Conditional distributions, expectations

When X and Y are discrete we have

$$\mathsf{E}(Y|X=x) = \sum_y yP(Y=y|X=x)$$

for any x for which $P(X=x)$ is positive.

Defines a function of x.

This function evaluated at X gives rv which is ftn of X denoted

$$\mathsf{E}(Y|X).$$

Example: $Y|X = x \sim$ Binomial(x, p). Since mean of a Binomial(n, p) is np we find

$$\mathsf{E}(Y|X = x) = px$$

and

$$\mathsf{E}(Y|X) = pX$$

Notice you simply replace x by X.

Here are some properties of the function

$$\mathsf{E}(Y|X = x)$$

1) Suppose A is a function defined on the range of X. Then

$$\mathsf{E}(A(X)Y|X = x) = A(x)\mathsf{E}(Y|X = x)$$

and so

$$\mathsf{E}(A(X)Y|X) = A(X)\mathsf{E}(Y|X)$$

Second assertion follows from first. Note that if $Z = A(X)Y$ then Z is discrete and

$$P(Z = z) = \sum_{x,y} P(Y = y, X = x)1(z = A(x)y)$$

Also

$$\begin{aligned} &P(Z = z|X = x) \\ &\quad = \frac{\sum_y P(Y = y, X = x)1(z = A(x)y)}{P(X = x)} \\ &\quad = \sum_y P(Y = y|X = x)1(z = A(x)y) \end{aligned}$$

Thus

$$
\begin{aligned}
&\mathsf{E}(Z|X=x)\\
&\quad = \sum_z zP(Z=z|X=x)\\
&\quad = \sum_z \sum_y zP(Y=y|X=x)1(z=A(x)y)\\
&\quad = \sum_z \sum_y A(x)yP(Y=y|X=x)1(z=A(x)y)\\
&\quad = A(x)\sum_y yP(Y=y|X=x)\sum_z 1(z=A(x)y)\\
&\quad = A(x)\sum_y yP(Y=y|X=x)
\end{aligned}
$$

2) Repeated conditioning: if X, Y and Z discrete then

$$\mathsf{E}\left\{\mathsf{E}(Z|X,Y)|X\right\} = \mathsf{E}(Z|X)$$
$$\mathsf{E}\left\{\mathsf{E}(Y|X)\right\} = \mathsf{E}(Y)$$

3) Additivity

$$\mathsf{E}(Y+Z|X) = \mathsf{E}(Y|X) + \mathsf{E}(Z|X)$$

4) Putting the first two items together gives

$$\begin{aligned} &\mathsf{E}\left\{\mathsf{E}(A(X)Y|X)\right\} = \qquad (1)\\ &\mathsf{E}\left\{A(X)\mathsf{E}(Y|X)\right\} = \mathsf{E}(A(X)Y) \end{aligned}$$

Definition of E$(Y|X)$ when X and Y are not assumed to discrete:

E$(Y|X)$ is rv which is measurable function of X satisfying(1).

Existence is measure theory problem.

Aside on "measurable": what sorts of events can be defined in terms of a family $\{Y_i : i \in I\}$?

Natural: any event of form $(Y_{i_1}, \ldots, Y_{i_k}) \in C$ is "defined in terms of the family" for any finite set $i_1, \ldots, i_k$ and any (Borel) set C in S^k.

For countable S: each singleton $(s_1, \ldots, s_k) \in S^k$ Borel. So every subset of S^k Borel.

Natural: if you can define each of a sequence of events A_n in terms of the Ys then the definition "there exists an n such that (definition of A_n) …" defines $\cup A_n$.

Natural: if A is definable in terms of the Ys then A^c can be defined from the Ys by just inserting the phrase "It is not true that" in front of the definition of A.

So family of events definable in terms of the family $\{Y_i : i \in I\}$ is a σ-field which includes every event of the form $(Y_{i_1}, \ldots, Y_{i_k}) \in C$. We call the smallest such σ-field, $\mathcal{F}(\{Y_i : i \in I\})$, the σ-field generated by the family $\{Y_i : i \in I\}$.

Suppose X is discrete and $X^* = g(X)$ is a one to one transformation of X. Since $X = x$ is the same event as $X^* = g(x)$ we find

$$\mathsf{E}(Y|X = x) = \mathsf{E}(Y|X^* = g(x))$$

Let $h^*(u)$ denote the function $\mathsf{E}(Y|X^* = u)$ and $h(u) = \mathsf{E}(Y|X = u)$. Then

$$h(x) = h^*(g(x))$$

Thus

$$h(X) = h^*(g(X)) = h^*(X^*)$$

This just means

$$\mathsf{E}(Y|X) = \mathsf{E}(Y|X^*)$$

Interpretation.

Formula is "obvious".

Example: Toss coin $n = 20$ times. Y is indicator of first toss is a heads. X is number of heads and X^* number of tails. Formula says:

$$\mathrm{E}(Y|X = 17) = \mathrm{E}(Y|X^* = 3)$$

In fact for a general k and n

$$\mathrm{E}(Y|X = k) = \frac{k}{n}$$

so

$$\mathrm{E}(Y|X) = \frac{X}{n}$$

At the same time

$$\mathrm{E}(Y|X^* = j) = \frac{n-j}{n}$$

so

$$\mathrm{E}(Y|X^*) = \frac{n-X^*}{n}$$

But of course $X = n - X^*$ so these are just two ways of describing the same random variable.

Another interpretation: Rv X partitions Ω into countable set of events of the form $X = x$.

Other rv X^* partitions Ω into the same events.

Then values of $\mathsf{E}(Y|X^* = x^*)$ are same as values of $\mathsf{E}(Y|X = x)$ but labelled differently.

To form $\mathsf{E}(Y|X)$ take value ω, compute $X(\omega)$ to determine member A of the partition we being conditionsed on, then write down corresponding $\mathsf{E}(Y|A)$.

Hence conditional expectation depends only on partition of Ω.

X not discrete: replace partition with σ-field. Suppose X and X^* 2 rvs such that $\mathcal{F}(X) = \mathcal{F}(X^*)$. Then:

- There is g Borel,one to one with one to one Borel inverse s.t. $X^* = g(X)$.

- $\text{E}(Y|X) = \text{E}(Y|X^*)$ almost surely.

In other words $\text{E}(Y|X)$ depends *only* on the σ-field generated by X. We write

$$\text{E}(Y|\mathcal{F}(X)) = \text{E}(Y|X)$$

Def'n: Suppose $\mathcal{G}$ is sub-σ-field of $\mathcal{F}$. X is $\mathcal{G}$ measurable if, for every Borel B

$$\{\omega : X(\omega) \in B\} \in \mathcal{G}\,.$$

Def'n: $\mathsf{E}(Y|\mathcal{G})$ is any $\mathcal{G}$ measurable rv s.t. for every $\mathcal{G}$ measurable rv variable A we have

$$\mathsf{E}(AY) = \mathsf{E}\left\{A\mathsf{E}(Y|\mathcal{G})\right\}\,.$$

Again existence is measure theory problem.