Math 170E: Winter 2023

Lecture 18, Fri 24th Feb

Bivariate distributions of the discrete type continued

Last time: (X,Y)

Definition 4.1: Let X, Y be a pair of discrete random variables taking values in sets $S_X, S_Y \subset \mathbb{R}$, respectively and let $S = S_X \times S_Y$.

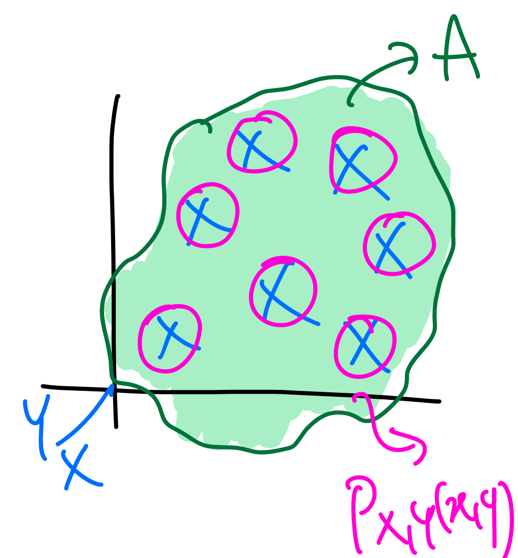
- We define the joint probability mass function of X, Y to be the function $p_{X,Y} : S \to [0,1]$ by

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

Proposition 4.2: Let X, Y be a pair of discrete random variables taking values in sets $S_X, S_Y \subset \mathbb{R}$, respectively and let $S = S_X \times S_Y$.

If X, Y have joint PMF $p_{X,Y}(x,y)$ and $A \subseteq \mathbb{R}^2$, then

$$\mathbb{P}((X,Y) \in A) = \sum_{(x,y) \in A \cap S} p_{X,Y}(x,y)$$



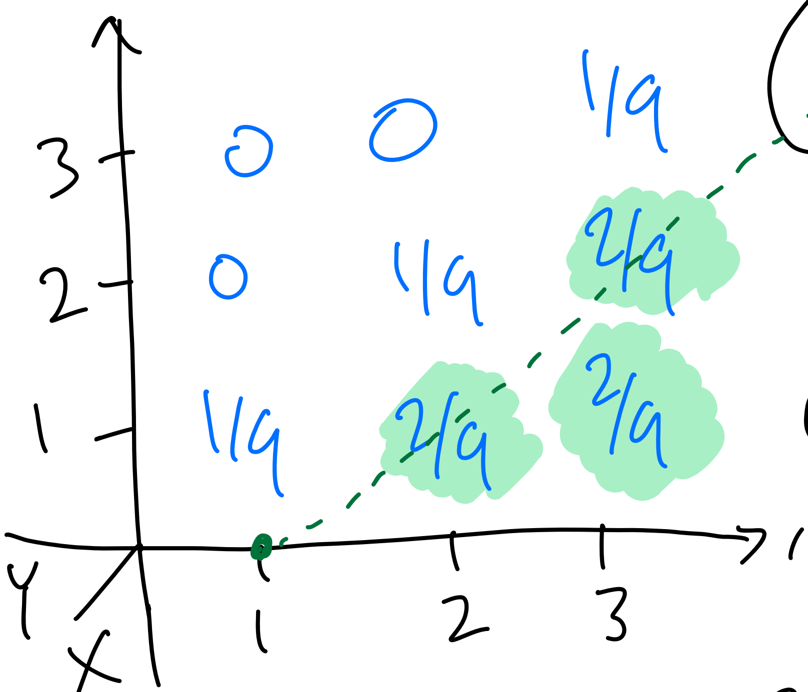
Normalisation condition:

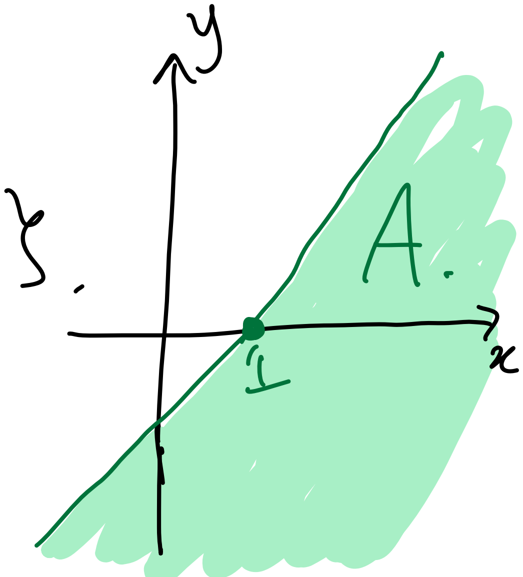
$$\sum_{(x,y) \in S} p_{X,Y}(x,y) = 1$$

Example 2:

- You choose two numbers at random from the set $\{1, 2, 3\}$
- Let X be the larger and Y be the smaller of these two numbers
- What is $\underbrace{\mathbb{P}(X - Y \geq 1)}$? $\rightsquigarrow \mathbb{P}((X,Y) \in A)$

$$A = \{(x,y) \in \mathbb{R}^2 : x - y \geq 1\}.$$





$$= \sum_{(x,y) \in A \cap S} p_{X,Y}(x,y)$$

$$= \sum_{(x,y) \in \{(2,1),(3,2),(3,1)\}} p_{X,Y}(x,y)$$

$$A \cap S = \{(2,1),(3,2),(3,1)\}$$

$$= p_{X,Y}(2,1) + p_{X,Y}(3,2) + p_{X,Y}(3,1)$$

$$= 2/9 + 2/9 + 2/9 = 2/3.$$

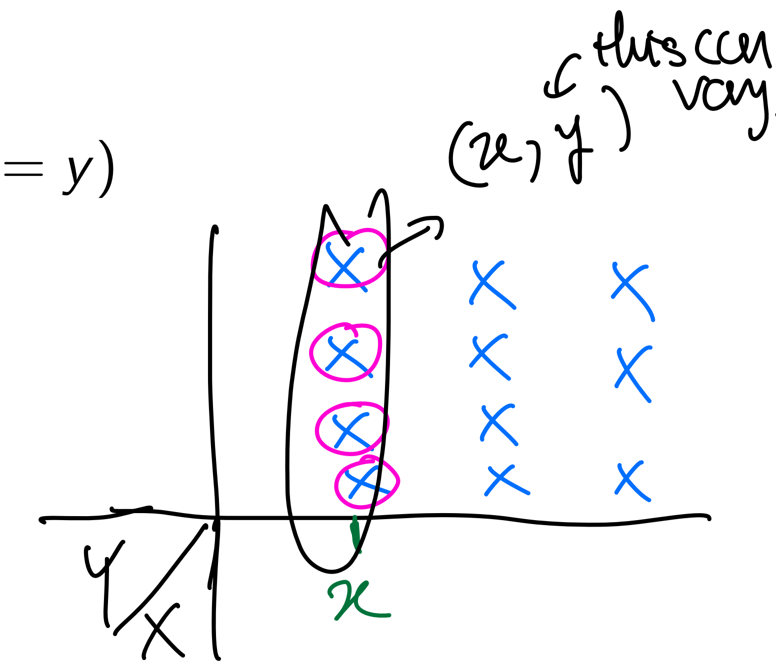
- Let X, Y be a pair of discrete random variables taking values in sets $S_X, S_Y \subset \mathbb{R}$, respectively and let $S = S_X \times S_Y$.

- We define the marginal probability mass function of X to be the function $p_X : S_X \to [0, 1]$ given by

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

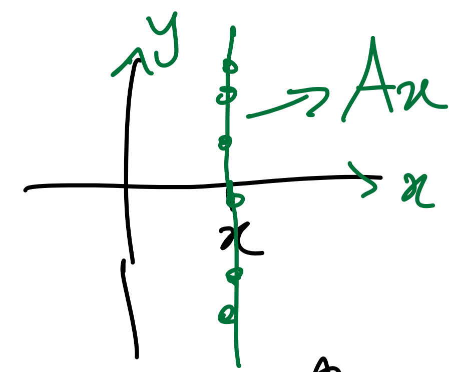
- We define the marginal probability mass function of Y to be the function $p_Y : S_Y \to [0, 1]$ given by

$$p_Y(y) = \mathbb{P}(Y = y)$$



Proposition 4.4: Let X, Y be discrete r.vs taking values in sets $S_X, S_Y \subset \mathbb{R}$.

If X, Y have joint PMF $p_{X,Y}(x,y)$, then



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

$$p_Y(y) = \sum_{x \in S_X} p_{X,Y}(x,y)$$

Proof: Focus on p_X as the case for p_Y is similar.

Fix $x \in S_X$, let $A_x = \{(x,y) : y \in \mathbb{R}\} \subseteq \mathbb{R}^2$. this is a line

$$p_X(x) = \mathbb{P}(X = x) = \mathbb{P}((X,Y) \in A_x).$$

$$= \mathbb{P}((X,Y) \in A_x \cap S).$$

$A_x \cap S$
$= \{(x,y) : y \in S_Y\}.$

$$= \sum_{(t,y) \in \underbrace{A_x \cap S}} p_{X,Y}(t,y).$$

if

$(t,y) \in A_x \cap S$, then $y \in S_Y$, $t = x$

$$= \sum_{y \in S_Y} P_{X,Y}(x,y).$$

Proposition 4.5: Let X, Y be discrete r.vs taking values in sets $S_X, S_Y \subset \mathbb{R}$.

If X has marginal PMF $p_X(x)$ and Y has marginal PMF $p_Y(y)$, then:

$$\sum_{x \in S_X} p_X(x) = 1$$

$$\sum_{y \in S_Y} p_Y(y) = 1$$

Proof:

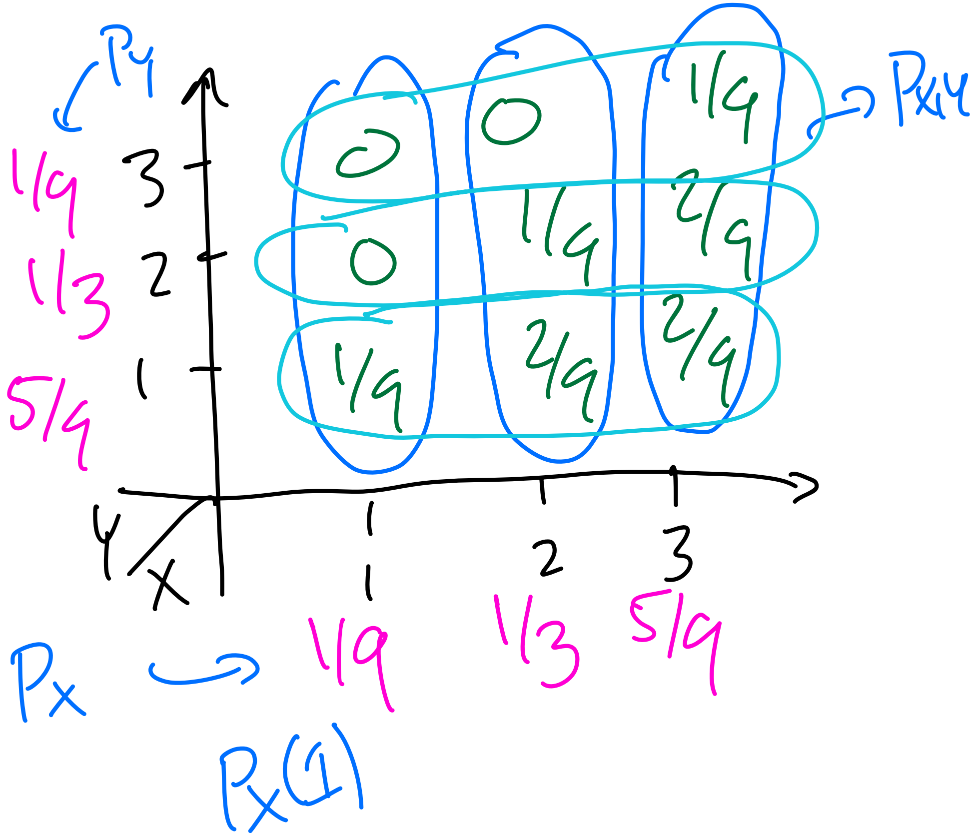
$$\sum_{x \in S_X} p_X(x) = \sum_{x \in S_X} \left(\sum_{y \in S_Y} p_{X,Y}(x,y) \right) = \sum_{(x,y) \in \underbrace{S_X \times S_Y}_{= S}} p_{X,Y}(x,y)$$

$$= 1$$

normalisation condition.

Example 3:

- You choose two numbers at random from the set $\{1, 2, 3\}$
- Let X be the larger and Y be the smaller of these two numbers
- What are the marginal PMFs of X and Y?



$$P_X(x) = \begin{cases} 1/9 & \text{if } x = 1 \\ 1/3 & \text{if } x = 2 \\ 5/9 & \text{if } x = 3. \end{cases}$$

$$P_Y(y) = \begin{cases} 5/9 & \text{if } y = 1 \\ 1/3 & \text{if } y = 2 \\ 1/9 & \text{if } y = 3. \end{cases}$$

Definition 4.6:

- Let X, Y be discrete r.vs taking values in sets $S_X, S_Y \subset \mathbb{R}$ and let $S = S_X \times S_Y$.

- We say that random variables X, Y are independent ~~if~~ if the events $\{X = x\}$ and $\{Y = y\}$ are independent for all $(x, y) \in S$.

- Equivalently, we have

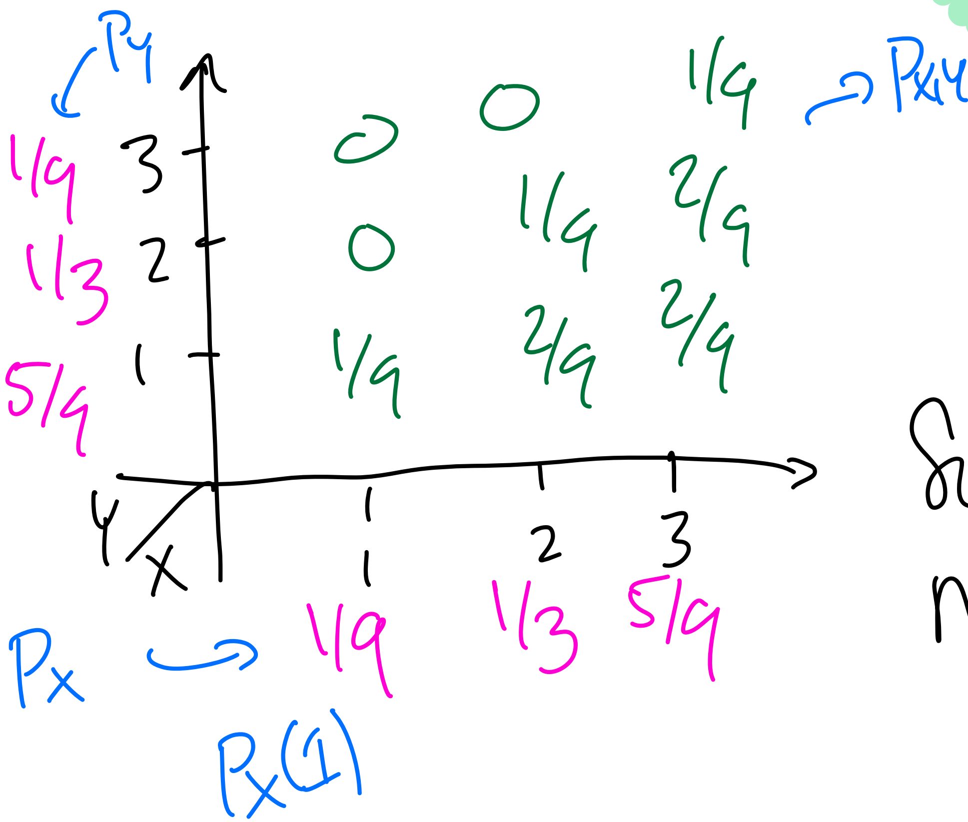
$$\boxed{p_{X,Y}(x, y) = p_X(x)p_Y(y) \quad \text{for all } (x, y) \in S.}$$

$$p_{X,Y}(x,y) = \mathbb{P}(X=x, Y=y) = \mathbb{P}(\{X=x\} \cap \{Y=y\})$$
$$\overset{\text{indep}}{=} \mathbb{P}(X=x)\,\mathbb{P}(Y=y)$$
$$= p_X(x)\,p_Y(y).$$

$\triangle$ To prove (X,Y) are <u>not</u> indep, we just find particular value $(x_*, y_*) \in S$ for which $p_{X,Y}(x_*, y_*) \neq p_X(x_*)\,p_Y(y_*)$

Example 4:

- You choose two numbers at random from the set $\{1, 2, 3\}$
- Let X be the larger and Y be the smaller of these two numbers
- Are X and Y independent?



$$\frac{2}{9} = P_{X,Y}(3,3)$$

$$\neq P_X(3)P_Y(3) = \frac{5}{9} \times \frac{1}{9}$$

$$= \frac{5}{81}$$

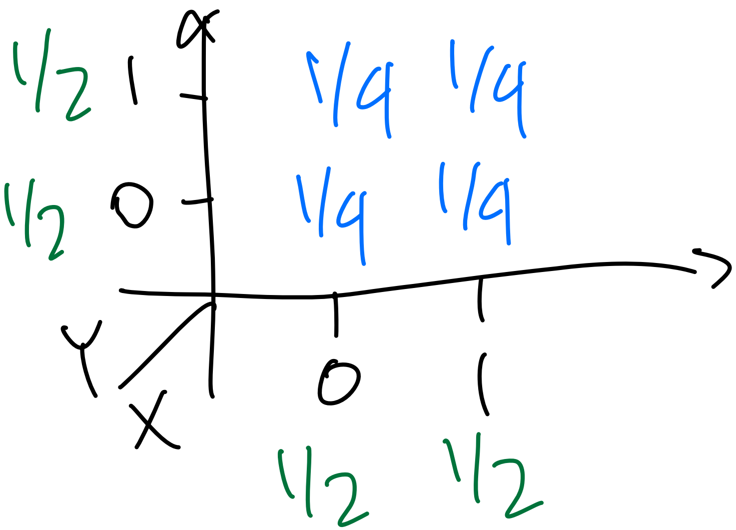
So X & Y are not independent!

Example 5:

- You flip two fair coins
- Let X be 1 if the first flip is HEADS and 0 if TAILS
- Let Y be 1 if the second flip is HEADS and 0 if TAILS
- Are X and Y independent?

$X \sim \text{Bernoulli}(1/2)$

$Y \sim$



So X & Y are independent because

$$P_{X,Y}(x,y) = P_X(x)P_Y(y)$$

for all

$$(x,y) \in \{0,1\} \times \{0,1\}.$$

Definition 4.7 Let X, Y be a pair of discrete r.v.s taking values in sets $S_X, S_Y \subset \mathbb{R}$, let $S = S_X \times S_Y$ and $p_{X,Y}(x, y)$ be the joint PMF of X, Y.

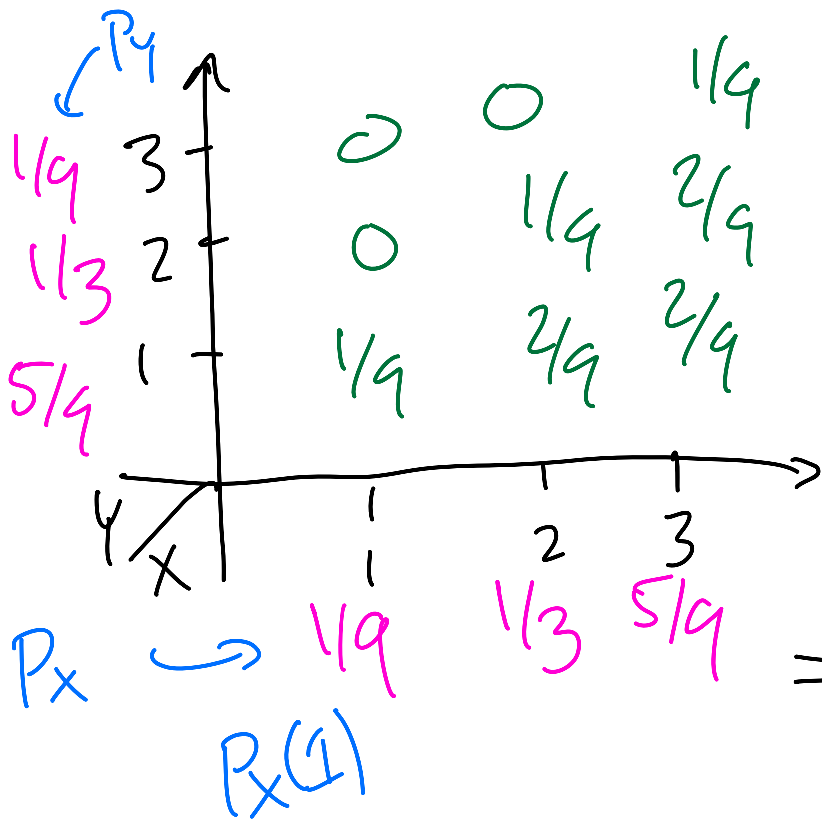
If $g : S \to \mathbb{R}$, then the expected value of g(X,Y) is

$$\mathbb{E}[g(X, Y)] = \sum_{(x,y)\in S} g(x, y)p_{X,Y}(x, y)$$

1-variable: $\mathbb{E}[g(X)] = \sum_{x \in S_X} g(x) p_X(x)$.

Example 6:

- You choose two numbers at random from the set $\{1, 2, 3\}$
- Let X be the larger and Y be the smaller of these two numbers
- What is $\mathbb{E}[XY]$? $\rightsquigarrow g(x,y) = xy.$



$$\mathbb{E}[XY] = \sum_{(x,y)\in S} xy \, P_{X,Y}(x,y).$$

$$= 1\times 1\times \tfrac{1}{9} + 2\times 2\times \tfrac{1}{9} + 3\times 3\times \tfrac{1}{9}$$

$$+ 2\times 1\times \tfrac{2}{9} + 3\times 1\times \tfrac{2}{9} + 3\times 2\times \tfrac{2}{9}$$

$$= \frac{1+4+9+4+6+12}{9}$$

$$= 36/9 = 4.$$

Proposition 4.8: Let X, Y be a pair of discrete r.vs taking values in sets $S_X, S_Y \subset \mathbb{R}$, respectively and let $S = S_X \times S_Y$ and $p_{X,Y}(x, y)$ be the joint PMF of X, Y.

Let $a, b \in \mathbb{R}$ and $g : S \to \mathbb{R}$ and $h : S \to \mathbb{R}$. Then,

$$\mathbb{E}[ag(X, Y) + bh(X, Y)] = a\mathbb{E}[g(X, Y)] + b\mathbb{E}[h(X, Y)]$$

If $g(x, y) \leq h(x, y)$ for all $(x, y) \in S$, then

$$\mathbb{E}[g(X, Y)] \leq \mathbb{E}[h(X, Y)]$$

Proof: Same as in 1-variable case.

Proposition 4.9: Let X, Y be discrete r.vs taking values in sets $S_X, S_Y \subset \mathbb{R}$.

Let $g : S_X \to \mathbb{R}$ and $h : S_Y \to \mathbb{R}$. Then,

$$\mathbb{E}[g(X)] = \sum_{x \in S_X} g(x) p_X(x)$$

$$\mathbb{E}[h(Y)] = \sum_{y \in S_Y} h(y) p_Y(y)$$

Proof:

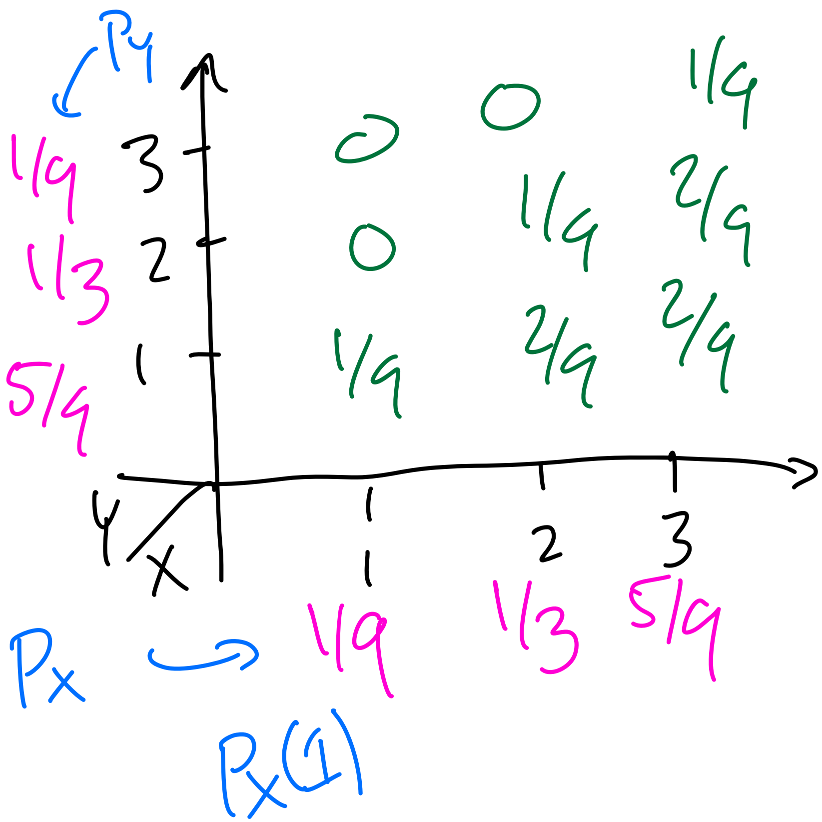
$$\mathbb{E}[g(X)] = \sum_{(x,y) \in S_X \times S_Y} g(x)\, p_{X,Y}(x,y)$$

$$= \sum_{x \in S_X} \sum_{y \in S_Y} g(x)\, p_{X,Y}(x,y)$$

$$= \sum_{x \in S_X} g(x) \underbrace{\sum_{y \in S_Y} p_{X,Y}(x,y)}_{= p_X(x)} = \sum_{x \in S_X} g(x) p_X(x).$$

Example 6:

- You choose two numbers at random from the set $\{1, 2, 3\}$
- Let X be the larger and Y be the smaller of these two numbers
- What is $\mathbb{E}[X]$?



$$\mathbb{E}[X] = \sum_{x \in \{1,2,3\}} x\, p_X(x)$$

$$= 1 \times \frac{1}{9} + 2 \times \frac{3}{9} + 3 \times \frac{5}{9}$$

$$= \frac{1 + 6 + 15}{9} = \frac{22}{9}.$$

Proposition 4.10: Let X, Y be independent discrete r.vs taking values in sets $S_X, S_Y \subset \mathbb{R}$.

$g(x) = h(x) = x$

Let $g : S_X \to \mathbb{R}$ and $h : S_Y \to \mathbb{R}$. Then,

$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$.

if indep

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)]$$

Proof:

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

$f(x,y) = g(x)h(y)$

$$= \sum_{(x,y)\in S} f(x,y)\, P_{X,Y}(x,y).$$

$$= \sum_{(x,y)\in S} g(x)\, h(y)\, P_{X,Y}(x,y)$$

$$= \sum_{x\in S_X} g(x) \sum_{y \in S_Y} h(y) \boxed{P_{X,Y}(x,y)}.$$

X, Y are indep

$= P_X(x)\, P_Y(y)$

$$= \sum_{x \in S_X} g(x) p_X(x) \underbrace{\sum_{y \in S_Y} h(y) p_Y(y)}_{\mathbb{E}[h(Y)]}.$$

$$= \sum_{x \in S_X} g(x) p_X(x) \, \mathbb{E}[h(Y)]$$

$$= \mathbb{E}[h(Y)] \underbrace{\left(\sum_{x \in S_X} g(x) p_X(x) \right)}_{\mathbb{E}[g(X)]}.$$

$$= \mathbb{E}[h(Y)] \, \mathbb{E}[g(X)].$$

Example 7:

- You flip a fair coin twice
- Let X be 1 if the first flip is HEADS and 0 if TAILS
- Let Y be 1 if the second flip is HEADS and 0 if TAILS
- What is $\text{var}(X+Y)$?

$X, Y \sim \text{Bernoulli}(1/2)$

X, Y are $\underline{\underline{\text{indep}}}$

$X+Y \sim \text{Bin}(2, 1/2)$

$\text{var}(X+Y) = 2 \times 1/2 \times 1/2$
$= 1/2.$

$$\text{var}(X+Y) = \mathbb{E}[(X+Y)^2] - \mathbb{E}[X+Y]^2$$

$$= \mathbb{E}[X^2 + 2XY + Y^2] - (\mathbb{E}[X] + \mathbb{E}[Y])^2$$

$$= \mathbb{E}[X^2] + 2\mathbb{E}[XY] + \mathbb{E}[Y^2] - \mathbb{E}[X]^2 - 2\mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[Y]^2$$

$$= \text{var}(X) + \text{var}(Y) + 2\left(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\right).$$

If X, Y are independent, then
$$E[XY] = E[X]E[Y]$$

If X, Y are indep,

$$Var(X+Y) = Var(X) + Var(Y)$$

$$\left(Var(X+Y) = 1/4 + 1/4 = 1/2\right).$$

$X, Y \sim Bernoulli(p)$ indep, why if $Z \sim Bin(2, p)$,

$$E[Z] = 2p = 2E[X]$$

$$Var(Z) = 2p(1-p)$$
$$= 2Var(X).$$

Proposition 4.11: (The Cauchy-Schwarz inequality)

Let X, Y be discrete random variables. Then,

$$\left|\mathbb{E}[XY]\right| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}$$

Proof:

$$\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \underbrace{\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]}_{\text{if } X, Y \text{ are not indep, how big can this be?}}.$$

Application:

$$\left|\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]\right| \leq \sqrt{\mathbb{E}[(X-\mathbb{E}[X])^2]\,\mathbb{E}[(Y-\mathbb{E}[Y])^2]}$$

$$\le \overbrace{\sqrt{\underbrace{\mathrm{var}(x)}_{}\mathrm{var}(y)}}$$

$\mathrm{Var}(X)$ $\mathrm{Var}(Y)$

$$\le \sqrt{\mathrm{var}(x)\,\mathrm{var}(y)}\,.$$

$$\sum_{(x,y)\in S_X\times S_Y} a(x,y) = \sum_{x\in S_X}\left(\sum_{y\in S_Y} a(x,y)\right) = \sum_{y\in S_Y}\left(\sum_{x\in S_X} a(x,y)\right).$$

$$= \sum_{x\in S_X}\left(a(x,y_1)+a(x,y_2)+\cdots\right)$$

$$\begin{aligned} &= a(x_1,y_1)+a(x_1,y_2)+\cdots \\ &\quad + a(x_2,y_1)+a(x_2,y_2)+\cdots \\ &\quad + a(x_3,y_1)+\cdots \end{aligned}$$

Proof of Cauchy-Schwarz inequality:

There are many proofs but here is a simple one.

- If $\mathbb{E}[Y^2]=0$, then $Y=0$ and the inequality is clearly true.

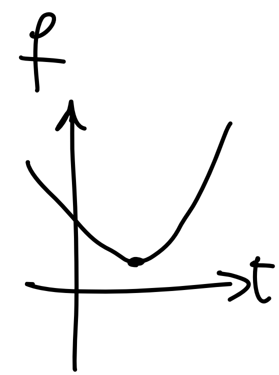
So we may assume that $\mathbb{E}[Y^2]\neq 0$.

Consider the one-variable function:

$$f(t) := \mathbb{E}[(X-tY)^2], \quad t\in\mathbb{R}.$$

We have:

- $f(t) = \mathbb{E}[X^2] - 2t\,\mathbb{E}[XY] + t^2\,\mathbb{E}[Y^2]$ ← a quadratic in t



- $f(t) \geq 0$ for all $t\in\mathbb{R}$.

- $f(0) = \mathbb{E}[X^2]$.

We look for the minimum of f. ?

$$f'(t) = -2\mathbb{E}[XY] + 2t\,\mathbb{E}[Y^2]$$

Critical points for f: $t_* = \dfrac{\mathbb{E}[XY]}{\mathbb{E}[Y^2]}$ → non-zero

So $$0 \leq f(t_*) = \mathbb{E}[X^2] - \frac{2\mathbb{E}[XY]^2}{\mathbb{E}[Y^2]} + \frac{\mathbb{E}[XY]^2}{\mathbb{E}[Y^2]}$$

$$= \mathbb{E}[X^2] - \frac{\mathbb{E}[XY]^2}{\mathbb{E}[Y^2]}$$

$$\Rightarrow \quad \mathbb{E}[XY]^2 \leq \mathbb{E}[X^2]\,\mathbb{E}[Y^2] \quad //$$