

Statistika

Hun Fauser

19. února 2015

1 LEKCE 01

Pravděpodobnost jevu A ($P(A)$) je rovna poměru počtu případů jevu A (m) k počtu případu možných (n). $P \in < 0; 1 >$

$$P_A = \frac{m}{n}$$

Ke každému náhodnému jevu A existuje opačný jev \bar{A} , který je nenastoupením jevu A, platí:

$$P(\bar{A}) = 1 - P(A)$$

PŘÍKLAD

S jakou pravděpodobností padne po jednom hodu kostkou číslo 6?

ŘEŠENÍ

$$P = \frac{1}{6}, \text{ respektive } P = 1.167$$

Statistická definice pravděpodobnosti

Statistická definice pravděpodobnosti je založená na pojmu relativní četnosti, kdy pravděpodobnost chápeme jako kvantitativní vlastnost náhodného jevu, kterou měříme pomocí relativní četnosti výskytu náhodného jevu, přičemž přesnost měření roste s počtem opakování nezávislých pokusů.

PŘÍKLAD

Určitý pokus (hod kostkou) může být mnohonásobně (N-krát) opakován za neměnných podmínek - N pokusů. Jestliže se u N pokusů vyskytne jev A (padne číslo 6) M-krát, poměr $\frac{M}{N}$ je relativní četností jevu A.

Seminar

Matematika - z řeckého jazyka

Matematika je věda zabývající se z formálního hlediska kvantitou, strukturou, prostorem a změnou.

Charakteristickou vlastností matematiky je její důraz na absolutní přesnost metod a nezpochybnitelnost výsledků

Historie matematiky sahá až do pravěku, kdy vznikly první abstraktní matematické pojmy - čísla.

Nula - nepřírozené číslo

Eukleides z Megary - řecký matematik, 4-3st. před n.l. Zakladatel matematiky.

Leonardo Fibonacci - arabské číslice. 13st.

Matematika je univerzální jazyk.

Existuje jenom jedna forma matematiky.

Elementární matematika

- Operace s čísly
- Rovnice
- Geometrické objekty

Aplikovaná matematika

- Ekonomie
- Statistika
- Chemie

Čistá matematika - vysoce abstraktní věda. Nepraktická.

V matematice se využívá řecká abeceda a symboly.

Konstanta - dané číslo. Nemusí být známá.

Opakem konstanty je proměnná

$$\pi = 3.141592654$$

$$e = 2.718281828$$

$$\phi = 1.618033989$$

Eulerovo číslo

Eulerovo číslo jako limita následující posloupnosti:

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

Eulerovo číslo jako součet následující nekonečné řady.

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots$$

Eulerovo číslo jako jediné číslo $x > 0$, pro které platí, že:

$$\ln x = \int_1^x \frac{dt}{t} = 1$$

Výrok

- Je tokové tvrzení, o kterém lze jednoznačně rozhodnout, zda je to pravdivě nebo ne
- Označuje se malými písmeny latinské abecedy
- Jeho platnost se označuje číslem 1
- Neplatnost 0

Výroky

- Jednoduché (nedelitelné)
- Složené

Jednoduché výroky

- $3 < 5$ (1) - platí
- $2 > 3$ (0) - neplatí

Symbol negace - \neg

Slovně "není pravda"

Logické operace

Konjunkce $a \wedge b$

Disjunkce $a \vee b$

Implikace $a \Rightarrow b$

Ekvivalence $a \Leftrightarrow b$

Výroková forma - výraz, který obsahuje jednu nebo více proměnných z daného, předem určeného souboru, který se po dosažení konkrétních hodnot proměnných stane výrokem.

Kvantifikatory

- Velký
- Malý

Obecný

- Daná vlastnost platí pro každý prvek
- Výroky, které obsahují obecný kvantifikator se nazývají obecné výroky

Výrok	Negace výroku
Každý...je	Aspoň jeden...není
Aspoň jeden...je	Žádný...není
Aspoň n...je	Nejvýše...je
Nejvýše n...je	Aspoň...je
Právě n...je	

Statistika

Status - stav

- Číselné udaje
- Praktická činnost
- Vědní disciplína

Český Statistický Úřad

Statistické charakteristiky

- Míry polohy
- Míry variability (s nedefinovanou proměnlivostí uvnitř souboru dat, s definovanou proměnlivostí uvnitř souboru dat)

$$a(x) : (x + 1)^2 = x^2 + 2x + 1; x \in R$$

$$\forall((x^2 - 1) - (x + 1)(x - 1)); x \in R$$

$$\exists(x^2 + 1 = 0); x \in R$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Aritmetický průměr \bar{x} střední hodnota kvantitativního statistického znaku (součet hodnot, dělený jejích počtem)

$$\bar{x} = \frac{\sum x_i}{n}$$

, kde n je rozsah souboru

Modus \hat{x} - hodnota nejčastěji se v souboru vyskytující

Median \tilde{x} = prostřední hodnota \Rightarrow je-li n (rozsah souboru) liché číslo, median je prostřední hodnota, je-li n sudé číslo je median aritmetickým průměrem dvou prostředních hodnot.

100P% kvantil x_p je číslo, které odděluje 100P% nejmenších hodnot náhodné veličiny X .

Tedy 50% kvantil $x_{0.50}$ je totéž co medián

Dobrý popis rozdělení pravděpodobnosti dostaneme stanovením dostatečného počtu kvantilů.

Kvantily zaznamenané po dvaceti pěti procentech nazýváme kvartily, po deseti procentech decily a po jednom procentu percentily.

Tedy 25% kvantil je 1. kvartil (dolní kvartil), 10% kvantil je 1. decil a podobně 1% kvantil je 1. percentil

Medián je totéž co 50% kvantil, 2. kvartil, 5. decil nebo 50. percentil.

S použitím kvartilů, decilů a percentilů se často setkáváme při prezentaci výsledků antropometrických studií.

2 LEKCE 02

V hotelu je ubytováno 185 hostu

Není pravda, že v hotelu je ubytováno 185 hostu

Alespoň 4 zaměstnanci hotelu získají ocenění

Nejvýše 3 zaměstnanci hotelu získají ocenění

Žádný klient hotelu není z Ostravy

Alespoň jeden klient našeho hotelu je z Ostravy

Právě jeden z našich hostů je v sauně

Nikdo z našich hostů není v sauně, nebo jsou tam alespoň dva

Číslo 29 je prvočíslo - výrok, pravdivý

Vltava je nejdelší řeka světa - výrok, nepravdivý

Pro reálná čísla x, y platí $x+y=10$ - výroková forma

Univerzita Karlova byla založena v roce 1492 - výrok, nepravdivý

Součin dvou libovolných záporných čísel je kladné číslo - výrok, pravdivý

Přímka p je rovnoběžná s přímkou q - výroková forma

Přejí Vám dobrý večer - zdělení

Prezident Václav Havel zemřel v roce 2011 - výrok, pravdivý

Trojúhelník ABC je pravoúhlý - výroková forma

Rozdíl dvou libovolných kladných čísel je kladné číslo - výrok, nepravdivý

Operator	Slovní vyjádření	Symbol
Negace	není pravda, že	\neg
Konjunkce	a, a zároveň	\wedge
Disjunkce	nebo	\vee
Implikace	jestliže...pak	\Rightarrow
Ekvivalence	právě když, právě tehdy, když	\Leftrightarrow

Velký kvantifikátor - \forall

Malý kvantifikátor - \exists

Příklad

11, 20, 11, 9, 25, 6, 10, 11, 18, 19, 18

Sort

6 9 10 11 11 11 18 18 19 20 25

Min - 6

Max - 25

Median (\tilde{x}) - 11 (středina)

Modus (\hat{x}) - 11 (nejčastější číslo)

Aritmetický průměr (\bar{x}) - 14.36 (střední)

Teorie množin

Množina je takový souhrn objektů, že o každém objektu můžeme rozhodnout, zda patří nebo nepatří do uvažovaného souhrnu objektu

Množina - A, B, C

Prvky množiny a, b, c

Množiny

Konečné - má konečně mnoho prvků

Nekonečné - počet prvků není konečný

Prázdné - množiny nemají žádný prvek \emptyset

Sjednocení množin A a B je množina prvků, které buď patří do množiny A nebo do množiny B ($A \cup B$)

Průnik množin A a B je množina těch prvků, které patří oběma množinám A a B ($A \cap B$)

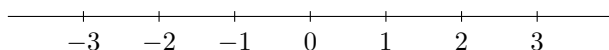
Rozdíl množin A a B je množina těch prvků, které patří do množiny A, ale nepatří do množiny B ($A - B$)

Karteský součin množin A a B je množina všech uspořádaných dvojic (a, b), takových, že a patří do množiny A a b do množiny B ($A \times B$).

	1	2	3
x	(x, 1)	(x, 2)	(x, 3)
y	(y, 1)	(y, 2)	(y, 3)
z	(z, 1)	(z, 2)	(z, 1)

Přirozená čísla (\mathbb{N}) - 1, 2, 3, 4, ...
 Celá čísla (\mathbb{Z}) - -2, -1, 0, 1, ...
 Racionální čísla (\mathbb{Q}) - $1/2$, 0, 3.6, ...
 Iracionální čísla - π , e, ...

Reálná osa



Přímka, na níž je zvolen Počátek jako obraz čísla 0 a jednotka
 Každému reálnému číslu potom odpovídá právě jeden bod reálné osy a naopak, každému bodu reálné osy je přiřazeno právě jedno reálné číslo

Komplexní čísla (\mathbb{K} nebo \mathbb{C})

$a+bi$

Číslo složené z reálné a imaginární částí

Číselné obory \mathbb{N} , \mathbb{Z} , \mathbb{Q} , \mathbb{R} jsou v tomto vzájemném vztahu:

$\mathbb{N} \in \mathbb{Z} \in \mathbb{Q} \in \mathbb{R}$

Absolutní hodnota vyjadřuje vzdálenost obrazu čísla na číselné ose od nuly Absolutní hodnota je ve skutečnosti vzdálenost. Stejně jako vzdálenost od bodu na obě strany nemůže být záporná nemůže ani absolutní hodnota být záporná

Interval je každá taková množina reálných čísel, jejichž obrazy na číselné ose vyplňují její souvislou podmnožinu

Rozptyl je průměrná hodnota za součtu čtverců odchylek jednotlivých hodnot souboru od aritmetického průměru

Směrodatná odchylka je kladná hodnota druhé odmocniny rozptylu

Grafické znázornění normálního rozdělení je dáno touto symetrickou jednovrcholovou hustotou, která je zvonovitého tvaru a nikde neprotíná vodorovnou osu

$a \in B$, a elementem B

$b \notin B$, b není elementem B

Uzavřený interval - $\langle a, b \rangle$, $a \leq x \leq b$

Otevřený interval - (a, b) , $a < x < b$

Zleva otevřený - $(a, b \rangle$, $a < x \leq b$

Zleva uzavřený - $\langle a, b)$, $a \leq x < b$

Intervaly

Množinový zápis - $x \in R; a < x \leq b$

Pomocí intervalu - $(a; b>$

Slovní vyjádření

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$
$$S^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}$$

Odchylka

$$\sigma = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2}$$
$$S = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}}$$

Variabilní koeficient

$$V_k = \frac{S}{\bar{x}} \times 100$$

V_k - variabilní koeficient

S - směrodatná odchylka

\bar{x} - aritmetický průměr

3 MATH TWEAKS

3.1 Number 01

$$A = 100\% \quad B = x$$
$$x = \frac{B \times 100}{A}$$

3.2 Number 02

$$A = 100\% \quad x = y\%$$
$$x = \frac{A}{B} \times y$$

3.3 Number 03

$$A + B = 110$$
$$A = B + 100$$
$$B + 100 + B = 110$$
$$2B = 10$$
$$B = \frac{10}{2} = 5$$
$$A = 5 + 100 = 105$$
$$: B = 5; A = 105;$$