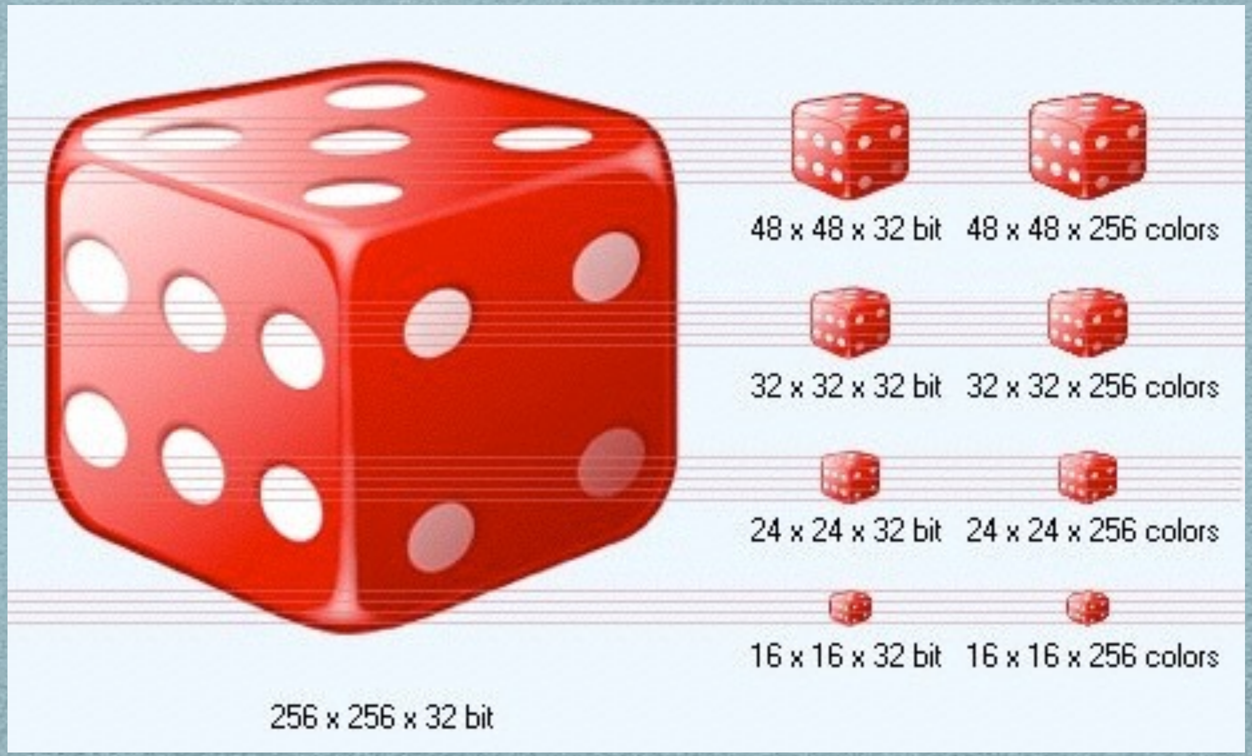


TEORIJA VJEROJATNOSTI

Proučavanje nasumičnosti i neodređenosti.

Jezik vjerojatnosti u svakodnevnom životu:

- cijena Plivinih dionica vjerojatno će porasti
- imamo 90% šansi da dobijemo utakmicu
- ovdje ima toliko osigurača da će sigurno barem jedan pregorjeti
- naš predsjednički kandidat ima 80% šanse za pobjedu
- položit ću Statistiku 100%

Osnovni pojmovi i definicije

Terminologija

Pokus = bilo koji postupak ili proces koji rezultira opažanjem

pokus 1: bacanje dvije kocke

pokus 2: mjerenje mase nekog predmeta. Ponavljanje pokusa

pokus 3: izvlačenje kuglice iz kutije. Ponavljanje pokusa

Ishod ili elementarni događaj = rezultat pokusa

primjer 1: izvučena je crvena kuglica

primjer 2: dvaput je palo pismo

Osnovni pojmovi i definicije

Prostor elementarnih događaja Ω = skup svih ishoda nekog pokusa

Primjer 1: bacamo dva različita novčića. Prostor elementarnih događaja je $\Omega = \{(P,P), (P,G), (G,P), (G,G)\}$

Primjer 2: bacamo novčić dok ne padne pismo. Prostor elementarnih događaja je $\Omega = \{(P), (GP), (GGP), (GGGP), ...\}$

Događaj = svaki podskup od Ω

- elementarni događaj = jednočlan podskup

- složeni događaj = višečlani podskup

Osnovni pojmovi i definicije

Primjer 1: bacamo dva različita novčića. Definiramo događaje A i B:

A = 'pali su jednako' = {(P,P); (G,G)}

B = 'pao je barem jedan grb' = {(P,G); (G,P); (G,G)}

Događaji A i B se međusobno ne isključuju.

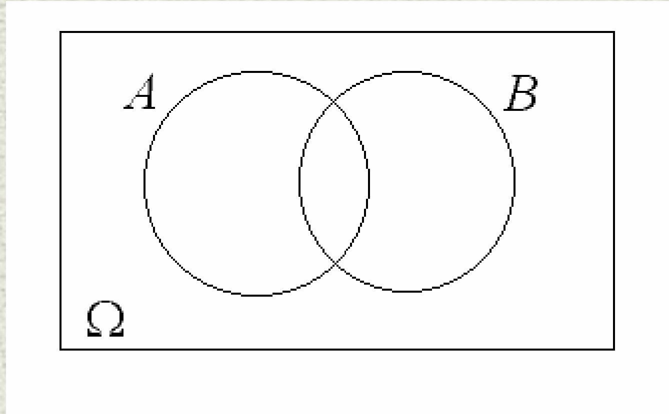
Primjer 2: pokus bacanja novčića dok ne padne pismo. Definiramo događaj A:
A = 'bacali smo barem četiri puta' = {(GGGP);, (GGGGP); {GGGGGP), ...}

Siguran događaj = Ω (prostor elementarnih događaja)

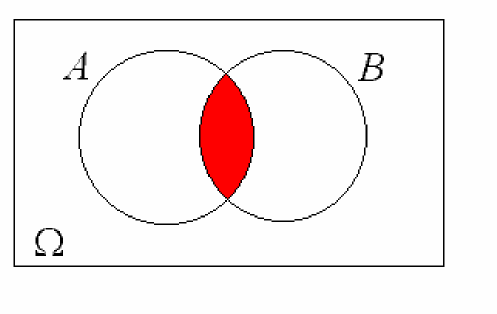
Nemoguć događaj = $\varnothing$ (prazan skup)

Osnovni pojmovi i definicije - relacije iz teorije skupova

Budući da smo događaje definirali kao skupove, prikladno je primjenjivati operacije iz teorije skupova. Za prikazivanje tih relacija koristimo Vennove dijagrame. Prostor elementarnih događaja označimo s Ω (pravokutnik), a ishodi tog pokusa predstavljeni su točkama unutar pravokutnika.

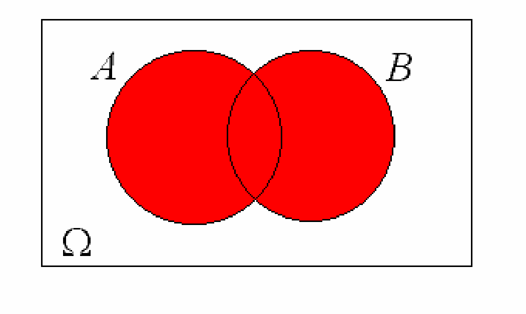


Složeni događaji A i B



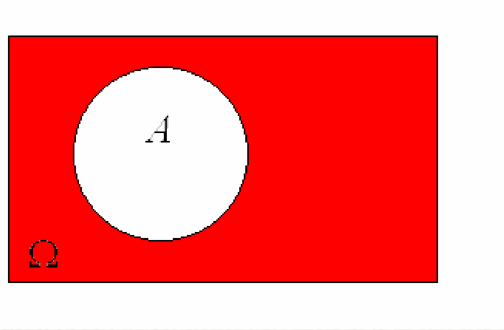
Presjek događaja A i B

$A \cap B$



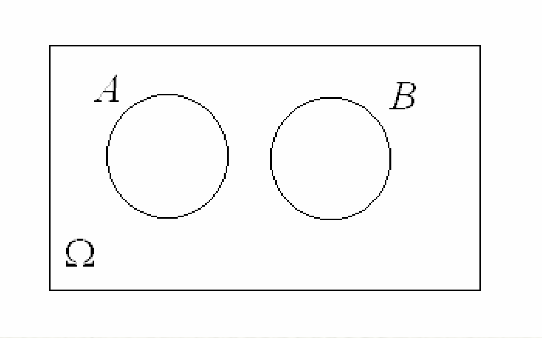
Unija događaja A i B

$A \cup B$



Komplement događaja A

$\overline{A}$



Međusobno isključivi događaji A i B

$A \cap B = \emptyset$

Definicija vjerojatnosti

Definicija **a priori**

Neka je *m(A)* broj ishoda koji realiziraju svojstvo *A* i neka su svi oni jednako mogući. Neka je *n* ukupan broj ishoda. Vjerojatnost događaja *A* je:

$$P(A) = \frac{m(A)}{n}; \quad 0 \leq P(A) \leq 1$$

Manjkavosti ove definicije:

- primjenjiva je samo na konačne skupove
- definicija je kružna. Koristimo izraz 'jednako mogući' što znači 'jednako vjerojatni' da bismo definirali vjerojatnost

Definicija vjerojatnosti

Primjer 1: bacamo novčić. Kolika je vjerojatnost da padne pismo?

$P = 1/2$

Na osnovi čega to zaključujemo:
- iskustvo
- pretpostavljena simetrija novčića

Primjer 2: bacamo kocku. A priori vjerojatnost da padne trica je $P(3) = 1/6$.

Oprez! Još u staroegipatskim piramidama pronađene su “nepoštene” kocke.

Postoji i druga definicija vjerojatnosti

Za nju je potrebno definirati još dva pojma

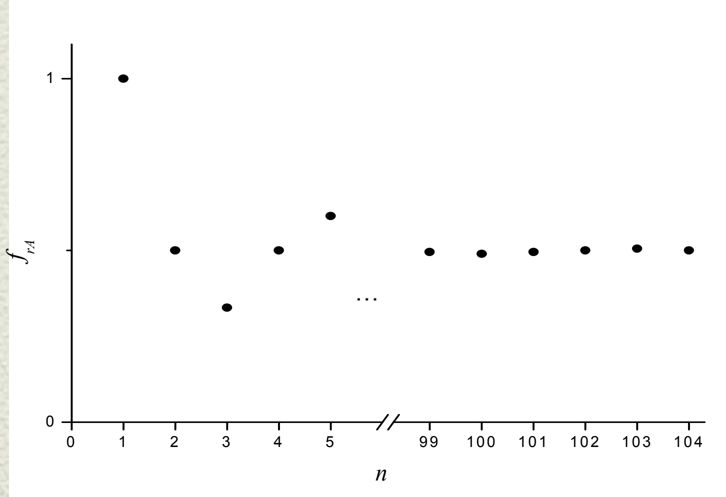
Frekvencija

Ponovimo slučajni pokus n puta. Neka se događaj A pojavio točno n_A puta. Tada broj n_A nazivamo frekvencijom događaja $A(f_A)$, a broj

$f_{rA} = \frac{n_A}{n} = \frac{f_A}{n}$ je relativna frekvencija tog događaja. Vrijedi $0 \leq f_{rA} \leq 1$

Statistička stabilnost relativnih frekvencija

Ako se prilikom velikog broja ponavljanja slučajnog pokusa relativne frekvencije događaja A grupiraju oko nekog broja, kažemo da je relativna frekvencija statistički stabilna.



Definicija vjerojatnosti **a posteriori**

Ako slučajni pokus zadovoljava uvjet statističke stabilnosti relativnih frekvencija, onda se vjerojatnost a posteriori proizvoljnog događaja *A* definira kao realan broj P(A)∈[0,1] oko kojeg se grupiraju relativne frekvencije tog događaja

Ovu definiciju još zovemo i ‘zakon velikih brojeva’

Problemi:

- kako provjeriti stabilnost?
- kako odrediti vjerojatnost jednog događaja? (Velika je vjerojatnost da sutra neće padati kiša)

Suprotna vjerojatnost je vjerojatnost da se svojstvo *A* ne realizira:

$$P(\bar{A}) = 1 - P(A) = Q(A)$$
$$P(A) + Q(A) = 1$$

Aksiomatsko zasnivanje teorije vjerojatnosti

A.N. Kolmogorov, 1933.

Poznajemo li prostor elementarnih događaja Ω za neki pokus, svrha teorije vjerojatnosti je da se svakom događaju $A \subseteq \Omega$ pridruži broj $P(A)$ koji će biti precizna mjera šanse da se A ostvari. Svako takvo pridruživanje mora zadovoljavati sljedeće aksiome:

A1. Za svaki događaj $A \Rightarrow P(A) \geq 0$

A2. $P(\Omega) = 1$

A3. a) ako se konačan broj događaja $A_1, A_2, \ldots, A_n$ međusobno isključuje, vrijedi: $$P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} P(A_i)$$

b) ako se prebrojivo beskonačan broj događaja $A_1, A_2, \ldots, A_n$ međusobno isključuje, vrijedi: $$P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{\infty} P(A_i)$$

Svojstva vjerojatnosti
Propozicije:

$$\text{P1. } \forall A \Rightarrow P(A) = 1 - P(\bar{A})$$

dokaz: Iz trećeg aksioma. Neka je $n = 2$. Neka je $A_1 = A \Rightarrow A_2 = \bar{A}$

$$A \cup \bar{A} = \Omega$$

$$1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$$

primjer: znamo da je 85% ljudi Rh-pozitivno. Liječnik čini krvni test na novorođenčadi tako dugo dok ne nađe Rh-negativno. Kolika je vjerojatnost da će morati učiniti barem 2 testa?

rješenje:
zanima nas događaj A = 'barem 2 testa' = {(+,-), (+,+,-), (+, +, +, -),...}
njegov komplement $\bar{A}$ = 'prvi Rh je negativan' = {(-)}
Lako izračunamo vjerojatnost $P(\bar{A}) = 0.15$
Dakle $P(A) = 0.85$

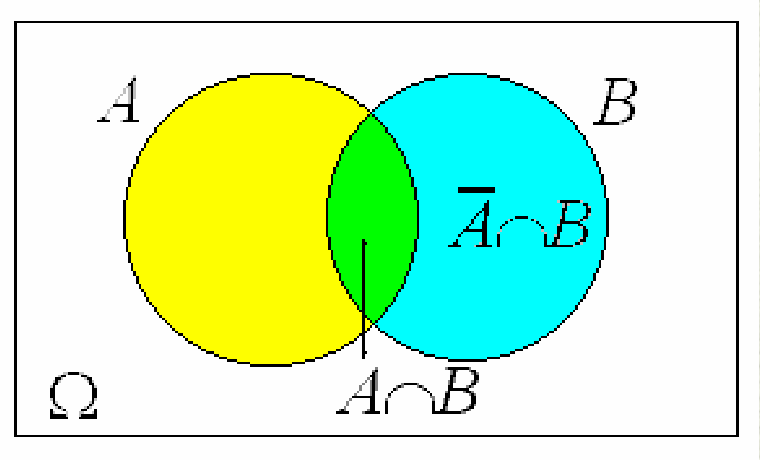
Svojstva vjerojatnosti
Propozicije:

P2. Ako se A i B isključuju, onda je $P(A \cap B) = 0$

$$A \cap B = \varnothing \Rightarrow \overline{A \cap B} = \Omega \Rightarrow 1 = P(\Omega) = P\left(\overline{A \cap B}\right) = 1 - P(A \cap B)$$

P3. Za bilo koja dva događaja A i B vrijedi:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Sustavno određivanje vjerojatnosti

Neka neki pokus ima velik broj mogućih ishoda (elementarnih događaja) E_i, koji se međusobno isključuju. Za određivanje vjerojatnosti složenog događaja A, najbolje je najprije odrediti vjerojatnosti svih elementarnih događaja E_i. Mora vrijediti: $\forall E_i \Rightarrow P(E_i) \geq 0$,

$$\sum_i P(E_i) = 1$$

Tada je vjerojatnost bilo kojeg složenog događaja A dana zbrajanjem

$$P(A) = \sum_{\text{svi } E_i \text{ iz } A} P(E_i)$$

Primjer: nepoštena kocka

$P(E_1)=P(E_3)=P(E_5)=1/9$; $P(E_2)=P(E_4)=P(E_6)=2/9$

Zanima nas događaj A; A = ‘broj manji od 4’

$P(A) = P(E_1)+P(E_2)+P(E_3) = 4/9$

Događaji koji se isključuju
Zbrajanje vjerojatnosti ("ili")

Događaji A_1 i A_2 se isključuju ako istovremeno ne mogu nastupiti oba.
Za događaje koji se isključuju vrijedi da je vjerojatnost da se dogodi ili A_1 ili A_2 ili ... A_n

$$P\left(A_1 + A_2 + \ldots + A_n\right) = \frac{m\left(A_1\right) + m\left(A_2\right) + \ldots + m\left(A_n\right)}{n} = P\left(A_1\right) + P\left(A_2\right) + \ldots + P\left(A_n\right)$$

Primjer: kolika je vjerojatnost da se pri jednom bacanju jedne kocke okrene ili 2 ili 3?
A_1 = 'okrene se 2' A_2 = 'okrene se 3' (isključuju se)
$P(A_1) = 1/6$, $P(A_2) = 1/6$
$P(A_1+A_2) = P(A_1)+P(A_2) = 1/3$

Događaji koji se NE isključuju
Množenje vjerojatnosti ("i")

Razmotrimo A_1 i A_2 - događaji koji se međusobno ne isključuju
Npr. istovremeno bacanje kocke i novčića, događaji A_1 = 'okrene se P', A_2 = 'okrene se 2 na kocki'

Definicija: svojstva A_1, A_2, ..., A_n su nezavisna ako je vjerojatnost da istovremeno dogode i A_1 i A_2 i ... i A_n:

$$P\left(A_1 \cap A_2 \cap \ldots \cap A_n\right) = P\left(A_1 \cdot A_2 \cdot \ldots \cdot A_n\right) = P\left(A_1\right) \cdot P\left(A_2\right) \cdot \ldots \cdot P\left(A_n\right)$$

Primjer: kolika je vjerojatnost da se u istovremenom bacanju novčića i kocke okrene pismo i dvojka?

$n = 2 \times 6 = 12$, $m = 1 \Rightarrow P = 1/12$

A_1 = 'okrene se P' A_2 = 'okrene se 2'
$P(A_1 \cdot A_2) = P(A_1) \cdot P(A_2) = 1/2 \cdot 1/6 = 1/12$

Geometrijska vjerojatnost

Definicija:
P = (područje u kojem se svojstvo ostvaruje)/(ukupno područje) = M/N

područje: duljina, površina, volumen...

Primjer: kolika je vjerojatnost da je slučajno odabran realni broj iz intervala [0, 5] veći od 1 a manji od 3?
N = duljina intervala = 5
$M = 3 - 1 = 2$
$P = 2/5$

Jednako vjerojatni ishodi

Primjeri:
izvlačenje karata iz novog špila, bacanje poštene kocke, popunjavanje energetskih nivoa,...

Imamo n mogućih ishoda, a vjerojatnost svakog od njih je

$$P(E_i) = \frac{1}{n}, \forall i$$

Vjerojatnost događaja A je $P(A) = \frac{n(A)}{n}$ (vratili smo se na definiciju a priori)

Uvjetna vjerojatnost i statistika

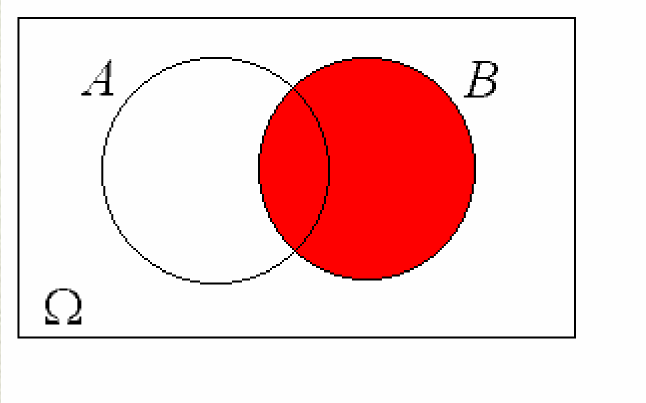
Uvjetna vjerojatnost

Često neka vjerojatnost ovisi o tome što znamo o tijeku pokusa. Zamislimo da u nekom pokusu gledamo događaje A i B, čije su vjerojatnosti $P(A)$ i $P(B)$. Međutim, ako smo saznali da se već dogodio događaj A, vjerojatnost događaja B ne mora biti ista kao kad nismo znali da se događaj A desio. Tu novu vjerojatnost događaja B nazivamo:

Uvjetna vjerojatnost događaja B ako se dogodio A i označavamo

$$P(B|A)$$

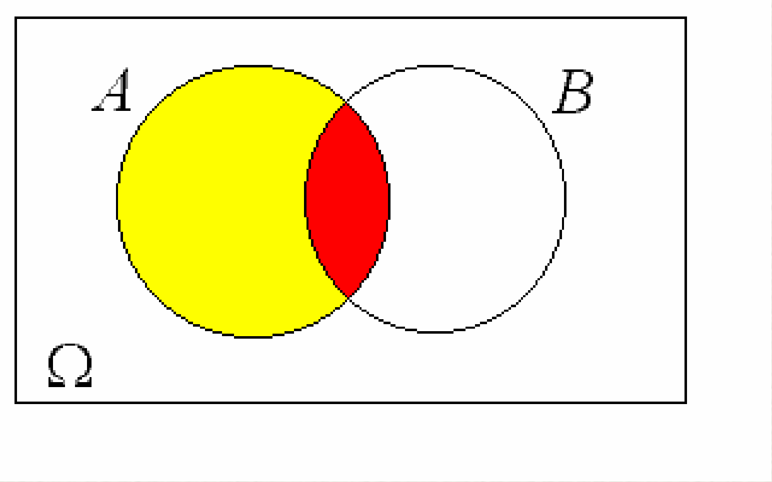
Prije nego što znamo ishod događaja A, vjerojatnost za B, je $P(B)$. To je vjerojatnost svih elementarnih događaja u skupu B podijeljena s vjerojatnošću sigurnog događaja $P(\Omega)$.



$P(B)$

Uvjetna vjerojatnost

Međutim, ako se A već dogodio, onda on predstavlja novi siguran događaj pa je vjerojatnost događaja B u stvari vjerojatnost svih elementarnih događaja iz skupa $A \cap B$ podijeljena s vjerojatnošću događaja A.



$P(B|A)$

Tako dobivamo izraz za uvjetnu vjerojatnost:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Uvjetna vjerojatnost
Primjer:

Dva stroja proizvela su 19 proizvoda i bacila ih u istu kutiju. Stroj A proizveo je 9 ispravnih i 1 neispravan proizvod, dok je stroj B proizveo 7 ispravnih i 2 neispravna proizvoda.
Kontrolor izvlači jedan proizvod iz kutije. a) Kolika je vjerojatnost da izvuče proizvod iz stroja B? b) Ako je izvukao neispravan proizvod, kolika je vjerojatnost da je on iz stroja B?

Rješenje: Napravimo tablicu

	I	N
A	9	1
B	7	2

a) $P(B) = 9/19 = 47\%$
b) $P(B|N) = P(B \cap N)/P(N) = 2/3 = 67\%$

Potpuni sustav događaja

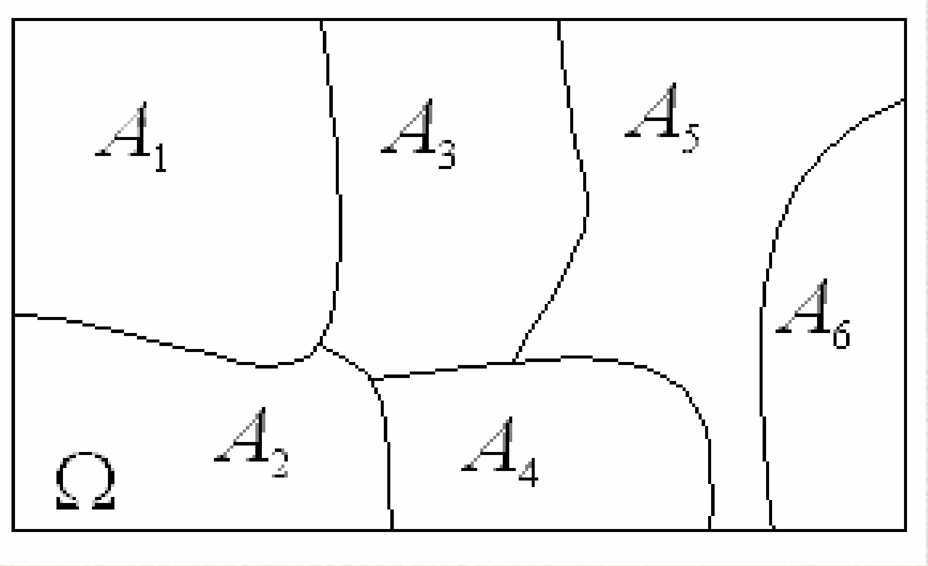
Neka za događaje A_i ($i = 1, 2, \ldots$) vrijedi:

$A_i \neq \emptyset$, $\forall i, j\ i \neq j$

$A_i \cap A_j = \emptyset$, $\forall i,j\ i \neq j$ (međusobno se isključuju)

$\cup A_i = \Omega$ (prekrivaju cijeli prostor)

Tada kažemo da skupovi A_i čine potpuni sustav događaja

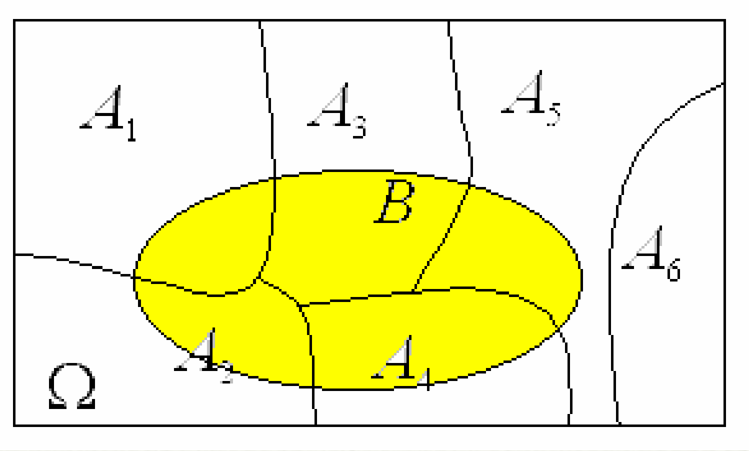


Zakon totalne vjerojatnosti

Neka A_i čine potpuni sustav događaja. Tada za bilo koji događaj B vrijedi:

$$P(B) = \sum_i P(B|A_i) \cdot P(A_i)$$

Dokaz:



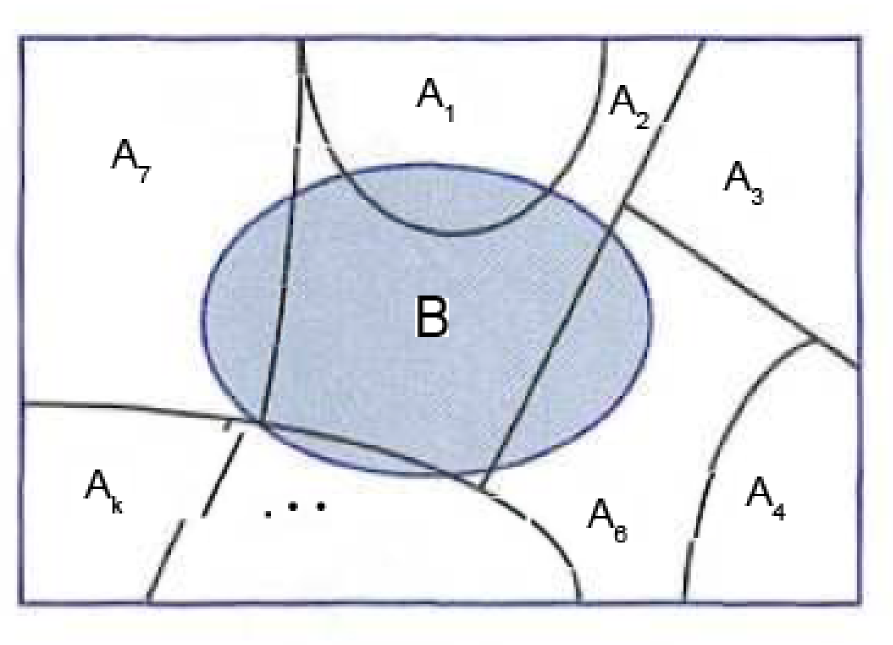
Iz slike vidimo da je $P(B) = \sum P(A_i \cap B)$ pa iz definicije uvjetne vjerojatnosti slijedi gornja tvrdnja.

Bayesov teorem

Neka A_i čine potpuni sustav događaja. Neka je B neki događaj. Tada vrijedi:

$$P(A_k|B) = \frac{P(A_k \cap B)}{P(B)} = \frac{P(B|A_k)\cdot P(A_k)}{P(B)} = \frac{P(B|A_k)\cdot P(A_k)}{\sum_i P(B|A_i)\cdot P(A_i)}$$

(ako se dogodi B, koja je vjerojatnost da se dogodi A_k?)



$A_1, A_2, \ldots, A_k$ se međusobno isključuju!

primjer:

$$P(A_k|B) = \frac{P(A_k \cap B)}{P(B)} = \frac{P(B|A_k)\cdot P(A_k)}{P(B)} = \frac{P(B|A_k)\cdot P(A_k)}{\sum_i P(B|A_i)\cdot P(A_i)}$$

U jednoj školi 40% učenika čine cure, a 60% dečki. Cure podjednako nose hlače i suknje, dok svi dečki nose hlače. Promatrač iz daljine opaža učenika kojem ne može razaznati spol, ali uočava da nosi hlače. Kolika je vjerojatnost da je učenik ženskog spola?

Rješenje:
definirajmo događaje i vjerojatnosti
A = 'učenik je žensko', $P(A) = 0.4$
B = 'učenik nosi hlače', $P(B) = 0.5 \cdot 0.4 + 0.6 = 0.8$ $P(B|A) = 0.5$

Zanima nas $P(A|B)$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0.5\cdot 0.4}{0.8} = 0.25$$

Slučajna varijabla

Pokus: test tri elektronske komponente na ispravnosti

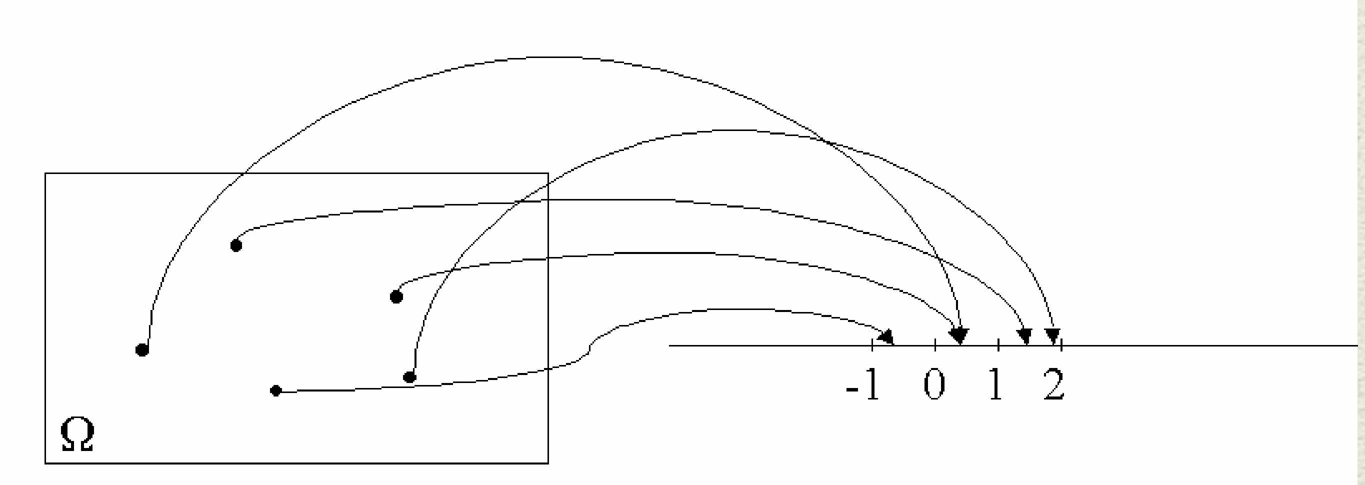
$\Omega = \{NNN, NNI, NIN, INN, NII, INI, IIN, III\}$ (n-neispravan, i-ispravan)

Zanima nas koliko ima neispravnih komponenti. Svakom elementu od Ω možemo pridružiti broj 0, 1, 2 ili 3.

Def: Za dani prostor događaja Ω nekog pokusa, slučajna varijabla jest bilo koje pravilo kojim svakom ishodu pokusa pridružujemo neki broj.

D - skup svih mogućih vrijednosti slučajne varijable X. Vrijedi $D \subseteq \Re$

Sa X označavamo slučajnu varijablu, a sa x njenu vrijednost



Primjeri:

1. Bacamo obojenu kocku. Svakoj boji pridružimo neki broj. Npr. žuta → 2.
Ω ={bijela, žuta, crvena, narančasta, zelena, bijela}; $D = \{1, 2, 3, 4, 5, 6\}$.
X(bijela) = 1, X(žuta) = 2, itd...

2. Bacamo dvije kocke, crvenu i zelenu. Prostor elementarnih događaja je
$\Omega = \{(1,1), (1,2), (1,3), ..., (5,6), (6,6)\} = \{a,b\}: a,b \in \{1, 2, 3, 4, 5, 6\}\}$

Definiramo 3 različite slučajne varijable:

a) X = zbroj točkica na kockama
$D = \{2, 3, 4, ..., 12\}$; npr. $X((1,3)) = 4$

b) Y = razlika broja točkica na crvenoj i zelenoj kocki
$Y = \{-5, -4, ..., 5\}$; npr. $Y((1,3)) = -2$

c) Z = umnožak broja točkica na kockama
$Z = \{1, 2, 3, ..., 36\}$; npr. $Z((5,3)) = 15$

Primjeri:

3. Bacanje novčića

a) jednom $\Omega = \{P,G\}$; $D = \{0,1\}$

$X(P) = 0$, $X(G) = 1$

Def. Događaj koji ima samo dva moguća ishoda zove se Bernoullijev događaj, a slučajna varijabla koja ga opisuje zove se Bernoullijeva slučajna varijabla.

b) dok ne padne pismo

$\Omega = \{(P), (GP), (GGP), \ldots\}$; $D = \mathbf{N} = \{1, 2, 3, \ldots\}$

slučajna varijabla: Y = broj bacanja dok ne padne pismo

$Y((P)) = 1$, $Y((GP)) = 2$, $Y((GGP)) = 3$, ...

Slučajne varijable mogu biti diskretne ili kontinuirane!

Diskretna slučajna varijabla i raspodjela vjerojatnosti

Def: Raspodjela vjerojatnosti diskretne slučajne varijable X definirana je za svaki broj x relacijom:
$p(x) = P(X=x) = P(\text{svi ishodi } s \in \Omega \text{: } X(s) = x)$
Ovdje se P(X=x) čita 'vjerojatnost da slučajna varijabla X poprimi vrijednost x'

Raspodjele vjerojatnosti možemo prikazati tablično ili funkcionalno.

Primjeri: Raspodjele vjerojatnosti za slučajne varijable iz primjera 1, 2a, 2b, 2c, 4a i 4b. To su *a priori* vjerojatnosti uz pretpostavku jednako vjerojatnih ishoda.

1.

$$p(x) = \begin{cases} 1/6 & , \quad x \in D = \{1,2,3,4,5,6\} \\ 0 & , \quad x \notin D \end{cases}$$

2.a)

$$p(x) = \begin{cases} \frac{1}{36}\left(6-|x-7|\right) & , \quad x \in D = \{2,3,4,\ldots,12\} \\ 0 & , \quad x \notin D \end{cases}$$

Primjeri:

2.b)

$$p(y)=\begin{cases}\frac{1}{36}\left(6-|y|\right) & , \quad y\in D=\{-5,-4,\ldots,0,\ldots,5\}\\ 0 & , \quad y\notin D\end{cases}$$

2.c)

z	$p(z)$
1	1/36
2	2/36
3	2/36
4	3/36
5	2/36
6	4/36
8	2/36
9	1/36
10	2/36
12	4/36
15	2/36
16	1/36
18	2/36
20	2/36
24	2/36
25	1/36
30	2/36
36	1/36

Primjeri:

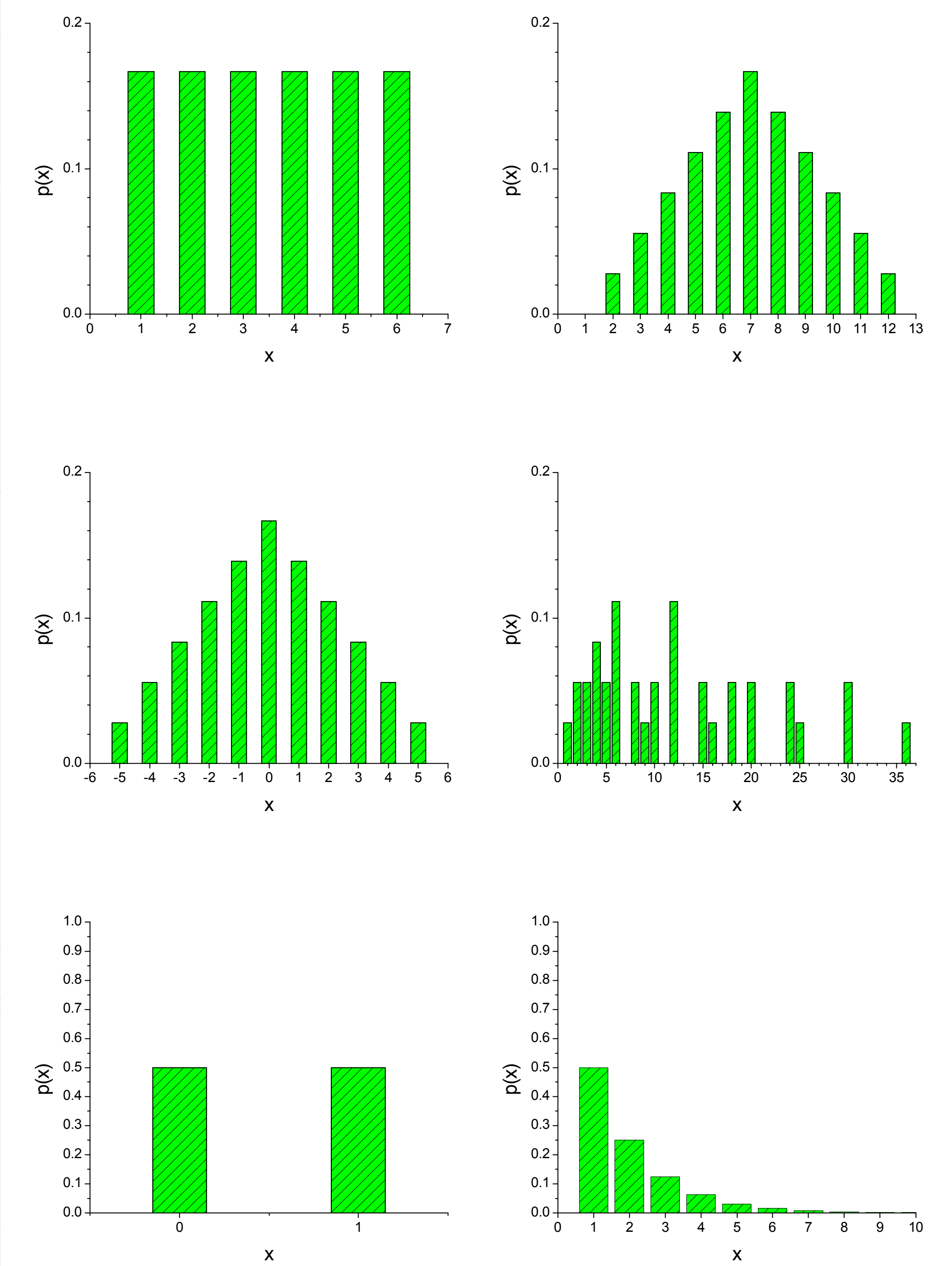
4.a)

$$p(x) = \begin{cases} 1/2 & , \quad x \in D = \{0,1\} \\ 0 & , \quad x \notin D \end{cases}$$

4.b)

$$p(y) = \begin{cases} 1/2^{y} & , \quad y \in \mathbf{N} = \{1,2,3,4,.....\} \\ 0 & , \quad y \notin \mathbf{N} \end{cases}$$

Raspodjele vjerojatnosti za slučajne varijable iz primjera 1, 2a, 2b, 2c, 4a i 4b:



Važno!!

$$\boxed{\sum_{x \in D} p(x) = 1}$$

Funkcija raspodjele vjerojatnosti - kumulativna funkcija distribucije

Def. Za diskretnu slučajnu varijablu koja ima raspodjelu vjerojatnosti $p(x)$ definira se funkcija raspodjele $F(x)$ na sljedeći način:

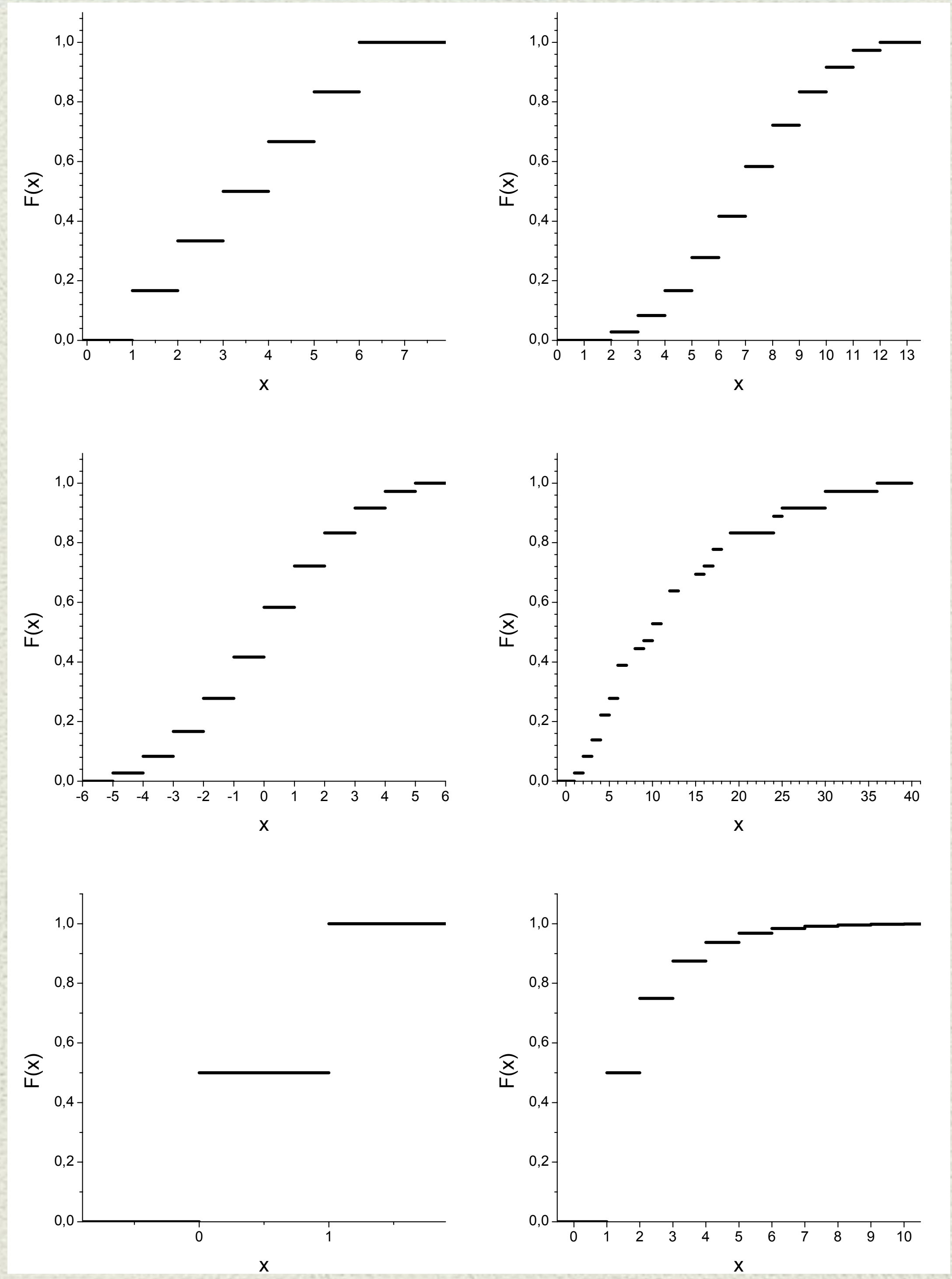
$$F(x) = P(X \leq x) = \sum_{y \leq x} p(y)$$

Za svaki x, $F(x)$ predstavlja vjerojatnost da X poprimi vrijednost ne veću od x.

Vrijedi:

$$\boxed{F(-\infty) = 0; \quad F(+\infty) = 1}$$

Primjer: funkcije raspodjele za slučajne varijable iz primjera 1, 2a, 2b, 2c, 4a i 4b:



Za bilo koja dva broja a i b ($a < b$) vrijedi sljedeće:

$$P(a \leq X \leq b) = F(b) - F(a-),$$

gdje a- predstavlja najveću vrijednost varijable X koja je strogo manja od a

Primjer 4 b) Kolika je vjerojatnost da ćemo morati bacati najmanje dva a najviše 5 puta?

$$P(2 \le Y \le 5) = F(5) - F(1)$$

$$F(1) = \sum_{y'=1}^{1} \frac{1}{2^{y'}} = \frac{1}{2},$$

$$F(5) = \sum_{y'=1}^{5} \frac{1}{2^{y'}} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} = \frac{31}{32}$$

$$P(2 \le Y \le 5) = \frac{15}{32}$$

Karakterizacija raspodjela
Očekivana vrijednost diskretne slučajne varijable

Zanima nas srednja vrijednost raspodjele!

Def. Neka je *X* diskretna slučajna varijabla sa skupom mogućih vrijednosti *D* i neka je *p(x)* njezina funkcija vjerojatnosti. Tada je srednja vrijednosti ili očekivanje varijable *x* dano s:

$$E(X) = \mu_X = \sum_{x \in D} x \cdot p(x)$$

Očekivanja za raspodjele vjerojatnosti iz primjera 1, 2a, 2b, 2c, 4a i 4b:

1. $E(X) = \sum_{x \in D} x \cdot p(x) = \sum_{x=1}^{6} \frac{1}{6} x = \frac{21}{6} = 3.5$

2.a) $E(X) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + ... + 7 \cdot \frac{6}{36} + ... + 12 \cdot \frac{1}{36} = 7$

2.b) $E(Y) = -5 \cdot \frac{1}{36} - 4 \cdot \frac{2}{36} + ... + 0 \cdot \frac{6}{36} + ... + 5 \cdot \frac{1}{36} = 0$

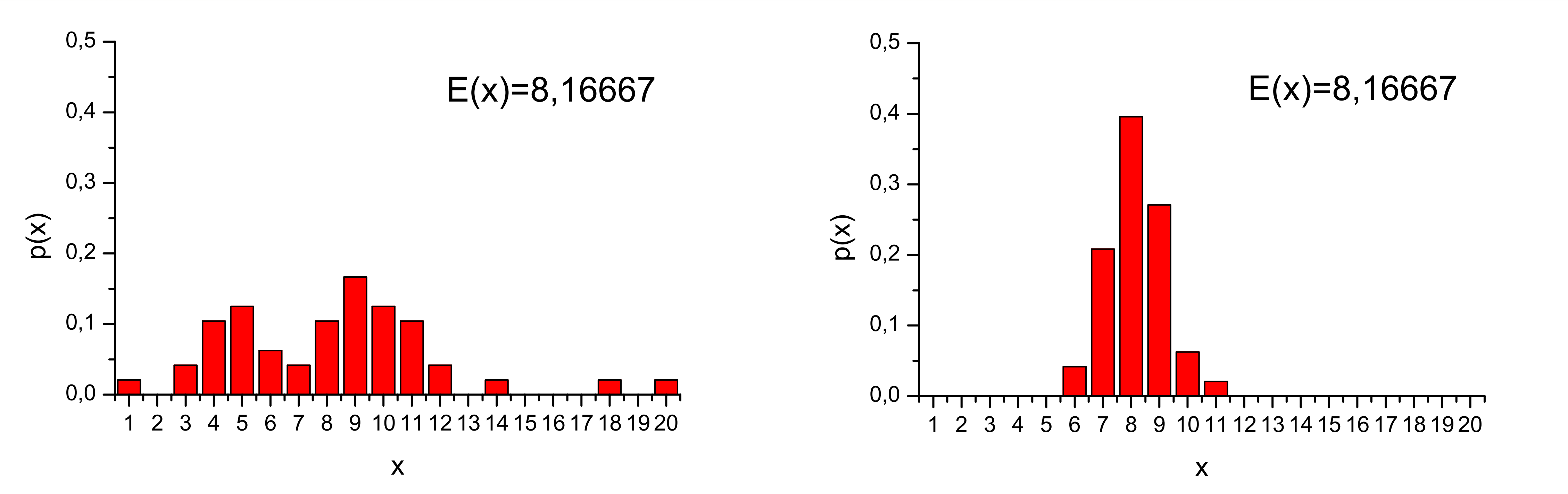
2.c) $E(Z) = 12.25$

4.a) $E(X) = \sum_{x=0}^{1} x \cdot \frac{1}{2} = \frac{1}{2}$

4.b) $E(Y) = \sum_{y=1}^{\infty} y \cdot \frac{1}{2^y} = 2$

Varijanca diskretne slučajne varijable

Sjetimo se primjera djece iz parka! Prosječna starost djece je bila 8.17 godina. Zanima nas raspršenost dječjih uzrasta u parku.



raspodjela starosti djece

neka druga raspodjela

jednake srednje vrijednosti!

Mjera raspršenja - srednje kvadratično odstupanje, tj. aritmetička sredina kvadrata odstupanja

Neka je *X* diskretna slučajna varijabla i neka je *p(x)* njezina raspodjela vjerojatnosti. Tada je varijanca od *X* dana s očekivanjem kvadrata odstupanja od μ.

$$V(X) = \sigma_X^2 = \sum_{x \in D} (x - \mu_X)^2 p(x) = E(X - \mu_X)^2$$

Veličinu nazivamo $\sigma_X = \sqrt{V(X)}$ standardnim odstupanjem ili standardnom devijacijom slučajne varijable *X*.

Izračunamo li standardne devijacije za raspodjele s prethodne slike, dobivamo $\sigma = 3.64$ za lijevu raspodjelu, i $\sigma = 1.03$ za desnu

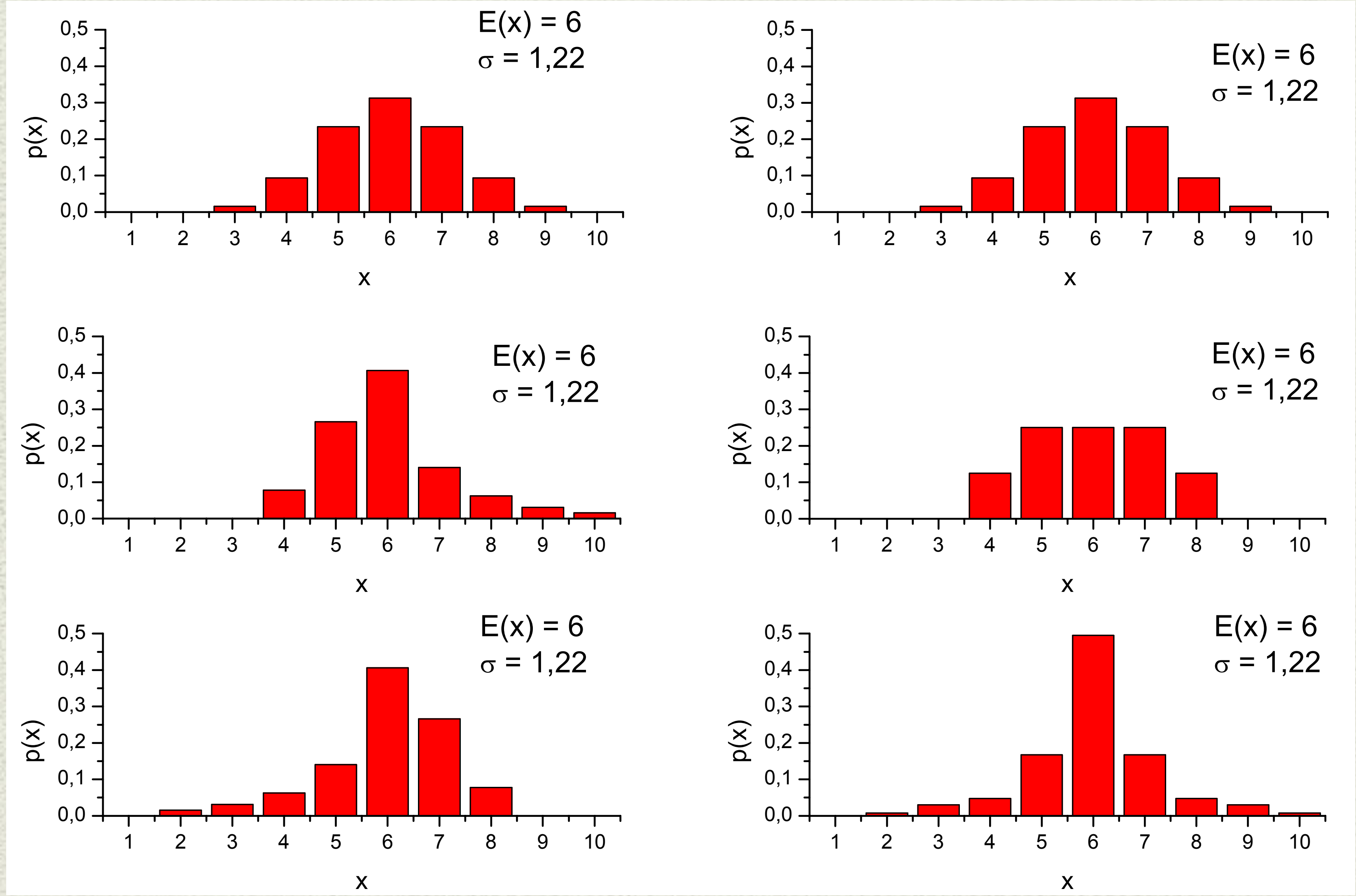
Važno svojstvo koje često olakšava račune je:

$$\boxed{V(X) = \sigma^2 = \sum_{x \in D} \left(x^2 p(x) - \mu_X^2\right) = E(X^2) - [E(X)]^2}$$

$$\overline{(x^2)} - (\bar{x})^2$$

Momenti višeg reda

- dodatne veličine koje opisuju raspodjelu
- sljedeće raspodjele imaju istu srednju vrijednost i standardnu devijaciju



$$E(x) = \mu = \sum_x x p(x)$$

$$V(x) = \sum_x (x-\mu)^2 p(x)$$

Momenti višeg reda

Def. Središnji (centralni) moment r-tog reda je:

$$M_r = \sum_{x \in D} (x - \mu)^r \, p(x)$$

Def. Pomoćni (ishodišni) moment r-tog reda je:

$$m_r = \sum_{x \in D} x^r \, p(x)$$

- središnji moment nultog reda totalnu vjerojatnost:

$$M_0 = m_0 = 1$$

- središnji moment prvog reda predstavlja srednje odstupanje od aritmetičke sredine:

$$M_1 = 0$$

Momenti višeg reda

- pomoćni moment prvog reda predstavlja očekivanje ili aritmetičku sredinu:

$$m_1 = E(X) = \overline{x} = \mu$$

- središnji moment drugog reda predstavlja varijancu:

$$M_2 = V(X) = \sigma^2$$

- za izračun središnjeg momenta drugog reda možemo upotrijebiti relaciju: $M_2 = m_2 - m^2{}_1$

Općenita formula za središnji moment *r*-tog reda je:

$$\boxed{M_r = \sum_k (-1)^k \binom{r}{k} m_{r-k} m_1^k}$$

Momenti višeg reda

Moment trećeg reda govori o asimetriji. Primjeri iz lijevog stupca prethodne slike imaju sljedeće momente trećeg reda:
$M_3 = 0$, $M_3 = 1.594 > 0$, $M_3 = -1.594 < 0$

Definiramo veličinu koeficijent asimetrije kao $\alpha_3 = \dfrac{M_3}{\sigma^3}$

Ako je $\alpha_3 = 0$ raspodjela je simetrična, ako je $\alpha_3 > 0$ onda je nagnuta udesno a ako je $\alpha_3 < 0$ onda je nagnuta ulijevo.

Gornji primjeri imaju: $\alpha_3 = 0$, $\alpha_3 = 0.868 > 0$, $\alpha_3 = - 0.868 <0$

Momenti višeg reda

Moment četvrtog reda govori o spljoštenosti. Primjeri iz desnog stupca prethodne slike imaju sljedeće momente četvrtog reda: $M_4 = 6$, $M_4 = 4.5$, $M_4 = 10.67$

Definiramo veličinu koeficijent spljoštenosti kao $\alpha_4 = \dfrac{M_4}{\sigma^4}$

Ako je $\alpha_4 = 3$ raspodjela je 'normalno spljoštena', ako je $\alpha_4 > 3$ onda je 'šiljata' a ako je $\alpha_4 < 3$ onda je 'široka'.

Gornji primjeri imaju: $\alpha_4 = 2.67$, $\alpha_4 = 2$, $\alpha_4 = 4.74$

Binomna ili Bernoullijeva raspodjela

Bernoullijev (binomni) pokus: pokus koji ima samo dva moguća ishoda
- 'uspjeh' = A, 'neuspjeh' = $\bar{A}$
- vjerojatnost uspjeha označavamo s p, a neuspjeha s q:

  $p = P(A)$, $q = P(\bar{A})$

Očito je $p + q = 1$

Bernoullijev (binomni) pokus:
1. sastoji se od n pokušaja, a n je određen unaprijed
2. pokušaji su identični Bernoullijevi pokusi s mogućim ishodima 'uspjeh' = A i 'neuspjeh' = $\bar{A}$
3. Pokušaji su nezavisni. Ishod bilo kojeg pokušaja ne utječe na ishod drugog
4. Vjerojatnost '*uspjeha*' jednaka je za sve pokušaje i označavamo ju s p

Primjeri:

1. Bacamo novčić četiri puta (ili bacimo četiri novčića). Uspjeh je kad padne 'glava':

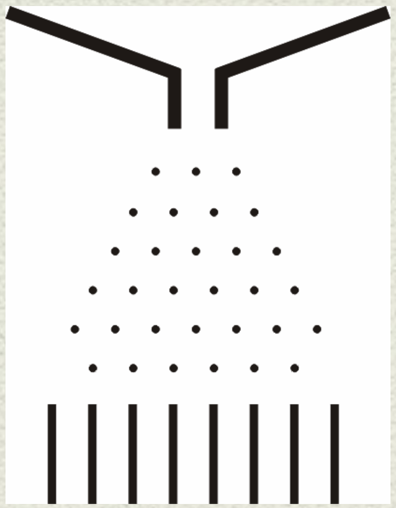
$$n = 4,\ p = 1/2$$

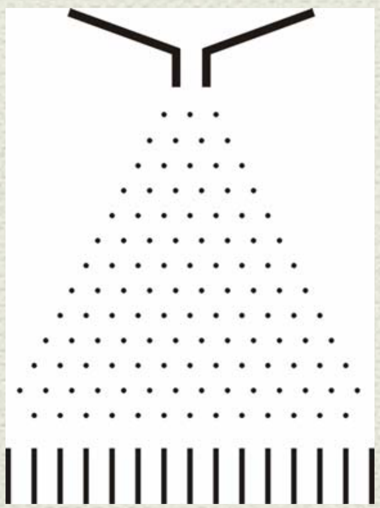
2. Bacamo kocku 10 puta. Uspjeh je kad padne šestica:

$$n = 10,\ p = 1/6$$

3. Galtonova daska: kuglicu spuštamo na čavliće koji su složeni u pravilnu trokutastu rešetku. Padom na čavlić, kuglica može skrenuti ulijevo (*neuspjeh*) ili udesno (*uspjeh*). Ako je daska pravilna i vodoravna, ti su ishodi jednako vjerojatni.

$$n = \text{broj redova čavlića},\ p = 1/2$$





Binomna slučajna varijabla

Def. Ako imamo binomni eksperiment koji se sastoji od n pokušaja, slučajnu varijablu X definiranu kao X = broj uspjeha u n pokušaja nazivamo binomna slučajna varijabla i označavamo ju $X \sim \mathrm{Bin}(n, p)$

Binomna raspodjela vjerojatnosti

Ako imamo binomni eksperiment i binomnu slučajnu varijablu $X \sim \mathrm{Bin}(n, p)$ njezinu raspodjelu vjerojatnosti $p(x)$ nazivamo binomna raspodjela i označavamo ju s $b(x; n, p)$

Odredimo vrijednosti $b(x; n,p)$

- pogledajmo Galtonovu dasku s $n = 4$

Galtonova daska - binomni pokus s $n = 4$

A = ‘kuglica se odbila udesno’, $\bar{A}$ = ’ kuglica se odbila ulijevo’

x = broj odbijanja udesno

Svi mogući ishodi binomnog pokusa s $n = 4$

p=0.5, q=0.5

ishod	slučajna varijabla x	vjerojatnost
$\bar{A}\bar{A}\bar{A}\bar{A}$	0	q^4
$\bar{A}\bar{A}\bar{A}A$	1	$p \cdot q^3$
$\bar{A}\bar{A}A\bar{A}$	1	
$\bar{A}A\bar{A}\bar{A}$	1	
$A\bar{A}\bar{A}\bar{A}$	1	
$\bar{A}\bar{A}AA$	2	$p^2 \cdot q^2$
$\bar{A}A\bar{A}A$	2	
$\bar{A}AA\bar{A}$	2	
$A\bar{A}\bar{A}A$	2	
$A\bar{A}A\bar{A}$	2	
$AA\bar{A}\bar{A}$	2	
$\bar{A}AAA$	3	$p^3 \cdot q$
$A\bar{A}AA$	3	
$AA\bar{A}A$	3	
$AAA\bar{A}$	3	
$AAAA$	4	p^4

vrijednost 0 za 1 ishod, tj. $\binom{4}{0}$

vrijednost 1 za 4 ishoda, tj. $\binom{4}{1}$

vrijednost 2 za 6 ishoda, tj. $\binom{4}{2}$

vrijednost 3 za 4 ishoda, tj. $\binom{4}{3}$

vrijednost 4 za 1 ishod, tj. $\binom{4}{4}$

Binomna raspodjela vjerojatnosti

Općenito:

- ishod događaja je *AAĀ…ĀA*, ukupno ima *n* događaja
- *A* se pojavljuje *x* puta, a *Ā* (*n-x*) puta
- ukupno postoji $\frac{n!}{x!(n-x)!} = \binom{n}{x}$ permutacija gornjeg niza
- svaki gornji niz ostvaruje se s vjerojatnošću (nezavisni događaji)

$$p^x(1-p)^{n-x} = p^x q^{n-x}$$

Stoga je:

$$b(x;n,p) = \begin{cases} \binom{n}{x} p^x q^{n-x} & , x = 0, 1, 2, \ldots, n \\ 0 & , \text{inače} \end{cases}$$

Primjer:

Neki uređaj daje 30% loših proizvoda. Kolika je vjerojatnost da će od 10 proizvoda njih 3 biti loših?

- vjerojatnost da izvučemo loš proizvod: p = 0.3
- vjerojatnost da izvučemo dobar proizvod: q = 0.7

$n = 10$, $x = 3$

- P(x = 3) - vjerojatnost da od 10 proizvoda 3 budu loša

$$b(3; 10, 0.3)$$

$$P(x=3) = \binom{10}{3} 0.3^3 0.7^7 = 0.2668$$

$$\boxed{P(x) = \binom{n}{x} p^x q^{n-x}}$$

Rekurzivna formula

- računanje koeficijenata $\binom{n}{x}$ nije jednostavno

-zato se služimo rekurzivnom formulom

- rekurzivna formula: izračunavanje vrijednosti vjerojatnosti *P*(*x*) iz već poznate *P*(*x*-1)

- za binomnu raspodjelu vrijedi:

$$\boxed{P(x) = \frac{n-x+1}{x}\frac{p}{q}P(x-1)}$$

Izvod: $P(0) = \binom{n}{0} p^0 q^{n-0} = q^n$

$$\frac{P(x)}{P(x-1)} = \frac{\binom{n}{x} p^x q^{n-x}}{\binom{n}{x-1} p^{x-1} q^{n-x+1}} = \frac{\frac{n!}{x!(n-x)!}}{\frac{n!}{(x-1)!(n-x+1)!}} \cdot \frac{1}{p^{-1}q} = \frac{n-x+1}{x}\frac{p}{q}$$

Očekivanje i varijanca za binomnu raspodjelu

- očekivanje

$$E(x) = \sum_{x=0}^{n} xP(x) = \sum_{x=1}^{n} x\binom{n}{x} p^x q^{n-x}$$

vrijedi sljedeće:

$$\binom{n}{x} = \frac{n(n-1)!}{x(x-1)![(n-1)-(x-1)]!} = \frac{n}{x}\frac{(n-1)!}{(x-1)![(n-1)-(x-1)]!} = \frac{n}{x}\binom{n-1}{x-1}$$

pa imamo:

$$E(x) = \sum_{x=1}^{n} n\binom{n-1}{x-1} p^x q^{n-x} = np\sum_{x=1}^{n}\binom{n-1}{x-1} p^{x-1} q^{n-x} = \left\{x-1 = y, n-1 = m, y_{\max} = n-1 = m, y_{\min} = 0\right\} =$$

$$= np\sum_{y=0}^{m}\binom{m}{y} p^y q^{m-y} = np$$

$$\sum_{y} P(y) = 1$$

Očekivanje i varijanca za binomnu raspodjelu

Konačno, očekivanje binomne raspodjele:

$$\boxed{E(x) = n \cdot p}$$

Varijanca:

$$V(X) = E(X^2) - [E(X)]^2 = E(X)^2 - n^2 p^2$$

prvi član:

$$E(X^2) = \sum_{x=0}^{n} x^2 \cdot b(x; n, p) = \sum_{x=1}^{n} \binom{n}{x} p^x q^{n-x} x^2 = np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} q^{n-x} x$$

uvodimo supstituciju: $y = x - 1,\ m = n - 1$

$$E(X^2) = np \sum_{y=0}^{n} \binom{m}{y} p^y q^{m-y} (y+1) = np \left[\sum_{y=0}^{m} y \binom{m}{y} p^y q^{m-y} + 1 \right]$$

Očekivanje i varijanca za binomnu raspodjelu

$$E\left(X^2\right) = np\left\{1 + E\left(y\right)\right\} = np\left\{1 + mp\right\} = np\left\{1 + np - p\right\} =$$

$$= \left(np\right)^2 + np\left(1 - p\right) = \left[E\left(x\right)\right]^2 + np - np^2$$

Konačno, varijanca je:

$$V\left(X\right) = E\left(X^2\right) - \left[E\left(X\right)\right]^2 = np - np^2 = np\left(1 - p\right) = npq$$

Na sličan način kao za očekivanu vrijednost i varijancu može se pokazati:

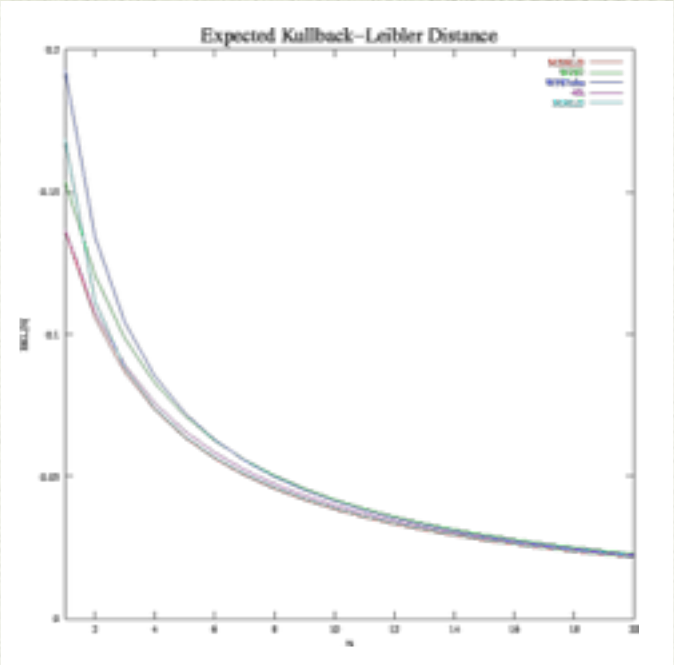
$$\alpha_3 = \frac{q - p}{\sqrt{npq}}; \quad \alpha_4 = 3 + \frac{1 - 6pq}{npq}$$

Izgled binomne raspodjele

(i) može se desiti da raspodjela tvori padajući niz:

P(0) ≥ P(1) ≥ P(2) ≥ ... ≥ P(n)

$$\boxed{\frac{P(x)}{P(x-1)} = \frac{n-x+1}{x}\frac{p}{q}}$$



iz rekurzivne relacije:

$$P(x) \le P(x-1) \quad \Rightarrow \quad \frac{n-x+1}{x}\frac{p}{q} \le 1, \text{za sve } x$$

$$x(p+q) \ge p(n+1) \text{ tj. } p \le \frac{x}{n+1}, \text{za sve } x$$

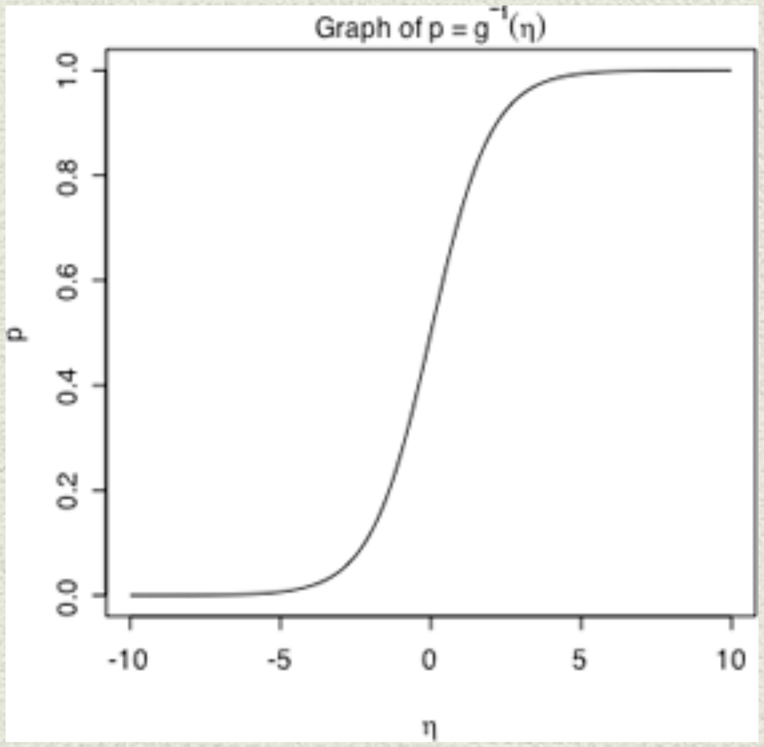
- ova relacija je ispunjena za sve *x* ako je zadovoljena za *x* = 1, pa je uvjet da bi binomne vjerojatnosti činile padajući niz:

$$p \le \frac{1}{n+1}$$

Izgled binomne raspodjele

(ii) može se desiti da raspodjela tvori rastući niz:

$P(0) \leq P(1) \leq P(2) \leq \ldots \leq P(n)$



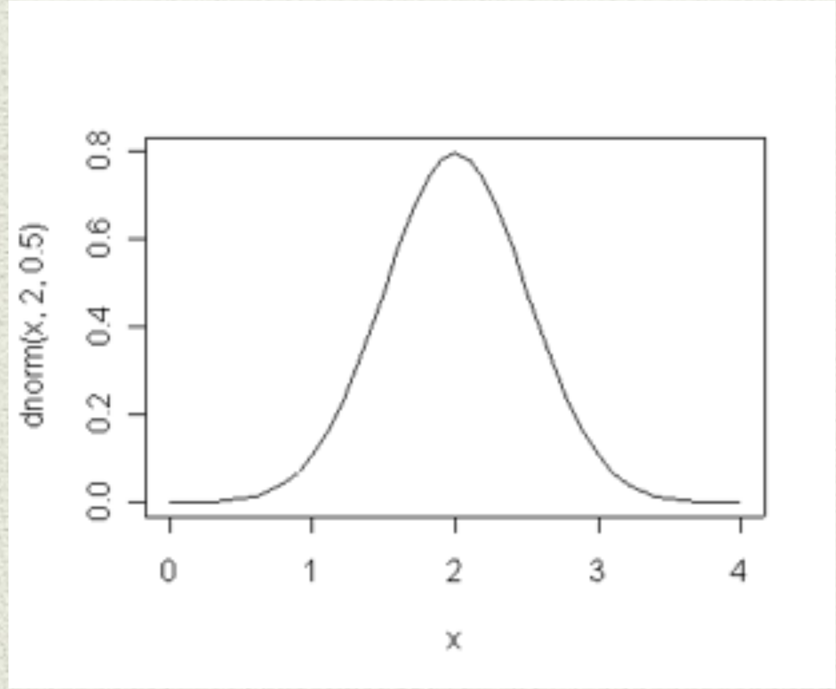
- isto kao i u prethodnom slučaju, dobiva se uvjet:

$$p \geq \frac{n}{n+1}$$

Izgled binomne raspodjele

(iii) kad ne tvori ni rastući, ni padajući niz, binomna raspodjela ima maksimum u nekoj točki x_0, $x_0 \neq 0$, $x_0 \neq n$

$$P(x_0-1) \leq P(x_0) \geq P(x_0+1)$$



iz rekurzivne relacije:

$$P(x_0-1) \leq P(x_0) \quad \Rightarrow \quad qx_0 \leq np - px_0 + p$$

$$x_0 \leq np + p$$

$$P(x_0) \geq P(x_0+1) \quad \Rightarrow \quad x_0 \geq np - q$$

Maksimum je definiran sa: $np - q \leq x_0 \leq np + p$

- ako np-q nije cijeli broj, samo jedan x_0 zadovoljava gornju relaciju
- ako np-q je cijeli broj, postoje dva x_0 koji zadovoljavaju gornju relaciju

Poissonova raspodjela

- Binomna raspodjela polazi od osnovnih postavki kombinatorike i teorije vjerojatnosti, i stoga je egzaktana, tj. vrijedi za proizvoljne p i n
- međutim, kao što smo već spomenuli, česti su problemi s računanjem binomnih koeficijenata, pa zato gledamo njezin graničan slučaj

pretpostavke:

1. $n \to \infty$ (velik broj pokusa)
2. $m = np =$ const., tj. $p = m/n \to 0$ (mala vjerojatnost događaja)

$$P(x,m) = \binom{n}{x} p^x q^{n-x} = \frac{n(n-1)\ldots(n-x+1)}{x!} p^x q^{n-x} =$$

$$= \frac{n(n-1)\ldots(n-x+1)}{x!} \frac{m^x}{n^x} \left(1 - \frac{m}{n}\right)^{n-x} =$$

Poissonova raspodjela

$$P(x,m)=\frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{x-1}{n}\right)}{x!}m^x\left(1-\frac{m}{n}\right)^{n-x}$$

- stavimo $n\to\infty$

$$\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{x-1}{n}\right)\to 1$$

$$\left(1-\frac{m}{n}\right)^{n-x}\to\left(1-\frac{m}{n}\right)^{n}$$

$$\Rightarrow \quad P(x,m)=\frac{m^x}{x!}\lim_{\substack{n\to\infty\\ np=m}}\left(1-\frac{m}{n}\right)^{n}$$

$$e^t=\sum_{n=0}^{\infty}\frac{t^n}{n!}$$

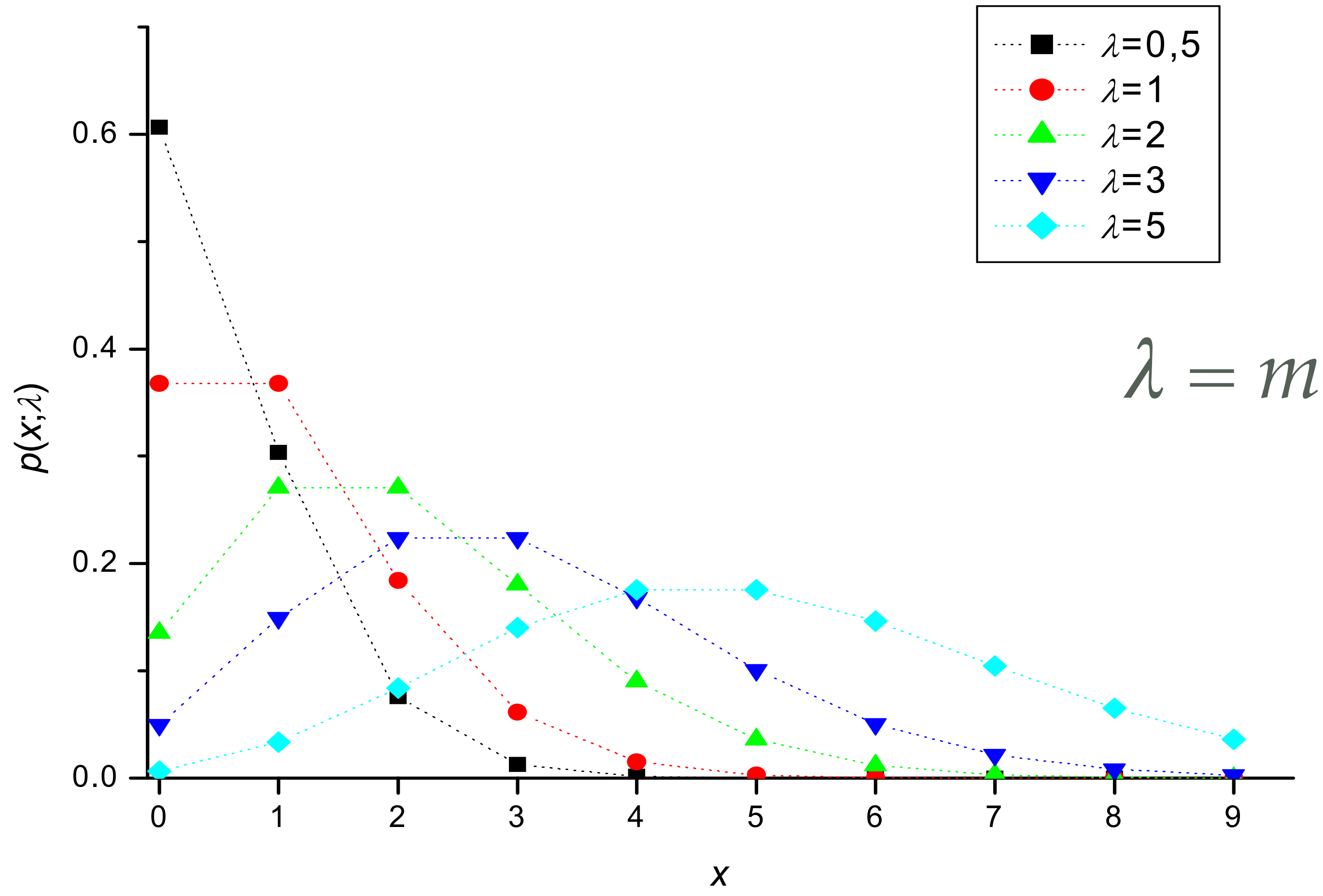
$$e^{-t}=\lim_{n\to\infty}\left(1-\frac{t}{n}\right)^{n}$$

I konačno:

$$P(x,m)=\frac{m^x}{x!}e^{-m}$$

Poissonova raspodjela

$$P(x;m) = \frac{m^x}{x!} e^{-m}$$



$\lambda = m$

Slika 1: Poissonova raspodjela za nekoliko različitih λ.

Poissonova raspodjela

- Poissonova raspodjela je granični slučaj binomne raspodjele za veliki n i mali p

- može se pokazati da je Poissonova raspodjela adekvatna zamjena za binomnu ako je $p \leq 0.1$ i $m = np \leq 5$

- u praksi je često zadovoljavajuća i za veće vrijednosti

Očekivanje i varijanca

- u binomnoj raspodjeli imali smo:

$$E(x) = np \text{ i } V(x) = npq = np(1-p)$$

- ako je $p << 1$ očekivali bismo za Poissonovu raspodjelu:

$$E(x) = np = m \ i \ V(x) = np = m$$

Da li je to uistinu tako?

Poissonova raspodjela

Očekivana vrijednost:

$$E(x) = \{n \to \infty\} = \sum_{x=0}^{\infty} xP(x) = \sum_{x=1}^{\infty} x\frac{m^x}{x!}e^{-m} =$$

$$= e^{-m}\sum_{x=1}^{\infty}\frac{m^{x-1}m}{(x-1)!} = \{x-1=y\} = me^{-m}\sum_{y=1}^{\infty}\frac{m^y}{y!} = me^{-m}e^m = m$$

Podsjetnik:

$$V(x) = E(x^2) - [E(x)]^2$$

Varijanca:

$$E(x^2) = \{n \to \infty\} = \sum_{x=0}^{\infty} x^2P(x) = \sum_{x=1}^{\infty} x^2\frac{m^x}{x!}e^{-m} =$$

$$= \sum_{x=1}^{\infty} xe^{-m}m\frac{m^{x-1}}{(x-1)!} = me^{-m}\left[\sum_{x=1}^{\infty}\frac{m^{x-1}}{(x-1)!} + \sum_{x=1}^{\infty}(x-1)\frac{m^{x-1}}{(x-1)!}\right] =$$

Poissonova raspodjela

$$= me^{-m}\left[\sum_{y=0}^{\infty}\frac{m^y}{y!}+\sum_{x=2}^{\infty}\frac{m^{x-1}}{(x-2)!}\right]= me^{-m}\left[e^m + m\sum_{x=2}^{\infty}\frac{m^{x-2}}{(x-2)!}\right]= me^{-m}\left[e^m + m\sum_{y=2}^{\infty}\frac{m^y}{y!}\right]=$$

$$= me^{-m}\left[e^m + me^m\right] = m + m^2$$

$$V(x) = E(x^2) - [E(x)]^2 = m + m^2 - m^2 = m$$

Sličnim načinom pokazuje se da je:

$$\alpha_3 = \frac{1}{\sqrt{m}}; \quad \alpha_4 = 3 + \frac{1}{m}$$

Poissonova raspodjela - rekurzijska formula

$$\frac{P(x)}{P(x-1)} = \frac{\dfrac{m^x e^{-m}}{x!}}{\dfrac{m^{x-1} e^{-m}}{(x-1)!}} = \frac{m}{x}$$

Dakle, vrijedi: $$P(x) = \frac{m}{x} P(x-1); \quad P(0) = e^{-m}$$

- puno jednostavnije nego u binomnoj raspodjeli!

Zadaća: dokažite da je najvjerojatnija vrijednost x_0 slučajne varijable x u Poissonovoj raspodjeli dana s $(m-1) \leq x_0 \leq m$

Poissonova raspodjela

Primjena Poissonove raspodjele:

a) aproksimacija binomne raspodjele

b) proučavanje rijetkih događaja

npr.: - prometne nesreće na određenoj dionici autoceste u jednom danu
- telefonski pozivi centrali u jednoj minuti
- defekti po jedinici duljine bakrene žice
- broj čestica kozmičkog zračenja detektiranih u sekundi

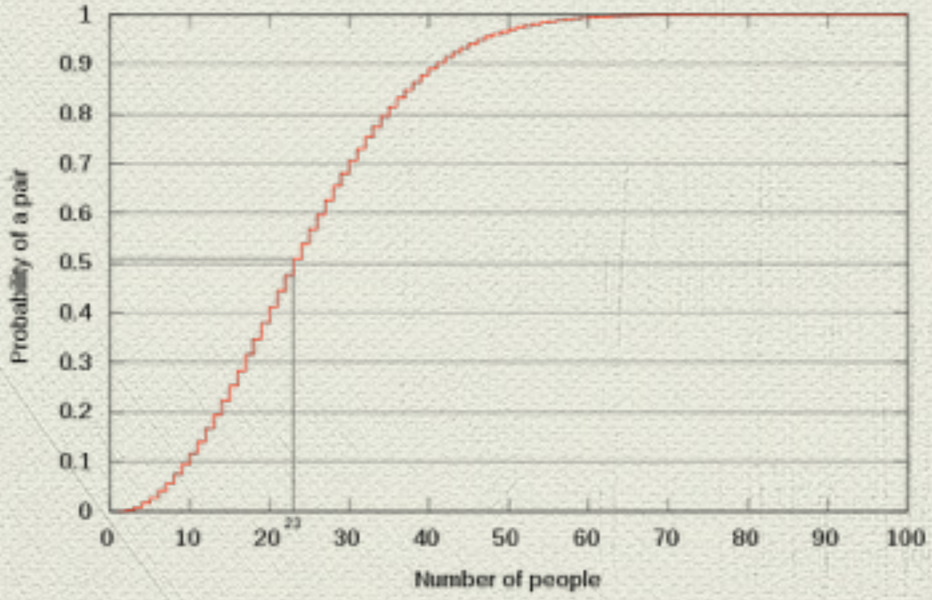
Poissonova raspodjela - primjer

‘Rođendanski’ problem

Kolika je vjerojatnost da u skupu od n nasumično odabranih ljudi dvoje imaju rođendan isti dan?

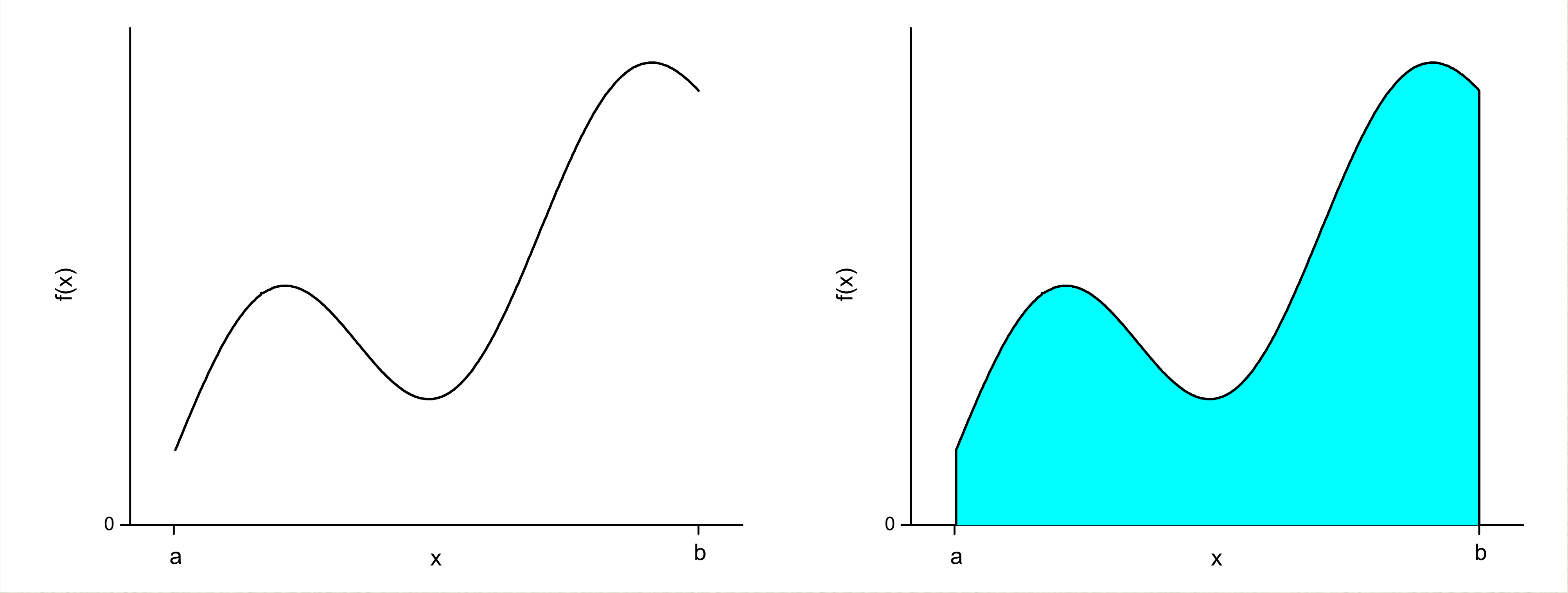
- očito, ukoliko je $n > 366$, vjerojatnost je 100%
- ali, ukoliko je $n = 57$, vjerojatnost je 99%
- za $n = 23$, vjerojatnost je 50%

Provjerite za zadaću



Pojam integrala

Prilikom razmatanja kontinuiranih raspodjela često ćemo se susretati s problemom određivanja površine ispod funkcije.



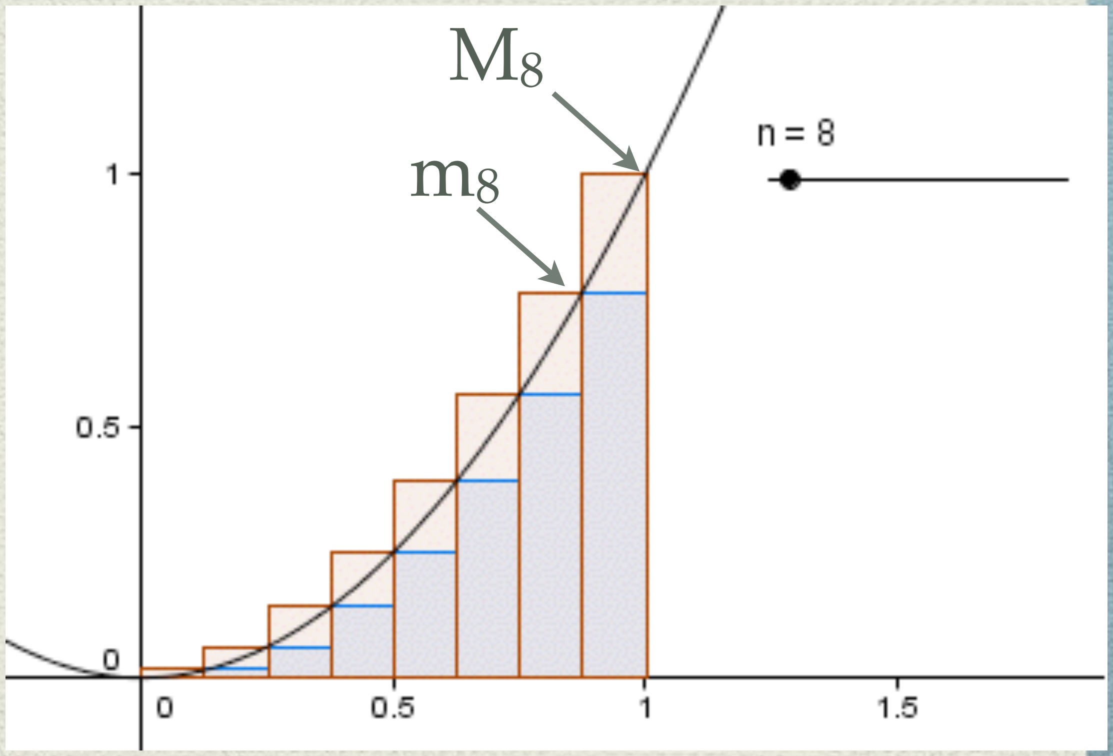
Problem: funkcija je prikazana zakrivljenom crtom i nema jednostavnog geometrijskog načina da točno izračunamo obojenu površinu

Pojam integrala

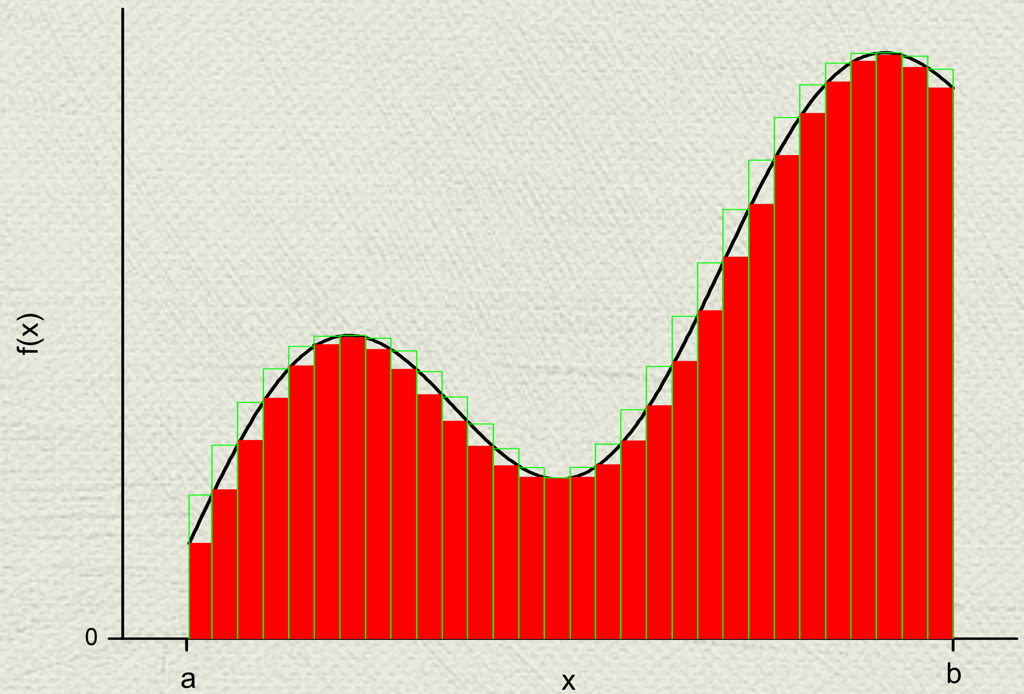
Ideja: promatrana ploha podijeli se na manje dijelove koji se onda aproksimiraju pravokutnicima. Podijelimo promatrani interval $[a, b]$ na n ekvidistantnih segmenata $[x_{k-1}, x_k]$. Pri tome je $x_0 = a$ i $x_n = b$. Na svakom segmentu pronađemo najveću (M_k) i najmanju (m_k) vrijednost funkcije f. Definirajmo sume:

$$s = \sum_{k=1}^{n} m_k \left(x_k - x_{k-1}\right) = \sum_{k=1}^{n} m_k \Delta x$$

$$S = \sum_{k=1}^{n} M_k \left(x_k - x_{k-1}\right) = \sum_{k=1}^{n} M_k \Delta x$$



Pojam integrala



Očito je tražena površina *I* veća od *s*, a manja od *S*. Što je veći broj segmenata *n* na koje smo podijelili promatrani interval, to je razlika između *s* i *S* manja, pa je procjena iznosa površine *I* to bolja.

Ako je promatrana funkcija *f* neprekinuta i konačna na promatranom intervalu, onda će se u granici $n \to \propto$ sume *s* i *S* izjednačiti.

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} S_n = I$$

$$\boxed{I = \lim_{n\to\infty} \sum_{k=1}^{n} f\left(x_k\right)\Delta x = \int_{[a,b]} f\left(x\right)\mathrm{d}x = \int_a^b f\left(x\right)\mathrm{d}x}$$

Pojam integrala

Def.: Primitivna funkcija od *f* je svaka funkcija F sa svojstvom

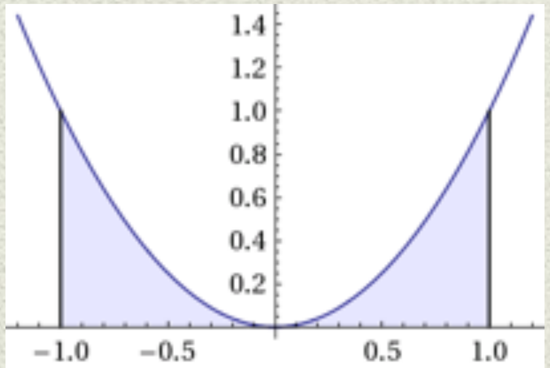
$$F'(x) = f(x)$$

Sve primitivne funkcije neke funkcije *f* razlikuju se samo za konstantu!

Za neprekidne funkcije vrijedi Leibniz-Newtonova formula:

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

Primjer: $\int_{-1}^{1} x^2\,dx = \left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{(1)^3}{3} - \frac{(-1)^3}{3} = \frac{2}{3}$



Pojam integrala

Tablica osnovnih “tabličnih” primitivnih funkcija:

$f(x)$	$F(x)$
$x^n \quad , \quad n \neq -1$	$\dfrac{x^{n+1}}{n+1}$
x^{-1}	$\ln\lvert x\rvert$
e^x	e^x
a^x	$\dfrac{a^x}{\ln a}$
$\sin x$	$-\cos x$
$\cos x$	$\sin x$
$\dfrac{1}{\cos^2 x}$	$\mathrm{tg}\, x$
$\dfrac{1}{\sin^2 x}$	$-\mathrm{ctg}\, x$
$\vdots$	$\vdots$

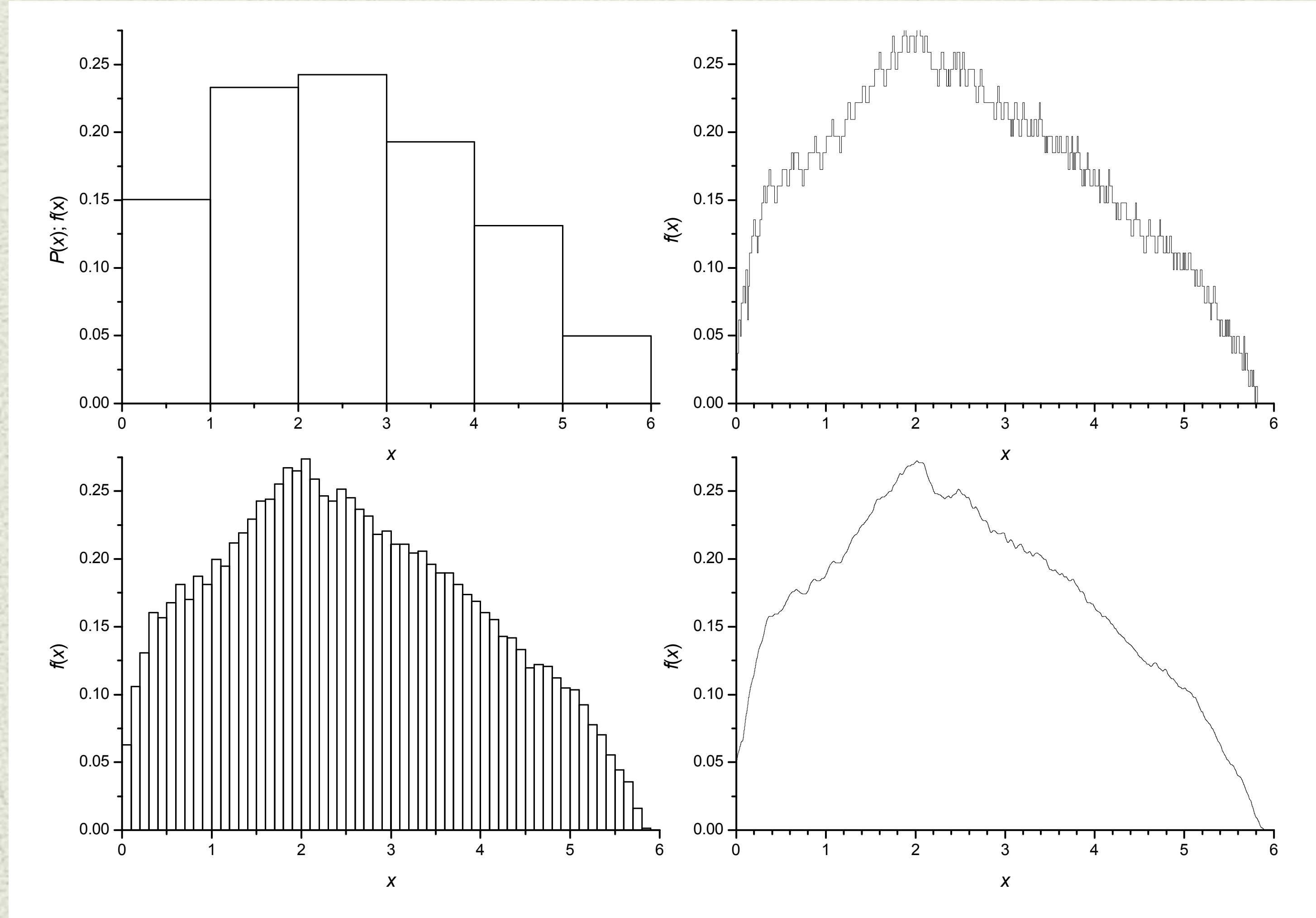
Kontinuirana slučajna varijabla i raspodjela vjerojatnosti

- do sada smo razmatrali diskretne slučajne varijable
- moguće vrijednosti čine konačan skup ili se mogu služiti u uredan beskonačan niz
- u slučaju kontinuirane slučajne varijable, moguće vrijednosti čine cijeli interval brojeva

Def.: Slučajna varijabla X je kontinuirana ako skup njezinih vrijednosti čini cijeli interval brojeva, tj. ako je za neke A i B ($A < B$), svaki $x \in [A, B]$ moguć.

Kontinuirana slučajna varijabla i raspodjela vjerojatnosti

Primjer: mjerimo dubinu nekog jezera na 8115 mjesta. Slike prikazuju: 1. raspodjelu dubina zaokruzenih na cijele metre; 2. zaokruženih na cijele decimetre; 3. zaokruženih na cijele centimetre; 4. kontinuiranu raspodjelu dubina



Za svaki interval Δx definiramo funkciju $f\,(x)$ na sljedeći način:

$P(x \leq X \leq x+\Delta x) = f(x)\Delta x$

Na četvrtoj slici intervali su infinitezimalni ($\Delta x \rightarrow \mathrm{d}x$)

$P(x \leq X \leq x+\Delta x) = f(x)\mathrm{d}x$

Površina ispod svih grafova = 1!

Funkcija gustoće vjerojatnosti

Def.: Neka je *X* kontinuirana varijabla. Raspodjela vjerojatnosti ili funkcija gustoće vjerojatnosti varijable *X* je funkcija *f(x)* takva da za bilo koja dva broja $a \leq b$ vrijedi:

$$P(a \leq X \leq b) = \int_a^b f(x)\mathrm{d}x$$

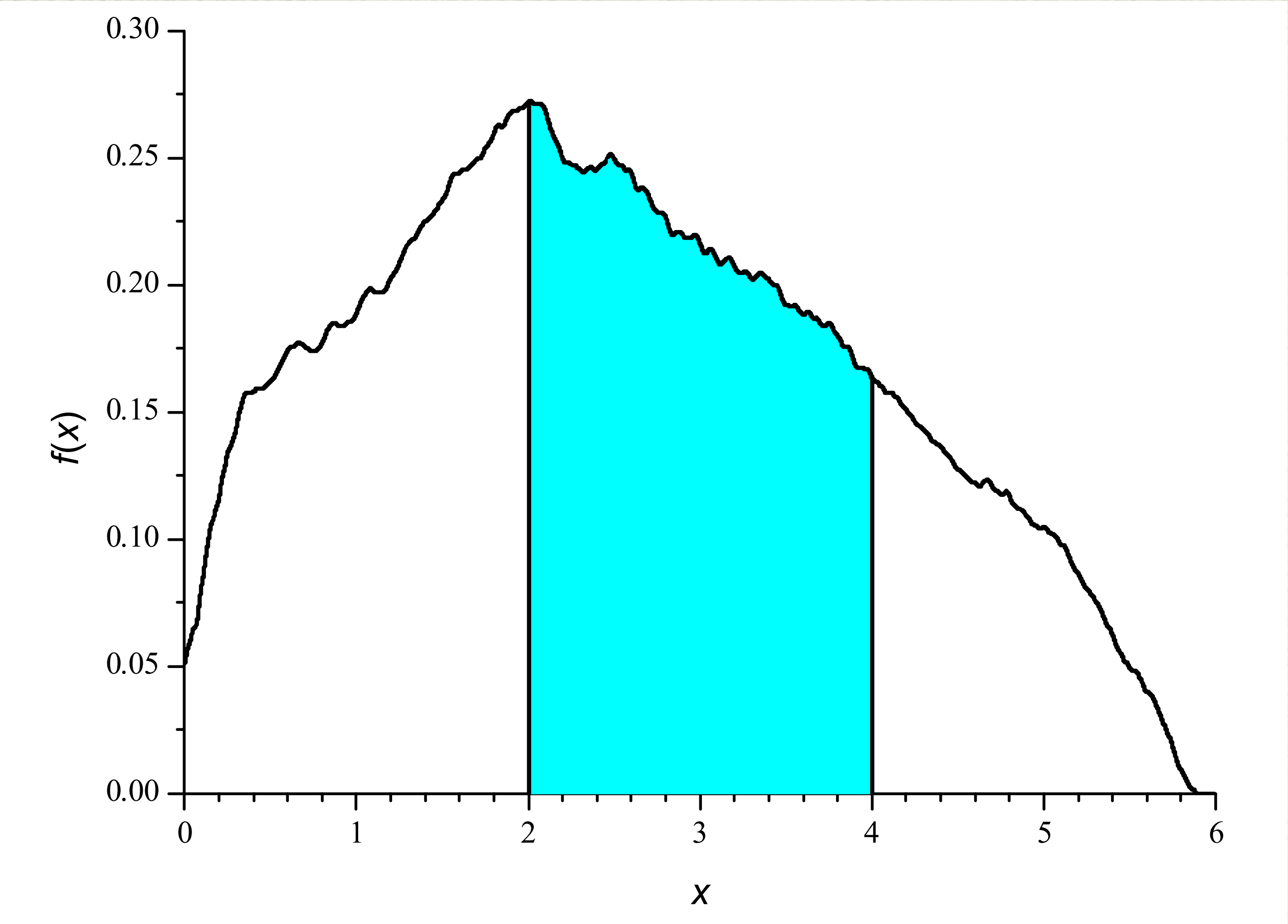
tj. vjerojatnost da *X* poprimi vrijednost u intervalu [*a*, *b*] dana je površinom ispod funkcije gustoće vjerojatnosti u tom intervalu.

Vrijedi sljedeće:

1. $f(x) \geq 0, \forall x$
2. $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$

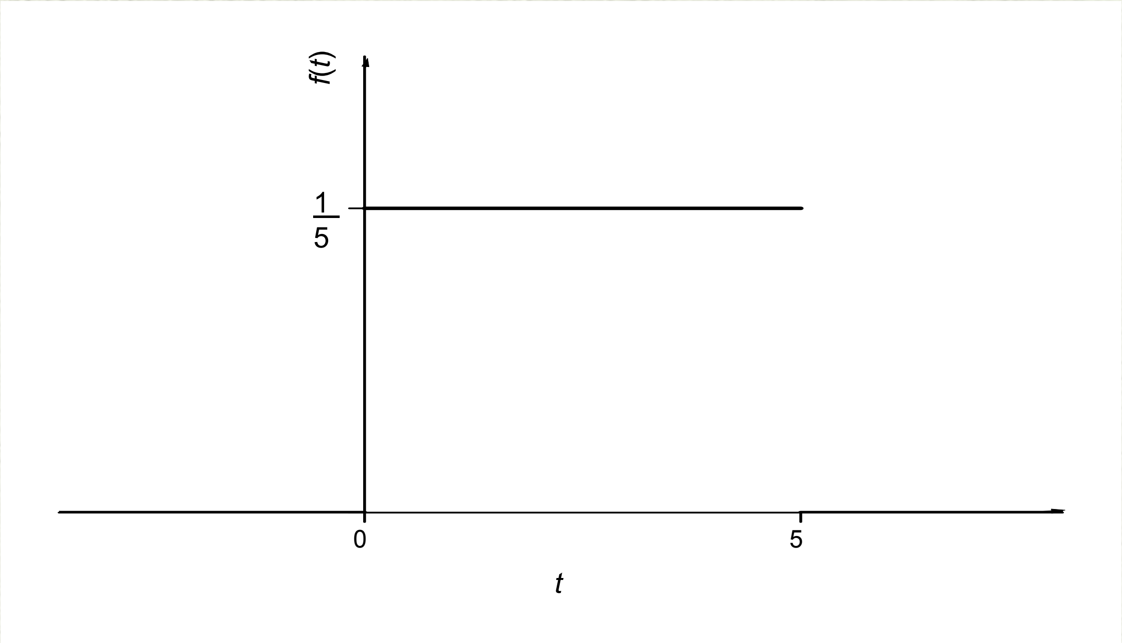
Funkcija gustoće vjerojatnosti

Primjer: za raspodjelu dubine jezera - $P\left(2 \leq X \leq 4\right) = \int_{2}^{4} f\left(x\right) \mathrm{d}x$



Funkcija gustoće vjerojatnosti

Primjer: čekanje tramvaja koji vozi svakih 5 min



$$P\left(X \leq 3\right) = \int_{-\infty}^{3} f\left(x\right) \mathrm{d}x = \frac{3}{5}$$

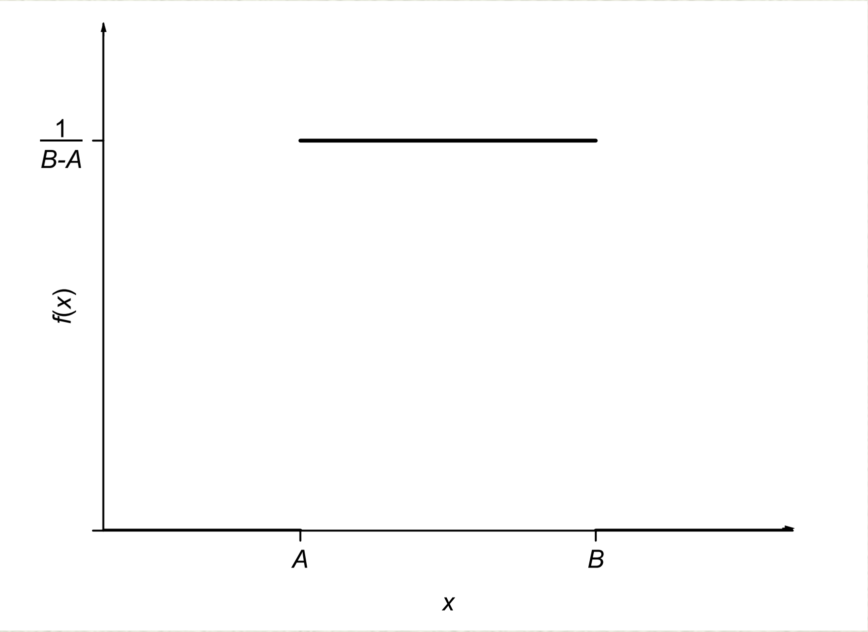
Uniformna ili pravokutna raspodjela!

Funkcija gustoće vjerojatnosti

Def.: Kontinuirana slučajna varijabla *X* ima uniformnu raspodjelu u intervalu [*A*, *B*] ako je funkcija gustoće vjerojatnosti dana s

$$f(x) = \begin{cases} \dfrac{1}{B - A} & A \le X \le B \\ 0 & \text{inace} \end{cases}$$

Ilustracija pravokutne raspodjele:



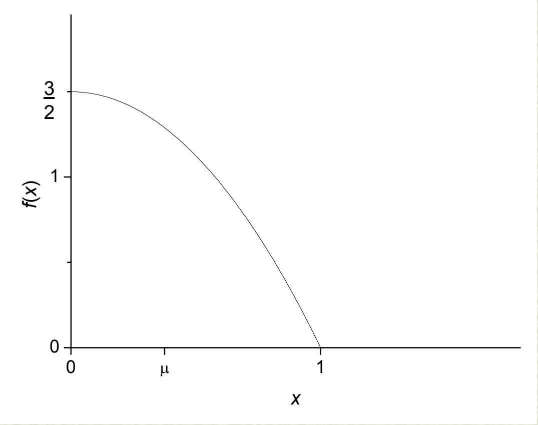
Očekivanje kontinuirane slučajne varijable

Def.: Neka je *X* kontinuirana varijabla i *f(x)* njezina funkcija gustoće vjerojatnosti. Onda je njezino očekivanje

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

Primjer: neka je

$$f(x) = \begin{cases} \frac{3}{2}\left(1 - x^2\right) & 0 \leq x \leq 1 \\ 0 & \text{inace} \end{cases}$$



Očekivanje je:

$$E(X) = \int_0^1 x \cdot f(x) \mathrm{d}x = \frac{3}{2} \int_0^1 \left(x - x^3\right) \mathrm{d}x = \frac{3}{4} - \frac{3}{8} = \frac{3}{8}$$

Varijanca kontinuirane slučajne varijable

Def.: Neka je X kontinuirana varijabla i $f(x)$ njezina funkcija gustoće vjerojatnosti, a μ njezino očekivanje. Onda je njezina varijanca dana relacijom:

$$V(X) = \sigma_X^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\,\mathrm{d}x = E\left((X-\mu)^2\right)$$

Standardna devijacija je $\sigma_X = \sqrt{V(X)}$

Vrijedi: $V(X) = E\left(X^2\right) - \left(E(X)\right)^2$

Momenti:

$$m_r = \int_{-\infty}^{\infty} x^r f(x)\,\mathrm{d}x$$

$$M_r = \int_{-\infty}^{\infty} (x-\mu)^r f(x)\,\mathrm{d}x$$

Normalna (Gaussova) raspodjela

- premda postoje mnoge druge raspodjele, normalna raspodjela objašnjava najveći broj statističkih opažanja
- npr. raspodjela visina odraslih ljudi ili rezultati mjerenja fizikalnih veličina

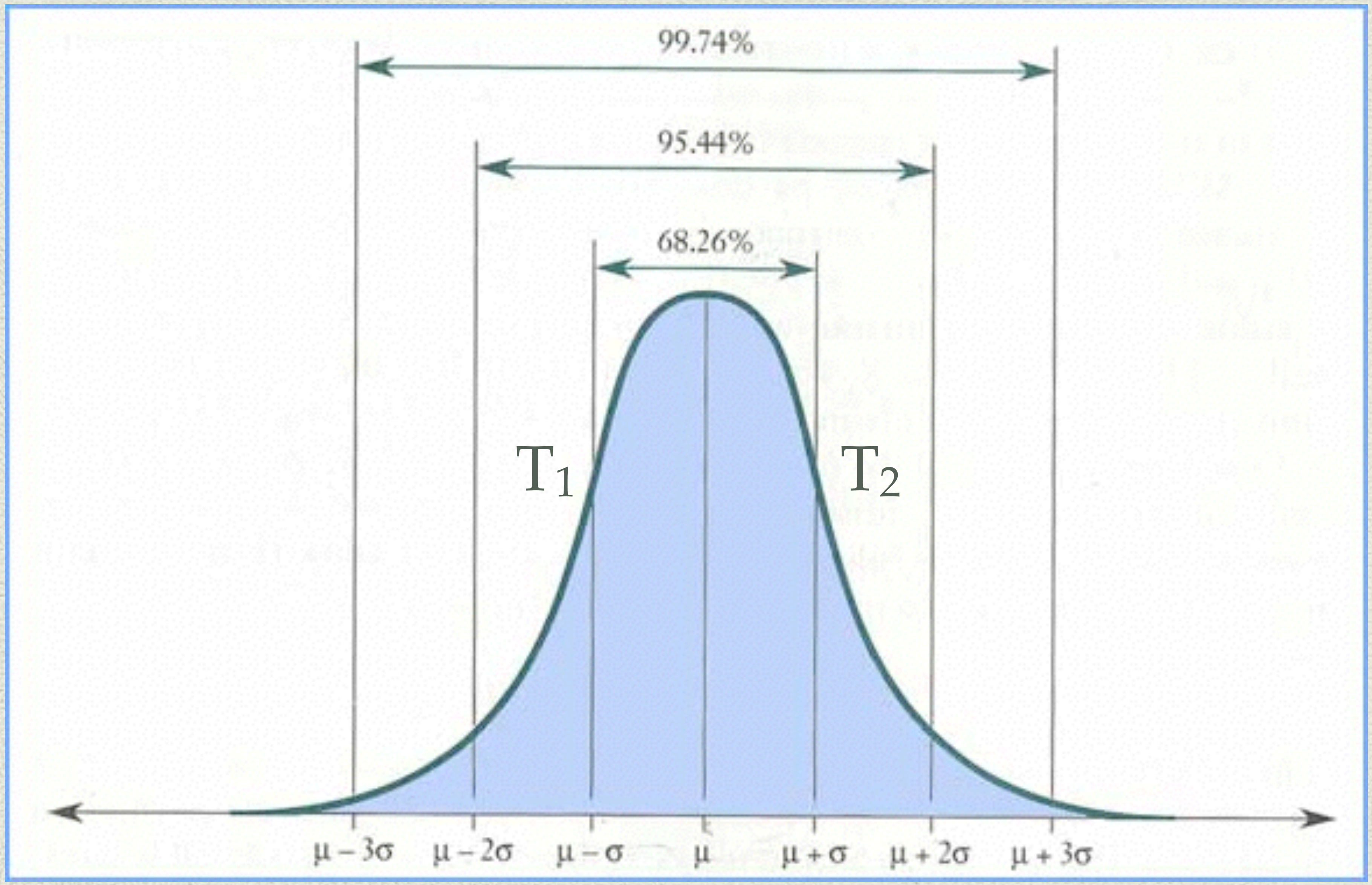
Def.: Za slučajnu varijablu *X* kažemo da je distribuirana po zakonu normalne vjerojatnosti ako je područje njenih vrijednosti $(-\infty, +\infty)$, a funkcija gustoće vjerojatnosti

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

- gdje je $\mu = E(x)$, $\sigma = \sqrt{V(x)}$

- također je $\alpha_3 = 0$ (simetrična raspodjela)
$\alpha_4 = 3$ (normalna spljoštenost)

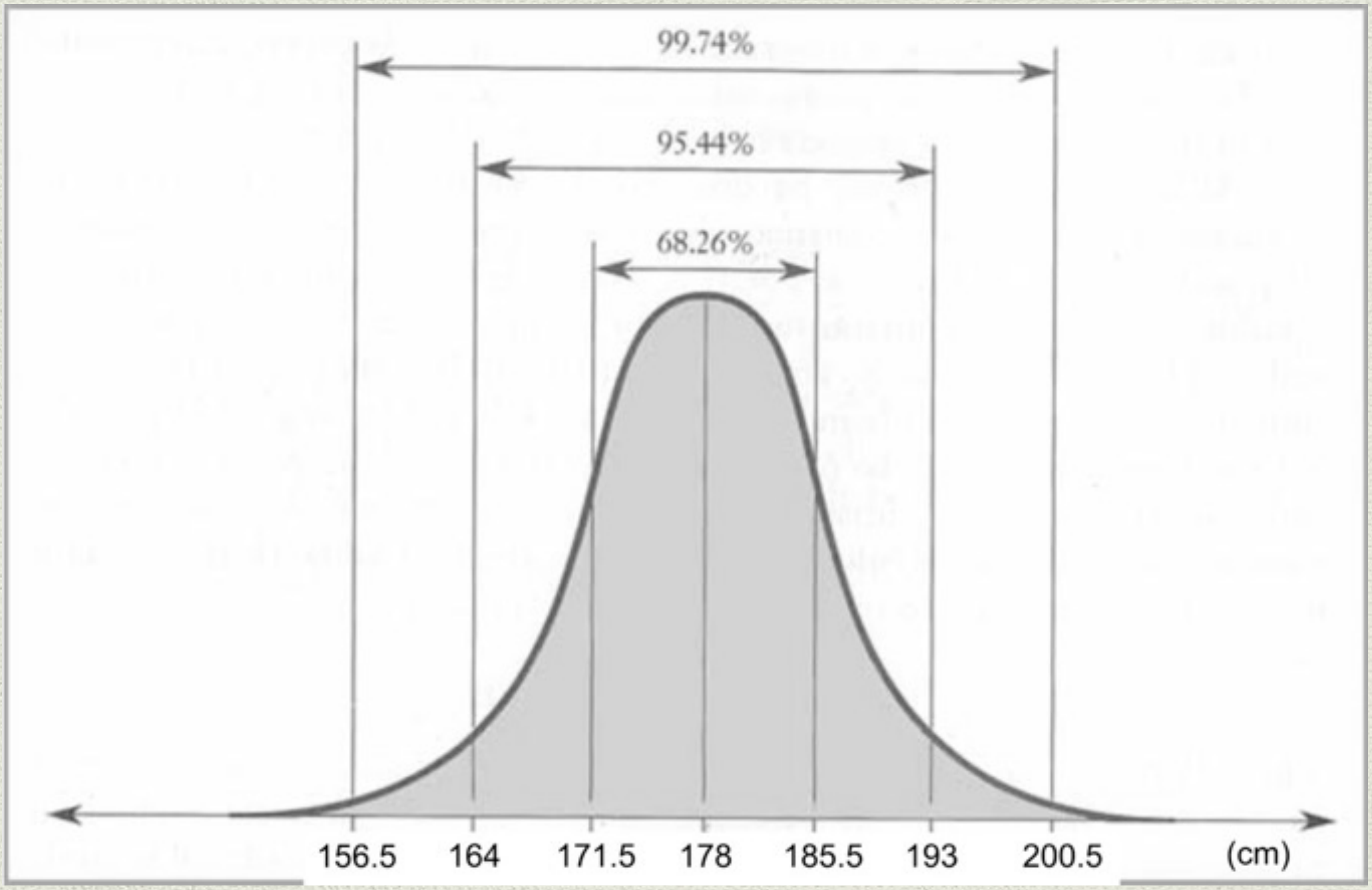
Normalna (Gaussova raspodjela)



T1 i T2 - točke infleksije!

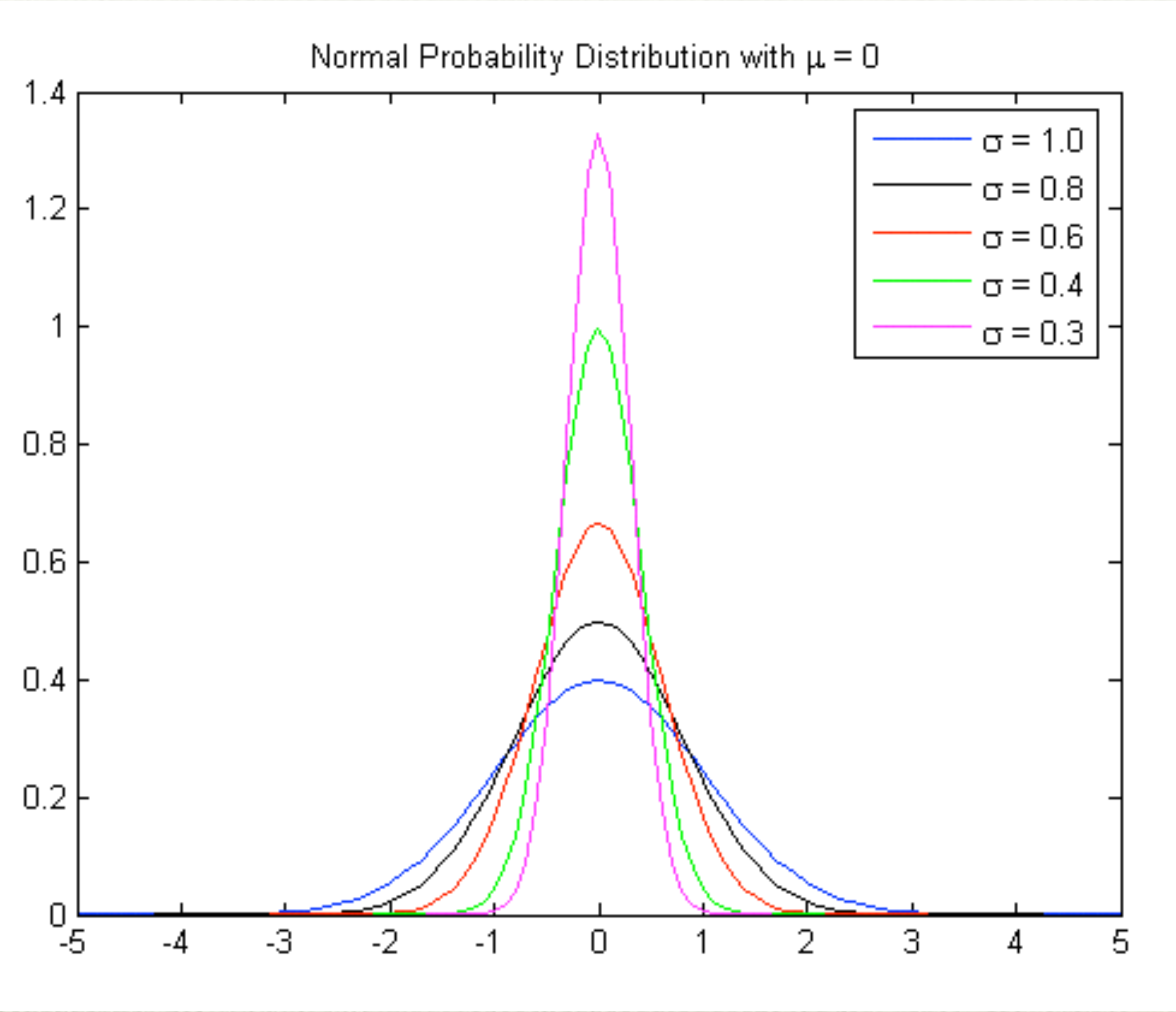
Npr.: prosječna visina muškaraca u Hr 178 cm
st. devijacija 7.5 cm

Normalna (Gaussova raspodjela)



Npr.: prosječna visina muškaraca u Hr 178 cm
st. devijacija 7.5 cm

Normalna (Gaussova raspodjela)



Što veći σ to je funkcija gustoće vjerojatnosti šira!

Vjerojatnost pri normalnoj raspodjeli

Znamo otprije:

$$P\left(x_1 < x < x_2\right) = \int_{x_1}^{x_2} f\left(x\right)\mathrm{d}x = F\left(x_2\right) - F\left(x_1\right)$$

U slučaju Gaussove raspodjele:

$$P\left(x_1 < x < x_2\right) = \frac{1}{\sigma\sqrt{2\pi}}\int_{x_1}^{x_2} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\mathrm{d}x$$

uvodimo sljedeće supstitucije:

$$u = \frac{x-\mu}{\sigma}; \quad \mathrm{d}u = \frac{1}{\sigma}\mathrm{d}x$$

$$u_1 = \frac{x_1-\mu}{\sigma}; \quad u_2 = \frac{x_2-\mu}{\sigma}$$

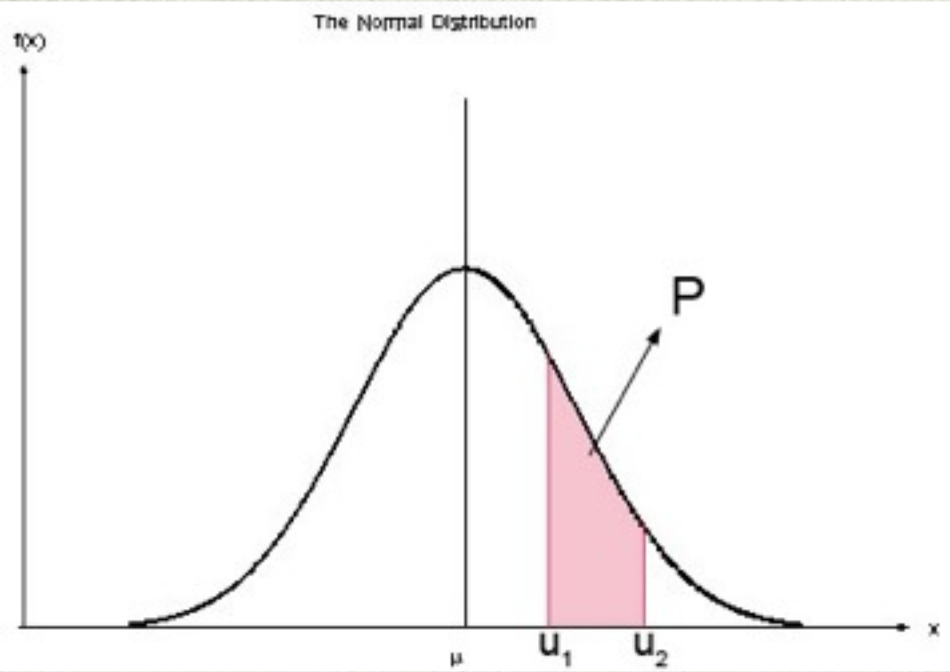
Vjerojatnost pri normalnoj raspodjeli

To daje:

$$P\left(x_1 < x < x_2\right) = \frac{1}{\sqrt{2\pi}} \int_{u_1}^{u_2} e^{-\frac{1}{2}u^2}\, \mathrm{d}u$$

Kako riješiti gornji izraz?

1) $u_1, u_2 \geq 0$ $(x_1, x_2 \geq \mu)$

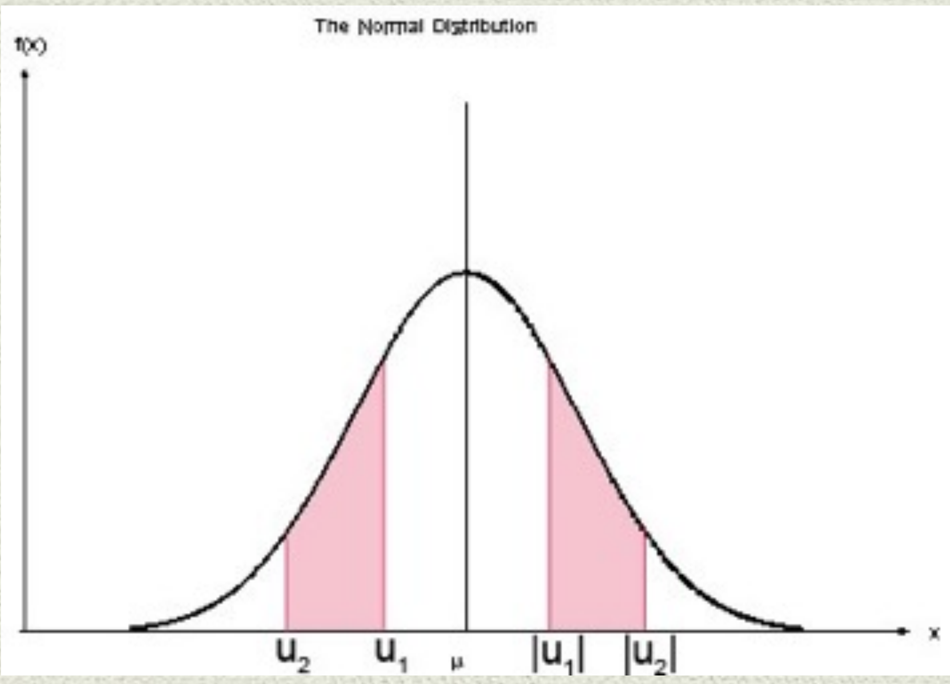


$$\boxed{P = \Phi\left(u_2\right) - \Phi\left(u_1\right)}$$

$$\Phi\left(t\right) = \frac{1}{\sqrt{2\pi}} \int_{0}^{t} e^{-\frac{1}{2}u^2}\, \mathrm{d}u$$

Vjerojatnost pri normalnoj raspodjeli

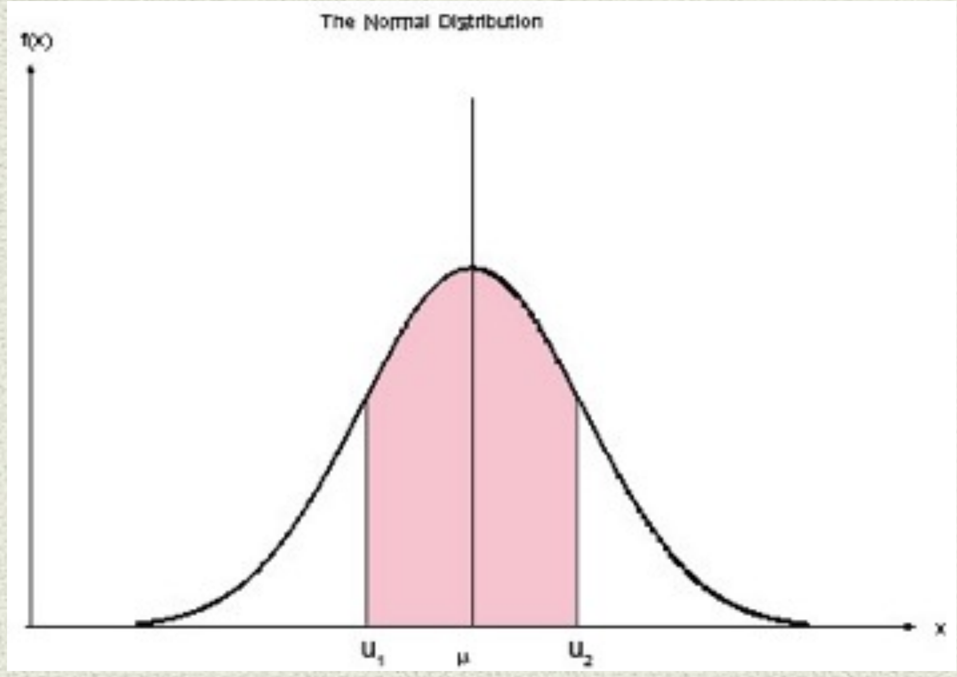
2) $u_1, u_2 \leq 0$ $(x_1, x_2 \leq \mu)$



$$\boxed{P = \Phi(|u_2|) - \Phi(|u_1|)}$$

Vjerojatnost pri normalnoj raspodjeli

3) $u_1 \leq 0$, $u_2 \geq 0$ ($x_1 \leq \mu$, $x_2 \geq \mu$)



$$\boxed{P = \Phi\left(|\,u_1\,|\right) + \Phi\left(|\,u_2\,|\right)}$$

Vjerojatnost pri normalnoj raspodjeli

Svojstva funkcije $\phi(t)$

(i) $\phi(-t) = -\phi(t)$

$$\Phi(-t) = \frac{1}{\sqrt{2\pi}}\int_0^{-t} e^{-\frac{1}{2}u^2}\,\mathrm{d}u = \{u = -v, \mathrm{d}u = -\mathrm{d}v\} = -\frac{1}{\sqrt{2\pi}}\int_0^{t} e^{-\frac{1}{2}v^2}\,\mathrm{d}v = -\Phi(t)$$

Uz to svojstvo, prethodna tri slučaja postaju:

$$P(x_1 < x < x_2) = \Phi(u_2) - \Phi(u_1)$$

Vjerojatnost pri normalnoj raspodjeli

(ii) $\phi(\infty) = 1/2$

To je intuitivno jasno jer $P\left(0 < x < \infty\right) = \dfrac{1}{2}$

$$\Phi\left(\infty\right) = \frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-\frac{1}{2}u^2}\,\mathrm{d}u = \left\{\frac{u^2}{2} = z, \mathrm{d}u = \sqrt{2}\mathrm{d}z\right\} = \frac{1}{\sqrt{\pi}}\int_0^\infty e^{-z^2}\,\mathrm{d}z$$

$$\Phi^2 = \frac{1}{\pi}\int_0^\infty\int_0^\infty e^{-x^2}e^{-y^2}\,\mathrm{d}x\,\mathrm{d}y = \left\{\begin{array}{c} x^2 + y^2 = r^2, x = r\cos\varphi, y = r\sin\varphi, \mathrm{d}x\mathrm{d}y = r\mathrm{d}r\mathrm{d}\varphi \\ 0 < x,y < \infty \Rightarrow 0 < r < \infty, \quad 0 < \varphi < \dfrac{\pi}{2} \end{array}\right\}$$

$$\Phi^2 = \frac{1}{\pi}\int_0^{\frac{\pi}{2}}\mathrm{d}\varphi\int_0^\infty e^{-r^2}r\,\mathrm{d}r = \frac{1}{\pi}\times\frac{\pi}{2}\times\int_0^\infty e^{-r^2}\,\frac{\mathrm{d}\left(r^2\right)}{2} = \frac{1}{4}\int_0^\infty e^{-s}\,\mathrm{d}s = \frac{1}{4}$$

$$\Phi = \sqrt{\Phi^2} = \frac{1}{2}$$

Vjerojatnost pri normalnoj raspodjeli

(iii) $\phi(0) = 0$ - očito

Sumirajmo svojstva normalne raspodjele

- $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
- $\mu = E(x), \sigma = V(x), \alpha_3 = 0, \alpha_4 = 0$
- $P(x_1 < x < x_2) = \Phi\left(\dfrac{x_2 - \mu}{\sigma}\right) - \Phi\left(\dfrac{x_1 - \mu}{\sigma}\right)$
- $\Phi(t) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_0^t e^{-\frac{1}{2}u^2} \, \mathrm{d}u$
- $\Phi(-t) = \Phi(t); \Phi(0) = 0; \Phi(\infty) = \dfrac{1}{2}$

MURPHY'S LAW

What can go wrong, will go wrong.

Essentially, the laws of nature always work, whether we are paying attention or not.

(Equipment blows to protect fuses.)

(Interchangeable parts aren't & fail-safes don't.)

Mrs MURPHY'S COROLLARY

Murphy is too much of an optimist.