

ANALYTICKÉ POSTUPY, FORMULACE VÝSLEDKŮ A LIMITŮ V PSYCHOLOGII (KVANTITATIVNÍ)

Jitka Lindová

rozložení stránky Vzorce Data			
fx		3,79130434782609	
	C	D	E
	Age	Education	WorkStatusC pravene
	24	6	4
	24	6	2
	23	6	2
	30	4	5
	26	6	1
	22	4	2
	21	4	2
	19	4	3
	23	6	2
	22	6	3
	22	3	5
	20	4	4
	27	7	4
	24	6	2
	23	6	3
	23	6	2
	23	4	5
	23	4	3

PROČ POTŘEBUJEME STATISTIKU



rozložení stránky

Vzorce

Data

f_x	3,79130434782609		
	C	D	E
	Age	Education	WorkStatus pravene
	24	6	4
	24	6	2
	23	6	2
	30	4	5
	26	6	1
	22	4	2
	21	4	2
	19	4	3
	23	6	2
	22	6	3
	22	3	5
	20	4	4
	27	7	4
	24	6	2
	23	6	3
	23	6	2
	23	4	5
	23	4	3

K ČEMU SLOUŽÍ STATISTIKA?

- *popisuje velké soubory údajů pomocí charakteristických čísel*
 - *deskriptivní statistika*
- *ověřuje platnost našich tvrzení (hypotéz) na základě omezeného vzorku dat*
 - *testová statistika*
- *hledá skryté zákonitosti v nepřehledně velkých souborech údajů*
 - *explorační statistika*

rozložení stránky Vzorce Data

fx

3,79130434782609

	C	D	E
	Age	Education	WorkStatusC pravene
	24	6	4
	24	6	2
	23	6	2
	30	4	5
	26	6	1
	22	4	2
	21	4	2
	19	4	3
	23	6	2
	22	6	3
	22	3	5
	20	4	4
	27	7	4
	24	6	2
	23	6	3
	23	6	2
	23	4	5
	23	4	3

DESKRIPTIVNÍ (POPISNÁ) STATISTIKA

- Charakteristiky *centrální tendence* – čísla, která nejlépe zastupují celý soubor dat
 - *aritmetický průměr*: $\sum x_n/n$
 - *medián*: hodnota ležící právě uprostřed
- Charakteristiky *variability*
 - *rozptyl, VAR* ($s^2 = \sum (X_i - X_m)^2 / n - 1$)
 - *směrodatná odchylka, SD* ($SD = \sqrt{s^2}$)

rozložení stránky Vzorce Data

f_x

3,79130434782609

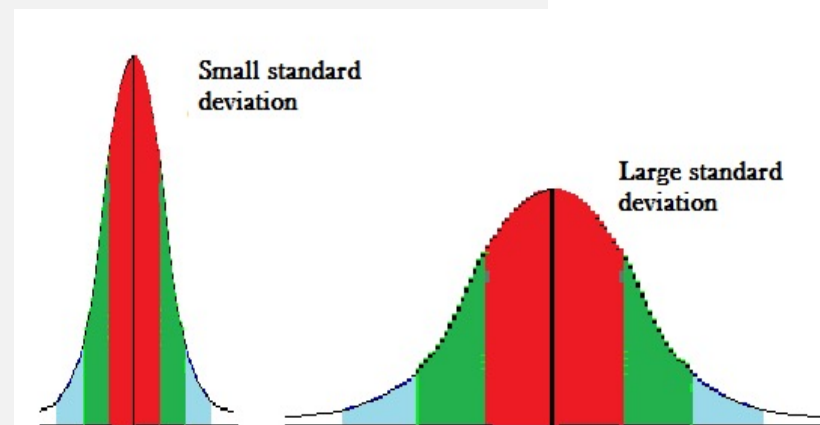
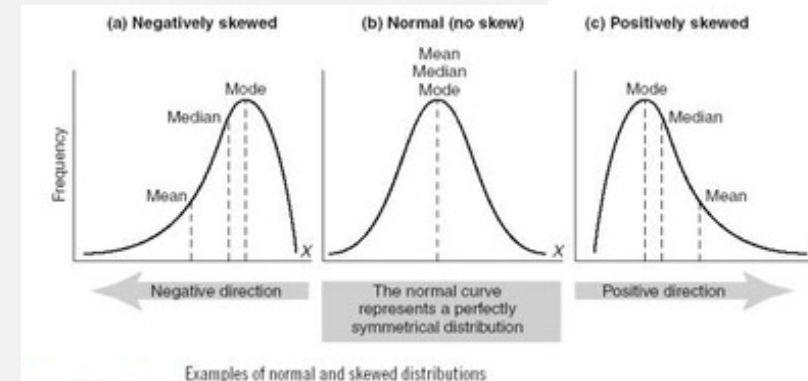
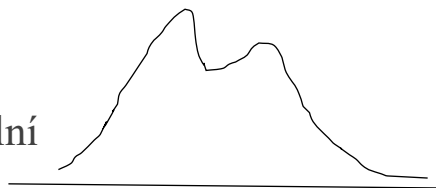
	C	D	E
	Age	Education	WorkStatusC pravene
	24	6	4
	24	6	2
	23	6	2
	30	4	5
	26	6	1
	22	4	2
	21	4	2
	19	4	3
	23	6	2
	22	6	3
	22	3	5
	20	4	4
	27	7	4
	24	6	2
	23	6	3
	23	6	2
	23	4	5
	23	4	3

DESKRIPTIVNÍ STATISTIKA

Popisuje četnosti (počty), relativní četnosti (%)

Popisuje rozložení pomocí údajů o

- Centrální tendenci
 - Průměr
 - Medián
 - Modus – bimodální, trimodální
- Variabilitě
 - Rozsah – rozdíl mezi mezními hodnotami
 - Směrodatná odchylka – průměrná odchylka od průměru

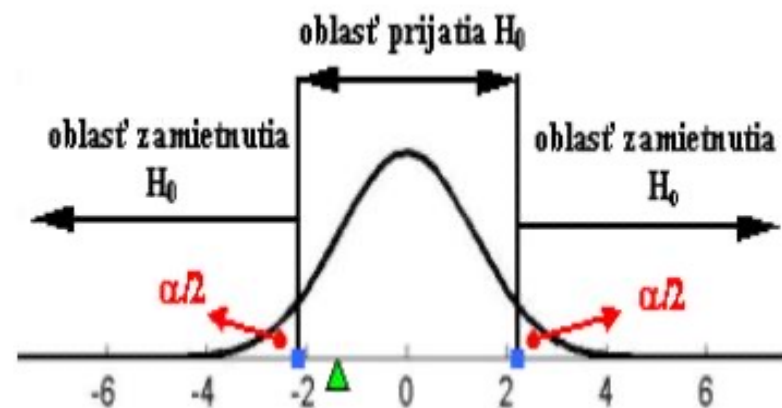


TESTOVÁ STATISTIKA

Testuje pravděpodobnost, s jakou mohu zamítnout hypotézu H_0

- H_0 – tvrzení o rozdělení dat, prakticky např. „není rozdíl od dalšího souboru“, „není vztah s další proměnnou“ atd.
- H_A nebo H_1 - je rozdíl, vztah
- α – hladina významnosti – hranice případů stejně daleko nebo vzdálenějších od středu předpokládaného rozdělení (odpovídajícího H_0), při které zamítám hypotézu
 - Tu si předem stanovím, např. 0,05 (= 5%)
- p-hodnota – s jakou pravděpodobností se stane, že při předpokladu H_0 dostanu náhodou stejný (nebo ještě méně H_0 odpovídající) výsledek, jako ten, který mi vyšel

- A. chí kvadrát
- B. t-test dvouvýběrový, ANOVA, Mann-Whitney U test (nepar.), Kruskal-Wallis ANOVA (nepar.)
- C. párový t-test, Wilcoxon matched-pairs signed-rank test
- D. Pearsonova korelace (Spearman, Kendall), lineární regrese



STATISTICKÉ TESTY

- Odpovídá délka vypracování testu u dětí v dané škole udávané hodnotě 60 minut?
- Je poměr pohlaví v populaci kmene Hadza vyrovnaný?
- Liší se délka života u žen a mužů?
- Mají introvertní jedinci menší počet potomků?
- Ve všech těchto případech nás zajímá situace v rozsáhlejším souboru, nikoli jen náš vzorek dat.
- Pokud je náš vzorek náhodným výběrem dat, jsou naše výsledku ovlivněny náhodnou chybou
- Vliv náhody musíme zhodnotit, než učiníme závěr pro celý soubor.

rozložení stránky Vzorce Data

fx

3,79130434782609

	C	D	E
	Age	Education	WorkStatusC pravene
	24	6	4
	24	6	2
	23	6	2
	30	4	5
	26	6	1
	22	4	2
	21	4	2
	19	4	3
	23	6	2
	22	6	3
	22	3	5
	20	4	4
	27	7	4
	24	6	2
	23	6	3
	23	6	2
	23	4	5
	23	4	3

INTROVERTI, EXTRAVERTI A POČET POTOMKŮ

- Introverti: 1,0,0,3,2,2,3,1,2,1,2
- Extraverti: 2,3,3,1,0,1,1,1,2,3

Vytvoření stránky Vzorce Data			
fx		3,79130434782609	
	C	D	E
	Age	Education	WorkStatusC pravene
	24	6	4
	24	6	2
	23	6	2
	30	4	5
	26	6	1
	22	4	2
	21	4	2
	19	4	3
	23	6	2
	22	6	3
	22	3	5
	20	4	4
	27	7	4
	24	6	2
	23	6	3
	23	6	2
	23	4	5
	23	4	3

INTROVERTI, EXTRAVERTI A POČET POTOMKŮ

- Introverti: 1,0,0,3,2,2,3,1,2,1
- Extraverti: 2,3,3,1,0,1,1,1,2,3
- Introverti průměr: 1,5
- Extraverti průměr: 1,7

Vytvoření stránky Vzorce Data			
fx		3,79130434782609	
	C	D	E
	Age	Education	WorkStatusC pravene
	24	6	4
	24	6	2
	23	6	2
	30	4	5
	26	6	1
	22	4	2
	21	4	2
	19	4	3
	23	6	2
	22	6	3
	22	3	5
	20	4	4
	27	7	4
	24	6	2
	23	6	3
	23	6	2
	23	4	5
	23	4	3

INTROVERTI, EXTRAVERTI A POČET POTOMKŮ

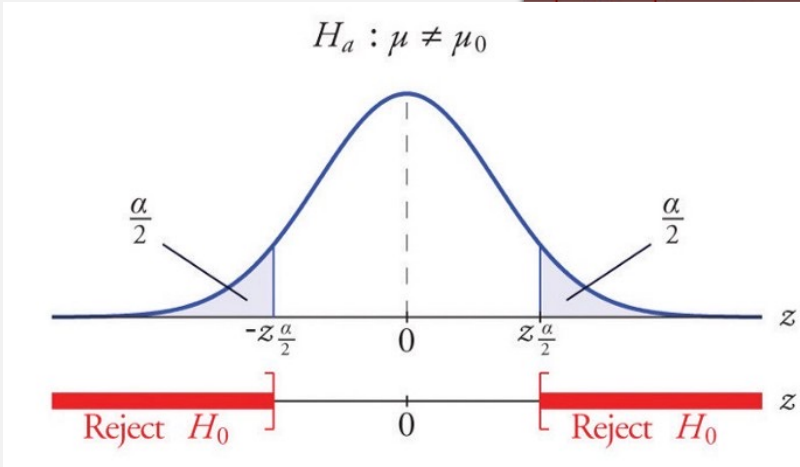
- Introverti: 1,0,0,3,2,2,3,1,2,1
- Extraverti: 2,3,3,1,0,1,1,1,2,3
- Introverti průměr: 1,5
- Extraverti průměr: 1,7
- $t_{18} = 0,41, p = 0,68$

Vzorce			
3,79130434782609			
	C	D	E
	Age	Education	WorkStatusC pravene
	24	6	4
	24	6	2
	23	6	2
	30	4	5
	26	6	1
	22	4	2
	21	4	2
	19	4	3
	23	6	2
	22	6	3
	22	3	5
	20	4	4
	27	7	4
	24	6	2
	23	6	3
	23	6	2
	23	4	5
	23	4	3

CHYBA I. A II. DRUHU, HLADINA VÝZNAMNOSTI A P-HODNOTA



Decision Made	State of Nature	
	Null true	Null false
Reject Null	Type I error (α)	Correct decision ($1 - b$) POWER!
Fail to reject null	Correct decision	Type II error (b)

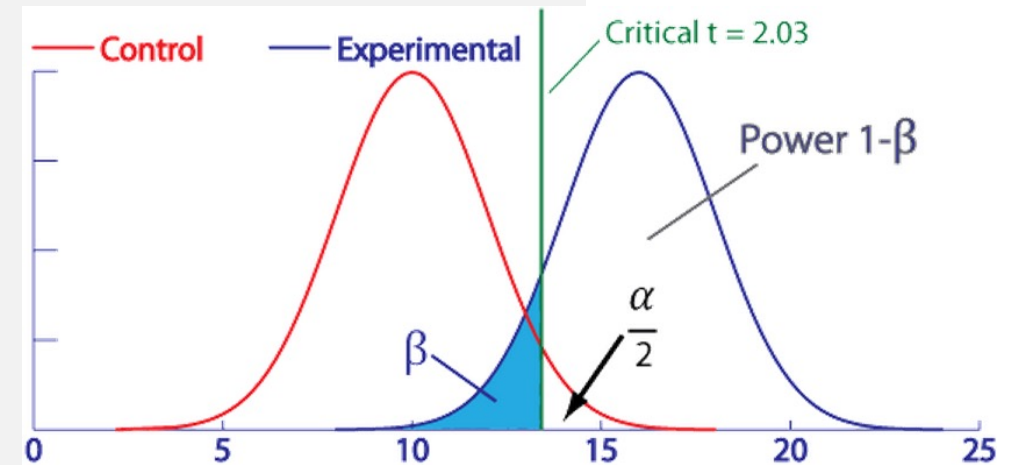


E	WorkStatusC	pravene
4		
2		
2		
5		
1	26	6
2		
2		
3		
2		
3		
5		
4		
4		
2		
3		
2	23	6
5	23	4
3	23	4

CHYBA I. A II. DRUHU, HLADINA VÝZNAMNOSTI A P-HODNOTA



Decision Made ↓	State of Nature	
	Null true	Null false
Reject Null	Type I error (α)	Correct decision ($1 - b$) POWER!
Fail to reject null	Correct decision	Type II error (b)



ANALÝZA DAT

Volba testu, tvorba modelu

Deskriptivní statistika

Testová statistika

ŠKÁLY

Jakému typu škály odpovídá moje proměnná?

Spojité (continuous)

- Číselná řada, kde stejný číselný rozdíl reprezentuje stejný reálný rozdíl
 - Intervalová nebo poměrová
 - Lidé měřící 158 a 160 cm mají stejný rozdíl jako lidé měřící 177 a 179 cm
- Hodnoty měrných jednotek - délka, váha, teplota, čas...; četnost, podíl, počet ...
 - Může a nemusí zahrnovat jen celá čísla
 - Může a nemusí mít normální rozdělení
- Vždy zapisuju jako číslo.
- Stejná (jedna) jednotka
 - Jednotka v legendě

Ordinální

- Hodnoty je možné seřadit
 - Ale není jisté, že intervaly mezi následujícími hodnotami jsou reálně stejné
 - Vůbec – málo – středně – hodně – velmi (1-2-3-4-5)
 - Jednou týdně – 2x týdně – více než 2x týdně – více než 4x týdně – denně
- Dotazníkové hodnotící škály, známky ve škole, IQ?, stupeň vzdělání...
- Vždy zapisuju jako číslo
- Jednoznačné určení čísel na škále

Nominální (kategorální)

- Hodnoty není možné seřadit
- Reprezentují kvalitativní kategorie
 - Studenti – ženy na mateřské – pracující
 - Sangvinik – cholerik – flegmatik – melancholik
 - Ano – ne – nevím
- Můžu zapsat jako text nebo jako číslo
 - Text musí být přesně stejný pro kategorii
 - Když použiju číslo, legenda

Binární

- Speciální případ nominální škály (ale i ordinální a intervalové)
 - Muži - ženy
 - Ano – ne
 - Úspěch - neúspěch
- Zápis číslem má řadu výhod - lze provádět statistiku
 - Text možno použít, musí být přesně stejný pro kategorii
 - Když použiju číslo, legenda

Volba testu, tvorba modelu

- *Povaha testu*
 - A. *Porovnání četnosti*
 - B. *Porovnání skupin*
 - C. *Párové srovnání:*
 - D. *Korelační studie:*

Sex \ Handed- ness	Handed- ness		
	Right-handed	Left-handed	Total
Male	43	9	52
Female	44	4	48
Total	87	13	100

POJMENOVÁNÍ
PROMĚNNÝCH
ve statistice

Závislá proměnná

Nezávislá proměnná

Testová proměnná

Prediktor

Faktor

Grouping variable

Kovariáta

Kontrolní proměnná

Mediátor

Volba testu, tvorba modelu

- *Povaha testu*
 - A. *Porovnání četnosti*
 - B. *Porovnání skupin*
 - C. *Párové srovnání:*
 - D. *Korelační studie:*

Sex \ Handed-ness			
	Right-handed	Left-handed	Total
Male	43	9	52
Female	44	4	48
Total	87	13	100

	id	gender	ad1
1	1	Female	94
2	4	Female	92
3	5	Female	93
4	2	Male	92
5	6	Male	49
6	7	Male	53

INDEPENDENT SAMPLES T-TEST

Population Means Equal?

1 metric outcome variable
2 groups of cases

POJMENOVÁNÍ
PROMĚNNÝCH
ve statistice

Závislá proměnná

Nezávislá proměnná

Testová proměnná

Prediktor

Faktor

Grouping variable

Kovariáta

Kontrolní proměnná

Mediátor

Volba testu, tvorba modelu

- *Povaha testu*
 - A. *Porovnání četnosti*
 - B. *Porovnání skupin*
 - C. *Párové srovnání:*
 - D. *Korelační studie:*

Sex \ Handed-ness			
	Right-handed	Left-handed	Total
Male	43	9	52
Female	44	4	48
Total	87	13	100

	id	gender	ad1
1	1	Female	94
2	4	Female	92
3	5	Female	93
4	2	Male	92
5	6	Male	49
6	7	Male	53

	Before	After
1	18.6	13.5
2	20.1	14.0
3	17.4	12.9
4	19.0	13.6
5	19.1	14.5
6	19.6	13.8
7	19.4	14.4
8	18.5	14.9
9	20.0	15.3
10	21.3	15.3
11	19.6	15.1

POJMENOVÁNÍ
PROMĚNNÝCH
ve statistice

Závislá proměnná

Nezávislá proměnná

Testová proměnná

Prediktor

Faktor

Grouping variable

Kovariáta

Kontrolní proměnná

Mediátor

Volba testu, tvorba modelu

- *Povaha testu*
 - A. *Porovnání četnosti*
 - B. *Porovnání skupin*
 - C. *Párové srovnání:*
 - D. *Korelační studie:*

POJMENOVÁNÍ
PROMĚNNÝCH
ve statistice

Závislá proměnná

Nezávislá proměnná

Testová proměnná

Prediktor

Faktor

Grouping variable

Kovariáta

Kontrolní proměnná

Mediátor

Sex \ Handed-ness			
	Right-handed	Left-handed	Total
Male	43	9	52
Female	44		
Total	87		

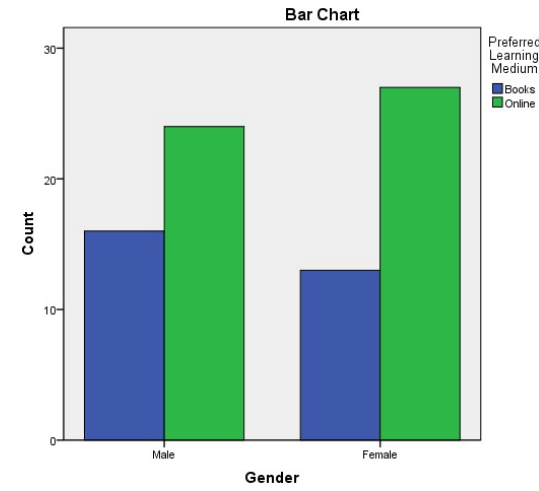
	id	gender	ad1
1	1	Female	94
2	4	Female	92
3	5	Female	93
		Male	92
		Male	49
		Male	53

	Age	Cholesterol
1	18	5.4
2	20	6.5
3	25	10.7
4	24	11.8
5	21	9.9
6	20	7.4
7	24	8.9
8	22	7.2
9	17	4.3
10	20	6.3
11		

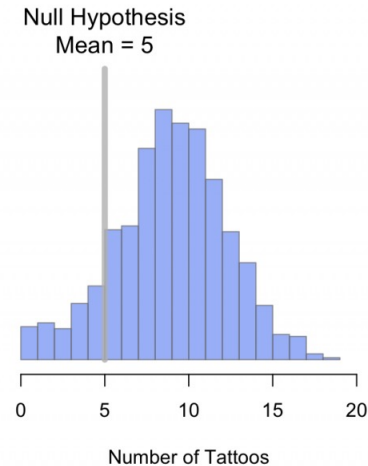
	Before	After
1	18.6	13.5
2	20.1	14.0
3	17.4	12.9
4	19.0	13.6
5	19.1	14.5
6	19.6	13.8
7	19.4	14.4
8	18.5	14.9
9	20.0	15.3
10	21.3	15.3
11	19.6	15.1

Volba testu, tvorba modelu

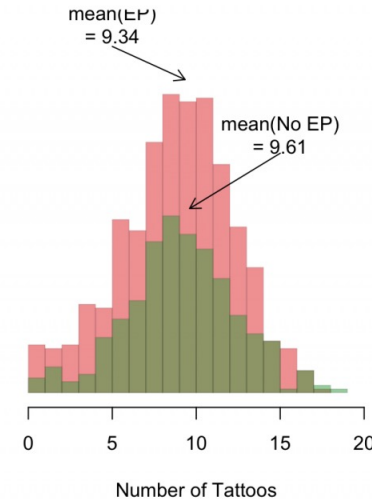
- *Povaha testu*
 - A. *Porovnání četnosti*
 - B. *Porovnání skupin*
 - C. *Párové srovnání:*
 - D. *Korelační studie:*



1-Sample t-test



2-Sample t-test



POJMENOVÁNÍ
PROMĚNNÝCH
ve statistice

Závislá proměnná

Nezávislá proměnná

Testová proměnná

Prediktor

Faktor

Grouping variable

Kovariáta

Kontrolní proměnná

Mediátor

Volba testu, tvorba modelu

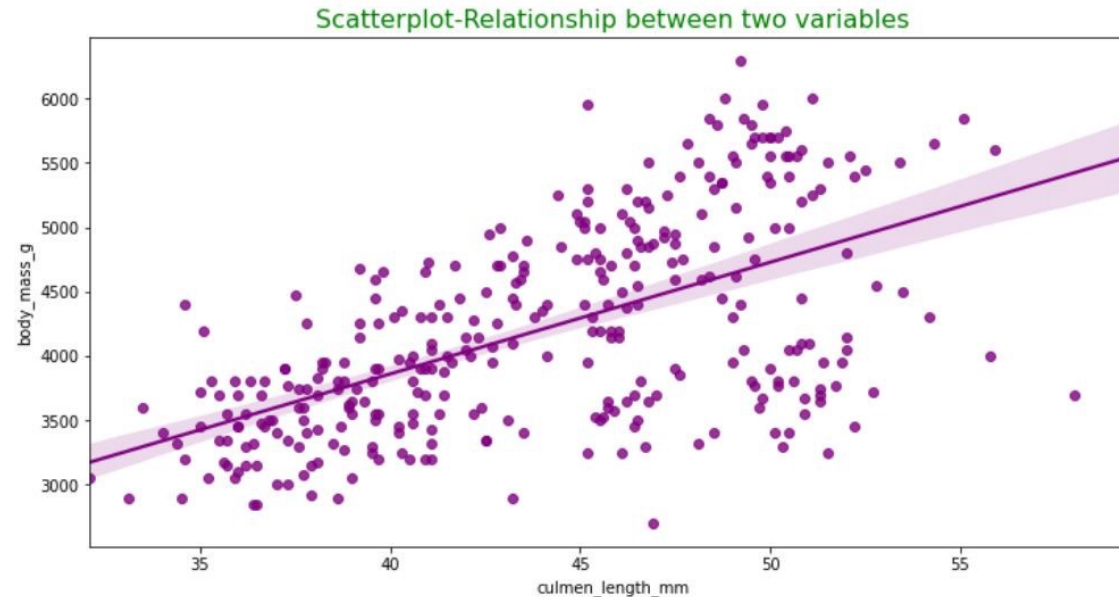
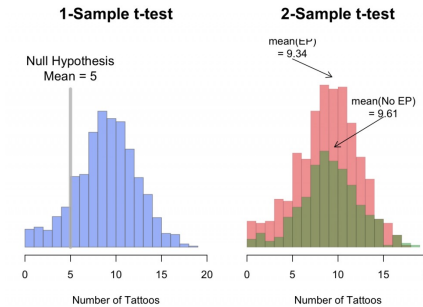
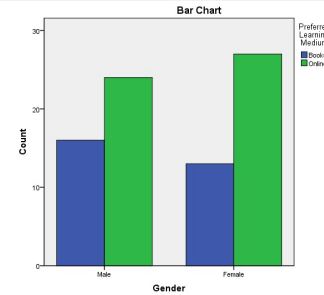
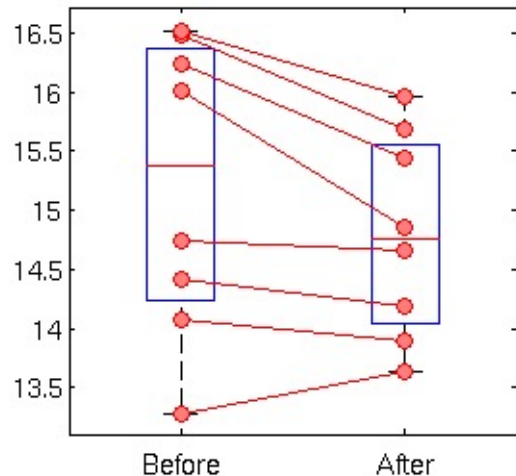
- *Povaha testu*

A. Porovnání četnosti

B. Porovnání skupin

C. Párové srovnání:

D. Korelační studie:



POJMENOVÁNÍ
PROMĚNNÝCH
ve statistice

Závislá proměnná

Nezávislá proměnná

Testová proměnná

Prediktor

Faktor

Grouping variable

Kovariáta

Kontrolní proměnná

Mediátor