In [Lectures 33.1, 33.2, 34.1] we introduced probability kernels $Q$, regular conditional distributions, and associated integral operators $L$.

We're going to change some notation and terminology for these now (sorry!) to avoid some later confusion.

Summary: $Q \to q$ probability kernel

$L \to Q$ Markov (transition) ~~generator~~ operator.

A probability kernel over $(S_1, \mathcal{B}_1), (S_2, \mathcal{B}_2)$ is a function

$$q : S_1 \times \mathcal{B}_2 \to [0,1]$$

s.t. $\quad q(x, \cdot) \in \mathrm{Prob}(S_2, \mathcal{B}_2) \quad \forall x \in S_1$

& $\quad q(\cdot, B)$ is $\mathcal{B}/\mathcal{B}(\mathbb{R})$-measurable $\forall B \in \mathcal{B}_2$

E.g. $\quad q(x, B) = \int_B q(x,y)\, \nu(dy) \quad \nu \in \mathrm{Prob}(S_2, \mathcal{B}_2)$

For $f : (S_1 \times S_2, \mathcal{B}_1 \otimes \mathcal{B}_2) \to (\mathbb{C}, \mathcal{B}(\mathbb{C}))$ bounded or non-negative,

$$x \mapsto \int_{S_2} f(x,y)\, q(x, dy) \quad \text{is } \mathcal{B}_1\text{-measurable.}$$

Given random variables $X:(\Omega,\mathcal{F})\to(S_1,\mathcal{B}_1)$, $Y:(\Omega,\mathcal{F})\to(S_2,\mathcal{B}_2)$

a regular conditional distribution of $Y|X$ is a probability kernel $q=q_{Y|X}$

s.t.

$$\mathbb{P}(Y\in B\mid X) = q(X,B) \quad \forall B\in\mathcal{B}_2$$

$$\mathbb{E}[f(X,Y)\mid X] = \mathbb{E}[f(x,Y)]\big|_{x=X}$$

Always exists if $(S_j,\mathcal{B}_j)$ are regular Borel spaces.

$\Rightarrow$ If $f\in B(S_1\times S_2, \mathcal{B}_1\otimes\mathcal{B}_2)$,

$$\mathbb{E}[f(X,Y)\mid X] = \int_{S_2} f(x,y)\,q(x,dy)\Big|_{x=X}.$$

Special cases:

- If $X,Y$ are independent, $q_{Y|X}(x,\cdot)=\mu_Y$

  I.e. $P(Y\in B\mid X) = \mu_Y(B) = P(Y\in B)$

$$\rho_{Y|X}(y|x) = \frac{\rho_{X,Y}(x,y)}{\rho_X(x)}$$

$$\int_{S_2}\rho_{X,Y}(x,y)\,\nu_2(dy)$$

- If $(S_1,\mathcal{B}_1,\nu_1)$, $(S_2,\mathcal{B}_2,\nu_2)$ are $\sigma$-finite measure spaces, and $(X,Y)$ have a joint density $\rho_{X,Y}$ wrt $\nu_1\otimes\nu_2$, then

$$q_{Y|X}(x,B) = \int_B \rho_{Y|X}(y|x)\,\nu_2(dy)$$

- If $S_1,S_2$ are discrete, $q(x,B)=\sum_{y\in B} P(Y=y\mid X=x)$.

Given a probability kernel $q : S_1 \times \mathcal{B}_2 \to [0,1]$, there is an associated operator

$$Q : \mathbb{B}(S_2, \mathcal{B}_2) \to \mathbb{B}(S_1, \mathcal{B}_1)$$

formerly $Lf$ $\rightarrow$ $(Qf)(x) = \int_{S_2} f(y)\, q(x, dy)$ "Markov transition operator"

$$\|Qf\|_\infty \leq \|f\|_\infty = \sup|f|.$$

If $S_1 = S_2$, such transition operators are characterized by:

0. $Q : \mathbb{B}(S, \mathcal{B}) \to \mathbb{B}(S, \mathcal{B})$ is linear.
1. $Q(1) = 1$.
2. $Q(f) \geq 0$ if $f \geq 0$.
3. $f_n \to f$ boundedly $\Rightarrow$ $Qf_n \to Qf$ boundedly.

In this case, $q(x, B) = Q(\mathbb{1}_B)(x)$. $\int f(y)\, q(x, dy) = (Qf)(x)$.

Composing transition operators:

$$(Q_1 Q_2 f)(x) = \int f(z) \int q_1(x, dy)\, q_2(y, dz)$$

If $q$ is a propability kernel on $(S_1, \mathcal{B}_1), (S_2, \mathcal{B}_2)$ and $\nu \in \mathrm{Prob}(S_1, \mathcal{B}_1)$,
then $\nu \otimes q \in \mathrm{Prob}(S_1 \times S_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$ is defined by

$$(\nu \otimes q)(B) = \int_{S_1} \nu(dx) \int_{S_2} \mathbb{1}_B(x,y)\, q(x,dy) \qquad \int f \, d(\nu \otimes q) = \int_{S_1} \nu(dx) \int_{S_2} f(x,y)\, q(x,dy)$$

If $\mathcal{B}_1, \mathcal{B}_2$ are countably generated, and $\tilde{q}$ is another probability kernel,
then $\nu \otimes q = \nu \otimes \tilde{q} \iff q(x, \cdot) = \tilde{q}(x, \cdot)$ for $\nu$-a.e. $x \in S_1$.

More generally, if $q_1, \ldots, q_n : (S, \mathcal{B})^2 \to [0,1]$, $\nu \in \mathrm{Prob}(S, \mathcal{B})$,

$$\int_{S^{n+1}} f \, d\mu := \int_S \nu(dx_0) \int_S q_1(x_0, dx_1) \int_S q_2(x_1, dx_2) \cdots \int_S q_n(x_{n-1}, dx_n)\, f(x_0, \ldots, x_n)$$

defines a probability measure $\mu \in \mathrm{Prob}(S^{n+1}, \mathcal{B}^{\otimes n+1})$.

If $f = f_0 \otimes f_1 \otimes \cdots \otimes f_n$,

$$\int_{S^{n+1}} f \, d\mu = E_\nu [ f_0 Q_1 (f_1 Q_2 ( \cdots (f_{n-1} (Q_n f_n)) \cdots ))].$$