Notes on Probability

1 Radon-Nikodym Theorem

1.1 Statement

Definition 1.1.

- ***Indefinite Integrals***.

  Let $(\Omega, \mathcal{F}, \mu)$ *be a probability space.* $p \in \mathcal{F}_+$. *Define*

  $$\nu(A) = \int_A \mu(\mathrm{d}x)p(x).$$

  ν *is a measure on* $(\Omega, \mathcal{F})$, *and is called* *the indefinite integral of* p *w.r.t* μ.

- ***Absolutely continuous***.

  Let μ *and* ν *be measures on* $(\Omega, \mathcal{F})$. v *is said to be* *absolutely continuous w.r.t* μ *if* $\forall A \in \mathcal{F}$,

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

Remark.

- *The indefinite integral gives us a convenient tool to construct new measures from the old.*
- *If* ν *is the indefinite integral of* p *w.r.t* μ, *for every* $f \in \mathcal{F}_+$, *we have*

  $$\int_\Omega \nu(\mathrm{d}x)f(x) = \int_\Omega \mu(\mathrm{d}x)p(x)f(x);$$

  sometimes we write

  $$v(\mathrm{d}x) = \mu(\mathrm{d}x)p(x).$$

- *If* ν *is the indefinite integral of* p *w.r.t* μ, ν *is absolutely continuous w.r.t* μ. *(The Radon-Nikodym Theorem says the converse is true as well.)*

Theorem 1.2 (Radon-Nikodym Theorem)**.** *Let* μ *be a* σ*-finite measure and* ν *be absolutely continuous w.r.t* μ. *Then there exists* $p \in \mathcal{F}_+$ *s.t.*

$$\int_\Omega \nu(\mathrm{d}x)f(x) = \int_\Omega \mu(\mathrm{d}x)p(x)f(x), \quad f \in \mathcal{F}_+.$$

Moreover, p *is unique for* μ*-almost every* $\omega \in \Omega$.

Remark. p *can be denoted by* $\mathrm{d}\nu/\mathrm{d}\mu$, *and be called* *the Radon-Nikodym derivative of* ν *w.r.t* μ.

1.2 Application of Radon-Nikodym Theorem

Theorem 1.3 (Existence and Uniqueness of Conditional Expectation)**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space,* $X \in \mathcal{F}_+$, *and* $\mathcal{F}_1 \subset \mathcal{F}$ *be a sub-*σ*-algebra.*
$\implies$ $\mathbb{E}(X|\mathcal{F}_1)$ *exists and is unique.*

Proof. For each $A \in \mathcal{F}_1$, define $P(A) = \mathbb{P}(A)$, $Q(A) = \int_A \mathbb{P}(\mathrm{d}\omega)X(\omega)$. ($P$ is the restriction of $\mathbb{P}$ to $\mathcal{F}_1$; Q is the indefinite integral of X w.r.t. P.)

Check conditions: P is a finite measure; Q is absolutely continuous w.r.t P.

By the Radon-Nikodym theorem, $\exists \bar{X} \in \mathcal{F}_+$ s.t.

$$\int_\Omega Q(\mathrm{d}\omega)V(\omega) = \int_\Omega P(\mathrm{d}\omega)\bar{X}(\omega)V(\omega)$$

for every $V \in \mathcal{F}_+$; i.e.

$$\mathbb{E}XV = \mathbb{E}\bar{X}V$$

for every $V \in \mathcal{F}_+$.

$\implies \bar{X} = \mathbb{E}(X|\mathcal{F}_1)$, by the definition of conditional expectations. □

2 Riesz Representation Theorem (RRT)

2.1 Setting the Stage

Theorem 2.1 (Riesz Representation Theorem)**.** *Every continuous linear functional $\phi : C(X) \to \mathbb{R}$ can be represented by a unique signed Baire measure μ on X via*

$$\phi(f) = \int_X f \ \mathrm{d}\mu.$$

Remark.

- *Fix a compact Hausdorff space X. For each μ, we can define a bounded linear functional $\phi_\mu :\ f \mapsto \int_X f \ \mathrm{d}\mu$.*
- *$\mu \neq \mu' \implies \phi_\mu \neq \phi_{\mu'}$. That is, $\mu \mapsto \phi_\mu$ is injective.*
- *The Riesz representation theorem says that $\mu \mapsto \phi_\mu$ is also surjective.*

Definition 2.2. *We will use the following notations to re-state the RRT:*

- *Comp - Compact Hausdorff spaces + continuous maps*
- *Ban - Banach spaces + continuous linear operators*
- *$C(X)$ - all continuous real-valued functions on X*
- *$C(X)^*$ - dual space of $C(X)$*
- *$M(X)$ - all Baire measures on X*

Lemma 2.3. *It is easy to check:*

- *C is a contravariant functor from Comp to Ban.*

$$\begin{aligned} C\colon X &\longmapsto C(X) \\ \alpha &\longmapsto \alpha^\sharp \end{aligned}$$

  where $\alpha^\sharp : g \mapsto g \circ \alpha$.
- *M is a convariant functor form Comp to Ban.*

$$\begin{aligned} M\colon X &\longmapsto M(X) \\ \alpha &\longmapsto \alpha^* \end{aligned}$$

  where $\alpha^* : \mu \mapsto \mu \circ \alpha^{-1}$.
- *Define $\iota_X : M(X) \to C(X)^*$ by $\mu \mapsto \phi_\mu$ where $\phi_\mu(f) = \int_X f \ \mathrm{d}\mu$.*
  Then ι is a natural transformation from M to C^*.

Theorem 2.4 (Riesz Representation Theorem Revisited)**.** *ι is a natural equivalent from M to C^*.*

2.2 Proof of RRT

Definition 2.5.

- *A topological space is called* extremally disconnected *if each open set has an open closure.*
- $\mathscr{C}$ *- all clopen sets*

Lemma 2.6. *Let* $X \in Comp$ *be extremally disconnected. Then*

a) the Baire sets in X *are generated by* $\mathscr{C}$*;*

b) the simple functions based on $\mathscr{C}$ *are uniformly dense in* $C(X)$*.*

Lemma 2.7. *The Stone-Čech compactification of a discrete space is extremally disconnected.*

Lemma 2.8. *For each object* $Y \in Comp$*, there is a morphism* $\alpha : X \to Y$ *such that* X *is extremally disconnected.*

Proof of RRT. Let Y be an object in *Comp*, and $\alpha : X \to Y$ be the morphism starting at an extremally disconnected object X.

- $\alpha^\sharp$ is a norm isomorphism.
- By the Hahn-Banach Theorem, $\alpha^{\sharp\sharp}$ is also surjective.
- ι_X is surjective $\implies$ ι_Y is surjective.

□

2.3 Application of RRT in the Moment Problem

Exsercise 2.3.1. *Show that the moments completely determine the probability distribution if it is concentrated on a finite interval (that if,* $\mathbb{P}(X_n \in [a,b]) = 1$*, for all* n*).*

Proof. **The first proof is based on the Riesz representation theorem.**

- Assume $F,\ G$ are concentrated on $[a,b]$ and
$$\int_a^b x^k \ \mathrm{d}F = \int_a^b x^k \ \mathrm{d}G \quad \forall k \in \mathbb{R}.$$
We want to prove $F = G$ on $[a,b]$.

- Let $C[a,b]$ be the space of all continuous functions on $[a,b]$ with the uniform norm. Define
$$\begin{aligned} T_F \colon C[a,b] &\longrightarrow \mathbb{R} \\ h &\longmapsto \int_{[a,b]} h \ \mathrm{d}F \\ T_G \colon C[a,b] &\longrightarrow \mathbb{R} \\ h &\longmapsto \int_{[a,b]} h \ \mathrm{d}G \end{aligned}$$
It is easy to check that T_F and T_G are both continuous linear functionals.

- Obviously, for each polynomial $h(x) = \sum_{i=0}^n a_i x^i$, $T_F(h) = T_G(h)$. Thus, $T_F = T_G$ on $C[a,b]$, since polynomials are a dense set of $C[a,b]$ by the Stone-Weiestrass theorem.

- (Riesz representation theorem) Every continuous linear functional ψ over $C[a,b]$ can be uniquely represented as

$$\psi(f) = \int_{[a,b]} f \, \mathrm{d}\mu$$

where μ is .

It suffices to check F and G are bounded variation funcions.

$\square$

3 Poisson Ramdom Measures

3.1 Characterization via Laplace functionals

Definition 3.1. *A random measure N on $(E, \mathcal{E})$ is called* Poisson with mean ν *if*

a) for every $A \in \mathcal{E}$, $N(A) \sim \mathrm{Poisson}(\nu A)$.

b) for disjoint $A_1, \ldots, A_n \in \mathcal{E}$, $N(A_1), \ldots, N(A_n)$ are independent.

Remark.

- *For every $f \in \mathcal{E}_+$, define*

$$Nf(\omega) = \int_E N(\omega, \mathrm{d}x) f(x).$$

- *$f \mapsto \mathbb{E}e^{-Nf}$ is called* the Laplace functional of N.
- *We say a sequence of RVs X taking values in $(E, \mathcal{E})$ forms a Poisson random measure on $(E, \mathcal{E})$, if*

$$Nf = \sum_{i \in I} f \circ X_i, \quad f \in \mathcal{E}_+$$

is Poisson with some mean meaure ν.

Theorem 3.2. *Let N be a random measure on $(E, \mathcal{E})$. It is Poisson with mean ν if and only if*

$$\mathbb{E}e^{-Nf} = e^{-\nu(1-e^f)}, \quad f \in \mathcal{E}_+,$$

where $\nu f = \int_E \nu(\mathrm{d}x) f(x) = \int_E f \mathrm{d}\nu$.

Proof. $\Rightarrow$: It suffices to prove this result for all simple functions.

Assume $f = \sum a_i \mathbf{1}_{A_i}$. Then $Nf = \sum a_i N(A_i)$. By the defnition, all A_i are independent, and the distribution of $N(A_i)$ is $\mathrm{Poisson}(\nu A_i)$, so

$$\begin{aligned} \mathbb{E}e^{-Nf} &= \mathbb{E}e^{-\sum a_i N(A_i)} \\ &= \prod \mathbb{E}e^{-a_i N(A_i)} \\ &= \prod \exp\left[\nu A_i \cdot (e^{-a_i} - 1)\right] \\ &= \exp\left[-\sum \nu A_i \cdot (1 - e^{-a_i})\right] = e^{-\nu(1-e^f)} \end{aligned}$$

If $f \in \mathcal{E}_+$, we can use a sequence of simple functions in $\mathcal{E}_+$ increasing to f. Then use the continuity of Laplace functionals and the monotone convergence theorem.

$\Leftarrow$: Note that $\{\mathbb{E}e^{-Nf}\}$ uniquely determines the probability law of N. And by the necessity part, it must be Poisson. $\square$