Villamosmérnök A4 — 12. hét
Kétdimenziós normális eloszlás, cht

Kétdimenziós normális összefoglalás

Egy kétdimenziós (X,Y) valószínűségi változó **kovariancia mátrixa**:

$$\Sigma = \begin{pmatrix} \mathbb{V}\mathrm{ar}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y) & \mathbb{V}\mathrm{ar}(Y) \end{pmatrix}$$

Korrelációs együttható: $\mathrm{r}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\mathbb{SD}(X)\mathbb{SD}(Y)}$.

Kétdimenziós normális eloszlás: Standard kétdimenziós normális eloszlású egy (U,V) pár, ha sűrűségfüggvénye

$$\varphi(u,v) := \frac{1}{2\pi}\exp\left(-\frac{u^2+v^2}{2}\right) \qquad (u,v\in\mathbb{R}).$$

(U,V) zérus várható érték vektorú, egységmátrix kovarianciájú pár. Azt mondjuk, hogy az $(X,Y) = (U,V)\mathbf{A}+\mu$ pár kétdimenziós normális eloszlású $\mu\in\mathbb{R}^2$ várható érték vektorral és $\Sigma\in\mathbb{R}^{2\times 2}$ (invertálható) kovariancia mátrixszal, ha (U,V) standard normális pár, és $\mathbf{A}^\top\mathbf{A} = \Sigma$, ahol $\mathbf{A}\in\mathbb{R}^{2\times 2}$.
Egy kétdimenziós normális (X,Y) pár sűrűségfüggvénye:

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2r\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\},$$

ahol $\mathbb{E}X = \mu_X$, $\mathbb{E}Y = \mu_Y$, $\mathbb{SD}(X) = \sigma_X$, $\mathbb{SD}(Y) = \sigma_Y$, és $r = r(X,Y)$ X és Y korrelációs egyyüthatója. Kétdimenziós normális eloszlás esetében a perem- illetve feltételes eloszlások is normálisok: a marginálisok eloszlása $N(\mu_X,\sigma_X)$ és $N(\mu_Y,\sigma_Y)$, az $X|Y=y$ (a másik szimmetrikusan hasonlóan) feltételes eloszlás pedig $N(\mu_{X|Y=y},\sigma_{X|Y=y})$, ahol

$$\mu_{X|Y=y} = \mu_X + (y-\mu_Y)\,r\,\frac{\sigma_X}{\sigma_Y} \qquad \text{és} \qquad \sigma_{X|Y=y} = \sqrt{1-r^2}\,\sigma_X\,.$$

Kétdimenziós normális feladatok

1. Tegyük fel, hogy egy jólmenő étterem heti összbevétele normális eloszlást követ 1 millió forint várható haszonnal, és 700000 forint szórással. Mi annak a valószínűsége, hogy kevesebb, mint 1.5 millió forint a bevétele két, egymást követő héten? Itt tegyünk még fel függetlenséget! Majd nézzük meg, hogyan változik annak a valószínűsége, hogy a második héten több mint 2 millió forint a bevétel feltéve, hogy az első héten 1.5 millió forint volt a bevétel és a korreláció -0.5! Mi a második hét várható bevétele ugyanezen feltétel mellett? Mennyi a két hét várható összbevétele? Mi a szórás?

2. Budapesten májusban az átlagos hőmérséklet 25 °C, 7 °C szórással, valamint az átlagnyomás 10^5 Pa, 2×10^4 Pa szórással. A hőmérséklet/nyomás változása szoros összhangban van, köztük lévő korreláció 0.7. Írjuk fel a kovariancia mátrixot majd határozzuk meg a következőket:
   (a) Mi a valószínűsége annak, hogy egy nap melegebb lesz, mint 40 °C? És, hogy alacsonyabb a nyomás 6×10^4 Pa-nál?
   (b) Egy nap 20 °C-ot mértünk. Mi annak a valószínűsége, hogy a légnyomás 1.2×10^5 Pa fölött járt? Átlagosan mekkora volt a légnyomás? Mekkora a szórás?
   (c) Feltéve, hogy egy nap 10^5 Pa volt a légnyomás, mi annak a valószínűsége, hogy melegebb volt, mint 35 °C? Átlagosan hány fok volt aznap? Mekkora a szórás?

3. Magyarországon a felnőtt férfiak testmagassága átlagosan 178 cm, 9 cm szórással, míg testsúlyuk 85 kg, 10 kg szórással. A korrelációs együttható 0.7, azaz a magasabb emberek általában súlyosabbak is. Írjuk fel a kovariancia mátrixot!
   (a) Mi a valószínűsége annak, hogy egy férfi magasabb 2 méternél? És, hogy nehezebb 100 kg-nál?
   (b) Feltéve, hogy egy férfi 80 kg, mi annak a valószínűsége, hogy magasabb, mint 180 cm? Várhatóan hány cm magas egy ilyen férfi? Mekkora a szórás?
   (c) Átlagosan mekkora súlyú egy 190 cm magas férfi?
   (d) Átlagosan milyen magas egy 94.3 kg-os férfi?
   (e) Hasonlítsuk össze az utolsó két eredményt.

4. Az X áramerősség normális $N(230,6)$ eloszlású, a mérőkészülék Z hibája ettől független $N(0,8)$ eloszlású, mi az $Y = X+Z$ értéket mérjük. Mi a valószínűsége, hogy $Z > X/20$?

5. Legyen a (X,Y) pár kétdimenziós normális eloszlású, r korrelációval. Mi az eloszlása $U = X+Y$-nak és $V = X-Y$-nak? Független-e U a V-től? Számoljuk ki a várható értékeket és szórásokat is!

6. Hogyan generálna le kétdimenziós normális eloszlású véletlen pontokat a síkon, melyek várható értéke (μ_1, μ_2), szórása (σ_1, σ_2), korrelációs együtthatója pedig r. Függetlenek-e a koordináták, ha $r = 0$?

CHT

Centrális határeloszlás-tétel (CHT) (Akinek ez elsőre nehéznek tűnik nyugodtan ugorjon a "mindez szóban" részhez): legyenek $X_1, X_2, \dots$ független, azonos eloszlású valószínűségi változók, $\mu := \mathbb{E}X_i \in \mathbb{R}$ és $\sigma := \mathbb{SD}(X_i) \in \mathbb{R}^+$. Ekkor

$$\mathbb{P}\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} < x\right) \longrightarrow \Phi(x), \text{ amint } n \to +\infty, \ (x \in \mathbb{R}).$$

Mindez szóban: elég nagy n esetén a – FAE – valószínűségi változók összege normális eloszlással közelíthető. A fenti képlet magyarázata kicsit pontosabban: a standardizált összeg közelítőleg (standard) normális eloszlású. Speciálisan: ha az X_i változókat azonos, p paraméterű Bernoulli változóknak választjuk, akkor jutunk el a de Moivre–Laplace tételhez (avagy binomiális CHT).

7. Mennyi annak a valószínűsége, hogy 12 000 kockadobás során előforduló hatosok száma 1900 és 2150 közé esik?

8. Egy gyár két fajta érmét gyárt: egy igazságosat, és egy hamisat ami 55% eséllyel mutat fejet. Van egy ilyen érménk, de nem tudjuk igazságos-e vagy pedig hamis. Ennek eldöntésére a következő statisztikai tesztet hajtjuk végre: Feldobjuk az érmét 1000-szer. Ha legalább 525-ször fejet mutat, akkor hamisnak nyilvánítjuk, ha 525-nél kevesebb fej lesz a dobások között, akkor az érmét igazságosnak tekintjük. Mi a valószínűsége, hogy a tesztünk téved abban az esetben, ha az érme igazságos volt? És ha hamis volt?

9. Határozzuk meg azt a k egész számot, amelyre igaz, hogy annak a valószínűsége, hogy 1000 érmedobás során a fejek száma $500 - k$ és $500 + k$ közé esik, kb. 0.9.

10. Mennyi a valószínűsége annak, hogy 50 darab független és azonos eloszlású valószínűségi változó összege a [0, 30] intervallumba esik, ha egy ilyen változó eloszlása a [0, 1] intervallumon

    (a) egyenletes;

    (b) $f(x) = 2x$ sűrűségfüggvény szerint alakul?

11. Becsüljük meg annak valószínűségét, hogy 10 000 kockadobás összege 34 800 és 35 200 közé esik.

12. Egy kockát folyamatosan feldobunk addig, amíg a dobások összege meghaladja a 300-at. Becsüljük meg annak valószínűségét, hogy legalább 80 dobásra van ehhez szükség.

13. Adott 100 égőnk, melyek élettartama egymástól független exponenciális eloszlású, 5 óra várható értékkel. Tegyük fel, hogy az égőket egymás után használjuk, azonnal kicserélve azt, amelyik kiégett. Becsüljük meg annak valószínűségét, hogy 525 óra után még van működő égőnk.

14. A jegyiroda előtt a fiatalok hosszú sorban állnak egy koncertjegyért. Ebben a pillanatban éppen 18-an állnak az egyik pénztár előtt. Megfigyeltem, hogy egy vásárló kiszolgálási ideje memória nélküli (más néven örökifjú) valószínűségi változó 3 perc átlaggal és a kiszolgálási idők függetlenek. Becsülje meg annak a valószínűségét, hogy a most utolsóként álló fiatal több mint 60 percet fog a pénztár előtt eltölteni!

15. Egy szabályos érmét 40-szer feldobunk, és X-szel jelöljük a kapott fejek számát. Határozzuk meg annak valószínűségét, hogy $X = 20$

    - a binomiális eloszlás segítségével,
    - a deMoivre–Laplace-tételt használva. Ez utóbbihoz segítség: $\mathbb{P}\{X = 20\} = \mathbb{P}\{19.5 \le X < 20.5\}$, ami persze nem számít amíg X-et binomiálisnak (azaz egész értékűnek) tekintjük, de fontos lesz a deMoivre–Laplace-tétel alkalmazásánál.