Sigma-algebras and random variables

Anders Muszta

June 17, 2005

1 Measurable space

Let Ω be any nonempty set. A collection, $\mathcal{F}$, of subsets of Ω is said to be a *σ-algebra* of subsets of Ω if it has the following properties:

1. $\Omega \in \mathcal{F}$;
2. If $A \in \mathcal{F}$, then $A^c \equiv \Omega \setminus A \in \mathcal{F}$;
3. If $\{A_n\}_{n=1}^{\infty} \in \mathcal{F}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

The pair $(\Omega, \mathcal{F})$ is called a *measurable space*.

Note 1. *If you want to understand how something works, let's say a clock, then a good idea is to disassemble the clock into the pieces of which it is built and then putting the pieces back together. Now, of course you can't put the pieces together anyway you like but they have to fit together in a certain way.*

In this context the pieces of the clock is represented by the set Ω and the ways in which you are allowed to combine the pieces of the clock is represented by the sigma-algebra, $\mathcal{F}$, of Ω. Just as you can learn about the workings of different mechanisms of the clock by combining some of its parts, so you can learn about the set Ω by considering the subsets of the sigma-algebra, $\mathcal{F}$.

The sole purpose of the concept of a sigma-algebra is to tell us exactly which subsets of Ω we are allowed to consider in our investigation.

The pair $(\Omega, \mathcal{F})$ is called a *measurable space*.

2 Random variable

Definition 1 (Random variable). *A function $X : \Omega \to \mathbb{R}$ is called a (real-valued)* random variable *if is has the following property: For every real number $x \in \mathbb{R}$, the subset $\{\omega \in \Omega : X(\omega) \leqslant x\}$ of Ω is a element of the sigma-algebra $\mathcal{F}$,i.e.,*

$$\forall\, x \in \mathbb{R},\ \{\omega \in \Omega : X(\omega) \leqslant x\} \in \mathcal{F}.$$

Example 1 (Indicator function). *Let $A \subseteq \Omega$ be any subset of Ω and define the* indicator function $1_A : \Omega \to \mathbb{R}$ *by*

$$1_A(\omega) = \begin{cases} 1, & \omega \in A; \\ 0, & \omega \notin A. \end{cases}$$

For the indicator function we observe that

$$\{\omega \in \Omega : 1_A(\omega) \leqslant x\} = \begin{cases} \Omega, & x \geqslant 1; \\ A^c, & 0 \leqslant x < 1; \\ \emptyset, & x < 0. \end{cases}$$

In order for 1_A to be a random variable we must have all three sets Ω, A^c and $\emptyset$ be elements of the sigma-algebra $\mathcal{F}$. Since $\Omega \in \mathcal{F}$ by definition and $\emptyset = \Omega^c \in \mathcal{F}$, we see that 1_A is a random variable if and only if $A^c \in \mathcal{F}$. But $A^c \in \mathcal{F}$ if and only if $A \in \mathcal{F}$. Thus we have established the fact that for any set $A \subseteq \Omega$,

$$1_A \textit{ is a random variable } \Leftrightarrow A \in \mathcal{F}.$$

Example 2 (Linear combination of indicator functions). *Consider two subsets $A, B \subseteq \Omega$ of Ω and their associated indicator functions 1_A and 1_B. For any two real numbers $0 \leqslant a \leqslant b \in \mathbb{R}$, define a function $X : \Omega \to \mathbb{R}$ by*

$$X(\omega) = a1_A(\omega) + b1_B(\omega).$$

For this function we observe that

$$X(\omega) = \begin{cases} a+b, & \omega \in A \cap B; \\ b, & \omega \in B \cap A^c; \\ a, & \omega \in A \cap B^c; \\ 0, & \omega \in A^c \cap B^c. \end{cases}$$

Thus we find that

$$\{\omega \in \Omega : X(\omega) \leqslant x\} = \begin{cases} \Omega, & x \geqslant a+b; \\ \Omega \setminus (A \cap B), & b \leqslant x < a+b; \\ \Omega \setminus B, & a < x \leqslant b; \\ \Omega \setminus (A \cup B), & 0 \leqslant x \leqslant a; \\ \emptyset, & x < 0. \end{cases}$$

In order for X to be a random variable we need to have $A \cup B$, B and $A \cap B$ be elements of the sigma-algebra $\mathcal{F}$. However, since $A = (A \cup B) \cup B^c$, we see that A is also an element of the sigma-algebra. We have therefore established that fact that

$$a1_A + b1_B \textit{ is a random variable } \Leftrightarrow A, B \in \mathcal{F}.$$

Note that we assumed that the real numbers a and b were such that $0 \leqslant a \leqslant b$. This we did for convenience only; Hence the result, $a1_A + b1_B$ is a random variable if and only if A and B are elements of the sigma-algebra $\mathcal{F}$, holds for any *real numbers a and b.*

Let $\{X_i\}_{i \in I}$ be any collection of random variables, $X_i : \Omega \to \mathbb{R}$, on a measurable space $(\Omega, \mathcal{F})$. Is the function $S : \Omega \to \mathbb{R}$ defined by

$$S(\omega) \equiv \sup_{i \in I} X_i(\omega)$$

also a random variable? For general index sets I, the answer is: We don't know! The reason for this depends on the way we defined a sigma-algebra, and will now be demonstrated.

In order to determine whether or not S is a random variable we need to consider subsets of Ω of the form

$$\{\omega \in \Omega : S(\omega) \leqslant x\} \text{ for an arbitrary } x \in \mathbb{R}.$$

Now,

$$\{\omega \in \Omega : S(\omega) \leqslant x\} = \{\omega \in \Omega : \sup_{i \in I} X_i(\omega) \leqslant x\} = \bigcap_{i \in I} A_i,$$

where we have defined the subsets A_i of Ω as $A_i \equiv \{\omega \in \Omega : X_i(\omega) \leqslant x\}$. Because we know that each X_i is a random variable, every set A_i is an element of the sigma-algebra $\mathcal{F}$. So we have a collection $\{A_i\}_{i \in I}$ of elements in the sigma-algebra $\mathcal{F}$ and we are asking whether their intersection $\bigcap_{i \in I} A_i$ also is an element of $\mathcal{F}$. If we recall the definition of a sigma-algebra it was stated that if $\{A_n\}_{n=1}^{\infty}$ is a *countable* collection of elements of $\mathcal{F}$, then their union $\bigcup_{n=1}^{\infty} A_n$ is also an element of $\mathcal{F}$. In our case we have a general index set, I, which need not be countable. Hence in general we cannot say whether $\bigcap_{i \in I} A_i \in \mathcal{F}$ or not.

However, if we have a *countable* collection $\{X_n\}_{n=1}^{\infty}$ of random variables, then

$$\sup_{n \geqslant 1} X_n$$

will be a random variable.

By similar reasoning, we may establish that in general we cannot say whether

$$\inf_{i \in I} X_i$$

is a random variable. But, for a countable collection, $\{X_n\}_{n=1}^{\infty}$, of random variables,

$$\inf_{n \geqslant 1} X_n$$

will be a random variable.

We may use this knowledge that $\sup_{n \geqslant 1} X_n$ and $\inf_{n \geqslant 1} X_n$ are random variables, whenever $\{X_n\}_{n=1}^{\infty}$ are, to deduce that $\lim_{n \to \infty} X_n$ also is a random variable. The limit $\lim_{n \to \infty} X_n$ can be expressed in terms of $\sup X_n$ and $\inf X_n$ as [1]

$$\lim_{n \to \infty} X_n = \inf_{n \geqslant 1} \{ \sup_{m \geqslant n} X_m \} = \sup_{n \geqslant 1} \{ \inf_{m \geqslant n} X_m \}.$$

Either of these expressions show that $\lim_{n \to \infty} X_n$ is a random variable.

The previous Example 2 on page 2, shows us that any finite linear combination

$$\sum_{k=1}^{n} a_k 1_{A_k}(\omega)$$

of indicator random variables, $1_{A_k}(\omega)$, is again a random variable. It is an essential fact that *any* random variable can be approximated arbitrarily well by a finite linear combination of indicator random variables.

[1] The expression $\inf_{n \geqslant 1}\{\sup_{m \geqslant n} X_m\}$ is called *limit superior* of the sequence $\{X_n\}_{n=1}^{\infty}$ and is denoted by $\limsup_{n \to \infty} X_n$. The expression $\sup_{n \geqslant 1}\{\inf_{m \geqslant n} X_m\}$ is called *limit inferior* of $\{X_n\}_{n=1}^{\infty}$ and is denoted by $\liminf_{n \to \infty} X_n$.

Theorem 1 (Random variable approximation). *Let $X : \Omega \to \mathbb{R}$ be a random variable on the measurable space $(\Omega, \mathcal{F})$. Then there exists a sequence $\{a_k\}_{k=1}^{\infty} \in \mathbb{R}$ of real numbers and a sequence $\{A_k\}_{k=1}^{\infty} \in \mathcal{F}$ of elements of the sigma-algebra $\mathcal{F}$, such that for every $\omega \in \Omega$,*

$$X(\omega) = \lim_{n\to\infty} \sum_{k=1}^{n} a_k 1_{A_k}(\omega).$$

Proof. Define a sequence of functions $\{X_n\}_{n=1}^{\infty}$, $X_n : \Omega \to \mathbb{R}$ by

$$X_n(\omega) \equiv \sum_{k=-n2^n}^{n2^n-1} \frac{k}{2^n} 1_{[k2^{-n},(k+1)2^{-n})}\{X(\omega)\}.$$

Because we have $\frac{k}{2^n}$ in front of the indicator functions $1_{[k2^{-n},(k+1)2^{-n})}\{X(\omega)\}$, for every $n \geqslant 1$ and for every $\omega \in \Omega$, $X_n(\omega) \leqslant X(\omega)$. This implies that $\sup_{n\geqslant 1} X_n(\omega) = X(\omega)$. Also, because $X_n(\omega) \leqslant X_{n+1}(\omega)$, $\lim_{n\to\infty} X_n(\omega) = \sup_{n\geqslant 1} X_n(\omega)$.

All that remains for us to show is that for every k and n, the indicator functions $1_{[k2^{-n},(k+1)2^{-n})}\{X(\omega)\}$ are random variables. We will now proceed to do this.

Let $x \in \mathbb{R}$ be an arbitrary real number, and consider the subset of Ω

$$\Big\{\omega \in \Omega : 1_{[k2^{-n},(k+1)2^{-n})}\{X(\omega)\} \leqslant x\Big\}.$$

Depending on the value of the real number x, this set is one of three possible sets:

$$\begin{aligned} &\Big\{\omega \in \Omega : 1_{[k2^{-n},(k+1)2^{-n})}\{X(\omega)\} \leqslant x\Big\} \\ &= \begin{cases} \Omega, & x \geqslant 1; \\ \{\omega \in \Omega : X(\omega) \notin [\frac{k}{2^n}, \frac{k+1}{2^n})\}, & 0 \leqslant x < 1; \\ \emptyset, & x < 0. \end{cases} \end{aligned} \tag{1}$$

In order for the indicator function to be a random variable, the second set in (1) need to be an element of the sigma-algebra $\mathcal{F}$.

$$\begin{aligned} &\Big\{\omega \in \Omega : X(\omega) \notin [\frac{k}{2^n}, \frac{k+1}{2^n})\Big\} \\ &= \{\omega \in \Omega : X(\omega) < k2^{-n}\} \cup \{\omega \in \Omega : X(\omega) < (k+1)2^{-n}\}^c. \end{aligned}$$

Each of these two sets are elements of the sigma-algebra $\mathcal{F}$ because X is a random variable. Thus we see that indeed the indicator function $1_{[k2^{-n},(k+1)2^{-n})}\{X(\omega)\}$ *is* a random variable. □

3 Conditional expectation

Definition 2 (Heuristic definition). *Let $X : \Omega \to \mathbb{R}$ be any random variable and let $\mathcal{F}$ be any sigma-algebra of subsets of Ω. The* conditional expectation *of X, given the information contained in the sigma-algebra $\mathcal{F}$ is denoted by $\mathbb{E}\{X|\mathcal{F}\}$. The conditional expectation is defined to be the best* guess *of the value of the random variable X, if we have available the information contained in the sigma-algebra $\mathcal{F}$.*

Definition 3 (Mathematical definition). *Let $X : \Omega \to \mathbb{R}$ be any random variable and let $\mathcal{F}$ be any sigma-algebra of subsets of Ω. The* conditional expectation *of X, given the information contained in the sigma-algebra $\mathcal{F}$ is denoted by $\mathbb{E}\{X|\mathcal{F}\}$. The* conditional expectation *is defined to be the random variable $Y = \mathbb{E}\{X|\mathcal{F}\}$ such that*

$$\text{for every set } A \in \mathcal{F}, \mathbb{E}\{X 1_A\} = \mathbb{E}\{Y 1_A\},$$

and the conditional expectation, $\mathbb{E}\{X|\mathcal{F}\}$, does not provide us with any more information than that contained in the sigma-algebra $\mathcal{F}$, i.e.,

$$\sigma(\mathbb{E}\{X|\mathcal{F}\}) \subset \mathcal{F},$$

where $\sigma(\mathbb{E}\{X|\mathcal{F}\})$ is the smallest sigma-algebra containing complete information on the conditional expectation $\mathbb{E}\{X|\mathcal{F}\}$.

Example 3. *Let the sigma-algebra $\mathcal{F}$ contain complete information on the random variable X. Then the best guess of the value of X, given this information is X, since $\mathcal{F}$ tells us everything there is to know about the random variable X. Thus, $\mathbb{E}\{X|\mathcal{F}\} = X$.*