Introduction to Stochastic Calculus

Yuewu Xu

School of Business

Fordham University

Fall, 2022

Important Note: This material is for MSQF students use only. Do NOT distribute outside of class!

0. Conditional Expectations – A Review

Probability Space (Ω, F, P)

- Ω : sample space
- F : σ-algebra (also called σ-field). It is a collection of subsets of Ω satisfying the "usual conditions". See below.
- P : the probability measure. It is a function mapping the sets in F into the unit interval, satisfying the "usual conditions".

Example: (Ω, F, P) for

- $\Omega = \mathbb{R}$
- $F = B(\mathbb{R})$, which includes all intervals. See discussion below.
- P defined by: for $A \in F$,

$$P(A) = \int_A \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Sigma-algebra (aka, Sigma Field)

Definition: F is a sigma-algebra if it satisfies

- $\Omega \in F$
- If $A \in F$, then its complement $A^c \in F$
- If $A_n \in F$ for $n = 1, 2, ...$, then $\cup_n A \in F$

Important: The Borel-algebra $F = B(\mathbb{R})$ on the real line. It is σ-algebra generated by the open sets in R. It exists – why?

Definition: A probability is a function defined on sets in F satisfing

- $P(A) \geq 0$, for $A \in F$
- $P(\Omega) = 1$
- For sequence A_n $(n = 1, 2, ...)$ of disjoint events: $P(\cup_n A_n) = \sum_n P(A_n)$.

Example [Coin toss] The simplist non-trivial σ-field

- $\Omega = \{H, T\}$
- $F = \{\phi, \{H\}, \{T\}, \Omega\}$
- $P(\phi) = 0, P(\Omega) = 1, P(\{H\}) = p, P(\{T\}) = 1 - p$

Define map X, which maps Ω to the real line

$$X(\omega) = \begin{cases} 1 & if\ \omega = H \\ 0 & if\ \omega = T \end{cases}$$

Random variable: "measurable" maps that maps Ω to the real line R.

- Important!!! Any random variable X induces a measure μ_X on $R : \mu_X(B) = P(X(\omega) \in B)$. It is called the distribution of X.

The usual wording "Let (Ω, F, P) be a probability space. Let X be a (standard normal) random variable..."

Why people seldom mention how the random variable $X(\omega)$ is defined?

There is no "randomness" or "uncertainty" in a random variable: it is just a function defined on Ω that is measurable. Only when viewed at a higher level and with P, objects like X can be used to "model" uncertainty.

Expectations of Random Variables: When we have a random variable X defined on a probability space (Ω, F, P), we can define the Expectations of random variable X, which is denoted by

$$\mathbb{E}X = \int_{\Omega} X(\omega) dP\left(\omega\right).$$

In the case that the random variable X with density $f(x)$, it's expectation is defined by (if exist)

$$\mathbb{E}X = \int x f(x) dx$$

and the discrete case is

$$\mathbb{E}X = \sum x_i p_i,$$

where $p_i = P(X = x_i)$.

The Concept of Conditional Expectations

The Elementary Definition. In elementary probability, the conditional expectation $\mathbb{E}[X|Y]$ can be calculated by first determining the conditional density of $X|Y = y$, and then computing the expectation of that conditional density.

Example: Assume that the joint density of X and Y : $f_{X,Y}(x, y)$.

The conditional density of Y given $X = x$ is defined by

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)},$$

where

$$f_X(x) = \int f_{X,Y}(x, y) dy$$

is the marginal density of X. The conditional expectation of $Y|X = x$ is

$$\mathbb{E}\left[Y|X = x\right] = \int y f_{Y|X}(y|x) dy$$

Modern Definition (Mesure Theoretic). Let (Ω, F, P) be a probability space. Let G be a sub-σ-algebra of F, and let X be a random variable that either nonnegative or integrable. The conditional expectation of X given G, denoted by $\mathbb{E}[X|G]$, is any random variable Y that satisfies

(1). (Measurability) Y is G measurable

(2). (Partial averaging)

$$\mathbb{E}[Y \cdot 1_A] = \mathbb{E}[X \cdot 1_A] \text{ for all } A \in G.$$

If G is the σ-algebra generated by some other random variable W, we generally write $\mathbb{E}[X|W]$ rather than $\mathbb{E}[X|\sigma(W)]$.

Natural Question: How to reconcile it with the traditional definition of Conditional Expectation? We know that $\mathbb{E}[X|Y]$ is a random variable (modern definition), but the traditional conditional expectation is a function depending on x.

Example. Let $\Omega = \{a, b, c, d\}$. The σ-algebra F is the collection of all subsets of Ω, i.e., the sets in F are

Ω,

$\{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\}$,

$\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}$

$\{a\}, \{b\}, \{c\}, \{d\}$,

ϕ.

We define probability measure by

$P\{a\} = \frac{1}{6}, P\{b\} = \frac{1}{3}, P\{c\} = \frac{1}{4}, P\{d\} = \frac{1}{4}$.

Define two random variables X and Y :

$X(a) = 1, X(b) = 1, X(c) = -1, X(d) = -1$

$Y(a) = 1, Y(b) = -1, Y(c) = 1, Y(d) = -1$

Set $Z = X + Y$

Questions:

(i). What are the sets of $\sigma(X)$.

(ii). Determine $\mathbb{E}[Y|X]$. verify the partial-averaging property

(iii). Determine $\mathbb{E}[Z|X]$. Verify the partial-averaging property

(iv). Compute $\mathbb{E}[Z|X] - \mathbb{E}[Y|X]$. Why you get X?

There are two approaches: traditional and measure-theoretic.

Note: The purpose of this exercise is to show that we will get the same result obtained using the "measure-theoretic" approach.

The Traditional Approach

Using this approach, we need to first find the joint distribution for X and Y. Then computing the conditional distribution of Y given $X = x$ for values of $x = 1$ and -1 (these are the only two values the random variable X takes). Once we have the conditional distribution of Y given $X = x$, we can then compute the conditional expectation $\mathbb{E}[Y|X = x]$.

The joint distribution of X and Y is defined by

$P(X=1, Y=1) = P\{a\} = \frac{1}{6}$

$P(X=1, Y=-1) = P\{b\} = \frac{1}{3}$

$P(X=-1, Y=1) = P\{c\} = \frac{1}{4}$

$P(X=-1, Y=-1) = P\{d\} = \frac{1}{4}$

The conditional distribution of $Y|X=x$ is computed as

$P[Y=1|X=1] = \frac{P[Y=1,X=1]}{P[X=1]} = \frac{\frac{1}{6}}{\frac{1}{6}+\frac{1}{3}} = \frac{1}{3}$

$P[Y=-1|X=1] = \frac{P[Y=-1,X=1]}{P[X=1]} = \frac{\frac{1}{3}}{\frac{1}{6}+\frac{1}{3}} = \frac{2}{3}$

$P[Y=1|X=-1] = \frac{P[Y=1,X=-1]}{P[X=-1]} = \frac{\frac{1}{4}}{\frac{1}{4}+\frac{1}{4}} = \frac{1}{2}$

$P[Y=-1|X=-1] = \frac{P[Y=-1,X=-1]}{P[X=1]} = \frac{\frac{1}{4}}{\frac{1}{4}+\frac{1}{4}} = \frac{1}{2}$

Note that above conditional probabilities sum to 1, as it should.

With the above conditional probability for $Y|X=x$, we can readily compute the the conditional expectation $\mathbb{E}[Y|X=x]$:

$\mathbb{E}[Y|X=1] = 1 * P(Y=1|X=1) + (-1) * P(Y=-1|X=1)$

$= 1 * \frac{1}{3} + (-1) * \frac{2}{3} = -\frac{1}{3}.$

and

$\mathbb{E}[Y|X=-1] = 1 * P(Y=1|X=-1) + (-1) * P(Y=-1|X=-1)$

$= 1 * \frac{1}{2} + (-1) * \frac{1}{2} = 0.$

To summarize, the conditional expectation computed from the "traditional approach" gives

$$\mathbb{E}[Y|X=x] = \begin{cases} -\frac{1}{3} & if\ x = 1 \\ 0 & if\ x = -1 \end{cases} \stackrel{define}{=} g(x).$$

You need to understand that the above conditional expectation is a function of the "dummy" variable x. It agrees with our "measure-theoretic" result:

$$V(\omega) = g(x)|_{x=X(\omega)} = \begin{cases} -\frac{1}{3} & if\ X(\omega) = 1 \\ 0 & if\ X(\omega) = -1 \end{cases}$$

This is just

$$\begin{aligned} V(a) &= -\frac{1}{3} \\ V(b) &= -\frac{1}{3} \\ V(c) &= 0 \\ V(d) &= 0 \end{aligned}$$

Mesure-Theoretic Approach

(i). Recall that $\sigma(X)$ is the σ-algebra generated by all the reverse images of all the Borell sets on R. So we first need to figure out all the sets generated by $X^{-1}(B)$, where B is a Borell set (all open sets is enough). This is task is made easy because the random variable X takes only values of 1 and -1. Any interval B that includes only the number 1 but NOT -1, will produce the set $\{a, b\}$. Similarly, Any interval B that includes only the number -1 but NOT 1, will produce the set $\{c, d\}$. Any interval B that contains both -1 and 1 will product $X^{-1}(B) = \{a, b, c, d\} = \Omega$. And of course, the empty set ϕ is always in the $\sigma(X)$.

So we obtain $\sigma(X) = \{\phi, \{a, b\}, \{c, d\}, \Omega\}$, which contains 4 sets.

(ii) We need to compute $\mathbb{E}[Y|X]$, which is short-hand for $\mathbb{E}[Y|G]$, where $G = \sigma(X)$. To simplify notation, let's denote by $V = \mathbb{E}[Y|X]$.

To this end, we use the measure-theoretic definition. First, the conditional expectation V is a random variable defined on Ω. So as long as we can determine the 4 numbers $V(a), V(b), V(c), V(d)$ we are fine.

For V to be the conditional expectation $\mathbb{E}(Y|X)$, the first condition the definition of Conditional Expectation says that V need to be $G = \sigma(X)$ measurable. This put a restriction on the values $V(a), V(b), V(c), V(d)$: it must hold that $V(a) = V(b), V(c) = V(d)$. Why? because if it is not true, then using the argument that we used to obtain $\sigma(X)$, we will find such V would not be $\sigma(X)$ measurable.

Hence to figure out V, we only need to find two numbers: $\alpha = V(a) = V(b)$, and $\beta = V(c) = V(d)$.

Thse two numbers are determined by the second requirement for V to be the conditional expectation $\mathbb{E}[Y|X]$

$$\mathbb{E}\left[V \cdot 1_A\right] = \mathbb{E}\left[Y \cdot 1_A\right] \quad \text{for all } A \in \sigma(X).$$

The above requirement needs to hold for all $A \in \sigma(X)$. So we can set $A = \{c, d\}$ to obtain

$$\mathbb{E}\left[V \cdot 1_{\{a,b\}}\right] = \mathbb{E}\left[Y \cdot 1_{\{a,b\}}\right]$$

Since the sample space is discrete, this is just

$$\begin{aligned}
V(a) * P\{a\} + V(b) * P\{b\} &= Y(a) * P\{a\} + Y(b) * P\{b\} \\
\alpha[P\{a\} + P(b)] &= Y(a) * P\{a\} + Y(b) * P\{b\} \\
\alpha &= \frac{Y(a) * P\{a\} + Y(b) * P\{b\}}{[P\{a\} + P(b)]} \\
&= \frac{1 * \frac{1}{6} - 1 * \frac{1}{3}}{\frac{1}{6} + \frac{1}{3}} = \frac{-\frac{1}{6}}{\frac{1}{2}} = -\frac{1}{3}.
\end{aligned}$$

Similarly, we can set $A = \{c, d\}$ to obtain

$$\mathbb{E}\left[V \cdot 1_{\{c,d\}}\right] = \mathbb{E}\left[Y \cdot 1_{\{c,d\}}\right]$$

Since the sample space is discrete, this is just

$$\begin{aligned} V(c) * P\{c\} + V(d) * P\{d\} &= Y(c) * P\{c\} + Y(d) * P\{d\} \\ \beta[P\{c\} + P(d)] &= Y(c) * P\{c\} + Y(d) * P\{d\} \\ \beta &= \frac{Y(c) * P\{c\} + Y(d) * P\{d\}}{[P\{c\} + P(d)]} \\ &= \frac{1 * \frac{1}{4} - 1 * \frac{1}{4}}{\frac{1}{4} + \frac{1}{4}} = \frac{0}{\frac{1}{2}} = 0. \end{aligned}$$

Hence the conditional expectation

$\mathbb{E}[Y|X] = V$, defined by $V(a) = -\frac{1}{6}, V(b) = -\frac{1}{6}, V(c) = 0, V(d) = 0.$

Remark: You need to make sure that you can check that the V defined above is indeed $\sigma(X)$ measurable.

(iiI). To compute $\mathbb{E}[Z|X]$, we first compute random variable Z.

$Z(a) = X(a) + Y(a) = 2,$

$Z(b) = X(b) + Y(b) = 0,$

$Z(c) = X(c) + Y(c) = 0,$

$Z(d) = X(d) + Y(d) = -2.$

Again, denote the conditional expection $\mathbb{E}[Z|X] = U$. Then an argument similar to that in (ii) will show that U has to satisfy

$U(a) = U(b) = \gamma$

$U(c) = U(d) = \delta.$

The exact values for γ and δ can be determined by the averaging condition:

$$\mathbb{E}\left[U \cdot 1_A\right] = \mathbb{E}\left[Z \cdot 1_A\right] \ \textit{for all } A \in \sigma(X).$$

Setting $A = \{a, b\}$ yields:

$$\mathbb{E}\left[U \cdot 1_{\{a,b\}}\right] = \mathbb{E}\left[Z \cdot 1_{\{a,b\}}\right]$$

Since the sample space is discrete, this is just

$$\begin{aligned}
U(a) * P\{a\} + U(b) * U\{b\} &= Z(a) * P\{a\} + Z(b) * P\{b\} \\
\gamma[P\{a\} + P(b)] &= Z(a) * P\{a\} + Z(b) * P\{b\} \\
\alpha &= \frac{Z(a) * P\{a\} + Z(b) * P\{b\}}{[P\{a\} + P(b)]} \\
&= \frac{2 * \frac{1}{6} + 0 * \frac{1}{3}}{\frac{1}{6} + \frac{1}{3}} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}.
\end{aligned}$$

Setting $A = \{c, d\}$ yields:

$$\mathbb{E}\left[U \cdot 1_A\right] = \mathbb{E}\left[Z \cdot 1_A\right] \ \textit{for all } A \in \sigma(X).$$

Since the sample space is discrete, this is just

$$\begin{aligned}
U(c) * P\{c\} + U(d) * U\{d\} &= Z(c) * P\{c\} + Z(d) * P\{d\} \\
\delta[P\{c\} + P(d)] &= Z(c) * P\{c\} + Z(d) * P\{d\} \\
\gamma &= \frac{Z(c) * P\{c\} + Z(d) * P\{d\}}{[P\{c\} + P(d)]} \\
&= \frac{0 * \frac{1}{4} - 2 * \frac{1}{4}}{\frac{1}{4} + \frac{1}{4}} = \frac{-\frac{1}{2}}{\frac{1}{2}} = -1.
\end{aligned}$$

So we find that U is defined by

$U(a) = \frac{2}{3}$

$U(b) = \frac{2}{3}$

$U(c) = -1$

$U(d) = -1$

(iv). $W = \mathbb{E}[Z|X] - \mathbb{E}[Y|X] = U - V$, which is just

$W(a) = \frac{2}{3} - (-\frac{1}{3}) = 1$

$W(b) = \frac{2}{3} - (-\frac{1}{3}) = 1$

$W(c) = -1 - 0 = -1$

$W(d) = -1 - 0 = -1.$

Hence W is indentical to X. This is a result verifying the "linearity" of conditional expectations.

Properties of Conditional Expectations

Assume that G is a sub-σ-algebra of F. If X and Y are integrable random variables. Then

(i). (Linearity) $\mathbb{E}[(aX + bY)|G] = a\mathbb{E}[X|G] + b\mathbb{E}[Y|G]$ for constants a and b.

(ii). (Taking out what is known) If in addition XY are integrable and X is G-measurable, then $\mathbb{E}[XY|G] = X\mathbb{E}[Y|G]$.

(iii). (Iterated conditioning) If H is a sub-σ-algebra of G, then $\mathbb{E}[\mathbb{E}[X|G]|H] = \mathbb{E}[X|H]$.

(iv). (Independence)If X is independent of G, $\mathbb{E}[X|G] = \mathbb{E}X$.

(v). (Conditional Jensen Inequality) If $\varphi(x)$ is a convex function, then $\mathbb{E}[\varphi(X)|G] \geq \varphi(\mathbb{E}[X|G])$.

Filtration

Definition. Let Ω be a nonempty set. Let T be a fixed positive number, and assume that for each $t \in [0, T]$

there is a σ-algebra $F(t)$. Assume further that is $s \le t$, then every set in $F(s)$ is also in $F(t)$. Then we call the collection of σ-algebras $F(t), 0 \le t \le T$, a filtration.

Example:

(i). The sequence $F(n)$ of σ-algebras generated by a sequence of independent random variables is a filtration: $F(n) = \sigma(X_1, X_2, ...X_n)$.

(ii). Filtration generated by Brownian Motion $W(t) : F(t) := \sigma(X_s : 0 \le s \le t)$.

Martingale

Definition. Let (Ω, F, P) be a probability space, let $F(t)$ $(0 \le t \le T)$ be filtration of sub-σ-algebra of F. An adapted stochastic processes $M(t)$ $(0 \le t \le T)$ is

(i) a martingale if $\mathbb{E}[M(t)|F(s)] = M(s)$ for all $0 \le s \le t \le T$.

(ii) a submartingale if $\mathbb{E}[M(t)|F(s)] \ge M(s)$ for all $0 \le s \le t \le T$.

(iii) a supermartingale if $\mathbb{E}[M(t)|F(s)] \le M(s)$ for all $0 \le s \le t \le T$.

Discussion:

Sample path.

Example: The Brownian Motion $W(t)$ can be constructed as

$$W(t) = \frac{t}{\sqrt{\pi}} X_0 + \sqrt{\frac{2}{\pi}} \sum_{m=1}^{\infty} \frac{\sin(mt)}{m} X_m,\ t \in [0, 2\pi]$$

where X_m $(m = 0, 1, 2, ...)$are *i.d.d.* standard normal random variables.