Supplementary:
Basics of Stochastic Processes

Hyunwoo Gu

1 Basic measure theory

Lebesgue integration

We say f is **Lebesgue integrable** if

$$\int |f| d\mu < \infty$$

We can **construct** the Lebesgue integral by using **simple functions**, which approximate a measurable funcion. They are basically a finite linear combination of indicator functions,

$$\sum_k a_k \mathbf{1}_{S_k}$$

where S_k : measurable. Under $0 \times \infty = 0$, we can define

$$\int \left(\sum_k a_k \mathbf{1}_{S_k} \right) d\mu = \sum_k a_k \int \mathbf{1}_{S_k} d\mu = \sum_k a_k \mu(S_k)$$

If B: a measurable subset of Ω and s is a measurable simple function, one defines

$$\int_B s d\mu = \int 1_B s d\mu = \sum_k a_k \mu(S_k \cap B)$$

Let f be a nonnegative measurable function on Ω, a subset of the extended real number line. We **define**

$$\int_\Omega f d\mu = \sup \left\{ \int_\Omega s d\mu : 0 \leq s \leq f \right\}$$

where s is constrained to be simple. Alternatively, let $s_n(x)$ be the simple function whose value is $k/2^n$ whenever $k/2^n \leq f(x) < (k+1)/2^n$ for k a non-negative integer less than, say 4^n.

$$\int f d\mu = \lim_{n \to \infty} \int s_n d\mu$$

Note that we can define, for general f,

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu$$

Lebesgue-Stieltjes integration

A generaliation of Riemann-Stieltjes integration and Lebesgue integration, which is just the ordinary Lebesgue integral with respect to a measure known as the **Lebesgue–Stieltjes measure**.

The Lebesgue-Stieltjes integration

$$\int_a^b f(x)dG(x)$$

is defined when $f : [a, b] \to \mathbb{R}$ is **Borel-measurable** and **bounded**, and $G : [a, b] \to \mathbb{R}$ is of bounded variation

Assume $G(\cdot)$ is a right continuous step function having jumps at $x_1, x_2, \cdots$. Then

$$\int_a^b f(x)dG(x) = \sum_{a<x_j\le b} f(x_j)\{G(x_j) - G(x_j^-)\} = \sum_{a<x_j\le b} f(x_j)\Delta G(x_j)$$

Let $F : \mathbb{R} \to \mathbb{R}$ be a distribution function. Then **there uniquely exists** a measure μ_F on $(\mathbb{R}, \mathcal{B}(\mathbb{R})$ s.t.

$$\mu_F\left((a, b]\right) = F(b) - F(a), \quad \forall a < b \in \mathbb{R}$$

Thus,

$$\int xdF(x) = \int x\mu_F(dx) = \int xd\mu_F(x)$$

where the last equality is a **Lebesgue integral**. Note that $P_X\left((a, b]\right) = \mu_F\left((a, b]\right)$ for $X \sim F$. Thus

$$P_X = \mu_F$$

Thus we have

$$\mathbb{E}(X) = \int_\Omega X dP_X = \int_\mathbb{R} x P_X(dx) = \int_\mathbb{R} x\mu_F(dx) = \int_\mathbb{R} x dF(x)$$

Note that

$$\mathbb{E}[X^k] = M(0)$$

where $M(t) = \sum_{l=0}^{\infty} \frac{\mathbb{E}(X^l)}{l!} t^k, \forall t \in (-\epsilon, \epsilon), \exists \epsilon > 0$.

Radon-Nikodym theorem

If $\nu << \mu$, then there is a **measurable function** $f : X \to [0, \infty)$ such that for any measurable set $A \subset X$,

$$\nu(A) = \int_A f d\mu$$

For example, KL divergence from $\mu to \nu$ is defined

$$D_{KL}(\mu \| \nu) = \int_\mathcal{X} \log \left(\frac{d\mu}{d\nu} d\mu \right)$$

For another example, let $\Omega = [0, 1]$, and $\mathbb{P}$ is uniform on Ω. Define $X : [0, 1] \to \mathbb{R}$ by $X(w) = -\log w$. Under $\mathbb{P}$ the rv X has an Exp(1) distribution. Alternative measure Q assigns $Q[a, b] = b^2 - a^2$. Under Q the rv X has an Exp(2) distribution:

$$Q\{X \leq x\} = Q\{w : -\log w \leq x\} = Q[e^{-x}, 1] = 1^2 - (e^{-x})^2 = 1 - e^{-2x}$$

The Radon-Nikodym derivative of Q wrt $\mathbb{R}$ is $\frac{dQ}{d\mathbb{P}}(\omega) = 2\omega$.

Generally, the Radon Nikodym derivative f of Q with respect to P is defined by the equation

$$Q(E) = \int_E f dP$$

for every measurable set E, if $Q << P$ In order find what this f is it is enough consider the sets $E = [0, x]$ where $0 \leq x \leq 1$.

Thus we have to find f such that $1 - e^{-2x} = Q(X \leq x) = \int_{[0,x]} f(t)dt$. [Note that P is just the uniform measure (i.e. the Lebesgue measure on $[0, 1]$ so $\int_E f dP = \int_E f(y)dy$).

To find f from the equation $1 - e^{-2x} = \int_0^x f(y)dy$ simply differentiate both sides with respect to x. Hence $f(x) = 2e^{-2x}$.

Note that for the **cumulative intensity process**, we have

$$d\Lambda(t) = \lambda(t)dt$$

However, generally we cannot define

$$\frac{dN(t)}{dt}$$

since $N(t)$ is not absolutely continuous wrt t.