

Rozložení četnosti

Mgr. Zuzana Szabó Lenhartová

Rozložení četnosti

- ověř, co jsme ve výzkumu získali (náhodná data, je správně interpretováno) (statistická), která má za cíl, četnost výskytu jednotlivých hodnot (resp. hodnot uvnitř jednotlivých intervalů)
- konfiguraci četností nazýváme rozložení (rozložení)
- rozložení popisujeme
 - tabulkou četností
 - graficky – histogram, skupkový diagram
 - pomocí percentilů
- o typu, tvaru rozložení hodnot proměnné využíváme většinou graficky – histogram, skupkový diagram.

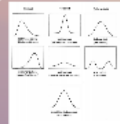
Rozložení četnosti

- Mírně jsou popsány vzhledem k rozložení (početnosti) a jeho tvaru (je to důležité pro interpretaci)
- Číslo rozložení, které má vliv na výskyt proměnných uvnitř intervalů (statistická) – pokud rozložení je správně interpretováno, má to velký vliv na výskyt proměnných uvnitř intervalů (statistická)



Tvar rozložení četností

- Normální
- Uniformní
- Počet vrcholů
- Unimodální, bimodální, multimodální
- Zešikmení
- Zešikmeně zprava (pozitivně), efekt podlahy
- Zešikmeně vlevo (negativně), efekt stropu
- Šířkost
- Lepičkovitě
- Platykurtické

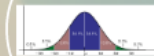


Normální rozložení (Gaussovo rozložení)

- „Normální“ se týká „světlá bílá“
- Je to rozložení, které nalezáme u lidí, které jsou statisticky velmi rozšířené (bílé) – to znamená, že máme dost výzkumu, se statisticky velmi podobáme ostatním
- Normální je „světlá bílá“ dat
- I parametrizované normální rozložení, jsou normální
- v angličtině: bell curve (zvonek křivka)
- při výšce, váze, IQ, měření výskytu žlout v tělesné výšce, ...



Normální rozložení



Uniformní rozložení

- všechny hodnoty jsou stejné



Počet vrcholů

- vzhledem ke počtu prvků (jednotek) v datové sadě, rozložení má, ale může být i 0
- počet = 0 vrcholů = bimodální
- počet = 1 vrchol = unimodální
- počet = 2 vrcholy = multimodální
- počet = 3 vrcholy = multimodální

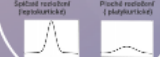
Šířkost

- širší rozložení znamená větší rozptýlení (širší rozložení má více rozptýlené hodnoty)
- širší rozložení znamená větší rozptýlení (širší rozložení má více rozptýlené hodnoty)
- širší rozložení znamená větší rozptýlení (širší rozložení má více rozptýlené hodnoty)
- širší rozložení znamená větší rozptýlení (širší rozložení má více rozptýlené hodnoty)



Špičatost

- špičatost (leptokurtické) znamená, že rozložení má špičatější vrchol (špičatější vrchol)
- špičatost (leptokurtické) znamená, že rozložení má špičatější vrchol (špičatější vrchol)
- špičatost (leptokurtické) znamená, že rozložení má špičatější vrchol (špičatější vrchol)



Popis rozložení pomocí percentilů

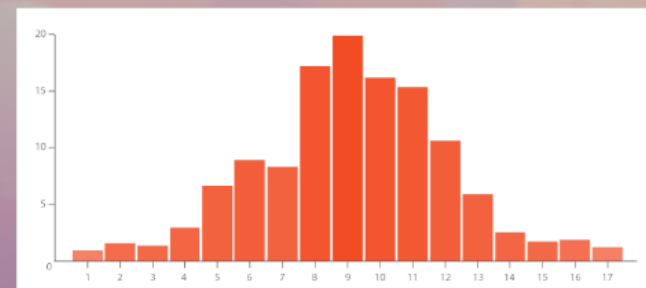
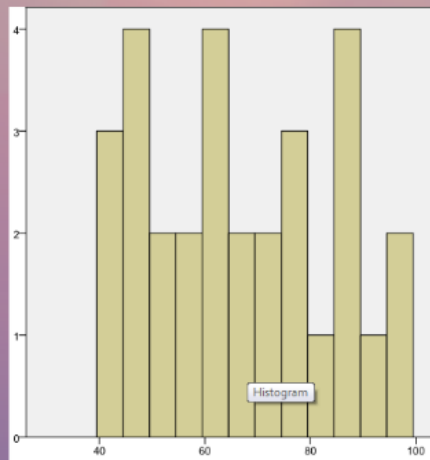
- popis rozložení pomocí percentilů (popis rozložení pomocí percentilů)
- popis rozložení pomocí percentilů (popis rozložení pomocí percentilů)
- popis rozložení pomocí percentilů (popis rozložení pomocí percentilů)
- popis rozložení pomocí percentilů (popis rozložení pomocí percentilů)

Rozložení četnosti

- poté, co jsme ve výzkumu získali nějaká data, je první informací (statistikou), která nás zajímá, četnost výskytu jednotlivých hodnot (resp. hodnot uvnitř jednotlivých intervalů)
- konfiguraci četností nazýváme rozložení (rozdělení).
- rozložení popisujeme
 - tabulkou četností
 - graficky – histogram, sloupcový diagram
 - pomocí percentilů
- o typu, tvaru rozložení hodnot proměnné uvažujeme většinou graficky – histogram, sloupcový diagram.

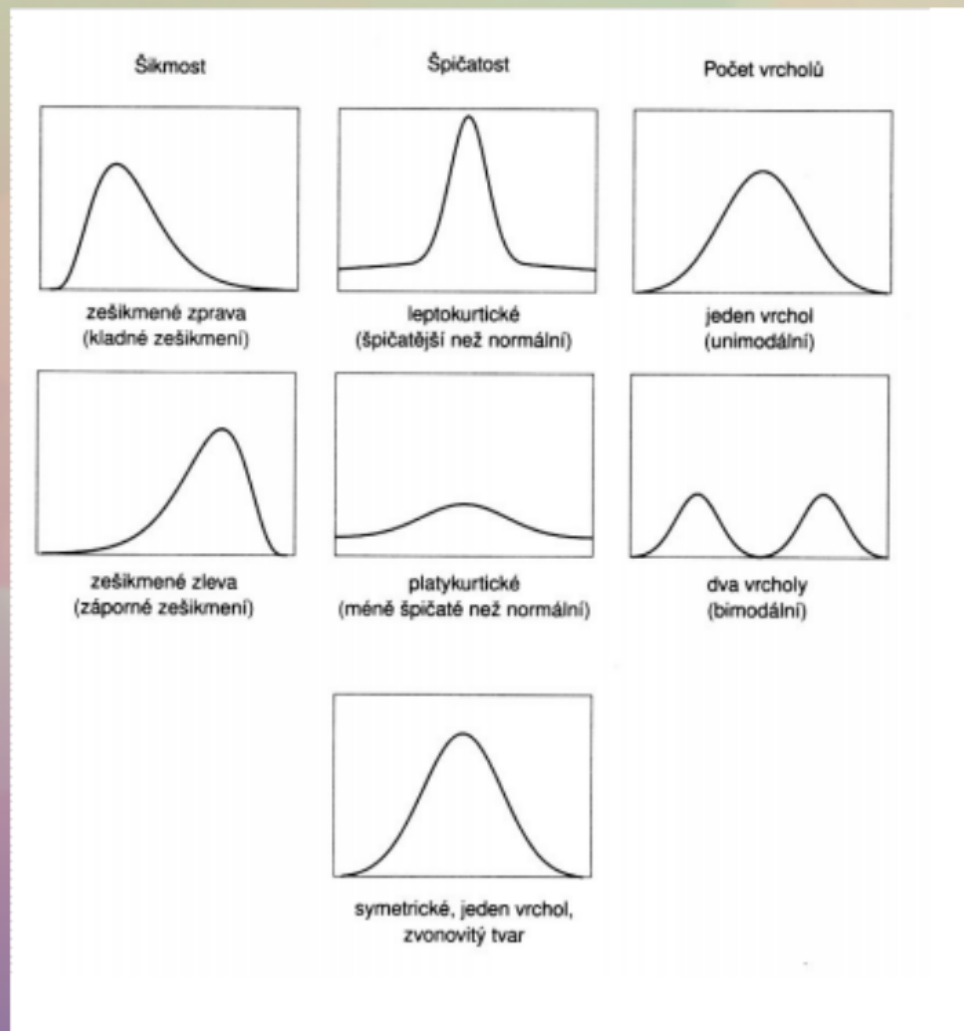
Rozložení četnosti

- Měřené jevy jsme si nějak rozdělili do kategorií (či intervalů) a tyto kategorie jsou různě „populární“ – četné.
- Četnosti u reálných ordinálních a vyšších proměnných obvykle nebývají distribuovány nahodile – jejich rozdělení zobrazené histogramem má popsateľný tvar
- Rozložení četnosti je často možné popsat slovy a nejlépe je vidět v grafech



Tvar rozložení četnosti

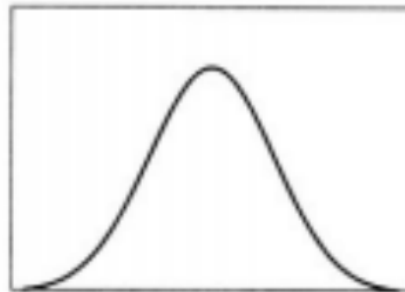
- Normální
- Uniformní
- Počet vrcholů
 - Unimodální, bimodální, multimodální
- Zešikmení
 - Zešikmené zprava (pozitivně), efekt podlahy
 - Zešikmené zleva (negativně), efekt stropu
- Strmost
 - Leptokurtické
 - Platykurtické



Normální rozložení

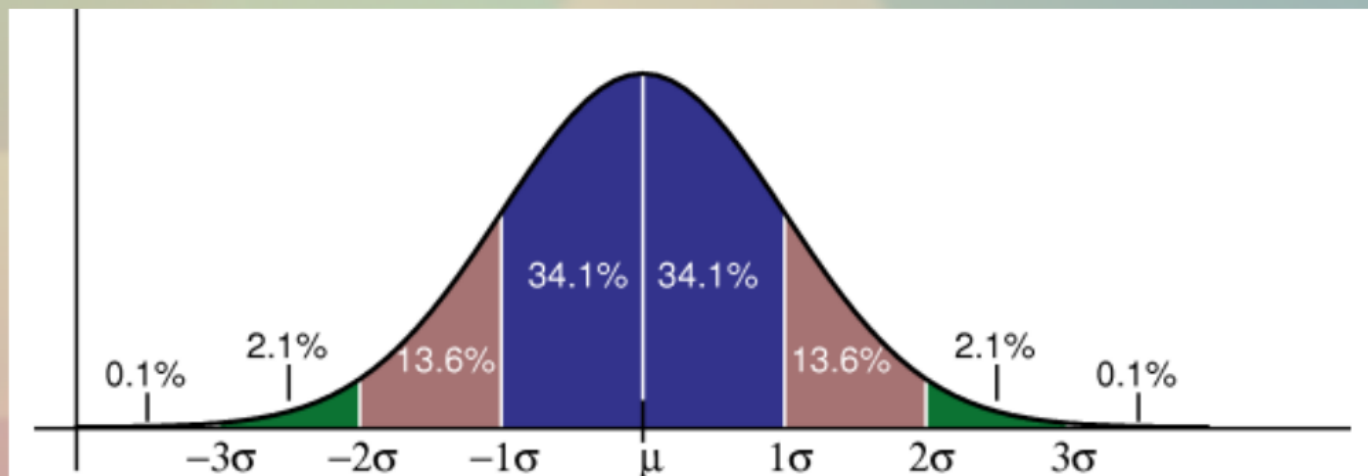
(Gaussovo rozložení)

- „Normální“ ve smyslu „velmi běžné“
- Je to rozložení, které očekáváme u jevů, které jsou způsobeny velkým množstvím faktorů - to způsobuje, že značná část výsledků se soustřeďuje kolem průměrné hodnoty
- Nesouvisí s „kvalitou“ dat
- I proměnné, které nemají normální rozložení, jsou normální
- v angličtině: bell curve (zvonová křivka)
- př. výška, váha, IQ, měření výkonu žáků v tělesné výchově, ...



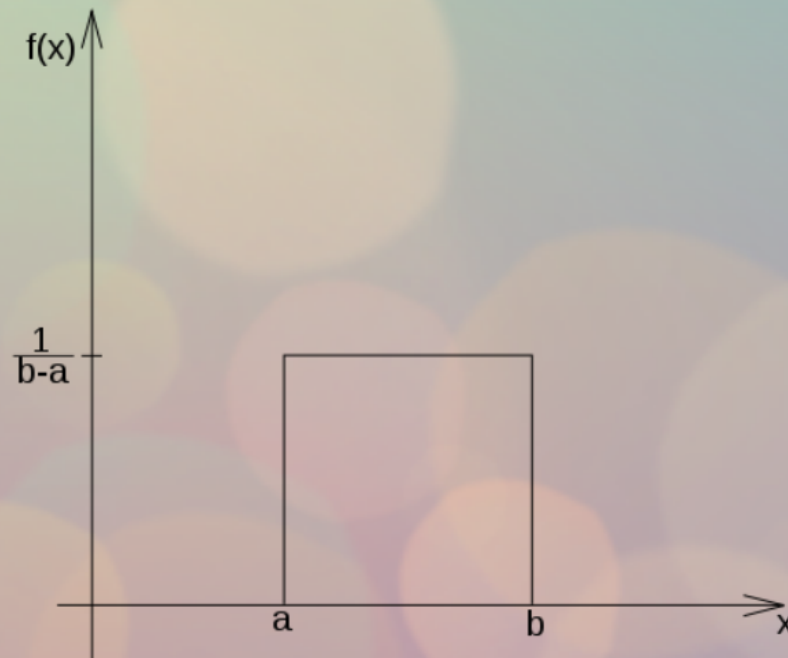
symetrické, jeden vrchol,
zvonovitý tvar

Normální rozložení



Uniformní rozložení

- všechny hodnoty jsou stejné



Počet vrcholů

- obvykle je pouze jeden vrchol (středních hodnot je hodně, extrémních málo), ale může jich být více

1 vrchol = unimodální

2 vrcholy = bimodální

více vrcholů = multimodální

(většinou spíše chyba měření)

Šikmost

- míra symetrie rozložení hodnot proměnné
- šikmost rovnající se nule indikuje normální rozložení

- zešikmení zprava je pozitivní
 - polovina menších hodnot je méně rozptýlena než polovina větších hodnot
 - pravá strana má delší konec než strana levá

- zešikmení zleva je negativní
 - polovina menších hodnot je více rozptýlena než polovina větších hodnot
 - levý konec je delší než pravý

- pochybné zešikmení:
 - **efekt stropu** - většina výsledků se pohybuje kolem maximální hodnoty, protože škála málo diferencuje v horní polovině výkonu (zešikmení zleva, "test" je velmi snadný, hodně lidí dosáhne nejlepšího výsledku)
 - **efekt podlahy** - většina výsledků se pohybuje kolem nuly, protože škála měření je málo senzitivní vůči slabším výkonům (asi jsme použili špatné měřítko)



Špičatost

- vyjadřuje stupeň koncentrace hodnot kolem střední hodnoty
- rozložení špičaté /ploché
- čím špičatější rozdělení, tím více jsou hodnoty soustředěny kolem jeho středu
- čím plošší rozdělení, tím častěji obsahuje hodnoty vzdálené od středu

**Špičaté rozložení
(leptokurtické)**



**Ploché rozložení
(platykurtické)**



Popis rozložení pomocí percentilů

- percentilová škála
každému dosaženému počtu bodů se přiřadí tzv. percentilové pořadí, které udává, kolik % testovaných osob dosáhlo horšího výkonu
umožňuje posoudit relativní pořadí určitého jedince ve skupině

X-tý percentil

- hodnota, pro kterou platí, že X % lidí (jevů) ve vzorku má/ získalo tuto nebo menší hodnotu
- lze snadno odečíst z kumulativního histogramu či patřičného sloupce tabulky četností

Typicky rozložení popisujeme

- 10., 20., ..., 80., 90. percentilem – obecně
- min, 25., 50., 75., max – nejčastěji
- min., 1., 5., 10., 25., 50., 75., 90., 95., 99. – v normách

