Probability Theory for EOR

Expectations of functions of random variables
(with discrete random variables examples)

I

Expectation maps a subset of the collection of random variables to numbers (expectations/expected values of associated random variables).

We know some functions of random variables are still random variables,

and we look into **expectations of functions of random variables and how they are linked with (the expectations of) the original random variables.**

Law of the unconscious statistician (LOTUS)

Distributions determine expectations.

The **distribution** of X determines the **distribution** of $g(X)$.

The **distribution** of X determines the **expectation** of $g(X)$.

LOTUS:

The **distribution** of X
determines
the **expectation** of $g(X)$.

Example with discrete random r.v.'s.

If X is a discrete random variable with support $\text{supp}(X) = \{x_1, x_2, x_3. \cdots\}$, and $g(x)$ is a function from $R \to R$ such that $g(X)$ is a discrete r.v., then

$$\mathrm{E}[g(X)] = \sum_{x \in \text{supp}(X)} g(x) P(X = x) \quad (= \sum_{x \in \text{supp}(X)} g(x) \Delta F_X(x))$$

with $\Delta F_X(x) = F(x) - \lim_{y \uparrow x} F(y)$.

Proof with discrete r.v.'s:

$$
\begin{aligned}
E(g(X)) &= \sum_c cP(g(X) = c) \\
&= \sum_c c \left(\sum_{a \in \text{supp}(X): g(a) = c} P(X = a) \right) \\
&= \sum_c \sum_{a \in \text{supp}(X): g(a) = c} cP(X = a) \\
&= \sum_c \sum_{a \in \text{supp}(X): g(a) = c} g(a)P(X = a) \\
&= \sum_{a \in \text{supp}(X)} g(a)P(X = a)
\end{aligned}
$$

Example:

- Suppose a random variable X has outcomes $\{0, 1, 2, 3\}$ with probabilities given by the probability mass function $p_X(x)$.
- Now consider X^2. This has outcomes $\{0, 1, 4, 9\}$.
- The probability of seeing 9 from X^2 is the same as seeing 3 from X; ...

So

$$E[X^2] = \sum_{k=0}^{3} k^2 P(X^2 = k^2) = \sum_{k=0}^{3} k^2 p_X(k)$$

Another example:

- Suppose a random variable X has outcomes $\{-2,-1,1,2\}$ with probabilities given by the probability mass function $p_X(x)$.
- Now consider X^2. This has outcomes $\{1,4\}$, occurring with the PMF:

$$p_{X^2}(1) = P(X^2 = 1) = P(X = 1) + P(X = -1) = p_X(1) + p_X(-1)$$
$$p_{X^2}(4) = P(X^2 = 4) = P(X = 2) + P(X = -2) = p_X(2) + p_X(-2).$$

- For the expected value we have

$$\begin{aligned} E[X^2] &= \sum_{k=1,4} k p_{X^2}(k) = \sum_{j=\{-2,-1,1,2\}} j^2 p_X(j) \\ &= \sum_{j=\{-1,1\}} j^2 p_X(j) + \sum_{j=\{-2,2\}} j^2 p_X(j) \end{aligned}$$

Variance

Definition

The **variance** of an **real-valued** r.v. X (if exists) is

$$\mathrm{Var}(X) = E(X - EX)^2$$

and the **standard deviation** is

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$$

- Variance of X is essentially the expectation of a function of X, $g(X)$, with $g(x) = (x - \mu)^2, \mu = E(X)$.
- The distribution of X determines the variance of X.

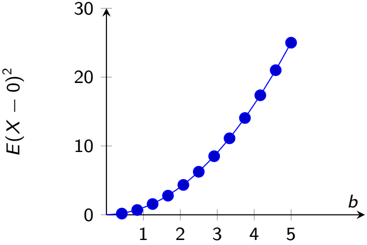
For any **real-valued** r.v. X,

$$\begin{aligned}
\mathrm{Var}(X) &= E(X - EX)^2 \\
&= E\left(X^2 - 2XE(X) + (E(X))^2\right) \\
&= E\left(X^2\right) - E\left(2XE(X)\right) + E\left((E(X))^2\right) \\
&= E\left(X^2\right) - 2E(X)E\left(X\right) + (E(X))^2 \\
&= E(X)^2 - (EX)^2
\end{aligned}$$

Example:

Mysterious door $\Rightarrow$ A: consequence X_b $\Rightarrow$ B: $X_b = b$ or $-b$ by a fair coin $\Rightarrow$ C: Expected Outcome always 0

However, it feels different when b varies from **5** to **0**. How to describe the difference?



$V(X_b) = E(X_b - 0)^2 = b^2.$

It describes the "spread" of the random variable from its "average".

Some properties of variance:

- For scalars $a, c \in \mathbb{R}$, for **any** r.v. X (if varaince well-defined):

$$\mathrm{Var}(a + cX) = \mathrm{Var}(cX) = c^2 \mathrm{Var}(X).$$

- For **independent** r.v.'s, $X_1, X_2, \cdots, X_n$ (if exists):

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}\left(X_i\right).$$

Different from the expectation, as for expectation for **any** r.v.'s $X_1, X_2, \cdots, X_n$ (if exists):

$$E(a + cX_1) = a + cE(X_1).$$
$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E\left(X_i\right).$$

Example: $X_1, X_2 \sim$i.i.d. Bern(p).

- Calculate $E(X_1), E(1 + 2X_1), E(X_1 + X_2), \mathrm{Var}(X_1), \mathrm{Var}(1 + 2X_1), \mathrm{Var}(2X_1), \mathrm{Var}(X_1 + X_2)$:
- $E(X_1) = p, E(1 + 2X_1) = 1 + 2p, E(X_1 + X_2) = 2p, \mathrm{Var}(X_1) = p(1 - p), \mathrm{Var}(1 + 2X_1) = 4p(1 - p), \mathrm{Var}(2X_1) = 4p(1 - p), \mathrm{Var}(X_1 + X_2) = 2p(1 - p)$.