Alapfogalmak

1. *Statisztikai mező*: (**Ω**, ***A***, ***P***), ahol (**Ω**, ***A***) mérhető tér és ***P*** valószínűségi mértékek egy családja. megfigyelés(eke)t végzünk.
2. *Minta*: X: (**Ω**, ***A***) → (***X***, ***B***) mérhető leképezés.
   Mintatér: az (***X***, ***B***) mérhető tér.
   A mintát gyakran úgy kapjuk, hogy ugyanazt a kísérletet azonos körülmények között *n*-szer megismételjük (*n elemű független minta*) Ekkor X koordinátái függetlenek és azonos eloszlásúak.
3. A mértékcsaládra gyakran $\boldsymbol{P} = \{P_\vartheta : \vartheta \in \Theta\}$ alakban hivatkozunk, ahol Θ paramétertér.
   *Paraméteres statisztikai problémá*ról beszélünk, ha $\Theta \subset \boldsymbol{R}^s$, azaz a család leírható néhány paraméterrel (pl. normális eloszlás stb.).
   Nemparaméteres probléma: ha ***P*** „bő" (pl. a folytonos eloszlások családja)
4. *Becslési probléma* esetén feladatunk, hogy a paramétert vagy annak valamilyen függvényét a minta segítségével közelítsük
   Hipotézisvizsgálati probléma esetén csak arról kell döntenünk, hogy egy bizonyos állítás (hipotézis) teljesül-e az igazi paraméterértékre.
5. *Statisztikák*nak nevezzük a minta különböző függvényeit.

Definíció: Az X_1, X_2, ... X_n mintavételi változókból valamilyen algebrai vagy egyéb utasítással képzett valószínűségi változókat ***statisztikai függvények***nek, vagy röviden ***statisztikák***nak nevezzük.

A valószínűségi változóknál megismert jellemzők (eloszlás függvény, várható érték, szórás stb) mindegyike előállítható ***empírikus alakban*** a mintából, alkalmas statisztikai függvények segítségével. Az így kapható empírikus jellemzők tehát maguk is valószínűségi változók, értékük mintáról mintára más és más lehet.

4.2. A minta néhány statisztikai jellemzője

4.2.1. A tapasztalati várható érték, a minta középértéke

Definíció: A mintaelemek számtani átlagát, azaz az

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

valószínűségi változót a ***minta középértékének***, vagy ***empírikus (tapasztalati) várható értéknek*** nevezzük.

u-próba
, röviden statisztika:

$$u = \frac{\overline{\xi} - m}{\sigma}\sqrt{n}$$

Az u statisztika, mivel egy normális eloszlású valószínűségi

változóra vonatkozó minta mintaközepének – $\overline{\xi}$ $N\left(m, \frac{\sigma}{\sqrt{n}}\right)$ – standardizáltja,

standard normális eloszlású.

DEFINÍCIÓ: χ^2-ELOSZLÁS

Legyen $\xi_1, \xi_2, ..., \xi_n$ n darab független, $N(0,1)$ eloszlású valószínűségi változó. Ezekből képezünk egy új valószínűségi változót:

$$\chi^2 = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2 .$$

χ^2 eloszlását n szabadsági fokú χ^2-eloszlásnak nevezzük.

A χ^2-eloszlás sűrűségfüggvénye

$$f_n(x) = \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}, \quad x>0.$$

A χ^2-eloszlás várható értéke

$$M\left(\chi^2\right) = n .$$

A χ^2-eloszlás szórásnégyzete

$$D^2(\chi^2) = 2n .$$

DEFINÍCIÓ: STUDENT-ELOSZLÁS

Legyenek η, ξ_1, $\xi_{2,\dots,}$ ξ_n független, $N(0,1)$ eloszlású valószínűségi változók. Ezekből képezünk egy új valószínűségi változót az alábbi képlet szerint:

$$t = \frac{\sqrt{n}\eta}{\sqrt{\xi_1^2 + \xi_2^2 + \ldots + \xi_n^2}}.$$

A t valószínűségi változó eloszlását n szabadsági fokú Student-eloszlásnak nevezzük. Az eloszlást néha *t-eloszlásnak* is nevezik.

A t-eloszlás sűrűségfüggvénye

$$h(x) = \frac{1}{\sqrt{\pi n}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(\frac{x^2}{n}+1\right)^{\frac{n+1}{2}}}.$$

A függvény értelmezési tartománya: $-\infty < x < \infty$.

A Student-eloszlás várható értéke

A Student-eloszlás egy szimmetrikus eloszlás a *0* pontra nézve, tehát a Student-eloszlás várható értéke (csak $n \geq 2$ esetén létezik) $M(t)=0$. Könnyen belátható, hogy $n=1$ esetén nem létezik várható érték.

A Student-eloszlás szórásnégyzete

$$D^2(t) = M(t^2) - M^2(t) = M(t^2) =$$

$$M(n\eta^2)M\left(\frac{1}{\xi_1^2+\xi_2^2+\ldots+\xi_n^2}\right) = \frac{n}{n-2}, \text{ ha } n \geq 3.$$

DEFINÍCIÓ: *F*-ELOSZLÁS

Legyenek a $\xi_1, \xi_2, \ldots, \xi_m$ és az $\eta_1, \eta_2, \ldots, \eta_n$ valószínűségi változók függetlenek és $N(0,1)$ eloszlásúak. Az ezekből képezett F új valószínűségi változót az alábbi összefüggés definiálja:

$$F = \frac{\frac{1}{m}\sum_{i=1}^{m}\xi_i^2}{\frac{1}{n}\sum_{i=1}^{n}\eta_i^2}.$$

Az F valószínűségi változó eloszlását m, n szabadsági fokú F-eloszlásnak nevezzük.

Az F eloszlás sűrűségfüggvénye:

$$f_{m,n}(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \frac{x^{\frac{m}{2}-1}}{(1+x)^{\frac{m}{2}-1}}, \quad x > 0.$$