakování pokusu (0) dokud nenastane úspěch (1) (čekání na první úspěch) má **geometrické** rozdělení pravděpodobnosti s parametrem *p*.

Geometrické rozdělení

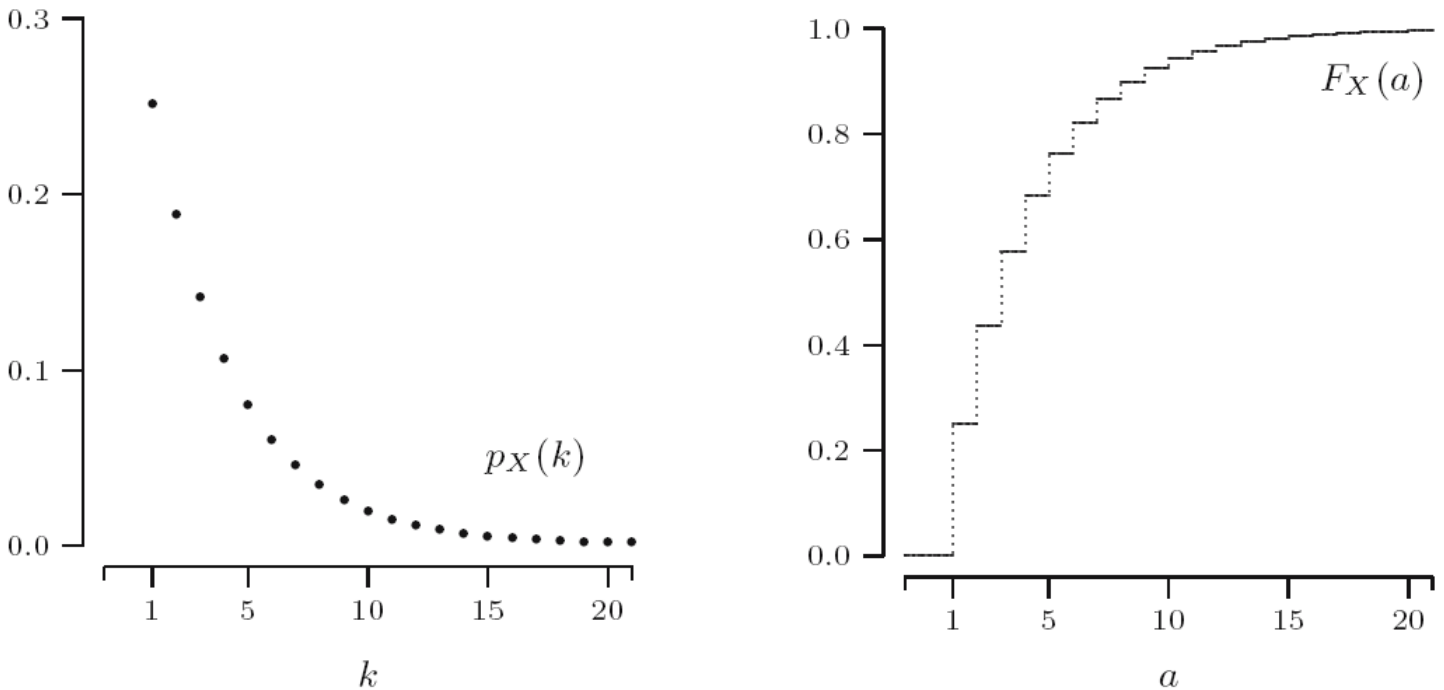


Fig. 4.3. Probability mass function and distribution function of the $Geo(\frac{1}{4})$ distribution.

Hypergeometrické rozdělení

- Provádíme **závislé** opakování náhodného experimentu.
- Typickým příkladem je náhodný výběr ze souboru bez vracení prvku zpátky a tím tedy zabráníme jeho opětovnému výběru.
- Tedy pravděpodobnost vybrání prvku ze souboru je závislá na předchozích náhodných experimentech.
- Př.: nechť máme soubor s N prvky, z nichž M vykazuje nějaký sledovaný znak V. Vybereme náhodně *n* jednotek a nevracíme je zpět. Pravděpodobnost vytažení jednotky je vždy stejná. Jaká je pravděpodobnost, že vybereme *k* prvků s vlastností V a *n-k* prvků, které znak V nemají?
- Počet všech možných případů je dán počtem všech možností výběru *n* prvků z N. Tedy je to kombinace *n*-té třídy bez opakování z N prvků (každá *n*-tice prvků) → $\binom{N}{n}$. Počet příznivých případů je $\binom{M}{k}$ a z počtu zbylých N-M prvků lze vybrat *n-k* prvků $\binom{N-M}{n-k}$ způsoby.
- Diskrétní náhodná veličina X má hypergeometrickou pravděpodobnostní distribuci s parametry N, M, n, pro něž platí $1 \leq \mathrm{n} < \mathrm{N}$ a $1 \leq \mathrm{M} < \mathrm{N}$

$$p_X(k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

Spojitá náhodná proměnná

Spojitá náhodná proměnná

- Existuje mnoho náhodných experimentů, kde jejich výsledkem je spojitá veličina.
- Potom tyto výsledky musí mít přirozeně i spojitou náhodnou proměnnou.
- Jak získat spojitou X z diskrétní X?
- Můžeme jít cestou **neustálého zpřesňování** diskrétní náhodné proměnné X spojené s daným náhodným experimentem s pravděp. funkcí *p* = P(X=k). Tedy X bude nabývat nějaké konkrétní číselné hodnoty (třeba 1,234).
- Tj. budeme velikost diskrétní náhodné proměnné znát stále s větší přesností (s větším počtem desetinných míst→1,2340…1,2349).
- Tedy pravděpodobnost P’, že X bude nabývat číselné hodnoty udané s větší přesností bude menší. Součet všech P’ musí být roven P.
- **Pro nekonečně přesně udanou hodnotu X bude pravděpodobnost nulová**.
- Nicméně pravděpodobnost, že X bude nabývat číselné hodnoty v nějakém pevném intervalu X = [a, b] bude daná → P(X=[a, b]) = p.
- Matematicky to odpovídá způsobu konvergence sumy v integrál.

Hustota pravděpodobnosti

- Rozdělení pravděpodobnosti spojité náhodné proměnné určujeme pomocí spojité funkce nazývané **hustota pravděpodobnosti**.

DEFINITION. A random variable X is *continuous* if for some function $f : \mathbb{R} \to \mathbb{R}$ and for any numbers a and b with $a \leq b$,

$$\mathrm{P}(a \leq X \leq b) = \int_a^b f(x)\,\mathrm{d}x.$$

The function f has to satisfy $f(x) \geq 0$ for all x and $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$. We call f the *probability density function* (or *probability density*) of X.

Hustota pravděpodobnosti

- Pravděpodobnost, že náhodná proměnná X leží v intervalu [a, b] je daná plochou pod funkcí hustoty pravděpodobnosti *f* na intervalu [a, b].
- Pokud interval konverguje k nule (a -> b), pak P(X=a) = 0.
- Jaká je interpretace f(a)?
- Je to relativní míra, jak je pravděpodobné, že X je „blízko“ k *a*. Tedy X ∈ <a-ε, a+ε>, kde *ε* je infinitezimálně malé.
- **Pozor**: f(a) není pravděpodobnost!!! na rozdíl od pravděpodobnostní funkce diskrétní náhodné proměnné.

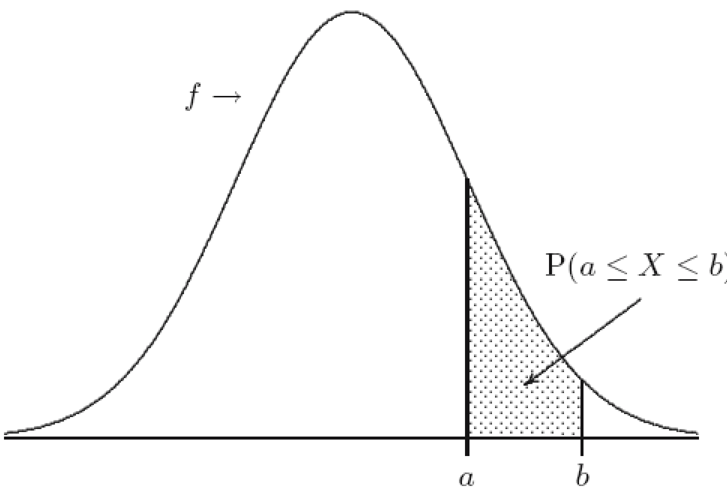


Fig. 5.1. Area under a probability density function f on the interval $[a, b]$.

Hustota pravděpodobnosti

- Důležité si zapamatovat:
  - Diskrétní náhodná proměnná nemá hustotou pravděpodobnosti *f*.
  - Spojitá náhodná proměnná nemá pravděpodobnostní funkci *p*.
  - Ale obě náhodné proměnné mají svou distribuční funkci $F(a) = P(X \leq a)$.

Distribuční funkce

- Pro distribuční funkci spojité náhodné proměnné platí, že 1) F(a) musí ležet v intervalu [0, 1] a 2) je to neklesající funkce na prostoru reálných čísel.
- Důkaz 2): Nechť a<b, potom jev, že X≤*b* lze najít jako sloučení dvou jevů: $X \leq a \cup a < X \leq b$. Pravděpodobnost, že X bude v intervalu (a, b> lze vyjádřit pomocí distribuční funkce: $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$.
- Z matematické analýzy plyne, že: $F(b) = \int_{-\infty}^{b} f(x)\,\mathrm{d}x$ and[1] $f(x) = \frac{\mathrm{d}}{\mathrm{d}x}F(x)$
- Dostali jsme vztah mezi distribuční funkcí a hustotou pravděpodobnosti spojité náhodné proměnné.
- Tedy *F(x)* a *f(x)* obsahují kompletní pravděpodobnostní informaci o X.

Rovnoměrné rozdělení

- Výsledek náhodného experimentu je úplně libovolně rozložen v nějakém intervalu.
- Typickým příkladem je detekce radioaktivního záření Geiger-Müllerovým čítačem během třeba jedné hodiny. Tedy interval je [0, 60] minut.
- Záznamy detekce záření budou rovnoměrně rozloženy v intervalu 60 minut.
- Tedy stejné libovolné podintervaly měly by mít stejnou pravděpodobnost, že tam nalezneme záznam.
- Tedy hustota pravděpodobnosti spojená s tímto experimentem musí být konstantní na celém intervalu 60 minut.

Rovnoměrné rozdělení

DEFINITION. A continuous random variable has a *uniform distribution* on the interval $[\alpha, \beta]$ if its probability density function f is given by $f(x) = 0$ if x is not in $[\alpha, \beta]$ and

$$f(x) = \frac{1}{\beta - \alpha} \quad \text{for } \alpha \leq x \leq \beta.$$

We denote this distribution by $U(\alpha, \beta)$.

- Distribuční funkce je:

  F(x) = (x-α)/(β-α)

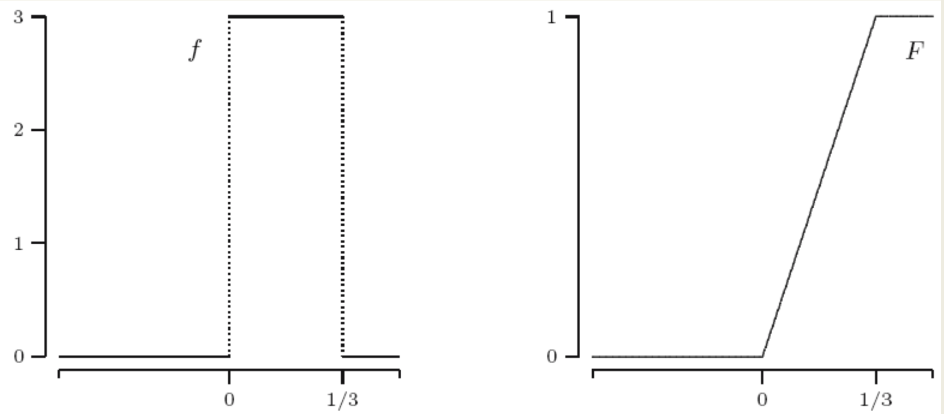


Fig. 5.3. The probability density function and the distribution function of the $U(0, \frac{1}{3})$ distribution.

Exponenciální rozdělení

- Je to obdoba geometrického rozdělení pro spojitou náhodnou proměnnou. Tedy „počet náhodných experimentů“ limitně jde k nekonečnu: $n \to \infty$.
- Příkladem je třeba doba setrvání částice v kapalině (částice jsou v kapalině rovnoměrně rozptýlené), která proudí nějakou nádobou.
- Existují dva stavy: opustí nádobu (1) nebo neopustí nádobu (0) za dobu *n* časových intervalů (třeba 1 sekunda).
- Náhodná proměnná *T* je čas setrvání v kapalině, neboli kolik nastane neúspěšných časových intervalů, kdy částice neopustí nádobu, než dojde k úspěchu (1).
- Dostaneme geometrické rozdělení. Pokud časový interval budeme nekonečně zkracovat, nebo-li počet intervalů limitně poroste do nekonečna, pak dostaneme pravděpodobnostní rozdělení pro spojitou náhodnou proměnnou.
- Jiným příkladem je rozpadový zákon radioaktivního materiálu, kdy studuji časovou závislost mezi náhodně se opakujícími jevy (radioaktivními rozpady), které mají všechny stejnou pravděpodobnost.

Exponenciální rozdělení

DEFINITION. A continuous random variable has an *exponential distribution* with parameter λ if its probability density function f is given by $f(x) = 0$ if $x < 0$ and

$$f(x) = \lambda e^{-\lambda x} \quad \text{for } x \geq 0.$$

We denote this distribution by $Exp(\lambda)$.

The distribution function F of an $Exp(\lambda)$ distribution is given by

$$F(a) = 1 - e^{-\lambda a} \quad \text{for} \quad a \geq 0.$$

Exponenciální rozdělení

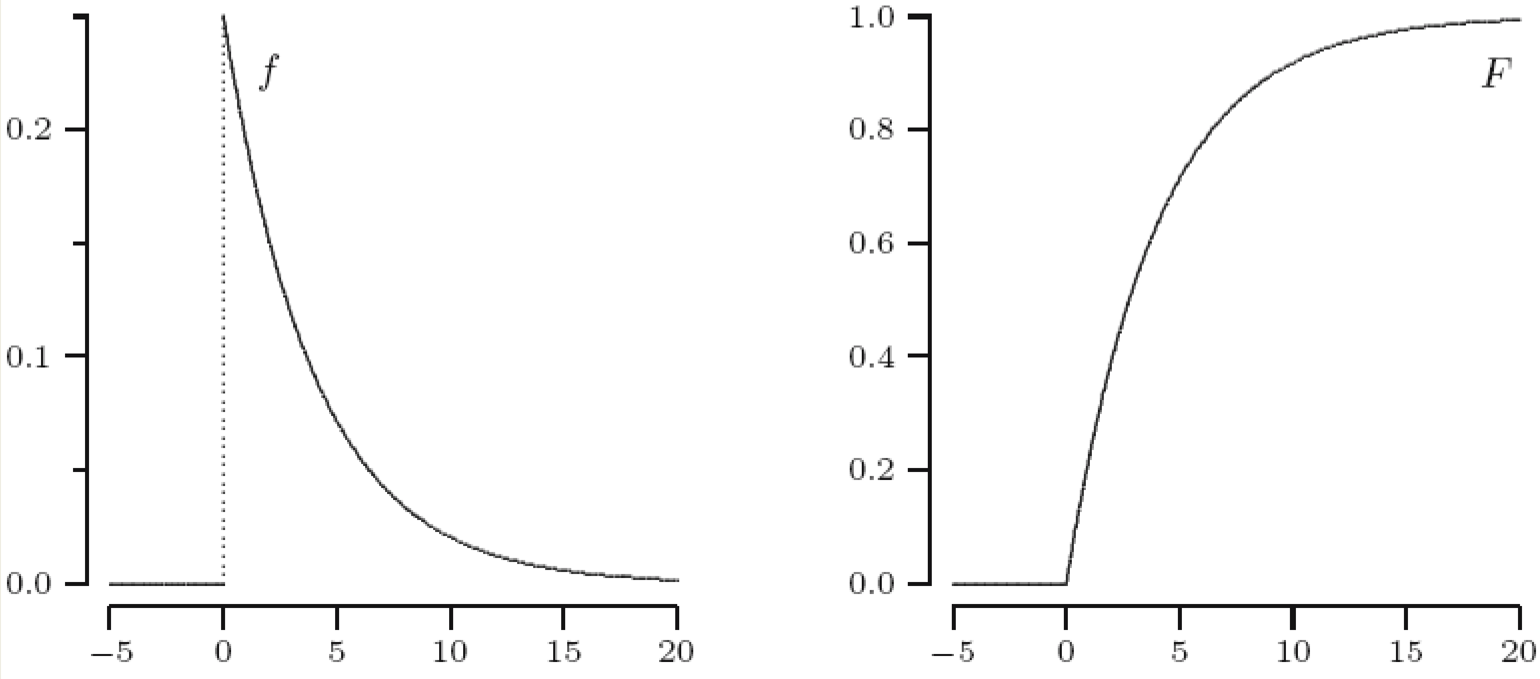


Fig. 5.4. The probability density and the distribution function of the *Exp*(0.25) distribution.

Normální rozdělení

- Nazývá se také jako Gaussovo rozdělení.
- Je to hlavně myšlenkový model. Podstata spočívá ve skutečnosti, že většina kvantitativních náhodných proměnných je rozložena tak, že jeden náhodný jev je nejčetnější. Ostatní náhodné jevy jsou náhle méně četné až konečně vykazují jen ojedinělou extrémní hodnotu. Příkladem je třeba rozložení hodnoty IQ mezi obyvatelstvem.
- První příklady pozorování normálního rozdělení podal matematik Quételet, který měřil objem hrudníku francouzských vojáků. Matematik Gauss aplikoval toto rozdělení na model pozorovacích chyb v astronomii.

Normální rozdělení

Definition. A continuous random variable has a *normal distribution* with parameters μ and $\sigma^2 > 0$ if its probability density function f is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{for} \quad -\infty < x < \infty.$$

We denote this distribution by $N(\mu, \sigma^2)$.

$$F(a) = \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \mathrm{d}x \quad \text{for } -\infty < a < \infty$$

- Parametr μ mi určuje posunutí funkce na ose *x* a parametr σ mi určuje strmost a tedy pološířku křivky.

Normální rozdělení

- Bohužel distribuční funkce normálního rozdělení N(μ, σ²) nemá neurčitý integrál a nedá se vyjádřit analyticky.

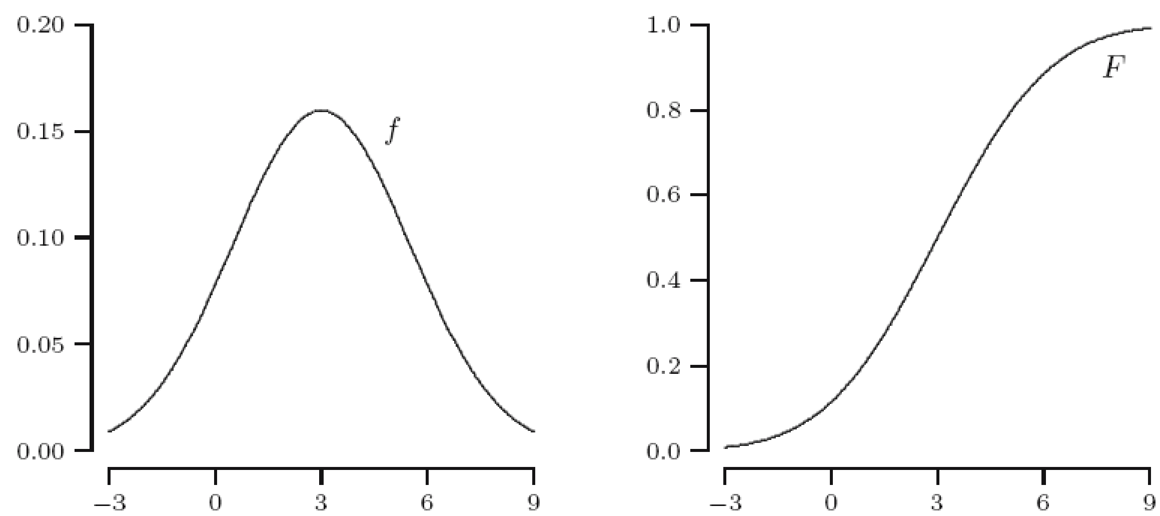
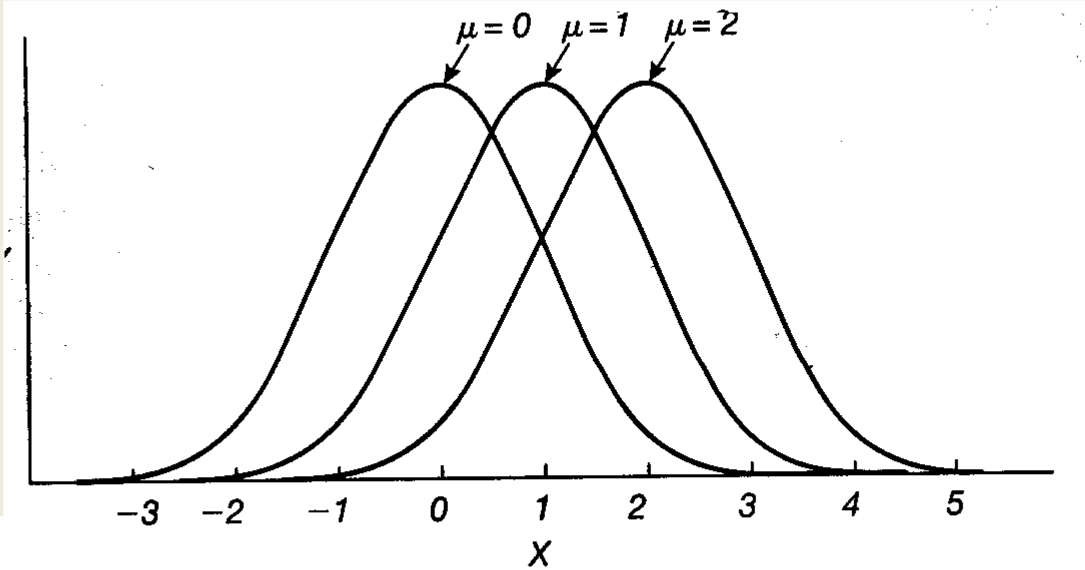
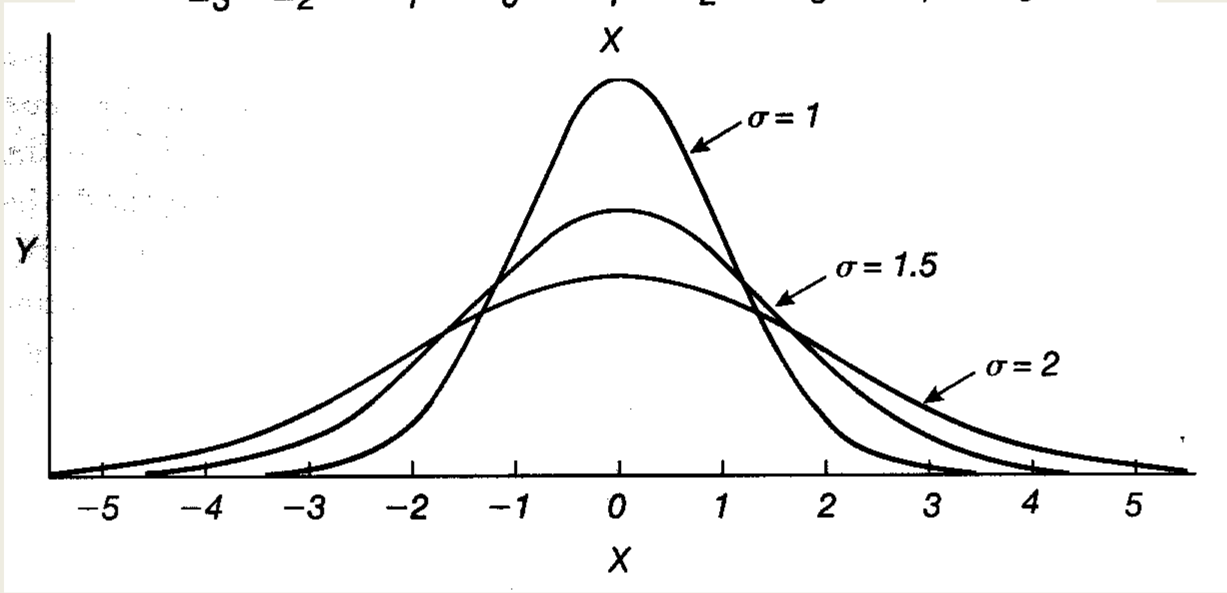


Fig. 5.6. The probability density and the distribution function of the $N(3, 6.25)$ distribution.

Normální rozdělení





Normální rozdělení

- V praxi se často používá tzv. **standardizované normální rozdělení** - N(0, 1).
- Výraz pro hustota pravděpodobnosti pak má jednodušší tvar:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2} \quad \text{for } -\infty < x < \infty$$

- N(0,1) hustota pravděpodobnosti je symetrická kolem nuly.
- Díky jednoduché transformaci, lze každé normální rozdělení zapsat jako N(0,1).

Normální rozdělení

Table B.1. Right tail probabilities $1 - \Phi(a) = P(Z \geq a)$ for an $N(0,1)$ distributed random variable Z.

a	0	1	2	3	4	5	6	7	8	9
0.0	5000	4960	4920	4880	4840	4801	4761	4721	4681	4641
0.1	4602	4562	4522	4483	4443	4404	4364	4325	4286	4247
0.2	4207	4168	4129	4090	4052	4013	3974	3936	3897	3859
0.3	3821	3783	3745	3707	3669	3632	3594	3557	3520	3483
0.4	3446	3409	3372	3336	3300	3264	3228	3192	3156	3121
0.5	3085	3050	3015	2981	2946	2912	2877	2843	2810	2776
0.6	2743	2709	2676	2643	2611	2578	2546	2514	2483	2451
0.7	2420	2389	2358	2327	2296	2266	2236	2206	2177	2148
0.8	2119	2090	2061	2033	2005	1977	1949	1922	1894	1867
0.9	1841	1814	1788	1762	1736	1711	1685	1660	1635	1611
1.0	1587	1562	1539	1515	1492	1469	1446	1423	1401	1379
1.1	1357	1335	1314	1292	1271	1251	1230	1210	1190	1170
1.2	1151	1131	1112	1093	1075	1056	1038	1020	1003	0985
1.3	0968	0951	0934	0918	0901	0885	0869	0853	0838	0823
1.4	0808	0793	0778	0764	0749	0735	0721	0708	0694	0681
1.5	0668	0655	0643	0630	0618	0606	0594	0582	0571	0559
1.6	0548	0537	0526	0516	0505	0495	0485	0475	0465	0455
1.7	0446	0436	0427	0418	0409	0401	0392	0384	0375	0367
1.8	0359	0351	0344	0336	0329	0322	0314	0307	0301	0294
1.9	0287	0281	0274	0268	0262	0256	0250	0244	0239	0233
2.0	0228	0222	0217	0212	0207	0202	0197	0192	0188	0183
2.1	0179	0174	0170	0166	0162	0158	0154	0150	0146	0143
2.2	0139	0136	0132	0129	0125	0122	0119	0116	0113	0110
2.3	0107	0104	0102	0099	0096	0094	0091	0089	0087	0084
2.4	0082	0080	0078	0075	0073	0071	0069	0068	0066	0064
2.5	0062	0060	0059	0057	0055	0054	0052	0051	0049	0048
2.6	0047	0045	0044	0043	0041	0040	0039	0038	0037	0036
2.7	0035	0034	0033	0032	0031	0030	0029	0028	0027	0026
2.8	0026	0025	0024	0023	0023	0022	0021	0021	0020	0019
2.9	0019	0018	0018	0017	0016	0016	0015	0015	0014	0014
3.0	0013	0013	0013	0012	0012	0011	0011	0011	0010	0010
3.1	0010	0009	0009	0009	0008	0008	0008	0008	0007	0007
3.2	0007	0007	0006	0006	0006	0006	0006	0005	0005	0005
3.3	0005	0005	0005	0004	0004	0004	0004	0004	0004	0003
3.4	0003	0003	0003	0003	0003	0003	0003	0003	0003	0002

Charakteristiky náhodných proměnných

Charakteristiky náhodných proměnných

- Pravděpodobnostní funkce (hustota pravděpodobnosti) nebo distribuční funkce podává o náhodné proměnné úplný obraz, ale často nepřehledný, komplikovaný, příliš mnoho informací atp.
- Často je vhodné shrnout informaci o náhodné proměnné do jediného čísla, které bude proměnnou dobře charakterizovat. Jinými slovy toto číslo bude nepochybně *očekávaná* hodnota náhodného experimentu.
- Výpočet charakterizujícího čísla je jednoznačně definován. Tyto čísla nazýváme jako tzv. *charakteristiky* náhodných proměnných.
- Charakteristiky:
  - Polohy: určují mi polohu očekávaných hodnot (maximum hustoty pravděpodobnosti, střední hodnota, průměr…)
  - Variability: pokud známe očekávanou hodnotu bude nás dále zajímat šířka rozdělovací funkce, nebo-li jak moc nebo málo se hodnoty rozdělení odchylují od očekávání.

Střední hodnota

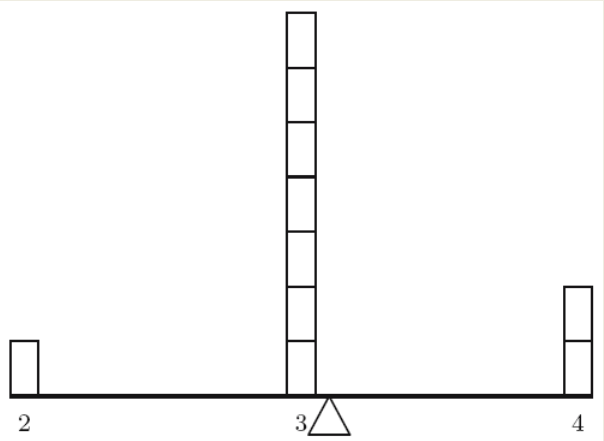
- Nazývá se také někdy jako očekávaná hodnota.
- Př.: vrták se opotřebuje s P=0,1 za 2 hodiny s P=0,7 za 3 hodiny a s P=0,2 za 4 hodiny. Jaká bude očekávaná hodnota opotřebení náhodně vybraného vrtáku?
- Srovnej s váženým průměrem.

DEFINITION. The *expectation* of a discrete random variable X taking the values $a_1, a_2, \ldots$ and with probability mass function p is the number

$$\mathrm{E}[X] = \sum_i a_i \mathrm{P}(X = a_i) = \sum_i a_i p(a_i).$$

Střední hodnota

- Tedy střední hodnota rozdělení náhodné proměnné (diskrétní nebo spojité) je vlastně vážený průměr rozdělení náhodné proměnné.
- Fyzikální význam střední hodnoty – **těžiště** pravděpodobnostního rozdělení $p(a_i)$.



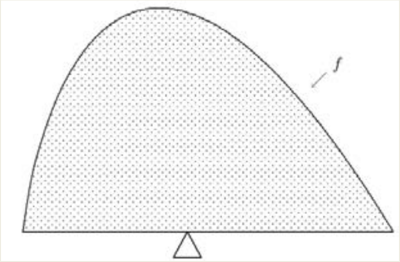
Střední hodnota

- Spojitá náhodná proměnná má střední hodnotu danou:

DEFINITION. The *expectation of a continuous random variable* X with probability density function f is the number

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x.$$

- E[X] je opět střední hodnota – těžiště plochy pod pravděpodobnostní funkcí.



Střední hodnota

- Pozn.: existují i takové pravděpodobnostní rozdělení, pro které neexistuje střední hodnota.
- Typický příklad je Cauchyho rozdělení (popisuje rozložení rezonančních stavů ve fyzice částic) s pravděpodobnostní funkcí:

$$f(x) = \frac{1}{\pi(1+x^2)} \quad \text{for } -\infty < x < \infty$$

- E[X] pro tento případ se rovná ∞.

Střední hodnota

- Základní matematické vlastnosti střední hodnoty pravděpodobnostního rozdělení
  - $E[\text{konst}] = \text{konst}$
  - $E[kX] = k\,E[X]$
  - $E[X+Y] = E[X] + E[Y]$
  - $E[X \cdot Y] = E[X] \cdot E[Y]$ (platí pro nezávislé proměnné)
  - $E[X+c] = E[X] + c$
- Důkaz na základě znalostí z matematické analýzy pro operace s integrály.

Střední hodnota složená proměnné

- Zajímá nás, jak nalézt hustotu pravděpodobnosti ne náhodné proměnné X ale nové proměnné $Y = X^2$.
- Př.: čtvercové budovy mají délku strany X – náhodná proměnná s rovnoměrným rozdělením v intervalu 0 až 10. Jaké bude rozdělení plochy budov X^2? X^2 musí tedy nabývat hodnot od 0 do 100.

$$F_Y(a) = \mathrm{P}\left(X^2 \le a\right) = \mathrm{P}\left(X \le \sqrt{a}\right) = \frac{\sqrt{a}}{10} \qquad f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y}F_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y}\frac{\sqrt{y}}{10} = \frac{1}{20\sqrt{y}}$$

- Hustota pravděpodobnosti i distribuční funkce se liší od f(X) a F(X).
- Jaká je střední hodnota plochy domů?

$$\mathrm{E}\left[X^2\right] = \mathrm{E}[Y] = \int_0^{100} y \cdot \frac{1}{20\sqrt{y}}\,\mathrm{d}y = \int_0^{100} \frac{\sqrt{y}}{20}\,\mathrm{d}y = \left[\frac{1}{20}\frac{2}{3}y^{\frac{3}{2}}\right]_0^{100} = 33\tfrac{1}{3}\ \mathrm{m}^2$$

Střední hodnota složená proměnné

- Je zřejmé, že pro výpočet střední hodnoty složené náhodné proměnné je potřeba použít správný výpočet, protože střední hodnota E[X] = 5 a $E^2[X]$ = 25. Což je rozdílné od $E[X^2]$ = 33,3.

THE CHANGE-OF-VARIABLE FORMULA. Let X be a random variable, and let $g : \mathbb{R} \to \mathbb{R}$ be a function.
If X is discrete, taking the values $a_1, a_2, \ldots$, then

$$\mathrm{E}\left[g(X)\right] = \sum_i g(a_i)\mathrm{P}(X = a_i)\,.$$

If X is continuous, with probability density function f, then

$$\mathrm{E}\left[g(X)\right] = \int_{-\infty}^{\infty} g(x)f(x)\,\mathrm{d}x.$$

Střední hodnota

- Střední hodnoty vybraných rozdělení

THE EXPECTATION OF A GEOMETRIC DISTRIBUTION. Let X have a geometric distribution with parameter p; then

$$\mathrm{E}[X] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \frac{1}{p}.$$

THE EXPECTATION OF AN EXPONENTIAL DISTRIBUTION. Let X have an exponential distribution with parameter λ; then

$$\mathrm{E}[X] = \int_0^{\infty} x\lambda \mathrm{e}^{-\lambda x}\, \mathrm{d}x = \frac{1}{\lambda}.$$

THE EXPECTATION OF A NORMAL DISTRIBUTION. Let X be an $N(\mu, \sigma^2)$ distributed random variable. Then

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \mathrm{d}x = \mu.$$

Aritmetický průměr

- Je to speciální případ střední hodnoty náhodné proměnné v případě, že pravděpodobnost každého jevu, kterého nabývá náhodná proměnná X je stejná.
- Jinými slovy X nabývá hodnot a_i právě jednou.
- $\bar{X} = \sum_{i=1}^{n} a_i P_i = \frac{1}{n} \sum_{i=1}^{n} a_i$

Modus

- Modus je definován jako nejpravděpodobnější hodnota náhodné proměnné.
- Hodnota modus se tedy v náhodném rozdělení vyskytuje s největší četností a představuje typickou hodnotu množiny náhodných jevů.

$$P[X=\tilde{x}] \geq P[X=x_i]$$

- Modus je definován stejnou nerovností i pro hustotu pravděpodobnosti spojité náhodné proměnné f(x).

Kvantily

- Jsou to charakteristiky, které mi rozdělují množinu náhodných proměnných na určité části. Tedy kvantil mi dělí prostor náhodných proměnných na část jevů, které nastanou s pravděpodobností menší nebo rovnou pravděpodobnosti odpovídající danému kvantilu a druhá část jevů nastane s pravděpodobností doplňkovou.
- Tedy kvantil *q* je taková hodnota X s odpovídající P, že všechny náhodné proměnné nabývajících hodnot menších než *q*, nastanou s pravděpodobností P (odpovídající hodnotě *p*).

DEFINITION. Let X be a continuous random variable and let p be a number between 0 and 1. The pth *quantile* or 100pth *percentile* of the distribution of X is the smallest number q_p such that

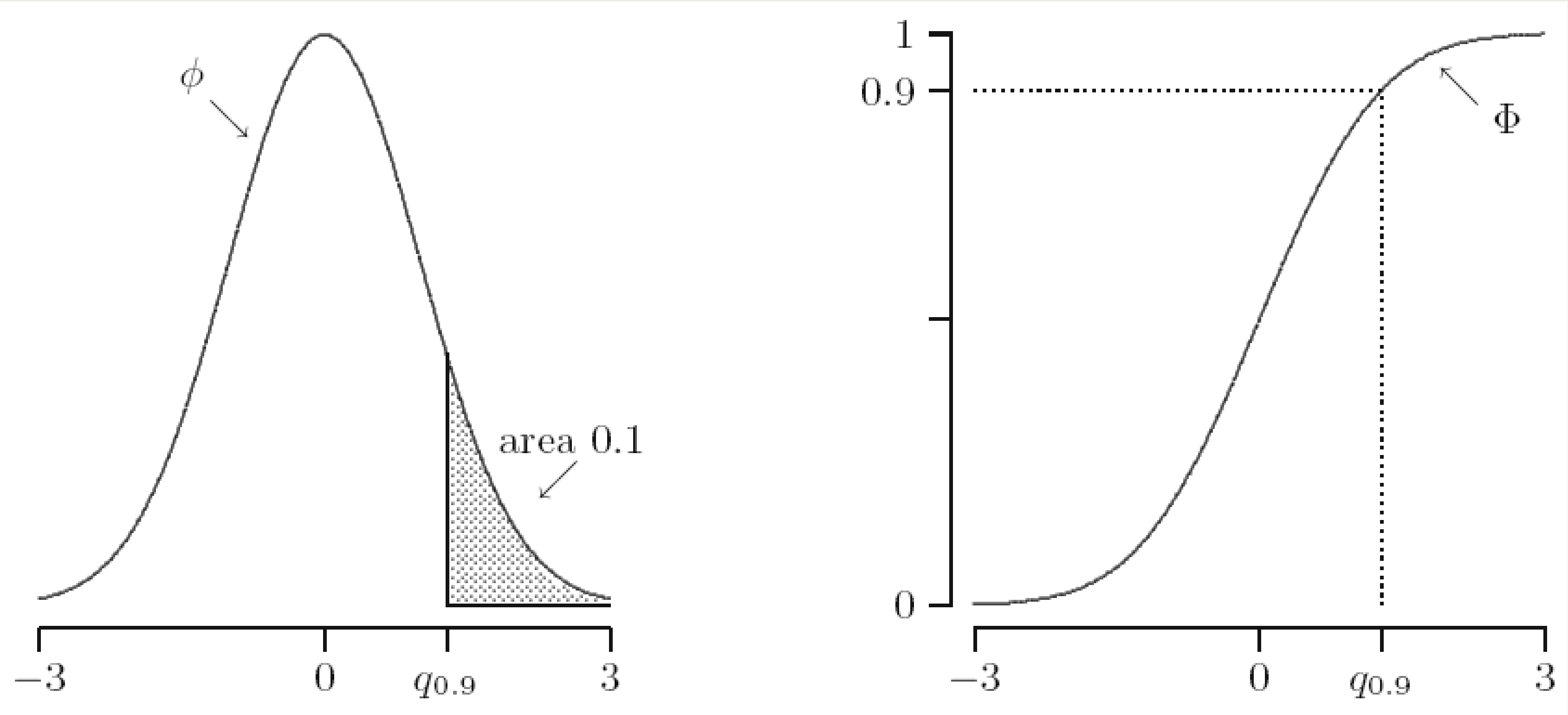
$$F(q_p) = \mathrm{P}(X \leq q_p) = p.$$

Kvantily

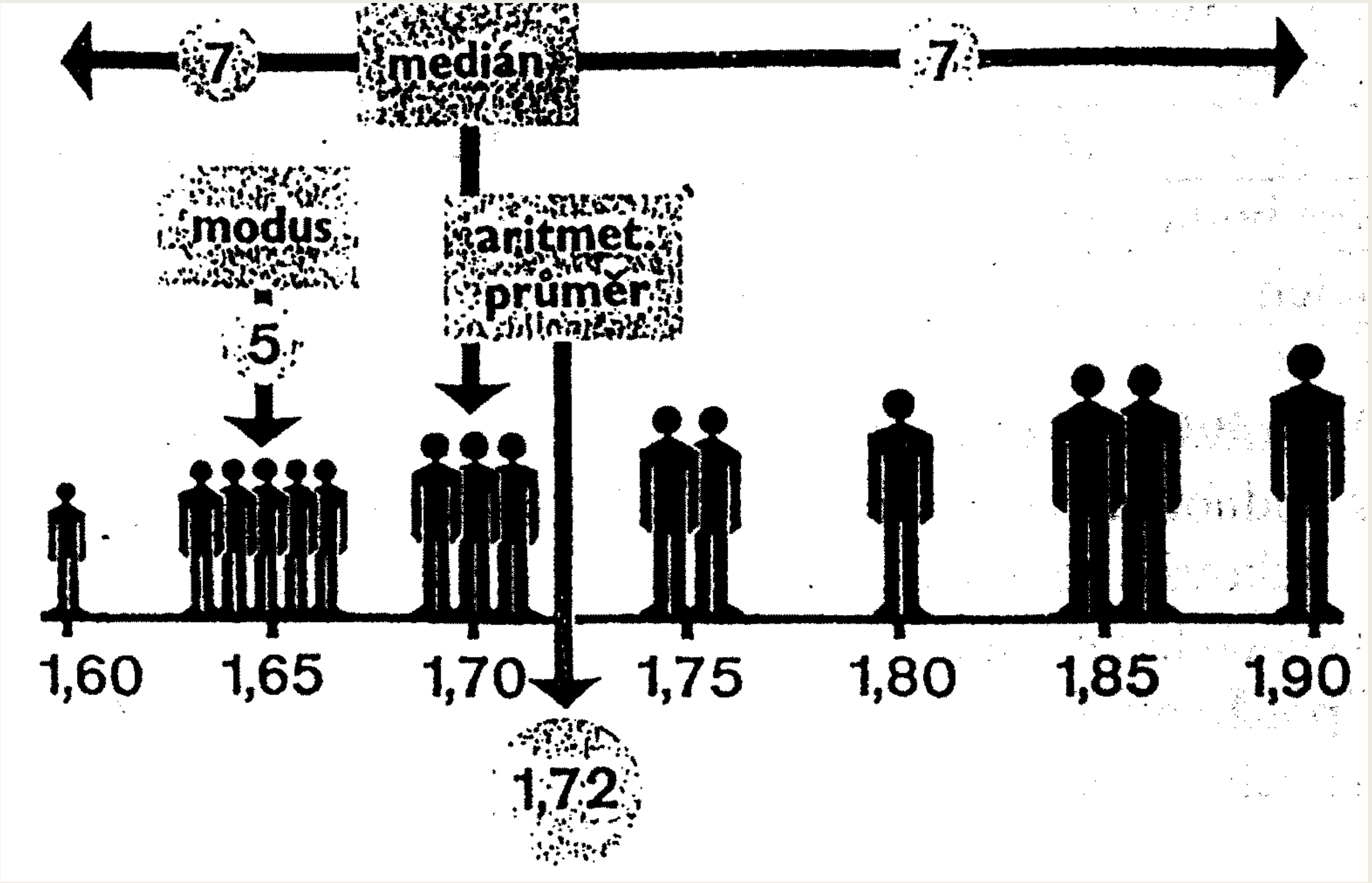
- Pokud mluvíme o kvantilu 0,25 nebo 0,50 nebo 0,75 nazýváme je **kvartily**.
- Pokud mluvíme o kvantilu 0,1 nebo 0,2 až 0,9 nazýváme je **decily**.
- Pokud mluvíme kvantilu 0,01 nebo 0,02 až 0,99 nazýváme je **centily**.
- **Medián** je 0,5 kvantil. Dělí tedy prostor náhodných proměnných na dvě stejně pravděpodobné části.
- Nabývá-li náhodná proměnná lichého počtu hodnot, pak je medián její prostřední hodnota.
- Nabývá-li náhodná proměnná sudého počtu hodnot, pak je medián roven aritmetickému průměru dvou prostředních hodnot.

Kvantily

- Příklad kvantilu 0,9 (90%) pro normální rozdělení.



Charakteristiky náhodných proměnných



- Aritmetický průměr je roven střední hodnotě.

Rozptyl

- Střední hodnota je obecně nazývána také jako tzv. obecný moment prvního řádu.
- **Obecné momenty** k-tého řádu $M_k[X]$ jsou charakteristiky náhodné proměnné X. Ve vztahu pro výpočet $M_k[X]$ je parametr a_i umocněn na *k*.
- $E[X]=M_{k=1}[X]$, tedy střední hodnota náhodné proměnné X je nejjednodušší obecný moment.
- Obecný moment k-tého řádu je definován vzhledem k nule.
- Pokud známe E[X], můžeme definovat momenty vzhledem ke střední hodnotě. Nazýváme je jako tzv. **centrální momenty**.
- Obecně je definován: $M_k[X] = M[X] – M[X]^k$. Pro k=1 je centrální moment roven nule. Tedy náhodná proměnná nabývá jen středních hodnot → to už není vůbec náhodná veličina, ale konstanta.
- Pro k=2 nazýváme centrální moment jako **rozptyl** (disperze).

Rozptyl

DEFINITION. The *variance* Var(X) of a random variable X is the number

$$\mathrm{Var}(X) = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right].$$

- Disperze náhodné proměnné představuje rozptýlení hodnot náhodné proměnné kolem její střední hodnoty.
- Var (X) je vždy kladná.
- V praktických úlohách se můžeme často setkat s tzv. **směrodatnou odchylkou** definovanou jako √Var(X), protože má stejný rozměr jako střední hodnota (je vyjádřena ve stejných jednotkách).

Rozptyl

- Protože je často nepraktické používat k výpočtu Var(X) definiční vztah, lze odvodit užitečnější rovnici.

AN ALTERNATIVE EXPRESSION FOR THE VARIANCE. For any random variable X,

$$\mathrm{Var}(X) = \mathrm{E}\left[X^2\right] - (\mathrm{E}[X])^2 .$$

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] \\
&= \int_{-\infty}^{\infty} (x - \mathrm{E}[X])^2 f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \left(x^2 - 2x\mathrm{E}[X] + (\mathrm{E}[X])^2\right) f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x - 2\mathrm{E}[X] \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x + (\mathrm{E}[X])^2 \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x \\
&= \mathrm{E}\left[X^2\right] - 2(\mathrm{E}[X])^2 + (\mathrm{E}[X])^2 \\
&= \mathrm{E}\left[X^2\right] - (\mathrm{E}[X])^2 .
\end{aligned}$$