Teorija verovatnoće i teorijski rasporedi verovatnoća

Predavač: Dr Mirko Savić

savicmirko@ef.uns.ac.rs

www.ef.uns.ac.rs

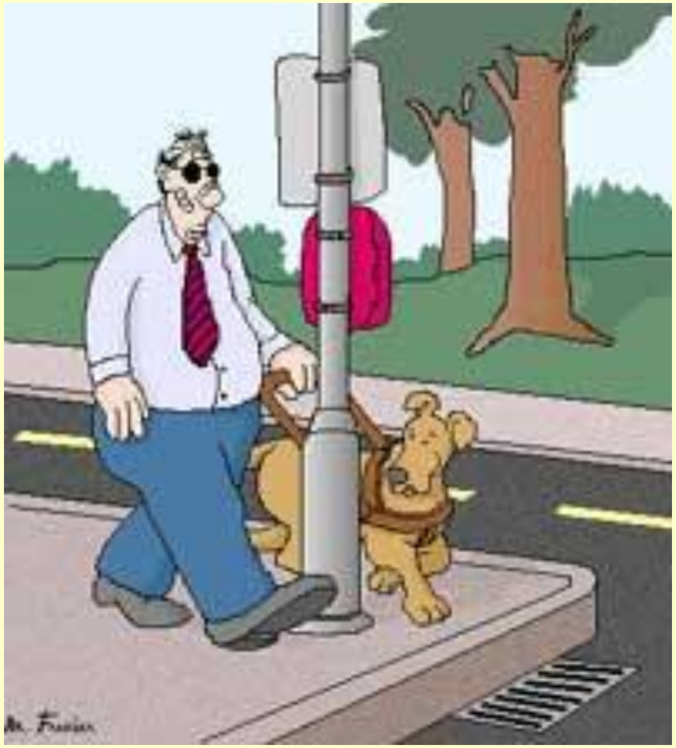
Šta je podstaklo razvoj teorije verovatnoće?

Blaise Pascal

Osnovni pojmovi

Teorija verovatnoće je matematička disciplina koja izučava zakonitosti slučajnih pojava.

Slučajni događaj je onaj događaj čija se realizacija ne može pouzdano predvideti.



Skup svih elementarnih događaja se naziva sigurni događaj:

$\Omega=\{ \omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$

Verovatnoće slučajnog događaja

Verovatnoće mogu biti:
- klasična verovatnoća*,
- geometrijska verovatnoća,
- statistička verovatnoća,
- aksiomatska verovatnoća.
- itd

Klasična (Laplasova) verovatnoća

$$p(A) = \frac{m}{n}$$

gde je:
A – slučajni događaj kao skup povoljnih elementarnih događaja D_1, D_2,..., D_n,
m – broj povoljnih realizacija događaja A,
n – broj eksperimenata.

Suprotna verovatnoća

$$q(A) = 1 - p(A)$$

Osobine klasične verovatnoće:

Klasična verovatnoća je uvek nenegativna vrednost: $p(A) \geq 0$.

Ako je $m=0$, događaj je nemoguć: $p(A)=0$.

Ako je $m=n$, onda je događaj A pouzdan: $p(A)=1$.

Vrednost klasične verovatnoće je u granicama: $0 \leq p(A) \leq 1$.

Verovatnoća da se događaj A ne realizuje je suprotna verovatnoća: $q(A)=1-\frac{m}{n}$.

Statistička definicija verovatnoće

Ako je n broj ponavljanja eksperimenta a m broj uspešnih realizacija događaja A, tada relativna frekvencija m/n predstavlja statističku verovatnoću događaja A, tj

$$P(A) = \frac{m}{n}.$$

Primer. Kockica se baca 1200 puta. Rezultat eksperimenata prikazan je u tabeli:

x_i	1	2	3	4	5	6
f_i	205	196	193	195	207	204

Ako događaj A predstavlja „kockica pokazuje paran broj“, onda je

$$P(A) = \frac{196+195+204}{1200} = \frac{595}{1200} \approx \frac{1}{2}$$

Nezavisni događaji

Definicija: Događaji A i B su nezavisni ako važi

$$P(A \cap B) = P(A) \cdot P(B)$$

Primer. Kockica se baca dva puta. Koliko iznosi verovatnoća da prvi put pokazuje broj 5 a drugi put broj 2? (Da li se logika menja ako prvi put pokazuje broj 5 a drugi put broj 5?)

Uslovna verovatnoća

$$P(A|B)=\frac{P(A\cap B)}{P(B)},\ P(B)>0$$

Princip proizvoda za uslovne verovatnoće.

$$P(AB)=P(B)\cdot P(A|B)$$ ili

$$P(AB)=P(A)\cdot P(B|A)$$

Primer: U kutiji ima *r* crvenih lopti i *b* plavih.
a) Koliko iznosi verovatnoća da pri izboru dve lopte bez vraćanja, prva izabrana lopta bude crvena a druga plava.
b) Koliko iznosi verovatnoća da pri izboru dve lopte bez vraćanja, prva izabrana lopta bude plava a druga crvena.

a) $p=\frac{r}{b+r}\cdot\frac{b}{b+r-1}$; b) $p=\frac{b}{b+r}\cdot\frac{r}{b+r-1}$

Bejsova formula

Iz relacije

$$P(A)\cdot P(B\mid A)=P(AB)=P(B)\cdot P(A\mid B)$$

se lako izvodi Bejsova formula:

$$P(A\mid B)=\frac{P(B\mid A)P(A)}{P(B)}.$$

U ekonomiji se Bejsova formula često koristi za ažuriranje verovatnoće novim informacijama.

Primer. Pretpostavimo da cena akcija raste u 900 dana i pada u 1100. Sa kojom verovatnoćom cena akcija pada? ($P(A) = 0{,}55$ – a priori verovatnoća)
Pretpostavimo da raspolažemo sa dodatnim informacijama o kretanju kamatne stope:

Cena akcija	Kamatna stopa	
	Pada	Raste
Pada	200	900
Raste	800	100

Sa kojom verovatnoćom cena akcija pada ako imamo informaciju da kamatna stopa raste?

Označimo: B – kamatna stopa raste,

$\overline{B}$ – kamatna stopa pada,

A – cena akcija pada.

Treba izračunati $P(A \mid B) = ?$

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}.$$

$P(B) = 0{,}5;$

$P(B \mid A) = (900/1100) = 0{,}81818$

$P(A \mid B) = (0{,}81818 \cdot 0{,}55)/0{,}5 = 0{,}9$ – a posteriori verovatnoća

Permutacije i kombinacije

Permutacija *k* elemenata iz skupa *n* elemenata predstavlja uređenu sekvencu *k* elemenata. Broj permutacija računa se prema obrascu

$$P_n^k = \frac{n!}{(n-k)!} = n \cdot (n-1) \cdot \ldots \cdot (n-k+1)$$

Redosled je bitan!

Primer. Čovek svakog dana čita novine „A“, „B“, „C“, „D“ . Na koliko načina se može odrediti redosled čitanja

a) sve 4 novine,
b) 3 novine,
c) 2 novine,
d) 1 novine?

Rešenje:

a) $P_4^4 = 24;$

b) $P_4^3 = 24;$

c) $P_4^2 = 12;$

d) $P_4^1 = 4.$

Primer: Na koliko načina se 5 ljudi može rasporediti u jednom vozilu?

$$P = 120$$

Primer: Na koliko načina se 5 ljudi može rasporediti u jednom vozilu s tim da se vozač ne menja?

$$P = 24$$

Kombinacija podrazumeva formiranje podskupa od k elemenata iz skupa od n elemenata ($0 \leq k \leq n$), pri čemu redosled izabranih elemenata nije bitan. Broj kombinacija određuje se prema obrascu

$$C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k \cdot (k-1) \cdot \ldots \cdot 1}.$$

Primer. U prethodnom primeru na koliko načina se može izabrati X novina koje će biti pročitane (redosled nije bitan), ako je

a) $X = 4$,
b) $X = 3$,
c) $X = 2$,
d) $X = 1$,
e) $X = 0$?

Rešenje:

a) $C_4^4 = \binom{4}{4} = 1$; b) $C_4^3 = \binom{4}{3} = 4$; c) $C_4^2 = \binom{4}{2} = 6$; d) $C_4^1 = \binom{4}{1} = 4$; d) $C_4^0 = \binom{4}{0} = 1$.

Verovatnoća određenog broja realizacija događaja pri višestrukom ponavljanju eksperimenta

- binomna verovatnoća*,
- Puasonova (Poisson) verovatnoća*,
- hipergeometrijska verovatnoća.

Binomna verovatnoća

Pokazuje kolika je verovatnoća da će se u n nezavisnih eksperimenata događaj A realizovati x puta, a da se neće realizovati $n-x$ puta, pod uslovom da je $0 \leq x \leq n$.

Kada se izračunava verovatnoća $p(n;x)$:

- kada je verovatnoća realizacije događaja A, $p(A)$, velika ($p(A)>0,05$)
- broj nezavisnih eksperimenata mali ($n \leq 20$)

Formule:

Binomna verovatnoća: $p(n;x)=\binom{n}{x}p^x q^{n-x};\ x=0,1,2,\dots,n,$

Rekurentni obrazac: $p(n;x)=\frac{n-x+1}{x}\cdot\frac{p}{q}\cdot p_{(n;x-1)},$

Binomna verovatnoća ako je $p=q$: $p(n;x)=\frac{\binom{n}{x}}{2^n},$

Binomni obrazac x-te klase od n elemenata: $\binom{n}{x}=\frac{x!}{x!(n-x)!},$

Napomena za binomni obrazac: $\binom{n}{n}=\binom{n}{0}=1.$

VER-040
Binomna verovatnoća

VER-004 K:3-1
Binomna verovatnoća

VER-027 Z(06)5-1
Binomna verovatnoća

Puasonova verovatnoća

Puasonova verovatnoća se izračunava kada je verovatnoća $p \leq 0,05$ i $n > 20$.

Puasonova verovatnoća: $p(n;x) = \frac{m^x}{x!} e^{-m};\ x = 0, 1, 2, \ldots;\quad m > 0$

Parametar m: $m = np = const,\ m > 0$

Rekurentni obrazac: $p(n;x) = \frac{m}{x} \cdot p_{(n;x-1)}$

Poslednja verovatnoća: $p(n;n) = 1 - \sum_{i=1}^{n-1} p(n;x)$

VER-041
Puasonova verovatnoća

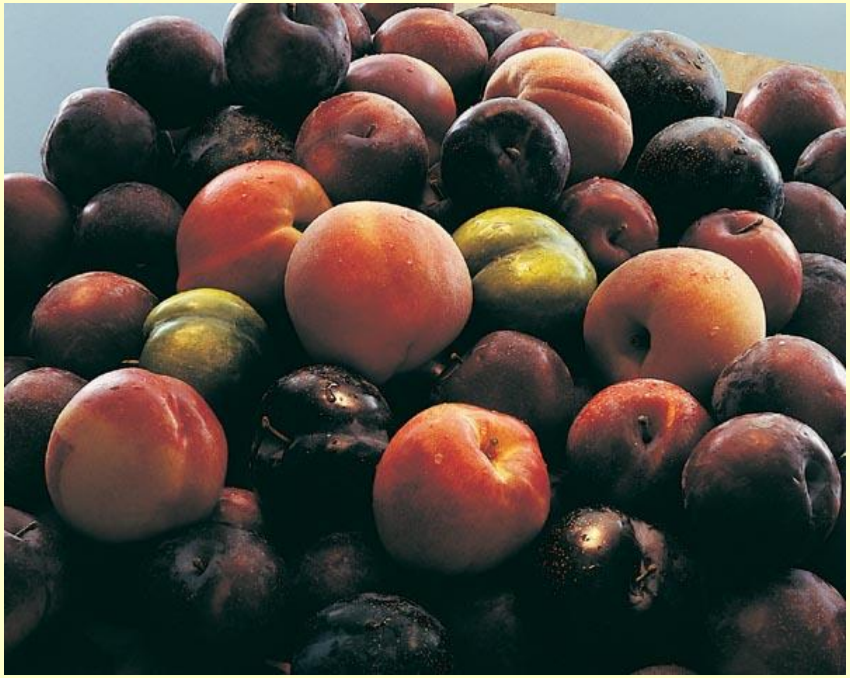
VER-006 K:3-2
Puasonova verovatnoća

VER-029 Z(06)5-4
Puasonova verovatnoća

Zakoni velikih brojeva

Zakon velikih brojeva se zasniva na pretpostavci da u velikom broju slučajnih pojava njihova srednja vrednost prestaje da bude slučajna veličina i da se može predvideti sa velikom pouzdanošću.

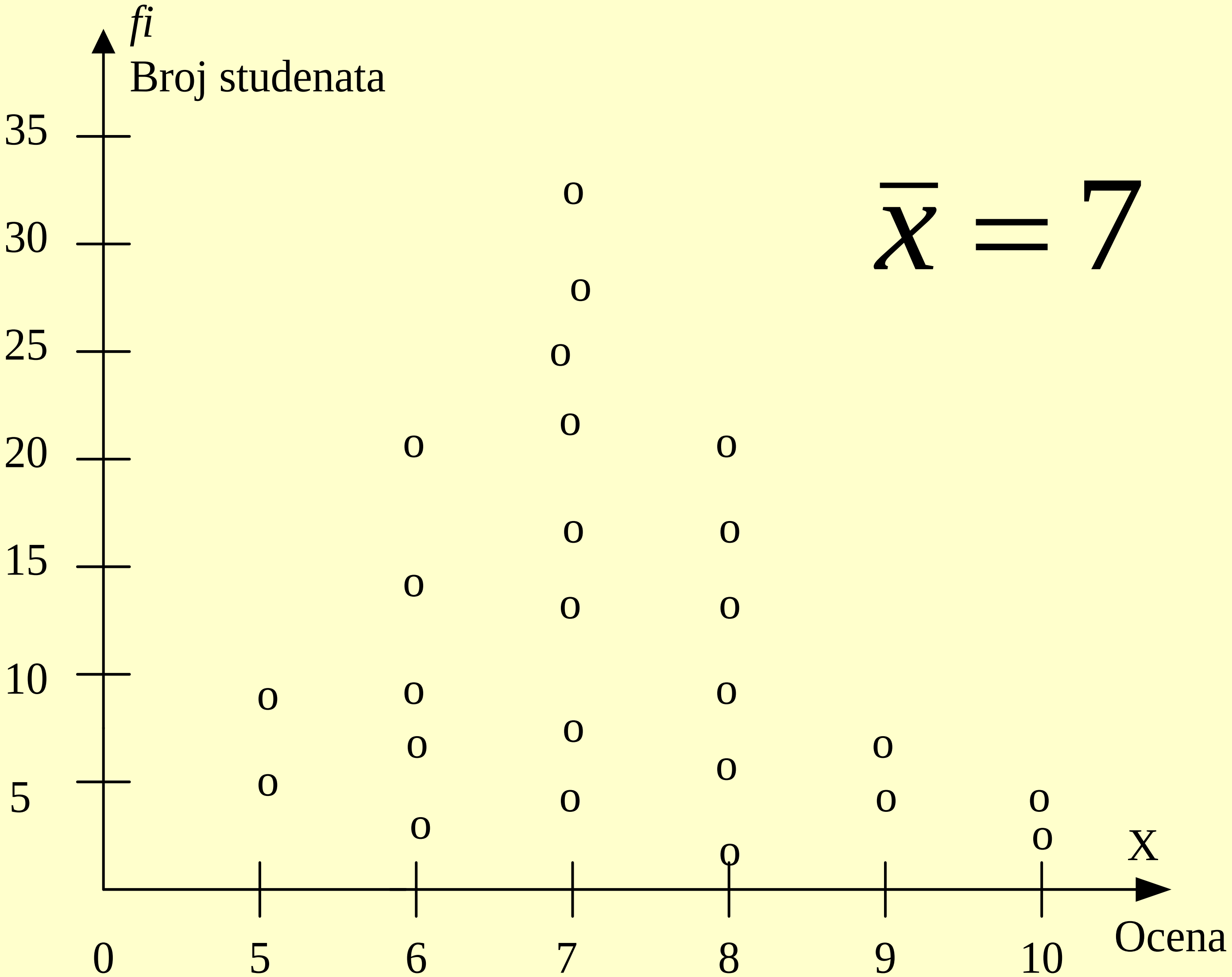
Primer: Bacanje novčića



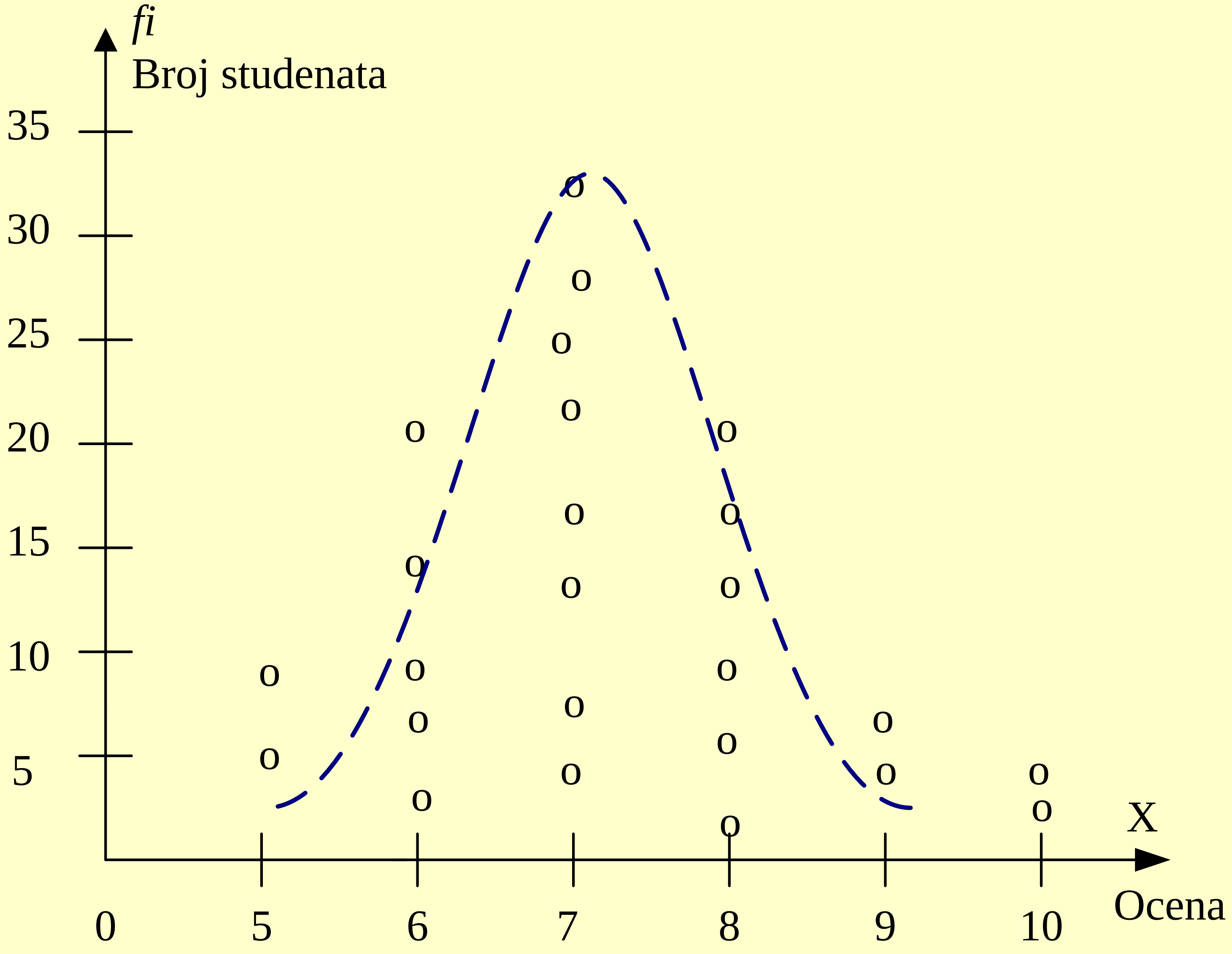
Centralna granična teorema

Centralna granična teorema je skup uslova pri kojima se empirijski rasporedi približavaju normalnom zakonu rasporeda verovatnoća.

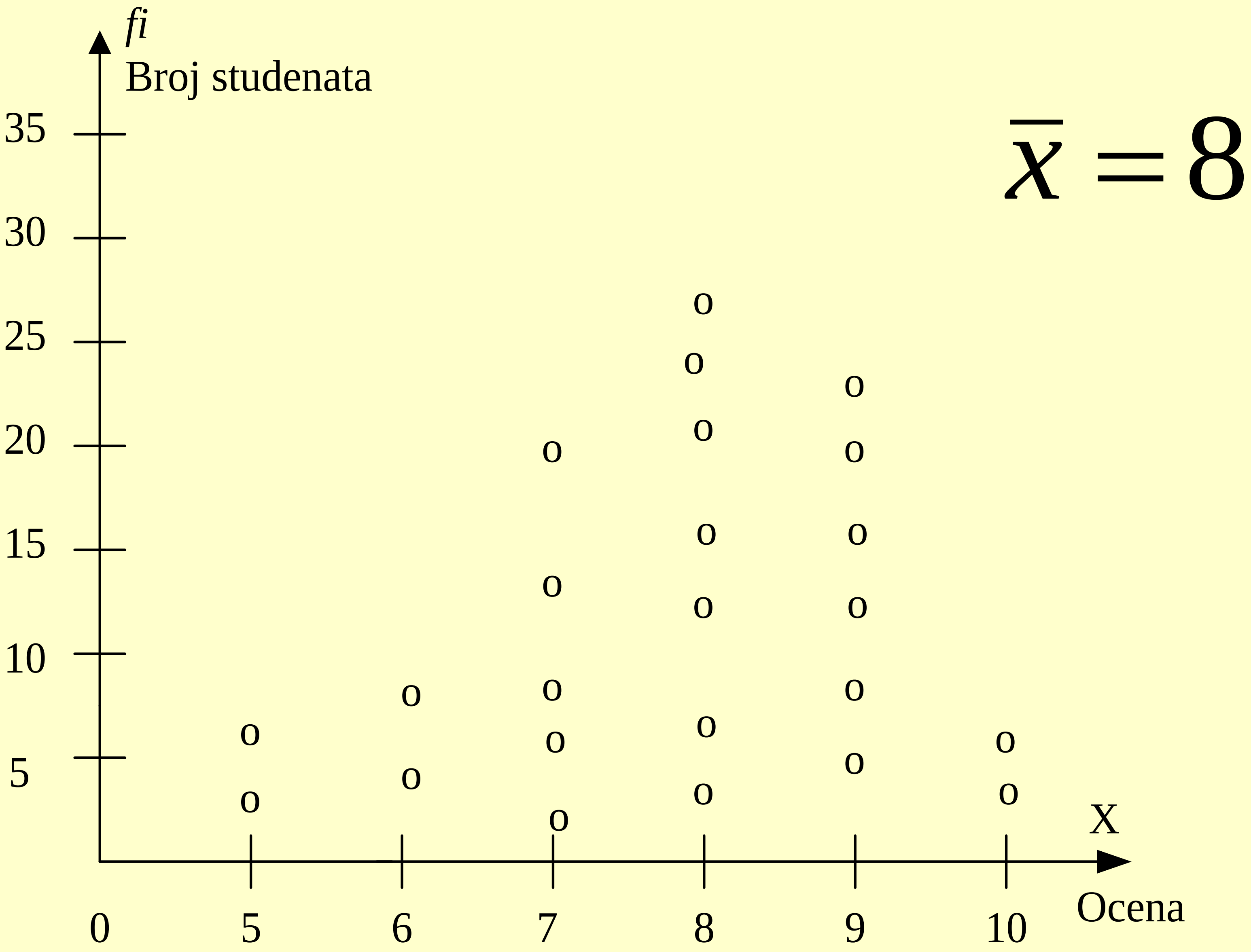
Primer:Ocene studenata po ispitnim rokovima:

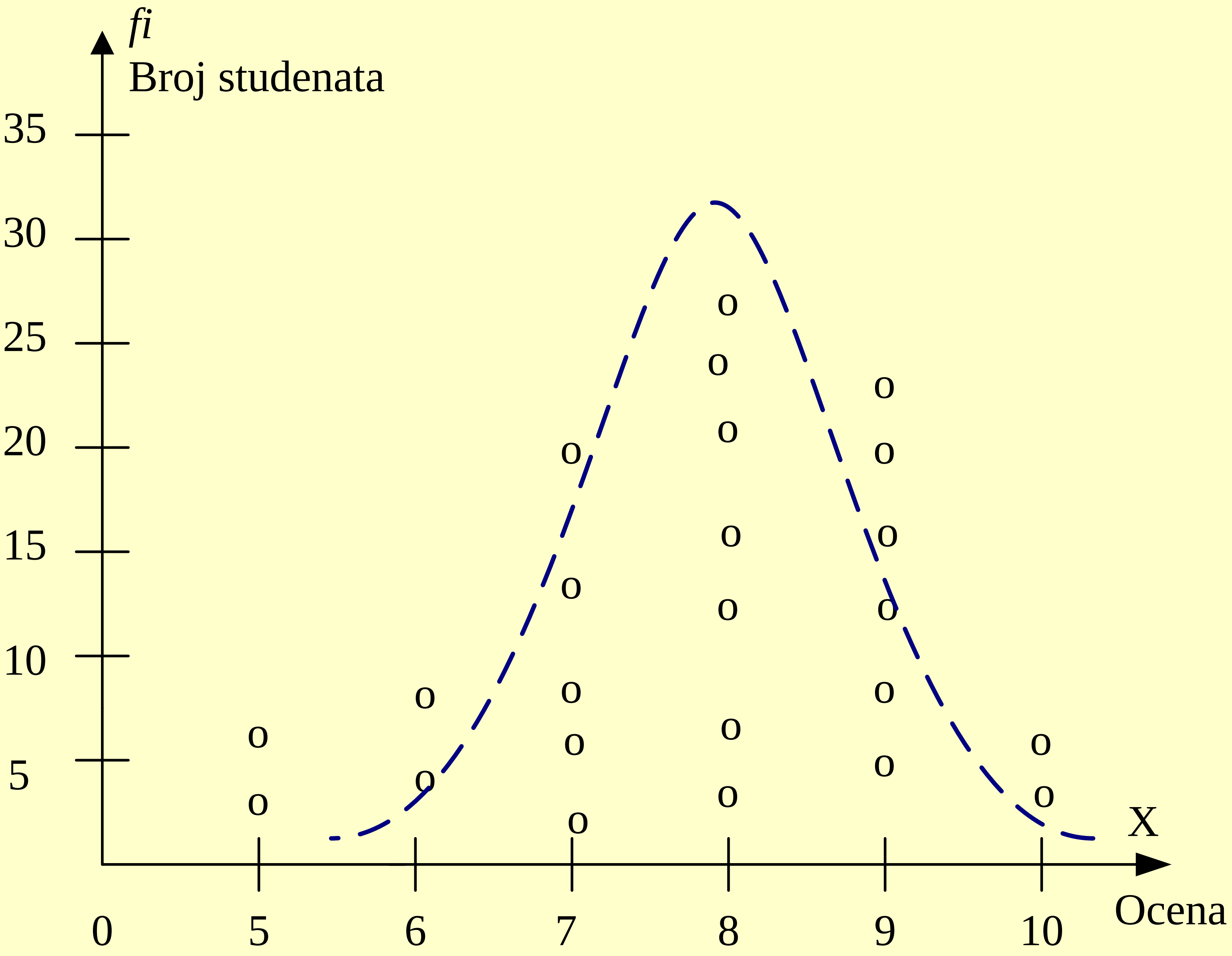


$$\bar{x} = 7$$



fi
Broj studenata
35
30
25
20
15
10
5
0
5
6
7
8
9
10
X
Ocena


fi
Broj studenata
35
30
25
20
15
10
5
0
5
6
7
8
9
10
X
Ocena
$\bar{x} = 8$


fi
Broj studenata
35
30
25
20
15
10
5
X
0
5
6
7
8
9
10
Ocena

Slučajne promenljive

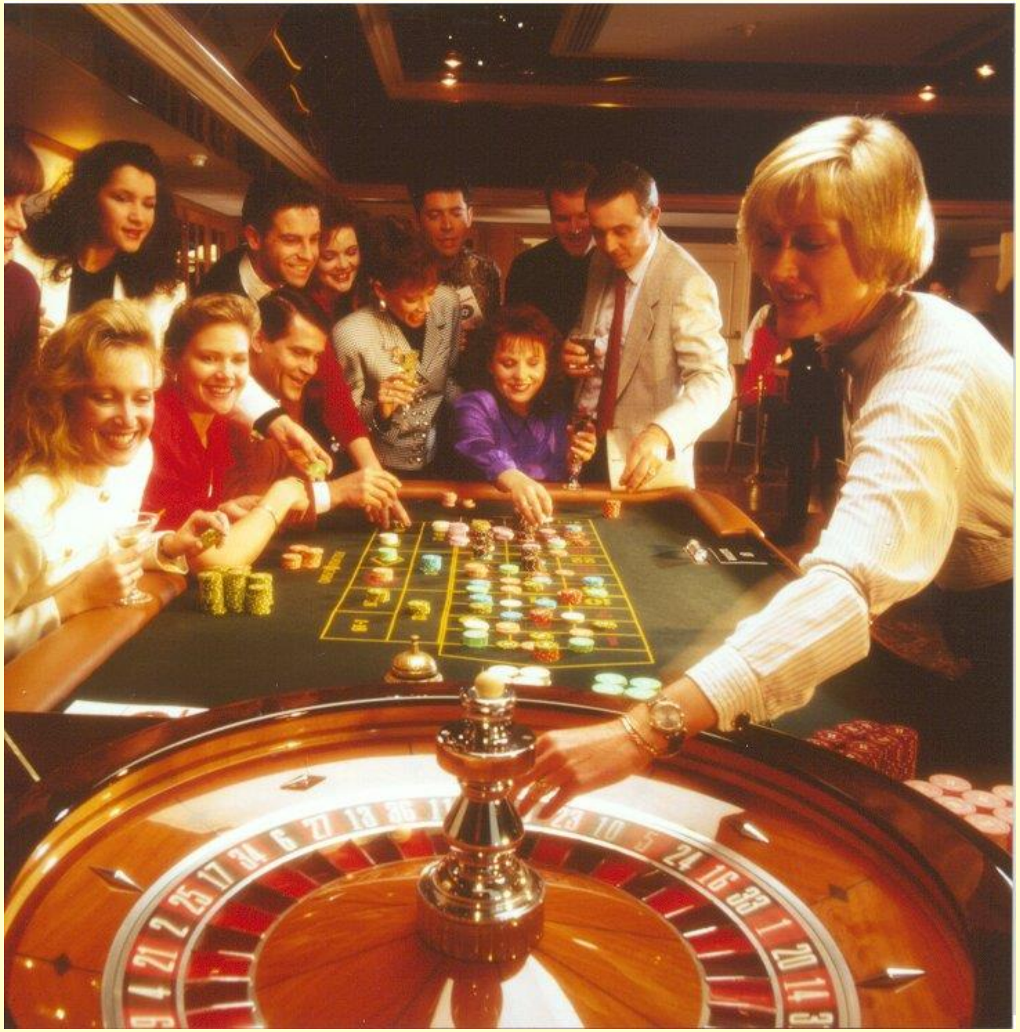
Slučajna promenljiva X je takva promenljiva koja može primati različite brojčane vrednosti sa određenom verovatnoćom.

Ona je u potpunosti definisana ako se pored saznanja o tome koje vrednosti ona može uzimati, zna i sa kojim verovatnoćama ona može te vrednosti uzimati.

Razlikuju se:

•**prekidna** (diskontinualna, diskretna) promenljiva,

•**neprekidna** (kontinualna) promenljiva.



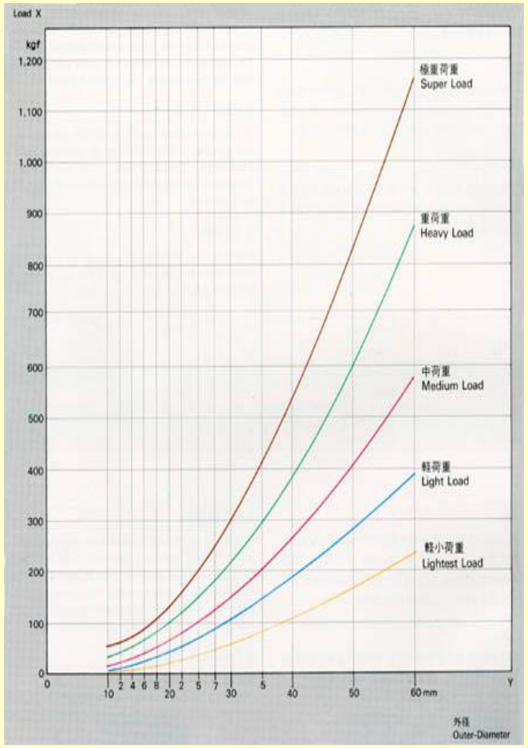
Jednodimenzionalna prekidna slučajna promenljiva

Prekidna slučajna promenljiva X je ona slučajna promenljiva koja za svoju vrednost uzima cele nenegativne brojčane vrednosti, sa pozitivnim verovatnoćama p_i, $i=1,2,\ldots,n$.

$$X=\begin{Bmatrix} x_1 & x_2 & \ldots & x_i & \ldots & x_n \\ p_1 & p_2 & \ldots & p_i & \ldots & p_n \end{Bmatrix}; \qquad \sum_{i=1}^{n} p_i = 1$$

VER-031 Jednodimenzionalna prekidna slučajna promenljiva

Podela svih rasporeda verovatnoća u dve velike grupe: empirijski i teorijski rasporedi.



Teorijski prekidni rasporedi (nema u knjizi):

- Bernulijev raspored verovatnoća.
- Binomni raspored verovatnoća*.
- Puasonov raspored verovatnoća*.
- Hipergeometrijski raspored verovatnoća.
- Negativni binomni raspored verovatnoća.
- Geometrijski raspored verovatnoća.

Binomni raspored verovatnoća

Verovatnoće binomnog rasporeda se izračunavaju na osnovu već pominjanih formula!

Uslovi:

$$0<p<1,$$

$$p+q=1,$$

$$\sum_{x=0}^{n} P(n;x)=1.$$

$$(p>0{,}05);\ n\leq 20.$$

Puasonov raspored verovatnoća

$$(n \to \infty);\ (p(A) \to 0)$$

$$(p \leq 0{,}05);\ n > 20.$$

Jednodimenzionalna neprekidna slučajna promenljiva

Ako slučajna promenljiva X može na slučajan način uzimati bilo koju proizvoljnu brojčanu vrednost, u određenom intervalu, na primer, u intervalu (a,b).

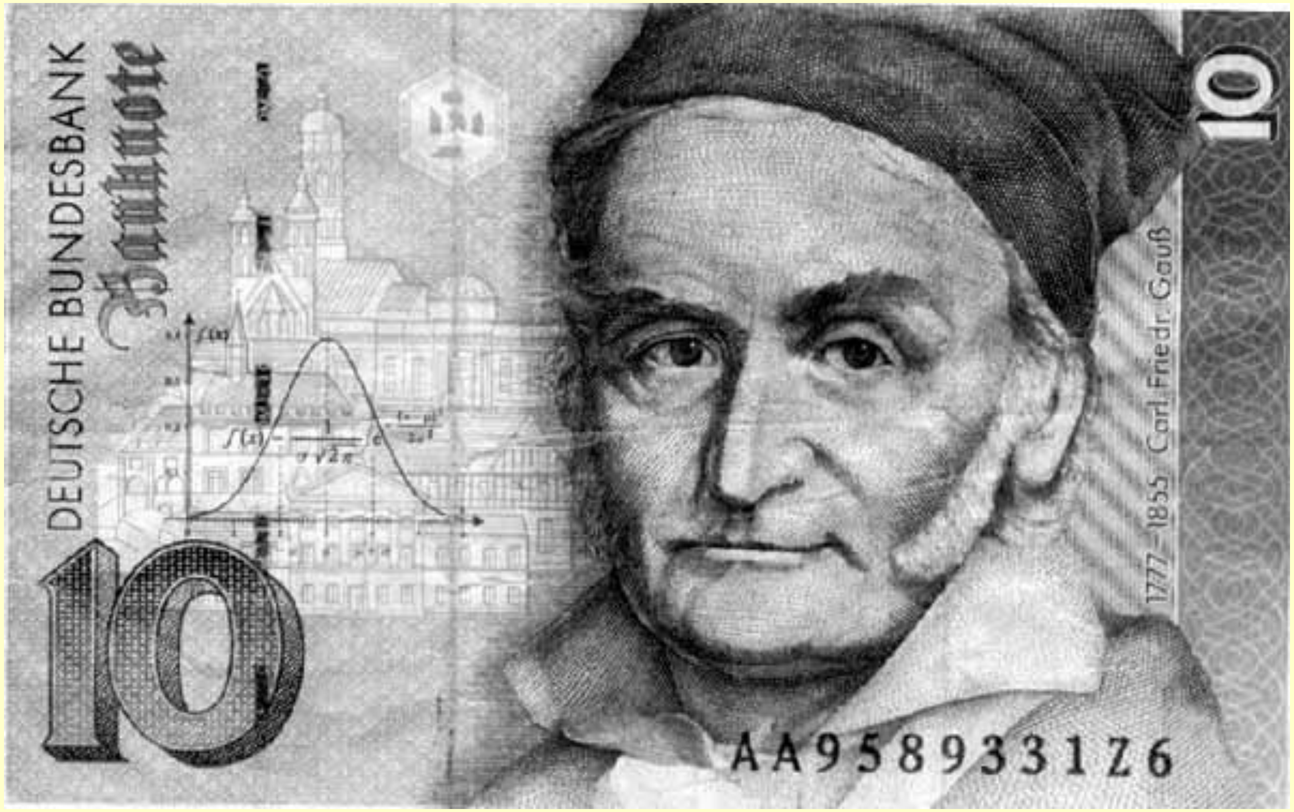
Teorijski neprekidni rasporedi verovatnoća

Vrste:

- Normalni raspored verovatnoća*.
- Studentov raspored verovatnoća*.
- Hi-kvadrat raspored verovatnoća*.
- Snedekorov F-raspored verovatnoća (Fišerov raspored)*.
- Ravnomerni raspored verovatnoća.
- Eksponencijalni raspored verovatnoća.
- Lognormalni raspored verovatnoća.

Normalan raspored verovatnoća

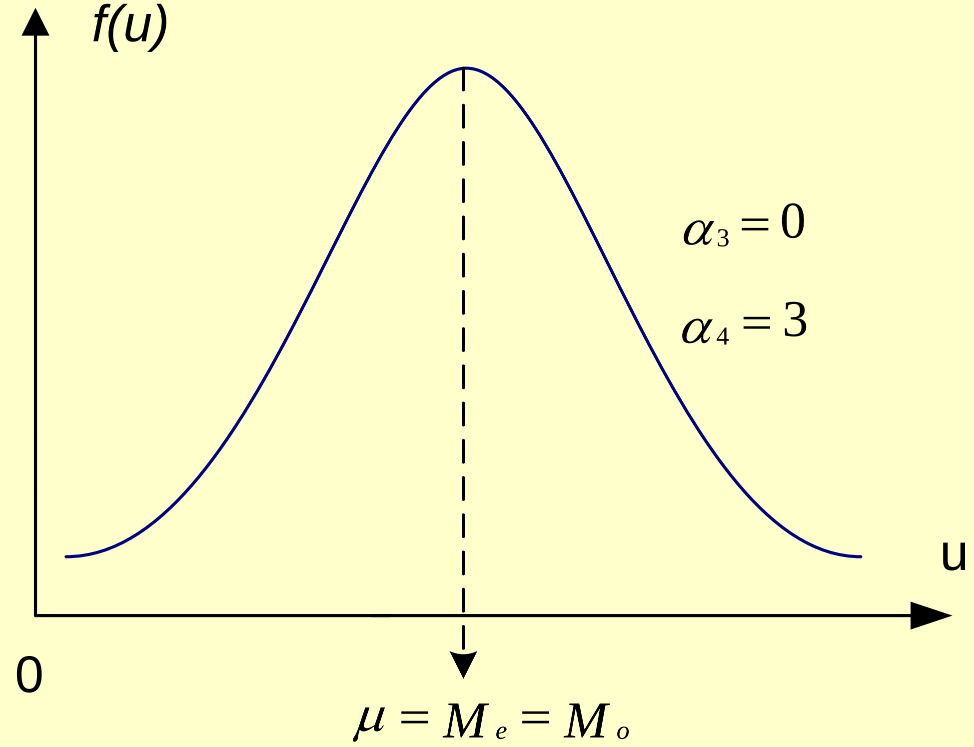
Karl Fridrih Gaus
XIX vek



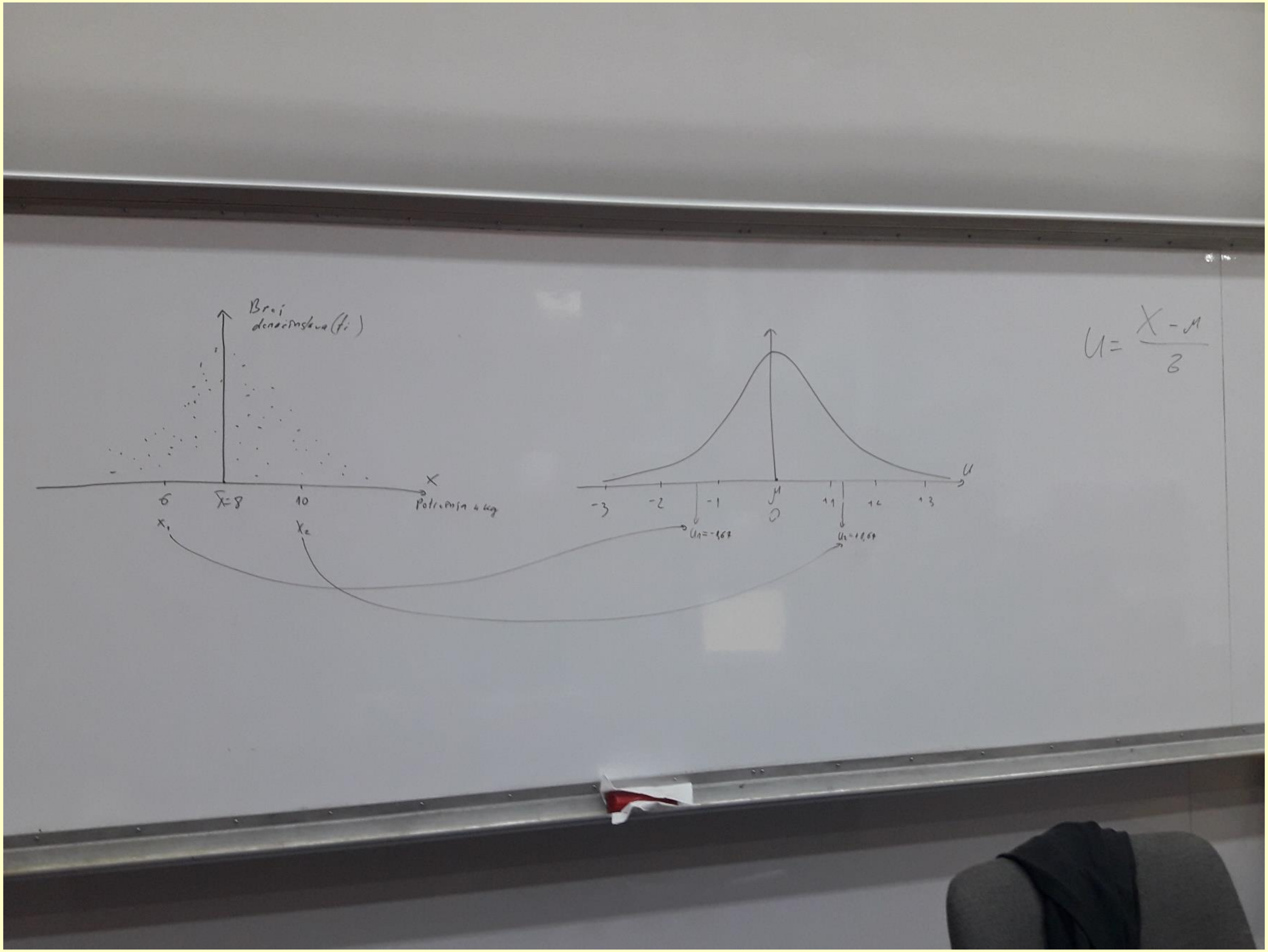
Zašto je normalni raspored toliko važan?

- Veliki broj masovnih pojava ima približno normalan raspored verovatnoća.
- Mnogi prekidni rasporedi verovatnoća mogu se aproksimirati normalnim rasporedom verovatnoća.
- Na osnovu njega razvijen je veći broj neprekidnih rasporeda, kao što su Studentov, Hi-kvadrat, Snedekorov F-raspored itd.
- Predstavlja osnovu za parametarsko zaključivanje.
- Veliki broj statističkih problema se rešava ili se može rešavati samo uz pretpostavku da osnovni skup, iz koga je formiran uzorak, ima normalan raspored.

Osobine normalnog rasporeda verovatnoća



- Oblik simetričnog zvona.
- Koeficijent asimetrije je α_3=0 i koeficijent spljoštenosti je α_4=3.
- Ukupna površina ispod krive normalnog rasporeda jednaka je jedinici.
- Frekvencije rasporeda opadaju istim tempom i potpuno su simetrično raspoređene.
- Normalni raspored je u potpunosti definisan parametrima μ i σ^2.
- Ukoliko su poznate μ i σ, može se izračunati površina ispod normalne krive.



Broj
X
6
10
X_1
X_2
-3
-2
-1
μ
0
u
$u = \frac{X - \mu}{\sigma}$

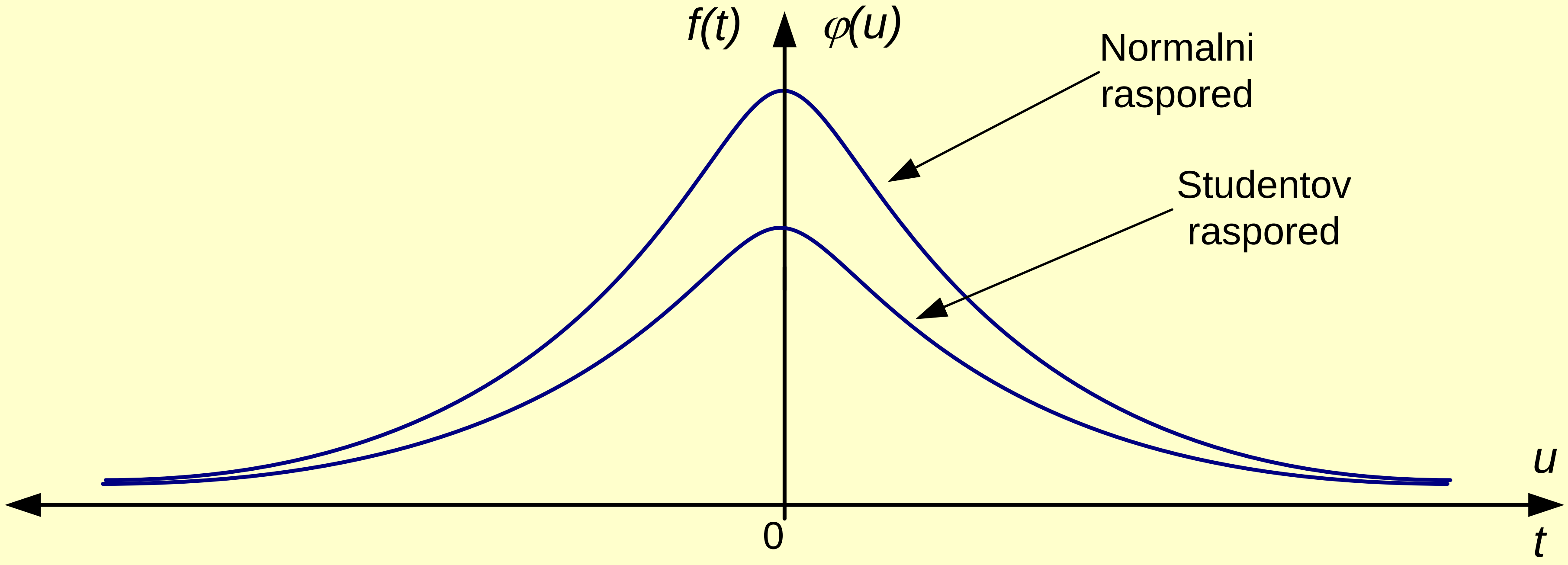
VER-042 Normalizovani raspored

VER-017 Z(06)6-1 Normalizovani raspored

VER-022 Normalizovani raspored

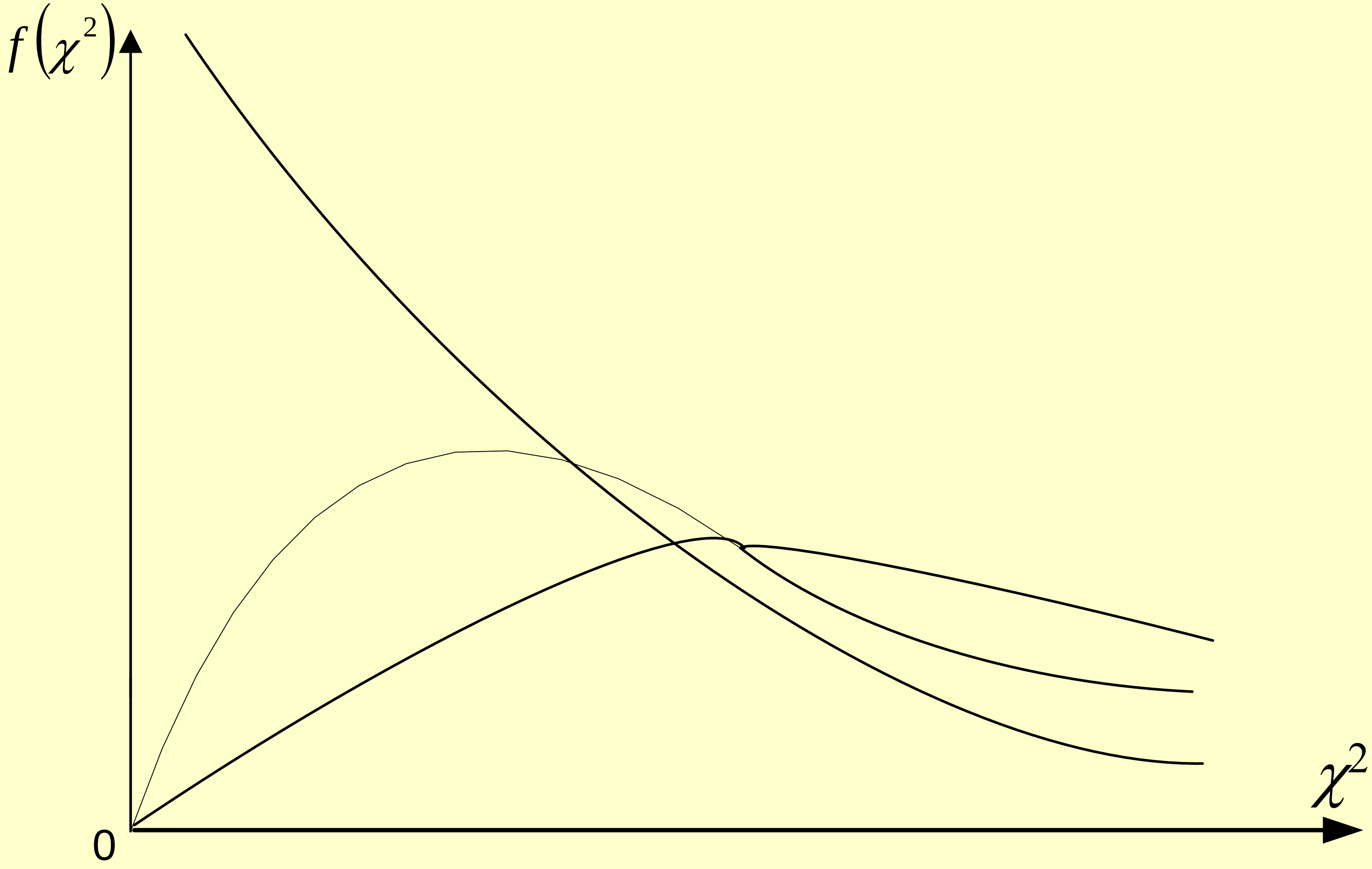
Studentov t-raspored verovatnoća

W. S. Gosset (1908)



Ovaj raspored se koristi kod malih uzoraka za ocenjivanje parametara osnovnog skupa i za testiranje hipoteza o parametrima osnovnog skupa. Mali uzorak je onaj koji ima manje od 30 jedinica (n<30).

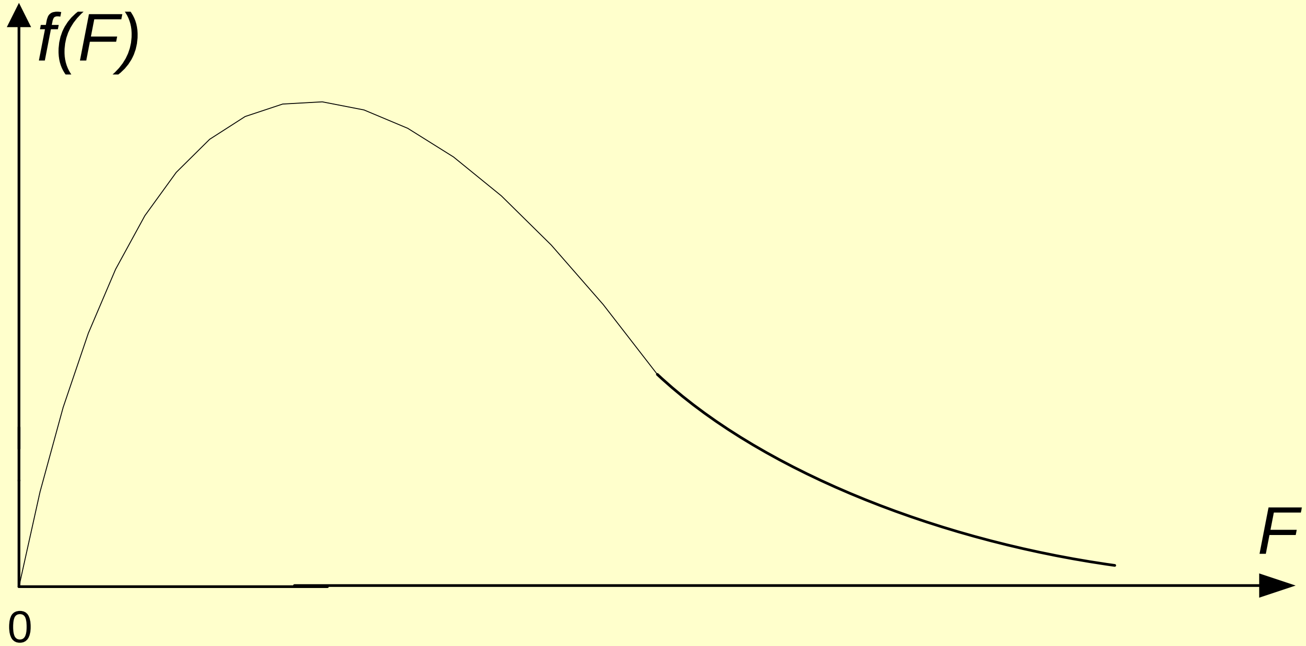
χ^2 (hi-kvadrat) raspored verovatnoća



Hi-kvadrat raspored se najviše koristi u oblasti neparametarskih testova.

Snedekorov F-raspored verovatnoća

R. A. Fisher



Ovaj raspored se najviše koristi u statističkom metodu pod nazivom "analiza varijanse".

Izbor odgovarajućeg teorijskog rasporeda

$$\sigma^2_{rasp} = \frac{\sum_{i=1}^{n}\left(f_i - f_i^{rasp}\right)^2}{n+1}$$

Najbolji teorijski raspored je onaj koji ima najmanju varijansu!

Aproksimacija normalnim rasporedom

Prvi uslov za aproksimaciju: $\alpha_3 \approx 0.$

Drugi uslov za aproksimaciju: $\alpha_4 \approx 3.$

VER-008 K:3-3
Aproksimacija normalnim rasporedom

VER-009 Z(06)7-1
Aproksimacija normalnim rasporedom

VER-003; K(05)z 3-1 Binomna verovatnoća

VER-005; K(05)z 3-2 Puasonova verovatnoća

VER-007; K(05)z 3-3 Aproksimacija norm. rasporedu

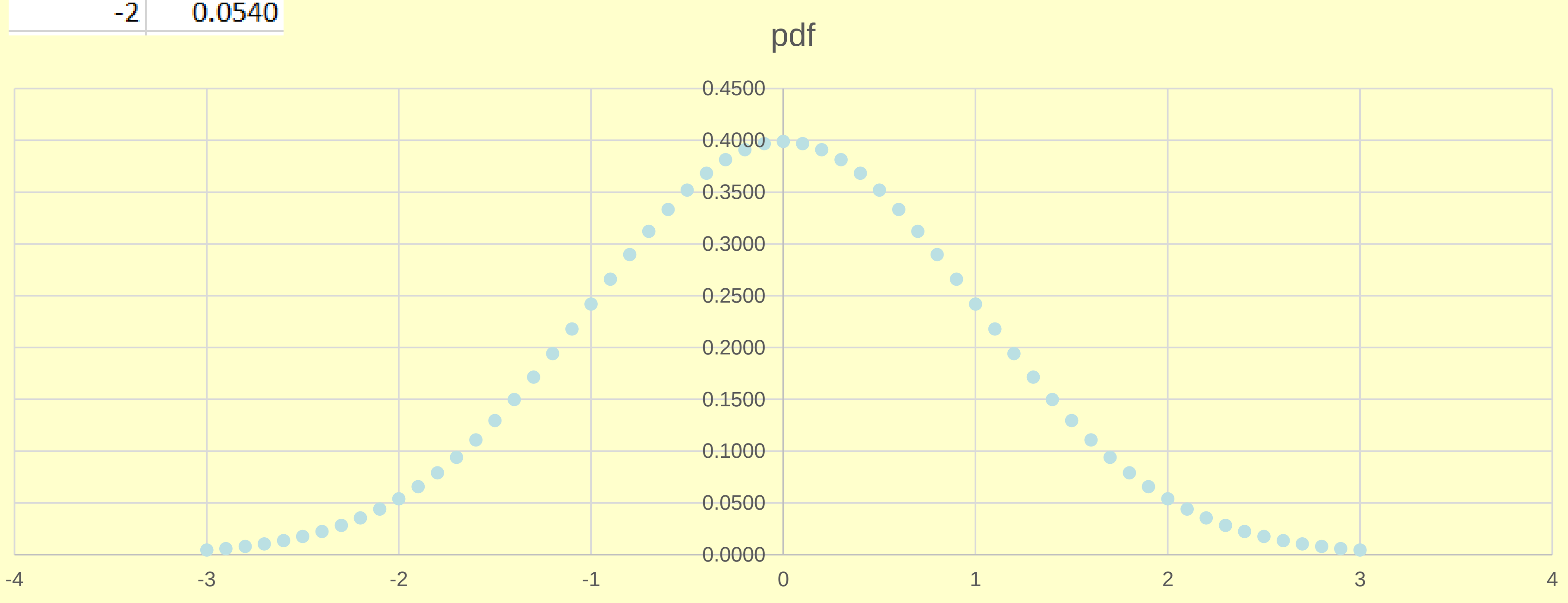
=NORM.S.DIST(B2,FALSE)

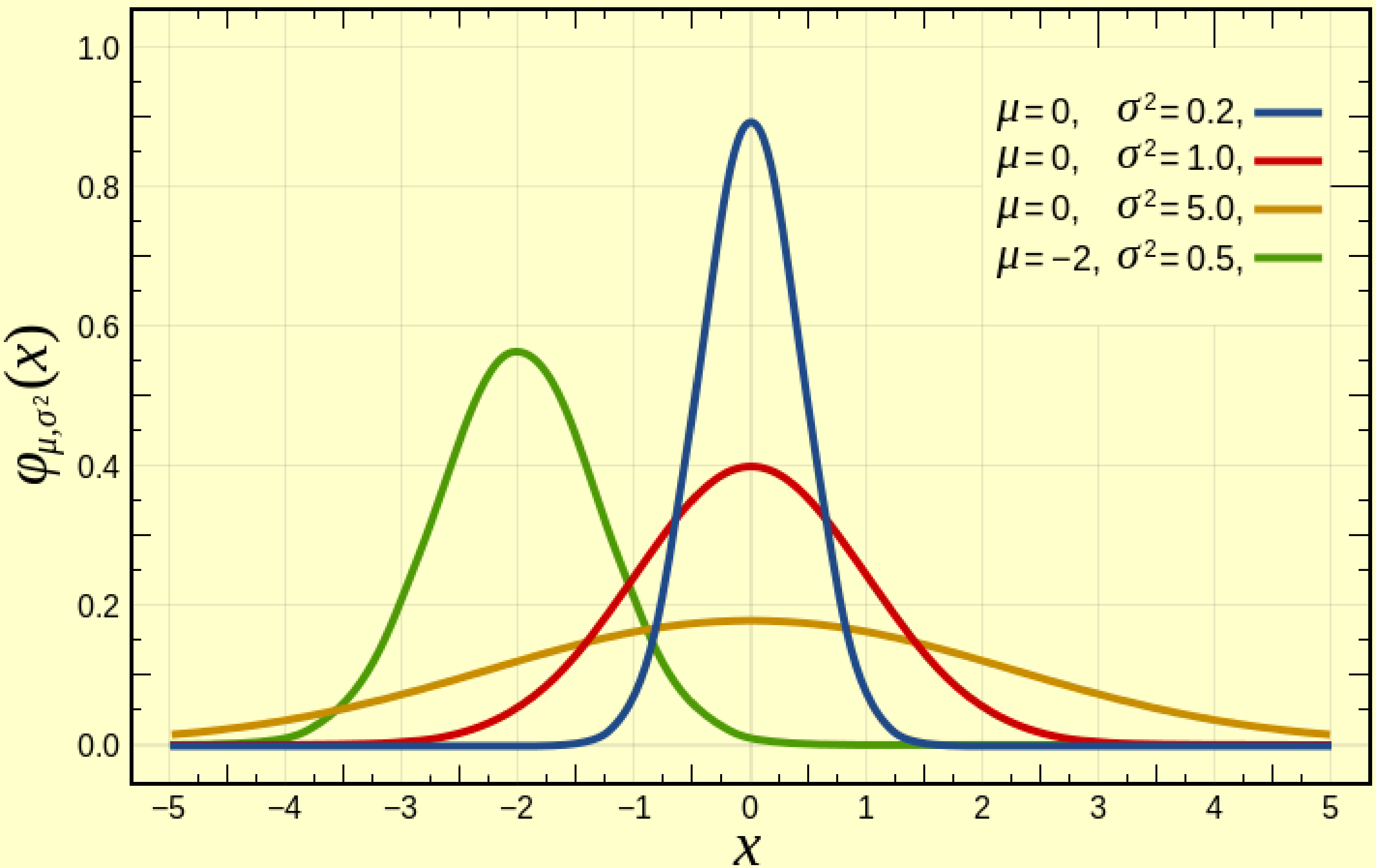
B	C	D	E	F
z	pdf	Kumulativna funkcija		
-3	0.0044	0.0013		
-2.9	0.0060	0.0019		
-2.8	0.0079	0.0026		
-2.7	0.0104	0.0035		
-2.6	0.0136	0.0047		
-2.5	0.0175	0.0062		
-2.4	0.0224	0.0082		
-2.3	0.0283	0.0107		
-2.2	0.0355	0.0139		
-2.1	0.0440	0.0179		
-2	0.0540	0.0228		
-1.9	0.0656	0.0287		
-1.8	0.0790	0.0359		
-1.7	0.0940	0.0446		
-1.6	0.1109	0.0548		
-1.5	0.1295	0.0668		

z	pdf
-3	0.0044
-2.9	0.0060
-2.8	0.0079
-2.7	0.0104
-2.6	0.0136
-2.5	0.0175
-2.4	0.0224
-2.3	0.0283
-2.2	0.0355
-2.1	0.0440
-2	0.0540

=NORM.S.DIST(B2,FALSE)

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

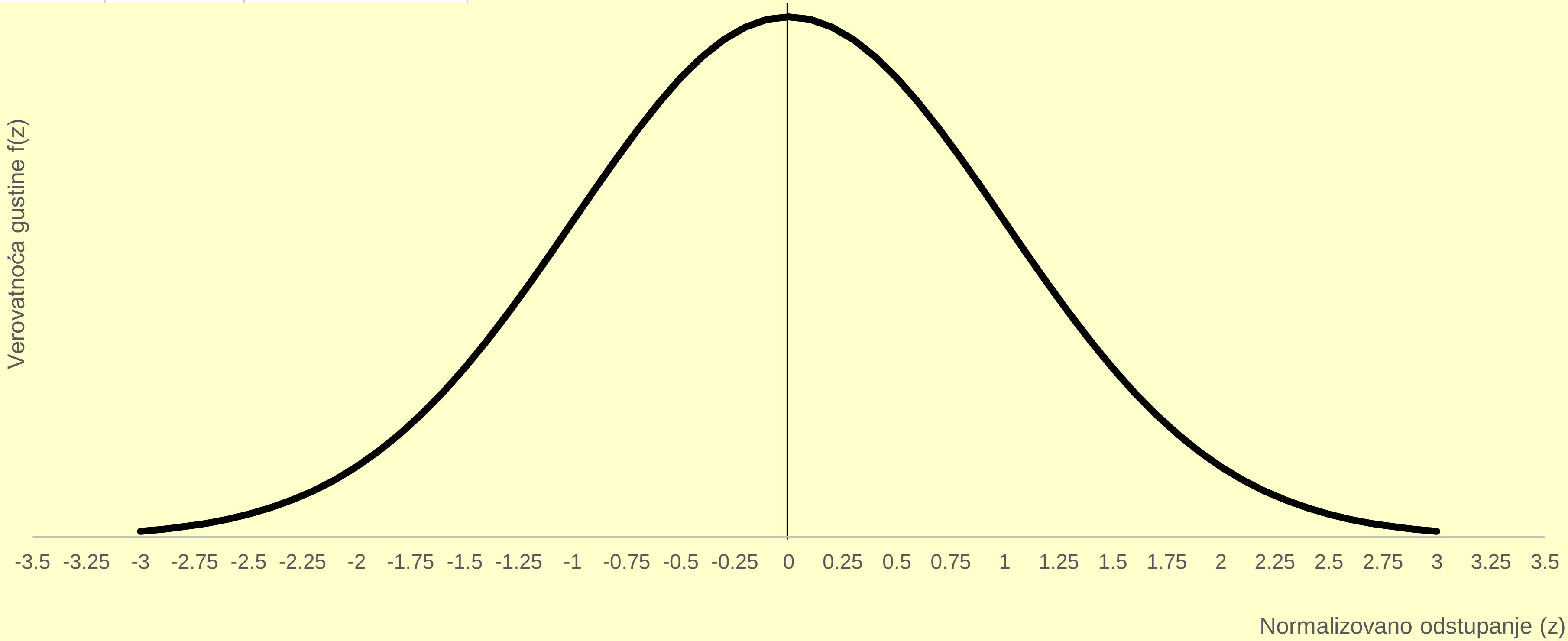


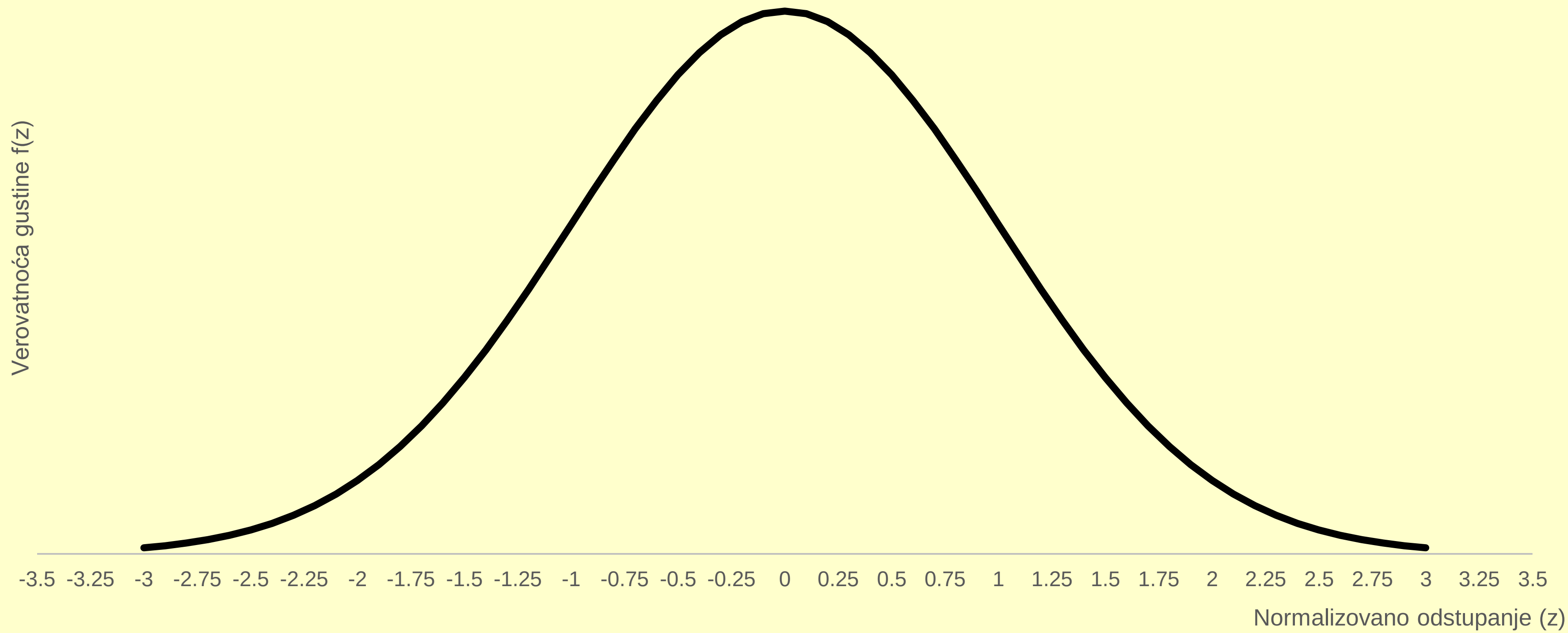

$\mu=0, \quad \sigma^2=0.2,$
$\mu=0, \quad \sigma^2=1.0,$
$\mu=0, \quad \sigma^2=5.0,$
$\mu=-2, \quad \sigma^2=0.5,$
$\varphi_{\mu,\sigma^2}(x)$
1.0
0.8
0.6
0.4
0.2
0.0
−5 −4 −3 −2 −1 0 1 2 3 4 5
x

z	pdf	Kumulativna funkcija
-3	0.0044	0.0013
-2.9	0.0060	0.0019
-2.8	0.0079	0.0026
-2.7	0.0104	0.0035
-2.6	0.0136	0.0047
-2.5	0.0175	0.0062
-2.4	0.0224	0.0082
-2.3	0.0283	0.0107
-2.2	0.0355	0.0139
-2.1	0.0440	0.0179
-2	0.0540	0.0228

=NORM.S.DIST(B3,TRUE)

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$




Verovatnoća gustine f(z)
-3.5 -3.25 -3 -2.75 -2.5 -2.25 -2 -1.75 -1.5 -1.25 -1 -0.75 -0.5 -0.25 0 0.25 0.5 0.75 1 1.25 1.5 1.75 2 2.25 2.5 2.75 3 3.25 3.5
Normalizovano odstupanje (z)