Lecture 19: Radon-Nikodym Derivatives

Recall that if μ, ν are two measures on $(\Omega, \mathcal{A})$ then $\frac{d\nu}{d\mu}$ was notation for the density of ν w.r.t μ when such a thing exists.

$$\underset{\mu,\nu}{\Omega}\ \mathcal{A} \xrightarrow{\frac{d\nu}{d\mu} \in \mathcal{M}(\Omega,\mathcal{A})} \overline{\mathbb{R}}\ \ \mathcal{B}(\overline{\mathbb{R}})$$

such that $\nu(A) = \int_A \frac{d\nu}{d\mu}\, d\mu \quad \forall A \in \mathcal{A}$.

We never had a general thm to show when $\frac{d\nu}{d\mu}$ exists. This will come from the Radon-Nikodym Thm.

This theorem is also related to the existance of conditional expected value. Here is the heuristic:

Let X and Y be two r.v.s on $(\Omega, \mathcal{A}, P)$. Suppose $X \in \mathcal{M}(\Omega, \mathcal{A})$.

In undergrad we learned

$$E(X) = E(E(X|Y))$$

Indeed for any $A \in \mathcal{A}$ we have

$$E(I_A X) = E(E(I_A X|Y))$$

So that

$$\int_A X\,dP = \int_\Omega E(I_A X|Y)\,dP \quad \forall A \in \mathcal{A}$$

Also notice that the result "characterizing ⓜ funs of ⓜ funs" from lecture 9 implies

$$\begin{aligned} A \in \sigma\langle Y\rangle &\iff I_A \text{ ⓜ } \sigma\langle Y\rangle \\ &\iff I_A(\omega) = g(Y(\omega)) \\ &\qquad \text{for some ⓜ } g \end{aligned}$$

$\therefore$ If $A \in \sigma\langle Y\rangle$ then I_A can be pulled out of $E(I_A X|Y)$ and we have

$$\int_A X\,dP = \int_A E(X|Y)\,dP \quad \forall A \in \sigma\langle Y\rangle$$

In otherwords $E(X|Y)$ appears to be the density of the measure $\int_\cdot X\,dP$ on $(\Omega, \sigma\langle Y\rangle)$ w.r.t. $P|_{\sigma\langle Y\rangle}$

$$\frac{d \int_\cdot X\,dP|_{\sigma\langle Y\rangle}}{d\, P|_{\sigma\langle Y\rangle}} = E(X|Y).$$

Definition: If ν & μ are measures on a measurable space $(\Omega, \mathcal{A})$ then

(i) $\nu \perp \mu$ iff $\exists A \in \mathcal{A}$ s.t.

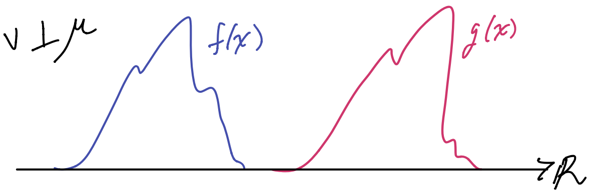
$$\nu(A^c) = 0 = \mu(A)$$

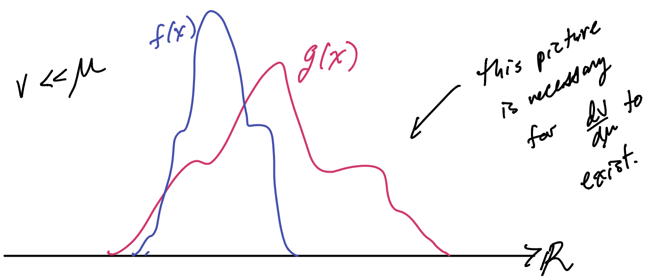
(ii) $\nu \ll \mu$ iff $\forall A \in \mathcal{A}$

$$\mu(A) = 0 \implies \nu(A) = 0.$$

Here is the pictures when

$$d\nu(x) = f(x)dx \quad \& \quad d\mu(x) = g(x)dx$$





Before we prove the Radon-Nikodym result lets recall the following result from Lecture 11:

"Probabilits' world view"

If μ is a non-trivial & σ-finite measure on $(\Omega, \mathcal{A})$ then $\exists$ a prob measure P on $(\Omega, \mathcal{A})$ s.t. $\frac{d\mu}{dP}$ exists with the addition property that $\frac{d\mu}{dP}$ takes values in $(0,\infty)$.

Theorem: (Radon-Nikodym)

If μ & ν are two measures on $(\Omega, \mathcal{A})$ s.t. $\nu \ll \mu$ and both are σ-finite then $\frac{d\nu}{d\mu} \in \mathcal{M}(\Omega, \mathcal{A})$ exists and is μ-unique.

Proof: If μ or $\nu \equiv 0$ the theorem is true so suppose both are non-trivial.

Since μ & σ are σ-finite the "probabilits world" implies $\exists$ probs P, Q on $(\Omega, \mathcal{A})$ s.t.

$$\frac{d\nu}{dQ} \ \& \ \frac{d\mu}{dP} \text{ exist \& take values in } (0,\infty).$$

which means exercize 3 in HWK 7 (from 235A) applies & gives

$$\therefore \quad \frac{dQ}{d\nu} = \frac{1}{d\nu/dQ} \quad \& \quad \frac{dP}{d\mu} = \frac{1}{d\mu/dP}$$

Now if $\frac{dQ}{dP}$ exists then we have

$$\frac{d\nu}{d\mu} = \frac{d\nu}{dQ}\frac{dQ}{dP}\frac{dP}{d\mu} \quad \mu\text{-a.e.}$$

by the "chain rule Thm" of Lecture 11. Therefor all we need to do is show $\frac{dQ}{dP}$ exists.

The main idea is to define

$$W = \frac{P+Q}{2}$$

and use Riesz to get $\frac{dQ}{dW}$ & $\frac{dP}{dW}$.

Then show

$$\frac{dQ}{dP} = \frac{dQ}{dW} \Big/ \frac{dP}{dW}.$$

(5)

(show dQ/dP exists):

For all $X \in L_2(\Omega, \mathcal{A}, W)$ define the following continuous linear functionals:

$$f_P(X) := \int_\Omega X\,dP = E_P(X) \overset{\text{Riesz}}{=} \langle Y_P, X \rangle_{L_2(W)}$$

$$f_Q(X) := \int_\Omega X\,dQ = E_Q(X) \overset{\text{Riesz}}{=} \langle Y_Q, X \rangle_{L_2(W)}$$

For some $Y_P, Y_Q \in L_2(\Omega, \mathcal{A}, W)$.

To see why f_P & f_Q are continuous linear functionals over $L_2(\Omega, \mathcal{A}, W)$ notice first that

$$\int_\Omega |X|^2 dP,\ \int_\Omega |X|^2 dQ \le 2 \int_\Omega |X|^2 d\underbrace{\left(\tfrac{P+Q}{2}\right)}_{W} \quad (*)$$

So $X \in L_2(W) \Rightarrow X \in L_2(P) \cap L_2(Q)$

$\therefore$ f_P & f_Q are defined over $L_2(W)$ & clearly linear.

For continuity notice that

$$X_n \xrightarrow{L_2(W)} X \implies \begin{array}{l} X_n \xrightarrow{L_2(P)} X \\ X_n \xrightarrow{L_2(Q)} X \end{array} \quad \text{by } (*)$$

$$\implies \begin{array}{l} E_P(X_n) \to E(X) \\ E_Q(X_n) \to E(X) \end{array} \quad \text{by } L_p \text{ Convergence Thm}$$

$$\implies \begin{array}{l} f_P(X_n) \to f(X) \\ f_Q(X_n) \to f(X) \end{array}$$

$\therefore$ Indeed, f_P & f_Q are continuous linear functionals over $L_2(W)$.

Now plug in I_A for X $(A \in \mathcal{A})$ to get

$$f_P(I_A) = P(A) = \langle Y_P, I_A \rangle_{L_2(W)} = \int_A Y_P\,dW$$

$$f_Q(I_A) = Q(A) = \langle Y_Q, I_A \rangle_{L_2(W)} = \int_A Y_Q\,dW$$

$\therefore$ $Y_P = \dfrac{dP}{dW}$ & $Y_Q = \dfrac{dQ}{dW}$

modified on W-null sets so they are in $\mathcal{M}(\Omega, \mathcal{A})$. Possible since $\int_\cdot 0\,dW \le \int_\cdot Y_P\,dW \Leftrightarrow 0 \le Y_P$ W-a.e. (by Thm in Lecture 11 which requires 0 or $Y_P \in L_1$ or W σ-finite)

(6)

Now we simply check that

$$\frac{dQ}{dP} := \begin{cases} Y_Q/Y_P & \text{on } \{Y_P \neq 0\} \\ 0 & \text{on } \{Y_P = 0\} \end{cases}$$

serves as a density of Q w.r.t. P.

Indeed let $N = \{Y_P = 0\}$ & $A \in \mathcal{A}$ so that

$$\underbrace{\int_A \frac{dQ}{dP}\,dP}_{\text{defined since } \frac{dQ}{dP} \in \mathcal{M}(\Omega, \mathcal{A})} = \int_{A \cap N^c} (Y_Q/Y_P)\,dP$$

$$= \int_{A \cap N^c} (Y_Q/Y_P)\, Y_P\,dW \quad \text{by slap-in-the density}$$

$$= \int_{A \cap N^c} Y_Q\,dW$$

$$= Q(A \cap N^c) \quad \text{since } Y_Q = \frac{dQ}{dW}$$

$$= Q(A)$$

since $P(A \cap N) \le P(N) = P(Y_P = 0) = \int_{\{Y_P = 0\}} Y_P\,dW = 0$

and $P \gg \mu$ since $\frac{dP}{d\mu}$ exists

$\gg \nu$ by assumption

$\gg Q$ since $\frac{dQ}{d\nu}$ exists

so that $Q(A \cap N) = 0$

QED

To recap the proof we showed

$$\frac{d\nu}{d\mu} = \frac{d\nu}{dQ}\frac{dQ}{dP}\frac{dP}{d\mu}$$

where P & Q are from "probabilists world view" which requires μ & ν σ-finite

$$= \frac{d\nu}{dQ}\,\frac{dQ/dW}{dP/dW}\,\frac{dP}{d\mu} \quad \text{for } W = \frac{P+Q}{2}$$

found by Riesz on $L_2(W)$ for $f_P(X) = E_P(X)$ & $f_Q(X) = E_Q(X)$

The following example suggests we can possibly extend the Radon-Nikodym result to the assumption μ is σ-finite rather than both μ & ν are σ-finite.

Example:

$\Omega = \mathbb{R}$

$\mathcal{A} = \mathcal{B}(\mathbb{R})$

$\mu = \mathcal{L}' :=$ Lebesgue measure

$$\nu = \infty \cdot \mathcal{L}' = \begin{cases} 0 & \text{when } \mathcal{L}'(A) = 0 \\ \infty & \text{o.w.} \end{cases}$$

$\therefore$ $\nu \ll \mu$ and μ is σ-finite

but ν is not σ-finite.

Yet $\nu(A) = \int_A \infty \, d\mu$ so $\frac{d\nu}{d\mu}$ exists.

Theorem: (improved Radon-Nikodym)

If μ & ν are two measures on $(\Omega, \mathcal{A})$ s.t. $\nu \ll \mu$ and μ is σ-finite then

$\frac{d\nu}{d\mu} \in \mathcal{N}(\Omega, \mathcal{A})$ exists and is μ-unique.

Proof:

The problem here is we cannot use the "probabilist world view" to get the existance of $\mathbb{Q}$. The plan is to find $\frac{d\nu}{dP}$ s.t.

$$\frac{d\nu}{d\mu} = \frac{d\nu}{dP}\frac{dP}{d\mu}$$

where the existance of $\frac{d\nu}{dP}$ will come from the fact that P is a finite measure.

($d\nu/dP$ exists):

Note that for any $F \in \mathcal{A}$ we can write $\nu(\cdot) = \nu(\cdot \cap F) + \nu(\cdot \cap F^c)$

we will want to find F s.t.

(i) $\nu(\cdot \cap F)$ is σ-finite

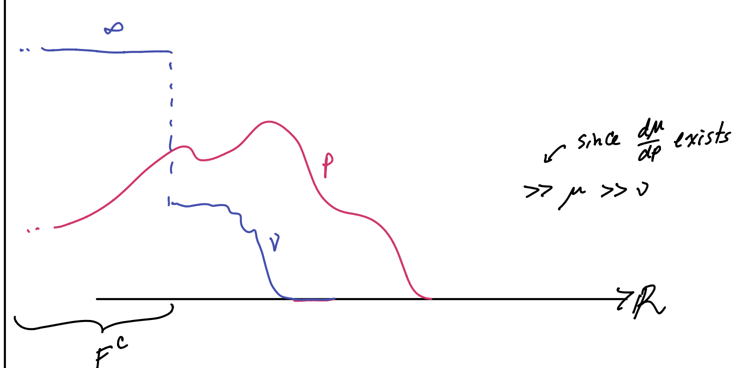
so old RD Thm applies to give $\frac{d\,\nu(\cdot \cap F)}{dP}$

(ii) $\nu(\cdot \cap F^c) = \infty P(\cdot \cap F^c)$

where this "bad piece" is covered by the last example so that $\frac{d\,\nu(\cdot \cap F^c)}{dP} = \infty I_{F^c}$

since $\int_A \infty I_{F^c} \, dP = \infty P(A \cap F^c) = \nu(A \cap F^c)$.

Here is the picture



Lets find F as the "P-biggest set s.t. ν is σ-finite over F."

Set

$$\mathcal{F} := \left\{ \bigcup_{k=1}^{\infty} A_k : \nu(A_k) < \infty, A_k \in \mathcal{A}, \forall k \right\}$$

and notice that $\mathcal{F}$ is closed under countable union.

Let $m = \sup\{P(F) : F \in \mathcal{F}\}$ and choose $F = \bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$ to attain the above sup.

9

The existance of such an F holds since

$$\left\{\begin{array}{l} F_n \in \mathcal{F} \text{ s.t. } P(F_n) \to m \text{ implies} \\ m \xleftarrow{n\to\infty} P(F_n) \le P\Big(\underbrace{\bigcup_{n=1}^\infty F_n}_{\in \mathcal{F}}\Big) \le m \\ \therefore \text{ sup is attained at } \bigcup_{n=1}^\infty F_n. \end{array}\right.$$

Now we just check that

(i) $\nu(\cdot \cap F)$ is σ-finite

(ii) $\nu(\cdot \cap F^c) = \infty P(\cdot \cap F^c)$.

For (i) notice that since $F \in \mathcal{F}$ we have

$$F = \bigcup_{k=1}^\infty A_k \quad \text{for } \nu(A_k) < \infty \ \forall k \text{ and}$$

$\therefore$ $\nu(F^c \cap F), \nu(A_1 \cap F), \nu(A_2 \cap F), \dots$

are all finite & $F^c \cup A_1 \cup A_2 \cup \dots = \Omega$

$\therefore$ $\nu(\cdot \cap F)$ is σ-finite

For (ii) notice that $\forall A \in \mathcal{A}$

$$P(A \cap F^c) = 0 \implies \nu(A \cap F^c) = 0$$

by $P \gg \mu \gg \nu$. Also

$$P(A \cap F^c) > 0 \implies \nu(A \cap F^c) = 0$$

For suppose not.

$\therefore$ $\exists A \in \mathcal{A}$ s.t. $P(A \cap F^c) > 0$ (a)

$\nu(A \cap F^c) < \infty$ (b)

$\therefore$ $A \cap F^c \in \mathcal{F}$ by (b)

$\therefore$ $F \cup (A \cap F^c)$ since $F \in \mathcal{F}$ & $\mathcal{F}$ is closed under countable union

$\therefore$ $m = P(F) \overset{(a)}{<} P(F) + P(A \cap F^c)$

$= P(F \cup (A \cap F^c)) \le m$

$\therefore$ contradiction

QED

10

Remark: the strict inequality $\overset{(a)}{<}$ above is where I needed P to be a finite measure.

11

12