Přednáška 2 – Diskrétní náhodná veličina, rozdělení

Náhodná veličina – veličina, kterou měříme a při opakováném měření dostáváme různé hodnoty

Realizace – naměřené hodnoty náhodné veličiny

Příklad:

Náhodná veličina	Realizace
výška	175, 188.3, 160.5, ...
známka	A, B, 2, D, 1, C, 4, . . .
váha	55, 100.5, 789.39, . . .
teplota	21.5, 36.6, 38, 99, . . .

Proč potřebujeme náhodnou veličinu?

- nejde naměřit všechna data
- popisná statistika – neúplný popis
- náhodná veličina – jako soubor úplně všech dat
  - úplný popis – pravděpodobnostní a distribuční funkce N.V.
- pravděpodobnosti jednotlivých realizací (nebo jejich intervalů)
  - známka A? známka B? teplota 20 – 25°C? cena cca 1000 Kč?

Diskrétní náhodná veličina $x \in \{1, 2, \ldots, n\}$

Naměříme pouze konečný nebo spočetný počet možných realizací

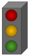
Příklad:

 kostka $\in \{1, 2, 3, 4, 5, 6\}$ – $\cancel{8}$, $\cancel{3.5}$

 bankovka $\in \{100, 200, 500, 1000, 2000\}$ – $\cancel{38}$, $\cancel{150}$

 mince $\in \{1, 2, 5, 10, 20, 50\}$ – $\cancel{3}$, $\cancel{8}$

známka $\in \{A, B, C, D, E\}$ – $\cancel{M}$, $\cancel{P}$

 semafor $\in$ {zelená, žlutá, červená} ~~modrá~~

 házení mincí $\in$ {rub, líc}

Ordinální vs. nominální D.N.V.

Ordinální – můžeme uspořádat

- kostka, bankovka, mince, známka

Nominální – nemůžeme uspořádat

- semafor, házení mincí

Pravděpodobnosti jednotlivých hodnot – rozdělení

- Rozdělení – funkce, která přiřazuje pravděpodobnosti jednotlivým hodnotám náhodné veličiny
- v případě diskrétní náhodné veličiny:

  rozdělení – pravděpodobnostní funkce

Pravděpodobnostní funkce (pf) diskrétní náhodné veličiny

$$f(x) = P(x = i), \quad x \in \{1, 2, \ldots, n\}$$

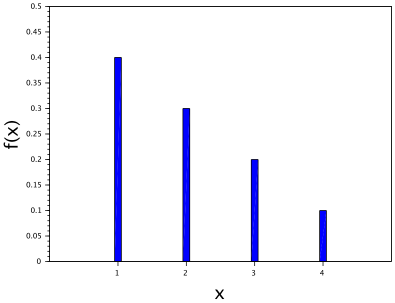
Tabulka:

Příklad: známka $x \in \{A = 1, B = 2, C = 3, D = 4\}$

x	1	2	3	4
$f(x)$	0.4	0.3	0.2	0.1

$\forall f(x_i) \geq 0, \quad \sum_{i=1}^{n} f(x_i) = 1$

Graf:



Předpis:

např., alternativní rozdělení:

$f(x) = p^x(1-p)^{1-x}$

Distribuční funkce (DF) diskrétní náhodné veličiny

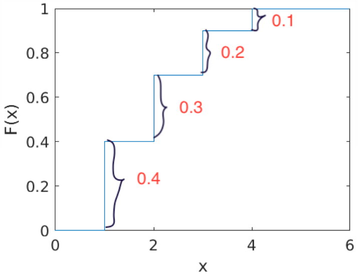
$$F(x) = P(x \le i), \quad x \in \{1, 2, \dots, n\}$$

neklesající, $\quad F(x) = 0$ pro $x < 1, \quad F(x) = 1$ pro $x \ge n$

Příklad: Přepočet $f(x) \to F(x)$

x	1	2	3	4
$f(x)$	0.4	0.3	0.2	0.1

$$F(x) = \begin{cases} 0 & x < 1 \\ 0 + 0.4 = 0.4 & 1 \le x < 2 \\ 0.4 + 0.3 = 0.7 & 2 \le x < 3 \\ 0.7 + 0.2 = 0.9 & 3 \le x < 4 \\ \underbrace{0.9 + 0.1 = 1}_{\text{vždy } 1} & x \ge 4 \end{cases}$$



Střední hodnota, rozptyl, směrodatná odchylka, modus

Příklad:

x	1	2	3	4
$f(x)$	0.4	0.3	0.2	0.1

Střední hodnota

pozor na ordinální data

$$E[x] = \sum_{i=1}^{N} x_i f(x_i) =$$

$$= 1 * 0.4 + 2 * 0.3 + 3 * 0.2 + 4 * 0.1 = 2$$

Rozptyl

$$D[x] = \sum_{i=1}^{N} (x_i - E[x])^2 f(x_i) = \sum_{i=1}^{N} x_i^2 f(x_i) - (E[x])^2 =$$

$$= 1 * 0.4 + 4 * 0.3 + 9 * 0.2 + 16 * 0.1 - 4 = 1$$

Odvození: $D[x] = E\left[(x - E[x])^2\right] = E\left[x^2 - 2xE[x] + (E[x])^2\right] =$
$= E[x^2] - 2E[x]E[x] + (E[x])^2 = E[x^2] - (E[x])^2$

Směrodatná odchylka

$$\sigma = \sqrt{D[x]} = 1$$

Modus

$$\hat{x} = \arg\max(f(x)) = 1$$

Kvantily

přes DF

hodnota |, průměr –

Významná rozdělení diskrétní náhodné veličiny

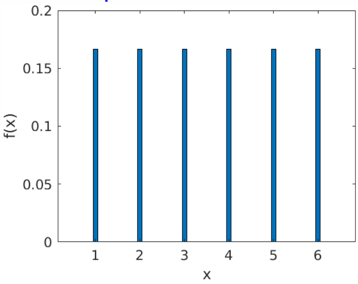
Rovnoměrné rozdělení – stejné pravděpodobnosti

$$f(x) = \frac{1}{n}, \quad x \in \{1, 2, \ldots, n\}, \quad n \text{ – počet hodnot}$$

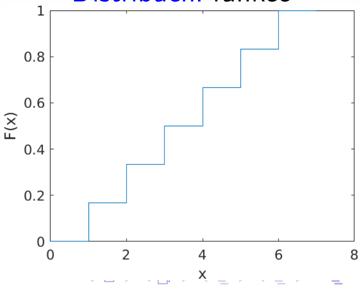
Příklad: hod kostkou 

x	1	2	3	4	5	6
$f(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Pravděpodobnostní funkce



Distribuční funkce



Alternativní rozdělení – úspěch/neúspěch (protiklady)

$$f(x) = p^x(1-p)^{1-x}, \quad x \in \{0,1\} \text{ – pouze 2 možné hodnoty}$$

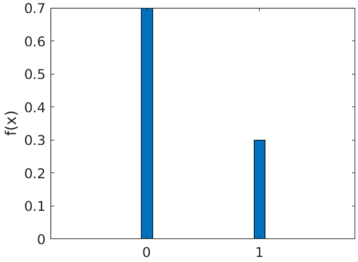
1 - úspěch, 0 - neúspěch, p – pravděpodobnost úspěchu

Příklad:
vadný/kvalitní výrobek, odbočení doleva/doprava, uspěl/neuspěl, ANO/NE

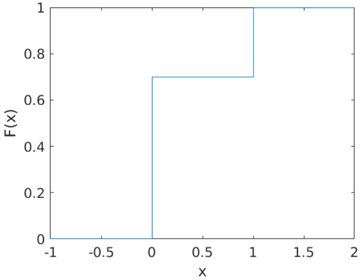
$$0.3^1(1-0.3)^{1-1} = 0.3$$

x	0	1
$f(x)$	0.7	0.3

Pravděpodobnostní funkce



Distribuční funkce



Základ pro další rozdělení

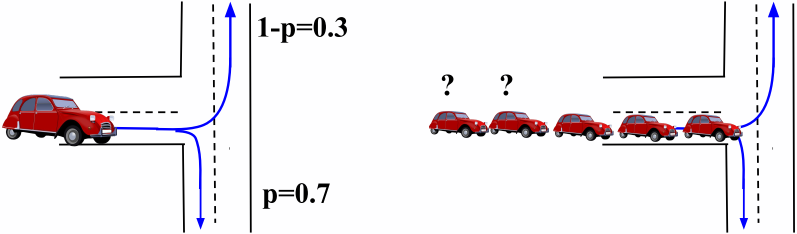
Binomické rozdělení – počet úspěchů v několika pokusech

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad \binom{n}{x} = \frac{n!}{(n-x)!x!}$$

x – počet úspěchů v n alternativních pokusech

p – pravděpodobnost úspěchu v jednotlivém pokusu

Příklad: Auto odbočí doprava s pravděpodobností $p = 0.7$.



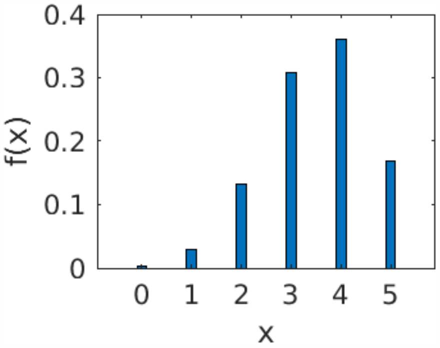
S jakou pravděpodobností 2 auta z 5 odbočí doprava?

$$\binom{5}{2} 0.7^2 (1-0.7)^{5-2} = \frac{5!}{(5-2)!2!} 0.7^2 0.3^3 = 0.1323$$

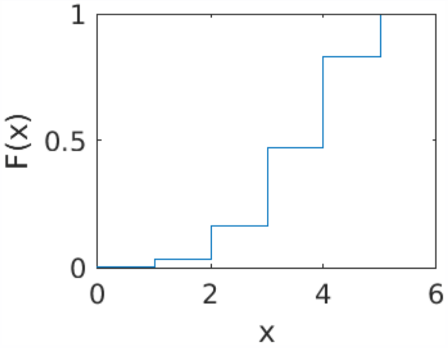
Příklady: 5 dívek z 10 novorozenců, 12 vadných výrobků z 50, atd.

Binomické rozdělení – pf a DF

Pravděpodobnostní funkce



Distribuční funkce



Poissonovo rozdělení – náhodné události za jednotku času

$$f(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad x \in \{0, 1, 2, \dots, n \to \infty\}, \; \lambda = E[x] = D[x]$$

- x – počet výskytů náhodných nezávislých událostí za jednotku času
- $n \Rightarrow \infty$ – velký spočetný počet pokusů
- $p \Rightarrow 0$ – malá pravděpodobnost události
- $n * p = \lambda$ – intenzita výskytu události (konstantní)

Příklad: Za hodinu projede křižovatkou v průměru 360 aut. S jakou pravděpodobností za minutu projede nejvýše 3 auta?

$\lambda = 360$ – intenzita za hodinu, $\quad \lambda = \frac{360}{60} = 6$ – intenzita za minutu

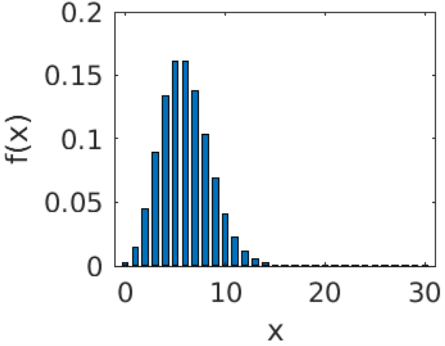
$f(0) = e^{-6}\frac{6^0}{0!} = e^{-6}$, $f(1) = e^{-6}\frac{6^1}{1!} = 6e^{-6}$,

$f(2) = e^{-6}\frac{6^2}{2!} = 18e^{-6}$, $f(3) = e^{-6}\frac{6^3}{3!} = \frac{216}{6}e^{-6} = 36e^{-6}$

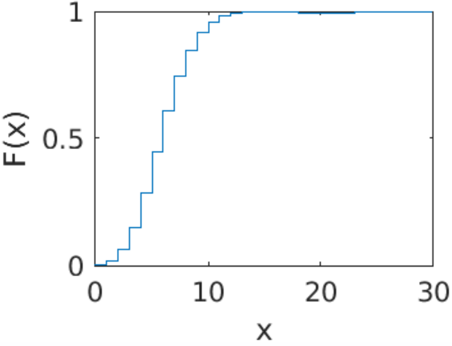
 $P(x \leq 3) = f(0) + f(1) + f(2) + f(3) = e^{-6}61$

Poissonovo rozdělení – pf a DF

Pravděpodobnostní funkce



Distribuční funkce



Geometrické rozdělení – neúspěch před prvním úspěchem

$$f(x) = p(1-p)^x$$

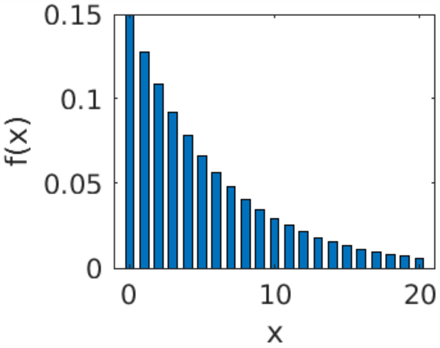
x – počet neúspěchů před prvním úspěchem
p – pravděpodobnost úspěchu

Příklad:  Voláme na zákaznickou linku. Pravděpodobnost, že se dovoláme je $p = 0.15$.

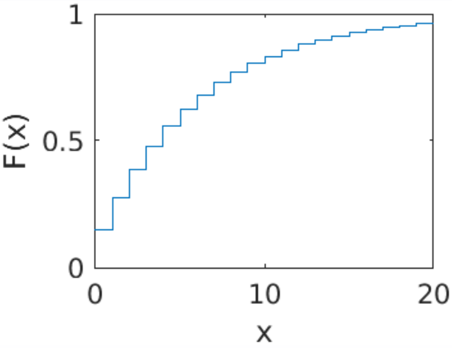
- S jakou pravděpodobností se dovoláme hned?
  $x = 0$ (zádný neúspěch), $\quad f(0) = 0.15(1-0.15)^0 = 0.15$
- Na druhý pokus?
  $x = 1$ (1 neúspěch), $\quad f(1) = 0.15(1-0.15)^1 = 0.128$
- Na 6.pokus?
  $x = 5$ (5 neúspěchů), $\quad f(5) = 0.15(1-0.15)^5 = 0.067$

Geometrické rozdělení – pf a DF

Pravděpodobnostní funkce



Distribuční funkce



Kategorické rozdělení – minimálně 3 možné realizace

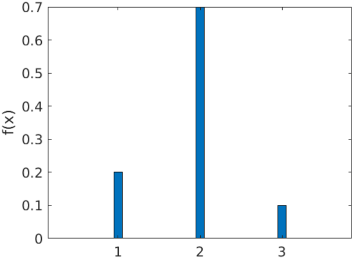
$$f(x) = p_x, \quad x \in \{1, 2, \dots, n\}, \text{ minimálně 3 možné hodnoty}$$

x	1	2	...	n
$f(x)$	p_1	p_2	...	p_n

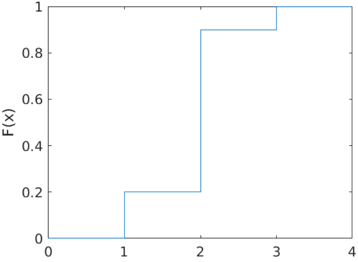
$$\forall f(x_i) \geq 0, \quad \sum_{i=1}^{n} f(x_i) = 1$$

Příklad: $x \in \{\text{kámen} = 1, \text{nůžky} = 2, \text{papír} = 3\}$

Pravděpodobnostní funkce



Distribuční funkce



Příklad: známka u zkoušky, druh dopravního prostředků,
typ motoru, atd

Jak poznám rozdělení diskrétní náhodné veličiny?

Rovnoměrné rozdělení	stejné pravděpodobnosti
Alternativní rozdělení	úspěch/neúspěch (protiklady)
Binomické rozdělení	počet úspěchů v několika alternativních pokusech
Poissonovo rozdělení	počet výskytů náhodných nezávislých událostí za jednotku času
Geometrické rozdělení	počet neúspěchů před prvním úspěchem
Kategorické rozdělení	minimálně 3 možné realizace