

# Measure

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

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

**Definition** A set function  $\mu$  on  $\mathcal{F}$  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  $\mathcal{F}$  and  $\bigcup_k^\infty A_k \in \mathcal{F}$

**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:  $\mu^*(A) \leq \mu^*(B)$  if  $A \subset B$
- $\mu^*$  is countably subadditive:  $\mu^*(\bigcup_n A_n) \leq \sum_n \mu^*(A_n)$

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

- $\lambda_k(\{(x_1, \dots, x_k) | a_i < x_i < b_i\}) = \prod_{i=1}^k (b_i - a_i)$
- $\lambda_k(A) = \lambda_k(A + x)$  (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  $(\Omega, \mathcal{F})$  and  $(\Omega', \mathcal{F}')$  is *measurable*  $\mathcal{F}/\mathcal{F}'$  if  $\forall A \in \mathcal{F}', T^{-1}A \in \mathcal{F}$ . We say  $T$  is *measurable*  $\mathcal{F}$  if it is measurable  $\mathcal{F}/\mathcal{R}^1$ .

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

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

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

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

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

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

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

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

# Integration

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

**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  $\mathcal{F}$ -sets.

**Properties** General integral:

- Monotonicity
- Linearity
- (Monotone Convergence)  
if  $0 \leq f_n \uparrow f$  almost everywhere then  $\int f_n d\mu \uparrow \int f d\mu$
- (Fatou's lemma)  
 $\int \liminf_n f_n d\mu \leq \liminf_n \int f_n d\mu, f_n \geq 0$ .
- (Dominated Convergence THM)  
 $|f_n| \leq g$  almost everywhere ( $g$  integrable),  $f_n \rightarrow f$  almost everywhere, then  $f, f_n$  integrable and  $\int f_n d\mu \rightarrow \int f d\mu$
- $f, g$  integrable and  $\int_A f d\mu = \int_A g d\mu$  for all  $A \in \mathcal{F}$ , then  $f = g$  almost everywhere.

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

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

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

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

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

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

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

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

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

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

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

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

# 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  $p(B) > 0$
- $p(X = x|Y = y) = \frac{p(X=x, Y=y)}{p(Y=y)} = \frac{p(X=x)p(Y=y|X=x)}{\sum_{x'} p(X=x')p(Y=y|X=x')}$
- $X \perp Y \iff p(X, Y) = p(X)p(Y)$
- $X \perp Y|Z \iff p(X, Y|Z) = p(X|Z)p(Y|Z)$
- $X \perp Y|Z \iff p(x, y|z) = g(x, z)h(y, z) \forall x, y, z \text{ s.t } p(z) > 0$
- $F(q) = p(X \leq q)$
- $f(x) = \frac{d}{dx}F(x)$
- $P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x)dx$
- $Unif(x|a, b) = \frac{1}{b-a}I(a \leq x \leq b)$
- $cov[X, Y] = E[(X - EX)(Y - EY)] = E[XY] - E[X]E[Y]$