Probability for Computer Science

Spring 2021

Lecture 9

Prof. Claire Monteleoni

Today

- Discrete Random Variables
  - Some common random variables
  - Functions of a random variable
- Expectation and Variance
- Multiple random variables

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, $E[X]$.

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

Joint PMF

$$\boxed{p_{X,Y}(a,b) = P(\{X=a\} \cap \{Y=b\}) = P(X=a, Y=b)}$$

Ex. Pick a random student from the class. Define random variables
X = {eye color}, Y = {birthday time of year} (using integer values).

Y

X	0 (Winter)	1 (Spring)	2 (Summer)	3 (Fall)
0 (Brown)	0.1	0.1	0	0.2
1 (Blue)	0.05	0.05	0.1	0
2 (Green)	0	0.1	0.2	0.1

$p_{X,Y}(X=2, Y=1)$

Joint PMF

$$\boxed{p_{X,Y}(a,b) = P(\{X=a\} \cap \{Y=b\}) = P(X=a, Y=b)}$$

Is this a valid PMF? How can you check?

Y

X

	0 (Winter)	1 (Spring)	2 (Summer)	3 (Fall)
0 (Brown)	0.1	0.1	0	0.2
1 (Blue)	0.05	0.05	0.1	0
2 (Green)	0	0.1	0.2	0.1

Joint PMF

$$p_{X,Y}(a,b) = P(\{X = a\} \cap \{Y = b\}) = P(X = a, Y = b)$$

How do you compute $p_X(x)$?

Y

X

	0 (Winter)	1 (Spring)	2 (Summer)	3 (Fall)
0 (Brown)	0.1	0.1	0	0.2
1 (Blue)	0.05	0.05	0.1	0
2 (Green)	0	0.1	0.2	0.1

Joint PMF

$$p_{X,Y}(a,b) = P(\{X=a\} \cap \{Y=b\}) = P(X=a, Y=b)$$

The marginal PMF of X is $\boxed{p_X(x) = \sum_y p_{X,Y}(x,y)}$

Y

X		0 (Winter)	1 (Spring)	2 (Summer)	3 (Fall)	$p_X(x)$
	0 (Brown)	0.1	0.1	0	0.2	0.4
	1 (Blue)	0.05	0.05	0.1	0	0.2
	2 (Green)	0	0.1	0.2	0.1	0.4

Joint PMF

$$p_{X,Y}(a,b) = P(\{X=a\} \cap \{Y=b\}) = P(X=a, Y=b)$$

The marginal PMF of Y is $\boxed{p_Y(y) = \sum_x p_{X,Y}(x,y)}$

Y

X		0 (Winter)	1 (Spring)	2 (Summer)	3 (Fall)	$p_X(x)$
X	0 (Brown)	0.1	0.1	0	0.2	0.4
X	1 (Blue)	0.05	0.05	0.1	0	0.2
X	2 (Green)	0	0.1	0.2	0.1	0.4
	$p_Y(y)$	0.15	0.25	0.3	0.3	

Functions of multiple random variables

Given r.v.s X, Y, a function Z = g(X,Y) defines another r.v.

The PMF of Z can be computed from the joint PMF of X and Y.

$$\boxed{p_Z(z) = \sum_{\{(x,y) | g(x,y)=z\}} p_{X,Y}(x,y)}$$

Functions of multiple random variables

Given r.v.s X, Y, a function Z = g(X,Y) defines another r.v.

The expectation of Z can be computed from the joint PMF of X and Y.

$$\boxed{E[g(X,Y)] = \sum_x \sum_y g(x,y) p_{X,Y}(x,y)}$$

Linearity of Expectation:

If g is a linear function, i.e. Z = aX + bY + c, then its expectation is:

$$\boxed{E[aX + bY + c] = aE[X] + bE[Y] + c}$$

Joint PMFs of three or more r.v.s

Joint PMFs can be extended to any number of r.v.s

Ex. 3 random variables X, Y, Z:

- Joint PMF:

$$\boxed{p_{X,Y,Z}(x, y, z) = P(X = x, Y = y, Z = z)}$$

- Marginal PMF of X:

$$\boxed{p_X(x) = \sum_y \sum_z p_{X,Y,Z}(x, y, z)}$$

Example

Hat Problem: *n* people throw their hats in a box, and then each person picks one hat from the box, uniformly at random.

- Each hat can only be taken by one person, and every assignment of people to hats is equally likely.

Let X be the number of people who end up with their original hat. What is the expected value of X?

Bonus Slides

Discrete random variables

- Conditional PMFs
- Conditional Expectation

Conditional PMF

The conditional Probability Mass Function (PMF) of a random variable X, conditioned on an event A with P(A) > 0, is defined, for each x, as:

$$\boxed{p_{X|A}(x) = P(X = x \mid A) = \frac{P(\{X = x\} \cap A)}{P(A)}}$$

Conditional PMF

As x varies over all possible values of X, the events

$\{X = x\} \cap A$ are disjoint, and their union is A.

So, by Total Probability Theorem: $$P(A) = \sum_x P(\{X = x\} \cap A)$$

The definition of conditional PMF is:

$$p_{X|A}(x) = \frac{P(\{X = x\} \cap A)}{P(A)}$$

to verify the Normalization property of the conditional PMF:

$$\sum_x p_{X|A}(x) = \frac{\sum_x P(\{X = x\} \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

Conditional PMF

To compute the PMF of a random variable X, conditioned on an event A with P(A) > 0:

For each possible value x of X:

Collect all the possible outcomes in the event $\{X = x\} \cap A$

Sum their probabilities and normalize, by dividing by P(A), to obtain $p_{X|A}(x)$

Conditional PMF

The conditional PMF of a random variable X, conditioned on another random variable Y, is defined as:

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

$$p_{X|Y}(x \mid y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

$$\boxed{p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}}$$

Note: just apply the original definition, but now the event to condition on is {Y = y} (for y s.t. $p_Y(y) > 0$).

Conditional PMF

$$p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

So from the joint PMF, we can compute conditional PMFs by normalizing the values in a particular row or column (divide by row/column total).

Y

X

	0 (Winter)	1 (Spring)	2 (Summer)	3 (Fall)
0 (Brown)	0.1	0.1	0	0.2
1 (Blue)	0.05	0.05	0.1	0
2 (Green)	0	0.1	0.2	0.1

0.2/0.3 = 2/3

0

0.1/0.3 = 1/3

$$p_{X|Y}(x \mid 3) = \frac{p_{X,Y}(x,3)}{p_Y(3)}$$

Conditional PMF

Conditional PMFs of one random variable conditioned on another r.v. provide ways to calculate the joint PMF:

$$p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

so $\boxed{p_{X,Y}(x,y) = p_Y(y)p_{X|Y}(x \mid y)}$ by Multiplication Rule.

And $\boxed{p_{X,Y}(x,y) = p_X(x)p_{Y|X}(y \mid x)}$ by Multiplication Rule,

and definition of $p_{Y|X}(y|x)$.

Conditional PMF

Conditional PMFs of one random variable conditioned on another r.v. provide ways to calculate the marginal PMFs:

$$p_X(x) = \sum_y p_Y(y) p_{X|Y}(x|y)$$

$$p_Y(y) = \sum_x p_X(x) p_{Y|X}(y|x)$$

by Total Probability Theorem.

Conditional Expectation

- The conditional expectation of r.v. X, given an event A with P(A)>0 is defined as: $E[X|A] = \sum_x x \cdot p_{X|A}(x)$

- For a function g(X), $E[g(X)|A] = \sum_x g(x) \cdot p_{X|A}(x)$

- Given r.v.s X and Y associated with the same experiment, the conditional expectation of X given a value y of Y is:

$$E[X|Y = y] = \sum_x x \cdot p_{X|Y}(x|y)$$

Total Expectation Theorem

- Given disjoint events A_1, … , A_n that partition the sample space, with $P(A_i) > 0$ for all i,

$$E[X] = \sum_{i=1}^{n} P(A_i) E[X|A_i]$$

Example

Messages are sent from a computer in Boston, over the internet to the following destinations, with the following probabilities:

- NYC with probability 0.5
- DC with probability 0.3
- SF with probability 0.2

The transit time is a random variable, T. Its expectation, conditioned on each city, is:

- 0.05 if message destination is NYC
- 0.1 if message destination is DC
- 0.3 if message destination is SF

Q: What is E[T]?