

Lecture 19:

(1)

Radon-Nikodym Derivatives and Lebesgue decomposition

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

$$\begin{array}{ccc} \nu & \xrightarrow{\frac{d\nu}{d\mu} \in \mathcal{N}(\Omega, \mathcal{Q})} & \bar{\mu} \\ \mu, \nu \in \mathcal{Q} & & B(\bar{\mu}) \end{array}$$

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

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 existence of conditional expected value. Here is the heuristic:

Let X and Y be two r.v.s on (Ω, \mathcal{Q}, P) .

Suppose $X \in \mathcal{N}(\Omega, \mathcal{Q})$.

In undergrad we learned

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

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

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

So that

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

Also notice that the result "characterizing \mathcal{C} functions" from lecture 9 implies

$$\begin{aligned} A \in \sigma(Y) &\iff I_A \in \sigma(Y) \\ &\iff I_A(w) = g(Y(w)) \\ &\text{for some } g \in \mathcal{G} \end{aligned}$$

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

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

In other words $E(X|Y)$ appears to be the density of the measure $\int_X dP$ on $(\Omega, \sigma(Y))$ w.r.t. $P|_{\sigma(Y)}$

$$\frac{d \int_X dP|_{\sigma(Y)}}{d P|_{\sigma(Y)}} = E(X|Y).$$

Definition: if $\nu \ll \mu$ are measures on a measurable space (Ω, \mathcal{Q}) then

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

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

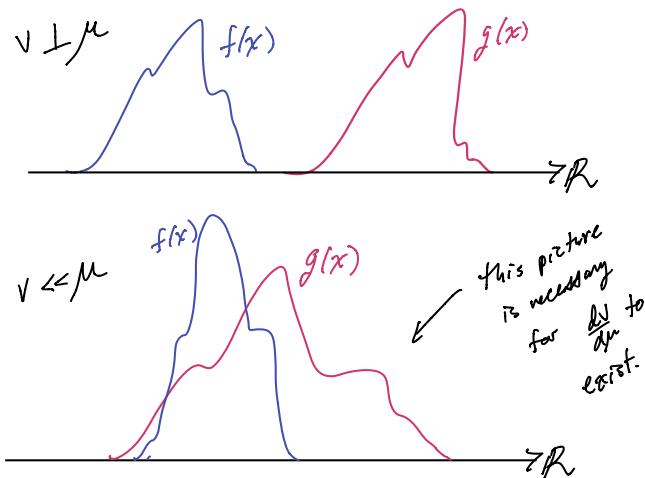
(ii) $\nu \ll \mu$ iff $\forall A \in \mathcal{Q}$

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

Here is the pictures when

(3)

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



Before we prove the Radon-Nikodym result let's recall the following result from Lecture 11:

"Probabilists' world view"

If μ is a non-trivial & σ -finite measure on (Ω, \mathcal{A}) then \exists a prob measure P on (Ω, \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 (Ω, \mathcal{A}) s.t. $\nu \ll \mu$ and both are σ -finite then

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

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

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

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

which means exercise 3 in Hulk 7 (from 235A) applies & gives

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

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.

Therefore 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

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

(show dQ/dP exists):

(5)

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

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

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

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

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

$$\int_{\Omega} |X|^2 dP, \int_{\Omega} |X|^2 dQ \leq \int_{\Omega} |X|^2 d\left(\frac{P+Q}{2}\right) (*)$$

$$\text{So } X \in L_2(\nu) \Rightarrow X \in L_2(P) \cap L_2(Q)$$

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

For continuity notice that

$$\begin{aligned} X_n \xrightarrow{L_2(\nu)} X &\implies X_n \xrightarrow{L_2(P)} X \quad \text{by } (*) \\ &\quad X_n \xrightarrow{L_2(Q)} X \\ &\implies E_P(X_n) \rightarrow E(X) \quad \text{by } L_p \\ &\quad E_Q(X_n) \rightarrow E(X) \quad \text{convergence} \\ &\implies f_P(X_n) \rightarrow f(X) \\ &\quad f_Q(X_n) \rightarrow f(X) \end{aligned}$$

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

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

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

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

$$\therefore Y_P = \frac{dP}{d\nu} \quad \text{and} \quad Y_Q = \frac{dQ}{d\nu}$$

Modifed on ν -null sets so they

are in $\eta(\Omega, \mathcal{A})$. Possible since

$\int \nu d\nu \leq \int Y_P d\nu \iff 0 \leq Y_P \text{ } \nu\text{-a.e. (by Thm in Lecture 11 which requires } 0 \text{ or } Y_P \in L_1 \text{ or } \nu \text{ } \sigma\text{-finite)}$

Now define

$$\frac{dQ}{dP} := \frac{dQ/d\nu}{dP/d\nu} I_{\{dP/d\nu \neq 0\}}$$

and simply check that it serves as the density of Q w.r.t P .

Indeed, let $A \in \mathcal{Q}$ and notice

$$\begin{aligned} \int_A \frac{dQ}{dP} dP &= \underbrace{\int_A \frac{dQ/d\nu}{dP/d\nu} I_{\{\frac{dP}{d\nu} \neq 0\}} dP}_{\text{defined since } \frac{dQ}{dP} \in \eta(\Omega, \mathcal{A})} \\ &\stackrel{\text{"step"} \atop \frac{dQ}{dP} \in \eta(\Omega, \mathcal{A})}{=} \int_A \frac{dQ/d\nu}{dP/d\nu} I_{\{\frac{dP}{d\nu} \neq 0\}} \frac{dP}{d\nu} d\nu \\ &= \int_{A \cap \{\frac{dP}{d\nu} \neq 0\}} dQ/d\nu d\nu \\ &\stackrel{\text{"step"} \atop = Q(A \cap \{\frac{dP}{d\nu} \neq 0\})}{=} Q(A \cap \{\frac{dP}{d\nu} \neq 0\}) \\ &= Q(A) \end{aligned}$$

↗

Since $P(A \cap \{\frac{dP}{d\nu} = 0\}) \leq P(\frac{dP}{d\nu} = 0) = \int_{\Omega} \frac{dP}{d\nu} d\nu = 0$

and $P \gg \nu$ since $\frac{dP}{d\nu}$ exists
 $\gg \nu$ by assumption
 $\gg Q$ since $\frac{dQ}{d\nu}$ exists

implies $Q(A \cap \{\frac{dP}{d\nu} = 0\}) = 0$

QED

To recap the proof we showed

$$\frac{d\nu}{d\mu} = \frac{d\nu}{dQ} \frac{dQ}{dP} \frac{dP}{d\mu} \quad \text{where } P \text{ & } Q \text{ are from "probabilists world view" which requires } \mu \text{ & } \nu \text{ } \sigma\text{-finite}$$

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

↗ found by Riesz in $L_2(\nu)$
 for $f_P(X) = E_P(X) \wedge f_Q(X) = E_Q(X)$

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

(7)

Example:

$$\Omega = \mathbb{R}$$

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

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

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

$\therefore v \ll \mu$ and μ is σ -finite
but v is not σ -finite.

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

Theorem: (improved Radon-Nikodym)

If μ & v are two measures on (Ω, \mathcal{Q})
s.t. $v \ll \mu$ and μ is σ -finite then

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

Proof:

The problem here is we cannot use the "probabilist world view" to get the existence of $\frac{dv}{d\mu}$. The plan is to find $\frac{dv}{dP}$ s.t.

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

↑ where the existence of $\frac{dv}{dP}$ will come from the fact that P is a finite measure.

($\frac{dV}{dP}$ exists):

(8)

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

We will want to find F s.t.

(i) $v(\cdot \cap F)$ is σ -finite

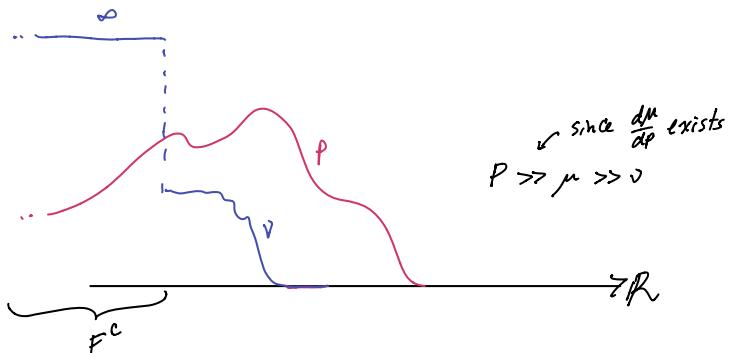
so old RD Thm applies to give $\frac{dV(\cdot \cap F)}{dP}$

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

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

Since $\int_A \infty I_{F^c} dP = \infty P(A \cap F^c) = v(A \cap F^c)$.

Here is the picture



Let's find F as the " P -biggest set s.t. v is σ -finite over F ".

Set

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

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.

The existence of such an F
holds since

$$\left\{ \begin{array}{l} F_n \in \mathcal{F} \text{ s.t. } P(F_n) \rightarrow m \text{ implies} \\ m \stackrel{n \rightarrow \infty}{\longleftarrow} P(F_n) \leq P\left(\bigcup_{n=1}^{\infty} F_n\right) \leq m \\ \therefore \text{sup is attained at } \bigcup_{n=1}^{\infty} F_n. \end{array} \right.$$

Now we just check that

- (i) $v(\cdot \cap F)$ is σ -finite
- (ii) $v(\cdot \cap F^c) = \infty P(\cdot \cap F^c)$.

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

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

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

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

$$\therefore v(\cdot \cap F) \text{ is } \sigma\text{-finite}$$

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

$$P(A \cap F^c) = 0 \Rightarrow v(A \cap F^c) = 0$$

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

$$P(A \cap F^c) > 0 \Rightarrow v(A \cap F^c) = \infty$$

For suppose not.

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

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

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

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

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

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

\therefore contradiction

QED

Remark: The strict inequality $\stackrel{(a)}{<}$

above is where we needed that P be a finite measure.

Properties of Radon-Nikodym derivatives

In this section it will be convenient to use the following (totally not standard) notation

$v \ll \mu$ means $v \ll \mu$ & μ is σ -finite

Theorem: (RND props)

Let $v, \mu, \nu, v_1, v_2, \dots$ be measures on a measurable space (Ω, \mathcal{Q}) .

(1) If $v_1, v_2 \ll \mu$ & $c_1, c_2 \geq 0$ then

$$(c_1 v_1 + c_2 v_2) \ll \mu \text{ and}$$

$$\frac{d(c_1 v_1 + c_2 v_2)}{d\mu} = c_1 \frac{dv_1}{d\mu} + c_2 \frac{dv_2}{d\mu} \quad \mu\text{-a.e.}$$

(2) If $v_1, v_2 \ll \mu$ then

$$v_1 \leq v_2 \text{ on } \Omega \text{ iff } \frac{dv_1}{d\mu} \leq \frac{dv_2}{d\mu} \quad \mu\text{-a.e.}$$

(3) If $v_n \ll \mu$ & $v_n(A) \uparrow$ & $A \in \mathcal{Q}$ then

$$v(\cdot) := \lim_n v_n(\cdot) \ll \mu \text{ and}$$

$$\frac{dv_n}{d\mu} \xrightarrow{\mu\text{-a.e.}} \frac{dv}{d\mu}$$

(4) If $v \ll \mu$ then

$$v \text{ is finite} \iff \frac{dv}{d\mu} \in L_1(\Omega, \mathcal{A}, \mu) \text{ and}$$

$$v \text{ is } \sigma\text{-finite} \iff \frac{dv}{d\mu} < \infty \text{ } \mu\text{-a.e.}$$

(5) If $v \ll \sigma \ll \mu$ then $v \ll \mu$ and

$$\frac{dv}{d\mu} = \frac{dv}{d\sigma} \frac{d\sigma}{d\mu} \quad \mu\text{-a.e.}$$

$$\text{and } \frac{dv}{d\sigma} = \frac{dv/d\mu}{d\sigma/d\mu} \mathbb{I}_{\left\{\frac{d\sigma}{d\mu} > 0\right\}} \text{ } \sigma\text{-a.e.}$$

(6) If both $\mu \ll v$ & $v \ll \mu$ then

$$\frac{dv}{d\mu} > 0 \text{ } \mu\text{-a.e. and}$$

$$\frac{d\mu}{dv} = \frac{1}{dv/d\mu} \quad v\text{-a.e. \&} \quad \mu\text{-a.e.}$$

Proof:

For (1): Just check RHS integrates correctly.

For (2): Since μ is σ -finite our results on indefinite integrals applies & can be re-stated to say $\int \frac{dV_1}{d\mu} d\mu = \int \frac{dV_2}{d\mu} d\mu + \forall A \in \mathcal{A}$

\Downarrow Lecture 11

$$\frac{dV_1}{d\mu} \leq \frac{dV_2}{d\mu} \quad \mu\text{-a.e.}$$

This proves (2).

(11)

For (3): First note that $V(A)$ is defined by monotonicity & is a measure since clearly $V(\emptyset) = 0$ and

$$\begin{aligned} V\left(\bigcup_k \underbrace{A_k}_{\text{disjoint}}\right) &= \lim_n V_n\left(\bigcup_k A_k\right) \\ &= \lim_n \sum_k V_n(A_k) \\ &= \sum_k \lim_n V_n(A_k) \quad \text{by Monotone} \\ &\quad \text{Convergence Thm (MCT)} \\ &= \sum_k V(A_k). \end{aligned}$$

Clearly $v \ll \mu$ and by (2)

$$0 \leq \frac{dV_n}{d\mu} \leq \frac{dV_{n+1}}{d\mu} \quad \mu\text{-a.e.}$$

$$\begin{aligned} \therefore V(A) &:= \lim_n V_n(A) \\ &= \lim_n \int_A \frac{dV_n}{d\mu} d\mu \\ &= \int_A \lim_n \underbrace{\frac{dV_n}{d\mu}}_{\text{exists by monotonicity}} d\mu \quad \text{by MCT} \end{aligned}$$

For (4)-(6): These follow from our old results on densities, HWK & from Stat 235A, and similar arguments to the RND Thm.

QED

(12)

Lebesgue Decomposition

(13)

In some sense the Lebesgue decomposition result is a tweak to the RND Thm that allows us to say something when $P \not\ll Q$. We need it for studying likelihood ratios with martingales.

Theorem: (Lebesgue Decomposition)

Let P and Q be two probability measures on (Ω, \mathcal{A}) . Then there exists two measures $Q_{\ll P}$, Q_{\perp} on (Ω, \mathcal{A}) st.

$$Q = Q_{\ll P} + Q_{\perp} \quad (*)$$

where this is the P -largest measure $\leq Q$ that is $\ll P$

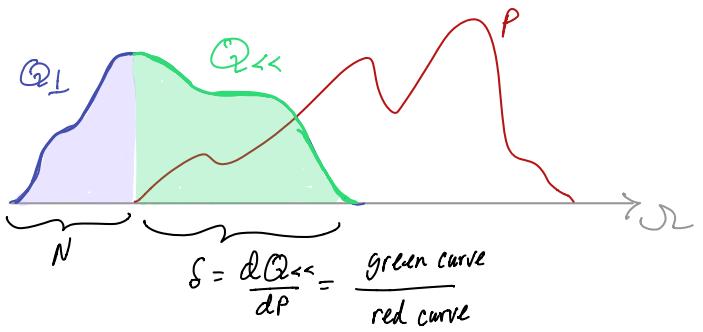
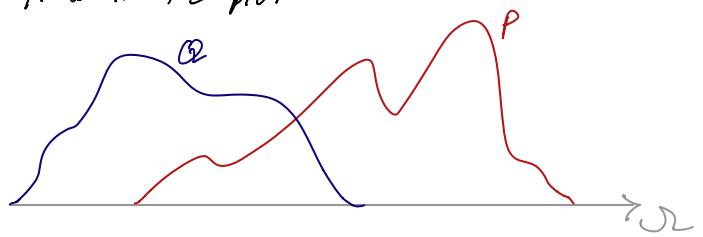
where $Q_{\perp}(\cdot) = Q(\{-n\})$
for $P(N) = 0$.
 $\therefore Q_{\perp} \perp P$

Remark: $Q_{\ll P}$ is the P -largest measure $\leq Q$ that is $\ll P$ means that for any other \tilde{Q} measure which satisfies $\tilde{Q}(\cdot) \leq Q(\cdot)$ & $\tilde{Q} \ll P$ then it must be the case that

$$\frac{d\tilde{Q}}{dP} \leq \frac{dQ_{\ll P}}{dP} \quad P\text{-a.e.}$$

(14)

Here is the picture:



Proof:

Recall in the proof of the first RND Thm to show $\frac{dQ}{dP}$ existed we set

$$W = \frac{P+Q}{2} \quad \text{and showed}$$

$$\frac{dQ}{dP} = \frac{dQ/dW}{dP/dW} I_{\{dP/dW \neq 0\}}$$

$$\begin{aligned} &\stackrel{P\text{-a.e.}}{=} \frac{dQ/dW}{dP/dW} \quad \text{by } P\left(\frac{dP}{dW} = 0\right) = 0 \\ &\stackrel{Q \ll P}{=} Q\left(\frac{dP}{dW} = 0\right) = 0 \end{aligned}$$

But for this Thm we don't have $Q \ll P$ so let's simply define

$$Q_{\ll P}(\cdot) = \int \frac{dQ/dW}{dP/dW} I_{\{dP/dW \neq 0\}} dP$$

where dQ/dW & dP/dW exists from the RND Thm by the fact that $P, Q \ll W$.

Now $\forall A \in \mathcal{Q}$

$$\begin{aligned} Q(A) &= Q\left(A \cap \left\{\frac{dP}{d\mu} \neq 0\right\}\right) + Q\left(A \cap \left\{\frac{dP}{d\mu} = 0\right\}\right) \\ &= \underbrace{\int_A I_{\left\{\frac{dP}{d\mu} \neq 0\right\}} \frac{dQ}{d\mu} d\mu}_{=: Q_{\perp}(A)} \quad \text{which is } \perp \text{ to } P \\ &\quad \text{since } \left\{\frac{dP}{d\mu} = 0\right\} \text{ is } P\text{-null.} \\ &= \int_A I_{\left\{\frac{dP}{d\mu} \neq 0\right\}} \frac{dQ/d\mu}{dP/d\mu} \frac{dP}{d\mu} d\mu \\ &= Q_{\ll}(A) \end{aligned}$$

This proves $Q = Q_{\ll} + Q_{\perp}$ where

$Q_{\ll} \ll P$ & $Q_{\perp} \perp P$

To show Q_{\ll} is P -largest let \tilde{Q}

be a measure s.t. $\tilde{Q}(\cdot) \leq Q(\cdot)$ &
 $\tilde{Q} \ll P$. Now let $N = \left\{\frac{dP}{d\mu} = 0\right\}$ and notice

$$\begin{aligned} \int_A \frac{d\tilde{Q}}{dP} dP &= \int_{A \cap N} \frac{d\tilde{Q}}{dP} dP + \int_{A \cap N^c} \frac{d\tilde{Q}}{dP} dP \\ &\leq \int_{A \cap N} \frac{d\tilde{Q}}{dP} dP + Q(A \cap N^c) \end{aligned}$$

since $\tilde{Q}(C) \leq Q(C)$

$$\begin{aligned} &= \underbrace{0}_{\text{since } P(N)=0} + \underbrace{Q_{\ll}(A \cap N^c)}_{\text{since } Q_{\perp}(A \cap N^c)} \\ &= Q(A \cap N^c) = 0 \end{aligned}$$

$$= \int_{A \cap N^c} \frac{dQ_{\ll}}{dP} dP$$

$$= \int_A I_{N^c} \frac{dQ_{\ll}}{dP} dP$$

but this is 1 P -a.e.
 since $P(N)=0$

(15)

$$\therefore \int_A \frac{d\tilde{Q}}{dP} dP \leq \int_A \frac{dQ_{\ll}}{dP} dP \quad \forall A \in \mathcal{Q}$$

$\therefore \frac{d\tilde{Q}}{dP} \leq \frac{dQ_{\ll}}{dP}$ P -a.e. by our result
 on indefinite integrals in Lecture 11.

QED

Example

$$(R, \mathcal{Q}) = (R, \mathcal{B}(R))$$

$$dP = e^{-x} I_{(0, \infty)}(x) dx$$

$$dQ = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

Let's find Q_{\ll} & $Q_{\perp}(\cdot) = Q(\cdot \cap N)$.

Set $N = [-\infty, 0]$ which is P -null and

$$S(x) = \begin{cases} e^x e^{-x^2/2} (2\pi)^{-1/2} & \text{for } x > 0 \\ 0 & \text{o.w.} \end{cases}$$

Now

$$\int_A S dP + Q(A \cap N)$$

$$= \int_A S(x) e^{-x} I_{(0, \infty)}(x) dx + Q(A \cap N)$$

by step in the density

$$= \int_{A \cap N^c} e^{-x^2/2} (2\pi)^{-1/2} dx + Q(A \cap N)$$

$$= Q(A \cap N^c) + Q(A \cap N)$$

$$= Q(A)$$

$\therefore Q_{\ll}(\cdot) = \int_S S dP$ & $Q_{\perp}(\cdot) = Q(\cdot \cap N)$
 satisfies the LD Thm.

(Note the LD is unique but we only need
 that Q_{\ll} is P -largest later)

(16)