Probability for Computer Science

Spring 2021

Lecture 8

Prof. Claire Monteleoni

Today

- Independent trials and Binomial probabilities
- Discrete Random Variables
  - Some common random variables
  - Expectation and Variance

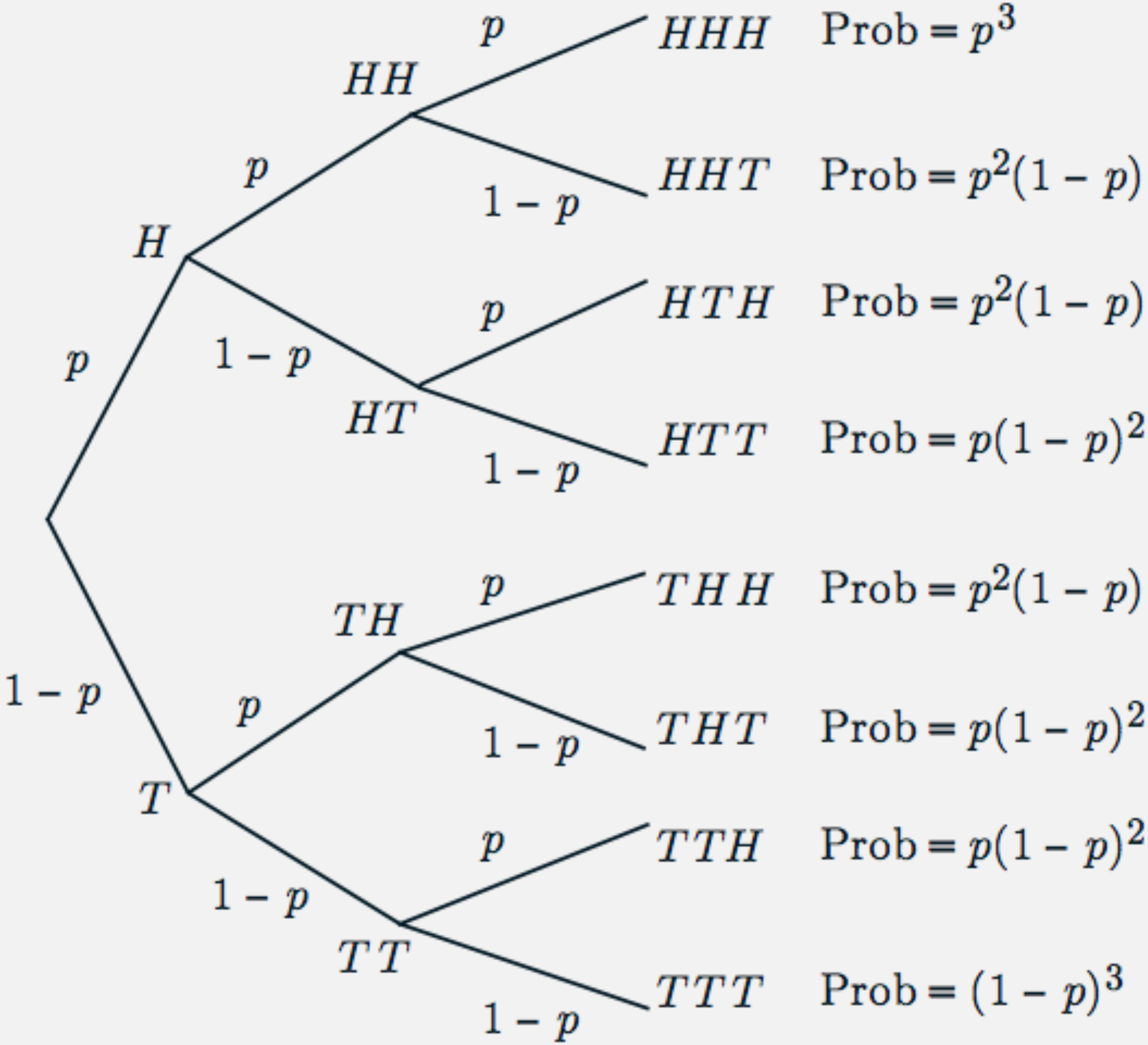
Independent Trials

An experiment that consists of a sequence of independent but identical stages is called a sequence of independent trials. E.g.

  - Repeatedly flipping the same coin
  - Repeatedly rolling the same die

In the special case where there are only 2 possible outcomes this is called a sequence of independent Bernoulli trials. E.g.

- Repeatedly flipping the same coin
- Repeatedly receiving emails that are either spam or not spam



Credit: Bertsekas & Tsitsiklis, 2008

Binomial probabilities

What is the probability that exactly k heads come up in a sequence of n independent coin tosses?

- Define the event A = {the sequence contains exactly k heads}.
- Define the coin's probability of heads, P[H] = p.
- The probability of each outcome in A:

  For any particular sequence containing exactly k heads, its probability is: $p^k(1-p)^{n-k}$

- |A| = the number of sequences containing exactly k heads = $\binom{n}{k}$
- P(A) = |A|× (Probability of each outcome in A)

The answer is therefore: $P(A) = \boxed{p(k) = \binom{n}{k} p^k (1-p)^{n-k}}$

Example

- Assume a cell phone provider can handle up to d data requests at once.
- Assume that every minute, each of its n customers makes a data request with probability p, independent of the other customers.

1. What is the probability that exactly c customers will make a data request during a particular minute?
2. What is the probability that more than d customers will make a data request during a particular minute?

Random Variables

A random variable (r.v.) is a real-valued function of the outcome of an experiment. $f : \Omega \rightarrow \mathbb{R}$

Ex. A) If the experiment is rolling a die twice, here are some r.v.s:

- The sum of the two rolls
- The number of 3's rolled
- (Value of the second roll)^3

Ex. B) If the experiment is transmitting a message, here are some r.v.s:

- The number of bits that are transmitted incorrectly
- The time the message takes to be transmitted on the network

Discrete Random Variables

A random variable is discrete if the set of values it can take is finite, or countably infinite.

Ex. A) all the r.v.s listed are discrete.

Ex. B) if the r.v.s listed here take integer values (bits, milliseconds), then they are discrete.

Ex. C) Experiment is picking a point a in [-1, 1].

- Define r.v. as a^2. Is this a discrete r.v.?
- Define r.v. as sign(a). Is this a discrete r.v.?

Discrete Random Variables

Given a probabilistic model of an experiment:

- A discrete random variable is a real-valued function of the outcome of the experiment, that can take a finite or countably infinite number of values.

- A discrete r.v. has a probability mass function (PMF) which gives the probability of each numerical value that the r.v. can take.

- A function of a discrete random variable is itself a discrete random variable.
  - Its PMF can be computed from the PMF of the original r.v.

Probability Mass Function (PMF)

Each r.v., X, has an associated PMF, defined as follows, for each value x that X can take:

$$p_X(x) = P(\{X = x\})$$

To compute the PMF of a random variable X:

For each possible value x of X:

Collect all the possible outcomes in the event {X = x}

Sum their probabilities to obtain $p_X(x)$

For simplicity, we will use the notation:

$$p_X(x) = P(X = x)$$

Probability Mass Function (PMF)

By additivity and normalization axioms: $$\sum_{x} p_X(x) = 1$$

The events {X = x} are disjoint and form a partition of the sample space.

For any set S of possible values of X, $$P(X \in S) = \sum_{x \in S} p_X(x)$$

Common random variables

1. The Bernoulli random variable
   – A binary r.v. with “success” probability p.
   – Takes values 0 and 1.
   – PMF:

$$\boxed{\begin{aligned} p_X(1) &= p \\ p_X(0) &= 1 - p \end{aligned}}$$

Common random variables

2. The Binomial random variable

- The number of “successes” in n independent Bernoulli trials, each with probability of success p.
- Possible values are 0, …, n
- PMF: $\boxed{p_X(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}}$
- The normalization property is therefore:

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$$

Common random variables

3. The Geometric random variable

   – The number of trials until the first “success,” in repeated independent Bernoulli trials, each with success probability $0 < p < 1$.

   – Takes as possible values: all the positive integers.

   – PMF: $\boxed{p_X(k) = P(X = k) = (1-p)^{k-1}p}$

   i.e., we have k - 1 trials that don’t yield a success and then we get a success on the k-th trial.

Common random variables

4. The Poisson random variable

- The number of “rare” events (p small) in a large number of independent Bernoulli trials (n large).
- Possible values: all the non-negative integers
- PMF: $$\boxed{p_X(k) = e^{-\lambda}\left(\frac{\lambda^k}{k!}\right)}$$ where λ is a parameter.

Ex. Consider a book with n words. For each word in the book, the probability it is misspelled is p, (independent of whether any other word is misspelled). Let X be the number of misspelled words in the book.

We could use a Binomial r.v. for X. But when n is very large and p is very small, the Poisson r.v. is a good approximation to the Binomial (their PMFs are similar for large n and small p), and is simpler to compute.

Functions of a random variable

If X is a random variable, then any function of X, Y = g(X), is also a random variable.

The PMF of Y can be computed from the PMF of X as follows:

$$\boxed{p_Y(y) = \sum_{\{x \mid g(x)=y\}} p_X(x)}$$

Expectation of a random variable

The expectation of a random variable, X, is a weighted average of the possible values of X.

– The weights are the probabilities of each possible value.

Formally, the expectation of a random variable, X, is defined as:

$$\boxed{E[X] = \sum_k k \cdot p_X(k)}$$

where $p_X(k)$ is from the PMF of X.

Other names for expectation are expected value, and mean.

Expectation of a function of a r.v.

If X is a random variable with PMF p_X, and g(X) is a function of X, then the expectation of the random variable g(X) is:

$$\boxed{E[g(X)] = \sum_x g(x) p_X(x)}$$

Moments of a random variable

The *n*-th moment of a random variable X is defined as: $E[X^n]$

- The first moment of X is E[X], the expectation.
- The second moment of X is $E[X^2]$.

- Since X^n is a function of X, can compute the *n*-th moment using definition of expectation of g(X):

$$E[X^n] = \sum_x x^n p_X(x)$$

Variance

For an r.v., X, consider $(X - E[X])^2$. This is also an r.v. because it's a function of X. The variance of X is defined as the expected value of the r.v. $(X - E[X])^2$.

$$\mathrm{var}(X) = E[(X - E[X])^2]$$

$$= \sum_x (x - E[X])^2 p_X(x)$$

Linearity of Expectation

Given r.v. X, if Y = a X + b, for constants a, b, then the expectation of Y can be computed as follows:

$$E[Y] = E[aX + b] = aE[X] + b$$

WARNING: this is true for linear functions, but

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

does NOT hold in general.

Variance of a linear function of r.v.

Given r.v. X, if Y = a X + b, for constants a, b, then the variance of Y can be computed as follows:

$$\mathrm{var}(Y) = \mathrm{var}(aX + b) = a^2\mathrm{var}(X)$$

Variance

$$\mathrm{var}(X) = E[(X - E[X])^2]$$

The variance can also be expressed in terms of the second moment, $E[X^2]$, and the expectation, or first moment, $E[X]$.

$$\mathrm{var}(X) = E[X^2] - (E[X])^2$$