TAM 598 Lecture 3 : DISCRETE RANDOM VARIABLES

Announcements:

- HW 1 covers lectures 1-4 ; due on Feb 12
- to be submitted via CANVAS

I Some definitions and concepts:

- We condition probabilities on our current information I. Because this info is always in the background, we will not explicitly show it in our notation.

- Say we are doing an experiment. Result of the experiment depends on a bunch of things denoted by ω, which we call the <u>state of nature</u>.
  - → ω might be overcomplete, but it includes the variables that determine the result of the experiment

- The state of nature takes values in an enormous set Ω, the set of all possible $\omega_1, \omega_2, \ldots$. Since we don't know which ω will happen, we describe this uncertainty using a <u>probability measure</u>.

Definitions and concepts, cont'd:

- the probability measure is a function that takes a subset A of Ω and tells us how probable it is that the state of nature ω will be in A.

  → We denote this probability measure by $p(A)$

- When the set Ω is equipped with a probability measure, we call it a probability space

II Definition of a random variable:

Let X (capital letter) be the result of an experiment, read by some measuring device.

The measuring device takes the state of nature ω and maps it to a number $X(\omega)$

So we define a random variable X as a function from Ω to some set of values

Depending on the values that X takes, we can classify it into one of several types:

- discrete random variable - $X(\omega)$ takes discrete values eg, heads or tails or $0, 1, 2, \ldots$
- continuous random variable - eg, values between 0 and 1, or all positive reals
- random vector - when $X(\omega)$ is a vector
- random matrix - when $X(\omega)$ is a matrix
- random process - when $X(\omega)$ itself is a function (also "stochastic process", "random field")

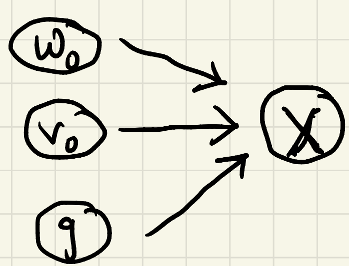
Note: X, Y, Z (uppercase) are random variables
x, y, z (lowercase) are specific values of random variables

Example : coin toss from previous lecture

state of nature $\omega = (v_0, w_0)$

initial velocity   angular velocity

Then
$$X(v_0, w_0) = \begin{cases} T & \text{if } \dfrac{2 v_0 w_0}{g} \bmod 2\pi \in \left(\frac{\pi}{2}, \frac{3\pi}{2}\right) \\ H & \text{else} \end{cases}$$



The **probability mass function** of X is a function that gives the probability that X takes a certain value $X = x$.

- we write $p(X = x)$ or just $p(x)$
- if we want to be explicit about background info, we could write $p(x \mid I)$
- mathematicians write $f_X(x)$ to remember which random variable X we are concerned with

To define $p(x)$, first define the set of states of nature that yield the specific outcome $X = x$

$$(X = x) \equiv \{ \omega \in \Omega : X(\omega) = x \}$$

Then we define the probability mass function (pmf) as

$$p(x) \equiv f_X(x) := p(X = x)$$

properties of the probability mass function $p(x)$

1) It is non-negative

2) It is normalized

3) It is additive. For any set A of possible values of X, the probability that X takes values in A is

8

Example: X is the outcome of a coin toss. Let $X = 0$ denote heads, and $X = 1$ denote tails.

Then for a fair coin:

$$p(X = 0) = \frac{1}{2}$$

$$p(X = 1) = \frac{1}{2}$$

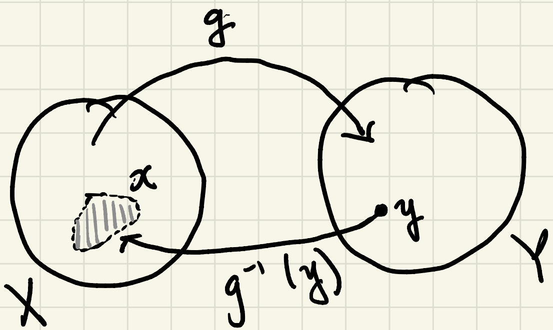
III Functions of discrete random variables

→ let X be a discrete random variable

→ let g be a function from the state space of X to some other set

→ define a new random variable $Y = g(X)$

→ let $g^{-1}(y)$ be the set of values of X that map to y through g



THEN: the probability mass function of Y is

$$\boxed{p(y) = p\left(X \in g^{-1}(y)\right) = \sum_{x \in g^{-1}(y)} p(x)}$$

This is the formal definition of the uncertainty propagation problem and the model calibration problem.

X are parameters of a causal model (inputs)

$Y = g(X)$ is the uncertain result of the causal model (output)

IV Expectation of a random variable : the value we get "on average"

$$\mathbb{E}[X] =$$

example : coin toss

$$\mathbb{E}[X] =$$

properties of the expectation:

① $\mathbb{E}[X+c] =$

② $\mathbb{E}[\lambda X] =$

③ For any function $g(x)$

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

④ Jensen's Inequality.

examples:

① $g(x) = x^2$

x drawn uniformly from $[-1, 1]$

② $g(x) = \log x$

V. **Variance** of random variables is the expected value of the squared deviation from its expectation

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

$$=$$

example: variance of a coin toss

properties of the variance:

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

2. $\mathbb{V}[X+c] = \mathbb{V}[X]$ for any constant c

3. $\mathbb{V}[\lambda X] = \lambda^2 \, \mathbb{V}[X]$

Let's get familiar w/ a few common pmf's that describe discrete random variables, and how to use them in python.

① Bernoulli distribution - outcomes 0 or 1

② Categorical distribution aka "multinulli", K possible different outcomes

③ Binomial distribution

Say you toss a coin n times, with θ = prob. of tossing heads.
Let X = the number of times you toss heads.

X is the binomial random variable

It has PMF

④ Poisson distribution - a model of the number of times an event occurs in an interval of space or time.

For instance:

⊙ X = # earthquakes > 6 Richter in next 100 years

= # major floods over next 100 years

= # patients arriving at the emergency room during the night shift

= # electrons hitting a detector in a specific time interval

Poisson is a good model when:

- → # of times takes discrete values 0, 1, 2, ...
- → events occur independently
- → the probability that an event occurs is constant per unit time, ie
  the average rate at which events occur is constant
- → events cannot occur at the same time

we say

The PMF is

Then $\mathbb{E}[X] =$

$\mathbb{V}[X] =$