CSE 312
Section 4

Random Variables

Slide design by Claris Winston, Charles Immendorf

01 - Reminders!

PSet 2 grades released

PSet 3 was due

PSet 4 is released! Start early :D

Review

02 - Random Variables

An outcome from a random experiment → **X** **a random variable** → *Some number* (the **range of X** is the set of possible values X can take on)

Probability Mass Function (PMF)

$P(X=k)$

probability that the random variable X will take on the value k

what is the probability of an outcome that will result in X being k

for discrete random variables (random variables with afinite, countably infinite range), this may sometimes be a piecewise function

Random Variables

01
02
03
04

Random variable

Captures a *quantitative property* *(some numerical value that describes the outcome)* of the outcome in a random experiment

e.g., sum of the dices on an random experiment where we roll 2 dice

Random Variables

01
02
03
04

An outcome from a random experiment → **X** **a random variable** → *Some number*

(the **range (or support) of X** (sometimes denoted as $\mathbf{\Omega_X}$) is the set of possible values X can take on)

Probability Mass Function (PMF)

$$p_X(k) = P(X=k)$$

probability that the random variable X will take on the value k

what is the probability of an outcome that will result in X being k

for discrete random variables (random variables with a finite, countably infinite range), this may sometimes be a piecewise function

Random Variables

01
02
03
04

An outcome from a random experiment → **X** **a random variable** → *Some number*
(the **range (or support) of X** (sometimes denoted as $\boldsymbol{\Omega_X}$) is the set of possible values X can take on)

Cumulative Distribution Function

$$F_X(k) =$$

P(X <= k) –> probability that the value X takes on is less than or equal to k

what is the probability of an outcome that will result in X being <= k

often can be derived from the PDF

Random Variables

01

02

03

04

An outcome from a random experiment → **X** **a random variable** → *Some number* (the **range (or support) of X** (sometimes denoted as $\boldsymbol{\Omega_X}$) is the set of possible values X can take on)

Expectation

$$E[X] = \Sigma(k \cdot P(X=k))$$

sum of values in the range of X, weighted by the probability

on average, what value can we "expect" X to take?

think about it like a weighted average of all the possible values X could be (weighted by the P(X=k))

Random Variables

01

02

03

04

An outcome from a random experiment → **X** **a random variable** → *Some number*

(the **range (or support) of X** (sometimes denoted as $\boldsymbol{\Omega_X}$) is the set of possible values X can take on)

Expectation

$$E[X] = \Sigma(k \cdot P(X=k))$$

sum of values in the range of X, weighted by the probability

on average, what value can we "expect" X to take?

just averaging all the possible values of X wouldn't work since each outcome isn't necessarily equally likely

Random Variables

01

02

03

04

An outcome from a random experiment → **X** **a random variable** → *Some number*

(the **range (or support) of X** (sometimes denoted as $\mathbf{\Omega_X}$) is the set of possible values X can take on)

Expectation of a function of X (aka "Law of the Unconscious Statistician" (aka "LOTUS"))

$$E[f(X)] = \Sigma(f(k) \cdot P(X=k))$$

(note that the probabilities are still weighted using X (not f(X))

on average, what value can we "expect" f(X) to take?

LoE

Linearity **o**f **E**xpectation is a powerful property of random variables!

02 - Linearity of Expectation

Random Variables

allow us to represent a quantitative property of a random experiment

EXPECTATION - weighted average of possible outcomes

you could use “brute force” and use the formula for expectation ($E[X]=\sum(x*P(x))$)

sometimes, just applying the formula can be messy, so LoE comes in handy

LINEARITY OF EXPECTATION (LoE) *is one important property*

$$E(X+Y) = E(X) + E(Y)$$

the expected value of the sum of 2 random variables is the sum of their expected values

02 - Linearity of Expectation

$$E(X+Y) = E(X) + E(Y)$$

the expected value of the sum of 2 random variables is the sum of their expected values

this gives us a helpful ***tool to calculate expectations of complex RVs***

02 - Linearity of Expectation

$$E(X+Y) = E(X) + E(Y)$$

the expected value of the sum of 2 random variables is the sum of their expected values

*this gives us a helpful **tool to calculate expectations of complex RVs***

DECOMPOSE into a sum of random variables

$$X = X1 + X2 + \ldots + Xn$$

02 - Linearity of Expectation

$$E(X+Y) = E(X) + E(Y)$$

the expected value of the sum of 2 random variables is the sum of their expected values

this gives us a helpful ***tool to calculate expectations of complex RVs***

DECOMPOSE into a sum of random variables $X = X1 + X2 + \ldots + Xn$

APPLY linearity of expectation $E[X] = E[X1] + E[X2] + \ldots + E[Xn]$

02 - Linearity of Expectation

$$E(X+Y) = E(X) + E(Y)$$

the expected value of the sum of 2 random variables is the sum of their expected values

this gives us a helpful ***tool to calculate expectations of complex RVs***

DECOMPOSE into a sum of random variables $X = X1 + X2 + \ldots + Xn$

APPLY linearity of expectation $E[X] = E[X1] + E[X2] + \ldots + E[Xn]$

CONQUER and calculate each value $E[X1] = \ldots, E[X2] = \ldots, \ldots$

02 - Linearity of Expectation

$$E(X+Y) = E(X) + E(Y)$$

the expected value of the sum of 2 random variables is the sum of their expected values

---

this gives us a helpful ***tool to calculate expectations of complex RVs***

sometimes, these Xi variables we "decompose" X into are **indicator random variables**

DECOMPOSE into a sum of random variables $X = X1 + \boxed{X2} + \ldots + Xn$

APPLY linearity of expectation $E[X] = E[X1] + E[X2] + \ldots + E[Xn]$

CONQUER and calculate each value $E[X1] = \ldots, E[X2] = \ldots, \ldots$

02 - Linearity of Expectation

***X and Y DON'T have to be independent!

$$E(X+Y) = E(X) + E(Y)$$

the expected value of the sum of 2 random variables is the sum of their expected values

sometimes, these Xi variables we "decompose" X into are **indicator random variables**

*this gives us a helpful **tool to calculate expectations of complex RVs***

DECOMPOSE into a sum of random variables

$$X = X1 + \boxed{X2} + \ldots + Xn$$

APPLY linearity of expectation

$$E[X] = E[X1] + E[X2] + \ldots + E[Xn]$$

CONQUER and calculate each value

$$E[X1] = \ldots, E[X2] = \ldots, \ldots$$

Indicator Random Variables

we can define a *indicator random variable* X for an event A

$$X = \begin{cases} 1 & \text{if event A happens} \\ 0 & \text{if event A doesn't happen} \end{cases}$$

^ X tells us whether event A will happen → so, $P(X = 1) = P(A)$

Note that **E[X]** $= 1 * P(X=1) + 0 * P(X=0)$ **= P(X=1)**

this is why indicator RVs can be really useful when applying linearity of expectation!

Additional slides for content that will be covered later in the week!

02 - Linearity of Expectation

linearity of expectation is special!

E[X+Y] = E[X] + E[Y] *but* E[X^2] ≠ (E[X])^2

instead...

E[g(X)] = Σ(g(x) * P(X=x))

Variance

Variance is a another property of RVs (like expectation) that measures how much the values in the RV "vary"

03 - Variance

Random Variables

allow us to represent a quantitative property of a random experiment

VARIANCE - how "different" are values from the expectectation "on average"

every random variable has some variance

$$Var(X) = E[(X-E(X))^2] = \Sigma_x (P(X=x) * (x-E(X))^2)$$

expected value of the squared distance between each RV outcome and the expected value of RV

add up all the squared distances weighted by their probabilities

variance = (standard deviation)^2

03 - Variance

Random Variables

allow us to represent a quantitative property of a random experiment

VARIANCE - how "different" are values from the expectectation "on average"

every random variable has some variance

$$Var(X) = E[(X-E(X))^2] = \Sigma_x (P(X=x) * (x-E(X))^2)$$

Properties

$$Var(a \cdot X + b) = a^2 \cdot Var(X)$$
$$Var(X) = E[X^2] - (E[X])^2$$

03 - Variance

Random Variables

allow us to represent a quantitative property of a random experiment

VARIANCE - how "different" are values from the expectectation "on average"

every random variable has some variance

$$Var(X) = E[(X-E(X))^2] = \Sigma_x(P(X=x) * (x-E(X))^2)$$

Properties

$$Var(a \cdot X + b) = a^2 \cdot Var(X)$$

$$Var(X) = E[X^2] - (E[X])^2$$

INDEPENDENT RV

What does independence mean for random variables?

Random Variable Independence

Random variables X and Y are **independent** if, *for all x, y in the ranges of X and Y*:

$$P(X=x, Y=y) = P(X=x) \cdot P(Y=y)$$

Knowing the value of X doesn't help "guess" what Y is

Random Variable Independence

Random variables X and Y are **independent** if –

$$P(X=x, Y=y) = P(X=x) \cdot P(Y=y)$$

Knowing the value of X doesn't help "guess" what Y is

it's a useful property! if X and Y are independent random variables then —

$$E(X \cdot Y) = E[X] \cdot E[Y]$$

$$Var(X + Y) = Var[X] + Var[Y]$$ *Linearity of variance holds*

Random Variable Independence

Random variables X and Y are **independent** if –

$$P(X=x, Y=y) = P(X=x) \cdot P(Y=y)$$

Knowing the value of X doesn't help "guess" what Y is

Additionally, there's **independent and identically distributed** (aka, "**i.i.d.**") random variables

In addition to independence, i.i.d. random variables also **have the same pmf.**

For example, rolling a die twice, where X is the first roll number and Y is the second roll number

Problems!

CONTENT REVIEW

CONTENT REVIEW

True or false: the range of a random variable X is the set of probabilities corresponding to
the possible values X can take on.

CONTENT REVIEW

FALSE. The range (support) of an RV is the set of all possible values it can take on

CONTENT REVIEW

What is the relationship between the stdev and the variance of a random variable X?

- ○ $\sigma = (\mathrm{Var}(X))^2$
- ○ $\sigma = \mathrm{Var}(X^2)$
- ○ $\mathrm{Var}(X) = \sigma^2$

CONTENT REVIEW

C. Var(X) = sigma^2 :)

CONTENT REVIEW

Let X be the random variable representing the the sum of a 3-dice roll of 6-sided dice. Which function would you use to determine the probability that $X = 7$?

A) CDF
B) PMF

CONTENT REVIEW

PMF. We use the PMF when we want to find the probability of a specific, exact value of a discrete random variable occurring.

CONTENT REVIEW

Let X be the random variable representing the outcome of taking the sum of a 3-dice roll of 6-sided dice. Which function would you use to determine the probability that $X \leq 7$?

A) CDF
B) PMF

CONTENT REVIEW

CDF. The CDF directly gives us the cumulative probability $P(X \leq x)$.

A random variable X has the PMF

$$p_X(x) = \begin{cases} 1/4 & x = -1 \\ 1/4 & x = 0 \\ 1/2 & x = 2 \\ 0 & \text{otherwise} \end{cases}$$

Find E[X]?

CONTENT REVIEW

¾, by unrolling the PMF

$$\mathbb{E}[X] = \sum_{x \in \Omega_X} x p_X(x) = (-1)\left(\frac{1}{4}\right) + (0)\left(\frac{1}{4}\right) + (2)\left(\frac{1}{2}\right) = -\frac{1}{4} + 0 + 1 = \frac{3}{4}$$

CONTENT REVIEW

A random variable X has the PMF

$$p_X(x) = \begin{cases} 1/4 & x = -1 \\ 1/4 & x = 0 \\ 1/2 & x = 2 \\ 0 & \text{otherwise} \end{cases}$$

Find Var(X)?

CONTENT REVIEW

27/16.

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \left(\sum_{x\in\Omega_X} x^2 p_X(x)\right) - \left(\frac{3}{4}\right)^2 \\
&= \left((-1)^2\cdot\frac{1}{4} + (0)^2\cdot\frac{1}{4} + (2)^2\cdot\frac{1}{2}\right) - \frac{9}{16} \\
&= \left(\frac{1}{4} + 0 + 2\right) - \frac{9}{16} \\
&= \frac{9}{4} - \frac{9}{16} = \frac{27}{16}
\end{aligned}$$

CONTENT REVIEW

A random variable X has the CDF

$$F_X(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{5} & 1 \leq x < 2 \\ \frac{1}{2} & 2 \leq x < 3 \\ \frac{3}{5} & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

- ○ $3/10$
- ○ $1/2$
- ○ $7/10$
- ○ $1/5$

What is p_X (2)? (PMF)

CONTENT REVIEW

$3/10$.

$$
\begin{aligned}
p_X(2) &= \mathbb{P}_X(X \leq 2) - \mathbb{P}_X(X \leq 1) \\
&= F_X(2) - F_X(1) \\
&= \frac{1}{2} - \frac{1}{5} = \frac{3}{10}
\end{aligned}
$$

CHEAT SHEET

- **Probability Mass Function (pmf) for a discrete random variable** X: a function $p_X : \Omega_X \to [0,1]$ with $p_X(x) = \mathbb{P}(X = x)$ that maps possible values of a discrete random variable to the probability of that value happening, such that $\sum_x p_X(x) = 1$.
- **Cumulative Distribution Function (CDF) for a random variable** X: a function $F_X : R \to R$ with $F_X(x) = \mathbb{P}(X \leq x)$
- **Expectation (expected value or mean)**: The expectation of a discrete random variable is defined to be $\mathbb{E}[X] = \sum_x x p_X(x) = \sum_x x\mathbb{P}(X = x)$.
- **Law of the unconscious statistician (LOTUS)**: The expectation of a function of a discrete random variable $g(X)$ is $\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$.
- **Linearity of Expectation (LoE)**: Let X and Y be random variables, and $a, b, c \in \mathbb{R}$. Then, $\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$. Also, for any random variables $X_1, \dots, X_n$,

$$\mathbb{E}[X_1 + X_2 + \dots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \dots + \mathbb{E}[X_n].$$

- **Variance**: Let X be a random variable and $\mu = \mathbb{E}[X]$. The variance of X is defined to be $\mathrm{Var}(X) = \mathbb{E}[(X - \mu)^2]$. Notice that since this is an expectation of a non-negative random variable $((X - \mu)^2)$, variance is always non-negative. With some algebra, we can simplify this to $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
- **Standard Deviation**: Let X be a random variable. We define the standard deviation of X to be the square root of the variance, and denote it $\sigma = \sqrt{\mathrm{Var}(X)}$.
- **Property of Variance**: Let $a, b \in \mathbb{R}$ and let X be a random variable. Then, $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$.

3 - SIDED DIE

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

a) What is the probability mass function of X?

b) Find E[X] directly from the definition of expectation.

c) Find E[X] again, but this time using Linearity of Expectation.

d) What is Var(X)

8- BALLS IN BINS

Let X be the number of bins that remain empty when m balls are distributed into n bins randomly and independently. For each ball, each bin has an equal probability of being chosen. (Notice that two bins being empty are not independent events: if one bin is empty, that decreases the probability that the second bin will also be empty. This is particularly obvious when $n = 2$ and $m > 0$.)

Find $E[X]$.

II- PRACTICE

a) Let X be a random variable with $p_X(k) = ck$ for $k \in \{1, \dots, 5\} = \Omega_X$, and 0 otherwise. Find the value of c that makes X follow a valid probability distribution and compute its mean and variance ($\mathbb{E}[X]$ and $\mathrm{Var}(X)$).

11- PRACTICE

b) Let X be *any* random variable with mean $\mathbb{E}[X] = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$. Find the mean and variance of $Z = \dfrac{X - \mu}{\sigma}$. (When you're done, you'll see why we call this a "standardized" version of X!)

11- PRACTICE

c) Let X, Y be independent random variables. Find the mean and variance of $X - 3Y - 5$ in terms of $\mathbb{E}[X], \mathbb{E}[Y], \mathrm{Var}(X)$, and $\mathrm{Var}(Y)$.

II- PRACTICE

d) Let $X_1, \dots, X_n$ be independent and identically distributed (iid) random variables each with mean μ and variance σ^2. The sample mean is $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Find the mean and variance of $\bar{X}$. If you use the independence assumption anywhere, **explicitly label** at which step(s) it is necessary for your equalities to be true.

5: HUNGRY WASHING MACHINE

You have 10 pairs of socks (so 20 socks in total), with each pair being a different color. You put them in the washing machine, but the washing machine eats 4 of the socks chosen at random. Every subset of 4 socks is equally probable to be the subset that gets eaten. Let X be the number of complete pairs of socks that you have left.

a) What is the range of X, ΩX (the set of possible values it can take on)? What is the probability mass function of X?
b) find E[X] using the definition of expectation.
c) find E[X] using LOE.

6: HAT CHECK

At a reception, n people give their hats to a hat-check person. When they leave, the hat-check person gives each of them a hat chosen at random from the hats that remain. What is the expected number of people who get their own hats back? (Notice that the hats returned to two people are not independent events: if a certain hat is returned to one person, it cannot also be returned to the other person.)

see this video for a walkthrough