METODY V GEOGRAFII
Mgr. Darina MÍSAŘOVÁ, Ph.D.

Sylabus přednášky 5: Teoretická rozdělení

Sylabus slouží jako přehled základních pojmů zmiňovaných na přednášce. Není dostačující pro úspěšné zvládnutí zkoušky z Metod v geografii. Sylabus je nezbytné doplit informacemi z přednášky.

K čemu je to dobré?

Popisné a průzkumové metody umožňují přehledné shrnutí informací, které se týkají jen objektů měřených či pozorovaných.

- **výběrový soubor - popisujeme jen to, co bylo zjištěno, naměřeno.**
- **zobecňující úsudky**

Příklady:

- Jak často se takováto povodeň může opakovat?
- Jakou hodnotu měřené veličiny nejpravděpodobněji získáme
- opakovaným měřením?
- Je vysoký počet dvojčat narozených v určitém okrese „normální"?
- Je rozdíl mezi dvěma jevy významný?

Náhodný jev, náhodná proměnná

Náhodný jev - za určitého souboru podmínek může nastat jeden z množiny výsledků, který závisí nejen na vstupních podmínkách, ale obsahuje i **prvek náhody (tahání karet, měření teploty vzduchu, …).**

Náhodná proměnná – proměnná, u které nelze na základě určité zákonitosti předem stanovit její konkrétní hodnotu.

Náhodná veličina

a) náhodná veličina **spojitá**
Může teoreticky nabývat nekonečného množství hodnot z určitého intervalu
Př:

b) náhodná veličina **nespojitá**
Nabývá jen konečného množství hodnot urč. intervalu.
Př:

- Každé hodnotě je možno přiřadit pravděpodobnost jejího výskytu, součet všech dílčích pravděpodobností je 1

Pravděpodobnost

- Vyjadřuje míru nejistoty, s jakou určitý náhodný jev může nastat.
- Vyjadřuje míru nejistoty s jakou může náhodná veličina nabývat určité hodnoty

- **Tuto míru nejistoty (pravděpodobnost) můžeme kvantifikovat.**
- Řada jevů a procesů studovaných v geografických disciplínách má charakter **náhodné proměnné, má pravděpodobnostní charakter** (mohu nastat s určitou pravděpodobností) – např. výsledky prognóz (demografie, meteorologie apod.)
- Pravděpodobnost jako vyjádření míry nejistoty o výskytu náhodného jevu, o výsledku náhodného jevu.
- Pravděpodobnost, že nastane určitý náhodný jev se pohybuje v intervalu:

Jev možný – množina všech možných výsledků - náhodný jev
Jev jistý – padne něco mezi 1 až 6
Jev nemožný – padne 7
Jev elementární – padne 6
Jev složený – více možných výsledků (padne sudé číslo)

Pravděpodobnost

*P(A) - Určení pravděpodobnosti P, s jakou náhodný jev A nastane,*můžeme povést dvěma způsoby:

a) Určení pravděpodobnosti „a priori“:

Podíl počtu požadovaných výsledků a počtu všech možných výsledků:
Př.: S jako pravděpodobností padne při házení kostkou šestka:

b) Určení pravděpodobnosti „a posteriori“:

Pomocí relativní četnosti výskytu studovaného jevu:

ni – počet požadovaných výsledků, které nastaly při realizaci jevu
(absolutní četnost)
n – celkový počet pokusů (rozsah souboru)

Příklad: Z deseti hodů kostkou (n=10) jsme získali následující výsledky:
2,4,6,1,6,3,5,6,2,1. Spočteme frekvenci výskytu jednotlivých výsledků a následně relativní četnost výsledku, při kterém padla šestka, tedy počet případů příznivých jevu A k počtu případů možných.

ROZDĚLENÍ NÁHODNÉ PROMĚNNÉ

- Každý výsledek náhodného jevu - určitá pravděpodobnost
- Můžeme určit s jakou pravděpodobností náhodný jev nabývá určité výsledné hodnoty či hodnoty z určitého intervalu.
- model umožňující zobecnění našich poznatků o chování hromadných náhodných jevů – **teoretické rozdělení pravděpodobnosti**

Teoretická rozdělení – základní pojmy

- Teoretická rozdělení ve statistice charakterizujeme:

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

- Neznámé hodnoty základních statistických charakteristik základního souboru, které můžeme jen **odhadnout z charakteristik výběrových**

Teoretická rozdělení spojité náhodné veličiny

- Frekvenční funkce f(x) představuje teoretické rozdělení četností základního souboru o parametrech μ, σ.
- Cíl – nahradit výběrové soubory základními a pro ně odvozovat potřebné charakteristiky
- Analogicky lze ze součtové čáry definovat tzv. **distribuční funkci *F(x).***
- kumulativní relativní četnost tj. součtová čára - distribuční funkce Fx
- Distribuční funkce udává pravděpodobnost, se kterou náhodná proměnná nabývá hodnoty menší nebo rovné určité konkrétní velikosti x.

NORMÁLNÍ ROZDĚLENÍ / GAUSSOVO, LAPLACEOVO- GAUSSOVO

- **Normální rozdělení se univerzálně používá k aproximaci (k přibližnému vyjádření) rozdělení pravděpodobnosti velkého množství náhodných veličin (v biologii, technice, ekonomii atd.)**
- Nejčastěji používané rozdělení spojité náhodné veličiny.
- Opakované měření stejné veličiny za stejných podmínek.
- Naměřené veličiny více méně kolísají kolem skutečné hodnoty
- Má dva parametry:

- **Hustota pravděpodobnosti normálního rozdělení je symetrická zvonovitá Gaussova křivka.**
- **Vlastnosti:**

- Pomocí násobků směrodatné odchylky lze stanovit pravděpodobnosti, s nimiž leží hodnoty v určitém intervalu:

- Normální křivka a osa x **vymezují plochu 100%,**
- tj. lze stanovit pravděpodobnosti, s nimiž leží hodnoty v určitém intervalu,
- hranice intervalu tvoří průměr a násobky směrodatné odchylky
  obr.

- Normální rozdělení s parametry:
  - stejný průměr, různé směrodatné odchylky
  - čím větší odchylka , tím „plošší" tvar rozdělení

Příklady

Př.1

Populace má v daném testu průměr 100, směrodatnou odchylku 15.
Vypočítejte hranice intervalů, v kterém se nachází 68 % populace.

Př.2

Výška v populaci chlapců ve věku 3,5 - 4 roky má normální rozdělení s průměrem 102 cm a směrodatnou odchylkou 4,5 cm.

Vypočítejte hranice intervalu hodnot výšky , ve kterých se nachází
A)70% (68%)
B) 95%
C)99%
příslušné populace

Př.3

Výška v populaci chlapců ve věku 3,5 - 4 roky má normální rozdělení s průměrem 102 cm a směrodatnou odchylkou 4,5 cm.
Spočtěte, jaké procento chlapců v uvedeném věku má výšku menší nebo rovnou 93 cm.

Př. 4

Psychologickými testy bylo zjištěno, že hodnota IQ populace je náhodnou veličinou s normálním rozdělením, jehož střední hodnota je 104 a směrodatná odchylka 8.
Určete hodnotu IQ:
meze, ve kterých bude 50% populace

BINOMICKÉ ROZDĚLENÍ

- pro diskrétní náhodné proměnné, které mohou nabývat pouze dvou hodnot (např. ano, ne)
- pravděpodobnost, že nastane alternativa ANO označme π
- pravděpodobnost, že nastane NE …q = 1 – π), protože
- platí π +q = 1 (100 %)
- k výpočtu se používá **binomický rozvoj**

Př. 1 – binomické rozdělení
Předpokládejme, že pravděpodobnost narození dívky je 0,49.
Jaká je pravděpodobnost toho, že mezi třemi dětmi v rodině je právě jedna dívka?

Př. 2
Jaká je pravděpodobnost, že v rodině s 8 dětmi jsou právě 3 dívky? Pravděpodobnost narození dívky je 0,49.

Př. 3
Vypočítejte pravděpodobnost, se kterou se vyskytne určitý počet měsíců v roce hodnocených jako „ suché“.
Konkretizace:
- oblast Oxford,
- období 1851 – 1943, tj. 1116 měsíců
- Suchý měsíc - tj. méně srážek v měsíci než je dlouhodobý průměr tohoto měsíce.
- 617 měsíců hodnocených jako suché
- 499 – vlhké měsíce
- Jak bude vypadat situace pro „vlhké“ měsíce?

Příklady použití binomického rozdělení
- rozdělení počtu dní s určitým meteorologickým jevem za měsíc
- pravděpodobnost narození dvou chlapců v rodinách se třemi dětmi
- pravděpodobnost pozdního příchodu na jednu ze 12 přednášek ze statistiky

Příklad: Pravděpodobnost, že se v určitém roce vyskytne na studovaném toku povodeň je 0,25. Jaká je pravděpodobnost, že se během příštích čtyř let vyskytnou 3 povodně?

POISSONOVO ROZDĚLENÍ

- pro rozdělení vzácných případů
- např: zimní bouřka, výskyt mutace apod.

- Je-li pravděpodobnost nějaké výjimečné události (např. určité mutace genu) relativně malá a rozsah výběru poměrně velký, pak Poissonovo rozdělení v podstatě splývá s binomickým, ale je mnohem výhodnější pro počítání .
- Náhodná veličina s Poissonovým rozdělením může nabývat hodnot *x = 0,1,2,3,... (kolikrát jev nastal v určitém časovém úseku) a to s* rozdělením pravděpodobnosti:

- Pro aritmetický průměr a rozptyl platí:

- kde lambda (λ) je očekávaná hodnota a jediný parametr Poissonova rozdělení:

- Označuje se jako rozdělení vzácných případů (bouřky v zimě, výskyt krupobití v roce, …). Jeho použití se doporučuje, pokud n > 30 (resp. 50) a *p ≤0,1 nebo p ≥ 0,9 .*

Příklady použití Poissonova rozdělení

- počet dětí ztracených v obchodním domě v určité časovém úseku
- počet telefonních hovorů v určitém časovém úseku
- počet borovic na jednotku plochy smíšeného lesa

Př.

Předpokládejme, že v určité populaci krys se vyskytuje albín s pravděpodobností p = 0,001 , ostatní krysy jsou normálně pigmentované.

Ve vzorku 100 krys náhodně vybraných z této populace určete pravděpodobnost, že vzorek

a) neobsahuje albína,

b) obsahuje právě jednoho albína.

PEARSONOVA KŘIVKA III. TYPU

- Na empirické rozdělení mnoha statistických souborů s nimiž v geografii pracujeme, nelze aplikovat normální rozdělení.
- Platí to například v těch případech, kdy studovaná náhodná veličina nemá teoreticky zdůvodněnou možnost nabývat nekonečných hodnot nebo je-li omezena konečnými čísly
- V takovýchto případech lze aplikovat na studovaný soubor některou ze dvanácti křivek Pearsonova systému.

- Především v meteorologii a klimatologii se ke konstrukci tzv. čar překročení využívá Pearsonovy křivky III. typu.
- obvykle pro veličiny s omezeným množstvím hodnot, které může nabývat
- z křivky lze např. vyčíst pravděpodobnost se kterou bude hodnota sledovaného statistického znaku dosažena
- Průběh křivky je určen třemi parametry:
1. aritmetickým průměrem
2. variačním koeficientem
3. koeficientem asymetrie
- v hydrologii se počítá Pearsonova křivka - součtová čára četností jako tzv. čára překročení
- Čára překročení je součtová čára četností a lze z ní stanovit pravděpodobnost, se kterou bude znak určité hodnoty dosažený a překročený (či nebude dosažený).

ROZDĚLENÍ CHÍ – KVADRÁT

- Ze základního souboru s normovaným normálním rozdělením provedeme náhodný výběr *n prvků, které označíme X1,X2, X3,….Xn*
- Součet čtverců těchto hodnot se označuje jako („chí – kvadrát“):
- Hodnota chí-kvadrát může nabývat v různých výběrech různých hodnot v intervalu (0, ∞) a má své vlastní rozdělení (frekvenční a distribuční funkci)
- Symbol **v** značí počet stupňů volnosti a je jediným parametrem rozdělení. Je roven rozsahu náhodného výběru.
- Každé hodnotě **v = n** přísluší jiná křivka. S rostoucím **v** se rozdělení blíží rozdělení normálnímu.

Použití:

- v teorii odhadu a testování hypotéz
- při ověřování předpokladu zda empirické rozdělení četností má určité teoretické pravděpodobností rozdělení
- testování rozptylu dvou výběrových souborů při neznámé střední hodnotě
- při ověřování nezávislosti kvalitativních znaků
- pro testy nezávislosti v kontingenčních tabulkách.

STUDENTOVO/T/ ROZDĚLENÍ

- Využívá se především pro hodnocení odchylek hodnot aritmetického průměru základního souboru a aritmetického průměru výběrového souboru .
- Pro hodnocení odchylek se definuje náhodná veličina t