2. Conditional Expectation (9/10/04; cf. Ross)

Intro / Definition

Examples

Conditional Expectation

Computing Probabilities by Conditioning

Intro / Definition

Recall conditional probability: $\text{Pr}(A|B) = \text{Pr}(A \cap B)/\text{Pr}(B)$ if $\text{Pr}(B) > 0$.

Suppose that X and Y are jointly discrete RV's. Then if $\text{Pr}(Y = y) > 0$,

$$\text{Pr}(X = x|Y = y) = \frac{\text{Pr}(X = x \cap Y = y)}{\text{Pr}(Y = y)} = \frac{f(x, y)}{f_Y(y)}$$

$\text{Pr}(X = x|Y = 2)$ defines the probabilities on X given that $Y = 2$.

Definition: If $f_Y(y) > 0$, then $f_{X|Y}(x|y) \equiv \frac{f(x,y)}{f_Y(y)}$ is the **conditional pmf/pdf of X given $Y = y$.**

Remark: Usually just write $f(x|y)$ instead of $f_{X|Y}(x|y)$.

Remark: Of course, $f_{Y|X}(y|x) = f(y|x) = \frac{f(x,y)}{f_X(x)}$.

Old Discrete Example: $f(x,y) = \text{Pr}(X = x, Y = y)$.

	$X=1$	$X=2$	$X=3$	$X=4$	$f_Y(y)$
$Y=1$	.01	.07	.09	.03	.2
$Y=2$	**.20**	**.00**	**.05**	**.25**	**.5**
$Y=3$	.09	.03	.06	.12	.3
$f_X(x)$	.3	.1	.2	.4	1

Find $f(x|2)$.

Then

$$f(x|2) = \frac{f(x,2)}{f_Y(2)} = \frac{f(x,2)}{0.5} = \begin{cases} 0.4 & \text{if } x = 1 \\ 0 & \text{if } x = 2 \\ 0.1 & \text{if } x = 3 \\ 0.5 & \text{if } x = 4 \\ 0 & \text{otherwise} \end{cases}$$

Old Cts Example:

$$f(x,y) = \frac{21}{4}x^2y, \quad \text{if } x^2 \le y \le 1$$

$$f_X(x) = \frac{21}{8}x^2(1-x^4), \quad \text{if } -1 \le x \le 1$$

$$f_Y(y) = \frac{7}{2}y^{5/2}, \quad \text{if } 0 \le y \le 1$$

Find $f(y|X = 1/2)$.

$$
\begin{aligned}
f(y|\frac{1}{2}) &= \frac{f(\frac{1}{2}, y)}{f_X(\frac{1}{2})} \\
&= \frac{\frac{21}{4} \cdot \frac{1}{4} y}{\frac{21}{8} \cdot \frac{1}{4} \cdot (1 - \frac{1}{16})}, \quad \text{if } \tfrac{1}{4} \le y \le 1 \\
&= \frac{32}{15} y, \quad \text{if } \tfrac{1}{4} \le y \le 1
\end{aligned}
$$

More generally,

$$\begin{aligned} f(y|x) &= \frac{f(x,y)}{f_X(x)} \\ &= \frac{\frac{21}{4}x^2 y}{\frac{21}{8}x^2(1-x^4)}, \quad \text{if } x^2 \le y \le 1 \\ &= \frac{2y}{1-x^4} \quad \text{if } x^2 \le y \le 1. \end{aligned}$$

Note: $2/(1-x^4)$ is a constant with respect to y, and we can check to see that $f(y|x)$ is a legit condl pdf:

$$\int_{x^2}^{1} \frac{2y}{1-x^4}\, dy = 1.$$

Typical Problem: Given $f_X(x)$ and $f(y|x)$, find $f_Y(y)$.

Steps: (1) $f(x,y) = f_X(x)f(y|x)$

(2) $f_Y(y) = \int_{\Re} f(x,y)\,dx$.

Example: $f_X(x) = 2x, \;\; 0 < x < 1$.

Given $X = x$, suppose that $Y|x \sim \mathrm{U}(0,x)$. Now find $f_Y(y)$.

Solution: $Y|x \sim \mathrm{U}(0, x) \;\Rightarrow\; f(y|x) = 1/x, \;\; 0 < y < x$.

So

$$\begin{aligned} f(x,y) &= f_X(x) f(y|x) \\ &= 2x \cdot \frac{1}{x}, \;\; \text{if } 0 < x < 1 \text{ and } 0 < y < x \\ &= 2, \;\; \text{if } 0 < y < x < 1. \end{aligned}$$

Thus,

$$f_Y(y) = \int_{\Re} f(x,y)\,dx = \int_y^1 2\,dx = 2(1-y), \;\; 0 < y < 1.$$

Conditional Expectation

Usual definition of expectation:

$$\mathsf{E}[Y] = \begin{cases} \sum_y y f(y) & \text{discrete} \\ \int_{\Re} y f(y)\, dy & \text{continuous} \end{cases}$$

$f(y|x)$ is the conditional pdf/pmf of Y given $X = x$.

Definition: The **conditional expectation** of Y given $X = x$ is

$$\mathsf{E}[Y|X = x] \equiv \begin{cases} \sum_y y f(y|x) & \text{discrete} \\ \int_{\Re} y f(y|x)\, dy & \text{continuous} \end{cases}$$

Note that $\mathsf{E}[Y|X=x]$ is a function of x.

Example: Suppose that

$$f(y|X=2) \;=\; \begin{cases} 0.2 & \text{if } y=1 \\ 0.3 & \text{if } y=2 \\ 0.5 & \text{if } y=3 \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\mathsf{E}[Y|X=2] \;=\; \sum_y y f(y|2) \;=\; 1(.2)+2(.3)+3(.5) \;=\; 2.3.$$

Old Cts Example:

$$f(x,y) = \frac{21}{4}x^2y, \quad \text{if } x^2 \le y \le 1.$$

Recall that

$$f(y|x) = \frac{2y}{1-x^4} \quad \text{if } x^2 \le y \le 1.$$

Thus,

$$\mathsf{E}[Y|x] = \int_{\Re} y f(y|x)\,dy = \frac{2}{1-x^4}\int_{x^2}^{1} y^2\,dy = \frac{2}{3}\cdot\frac{1-x^6}{1-x^4}.$$

Theorem (double expectations): $\mathsf{E}[\mathsf{E}(Y|X)] = \mathsf{E}[Y]$.

Remarks: Yikes, what the heck is this!? The exp value (averaged over all X's) of the conditional exp value (of $Y|X$) is the plain old exp value (of Y).

Think of the outside exp value as the exp value of $h(X) = \mathsf{E}(Y|X)$. Then the Law of the Unconscious Statistician miraculously gives us $\mathsf{E}[Y]$.

Proof (cts case): By the Unconscious Statistician,

$$\begin{aligned}
\mathsf{E}[\mathsf{E}(Y|X)] &= \int_{\Re} \mathsf{E}(Y|x) f_X(x)\, dx \\
&= \int_{\Re} \left(\int_{\Re} y f(y|x)\, dy \right) f_X(x)\, dx \\
&= \int_{\Re} \int_{\Re} y f(y|x) f_X(x)\, dx\, dy \\
&= \int_{\Re} y \int_{\Re} f(x,y)\, dx\, dy \\
&= \int_{\Re} y f_Y(y)\, dy \;=\; \mathsf{E}[Y].
\end{aligned}$$

Old Example: Suppose $f(x,y) = \frac{21}{4}x^2y$, if $x^2 \le y \le 1$. Find $\mathsf{E}[Y]$ **two ways**.

By previous examples, we know that

$$f_X(x) = \frac{21}{8}x^2(1-x^4), \quad \text{if } -1 \le x \le 1$$

$$f_Y(y) = \frac{7}{2}y^{5/2}, \quad \text{if } 0 \le y \le 1$$

$$\mathsf{E}[Y|x] = \frac{2}{3} \cdot \frac{1-x^6}{1-x^4}.$$

Solution #1 (old, boring way):

$$\mathsf{E}[Y] = \int_{\Re} y f_Y(y)\, dy = \int_0^1 \frac{7}{2} y^{7/2}\, dy = \frac{7}{9}.$$

Solution #2 (new, exciting way):

$$\begin{aligned}\mathsf{E}[Y] &= \mathsf{E}[\mathsf{E}(Y|X)] \\ &= \int_{\Re} \mathsf{E}(Y|x) f_X(x)\, dx \\ &= \int_{-1}^{1} \left(\frac{2}{3} \cdot \frac{1 - x^6}{1 - x^4}\right)\left(\frac{21}{8} x^2 (1 - x^4)\right) dx = \frac{7}{9}.\end{aligned}$$

Notice that both answers are the same (good)!

Believe it or not, sometimes it's easier to calculate E$[Y]$ indirectly by using our double expectation trick.

Example: An alternative way to calculate the mean of the Geom(p).

Let $N \sim$ Geom(p), e.g., N could be the number of coin flips before H appears.

Let

$$Y = \begin{cases} 1 & \text{if first flip is H} \\ 0 & \text{otherwise} \end{cases}.$$

We'll apply a "standard conditioning argument" (in the discrete case) to compute E[N].

$$\begin{aligned}
\mathsf{E}[N] &= \mathsf{E}[\mathsf{E}(N|Y)] \\
&= \sum_y \mathsf{E}(N|y) f_Y(y) \\
&= \mathsf{E}(N|Y=0)\mathsf{Pr}(Y=0) + \mathsf{E}(N|Y=1)\mathsf{Pr}(Y=1) \\
&= (1+\mathsf{E}[N])(1-p) + 1(p).
\end{aligned}$$

Solving, we get $\mathsf{E}[N] = 1/p$.

Theorem (expectation of a random number of RV's):

Suppose that $X_1, X_2, \ldots$ are independent RV's, all with the same mean. Also suppose that N is a nonnegative, integer-valued RV, that's independent of the X_i's. Then

$$\mathsf{E}\left(\sum_{i=1}^{N} X_i\right) = \mathsf{E}[N]\mathsf{E}[X_1].$$

Proof: By double expectation, we have

$$\begin{aligned}
\mathsf{E}\left(\sum_{i=1}^{N} X_i\right) &= \mathsf{E}\left[\mathsf{E}\left(\sum_{i=1}^{N} X_i \middle| N\right)\right] \\
&= \sum_{n=1}^{\infty} \mathsf{E}\left(\sum_{i=1}^{N} X_i \middle| N = n\right) \mathsf{Pr}(N = n) \\
&= \sum_{n=1}^{\infty} \mathsf{E}\left(\sum_{i=1}^{n} X_i \middle| N = n\right) \mathsf{Pr}(N = n) \\
&= \sum_{n=1}^{\infty} \mathsf{E}\left(\sum_{i=1}^{n} X_i\right) \mathsf{Pr}(N = n) \\
&= \sum_{n=1}^{\infty} n\mathsf{E}[X_1]\mathsf{Pr}(N = n) \\
&= \mathsf{E}[X_1] \sum_{n=1}^{\infty} n\mathsf{Pr}(N = n). \quad \diamondsuit
\end{aligned}$$

Example: Suppose the number of times we roll a die is $N \sim$ Pois(10). If X_i denotes the value of the ith toss, then the expected number of rolls is

$$\mathsf{E}\left(\sum_{i=1}^{N} X_i\right) = \mathsf{E}[N]\mathsf{E}[X_1] = 10(3.5) = 35. \quad \diamondsuit$$

Theorem: Under the same conditions as before,

$$\mathsf{Var}\left(\sum_{i=1}^{N} X_i\right) = \mathsf{E}[N]\mathsf{Var}(X_1) + (\mathsf{E}[X_1])^2\mathsf{Var}(N).$$

Proof: See, for instance, Ross. $\diamondsuit$

Computing Probabilities by Conditioning

Let A be some event, and define the RV Y as:

$$Y = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}.$$

Then

$$\mathsf{E}[Y] = \sum_y y f_Y(y) = \Pr(Y = 1) = \Pr(A).$$

Similarly, for any RV X, we have

$$\begin{aligned} \mathsf{E}[Y|X=x] &= \sum_y y f_Y(y|x) \\ &= \mathsf{Pr}(Y=1|X=x) \\ &= \mathsf{Pr}(A|X=x). \end{aligned}$$

Further, since $\mathsf{E}[Y] = \mathsf{E}[\mathsf{E}(Y|X)]$, we have

$$\begin{aligned} \mathsf{Pr}(A) &= \mathsf{E}[Y] \\ &= \mathsf{E}[\mathsf{E}(Y|X)] \\ &= \int_{\Re} \mathsf{E}[Y|x] dF_X(x) \\ &= \int_{\Re} \mathsf{Pr}(A|X = x) dF_X(x). \end{aligned}$$

Example/Theorem: If X and Y are independent continuous RV's, then

$$\Pr(Y < X) = \int_{\Re} F_Y(x) f_X(x)\, dx,$$

where $F_Y(\cdot)$ is the c.d.f. of Y and $f_X(\cdot)$ is the p.d.f. of X.

Proof: (Actually, there are many proofs.) Let the event $A = \{Y < X\}$. Then

$$\begin{aligned}\text{Pr}(Y < X) &= \int_{\Re} \text{Pr}(Y < X | X = x) f_X(x)\, dx \\ &= \int_{\Re} \text{Pr}(Y < x | X = x) f_X(x)\, dx \\ &= \int_{\Re} \text{Pr}(Y < x) f_X(x)\, dx \\ &\quad \text{(since } X, Y \text{ are indep).} \quad \diamondsuit\end{aligned}$$

Example: If $X \sim \text{Exp}(\mu)$ and $Y \sim \text{Exp}(\lambda)$ are independent RV's. Then

$$\begin{aligned} \Pr(Y < X) &= \int_{\Re} F_Y(x) f_X(x)\, dx \\ &= \int_0^\infty (1 - e^{-\lambda x}) \mu e^{-\mu x}\, dx \\ &= \frac{\lambda}{\lambda + \mu}. \quad \diamondsuit \end{aligned}$$

Example/Theorem: If X and Y are independent continuous RV's, then

$$\Pr(X+Y<a) = \int_{\Re} F_Y(a-x) f_X(x)\, dx,$$

where $F_Y(\cdot)$ is the c.d.f. of Y and $f_X(\cdot)$ is the p.d.f. of X. The quantity $X+Y$ is called a *convolution*.

Proof:

$$\begin{aligned}
\Pr(X+Y<a) &= \int_{\Re} \Pr(X+Y<a|X=x) f_X(x)\,dx \\
&= \int_{\Re} \Pr(Y<a-x|X=x) f_X(x)\,dx \\
&= \int_{\Re} \Pr(Y<a-x) f_X(x)\,dx
\end{aligned}$$

(since X, Y are indep). $\diamondsuit$

Example: Suppose $X, Y \overset{\text{iid}}{\sim} \text{Exp}(\lambda)$. Note that

$$F_Y(a-x) = \begin{cases} 1 - e^{-\lambda(a-x)} & \text{if } a - x \geq 0 \text{ and } x \geq 0 \\ & \quad \text{(i.e., } 0 \leq x \leq a) \\ 0 & \text{if otherwise} \end{cases}$$

$$\begin{aligned} \Pr(X + Y < a) &= \int_{\Re} F_Y(a-x) f_X(x)\, dx \\ &= \int_0^a (1 - e^{-\lambda(a-x)}) \lambda e^{-\lambda x}\, dx \\ &= 1 - e^{-\lambda a} - \lambda a e^{-\lambda a}, \quad \text{if } a \geq 0. \end{aligned}$$

$$\frac{d}{da} \Pr(X + Y < a) = \lambda^2 a e^{-\lambda a}, \quad a \geq 0.$$

This implies that $X + Y \sim \text{Gamma}(2, \lambda)$. $\diamondsuit$