Инжењерска статистика
Вежба 2

Статистике

Нумеричке мере које описују карактеристике узорка, односно популације, називају се статистике и деле се на:

- Мере локације података
  - Аритметичка средина
  - Медијана
  - Мода
- Мере варијације података
  - Ранг
  - Варијанса
  - Стандардна девијација

Пример

- Нека су дата два скупа података
  А: 130, 140, 145, 150, 158, 165
  Б: 90, 128, 140, 160, 165, 205

Аритметичка средина свих мерења:

$$\bar{x} = \frac{1}{12}(130 + 140 + 145 + \ldots + 165 + 205) = 148$$

Медијана: 90, 128, 130, 140, 140, 145, 150, 158, 160, 165, 165, 205

$$\mathrm{Me} = \frac{145 + 150}{2} = 147{,}5$$

Мода: 90, 128, 130, 140, 140, 145, 150, 158, 160, 165, 165, 205

2x140, 2x165

Пример

А: 130, 140, 145, 150, 158, 165

Б: 90, 128, 140, 160, 165, 205

Ранг: $R_A = \max(x_i) - \min(x_i) = x_5 - x_1 = 165 - 130 = 35$

$R_B = 205 - 90 = 115$

Варијанса:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} \qquad n < 30$$

$$s_A^2 = \frac{1}{6-1}\left[(130-148)^2 + \ldots + (140-148)^2\right] = 158$$

$$s_B^2 = \frac{1}{6-1}\left[(90-148)^2 + \ldots + (160-148)^2\right] = 1502$$

Стандардна девијација: $S_A = \sqrt{158} = 12.57$ $\qquad S_B = \sqrt{1502} = 38.76$

	sredina	varijansa	st.devijacija
uzorak	$\bar{x}$	S^2	S
populacija	μ	σ^2	σ

Случајна променљива

- Варијабилност се описује функцијом која се назива случајна променљива

$$X = \mu + e$$

μ - је константа
e - случајно одступање константе

- Расподела променљиве X може бити:
  - Дискретна
  - Континуална

Дискретне случајне променљиве

Дискретна случајна променљива *X* може имати само одређене вредности.

Расподела случајне променљиве *X* дефинише се функцијом расподеле.

Функција расподеле је **опис вероватноћа** које се везују за могуће вредности случајне променљиве *X*.

Особине функције расподеле дискретне случајне величине су:

$$F(X) = P(X = x)$$

$$F(X) \geq 0$$

$$\sum_{i=1}^{n} F(X_i) = 1$$

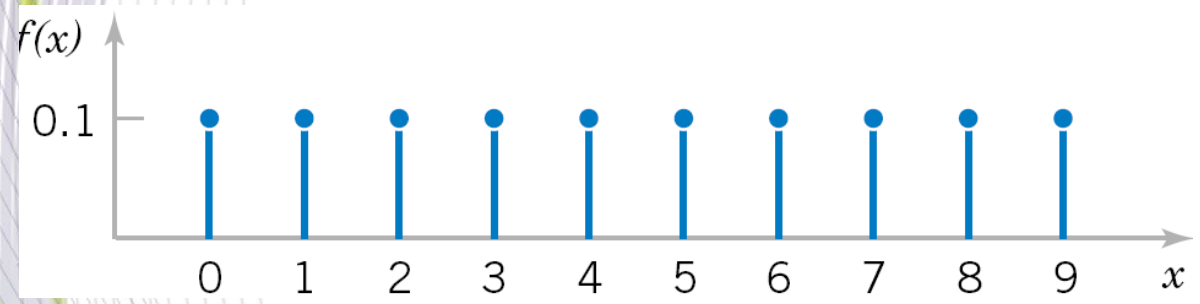
- Средина очекиване вредности случајне променљиве *X* је:

$$\mu = E(X) = \sum X \cdot F(X)$$

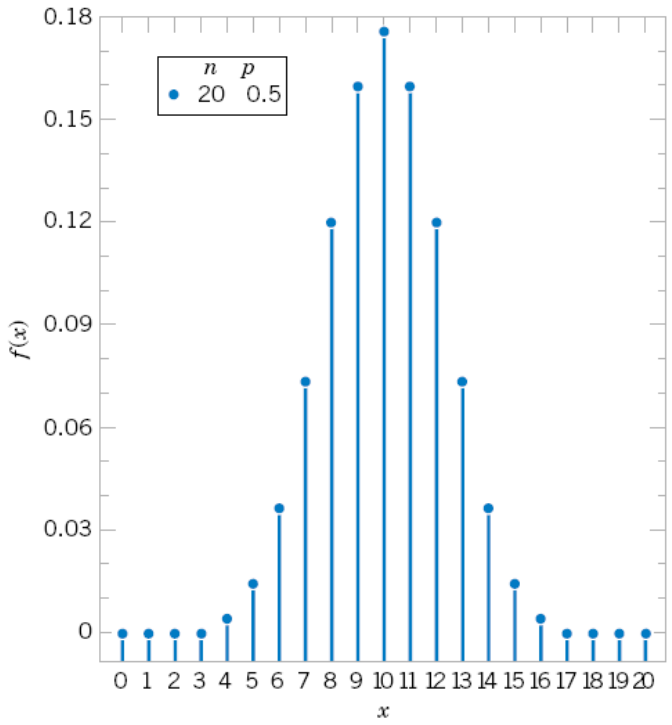
- Варијанса случајне променљиве *X* је:

$$\sigma^2 = D(X) = \sum (X - \mu)^2 \cdot F(X)$$

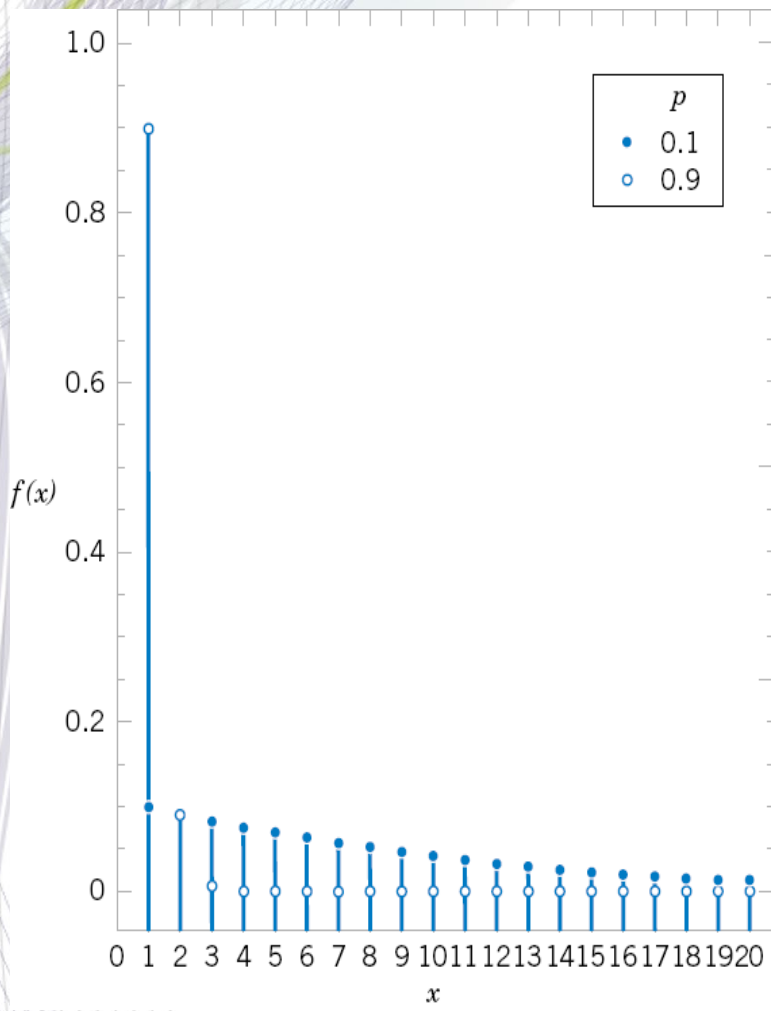
Примери дискретних расподела:



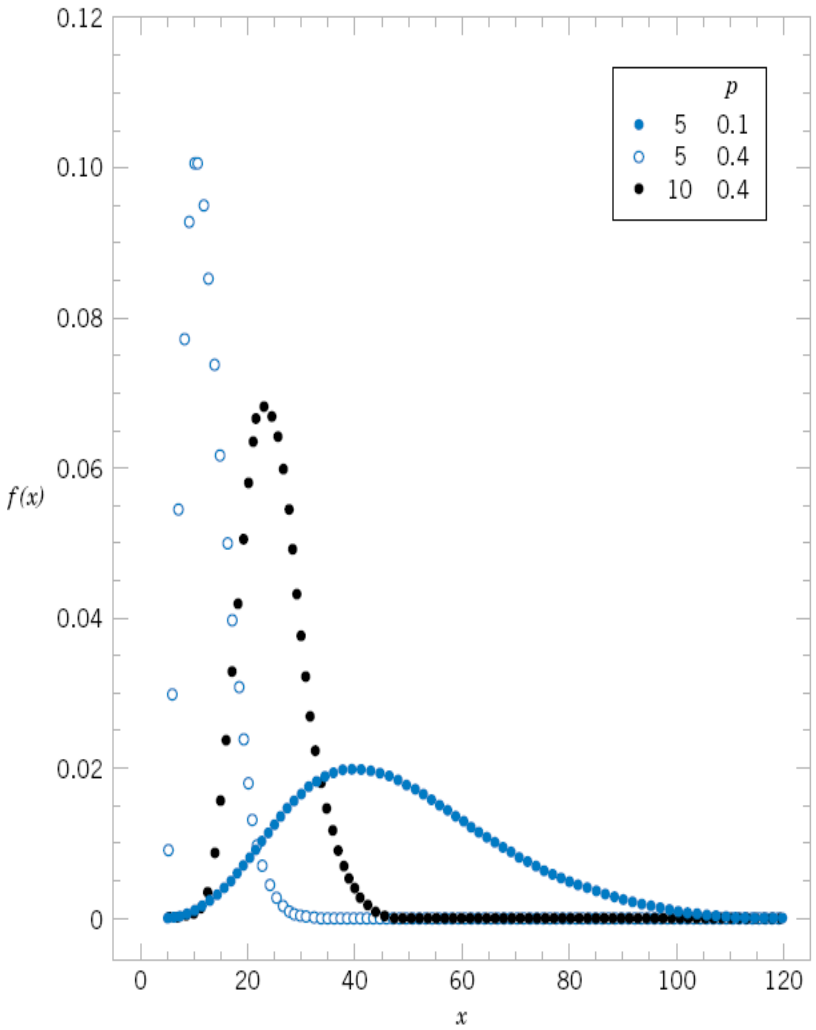
Униформна дискретна расподела



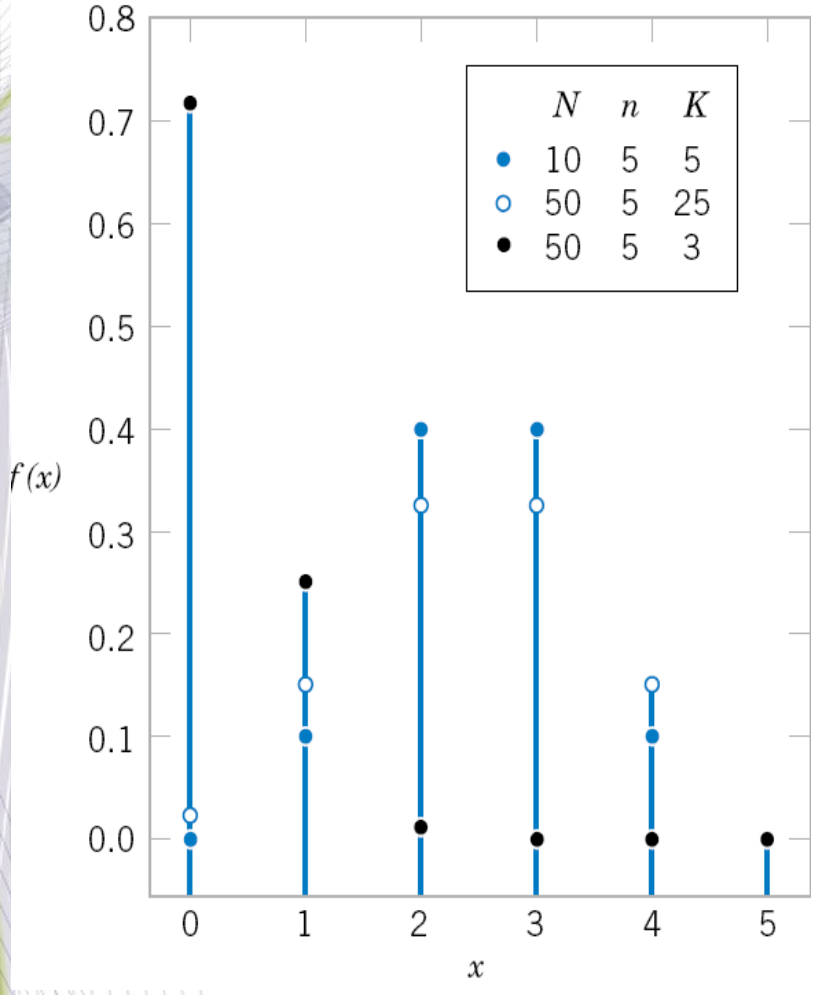
Биномна расподела



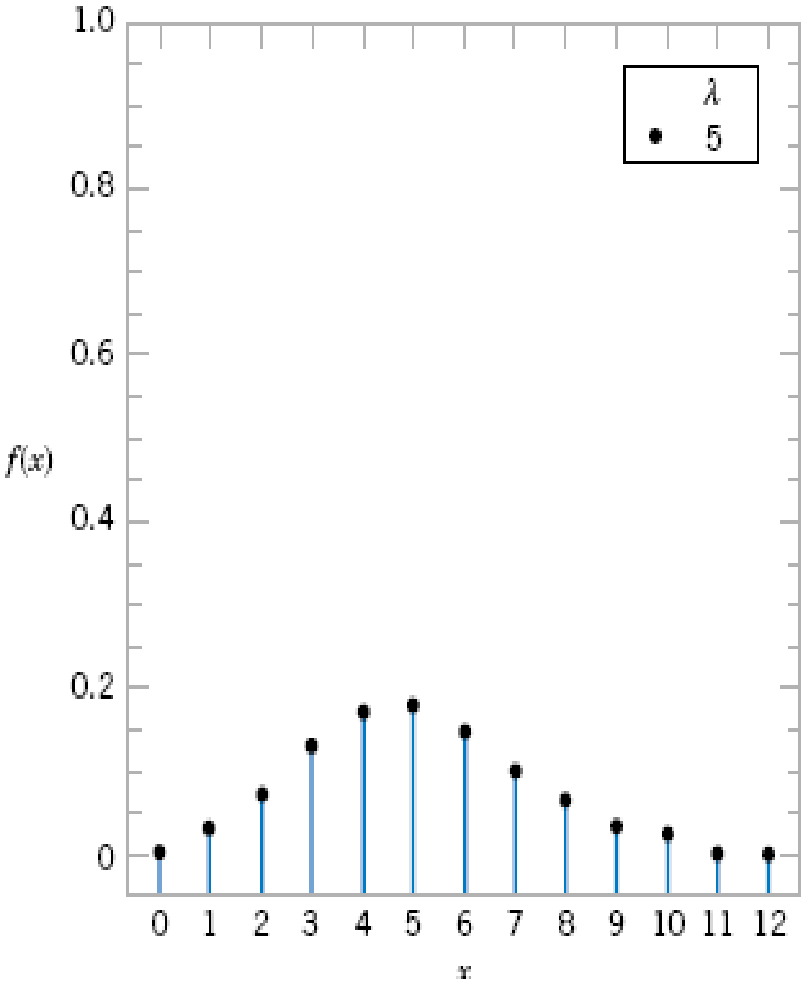
Геометријска



Негативна биномна



Хипергеометријска



Poasonova

Задатак 1

Оцењују се квалитет боје и скупљање изливене пластике која се добија при новом производном процесу. Испитивања су показала следеће информације које су дате у табели, а потребно је испитати расподелу:

Боја	Скупљање	p	x
одговара	одговара	0.54	2
одговара	не одговара	0.20	1
не одговара	одговара	0.20	1
не одговара	не одговара	0.06	0

Функција расподеле појединих вредности случајне променљиве *X* је:

$$x : \begin{pmatrix} 0 & 1 & 2 \\ 0.06 & 0.40 & 0.54 \end{pmatrix}$$

За дати пример, важе све особине функција расподеле:

$$F(X) = P(X = x) \Rightarrow$$

$$F(x_1) = F(0) = P(x = 0) = 0.06$$

$$F(x_2) = F(1) = P(x = 1) = 0.40$$

$$F(x_3) = F(2) = P(x = 2) = 0.54$$

$$F(X) \geq 0 \Rightarrow F(x_1) = 0.06 > 0;\ F(x_2) = 0.40 > 0;\ F(x_3) = 0.54$$

$$\sum_{i=1}^{n} F(X_i) = 1 \Rightarrow F(x_1) + F(x_2) + F(x_3) = 0.06 + 0.40 + 0.54 = 1$$

$$\mu = \sum XF(X)$$

$$\bar{x} = 0 \cdot 0.06 + 1 \cdot 0.40 + 2 \cdot 0.54 = 0.40 + 1.08 = 1.48$$

$$\sigma^2 = \sum (X - \mu)^2 F(X)$$

$$s^2 = \left[(0 - 1.48)^2 \cdot 0.06 + (1 - 1.48)^2 \cdot 0.40 + (2 - 1.48)^2 \cdot 0.54 \right] =$$

$$= 0.131424 + 0.09216 + 0.146016 = 0.3696$$

$$s = \sqrt{s^2} = \sqrt{0.3696} = 0.608$$

Задатак 2

Постоји вероватноћа од 0.2 да ће настати грешка при преносу 1 податка кроз дигитални преносни кабл. Претпоставити да су испитивања тачности преноса независна. Случајна променљива *X* је дефинисана као број нетачно пренесених информација у наредна 3 преноса. Испитати особине расподеле.

Излаз	*x*
TTT	0
TTN	1
TNT	1
NTT	1
TNN	2
NTN	2
NNT	2
NNN	3

Имамо 4 могуће вредности случајне променљиве *x* и оне су 0, 1, 2 и 3.
Вероватноћа да ће податак бити тачно пренесен је 0.8, а мерења су међусобно независна.
Вероватноће појављивања сваке од вредности случајне променљиве су:

$$P(x=0)=P(TTT)=P(T)\cdot P(T)\cdot P(T)=0.8^3=0.512$$

$$\begin{aligned}P(x=1)&=P(TTN)+P(TNT)+P(NTT)=\\&=3\cdot P(T)\cdot P(T)\cdot P(N)=\\&=3\cdot 0.8^2\cdot 0.2=0.384\end{aligned}$$

$$\begin{aligned}P(x=2)&=P(TNN)+P(NTN)+P(NNT)=\\&=3P(T)\cdot P(N)\cdot P(N)=\\&=3\cdot 0.8\cdot 0.2^2=0.096\end{aligned}$$

$$P(x=3)=P(NNN)=P(N)\cdot P(N)\cdot P(N)=0.2^3=0.008$$

$$\forall F(X) \geq 0 \quad \wedge \quad F(x_1)+F(x_2)+F(x_3)+F(x_4)=$$
$$=0.512+0.384+0.096+0.008=1$$

Вероватноћа да ће податак бити погрешно пренесен је *p*=0.2
Тада је вероцатноћа да ће бити исправно пренесен (1-*p*)=0.8.

Функција расподеле је:

$$F(x)=\binom{4}{x}\cdot p^x\cdot(1-p)^{4-x}$$

Binomna raspodela

$$\binom{n}{r}=\frac{n!}{r!(n-r)!}$$

$$\bar{x}=np=4\cdot 0.2=0.8$$

$$s^2=np(1-p)=4\cdot 0.2\cdot 0.8=0.64\Rightarrow$$

$$s=\sqrt{s^2}=\sqrt{0.64}=0.8$$

Континуалне случајне променљиве

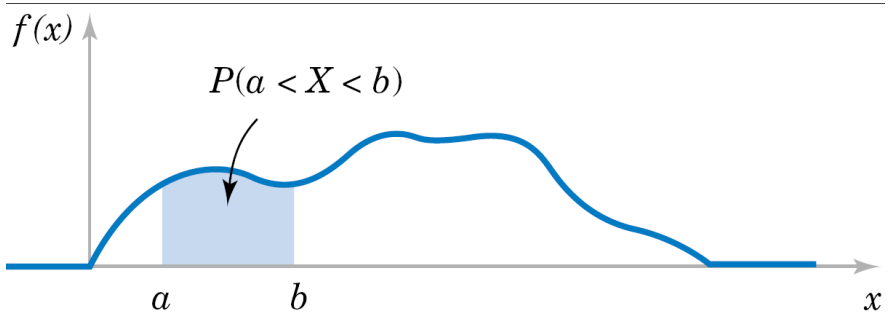
Код континуалних променљувих, број вредности које може имати случајна променљива *X* је бесконачан. *X* може обухватити све вредности реалних бројева, па се може сматрати континуалном.

Вероватноћа појављивања случајне променљиве *X* која је континуална величина, има функцију густине расподеле, која се налази у неком интервалу[*x1,x2*] реалних бројева са особинама:

$$F(X) \geq 0$$

$$\int_{-\infty}^{\infty} F(X)\,dX = 1$$

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} F(u)\,du$$



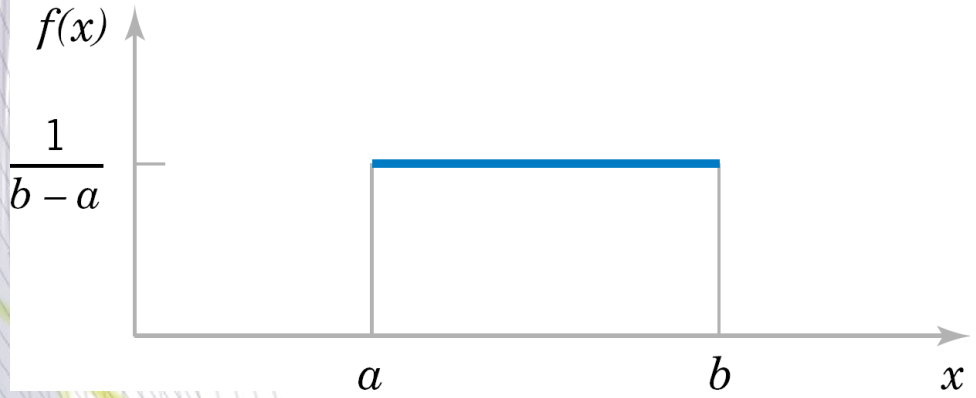
- Аритметичка средина очекиване вредности случајне променљиве *X* је:

$$\mu = E(X) = \int_{-\infty}^{\infty} X \cdot F(X)\,dX$$

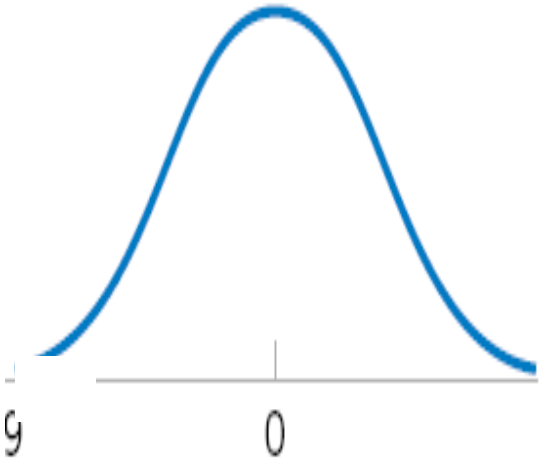
- Варијанса случајне променљиве *X* је:

$$\sigma^2 = D(X) = \int_{-\infty}^{\infty} (X-\mu)^2 \cdot F(X)\,dX$$

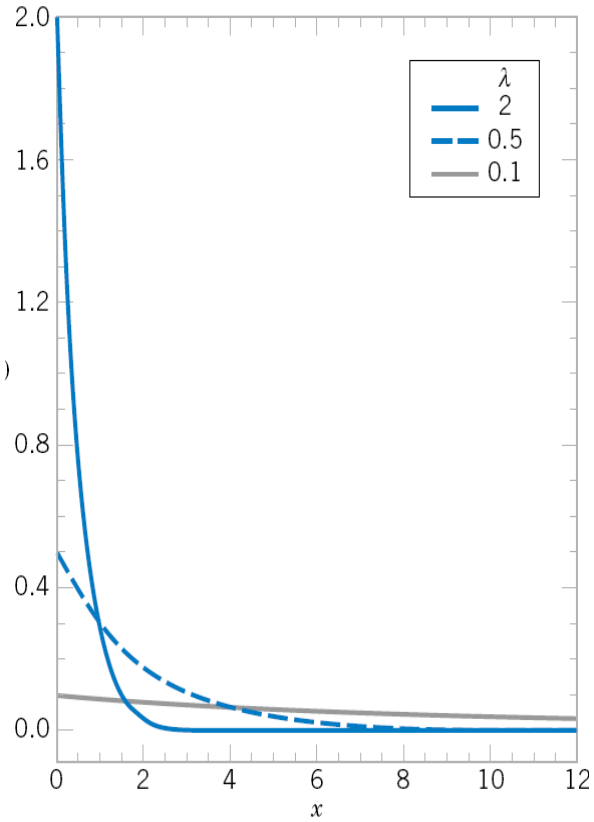
Примери дискретних расподела:



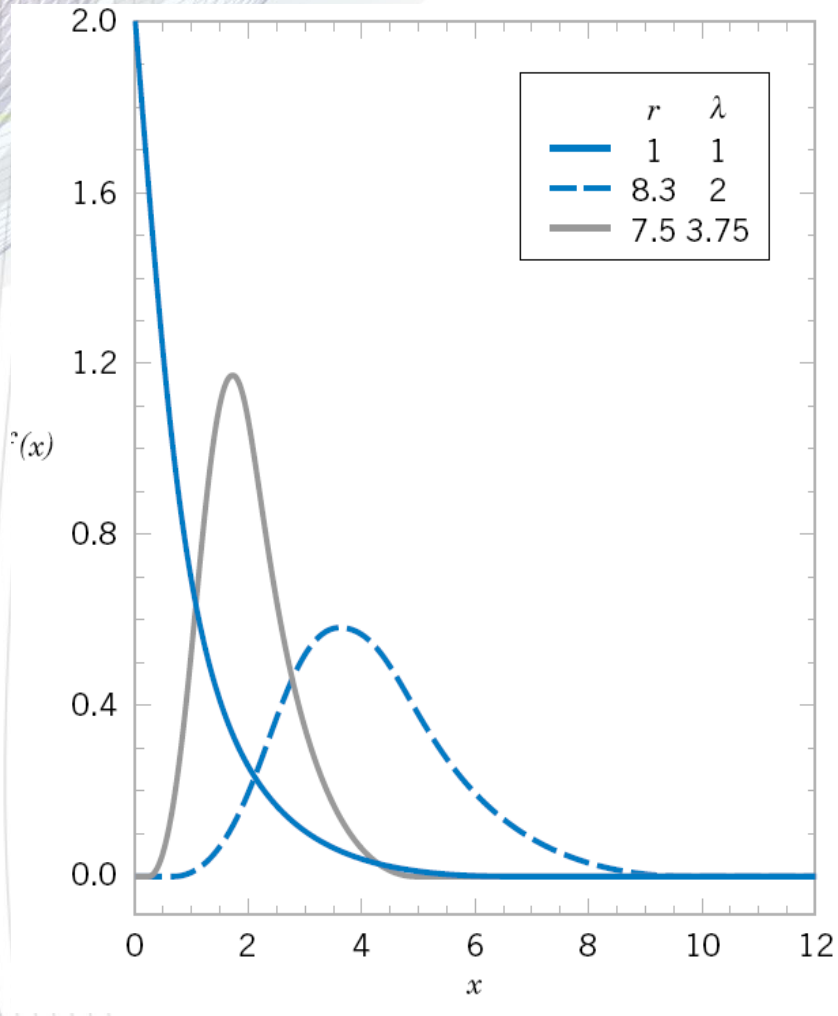
Униформна



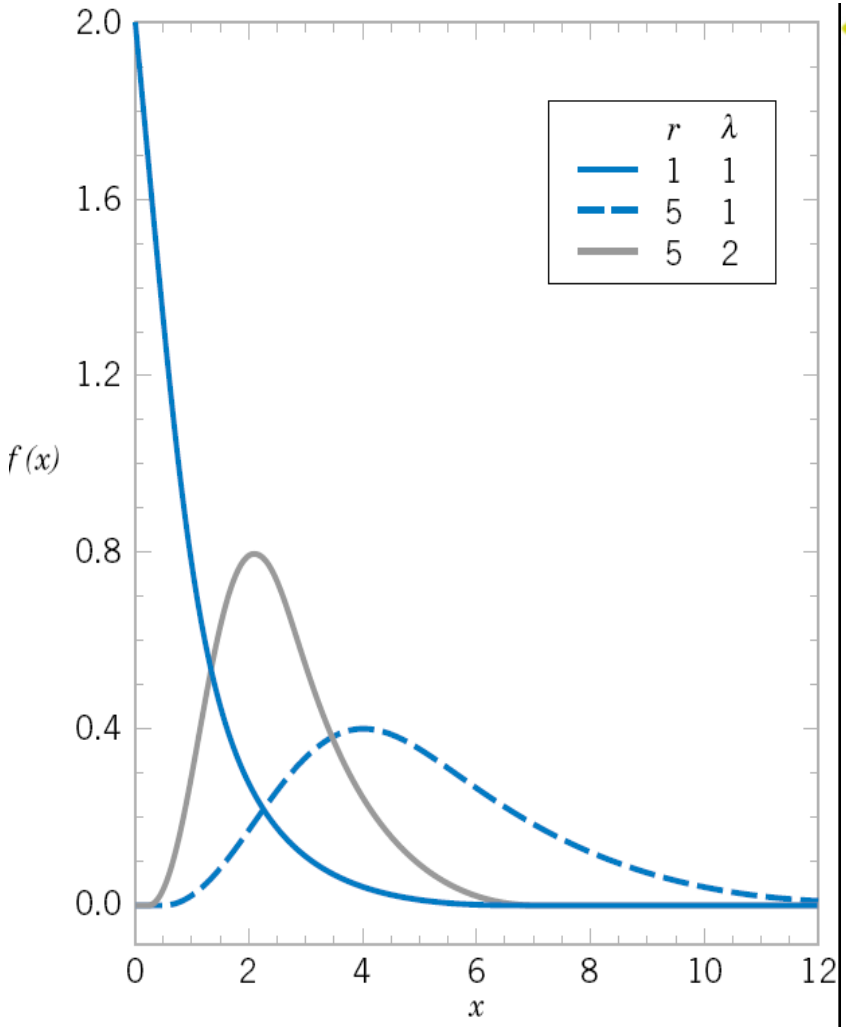
Нормална



Експоненцијална



Гама



Erlangova

Задатак 3

Нека је *X* случајна променљива, која описује континуалну величину на интервалу од [0, 40]. Нека се, даље, претпостави да је теоријска функција по којој се *X* распоређује: $F(x) = 0.025$

Испитати особине функције и одредити вероватноћу да се променљива нађе у области мањој од 10, и у области мањој од 10, а већој од 5.

$$\bar{x} = E(x) = \int_0^{40} xF(x)\,dx = 0.025 \left.\frac{x^2}{2}\right|_0^{40} = 20$$

$$s^2 = \int_0^{40} (x-20)^2 F(x)\,dx = 0.025 \left.\frac{(x-20)^3}{3}\right|_0^{40} = 66.67$$

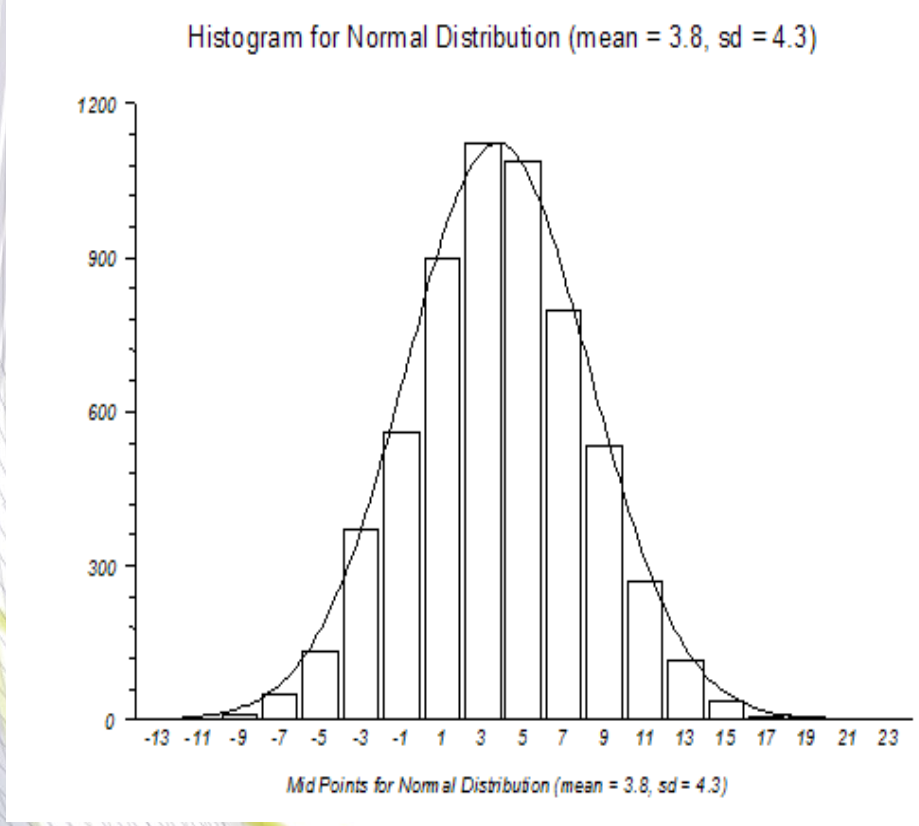
$$s = \sqrt{s^2} = \sqrt{66.67} = 8.16$$

$$P(X < 10) = \int_0^{10} F(x)\,dx = 0.025\, x\Big|_0^{10} = 0.25$$

$$P(5 < X < 10) = \int_5^{10} F(x)\,dx = 0.025\, x\Big|_5^{10} = 0.03 \cdot 10 - 0.025 \cdot 5 = 0.125$$

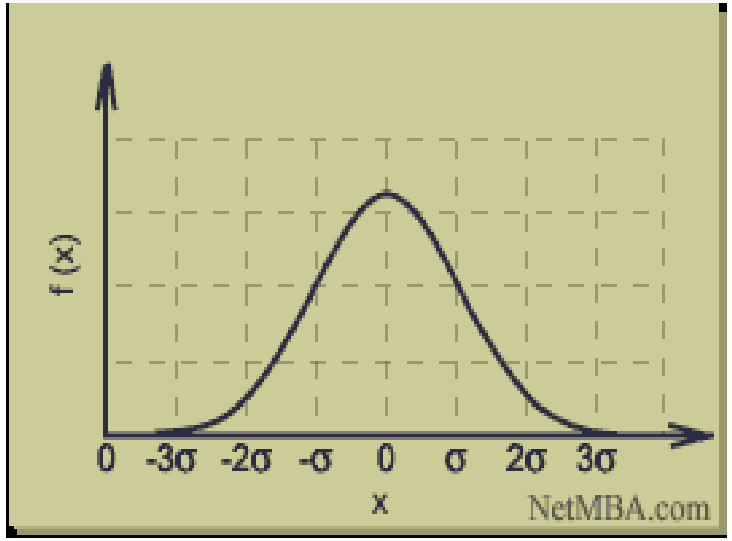
Нормална расподела

$$F(X) = N\left(\mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(X-\mu)^2}{2\sigma^2}}, -\infty < X < \infty$$



Стандардизована нормална расподела

$$Z = \frac{X - \mu}{\sigma} \sim N\left(0, 1^2\right)$$



Стандардизована нормална расподела – Правила за коришћење таблица

а) $P(Z < 3.64) = 0.9998636$

б) $P(Z < -0.86) = 0.19489$

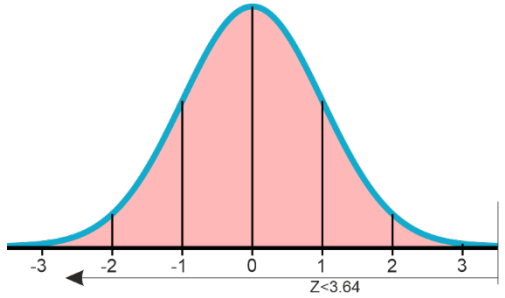
в) $P(Z > 1.26) = 1 - P(Z < 1.26) = P(Z < -1.26) = 0.10383$

г) $P(Z > -1.37) = P(Z < 1.37) = 0.91466$

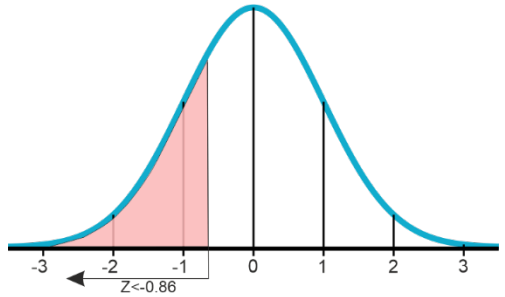
д) $P(-1.25 < Z < 0.37) = P(Z < 0.37) - P(Z < -1.25) = 0.6443087 - 0.1056498 = 0.53866$

е) $P(Z < u) = 0.975 \Rightarrow u = \Phi^{-1}(0.975) = 1.96$

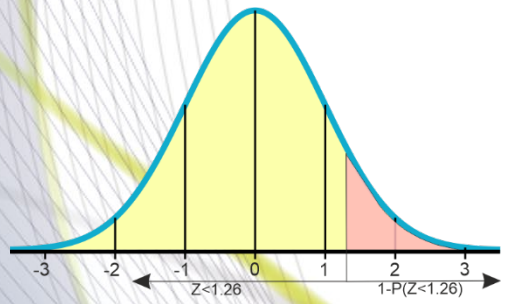
а)



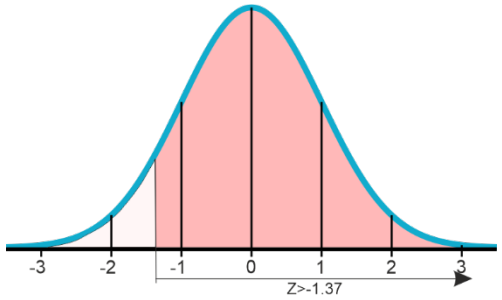
б)



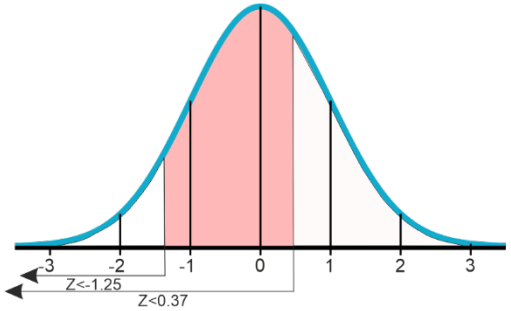
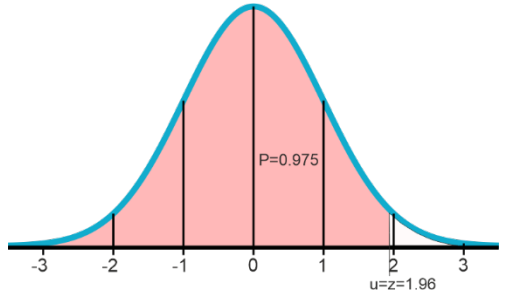
в)



г)



д)



е)