Az $X$ eloszlásfüggvénye a következőképpen számolható:

$$\begin{aligned} F(x) &= \mathbf{P}(X \le x) \\ &= \int_{-\infty}^{x} f_{\mu,\sigma}(y)\mathrm{d}y \\ &= \int_{-\infty}^{(x-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\mathrm{d}y \\ &= \Phi((x-\mu)/\sigma), \end{aligned} \tag{5}$$

ahol

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}}\mathrm{d}y$$

a standard normális eloszlás eloszlásfüggvénye. Ebből a számolásból világos, hogy elég a $\Phi$ függvény értékeit ismerni, és ebből tetszőleges paraméterű normális eloszlás eloszlásfüggvénye számolható.

Ugyancsak (5) egyszerű következménye az alábbi állítás.

**8.2. Állítás.** *Ha $X \sim N(\mu, \sigma^2)$, akkor $(X-\mu)/\sigma \sim N(0,1)$.*

Sőt, ez kicsit általánosabban is igaz: ha $X \sim \mathrm{N}(\mu, \sigma^2)$, és $a, b$ valós állandók, $a \neq 0$, akkor $aX + b \sim \mathrm{N}(a\mu + b, a^2\sigma^2)$.

A normális eloszlás nagyon erősen koncentrálódik a várható értéke körül. Valóban, ha $X \sim \mathrm{N}(\mu, \sigma^2)$ és $Z \sim \mathrm{N}(0,1)$, akkor

$$\begin{aligned} \mathbf{P}(|X-\mu| \le \lambda\sigma) &= \mathbf{P}(|Z| \le \lambda) \\ &= \Phi(\lambda) - \Phi(-\lambda) \\ &= 2\Phi(\lambda) - 1 \\ &= \begin{cases} 0,6827, & \lambda = 1, \\ 0,9545, & \lambda = 2, \\ 0,9973, & \lambda = 3, \\ 0,9999, & \lambda = 4. \end{cases} \end{aligned}$$

# 9. Véletlen változók konvergenciája

$E(|X|) < \infty$

## 9.1. Markov és Csebisev egyenlőtlenségei

**9.1. Tétel.** *Markov-egyenlőtlenség Legyen $X$ egy véletlen változó $(\Omega, \mathcal{A}, \mathbf{P})$ valószínűségi mezőn, melynek véges a várható értéke. Ekkor tetszőleges pozitív $c$ konstansra*

$$\mathbf{P}(|X| \ge c) \le \frac{\mathbf{E}(|X|)}{c}.$$

$=$

$P(\{\omega : |X(\omega)| \geq c\})$

Ha $c \leq E(|X|)$

akkor $\frac{E(|X|)}{c} \geq 1$

$h(x)=|x|$

$$\mathbf{P}(|X|\geq c)\leq \frac{\mathbf{E}(|X|)}{c} \qquad \mathbf{E}(|X|)=\sum_i{}' |x_i|\cdot \mathbf{P}(X=x_i)$$

$$\mathbf{P}(|X|\geq c)=\mathbf{P}(\{\omega : |X(\omega)|>c\})=\mathbf{P}(\{\omega : X(\omega)=x_i,\ |x_i|>c\})$$

(véges, vagy megszámlálhatóan végtelen)

*Bizonyítás.* Ha $X$ diszkrét $x_1, x_2, \ldots$ lehetséges értékekkel, akkor

$$\mathbf{P}(|X| \geq c) \overset{\text{megszámlálható additív}}{=} \sum_{i:|x_i|\geq c} 1\cdot\mathbf{P}(X = x_i) \overset{(*)}{\leq} \sum_{i:|x_i|\geq c} \frac{|x_i|}{c}\mathbf{P}(X = x_i)$$
$$\leq \sum_{i} \frac{|x_i|}{c}\mathbf{P}(X = x_i) = \frac{\mathbf{E}(|X|)}{c}.$$

$(*)$: ha $|x_i|\geq c$ akkor $\frac{|x_i|}{c}\geq 1$

Ha $X$ folytonos $f$ sűrűségfüggvénnyel, akkor

$$\mathbf{P}(|X| \geq c) = \int_{|y|\geq c} 1\cdot f(y)\mathrm{d}y \leq \int_{|y|\geq c} \frac{|y|}{c} f(y)\mathrm{d}y$$
$$\leq \int_{-\infty}^{\infty} \frac{|y|}{c} f(y)\mathrm{d}y = \frac{\mathbf{E}(|X|)}{c}.$$

$\mathbf{P}(X\in A)=\int_A f(y)\,dy$ akkor $f$ az $X$ sfgve.

$A=\{x: |x|\geq c\}$.

$\mathbf{E}(|X|)=\int_{-\infty}^{\infty}|y|\cdot f(y)\,dy$

□

A Markov-egyenlőtlenség egyszerű alkalmazásával adódik a

**9.2. Tétel.** *Csebisev-egyenlőtlenség Legyen $X$ egy véletlen változó $(\Omega, \mathcal{A}, \mathbf{P})$ valószínűségi mezőn, melynek véges a szórása. Ekkor tetszőleges pozitív $c$ konstansra*

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq c) \leq \frac{\mathbf{D}^2(X)}{c^2}.$$

Másik alak: $c=\lambda\cdot\mathbf{D}(X)$

$$\mathbf{P}(|X-\mathbf{E}(X)|\geq \lambda\cdot\mathbf{D}(X))\leq \frac{\mathbf{D}^2(X)}{\lambda^2\mathbf{D}^2(X)}=\frac{1}{\lambda^2}$$

*Bizonyítás.* A Markov-egyenlőtlenség szerint

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq c) = \mathbf{P}(\underbrace{(X - \mathbf{E}(X))^2}_{Y} \geq c^2) \leq \frac{\mathbf{D}^2(X)}{c^2}.$$

$Y=(X-\mathbf{E}(X))^2\geq 0$ □

$$\mathbf{P}(Y\geq c^2)\underset{\text{Markov}}{\leq}\frac{\mathbf{E}(Y)}{c^2}=\frac{\mathbf{E}((X-\mathbf{E}(X))^2)}{c^2}=\frac{\mathbf{D}^2(X)}{c^2}.$$

## 9.2. Nagy számok gyenge törvénye

**9.3. Tétel.** *Csebisev-féle nagy számok gyenge törvénye Legyenek $X_1, X_2, \ldots$ páronként független, véges szórású véletlen változók, melyek közös várható értéke $\mu$ és szórásnégyzete $\sigma^2$. Ekkor tetszőleges $\varepsilon > 0$ esetén*

$$\lim_{n\to\infty} \mathbf{P}\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| > \varepsilon\right) = 0.$$

$X, X_1, X_2, \ldots$

$\mathbf{E}(X_i)=\mu \quad i=1,2,\ldots$

$\mathbf{D}^2(X_i)=\sigma^2 \quad i=1,2,\ldots$

*Bizonyítás.* A páronkénti függetlenség miatt

$$\mathbf{D}^2(X_1 + \ldots + X_n) = n\sigma^2.$$

$$\mathbf{E}(X_1+\ldots+X_n)\overset{\text{mindig}}{=}\sum_{i=1}^{n}\mathbf{E}(X_i)=\sum_{i=1}^{n}\mu=n\mu$$

$$\mathbf{E}\left(\frac{X_1+\ldots+X_n}{n}\right)=\frac{1}{n}\cdot\mathbf{E}(X_1+\ldots+X_n)=\frac{1}{n}\cdot n\mu=\mu.$$

Várható érték motivációja:

független ismétlések az eredményekkel: $X_1, X_2, \dots$

pl.: kockadobás, $X_i$: $i$-edik dobás eredménye

$X_1 + X_2 + \dots + X_n$ : az első $n$ dobás összege

$$\frac{X_1 + \dots + X_n}{n} \quad \text{átlag} \approx E(X)$$

nem definiált

$\approx$ jelentése:

$$P\left(\left| \frac{X_1 + \dots + X_n}{n} - E(X) \right| > \varepsilon\right) = P\left(\left\{\omega : \left| \frac{X_1(\omega) + \dots + X_n(\omega)}{n} - E(X)\right| > \varepsilon \right\}\right) \xrightarrow[n \to \infty]{} 0$$

$(\Omega, \mathcal{A}, P)$ val.m., $X_i : \Omega \to \mathbb{R}$

Csebisev

$$P\left(\left|\frac{X_1+\dots+X_n}{n}-\mu\right|>\varepsilon\right) = P\left(|X_1+\dots+X_n-n\mu|>n\varepsilon\right) \le \frac{D^2(X_1+\dots+X_n)}{(n\varepsilon)^2}$$

korrelálatlanok

$$D^2(X_1+\dots+X_n) = \sum_{i=1}^{n} D^2(X_i) = n\sigma^2$$

bilineáris

$$\left[\ \mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i,X_j) = \sum_{i=1}^{n} D^2(X_i)\right.$$

$$\mathrm{Cov}(X_i,X_j) = \begin{cases} 0 & i \ne j \\ D^2(X_i) & i = j \end{cases}$$

A Csebisev-egyenlőtlenséget az $X = X_1 + \ldots + X_n$ változóra fölírva kapjuk, hogy

$$\mathbf{P}\left(\left|\frac{X_1+\ldots+X_n}{n}-\mu\right|>\varepsilon\right) \le \frac{\mathbf{D}^2(X_1+\ldots+X_n)}{n^2\varepsilon^2} = \frac{n\sigma^2}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}, \to 0$$

ha $n\to\infty$ hiszen $\varepsilon>0$ fix!!

ami tart 0-hoz. □

A bizonyításból látjuk, hogy a páronkénti függetlenség helyett elég korrelálatlanságot feltenni.

Speciális esetként adódik a

**9.4. Tétel.** *Bernoulli-féle nagy számok gyenge törvénye (1713) Jelölje $S_n$ egy $p$ valószínűségű $A$ esemény bekövetkezéseinek a számát egy kísérlet $n$ független ismétlése során. Ekkor tetszőleges $\varepsilon > 0$ esetén*

$$\lim_{n\to\infty} \mathbf{P}\left(\left|\frac{S_n}{n}-p\right|>\varepsilon\right) = 0.$$

$S_n = I_1 + I_2 + \dots + I_n$

$I_i = \begin{cases} 1 & \text{ha az } i\text{-edik kísérletben bekövetkezett a vizsgált esemény} \\ 0 & \text{kül.} \end{cases}$

Ekkor $I_1, I_2, \dots$ független azonos eloszlású vv. Bernoulli$(p)$

$E(I) = p$

$D^2(I) = p(1-p)$

A tétel szerint a relatív gyakoriságok a fenti értelemben konvergálnak az igazi valószínűséghez. Mivel a valószínűség definícióját a relatív gyakoriságok tulajdonságai motiválták (additivitás), ezért a fenti tétel szerint a valószínűség tényleg az, amit akarunk.

A fenti tételekben szereplő konvergencia a sztochasztikus konvergencia, melynek általános definíciója a következő.

$(\Omega, \mathcal{A}, P)$ vm.
$X_n : \Omega \to \mathbb{R}$ vv
$X : \Omega \to \mathbb{R}$ vv.

**9.5. Definíció.** Az $(X_n)_{n\in\mathbb{N}}$ véletlen változók sorozata *sztochasztikusan konvergál $X$-hez*, ha minden $\varepsilon > 0$ esetén

$$\lim_{n\to\infty} \mathbf{P}\left(|X_n - X| > \varepsilon\right) = 0.$$

*Megjegyzés.* A fenti tételekben szereplő gyenge jelző arra utal, hogy a konvergencia sztochasztikusan teljesül. Erős konvergenciáról akkor beszélünk, ha a véletlen változók majdnem biztosan konvergálnak. Pontosabban, az $(X_n)_{n\in\mathbb{N}}$ véletlen változók sorozata *majdnem biztosan, vagy egy valószínűséggel konvergál $X$-hez*, ha

$$\mathbf{P}\left(\left\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = \mathbf{P}(\lim_{n\to\infty} X_n = X) = 1.$$

Itt persze már az is magyarázatra szorul, hogy az $\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}$ halmaz valóban esemény, azaz eleme a megfelelő $\sigma$-algebrának. Ez a $\sigma$-algebra tulajdonságaiból következik. Erre részletesebben nem térünk ki.

A nagy számok erős törvénye a következő.

$\frac{S_n}{n}$: relatív gyakoriság : n kísérlet során a 6-osok aránya

$P(6\text{-os}) = \frac{1}{6}$ ≈ sokszor dobunk, akkor a 6-osok aránya kb. $\frac{1}{6}$.

$\frac{S_n}{n} \underbrace{\approx}_{} \frac{1}{6}$ precízen: $(\forall \varepsilon > 0): \lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \frac{1}{6}\right| > \varepsilon\right) = 0.$

nem definiált
≈ nagyjából

---

Erős törvények.

$$\left\{ \lim_{n\to\infty} \frac{X_1+X_2+\ldots+X_n}{n} = E(X) \right\} = \left\{ \omega : \lim_{n\to\infty} \frac{X_1(\omega)+\ldots+X_n(\omega)}{n} = E(X) \right\} \subseteq \Omega$$

$\boxed{\in A}$

dolgozni kell
HF

rögzített $\omega \in \Omega$ esetén $X_1(\omega), X_2(\omega), \ldots$ számsorozat

$\frac{1}{n} \sum_{i=1}^{n} X_i(\omega)$ valós szám determinisztikus

$$U_n(\omega) = \frac{1}{n} \sum_{i=1}^{n} X_i(\omega) \qquad \left(U_n(\omega)\right)_{n=1}^{\infty}$$

$$P\left( \left\{ \omega : \lim_{n\to\infty} \frac{X_1(\omega)+\ldots+X_n(\omega)}{n} = E(X) \right\} \right) = 1$$ nagy számok erős törvénye

$P: A \to [0,1]$

---

első rendű közelítés

NSZT: $\frac{S_n}{n} \approx E(X)$

n-szer dobunk érmével (szabályos) $S_n$: fejek száma

NSZT: $S_n \approx \frac{n}{2}$

1000-szer dobunk: kb. 500 fej lesz.

másodrendű közelítés

CHT: Mennyi a kb? Mit jelent a kb?

1000-szer dobunk 495-505 fejek száma

450-550 fejek száma

300-700 ?

9.6. *Tétel. Nagy számok erős törvénye Legyenek $X, X_1, X_2, \ldots$ független, azonos eloszlású véletlen változók, véges $\mathbf{E}(X)$ várható értékkel. Ekkor*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n X_i}{n} = \mathbf{E}(X) \quad \textit{majdnem biztosan.}$$

$(\Leftrightarrow) \; P\left( \lim \frac{\sum X_i}{n} = E(X) \right) = 1$

## 9.3. Centrális határeloszlás-tétel

A nagy számok törvénye azt állítja, hogy független, azonos eloszlású véletlen változók átlagai közel vannak a várható értékhez. Az alábbiakban ezt a közelséget tesszük precízzé.

**9.7. Tétel.** *Centrális határeloszlás-tétel Legyenek $X, X_1, X_2, \ldots$ független, azonos eloszlású véletlen változók közös $\mathbf{E}(X) = \mu$ várható értékkel, és véges $\mathbf{D}(X) = \sigma$ szórással. Ekkor tetszőleges $x \in \mathbb{R}$ esetén*

$$\lim_{n\to\infty} \mathbf{P}\left( \frac{\sum_{i=1}^n (X_i - \mu)}{\sqrt{n}\sigma} \le x \right) = \Phi(x),$$

*ahol*

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} \mathrm{d}y$$

*a standard normális eloszlás eloszlásfüggvénye.*

A tétel bizonyítása már komolyabb eszközökkel, a karakterisztikus függvények módszerével történik.

A tétel indikátorváltókra vonatkozó speciális esete a

**9.8. Tétel.** *de Moivre–Laplace tétel Jelölje $S_n$ egy $p$ valószínűségű $A$ esemény bekövetkezéseinek a számát egy kísérlet $n$ független ismétlése során. Ekkor tetszőleges $x \in \mathbb{R}$ esetén*

$$\lim_{n\to\infty} \mathbf{P}\left( \frac{S_n - n\overset{E(I)}{p}}{\underbrace{\sqrt{np(1-p)}}_{\text{Bernoulli szórásnégyzet}}} \le x \right) = \Phi(x).$$

$S_n = I_1 + \ldots + I_n$ ahol $I_1, I_2, \ldots$ …

$D^2(I) = p(1-p)$

$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$

$\frac{n!}{k!(n-k)!}$

Stirling-formula

Valóban, korábban láttuk, hogy a $p$-paraméterű Bernoulli-eloszlás várható értéke $p$ és szórása $\sqrt{p(1-p)}$. A speciális eset bizonyítása a binomiális együtthatók pontos aszimptotikájának meghatározásával történhet.

9.9. *Példa.* A kakucsretyegei polgármesterválasztáson két jelölt van: A és B. Kakucsretyege 40000 szavazója egymástól függetlenül, 1/2−1/2 valószínűséggel szavaz a két jelölt egyikére. A feszült politikai helyzet miatt a szavazatok újraszámlálását rendelik el, ha a két jelöltre leadott szavazatok száma között

$$E(Y_i) = E\left(\frac{X_i-\mu}{\sigma}\right) = \frac{1}{\sigma}E(X_i-\mu) = 0.$$

$$D^2(Y_i) = D^2\left(\frac{X_i-\mu}{\sigma}\right) = \frac{1}{\sigma^2}D^2(X_i) = 1$$

$$\frac{\sum_{i=1}^{n}(X_i-\mu)}{\sqrt{n}\,\sigma} = \frac{1}{\sqrt{n}}\cdot\sum_{i=1}^{n}\overbrace{\frac{X_i-\mu}{\sigma}}^{=:Y_i}$$

$$D^2\left(\frac{\sum_{i=1}^{n}(X_i-\mu)}{\sqrt{n}\,\sigma}\right) = \frac{1}{n}\sum_{i=1}^{n}D^2\left(\frac{X_i-\mu}{\sigma}\right) = \frac{1}{n}\sum_{i=1}^{n}1 = \frac{n}{n} = 1.$$

Tehát

$\frac{\sum(X_i-\mu)}{\sqrt{n}\,\sigma}$ várható értéke 0, szórásnégyzete 1.
(standardizált v.v.)

$$P\Bigg(\underbrace{\frac{\sum(X_i-\mu)}{\sqrt{n}\,\sigma}}_{\approx Z} \le x\Bigg) \xrightarrow[n\to\infty]{} P(Z \le x) = \Phi(x)$$

$$\frac{\sum(X_i-\mu)}{\sqrt{n}\,\sigma} = \frac{\sum X_i - n\mu}{\sqrt{n}\,\sigma} \approx Z \qquad Z \sim N(0,1)$$

$$\sum X_i \approx \underbrace{n\mu}_{\substack{\text{elsőrendű}\\ \text{közelítés}\\ \text{NSzT}}} + \underbrace{\sqrt{n}\,\sigma\cdot Z}_{\substack{\text{másodrendű}\\ \text{közelítés}\\ \text{CHT}}}$$

$n \gg \sqrt{n}$

$$\sqrt{1000}\cdot\frac{1}{2} \approx 16$$

Bernoulli$\left(\frac{1}{2}\right)$ $\quad \sigma^2 = \frac{1}{2}\left(1-\frac{1}{2}\right) = \frac{1}{4}$

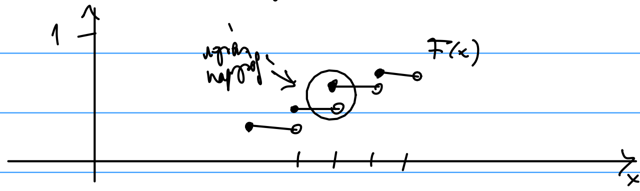
<u>Eloszlásbeli konvergencia</u>: $X, X_1, X_2, \dots$ véletlen változók

$X_n \xrightarrow{\mathcal{D}} X$ eloszlásban konvergálnak $X$-hez, ha minden olyan $x$-re, ami folytonossági pontja $P(X \leq x)$-nek.

(*) $\lim\limits_{n\to\infty} P(X_n \leq x) = P(X \leq x)$

Spec: $X$ folyt. vv. $\Rightarrow$ $F(x) = P(X \leq x)$ folytonos
ekkor (*) teljesül $\forall x \in \mathbb{R}$.

Korábban. (Poisson konvergencia): nemnegatív egész értékű vv esetén



$n=40\,000$

$S_n \sim$ Binomiális$(n, \frac{1}{2})$ : A-ra szavazók száma

legfeljebb 100 a különbség. Mi a valószínűsége, hogy újraszámlálásra kerül sor?

Ekkor tehát $n = 40000$, $p = \mathbf{P}(\text{A-ra szavaz valaki}) = 1/2$. Legyen $S_n$ az A-ra szavazók száma, ekkor $n - S_n$ a B-re szavazók száma. A kérdés $\mathbf{P}(|S_n - (n - S_n)| \le 100)$. A CHT-ban előforduló mennyiségek $np = 20000$ és $\sqrt{np(1-p)} = 100$. Így a CHT szerint

$\frac{S_n - np}{\sqrt{np(1-p)}} = \frac{S_n - 20000}{100}$

$a<b$

$$\begin{aligned}\mathbf{P}(|S_n - (n - S_n)| \le 100) &= \mathbf{P}(-100 \le 2S_n - n \le 100)\\ &= \mathbf{P}\left(-0,5 \le \frac{S_n - np}{\sqrt{npq}} \le 0,5\right)\\ &\approx \Phi(0,5) - \Phi(-0,5)\\ &= 2\Phi(0,5) - 1 \approx 0,38.\end{aligned}$$

$P(a \le X \le b) = P(X \le b) - P(X < a)$

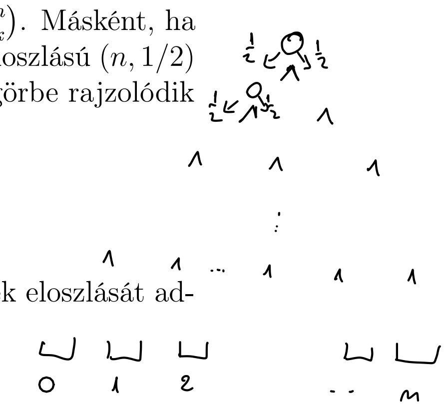
$\approx \Phi(b) - \Phi(a)$

$-100 \le 2S_n - n \le 100$

$-\frac{1}{2} \le \frac{S_n - 20000}{100} \le \frac{1}{2}$

**Galton deszkája.** *Sir Francis Galton (1822–1911): polihisztor, Darwin unokatestvére.* A centrális határeloszlás szemléltetése. Az első sorban 1 ék van, alatta 2, …, az $n$-edik sorban $n$. Az $n$-edik éksor alatt van $n+1$ tartály, 0-tól $n$-ig sorszámozva. Egy golyót elindítunk az első éknél, és a golyót minden ék $1/2 - 1/2$ valószínűséggel téríti el jobbra vagy balra. Annak a valószínűsége, hogy a golyó a $k$-adik tartályban landol $= \frac{1}{2^n} \cdot$ azon útvonalak száma, ahol a golyó $k$-szor megy jobbra és $(n-k)$-szor balra $= \frac{1}{2^n}\binom{n}{k}$. Másként, ha $S_n$ a golyó jobbra eltérítéseinek száma, akkor $S_n$ binomiális eloszlású $(n, 1/2)$ paraméterekkel. Ha sok golyót engedünk le, akkor a haranggörbe rajzolódik ki a tartályokban.

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

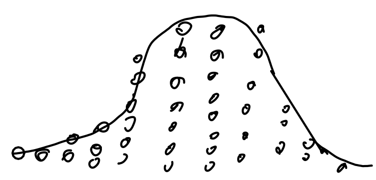



# 10. Konvolúció

A konvolúciós formulák független véletlen változók összegének eloszlását adják meg.

## 10.1. Diszkrét eset

Legyenek $X, Y$ független diszkrét véletlen változók $x_1, x_2, \ldots$, és $y_1, y_2, \ldots$ lehetséges értékekkel. Ekkor $Z = X + Y$ véletlen változó is diszkrét, lehetséges értékei $\{z_1, z_2, \ldots\} = \{x_i + y_j : i, j \in \mathbb{N}\}$. Továbbá, $Z$ eloszlása



$$\begin{aligned}\mathbf{P}(Z = z) &= \mathbf{P}(X + Y = z)\\ &= \sum_i \mathbf{P}(X = x_i, Y = z - x_i)\\ &= \sum_i \mathbf{P}(X = x_i)\mathbf{P}(Y = z - x_i).\end{aligned}$$