Last time, we saw that if (X,Y) has a joint density ρ

I.e. $\mathbb{E}[f(X,Y)] = \iint f(x,y)\,\rho(x,y)\,dx\,dy \quad \forall f \in B(\mathbb{R}^2)$

then
$$\mathbb{E}[f(X,Y) \mid X = x] = \int f(x,y)\,\rho_{Y|X}(y|x)\,dy$$
$$\mathbb{E}[f(X,Y) \mid X] = g(X) \qquad g(x)$$

Here $\rho_{Y|X}(y|x) = \dfrac{\rho(x,y)}{\underbrace{\int \rho(x,y)\,dy}_{\rho_X(x)}} \mathbb{1}_{0<\rho_X(x)<\infty}$.

That is: we can define, for each x, a probability measure Q_x by
$$Q_x(B) = \int_B \rho_{Y|X}(y|x)\,dy$$

Then $\mathbb{E}[f(X,Y) \mid X = x] = \int f(x,y)\,Q_x(dy)$

I.e. $(x,B) \longmapsto Q_x(B)$ is the conditional distribution of (X,Y) given X.

This Q is an example of a probability kernel.

Probability Kernels

Given two measurable spaces

$$(S_1, \mathcal{B}_1), \quad (S_2, \mathcal{B}_2)$$

a probability kernel Q is a function $Q : S_1 \times \mathcal{B}_2 \to [0,1]$

s.t. 1. $Q(x, \cdot) : \mathcal{B}_2 \to [0,1]$ is a probability measure on $(S_2, \mathcal{B}_2)$ $\forall x \in S_1$

2. $Q(\cdot, B) : S_1 \to [0,1]$ is $\mathcal{B}_1 / \mathcal{B}(\mathbb{R})$ - measurable $\forall B \in \mathcal{B}_2$.

Notation: $\int_{S_2} f(y)\, Q(x, dy)$ $\qquad \mu(dx) \qquad d\mu(x)$

$$d\,Q\cancel{(x, y)}$$

E.g. If $\nu \in \mathrm{Prob}(S_2, \mathcal{B}_2)$, $Q(x, B) := \nu(B)$ $\forall x \in S_1$ is a probability kernel.

E.g. If (X, Y) have a joint density, $Q(x, B) = \int_B \rho_{Y|X}(y|x)\, dy$ is a probability kernel.

$$= \frac{\int_B \rho_{X,Y}(x,y)\, dy}{\int_{\mathbb{R}} \rho_{X,Y}(x,y)\, dy} \mathbb{1}_{\{\text{denom} \in (0,\infty)\}}(x)$$

Properties of Probability Kernels

Let Q be a probability kernel over $(S_1, \mathcal{B}_1), (S_2, \mathcal{B}_2)$.

Lemma: If $f : S_1 \times S_2 \to \mathbb{R}$ is $\mathcal{B}_1 \otimes \mathcal{B}_2$-measurable, and is either bounded or non-negative, then

$$x \longmapsto \int_{S_2} f(x,y)\, Q(x, dy) \quad \text{is measurable.} \quad (\star)$$

Pf. Let $\mathbb{H} = \{ f \in \mathbb{B}(S_1 \times S_2, \mathcal{B}_1 \otimes \mathcal{B}_2) : (\star) \text{ holds} \} \ni 1$

Let $\mathbb{M} \subseteq \mathbb{H}$, $\mathbb{M} = \{ \mathbb{1}_{B_1 \times B_2} : B_1 \in \mathcal{B}_1, B_2 \in \mathcal{B}_2 \}$ mult. system, contains $\mathbb{1}$

$$\int_{S_2} \underbrace{\mathbb{1}_{B_1 \times B_2}(x,y)}_{\mathbb{1}_{B_1}(x)\, \mathbb{1}_{B_2}(y)} Q(x, dy) = \mathbb{1}_{B_1}(x)\, Q(x, B_2) \quad \text{measurable } (\star) \checkmark$$

$\mathbb{H}$ is a subspace by linearity of $\int$.

If $f_n \to f$ and $\underline{|f_n| \leq M}$ $\forall n$, $\forall x$ $\int f_n(x,y)\, Q(x,dy) \xrightarrow{\text{DCT}} \int f(x,y)\, Q(x,dy)$

measurable in x

$\therefore$ b/c of robustness of measurability under limits.

$\mathbb{H}$ is closed under bdd convergence.

$\therefore$ by Dynkin $\mathbb{B}(\sigma(\mathbb{M})) \subseteq \mathbb{H}$.

$\sigma(\mathbb{M}) = \mathcal{B}_1 \otimes \mathcal{B}_2$.

Thus, (☆) holds for bounded f.

If $f \geq 0$, apply this to $f \mathbb{1}_{|f| \leq n}$. Then $f \mathbb{1}_{|f| \leq n} \uparrow$

$$\underbrace{\int (f \mathbb{1}_{|f| \leq n})(x,y)\, Q(x, dy)}_{\uparrow x \text{ meas.}} \xrightarrow{MCT} \underbrace{\int f(x,y)\, Q(x,dy)}_{\uparrow x \ \therefore \text{ meas.}}$$

///

This allows us to take the "product" $\mu \otimes Q$, for $\mu \in \mathrm{Prob}(S_1, \mathcal{B}_1)$, to form a probability measure $\mu \otimes Q \in \mathrm{Prob}(S_1 \times S_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$ as follows:

For $C \in \mathcal{B}_1 \otimes \mathcal{B}_2$, $(\mu \otimes Q)(C) := \int_{S_1} \mu(dx) \int_{S_2} \mathbb{1}_C(x,y)\, Q(x, dy)$

If $Q(x, dy) = \nu(dy)$

$\mu \otimes Q = \mu \otimes \nu$.

$$(\mu \otimes Q)(S_1 \times S_2) = \int_{S_1} \mu(dx) \underbrace{\int_{S_2} 1\, Q(x, dy)}_{1} = 1.$$

$\mu \otimes Q \in \mathrm{Prob}(S_1 \times S_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$.

Lemma: If $f : S_1 \times S_2 \to \mathbb{R}$ is $\mathcal{B}_1 \otimes \mathcal{B}_2 / \mathcal{B}(\mathbb{R})$-measurable, and either bounded or ≥ 0, then for $\mu \in \mathrm{Prob}(S_1, \mathcal{B}_1)$

$$\int_{S_1 \times S_2} f \, d(\mu \otimes Q) = \int_{S_1} \mu(dx) \int_{S_2} f(x,y)\, Q(x, dy). \qquad (☺)$$

Pf. Basically same as the proof that if $\nu(B) = \int_B \rho \, d\mu$ then $\int g \, d\nu = \int g\rho \, d\mu$.
First verify for simple f, then extend to bounded / non-negative f with limit arguments (using DCT / MCT). ///

Cor: If $f \in L^1(\mu \otimes Q)$, then $\underbrace{\int_{S_2} |f(x,y)| \, Q(x, dy) < \infty}_{\text{If } = \infty, \text{ define } \int_{S_2} f(x,y) Q(x,y) := 0.}$ for $[\mu]$-a.e. $x \in S_1$.

Then (☺) still holds.

Pf. $f \in L^1(\mu \otimes Q) \Rightarrow \infty > \int_{S_1 \times S_2} |f| \, d(\mu \otimes Q) = \int_{S_1} \mu(dx) \int_{S_2} |f(x,y)| \, Q(x, dy)$

$\therefore \mu \{ \underbrace{x : \int_{S_2} |f(x,y)| \, Q(x, dy) = \infty}_{(S_1')^c \subset S_1} \} = 0.$

$\int f(x,y) \, Q(x, dy)$
$= \int f^+(x,y) Q(x,y) - \int f^-(x,y) Q(x,y)$

For $x \in S_1'$ $(\mu(S_1)=1)$ I.e. $\int_{S_2} |f(x,y)|\, Q(x,dy) < \infty$,

$$\int_{S_2} f(x,y)\, Q(x,dy) = \int_{S_2} f^+(x,y)\, Q(x,dy) - \int_{S_2} f^-(x,y)\, Q(x,dy)$$

$$\therefore \int_{S_1} \mu(dx) \int_{S_2} f(x,y)\, Q(x,dy) = \int_{S_1} \mu(dx) \int_{S_2} f^+(x,y)\, Q(x,dy) - \int_{S_1} \mu(dx) \int_{S_2} f^+(x,y)\, Q(x,dy)$$

$$= \int f^+ \, d(\mu \otimes Q) - \int f^- \, d(\mu \otimes Q)$$

$$= \int f \, d(\mu \otimes Q).$$

///

Lemma: Let $Q, \tilde{Q}$ be probability kernels on $(S_1, \mathcal{B}_1)$, $(S_2, \mathcal{B}_2)$, and let $\mu \in \mathrm{Prob}(S_1, \mathcal{B}_1)$. If $Q(x,\cdot) = \tilde{Q}(x,\cdot)$ for $[\mu]$-a.e. $x \in S_1$, then $\mu \otimes Q = \mu \otimes \tilde{Q}$. Conversely, if $\mathcal{B}_2$ is countably generated, then if $\mu \otimes Q = \mu \otimes \tilde{Q}$, it follows that $Q(x,\cdot) = \tilde{Q}(x,\cdot)$ for $[\mu]$-a.e. $x \in S_1$.

Pf. ($\Longrightarrow$) It suffices to check that (by Dynkin)

$$\mu\otimes Q(B_1\times B_2)=\mu\otimes\tilde{Q}(B_1\times B_2)\quad\forall B_1\in\mathcal{B}_1,\ B_2\in\mathcal{B}_2$$

$$\mu\otimes Q(B_1\times B_2)=\int_{S_1}\mu(dx)\int_{S_2}\underbrace{\mathbb{1}_{B_1\times B_2}(x,y)}_{\mathbb{1}_{B_1}(x)\,\mathbb{1}_{B_2}(y)}Q(x,dy)$$

$$=\int_{B_1}\mu(dx)\,Q(x,B_2)=\int_{B_1}\mu(dx)\,\hat{Q}(x,B_2)$$

($\Longleftarrow$) If $\mu\otimes Q=\mu\otimes\tilde{Q}$, the above computation shows that

$$\int_{B_1}\mu(dx)\,Q(x,B_2)=\int_{B_1}\mu(dx)\,\hat{Q}(x,B_2)\quad\forall B_1\in\mathcal{B}_1,\ B_2\in\mathcal{B}_2.$$

$$\therefore\int_{S_1}[Q(x,B_2)-\tilde{Q}(x,B_2)]\,\mathbb{1}_{B_1}(x)\,\mu(dx)=0\quad\forall B_1\in\mathcal{B}_1$$

by Dynkin

$$\therefore\int_{S_1}[Q(x,B_2)-\tilde{Q}(x,B_2)]\,h(x)\,\mu(dx)=0\quad\forall h\in B(\mathcal{B}_1)$$

$$h(x)=\operatorname{sgn}(Q(x,B_2)-\tilde{Q}(x,B_2))$$

$\therefore\ \forall B_2\in\mathcal{B}_2,\ \exists\, S_2(B_2)\subseteq S_2,\ \mu(S_2(B_2))=1$, s.t. $Q(x,B_2)=\tilde{Q}(x,B_2)\ \forall x\in S_2(B_2)$.