Ch1. Probability and Measure

- Set: $\mathcal{X}$.

- σ-field: A collection $\mathcal{A}$ of subsets of a set $\mathcal{X}$.
  1. $\mathcal{X} \in \mathcal{A}$ and $\phi \in \mathcal{A}$
  2. if $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
  3. if $A_1, A_2, \cdots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

  ex. $\mathcal{A} = \{\phi, \mathcal{X}\}$ $\quad$ $\mathcal{A} = \{\phi, A, A^c, \mathcal{X}\}$

  $\mathcal{A}$ = power set of $\mathcal{X}$

- Measure: A function μ on a σ-field of $\mathcal{X}$.
  1. $\mu: \mathcal{A} \to \mathbb{R}$. $\quad \forall A \in \mathcal{A}, 0 \le \mu(A) \le +\infty$
  2. if $A_1, A_2, \cdots$ are disjoint, $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$

ex. $B_1 \subset B_2 \subset \cdots \qquad B = \bigcup_{n=1}^{\infty} B_n$

$$\mu(B) = \mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n - B_{n-1}) = \lim_{n\to\infty} \mu(B_n)$$

ex. Counting measure if $\mathcal{X}$ is countable.

$\mu(A) = \# A$.

ex. Lebesgue measure : $\mathcal{X} = \mathbb{R}^n$

$$\mu(A) = \int \cdots \int_A dx_1 \cdots dx_n$$

- $(\mathcal{X}, \mathcal{A})$: measurable space

  $(\mathcal{X}, \mathcal{A}, \mu)$: measure space

- A measure μ is finite if $\mu(\mathcal{X}) < \infty$

  μ is σ-finite if there exist $A_1, A_2, \cdots$ in $\mathcal{A}$ s.t. $\mu(A_i) < \infty$ for all i and $\bigcup_{i=1}^{\infty} A_i = \mathcal{X}$.

- Probability measure $\mu(\mathcal{X}) = 1$

  $(\mathcal{X}, \mathcal{A}, \mu)$: Probability space.

ex. Borel sets of $\mathbb{R}^n$. $\mathcal{A}$ is the smallest σ-field that contains all "rectangles" $(a_1, b_1) \times \cdots \times (a_n, b_n)$

- Integration $\int f \, d\mu$ or $\int f(x) \, d\mu(x)$

ex. μ counting measure on $\mathcal{X}$.

$$\int f \, d\mu = \sum_{x \in \mathcal{X}} f(x)$$

ex. μ is Lebesgue measure on R^n.

$$\int f \, d\mu = \int \cdots \int f(x_1 \cdots x_n) \, dx_1 \cdots dx_n$$

- $(\mathcal{X}, \mathcal{A})$ — measurable space. f — a real-value function

f is measurable if

$$f^{-1}(B) \triangleq \{ x \in \mathcal{X} : f(x) \in B \} \in \mathcal{A}$$ for every

Borel set B.

- Indicator function:

$$I_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

For any $B \subset R$

$$I_A^{-1}(B) = \begin{cases} \phi & 0 \notin B \quad 1 \notin B \\ A & 0 \notin B \quad 1 \in B \\ A^c & 0 \in B \quad 1 \notin B \\ \mathcal{X} & 0 \in B \quad 1 \in B \end{cases}$$

- basic properties for integrals.

1. $\int I_A(x)\, d\mu(x) = \mu(A)$ $\longleftrightarrow$ link $\int f d\mu$ and μ

2. $f \cdot g$ — nonnegative measurable functions $a, b > 0$

$$\int (af + bg)\, d\mu = a \int f d\mu + b \int g d\mu \longleftrightarrow \text{linearity}$$

3. $f_1 \le f_2 \le \cdots$ nonnegative measurable functions,

$$f(x) = \lim_{n\to\infty} f_n(x) \Rightarrow \int f\, d\mu = \lim_{n\to\infty} \int f_n d\mu$$

$\longleftrightarrow$ taking limit

- Simple function: $\sum_{i=1}^{n} a_i I_{A_i}$ $\underline{a_i > 0}$.

ex. $1_{(\frac{1}{2}, \pi)} + 2\, 1_{(1,2)}$

Thm: if f is nonnegative and measurable, then there exist nonnegative simple functions $f_1 \le f_2 \le \cdots$ with $f = \lim_{n\to\infty} f_n$.

- $f^+(x) = \max\{f(x), 0\}$. $f^-(x) = -\min\{f(x), 0\}$

$f = f^+ - f^-$. $\underline{|f| = f^+ + f^-}$ $\int f d\mu = \int f^+ d\mu - \int f^- d\mu$.

$\int |f| d\mu < \infty \Rightarrow f$ is integrable

Events. Probabilities, and RVs.

- $(\mathcal{E}, \mathcal{B})$: measurable space

  $(\mathcal{E}, \mathcal{B}, P)$: probability space.

  $B \in \mathcal{B}$: event. $\quad e \in \mathcal{E}$: outcome.

  $P(B)$.

- RV: $\quad X : \mathcal{E} \to \mathbb{R}$.

  $$P_X(A) = P\left(\{ e \in \mathcal{E} : X(e) \in A \} \right) \triangleq P(X \in A)$$

  for Borel sets A.

  $X \sim Q$.

  distribution function: $F_X(x) = P(X \le x) = P_X\left((-\infty, x] \right)$

Null sets.

- A set N is called <u>null</u> if $\mu(N) = 0$

  If a statement holds for $x \in X - N$ with N null, the statement is said to hold <u>almost everywhere</u> (a.e.)

  The statement holds with probability one: if hold for $x \in B$

  $\mu(B) = 1$

  Null set $\to$ does not affect its integral. $\quad \mu(B^c) = 0$

Densities

- P and μ — measures on a σ-field $\mathcal{A}$ of $\mathcal{X}$.

P is called absolutely continuous w.r.t. μ. $P \ll \mu$.

if $P(A) = 0$ whenever $\mu(A) = 0$

Thm: (Radon-Nikodym): if $P \ll \mu$, there exists a nonnegative measurable function f. s.t.

$$P(A) = \int_A f \, d\mu \triangleq \int f 1_A \, d\mu.$$

f: Radon-Nikodym derivative of P w.r.t. μ.

$$f = \frac{dP}{d\mu}.$$

if $X \sim P_X$. $P_X \ll \mu$. $p = \frac{dP_X}{d\mu}$. — density.

ex: X — density.

$$F_X(x) = P(X \le x) = P_X((-\infty, x]) = \int_{-\infty}^{x} p(u)\,du$$

$$p(x) = F_X'(x).$$

ex. Discrete $\mathcal{X}_0 \subseteq \mathbb{R}$

$$\mu(B) = \#(\mathcal{X}_0 \cap B)$$

$$\int f \, d\mu = \sum_{x \in \mathcal{X}_0} f(x). \qquad P(X \in \mathcal{X}_0) = P_X(\mathcal{X}_0) = 1.$$

Expectation

- $$EX = \int X \, dP = \int x \, dP_X(x)$$
- $$Y = f(X) \Rightarrow EY = E f(X) = \int f \, dP_X$$

ex. $$EX = \int x \, dP_X(x) = \int x \, p(x) \, dx$$

$$EX = \int x \, dP_X(x) = \int x \, p(x) \, d\mu(x) = \sum_{x \in \mathcal{X}_0} x \, p(x).$$

- linear operator: $E(aX + bY) = aEX + bEY$
- variance: $Var(X) = E(X - EX)^2 = EX^2 - (EX)^2.$
- covariance $cov(X, Y) = E((X-EX)(Y-EY)) = EXY - EXEY$

correlation. $$cor(X, Y) = \frac{cov(X, Y)}{\sqrt{Var(X) Var(Y)}} \in [-1, 1].$$

Random Vectors.

- $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} : \mathcal{E} \to \mathbb{R}^n$.

$EX = \begin{pmatrix} EX_1 \\ \vdots \\ EX_n \end{pmatrix}$ $\quad cov(X) = \begin{pmatrix} cov(X_i, X_j) \end{pmatrix} = E(XX^T) - \mu\mu^T$.

$E(v + AX) = v + A\,EX$.

$cov(v + AX) = A\,cov(X)\,A^T$.

product measures and Independence

- $(\mathcal{X}.\mathcal{A}.\mu)$, $(\mathcal{Y}.\mathcal{B}.\nu)$ — measure spaces.

product measure: $\mu \times \nu$ on $(\mathcal{X} \times \mathcal{Y}, \mathcal{A} \vee \mathcal{B})$

$(\mu \times \nu)(A \times B) = \mu(A)\,\nu(B)$

σ-field $\mathcal{A} \vee \mathcal{B}$ — smallest σ-field containing all sets $A \times B$ with $A \in \mathcal{A}$, $B \in \mathcal{B}$.

ex. $\mu . \nu$ — lebesgue measures on $\mathbb{R}^n$ and $\mathbb{R}^m$ resp.

$\mu \times \nu$ — lebesgue measure on $\mathbb{R}^{n+m}$.

$\mu . \nu$ — countable sets $\mathcal{X}_0$ and $\mathcal{Y}_0$

$\mu \times \nu$ — counting measure on $\mathcal{X}_0 \times \mathcal{Y}_0$

Thm. (Fubini): if $f \geq 0$. then

$$\int f \, d(\mu \times \nu) = \int \Big[\int f(x.y) \, d\nu(y) \Big] \, d\mu(x)$$
$$= \int \Big[\int f(x.y) \, d\mu(x) \Big] \, d\nu(y).$$

dropping $f \geq 0$. if $\int |f| \, d(\mu \times \nu) < \infty$ Equations hold.

Taking $f = 1_S$. $\Rightarrow$ compute $(\mu \times \nu)(S)$. when S is not the cartesian product.

- Independence. $X \perp Y$ iff $P(X \in A, Y \in B) = P(X \in A) P(Y \in B)$ for all Borel sets A and B.

$Z = \binom{X}{Y}$. $P_Z(A \times B) = P_X(A) P_Y(B)$.

$\Rightarrow$ the distribution of Z is the product measure.

$P_Z = P_X \times P_Y$.

$$P(Z \in S) = \int 1_S \, d(P_X \times d_Y) = \int \Big[\int 1_S(x.y) \, dP_X(x) \Big] \, dP_Y(y)$$
$$= \int \Big[\int 1_S(x.y) \, p_X(x) \, d\mu(x) \Big] \, p_Y(y) \, d\nu(y)$$
$$= \int 1_S(x.y) \, p_X(x) \, p_Y(y) \, d(\mu \times \nu)(x.y)$$

density $p_X(x) \, p_Y(y)$.