



# Verjetnost in verjetnostne porazdelitve

---

Iztok Grabnar

Univerza v Ljubljani, Fakulteta za farmacijo

*Oktober, 2011*

# What is the Normal Distribution?

Statistician and Chemist Jack Youden's Graphic:

**THE  
NORMAL  
LAW OF ERROR  
STANDS OUT IN THE  
EXPERIENCE OF MANKIND  
AS ONE OF THE BROADEST  
GENERALIZATIONS OF NATURAL  
PHILOSOPHY ♦ IT SERVES AS THE  
GUIDING INSTRUMENT IN RESEARCHES  
IN THE PHYSICAL AND SOCIAL SCIENCES AND  
IN MEDICINE AGRICULTURE AND ENGINEERING ♦  
IT IS AN INDISPENSABLE TOOL FOR THE ANALYSIS AND THE  
INTERPRETATION OF THE BASIC DATA OBTAINED BY OBSERVATION AND EXPERIMENT**

# Verjetnost

---

“Nothing in life is certain. In everything we do, we gauge the chances of successful outcomes, from business to medicine to the weather. But for most of human history, **probability**, the formal study of the laws of chance, was used for only one thing: **gambling.**”



# Teorija verjetnosti

---

- Obravnava situacije (poskuse) pri katerih je izid odvisen od naključja (naključni eksperiment).
  - Prostor izidov je množica vseh izidov.
  - Podmnožice prostora izidov so dogodki.
-

# Met dveh igralnih kock

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

# Definicije

---

- Izid
  - Dogodek
  - Vsota
  - Produkt
  - Komplementaren dogodek
  - Nemogoč dogodek
  - Gotov dogodek
  - Nezdružljiva dogodka
  - Neodvisna dogodka
-

# Vejetnost

---

- Relativna frekvenca dogodka pri  $n$  ponovitvah

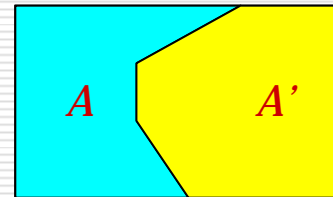
$$\frac{\textit{frekvenca}(A)}{N}$$

- $P(A)$  je limita relativne frekvence dogodka  $A$  pri velikem številu ponovitev poskusa.
-

- 
- Na dogodkih lahko izvajamo iste operacije kot na množicah.
  - Verjetnost je funkcija, ki dogodku  $A$  priredi število  $P(A) \in [0, 1]$ . Pri tem velja:
    - $P(G) = 1$
    - $AB = \emptyset \Rightarrow P(A+B) = P(A) + P(B)$
-



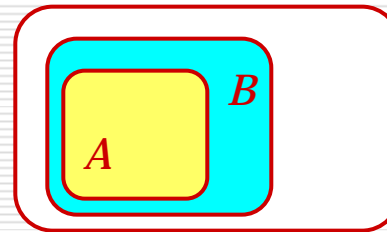
- $P(A') = 1 - P(A)$



$$P(A) + P(A') = P(G) = 1$$

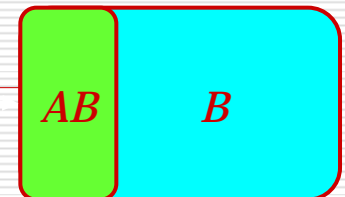
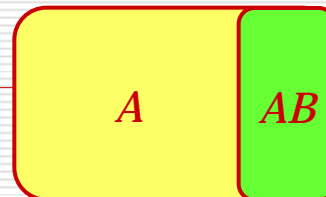
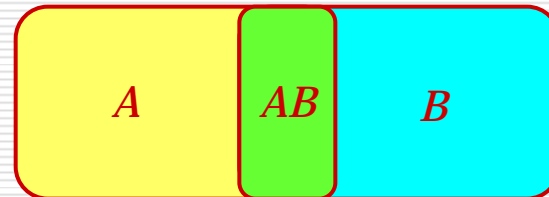
- $P(N) = 0$

- $A \subseteq B \Rightarrow P(A) \leq P(B)$



$$P(B) = P(A + (B - A)) = P(A) + P(B - A) \geq P(A)$$

- $P(A + B) = P(A) + P(B) - P(AB)$



# Pogojna verjetnost

---

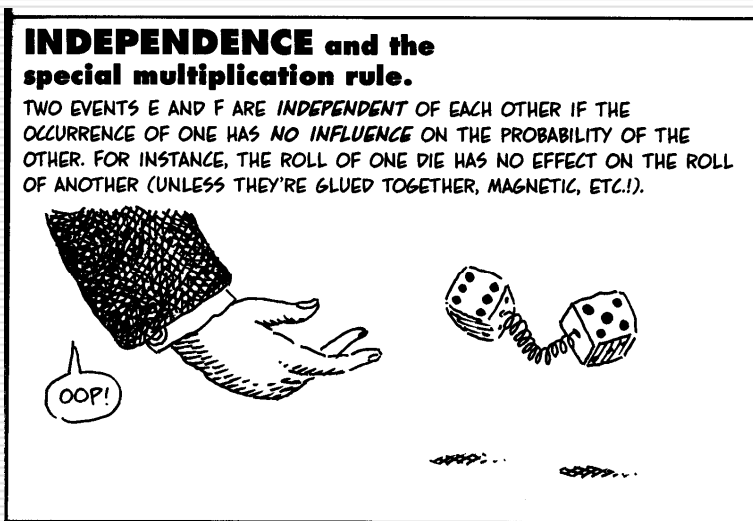
- $A, B$  dogodka (  $P(B) \neq 0$  )
- **Pogojna verjetnost** dogodka  $A$  pri pogoju  $B$  je delež dogodka  $A$  med poskusi, pri katerih se zgodi dogodek  $B$ .

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

# Neodvisna dogodka

---

- Če je  $P(A|B)=P(A)$  sta dogodka A in B neodvisna.
- $P(AB)=P(A) P(B)$



# Slučajna spremenljivka

---

- Funkcija katere vrednost je odvisna od slučaja.
    - Zaloga vrednosti (vrednosti ki jih spremenljivka lahko zajame)
    - Porazdelitev (verjetnosti s katerimi zavzame posamezne vrednosti)
  - Slučajne spremenljivke
    - Diskretne
    - Zvezne
  - Funkcija  $p_V(n) = P(V=n)$  je **verjetnostna gostota** slučajne spremenljivke  $V$ .
-

# Diskretne slučajne spremenljivke

---

- Števno mnogo izidov ( $x_1, x_2, x_3, \dots, x_n$ )
- Porazdelitev je povsem določena z gostoto  $p(x_i) = P(X = x_i)$ .

$$0 \leq p(x_i) \leq 1 \quad \sum_i p(x_i) = 1$$

# Zvezne slučajne spremenljivke

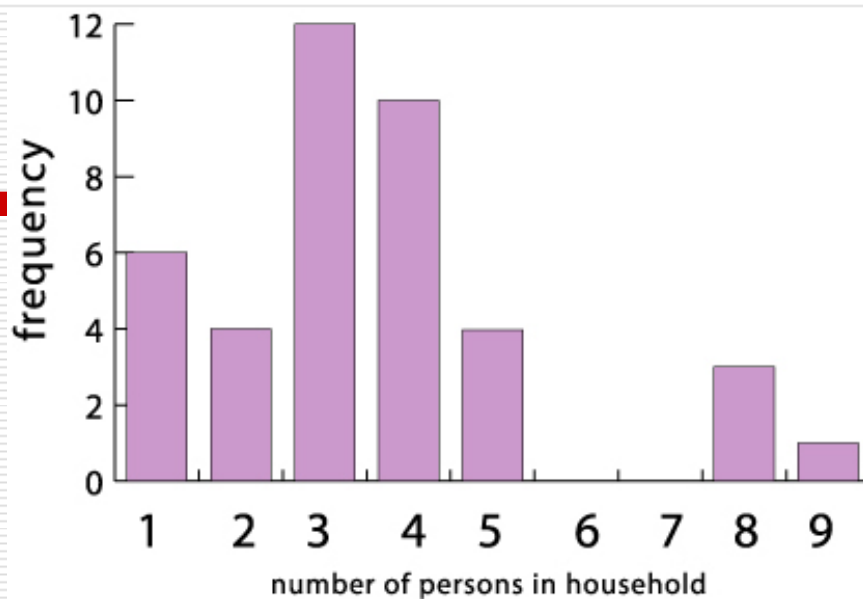
---

- Zaloga vrednosti slučajne spremenljivke je neštevna.
- Posamezni vrednosti zato ne moremo pripisati pozitivne verjetnosti.
- Verjetnostna gostota slučajne spremenljivke.

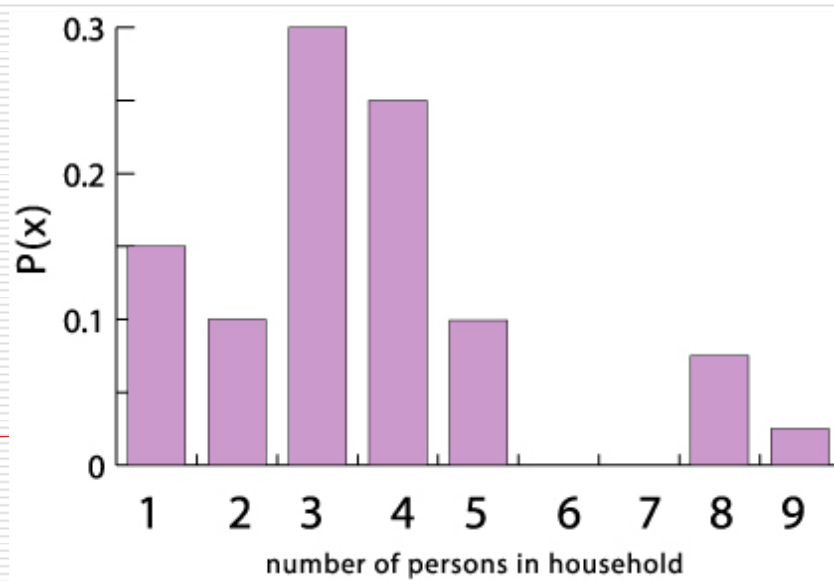
$$0 \leq p(x_i) \leq 1 \quad \text{in} \quad \int_{-\infty}^{+\infty} p(x) dx = 1$$

- Obravnavamo podobno kot diskretne slučajne spremenljivke, le da vsoto nadomestimo z integralom.
  - $P(a \leq x \leq b)$  verjetnost, da  $x$  zavzame vrednost med  $a$  in  $b$ .
-

# frekvenčni histogram

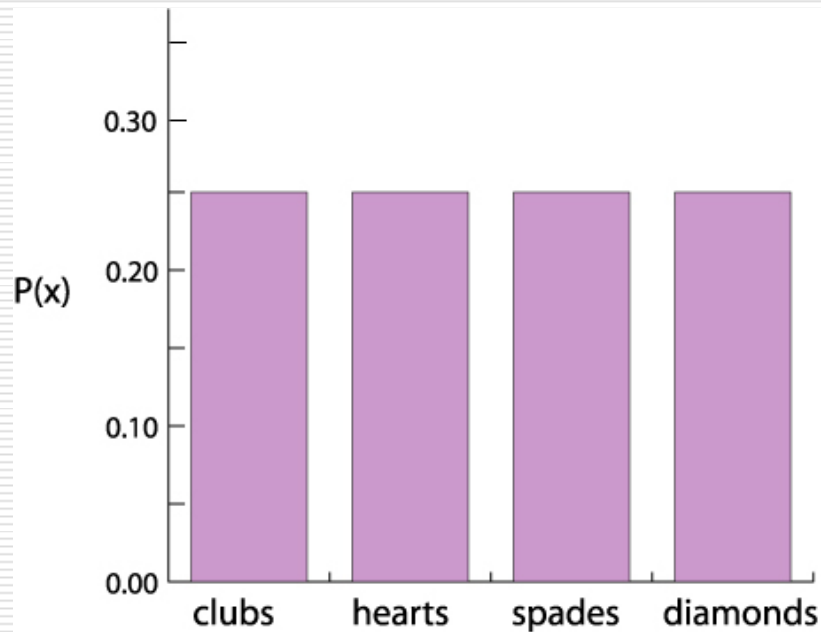


# Gostota verjetnosti



# Enakomerna porazdelitev

---



Verjetnosti k možnih dogodkov so enake  
 $P(x) = 1/k$

---



# Binomska porazdelitev

Bernoulli

- Enostaven naključni eksperiment z dvema mogočima izidoma
  - Uspeh ( $U$ )
  - Neuspeh ( $N$ )
  
- Primeri
  - Met kovanca
  - Spol novorojenčka
  - Izid zdravljenja



Jacob Bernoulli (1654-1705)

# Binomska porazdelitev

---

- Če je verjetnost uspeha  $p$
  - Kakšna je verjetnost neuspeha (komplementarni dogodek)?
    - $q = 1 - p$
  - Primeri
    - Met kovanca ( $U = \text{cifra}$ ):  $p = 0.5 \Rightarrow q = 0.5$
    - Met kocke ( $U = 1$ ):  $p = 0.1667 \Rightarrow q = 0.8333$
    - Verjetnost ozdravitve okužbe ( $U = \text{ozdravitev}$ ):  $p = 0.8 \Rightarrow q = 0.2$
-

# Binomska porazdelitev

---

- n ponovitev poskusa
  
  - Primeri
    - 5 metov kovanca
    - 25 metov igralne kocke
    - Zdravljenje okužbe pri 50 bolnikih
  
  - **Predpostavke:** 1)  $p$  se od poskusa do poskusa ne spreminja in 2) poskusi so med seboj neodvisni
-

# Binomska porazdelitev

---

- Kakšna je verjetnost  $X$  uspehov v seriji  $n$  poskusov?
  
- Primer
  - Kakšna je verjetnost, da pri 5 metih kovanca dvakrat vržemo cifro (C) in trikrat grb (G)?

$$P(\text{CCGGG}) = (1/2)^5 = 1/32$$

---

# Binomska porazdelitev

---

- Kombinacije dogodkov:

CCGGG CGCGG CGGCG CGGGC

GCCGG GCGCG GCGGC

GGCCG GGCGC

GGGCC

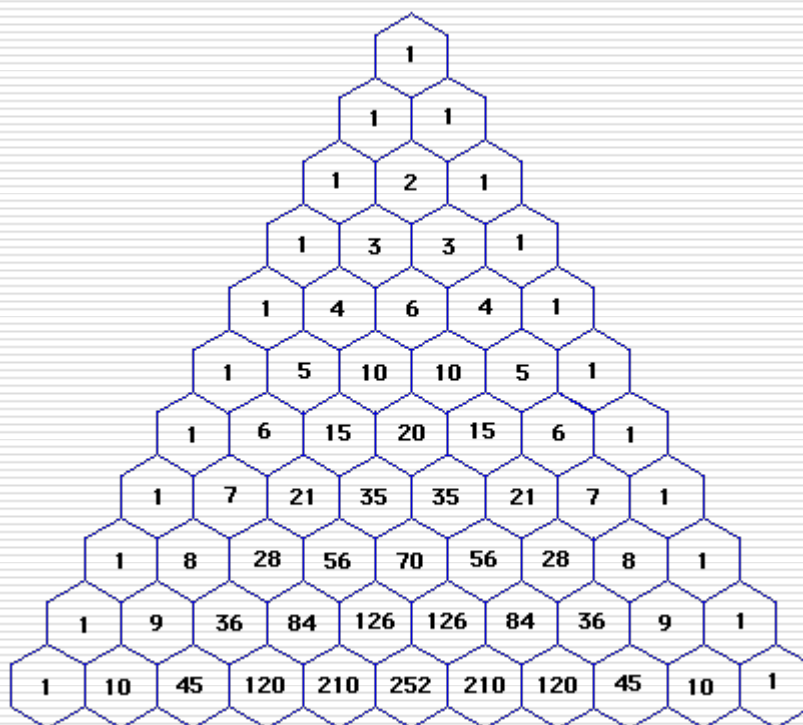
$$P(2 \text{ C}) = 10 \times 1/32 = 10/32$$

$$\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 3 \cdot 2 \cdot 1} = 10$$

---

# Pascalov trikotnik

---



# Binomska porazdelitev

---

- Uspeh / Neuspeh

*NNUNNNNUNUNUUNNNNNUN...*

Verjetnost, da bo  $X$  poskusov uspešnih v tem vrstnem redu je

$$\begin{aligned} P(X) &= q \cdot q \cdot p \cdot q \cdot \dots \\ &= p^X \cdot q^{n-X} \end{aligned}$$

---

# Binomska porazdelitev

---

- Če zaporedje dogodkov ni pomembno

$$P(X) = \frac{n!}{X!(n-X)!} \cdot p^X \cdot q^{n-X}$$

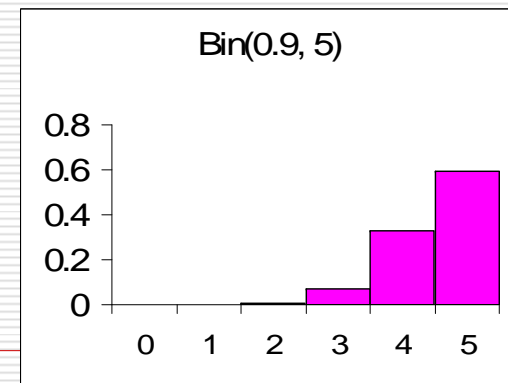
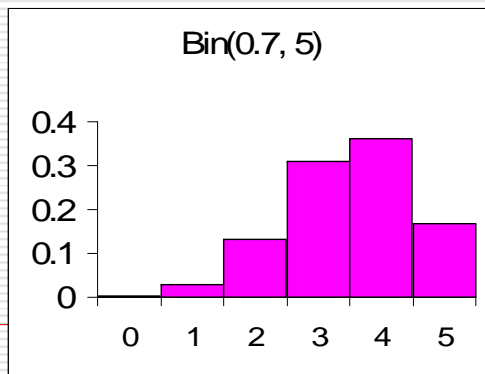
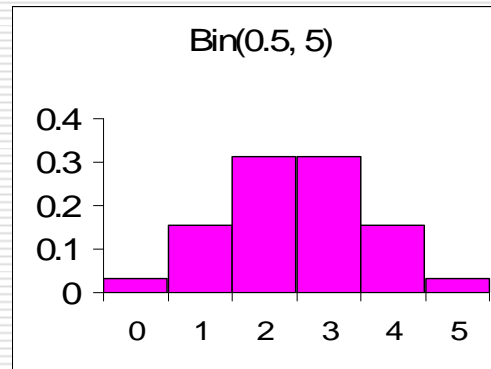
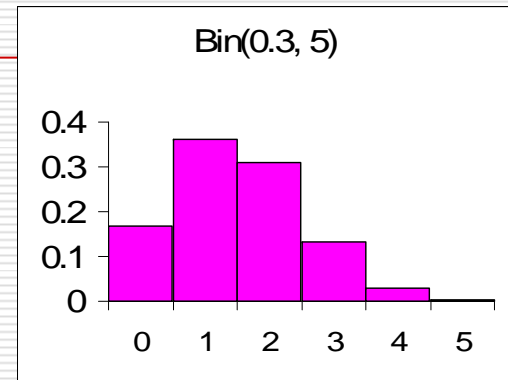
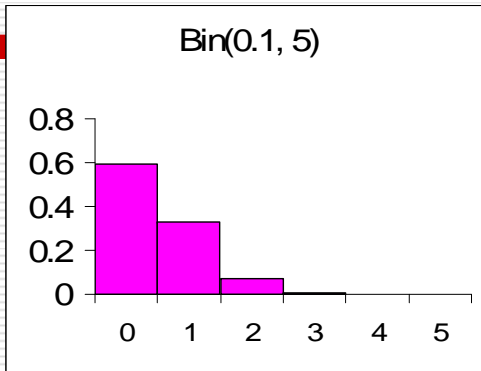
kjer  $\frac{n!}{X!(n-X)!}$  število možnih kombinacij dogodkov,  
od katerih je  $X$  uspešnih

pri  $n$  poskusih, kjer je  $n! = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 2 \cdot 1$

---



# Binomska porazdelitev



# Povprečna lega in razpršenost

---

□  $E(x) = n p$

□  $D(x) = n p (1-p)$

□  $E(x) = p$

□  $D(x) = p (1-p)/n$

$$\sigma = \sqrt{np(1-p)}$$

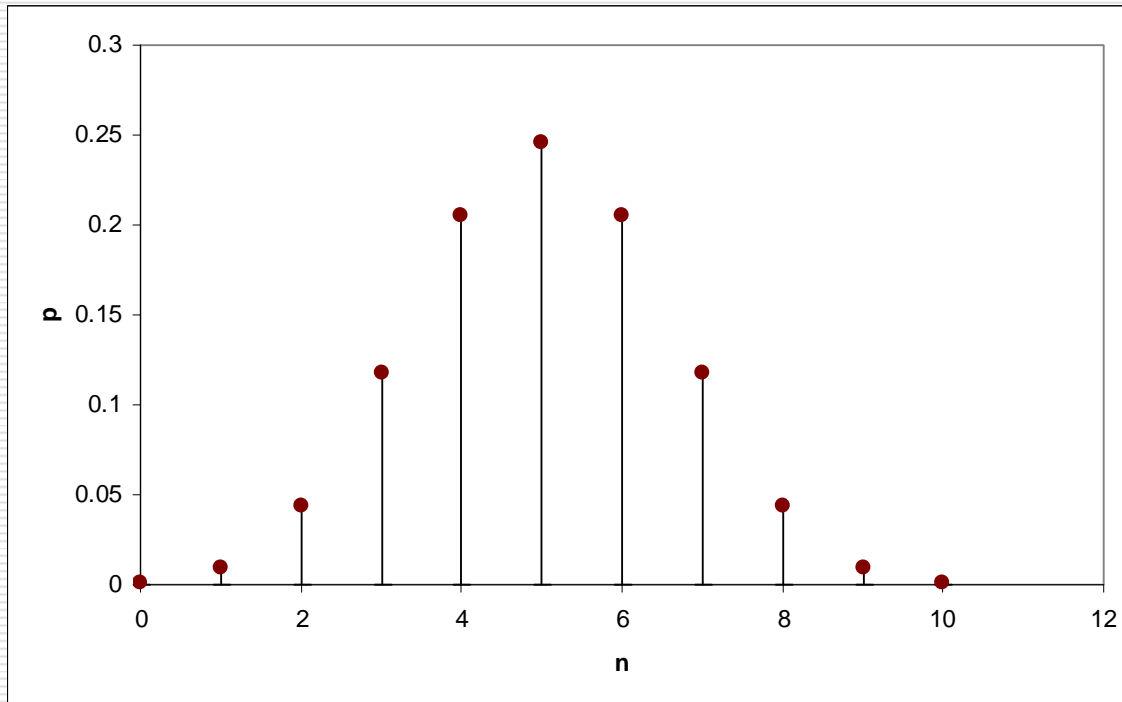
$$\sigma = \sqrt{\frac{p(1-p)}{n}}$$

p	1-p	$\sigma$
0.5	0.5	0.500
0.4	0.6	0.490
0.3	0.7	0.458
0.2	0.8	0.400
0.1	0.9	0.300

---

# Velikost vzorca

---

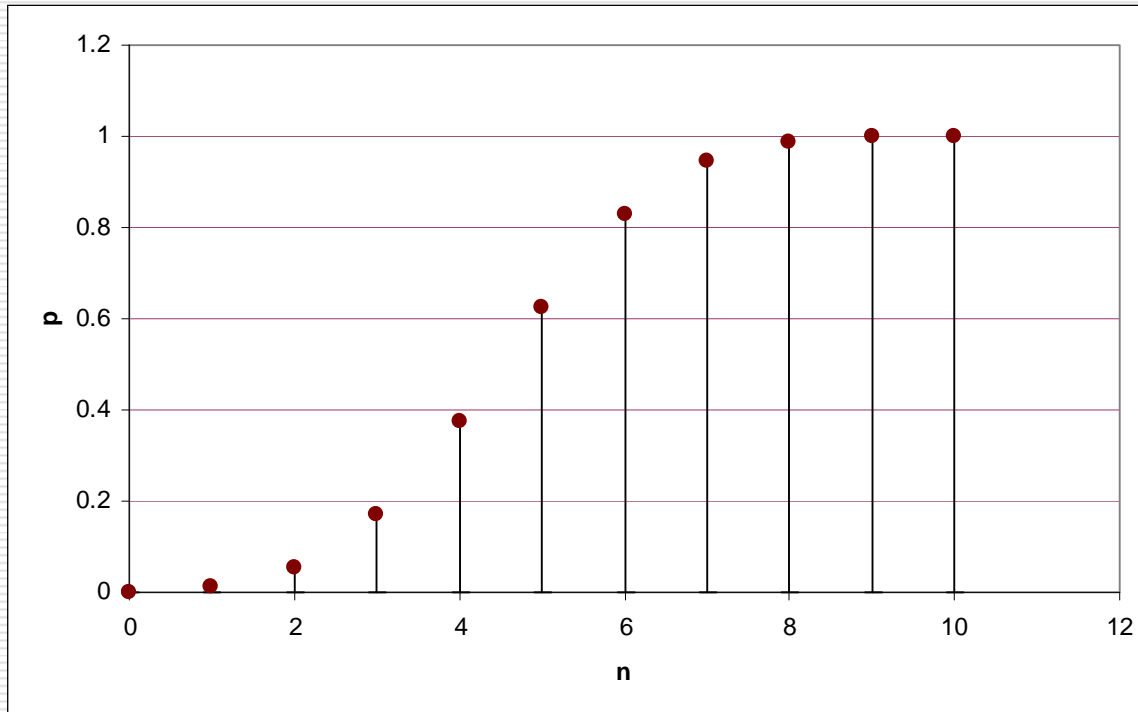


$N=10$   
 $P=0.5$

---

# Kumulativna gostota verjetnosti

---



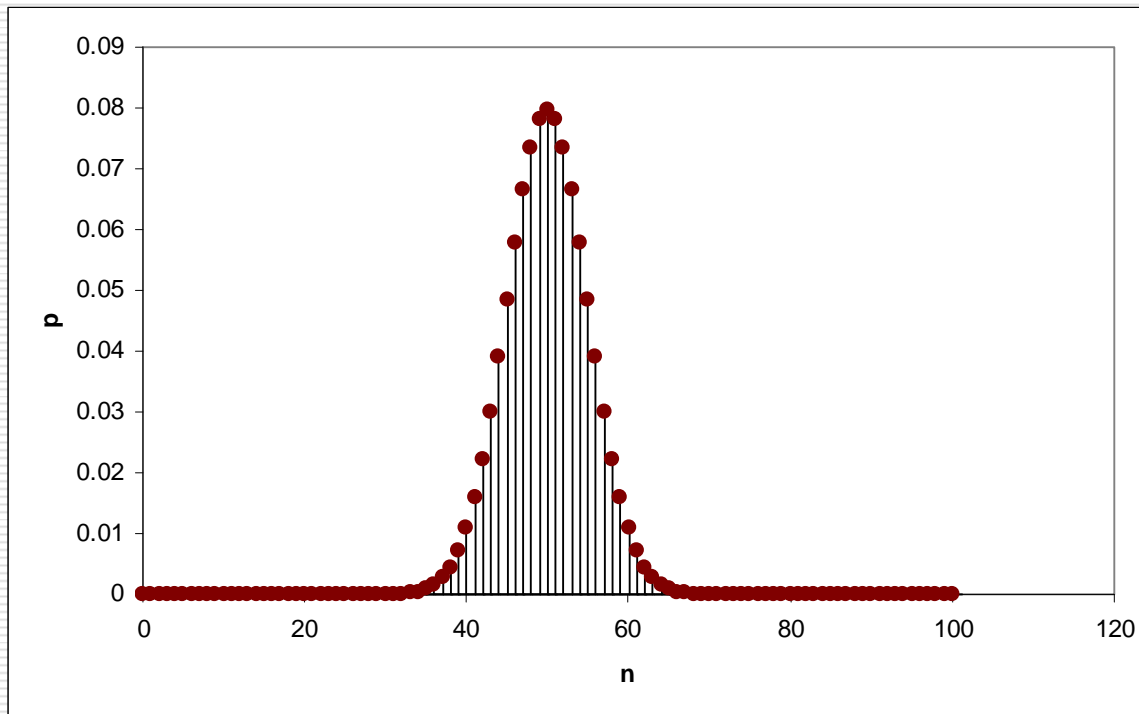
$N=10$

$P=0.5$

---

# Veliki vzorci

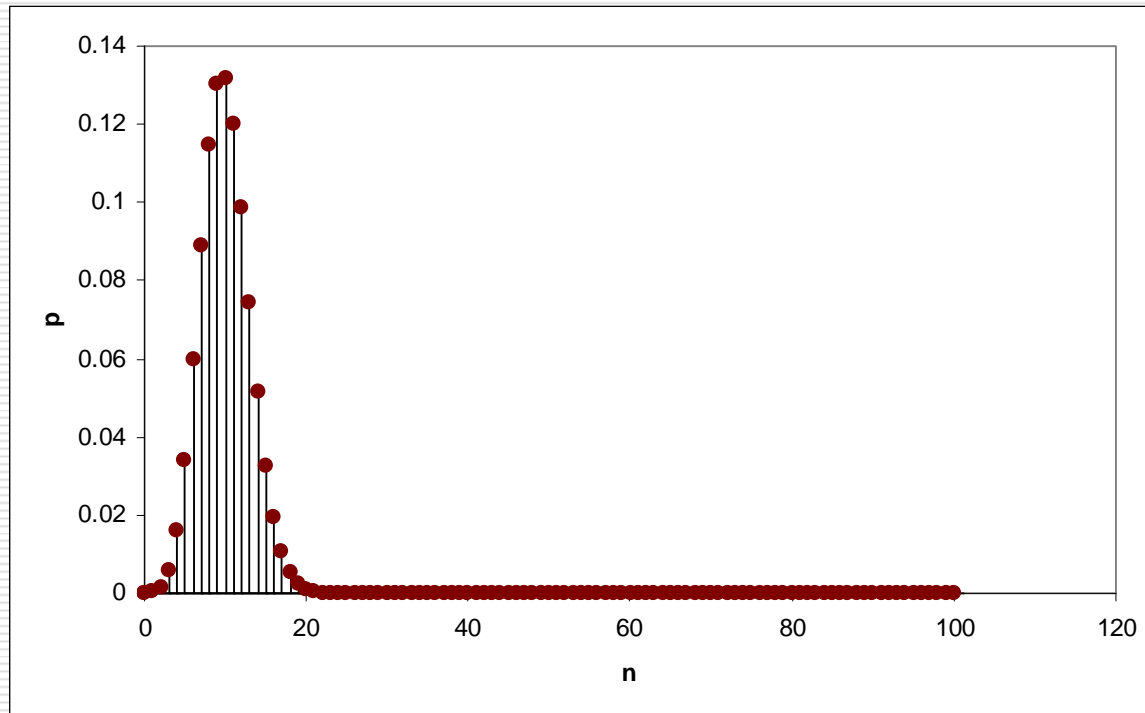
---



$N=100$

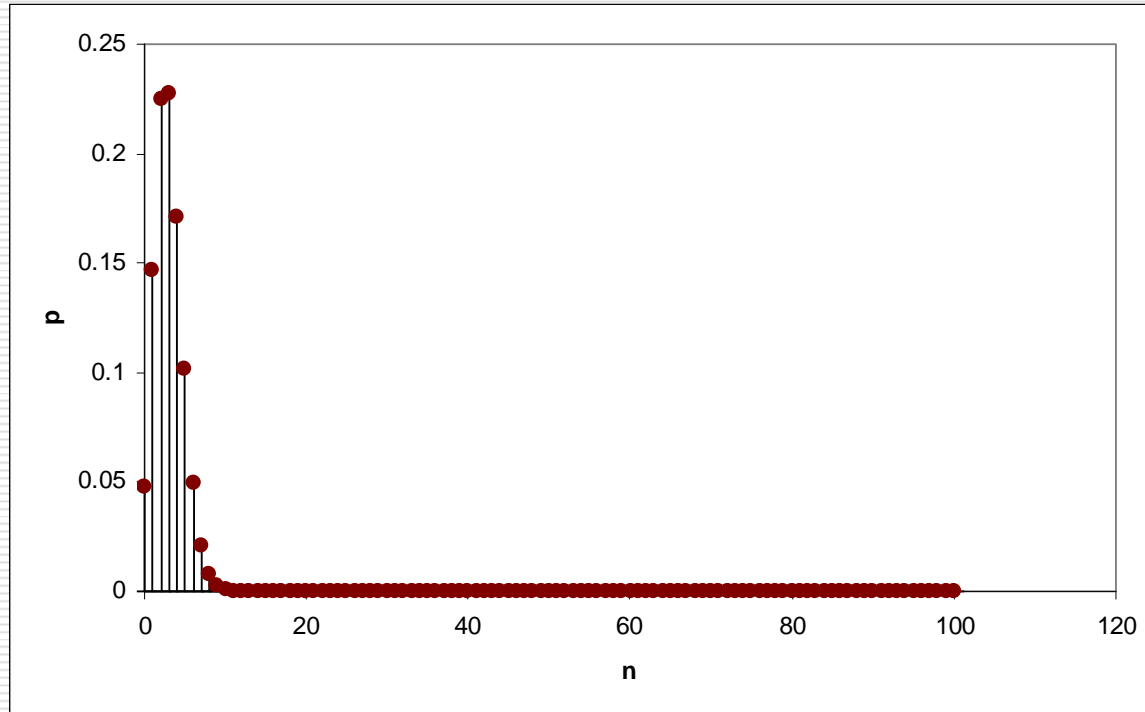
$P=0.5$

---



$N=100$   
 $P=0.1$

---



$N=100$   
 $P=0.03$

---

# Poissonova porazdelitev

- Ko je število poskusov veliko in je verjetnost dogodka majhna postane računanje binomskih verjetnosti zelo zahtevno
- Povprečno število uspehov v  $n$  poskusih je  $\lambda = np$ 
  - Primer: 64 smrti v 20 letih na 1000 vojakov



Simeon D. Poisson (1781-1840)



# Poissonova porazdelitev

---

- S substitucijo  $\lambda/n$  za  $p$  in pri velikem številu poskusov, postane Poissonova porazdelitev zelo dober približek za binomsko porazdelitev:

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

---

# Poissonova porazdelitev

## Emisija $\alpha$ -delcev

---

- Rutherford, Geiger, and Bateman (1910) so šteli  $\alpha$ -delce, ki jih je emitiral film iz polonija v 2608 zaporednih intervalih po 1/8 minute
    - $n?$
    - $p?$
  
  - Ali lahko rezultate poskusa opišemo s Poissonovo porazdelitvijo?
-

# Poissonova porazdelitev

## Emissija $\alpha$ -delcev

□ Izračun  $\lambda$ :

$$\begin{aligned}\lambda &= \text{Povprečno število delcev} \\ &\text{na interval} \\ &= 10097/2608 \\ &= 3.87\end{aligned}$$

□ Pričakovane vrednosti:

$$2608 \times P(x) = 2608 \times \frac{e^{-3.87}(3.87)^x}{x!}$$

No. $\alpha$ -particles	Observed
0	57
1	203
2	383
3	525
4	532
5	408
6	273
7	139
8	45
9	27
10	10
11	4
12	0
13	1
14	1
Over 14	0
Total	2608

# Poissonova porazdelitev

Emisija  $\alpha$ -delcev

---

No. $\alpha$ -particles	Observed	Expected
0	57	54
1	203	210
2	383	407
3	525	525
4	532	508
5	408	394
6	273	255
7	139	140
8	45	68
9	27	29
10	10	11
11	4	4
12	0	1
13	1	1
14	1	1
Over 14	0	0
Total	2608	2608

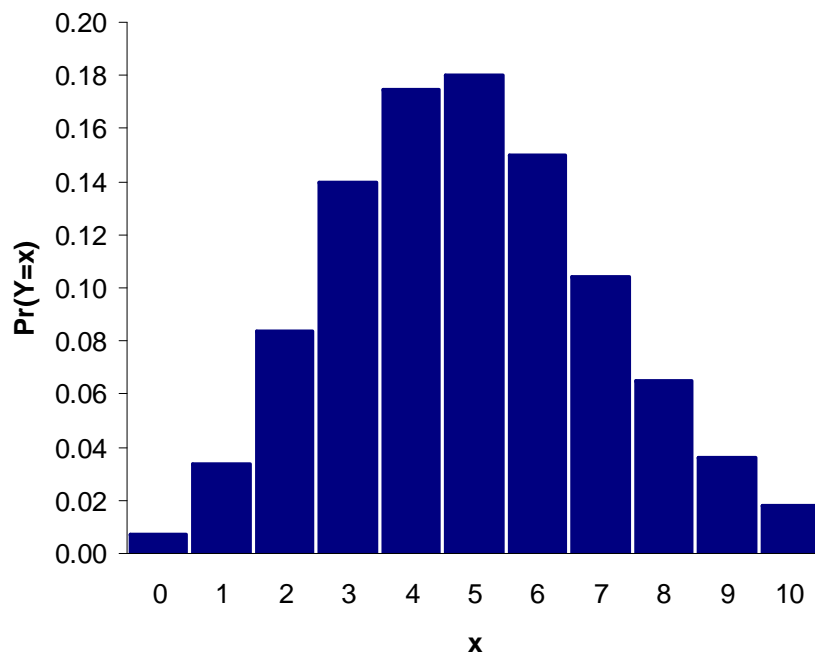
---

Tipično število epileptičnih napadov

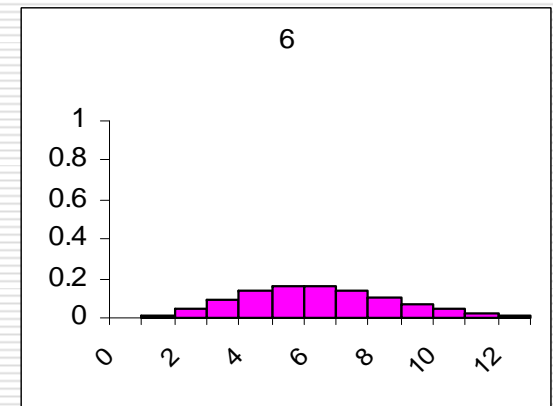
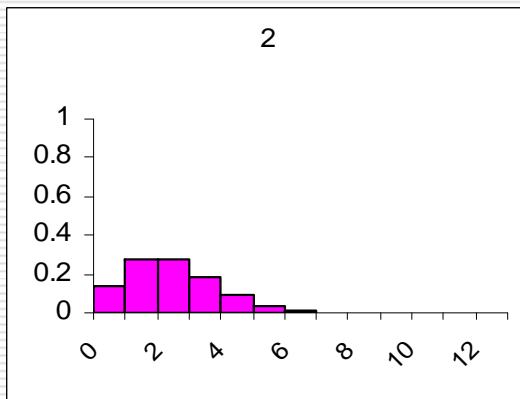
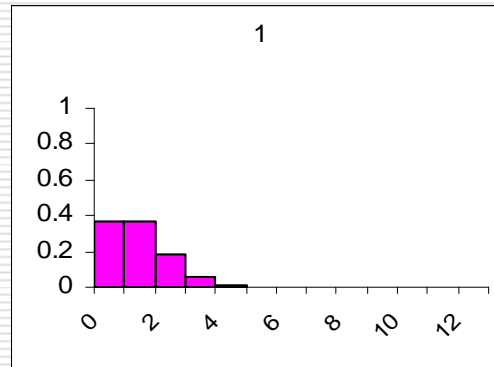
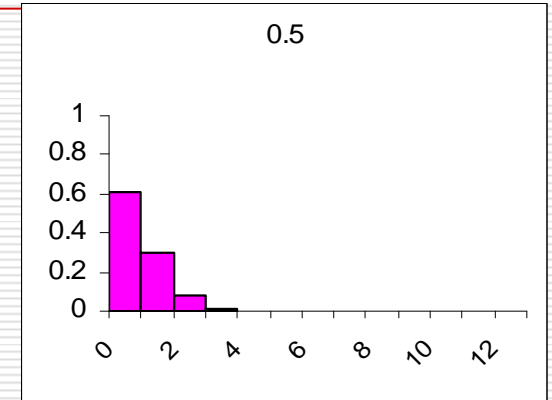
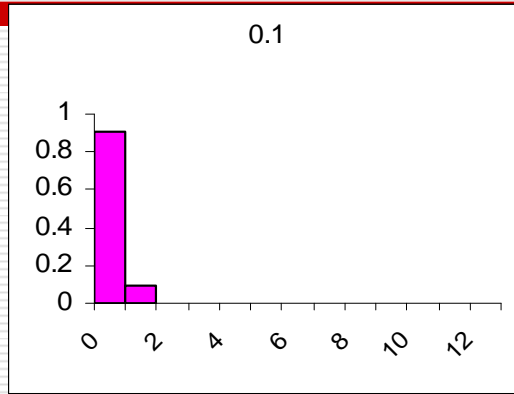
na mesec:  $\lambda=5$

$$P(Y_i = x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

<b>x=</b>	<b>Pr(Y=x)</b>
<b>0</b>	<b>0.007</b>
<b>1</b>	<b>0.034</b>
<b>2</b>	<b>0.084</b>
<b>3</b>	<b>0.140</b>
<b>4</b>	<b>0.175</b>
<b>5</b>	<b>0.180</b>
<b>6</b>	<b>0.150</b>
<b>7</b>	<b>0.104</b>
<b>8</b>	<b>0.065</b>
<b>9</b>	<b>0.036</b>
<b>10</b>	<b>0.018</b>



# Poissonova porazdelitev



# Pričakovana vrednost diskretne slučajne spremenljivke

---

$$E(X) = \sum_{i=1}^n a_i p_i = a_1 p_1 + a_2 p_2 + \dots + a_n p_n$$

---

# Varianca slučajne diskretne spremenljivke

---

$$\sigma^2(X) = E[X - E(X)]^2$$

$$= \sum_{i=1}^n p_i \left( a_i - \sum_{i=1}^n a_i p_i \right)^2$$

---



# Pričakovana vrednost zvezne slučajne spremenljivke

---

$$E(X) = \int xf(x)dx$$

Za enakomerno porazdelitev spremenljivke  $x$ , kjer je  $f(x)$  definiran na intervalu  $[a,b]$  in je  $a < b$ :

$$E(X) = (b + a) / 2 \quad \text{in} \quad \sigma^2(X) = \frac{(b - a)^2}{12}$$

---

# Normalna porazdelitev

- 1733 de Moivre (aproksimacija binomske porazdelitve za velike  $n$ )
- 1809 Gauss
- Pomembna zaradi **centralnega limitnega izreka**, ki pravi, da je vsota velikega števila neodvisnih slučajnih spremenljivk (binomska porazdelitev, Poissonova porazdelitev, ...) porazdeljena normalno
- Primer: Telesna teža človeka je odvisna od številnih dejavnikov (genetski in okoljski, njihovi vplivi so aditivni. Telesna teža je zato porazdeljena normalno.



Abraham de Moivre  
(1667-1754)



Karl F. Gauss  
(1777-1855)

# Normalna porazdelitev

---

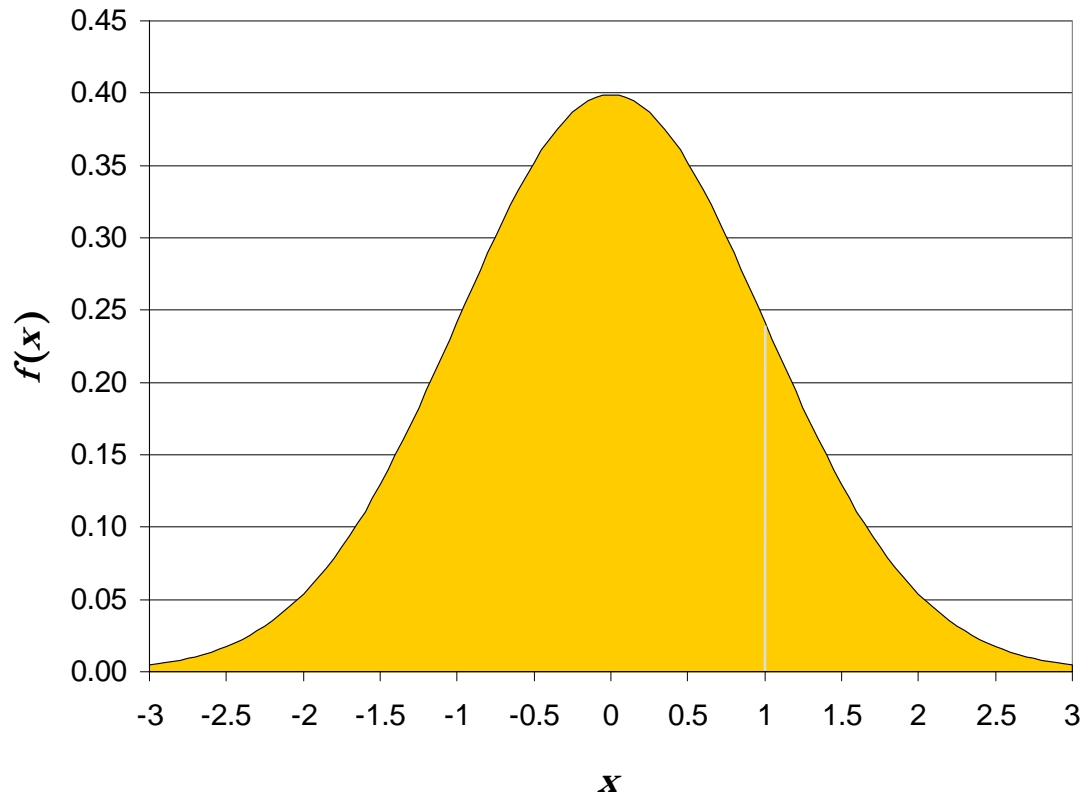
- Zvezna slučajna spremenljivka je porazdeljena normalno z aritmetično sredino  $\mu$  in varianco  $\sigma^2$  če je njena gostota verjetnosti:

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

- $f(x)$  ni isto kot  $P(x)$ 
    - $P(x) = 0$  za vsak  $x$ , ker je normalna porazdelitev zvezna
    - Toda,  $P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f(x) dx$
-

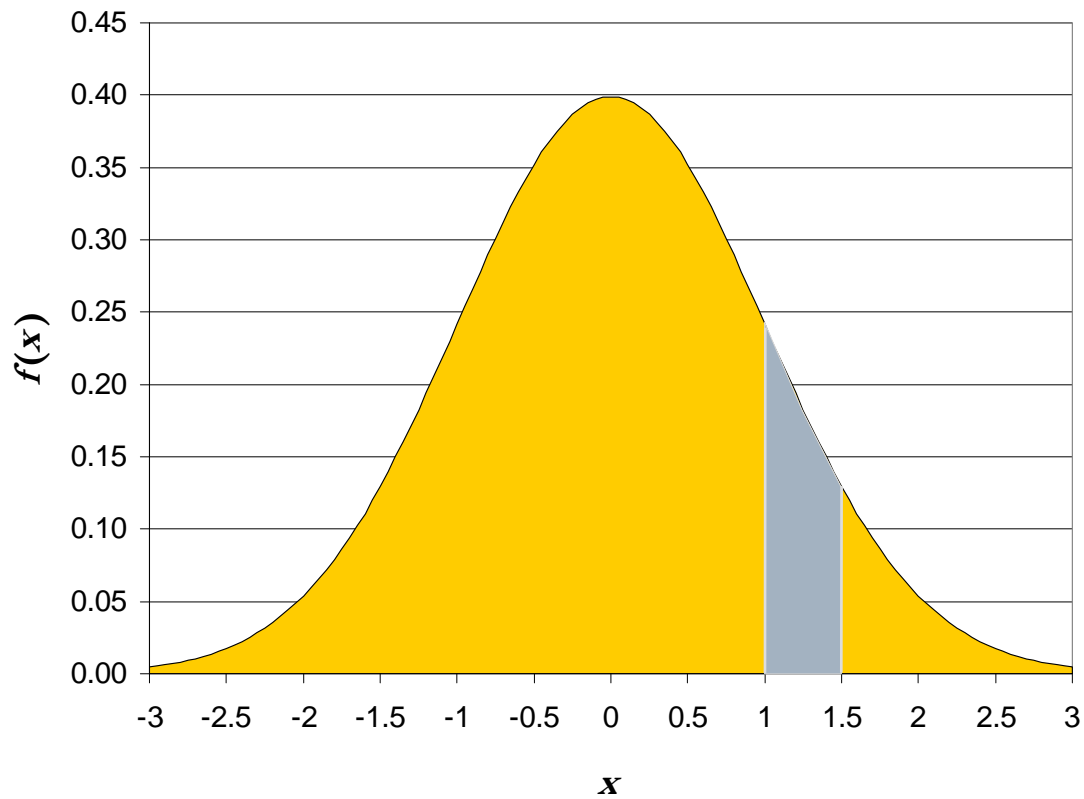
# Normalna porazdelitev

---



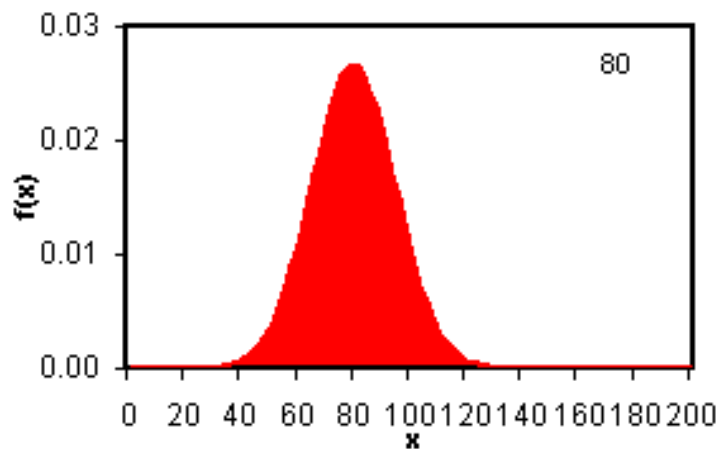
# Normalna porazdelitev

---

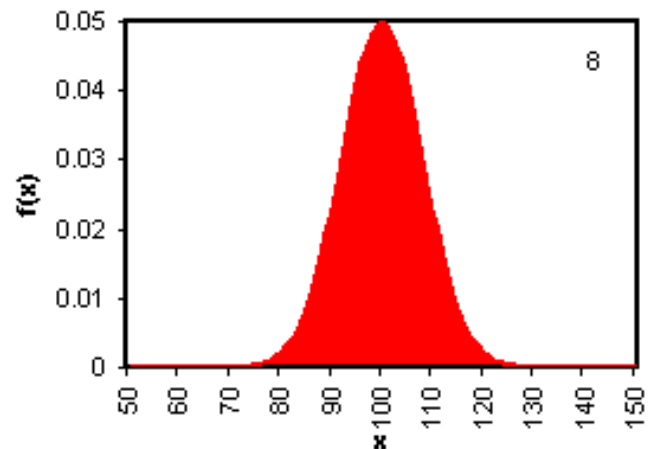


# Normalna porazdelitev

---



Aritmetična sredina

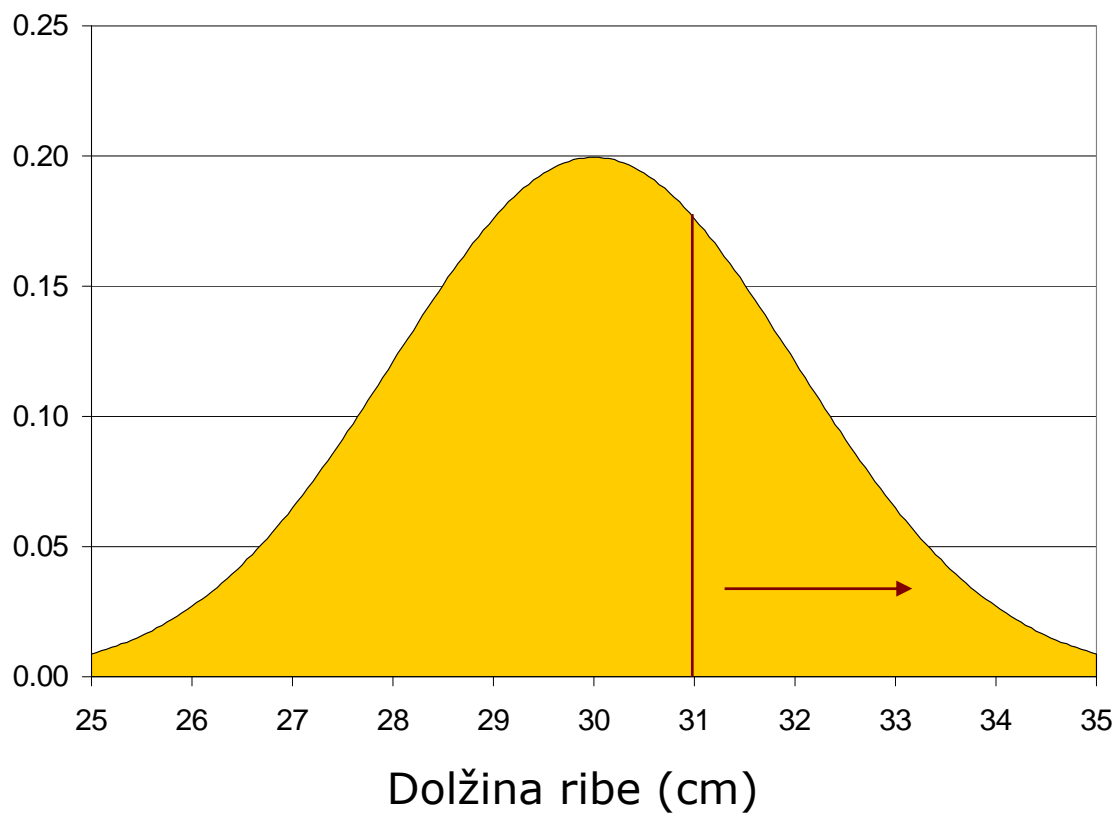


Varianca

---

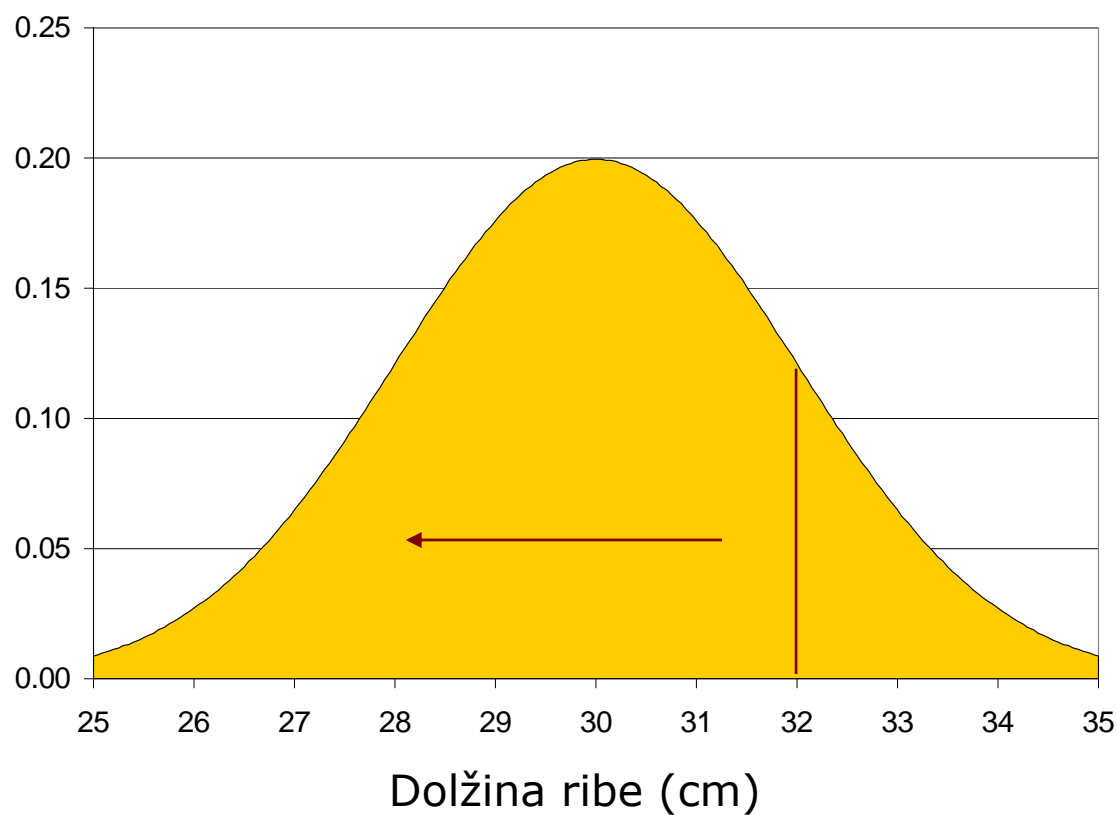
# Normalna porazdelitev

- Kakšna je verjetnost, da bo ujeta postrv dolga vsaj 31 cm?



# Normalna porazdelitev

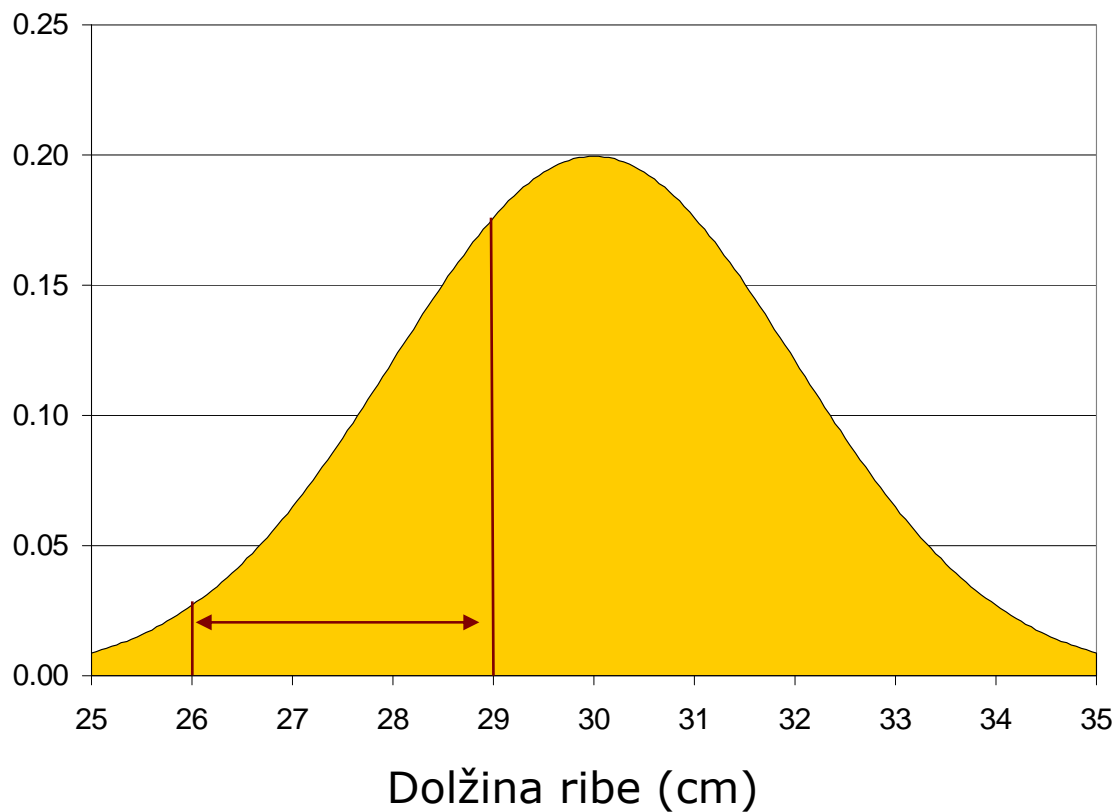
- Kakšna je verjetnost, da bo ujeta postrv krajša od 32 cm?





# Normalna porazdelitev

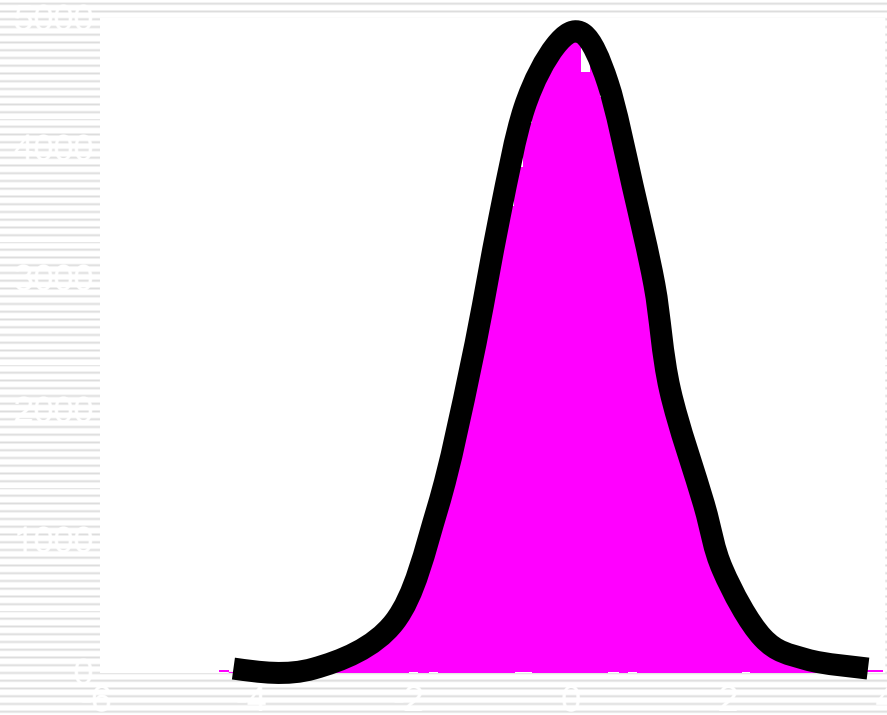
- Kakšna je verjetnost, da bo ujeta postrv dolga med 26 in 29 cm?



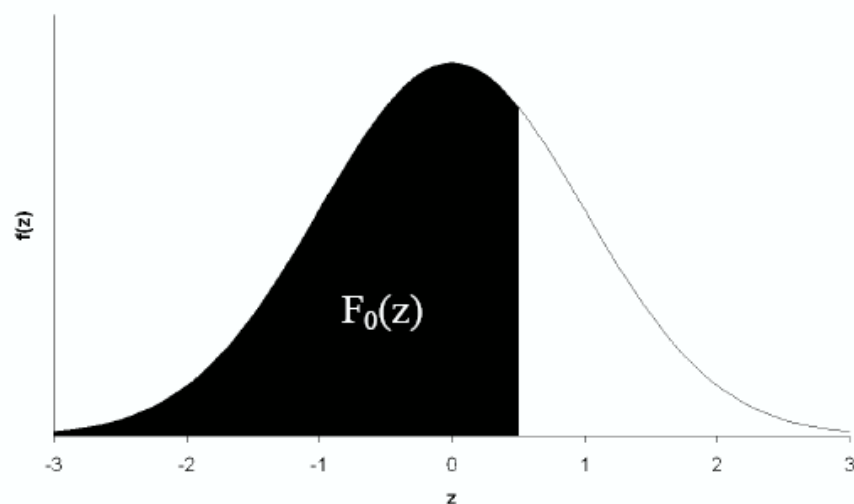
# Standardizirana normalna porazdelitev

---

□  $\mu=0$  and  $\sigma^2=1$



## Tabela standardizirane normalne distribucije.



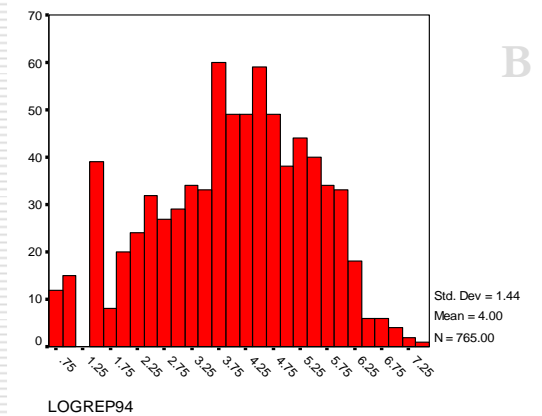
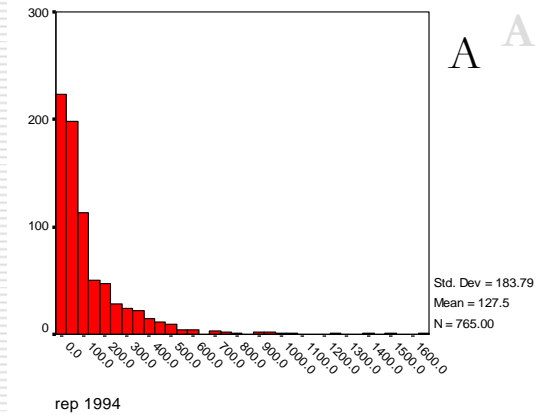
Stopnja tveganja ( $\alpha$ )	Enostransko tveganje	Dvostransko tveganje
0.1	1.282	1.645
0.05	1.645	1.960
0.025	1.960	2.241
0.01	2.326	2.576
0.005	2.576	2.807

$F_0(z)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879

# Log-normalna porazdelitev

- X je **log-normalno porazdeljena slučajna spremenljivka**, če so njene logaritemske vrednosti  $\ln(X)$ , porazdeljene normalno.
- Učinki posameznih dejavnikov se množijo (multiplikativni model)
- Porazdelitev spremenljivke je asimetrična v desno (A), njene logaritemske vrednosti pa se porazdeljujejo normalno (B).

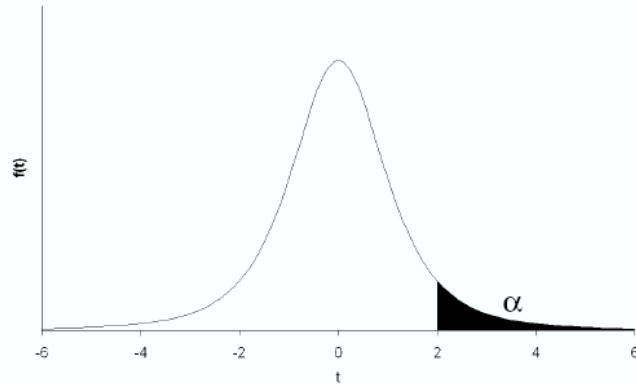


# Porazdelitev

---

- Binomska
  - Poissonova
  - Normalna
    - Standardizirana normalna
  - Studentova porazdelitev (t porazdelitev)
  - $\chi^2$  porazdelitev
  - F porazdelitev
-

# Studentova porazdelitev



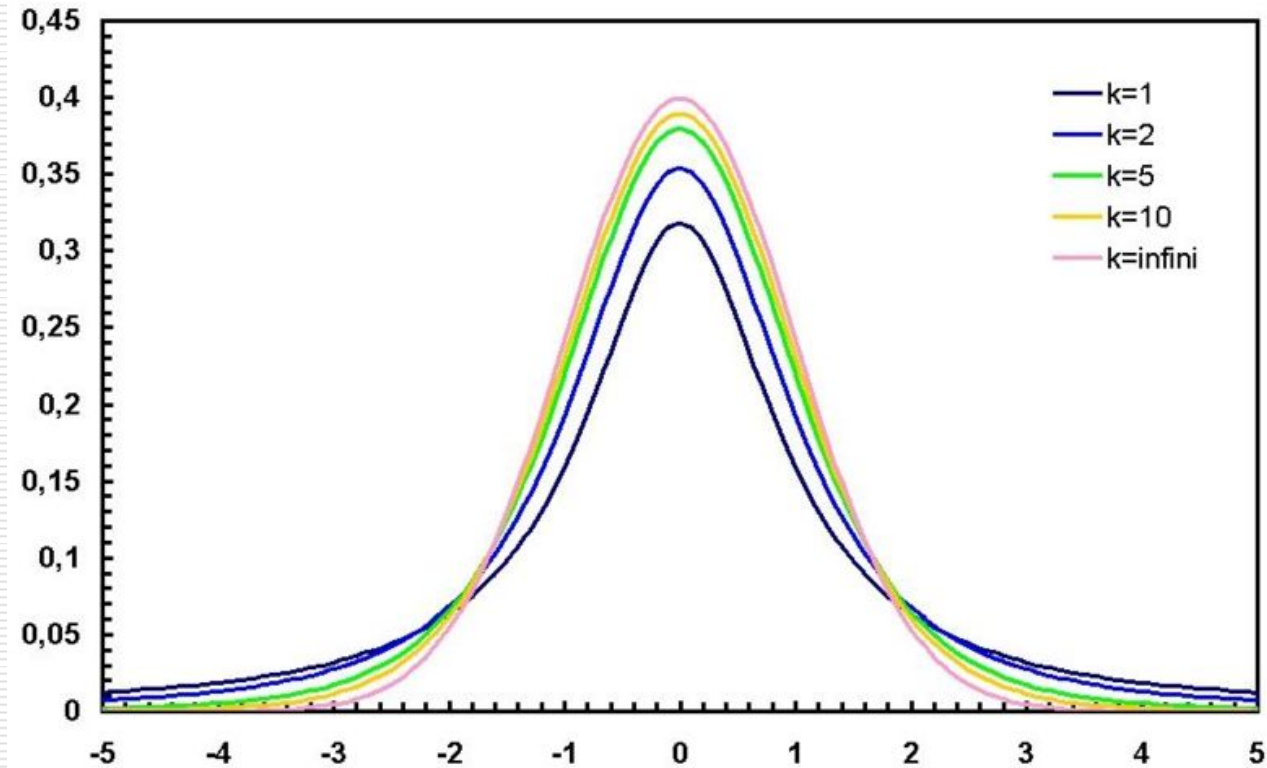
$$t = \frac{\bar{X} - \mu}{s / \sqrt{N}}$$

t

Število prostostnih stopenj	Stopnja tveganja (α)				
	0.4	0.2	0.1	0.05	0.01
Dvostransko tveganje	0.4	0.2	0.1	0.05	0.01
Enostransko tveganje	0.2	0.1	0.05	0.025	0.005
1	1.376	3.078	6.314	12.706	63.656
2	1.061	1.886	2.920	4.303	9.925
3	0.978	1.638	2.353	3.182	5.841
4	0.941	1.533	2.132	2.776	4.604
5	0.920	1.476	2.015	2.571	4.032
6	0.906	1.440	1.943	2.447	3.707
7	0.896	1.415	1.895	2.365	3.499
8	0.889	1.397	1.860	2.306	3.355
9	0.883	1.383	1.833	2.262	3.250

# Studentova porazdelitev

---

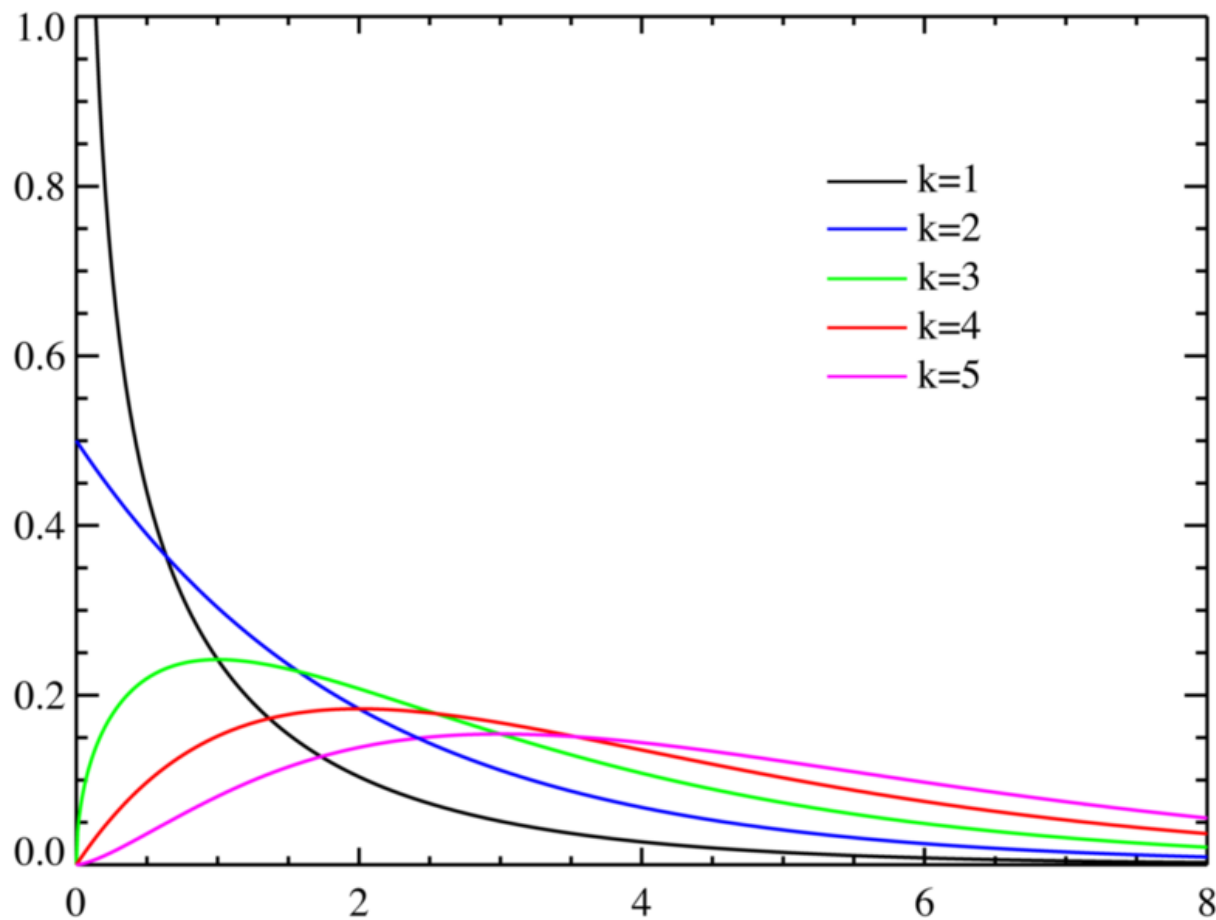


William Seally Gosset, Guinness Brew., Dublin

---

# $\chi^2$ porazdelitev

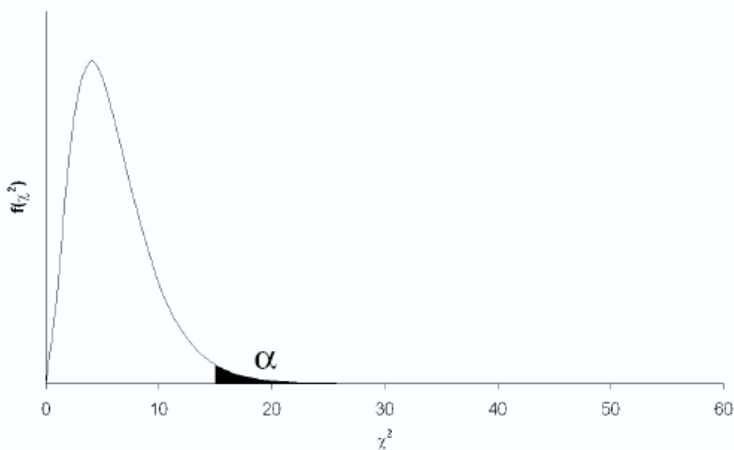
---



$$\chi^2 = \sum Z_i^2$$



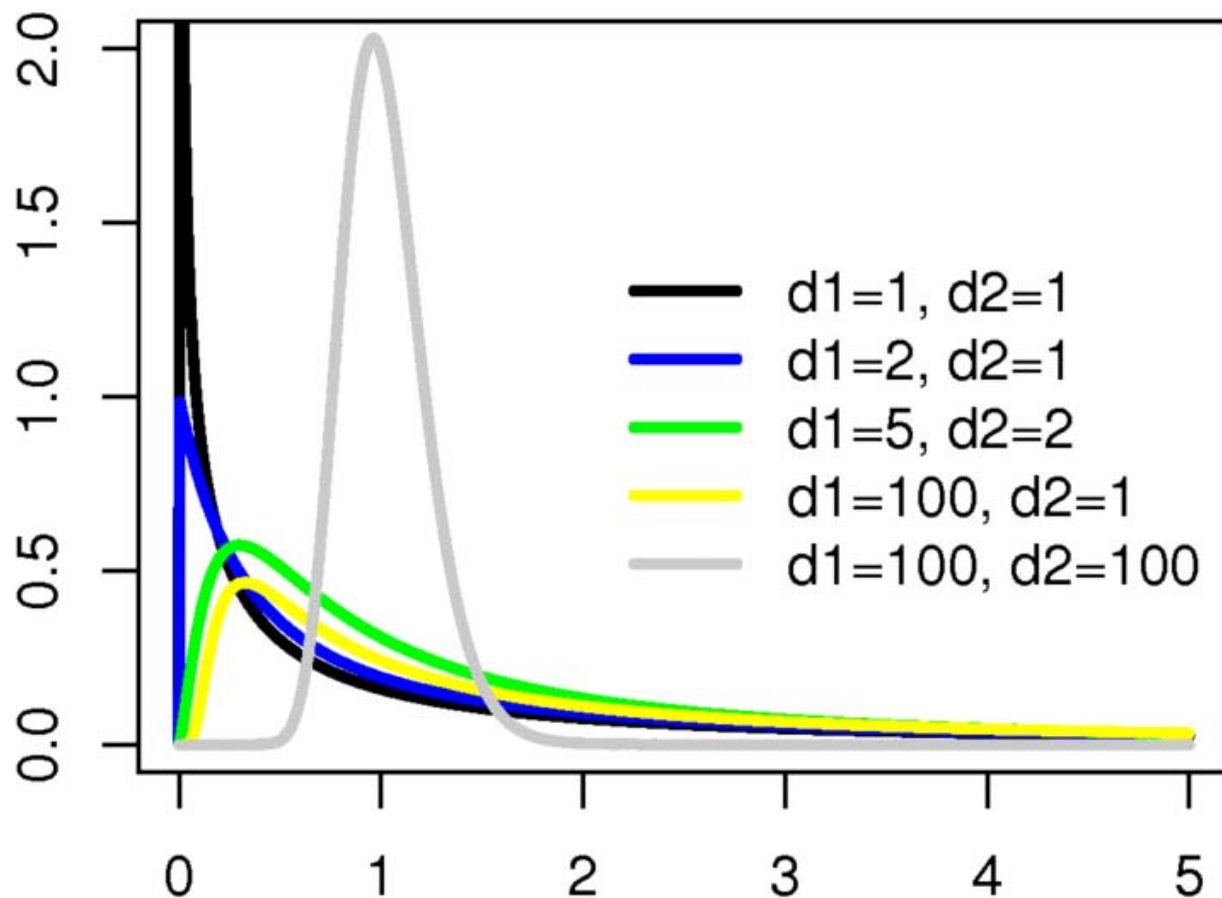
# $\chi^2$ porazdelitev



$\chi^2$

Število prostostnih stopenj	Stopnja tveganja ( $\alpha$ )				
	0.001	0.01	0.025	0.05	0.1
1	10.8274	6.6349	5.0239	3.8415	2.7055
2	13.8150	9.2104	7.3778	5.9915	4.6052
3	16.2660	11.3449	9.3484	7.8147	6.2514
4	18.4662	13.2767	11.1433	9.4877	7.7794
5	20.5147	15.0863	12.8325	11.0705	9.2363
6	22.4575	16.8119	14.4494	12.5916	10.6446
...	...	...	...	...	...

# F porazdelitev Fisher-Snedecor



$$F = \frac{s_1^2}{s_2^2}$$



# Centralni limitni izrek in intervalna ocena

---

Iztok Grabnar

Univerza v Ljubljani, Fakulteta za farmacijo

*Oktober, 2011*

# Diskretna spremenljivka

## Enakomerna porazdelitev

$$X = \begin{cases} 1 & \text{with probability } 1/3, \\ 2 & \text{with probability } 1/3, \\ 3 & \text{with probability } 1/3. \end{cases}$$

Enakomerna porazdelitev

o	o	o
-----		
1	2	3

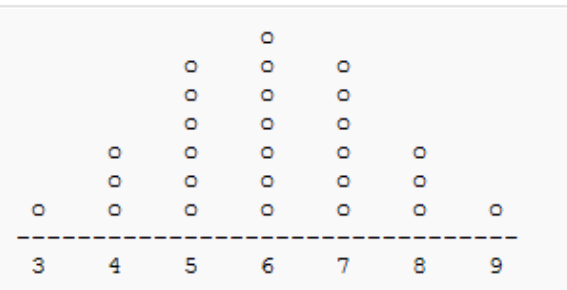
$$X_1 + X_2$$

$$\left. \begin{array}{l} 1 + 1 = 2 \\ 1 + 2 = 3 \\ 1 + 3 = 4 \\ 2 + 1 = 3 \\ 2 + 2 = 4 \\ 2 + 3 = 5 \\ 3 + 1 = 4 \\ 3 + 2 = 5 \\ 3 + 3 = 6 \end{array} \right\} = \begin{cases} 2 & \text{with probability } 1/9 \\ 3 & \text{with probability } 2/9 \\ 4 & \text{with probability } 3/9 \\ 5 & \text{with probability } 2/9 \\ 6 & \text{with probability } 1/9 \end{cases}$$

		o		
o	o	o	o	o
-----				
2	3	4	5	6

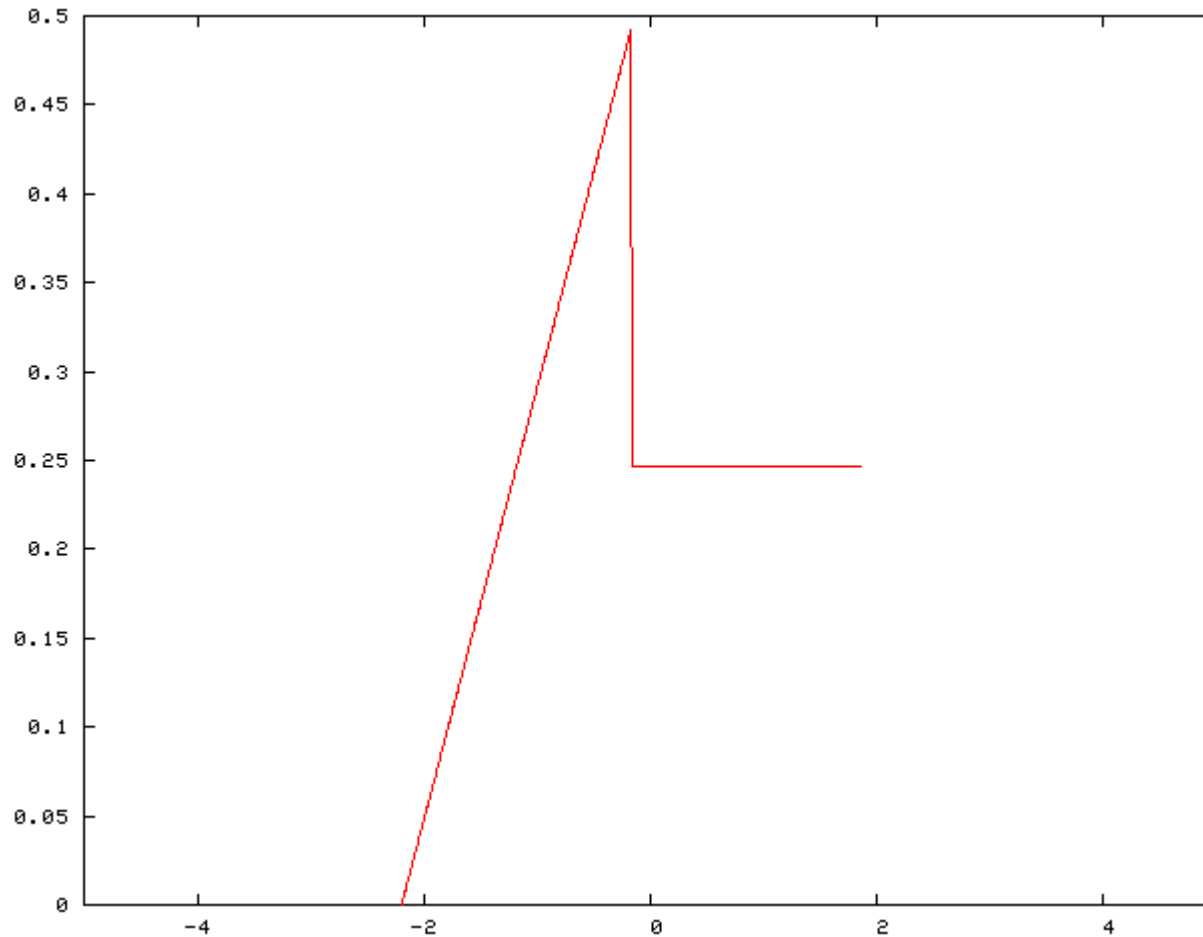
$$X_1 + X_2 + X_3$$

$$\left. \begin{array}{l}
 1+1+1 = 3 \\
 1+1+2 = 4 \\
 1+1+3 = 5 \\
 1+2+1 = 4 \\
 1+2+2 = 5 \\
 1+2+3 = 6 \\
 1+3+1 = 5 \\
 1+3+2 = 6 \\
 1+3+3 = 7 \\
 2+1+1 = 4 \\
 2+1+2 = 5 \\
 2+1+3 = 6 \\
 2+2+1 = 5 \\
 2+2+2 = 6 \\
 2+2+3 = 7 \\
 2+3+1 = 6 \\
 2+3+2 = 7 \\
 2+3+3 = 8 \\
 3+1+1 = 5 \\
 3+1+2 = 6 \\
 3+1+3 = 7 \\
 3+2+1 = 6 \\
 3+2+2 = 7 \\
 3+2+3 = 8 \\
 3+3+1 = 7 \\
 3+3+2 = 8 \\
 3+3+3 = 9
 \end{array} \right\} = \left\{ \begin{array}{l}
 3 \text{ with probability } 1/27 \\
 4 \text{ with probability } 3/27 \\
 5 \text{ with probability } 6/27 \\
 6 \text{ with probability } 7/27 \\
 7 \text{ with probability } 6/27 \\
 8 \text{ with probability } 3/27 \\
 9 \text{ with probability } 1/27
 \end{array} \right.$$



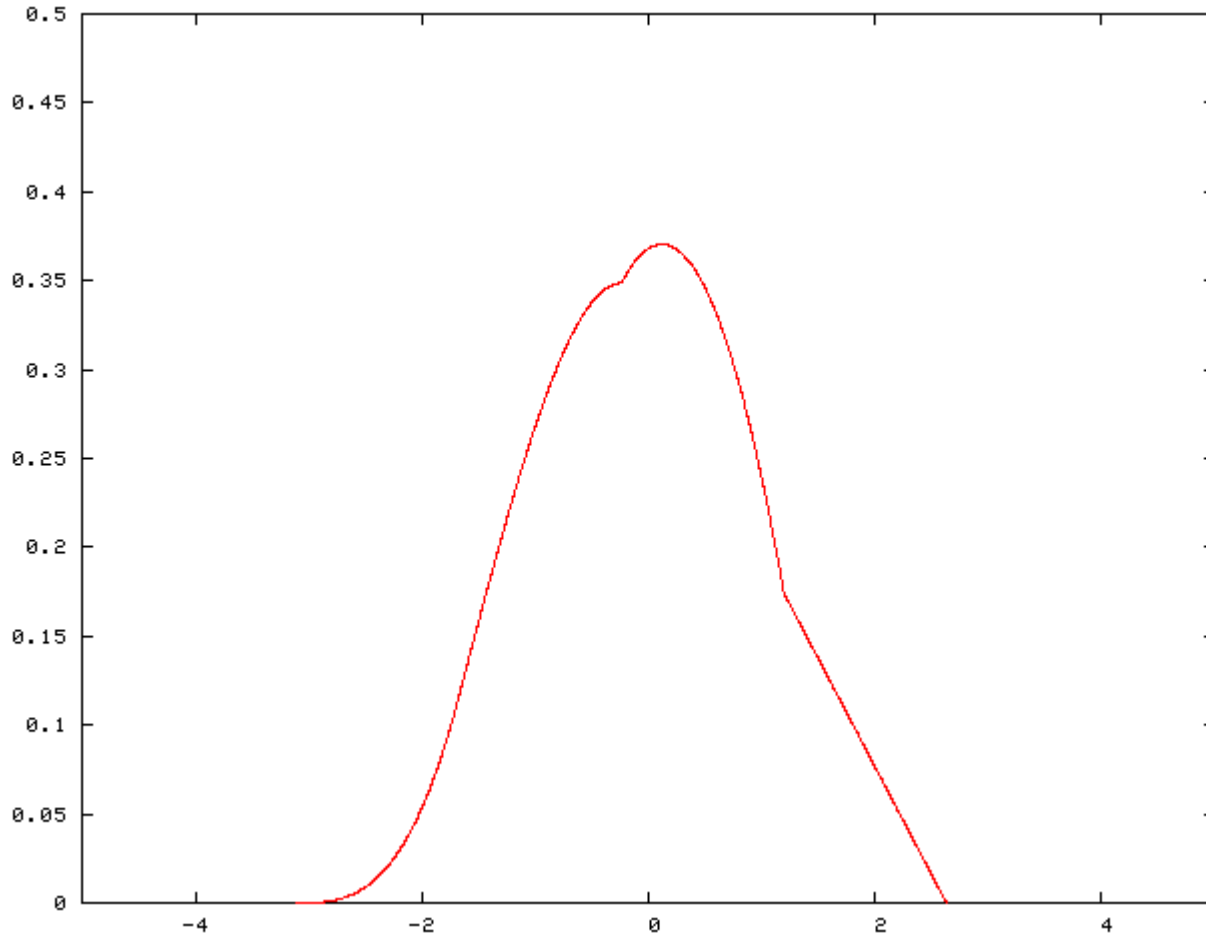
# Gostota verjetnosti X

---



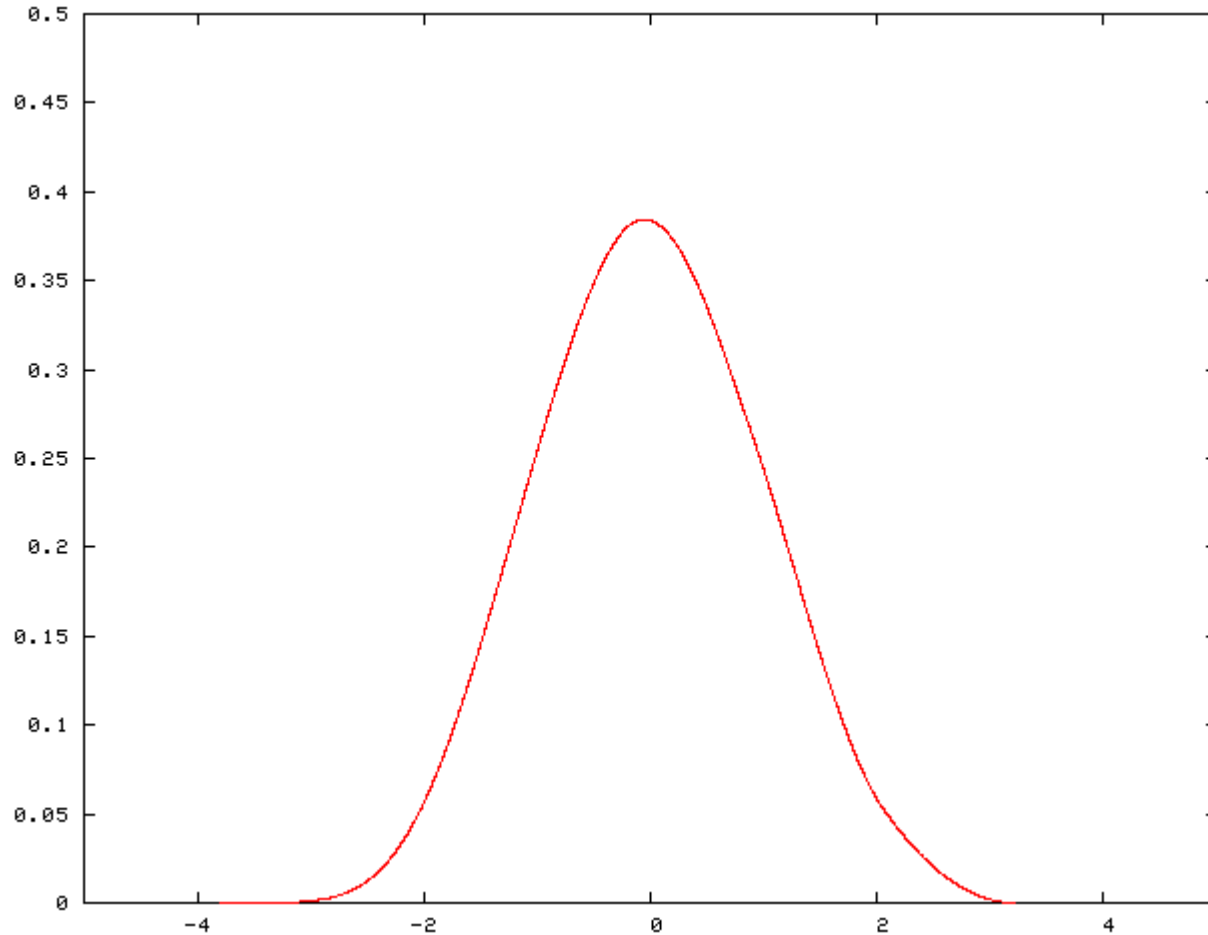
# Gostota verjetnosti $X_1 + X_2$

---



# Gostota verjetnosti $X_1 + X_2 + X_3$

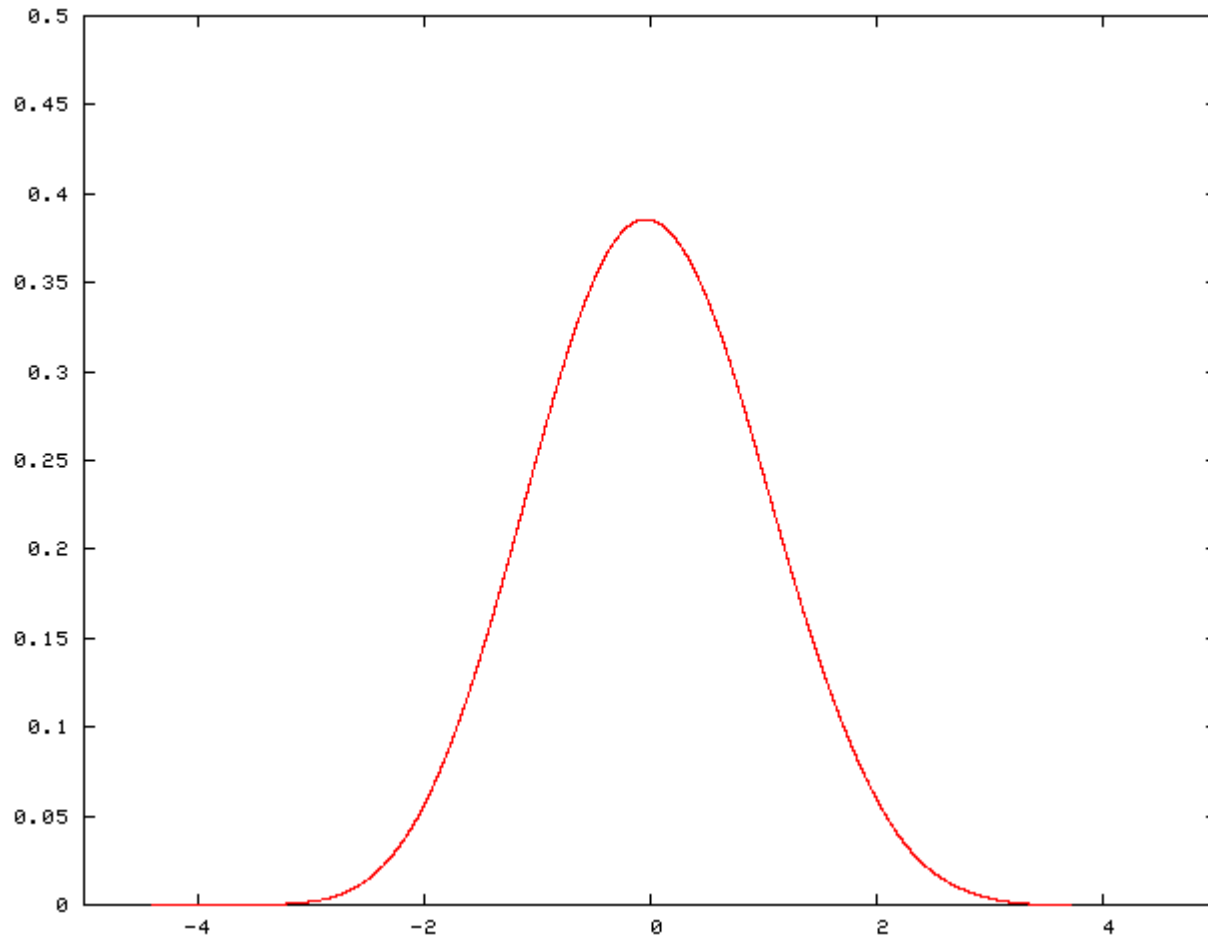
---





# Gostota verjetnosti $X_1 + X_2 + X_3 + X_4$

---



# Porazdelitev vzorčnih aritmetičnih sredin

## Vzorčna porazdelitev

---

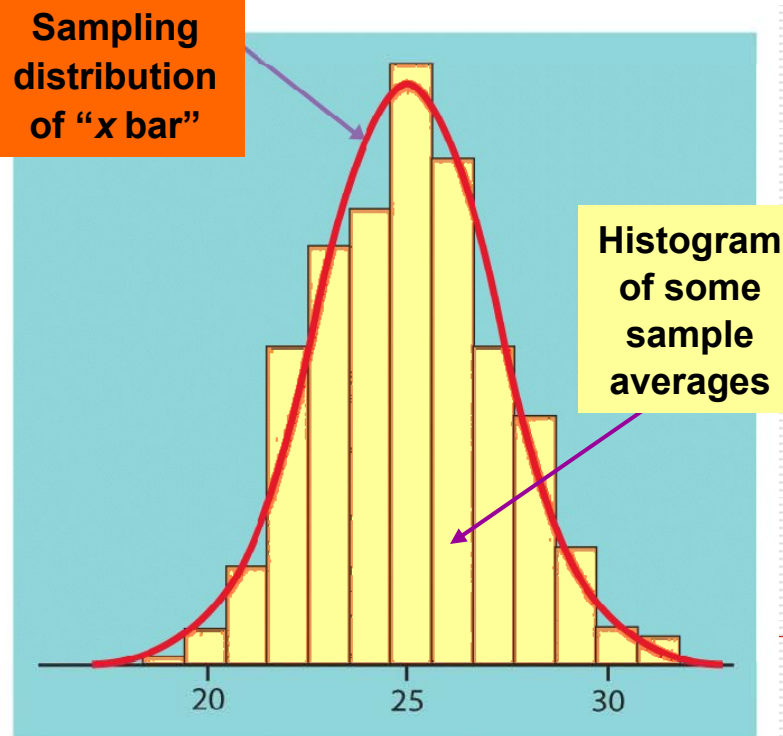
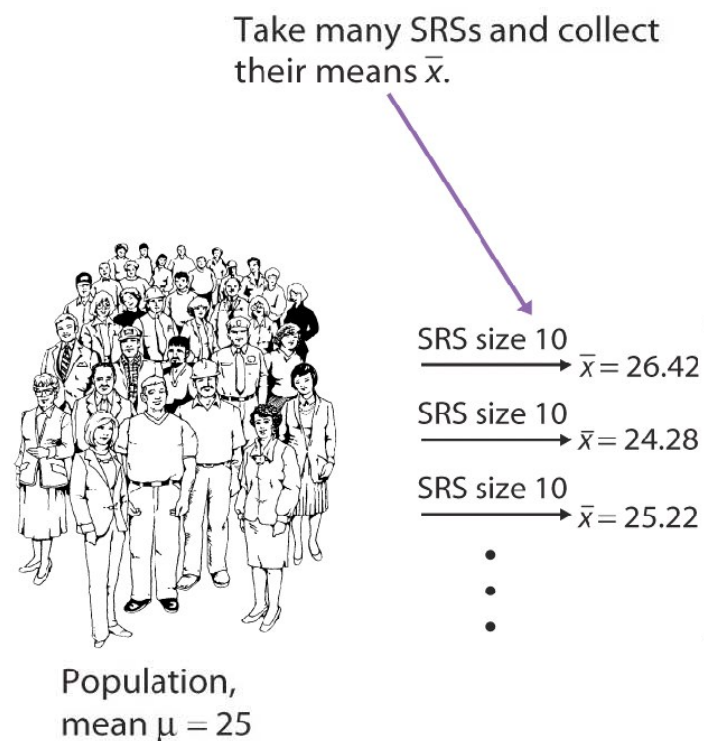
Vzorčna porazdelitev **statistike** je njena gostota verjetnosti za neskončno veliko vzorcev velikosti  $N$  iz populacije.

---

# Porazdelitev vzorčnih aritmetičnih sredin

We take many random samples of a given size  $n$  from a population with mean  $\mu$  and standard deviation  $\sigma$ .

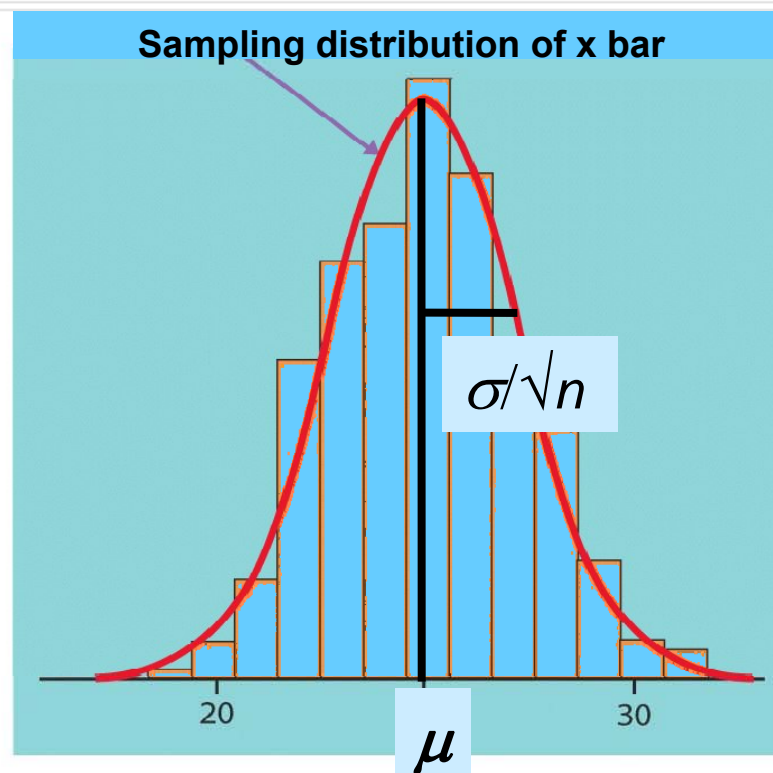
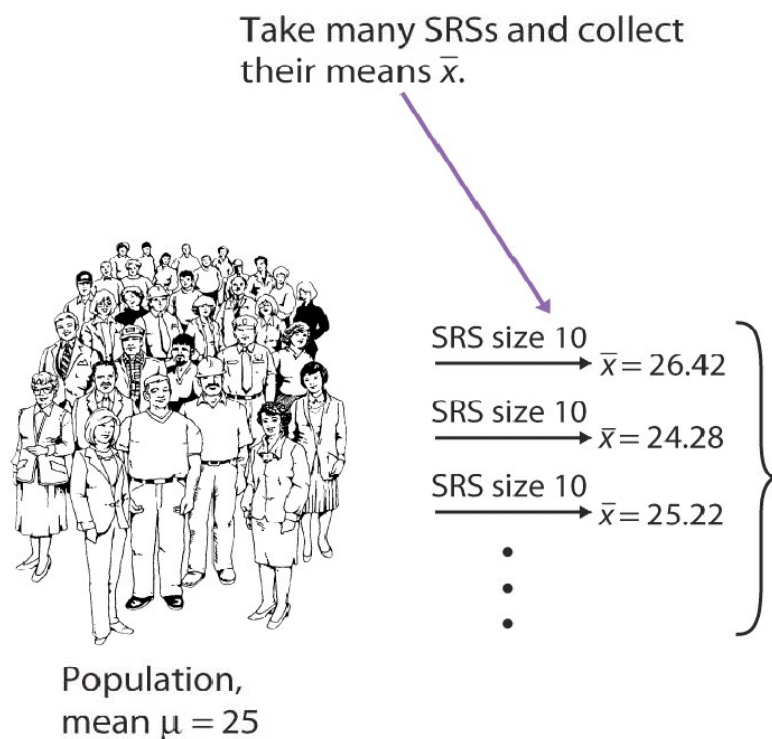
Some sample means will be above the population mean  $\mu$  and some will be below, making up the sampling distribution.



# Za vsako populacijo z aritmetično sredino $\mu$ in standardnim odklonom $\sigma$ velja

□ **Aritmetična sredina** ali centralna lega vzorčne porazdelitve, je enaka aritmetični sredini populacije  $\mu$ :  $\mu_x = \mu$ .

□ **Standardni odklon** vzorčne porazdelitve pa je  $\sigma/\sqrt{n}$ , kjer je  $n$  velikost vzorca in  $\sigma$  standardni odklon v populaciji:  $\sigma_x = \sigma/\sqrt{n}$ .



□ **Aritmetična sredina vzorčne porazdelitve:**

---

Je nepristranska ocena za **aritmetično sredino populacije**  $\mu$ .

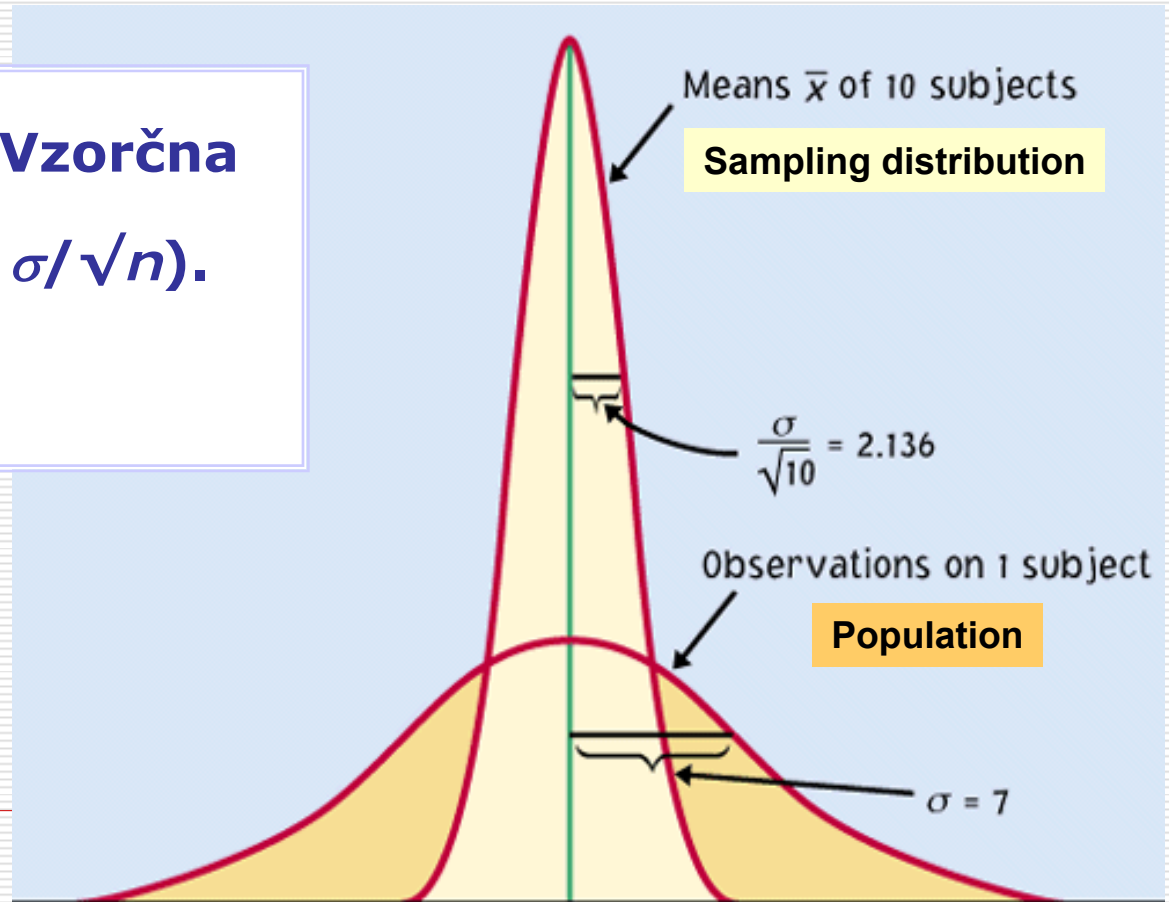
□ **Standardni odklon vzorčne porazdelitve:**

Standardna napaka  $\sigma / \sqrt{n}$ . → **Variabilnost povprečij je manjša od variabilnosti posameznih opazovanj.**

---

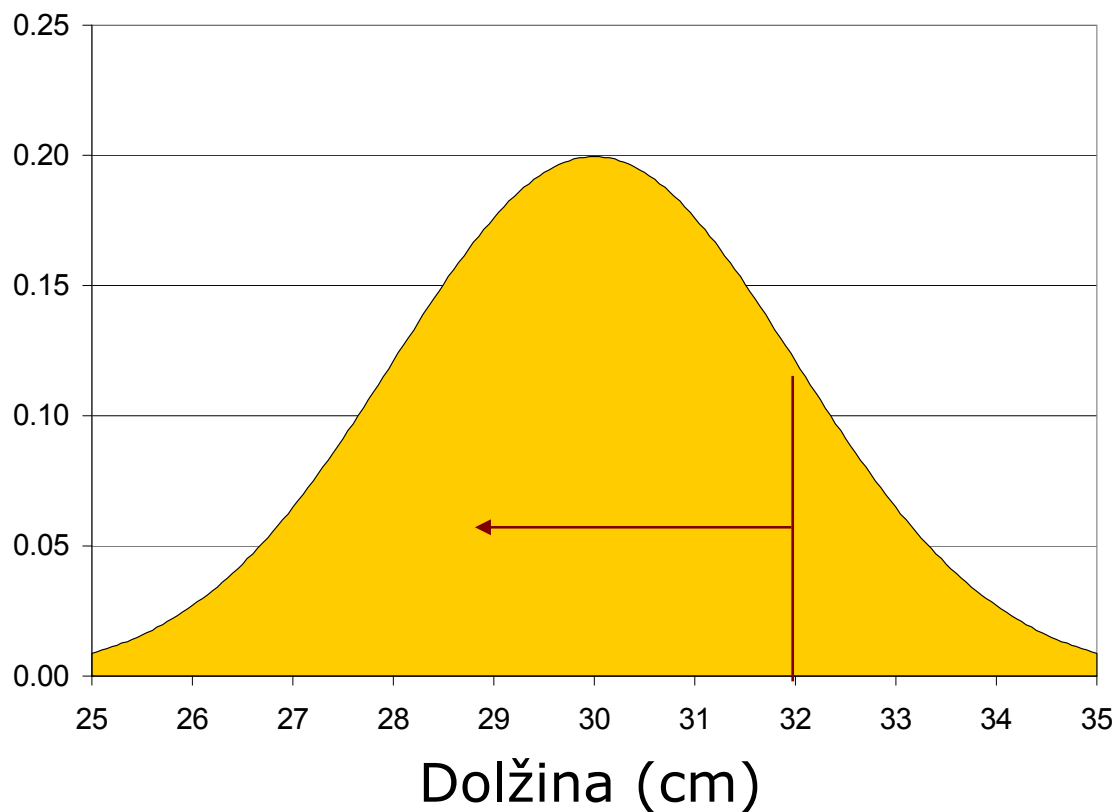
# Normalno porazdeljena spremenljivka

Populacija  $N(\mu, \sigma)$  Vzorčna porazdelitev  $N(\mu, \sigma/\sqrt{n})$ .



# Normalna porazdelitev

- Kakšna je verjetnost, da bo ujeta postrv krajša od 32 cm?



$$\mu = 30 \text{ cm}$$
$$\sigma = 1.5 \text{ cm}$$

# Primer

---

A certain brand of tires has a mean life of 25,000 km with a standard deviation of 1,600 km.

What is the probability that the mean life of 64 tires is less than 24,600 km?

---



---

The sampling distribution of the means has a mean of 25,000 km (the population mean)

$$\mu = 25000 \text{ km}$$

and a standard deviation (i.e.. standard error) of:

$$1600/8 = 200 \text{ km}$$

---

---

Convert 24,600 km to a z-score and use the normal table to determine the required probability.

$$z = (24600 - 25000) / 200 = -2$$

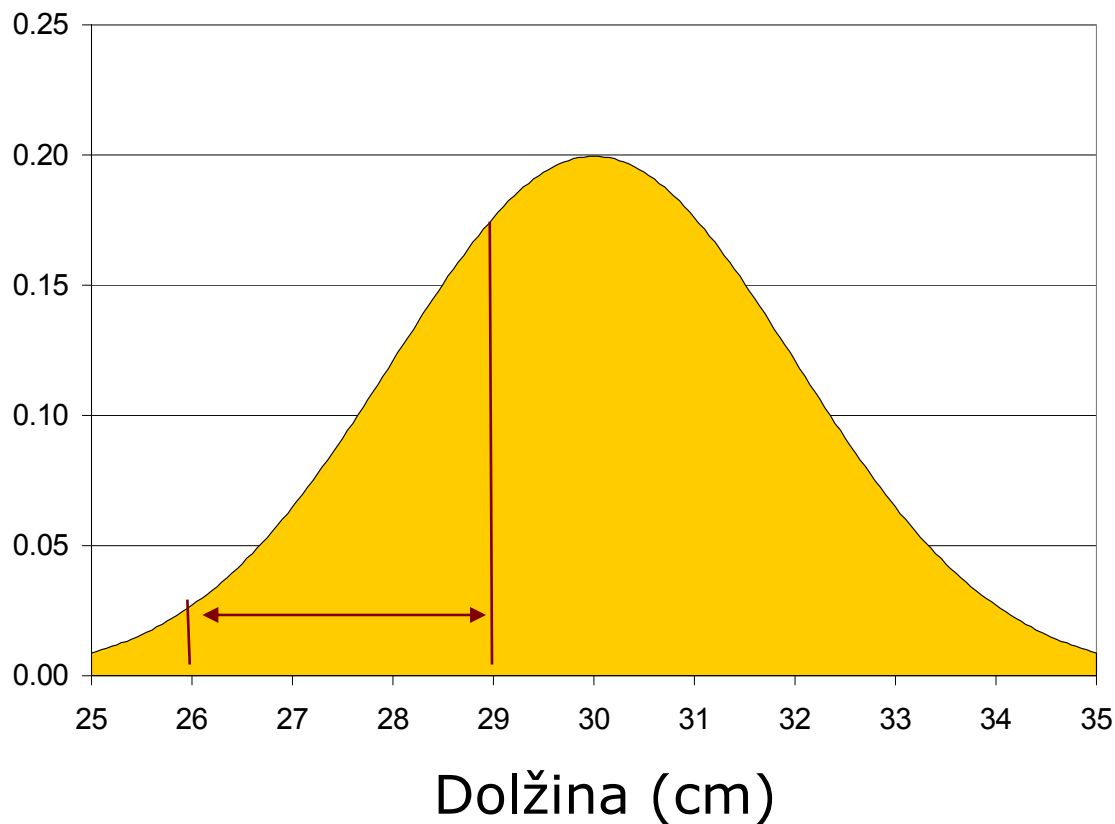
$$P(z < -2) = 0.0228$$

or 2.28% of the sample means will be less than 24,600 km.

---

# Normalna porazdelitev

- Kakšna je verjetnost, da bo ujeta postrv dolga med 26 in 29 cm?



# Ocenjevanje parametrov

---

- Točkovne ocene
  - Intervalne ocene
-

# Intervalne ocene za velike vzorce

---

- Naključen vzorec
  - Standardni odklon v populaciji je poznan, ali najmanj 30 enot v vzorcu.
-

# Interval zaupanja za aritmetično sredino populacije

---

$$\bar{X} - z \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z \frac{\sigma}{\sqrt{n}}$$

—

$\bar{X}$ : Vzorčna aritmetična sredina

$\sigma$ : standardni odklon

$n$ : velikost vzorca

---

# Statistical Inference: Estimation and Hypothesis Testing

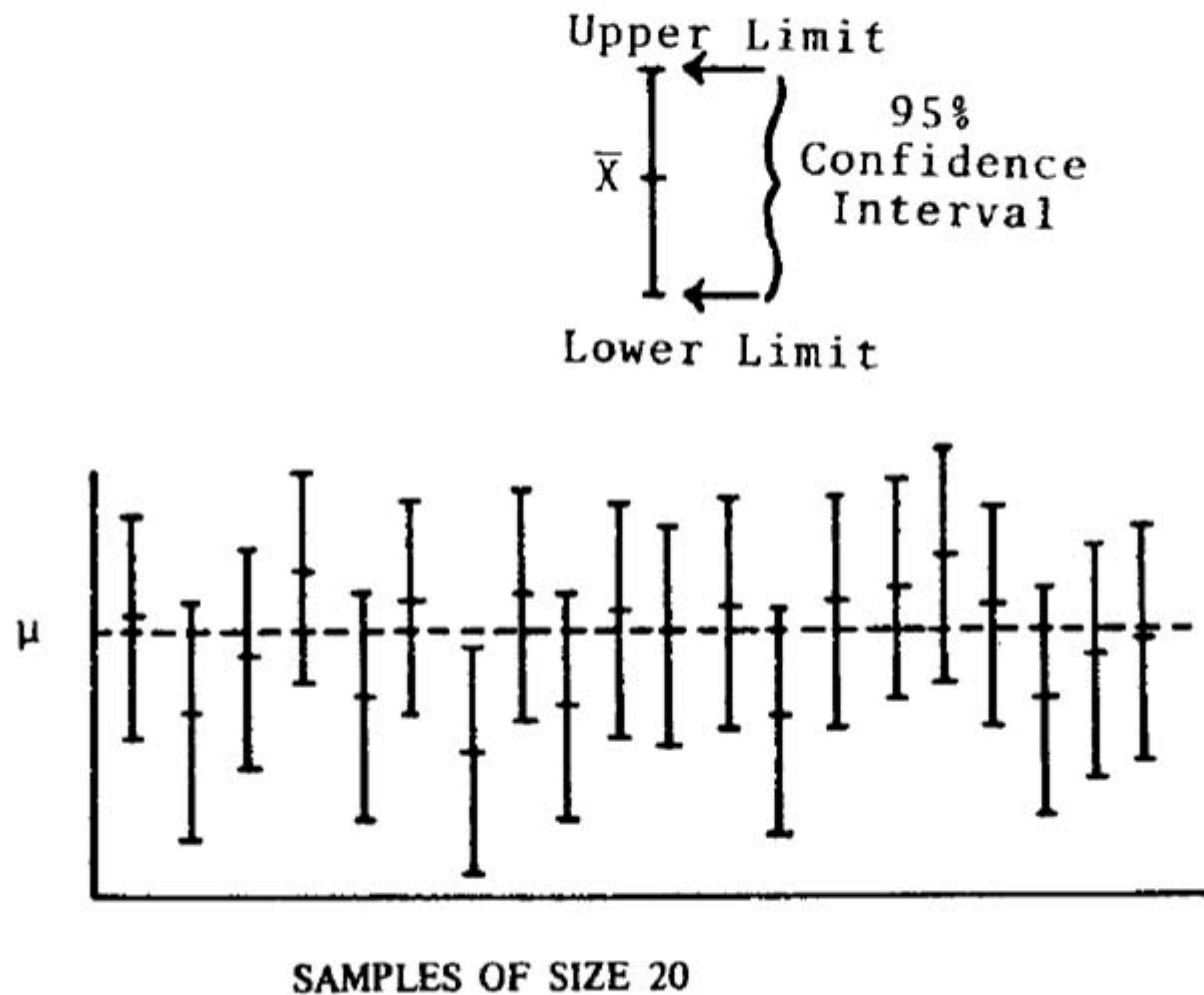
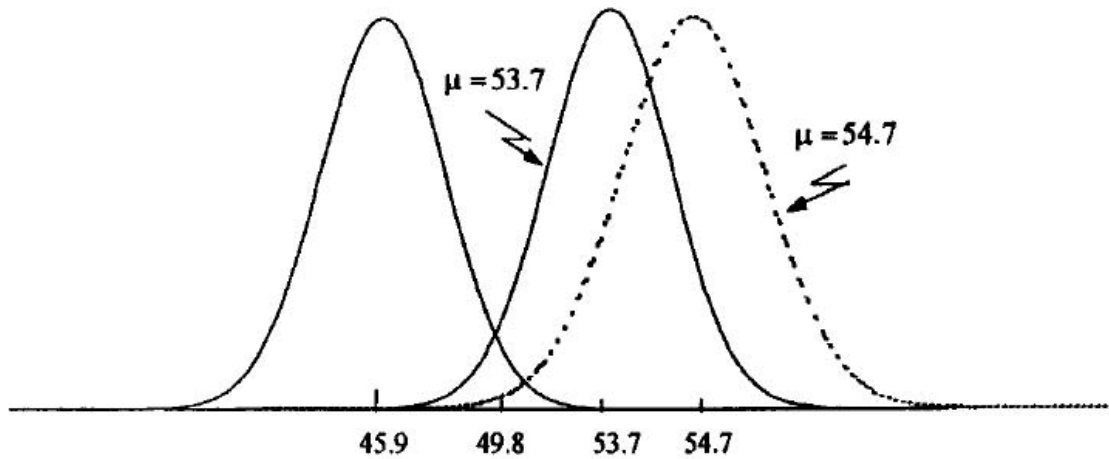


Figure 5.2 Concept of the confidence interval.



**Figure 5.3** This figure shows that a mean of 49.8 is unlikely to be observed if the true mean is 54.7 (confidence interval = 45.9 to 53.7).

$$\bar{x} = 49.8$$

$$N = 20$$

$$\sigma_{\bar{x}} = 2$$



# Primer

Estimate, with 95% confidence, the lifetime of nine volt batteries using a randomly selected sample where:

--

$\bar{X}$  = 49 hours

$s$  = 4 hours

$n$  = 36

---

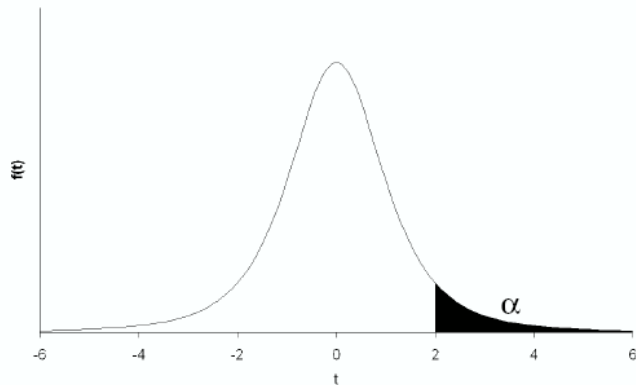
Lower Limit:  $49 - (1.96)(4/6)$   
 $49 - (1.3) = 47.7$  hrs

Upper Limit:  $49 + (1.96)(4/6)$   
 $49 + (1.3) = 50.3$  hrs

We are 95% confident that the mean lifetime of the population of batteries is between 47.7 and 50.3 hours.

---

# Majhni vzorci, standardni odklon v populaciji ni poznan



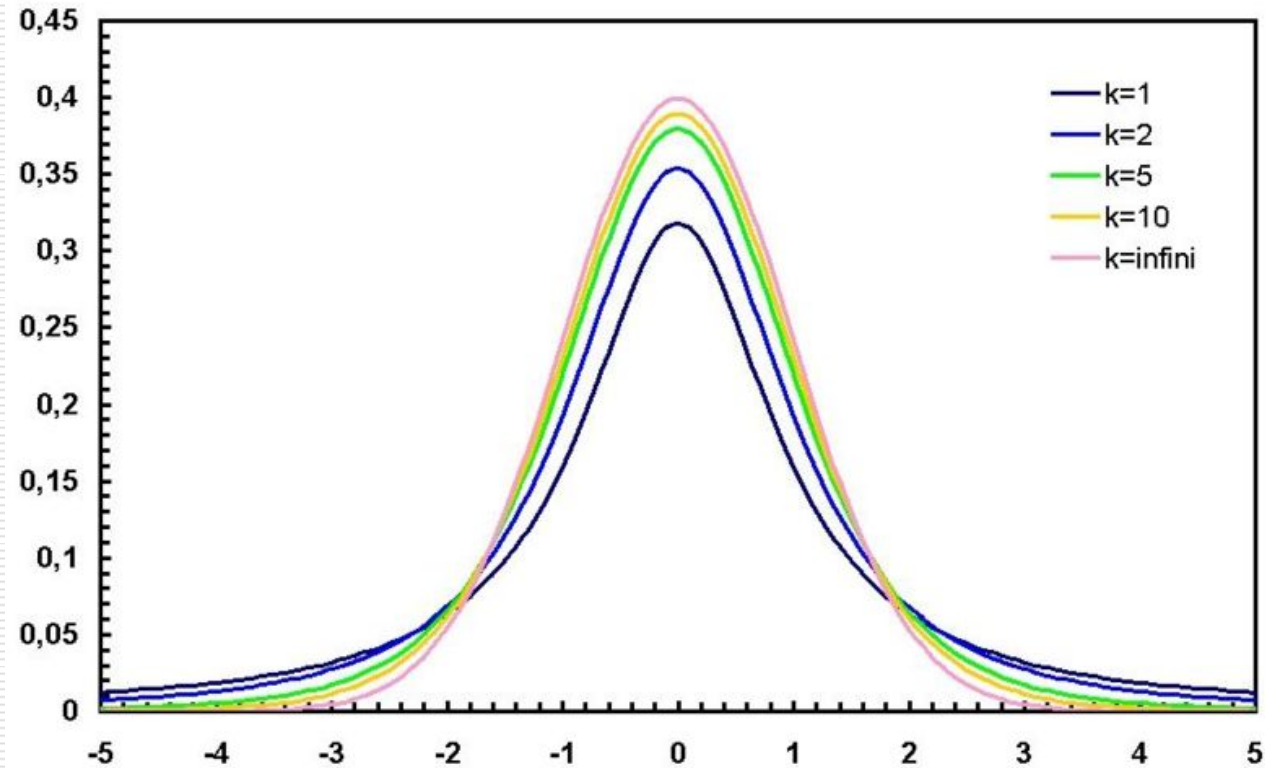
$$t = \frac{\bar{X} - \mu}{s / \sqrt{N}}$$

t

Število prostostnih stopenj	Stopnja tveganja ( $\alpha$ )				
	0.4	0.2	0.1	0.05	0.01
Dvostransko tveganje	0.4	0.2	0.1	0.05	0.01
Enostransko tveganje	0.2	0.1	0.05	0.025	0.005
1	1.376	3.078	6.314	12.706	63.656
2	1.061	1.886	2.920	4.303	9.925
3	0.978	1.638	2.353	3.182	5.841
4	0.941	1.533	2.132	2.776	4.604
5	0.920	1.476	2.015	2.571	4.032
6	0.906	1.440	1.943	2.447	3.707
7	0.896	1.415	1.895	2.365	3.499
8	0.889	1.397	1.860	2.306	3.355
9	0.883	1.383	1.833	2.262	3.250

# Studentova porazdelitev

---



William Seally Gosset, Guinness Brew., Dublin

---

# Primer

---

Pri kliničnem preskušanju novega postopka zdravljenja na 38 bolnikih so umrli trije bolniki.

Določite 90% interval zaupanja za delež umrlih bolnikov pri zdravljenju z novim postopkom!

Kakšna je verjetnost, da bi pri kliničnem preskušanju novega postopka zdravljenja na desetih bolnikih umrli natanko štirje?

---