

Valószínűségszámítás

2020. november 4.
Mészáros Szabolcs

Tárgyhonlap:
cs.bme.hu/valszam
Diasor: shorturl.at/gxCJV

A prezentáció anyagát és az abból készült videofelvételt a tárgy hallgatói jogosultak használni, kizárólag saját célra. A felvétel másolása, videómegosztókra való feltöltése részben vagy egészben tilos, illetve csak a tantárgyfelelős előzetes engedélyével történhet.

Copyright © 2020, BME VIK

Ismétlés: szórásnégyzet

Definíció: Egy X valószínűségi változó *szórásnégyzete*

$$\mathbb{D}^2(X) = \text{cov}(X, X) = \mathbb{E}((X - \mathbb{E}X)^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

(Akkor is értelmes, ha X folytonos.)

Kiszámolása: transzformált várható értékével

$$\mathbb{E}(X^2) - \mathbb{E}(X)^2 = \int_{-\infty}^{\infty} x^2 f_X(x) dx - \left(\int_{-\infty}^{\infty} x f_X(x) dx \right)^2$$

Szórásnégyzet, folyt. példa

Példa: $Z \sim \text{Exp}(\lambda)$ $\mathbb{D}^2(Z) = ?$

$$\begin{aligned}\mathbb{E}(Z^2) &= \int_0^{\infty} z^2 \lambda e^{-\lambda z} dz = & \mathbb{D}^2(Z) &= \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2} \\ &= \left[-e^{-\lambda z} z^2 \right]_0^{\infty} - \int_0^{\infty} -e^{-\lambda z} 2z dz & & \uparrow \\ &= 0 + \left[2z \left(-\frac{1}{\lambda} \right) e^{-\lambda z} \right]_0^{\infty} - \int_0^{\infty} 2 \left(-\frac{1}{\lambda} \right) e^{-\lambda z} dz = \frac{2}{\lambda^2}\end{aligned}$$

“A” nevezetes eloszlás

Mi a közös az alábbi véletlen mennyiségek eloszlásában?

Emberek magasságai, lábmérete, születési súlya, vérnyomása, napi átlaghőmérséklet (az év adott napján), páratartalom, mérési hiba egy kísérletben, zaj.

Sok tényező apró, független tényező folytonos eredménye.

Elnevezések: normális eloszlás, Gauss-eloszlás, haranggörbe (a sűrűségfüggvénye).

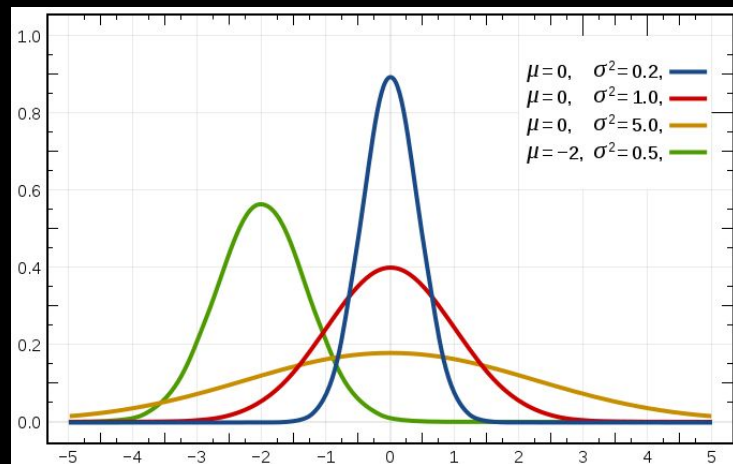
Normális eloszlás, def.

Definíció: Egy Y val. változó *normális eloszlású* μ és σ^2 paraméterekkel (ahol $\sigma^2 > 0$), ha sűrűségfüggvénye:

$$f_Y(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (x \in \mathbb{R})$$

Jelölés: $Y \sim N(\mu; \sigma^2)$

Speciális eset: $N(0; 1)$ Neve: standard normális



Standard normális eloszlás

Jelölés: standard normális sűrűségfüggvényére

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

Állítás: φ sűrűségfüggvény, vagyis a teljes \mathbb{R} -en vett integrálja 1.

Ötlet:

$$\left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \right)^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dx dy$$

Standard normális eloszlás

Biz. vázlat:

$$= \lim_{R \rightarrow \infty} \int_0^{2\pi} \underbrace{\int_0^R e^{-\frac{r^2}{2}} r \, dr}_{\left[-e^{-\frac{r^2}{2}} \right]_0^R} d\alpha$$

$$\left[-e^{-\frac{r^2}{2}} \right]_0^R = -e^{-\frac{R^2}{2}} + 1$$

$$= \lim_{R \rightarrow \infty} 2\pi(1 - e^{-\frac{R^2}{2}}) = 2\pi \implies \left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \right)^2 = 2\pi$$

Standardizálás, lemma

Kérdés: A nem-standard normális eloszlásra megadott függvények is sűrűségfüggvények?

Direkt számolás helyett:

Lemma: Legyen $\mu, \sigma \in \mathbb{R}$, $\sigma > 0$ és X sűrűségfüggvénye f_X .

Ekkor $Y = \sigma X + \mu$ sűrűségfüggvénye

$$f_Y(x) = \frac{1}{\sigma} f_X\left(\frac{x - \mu}{\sigma}\right) \quad (x \in \mathbb{R})$$

Standardizálás, lemma biz.

Biz: A megadott függvény nemnegatív, így már csak az integrálját kell kiszámoljuk minden a -ra:

$$\begin{aligned} \int_{-\infty}^a \frac{1}{\sigma} f_X\left(\frac{x - \mu}{\sigma}\right) dx &\stackrel{z = \frac{x - \mu}{\sigma}}{=} \int_{-\infty}^{\frac{a - \mu}{\sigma}} \frac{1}{\sigma} f_X(z) \sigma dz = \\ &= F_X\left(\frac{a - \mu}{\sigma}\right) = \mathbb{P}\left(X < \frac{a - \mu}{\sigma}\right) \\ &= \mathbb{P}(\sigma X + \mu < a) = \mathbb{P}(Y < a) = F_Y(a) \end{aligned}$$

Standardizálás normálisra

Következmény:

- Y pontosan akkor $N(\mu; \sigma^2)$ eloszlású, ha létezik $X \sim N(0; 1)$, amire $Y = \sigma X + \mu$
- Tehát a normális eloszlás valóban valószínűség-eloszlás.

Biz: Ha $X \sim N(0; 1)$ és $Y = \sigma X + \mu$, akkor a lemma miatt

$$f_Y(x) = \frac{1}{\sigma} f_X\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\frac{x - \mu}{\sigma}\right)^2}{2}} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x - \mu)^2}{2\sigma^2}}$$

Standardizálás normálisra, biz.

Visszafelé: ha $Y \sim N(\mu; \sigma^2)$ akkor legyen

$$X = \frac{1}{\sigma}(Y - \mu)$$

Ismét a lemma miatt

$$f_X(z) = \sigma f_Y(\sigma z + \mu) = \sigma \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\sigma z + \mu - \mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Kérdések: Mi a normális eloszlás

1. eloszlásfüggvénye,
2. várható értéke,
3. szórása?

Norm. elo., eloszlásfüggvény

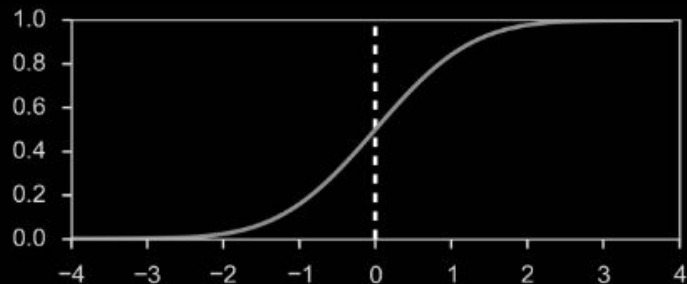
Jelölés: A standard normális eloszlás eloszlásfüggvénye:

$$\Phi(x) = \int_{-\infty}^x \varphi(z) dz = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

Kérdés: *Oké, és tengeri csikók nélkül?*

Ez elemi függvényekkel nem kifejezhető.

Ehelyett közelítéseket használunk a fenti integrál(ok)ra.



Hasznos összefüggés: $\Phi(-x) = 1 - \Phi(x) \quad (x \in \mathbb{R})$

Norm. elo., eloszlásfüggvény

A standard normális eloszlás eloszlásfüggvényének táblázata

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

x	0,00	0,01	0,02
0,0	0,5000	0,5040	0,5080
0,1	0,5398	0,5438	0,5478
0,2	0,5793	0,5832	0,5871
0,3	0,6179	0,6217	0,6255
0,4	0,6554	0,6591	0,6628
0,5	0,6915	0,6950	0,6985

$$\Phi(0,21) = 0,5832$$

$$\Phi(0,315) = 0,6236$$

$$\Phi(-0,11) = 1 - 0,5438 = 0,4562$$

Norm. elo., eloszlásfüggvény

Példa: Egy zajos csatornán egy +/-1 értékű jelet próbálunk átjuttatni. A zaj miatt a megkapott jel értéke nem feltétlenül +/-1, hanem a jel értéke plusz egy $N(0; \sigma^2)$ eloszlású véletlen szám. Határozzuk meg σ -t, ha tudjuk, hogy annak az esélye, hogy egy “+1” jel esetén negatív érték jut át a csatornán, 0,0226?

Zaj: $Y \sim N(0; \sigma^2)$

$$\begin{aligned}\mathbb{P}(+1 + Y < 0) &= \mathbb{P}(Y < -1) = \mathbb{P}\left(\frac{Y}{\sigma} < \frac{-1}{\sigma}\right) = \\ &= \Phi\left(-\frac{1}{\sigma}\right) = 0,0226 \implies \sigma = \frac{1}{2,00}\end{aligned}$$

Norm. elo., várható érték

Állítás: Legyen $Y \sim N(\mu; \sigma^2)$. Ekkor

$$\mathbb{E}(Y) = \mu \qquad \mathbb{D}^2(Y) = \sigma^2$$

Normális eloszlás standardizálása:

$$\frac{Y - \mathbb{E}Y}{\mathbb{D}(Y)} \sim N(0; 1)$$

Megjegyzés: Más eloszlású val. változót is standardizálhatunk (levonva a várható értékét, és leosztva a szórásával). Ez a standardizált persze tipikusan nem lesz standard normális eloszlású.

Norm. elo., várható érték

Állítás bizonyítása:

$$\mathbb{E}(Y) = \mathbb{E}(\sigma X + \mu) = \sigma \mathbb{E}(X) + \mu$$

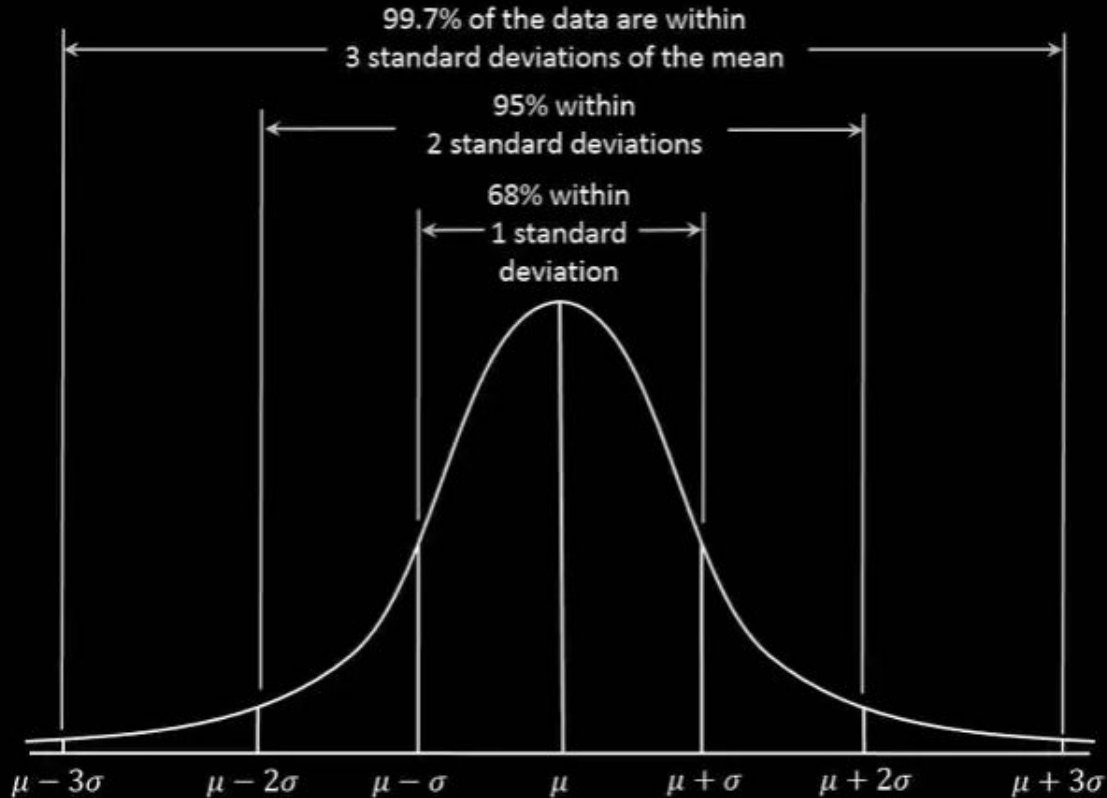
$$\mathbb{D}^2(Y) = \mathbb{D}^2(\sigma X + \mu) = \mathbb{D}^2(\sigma X) = \sigma^2 \mathbb{D}^2(X)$$

Tehát elég a standard normálisra kiszámolni az állítást.

$$\mathbb{E}(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \left[-e^{-\frac{x^2}{2}} \right]_{-\infty}^{\infty} = 0$$

$$\mathbb{E}(X^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx = (\text{parciális integrálás}) = 1$$

Norm. elo., szórás



Normális eloszlás
szórása szemléletesen:

- az esetek kb. 68%-át lefedi a $[\mu - \sigma, \mu + \sigma]$ intervallum,
- a sűrűségfüggvény a $\mu \pm \sigma$ pontoknál vált konvexitást.

Norm. elo., példa

Példa: minta hőmérséklete (°C) Jel.: Y
 $\sim N(-2; 1,69)$ Mi a valószínűsége, hogy
a minta hőmérséklete nagyobb, mint 0 °C?

Kérdés: $\mathbb{P}(Y > 0) = ?$ Standardizált: $X = \frac{Y + 2}{\sqrt{1,69}} \sim N(0; 1)$

$$\mathbb{P}\left(\frac{Y + 2}{1,3} < x\right) = \mathbb{P}(X < x) = \Phi(x)$$

$$\mathbb{P}(Y > 0) = 1 - \mathbb{P}(Y < 0) = 1 - \mathbb{P}\left(\frac{Y + 2}{1,3} < \frac{2}{1,3}\right) = 1 - \Phi\left(\frac{2}{1,3}\right)$$

$$\mathbb{P}(Y > 0) \approx 0,0620 \approx 6\%$$

de Moivre–Laplace-tétel

Kérdés: Miért jön elő alkalmazásokban, mérési eredményeknél?

Tétel: Legyen $p \in (0, 1)$ és $S_n \sim B(n; p)$.
Ekkor minden $a < b$ valós számokra:

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left(a < \frac{S_n - \mathbb{E}(S_n)}{\mathbb{D}(S_n)} < b \right) &= \\ &= \int_a^b \varphi(x) dx = \Phi(b) - \Phi(a) \end{aligned}$$

de Moivre–Laplace-tétel, példa

Példa: A matematikusok 31,4% százaléka szandált hord. Száz találomra választott matematikust nézve, közelítőleg mi az esélye, hogy kevesebb, mint 25 pár szandált találunk rajtuk?

Szandálok száma: $S_n \sim B(n; p)$ $n = 100$ $p = 0,314$

$$\mathbb{E}(S_n) = np = 31,4 \quad \mathbb{D}(S_n) = \sqrt{np(1-p)} = 4,6412$$

$$\frac{1}{4,6412}(S_n - 31,4) \text{ közelítőleg } N(0; 1)$$

$$\mathbb{P}(S_n < 25) = \mathbb{P}\left(\frac{S_n - 31,4}{4,6412} < \frac{25 - 31,4}{4,6412}\right) \approx \Phi(-1,3790)$$

← -1,3790

de Moivre–Laplace-tétel, megj.

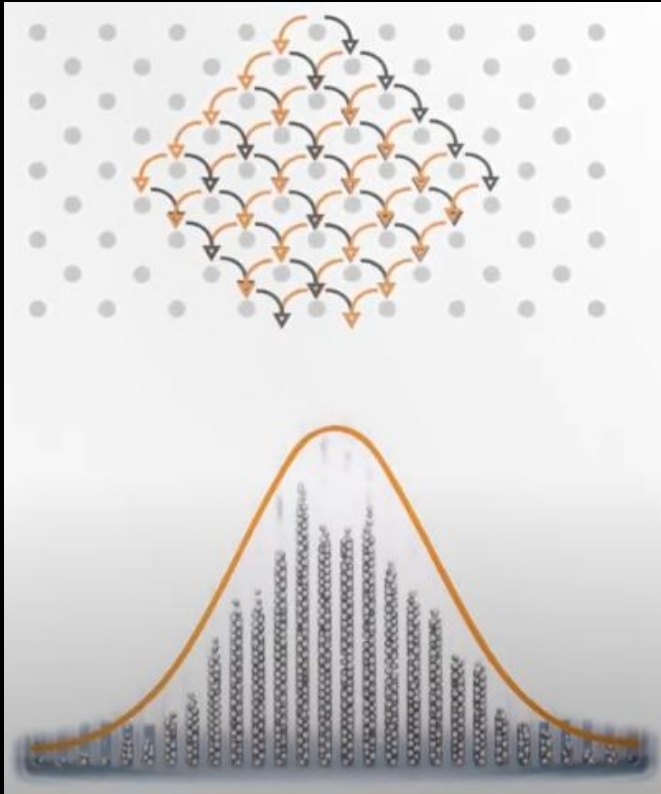
Kérdés: Nem ezt mondtuk a Poisson eloszlásnál is, hogy a binomiális határeloszlása a Poisson? Most meg azt mondjuk, hogy a binomiális határeloszlása a normális?

Válasz: A részletek eltérnek.

- Ott nem vontunk le (n -től függő) várható értéket, itt igen.
- Ott nem osztottunk le (n -től függő) szórással, itt igen.
- Ott $p_n \rightarrow 0$ és $np_n = \lambda$ volt feltétel, itt p konstans.

Megjegyzés: A konvergencia sebességéről is lehet tudni konkrétumot, lásd Berry–Esseen tétel.

Galton-deszka



Egy golyó vég-pozíciójának eloszlása:

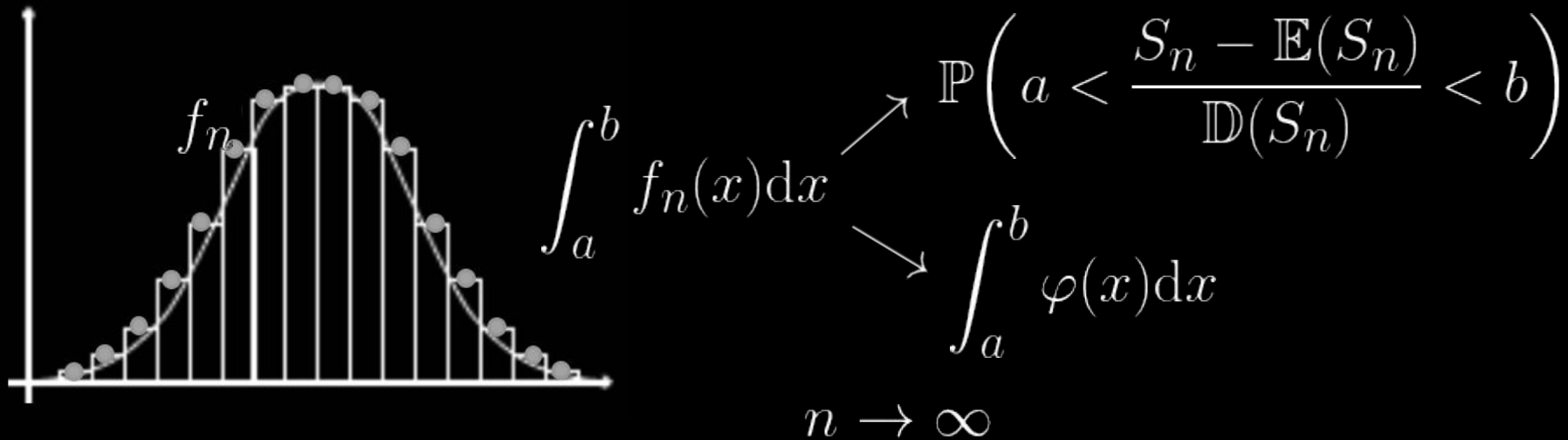
$$B\left(n; \frac{1}{2}\right) \text{ (a bal széléről számolva)}$$

Mit szemléltet? A binomiális eloszlás közelíti a normális eloszlás sűrűségfüggvényét.

de Moivre–Laplace-tétel, levezetés

Biz vázlat: A standardizált $\frac{S_n - \mathbb{E}(S_n)}{\mathbb{D}(S_n)}$ változó eloszlásából

készítsünk f_n sűrűségfüggvény-szerűséget:



de Moivre–Laplace-tétel, levezetés

Honnan tudjuk, hogy $f_n \rightarrow \varphi$?

Az f_n függvények közelítőleg teljesítenek egy diff-egyenletet:

$$\lim_{n \rightarrow \infty} \frac{f'_n(x)}{f_n(x)} = -x$$

Miért jó ez? Mert ebből a $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ függvényre belátható, hogy

$$-x = f'(x)/f(x) \implies f(x) = e^{-\frac{x^2}{2}} e^c$$

Nem-triviális részletek:

- létezik-e a fenti határérték?
- a határértékfüggvény deriválható-e?
- igaz-e a határértékfüggvényre a diff-egyenlet?

Köszönöm a figyelmet!
