MA400. September Introductory Course
(Financial Mathematics, Risk & Stochastics)
P4. Expectation

1. In what follows, we assume that an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is fixed.

Preliminary considerations

2. Consider the toss of a coin that lands heads with probability $p \in (0, 1)$ and tails with probability $1 - p$. Also, denote by X the random variable that takes the value 1 if tails are observed and the value 0 if heads are observed. Now, consider two parties, say A and B, that bet on the coin's toss: once the coin lands, party A will pay £X to party B (i.e., A will pay B £1 if tails occur and £0 if heads occur). What is the value $\mathbb{E}[X]$ of this game? In other words, how much money $\mathbb{E}[X]$ should B pay to A in advance for both parties to feel that they engage in a fair game? Intuition suggest that

$$\mathbb{E}[X] = 1 \times (1 - p) + 0 \times p = 1 - p.$$

The number $\mathbb{E}[X]$ is the expectation of X.

3. Generalising the example above, the expectation $\mathbb{E}[\mathbf{1}_A]$ of the random variable

$$\mathbf{1}_A = \begin{cases} 1, & \text{if } \omega \in A, \\ 0, & \text{if } \omega \notin A, \end{cases}$$

where A is an event in $\mathcal{F}$, is given by

$$\mathbb{E}[\mathbf{1}_A] = 1 \times \mathbb{P}(A) + 0 \times \mathbb{P}(A^c) = \mathbb{P}(A).$$

this idea and the requirement that expectation should be a linear operator provide the starting point of this chapter's theory.

Definitions

4. *Definition.* We say that X is a *simple random variable* if there exist distinct real numbers $x_1, x_2, \ldots, x_n$ and a measurable partition $A_1, A_2, \ldots, A_n$ of the sample space Ω (i.e., events $A_1, A_2, \ldots, A_n \in \mathcal{F}$ satisfying $A_i \cap A_j = \emptyset$, for $i \neq j$, and $\bigcup_{i=1}^n A_i = \Omega$) such that

$$X(\omega) = \sum_{i=1}^n x_i 1_{A_i}(\omega) \quad \text{for all } \omega \in \Omega. \tag{1}$$

5. *Definition.* The *expectation* of the simple random variable X given by (1) is defined by

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \mathbb{P}(A_i).$$

6. *Definition.* Suppose that X is a $\big([0,\infty], \mathcal{B}([0,\infty])\big)$-valued random variable. The *expectation* of X is defined by

$$\mathbb{E}[X] = \sup\big\{\mathbb{E}[Y] \mid Y \text{ is a simple random variable with } 0 \leq Y \leq X\big\}.$$

Note that $\mathbb{E}[X] \geq 0$, but we may have $\mathbb{E}[X] = \infty$ even if $X(\omega) < \infty$ for all $\omega \in \Omega$.

7. *Definition.* Given a real-valued random variable X, define

$$X^+ = \max(0, X) \quad \text{and} \quad X^- = -\min(0, X)$$

and observe that X^+, X^- are positive random variables such that $X = X^+ - X^-$ and $|X| = X^+ + X^-$.

A random variable X has *finite expectation* (is *integrable*) if both $\mathbb{E}[X^+] < \infty$ and $\mathbb{E}[X^-] < \infty$. In this case, the expectation of X is defined by

$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-].$$

We often write $\int_\Omega X(\omega)\,\mathbb{P}(d\omega)$ or $\int_\Omega X\,d\mathbb{P}$ instead of $\mathbb{E}[X]$.

8. *Definition.* We denote by $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, or just $\mathcal{L}^1$ if there is no ambiguity, the set of all integrable random variables.

For $1 \leq p < \infty$, we denote by $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$, or just $\mathcal{L}^p$ if there is no ambiguity, the set of all random variables X such that $|X|^p \in \mathcal{L}^1$.

9. For every positive random variable X, there exists a sequence (X_n) of positive simple random variables such that X_n increases to X as n increases to infinity. An *example* of such a sequence is given by

$$X_n(\omega) = \begin{cases} k2^{-n}, & \text{if } k2^{-n} \leq X(\omega) < (k+1)2^{-n} \text{ and } 0 \leq k \leq n2^n - 1, \\ n, & \text{if } X(\omega) \geq n. \end{cases}$$

Properties of expectation

10. We say that a property holds $\mathbb{P}$-a.s. if it is true for all ω in a set of probability 1. For example, we say that $X = Y$, $\mathbb{P}$-a.s., if

$$\mathbb{P}(X = Y) \equiv \mathbb{P}\big(\{\omega \in \Omega \mid X(\omega) = Y(\omega)\}\big) = 1.$$

Similarly, we say that a sequence of random variables (X_n) converges to a random variable X, $\mathbb{P}$-a.s., if

$$\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) \equiv \mathbb{P}\left(\left\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1.$$

11. The following results hold true.

(i) $\mathcal{L}^1$ is a vector space, i.e.,

$$X, Y \in \mathcal{L}^1 \text{ and } a, b \in \mathbb{R} \quad \Rightarrow \quad aX + bY \in \mathcal{L}^1.$$

Expectation is a positive, linear map on $\mathcal{L}^1$, i.e.,

$$\begin{aligned} X \geq 0 \quad &\Rightarrow \quad \mathbb{E}[X] \geq 0, \\ X, Y \in \mathcal{L}^1 \text{ and } a, b \in \mathbb{R} \quad &\Rightarrow \quad \mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]. \end{aligned}$$

(ii) If $X = Y$, $\mathbb{P}$-a.s., then $\mathbb{E}[X] = \mathbb{E}[Y]$.

(iii) (*Monotone convergence theorem*) If (X_n) is an increasing sequence of positive random variables (i.e., $0 \leq X_1 \leq X_2 \leq \cdots \leq X_n \leq \cdots$) such that $\lim_{n\to\infty} X_n = X$, $\mathbb{P}$-a.s., for some random variable X, then

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

Note that we may have $\mathbb{E}[X] = \infty$ here.

(iv) (*Dominated convergence theorem*) If (X_n) is a sequence of random variables that converges to a random variable X, $\mathbb{P}$-a.s., and is such that $|X_n| \leq Y$, $\mathbb{P}$-a.s., for all $n \geq 1$, for some $Y \in \mathcal{L}^1$, then

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

(v) (*Fatou's lemma*) If (X_n) is a sequence of random variables such that $X_n \geq Y$, $\mathbb{P}$-a.s., for all $n \geq 1$, for some $Y \in \mathcal{L}^1$, then

$$\mathbb{E}\left[\liminf_{n\to\infty} X_n\right] \leq \liminf_{n\to\infty} \mathbb{E}[X_n].$$

Similarly, if (X_n) is a sequence of random variables such that $X_n \leq Y$, $\mathbb{P}$-a.s., for all $n \geq 1$, for some $Y \in \mathcal{L}^1$, then

$$\mathbb{E}\left[\limsup_{n\to\infty} X_n\right] \geq \limsup_{n\to\infty} \mathbb{E}[X_n].$$

(vi) (*Jensen's inequality*) Given a random variable X and a convex function $g : \mathbb{R} \to \mathbb{R}$, such that $X, g(X) \in \mathcal{L}^1$,

$$\mathbb{E}\big[g(X)\big] \geq g\big(\mathbb{E}[X]\big).$$

(vii) If X and Y are independent random variables, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

12. *Example*. Suppose that $\Omega = (0,1)$, $\mathcal{F} = \mathcal{B}\big((0,1)\big)$ and $\mathbb{P}$ is the Lebesgue measure on $\big((0,1), \mathcal{B}\big((0,1)\big)\big)$. Consider the sequence $(X_n,\ n \geq 1)$ of the random variables given by

$$X_n(\omega) = (n+1)\mathbf{1}_{\left(\frac{1}{2}, \frac{1}{2}+\frac{1}{n+1}\right)}(\omega) \equiv \begin{cases} n+1, & \text{if } \omega \in \left(\frac{1}{2}, \frac{1}{2}+\frac{1}{n+1}\right), \\ 0, & \text{otherwise.} \end{cases}$$

Given any $n \geq 1$, we calculate

$$\begin{aligned} \mathbb{E}\left[X_n\right] &= (n+1)\mathbb{P}\big(\left(\tfrac{1}{2}, \tfrac{1}{2}+\tfrac{1}{n+1}\right)\big) + 0\mathbb{P}\big(\left(0, \tfrac{1}{2}\right] \cup \left[\tfrac{1}{2}+\tfrac{1}{n+1}, 1\right)\big) \\ &= 1. \end{aligned}$$

Moreover, we can see that

$$\lim_{n\to\infty} X_n(\omega) = 0 \quad \text{for all } \omega \in \Omega.$$

These observations imply that

$$\lim_{n\to\infty} \mathbb{E}\left[X_n\right] = 1 > 0 = \mathbb{E}\left[\lim_{n\to\infty} X_n(\omega)\right].$$

Note that the sequence of random variables considered in this example does not satisfy the assumptions of either the monotone convergence theorem or the dominated convergence theorem.

Also, this example shows that the inequalities in Fatou's lemma can be strict.

13. *Lemma*. The expectation of a discrete random variable X is given by

$$\mathbb{E}[X] = \sum_{x_i} x_i \mathbb{P}(X = x_i).$$

The expectation of a continuous random variable X with probability density function f is given by

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\, dx.$$

14. *Example*. Suppose that X has the Poisson distribution. We can calculate the mean of X as follows:

$$\begin{aligned} \mathbb{E}[X] &= \sum_{n=0}^{\infty} n e^{-\lambda} \frac{\lambda^n}{n!} \\ &= \lambda e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} \\ &= \lambda. \end{aligned}$$

15. *Example*. Suppose that X is a Gaussian random variable with mean m and variance σ^2. We can calculate the mean of X as follows:

$$
\begin{aligned}
\mathbb{E}[X] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} a \exp\left(-\frac{(a-m)^2}{2\sigma^2}\right) da \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (a-m) \exp\left(-\frac{(a-m)^2}{2\sigma^2}\right) da \\
&\quad + m\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{(a-m)^2}{2\sigma^2}\right) da \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} a \exp\left(-\frac{a^2}{2\sigma^2}\right) da \\
&\quad + m\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{(a-m)^2}{2\sigma^2}\right) da \\
&= 0 + m \times 1 \\
&= m.
\end{aligned}
$$

Here, we have used the fact that the first integral is 0 because the integrand is an odd function, and the fact that the second integral is equal to one because the integrand is a probability density function.