Kiegyenlítő számítások: Valószínűségelméleti alapok

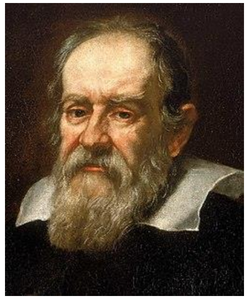
Barsi Árpád

A modell

- A geodéziai mérések célja: a Föld alakjának, nehézségi erőterének, a felszíni alakzatok helyzetének, méretének meghatározása és ábrázolása
- Bonyolult folyamatok és rendszerek (a méréseknél is) kezelése → modell
- Definíció: „fizikai rendszerek, folyamatok egyszerűsített, matematikailag szabatosan leírható idealizált mása, amely annak geometriai, fizikai stb. tulajdonságait szemlélteti”
- Fajtái:
  - Funkcionális: determinisztikus összefüggések
  - Sztochasztikus: véletlenre vonatkozó összefüggések

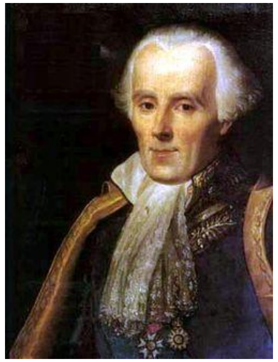
Egy kis történelem…

- Gerolamo Cardano (1501 – 1576)
  - csillagjós, orvos, 131 könyv + 170 elégetett kézirat + 111 hagyaték
  - „A szerencsejátékról” és „A kockajátékról” $\rightarrow$ bizonytalanság, eseménytér
- Galileo Galilei (1564 – 1642)
  - „Gondolatok a kockajátékról”
- Blaise Pascal (1623 – 1662)
  - Orvosi tanács, „osztozkodás problémája”, levelezés Pierre de Fermat-val
  - Várható érték, rulettkerék



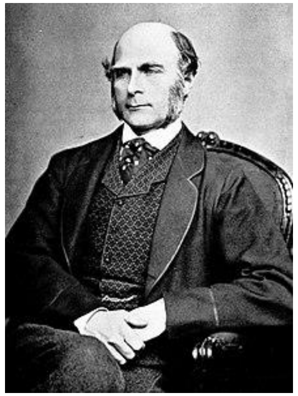
Még egy kis történelem

- Jakob Bernoulli (1654 – 1705)
  - Valószínűség meghatározása, a priori valószínűség, „nagy számok törvénye”
- Thomas Bayes (1702 – 1761)
  - Feltételes valószínűség („Tanulmány a valószínűség tana egyik problémájának megoldásáról”)
- Pierre-Simon de Laplace (1749 – 1827)
  - Mérések elemzése, majd cikk; 1812 és 1814 „A valószínűség analítikus elmélete”
- Antoine Lavoisier (1743 – 1794)
  - méréstechnika



S végül ne maradjon ki

- Jules Henri Poincaré (1854 – 1912)



- Francis Galton (1822 – 1911)
- John Graunt (1620 – 1674) és William Petty (1623 – 1687)
  - „A statisztikai atyjai”

- Carl Friedrich Gauss (1777 – 1855)

Alapfogalmak

- Véletlen (*random*):
  - a köznyelvben a meghatározhatatlan, megjósolhatatlan vagy szabálytalan dolgok
  - a filozófiában a nem-determinisztikus, vagyis nem-elrendelt esemény
  - a tudományban ha az esemény kimenete nem jósolható meg vagy az nagyon bonyolult
- Véletlen tömegjelenség: adott körülmények között végbemenő folyamat, ami sokszor megismételhető
- Valószínűségszámítás (valószínűségelmélet) tárgya – (*Probability theory*)

Alapfogalmak

- Elemi esemény (ω): egy jelenség kimenetele
- Esemény: egy konkrét elemi esemény, jelölésük: A, B, …
- Eseménytér (Ω): elemi események összessége
- Összetett esemény: elemi események bizonyos összessége, az eseménytér részhalmaza

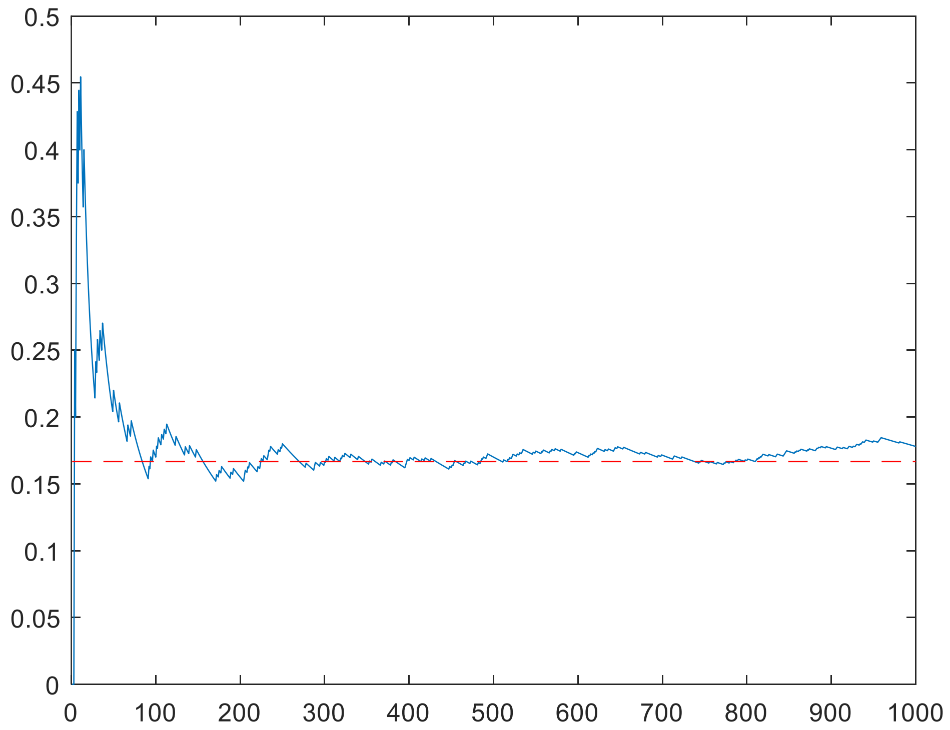
Műveletek az eseményekkel

- Események, mint halmazok
- Esemény ellentéte: $\overline{A}$
- Események összege: $A + B$
- Események szorzata: $A \cdot B$
- Kizáró események: $A \cdot B = 0$
- Események különbsége: $A - B = A \cdot \overline{B}$
- Események összessége: $\sum_i A_i = 1$
- Eseményalgebra

Relatív gyakoriság

- Az összes esemény bekövetkezésének száma: n
- A esemény bekövetkezésének száma: k_A
- Relatív gyakoriság: $\dfrac{k_A}{n}$
- Erre igaz: $0 \leq \dfrac{k_A}{n} \leq 1$
- A valószínűség (*probability*):

$$0 \leq P(A) \leq 1$$



A valószínűség néhány tulajdonsága

$$\frac{k_{A+B}}{n} = \frac{k_A + k_B}{n} = \frac{k_A}{n} + \frac{k_B}{n}$$

$$P(A+B) = P(A) + P(B)$$

$$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n) = \sum_{i=1}^{n} P(A_i)$$

$$P(\Omega) = 1 \qquad 0 \le P(A) \le 1 \qquad P(\Omega) = P(A) + P(\overline{A}) \qquad P(\overline{A}) = 1 - P(A)$$

- Kolmogorov axiomái, valamint néhány tétel

Feltételes valószínűség

- Definíció: $$P(A|B) = \frac{P(AB)}{P(B)} = \frac{\frac{k_{AB}}{n}}{\frac{k_B}{n}}$$

- Továbbá igaz

$$0 \le P(A|B) \le 1$$

$$P(B|B) = 1$$

$$P(A|B) = \frac{P(AB)}{P(B)} \neq P(B|A) = \frac{P(BA)}{P(A)} = \frac{P(AB)}{P(A)}$$

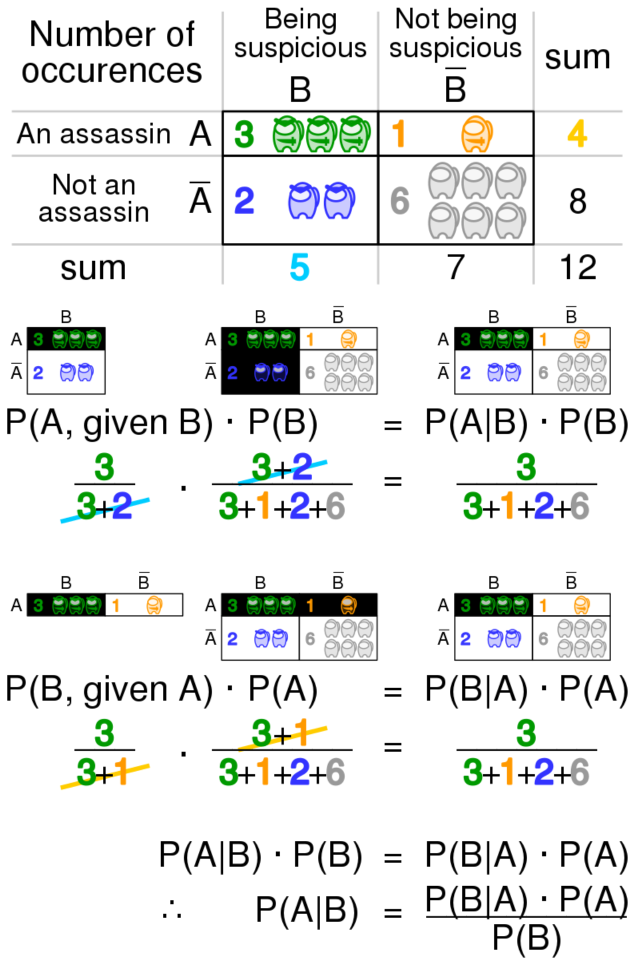
Valószínűségi tételek

- Teljes valószínűség tétele

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

- Bayes-tétel

$$P(B_k|A) = \frac{P(A|B_k) \cdot P(B_k)}{\sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)}$$



Gyakorlati példa következik!

Bayes-tétel és gyakorlati példa (thx ChatGPT)

- A Bayes-tétel egy matematikai képlet, amely segít megérteni, hogyan változik valaminek a valószínűsége, ha új információ áll rendelkezésre. Képzeld el, hogy detektívként dolgozol, és megpróbálod kideríteni, ki követte el a bűncselekményt. Kezdetben minden gyanúsítottról van egy előzetes véleményed vagy "előzetes valószínűség" arról, hogy ő lehet-e a tettes. Amikor új bizonyítékot találsz, a Bayes-tétel segítségével frissítheted a gyanúsítottakkal kapcsolatos véleményedet, hogy jobban megértsd, ki lehet a valódi bűnös.

- Tehát, nagyon egyszerűen szólva, a Bayes-tétel segít nekünk "frissíteni" a valószínűségeket azáltal, hogy figyelembe veszi az új információkat. Ez olyan, mint amikor több nyomot találsz, és ezek alapján finomítod a gyanúsítottaid listáját.

A példa konkrét számokkal

Tegyük fel, hogy van egy város, ahol három személy gyanúsítható egy bűncselekmény elkövetésével: Anna, Béla, és Csaba. Kezdetben egyenlően valószínűnek tartjuk, hogy bármelyikük lehet a tettes, tehát mindenki esélye 1/3, vagyis körülbelül 33.3%.

Most kiderül egy új információ: a bűncselekmény elkövetője viselt kesztyűt. A rendelkezésre álló adatok alapján tudjuk, hogy:

- Anna 80%-ban hord kesztyűt hideg időben.
- Béla csak 30%-ban hord kesztyűt, mert nem szereti, ha korlátozzák a kezét.
- Csaba 50%-ban hord kesztyűt, mert néha elfelejti otthon.

A bűncselekmény napján hideg volt, tehát a kesztyű viselése releváns információ.

A Bayes-tétel segítségével frissíthetjük a gyanúsítottak esélyeit. Ehhez először kiszámoljuk az úgynevezett "teljes valószínűségét" annak, hogy a bűncselekmény elkövetője kesztyűt viselt, minden gyanúsított esetében figyelembe véve az előzetes valószínűségeket és az új információt:

$$P(\text{Kesztyű}) = P(\text{Anna}) \times P(\text{Kesztyű}|\text{Anna}) + P(\text{Béla}) \times P(\text{Kesztyű}|\text{Béla}) + P(\text{Csaba}) \times P(\text{Kesztyű}|\text{Csaba})$$

$$P(\text{Kesztyű}) = \frac{1}{3} \times 0.8 + \frac{1}{3} \times 0.3 + \frac{1}{3} \times 0.5$$

Ezután kiszámoljuk az új valószínűségeket, hogy ki lehetett a tettes, figyelembe véve, hogy kesztyűt viseltek:

$$P(\text{Anna}|\text{Kesztyű}) = \frac{P(\text{Kesztyű}|\text{Anna}) \times P(\text{Anna})}{P(\text{Kesztyű})}$$

$$P(\text{Béla}|\text{Kesztyű}) = \frac{P(\text{Kesztyű}|\text{Béla}) \times P(\text{Béla})}{P(\text{Kesztyű})}$$

$$P(\text{Csaba}|\text{Kesztyű}) = \frac{P(\text{Kesztyű}|\text{Csaba}) \times P(\text{Csaba})}{P(\text{Kesztyű})}$$

Vegyük a számításokat, hogy lássuk, hogyan változnak a valószínűségek.

Az új információ fényében a gyanúsítottakkal kapcsolatos valószínűségek a következőképpen frissültek:

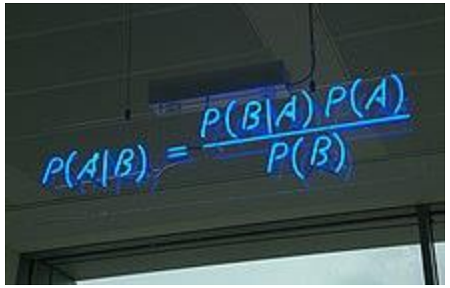
- A teljes valószínűsége annak, hogy a bűncselekmény elkövetője kesztyűt viselt, 53.33%.
- Anna valószínűsége, mint a bűncselekmény elkövetője, figyelembe véve, hogy kesztyűt viselt, 50% lett.
- Béla valószínűsége, mint a bűncselekmény elkövetője, 18.75% lett.
- Csaba valószínűsége, mint a bűncselekmény elkövetője, 31.25% lett.

Ez azt jelenti, hogy az új bizonyíték (kesztyű viselése) alapján Anna valószínűsége, hogy ő a tettes, a legmagasabb, míg Béla esélye a legalacsonyabb lett. [>-]

Egy sheffieldi lelkész

- Thomas Bayes (1702-1761)
  - Két művet publikált életében
    - Teológiai (1731)
    - Matematikai (1736, névtelenül $\rightarrow$ Newton védelme és Royal Society tagja 1742)
  - Tételét és igazi munkásságát a jegyzeteiből Richard Price adja ki halála után 1763-ban

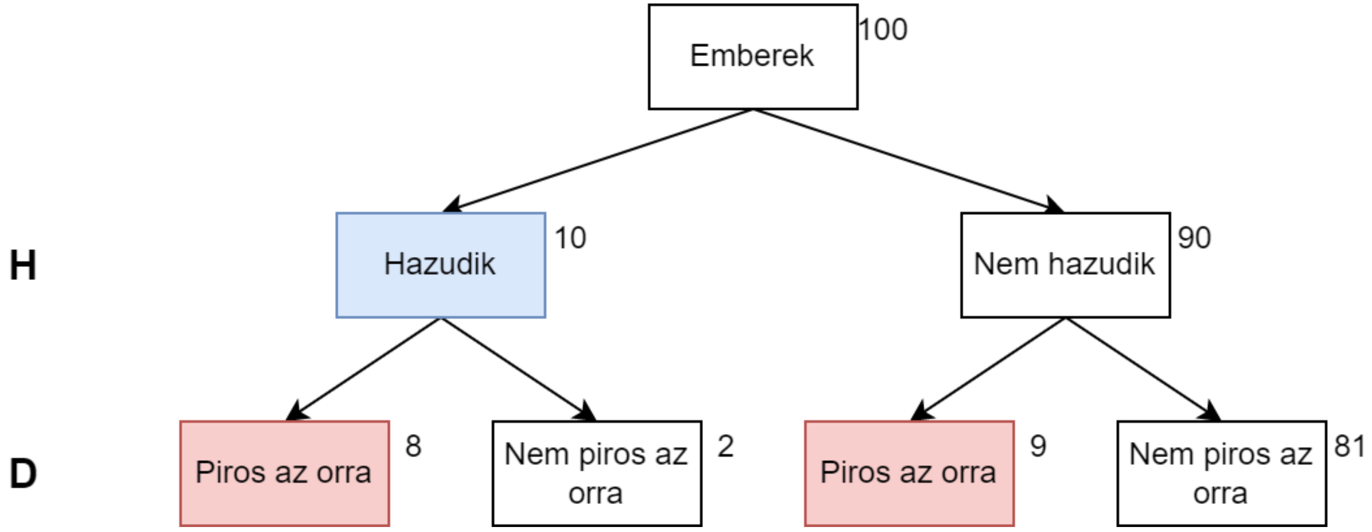
Bayes-tétel alkalmazása



- (*reasoning*)
- Egy faluban 100 emberből 10 hazudik (H). A 10 hazudósból 8-nak piros az orra (D). A 90 nemhazudósból is 9-nek piros az orra. Ha meglátogatjuk a falut és piros orrúval találkozunk, mennyi az esélye, hogy hazudik?

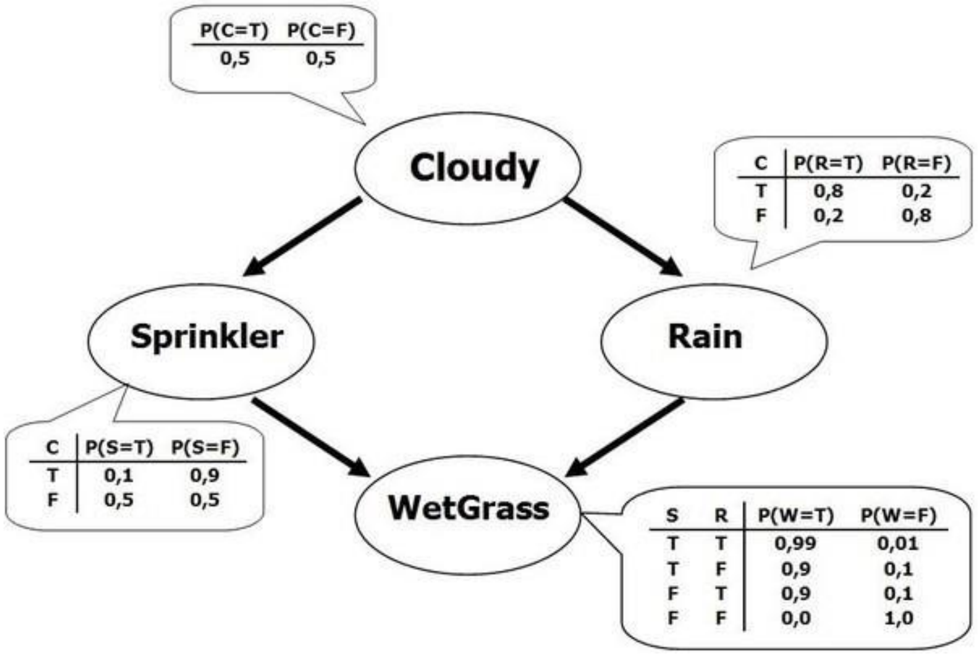
$$p(H|D) = \frac{p(H) \cdot p(D|H)}{p(D)} = \frac{\frac{10}{100} \cdot \frac{8}{10}}{\frac{17}{100}} = 47\%$$

$$p(H|D) = \frac{8}{9+8} = 47\%$$



Még egyszer a Bayes-tétel alkalmazása

- Bayes-háló (valószínűségi háló)



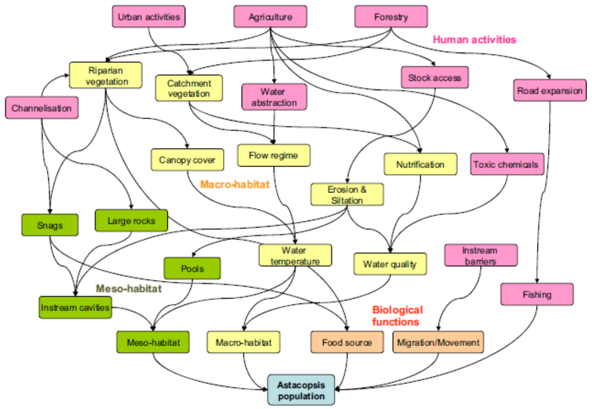
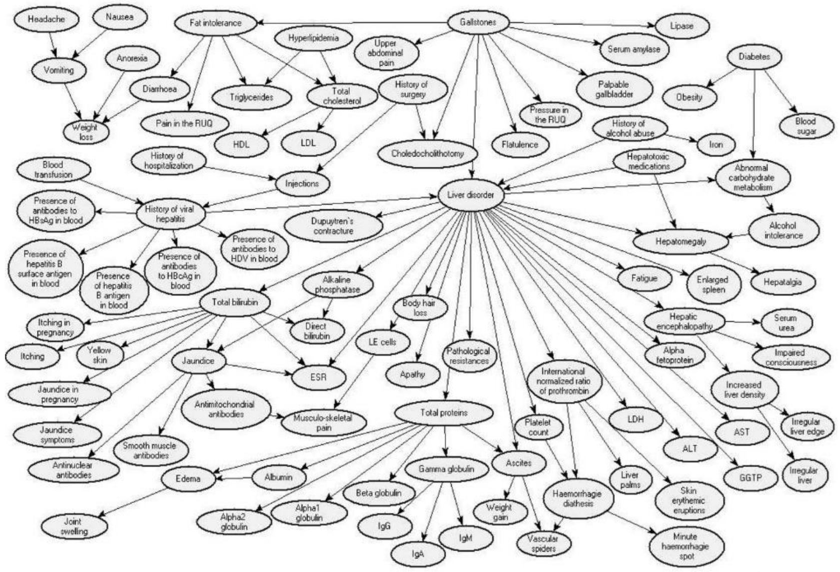


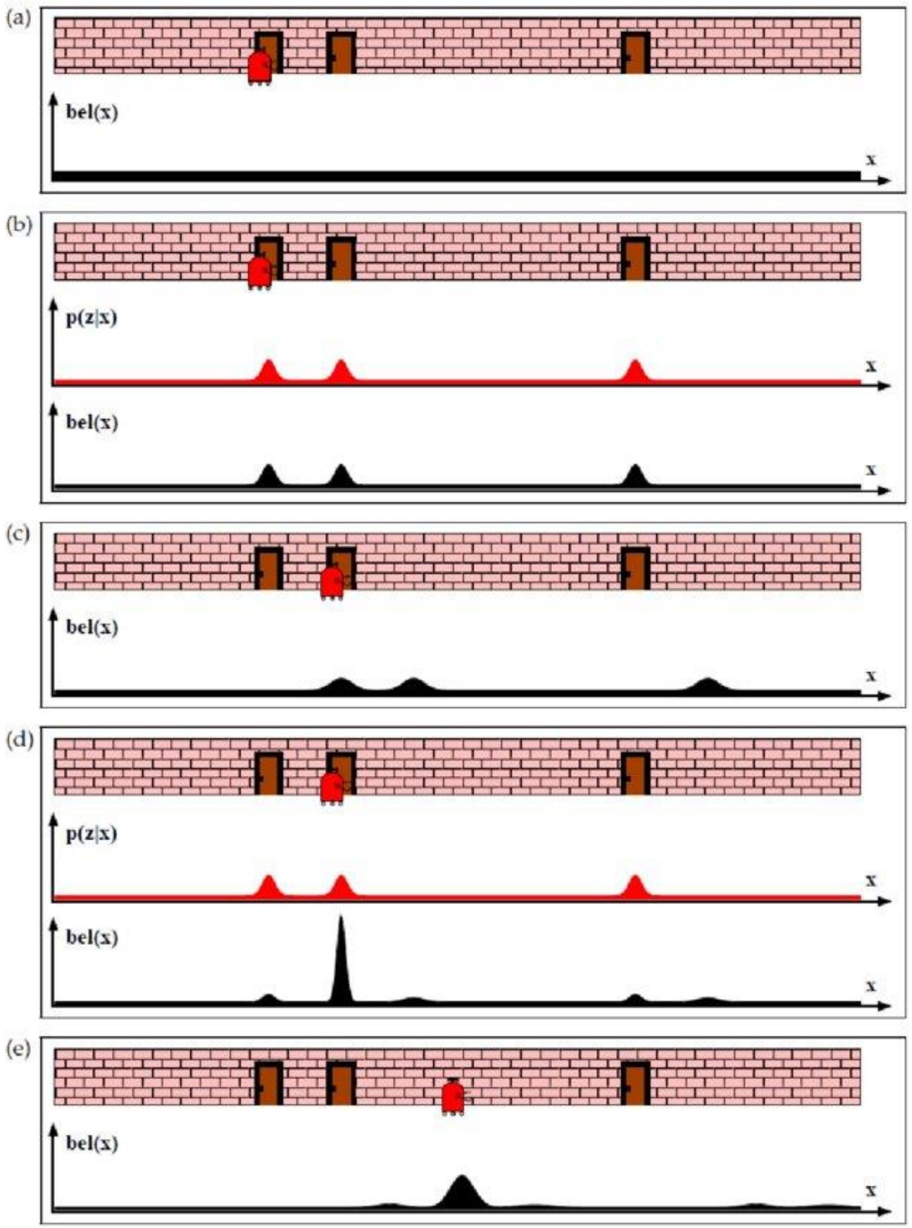
Fig. 2. Conceptual model of the key factors affecting *Astacopsis gouldi* populations.



- Alkalmazások: robotika, (műszaki, orvosi…) diagnosztika, kockázatelemzés…

Utolsó példa

- Robot lokalizáció



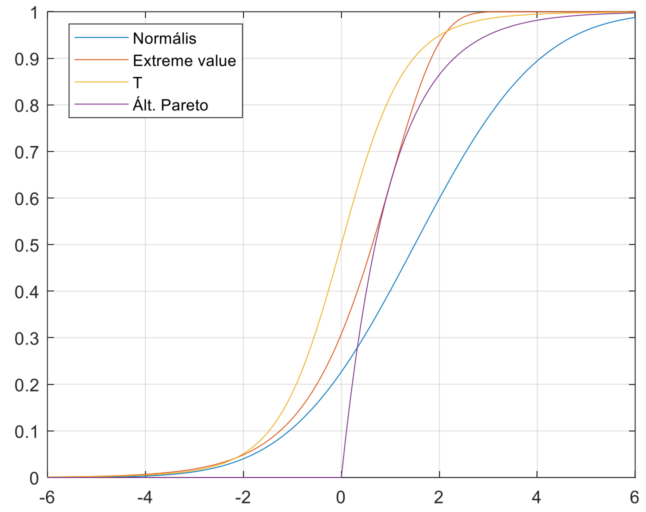
Valószínűségi változó

- *Random variable*
- Definíció: az elemi események kimeneteléhez rendelt valós számérték
- Jelölése: ξ
- Fajtái:
  - Folytonos
  - Diszkrét
- Valószínűségi vektorváltozó: valószínűségi változók vektora
- Jelölése: $\boldsymbol{\xi} = [\xi_1, \xi_2, \ldots, \xi_n]$

Valószínűségi változók eloszlása

- Eloszlásfüggvény – *cumulative distribution function* (CDF)

$$F(x) = P(\xi < x)$$



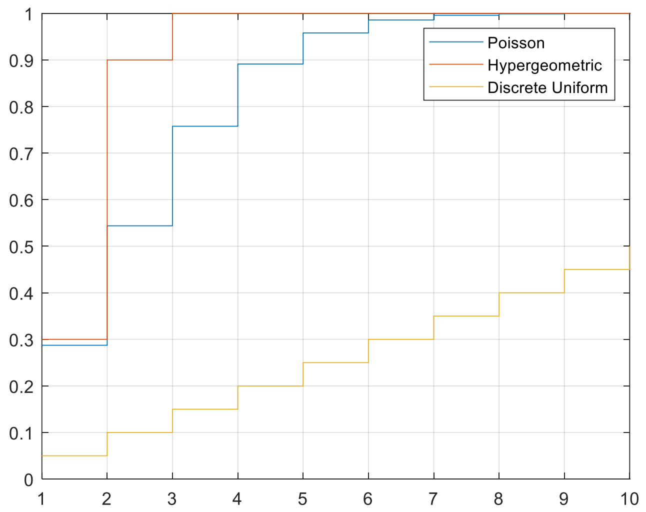
- Jellemzői:

$$0 \le F(x) \le 1$$

$$F(a) \le F(b) \quad ha\ a < b$$

$$\lim_{x \to -\infty} F(x) = 0 \qquad \lim_{x \to +\infty} F(x) = 1$$

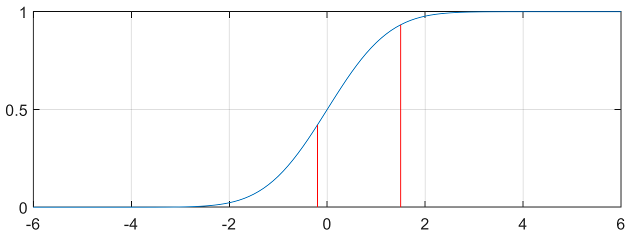
$$P(a \le \xi < b) = F(b) - F(a)$$



Valószínűségi sűrűségfüggvény

- *probability density function* (PDF)

$$\int_{-\infty}^{x} f(t)dt = F(x)$$

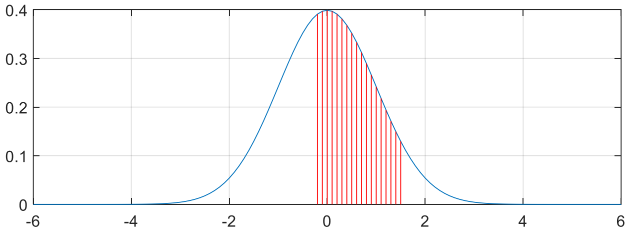


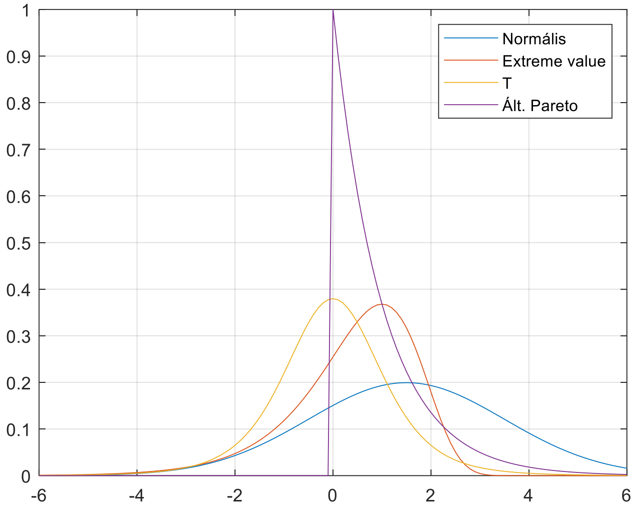
- Jellemzői:

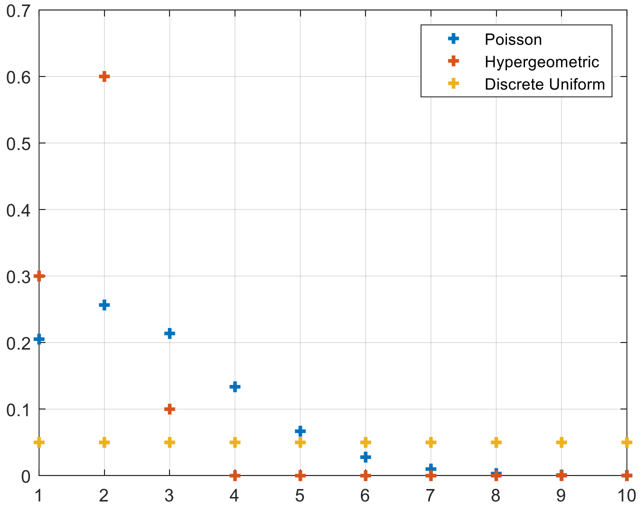
$$f(x) = F'(x)$$

$$f(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(t)dt = 1$$







$$P(a \leq \xi < b) = F(b) - F(a) = \int_{a}^{b} f(t)dt$$

A valószínűségi változó jellemzői

- Várható érték: a lehetséges értékek valószínűséggel súlyozott közepe

$$M(\xi) = \sum_i x_i p_i$$

- Szórásnégyzet: a várható értéktől való négyzetes eltérés várható értéke

$$D^2(\xi) = M\left([\xi - M(\xi)]^2\right)$$

- Jellemzőik:

$$M(c) = c \qquad M(c\xi) = cM(\xi) \qquad M(\xi_1 + \xi_2) = M(\xi_1) + M(\xi_2) \qquad M(\xi_1\xi_2) = M(\xi_1) \cdot M(\xi_2)$$

$$D^2(\xi) = M(\xi^2) - M^2(\xi)$$

$$D(c) = 0 \qquad D(c\xi) = |c| \cdot D(\xi) \qquad D^2(\xi_1 + \xi_2) = D^2(\xi_1) + D^2(\xi_2)$$

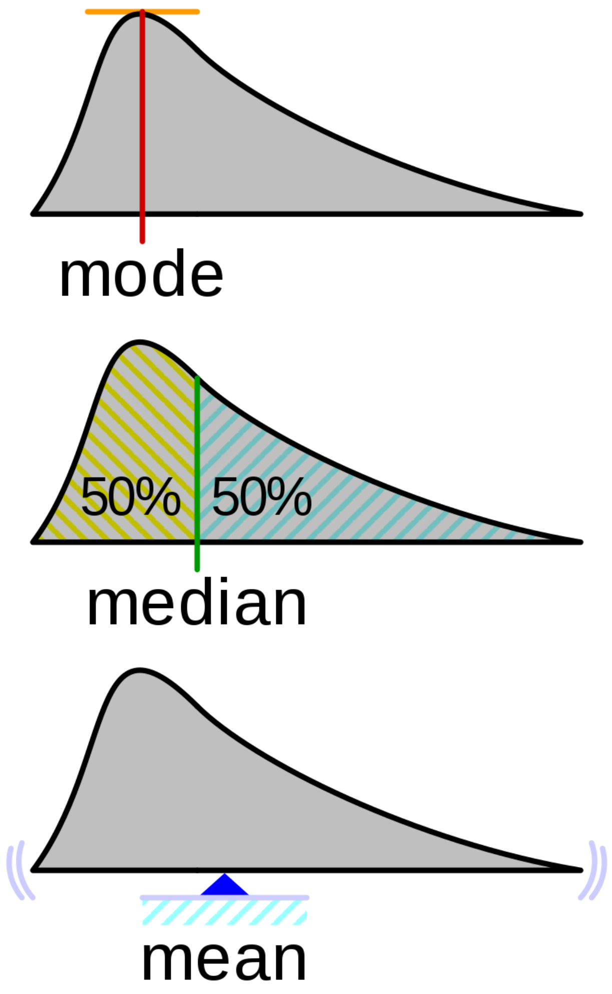
A valószínűségi változó jellemzői

- Medián

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

- Módusz

$$f(m_o) = \max(f(x))$$



A normális eloszlás

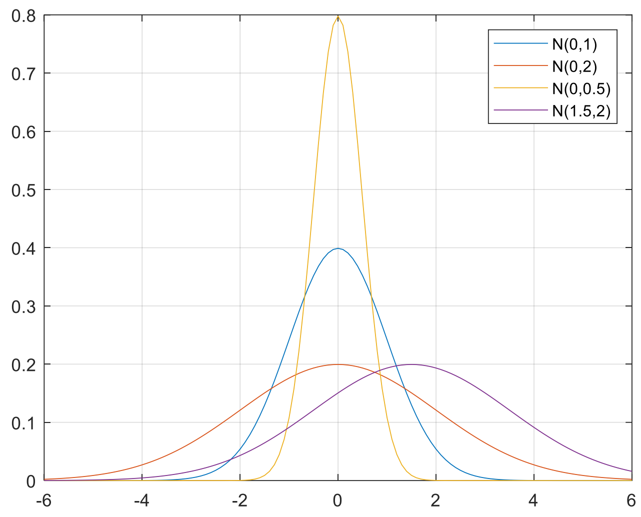
$$N(m,\sigma)$$

- Valószínűségi sűrűségfüggvény

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

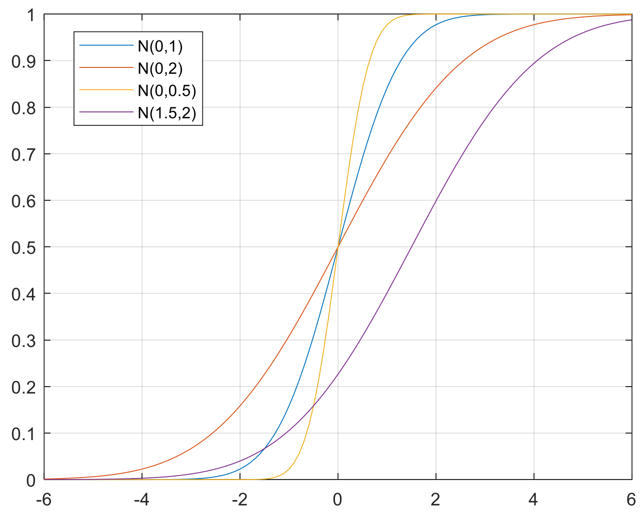
$$M(\xi)=m$$

$$D(\xi)=\sigma \qquad \sigma\neq 0$$



- „haranggörbe”
- Eloszlásfüggvény

$$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{(t-m)^2}{2\sigma^2}}\,dt$$



Standardizált normális eloszlás $N(0{,}1)$

- Valószínűségi változó standardizálása

$$M(\xi) = m$$

$$D(\xi) = \sigma$$

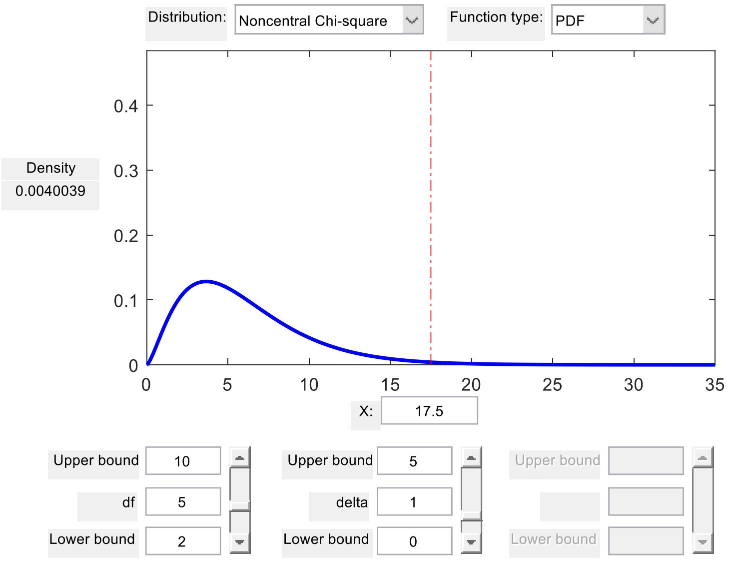
$$\xi' = \frac{\xi - M(\xi)}{D(\xi)}$$

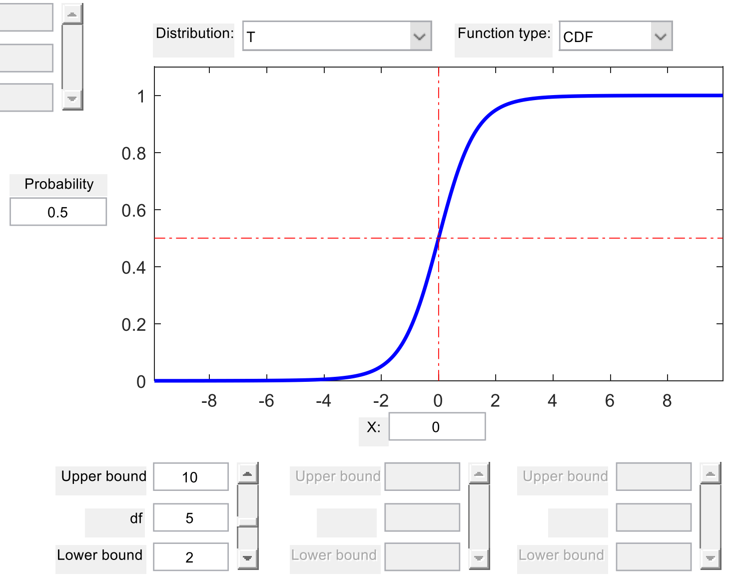
$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

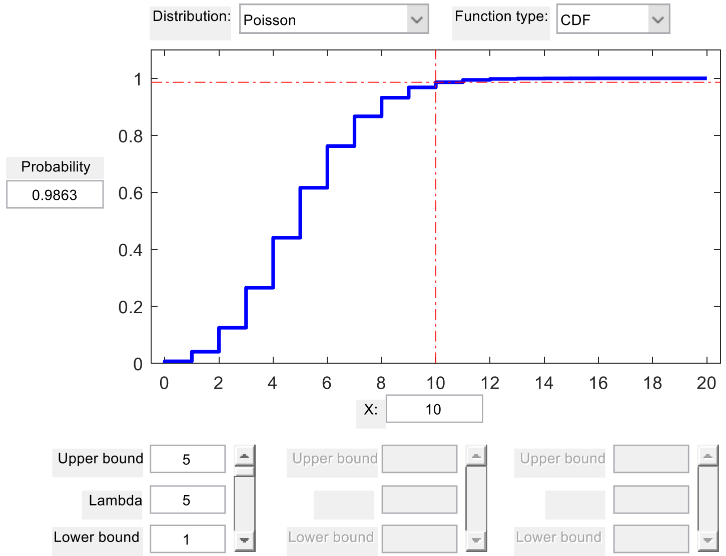
$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

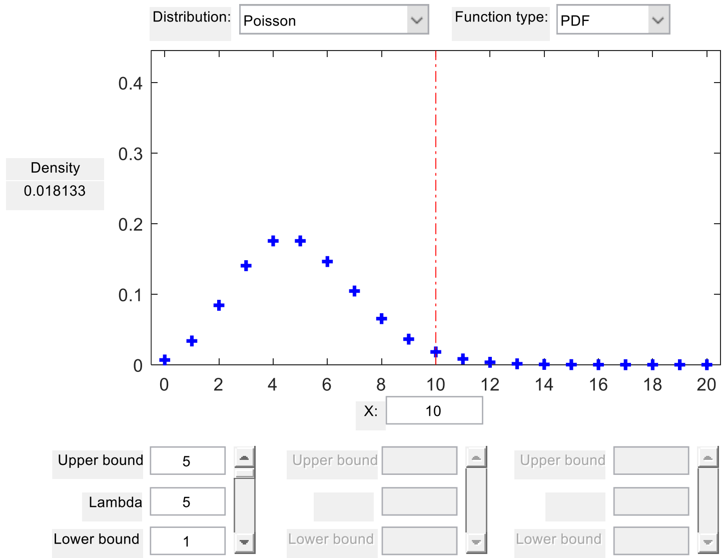
További eloszlások

- Matlab disttool eszköze









Több valószínűségi változó

- Valószínűségi vektorváltozó $\boldsymbol{\xi} = [\xi_1, \xi_2, \ldots, \xi_n]$
- Várható érték-vektor

$$m_i = M(\xi_i) \qquad \mathbf{m} = M(\boldsymbol{\xi})$$

- Szórásnégyzet-vektor

$$\sigma_i^2 = D^2(\xi_i) = M\left([\xi_i - M(\xi_i)]^2\right)$$

- Általánosítás: kovariancia

$$c_{ij} = M\left([\xi_i - M(\xi_i)][\xi_j - M(\xi_j)]\right)$$

$$\underset{(n,n)}{\mathbf{C}} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

A kovarianciáról még

- Jellemzők:

$$c_{ij} = M\left(\left[\xi_i - M\left(\xi_i\right)\right]\left[\xi_j - M\left(\xi_j\right)\right]\right)$$

$$c_{ii} = \sigma_i^2 \qquad c_{ij} = c_{ji}$$

- Korreláció:

$$r_{ij} = \frac{c_{ij}}{\sigma_i \cdot \sigma_j} \qquad 0 \le r_{ij} \le 1$$

$$r_{ii} = 1 \qquad r_{ij} = r_{ji}$$

$$\underset{(n,n)}{\mathbf{C}} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

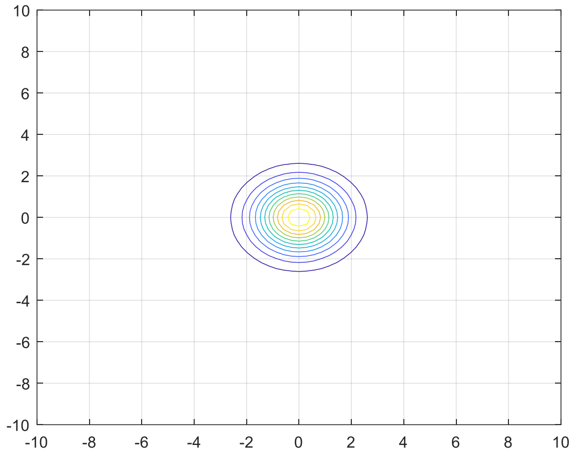
Variancia?

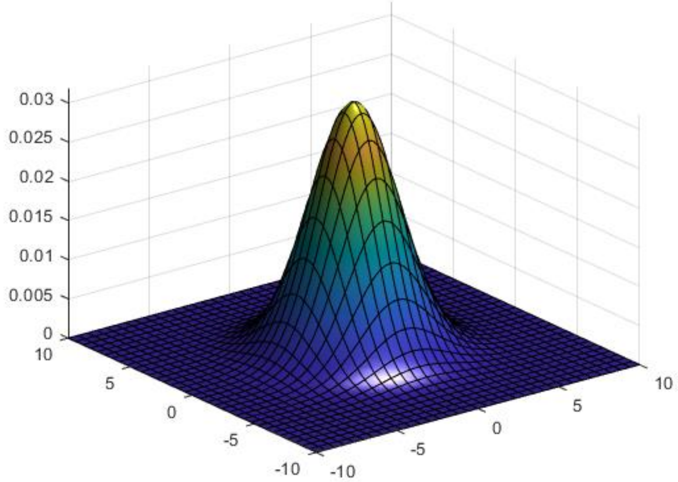
$$\underset{(n,n)}{\mathbf{R}} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix}$$

A normális eloszlás általánosítása

- Valószínűségi sűrűségfüggvény

$$f\left(x\right)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{\left(x-m\right)^2}{2\sigma^2}}$$





- Kétdimenziós esetben

$$f\left(x_1,x_2\right)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\exp\left\{-\frac{1}{2\left(1-r^2\right)}\left[\frac{\left(x_1-m_1\right)^2}{\sigma_1^2}-2r\frac{\left(x_1-m_1\right)\left(x_2-m_2\right)}{\sigma_1\sigma_2}+\frac{\left(x_2-m_2\right)^2}{\sigma_2^2}\right]\right\}$$

- N-dimenziós esetben

$$f\left(\mathbf{x}\right)=\frac{1}{\left(2\pi\right)^{n/2}\cdot\left|\mathbf{C}\right|}\exp\left[-\frac{1}{2}\left(\mathbf{x}-\mathbf{m}\right)^T\mathbf{C}^{-1}\left(\mathbf{x}-\mathbf{m}\right)\right] \qquad \left|\mathbf{C}\right|\neq 0$$

Köszönöm a figyelmet!