PSY117/454
Statistická analýza dat v psychologii
Přednáška 7

Počet pravděpodobnosti

Je známo, že když muž použije jeden z okrajových pisoárů, sníží se pravděpodobnost, že bude pomočen o 50%.

anonym

Pravděpodobnost jevu

- Pravděpodobnost, že nastane jev A
  - jistý jev: P = 1
  - nemožný jev: P = 0
  - jisté a nemožné jevy se vyskytují pouze v teorii
- 2 pojetí pravděpodobnosti
  - subjektivní jistota
  - četnostní (statistická): z m náhodných pokusů nastal jev A n-krát
    - P(A) = n / m , blíží-li se počet pokusů ∞ (populaci)

AJ: probability, event, random trial

Jevy a náhodné pokusy

- Jevy
  - ≈ hodnoty proměnných – např. Petr má IQ = 150
  - vzorek 15 IQ (lidí) – 15 jevů
  - ...a jejich kombinace (složené jevy)
    - náhodné vs. deterministické, 2: neslučitelné(disjunktní), ekvivalentní
    - doplňkový jev ($\overline{A}$)
- Pole jevů
  - množina hodnot, kterých může proměnná/é nabývat
  - ≈ proměnná
- Náhodný pokus
  - situace, kdy z pole jevů může nastat jeden nebo více jevů
  - ≈ výběr a změření člověka, hod kostkou
  - nelze určit, který jev nastane & lze opakovat <u>bez vzájemného ovlivňování</u>

Náhodným pokusem získáváme z pole jevů jev.

AJ: event, sample space, random trial, random vs. deterministic events, mutally exclusive events, equivalent events

Počítání s pravděpodobnostmi

- „NEBO“ – součet jevů - nastane jev A nebo jev B [nebo oba nejsou-li disjunktní]
  - $P(A \cup B) = P(A) + P(B) - [P(A \cap B)]$
    - *př.* náhodně vybraný člověk je žena nebo psycholog
    - *př. disj.* náhodně vybraný člověk má základní vz. nebo je vyučen .
- „A“ – součin jevů - nastane jev A a zároveň nastane jev B
  - $P(A \cap B) = P(A) . P(B)$
    - *př.* náhodně vybraný člověk je psycholožka (pohlaví=žena, povolání=psychologie)
- Kombinatorika – velikost pole jevů
  - permutace n prvků = n!
  - kombinace r prvků z n-prvkové množiny = n! / r!(n-r)!
- Šance – odds ratio - častý způsob vyjádření pravděpodobnosti
  - šance Komety na vítězství jsou 1:10
  - $= P(A) / P(\bar{A})$

Podmíněná pravděpodobnost

- Jaká je pravděpodobnost jevu A, pokud nastal jev B?
  - $P(A|B)$ … $P(A \cap B) = P(A|B) \cdot P(B)$
    - *př.* kouří-li člověk (jev B), je riziko onemocnění rakovinou (jev A) 40% … $P(A|B)$
    - Kuřáků je 30, tedy $P(B) = 0{,}3$.
    - P, že náhodně vybraný člověk je kuřák, kt. onemocní rakovinou je 12%
- Bayesův teorém
  - přepočet mezi $P(A|B)$ a $P(B|A)$
  - $P(A|B) = P(A).P(B|A) / [P(A).P(B|A) + P(\bar{A}).P(B|\bar{A})]$
    - *př.* v Hendlovi (s.120)
    - př. Test na LMD má 15% chybovost ($P(T+|L+)=0{,}85$; $P(T+|L-)=0{,}15$)
    - Prevalence LMD je 10%. ($P(L+)=0{,}1$)
    - P., že dítě s pozitivním výsledkem v testu má LMD? $P(L+|T+)=?$
    - $0{,}1 \cdot 0{,}85 / (0{,}1 \cdot 0{,}85 + 0{,}9 \cdot 0{,}15) =$ **0,39**

AJ: conditional probability, Bayes's theorem

Pravděpodobnostní rozložení náhodné proměnné

Je-li **proměnná náhodná** (tj. její hodnoty lze považovat za výsledek náhodných pokusů)… …jaká je *P* výskytu jednotlivých možných hodnot?

- Vzpomeňme si, že *P*(A) = n / m , blíží-li se počet pokusů ∞ (populaci)
- Máme-li tedy dost velký, náhodně vybraný vzorek, pak *P* výskytu jednotlivých hodnot ≈ jejich relativní četnost

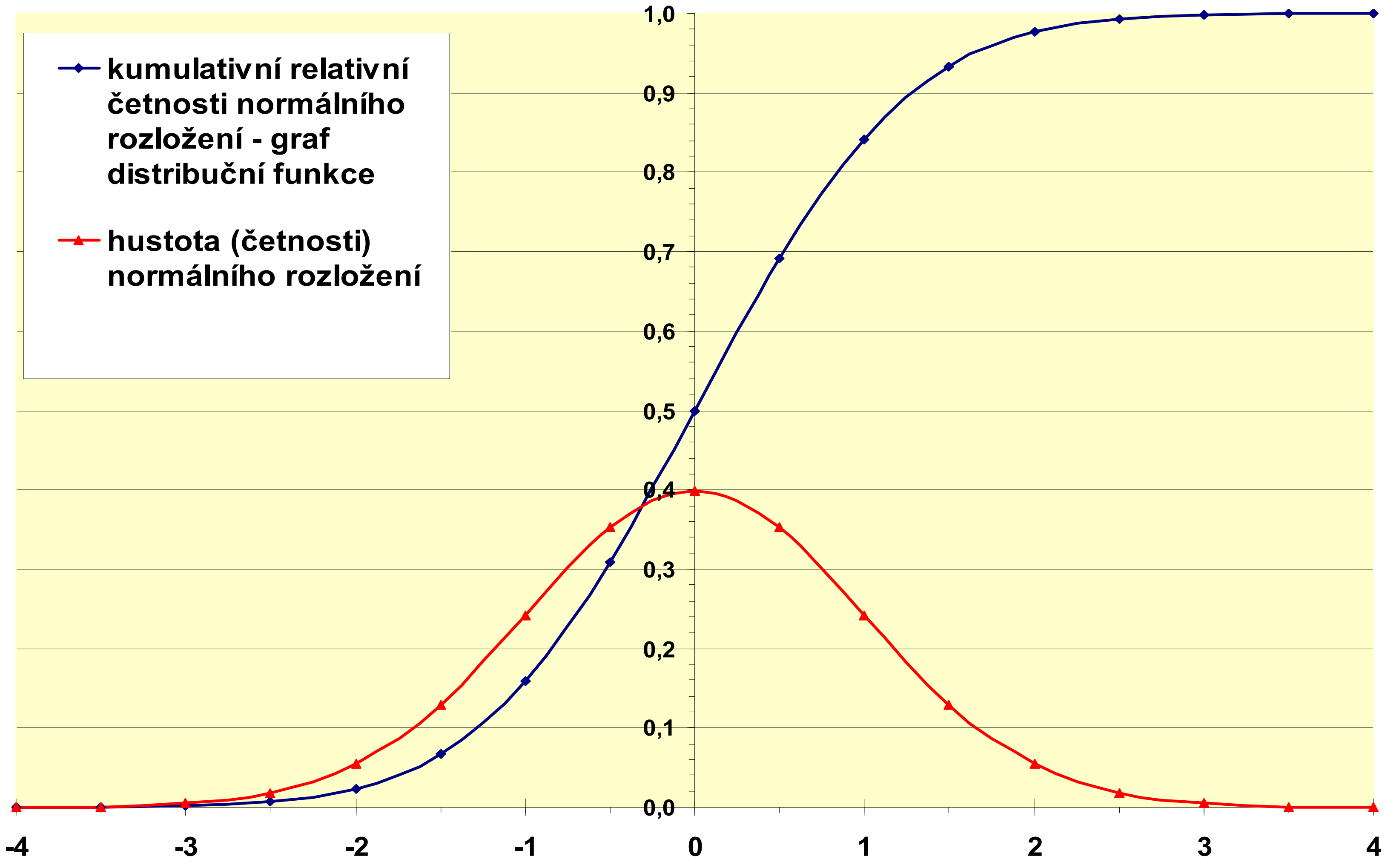
Pravděpodobnostní rozložení = <u>teoretické</u> rozložení relativních četností

- U spojitých proměnných neuvažujeme o *P* výskytu jednotlivých hodnot (∞), ale spíše o *p* výskytu hodnot v intervalech – **hustota pravděpodobnosti**
- U diskrétních proměnných uvažujeme o *P* výskytu jednotlivých hodnot.

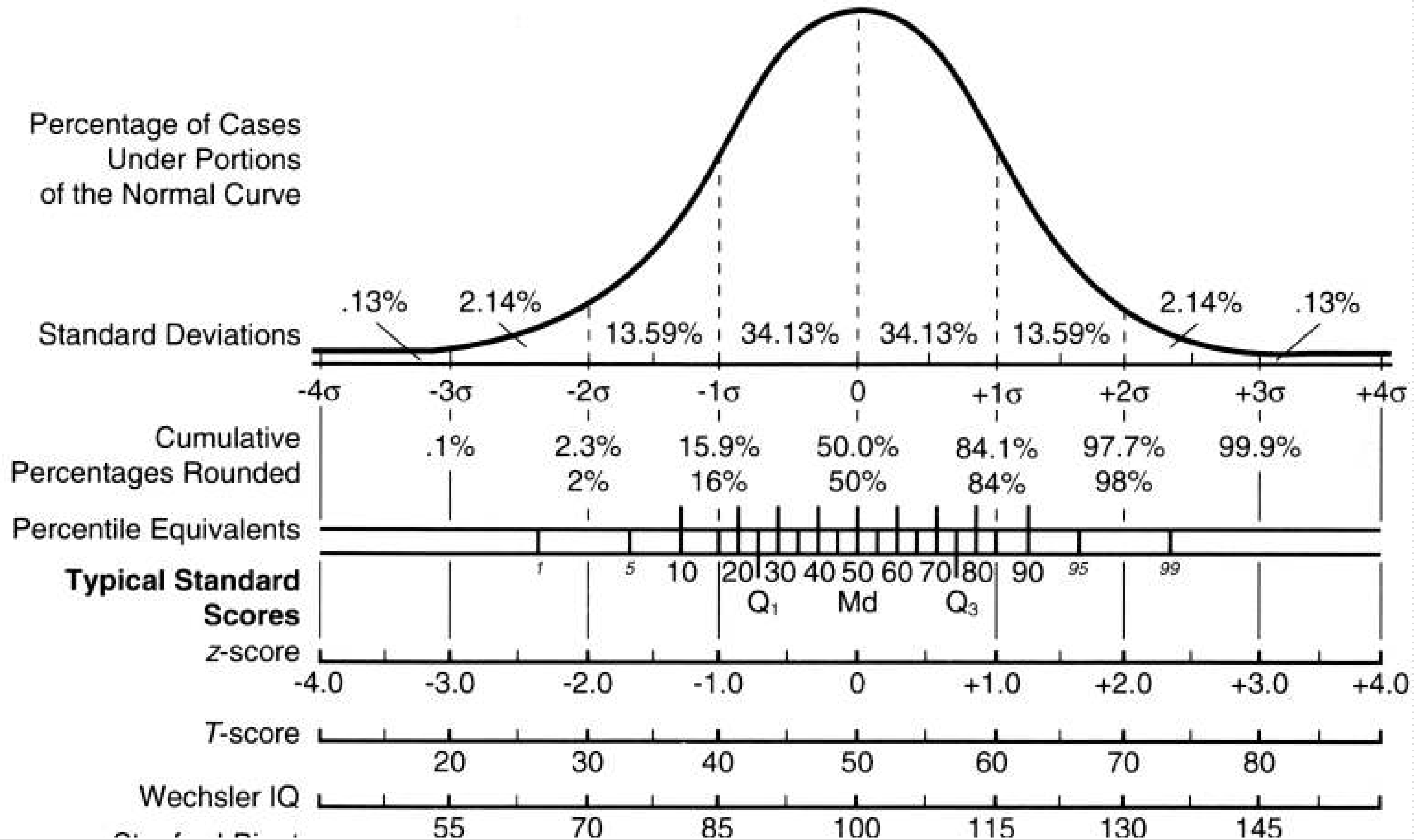
- P-nostní rozložení je popsáno **distribuční funkcí**
  - F(x) = *P*(X≤x) tj. *P* výskytu hodnot ≤ x
  - Tato *P* je rovna „ploše oblasti pod křivkou“

AJ: random variable, probability distribution, distribution function, probability density

Normální rozložení



P v normálním rozložení



Důležitá p-nostní rozložení

- Normální
- Poissonovo
- Studentovo t- rozložení
- Fisherovo F-rozložení
- χ^2-rozložení (chí-kvadrát)
- Binomické (Hendl, 134)

Vyjma binomického se všechna uvedená rozložení používají jako přibližné (asymptotické) ideály, jimž by se rozložení našich proměnných (statistik) blížilo, kdybychom měli obrovský a reprezentativní vzorek.

AJ: random variable, probability distribution, distribution function