# Measure and Random Variable

**Definition** A ***sigma algebra*** $\mathscr{F}$ of subsets of $\Omega$:

- $\Omega \in \mathscr{F}$
- close under union and intersection
- close under complement

**Definition** A set function $\mu$ on $\mathscr{F}$ is a ***measure*** if (Probability measure if $\mu(\Omega) = 1$):

- $\mu(A) \in [0, \infty]$ for all $A \subset \Omega$
- $\mu(\emptyset) = 0$
- close under complement
- $\mu(\bigcup_k^\infty A_k) = \sum_k^\infty \mu(A_k)$ where $\{A_k\}$ is disjoint sequence in $\mathscr{F}$ and $\bigcup_k^\infty A_k \in \mathscr{F}$

**Definition** ***Outer measure*** $\mu^*$ is function defined on all subsets of $\Omega$:

- $\mu^*(A) \in [0, \infty]$ for all $A \subset \Omega$
- $\mu^*(\emptyset) = 0$
- $\mu^*$ is monotone: $\mu^*(A) \le \mu^*(B)$ if $A \subset B$
- $\mu^*$ is countably subadditive: $\mu^*(\bigcup_n A_n) \le \sum_n \mu^*(A_n)$

**Definition** ***Lebesgue measure*** $\lambda_k$ on $\mathscr{R}^k$ *(Note: R denote euclidean space. $\mathscr{R}$ denotes sigma field on euclidean space.)*

- $\lambda_k(\{(x_1, ..., x_k) | a_i < x_i < b_i\}) = \prod_{i=1}^k (b_i - a_i)$
- $\lambda_k(A) = \lambda_k(A + x)$ (translation invariance)
- $\lambda_k(TA) = |\det T| \lambda_k(A)$, $T$ is linear and nonsingular.
- $\lambda_k$ is regular (finite measure to bounded set)

**Definition** Function $T$ between two measure spaces $(\Omega, \mathscr{F})$ and $(\Omega', \mathscr{F}')$ is ***measurable*** $\mathscr{F}/\mathscr{F}'$ if $\forall A \in \mathscr{F}', T^{-1}A \in \mathscr{F}$. We say $T$ is ***measurable*** $\mathscr{F}$ if it is measurable $\mathscr{F}/\mathscr{R}^1$.

**Definition** ***Probability Space*** is denoted as $(\Omega, \mathscr{F}, P)$

**Definition** A ***random variable*** on $(\Omega, \mathscr{F}, P)$is a real-valued function $X = X(\omega)$ measurable $\mathscr{F}$. A ***random vector*** is mapping from $\Omega$ to $R^k$ that is measurable $\mathscr{F}$. e.g. $X(\omega) = (X_1(\omega), ..., X_k(\omega))$.

**Definition** The ***distribution (law) of random variable*** $X$ is the probability measure $\mu = PX^{-1}$ on $(R^1, \mathscr{R}^1)$ defined by

$$\mu(A) = P[X \in A], \quad A \in \mathscr{R}^1$$

($P[X \in A]$ means $P[\omega : X(\omega) \in A]$.

**Definition** The ***distribution function of random variable*** $X$ is

$$F(x) = \mu(-\infty, x] = P[X \le x]$$

. If $F$ is right-continuous, and non-decreasing, there is a random variable $X$ on some $(\Omega, \mathscr{F}, P)$corresponding to $F$.

**Definition** If $X = (X_1, ..., X_k)$,

$$\mu(A) = P[(X_1, ..., X_k) \in A], \quad A \in \mathscr{R}^k$$

$$F(x_1, ..., x_k) = P[X_1 \le x_1, ..., X_k \le x_k]$$

$\mu, F$ are called ***joint distribution*** and ***join distribution function*** of $X$.

# Integration

## *Let $f, g$ be real measurable function on $(\Omega, \mathscr{F}, P)$*

**Definition** The ***definite integral*** is denoted:

$$\int f d\mu = \int_\Omega f(\omega) d\mu(\omega) = \int_\Omega f(\omega)\mu(d\omega)$$

is defined by

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

and

$$\int f^\pm d\mu = \sup \sum_i \left[ \inf_{\omega \in A_i} f^\pm(\omega) \right] \mu(A_i)$$

Where $\{A_i\}$ is a finite decomposition of $\Omega$ into $\mathscr{F}$-sets.

**Properties** General integral:

- Monotonicity
- Linearity
- (Monotone Convergence)
  if $0 \le f_n \uparrow f$ almost everywhere then $\int f_n d\mu \uparrow \int f d\mu$
- (Fatou's lemma)
  $\int \liminf_n f_n d\mu \le \liminf_n \int f_n d\mu, f_n \ge 0$.
- (Dominated Convergence THM)
  $|f_n| \le g$ almost everywhere ($g$ integrable), $f_n \to f$ almost everywhere, then $f$, $f_n$ integrable and $\int f_n d\mu \to \int f d\mu$
- $f, g$ integrable and $\int_A f d\mu = \int_A g d\mu$ for all $Ain\mathscr{F}$, then $f = g$ almost everywhere.

**Definition** If $\delta$ is nonnegative measurable function, define measure $v$

$$v(A) = \int_A \delta d\mu, A \in \mathscr{F}$$

. Then $v$ is said to have ***density*** $\delta$ wrt to $\mu$.

**Properties** $\int f dv = \int f \delta d\mu$

**Definition** ***Transformatin of Measure***: Given measurable mapping $T$ between $(\Omega, \mathscr{F})$ and $(\Omega', \mathscr{F}')$. For a measure $\mu$ on $\mathscr{F}$, define set function $\mu T^{-1}$ on $\mathscr{F}'$ by

$$\mu T^{-1}(A') = \mu(T^{-1}A'), \forall A' \in \mathscr{F}'$$

**Definition** ***Change of variable***: Suppose $f$ is real function on $\Omega'$ measurable $\mathscr{F}'$, so $fT$ is real function on $\Omega$ measurable $\mathscr{F}$. $f$ is nonnegative or (integratable wrt $\mu T^{-1} \Leftrightarrow fT$ integrable wrt $\mu$). The following hold:

$$\int_\Omega f(T\omega)\mu(d\omega) = \int_{\Omega'} f(\omega')\mu T^{-1}(d\omega')$$

$$\int_{T^{-1}A'} f(T\omega)\mu(d\omega) = \int_{A'} f(\omega')\mu T^{-1}(d\omega')$$

**Definition** Let $X \sim \mu$ and $Y \sim v$ be independent, ***Convolution*** of $\mu$ and $v$ is defined as:

$$\begin{aligned} (\mu * v)(H) &= P[X + Y \in H] \\ &= \int_{-\infty}^{\infty} v(H - x)\mu(dx), \quad H \in \mathscr{R}^1 \\ &= \int_{-\infty}^{\infty} P[Y \in H - x]\mu(dx) \end{aligned}$$

If $\mu$ and $v$ are distribution functions $F, G$. And $f, g$ are densities Then

$$(F * G)(y) = \int_{-\infty}^{\infty} G(y-x)dF(x)$$
$$= (f * g)(y) = \int_{-\infty}^{\infty} g(y-x)f(x)dx$$

**Definition** ***Expected value*** of $X$ on $(\Omega, \mathscr{F}, P)$is defined as:

$$E[X] = \int_{\Omega} X dP = \int_{\Omega} X(\omega)P(d\omega)$$

e.g. If $X \sim \mu$ and $g$ is real measurable $\mathscr{R}$, then

$$\begin{aligned} E[g(X)] &= \int_{\Omega} g(X)dP \\ &= \int_{\Omega} g(X(\omega))P(d\omega) \\ &= \int_{R} g(x)PX^{-1}(dx) \quad (chage\ of\ variable, X(\omega) = x) \\ &= \int_{R} g(x)\mu(dx) \quad (PX^{-1} = \mu) \end{aligned}$$

# Differentiaion

**Definition** Given two measures $\mu, v$, $v$ is said to be ***absolutely continuous*** with respect to $\mu$, denoted $v \ll \mu$, if $\mu(A) = 0 \Rightarrow v(A) = 0 \ \ \forall A \in \mathscr{F}$

**Definition** ***Radon-Nikodym theorem***: Given a measurable space $(\Omega, \mathscr{F})$, if two $\sigma$-finite measures $v, \mu$ and $v \ll \mu$, then there is a measurable function $f : X \to [0, \infty)$ such that:

$$v(A) = \int_{A} f d\mu \quad (\forall A \subset X, \quad f = \frac{dv}{d\mu})$$

# Probability Basics

**Properties** (Probability measure)

- $p(A \vee B) = p(A) + p(B) - p(A \wedge B)$
- $p(A, B) = p(A \wedge B) = p(A|B)p(B)$
- $p(A) = \sum_{b} p(A, B) = \sum_{b} p(A|B = b)p(B = b)$
- $p(X_{1:D}) = p(X_1)p(X_2|X_1)p(X_3|X_2, X_1)...p(X_D|X_{1:D-1})$
- $p(A|B) = \frac{p(A,B)}{p(B)}$ if $p(B) > 0$
- $p(X = x|Y = y) = \frac{p(X=x,Y=y)}{p(Y=y)} = \frac{p(X=x)p(Y=y|X=x)}{\sum_{x'} p(X=x')p(Y=y|X=x')}$
- $X \perp Y \iff p(X, Y) = p(X)p(Y)$
- $X \perp Y|Z \iff p(X, Y|Z) = p(X|Z)p(Y|Z)$
- $X \perp Y|Z \iff p(x, y|z) = g(x, z)h(y, z) \ \forall x, y, z \ s.t \ p(z) > 0$
- $F(q) = p(X \leq q)$
- $f(x) = \frac{d}{dx}F(x)$
- $P(a < X \leq b) = F(b) - F(a) = \int_{a}^{b} f(x)dx$
- $Unif(x|a, b) = \frac{1}{b-a} I(a \leq x \leq b)$
- $cov[X, Y] = E[(X - EX)(Y - EY)] = E[XY] - E[X][Y]$