Štatistika 1
Pravdepodobnosť II.

5. prednáška

Niektoré typy rozdelení

- diskrétnej náhodnej premennej:
  - binomické
  - hypergeometrické
  - Poissonovo
- spojitej náhodnej premennej:
  - normálne (Laplace-Gaussovo rozdelenie)
  - t-rozdelenie (Studentovo rozdelenie)
  - χ^2 rozdelenie (chí-kvadrát rozdelenie)
  - F-rozdelenie (Fisherovo rozdelenie)

Rozdelenia diskrétnej náhodnej premennej

Binomické rozdelenie ($X\sim Bi[n;p]$)

- predpoklady:
  - n nezávislých pokusov
  - v každom pokuse nastane jav A s pravdepodobnosťou p a jav A^C s pravdepodobnosťou $1\text{-}p$ (tzv. výber s vrátením)

$\Omega=\{0,1\}^n$, elementárny jav je $\omega=(\omega_1,\omega_2,\dots,\omega_n)$

kde ω_i je počet nastatí javu A v i-tom pokuse, teda 0 alebo 1

Náhodná veličina $X(\omega)=\sum_{i=1}^{n}\omega_i$

počet výskytov javu A pri n nezávislých pokusoch má **binomické rozdelenie s parametrami *n* a *p***

(n – počet pokusov, p – pravdepodobnosť nastatia javu A pri jednom pokuse)

Pravdepodobnosť, že pri n pokusoch nastane jav práve x –krát je daná vzťahom:

$$P(X=x)=\binom{n}{x}p^x(1-p)^{n-x},$$

- $E[X] = n.p$
- $D[X] = n.p.(1\text{-}p)$

Hypergeometrické rozdelenie $Hy[N;M;n]$

- je rozdelením pri **výbere bez vrátenia**: náhodne vybrané jednotky nevraciame späť do základného súboru, **jednotlivé pokusy sú závislé.**
- predpoklady:
  - máme konečný súbor N prvkov, z ktorých M má určitú sledovanú vlastnosť V
  - z tohto súboru vyberieme naraz alebo postupne n prvkov bez vrátenia

Náhodná veličina X je počet prvkov vo výbere s vlastnosťou V.

Pravdepodobnosť, že medzi *n* vybranými prvkami bude práve *x* prvkov so sledovanou vlastnosťou vypočítame

$$P(X=x)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}},$$

- pre strednú hodnotu a disperziu platí

$$E[X]=n\frac{M}{N},\quad D[X]=\frac{nM(N-M)}{N^2}\frac{N-n}{N-1}.$$

Poissonovo rozdelenie ($X\sim$ Po(λ))

sa používa v súvislosti s **poissonovskými prúdmi javov.**

Prúd javov je postupnosť náhodných javov, ktoré:

1. nastanú v náhodných časových okamihoch určitého časového intervalu,
2. pravdepodobnosť výskytu javu v určitom časovom intervale závisí len od dĺžky tohto intervalu,
3. výskyt dvoch alebo viacerých javov počas veľmi malého časového intervalu je prakticky nemožný,
4. pravdepodobnosť nastatia určitého počtu javov za určitý časový interval nezávisí od počtu nastatia javov v iných časových intervaloch

- Náhodná premenná X má Poissonovo rozdelenie pravdepodobnosti s parametrom $\lambda>0$, ak nadobúda hodnoty x = 0, 1,2,... s pravdepodobnosťou

$$P(X=x)=\frac{\lambda^x}{x!}e^{-\lambda}$$

- platí

$$E[X]=\lambda \quad \text{a} \quad D[X]=\lambda$$

Siméon Denis Poisson
(1781-1840)

Binomické rozdelenie môžeme **aproximovať** Poissonovým rozdelením s parametrom $\lambda = n.p$,

ak je počet pokusov (rozsah výberu) n dostatočne veľký ($n > 30$)

a pravdepodobnosť p veľmi malá ($p \leq 0{,}1$).

Rozdelenia spojitej náhodnej premennej

Normálne rozdelenie ($X\sim$ N[$\mu;\sigma^2$])

- náhodná premenná X má normálne rozdelenie s parametrami $\mu\in R$ a $\sigma^2>0$, ak jej funkcia hustoty pravdepodobnosti je daná vzťahom

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

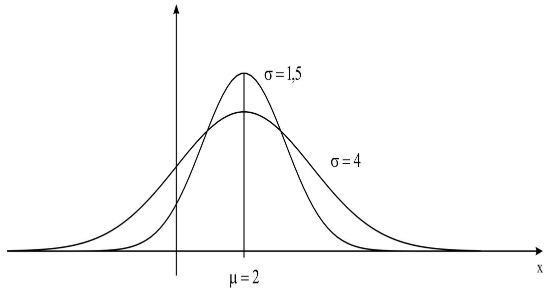
- platí

$$E[X]=\mu \quad \text{a} \quad D[X]=\sigma^2$$

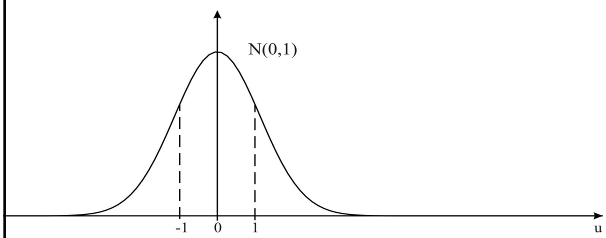
- grafom funkcie hustoty je tzv. Gaussova krivka
- je symetrická okolo osi v bode μ
- inflexné body sú $\mu \pm \sigma$
- obojstrannou asymptotou so smernicou je priamka $y = 0$

Karl Friedrich Gauss (1777-1855)

Graf hustoty normálneho rozdelenia $N[2;1,5^2]$ a $N[2;4^2]$



Normované normálne rozdelenie N[0;1]



- normovaná náhodná premenná: $U = \dfrac{X - \mu}{\sigma}$

Štatistické tabuľky

V štatistických tabuľkách sa uvádzajú kvantily (kritické hodnoty) normovaného normálneho rozdelenia:

pre $\alpha \in (0;1)$ je číslo $u(\alpha)$ α -kvantilom normovaného normálneho rozdelenia, ak platí

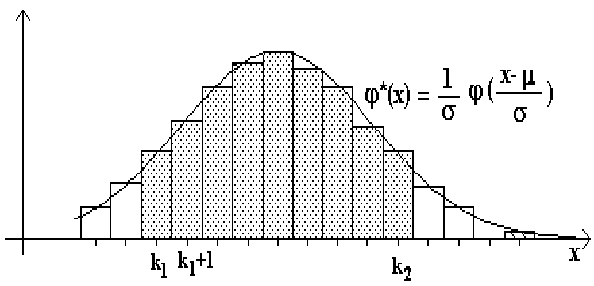
$$P\big(X < u(\alpha)\big) = \alpha$$

Moivrova – Laplaceova integrálna veta

- Nech X je náhodnou premennou, ktorá má binomické rozdelenie s parametrami n a p, potom pravdepodobnosť toho, že X nadobúda hodnoty z intervalu $< a; b>$ môžeme približne vypočítať podľa vzťahu

$$P(a \le X \le b) \approx \Phi(\beta) - \Phi(\alpha)$$

kde $\alpha = \dfrac{a - 0,5 - np}{\sqrt{npq}}$ a $\beta = \dfrac{b + 0,5 - np}{\sqrt{npq}}$



χ^2- rozdelenie $\chi^2[\nu]$

- Náhodná premenná X má **χ^2- rozdelenie** s ν-stupňami voľnosti ($\upsilon \geq 1$), ak pre hustotu jej rozdelenia platí

$$f(x) = \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}, \quad \text{pre } x>0 \text{ a } \quad f(x)=0 \ \text{ inak}$$

- kde $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$ je tzv. gama-funkcia,

definovaná pre kladné hodnoty a

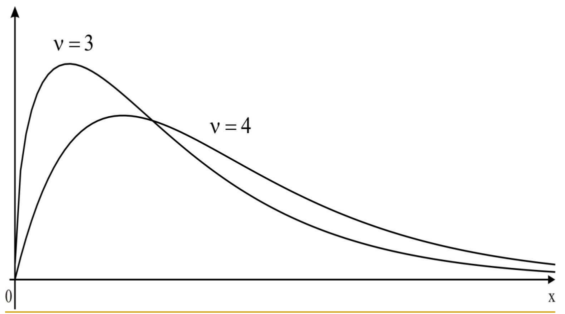
- Majme náhodné premenné X_1, X_2, ….,X_ν, ktoré majú **normované normálne rozdelenie**, potom náhodná premenná Y , ktorá vznikne ako súčet štvorcov premenných X_i

$$Y = \sum_{i=1}^{\nu} X_i^2$$

bude mať χ^2 (chí-kvadrát) rozdelenie s υ stupňami voľnosti, υ – stupne voľnosti (degrees of freedom) sú jediným parametrom a vyjadrujú počet nezávislých sčítancov

- platí E[X] = υ a D[X] = 2υ.

Graf hustoty χ^2 rozdelenia



Kritická hodnota χ^2 rozdelenia

- pre $\alpha \in (0;1)$ je číslo $\chi^2_\upsilon(\alpha)$ α-kritickou hodnotou rozdelenia, ak platí

$$P(X > \chi^2_\nu(\alpha)) = \alpha$$

Studentovo rozdelenie (t–rozdelenie) t[ν]

- Náhodná premenná X má ***t*- rozdelenie** s ν- stupňami voľnosti ($\upsilon \geq 1$), ak pre hustotu jej rozdelenia platí

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}} \left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

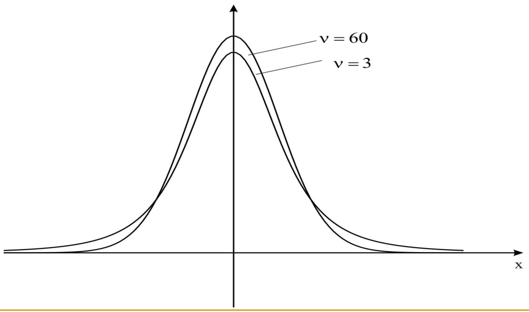
- ďalej platí $E[X] = 0$ (pre $\nu > 1$), $D[X] = \frac{\nu}{\nu-2}$ (pre $\nu > 2$)

Majme náhodnú premennú X s normovaným normálnym rozdelením a náhodnú premennú Y, ktorá má χ^2 rozdelenie s ν stupňami voľnosti, pričom tieto náhodné premenné sú navzájom nezávislé. Potom náhodná premenná

$$T = \frac{X}{\sqrt{\frac{Y}{\nu}}}$$

má Studentovo rozdelenie s ν stupňami voľnosti.

Graf hustoty t-rozdelenia



Kritická hodnota t-rozdelenia

- pre $\alpha \in (0;1)$ je číslo $t_\nu(\alpha)$ α-kritickou hodnotou rozdelenia, ak platí

$$P(|T| > t_\nu(\alpha)) = \alpha$$

Fisherovo – Snedecorovo rozdelenie (F - rozdelenie) F[ν_1; ν_2]

- Náhodná premenná má Fisherovo - Snedecorovo rozdelenie, F-rozdelenie s υ_1, υ_2 stupňami voľnosti ($\upsilon_1 \geq 1, \ \upsilon_2 \geq 1$), ak je jeho hustota vyjadrená nasledujúcim vzťahom

$$f_{\upsilon_1,\upsilon_2}(x) = \frac{\Gamma\left(\frac{\upsilon_1+\upsilon_1}{2}\right)}{\Gamma\left(\frac{\upsilon_1}{2}\right)\Gamma\left(\frac{\upsilon_2}{2}\right)}\left(\frac{\upsilon_1}{\upsilon_2}\right)^{\frac{\upsilon_1}{2}} x^{\frac{\upsilon_1}{2}-1}\left(1+\frac{\upsilon_1}{\upsilon_2}x\right)^{\frac{-\upsilon_1+\upsilon_2}{2}}$$

pre $x > 0$ a je rovná nule inak

Majme dve nezávislé náhodné premenné Y_1 s χ^2 rozdelením s ν_1 stupňami voľnosti a Y_2 s ν_2 stupňami voľnosti.

Náhodná premenná

$$F = \frac{\frac{Y_1}{\nu_1}}{\frac{Y_2}{\nu_2}}$$

má F –rozdelenie s ν_1 a ν_2 stupňami voľnosti

Graf hustoty F-rozdelenia