RVs Continued: Joint Distribution and Intro to Expectation

CS 70, Summer 2019

Lecture 19, 7/25/19

From Yesterday...

- RVs **assign numbers** to outcomes.
- **Treat $X = i$ as any ordinary event.**
- Bernoulli, Binomial, Geometric, Poisson RVs.

Today:

- Joint Distributions, Independent RVs, Conditional Probability
- Introduction to expectation and linearity of expectation

Joint Distribution

Lets you work with **multiple** random variables.
No different from **intersections of events!**

RV X: takes **values** a in **set** A
RV Y: takes **values** b in **set** B } Defining sets of events.

Joint Distribution:
Values: $\{(\overset{X=a}{a}, \overset{Y=b}{b}) : a \in A, b \in B\}$

Specify the Probabilities:

often see comma

$\mathbb{P}[(a,b)] = \mathbb{P}[X=a \cap Y=b] = \mathbb{P}[X=a, Y=b]$

Joint Distribution: Example I

	$X=1$	$X=2$	$X=3$
$Y=2$	0.2	0.2	0.1
$Y=3$	0.1	0	0.3
$Y=4$	0	0.1	0

$\mathbb{P}[X=3, Y=2]$

$\mathbb{P}[(X \text{ even}) \cap (Y \text{ even})] = \mathbb{P}[X=2, Y=2] + \mathbb{P}[X=2, Y=4]$
$= 0.2 + 0.1 = 0.3$

Joint Distribution: Example II

	$X=1$	$X=2$	$X=3$
$Y=2$	0.2	0.2	0.1
$Y=3$	0.1	0	0.3
$Y=4$	0	0.1	0

$\mathbb{P}[X=2] = 0.2 + 0 + 0.1 = 0.3$

$\mathbb{P}[Y=2] = 0.2 + 0.2 + 0.1 = 0.5$

$\mathbb{P}[(X=2) \cap (Y=2)] = 0.2$

Are the events $X=2$, $Y=2$ independent?
$(0.3)(0.5) \overset{?}{=} 0.2$
✗

Joint Distribution: Example III

	$X=1$	$X=2$	$X=3$
$Y=2$	0.2	0.2	0.1
$Y=3$	0.1	0	0.3
$Y=4$	0	0.1	0

Are the events $X=1$ and $Y=2$ **independent**?

$\mathbb{P}[X=1] = 0.2 + 0.1 = 0.3$
$\mathbb{P}[Y=2] = 0.5$ ← Last slide.
$\mathbb{P}[(X=1) \cap (Y=2)] = 0.2$
✗

Independent Random Variables

RVs X (values in A) and Y (values in B) are **independent** if:

for all $a \in A, b \in B$:

$$\mathbb{P}[X = a, Y = b] = \mathbb{P}[X=a]\,\mathbb{P}[Y=b]$$

Essentially the same story as **ordinary events**!!

Conditional Distributions

Also the **same exact story**:

	$X = 1$	$X = 2$	$X = 3$
$Y = 2$	0.2	0.2	0.1
$Y = 3$	0.1	0	0.3
$Y = 4$	0	0.1	0

$$\mathbb{P}[(Y \text{ even}) | X \leq 2] = \frac{\mathbb{P}[Y \text{ even} \cap X \leq 2]}{\mathbb{P}[X \leq 2]}$$

$$= \frac{0.2 + 0.2 + 0 + 0.1}{0.2 + 0.2 + 0.1 + 0 + 0 + 0.1} = \frac{0.5}{0.6} = \frac{5}{6}$$

Memorylessness of Geometrics

Memoryless: For all positive integers s, t:

$$\mathbb{P}[X \geq s + t | X > t] = \mathbb{P}[X \geq s]$$

redundant condition.

Let $X \sim \text{Geometric}(p)$. X is memoryless:

LHS: $\mathbb{P}[X \geq s+t | X > t] = \dfrac{\mathbb{P}[(X \geq s+t) \cap (X > t)]}{\mathbb{P}[X > t]}$

$= \dfrac{\mathbb{P}[X \geq s+t]}{\mathbb{P}[X > t]}$

$= \dfrac{(1-p)^{s+t-1}}{(1-p)^t} = (1-p)^{s-1}$

RHS: $\mathbb{P}[X \geq s] = (1-p)^{s-1}$

Sum of Two Independent Poissons

Let $X \sim \text{Poisson}(\lambda_1)$, $Y \sim \text{Poisson}(\lambda_2)$. ($0, 1, 2, \ldots$) ($0, 1, 2, \ldots$)
X and Y are independent.

$$\mathbb{P}[X + Y = k] = \sum_{i=0}^{k} \mathbb{P}[X = i, Y = k-i]$$

cases

$$= \sum_{i=0}^{k} \mathbb{P}[X=i]\,\mathbb{P}[Y=k-i].$$

defn. of Poisson.

$$= \sum_{i=0}^{k} \left(\frac{\lambda_1^{i}}{i!} e^{-\lambda_1}\right)\left(\frac{\lambda_2^{k-i}}{(k-i)!} e^{-\lambda_2}\right)\left(\frac{k!}{k!}\right)$$

$$= e^{-(\lambda_1+\lambda_2)}\left[\sum_{i=0}^{k} \underbrace{\frac{k!}{i!(k-i)!}}_{\binom{k}{i}} \lambda_1^{i} \lambda_2^{k-i}\right]\frac{1}{k!}$$

use Binomial Thm.

$$= e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^k \frac{1}{k!}$$

$\mathbb{P}[Z = k]$ if $Z \sim \text{Poi}(\lambda_1+\lambda_2)$

$\Rightarrow \boxed{X + Y \sim \text{Poi}(\lambda_1 + \lambda_2)}$

Break

Which building on or near campus is your "spirit building"?

Expectation of a RV

Also called the **mean** or **average** of a RV.
Let X be a RV with values in A.

Its **expectation** is defined as:

$$\mathbb{E}[X] = \sum_{a \in A} a \cdot \mathbb{P}[X = a]$$

Expectation of a RV: Example I

$$X = \begin{cases} 1 & \text{wp } 0.4 \\ \frac{1}{2} & \text{wp } 0.25 \\ -\frac{1}{2} & \text{wp } 0.35 \end{cases}$$

$\mathbb{E}[X] =$ $1 \cdot \mathbb{P}[X=1] + \frac{1}{2}\mathbb{P}[X=\frac{1}{2}] + (-\frac{1}{2})\mathbb{P}[X=-\frac{1}{2}]$.

$= 1 \cdot 0.4 + \frac{1}{2} \cdot 0.25 + -\frac{1}{2} 0.35$

$= 0.4 + \underbrace{0.125 - 0.175}_{-0.05}$

$= 0.35$

Expectation of a RV: Example III

	$X=1$	$X=2$	$X=3$
$Y=2$	0.2	0.2	0.1
$Y=3$	0.1	0	0.3
$Y=4$	0	0.1	0

$\mathbb{E}[X] =$ $1 \cdot \mathbb{P}[X=1] + 2 \cdot \mathbb{P}[X=2] + 3 \cdot \mathbb{P}[X=3]$

$= 1(0.3) + 2(0.3) + 3(0.4) = 2.1$

$\mathbb{E}[Y] =$ $2(0.5) + 3(0.4) + 4(0.1)$

$= 1 + 1.2 + 0.4 = 2.6$

Expectation of a Bernoulli

Recall that if $X \sim \text{Bernoulli}(p)$

$$\mathbb{P}[X=1] = p$$

$$\mathbb{P}[X=0] = 1 - p$$

Then: $\mathbb{E}[X] =$ $\cancel{0 \cdot (1-p)} + 1 \cdot p$

$= p$

Mixing Up HW

n students turn in their HW, but I accidently mix them up. I return HW to the students, so that each mixup (**permutation**) is equally likely.

What is the expected number of students who get their own HW back?

cases: $0, 1, 2, \ldots, n-2, n$.

Size of prob space: $n!$

Finding $\mathbb{P}[\#\text{ fixed points} = i]$ is hard!

$\Rightarrow$ Def'n is not enough!

Linearity of Expectation

The definition of expectation isn't always easy to use. **Linearity** remedies this.

Theorem: Let $X_1, X_2, \ldots, X_n$ be RVs over the same probability space. ✱
They are **not necessarily independent.** Then:

1. $\mathbb{E}[X_1 + \ldots + X_n] = \mathbb{E}[X_1] + \ldots + \mathbb{E}[X_n]$
2. For **constant** c, $\mathbb{E}[cX_i] = c \cdot \mathbb{E}[X_i]$

Proof: Notes. Out of scope, but not a hard proof.
~~Maybe~~ Probably formally go through it next lecture.

Linearity: Example I

	$X=1$	$X=2$	$X=3$
$Y=2$	0.2	0.2	0.1
$Y=3$	0.1	0	0.3
$Y=4$	0	0.1	0

From previous: $\mathbb{E}[X] = 2.1$, $\mathbb{E}[Y] = 2.6$.

$\mathbb{E}[3X + 7Y] =$ $\mathbb{E}[3X] + \mathbb{E}[7Y]$

$= 3\,\mathbb{E}[X] + 7\,\mathbb{E}[Y]$

$= 3 \cdot 2.1 + 7 \cdot 2.6$

$= 6.3 + 18.2 = 24.5$

Expectation of a Binomial

Let $X_1, \ldots, X_n$ be i.i.d. Bernoulli(p) RVs. → ind., identically distributed.

Let $X = X_1 + \ldots + X_n$.

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

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

$$\mathbb{E}[X] = \mathbb{E}[X_1 + X_2 + \ldots + X_n]$$
$$= \mathbb{E}[X_1] + \mathbb{E}[X_2] + \ldots + \mathbb{E}[X_n]$$
$$= n \cdot \underbrace{\mathbb{E}[X_1]}_{=p \text{ from previous slide}} = \boxed{np}$$

Linearity: Example II

I draw two cards from a standard deck.
What is the expected number of aces I get?

$A = \#$ aces.

Attempt #1: Use the definition.

$$\mathbb{E}[A] = \cancel{0 \cdot \mathbb{P}[A=0]} + 1 \cdot \mathbb{P}[A=1] + 2 \cdot \mathbb{P}[A=2]$$

$$\mathbb{P}[A=1] = \mathbb{P}[\text{not ace}, \text{ace}] + \mathbb{P}[\text{ace}, \text{not ace}]$$
$$= \left(\frac{48}{52}\right)\left(\frac{4}{51}\right) + \left(\frac{4}{52}\right)\left(\frac{48}{51}\right) = \frac{384}{52 \cdot 51}$$

$$\mathbb{P}[A=2] = \left(\frac{4}{52}\right)\left(\frac{3}{51}\right) = \frac{12}{52 \cdot 51}$$

$$\mathbb{E}[A] = 1 \cdot \frac{384}{52 \cdot 51} + 2 \cdot \frac{12}{52 \cdot 51} = \frac{\cancel{408}^{\,8}}{52 \cdot \cancel{51}} = \frac{8}{52}$$
$$= \frac{2}{13}$$

Linearity: Example II

Attempt #2: Use linearity of expectation.

$C_1 =$ indicator variable for card 1 = Ace.

$C_2 =$ " " " card 2 = Ace.

$A = C_1 + C_2$

$C_1 \sim \text{Ber}\left(\frac{4}{52}\right) = \text{Ber}\left(\frac{1}{13}\right)$

$C_2 \sim \text{Ber}\left(\frac{1}{13}\right)$

$\mathbb{E}[C_1] = \frac{1}{13}$, $\mathbb{E}[C_2] = \frac{1}{13}$

$$\mathbb{E}[A] = \mathbb{E}[C_1 + C_2] = \mathbb{E}[C_1] + \mathbb{E}[C_2]$$
$$= \frac{1}{13} + \frac{1}{13} = \frac{2}{13}$$

A Note on Symmetry

$C_i =$ indicator for the i-th card being an ace.

$$\mathbb{P}[C_i = 1] = \frac{1}{13}$$

Now, imagine I draw the entire deck.

$$\mathbb{E}[\underbrace{C_1 + C_2 + \ldots + C_{52}}_{\substack{\#\text{ of aces I get.} \\ =4,\text{ always}}}] = 4$$

Using this, **for any** i, what is $\mathbb{E}[C_i]$?

$$\frac{4}{52} = \frac{1}{13}$$

Linearity: Mixing Up HW

(From notes.)

Same HW setup as before with n students.

Want: Expected # of fixed points.

$S_i =$ indicator variable for student i getting their own HW.

$S_i \sim \text{Ber}\left(\frac{1}{n}\right)$

$S = \#$ of fixed points (students who get own HW back)

$S = S_1 + S_2 + \ldots + S_n$

$$\mathbb{E}[S] = \mathbb{E}[S_1 + S_2 + \ldots + S_n]$$
$$= \mathbb{E}[S_1] + \mathbb{E}[S_2] + \ldots + \mathbb{E}[S_n]$$
$$= n \cdot \mathbb{E}[S_1] = n\left(\frac{1}{n}\right) = 1$$

Summary

- Joint distribution: **multiple** RVs. Can still be defined for **non-independent** RVs.
- Ideas of independence, conditional probability **same as before**.
- **Expectation** describes the **weighted average** of a RV.
- For more complicated RVs, break down into smaller parts (e.g. **indicator variables**) and **use linearity**