

Dispersioon

Keskväärtuse ja dispersiooni
omadused

Pidevad juhuslikud suurused.
Tihedusfunktsioon.

Dispersioonist ja standardhälbest

diskreetse juhusliku
suuruse näitel

Juhusliku suuruse funktsiooni keskväärtus

Olgu uuritava tunnuse X jaotus järgmine:

x	1	2	5	
$P(X=x)$	0,1	0,2	0,7	$E(X)=\dots$

siis juhusliku suuruse X^2 jaotus on

x	1	4	25
$P(X^2=x)$	0,1	0,2	0,7

ja juhusliku suuruse X^2 keskväärtus $E(X^2)$ on

$$\begin{aligned} E(X^2) &= 1 \cdot 0,1 + 4 \cdot 0,2 + 25 \cdot 0,7 \\ &= 18,4 \end{aligned}$$

Dispersioon $D(X) = E[(X - EX)^2]$

Olgu uuritava tunnuse X jaotus järgmine:

x	1	2	5	
$P(X=x)$	0,1	0,2	0,7	$E(X)=4$

siis juhusliku suuruse $X-E(X)$ jaotus on

x	-3	-2	1
$P(X-E(X)=x)$	0,1	0,2	0,7

Dispersioon $D(X) = E[(X - EX)^2]$

Olgu uuritava tunnuse X jaotus järgmine:

x	1	2	5	
$P(X=x)$	0,1	0,2	0,7	$E(X)=4$

siis juhusliku suuruse $(X-E(X))^2$ jaotus on

x	9	4	1
$P((X-E(X))^2 = x)$	0,1	0,2	0,7

ja juhusliku suuruse $(X-E(X))^2$ keskvaartus ehk dispersioon on

$$\begin{aligned} D(X) &= E(X-E(X))^2 \\ &= 9 \cdot 0,1 + 4 \cdot 0,2 + 1 \cdot 0,7 = 2,4 \end{aligned}$$

Dispersioon $D(X) = E[(X - EX)^2]$

inglise keeles: *Variance*
vene keeles: *Дисперсия*

Tähistustest:

inglisekeelses kirjanduses kasutatakse sageli tähistusi ***Var(X)*** või ***V(X)***.

Vene ja eesti keeles on kasutusel tähistus ***D(X)***

Lisaks kasutatakse: σ^2 , näiteks juhusliku suuruse X dispersiooni tähistatakse sageli kujul

$$\sigma_X^2 \text{ või } \sigma^2(X).$$

Standardhälve

Juhusliku suuruse X standardhälve σ on ruutjuur dispersioonist:

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

Tõenäosus, et juhusliku suuruse väärtus jääb kahe standardhälbe kaugusele keskvaartusest on *vähemalt* $\frac{3}{4} = 0,75$ (enamasti $\sim 0,95$).

Tõenäosus, et juhusliku suuruse väärtus jääb kahe standardhälbe kaugusele keskvaartusest on *vähemalt* $\frac{8}{9} = 0,88\dots$ (enamasti $\sim 0,997$).

Keskvaartuse ja dispersiooni omadused

$$1. E(cX) = cE(X)$$

c – konstant;
 X, Y – juhuslikud suurused

$$2. E(c+X) = c+E(X)$$

$$3. E(X+Y) = E(X)+E(Y)$$

Aga enamasti
 $E(f(X)) \neq f(E(X))$!!

$$4. E(c)=c$$

$$1. D(c+X) = D(X)$$

$$2. D(cX) = c^2 D(X)$$

$$3. D(X+Y) = D(X) + D(Y), \text{ kui } X \text{ ja } Y \text{ on sõltumatud}$$

$$4. D(c) = 0$$

Pideva tunnuse jaotus

Problem

Pideva tunnuse jaotust ei saa (täpselt) esitada tõenäosusfunktsiooni abil.

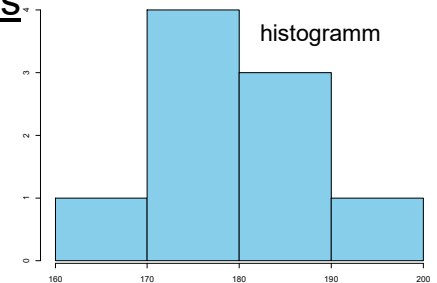
$P(X = x)$	x
0	170,0000000000000000001 ← ?
0	170,0000000000000000002
0	170,0000000000000000003
....	

Pideva tunnuse jaotus

Praktikas vaatluste korral:

172,122392...	165,45...	178,09..
197,998783...	177,56...	181,12..
183,6778..	179,001...	180,01..

<u>vahemik</u>	<u>sagedus</u>
(160..170]	1
(170..180]	4
(180..190]	3
(190..200]	1



Pideva tunnuse jaotus

Probleem

Pideva tunnuse jaotust ei saa (täpselt) esitada tõenäosusfunktsiooni abil.

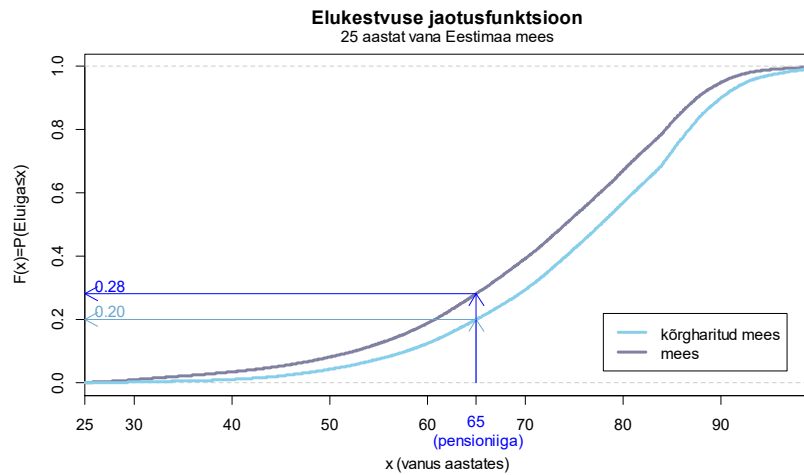
Võimalikud alternatiivid:

- kumulatiivne jaotusfunktsioon;
- tihedusfunktsioon.

Kumulatiivne jaotusfunktsioon

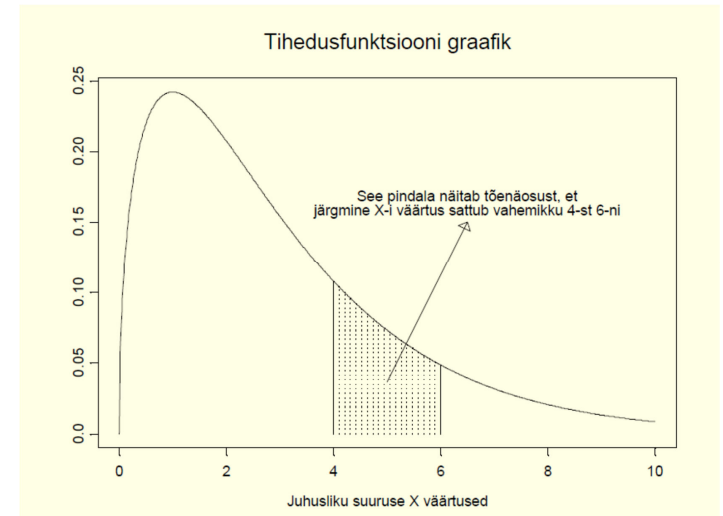
Kumulatiivne jaotusfunktsioon $F(t)$ näitab mistahes väärtuse t korral tõenäosust, et juhuslik suurus X omandab väärtuse, mis on võrdne või väiksem kui t :

$$F(t) := P(X \leq t)$$

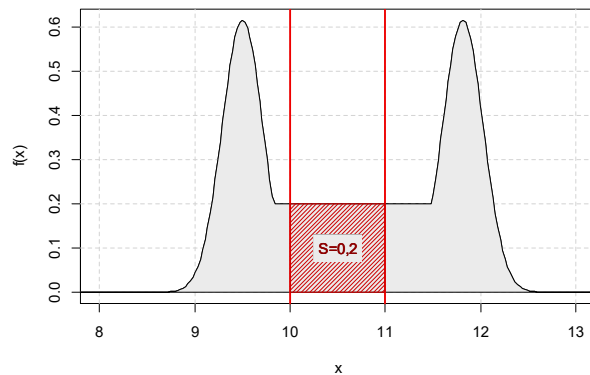


Tihedusfunktsioon

Tihedusfunktsiooni graafiku alune pindala näitab tõenäosust, et juhusliku suuruse väärtus sattub vaadeldavasse vahemikku.



Tihedusfunktsioon



20% uuritava tunnuse väärtustest jääb vahemikku 10..11 (üks juhuslikult valitud vaatlus sattub sinna vahemikku tõenäosusega 0,2)

Tihedusfunktsioon

