Week 07 notes

Joint distributions

01 Theory

Joint distributions describe the probabilities of events associated with multiple random variables simultaneously.

In this course we consider only two variables at a time, typically called X and Y. It is easy to extend this theory to vectors of n random variables.

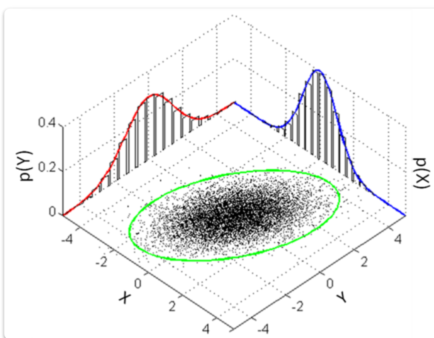
Joint PMF and joint PDF

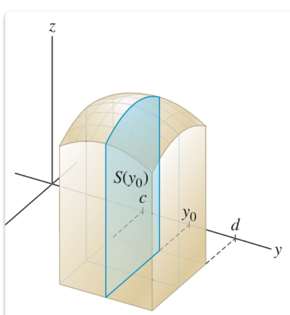
Discrete joint PMF:

$$P_{X,Y}(k,\ell) \quad = \quad P_{X,Y}\big[X=x,\, Y=\ell\big]$$

Continuous joint PDF:

$$f_{X,Y}(x,y) \quad = \quad \text{density at } (x,y)$$





Probabilities of events: Discrete case

If $B \subset \mathbb{R}^2$ is a set of points in the plane, then an *event $\mathcal{B}$* is formed by the set of all outcomes s mapped by X and Y to points in B:

$$\mathcal{B} = \Big\{ s \in S \,\Big|\, \big(X(s),\, Y(s)\big) \in B \Big\}$$

The probabilities of such events can be measured using the joint PMF:

$$P\big[(X,Y) \in B\big] = P[\mathcal{B}] \quad = \quad \sum_{(k,\ell)\in B} P_{X,Y}(k,\ell)$$

Probabilities of events: Continuous case

Let $\mathcal{V} = [a,b] \times [c,d] \subset \mathbb{R}^2$ be the rectangular region defined by $(x,y) \in \mathbb{R}^2$ such that $a \leq x \leq b$ and $c \leq y \leq d$. Then:

$$P\Big[(x,y) \in \mathcal{V}\Big] = P\Big[a \leq X \leq b,\ c \leq Y \leq d\Big] \quad = \quad \int_c^d \int_a^b f_{X,Y}(x,y)\,dx\,dy$$

For more general regions $\mathcal{V} \subset \mathbb{R}^2$:

$$P\Big[(X,Y) \in \mathcal{V}\Big] \quad = \quad \iint_{\mathcal{V}} f_{X,Y}(x,y)\,dA$$

The existence of a variable Y does not change the theory for a variable X considered by itself.

However, it is possible to *relate* the theory for X to the theory for (X,Y), in various ways.

The simplest relationship is the **marginal distribution** for X, which is merely the distribution of X *itself*, considered as a single random variable, but in a context where it is *derived from the joint* distribution for (X,Y).

Marginal PMF, marginal PDF

Marginal distributions are obtained from joint distributions by *summing* the probabilities over all possibilities of the *other* variable.

Discrete marginal PMF:

$$P_X(k) \quad = \quad \sum_\ell P_{X,Y}(k,\ell)$$

$$P_Y(\ell) \quad = \quad \sum_k P_{X,Y}(k,\ell)$$

Continuous marginal PMF:

$$f_X(x) \quad = \quad \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dy$$

$$f_Y(y) \quad = \quad \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx$$

Infinitesimal method

Suppose X has density $f_X(x)$ that is continuous at x_0. Then for small $dx > 0$:

$$P[x_0 < X \leq x_0 + dx] \quad \approx \quad f_X(x)\,dx$$

Suppose X and Y have joint density $f_{X,Y}(x,y)$ that is continuous at (x_0, y_0). Then for small dx, $dy > 0$:

$$P\Big[x_0 < X \leq x_0 + dx,\ y_0 < Y \leq y_0 + dy\Big] \quad \approx \quad f_{X,Y}(x_0,y_0)\,dxdy$$

⚠ Joint densities depend on coordinates

The density $f_{X,Y}(x,y)$ in these integration formulas depends on the way X and Y act as Cartesian coordinates and determine differential areas $dx\,dy$ as little rectangles.

To find a density $f_{R,\Theta}(r,\theta)$ in polar coordinates, for example, it is *not enough* to solve for $x(r,\theta)$ and $y(r,\theta)$ and plug into $f_{X,Y}$. We must consider the differential area $dx\,dy$ vs. $dr\,d\theta$. We find that $dx\,dy = r\,dr\,d\theta$. So we will add a factor of r. See an example below for details.

Joint densities may not exist

It is not always *possible* to form a joint PDF $f_{X,Y}$ from any two continuous RVs X and Y .

For example, if $X = Y$, then (X,Y) cannot have a joint PDF, since $P[X = Y] = 1$ but the integral over the region $X = Y$ will always be 0. (The area of a line is zero.)

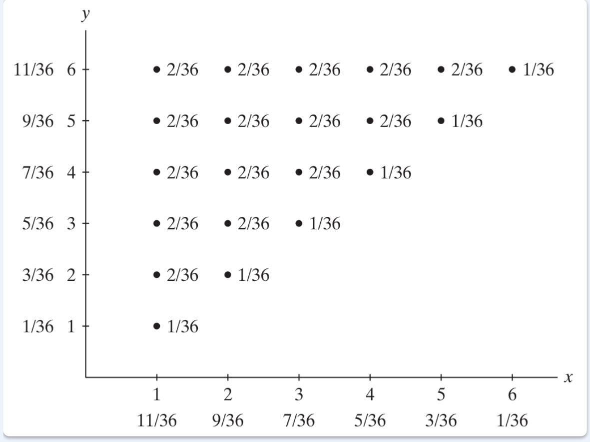
02 Illustration

Example - Smaller and bigger rolls

Roll two dice, and let X indicate the smaller of the numbers rolled, and let Y indicate the bigger number.

Make a chart showing the PMF. Compute the marginal probabilities, and write them in the margins of the chart.

Solution



Exercise - Reading a PMF table

Here is a joint PMF table:

$P_{Q,G}(q,g)$	$g=0$	$g=1$	$g=2$	$g=3$
$q=0$	0.06	0.18	0.24	0.12
$q=1$	0.04	0.12	0.16	0.08

Using the table, compute the following event probabilities:
(a) $P[Q=0]$
(b) $P[Q=G]$
(c) $P[G>1]$
(d) $P[G>Q]$

Exercise - Coin flipping

Flip a fair coin four times. Let X measure the number of heads in the first two flips, and let Y measure the total number of heads.

Make a chart showing the PMF. Compute the marginal probabilities, and write them in the margins of the chart.

Example - Marginal and event probability from joint density

Suppose the joint density of X and Y is given by:

$$f_{X,Y}(x,y) \quad = \quad \begin{cases} 2xe^{x^2-y} & y > x^2,\, x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

Find $f_Y(y)$ and $P\big[Y < 3X^2\big]$.

Solution
Compute the marginal PDF:

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx$$

$$\gg\gg \quad \int_0^{\sqrt{y}} 2xe^{x^2}e^{-y}\,dx \quad \gg\gg \quad 1 - e^{-y}$$

Find probability of the event $Y < 3X^2$:

$$P\big[Y < 3X^2\big] = \int_0^1 \int_{x^2}^{3x^2} 2xe^{x^2-y}\,dy\,dx$$

$$\gg\gg \quad \int_0^1 2xe^{x^2}\left(e^{-x^2} - e^{-3x^2}\right)dx$$

$$\gg\gg \quad \tfrac{1}{2}\left(1 + e^{-2}\right)$$

Exercise - Marginals from joint density

The joint PDF for X and Y is given by:

$$f_{X,Y}(x,y) = \begin{cases} 6\left(x+y^2\right)/5 & 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $f_X(x)$ and $f_Y(y)$.

Exercise - Event probability from joint density

The joint PDF for X and Y is given by:

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y} & x, y > 0 \\ 0 & \text{else} \end{cases}$$

Compute $P[X < Y]$.

03 Theory

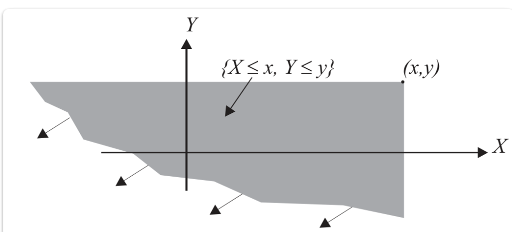
Joint CDF

The joint CDF of X and Y is defined by:

$$F_{X,Y}(x,y) \quad = \quad P\Big[X \leq x,\, Y \leq y\Big]$$

We can relate the joint CDF to the joint PDF using integration:

$$F_{X,Y}(x,y) \quad = \quad \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(s,t)\,ds\,dt$$



Conversely, if X and Y have a continuous joint PDF $f_{X,Y}(x,y)$ that is also *differentiable*, we can obtain the PDF from the CDF using partial derivatives:

$$f_{X,Y}(x,y) \quad = \quad \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$$

There is also a **marginal CDF** that is computed using a limit:

$$F_X(x) = \lim_{y \to +\infty} F_{X,Y}(x,y)$$

This could also be written, somewhat abusing notation, as $F_X(x) = F_{X,Y}(x,+\infty)$.

04 Illustration

Exercise - Properties of joint CDFs

(a) Show with a drawing that if both $x < x'$ and $y < y'$, we know:

$$F_{X,Y}(x,y) \leq F_{X,Y}(x',y')$$

(b) Explain why:

- $F_X(x) = F_{X,Y}(x,\infty)$
- $F_Y(y) = F_{X,Y}(\infty,y)$

(c) Explain why:

- $F_{X,Y}(x,-\infty) = 0$
- $F_{X,Y}(-\infty,y) = 0$

Independent random variables

05 Theory

Independent random variables

Random variables X, Y are **independent** when they satisfy the *product rule* for all valid subsets B_1, $B_2 \subset \mathbb{R}$:

$$P\big[X \in B_1,\, Y \in B_2\big] = P\big[X \in B_1\big] \cdot P\big[Y \in B_2\big]$$

Since $\{X \in B_1,\, Y \in B_2\} = \{X \in B_1\} \cap \{Y \in B_2\}$, this definition is equivalent to independence of *all events* constructible using the variables X and Y.

For discrete random variables, it is enough to check independence for simple events of type $\{X = k\}$ and $\{Y = \ell\}$ for k and ℓ any *possible values* of X and Y.

The independence criterion for random variables can be cast entirely in terms of their distributions and written using the PMFs or PDFs.

Independence using PMF and PDF

Discrete case:

$$P_{X,Y}(k,\ell) \quad = \quad P_X(k) \cdot P_Y(\ell)$$

Continuous case:

$$f_{X,Y}(x,y) \quad = \quad f_X(x) \cdot f_Y(y)$$

Independence via joint CDF

Random variables X and Y are independent when their CDFs obey the product rule:

$$F_{X,Y}(x,y) \quad = \quad F_X(x) \cdot F_Y(y)$$

06 Illustration

Example - Meeting in the park

A man and a woman arrange to meet in the park between 12:00 and 1:00 am. They both arrive at a random time with uniform distribution over that hour, and do not coordinate with each other.

Find the probability that the first person to arrive has to wait longer than 15 minutes for the second person to arrive.

Solution

Let X denote the time the man arrives. Use minutes starting from 12:00, so $X \in (0, 60)$. Let Y denote the time the woman arrives, using the same interval.

The probability we seek is:

$$P[X + 15 < Y] + P[Y + 15 < X]$$

Because X and Y are symmetrical in probability, these terms have the same value, so we just double the first one for our answer.

Since the arrivals are independent of each other, we have $f_{X,Y} = f_X \cdot f_Y$.

Since each arrival time is uniform over the interval, we have:

$$f_X(x) = \begin{cases} 1/60 & x \in (0, 60) \\ 0 & \text{otherwise,} \end{cases} \qquad f_Y(y) = \begin{cases} 1/60 & y \in (0, 60) \\ 0 & \text{otherwise} \end{cases}$$

Therefore the joint density is $f_{X,Y} = \left(\frac{1}{60}\right)^2$. Calculate:

$$\begin{aligned} 2P\big[X + 15 < Y\big] &= 2 \iint_{x+15<y} f(x, y)\, dx\, dy \\ &= 2 \iint_{x+15<y} f_X(x) f_Y(y)\, dx\, dy \\ &= 2 \int_{15}^{60} \int_0^{y-15} \left(\frac{1}{60}\right)^2 dx\, dy \\ &= \frac{2}{(60)^2} \int_{15}^{60} y - 15\, dy \\ &= \frac{9}{16} \end{aligned}$$

Example - Uniform disk: Cartesian vs. polar

Suppose that a point is chosen uniformly at random on the unit disk.

(a) Let X and Y be the Cartesian coordinates of the chosen point. Are X and Y independent?

(b) Let R and Θ give the polar coordinates of the chosen point. Are R and Θ independent?

Solution

(a)

Write $f_{X,Y}$ for the joint distribution of X and Y. We have:

$$f_{X,Y} = \begin{cases} 1/\pi & x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Then computing $f_X(x)$, we obtain:

$$\int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}} \frac{1}{\pi}\, dy \qquad \gg\gg \qquad \frac{2}{\pi}\sqrt{1-x^2}$$

$$\gg\gg \qquad f_X(x) \quad = \quad \begin{cases} \frac{2}{\pi}\sqrt{1-x^2} & x \in [-1,1] \\ 0 & \text{otherwise} \end{cases}$$

By similar reasoning, $f_Y(y) = \frac{2}{\pi}\sqrt{1-y^2}$ for $y \in [0,1]$.

The product $f_X(x) f_Y(y)$ is not equal to $f_{X,Y}(x,y)$, so X and Y are *not* independent. Information about the value of X does provide constraints on the possible values of Y, so this result makes sense.

(b)

To find the marginals $f_R(r)$ and $f_\Theta(\theta)$, the standard method is to integrate the density $f_{R,\Theta}$ in the opposite variables.

- ⚠ The probability density $f_{R,\Theta}(r,\theta)$ is not constant! The area of a differential sector $dr\, d\theta$ depends on r.

We can take two approaches to finding the density $f_{R,\Theta}$.

- (i) Area of a differential sector divided by total area $\quad = \dfrac{r\, dr\, d\theta}{\pi} \quad \gg\gg \quad \dfrac{r}{\pi} dr\, d\theta$
  - So the density is $f_{R,\Theta} = \frac{r}{\pi}$
- (ii) Compute CDF first, then use partial derivatives.

Elaborating on (ii): joint CDF then partials for joint PDF

The region 'below' a given point (r,θ), in polar coordinates, is a sector with area $\frac{\theta}{2\pi} \cdot \pi r^2$. The factor $\frac{\theta}{2\pi}$ is a percentage of the circle with area πr^2.

The density is a constant $\frac{1}{\pi}$ across the disk, so the CDF at (r,θ) is this same area times $\frac{1}{\pi}$. Thus:

$$F_{R,\Theta} = \frac{\theta r^2}{2\pi}$$

Then in polar coordinates the density is given by taking partial derivatives:

$$f_{R,\Theta}(r,\theta) \quad = \quad \frac{\partial^2}{\partial r \partial \theta}\left(\frac{1}{2\pi}\theta r^2\right) \quad \gg\gg \quad \frac{r}{\pi}$$

Once we have $f_{R,\Theta}$, integrate to get the marginals:

$$f_R(r) = \int_{\theta=0}^{2\pi} f_{R,\Theta}\, d\theta \quad \gg\gg \quad \int_0^{2\pi} \frac{r}{\pi}\, d\theta \quad \gg\gg \quad 2r$$

$$f_\Theta(\theta) = \int_{r=0}^{1} f_{R,\Theta}\, dr \quad \gg\gg \quad \int_0^{1} \frac{r}{\pi}\, dr \quad \gg\gg \quad \frac{1}{2\pi}$$

Check independence:

$$f_{R,\Theta} = \frac{r}{\pi} \qquad \overset{\checkmark}{=} \qquad (2r)\left(\frac{1}{2\pi}\right) = f_R \cdot f_\Theta$$

In this problem it is feasible to find the marginals *directly*, without computing the new density, only using some geometric reasoning.

Direct geometric approach

The probability $P\big[R \in (r, r+dr)\big]$ is the area (over π) of a thickened circle with radius r and thickness dr. The *circumference* of a circle at radius r is $2\pi r$. So the area of the thickened circle is $2\pi r\, dr$. So the probability is $2r\, dr$. This tells us that the marginal probability density is $P_R(r) = 2r$.

The probability $P\big[\Theta \in (\theta, \theta + d\theta)\big]$ is the area (over π) of a thin sector with radius 1 and angle $d\theta$. This area is $\frac{1}{2}1^2\, d\theta$. So the probability is $\frac{1}{2\pi}\, d\theta$. This tells us that the marginal probability density is $P_\Theta(\theta) = \frac{1}{2\pi}$.

These results agree with those of the ‘calculus’ approach above!