Klassikalised jaotused

Eelmistes peatükkides tutvusime juhusliku suuruse ja seda iseloomustavate karakteristikutega (jaotusfunktsioon, tihedusfunktsioon, keskväärtus jne.). Käsitlus toimus üldises plaanis, käsitlesime neid omadusi, mis on omased kõikidele juhuslikele suurustele. Käesolevas loengus vaatleme nimelisi juhuslikke suurusi. Nimelt on tõenäosusteoorias defineeritud suur hulk erinevaid jaotusi, mis iseloomustavad teoreetiliste juhuslike suuruste tõenäosuslikku käitumist. Paljud nendest on leidnud rakendust reaalsete protsesside või nähtuste kirjeldamisel. Mitmeid jaotusseadusi kasutatakse abivahendina statistiliste otsuste tegemisel. Klassikalisi jaotusseadusi on nii pidevate kui ka diskreetsete juhuslike suuruste jaoks. Kõigepealt vaatleme diskreetseid juhuslikke suurusi.

Binoomjaotus

Binomiaalne juhuslik suurus tekib sõltumatute katsete korral. Juhuslikuks suuruseks on meid huvitava sündmuse A toimumiste arv. Binoomjaotusega juhuslik suuurus on näiteks märki tabanud laskude arv korduval tulistamisel.

Juhuslikku suurust X, mille võimalike väärtuste hulgaks on naturaalarvud 0, 1, ..., n ja millele vastavad tõenäosused arvutatakse Bernoulli valemiga

$$P(X = m) = P_{m,n} = C_n^m p^m q^{n-m}$$

nimetatakse **binoomjaotusega juhuslikuks suuruseks**.

Asjaolu, et juhuslik suurus X on binoomjaotusega parameetritega n ja p, märgitakse lühidalt $\boldsymbol{X \sim B(n, p)}$.
Binoomjaotusega juhuslik suurus on diskreetne. Seetõttu ei eksisteeri sellel ka tõenäosuse tihedust (tihedusfunktsiooni). Binoomjaotusega suuruse jaotuspolügooni nimetatakse **binoompolügooniks**.

Näide. Laskur tulistab märklauda 6 korda. Tõenäosus tabada märki igal lasul on 0,7. Koostada tabamuste arvu kui juhusliku suuruse jaotustabel ja vastav binoompolügoon.

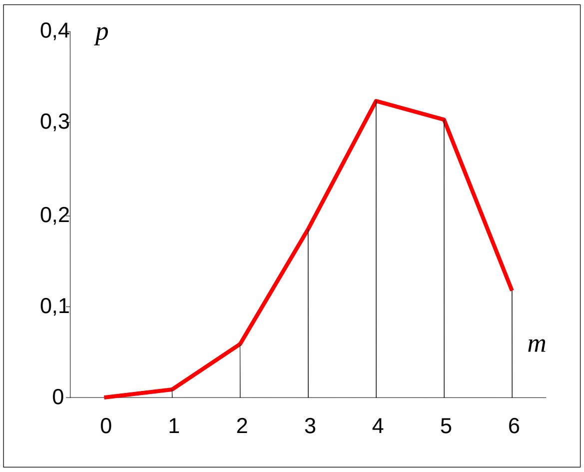
Lahendus. Ülesande kohaselt on tabamuste arv X binoomjaotusega juhuslik suurus, parameetritega n = 6 ja p = 0,7, sest juhusliku suuruse X võimalike väärtuste hulk on naturaalarvud 0, 1, ..., 6 ja üksikväärtuste tõenäosused saame leida Bernoulli valemist:

$$P_{m,6} = C_6^m \cdot 0{,}7^m \cdot 0{,}3^{6-m}\,.$$

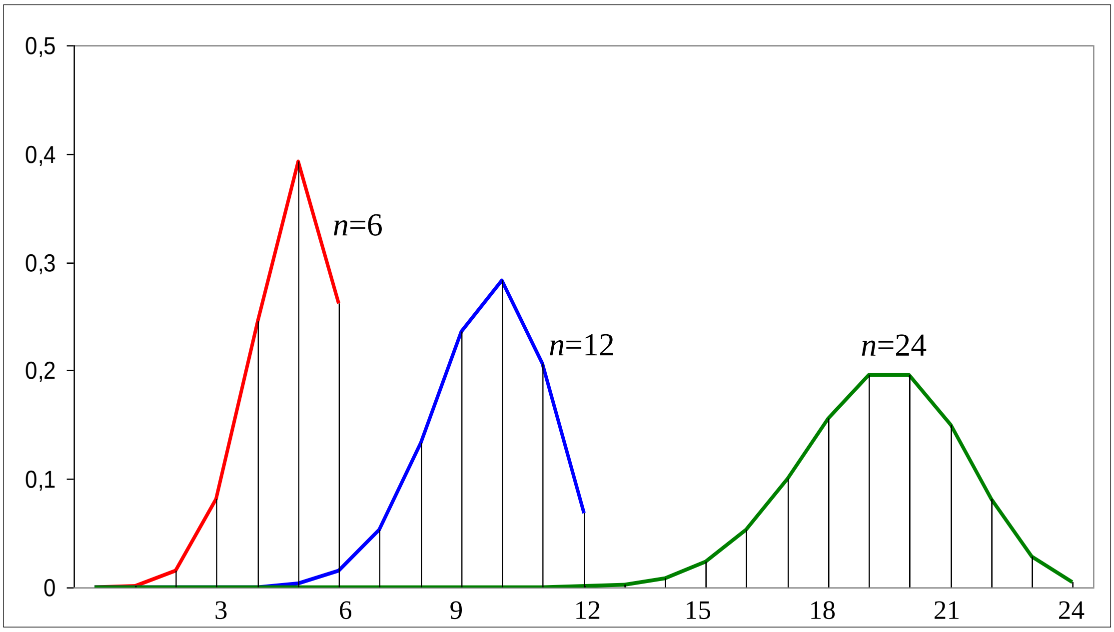
Tehes vastavad arvutused $m = 0,1,\ldots,6$ korral, saame jaotustabeli.

X	0	1	2	3	4	5	6
p	0,0007	0,0102	0,0595	0,1852	0,3241	0,3025	0,1176

Vastav binoompolügoon on näha järgneval joonisel.



Joonistame veel binoompolügoonid juhul kui p = 0,8 kolme erineva n väärtuse jaoks.



Jooniselt võib tähele panna, et:

- Sageduse tõenäosused $P_{m,n}$ üldiselt kasvavad maksimaalse väärtuseni ja siis jälle kahanevad.
- Katsete arvu n suurenemisel sama p korral tõenäosuse $P_{m,n}$ väärtused vähenevad.
- Kui $p = q$, siis jaotus on sümmeetriline iga n korral. Kui $p \neq q$, siis ebasümmeetriline. Aga katsete arvu n suurenemisel läheneb jaotus sümmeetrilisele isegi teineteisest erinevate p ja q korral.

Binoomjaotuse parameetrid

Binoomjaotuse keskväärtus: $EX = np$.

Binoomjaotuse dispersioon: $DX = npq$.

Binoomjaotuse standardhälve: $\sigma(X) = \sqrt{DX} = \sqrt{npq}$.

Järgnevalt veendume, et see tõesti nii on.

Binoomjaotusega juhusliku suuruse keskväärtus

Olgu juhuslikuks suuruseks X sündmuse A toimumiste arv n ühesugusel ja sõltumatul katsel. Kui X_1 on sündmuse toimumiste arv esimesel katsel, X_2 teisel katsel jne., siis üldine sündmuse toimumiste arv on

$$X = X_1 + X_2 + \ldots + X_n .$$

Ja vastav keskväärtus

$$EX = EX_1 + EX_2 + \ldots + EX_n .$$

Vaatleme kõigepealt suurust X_1:

$$X_1 = \begin{cases} 1, & \text{kui } A \text{ toimub (tõenäosusega } p) \\ 0, & \text{kui } A \text{ ei toimu (tõenäosusega } 1-p) \end{cases}$$

Seega X_1 jaotusrida on

x_i	1	0
p_i	p	$1 - p$

Arvestades valemit $EX = \sum_{i=1}^{n} x_i p_i$ saame $EX_1 = 1 \cdot p + 0 \cdot (1-p) = p$. Arvestades nüüd, et $EX = EX_1 + EX_2 + \ldots + EX_n$ saame $EX = np$.

Binoomjaotusega juhusliku suuruse dispersioon

Kuna $X_1, X_2, \ldots, X_n$ on sõltumatud, siis

$$DX = DX_1 + DX_2 + \ldots + DX_n .$$

Vaatleme kõigepealt suurust X_1. Arvestades valemit $DX = \sum_{i=1}^{n} (x_i - EX)^2 \cdot p_i$ saame

$$DX_1 = (1-p)^2 \cdot p + (0-p)^2 \cdot q = q^2 p + p^2 q = pq(q+p) = pq$$

Arvestades nüüd, et $DX = DX_1 + DX_2 + \ldots + DX_n$ saame $DX = npq$.

Poissoni jaotus

Kui Bernoulli valemis on n suur, on tema kasutamine komplitseeritud. Seetõttu kasutatakse sageli ka binoomjaotuse piirjaotusi, millest üheks on **Poissoni jaotus** (harva esinevate sündmuste jaotusseadus).

Poissoni jaotus on binoomjaotuse piirjaotuseks sel juhul, kui katseseeria pikkus $n \to \infty$, tõenäosus $p \to 0$ selliselt, et korrutis $\lambda = np$ püsib konstantsena (läheneb konstandile).

Osutub, et sel korral

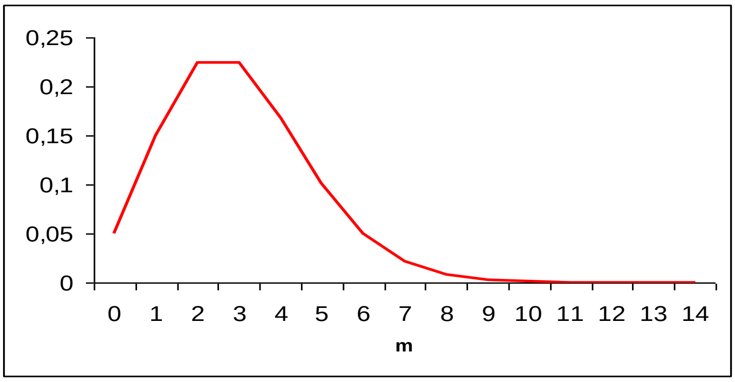
$$\lim_{n\to\infty} P(X = m) = \lim_{n\to\infty} C_n^m p^m q^{n-m} = \frac{\lambda^m}{m!} e^{-\lambda} .$$

Juhuslikku suurust, mille võimalike väärtuste hulgaks on täisarvud 0, 1, 2, ... ja mille jaotus on määratud valemiga

$$P(X = m) = \frac{\lambda^m}{m!} e^{-\lambda}$$

nimetatakse **Poissoni jaotusega juhuslikuks suuruseks**.

Sümboolselt tähistatakse asjaolu, et juhuslik suurus on Poissoni jaotusega: ***X* ~ *P*(λ)**. Poissoni jaotuse jaotuspolügooni põhimõtteline kuju on esitatud järgneval joonisel.



Poissoni jaotusega juhuslik suurus esineb tegelikus elus väga sageli oma tekkemehhanismi tõttu. Poissoni jaotusega on näiteks:

- radioaktiivses aines lagunevate aatomite arv ajaühikus;
- kosmoselaeva tabavate meteoriitide arv;
- igat liiki õnnetuste (liiklusõnnetused, tulekahjud, tööõnnetused, masinate rikked) toimumise arv ajaühikus, kui see ajaühik on küllalt lühike (näiteks päev).

Poissoni jaotuse parameetrid

Poissoni jaotuse keskväärtus: $E(X) = \lambda = np$.

Poissoni jaotuse dispersioon: $D(X) = \lambda = np$.

Poissoni jaotuse standardhälve: $\sigma(X) = \sqrt{\lambda}$.

Jämeda hinnanguna on Poissoni jaotuse kasutamine õigustatud, kui $np \le 5$ ja $n \ge 30$.

Näide. Aparatuur koosneb 1000 elemendist, kusjuures elemendid töötavad üksteisest sõltumatult. Tõenäosus elemendi riknemiseks ajavahemiku T jooksul olgu kõikide elementide jaoks ühesugune ja võrdne 0,0005-ga. Leida tõenäosus, et ajavahemiku T jooksul rikneb mitte rohkem kui 2 elementi.

Lahendus. Tehtud eeldustel on ajavahemiku T jooksul riknenud elementide arv X binoomjaotusega,

$$X \sim B(1000;\, 0{,}0005).$$

Kuna $n \ge 30$ ja $np = 1000 \cdot 0{,}0005 = 0{,}5 \le 5$, siis on täidetud ka Poissoni jaotuse kasutamise eeldused.

Arvutame Poissoni jaotuse parameetri

$$\lambda = np = 1000 \cdot 0{,}0005 = 0{,}5$$

ja küsitud tõenäosuse

$$P(X \le 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{0{,}5^0}{0!}e^{-0{,}5} + \frac{0{,}5^1}{1!}e^{-0{,}5} + \frac{0{,}5^2}{2!}e^{-0{,}5} \approx 0{,}986$$

Ühtlane jaotus

Praktikas esineb mõnikord olukordi, kus juhusliku suuruse kohta on teada vaid see, et tema väärtused paiknevad mingis vahemikus, kuid midagi lähemat selle juhusliku suuruse kohta teada ei ole. Näiteks oletame, et läheme juhuslikul ajamomendil bussipeatusse, teades vaid, et sellel liinil liiguvad bussid 10 minutiliste vahedega. Sellisel juhul tuleb meil bussi oodata 0 kuni 10 minutit, kusjuures mingit ooteaega ei saa ühestki teisest tõenäosemaks lugeda. Kirjeldatud olukorda iseloomustab ühtlane jaotus.

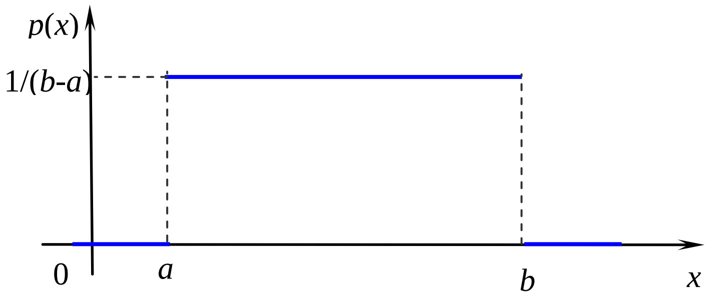
Pidev juhuslik suurus X on **ühtlase jaotusega** lõigul [a; b], kui sellel lõigul on tema tihedusfunktsioon konstantne ja väljapool seda lõiku võrdne nulliga:

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

Ühtlast juhuslikku suurust kasutatakse ka keerulisemate juhuslike suuruste lähendamisel. Nii on võimalik mitme erineva ühtlase juhusliku suuruse abil kirjeldada väga keeruka tihedusfunktsiooniga juhuslikke suurusi. Looduslikus keskkonnas esineb ühtlane juhuslik suurus väga harva. Ühtlase jaotusega juhuslikuks suuruseks on näiteks trammi ooteaeg, ooteaeg valgusfoori taga ning mõõteriista näidu ümardamise viga skaala täisjaotuseni.

Asjaolu, et juhuslik suurus X on ühtlase jaotusega lõigul [a; b], tähistatakse sümboolselt $\boldsymbol{X \sim U(a; b)}$.

Järgnevalt on esitatud ühtlase jaotuse tihedusfunktsiooni graafik.



Ühtlase jaotuse parameetrid

Ühtlase jaotusega juhusliku suuruse keskväärtus:

$$EX = \int_{-\infty}^{\infty} x \cdot p(x)\,dx = \int_a^b \frac{x}{b-a}\,dx = \frac{a+b}{2}.$$

Ühtlase jaotusega juhusliku suuruse dispersioon:

$$DX = E(X^2) - (EX)^2 = \int_a^b \frac{x^2}{b-a}\,dx - \left(\frac{b+a}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

Ühtlase jaotusega juhusliku suuruse standardhälve:

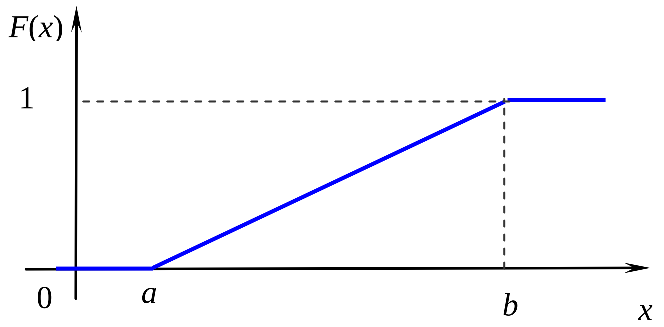
$$\sigma(X) = \sqrt{DX} = \frac{b-a}{2\sqrt{3}}.$$

Ühtlase jaotuse jaotusfunktsioon

Ühtlase jaotuse jaotusfunktsioon avaldub (jaotusfunktsiooni avaldise tuletasime peatükis „Juhuslikud suurused”):

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

Jaotusfunktsiooni graafik on toodud järgneval joonisel.



Tõenäosus, et juhusliku suuruse X väärtus on vahemikus $(\alpha,\beta)\subset[a,b]$ avaldub

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha) = \frac{\beta-\alpha}{b-a}$$

Näide. Tramm liigub regulaarselt intervalliga 2 minutit. Reisija astub ooteplatvormile juhuslikul ajamomendil, mis ei ole seotud trammide liikumisgraafikuga. Leida reisija ooteaja jaotustihedus, keskväärtus, dispersioon ja standardhälve. Leida tõenäosus, et reisijal ei tule oodata rohkem kui 0,5 minutit.

Lahendus. Tihedusfunktsioon (jaotustihedus)

$$p(x) = \begin{cases} 0, & x < 0, \\ \dfrac{1}{b-a} = \dfrac{1}{2-0} = \dfrac{1}{2}, & 0 \le x \le 2, \\ 0, & x > 2. \end{cases}$$

Vastavad arvkarakteristikud:

$$EX = \frac{a+b}{2} = \frac{0+2}{2} = 1$$

$$DX = \frac{(b-a)^2}{12} = \frac{(2-0)^2}{12} = \frac{4}{12} = \frac{1}{3}$$

$$\sigma(X) = \sqrt{DX} = \sqrt{\frac{1}{3}} \approx 0{,}58$$

Tõenäosus, et reisijal ei tule oodata üle poole minuti

$P(X < 0{,}5) = F(0{,}5) = \dfrac{0{,}5}{2} = \dfrac{1}{4}$.

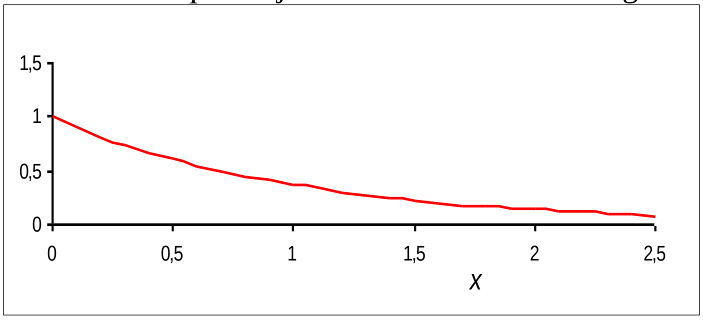
Eksponentjaotus

Eksponentjaotust kasutatakse enamasti massiteeninduse teooria reaalsete teenindussüsteemide teenindusaja modelleerimisel. Juhuslikuks suuruseks X on teenindusaja pikkus. Jaotusfunktsioon $P(x) = P(X < x)$ näitab, millise tõenäosusega kestab teenindamine vähem kui x ajaühikut. Parameeter λ näitab, palju "kliente" teenindatakse keskmiselt ühe ajaühiku kohta (s.t. teenindamise kiirus). Selle pöördväärtus $1/\lambda$ on ühe tellimuse keskmine teenindusaeg. Eksponentjaotus sobib ka süsteemi tõrketa tööaja kirjeldamiseks, eeldusel, et tõrgete intensiivsus ajas on muutumatu.

Öeldakse, et juhuslik suurus X on **eksponentjaotusega**, kui tema tihedusfunktsiooniks on

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

Asjaolu, et juhuslik suurus X on eksponentjaotusega, tähistatakse sümboolselt $\boldsymbol{X \sim Exp(\lambda)}$. Järgnevalt on esitatud eksponentjaotuse tihedusfunktsiooni graafik, kui $\lambda = 1$.



Eksponentjaotuse jaotusfunktsioon

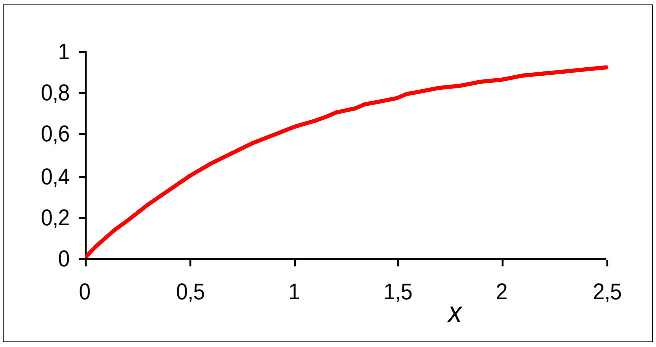
Leiame eksponentjaotusega juhusliku suuruse jaotusfunktsiooni tihedusfunktsiooni abil. Kui $x < 0$, siis $F(x) = 0$. Kui $x \geq 0$, siis

$$F(x) = \int_{-\infty}^{x} p(x)\,dx = \int_{-\infty}^{0} 0dx + \int_{0}^{x} \lambda e^{-\lambda x}\,dx = \lambda \int_{0}^{x} e^{-\lambda x}\,dx = -e^{-\lambda x}\Big|_0^x = 1 - e^{-\lambda x}.$$

Seega

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Jaotusfunktsiooni graafik $\lambda = 1$ korral on kujutatud järgneval joonisel.



Eksponentjaotuse arvkarakteristikud

Keskväärtus $EX = \dfrac{1}{\lambda}$.

Dispersioon $DX = \dfrac{1}{\lambda^2}$.

Standardhälve $\sigma(X) = \sqrt{DX} = \dfrac{1}{\lambda}$.

<u>Näide.</u> Protsessori keskmine tööiga on 1000 tundi. Millise tõenäosusega töötab protsessor tõrgeteta vähemalt 1500 tundi?

<u>Lahendus.</u>
X – protsessori tõrketa tööiga
$X \sim Exp(\lambda)$
Kuna $EX = \frac{1}{\lambda} = 1000$, siis $\lambda = 1/1000$.
Otsitav tõenäosus avaldub nüüd jaotusfunktsiooni kaudu järgnevalt:

$$P(X \geq 1500) = 1 - P(X < 1500) = 1 - F(1500) = 1 - (1 - e^{-\frac{1}{1000} \cdot 1500}) = e^{-1,5}.$$

Weibulli jaotus

Kui pideva juhusliku suuruse tihedusfunktsiooniks on funktsioon

$$p(x) = \begin{cases} \frac{\delta}{\eta}\left(\frac{x - x_0}{\eta}\right)^{\delta - 1} \cdot e^{(-\frac{x - x_0}{\eta})^{\delta}}, & x > x_0; \\ 0, & x \leq x_0, \end{cases}$$

siis öeldakse, et see suurus on **Weibulli jaotusega**.
Weibulli jaotust kasutatakse toodete ja konstruktsioonide tööea ning usaldatavuse analüüsis. Parameeter δ iseloomustab jaotuse kuju, η mastaapi ning x_0 asukohta x-teljel. Eksponentjaotus on Weibulli jaotuse erijuht, kui $x_0 = 0$, $\delta = 1$, $\eta = \frac{1}{\lambda}$.