A46

2021. október 21., csütörtök 12:07

A46

A4 Valószínűségszámítás — VI. EA

dr. Keszthelyi Gabriella
Sztochasztika és Dinamikai rendszerek tanszék

2021. október 21.

Még Béta

Példa
Legyenek $X_1, X_2, \dots X_n$ független, egyenletes eloszlású valváltozók. $Y = \min(X_1, X_2, \dots, X_n)$.
Mi Y eloszlás- és sűrűségfv-e?

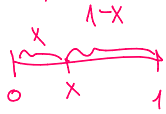
$X \sim Uni(0,1)$

komplementer, függetlenek

$$P(Y < x) = 1 - \underbrace{P(Y \geq x)}_{\text{komplementer}} = 1 - \Pi_{i=1}^{n}(P(X_i \geq x)) = 1 - \underbrace{P(X \geq x)^n}_{(1-x)^n}$$

$$= 1 - ((1 - \underbrace{P(X < x)}_{x})^n) = 1 - (1 - x)^n$$

$$f(y) = (1 - (1 - x)^n)' = n(1 - x)^{n-1}$$



$P(X>x) = \frac{1-x}{1}$

$P(X<x) = \frac{x}{1}$

Általában (k. legkisebb):

$k=1$

$$g(x) = n \cdot \binom{n-1}{k-1} x^{k-1}(1-x)^{n-k}$$

Paraméterbecslés

Még általánosabban:

a−1, b−1 siker sikertelen

$$g(x) = \frac{x^{a-1}(1-x)^{b-1}}{\beta(a,b)}$$

ahol

$a, b \in \mathbb{N}$

$$\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1} = \frac{(a-1)!(b-1)!}{(a+b-1)!}$$

Binomiális: $f(k) = \binom{n}{k} p^k (1-p)^{n-k}$ itt az n, p paraméter, k változó

Béta: $g(p) = n \cdot \binom{n-1}{k-1} p^{k-1}(1-p)^{n-k}$ itt a p változó, n, k paraméter

Példa
Tegyük fel, hogy van egy cinkelt érménk, a fej valószínűsége ismeretlen p. 12-szer dobjuk fel, 8-szor kapunk fejet. Becsüljük meg p valószínűségét!

$? \; \frac{8}{12} = \frac{2}{3}$ a fej valószínűsége?

Példa

Ez egy feltételes valószínűség lesz, ahol a p eloszlását próbáljuk megbecsülni a rendelkezésre álló adatokból (**posterior**).

- Első körben (adatok nélkül) p eloszlását egyenletesnek tekintjük (**prior**).
- Ezt update-eljük egy **likelihood**-dal ami jelen esetben egy binomiális eloszlás (k kísérlet volt sikeres n-ből egy rögzített p-re).
- Végül pedig leosztunk egy normalizáló faktorral.

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

binomiális feltéve $p = x$ x egyenletes eloszlású

$$\underbrace{(n+1)\binom{n}{k}x^k(1-x)^{n-k}}_{\text{sűrűségfv}} = \frac{\cancel{\binom{n}{k}}x^k(1-x)^{n-k} \cdot 1}{\cancel{\binom{n}{k}}\int_0^1 x^k(1-x)^{n-k}dx}$$

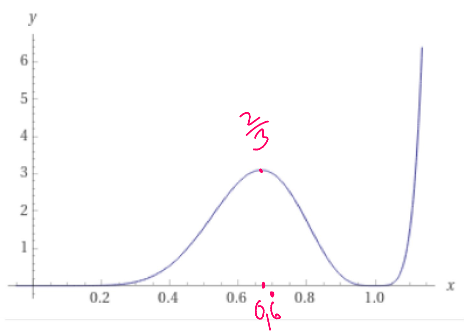
→ x-nek a sűrűségfv-e

$B(k+1, n-k+1) = \frac{k!\,(n-k)!}{(n+1)!}$

$n = 12, k = 8$, x sűrűségfüggvénye:

$$13\binom{12}{8}x^8(1-x)^4 = 6435x^8(1-x)^4$$

ha végzek egy új kísérletet pl. 20 dobás 15 fej → új prior



egyenletes eloszlás = speciális béta $n=1$ $k=1$

Nagy Számok Törvénye

"Ha egy szabályos pénzérmét n-szer feldobunk és az n feldobás során k-szor kapunk fejet, akkor dobások számának növelésével a k/n hányados (a fej dobásának relatív gyakorisága) egyre nagyobb valószínűséggel egyre pontosabban megközelíti az 1/2-et. "

Jacob Bernoulli *Ars conjectandi*

> *"Ha már sokszor volt fekete, akkor szükségképpen nagyobb lesz a piros valószínűsége, különben nagyszámú pörgetés után nem lehetne kb. ugyanannyi a pirosak és feketék száma."*

Géza, 3 óra rulettezés után a Tropicana Casinoban

Szerencsejáték

Ugyanakkor világos, hogy a rulettnek (vagy a pénzérémének) nincs emlékezőtehetsége és a minden pörgetésnél ugyanannyi a piros és a fekete valószínűsége.
Mit jelent a kb.?

> *Nem a piros és feketék számának különbsége nulla, hanem a piros/fekete és az összes dobás aránya 1/2 (vagy piros/fekete=fej/írás ≈ 1).*

Sőt, a pörgetések számának növelésével a piros/fekete aránya tart 1-hez, de annak valószínűsége, hogy *pontosan ugyanannyi a számuk* 0-hoz tart.

Várható érték – folytonos eset

Emlékeztető (diszkrét eset):

$$E(X) = \sum_{k=0}^{\infty} P(X = k) \cdot k$$

Folytonos eset:

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

$$E(X^2) = \int_{-\infty}^{\infty} f(x) \cdot x^2 \, dx$$

$$E(g(X)) = \int_{-\infty}^{\infty} f(x) \cdot g(x) \, dx$$

$$D^2(X) = \int_{-\infty}^{\infty} f(x) \cdot (x - E(X))^2 \, dx = E(X^2) - (E(X))^2 = \int f(x) \cdot x^2 dx - \left(\int f(x) \cdot x\right)^2$$

Nagy Számok Törvényei

Gyenge
Legyenek $X_1, X_2, \ldots X_n$ azonos eloszlású, független valváltozók véges $E(X)$ várható értékkel . Ekkor

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - E(X)\right| < \epsilon\right) = 1$$

Erős
Legyenek $X_1, X_2, \ldots X_n$ azonos eloszlású, független valváltozók véges $E(X)$ várható értékkel és véges $D^2(X)$ varianciával. Ekkor

$$P\left(\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = E(X)\right) = 1$$

De Moivre – Laplace tétele

$$\binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \approx$$

A Stirling formula ($n! \sim n^n e^{-n}\sqrt{2\pi n}$ ha $n \to \infty$) alapján:

$$\approx \frac{n^n e^{-n}\sqrt{2\pi n}}{k^k e^{-k}\sqrt{2\pi n}(n-k)^{n-k} e^{-(n-k)}\sqrt{2\pi(n-k)}} p^k (1-p)^{n-k} \approx$$

$$\approx \sqrt{\frac{1}{2\pi n \frac{k}{n}\left(1-\frac{k}{n}\right)}} \left(\frac{np}{k}\right)^k \left(\frac{n(1-p)}{k}\right)^{n-k} \approx$$

$k/n \to p$ + Taylor sorfejtés

$np = E(x)$ $np(1-p) = D^2(x)$

$$\approx \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{(k-np)^2}{2np(1-p)}}$$

NORMÁLIS ELOSZLÁS

Normális eloszlás

Binomiálisnál $E(X) = np, \quad D^2(X) = np(1-p)$

$$\frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{(k-np)^2}{2np(1-p)}} = \frac{1}{\sqrt{2\pi D^2(X)}} e^{-\frac{(X-E(X))^2}{2D^2(X)}}$$

Jelölés: $E(X) = \mu, \quad D(X) = \sigma, \quad X \sim \mathcal{N}(\mu, \sigma^2)$

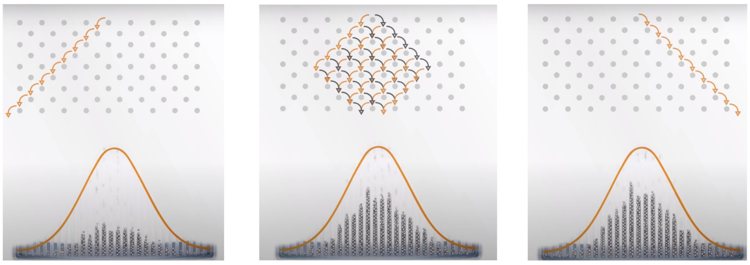
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

Eloszlásfüggvénye:

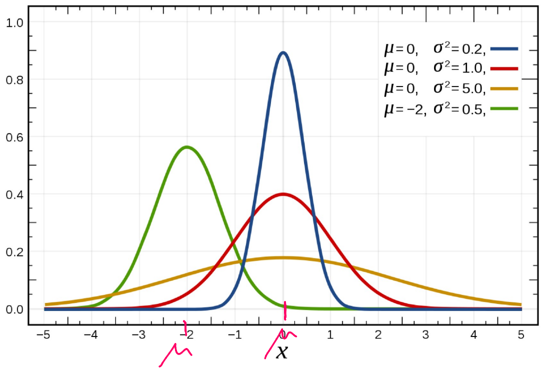
$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(t-\mu)^2} dt$$

Normális eloszlás család

Galton deszka

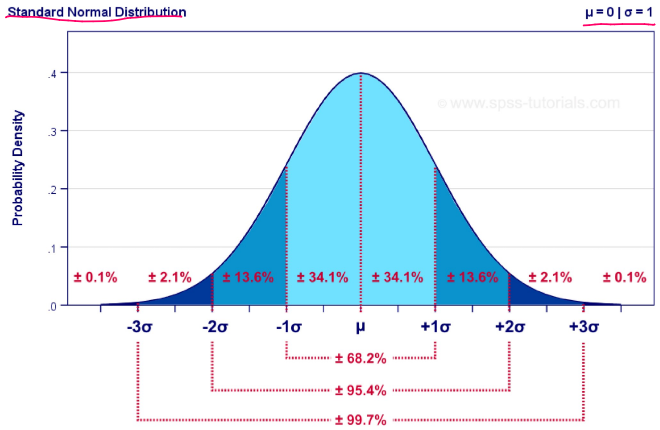


Normális eloszlás



Emberek magassága, cipőmérete, IQ-ja. Gépek, alkatrészek élettartama (ha kopásból eredően mennek tönkre). Mérési hibák eloszlása.

Normális eloszlás



tetszőleges szigmára:

$P(-\sigma < x < \sigma) = 0{,}682$

$P(-2\sigma < x < 2\sigma) = 0{,}954$

$P(-3\sigma < x < 3\sigma) = 0{,}997$

Standard normális eloszlás

Ha $\mu = 0, \sigma = 1$ akkor standard normális eloszlásról beszélünk:

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

Eloszlásfüggvény:

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

NEM LÉTEZIK A PRIMITÍV FV

Szimmetria:

$$\Phi(-x) = 1 - \Phi(x)$$

$[-r, r]$ intervallum valószínűsége:

$$\Phi(r) - \Phi(-r) = \Phi(r) - (1 - \Phi(r)) = 2\Phi(r) - 1$$

p valószínűséghez szimmetrikus intervallum:

$$2\Phi(r) - 1 = p$$
$$r = \Phi^{-1}((p+1)/2)$$

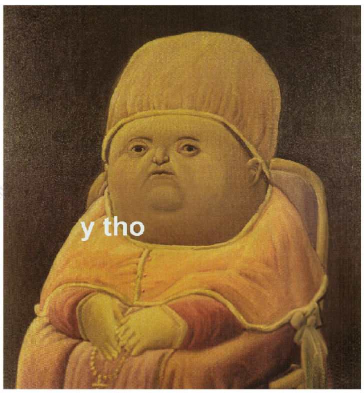
Standard Normális Eloszlástábla

Numerikus integrál közelítése egy 200 éves táblázat alapján

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.0	0.500	1.0	0.841	2.0	0.977	3.0	0.9987
0.1	0.540	1.1	0.864	2.1	0.982	3.1	0.9990
0.2	0.579	1.2	0.885	2.2	0.986	3.2	0.9993
0.3	0.618	1.3	0.903	2.3	0.989	3.3	0.9995
0.4	0.655	1.4	0.919	2.4	0.992	3.4	0.9997
0.5	0.691	1.5	0.933	2.5	0.994	3.5	0.9998
0.6	0.726	1.6	0.945	2.6	0.995	3.6	0.9998
0.7	0.758	1.7	0.955	2.7	0.997	3.7	0.9999
0.8	0.788	1.8	0.964	2.8	0.997	3.8	0.9999
0.9	0.816	1.9	0.971	2.9	0.998	3.9	1.0000
1.0	0.841	2.0	0.977	3.0	0.999	4.0	1.0000

Standardizálás

$$X \sim \mathcal{N}(\mu, \sigma^2)$$
$$Z = \frac{X - \mu}{\sigma}$$



WolframAlpha

Példa

Egy kommunikációs rendszer V feszültséget kap inputnak és $Y = \alpha V + N$ az outputja, ahol $\alpha = 10^{-2}$ és N normális eloszlású $\mu = 0$ és $\sigma = 2$-vel. Milyen V-re lesz $P(Y < 0) = 10^{-6}$?

$\sim N(0,2)$

$$P(Y < 0) = P(\alpha V + N < 0) = P(N < -\alpha V) =$$

STANDARDIZÁLÁS $\Phi(-\alpha V/\sigma) = 1 - \Phi(\alpha V/\sigma) = 10^{-6}$ $\qquad \Phi(\alpha V/\sigma) = 1 - 10^{-6}$

$N(0,1) \sim$ $\alpha V/\sigma = 4.753 \Rightarrow V = 950.6$

De Moivre – Laplace tétele

STANDARDIZÁLÁS

$$\lim_{n\to\infty} P\left(a \leq \frac{k - np}{\sqrt{np(1-p)}} \leq b\right) = \Phi(b) - \Phi(a)$$

Folytonossági korrekció:

$$P(X = k) \text{ helyett } P(k - 1/2 < X < k + 1/2)$$

Példa
Legyen X valváltozó a fejek száma, ha 40-szer feldobunk egy szabályos érmét. Mi a valószínűsége, hogy 20 fejet kaptunk?

$X \sim Binom(40, 0.5)$

$E(X) =$ $np = 40 \cdot 0.5 = 20 \quad \sqrt{np(1-p)} = \sqrt{10}$ $= D(x)$

$$P(X = 20) = \binom{40}{20}\left(\frac{1}{2}\right)^{40} \approx 0.1254$$

Példa

STANDARDIZÁLÁS

$$P(19.5 < X < 20.5) = P\left(\underbrace{\frac{19.5-20}{\sqrt{10}}}_{-0,16} < \frac{X-20}{\sqrt{10}} < \underbrace{\frac{20.5-20}{\sqrt{10}}}_{0,16}\right) =$$

$$= \Phi(0.16) - \Phi(0.16) \approx 0.1272$$

$2\Phi(0,16)-1$

Ugyanez Poissonnal:

$\lambda t \approx np = 20$ $\lambda = 20$

$$P(X = 20) = \frac{20^{20}}{20!}e^{-20} \approx 0.0888$$

:(

Poissonos közelítés csak kis λ-ra működik

$\lambda < 5$

De Moivre – Laplace tétele

Poisson közelítése normálissal

Példa
Memphis 40 km-es körzetében a regisztrált földrengések száma Poisson eloszlást követ 6.5-ös évi átlaggal. Mi a valószínűsége, hogy legalább 9 földrengés lesz a következő évben?

$$P(X \geq 9) = 1 - \sum_{k=0}^{8} \frac{6.5^8}{k!}e^{-6.5} \approx 0,208$$

λ elég nagy akkor szabad, ha $\lambda \geq 10$

Ugyanez normálissal:

$1-\Phi\left(\frac{2}{\sqrt{6.5}}\right)$

$$P(X \geq 9) = P(X > 8.5) = P\left(Z > \frac{8.5-6.5}{\sqrt{6.5}}\right) \approx 0.218$$

Minden közelítése normálissal

Centrális Határeloszlás Tétel (CHT)

Legyenek $X_1, X_2, \dots X_n$ azonos eloszlású, független valváltozók μ várható értékkel és σ^2 varianciával. Ekkor

$$\lim_{n\to\infty} P\left(\frac{\overbrace{X_1 + X_2 + \dots + X_n}^{S_n} - n\mu}{\sqrt{n}\sigma} \le a\right) = \Phi(a)$$

$E(\sum X_n) = E(S_n) = n \cdot \mu$

várható érték additivitása

$D^2(\sum X_n) = D^2(S_n) = n \cdot \sigma^2$

$D(S_n) = \sqrt{n}\sigma$

$S_n \sim N(n\mu, n\sigma^2)$

$E\left(\frac{S_n}{n}\right) = \frac{n \cdot \mu}{n} = \mu$

$D\left(\frac{S_n}{n}\right) = \frac{\sigma}{\sqrt{n}}$

Tehát a Poisson közelítése normálissal nagy λ-ra azért működik, mert felfogható sok kis lambda összegeként: $X_i \sim Poisson(1)$, $\sum_{i=1}^{100} X_i \sim Poisson(100)$

$$N(n\lambda, n\lambda) \approx Poisson(n\lambda)$$

Példa

Távoli csillag távolságának fényévekben való mérésekor minden mérésnél az atmoszféra változó állapota miatt hibát vétünk. Ezért több mérést hajtunk végre, és ezek eredményét átlagoljuk. A mérési eredményeket független azonos eloszlású változóknak tekintjük, melyek várható értéke a csillag (a megbecsülni kívánt) távolsága és varianciája 4 fényév. Hány mérést kell végeznünk, hogy 0.95 *valószínűséggel a méréseink eredményeinek átlaga a valódi távolságtól nem több mint 0.5 fényévvel térjen el?*

$E(X) = \mu = d$ a távolság $\quad D^2(X) = \sigma^2 = 4 \quad \sigma = 2$

$P\left(\left| \frac{S_n}{n} - d \right| < 0{,}5 \right) = 0{,}95$

$P\left(\frac{S_n - n \cdot d}{\sqrt{n} \cdot 2} < a \right) = \Phi(a)$

$P\left(-0{,}5 < \frac{S_n}{n} - d < 0{,}5 \right) = P\left(-0{,}5\sqrt{n} < \frac{S_n}{\sqrt{n}} - \sqrt{n} d < 0{,}5\sqrt{n} \right)$

$X_1, X_2, \ldots, X_n$ mérési eredmények

$S_n = \sum_{i=1}^{n} X_i$

$\frac{S_n}{n}$ = mérések átlaga

STANDARD NORM.

$= P\left(\underbrace{-\frac{0{,}5}{2}\sqrt{n} < \frac{S_n - nd}{2\sqrt{n}} < \frac{0{,}5}{2}\sqrt{n}}_{\text{STANDARD}} \right) =$

$P\left(-\frac{\sqrt{n}}{4} < \frac{S_n - nd}{2\sqrt{n}} < \frac{\sqrt{n}}{4} \right) =$

$\Phi\left(\frac{\sqrt{n}}{4}\right) - \Phi\left(-\frac{\sqrt{n}}{4}\right) =$

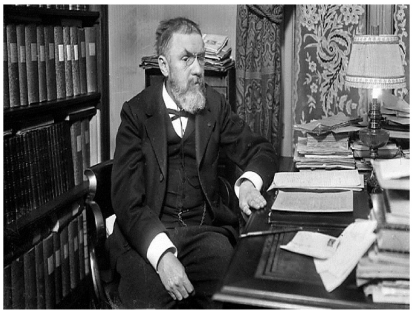
$= 2\Phi\left(\frac{\sqrt{n}}{4}\right) - 1 = 0{,}95$

$\Phi\left(\frac{\sqrt{n}}{4}\right) = \frac{1{,}95}{2} = 0{,}975$

$\frac{\sqrt{n}}{4} = 1{,}96$

$n \approx 61{,}44$

Tényleg olyan gyakori-e a normális?



"Mindenki hisz a normális eloszlásban: a fizikusok azért, mert azt hiszik a matematikusok igazolták logikai szükségességét, a matematikusok pedig azért, mert azt hiszik, egy kísérleti tény. "

Henri Poincaré

Irodalomjegyzék

- Vetier András—Valószínűségszámítás
- Alberto Leon-Garcia—Probability, Statistics, and Random Processes for Electrical Engineering
- Sheldon M. Ross —Introduction To Probability and Statistics for Engineers and Scientists

Köszönöm a figyelmet!