

Valószínűségszámítás
3. RÉSZ
Kétdimenziós folytonos valószínűségi változók

jegyzet

Vetier András

2017. augusztus 8.

Tartalomjegyzék

1. Kétdimenziós folytonos valószínűségi változók	5
1.1. Sűrűségfüggvény	5
1.2. Feltételes valószínűség	6
1.3. Feltételes sűrűségfüggvény egy pozitív valószínűségű eseményen belül	6
1.4. Szorzási szabály független valószínűségi változókra	6
1.5. Általános szorzási szabály	7
1.6. Eloszlásfüggvény	8
1.7. Gyakorló feladatok	8
2. Kétdimenziós egyenletes eloszlás	13
2.1. Egyenletes eloszlás egy véges területű halmazon	13
2.2. Gyakorló feladatok	13
3. Béta eloszlások kétdimenzióban	14
3.1. A sűrűségfüggvény levezetése a $[0, 1]$ intervallum esetére	14
3.2. Speciális esetek	15
3.3. Feladatok megoldásokkal	23
3.4. Egyenletes eloszlás tetszőleges intervallumon (<i>Extra tananyag</i>)	25
3.5. Nem-egyenletes eloszlás esete (<i>Extra tananyag</i>)	25
3.6. Kovariancia (<i>Extra tananyag</i>)	26
3.7. Gyakorló feladatok	27
4. Kísérleti eredmények függvényének a várható értéke	29
4.1. Diszkrét eset	29
4.2. Folytonos eset	29
4.3. Szorzat várható értéke	30
4.4. Gyakorló feladatok	31
5. NSZT a kísérleti eredmények függvényének az átlagára	32
5.1. Diszkrét eset	32
5.2. Folytonos eset	33
5.3. Gyakorló feladatok	34

6. Vetület- és feltételes eloszlások	35
6.1. Vetület eloszlások	35
6.2. Feltételes sűrűségfüggvények	35
6.3. Feltételes sűrűségfüggvények rendszere	36
6.4. Sűrűségfüggvények keverése	37
6.5. Feltételes eloszlásfüggvény	38
6.6. Feltételes valószínűség	39
6.7. Feltételes medián	39
6.8. Feltételes várható érték	40
6.9. Feltételes variancia	41
6.10. Feltételes szórás	41
6.11. Egy számolós minta példa	42
6.12. Gyakorló feladatok	44
7. Kétdimenziós folytonos valószínűségi változó szimulációja	45
7.1. A módszer	45
7.2. Gyakorló feladatok	48
8. Transzformáció síkról síkra	49
8.1. Jacobi mátrix	49
8.2. Sűrűségfüggvények közötti kapcsolat	49
8.3. Gyakorló feladatok	50
9. Lineáris transzformáció síkról síkra	51
9.1. Mátrixszal való szorzás	51
9.2. Mátrixszal való szorzás és eltolás	51
9.3. Gyakorló feladatok	52
10. Transzformáció síkról egyenesre	53
10.1. Általános eset	53
10.2. Összeg sűrűségfüggvénye	53
10.3. Konvolúció	55
10.4. Gyakorló feladatok	55
11. Regresszió a mediánnal és a várható értékkel	57
11.1. Regresszió egydimenzióban	57
11.1.1. A medián minimál tulajdonsága	57
11.1.2. A távolság várható értékének minimalizálása	58
11.1.3. A várható érték minimál tulajdonsága	58
11.1.4. A távolság négyzete várható értékének minimalizálása	59
11.2. Regresszió kétdimenzióban	60
11.2.1. A hiba abszolút értéke várható értékének minimalizálása	60
11.2.2. A hiba négyzete várható értékének minimalizálása	61
11.3. Egy számolós minta példa	61
11.4. Gyakorló feladatok	64
12. Normális eloszlások a síkon	65
12.1. Standard normális eloszlás	65
12.2. Normális eloszlás $\mu_1, \mu_2, \sigma_1, \sigma_2, r$ paraméterekkel	66
12.3. Vetület eloszlások	74
12.4. Korrelációs együttható	74
12.5. Feltételes sűrűségfüggvény	81
12.6. Feltételes medián és várható érték	82
12.7. Feltételes szórás	82
12.8. Példa: testmagasság és testsúly	83

12.9. Példa: A műszer hibáját korigáljuk	86
12.10. Gyakorló feladatok	87
13. Kovariancia adatrendszerekre	88
13.1. Kétdimenziós adatrendszer fogalma	88
13.2. Kovariancia és kovariancia mátrix	89
13.3. Kovariancia mátrix transzformálódása	90
13.4. Gyakorló feladatok	90
14. Kovariancia valószínűségi változókra	91
14.1. A kovariancia fogalma	91
14.2. A kovariancia tulajdonságai	91
14.3. A korrelációs együttható fogalma és tulajdonságai	92
14.4. A kovariancia mátrix fogalma és tulajdonságai	93
14.5. Gyakorló feladatok	93
15. NSZT a vegyes momentumra, kovarianciára, korrelációs együtthatóra	94
15.1. Gyakorló feladatok	94
16. Kétdimenziós normális eloszlások alakja	95
16.1. A korrelációs együttható szemléletes jelentése	95
16.2. A kovariancia mátrix sajátvektorai, sajátértékei	95
16.3. Gyakorló feladatok	95
17. Polinomiális eloszlás közelítése kétdimenziós normális eloszlással (Extra tananyag)	96
17.1. Gyakorló feladatok	96
18. Többdimenziós normális eloszlások transzformációi (Extra tananyag)	97
18.1. A standard normális eloszlás lineáris transzformációi	97
18.2. Tetszőleges normális eloszlás lineáris transzformációi	97
18.3. Khí-négyzet eloszlás	97
18.4. Gyakorló feladatok	97
19. Lineáris regresszió	98
19.1. Általános probléma	98
19.2. Számolós példa	98
19.3. Gyakorló feladatok	100
20. Statisztika	101
20.1. Khí-négyzet próba	101
20.1.1. Prüntyőke fiókák száma	103
20.1.2. Madárfiókák helyett színes golyók	104
20.1.3. Egyforma esélyűek a lottózszámok?	104
20.2. U-próba	105
20.3. Tapasztalati eloszlásfüggvény	110
20.4. Gyakorló feladatok	111
21. Folytonos, de nem abszolút folytonos eloszlás a síkon (Extra tananyag)	112

Tisztelt Hallgatók!

Kérem, hogy a jegyzetben talált hibákat a **vetier@math.bme.hu** email címen jelezzék nekem. A levél tárgya legyen: **Hibát találtam**. Egy hatékony módszer a hibák rögzítésére, ha valaki a számítógépen olvassa a jegyzetet:

- nyom egy PRINT SCREEN-t
- behívja a PAINT programot
- nyom egy PASTE (ctrl-V) utasítást
- pirossal bekarikázza a hibát, esetleg valamit oda ír
- elmenti a JPG fájl, a fájl neve legyen a hiba helyének az oldalszáma, vagy a nevében legyen benne az oldalszám
- a JPG fájlokat csatolt fájlként a fent megadott címre

Természetesen minden más módszerrel küldött hibajelzést is köszönök.

Az "*Extra feladat*", "*Extra tananyag*" címkékkel megjelölt feladatok és anyagrészek nem részei a kötelező vizsga anyagának. Ezeket a vizsgákon nem kérdezzük. Viszont ha valaki az "*Extra feladat*", "*Extra tananyag*" címkékkel megjelölt feladatokból és anyagrészekből

- összesen 10 oldalnyi mennyiséget szépen, okosan megtanult, és
- a vizsga utáni napon délig a szándékát emailben jelzi, és
- a szóbeli beszámoló elején deklarálja, hogy mik azok az anyagrészek, melyekből összeáll az ő extra 10 oldala, és
- utána a vizsgáztató által a deklarált anyagból feltett kérdésekre szépen, okosan válaszol, akkor

a vizsgán már elért a közepes (3) vagy a jó (4) osztályzatát egy jeggyel javíthatja.

A "szépen, okosan" itt azt jelenti, hogy tényleg érti a dolgokat, nem csupán visszamondja az olvasottakat. iMSc hallgatók a jegy javításán túl iMSc extra pontokat szerezhetnek, ha a megtanult anyag mennyisége 15 oldal. A küldendő email címzettje: **vetier@math.bme.hu**, tárgya: **Beszámoló Extra tananyagból**.

Eredményes, örömteli, jó tanulást kívánok! Üdvözlettel,

2017. január 31.

Vetier András

1. Kétdimenziós folytonos valószínűségi változók

Két valószínűségi változót véve – legyenek ezek X és Y – összerakhatjuk őket egyetlen (X, Y) véletlen számpárrá, véletlen ponttá. Ily módon egy kétdimenziós valószínűségi változót kapunk. Példák:

1. Amikor az ejtőernyős célba ugrik, igyekszik a célpontot eltalálni, de ez nem sikerül pontosan. Általában a célpont közelében ér talajt. Ez a pont, ami véletlentől (is) függ. A pont, illetve annak koordinátái egy $(X, Y) =$ kétdimenziós valószínűségi változót adnak.
2. Ha véletlenszerűen választunk egy magyar férfit, akkor az

$X =$ testmagssága centiméterekben mérve

$Y =$ testúlya (pontosabban: testtömege) kilogramokban mérve

jelölésekkel élve $(X, Y) =$ kétdimenziós valószínűségi változó.

3. Generáljunk három 0 és 1 közötti véletlen számot, és legyenek

$X =$ a legkisebb

$Y =$ a legnagyobb

Ekkor $(X, Y) = ($ legkisebb, legnagyobb $)$ kétdimenziós valószínűségi változó.

1.1. Sűrűségfüggvény

Folytonos kétdimenziós valószínűségi változókat általában a sűrűségfüggvényükkel jellemezzük. Az $f(x, y)$ sűrűségfüggvényt a síknak valamilyen A részhalmazán integrálva megkapjuk annak a valószínűségét, hogy az (X, Y) véletlen pont az A halmazba esik:

$$P(A) = P((X, Y) \in A) = \iint_A f(x, y) dx dy$$

Ahhoz, hogy egy $f(x, y)$ függvény valamilyen kétdimenziós (X, Y) valószínűségi változó sűrűségfüggvénye lehessen két feltételnek kell teljesülni:

$$f(x, y) \geq 0$$

$$\iint_{\mathbb{R}^2} f(x, y) dx dy = 1$$

Ezek a feltételek minden kétfimenziós sűrűségfüggvényre teljesülnek, és ha egy kétváltozós függvény teljesíti ezeket a tulajdonságokat, akkor lehet olyan kétdimenziós (X, Y) valószínűségi változót értelemezni, melynek sűrűségfüggvénye a szóbanforgó függvény.

A sűrűségfüggvény értéke, mint arányossági tényező. Ha A egy kicsike halmaz (x, y) közelében, akkor

$$P((X, Y) \in A) = \iint_A f(x, y) dx dy \approx f(x, y) \times A \text{ területe}$$

és

$$f(x, y) \approx \frac{P((X, Y) \in A)}{A \text{ területe}}$$

Hangsúlyozzuk, hogy az $f(x, y)$ sűrűségfüggvény értéke nem képviseli valaminek a valószínűségét. Ha (x, y) egy adott pont a síkon, akkor $f(x, y)$ jelentése a következő: az (x, y) pont közelében akármilyen kicsi A halmaz esetén annak a valószínűsége, hogy (X, Y) az A halmazba esik, közelítőleg $f(x, y) \times A$ területe:

$$P((X, Y) \in A) \approx f(x, y) \times A \text{ területe}$$

A sűrűségfüggvény értékének közelítése. Ha A egy kicsi téglalap az (x, y) pont közelében Δx , illetve Δy oldalhosszakkal, akkor

$$f(x, y) \approx \frac{P(x < X < x + \Delta x \text{ és } y < Y < y + \Delta y)}{\Delta x \Delta y}$$

Ez a formula sok esetben hasznos eszköz a sűrűségfüggvény képletének meghatározásához.

1.2. Feltételes valószínűség

Ha A és B a síknak részhalmazai, akkor $(X, Y) \in A$, illetve $(X, Y) \in B$ egy-egy eseményt jelentenek. Az $(X, Y) \in A$ feltétel mellett az $(X, Y) \in B$ esemény valószínűségét $P((X, Y) \in B | (X, Y) \in A)$ -vel, vagy rövidebben csak $P(B|A)$ -val jelöljük. A számláló és a nevező is egy-egy integrállal írható fel:

$$P(B|A) = \frac{\iint_{A \cap B} f(x, y) dx dy}{\iint_A f(x, y) dx dy}$$

Ha $B \subseteq A$, akkor $A \cap B = B$, ezért ezt kapjuk:

$$P(B|A) = \frac{\iint_B f(x, y) dx dy}{\iint_A f(x, y) dx dy}$$

1.3. Feltételes sűrűségfüggvény egy pozitív valószínűségű eseményen belül

Legyen A a síknak egy részhalmaza, melynek pozitív a valószínűsége. Tegyük fel, hogy az $(X, Y) \in A$ esemény bekövetkezett. Ilyen feltétel mellett az (X, Y) valószínűségi változó sűrűségfüggvénye nyilván

$$f(x, y | A) = \frac{f(x, y)}{P(A)} = \frac{f(x, y)}{\iint_A f(x, y) dx dy} \quad \text{ha } (x, y) \in A$$

1.4. Szorzási szabály független valószínűségi változókra

Ha X és Y függetlenek, és sűrűségfüggvényeik $f_1(x)$, illetve $f_2(y)$, akkor az (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye az $f_1(x)$ és $f_2(y)$ sűrűségfüggvények direktszorzata:

$$f(x, y) = f_1(x) \cdot f_2(y)$$

Vázlatos bizonyítás.

$$\begin{aligned} f(x, y) &\approx \frac{P(x < X < x + \Delta x \text{ és } y < Y < y + \Delta y)}{\Delta x \Delta y} = \\ &= \frac{P(x < X < x + \Delta x)}{\Delta x} \cdot \frac{P(y < Y < y + \Delta y)}{\Delta y} \approx f_1(x) \cdot f_2(y) \end{aligned}$$

1. Példa: Ha X és Y független exponenciális eloszlású valószínűségi változók $\lambda_1 = 3$, illetve $\lambda_2 = 4$ paraméterekkel, akkor (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye az $f_1(x)$ és $f_2(y)$ sűrűségfüggvények direktszorzata:

$$f(x, y) = 3e^{-3x} \cdot 4e^{-4y} = 12e^{-3x-4y} \quad (x, y \geq 0)$$

1.5. Általános szorzási szabály

Ha X sűrűségfüggvénye $f_1(x)$, és Y feltételes sűrűségfüggvénye az $X = x$ feltétel mellett $f_{2|1}(y|x)$, akkor (X, Y) sűrűségfüggvénye

$$f(x, y) = f_1(x) \cdot f_{2|1}(y|x)$$

Hasonlóan:

$$f(x, y) = f_2(y) \cdot f_{1|2}(x|y)$$

Vázlatos bizonyítás. A két formula közül az elsőt bizonyítjuk:

$$\begin{aligned} f(x, y) &\approx \frac{\mathbf{P}(x < X < x + \Delta x \text{ és } y < Y < y + \Delta y)}{\Delta x \Delta y} = \\ &= \frac{\mathbf{P}(x < X < x + \Delta x)}{\Delta x} \cdot \frac{\mathbf{P}(y < Y < y + \Delta y \mid x < X < x + \Delta x)}{\Delta y} \approx \\ &\approx \frac{\mathbf{P}(x < X < x + \Delta x)}{\Delta x} \cdot \frac{\mathbf{P}(y < Y < y + \Delta y \mid X \approx x)}{\Delta y} \approx \\ &\approx f_1(x) \cdot f_{2|1}(y|x) \end{aligned}$$

2. Példa: Ha X exponenciális eloszlású valószínűségi változó $\lambda_1 = 3$ paraméterrel, és az $X = x$ feltétel mellett Y exponenciális eloszlást követ x paraméterrel, akkor

$$f_1(x) = 3e^{-3x} \quad (x \geq 0)$$

$$f_{2|1}(y|x) = xe^{-xy} \quad (y \geq 0)$$

és így az (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye:

$$f(x, y) = 3e^{-3x} \cdot xe^{-xy} = 3xe^{-3x-xy} = 3xe^{-x(3+y)} \quad (x, y \geq 0)$$

3. Példa: Ha Y egyenletes eloszlású valószínűségi változó 0 és 1 között, és az $Y = y$ feltétel mellett X egyenletes eloszlású valószínűségi változó 0 és y között, akkor

$$f_2(y) = 1 \quad (0 < y < 1)$$

$$f_{1|2}(x|y) = \frac{1}{y} \quad (0 < x < y)$$

és így az (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye:

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

1.6. Eloszlásfüggvény

Egy kétdimenziós valószínűségi változó eloszlásfüggvényét az

$$F(x, y) = P(X < x, Y < y)$$

képlettel definiáljuk. Az eloszlásfüggvény a sűrűségfüggvényből integrálással adódik:

$$F(x, y) = \int_{-\infty}^x \left(\int_{-\infty}^y f(x, y) dy \right) dx = \int_{-\infty}^y \left(\int_{-\infty}^x f(x, y) dx \right) dy$$

A sűrűségfüggvényt az eloszlásfüggvényből deriválással kapjuk:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

Egy téglalap valószínűsége az eloszlásfüggvénynek a téglalap sarkain vett értékeiből kiszámolható:

$$\begin{aligned} P(x_1 < X < x_2 \text{ és } y_1 < Y < y_2) &= \\ &= P(X < x_2 \text{ és } Y < y_2) - P(X < x_1 \text{ és } Y < y_2) - \\ &\quad - P(X < x_2 \text{ és } Y < y_1) + P(X < x_1 \text{ és } Y < y_1) = \\ &= F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \end{aligned}$$

1.7. Gyakorló feladatok

1. Egy (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye

$$f(x, y) = 4xy \quad (0 < x, y < 1)$$

- Szemléltesse az eloszlást "festéssel"!
- Szemléltesse az eloszlást a sűrűségfüggvényt megadó felülettel!
- Határozza meg az eloszlásfüggvény képletét!
- Mennyi a valószínűsége annak, hogy $X + Y < 1$?
- Mennyi a valószínűsége annak, hogy $Y < X^2$?
- Mennyi a valószínűsége annak, hogy $Y < X^2$, feltéve, hogy $X + Y < 1$?
- Adja meg az $f_1(x)$ sűrűségfüggvény képletét!
- Adja meg az $f_{2|1}(y|x)$ feltételes sűrűségfüggvény képletét!
- Határozza meg az $f_2(y)$ sűrűségfüggvény képletét!
- Határozza meg az $f_{1|2}(x|y)$ feltételes sűrűségfüggvény képletét!

2. Egy (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye

$$f(x, y) = c(x + y) \quad (x > 0, y > 0, x + y < 1)$$

ahol c egy alkalmas konstans.

- Határozza meg a c konstans értékét!
- Szemléltesse az eloszlást "festéssel"!

- (c) Szemléltesse az eloszlást a sűrűségfüggvényt megadó felülettel!
- (d) Határozza meg az eloszlásfüggvény képletét!
- (e) Mennyi a valószínűsége annak, hogy $X + Y < 1$?
- (f) Mennyi a valószínűsége annak, hogy $Y < X^2$?
- (g) Mennyi a valószínűsége annak, hogy $Y < X^2$, feltéve, hogy $X + Y < 1$?

3. Egy (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye

$$f(x, y) = c(2x + y) \quad (x > 0, y > 0, x + y < 1)$$

ahol c egy alkalmas konstans.

- (a) Határozza meg a c konstans értékét!
- (b) Adja meg az $f_1(x)$ sűrűségfüggvény képletét!
- (c) Adja meg az $f_{2|1}(y|x)$ feltételes sűrűségfüggvény képletét!
- (d) Számolja ki a $P(Y < 0.1 | X = 0.25)$ feltételes valószínűséget!
- (e) Számolja ki a $P(Y < 0.1 | X = 0.50)$ feltételes valószínűséget!
- (f) Számolja ki a $P(Y < 0.1 | X = 0.75)$ feltételes valószínűséget!
- (g) Számolja ki a $P(Y < 0.1 | X = x)$ feltételes valószínűséget!
- (h) Határozza meg az $f_2(y)$ sűrűségfüggvény képletét!
- (i) Határozza meg az $f_{1|2}(x|y)$ feltételes sűrűségfüggvény képletét!
- (j) Számolja ki a $P(X < 0.1 | Y = 0.25)$ feltételes valószínűséget!
- (k) Számolja ki a $P(X < 0.1 | Y = 0.50)$ feltételes valószínűséget!
- (l) Számolja ki a $P(X < 0.1 | Y = 0.75)$ feltételes valószínűséget!
- (m) Számolja ki a $P(X < 0.1 | Y = y)$ feltételes valószínűséget!

Megoldás:

(a) $c = 2$, vagyis $f(x, y) = 2(2x + y) = 4x + 2y \quad (x > 0, y > 0, x + y < 1)$

(b) $f_1(x) = 2x - 3x^2 + 1 \quad (0 < x < 1)$

$$F_1(x) = x^2 - x^3 + x \quad (0 < x < 1)$$

(c) $f_{2|1}(y|x) = \frac{4x+2y}{2x-3x^2+1} \quad (0 < x < 1 - y < 1)$

$$F_{2|1}(y|x) = \frac{4xy+y^2}{2x-3x^2+1} \quad (0 < x < 1 - y < 1)$$

(d) $P(Y < 0.1 | X = 0.25) = F_{2|1}(y|x) \Big|_{x=0.25, y=0.1} = 0.080$

(e) $P(Y < 0.1 | X = 0.50) = F_{2|1}(y|x) \Big|_{x=0.50, y=0.1} = 0.164$

(f) $P(Y < 0.1 | X = 0.75) = F_{2|1}(y|x) \Big|_{x=0.75, y=0.1} = 0.375$

(g) $P(Y < 0.1 | X = x) = F_{2|1}(0.1|x) = \frac{0.4x+0.01}{2x-3x^2+1} \quad \text{ha } 0 < x < 0.9$

$$P(Y < 0.1 | X = x) = 1 \quad \text{ha } 0.9 < x < 1$$

(h) $f_2(y) = 2 - 2y \quad (0 < y < 1)$

$F_2(y) = 2y - y^2 \quad (0 < y < 1)$

(i) $f_{1|2}(x|y) = \frac{2x+y}{1-y} \quad (0 < y < 1 - x < 1)$

$F_{1|2}(x|y) = \frac{x^2+xy}{1-y} \quad (0 < y < 1 - x < 1)$

(j) $P(X < 0.1|Y = 0.25) = \left[F_{1|2}(x|y) = \frac{x^2+xy}{1-y} \right]_{x=0.1, y=0.25} = 0.047$

(k) $P(X < 0.1|Y = 0.50) = \left[F_{1|2}(x|y) = \frac{x^2+xy}{1-y} \right]_{x=0.1, y=0.50} = 0.120$

(l) $P(X < 0.1|Y = 0.75) = \left[F_{1|2}(x|y) = \frac{x^2+xy}{1-y} \right]_{x=0.1, y=0.75} = 0.340$

(m) $P(X < 0.1|Y = y) = F_{1|2}(0.1|y) = \frac{0.01+0.1y}{1-y} \quad \text{ha } 0 < y < 0.9$

$P(X < 0.1|Y = y) = 1 \quad \text{ha } 0.9 < y < 1$

4. Reggel taxival megyek az egyetemre. A várakozási időm percekben mérve X . Feltesszük, hogy X exponenciális eloszlást követ $\lambda_1 = 0.1$ paraméterrel. Este taxival megyek haza. A várakozási időm percekben mérve Y . Feltesszük, hogy Y exponenciális eloszlást követ $\lambda_2 = 0.2$ paraméterrel. A két várakozási idő független egymástól.

- Határozza meg az $f(x,y)$ síkbeli sűrűségfüggvény képletét!
- Mennyi annak a valószínűsége, hogy $X + Y < 15$?
- Mennyi annak a valószínűsége, hogy $X + Y < z$?
- Mennyi annak a valószínűsége, hogy $X < Y$?
- Mennyi annak a valószínűsége, hogy $X < Y$, feltéve, hogy $X + Y < 15$?

Az alábbi "pince-világítással kapcsolatos" feladatok egymásra épülnek. A feladat-sorozat lépésről-lépésre vezet a gamma eloszlások sűrűségfüggvényeinek a képletét.

5. Másodrendű gamma eloszlás sűrűségfüggvényének levezetése:

Sötét pincénkben állandóan fog égni a villany, amíg csak bírják az izzóim. Sajnos csak két izzóm van. Ezeket egymás után fogom használni. Élettartamaik függetlenek és exponenciális eloszlást követnek 2 hónap várható értékkel. Legyen X az izzócsere pillanata, Y pedig az a pillanat, amikor a második is kiég.

- Adja meg az $f_1(x)$ sűrűségfüggvény képletét!
- Adja meg az $f_{2|1}(y|x)$ feltételes sűrűségfüggvény képletét!
- Határozza meg az $f(x, y)$ sűrűségfüggvény képletét!
- Határozza meg az $f_2(y)$ sűrűségfüggvény képletét! (Ez a másodrendű gamma eloszlás sűrűségfüggvénye. Emlékeztetőül: itt most Y az a pillanat, amikor a második izzó kiég.)
- Az érdekesség kedvéért határozza meg az $f_{1|2}(x|y)$ feltételes sűrűségfüggvény képletét, és gondolja meg, mit jelent a kapott eredmény!

6. Harmadrendű gamma eloszlás sűrűségfüggvényének levezetése:

Szomszédom pincéjében is állandóan fog égni a villany, amíg csak bírják az izzói. Ő gazdagabb, neki három izzója van. Ezeket egymás után fogja használni. Élettartamaik függetlenek és exponenciális eloszlást követnek 2 hónap várható értékkel. Legyen X a második izzócsere pillanata, Y pedig az a pillanat, amikor a harmadik is kiég.

- (a) Adja meg az $f_1(x)$ sűrűségfüggvény képletét! (Segítség: ezt a sűrűségfüggvényt az előző feladatban határoztuk meg, és akkor $f_2(y)$ -nal jelöltük.)
- (b) Adja meg az $f_{2|1}(y|x)$ feltételes sűrűségfüggvény képletét!
- (c) Határozza meg az $f(x, y)$ sűrűségfüggvény képletét!
- (d) Határozza meg az $f_2(y)$ sűrűségfüggvény képletét! (Ez a harmadrendű gamma eloszlás sűrűségfüggvénye. Emlékeztetőül: itt most Y az a pillanat, amikor a harmadik izzó kiég.)
- (e) Az érdekesség kedvéért határozza meg az $f_{1|2}(x|y)$ feltételes sűrűségfüggvény képletét, és gondolja meg, mit jelent a kapott eredmény!

7. *Negyedrendű gamma eloszlás sűrűségfüggvényének levezetése:*

Másodsomszédom pincéjében is állandóan fog égni a villany, amíg csak bírják az izzói. Ő még gazdagabb, neki négy izzója van. Ezeket egymás után fogja használni. Élettartamaik függetlenek és exponenciális eloszlást követnek 2 hónap várható értékkel. Legyen X a harmadik izzócsere pillanata, Y pedig az a pillanat, amikor a negyedik izzó kiég.

- (a) Adja meg az $f_1(x)$ sűrűségfüggvény képletét! (Segítség: ezt a sűrűségfüggvényt az előző feladatban határoztuk meg, és akkor $f_2(y)$ -nal jelöltük.)
- (b) Adja meg az $f_{2|1}(y|x)$ feltételes sűrűségfüggvény képletét!
- (c) Határozza meg az $f(x, y)$ sűrűségfüggvény képletét!
- (d) Határozza meg az $f_2(y)$ sűrűségfüggvény képletét! (Ez a negyedrendű gamma eloszlás sűrűségfüggvénye. Emlékeztetőül: itt most Y az a pillanat, amikor a negyedik izzó kiég.)
- (e) Az érdekesség kedvéért határozza meg az $f_{1|2}(x|y)$ feltételes sűrűségfüggvény képletét, és gondolja meg, mit jelent a kapott eredmény!

8. *Ötödrendű gamma eloszlás sűrűségfüggvényének levezetése:*

Harmadik szomszédom pincéjében is állandóan fog égni a villany, amíg csak bírják az izzói. Ő még a másodsomszédomnál is gazdagabb, neki öt izzója van. Ezeket egymás után fogja használni. Élettartamaik függetlenek és exponenciális eloszlást követnek 2 hónap várható értékkel. Legyen X a negyedik izzócsere pillanata, Y pedig az a pillanat, amikor az ötödik izzó kiég.

- (a) Adja meg az $f_1(x)$ sűrűségfüggvény képletét! (Segítség: ezt a sűrűségfüggvényt az előző feladatban határoztuk meg, és akkor $f_2(y)$ -nal jelöltük.)
- (b) Adja meg az $f_{2|1}(y|x)$ feltételes sűrűségfüggvény képletét!
- (c) Határozza meg az $f(x, y)$ sűrűségfüggvény képletét!
- (d) Határozza meg az $f_2(y)$ sűrűségfüggvény képletét! (Ez a negyedrendű gamma eloszlás sűrűségfüggvénye. Emlékeztetőül: itt most Y az a pillanat, amikor aaz ötödik izzó kiég.)
- (e) Az érdekesség kedvéért határozza meg az $f_{1|2}(x|y)$ feltételes sűrűségfüggvény képletét, és gondolja meg, mit jelent a kapott eredmény!

9. *n -ed rendű gamma eloszlás sűrűségfüggvényének levezetése a teljes indukció módszerével:*

A jó hosszú utcánk végében lakó ember pincéjében is állandóan fog égni a villany, amíg csak bírják az izzói. Neki – mondjuk – n izzója van. Ezeket egymás után fogja használni. Élettartamaik függetlenek és exponenciális eloszlást követnek 2 hónap várható értékkel. Legyen X az $n - 1$ -edik izzócsere pillanata, Y pedig az a pillanat, amikor az n -edik izzó kiég.

- (a) Adja meg az $f_1(x)$ sűrűségfüggvény képletét! (Segítség: ezt a sűrűségfüggvényt a teljes indukció módszerének megfelelően – okosan – vegye fel!
- (b) Adja meg az $f_{2|1}(y|x)$ feltételes sűrűségfüggvény képletét!
- (c) Határozza meg az $f(x, y)$ sűrűségfüggvény képletét!
- (d) Határozza meg az $f_2(y)$ sűrűségfüggvény képletét! (Ez n -ed rendű gamma eloszlás sűrűségfüggvénye. Emlékeztetőül: itt most Y az a pillanat, amikor a n -edik izzó kiég.)

- (e) Az érdekesség kedvéért határozza meg az $f_{1|2}(x|y)$ feltételes sűrűségfüggvény képletét, és gondolja meg, mit jelent a kapott eredmény!

2. Kétdimenziós egyenletes eloszlás

2.1. Egyenletes eloszlás egy véges területű halmazon

Legyen S egy véges területű halmaz a síkon. Tekintsük azt a kétváltozós függvényt, melynek értéke a halmaz területének reciprokával egyenlő a halmazba tartozó pontokban, egyéb pontokban pedig 0:

$$f(x, y) = \frac{1}{S \text{ területe}} \quad \text{ha } (x, y) \in S \quad f(x, y) = 0 \quad \text{egyébként}$$

Vegyük most S -nek egy A részhalmazát. Ha a $P(A)$ valószínűség kiszámítása érdekében integráljuk az $f(x, y)$ függvényt az A halmazon, akkor az integrál értékül a szóbanforgó konstans és a halmaz területének a szorzatát kapjuk:

$$P(A) = \iint_A f(x, y) dx dy = \iint_A \frac{1}{S \text{ területe}} dx dy = \frac{A \text{ területe}}{S \text{ területe}}$$

Tehát az A halmaz valószínűsége egyenlő az A területe osztva az S területével. Ugyanahoz a képlethez jutottunk, mint az 1. rész 3. fejezetének 1. pontjában a "Folytonos egyenletes eloszlás" cím alatt.

2.2. Gyakorló feladatok

1. Egy bolha az origóból indulva a pozitív irányba ugrál. Ugrásainak nagysága független egymástól, és deciméterben mérve egyenletes eloszlást követ 30 és 60 között. Legyen X annak a pontnak a koordinátája, ahova először ugrik, Y pedig annak a pontnak a koordinátája, ahova másodszor ugrik.
 - (a) Adja meg az $f_1(x)$ sűrűségfüggvény képletét!
 - (b) Adja meg az $f_{2|1}(y|x)$ feltételes sűrűségfüggvény képletét!
 - (c) Határozza meg az $f(x, y)$ sűrűségfüggvény képletét!
 - (d) Határozza meg az $f_2(y)$ sűrűségfüggvény képletét!
 - (e) Határozza meg az $f_{1|2}(x|y)$ feltételes sűrűségfüggvény képletét!
2. Az origó középpontú, egységnyi sugarú körlapon egyenletes eloszlás szerint választunk egy (X, Y) pontot.
 - (a) Határozza meg az $f_1(x)$ sűrűségfüggvény képletét!
 - (b) Határozza meg az $f_{2|1}(y|x)$ feltételes sűrűségfüggvény képletét!
 - (c) Határozza meg az $f(x, y)$ sűrűségfüggvény képletét!
 - (d) Határozza meg az $f_2(y)$ sűrűségfüggvény képletét!
 - (e) Határozza meg az $f_{1|2}(x|y)$ feltételes sűrűségfüggvény képletét!

3. Béta eloszlások kétdimenzióban

3.1. A sűrűségfüggvény levezetése a $[0, 1]$ intervallum esetére

Tegyük fel, hogy n ember mindegyike – akik rendszeresen együtt ebédelnek – a többitől függetlenül dél és 1 óra között egyenletes eloszlás szerint érkezik egy vendéglőbe. Legyen X az i -ik, Y a j -ik érkezési időpont Könnyen szimulálhatjuk a problémát, ha generálunk n random számot 0 és 1 között, és

$$X = \text{az } i\text{-ik legkisebb}$$

$$Y = \text{a } j\text{-ik legkisebb}$$

Most levezetjük a kétdimenziós (X, Y) valószínűségi változó sűrűségfüggvényének a képletét. E célból először felvesszünk két x, y számot, melyekre igaz, hogy

$$0 < x < y < 1$$

Ezek után felvesszünk x mellett egy kis $[x, x + \Delta x]$ intervallumot és y mellett egy kis $[y, y + \Delta y]$ intervallumot úgy, hogy teljesüljön az $x + \Delta x < y$ egyenlőtlenség is. (X, Y) sűrűségfüggvényét hányadossal közelítjük:

$$f(x, y) \approx \frac{\mathbf{P}(x < X < x + \Delta x \text{ és } y < Y < y + \Delta y)}{\Delta x \Delta y}$$

A számlálóban álló

$$x < X < x + \Delta x \quad \text{és} \quad y < Y < y + \Delta y$$

esemény azt jelenti, hogy

az i -ik legkisebb szám, amit X jelöl, az $[x, x + \Delta x)$ intervallumba esik,

és

a j -ik legkisebb szám, amit Y jelöl, az $[y, y + \Delta y)$, intervallumba esik,

ami – részletesebben kifejtve a következőket jelenti:

legalább	1	X pont van az	$[x, x + \Delta x)$	intervallumban, és
legalább	1	Y pont van az	$[y, y + \Delta y)$	intervallumban, és
pontosan	$(i - 1)$	pont van az	$[0, X)$	intervallumban, és
pontosan	$(j - i - 1)$	pont van az	$[X, Y)$	intervallumban, és
pontosan	$(n - j)$	pont van az	$[Y, 1]$	intervallumban.

Ez – jó közelítéssel – azt jelenti, hogy

pontosan	$(i - 1)$	pont van a	$[0, x)$	intervallumban, és
pontosan	1	pont van az	$[x, x + \Delta x)$	intervallumban, és
pontosan	$(j - i - 1)$	pont van az	$[x + \Delta x, y)$	intervallumban, és
pontosan	1	pont van az	$[y, y + \Delta y)$	intervallumban, és
pontosan	$(n - j)$	pont van az	$[y + \Delta y, 1]$	intervallumban.

A polihipergeometrikus eloszlás formuláját felhasználva így azt kapjuk, hogy az

$$x < X < x + \Delta x \quad \text{és} \quad y < Y < y + \Delta y$$

esemény valószínűsége körülbelül

$$\frac{n!}{(i-1)! 1! (j-i-1)! 1! (n-j)!} (x)^{i-1} (\Delta x)^1 (y - (x + \Delta x))^{j-i-1} (\Delta y)^1 (1 - (y + \Delta y))^{n-j}$$

Eldobhatunk néhány fölösleges zárójelet, tényezőt és kitevőt, ami által a formula így egyszerűsödik:

$$\frac{n!}{(i-1)! (j-i-1)! (n-j)!} x^{i-1} \Delta x (y - (x + \Delta x))^{j-i-1} \Delta y (1 - (y + \Delta y))^{n-j}$$

Osztva $\Delta x \cdot \Delta y$ -nal osztva, majd $\Delta x \rightarrow 0$, $\Delta y \rightarrow 0$ határérték mellett megkapjuk a sűrűségfüggvényt:

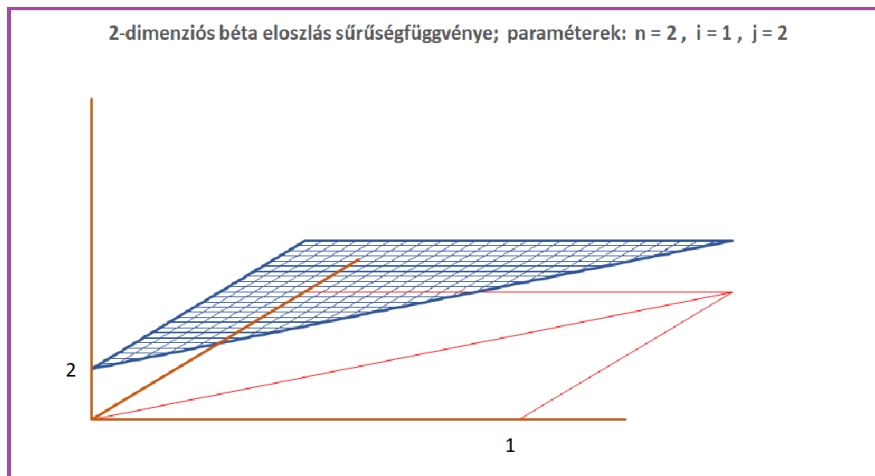
$$f(x, y) = \frac{n!}{(i-1)! (j-i-1)! (n-j)!} x^{i-1} (y-x)^{j-i-1} (1-y)^{n-j}$$

3.2. Speciális esetek

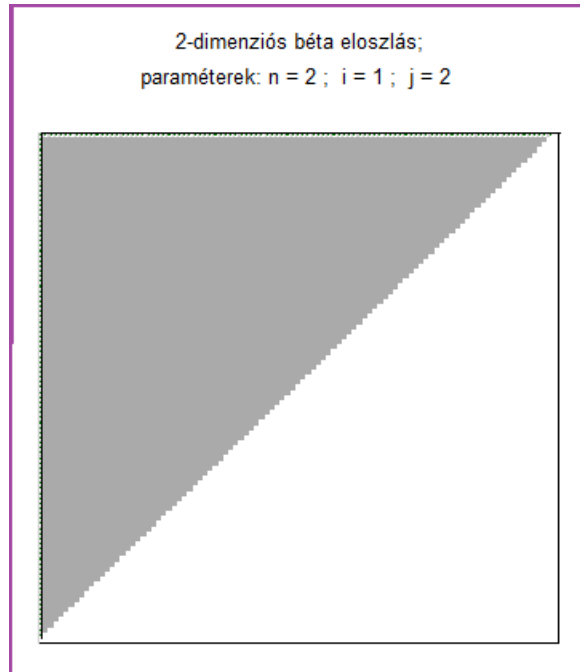
Két véletlen szám esete:

$X = \mathbf{a\ kisebbik}$ $Y = \mathbf{a\ nagyobbik}$ $n = 2, i = 1, j = 2$ esetén ezt kapjuk:

$$f(x, y) = 2 \quad \text{if } 0 < x < y < 1$$



1. ábra. 2-dimenziós béta eloszlás sűrűségfüggvénye; paraméterek: $n = 2, i = 1, j = 2$



2. ábra. 2-dimenziós béta eloszlás festéssel szemléltetve; paraméterek: $n = 2$, $i = 1$, $j = 2$

Három véletlen szám esete:

1. $X = \mathbf{a}$ legkisebb $Y = \mathbf{a}$ legnagyobb $n = 3, i = 1, j = 3$ esetén ezt kapjuk:

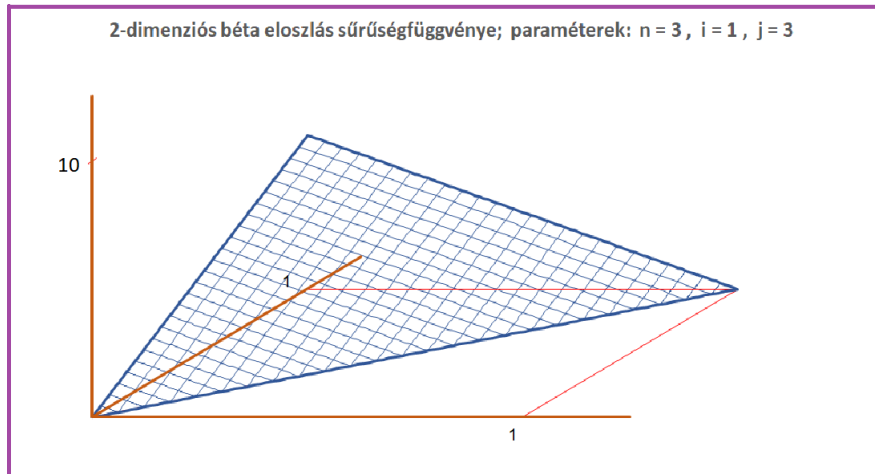
$$f(x, y) = 6(y - x) \quad \text{if } 0 < x < y < 1$$

2. $X = \mathbf{a}$ legkisebb $Y = \mathbf{a}$ középső $n = 3, i = 1, j = 2$ esetén ezt kapjuk:

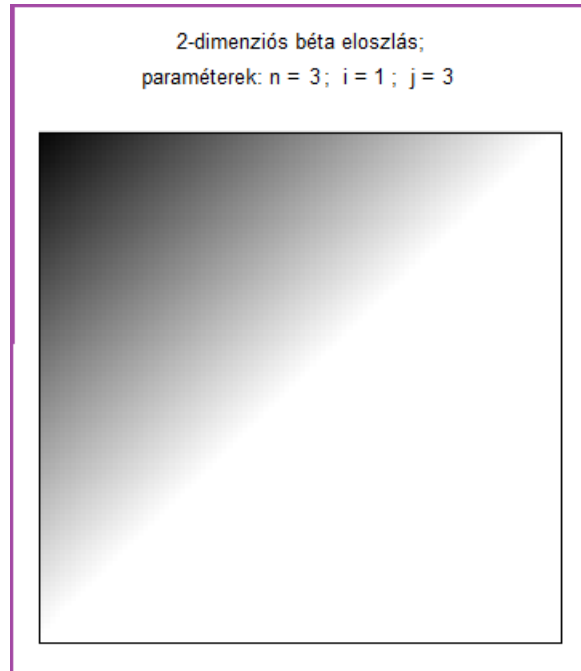
$$f(x, y) = 6(1 - y) \quad \text{ha } 0 < x < y < 1$$

3. $X = \mathbf{a}$ középső $Y = \mathbf{a}$ legnagyobb $n = 3, i = 2, j = 3$ esetén ezt kapjuk:

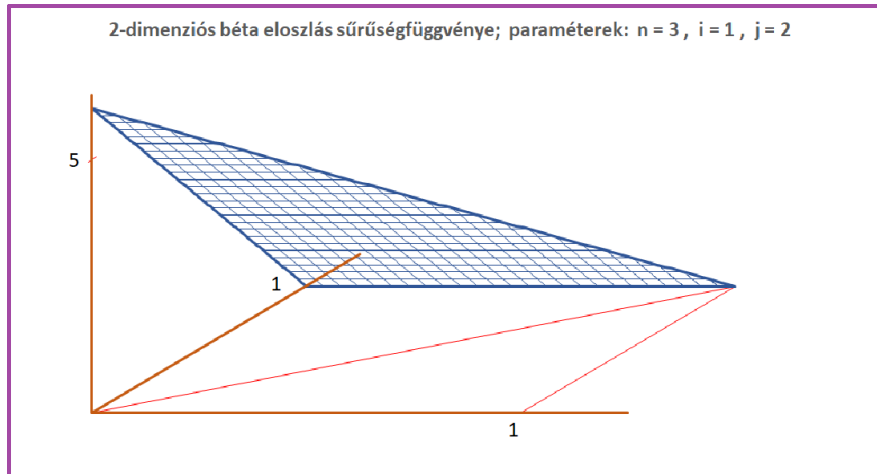
$$f(x, y) = 6x \quad \text{ha } 0 < x < y < 1$$



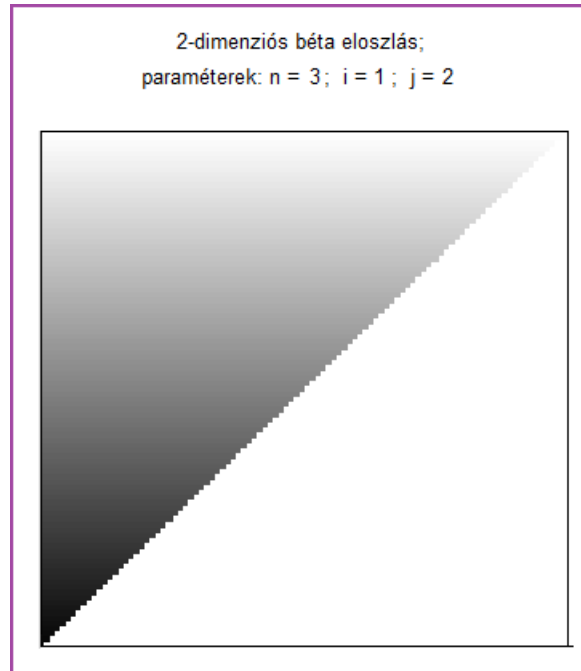
3. ábra. 2-dimenziós béta eloszlás sűrűségfüggvénye; paraméterek: $n = 3$, $i = 1$, $j = 3$



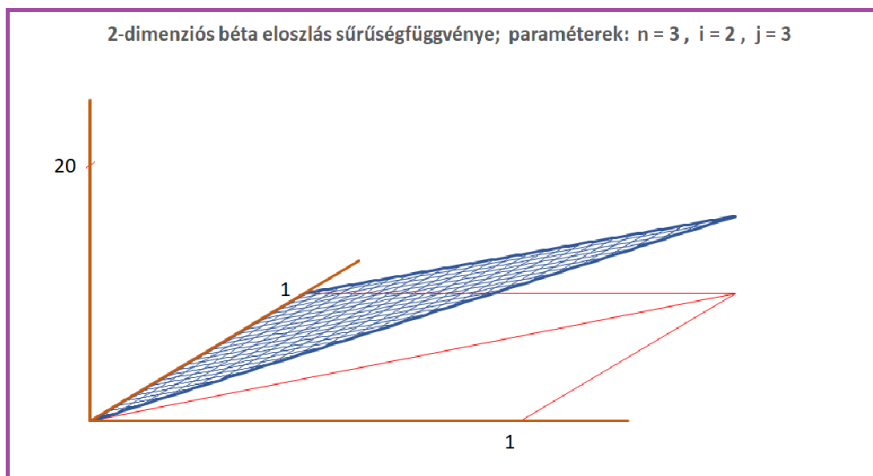
4. ábra. 2-dimenziós béta eloszlás festékekkel szemléltetve; paraméterek: $n = 3$, $i = 1$, $j = 3$



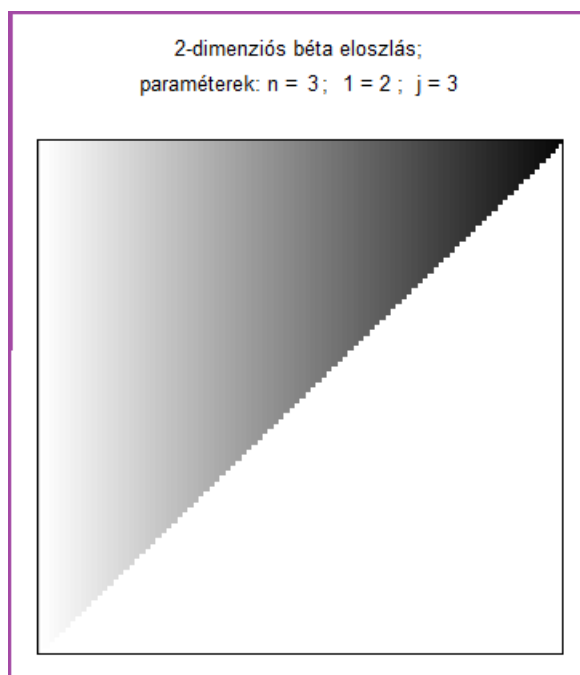
5. ábra. 2-dimenziós béta eloszlás sűrűségfüggvénye; paraméterek: $n = 3, i = 1, j = 2$



6. ábra. 2-dimenziós béta eloszlás festékekkel szemléltetve; paraméterek: $n = 3, i = 1, j = 2$



7. ábra. 2-dimenziós béta eloszlás sűrűségfüggvénye; paraméterek: $n = 3, i = 2, j = 3$



8. ábra. 2-dimenziós béta eloszlás festékkal szemléltetve; paraméterek: $n = 3, i = 2, j = 3$

Négy véletlen szám esete:

1. $X = \mathbf{a}$ legkisebb $Y = \mathbf{a}$ legnagyobb $n = 4, i = 1, j = 4$ esetén ezt kapjuk:

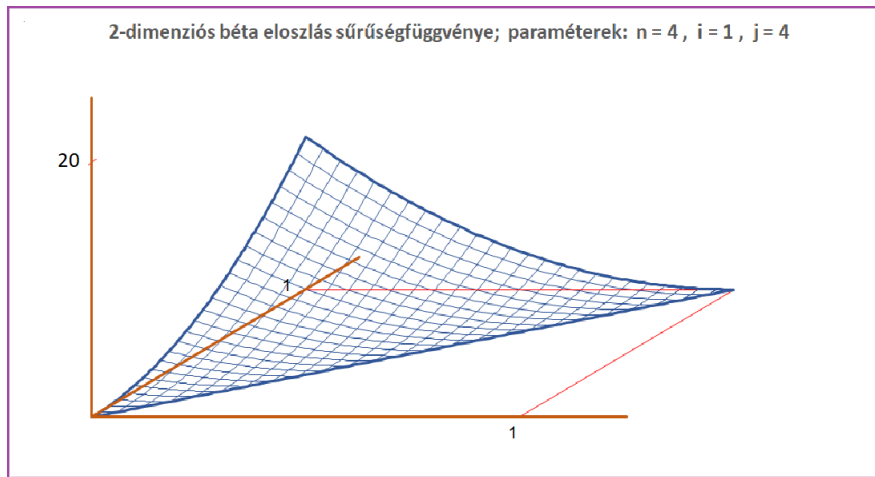
$$f(x, y) = 12 (y - x)^2 \quad \text{ha } 0 < x < y < 1$$

2. $X = \mathbf{a}$ legkisebb $Y = \mathbf{a}$ második legkisebb $n = 4, i = 1, j = 2$ esetén ezt kapjuk:

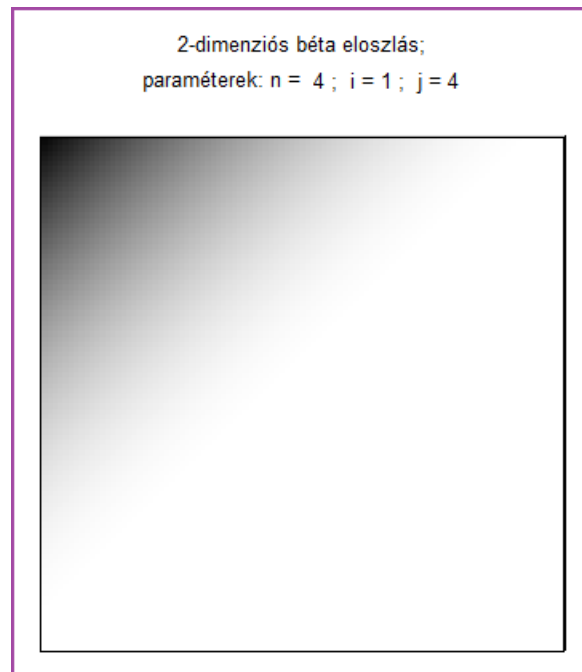
$$f(x, y) = 12(1 - y)^2 \quad \text{ha } 0 < x < y < 1$$

3. $X = a$ második legkisebb $Y = a$ harmadik legkisebb $n = 4, i = 2, j = 3$ esetén ezt kapjuk:

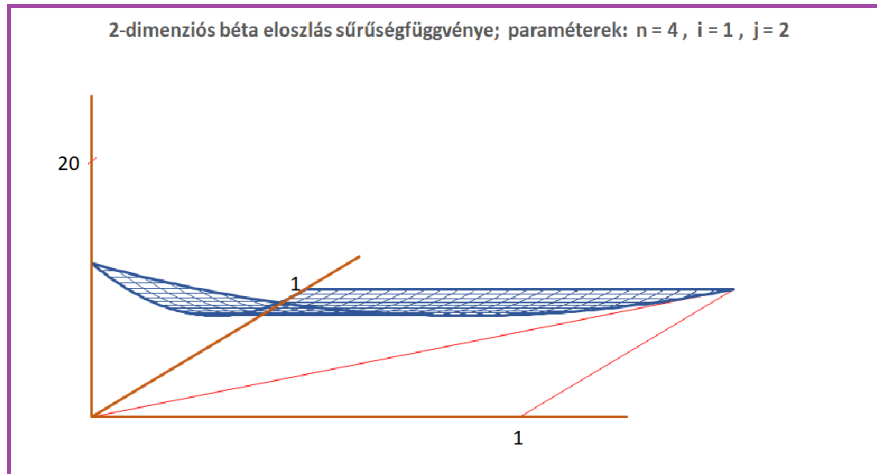
$$f(x, y) = 24x(1 - y)x \quad \text{ha } 0 < x < y < 1$$



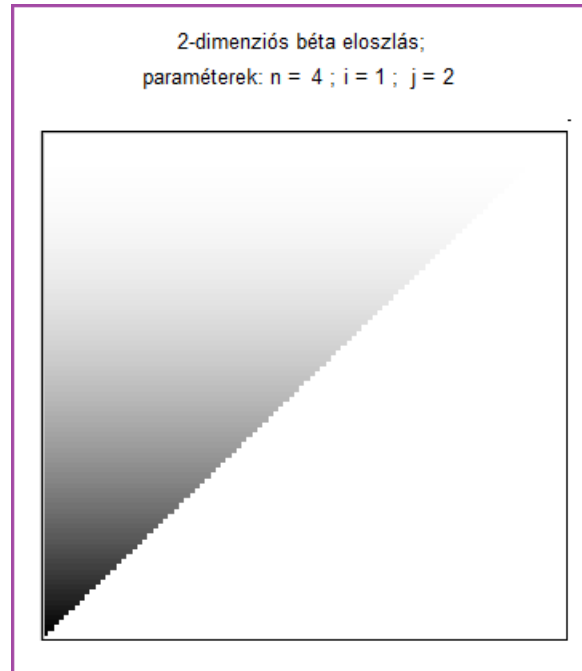
9. ábra. 2-dimenziós béta eloszlás sűrűségfüggvénye; paraméterek: $n = 4, i = 1, j = 4$



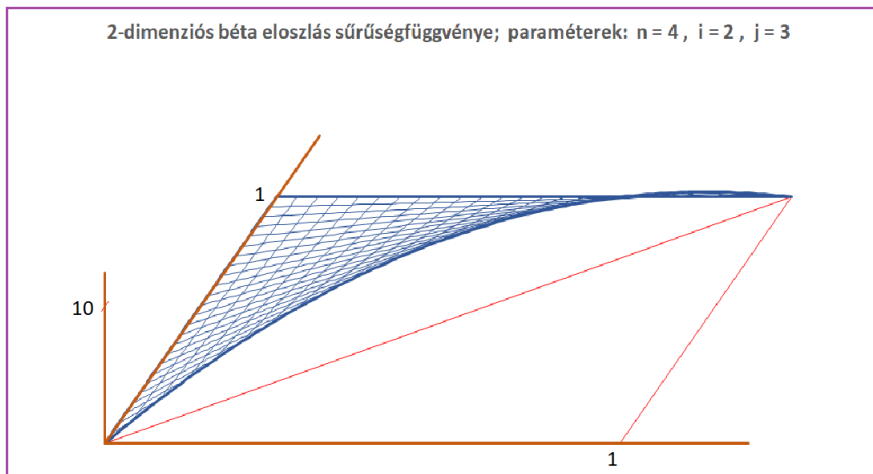
10. ábra. 2-dimenziós béta eloszlás festéssel szemléltetve; paraméterek: $n = 4, i = 1, j = 4$



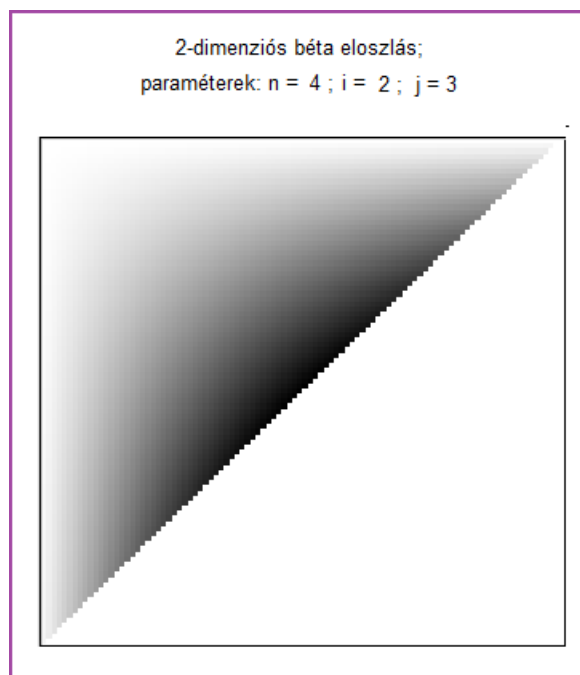
11. ábra. 2-dimenziós béta eloszlás sűrűségfüggvénye; paraméterek: $n = 4, i = 1, j = 2$



12. ábra. 2-dimenziós béta eloszlás festékkal szemléltetve; paraméterek: $n = 4, i = 1, j = 2$



13. ábra. 2-dimenziós béta eloszlás sűrűségfüggvénye; paraméterek: $n = 4, i = 2, j = 3$

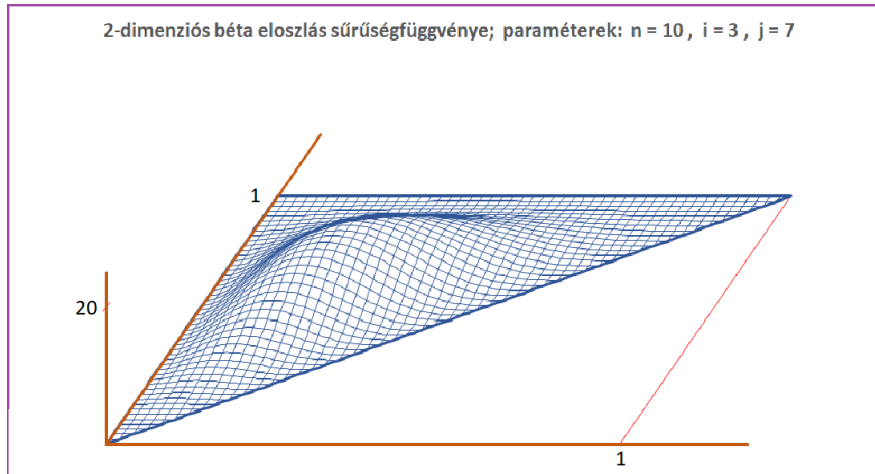


14. ábra. 2-dimenziós béta eloszlás festéssel szemléltetve; paraméterek: $n = 4, i = 2, j = 3$

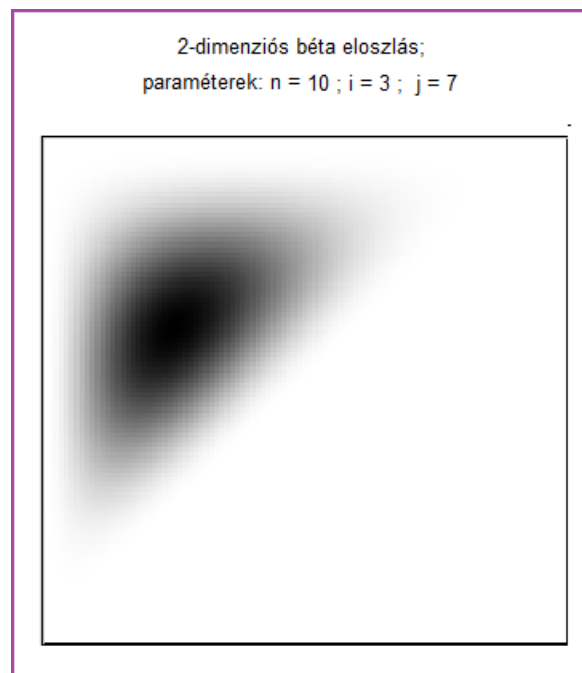
Tíz véletlen szám esete:

$X = \mathbf{a\ 3-ik\ legkisebb}$ $Y = \mathbf{a\ 7-ik\ legkisebb}$ $n = 10, i = 3, j = 7$ esetén ezt kapjuk:

$$f(x, y) = \frac{10!}{2! 3! 3!} x^2 (y - x)^3 (1 - y)^3 \quad \text{ha } 0 < x < y < 1$$



15. ábra. 2-dimenziós béta eloszlás sűrűségfüggvénye; paraméterek: $n = 10, i = 3, j = 7$



16. ábra. 2-dimenziós béta eloszlás festéssel szemléltetve; paraméterek: $n = 10, i = 3, j = 7$

3.3. Feladatok megoldásokkal

1. Feladat: Három jóbarát déli 12 óra és délután 1 óra között egymástól függetlenül (folytonos) egyenletes eloszlás szerint érkeznek a menzára. Mi a valószínűsége annak, hogy aki elsőnek ér oda, az fél egy előtt, aki utolsónak, az fél egy után érkezik?

Megoldás: Bizonyára az Olvasó emlékszik az előző alpontban említett "Három véletlen szám esete" példára, amikor X jelölte a három véletlen szám közül a legkisebbet, Y a legnagyobbat, és (X, Y) sűrűségfüggvénye az

$$f(x, y) = 6(y - x) \quad (0 \leq x \leq y \leq 1)$$

függvény volt. Ha most tekintjük az

$$N = \{(x, y) : 0 \leq x \leq 0.5 \quad 0.5 \leq y \leq 1\}$$

négyzetet, akkor a kérdéses valószínűség annyi, mint annak a valószínűsége, hogy az (X, Y) pont az N négyzetbe esik:

$$P((X, Y) \in N) = \iint_N 6(y-x) dx dy = \int_{0.5}^1 \left(\int_0^{0.5} 6(y-x) dx \right) dy = \dots = \frac{3}{4}$$

A képletben a három pont az jelenti, hogy az intergrál kiszámolása az Olvasó feladata.

2. Feladat: Három jóbarát déli 12 óra és délután 1 óra között egymástól függetlenül (folytonos) egyenletes eloszlás szerint érkezik a menzára. Mi a valószínűsége annak, hogy aki elsőnek ér oda, az fél egy előtt, aki utolsónak, az fél egy után érkezik feltéve, hogy érkezési időpontjaik között legfeljebb fél óra telik el?

Megoldás: Az első érkezési időpont legyen X , az utolsó Y . Ekkor (X, Y) sűrűségfüggvénye – mint tudjuk:

$$f(x, y) = 6(y-x) \quad (0 < x < y < 1)$$

A kért feltételes valószínűség:

$$\begin{aligned} P(X < 0.5 \text{ és } Y > 0.5 \mid Y - X < 0.5) &= \frac{P(X < 0.5 \text{ és } Y > 0.5 \text{ és } Y - X < 0.5)}{P(Y - X < 0.5)} = \\ &= \frac{P(0 < X < 0.5 \text{ és } 0.5 < Y < X + 0.5)}{1 - P(Y - X > 0.5)} \end{aligned}$$

A számláló:

$$P(0 < X < 0.5 \text{ és } 0.5 < Y < X + 0.5) = \int_0^{0.5} \left(\int_{0.5}^{x+0.5} 6(y-x) dy \right) dx = \dots = 0.25$$

A nevező:

$$1 - P(0 < X < 0.5 \text{ és } Y > X + 0.5) = \int_0^{0.5} \left(\int_{0.5}^1 6(y-x) dy \right) dx = \dots = 0.5$$

Így a kért feltételes valószínűség $0.25/0.5 = 0.5$ -del egyenlő.

3. Feladat: Most tegyük fel, hogy nem 3, hanem n jóbarát érkezik a menzára déli 12 óra és délután 1 óra között egymástól függetlenül (folytonos) egyenletes eloszlás szerint. **(a)** Mi a valószínűsége annak, hogy az utolsó és az első érkezési időpont között legalább z idő eltelik ($0 < z < 1$)? Az említett két időpont között eltelt időtartamnak mi **(b)** az eloszlásfüggvénye, **(c)** sűrűségfüggvénye?

Megoldás: Az első érkezési időpont legyen X , az utolsó Y , az X és Y között eltelt időtartam $Z = Y - X$. Ekkor (X, Y) sűrűségfüggvénye – mint tudjuk:

$$f(x, y) = n(n-1)(y-x)^{n-2} \quad (0 < x < y < 1)$$

(a) A kért valószínűség:

$$\begin{aligned} P(Z > z) &= P(X < Y - z) = \int_z^1 \left(\int_0^{y-z} n(n-1)(y-x)^{n-2} dx \right) dy = \dots \\ &= 1 - n z^{n-1} + (n-1) z^n \end{aligned}$$

(b) Az eloszlásfüggvény:

$$R(z) = 1 - P(Z > z) = n z^{n-1} - (n-1) z^n$$

(c) Az sűrűségfüggvény:

$$r(z) = (n z^{n-1} - (n-1) z^n)' = n(n-1) z^{n-2} - n(n-1) z^{n-1} = n(n-1)(z^{n-2} - z^{n-1})$$

3.4. Egyenletes eloszlás tetszőleges intervallumon (*Extra tananyag*)

Az előző gondolatmenetet könnyű általánosítani arra az esetre, ha az emberek az A és B időpontok között érkeznek (egyenletes eloszlás szerint, egymástól függetlenül), azaz n független, egyenletes eloszlású véletlen számot tekintünk A és B között, és

$X =$ az i -ik legkisebb

$Y =$ a j -ik legkisebb

ahol $(0 < i < j \leq n)$, akkor minden $A < x < y < B$ esetén

$$f(x, y) = \frac{1}{(B - A)^2} \cdot \frac{n!}{(i - 1)! (j - i - 1)! (n - j)!} \cdot \left(\frac{x - A}{B - A}\right)^{i-1} \cdot \left(\frac{y - x}{B - A}\right)^{j-i-1} \cdot \left(\frac{B - y}{B - A}\right)^{n-j}$$

3.5. Nem-egyenletes eloszlás esete (*Extra tananyag*)

Tegyük fel, hogy az emberek egyforma, de nem egyenletes eloszlás szerint érkeznek a vendéglőbe: mindenki – a többitől függetlenül – az $s(x)$ sűrűségfüggvényű, $S(x)$ eloszlásfüggvényű eloszlás szerint érkezik ($A < x < B$). Legyen X az i -ik, Y a j -ik érkezési időpont.

Most erre az általánosabb esetre is levezetjük a kétdimenziós (X, Y) valószínűségi változó sűrűségfüggvényének a képletét. E célból – ugyanúgy, mint korábban – felvesszünk két x, y számot, melyekre igaz, hogy

$$0 < x < y < 1$$

Ezek után felvesszünk x mellett egy kis $[x, x + \Delta x]$ intervallumot és y mellett egy kis $[y, y + \Delta y]$ intervallumot úgy, hogy teljesüljön az $x + \Delta x < y$ egyenlőtlenség is. (X, Y) sűrűségfüggvényét hányadossal közelítjük:

$$f(x, y) \approx \frac{P(x < X < x + \Delta x \text{ és } y < Y < y + \Delta y)}{\Delta x \Delta y}$$

A számlálóban álló

$$x < X < x + \Delta x \quad \text{és} \quad y < Y < y + \Delta y$$

esemény – ugyanúgy, mint korábban – azt jelenti, hogy

az i -ik legkisebb szám, amit X jelöl, az $[x, x + \Delta x)$ intervallumba esik,

és

a j -ik legkisebb szám, amit Y jelöl, az $[y, y + \Delta y)$ intervallumba esik,

ami – részletesebben kifejtve a következőket jelenti:

legalább	1	X pont van az	$[x, x + \Delta x)$	intervallumban, és
legalább	1	Y pont van az	$[y, y + \Delta y)$	intervallumban, és
pontosan	$(i - 1)$	pont van az	$[0, X)$	intervallumban, és
pontosan	$(j - i - 1)$	pont van az	$[X, Y)$	intervallumban, és
pontosan	$(n - j)$	pont van az	$[Y, 1]$	intervallumban.

Ez – jó közelítéssel – azt jelenti, hogy

pontosan	$(i-1)$	pont van a	$[0, x)$	intervallumban, és
pontosan	1	pont van az	$[x, x + \Delta x)$	intervallumban, és
pontosan	$(j-i-1)$	pont van az	$[x + \Delta x, y)$	intervallumban, és
pontosan	1	pont van az	$[y, y + \Delta y)$	intervallumban, és
pontosan	$(n-j)$	pont van az	$[y + \Delta y, 1]$	intervallumban.

A polihipergeometrikus eloszlás formuláját felhasználva most azt kapjuk, hogy az

$$x < X < x + \Delta x \quad \text{és} \quad y < Y < y + \Delta y$$

esemény valószínűsége körülbelül egyenlő a

$$\frac{n!}{(i-1)! 1! (j-i-1)! 1! (n-j)!}$$

konstans és

$$(S(x))^{i-1} (s(x) \Delta x)^1 (S(y) - S(x + \Delta x))^{j-i-1} (s(y) \Delta y)^1 (1 - S(y + \Delta y))^{n-j}$$

kifejezés szorzatával. Eldobhatunk néhány fölösleges zárójelet, tényezőt és kitevőt, ami által a teljes formula, vagyis maga szorzat így egyszerűsödik:

$$\frac{n!}{(i-1)! (j-i-1)! (n-j)!} (S(x))^{i-1} s(x) \Delta x (S(y) - S(x + \Delta x))^{j-i-1} s(y) \Delta y (1 - S(y + \Delta y))^{n-j}$$

Osztva $\Delta x \cdot \Delta y$ -nal osztva, majd $\Delta x \rightarrow 0$, $\Delta y \rightarrow 0$ határérték mellett megkapjuk a sűrűségfüggvényt:

$$f(x, y) = \frac{n!}{(i-1)! (j-i-1)! (n-j)!} S(x)^{i-1} s(x) (S(y) - S(x))^{j-i-1} f(y) (1 - S(y))^{n-j}$$

3.6. Kovariancia (*Extra tananyag*)

Először egy alapvető formulát állítunk fel a sűrűségfüggvény segítségével, utána kiszámoljuk a kovarianciát.

Mivel minden sűrűségfüggvénynek a teljes értelmezési tartományán vett integrálja 1 -gyel egyenlő, ezért fennáll az alábbi egyenlőség:

$$\int_0^1 \int_0^y \frac{n!}{(i-1)! (j-i-1)! (n-j)!} x^{i-1} (y-x)^{j-i-1} (1-y)^{n-j} dx dy = 1$$

Ha valaki ezt az egyenlőséget közvetlenül integrálással szeretné megkapni, akkor ennek sincs különösebb akadálya, hiszen parciális integrálásokkal és kellő türelemmel az $\int_0^1 \int_0^y x^{i-1} (y-x)^{j-i-1} (1-y)^{n-j} dx dy$ integrál értékére kijön:

$$\int_0^1 \int_0^y x^{i-1} (y-x)^{j-i-1} (1-y)^{n-j} dx dy = \frac{(i-1)! (j-i-1)! (n-j)!}{n!} \quad \text{alapvető formula}$$

Érdeemes ebben az *alapvető formulában* i , j és n helyett $(i+1)$ -et $(j+2)$ -t és $(n+2)$ -t írni. Ezt az összefüggést kapjuk:

$$\begin{aligned} \int_0^1 \int_0^y x^{(i+1)-1} (y-x)^{(j+2)-(i+1)-1} (1-y)^{(n+2)-(j+2)} dx dy = \\ = \frac{((i+1)-1)! ((j+2)-(i+1)-1)! ((n+2)-(j+2))!}{(n+2)!} \end{aligned}$$

amiből egyszerűsítés után ez jön ki:

$$\int_0^1 \int_0^y x^i (y-x)^{j-i} (1-y)^{n-j} dx dy = \frac{i! (j-i)! (n-j)!}{(n+2)!} \quad \text{származtatott formula}$$

Az itt következő várható értéket könnyen meg lehet határozni a most kapott származtatott formula segítségével:

$$\begin{aligned} E(X(Y-X)) &= \\ &= \int_0^1 \int_0^y x(y-x) \cdot \frac{n!}{(i-1)! (j-i-1)! (n-j)!} x^{i-1} (y-x)^{j-i-1} (1-y)^{n-j} dx dy = \\ &= \frac{n!}{(i-1)! (j-i-1)! (n-j)!} \int_0^1 \int_0^y x(y-x) \cdot x^{i-1} (y-x)^{j-i-1} (1-y)^{n-j} dx dy = \\ &= \frac{n!}{(i-1)! (j-i-1)! (n-j)!} \cdot \int_0^1 \int_0^y x^i (y-x)^{j-i} (1-y)^{n-j} dx dy = \\ &= \frac{n!}{(i-1)! (j-i-1)! (n-j)!} \cdot \frac{i! (j-i)! (n-j)!}{(n+2)!} = \frac{i(j-i)}{(n+1)(n+2)} \end{aligned}$$

Ezek után a szorzat várható értéke egyszerűen adódik:

$$E(XY) = E(X(Y-X)) + E(X^2) = \frac{i(j-i)}{(n+1)(n+2)} + \frac{i(i+1)}{(n+1)(n+2)} = \frac{i(j+1)}{(n+1)(n+2)}$$

Végül a kovariancia:

$$\text{COV}(XY) = E(XY) - E(X)E(Y) = \frac{i(j+1)}{(n+1)(n+2)} - \frac{i}{n+1} \frac{j}{n+1}$$

3.7. Gyakorló feladatok

1. Öt jóbarát mindegyike a többitől függetlenül dél és 1 óra között egyenletes eloszlás szerint érkezik a menzára. Legyen X az az időpont, amikor a már négyen vannak, és legyen Y az az időpont, amikor a már mind az öten ott vannak. Ha van rá lehetősége, szimulálja a jelenséget és a valószínűségi változókat! Az alábbi feladatok megértésében és megoldásában sokat segít, ha érti, látja, hogy mi történik, hogyan dolgozik a véletlen!

(a) Fókuszáljunk először csak az Y valószínűségi változóra!

i. Határozza meg az

$$Y < 0.75$$

esemény valószínűségét! (Vegye észre, hogy Vegye észre, hogy ez az esemény azt jelenti, esemény azt jelenti, hogy mindenki odaér háromnegyed egy előtt.)

ii. Határozza meg Y eloszlásfüggvényének a képletét!

iii. Határozza meg Y sűrűségfüggvényének a képletét az eloszlásfüggvény deriválásával!

iv. Határozza meg Y sűrűségfüggvényének a képletét az eloszlásfüggvény felhasználása nélkül!

(b) Fókuszáljunk most csak az X valószínűségi változóra!

i. Határozza meg az

$$X < 0.25$$

esemény valószínűségét! (Vegye észre, hogy Vegye észre, hogy ez az esemény azt jelenti, esemény azt jelenti, hogy legalább négyen odaérnek negyed egy előtt. A binomiális eloszlás képletét összegzőkkel kombinálva ügyesen lehet használni.)

- ii. Határozza meg X eloszlásfüggvényének a képletét!
 - iii. Határozza meg X sűrűségfüggvényének a képletét az eloszlásfüggvény deriválásával!
 - iv. Határozza meg X sűrűségfüggvényének a képletét az eloszlásfüggvény felhasználása nélkül!
- (c) Most foglalkozunk az (X, Y) kétdimenziós valószínűségi változóval!

- i. Határozza meg az

$$X < 0.25 \text{ és } Y < 0.75$$

esemény valószínűségét! (Vegye észre, hogy ez az esemény azt jelenti, hogy legalább négyen odaérnek negyed egy előtt, és mindenki odaér háromnegyed egy előtt. A polinomiális eloszlás összegzésekkel kombinálva ügyesen lehet használni.)

- ii. Határozza meg (X, Y) sűrűségfüggvényének a képletét az eloszlásfüggvény deriválásával!
- iii. Határozza meg (X, Y) eloszlásfüggvényének a képletét az eloszlásfüggvény felhasználása nélkül!

4. Kísérleti eredmények függvényének a várható értéke

4.1. Diszkrét eset

Ha egy (X, Y) kétdimenziós diszkrét valószínűségi változó mellett egy $z = t(x, y)$ transzformációt is tekintünk, és képezzük a $t(X, Y)$ valószínűségi változót, akkor az

$$\sum_{(x,y)} t(x, y) p(x, y)$$

számot a transzformációval nyert $t(X, Y)$ **valószínűségi változó várható értékének** nevezzük, és $E(t(X, Y))$ -vel jelöljük:

$$E(t(X, Y)) = \sum_{(x,y)} t(x, y) p(x, y)$$

Ha egy diszkrét eloszlást a síkon csak úgy önmagában – mindenféle valószínűségi változó nélkül – vizsgálunk, akkor a

$$\sum_{(x,y)} t(x, y) p(x, y)$$

számot a szóbanforgó eloszlásból a $z = t(x, y)$ **transzformációval kapott eloszlás várható értékének** nevezzük.

Megjegyzés. Természetesen a $Z = t(X, Y)$ valószínűségi változó várható értékét meghatározhatjuk úgy is, hogy meghatározzuk a $Z = t(X, Y)$ valószínűségi változó lehetséges z értékeit és az eloszlásának az $r(z)$ képletét, és a kapott eloszlás

$$E(Z) = \sum_z z r(z)$$

várható értékét kiszámoljuk.

4.2. Folytonos eset

Ha egy (X, Y) kétdimenziós folytonos valószínűségi változó mellett egy $z = t(x, y)$ folytonos transzformációt is tekintünk, és képezzük a $t(X, Y)$ valószínűségi változót, akkor az

$$\iint_{\mathbb{R}^2} t(x, y) f(x, y) dx dy$$

számot a $t(X, Y)$ **valószínűségi változó várható értékének** nevezzük, és $E(t(X, Y))$ -vel jelöljük:

$$E(t(X, Y)) = \iint_{\mathbb{R}^2} t(x, y) f(x, y) dx dy$$

Ha egy folytonos eloszlást a síkon csak úgy önmagában – mindenféle valószínűségi változó nélkül – vizsgálunk, akkor az

$$\iint_{\mathbb{R}^2} t(x, y) f(x, y) dx dy$$

számot a $z = t(x, y)$ **transzformációval kapott eloszlás várható értékének** nevezzük.

Megjegyzés. Természetesen a $Z = t(X, Y)$ valószínűségi változó várható értékét meghatározhatjuk úgy is, hogy meghatározzuk a $Z = t(X, Y)$ valószínűségi változó eloszlásának $r(z)$ sűrűségfüggvényét, és a kapott eloszlás

$$E(Z) = \int_{-\infty}^{\infty} z r(z) dz$$

várható értékét kiszámoljuk.

Feladat: Tegyük fel, hogy n jóbarát érkezik a menzára déli 12 óra és délután 1 óra között egymástól függetlenül (folytonos) egyenletes eloszlás szerint. Az utolsó és az első érkezési között eltelt időtartamnak mennyi a várható értéke?

Megoldás: Az első érkezési időpont legyen X , az utolsó Y , az X és Y között eltelt időtartam $Z = Y - X$. Ekkor (X, Y) sűrűségfüggvénye – mint korábban meghatároztuk:

$$f(x, y) = n(n-1)(y-x)^{n-2} \quad (0 < x < y < 1)$$

Z sűrűségfüggvényét is kiszámítottuk a korábbiakban:

$$r(z) = n(n-1)(z^{n-2} - z^{n-1}) \quad (0 < z < 1)$$

Most háromféleképpen is felírjuk Z várható értékét:

1. $Z = Y - X$ várható értéke az $f(x, y)$ sűrűségfüggvényéből az $y - x$ függvénnyel így írható fel:

$$\begin{aligned} E(Z) &= \iint_{\mathbb{R}^2} (y-x) \cdot f(x, y) dx dy = \iint_{0 < x < y < 1} (y-x) \cdot n(n-1)(y-x)^{n-2} dx dy \\ &= \int_0^1 \left(\int_0^y n(n-1)(y-x)^{n-1} dx \right) dy = \dots = \frac{n-1}{n+1} \end{aligned}$$

2. A várható érték a Z sűrűségfüggvényéből:

$$E(Z) = \int_{-\infty}^{\infty} z \cdot r(z) dz = \int_0^1 z \cdot (n(n-1)(z^{n-2} - z^{n-1})) dz = \dots = \frac{n-1}{n+1}$$

3. A legegyszerűbb, ha a várható értéket a korábbról már ismert

$$E(Y) = \frac{n}{n+1} \quad E(X) = \frac{1}{n+1}$$

tények alapján különbségként írjuk fel:

$$E(Z) = E(Y) - E(X) = \frac{n}{n+1} - \frac{1}{n+1} = \frac{n-1}{n+1}$$

4.3. Szorzat várható értéke

Fontossága miatt külön megemlítjük az XY szorzat várható értékének a képletét.

Diszkrét esetben:

$$E(XY) = \sum_{(x,y)} xy \cdot p(x, y)$$

Folytonos esetben:

$$\iint_{\mathbb{R}^2} xy \cdot f(x, y) dx dy$$

4.4. Gyakorló feladatok

1. Tegyük fel, hogy az (X, Y) kétdimenziós valószínűségi változó eloszlása táblázattal megadva így fest:

y					
4	0.060	0.120	0.090	0.030	
3	0.080	0.160	0.120	0.040	
2	0.030	0.060	0.045	0.015	
1	0.020	0.040	0.030	0.010	
0	0.010	0.020	0.015	0.005	
	0	1	2	3	x

Határozza meg

- (a) az XY szorzatának,
- (b) az Y/X hányados,
- (c) az X/Y hányados

a várható értékét!

2. Egy (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye $f(x, y) = x + y$ ($x > 0, y > 0, x + y < 1$). Határozza meg

- (a) az XY szorzatának,
- (b) az Y/X hányados,
- (c) az X/Y hányados

a várható értékét!

3. Egy (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye $f(x, y) = 4x + 2y$ ($x > 0, y > 0, x + y < 1$). Határozza meg

- (a) az XY szorzatának,
- (b) az Y/X hányados,
- (c) az X/Y hányados

a várható értékét!

5. NSZT a kísérleti eredmények függvényének az átlagára

5.1. Diszkrét eset

Képzeljük el, hogy egy (X, Y) kétdimenziós valószínűségi változóra sok kísérletet végzünk. A kísérletek számát jelölje N , a kísérleti eredményeket, melyek véletlen számpárok, jelölje

$$(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$$

Állítás: Ha kísérleti eredményeket egy $t(x, y)$ függvénybe helyettesíthetjük, majd az így kapott

$$t(X_1, Y_1), t(X_2, Y_2), \dots, t(X_N, Y_N)$$

értékeket átlagoljuk, akkor ez az átlag – nagy kísérletszám esetén – körülbelül a $t(X, Y)$ várható értékével egyenlő:

$$\frac{t(X_1, Y_1) + t(X_2, Y_2) + \dots + t(X_N, Y_N)}{N} \approx E(t(X, Y))$$

Vázlatos bizonyítás. (*Extra tananyag.*) Az (X, Y) kétdimenziós valószínűségi változó lehetséges értékeit most indexelve soroluk fel:

$$(x_1, y_1), (x_2, y_2), \dots$$

A lehetséges értékek valószínűségeit is indexekkel látjuk el:

$$p_1, p_2, \dots$$

Tehát minden i esetén p_i megadja az (x_i, y_i) számpár valószínűségét. Jelölje továbbá

$$N_1, N_2, \dots$$

a lehetséges értékek gyakoriságait N kísérlet során. Tehát minden i esetén N_i azt mutatja, hogy az (x_i, y_i) számpár hányszor következik be az N kísérlet során. Tudjuk, hogy – nagy N esetén – minden i -re az N_i/N relatív gyakoriság közel van a p_i valószínűséghez. A

$$\frac{t(X_1, Y_1) + t(X_2, Y_2) + \dots + t(X_N, Y_N)}{N}$$

kifejezés számlálója N darab tagból áll. Közöttük a t függvény argumentumaként az (x_1, y_1) számpár N_1 -szer, az (x_2, y_2) számpár N_2 -szer, ... fordul elő. Ezért a számláló értéke:

$$t(X_1, Y_1) + t(X_2, Y_2) + \dots + t(X_N, Y_N) = N_1 \cdot t(x_1, y_1) + N_2 \cdot t(x_2, y_2) + \dots$$

Az átlag öt jelentő tört értéke:

$$\begin{aligned} \frac{t(X_1, Y_1) + t(X_2, Y_2) + \dots + t(X_N, Y_N)}{N} &= \frac{N_1 \cdot t(x_1, y_1) + N_2 \cdot t(x_2, y_2) + \dots}{N} = \\ &= \frac{N_1}{N} \cdot t(x_1, y_1) + \frac{N_2}{N} \cdot t(x_2, y_2) + \dots \approx p_1 \cdot t(x_1, y_1) + p_2 \cdot t(x_2, y_2) + \dots = E(t(X, Y)) \end{aligned}$$

5.2. Folytonos eset

Képzeld el, hogy egy (X, Y) kétdimenziós folytonos valószínűségi változóra sok kísérletet végzünk. A kísérletek számát jelölje N , a kísérleti eredményeket, melyek véletlen számpárok, jelölje

$$(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$$

Állítás: Ha kísérleti eredményeket egy $t(x, y)$ folytonos függvénybe helyettesíthetjük, majd az így kapott

$$t(X_1, Y_1), t(X_2, Y_2), \dots, t(X_N, Y_N)$$

értékeket átlagoljuk, akkor ez az átlag – nagy kísérletszám esetén – körülbelül a $t(X, Y)$ várható értékével egyenlő:

$$\frac{t(X_1, Y_1) + t(X_2, Y_2) + \dots + t(X_N, Y_N)}{N} \approx E(t(X, Y))$$

Vázlatos bizonyítás (Extra tananyag): Vegyük fel az x -tengelyen az x_i , ($i = \dots, -1, 0, 1, \dots$) pontokat, az y -tengelyen az y_j , ($j = \dots, -1, 0, 1, \dots$) pontokat úgy, hogy szomszédos x_i, x_{i+1} és y_j, y_{j+1} pontok távolsága, amit Δx_i -vel illetve Δy_j -vel jeleljük, kicsi legyen minden i -re és j -re. Az X valószínűségi változó megfigyelt értékét az $[x_i, x_{i+1})$ intervallumban kerekítjük le x_i -re, és a kapott értéket jelöljük X^* -gal:

$$X^* = x_i \quad \text{ha} \quad x_i \leq X < x_{i+1}$$

Az Y valószínűségi változó megfigyelt értékét az $[y_j, y_{j+1})$ intervallumban kerekítjük le y_j -re, és a kapott értéket jelöljük Y^* -gal:

$$Y^* = y_j \quad \text{ha} \quad y_j \leq Y < y_{j+1}$$

Az (X^*, Y^*) diszkrét valószínűségi változó, melynek lehetséges értékei az (x_i, y_j) , ($i, j = \dots, -1, 0, 1, \dots$) számpárok, és minden i és j esetén-re az (x_i, y_j) számpár valószínűsége:

$$P(X^* = x_i \text{ és } Y^* = y_j) = \int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} f(x, y) dx dy$$

$t(X^*, Y^*)$ várható értéke:

$$E(t(X^*, Y^*)) = \sum_{(i,j)} t(x_i, y_j) P(X^* = x_i \text{ és } Y^* = y_j) = \sum_{(i,j)} t(x_i, y_j) \int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} f(x, y) dx dy$$

Mivel X^* és X , illetve Y^* és Y csak kicsit térnek el egymástól, $t(X, Y)$ értéke is csak kicsit tér el $t(X^*, Y^*)$ -tól, ezért a kísérleti eredmények átlagai is csak kicsit térnek el:

$$\frac{t(X_1, Y_1) + t(X_2, Y_2) + \dots + t(X_N, Y_N)}{N} \approx \frac{t(X_1^*, Y_1^*) + t(X_2^*, Y_2^*) + \dots + t(X_N^*, Y_N^*)}{N}$$

Mivel Δx_i minden i -re és Δy_j minden i -re és j -re kicsik, az alábbi közelítésekkel élhetünk:

$$P(X^* = x_i \text{ és } Y^* = y_j) = \int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} f(x, y) dx dy \approx f(x_i, y_j) \cdot \Delta x_i \Delta y_j \quad (i, j = \dots, -1, 0, 1, \dots)$$

$$\begin{aligned} E(t(X^*, Y^*)) &= \sum_{(i,j)} t(x_i, y_j) \int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} f(x, y) dx dy \approx \sum_{(i,j)} t(x_i, y_j) f(x_i, y_j) \cdot \Delta x_i \Delta y_j \approx \\ &\approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x, y) f(x, y) dx dy \end{aligned}$$

Mindezt felhasználva kapjuk, hogy az (X, Y) -ra végzett kísérleti eredmények $t(x, y)$ -ba való helyettesítésekor adódó értékeinek az átlaga közel van $t(X, Y)$ várható értékéhez:

$$\begin{aligned} \frac{t(X_1, Y_1) + t(X_2, Y_2) + \dots + t(X_N, Y_N)}{N} &\approx \frac{t(X_1^*, Y_1^*) + t(X_2^*, Y_2^*) + \dots + t(X_N^*, Y_N^*)}{N} \approx \mathbf{E}(t(X^*, Y^*)) \approx \\ &\approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x, y) f(x, y) dx dy = \mathbf{E}(t(X, Y)) \end{aligned}$$

Folytonos esetben:

$$\mathbf{E}(X Y) = \iint_{\mathbf{R}^2} x y f(x, y) dx dy$$

5.3. Gyakorló feladatok

- 1.
- 2.

6. Vetület- és feltételes eloszlások

6.1. Vetület eloszlások

Állítás: Ha a kétdimenziós (X, Y) valószínűségi változó sűrűségfüggvénye $f(x, y)$, akkor az X valószínűségi változó $f_1(x)$ sűrűségfüggvénye integrálként áll elő $f(x, y)$ -ből:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

Hasonlóképpen az Y valószínűségi változó $f_2(y)$ sűrűségfüggvénye is integrálként áll elő $f(x, y)$ -ből:

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Diszkrét esetben az ezekkel analóg szabályokat összegzési szabályoknak hívtuk. Most joggal nevezhetjük ezeket a szabályokat **integrálási szabályoknak**. **Vázlatos bizonyítás.** A vízszintes tengelyen felvett $[x, x + \Delta x]$ intervallum

egy sávot definiál a síkon:

$$S_{[x, x+\Delta x]} = \{(x, y) : x \in [x, x + \Delta x]\}$$

és az $X \in [x, x + \Delta x]$ esemény ekvivalens az $(X, Y) \in S_{[x, x+\Delta x]}$ eseménnyel. E tény alapján:

$$\begin{aligned} f_1(x) &\approx \frac{\mathbb{P}(X \in [x, x + \Delta x])}{\Delta x} = \frac{\mathbb{P}((X, Y) \in S_{[x, x+\Delta x]})}{\Delta x} = \\ &= \frac{\iint_{S_{[x, x+\Delta x]}} f(x, y) dx dy}{\Delta x} = \frac{\int_x^{x+\Delta x} \left(\int_{-\infty}^{\infty} f(x, y) dy \right) dx}{\Delta x} \approx \int_{-\infty}^{\infty} f(x, y) dy \end{aligned}$$

Ha a kétdimenziós (X, Y) valószínűségi változó sűrűségfüggvénye $f(x, y)$, akkor – mint korábban már láttuk, az X valószínűségi változó $f_1(x)$ sűrűségfüggvénye integrálként áll elő $f(x, y)$ -ből:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

Hasonlóképpen az Y valószínűségi változó $f_2(y)$ sűrűségfüggvénye is integrálként áll elő $f(x, y)$ -ből:

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

6.2. Feltételes sűrűségfüggvények

Amikor egy kétdimenziós (X, Y) valószínűségi változóval van dolgunk, előfordulhat, hogy X értékét megtudjuk: $X = x$, de a másik koordinátát, Y -t nem tudjuk: Y -t továbbra is valószínűségi változónak kell tekintenünk. Felmerül a kérdés: milyen eloszlással modellezzük Y -t az $X = x$ feltétel mellett? Belátjuk, hogy az $X = x$ feltétel mellett Y feltételes sűrűségfüggvényét hányadosként kaphatjuk meg:

$$f_{2|1}(y|x) = \frac{f(x, y)}{f_1(x)}$$

Hasonlóképpen, az $Y = y$ feltétel mellett X feltételes sűrűségfüggvénye:

$$f_{1|2}(x|y) = \frac{f(x, y)}{f_2(y)}$$

Vázlatos bizonyítás. A két formula közül az első vezetjük le:

$$\begin{aligned} f_{2|1}(y|x) &\approx \frac{\mathbb{P}(Y \in [y, y + \Delta y] \mid X = x)}{\Delta y} \approx \\ &\approx \frac{\mathbb{P}(Y \in [y, y + \Delta y] \mid X \in [x, x + \Delta x])}{\Delta y} = \\ &= \frac{\left(\frac{\mathbb{P}(X \in [x, x + \Delta x] \text{ és } Y \in [y, y + \Delta y])}{\mathbb{P}(X \in [x, x + \Delta x])} \right)}{\Delta y} = \\ &= \frac{\left(\frac{\mathbb{P}(X \in [x, x + \Delta x] \text{ és } Y \in [y, y + \Delta y])}{\Delta x \Delta y} \right)}{\left(\frac{\mathbb{P}(X \in [x, x + \Delta x])}{\Delta x} \right)} \approx \frac{f(x, y)}{f_1(x)} \end{aligned}$$

Példa: Ha Y egyenletes eloszlást követ 0 és 1 között, és az $Y = y$ feltétel mellett X egyenletes eloszlást követ 0 és y között, akkor

$$\begin{aligned} f_2(y) &= 1 \quad (0 < y < 1) \\ f_{1|2}(x|y) &= \frac{1}{y} \quad (0 < x < y) \end{aligned}$$

és – mint korábban meghatároztuk – az (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye:

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

X feltétel nélküli sűrűségfüggvényét is kiszámítottuk:

$$f_1(x) = -\ln(x) \quad (0 < x < 1)$$

Most – hányadosként – felírjuk Y feltételes sűrűségfüggvényét az $X = x$ feltétel mellett X :

$$f_{2|1}(y|x) = \frac{\frac{1}{y}}{-\ln(x)} \quad (0 < x < y < 1)$$

6.3. Feltételes sűrűségfüggvények rendszere

Megjegyzés. Nyilvánvaló, hogy az (első fejezetben tárgyalt)

$$f(x, y) = f_1(x) \cdot f_{2|1}(y|x)$$

szorzási szabály és az (előző pontban levezetett)

$$f_{2|1}(y|x) = \frac{f(x, y)}{f_1(x)}$$

osztási szabály egyenértékű egymással. Hasonlóképpen az

$$f(x, y) = f_2(y) \cdot f_{1|2}(x|y)$$

szorzási szabály és az

$$f_{1|2}(x|y) = \frac{f(x, y)}{f_2(y)}$$

osztási szabály egyenértékű egymással.

Ha minden x mellett tekintjük a $f_{2|1}(y|x)$ sűrűségfüggvényt, akkor sűrűségfüggvényeknek egy rendszerét kapjuk, amit az Y valószínűségi változó X -re vonatkozó **feltételes sűrűségfüggvény-rendszerének** nevezzük.

Ha minden y mellett tekintjük a $f_{1|2}(x|y)$ sűrűségfüggvényt, akkor sűrűségfüggvényeknek egy rendszerét kapjuk, amit az X valószínűségi változó Y -re vonatkozó **feltételes sűrűségfüggvény-rendszerének** nevezzük.

Egy adott $f(x, y)$ kétdimenziós sűrűségfüggvényből, az

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

integrálási szabállyal meghatározhatjuk az $f_1(x)$ vetület-sűrűségfüggvényt, majd pedig a

$$f_{2|1}(y|x) = \frac{f(x, y)}{f_1(x)}$$

hányados szabállyal az $f_{2|1}(y|x)$ feltételes sűrűségfüggvények rendszerét. Fordítva, az $f_1(x)$ vetület-sűrűségfüggvényből és a feltételes sűrűségfüggvények $f_{2|1}(y|x)$ rendszeréből a

$$f(x, y) = f_1(x) \cdot f_{2|1}(y|x)$$

szorzási szabállyal a megkaphatjuk az $f(x, y)$ kétdimenziós sűrűségfüggvényt. Ezért az $f(x, y)$ kétdimenziós sűrűségfüggvény ismerete egyenértékű azzal, hogy ismerjük az $f_1(x)$ vetület-sűrűségfüggvényt és a feltételes sűrűségfüggvények $f_{2|1}(y|x)$ rendszerét. Hasonlóképpen kiadódik az is, hogy a az $f(x, y)$ kétdimenziós sűrűségfüggvény ismerete egyenértékű azzal, hogy ismerjük az $f_2(y)$ vetület-sűrűségfüggvényt és a feltételes sűrűségfüggvények $f_{1|2}(x|y)$ rendszerét.

6.4. Sűrűségfüggvények keverése

A feltételes sűrűségfüggvényekkel kapcsolatos

$$f(x, y) = f_1(x) f_{2|1}(y|x)$$

szorzási szabály és a vetítésre vonatkozó

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

integrálási szabály kombinálásával adódik az alábbi formula:

$$f_1(x) = \int_{-\infty}^{\infty} f_2(y) f_{1|2}(x|y) dy$$

A koordináták szerepének felcserélése pedig a

$$f_2(y) = \int_{-\infty}^{\infty} f_1(x) f_{2|1}(y|x) dy$$

formulát adja.

Ezek a formulák a korábban tanult

$$p_1(x) = \sum_y p_2(y) p_{1|2}(x|y)$$

$$p_2(y) = \sum_x p_1(x) p_{2|1}(y|x)$$

diszkrét formulák folytonos megfelelői. Ezért ezekre az integrál-formulákra hivatkozhatunk úgy is mint a **teljes valószínűség formulái kétdimenziós folytonos valószínűségi változókra**. A jobboldali integrál-kifejezéseket egyfajta keverésnek is felfoghajuk: például az

$$f_2(y) = \int_{-\infty}^{\infty} f_1(x) f_{2|1}(y|x) dx$$

formula azt mutatja, hogy az $f_2(y)$ függvényértéket úgy kaphatjuk meg, hogy a az $f_{2|1}(y|x)$ függvényértékeket az $f_1(x)$ súlyfüggvény szerint keverjük. Tehát

- az Y valószínűségi változó sűrűségfüggvénye (eloszlása) nem más, mint az X feltételes sűrűségfüggvényeinek (feltételes eloszlásainak) keveréke az X sűrűségfüggvénye (eloszlása) szerint.

Hasonlóképpen

- az X valószínűségi változó sűrűségfüggvénye (eloszlása) nem más, mint az X feltételes sűrűségfüggvényeinek (feltételes eloszlásainak) keveréke az Y sűrűségfüggvénye (eloszlása) szerint.

6.5. Feltételes eloszlásfüggvény

A feltételes sűrűségfüggvény integráljaként a feltételes eloszlásfüggvényt kapjuk:

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

illetve

$$F_{1|2}(x|y) = \int_{-\infty}^x f_{1|2}(x|y) dx = P(X < x | Y = y)$$

Megfordítva: a feltételes eloszlásfüggvény deriváltjaként a feltételes sűrűségfüggvényt kapjuk meg

$$f_{2|1}(y|x) = \frac{\partial F_{2|1}(y|x)}{\partial y}$$

$$f_{1|2}(x|y) = \frac{\partial F_{1|2}(x|y)}{\partial x}$$

A feltételes eloszlásfüggvény jelentése:

- Ha (x_1, x_2) egy picike intervallum az x szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az X értéke x_1 és x_2 közé esik, akkor az ilyen kísérletekhez tartozó Y értékeknek körülbelül az

$$F_{2|1}(y|x) \text{ -ed része kisebb } y \text{ -nél}$$

- Ha (y_1, y_2) egy picike intervallum az y szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az Y értéke y_1 és y_2 közé esik, akkor az ilyen kísérletekhez tartozó X értékeknek körülbelül az

$$F_{1|2}(x|y) \text{ -ed része kisebb } x \text{ -nél}$$

6.6. Feltételes valószínűség

Az $y_1 < Y < y_2$ esemény valószínűsége az $X = x$ feltétel mellett a feltételes sűrűségfüggvényből integrálással adódik:

$$P(y_1 < Y < y_2 | X = x) = \int_{y_1}^{y_2} f_{2|1}(y|x) dy$$

Hasonlóan adódik az $x_1 < X < x_2$ esemény valószínűsége az $X = x$ feltétel mellett:

$$P(x_1 < X < x_2 | Y = y) = \int_{x_1}^{x_2} f_{1|2}(x|y) dx$$

Ugyanezeket a feltételes valószínűségeket a feltételes eloszlásfüggvényből különbséggént is megkaphatjuk:

$$P(y_1 < Y < y_2 | X = x) = F_{2|1}(y_2|x) - F_{2|1}(y_1|x)$$

$$P(x_1 < X < x_2 | Y = y) = F_{1|2}(x_2|y) - F_{1|2}(x_1|y)$$

A feltételes valószínűség jelentése:

- Ha (y_1, y_2) egy tetszőleges intervallum az y tengelyen, (x_1, x_2) pedig egy picike intervallum az x szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az X értéke x_1 és x_2 közé esik, akkor az ilyen kísérletekhez tartozó Y értékeknek körülbelül a

$$P(y_1 < Y < y_2 | X = x) \text{-ed része esik az } (y_1, y_2) \text{ intervallumba}$$

- Ha (x_1, x_2) egy tetszőleges intervallum az x tengelyen, (y_1, y_2) pedig egy picike intervallum az y szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az Y értéke y_1 és y_2 közé esik, akkor az ilyen kísérletekhez tartozó X értékeknek körülbelül a

$$P(x_1 < X < x_2 | Y = y) \text{-ed része esik az } (x_1, x_2) \text{ intervallumba}$$

Megjegyzés. A

$$P(y_1 < Y < y_2 | X = x)$$

feltételes valószínűségben a feltétel valószínűsége nulla

$$P(X = x) = 0$$

Ezért a feltételes valószínűség elemi

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

formuláját nem lehet alkalmazni a $P(y_1 < Y < y_2 | X = x)$ valószínűség definiálására.

6.7. Feltételes medián

Az

$$F_{2|1}(y|x) = \frac{1}{2}$$

egyenletet y -ra megoldva, azaz kifejezve az egyenletből y -t az x segítségével, megkapjuk Y mediánját az $X = x$ feltétel mellett.

Hasonlóképpen, az

$$F_{1|2}(x|y) = \frac{1}{2}$$

egyenletet x -re megoldva, azaz kifejezve az egyenletből x -et az y segítségével, megkapjuk X mediánját az $Y = y$ feltétel mellett.

A feltételes medián jelentése:

- Ha (x_1, x_2) egy picike intervallum az x szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az X értéke x_1 és x_2 közé esik, akkor az ilyen kísérletekhez tartozó Y értékeknek körülbelül a

fele kisebb, a fele nagyobb, mint az $X = x$ feltételhez tartozó medián

- Ha (y_1, y_2) egy picike intervallum az y szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az Y értéke y_1 és y_2 közé esik, akkor az ilyen kísérletekhez tartozó X értékeknek körülbelül a

fele kisebb, a fele nagyobb, mint az $Y = y$ feltételhez tartozó medián

6.8. Feltételes várható érték

A feltételes várható érték nem más mint a feltételes eloszlás várható értéke:

$$E(Y|X = x) = \int_{-\infty}^{\infty} y f_{2|1}(y|x) dy$$

illetve

$$E(X|Y = y) = \int_{-\infty}^{\infty} x f_{1|2}(x|y) dx$$

A feltételes várható érték jelentése:

- Ha (x_1, x_2) egy picike intervallum az x szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az X értéke x_1 és x_2 közé esik, akkor az ilyen kísérletekhez tartozó Y értékeknek az átlaga körülbelül az

$E(Y|X = x)$ feltételes várható értékkel egyenlő

- Ha (y_1, y_2) egy picike intervallum az y szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az Y értéke y_1 és y_2 közé esik, akkor az ilyen kísérletekhez tartozó X értékeknek az átlaga körülbelül az

$E(X|Y = y)$ feltételes várható értékkel egyenlő

6.9. Feltételes variancia

A feltételes variancia nem más mint a feltételes eloszlás varianciája:

$$\begin{aligned}\text{VAR}(Y|X = x) &= \int_{-\infty}^{\infty} (y - \text{E}(Y|X = x))^2 f_{2|1}(y|x) dy = \\ &= \left(\int_{-\infty}^{\infty} y^2 f_{2|1}(y|x) dy \right) - (\text{E}(Y|X = x))^2\end{aligned}$$

illetve

$$\begin{aligned}\text{VAR}(X|Y = y) &= \int_{-\infty}^{\infty} (x - \text{E}(X|Y = y))^2 f_{1|2}(x|y) dx = \\ &= \left(\int_{-\infty}^{\infty} x^2 f_{1|2}(x|y) dx \right) - (\text{E}(X|Y = y))^2\end{aligned}$$

A feltételes variancia jelentése:

- Ha (x_1, x_2) egy picike intervallum az x szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az X értéke x_1 és x_2 közé esik, akkor az ilyen kísérletekhez tartozó Y értékeknek a varianciája körülbelül az

$$\text{VAR}(Y|X = x) \text{ feltételes varianciával egyenlő}$$

- Ha (y_1, y_2) egy picike intervallum az y szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az Y értéke y_1 és y_2 közé esik, akkor az ilyen kísérletekhez tartozó X értékeknek a varianciája körülbelül az

$$\text{VAR}(X|Y = y) \text{ feltételes varianciával egyenlő}$$

6.10. Feltételes szórás

A feltételes szórás nem más mint a feltételes eloszlás szórása, vagyis a feltételes variancia négyzetgyöke:

$$\begin{aligned}\text{SD}(Y|X = x) &= \sqrt{\text{VAR}(Y|X = x)} = \\ &= \sqrt{\int_{-\infty}^{\infty} (y - \text{E}(Y|X = x))^2 f_{2|1}(y|x) dy} = \sqrt{\int_{-\infty}^{\infty} y^2 f_{2|1}(y|x) dy - (\text{E}(Y|X = x))^2}\end{aligned}$$

illetve

$$\begin{aligned}\text{SD}(X|Y = y) &= \sqrt{\text{VAR}(X|Y = y)} = \\ &= \sqrt{\int_{-\infty}^{\infty} (x - \text{E}(X|Y = y))^2 f_{1|2}(x|y) dx} = \sqrt{\int_{-\infty}^{\infty} x^2 f_{1|2}(x|y) dx - (\text{E}(X|Y = y))^2}\end{aligned}$$

A feltételes szórás jelentése:

- Ha (x_1, x_2) egy picike intervallum az x szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az X értéke x_1 és x_2 közé esik, akkor az ilyen kísérletekhez tartozó Y értékeknek a szórása körülbelül az

$$SD(Y|X = x) \quad \text{feltételes szórással egyenlő}$$

- Ha (y_1, y_2) egy picike intervallum az y szám körül, és az (X, Y) kétdimenziós valószínűségi változóra annyira sok kísérletet végzünk, hogy még az olyan kísérletekből is sok adódik, melyekre az Y értéke y_1 és y_2 közé esik, akkor az ilyen kísérletekhez tartozó X értékeknek a szórása körülbelül az

$$SD(X|Y = y) \quad \text{feltételes szórással egyenlő}$$

Megjegyzés: A feltételes várható érték fogalmát már a jegyzet első részében értelmeztük diszkrét eloszlásokra – integrálok helyett szummákkal. A feltételes variancia, feltételes szórás is nyilvánvalóan értelmezhető diszkrét eloszlásokra is: a fenti formulákban az integrálok a megfelelő formulákban értelemszerűen szummákkal cserélődnek fel.

6.11. Egy számolás minta példa

Feladat: Választunk egy Y számot 0 és 1 között egyenletes eloszlás szerint. Ha $Y = y$, akkor 0 és y között választunk egy X számot ugyancsak egyenletes eloszlás szerint. És most meghatározzuk a feltétel nélküli- és feltételes sűrűség- és eloszlásfüggvényeket és a feltételes eloszlások jellemzőit!

Megoldás:

- Y sűrűségfüggvényét Y definíciója alapján felírhatjuk:

$$f(y) = 1 \quad \text{ha} \quad 0 < y < 1$$

- A most következő képletek $0 < y < 1$ mellett értendők.

Az $Y = y$ feltétel mellett

X feltételes sűrűségfüggvényét X definíciója alapján fel tudjuk írni:

$$f_{1|2}(x|y) = \frac{1}{y} \quad \text{ha} \quad 0 < x < y$$

X feltételes eloszlásfüggvénye nyilván:

$$F_{1|2}(x|y) = \frac{x}{y} \quad \text{ha} \quad 0 < x < y$$

X feltételes mediánja és feltételes várható értéke nyilván:

$$\frac{y}{2}$$

a feltételes variancia, illetve szórás pedig:

$$\frac{y^2}{12} \quad \text{illetve} \quad \frac{y}{\sqrt{12}}$$

- (X, Y) sűrűségfüggvényét a szorzási szabály alapján kapjuk meg:

$$f(x, y) = f_2(y) f_{1|2}(x|y) = 1 \frac{1}{y} = \frac{1}{y} \quad \text{ha} \quad 0 < x < y < 1$$

- X sűrűségfüggvényét y szerinti integrálással kapjuk meg. $0 < x < 1$ esetén ez adódik :

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_x^1 \frac{1}{y} dy = -\ln(x)$$

- A most következő képletek $0 < x < 1$ mellett értendők.

Y -nak az $X = x$ feltétel mellett

feltételes sűrűségfüggvényét hányadosként kapjuk:

$$f_{2|1}(y|x) = \frac{f(x, y)}{f_1(x)} = \frac{\frac{1}{y}}{-\ln(x)} = -\frac{1}{y \ln(x)} \quad \text{ha } x < y < 1$$

A feltételes eloszlásfüggvényt integrálással kapjuk meg:

$$\begin{aligned} F_{2|1}(y|x) &= \int_{-\infty}^y f_{2|1}(y|x) dy = \\ &= \int_x^y \left(-\frac{1}{y \ln(x)} \right) dy = \frac{1}{-\ln(x)} \int_x^y \frac{1}{y} dy = \frac{1}{-\ln(x)} (\ln(y) - \ln(x)) = \\ &= 1 - \frac{\ln(y)}{\ln(x)} \quad \text{ha } x < y < 1 \end{aligned}$$

A feltételes medián az

$$F_{2|1}(y|x) = 1 - \frac{\ln(y)}{\ln(x)} = \frac{1}{2}$$

egyenlet y -re való megoldásából adódik:

$$\begin{aligned} \frac{\ln(y)}{\ln(x)} &= \frac{1}{2} \\ \ln(y) &= \frac{1}{2} \ln(x) \\ \ln(y) &= \ln(\sqrt{x}) \\ y &= \sqrt{x} \end{aligned}$$

Tehát a feltételes medián \sqrt{x} .

A feltételes várható érték:

$$\begin{aligned} E(Y|X = x) &= \int_{-\infty}^{\infty} y f_{2|1}(y|x) dy = \\ &= \int_x^1 y \left(-\frac{1}{y \ln(x)} \right) dy = \int_x^1 \left(-\frac{1}{\ln(x)} \right) dy = \left(-\frac{1}{\ln(x)} \right) (1 - x) = \\ &= \frac{x - 1}{\ln(x)} \end{aligned}$$

A feltételes második momentum:

$$\begin{aligned} E(Y^2|X = x) &= \int_{-\infty}^{\infty} y^2 f_{2|1}(y|x) dy = \\ &= \int_x^1 y^2 \left(-\frac{1}{y \ln(x)} \right) dy = \int_x^1 y \left(-\frac{1}{\ln(x)} \right) dy = \left(-\frac{1}{\ln(x)} \right) \frac{1 - x^2}{2} = \\ &= \frac{x^2 - 1}{2 \ln(x)} \end{aligned}$$

A feltételes varianciát a

$$\text{VAR}(Y|X = x) = E(Y^2|X = x) - (E(Y|X = x))^2$$

képlet alapján fel lehet írni, amiből gyökvonással a feltételes szórás képletét is megkaphatjuk:

$$\text{VAR}(Y|X = y) = \frac{x^2 - 1}{2 \ln(x)} - \left(\frac{x - 1}{\ln(x)} \right)^2$$
$$\text{SD}(X|Y = x) = \sqrt{\frac{x^2 - 1}{2 \ln(x)} - \left(\frac{x - 1}{\ln(x)} \right)^2}$$

6.12. Gyakorló feladatok

1. Tekintjük a $(0, 0)$, $(1, 1)$, $(1, 2)$, $(0, 1)$ pontok által meghatározott paralelogrammán vett egyenletes eloszlást. Határozza meg

- az $X = x$ feltétel mellett Y feltételes sűrűségfüggvényét, majd pedig
 - (a) a feltételes eloszlásfüggvényét,
 - (b) a feltételes mediánját,
 - (c) a feltételes várható értékét,
 - (d) a feltételes második momentumát,
 - (e) a feltételes varianciáját
 - (f) a feltételes szórását!
- az $Y = y$ feltétel mellett X feltételes sűrűségfüggvényét, majd pedig
 - (a) a feltételes eloszlásfüggvényét,
 - (b) a feltételes mediánját,
 - (c) a feltételes várható értékét,
 - (d) a feltételes második momentumát,
 - (e) a feltételes varianciáját
 - (f) a feltételes szórását!

2. A $0 < x < 1$, $x < y < \frac{1}{x}$ egyenlőtlenségekkel definiált tartományon ("vitorlán") vesszük azt az eloszlást, melynek sűrűségfüggvénye $f(x, y) = \frac{1}{2y}$. Határozza meg

- az $X = x$ feltétel mellett Y feltételes sűrűségfüggvényét, majd pedig
 - (a) a feltételes eloszlásfüggvényét,
 - (b) a feltételes mediánját,
 - (c) a feltételes várható értékét,
 - (d) a feltételes második momentumát,
 - (e) a feltételes varianciáját
 - (f) a feltételes szórását!
- az $Y = y$ feltétel mellett X feltételes sűrűségfüggvényét, majd pedig
 - (a) a feltételes eloszlásfüggvényét,
 - (b) a feltételes mediánját,
 - (c) a feltételes várható értékét,
 - (d) a feltételes második momentumát,

- (e) a feltételes varianciáját
 (f) a feltételes szórását!
3. A $0 < x < 1$, $x < y < \frac{1}{x}$ egyenlőtlenségekkel definiált tartományon ("vitorlán") vesszük azt az eloszlást, melynek sűrűségfüggvénye $f(x, y) = \frac{2x}{y}$. Határozza meg
- az $X = x$ feltétel mellett Y feltételes sűrűségfüggvényét, majd pedig
 - (a) a feltételes eloszlásfüggvényét,
 - (b) a feltételes mediánját,
 - (c) a feltételes várható értékét,
 - (d) a feltételes második momentumát,
 - (e) a feltételes varianciáját
 - (f) a feltételes szórását!
 - az $Y = y$ feltétel mellett X feltételes sűrűségfüggvényét, majd pedig
 - (a) a feltételes eloszlásfüggvényét,
 - (b) a feltételes mediánját,
 - (c) a feltételes várható értékét,
 - (d) a feltételes második momentumát,
 - (e) a feltételes varianciáját
 - (f) a feltételes szórását!

7. Kétdimenziós folytonos valószínűségi változó szimulációja

7.1. A módszer

Általános eset: Tegyük fel, hogy adott a síkon egy folytonos eloszlás, és nekünk az a feladatunk, hogy a kalkulátor vagy a számítógép által előállított 0 és 1 között egyenletes eloszlású véletlen számokra építve definiáljunk olyan kétdimenziós valószínűségi változót, mely a megadott eloszlást követi.

A megadott eloszlásnak az x -tengelyre vetett vetületének az eloszlásfüggvényét jelöljük $F_1(x)$ -szel, inverzét pedig $F_1^{-1}(u)$ -val. (Ha $F_1(x)$ nem nő szigorúan monoton módon az egész számegyenesen, és ebből kifolyólag nem invertálható, viszont $F_1(x)$ egy (A, B) intervallumon szigorúan monoton nő, akkor $F_1(x)$ -nek erre az intervallumra való megszorítása már invertálható. Ilyenkor $F_1^{-1}(u)$ ennek a megszorításnak az inverze.) Az inverz függvény formuláját úgy kapjuk meg, hogy az

$$u = F_1(x)$$

egyenletből kifejezzük x -t u segítségével:

$$x = F_1^{-1}(u)$$

A vízszintes számegyenes x pontjához tartozó függőleges egyenes mentén vett feltételes eloszlás $F_{2|1}(y|x)$ eloszlásfüggvényének inverzét jelöljük $F_{2|1}^{-1}(v|x)$ -szel. Az inverz függvény $F_{2|1}^{-1}(v|x)$ formuláját úgy kapjuk meg, hogy az

$$v = F_{2|1}(y|x)$$

egyenletet megoldjuk y -ra, vagyis az egyenletből y -t kifejezzük x -szel és v -vel:

$$y = F_{2|1}^{-1}(v|x)$$

Ebben a kifejezésben v a változó és x a paraméter.

Ezek után X definíciója:

$$X = F_1^{-1}(\text{RND}_1)$$

Nyilvánvaló, hogy X eloszlásfüggvénye $F_1(x)$. Y -t ezek után definiáljuk:

$$Y = F_{2|1}^{-1}(\text{RND}_2|X)$$

Nyilvánvaló, hogy Y feltételes eloszlásfüggvénye az $X = x$ feltétel mellett $F_{2|1}(y|x)$. E két tény azt adja, hogy (X, Y) a kívánt eloszlást követi.

Speciális eset: Ha $f(x, y)$ direktszorzat alakú, azaz $f(x, y) = f_1(x) \cdot f_2(y)$, akkor X -et és Y -t egymástól függetlenül definiálhatjuk az $F_1(x)$ és $F_2(y)$ eloszlásfüggvények inverzeivel:

$$X = F_1^{-1}(\text{RND}_1)$$

$$Y = F_2^{-1}(\text{RND}_2)$$

1. Példa: Általános eset. Tekintsük az

$$f(x, y) = 3x e^{-x(3+y)} \quad (x, y \geq 0)$$

sűrűségfüggvényű eloszlást a síkon. Ilyen eloszlást követő kétdimenziós valószínűségi változót az alábbiak szerint állíthatunk elő.

A definiálandó X valószínűségi változó sűrűségfüggvényét y szerinti integrálással kapjuk síkbeli sűrűségfüggvényből:

$$f_1(x) = \int_0^{\infty} 3x e^{-x(3+y)} dy = 3e^{-3x} \cdot \int_0^{\infty} x e^{-xy} dy = 3e^{-3x} \cdot [-e^{-xy}]_0^{\infty} = 3e^{-3x} \quad (x \geq 0)$$

Látjuk, hogy 3-paraméterű exponenciális eloszlásról van szó. Az eloszlásfüggvény:

$$F_1(x) = \int_0^x 3e^{-3x} dx = 1 - e^{-3x} \quad (x \geq 0)$$

Az

$$u = F_1(x) \quad \text{azaz} \quad u = 1 - e^{-3x}$$

egyenletből az $F_1(x)$ eloszlásfüggvény $x = F_1^{-1}(u)$ inverze:

$$x = -\frac{\ln(1-u)}{3}$$

Tehát az X valószínűségi változó szimulációja:

$$X = -\frac{\ln(1-\text{RND}_1)}{3}$$

A definiálandó Y valószínűségi változó feltételes sűrűségfüggvénye az $X = x$ ($x \geq 0$) feltétel mellett:

$$f_{2|1}(y|x) = \frac{3x e^{-x(3+y)}}{3e^{-3x}} = x e^{-xy} \quad (y \geq 0)$$

Látjuk, hogy x -paraméterű exponenciális eloszlás jött ki. A feltételes eloszlásfüggvény:

$$F_{2|1}(y|x) = \int_0^y x e^{-xy} dy = 1 - e^{-xy} \quad (y \geq 0)$$

Rögzített x mellett a

$$v = F_{2|1}(y|x) \quad \text{azaz} \quad v = 1 - e^{-xy}$$

egyenletből az $F_{2|1}(y|x)$ feltételes eloszlásfüggvény $y = F_{2|1}^{-1}(v|x)$ inverze:

$$y = -\frac{\ln(1-v)}{x}$$

Tehát az Y valószínűségi változó szimulációja:

$$Y = -\frac{\ln(1 - \text{RND}_2)}{X}$$

Tehát a szimuláció két lépésben történik. Először X -et állítjuk elő egy RND_1 véletlen szám segítségével:

$$X = -\frac{\ln(1 - \text{RND}_1)}{-3}$$

majd utána X és egy másik RND_2 véletlen szám segítségével Y -t:

$$Y = -\frac{\ln(1 - \text{RND}_2)}{X}$$

2. Példa: Speciális eset. Tekintsük az

$$f(x, y) = 2e^{-2x} \cdot 3e^{-3y} \quad (x, y \geq 0)$$

kétdimenziós sűrűségfüggvényt, mely direktszorzat alakú. Nyilván

$$f_1(x) = 2e^{-2x} \quad (x \geq 0) \quad f_2(y) = 3e^{-3y} \quad (y \geq 0)$$

X eloszlásfüggvénye:

$$F_1(x) = \int_0^x 2e^{-2x} dx = 1 - e^{-2x} \quad (x \geq 0)$$

Az $F_1(x)$ eloszlásfüggvény $x = F_1^{-1}(u)$ inverze:

$$x = -\frac{\ln(1-u)}{2}$$

Tehát az X valószínűségi változó szimulációja:

$$X = -\frac{\ln(1 - \text{RND}_1)}{2}$$

Az $F_2(y)$ eloszlásfüggvény:

$$F_2(y) = \int_0^y 3e^{-3y} dy = 1 - e^{-3y} \quad (y \geq 0)$$

Az $F_2(y)$ eloszlásfüggvény $y = F_2^{-1}(v)$ inverze:

$$y = -\frac{\ln(1-v)}{3}$$

Tehát az Y valószínűségi változó szimulációja:

$$y = -\frac{\ln(1 - \text{RND}_2)}{3}$$

7.2. Gyakorló feladatok

1. Kalkulátor vagy a számítógép által előállított 0 és 1 között egyenletes eloszlású véletlen számokra építve definiáljon olyan kétdimenziós valószínűségi változót, mely az alább megadott eloszlásokat követi. Az eloszlásokat a sűrűségfüggvényükkel adjuk meg:

(a) $f(x, y) = 4xy \quad (0 < x, y < 1)$

(b) $f(x, y) = x + y \quad (x > 0, y > 0, x + y < 1)$

(c) $f(x, y) = 4x + 2y \quad (x > 0, y > 0, x + y < 1)$

8. Transzformáció síkról síkra

8.1. Jacobi mátrix

Tekintsünk egy kölcsönösen egyértelmű t transzformációt az (x, y) -síkról az (u, v) -síkra, mely az alábbi két kétváltozós koordináta függvénnyel van megadva:

$$u = u(x, y)$$

$$v = v(x, y)$$

Az inverz transzformációt t^{-1} jelöli, ennek koordináta függvényei:

$$x = x(u, v)$$

$$y = y(u, v)$$

Az inverz transzformáció **Jacobi mátrixa** fontos szerepet játszik abban a formulában, melyet alább megadunk. Definiciója:

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{pmatrix} \frac{\partial x(u, v)}{\partial u} & \frac{\partial x(u, v)}{\partial v} \\ \frac{\partial y(u, v)}{\partial u} & \frac{\partial y(u, v)}{\partial v} \end{pmatrix}$$

8.2. Sűrűségfüggvények közötti kapcsolat

Ha az (x, y) -síkon tekintett $f(x, y)$ sűrűségfüggvényű eloszlást a t transzformációval az (u, v) -síkra transzformáljuk, akkor az új eloszlás sűrűségfüggvénye így fejezhető ki a régivel:

$$s(u, v) = f(x(u, v), y(u, v)) \left| \det \left(\frac{\partial(x, y)}{\partial(u, v)} \right) \right|$$

Vázlatos bizonyítás. Tekintsünk egy (x, y) pontot az (x, y) -síkon, mely a t transzformációnál (u, v) pontnak felel meg az (u, v) -síkon. Vegyünk fel egy kis A halmazt az (x, y) pont körül, mely a t transzformációnál egy B halmaznak felel meg az (u, v) pont körül. Mint ismeretes

$$\frac{A \text{ területe}}{B \text{ területe}} \approx \left| \det \left(\frac{\partial(x, y)}{\partial(u, v)} \right) \right|$$

Ezt felhasználva

$$s(u, v) \approx \frac{\mathbf{P}((U, V) \in B)}{B \text{ területe}} = \frac{\mathbf{P}((X, Y) \in A)}{A \text{ területe}} \frac{A \text{ területe}}{B \text{ területe}} \approx f(x, y) \left| \det \left(\frac{\partial(x, y)}{\partial(u, v)} \right) \right|$$

8.3. Gyakorló feladatok

1. Az (x, y) -sík egységnyezetén vett egyenletes eloszlást transzformáljuk az $u = xy, v = \frac{y}{x}$ transzformációval az (u, v) -síkra.
 - (a) Határozza meg az egységnyezet képét!
 - (b) Határozza meg a kapott eloszlás sűrűségfüggvényének a képletét!
2. Az (x, y) -sík egységnyezetén tekintjük azt az eloszlást, melynek sűrűségfüggvénye $f(x, y) = 4xy$. Az eloszlást az (u, v) -síkra transzformáljuk az $u = xy, v = \frac{y}{x}$ transzformációval. Határozza meg a kapott eloszlás sűrűségfüggvényének a képletét!

9. Lineáris transzformáció síkról síkra

9.1. Mátrixszal való szorzás

Tekintsük a

$$\begin{aligned}u &= a_{11}x + a_{12}y \\v &= a_{21}x + a_{22}y\end{aligned}$$

képletekkel definiált, azaz – mátrixokkal felírva – a

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

képlettel definiált lineáris transzformációt, mely az (x, y) -síkról az (u, v) -síkra képez. Bevezetve a

$$\mathbf{u} = \begin{pmatrix} u \\ v \end{pmatrix}$$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$$

jelöléseket, a transzformáció formálisan így is írható:

$$\mathbf{u} = \mathbf{A}\mathbf{x}$$

Az inverz transzformáció képlete:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{u}$$

ahol \mathbf{A}^{-1} az \mathbf{A} mátrix inverze.

Ha az (x, y) -síkon tekintett $f(\mathbf{x})$ sűrűségfüggvényű eloszlást a lineáris transzformációval az (u, v) -síkra transzformáljuk, akkor az új eloszlás sűrűségfüggvénye így fejezhető ki a régivel:

$$s(\mathbf{u}) = f(\mathbf{A}^{-1}\mathbf{u}) |\det(\mathbf{A}^{-1})|$$

A képletben a $|\det(\mathbf{A}^{-1})|$ szorzófaktor az inverzmátrix determinánsának abszolút értékét jelenti.

9.2. Mátrixszal való szorzás és eltolás

Tekintsük most a

$$\begin{aligned}u &= a_{11}x + a_{12}y + b_1 \\v &= a_{21}x + a_{22}y + b_2\end{aligned}$$

képlettel definiált lineáris transzformációt, mely az (x, y) -síkról az (u, v) -síkra képez. Bevezetve a

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

jelölést, a transzformáció formálisan így is írható:

$$\mathbf{u} = \mathbf{Ax} + \mathbf{b}$$

Az inverz transzformáció képlete:

$$\mathbf{x} = \mathbf{A}^{-1}(\mathbf{u} - \mathbf{b})$$

Ha az (x, y) -síkon tekintett $f(\mathbf{x})$ sűrűségfüggvényű eloszlást a lineáris transzformációval az (u, v) -síkra transzformáljuk, akkor az új eloszlás sűrűségfüggvénye így fejezhető ki a régivel:

$$s(\mathbf{u}) = f(\mathbf{A}^{-1}(\mathbf{u} - \mathbf{b})) |\det(\mathbf{A}^{-1})|$$

9.3. Gyakorló feladatok

1. A síkbeli standard normális eloszlást transzformáljuk az

$$\begin{aligned} u &= 2x + 3y \\ v &= x + 2y \end{aligned}$$

lineáris transzformációval.

- (a) Határozza meg a Jacobi mátrixot!
- (b) Számolja ki a Jacobi determinánst!
- (c) Írja fel a kapott eloszlás sűrűségfüggvényének a képletét!

2. A síkbeli standard normális eloszlást transzformáljuk az

$$\begin{aligned} u &= 2x + 5y \\ v &= x + 2y \end{aligned}$$

lineáris transzformációval.

- (a) Határozza meg a Jacobi mátrixot!
- (b) Számolja ki a Jacobi determinánst!
- (c) Írja fel a kapott eloszlás sűrűségfüggvényének a képletét!

3. A síkbeli standard normális eloszlást transzformáljuk az

$$\begin{aligned} u &= 2x + 6y + 10 \\ v &= x + 2y - 5 \end{aligned}$$

lineáris transzformációval.

- (a) Határozza meg a Jacobi mátrixot!
- (b) Számolja ki a Jacobi determinánst!
- (c) Írja fel a kapott eloszlás sűrűségfüggvényének a képletét!

10. Transzformáció síkról egyenesre

10.1. Általános eset

Tekintsünk egy (X, Y) folytonos kétdimenziós valószínűségi változót. Jelöljük a sűrűségfüggvényét $f(x, y)$ -nal. Legyen $z = t(x, y)$ egy folytonosan differenciálható kétváltozós függvény, mely a síkról a számegyenesre képez: az (x, y) számpárhoz hozzárendeli azt a z számot, melyet a $z = t(x, y)$ képlet definiál. Ha az X és Y valószínűségi változók értékeit behelyettesítjük a $t(x, y)$ függvénybe, akkor egy új Z valószínűségi változót kapunk, melyet $t(X, Y)$ -nal jelölünk. A Z , vagyis a $t(X, Y)$ valószínűségi változó eloszlásfüggvényét jelöljük $R(z)$ -vel, sűrűségfüggvényét $r(z)$ -vel.

Állítás:

$$R(z) = \iint_{A_z} f(x, y) dx dy$$

ahol az A_z halmaz a $(-\infty, z)$ intervallum ősképe a $z = t(x, y)$ transzformációval kapcsolatban:

$$A_z = \{(x, y) : t(x, y) < z\}$$

Bizonyítás. A $Z < z$ esemény ekvivalens az $(X, Y) \in A_z$ eseménnyel, így

$$R(z) = P(Z < z) = P((X, Y) \in A_z) = \iint_{A_z} f(x, y) dx dy$$

A kapott $R(z)$ függvényt z szerint deriválva a sűrűségfüggvényt kapjuk:

$$r(z) = R'(z)$$

10.2. Összeg sűrűségfüggvénye

Mielőtt ezt a részt elolvassa, érdemes átnézni az 1. részből a "Súlyfüggvény (eloszlás) transzformációja kétdimenzióból egydimenzióba" című fejezet "Összeg súlyfüggvénye" című alfejezetét, mert igen erős az analógia a diszkrét és folytonos esetekről szóló szabályok között:

- mint tanultuk, diszkrét esetben az összeg $r(z)$ súlyfüggvényének meghatározásához összegezni kellett a síkbeli $p(x, y)$ súlyfüggvény értékeit az $x + y = z$ egyenletű egyenes mentén,
- mint látni fogjuk, folytonos esetben az összeg $r(z)$ sűrűségfüggvényének meghatározásához integrálni kell a síkbeli $f(x, y)$ sűrűségfüggvény értékeit az $x + y = z$ egyenletű egyenes mentén.

Állítás. Tegyük fel, hogy az (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye $f(x, y)$. X és Y összegét jelöljük Z -vel: $Z = X + Y$. Ekkor Z sűrűségfüggvénye kétféleképpen is kiszámolható $f(x, y)$ -ből:

$$r(z) = \int_{-\infty}^{\infty} f(z - v, v) dv$$

$$r(z) = \int_{-\infty}^{\infty} f(u, z - u) du$$

1. Megjegyzés. Adott z esetén a $(z - v, v)$, illetve az $(u, z - u)$ pontok az $x + y = z$ egyenletű egyenest futják be, ezért mindkét formula azt fejezi ki, hogy $r(z)$ -t úgy kapjuk meg $f(x, y)$ -ből, hogy az értékeit az $x + y = z$ egyenletű egyenes mentén integráljuk.

2. Megjegyzés. Természetesen az integrál kiszámításánál az integrációs tartományból elhagyatók az olyan részek, ahol az integrandus 0-val egyenlő. Ezért ha az $f(x, y)$ sűrűségfüggvény az S tartományon kívül nulla, akkor egy adott z érték esetén a $(-\infty, \infty)$ integrációs tartomány

- az első integrálban az $A_z = \{v : (z - v, v) \in S\}$ halmazzal helyettesíthető:

$$r(z) = \int_{A_z} f(z - v, v) dv$$

- a második integrálban a $B_z = \{u : (u, z - u) \in S\}$ halmazzal helyettesíthető:

$$r(z) = \int_{B_z} f(u, z - u) du$$

Az első formula vázlatos bizonyítása. A $z = x + y$ leképezést több lépésben hajtjuk végre. Először tekintjük az (x, y) -síkról az (u, v) -síkra ható

$$\begin{aligned} u &= x + y \\ v &= y \end{aligned}$$

lineáris leképezést, majd az (u, v) -síkról vetítünk a vízszintes tengelyre. A vízszintes tengelyt általában u -tengelynek szoktunk nevezni, de most z tengelynek is nevezzük. (Egy tengelynek lehet több neve, jele is!) A síkról síkra való transzformáció a vetítéssel és az átnevezéssel együtt a $z = x + y$ leképezést adja.

Az (x, y) -síkról az (u, v) -síkra ható lineáris transzformáció inverze:

$$\begin{aligned} x &= u - v \\ y &= v \end{aligned}$$

Ennek Jacobi mátrixa:

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

A Jacobi determináns nyilván 1-gyel egyenlő. Ezért

$$s(u, v) = f(u - v, v) \cdot 1 = f(u - v, v)$$

Ezek után a vízszintes tengelyre való vetítéssel ezt kapjuk:

$$s_1(u) = \int_{-\infty}^{\infty} s(u, v) dv = \int_{-\infty}^{\infty} f(u - v, v) dv$$

Most u helyett z -t írunk, $s_1(u)$ helyett pedig $r(z)$ -t, és megkapjuk a fenti formulák közül az elsőt:

$$r(z) = \int_{-\infty}^{\infty} f(z - v, v) dv$$

A második formula vázlatos bizonyítása. A $z = x + y$ leképezést most is több lépésben hajtjuk végre, de egy kicsit másképpen. Először tekintjük az (x, y) -síkról az (u, v) -síkra ható

$$\begin{aligned} u &= x \\ v &= x + y \end{aligned}$$

lineáris leképezést, majd az (u, v) -síkról vetítünk a függőleges tengelyre. A függőleges tengelyt általában v -tengelynek szoktunk nevezni, de most z tengelynek is nevezzük. A síkról síkra való transzformáció a vetítéssel és az átnevezéssel együtt a $z = x + y$ leképezést adja.

Az (x, y) -síkról az (u, v) -síkra ható transzformáció inverze:

$$\begin{aligned} x &= u \\ y &= v - u \end{aligned}$$

A Jacobi determináns nyilván most is 1-gyel egyenlő. Ezért

$$s(u, v) = f(u, v - u) \cdot 1 = f(u, v - u)$$

Ezek után a függőleges tengelyre való vetítéssel ezt kapjuk:

$$s_2(v) = \int_{-\infty}^{\infty} s(u, v) du = \int_{-\infty}^{\infty} f(u, v - u) du$$

Most v helyett z -t írunk, $s_2(v)$ helyett pedig $r(z)$ -t, és megkapjuk a fenti formulák közül a másodikat:

$$r(z) = \int_{-\infty}^{\infty} f(u, z - u) du$$

10.3. Konvolúció

Független valószínűségi változók összege (folytonos eset). Ha az X és Y valószínűségi változók függetlenek, akkor $f(x, y) = f_1(x)f_2(y)$, ezért a kapott formula így írható:

$$r(z) = \int_{-\infty}^{\infty} f_1(z - v) f_2(v) dv$$

avagy

$$r(z) = \int_{-\infty}^{\infty} f_1(u) f_2(z - u) du$$

Felismerjük, hogy $r(z)$ nem más, mint $f_1(x)$ és $f_2(y)$ **konvolúciója**.

10.4. Gyakorló feladatok

1. Kör alakú céltáblára lövünk. A kör sugara 1 méter. A találat helye egyenletes eloszlást követ a körlapon. A találat helyének távolsága

- a kör középpontjától
- a kör kerületétől

véletlentől függ. Határozza meg ennek a valószínűségi változónak az

- (a) eloszlásfüggvényét!
- (b) sűrűségfüggvényét!

2. Hatalmas köralakú céltáblára lövünk. A találat helye standrd normális eloszlást követ a körlapon. A találat helyének

- a kör középpontjától való távolsága
- a kör középpontjától való távolságának a négyzete

véletlentől függ. Határozza meg ennek a valószínűségi változónak az

- (a) eloszlásfüggvényét!
- (b) sűrűségfüggvényét!

3. A $(0, 0)$, $(1, 1)$, $(1, 2)$, $(0, 1)$ pontok által meghatározott paralelogrammán vett egyenletes eloszlást veítjük

- (a) az x -tengelyre,
- (b) az y -tengelyre

Milyen eloszlásokat kapunk a tengelyeken?

4. A $0 < x < 1$, $x < y < \frac{1}{x}$ egyenlőtlenségekkel definiált tartományon ("vitorlán") vesszük azt az eloszlást, melynek sűrűségfüggvénye $f(x, y) = \frac{1}{2y}$. Határozza meg

- (a) az x -tengelyre,
- (b) az y -tengelyre

vetett vetületeloszlás sűrűségfüggvényének a képletét!

5. A $0 < x < 1$, $x < y < \frac{1}{x}$ egyenlőtlenségekkel definiált tartományon ("vitorlán") vesszük azt az eloszlást, melynek sűrűségfüggvénye $f(x, y) = \frac{2x}{y}$. Határozza meg

- (a) az x -tengelyre,
- (b) az y -tengelyre

vetett vetületeloszlás sűrűségfüggvényének a képletét!

6. Tegyük fel, hogy az (X, Y) valószínűségi változó sűrűségfüggvénye: $f(x, y) = 4x + 2y$ ($x > 0$, $y > 0$, $x + y < 1$) Határozza meg $X + Y$ sűrűségfüggvényének a képletét!

7. Tegyük fel, hogy X és Y függetlenek, és mindketten egyenletes eloszlást követnek 0 és 1 között. Határozza meg $X + Y$ sűrűségfüggvényének a képletét!

8. Tegyük fel, hogy X és Y függetlenek, és X egyenletes eloszlást követ 0 és 3 között, Y pedig egyenletes eloszlást követ 0 és 5 között. Határozza meg $X + Y$ sűrűségfüggvényének a képletét!

9. Tegyük fel, hogy X és Y függetlenek, és mindketten λ paraméterű exponenciális eloszlást követnek. Határozza meg $X + Y$ sűrűségfüggvényének a képletét!

10. Tegyük fel, hogy X és Y függetlenek, és X λ_1 paraméterű, Y pedig λ_2 paraméterű exponenciális eloszlást követ. Határozza meg $X + Y$ sűrűségfüggvényének a képletét!

11. Tegyük fel, hogy X_1, X_2, X_3 függetlenek, és mindhárman λ paraméterű exponenciális eloszlást követnek. Határozza meg $X_1 + X_2 + X_3$ sűrűségfüggvényének a képletét!

12. Tegyük fel, hogy X_1, X_2, X_3, X_4 függetlenek, és mind λ paraméterű exponenciális eloszlást követnek. Határozza meg $X_1 + X_2 + X_3 + X_4$ sűrűségfüggvényének a képletét!

13. Tegyük fel, hogy X_1, X_2, \dots, X_n függetlenek, és mind λ paraméterű exponenciális eloszlást követnek. Határozza meg $X_1 + X_2 + \dots + X_n$ sűrűségfüggvényének a képletét!

11. Regresszió a mediánnal és a várható értékkel

11.1. Regresszió egydimenzióban

11.1.1. A medián minimál tulajdonsága

Legyen X egy folytonos valószínűségi változó, melynek sűrűségfüggvénye $f(x)$, eloszlásfüggvénye $F(x)$, és c egy konstans. A véletlen X pontnak és a c pontnak az $|X - c|$ távolsága is valószínűségi változó. Mint tudjuk, $|X - c|$ várható értéke így számolható ki:

$$E(|X - c|) = \int_{-\infty}^{\infty} |x - c| f(x) dx$$

Ennek az integrálnak az értéke függ természetesen c -től.

Állítás: Az integrál értéke akkor minimális, ha c eleget tesz az $F(c) = \frac{1}{2}$ egyenletnek. Tehát az $|X - c|$ távolság várható értéke akkor minimális, ha c a valószínűségi változó mediánja.

Bizonyítás (*Extra tananyag*): Az integrál értékét jelöljük $h(c)$ -vel. Az alábbi átalakítások nyilvánvalóak:

$$\begin{aligned} h(c) &= \int_{-\infty}^{\infty} |x - c| f(x) dx = \int_{-\infty}^c |x - c| f(x) dx + \int_c^{\infty} |x - c| f(x) dx = \\ &= \int_{-\infty}^c (c - x) f(x) dx + \int_c^{\infty} (x - c) f(x) dx = \\ &= c \int_{-\infty}^c f(x) dx - \int_{-\infty}^c x f(x) dx + \int_c^{\infty} x f(x) dx - c \int_c^{\infty} f(x) dx \end{aligned}$$

A kapott négy tag mindegyikét külön-külön deriváljuk c szerint:

$$\begin{aligned} \left(c \int_{-\infty}^c f(x) dx \right)' &= 1 \int_{-\infty}^c f(x) dx + c f(c) = F(c) + c f(c) \\ - \left(\int_{-\infty}^c x f(x) dx \right)' &= -c f(c) \\ \left(\int_c^{\infty} x f(x) dx \right)' &= -c f(c) \\ - \left(c \int_c^{\infty} f(x) dx \right)' &= -1 \cdot \int_c^{\infty} f(x) dx + c f(c) = -1 + F(c) + c f(c) \end{aligned}$$

A kapott 6 tagot összeadva megkapjuk a deriváltat:

$$h'(c) = 2F(c) - 1$$

Mivel a

$$2F(c) - 1 = 0$$

egyenlet ekvivalens a

$$F(c) = \frac{1}{2}$$

egyenlettel,

$$h'(c) = 2F(c) - 1 = 0 \quad \text{ha } c = \text{medián}$$

$$h'(c) = 2F(c) - 1 < 0 \quad \text{ha } c < \text{medián}$$

$$h'(c) = 2F(c) - 1 > 0 \quad \text{ha } c > \text{medián}$$

ami tényleg azt jelenti, hogy $h(c)$ minimuma akkor adódik, ha $c = \text{medián}$.

11.1.2. A távolság várható értékének minimalizálása

Tegyük fel, hogy X valószínűségi változóra N kísérletet végeztünk. A kísérleti eredményeket jelöljük X_1, X_2, \dots, X_N -nel. Ha minden kísérleti eredményt egy c konstanssal helyettesítünk, akkor – természetesen – mindenegyes helyettesítésnél egy eltérés, egy "hiba" adódik. Ezeknek az eltéréseknek, hibáknak a nagyságai:

$$|X_1 - c|, |X_2 - c|, \dots, |X_N - c|$$

A hibák átlaga:

$$\frac{|X_1 - c| + |X_2 - c| + \dots + |X_N - c|}{N}$$

Nagy N esetén ez az átlag közel van az $|X - c|$ valószínűségi változó várható értékéhez:

$$\begin{aligned} \frac{|X_1 - c| + |X_2 - c| + \dots + |X_N - c|}{N} &\approx \\ &\approx \mathbb{E}(|X - c|) = \int_{-\infty}^{\infty} |x - c| f(x) dx \end{aligned}$$

Az előző pontban láttuk, hogy ez a várható érték akkor minimális, ha c az X valószínűségi változó mediánja.

11.1.3. A várható érték minimál tulajdonsága

Steiner egyenlőség: Ha c egy pont a számegyenesen, és tekintünk egy valószínűségi változót vagy eloszlást, akkor a c -re vonatkozó második momentum egyenlő a variancia (azaz szórásnégyzet) plusz a c pont és várható érték közötti távolság négyzete:

$$\sum_x (x - c)^2 p(x) = \sum_x (x - \mu)^2 p(x) + (\mu - c)^2 = \sigma^2 + (\mu - c)^2 \quad (\text{diszkrét esetben})$$

$$\int_{-\infty}^{\infty} (x - c)^2 f(x) dx = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx + (\mu - c)^2 = \sigma^2 + (\mu - c)^2 \quad (\text{folytonos esetben})$$

Felhívjuk a figyelmet a Steiner egyenlőség mechanikai jelentésére: a számegyenesen vett akármilyen tömegeloszlás esetén a bármely c esetén a c pontra vonatkozó tehetetlenségi nyomaték egyenlő a súlypontra vonatkozó tehetetlenségi nyomaték plusz az a c -re vonatkozó tehetetlenségi nyomaték, amit úgy kapnánk, mintha az összes tömeget a súlypontba helyeznénk.

Bizonyítás A Steiner egyenlőség bizonyítását a folytonos esetre adjuk meg:

$$\int_{-\infty}^{\infty} (x - c)^2 f(x) dx = \int_{-\infty}^{\infty} ((x - \mu) + (\mu - c))^2 f(x) dx =$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} ((x - \mu)^2 + 2(x - \mu)(\mu - c) + (\mu - c)^2) f(x) dx = \\
&= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx + \int_{-\infty}^{\infty} 2(x - \mu)(\mu - c)f(x) dx + \int_{-\infty}^{\infty} (\mu - c)^2 f(x) dx = \\
&= \sigma^2 + 2(\mu - c) \int_{-\infty}^{\infty} (x - \mu)f(x) dx + (\mu - c)^2 \int_{-\infty}^{\infty} f(x) dx = \\
&= \sigma^2 + 2(\mu - c) 0 + (\mu - c)^2 1 = \sigma^2 + 0 + (\mu - c)^2 = \sigma^2 + (\mu - c)^2
\end{aligned}$$

Ebben az okoskodásban kihasználtuk azt, hogy

$$\int_{-\infty}^{\infty} (x - \mu)f(x) dx = 0$$

ami azért igaz, mert

$$\begin{aligned}
\int_{-\infty}^{\infty} (x - \mu)f(x) dx &= \int_{-\infty}^{\infty} xf(x) dx - \int_{-\infty}^{\infty} \mu f(x) dx = \\
&= \int_{-\infty}^{\infty} xf(x) dx - \mu \int_{-\infty}^{\infty} f(x) dx = \mu - \mu 1 = 0
\end{aligned}$$

A diszkrét eset bizonyításának a gondolatmenete hasonló, csak éppen a sűrűségfüggvények helyett súlyfüggvényt kell venni, és az integrálok helyett összegzést kell végezni.

A Steiner egyenlőség nyilvánvaló következménye a Steiner egyenlőtlenség.

Steiner egyenlőtlenség: A c -re vonatkozó második momentum nagyobb vagy egyenlő, mint a variancia (azaz szórásnégyzet), és egyenlőség csak akkor áll fenn, ha a c pont egybeesik a várható értékkel. Más szavakkal: a c -re vonatkozó második momentum akkor minimális, ha c egyenlő a várható értékkel, és a minimum értéke egyenlő a varianciával (szórásnégyzettel).

Felhívjuk a figyelmet a Steiner egyenlőtlenség mechanikai jelentésére: a számegyenesen vett akármilyen tömegeloszlás esetén a bármely c esetén a c pontra vontató tehetlenségi nyomaték nagyobb, mint a súlypontra vonatkozó tehetlenségi nyomaték.

11.1.4. A távolság négyzete várható értékének minimalizálása

A kísérleti eredmények hibáit négyzetre emelve az alábbi értékeket kapjuk:

$$(X_1 - c)^2, (X_2 - c)^2, \dots, (X_N - c)^2$$

Ezek átlaga:

$$\frac{(X_1 - c)^2 + (X_2 - c)^2 + \dots + (X_N - c)^2}{N}$$

Nagy N esetén ez az átlag közelíthető az $(X - c)^2$ várható értékével:

$$\begin{aligned}
&\frac{(X_1 - c)^2 + (X_2 - c)^2 + \dots + (X_N - c)^2}{N} \approx \\
&\approx \mathbb{E}((X - c)^2) = \int_{-\infty}^{\infty} (x - c)^2 f(x) dx
\end{aligned}$$

Az előző pontban láttuk, hogy ez a várható érték akkor minimális, ha c az X valószínűségi változó várható értéke.

11.2. Regresszió kétdimenzióban

Tegyük fel, hogy az (X, Y) kétdimenziós valószínűségi változóval van dolgunk. Képzeljük el, hogy N kísérletet végzünk. A megfigyelt értékek legyenek:

$$(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$$

Ha valaki az X és Y értékek közül csak az X -et tudja megfigyelni, akkor feladata lehet, hogy Y -ra tippeljen X megfigyelt értékéből úgy, hogy ügyesen választ egy alkalmas $y = k(x)$ függvényt, amibe X megfigyelt értékét mindig behelyettesíti, és ezzel a függvényértékkel tippel Y -ra: Y -ra a $k(X)$ értékkel tippel. Ennél a tippelésnél a véletlen X értékből tippelünk a véletlen Y értékre, ezért az óhatatlanul adódó hiba is véletlentől függ.

11.2.1. A hiba abszolút értéke várható értékének minimalizálása

Az elkövetett hibák abszolút értékei:

$$|Y_1 - k(X_1)|, |Y_2 - k(X_2)|, \dots, |Y_N - k(X_N)|$$

Ezek átlaga

$$\frac{|Y_1 - k(X_1)| + |Y_2 - k(X_2)| + \dots + |Y_N - k(X_N)|}{N} \approx$$

Nagy N esetén ez az átlag az átlag közelíthető az $|Y - k(X)|$ várható értékével:

$$\approx \frac{|Y_1 - k(X_1)| + |Y_2 - k(X_2)| + \dots + |Y_N - k(X_N)|}{N} \approx$$

$$\approx E(|Y - k(X)|) = \iint_{R^2} |y - k(x)| f(x, y) dx dy =$$

$$\iint_{R^2} |y - k(x)| f_1(x) f_{2|1}(y|x) dx dy =$$

$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |y - k(x)| f_{2|1}(y|x) dy \right) f_1(x) dx$$

Minden rögzített x esetén a

$$\int_{-\infty}^{\infty} |y - k(x)| f_{2|1}(y|x) dy$$

belső integrálban $k(x)$ egy rögzített érték. Ezért a belső integrál akkor minimális, ha $k(x)$ egyenlő a belső integrálban szerepet játszó $f_{2|1}(y|x)$ sűrűségfüggvényű feltételes eloszlás mediánjával. Ezért az optimumot nyújtó $y = k(x)$ függvényt úgy kapjuk meg, hogy az

$$F_{2|1}(y|x) = \frac{1}{2}$$

egyenletből y -t kifejezzük x segítségével.

11.2.2. A hiba négyzete várható értékének minimalizálása

Az elkövetett hibák négyzetei:

$$(Y_1 - k(X_1))^2, (Y_2 - k(X_2))^2, \dots, (Y_N - k(X_N))^2$$

Ezek átlaga

$$\frac{(Y_1 - k(X_1))^2 + (Y_2 - k(X_2))^2 + \dots + (Y_N - k(X_N))^2}{N}$$

Nagy N esetén ez az átlag az átlag közelíthető az $(Y - k(X))^2$ várható értékével:

$$\begin{aligned} & \frac{(Y_1 - k(X_1))^2 + (Y_2 - k(X_2))^2 + \dots + (Y_N - k(X_N))^2}{N} \approx \\ & \approx \mathbb{E} \left((Y - k(X))^2 \right) = \iint_{R^2} (y - k(x))^2 f(x, y) dx dy = \\ & = \iint_{R^2} (y - k(x))^2 f_1(x) f_{2|1}(y|x) dx dy = \\ & = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} (y - k(x))^2 f_{2|1}(y|x) dy \right) f_1(x) dx \end{aligned}$$

Minden rögzített x esetén a

$$\int_{-\infty}^{\infty} |y - k(x)| f_{2|1}(y|x) dy$$

belső integrálban $k(x)$ egy rögzített érték. Ezért a belső integrál akkor minimális, ha $k(x)$ egyenlő a belső integrálban szerepet játszó $f_{2|1}(y|x)$ sűrűségfüggvényű feltételes eloszlásnak a várható értékével. Ezért az optimumot nyújtó $y = k(x)$ függvényt így kapjuk meg:

$$k(x) = m_1(x) = \int_{-\infty}^{\infty} y f_{2|1}(y|x) dy$$

11.3. Egy számolás minta példa

Példa: Először 0 és 1 között választunk egy Y számot egyenletes eloszlás szerint, majd pedig 0 és Y között választunk egy X számot ugyancsak egyenletes eloszlás szerint. Számítógéppel a szimuláció így történhet:

$$X = \text{RND}_1 \cdot \text{RND}_2 \quad Y = \text{RND}_1$$

A definíciót szavakban így mondhatjuk:

$$X = \text{két független random szám szorzata}$$

$$Y = \text{a két tényező közül az első tényező}$$

Tegyük fel, hogy valakinek az a dolga, hogy X -ből Y -ra tippeljen. Két célt is kitűznek neki:

1. A hiba abszolút értékének a várható értékét kell minimalizálnia.
2. A hiba négyzetének a várható értékét kell minimalizálnia.

Megoldás:

$$f(x, y) = f_{1|2}(x|y) \cdot f_2(y)$$

Y sűrűségfüggvénye:

$$f_2(y) = 1 \quad 0 < y < 1$$

X feltételes sűrűségfüggvénye az $Y = y$ feltétel mellett:

$$f_{1|2}(x|y) = \frac{1}{y} \quad 0 < x < y$$

(X, Y) sűrűségfüggvénye:

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

X sűrűségfüggvénye:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_x^1 \frac{1}{y} dy = -\ln(x) \quad 0 < x < 1$$

Y feltételes sűrűségfüggvénye az $X = x$ feltétel mellett:

$$f_{2|1}(y|x) = \frac{\frac{1}{y}}{-\ln(x)} \quad 0 < x < y < 1$$

Y feltételes eloszlásfüggvénye az $X = x$ feltétel mellett:

$$F_{2|1}(y|x) = \int_{-\infty}^y f_{2|1}(y|x) dy = \int_x^y \frac{\frac{1}{y}}{-\ln(x)} dy = \frac{\ln(y) - \ln(x)}{-\ln(x)} \quad 0 < x < y < 1$$

Y feltételes mediánja az $X = x$ feltétel mellett az

$$\frac{\ln(y) - \ln(x)}{-\ln(x)} = \frac{1}{2}$$

egyenlet y -ra történő megoldásából jön ki:

$$\ln(y) - \ln(x) = -\frac{1}{2} \ln(x)$$

$$\ln(y) = \frac{1}{2} \ln(x)$$

$$\ln(y) = \ln(\sqrt{x})$$

$$y = \sqrt{x}$$

Ezért az abszolút hiba várható értékének minimalizálásához az optimális tippelést nyújtó függvény:

$$k(x) = \sqrt{x}$$

Y feltételes várható értéke az $X = x$ feltétel mellett:

$$E(Y|X = x) = \int_{-\infty}^{\infty} y f_{2|1}(y|x) dy = \int_x^1 y \frac{\frac{1}{y}}{-\ln(x)} dy = \int_x^1 \frac{1}{-\ln(x)} dy = \frac{1-x}{-\ln(x)} = \frac{x-1}{\ln(x)}$$

Ezért a négyzetes hiba várható értékének minimalizálásához az optimális tippelést nyújtó függvény:

$$k(x) = \frac{x-1}{\ln(x)} \quad 0 < x < 1$$

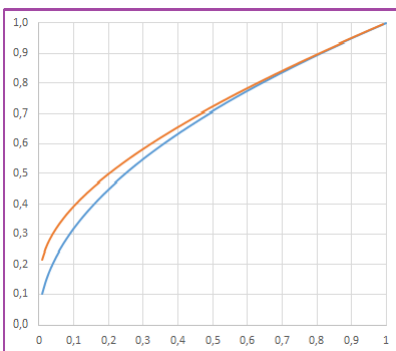
A két függvény összevetése céljából egy egyszerű táblázatot készítettünk az értékeikből:

x	\sqrt{x}	$\frac{1-x}{\ln(x)}$
0,1	0,32	0,39
0,2	0,45	0,50
0,3	0,55	0,58
0,4	0,63	0,65
0,5	0,71	0,72
0,6	0,77	0,78
0,7	0,84	0,84
0,8	0,89	0,90
0,9	0,95	0,95

Látható, hogy

$$y = \sqrt{x} \quad \text{kisebb, mint} \quad y = \frac{x-1}{\ln(x)}$$

ami azt jelenti, hogy az abszolút hiba várható értékének minimalizálásakor kisebb értékekkel kell tippelnünk, mint amikor a négyzetes hiba várható értékét minimalizáljuk. Ezt a tényt a "Regressziós függvények összevetése" feliratú ábrán is bemutatjuk.



17. ábra. Regressziós függvények összevetése: $y = \sqrt{x}$ görbe lejjebb van, mint az $y = \frac{x-1}{\ln(x)}$ görbe

11.4. Gyakorló feladatok

1. Ha a Hoppantó Bolhának azt mondjuk, hogy "HOPP!", akkor ugrik egy nagyot. Az ugrásának (méterben mért) nagysága véletlentől függ. Ezt a valószínűségi változót X -szel jelöljük. Ebben a feladatban feltesszük, hogy X egyenletes eloszlást követ 4 és 6 között. Ez – többek között – azt is jelenti, hogy

- az eseteknek kb. a negyedében 4 méternél nagyobb, de 4.5 méternél kisebbet ugrik,
- az eseteknek kb. a negyedében 4.5 méternél nagyobb, de 5 méternél kisebbet ugrik,
- az eseteknek kb. a negyedében 5 méternél nagyobb, de 5.5 méternél kisebbet ugrik,
- az eseteknek kb. a negyedében 5.5 méternél nagyobb, de 6 méternél kisebbet ugrik.

A bolha szeret a véletlennel és a pénzzel játszani. Ha olyan adakozó emberre akad, aki beígéri neki, hogy fizetni fog, és csak úgy kijelent egy c konstans számot is, akkor elkezd ugrálni, és jó sok (mondjuk kb. 100) ugrást végez. A szabályok szerint az adakozó ember minden egyes ugrás után fizet:

- ha $X > c$, akkor $(X - c)$ -szer 10 forintot,
- ha $X < c$, akkor $(c - X)$ -szer 20 forintot.

Ezért a bevállalós ember vesztesége ugrásonként:

- ha $X > c$, akkor $10 \cdot (X - c)$ forint,
- ha $X < c$, akkor $20 \cdot (c - X)$ forint.

Hogyan válassza meg a bevállalós ember a c konstans számot, ha azt akarja, hogy veszteségének ugrásonkénti várható értéke minimális legyen?

2. (Az előző feladat folytatása) Ebben a feladatban a Hoppantó Bolha ugrásának nagysága normális eloszlást követ 5 méter várható értékkel és 1 méter szórással. A kérdés ugyanaz, mint az előző feladatban.

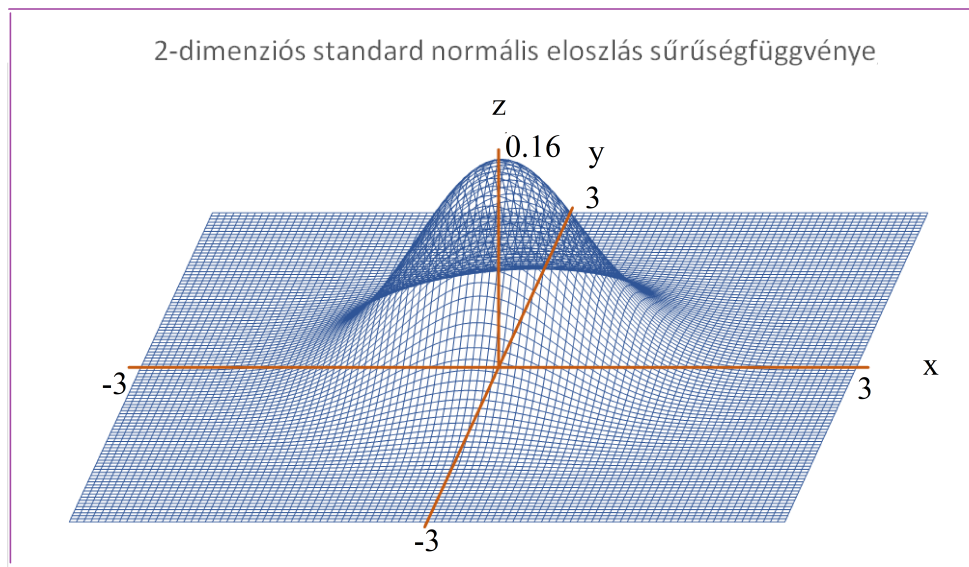
12. Normális eloszlások a síkon

12.1. Standard normális eloszlás

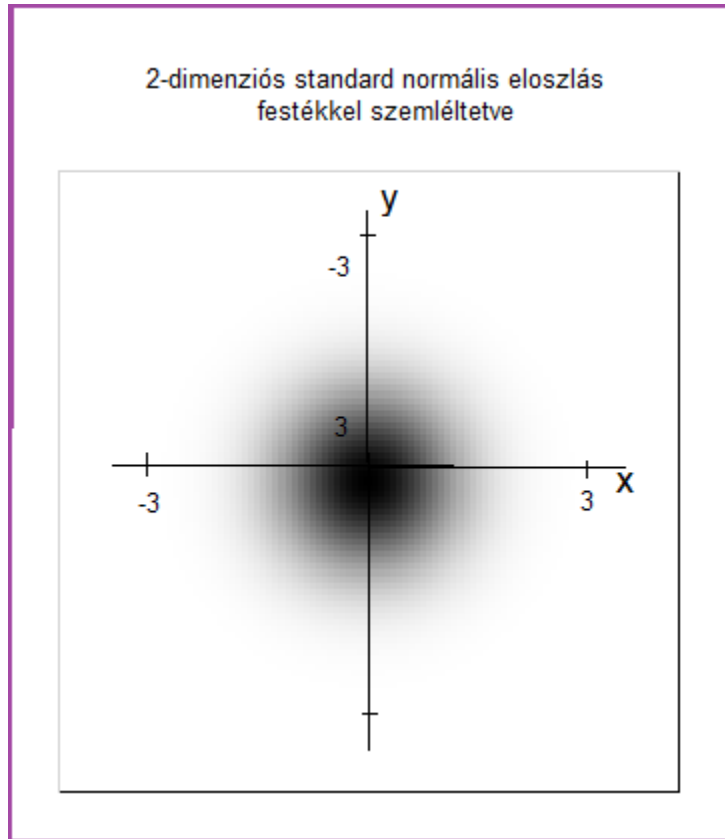
A kétdimenziós standard normális eloszlást az egész síkon az alábbi sűrűségfüggvénnyel definiáljuk:

$$f(x, y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}$$

Mivel a sűrűségfüggvény csak $(x^2 + y^2)$ -en keresztül függ x -től és y -től, az eloszlás körszimmetrikus. A függvénynek megfelelő felület egy harang alakjára emlékeztet.



18. ábra. 2-dimenziós standard normális eloszlás sűrűségfüggvénye



19. ábra. 2-dimenziós standard normális eloszlás sűrűségfüggvénye festékekkel szemléltetve

Könnyű észrevenni, hogy a kétdimenziós standard normális sűrűségfüggvény előállítható egydimenziós standard normális sűrűségfüggvények direktszorzataként is:

$$\frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$$

Ezért a kétdimenziós standard normális eloszlásnak a tengelyekre veteett vetületei egydimenziós standard normális eloszlások, és a feltételes eloszlások a függőleges és a vízszintes egyenesek mentén is egydimenziós standard normális eloszlások.

12.2. Normális eloszlás $\mu_1, \mu_2, \sigma_1, \sigma_2, r$ paraméterekkel

A kétdimenziós normális eloszlásokat az általános esetben is a sűrűségfüggvényükkel definiáljuk:

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2} \frac{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2r\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)}{(1-r^2)}$$

ahol μ_1, μ_2 valós számok, σ_1, σ_2 pozitív számok, az r pedig -1 és 1 közötti valós szám.

A sűrűségfüggvény kitevőjében szereplő tört számlálója, a

$$\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2r\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)$$

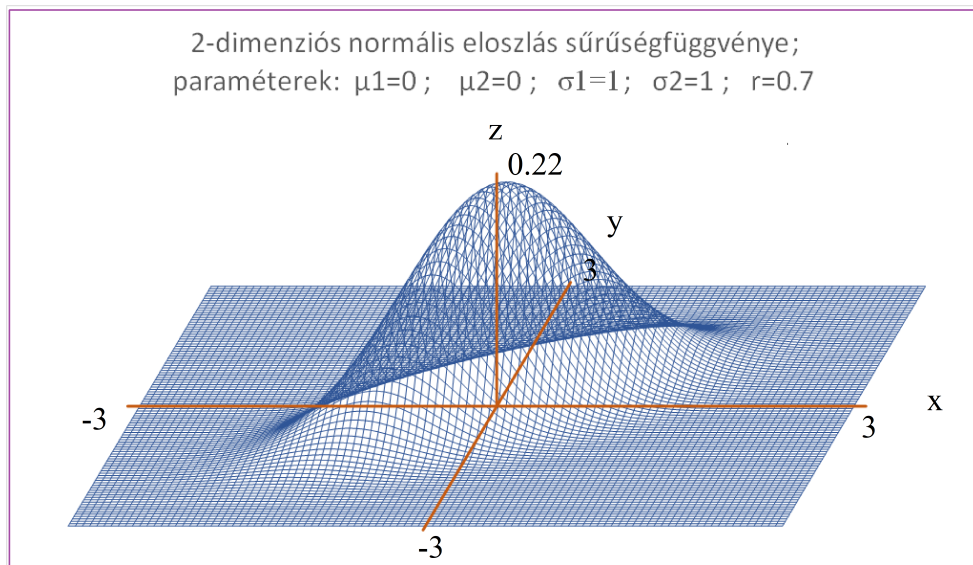
kifejezés egy kvadratikus képlet, aminek – mint ismeretes – a szintvonalai koncentrikus ellipszisek. Ezért a sűrűségfüggvénynek megfelelő felület egy deformált harang alakjára emlékeztet, mintha a harangot összenyomtuk volna úgy, hogy – felülnézetben – a kör helyett ellipszis legyen az alakja.

Megjegyzés. (*Extra tananyag*): Meg lehet mutatni, hogy

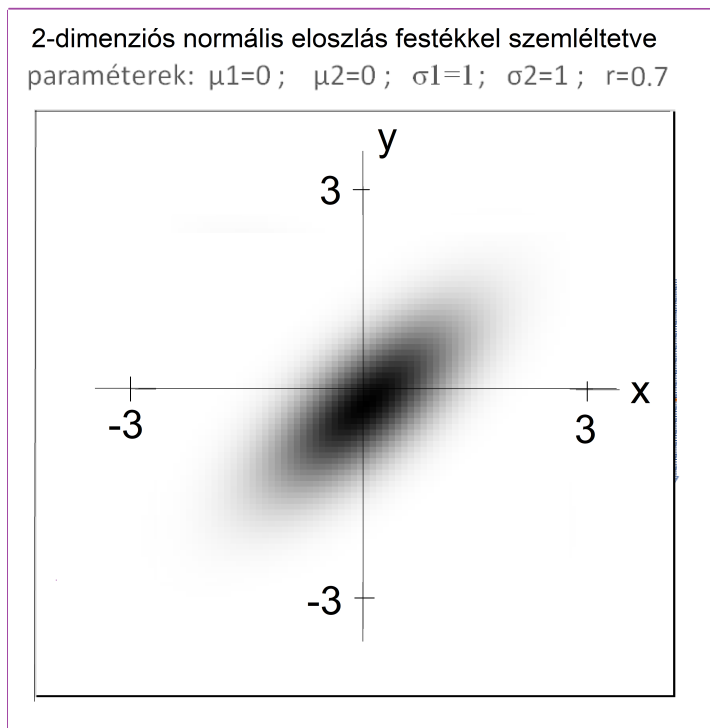
- a sűrűségfüggvény szintvonalai koncentrikus ellipszisek
- az ellipszisek közös középpontja a (μ_1, μ_2) pont
- az ellipszisek tengelyeinek irányát a (a később tárgyalandó) kovariancia mátrix sajátvektorai adják
- a tengelyek hossza pedig arányos a kovariancia mátrix sajátértékeinek négyzetgyökeivel

Normális sűrűségfüggvények felületekkel ábrázolva

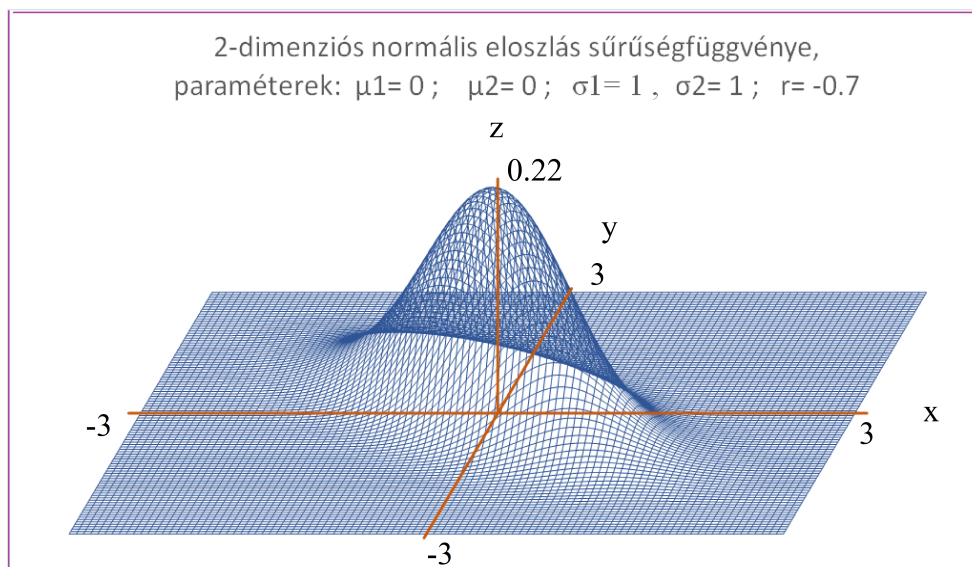
Néhány paraméter érték-rendszerre a megfelelő kétdimenziós normális sűrűségfüggvényt felülettel szemléltetjük:



20. ábra. 2-dimenziós normális eloszlás sűrűségfüggvénye; paraméterek: $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 1$, $\sigma_2 = 1$, $r = 0.7$

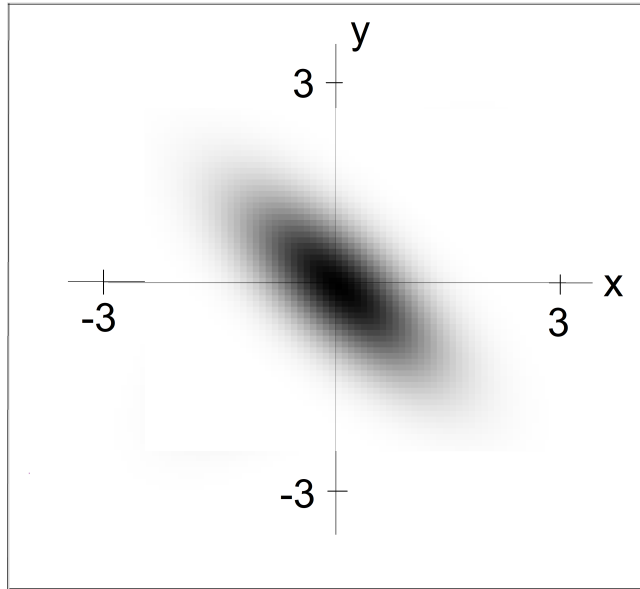


21. ábra. 2-dimenziós normális eloszlás festékkal szemléltetve; paraméterek: $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 1$, $\sigma_2 = 1$, $r = 0.7$



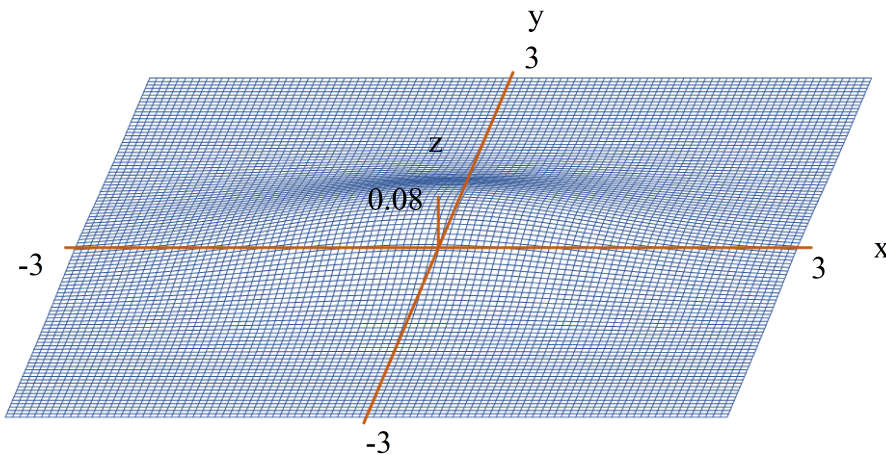
22. ábra. 2-dimenziós normális eloszlás sűrűségfüggvénye; paraméterek: $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 1$, $\sigma_2 = 1$, $r = -0.7$

2-dimenziós normális eloszlás festékekkel szemléltetve
 paraméterek: $\mu_1=0$; $\mu_2=0$; $\sigma_1=1$; $\sigma_2=1$; $r=-0.7$

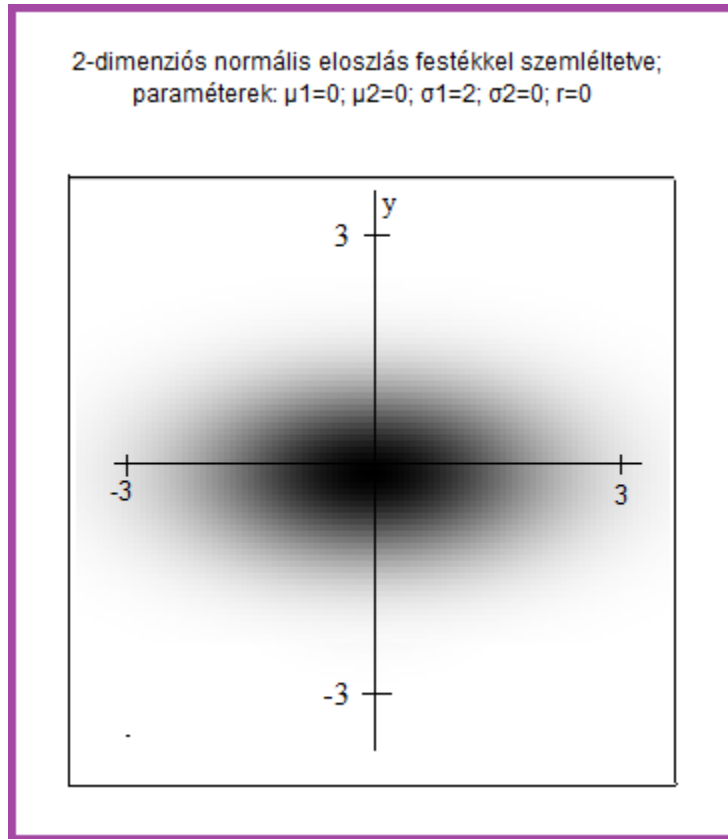


23. ábra. 2-dimenziós normális eloszlás festékekkel szemléltetve; paraméterek: $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 1$, $\sigma_2 = 1$, $r = -0.7$

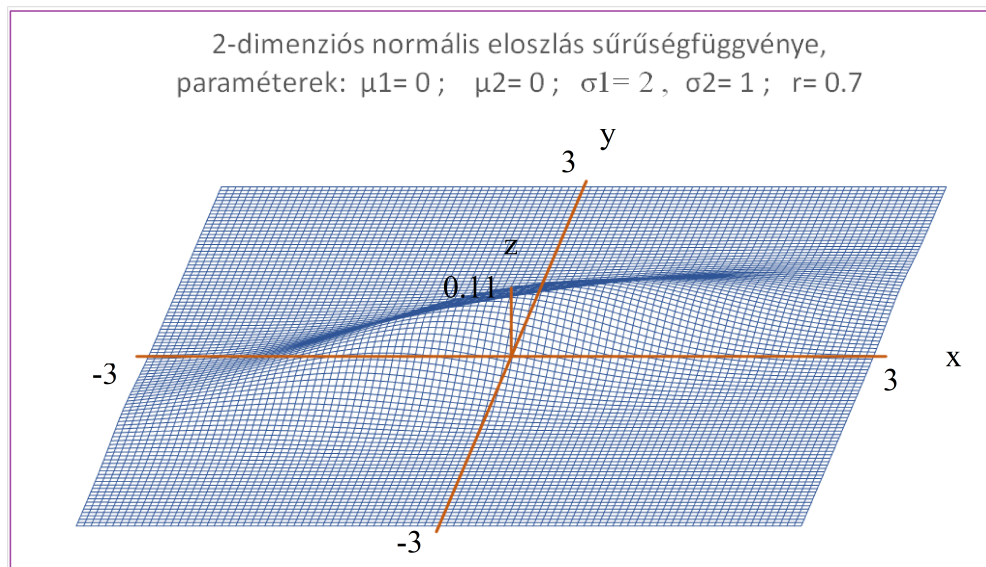
2-dimenziós normális eloszlás sűrűségfüggvénye,
 paraméterek: $\mu_1=0$; $\mu_2=0$; $\sigma_1=2$, $\sigma_2=1$; $r=0$



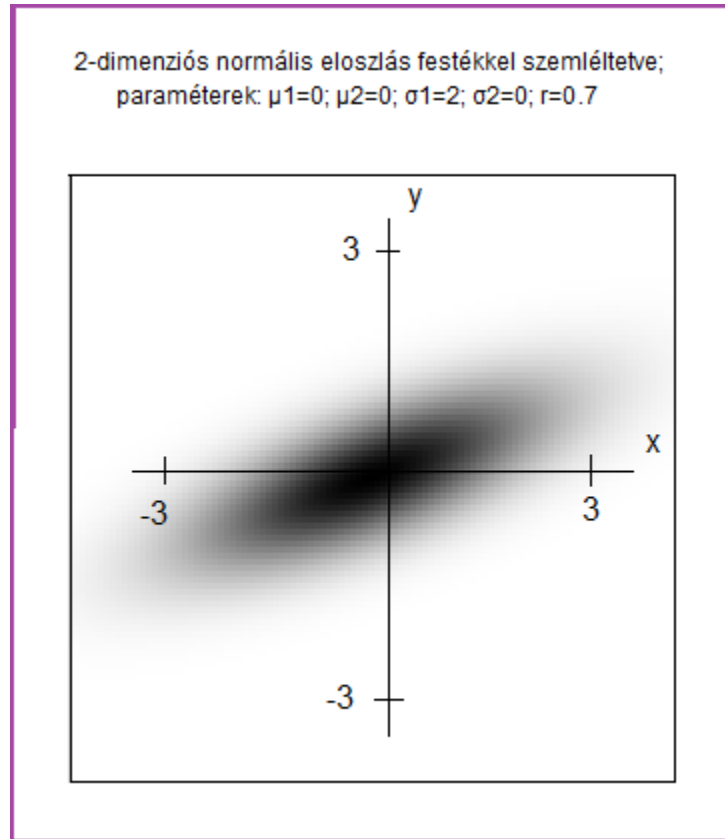
24. ábra. 2-dimenziós normális eloszlás sűrűségfüggvénye; paraméterek: $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 2$, $\sigma_2 = 1$, $r = 0$



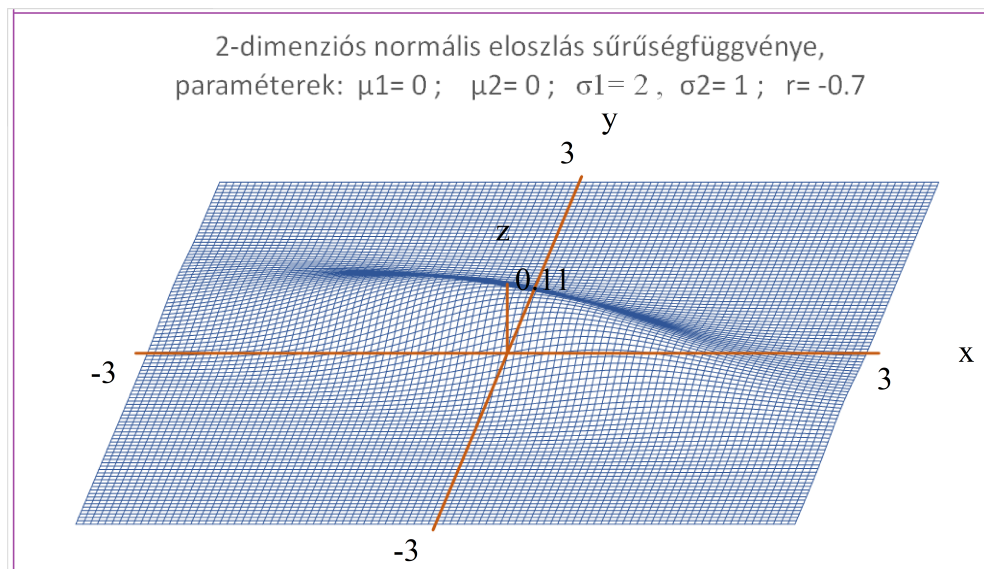
25. ábra. 2-dimenziós normális eloszlás festékekkel szemléltetve; paraméterek: $\mu_1 = 0, \mu_2 = 0, \sigma_1 = 2, \sigma_2 = 1, r = 0$



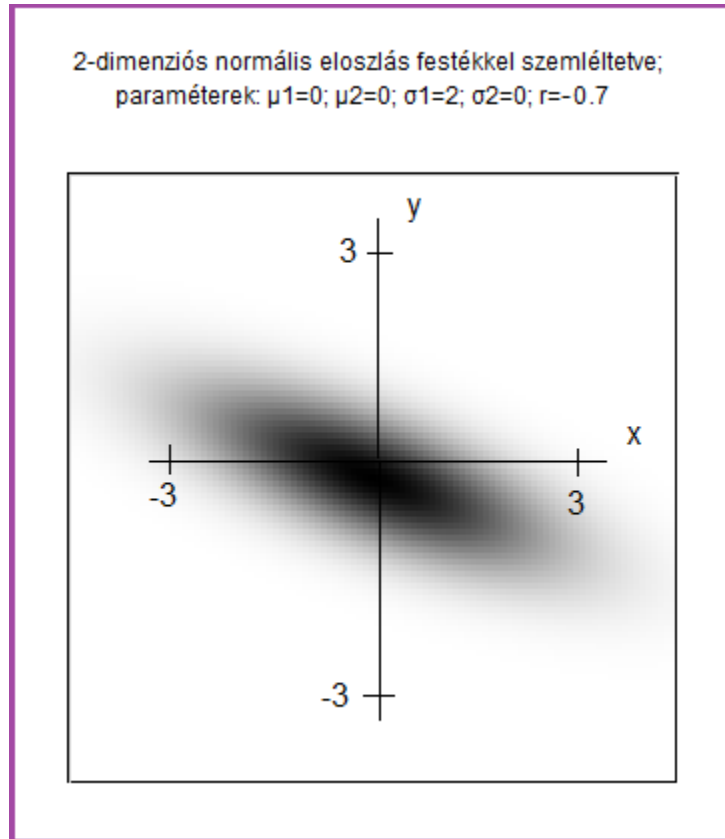
26. ábra. 2-dimenziós normális eloszlás sűrűségfüggvénye; paraméterek: $\mu_1 = 0, \mu_2 = 0, \sigma_1 = 2, \sigma_2 = 1, r = 0.7$



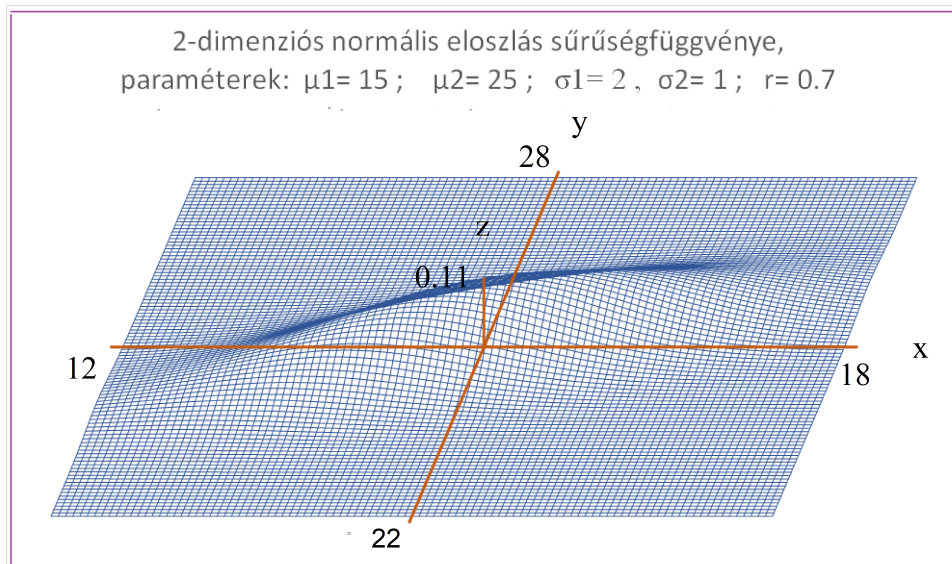
27. ábra. 2-dimenziós normális eloszlás festékekkel szemléltetve; paraméterek: $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 2$, $\sigma_2 = 1$, $r = 0.7$



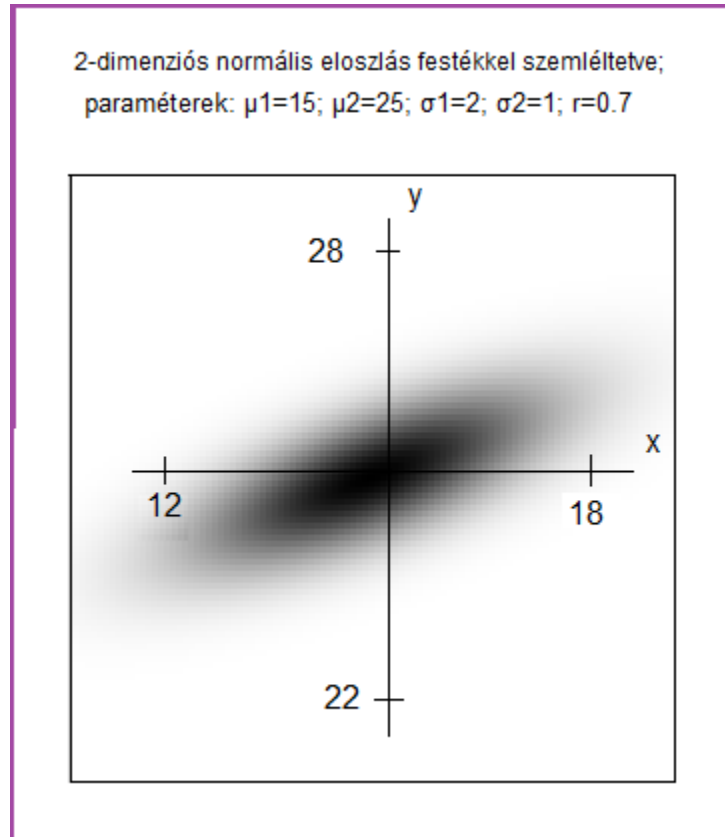
28. ábra. 2-dimenziós normális eloszlás sűrűségfüggvénye; paraméterek: $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 2$, $\sigma_2 = 1$, $r = -0.7$



29. ábra. 2-dimenziós normális eloszlás festékekkel szemléltetve; paraméterek: $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 2$, $\sigma_2 = 1$, $r = -0.7$



30. ábra. 2-dimenziós normális eloszlás sűrűségfüggvénye; paraméterek: $\mu_1 = 15$, $\mu_2 = 25$, $\sigma_1 = 2$, $\sigma_2 = 1$, $r = 0.7$



31. ábra. 2-dimenziós normális eloszlás festékekkel szemléltetve; paraméterek: $\mu_1 = 15$, $\mu_2 = 25$, $\sigma_1 = 2$, $\sigma_2 = 1$, $r = 0.7$

Mikor használunk kétdimenziós normális eloszlásokat?

1. Ha egy kétdimenziós valószínűségi változó sok, független, *kis szórású* koordinátákkal rendelkező kétdimenziós valószínűségi változó összege, akkor ez a kétdimenziós valószínűségi változó normális eloszlásúnak vehető valamilyen $\mu_1, \mu_2, \sigma_1, \sigma_2, r$ paraméterekkel.

Ezt a kijelentést *kétdimenziós centrális határeloszlás tételek* támasztják alá, melyek tárgyalása meghaladja lehetőségeinket. Ha egy valószínűségi változó kis értékeket vesz fel, akkor szórása is szükségképpen kicsi. Ezért az előző kijelentésnek következménye az alábbi:

2. Ha egy kétdimenziós valószínűségi változó sok, független, *kis értékű* koordinátákkal rendelkező kétdimenziós valószínűségi változó összege, akkor ez a kétdimenziós valószínűségi változó normális eloszlásúnak vehető valamilyen $\mu_1, \mu_2, \sigma_1, \sigma_2, r$ paraméterekkel, hiszen kis értékek esetén a szórások is kicsik.
3. Sokszor csak kényelmi okokból használjuk a kétdimenziós normális eloszlásokat, mert úgy látjuk, úgy gondoljuk, hogy a normális eloszlás jó közelítése az igazi eloszlásnak, és kétdimenziós problémák esetén normális eloszlásokkal nagyon kényelmesen lehet dolgozni. Ilyen alkalmazásra hamarosan mutatunk példát.
4. A standard normális eloszlás jelentősége abban rejlik, hogy
 - lineáris transzformációk segítségével tetszőleges normális eloszlás kapcsolatban hozható a standard normálissal, és
 - a standard normális eloszlással – speciális tulajdonságai miatt – az elméleti és a gyakorlati számítások könnyebben végezhetőek el.

12.3. Vetület eloszlások

Az x -tengelyre vetett vetület sűrűségfüggvényét a tengely tetszőleges x pontjában az alábbi integrál adja:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2} \frac{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2r\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)}{1-r^2}} dy$$

Ennek a – kissé bonyolult – integrálnak a kiszámolása a $z = \frac{y-\mu_2}{\sigma_2}$ helyettesítéssel elvégezhető. Az eredmény:

$$f_1(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

Hasonlóképp meghatározható a függőleges tengelyre vetett vetület eloszlás sűrűségfüggvénye is:

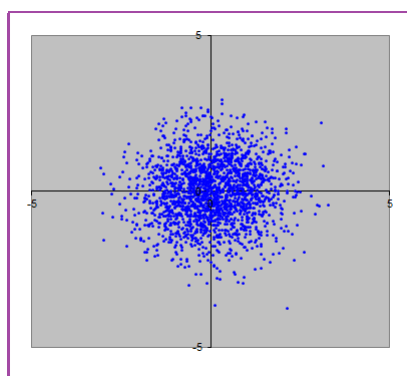
$$f_2(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

Látjuk, hogy a tengelyekre vetett vetületek normális eloszlások μ_1, σ_1 , illetve μ_2, σ_2 paraméterekkel.

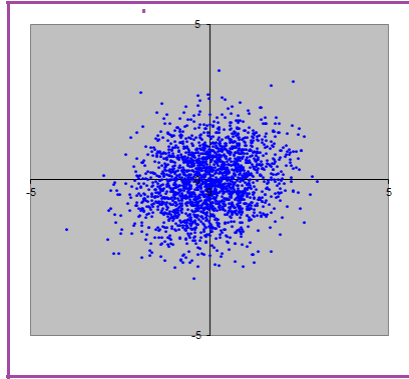
Ezzel a $\mu_1, \mu_2, \sigma_1, \sigma_2$ paraméterek jelentését megtaláltuk. μ_1 és μ_2 a vetületek várható értékét jelenti, σ_1, σ_2 pedig a vetületek szórását.

12.4. Korrelációs együttható

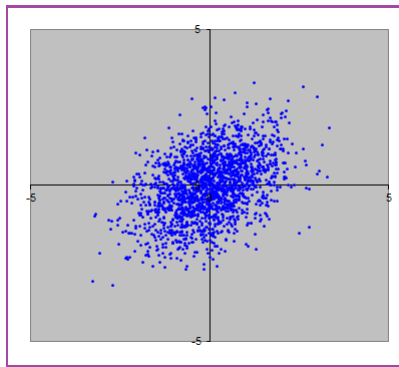
Ha a korrelációs együtthatót, melynek fogalmát később tanuljuk, a definíciójának megfelelő módon kiszámoljuk, akkor értékére az r paraméter adódik. Tehát az r paraméter "hivatalos nevét" is tudjuk már: ő a korrelációs együttható. A korrelációs együttható szemléletes jelentését – kétdimenziós adatrendszerekkel (pontfelhőkkel) kapcsolatban – a mellékelt ábrákról lehet leolvasni. Az ábra sorozatokat egyszerre kell áttekinteni úgy, mintha egy moziban lepörgetnénk előtűnk a képeket



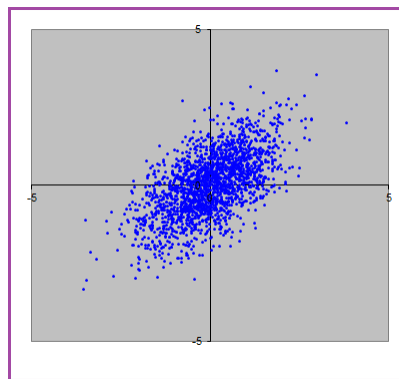
32. ábra. Korrelációs együttható 0.00



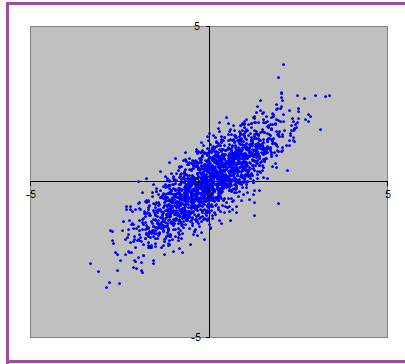
33. ábra. *Korrelációs együttható 0.20*



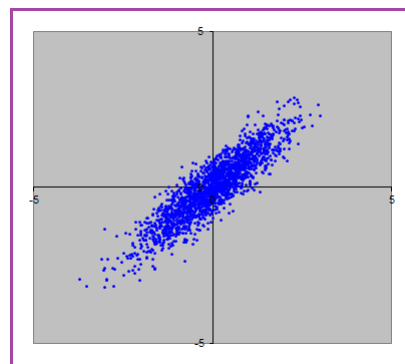
34. ábra. *Korrelációs együttható 0.40*



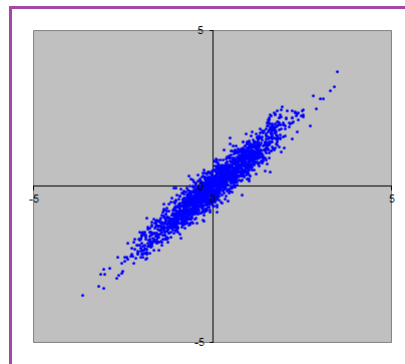
35. ábra. *Korrelációs együttható 0.60*



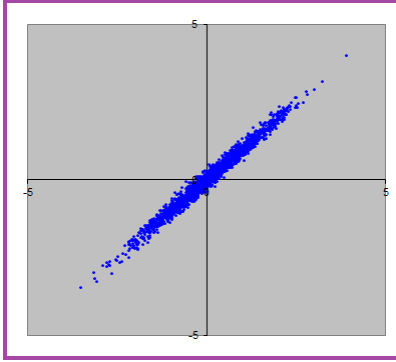
36. ábra. Korrelációs együttható 0.80



37. ábra. Korrelációs együttható 0.90



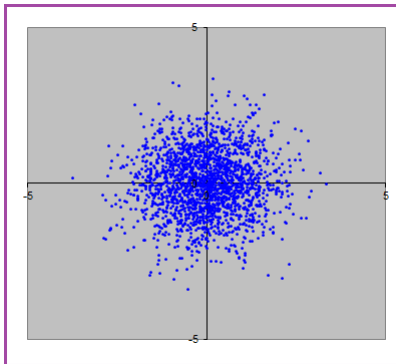
38. ábra. Korrelációs együttható 0.95



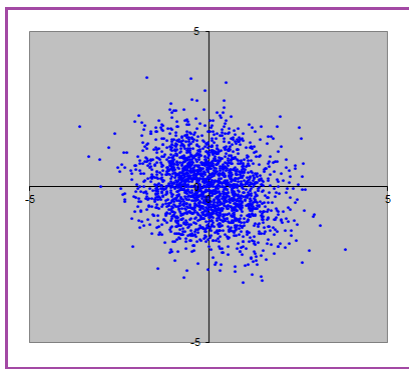
39. ábra. Korrelációs együttható 0.99



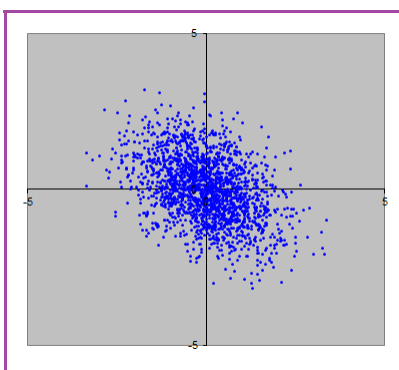
40. ábra. Korrelációs együttható 1.00



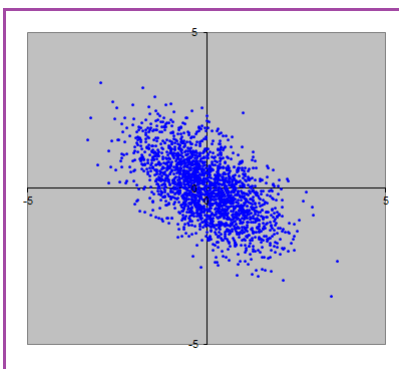
41. ábra. Korrelációs együttható: 0.00



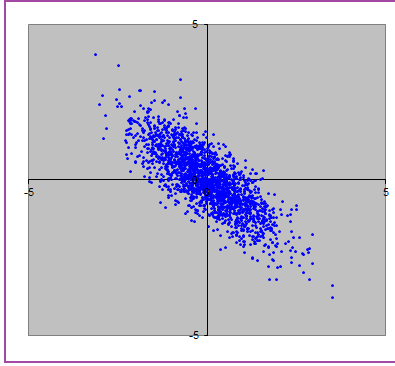
42. ábra. *Korrelációs együttható: -0.20*



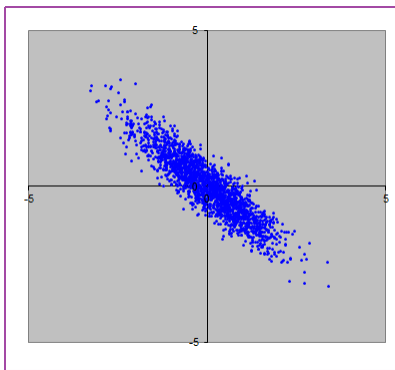
43. ábra. *Korrelációs együttható: -0.40*



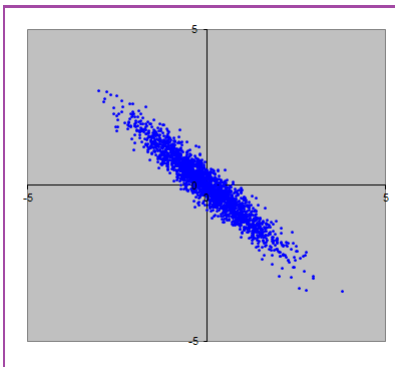
44. ábra. *Korrelációs együttható: -0.60*



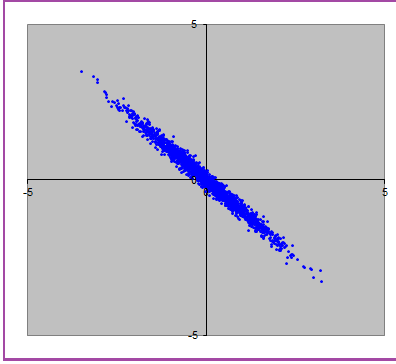
45. ábra. *Korrelációs együttható: -0.80*



46. ábra. *Korrelációs együttható: -0.90*



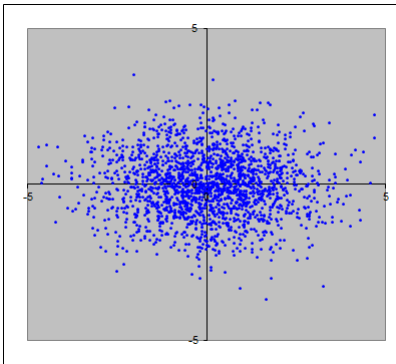
47. ábra. *Korrelációs együttható: -0.95*



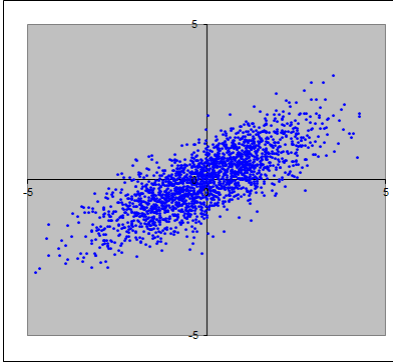
48. ábra. *Korrelációs együttható: -0.99*



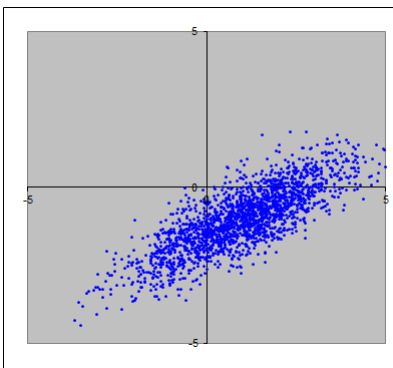
49. ábra. *Korrelációs együttható: -1.00*



50. ábra. *Szórások: 1.5 és 1.0 ; Korrelációs együttható: 0.00*



51. ábra. Szórások: 1.5 és 1.0 ; Korrelációs együttható: 0.80



52. ábra. Szórások: 1.5 és 1.0 ; Korrelációs együttható: 0.80 ; Várható értékek: 1.0 és -1.0

12.5. Feltételes sűrűségfüggvény

A feltételes sűrűségfüggvényt az $X = x$ feltétel mellett a megfelelő függőleges egyenes mentén az alábbi hányados adja:

$$f_{2|1}(y|x) = \frac{f(x, y)}{f_1(x)} = \frac{\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2} \frac{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2r\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)}{(1-r^2)}}{\frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}}$$

Egyszerűsítésekkel az alábbi eredmény adódik:

$$f_{2|1}(y|x) = \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2} \left(\frac{y - \left(\mu_2 + r\frac{\sigma_2}{\sigma_1}(x - \mu_1)\right)}{\sigma_2\sqrt{1-r^2}} \right)^2}$$

A kapott képletből kiolvashatjuk, hogy az $X = x$ feltétel mellett az Y feltételes eloszlása normális eloszlás, melynek várható értéke

$$\mu_2 + r\frac{\sigma_2}{\sigma_1}(x - \mu_1)$$

és szórása

$$\sigma_2\sqrt{1-r^2}$$

A feltételes sűrűségfüggvény az $Y = y$ feltétel mellett a megfelelő vízszintes egyenes mentén hasonló:

$$f_{1|2}(x|y) = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-r^2}} e^{-\frac{1}{2}\left(\frac{x - (\mu_1 + r\frac{\sigma_1}{\sigma_2}(y - \mu_2))}{\sigma_1\sqrt{1-r^2}}\right)^2}$$

A képletből látjuk, hogy az $Y = y$ feltétel mellett az X feltételes eloszlása normális eloszlás, melynek várható értéke

$$\mu_1 + r\frac{\sigma_1}{\sigma_2}(y - \mu_2)$$

szórása:

$$\sigma_1\sqrt{1-r^2}$$

12.6. Feltételes medián és várható érték

A fentiek szerint tehát az $X = x$ feltétel mellett az Y várható értéke és mediánja:

$$E(Y | X = x) = \mu_2 + r\frac{\sigma_2}{\sigma_1}(x - \mu_1)$$

Hasonlóképpen az $Y = y$ feltétel mellett az X várható értéke és mediánja:

$$E(X | Y = y) = \mu_1 + r\frac{\sigma_1}{\sigma_2}(y - \mu_2)$$

Látjuk, hogy a feltételes várható érték és medián lineárisan függ a feltételtől. A lineáris függésben a meredekség $r\frac{\sigma_2}{\sigma_1}$; illetve $r\frac{\sigma_1}{\sigma_2}$. Az

$$y = \mu_2 + r\frac{\sigma_2}{\sigma_1}(x - \mu_1)$$

egyenletű egyenest **első regressziós egyenesnek** nevezzük. Hasonlóképpen az

$$x = \mu_1 + r\frac{\sigma_1}{\sigma_2}(y - \mu_2)$$

egyenletű egyenest **második regressziós egyenesnek** nevezzük.

12.7. Feltételes szórás

Az $X = x$ feltétel mellett az Y szórása is kiolvasható a feltételes sűrűségfüggvény képletéből:

$$SD(Y | X = x) = \sigma_2\sqrt{1-r^2}$$

Hasonlóképpen az $Y = y$ feltétel mellett az X szórása:

$$SD(X | Y = y) = \sigma_1\sqrt{1-r^2}$$

Látjuk, hogy a feltételes szórás nem függ a feltételtől, és kisebb, mint a feltétel nélküli szórás.

$$SD(Y | X = x) = \sigma_2\sqrt{1-r^2} < \sigma_2$$

illetve

$$SD(X | Y = y) = \sigma_1\sqrt{1-r^2} < \sigma_1$$

12.8. Példa: testmagasság és testsúly

Véletlenszerűen választunk egy felnőtt férfit, és megmérjük testmagasságát centiméterekben és testsúlyát (pontosabban: testtömegét) kilogramokban:

$$X = \text{testmagasság} \quad Y = \text{testsúly}$$

Tegyük fel, hogy

- a testmagasság várható értéke 180 (cm)
- a testmagasság szórása 15 (cm)
- a testsúly várható értéke 80 (kg)
- a testsúly szórása 10 (kg)
- a korrelációs együttható pedig 0.6

Ha az (X, Y) kétdimenziós valószínűségi változóval kapcsolatos számításokban kétdimenziós normális eloszlással dolgozunk, az alábbi numerikus eredményeket kapjuk.

1. Annak a valószínűsége, hogy egy véletlenszerűen választott felnőtt férfi 90 kg-nál súlyosabb:

$$P(Y > 90) = 1 - \text{NORM.DIST}(90; 80; 10; \text{TRUE}) = 0.16$$

2. Tekintsünk egy tetszőleges x testmagasság értéket! Ha a véletlenszerűen választott felnőtt férfiak közül csak azokat tekintjük, akiknek a testmagassága körülbelül x (cm), akkor az ő testsúlyaik szórása nem 10 (kg), hanem csak:

$$\text{SD}(Y | X = x) = 10 \sqrt{1 - 0.6^2} = 8$$

Kiemeljük, hogy ennek a feltételes szórásnak az értéke nem függ x -től.

3. Ha a véletlenszerűen választott felnőtt férfiak testmagassága körülbelül egyforma, mondjuk x , akkor a testsúly várható értéke – a feltételezett x függvényében – így alakul:

- $x = 170$ esetén Y feltételes várható értéke:

$$E(Y | X = 170) = 80 + 0.6 \frac{10}{15}(170 - 180) = 76$$

Tehát a körülbelül 170 cm magas felnőtt férfiak átlagos testsúlya körülbelül 76 kg.

- $x = 180$ esetén Y feltételes várható értéke:

$$E(Y | X = 180) = 80 + 0.6 \frac{10}{15}(180 - 180) = 80$$

Tehát a körülbelül 180 cm magas felnőtt férfiak átlagos testsúlya körülbelül 80 kg.

- $x = 190$ esetén Y feltételes várható értéke:

$$E(Y | X = 190) = 80 + 0.6 \frac{10}{15}(190 - 180) = 84$$

Tehát a körülbelül 190 cm magas felnőtt férfiak átlagos testsúlya körülbelül 84 kg.

- $x = 200$ esetén Y feltételes várható értéke:

$$E(Y | X = 200) = 80 + 0.6 \frac{10}{15}(200 - 180) = 88$$

Tehát a körülbelül 200 cm magas felnőtt férfiak átlagos testsúlya körülbelül 88 kg.

A számadatokból kiolvashatjuk, hogy 10 cm-rel magasabb felnőtt férfiak átlagosan 4 kg-mal nehezebbek, mint a 10 cm-rel alacsonyabbak. Tehát 1 cm magasságnövekedés átlagosan 0.4 kg testsúly növekedést eredményez. Vegyük észre, hogy a testsúly szórása és a testmagasság szórásának aránya a korrelációs együtthatóval szorozva adja ki ezt az átlagos testsúly növekedést:

$$\frac{10}{15} \cdot 0.6 = 0.4$$

4. Annak a valószínűsége, hogy egy felnőtt férfi 90 kg-nál súlyosabb – a feltételezett testmagasság függvényében – így alakul:

- $x = 170$ esetén a feltételes valószínűség értéke:

$$P(Y > 90 | X = 170) = 1 - \text{NORM.DIST}(90; 76; 8; \text{TRUE}) = 0.04$$

Tehát a körülbelül 170 cm magas felnőtt férfiaknak körülbelül 4 %-a súlyosabb 90 kg-nál.

- $x = 180$ esetén a feltételes valószínűség értéke:

$$P(Y > 90 | X = 180) = 1 - \text{NORM.DIST}(90; 80; 8; \text{TRUE}) = 0.11$$

Tehát a körülbelül 180 cm magas felnőtt férfiaknak körülbelül 11 %-a súlyosabb 90 kg-nál.

- $x = 190$ esetén a feltételes valószínűség értéke:

$$P(Y > 90 | X = 190) = 1 - \text{NORM.DIST}(90; 84; 8; \text{TRUE}) = 0.23$$

Tehát a körülbelül 190 cm magas felnőtt férfiaknak körülbelül 23 %-a súlyosabb 90 kg-nál.

- $x = 200$ esetén a feltételes valószínűség értéke:

$$P(Y > 90 | X = 200) = 1 - \text{NORM.DIST}(90; 88; 8; \text{TRUE}) = 0.40$$

Tehát a körülbelül 200 cm magas felnőtt férfiaknak körülbelül 40 %-a súlyosabb 90 kg-nál.

5. Annak a valószínűsége, hogy egy véletlenszerűen választott felnőtt férfi 200 cm-nél magasabb:

$$P(X > 200) = 1 - \text{NORM.DIST}(200; 180; 15; \text{TRUE}) = 0.09$$

6. Tekintsünk egy tetszőleges y testsúly értéket! Ha a véletlenszerűen választott felnőtt férfiak közül csak azokat tekintjük, akiknek a testsúlya körülbelül y (kg), akkor az ő testmagasságaik szórása nem 15 (cm), hanem csak:

$$SD(X | Y = y) = 15 \sqrt{1 - 0.6^2} = 12$$

Kiemeljük, hogy ennek a feltételes szórásnak az értéke nem függ y -tól.

7. Ha a véletlenszerűen választott felnőtt férfiak testsúlya körülbelül egyforma, mondjuk y (kg), akkor a testmagasság várható értéke – a feltételezett y függvényében – így alakul:

- $y = 76$ esetén X feltételes várható értéke:

$$E(X | Y = 76) = 180 + 0.6 \frac{15}{10} (76 - 80) = 176.4$$

Tehát a körülbelül 76 kg testsúlyú felnőtt férfiak átlagos testmagassága körülbelül 176.4 cm.

- $y = 80$ esetén X feltételes várható értéke:

$$E(X | Y = 80) = 180 + 0.6 \frac{15}{10} (80 - 80) = 180.0$$

Tehát a körülbelül 80 kg testsúlyú felnőtt férfiak átlagos testmagassága körülbelül 180.0 cm.

- $y = 84$ esetén X feltételes várható értéke:

$$E(X | Y = 84) = 180 + 0.6 \frac{15}{10} (84 - 80) = 183.6$$

Tehát a körülbelül 84 kg testsúlyú felnőtt férfiak átlagos testmagassága körülbelül 183.6 cm.

- $y = 88$ esetén X feltételes várható értéke:

$$E(X | Y = 88) = 180 + 0.6 \frac{15}{10} (88 - 80) = 187.2$$

Tehát a körülbelül 88 kg testsúlyú felnőtt férfiak átlagos testmagassága körülbelül 187.2 cm.

A számadatokból kiolvashatjuk, hogy 4 kg-mal nehezebb felnőtt férfiak átlagosan 3.6 cm-rel magasabbak, mint a 4 kg-mal alacsonyabbak. Tehát 1 kg súlynövekedés átlagosan 0.9 cm testmagasság növekedést eredményez. Vegyük észre, hogy a testmagasság szórása és a testsúly szórásának aránya a korrelációs együtthatóval szorozva adja ki ezt az átlagos testsúly növekedést:

$$\frac{15}{10} \cdot 0.6 = 0.9$$

8. Annak a valószínűsége, hogy egy felnőtt férfi 200 cm-nél magasabb – a feltételezett testsúly függvényében – így alakul:

- $y = 76$ esetén a feltételes valószínűség értéke:

$$P(X > 200 | Y = 76) = 1 - \text{NORM.DIST}(200 ; 176.4 ; 12 ; \text{TRUE}) = 0.02$$

Tehát a körülbelül 76 kg testsúlyú felnőtt férfiaknak körülbelül 2 %-a súlyosabb magasabb 200 cm-nél.

- $y = 80$ esetén a feltételes valószínűség értéke:

$$P(X > 200 | Y = 80) = 1 - \text{NORM.DIST}(200; 180.0; 12; \text{TRUE}) = 0.05$$

Tehát a körülbelül 80 kg testsúlyú felnőtt férfiaknak körülbelül 5 %-a súlyosabb magasabb 200 cm-nél.

- $y = 84$ esetén a feltételes valószínűség értéke:

$$P(X > 200 | Y = 84) = 1 - \text{NORM.DIST}(200; 183.6; 12; \text{TRUE}) = 0.09$$

Tehát a körülbelül 84 kg testsúlyú felnőtt férfiaknak körülbelül 9 %-a súlyosabb magasabb 200 cm-nél.

- $y = 88$ esetén a feltételes valószínűség értéke:

$$P(X > 200 | Y = 88) = 1 - \text{NORM.DIST}(200; 187.2; 12; \text{TRUE}) = 0.14$$

Tehát a körülbelül 88 kg testsúlyú felnőtt férfiaknak körülbelül 14 %-a súlyosabb magasabb 200 cm-nél.

Kiemelünk két tényt a fentiek közül:

- $x = 190$ esetén Y feltételes várható értéke:

$$E(Y | X = 190) = 80 + 0.6 \frac{10}{15} (190 - 180) = 84$$

vagyis a körülbelül 190 cm magas felnőtt férfiak átlagos testsúlya körülbelül 84 kg.

- $y = 84$ esetén X feltételes várható értéke:

$$E(X | Y = 84) = 180 + 0.6 \frac{15}{10} (84 - 80) = 183.6$$

vagyis a körülbelül 84 kg testsúlyú felnőtt férfiak átlagos testmagassága körülbelül 183.6 cm.

Lehet, hogy furcsának tűnik, de mégis ez az igazság: a 190 cm magas férfiak átlagos testsúlya 84 kg, de a 84 kg testsúlyú férfiak átlagos testmagassága nem 190 cm, hanem csak 183.6 cm.

12.9. Példa: A műszer hibáját korigáljuk

Tegyük fel, hogy egy áramkörben a feszültség névleges értéke 230 Volt, de az igazság az, hogy a tényleges feszültség ingadozik, egyszer több, máskor kevesebb, mint 230 Volt. Tegyük fel, hogy a tényleges feszültség normális eloszlású valószínűségi változó 230 várható értékkel és 10 szórással.

Ha valakinek tökéletes műszere van, akkor a tényleges feszültséget pontosan ki tudja mérni. De ha a műszer tévedhet, és a leolvasott feszültség a tényleges feszültségnél több is kevesebb is lehet, akkor felmerül a kérdés: hogyan következtessünk a leolvasott feszültségből a tényleges feszültségre?

Tegyük fel, hogy a műszer hibája – a tényleges feszültségtől független, és normális eloszlást követ 0 Volt várható értékkel és 5 Volt szórással. Ha a leolvasott értéket a tényleges feszültség és a műszer hibájának az összege, akkor így okoskodhatunk.

Legyen X a tényleges feszültség, H a műszer hibája, Y a leolvasott érték: $Y = X + H$. A feltételek szerint

- X normális eloszlású 230 várható értékkel és 10 szórással,

• H független X -től, és normális eloszlású 0 várható értékkel és 5 szórással, amiből következik, hogy

- Y normális eloszlású 230 várható értékkel és $\sqrt{10^2 + 5^2} = \sqrt{125}$ szórással.

Y feltételes szórása X adott értéke esetén a hiba szórásával, vagyis 5 -tel egyenlő. Ezért az

$$\text{SD}(Y | X = x) = \sigma_2 \sqrt{1 - r^2}$$

$$5 = \sqrt{125} \sqrt{1 - r^2}$$

egyenletből a korrelációs együtthatóra az alábbi érték jön ki:

$$r = \sqrt{\frac{4}{5}}$$

Ezek alapján – a tanultak szerint – a legjobb tipp Y -ből X -re az alábbi képlettel adódik:

$$x = \mu_1 + r \frac{\sigma_1}{\sigma_2} (y - \mu_2)$$

$$x = 230 + \sqrt{\frac{4}{5}} \frac{10}{\sqrt{125}} (y - 230)$$

$$x = 230 + \frac{20}{25} (y - 230)$$

$$x = 230 + 0.8 (y - 230)$$

Ezek szerint, ha valaki rendszeresen következtet a leolvasott feszültségből a tényleges feszültségre, és a hiba négyzetének (vagy abszolút értékének) a várható értékét igyekszik minimalizálni, akkor – például – 240 Volt feszültség leolvassása esetén a tényleges feszültséget 238 Voltnak kell vennie.

12.10. Gyakorló feladatok

- 1.
- 2.

13. Kovariancia adatrendszerekre

13.1. Kétdimenziós adatrendszer fogalma

Számpároknak egy sorozatát *kétdimenziós adatrendszernek* hívjuk. Egy számpár egy vektort is jelent:

- ha a számpár elemeit egymás mellé, azaz "egy sorba" írjuk, akkor a *sorvektort* kapunk
- ha a számpár elemeit egymás alá, azaz "egy oszlopba" írjuk, akkor a *oszlopvektort* kapunk

Izlés dolga, és igazából egyremegy, hogy sor- vagy oszlopvektorokkal dolgozunk, de választani kell a két lehetőség közül, és a két lehetőséget nem szabad összekeverni. Mi itt ebben a könyvben oszlopvektorokat fogunk használni.

Vektorok összege, átlaga jól ismert fogalmak, ezért egy kétdimenziós adatrendszer *átlagának* fogalma sem szorul sok magyarázatra. Nyilvánvaló, hogy egy kétdimenziós adatrendszer átlagának koordinátái a koordináták külön-külön vett átlagaival egyenlő.

1. Példa: Kétdimenziós adatrendszer 5 elemből. Az alábbi adatrendszer 5 számpárból áll:

x	54	55	44	44	60
y	85	83	92	65	75

Vegyük most minden (x, y) adat esetén az $x \cdot y$ szorzatot, és tekintsük ezeknek a szorzatoknak az átlagát. Ezt az átlagot a kétdimenziós adatrendszer *vegyes (második) momentumának* nevezünk:

						átlag
$x \cdot y$	4 590	4 565	4 048	2 860	4 500	4 112.6
						vegyes momentum

A vegyes momentum fogalma akkor jut szerephez, amikor egy kétdimenziós adatrendszerből – például – az

x	54	55	44	44	60
y	85	83	92	65	75

adatrendszerből a megfelelő x és y értékek összeadásával egy új egydimenziós adatrendszert állítunk elő:

$x + y$	139	138	136	109	135
---------	-----	-----	-----	-----	-----

Ennek az egydimenziós adatrendszernek a második momentumát a szokásos módon kiszámíthatjuk. A második momentum értéke 17 393.4.

A középiskolai tanulmányainkból jól ismert $(a + b)^2 = a^2 + b^2 + 2ab$ azonosság mintájára egyszerűen adódik, hogy az összegzéssel kapott adatrendszer második momentuma egyenlő a külön-külön vett egydimenziós adatrendszerek második momentumainak az összegével plusz 2-szer a vegyes momentum. A mi numerikus példánkra ez a szabály teljesül:

$$17\,393.4 = 2\,682.6 + 6\,485.6 + 2 \cdot 4\,112.6$$

Az általános formula megfogalmazása és annak algebrai úton való bizonyítása legyen az Olvasó feladata.

Ha a kétdimenziós adatrendszer elemeit oszlopvektorként kezeljük, és az adatokat egymás mellé írjuk, akkor az egymás mellé írt oszlopvektorok egy mátrixot alkotnak. A mátrixnak két sora van, és annyi oszlopa, ahány számpárból áll az adatrendszer.

Ha a kétdimenziós adatrendszer által definiált mátrixot megszorozzuk jobbról az ő transzponáltjával, és osztunk még az adatok számával (ami itt 5), akkor egy olyan 2×2 -es mátrixot kapunk, melynek

- bal felső eleme az x -koordináták adatrendszerének második momentuma
- jobb alsó eleme az y -koordináták adatrendszerének második momentuma
- a bal alsó és a jobb felső eleme pedig a vegyes momentum

Az állítás igazságát a kétdimenziós példánkon ellenőrizhetjük:

$$\begin{array}{|c|c|c|c|c|} \hline 54 & 55 & 44 & 44 & 60 \\ \hline 85 & 83 & 92 & 65 & 75 \\ \hline \end{array} \cdot \begin{array}{|c|c|} \hline 54 & 85 \\ \hline 55 & 83 \\ \hline 44 & 92 \\ \hline 44 & 65 \\ \hline 50 & 75 \\ \hline \end{array} \cdot \frac{1}{5} = \begin{array}{|c|c|} \hline 2682,6 & 4112,6 \\ \hline 4112,6 & 6485,6 \\ \hline \end{array}$$

Ennek a mátrixnak a neve: a kétdimenziós adatrendszer *második momentum mátrixa*.

13.2. Kovariancia és kovariancia mátrix

Ha egy kétdimenziós adatrendszernek vesszük egy elemét, és ebből az elemből kivonjuk az adatrendszer átlagát, (vagyis az első koordinátájából kivonjuk az első koordináták átlagát, és a második koordinátájából kivonjuk a második koordináták átlagát), akkor a szóbanforgó elem *centralizáltjához* jutunk. A centralizálással kapott számpár megadja a szóbanforó elemnek az átlaghoz viszonyított helyzetét. Ha az adatrendszer minden elemének vesszük a centralizáltját, akkor a kapott kétdimenziós adatrendszert az eredeti *adatrendszer centralizáltjának* nevezzük. Az adatrendszer centralizáltja az adatrendszernek az átlaghoz viszonyított helyzetét adja meg.

Példa: A

						átlag
x	54	55	44	44	60	51.4
y	85	83	92	65	75	80

kétdimenziós adatrendszer centralizáltja:

$x - x$ átlag	2.6	3.6	-7.4	-7.4	8.6
$y - y$ átlag	5	3	12	-15	-5

A centralizált adatrendszer vegyes momentumát az eredeti adatrendszer *kovarianciájának* nevezzük. Egyszerűen igazolható tény, hogy

$$\text{kovariancia} = \text{vegyes momentum} - \text{első momentumok szorzata}$$

A centralizált adatrendszer momentum mátrixát az eredeti adatrendszer *kovariancia mátrixának* nevezzük. Nem nehéz belátni, hogy

$$\text{kovariancia mátrix} = \text{második momentum mátrix} - \text{várható értékek oszlopvektora} \cdot \text{várható értékek sorvektora}$$

Ha a kétdimenziós adatrendszer centralizáltja által definiált mátrixot megszorozzuk jobbról az ő transzponáltjával, és osztunk az adatok számával (ami itt 5), akkor egy 2×2 -es mátrixot kapunk, melyet *kovariancia mátrixnak* nevezzük. A kovariancia mátrix

- bal felső eleme az x -koordináták adatrendszerének varianciája
- jobb alsó eleme az y -koordináták adatrendszerének varianciája
- a bal alsó és a jobb felső eleme pedig a kovariancia

Az állítás igazságát a kétdimenziós példánkon ellenőrizhetjük:

$$\begin{bmatrix} 2.6 & 3.6 & -7.4 & -7.4 & 8.6 \\ 5 & 3 & 12 & -15 & -5 \end{bmatrix} \cdot \begin{bmatrix} 2.6 & 5 \\ 3.6 & 3 \\ -7.4 & 12 \\ -7.4 & -15 \\ 8.6 & -5 \end{bmatrix} \cdot \frac{1}{5} = \begin{bmatrix} 40.6 & 0.6 \\ 0.6 & 85.6 \end{bmatrix}$$

Mint fentebb tisztáztuk: az összegzéssel kapott adatrendszer második momentuma egyenlő a külön-külön vett egydimenziós adatrendszerek második momentumainak az összege plusz 2-szer a egyes momentum. Ezt alkalmazva centralizált kétdimenziós adatrendszerre azt kapjuk, hogy az összegzéssel kapott adatrendszer varianciája egyenlő a külön-külön vett egydimenziós adatrendszerek varianciáinak az összege 2-szer a kovariancia.

13.3. Kovariancia mátrix transzformálódása

Ha egy kétdimenziós oszlopvektort megszorozunk balról egy 2×2 -es ún. "transzformációs" mátrixszal, akkor egy másik oszlopvektort kapunk. Ha egy ilyen szorzást megcsiálunk egy kétdimenziós adatrendszer minden oszlopvektorával, akkor egy új kétdimenziós adatrendszert kapunk. Ha az új adatrendszerre előállítjuk a előző részben leírt módon a kovariancia mátrixot, akkor az egész procedúrából balról kiemelhető a 2×2 -es transzformációs mátrix, jobbról pedig a transzformációs mátrix transzponáltja, és kiadódik a következő szabály: az új kétdimenziós adatrendszer kovariancia mátrixa egyenlő az eredeti adatrendszer kovariancia mátrix balról megszorozva a transzformációs mátrixszal, jobbról pedig a transzformációs mátrix transzponáltjával.

Ha a transzformációs mátrix egy olyan sorvektor, mely egy egységvektor koordinátáiból áll, akkor transzformációs mátrixszal balról szorozva egy oszlopvektort a szorzat értéke az oszlopvektornak az egységvektor irányára vonatkozó vetületét adja. Ezért ha egy egységvektornak megfelelő sorvektorral balról, és az ő transzponáltjával jobbról szorzunk egy kovariancia mátrixot, akkor az eredmény annak az adatrendszernek a varianciáját adja, melyet akkor kapunk, ha a kétdimenziós adatrendszer elemeit a sorvektor irányára vetítjük.

Ezért a kovariancia mátrix segítségével a kétdimenziós adatrendszer mindenféle vetületeinek a varianciáját ki lehet számolni. Ilyen értelemben a kovariancia mátrix a variancia fogalmának többdimenzióra vonatkozó kiterjesztésének tekinthető.

13.4. Gyakorló feladatok

- 1.
- 2.

14. Kovariancia valószínűségi változókra

14.1. A kovariancia fogalma

Az X és az Y valószínűségi változók szorzatának $E(XY)$ várható értékét **vegyes (második) momentum**nak nevezzük. Az X és az Y valószínűségi változók centralizáltjai szorzatának a várható értékét **kovarianciának** nevezzük:

$$\text{COV}(X, Y) = E([X - E(X)][Y - E(Y)])$$

Könnyű belátni, hogy a kovarianciát úgy is megkapjuk, hogy a vegyes momentumból kivonjuk a várható értékek szorzatát:

$$\text{COV}(X, Y) = E(XY) - E(X)E(Y)$$

14.2. A kovariancia tulajdonságai

1. Egy valószínűségi változó önmagával vett kovarianciája a valószínűségi változó varianciájával egyenlő:

$$\text{COV}(X, X) = \text{VAR}(X)$$

2. A kovariancia szimmetria tulajdonsága:

$$\text{COV}(Y, X) = \text{COV}(X, Y)$$

3. A kovariancia linearitási tulajdonsága:

- Két tagra:

$$\text{COV}(aX + bY, Z) = a \text{COV}(X, Z) + b \text{COV}(Y, Z)$$

- Több tagra:

$$\text{COV}(a_1X_1 + a_2X_2 + \dots + a_nX_n, Y) = a_1 \text{COV}(X_1, Y) + a_2 \text{COV}(X_2, Y) + \dots + a_n \text{COV}(X_n, Y)$$

4. A kovariancia abszolút értéke mindig kisebb vagy egyenlő, mint a szórások szorzata:

$$|\text{COV}(X, Y)| \leq \text{SD}(X) \text{SD}(Y)$$

5. Független valószínűségi változók között a kovariancia 0:

$$\text{COV}(X, Y) = 0$$

6. Összeg varianciája:

$$\text{VAR}(X + Y) = \text{VAR}(X) + \text{VAR}(Y) + 2 \text{COV}(X, Y)$$

7. Egymástól független vektorváltozók összegének koordinátái közötti kovariancia egyenlő a tagok koordinátái közötti kovarianciák összegével:

- Két tagra:

Ha az (X_1, Y_1) , (X_2, Y_2) véletlen vektorok függetlenek egymástól, és

$$(X, Y) = (X_1, Y_1) + (X_2, Y_2)$$

akkor

$$\text{COV}(X, Y) = \text{COV}(X_1, Y_1) + \text{COV}(X_2, Y_2)$$

- Több tagra:
Ha az $(X_1, Y_1), \dots, (X_n, Y_n)$ véletlen vektorok függetlenek egymástól, és

$$(X, Y) = (X_1, Y_1) + \dots + (X_n, Y_n)$$

akkor

$$\text{COV}(X, Y) = \text{COV}(X_1, Y_1) + \dots + \text{COV}(X_n, Y_n)$$

Megjegyzés: A feltételben nem az X_1 független Y_1 -től, illetve X_2 független Y_2 -től, hanem az (X_1, Y_1) vektor független az (X_2, Y_2) vektortól.

8. *Független vektor valószínűségi változók összegének koordinátái közötti kovariancia, amikor a tagok koordinátái közötti kovarianciák azonosak:* Ha az $(X_1, Y_1), \dots, (X_n, Y_n)$ véletlen vektorok függetlenek egymástól, és mindegyik tag esetében a koordináták közötti kovariancia azonos, azaz

$$\text{COV}(X_i, Y_i) = c \quad \text{minden } i\text{-re}$$

és

$$(X, Y) = (X_1, Y_1) + \dots + (X_n, Y_n)$$

akkor az összeg koordinátái közötti kovariancia $(c n)$ -nel egyenlő:

$$\text{COV}(X, Y) = c n$$

14.3. A korrelációs együttható fogalma és tulajdonságai

1. Az X és az Y valószínűségi változók standardizáltjai szorzatának a várható értékét **korrelációs együtthatójának** nevezzük:

$$\text{CORR}(X, Y) = E \left(\frac{X - E(X)}{\text{SD}(X)} \cdot \frac{Y - E(Y)}{\text{SD}(Y)} \right)$$

2. A korrelációs együttható abszolút értéke mindig kisebb vagy egyenlő, mint 1:

$$|\text{CORR}(X, Y)| \leq 1$$

3. Egy valószínűségi változónak bármely lineáris transzormáltjával vett korrelációs együtthatója 1-gyel egyenlő:

$$\text{CORR}(X, aX + b) = 1$$

4. Ha X és Y között a korreláció -1 vagy $+1$, akkor vannak olyan a és b konstansok, hogy

$$Y = aX + b \quad \text{fennáll 1 valószínűséggel, azaz } P(Y = aX + b) = 1$$

vagyis az X és Y véletlen értékek között determinisztikus lineáris kapcsolat áll fenn.

5. Független valószínűségi változók között a korreláció 0:

$$\text{CORR}(X, Y) = 0$$

14.4. A kovariancia mátrix fogalma és tulajdonságai

1. Az (X, Y) kétdimenziós valószínűségi változó **kovariancia mátrixának** definíciója:

$$C = \begin{pmatrix} \text{COV}(X, X) & \text{COV}(X, Y) \\ \text{COV}(Y, X) & \text{COV}(Y, Y) \end{pmatrix}$$

azaz

$$C = \begin{pmatrix} \text{VAR}(X) & \text{COV}(X, Y) \\ \text{COV}(X, Y) & \text{VAR}(Y) \end{pmatrix}$$

A kovariancia mátrix a variancia fogalmának általánosítása kétdimenzióra.

2. *Várható érték vektor és kovariancia mátrix transzformálódása lineáris transzformáció esetén.*

Ha az \mathbf{x} -síkról egy kétdimenziós eloszlást az

$$\mathbf{u} = \mathbf{A} \mathbf{x} + \mathbf{b}$$

lineáris transzformációval az \mathbf{u} -síkra képezünk, akkor a kapott új eloszlás várható érték vektora a régi $\mathbf{m}_{\text{régi}}$ várható érték vektor lineáris transzformáltja:

$$\mathbf{m}_{\text{új}} = \mathbf{A} \mathbf{m}_{\text{régi}} + \mathbf{b}$$

Ha a régi eloszlás kovariancia mátrixa $\mathbf{C}_{\text{régi}}$, akkor a transzformációval kapott új eloszlás kovariancia mátrixa

$$\mathbf{C}_{\text{új}} = \mathbf{A} \mathbf{C}_{\text{régi}} \mathbf{A}^T$$

3. *Független vektor valószínűségi változók összegének kovariancia mátrixa.* Független vektor valószínűségi változók összegének kovariancia mátrixa egyenlő a tagok kovariancia mátrixainak összegével.
4. *Független vektor valószínűségi változók összegének kovariancia mátrixa, amikor a tagok kovariancia mátrixai azonosak.* Független vektor valószínűségi változók összegének kovariancia mátrixa, amikor a tagok kovariancia mátrixai azonosak, egyenlő a tagok közös kovariancia mátrixa szorozva a tagok számával.

14.5. Gyakorló feladatok

- 1.
- 2.

15. NSZT a vegyes momentumra, kovarianciára, korrelációs együtthatóra

Heurisztikus megfogalmazás: Tekintsünk egy (X, Y) kétdimenziós valószínűségi változót. A kétdimenziós valószínűségi változó eloszlásából ki lehet számolni a vegyes momentumot, a kovarianciát, a korrelációs együtthatót. Ha a kétdimenziós valószínűségi változóra N kísérletet végzünk, és a kísérleti eredmények

$$(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$$

akkor a kísérleti eredményekből adódó kétdimenziós adatrendszer vegyes momentuma, kovarianciája, korrelációs együtthatója szintén kiszámolható. Fontos – matematikailag precízen is igazolható – tény, hogy *nagy kísérletszám esetén a kísérleti eredményekből adódó kétdimenziós adatrendszer vegyes momentuma, kovarianciája, korrelációs együtthatója körülbelül egyenlő az (X, Y) kétdimenziós valószínűségi változó eloszlásából kiszámolt vegyes momentumával, kovarianciájával, korrelációs együtthatójával.*

15.1. Gyakorló feladatok

- 1.
- 2.

16. Kétdimenziós normális eloszlások alakja

16.1. A korrelációs együttható szemléletes jelentése

Ez itt csak rövid emlékeztetője az előadáson elhangzottaknak. A készülő jegyzetben alaposabban lesz kifejtve. Az előadás gondolatmenetének fontos részét képezi az idevonatkozó Excel fájl, ami a tárgy honlapján elérhető.

A kétdimenziós normális eloszlásokkal kapcsolatban a korrelációs együttható és az eloszlás alakja között kapcsolatot áll fenn. A kapcsolat leírásának céljából tekintsük a

$$(\mu_1 - \sigma_1 ; \mu_1 + \sigma_1) \text{ és a } (\mu_2 - \sigma_2 ; \mu_2 + \sigma_2)$$

intervallumok direktszorzataként adódó téglalapot, vagy általánosabban a

$$(\mu_1 - s \cdot \sigma_1 ; \mu_1 + s \cdot \sigma_1) \text{ és a } (\mu_2 - s \cdot \sigma_2 ; \mu_2 + s \cdot \sigma_2)$$

intervallumok direktszorzataként adódó téglalapot, ahol $s = 1$ vagy $s = 2$ vagy $s = 3$. Ezeket a téglalapokat a **szórások alapján felvett téglalapoknak** nevezzük. Képzeljünk el egy-egy skálát -1 -től 1 -ig a téglalapok minegyik oldalán úgy, hogy

- a vízszintes oldalak bal végében a -1 , a jobb végében pedig a 1 helyezkedik el
- a függőleges oldalak alsó végében a -1 , a felső végében pedig a 1 helyezkedik el

Mindegyik oldalon tekintsük most az r értékhez tartozó pontot, ahol r a korrelációs együttható. Így négy pontot jelöltünk ki a síkon. Ezen a négy ponton át nem nehéz egy olyan ellipszist rajzolni, mely belülről érinti a téglalapot. **Ez az ellipszis a normális eloszlás sűrűségfüggvényének egy szintvonala.** Ez az ellipszis meghatározza a normális eloszlás alakját, mert a többi szintvonal ennek az ellipszisnek a kicsinyítésével, vagy nagyításával keletkezik.

16.2. A kovariancia mátrix sajátvektorai, sajátértékei

Egy kétdimenziós normális eloszlás sűrűségfüggvényének szintvonalai olyan ellipszisek, melyek tengelyinek irányát a kovariancia mátrix sajátvektorai adják meg. Az ellipszisek tengelyeinek hosszai pedig arányosak a sajátértékek négyzetgyökeivel.

16.3. Gyakorló feladatok

- 1.
- 2.

17. Polinomiális eloszlás közelítése kétdimenziós normális eloszlással (*Extra tananyag*)

Ha az n paraméter elég nagy, és a p_1, p_2 paraméterek se a 0-hoz se az 1-hez nincsenek túl közel, akkor a síkon vett n -ed rendű, (p_1, p_2) paraméterű polinomiális eloszlás közelíthető kétdimenziós normális eloszlással. A normális eloszlás paramétereit a polinomiális eloszlás paramétereire úgy kell hozzáigazítani, hogy a megfelelő paraméterek (várható értékek, szórások, korrelációs együttható) megegyezzenek, vagyis

$$\mu_1 = np_1 \quad \mu_2 = np_2 \quad \sigma_1 = \sqrt{np_1(1-p_1)} \quad \sigma_2 = \sqrt{np_2(1-p_2)} \quad r = -\sqrt{\frac{p_1}{(1-p_1)} \frac{p_2}{(1-p_2)}}$$

17.1. Gyakorló feladatok

- 1.
- 2.

18. Többdimenziós normális eloszlások transzformációi (*Extra tananyag*)

18.1. A standard normális eloszlás lineáris transzformációi

Jön majd ide

18.2. Tetszőleges normális eloszlás lineáris transzformációi

Jön majd ide

18.3. Khí-négyzet eloszlás

Ez itt csak rövid emlékeztetője az előadáson elhangzottaknak. A készülő jegyzetben alaposabban lesz kifejtve. Az előadás gondolatmenetének fontos részét képezi az idevonatkozó Excel fájl, ami a tárgy honlapján elérhető.

Ha n darab független standard normális eloszlású valószínűségi változó négyzetét összeadjuk, akkor egy valószínűségi változóhoz jutunk. Ennek a valószínűségi változónak az eloszlását n **szabadságfokú khí-négyzet eloszlásnak** hívjuk.

18.4. Gyakorló feladatok

- 1.
- 2.

19. Lineáris regresszió

19.1. Általános probléma

Tekintsünk egy (X, Y) kétdimenziós valószínűségi változót. Célunk hogy az, hogy megtaláljuk azokat az A és B konstansokat, melyekre az

$$E\left((Y - (AX + B))^2\right) = \iint_{R^2} (y - (Ax + B))^2 f(x, y) dx dy$$

várható érték minimális.

Megoldás: Ha a fenti integrálban elvégezzük a négyzetre emelést, és az integrált tagokra bontjuk, akkor hat tagot kapunk:

$$\begin{aligned} E\left((Y - (AX + B))^2\right) &= \\ &= \iint_{R^2} y^2 f(x, y) dx dy + A^2 \iint_{R^2} x^2 f(x, y) dx dy + B^2 \iint_{R^2} f(x, y) dx dy - \\ &- 2A \iint_{R^2} xy f(x, y) dx dy - 2B \iint_{R^2} y f(x, y) dx dy + 2AB \iint_{R^2} x f(x, y) dx dy \end{aligned}$$

Az itt szereplő integrálok értékei konstansokat adnak, ezért A -ra, B -re nézve egy kvadratikus formulát kaptunk. Azokat az A, B értékeket, melyekre ez a kvadratikus formula minimális, megkaphatjuk például úgy, hogy parciálisan deriválunk A és B szerint, majd a parciális deriváltakat nullával tesszük egyenlővé, és az így kiadódó egyenletrendszert A -ra, B -re megoldjuk. A számolás részleteitől eltekintünk, csak az eredményt közöljük:

$$A_{\text{opt}} = r \frac{\sigma_2}{\sigma_1}$$

$$B_{\text{opt}} = \mu_2 - r \frac{\sigma_2}{\sigma_1} \mu_1$$

Az optimumot nyújtó lineáris függvény tehát így fest:

$$y = \mu_2 + r \frac{\sigma_2}{\sigma_1} (x - \mu_1)$$

Az egyenletet átírhatjuk erre a – könnyebben megjegyezhető – formára is:

$$\frac{y - \mu_2}{\sigma_2} = r \frac{x - \mu_1}{\sigma_1}$$

A kapott egyenes neve: **regressziós egyenes**.

19.2. Számolós példa

Legyen $X = \text{RND}_1 \cdot \text{RND}_2$ és $Y = \text{RND}_1$. Határozzuk meg

- X, Y, XY várható értékét,
- X, Y varianciáját, szórását,
- az X és közötti Y kovarianciáját és a korrelációs együtthatót!
- Írjuk fel a regressziós egyenes egyenletét, ha Y -ből akarunk X -re tippelni!
- Írjuk fel a regressziós egyenes egyenletét, ha X -ből akarunk Y -ra tippelni!

Megoldás:

$$E(X) = E(\text{RND}_1 \cdot \text{RND}_2) = E(\text{RND}_1) \cdot E(\text{RND}_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$E(X^2) = E(\text{RND}_1^2 \cdot \text{RND}_2^2) = E(\text{RND}_1^2) \cdot E(\text{RND}_2^2) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$$

$$\text{VAR}(X) = E(X^2) - (E(X))^2 = \frac{1}{9} - \left(\frac{1}{4}\right)^2 = \frac{7}{144}$$

$$\text{SD}(X) = \sqrt{\text{VAR}(X)} = \sqrt{\frac{7}{144}}$$

$$E(Y) = E(\text{RND}_1) = \frac{1}{2}$$

$$E(Y^2) = E(\text{RND}_1^2) = \frac{1}{3}$$

$$\text{VAR}(Y) = E(Y^2) - (E(Y))^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}$$

$$\text{SD}(Y) = \sqrt{\text{VAR}(Y)} = \sqrt{\frac{1}{12}}$$

$$E(X \cdot Y) = E(\text{RND}_1 \text{RND}_2 \cdot \text{RND}_1) = E(\text{RND}_1^2 \cdot \text{RND}_2) = E(\text{RND}_1^2) \cdot E(\text{RND}_2) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$$

$$\text{COV}(X; Y) = E(X \cdot Y) - E(X) \cdot E(Y) = \frac{1}{6} - \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{24}$$

$$\text{CORR}(X; Y) = \frac{\text{COV}(X; Y)}{\text{SD}(X) \text{SD}(Y)} = \frac{\frac{1}{24}}{\sqrt{\frac{7}{144}} \sqrt{\frac{1}{12}}} = \sqrt{\frac{3}{7}}$$

$$y = \mu_2 + r \frac{\sigma_2}{\sigma_1} (x - \mu_1)$$

$$y = \frac{1}{2} + \sqrt{\frac{3}{7}} \frac{\sqrt{\frac{1}{12}}}{\sqrt{\frac{7}{144}}} \left(x - \frac{1}{4}\right)$$

$$y = \frac{6}{7}x + \frac{2}{7}$$

19.3. Gyakorló feladatok

1. Legyen $X = \text{RND}_1 \cdot \text{RND}_2$ és $Y = \text{RND}_1^2$. Határozza meg

- (a) $X, Y, (X, Y)$ sűrűségfüggvényének képletét, majd ezekből számolja ki
- (b) X, Y, XY várható értékét,
- (c) X, Y varianciáját, szórását,
- (d) az X és közötti Y kovarianciát és a korrelációs együtthatót!
- (e) Írja fel a regressziós egyenes egyenletét, ha Y -ből akarunk X -re tippelni!
- (f) Írja fel a regressziós egyenes egyenletét, ha X -ből akarunk Y -ra tippelni!

2.

20. Statisztika

20.1. Khí-négyzet próba

Remélem, hogy az Olvasó örömmel tapasztalja, hogy ez a fejezet fejlődött (de nem bővült) a 2016. december 16 -i verzióhoz képest.

Tekintsünk egy r elemű teljes eseményrendszert. Az egyes események valószínűségei legyenek

$$p_1, p_2, \dots, p_r$$

Ha N kísérletet végzünk, akkor nézhetjük, hogy a teljes eseményrendszer egyes elemei hányszor következnek be. Legyenek ezek a gyakoriság értékek:

$$X_1, X_2, \dots, X_r$$

Legyenek

$$p_1^{\text{hip}}, p_2^{\text{hip}}, \dots, p_r^{\text{hip}}$$

valamilyen **hipotetikus valószínűségek**, vagyis olyan valószínűségek, melyeket valaki csak úgy gondol, képzeli, és épp a kísérleti eredményekből akarja eldönteni, hogy az események igazi valószínűségei mind megegyeznek-e az általa gondolt hipotetikus valószínűségekkel:

$$p_1 = p_1^{\text{hip}}, \quad p_2 = p_2^{\text{hip}}, \quad \dots, \quad p_r = p_r^{\text{hip}} \quad ?$$

Ha a tényleges valószínűségek mind megegyeznek a hipotetikus valószínűségekkel, akkor az

$$X_1, X_2, \dots, X_r$$

gyakoriságok várható értékei:

$$Np_1^{\text{hip}}, \quad Np_2^{\text{hip}}, \quad \dots, \quad Np_r^{\text{hip}}$$

Ezért ezeket a szorzatokat **hipotetikus várható értékeknek** nevezzük.

Tekintsük a gyakoriságoknak a hipotetikus várható értékektől való eltérésének a négyzetét:

$$\left(X_1 - Np_1^{\text{hip}}\right)^2, \quad \left(X_2 - Np_2^{\text{hip}}\right)^2, \quad \dots, \quad \left(X_r - Np_r^{\text{hip}}\right)^2$$

aztán az alábbi hányadosokat:

$$\frac{\left(X_1 - Np_1^{\text{hip}}\right)^2}{Np_1^{\text{hip}}}, \quad \frac{\left(X_2 - Np_2^{\text{hip}}\right)^2}{Np_2^{\text{hip}}}, \quad \dots, \quad \frac{\left(X_r - Np_r^{\text{hip}}\right)^2}{Np_r^{\text{hip}}}$$

s végül ezeknek az összegét, amit $\chi_{\text{megfigyelt}}^2$ -tel jelölünk, és **megfigyelt kí-négyzet értéknek** nevezzük:

$$\chi_{\text{megfigyelt}}^2 = \frac{\left(X_1 - Np_1^{\text{hip}}\right)^2}{Np_1^{\text{hip}}} + \frac{\left(X_2 - Np_2^{\text{hip}}\right)^2}{Np_2^{\text{hip}}} + \dots + \frac{\left(X_r - Np_r^{\text{hip}}\right)^2}{Np_r^{\text{hip}}}$$

Tegyük fel, hogy az N kísérletszám annyira nagy, hogy hipotetikus várható értékek mindegyike nagyobb 10 -nél. (Ha ez a feltétel nem teljesül, akkor az alábbi eljárás – pontatlansága miatt – nem használható.)

Be lehet látni, hogy az alábbiak igazak:

- Ha a tényleges valószínűségek mind megegyeznek a hipotetikus valószínűségekkel, akkor a megfigyelt kí-négyzet érték közelítőleg $r - 1$ szabadságfokú kí-négyzet eloszlást követ.

- Ha pedig a tényleges valószínűségek nem mind egyeznek meg a hipotetikus valószínűségekkel, akkor a megfigyelt χ^2 -négyzet érték eloszlása jelentősen eltér az $r - 1$ szabadságfokú χ^2 -négyzet eloszlástól, és a megfigyelt χ^2 -négyzet érték általában nagy.

Ezeket a tényeket érezhetővé tehetjük, ha a megfigyelt χ^2 -négyzet értéket az alábbi alakra hozzuk:

$$\chi_{\text{megfigyelt}}^2 = N \cdot \left[\frac{\left(\frac{X_1}{N} - p_1^{\text{hip}}\right)^2}{p_1^{\text{hip}}} + \frac{\left(\frac{X_2}{N} - p_2^{\text{hip}}\right)^2}{p_2^{\text{hip}}} + \dots + \frac{\left(\frac{X_r}{N} - p_r^{\text{hip}}\right)^2}{p_r^{\text{hip}}} \right]$$

Ugyanis ebből az alakból láthatóak, hihetőek az alábbiak:

- Ha a tényleges valószínűségek mind megegyeznek a hipotetikus valószínűségekkel, akkor minden i -re az $\frac{X_i}{N}$ relatív gyakoriság – nagy N esetén – közel van a p_i^{hip} hipotetikus valószínűséghez, és így a most felírt kifejezés szögletes zárójelében álló tagok mind kicsik. N növekedtével a tagok még kisebbek, hiszen a relatív gyakoriságok még közelebb kerülnek a valószínűséghez. Emiatt elhíhet, elképzelhető, hogy a szögletes zárójelben lévő összeg értéke N -nel megszorozva (ahogy a megfigyelt χ^2 -négyzet értéket kapjuk) olyan véletlen érték, ami közelítőleg egy jól meghatározott eloszlást követ. Ez az eloszlás történetesen az $r - 1$ szabadságfokú χ^2 -négyzet eloszlás. Tehát ilyenkor a megfigyelt χ^2 -négyzet érték közelítőleg $r - 1$ szabadságfokú χ^2 -négyzet eloszlást követ.
- Ha pedig a tényleges valószínűségek nem mind egyeznek meg a hipotetikus valószínűségekkel, akkor van olyan i , hogy az $\frac{X_i}{N}$ relatív gyakoriság – nagy N esetén – a p_i^{hip} hipotetikus valószínűségtől különböző p_i -hez lesz közel, és így az

$$\frac{\left(\frac{X_i}{N} - p_i^{\text{hip}}\right)^2}{p_i^{\text{hip}}}$$

kifejezés értéke a pozitív

$$\frac{\left(p_i - p_i^{\text{hip}}\right)^2}{p_i^{\text{hip}}}$$

szám közelében lesz. Mivel a szögletes zárójelben álló kifejezés a $\frac{\left(\frac{X_i}{N} - p_i^{\text{hip}}\right)^2}{p_i^{\text{hip}}}$ tagnál nagyobb, a szögletes zárójel előtt álló N -nel való szorzás – nagy N esetén – a megfigyelt χ^2 -négyzet értéket nagyra teszi. Ha N nagyobb, akkor még nagyobb. Tehát ilyenkor – nagy N esetén – a megfigyelt χ^2 -négyzet érték nagy szeret lenni.

Az $r - 1$ szabadságfokú χ^2 -négyzet eloszlás segítségével tudunk választani egy χ_{kritikus}^2 -sal jelölt ún. **kritikus χ^2 -négyzet értéket**, mely arról nevezetes, hogy az $r - 1$ szabadságfokú χ^2 -négyzet eloszlást követő valószínűségi változó 0.95 valószínűséggel kisebb, 0.05 valószínűséggel nagyobb ennél a χ_{kritikus}^2 -sal jelölt értéknél. (Arról, hogy itt miért éppen a 0.05 valószínűség érték áll, és nem valami más, később ejtünk szót.) A χ_{kritikus}^2 értéket az $F(x)$ baloldali- vagy a $T(x)$ jobboldali eloszlásfüggvény inverzével kapjuk a

$$\chi_{\text{kritikus}}^2 = F^{-1}(0.95) \quad \chi_{\text{kritikus}}^2 = T^{-1}(0.05)$$

összefüggések valamelyikéből.

Ennek segítségével így dönthetünk afelől, hogy a tényleges valószínűségek mind megegyeznek-e a hipotetikus valószínűségekkel:

- Ha a megfigyelt χ^2 -négyzet érték kisebb, mint a kritikus χ^2 -négyzet érték, $\chi_{\text{megfigyelt}}^2 < \chi_{\text{kritikus}}^2$, akkor arra tippelünk, hogy a tényleges valószínűségek mind megegyeznek a hipotetikus valószínűségekkel.

- Ha a megfigyelt χ^2 -négyzet érték nagyobb mint a kritikus χ^2 -négyzet érték, $\chi_{\text{megfigyelt}}^2 > \chi_{\text{kritikus}}^2$, akkor arra tippelünk, hogy a tényleges valószínűségek nem mind egyeznek meg a hipotetikus valószínűségekkel.

Ez az eljárás – a fentebb mondottak miatt – ilyen:

- Ha a tényleges valószínűségek mind megegyeznek-e a hipotetikus valószínűségekkel, akkor kb. 0.95 valószínűséggel a megfigyelt χ^2 -négyzet érték kisebb, mint a kritikus χ^2 -négyzet érték, és így az eljárás helyesen dönt, és csak kb. 0.05 valószínűséggel lesz a megfigyelt χ^2 -négyzet érték nagyobb, mint a kritikus χ^2 -négyzet érték, ami miatt az eljárás hibázik.
- Ha pedig a tényleges valószínűségek nem mind egyeznek meg a hipotetikus valószínűségekkel, akkor nagy valószínűséggel a megfigyelt χ^2 -négyzet érték nagyobb, mint a kritikus χ^2 -négyzet érték, és így az eljárás helyesen dönt, és csak kis valószínűséggel lesz a megfigyelt χ^2 -négyzet érték kisebb, mint a kritikus χ^2 -négyzet érték, ami miatt az eljárás hibázik. (Azt, hogy a "nagy valószínűséggel", "kis valószínűséggel" kifejezések pontosan mit jelentenek, itt nem vizsgáljuk.)

Eléggé elterjedt az alkalmazók körében, hogy a kritikus értéket úgy választják meg, hogy a hipotézis fennállása esetén a hipotézis elfogadásának (a helyes döntésnek) a valószínűsége 0.95, az elutasítás (a hibás döntésnek) a valószínűsége 0.05 legyen. A 0.95, 0.05 értékek kitüntetett szerepét igazából semmi sem indokolja. Egyes esetekben más **elfogadási- és elutasítási valószínűségek** használata is indokolt lehet. Az elfogadási- és elutasítási valószínűségek megválasztása nem matematikai kérdés, az alkalmazók szubjektív választásán múlik.

- Ha valakinek az a fontos, hogy az eljárás alkalmazásakor a hipotézis fennállása esetén ritkábban tévedjen, akkor az eljárást a 0.95 helyett nagyobb – mondjuk – 0.99 elfogadási- és 0.01 elutasítási valószínűség értékkel – és ennek megfelelően – nagyobb kritikus értékkel kell használnia.
- Ha pedig az a fontos, hogy lehetőleg ne tévedjünk, amikor a hipotézis nem áll fenn, akkor 0.95 helyett kisebb – mondjuk – 0.90 elfogadási és 0.10 elutasítási valószínűség értékkel és így kisebb kritikus értékkel kell dolgoznunk.

A két igény, hogy se a hipotézis fennállása, se a hipotézis nem fennállása esetén ne tévedjünk, egymásnak ellentmond. Az egyik igény javítása csak a másik kárára történhet.

20.1.1. Prüntyőke fiókák száma

Már rég megfigyelték, hogy a prüntyőke madarak fészkeiben tavaszi költéskor a tojások száma 2, 3 vagy 4. Mostanában olvatom egy újságban, hogy valaki – elméleti kutatómunkája eredményeként – azt hozta ki, hogy

$$P(2 \text{ tojás}) = 0.2 \quad P(3 \text{ tojás}) = 0.5 \quad P(4 \text{ tojás}) = 0.3$$

Én – gyakorlati oldalról – kissé utánanéztam a dolognak. Nem kis munkával 90 fészket kutattam fel. 12 esetben 2 tojást, 43 esetben 3 tojást, 35 esetben 4 tojást láttam. Tehát a korábbi jelölésekkel:

$$r = 3, \quad p_1^{\text{hip}} = 0.2, \quad p_2^{\text{hip}} = 0.5, \quad p_3^{\text{hip}} = 0.3, \quad N = 90, \quad X_1 = 12, \quad X_2 = 43, \quad X_3 = 35$$

A megfigyelt χ^2 -négyzet értéket könnyűszerrel kiszámoltam: $\chi_{\text{megfigyelt}}^2 = 4.46$. Ezt összevetve a 0.95 -ös elfogadási valószínűséghez tartozó $\chi_{\text{kritikus}}^2 = 5.99$ -es kritikus értékkel, levontam a következtetést. Mivel $\chi_{\text{megfigyelt}}^2 < \chi_{\text{kritikus}}^2$, a megfigyelésem nem mond ellent a tudós állításának, ezért elfogadtam a tudós által kijelentett valószínűség értékeket.

1. Megjegyzés: Fontos, hogy a 0.95 -ös elfogadási valószínűség jelentését ne értsük félre. A 0.95 nem azt jelenti, hogy a tudós állítása 0.95 valószínűséggel igaz. Azt sem jelenti, hogy az én ítéletem 0.95 valószínűséggel helyes. A 0.95 -ös elfogadási valószínűség az általam használt döntési eljárásnak a jellemzője: a döntési eljárás, amit én itt most egyszer használtam olyan, hogy ha igazak a

$$P(2 \text{ tojás}) = 0.2 \quad P(3 \text{ tojás}) = 0.5 \quad P(4 \text{ tojás}) = 0.3$$

valószínűségek, akkor

- 0.95 a valószínűsége annak, hogy a megfigyelt χ^2 -négyzet érték az 5.99 -es kritikus értéknél kisebbnek adódik
- 0.05 a valószínűsége annak, hogy a megfigyelt χ^2 -négyzet érték az 5.99 -es kritikus értéknél nagyobbak adódik

Mással

- 0.95 a valószínűsége annak, hogy az eljárás helyesen dönt
- 0.05 a valószínűsége annak, hogy az eljárás tévesen dönt

2. Megjegyzés: Megjegyezzük, hogy ha a 90 fészek közül 8 -ban 2 tojást, 47 -ben 3 tojást, 35 -ben 4 tojást találtam volna, akkor a megfigyelt χ^2 -négyzet érték 8.01 -nek adódott volna, ami – 0.95 -ös elfogadási valószínűség mellett – már a tudós állításának az elutasítást indikálná.

3. Megjegyzés: Amde a 0.95 -nél nagyobb 0.99 -es elfogadási valószínűség mellett ugyanez a 8.01 -es megfigyelt χ^2 -négyzet érték – mivel a kritikus érték 9.21 – a tudós állításának az elfogadást indikálná. Nem meglepő, hogy ha növeljük az elfogadási valószínűséget, akkor elfogadóbbá válunk.

20.1.2. Madárfiókák helyett színes golyók

Lehet, hogy van, akinek segít a fenti probléma megemésztésében az alábbi könnyen elképzelhető modell, mely ekvivalens az előző pontban tált problémával.

Tegyük fel, hogy egy dobozban 10 színes golyó van: pirosak, fehérek és zöldek. Valaki azt állítja, hogy

2 piros, 5 fehér és 3 zöld golyó van a dobozban

Én nem nézhetek bele a dobozba, viszont csukott szemmel húzhatok a dobozból – mondjuk – 90 -szer, és – mielőtt visszateszem a golyókat, mindig megnézhetem a színüket.

Tegyük fel, hogy a 90 húzásból 12 esetben pirosat, 43 esetben fehéret, 35 esetben zölDET húztam. Elfogadjam-e, hogy 2 piros, 5 fehér és 3 zöld golyó van a dobozban?

Az okoskodás pontosan ugyanaz lenne, mint a prüntyőke fiókák számával kapcsolat fentebb volt. A

$$r = 3, \quad p_1^{\text{hip}} = 0.2, \quad p_2^{\text{hip}} = 0.5, \quad p_3^{\text{hip}} = 0.3, \quad N = 90, \quad X_1 = 12, \quad X_2 = 43, \quad X_3 = 35$$

adatokból kiszámolt megfigyelt χ^2 -négyzet érték: $\chi_{\text{megfigyelt}}^2 = 4.46$, amit összevetve a 0.95 -ös elfogadási valószínűséghez tartozó $\chi_{\text{kritikus}}^2 = 5.99$ -es kritikus értékkel, elfogadtam a doboz állított összetételét.

20.1.3. Egyforma esélyűek a lottózszámok?

1956 óta 3120 húzás volt az ötöslottón. Az alábbi táblázat azt mutatja, hogy az eltelt 60 év során az egyes számokat hányszor húzták ki:

1	188	16	166	31	156	46	171	61	166	76	185
2	159	17	158	32	172	47	186	62	162	77	195
3	202	18	187	33	171	48	153	63	145	78	179
4	176	19	178	34	174	49	177	64	180	79	166
5	147	20	178	35	178	50	169	65	174	80	154
6	163	21	163	36	176	51	185	66	195	81	179
7	180	22	185	37	172	52	177	67	173	82	157
8	156	23	183	38	176	53	164	68	159	83	179
9	160	24	185	39	157	54	178	69	191	84	176
10	193	25	177	40	158	55	181	70	158	85	176
11	178	26	168	41	171	56	192	71	181	86	194
12	185	27	174	42	194	57	175	72	183	87	148
13	191	28	162	43	178	58	155	73	185	88	135
14	163	29	201	44	161	59	177	74	164	89	154
15	190	30	157	45	175	60	177	75	197	90	171

Látjuk, hogy a 3-as számot 202 -szer, a 88 -at pedig csak 135 -ször. Mivel a 202 a 135 -nek másfélszerese, felmerülhet bennünk a gyanu, hogy a lottószámok nem egyformán valószínűek. Pedig úgy illene, hogy a 90 lottó szám egyformán valószínű legyen!

Ha a hipotézisünk az, hogy a lottószámok húzásánál minden szám valószínűsége $\frac{1}{90}$, akkor a fenti táblázatban olvasható gyakoriságokkal számolva (aki nem hiszi, számoljon utána) a megfigyelt hí-négyzet érték:

$$\chi_{\text{megfigyelt}}^2 = 95.91$$

A 89 szabadságfogó hí-négyzet eloszlásból a 0.95 elfogadási szinthez tartozó kritikus érték

$$\chi_{\text{kritikus}}^2 = F^{-1}(0.95) = 112.02$$

Mivel a megfigyelt hí-négyzet érték kisebb, mint a kritikus érték, a táblázat adatai nem mondanak ellent a hipotézisnek, elfogadjuk, hogy a lottószámok egyformán valószínűek.

20.2. U-próba

Tegyük fel, hogy a férfiak testmagassága a Föld minden országban normális eloszlást követ 10 cm szórással és valamilyen – az országra jellemző – várható értékkel. A várható érték más-más lehet: pl. a svédek magasabbak, a pigmeusok alacsonyabbak, de a szórásról azt feltételezzük, hogy minden országban $\sigma = 10$ cm.

Egy marslakó leszáll a Földre, kinéz az űrhajóból, és látja a nagy plakátot: "Ebben az országban a férfiak testmagasságának a várható értéke 180 cm". A marslakó el akarja dönteni, hogy a plakát igazat beszél-e. Tanult valószínűség-számítást, tudja, hogy a várható érték az országos átlagot jelenti. Tisztában van vele, hogy az ország minden férfi emberének magasságát nem tudja megmérni, ezért 4 véletlenül odaverődő, bámészkodó férfi testmagasságát megméri. Tegyük fel, hogy ezeket az adatokat kapja: 177.3, 174.8, 184.2, 178.5 cm. Kiszámolja az átlagot, ami 178.7.

Vakarja is ám a fejét a marslakó: mert a 178.7 nem 180, ahogy a plakáton látja. Először – elhamarkodottan – azt gondolja: lám, hazudnak! De hamar rájön, hogy nagy butaság lenne, ha csak akkor adna igazat a plakáton látottaknak, ha az átlag pontosan 180 lenne! Hiszen annak a valószínűsége, hogy egy folytonos valószínűségi változó, mint amilyen a négy mérési eredmény átlaga, pontosan egyenlő egy adott értékkel – mint azt jól megtanulta – nullával egyenlő.

Gondolkodni kezd. A 178.7 -nek és a 180 -nak az eltérése nem túl nagy, ezért elképzelhető, hogy a várható érték mégis 180. Vagy mégse? Hiszen a 178.7 -nek és a 180 -nak az eltérése azért mégse olyan kicsike, hiszen 1 -nél azért csak nagyobb.

Segítsünk a nagy hezitálásban szenvedő marslakónak! Szeretné elkerülni, hogy hibázzon.

- Hiba lenne a 180 -as országos átlag teljesülése esetén elutasítani a 180 -as átlagot.
- De hiba lenne az is, hogy ha az országos átlag nem 180, és ő mégis elfogadja a plakáton reklámozott 180 -as átlagot.

Mivel a marslakó döntési eljárásába a véletlen is beleszól (éppen ki az a 4 báméskodó ember, akiket történetesen lemér), nyilvánvaló, hogy bele kell törödni abba, hogy esetleg hibázik.

Naív álmai szerint azt szeretné, hogy a döntési eljárás olyan legyen, hogy

- ha az országos átlag 180, akkor a döntési eljárás 1 valószínűséggel a 180 mellett döntsön,
- ha az országos átlag különbözik 180 -tól, akkor a döntési eljárás 0 valószínűséggel döntsön a 180 mellett.

Hamar ráébred, hogy meg kell elégednie egy olyan döntési eljárással, ami olyan, hogy

1. ha az országos átlag 180, akkor a döntési eljárás 1-hez közeli p_0 valószínűséggel a 180 mellett dönt,
2. ha az országos átlag különbözik 180 -tól, akkor a döntési eljárás p_0 -nál kisebb valószínűséggel dönt a 180 mellett.
3. ha az országos átlag nagyon különbözik 180 -tól, akkor a döntési eljárás csak kis valószínűséggel dönt a 180 mellett.

Mindebből kirajzolódik, hogy választani kell egy alkalmas $\Delta\mu$ értéket, és

- ha az N kísérleti eredmény \bar{X}_N átlagának és a hipotetikus $\mu_0 = 180$ értéknek az egymástól való távolsága kisebb $\Delta\mu$ -nél, azaz

$$|\bar{X}_N - \mu_0| < \Delta\mu$$

akkor elfogadjuk a hipotézist, vagyis azt, hogy a plakát igazat mond,

- ha pedig

$$|\bar{X}_N - \mu_0| > \Delta\mu$$

akkor elutasítjuk a hipotézist, vagyis azt mondjuk, hogy a plakát nem mond igazat.

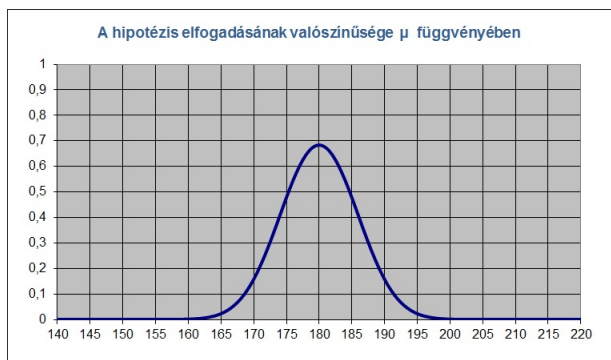
A hipotézis elfogadásának a

$$P(|\bar{X}_N - \mu_0| < \Delta\mu)$$

valószínűsége nyilván a μ paraméter függvénye annak ellenére, hogy a felírt képletben nem jelenik meg a μ betű, de – ne feledjük el – az \bar{X}_N átlag olyan valószínűségi változók átlaga, melyek μ várható értékű normális eloszlást követnek! A függvény képletének felírása céljából említjük, hogy \bar{X}_N normális eloszlást követ μ várható értékkel és $\frac{\sigma}{\sqrt{N}}$ szórással, ezért a hipotézis elfogadásának a valószínűsége:

$$P(|\bar{X}_N - \mu_0| < \Delta\mu) = P(\mu_0 - \Delta\mu < \bar{X}_N < \mu_0 + \Delta\mu) = \Phi\left(\frac{\mu_0 + \Delta\mu - \mu}{\left(\frac{\sigma}{\sqrt{N}}\right)}\right) - \Phi\left(\frac{\mu_0 - \Delta\mu - \mu}{\left(\frac{\sigma}{\sqrt{N}}\right)}\right)$$

Ennek a függvénynek a grafikonját a mellékelt ábrán a $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu = 5$ paraméter értékek mellett megadjuk $140 < \mu < 220$ -re.



53. ábra. Elfogadás valószínűsége $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu = 5$ esetén

Az ábráról leolvassuk, hogy a hipotézis elfogadásának a valószínűsége $\mu = 180$ esetén a legnagyobb. Azt is látjuk, hogy ha μ távolodik μ_0 -tól, akkor a hipotézis elfogadásának a valószínűsége egyre csökken és 0-hoz tart.

A maximális p_0 valószínűség könnyen kifejezhető a σ , N és $\Delta\mu$ paraméterekkel, hiszen ha $\mu = \mu_0$, akkor

$$\Phi\left(\frac{\mu_0 + \Delta\mu - \mu}{\left(\frac{\sigma}{\sqrt{N}}\right)}\right) - \Phi\left(\frac{\mu_0 - \Delta\mu - \mu}{\left(\frac{\sigma}{\sqrt{N}}\right)}\right) = \Phi\left(\frac{\Delta\mu}{\left(\frac{\sigma}{\sqrt{N}}\right)}\right) - \Phi\left(\frac{-\Delta\mu}{\left(\frac{\sigma}{\sqrt{N}}\right)}\right) = 2\Phi\left(\frac{\Delta\mu}{\left(\frac{\sigma}{\sqrt{N}}\right)}\right) - 1$$

Tehát

$$p_0 = 2\Phi\left(\frac{\Delta\mu}{\left(\frac{\sigma}{\sqrt{N}}\right)}\right) - 1$$

Ennek az összefüggésnek az alapján, ha a négy paraméter közül ha három adott, akkor a negyediket meg tudjuk határozni.

Ezen a ponton az internet és a számítógép adta lehetőségeket hívjuk segítségül. Az Olvassó nyissa meg a

http://math.bme.hu/~vetier/2016_osz/A4_vill_2016_osz.html címen a lap felső részében a mellékletek között található

U-próba Elfogadás valószínűsége [linken](#) található

U-proba__Elfogadas_vsz-e.xlsx című Excel fájl,

és a zöld háttérű cellákban található paraméterek változtatásával vizsgálja meg hogyan változik a hipotézis elfogadásának valószínűségének az ábrája. Ebben a nyomtatott jegyzetben az ábrák (sajnos a szövegszerkesztő által nem éppen ideálisan elhelyezve) az alábbi paraméter értékek mellett láthatóak:

1. $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu = 5$
2. $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu = 10$
3. $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu = 2$
4. $\mu_0 = 180$, $\sigma = 10$, $N = 20$, $\Delta\mu = 2$
5. $\mu_0 = 180$, $\sigma = 10$, $N = 100$, $\Delta\mu = 2$

Az első, a második és a harmadik esetben $\mu_0 = 180$, $\sigma = 10$, $N = 4$, és csak a $\Delta\mu$ értéke változik:

- $\Delta\mu = 5$
- $\Delta\mu = 10$
- $\Delta\mu = 2$

A harmadik, a negyedik és az ötödik esetben $\mu_0 = 180$, $\sigma = 10$, $\Delta\mu = 2$, és csak az $N = 4$ értéke változik:

- $N = 4$
- $N = 20$
- $N = 100$

Az Olvasó feladata észrevenni, hogy a N növelésével és $\Delta\mu$ csökkentésével elérhető, hogy a grafikon a $\mu_0 = 180$ helyen megközelítse az 1 értéket, egyéb helyeken pedig a 0 -t.

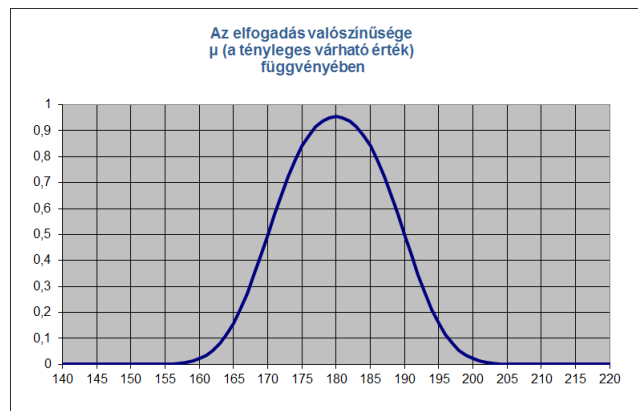
Bár a marslakó naív álmait nem lehet teljesen kielégíteni, az álmokat megközelíteni egészen jól lehet! Ennek persze ára van: a kísérletek számát kell kellően nagyra venni, amit kimondani könnyű, de a gyakorlatban megvalósítani sok esetben nem lehet. Bele kell törődni, hogy a valóság az álmoktól néha messze van. Ilyenkor általában azt szokták csinálni, hogy szubjektív igények alapján először megválasztják a p_0 értéket. A p_0 megválasztásával azt szabályozzuk be, hogy a hipotézis fennállása esetén a döntési eljárásunk ilyen valószínűséggel fog helyesen dönteni, $1 - p_0$ valószínűséggel hibázni fog. Valamiért a $p_0 = 0.95$ a legnépszerűbb érték. Ezek után a $\Delta\mu$ -t választjuk meg a

$$p_0 = 2 \Phi \left(\frac{\Delta\mu}{\left(\frac{\sigma}{\sqrt{N}}\right)} \right) - 1$$

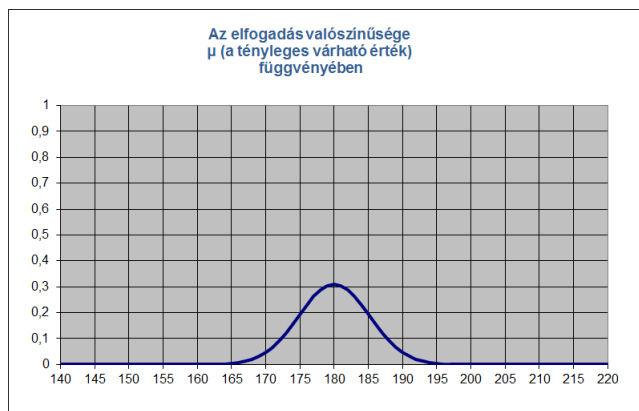
összefüggés segítségével a megválasztott p_0 és az adott σ és N -hez. És kénytelen-kelletlen beletörődünk abba, hogy a hipotézis elfogadásának a valószínűsége a μ tényleges értékétől úgy függ, ahogy azt a fenti linken elérhető

U-proba__Elfogadas_vsz-e.xlsx című Excel fájl

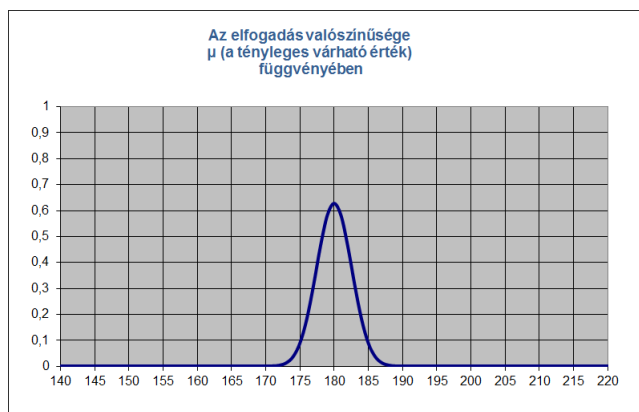
kirajzolja.



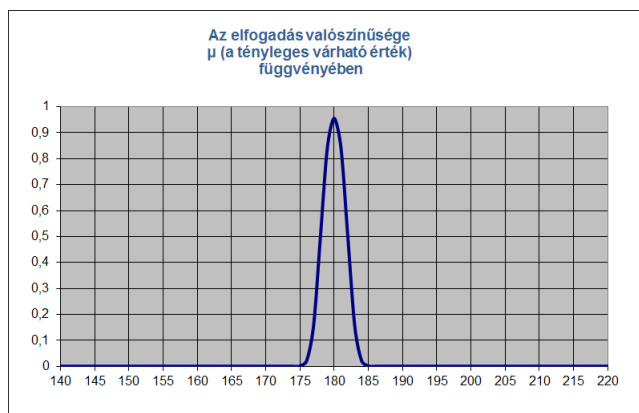
54. ábra. Elfogadás valószínűsége $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu = 10$ esetén



55. ábra. Elfogadás valószínűsége $\mu_0 = 180$, $\sigma = 10$, $N = 4$, $\Delta\mu = 2$ esetén



56. ábra. Elfogadás valószínűsége $\mu_0 = 180$, $\sigma = 10$, $N = 20$, $\Delta\mu = 2$ esetén



57. ábra. Elfogadás valószínűsége $\mu_0 = 180$, $\sigma = 10$, $N = 100$, $\Delta\mu = 2$ esetén

20.3. Tapasztalati eloszlásfüggvény

Ha egy valószínűségi változóra végzünk mondjuk N kísérletet, akkor minden x valós szám esetén megnézhetjük, hogy a kísérleti eredmények alapján mi a relatív gyakorisága annak az eseménynek, hogy a véletlen X érték kisebb mint a szóbanforgó x . Ha a relatív gyakoriságot x függvényében ábrázoljuk, akkor egy olyan monoton növekedő "lépcsős" függvényt kapunk, aminek grafikonja vízszintes szakaszokból áll, és a vízszintes szakaszok közötti ugrások nagysága $\frac{1}{N}$. Ezt a függvényt *tapasztalati eloszlásfüggvénynek* nevezzük.

Ha például a kísérleti eredmények:

0.73 , 24.50 , 10.46 , 9.81 , 22.78

amik sorba rendezve így festenek:

0.73 , 9.81 , 10.46 , 22.78 , 24.50

akkor a tapasztalati eloszlásfüggvény grafikonja úgy fest, ahogy a

A tapasztalati eloszlásfüggvény 5 kísérletből

felíratú ábrán láthatjuk. A rajz elkészítését megkönnyíti, ha az ugrásoknál függőleges szakaszokat húzunk, ahogy

Az ugrásoknál függőleges szakaszok

felíratú ábrán láthatjuk. Végül

A tapasztalati eloszlásfüggvény közelíti az elméleti eloszlásfüggvényt

felíratú ábrán pedig azt a fontos tényt szemléltetjük a $\lambda = 0.1$ paraméterű exponenciális eloszlás példáján, hogy 100 kísérlet esetén a tapasztalati eloszlásfüggvény milyen jól közelíti az elméleti eloszlásfüggvényt.

Az Olvassó megnyithatja a

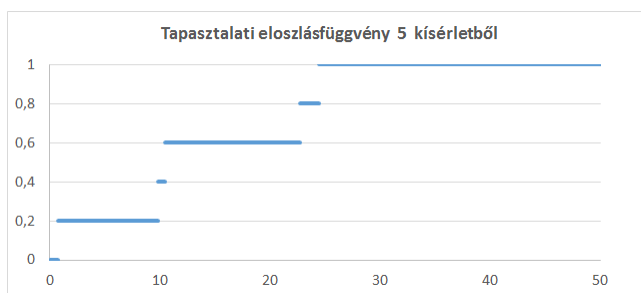
http://math.bme.hu/~vetier/2016_osz/A4_vill_2016_osz.html címen a lap felső részében a mellékletek között található

Tap_el_fv.xlsx című Excel fájlt,

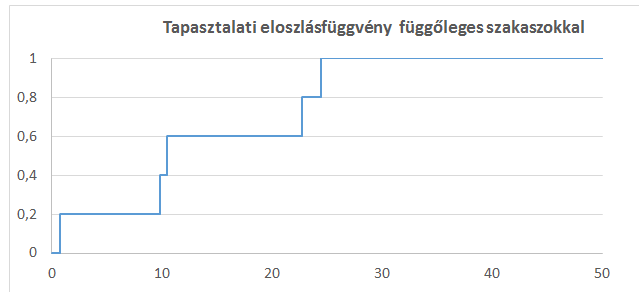
és a számítógép F9 gombjának ismételt nyomogatásával meggyőződhet, hogy a 100 kísérletre épített tapasztalati eloszlásfüggvény mennyire jól közelíti az elméleti eloszlásfüggvényt.

Igaz a következő fontos állítás:

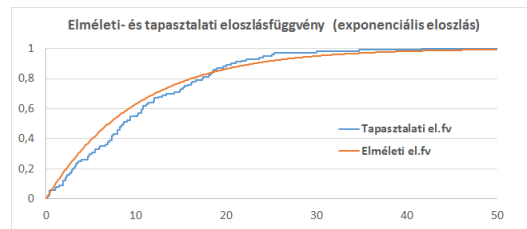
Matematikai statisztika alaptétele: Ha a kísérletek száma tart a végtelenhez, akkor tetszőleges eloszlás esetén a tapasztalati eloszlásfüggvények sorozata 1 valószínűséggel egyenletesen konvergál az elméleti eloszlásfüggvényhez.



58. ábra. A tapasztalati eloszlásfüggvény 5 kísérletből



59. ábra. Az ugrásoknál függőleges szakaszok



60. ábra. A tapasztalati eloszlásfüggvény közelíti az elméleti eloszlásfüggvényt

20.4. Gyakorló feladatok

- 1.
- 2.

21. Folytonos, de nem abszolút folytonos eloszlás a síkon (*Extra tananyag*)

Feladat: Három izzó az áramkörben, de csak kettő érdekel. Tegyük fel, két párhuzamosan kötött – mondjuk – piros és kék izzó mellett, velük sorba kötve, van még egy harmadik, ami – mondjuk – fehér. Tegyük fel, hogy a három élettartam egymástól független, és mindegyik élettartam exponenciális eloszlást követ. Azt is feltesszük, hogy a három izzó egyforma típusú, és ezért várható értékeik egyenlők – mondjuk – 1000 óra várható értékkel.

Az élettartamok függetlensége és azonos eloszlása miatt

- annak a valószínűsége, hogy a fehér lámpa ég ki elsőnek, $1/3$ -dal egyenlő,
- annak a valószínűsége, hogy az egyik színes lámpa ég ki elsőnek, $2/3$ -dal egyenlő.

Jelöljük X -szel azt az időtartamot, amíg a piros lámpa világít, Y -nal azt, amíg a kék lámpa világít, Z -vel azt, amíg a fehér lámpa világít. A piros lámpa két okból fejezheti be fényes működését:

- az egyik az, hogy – puff neki! – kiég, és így nem világít tovább,
- a másik pedig az, hogy a fehér lámpa ég ki, és akkor a soros kötés miatt a piros lámpa se világít tovább.

Ezért nyilván igaz, hogy

$$X = \text{a piros lámpa élettartama és a fehér lámpa élettartama közül a kisebbik}$$

Hasonló okok miatt igaz, hogy

$$Y = \text{a kék lámpa élettartama és a fehér lámpa élettartama közül a kisebbik}$$

Nyilvánvaló, hogy

- ha a piros lámpa ég ki elsőnek, akkor $X < Y$,
- ha a kék lámpa ég ki elsőnek, akkor $X > Y$,
- ha a fehér lámpa ég ki elsőnek, akkor $X = Y$.

Mivel mindegyik lámpa $1/3$ valószínűséggel ég ki elsőnek, az

$$X < Y \qquad X > Y \qquad X = Y$$

események mindegyikének $1/3$ a valószínűsége.

Fókuszáljunk az (X, Y) kétdimenziós valószínűségi változóra, vagyis figyeljük meg mennyi ideig világít a piros és mennyi ideig világít a kék lámpa! A fehér lámpával nem törődünk. Nyilván igaz, hogy (X, Y) minden (x, y) értéket 0 valószínűséggel vesz, ezért (X, Y) folytonos valószínűségi változó.

Viszont ennek az (X, Y) valószínűségi változónak – mindjárt látjuk – a korábban folytonosnak nevezett valószínűségi változókkal szemben van valami furcsasága. A korábbi folytonos valószínűségi változók esetében bármely halmaz valószínűsége a kétdimenziós sűrűségfüggvényből kettősintegrállal volt kiszámítható. Egy nulla területű halmazon vett kettős integrál értéke természetesen nulla, ezért a korábbi folytonos valószínűségi változók esetében bármely olyan esemény valószínűsége, mely nulla területű halmazhoz kötődött, nulla volt. Most viszont az az esemény, hogy (X, Y) a nulla területű $x = y$ egyenletű egyenesre esik, nem nulla: mint fentebb leszögeztük, $1/3$ -dal egyenlő!

És akkor most töredelmesen be kell ismernünk, hogy azokat a síkbeli folytonos eloszlásokat, amilyenekkel eddig foglalkoztunk, és csak egyszerűen folytonosnak neveztünk, igazából *a síkon abszolút folytonos* vagy *síkbeli sűrűségfüggvénnyel definiált* eloszlásoknak is szokás hívni. Tehát – ezzel a terminológiával élve azt kell mondanunk, hogy – eddig mi csak abszolút folytonos síkbeli eloszlásokkal foglalkoztunk, de a rövideg kedvéért az "abszolút" jelzőt elhagytuk. Eme pongyolaság utólagos magyarázatául még az alábbi érveket hozzuk fel:

- a "folytonos" jelzőt a "diszkrét" ellenpólusaként használtuk,
- amiket folytonosnak neveztünk, azok tényleg folytonosak voltak, sőt mi több,
- abszolút folytonosak is, de a kezdő diákat nem akartuk az "abszolút" szó – talán akkor még nehezebben emészthető – jelentésével terhelni.

Ezzel a most bevezetett terminológiával élve azt mondhatjuk, hogy az előbb – lámpák segítségével – definiált (X, Y) kétdimenziós valószínűségi változó eloszlása a síkon – bár folytonos –, de nem abszolút folytonos,