Jak modelovat výsledky náhodných pokusů? (část I.)

Martina Litschmannová

VŠB TECHNICKÁ UNIVERZITA OSTRAVA | FAKULTA ELEKTROTECHNIKY A INFORMATIKY | KATEDRA APLIKOVANÉ MATEMATIKY

Opakování: Základní pojmy z teorie pravděpodobnosti

- **Co je to náhodný pokus?**

  Děj, jehož výsledek není předem jednoznačně určen podmínkami, za nichž probíhá.

- **Co je to náhodný jev?**

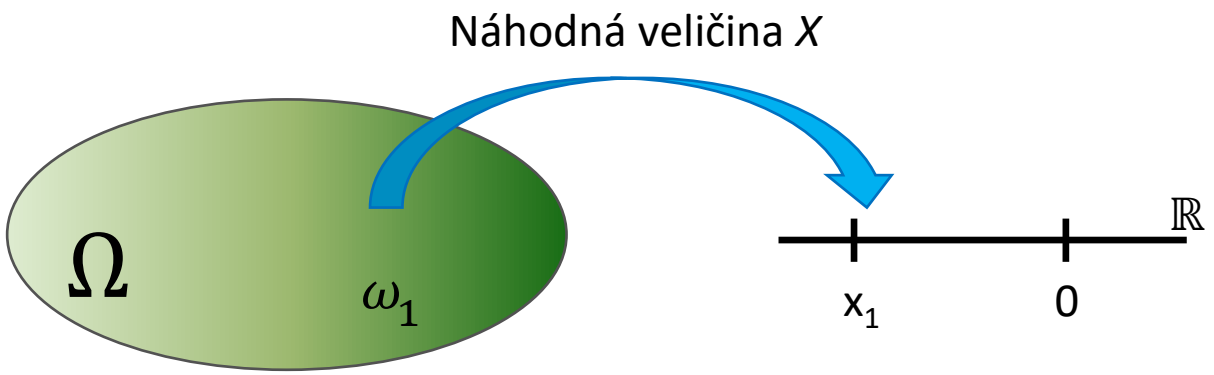
  Tvrzení o výsledku náhodného pokusu, přičemž o pravdivosti tohoto tvrzení lze po ukončení pokusu rozhodnout.

Jak modelovat výsledky náhodných pokusů? (část I.)

Definice

Náhodná veličina je funkce, která zobrazuje elementární jevy $\omega \in \Omega$ na reálná čísla.

- NV obvykle značíme velkými písmeny (X, Y, Z).
- NV přiřadí každému elementárnímu jevu ω reálné číslo (převádí elementární jevy (abstraktními objekty) na čísla).
- Hodnota $X(\omega)$ náhodné veličiny X závisí na tom, který elementární jev ω nastal.
- Víme-li, který elementární jev ω nastal, známe hodnotu $X(\omega)$ náhodné veličiny X.



Zápis jevů pomocí náhodných veličin

Zápisem $(X = a)$ rozumíme jev složený ze všech elementárních jevů $\omega \in \Omega$, pro které $X(\omega) = a$,

$$\text{tj. } (X = a) = \{\omega \in \Omega : X(\omega) = a\}.$$

Podobně:

- $(X < a) = \{\omega \in \Omega : X(\omega) < a\}$,
- $(X > a) = \{\omega \in \Omega : X(\omega) > a\}$,
- $(a < X < b) = \{\omega \in \Omega : a < X(\omega) < b\}$, …

Náhodná veličina (NV)

Náhodná veličina – **číselné** vyjádření výsledku náhodného pokusu

$\boldsymbol{X}$ **… výsledek hodu kostkou**

Náhodný pokus: **Hod kostkou**

Náhodný jev: **Padne sudé číslo**. $(X \in \{2; 4; 6\})$

Náhodná veličina (NV)

Náhodná veličina – **číselné** vyjádření výsledku náhodného pokusu

X **… počet dívek mezi 1 000 náhodně vybranými dětmi**

Náhodný pokus: **Náhodný výběr 1 000 dětí a zjištění počtu dívek mezi nimi.**

Náhodný jev: **Mezi 1 000 náhodně vybranými dětmi bude více než 500 dívek**.

$(X > 500)$

Náhodná veličina (NV)

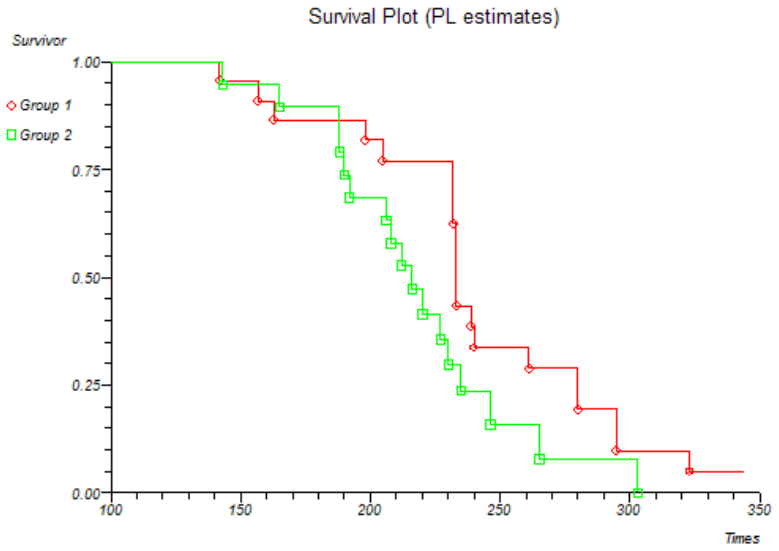
Náhodná veličina – **číselné** vyjádření výsledku náhodného pokusu

X **… doba do remise onemocnění (měsíc)**

Náhodný pokus: **Pozorování doby do remise onemocnění**

Náhodný jev: **Doba do remise nepřekročí 60 měsíců.**

$(X \leq 60)$



Náhodná veličina (NV)

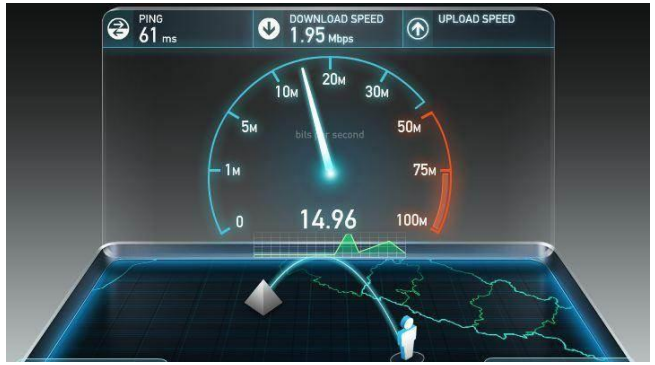
Náhodná veličina – **číselné** vyjádření výsledku náhodného pokusu

X **… rychlost připojení k internetu (Mb/s)**

Náhodný pokus: **Měření rychlosti připojení k internetu (download)**

Náhodný jev: **Rychlost připojení k internetu je vyšší než 20 Mb/s.**

$(X > 20)$



Náhodná veličina (NV)

Náhodná veličina – **číselné** vyjádření výsledku náhodného pokusu

POZOR!

Číselně vyjádření lze použít i u veličin, které nejsou kvantitativní ze své podstaty (např. pohlaví – muž (0), žena (1)).

Jak popsat chování náhodné veličiny?

Rozdělení pravděpodobnosti náhodné veličiny - předpis, který jednoznačně určuje všechny pravděpodobnosti typu

$$P(X \in M), \text{ kde } M \subset \mathbb{R}.$$

$$(\text{tj. } P(X = a), P(X < a), P(X > a), P(a < X < b), \ldots, \text{ kde } a, b \in \mathbb{R})$$

Jak zadat rozdělení pravděpodobnosti?

- funkci zadanou analyticky,
- výčtem hodnot NV a příslušných pravděpodobností,
- graficky.

Jak popsat chování náhodné veličiny?

Rozdělení pravděpodobnosti náhodné veličiny - předpis, který jednoznačně určuje všechny pravděpodobnosti typu

$$P(X \in M), \text{ kde } M \subset \mathbb{R}.$$

$$(\text{tj. } P(X = a), P(X < a), P(X > a), P(a < X < b), \ldots, \text{ kde } a, b \in \mathbb{R})$$

Jak zadat rozdělení pravděpodobnosti?

- distribuční funkcí,
- pravděpodobnostní funkcí (diskrétní NV),
- hustotou pravděpodobnosti (spojitá NV).

Distribuční funkce $F(x)$

Distribuční funkce $F(t)$ je pravděpodobnost, že náhodná veličina X bude **menší než** dané reálné číslo t.

$$\boxed{F(t) = P(X < t)}$$

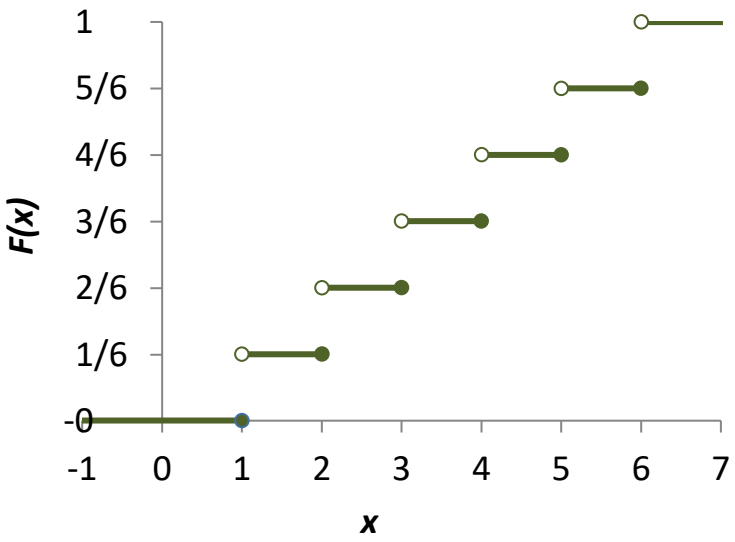
- $F(t)$ udává pravděpodobnost, že náhodná veličina nabývá hodnot menších než t.
- Distribuční funkce **jednoznačně určuje** rozdělení NV, tj. známe-li distribuční funkci, umíme určit pravděpodobnost $P(X \in M)$ pro libovolnou $M \subset \mathbb{R}$.
- Distribuční funkci náhodné veličiny X někdy značíme $F_X(t)$.

Někteří autoři definuji $F(t) = P(X \leq t)$!!!

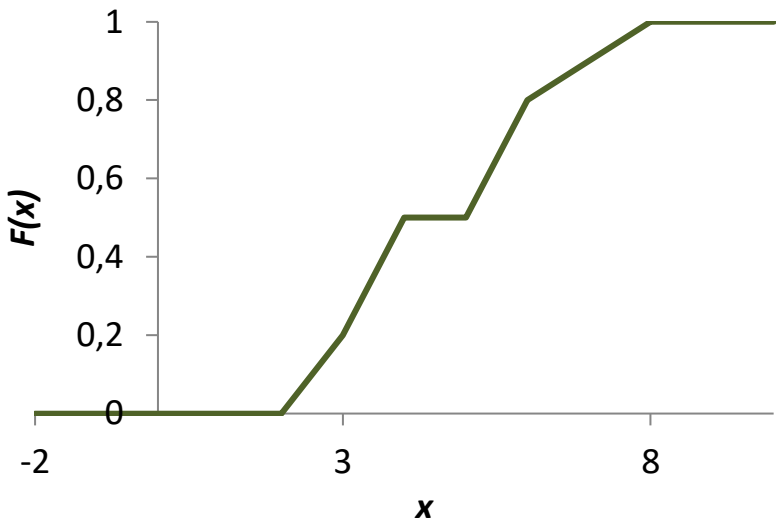
Distribuční funkce $F(x)$

Distribuční funkce $F(t)$ je pravděpodobnost, že náhodná veličina X bude **menší než** dané reálné číslo t.

$$F(t) = P(X < t)$$



Ukázka distribuční funkce diskrétní náhodné veličiny



Ukázka distribuční funkce spojité náhodné veličiny

Základní typy náhodných veličin

Diskrétní NV – mohou nabývat spočetně mnoha hodnot

Příklady: počet průjezdů automobilů Klimkovickým tunelem mezi 13:00 – 14:00, počet kybernetických útoků na určitý web během dne, počet dní nemocenské, počet zákazníků v lékárně během jednoho dne, …

Spojité NV – mohou nabývat všech hodnot na nějakém intervalu (mají spojitou distribuční funkci)

Příklady: doba do remise onemocnění, výška, váha, BMI, IQ, doba mezi dvěma následujícími kybernetickými útoky na web, chyba měření, …

Základní typy náhodných veličin

Diskrétní náhodná veličina (dále DNV)

- může nabývat spočetně mnoha hodnot
- DNV X s distribuční funkcí $F_X(t)$ je charakterizována pravděpodobnostní funkcí $P(X = x_i)$, tj. funkcí pro níž platí:

$$F_X(t) = \sum_{x_i<t} P(X = x_i) = \sum_{x_i<t} P(x_i)$$

Spojitá náhodná veličina (dále SNV)

- může nabývat všech hodnot na nějakém intervalu (má spojitou distribuční funkci)
- SNV X s distribuční funkcí $F_X(t)$ je charakterizována hustotou pravděpodobností $f(x)$, tj. funkcí pro níž platí:

$$F_X(t) = \int_{-\infty}^{t} f(x)\, dx$$

Diskrétní náhodná veličina

Definice

Náhodná veličina *X* má diskrétní rozdělení pravděpodobnosti (zkráceně: „je diskrétní“), právě když nabývá spočetně mnoha hodnot.

DNV lze charakterizovat: pravděpodobnostní funkcí, distribuční funkcí

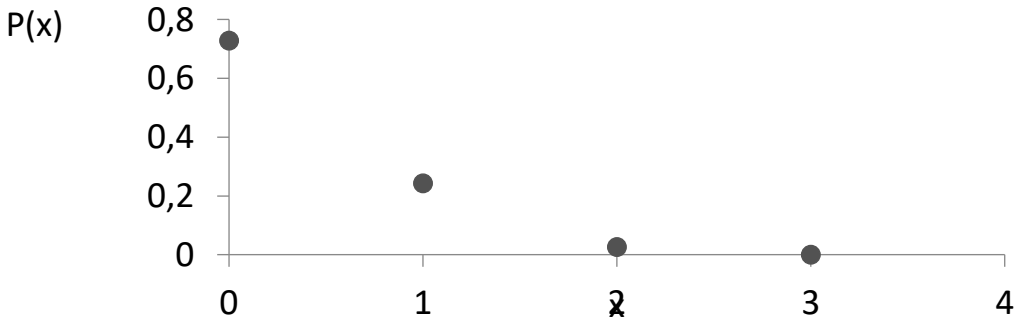
Pravděpodobnostní funkce

- $P(x_i) \geq 0$
- $\sum_i P(X = x_i) = 1$

- Lze zadat:
  - ✓ předpisem, $\forall x \in \{0; 1; 2; 3\}: P(X = x) = \binom{3}{x} \cdot 0{,}1^x \cdot 0{,}9^{3-x}$
  - ✓ tabulkou,

x	0	1	2	3
$P(x)$	0,729	0,243	0,027	0,001

  - ✓ grafem.



Příklad 1

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličinou X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

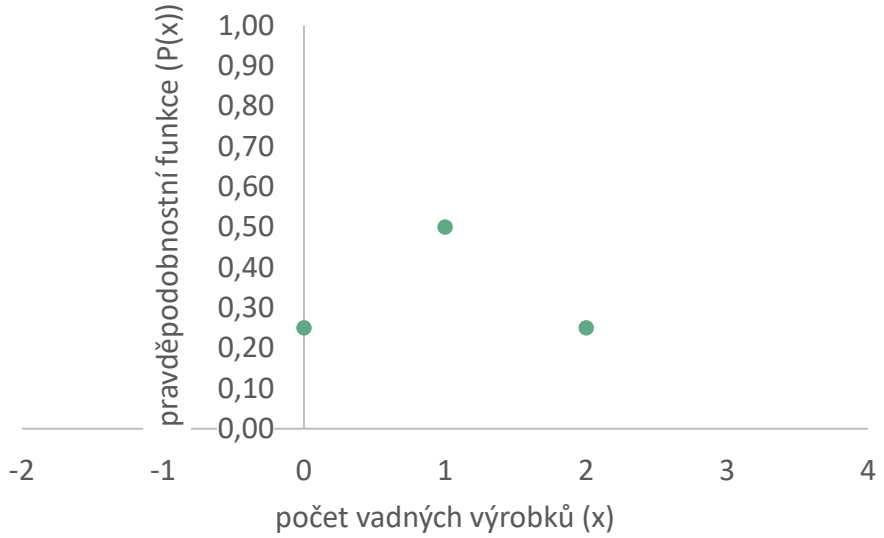
Počet vadných výrobků	Pravděpodobnost
0	0,25
1	0,50
2	0,25

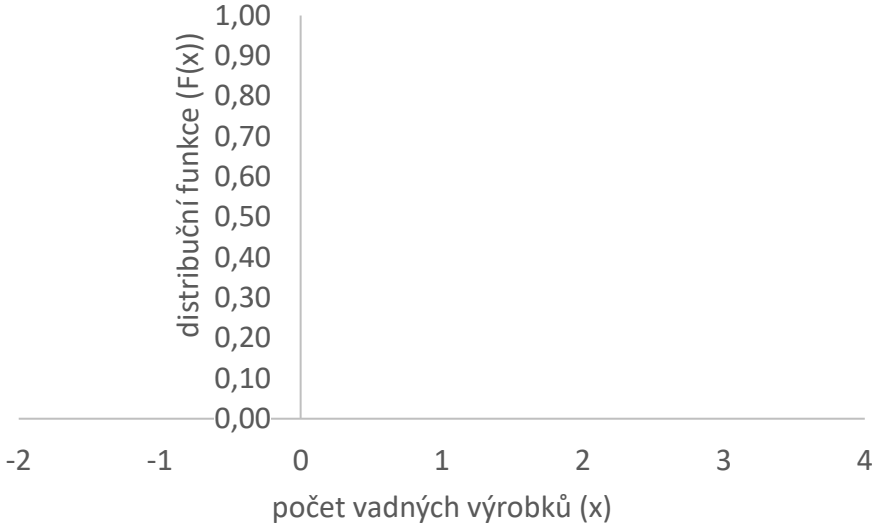
Určete pravděpodobnostní a distribuční funkci počtu vadných výrobků v testovaném vzorku.

Příklad 1

Počet vadných výrobků x	Pravděpodobnost $P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

Počet vadných výrobků x	Distribuční funkce $F(x) = P(X < x)$

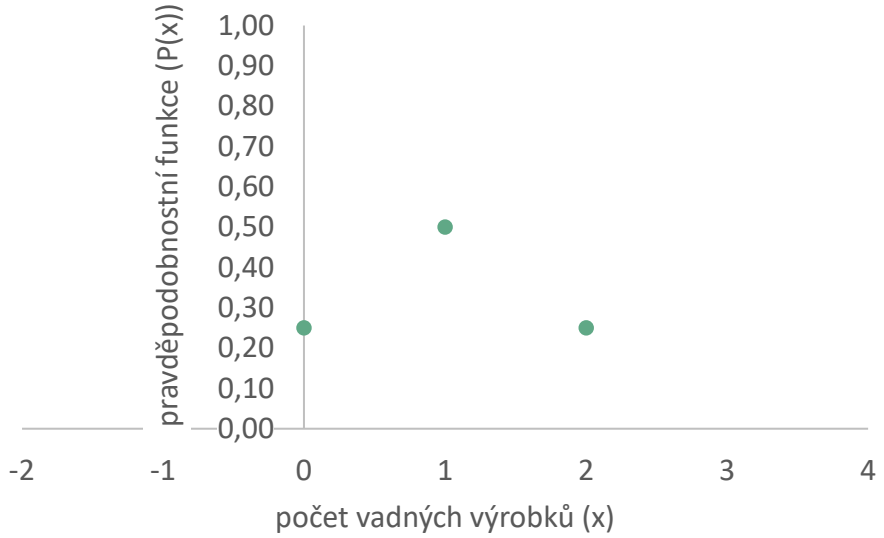


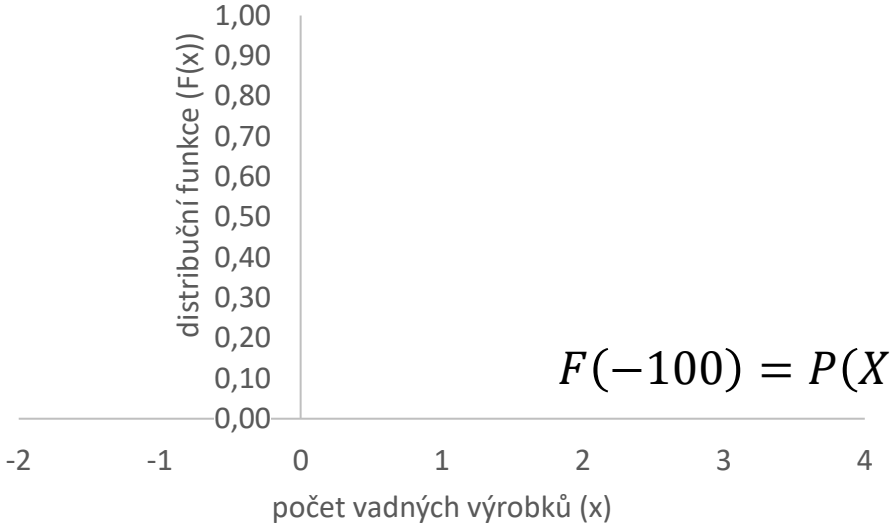


Příklad 1

Počet vadných výrobků x	Pravděpodobnost $P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

Počet vadných výrobků x	Distribuční funkce $F(x) = P(X < x)$



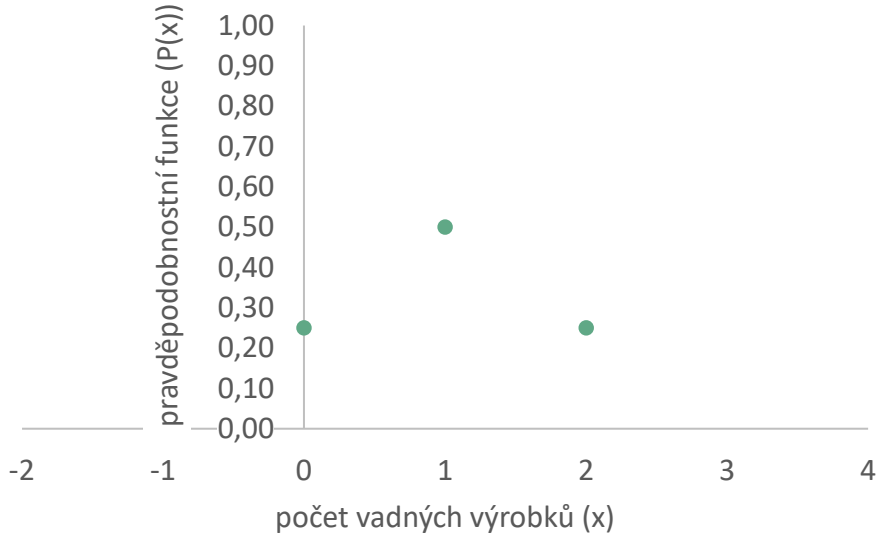


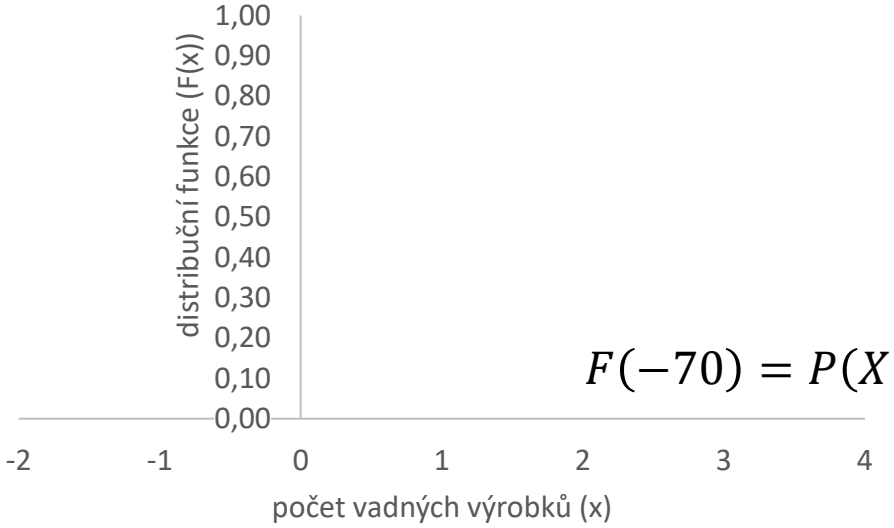
$$F(-100) = P(X < -100) = 0$$

Příklad 1

Počet vadných výrobků x	Pravděpodobnost $P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

Počet vadných výrobků x	Distribuční funkce $F(x) = P(X < x)$



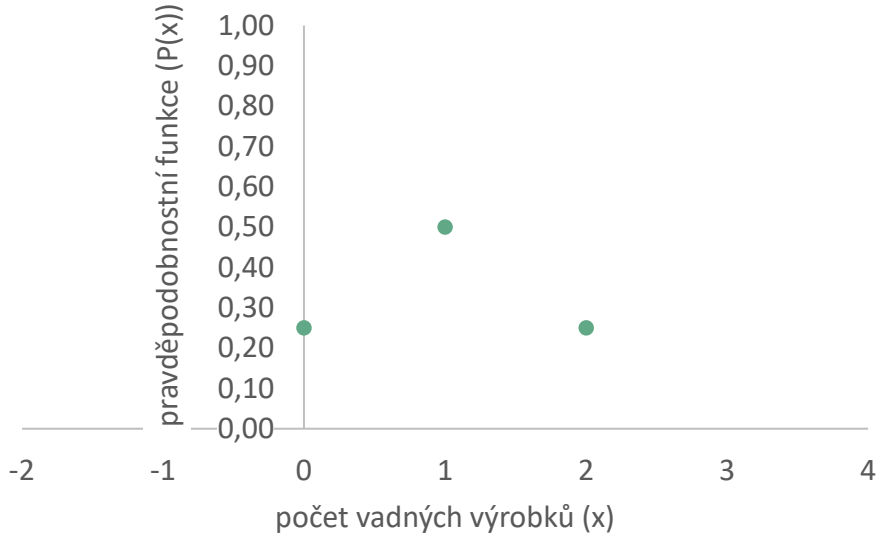


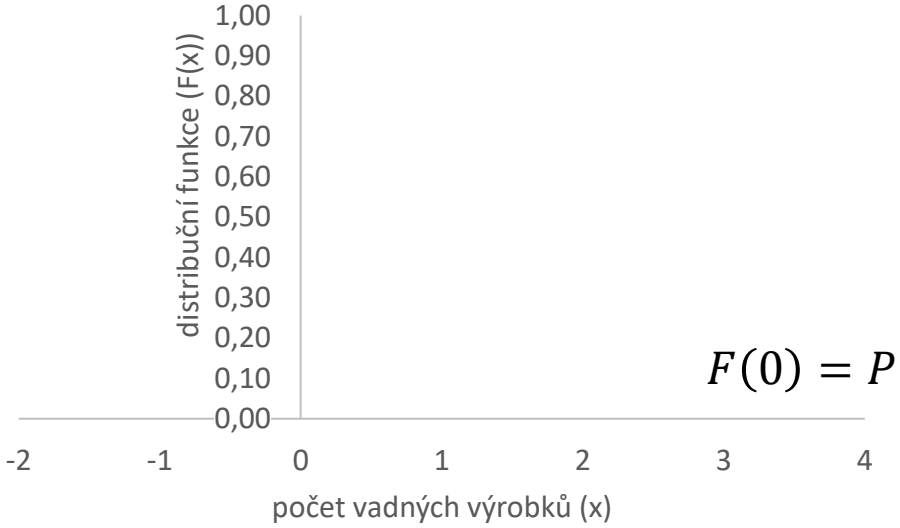
$$F(-70) = P(X < -70) = 0$$

Příklad 1

Počet vadných výrobků x	Pravděpodobnost $P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

Počet vadných výrobků x	Distribuční funkce $F(x) = P(X < x)$



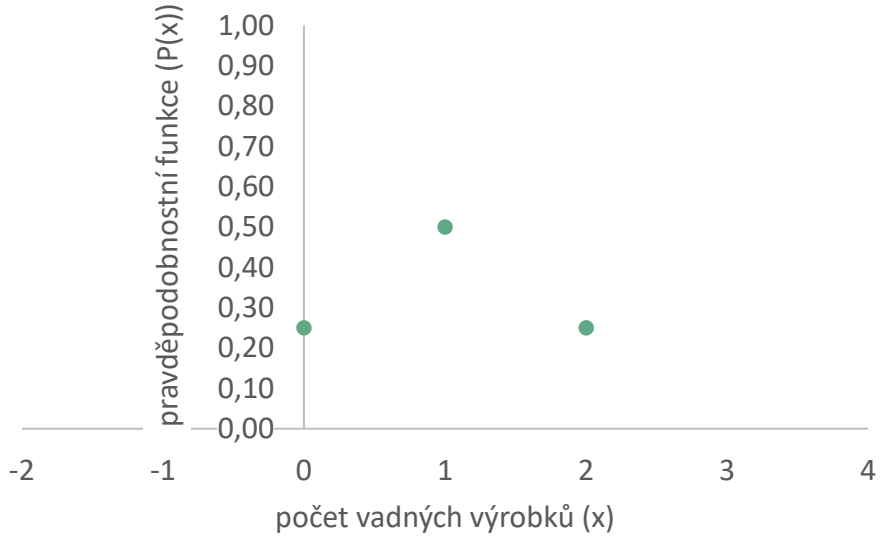


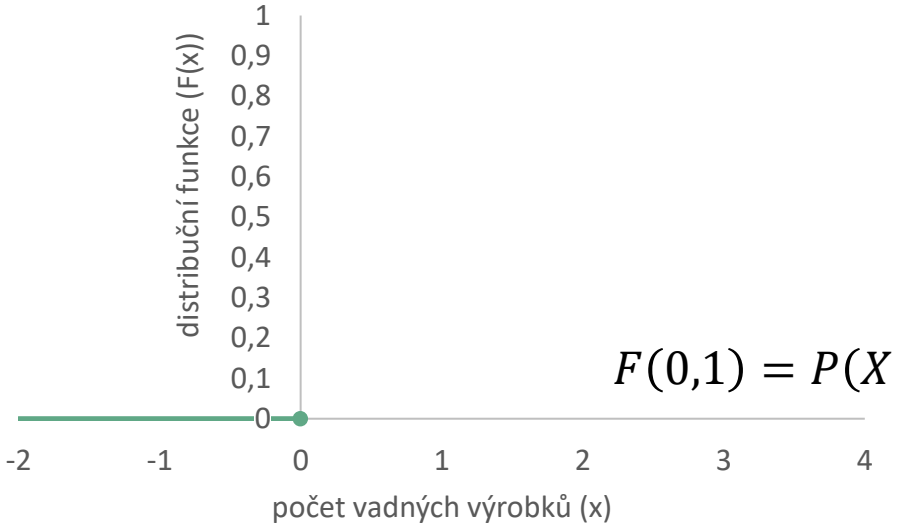
$$F(0) = P(X < 0) = 0$$

Příklad 1

Počet vadných výrobků x	Pravděpodobnost $P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

Počet vadných výrobků x	Distribuční funkce $F(x) = P(X < x)$
$(-\infty; 0\rangle$	0,00



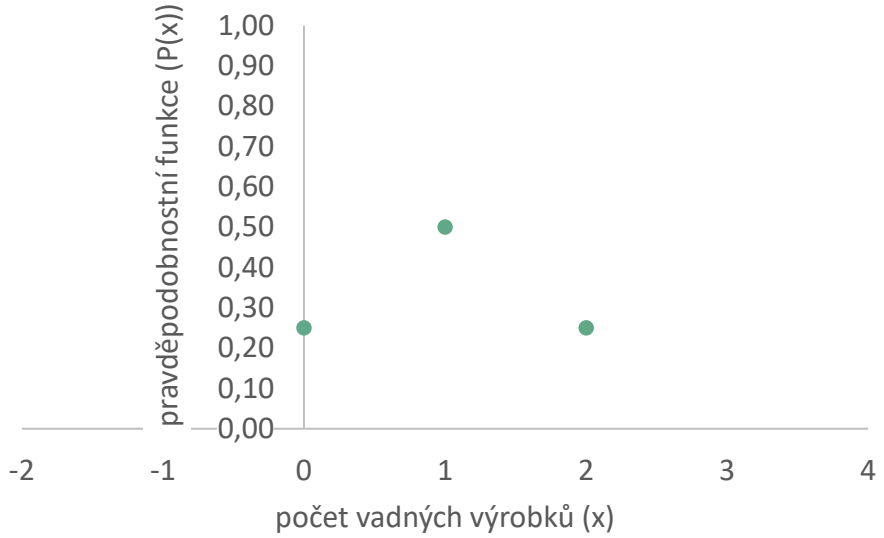


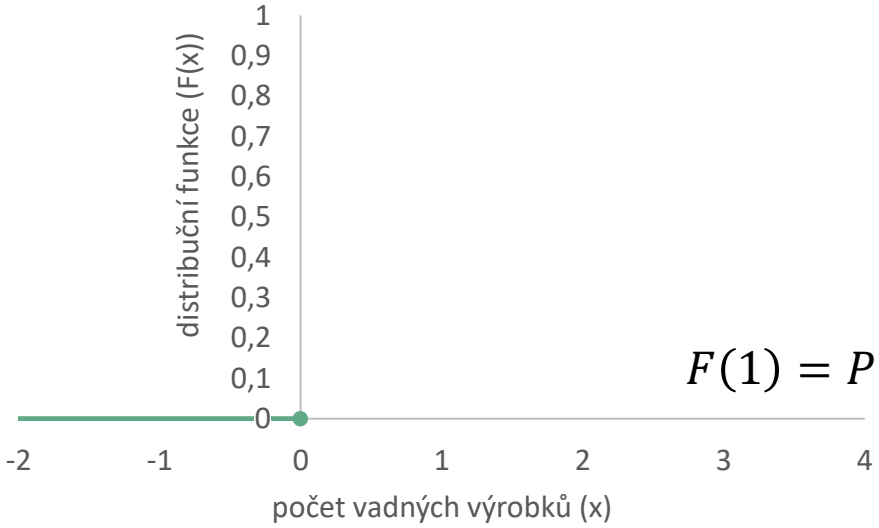
$$F(0{,}1) = P(X < 0{,}1) = 0{,}25$$

Příklad 1

Počet vadných výrobků x	Pravděpodobnost $P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

Počet vadných výrobků x	Distribuční funkce $F(x) = P(X < x)$
$(-\infty; 0\rangle$	0,00



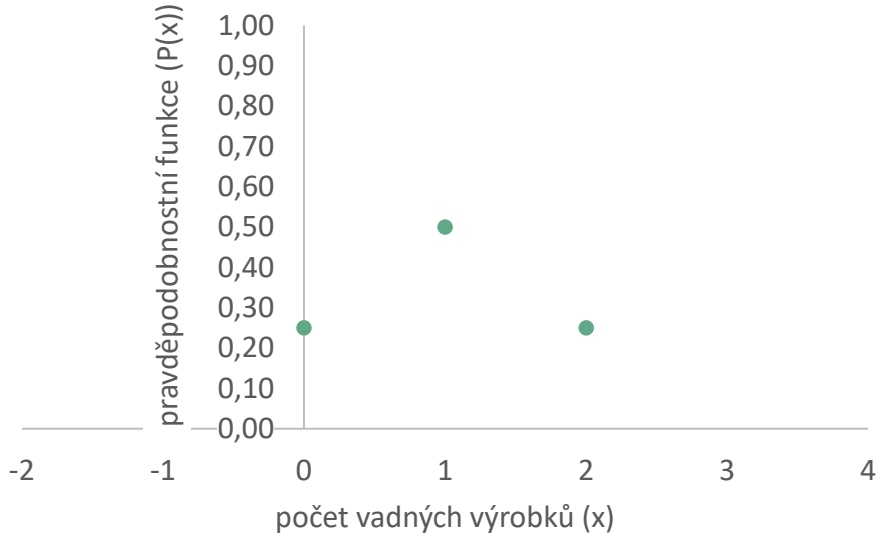


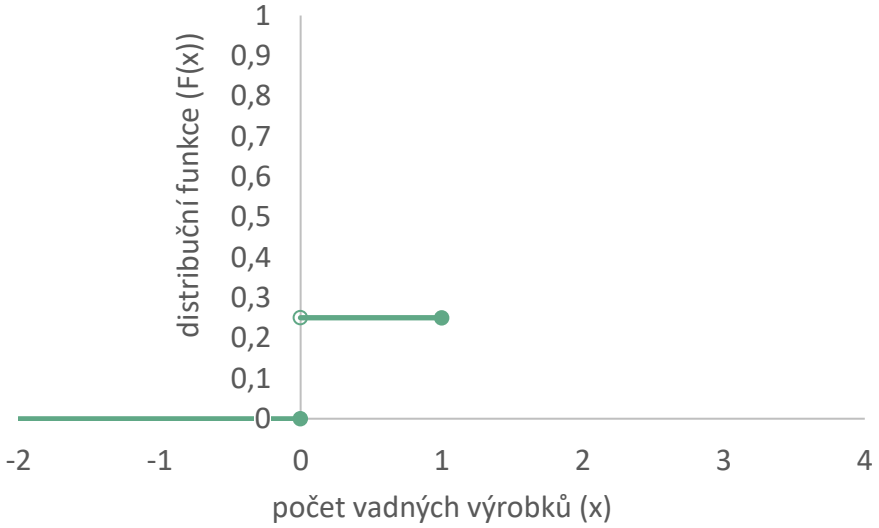
$$F(1) = P(X < 1) = 0{,}25$$

Příklad 1

Počet vadných výrobků x	Pravděpodobnost $P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

Počet vadných výrobků x	Distribuční funkce $F(x) = P(X < x)$
$(-\infty; 0\rangle$	0,00
$(0; 1\rangle$	0,25

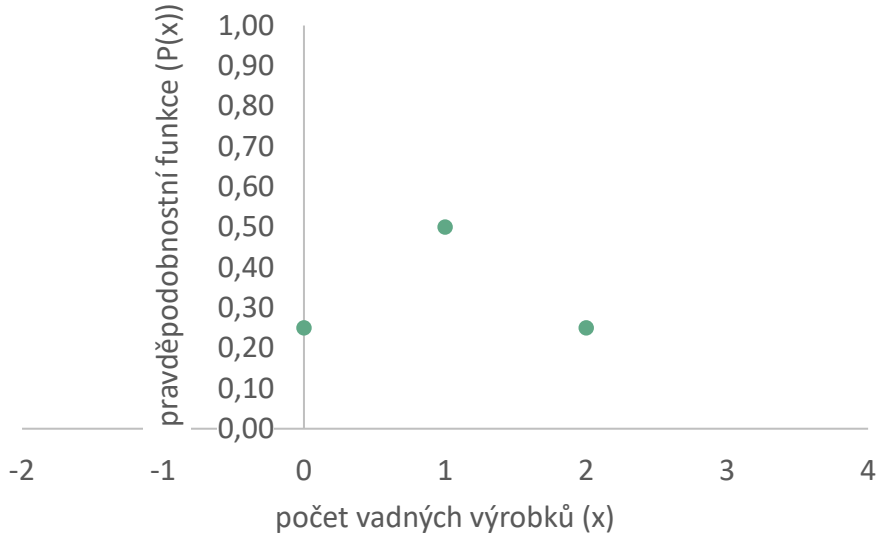


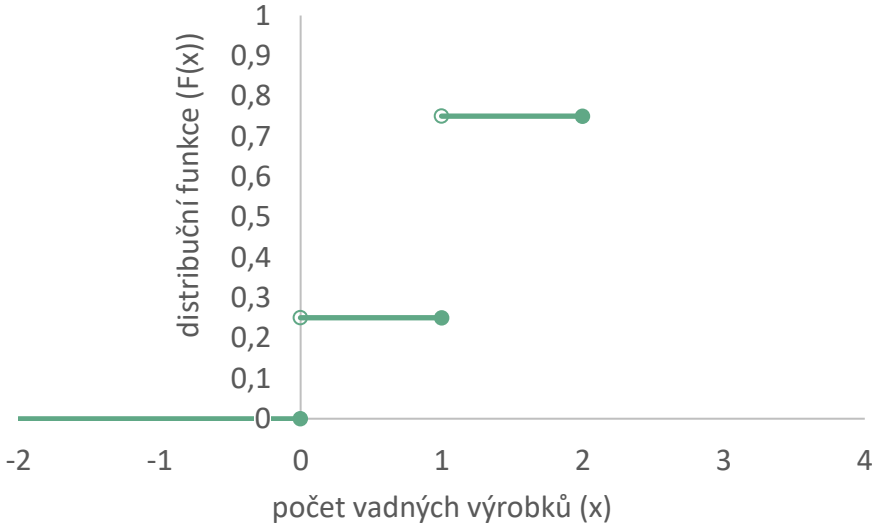


Příklad 1

Počet vadných výrobků x	Pravděpodobnost $P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

Počet vadných výrobků x	Distribuční funkce $F(x) = P(X < x)$
$(-\infty; 0\rangle$	0,00
$(0; 1\rangle$	0,25
$(1; 2\rangle$	0,75

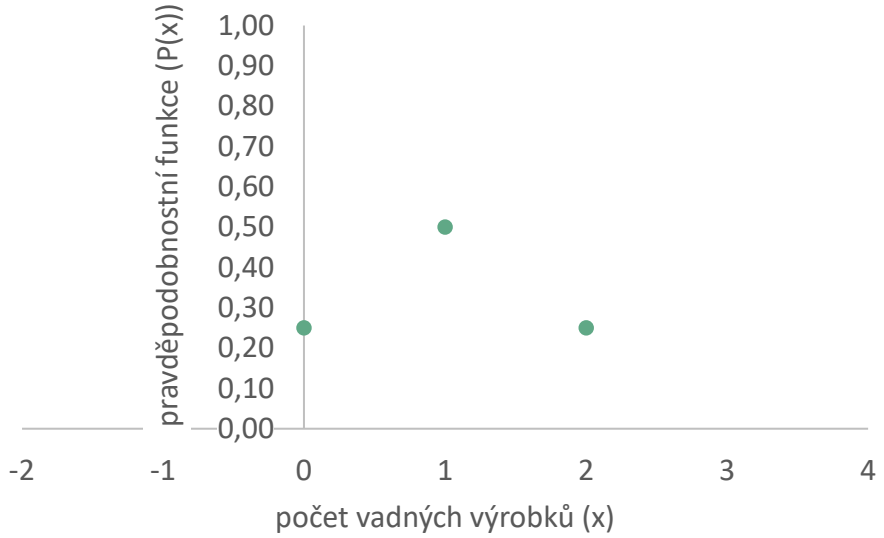


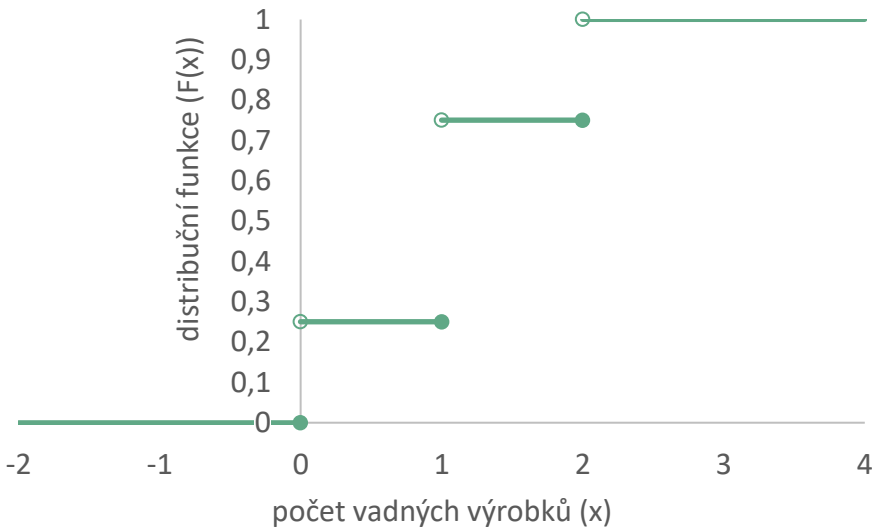


Příklad 1

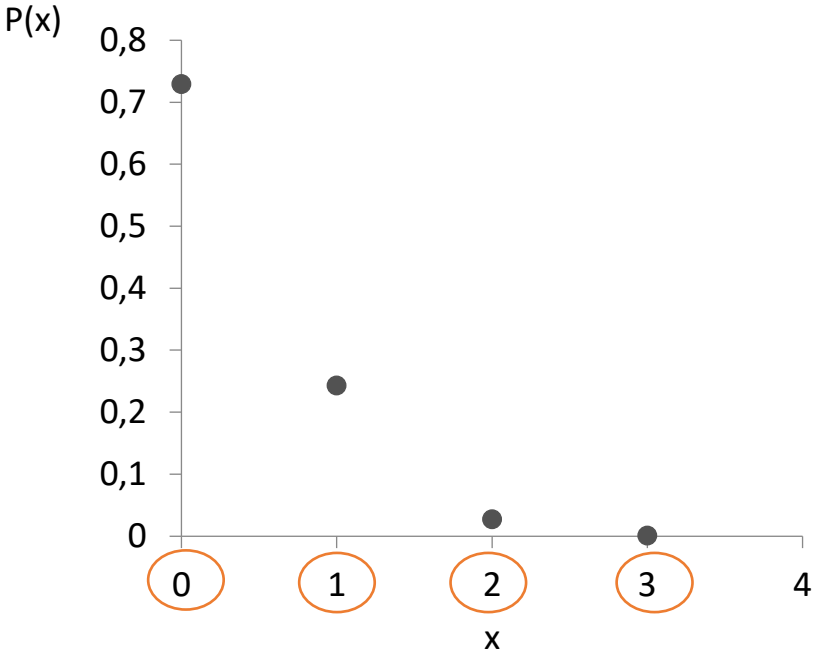
Počet vadných výrobků x	Pravděpodobnost $P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

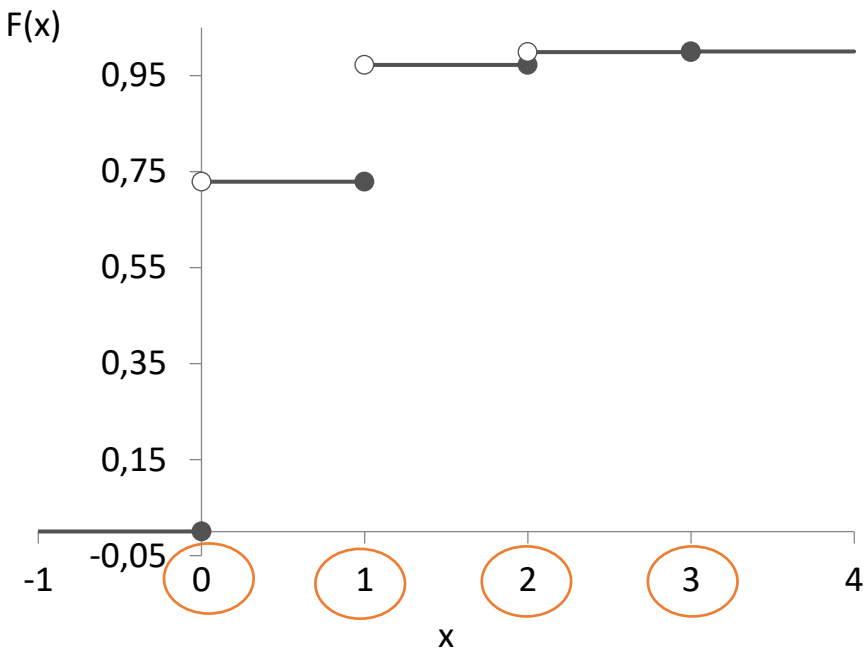
Počet vadných výrobků x	Distribuční funkce $F(x) = P(X < x)$
$(-\infty; 0\rangle$	0,00
$(0; 1\rangle$	0,25
$(1; 2\rangle$	0,75
$(2; \infty)$	1,00





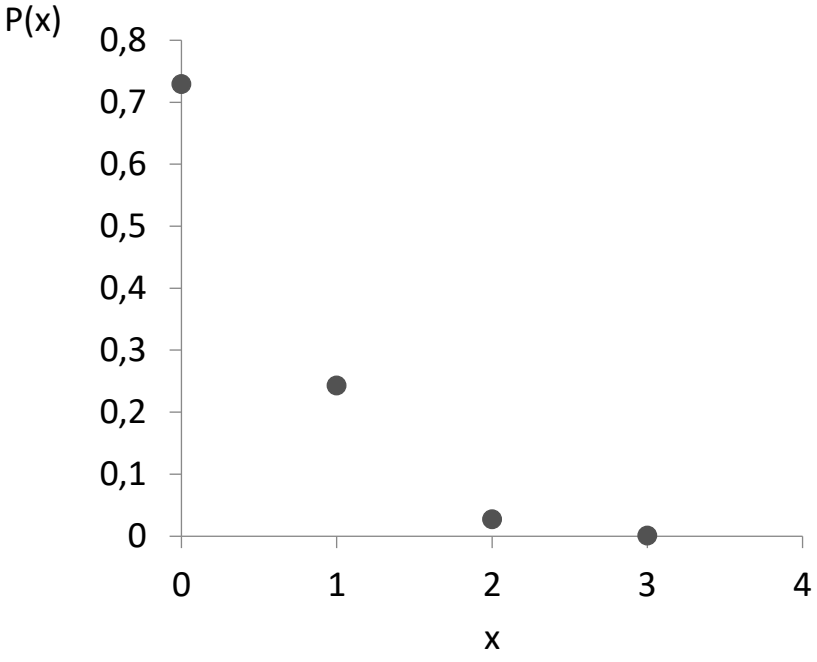
Vztah mezi pravděpodobnostní a distribuční funkci

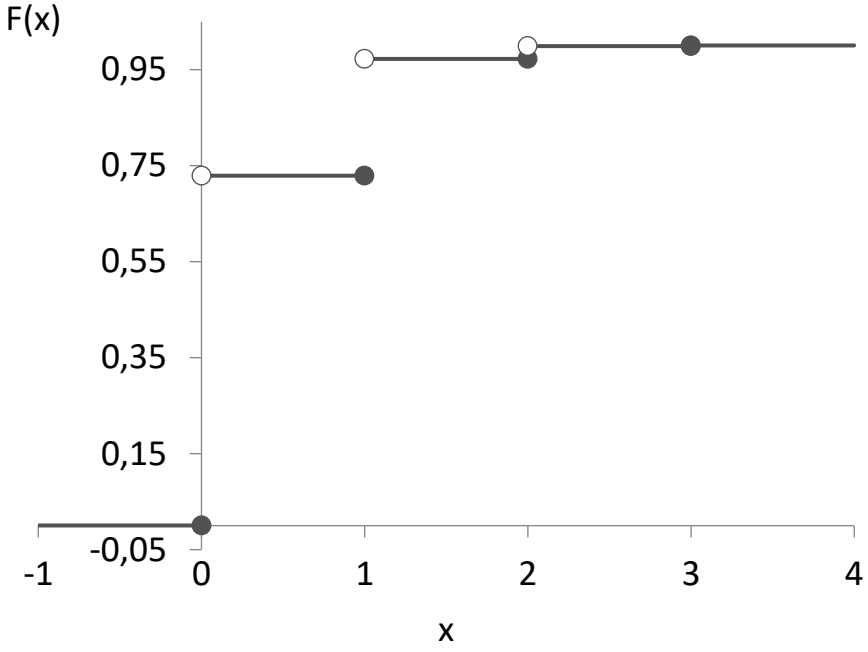




- Body nespojitosti distribuční f-ce jsou body, v nichž je pravděpodobnostní f-ce nenulová.

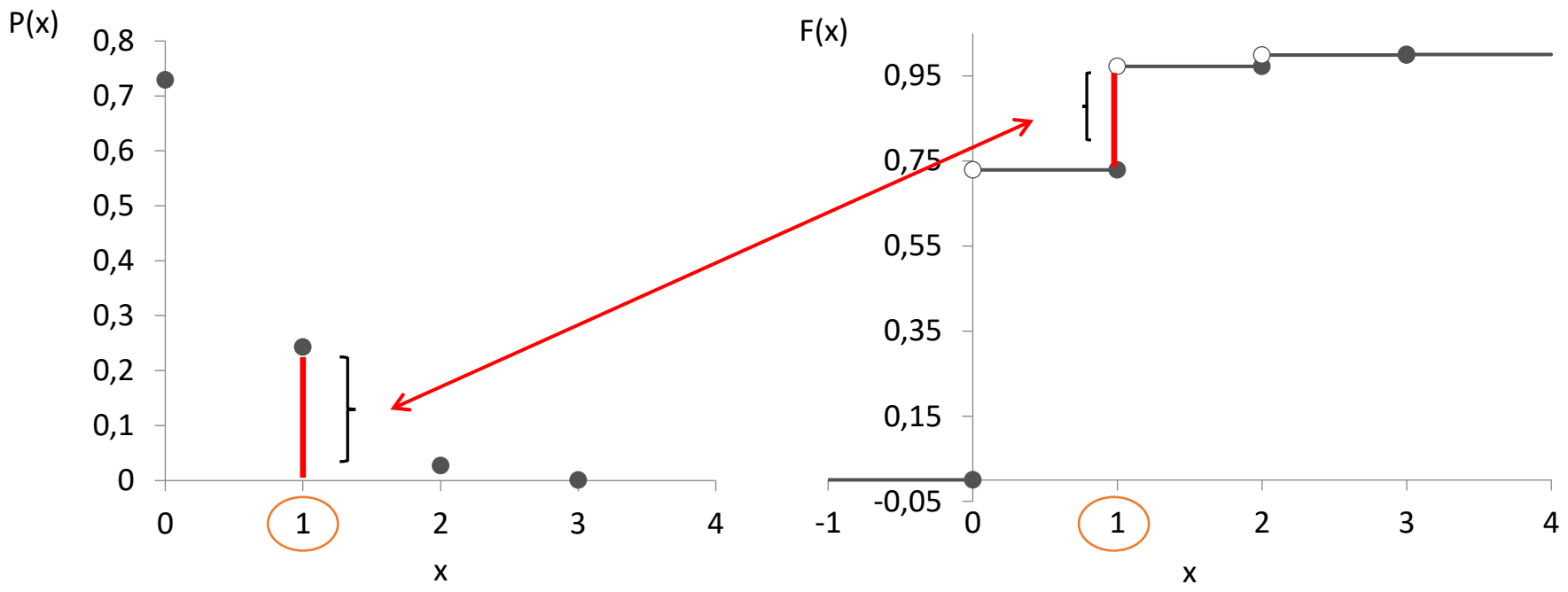
Vztah mezi pravděpodobnostní a distribuční funkci





- Body nespojitosti distribuční f-ce jsou body, v nichž je pravděpodobnostní f-ce nenulová.
- $P(X = a) = \lim_{x \to a+} F(x) - F(a)$

Vztah mezi pravděpodobnostní a distribuční funkci



- Body nespojitosti distribuční f-ce jsou body, v nichž je pravděpodobnostní f-ce nenulová.
- $P(X = a) = \lim_{x \to a+} F(x) - F(a)$, tj. velikost „skoku“ distribuční funkce v bodech nespojitosti je rovna příslušným hodnotám pravděpodobnostní funkce.

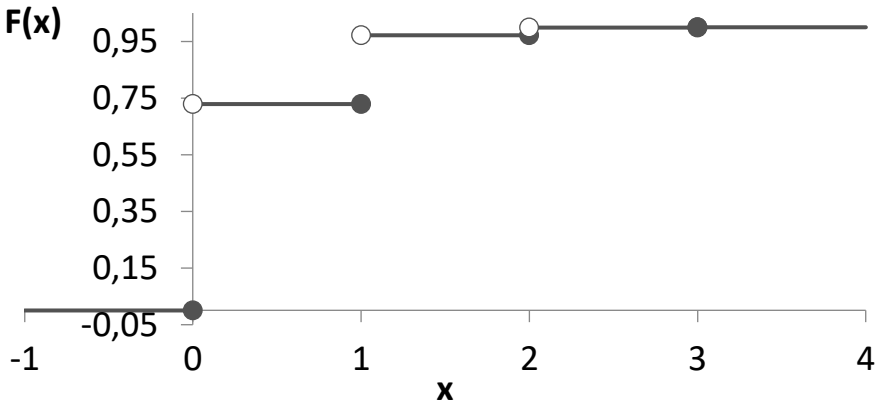
Distribuční funkce

- $F(t) = \sum_{x_i < t} P(x_i)$

- Lze zadat:
  - ✓ předpisem,
  - ✓ tabulkou,
  - ✓ grafem.

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 0{,}729 & 0 < x \leq 1 \\ 0{,}972 & 1 < x \leq 2 \\ 0{,}999 & 2 < x \leq 3 \\ 1 & 3 < x \leq \infty \end{cases}$$

x	$F(x)$
$(-\infty; 0\rangle$	0
$(0; 1\rangle$	0,729
$(1; 2\rangle$	0,972
$(2; 3\rangle$	0,999
$(3; \infty)$	1



Distribuční funkce $F(x)$

Distribuční funkce $F(t)$ je pravděpodobnost, že náhodná veličina X bude **menší než** dané reálné číslo t.

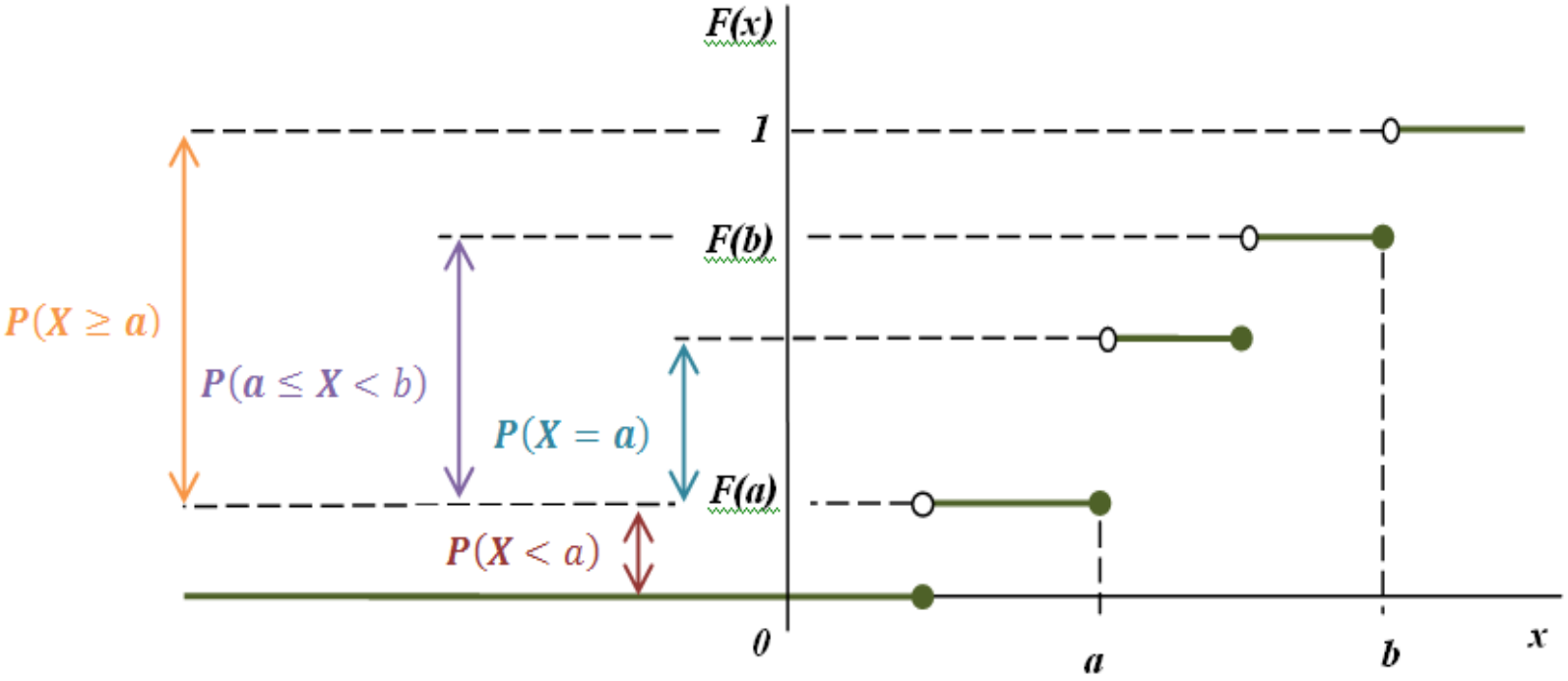
$$\boxed{F(t) = P(X < t)}$$

Vlastnosti distribuční funkce:

- $0 \leq F(t) \leq 1$,
- je neklesající,
- je zleva spojitá,
- má nejvýše spočetně mnoho bodů nespojitosti,
- $\lim\limits_{t \to -\infty} F(t) = 0$ („začíná“ v 0),
- $\lim\limits_{t \to \infty} F(t) = 1$ („končí“ v 1).

Někteří autoři definuji $F(t) = P(X \leq t)$!!!

Vztah mezi pravděpodobnostní a distribuční funkci



$$P(X = a) = \lim_{x \to a+} F(x) - F(a)$$

Vztah mezi pravděpodobnostní a distribuční funkci

- $P(X < a) = F(a) = \sum_{x_i < a} P(x_i)$
- $P(X \geq a) = 1 - P(X < a) = 1 - F(a) = \sum_{x_i \geq a} P(x_i)$
- $P(a \leq X < b) = P(X < b) - P(X < a) = F(a) - F(b) = \sum_{a \leq x_i < b} P(x_i)$
- $P(X = a) = \lim_{x \to a+} F(x) - F(a)$

Příklad 2

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X, která má níže uvedenou distribuční funkci.

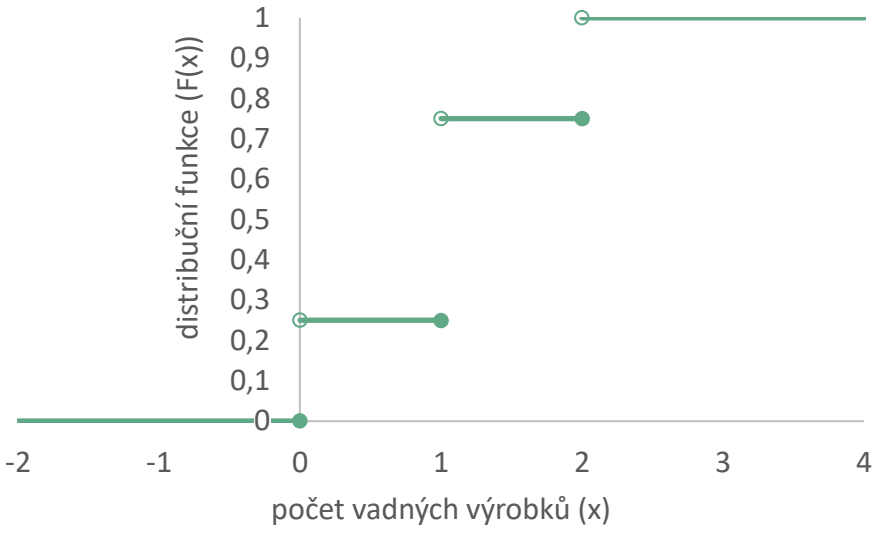
$$F(x) = \begin{cases} 0{,}00, x \in (-\infty; 0\rangle \\ 0{,}25, x \in (0; 1\rangle \\ 0{,}75, x \in (1; 2\rangle \\ 1{,}00, x \in (2; \infty). \end{cases}$$

Určete pravděpodobnostní funkci počtu vadných výrobků v testovaném vzorku.

Příklad 2

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X, která má níže uvedenou distribuční funkci.

$$F(x) = \begin{cases} 0{,}00, x \in (-\infty; 0\rangle \\ 0{,}25, x \in (0; 1\rangle \\ 0{,}75, x \in (1; 2\rangle \\ 1{,}00, x \in (2; \infty). \end{cases}$$



x	$P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

Určete pravděpodobnostní funkci počtu vadných výrobků v testovaném vzorku.

K čemu potřebujeme popisovat chování NV?

- NV převádějí abstraktní základní prostor Ω na čísla (s čísly se lépe pracuje).

- NV slouží jako model pro naše empirická pozorování (data).

  - V teorii pravděpodobnosti s nimi pracujeme teoreticky – jejich rozdělení považujeme za dané a zkoumáme jejich vlastnosti.

  - Ve statistice se snažíme cosi usoudit o jejich neznámém rozdělení na základě konkrétních realizací.

Proč potřebujeme číselné charakteristiky NV?

- Distribuční funkce, resp. pravděpodobnostní funkce, popisují rozdělení NV jednoznačně, do všech podrobností.

- Někdy nás zajímá pouze některý aspekt NV, který se dá popsat jedním číslem:
  - očekávaná hodnota NV,
  - variabilita možných hodnot,
  - ...

Číselné charakteristiky NV

- Obecný moment r-tého řádu (značí se μ_r nebo $E(X^r)$ pro $r = 1, 2, \ldots$)
  - pro diskrétní NV: $\mu_r = \sum_{(i)} x_i^r \cdot P(x_i)$
  - Střední hodnota (angl. expected value, mean; značí se $E(X)$nebo μ)
    - pro diskrétní NV: $\mathrm{E}(X) = \mu = \sum_{(i)} x_i \cdot P(x_i)$

Význam střední hodnoty

Střední hodnotu $E(X)$ náhodné veličiny X lze chápat jako:

- průměrnou (očekávanou) hodnotu NV X, kolem níž hodnoty NV kolísají,
- míru polohy, populační průměr,
- vážený průměr všech možných hodnot $\left(\mathrm{E}(X) = \sum_{(i)} x_i \cdot P(x_i)\right)$.

Příklad 3

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

Počet vadných výrobků	Pravděpodobnost
0	0,25
1	0,50
2	0,25

Určete střední hodnotu počtu vadných výrobků v testovaném vzorku.

Příklad 3

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

$$\mathrm{E}(X) = \mu = \sum_{(i)} x_i \cdot P(x_i)$$

x	$P(x)$	$x \cdot P(x)$
0	0,25	
1	0,50	
2	0,25	
Celkem	**1,00**	

Určete střední hodnotu počtu vadných výrobků v testovaném vzorku.

Příklad 3

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

$$\mathrm{E}(X) = \mu = \sum_{(i)} x_i \cdot P(x_i)$$

x	$P(x)$	$x \cdot P(x)$
0	0,25	$0 \cdot 0{,}25$
1	0,50	$1 \cdot 0{,}50$
2	0,25	$2 \cdot 0{,}25$
Celkem	**1,00**	

Určete střední hodnotu počtu vadných výrobků v testovaném vzorku.

Příklad 3

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

$$\mathrm{E}(X) = \mu = \sum_{(i)} x_i \cdot P(x_i)$$

x	$P(x)$	$x \cdot P(x)$
0	0,25	0,00
1	0,50	0,50
2	0,25	0,50
Celkem	**1,00**	

Určete střední hodnotu počtu vadných výrobků v testovaném vzorku.

Příklad 3

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

$$\mathrm{E}(X) = \mu = \sum_{(i)} x_i \cdot P(x_i) = 1{,}00$$

x	$P(x)$	$x \cdot P(x)$
0	0,25	0,00
1	0,50	0,50
2	0,25	0,50
Celkem	**1,00**	**1,00**

Mezi vybranými výrobky lze očekávat průměrně 1 vadný výrobek.

Vlastnosti střední hodnoty

- $\forall a, b \in \mathbb{R}$: $E(aX + b) = aE(X) + b$,

- $E(\sum_i^n X_i) = \sum_i^n E(X_i)$,

  tj. střední hodnota součtu NV je rovna součtu jejich středních hodnot (vždy),

- $X_1, \cdots, X_n$ nezávislé $\Rightarrow$ $E(\prod_{i=1}^n X_i) = \prod_{i=1}^n E(X_i)$,
  tj. jsou-li NV nezávislé, pak střední hodnota součinu NV je rovna součinu jejich středních hodnot (obecně to neplatí).

Příklad 4

Nezávislé náhodné veličiny X a Y mají následující střední hodnotu:

$$E(X) = 1, E(Y) = 3.$$

Definujme náhodné veličiny Z a Q jako

$$Z = 4X - 2Y + 12,$$

$$Q = -2XY + Y - 7.$$

Určete střední hodnotu náhodných veličin Z a Q.

$$E(Z) = E(4X - 2Y + 12) = E(4X) + E(-2Y) + E(12) = 4E(X) - 2E(Y) + 12 = 4 \cdot 1 - 2 \cdot 3 + 12 = 10$$

$$E(Q) = E(-2XY + Y - 7) = -2\underbrace{E(X) \cdot E(Y)}_{} + E(Y) - E(7) = -2 \cdot 1 \cdot 3 + 3 - 7 = -10$$

toto je možné pouze proto, že X, Y jsou nezávislé NV

Modus

Modus $\widehat{X}$ – typická hodnota náhodné veličiny

pro diskrétní NV: $\forall x_i$: $P(X = \hat{x}) \geq P(X = x_i)$

(tzn. modus je taková hodnota DNV, v níž $P(x_i)$ nabývá svého maxima)

- Modus není těmito podmínkami určen jednoznačně, tzn. náhodná veličina může mít několik modů (např. výsledek hodu kostkou).
- Má-li NV právě jeden modus, mluvíme o unimodálním rozdělení NV.
- Má-li NV unimodální symetrické rozdělení, pak $E(X) = \widehat{X}$.

Příklad 5

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

Počet vadných výrobků	Pravděpodobnost
0	0,25
1	0,50
2	0,25

Určete modus počtu vadných výrobků v testovaném vzorku.

Příklad 5

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

Počet vadných výrobků	Pravděpodobnost
0	0,25
1	0,50
2	0,25

$\hat{X} = 1$

Nejčastěji lze mezi vybranými výrobky očekávat 1 vadný výrobek.

Číselné charakteristiky NV

- Centrální moment r-tého řádu (značí se μ'_r nebo $E\left(\left(X - E(X)\right)^r\right)$ pro $r = 1, 2, \ldots$)
  - pro diskrétní NV: $\mu'_r = \sum_{(i)}\left(x_i - E(X)\right)^r \cdot P(x_i)$
  - Rozptyl (angl. dispersion, resp. variance, značí se $\mu_2{}'$ nebo $D(X)$ nebo σ^2)
    - pro diskrétní NV: $D(X) = \sigma^2 = \sum_{(i)}(x_i - E(X))^2 \cdot P(x_i)$

Výpočetní vztah pro rozptyl:

$$\boxed{D(X) = E(X^2) - \left(E(X)\right)^2}$$

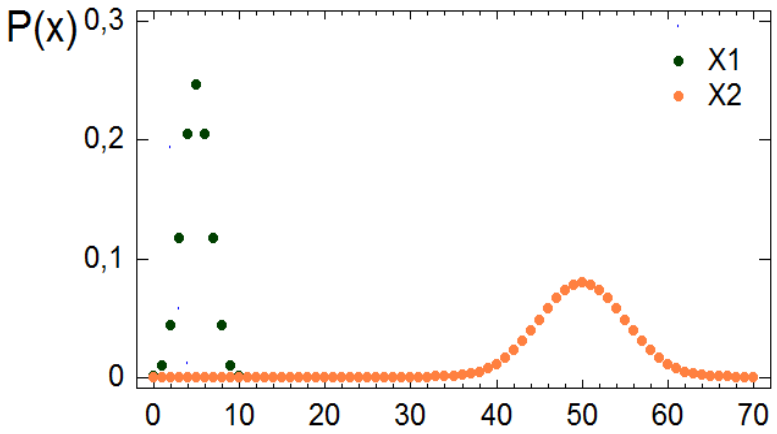
Důkaz:

$$D(X) = E\left((X - E(X))^2\right) = E\left(X^2 - 2X \cdot E(X) + \left(E(X)\right)^2\right) =$$

$$= E\left(X^2\right) - 2E(X) \cdot E(X) + \left(E(X)\right)^2 = E(X^2) - \left(E(X)\right)^2$$

Význam rozptylu

- Míra variability dat kolem střední hodnoty.
- Střední kvadratická odchylka od střední hodnoty $\left(D(X) = E(X - E(X))^2\right)$.
- Malý rozptyl ≈ hodnoty NV se s vysokou pravděpodobností objevují blízko $E(X)$.
- Velký rozptyl ≈ hodnoty NV se často objevují ve velké vzdálenosti od $E(X)$.



$$\boldsymbol{D(X1) < D(X2)}$$

Jednotka rozptylu je kvadrátem jednotky náhodné veličiny!!!

Příklad 6

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

Počet vadných výrobků	Pravděpodobnost
0	0,25
1	0,50
2	0,25

Určete rozptyl počtu vadných výrobků v testovaném vzorku.

Příklad 6

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

$$\mathrm{D}(X) = E(X^2) - \big(E(X)\big)^2$$

x	$P(x)$	$x \cdot P(x)$
0	0,25	0,00
1	0,50	0,50
2	0,25	0,50
Celkem	**1,00**	**1,00**

Určete rozptyl počtu vadných výrobků v testovaném vzorku.

Příklad 6

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

$$\mathrm{D}(X) = E(X^2) - \big(E(X)\big)^2$$

x	$P(x)$	$x \cdot P(x)$	$x^2 \cdot P(x)$
0	0,25	0,00	
1	0,50	0,50	
2	0,25	0,50	
Celkem	**1,00**	**1,00**	

Určete rozptyl počtu vadných výrobků v testovaném vzorku.

Příklad 6

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

$$\mathrm{D}(X) = E(X^2) - \big(E(X)\big)^2$$

x	$P(x)$	$x \cdot P(x)$	$x^2 \cdot P(x)$
0	0,25	0,00	$0^2 \cdot 0{,}25$
1	0,50	0,50	$1^2 \cdot 0{,}50$
2	0,25	0,50	$2^2 \cdot 0{,}25$
Celkem	**1,00**	**1,00**	

Určete rozptyl počtu vadných výrobků v testovaném vzorku.

Příklad 6

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

$$\mathrm{D}(X) = E(X^2) - \left(E(X)\right)^2 = 1{,}50 - 1{,}00^2 = 0{,}50$$

x	$P(x)$	$x \cdot P(x)$	$x^2 \cdot P(x)$
0	0,25	0,00	0,00
1	0,50	0,50	0,50
2	0,25	0,50	1,00
Celkem	**1,00**	**1,00**	**1,50**

Rozptyl počtu vadných výrobků v testovaném vzorku je 0,5 ks^2.

Vlastnosti rozptylu

- $\forall a, b \in \mathbb{R}$: $\quad D(aX + b) = a^2 D(X)$,

Důkaz:

$$D(aX+b) = E\big((aX+b)^2\big) - \Big(E\big((aX+b)\big)\Big)^2 =$$

$$= a^2 E\big(X^2\big) + 2abE(X) + b^2 - \Big(a^2\big(E(X)\big)^2 + 2abE(X) + b^2\Big) =$$

$$= a^2\left(E\big(X^2\big) - \big(E(X)\big)^2\right) = a^2 D(X)$$

- $X_1, \cdots, X_n$ nezávislé $\Rightarrow D(\sum_i^n X_i) = \sum_i^n D(X_i)$,

  tj. jsou-li NV nezávislé, pak rozptyl součtu NV je roven součtu jejich rozptylů (obecně to neplatí).

Příklad 7

Nezávislé náhodné veličiny X a Y mají následující rozptyly:

$$D(X) = 1, D(Y) = 2.$$

Definujme náhodné veličiny Z a Q jako

$$Z = 4X - 2Y + 12,$$

$$Q = -2XY + Y - 7.$$

Určete rozptyly náhodných veličin Z a Q.

$$D(Z) = D(4X - 2Y + 12) = \underbrace{D(4X) + D(-2Y)}_{\text{toto je možné pouze proto, že X, Y jsou nezávislé NV}} + D(12) = 4^2 D(X) + (-2)^2 D(Y) = 16 \cdot 1 + 4 \cdot 2 = 24$$

$$D(Q) = D(-2XY + Y - 7) = 4D\underbrace{(XY) + D(Y)}_{\text{pro výpočet nemáme dostatek informací}}$$

Směrodatná odchylka

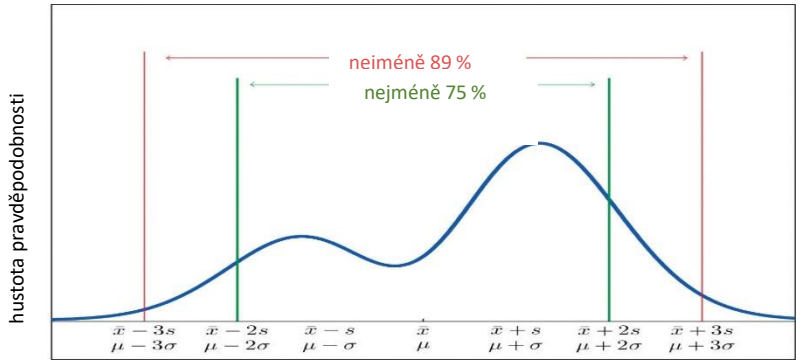
Směrodatná odchylka σ

$$\boxed{\sigma(X) = \sqrt{D(X)}}$$

Jakou představu o náhodné veličině X si lze udělat
na základě její stř. hodnoty μ a sm. odchylky σ?

$$\forall k > 0: P(\mu - k\sigma < X < \mu + k\sigma) > 1 - \frac{1}{k^2}$$ (Čebyševova nerovnost)

$\boldsymbol{k}$	$P(\mu - k\sigma < X < \mu + k\sigma)$
1	>0
2	>0,75
3	>0,89



Směrodatná odchylka neumožňuje srovnávat variabilitu náhodných veličin měřených v různých jednotkách!

Variační koeficient

Variační koeficient $\gamma(X)$ definujme pouze pro nezáporné náhodné veličiny.

$$\boxed{\gamma(X) = \frac{\sigma(X)}{E(X)}, \text{ resp. } \frac{\sigma(X)}{E(X)} \cdot 100\ (\%)}$$

- Variační koeficient – směrodatná odchylka v procentech střední hodnoty
- Čím nižší variační koeficient, tím homogennější soubor.
- $\gamma(X) > 50\%$ $\approx$ silně rozptýlený soubor.

Příklad 8

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

Počet vadných výrobků	Pravděpodobnost
0	0,25
1	0,50
2	0,25

Určete směrodatnou odchylku a variační koeficient počtu vadných výrobků v testovaném vzorku.

Příklad 8

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

$$\mathrm{E}(X) = \sum_{(i)} x_i \cdot P(x_i) = 1{,}00, \quad \mathrm{D}(X) = E(X^2) - \left(E(X)\right)^2 = 1{,}50 - 1{,}00^2 = 0{,}50$$

x	$P(x)$	$x \cdot P(x)$	$x^2 \cdot P(x)$
0	0,25	0,00	0,00
1	0,50	0,50	0,50
2	0,25	0,50	1,00
Celkem	**1,00**	**1,00**	**1,50**

Určete směrodatnou odchylku a variační koeficient počtu vadných výrobků v testovaném vzorku.

Příklad 8

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

$$\mathrm{E}(X) = \sum_{(i)} x_i \cdot P(x_i) = 1{,}00, \quad \mathrm{D}(X) = 0{,}50 \Rightarrow \sigma(X) = \sqrt{D(X)} = \frac{\sqrt{2}}{2} \cong 0{,}71$$

x	$P(x)$	$x \cdot P(x)$	$x^2 \cdot P(x)$
0	0,25	0,00	0,00
1	0,50	0,50	0,50
2	0,25	0,50	1,00
Celkem	**1,00**	**1,00**	$\mathbf{1,50}$

Směrodatná odchylka počtu vadných výrobků v testovaném vzorku je 0,71 ks.

Příklad 8

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

$$\mathrm{E}(X) = \sum_{(i)} x_i \cdot P(x_i) = 1{,}00, \quad \mathrm{D}(X) = 0{,}50, \quad \sigma(X) \cong 0{,}71 \Rightarrow \gamma(X) = \frac{\sigma(X)}{E(X)} \cong 0{,}71 \approx 71\ \%$$

x	$P(x)$	$x \cdot P(x)$	$x^2 \cdot P(x)$
0	0,25	0,00	0,00
1	0,50	0,50	0,50
2	0,25	0,50	1,00
Celkem	**1,00**	**1,00**	$\mathbf{1,50}$

Variační koeficient počtu vadných výrobků v testovaném vzorku je cca 71 %, tj. lze mluvit o nehomogenním (silně rozptýleném) rozdělení.

Transformace náhodné veličiny

- Transformací NV *X* rozumíme aplikaci prosté reálné funkce *g* tak, že vznikne nová náhodná veličina
$$Y = g(X).$$

- NV *Y* může nabývat jiných hodnot než NV *X*,

  může mít také jiné rozdělení pravděpodobnosti, které chceme najít,

  tj. hledáme pravděpodobnostní f-ci, resp. hustotu pravděpodobnosti NV *Y*.

Transformace náhodné veličiny

$$Y = g(X)$$

$$\forall y \in \mathbb{R}: \quad F_Y(y) = P(Y < y) = P(g(X) < y)$$

Pro diskrétní NV:

$$P_Y(y) = P(Y = y) = P(g(X) = y) = P\left(X = g^{-1}(y)\right)$$

Příklad 9

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

Počet vadných výrobků	Pravděpodobnost
0	0,25
1	0,50
2	0,25

Na evidenci výsledků kontroly kvality testovaných výrobků potřebujeme 10 minut plus dalších 5 minut na evidenci podrobného výsledku v případě, že je výrobek vadný. Určete pravděpodobnostní funkci celkového času potřebného k evidenci výsledků kontroly kvality.

Příklad 9

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

x	$P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

Y … čas potřebný k evidenci výsledků kontroly kvality (min)

$$Y = 10 + 5X$$

$$P_Y(y) = P(Y = y) = P(10 + 5X = y) = P\left(X = \frac{y-10}{5}\right) = P_X\left(\frac{y-10}{5}\right)$$

y	$P(y)$
$10 + 5 \cdot 0$	0,25
$10 + 5 \cdot 1$	0,50
$10 + 5 \cdot 2$	0,25
Celkem	**1,00**

Na evidenci výsledků kontroly kvality testovaných výrobků potřebujeme 10 minut plus dalších 5 minut na evidenci podrobného výsledku v případě, že je výrobek vadný. Určete pravděpodobnostní funkci celkového času potřebného k evidenci výsledků kontroly kvality.

Příklad 9

Při ověřování kvality výroby jsou náhodně vybrány dva výrobky a je testována jejich kvalita. Počet vadných výrobků mezi vybranými modelujeme náhodnou veličino X. Z dlouhodobého pozorování jsou známy údaje uvedené v následující tabulce.

X … počet vadných výrobků

x	$P(x)$
0	0,25
1	0,50
2	0,25
Celkem	**1,00**

Y … čas potřebný k evidenci výsledků kontroly kvality (min)

$$Y = 10 + 5X$$

$$P_Y(y) = P(Y = y) = P(10 + 5X = y) = P\left(X = \frac{y-10}{5}\right) = P_X\left(\frac{y-10}{5}\right)$$

y	$P(y)$
10	0,25
15	0,50
20	0,25
Celkem	**1,00**

Na evidenci výsledků kontroly kvality testovaných výrobků potřebujeme 10 minut plus dalších 5 minut na evidenci podrobného výsledku v případě, že je výrobek vadný. Určete pravděpodobnostní funkci celkového času potřebného k evidenci výsledků kontroly kvality.

Děkuji za pozornost!

martina.litschmannova@vsb.cz

VŠB TECHNICKÁ UNIVERZITA OSTRAVA | FAKULTA ELEKTROTECHNIKY A INFORMATIKY | KATEDRA APLIKOVANÉ MATEMATIKY