STA257

Neil Montgomery | HTML is official | PDF versions good for in-class use only
Last edited: 2016-10-31 14:57

probabilities involving more than one random variable at a time

motivation

A random variable is a function of a sample space, and we care about its distribution.

So far we've focussed on $X : S \to \mathbb{R}$.

Now we will look at $X : S \to \mathbb{R}^n$, which arise quite naturally. We've actually been doing this already sometimes, without saying so explicitly (see Case 2).

Case 1: you *actually observe* multiple things about a particular random outcome (e.g. you measure the weight and blood pressure of a randomly selected study participant.)

Case 2: you are considering a sequence of random variables that "replicate" the same "experiment" (e.g. repeat a Bernoulli(p) trial n times$\ldots$)

discrete motivating example

Toss two fair six-sided dice. The sample space has 36 elements.

Observe the total and the difference.

Denote the total by X_1 and the difference by X_2, and $X = (X_1, X_2)$

$X : S \to \mathbb{R}^2$ is a random variable and we can consider probabilities of the form $P(X \in A_1 \times A_2) = P(X_1 \in A_1, X_2 \in A_2)$ for $A_i \subset \mathbb{R}$. All such probabilities together form the "distribution" of X.

But this is an excess of formality. Normally we work directly with the components, in this case X_1 and X_2, which are two discrete rvs and we can put probabilities of combinations of outcomes in a table.

table of probabilities

		X_1										
		2	3	4	5	6	7	8	9	10	11	12
X_2	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$
	1	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0
	2	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0
	3	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	0
	4	0	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	0	0
	5	0	0	0	0	0	$\frac{2}{36}$	0	0	0	0	0

"joint" distribution

<Insert cannabis joke here>

Such a table summarizes the distribution of X, which we'll just call the *joint distribution* of X_1 and X_2.

The table has all the values of the form $p(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$ and is called the *joint* probability mass function.

A joint pmf is non-negative, and its positive values sum to 1 (just like before).

The joint cdf is defined as: $P(X_1 \leq x_1, X_2 \leq x_2)$ for all $(x_1, x_2) \in \mathbb{R}^2$.

It is non-decreasing and right-continuous in both variables, and goes to 0 and 1 as both dimensions go to $\pm\infty$ respectively.

marginal distributions

The pmf of any of the component random variables can be recovered by summing over all the others, e.g.:

$$p_{X_1}(x_1) = \sum_{x_2} p(x_1, x_2),$$

(where $\sum_{x_2}$ denotes summing over all values that X_2 takes on.)

This is called a *marginal* pmf, which characterizes the distribution of that random variable.

example: marginal for X_2

		X_1											
		2	3	4	5	6	7	8	9	10	11	12	$p_{X_2}(x_2)$
X_2	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	$\frac{6}{36}$
	1	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{10}{36}$
	2	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	$\frac{8}{36}$
	3	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	0	$\frac{6}{36}$
	4	0	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	0	0	$\frac{4}{36}$
	5	0	0	0	0	0	$\frac{2}{36}$	0	0	0	0	0	$\frac{2}{36}$

volumes under surfaces

discrete analogue

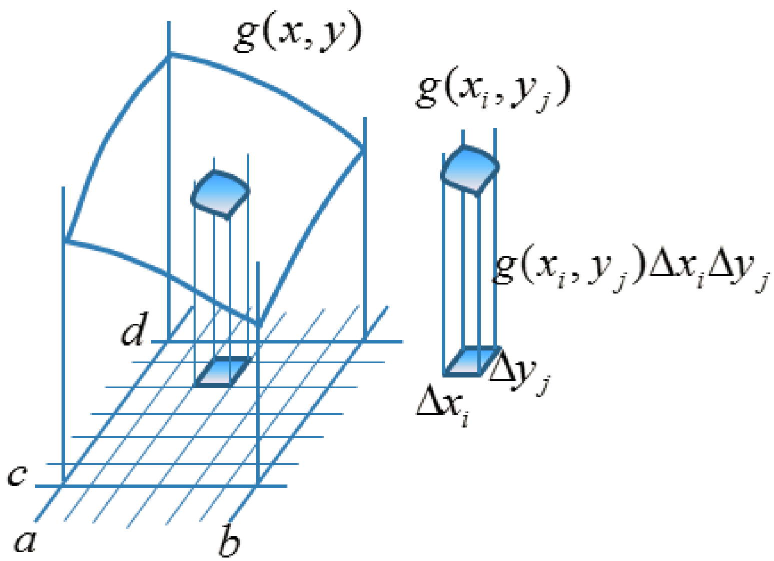
First, consider calculating in the discrete example:

$$P(4 \le X_1 \le 8,\ 2 \le X_2 \le 5)$$

		X_1				
		4	5	6	7	8
X_2	2	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$
	3	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0
	4	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$
	5	0	0	0	$\frac{2}{36}$	0

"double integral" crash course - I

A double ("Riemann") integral calculates the volume between the a rectangle $[a, b] \times [c, d]$ in the xy–plane and a function $g(x, y)$ in pretty much the same way $\int g(x)$ calculates the area under a curve.

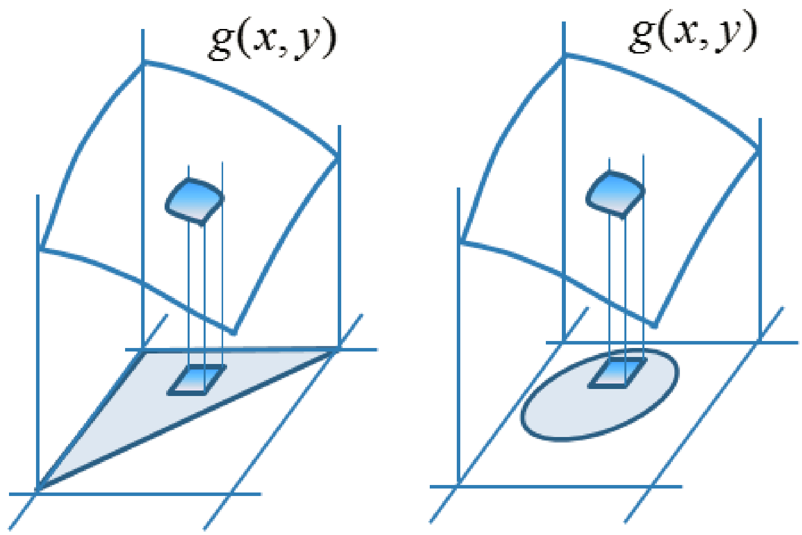


$$\int_a^b \int_c^d g(x, y)\, dx\, dy \approx$$

$$\sum_{i,j} g(x_i, y_i) \Delta x_i \Delta y_i$$

double integral crash course - II

In general the same idea applies to a volume between any region $B \subset \mathbb{R}^2$ and $g(x,y)$.



$$\iint_B g(x,y)\,dx\,dy$$

double integrals crash course - III

Two basic properties:

$$\iint\limits_B \left[\alpha_1 g_1(x,y) + \alpha_2 g_2(x,y)\right]\,dx\,dy =$$

$$\alpha_1 \iint\limits_B g_1(x,y)\,dx\,dy + \alpha_2 \iint\limits_B g_2(x,y)\,dx\,dy$$

When B_1 and B_2 are disjoint:

$$\iint\limits_{B_1 \cup B_2} g(x,y)\,dx\,dy =$$

$$\iint\limits_{B_1} g(x,y)\,dx\,dy + \iint\limits_{B_2} g(x,y)\,dx\,dy$$

double integrals crash course - IV

The actual calculation is just done one variable at a time, from the inside out (first holding the outside variable "constant")

Simple example (B is a rectangle): $g(x, y) = xy^2$. Consider:

$$\int_0^2 \int_1^3 xy^2 \, dx \, dy$$

And it makes no difference which order, but be careful when B is not a rectangle. Consider the integral of g over the region bounded by $x = 0$, $y = 0$, and $x + y = 1$.

joint continuous distributions

Reconsider measuring the weight X and (systolic) blood pressure Y of a randomly selected study participant. We'll be interested in things like:

$$P(70 < X < 80, 120 < Y < 140)$$

In this case X and Y are both continuous random variables. A fancy way to re-write the above would be to let $A = [70, 80] \times [120, 140]$ and get:

$$P((X, Y) \in A)$$

Definition: X and Y are *jointly continuous* if there is a $f(x, y)$ such that, for "all" $A \subset \mathbb{R}^2$

$$P((X, Y) \in A) = \iint_A f(x, y)\, dx\, dy$$

joint density

The function $f(x, y)$ is called the *joint density*, and is used to calculate probabilities.

Properties: $f \geq 0$ and $\iint\limits_{\mathbb{R}^2} f(x, y)\, dx\, dy = 1$

Example 1: $f(x, y) = 1$ on $0 < x,\ y < 1$ and 0 otherwise. ("Joint Uniform") Is this a density? Calculate: $P(X < 1/2, Y < 1/2)$ and $P(X < Y)$.

Example 2: (artificial) $f(x, y) = cxy$ on $0 < x < 1$, $0 < y < 2$, and 0 otherwise. Determine c. Calculate $P(X > 0.5, 0 < Y < 1)$ and $P(Y > X)$.

joint cdf - continuous case

The joint cdf $F(x, y) = P(X \leq x, Y \leq y)$ can be calculated:

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u, v) \, du \, dv$$

It is a mystery of multivariable calculus how to obtain f from F

"partial" derivatives crash course - I

Maybe you got this far in your co-requisite! But your lives are a mystery to me.

With a function $g(x, y)$ you can take the derivative with respect to one variable at a time, holding the other variable constant. Notation:

$$\frac{\partial}{\partial x} g(x, y) \quad \text{and} \quad \frac{\partial}{\partial y} g(x, y).$$

When g is "smooth" you get the nice result:

$$\frac{\partial}{\partial y}\left[\frac{\partial}{\partial x} g(x, y)\right] = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial y} g(x, y)\right],$$

and we just call this:

$$\frac{\partial^2}{\partial x \partial y} g(x, y).$$

joint cdf to joint density

Just take all the "partial" derivatives in any order you like.

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

"Proof: …"

Examples can be challenging! Consider $f(x, y) = xy$ on $0 < x < 1, 0 < y < 2$
(…to be revisited.)

marginal cdf and marginal density

Just like in the discrete case we can recover information about X and Y individually by "integrating out" the other variable. The marginal densities are:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

What about the marginal cdfs?

Continue the example on the previous slide.