

**Zpracování**



# Obsah

Předmět v 10 bodech	03	Indexní analýza	135
Přednášející v 10 bodech	04	Analýza časových řad	161
Přehled přednášek	05	Zdroje	186
Úvod do předmětu	07		
Kombinatorika	09		
Teorie pravděpodobnosti	33		
Základní statistické pojmy	50		
Popisná statistika	68		

# Předmět v 10 bodech

3

K rychlejší orientaci při studiu poslouží pevná struktura všech kapitol, které obsahují příklady s podrobným a názorným řešením.

4

Prezentace je členěna do dvou základních částí. V první z nich je věnován prostor kombinatorice a výpočtu pravděpodobnosti, v druhé části studenti proniknou do problematiky matematické statistiky.

7

Snahou autorky je postupovat od věcí jednoduchých ke složitějším při zachování jednotlivých návazností.

8

K předmětu Exaktní myšlení jsou studentům k dispozici skripta s názvem Metody zpracování dat v ekonomické praxi.

1

Každý vysokoškolsky vzdělaný manažer by měl zvládnout popis datových souborů prostřednictvím souhrnných charakteristik.

2

Dle zkušeností lektorky vyvolává slovo statistika, matematika nebo pravděpodobnost u studentů určitou nechuť. Ta je však způsobena tím, že si studenti nedovedou pod jednotlivými pojmy představit nic konkrétního.

5

Ambicí předmětu je rozvoj základních výpočetních postupů z oblasti popisné statistiky, základních znalostí z teorie pravděpodobnosti, rozvoj znalostí při sběru dat, jejich zpracování a elementární analýze při jejich vyhodnocování především v oblasti kvantitativních dat.

6

Prezentace poskytne čtenářům dostatečný teoretický základ pro studium pokročilých metod matematické statistiky. Důraz je kladen taktéž na upevnění početních dovedností.

9

Závěrem by lektorka ráda podotkla, že uvítá jakékoli připomínky, poznámky a názory.

10

Huff (2013) tuto disciplínu výstižně shrnuje následovně: „*Tajemná řeč statistiky, obzvláště přitažlivá v době holdující průkazným tvrzením, je využívána k vytváření senzací, nafukování, matení a k přehnanému zjednodušování. Statistické metody a pojmy jsou nezbytné při popisování hromadných dat o ekonomických trendech. Ale bez autorů, kteří užívají slova poctivě a bez čtenářů, kteří chápou jejich význam, z toho nemůže být nic jiného než významový nesmysl.*“

# Přednášející Pavla Vrabcová v 10 bodech

3

Pavla Vrabcová je členkou Finanční akademie Zlaté koruny, České podnikatelské rady pro udržitelný rozvoj a také členkou odborné sekce Rady kvality pro Společenskou odpovědnost a udržitelný rozvoj.

7

Je hrdou držitelkou ocenění Pedagog roku 2017, 2022 za vynikající pedagogickou činnost.

4

Je tajemnicí Platformy pro bioekonomiku České republiky a v této vědecké oblasti publikuje vědecké články a učební texty.

8

Odvedla více než 120 úspěšných bakalářských a diplomových prací.

1

Absolventka magisterského studijního oboru Pojišťovnictví a stejnojmenného doktorského studijního oboru na Ekonomické fakultě Technické univerzity v Liberci. Habilitována v roce 2022 docentkou v oboru Podniková ekonomika a management.

5

Pavla Vrabcová je členkou Odboru ekonomiky, řízení, sociologie a informatiky v rámci České akademie zemědělských věd a podílí se na přípravě vybraných strategických dokumentů.

9

Pedagogická činnost:  
Česká zemědělská univerzita v Praze: Lesnická bioekonomika, Forestry enterprise economics, Řízení jakosti, Kvalita výroby  
Technická univerzita v Liberci: Základy pojišťovnictví, Pojištění osob, Penzijní fondy

2

V roce 2015 úspěšně dokončila Kurz vysokoškolské pedagogiky v rámci celoživotního vzdělávání.

6

Pavla Vrabcová má na svém kontě více než 100 odborných recenzovaných článků se zahraničními ohlasy a 6 učebních textů.  
Je členkou řady redakčních rad odborných časopisů.

10

V současné době je součástí týmu, který řeší projekt s názvem Empowering the Central and Eastern European Countries and beyond to Develop Bioeconomy Strategies and Action Plans.

# Přehled přednášek

1. Úvod
2. Kombinatorika
3. Teorie pravděpodobnosti
4. Základní statistické pojmy
5. Popisná statistika + excel
6. Indexní analýza
7. Časové řady

# Úvod do předmětu



Termín statistika je odvozen od latinského „status“, což v latině znamená „stav“ a ve slovním spojení „status rei republice“ je to „stav věci veřejné“, neboli stát.

Od tohoto významu vzniklo v 16. a 17. století italské slovo „statistica“ pro označení souhrnu znalostí o státních záležitostech.

Tento termín se pak rozšířil v podobném významu i mezinárodně.

Pojem statistika se používá ve třech významech pro:

- 1. vyjádření souhrnu dat o hromadných jevech,**
- 2. činnost směřující k získávání statistických dat, jejich uspořádání a zpracování a následnou prezentaci,**
- 3. vědu, jejímž cílem je zkoumání zákonitostí hromadných jevů a kterou tvoří metodologie zjišťování, zpracování a analýzy dat.**

# Úvod do předmětu

Pro statistiku chápanou jako věda, je charakteristické:

- Jejím předmětem jsou soubory **hromadného pozorování** zkoumaných jevů.
- Znamená to, že statistiku nezajímá konkrétní jedinec (předmět, objekt, událost) sám o sobě, ale jen jako součást souboru jedinců.
- Zkoumané poznatky o **hromadných jevech** vyjadřuje statistickými daty.

## Hromadný jev

- výskyt ve velkém počtu,
- týká se velkého počtu jedinců nebo prvků,
- jev se v čase opakuje.



# Kombinatorika

- Část matematiky zabývající se kolekcemi prvků množin s definovanou vnitřní strukturou.
- Zkoumá skupiny (podmnožiny) prvků vybraných z jisté základní množiny.
- Podle toho, zda se prvky v jednotlivých skupinách **mohou či nemohou opakovat**, rozdělujeme skupiny prvků na **skupiny s opakováním a skupiny bez opakování**.
- Skupiny, kde se prvky **nemohou opakovat**, si lze tedy představit tak, že prvky, které vybíráme ze základní skupiny do ní nevracíme zpět a nemůžeme je tedy použít při dalším výběru.
- Naopak skupiny, kde se prvky **mohou opakovat**, vznikají tak, že vybrané prvky vracíme do základní skupiny a v dalším výběru je můžeme znovu použít.

# Faktoriál

Pokud uvažujeme faktoriál určitého čísla, je možné jej rozepsat na součin všech jeho členů:

$$0! = 1,$$

$$1! = 1,$$

$$2! = 2 \cdot 1 = 2,$$

$$3! = 3 \cdot 2 \cdot 1 = 6,$$

$$4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24,$$

$$10! = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 3\,628\,800.$$



## Příklad



Vypočítejte  $\frac{15!}{13!}$ .

## Řešení:

U prvního příkladu je zřejmé, že můžeme vypočítat každý faktoriál zvlášť a poté zlomek

zkrátit. Výraz  $\frac{15!}{13!}$  lze zjednodušit následovně:

$$\frac{15!}{13!} = \frac{15 \cdot 14 \cdot 13!}{13!} = 15 \cdot 14 = 210.$$

Odpověď: Uvedený výraz je po vykrácení  $13!$  roven 210.

## Příklad

Zjednodušte výraz v  $N$ :



$$\frac{(n + 4)!}{(n + 1)!}$$

Definiční obor:

$$D(f) = \{n; n \geq -1; n \in N\}.$$

## Řešení:

Výraz  $\frac{(n+4)!}{(n+1)!}$  lze zjednodušit následovně:

$$\frac{(n + 4)!}{(n + 1)!} = \frac{(n + 4) \cdot (n + 3) \cdot (n + 2) \cdot (n + 1)!}{(n + 1)!} = (n + 4) \cdot (n + 3) \cdot (n + 2) = n^3 + 9n^2 + 26n + 24.$$

Odpověď: Zjednodušíme faktoriály tak, aby mohly být vykráceny, výraz se rovná  $n^3 + 9n^2 + 26n + 24$ .



# Kombinační číslo

## Definice

Pro všechna celá nezáporná čísla  $n, k$ , kde  $k \leq n$ , definujeme kombinační číslo  $\binom{n}{k}$ . Čteme  $n$  nad  $k$ .

Pro  $\binom{n}{k}$  platí následující vztah

$$\binom{n}{k} = \begin{cases} \frac{n!}{k!(n-k)!} & \text{pro } n \geq k \geq 0; \\ 0 & \text{jinak.} \end{cases}$$


Dále platí rovnost:

$$\binom{0}{0} = \binom{n}{0} = \binom{n}{n} = 1.$$





## Příklad

Vypočítejte výraz:

$$\binom{26}{2}.$$

## Řešení:

Výraz lze zjednodušit následovně dle vztahu

$$\binom{26}{2} = \frac{26!}{2!(26-2)!} = \frac{26!}{2! \cdot 24!} = \frac{26 \cdot 25 \cdot 24!}{2! \cdot 24!} = \frac{26 \cdot 25}{2 \cdot 1} = 325.$$

Odpověď: Uvedený výraz je roven 325.

# Kombinatorika

- Tři základní způsoby výběru: **variace, permutace, kombinace.**
- Předpokládané znalosti: množiny, faktoriál, kombinační číslo



# Variace

## Definice

### Variace bez opakování

$k$ -členná variace z  $n$  prvků je **uspořádaná**  $k$ -tice sestavená z těchto prvků tak, že se v ní každý prvek vyskytuje nejvýše jednou.

Počet variací  $k$ -té třídy z  $n$  prvků **bez opakování** lze vypočítat jako

$$V_k(n) = \frac{n!}{(n-k)!}$$



Tato definice bude názorně vysvětlena na následujícím příkladu. V tomto případě se předpokládá, že se prvky v  $k$ -tici **neopakují**.



## Příklad

Je dána množina  $M = \{1, 2, 3, 4, 5, 6\}$ . Z prvků této množiny  $M$  vytvořte dvojčíferná čísla. Zároveň je dáno, že se prvky v dvojici nebudou opakovat.

## Řešení:

Vzhledem k tomu, že se prvky v dvojici nebudou opakovat a jedná se o uspořádanou  $k$ -tici, lze dosadit do vztahu:

$$V_2(6) = \frac{6!}{(6-2)!} = \frac{6!}{4!} = \frac{6 \cdot 5 \cdot 4!}{4!} = 30.$$



## Příklad

Na závodě startuje 7 biatlonistů. Kolika způsoby mohou být obsazeny 3 stupně vítězů (zlatá, stříbrná a bronzová medaile)?

## Řešení:

V tomto případě je evidentní, že se prvky v  $k$ -tici nebudou opakovat (není uvažován případ, že by jeden biatlonista obsadil 2 nebo 3 místa najednou). Zároveň **záleží na pořadí prvků**, proto se jedná o variaci bez opakování a lze dosadit do vztahu:

$$V_3(7) = \frac{7!}{(7-3)!} = \frac{7!}{4!} = \frac{7 \cdot 6 \cdot 5 \cdot 4!}{4!} = 7 \cdot 6 \cdot 5 = 210.$$

Odpověď: Existuje 210 způsobů, jakými může 7 sportovců obsadit stupně vítězů.



## Definice

### Variace s opakováním

$k$ -členná variace s opakováním z  $n$  prvků je uspořádaná  $k$ -tice sestavená z těchto prvků tak, že se v ní každý prvek vyskytuje nejvýše  $k$ -krát.

Počet  $V'(k, n)$  všech  $k$ -členných variací s opakováním z  $n$  prvků je roven vztahu:

$$V'(k, n) = n^k.$$



Následující příklad bude velmi podobný jako příklad uvedený na začátku výkladu variací bez opakování.



## Příklad

Je dána množina  $M = \{1, 2, 3, 4, 5, 6\}$ . Z prvků této množiny  $M$  vytvořte dvojčíferná čísla. Zároveň je dáno, že se prvky v dvojici mohou opakovat.

## Řešení:

V tomto případě záleží na pořadí první v  $k$ -tici a prvky se mohou opakovat, proto lze dosadit do vztahu:

$$V'(2, 6) = 6^2 = 36.$$

Odpověď: Existuje 36 dvojčíferných čísel, která lze sestavit z uvedené množiny  $M$ . Všimněte si zároveň, že je oproti předchozímu příkladu uvažováno 6 dalších možností: 11, 22, 33, 44, 55, 66.



# Permutace

## Permutace bez opakování

Permutace z  $n$  prvků je uspořádaná  $n$ -tice sestavená z těchto prvků tak, že se v ní každý prvek vyskytuje právě jednou.

Počet  $P(n)$  všech permutací z  $n$  prvků je dán vztahem

$$P(n) = n!$$





## Příklad

Určete, kolika způsoby může nastoupit 5 vojáků při nástupu na ranní rozcvičku.

## Řešení:

Uvedený příklad odpovídá definici permutace bez opakování (žádný z vojáků se nemůže objevit v řadě víckrát) a do řady nastupují všichni (nevybíráme žádnou podmnožinu  $k$ ). Lze tedy dosadit do vztahu:

$$P(5) = 5! = 120.$$

Odpověď: 5 vojáků může nastoupit do řady celkem 120 způsoby.



## Definice

### Permutace s opakováním

Permutace s opakováním z  $n$  prvků je uspořádaná  $k$ -tice sestavená z těchto prvků tak, že se v ní každý prvek vyskytuje aspoň jednou.

Počet  $P'(k_1, k_2, \dots, k_n)$  permutací z  $n$  prvků, v nichž se jednotlivé prvky opakují  $k_1, k_2, \dots, k_n$ -krát, je dán vztahem:



$$P'(k_1, k_2, \dots, k_n) = \frac{(k_1 + k_2 + \dots + k_n)!}{k_1! k_2! \dots k_n!}.$$



## Příklad

Rozhodněte, kolika způsoby je možné srovnat do řady 2 červené, 2 oranžové a 2 zelené košile.

## Řešení:

Do řady jsou rovnány všechny košile, nevybíráme podmnožinu prvků. Musí se tedy jednat o permutace. Vzhledem k tomu, že se prvky mohou opakovat (vyskytují se zde 2 košile stejné barvy), jedná se o permutace s opakováním.

Pro zjištění počtu způsobů seřazení košil do řady lze dosadit do vzorce pro počet permutací ze tří prvků s opakováním  $k_1 = k_2 = k_3 = 2$ :

$$P'(2, 2, 2) = \frac{(2 + 2 + 2)!}{2! 2! 2!} = \frac{6!}{8} = \frac{720}{8} = 90.$$



# Kombinace

## Definice

### Kombinace bez opakování

$k$ -členná kombinace z  $n$  prvků je **neuspořádaná**  $k$ -tice sestavená z těchto prvků tak, že se v ní každý prvek vyskytuje nejvýše jednou.

Počet  $K(k, n)$  všech  $k$ -členných kombinací z  $n$  prvků je dán vztahem:



$$K(k, n) = \binom{n}{k} = \frac{n!}{k! (n - k)!}$$



## Příklad

Najděte všechny kombinace druhé třídy z množiny  $M = \{1, 2, 3, 4, 5, 6\}$ . Uvažujte dále, že se prvky v  $k$ -tici nebudou opakovat.

## Řešení:

Počet všech kombinací bez opakování lze zjistit dosazením do vztahu:

$$K(2,6) = \binom{6}{2} = \frac{6!}{2!(6-2)!} = \frac{6!}{2!4!} = \frac{6 \cdot 5 \cdot 4!}{2!4!} = \frac{6 \cdot 5}{2} = 15.$$

Odpověď: Počet všech možných kombinací bez opakování je 15.



## Definice

### Kombinace s opakováním

$k$ -členná kombinace s opakováním z  $n$  prvků je neuspořádaná  $k$ -tice sestavená z těchto prvků tak, že se v ní každý vyskytuje nejvýše  $k$ -krát.

Počet  $K'(k, n)$  všech  $k$ -členných kombinací s opakováním z  $n$  prvků je dán vztahem:



$$K'(k, n) = \binom{n+k-1}{n-1} = \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}.$$



## Příklad

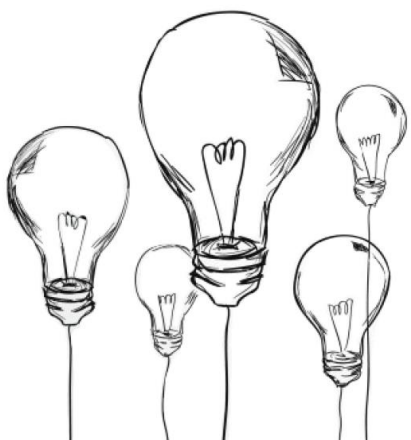
Najděte všechny kombinace druhé třídy z množiny  $M = \{1, 2, 3, 4, 5, 6\}$ . Uvažujte dále, že se prvky v  $k$ -tici budou opakovat.

## Řešení:

Pro zjištění všech kombinací s opakováním lze dosadit do vztahu:

$$K'(2,6) = \binom{6+2-1}{2} = \frac{(6+2-1)!}{2!(7-2)!} = \frac{7!}{2!5!} = \frac{7 \cdot 6 \cdot 5!}{2!5!} = \frac{7 \cdot 6}{2} = 21.$$

Odpověď: Existuje celkem 21 možností, kterými je možno sestavit kombinace druhé třídy s opakováním. Porovnejte výsledek s předchozím příkladem, v tomto případě je nutné ještě navíc uvažovat dalších 6 možností –  $(1,1)$ ,  $(2,2)$ ,  $(3,3)$ ,  $(4,4)$ ,  $(5,5)$ ,  $(6,6)$ .



Uspořádané výběry		
Bez opakování	Variace	$V(k, n) = \frac{n!}{(n - k)!}$
	Permutace	$P(n) = n!$
S opakováním	Variace	$V'(k, n) = n^k$
	Permutace	$P'(k_1, k_2, \dots, k_n)$ $= \frac{(k_1 + k_2 + \dots + k_n)!}{k_1! k_2! \dots k_n!}$
Neuspořádané výběry		
Bez opakování	Kombinace	$K(k, n) = \binom{n}{k} = \frac{n!}{k! (n - k)!}$
S opakováním	Kombinace	$K'(k, n) = \binom{n + k - 1}{n - 1}$ $= \binom{n + k - 1}{k}$ $= \frac{(n + k - 1)!}{k! (n - 1)!}$



1. Do školního výboru zvolili 7 žáků. Kolika způsoby se z nich dá vybrat předseda, místopředseda, pokladník a tajemník? (840)
2. 8 studentů si slíbilo, že si z prázdninových cest vzájemně vymění pohlednice. Kolik pohlednic poslali? (56)
3. V obchodě mají tři druhy sirupu: J, P, M. Určete počet všech možností nákupu 5 lahví. (21)
4. Kolika způsoby je možné srovnat do řady 15 svatebčanů na fotografii? (15!)

# Přestávka

# 2 Teorie pravděpodobnosti





# Základní pojmy – definice

- Teorie pravděpodobnosti vychází ze studia náhodných pokusů.

## Náhodný pokus

- Proces, který při opakování dává ze stejných podmínek rozdílné výsledky.
- Výsledek pokusu není předem znám (výsledek není jednoznačně určen jeho podmínkami), ale je předem dána množina možných výsledků.
- Každý možný výsledek náhodného pokusu nazýváme **elementárním náhodným jevem** (značíme  $E_1, E_2, \dots, E_n$ ).
- Všechny elementární jevy tvoří tzv. **základní prostor elementárních jevů**; značí se  $\Omega$ .
- Každá podmnožina základního prostoru  $\Omega$  se nazývá **náhodný jev** (značíme  $A, B, \dots$ ), přičemž prázdná podmnožina se nazývá **jev nemožný**, označujeme  $\emptyset$  a celý základní prostor **jev jistý**, označujeme  $I$ .

**Pravděpodobnost jevu  $A$**  je reálné číslo  $P(A)$ , pro něž platí:

1.  $P(A) \geq 0$  . . . axiom nezápornosti
2.  $P(I) = 1$  . . . axiom jednotky





# Ukázkový příklad

Klasickým příkladem náhodného pokusu je hod hrací kostkou, tedy:

Náhodný pokus . . .	Hod hrací kostkou
Elementární jevy . . .	"padne 1" ... $E_1$ "padne 2" ... $E_2$ . . . "padne 6" ... $E_6$

Jevy  $E_1, E_2, \dots, E_6$  vymezují základní prostor  $\Omega$ .

V tomto základním prostoru mohou být například následující jevy:

náhodný jev  $A$  . . . "padne liché číslo" . . .  $A = E_1 + E_3 + E_5$

náhodný jev  $B$  . . . "padne číslo  $\geq 4$ " . . .  $A = E_4 + E_5 + E_6$

jev nemožný . . . . "padne číslo  $> 6$ "

jev jistý . . . . . "padne číslo  $< 7$ "

neslučitelné jevy. . . "padne sudé číslo", "padne liché číslo"



# Operace s jevy

## Součet jevů $A, B$

- jev, který nastane právě tehdy, když nastane alespoň jeden z jevů  $A, B$ . Zavádíme označení  $A+B$  nebo množinově.

## Součin jevů $A, B$

- jev, který nastane právě tehdy, když nastanou oba jevy současně. Zavádíme označení  $A.B$  nebo množinově  $A \cap B$ .

## Rozdíl jevů $A, B$

- jev, který nastane právě tehdy, když nastane jev  $A$  a nenastane jev  $B$ . Zavádíme označení  $A - B$ .
- Jev nazýváme **jevem opačným** k jevu  $A$ , je-li  $\Omega - A$ .
- Náhodné jevy se nazývají **neslučitelné** (disjunktní), jestliže platí  $A.B = \emptyset$ .



# Klasická definice pravděpodobnosti

Nechť je dáno  $n$  elementárních jevů  $E_1, E_2, \dots, E_n$ , které tvoří úplný systém neslučitelných jevů a jsou **stejně možné**. Rozkládá-li se jev  $A$  na  $m$  ( $m \leq n$ ) elementárních jevů z tohoto systému, pak pravděpodobnost jevu  $A$  je reálné číslo

$$P(A) = \frac{\text{počet všech příznivých případů}}{\text{počet všech možných případů}} = \frac{m}{n}.$$



Všechny elementární jevy se obvykle označují jako všechny možné případy.

Všechny elementární jevy, na které se rozkládá jev  $A$ , se nazývají všechny příznivé případy.

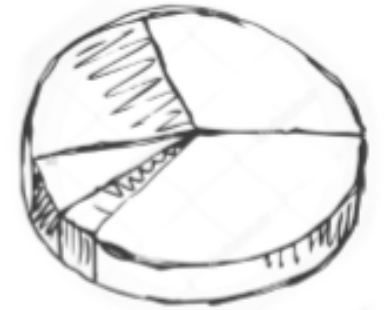


## Příklad

Při hodu kostkou určete pravděpodobnost jevů:

- a) jev  $A$ : padne číslo 5
- b) jev  $B$ : padne číslo  $\leq 3$

## Řešení



a) Dle klasické definice pravděpodobnosti lze jednoduše dosadit do vztahu:

$$P(A) = \frac{\text{počet příznivých případů}}{\text{počet všech možných případů}} = \frac{1}{6}.$$

b) Dle klasické definice pravděpodobnosti platí, že počet příznivých případů je 3 (padne 1, 2, 3), jmenovatel zůstává stejný:

$$P(B) = \frac{3}{6} = \frac{1}{2}.$$



## Příklad

V klobouku se nachází  $a$  bílých a  $b$  černých králíků. Najděte pravděpodobnost následujících jevů:

a) Náhodný jev A: Vytažený králík bude bílý.

b) Náhodný jev B: Vytáhne se jednou bílý králík, odloží se, pak se vytáhne druhý. Jaká je pravděpodobnost, že druhý vytažený bude opět bílý?

c) Náhodný jev C: Vytáhnou se všichni králíci kromě jednoho. Jaká je pravděpodobnost, že poslední vytažený bude bílý?

d) Náhodný jev D: Najděte pravděpodobnost, že postupně vytáhneme 1 bílého králíka, 1 černého králíka, 1 bílého králíka, když je nevracíme zpět do klobouku.

## Řešení

$$P(A) = \frac{a}{a+b};$$

$$P(B) = \frac{a-1}{a+b-1};$$

$$P(C) = \frac{a}{a+b};$$

$$P(D) = \frac{a}{a+b} \cdot \frac{b}{a+b-1} \cdot \frac{a-1}{a+b-2}.$$





## Příklad

Karetní hru tvoří 32 karet. Po vytažení a vrácení karty do balíčku se balíček celý promíchá a znovu se vytáhne karta. Určete pravděpodobnost následujících jevů:

- a) Vytažená karta je srdcové eso;
- b) Jako první karta je vytaženo srdce, druhá karta je figura (svršek, spodek, král, eso);
- c) Jako první karta je vytaženo srdce, druhá karta je žaludová. Uvažujte zároveň, že jsme první kartu do balíčku **nevrátili**;
- d) 2 vytažené karty jsou esa (1. vytažené eso se do balíčku nevrátilo zpět).

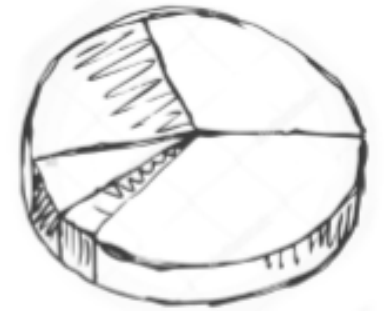
## Řešení

$$P(A) = \frac{1}{32};$$

$$P(B) = \frac{8}{32} \cdot \frac{16}{32} = \frac{1}{8};$$

$$P(C) = \frac{8}{32} \cdot \frac{8}{31} = \frac{2}{31};$$

$$P(D) = \frac{4}{32} \cdot \frac{3}{31} = \frac{3}{248}.$$





## **Příklad**

Uvažujte, že na 8 strojírenských výrobců připadají 2 vadné.

Z těchto 8 výrobců náhodně vybereme 3.

Určete pravděpodobnost, že se ve 3 náhodně vybraných kusech bude vyskytovat 1 dobrý a 2 vadné.



## Řešení:

Uvedenou úlohu budeme opět řešit prostřednictvím klasické definice pravděpodobnosti. Jak je známo, ve jmenovateli se nachází počet všech možných případů.

To lze vypočítat prostřednictvím kombinací bez opakování, viz vzorec, konkrétně počet všech kombinací  $K(3, 8)$ . V čitateli je nutné uvažovat dva případy, a to výběr 1 dobrého kusu ze 6 a výběr 2 vadných z celkového počtu vadných kusů (2). Výše uvedené lze vyjádřit následovně:

$$P(A) = \frac{\binom{6}{1}\binom{2}{2}}{\binom{8}{3}} = \frac{6 \cdot 1}{\frac{8!}{3!5!}} = \frac{6}{\frac{8 \cdot 7 \cdot 6 \cdot 5!}{3!5!}} = \frac{6}{8 \cdot 7} = \frac{6}{56} = \frac{3}{28}$$

Odověď: Pravděpodobnost, že ze 3 náhodně vybraných kusů bude 1 dobrý a 2 vadné, je  $3/28$  (cca 10,7 %).



## Příklad

Uvažujte, že na 10 výrobků připadají 3 vadné. Náhodně vybereme skupinu 5 výrobků. Určete pravděpodobnost, že v této skupině bude:

- a) 1 vadný;
- b) 2 vadné;
- c) maximálně 2 vadné.

## Řešení:

a) Při výpočtu budeme postupovat stejným způsobem jako u předchozího příkladu při využití kombinací bez opakování:

$$P(A) = \frac{\binom{3}{1}\binom{7}{4}}{\binom{10}{5}} = 0,4167.$$

b) I v tomto případě budeme postupovat stejným způsobem jako u předchozího příkladu při využití kombinací bez opakování:

$$P(B) = \frac{\binom{3}{2}\binom{7}{3}}{\binom{10}{5}} = 0,4167.$$

c) Pokud budeme vytahovat maximálně 2 vadné kusy, nastávají tyto možnosti: nevytáhne se žádný vadný kus, nebo se vytáhne jeden vadný kus, nebo se vytáhnou 2 vadné kusy.

$$P(C) = \frac{\binom{7}{5}\binom{3}{0} + \binom{3}{2}\binom{7}{3} + \binom{3}{1}\binom{7}{4}}{\binom{10}{5}} = 0,9167.$$





## Příklad s řešením k procvičení

Ve sportovním týmu je 21 mužů a 9 žen. Jaká je pravděpodobnost, že při výběru 3členné skupiny v ní budou:

a) právě 3 muži  $(\frac{1\ 330}{4\ 060} = 0,3276),$

b) nejvýše 2 muži  $(\frac{2\ 730}{4\ 060} = 0,6724),$

c) alespoň 2 muži  $(\frac{3\ 220}{4\ 060} = 0,7931).$

# Přestávka

# 3 Základní statistické pojmy





# Hromadný jev – definice

- Předmětem vědního oboru statistika jsou soubory hromadného pozorování zkoumaných jevů.
- Znamená to, že statistiku nezajímá konkrétní jedinec sám o sobě, ale jen jako součást souboru těchto jedinců.
- **Hromadný jev** je takový, který lze za daného systému podmínek libovolně krát opakovat nebo který lze pozorovat na hromadně se vyskytujících předmětech téhož druhu.



# Statistická jednotka, statistický soubor, znak

- **Statistická jednotka** je prvek souboru, u něhož jsou sledovány určité vlastnosti.
- **Statistický soubor** je soubor vytvořený statistickými jednotkami, u nichž se sledují statistické znaky.
- Vlastnosti statistických souborů sleduje vědní disciplína statistika prostřednictvím vlastností statistických jednotek, které postihuje **statistickými znaky**.

Obecně lze rozlišit dva typy souborů, a to:

- základní soubor (soubor všech statistických jednotek) a
- výběrový soubor obsahující pouze část jednotek základního souboru.

# Rozsah souboru, hodnoty statistického znaku



- **Rozsah souboru** je počet statistických jednotek ve statistickém souboru. Bývá označován písmenem  $n$ .
- Pokud jsou hodnoty statistického znaku vyjádřeny čísly, mluvíme o znacích **kvantitativních** (např. tělesná výška, hmotnost, počet obyvatel).
- V druhém případě mluvíme o znacích **kvalitativních**.
- Jestliže existuje statistický znak udávající hmotnost výrobku, pak hodnoty tohoto statistického znaku mohou být 10 kg, 20 kg, ...



# Klasifikace statistických proměnných

Podle hlediska typů vztahů mezi hodnotami rozlišujeme proměnné:

- nominální,
- ordinální,
- intervalové a
- poměrové.



# Klasifikace proměnných

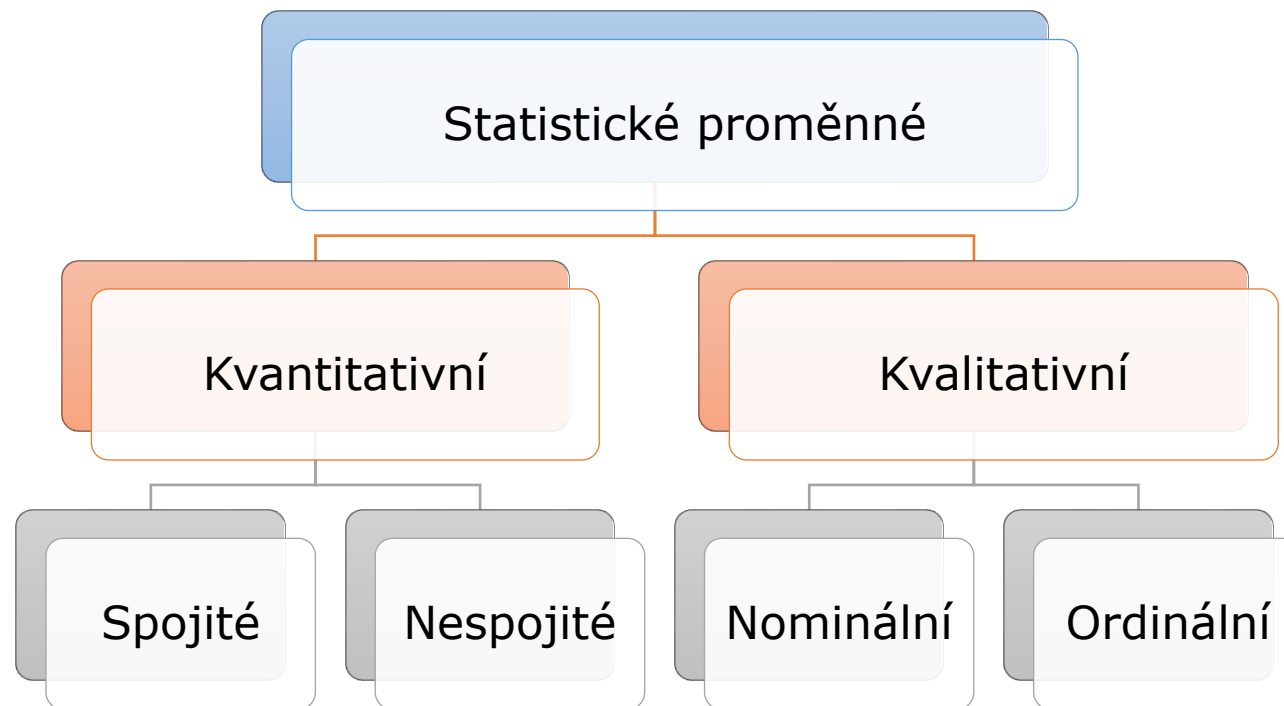


Schéma 3.1 Klasifikace proměnných



# Nominální proměnné

- Název proměnné je odvozen z latinského slova nomen ve smyslu jméno, název, pojmenování.
- Nominální proměnná je kvalitativní proměnná, o jejíchž dvou hodnotách můžeme pouze říci, zda jsou stejné či různé (lze interpretovat rovnost, případně nerovnost).
- Hodnotami bývají texty či číselné kódy (např. EAN kód u výrobků). U nominální proměnné lze zjišťovat rozdělení četností, nelze (a nemá smysl) je sčítat či provádět další aritmetické operace.



# Ordinální proměnné

- Ordinální proměnná je odvozena od latinského slova ordatio – pořadí.
- Na rozdíl od nominální proměnné lze u dvou hodnot navíc určit pořadí (např. stupeň spokojenosti, klasifikace, vzdělání atd.).
- Pro statistické analýzy (s výjimkou zjišťování četností) je třeba texty převést na čísla.



# Intervalové proměnné

- Intervalová proměnná je taková, pro jejíž dvě hodnoty můžeme navíc k možnostem ordinální (pořadové) proměnné vypočítat, o kolik je jedna hodnota větší (resp. menší) než druhá.
- Jsme u nich schopni interpretovat rozdíl dvou hodnot, proto jsou také někdy nazvané jako rozdílové proměnné.
- Pro tento typ znaků jsou typické číselné hodnoty, které vyjadřují nejen totožnost a rozdílnost hodnot jako u nominálních proměnných, a nejen pořadí, jako u ordinálních proměnných, ale také vzdálenosti.
- *Příkladem může být rok nástupu do zaměstnání, počet dětí v rodině, roční příjem domácnosti atd.*



# Poměrové proměnné

- Poměrová proměnná je ta, pro jejíž dvě hodnoty můžeme navíc k možnostem intervalové proměnné vypočítat, kolikrát je jedna hodnota větší (resp. menší) než druhá.
- Mimo rozdílů jsme tedy schopni interpretovat i podíl 2 hodnot.
- Jedná se pouze o kladné hodnoty.

Podle formálního hlediska je možno dále rozlišit kvantitativní proměnné na:

- nespojité (diskrétní),
- spojité.



# Poměrové proměnné

- Nespojité proměnné nabývají jen celočíselných obměn, např.: počet dětí, počet nemocných v určité lokalitě atd.
- Speciálním případem nespojitých proměnných, nominálních a ordinálních jsou proměnné **alternativní (dichotomické)** nabývající pouze 2 hodnot, např.: ekonomicky aktivní nebo ekonomicky neaktivní, muž nebo žena, kuřák nebo nekuřák.
- Nebo se může jednat o hodnoty **vícekategoriální (množné)** nabývající více než dvou kategorií (obor studia, hodnota...).
- Spojité proměnné mohou nabývat všech hodnot v rámci reálného intervalu (tělesná hmotnost, výška, teplota, věk atd.).

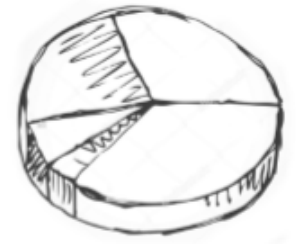


## Příklad

Vyberte z následující nabídky správný typ proměnné:

- pohlaví,
- profese,
- úroveň spokojenosti s výrobkem,
- obor studia,
- věk,
- hmotnost výrobku.

## Řešení:



- Pohlaví je znak kvalitativní, nominální, alternativní (muž, žena);
- Profese je také znak kvalitativní, nominální, nicméně nabývá více než 2 hodnot, tedy vícekategoriální;
- Úroveň spokojenosti zákazníka je ordinální proměnnou kvalitativní, která může a nemusí nabývat 2 hodnot.
- Obor studia je kvalitativní, nominální, množná proměnná.
- Věk je znak kvantitativní, intervalový.
- Hmotnost výrobku je kvantitativní, poměrová proměnná (spojitá).



# Statistické šetření

- Základem každého statistického zkoumání jsou data, která lze získat v zásadě dvěma způsoby, buď je převzít, nebo je sami zjistit.
- **Sekundární sběr dat** je výzkum zaměřený na získávání, analýzu a vyhodnocení informací, které již existují (byly sesbírány).
- **Primární sběr dat** je výzkum, který se zaměřuje na získávání, analýzu a vyhodnocování nových konkrétních informací.

# 1. Příprava statistického šetření – formulace problému



- V rámci této etapy je nutné ujasnit si základní otázky ohledně vymezení statistického souboru, jednotek, znaků a způsobu zjišťování.
- **Rozhodný okamžik** je datum, ke kterému se provádí statistické šetření – nepodceňovat.
- Na konci této etapy je zodpovězena otázka týkající se „*co, kdy kde a jak bude zjišťováno*“.
- Bude rozhodnuto, zda bude uskutečněno přímé zjišťování, rozhovor (přímý nebo telefonický) nebo dotazníkové šetření.
- Základním pravidlem dobře provedeného výběrového šetření je zásada, že **výběr musí být reprezentativní (všechny statistické jednotky musí mít stejnou šanci na zařazení do zkoumaného výběru).**



## 2. Pořizování dat

- Nejpoužívanější technikou je však dotazování – komunikace přímá nebo zprostředkovaná.
- Kvalita dotazníku významně ovlivňuje každé statistické šetření.
- To se týká i samotných formulací bez sugestivních otázek, zajištění **validity** otázek (ptát se skutečně na to, co chceme zjistit) a **reliability** (spolehlivosti) odpovědí.



# 3. Zpracování a analýza dat

- Znalost jednotlivých metod a postupů!
- Vypočítat souhrnné charakteristiky, uspořádat data do tabulky četností, vyhotovit grafické znázornění atd.
- Není-li výběrové šetření provedeno metodicky správně, je zatíženo tzv. systematickou chybou, např. nedostatečně velký nebo nereprezentativní vzorek, nesprávně formulované otázky, nesprávný způsob dotazování (viz dotazníková šetření rozesílaná pouze prostřednictvím „přátel na Facebooku“).

# 4. Vyhodnocení a publikování výsledků



- Popis ve formě souvislého textu – použitelné, sdělujeme-li výsledky získané zpracováním malých souborů.
- Statistické tabulky pro zabezpečení stručnosti a srozumitelnosti s předmětem zkoumání, doby a území, jichž se údaje týkají.
- Statistické grafy poskytující rychlejší a přehlednější představu o důležitých skutečnostech, tendencích a charakteristických rysech analyzovaných jevů.
- Grafické znázornění odhaluje v řadě údajů některé skutečnosti na první pohled, tzn. že ukáže rostoucí nebo klesající trend, jaké jsou maximální, minimální nebo nejčastější hodnoty, zda vývoj probíhá rovnoměrně nebo je ovlivněna sezónním kolísáním apod.
- Každý graf by měl obsahovat název a vysvětlivky. Z hlediska konstrukce rozeznáváme grafy spojnicové, bodové, sloupkové, výsečové, plošné atd.

# 4

## Popisná statistika



# Statistické zpracování kategoriální proměnné



- Kategoriální proměnné nabývají určitého počtu obměn.
- K popisu počtu těchto kategorií použijeme četnost.
- Absolutní četnost  $n_i$  je definována jako počet výskytu dané varianty kvalitativní proměnné.
- Relativní četnost je vyjádřena vztahem:

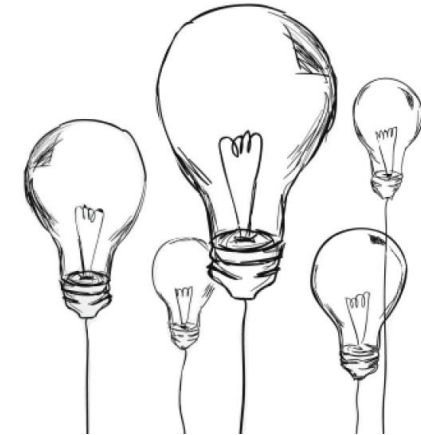
$$p_i = \frac{n_i}{n}.$$

$$p_1 + p_2 + \dots + p_k = \sum_{i=1}^k p_i = 1.$$



- Při zpracování kvalitativní proměnné je vhodné absolutní četnosti i relativní četnosti uspořádat do **tabulky rozdělení četností**.

Hodnoty $x_i$	Absolutní četnost $n_i$	Relativní četnost $p_i$
$x_1$	$n_1$	$p_1$
$x_2$	$n_2$	$p_2$
...	...	...
$x_k$	$n_k$	$p_k$
Celkem	$\sum_{i=1}^k n_i = n$	$\sum_{i=1}^k p_i = 1$



Tabulka 4.1 Statistické zpracování kategoriální proměnné

Další charakteristikou, kterou si pro popis kategoriální proměnné uvedeme, je **modus**, vykazující nejvyšší absolutní četnost.

- V případech, kdy máme velké množství obměn a tabulka je tedy velmi rozsáhlá, používáme **grafy a diagramy** pro přehlednější zobrazení výsledků (**sloupcové diagramy, histogramy a řádkové diagramy**).
- Jednotlivé obměny jsou vyznačeny sloupci, které jsou stejně široké, a výška těchto sloupců značí počet statistických jednotek.
- Obměnu, jejíž absolutní i relativní četnost je největší, nazýváme **modus**..
- Sloupce u histogramu se vzájemně dotýkají (na rozdíl od sloupcového grafu), jelikož se jedná o spojité výsledky.
- **Výsečové grafy (koláčové grafy)** zobrazují relativní četnosti – obsah zvoleného geometrického obrazce tvoří 100 % a jednotlivě vyznačené části určují relativní četnosti.





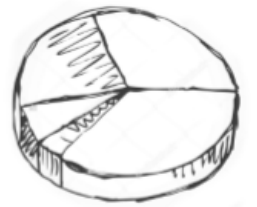
# Příklad

V níže uvedené tabulce jsou uvedeny absolutní četnosti k jednotlivým způsobům dopravy studentů do školy.

Vyčíslete modus, relativní četnosti a výsledky zakreslete graficky.

Frekvence sportovních aktivit	$n_i$
jízda na kole	15
jízda MHD	25
pešky	57
jízda na koni	1
Celkem	98

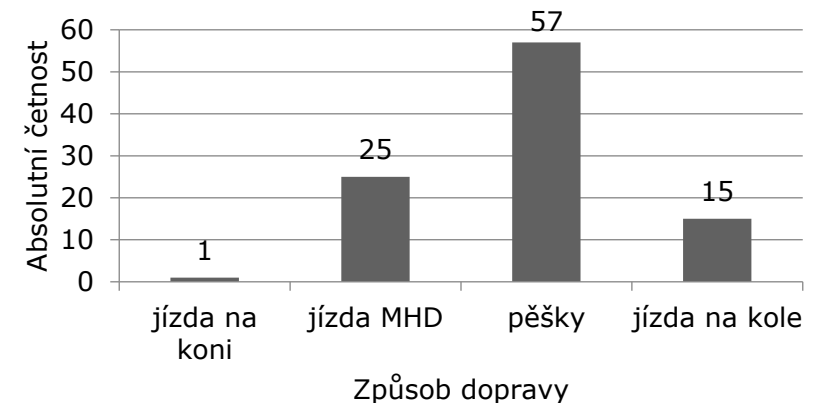
Tabulka 4.2 Zadání příkladu



## Řešení:

V tomto případě se jedná o vícekategoriální proměnnou, a to nominální. Relativní četnosti jsou vyčísleny v tabulce a modus (obměna s nejvyšší absolutní četností) je zvýrazněn tučně.

Frekvence sportovních aktivit	$n_i$	$p_i$
jízda na kole	15	15 %
jízda MHD	25	26 %
<b>pešky</b>	<b>57</b>	<b>58 %</b>
jízda na koni	1	1 %
celkem	98	100 %



Tabulka 4.3 Řešení příkladu – tabulka rozdělení četností

Graf 4.1 Řešení příkladu – sloupcový graf

# Přestávka

# Zpracování numerické proměnné



- V první řadě je nutné rozhodnout, zda se jedná o **diskrétní** či **spojitou proměnnou**.
- U diskrétní proměnné lze použít **prosté rozdělení četností** (jako v případě kategoriální proměnné).
- V případě, že proměnná obsahuje velké množství obměn, nebo se jedná o proměnnou spojitou, využíváme k jejímu zpracování **intervalové rozdělení četností**.

- Pokud využíváme **prosté rozdělení četností**, postupujeme podobně jako v případě kategoriálních proměnných s tím rozdílem, že obměny uspořádáme podle velikosti.
- Následně vypočítáme kumulativní absolutní a kumulativní relativní četnosti.
- Vzorec pro výpočet absolutní kumulativní četnosti, která udává, u kolika statistických jednotek byla zkoumaná proměnná menší nebo stejná jako  $x_i$ , je:



$$N_i = \sum_{j=1}^i n_j.$$

- Počet obměn je značen  $k$ , jednotlivé obměny označíme  $x_i$ , kde  $i = 1, 2, \dots, k$ . Rozsah statistického souboru značíme  $n$ .



- Hodnota kumulované absolutní četnosti poslední obměny  $x_k$  musí být rovna  $n$ .
- Kumulativní relativní četnost  $P_i$  udává, jaký poměr statistických jednotek má zkoumanou proměnnou rovnu nebo menší než zvolená hodnota  $x_i$ :



$$P_i = \sum_{j=1}^i p_j.$$

- Kumulativní relativní četnost vypočteme jako poměr mezi kumulativní absolutní četností a rozsahem statistického souboru:

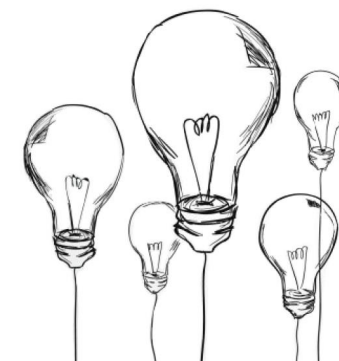


$$P_i = \frac{N_i}{n}.$$



# Tabulka prostého rozdělení četnosti

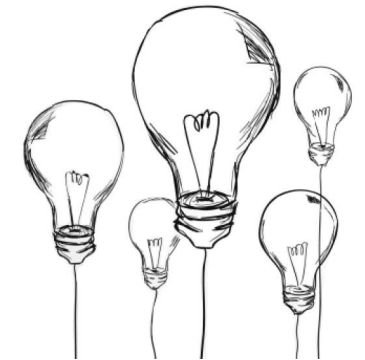
Hodnoty $x_i$	Četnost		Kumulativní četnost	
	Absolutní $n_i$	Relativní $p_i$	Absolutní $N_i$	Relativní $P_i$
$x_1$	$n_1$	$p_1$	$N_1 = n_1$	$P_1 = \frac{N_1}{n}$
$x_2$	$n_2$	$p_2$	$N_2 = N_1 + n_2$	$P_2 = \frac{N_2}{n}$
...	...	...	...	...
$x_k$	$n_k$	$p_k$	$N_k = N_{k-1} + n_k$	$P_k = \frac{N_k}{n}$
Celkem	$\sum_{i=1}^k n_i = n$	$\sum_{i=1}^k p_i = 1$		



Tabulka 4.4 Tabulka prostého rozdělení četnosti

# Intervalové rozdělení četnosti

Interval	Četnost		Kumulativní četnost	
	Absolutní $n_i$	Relativní $p_i$	Absolutní $N_i$	Relativní $P_i$
$(a_1, a_2)$	$n_1$	$p_1$	$N_1 = n_1$	$P_1 = \frac{N_1}{n}$
$(a_2, a_3)$	$n_2$	$p_2$	$N_2 = N_1 + n_2$	$P_2 = \frac{N_2}{n}$
...	...	...	...	...
$(a_k, a_{k+1})$	$n_k$	$p_k$	$N_k = N_{k-1} + n_k$	$P_k = \frac{N_k}{n} = 1$
Celkem	$\sum_{i=1}^k n_i = n$	$\sum_{i=1}^k p_i = 1$		



Tabulka 4.5 Intervalové rozdělení četnosti



## Příklad

V prodejně bot prodali během jednoho týdne 62 párů. Velikosti prodaných bot jsou uvedeny níže. Zpracujte dané údaje do tabulky četností, vypočtete také kumulativní četnosti. Sestrojte pro tato data sloupcový graf a polygon četností. Určete zároveň, kolik lidí si koupilo boty velikosti nejvýše 40.

Velikosti prodaných bot jsou následující:

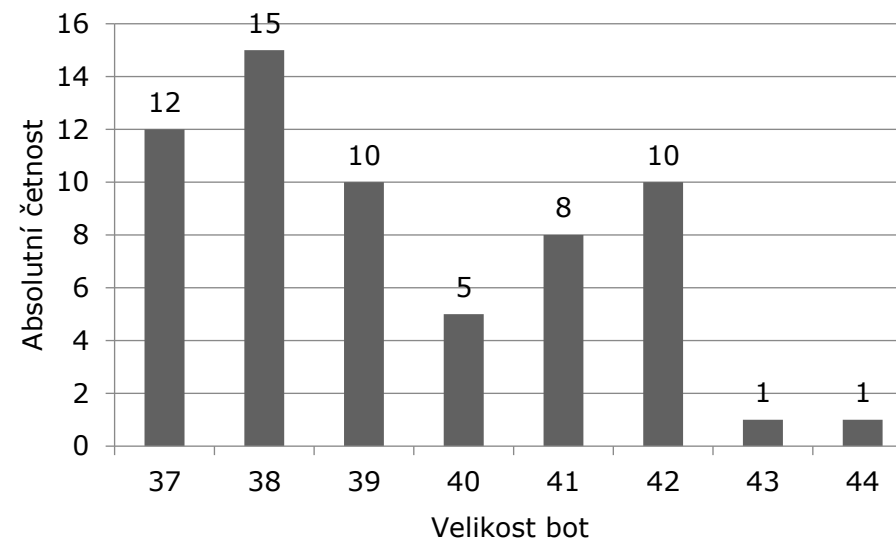
37 38 41 41 38 38 41 44 40 39 39 37 40 42 41 37 40 37 41 39 38 38  
39 42 37 43 39 39 37 38 38 41 42 41 37 41 42 37 42 37 42 37 42 38  
38 37 42 38 38 42 37 39 38 39 42 40 40 39 39 38 38 38.

# Řešení:

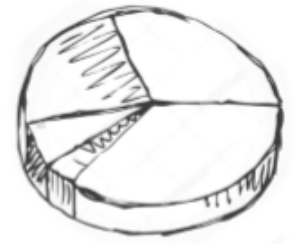


Hodnoty $x_i$	Četnost		Kumulativní četnost	
	Absolutní $n_i$	Relativní $p_i$	Absolutní $N_i$	Relativní $P_i$
37	12	0,193548	12	0,193548
38	15	0,241935	27	0,435484
39	10	0,161290	37	0,596774
40	5	0,080645	42	0,677419
41	8	0,129032	50	0,806452
42	10	0,161290	60	0,967742
43	1	0,016129	61	0,983871
44	1	0,016129	62	1
Celkem	62	1		

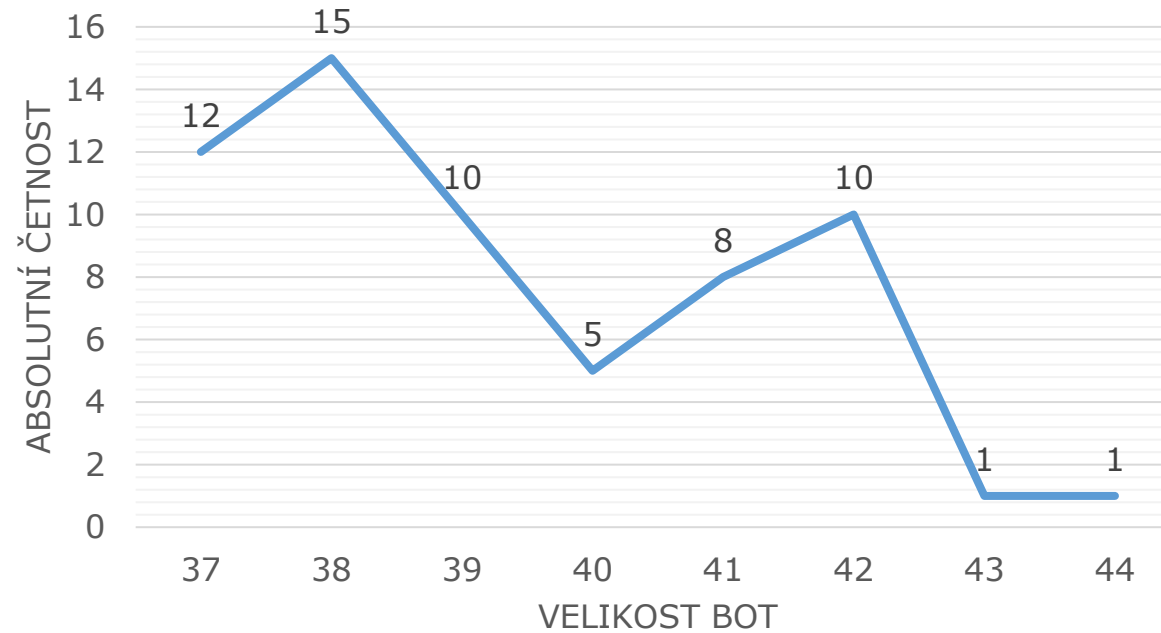
Tabulka 4.6 Tabulka prostého rozdělení četnosti



Graf 4.2 Sloupcový graf



# Polygon četností – řešení

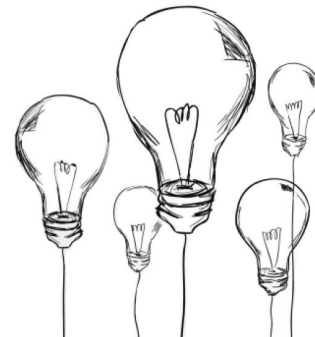


Graf 4.3 Polygon četností

# Statistické charakteristiky numerických proměnných



- Pro popis numerické proměnné můžeme použít většinu statistických charakteristik určené pro popis proměnné ordinální (četnost, relativní četnost, kumulativní četnost, kumulativní relativní četnost).
- Uvedené lze doplnit dalšími charakteristikami, mezi které patří:
- **charakteristiky polohy,**
- **charakteristiky variability či**
- **charakteristiky tvaru rozdělení.**



# Charakteristiky polohy



- Míry polohy určují typické rozložení hodnot proměnné (jejich rozmístění na číselné ose).
- Tyto statistiky nazýváme **střední hodnoty**.

Střední hodnoty rozdělujeme na dvě velké skupiny, a to:

- **průměry,**
- **poziční střední hodnoty** (modus, kvantily a medián, minimální či maximální hodnota).

- Nejznámější z průměrů je **aritmetický průměr**  $\bar{x}$  , ale ve speciálních případech mají využití geometrický  $\bar{x}_G$  a harmonický  $\bar{x}_H$  průměr.



- **Prostý aritmetický průměr** se používá v případech, kdy nejsou hodnoty souboru utříděny. Výpočet se provede následovně:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} .$$



- **Vážený aritmetický průměr** se využívá v případech, kdy máme hodnoty seříděny do tabulky četností.
- Platí, že  $x_1, x_2, \dots, x_k$  značí obměny proměnné  $x$ .
- Proměnné  $n_1, n_2, \dots, n_k$  značí absolutní četnosti obměn  $x$  (tzv. váhy).
- Hodnoty  $p_i$  značí relativní četnost obměny  $x_i$ :



$$\bar{x} = \frac{x_1 n_1 + x_2 n_2 + \dots + x_k n_k}{n} = \frac{\sum_{i=1}^k x_i n_i}{n} = \sum_{i=1}^k x_i p_i.$$



## Příklad

Deset dětí odpovědělo na otázku, jaká je jejich výše měsíčního kapesného. Odpovědi jsou zaznamenány níže:

300, 400, 50, 100, 200, 300, 300, 500, 100, 300.

Vypočítejte prostý aritmetický průměr a také určete průměrné kapesné pro soubor dětí (vážený aritmetický průměr).



## Řešení:

Prostý aritmetický průměr lze vypočítat dle vztahu:

$$\bar{x} = \frac{300 + 400 + 50 + 100 + 200 + 300 + 300 + 500 + 100 + 300}{10} = 255.$$

Průměrné měsíční kapesné tohoto souboru dětí je 255 Kč.

Varianta znaku – kapesné	Absolutní četnost $n_i$
50	1
100	2
200	1
300	4
400	1
500	1
Celkem	10

Vážený aritmetický průměr:

$$\begin{aligned}\bar{x} &= \frac{50 \cdot 1 + 100 \cdot 2 + 200 \cdot 1 + 300 \cdot 4 + 400 \cdot 1 + 500 \cdot 1}{10} \\ &= 255.\end{aligned}$$

Tabulka 4.7 Absolutní četnosti



## Příklad

Vypočítejte vážený aritmetický průměr mezd mužů a žen.

Interval hodinových mezd v Kč	Relativní četnosti	
	Muži	Ženy
100–120	35	58
121–140	37	19
141–160	21	13
161 a více	7	10
Celkem	100	100

Tabulka 4.8 Zadání příkladu

# Řešení

Muži:  $\bar{x} = \frac{13\,000}{100} = 130$

Ženy:  $\bar{x} = \frac{12\,500}{100} = 125$



Interval hodinových mezd v Kč	Relativní četnosti		Střed intervalů $x_i$	$x_i p_i$	
	Muži	Ženy		Muži	Ženy
100–120	35	58	110	3 850	6 380
121–140	37	19	130	4 810	2 470
141–160	21	13	150	3 150	1 950
161 a více	7	10	170	1 190	1 700
Celkem	100	100		13 000	12 500

Tabulka 4.9 Řešení příkladu



- **Harmonický průměr**  $\bar{x}_H$  se používá v případech, kdy proměnná má charakter části z celku a má smysl součet převrácených hodnot proměnné.
- Nejčastěji se využívají v příkladech počítajících s časem.
- **Prostý harmonický průměr** lze vypočítat prostřednictvím následujícího vztahu:

$$\bar{x}_H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_k}} = \frac{n}{\sum_{i=1}^k \frac{1}{x_i}}$$



**Vážený harmonický průměr** lze vypočítat následujícím způsobem:



$$\bar{x}_H = \frac{n_1 + n_2 + \dots + n_k}{\frac{n_1}{x_1} + \frac{n_2}{x_2} + \dots + \frac{n_k}{x_k}} = \frac{\sum_{i=1}^k n_i}{\sum_{i=1}^k \frac{n_i}{x_i}}$$



## Příklad

Vypočtete průměrnou rychlost řidiče, který na trasách :

- z A do B (200 km) jel rychlostí 110 km/h
- z B do C (160 km) jel rychlostí 80 km/h
- z C do D (300 km) jel rychlostí 120 km/h

**Řešení: 104,46 km/h**



- **Geometrický průměr** je možné použít jen v případě, kdy jsou všechny hodnoty kladné a pokud má smysl násobení hodnot proměnných, především pracujeme-li s kladnou proměnnou představující relativní změny (růstové indexy, cenové indexy, ...).

- **Prostý geometrický průměr**



$$\bar{x}_G = \sqrt[n]{x_1 x_2 \dots x_n} = \sqrt[n]{\prod_{i=1}^n x_i}$$



- **Vážený geometrický průměr**

$$\bar{x}_G = \sqrt{(n_1+n_2+\dots+n_k) x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}} = \sqrt{\sum_{i=1}^k n_i \prod_{i=1}^k x_i^{n_i}}$$

# Příklad

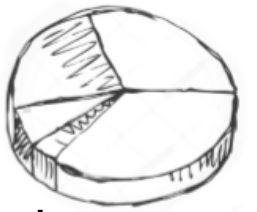


V následující tabulce je uveden vývoj HDP v mld. Kč od roku 2010–2016. Spočítejte průměrné tempo růstu tohoto ukazatele.

Rok	HDP v mld. Kč
2010	3 667,6
2011	3 807,2
2012	4 041,9
2013	4 077,3
2014	4 261,1
2015	4 477,0
2016	4 712,9

Tabulka 4.10 Zadání příkladu – HDP v mld. Kč v letech 2010–2016

## Řešení



Nejdříve je třeba vypočítat koeficienty růstu, které následně použijeme pro výpočet průměrného tempa růstu.

Rok	HDP v mld. Kč	Koeficienty růstu	Tempo růstu v %
2010	3 667,6	×	×
2011	3 807,2	1,038063038	3,806303850
2012	4 041,9	1,061646354	6,164635428
2013	4 077,3	1,008758257	0,875825726
2014	4 261,1	1,045078851	4,50788512
2015	4 477,0	1,050667668	5,066766797
2016	4 712,9	1,052691535	5,269153451

Tabulka 4.11 Řešení příkladu – HDP v mld. Kč v letech 2010–2016

# Řešení



$$\bar{x}_G = \sqrt[6]{1,038063038 \cdot 1,061646354 \cdot 1,008758257 \cdot 1,045078851 \cdot 1,050667668 \cdot 1,052691535} = 1,04268.$$

Průměrné tempo tedy činí:

$$1,04268 \cdot 100 - 100 = 4,268.$$

Odpověď: HDP v ČR mezi lety 2010 a 2016 průměrně rostl tempem 4,268 %.

**Modus** je ta hodnota argumentu  $X$ , která má největší absolutní četnost (vyskytuje se v souboru nejčastěji).

Pro diskrétní proměnnou definujeme modus jako hodnotu nejčetnější varianty proměnné.

U spojité proměnné považujeme modus za hodnotu, kolem níž je největší koncentrace hodnot proměnné



**Kvantil** je chápán jako hodnota, která rozděluje výběrový soubor na dvě části.

První z nich obsahuje hodnoty, které jsou menší než daný kvantil, druhá část obsahuje hodnoty, které jsou větší nebo rovny danému kvantilu.

Pro určení kvantilu je proto nutné výběr uspořádat od nejmenší hodnoty k největší.

Nejznámějším kvantilem je **medián**, který označujeme  $\tilde{x}$ , a který rozděluje hodnoty na dva stejně pravděpodobné intervaly.



## Mezi další nejužívanější kvantily patří



- **kvartily**:  $x_{0,25}$ ,  $x_{0,50}$ ,  $x_{0,75}$ , které dělí obor hodnot na čtyři části, v nichž se náhodná veličina nachází s pravděpodobností 0,25.
- **decily**:  $x_{0,1}$ ,  $x_{0,2}$ , ...,  $x_{0,9}$ , které dělí obor možných hodnot na deset částí se stejnou pravděpodobností výskytu.
- **percentily**:  $x_{0,01}$ ,  $x_{0,02}$ , ...,  $x_{0,99}$ , které dělí obor možných hodnot na sto částí se stejnou pravděpodobností výskytu.



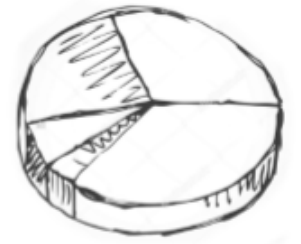
# Výpočet kvantilů

- Z povahy kvantilů je zřejmé, že prvním krokem při jejich výpočtu je **uspořádání** všech hodnot sledovaného znaku v souboru vzestupně podle velikosti.
- Pak stanovíme **pořadové číslo** statistické jednotky, jejíž hodnota je hledaným kvantilem.
- Označme-li toto pořadové číslo  $z_p$ , bude jím první přirozené číslo, pro něž platí  $z_p > np$ , kde  $n$  je rozsah souboru a  $p$  je relativní četnost nejnižších hodnot.
- Je-li součin  $np$  přirozeným číslem, pak je za kvantil obvykle považován průměr z hodnoty na místě s pořadovým číslem  $np$  a na místě s pořadovým číslem  $np + 1$ .



## **Příklad**

V souboru o rozsahu  $n = 126$  určete pořadové číslo pro dolní kvartil, horní kvartil, medián, 10% decil.



## Řešení

- Pořadové číslo pro dolní kvartil je  $z_{0,25} = 32$ , poněvadž  $np = 31,5$ .
- Pořadové číslo pro horní kvartil je  $z_{0,75} = 95$ , poněvadž  $np = 94,5$ .
- Medián je průměr z hodnot na místech s pořadovými čísly 63 a 64, poněvadž  $np = 63$ .
- Pořadové číslo pro 10% decil je  $z_{0,1} = 13$ , poněvadž  $np = 12,6$ .



## Příklad

Určete intervaly obsahující kvartily  $\tilde{x}_{0,25}$ ,  $\tilde{x}_{0,50}$ ,  $\tilde{x}_{0,75}$  v rozdělení četností hodinových mezd naznačeném v tabulce

Interval hodinových mezd v Kč	Relativní četnosti	
	Muži	Ženy
100–120	35	58
121–140	37	19
141–160	21	13
161 a více	7	10
Celkem	100	100

Tabulka 4.12 Zadání příkladu



## Řešení

Interval hodinových mezd v Kč	Relativní četnosti		Kumulativní relativní četnosti	
	Muži	Ženy	Muži	Ženy
100–120	35	58	35	58
121–140	37	19	72	77
141–160	21	13	93	90
161 a více	7	10	100	100
Celkem	100	100	×	×

Tabulka 4.13 Řešení příkladu – dopočtené kumulativní četnosti

## Řešení

Muži: Využijeme sloupce obsahující v procentech vyjádřené kumulativní četnosti. Z vlastností kvantilů vyplývá, že v rozdělení četností můžeme využít kumulativních relativních četností.

Zjistíme, která z nich překročí jako první hodnotu  $p$  hledaného kvartilu.

- dolní kvartil se nachází v intervalu 100–120 Kč (kumulativní relativní četnost je 0,35),
- medián v intervalu 121–140 Kč (kumulativní relativní četnost je 0,72),
- horní kvartil v intervalu 141–160 (kumulativní relativní četnost je 0,93).



Ženy: Hodnotu  $p = 0,25$  jako první převyšuje v tomto případě rovněž kumulativní relativní četnost intervalu 100–120 Kč, která je u žen 0,58.

- Medián se však na základě kumulativní relativní četnosti 0,58 nachází ve stejném intervalu.
- Horní kvartil pak je v intervalu 121–140 Kč (kumulativní relativní četnost 0,77).

Všimněme si, že ze srovnání kvantilů plyne nižší úroveň mezd u žen.



## Příklad

Jsou k dispozici údaje o hodinových mzdách 10 pracovníků jednoho oddělení firmy: 54, 59, 70, 61, 61, 50, 55, 51, 66 a 148. Charakterizujte vhodnou charakteristikou úroveň mezd v daném oddělení.

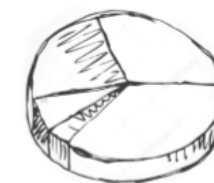
## Řešení:

Nabízí se především zjištění aritmetického průměru hodinové mzdy: Ze srovnání získané hodnoty aritmetického průměru (67,5) s výchozími daty však vyplývá, že prakticky všichni pracovníci (až na dva) mají podprůměrný plat.

Přitom se na výši průměru výrazně podepsala nejvyšší hodinová mzda 148 Kč, která je však v daném souboru netypická (odlehlá, extrémní).

Charakteristikou, která by v daném případě lépe informovala o typické úrovni mezd v souboru, je medián, protože ten obecně není ovlivněn extrémními hodnotami v souboru. Následující tabulka ukazuje jednotlivé hodnoty znaku podle velikosti a ke každé přiřazuje pořadové číslo.

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
50	51	54	55	59	61	61	66	70	148



Tabulka 4.14 Řešení příkladu – pořadová čísla

Jelikož  $np = 5$ , určíme medián jako průměr z páté a šesté hodnoty v řadě (60).

Odpověď: Zjištěná hodnota mediánu 60 Kč v tomto případě vystihuje lépe typickou úroveň platů v oddělení než aritmetický průměr.



# Charakteristiky variability

- Průměry, modus, stejně jako medián vyjadřují pomyslný „střed“ proměnné, neříkají však nic o rozložení jednotlivých hodnot proměnné kolem tohoto „středu“.
- Variabilita obecně značí míru rozdílnosti hodnot statistického souboru.
- Konkrétně se jedná o vymezení následujících charakteristik:

1. charakteristiky absolutní variability,

2. charakteristiky relativní variability.

Charakteristiky **absolutní variability** (variační rozpětí, kvantilová rozpětí, rozptyl, směrodatná odchylka) měří variabilitu v absolutních hodnotách.

Charakteristiky **relativní variability** (např. variační koeficient) mohou vyjadřovat variabilitu v poměru k poloze sledované proměnné.

**Variační rozpětí  $R$**  poskytuje hrubý odhad variability, protože jeho výpočet vzniká vlastně z extrémních hodnot, viz následující vztah:

$$R = x_{max} - x_{min}.$$



**Kvantilová rozpětí** jsou mnohem vhodnějšími charakteristikami pro určení variability, než je variační rozpětí. Jedná se o rozdíly krajních kvantilů, které většinou nejsou zatíženy extrémními hodnotami.

Určujeme **kvartilové rozpětí** (rozdíl mezi horním a dolním kvartilem):

$$R_q = Q_{0,75} - Q_{0,25}.$$



Dále určujeme **decilové rozpětí** (rozdíl krajních decilů):

$$R_d = Q_{0,9} - Q_{0,1}.$$

A v neposlední řadě **percentilové rozpětí** (rozdíl krajních percentilů):

$$R_p = Q_{0,99} - Q_{0,01}.$$





## **Příklad**

Deset dětí odpovědělo na otázku, jaká je jejich výše měsíčního kapesného. Odpovědi jsou zaznamenány níže:

300, 400, 50, 100, 200, 300, 300, 500, 100, 300.

Vypočítejte variační rozpětí a kvartilové rozpětí.

## Řešení:

Variační rozpětí  $R$  lze vypočítat ze vztahu

$$R = 500 - 50 = 450.$$

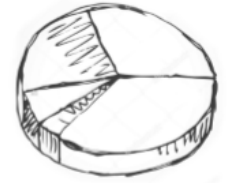
Pro určení kvartilového rozpětí  $R_q$  je nutné vypočítat horní a dolní kvartil. To uděláme tak, že hodnoty seřadíme vzestupně a přiřadíme pořadí, viz následující tabulka

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
50	100	100	200	300	300	300	300	400	500

Tabulka 4.15 Řešení příkladu – pořadová čísla

Pro výpočet dolního kvartilu  $x_{0,25}$  určíme pořadové číslo  $z_{0,25}$ ,  $np$  se rovná:

$$np = 10 \cdot 0,25 = 2,5.$$



## Řešení

Pokud vyjde desetinné číslo, pak  $z_{0,25} = 3$ ,  $x_{0,25}$  je tedy roven 100. Obdobně budeme postupovat pro vyjádření  $x_{0,75}$ .  $Z_{0,75}$  je v tomto případě rovno 8, tedy hodnota horního kvartilu je 300. Nyní lze vypočítat kvartilové rozpětí dle vztahu

$$R_q = 300 - 100 = 200.$$



Odpověď: Variační rozpětí je rovno 450 a kvartilové rozpětí činí 200.

# Přestávka

**Rozptyl**  $s_x^2$  je nejčastěji používaná charakteristika variability numerické proměnné.



Rozptyl je aritmetický průměr čtverců odchylek hodnot sledované proměnné od aritmetického průměru této proměnné:

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n},$$



kde:

$n$  je rozsah statistického souboru,

$x_i$  jsou hodnoty jednotlivých proměnných a

$\bar{x}$  je aritmetický průměr.

Pokud máme hodnoty rozdělené v tabulce četností, pak místo prostého rozptylu použijeme vážený rozptyl, viz následující vztah:

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 n_i}{n}.$$



Prostý rozptyl a vážený rozptyl se využívají v případě, kdy známe všechny hodnoty statistického souboru. Pokud provádíme výběr ze základního souboru, je nutné počítat s výběrovým rozptylem, viz následující vztah:

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}.$$

Výběrový vážený rozptyl má pak tvar

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 n_i}{n - 1}.$$



## Příklad



Z výsledků přijímacích zkoušek byly u 5 studentů z určité střední školy zjišťovány dosažené body v testech z matematiky (znak  $x$ ) a angličtiny (znak  $y$ ), viz následující tabulka. Vypočítejte rozptyl výsledků.

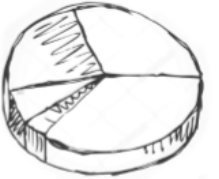
Student	$x_i$	$y_i$
1	30	10
2	10	20
3	60	60
4	80	60
5	20	50
Celkem	200	200

Tabulka 4.16 Zadání příkladu

## Řešení:

Pro výpočet rozptylu je nutné nejdříve výše uvedenou tabulku upravit a vypočítat potřebné mezivýpočty.

Student	$x_i$	$y_i$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	30	10	100	900
2	10	20	900	400
3	60	60	400	400
4	80	60	1 600	400
5	20	50	400	100
Celkem	200	200	3 400	2 200



Tabulka 4.17 Výpočet příkladu s rozptylem

Do mezivýpočtů vstupují aritmetické průměry  $\bar{x}$  a  $\bar{y}$ , které lze jednoduše spočítat a vyjde 40 (200/5). Nyní lze dosadit do vztahu

$$s_x^2 = \frac{3\,400}{5} = 680.$$

$$s_y^2 = \frac{2\,200}{5} = 440.$$

- Jednotka této charakteristiky je druhou mocninou jednotky proměnné, což je značnou nevýhodou při interpretaci rozptylu. Protože je přehlednější popis pomocí základních jednotek, často rozptyl doplňujeme směrodatnou odchylkou, což je odmocnina z rozptylu.
- **Výběrová směrodatná odchylka** (kladná odmocnina výběrového rozptylu) má následující vztah:

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



- **Variační koeficient  $V_x$**  patří do charakteristik relativní variability, poněvadž vyjadřuje relativní míru variability proměnné  $x$  podle níže uvedeného vzorce:

$$V_x = \frac{s_x}{\bar{x}}$$



## Příklad



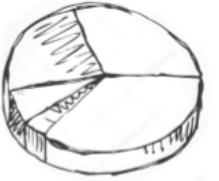
Vypočtete popisné charakteristiky souboru, a to variační rozpětí, modus, vážený aritmetický průměr, rozptyl, směrodatnou odchylku a variační koeficient. Zadání je uvedeno v tabulce.

$i$	$x_i$	$n_i$
1	13	8
2	22	7
3	41	9
4	53	8
5	54	4
6	56	4
7	65	3
8	72	4
9	78	1
10	89	12
$\Sigma$	$\times$	60

## Řešení

Variační rozpětí lze vypočítat jako:

$$R = 89 - 13 = 76.$$



Modus (hodnota s nejvyšší četností) je:

$$\hat{x} = 89.$$

Pro výpočet váženého aritmetického průměru postupujeme následovně:

$$\bar{x} = \frac{13 \cdot 8 + 22 \cdot 7 + 41 \cdot 9 + 53 \cdot 8 + 54 \cdot 4 + 56 \cdot 4 + 65 \cdot 3 + 72 \cdot 4 + 78 \cdot 1 + 89 \cdot 12}{60} = \frac{3\,120}{60} = 52.$$

Pro výpočet rozptylu je vhodné tabulku rozšířit o další sloupec s pomocnými výpočty, viz tabulka.

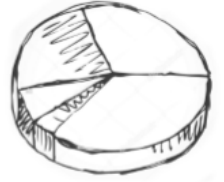
$i$	$x_i$	$n_i$	$(x_i - \bar{x})^2 n_i$	$x_i n_i$
1	13	8	12 168	104
2	22	7	6 300	154
3	41	9	1 089	369
4	53	8	8	424
5	54	4	16	216
6	56	4	64	224
7	65	3	507	195
8	72	4	1 600	288
9	78	1	676	78
10	89	12	16 428	1 068
$\Sigma$	$\times$	60	38 856	3 120



Tabulka 4.19 Řešení příkladu

Pro výpočet rozptylu lze tedy jen dosadit:

$$s_x^2 = \frac{38\,856}{60} = 647,6.$$



Směrodatnou odchylku vypočítáme jako odmocninu z rozptylu:

$$s_x = \sqrt{647,6} = 25,45.$$

A konečně variační koeficient je roven:

$$V_x = \frac{25,45}{52} = 0,49.$$



# Charakteristiky koncentrace

Koncentrací se rozumí jev, kdy se více hodnot proměnné vyskytuje v jisté části variačního rozpětí a naopak. Mezi základní charakteristiky koncentrace patří šikmost a špičatost.

U **šikmosti** se porovnávají počty výskytů velkých a malých hodnot, nejčastěji se pro určení druhu šikmosti provádí porovnání aritmetického průměru a mediánu, což nemusí být zcela spolehlivé.

Šikmost se určuje pomocí koeficientu šikmosti (třetí normovaný moment), který je definován vzorcem

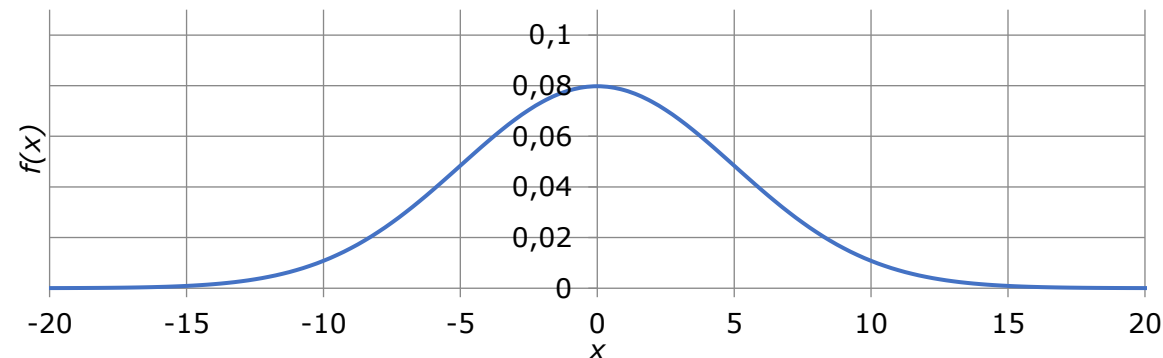
$$\gamma_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{n s_x^3}.$$



Jsou dána tři rozdílná rozdělení četností podle šikmosti, a to:



- souměrné rozdělení četnosti,
  - nesouměrné rozdělení četností zešikmené kladně,
  - nesouměrné rozdělení četností zešikmené záporně.
- 
- Souměrné rozdělení četnosti je typické tím, že koeficient šikmosti je roven 0, přičemž hodnoty větší než průměr a hodnoty menší než průměr mají stejný variační rozsah.
  - U normálního rozdělení má stejnou hodnotu aritmetický průměr i modus a medián (nejčetnější hodnota a prostřední hodnota).
  - Z toho lze vyvodit, že normální rozdělení s největší četností „uprostřed“ musí být symetrické.

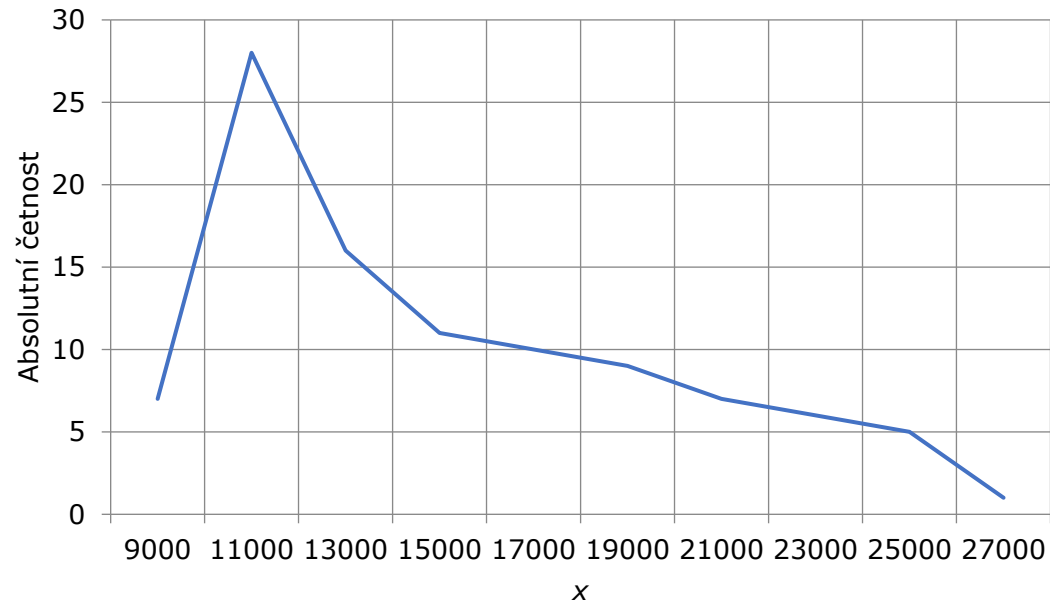


Graf 4.4 Normální rozdělení

Nesouměrné rozdělení četností zešikmené kladně má kladný koeficient šikmosti, což znamená, že vpravo od průměru se vyskytují odlehlejší hodnoty než vlevo.

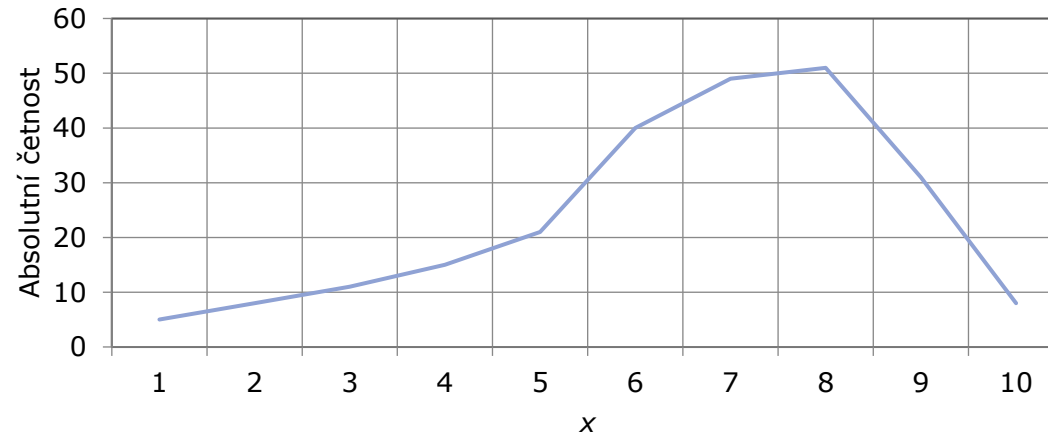
Medián je obvykle menší než průměr.

Jinými slovy: pro **kladně zešikmené** rozdělení je obvyklé, že **aritmetický průměr je větší než medián a modus.**



Graf 4.5 Nesouměrné rozdělení – kladně zešikmené

- Nesouměrné rozdělení četností **zešikmené záporně** má záporný koeficient šikmosti, což znamená, že vlevo od průměru se vyskytují odlehlejší hodnoty než vpravo.
- Medián je obvykle větší než průměr.
- Platí tedy: aritmetický průměr je menší než medián a modus.



Graf 4.6 Nesouměrné rozdělení – záporně zešikmené

**Špičatost** vyjadřuje koncentraci hodnot proměnné kolem jejího průměru. Koeficient špičatosti je nejčastěji používaná charakteristika špičatosti, lze ji vypočítat následovně:

$$\gamma_2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{n s_x^4} - 3,$$



kde:

$n$  je četnost statistického souboru,

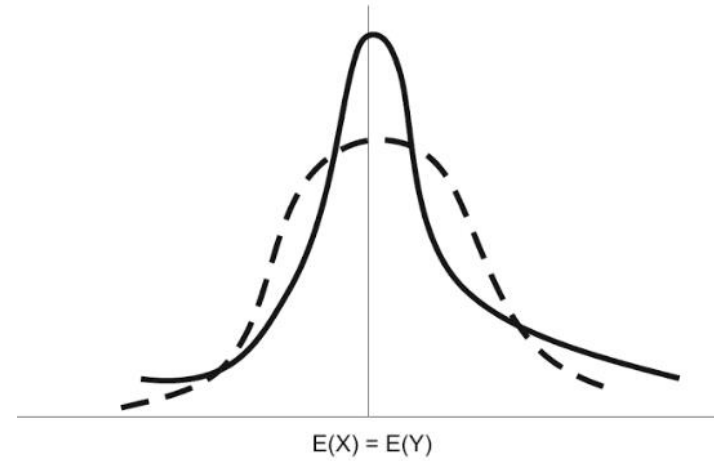
$x_i$  jsou hodnoty proměnné a

$s_x$  je směrodatná odchylka.

Mohou nastat následující případy:

- $\gamma_2 = 0$ , špičatost odpovídá normálnímu rozdělení;
- $\gamma_2 > 0$ , jde o špičaté rozdělení proměnné;
- $\gamma_2 < 0$ , jde o ploché rozdělení proměnné.





Graf 4.7 Špičatost



## Funkce v MS Excel pro charakteristiky polohy

Tabulkový procesor MS Excel 2013 má následující funkce pro určení základních charakteristik polohy:

- PRŮMĚR(číslo1;číslo2;...) – prostý aritmetický průměr;
- GEOMEAN(číslo1;číslo2;...) – geometrický průměr;
- MEDIAN(číslo1;číslo2;...) – medián;
- MODE(číslo1;číslo2;...) – modus;
- PERCENTIL(pole;k), přičemž parametr  $k$  označuje kolika procentní kvantil se bude počítat, zadáváme hodnotu v intervalu 0 a 1;
- QUARTIL(pole;kvartil) – kvartil atd.

Je nutné dávat pozor na to, že MS Excel ignoruje při výpočtu prázdné buňky.



## Funkce v MS Excel pro charakteristiky variability

Excel samozřejmě počítá i základní charakteristiky rozptýlenosti (variability), maximum, minimum, rozptyl, směrodatnou odchylku prostřednictvím následujících funkcí:

- MAX(číslo1;číslo2;...) – maximální hodnota;
- MIN(číslo1;číslo2;...) – minimální hodnota;
- VAR(číslo1;číslo2;...) – rozptyl;
- VAR.VÝBĚR(číslo1;číslo2;...) – výběrový rozptyl;
- SMODCH(číslo1;číslo2;...) – směrodatná odchylka;
- SMODCH.VÝBĚR(číslo1;číslo2;...) – výběrová směrodatná odchylka.



## Funkce v MS Excel pro charakteristiky koncentrace

Pro charakterizování šikmosti a špičatosti lze užit též samostatných funkcí:

- SKEW (číslo1;číslo2;...) – šikmost,
- KURT (číslo1;číslo2;...) – špičatost.

# Přestávka

# 6

## Indexní analýza



Index je bezrozměrné číslo vyjadřující změnu sledovaného ukazatele mezi dvěma obdobími nebo místech srovnání v relativním vyjádření.

Ukazatel je statistická veličina popisující některý ze sociálně-ekonomických jevů. Lze je rozdělit do následujících 2 skupin, a to:

1. extenzivní,
2. intenzivní.



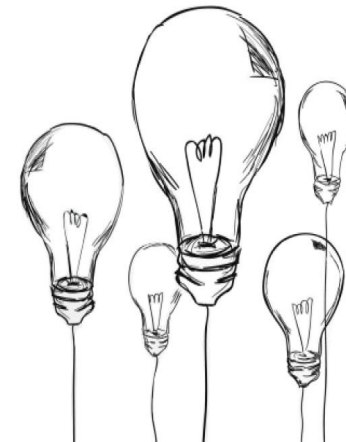
**Extenzivní ukazatele** vyjadřují velikost zkoumaného jevu, poněvadž charakterizují množství, rozsah, objem atd. Tyto ukazatele lze při počítání souhrnů sčítat. Pro označení těchto ukazatelů je používáno písmeno  $q$  (případně  $Q$ ). Extenzivní ukazatele lze označit jako ukazatele množství.

**Intenzivní ukazatele** vznikají jako podíl dvou extenzivních ukazatelů (např. cenu za kus, průměrná mzda atd.). Pro jejich označení se používá znak  $p$ , případně  $c$ . Označují se také jako ukazatele úrovně. Intenzivní ukazatele nelze při souhrnech sčítat.

Hodnoty ukazatele se v čase mění, a proto je nutné je rozlišovat z časového hlediska. Ukazatele se číslují dolními indexy, přičemž 0 značí původní časový údaj, nebo období, vzhledem ke kterému se porovnávání děje (báze) a pozdější údaje postupně označujeme 1, 2, ...

Veličiny	Období	
	Základní	Běžné
Extenzivní	$q_0$	$q_1$
Intenzivní	$p_0$	$p_1$

Tabulka 6.1 Základní a běžné období



Lze rozlišit 3 druhy indexů, a to:

1. individuální jednoduché indexy,
2. individuální složené indexy
3. souhrnné indexy.

Individuální jednoduché indexy sledují vývoj stejnorodé veličiny extenzitní ( $q$ ) nebo intenzitní ( $p$ ) na jednom místě. Pokud porovnáваме extenzivní ukazatel, pak počítáme **jednoduchý individuální index množství**, který vyjadřuje změny extenzitní veličiny, viz následující vztah:

$$I_q = \frac{q_1}{q_0}.$$



Index, který sleduje změny intenzitní veličiny, se nazývá **individuální jednoduchý index úrovně**, který lze vypočítat tímto způsobem:

$$I_p = \frac{p_1}{p_0}.$$





## Příklad

Firma XY prodává jeden druh másla. Cena a prodaná množství jsou naznačeny v následující tabulce. Vyčíslete individuální jednoduchý index úrovně, jednoduchý individuální index množství, tržby. Zároveň vysvětlete vazbu indexu tržby, indexu ceny a množství.

Měsíc	1	2
Cena v Kč/kg	150	180
Prodané množství v kg	100	90

Tabulka 6.2 Zadání příkladu

## Řešení:

Nejdříve začneme výpočtem individuálního jednoduchého indexu úrovně (cenový index) dle vztahu:

$$I_p = \frac{180}{150} = 1,2 = 120 \%$$



V druhém měsíci se zvýšila cena másla na 120 %, tj. o 20 % neboli 1,2× oproti předchozímu měsíci.

Dále bude vypočítán jednoduchý individuální index množství dle vztahu:

$$I_q = \frac{90}{100} = 0,9 = 90 \%$$

V druhém měsíci se snížilo prodané množství másla na 90 %, tj. o 10 %, neboli 0,9× oproti předchozímu měsíci.

Následně dopočítejme tržby, přičemž víme, že jsou násobkem ceny a prodaného množství:

$$Q_1 = 150 \cdot 100 = 15\,000 \text{ Kč.}$$

$$Q_2 = 180 \cdot 90 = 16\,200 \text{ Kč.}$$

Index tržeb lze pak jednoduše vypočítat jako podíl:

$$I_Q = \frac{Q_2}{Q_1} = \frac{16\,200}{15\,000} = 1,08 = 108 \%$$

V druhém měsíci se zvýšily tržby na 108 %, tj. o 8 %, neboli 1,08× oproti předchozímu měsíci.

Odpověď: Výše uvedené indexy spolu bezesporu souvisí. Platí totiž následující: Index tržby = index ceny · index množství, viz níže:

$$1,08 = 1,2 \cdot 0,9.$$

Individuální složené indexy posuzují vývoj stejnorodé veličiny extenzitní ( $q$ ) nebo intenzitní ( $p$ ) na více místech určitého souboru. V případě výpočtů složených individuálních indexů je třeba důkladně hlídat, zda se počítají složené indexy extenzivních, nebo intenzivních ukazatelů. Index, který vyjadřuje vývojové změny extenzitních veličin, jež byly sumarizovány, se nazývá **individuální složený index množství**, který lze vyjádřit následovně:

$$I_q = \frac{\sum q_1}{\sum q_0}.$$



Vliv obou veličin sleduje **index proměnlivého složení**, který lze vyjádřit jako:

$$I_{ps} = \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}}.$$



Někdy potřebujeme znát pouze změnu jedné z těchto složek při konstantní hodnotě druhé složky, je tedy nutné výpočet vztáhnout k jednomu období a počítáme **index stálého složení**, který vyjadřuje vliv změny intenzivní složky při konstantním působení složky extenzivní, viz následující vztah pro běžné období:

$$I_{SS} = \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_1}{\sum q_1}} = \frac{\sum p_1 q_1}{\sum p_0 q_1}.$$



Pro základní období však bude platit následující vztah:

$$I_{SS} = \frac{\frac{\sum p_1 q_0}{\sum q_0}}{\frac{\sum p_0 q_0}{\sum q_0}} = \frac{\sum p_1 q_0}{\sum p_0 q_0}.$$



Podobně lze vymezit **index struktury**, který charakterizuje vliv extenzivního ukazatele při konstantním působení intenzivního ukazatele. Pro běžné období platí následující vztah:

$$I_{STR} = \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_1 q_0}{\sum q_0}}.$$



Pro základní období platí následující:

$$I_{STR} = \frac{\frac{\sum p_0 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}}.$$

Vztahy mezi individuálními složenými indexy

$$I_{ps} = I_{SS} \cdot I_{STR}.$$



## Příklad

V tabulce jsou uvedeny ceny a prodeje téhož druhu zboží v pobočkách firmy A, B, C. Vyčíslete jednoduché individuální indexy a individuální složené indexy.



	Cena Kč/ks		Prodej [100 ks]	
	leden	únor	leden	únor
	$p_0$	$p_1$	$q_0$	$q_1$
Pobočka A	6	8	60	40
Pobočka B	5	3	70	90
Pobočka C	9	10	50	40
Součet	×	×	180	170

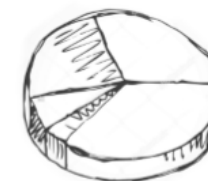
Tabulka 6.3 Zadání příkladu

## Řešení

### Pobočka A:

Začneme výpočtem individuálního jednoduchého indexu úrovně (cenový index) dle vztahu:

$$I_p = \frac{8}{6} = 1,333 = 133,33 \%$$



Dále bude vypočítán jednoduchý individuální index množství dle vztahu:

$$I_q = \frac{40}{60} = 0,66 = 66,66 \%$$

Index tržeb lze pak jednoduše vypočítat jako podíl:

$$I_Q = \frac{8 \cdot 40}{6 \cdot 60} = 0,88 = 88,88 \%$$

## Pobočka B:

Začneme výpočtem individuálního jednoduchého indexu úrovně (cenový index) dle vztahu:

$$I_p = \frac{3}{5} = 0,6 = 60 \%$$

Dále bude vypočítán jednoduchý individuální index množství dle vztahu:

$$I_q = \frac{90}{70} \doteq 1,29 = 128,57 \%$$

Index tržeb lze pak jednoduše vypočítat jako podíl:

$$I_Q = \frac{3 \cdot 90}{5 \cdot 70} = 0,77 = 77,14 \%$$



## Pobočka C:

Opět začneme výpočtem individuálního jednoduchého indexu úrovně (cenový index) dle vztahu:

$$I_p = \frac{10}{9} = 1,11 = 111,11 \%$$

Dále bude vypočítán jednoduchý individuální index množství dle vztahu:

$$I_q = \frac{40}{50} = 0,8 = 80 \%$$

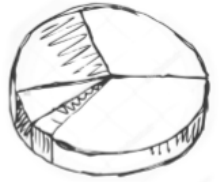
Index tržeb lze pak jednoduše vypočítat jako podíl:

$$I_Q = \frac{10 \cdot 40}{9 \cdot 50} = 0,88 = 88,88 \%$$



	Cena [Kč/ks]		Prodej [100 ks]		Tržby [Kč]		Pomocný
	leden	únor	leden	únor	leden	únor	
	$p_0$	$p_1$	$q_0$	$q_1$	$Q_0$	$Q_1$	$p_1 \cdot q_0$
Pobočka A	6	8	60	40	360	320	480
Pobočka B	5	3	70	90	350	270	210
Pobočka C	9	10	50	40	450	400	500
Součet	x	x	180	170	1 160	990	1 190

Tabulka 6.4 Řešení příkladu



V tuto chvíli již lze vypočítat celkový index tržeb:

$$I_Q = \frac{990}{1160} \doteq 0,85 = 85,34 \%$$

Výše uvedený výpočet znamená, že celkové tržby klesly o 14,66 % (100 % – 85,34 %).

Individuální složený index množství vypočítáme dle vztahu:

$$I_q = \frac{170}{180} = 0,94 = 94,44 \%$$

Znamená to, že prodej celé firmy klesl o 5,56 %.

Index proměnlivého složení vypočítáme jako součin indexu stálého složení a indexu struktury:

$$I_{ps} = I_{SS} \cdot I_{STR} = \frac{1190}{1160} \cdot \frac{990 \cdot 180}{1190 \cdot 170} \doteq 1,03 \cdot 0,88 = 0,904 = 90,37 \%$$

Průměrná cena celkově poklesla o 9,6 %, z toho v důsledku změn cen vzrostla o 3 % a v důsledku změn v prodeji poklesla o 11,91 %.

Souhrnné indexy charakterizují vývojové změny různorodých extenzivních a intenzivních veličin. Extenzitní veličiny jsou zde neslučitelné (jde o různorodé veličiny a lze je sčítat pouze mezi obdobími). Pro srovnání je nutné použít společný intenzivní ukazatel, vzhledem k němuž lze provést souhrn nestejnorodých ukazatelů.

Souhrnný index, který udává změnu vytvořené hodnoty lze nazvat **hodnotový index**, který má tvar:

$$I_h = \frac{\sum q_1 p_1}{\sum p_0 q_0}.$$



**Cenových indexů** existuje celá řada, nicméně rozebereme 4 nejdůležitější:

- Laspeyresův cenový index,
- Paascheho cenový index,
- Loweho cenový index,
- Fisherův cenový index.

Laspeyresův cenový index využívá jako váhu množství základního období, viz následující vzorec:

$$I_c^L = \frac{\sum p_1 q_0}{\sum p_0 q_0}.$$



Paascheho cenový index využívá jako váhu množství běžného období, viz následující vzorec:

$$I_c^P = \frac{\sum p_1 q_1}{\sum p_0 q_1}.$$



Loweho cenový index využívá jako váhu předem zvolené  $q$ :

$$I_c^{LO} = \frac{\sum p_1 q}{\sum p_0 q}.$$

Fisherův cenový index je definován jako geometrický průměr Laspeyresova a Paascheho indexu:

$$I_c^F = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_1 q_1}{\sum p_0 q_1}}.$$



Jak bylo uvedeno výše, podobně jako cenové indexy udávají objemové indexy vliv změny množství na hodnotový index (za předpokladu konstantní ceny).

**Objemové indexy** se také nazývají souhrnné indexy množství.

Laspeyresův objemový index využívá jako váhu cenu základního období:

$$I_q^L = \frac{\sum p_0 q_1}{\sum p_0 q_0}.$$



Paascheho objemový index využívá jako váhu cenu běžného období, viz následující vztah:

$$I_q^P = \frac{\sum p_1 q_1}{\sum p_1 q_0}.$$



Loweho objemový index využívá jako váhu předem zvolené číslo  $p$ :

$$I_q^{LO} = \frac{\sum q_1 p}{\sum q_0 p}.$$

Fisherův objemový index je počítán jako geometrický průměr Laspeyresova a Paascheho objemového indexu:

$$I_q^F = \sqrt{\frac{\sum p_0 q_1}{\sum p_0 q_0} \frac{\sum p_1 q_1}{\sum p_1 q_0}}.$$



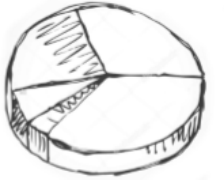
## Příklad

V následující tabulce jsou k dispozici údaje o prodeji tří druhů zboží a jejich cenách v lednu a únoru 2019. Určete celkovou změnu cen prodávaného zboží prostřednictvím souhrnných indexů.

	Cena [Kč/ks]		Prodej [1 000 ks]	
	leden	únor	leden	únor
	$p_0$	$p_1$	$q_0$	$q_1$
zboží A	16	18	90	80
zboží B	15	13	70	90
zboží C	7	10	50	30
Součet	x	x	210	200

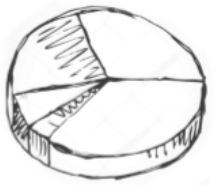


Tabulka 6.5 Zadání příkladu



	Cena [Kč/ks]		Prodej [1 000 ks]		Tržby		Pomocný sloupec	Pomocný sloupec
	leden	únor	leden	únor	leden	únor		
	$p_0$	$p_1$	$q_0$	$q_1$	$Q_0$	$Q_1$	$p_0 \cdot q_1$	$p_1 \cdot q_0$
zboží A	16	18	90	80	1 440	1 440	1 280	1 620
zboží B	15	13	70	90	1 050	1 170	1 350	910
zboží C	7	10	50	30	350	300	210	500
Součet	x	x	210	200	2 840	2 910	2 840	3 030

Tabulka 6.6 Řešení příkladu



Začneme Laspeyresovým cenovým indexem

$$I_c^L = \frac{\sum p_1 q_0}{\sum p_0 q_0} = \frac{3\,030}{2\,840} = 1,067.$$

Vezmeme-li v úvahu vyrobené množství výrobků jako v základním období, tj. v lednu 2019, vzrostly ceny proti základnímu období o 6,7 %, tj. o 190 Kč.

Paascheho cenový index

$$I_c^P = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{2\,910}{2\,840} = 1,025.$$

Vezmeme-li neměnnost objemu prodeje běžného období, ceny prodáváného zboží vzrostly ve sledovaném období o 2,5 %. Kupující museli při nákupu stejného množství zboží jako v běžném období vydat o 70 tis. Kč více.

Fisherův cenový index dle vztahu je uveden níže:

$$I_c^F = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_1 q_1}{\sum p_0 q_1}} = \sqrt{\frac{3\,030}{2\,840} \frac{2\,910}{2\,840}} = 1,046$$

Průměrná změna cen tedy činila 4,6 %.

# Přestávka



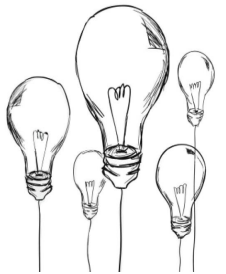
# Bazické a řetězové indexy

Je nutné rozlišit, zda provádíme porovnání vzhledem k bázi. V takovém případě dostáváme bazické indexy.

Obecně bazický index říká, na kolik procent se změnila veličina v běžném období, tj. v následném roce, měsíci, čtvrtletí, pololetí aj. oproti veličině v základním období, tj. prvním uvedeném roce, měsíci, čtvrtletí, pololetí aj.

Výpočet lze provést následovně:

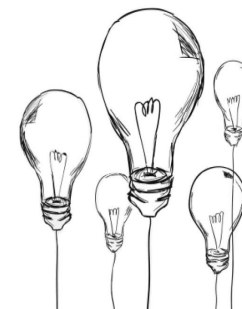
$$I_{\frac{i}{0}} = \frac{q_i}{q_0}.$$



Když	pak hodnota v následných (běžných) obdobích
$I_{\frac{i}{0}} < 100 \%$	se snížila vůči hodnotě v prvním (základním) období.
$I_{\frac{i}{0}} > 100 \%$	se zvýšila vůči hodnotě v prvním (základním) období.
$I_{\frac{i}{0}} = 100 \%$	zůstala stejná, jako hodnota v prvním (základním) období.

Druhou možností je porovnání vzhledem k sousedním obdobím. V takovém případě se hovoří o indexech řetězových, které lze vyjádřit následovně

$$I_{\frac{i}{i-1}} = \frac{q_i}{q_{i-1}}$$



Řetězové indexy, jak je vidět z předchozího vzorce, značí tempo růstu (či poklesu) za každý interval zvlášť.

Hodnota řetězového indexu nabývá jedné ze tří možností, jak ukazuje následující tabulka

Když:	pak hodnota v následných (běžných) obdobích:
$I_{\frac{i}{i-1}} < 100 \%$	se snížila vůči hodnotě v předchozím období.
$I_{\frac{i}{i-1}} > 100 \%$	se zvýšila vůči hodnotě v předchozím období.
$I_{\frac{i}{i-1}} = 100 \%$	zůstala stejná, jako v předchozím období.

Tabulka 6.8 Řetězové indexy

Je možné vyčíslit průměrný koeficient růstu, který lze vypočítat pomocí geometrického průměru:

$$k = \sqrt[n]{\frac{q_n}{q_0}}$$



## Příklad

V následující tabulce lze shlédnout vývoj kurzu dolaru vůči české koruně v letech 2010–2017. Posudte vývoj prostřednictvím bazických a řetězových indexů a také spočítejte průměrný koeficient růstu.

Rok	2010	2011	2012	2013	2014	2015	2016	2017
$x_i$	19,10	17,70	19,58	19,57	20,75	24,60	24,43	23,38



Tabulka 6.9 Řetězové a bazické indexy – zadání



Rok	2010	2011	2012	2013	2014	2015	2016	2017
$x_i$	19,1100	17,6900	19,5800	19,5700	20,7500	24,6000	24,4300	23,3800
$I_{\frac{i}{0}}$	×	0,9257	1,0246	1,0241	1,0858	1,2873	1,2784	1,2234
$I_{\frac{i}{i-1}}$	×	0,9257	1,1068	0,9995	1,0603	1,1855	0,9931	0,9570

Tabulka 6.10 Řetězové a bazické indexy – řešení

Z vývoje uvedeného v tabulce lze např. konstatovat, že kurz v roce 2014 vzrostl oproti výchozímu roku 2010 o 8,58 %. Nebo např. v roce 2016 kurz klesl oproti předchozímu roku 2015 o 0,69 % (99,31 % – 100 %).

Dle vztahu lze také spočítat průměrný koeficient vývoje:

$$k = \sqrt[7]{0,9257 \cdot 1,1068 \cdot 0,9995 \cdot 1,0603 \cdot 1,1855 \cdot 0,9931 \cdot 0,9570} = 1,0292.$$

Znamená to tedy, že průměrně každý rok kurz rostl o 2,92 % (102,92 % – 100 %).

# 7

## Analýza časových řad



**Časová řada** je numerická proměnná, jejíž hodnoty závisí na čase, v němž byly získány. Jedná se o posloupnost chronologicky uspořádaných pozorování. Časové okamžiky, kdy byla data získána, jsou od sebe většinou stejně vzdáleny

Časové řady lze rozdělit pomocí několika hledisek. Podle periodicity lze časové řady rozdělit na:

- krátkodobé (periodicita je kratší než jeden rok),
- dlouhodobé (jejich periodicita je jeden rok a více).



Podle charakteru hodnot časové řady lze rozdělit na:

- Časové řady **intervalové** (např. počet ks zboží vyrobeného za daný měsíc);
- Časové řady **okamžikové** (např. počet strojů, které firma vlastní k určitému datu).

Intervalový ukazatel je ukazatel, jehož velikost závisí na délce intervalu, za který je sledován. Pokud jsou intervaly různě dlouhé, je třeba provést přepočítání na jednotkový interval pomocí tzv. očišťování. V případě, že údaje časové řady toto **zkreslení** v sobě nesou, očišťují se tak, že se transformují na stejně dlouhý interval. Např. průměrná délka měsíce je  $365/12 = 30,42$  dní, potom např. lednovou hodnotu časové řady musíme násobit číslem  $30,42/31$  a únor  $30,42/28$ . Intervalové časové řady lze sčítat.



Časové řady okamžikové jsou řady okamžikových (stavových) ukazatelů. Řady tohoto typu se shrnují pomocí **prostého chronologického průměru**, viz následující vztah

$$\bar{y} = \frac{\frac{y_1 + y_2}{2} + \frac{y_2 + y_3}{2} + \dots + \frac{y_{n-1} + y_n}{2}}{n - 1} = \frac{y_1 + y_2 + \dots + y_{n-1} + \frac{y_n}{2}}{n - 1}.$$



V případě, že délka mezi jednotlivými časovými okamžiky není konstantní, počítáme **vážený chronologický průměr**:

$$\bar{y} = \frac{\frac{y_1 + y_2}{2} d_1 + \frac{y_2 + y_3}{2} d_2 + \dots + \frac{y_{n-1} + y_n}{2} d_{n-1}}{d_1 + d_2 + \dots + d_{n-1}},$$

kde:

$y_1, y_2, \dots, y_n$  jsou jednotlivé hodnoty okamžikového ukazatele,

$d_1, d_2, \dots, d_n$  jsou délky jednotlivých časových intervalů.



## Příklad

Údaje v následující tabulce o množství vyrobených kusů ve firmě XY za 1. pololetí očistěte od časového zkreslení.

Měsíc	Počet dní v měsíci	Množství vyrobených kusů na 1. pololetí
Leden	31	1 500
Únor	28	1 450
Březen	31	2 400
Duben	30	1 800
Květen	31	1 750
Červen	30	1 500
Součet	181	10 400



Tabulka 7.1 Zadání příkladu

Měsíc	Počet dní v měsíci	Množství vyrobených kusů na 1. pololetí	Očištěné hodnoty
Leden	31	1 500	$1\,500 \cdot 30,42/31 \doteq 1\,472$
Únor	28	1 450	$1\,450 \cdot 30,42/28 \doteq 1\,575$
Březen	31	2 400	$2\,400 \cdot 30,42/31 \doteq 2\,355$
Duben	30	1 800	$1\,800 \cdot 30,42/30 \doteq 1\,825$
Květen	31	1 750	$1\,750 \cdot 30,42/31 \doteq 1\,717$
Červen	30	1 500	$1\,500 \cdot 30,42/30 = 1\,521$
Součet	181	10 400	10 465



Tabulka 7.2 Řešení příkladu

## Příklad

Ze záznamů o denních počtech zaměstnanců ve firmě s celotýdenním provozem (viz tabulka) zjistěte průměrnou docházku v daném týdnu.

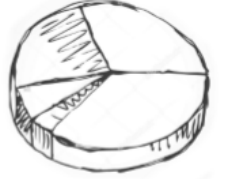
Dny	Pondělí	Úterý	Středa	Čtvrtek	Pátek	Sobota	Neděle
Počet zaměstnanců	50	45	45	50	55	40	30



Tabulka 7.3 Zadání příkladu

## Řešení

Jak již bylo uvedeno výše, okamžikové časové řady shrnujeme prostřednictvím prostého chronologického průměru:



$$\bar{y} = \frac{\frac{y_1}{2} + y_2 + \dots + y_{n-1} + \frac{y_n}{2}}{n - 1} = \frac{\frac{50}{2} + 45 + 45 + 50 + 55 + 40 + \frac{30}{2}}{6} = 45,83.$$

## Příklad

Ze záznamů o počtech zaměstnanců v jednotlivých měsících ve firmě s celotýdenním provozem zjistěte průměrnou docházku.

Dny	leden 2018	únor 2018	březen 2018	duben 2018	květen 2018	červen 2018	červenec 2018
Počet zaměstnanců	150	145	155	140	155	130	120
Délka časové mezery $d_t$	31	28	31	30	31	30	31



Tabulka 7.4 Zadání příkladu

## Řešení

Délka mezi jednotlivými časovými okamžiky není konstantní, počítáme tedy vážený chronologický průměr:



$$\begin{aligned}\bar{y} &= \frac{\frac{y_1 + y_2}{2} d_1 + \frac{y_2 + y_3}{2} d_2 + \dots + \frac{y_{n-1} + y_n}{2} d_{n-1}}{d_1 + d_2 + \dots + d_{n-1}} \\ &= \frac{\frac{150 + 145}{2} 31 + \frac{145 + 155}{2} 28 + \frac{155 + 140}{2} 31 + \frac{140 + 155}{2} 30 + \frac{155 + 130}{2} 31 + \frac{130 + 120}{2} 30}{31 + 28 + 31 + 30 + 31 + 30} \\ &= \frac{4\,572,5 + 4\,200 + 4\,572,5 + 4\,425 + 4\,417,5 + 3\,750}{181} = \frac{25\,937,5}{181} = 143,3.\end{aligned}$$

Odpověď: Průměrná docházka v daném období činí cca 143 zaměstnanci.

Mezi základní charakteristiky časových řad lze zařadit:

- difference jednotlivých řádů (především 1. a 2. řádu),
- tempa růstu (řetězové indexy),
- průměrné tempo růstu,
- průměry hodnot časové řady.



Mějme hodnoty sledovaného ukazatele  $y_t$  pro  $t = 1, 2, \dots, n$ . Diferenci 1. řádu (absolutní přírůstek) lze vyjádřit následovně:

$$D1_t = y_t - y_{t-1}.$$

Absolutní přírůstek vyjadřuje změnu hodnoty v čase  $t$  proti času  $t - 1$ .

Průměrný absolutní přírůstek lze vyjádřit následujícím vztahem:

$$\bar{\Delta} = \frac{1}{n-1} \sum_{t=2}^n D1_t = \frac{y_n - y_1}{n-1}.$$



Výpočet relativních přírůstků je uveden níže:

$$\delta_t = \frac{D1_t}{y_{t-1}} = \frac{y_t - y_{t-1}}{y_{t-1}}, \quad t = 2, 3, \dots, n.$$

Velmi důležitou mírou dynamiky časových řad je **koeficient růstu**, viz níže:



$$k_t = \frac{y_t}{y_{t-1}}, \quad t = 2, 3, \dots, n.$$



**Průměrný koeficient růstu** (neboli průměrné tempo růstu) se vypočítá jako geometrický průměr jednotlivých koeficientů růstu

$$\bar{k} = \sqrt[n-1]{k_2 k_3 \dots k_n} = \sqrt[n-1]{\frac{y_n}{y_1}}.$$



## Příklad

Určete elementární charakteristiky růstu časové řady sledující objem výroby konkrétního zboží ve firmě XY v letech 2013–2018. Údaje jsou zobrazeny v tabulce.

2013	2014	2015	2016	2017	2018
3 256	5 896	4 577	5 423	2 145	2 146

Tabulka 7.5 Zadání příkladu



rok	výroba $y_t$	absolutní přírůstky	koeficienty růstu
2013	3 256	×	×
2014	5 896	2 640	1,811
2015	4 577	-1 319	0,776
2016	5 423	846	1,185
2017	2 145	-3 278	0,396
2018	2 146	1	1,000



## Tabulka 7.6 Řešení příkladu

Průměrný absolutní přírůstek lze vypočítat dle vztahu:

$$\bar{\Delta} = \frac{2\,146 - 3\,256}{5} = -222.$$

Průměrný absolutní přírůstek uvádí, o kolik (222) se každý rok v průměru vyrobí méně zboží.

Další charakteristikou je průměrný koeficient růstu dle vztahu:

$$\bar{k} = \sqrt[5]{\frac{2\,146}{3\,256}} = 0,92.$$

Znamená to tedy, že ročně klesá výroba v průměru o 8 % (100 % - 92 %).



Časová řada se rozkládá na **čtyři základní složky**:

- trend ( $Tr$ ),
- sezónní složka ( $Sz$ ),
- cyklická složka ( $C$ ) a
- náhodná (reziduální) složka ( $\varepsilon$ ).

**Trend**  $Tr$  reprezentuje dlouhodobé změny v průměrném chování řady, poněvadž je způsoben faktory, jež působí systematicky.

**Sezónní složka**  $Sz$  představuje periodické změny, které se odehrávají v průběhu roku a které se opakují. Změny mohou často souviset se střídáním ročních období. Pro analýzu sezónních vlivů se obecně doporučuje pracovat s řadami měsíčních a (nejvýše) kvartálních měření.

**Cyklickou složku**  $C$  lze chápat jako fluktuaci (kolísání) kolem trendu. Délka cyklu i intenzita jednotlivých fází mohou v průběhu času měnit.

**Náhodná (reziduální) složka**  $\varepsilon$  představuje náhodné fluktuace. Náhodná složka nemá systematický charakter, protože zahrnuje též chyby měření či chyby ve statistickém zpracování dat. Znalost každé složky časové řady umožní také lepší odhad vývoje daného procesu do budoucna (predikci).

# Příklady k procvičení



## Cvičení ke kapitole **kombinatorika**:

- Kolika způsoby si můžeme nastavit PIN na kartě? (10 000)
- Určete počet prvků množiny všech možných výsledků, pokud ze skupiny 10 výrobků chceme vybrat libovolné 2. (45)
- Kolik různých hodů můžeme provést dvěma spravedlivými šestistěnnými kostkami? (36)
- Kolik různých hodů můžeme provést třemi spravedlivými šestistěnnými kostkami? (216)
- Určete, kolik je možností, jak sestavit 6místné telefonní číslo. Náповěda: telefonní číslo nemůže začínat 0. (900 000)



## Cvičení ke kapitole **pravděpodobnost**:

- Ve skupině je 20 mužů a 12 žen. Losem jsou určeni 2 vedoucí skupiny. Jaká je pravděpodobnost, že oba vedoucí budou různého pohlaví? (0,484)
- Jaká je pravděpodobnost, že při hodu dvěma šestistěnnými spravedlivými kostkami obsahujícími čísla 1 až 6, padnou na obou z nich čísla větší než 4? (0,111)
- Je známo, že na 10 výrobků připadají 3 vadné. Náhodně vybereme skupinu 5 výrobků. Určete pravděpodobnost, že v této skupině budou právě 3 vadné. (0,083)
- Z 25 lístků označených čísly 1–25 vytáhneme náhodně jeden lístek. Jaká je pravděpodobnost, že na vytaženém lístku bude liché číslo? ( $13/25 = 52\%$ )
- V hazardní číselné hře se losuje 6 čísel ze 49 čísel. Jaká je pravděpodobnost uhodnutí 3 čísel ze 6 losovaných? (0,0177)



## Cvičení ke kapitole **popisná statistika**:

- Spočtete harmonický průměr z následujících 5 hodnot: 4, 8, 16, 16, 64. (9,69)
- Při hodnocení úspěšnosti přijímacích zkoušek bylo vybráno 15 náhodných testů, kde byl zjištěn počet bodů z angličtiny. Vypočtete medián. (38)

25	42	45	36	38	31	29	35	40
42	39	37	33	41	45			

- V tabulce hodnocení úspěšnosti přijímacích zkoušek v předchozím příkladu vypočítejte prostý aritmetický průměr. (37,2)
- V tabulce hodnocení úspěšnosti přijímacích zkoušek v předchozím příkladu vypočítejte rozptyl a směrodatnou odchylku. (33,74; 5,808)



## Cvičení ke kapitole **indexní analýza**:

- V tabulce je uveden počet servisních oprav výrobků v záruční době za rok.

Rok	2006	2007	2008	2009	2010
Počet oprav	10	15	13	10	14

Určete: Hodnota řetězového indexu v roce 2010 je ..., a tedy došlo k ....: (1,4; zvýšení o 40 %)

- Firma má 3 pobočky. Chceme zjistit změnu ceny výrobku A mezi rokem 2009 a 2010 za všechny pobočky dohromady. Použijeme: (Individuální index složený – cenový)
- Uvažujme řadu tržeb (v tis. Kč) v letech 2005–2008. Vypočítejte bazický index v roce 2008 se základem v roce 2005 a meziroční řetězový index v roce 2008. Rok 2005: 17, rok 2006: 19, rok 2007: 19, rok 2008: 21. (Indexy: bazický index: 1,24 a řetězový index: 1,11)



### Cvičení ke kapitole **časové řady**:

- Očistěte od časového zkreslení březnové (31 dní) a listopadové (30 dní) hodnoty – březen: 1 200 a listopad: 1 090. (březen: 1 178, listopad: 1 105)
- Určete elementární charakteristiky růstu časové řady sledující objem výroby konkrétního zboží ve firmě XY v letech 2013–2018. Údaje jsou zobrazeny v tabulce. (řešení na dalším snímku)

rok	výroba $y_t$
2013	4 521
2014	7 458
2015	4 122
2016	4 511
2017	622
2018	9 875

rok	výroba $y_t$	absolutní přírůstky	koeficienty růstu
2013	4 521		
2014	7 458	2 937	1,650
2015	4 122	-3 336	0,553
2016	4 511	389	1,094
2017	622	-3 889	0,138
2018	9 875	9 253	15,876

Průměrný absolutní přírůstek je 1070,8. Průměrný koeficient růstu je 1,169.

# Chyby ve skriptech Metody zpracování dat v ekonomické praxi (Kubová, 2018)

- Str. 31 a 33 – vzorec pro permutace s opakováním, správně:

$$P'(k_1, k_2, \dots, k_n) = \frac{(k_1 + k_2 + \dots + k_n)!}{k_1! k_2! \dots k_n!}.$$

- Str. 47 – otázka 4, správná odpověď a) 0,5**2**
- Str. 119, tabulka 5.27

Pro výpočet rozptylu lze tedy jen dosadit:

$$s_x^2 = \frac{38\,856}{60} = 647,6.$$

Směrodatnou odchylku vypočítáme jako odmocninu z rozptylu:

$$s_x = \sqrt{647,6} = 25,45.$$

Variační koeficient:

$$V_x = \frac{25,45}{52} = 0,49.$$

# Chyby ve skriptech Metody zpracování dat v ekonomické praxi (Kubová, 2018)

- Str. 119, tabulka 5.27

$i$	$x_i$	$n_i$	$(x_i - \bar{x})^2 n_i$	$x_i n_i$
1	13	8	12 168	104
2	22	7	6 300	154
3	41	9	1 089	369
4	53	8	8	424
5	54	4	16	216
6	56	4	<b>64</b>	224
7	65	3	507	195
8	72	4	1 600	288
9	78	1	676	78
10	89	12	16 428	1 068
$\Sigma$	$\times$	60	38 856	3 120

## Chyby ve skriptech Metody zpracování dat v ekonomické praxi (Kubová, 2018)

- Str. 215

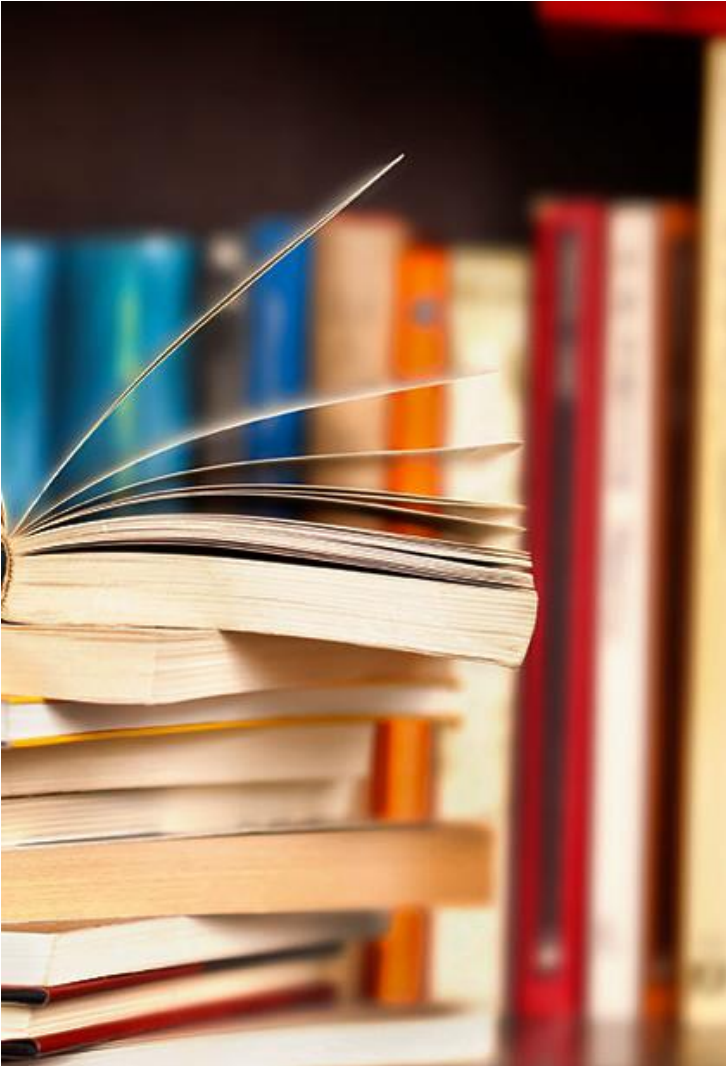
Fisherův cenový index vychází nikoli 1,093, ale **1,046**.

Průměrná změna cen tedy činila **4,6 %** (nikoli 9,3 %).

# Chyby ve skriptech Metody zpracování dat v ekonomické praxi (Kubová, 2018)

- Str. 217:

Rok	2010	2011	2012	2013	2014	2015	2016	2017
$x_i$	19,1100	17,6900	19,5800	19,5700	20,7500	24,6000	24,4300	23,3800
$I_{\frac{i}{0}}$	×	0,9257	1,0246	<b>1,0241</b>	1,0858	1,2873	1,2784	1,2234
$I_{\frac{i}{i-1}}$	×	0,9257	1,1068	0,9995	1,0603	1,1855	0,9931	0,9570



ANDĚL, J. *Základy matematické statistiky*. 2., opr. vyd. Praha : Matfyzpress, 2007. ISBN 80-7378-001-1.

ARLT, J., ARLTOVÁ, M., RUBLÍKOVÁ, E. *Analýza ekonomických časových řad s příklady*. 2. vyd. Praha : Oeconomica, 2004. ISBN 80-245-0777-3.

BARTSCH, H. *Matematické vzorce*. Vyd. 4., Academia 1. (reprint). Praha : Academia, 2006. ISBN 80-200-1448-9.

BAUER, L. *Matematika v ekonomii a ekonomice*. Praha : Grada, 2015. Expert (Grada). ISBN 978-80-247-4419-3.

BÍLKOVÁ, D., BUDINSKÝ, P., VOHÁNKA, V. *Pravděpodobnost a statistika*. Plzeň : Vydavatelství a nakladatelství Aleš Čeněk, 2009. ISBN 978-80-7380-224-0.

BRISĚ, R. *Probability and Statistics for Engineers*. Ostrava, 2011. Project CZ.1.07/2.2.00/15.0132. Dostupné z WWW: <http://homel.vsb.cz/~bri10/Teaching/Prob%20&%20Stat.pdf>.

BRISĚ, R., LITSCHMANNOVÁ, M. *Statistika II* [online]. Ostrava : Vysoká škola báňská – Technická univerzita, 2008 [cit. 2017-09-26]. ISBN 978-80-248-1482-7.

BUDÍKOVÁ, M., KRÁLOVÁ, M., MAROŠ, B. *Průvodce základními statistickými metodami*. Praha : Grada, 2010. Expert (Grada). ISBN 978-80-247-3243-5.

COUFAL, J. *Logika a logické myšlení*. Praha: Vysoká škola ekonomie a managementu, 2014. ISBN 9788087839591.

COUFAL, J., LÍNEK, V. *Logika a matematika pro ekonomy*. Praha : Vysoká škola ekonomie a managementu, 2009. ISBN 978-80-86730-48-6.

CYHELSKÝ, L., KAHOUNOVÁ, J., HINDLS, R. *Elementární statistická analýza*. Vyd. 2., dopl. Praha : Management Press, 1999. ISBN 80-7261-003-1.

FOTR, J., HNILICA, J. *Aplikovaná analýza rizika ve finančním managementu a investičním rozhodování*. 2., aktualiz. a rozš. vyd. Praha : Grada, 2014. Expert (Grada). ISBN 978-80-247-5104-7.

HÁTLE, J., LIKEŠ, J. *Základy počtu pravděpodobnosti a matematické statistiky*. Vyd. 2. Praha : SNTL – Nakladatelství technické literatury, 1974.

HÁTLE, J., KAHOUNOVÁ, J. *Úvod do teorie pravděpodobnosti : celostátní vysokoškolská učebnice pro studenty ekonomických fakult studijních oborů 62*. Praha : Státní nakladatelství technické literatury, 1987.

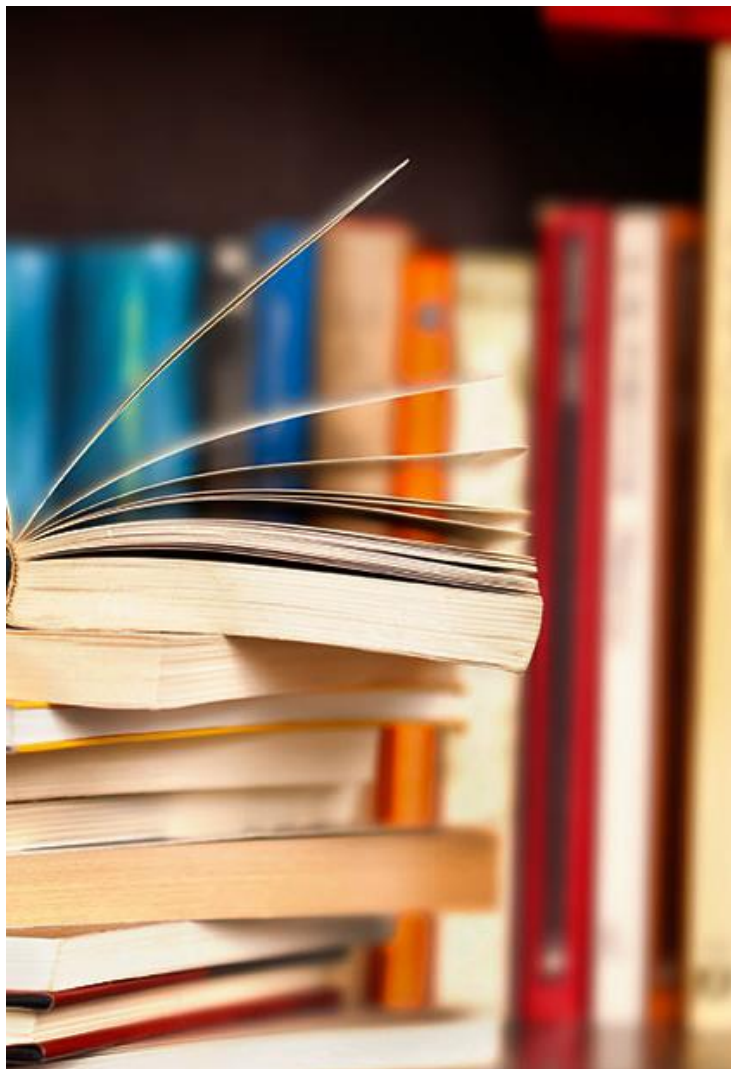
HEBÁK, P. *Statistické myšlení a nástroje analýzy dat*. 2. vydání. Praha: Informatorium, 2015. ISBN 978-80-7333-118-4.

HINDLS, R., SEGER, J., HRONOVÁ, S. *Statistika pro ekonomy*. Brno : Professional Publishing, 2002. ISBN 80-86419-26-6.

HOŠKOVÁ, P. *Statistika I*. V Praze : Česká zemědělská univerzita, Provozně ekonomická fakulta, 2013. ISBN 978-80-213-2341-4.

HUFF, D. *Jak Ihát se statistikou*. Ilustroval Petra BOBALOVÁ. Praha : Brána, 2013. ISBN 978-80-7243-623-1.

JAROŠOVÁ, E., PECÁKOVÁ, I. *Příklady k předmětu Statistika B*. 2. vyd. Praha : Oeconomica, 2004. ISBN 80-245-0680-7.



KÁBA, B., SVATOŠOVÁ, L. *Statistické nástroje ekonomického výzkumu*. Plzeň : Vydavatelství a nakladatelství Aleš Čeněk, 2012. ISBN 978-80-7380-359-9.

KOVAŘÍK, P. *Aplikovaná statistika*. Brno : Vysoká škola Karla Engliš, 2011. ISBN 978-80-86710-28-0.

KŘÁPEK, M. *Statistika I*. Distanční studijní opora. SVŠE Znojmo, 2013. ISBN 978-80-87314-40-1.

KŘIVÝ, I. *Analýza časových řad*. Ostrava : OU, 2006, [online], [cit. 2017-12-12], Dostupné z WWW: [http://prf.osu.cz/doktorske\\_studium/dokumenty/Time\\_Series\\_Analysis.pdf](http://prf.osu.cz/doktorske_studium/dokumenty/Time_Series_Analysis.pdf).

KŘIVÝ, I. *Analýza časových řad*. Ostrava : OU, 2012, [online], [cit. 2017-12-12], Dostupné z WWW: <https://publi.cz/download/publication/45?pc=1>.

KUBANOVÁ, J. *Statistické metody pro ekonomickou a technickou praxi*. Statis Bratislava, 2004a. ISBN. 80-85659-379.

KUBANOVÁ, J. *Sbírka příkladů z pravděpodobnosti*. Bratislava : Statis, 2004b. ISBN 80-85659-36-0.

KUBESA, M. *Základy diskrétní matematiky, výukový text I (skripta online)*, FEI VŠB – TUO, 2011.

KURZY. HDP 2018, vývoj HDP v ČR [online]. 2018 [cit. 2018-01-31]. Dostupné z WWW: <https://www.kurzy.cz/makroekonomika/hdp/?G=3&A=2&page=1>.

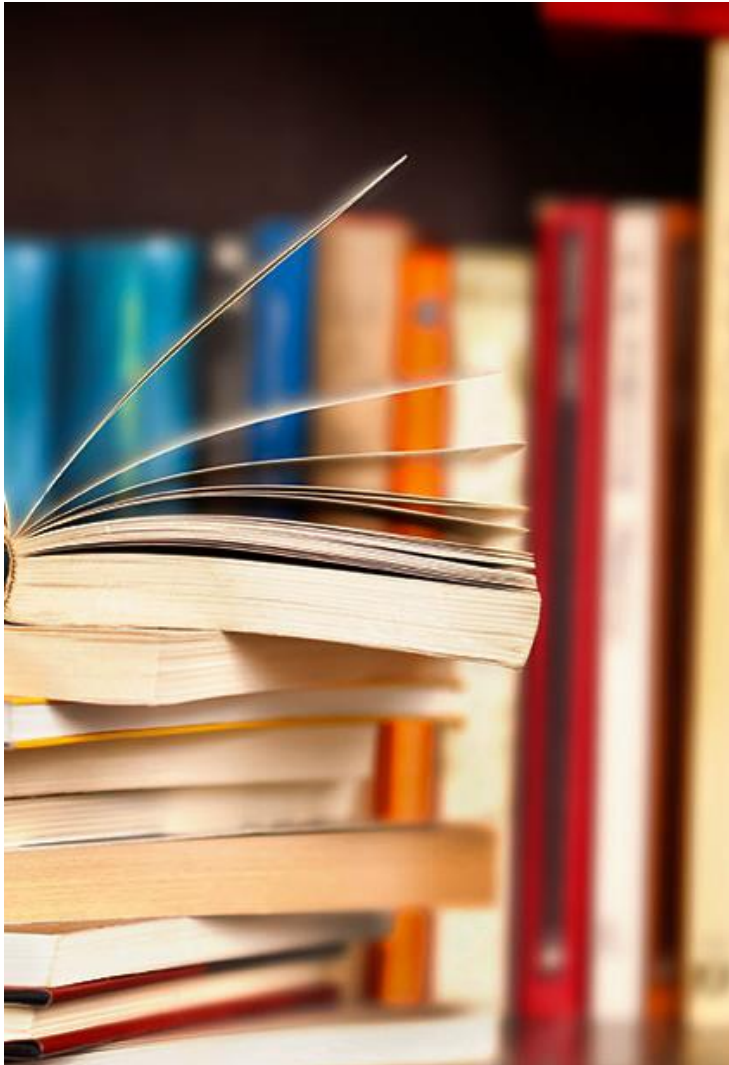
LEGIERSKÁ, Y. *Statistika v cestovním ruchu* [online]. 2007 [cit. 2017-10-31] Dostupné z WWW: <http://www.mmr.cz/getmedia/d99c234e-7624-4a22-9638-7688dd1040e8/Statistika-v-cestovnim-ruchu.pdf>.

LITSCHMANNOVÁ, M. Úvod do analýzy časových řad [online]. Ostrava : VŠB, 2010, [cit. 2017-12-16]. Dostupné z WWW: [http://homel.vsb.cz/~lit40/SMAD/Casove\\_rady.pdf](http://homel.vsb.cz/~lit40/SMAD/Casove_rady.pdf).

LITSCHMANNOVÁ, M. *Úvod do statistiky, elektronická skripta a doplňkové interaktivní materiály*. 2012.

LITSCHMANNOVÁ, M. *Vybrané kapitoly z pravděpodobnosti*. VŠB-TU Ostrava, 2011.

LÖSTER, T., ŘEZANKOVÁ, H.H.H, LANGHAMROVÁ, J. *Statistické metody a demografie*. Praha : Vysoká škola ekonomie a managementu, 2009. ISBN 978-80-86730-43-1.



MACEK, J. a kol. *Ekonomická statistika*. Ostrava : Vysoká škola báňská – TU, 2002. ISBN 80-7082-841-2.

MAREK, L. a kol. *Statistika pro ekonomy : aplikace*. 2. vyd. Praha : Professional Publishing, 2007. ISBN 978-80-86946-40-5.

NEUBAUER, J. SEDLAČÍK, M. A KŘÍŽ O. *Základy statistiky: Aplikace v technických a ekonomických oborech*. Praha : Grada, 2012. ISBN 978-80-247-4273-1.

NEUBAUER, J., SEDLAČÍK, M., KŘÍŽ, O. *Základy statistiky : aplikace v technických a ekonomických oborech*. 2., rozšířené vydání. Praha : Grada, 2016. ISBN 978-80-247-5786-5.

OTIPKA, P., ŠMAJSTRLA, V. *Pravděpodobnost a statistika*. Ostrava : Technická univerzita, 2008. ISBN 978-80-248-1302-8.

PŁOCKI, A., TLUSTÝ, P. *Pravděpodobnost a statistika pro začátečníky a mírně pokročilé*. Praha : Prometheus, 2007. ISBN 978-80-7196-330-1.

ROBOVÁ, J., HÁLA, M., CALDA, E. *Komplexní čísla, kombinatorika, pravděpodobnost a statistika : matematika pro střední školy*. Praha : Prometheus, 2013. Učebnice pro střední školy (Prometheus). ISBN 978-80-7196-425-4.

SEGER, J., HINDLS, R. *Statistické metody v tržním hospodářství*. Praha : Victoria Publishing, 1995. ISBN 80-7187-058-7.

TOŠENOVSKÝ, J. *Teorie pravděpodobnosti*. Ostrava : Vysoká škola báňská – Technická univerzita, 2012. ISBN 978-80-248-2591-5.

ZVÁROVÁ, J. a MAZURA, I. *Stochastická genetika*. Praha : Karolinum, 2002. ISBN 80-246-0264-4.

**DĚKUJI ZA  
POZORNOST**