NMAI059 – Probability and Statistics 1

Mykhaylo Tyomkyn

Lecture 7 - Expectation, variance, independence of continuous rv's.

The concepts we have developed for discrete random variables extend to the continuous setting analogously, whereby the pmf is to be replaced with the pdf, and the sums with the integrals.

Definition 1 (Expected value of a continuous rv) *Let X be a continuous random variable with probability density function f_X, and suppose*[1] *that*

$$\int_{-\infty}^{\infty} |x| f_X(x)\, dx < \infty.$$

The expected value/expectation/mean *of X is defined as*

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

As in the discrete case, the expectation is a function of the distribution/the pdf of X ('does not care where X is coming from'). If $\int_{-\infty}^{\infty} |x| f(x) = \infty$, we say that X does not have an expectation.

Exercise 1 *Prove that the Cauchy distribution does not have a (finite) expectation.*

If $\mathbb{E}(X)$ does exist ('is finite') we write $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, or simply (as the origin of X is often not important) $X \in L^1$ and call X an L^1*-random variable.*[2]

Example 1 *For $X \sim Exp(\lambda)$ we have (integrating by parts)*

$$\mathbb{E}(X) = \int_0^{\infty} x \cdot \lambda e^{-\lambda x}\, dx = \frac{1}{\lambda}.$$

Notice that this is analogous to $\mathbb{E}(X) = 1/p$ for $X \sim Geom(p)$.

Remark 1 *If X is 'discretized along the y-axis', that is if we set $X_n = \lfloor nX \rfloor / n$ for $n = 1, 2, \ldots,$ then each X_n is discrete, and by the properties of the Lebesgue integral, $\mathbb{E}(X) = \lim_{n\to\infty} \mathbb{E}(X_n)$.*

Theorem 1 (LOTUS for continuous rv's) *Let X be a continuous random variable with the pdf f_X, and let $g\colon \mathbb{R} \to \mathbb{R}$ be a function such that $g(X)$ is continuous.*[3] *Then (assuming convergence)*

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$$

[1] This is analogous to the $\sum |x| p_X < \infty$ condition in the definition of $\mathbb{E}(X)$ for discrete X.

[2] We will adopt this notation also with discrete random variables

[3] This is known to be the case when g is continuously differentiable and piecewise monotone, e.g. $g(x) = x^2$. We will take this condition for granted in applications.

We skip the proof, as it would require measure theory, but you can compare the statement with the discrete case. The two can be related via the discretization procedure described in Remark 1.

Applying LOTUS to a linear function $g(x) = ax + b$ we obtain

Corollary 1 *For any continuous X and $a, b \in \mathbb{R}$ we have $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$.*

Let us now discuss the joint distribution. We start with a definition that applies to arbitrary random variables.

Definition 2 (Joint cdf) *Let $X, Y\colon (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ be random variables. The* joint cumulative distribution function *of X and Y is defined as $F_{X,Y} : \mathbb{R}^2 \to [0,1]$,*

$$F_{X,Y}(x,y) = \mathbb{P}(X \leq x, Y \leq y).$$

Definition 3 (Jointly continuous rv's) *Two continuous random variables $X, Y\colon (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ are called* jointly continuous *if there exists a function $f = f_{X,Y}\colon \mathbb{R}^2 \to [0,\infty)$ such that for all $x, y \in \mathbb{R}$*

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s,t)\, dt\, ds.$$

In such case $f_{X,Y}$ is called the joint probability distribution function *(joint pdf) of X and Y.*

Note that, by Fubini's theorem (which remains valid for Lebesgue integrals), the order of integration can be reversed:

$$\int_{-\infty}^{x} \int_{-\infty}^{y} f(s,t)\, dt\, ds = \int_{-\infty}^{y} \int_{-\infty}^{x} f(s,t)\, ds\, dt.$$

Moreover, also by Fubini's theorem, f can be integrated over any Lebesgue measurable (i.e. any 'reasonably defined') shape $A \subseteq \mathbb{R}^2$, resulting in

$$\mathbb{P}((X,Y) \in A) = \int_A f(x,y)\, dx\, dy = \int_A f(x,y)\, dy\, dx.$$

Similarly to the discrete case, the *marginals* of $f_{X,Y}$ are given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy \quad \text{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\, dx,$$

and the *conditionals* are given by

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad \text{and} \quad f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

Theorem 2 (LOTUS for joint distribution) *Let $f_{X,Y}$ be the joint pdf of X and Y, and let $g\colon \mathbb{R}^2 \to \mathbb{R}$ be a function such that $g(X,Y)$ is continuous.*[4] *Then (assuming convergence)*

$$\mathbb{E}(g(X,Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\, dx\, dy.$$

We again skip the proof, but observe that the statement is analogous to the discrete case.

[4] This is the case for functions such as the sum and the product. We will take this condition for granted in applications.

Theorem 3 (Linearity of expectation) *For jointly continuous random variables $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ we have $X + Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and*

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

Proof Using LOTUS with the function $g(x, y) = x + y$ and Fubini's theorem gives

$$\begin{aligned}
\mathbb{E}(X + Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f_{X,Y}(x, y)\, dx\, dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X,Y}(x, y)\, dx\, dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_{X,Y}(x, y)\, dx\, dy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy\, dx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy \\
&= \int_{-\infty}^{\infty} x f_X(x)\, dx + \int_{-\infty}^{\infty} y f_Y(y)\, dy = \mathbb{E}(X) + \mathbb{E}(Y).
\end{aligned}$$

□

Just like in the discrete case the *variance* of a continuous random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is defined (subject to convergence) as $Var(X) = \mathbb{E}((X - \mathbb{E}(X)^2)$. If $Var(X) < \infty$, we call X an L^2-*random variable* and[5] write $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, or simply $X \in L^2$. By the linearity of expectation just established, we have (via the same proofs as in the discrete case) $Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$ and $Var(aX + b) = a^2 Var(X)$. Note that, by LOTUS, we have $\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx$. The *standard deviation* of $X \in L^2$ is, as before, $\sigma(X) = \sqrt{Var(X)}$.

Definition 4 (Independence or rv's in general) *$X, Y : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{L})$ are* independent *if for all $I, J \in \mathcal{L}$ we have*

$$\mathbb{P}(X \in I \wedge Y \in J) = \mathbb{P}(X \in I)\mathbb{P}(Y \in J).$$

Fact 1 *To ensure independence it suffices to check the above condition 'just' for the sets $\{(I, J) = ((-\infty, x], (-\infty, y]) : x, y \in \mathbb{R}\}$.*

In other words, it suffices to establish that $F_{X,Y}(x, y) = F_X(x)F_Y(y)$ for all $x, y \in \mathbb{R}$.

Lemma 1 *If X and Y are jointly continuous $f_{X,Y} = f_X \cdot f_Y$ then X and Y are independent*

Proof

$$\begin{aligned}
F_{X,Y}(x, y) &= \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s, t)\, dt\, ds = \int_{-\infty}^{x} \int_{-\infty}^{y} f_X(s) f_Y(t)\, dt\, ds \\
&= \int_{-\infty}^{x} f_X(s) \int_{-\infty}^{y} f_Y(t)\, dt\, ds = F_Y(y) \int_{-\infty}^{x} f_X(s)\, ds \\
&= F_Y(y) F_X(x).
\end{aligned}$$

□

[5] Note that $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ automatically implies $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$.

Fact 2 *The (essential) converse to this statement is also true: if X and Y are independent and jointly continuous then $f_{X,Y} = f_X f_Y$ a.s. (that is, $\mathbb{P}(f_{X,Y} \neq f_X \cdot f_Y) = 0$).*

Example 2 (Multivariate uniform distribution) *Take $f = \mathbb{1}_{[0,1]}$, the pdf of a $Unif([0,1])$ variable. Then the joint pdf of n independent $Unif([0,1])$ variables is*

$$f(t_1) \cdots f(t_n) = \mathbb{1}_{[0,1]}(t_1) \cdots \mathbb{1}_{[0,1]}(t_n) = \mathbb{1}_{[0,1]^n}(t_1, \ldots, t_n).$$

Example 3 (Multivariate normal distribution) *Take $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$, the pdf of a $\mathcal{N}(0,1)$ variable. Then the joint pdf of n independent $\mathcal{N}(0,1)$ variables is*

$$f(t_1, \ldots, t_n) = \phi(t_1) \cdots \phi(t_n) = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{t_1^2 + \cdots + t_n^2}{2}} = (2\pi)^{-\frac{n}{2}} e^{-\frac{r^2}{2}},$$

where $r = \sqrt{t_1^2 + \cdots + t_n^2}$ is the euclidean norm $||\cdot||_2$ of the vector $(t_1, \ldots, t_n) \in \mathbb{R}^n$.

Thus, f is a *radially symmetric* function (for the 'random vector' $X = (X_1, \ldots, X_n)$ every direction is 'equally likely'). In particular, for every vector $u \in \mathbb{R}^n$, $u = (u_1, \ldots, u_n)$ with $||u||_2 = 1$ we have

$$u \cdot X = \sum_{i=1}^{n} u_i X_i \sim \mathcal{N}(0,1).$$

Theorem 4 *For independent jointly continuous $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ we have $XY \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and*

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Proof By Theorem 2 with $g(x,y) = xy$, and by independence, we have

$$\begin{aligned}
\mathbb{E}(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{X,Y}(x,y)\, dx\, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y)\, dx\, dy \\
&= \int_{-\infty}^{\infty} x f_X(x) \int_{-\infty}^{\infty} y f_Y(y)\, dy\, dx = \int_{-\infty}^{\infty} x f_X(x) \mathbb{E}(Y)\, dx \\
&= \mathbb{E}(Y) \int_{-\infty}^{\infty} x f_X(x)\, dx = \mathbb{E}(Y)\mathbb{E}(X).
\end{aligned}$$

□

Note that the converse does not hold: we can have $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ (in which case we say X and Y are *uncorrelated*), when X and Y are not independent.

Example 4 *Let $X \in \{-1, 0, 1\}$ with $p_X(-1) = p_X(0) = p_X(1) = 1/3$, and let $Y = 1$ if $X = 0$, and $Y = 0$ otherwise. Then X and Y are uncorrelated but not independent.*

Corollary 2 *For independent jointly continuous $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ we have $X + Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ and*

$$Var(X + Y) = Var(X) + Var(Y).$$

The proof is exactly the same as in the discrete case.