Probability for Computer Science

Spring 2021

Lecture 10

Prof. Claire Monteleoni

Boulder

Today

- Variance (continued)
- Linearity of Expectation
- Multiple random variables
  - Joint PMF
  - Conditional PMF
  - Conditional Expectation
  - Indep. of multiple r.v.s
- The Sample Mean

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. $(X - E[X])^2$.

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

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

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

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

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

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

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	$\boxed{0.1}$	0.2	0.1

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

Joint PMF

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

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)	$p_X(x)$
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
$p_Y(y)$	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.

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

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?

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
- Denver 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 Denver

Q: What is E[T]?

Independence of a r.v. from an event

A random variable, X, is independent of an event, A, if, for all x,

$$
\begin{aligned}
P(X = x \text{ and } A) &= P(X = x)P(A) \\
&= p_X(x)P(A)
\end{aligned}
$$

i.e., X is independent of A if the events {X=x} and A are independent, for every value of x.

Independence of random variables

To prove or disprove that r.v.s X and Y are independent, it is enough to prove or disprove any of the following statements (as they are equivalent):

- $p_{X,Y}(a,b) = p_X(a)p_Y(b)$ *for all a and b*
- $p_X(a) = p_{X|Y}(a \mid b)$ *for all a and b s.t.* $p_Y(b) > 0$
- $p_Y(b) = p_{Y|X}(b \mid a)$ *for all a and b s.t.* $p_X(a) > 0$

Independence of random variables

Are X and Y independent?

Y

X

	y_1	y_2	y_3	y_4
x_1	0.05	0.15	0	0.2
x_2	0.025	0.075	0	0.1
x_3	0.05	0.15	0	0.2

Independence of random variables

Are X and Y independent? Yes.

How can we tell?

Y

	y_1	y_2	y_3	y_4
x_1	0.05	0.15	0	0.2
x_2	0.025	0.075	0	0.1
x_3	0.05	0.15	0	0.2

X

Independence of random variables

How can we tell that X and Y are independent?

The columns are multiples of each other. Therefore $p_{X|Y}(x \mid y)$ is the same for every value of y, and therefore does not depend on y, so $p_{X|Y}(x \mid y) = p_X(x)$.

Y

X

	y_1	y_2	y_3	y_4
x_1	0.05	0.15	0	0.2
x_2	0.025	0.075	0	0.1
x_3	0.05	0.15	0	0.2

Conditional independence of r.v.s

Random variables X and Y are conditionally independent, given event A if, for all x, y,

$$P(X = x, Y = y \mid A) = P(X = x \mid A)P(Y = y \mid A)$$

$$= p_{X|A}(x)p_{Y|A}(y)$$

Equivalently, for all x and y s.t. $p_{Y|A}(y) > 0$,

$$p_{X|Y,A}(x|y) = p_{X|A}(x)$$

Properties of independent r.v.s

If X and Y are independent random variables, then:

- $E[XY] = E[X]E[Y]$
- $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$
- $\mathrm{var}(X + Y) = \mathrm{var}(X) + \mathrm{var}(Y)$

NOTE: Not true in general for arbitrary r.v.s!

Independence of multiple r.v.s

- Random variables X, Y, and Z are independent if:

  For all x, y, z:

$$p_{X,Y,Z}(x,y,z) = p_X(x)p_Y(y)p_Z(z)$$

- Let X_1 ,…, X_n be independent random variables. Then:

$$\mathtt{var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathtt{var}(X_i)$$

Note: These properties need not hold in general – need independence.

The Sample Mean

- Suppose we want to estimate the approval rating of a public figure, B.
- We ask *n* people drawn uniformly at random from the population.
- Define X_i as an indicator random variable for whether the i-th person approves of B.
- We model $X_1, X_2, \ldots, X_n$ as independent Bernoulli random variables, with common mean, *p*, and variance *p(1-p)*.
- That is, we assume *p* is the true approval rating of B. It is unknown, so we try to estimate it.
- We compute the Sample Mean from the *n* responses, i.e. the average approval rating in the *n*-person sample:

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Mean and Variance of the Sample Mean

- The Sample Mean is:

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

- What are $E[S_n]$ and $\mathrm{var}(S_n)$?