

CSE 312

Foundations of Computing II

Lecture 18: Joint Distributions

Midterm

- Grading is not finished in time for today
 - After grading I will review for fairness before releasing grades
 - With the Friday holiday, grades won't be ready until Monday
 - Solutions will be posted on Canvas pages for Monday

- Please focus on the course content and problem sets while you are waiting...

Agenda

- Joint Distributions ◀
 - Cartesian Products
 - Joint PMFs and Joint Range
 - Marginal Distribution
- Conditional Expectation and Law of Total Expectation

Why joint distributions?

- Given all of its user's ratings for different movies, and any preferences you have expressed, Netflix wants to recommend a new movie for you.
- Given a large amount of medical data correlating symptoms and personal history with diseases, predict what is ailing a person with a particular medical history and set of symptoms.
- Given current traffic, pedestrian locations, weather, lights, etc. decide whether a self-driving car should slow down or come to a stop

Review Cartesian Product

Definition. Let A and B be sets. The **Cartesian product** of A and B is denoted

$$A \times B = \{(a, b) : a \in A, b \in B\}$$

Example.

$$\{1,2,3\} \times \{4,5\} = \{(1,4), (1,5), (2,4), (2,5), (3,4), (3,5)\}$$

If A and B are finite sets, then $|A \times B| = |A| \cdot |B|$.

The sets don't need to be finite! You can have $\mathbb{R} \times \mathbb{R}$ (often denoted \mathbb{R}^2)

Joint PMFs and Joint Range

Definition. Let X and Y be discrete random variables. The **Joint PMF** of X and Y is

$$p_{X,Y}(a, b) = P(X = a, Y = b)$$

Definition. The **joint range** of $p_{X,Y}$ is

$$\Omega_{X,Y} = \{(c, d) : p_{X,Y}(c, d) > 0\} \subseteq \Omega_X \times \Omega_Y$$

Note that

$$\sum_{(s,t) \in \Omega_{X,Y}} p_{X,Y}(s, t) = 1$$

Example – Weird 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.

$$\Omega_X = \{1,2,3,4\} \text{ and } \Omega_Y = \{1,2,3,4\}$$

In this problem, the joint PMF is if

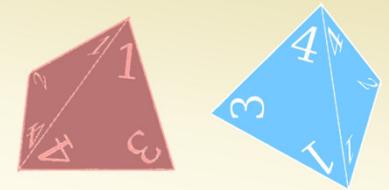
$$p_{X,Y}(x,y) = \begin{cases} 1/16 & \text{if } x,y \in \Omega_{X,Y} \\ 0 & \text{otherwise} \end{cases}$$

| $X \setminus Y$ | 1 | 2 | 3 | 4 |
|-----------------|------|------|------|------|
| 1 | 1/16 | 1/16 | 1/16 | 1/16 |
| 2 | 1/16 | 1/16 | 1/16 | 1/16 |
| 3 | 1/16 | 1/16 | 1/16 | 1/16 |
| 4 | 1/16 | 1/16 | 1/16 | 1/16 |

and the joint range is (since all combinations have non-zero probability)

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

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 $W = \max(X, Y)$

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

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

Poll: pollev.com/paulbeame028

What is $p_{U,W}(1, 3) = P(U = 1, W = 3)$?

- a. $1/16$
- b. $2/16$
- c. $1/2$
- d. Not sure

| $u \setminus w$ | 1 | 2 | 3 | 4 |
|-----------------|---|---|---|---|
| 1 | | | | |
| 2 | | | | |
| 3 | | | | |
| 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 $W = \max(X, Y)$

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

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

The joint PMF $p_{U,W}(u, w) = P(U = u, W = w)$ is

$$p_{U,W}(u, w) = \begin{cases} 2/16 & \text{if } (u, w) \in \Omega_U \times \Omega_W \text{ where } w > u \\ 1/16 & \text{if } (u, w) \in \Omega_U \times \Omega_W \text{ where } w = u \\ 0 & \text{otherwise} \end{cases}$$

| $u \setminus w$ | 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 |

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 $W = \max(X, Y)$

Suppose we didn't know how to compute $P(U = u)$ directly. Can we figure it out if we know $p_{U,W}(u, w)$?

Just apply LTP over the possible values of W :

$$p_U(1) = 7/16$$

$$p_U(2) = 5/16$$

$$p_U(3) = 3/16$$

$$p_U(4) = 1/16$$

| $u \setminus w$ | 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 |

Marginal PMF

Definition. Let X and Y be discrete random variables and $p_{X,Y}(a, b)$ their joint PMF. The **marginal PMF** of X

$$p_X(a) = \sum_{b \in \Omega_Y} p_{X,Y}(a, b)$$

Similarly, $p_Y(b) = \sum_{a \in \Omega_X} p_{X,Y}(a, b)$

Continuous distributions on $\mathbb{R} \times \mathbb{R}$

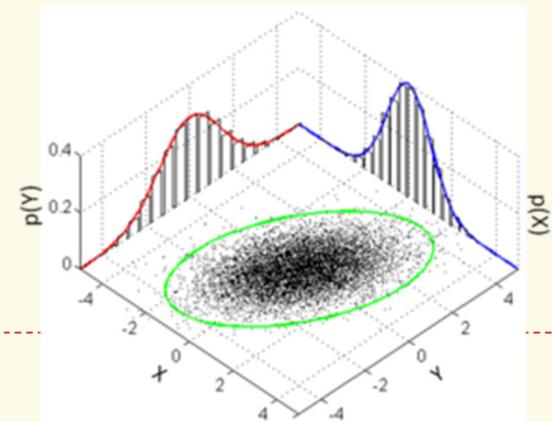
Definition. The **joint probability density function (PDF)** of continuous random variables X and Y is a function $f_{X,Y}$ defined on $\mathbb{R} \times \mathbb{R}$ such that

- $f_{X,Y}(x, y) \geq 0$ for all $x, y \in \mathbb{R}$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$

for $A \subseteq \mathbb{R} \times \mathbb{R}$ the probability that $(X, Y) \in A$ is $\iint_A f_{X,Y}(x, y) dx dy$

The **(marginal) PDFs** f_X and f_Y are given by

- $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$
- $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$



Independence and joint distributions

Definition. Discrete random variables X and Y are **independent** iff

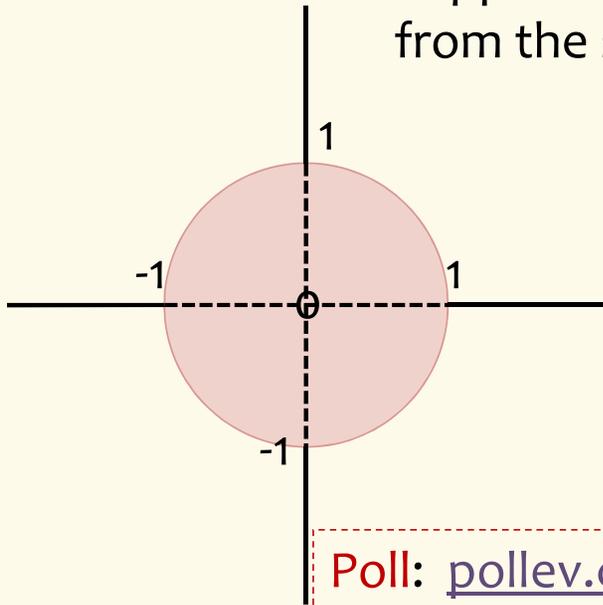
- $p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$ for all $x \in \Omega_X, y \in \Omega_Y$

Definition. Continuous random variables X and Y are **independent** iff

- $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$ for all $x, y \in \mathbb{R}$

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

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

Poll: pollev.com/paulbeame028

Are X and Y independent?

- a. Yes
- b. No

$$\begin{aligned} f_X(x) &= \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy \\ &= 2\sqrt{1-x^2}/\pi \end{aligned}$$

Joint Expectation

Definition. Let X and Y be discrete random variables and $p_{X,Y}(a, b)$ their joint PMF. The **expectation** of some function $g(x, y)$ with inputs X and Y

$$\mathbb{E}[g(X, Y)] = \sum_{a \in \Omega_X} \sum_{b \in \Omega_Y} g(a, b) \cdot p_{X,Y}(a, b)$$

Brain Break



Agenda

- Joint Distributions
 - Cartesian Products
 - Joint PMFs and Joint Range
 - Marginal Distribution
- Conditional Expectation and Law of Total Expectation ◀

Conditional Expectation

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

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

Notes:

- Can be phrased as a “random variable version”

$$\mathbb{E}[X | Y = y]$$

- Linearity of expectation still applies here

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

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,

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

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

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

Proof of Law of Total Expectation

Follows from Law of Total Probability and manipulating sums

$$\begin{aligned}\mathbb{E}[X] &= \sum_{x \in \Omega_X} x \cdot P(X = x) \\ &= \sum_{x \in \Omega_X} x \cdot \sum_{i=1}^n P(X = x | A_i) \cdot P(A_i) && \text{(by LTP)} \\ &= \sum_{i=1}^n P(A_i) \sum_{x \in \Omega_X} x \cdot P(X = x | A_i) && \text{(change order of sums)} \\ &= \sum_{i=1}^n P(A_i) \cdot \mathbb{E}[X | A_i] && \text{(def of cond. expect.)}\end{aligned}$$

Example – Flipping a Random Number of Coins

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.

By the Law of Total Expectation

$$\begin{aligned}\mathbb{E}[X] &= \sum_{i=0}^{\infty} \mathbb{E}[X | Y = i] \cdot P(Y = i) = \sum_{i=0}^{\infty} \frac{i}{2} \cdot P(Y = i) \\ &= \frac{1}{2} \cdot \sum_{i=0}^{\infty} i \cdot P(Y = i) \\ &= \frac{1}{2} \cdot \mathbb{E}[Y] = \frac{1}{2} \cdot 5 = 2.5\end{aligned}$$

Example – Computer Failures (a familiar example)

Suppose your computer operates in a sequence of steps, and that at each step i your computer will fail with probability p (independently of other steps).

Let X be the number of steps it takes your computer to fail.

What is $\mathbb{E}[X]$?

Let Y be the indicator random variable for the event of failure in step 1

$$\begin{aligned}\text{Then by LTE, } \mathbb{E}[X] &= \mathbb{E}[X | Y = 1] \cdot P(Y = 1) + \mathbb{E}[X | Y = 0] \cdot P(Y = 0) \\ &= 1 \cdot p + \mathbb{E}[X | Y = 0] \cdot (1 - p) \\ &= p + (1 + \mathbb{E}[X]) \cdot (1 - p)\end{aligned}$$

since if $Y = 0$ experiment starting at step 2 looks like original experiment

Solving we get $\mathbb{E}[X] = 1/p$

Covariance: How correlated are X and Y ?

Recall that if X and Y are independent, $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

Definition: The **covariance** of random variables X and Y ,
$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

Unlike variance, covariance can be positive or negative. It has value 0 if the random variables are independent.

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

Two Covariance examples:

Suppose $X \sim \text{Bernoulli}(p)$

If random variable $Y = X$ then

$$\text{Cov}(X, Y) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \text{Var}(X) = p(1 - p)$$

If random variable $Z = -X$ then

$$\begin{aligned} \text{Cov}(X, Z) &= \mathbb{E}[XZ] - \mathbb{E}[X] \cdot \mathbb{E}[Z] \\ &= \mathbb{E}[-X^2] - \mathbb{E}[X] \cdot \mathbb{E}[-X] \\ &= -\mathbb{E}[X^2] + \mathbb{E}[X]^2 = -\text{Var}(X) = -p(1 - p) \end{aligned}$$

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 | $F_{X,Y}(x, y) = \sum_{t \leq x} \sum_{s \leq y} p_{X,Y}(t, s)$ | $F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(t, s) ds dt$ |
| 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 | $p_{X Y}(x y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$ | $f_{X Y}(x y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$ |
| Conditional Expectation | $E[X Y = y] = \sum_x x p_{X Y}(x y)$ | $E[X Y = y] = \int_{-\infty}^{\infty} x f_{X Y}(x y) dx$ |
| Independence | $\forall x, y, p_{X,Y}(x, y) = p_X(x) p_Y(y)$ | $\forall x, y, f_{X,Y}(x, y) = f_X(x) f_Y(y)$ |