Normalna distribucija

Varijabilnosti frekvencija distribucija srednjih vrijednosti uzoraka iz populacije normalne distribucije

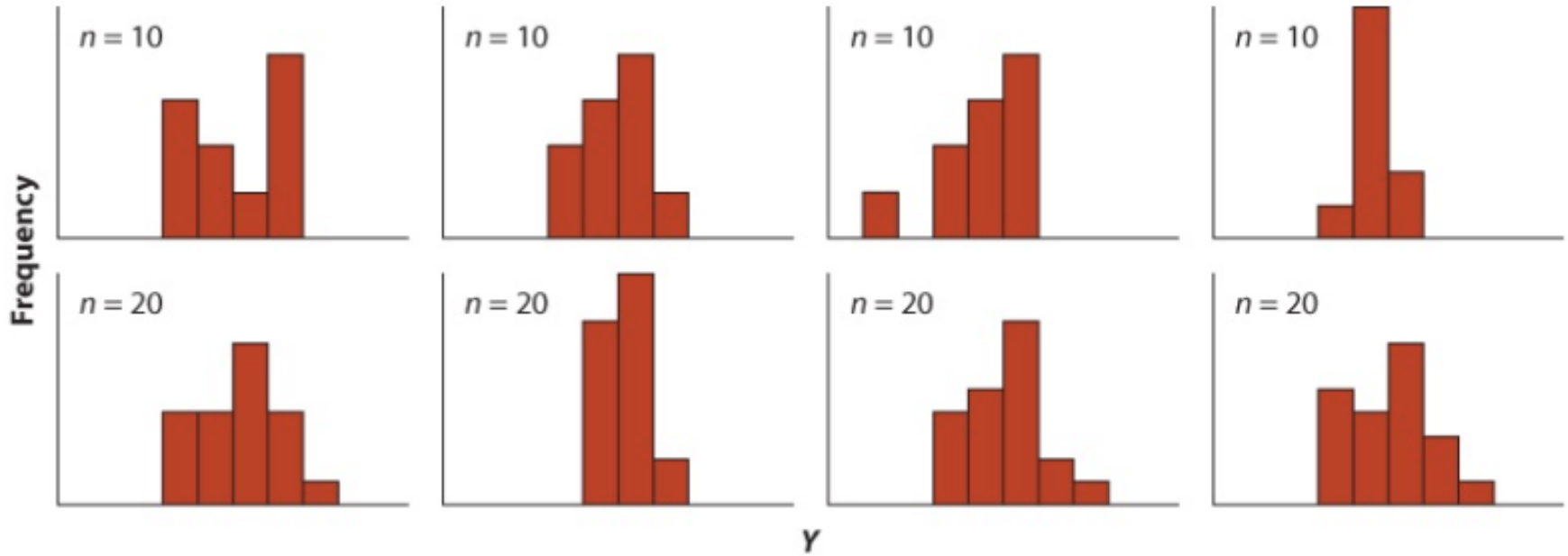


Figure 13.1-1
Whitlock et al., *The Analysis of Biological Data*, 2e, © 2015 W. H. Freeman and Company

FIGURE 13.1-1 *Top row*: Histograms of four random samples of size $n = 10$ from the same normal distribution. *Bottom row*: Histograms of four random samples of size $n = 20$ from the same normal distribution.

Normalna distribucija

Frekvencija distribucija uzoraka iz populacije koja nema normalnu distribuciju

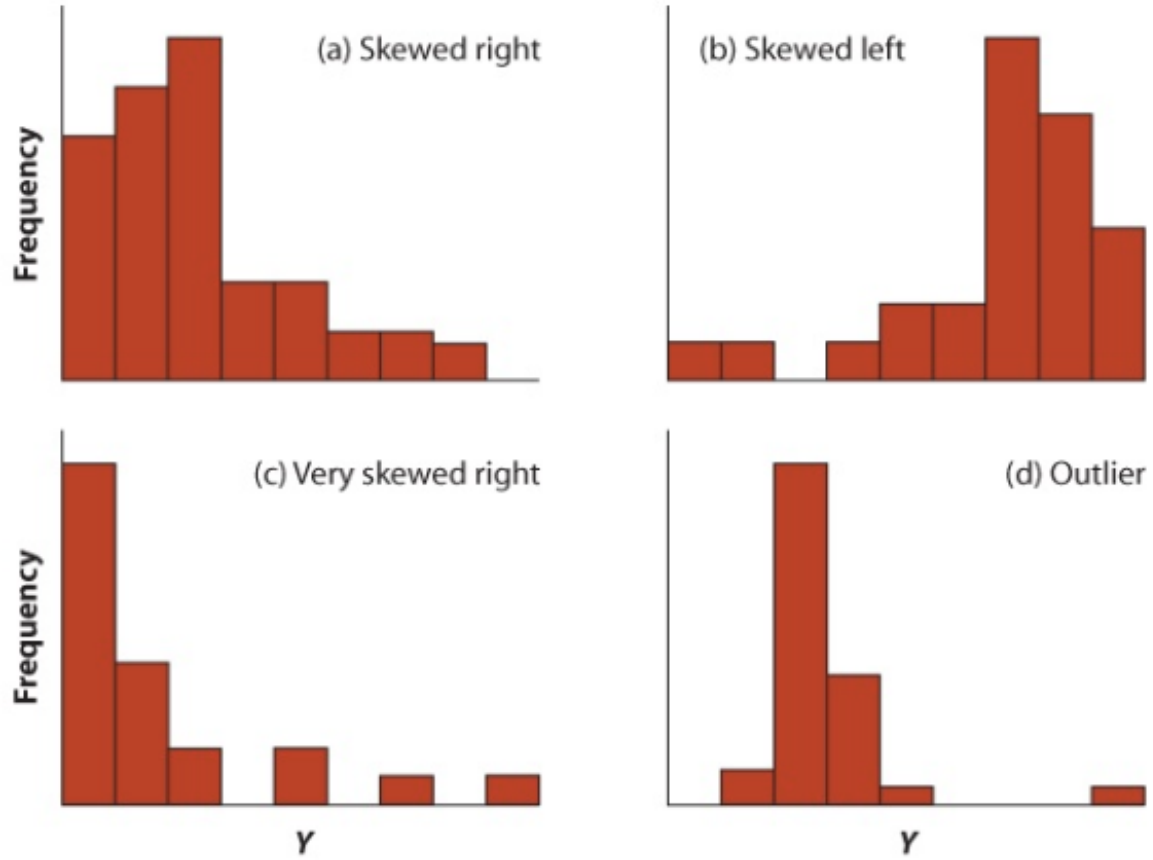
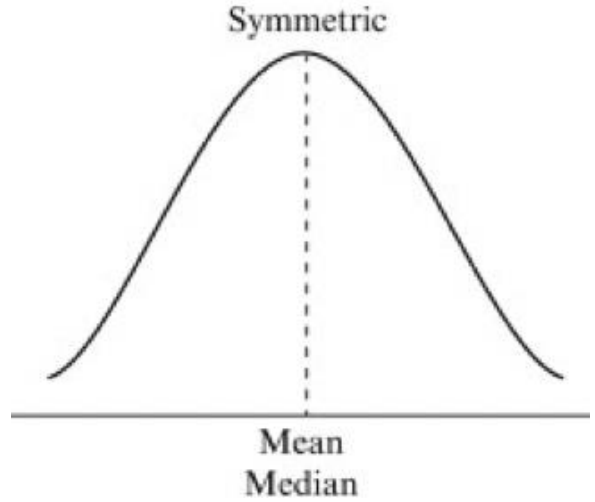


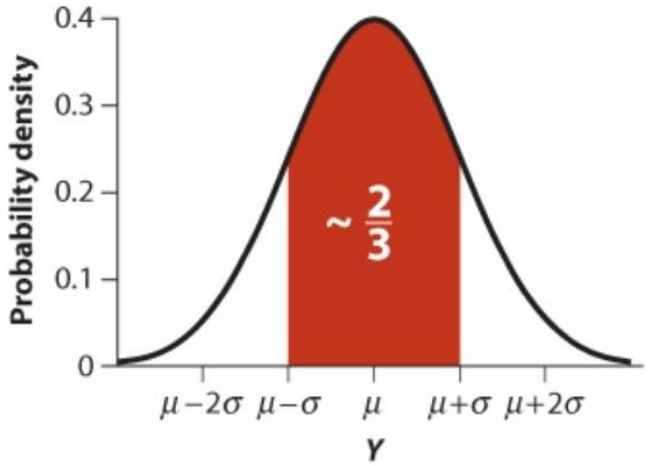
Figure 13.1-2
Whitlock et al., *The Analysis of Biological Data*, 2e, © 2015 W. H. Freeman and Company

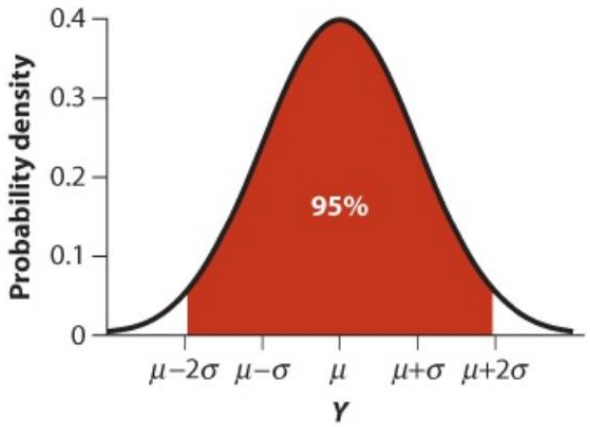
FIGURE 13.1-2 Frequency distributions of four data sets drawn from non-normal distributions. These distributions are (a) skewed right, (b) skewed left, (c) more extreme skew to the right, and (d) with an outlier.

Normalna distribucija

- **srednje vrijednosti, medijan, mode i aritmetička sredina u sredini**
- **simetrične standardne devijacije**
- otprilike 2/3 površine nalazi se unutar jedne standardne devijacije
- 95% površine nalaze se unutar 1.96 standardne devijacije

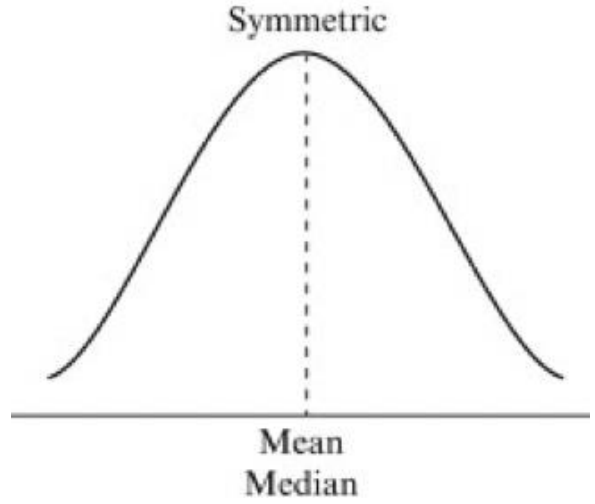


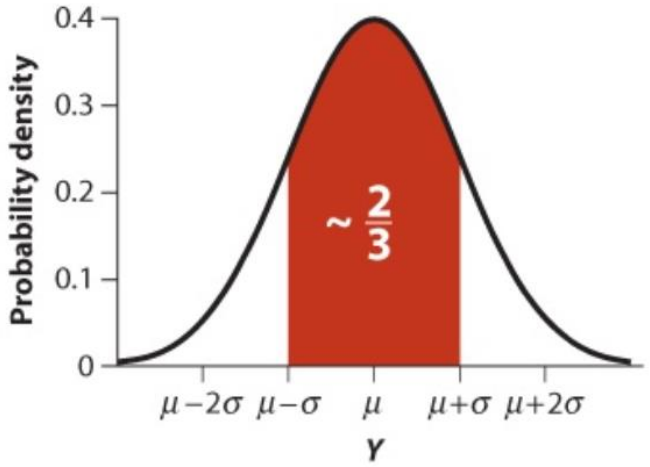


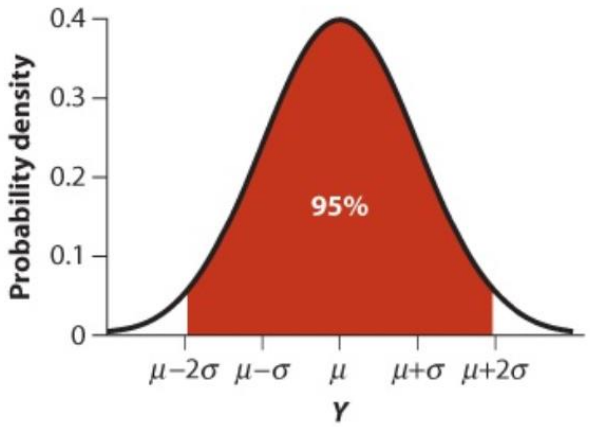


Normalna distribucija

- **srednje vrijednosti, medijan, mode i aritmetička sredina u sredini**
- **simetrične standardne devijacije**
- otprilike 2/3 površine nalazi se unutar jedne standardne devijacije
- 95% površine nalaze se unutar 1.96 standardne devijacije





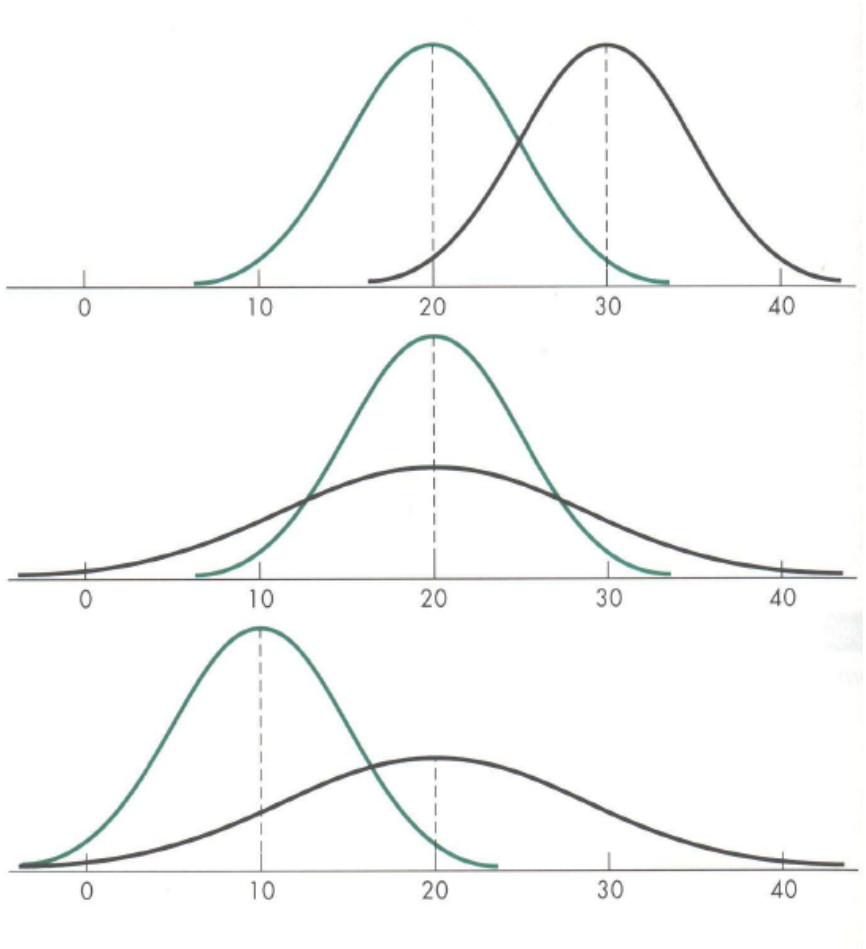


- **μ – „vrh” krivulje**
- **Med = Mod = μ**
- **Većina podataka je „oko” μ**
- **50% podataka je veće ili manje od μ**

Normalna distribucija

Standardna devijacija određuje širinu krivulje

- iste varijabilnosti različite sredine
- iste sredine različite varijabilnosti
- različite sredine različite varijabilnosti

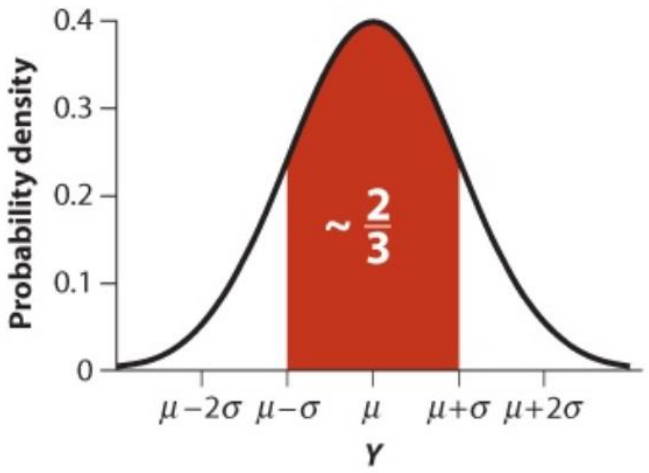


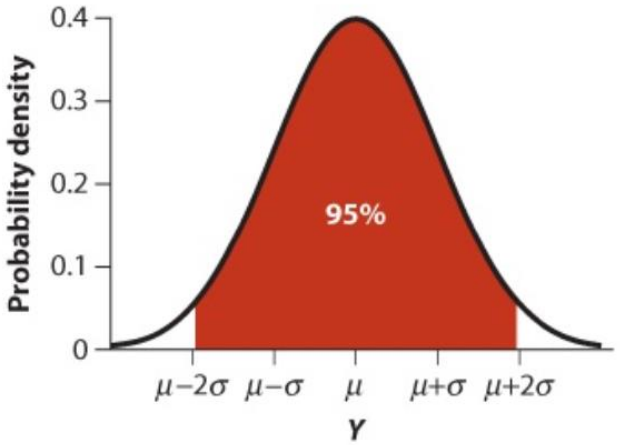
Normalna distribucija intervali pouzdanosti (CI)

- Raspon unutar određene raspodjele u kojem se očekuje zadana količina podataka
- Omeđen granicama (donja i gornja granica pouzdanosti)

- Vjerojatnost da će normalno distribuirana varijabla poprimiti vrijednost između μ–σ i μ+σ je 68.26%
- Vjerojatnost da će normalno distribuirana varijabla poprimiti vrijednost između μ–2σ i μ+2σ je 95.46%
- Vjerojatnost da će normalno distribuirana varijabla poprimiti vrijednost između μ–3σ i μ+3σ je 99.72%

- Granice pouzdanosti (confidence limits – CI; confidence interval)





Standardizirana normalna distribucija

Standardizirana normalna distribucija
μ=0 σ=1

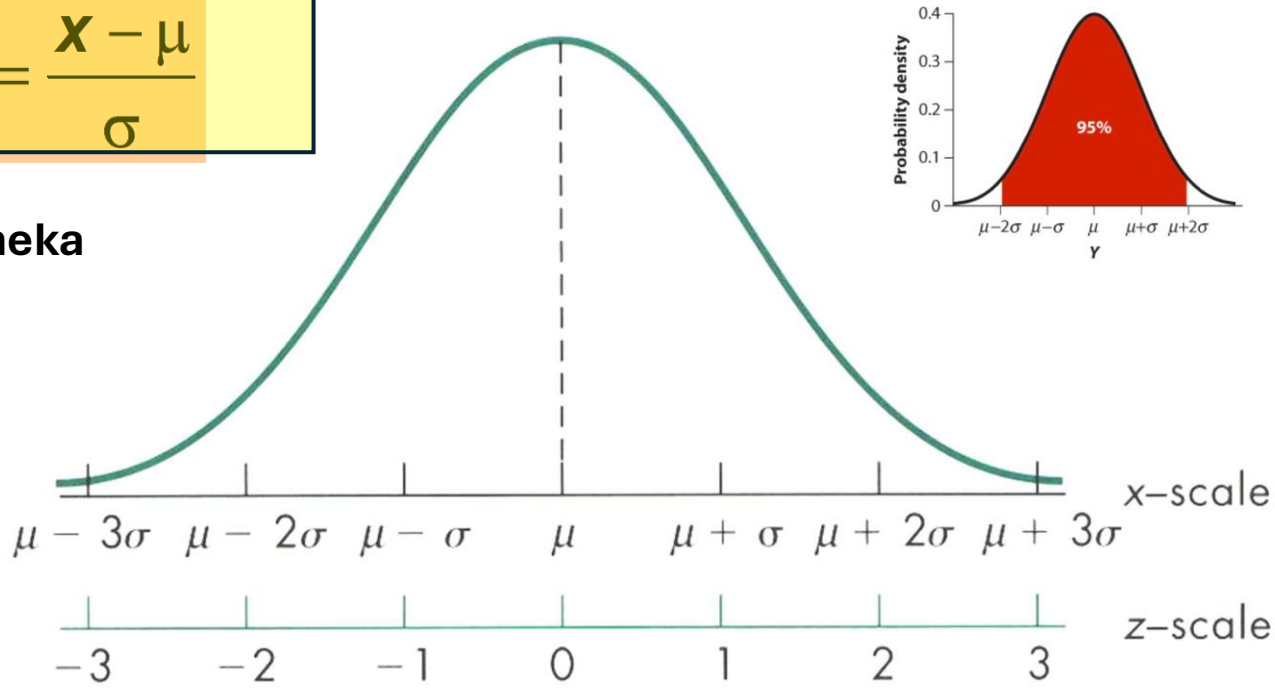
- "Vrh" svake normalne distribucije ("čvrsta točka") – μ
- "Širina" distribucije ("korak") - σ
- Svaku vrijednost varijable (svaku točku na osi x) možemo prikazati kao udaljenost od aritmetičke sredine izraženu u standardnim devijacijama

$$x=\mu \pm z*\sigma$$

$\mu = \mu \pm 0*\sigma$ – za srednju vrijednost Z=0

$$z = \frac{x - \mu}{\sigma}$$

- **z vrijednost – za koliko standardnih devijacija (σ) je neka vrijednost (x) udaljena (različita) od sredine (μ)**
- **negativan z – x je manji od μ**
- **pozitivan z – x je veći od μ**



- *99% podataka ima z vrijednost = 2.58 ili manju*

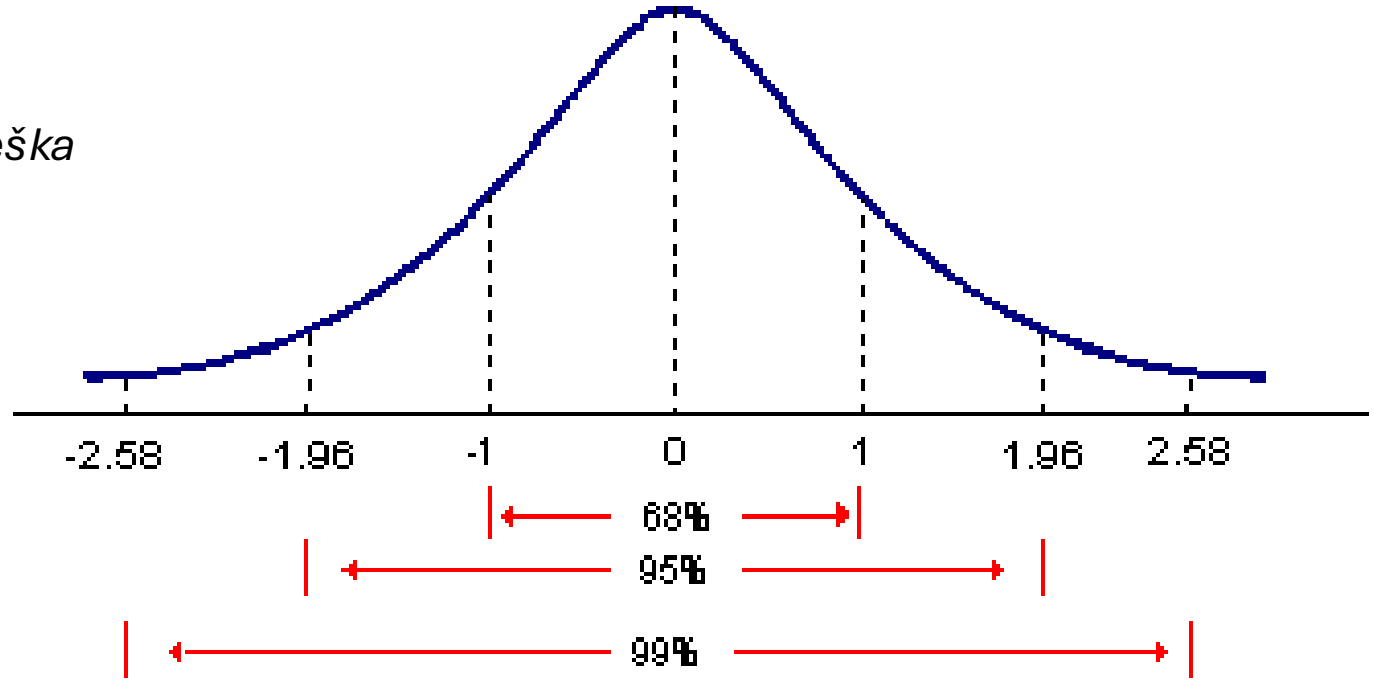
Z tablice – Površine ispod normalne krivulje od x do bližeg kraja krivulje

Normalna distribucija intervali pouzdanosti (CI)

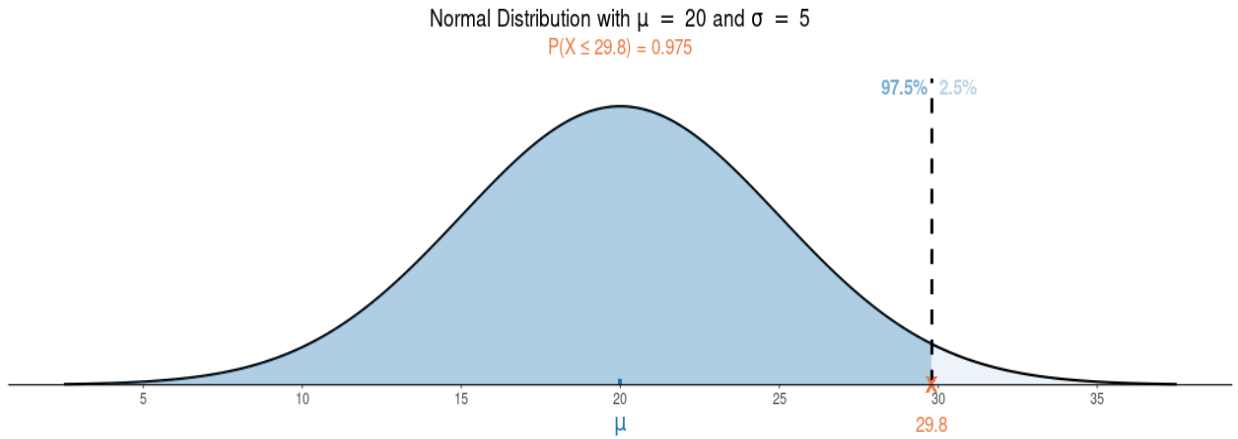
- Statističko zaključivanje – intervali pouzdanosti od 95% i 99%
  - 95% - između μ–1.96σ i μ+1.96σ
    - *u 19 od 20 slučajeva 95% CI sadrži populacijski parametar*
  - 99% - između μ–2.58σ i μ+2.58σ
  - 100% - ????

$$CI = \bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

Standardna greška



Normalna distribucija intervali pouzdanosti (CI)



Duljina ribe

n=365
x=20cm
s=5 cm

Vjerojatnost da se odabere riba dulji od 20 cm?

50%

Vjerojatnost da se odabere riba dulja od 29.8cm?

*29.8=20+**1.96***5 - vrijednost je 1.96 σ veća od μ*

x=μ ±z∗σ

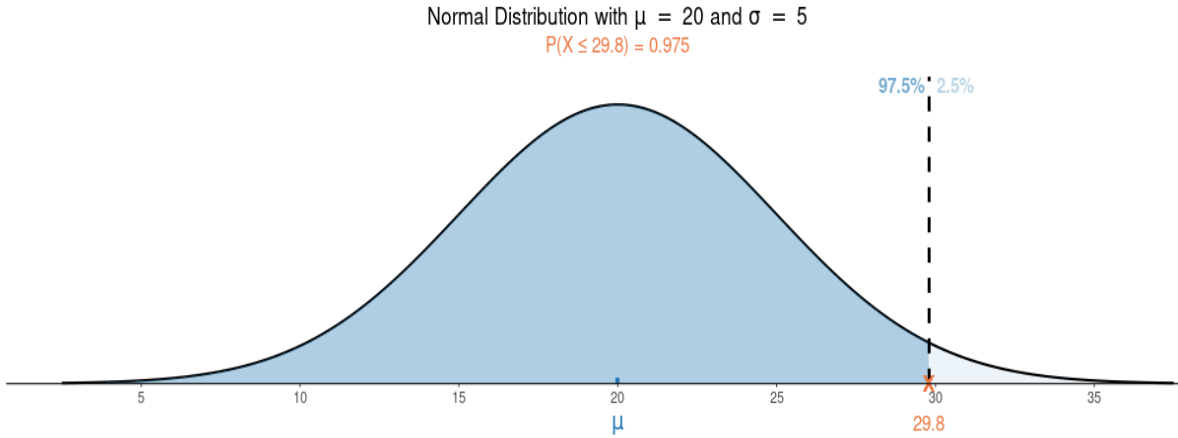
p=2.5%

365 x 0.025=9.125 – toliko je riba veće od 29.8 cm

Normalna distribucija intervali pouzdanosti (CI)

Duljina ribe

n=365
x=20cm
s=5 cm



365 x *0.025*=9.125 – toliko je riba veće od 29.8 cm

$$SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad SE = \frac{5}{\sqrt{365}} \approx 0.262 \text{ standarna greška}$$

$$CI = \bar{x} \pm 1.96\frac{\sigma}{\sqrt{n}} \qquad 1.96 \times 0.262 \approx 0.514 \text{ cm} \qquad CI = 20 \pm 0.514 \qquad \boxed{CI = [19.486, 20.514] \text{ cm}}$$

S 95% pouzdanosti procjenjujemo da se prava populacijska sredina nalazi između 19,49 i 20,51 cm
– *ako* bismo mnogo puta uzorkovali po 365 riba i svaki put izračunali interval, približno 95% tih intervala obuhvatilo bi pravu populacijsku sredinu.
Interval pouzdanosti procjenjuje sredinu populacije, a ne raspon duljina pojedinačnih riba!

Becoming a scientist...

https://www.youtube.com/shorts/UKc5Np8H3Hk