

TIPOVI ISTRAŽIVANJA U BIOLOGIJI

EKSPERIMENT – istraživač kontrolira/određuje vrijednosti nezavisnih varijabli nasumično za jedinke u istraživanju (najčešće laboratorijska istraživanja)

- Npr. klinička ispitivanja, testiranja na laboratorijskim životinjama, testna polja s klijancima...

OPAŽANJE – istraživač nema kontrolu nad tim koji subjekti istraživanja su u kojoj grupi

- *Npr. istraživanja prirodne populacije šišmiša u različitim staništima, usporedbe zdravstvenih aspekata pušača i nepušača*

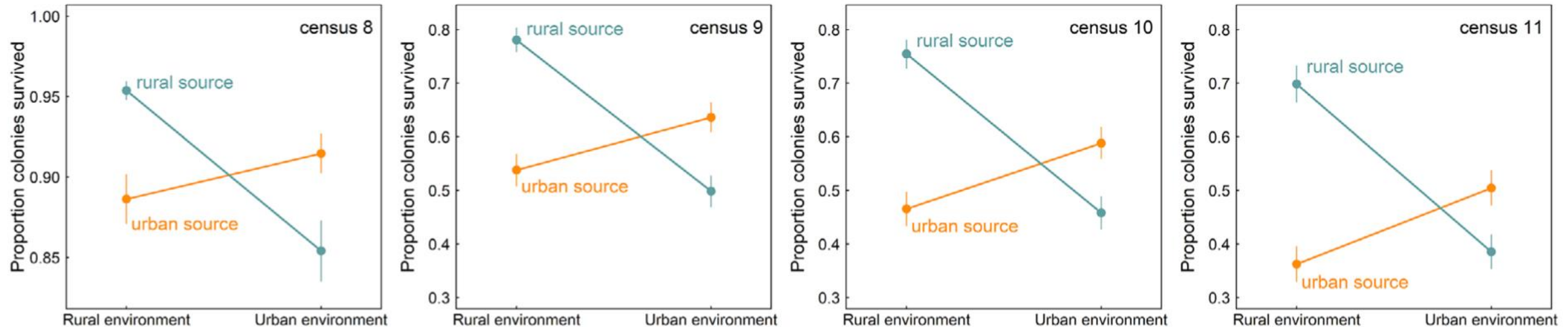
Eksperimentom možemo utvrditi uzročno-posljedičnu vezu, opažanjima samo povezanost među varijablama.

Eksperimenti su češće laboratorijska istraživanja, opažanja češće na prirodnim populacijama, ali to nije pravilo... mogu se vršiti i eksperimenti s jedinkama u prirodi, mesokozmosu, transplant eksperimenti etc...

In a nutshell, a reciprocal transplant experiment reveals local adaptation and fitness trade-offs in response to urban evolution in an acorn-dwelling ant

[Ryan A. Martin](#),¹ [Lacy D. Chick](#),^{1,2} [Matthew L. Garvin](#),^{1,3} and [Sarah E. Diamond](#)¹

► [Author information](#) ► [Article notes](#) ► [Copyright and License information](#) [PMC Disclaimer](#)



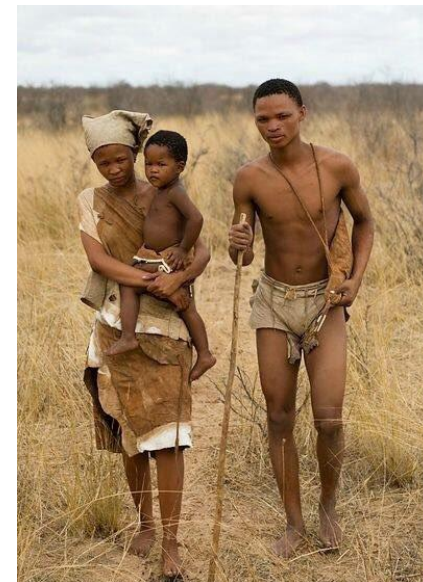
preživljanje - tolerancija niskih i visokih temperatura – evolucijski trade-off!

UZORAK I POPULACIJA

POPULACIJA – **sve jedinice** (na svim razinama biološke organizacije- molekula, stanica, jedinka, populacija) **koje su interes istraživanja**

Populacija u biostatističkom smislu nije isto kao populacija neke vrste u biološkom smislu!

- *Sposobnost graha tetovca za vezivanje dušika iz tla – sve biljke te sorte*
- *Fagocitna aktivnost limfocita u pacijenata koji su preboljeli mononukleozu – svi limfociti kod svih pacijenata koji su preboljeli mononukleozu*
- *Boja kože u narodu KhoeSan u Africi – sve jedinice u toj populaciji*
- *Bazalna ekspresija CyP1A2 gena u stanici jetre zdravog muškarca – svi zdravi muškarci (i sve njihove stanice jetre)*
- *Brzina leta šišmiša kasni noćnjak Eptesicus serotinus– sve jedinice te vrste*



UZORAK I POPULACIJA

UZORAK – stvarni broj jedinica (na svim razinama biološke organizacije - molekula, stanica, jedinka, populacija) **koje su predmet istraživanja**

Uzorak je uvijek puno manji od populacije!

UZORAK – jedinice mogu biti i grupe (npr. kolonija mikroba, obitelj, akvarij s ribama, kavez s miševima, polje raži, cret..)

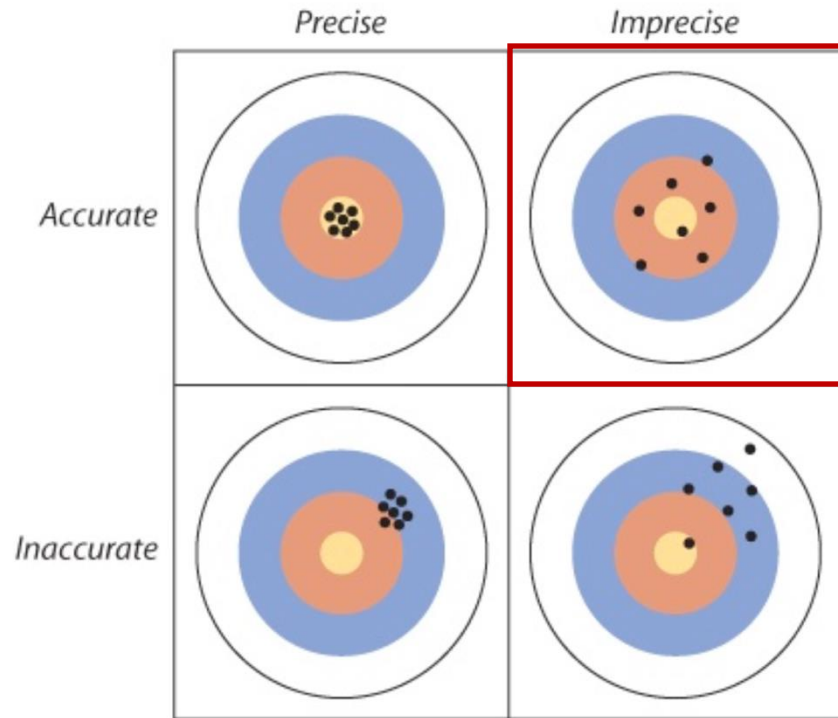
- *Sposobnost graha tetovca za vezivanje dušika iz tla – 100 biljaka te sorte sakupljenih na 10 lokacija (10 „uzoraka“)*
- *Fagocitna aktivnost limfocita u pacijenata koji su preboljeli mononukleozu – po 100 limfocita 300 ženskih i 300 muških pacijenata koji su preboljeli mononukleozu (podjednako zastupljenih ljudi zapadno europskog, azijskog i afričkog podrijetla)*
- *Boja kože u narodu KhoeSan u Africi – 40 muških i 40 ženskih jedinki u toj populaciji*
- *Bazalna ekspresija CyP1A2 gena u stanici jetre zdravog muškarca – 300 muškaraca koji su preboljeli mononukleozu (podjednako zastupljenih ljudi zapadno europskog, azijskog i afričkog podrijetla)*
- *Brzina leta šišmiša kasni noćnjak Eptesicus serotinus u zapadnoj Hrvatskoj– 20 jedinki oba spola na 4 lokacije*



UZORAK I POPULACIJA

Svojstva koja mjerimo ili promatramo na uzorku uvijek se malo razlikuju od svojstva cijele populacije (npr. visina studenata biologije na PMF-u na uzorku od 30 ili svih studenata biologije na PMF-u) – **greška uzorkovanja** (*sampling error*). Greška uzorkovanja dešava se zbog **slučajnih** odstupanja.

Greška uzorkovanja dovodi do **nepreciznosti** (koja se smanjuje što je uzorak i uzorkovanje veće)



Greška uzorkovanja dovodi do **nepreciznosti** (koja se smanjuje što su uzorak i uzorkovanje veći)

Figure 1.2-2

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

<https://www.zoology.ubc.ca/~whitlock/Kingfisher/SamplingNormal.htm>

SAMPLE 1 INDIVIDUAL

COMPLETE SAMPLE OF 10

CALCULATE MEAN

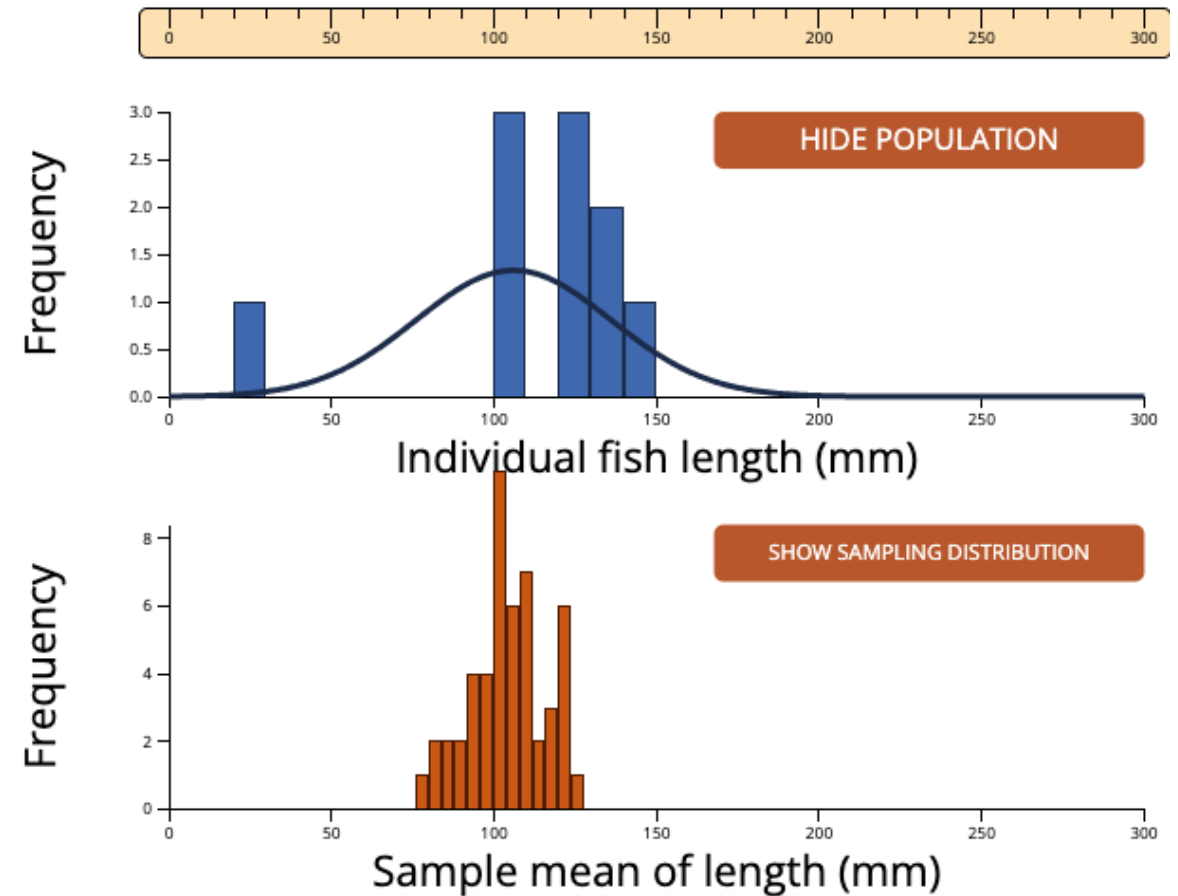
MEANS FOR MANY SAMPLES

n 10 μ 106 σ 30



Sampling from a normally distributed population

TUTORIAL



UZORAK I POPULACIJA

Greška pristranosti (“*bias*”) se događa zbog **nereprezentativnih** uzoraka (npr. mjerimo srdele ulovljene mrežama s velikim očima i na osnovu njih procijenjujemo veličinu jedinki u populaciji). Greška pristranosti dovodi do **netočnosti**, te **izmjereni parametri srednjih vrijednosti značajno odstupaju od stvarne srednje vrijednosti populacije – NETOČNOST MJERENJA**

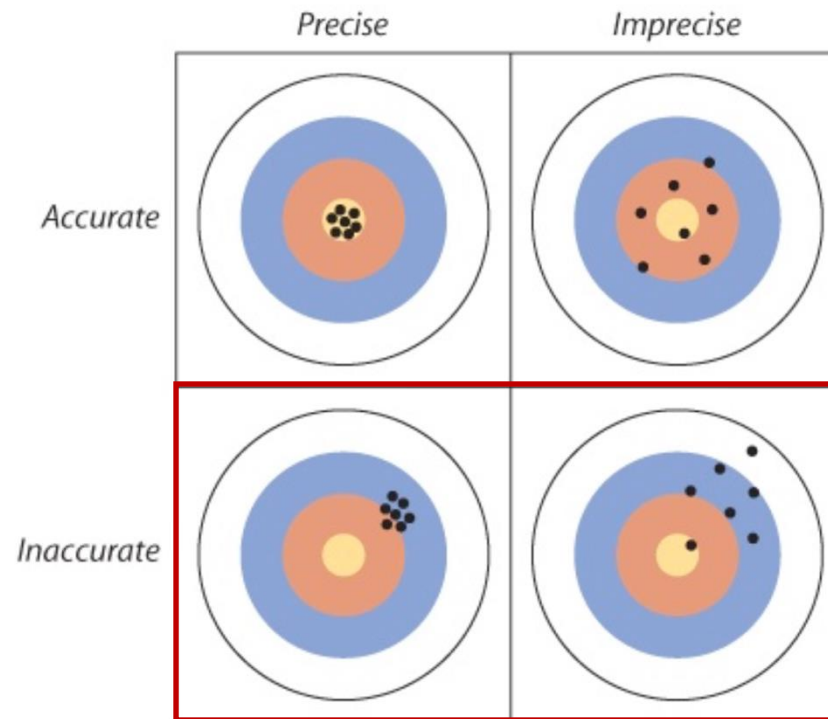


Figure 1.2-2

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

GREŠKE UZORKOVANJA

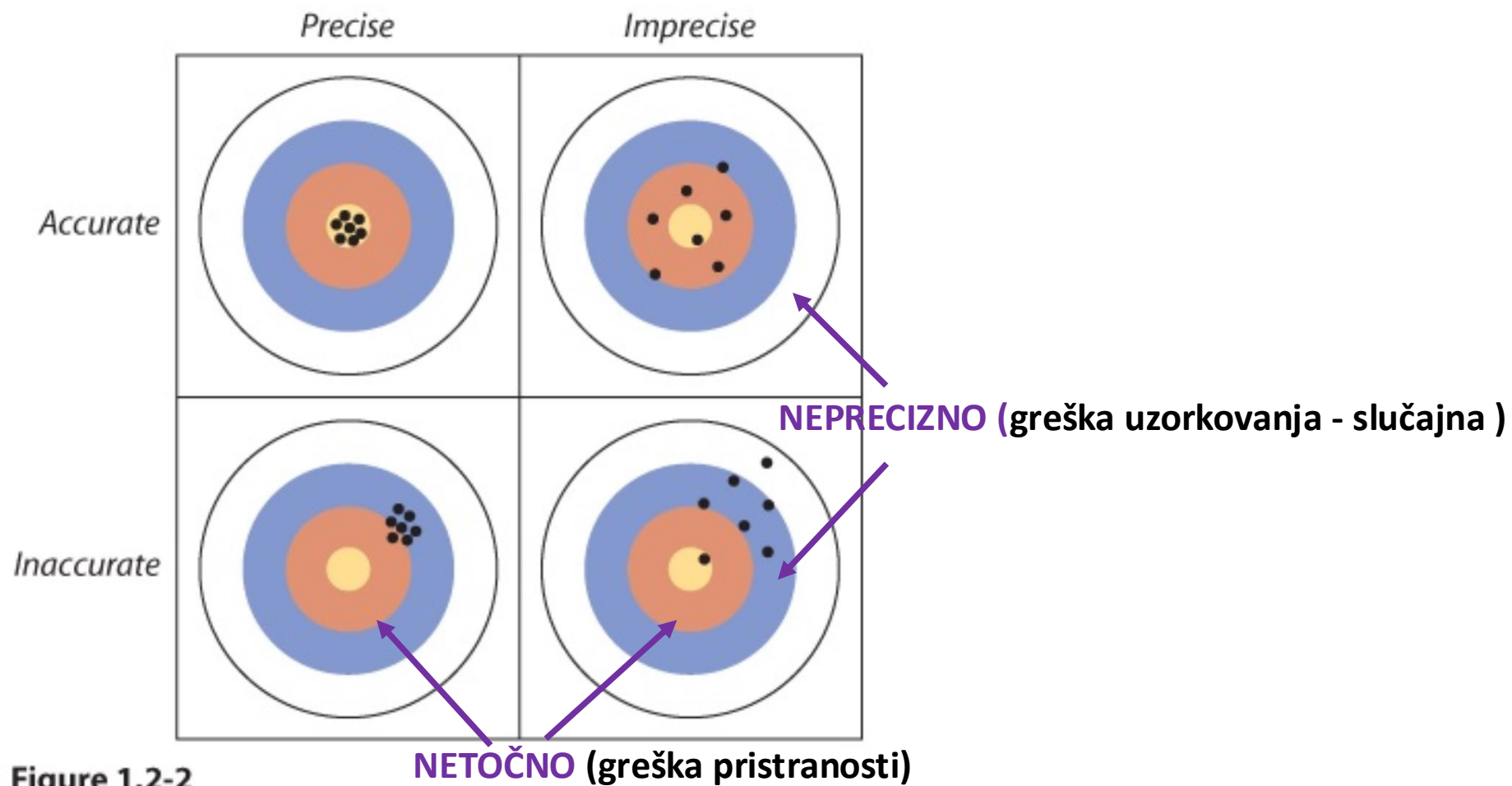


Figure 1.2-2

Whitlock et al., *The Analysis of Biological Data*, 2e, © 2015

W. H. Freeman and Company

UZORAK I POPULACIJA

UZORAK – u biološkim istraživanjima vrlo je važno da je kontrolni uzorak reprezentativan!

- *npr. gledamo utjecaj onečišćenja u lukama na fagocitoznu aktivnost hemocita dagnji. Iznimno je važno da **kontrolne postaje** imaju zanemarivu količinu toksikanata.*
- *npr. mjerimo gensku ekspresiju peroksidaza u staničnoj liniji štakora Štaky33 nakon izlaganja citostatiku. Ukoliko će kontrolne stanice imati onečišćenje rezultati neće biti reprezentativni.*

UZORAK – treba biti „slučajan“ (*random*) i neovisan!

Slučajan uzorak znači da svaka jedinka u grupi ima istu vjerojatnost biti uključena u uzorak

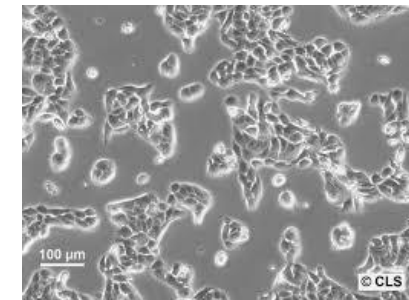
- *npr. srdele i oko mreže, ili dobrovoljci u istraživanju lijeka koji mogu biti motivirani slabijim socijalnim statusom – NIJE SLUČAJAN UZORAK*

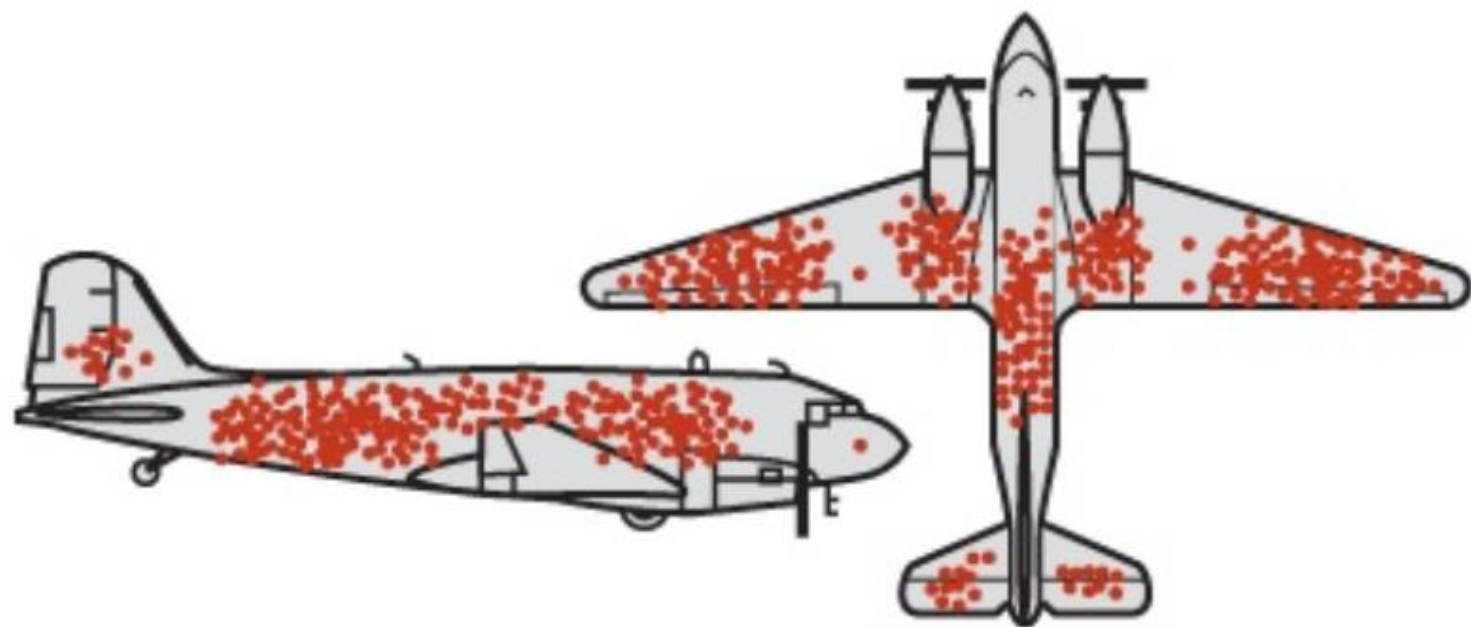
Neovisan uzorak podrazumijeva da su sve jedinke uzorkovane neovisno

- *npr. prilikom uzorkovanja žireva za analizu tanina uzeti su svi plodovi ispod istog drveta – NIJE NEOVISNO*
- *npr. za analizu učestalosti sindroma mladog goluba u Zagrebu uzrokovani su mladi iz tri gnijezda – NIJE NEOVISNO*

Često se jedinke u istraživanju randomizirano (nasumice) dijele u grupe, kako bi se smanjila pristranost i netočnost istraživanja.

SLUČAJAN I NEOVISAN UZORAK SMANJUJE GREŠKU PRISTRANOSTI (POVEĆAVA TOČNOST)

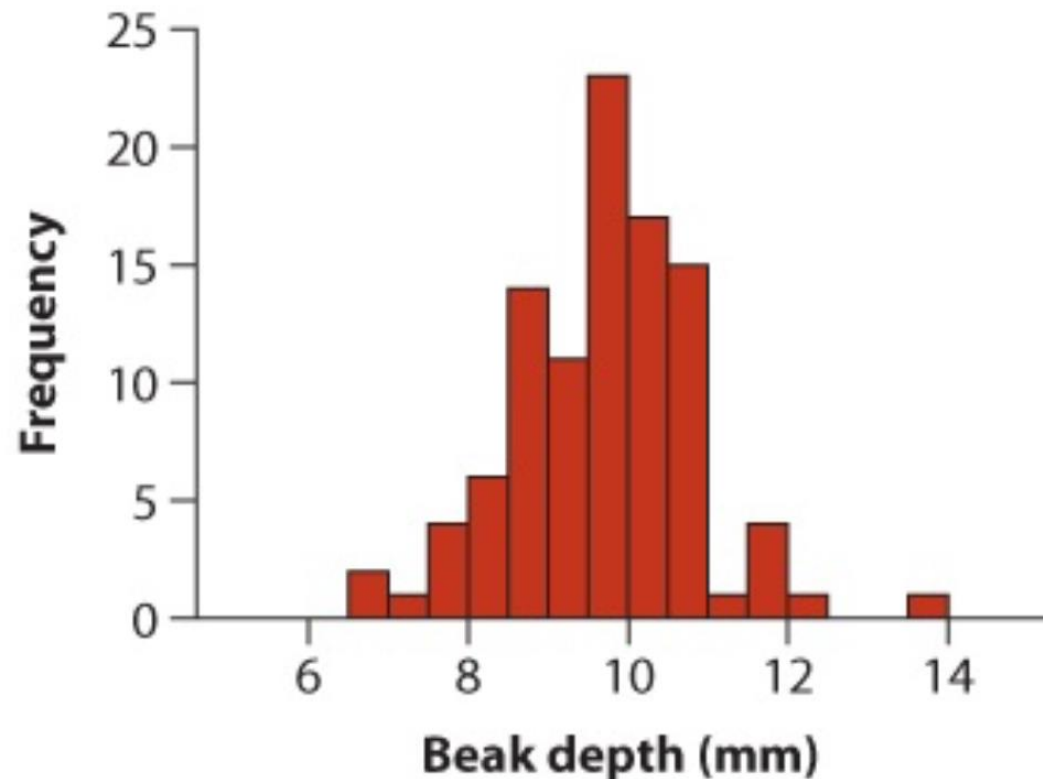




Whitlock et al., *The Analysis of Biological Data*, 2e, © 2015
W. H. Freeman and Company

Distribucije

Distribucija frekvencija (učestalosti, pojavnosti) opisuje koliko se puta vrijednost varijable pojavljuje u uzorku.



Distribucija frekvencija duljine kljuna u uzorku od 100 zebica *Geospiza* sp. na Galapagosu (interval 0.5 mm)
(Boag i Grant 1984, Biol. J. Linn. Soc)

Figure 1.4-1

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

Distribucije

Distribucija vjerojatnosti (*probability distribution*) opisuje teoretsku razdiobu podataka u populaciji.

Najčešća distribucija podataka u biološkim sustavima je **normalna distribucija**.

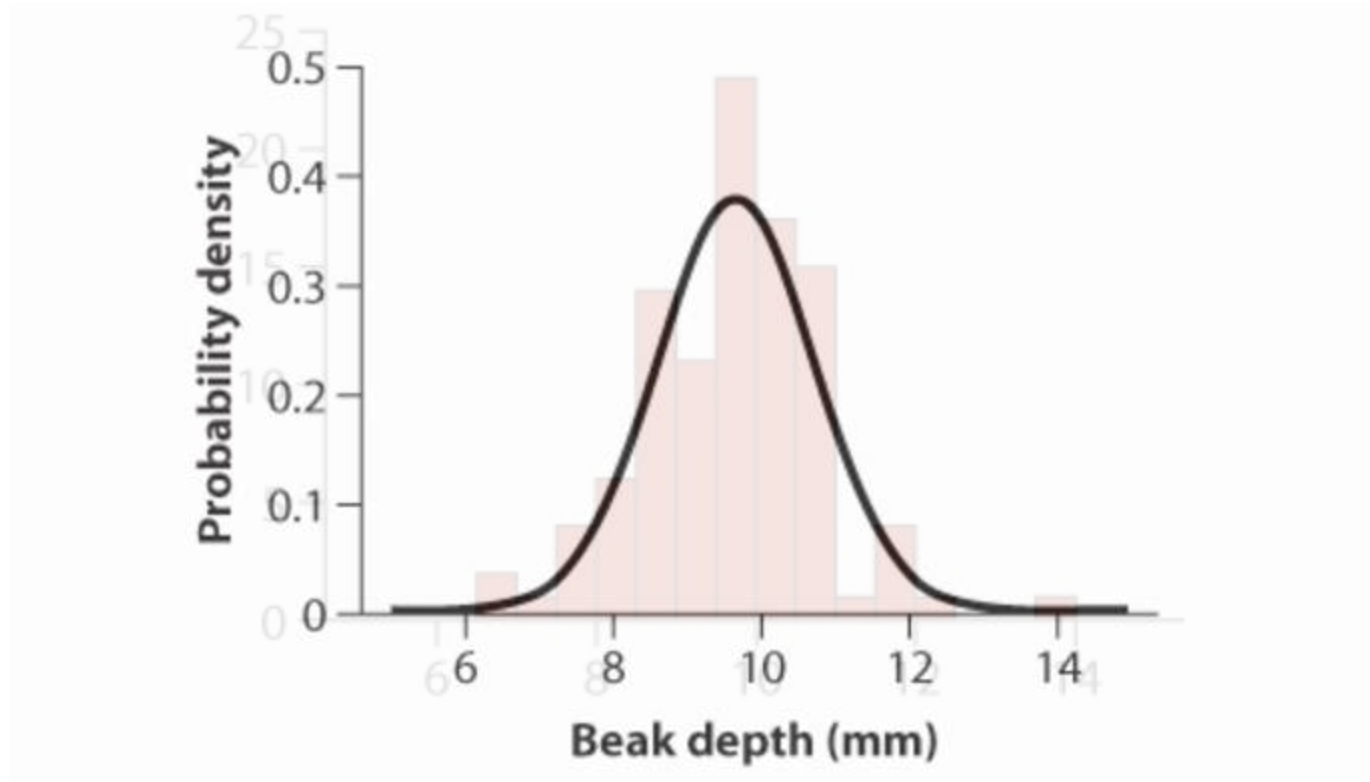


Figure 1.4-2

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

teško je imati uvid u stvarnu razdiobu podataka za cijelu populaciju!

Distribucije - normalna distribucija

Distribucija vjerojatnosti (*probability distribution*) opisuje teoretsku razdiobu podataka u populaciji.

Najčešća distribucija kontinuiranih podataka u biološkim sustavima je **normalna distribucija**.

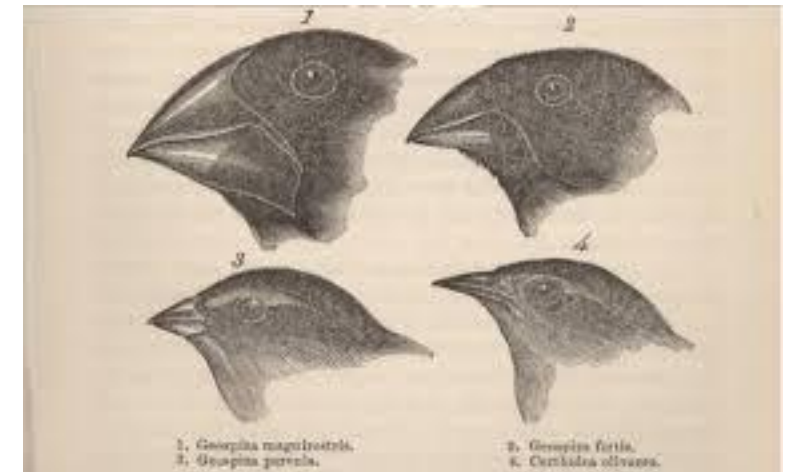
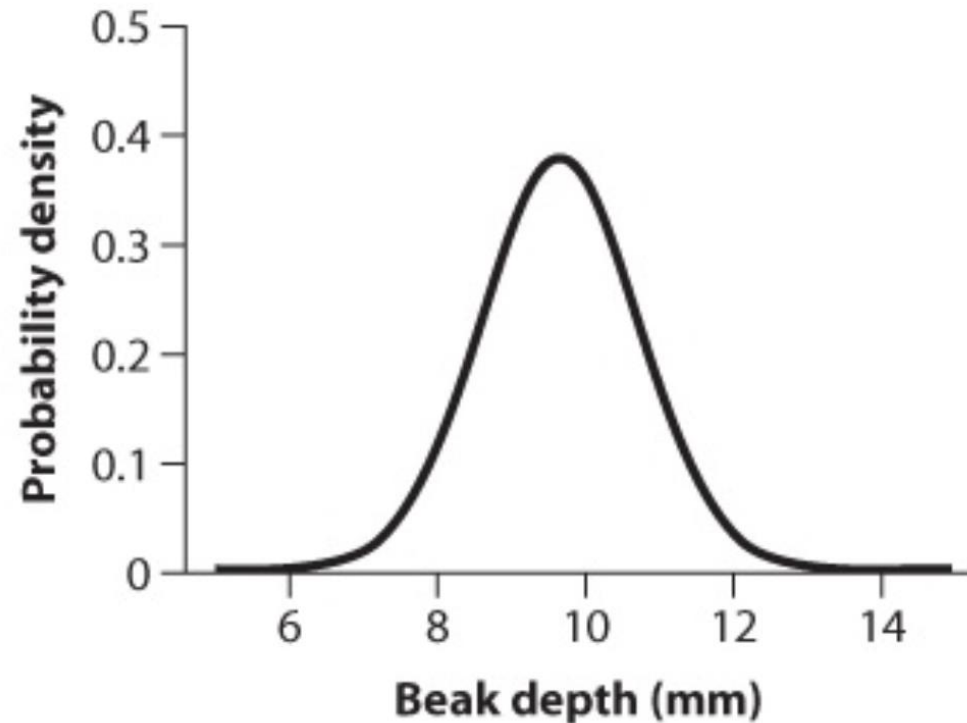


Figure 1.4-2

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

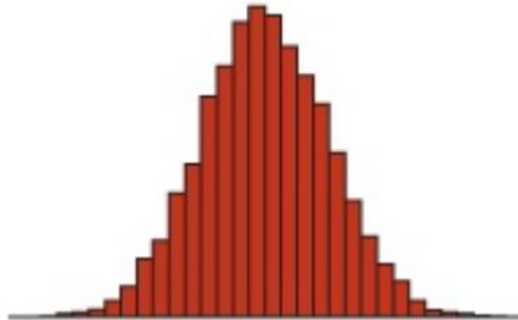
teško je imati uvid u stvarnu razdiobu podataka za cijelu populaciju

Distribucije

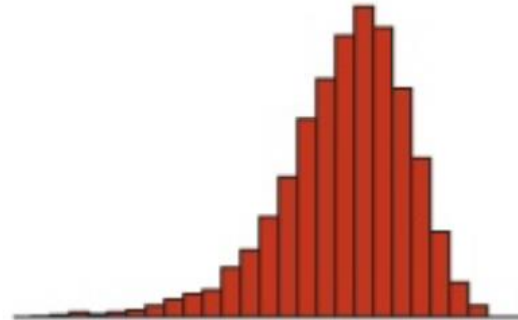
UNIFORMNA



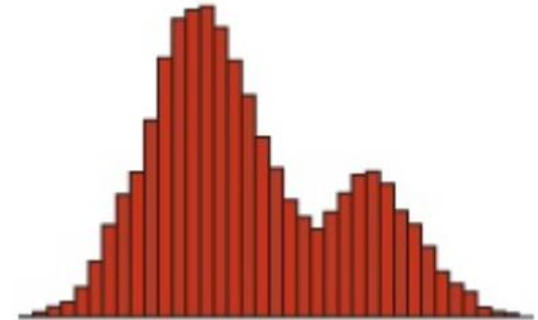
NORMALNA (z)



ASIMETRIČNA (F)



BIMODALNA



https://gallery.shinyapps.io/CLT_mean/

Figure 2.2-4

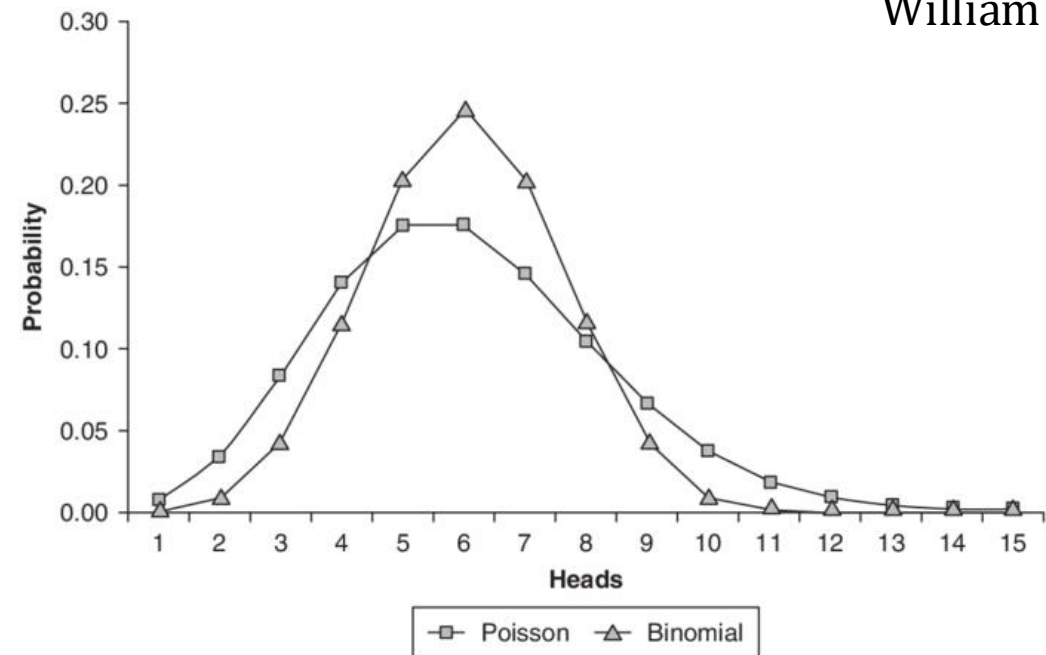
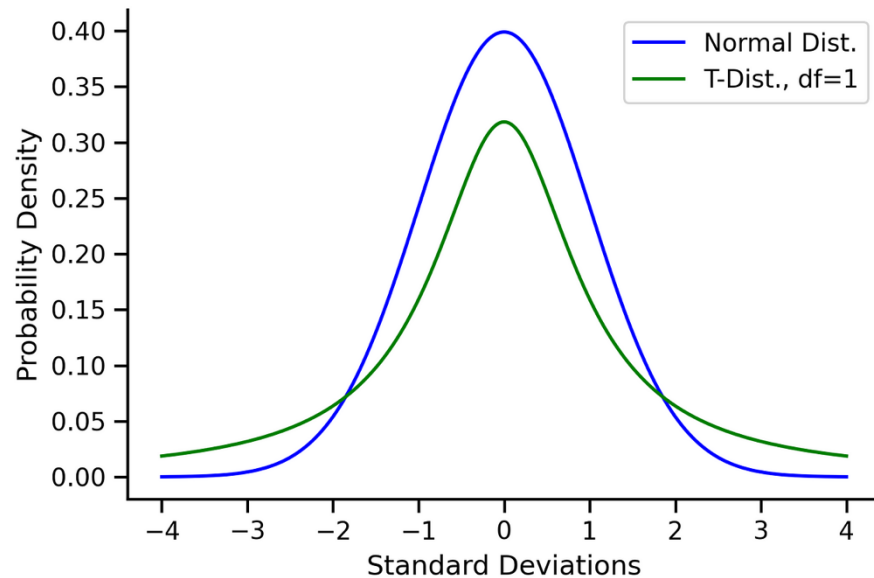
Whitlock et al., *The Analysis of Biological Data*, 2e, © 2015 W. H. Freeman and Company

Distribucije

- velik broj specifičnih distribucija
- kontinuirani podaci: normalna, lognormal, gamma, Weibull, t (Student)...
- diskretni podaci: binomijalna (proporcije), Poissonova, negativna binomijalna

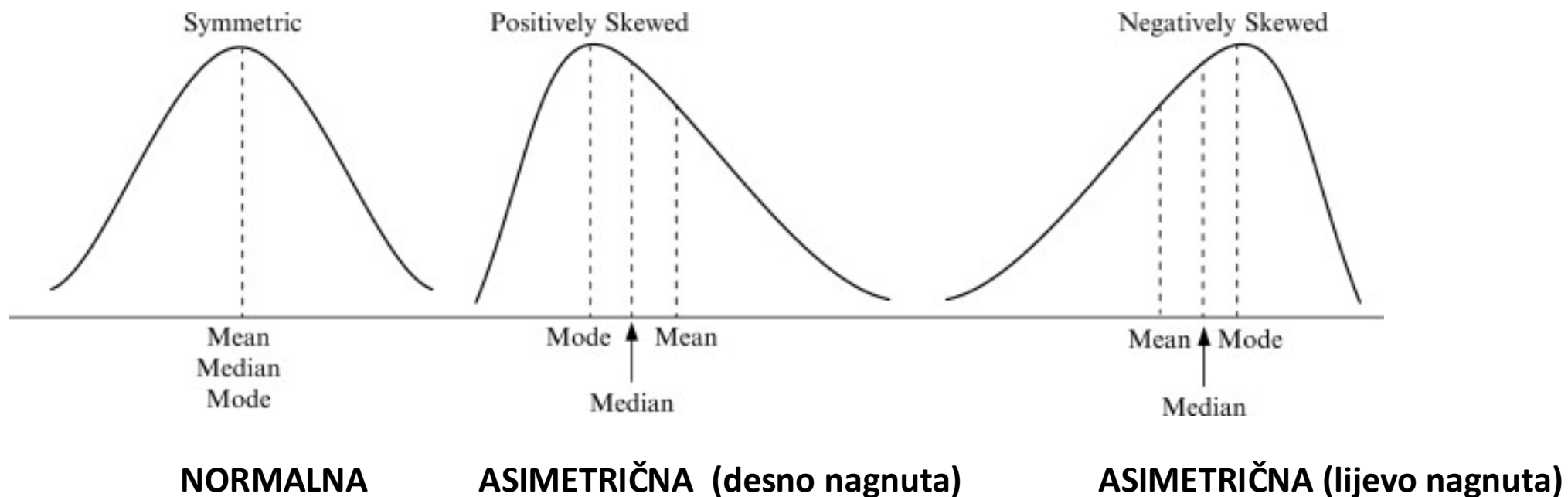


William Sealy Gosset



Srednje vrijednosti

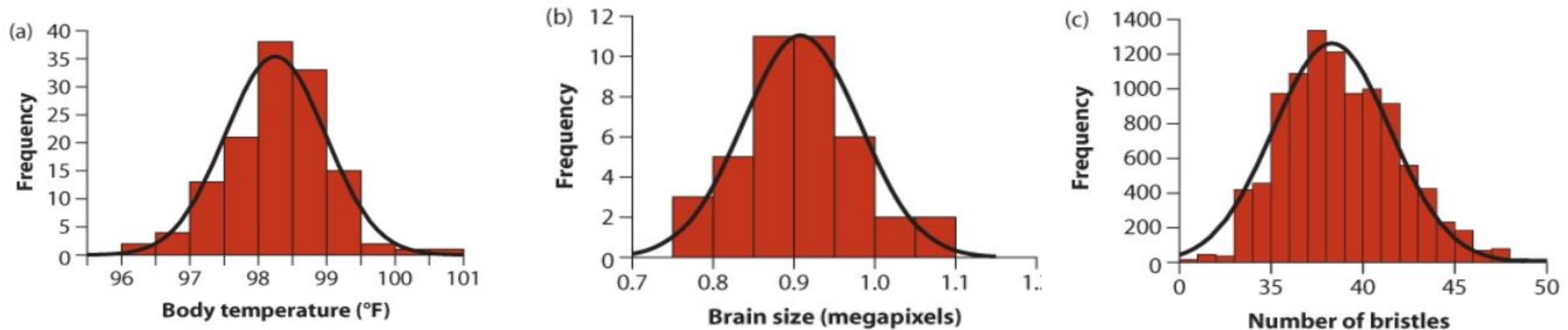
- **aritmetička sredina**, prosjek (*mean*) - μ
- **medijan** – vrijednost u sredini - od koje je 50 % podataka manja i 50% podatka veća
- **mode** – najčešća vrijednost



nagnutost distribucije – u kojem smjeru je “teži rep”

Normalna distribucija

- velik broj bioloških kontinuiranih varijabli u populacijama imaju normalnu distribuciju

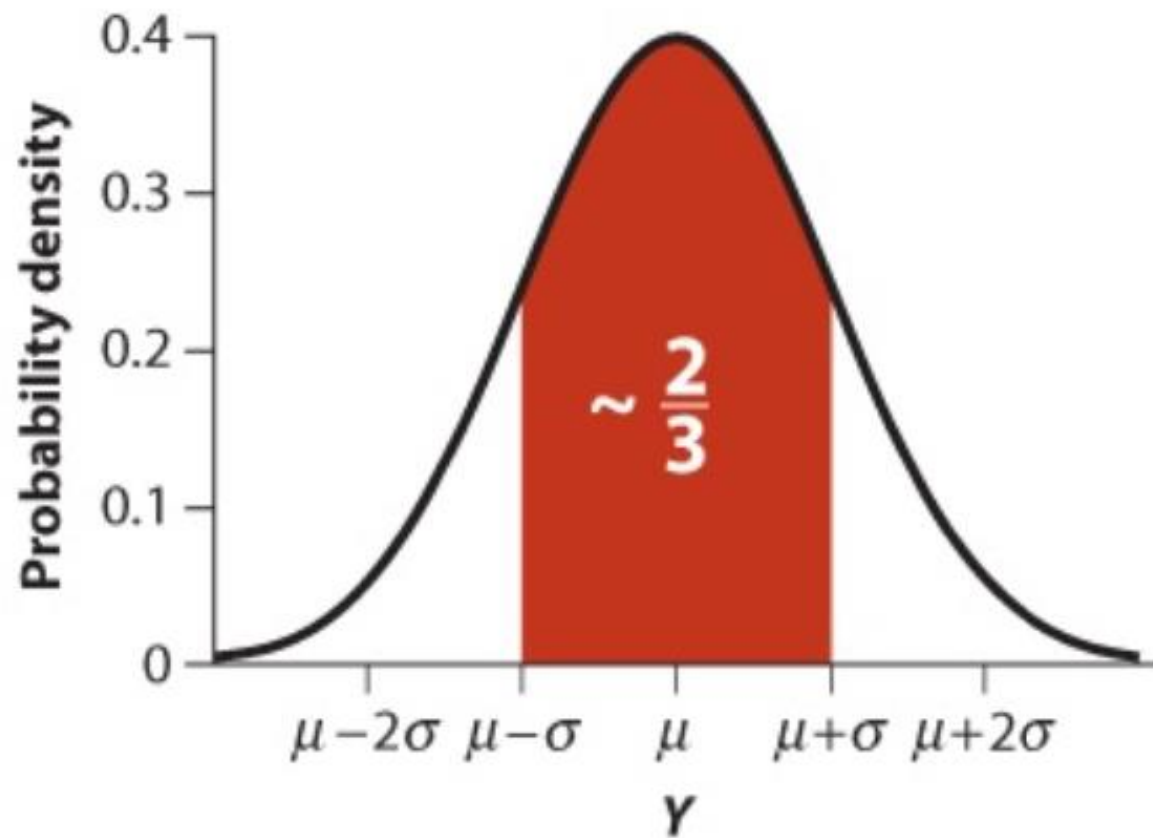


a) ljudska temperatura tijela (b) veličina mozga studenata (broj pixela na MRI) (c) broj dlačica na 4 i 5 segmentu vinske mušice.

CRNA LINIJA POKAZUJE NORMALNU DISTRIBUCIJU (gustoća vjerojatnosti, *probability density*) POPULACIJE ISTE SREDNJE VRIJEDNOSTI I STANDARDNE DEVIJACIJE KAO U PODACIMA DOBIVENIH IZ UZORKA (

Normalna distribucija

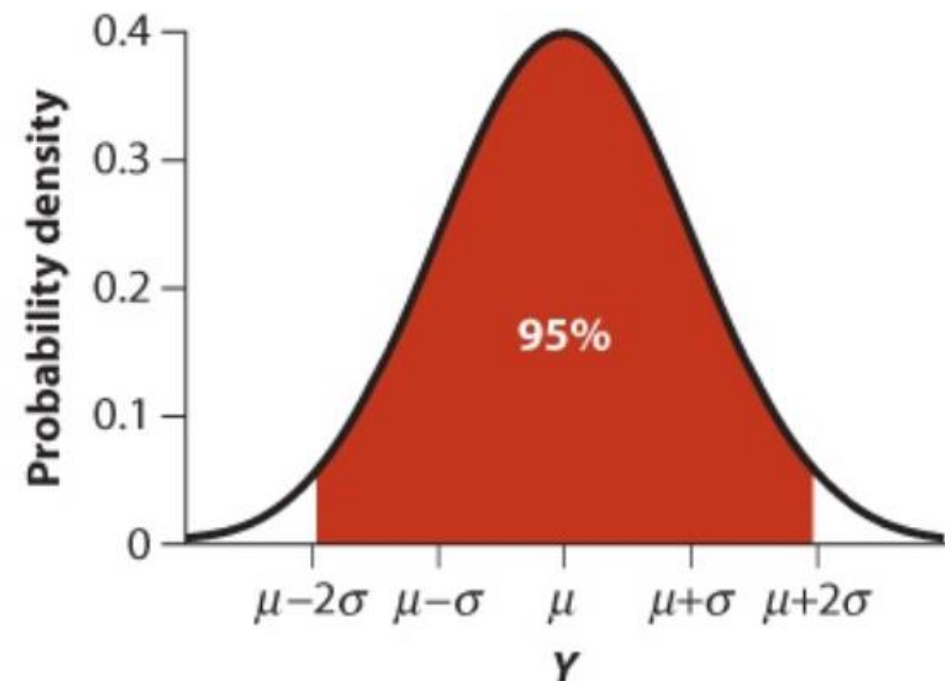
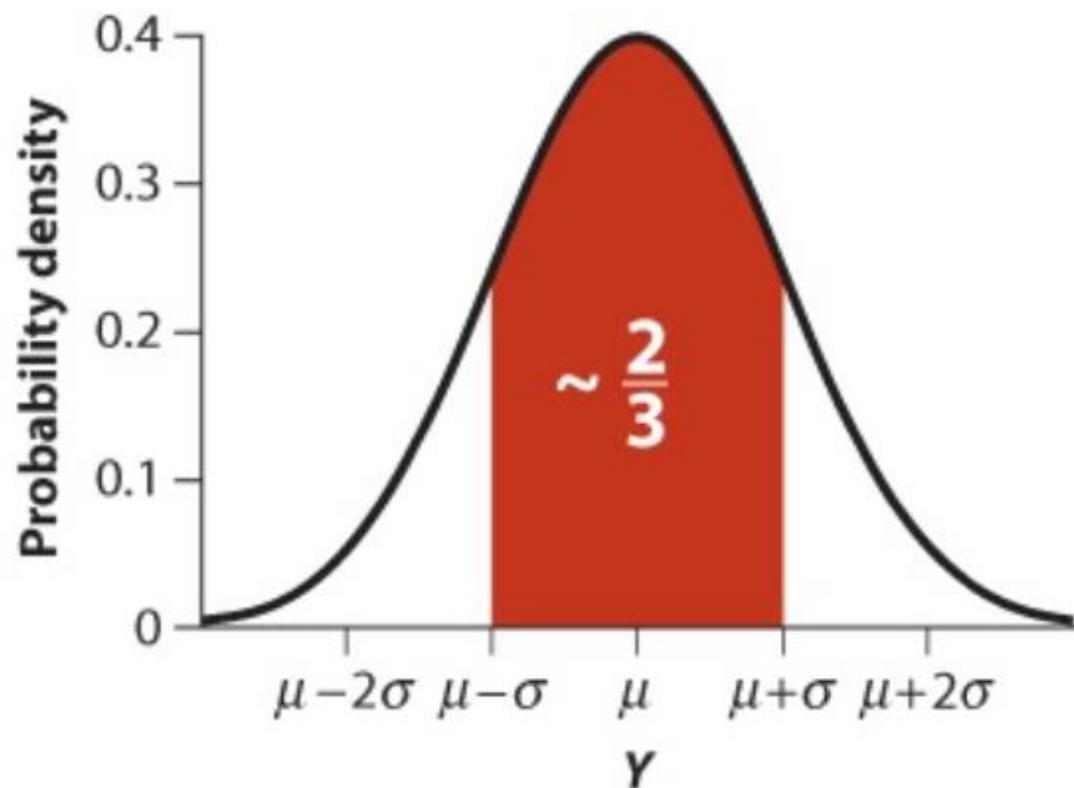
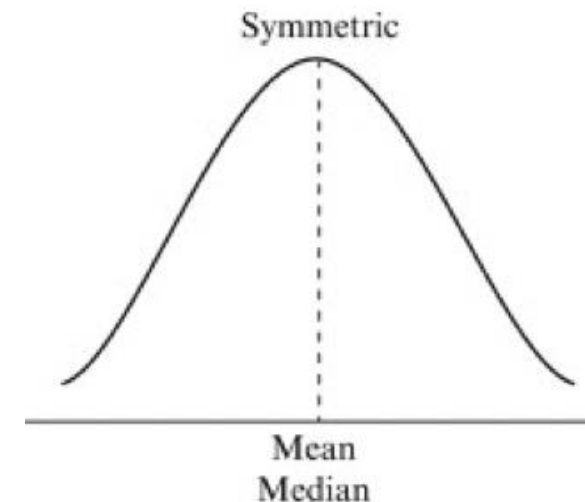
- μ – aritmetička sredina, prosjek (*mean*)
- s^2 – varijanca – mjera raspršenosti - prosječno kvadratno odstupanje podataka od aritmetičke sredine.
- σ – standardna devijacija – mjera raspršenosti uzorka, korijen varijance (iste jedinice kao podaci)



$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

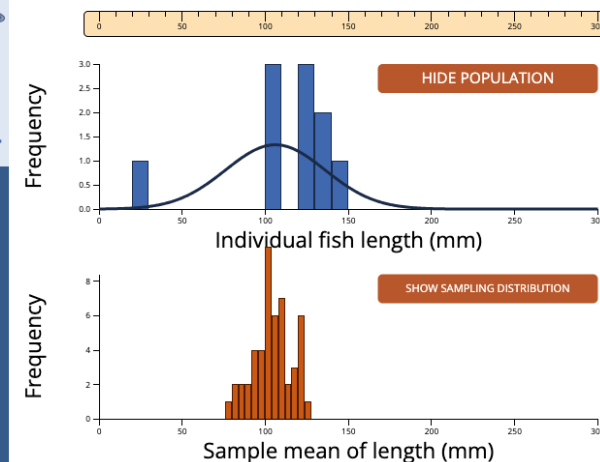
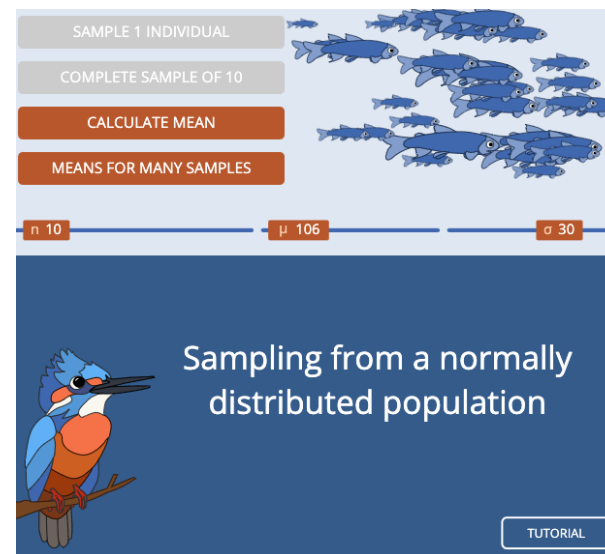
Normalna distribucija

- srednje vrijednosti, medijan, mode i aritmetička sredina u sredini
- jednake standardne devijacije
- otprilike $\frac{2}{3}$ površine nalazi se unutar jedne standardne devijacije
- 95% površine nalaze se unutar 1.96 standardne devijacije



Centralni teorem normalne distribucije

- vjerojatnost distribucije
 - ako opetovano uzorkovanje daje slične parametre (npr. srednja vrijednost), onda je metoda uzorkovanja **precizna**
 - **distribucija srednjih vrijednosti pojedinih uzoraka je normalno distribuirana**
 - srednja vrijednost srednjih vrijednosti više uzoraka trebala bi biti ista kao srednja vrijednost populacije
-
- *aritmetička sredina više uzorka sličí aritmetičkoj sredini populacije normalne razdiobe!*



<https://www.zoology.ubc.ca/~whitlock/Kingfisher/CLT.htm>

https://gallery.shinyapps.io/CLT_mean/

Parent distribution (population):

☐ Normal

☐ Uniform

☒ Right skewed

☐ Left skewed

Skew:

Low skew ▼

Sample size:

2 30 500

2 52 102 152 202 252 302 352 402 452 500

Number of samples:

10 200 1,000

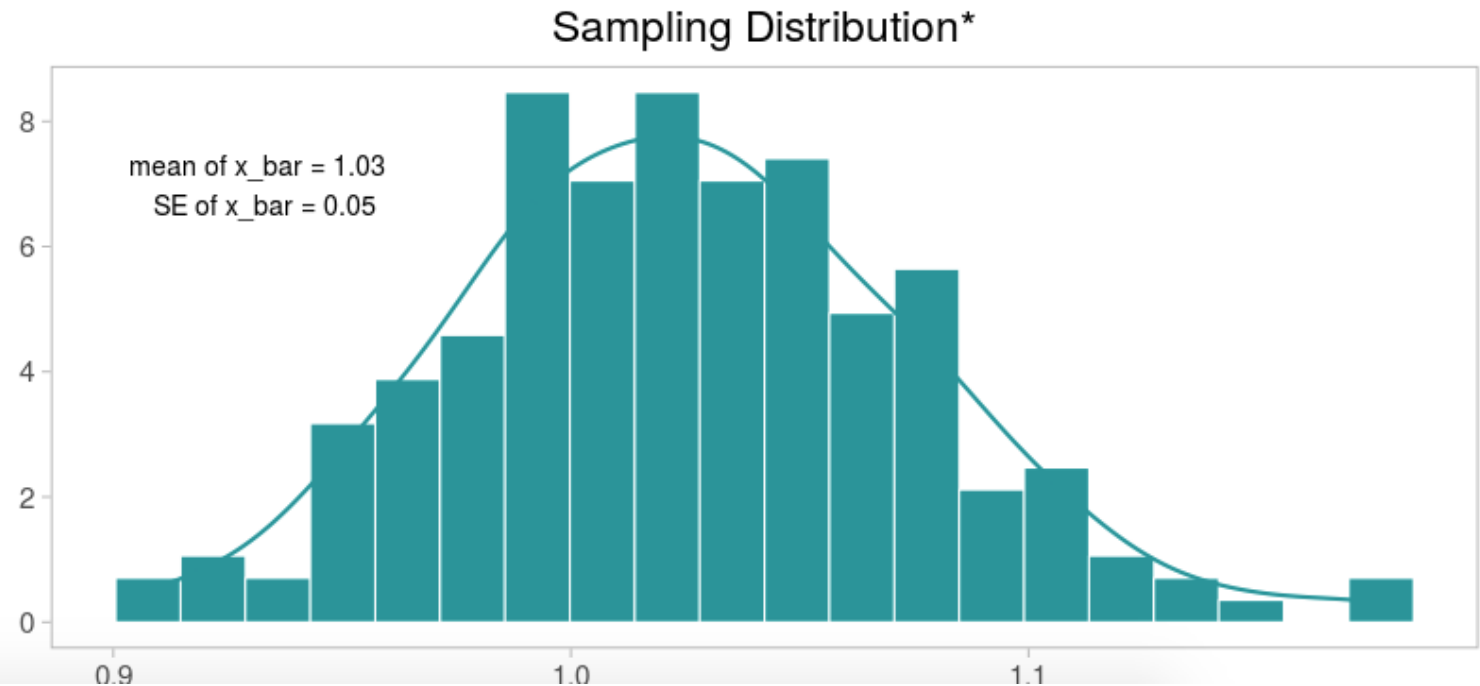
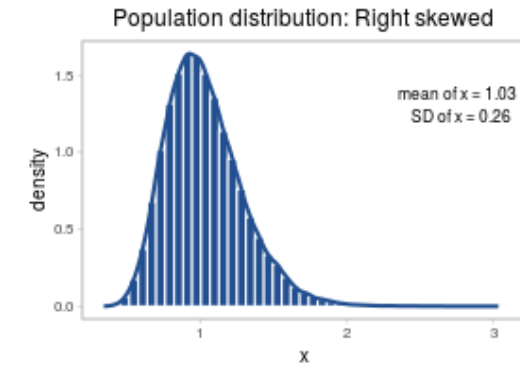
10 109 208 307 406 505 604 703 802 901 1,000

[View the code](#)

[Check out other apps](#)

[Learn more for free!](#)

According to the Central Limit Theorem (CLT), the distribution of sample means (the sampling distribution) should be nearly normal. The mean of the sampling distribution should be approximately equal to the population mean (1.03) and the standard error (the standard deviation of sample means) should be approximately equal to the SD of the population divided by square root of sample size ($0.26/\sqrt{30} = 0.05$). Below is our sampling distribution graph. To help compare, population distribution plot is also displayed on the right.



Mjerila raspršenosti

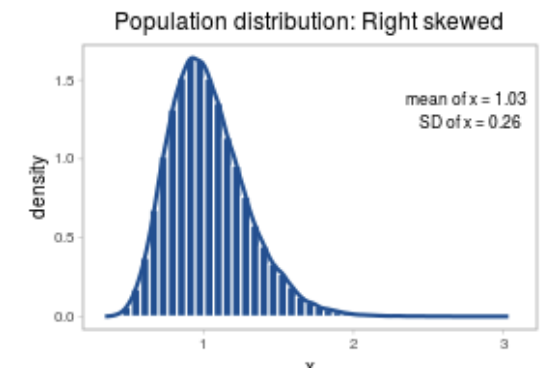
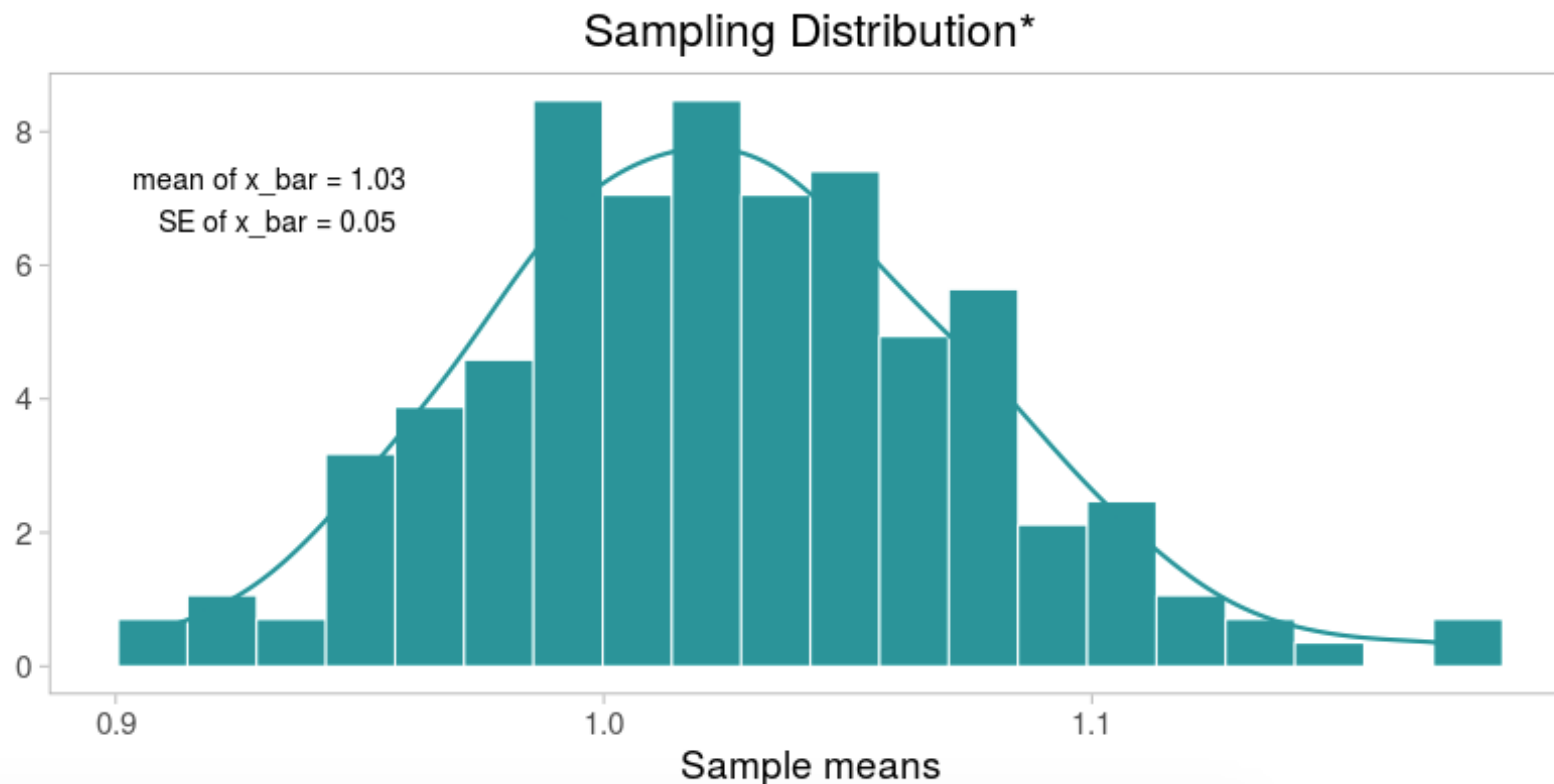
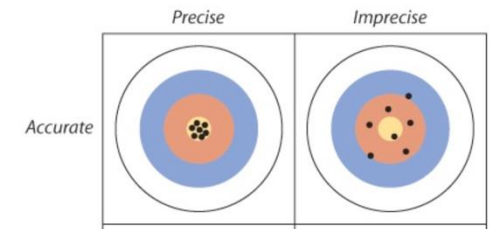
- varijanca i standardna devijacija
- standardna greška – standardna devijacija srednjih vrijednosti više uzoraka
- kvartili
- minimum i maksimum
- intervali pouzdanosti

Mjerila raspršenosti – standardna greška

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

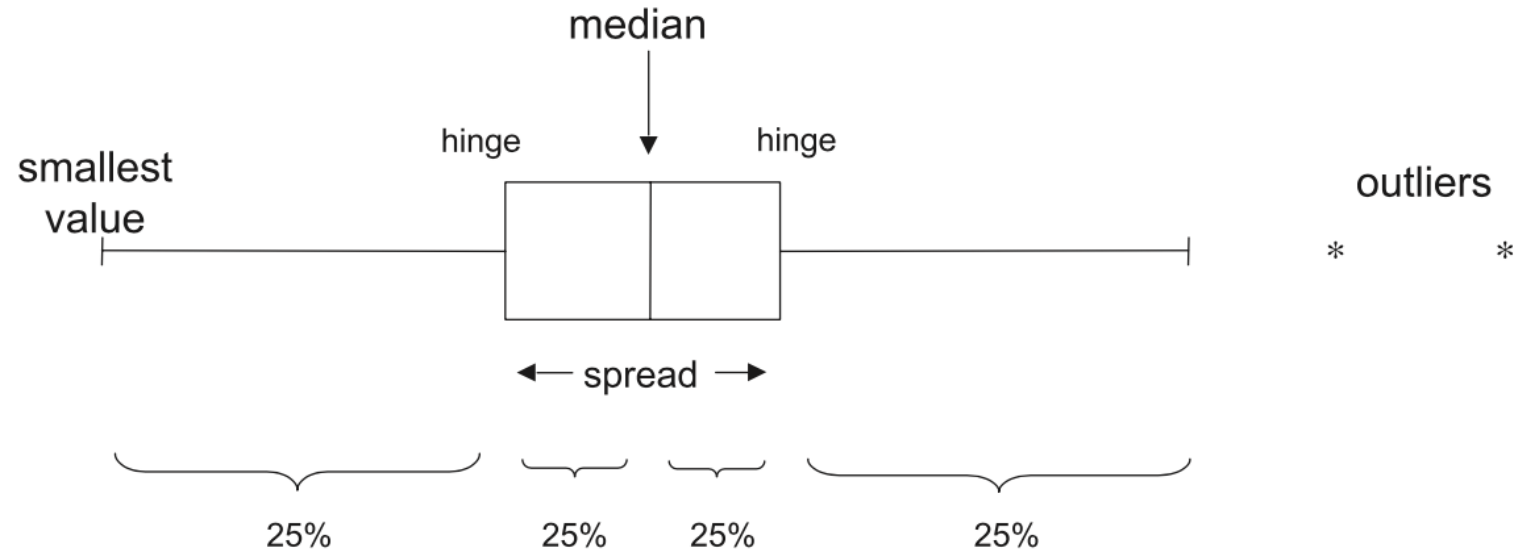
Standardna greška (SE) je standardna devijacija srednjih vrijednosti više uzoraka

- mala SE znači da svaki uzorak dobro opisuje srednju vrijednost populacije (PRECIZNOST)
- **SE se smanjuje s povećanjem veličine uzorka!**
- SE ne opisuje prvenstveno raspršenost (varijabilnost), već preciznost podataka!



Mjerila raspršenosti – kvartili, minimum i maximum

- četiri kvartila, svaki sadrži 25% podataka mjerenja



- *“outliers” – ekstremne vrijednosti van raspona ostalih podataka – uvijek provjeriti mjerenje!*
- *jako utječu na aritmetičku sredinu, ali na medijan ne*
- *više načina izračunavanja, ali može se i subjektivno odrediti*

Mjerila raspršenosti – kvartili, minimum i maximum

- četiri kvartila, svaki sadrži 25% podataka mjerenja



Ramos et al. PNAS 2004

I'd give my right arm for a female

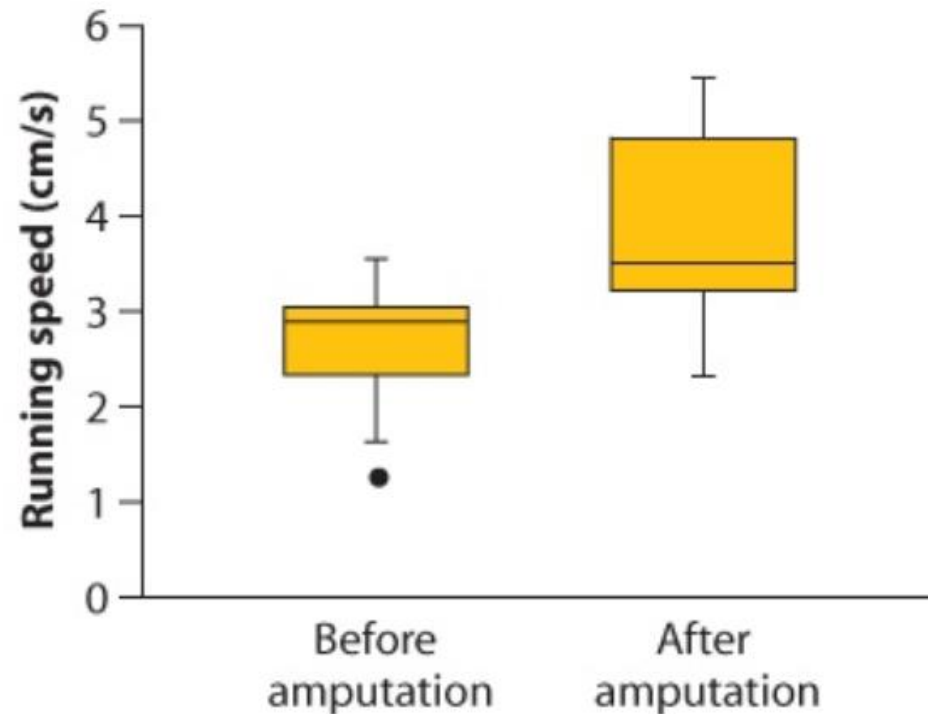


Figure 3.2-2

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

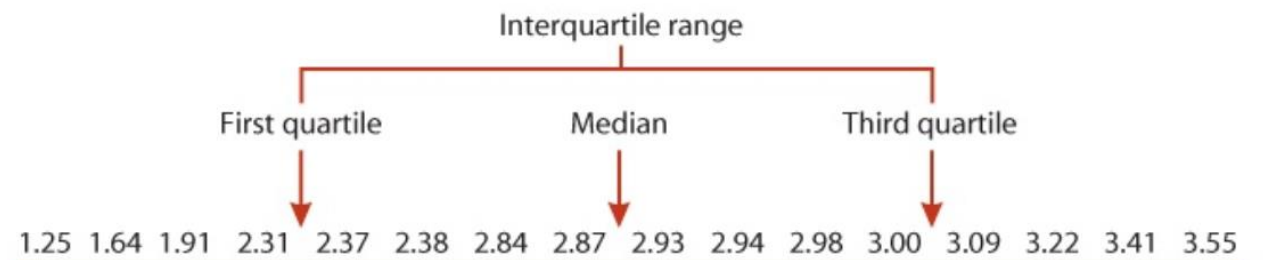


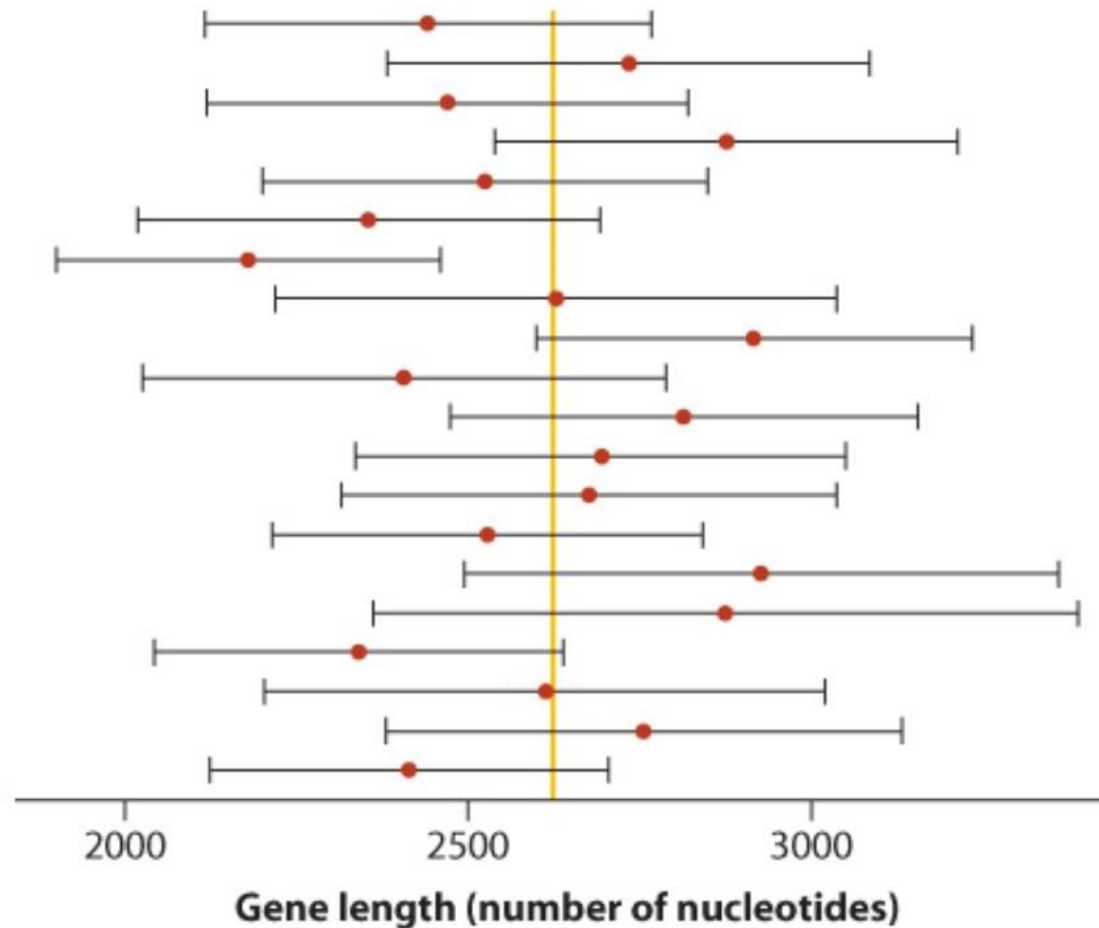
Figure 3.2-1

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

outlier – odmaknut više od 1.5 interkvantilnog raspona od “kutije”

Mjerila raspršenosti – intervali pouzdanosti

- raspon vrijednosti koji vjerojatno (npr. s vjerojatnošću 95%) uključuje srednju vrijednost populacije (ili drugi parametar)
- informiraju nas o **pouzdanosti srednje vrijednosti uzorka** – koliko je srednja vrijednost uzorka blizu srednjoj vrijednosti populacije



95% intervali pouzdanosti za aritmetičku sredinu 20 odvojenih slučajnih uzoraka po 100 gena jednog ljudskog genoma. *Točke su aritmetičke sredine, okomita crta aritmetička sredina uzorka.*

Figure 4.3-1

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

99% confidence intervals for the mean

Successes: 19 Failures: 1 Success rate: 95.0%

MAKE A SAMPLE

PAUSE SAMPLING

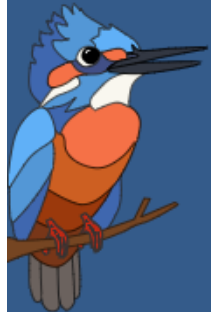
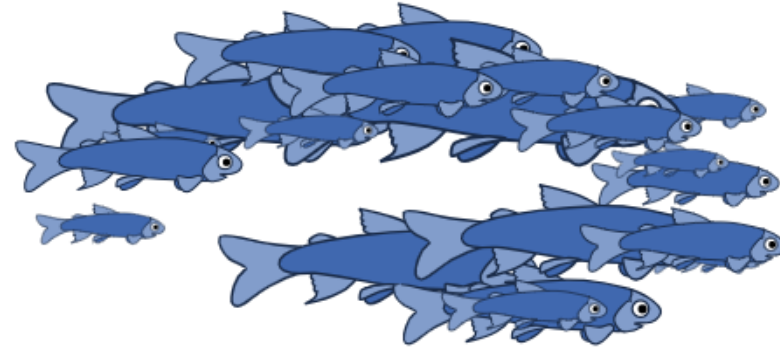
FASTER!

USE 95% CONFIDENCE

n 10

μ 105

σ 30



Confidence intervals for the mean

TUTORIAL

