Conditional Probability

$(\Omega, \mathcal{F}, \mathbb{P})$

Previously: for $X \in L^1$, $\mathcal{G} \subseteq \mathcal{F}$ constructed $\mathbb{E}[X | \mathcal{G}]$
(Radon-Nikodym derivative)

Q: $\mathbb{P}[A | B] = ?$ (for $A, B \in \mathcal{F}$)

Natural choice: $\mathbb{P}[A | \mathcal{G}] := \mathbb{E}[\mathbb{1}_A | \mathcal{G}] = Z_A$

($A \in \mathcal{F}$; Z_A: $\mathcal{G}$-meas r.v. well-defined $\mathbb{P}$-a.s.)

Consider $\mathcal{G} = \sigma(Y)$. $\rightsquigarrow \mathbb{P}[A | Y]$. (r.v. $\in \mathbb{R}$; $\sigma(Y)$)

Recall motivating examples:

Ex 1: (Discrete) pmf $p(x,y) > 0$, i.e. $\mathbb{P}[X=x, Y=y] = p(x,y)$. ($(x,y) \in \mathcal{X} \times \mathcal{Y}$ finite)

$$\rightsquigarrow \mathbb{P}[X=x | Y=y] := \frac{p(x,y)}{\sum_{x'} p(x',y)} = p_y(x) \text{ pmf param by } y \in \mathcal{Y}$$

Ex 2: (Continuous) pdf $p(x,y)$ cts. > 0, i.e. $\mathbb{P}[(X,Y) \in A] = \int_A p(x,y)\, dx\, dy$ ($(x,y) \in \mathbb{R}^2$, $\mathbb{R}^2 \supseteq A$)

$$\rightarrow p(x|y) := \frac{p(x,y)}{\int_{\mathbb{R}} p(x',y)\, dx'} = p_y(x) \text{ pdf param by } y \in \mathbb{R}.$$

(or $p(x | Y=y)$)

Q: Does / how does measure-theoretic $\mathbb{P}[A | Y]$ recover Ex 1-2? ($= Z_A$)

Z_A is $\sigma(Y)$-meas. r.v.

Doob-Dynkin Lem $\Longrightarrow \exists f_A : \mathbb{R} \to \mathbb{R}$ meas. s.t.

$$Z_A = f_A(Y) \text{ a.s.}$$

i.e. $\forall A \in \mathcal{F}$, $\exists f_A$ meas, and $N_A \in \mathcal{F}$ ("bad set") s.t. $\begin{cases} \mathbb{P}[N_A] = 0; \\ \forall \omega \notin N_A, \\ Z_A(\omega) = f_A(Y(\omega)) \end{cases}$

Reasonable proposal: $\mathbb{P}\{A \mid Y = y\} \overset{(?)}{:=} f_A(y)$

$\quad = \mathbb{P}_y[A]$ (notation)

Does this give "coherent" family $(\mathbb{P}_y)$ of prob. meas, compatible w/ Ex 1-2?

Are $\mathbb{P}_y$ even necessarily prob. meas.?

E.g. are they additive? Fix $A, B \in \mathcal{F}$, $A \cap B = \emptyset$.

Expect: $\mathbb{P}_y[A \cup B] \overset{(?)}{=} \mathbb{P}_y[A] + \mathbb{P}_y[B]$.

Suppose $\omega \notin N_A \cup N_B \cup N_{A\cup B} \cup N_{A,B}^{(add)}$, $y = Y(\omega)$

$$\underset{\in \mathbb{R}}{\mathbb{P}_y[A]} + \underset{\in \mathbb{R}}{\mathbb{P}_y[B]} = \mathbb{P}[A \mid Y](\omega) + \mathbb{P}[B \mid Y](\omega)$$

(r.v. func: $\Omega \to \mathbb{R}$)

$$= \mathbb{E}[\mathbb{1}_A \mid Y](\omega) + \mathbb{E}[\mathbb{1}_B \mid Y](\omega)$$

$$\overset{\circledast}{=} \mathbb{E}[\mathbb{1}_A + \mathbb{1}_B \mid Y](\omega) \quad \leftarrow \text{for } \omega \notin N_{A,B}^{(add)}, \ \mathbb{P}[N_{A,B}^{(add)}] = 0$$

$$= \mathbb{E}[\mathbb{1}_{A\cup B} \mid Y](\omega)$$

$$= \mathbb{P}[A \cup B \mid Y](\omega)$$

$$= \mathbb{P}_y[A \cup B]$$

Upshot: we have $\mathbb{P}_y[A] + \mathbb{P}_y[B] = \mathbb{P}_y[A \cup B]$ provided that $y = Y(\omega)$ for $\omega \notin \tilde{N}_{A,B} := N_A \cup N_B \cup N_{A\cup B} \cup N_{A,B}^{(add)}$.

$\Updownarrow \qquad \qquad \mathbb{P}[\tilde{N}_{A,B}] = 0.$

$y \in Y(\tilde{N}_{A,B}^c) =: \tilde{F}_{A,B} \subseteq \mathbb{R}$

$Q := \mathrm{Law}(Y)$ prob. meas. on $\mathbb{R}$, then $Q[\tilde{F}_{A,B}] = 1$.

Will get $\mathbb{P}_y$ finitely additive (i.e. $\mathbb{P}_y[A] + \mathbb{P}_y[B] = \mathbb{P}_y[A \cup B]$ $\forall A, B \in \mathcal{F},\ A \cap B = \emptyset$) if

$$y \in \bigcap_{\substack{A,B \in \mathcal{F} \\ A \cap B = \emptyset}} \tilde{F}_{A,B} =: \tilde{F}^{(add)}$$

- If $\mathcal{F}$ countable (e.g. in Ex 1, all subsets of $\mathcal{X} \times \mathcal{Y}$)

  $Q[\tilde{F}_{A,B}] = 1 \ \forall A, B \implies Q[\tilde{F}^{(add)}] = 1$

  Similarly $\rightsquigarrow$ $\mathbb{P}_y$ is prob. meas. for Q-almost-all y.
- If $\mathcal{F}$ uncountable (e.g. in Ex 2) $\rightsquigarrow$ no control of $\tilde{F}^{(add)}$, could even have $Q[\tilde{F}^{(add)}] = 0$ or even $\tilde{F}^{(add)} = \emptyset$, i.e. no $\mathbb{P}_y$ is prob. meas.

Regular Conditional Prob / Dist: issues came from bad $(\Omega, \mathcal{F}, \mathbb{P})$.

With structural (topological) assumptions on $(\Omega, \mathcal{F})$, can repair construction.

Def: $(\Omega, \mathcal{F}, \mathbb{P})$ prob space, $(\mathcal{Y}, \mathcal{B})$ meas. space (previously $\mathbb{R}$, Borel σ-algebra), $Y : \Omega \to \mathcal{Y}$ meas. (i.e. $\mathcal{Y}$-valued r.v.), $Q := \text{Law}(Y)$ (prob. meas. on $(\mathcal{Y}, \mathcal{B})$).

We say $(\mathbb{P}_y)_{y \in \mathcal{Y}}$ is (strong) regular conditional prob. (SRCP) of $\mathbb{P}$ on Y if:

(1) $\mathbb{P}_y$ prob. meas on Ω for each $y \in \mathcal{Y}$ ("$\mathbb{P}_y[A] = \mathbb{P}[A \mid Y = y]$")

(2) $\mathbb{P}_y[Y \neq y] = 0$ for Q-almost-all $y \in Y$ ← if $Y(\omega)$ never $= y$, some such cond. necessary

and $\forall$ bounded meas $f : \Omega \to \mathbb{R}$ ($\mathbb{R}$-valued r.v.)

(3) $y \longmapsto \int f \, d\mathbb{P}_y$ measurable fn.

(✱) (4) $\int f \, d\mathbb{P} = \int \left(\int f \, d\mathbb{P}_y \right) dQ(y)$

For (3), (4): consider $f = \mathbb{1}_A$ for $A \in F$. Then

(3): $y \longmapsto \mathbb{P}_y[A]$ meas.

(4): $\mathbb{P}[A] = \underset{y \sim Q}{\mathbb{E}} \mathbb{P}_y[A] = \mathbb{E}\,\mathbb{P}_Y[A]$.

Intuitively: disintegration of $\mathbb{P}$ into $\mathbb{P}_y$ along Q.

(e.g. $\Omega = \mathbb{R}^n$)

Thm: ("Disintegration Thm") Suppose that:

(1) $(\Omega, \mathcal{F})$ is metric space (Polish space) Ω with its Borel σ-alg. $\mathcal{F}$. (topological conditions)

(2) $\mathbb{P}$ is countable sum of compactly supported fin. non-neg. measures

(3) $\mathrm{Graph}(Y) = \{(\omega, Y(\omega)) : \omega \in \Omega\}$ is $\mathcal{F} \otimes \mathcal{B}$ in $\Omega \times \mathcal{Y}$.

Then:

- Existence: an SRCP $(\mathbb{P}_y)_{y \in \mathcal{Y}}$ exists.
- Uniqueness: if $(\mathbb{P}_y)$, $(\mathbb{P}_y')$ are SRCP, then $\mathbb{P}_y = \mathbb{P}_y'$ for Q-almost all y.

Rk: There exist $(\Omega, F, \mathbb{P})$ and Y s.th. SRCP does not exist! Some extra structure required.

Exc: (1) $(\mathbb{P}_y)$ (on $\mathcal{X}$) with pmf's p_y from Ex 1 give SRCP.

(2) $(\mathbb{P}_y)$ (on $\mathbb{R}$) with pdf's p_y from Ex 2 give SRCP.

Warnings:

(1) $\mathbb{P}[A \mid B]$ for $\mathbb{P}[B] = 0$ not well-defined on its own, but rather with $B = \{Y = y\}$ for some r.v. Y.

(2) Even so, $\mathbb{P}\{A \mid Y = y\}$ not well-defined for specific values of y (say $y = 0$).

(Cf. "Q-almost-all" uniqueness.)

See notes: can extend to coherent theory of individual $\mathbb{P}_y$.

General idea: to ask $y \longmapsto \mathbb{P}_y$ to have "continuity".

$y \in \mathcal{Y}$, say $\mathbb{R}$ $\qquad$ $\mathbb{P}_y \in \{\text{measures on } \Omega\}$

say $\{\text{measures on } \mathbb{R}\}$.