CSE 312

Foundations of Computing II

19: More Joint Distributions; Law of Total Expectation

Agenda

- Bit more on joint distributions
- Law of Total Expectation and LTP (continuous)

Discrete to Continuous

	Discrete	Continuous
PMF/PDF	$p_X(x) = P(X = x)$	$f_X(x) \neq P(X = x) = 0$
CDF	$F_X(x) = \sum_{t \leq x} p_X(t)$	$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$
Normalization	$\sum_x p_X(x) = 1$	$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$
Expectation	$\mathbb{E}[g(X)] = \sum_x g(x)\, p_X(x)$	$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx$

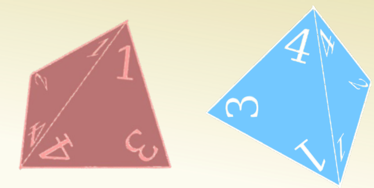
Support/range $\Omega_X = \{x \mid p_X(x) > 0\}$ $\Omega_X = \{x \mid f_X(x) > 0\}$

$P(X \in [x, x+dx]) \approx f_X(x)\,dx$

Jointly distributed random variables

X	Y
Grade on exam	Amount of sleep the night before
Performance of Microsoft stock	Performance of Amazon stock
Grade of person A on exam in 312	Grade of person B on exam in 332
Number of job interviews to get a job	State of the economy

X_1 blood pressure
X_2 temperature
X_3 blood glucose
X_4 kidney function



X is the number of of job interviews you need to do to get a job.

Y is the state of the economy. (Bad = 1, Below Avg = 2, Above Avg =3, Excellent = 4

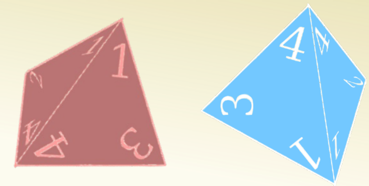
1 2 3 4

X/Y	Bad	Below Avg	Above Avg	Excellent
1	0.01	0.05	0.08	0.12
2	0.03	0.05	0.06	0.12
3	0.03	0.12	0.04	0.05
4	0.08	0.12	0.03	0.01

- Valid joint probability mass function?

Each entry gives prob mass fn at that (x, y) pair

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



X is the number of of job interviews you need to do to get a job.

Y is the state of the economy. (Bad = 1, Below Avg = 2, Above Avg =3, Excellent = 4

X/Y	Bad	Below Avg	Above Avg	Excellent
1	0.01	0.05	0.08	0.12
2	0.03	0.05	0.06	0.12
3	0.03	0.12	0.04	0.05
4	0.08	0.12	0.03	0.01

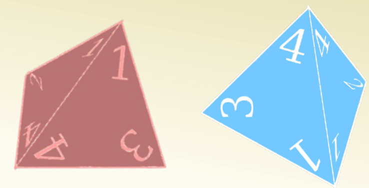
- Valid joint probability mass function?

Need to check:

1) $P_{X,Y}(x,y) \geq 0 \quad \forall x,y$

2) $\underbrace{\sum_{x,y \in \Omega_{X,Y}} P_{X,Y}(x,y) = 1}_{\Sigma \text{ of all entries in table} = 1}$

$$\Omega_{X,Y} = \{ (x,y) \mid \underbrace{P_{X,Y}(x,y)}_{P(X=x,\ Y=y)} > 0 \}$$



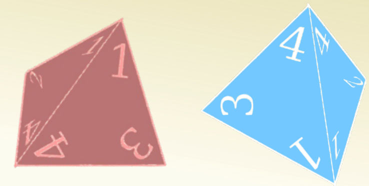
X is the number of of job interviews you need to do to get a job.

Y is the state of the economy. (Bad = 1, Below Avg = 2, Above Avg =3, Excellent = 4

	1	2	3	4
X/Y	**Bad**	**Below Avg**	**Above Avg**	**Excelle nt**
1	0.01	0.05	0.08	0.12
2	0.03	0.05	0.06	0.12
3	0.03	0.12	0.04	0.05
4	0.08	0.12	0.03	0.01

- Probability a student gets a job in one interview and economy is excellent?

$$P(X=1, Y=4) = 0.12$$



X is the number of of job interviews you need to do to get a job.

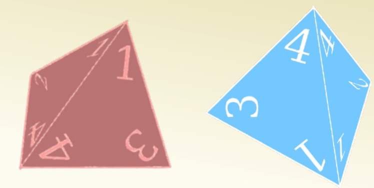
Y is the state of the economy. (Bad = 1, Below Avg = 2, Above Avg =3, Excellent = 4

	1	2	3	4
X/Y	**Bad**	**Below Avg**	**Above Avg**	**Excellent**
1	0.01	0.05	0.08	0.12
2	0.03	0.05	0.06	0.12
3	0.03	0.12	0.04	0.05
4	0.08	0.12	0.03	0.01

- Probability a student gets a job in at most 3 interviews and economy below avg?

$$P(X \leq 3 \cap Y = 2)$$

$$= 0.22$$



X is the number of of job interviews you need to do to get a job.

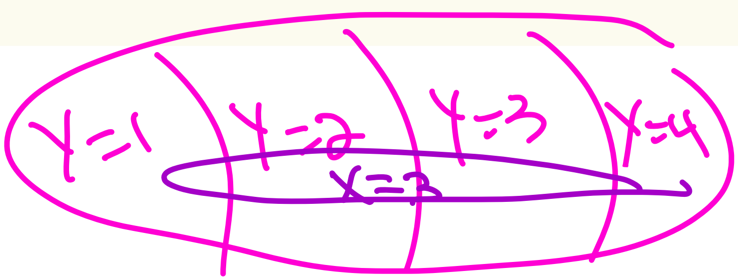
Y is the state of the economy. (Bad = 1, Below Avg = 2, Above Avg =3, Excellent = 4

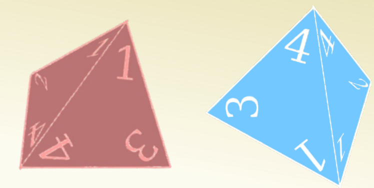
	1	2	3	4
X/Y	**Bad**	**Below Avg**	**Above Avg**	**Excellent**
1	0.01	0.05	0.08	0.12
2	0.03	0.05	0.06	0.12
3	0.03	0.12	0.04	0.05
4	0.08	0.12	0.03	0.01

- A student ~~needs~~ takes exactly 3 interviews to get a job?

$$P(X = 3)$$
$$= P(X=3 \cap Y=1)$$
$$+ P(X=3 \cap Y=2)$$
$$+ P(X=3 \cap Y=3)$$
$$+ P(X=3 \cap Y=4)$$

marginal distn of X





X is the number of of job interviews you need to do to get a job.

Y is the state of the economy. (Bad = 1, Below Avg = 2, Above Avg =3, Excellent = 4

	1	2	3	4
X/Y	**Bad**	**Below Avg**	**Above Avg**	**Excellent**
1	0.01	0.05	0.08	0.12
2	0.03	0.05	0.06	0.12
3	0.03	0.12	0.04	0.05
4	0.08	0.12	0.03	0.01

- A student gets a job in one interview given that the economy is below average?

$$P(X=1 \mid Y=2)$$

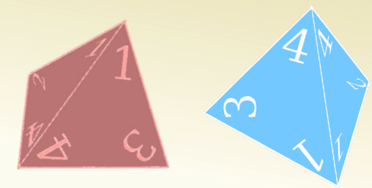
$$= \frac{P(X=1 \cap Y=2)}{P(Y=2)}$$

$$= \frac{0.05}{0.34}$$

$$P(X=x \mid Y=2)$$

$$= p_{X|Y=2}(x)$$

conditional prob mass fn



X is the number of of job interviews you need to do to get a job.

Y is the state of the economy. (Bad = 1, Below Avg = 2, Above Avg =3, Excellent = 4

X/Y	Bad (1)	Below Avg (2)	Above Avg (3)	Excellent (4)
1	0.01	0.05	0.08	0.12
2	0.03	0.05	0.06	0.12
3	0.03	0.12	0.04	0.05
4	0.08	0.12	0.03	0.01

- A student gets a job in one interview given that the economy is below average?

$$P(X=1 \mid Y=2)$$

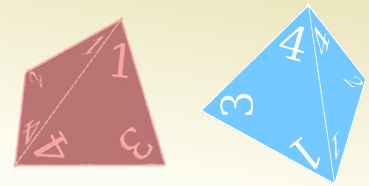
$$= \frac{P(X=1 \cap Y=2)}{P(Y=2)} = \frac{0.05}{\text{pink}}$$

$$= \frac{0.05}{0.34}$$

$P(X=x \mid Y=2)$ called conditional p.m.f.

$p_{X|Y=2}(x)$

$$\Omega_{X,Y} = \Omega_X * \Omega_Y$$



X is the number of of job interviews you need to do to get a job.

Y is the state of the economy. (Bad = 1, Below Avg = 2, Above Avg =3, Excellent = 4

	1	2	3	4
X/Y	**Bad**	**Below Avg**	**Above Avg**	**Excelle nt**
1	0.01	0.05	0.08	0.12
2	0.03	0.05	0.06	0.12
3	0.03	0.12	0.04	0.05
4	0.08	0.12	0.03	0.01

- Are these random variables independent?

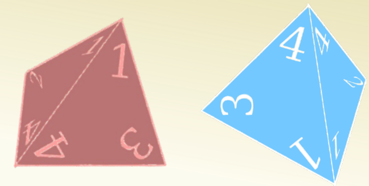
$\boxed{\text{Need to check that } P_{X,Y}(x,y) = P_X(x) \cdot P_Y(y) \quad \forall x,y}$

$$P_X(1) = 0.01 + 0.05 + 0.08 + 0.12 = 0.26$$

$$P_Y(4) = 0.12 + 0.12 + 0.05 + 0.01 = 0.3$$

$$0.26 \cdot 0.3 = 0.078 \neq 0.12 = P_{X,Y}(1,4)$$

Example – Weirder Dice



Suppose I roll two fair 4-sided die independently. Let X be the value of the first die, and Y be the value of the second die. Let $U = \min(X, Y)$ and $W = \max(X, Y)$

$\Omega_U = \{1,2,3,4\}$ and $\Omega_W = \{1,2,3,4\}$

$\Omega_{U,W} = \{(u, w) \in \Omega_U \times \Omega_W : u \le w\} \ne \Omega_U \times \Omega_W$

Are U and W independent?

No

U\W	1	2	3	4
1	1/16	2/16	2/16	2/16
2	0	1/16	2/16	2/16
3	0	0	1/16	2/16
4	0	0	0	1/16

$\Omega_U = \{1,2,3,4\}$

$\Omega_V = \{1,2,3,4\}$

A quick check for independence

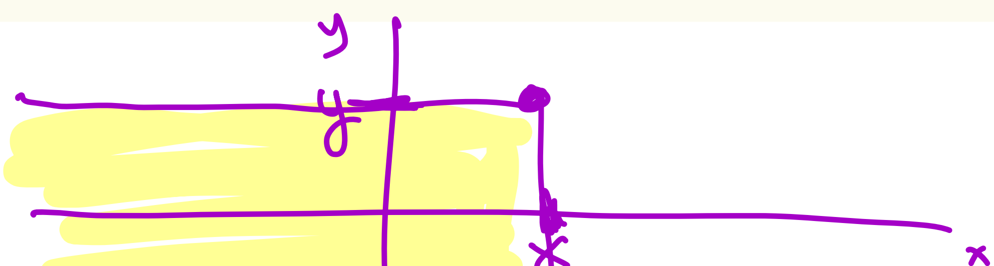
The check: $\Omega_X \times \Omega_Y$ must equal $\Omega_{X,Y}$ for independence.

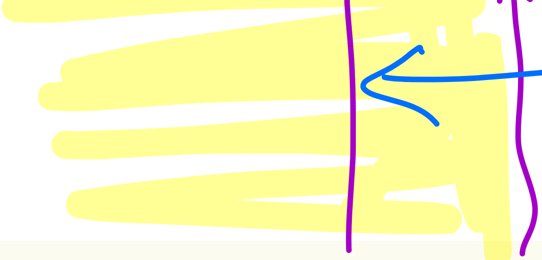
Suppose that there is some $(a, b) \in \Omega_X \times \Omega_Y$, but not in $\Omega_{X,Y}$.

Then: $p_{X,Y}(a, b) = 0$

But: $p_X(a) > 0$ and $p_Y(b) > 0$

But beware, the converse is not true: $\Omega_X \times \Omega_Y = \Omega_{X,Y}$ does not imply independence!





$P(X \le x, Y \le y)$

Joint distributions

$\Omega_{X,Y} = \{(x,y) \mid p_{X,Y}(x,y) > 0\}$ $\Omega_{X,Y} = \{(x,y) \mid f_{X,Y}(x,y) > 0\}$

	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x,y) = P(X = x, Y = y)$	$f_{X,Y}(x,y) \neq P(X = x, Y = y)$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \le x} \sum_{s \le y} p_{X,Y}(t,s)$ $P(X \le x, Y \le y)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s)\,ds\,dt$
Normalization	$\sum_x \sum_y p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_y p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$
Expectation	$E[g(X,Y)] = \sum_x \sum_y g(x,y) p_{X,Y}(x,y)$	$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy$
Independence	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$

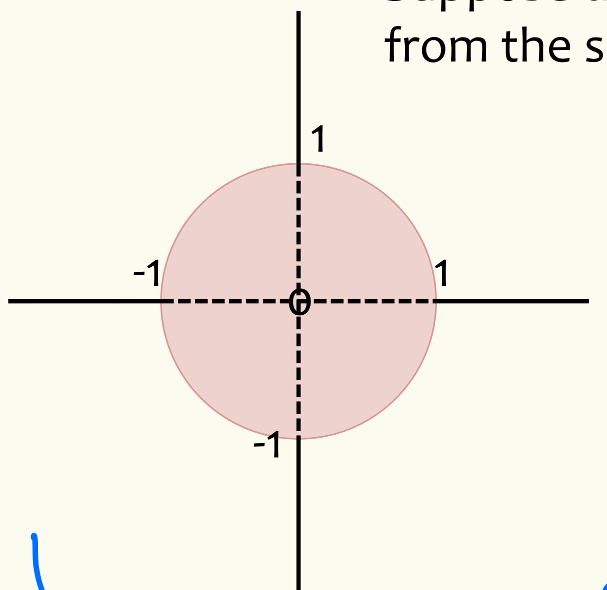
$P(X \in [x, x+dx], Y \in [y+dy]) \approx f_{X,Y}(x,y)\,dx\,dy$

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

set of outcomes

$P((X,Y) \subseteq R) = \iint_R f_{X,Y}(x,y)\,dx\,dy$

Example – Uniform distribution on a unit disk



Suppose that a pair of random variables (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \le 1$

What is the joint density?

$$f_{X,Y}(x,y) = \begin{cases} c & \text{if } x^2 + y^2 \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(X, Y)

$$\iint_{x^2+y^2\le 1} c \, dx\, dy = c \underbrace{\boxed{\text{Red Area}}}_{\pi} = 1$$

$$\Rightarrow c = \frac{1}{\pi}$$

Joint distributions

$\Omega_{X,Y} = \{(x,y) \mid p_{X,Y}(x,y) > 0\}$ $\Omega_{X,Y} = \{(x,y) \mid f_{X,Y}(x,y) > 0\}$

	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x,y) = P(X = x, Y = y)$	$f_{X,Y}(x,y) \neq P(X = x, Y = y)$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \le x} \sum_{s \le y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s)\,ds\,dt$
Normalization	$\sum_{x} \sum_{y} p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_{y} p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$
Expectation	$E[g(X,Y)] = \sum_{x} \sum_{y} g(x,y) p_{X,Y}(x,y)$	$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy$
Independence	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$

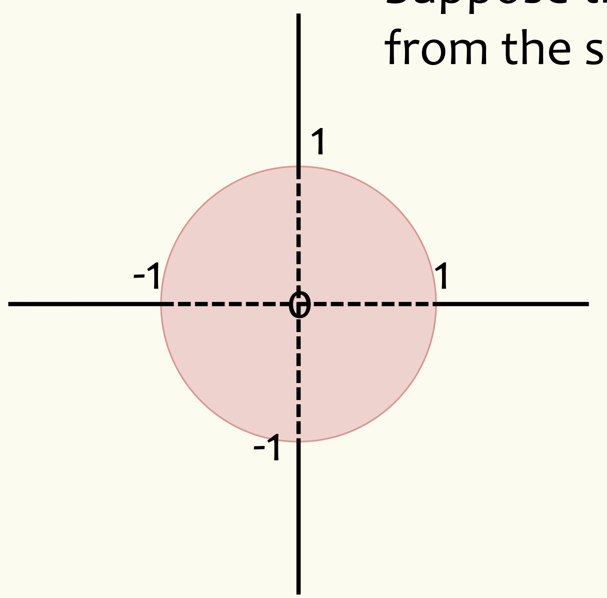
$P(X \in [x, x+dx], Y \in [y+dy]) \approx f_{X,Y}(x,y)\,dx\,dy$

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

set of outcomes

$P((X,Y) \subset R) = \iint_R f_{X,Y}(x,y)\,dx\,dy$

Example – Uniform distribution on a unit disk

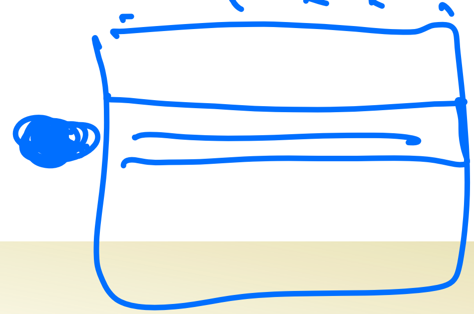


Suppose that a pair of random variables (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$

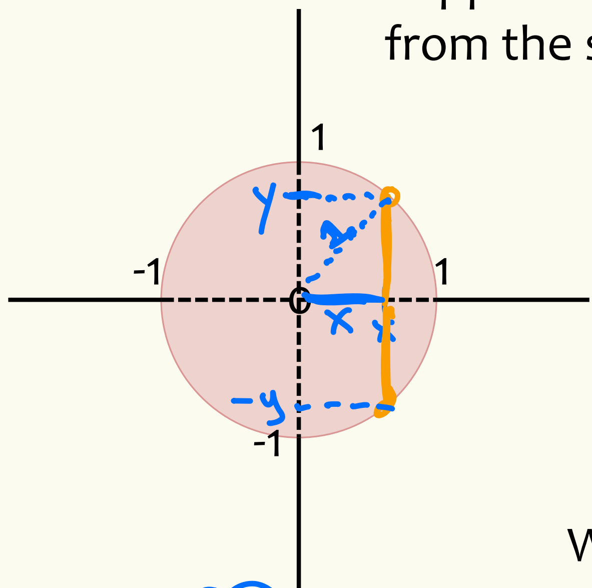
What is the joint density?

$$f_{X,Y}(x,y) = \begin{cases} c & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$c \cdot \underbrace{\text{Area of unit circle}}_{\pi} = \iint_{x^2+y^2 \leq 1} c \, dx\, dy = 1 \quad \Rightarrow c = \frac{1}{\pi}$



Example – Uniform distribution on a unit disk



Suppose that a pair of random variabes (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$

This is a disk of radius 1 which has area π

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{\pi} & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

What is $f_X(x)$?

$x^2 + y^2 = 1$

$y^2 = 1 - x^2$

y

$$\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_{-y}^{y} \frac{1}{\pi}\,dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\,dy$$

Joint distributions

$\Omega_{X,Y} = \{(x,y) \mid p_{X,Y}(x,y) > 0\}$ $\Omega_{X,Y} = \{(x,y) \mid f_{X,Y}(x,y) > 0\}$

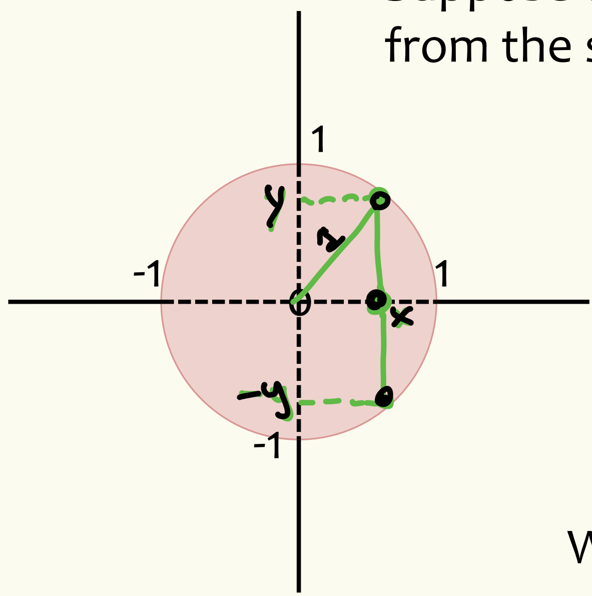
	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x,y) = P(X=x, Y=y)$	$f_{X,Y}(x,y) \neq P(X=x, Y=y)$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \le x} \sum_{s \le y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s)\,ds\,dt$
Normalization	$\sum_x \sum_y p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_y p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$
Expectation	$E[g(X,Y)] = \sum_x \sum_y g(x,y) p_{X,Y}(x,y)$	$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy$
Independence	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$

$P(X \in [x, x+dx], Y \in [y+dy]) \approx f_{X,Y}(x,y)\,dx\,dy$

$P((X,Y) \in \underset{\text{set of outcomes}}{A}) = \sum_{(x,y) \in A} p_{X,Y}(x,y)$

$P((X,Y) \subset R) = \iint_R f_{X,Y}(x,y)\,dx\,dy$

Example – Uniform distribution on a unit disk



Suppose that a pair of random variabes (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$

This is a disk of radius 1 which has area π

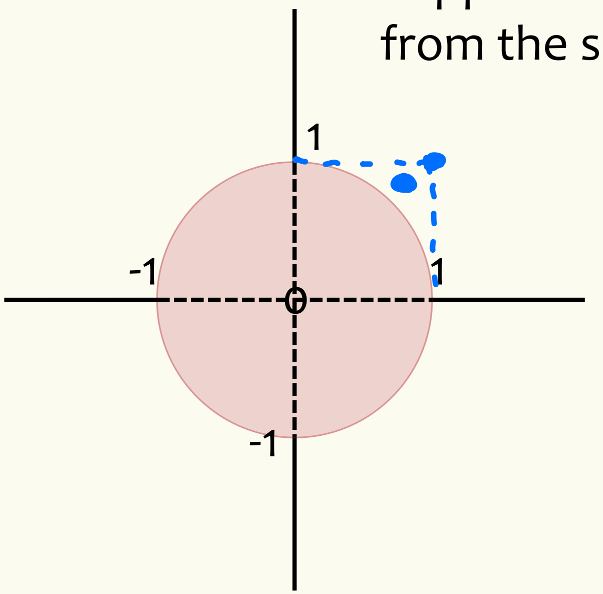
$$f_{X,Y}(x, y) = \begin{cases} \dfrac{1}{\pi} & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

What is $f_X(x)$?

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} \, \mathrm{d}y$$

$$= 2\sqrt{1 - x^2}/\pi$$

Example – Uniform distribution on a unit disk



Suppose that a pair of random variables (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$

This is a disk of radius 1 which has area π

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{\pi} & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$\forall x, y$

$$\text{indp} \equiv \underbrace{f_{X,Y}(x,y)}_{0} = \underbrace{f_X(x)}_{>0}\,\underbrace{f_Y(y)}_{>0}$$

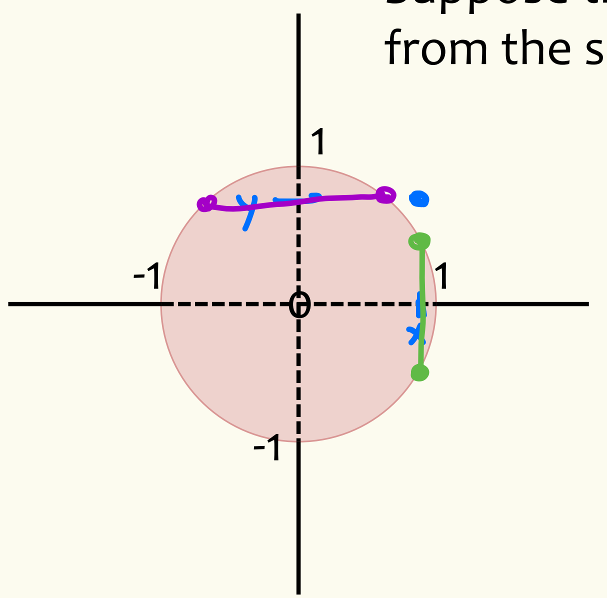
Are X and Y independent?

No

$$\Omega_{X,Y} \neq \Omega_X \times \Omega_Y$$

Example – Uniform distribution on a unit disk

Suppose that a pair of random variables (X, Y) is chosen uniformly from the set of real points (x, y) such that $x^2 + y^2 \leq 1$



This is a disk of radius 1 which has area π

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{\pi} & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Are X and Y independent?

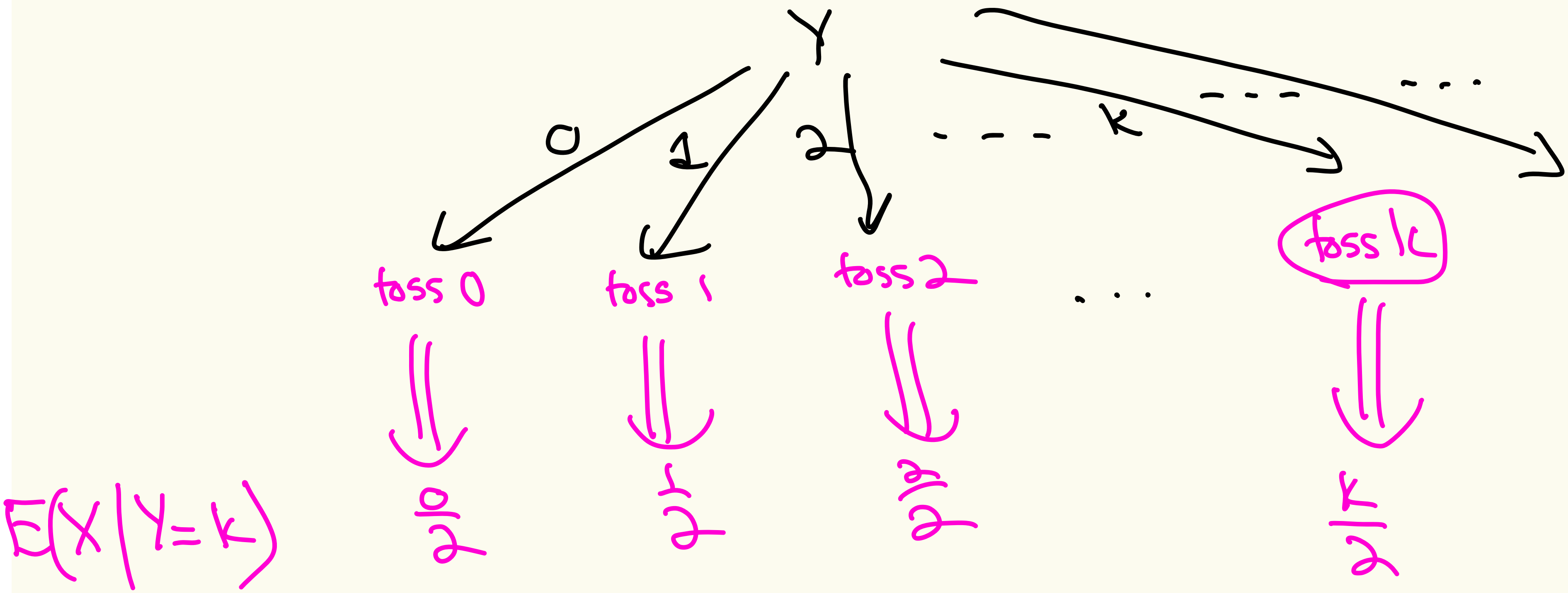
No $\Omega_{X,Y} \neq \Omega_X \times \Omega_Y$

Agenda

- Bit more on joint distributions
- Law of Total Expectation and LTP (continuous)

Conditional Expectation and Law of Total Expectation

Suppose someone gave us $Y \sim \text{Poi}(5)$ fair coins and we wanted to compute the expected number of heads X from flipping those coins.



$$E(X) = \sum_{k=0}^{\infty} E(X \mid Y=k)\, P(Y=k)$$

Conditional Expectation

Definition. If X is a discrete random variable then the **conditional expectation** of X given event A is

$$\underline{\mathbb{E}[X \mid A]} = \sum_{x \in \Omega_X} x \cdot P(X = x \mid A)$$

condition p.m.f.

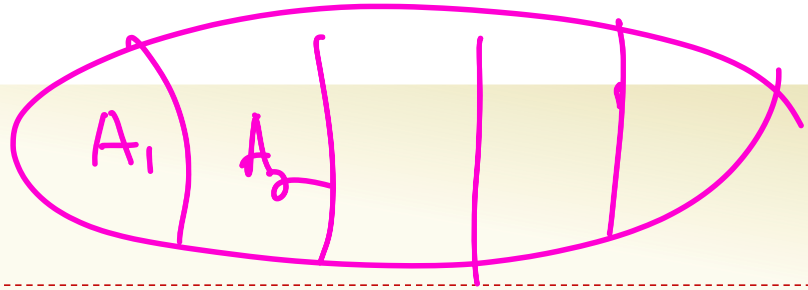
X|A

Note:

- Linearity of expectation still applies here

$$\mathbb{E}[aX + bY + c \mid A] = a\,\mathbb{E}[X \mid A] + b\,\mathbb{E}[Y \mid A] + c$$

Law of Total Expectation



Law of Total Expectation (event version). Let X be a random variable and let events $A_1, \dots, A_n$ partition the sample space. Then,

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X \mid A_i] \cdot P(A_i)$$

Law of Total Expectation (random variable version). Let X be a random variable and Y be a discrete random variable. Then,

$$\mathbb{E}[X] = \sum_{y \in \Omega_Y} \mathbb{E}[X \mid Y = y] \cdot P(Y = y)$$

Proof of Law of Total Expectation

Follows from Law of Total Probability and manipulating sums

$$\mathbb{E}[X] = \sum_{x \in \Omega_X} x \cdot P(X = x)$$

$$= \sum_{x \in \Omega_X} x \cdot \sum_{i=1}^{n} P(X = x \mid A_i) \cdot P(A_i)$$ (by LTP)

$$= \sum_{i=1}^{n} P(A_i) \sum_{x \in \Omega_X} x \cdot P(X = x \mid A_i)$$ (change order of sums)

$$= \sum_{i=1}^{n} P(A_i) \cdot \mathbb{E}[X \mid A_i]$$ (def of cond. expect.)

Example – Flipping a Random Number of Coins

Suppose someone gave us $Y \sim \mathrm{Poi}(5)$ fair coins and we wanted to compute the expected number of heads X from flipping those coins.

By the Law of Total Expectation

$$\mathbb{E}[X] = \sum_{i=0}^{\infty} \mathbb{E}[X \mid Y = i] \cdot P(Y = i) = \sum_{i=0}^{\infty} \frac{i}{2} \cdot P(Y = i)$$

$$= \frac{1}{2} \cdot \sum_{i=0}^{\infty} i \cdot P(Y = i)$$

$$= \frac{1}{2} \cdot \mathbb{E}[Y] = \frac{1}{2} \cdot 5 = 2.5$$

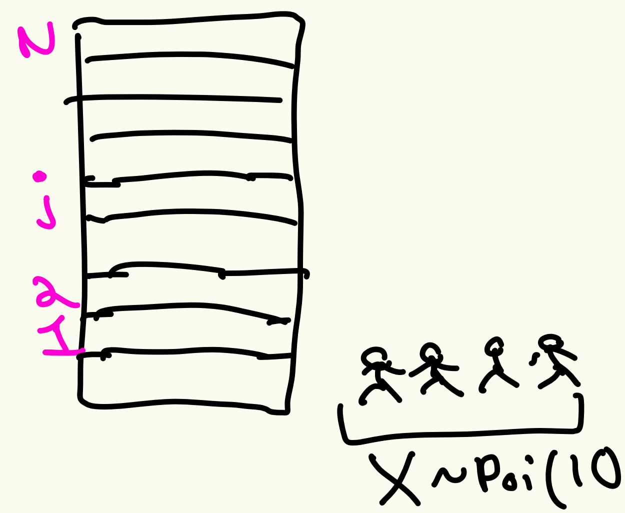
Example -- Elevator rides

The number X of people who enter an elevator on the ground floor is a Poisson random variable with mean 10. If there are N floors above the ground floor, and if each person is equally likely to get off at any one of the N floors, independently of where others get off, compute the expected number of stops the elevator will make before discharging all the passengers.

Example -- Elevator rides

The number X of people who enter an elevator on the ground floor is a Poisson random variable with mean 10. If there are N floors above the ground floor, and if each person is equally likely to get off at any one of the N floors, independently of where others get off, compute the expected number of stops the elevator will make before discharging all the passengers.

S # $\underbrace{\text{stops}}_{\text{someone gets off}}$ elevator makes



$$E(S) = \sum_{i=0}^{\infty} \underbrace{E(S \mid X=i)}_{?} \underbrace{P(X=i)}_{\frac{e^{-10}\,10^i}{i!}}$$

$$S = Y_1 + Y_2 + \dots + Y_N$$

$$Y_j = \begin{cases} 1 & \text{someone gets off at floor } j \\ 0 & \text{o.w.} \end{cases}$$

$$E(S \mid X=i) = E(Y_1 + \dots + Y_N \mid X=i) = E(Y_1 \mid X=i) + E(Y_2 \mid X=i) + \dots + E(Y_N \mid X=i)$$

$$E(Y_j \mid X=i) = \Pr(\text{at least one of } i \text{ ~~people~~ gets off at floor } j)$$

$$= 1 - \Pr(\text{nobody gets off at floor } j)$$

$$= 1 - \left(1 - \frac{1}{N}\right)^i$$

$$E(S \mid X=i) = N\left(1 - \left(1 - \frac{1}{N}\right)^i\right)$$

Law of total probability

Definition. Let A be an event and Y a discrete random variable. Then

$$P[A] = \sum_{y \in \Omega_Y} P(A|Y = y) p_Y(y)$$

Definition. Let A be an event and Y a continuous random variable. Then

$$P[A] = \int_{-\infty}^{\infty} P(A|Y = y) f_Y(y) \mathrm{d}y$$

Example use of law of total probability

Suppose that the time until server 1 crashes is $X \sim Exp\ (\lambda)$ and the time until server 2 crashes is independent, with $Y \sim Exp\ (\mu)$.

What is the probability that server 1 crashes before server 2?

Example use of law of total probability

$X \sim Exp\ (\lambda), Y \sim Exp\ (\mu)$.

What is the probability that $Y > X$?

$$P(Y > X) = \int_0^\infty \Pr(Y > X \mid X = x)\ f_X(x)dx$$

$$= \int_0^\infty \Pr(Y > x \mid X = x)\ \lambda e^{-\lambda x}\ \ dx$$

$$= \int_0^\infty \Pr(Y > x) \lambda e^{-\lambda x}\ \ dx$$

$$= \int_0^\infty e^{-\mu x}\ \lambda\ e^{-\lambda x}\ \ dx$$

$$= \frac{\lambda}{\lambda + \mu} \int_0^\infty (\lambda + \mu) \cdot e^{-\mu x}\ e^{-\lambda x}\ dx$$

$$= \frac{\lambda}{\lambda + \mu}$$

Alternative approach

$X \sim Exp\ (\lambda), Y \sim Exp\ (\mu)$.

What is the probability that $Y > X$?

$$P(Y > X) = \int_{x=0}^{\infty} \int_{y=x}^{\infty} f_{X,Y}(x, y) \mathrm{dy}\ \mathrm{d}x$$

$$= \int_{x=0}^{\infty} \int_{y=x}^{\infty} f_X(x) \cdot f_Y(y) \mathrm{dy}\ \mathrm{d}x$$

Reference Sheet (with continuous RVs)

	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x,y) = P(X = x, Y = y)$	$f_{X,Y}(x,y) \neq P(X = x, Y = y)$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \le x} \sum_{s \le y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s)\,ds\,dt$
Normalization	$\sum_{x} \sum_{y} p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_{y} p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$
Expectation	$E[g(X,Y)] = \sum_{x} \sum_{y} g(x,y) p_{X,Y}(x,y)$	$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy$
Conditional PMF/PDF	$p_{X \mid Y}(x \mid y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$	$f_{X \mid Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
Conditional Expectation	$E[X \mid Y = y] = \sum_{x} x p_{X \mid Y}(x \mid y)$	$E[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X \mid Y}(x \mid y)\,dx$
Independence	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$