February 12, 2018

CS 361: Probability & Statistics

Random Variables

Monty hall problem

- Recall the setup, there are 3 doors, behind two of them are indistinguishable goats, behind one is a car. You pick a door and win what's behind it. You prefer to win a car to a goat
- Let's suppose you pick a door at random and before you open it, Monty announces that he will now open a door and show you a goat from among the doors you didn't pick.
- After he does this, should you switch doors from your original pick to the one that you didn't pick that is still closed?

Monty hall

- Let's call the door you picked door #1, the one the host opened door #2, and the one that you didn't pick that is still closed door #3
- Let C_i be the event that the car is behind door i and H_j be the event that the host opened door j
- We want to compute P(C_1 | H_2) and compare it to P(C_3 | H_2) to see if we should switch

Monty hall

First we compute P(C_1 | H_2)

$$P(C_1|H_2) = \frac{\overset{1/2}{P(H_2|C_1)}\overset{1/3}{P(C_1)}}{\underset{1/2}{P(H_2|C_1)}\underset{1/3}{P(C_1)} + \underset{0}{P(H_2|C_2)}\underset{1/3}{P(C_2)} + \underset{1}{P(H_2|C_3)}\underset{1/3}{P(C_3)}} \quad \textbf{=1/3}$$

Now let's compute P(C_3 | H_2)

$$P(C_3|H_2) = \frac{\overset{1}{P(H_2|C_3)}\overset{1/3}{P(C_3)}}{\underset{1/2}{P(H_2|C_1)}\underset{1/3}{P(C_1)} + \underset{0}{P(H_2|C_2)}\underset{1/3}{P(C_2)} + \underset{1}{P(H_2|C_3)}\underset{1/3}{P(C_3)}} \quad \textbf{=2/3}$$

Takeaway

- See the text for other ways to set up Monty Hall and why it matters
- Conditional probabilities can be quite counterintuitive

Random Variables

Random variables

- We have figured out how to talk about assigning probabilities to outcomes and events
- We now look at a way of associating numbers with experiments, numbers that change as a function of the outcome of the experiment

Random variables

Definition: 5.1 *Discrete random variable*

Given a sample space Ω, a set of events $\mathcal{F}$, and a probability function P, and a countable set of of real numbers D, a discrete random variable is a function with domain Ω and range D.

Thus, for every outcome ω a random variable X associates to that outcome a real number $X(\omega)$

Example

- Flip a coin and observe the result. If it is heads, we report 1, if it is tails, we report 0. This is a random variable

Example

- We flip a coin 32 times, recording a 1 when we see heads and a 0 when we see tails. This produces a 32 bit random number which is a random variable

Example

- We flip a coin 32 times, reporting 1 for heads and 0 for tails. The parity of this 32 bit number is a random variable

Example

- We draw a hand of 5 cards. The number of pairs in the hand is a random variable (0, 1, or 2)

The relationship to events

Any value x of a random variable determines
a set of outcomes, i.e. an event

$$\{\omega : X(\omega) = x\}$$

So we will make reference to the probability
that the random variable X is equal to x

$$P(\{\omega : X(\omega) = x\})$$

And we will use shorthand **P(X=x)** or just **P(x)** to express this

Example

We draw a hand of 5 cards, let X be the random variable that indicates how many pairs are in the hand

So X maps from outcomes (which hand we draw) to a countable set D of numbers {0,1,2}

Saying X=1 defines a set of outcomes or an event: in this event are any outcome or hand with one pair in it

Thus P(X=1) is exactly the kind of thing we've already been studying. P(X=1) asks what is the probability of some event

Relationship to events

Another event of interest is

$$\{\omega : X(\omega) \leq x\}$$

We usually write the probability of this event

$$P(\{\omega : X(\omega) \leq x\})$$

With the following shorthand

$$P(X \leq x)$$

Random variables: two observations

Since X is a function from the sample space to the set D, every possible outcome is mapped to a number in D

Also observe that if $x_0 \neq x_1$ then the events

$$\{X = x_0\} \qquad \{X = x_1\}$$

are disjoint

And if we were to iterate over every number in D, the union of these disjoint events is the sample space, e.g.

$$\cup_{x \in D} \{X = x\} = \Omega$$

Distributions

The function of x given by

$$P(X = x)$$

is called the **probability distribution** of the discrete random variable X

P is defined for each value that X can take and is 0 everywhere else

This function might also be written as p(x)

Distributions

We also give a special name to the function of x given by

$$P(X \leq x)$$

We call this the **cumulative distribution function** of the discrete random variable X

Note that this is a non-decreasing function of x

Cumulative distributions are often written with an *f*. So we may have

$$f(x) = P(\{X \leq x\})$$

Example

Flip a biased coin 2 times.

P(H) = p and P(T) = 1-p.

Record a 1 when a toss comes up heads and a 0 when it comes up tails.

This gives a random variable that is a binary number taking on values 00,01,10,11 or 0, 1, 2, 3

What is the probability distribution and the cumulative distribution for this random variable?

x	p(x)	f(x)
0	$(1-p)^2$	$(1-p)^2$
1	$(1-p)p$	$(1-p)$
2	$p(1-p)$	$1-p^2$
3	p^2	1

Example

Flip a biased coin.

$P(H) = p$ and $P(T) = 1-p$.

If the coin comes up heads, you pay me q

If it comes up tails, I pay you r

The amount of money that changes hands is an RV

What is the probability distribution for this random variable? From my perspective

$P(X=q) = p$
$P(X=-r) = (1-p)$

Random variables

- Using our "two observations" from a few slides back, notice that since the values in D give disjoint events whose union is the sample space, we can get the following result from what we know about probabilities

$$\sum_{x \in D} P(x) = 1$$

Joint probability

- If we consider two random variables X and Y, it makes sense to think about the event generated by asking which outcomes give X=x and Y=y
- We could write the probability of this event as

$$P(\{X = x\} \cap \{Y = y\})$$

- But we will usually use shorthand and write $P(x, y)$

Joint probability

- For a discrete X and Y, we can think of P(x,y) as a table with an entry for each value that X and Y can take on
- This table is referred to as the **joint probability distribution** of X and Y

Conditional probability

- Using our prior formulation of conditional probability we know that

$$P(\{X = x\} \mid \{Y = y\})P(\{Y = y\}) = P(\{X = x\} \cap \{Y = y\})$$

- Using our shorthand we write this as

$$P(x|y)P(y) = P(x, y)$$

Bayes' rule

- An equation that we derived before for events can also be written in terms of random variables and has a special name due to its wide applicability

Definition: 5.4 *Bayes' rule*

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}$$

Joint probability

- Again using our two observations about the values of a random variable giving us disjoint events whose union is the sample space, observe the following about joint probabilities for a given value of y

$$\sum_{x \in D_x} P(x|y) = 1$$

- Write this out in terms of events if you want to convince yourself

Marginal distribution

- Another insight we can derive from our two observations is that if we have a joint distribution, we can recover the individual distributions of X and Y

$$\sum_y P(x, y) = P(x)$$

- This is often referred to as the **marginal distribution** of X, when we derive it from a joint distribution of X and Y

Random variable independence

Definition: 5.5 *Independent random variables*

The random variables X and Y are **independent** if the events $\{X = x\}$ and $\{Y = y\}$ are independent. This means that

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

which we can rewrite as

$$P(x, y) = P(x)P(y)$$

Sum and difference

Throw two dice. The number of spots on the first die is a random variable X, the number on the second is a random variable Y. Let S be the random variable given by S = X + Y and D = X - Y

What is the probability distribution of S? And what is the probability distribution of D?

S	P(s)
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

D	P(d)
-5	1/36
-4	2/36
-3	3/36
-2	4/36
-1	5/36
0	6/36
1	5/36
2	4/36
3	3/36
4	2/36
5	1/36

Sum and difference, joint distribution

Throw two dice. The number of spots on the first die is a random variable X, the number on the second is a random variable Y. Let S be the random variable given by S = X + Y and D = X - Y

What does is the joint distribution of S and D?

$$\frac{1}{36} \times \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

TABLE 5.1: *A table of the joint probability distribution of S (vertical axis; scale* $2, \ldots, 12$*) and D (horizontal axis; scale* $-5, \ldots, 5$*) from example 5.4*

Sum and difference, independence

Throw two dice. The number of spots on the first die is a random variable X, the number on the second is a random variable Y. Let S be the random variable given by $S = X + Y$ and $D = X - Y$

Are X and Y independent?

How about S and D?

Sum and difference, conditional

Throw two dice. The number of spots on the first die is a random variable X, the number on the second is a random variable Y. Let S be the random variable given by S = X + Y and D = X - Y

What is P(S | D=0)?

S	P(s)
2	1/6
4	1/6
6	1/6
8	1/6
10	1/6
12	1/6

What is P(D | S=11)

D	P(d)
-1	1/2
1	1/2