General definition of expectation

atmadeep sengupta

March 2024

1 Introduction

In this note, we aim to generalise the notion of expectation of a random variable, more specifically extended real random variable(rv's taking values in $[-\infty, \infty]$). Our definition should capture all the other existing notions of expectation, and connect the different formulae for different cases. We will define in three steps, and in each we will check two important characteristics of expectation each time, i.e., monotonicity and linearity.

2 Step 1

real simple rv: X is a real simple rv if it takes finitely many values.
At first, we will define expectation of simple real rv's.
If 1_A is the indicator variable of set A, i.e. takes value 1 if the event belongs to A, and 0 otherwise, then for some partition $\{A_1, \ldots A_n\}$ of R, and for some real numbers $c_1, \ldots c_n$, any real simple rv X can be represented by $X = \sum_{i=1}^{n} c_i \cdot 1_{A_i}$.
We call this representation the **canonical representation** of X.
Suppose $X = \sum_{i=1}^{n} c_i \cdot 1_{A_i} \sum_{j=1}^{m} d_j \cdot 1_{B_j}$, where $\{A_1, \ldots A_n\}$ and $\{B_1, \ldots B_m\}$ are two partitions of R.
$X = \sum_{i=1}^{n} c_i \cdot 1_{A_i} = X = \sum_{i=1}^{n} c_i \cdot 1_{A_i \cap R} = \sum_{i=1}^{n} c_i \cdot 1_{A_i \cap (\bigcup_{j=1}^{m} B_j)}$
$= \sum_{i=1}^{n} c_i \cdot 1_{\bigcup_{j=1}^{m} (A_i \cap B_j)} = \sum_{i=1}^{n} \sum_{j=1}^{m} c_i 1_{A_i \cap B_j}$
Similarly, $X = \sum_{i=1}^{n} \sum_{j=1}^{m} d_j 1_{A_i \cap B_j}$.
Now, if $A_i \cap B_j = \phi$, then $1_{A_i \cap B_j} = 0$.
If $A_i \cap B_j \neq \phi$, take any $w \in A_i \cap B_j$, $X(w) = c_i = d_j$. Hence m=n, $A_i = B_i$ and $d_i = c_i$[As WLOG we can take c_i's all distinct] Hence X has unique canonical representation.
$E(X) := \sum_{i=1}^{n} c_i P(A_i)$.
As αX is also a simple random variable, so $E(\alpha X) = \sum_{i=1}^{n} \alpha c_i P(A_i) = \alpha E(x)$.
Let $X = \sum_{i=1}^{n} c_i \cdot 1_{A_i}, Y = \sum_{j=1}^{m} d_j \cdot 1_{B_j}$. Clearly X+Y is a simple real rv.
$X+Y = \sum_{i=1}^{n} \sum_{j=1}^{m} (c_i+d_j) 1_{A_i \cap B_j} = \sum_{i=1}^{n} c_i \sum_{j=1}^{m} 1_{A_i \cap B_j} + \sum_{j=1}^{m} d_j \sum_{i=1}^{n} 1_{A_i \cap B_j}$
$E(X+Y) = \sum_{i=1}^{n} c_i \sum_{j=1}^{m} P(A_i \cap B_j) + \sum_{j=1}^{m} d_j \sum_{i=1}^{n} P(A_i \cap B_j) = \sum_{i=1}^{n} c_i P(A_i) + \sum_{j=1}^{m} d_j P(B_j) = E(X) + E(Y)$
$X \leq Y \Rightarrow c_i \leq d_j$ whenever $A_i \cap B_j \neq \phi$. Hence $E(X) \leq E(Y)$.

We have proved both the necessary properties mentioned earlier in this case. Now we proceed to step 2.

3 Step 2

Now we consider X any non negative rv.
$E(X) = sup\{E(Y) :$ Y simple, real s.t. $0 \leq Y \leq X\}$. Hence $0 \leq E(X) \leq +\infty$
Observe, if X is simple, then E(X) has two equivalent definitions.
$X_1 \geq X_2 \Rightarrow E(X_1) \geq E(X_2)$, as $\{E(Y) :$ Y simple, real s.t. $0 \leq Y \leq X_2\} \subseteq \{E(Y) :$ Y simple, real s.t. $0 \leq Y \leq X_2\}$
Result: If X be any non negative rv. If X_n, $n \geq 1$ be any sequence of non negative simple rv's with $X_n \uparrow X$, then $E(X_n) \uparrow E(X)$.
Proof As $X_n \geq X_{n-1} \Rightarrow E(X_n) \geq E(X_{n-1}$, hence lim $E(X_n)$ exists.
As $X_n \uparrow X$; $E(X) \geq E(X_n)$ $\forall n$, hence $E(X) \geq limE(X_n)$. To show, $E(X) \leq E(X_n)$
i.e. enough to show for ANY real simple rv Y with $0 \leq Y \leq X$, lim $E(X_n) \geq E(Y)$.
Fix $0 < \alpha < 1$, will prove lim $E(X_n) \geq \alpha E(Y)$.
Now for every n, define $\Omega_n = \{w \in \Omega; X_n(w) \geq \alpha Y(w)$
Observe $\Omega_n \subseteq \Omega_{n+1}$. If $Y(w) = 0$, $w \in \Omega_n \forall n$.
If $Y(w) > 0$ then $\alpha Y(w) < Y(w) \leq X(w)$. For some n, $X_n(w) \geq \alpha Y(w)$. Hence $\Omega_n \uparrow \Omega$.
$\forall n$, $X_n \geq X_n 1_{\Omega_n} \geq \alpha Y 1_{\Omega_n} = \sum_{i=1}^{m} \alpha c_i 1_{A_i \cap \Omega_n}$
And as $A_i \cap \Omega_n \uparrow A_i \uparrow \Omega = A_i$, hence $\alpha Y 1_{\Omega_n} \uparrow Y$, i.e. lim $X_n \geq \alpha Y$, and as α is arbitrary; lim $X_n \geq Y$, i.e., lim $E(X_n) \geq E(y)$, hence proved.
Take $X_n \uparrow X, Y_n \uparrow Y$, $E(X + Y) = limE(X_n + Y_n) = limE(X_n) + limE(Y_n) = E(X) + E(Y)$.
WLOG $\alpha > 0$, $E(\alpha X) = limE(\alpha X_n) = \alpha limE(X_n) = \alpha E(X)$, hence E is linear
Now we proceed to all possible rvs.

4 Step 3

Let X be any rv.
Define $X^+ = max\{X, 0\}$, $X^- = max\{-X, 0\}$. So both X^+, X^- are non negative rv.
Also, $X = X^+ - X^-$. $E(X) := E(X^+) - E(X^-)$, whenever $E(X^+), E(X^-)$ not both ∞. Else, we say, E(X) doesn't exist.
If $X \geq 0$; $X^- = 0$ i.e., $E(X) = E(X^+)$ i.e. consistent with step 2. If $X \leq Y$ and $E(X), E(Y)$ both exists.
To show, $E(X) \leq E(Y)$
Observe $X^+ \leq Y^+$, $X^- \geq Y^-$. If either $E(X) = -\infty, E(Y) = \infty$, we are done.
Else $E(X) > -\infty \Rightarrow E(X^-) < \infty \Rightarrow E(Y^-)$ and $E(Y) < \infty \Rightarrow E(Y^+) < \infty \Rightarrow E(X+) < \infty$, and hence $E(X) = E(X^+) - E(X^-) \leq E(Y^+) - E(Y^-) = E(Y)$

If E(X) exists; $\alpha \in R$. WLOG $\alpha > 0$.
αX^+=$(\alpha X)^+$, and $\alpha X^- = (\alpha X)^-$.
Hence $E(\alpha X) = E(\alpha X^+) - E(\alpha X^-) = \alpha(E(X^+) - E(X^-)) = \alpha E(X)$
X,Y r.v.s, To show $E(X+Y) = E(X) + E(Y)$ whenever X+Y is defined pointwise, E(X), E(Y) exists and $E(X) + E(Y)$ exists. Suppose $E(X^+) = \infty \Rightarrow E(X^-) < \infty \Rightarrow E(X) = \infty \Rightarrow E(Y) > -\infty \Rightarrow E(Y^-) < \infty$.
Hence either both $E(X^+), E(Y^+)$ both $< \infty$ or both $E(X^-), E(Y^-)$ both $< \infty$.
Now, to show $(X+Y)^+ + X^- + Y^- = (X+Y)^- + X^+ + Y^+$.
Observe if $X(w), Y(w)$ both have same sign, WLOG positive, then $(X{+}Y)^-, X^-, Y^-$=0 and $(X+Y)^+ = X^+ + Y^+$.
If both have different sign,WLOG $X(w) \geq 0 \geq Y(w)$ and $|X(w)| > |Y(w)|$, i.e. atleast Y(w) is finite.
Hence $(X+Y)^+ = X - Y$, $X^+ = X, Y^- = -Y$, and all other 0. Again, the equality holds, hence proved.
Observe $0 \leq (X+Y)^+ \leq X^+ + Y^+$ and $0 \leq (X+Y)^- \leq X^- + Y^-$. So if $E(X^+), E(Y^+)$ finite, then $E((X+Y)^+)$ is finite. If $E(Y^-), E(X^-)$ finite, then $E((X+Y)^-)$ is finite. Hence, in any case, $(X+Y)^+ - (X+Y)^- = X^+ - X^- + Y^+ - Y^- \Rightarrow E(X+Y) = E(X) + E(Y)$. Hence E is linear.

5 Conclusion

Note that, if we dealt with only real random variables and finite expectations, many of the troubles could have been avoided. But as that is not the case, we have to go through all these. Anyways, we can see that we have a general definition of expectation. This captures every possible kind of rv's, much more than only discrete and only absolutely continuous rv's.