1 Moments

Definition 1.1. Let $X$ be a random variable and $k \geq 0$. If $\mathbb{E}(|X|^k) < \infty$, we say that the moment of order $k$ of the random variable $X$ is finite and defined as

$$m_k = \mathbb{E}(X^k).$$

Notice that if a random variable has a finite moment of order $k$, then all moments of order $0 \leq j \leq k$ are also finite. Indeed, using Jensen’s inequality along with the fact that $f(x) = x^{j/k}$ is concave for $x \geq 0$ and $j \leq k$, we obtain

$$\mathbb{E}(|X|^j) = \mathbb{E}(f(|X|^k)) \leq f(\mathbb{E}(|X|^k)) < \infty.$$

Notice also that the moment of order 0 is always finite and equal to $m_0 = 1$.

We give below a list of interesting properties of moments.

a) If all the moments $(m_k, k \geq 0)$ of a given random variable $X$ are finite, then the infinite matrix $M$ whose entries are given by

$$M_{jk} = m_j + k, \quad j,k \geq 0$$

is positive semi-definite, that is,

$$\sum_{j,k=1}^n c_j c_k M_{jk} \geq 0, \quad \forall n \geq 1, \ c_1, ..., c_n \in \mathbb{C}.$$

Indeed,

$$\sum_{j,k=1}^n c_j c_k m_{j+k} = \sum_{j,k=1}^n c_j c_k \mathbb{E}(X^{j+k}) = \mathbb{E} \left( \left( \sum_{j=1}^n c_j X^j \right)^2 \right) \geq 0.$$

It is however not clear whether the distribution of a random variable $X$ is completely determined by its moments, even if they are all finite. In order to answer this question, we need to relate the moments to the characteristic function (to which we know that a unique distribution is associated, thanks to the inversion formula).

b) Let $k \geq 1$. If $\mathbb{E}(|X|^k) < \infty$, then $\phi_X$ is $k$ times continuously differentiable on $\mathbb{R}$ and

$$\left. \frac{d^k \phi_X}{dt^k} \right|_{t=0} = i^k m_k, \quad \text{so} \quad \phi_X(t) = \sum_{j=0}^k \frac{i^j m_j}{j!} t^j + o(t^k),$$

where $g(t) = o(t^k)$ means that $\lim_{t \to 0} |g(t)/t^k| = 0$. The first relation above actually says that the characteristic function $\phi_X$ is a moment generating function.

c) If $\phi_X$ is $k$ times differentiable in $t = 0$, then

$$\mathbb{E}(|X|^{2p}) < \infty, \quad \forall p \in \mathbb{N} \text{ such that } 2p \leq k.$$
Remark 1.3. Condition (1) actually limits the growth of the sequence \((m_k, k \geq 0)\). It only involves even moments \(m_{2k}\), because the growth of the odd moments is controlled by that of the even moments. Cauchy-Schwarz' inequality indeed guarantees that

\[
m_{2k+1} = (\mathbb{E}(X^{2k+1}))^2 \leq \mathbb{E}(X^{2k}) \mathbb{E}(X^{2k+2}) = m_{2k} m_{2k+2}, \ \forall k \geq 0.
\]

Proof of Theorem 1.2 in a particular case. Let us assume that the following stronger condition is met:

\[
\exists C > 0 \text{ such that } m_{2k} \leq C^{2k}, \ \forall k \geq 0. \tag{2}
\]

One can check that this condition indeed implies condition (1).

Under this condition, the characteristic function \(\phi_X\) is given by the power series

\[
\phi_X(t) = \sum_{k \geq 0} \frac{i^k m_k}{k!} t^k, \quad t \in \mathbb{R},
\]

which is analytic on \(\mathbb{R}\). In particular, it is completely determined by the value of all its derivatives in 0, i.e. by the moments, as we have seen above that:

\[
\left. \frac{d^k \phi_X}{dt^k} \right|_{t=0} = i^k m_k, \quad \forall k \geq 0.
\]

The proof then follows from the inversion formula, which guarantees that to a given characteristic function corresponds a unique distribution. □

Corollary 1.4. If \(X\) is a bounded random variable (i.e. if there exists \(C > 0\) such that \(|X(\omega)| \leq C\) for all \(\omega \in \Omega\)), then its distribution is completely determined by its moments.

Proof. If \(X\) is known to be bounded, then all its moments are finite; moreover,

\[|m_k| = |\mathbb{E}(X^k)| \leq C^k, \quad k \geq 0,
\]

so condition (2) is met, and the corollary follows. □

Remark 1.5. In general, we have the following correspondence:

weight of the tail of the distribution of \(X\) ↔ regularity of \(\phi_X\) ↔ growth of the moments of \(X\).

Here are a few examples:

In the case of a bounded random variable, the weight of the tail is zero, the characteristic function \(\phi_X\) is analytic on \(\mathbb{R}\), and the moments satisfy condition (2).

In the case of an exponential random variable, the tail of the distribution is \(\mathbb{P}(\{|X| > x\}) = \exp(-x)\), the characteristic function \(\phi_X(t) = \frac{1}{1-2it}\) is analytic in a neighborhood of \(t = 0\), and the moments \(m_k = k! \sim \exp(k \log k)\) satisfy condition (1).

In the case of a Cauchy random variable, the tail of the distribution is \(\mathbb{P}(\{|X| > x\}) \sim 1/x\), the characteristic function \(\phi_X(t) = \exp(-|t|)\) is only continuous in \(t = 0\), but not differentiable, and all the moments (except \(m_0\)) are infinite.

Moments and convergence in distribution. The above analysis allows us to state the following nice criterion for checking convergence in distribution.

Theorem 1.6. Let \((m_k, k \geq 0)\) be a sequence of real numbers satisfying condition (1). If \((X_n, n \geq 1)\) is a sequence of random variables such that for all \(k \geq 0,
\[
\mathbb{E}(X_n^k) \rightarrow m_k, \quad \text{as } n \rightarrow \infty,
\]

then \(X_n\) converges in distribution towards a random variable \(X\), whose moments are \((m_k, k \geq 0)\).

\[1\]This is a notion from complex analysis: being “analytic” actually means being “expandable in an infinite power series”: in particular, all analytic functions are \(C^\infty\), but the reciprocal statement does not hold.