MMAT 5340 - Probability and Stochastic Analysis

I. Probability Theory Review

Probability space

The probability space: $(\Omega, \mathcal{F}, \mathbb{P})$:

- Ω is a set.
- $\mathcal{F}$ is a space of subsets of Ω satisfying
  - $\Omega \in \mathcal{F}$,
  - $A \in \mathcal{F} \implies A^C \in \mathcal{F}$,
  - $A_n \in \mathcal{F},\ n \geq 1 \implies \cup_{n\geq 1} A_n \in \mathcal{F}$.

  The space $\mathcal{F}$ is called a σ-field, a set $A \in \mathcal{F}$ is called an event.
- A probability measure is a map $\mathbb{P} : \mathcal{F} \longrightarrow [0, 1]$ satisfying:
  - $\mathbb{P}[\Omega] = 1$,
  - If $\{A_n,\ n \geq 1\} \subset \mathcal{F}$ be such that $A_i \cap A_j = \emptyset$ for all $i \neq j$, then $\mathbb{P}[\cup_{n\geq 1} A_n] = \sum_{n\geq 1} \mathbb{P}[A_n]$.

Example 1: $\Omega = \{1, 2, \cdots, n\}$, $\mathcal{F} := \sigma(\{1\}, \cdots, \{n\})$, $\mathbb{P}[\{i\}] = \frac{1}{n}$, for each $i = 1, \cdots, n$.

Example 2: $\Omega = \mathbb{R}$, $\mathcal{F} := \mathcal{B}(\mathbb{R})$, for some density function $\rho : \mathbb{R} \longrightarrow \mathbb{R}_+$, $\mathbb{P}[(a, b)] = \int_a^b \rho(x)dx$, for all $a \leq b$.

Random variable, distribution

A *Random variable* is a map $X : \Omega \longrightarrow \mathbb{R}$ satisfying

$$X^{-1}(A) := \{\omega \in \Omega \ : X(\omega) \in A\} \in \mathcal{F}, \text{ for all } A \in \mathcal{B}(\mathbb{R})$$
$$\iff \{X \leq x\} \in \mathcal{F}, \text{ for all } x \in \mathbb{R}\,.$$

The *distribution function* of X is given by

$$F(x) \ := \ \mathbb{P}[X \leq x], \ x \in \mathbb{R}\,.$$

- *discrete random variable* X:

$$p_i = \mathbb{P}[X = x_i], \ i \in \mathbb{N}, \quad \sum_{i \in \mathbb{N}} p_i = 1.$$

- *continuous random variable* X (with continuous probability distribution), one has the *density function*

$$\rho(x) = F'(x), \ x \in \mathbb{R}\,.$$

- A distribution neither discrete nor continuous exists.

Random variable, distribution

The *expectation* $\mathbb{E}[f(X)]$ is defined by $\int_\Omega f(X(\omega))d\mathbb{P}(\omega)$ whenever the integral is well defined (measure theory needed to define it regorously).

- *discrete random variable* X:

$$\mathbb{E}[f(X)] \;:=\; \sum_{i\in\mathbb{N}} f(x_i)\mathbb{P}[X=x_i] \;=\; \sum_{i\in\mathbb{N}} f(x_i)p_i.$$

- *continuous random variable* X with density ρ:

$$\mathbb{E}[f(X)] \;:=\; \int_{\mathbb{R}} f(x)\rho(x)dx.$$

For two (*square integrable*) random variables X and Y, their *variance* and *co-variance* are defined by

$$\mathrm{Var}[X] := \mathbb{E}\Big[\big(X-\mathbb{E}[X]\big)^2\Big], \quad \mathrm{Cov}[X,Y] := \mathbb{E}\Big[\big(X-\mathbb{E}[X]\big)\big(Y-\mathbb{E}[Y]\big)\Big].$$

The *characteristic function* of X is defined by $\Phi(\theta) := \mathbb{E}[e^{i\theta X}]$.

Independence

The *events* $A_1, \cdots, A_n$ are *independent* if

$$\mathbb{P}[A_1 \cap \cdots \cap A_n] = \prod_{i=1}^{n} \mathbb{P}[A_i].$$

The *σ-fields* $\mathcal{F}_1, \cdots, \mathcal{F}_n$ are *independent* if

$$\mathbb{P}[A_1 \cap \cdots \cap A_n] = \prod_{i=1}^{n} \mathbb{P}[A_i], \text{ for all } A_1 \in \mathcal{F}_1, \cdots, A_n \in \mathcal{F}_n.$$

The *random variables* $X_1, \cdots, X_n$ are *independent* if

$$\sigma(X_1), \cdots, \sigma(X_n) \text{ are independent.}$$

Remarks: How is $\sigma(X_1)$ defined? What does it mean if we say X_1 is independent of $\mathcal{F}_2$? Concrete examples ...

Independence

Lemma

If $X_1, \cdots, X_n$ are independent, f_i are measurable functions. Then $f_1(X_1), \cdots, f_n(X_n)$ are independent.

Lemma

If $X_1, \cdots, X_n$ are independent, then

$$\mathbb{E}[f_1(X_1)\cdots f_n(X_n)] \;=\; \mathbb{E}[f_1(X_1)]\cdots\mathbb{E}[f_n(X_n)].$$

Consequently,

$$\mathrm{Var}[X_1 + \cdots + X_n] = \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_n].$$

$$\mathrm{Cov}[f_i(X_i), f_j(X_j)] = 0,\ \ i \neq j.$$

Remarks: The inverse may not be correct. Concrete examples ...

Convergence of random variables

Almost sure convergence: We say X_n converges almost surely to X if $\mathbb{P}[\lim_{n\to\infty} X_n = X] = 1$.

Convergence in probability: We say X_n converges to X in probability if, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}[|X_n - X| \geq \varepsilon] \;=\; 0.$$

Convergence in distribution: We say X_n converges to X in distribution if, for any bounded continuous function f,

$$\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)].$$

Convergence in L^p $(p \geq 1)$ space Let $X_n, n \geq 1$ satisfy $\mathbb{E}[|X_n|^p] < \infty$, we say X_n converges to X in L^p space if

$$\lim_{n\to\infty} \mathbb{E}\big[|X_n - X|^p\big] \;=\; 0.$$

Convergence of random variables

Lemma (Relations between the different notions of the convergence)

One has

$$\text{Cvg a.s.} \implies \text{Cvg in prob.} \implies \text{Cvg in dist.},$$

$$\text{Cvg in } L^p \implies \text{Cvg in prob.}$$

$$\text{Cvg in prob.} \implies \text{Cvg a.s. along a subsequence.}$$

Lemma (Monotone convergence theorem)

Assume that $0 \le X_n \le X_{n+1}$ *for all* $n \ge 1$, *then*

$$\mathbb{E}\big[\lim_{n\to\infty} X_n\big] = \lim_{n\to\infty} \mathbb{E}[X_n].$$

Remark: In practice, we may have $X_n := f_n(X)$ for a sequence $(f_n)_{n\ge 1}$ satisfying $0 \le f_1 \le f_2 \le \cdots$.

Limit theorems

Theorem (Law of Large Number)

Assume that $(X_n)_{n\geq 1}$ *is an i.i.d. sequence with the same distribution of* X *and such that* $\mathbb{E}[|X|] < \infty$*. Then*

$$\lim_{n\to\infty} \overline{X}_n \;:=\; \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k \;=\; \mathbb{E}[X], \text{ a.s.}$$

Theorem (Central Limit Theorem)

Assume that $(X_n)_{n\geq 1}$ *is an i.i.d. sequence with the same distribution of* X *and such that* $\mathbb{E}[|X|^2] < \infty$*. Then*

$$\frac{\sqrt{n}(\overline{X}_n - \mathbb{E}[X])}{\sqrt{\mathrm{Var}[X]}} \text{ converges in distribution to } N(0,1).$$

Inequalities

Lemma (Jensen inequality)

Let X be a r.v., ϕ be a convex function. Assume that $\mathbb{E}[|X|] < \infty$ and $\mathbb{E}[|\phi(X)|] < \infty$. Then

$$\phi(\mathbb{E}[X]) \ \leq \ \mathbb{E}\big[\phi(X)\big].$$

Lemma (Chebychev inequality)

Let X be a r.v., $f : \mathbb{R} \to \mathbb{R}_+$ be an increasing function. Assume that $\mathbb{E}[f(X)] < \infty$ and $f(a) > 0$. Then

$$\mathbb{P}[X \geq a] \ \leq \ \frac{\mathbb{E}[f(X)]}{f(a)]}.$$

Inequalities

Lemma (Cauchy-Schwarz inequality)

Let X *and* Y *be two r.v. Assume that* $\mathbb{E}[|X|^2] < \infty$ *and* $\mathbb{E}[|Y|^2] < \infty$*. Then*

$$\mathbb{E}[XY] \leq \sqrt{\mathbb{E}[|X|^2]\mathbb{E}[|Y|^2]}.$$

Conditional expectation

Theorem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$, X a random variable. Assume that $\mathbb{E}[|X|] < \infty$. Then there exists a random variable Z satisfying the following:

- $\mathbb{E}[|Z|] < \infty$.
- Z *is* $\mathcal{G}$*-measurable.*
- $\mathbb{E}[XY] = \mathbb{E}[ZY]$, *for all* $\mathcal{G}$*-measurable bounded random variables* Y.

Moreover, the random Z is unique in the sense of almost sure.

We say such a random variable Z is the *conditional expectation* of X knowing $\mathcal{G}$, and denote

$$\mathbb{E}[X|\mathcal{G}] := Z.$$

Conditional expectation, examples

When $\mathcal{G} = \sigma(Y_1, \cdots, Y_n)$, for $Y = (Y_1, \cdots .Y_n)$, we also write

$$\mathbb{E}[f(X)|Y_1, \cdots, Y_n] \;:=\; \mathbb{E}[f(X)|\mathcal{G}].$$

In this case, there exists a *measurable function* $g : \mathbb{R}^n \to \mathbb{R}$ such that $\mathbb{E}[X|Y] = g(Y)$. To compute $\mathbb{E}[f(X)|Y]$, it is enough to compute the function:

$$g(y) \;:=\; \mathbb{E}[f(X)|Y = y], \text{ for all } y \in \mathbb{R}^n .$$

Discrete case: $\mathbb{P}[X = x_i, Y = y_j] = p_{i,j}$ with $\sum_{i,j} p_{i,j} = 1$. Then

$$\mathbb{E}[f(X)|Y = y_j] = \frac{\mathbb{E}[f(X)\mathbf{1}_{Y=y_j}]}{\mathbb{E}[\mathbf{1}_{Y=y_j}]} = \frac{\sum_{i\in\mathbb{N}} f(x_i)p_{i,j}}{\sum_{i\in\mathbb{N}} p_{i,j}}.$$

Continuous case: Let $\rho(x, y)$ be the density function of (X, Y). Then

$$\mathbb{E}[f(X)|Y = y] = \frac{\int_{\mathbb{R}} f(x)\rho(x, y)dx}{\int_{\mathbb{R}} \rho(x, y)dx}.$$

Conditional distribution

Let A be an event such that $\mathbb{P}[A] > 0$, The conditional probability knowing event A is given by

$$\mathbb{P}\left[B \middle| A\right] \;:=\; \frac{\mathbb{P}\left[B \cap A\right]}{\mathbb{P}[A]},$$

and the conditional expectation is given by

$$\mathbb{E}\left[f(X) \middle| A\right] \;:=\; \frac{\mathbb{E}\left[f(X)1_A\right]}{\mathbb{P}[A]}.$$

Conditional expectation

Lemma

Let X and Y be two r.v. such that $\mathbb{E}[|X|] < \infty$ and $\mathbb{E}[|Y|] < \infty$, a, b be two real numbers. Then

$$\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}].$$

Lemma

Let X, Y be r.v. such that $\mathbb{E}[|X|] < \infty$, Y is $\mathcal{G}$-measurable and $\mathbb{E}[|XY|] < \infty$, then

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X], \qquad \mathbb{E}[\mathbb{E}[X|\mathcal{G}]Y] = \mathbb{E}[XY],$$

and

$$\mathbb{E}[XY|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]Y.$$

If X is independent of $\mathcal{G}$, then

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X].$$

Conditional expectation

Lemma

Let X be a random variable, φ be a convex function. Then

$$\mathbb{E}[\varphi(X)|\mathcal{G}] \geq \varphi(\mathbb{E}[X|\mathcal{G}]), \text{ a.s.}$$

Lemma (Monotone convergence theorem)

Let $(X_n, n \geq 1)$ be a sequence of integrable random variable such that $0 \leq X_n \leq X_{n+1}$, a.s. Then

$$\lim_{n\to\infty} \mathbb{E}\big[X_n\big|\mathcal{G}\big] = \mathbb{E}\big[\lim_{n\to\infty} X_n\big|\mathcal{G}\big].$$

Conditional expectation

Lemma

Let X be an integrable random variable, and $\mathcal{G} := \{\emptyset, \Omega\}$. Then

$$\mathbb{E}\big[X\big|\mathcal{G}\big] = \mathbb{E}\big[X\big].$$

Lemma

Let X be an integrable random variable, and $\mathcal{G}_1 \subset \mathcal{G}_2$ be two sub-σ-filed of $\mathcal{F}$. Then

$$\mathbb{E}\big[\mathbb{E}\big[X\big|\mathcal{G}_2\big]\big|\mathcal{G}_1\big] \;=\; \mathbb{E}\big[X\big|\mathcal{G}_1\big].$$