

# Statistika II

**PAA / LS 2016**

# Kontakt

Ing. Andrea Jindrová, Ph.D.

Katedra statistiky PEF ČZU

3. patro, dveře č. 427

tel.: 224 382 299

e-mail: [jindrova@pef.czu.cz](mailto:jindrova@pef.czu.cz)

# Cíl předmětu

**Prohloubit znalosti z kurzu  
Statistika I a zvládnut složitější  
statistické analýzy.**

Testy dobré shody

# Testy dobré shody

**Testy dobré shody** umožňují srovnání empirického (výběrového) rozdělení s jistým rozdělením teoretickým.

$\chi^2$  – test dobré shody

Kolmogorov – Smirnovův test

# Analýza kvalitativních znaků

# Cílem analýzy

## Otestovat závislosti mezi proměnnými

$\chi^2$  test nezávislosti

Exaktní testy: Fisherův faktoriálový test  
test Monte Carlo

## Změřit sílu závislosti

koeficienty kontingence

koeficienty asociace

# Kontingence

**Kontingence** je vztahem dvou či více **kvalitativních statistických znaků**, z nichž alespoň jeden je znakem **množným** (znaky, které mají větší počet obměn - barva očí, stupeň vzdělání).

Vztah mezi těmito znaky je zachycen v **kontingenční tabulce**.



# Kontingenční tabulka

Znak A / Znak B	$b_1$	$b_2$	.....	$b_s$	Celkem
$a_1$	$(n_{11})$	$(n_{12})$	.....	$(n_{1s})$	$n_1$
$a_2$	$(n_{21})$	$(n_{22})$	.....	$(n_{2s})$	$n_2$
.....	.....	.....	.....	.....	.....
$a_r$	$(n_{r1})$	$(n_{r2})$	.....	$(n_{rs})$	$n_r$
Celkem	$n_1$	$n_2$	.....	$n_s$	$n$

Okrajové četnosti

Celková četnost

# Asociace

**Asociace** zkoumá vztah dvou kvalitativních proměnnými, které jsou nositeli znaků **alternativních** (znaky, které nabývají jen dvou obměn - ano/ne, muž/žena).

Tento vztah zachycuje speciální typ kontingenční tabulky **2x2** - asociační tabulka, čtyřpolní tabulka.

# Kontingenční tabulka 2x2

Znak A / Znak B	b	$\beta$	
a	a	b	a+b
$\alpha$	c	d	c+d
	a+c	b+d	n

Okrajové četnosti

Celková četnost

a, b, c, d → skutečné (empirické) četnosti

# Regresní a korelační analýza

# Základní pojmy

**Závislost příčinná (kauzální)**

**Závislost pevná (funkční)**

**Volná závislost ( statistická)**

# Regrese a korelace

**Regrese** charakterizuje **průběh** závislosti mezi kvantitativními statistickými znaky pomocí matematického modelu (regresní funkce).

**Korelace** měří **těsnost** (sílu, míru, intenzitu) statistické závislosti mezi kvantitativními statistickými znaky pomocí koeficientů.

# Druhy závislostí

## Podle počtu kvantitativních znaků

- o závislost jednoduchá
- o závislost vícenásobná

# Druhy závislostí

## Podle typu regresní funkce

- **lineární** závislost
- **nelineární** závislost

## Podle směru změn kvan. znaků

- závislost **pozitivní** (kladná, přímá)
- závislost **negativní** (záporná, nepřímá)



# Regresní analýza

V regresní analýze obecně analyzujeme vztah mezi jednou proměnnou zvanou cílová nebo **závislá proměnná** ( $Y$ ) a několika dalšími, které nazýváme **nezávislé proměnné** ( $X$ ).

Závisle proměnná je spojena s nezávisle proměnnými **regresní funkcí**, jež obsahuje několik neznámých parametrů.

# Regresní analýza

## Základní úkoly regresní analýzy

- získání statistických odhadů neznámých parametrů regresní funkce na základě výběru
- testování hypotéz o těchto parametrech
- ověřování předpokladů regresního modelu

# Korelační analýza

**Korelace** obecně označuje míru stupně (sílu) závislosti dvou proměnných.

**Měření těsnosti - síly** závislosti - spočívá ve zjištění, jak těsně se jednotlivé skutečně napozorované hodnoty přibližují k regresní čáře, která vystihuje průběh závislosti.