§ 3.7 Conditional Distributions

In ch 2, defined the conditional prob. of an event A, given an event B, as

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

provided $P(B) \neq 0$.

Suppose now X & Y are discrete r.v.'s with joint density $f(x,y)$ & marginals $g(x)$, $h(y)$ resp. and let $A = \{X = x\}$, $B = \{Y = y\}$. Then

$$P(X = x \mid Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

$$= \frac{f(x,y)}{h(y)}$$

provided $h(y) = P(Y = y) \neq 0$.

If we call this cond. prob. $f(x|y)$ (to indicate that x is a variable & y is fixed), we get the defn.

<u>Defn</u>. Let X, Y be discrete r.v.'s with joint PDF $f(x,y)$ and let $g(x), h(y)$ be the marginal PDF's of X, Y (resp.).

The fn.

$$f(x|y) = \frac{f(x,y)}{h(y)}, \quad h(y) \neq 0$$

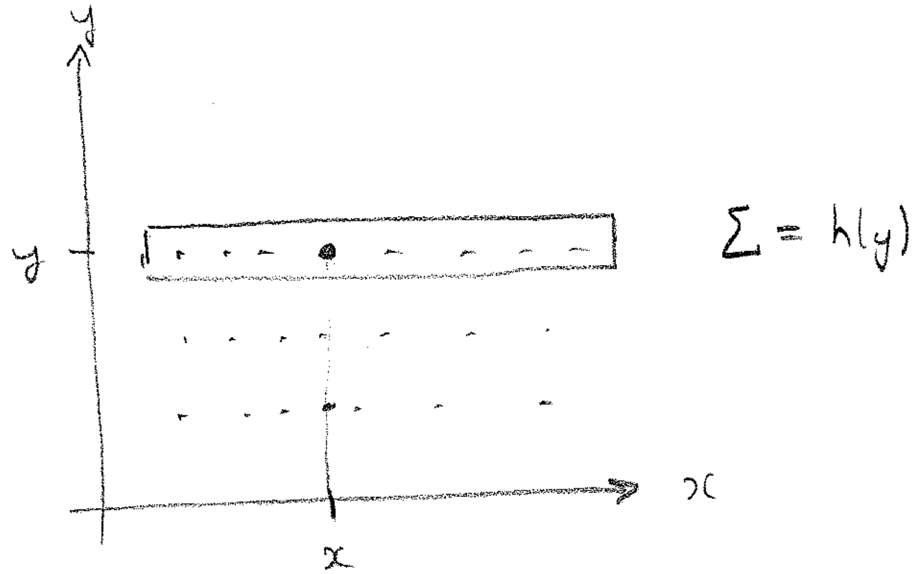
for each x within the range of X is called the <u>conditional distribution of X given $Y=y$</u>. Similarly the fn.

$$w(y|x) = \frac{f(x,y)}{g(x)}, \quad g(x) \neq 0$$

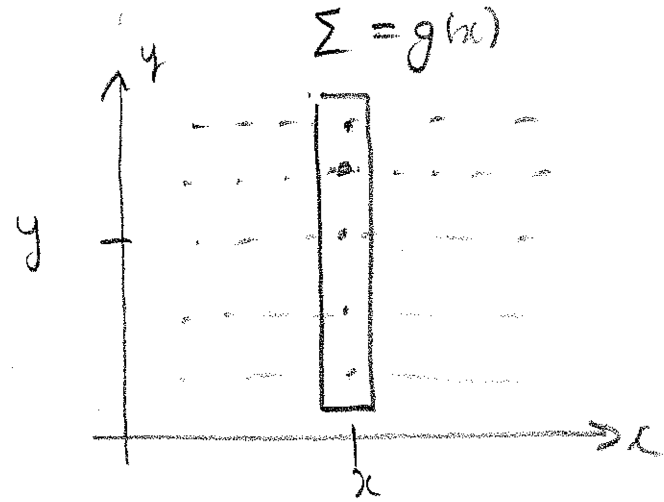
for each y within the range of Y is called the <u>conditional distribution of Y given $X=x$</u>

Pictures

$f(x|y)$



$w(y|x)$



Note $\quad f(x|y) \geq 0 \quad$ and

$$\sum_x f(x|y) = \sum_x \frac{f(x,y)}{h(y)} = \frac{1}{h(y)} \sum_x f(x,y) = \frac{h(y)}{h(y)} = 1$$

Similarly $\quad w(y|x) \geq 0 \quad$ and

$$\sum_y w(y|x) = \sum_y \frac{f(x,y)}{g(x)} = \frac{1}{g(x)} \sum_y f(x,y) = \frac{g(x)}{g(x)} = 1$$

Thus $f(x|y)$ and $g(y|x)$ can both serve as the PDF's of discrete r.v.'s by Thm 3.1 (p.74).

<u>Ex</u>. The tablets again!

		$x=0$	$x=1$	$x=2$	$h(y)$
y	0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{7}{12}$
	1	$\frac{2}{9}$	$\frac{1}{6}$		$\frac{7}{18}$
	2	$\frac{1}{36}$			$\frac{1}{36}$
	$g(x)$	$\frac{5}{12}$	$\frac{1}{2}$	$\frac{1}{12}$	

$$f(0|1) = \frac{f(0,1)}{h(1)} = \frac{\frac{2}{9}}{\frac{7}{18}} = \frac{4}{7}$$

$$f(1|1) = \frac{f(1,1)}{h(1)} = \frac{\frac{1}{6}}{\frac{7}{18}} = \frac{3}{7}$$

$$f(2|1) = \frac{f(2,1)}{h(1)} = \frac{0}{\frac{7}{18}} = 0$$

Note. $f(0|1) + f(1|1) + f(2|1) = 1$

as we saw earlier.

$$w(0|0) = \frac{f(0,0)}{g(0)} = \frac{\frac{1}{6}}{5/12} = \frac{2}{5}$$

NOTE SWAP IN ORDER OF VARIABLES !!

$$w(1|0) = \frac{f(0,1)}{g(0)} = \frac{\frac{2}{9}}{\frac{5}{12}} = \frac{8}{15}$$

$$w(2|0) = \frac{f(0,2)}{g(0)} = \frac{\frac{1}{36}}{\frac{5}{12}} = \frac{1}{15}$$

and

$$w(0|0) + w(1|0) + w(2|0) = 1$$

as we'd expect.

For cts. r.v.'s we have.

Defn. If $f(x,y)$ is the value of a joint PDF of the cts. r.v.'s X, Y at (x,y) and $h(y)$ is the value of the marginal density of Y at y, the f^n given by

$$f(x|y) = \frac{f(x,y)}{h(y)}, \quad h(y) \neq 0, \quad -\infty < x < \infty$$

is called the <u>conditional density</u> of X given $Y=y$.

Correspondingly, if $g(x)$ is the value of the marginal density of X at x, the f^n given by

$$w(y|x) = \frac{f(x,y)}{g(x)}, \quad g(x) \neq 0, \quad -\infty < y < \infty$$

is called the <u>conditional density</u> of Y given $X=x$.

Note $f(x|y) \geq 0$ and

$$\int_{-\infty}^{\infty} f(x|y)\,dx = \int_{-\infty}^{\infty} \frac{f(x,y)}{h(y)}\,dx = \frac{1}{h(y)} \int_{-\infty}^{\infty} f(x,y)\,dx$$

$$= \frac{h(y)}{h(y)} = 1.$$

Also $w(y|x) \geq 0$ and

$$\int_{-\infty}^{\infty} w(y|x)\,dy = \int_{-\infty}^{\infty} \frac{f(x,y)}{g(x)}\,dy = \frac{1}{g(x)} \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$= \frac{g(x)}{g(x)} = 1.$$

Thus (provided $h(y) \neq 0$, $g(x) \neq 0$), $f(x|y)$ and $g(y|x)$ can serve as the PDFs of cts r.v.'s by Thm 3.5 (P. 86).

Ex. For the density we had earlier

$$f(x,y) = \begin{cases} \frac{2}{3}(x+2y), & 0 < x,y < 1 \\ 0 & \text{elsewhere} \end{cases},$$

find the conditional density of X given $Y=y$ and use it to evaluate $P(X \le \frac{1}{2} \mid Y = \frac{1}{2})$.

Soln: Recall we had the marginal PDF for Y

$$h(y) = \begin{cases} \frac{1}{3}(1+4y), & 0 < y < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

Then for $0 < y < 1$ and $0 < x < 1$

$$f(x \mid y) = \frac{f(x,y)}{h(y)} = \frac{\frac{2}{3}(x+2y)}{\frac{1}{3}(1+4y)}$$

$$= \frac{2x+4y}{1+4y}$$

with $f(x \mid y) = 0$ elsewhere $(-\infty < x \le 1,\ 1 \le x < \infty)$.
(No info for $-\infty < y \le 1,\ 1 \le y < \infty$).

Then

$$f(x\,|\,\tfrac{1}{2}) = \frac{2x + 4\cdot\frac{1}{2}}{1 + 4\cdot\frac{1}{2}}$$

$$= \frac{2x+2}{3}$$

and so

$$P(X \le \tfrac{1}{2}\,|\,Y = \tfrac{1}{2}) = \int_0^{\frac{1}{2}} \frac{2x+2}{3}\,dx = \frac{5}{12}.$$

Ex. Given the joint PDF

$$f(x,y) = \begin{cases} 4xy, & 0 < x, y < 1 \\ 0 & \text{elsewhere} \end{cases}$$

find the marginal densities of X and Y and the conditional density of X given Y=y.

Soln. For $0<x<1$,

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 4xy\,dy = \left[2xy^2\right]_0^1 = 2x$$

while $g(x) = 0$ elsewhere.

For $0<y<1$

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^1 4xy\,dx = \left[2x^2y\right]_0^1 = 2y$$

while $h(y) = 0$ elsewhere.

Thus, if $0<x<1$ $(0<y<1)$

$$f(x|y) = \frac{f(x,y)}{h(y)} = \frac{4xy}{2y} = 2x$$

with $f(x|y) = 0$ elsewhere (in x).

For more than 2 r.v.'s the situation can be a bit more complicated.

E.g. if $f(x_1, x_2, x_3, x_4)$ is the joint PDF of the four discrete r.v.'s X_1, X_2, X_3, X_4,

$$p(x_3 \mid x_1, x_2, x_4) = \frac{f(x_1, x_2, x_3, x_4)}{g(x_1, x_2, x_4)},$$

$$g(x_1, x_2, x_4) := \sum_{x_3} f(x_1, x_2, x_3, x_4) \neq 0$$

gives the cond. distr. of X_3 at x_3 given $X_1 = x_1$, $X_2 = x_2$, $X_4 = x_4$.

$$q(x_2, x_4 \mid x_1, x_3) = \frac{f(x_1, x_2, x_3, x_4)}{m(x_1, x_3)},$$

$$m(x_1, x_3) := \sum_{x_2, x_4} f(x_1, x_2, x_3, x_4) \neq 0$$

gives the values of the joint conditional distribution of X_2, X_4 at (x_2, x_4) given $X_1 = x_1$, $X_3 = x_3$.

$$r(x_2, x_3, x_4 \mid x_1) = \frac{f(x_1, x_2, x_3, x_4)}{b(x_1)},$$

$$b_1(x_1) = \sum_{x_2, x_3, x_4} f(x_1, x_2, x_3, x_4) \neq 0$$

gives the joint conditional PDF of X_2, X_3, X_4 at (x_2, x_3, x_4) given $X_1 = x_1$.

RULE OF THUMB

The factor one needs to divide the joint density (e.g. $f(x_1, x_2, x_3, x_4)$) is the sum of the joint PDF over all those variables over which one is <u>NOT</u> conditioning.

Independence

Recall that in the last example, we found

$$f(x|y) = 2x \quad \text{- does not depend on } y$$

while in the example before that

$$f(x|y) = \frac{2x+4y}{1+4y} \quad \text{- depends on } y.$$

Suppose $f(x|y)$ does not depend on y.

Write $f(x|y) = k(x)$.

Then (in the discrete case), recalling the marginal PDF $g(x) = \sum_y f(x,y)$, we get

$$g(x) = \sum_y f(x,y) = \sum_y f(x|y)\, h(y)$$

$$= \sum_y k(x) h(y)$$

$h(y)$ is a PDF!

$$= k(x) \sum_y h(y) = k(x) \cdot 1 = k(x)$$

Thus $k(x) = p(x|y) = g(x)$

and so

$$f(x,y) = f(x|y)h(y) = g(x)h(y) \qquad (h(y) \neq 0).$$

This motivates the following def$\underline{n}$.

<u>Defn</u>. If $f(x_1, x_2, \ldots, x_n)$ is the joint PDF of the n discrete r.v.'s $X_1, X_2, \ldots X_n$ at $(x_1, \ldots, x_n)$, and $f_i(x_i)$ is the value of the marginal distribution of X_i at x_i for $1 \leq i \leq n$, then the n r.v.'s are <u>independent</u> iff

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1) \cdot f_2(x_2) \cdots f_n(x_n)$$

for all $x_1, x_2, \ldots, x_n$ within their range.

<u>Note</u>: A similar defn holds for cts. r.v's.

Let's see how this relates to the independence we are familiar with $(P(A \cap B) = P(A)P(B))$ for the simplest case of two discrete r.v.'s.

Basic idea is that for two r.v.'s, events associated with each of them $(\{X \in C\}, \{Y \in D\})$ should be independent in the usual sense.

If X_1, X_2 are ind., so that $f(x,y) = g(x)\,h(y)$.

Then

$$P(X=x, Y=y) = f(x,y) = g(x)\,h(y)$$
$$= P(X=x) \cdot P(Y=y)$$

or

$$P(\{X=x\} \cap \{Y=y\}) = P(X=x) \cdot P(Y=y). \quad (*)$$

This argument is reversible and if we were to take $(*)$ instead as our defn of independence, we would again find that $f(x,y) = g(x)\,h(y)$ and so we would recover the earlier defn. of independence for two discrete r.v.'s.

Ex For n independent flips of a coin, let X_i be the no. of heads (0 or 1) obtained in the ith flip for $1 \le i \le n$. Find the joint PDF of $X_1, \ldots, X_n$.

Soln.

Each X_i has prob distn $f_i(x_i)$ with

$$f_i(x_i) = \begin{cases} \frac{1}{2}, & x_i = 0, 1 \\ 0 & \text{otherwise.} \end{cases}$$

Since these r.vs. are ind. the joint PDF is $f(x_1, \ldots, x_n) = f_1(x_1) f_2(x_2) \cdots f_n(x_n)$

so that

$$f(x_1, \ldots, x_n) = \begin{cases} \frac{1}{2^n}, & x_i = 0 \text{ or } 1 \text{ for each } 1 \le i \le n \\ 0 & \text{otherwise.} \end{cases}$$

Ex. Given the independent r.v.'s X_1, X_2, X_3 with PDFs

$$f_1(x_1) = \begin{cases} e^{-x_1}, & x_1 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$f_2(x_2) = \begin{cases} 2e^{-2x_2}, & x_2 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$f_3(x_3) = \begin{cases} 3e^{-3x_3}, & x_3 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

find the joint PDF and use it to get $P(X_1 + X_2 \leq 1, X_3 > 1)$.

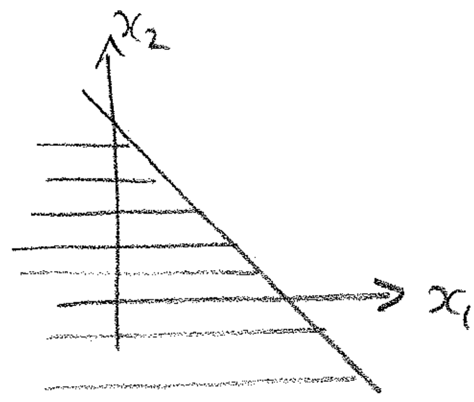
Soln. Since X_1, X_2, X_3 are ind., the joint PDF $f(x_1, x_2, x_3)$ is given by

$f(x_1, x_2, x_3) = f_1(x_1) f_2(x_2) f_3(x_3)$ or

$$f(x_1, x_2, x_3) = \begin{cases} 6e^{-x_1 - 2x_2 - 3x_3}, & x_1, x_2, x_3 > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Then

$P(X_1 + X_2 \leq 1, X_3 > 1)$



$$= \int_1^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{1-x_2} f(x_1, x_2, x_3)\, dx_1\, dx_2\, dx_3$$

$$= \int_1^{\infty}\int_0^{\infty}\int_0^{1-x_2} f(x_1, x_2, x_3)\, dx_1\, dx_2\, dx_3$$

as $f(x_1, x_2, x_3) = 0$ if $x_1 < 0$ or $x_2 < 0$

$$= \int_1^{\infty}\int_0^{1}\int_0^{1-x_2} f(x_1, x_2, x_3)\, dx_1\, dx_2\, dx_3$$

as if $x_2 > 1$, then $1 - x_2 < 0$ and $x_1 < 1 - x_2$ $\Rightarrow x_1 < 0$ and so $f(x_1, x_2, x_3) = 0$

$$= \int_1^{\infty}\int_0^{1}\int_0^{1-x_2} 6e^{-x_1 - 2x_2 - 3x_3}\, dx_1\, dx_2\, dx_3$$

$$= \int_1^{\infty}\int_0^{1} 6e^{-2x_2} e^{-3x_3} \int_0^{1-x_2} e^{-x_1}\, dx_1$$

$$= \int_1^\infty \int_0^1 6 e^{-2x_2} e^{-3x_3} \left[-e^{-x_1} \right]_0^{1-x_2} dx_2\, dx_3$$

$$= \int_1^\infty \int_0^1 6 e^{-2x_2} e^{-3x_3} \left(1 - e^{x_2-1} \right) dx_2\, dx_3$$

$$= \int_1^\infty 6 e^{-3x_3} \left\{ \int_0^1 \left(e^{-2x_2} - e^{-x_2-1} \right) dx_2 \right\} dx_3$$

$$= \int_1^\infty 6 e^{-3x_3} \left[-\frac{1}{2} e^{-2x_2} + e^{-x_2-1} \right]_0^1 dx_3$$

$$= \int_1^\infty 6 e^{-3x_3} \left(-\frac{1}{2} e^{-2} + e^{-2} - \left(-\frac{1}{2} + e^{-1} \right) \right) dx_3$$

$$= 6 \left(\frac{1}{2} - e^{-1} + \frac{1}{2} e^{-2} \right) \cdot \lim_{z \to \infty} \int_1^z e^{-3x_3}\, dx$$

$$= 6 \left(\frac{1}{2} - e^{-1} + \frac{1}{2} e^{-2} \right) \lim_{z \to \infty} \left[-\frac{e^{-3x_3}}{3} \right]_1^z$$

$$= 6\left(\frac{1}{2} - e^{-1} + \frac{1}{2}e^{-2}\right) \lim_{z \to \infty}\left(-\frac{e^{-3z}}{3} + \frac{e^{-3}}{3}\right)$$

$$= 6\left(\frac{1}{2} - e^{-1} + \frac{1}{2}e^{-2}\right) \cdot \frac{e^{-3}}{3}$$

$$= \left(1 - 2e^{-1} + \frac{1}{2}e^{-2}\right) \cdot e^{-3}.$$

Ex: The tablets (yet) again.

Recall that X (# of aspirin) and Y (# of sedatives) had the following joint distribution and marginals

		x = 0	1	2	$h(y)$
y	0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{7}{12}$
	1	$\frac{2}{9}$	$\frac{1}{6}$		$\frac{7}{18}$
	2	$\frac{1}{36}$			$\frac{1}{36}$
	$g(x)$	$\frac{5}{12}$	$\frac{1}{2}$	$\frac{1}{12}$	

Q. Are X and Y indep?

A. No. e.g. $P(X=2, Y=2) = 0$, but

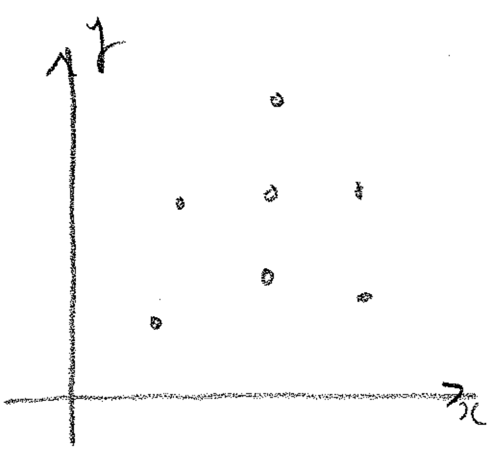
$$P(X=2)\cdot P(Y=2) = \frac{1}{12} \times \frac{1}{36} = \frac{1}{432}$$

FACT If the set of values where the joint PDF is non-zero is not a 'rectangle' (ie. a cartesian product of a set of x-values and a set of y-values), then X and Y are not indep.

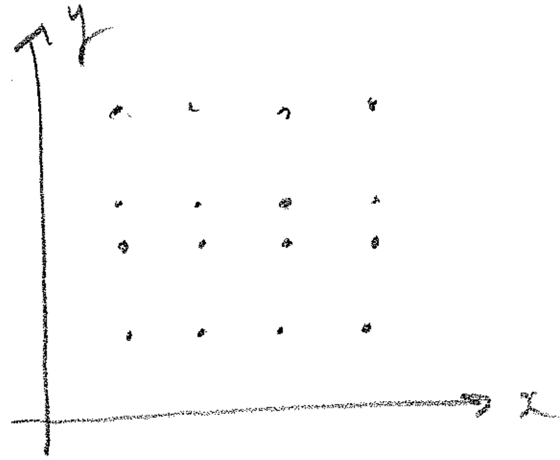
True for both discrete and cts. r.v.'s.

Similar conclusions hold for the joint PDF of any number of r.v.'s.

Eg. (Discrete case).



Definitely not independent



May (or may not) be independent. Needs further investigation.