Conditional distributions

Introduction

Recall $P(A|B)$ is the probability that A occurs given that B occurs:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{for } P(B) > 0$$

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \quad \text{for } \mathbb{P}(A) > 0$$

We want to apply this idea to random variables

Already seen one example: truncated distributions, e.g, fix x_0

$$P(\underbrace{X \le x}_{A} | \underbrace{X > x_0}_{B}) = \frac{P(\overbrace{X \le x}^{A}, \overbrace{X > x_0}^{B})}{\underbrace{P(X > x_0)}_{B}} = \frac{P(x_0 < X \le x)}{P(X > x_0)} = \frac{F(x) - F(x_0)}{1 - F(x_0)} \qquad x > x_0$$

Definitions: Conditional distribution (given general event A)

Suppose we observe A with $P(A) > 0$

1. **conditional cdf**

$$F(x|A) = P(X \le x|A) = \frac{P(A, X \le x)}{P(A)}$$

$$F(x) = \mathbb{P}(X \le x)$$

discrete

2. **conditional pmf/pdf**

pmf

$$f(x|A) = P(X = x|A) = \frac{P(A, X = x)}{P(A)}, \quad x \in \mathbb{R} \qquad \text{pmf if } X \text{ is discrete}$$

pdf

$$f(x|A) = \frac{dF(x|A)}{dx}, \quad x \in \mathbb{R} \qquad \text{pdf if } X \text{ is continuous}$$

event

$$f(x) = \frac{dF(x)}{dx}$$

For bivariate (X, Y), we're interested in conditional pmf/pdf $f(x|Y = y)$

Conditional distributions

Bivariate case: discrete distributions

- Discrete case: if both X and Y are discrete, we are interested mostly in events

$$\boxed{\text{“}X = x\text{” and “}Y = y\text{”}}$$

and the idea of conditional probability of events suggests the importance of the ratio:

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

→ joint pmf of (X,Y)

→ marginal of X or pmf of X

- *Definition:* For jointly discrete r.v.s X, Y with pmf $f(x, y)$, for each x with $f_X(x) > 0$, we call

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

the **conditional pmf of** Y **given** $X = x$ (that specifies a distribution of Y given $X = x$).

Similarly, for each y with $f_Y(y) > 0$, we call

$$\boxed{f(x|y)} = P(X = x|Y = x) = \frac{f(x, y)}{f_Y(y)}$$

the **conditional pmf of** X **given** $Y = y$

- One can verify that $f(x|y)$ is a pmf for a given y with $f_Y(y) > 0$

$\boxed{f(x|y)}$ is Conditional pmf/pdf

(i) $f(x|y) = \dfrac{\mathbb{P}(X=x \cap Y=y)}{\mathbb{P}(Y=y)} \geq 0$

(ii) $\displaystyle\sum_x f(x|y) = \sum_x \frac{f(x,y)}{f(y)} = \frac{1}{f_Y(y)} \sum_x f_{X,Y}(x,y) = \frac{f_Y(y)}{f_Y(y)} = 1$

↑ fix

Conditional distributions

Bivariate case: discrete distributions

- Note on notation/interpretation:
  1. In the jointly discrete case, for a given/fixed y with $P(Y = y) > 0$, $f(x|y)$ is a single (conditional) pmf
  2. BUT, $f(x|y)$ can also denote a family of pmfs, one conditional pmf for each given value of y
  3. $f(x|y)$ can be used to represent either situation

Example: simple discrete one again

		x		
		$\boxed{1}$	2	3
y	3	1/12	1/12	1/6
	2	1/12	1/6	1/12
	1	1/6	1/12	1/12
		1/3		

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

Y	$f(y\|1)$
3	$\frac{f(3,1)}{1/3} = \frac{1/12}{1/3} = 1/4$
2	$\frac{f(2,1)}{1/3} = \frac{1/12}{1/3} = 1/4$
1	$\frac{f(1,1)}{1/3} = \frac{1/6}{1/3} = 1/2$

$$P(Y \geq 2 | X = 1)$$
$$= \underbrace{P(Y=2 | X=1)} + P(Y=3 | X=1)$$
$$= 1/4 + 1/4 = 1/2$$

Y	$f(y\|2)$
3	1/4
2	1/2
1	1/4

Y	$f(y\|3)$
3	1/2
2	1/4
1	1/4

$$P(Y \geq 2 | X = 2)$$
$$= 1/2 + 1/4 = 3/4$$

Conditional distributions

Bivariate case: continuous distributions

Definition: For jointly continuous r.v.s X, Y with pdf $f(x,y)$, for each x with $f_X(x) > 0$, we define the **conditional pdf of** Y **given** $X = x$ as

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

which specifies the pdf for the continuous distribution of Y given $X = x$.

For y with $f_Y(y) > 0$, we define the **conditional pdf of** X **given** $Y = y$ as

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

Note: The conditional cdf* of X given $Y = y$ is then given by

$$F(x|y) = \boxed{P(X \le x|Y = y)} \int_{-\infty}^{x} f(t|y)dt = \boxed{\int_{-\infty}^{x} \frac{f(t,y)}{f_Y(y)}dt},$$

$F(x) = P(X \le x)$

$P(X \le x | Y = y)$ $= \int_{-\infty}^{x} f_{X,Y}(t|y)\,dt$

just the usual integral of a pdf which is $f(x|y) = f(x,y)/f_Y(y)$ here

$P(X \le x) = \int_{-\infty}^{x} f_X(t)\,dt$

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

*Technically, for a continuous r.v. Y, it holds that $P(Y = y) = 0$ for any y so that *one needs to* $\boxed{\textit{define}}$ $F(x|y) = P(X \le x|Y = y)$ *as a limit* for the conditioning to truly make sense: when $f_Y(y) > 0$,

$$\begin{aligned}
F(x|y) &= \lim_{h\to 0} P(X \le x|y \le Y \le y+h) \overset{\text{def}}{=} \lim_{h\to 0} \frac{P(X \le x, y \le Y \le y+h)}{P(y \le Y \le y+h)} \\
&= \lim_{h\to 0} \frac{\int_y^{y+h} \int_{-\infty}^{x} f(t,s)dt ds}{F_Y(y+h) - F_Y(y)} \\
&= \lim_{h\to 0} \frac{[\int_{-\infty}^{y+h} g(s)ds - \int_{-\infty}^{y} g(s)ds]/\cancel{h}}{[F_Y(y+h) - F_Y(y)]/\cancel{h}} \qquad g(s) = \int_{-\infty}^{x} f(t,s)dt \\
&= \frac{g(y)}{f_Y(y)} = \frac{\int_{-\infty}^{x} f(t,y)dt}{f_Y(y)} = \int_{-\infty}^{x} \frac{f(t,y)\,dt}{f_Y(y)}
\end{aligned}$$

$P(X \le x | Y = y)$

$P[(X,Y) \in R]$
where
$R = (-\infty, x] \times [y, y+h]$

Conditional distributions

Example in continuous case

Consider our continuous example: pdf $\underline{f(x,y) = \boxed{1/x}}$, $0 < y < x < 1$.

We've previously found $\underline{f_X(x) = 1}$, $\underbrace{0 < x < 1}$, and $\boxed{f_Y(y) = -\log y}$ $\underbrace{0 < y < 1}$.

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{1/x}{-\log y} = \frac{1}{x}\underbrace{\frac{1}{-\log y}} \quad \text{for } y<x<1$$

fixed

$$f_{Y|X}(y|x) = \frac{f_{Y,X}(y,x)}{f_X(x)} = \frac{1/x}{1} = \frac{1}{x} \quad \text{for } \underbrace{0<y<x}$$

$\Longrightarrow$ What is the distribution of $Y|X=x$?

Ans: $Y|X=x \sim \text{Uniform}(0,x)$

$H \sim \text{Uni}(a,b)$ If $f_H(h) = \begin{cases} \frac{1}{b-a} & \forall\, a<h<b \\ 0 & \text{ow} \end{cases}$

$$Y=1/2 \Rightarrow \mathbb{P}(X \le 3/4 \mid Y=1/2) = \int_{-\infty}^{3/4} f_{X|Y}(t,1/2)\,dt$$

$$= \int_{-\infty}^{3/4} \frac{f_{X,Y}(t,1/2)\,dt}{f_Y(1/2)}$$

$$= \int_{1/2}^{3/4} \frac{1/t}{-\log(1/2)}\,dt$$

$$= \frac{1}{-\log(1/2)} \log t \Big]_{1/2}^{3/4}$$

$$= -\left[\frac{\log 3/4 - \log 1/2}{\log 1/2}\right] = 1 - \frac{\log 3/4}{\log 1/2}$$