Layman's talk : what is continuous random variable

Random variables are functions

X maps elements from S to numbers! Now we can work with functions/numbers!

$$
\begin{tikzcd}
s \in S \arrow[r, "X"] \arrow[dr, dashed, "P : A \subseteq S"'] & Numbers \in \mathcal{C} \arrow[d, dashed, "P : \{X \in C \subseteq \mathcal{C}\} \subseteq S"] \\
& {[0,1]}
\end{tikzcd}
$$

X essentially is a function to numbers!

E.g., X may assign random outcomes with real numbers, and thus it may take values from the real line.

To describe how the probability values are assigned to different events generated by a **real-valued** random variable X, we can look at the cumulative distribution function F_X such that

$$F_X(x) = P(X \leq x), x \in \mathbb{R}.$$

Continuous r.v. example

Let's keep throwing a ten-sided fair die ((with numbers 0,1,...,9)) without stopping!

- We have outcomes such as

$$a = 010310819.....$$

- Let's map these outcomes to numbers with [0,1] via the decimal expansion:

$$X(a) = 0.a.$$

  Here we follow the convention such that $0.99999\cdots = 1$.
- What is the probability of the event $\{X = 1\}$?
  ZERO!
- $P(X = 1) = P(\{99999999\cdots\}) = \frac{1}{10} \times \frac{1}{10} \times \frac{1}{10} \times \frac{1}{10} \times \frac{1}{10} \times \cdots = 0.$
- $P(X = a) = 0.$

Let's keep throwing a ten-sided fair die (with numbers 0,1,...,9) without stopping!

- We have outcomes such as

$$a = 010310819.....$$

- Let's map these outcomes to numbers with [0,1]:$X(a) = 0.a$.

- **How do we characterize the distribution of X:** e.g., how likely is the event $\{X \leq 0.5\}$.
  $\{X \leq 0.5\}$ contains all the outcomes such that the first throw is strictly smaller than five (0,1,2,3,4) and the specific outcome (only the first throw is a five the rests are all 0's) $5000000000\cdots$:

$$P(X \leq 0.5) = 0.5 + P(\{5000000000\cdots\}) = 0.5$$

$$P(X \leq a) = a, a \in [0, 1].$$

$$P(X \leq a) = a, a \in [0, 1].$$

- Let's cut $[0, 1]$ into 10^1 grids such that each grid can be expressed as

  $$I_i = [x_{i-1,1}, x_{i,1}], i \in \{1, \cdots, 10^1\}$$

  with $x_j = j10^{-1}, j \in \{0, 1, \cdots, 10^1\}$. E.g., $i = 2$, $[0.1, 0.2]$.
  It is easy to check that

  $$P(X \in \cup_{j=1}^{i} I_j) = P(X \leq x_i) = x_i = i * 10^{-1}, i \in \{1, \cdots, 10\}$$

  and thus for each grids $P(X \in I_j) = 10^{-1}$.
  Let's approximate the probability value of $\{X \leq a\}$ with

  $$P(X \in \cup_{j=1}^{j_a} I_j) = \sum_{i=1}^{j_a} \Delta x_i, \Delta x_i = x_i - x_{i-1}, j_a = \min\{i \in \{1, \cdots, 10^1\} : a \in I_i\}.$$

- $$\left| P(X \leq x) - \sum_{i=1}^{j_a} \Delta x_i \right| \leq 10^{-1}.$$

$$P(X \leq a) = a, a \in [0, 1].$$

- Let's cut $[0, 1]$ into 10^n grids such that each grid can be expressed as

$$I_i = [x_{i-1,n}, x_{i,n}], i \in \{1, \cdots, 10^n\}$$

with $x_j = j10^{-n}, j \in \{0, 1, \cdots, 10^n\}$. E.g., $i = 2$, $[10^{-n}, 2 * 10^{-n}]$.
It is easy to check that

$$P(X \in \cup_{j=1}^{i} I_j) = P(X \leq x_i) = x_i = i * 10^n, i \in \{1, \cdots, 10^n\}$$

and thus for each grids $P(X \in I_j) = 10^{-n}$.
Let's approximate the probability value of $\{X \leq a\}$ with

$$\sum_{i=1}^{j_a} \Delta x_i, \Delta x_i = x_i - x_{i-1}, j_a = \min\{i \in \{1, \cdots, 10^n\} : a \in I_i\}.$$

- $$\left| P(X \leq a) - \sum_{i=1}^{j_a} \Delta x_i \right| \leq 10^{-n}.$$

- $$\left| P(X \leq a) - \sum_{i=1}^{j_a} \Delta x_i \right| \leq 10^{-n}.$$

-

$$F_X(a) = P(X \leq a) = a, a \in [0,1].$$

- Let's cut $[0,1]$ into 10^n grids such that each grid can be expressed as

$$I_i = [x_{i-1,n}, x_{i,n}], i \in \{1, \cdots, 10^n\}$$

with $x_j = j10^{-n}, j \in \{0, 1, \cdots, 10^n\}$. E.g., $i = 2$, $[10^{-n}, 2 * 10^{-n}]$.
It is easy to check that

$$P(X \in \cup_{j=1}^{i} I_j) = P(X \leq x_i) = x_i = i * 10^n, i \in \{1, \cdots, 10^n\}$$

and thus for each grids $P(X \in I_j) = 10^{-n}$.
Let's approximate the probability value of $\{X \leq a\}$ with

$$\sum_{i=1}^{j_a} \Delta x_i, \Delta x_i = x_{i-1}, j_a = \min\{i \in \{1, \cdots, 10^n\} : a \in I_i\}.$$

-

$$\sum_{i=1}^{j_a} \Delta x_i \to \int_0^a f_X(x) dx = a; f_X(x) = F_X'(x).$$

$$F_X(a) = P(X \leq a) = a, a \in [0, 1].$$

$$\int_{-\infty}^{a} f_X(x)dx = P(X \leq a) = a; f_X(x) = F_X'(x) = 1, x \in [0, 1] \text{ otherwise zero.}$$

- Two different ways describing the distribution of X, and particularly, there exists a probability density function f_X (integrate the (probability) *density*, we have the (probability) *mass*).

We have a probability density function f_X (integrate the (probability) density via Riemann integral, we have the (probability) value). If an interval has positive density function values, then there are uncountably many possible values in the interval taken by X.

$$\int_{-\infty}^{a} f_X(x)dx = P(X \leq a); a \in \mathbb{R}.$$

Any real-valued random variable that has a non-negative function f_X satisfying the above Riemann integral condition is a continuous real-valued random variable.

Let's compare.

- In the example, we see that we can choose $f_X(x) = F'_X(x)$ and $F'_X(x) = 1, x \in [0, 1]$.

  There are **uncountably infinite** possible values in [0,1] taken by X.

- For any discrete random variable Y, almost everywhere F'_Y is simply **zero**, there is no f_X satisfying the above identity (via Riemann integral).

  There are **at most countably many** possible values taken by Y.