

Slijedi 5 pitanja za vašu diskusiju

- U sljedećih pola sata pokušajte pismeno na njih odgovoriti te mi po želji možete predati odgovore.
- Na sljedećem satu prodiskutirati ćemo najpotpunije odgovore (za svako pitanje 1 do 2). Studenti po svakom odgovoru odabranom za diskusiju dobivaju bodove koji se pribrajaju pismenom dijelu ispita.
- Ukoliko netko od studenata sakupi 4 boda (a osim ovog biti će još jedan set koji će nositi 5 bodova) biti će oslobođen pismenog ispita u dijelu koji se tiče eksperimentalnog dizajna
- Studentima koji sakupe manje od 4 boda, ostvareni bodovi će biti pribrajani bodovi ostvarenim na ispitu

SVAKI ODGOVOR MORA BITI ARGUMENTIRAN!

1. Skupina lososa jede manje dnevno zimi nego ljeti. Predložite dvije hipoteze koje bi to mogle objasniti. Za svaku od njih predložite predviđanje i način testiranja. (0,5b)
2. Objasnite što detaljnije, zašto je replikacija (broj replika u uzorku) važan dio eksperimentalnog dizajna. (0,5b)
3. Ispitujete učinak upotrebe kompjutera na kvalitetu vida u ljudi. Hoćete li sprovesti klasifikaciju prije uzorkovanja? Ako da, opišite kako biste definirali klasu. Ako ne, kakvo bi uzorkovanje proveli, randomizirano ili neki drugi tip uzorkovanja? Obrazložite svoj odgovor. (1b)
4. Pretpostavite da želite usporediti masu poljskih miševa na području Medvednice sa masom onih na području Biokova. Koji tip uzorkovanja biste proveli? Ovisno o tipu uzorkovanja, kako biste odredili broj replika u svakom uzorku? (1b)
5. Analizirate učinak novog hormonskog tretmana na izduživanje korijena vrijednog varijeteta rajčice. Imate samo 10 sjemenki sto vam omogućuje rad na 10 klijanaca, po 5 tretiranih i 5 netretiranih. Alternativa je da klijance razmnožite u kulturi tkiva tehnikom nodalnih odsječaka te da iz svakog u 30 dana dobijate 6 zakorijenjenih biljčica. Sada imate 60 biljaka te ih 30 možete koristiti za tretman a 30 za kontrolu. Postoji li problem s pseudoreplikacijom, zašto? Kako biste ga kompenzirali kroz uzorkovanje? Kako biste pilot studijom unaprijed provjerili ima li uopće smisla raditi takav eksperiment? Imate li neku svoju ideju(2b)

Odnos između eksperimentalnog dizajna i statistike

- eksperimentalni dizajn i statistika blisko su povezani i bitno je da razmislite o statistici koju ćete koristiti za analizu podataka prije nego što ih prikupite.
- svaki statistički test ima malo drugačije pretpostavke o vrsti podataka koje zahtijeva ili o vrsti hipoteze koju može testirati, stoga je bitno biti siguran da se podaci koje prikupljate mogu analizirati testom koji će testirati hipotezu koja vas zanima.
- Jedini način da budete sigurni u to je da unaprijed odlučite kako ćete analizirati svoje podatke.



Eksperimentalni dizajn

Mislite na statistiku!!

Prevedite hipotezu u statistička pitanja:

- 1) Koju vrstu podataka?
- 2) Kakav statistički test?
- 3) Koja veličina uzorka?**

VAŽNO PITANJE – VELIČINA UZORKA – BROJ REPLIKA

- O čemu ovisi veličina uzorka?
- Primjer predizbornih anketa (broj ispitanika koji obećava točnost predikcije ovisi o mnogo faktora, i nije uvijek 30 niti je uvijek potreban isti broj ispitanika)
- Veličina i kvaliteta uzorka određuje točnost procjene!

Prije eksperimenta – u svrhu procjene velicine uzorka (DESKRIPTIVNA STATISTIKA)

- Deskriptivna statistika opisuje prirodu uzorkovanih podataka

Analiza prije eksperimenta

- Najvažniji dio statistike – ima najviše utjecaja na rezultate istraživanja
- Matematički gledano najlakši dio statistike
- Jako utječe na složenost statističke analize nakon eksperimenta
- Nitko to ne radi!

Glavni parametri descriptive statistike

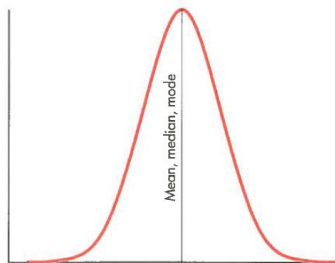
- Procjena središnje tendencije (da li je opisuje srednja vrijednost ili median)
- Disperzija tj. varijacije mjerenja (standardna devijacija)

Srednja vrijednost ili median

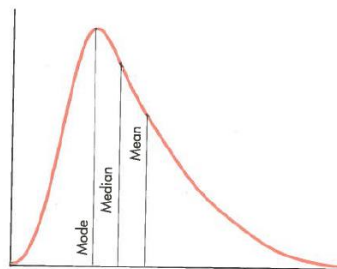
Odluka ovisi o distribuciji podataka:

- Simetrični podaci
- Asimetrični podaci

Simetrična distribucija



Asimetrična distribucija

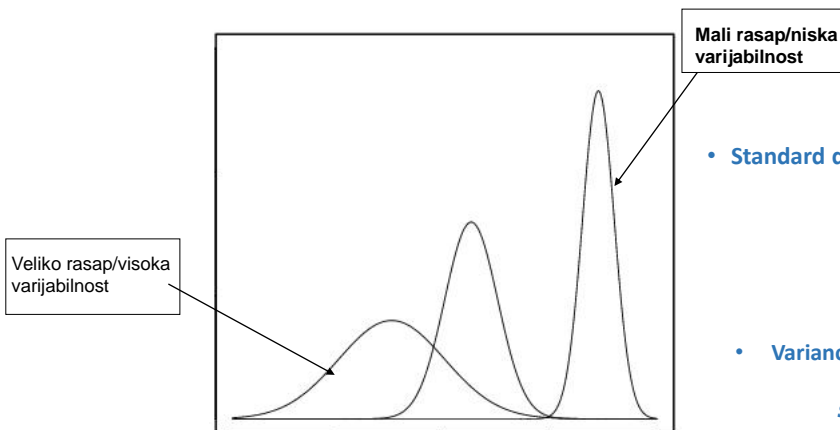


2023-11-03

9

Mjera disperzije podataka

Opisuje varijabilnost oko središnjeg dijela distribucije



- **Standard deviation – variability around the mean**

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2}$$

- **Variance = (standard deviation)²**

$$s^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2$$

2023-11-03

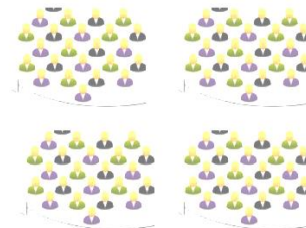
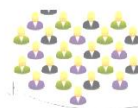
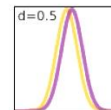
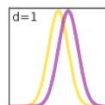
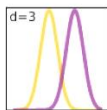
10

Pilot istraživanje

- Pilot istraživanje treba provesti prije prikupljanja svih podataka na manjem broju uzoraka, treba **procijeniti varijabilnost i snagu efekta**
- Pilot istraživanje je korisno jer će poslužiti za provjeru **odabira statističkog testa kao i izračunavanje statističke snage**
- **Snaga našeg testiranja je mjera da nećemo previdjeti razliku koja postoji**

Analiza statističke snage

- Veličina uzorka i veličina efekta (veličina učinka)
- Koliki nam broj replika treba da bi mogli vidjeti efekt određene veličine (snage efekta, učinka)?
- Koliko velik/snažan efekt možemo vidjeti na uzorku određene veličine?



Odabir primjerenog broja replikata (ili test snage efekta)

- 1. **snaga efekta tj. veličina učinka** (npr. koliko će dodatak prehrani povećati masu pilića tj. razlika između srednje vrijednosti mase pilića hranjenih normalnom hranom i hranjenih isto + dodatak), **odražava biologiju sustava koji proučavamo, na to ne možemo utjecati**
- 2. **varijabilnost** - odražava biologiju sustava koji proučavamo, **na to ne možemo utjecati**
- 3. **buka u eksperimentu** (nastaje zbog svih ostalih čimbenika koji utječu na brzinu rasta mase pilića), **odražava biologiju sustava koji proučavamo, na to ne možemo utjecati (odnosno možemo djelomično) ali ne možemo potpuno kontrolirati**
- 4. dizajn pokusa – kako uzorkovati, dodatni kriteriji pri uzorkovanju npr. klasifikacija, grupiranje itd... (i odabrani statistički test)
- 5. broj neovisnih replika (veličina uzorka) – **potpuno u kontroli istraživača**

Primjer varijacija i replika

- **Utječe li primjena GM žitarica u ishrani na debljinu ljuske jajeta kokoši?**
- Problem varijabilnosti!
- Zadatak je vidjeti koliko je varijabilnost povezana s GM prehranom
- NULL hipoteza: debljina ljuske nije u vezi s GM prehranom

Suočavamo se s činjenicom da je debljina ljuske uvjetovana različitim faktorima od genetičkih, porijekla, starosti, količine vode, vlažnosti staništa, vremenom nošenje jaja, temperaturom i slicno. Vjerojatnost da ćemo dobiti istu vrijednost ako uzmemo dva jaja različitih kokoši hranjenih istom hranom je niska. Isto tako ista kokoš danas i sutra ne nese jaja iste debljine ljuske (varijabilnost). Stoga provodimo eksperiment na istoj liniji kokoši, isto starih, mjerimo jaja snesena u isto vrijeme itd. Još uvijek ćemo imati varijabilnost zbog koje moramo imati replikate, no djelomično ćemo stvoriti buku.

- Kontrola eksperimenta (identični uvjeti uz razliku u dodatku GM žitarica)
- Veći broj jaja

Ako imamo perfektnu Gaussovu distribuciju tada je 95% svih vrijednosti unutar dvije SD vrijednosti, a 68% unutar 1 SD

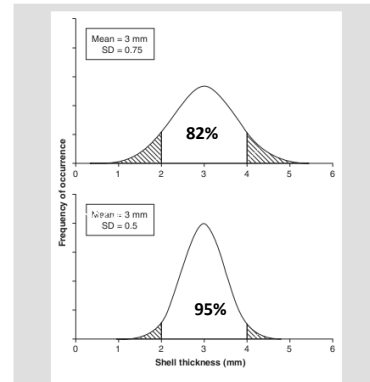
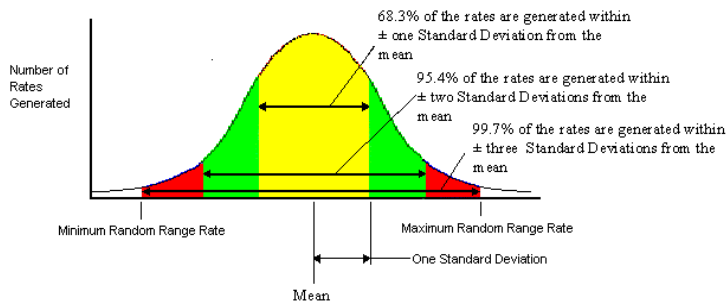


Figure 3.1 Two Gaussian or normal distributions to illustrate the kind of variation we are likely to see if we were to measure the shell thickness of many hundreds of thousands of eggs from hens fed a standard diet. Both distributions cluster around the same mean value (3 mm), but the distribution in the upper panel has a larger standard deviation, because the variation of the values around the mean value is greater: in the upper distribution, approximately 18% of the eggs have shells that are thicker than 4 mm or thinner than 2 mm, whilst in the lower distribution this is only 5%.

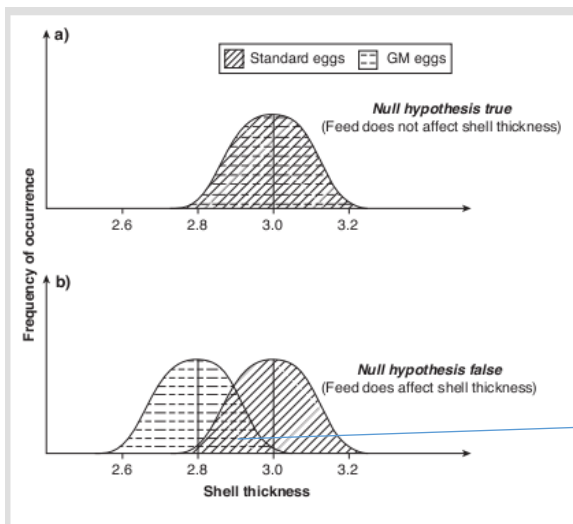


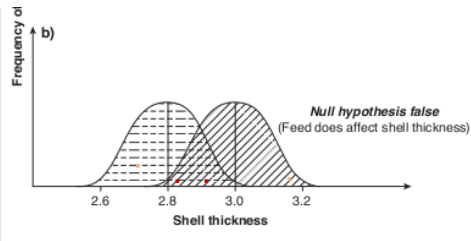
Figure 3.2 The distributions of egg shell thicknesses under two different scenarios. Under both scenarios shell thicknesses follow Gaussian distributions but in the top panel, the two populations have the same mean because the treatment is having no effect on shell thickness (i.e. the null hypothesis is true), whilst in the bottom panel the treatment is affecting shell thickness (i.e. the null hypothesis is false) and the means of the distributions differ.

GM hranjene:

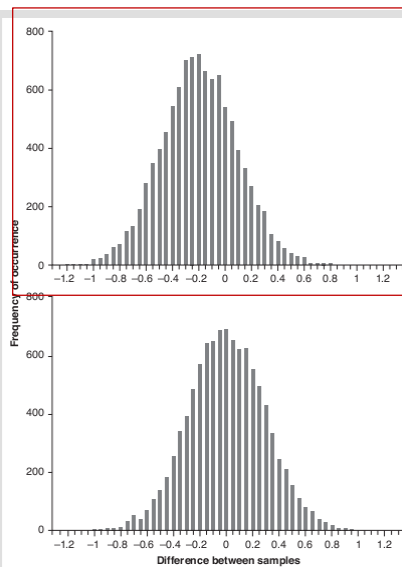
- ista distribucija varijabilnosti
- ako nema razlike – oko iste srednje vrijednosti
- ako ima razlike – oko razl. srednje vrijednosti

U preklapajućoj zoni su vrijednosti koje nam pokazuju da efekta nema

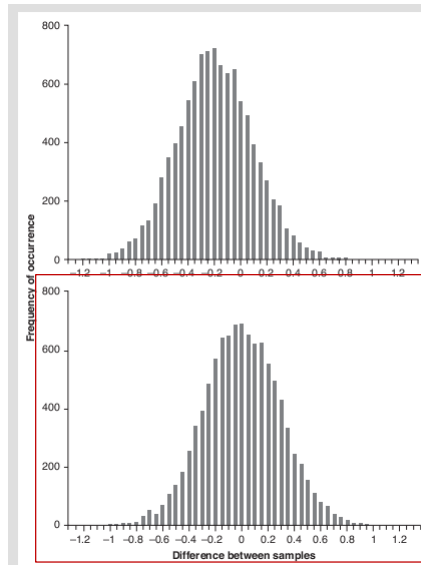
Kada bismo izmjerili sva jaja u populacijama ne bi nam trebala statistika međutim u svakom drugom slučaju potrebna je statistika da pokazuje značaj razlike od 0,2 mm



- Pojedinačna usporedba jednog GM i jednog normalnog jajeta – daje različite razlike
- Ako napravimo veliki broj (10000) nasumičnih usporedba i mjerimo razlike – dobijamo grafove (sljedeći slajd)

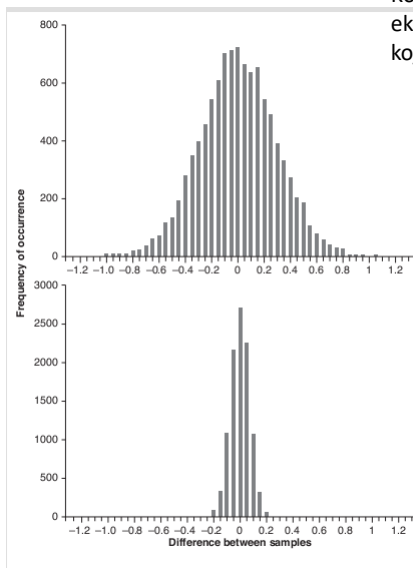


- Većina vrijednosti na gornjem grafu klasterira oko -0,2, no još uvijek ih ima puno koji pokazuju da GM hrana povećava debljinu ljuske (2427) dok 2351 eksperiment pokazuje da GM hrana smanjuje debljinu za više od 0,4mm .



- Većina vrijednosti na donjem grafu klasterira oko 0 ukazujući da nema razlike između GM ljsaka i neGM
- Unatoč tome puno pojedinačnih usporedbi moglo bi dati drugačije rezultate u bilo kojem smislu, npr razliku -0.32 i veću daje 1274

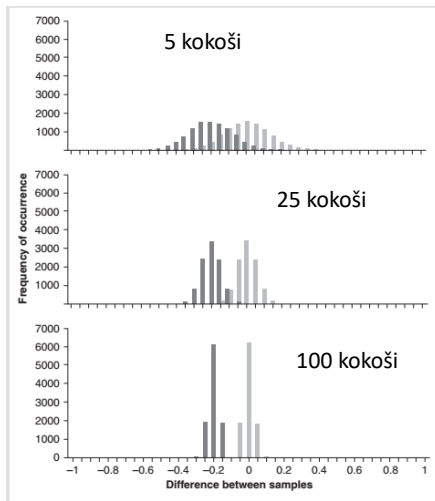
Odnos varijabilnosti klase/populacije i broja replika



Količina varijacija u usporedbama koju vidimo u opisanim eksperimentima ovisi o tome koliko su varijabilne populacije koje uzorkujemo

- Razliku od 0.32 možemo lako očekivati u mjerenjima u gornjem eksperimentu (SD 0,2)
- Kada je SD 0,05 ova se razlika uopće ne može pojaviti

Od 10 000 imaginarnih eksperimenata, niti jedan nije predložio smanjenje debljina ljsuke u GM uzorcima za 0,32 mm ili više.



Snaga efekta procijenjena iz 10 000 imaginarnih eksperimenata u svijetu nul hipoteze (lijevo na svakom grafu) i alternativne hipoteze (desno), ali je veličina uzorka različita.

Kako veličina uzorka raste (broj replikata) diskriminacija dva svijeta je jasnija.

Razlikovanje između nulte i alternativne hipoteze – snaga statistike

Zašto je kokoš u središtu pažnje, a ne jaje?



Broj replika utječe na SD

The screenshot shows a Microsoft Excel spreadsheet with the following data:

	A	B	C	D	E
1			4	4	4
2			1	1	1
3			0	0	0
4			6	6	6
5				4	4
6			1	1	1
7			0	0	0
8			6	6	6
9			4	4	4
10			1	1	1
11			0	0	0
12			6	6	6
13				4	4
14				1	1
15				0	0
16				6	6
17				4	4
18				1	1
19				0	0
20				6	6
21				4	4
22				1	1
23				0	0
24				6	6
25					
26	zbroj		11	33	66
27	srednja vrijednost		2.75	2.75	2.75
28	standardna devijacija		2.75	2.49	2.44
29					
30					

The help window for the STDEV function is open, showing the following information:

STDEV
Estimates standard deviation based on a sample. The standard deviation is a measure of how widely values are dispersed from the average value (the mean).

Syntax
STDEV(number1,number2,...)

Number1, number2, ... are 1 to 30 number arguments corresponding to a sample of a population. You can also use a single array or a reference to an array instead of arguments separated by commas.

Remarks

- STDEV assumes that its arguments are a sample of the population. If your data represents the entire population, then compute the standard deviation using STDEVP.
- The standard deviation is calculated using the "unbiased" or "n-1" method.
- STDEV uses the following formula:

$$\sqrt{\frac{\sum (x - \bar{x})^2}{(n - 1)}}$$
 where \bar{x} is the sample mean AVERAGE(number1,number2,...) and n is the sample size.
- Logical values such as TRUE and FALSE and text are ignored. If logical values and text must not be ignored, use the STDEVPA worksheet function.

Example

Stupičasti prikazi sa SD

Analiza statističke snage

- **Definicija snage:** vjerojatnost da će statistički test odbaciti lažnu nultu hipotezu (H_0).
- **Prijevod:** vjerojatnost otkrivanja učinka, s obzirom da učinak stvarno postoji.
- **Ukratko:** što je veći eksperiment (veličina uzorka), to je veća snaga (veća je vjerojatnost da će se uočiti razlika).
- **Glavni rezultat analize snage:** Procjena odgovarajuće veličine uzorka
 - Prevelik: rasipanje resursa,
 - Premalo: može izostati učinak ($p > 0,05$) + gubitak resursa,
 - Projekti: obrazloženje veličine uzorka,
 - Publikacije: traže izračun snage,
 - 3R : zamjena, smanjenje i usavršavanje

Methods which
avoid or replace
the use of animals

Replacement

Methods which
minimise the
number of animals
used per experiment

Reduction

Methods which
minimise suffering
and improve animal
welfare

Refinement