# – 2 –

## – DISCRETE RANDOM VARIABLES –

Often $S = \mathbb{R}, \mathbb{C}, \mathbb{R}^n, \mathbb{Z}$

<u>Def</u> A random variable (r.v.) $X$ on a prob. space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in a metric space $S$ is a <u>measurable</u> function

$X : \Omega \to S$

preimages of measurable sets (= Borel) in $S$ are events

e.g. $X^{-1}((-\infty, t)) \in \mathcal{F} \quad \forall t$

$$X^{-1}((-\infty,t)) = \{\omega \in \Omega, \quad X(\omega) \in (-\infty, t)\}$$

<u>E.g.</u> $\Omega = \{$ people in 21-325 $\} = \{$Tomasz, ... $\}$

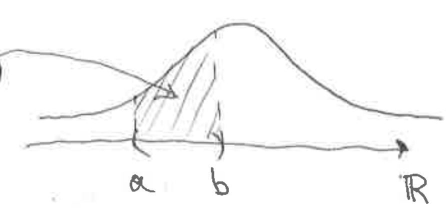
$X(\omega) =$ height of $\omega$

$\mathbb{P}($ a random person is taller than 7.7 ft $)$

$= \mathbb{P}(\{\omega \in \Omega, \quad X(\omega) > 7.7 \text{ ft}) \overset{\text{in short}}{=} \mathbb{P}(X > 7.7)$

The <u>law (distribution)</u> of $X$ is the prob. measure $\mu_X$ on $S$ defined by $\mu_X(A) = \mathbb{P}(X \in A)$

$\mu_X((a,b)) = \mathbb{P}(a < X < b)$
$= \%$ people with height between a and b



**Def** A discrete r.v. $X$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is a function $X : \Omega \to \mathbb{R}$ s.t. (i) $X(\Omega)$ is countable

(ii) $\forall x \in \mathbb{R}$ $X^{-1}(\{x\}) \in \mathcal{F}$.

**E.g.** roll a die, $X(\omega) = \omega$ (number rolled)
$\Omega = \{1,2,3,4,5,6\}$

- $\mathcal{F} = 2^\Omega$, $X$ is a discrete r.v.
- $\mathcal{F} = \{\emptyset, \Omega, \{1,3,5\}, \{2,4,6\}\}$, $X$ is not a r.v.
  (b/c $X^{-1}(\{1\}) = \{1\} \notin \mathcal{F}$).

The probability mass function (p.m.f.) of $X$ is $p_X : \mathbb{R} \to [0,1]$

$$p_X(x) = \mathbb{P}(X = x)$$

⚠

- $p_X(x) = 0$ unless $x \in X(\Omega) \equiv \operatorname{Im} X$ (Image)
- $$\sum_{x \in \operatorname{Im} X} p_X(x) = \sum_{x \in \operatorname{Im} X} \underbrace{\mathbb{P}(X = x)}_{\text{disjoint events}} \underset{\operatorname{Im} X \text{ countable}}{=} \mathbb{P}\left(\bigcup \{X = x\}\right)$$

$$= \mathbb{P}(\Omega) = 1.$$

Characterisation of p.m.f. **Thm** Let $q : \mathbb{R} \to [0,1]$ be s.t.

- the set $\{x \in \mathbb{R},\ q(x) > 0\}$ is countable
- $\sum q(x) = 1$.

Then there is a discrete r.v. $X$ whose p.m.f. is $q$ ( $p_X \equiv q$ ).

<u>Proof</u> Let $\{x \in \mathbb{R},\ q(x) > 0\} = \{x_1, x_2, \dots\}$.

Set $\Omega = \{1, 2, 3, \dots\}$, $\mathcal{F} = 2^\Omega$,

$$\mathbb{P}(A) = \sum_{i \in A} q(x_i), \qquad (\ \mathbb{P}(\{i\}) = q(x_i)\ )$$

$$X(i) = x_i$$

Then $X(\Omega) = \{x_1, x_2, \dots\}$ countable, since $\mathcal{F} = 2^\Omega$, regardless the definition of $X$, $X^{-1}(\{x\}) \in \mathcal{F}$, so $X$ is a discrete r.v. Its p.m.f. is

$$p_X(x) = \mathbb{P}(X = x) = \begin{cases} 0, & \text{if } x \notin \{x_1, x_2, \dots\} \\ \underbrace{q(x_i)}_{\mathbb{P}(X = x_i) = \mathbb{P}(\{i\})}, & \text{if } x = x_i \end{cases}$$

$\square$

This thm allows to forget about $\Omega$ – it suffices to specify the values of $X$ and their probabilities ( $p_X$ ).

<u>Important examples</u> (of discrete r.vs)

o) random sign (Rademacher r.v.) $\varepsilon$,
$\mathbb{P}(\varepsilon = -1) = \frac{1}{2} = \mathbb{P}(\varepsilon = +1)$.

1) Bernoulli distribution (biased coin) with param. $p \in [0,1]$

$$\mathbb{P}(X=0) = 1-p \quad , \quad \mathbb{P}(X=1) = p$$

(failure) (success)

Notation : $X \sim \text{Ber}(p)$

2) Binomial dist. with param's $p$ and $n$

$$X \in \{0, 1, 2, \dots, n\}$$

$$\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 \le k \le n$$

( this defines a pmf by the binomial thm:

$$\sum_{k=0}^{n} \mathbb{P}(X=k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p+(1-p))^n = 1 )$$

$$\underbrace{\text{HTHHTTT}}_{\text{exactly } k \text{ H's}} \quad \mathbb{P}(\text{ in } n \text{ tosses of a biased coin we get } k \text{ heads})$$

$$= \underbrace{\binom{n}{k}}_{\text{positions of } k \text{ H's}} \cdot p^k \cdot (1-p)^{n-k}$$

So $X$ = number of successes in $n$ indep. Bernoulli trials

Notation : $X \sim \text{Bin}(n,p)$

4) Geometric distribution with param$^{s}$ $p \in [0,1]$

$$X \in \{0, 1, 2, \dots\}$$

$$\mathbb{P}(X = k) = p(1-p)^{k-1}, \quad k \geq 1$$

(check $\sum_{k \geq 1} \mathbb{P}(X=k) = 1$).

$\mathbb{P}$( when tossing a biased coin, the first heads occurs only in $k^{th}$ toss )

$$= \mathbb{P}(\underbrace{TT\dots T}_{k-1} H) \overset{*)}{=} (1-p)^{k-1} \cdot p,$$

so $X$ = first success in Bernoulli trials

Notation $X \sim \text{Geom}(p)$

5) Poisson distribution with param $\lambda > 0$

$$X \in \{0, 1, 2, \dots\}$$

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

(check $\sum_{k=0}^{\infty} \mathbb{P}(X=k) = 1$ )

Notation: $X \sim \text{Poiss}(\lambda)$.

---

independent events

$$*) \quad \mathbb{P}(\underbrace{TT\dots T}_{k-1} H) = \mathbb{P}(\{1^{st} \text{ toss } T\} \cap \{2^{nd} \text{ toss } T\} \cap \dots \{k\text{-}1^{th} \text{ toss } T\} \cap \{k^{th} \text{ toss } H\})$$

6) Negative Binomial distribution with param's $n, p$

$X \in \{n, n+1, n+2, \dots\}$

$$\mathbb{P}(X=k) = \binom{k-1}{n-1} p^n (1-p)^{k-n}, \quad k \geq n$$

$\mathbb{P}$( when tossing a biased coin, $n^{th}$ heads at $k^{th}$ toss)

$$= \binom{k-1}{n-1} \cdot \underbrace{p^n}_{n \text{ heads}} \cdot \underbrace{(1-p)^{k-n}}_{k-n \text{ tails}}$$

need positions for $n-1$ heads → $\underbrace{\qquad\qquad}_{k-1}$ H

so $X$ = moment of $n^{th}$ success in Bernoulli trials

7) in general, if $A$ is an event,

then the indicator function

$$\mathbb{1}_A : \Omega \to \{0,1\}, \qquad \mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

is a $\{0,1\}$-valued discrete r.v. with p.m.f.

$$\mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(A)$$

$$\mathbb{P}(\mathbb{1}_A = 0) = 1 - \mathbb{P}(A),$$

so $\mathbb{1}_A \sim \text{Ber}(\mathbb{P}(A))$.

## Transformations

Suppose $X$ is a discrete r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$ and $g : \mathbb{R} \to \mathbb{R}$, Then

$$Y(\omega) = g(X(\omega)), \qquad \omega \in \Omega$$

defines a discrete r.v. (Check!)

E.g. • $g(x) = ax + b$, $\quad g(X) = aX + b$

In part. $\varepsilon$ = random sign, $X \sim \mathrm{Ber}(\frac{1}{2})$, $\quad \varepsilon = 2X - 1$

• $X$ = number on a fair die

$$Y = X \bmod 2 = \begin{cases} 0 & \text{if } X \text{ even} \\ 1 & \text{if } X \text{ odd} \end{cases}$$

$$Y \sim \mathrm{Ber}(\tfrac{1}{2})$$

The pmf of $Y = g(X)$:

$$p_Y(y) = \mathbb{P}(Y = y) = \mathbb{P}(g(X) = y) = \mathbb{P}(X \in g^{-1}(\{y\}))$$

$$= \sum_{x \in g^{-1}(\{y\})} \overset{p_X(x)}{\cancel{\mathbb{P}(X \in \{g}}} \,.$$

More general, $X_1, X_2, \dots, X_n$ discrete r.vs, $g : \mathbb{R}^n \to \mathbb{R}$,

$$Y = g(X_1, \dots, X_n) \quad \text{is a discrete r.vs}$$

In part., $Y = X_1 + \dots + X_n$, $\quad Y = X_1 \cdot \ldots \cdot X_n$.

## Expectation

The *expectation* of a discrete r.v. $X$ is

$$\mathbb{E}X = \sum_{x \in \operatorname{Im} X} x \cdot \mathbb{P}(X = x)$$

whenever this sum converges absolutely, i.e. $\sum |x| \cdot \mathbb{P}(X=x) < \infty$

E.g.
- $X \sim \operatorname{Ber}(p)$, $\quad \mathbb{E}X = 0 \cdot (1-p) + 1 \cdot p = p$.
- $X = \text{const}$ $\quad \mathbb{E}X = \text{const}$
- $\mathbb{E}\mathbb{1}_A = \mathbb{P}(A)$

**Thm** If $X, Y$ are discrete r.vs with expectations $\mathbb{E}X, \mathbb{E}Y$, then

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

**Proof** $X+Y$ takes values $x+y$, $x \in \operatorname{Im} X$, $y \in \operatorname{Im} Y$, so

$$\mathbb{E}(X+Y) = \sum_{\substack{x \in \operatorname{Im} X \\ y \in \operatorname{Im} Y}} (x+y)\, \mathbb{P}(X = x, Y = y)$$

$$= \sum_x \sum_y x\, \mathbb{P}(X=x, Y=y) + \sum_x \sum_y y\, \mathbb{P}(X=x, Y=y)$$

$\sum_y \mathbb{P}(X=x, Y=y) = \mathbb{P}(X=x)$

$$= \sum_x x\, \mathbb{P}(X=x) + \sum_y y\, \mathbb{P}(Y=y) = \mathbb{E}X + \mathbb{E}Y.$$

**Thm** $$\mathbb{E}g(X) = \sum_{x \in \operatorname{Im} X} g(x)\, \mathbb{P}(X = x).$$

E.g. $\mathbb{E}(aX+b) = a \cdot \mathbb{E}X + b$

**Proof** $g(X)$ takes the value $g(x)$ with prob. $\mathbb{P}(X=x)$. $\square$

⚠ In general $\mathbb{E}(X \cdot Y) \neq \mathbb{E}X \cdot \mathbb{E}Y$

E.g. $X = Y \sim \operatorname{Ber}(p)$ $\quad \mathbb{E}X \cdot Y = \mathbb{E}X^2 = 0^2 \cdot (1-p) + 1^2 \cdot p = p$

$\mathbb{E}X \cdot \mathbb{E}Y = (\mathbb{E}X)^2 = p^2$.

E.g. $\quad S \sim \mathrm{Bin}(n,p)$

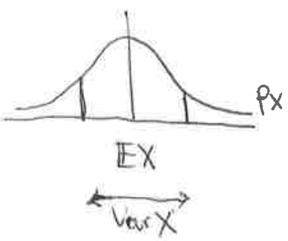
by def $\quad \mathbb{E}S = \sum_{k=0}^{n} k \cdot \mathbb{P}(S=k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$

$\overset{?}{=}$ (not difficult but not instant)

instant $\qquad S =$ number of successes $= X_1 + \dots + X_n$, $\quad X_i \sim \mathrm{Ber}(p)$

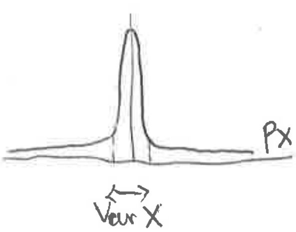
$$\mathbb{E}S = \mathbb{E}(X_1 + \dots + X_n) = \mathbb{E}X_1 + \dots + \mathbb{E}X_n = n \cdot \mathbb{E}X_1 = n \cdot p.$$

Variance $\quad \mathrm{Var}X = \mathbb{E}\underbrace{(X - \mathbb{E}X)^2}_{\geq 0,\ \text{small when } X \text{ close to } \mathbb{E}X}$




so $\quad \mathbb{E}X =$ " centre / average / mean of $X$ "

$\mathrm{Var}X =$ " dispersion of $X$ from $\mathbb{E}X$ "




Fact $\quad \mathrm{Var}X = \mathbb{E}\left(X^2 - 2X \cdot \mathbb{E}X + (\mathbb{E}X)^2\right)$

$$= \mathbb{E}X^2 - \mathbb{E}(\underbrace{2\mathbb{E}X}_{\text{const}} \cdot X) + \mathbb{E}\underbrace{(\mathbb{E}X)^2}_{\text{const}}$$

$$= \mathbb{E}X^2 - 2\,\mathbb{E}X \cdot \mathbb{E}X + (\mathbb{E}X)^2$$

$$= \mathbb{E}X^2 - (\mathbb{E}X)^2.$$

## Conditional expectation

If $X$ is a discrete r.v. and $B$ an event s.t. $\mathbb{P}(B)>0$,

then the conditional expectation of $X$ given $B$ is

$$\mathbb{E}(X \mid B) = \sum_{x \in \operatorname{Im} X} x \, \mathbb{P}(X=x \mid B)$$

**Thm** If $B_1, B_2, \dots$ form a partion, $\forall i \; \mathbb{P}(B_i)>0$, then

$$\mathbb{E}X = \sum_{i=1}^{\infty} \mathbb{E}(X \mid B_i) \cdot \mathbb{P}(B_i)$$

**Proof** 

$$\text{RHS} = \sum_i \sum_x x \cdot \mathbb{P}(X=x \mid B_i) \cdot \mathbb{P}(B_i)$$

$$= \sum_x x \cdot \underbrace{\sum_i \mathbb{P}(X=x \mid B_i) \cdot \mathbb{P}(B_i)}_{\mathbb{P}(X=x)} = \mathbb{E}X. \; \square$$

**E.g.** A biased coin is tossed repeatedly (heads with prob. $p$)

Find the expected length of the initial run.

$H$ = $1^{st}$ toss heads, $X$ = length of the initial run

$\{H, H^c\}$ - partition

$H \overbrace{H \dots H}^{k-1} T \qquad \mathbb{P}(X=k \mid H) = p^{k-1}(1-p) \qquad k = 1, 2, \dots$

$T \overbrace{T \dots T}^{k-1} H \qquad \mathbb{P}(X=k \mid H^c) = p(1-p)^{k-1} \qquad \left(\sum_{k=0}^{\infty} p^k\right)'$

$$\mathbb{E}(X \mid H) = \sum k \cdot \mathbb{P}(X=k \mid H) = (1-p) \sum_{k=1}^{\infty} kp^{k-1} = (1-p)\frac{1}{(1-p)^2}$$

$$= \frac{1}{1-p}$$

Similarly, $\mathbb{E}(X \mid H^c) = \frac{1}{p}$, so

$$\mathbb{E}X = \mathbb{E}(X \mid H)\,\mathbb{P}(H) + \mathbb{E}(X \mid H^c)\,\mathbb{P}(H^c)$$

$$= \frac{1}{1-p} \cdot p + \frac{1}{p}(1-p) = \frac{p}{1-p} + \frac{1-p}{p} \quad (\geq 2)$$

## -3-

## - Multivariate distributions and independence -

If $X, Y$ are discrete r.vs (on $(\Omega, \mathcal{F}, \mathbb{P})$) then the pmf of the vector $(X,Y)$ is

$$p_{(X,Y)}(x,y) = \mathbb{P}(\underbrace{X=x,\ Y=y}_{\{\omega \in \Omega,\ X(\omega)=x \text{ and } Y(\omega)=y\}})$$

⚠ Two properties

- $p_{(X,Y)} \geq 0$, if $x \notin \operatorname{Im} X$ or $y \notin \operatorname{Im} Y$, then $p_{(X,Y)}(x,y) = 0$
- $$\sum_{\substack{x \in \operatorname{Im} X \\ y \in \operatorname{Im} Y}} p_{(X,Y)}(x,y) = 1$$

Marginals $p_X, p_Y$ are found by taking sums

$$p_X(x) = \sum_{y \in \operatorname{Im} Y} p_{(X,Y)}(x,y)$$

because: $p_X(x) = \mathbb{P}(X=x) = \mathbb{P}(X=x, \Omega) = \mathbb{P}(X=x, \bigcup_{y \in \mathrm{Im}Y} \{Y=y\})$

$$= \mathbb{P}\Big(\bigcup_{y\in \mathrm{Im}Y} \{X=x, Y=y\}\Big) = \sum_{y \in \mathrm{Im}Y} \mathbb{P}(X=x, Y=y).$$

The same for $p_Y$.

The same for more variables

$\vec{X} = \underbrace{(X_1, \ldots, X_n)}_{\text{discrete r.vs}}$ discrete random vector

$$p_{\vec{X}}(x_1, \ldots, x_n) = \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n).$$

Thm If $g: \mathbb{R}^n \to \mathbb{R}$, $(X_1, \ldots, X_n)$ is a dis. r. vec, then

$$\mathbb{E}g(X_1, \ldots, X_n) = \sum_{\substack{x_1 \in \mathrm{Im}X_1 \\ \vdots \\ x_n \in \mathrm{Im}X_n}} g(x_1, \ldots, x_n)\, \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n).$$

Cor $\mathbb{E}\left(\sum_1^n a_i X_i\right) = \sum_1^n a_i \mathbb{E}X_i$.

E.g. Planet with $n$ days / year, $k$ people

$X =$ number of pairs of people sharing birthday

$\mathbb{E}X$ ? $\qquad X_{ij} = \begin{cases} 1 & i, j \text{ persons share b-day} \\ 0 & \text{o/w} \end{cases}$

$X = \sum_{i<j} X_{ij} \qquad \mathbb{E}X_{i,j} = \mathbb{P}(X_{ij} = 1) = \frac{1}{n}$

$$\mathbb{E}X = \sum_{i<j} \mathbb{E}X_{ij} = \sum_{i<j} \frac{1}{n} = \binom{k}{2}\frac{1}{n}.$$

Recall: events are indep if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

Discrete r.vs $X, Y$ are indep. if

$\forall x, y \in \mathbb{R}$ events $\{X = x\}$ and $\{Y = y\}$ are indep, i.e.

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\,\mathbb{P}(Y = y).$$

In terms of pmf

The joint pmf factorises

$$\forall x, y \in \mathbb{R} \quad p_{(X,Y)}(x, y) = p_X(x)\, p_Y(y)$$

The converse true!

<u>Thm</u> Discrete r.vs $X, Y$ are indep. if and only if there exist functions $f, g : \mathbb{R} \to \mathbb{R}$ s.t.

(*) $\quad p_{X,Y}(x, y) = f(x) g(y) \quad$ for all $x, y \in \mathbb{R}$.

<u>Proof</u> "$\Rightarrow$" clear by def.

"$\Leftarrow$" by taking in (*) $\left\{ \begin{array}{ll} \sum_{y \in \mathrm{Im} Y} : & p_X(x) = f(x) \sum_{y \in \mathrm{Im} Y} g(y) \\ \sum_{x \in \mathrm{Im} X} : & p_Y(y) = g(y) \sum_{x \in \mathrm{Im} X} f(x) \\ \sum_{x \in \mathrm{Im} X, y \in \mathrm{Im} Y} : & 1 = \sum_x f(x) \cdot \sum_y g(y). \end{array} \right.$

Therefore,
$$p_X(x)\, p_Y(y) = f(x) g(y) \sum_{y'} g(y') \sum_{x'} f(x')$$
$$= f(x) g(y) = p_{(X,Y)}(x, y). \quad \square$$

E.g. $\mathbb{P}(X=k, Y=\ell) = \frac{1}{2^{k+\ell}} = \frac{1}{2^k} \cdot \frac{1}{2^\ell}$ , $k,\ell \geq 1$

$\Rightarrow$ $X, Y$ indep. ( notation $X \perp\!\!\!\perp Y$ )

Thm If $X, Y$ are discrete r.v. with mean $\mathbb{E}X, \mathbb{E}Y$, then

$$\mathbb{E}XY = \mathbb{E}X \cdot \mathbb{E}Y.$$

Proof
$$\mathbb{E}XY = \sum xy\, \mathbb{P}(X=x, Y=y) = \sum xy\, \mathbb{P}(X=x)\, \mathbb{P}(Y=y)$$
$$= \sum x \mathbb{P}(X=x) \sum y \mathbb{P}(Y=y) = \mathbb{E}X \cdot \mathbb{E}Y. \quad \square$$

⚠ The converse false ! $X = 0, -1, +1$ with prob. $1/3$
$Y = |X|$

- $\mathbb{E}XY = 0 = \mathbb{E}X \cdot \mathbb{E}Y$
- $X, Y$ not indep.

if and only if

Thm Discrete r.vs $X, Y$ are indep. iff for all $f, g: \mathbb{R} \to \mathbb{R}$

$$\mathbb{E} f(X) g(Y) = \mathbb{E} f(X)\, \mathbb{E} g(Y)$$

Proof "$\Rightarrow$" the same comp. as for $\mathbb{E}XY = \mathbb{E}X\, \mathbb{E}Y$

"$\Leftarrow$" Fix $a, b \in \mathbb{R}$. Set $f(x) = 1_{\{a\}}(x)$
$g(x) = 1_{\{b\}}(x)$

$\mathbb{E} f(X) = \mathbb{E}\, 1_{\{a\}}(X) = \mathbb{P}(X=a)$ , $\mathbb{E} g(Y) = \mathbb{P}(Y=b)$,

$\mathbb{E} f(X) g(Y) = \mathbb{P}(X=a, Y=b)$.

The same for more variables

$X_1, X_2, \dots, X_n$ indep if $p_{(X_1,\dots,X_n)}(x_1,\dots,x_n) = \prod_1^n p_{X_i}(x_i)$,

then $\mathbb{E}\left(\prod_1^n X_i\right) = \prod_1^n \mathbb{E}X_i$ , etc.

<u>Sums of r.vs</u>

If $X, Y$ are two discrete r.vs what is the pmf of $Z = X+Y$?

$$p_Z(z) = \mathbb{P}(Z=z) = \mathbb{P}(\underbrace{X+Y=z}_{\text{for some } x \;\; X=x,\, Y=z-x}) = \sum_{x \in \operatorname{Im} X} \mathbb{P}(X=x, Y=z-x)$$

$$= \sum_{x \in \operatorname{Im} X} p_{(X,Y)}(x, z-x)$$

$$= \dots = \sum_{y \in \operatorname{Im} Y} p_{(X,Y)}(z-y, y).$$

If $X, Y$ indep, then

$$p_Z(z) = \sum_x p_X(x)\, p_Y(z-x)$$

so $p_Z = p_X \underset{\text{convolution}}{*} p_Y$

$(a_n)_{n=1}^\infty$, $(b_n)_{n=1}^\infty$ two seqs

$(a_n * b_n)_m = \sum_\ell a_\ell b_{m-\ell}$.

<u>Thm</u> $X, Y$ indep. $\Rightarrow$ $\operatorname{Var}(X+Y) = \operatorname{Var} X + \operatorname{Var} Y$

<u>Proof</u> $\operatorname{Var}(X+Y) = \mathbb{E}\left((X+Y) - \mathbb{E}(X+Y)\right)^2 = \mathbb{E}\left((X-\mathbb{E}X) + (Y-\mathbb{E}Y)\right)^2$

$$= \mathbb{E}(X-\mathbb{E}X)^2 + \mathbb{E}(Y-\mathbb{E}Y)^2 + 2\,\underbrace{\mathbb{E}\,\underbrace{(X-\mathbb{E}X),(Y-\mathbb{E}Y)}_{\text{indep}}}_{= \mathbb{E}(X-\mathbb{E}X)\cdot\mathbb{E}(Y-\mathbb{E}Y) = 0.}$$