

# Statistické zpracování dat

## 2.prezentace

# Analýza rozptylu

Mgr. Radmila Krkošková, Ph.D.



**SLEZSKÁ  
UNIVERZITA**

OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ



- Co je analýza rozptylu - ANOVA
- ANOVA v marketingu a managementu
- Jednofaktorová ANOVA
- Míry závislosti: determinační a korelační poměr
- Použití ANOVA v případových studiích

# ANOVA řeší 2 problémy (dvojí interpretace):

---



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

1. Ovlivňuje kvalitativní faktor kvantitativní hodnoty znaku?
2. Pochází  $k (\geq 3)$  vzorků ze stejné populace?

## Příklady:

- (1) Má věk respondentů vliv na konzumaci daného nápoje?
- (2) Ovlivňuje vzdělání respondentů jejich názor na poslance v parlamentu?

# Jednofaktorová ANOVA - postup

---



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

## Krok 1.

Uspořádání dat a výpočty

$X$  – kvalitativní znak (faktor),  $Y$  - kvantitativní znak

$$\bar{y} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}$$

Stanovení nulové hypotézy

$H_0: \mu_i = \mu_0$  pro všechna  $i = 1, 2, \dots, k$

## Krok 2. Výpočet

---



Celkový součet čtverců: 
$$S_y = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$$

Meziskupinový součet čtverců:

$$S_{y,m} = \sum_{i=1}^k n_i (\bar{y}_i - \bar{y})^2$$

Vnitroskupinový součet čtverců:

$$S_{y,v} = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

## Krok 2. Výpočet

---



Základní vztah ANOVA:

$$S_y = S_{y,m} + S_{y,v}$$

Výpočet testového kritéria:

$$F = \frac{\frac{S_{y,m}}{k-1}}{\frac{S_{y,v}}{n-k}}$$

## Krok 3. Testování

---



Jestliže platí:  $F > F_{1-\alpha}(k-1, n-k)$  pak

$H_0$  zamítáme (faktor má vliv)

jinak  $H_0$  nezamítáme (faktor nemá vliv),

přitom  $F_{1-\alpha}(df_1, df_2)$  je kritická hodnota Fisherova rozdělení

(tabelováno pro různé hodnoty  $df_1, df_2$  a  $\alpha$ ).



---

## Alternativně:

Pro hodnotu kritéria  $F$  vypočítáme  $p$ -hodnotu (signifikanci)

Jestliže platí:

$p$ -hodnota  $< \alpha$  pak  $H_0$  zamítáme (faktor má vliv)  
jinak  $H_0$  nezamítáme (faktor nemá vliv)

Předpoklad ANOVA: normálně rozdělený znak  $Y$



# Řešení příkladu v Excelu



SLEZSKÁ  
UNIVERZITA  
OBCHODNĚ PODNIKATELSKÁ  
FAKULTA V KARVINĚ

Data → Analýza dat → ANOVA: Jeden faktor...

Anova: jeden faktor

Faktor

Výběr	Počet	Součet	Průměr	Rozptyl
18	9	46	5,111111	54,11111
19	18	131	7,277778	70,80065
20	21	149	7,095238	36,89048
21	15	81	5,4	44,82857
22	8	71	8,875	222,6964

ANOVA

Zdroj variability	SS	Rozdíl	MS	F	Hodnota P	F krit
Mezi výběry	95,13097	4	23,78274	0,344165	0,847153	2,510831
Všechny výběry	4560,785	66	69,1028			
Celkem	4655,915	70				

$p$ -hodnota = 0,847 > 0,05  $\Rightarrow H_0$  nezamítáme, tzn. věk nemá na konzumaci limonád vliv!

Faktor  $X$  má  $k$  kategorií, sledovaný znak  $Y$  je kvantitativní

*Poměr korelace  $P$ :* 
$$P = \sqrt{\frac{S_{y,m}}{S_y}}$$

kde  $S_y$  - celkový součet čtverců

$$S_y = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$$

$S_{y,m}$  - meziskupinový součet čtverců

$$S_{y,m} = \sum_{i=1}^k n_i (\bar{y}_i - \bar{y})^2$$



## *Poměr determinace $P^2$*

**Čím je  $P$  bližší k 1, tím je závislost sledovaného znaku na daném faktoru silnější,  
čím je blíže k 0, tím je závislost slabší.**



Děkuji Vám za pozornost!!!