36710 - 36752

ADVANCED PROBABILITY OVERVIEW

FALL 2020

LECTURE 9 : WED, SEP 30, 2020

ASYNCHRONOUS LECTURE

RIEMANN VS LEBESGUE INTEGRAL

LEBESGUE INTEGRATION IS MORE GENERAL

1) IT ALLOWS FOR UNBOUNDED FUNCTIONS
2) IT ALLOWS FOR GENERAL FUNCTIONS
3) HAS GOOD LIMIT THEORY

<u>RESULT</u>:
- IF f IS CONTINUOUS ON $[a,b]$ $-\infty < a < b < +\infty$ THEN RIEMANN INTEGRAL OF f IS EQUAL TO THE LEBESGUE INTEGRAL OF f.
- IF f IS BOUNDED, IT IS RIEMANN INTEGRABLE IIF ITS POINTS OF DISCONTINUITY HAS LEBESGUE MEASURE ZERO

<u>RESULT</u> IF $f : I = [a, \infty) \to \mathbb{R}$ IS LEBESGUE INTEGRABLE ON $[a,b]$ $\forall b \geq a$ AND $\int_a^b |f(x)| \leq M$

FOR ALL $b \geq a$ AND SOME $M > 0$ THEN f IS LEB. INTEGRABLE ON I AND ITS INTEGRAL IS

$$\lim_{b \to \infty} \int_a^b f \, d\lambda$$

- IT IS POSSIBLE THAT INTEGRABILITY FAILS BUT THE ABOVE LIMIT EXISTS AS AN IMPROPER RIEMANN INTEGRAL

<u>EXAMPLE</u> 1) $f(x) = \frac{1}{1+x^2}$ $x \in \mathbb{R}$. THEN

$$\int_a^b f(x) \underbrace{d\lambda(x)}_{dx} = \arctan b - \arctan a \leq \pi$$

AND

$$\int_{-\infty}^{+\infty} f(x) \, dx = \lim_{a \to -\infty} \int_a^0 f(x) \, dx + \lim_{b \to \infty} \int_0^b f(x) \, dx$$

$$= \pi/2 + \pi/2 = \pi$$

$\Downarrow$

f IS LEBESGUE INTEGRABLE OVER $\mathbb{R}$

2) $f(x) = \frac{(-1)^n}{n}$ FOR $n-1 \leq x < n$ $x \geq 0$.

FIX $b > 0$ AND LET $m = \lfloor b \rfloor$. THEN

$$\int_0^b f(x) \, dx = \int_0^m f(x) \, dx + \int_m^b f(x) \, dx$$

$$= \sum_{n=1}^{m} \frac{(-1)^n}{n} + \frac{(b-m)(-1)^{m+1}}{m+1}$$

AS $b \to \infty$ THE SECOND TERM VANISHES AND THE

FIRST TERM CONVERGES TO $-\log 2$

$$\int_0^\infty f(x)\,dx = -\log 2$$ NOT QUITE...

$\hookrightarrow$ IMPROPER RIEMANN INTEGRAL

BUT f IS NOT LEBESGUE INTEGRABLE ON $[0,\infty)$

BECAUSE
$$\lim_{b\to\infty} \int_0^b |f|\,dx = \infty$$

SIMILARLY,
$$\lim_{b\to\infty} \int_0^b \frac{\sin x}{x}\,dx = \pi/2$$ IMPROPER RIEMANN INTEGRAL

BUT $$\lim_{b\to\infty} \int_0^b \left|\frac{\sin x}{x}\right| dx = +\infty$$ NOT LEB. INTEGRABLE

3) LET $f: [0,1] \to \mathbb{R}$ $\quad f(x) = 1_{\mathbb{Q}}(x)$

$\hookrightarrow$ RATIONAL

$$\int_0^1 f(x)\,dx = 0$$ AS A LEBESGUE INTEGRAL

BECAUSE $\lambda(\mathbb{Q}) = 0$ SO

$f = 0$ a.e $[\lambda]$

BUT f IS NOT RIEMANN INTEGRABLE

ABSOLUTE CONTINUITY AND RADON-NYKODIM DERIVATIVE

Def LET ν AND μ BE TWO MEASURES ON $(\Omega, \mathcal{F})$. WE SAY THAT ν IS ABSOLUTELY CONTINUOUS WRT μ, $\nu \ll \mu$, WHEN $\mu(A) = 0$ IMPLIES $\nu(A) = 0$, $A \in \mathcal{F}$.

EXAMPLE $(\Omega, \mathcal{F}, \mu)$ LET $f : \Omega \xrightarrow{\text{MEAS.}} \mathbb{R}_{\geq 0}$ AND DEFINE

$$\nu(A) = \int_A f \, d\mu. \quad \text{THEN} \quad \nu \ll \mu.$$

Def (MUTUALLY SINGULAR MEASURES) $(\Omega, \mathcal{F})$ TWO MEASURES μ AND ν ARE MUTUALLY SINGULAR, $\mu \perp \nu$, WHEN THERE EXIST DISJOINT MEAS. SETS S_μ AND S_ν S.T.

$$\mu(S_\mu^c) = 0 = \nu(S_\nu^c).$$

Thm (RADON-NYKODIM) LET μ AND ν BE σ-FINITE MEASURES ON $(\Omega, \mathcal{F})$. THEN $\nu \ll \mu$ IIF $\exists$ NON-NEGATIVE AND MEASURABLE FUNCTION $f : \Omega \to \mathbb{R}_{\geq 0}$ S.T. $\nu(A) = \int_A f(\omega) \, d\mu(\omega)$ FOR ALL $A \in \mathcal{F}$. f IS UNIQUE a.e. $[\mu]$.

Def (RN-DERIVATIVE) THE FUNCTION f IS CALLED THE RADON-NYKODIM DERIVATIVE OF ν WRT μ. IT IS DENOTED WITH $\frac{d\nu}{d\mu}$. IF $f = \frac{d\nu}{d\mu}$ a.e. $[\mu]$ THEN f IS CALLED A VERSION OF $\frac{d\nu}{d\mu}$

- A CONTINUOUS RANDOM VARIABLE X IS A RANDOM VARIABLE WHOSE PROBABILITY DISTRIBUTION μ_X [i.e. $\mu_X(B) = \Pr(X \in B)$, B BOREL SET] IS ABSOLUTELY CONTINUOUS WRT LEBESGUE MEASURE.
  - ↳ cdf OF X IS ALSO CONTINUOUS
  - ↳ THE p.d.f. (PROB. DENSITY FUNCTION) IS JUST THE RN-DERIV. OF μ_X WRT λ (↳ LEBESGUE MEASURE)

- A DISCRETE RANDOM VARIABLE IS ONE WHOSE PROB. DISTR. IS ABSOLUTELY CONTINUOUS WRT THE COUNTING MEASURE OVER A COUNTABLE SUBSET OF $\mathbb{R}$. THEN ITS RN-DER. IS JUST THE p.m.f (PROBABILITY MASS FUNCTION).

- THE SUPPORT OF μ_X IS EQUAL TO $\overline{\{x,\ \frac{d\mu_X}{d\lambda}(x) > 0\}}$ (CLOSURE)

- <u>STATISTICAL MODEL</u>

  THEORY OF STATISTICS

  LET $(\mathcal{X}, \mathcal{B})$ BE A SAMPLE SPACE (TYPICALLY $\mathcal{X} = \mathbb{R}^s$, $\mathcal{B} = \mathcal{B}^s$)

  LET $\Theta \subseteq \mathbb{R}^d$ BE AN OPEN SET

  ↳ PARAMETER SPACE

  A STATISTICAL MODEL IS A COLLECTION OF PROBABILITY DISTRIBUTIONS ON $(\mathcal{X}, \mathcal{B})$ INDEXED BY Θ.

  $$\mathcal{P} = \{ \underbrace{P_\theta}_{\text{PROB. DISTR.}},\ \theta \in \Theta \}$$

  EXAMPLE $\mathcal{P} = \{ N(\mu, 1),\ \mu \in \mathbb{R} \}$

  WHERE $P_\theta \ll \mu$ FOR ALL θ AND SOME σ-FINITE MEASURE μ ON $(\mathcal{X}, \mathcal{B})$

  THE DENSITY OF P_θ IS $\frac{dP_\theta}{d\mu}$ → WELL-BEHAVE FUNCTIONS OF θ

- KULLBACK-LEIBER DIVERGENCE BETWEEN PROB. DISTRIBUTIONS.

$(\Omega, \mathcal{F})$ $\quad P$ AND Q ARE PROB. MEASURES $\quad [\mu = P + Q]$

THE KL DIVERGENCE OF P FROM Q IS

$$KL(P, Q) = \begin{cases} \int \log\left(\frac{dP}{dQ}\right) dP & \text{IF } P \ll Q \\ \infty & \text{OTHERWISE} \end{cases}$$

$KL(P, Q) \geq 0$ AND $KL(P, Q) = 0$ IFF $P = Q$

NOT A METRIC. IF $P \sim N_d(\mu_1, \Sigma)$ $\quad Q \sim N_d(\mu_2, \Sigma)$

THEN $KL(P, Q) = (\mu_1 - \mu_2)^T \Sigma^{-1} (\mu_1 - \mu_2)$

IF $P \ll \lambda$ AND $Q \ll \lambda$ THEN $KL(P, Q) = \int \log\left(\frac{p(x)}{q(x)}\right) p(x)\, dx$

WHERE $p = \frac{dP}{d\lambda}$ $\quad q = \frac{dQ}{d\lambda}$