November 7, 2021

EXPECTATION II

ABSTRACT. We will define the expectation of a continuous random variable and discuss basic properties.

1. Introduction

Let X be a continuous real-valued random variable with pdf f. Suppose that

$$\int |x| f(x) dx < \infty.$$

We define the **expectation** or **mean** of X via

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

Remark 1. *The law of large numbers as stated earlier for discrete random variable is still valid.*

Exercise 1.1. *Let X be uniformly distributed on the unit interval. Find $\mathbb{E}X$.*

Exercise 1.2. *Let X be exponential with parameter (mean) $\mu > 0$. Show that $\mathbb{E}X = \mu$.*

Exercise 1.3. *Let $X \sim N(\mu, \sigma^2)$. Show that $\mathbb{E}X = \mu$.*

2. Properties of Expectation

Theorem 1 (Change of variables, law of the unconscious statistician)**.** *Let X be a continuous random variable with pdf f. If $g : \mathbb{R} \to \mathbb{R}$ is Borel measurable, then*

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

whenever the integral is absolutely convergent.

The proof of Theorem 1 is harder for the case of continuous random variables. We will revisit this Theorem later when we discuss expectation in more generality. We will prove a simple case of Theorem 1.

Proof of Theorem 1—g strictly increasing and differentiable. Since g is increasing, let $h = g^{-1}$. We have that

$$\begin{aligned}\mathbb{P}(g(X) \leq y) &= \mathbb{P}(X \leq h(y)) \\ &= \int_{-\infty}^{h(y)} f_X(x)dx,\end{aligned}$$

so that the fundamental theorem of calculus gives that the density for $Y = g(X)$ is given by

$$f_Y(y) = f_X(h(y))h'(y).$$

Thus with the change of variables $y = g(x)$ and the inverse function theorem, we have

$$\mathbb{E}Y = \int y f_X(h(y))h'(y)dy = \int g(x)f_X(x)dx.$$

□

If X and Y are continuous random variables defined on same probability space, we say that they are **jointly continuous** with joint pdf $j : \mathbb{R} \times \mathbb{R} \to [0, \infty)$, if for all $x, y \in \mathbb{R}$, we have

$$\mathbb{P}(X \leq x, Y \leq y) = \int_{-\infty}^{y} \int_{-\infty}^{x} j(u, v)dudv;$$

note that the pdf of X is recovered by taking the whole integral in v,

$$f_X(x) = \int_{-\infty}^{\infty} j(u, v)dv$$

and in this context sometimes its referred to as a **marginal**.

Theorem 2. *If X and Y are jointly continuous random variables with joint pdf j then they are independent if and only if j factorizes as*

$$j(x, y) = f(x)g(y)$$

for some functions f and g.

Proof. Clearly, if j factorizes, then X and Y are independent. If X and Y are independent, then for (x, y) where $F(x, y) = \mathbb{P}(X \leq x, Y \leq y)$

is differentiable, we have

$$\begin{aligned} j(x,y) &= \frac{\partial^2}{\partial x \partial y}\mathbb{P}(X \leq x, Y \leq y) \\ &= \frac{\partial^2}{\partial x \partial y}\mathbb{P}(X \leq x)\mathbb{P}(Y \leq y) \\ &= \frac{\partial}{\partial x}\mathbb{P}(X \leq x)\frac{\partial}{\partial y}\mathbb{P}(Y \leq y) \\ &= f_X(x)f_Y(y). \end{aligned}$$

□

Theorem 3 (Change of variables, law of the unconscious statistician)**.** *Let X and Y be jointly continuous random variables with joint pdf j. If $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is Borel measurable, then*

$$\mathbb{E}g(X,Y) = \int\int g(x,y)j(x,y)dxdy.$$

whenever the integral is absolutely convergent.

Again, we will postpone the proof of Theorem 3. An easy corollary of Theorem 3 is again the linearity of expectation.

3. General properties of expectation

We will not have enough time to give a unified treatment of expectation. Using more of the machinery from measure theory it is possible, in several classes to end up with the following expectation operator. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let RV^+ denote the set of non-negative random variable. We can construct $\mathbb{E} : RV^+ \to [0, \infty]$, with the following properties.

(a) If X is a discrete random variable, then $\mathbb{E}X$ is the expectation in the elementary sense.
(b) Given a random variable $X \in RV^+$ there exists a sequence of discrete random variables such that for all $\omega \in \Omega$, we have $X_n(\omega) \to X(\omega)$ monotonically, so that $X_n(\omega) \leq X_{n+1}(\omega)$; furthermore, $\mathbb{E}X_n \to \mathbb{E}X$.
(c) The expectation operator is linear.

We define expectation for general random variables, by declaring:

$$\mathbb{E}X := \mathbb{E}X^+ - \mathbb{E}X^-,$$

where $X^+ = \max\{X, 0\}$ and $X^- = -\min\{X, 0\}$, so that $\mathbb{E}X$ is well defined as long as both the expectations in the difference are not both infinity. Notice that

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

The expectation operator agrees with all our previous notations for expectation of discrete and continuous random variables.

The expectation operator also satisfies certain continuity properties.

Theorem 4 (Convergence theorems)**.** *Let X_n be a sequence of random variables that converge to X (almost surely), so that for all $\omega \in \Omega$ (or an event Ω' with $\mathbb{P}(\Omega') = 1$), we have $X_n(\omega) \to X(\omega)$.*

(a) Monotone convergence theorem: If $0 \leq X_n$ is a non-decreasing sequence, then $\mathbb{E}X_n \to \mathbb{E}X$.

(b) Bounded convergence theorem: If there exists C such that $|X_n| \leq C$, then $\mathbb{E}|X| < \infty$ and $\mathbb{E}|X_n - X| \to 0$.

(c) Dominated convergence theorem: If there exists a random variable Y such that $\mathbb{E}|Y| < \infty$, and $|X_n| \leq Y$, then $\mathbb{E}|X| < \infty$ and $\mathbb{E}|X_n - X| \to 0$.

(d) Fatou's Lemma: If $X_n \geq 0$, then

$$\liminf_{n\to\infty} \mathbb{E}X_n \geq \mathbb{E}(\liminf_{n\to\infty} X_n).$$

Throughout the course, we will assume such an expectation operator exists. However, most of the time in our specific examples, we will operate in the discrete or continuous realm. In addition, we will not use any fine properties or details of the construction, and furthermore, will try to limit our use of Theorem 4.