SPR Day 1

Goals: Review fundamental rules of probability

Reading: Bishop PRML Sec 1.2

Contents:

1. discrete random variables
   sample space
   probability mass function
2. joint, conditional, marginal
3. 2 fundamental rules:
   - sum rule
   - product rule

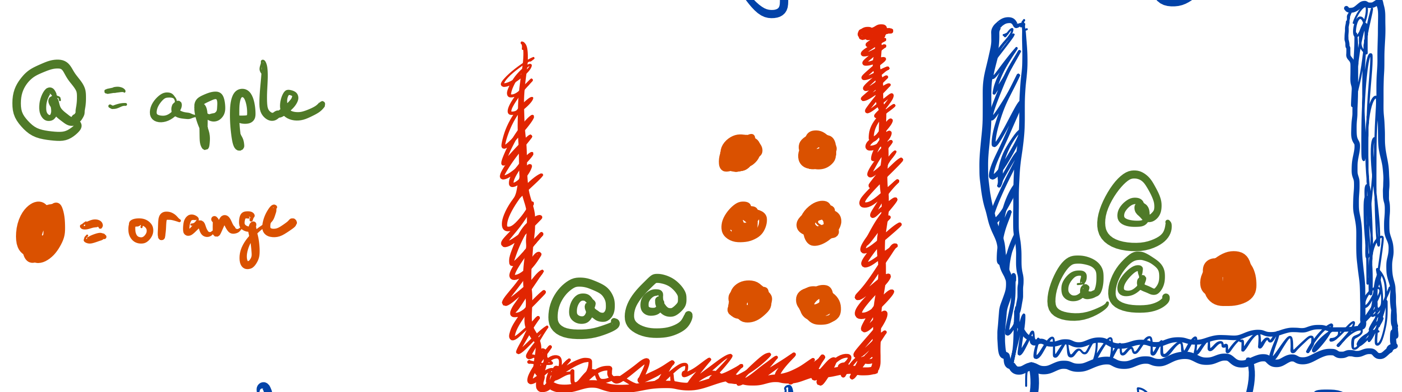
   Bayes rule
4. independence
   conditional independence
5. expectations, mean, variance

Motivating Toy Problem

Consider this stochastic procedure:

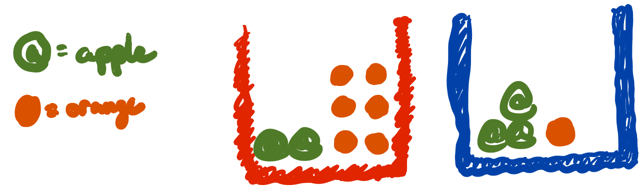
1) select a random bucket (red or blue)
   - red 40%
   - blue 60%
2) from that bucket, pick one fruit uniformly at random
3) report either "apple" or "orange" then replace the fruit

What is probability of picking "apple"?



Approach: - Break down into simple parts
- Clearly define random variables
- Apply relevant rules of probability

Events of Interest



- Which box is picked?

B

event	probability
B = red	40%
B = blue	60%

- Which fruit is picked in red box?

F

event	proba
F = @ | B = red	2/8
F = ● | B = red	6/8

- Which fruit is picked in blue box?

F

F = @ | B = blue	3/4
F = ● | B = blue	1/4

Key Concepts

random variable = variable whose value depends on outcomes of a random process

sample space = set of all possible outcome values for a random variable

sample space should satisfy:

(1) "mutually exclusive"
each outcome distinct
no values overlap

(2) "collectively exhaustive"
all possible values represented

probability mass function (PMF)
function that maps each sample space value to a probability between 0 and 1

Let X be random variable with sample space Ω
then p is a valid PMF if and only if

$$\sum_{x \in \Omega} p(x) = 1 \quad \text{and} \quad 0 \leq p(x) \leq 1 \quad \forall x \in \Omega$$

Joint Probability

Motivation: Ask question like: "what is proba. that basket is red and fruit is apple?"

Let X be random variable with sample space

$$\mathcal{X} = \{x_1, x_2, \cdots x_L\}$$

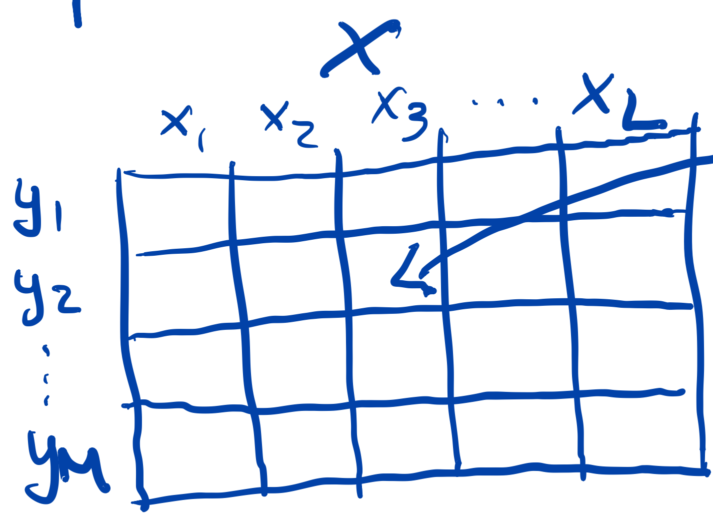
Let Y be random variable with sample space

$$\mathcal{Y} = \{y_1, y_2, \cdots y_M\}$$

Consider unordered pair of X, Y together

This is a random variable with $L \times M$ possible values

Can represent each value as a cell in a table



this cell is event $X = x_3, Y = y_2$

read as "the joint event that X equals x_3 and Y equals y_3"

The PMF for the pair r.v. would assign each cell a probability.

Conditional Probability

Motivation: Ask a question like

What is probability that basket is red given fruit is apple?

Write formally as:

$$p(B = \text{red} \mid F = @)$$

Read as:

what is probability of ______ event, given that ______ occurs.
conditioning event

Given a fixed condition, we say $p(B \mid F = @)$ is a conditional distribution.

That is, we can define a PMF over the sample space of B:

$$\sum_{b \in \Omega} p(B = b \mid F = @) = 1$$

Marginal Probability

Motivation: In context of a joint probability, ask a question about a subset of random variables:

"What is probability the fruit is apple?"

Visual

joint

F \ B	red	blue
O	0.3	0.15
@	0.1	0.45

sum over rows →

F	
O	0.45
@	0.55

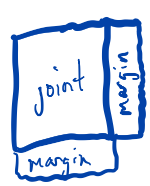
marginal distribution of F

↓ sum each column ↓

marginal of B

red	blue
0.40	0.60

Remember, the term "marginal" comes from the "margins" of the joint probability table.



2 fundamental rules

Formally relate the PMFs for joint, conditional, and marginal distributions.

$\boxed{\text{Sum Rule}}$

$$p(X = x_i) = \sum_{y \in \mathcal{Y}} p(X = x_i, Y = y)$$

marginal proba. of x_i equals sum over all joint event probabilities that involve x_i

$\boxed{\text{Product Rule}}$

$$p(Y = y \mid X = x)\, p(X = x) = p(X = x, Y = y)$$

conditional × marginal equals joint

$$p(X = x \mid Y = y)\, p(Y = y) = p(X = x, Y = y)$$

Bayes Rule

relates conditionals and marginals

$$p(Y=y \mid X=x) = \frac{p(X=x \mid Y=y)\,p(Y=y)}{p(X=x)}$$

Utility: if you know one conditional (right hand side) you can determine other (left hand side) using Bayes Rule

Derivation: Derive by product rule

$$p(Y=y \mid X=x) = \frac{p(X=x, Y=y)\,p(Y=y)}{p(X=x)}$$

Independence

Two random variables are independent if the PMF of conditional is same as PMF of the marginal

$$p(Y=y \mid X=x) = p(Y=y)$$

for all $x \in \mathcal{X}$, $y \in \mathcal{Y}$

and

$$p(X=x \mid Y=y) = p(X=x)$$

Intuitively, X is independent of Y if the probability of its value never depends on Y's value.

If X and Y are independent, the joint distribution PMF is:

$$p(X=x, Y=y) = p(X=x \mid Y=y)\, p(Y=y)$$ product rule

$$= p(X=x)\, p(Y=y)$$ using definition of independence

product of marginals

Conditional Independence

Let X, Y, Z be random variables.

We say X and Y are conditionally independent given Z if and only if

$$p(Y=y \mid X=x, Z=z) = p(Y=y \mid Z=z)$$

for all possible values x, y, z

Again, if we know X and Y are conditionally independent, we can write the joint as

$$p(x, y, z) = p(x \mid z)\, p(y \mid z)\, p(z)$$

Expectations

Given a random variable X,
we might be interested in a
function $h(x)$ of outcome x.

What is the "average" value of $h(x)$?
We can compute

$$\mathbb{E}[h(x)] \triangleq \sum_{x \in \Omega} p(x) h(x)$$

the "expected" value of h of x | is defined as | the sum over all possible values each weighted by the PMF at that value

Two useful expectations

Mean $\quad \mathbb{E}[x] = \sum_{x \in \Omega} p(x)\, x \quad$ average value of X

Variance $\quad \mathbb{E}\left[\left(x - \underbrace{\mathbb{E}[x]}_{\text{Let } \mu \text{ be this mean}}\right)^2\right] = \sum_{x \in \Omega} p(x)(x-\mu)^2 \quad$ average squared error / deviation of x from its mean