Matematika B4
XI. gyakorlat

2005. november 23., 25.

1. Sűrűségfüggvény a síkon

Sűrűségfüggvény tulajdonságai:

1. $f(x,y) \geq 0$, minden x, y valós számra.
2. $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$$

Az A tartományba esés valószínűsége:

$$P(A) = \iint_A f(x,y)dxdy$$

Feladatok:

1. Az alábbi függvények melyike sűrűségfüggvény? (amelyik tartomány nincs megadva, ott a függvény 0.)

   a)
   $$f(x,y) = \frac{4}{5}(x + xy + y) \quad , \quad \text{ha} \quad 0 < x < 1, \quad 0 < y < 1$$

   b)
   $$f(x,y) = \lambda^2 e^{-\lambda(x+y)} \quad , \quad \text{ha} \quad x > 0, \quad y > 0$$

   c)
   $$f(x,y) = 4xy - 10 \quad , \quad \text{ha} \quad x^2 + y^2 < 1$$

   d)
   $$f(x,y) = \frac{1}{x} \quad , \quad \text{ha} \quad 0 < y < x < 1$$

2. Határozzuk meg c-t úgy, hogy f(x,y) sűrűségfüggvény legyen:
   $$f(x,y) = cy \quad , \quad \text{ha} \quad x > 0, \quad y > 0, \quad x + y < 1$$

3. Vegyük az első feladat b.) részében megadott $f(x) = \lambda^2 e^{-\lambda(x+y)}$ függvényt. Számítsuk ki az alábbi események valószínűségét:

   a) $0 < X < 1$ és $0 < Y < 1$
   b) $1 < X < 5$ és $2 < Y < 8$
   c) $0 < X < 1$
   d) $3 < Y < 5$

4. Vegyük a következő 2-dimenziós valószínűségi változót:
   Első koordinátája legyen egy véletlen szám: $X = RND_1$. A másik koordinátája pedig ez az érték beszorozva egy másik véletlen számmal: $Y = RND_1 \cdot RND_2$. Valamint definiáljuk a következő eseményeket: $A = \{X > \frac{1}{2}\}$ és $B = \{Y < \frac{1}{2}\}$.

   a) Ellenőrizzük, hogy ennek a sűrűségfüggvénye:
   $$f(x,y) = \frac{1}{x} \quad , \quad \text{ha} \quad 0 < y < x < 1$$
   b) $P(A) =?$
   c) $P(B) =?$
   d) $P(A \text{ és } B) =?$
   e) $P(A|B) =?$
   f) $P(B|A) =?$

5. Vegyük a következő 2-dimenziós valószínűségi változót:
   Első koordinátája legyen $X = \sqrt{RND_1}$. A másik koordinátája pedig ez az érték beszorozva egy véletlen számmal: $Y = \sqrt{RND_1} \cdot RND_2$.

   a) Számoljuk ki e 2-dimenziós valószínűségi változó sűrűségfüggvényét!
   b) $P(Y < \frac{1}{4}) =?$
   c) $P(Y < y) =?$

6. Vegyük az alábbi sűrűségfüggvényt:
   $$f(x,y) = \frac{1}{\sqrt{y}} \quad , \quad \text{ha} \quad 0 < x < 1, \quad 0 < y < x^2.$$
   a) Igazoljuk, hogy ez valóban sűrűségfüggvény!
   b) $P(X + Y < 1) =?$
   c) $P(X < x) =?$
   d) $P(Y < y) =?$

2. 2-dimenziós valószínűségi változó függvényének várható értéke

A t(X,Y) valószínűségi változó várható értéke:
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x,y) \cdot f(x,y) dx dy$$
Speciálisan X és Y szorzatának várható értéke:
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x,y) dx dy$$

Feladatok:

7. Vegyük a következő 2-dimenziós valószínűségi változót:
Első koordinátája legyen $X = \sqrt{RND_1}$. A másik koordinátája pedig ez az érték beszorozva egy másik véletlen szám négyzetgyökével: $Y = \sqrt{RND_1} \cdot \sqrt{RND_2}$.
   a) Számoljuk ki e 2-dimenziós valószínűségi változó sűrűségfüggvényét!
   b) Legyen $t(x, y) = xy$. Mennyi $t(X, Y)$ valószínűségi változó várható értéke?
   c) Legyen $t(x, y) = xy^2$. Mennyi $t(X, Y)$ valószínűségi változó várható értéke?
8. Mennyi az 1.a.), ill. 4. feladatban szereplő eloszlás várható értéke?

3. Feltételes eloszlás

A legfontosabb az alábbi összefüggés:

$$P(c < Y < d|X = x) = \int_c^d f_{2|1}(y|x)dy = F_{2|1}(d|x) - F_{2|1}(c|x)$$

Feladatok:

9. Vegyük ismét a 4.feladatban szereplő eloszlást!
   a) $P(Y \in (0.3, 0.4)|X = 0.5) =?$
   b) $P(Y \in (0.3, 0.4)|X = 0.8) =?$
   c) $P(Y \in (0.3, 0.4)|X = x) =?$
10. Vegyük ismét a 4.feladatban szereplő eloszlást!
   a) $P(X \in (0.5, 0.7)|Y = 0.1) =?$
   b) $P(X \in (0.5, 0.7)|Y = 0.4) =?$
   c) $P(X \in (0.5, 0.7)|Y = y) =?$
11. Legyenek X és Y független 2 paraméterű exponenciális eloszlású valószínűségi változók.
   a) $P(X + Y < 3) =?$
   b) $P(X + Y < z) =?$
   c) $P(X + Y < 3|X < 2) =?$
   d) $P(2 < X + Y < 3|Y > 1) =?$

4. Általános regresszió

Adott (X, Y) kétdimenziós valószínűségi változó az együttes eloszlásával (folytonos esetben sűrűségfüggvényével). X-et megfigyelve Y-t szeretnénk közelíteni egy $k(X)$ alakú tippelő függvénnyel. A közelítés azt jelenti, hogy az elkövetett négyzetes hiba – $(Y - k(X))^2$ – átlagát szeretnénk minimalizálni. Pontosabban azt az k függvényt keressük, amire

$$E\left((Y - k(X))^2\right)$$

minimális. A tétel azt mondja, hogy

$$k(x) = E(Y \mid X = x).$$

Ha pedig az elkövetett abszolút hibát, azaz $E(|Y - k(X)|)$ -t szeretnénk minimalizálni, akkor a lehető legjobb tippelés az Y mediánja, feltéve, ha már tudjuk, hogy mit vett fel az X, azaz éppen a feltételes medián függvény. A feltételes sűrűségfüggvény

$$f_{2|1}(x) = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dy}$$

amelynek az első esetben a várható értékét kell kiszámolni, a második esetben $\frac{1}{2}$-edel egyenlővé tenni és belőle y-t mint x függvényét kifejezni.

Feladatok:

12. A Duna holnaputáni budapesti vízállását akarjuk becsülni a mai bécsi vízállásból. Bár a két vízállás közt szoros kapcsolat van, azért pontosan nem lehet megmondani a vízállást, mindkettőt egy-egy valószínűségi változó írja le.Tegyük fel, hogy mindkét vízállást egy 0 és 1 közti számmal tudunk jellemezni, melynek legyen az együttes eloszlásfüggvénye
$f(x,y) = \frac{6}{5}(x + (y-1)^2)$ ahol $0 < x < 1$ és $0 < y < 1$.
    a) Határozzuk meg a budapeati vízállást a bécsi ismeretében, azaz. mi a feltételes sűrűségfüggvény?
    b) Mi annak a valószínűsége, hogy budapesten alacsonynak nevezhető (azaz 0 és $1/2$ közé esik) a vízállás, ha Bécsben x volt? (Mennyi ez $x = 1/3$-ra?)
    c) ha már ismerjük a bécsi vízállást, mire tippelünk a budapestire, ha a lehető legkisebb négyzetes hibát akarjuk elkövetni?
    d) és ha az abszolút hibát akarjuk minimalizálni?
13. $X = RND1$, $Y = RND1 * RND2$ eloszás esetén láttuk, hogy az együttes sűrűségfüggvény $f(x,y) = 1/x$, ha $0 < y < x < 1$ háromszögön vagyunk. Mi a regressziós görbe, ha az abszolút hibát, illetve ha a négyzetes hibát szeretnénk minimalizálni?
14. Az egységkörön választunk egyenletes eloszlás szerint egy (X,Y) pontot. Az X koordinata ismeretében hogyan kozelítené $|Y|$-t, feltéve, hogy a hiba abszolútértéknégyzetét szertné minimalizálni?
15. Többpártrendszer eseten az egyes pártokra leadott szavazatok százalékos aránya valószínűségi változó. A Zöldek az összes szavazatok X, a Demokraták az összes szavazatok Y hányadát kapják, együttes eloszlásuk $h(x,y) = 24xy$, ha $0 < x$, $0 < y$, $x + y < 1$.
Ha a Demokraták az összes szavazatok 40%-át kaptak, mire tippelünk, mennyit kaptak a zöldek?