Random Variable Defn. 9/15/~~2008~~ 2022

Defn. Given a prob space $(\mathcal{S}, \mathcal{F}, P)$, a random variable (rv) is a function $X : \mathcal{S} \longrightarrow \mathbb{R}$ with the property that $X^{-1}(A) \in \mathcal{F}$ for every $A \in \mathcal{B}(\mathbb{R})$

($\mathcal{R} = \mathbb{R}$ for a rv)

We will want to know $P(X \in A)$ for subsets $A \subset \mathbb{R}$

For example, what is $P(400K < X \le 1000K)$ if X is a temperature? Here $A = (400, 1000]$.

Note that $\{X \in A\} = \{\omega \in \mathcal{S} : X(\omega) \in A\} = X^{-1}(A)$

Note that $A \subset \mathbb{R}$, $X^{-1}(A) \subset \mathcal{S}$

We have a prob space $(S_X, \mathcal{F}_X, P_X)$ induced by rv X.

$S_X = \mathbb{R}$, then $\mathcal{F}_X = B(\mathbb{R})$.

What is P_X?

We will want

$$P_X(A) = P(X^{-1}(A)) \quad \forall A \in B(\mathbb{R})$$

We need the property that $X^{-1}(A) \in \mathcal{F} \quad \forall A \in B(\mathbb{R})$, so that the right-hand side is well-defined.

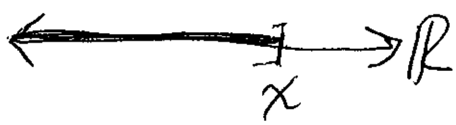
(The function X is said to be Borel-measurable if this is satisfied)

In practice, how do we find the measure P for a rv? Instead of finding $P(X \in A) \; \forall A \in B(\mathbb{R})$, we use:

1. The cumulative distribution function (cdf) F_X
2. The prob. density function (pdf) f_X
3. The prob. mass function (pmf) p_X (if X is a discrete rV — to be defined later)

Defn. The cumulative distr. fn., or distribution fn., or the cdf, of a rV X is

$$F_X(x) = \underbrace{P(X \le x)}_{P(\{\omega \in \mathcal{S} : X(\omega) \le x\})}, \quad \forall x \in \mathbb{R}$$



Notation: $F_X(x)$

↑ label to identify which rV the cdf belongs to

Some comments

- The cdf exists for any rv X.
- Since P must satisfy the axioms, we know that

① $\lim_{x \to \infty} F_X(x) = 1$, and

$\lim_{x \to -\infty} F_X(x) = 0$

Note that (informal proof)

$$\lim_{x \to \infty} F_X(x) = \lim_{x \to \infty} P(\{\omega \in S : X(\omega) \le x\})$$

$$= P\left(\lim_{x \to \infty} \{\omega \in S : X(\omega) \le x\}\right)$$

(assuming this can be done – which it can)

$$= P(S) = 1.$$

② For any $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$,

$$F_X(x_1) \leq F_X(x_2)$$

so F_X is non-decreasing

3 F_X is continuous from the right at all $x \in \mathbb{R}$, i.e.,

$$F_X(x^+) = \lim_{\substack{\epsilon \to 0, \\ \epsilon > 0}} F_X(x+\epsilon) = F_X(x) \quad \forall x \in \mathbb{R}$$

Proof of 3. Some prelim. results without proof:

(i) For a sequence of sets with $A_1 \supset A_2 \supset A_3 \supset \cdots$

$$\lim_{n \to \infty} A_n = \bigcap_{n=1}^{\infty} A_n$$

(ii) If $A_1 \supset A_2 \supset \cdots$, then

$$P(\lim_{n \to \infty} A_n) = \lim_{n \to \infty} P(A_n)$$

(from measure th.)

(iii) We can write $F_X(x^+) = \lim_{n \to \infty} F_X(x + \frac{1}{n})$

Now let $A_n = \{X \le x + \frac{1}{n}\}$ $\forall x \in \mathbb{R}$

Then $F_X(x^+) = \lim_{n\to\infty} P(X \le x + \frac{1}{n})$

$$= \lim_{n\to\infty} P(A_n)$$

$$= P(\lim_{n\to\infty} A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right)$$

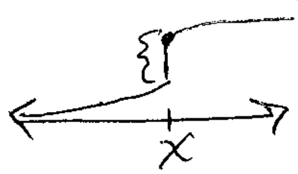
$$= P\left(\bigcap_{n=1}^{\infty} \{X \le x + \tfrac{1}{n}\}\right) = F_X(x),$$

since $\{X \le x\} = \bigcap_{n=1}^{\infty} \{X \le x + \frac{1}{n}\}$

QED

4 $P(X > x) = 1 - F_X(x)$ $\forall x \in \mathbb{R}$

5 $P(X = x) = F_X(x) - F_X(x^-)$



$\forall x \in \mathbb{R}$

Note: It can be shown that if a function F satisfies

Properties (1) - (3) above, then it is a valid cdf for some rv X. (so the measure P defined as

$$P(X^{-1}(A)) \quad \forall A \in \mathcal{B}(\mathbb{R})$$

~~with~~ ~~satis~~ where $P(X \leq x) = F_X(x)$ will satisfy the axioms

(Proof outside the scope of this class)