KGG/STG Statistika pro geografy

3. Teoretická rozdělení – úvod

Mgr. David Fiedor
2. března 2015

Osnova

1. Úvod
2. Historie
3. Základní pojmy
4. Teoretická rozdělení

Čím jsme se doposud zabývali

- definování základních pojmů
- grafická prezentace jevů a analýza zpracování
- typy statistických znaků
- charakteristiky polohy, variability, asymetrie a špičatosti

- shrnutí informací = *popisná statistika*

Matematická statistika

- hlavním úkolem statistiky „dnešní doby“ není pouze shrnovat data do přehlednějších forem, nýbrž výsledky zobecňovat

- počet pravděpodobnosti

Jak často nastane určitý jev, pokud experiment nebo výběr provedeme vícekrát?

Historie pravděpodobnosti

- Řekové - antiempirismus (nikoliv experimenty, pouze logická cesta)
- Blaise Pascal (1623–1662) - al zhar (kostka)
- Christian Huygens - Výpočty v hrách náhody (1657) - učební text z teorie pravděpodobnosti
- K. Pearson - pokus hodem mincí

Pojmy

Náhodný jev

Náhodným jevem rozumíme stav, kdy za určitého souboru podmínek může nastat jeden z množiny možných výsledků, který nezávisí pouze na vstupních podmínkách, ale obsahuje také prvek náhody.

Náhodný pokus

Pokus, jehož realizací lze obdržet různé výsledky a přitom:

- nelze předem určit, který z výsledků získáme,
- pokus lze libovolně často opakovat, aniž se jednotlivá opakování vzájemně ovlivňují,

nazýváme náhodný.

Pojmy

Náhodná proměnná

Předpis, který přiřazuje každému výsledku náhodného pokusu určité číslo, se nazývá náhodná proměnná.

- náhodné proměnné budeme značit velkými písmeny
- spojitá - může nabývat jakékoliv hodnoty z určitého intervalu hodnot
- diskrétní - může nabývat pouze konkrétních hodnot

Definice pravděpodobnosti

Statistická definice pravděpodobnosti

Pravděpodobnost náhodného jevu A je číslo P(A), k němuž se blíží relativní četnost jevu A, jestliže pokus dostatečně často opakujeme. Jestliže jsme provedli n pokusů a v m z nich nastal jev A, pak názorně vyjádřena:

$$lim_{n \to \infty} \frac{m}{n} = P(A)$$

Vlastnosti pravděpodobnosti

- normovanost - pravděpodobnost jakéhokoliv jevu je kladné číslo z intervalu $\langle 0, 1 \rangle$
- jev jistý má pravděpodobnost 1, jev nemožný pravděpodobnost 0
- komplementarita - je-li A´ opačný jev k jevu A, pak součet pravděpodobností těchto dvou jevů je 1
- monotonie - jestliže $A_1 \subseteq A_2$, potom $P(A_1) \leq P(A_2)$
- aditivita - jestliže $A_1 \cap A_2 = \emptyset$, potom $P(A_1 \cup A_2) = P(A_1) + P(A_2)$
- spočetná aditivita - $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

Příklad

Z karetní hry o 32 kartách vybereme náhodně bez vracení 4 karty. Jaká je pravděpodobnost, že aspoň jedna z nich je eso?

$$P(A) = 1 - P(A') = 1 - \frac{\binom{28}{4}}{\binom{32}{4}} = 0,4306$$

Podmíněná pravděpodobnost

- často závisí pravděpodobnost určitého jevu na tom, zda nastal či nenastal nějaký jiný jev
- $P(A|B)$ - pravděpodobnost jevu A za předpokladu, že nastal jev B
- $P(A \cap B) = P(A|B)P(B)$, kde P(B) není jev nemožný
- neboli $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

Stochasticky nezávislé jevy

Rovná-li se podmíněná pravděpodobnost pravděpodobnosti nepodmíněné, tedy $P(A|B) = P(A)$, říkáme, že jevy A a B jsou stochasticky nezávislé a platí $P(A \cap B) = P(A) \cdot P(B)$.

Určení pravděpodobnosti apriori

- podíl počtu požadovaných výsledků a počtu všech možných výsledků

Příklad

S jakou pravděpodobností padne při házení kostkou šestka?

Řešení:

$$P(A) = \frac{m}{n} = \frac{1}{6} = 0,1667$$

Určení pravděpodobnosti aposteriori

- určení pravděpodobnosti pomocí relativní četnosti výskytu studovaného jevu
- $P(A) = \frac{n_i}{n}$, kde n_i je počet požadovaných výsledků, které nastaly při realizaci jevu a n je celkový počet pokusů (rozsah souboru)

Příklad

Z deseti hodů kostkou jsme získali následující výsledky: 2, 4, 6, 1, 6, 3, 5, 6, 2, 1. Určete pravděpodobnost, s jakou padla u tohoto pokusu šestka.

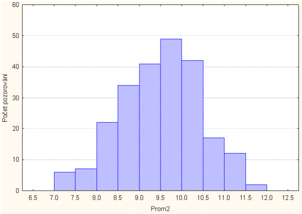
Řešení:

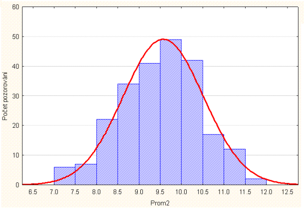
$$P(A) = \frac{n_i}{n} = \frac{3}{10} = 0,3$$

Rozdělení náhodné proměnné

Náhodná proměnná

Předpis, který přiřazuje každému výsledku náhodného pokusu určité číslo, se nazývá náhodná proměnná.





Teoretická rozdělení - pojmy

Princip

Budeme-li zvětšovat rozsah souboru (při předpokladu, že náhodná veličina je spojitá) a hodnoty třídit do stále menších intervalů, dostaneme histogramy, které se budou stále více blížit hladké křivce.

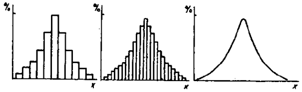
Teoretická rozdělení charakterizujeme:

1. průběhem *frekvenční* a *distribuční funkce*
2. parametry rozdělení

Frekvenční funkce

Frekvenční funkce $f(x)$ představuje teoretické rozdělení četností základního souboru o parametrech μ a σ^2.

- cílem je nahradit výběrové soubory základními a pro ně odvozovat potřebné charakteristiky

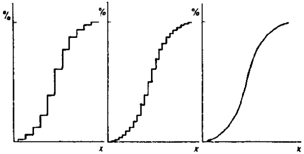


Distribuční funkce

Distribuční funkce $F(x)$ udává pravděpodobnost, se kterou náhodná proměnná nabývá hodnoty menší nebo rovné určité konkrétní velikosti x, tj.:

$$F(x) = P(X < x) = \sum_{x_i < X} P(X = x_i)$$

- lze jednoduše zkonstruovat pomocí součtové čáry (čáry kumulativních četností)



Parametry rozdělení

- neznámé hodnoty základních statistických charakteristik (μ a σ^2) základního souboru
- nutnost tyto parametry odhadnout z výběrového souboru ($\bar{x}$ a s^2)
- kapitola věnující se odhadům parametrů základního souboru bude tvořit samostatnou přednášku

Vybraná rozdělení náhodných proměnných

- normální rozdělení
- normované normální rozdělení
- binomické rozdělení
- Poissonovo rozdělení
- χ^2 rozdělení
- Studentovo t rozdělení
- Fisher-Snedecorovo rozdělení
- exponenciální, lognomální,...

Děkuji za pozornost...