AI1110 Probability and Random Variables

Lecture 4: Multiple Random Variables

Mar 4, 2024

Joint Distributions

Given two random variables X and Y, consider events of the form $P(a_1 < X \leq b_1, a_2 < Y \leq b_2)$.

The Joint CDF (written $F_{X,Y}(x,y)$) is defined as

$$F_{X,Y}(x,y) = P(X \leq x, Y \leq y).$$

This is a function of two variables x, y.

Note that $\lim_{x\to\infty} F_{X,Y}(x,y) = F_Y(y)$, $\lim_{x\to-\infty} F_{X,Y}(x,y) = 0$, and likewise for $\lim_{y\to\infty}$.

Discrete random variables

If X and Y are discrete random variables, the

joint pmf of X and Y (written $p_{X,Y}(x,y)$ is defined as

$$p_{X,Y}(x,y) = P(X = x, Y = y).$$

Once the joint pmf is known, we can obtain the pmfs of X and Y as

$$p_X(x) = \sum_y p_{X,Y}(x,y), \quad p_Y(y) = \sum_x p_{X,Y}(x,y).$$

Here $p_X(x)$ and $p_Y(y)$ are referred to as the marginal pmfs of X and Y.

Excercise 0: Let X and Y be independent outcomes of two fair-four sided-die rolls. Set $U = X + Y$ and $V = \min\{X, Y\}$. Find the joint and marginal pmfs of U and V.

U \ V	1	2	3	4	$p_U(u)$
2	1/16	0	0	0	1/16
3	1/8	0	0	0	1/8
4	1/8	1/16	0	0	3/16
5	1/8	1/8	0	0	1/4
6	0	1/8	1/16	0	3/16
7	0	0	1/8	0	1/8
8	0	0	0	1/16	1/16
$p_V(v)$	7/16	5/16	3/16	1/16	

Figure 1: The joint and marginal pmfs of U and V. Note that the marginal pmf of U is obtained as a row sum, and the marginal pmf of V as a column sum. The entries within the grid sum to 1.

U \ V	1	2	3	4
2	1/7	0	0	0
3	2/7	0	0	0
4	2/7	1/5	0	0
5	2/7	2/5	0	0
6	0	2/5	1/3	0
7	0	0	2/3	0
8	0	0	0	1

Figure 2: The marginal $p_{U|V}(u|v)$. Note that the column sum for any column is 1. We can for e.g. verify that some of the probabilities are as expected, for e.g. $P_{U|V}(8|4) = 1$.

The conditional pmf of X given Y is defined as

$$p_{X|Y}(x|y) = p_{X,Y}(x,y)/P_Y(y).$$

Note that this too, is a function of two variables x and y. Since $p_{X|Y}(.,y)$ is a valid pmf for any given y,

$$\sum_x p_{X|Y}(x|y) = 1.$$

Continuous random variables

The joint PDF $f_{X,Y}(x,y)$ of random variables X, Y satisfies

$$P(a_1 < X \leq b_1, a_2 < Y \leq b_2) = \int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x,y)\, dx\, dy$$

for any a_1, a_2, b_1, b_2.We see that in this case (similar to a single random variable)

$$f_{X,Y}(x,y) = \lim_{\Delta x \to 0, \Delta y \to 0} \frac{1}{\Delta x \Delta y} P(x < X \leq x + \Delta x, y < Y \leq y + \Delta y).$$

Similar to the discrete case, we have the marginal PDFs

$$f_X(x) = \int_y f_{X,Y}(x,y)dy, \quad f_Y(y) = \int_x f_{X,Y}(x,y)dx.$$

Here is a rough derivation for the marginal:

$$\begin{aligned} f_X(x) &\approx \frac{1}{\Delta x} P(x < X \leq x + \Delta x) \\ &= \frac{1}{\Delta x} \sum_{y_i} P(x < X \leq x + \Delta x, y_i < Y \leq y_i + \Delta y) \\ &= \frac{1}{\Delta x} \sum_{y_i} f_{X,Y}(x,y) \Delta x \Delta y \approx \int_y f_{X,Y}(x,y)dy \end{aligned}$$

.

The joint PDF is a second derivative of the joint CDF, given by

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

THE CONDITIONAL PDF of Y given X, written $f_{Y|X}(y|x)$ is defined by

$$f_{Y|X}(y|x) = f_{X,Y}(x,y)/f_X(x).$$

$$\begin{aligned} f_{Y|X}(y|x) &\approx P(y < Y \leq y + \Delta y | X = x)/\Delta y \\ &= \frac{P(y < Y \leq y + \Delta y, X = x)}{\Delta y P(X = x)} \\ &= \frac{P(y < Y \leq y + \Delta y, x < X \leq x + \Delta x)}{\Delta y P(x < X \leq x + \Delta x)} \\ &= \frac{f_{X,Y}(x,y)\Delta x \Delta y \Delta y}{\Delta y f_X(x) \Delta x = f_{X,Y}(x,y)/f_X(x).} \end{aligned}$$

The conditional pdf is a valid pdf. We have for e.g.

$$P(X \in A | Y = y) = \int_A f_{X|Y}(x|y)\,dx, \quad \int_x f_{X|Y}(x|y)dx = 1.$$

Excercise 1: Let X, Y be uniformly distributed on $[0,1] \times [0,1]$. Find the joint, marginal, and conditional pdfs.

Excercise 2: Let X, Y be uniformly distributed on a disc of radius 1 centered at the origin. Find the joint, marginal, and conditional pdfs.

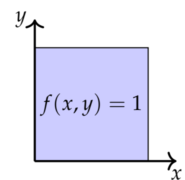

Figure 3: Joint PDF for uniformly distributed random variables on a square.

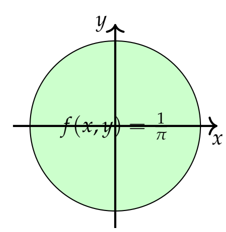

Figure 4: Joint PDF for uniformly distributed random variables in a disc. Refer to scribble for the calculations.

Conditional Expectation

The conditional expectation of X given Y (written $E(X|Y)$) is [1]

$$E(X|Y) = \int_x x f_{X|Y}(x|y)dx$$

Note that this is a function of Y.

[1] This is simply the expectation computed with the modified pdf $fX|Y(x|y)$.

ITERATED EXPECTATION often refers to the following:

$$E[X] = E[E[X|Y]]$$

Excercise 3: Suppose a stick of length L is broken by picking a point on the stick uniformly at random and throwing away the right fragment. This process is repeated another time on the surviving fragment. Let X be the length of the surviving fragment. What is $E(X)$? [2]

[2] Let Y be the position from the left of the first break. Then $E(Y) = L/2$, and $E(X|Y) = Y/2$. By iterated expectation, $E(X) = E(Y/2) = L/4$.

Independent Random Variables

Random variables X and Y are independent if the joint CDF can be expressed as the product of the marginal CDFs, i.e., $F_{X,Y}(x,y) = F_X(x)F_Y(y)$.

Events $X \leq a, Y \leq b$ are independent for all a, b.

Equivalently, if X and Y are continuous, then they are independent if and only if the joint pdf can be written as a product of marginal pdfs $f_{X,Y}(x) = f_X(x)f_Y(y)$. Likewise if X and Y are discrete then they are independent if and only if the joint pmf can be written as a product of marginal pmfs.

It is enough if the joint cdf (or joint pdf) is separable, i.e they can be written as a product of two single variable functions (one in x and the other in y)

For more than two random variables, we say that $X_1, X_2, \ldots, X_n$ are independent if the joint splits into the product of marginals

$$f_{X_1,X_2,\ldots}(x_1, x_2, \ldots) = \prod_i f_{X_i}(x_i).$$

It follows that if X_1, X_2, X_3 are independent, so that X_1 and X_2. So are $X_1 + X_2$ and X_3, and slightly more generally $h(X_1, X_2)$ and $g(X_3)$ are independent for any h, g.

IID: We say that $X_1, X_2, \ldots, X_n$ are i.i.d random variables with they are independent and have the same marginal pdfs i.e.

$$X_1, X_2, \ldots \text{ independent, and } f_{X_i} = f_{X_j} \text{ for all } i, j.$$

Sum of independent rvs:

If X and Y are independent random variables, the PDF of their sum $X + Y$ is the convolution of their individual PDFs

$$f_{X+Y}(z) = \int_x f_X(x)f_Y(z-x)dx,$$

Refer to the scribble for proof

See this or this for an illustration of this operation

and a similar expression for discrete random variables. *Excercise 4:* If X, Y are independent $U[0,1]$ random variables, what is the pdf of $X + Y$?

Excercise 5: If X, Y are independent Poisson(λ) random variables, what is the pdf of $X + Y$?

Excercise 6: If X, Y are independent Exp(λ) random variables, what is the pdf of $X + Y$?

Max, min of rvs: If X, Y are random variables, we can find the cdf of $U = \max\{X, Y\}$ by

$$F_U(u) = P(U \leq u) = P(X \leq u, Y \leq u) = F_{X,Y}(u, u).$$

Independence, if known, can be used to additionally simplify. Differentiating this w.r.t u gives the pdf of U.

Likewise, the tail distribution of $V = \min\{X, Y\}$ can be obtained by

$$P(V > v) = P(X > v, Y > v) = 1 - F_X(v) - F_Y(v) + F_{X,Y}(v, v).$$

Excercise 7: If X, Y are independent Exp(λ) random variables, what is the pdf of $\max(X, Y)$ and $\min(X, Y)$?

Computing expectation

The following problem was discussed in class: Let $X_1, X_2, \ldots$ be i.i.d uniform random variables, let

$$N = \min\{n : X_1 + X_2 + \ldots X_n > 1\}.$$

What is $E(N)$?

It is often useful to compute the expectation indirectly when possible, either by iterating expectation or by breaking up the random variable of interest into 'simpler' random variables with easy-to-compute expectations. The following problem was discussed in class

Given n i.i.d continuous random variables $X_1, X_2, \ldots, X_n$ we say that $n \geq 2$ is a record-till-date if $X_n > X_i$ for all $i < n$.

1. What is the expected number of records till date ?
2. Given an infinite sequence of i.i.d random variables $X_1, X_2, \ldots$, what is the expected time till first record ?

The key insight was exploiting symmetry. In particular, we used the following:

> If $X_1, X_2, \ldots$ are i.i.d continous random variables then $P(X_1 > X_k, k = 2, 3, \ldots, n) = 1/n$.

You can prove this by symmetry of the joint pdf, or directly by setting up an integral.

Suggested reading

Chapter 6 from [3] and Chapter 3 from [4]

These notes summarise the lectures, have not been proofread thoroughly, and are not a substitute for the notes you take in class. Please write to the teaching team about any errors you notice.

[3] Sheldon Ross. *First Course in Probability, A*. Pearson Higher Ed, 2019

[4] Dimitri P Bertsekas and John N Tsitsiklis. *Introduction to probability*, volume 1. 2002

References

[1] Dimitri P Bertsekas and John N Tsitsiklis. *Introduction to probability*, volume 1. 2002.

[2] Sheldon Ross. *First Course in Probability, A*. Pearson Higher Ed, 2019.