Arvkarakteristikud

Arvkarakteristikute roll

Juhusliku suuruse jaotusfunktsioon kirjeldab täielikult tema tõenäosuslikku käitumist. Sageli ei ole aga vajalik täielik informatsioon juhuslikust suurusest. Piisab, kui juhuslikust suurusest on mingi üldine ettekujutus. Juhusliku suuruse üldiseks iseloomustajaks kasutatakse tõenäosusteoorias **juhusliku suuruse arvkarakteristikuid**. Näiteks sagedamini kasutatavad on juhusliku suuruse paiknemist iseloomustavad karakteristikud (mingis mõttes "keskmist" väärtust väljendavad arvud) ja hajuvust ("laialipillutatust") iseloomustavad karakteristikud.

Diskreetse juhusliku suuruse keskväärtus

Üks tähtsam arvkarakteristik, mis iseloomustab juhusliku suuruse paiknemist, on keskväärtus.
Olgu juhuslik suurus X määratud jaotustabeliga:

$\boldsymbol{X}$	x_1	x_2	...	x_n
$\boldsymbol{p}$	p_1	p_2	...	p_n

Diskreetse juhusliku suuruse *X* keskväärtuseks (matemaatiliseks ootuseks) EX (kasutatakse ka tähistusi MX ja mx) nimetatakse suuruse võimalike väärtuste ja nende tõenäosuste korrutiste summat

$$EX = \sum_{i=1}^{n} x_i p_i \ .$$

Näide. Aadam ja Eeva istuvad paradiisis ja mängivad kulli ja kirja. Kui tuleb välja kull (sündmus A), annab Aadam Eevale krooni, kui kiri, siis Eeva Aadamale. Leida Aadama võidu keskväärtus.

Lahendus. Sündmuse toimumisele-mittetoimumisele seame vastavusse juhusliku suuruse X järgnevalt:

Kui A toimub – siis $X = 1$

A ei toimu – siis $X = 0$.

Saame jaotustabeli:

$\boldsymbol{X}$	0	1
$\boldsymbol{p}$	0,5	0,5

Leiame keskväärtuse $EX = \sum_{i=1}^{n} x_i p_i = 0 \cdot 0{,}5 + 1 \cdot 0{,}5 = 0{,}5$.

Sellist jaotust nimetatakse **kahepunktiliseks jaotuseks**.

Võib tähele panna, et:

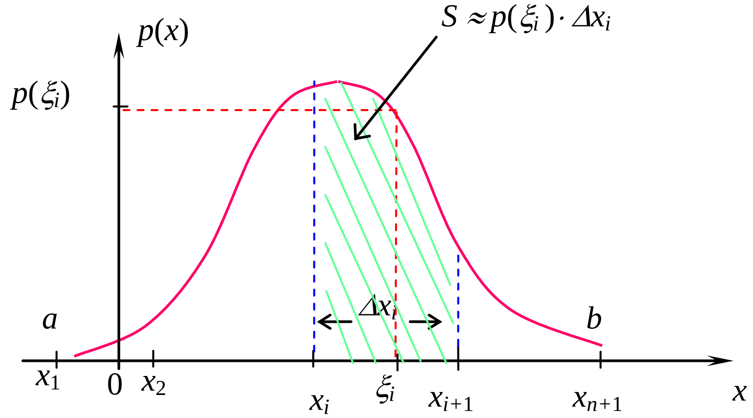
- Diskreetse juhusliku suuruse keskväärtus on ligikaudu võrdne katseseeria jooksul ilmnenud juhusliku suuruse väärtuste aritmeetilise keskmisega ning sealjuures seda täpsemalt, mida suurem on katsete arv.
- Kui viia läbi mitu katseseeriat, siis iga katseseeria jaoks leitud juhusliku suuruse väärtuste aritmeetilised keskmised kuhjuvad konstandi ümber, milleks on selle juhusliku suuruse keskväärtus.

Pideva juhusliku suuruse keskväärtus

Vaatleme nüüd pidevat juhuslikku suurust *X*, mille võimalike väärtuste hulk on lõik [*a*; *b*]. Olgu $p(x)$ selle juhusliku suuruse tihedusfunktsioon. Jaotame lõigu [*a*; *b*] *n* osalõiguks pikkustega $\Delta x_1, \Delta x_2, ..., \Delta x_n$. Valime igal osalõigul ühe punkti abstsissidega $\xi_1, \xi_2, \ldots, \xi_n$. Teame, et korrutis $p(\xi_i)\Delta x_i$ $(i = 1, \ldots, n)$ on ligikaudu võrdne tõenäosusega, et juhuslik suurus *X* satub osalõiku pikkusega Δx_i. Summa $\sum_{i=1}^{n} \xi_i p(\xi_i)\Delta x_i$ on aga analoogne diskreetse juhusliku suuruse keskväärtuse definitsiooniga. Minnes selles summas piirile nii, et suurim arvudest Δx_i $(i = 1, \ldots, n)$ läheneb nullile, saame määratud integraali

$$\lim_{\max \Delta x_i \to 0} \sum_{i=1}^{n} \xi_i p(\xi_i)\Delta x_i = \int_a^b x p(x)\, dx,$$

mis ongi pideva juhusliku suuruse keskväärtus.



Seega kui pideva juhusliku suuruse väärtused kuuluvad lõiku [*a*; *b*], siis selle juhusliku suuruse keskväärtuseks nimetatakse arvu

$$EX = \int_a^b x p(x)\, dx.$$

Kui pideva juhusliku suuruse väärtused paiknevad kogu arvteljel, siis keskväärtus on

$$EX = \int_{-\infty}^{\infty} x p(x)\, dx.$$

Näide. (Ühtlase jaotuse keskväärtus.) Olgu juhusliku suuruse *X* tihedusfunktsioon antud kujul

$$p(x) = \begin{cases} 0, & kui\ x < a, \\ \dfrac{1}{b-a}, & kui\ a \le x \le b, \\ 0, & kui\ x > b. \end{cases}$$

Leida juhusliku suuruse keskväärtus.

Lahendus. Vastavalt eelpool toodud valemile

$$EX = \int_{-\infty}^{\infty} x \cdot p(x)\, dx = \int_{-\infty}^{a} 0 dx + \int_a^b x \cdot \frac{1}{\mathrm{b-a}} dx + \int_b^{\infty} 0 dx =$$

$$= \int_a^b x \cdot \frac{1}{\mathrm{b-a}} dx = \frac{1}{b-a} \left. \frac{x^2}{2} \right|_a^b = \frac{1}{b-a}\left(\frac{b^2-a^2}{2}\right) = \frac{a+b}{2}$$

Keskväärtuse omadused

1. Keskväärtus asub kindlasti juhusliku suuruse vähima ja suurima võimaliku väärtuse vahel, kuid ei tarvitse alati kuuluda suuruse võimalike väärtuste hulka.

Näide. Täringuviskel saadava silmade arvu keskväärtus

$$E(X) = \left(1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6}\right) = 3{,}5$$

Näeme, et $3{,}5 \notin \{1;2;3;4;5;6\}$, aga $1 \le 3{,}5 \le 6$.

2. Konstandi keskväärtuseks on sama konstant ehk $E(C) = C$.

Põhjendus. Konstanti C võib vaadelda diskreetse juhusliku suurusena, mis tõenäosusega $p_1 = 1$ saavutab oma ainukese väärtuse, milleks on see konstant ise: $x_1 = C$. Seetõttu tema keskväärtus avaldub $E(C) = x_1 \cdot p_1 = C \cdot 1 = C$.

3. Konstantse kordaja võib tuua keskväärtuse tähise ette ehk $E(CX) = C \cdot EX$.

Põhjendus. (diskreetse juhusliku suuruse korral):

$$E(CX) = \sum_{i=1}^{n} C x_i p_i = C \sum_{i=1}^{n} x_i p_i = C \cdot EX$$

Põhjenda selle omaduse kehtivust ka pideva juhuslikul suurusel korral!

4. Juhuslike suuruste summa keskväärtus on võrdne nende suuruste keskväärtuste summaga ehk $E(X+Y) = EX + EY$.

5. Sõltumatute juhuslike suuruste korrutise keskväärtus on võrdne nende suuruste keskväärtuste korrutisega ehk $E(XY) = EX \cdot EY$.

6. Juhusliku suuruse hälbed kahele poole keskväärtust keskmiselt tasakaalustuvad

$$E(X - EX) = EX - E(EX) = EX - EX = 0.$$

See on seletatav ka sellega, et võimalikud hälbed on erinevate märkidega ja summaarselt kompenseeruvad. (Suuruse väärtuse ja mingi jääva arvu C vahet nimetatakse **väärtuse hälbeks** selle arvu suhtes.)

Järeldus. Keskväärtuse omaduste põhjal võib teha järelduse

$$E(AX + B) = A \cdot EX + B.$$

Põhjenda võrduse kehtivust!

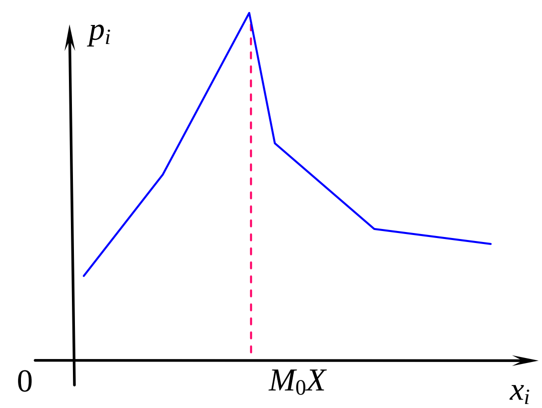
Näide. On teada, et $EX = 5$. Leida lineaaravaldise $4 \cdot X - 6$ keskväärtus.

Lahendus. $E(4 \cdot X - 6) = 4 \cdot EX - 6 = 4 \cdot 5 - 6 = 14$

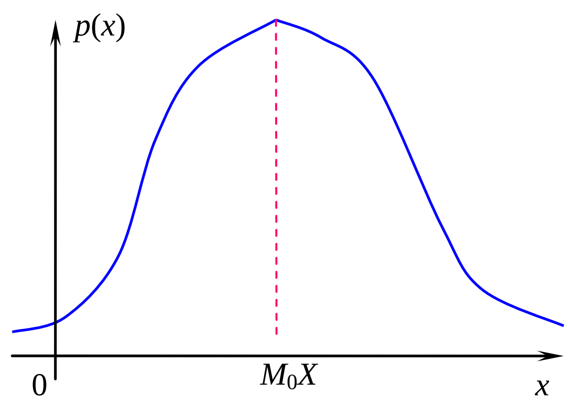
Juhusliku suuruse mood ja mediaan

Peale keskväärtuse, mis on juhusliku suuruse paiknemise põhiline arvkarakteristik, kasutatakse praktikas ka teisi paiknemise karakteristikuid nagu mood ja mediaan.

Diskreetse juhusliku suuruse moodiks M_0X nimetatakse selle suuruse kõige tõenäosemalt esinevat väärtust.
Pideva juhusliku suuruse mood on selle suuruse niisugune väärtus, mille korral tihedusfunktsioonil on maksimum.



Diskreetse juhusliku suuruse mood



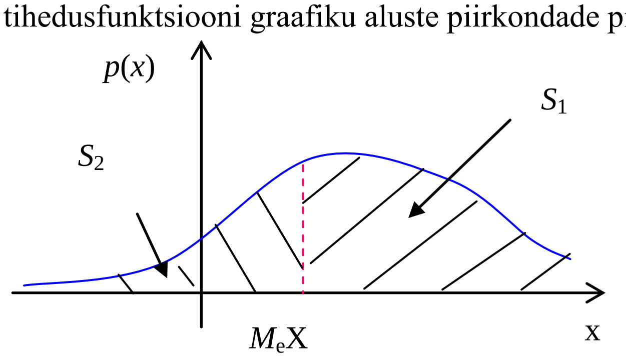
Pideva juhusliku suuruse mood

Kui jaotuspolügoonil või tihedusfunktsioonil on mitu maksimumi, siis on tegemist multimodaalse jaotusega. Kui aga maksimum puudub hoopis (selle asemel on miinimum), siis nimetatakse jaotust antimodaalseks.

Juhusliku suuruse mediaaniks M_eX nimetatakse niisugust arvu, millest suuremate ja väiksemate juhusliku suuruse väärtuste saamine on võrdtõenäone

$$P(X < M_e X) = P(X > M_e X) = 0{,}5\,.$$

Geomeetriliselt on mediaan sellise punkti abstsiss, millest vasakule ja paremale jäävate tihedusfunktsiooni graafiku aluste piirkondade pindalad on võrdsed.



$S_1 = S_2$

Kui jaotus on unimodaalne (ühe moodiga) ja sümmeetriline, siis langevad kõik kolm paiknemise karakteristikut – keskväärtus, mood ja mediaan kokku.

Juhusliku suuruse keskmine lineaarhälve

Juhusliku suuruse keskväärtus määrab suuruse keskmise nivoo. Ta ei kajasta aga suuruse hälbimist keskväärtusest. Nii näiteks on suurustel

X	-1	1
p	0,5	0,5

Y	-10	10
p	0,5	0,5

keskväärtused võrdsed: $EX = EY = 0$, kuid samas paigutuvad suuruse *Y* võimalikud väärtused keskväärtusest tunduvalt kaugemale kui suurusel *X*. Suuruse hajuvuse karakteristikuks ei kõlba ka $E(X - EX)$, kuna see on alati null. Üksikute hälvete märgist vabaneme, kui võtame absoluutväärtused $|X - EX|$. Juhusliku suuruse **keskmiseks lineaarhälbeks** *d* nimetatakse tema tsentreeritud hälbe absoluutväärtuse keskväärtust $E|X - EX|$:

$$d = \sum_{i=1}^{n} |x_i - EX| p_i$$

$$d = \int_{-\infty}^{\infty} |x - EX| p(x) dx$$

Teoreetilistes arutlustes on absoluutväärtusi sisaldavad avaldised tülikad. Üldiselt eelistatakse tsentreeritud hälbe absoluutväärtusele hälbe ruutu. Sel juhul kaob ka hälvete märk hälbimise suuna näitajana ning saab hinnata suuruse väärtuste keskmist hajuvust.

Dispersioon

Juhusliku suuruse dispersiooniks *DX* nimetatakse tema tsentreeritud hälbe ruudu keskväärtust: $DX = E(X - EX)^2$.

Keskväärtuse definitsiooni arvestades saame diskreetse juhusliku suuruse dispersiooni arvutamiseks järgmise eeskirja:

$$DX = \sum_{i=1}^{n} (x_i - EX)^2 \cdot p_i \ .$$

Pideva juhusliku suuruse dispersiooni arvutamise eeskiri on aga järgnev

$$DX = \int_{-\infty}^{\infty} (x - EX)^2 p(x) dx \ .$$

Dispersioon osutub väga sobivaks juhusliku suuruse väärtuste hajuvuse karakteristikuks. Puuduseks on asjaolu, et dispersiooni dimensioon on võrdne lähtesuuruse dimensiooni ruuduga. Selle puuduse kõrvaldamiseks on defineeritud veel üks karakteristik – standardhälve.

Dispersiooni ruutjuurt nimetatakse **standardhälbeks**:

$$\sigma(X) = \sqrt{DX}$$

Standardhälbe dimensioon ühtib lähtesuuruse dimensiooniga.

<u>Näide.</u> Juhuslikuks suuruseks on täringuviskel saadav silmade arv. Leida selle suuruse dispersiooni ja standardhälve.

Lahendus. Eelnevalt leidsime, et EX = 3,5. Arvestades diskreetse juhusliku suuruse dispersiooni valemit saame

$$DX = \sum_{i=1}^{6}(x_i - EX)^2 \cdot p_i = \sum_{i=1}^{6}(i-3{,}5)^2 \cdot \frac{1}{6} =$$

$$= \frac{1}{6}\left((1-3{,}5)^2 + (2-3{,}5)^2 + (3-3{,}5)^2 + (4-3{,}5)^2 + (5-3{,}5)^2 + (6-3{,}5)^2\right) \approx 2{,}92$$

$$\sigma(X) = \sqrt{DX} \approx 1{,}71$$

Näide. Leida ühtlase jaotuse dispersioon ja standardhälve.

Lahendus. Ühtlase jaotuse tihedusfunktsioon on:

$$p(x) = \begin{cases} 0, & x \notin [a,b] \\ \dfrac{1}{b-a}, & x \in [a,b] \end{cases}$$

Arvestades pideva juhusliku suuruse dispersiooni valemit saame

$$DX = \int_{-\infty}^{\infty}(x-EX)^2 p(x)dx = \int_a^b \frac{(x-(b+a)/2)^2}{b-a}dx = \frac{(b-a)^2}{12}$$

$$\sigma(X) = \sqrt{DX} = \frac{b-a}{2\sqrt{3}}$$

Dispersiooni omadusi

1. Konstandi dispersioon on null: $D(C) = 0$, kus C =const.

Põhjendus. Konstant C saab omada ainult ühte väärtust ja seega on tema hajuvus 0.

2. Konstantse teguri võib tuua dispersiooni märgi ette ruutu tõstetuna: $D(CX) = C^2 D(X)$.

3. Dispersiooni arvutamise teine eeskiri: $D(X) = E(X)^2 - (EX)^2$

4. Sõltumatute juhuslike suuruste puhul: $D(X+Y) = D(X) + D(Y)$

Omadustest 2 ja 4 järeldub: $D(X-Y) = D(X) + D(Y)$. *Põhjenda!*

Omadustest 4, 2 ja 1 järeldub: $D(CX+B) = C^2 D(X)$. *Põhjenda!*

Juhusliku suuruse momendid

Juhuslike suuruste põhiliste arvkarakteristikute üldistuseks on juhusliku suuruse momendi mõiste. Nimetus „moment“ tuleneb mehaanikast, kus seda mõistet kasutatakse massi jaotuse iseloomustamiseks.

Juhusliku suuruse ***r*-järku algmomendiks** nimetatakse suurust

$$M_r X = EX^r$$

Juhusliku suuruse keskväärtus on seega 1. järku algmoment.

Juhusliku suuruse ***r*-järku keskmomendiks** nimetatakse suurust

$$M_r^0 X = E(X-EX)^r$$

Juhusliku suuruse dispersioon on seega 2. järku keskmoment.

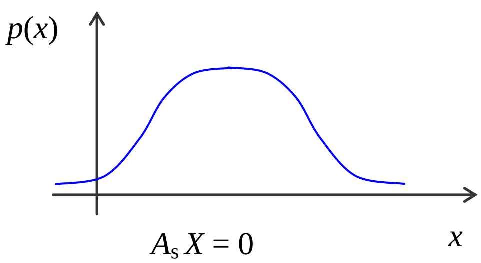
Teist järku keskmomendi kõrval vaadeldakse küllalt palju ka kolmandat ja neljandat järku keskmomente.

Juhusliku suuruse **asümmeetriateguriks** nimetatakse kolmandat järku keskmomendi ja standardhälbe kuubi jagatist:

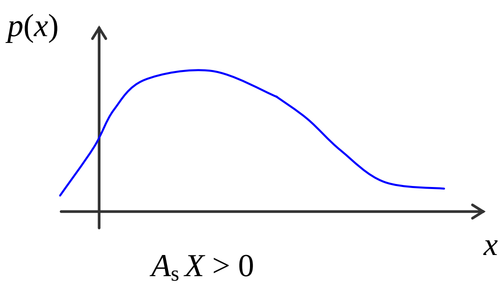
$$A_s X = \frac{M_3^0 X}{\sigma^3(X)}$$

(Kuna kolmandat järku keskmomendi dimensioon on võrdne juhusliku suuruse dimensiooni kuubiga, siis kasutatakse kolmandat järku keskmomendi ja standardhälbe kuubi suhet, mis on dimensioonita suurus.)

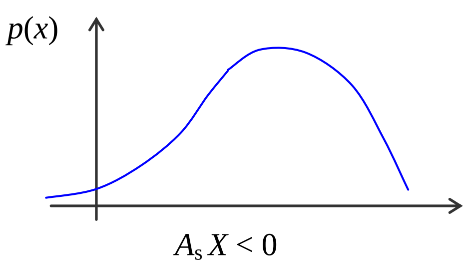
Sümmeetrilise jaotuse puhul on asümmeetriakordaja null.



Paremale väljavenitatud jaotuse puhul on asümmeetriakordaja positiivne.



Vasakule väljavenitatud jaotuse puhul on asümmeetriakordaja negatiivne.

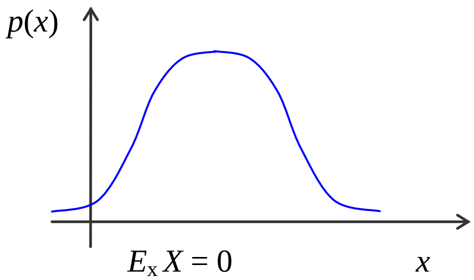


Juhusliku suuruse **ekstsessiks** nimetatakse suurust

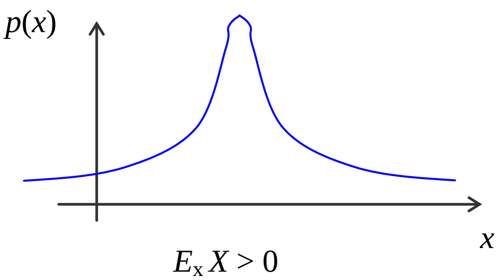
$$E_x X = \frac{M_4^0 X}{\sigma^4(X)} - 3.$$

Arv 3 lahutatakse sellepärast, et sageli kasutatava normaaljaotuse korral $\frac{M_4^0 X}{\sigma^4(X)} = 3$.

Normaaljaotus on seega võetud etaloniks, millega võrreldakse teisi jaotusi.
Juhusliku suuruse *X* ekstsess iseloomustab jaotuse järskust.



Ekstsess on positiivne siis, kui jaotusel on “rasked sabad” (ja enamasti ka terav tipp). Juhuslikul suurusel on üksikuid teistest liiga erinevaid väärtusi normaaljaotusega võrreldes.



Negatiivse ekstsessiga jaotus on suhteliselt lame, tihti tõkestatud väärtustega. Uuritaval suurusel on võrreldes normaaljaotusega vähem äärmuslikke väärtusi.