Multivariate Random Vectors

For simplicity, we mostly work with the bivariate case: X and Y are two rv's arising from the same experiment. An outcome $\omega \in \Omega$ leads to values $X(\omega)$ and $Y(\omega)$.

Think of (X, Y) as a random point in the plane whose location is known after performing the experiment.

The (joint) distribution of the bivariate random vector (X, Y) is determined by the (joint) cdf:

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y) \quad \text{for all } x, y \in \mathbb{R}.$$

More commonly used in calculations than the cdf are joint pmf's or pdf's.

The Jointly Discrete Case

If X and Y are discrete rv's, their distribution is determined by a (joint) pmf:

$$f_{X,Y}(x, y) = P(X = x, Y = y) \quad \text{for all } x, y \in \mathbb{R}$$

From now on, we shall drop the subscripts and write $f_{X,Y}(x, y)$ as $f(x, y)$, etc.

For any region $A \subset \mathbb{R}^2$,

$$P(\,(X, Y) \in A\,) = \sum_{(x,y) \in A} f(x, y)$$

For any function $g(\cdot, \cdot)$,

$$Eg(X,Y) = \sum_{(x,y)\in\mathbb{R}^2} g(x,y)f(x,y)$$

so long as $E|g(X,Y)| < \infty$.

Any function $f(\cdot, \cdot)$ which satisfies $f(x,y) \geq 0$ for all $x, y \in \mathbb{R}$, and

$$\sum_{(x,y)\in\mathbb{R}^2} f(x,y) = 1$$

can serve as a joint pmf.

The Jointly Continuous Case

If X and Y are "jointly continuous", their distribution is determined by a (joint) pdf $f_{X,Y}(\cdot, \cdot)$.

For any region $A \subset \mathbb{R}^2$,

$$P(\,(X,Y) \in A\,) = \iint_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

For any function $g(\cdot, \cdot)$,

$$Eg(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

so long as $E|g(X,Y)| < \infty$.

Any function $f(\cdot,\cdot)$ which satisfies $f(x,y) \geq 0$ for all $x, y \in \mathbb{R}$, and

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$$

can serve as a joint pdf.

<u>Connections between joint cdf and joint pdf</u>

If X and Y are jointly continuous with density $f(\cdot,\cdot)$, then

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(s,t)\mathrm{d}t\,\mathrm{d}s\,.$$

This is a special case of the formula for $P((X,Y) \in A)$ with $A = (-\infty, x) \times (-\infty, y)$.

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

at points (x,y) where $f(\cdot,\cdot)$ is continuous.

<u>Informal interpretation of joint pdf</u>

Let A be a very small rectangle with sides of length dx and dy which contains the point (x,y). Then

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

so long as f is continuous at the point (x,y).

The Mixed Case

If X is discrete and Y is continuous (or the other way around), I refer to (X, Y) as "mixed".

There are formulas for computing probabilities and expected values in the mixed case. These combine features of the jointly continuous and jointly discrete cases. The formulas are fairly intuitive. We won't cover them in detail, but they are used in a later example and one exercise (I think).

Example: X, Y discrete with joint pmf given in this table.

$Y \backslash X$	1	2	3	4
1	1/20	2/20	3/20	4/20
2	0	1/20	2/20	3/20
3	0	0	1/20	2/20
4	0	0	0	1/20

$f_{X,Y}(1,3) = 0$,
$f_{X,Y}(3,1) = 3/20$,
$f_{X,Y}(3,2) = 2/20$,
etc.

$$P(X=Y) = f(1,1) + f(2,2) + f(3,3) + f(4,4) = 4/20$$

$$P(X+Y \leq 4) = f(1,1) + f(2,1) + f(2,2) + f(3,1) = 1/20 + 2/20 + 1/20 + 3/20 = 7/20$$

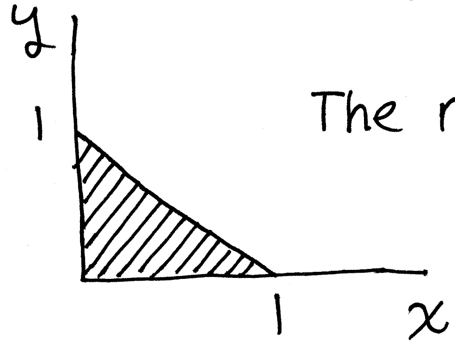
$$EXY = \sum xy\, f_{X,Y}(x,y)$$
$$= 1 \cdot 1 \cdot f(1,1) + 2 \cdot 1 \cdot f(2,1) + 3 \cdot 1 \cdot f(3,1) + \cdots + 4 \cdot 4 \cdot f(4,4)$$
$$= 1/20 + 4/20 + 9/20 + 16/20 + 4/20 + 12/20 + 24/20 + 9/20 + 24/20 + 16/20 = 119/20$$

Example (Jointly continuous distn.)

Joint pdf of (X,Y) is

$$f(x,y) = 6(1-x-y)\, I_D(x,y)$$

where $D = \{(x,y) : x>0, y>0, x+y<1\}$.



The region D

① Verify this is a pdf.

Clearly $f(x,y) > 0$ for $(x,y) \in D$,
$= 0$ for $(x,y) \notin D$.

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

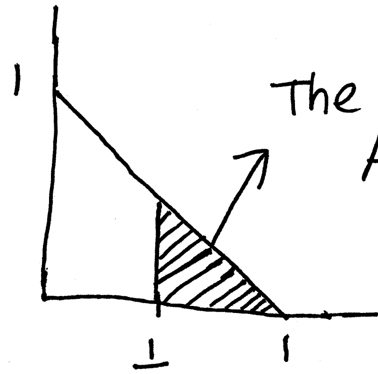
$$= \iint_D 6(1-x-y)\, dx\, dy$$

$$= \int_0^1 dx \int_0^{1-x} dy\; 6(1-x-y)$$

$$= \int_0^1 dx \left(-3(1-x-y)^2 \Big|_{y=0}^{1-x} \right)$$

$$= \int_0^1 3(1-x)^2\, dx = -(1-x)^3 \Big|_0^1 = 1$$

② Find $P(X > \frac{1}{2})$



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

$$= \iint_A f(x,y)\, dx\, dy$$

$$= \int_{\frac{1}{2}}^1 dx \int_0^{1-x} dy \ 6(1-x-y)$$

$$= \int_{\frac{1}{2}}^1 3(1-x)^2 dx = -(1-x)^3 \Big|_{\frac{1}{2}}^1 = \frac{1}{8}$$

③ Find EXY

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\, dx\, dy = \iint_D xy\, 6(1-x-y)\, dx\, dy$$

$$= \int_0^1 dx \int_0^{1-x} dy \overset{xy}{\uparrow} 6(1-x-y) = \frac{1}{20}$$

From a joint pmf or pdf

$$f_{X,Y}(x,y)$$

the marginal distributions are given by

Discrete Case (joint pmf)

$$f_X(x) = \sum_y f_{X,Y}(x,y)$$

$$f_Y(y) = \sum_x f_{X,Y}(x,y)$$

Continuous Case (joint pdf)

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

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

The conditional distributions are given by

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

(Valid for both discrete and continuous cases

Remarks:

For discrete rv's, since $f_{X,Y}(x,y) = P(X = x, Y = y)$ and $f_X(x) = P(X = x)$, the formulas for marginal and conditional distributions are immediate consequences of basic probability facts.

For a partition $\{B_k\}$,

$$P(A) = \sum_k P(A \cap B_k)\,.$$

Taking $A = \{X = x\}$ for some fixed x, and the partition to be $\{Y = y\}$ for $y \in \mathcal{Y}$ leads to

$$P(X = x) = \sum_{y \in \mathcal{Y}} P(X = x, Y = y)$$

which is the same as

$$f_X(x) = \sum_y f_{X,Y}(x,y)\,.$$

Similarly, from the basic definition of conditional probability we get

$$f_{Y|X}(y|x) = P(Y = y \,|\, X = x) = \frac{P(Y = y, X = x)}{P(X = x)} = \frac{f_{X,Y}(x,y)}{f_X(x)}\,.$$

For jointly continuous rv's, we may loosely argue the formulas for marginal and conditional distributions by analogy with the discrete case using

$$\begin{aligned} f_X(x)\,dx &= P\{X \in [x, x+dx)\} \quad \text{and} \\ f_{X,Y}(x,y)\,dx\,dy &= P\{X \in [x, x+dx), Y \in [y, y+dy)\} \end{aligned}$$

which hold in some vague limiting sense as $dx \to 0$ and $dy \to 0$.

Proper arguments in the jointly continuous case require more work. We may get the formula for the marginal pdf by differentiating the marginal cdf:

$$\begin{aligned} F_X(x) &= P(X \le x) \\ &= P(-\infty < X \le x, \; -\infty < Y < \infty) \\ &= \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f_{X,Y}(u,y)\, dy \right) du \end{aligned}$$

so that using the Fundamental Theorem of Calculus we obtain

$$f_X(x) = \frac{d}{dx} F_X(x) = \left(\int_{-\infty}^{\infty} f_{X,Y}(u,y)\, dy \right|_{u=x} = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy \,.$$

A proper derivation of the formula for the conditional density is harder and we omit it.

Note:

$f_{Y|X}(y \,|\, x)$ is a pdf (or pmf) as a function of y for every fixed value of x for which $f_X(x) > 0$.

For a fixed value x, the conditional pdf (or pmf) is just the joint pdf (or pmf) at the fixed value of x "normalized" to be a density in y. The denominator in

$$f_{Y|X}(y \,|\, x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

is just the normalization constant necessary to make the integral (or sum) over y equal to 1 for the fixed value of x.

Y \ X	1	2	3	4
1	1/20	2/20	3/20	4/20
2	0	1/20	2/20	3/20
3	0	0	1/20	2/20
4	0	0	0	1/20

Joint pmf

1/20	3/20	6/20	10/20

Marginal pmf of X

Conditional pmf's (for Y|X)

y	$f_{Y\|X}(y\|1)$	$f_{Y\|X}(y\|2)$
1	$\frac{1}{20}/\frac{1}{20} = 1$	$\frac{2}{20}/\frac{3}{20} = \frac{2}{3}$
2	$0/\frac{1}{20} = 0$	$\frac{1}{20}/\frac{3}{20} = \frac{1}{3}$
3	$0/\frac{1}{20} = 0$	$0/\frac{3}{20} = 0$
4	$0/\frac{1}{20} = 0$	$0/\frac{3}{20} = 0$

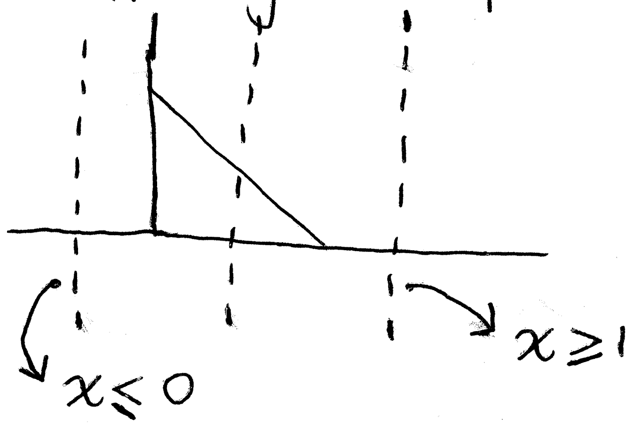
y	$f_{Y\|X}(y\|3)$	$f_{Y\|X}(y\|4)$
1	$\frac{3}{20}/\frac{6}{20} = \frac{1}{2}$	$\frac{4}{20}/\frac{10}{20} = \frac{4}{10}$
2	$\frac{2}{20}/\frac{6}{20} = \frac{1}{3}$	$\frac{3}{20}/\frac{10}{20} = \frac{3}{10}$
3	$\frac{1}{20}/\frac{6}{20} = \frac{1}{6}$	$\frac{2}{20}/\frac{10}{20} = \frac{2}{10}$
4	$0/\frac{6}{20} = 0$	$\frac{1}{20}/\frac{10}{20} = \frac{1}{10}$

Continuous Example: (X,Y) with joint pdf $f(x,y) = 6(1-x-y)\, I_D(x,y)$.

→ Find marginal $f_X(x)$

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

= integral over vertical line through the point $(x,0)$



$= 0$ for $x \leq 0$ or $x \geq 1$

$$= \int_0^{1-x} 6(1-x-y)\, dy \quad \text{for } 0 < x < 1$$

$$= -3(1-x-y)^2 \Big|_{y=0}^{1-x} = 3(1-x)^2$$

$f_X(x) = 3(1-x)^2$ for $0 < x < 1$
(= 0 otherwise)

$\rightarrow$ Find conditional density $f_{Y|X}(y|x)$.

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} \quad \text{when } f_X(x) > 0$$

$$= \frac{6(1-x-y)\,I(x>0, y>0, x+y<1)}{3(1-x)^2}$$

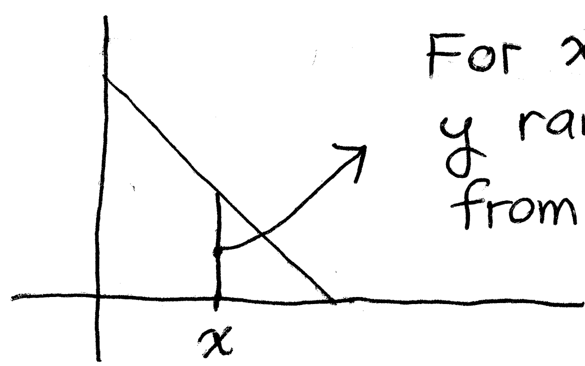
For $0<x<1$,

$$f_{Y|X}(y|x) = \frac{2(1-x-y)}{(1-x)^2}, \quad 0<y<1-x$$

$$(= 0 \text{ otherwise})$$

For $x \leq 0$ and $x \geq 1$,

$f_{Y|X}(y|x)$ is undefined.



For $x \in (0,1)$, y ranges from 0 to $1-x$.

Conditional Expected Values

$$\underbrace{E(g(Y) \mid X=x)}_{\text{function of } x} = \underbrace{\sum_y}_{\int_{-\infty}^{\infty} dy} g(y) f_{Y|X}(y|x)$$

Example (Table)

$$E(Y \mid X=1) = 1\cdot 1 + 2\cdot 0 + 3\cdot 0 + 4\cdot 0$$
$$= 1$$

$$E(Y \mid X=2) = 1\cdot 2/3 + 2\cdot 1/3 + 3\cdot 0 + 4\cdot 0$$
$$= 2/3 + 2/3 = 4/3$$

$$E(Y \mid X=3) = 1\cdot 3/6 + 2\cdot 2/6 + 3\cdot 1/6 + 4\cdot 0$$
$$= 10/6 = 5/3$$

$$E(Y \mid X=4) = 1\cdot 4/10 + 2\cdot 3/10 + 3\cdot 2/10 + 4\cdot 1/10$$
$$= 20/10 = 2$$

Can summarize: $E(Y \mid X=x) =$
$$1 + \tfrac{1}{3}(x-1)$$

Conditional Expected Values
(in jointly discrete example)

x	$E(Y\|X=x)$
1	1
2	4/3
3	5/3
4	2

Example of computation

y	$f_{Y\|X}(y\|4)$
1	.4
2	.3
3	.2
4	.1

$$E(Y|X=4) = 1\times.4 + 2\times.3 + 3\times.2 + 4\times.1$$
$$= 2.0$$

Notation:

$E(Y|X)$ is a random variable, which is a function of X.

= conditional expected value evaluated at the observed value of X.

$= \psi(X)$ where $\psi(x) = E(Y|X=x)$

In our example.

x	$E(Y\|X=x)$	$f_X(x)$
1	1	1/20
2	4/3	3/20
3	5/3	6/20
4	2	10/20

pmf of $E(Y|X)$

Example: $E(Y|X=x)$ (conditional mean) in continuous situation.

$$f_{X,Y}(x,y) = 6(1-x-y)\, I_D(x,y)$$

$$D = \{(x,y) : x>0, y>0, x+y<1\}$$

$$f_{Y|X}(y|x) = \frac{2(1-x-y)}{(1-x)^2} \quad \text{for } 0<y<1-x \quad (0<x<1)$$

$$E(Y|X=x) = \int_{-\infty}^{\infty} y\, f_{Y|X}(y|x)\, dy$$

$$= \int_0^{1-x} y\, \frac{2(1-x-y)}{(1-x)^2}\, dy$$

$$= \int_0^{1-x} \frac{y}{(1-x)}\, 2\left(1-\frac{y}{(1-x)}\right) \frac{dy}{(1-x)}(1-x)$$

$$= (1-x)\, 2\int_0^1 u(1-u)\, du = \frac{(1-x)}{3}$$

for $0<x<1$

$$E(Y|X=x) = \frac{(1-x)}{3} \quad \text{for } 0<x<1$$

(function of x)

$$E(Y|X) = \frac{(1-X)}{3}$$

(random variable)

Defn. (conditional variance)

$\mathrm{Var}(Y|X=x)$ is the variance of the conditional dist. of Y given $X=x$.

$$= \int y^2 f_{Y|X}(y|x)\,dy - \left(\int y f_{Y|X}(y|x)\,dy\right)^2 \quad \text{(continuous case)}$$

$$= E(Y^2|X=x) - (E(Y|X=x))^2$$

is a function of x.

$\mathrm{Var}(Y|X)$ is the conditional variance evaluated at the observed value of X

$= \psi(X)$ where $\psi(x) = \mathrm{Var}(Y|X=x)$.

is a r.v.

$$\mathrm{Var}(Y|X) = E(Y^2|X) - (E(Y|X))^2$$

Conditional Variances

(in jointly discrete example)

x	$E(Y^2 \mid X=x)$	$Var(Y \mid X=x)$
1	1	$1-1^2=0$
2	2	$2-(4/3)^2 = 2/9$
3	$10/3$	$\frac{10}{3}-\left(\frac{5}{3}\right)^2 = 5/9$
4	5	$5-2^2=1$

Example of Computation

$$E(Y^2 \mid X=4) = 1^2 \times .4 + 2^2 \times .3 + 3^2 \times .2 + 4^2 \times .1$$

$$= 5$$

Examples: Computing conditional variances

Example: (discrete distribution described by table)

Var$(Y \mid X = 4)$ is the variance of the pmf:

y	$f_{Y\mid X}(y \mid 4)$
1	4/10
2	3/10
3	2/10
4	1/10

From earlier work we know $E(Y \mid X = 4) = 2$ so that

$$\begin{aligned}
\text{Var}(Y \mid X = 4) &= E\left((Y - E(Y|X = 4))^2 \mid X = 4\right) \\
&= E((Y - 2)^2 \mid X = 4) \\
&= (1-2)^2 \cdot \frac{4}{10} + (2-2)^2 \cdot \frac{3}{10} \\
&\quad + (3-2)^2 \cdot \frac{2}{10} + (4-2)^2 \cdot \frac{1}{10} \\
&= 1
\end{aligned}$$

Alternatively,

$$\begin{aligned}
\text{Var}(Y \mid X = 4) &= E(Y^2 \mid X = 4) - (E(Y \mid X = 4))^2 \\
&= 1^2 \cdot \frac{4}{10} + 2^2 \cdot \frac{3}{10} + 3^2 \cdot \frac{2}{10} + 4^2 \cdot \frac{1}{10} - 2^2 \\
&= 1\,.
\end{aligned}$$

Example (jointly continuous distribution)

$$f(x,y) = 6(1-x-y)I_D(x,y)$$

For $0 < x < 1$,

$$\text{Var}(Y \,|\, X = x) = E(Y^2 \,|\, X = x) - (E(Y|X = x))^2$$

From earlier work $E(Y|X = x) = (1-x)/3$.

$$\begin{aligned}
E(Y^2 \,|\, X = x) &= \int_{-\infty}^{\infty} y^2 f_{Y|X}(y \,|\, x)\, dy \\
&= \int_0^{1-x} y^2 \cdot \frac{2(1-x-y)}{(1-x)^2}\, dy \\
&= 2(1-x)^2 \int_0^{1-x} \left(\frac{y}{1-x}\right)^2 \cdot \left(1 - \frac{y}{1-x}\right) \cdot \frac{dy}{1-x} \\
&= 2(1-x)^2 \int_0^1 u^2(1-u)\, du \quad (\text{Let } u = \frac{y}{1-x}.) \\
&= 2(1-x)^2 \cdot \frac{\Gamma(3)\Gamma(2)}{\Gamma(5)} \\
&= \frac{(1-x)^2}{6}.
\end{aligned}$$

Thus

$$\begin{aligned}
\text{Var}(Y \,|\, X = x) &= \frac{(1-x)^2}{6} - \left(\frac{(1-x)}{3}\right)^2 \\
&= \frac{(1-x)^2}{18} \quad \text{for } 0 < x < 1.
\end{aligned}$$

Note: $\text{Var}(Y \,|\, X) = \dfrac{(1-X)^2}{18}$ is a random variable.