Probability Theory for EOR

Discrete v.s. Continuous random variables

Discrete v.s. Continuous random variables: differences and connections

Real-valued random variables

	CDF $\mathbb{P}(Z \le z)$	PMF/PDF	Expectation $\mathbb{E}Z$
Discrete $R(X) = supp(X) =$ $\{x_1 < . < x_n < .\}$ $\mathring{R}(X) = \emptyset.$	Step function $F(x)$ increase from 0 to 1 **Jumps**	$p_X(x_i) = F(x_i) - F(x_{i-1})$ $= F(x_i) - \lim_{s \uparrow x_i} F(s)$ $F(x) = \sum_{x_i \in R(x);\, x_i \le x} p_X(x_i)$	$\sum_{x_i \in R(x);\, x_i < x_{i+1}} x_i \Delta F(x_i)$ $\sum_{x_i \in R(x);\, x_i < x_{i+1}} x_i p_X(x_i)$
		$p_X(x) \ge 0;\ \sum_{x_i \in R(X)} p_X(x_i) = 1$	$(\sum_{n=0}^{\infty} G_X(n))$
Continuous $R(Y) = supp(Y)$ $\mathring{R}(Y) \ne \emptyset$, e.g. (a, b)	Continuous $F(y)$ increase from 0 to 1 **No jumps** differentiable	$f_Y(y) = F'(y)$ $= \lim_{s \to y} \frac{F(y)-F(s)}{y-s}$ $F(y) = \int_{s<y} f_Y(s)ds$	$\int_{R(Y)} y dF(y)$ $\int_{R(Y)} y f(y) dy$
		$f_Y(y) \ge 0;\ \int_{R(Y)} f_Y(y)dy = 1$	$\left(\int_0^{\infty} G_Y(y)dy\right)$

Let X be a **non-negative integer-valued** r.v. then if its expectation exists

$$\mathbb{E}X = \sum_{i=0}^{\infty} G_X(i)$$

where G_X is the survival function of X such that $G_X(x) = 1 - F_X(x)$.

- Proof:

  $X = I_{\{X\geq 1\}} + \cdots + I_{\{X\geq n\}} + \cdots = \sum_{i\in\mathbb{N}_+} I_{\{X\geq i\}}$.

  The above holds true since X and $\sum_{i\in\mathbb{N}_+} I_{\{X\geq i\}}$ are the same function from S to the set of non-negative integers.

  Linearity of expectation, fundamental bridge, and the fact that $\{X \geq i\} = \{X > i-1\}, i \in \mathbb{N}_+$ give

  $EX = E\sum_{i=1}^{\infty} I_{\{X\geq i\}} = \sum_{i=1}^{\infty} EI_{\{X\geq i\}} = \sum_{i=1}^{\infty} P(X \geq i) = \sum_{i=0}^{\infty} P(X > i) = \sum_{i=0}^{\infty} G_X(i)$.

Let X be a **non-negative REAL-valued** (either discrete or continuous) r.v. then if its expectation exists

$$\mathbb{E}X = \int_0^{\infty} G_X(s)ds$$

where G_X is the survival function of X such that $G_X(x) = 1 - F_X(x)$.

- Proof:

  $X = \int_0^{\infty} I_{\{X>t\}}dt$.

  The above holds true since X and $\int_0^{\infty} I_{\{X>t\}}dt$ are the same function from S to the set of non-negative real values:
  when $X(s) = x$,
  $\int_0^{\infty} I_{\{X(s)>t\}}dt = \int_0^{\infty} I_{\{x>t\}}dt = x$
  since $I_{\{x>t\}} = 1$ only for $t \in [0, x]$.

  Linearity of expectation, fundamental bridge give

  $EX = E\int_0^{\infty} I_{\{X>t\}}dt = \int_0^{\infty} EI_{\{X>t\}}dt = \int_0^{\infty} P(X > t)dt = \int_0^{\infty} G_X(s)ds$.

Discussions via problems

I. what is the probability of $\{Z = z\}$?

X follows Bern(0.5) with $p_X(0) = p_X(1) = 1/2$.

- The support of X is $\{0, 1\}$ (**Finitely many**, at most countably infnite).
- $P(X = x) = 1/2$ for $x \in \{0, 1\}$, 0 otherwise.

X has $f_X(x) = 1, x \in [0, 1]$.

- The support of X is $[0, 1]$ (**Infinitely many**, uncountably infnite).
- $P(X = x) = 0$ for $x \in \mathbb{R}$. Since $P(X = x) = \lim_{\delta \downarrow 0} \int_{x-\delta}^{x+\delta} f_X(s)ds = 0$.

  Uncountably many possible values, so it is of zero probability that you choose a specific number. True for all continuous r.v.'s

  However, you could get outcomes in a neighborhood closer to the number $B_\epsilon(x) = (x - \epsilon, x + \epsilon) \subseteq [0, 1]$ (ϵ is an arbitrary and a very small positive number):

  $P(X \in (x - \epsilon, x + \epsilon))$
  $= \int_{x-\epsilon}^{x+\epsilon} f_X(s)ds = 2\epsilon.$

I. uncountably many possible values, so it is of zero probability that you get a specific number. True for all continuous r.v.'s

II. what is the probability of $\{Z_1 < Z_2 < Z_3\}$?

$X_i, i = 1, 2, 3$ follows i.i.d. Bern(1/2) with mass (PMF) $p(0) = p(1) = 1/2$.

- $P(Z_1 < Z_2 < Z_3) = 0$.
- **At least two of them would be equal to each other.**

X_i i.i.d. with density (PDF) $f(x) = 1, x \in [0, 1]$.

- First note that $X_i - X_j, i \neq j$ is also a continuous r.v. and thus $P(X_i = X_j) = P(X_i - X_j = 0) = 0$.
- **Any two of them will be equal to each other with zero probability.**
- Then one of the arbitrary order must happen $X_{a_1} < X_{a_2} < X_{a_3}$ for any permutations a_1, a_2, a_3 of $1, 2, 3$, and by symmetry

$$P(X_{a_1} < X_{a_2} < X_{a_3}) = \frac{1}{3!}.$$

Can be extended to cases of n i.i.d. continuous r.v.'s.

$$P(X_{a_1} < X_{a_2} < X_{a_3} < \cdots < X_{a_n}) = \frac{1}{n!}.$$

II. X_i i.i.d. from a continuous distribution, then

$$P(X_{a_1} < X_{a_2} < X_{a_3} < \cdots < X_{a_n}) = \frac{1}{n!}.$$

for any permutations $a_1, a_2, a_3, \cdots, a_n$ of $1, 2, 3, \cdots, n$.