Valószínűségszámítás és statisztika előadás informatika BSC/B szakosoknak

2021/2022 2. félév

Zempléni András
andras.zempleni@ttk.elte.hu
http://zempleni.elte.hu

Példák

$F_{aX+b}(z) = F_X((z-b)/a)$, ha $a>0$ és

$F_{aX+b}(z) = 1-F_X((z-b)/a)$, ha $a<0$.

Ebből adódik, hogy ha X abszolút folytonos, és $g(z)=az+b$, akkor $g(X)$ sűrűségfüggvénye

$f_{aX+b}(z)=f_X((z-b)/a)/|a|$.

Általános eredmény: ha g szigorúan monoton, folytonosan deriválható, $g' \neq 0$, akkor

$$f_{g(X)}(z) = \frac{f_X(g^{-1}(z))}{|g'(g^{-1}(z))|}$$

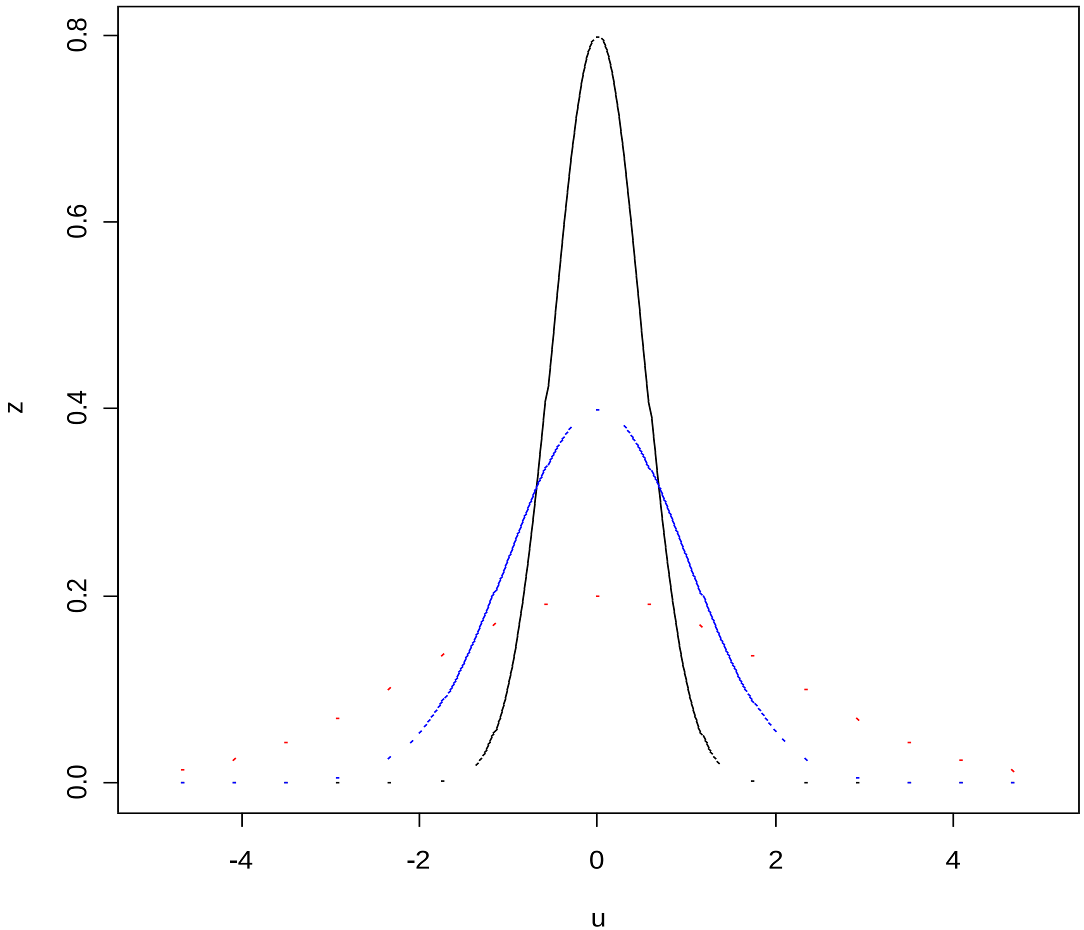
A normális eloszlás

- Legyen m tetszőleges, σ pedig pozitív valós szám. Ha X standard normális eloszlású, akkor az $Y = \sigma X + m$ valószínűségi változó (m,σ) paraméterű normális eloszlású. Ennek sűrűségfüggvénye az

  $f_{aX+b}(z)=f_X((z-b)/a)/|a|$ képletből

$$f_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

Normális eloszlások sűrűségfüggvénye (m=0)



Abszolút folytonos eloszlású valószínűségi változók várható értéke

- Egyre finomabb felosztással közelítve a folytonos eloszlást $E(X) \approx \Sigma y P(y<X<y+\delta) \approx \Sigma y \delta f(y) \approx \int y f(y) dy$
- Ebből a definíció: az abszolút folytonos

  eloszlású X várható értéke:
  ha az integrál létezik.

$$E(X) = \int_{-\infty}^{\infty} y f_X(y) dy$$

Tulajdonságok, példák

- Mivel a diszkrét esetből határátmenettel kaptuk a fogalmat, a tulajdonságok (pl. E(aX+b)=aE(X)+b, E(X+Y)=E(X)+E(Y) stb.) most is érvényben maradnak.
- Függvény várható értéke: Legyen X sűrűségfüggvénye f és Y=g(X). Ekkor

$$E(Y) = \int_{-\infty}^{\infty} g(y) f_X(y) dy$$

- Ha X egyenletes eloszlású az [a,b]-ben, akkor

$$E(X) = \int_a^b \frac{y}{b-a} dy = \left[\frac{y^2}{2(b-a)} \right]_a^b = \frac{a+b}{2}$$

További példák

- Ha X exponenciális eloszlású λ paraméterrel, akkor

$$E(X)=\int_0^\infty \lambda y e^{-\lambda y}\,dy=\left[-ye^{-\lambda y}\right]_0^\infty+\int_0^\infty e^{-\lambda y}\,dy=\frac{1}{\lambda}$$

- Ha X standard normális eloszlású, akkor

$$E(X)=\int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,dx=\left[-\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\right]_{-\infty}^{\infty}=0$$

- Ha a Z változó Q_Z eloszlása keverék-eloszlás (azaz pl. p valószínűséggel X-et, 1-p valószínűséggel Y-t figyeljük meg), akkor E(Z)=pE(X)+(1-p)E(Y).

További eloszlások

- Gamma eloszlás, sűrűségfüggvénye a pozitív x értékekre:
  (h pozitív egész, λ>0
  a paraméterek)
  $$\frac{\lambda\,(\lambda\,x)^{h-1}}{(h-1)!}\,e^{-\lambda\,x}.$$
- Lognormális eloszlás, sűrűségfüggvénye
  $$f_X(x;\mu,\sigma)=\frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}},\quad x>0$$
- Cauchy eloszlás , sűrűségfüggvénye
  $$f(x)=\frac{1}{\pi(1+x^2)}$$

A szórásnégyzet

- Mivel ez a várható értékből származtatott mennyiség, most is érvényes a

  $D^2(X):=E[(X-E(X))^2]$ definíció, illetve a

  $D^2(X)=E(X^2)-E^2(X)$ számítási módszer.

  A korábban látott tulajdonságok itt is érvényben maradnak.

Példák

- Ha X egyenletes eloszlású az [a,b] intervallumon, akkor

$$E(X^2)=\int_a^b \frac{y^2}{b-a}dy=\left[\frac{y^3}{3(b-a)}\right]_a^b=\frac{a^2+ab+b^2}{3}$$

$$D^2(X)=E(X^2)-E^2(X)=\frac{a^2+ab+b^2}{3}-\frac{(a+b)^2}{4}=\frac{a^2-2ab+b^2}{12}=\frac{(a-b)^2}{12}$$

- Ha X exponenciális eloszlású, akkor

$$E(X^2)=\int_0^\infty y^2\lambda e^{-\lambda y}dy=\left[-y^2e^{-\lambda y}\right]_0^\infty+\int_0^\infty 2ye^{-\lambda y}dy=\frac{2}{\lambda^2}$$

$$D^2(X)=E(X^2)-E^2(X)=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}$$

A normális eloszlás szórásnégyzete

- Legyen X standard normális eloszlású, akkor

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{\infty} x \cdot x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$$

parciálisan integrálva. Így $D^2(X)=1$. Ebből az (m,σ) paraméterű normális eloszlás szórásnégyzete: $D^2(\sigma X+m)= \sigma^2$, a szórása pedig σ.

A függetlenség karakterizációi

- Ha $\underline{X}$ koordinátái függetlenek, akkor definíció szerint
- $F_{\underline{X}}(\underline{z})=P(X_1<z_1, X_2< z_2,\dots, X_d<z_d)=F_1(z_1)F_2(z_2)\dots F_d(z_d)$ (minden $\underline{z} \in \boldsymbol{R}^d$ –re). Meg is fordítható: F szorzatelőállításából következik a függetlenség.
- Deriválva: a függetlenség abszolút folytonos változókra ekvivalens a sűrűségfüggvény $f_{\underline{X}}(\underline{z})=f_1(z_1)f_2(z_2)\dots f_d(z_d)$ alakú előállításával is.
- Példa: az egységnégyzeten egyenletes eloszlás sűrűségfüggvénye ($f(\underline{z})=1$ ha $0<\underline{z}<1$) előáll $f_1(z_1)f_2(z_2)$ alakban, ahol $f_i(z_i)=1$, ha $0<z_i<1$ ($i=1,2$), ez éppen a [0,1] intervallumon egyenletes eloszlás.

Tulajdonságok

- 1. Az $X_1,\ldots,X_n$ diszkrét valószínűségi változók függetlenek, ha $P(X_1 = x_1,\ldots, X_n = x_n) = P(X_1 = x_1)\cdot\ldots\cdot P(X_n = x_n)$ teljesül minden $x_1,\ldots,x_n$ értékre.
- 2. Ha az $X_1,\ldots,X_n$ valószínűségi változók függetlenek, a $g_1,\ldots,g_n$ függvények Borel-mérhetőek, akkor $g_1(X_1),\ldots, g_n(X_n)$ is függetlenek.
- 3. Ha az $X_1,\ldots,X_n$ valószínűségi változók függetlenek, a h k-változós Borel-mérhető fv., akkor $h(X_1,\ldots, X_k)$, $X_{k+1},\ldots,X_n$ is függetlenek.