Cvičenie 3a – Náhodná veličina – diskrétna

Doteraz sme sa zaoberali náhodnými udalosťami. Tie môžu buď nastať alebo nenastať.

Náhodná veličina je komplexnejšia, ide o systém viacerých náhodných udalostí. Každá z udalostí má svoju číselnú hodnotu, ktorá buď vyplýva „z podstaty veci“, alebo je priradená definíciou . Pri svojej realizácii (pokuse, meraní, ankete, teste) náhodná veličina nadobúda niektorú zo svojich možných hodnôt, a to vždy s nejakou pravdepodobnosťou.

Tabuľka možných hodnôt a ich pravdepodobností sa nazýva pravdepodobnostná funkcia a je diskrétnou obdobou distribučnej funkcie (pri spojitých veličinách).

X	x_1	x_2	...	x_n
P (X = x)	p_1	p_2	...	p_n

Súčet pravdepodobností všetkých možných hodnôt náhodnej veličiny musí byť 1.

Distribučná funkcia je kumulatívnym súčtom hodnôt P(x).

X	x_1	x_2	x_3	...	x_n
F (x)	p_1	p_1+p_2	$p_1+p_2+p_3$	...	$p_1+p_2+p_3+p_n$

Stredná hodnota

Strednú hodnotu veličiny X označíme E(X). V princípe ide o „vážený priemer“ hodnôt x_i, kde váhami sú pravdepodobnosti p_i.

$$E(X) = x_1p_1 + x_2p_2 + ... + x_np_n$$

Rozptyl (disperzia, variancia)

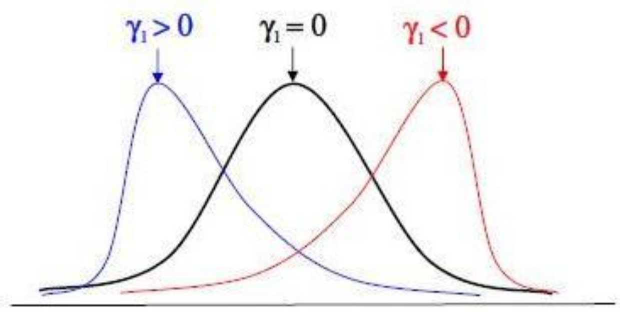
Rozptyl , označovaný D, var alebo σ^2, je hodnota, ktorá opisuje mieru „rozhádzanosti“ hodnôt x_i okolo E(X).

$$D(X) = var(X) = \sigma^2 (X) = E((X-E(X))^2) = E(X^2) - (E(X))^2$$

$$= x_1^2p_1 + x_2^2p_2 + ... + x_n^2p_n - (E(X))^2$$

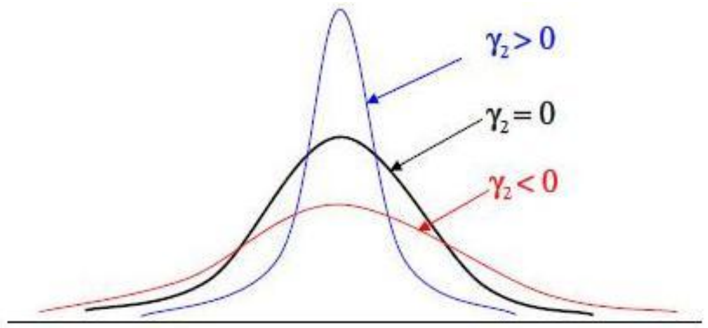
Koeficient šikmosti

$\gamma_1 = E((X - E(X))^3) / (D(X))^{3/2}$



Koeficient špicatosti

$\gamma_2 = E((X - E(X))^4) / (D(X))^2 - 3$



Obrázky sú zo stránky
https://wikisofia.cz/wiki/%C5%A0ikmost_a_%C5%A1pi%C4%8Datost

Cf.
https://en.wikipedia.org/wiki/Skewness
https://en.wikipedia.org/wiki/Kurtosis

Modus

Modus je „typická hodnota“, teda hodnota, ktorá sa vyskytuje najčastejšie.
V prípade náhodných veličín, ktoré sú zadané pravdepodobnostnou funkciou (diskrétne) alebo funkciou hustoty (spojité), ide o tú z možných hodnôt, ktorá má najvyššiu pravdepodobnosť.
Modus nemusí byť určený jednoznačne (!).

Kvantil

Číslo q_p, teda p-kvantil, kde p je číslo z intervalu (0,1), je taká spomedzi možných hodnôt náhodnej veličiny X, že

$$P(x \leq q_p) \geq p$$
$$P(x \geq q_p) \geq 1-p$$

Ak by vzorcu vyhovovali až dve hodnoty, za p-kvantil sa považuje ich aritmetický priemer.

Špeciálne kvantily:

0.5-kvantil = medián
0.25-kvantil = dolný kvartil
0.75-kvantil = horný kvartil

Riešené príklady

<u>Príklad 1:</u>

Hádžeme naraz 2 hracími kockami (červená a modrá). Ako náhodnú veličinu X budeme sledovať a zaznamenávať *maximum* z oboch „padnutých“ hodnôt. Tj. buď väčšie z oboch čísel, alebo ak je hodnota na oboch rovnaká, výstupom bude <u>tá</u> hodnota.

a) Zostavte tabuľku pravdepodobnostnej funkcie P a distribučnej funkcie F.

b) Zistite hodnotu P(x je párne), P(x≤3).

c) Nájdite X^{-1}.

d) Vypočítajte E(X) a D(X), koeficienty šikmosti a špicatosti.

Riešenie:

Všetky možné dvojice čísel na dvoch kockách (36 možností) si vieme usporiadať do tabuľky.

[čm]

č \ m	**1**	**2**	**3**	**4**	**5**	**6**
1	11	12	13	14	15	16
2	21	22	23	24	25	26
3	31	32	33	34	35	36
4	41	42	43	44	45	46
5	51	52	53	54	55	56
6	61	62	63	64	65	66

V ďalšej tabuľke každú z možností prepíšeme hodnotou veličiny X, teda ponecháme tam maximum z oboch čísel. Tento „prepis“ elementárnych udalostí na hodnoty X si môžeme predstaviť ako zobrazenie, a to konkrétne $X : \{1,2,3,4,5,6\}^2 \rightarrow \{1,2,3,4,5,6\}$.

[X]

č \ m	**1**	**2**	**3**	**4**	**5**	**6**
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

Poznámka: Uvedené tabuľky sú pomôckou, aby sme si vedeli lepšie predstaviť situáciu. Na riešenie príkladu však nie sú nutné a dá sa prejsť priamo k veci.

a) Na základe poslednej tabuľky X už vieme napísať pravdepodobnostnú funkciu náhodnej veličiny X. Tá nadobúda hodnoty od 1 do 6. Jednotlivé pravdepodobnosti budú mať v menovateli 36 (počet všetkých možností) a v čitateli počet výskytov danej hodnoty v tabuľke vyššie.

x	1	2	3	4	5	6
P (X=x)	1/36	3/36	5/36	7/36	9/36	11/36
F (x)	1/36	4/36	9/36	16/36	25/36	36/36

X je označenie náhodnej veličiny, x sú konkrétne hodnoty, ktoré X nadobúda.

X = x : Náhodná veličina X nadobudne hodnotu x.

b) P(x je párne) = (3+7+11)/36 = 21/36 = 7/12
$P(x \le 3) = (1+3+5)/36 = 9/36 = 1/4$ resp. z distrib. funkcie F(3) = 9/36

c) Zobrazenie X^{-1} (inverzné zobrazenie k X) nám ku každej hodnote x prezradí, ktoré elementárne udalosti za ňou stoja.

$X^{-1}(1) = \{11\}$
$X^{-1}(2) = \{12,22,21\}$
$X^{-1}(3) = \{13,23,33,32,31\}$
$X^{-1}(4) = \{14,24,34,44,43,42,41\}$
$X^{-1}(5) = \{15,25,35,45,55,54,53,52,51\}$
$X^{-1}(6) = \{16,26,36,46,56,66,65,64,63,62,61\}$

d) E(X) = 1*1/36 + 2*3/36 + 3*5/36 + 4*7/36 + 5*9/36 + 6*11/36 = 161/36 = 4.47222222222

$D(X) = 1^2*1/36 + 2^2*3/36 + 3^2*5/36 + 4^2*7/36 + 5^2*9/36 + 6^2*11/36 - (161/36)^2 =$

$= 2.0825617284$

Doplníme tabuľku:

X	1	2	3	4	5	6
X - E(X)	-3.4722	-2.4722	-1.4722	-0.4722	0.5278	1.5278
P	1/36	3/36	5/36	7/36	9/36	11/36

$\gamma_1 = E((X - E(X))^3) / (D(X))^{3/2} = -0.58539$

Výsledok dostávame dosadením do vzorca. Postup výpočtov uvedieme tak, ako bol vysvetlený aplikácii Octave (v Matlabe to ide podobne):

```
x=1:6                                   hodnoty, ktoré nadobúda X
xme = x-161/36                          X - E(X)
(xme.^3)*p'/ (2.0825617284)^(3/2)       dosadenie do vzorca
ans = -0.58539                          výsledok
```

$\gamma_2 = E((X - E(X))^4) / (D(X))^2 - 3 = -0.81032$

Postup výpočtov:

```
(xme.^4)*p'/((2.0825617284)^(2)) - 3
ans = -0.81032
```

Príklad 1++:

Pokračovanie príkladu 1 – nájdite modus, medián a oba kvartily.

Riešenie:

– Modus je hodnota s najvyššou pravdepodobnosťou (11/36), teda 6.

– Medián (= 0.5-kvantil) bude tá hodnota, kde distribučná funkcia prekročí hranicu 0.5. A to je <u>číslo 5</u>. Overme si, či „sedí vzorec“:

$$P(x \le 5) = P(1)+P(2)+P(3)+P(4)+P(5) = F(5) = 25/36 \ \ge\ 0.5$$
$$P(x \ge 5) = P(5)+P(6) = 1 - F(4) = 20/36 \ \ge\ 0.5$$

– Nájdime ďalej 0.75-kvantil (tzv. horný kvartil):

Číslo 25/36 = 0.69444 je menšie ako 0.75, až v 6-tke sa prekoná 0.75.

$$P(x \le 6) = P(1)+P(2)+P(3)+P(4)+P(5)+P(6) = F(6) = 1 \ \ge\ 0.75$$
$$P(x \ge 6) = P(6) = 11/36 = 0.30555 \ \ge\ 0.25$$

– Nájdime ešte 0.25-kvantil (tzv. dolný kvartil).

Zdanlivo potešujúci fakt, že F(3) = 9/36, teda presne 0.25, predstavuje v skutočnosti komplikáciu. Podľa vzorca totiž platí:

$$P(x \le 3) = F(3) = 0.25 \ \ge\ 0.25$$
$$P(x \ge 3) = 1\text{-}F(2) = 32/36 = 0.88888 \ \ge\ 0.75$$

$$P(x \le 4) = F(4) = 16/36 = 0.4444 \ \ge\ 0.25$$
$$P(x \ge 4) = 1\text{-}F(3) = 27/36 = 0.75 \ \ge\ 0.75$$

Slovne vyjadrené, vzorcu vyhovuje hodnota 3 aj 4, teda v zmysle prvotnej definície 0.25-kvantil nie je určený jednoznačne. Táto situácia sa rieši podľa široko akceptovanej dohody tak, že sa vezme presný stred medzi možnými hodnotami, tj. v tomto príklade (3+4)/2 = 3.5 .

Príklad 2:

Máme tri hracie kocky (červená, modrá, zelená) upravené tak, že na každej sa nachádzajú po dve jednotky, dve dvojky a dve trojky. Hádžeme naraz týmito 3 kockami. Ako náhodnú veličinu X budeme sledovať a zaznamenávať súčet dvoch väčších (v zmysle ≥) z troch „padnutých" hodnôt. Inými slovami, z troch hodnôt škrtneme jednu – minimum, a zvyšné dve sčítame.

a) Zostavte tabuľku pravdepodobnostnej funkcie P a distrib. funkcie F.

b) Nájdite X^{-1}.

c) Vypočítajte E(X) a D(X).

d) Nájdite medián, kvartily a modus.

Riešenie:

a) Najprv si v úspornej forme rozpíšeme elementárne udalosti a im prislúchajúce hodnoty X.

Elem. udalosti	X
111	1+1=2
211 (3 permutácie)	1+2=3
221 (3 permutácie)	2+2=4
222	2+2=4
311 (3 permutácie)	1+3=4
321 (6 permutácií)	2+3=5
322 (3 permutácie)	2+3=5
331 (3 permutácie)	3+3=6
332 (3 permutácie)	3+3=6
333	3+3=6

Veličina X nadobúda hodnoty 2,3,4,5,6. Všetkých elementárnych udalostí je 27 (3*3*3) – to bude menovateľ v pravdepodobnostiach. Čitateľ bude počet elem. udalostí s príslušnou hodnotou x.

x	2	3	4	5	6
P (X=x)	1/27	3/27	7/27	9/27	7/27
F (x)	1/27	4/27	11/27	20/27	27/27

b) $X^{-1}(2) = \{111\}$

$X^{-1}(3) = \{211, 121, 112\}$ = Perm(112)

$X^{-1}(4) = \{221, 212, 122, 222, 311, 131, 113\}$ = Perm(122) ∪{222}∪ Perm(113)

$X^{-1}(5) = \{321, 312, 213, 231, 132, 123, 322, 232, 223\}$ = Perm(123) ∪ Perm(223)

$X^{-1}(6) = \{331, 313, 133, 332, 323, 233, 333\}$ = Perm(133) ∪ Perm(233) ∪{333}

Zdĺhavé výpisy sme čiastočne zostručnili zhrnutím permutácií trojice čísel do jedného výrazu.

c) E(X) = 126/27 = 4.66666

D(X) = (4+9*3+16*7+25*9+36*7)/27 - (126/27)^2 = 1.18518518519

d) Jednoduché.

Príklad 3:

Máme dve dvojeurovky a tri jednoeurovky. Všetkých 5 mincí hodíme na stôl (tak, aby neuleteli na zem) a spočítame čísla, ktoré budú navrchu (opačná strana mince sa ráta za nulu). Tento súčet bude náhodná veličina X.

a) Zostavte tabuľku pravdepodobnostnej funkcie P a distr. f. F.
b) Nájdite X^{-1}.
c) Vypočítajte E(X) a D(X).
d) Nájdite medián, 0.15-kvantil a modus.

Riešenie:

a) Najprv si v úspornej forme rozpíšeme elementárne udalosti a im prislúchajúce hodnoty X.

	2€	1€		
x=0	0+0+	0+0+0	1	1
x=1	0+0+	1+0+0	3 permutácie	3
x=2	0+0+	1+1+0	3 permutácie	
	0+2+	0+0+0	2 permutácie	5
x=3	0+0+	1+1+1	1	
	2+0+	1+0+0	2*3 permutácie	7
x=4	2+0+	1+1+0	2*3 permutácie	
	2+2+	0+0+0	1	7
x=5	2+0+	1+1+1	2 permutácie	
	2+2+	1+0+0	3 permutácie	5
x=6	2+2+	1+1+0	3 permutácie	3
x=7	2+2+	1+1+1	1	1

Všetkých elementárnych udalostí je 32, na každej z 5 mincí sú 2 možnosti.

x	0	1	2	3	4	5	6	7
P (X=x)	1/32	3/32	5/32	7/32	7/32	5/32	3/32	1/32
F (x)	1/32	4/32	9/32	16/32	23/32	28/32	31/32	32/32

b) Všetko podstatné k tejto úlohe nájdeme v rozpise v bode a), stačí to len vhodne „naformátovať“.

c) Dosadíme do vzorca tak ako doteraz.

d) Modus v tomto príklade, to sú hodnoty 3 a 4. *(Modus nemusí byť určený jednoznačne.)*

Medián: ľahko sa presvedčíme, že vzorcu vyhovujú x=3 aj x=4.
Za medián preto vyhlásime kompromisnú hodnotu 3.5 .

Hľadajme 0.15-kvantil. V riadku s distribučnou funkciou vidíme, že 4/32 je ešte málo, ale už 9/32 hranicu 0.15 prekoná. Hľadaným kvantilom je $q_{0.15}$= 2. Ubezpečte sa, že „vzorec“ je splnený.

Príklad 4:

Máme jednoeurovku. Hodíme ju na stôl dvakrát po sebe a sčítame hodnoty. Aké číslo vyjde, toľkokrát ešte hodíme. Veličinou X bude celkový súčet.

Nájdite pravdepodobnostnú funkciu X.

Riešenie:

Rozpíšeme všetky možné scenáre vývoja do diagramu. Všetky vetvy majú pravdepodobnosť 0.5 – nebudeme to písať do obrázku.

```
                          štart
              /                                 \
              0                                 1
      /               \                /                 \
0 (x=0)               1                0                 1
                 /          \       /       \        /        \
              0 (x=1)   1 (x=2)  0 (x=1)  1 (x=2)    0         1
                                                   /   \     /    \
                                            0 (x=2) 1(x=3) 0 (x=3) 1 (x=4)
```

Pravdepodobnosti v tabuľke musíme vypočítať z diagramu, nebude to všade rovnaký menovateľ.

x	0	1	2	3	4
P (X=x)	0.25	0.25	0.3125	0.125	0.0625

Príklad 5:

Dvaja basketbalisti A, B na striedačku (začína A) hádžu z úctivej vzdialenosti na kôš. Pravdepodobnosť zásahu je pri hráčovi A 0.8, pri B je to 0.7. Náhodná veličina bude poradie hodu, pri ktorom prvýkrát padol kôš.

Nájdite pravdepodobnostnú funkciu pre X.

Riešenie:

Možných výsledkov je nekonečne veľa. Kto má chuť, nech vypíše všetky. Lepšie však bude skúsiť nájsť nejaký všezahŕňajúci šikovný vzorec.

Prvý hod:	P(1) = 0.8	(hádže A, len od neho to závisí)
Druhý hod:	P(2) = 0.2*0.7 = 0.14	(A netrafí, po ňom B trafí)
Tretí hod:	P(3) = 0.2*0.3*0.8 = 0.046	(A netrafí, B netrafí, A trafí)
Štvrtý hod:	P(4) = 0.2*0.3*0.2*0.7 = 0.0096	
	P(5) = 0.2*0.3*0.2*0.3*0.8 = 0.00288	
	P(6) = 0.2*0.3*0.2*0.3*0.2*0.7 = 0.000504	

....

$$P(2k) = 0.2^k * 0.3^{k-1} * 0.7$$
$$P(2k+1) = 0.2^k * 0.3^k * 0.8$$

Našli sme všeobecný vzorec pre pravdepodobnostnú funkciu na neohraničenom intervale. Ako však vidno z vyčíslených hodnôt, P(i) veľmi rýchlo klesá. Všetko podstatné sa takmer určite udeje na začiatku. Ak jednotlivé pokusy nie sú oddelené reklamnou prestávkou, kupovať k tomu pukance sa neoplatí.

Neriešené príklady

1. Hádžeme naraz 3 hracími kockami (červená, modrá, zelená). Ako náhodnú veličinu X budeme sledovať a zaznamenávať *maximum* z troch „padnutých" hodnôt.

a) Zostavte tabuľku pravdepodobnostnej funkcie P. ((k^3-(k-1)^3) / 216)

b) Nájdite X^{-1}. (...)

c) Vypočítajte E(X) a D(X). (4.9583, 1.30878)

d)** Hádžme n kockami a sledujme maximum z hodnôt. Nájdite funkciu P.

2. Riešte príklad 2 (z riešených príkladov) pre 3 bežné (neupravené) kocky. Optimalizujte postup tak, aby ste sa vyhli (kde to je možné) zdĺhavému vypisovaniu.

([1,3,7,12,19,27,34,36,34,27,16]/216)

3. Hádžeme dve hracie kocky (červená, zelená). Ako náhodnú veličinu X zaznamenávame rozdiel (v absolútnej hodnote) oboch „padnutých" hodnôt.

a) Zostavte tabuľku pravdepodobnostnej funkcie P. ([6,10,8,6,4,2]/36)

b) Nájdite X^{-1}. (...)

c) Vypočítajte E(X) a D(X). (70/36; 2.0525)

4. Hádžeme tri hracie kocky (červená, modrá, zelená). Ako náhodnú veličinu X zaznamenávame rozdiel (v absolútnej hodnote) maxima a minima „padnutých" hodnôt.

a) Zostavte tabuľku pravdepodobnostnej funkcie P. ([6,30,48,54,48,30]/216)

b) Vypočítajte E(X) a D(X). (...)

5. Máme nasledovné české mince: jednu 1-korunáčku, 2 dvojkoruny, jednu 5-korunu a dve 10-koruny. Hádžeme túto sadu mincí na stôl a sledujeme X = súčet padnutých hodnôt (znak = 0). Aké hodnoty nadobúda X a s akou pravdepodobnosťou?

([1,1,2,2,1,2,1,2,2,1,3,2,4,4,2,4,2,4,4,2,3,1,2,2,1,2,1,2,2,1,1] / 64)

6. Hádžeme naraz 2 jednoeurovkami, 2 dvojeurovkami a jednou hracou kockou (1...6). X = súčet padnutých hodnôt. Aké hodnoty nadobúda X a s akou pravdepodobnosťou?

([1,3,6,10,13,15,15,13,10,6,3,1] / 96)

7. Riešte príklad 4 (z riešených úloh) pre tri/štyri hody mincou na začiatku.

(...)

8. Dvaja basketbalisti A, B hádžu na kôš v takomto poradí: A B B A A A B B B B . Pravdepodobnosť zásahu je pri hráčovi A 0.4, pri B je to 0.2. Náhodná veličina bude poradie hodu, pri ktorom prvýkrát padol kôš.

Nájdite pravdepodobnostnú funkciu pre X. (triviálne)

9. Dvaja basketbalisti A, B hádžu striedavo na kôš, každý 2-krát, začína A. Pravdepodobnosť zásahu je pri hráčovi A 0.6, pri B je to 0.4. Náhodná veličina bude počet košov, ktoré sa zadaria počas tých 4 hodov.

Nájdite pravdepodobnostnú funkciu pre X. (0.0576, 0.2496, 0.3856, 0.2496, 0.0576)

10. Vypočítajte E(X), D(X), γ_1, γ_2 pre náhodnú veličinu danú tabuľkou.
Nájdite modus, medián a oba kvartily.

X	0	2	4	6	8
P	1/10	1/10	3/10	2/10	3/10

(5, 6.6, -0.42, -0.81,
modus 4 a 8, med. 5, kvartily 4 a 8)

11. Vypočítajte E(X), D(X), γ_1, γ_2 pre náhodnú veličinu danú tabuľkou:

X	-4	-2	0	2	4
P	1/10	2/10	1/10	3/10	3/10

(1, 7.4, -0.47, -1.1118
modus 2 a 4, med. 2, kvartily -2 a 4)

11. Vypočítajte γ_1, γ_2 a nájdite modus, medián, kvartily pre náhodnú veličinu
z neriešených príkladov 1, 3, 4.