भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probability and Stochastic Processes

Lecture 24: Expectations of Discrete Random Variables, Expectations of Functions of Random Variables, Expectations of Continuous Random Variables

Karthik P. N.
Assistant Professor, Department of AI
Email: pnkarthik@ai.iith.ac.in

31 October 2025

Property 6 (Linearity of Expectations)

For any two random variables X, Y,

$$\mathbb{E}[\alpha X] = \alpha\,\mathbb{E}[X] \quad \forall \alpha \in \mathbb{R}, \qquad\qquad \mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y],$$

whenever the right-hand sides are well-defined

Proof:

- Suppose X and Y are **simple** RVs with canonical representations

$$X = \sum_{i=1}^{m} a_i\,\mathbf{1}_{A_i}, \qquad\qquad Y = \sum_{j=1}^{n} b_j\,\mathbf{1}_{B_j},$$

  where $\{A_1, \ldots, A_m\}$ and $\{B_1, \ldots, B_m\}$ **each** partition Ω
- Then, $X + Y$ has the representation

$$X + Y = \sum_{i=1}^{m}\sum_{j=1}^{n}(a_i + b_j)\,\mathbf{1}_{A_i \cap B_j}$$

- We may combine similar $(a_i + b_j)$ terms to bring $X + Y$ to canonical form

Property 6 (Linearity of Expectations)

For any two random variables X, Y,

$$\mathbb{E}[\alpha X] = \alpha\,\mathbb{E}[X] \quad \forall \alpha \in \mathbb{R}, \qquad\qquad \mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y],$$

whenever the right-hand sides are well-defined

Proof:

- We then have

$$\mathbb{E}[X+Y] = \sum_{i=1}^{m}\sum_{j=1}^{n}(a_i + b_j)\,\mathbb{P}(A_i \cap B_j) = \sum_{i=1}^{m} a_i \underbrace{\sum_{j=1}^{n}\mathbb{P}(A_i \cap B_j)}_{=\mathbb{P}(A_i)} + \sum_{j=1}^{n} b_j \underbrace{\sum_{i=1}^{m}\mathbb{P}(A_i \cap B_j)}_{=\mathbb{P}(B_j)}$$

$$= \sum_{i=1}^{m} a_i\,\mathbb{P}(A_i) + \sum_{j=1}^{n} b_j\,\mathbb{P}(B_j) = \mathbb{E}[X] + \mathbb{E}[Y].$$

Property 6 (Linearity of Expectations)

For any two random variables X, Y,

$$\mathbb{E}[\alpha X] = \alpha\, \mathbb{E}[X] \quad \forall \alpha \in \mathbb{R}, \qquad\qquad \mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y],$$

whenever the right-hand sides are well-defined

Proof:

- Suppose X and Y are **non-negative** random variables
- Let $\{X_n\}$ and $\{Y_n\}$ be the associated quantization sequences, with

$$\forall\, \omega \in \Omega, \quad \lim_{n\to\infty} X_n(\omega) + Y_n(\omega) = X(\omega) + Y(\omega), \qquad\qquad \lim_{n\to\infty} \mathbb{E}[X_n + Y_n] = \mathbb{E}[X+Y],$$

- Then,

$$\mathbb{E}[X+Y] = \lim_{n\to\infty} \mathbb{E}[X_n + Y_n] = \lim_{n\to\infty} \mathbb{E}[X_n] + \lim_{n\to\infty} \mathbb{E}[Y_n] = \mathbb{E}[X] + \mathbb{E}[Y].$$

Property 6 (Linearity of Expectations)

For any two random variables X, Y,

$$\mathbb{E}[\alpha X] = \alpha\,\mathbb{E}[X] \quad \forall \alpha \in \mathbb{R}, \qquad\qquad \mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y],$$

whenever the right-hand sides are well-defined

Proof:

- Suppose X and Y are **arbitrary** random variables
- Define X_+, X_-, Y_+, Y_- as usual
- We have

$$X = X_+ - X_-, \qquad Y = Y_+ - Y_-, \qquad X + Y = X_+ - X_- + Y_+ - Y_-$$

- Then,

$$\begin{aligned}\mathbb{E}[X+Y] = \mathbb{E}[X_+ - X_- + Y_+ - Y_-] &= \mathbb{E}[X_+] - \mathbb{E}[X_-] + \mathbb{E}[Y_+] - \mathbb{E}[Y_-] \\ &= \mathbb{E}[X] + \mathbb{E}[Y].\end{aligned}$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Property 7

For any $x \in \mathbb{R}$,

$$\mathbb{E}[X \cdot \mathbf{1}_{\{X \geq x\}}] \geq x \cdot \mathbb{P}(\{X \geq x\}),$$
$$\mathbb{E}[X \cdot \mathbf{1}_{\{X \leq x\}}] \leq x \cdot \mathbb{P}(\{X \leq x\}),$$
$$\mathbb{E}[X \cdot \mathbf{1}_{\{X = x\}}] = x \cdot \mathbb{P}(\{X = x\}).$$

Proof:

- For any $x \in \mathbb{R}$,

$$\mathbb{E}[X \cdot \mathbf{1}_{\{X \geq x\}}] = \int_{\Omega} X(\omega) \cdot \mathbf{1}_{\{X \geq x\}}(\omega)\, \mathrm{d}\mathbb{P}(\omega) = \int_{\{X \geq x\}} X(\omega)\, \mathrm{d}\mathbb{P}(\omega) \geq x \int_{\{X \geq x\}} \mathrm{d}\mathbb{P}(\omega) = x\, \mathbb{P}(\{X \geq x\}).$$

- Along similar lines,

$$\mathbb{E}[X \cdot \mathbf{1}_{\{X \leq x\}}] = \int_{\Omega} X(\omega) \cdot \mathbf{1}_{\{X \leq x\}}(\omega)\, \mathrm{d}\mathbb{P}(\omega) = \int_{\{X \leq x\}} X(\omega)\, \mathrm{d}\mathbb{P}(\omega) \leq x \int_{\{X \leq x\}} \mathrm{d}\mathbb{P}(\omega) = x\, \mathbb{P}(\{X \leq x\}),$$

$$\mathbb{E}[X \cdot \mathbf{1}_{\{X = x\}}] = \int_{\Omega} X(\omega) \cdot \mathbf{1}_{\{X = x\}}(\omega)\, \mathrm{d}\mathbb{P}(\omega) = \int_{\{X = x\}} X(\omega)\, \mathrm{d}\mathbb{P}(\omega) = x \int_{\{X = x\}} \mathrm{d}\mathbb{P}(\omega) = x\, \mathbb{P}(\{X = x\}).$$

Monotone Convergence Theorem (MCT)

Suppose that X and $X_1, X_2, \ldots$ are **non-negative** RVs such that

$$\forall\, \omega \in \Omega, \quad X_1(\omega) \le X_2(\omega) \le \cdots, \qquad \lim_{n\to\infty} X_n(\omega) = X(\omega).$$

Then,

$$\mathbb{E}[X_1] \le \mathbb{E}[X_2] \le \cdots, \qquad \lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X].$$

Remarks:

- Because $X(\omega) = \lim_{n\to\infty} X_n(\omega)$ for each $\omega \in \Omega$, a compact way of stating MCT is as follows:

  $$\mathbb{E}\left[\lim_{n\to\infty} X_n\right] = \lim_{n\to\infty} \mathbb{E}[X_n].$$

  That is, **MCT is a statement about interchanging limits and expectation**.
- **An important corollary of MCT:**

  If $X_1, X_2, \ldots$ are a sequence of **non-negative** RVs, then

  $$\mathbb{E}\left[\sum_{n\in\mathbb{N}} X_n\right] = \sum_{n\in\mathbb{N}} \mathbb{E}[X_n] \qquad \text{(interchange of expectation and infinite sum for non-neg. RVs).}$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Property 8

If X is a **non-negative** RV with $0 < \mathbb{E}[X] < +\infty$, then the function $\mathbb{Q}^{(X)} : \mathscr{F} \to [0, 1]$ defined via

$$\mathbb{Q}^{(X)}(A) = \frac{\int_A X \, d\mathbb{P}}{\int_\Omega X \, d\mathbb{P}} = \frac{\mathbb{E}[X \cdot \mathbf{1}_A]}{\mathbb{E}[X]}, \qquad A \in \mathscr{F},$$

is a probability measure on $(\Omega, \mathscr{F})$. Consequently, for any $A, B \in \mathscr{F}$ such that $A \subseteq B$,

$$\mathbb{E}[X \cdot \mathbf{1}_A] = \int_A X \, d\mathbb{P} \leq \int_B X \, d\mathbb{P} = \mathbb{E}[X \cdot \mathbf{1}_B].$$

Proof:

- Clearly, $\mathbb{Q}^{(X)}(\Omega) = 1$
- For any $A \in \mathscr{F}$,

$$\mathbb{Q}^{(X)}(A^\complement) = \frac{\mathbb{E}[X \cdot \mathbf{1}_{A^\complement}]}{\mathbb{E}[X]} = \frac{\mathbb{E}[X \cdot (1 - \mathbf{1}_A)]}{\mathbb{E}[X]} = \frac{\mathbb{E}[X - X\,\mathbf{1}_A]}{\mathbb{E}[X]} = \frac{\mathbb{E}[X] - \mathbb{E}[X \cdot \mathbf{1}_A]}{\mathbb{E}[X]} = 1 - \mathbb{Q}^{(X)}(A).$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Property 8

If X is a **non-negative** RV with $0 < \mathbb{E}[X] < +\infty$, then the function $\mathbb{Q}^{(X)} : \mathscr{F} \to [0, 1]$ defined via

$$\mathbb{Q}^{(X)}(A) = \frac{\int_A X \, d\mathbb{P}}{\int_\Omega X \, d\mathbb{P}} = \frac{\mathbb{E}[X \cdot \mathbf{1}_A]}{\mathbb{E}[X]}, \qquad A \in \mathscr{F},$$

is a probability measure on $(\Omega, \mathscr{F})$. Consequently, for any $A, B \in \mathscr{F}$ such that $A \subseteq B$,

$$\mathbb{E}[X \cdot \mathbf{1}_A] = \int_A X \, d\mathbb{P} \leq \int_B X \, d\mathbb{P} = \mathbb{E}[X \cdot \mathbf{1}_B].$$

Proof:

- For any mutually disjoint $A_1, A_2, \ldots \in \mathscr{F}$,

$$\mathbb{Q}^{(X)}\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \frac{\mathbb{E}[X \cdot \mathbf{1}_{\bigsqcup_{n\in\mathbb{N}} A_n}]}{\mathbb{E}[X]} = \frac{\mathbb{E}\left[\sum_{n\in\mathbb{N}} X \, \mathbf{1}_{A_n}\right]}{\mathbb{E}[X]} \overset{\text{MCT}}{=} \sum_{n\in\mathbb{N}} \frac{\mathbb{E}[X \, \mathbf{1}_{A_n}]}{\mathbb{E}[X]} = \sum_{n\in\mathbb{N}} \mathbb{Q}^{(X)}(A_n).$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Some Corollaries of Properties

- If $X \sim \text{Exponential}(\lambda)$ for some fixed λ, then

$$\mathbb{P}(\{X \geq 0\}) = 1 \qquad \implies \qquad \mathbb{E}[X] \geq 0.$$

- If $X \sim \text{Uniform}[a, b]$ for some fixed $a, b \in \mathbb{R}$, $a < b$, then,

$$\mathbb{P}(\{X \in [a, b]\}) = 1 \qquad \implies \qquad a \leq \mathbb{E}[X] \leq b.$$

- Similar statements can be made for other distributions such as Bernoulli, Binomial, Poisson, Geometric, Gamma, Beta, Rayleigh, etc.

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Expectations of Discrete Random Variables

Expectation of a Discrete Random Variable

Lemma (Expectation of a Discrete Random Variable)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a **discrete** RV with $\mathbb{P}_X(E) = 1$ for some countable set $E = \{e_1, e_2, \ldots\}$, where $e_\ell \in \mathbb{R}$ for all $\ell \in \mathbb{N}$.
If p_X denotes the PMF of X, then

$$\mathbb{E}[X] \quad = \quad \int_\Omega X \, \mathrm{d}\mathbb{P} \quad = \quad \sum_{\ell \in \mathbb{N}} e_\ell \, p_X(e_\ell),$$

provided the summation in the right-most term is well-defined.

Proof:

- Suppose X is **simple** with the canonical representation

$$X = \sum_{\ell=1}^{n} e_\ell \, \mathbf{1}_{A_\ell}, \qquad A_\ell = \{X = e_\ell\},$$

where $e_1, \ldots, e_n \geq 0$ are distinct, and $A_1, \ldots, A_n$ partition Ω

Expectation of a Discrete Random Variable

Lemma (Expectation of a Discrete Random Variable)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a **discrete** RV with $\mathbb{P}_X(E) = 1$ for some countable set $E = \{e_1, e_2, \ldots\}$, where $e_\ell \in \mathbb{R}$ for all $\ell \in \mathbb{N}$.
If p_X denotes the PMF of X, then

$$\mathbb{E}[X] \quad = \quad \int_\Omega X \, \mathrm{d}\mathbb{P} \quad = \quad \sum_{\ell \in \mathbb{N}} e_\ell \, p_X(e_\ell),$$

provided the summation in the right-most term is well-defined.

Proof:

- Then, expectation of X is given by

$$\mathbb{E}[X] = \sum_{\ell=1}^{n} e_\ell \, \mathbb{P}(A_\ell) = \sum_{\ell=1}^{n} e_\ell \, \mathbb{P}(\{X = e_\ell\}) = \sum_{\ell=1}^{n} e_\ell \, p_X(e_\ell).$$

Expectation of a Discrete Random Variable

Lemma (Expectation of a Discrete Random Variable)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a **discrete** RV with $\mathbb{P}_X(E) = 1$ for some countable set $E = \{e_1, e_2, \ldots\}$, where $e_\ell \in \mathbb{R}$ for all $\ell \in \mathbb{N}$.
If p_X denotes the PMF of X, then

$$\mathbb{E}[X] = \int_\Omega X \, \mathrm{d}\mathbb{P} = \sum_{\ell \in \mathbb{N}} e_\ell \, p_X(e_\ell),$$

provided the summation in the right-most term is well-defined.

Proof:

- Suppose X is non-negative. Then, X can be represented as

$$X = \sum_{\ell \in \mathbb{N}} e_\ell \, \mathbf{1}_{A_\ell}, \qquad A_\ell = \{X = e_\ell\},$$

  where $e_1, e_2, \ldots \geq 0$ are distinct, and $A_1, A_2, \ldots$ partition Ω

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Expectation of a Discrete Random Variable

Lemma (Expectation of a Discrete Random Variable)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a **discrete** RV with $\mathbb{P}_X(E) = 1$ for some countable set $E = \{e_1, e_2, \ldots\}$, where $e_\ell \in \mathbb{R}$ for all $\ell \in \mathbb{N}$.
If p_X denotes the PMF of X, then

$$\mathbb{E}[X] = \int_\Omega X \, \mathrm{d}\mathbb{P} = \sum_{\ell \in \mathbb{N}} e_\ell \, p_X(e_\ell),$$

provided the summation in the right-most term is well-defined.

Proof:

- Then, the expectation of X is given by

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{\ell \in \mathbb{N}} e_\ell \, \mathbf{1}_{A_\ell}\right] \overset{\text{MCT}}{=} \sum_{\ell \in \mathbb{N}} \mathbb{E}[e_\ell \, \mathbf{1}_{A_\ell}] = \sum_{\ell \in \mathbb{N}} e_\ell \, \mathbb{P}(A_\ell) = \sum_{\ell \in \mathbb{N}} e_\ell \, \mathbb{P}(\{X = e_\ell\}) = \sum_{\ell \in \mathbb{N}} e_\ell \, p_X(e_\ell).$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Expectation of a Discrete Random Variable

Lemma (Expectation of a Discrete Random Variable)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a **discrete** RV with $\mathbb{P}_X(E) = 1$ for some countable set $E = \{e_1, e_2, \ldots\}$, where $e_\ell \in \mathbb{R}$ for all $\ell \in \mathbb{N}$.
If p_X denotes the PMF of X, then

$$\mathbb{E}[X] = \int_\Omega X \, \mathrm{d}\mathbb{P} = \sum_{\ell \in \mathbb{N}} e_\ell \, p_X(e_\ell),$$

provided the summation in the right-most term is well-defined.

Proof:

- Suppose X is an arbitrary discrete RV, i.e., some of the $\{e_\ell\}$ could be negative
- Write X_+ and X_- as

$$X_+ = \sum_{\ell : e_\ell \geq 0} e_\ell \, \mathbf{1}_{A_\ell}, \qquad X_- = \sum_{\ell : e_\ell < 0} -e_\ell \, \mathbf{1}_{A_\ell}, \qquad A_\ell = \{X = e_\ell\}.$$

Expectation of a Discrete Random Variable

Lemma (Expectation of a Discrete Random Variable)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a **discrete** RV with $\mathbb{P}_X(E) = 1$ for some countable set $E = \{e_1, e_2, \ldots\}$, where $e_\ell \in \mathbb{R}$ for all $\ell \in \mathbb{N}$.
If p_X denotes the PMF of X, then

$$\mathbb{E}[X] = \int_\Omega X \, \mathrm{d}\mathbb{P} = \sum_{\ell \in \mathbb{N}} e_\ell \, p_X(e_\ell),$$

provided the summation in the right-most term is well-defined.

Proof:

- Then, $\mathbb{E}[X_+]$ and $\mathbb{E}[X_-]$ can be written as

$$\mathbb{E}[X_+] = \sum_{\ell : e_\ell \geq 0} e_\ell \, \mathbb{P}(\{X = e_\ell\}) = \sum_{\ell : e_\ell \geq 0} e_\ell \, p_X(e_\ell), \qquad \mathbb{E}[X_-] = \sum_{\ell : e_\ell < 0} -e_\ell \, \mathbb{P}(\{X = e_\ell\}) = -\sum_{\ell : e_\ell < 0} e_\ell \, p_X(e_\ell)$$

- $\mathbb{E}[X] = \mathbb{E}[X_+] - \mathbb{E}[X_-]$, provided both $\mathbb{E}[X_+]$ and $\mathbb{E}[X_-]$ are not $+\infty$

Examples

Compute $\mathbb{E}[X]$ for the following cases.

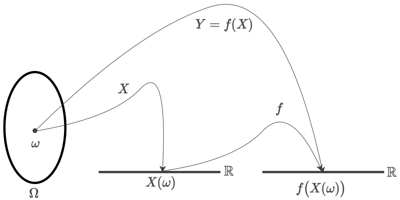
- $X \sim \text{Unif}\big(\{1, \ldots, n\}\big)$.
- $X \sim \text{Geom}(p), \qquad p \in (0, 1)$.
- $X \sim \text{Poisson}(\lambda), \qquad \lambda > 0$.
- $\mathbb{P}(\{X = k\}) = \dfrac{c}{k^2}, \qquad k \in \mathbb{N}$.
- $\mathbb{P}(\{X = k\}) = \dfrac{c}{k^2}, \qquad k \in \mathbb{Z}$.

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Expectations of Functions of Random Variables

Key Question: How to compute $\mathbb{E}[f(X)]$?

Expectation Over Different Spaces



Theorem (Expectations of Functions of Random Variables)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a RV, and let $f : \mathbb{R} \to \mathbb{R}$ be **measurable**. Let $Y = f(X)$.
Then, we have

$$\mathbb{E}[Y] = \mathbb{E}[f(X)] = \int_\Omega f(X(\omega))\, \mathrm{d}\mathbb{P}(\omega) = \int_\mathbb{R} f(x)\, \mathrm{d}\mathbb{P}_X(x) = \int_\mathbb{R} y\, \mathrm{d}\mathbb{P}_Y(y).$$

Proof of Theorem

- Suppose f is **simple** with a **finite range**, say $\mathrm{Range}(f) = \{y_1, \ldots, y_n\}$
- Then, $Y = f(X)$ is a **simple** RV having the canonical representation

$$Y(\omega) = f(X(\omega)) = \sum_{i=1}^{n} y_i \, \mathbf{1}_{A_i}(\omega), \qquad A_i = \{Y = y_i\} = \{f(X) = y_i\} = X^{-1}\left(f^{-1}(\{y_i\})\right),$$

  where $y_1, \ldots, y_n \geq 0$ are distinct and $A_1, \ldots, A_n$ partition Ω
- We then have

$$\mathbb{E}[f(X)] = \int_{\Omega} f(X(\omega)) \, \mathrm{d}\mathbb{P}(\omega) = \sum_{i=1}^{n} y_i \, \mathbb{P}(A_i) = \sum_{i=1}^{n} y_i \, \mathbb{P}(\{f(X) = y_i\}). \tag{1}$$

- $f : \mathbb{R} \to \mathbb{R}$ can also be represented in canonical form as

$$f(x) = \sum_{i=1}^{n} y_i \, \mathbf{1}_{B_i}(x), \qquad B_i = \{x' : f(x') = y_i\} = f^{-1}(\{y_i\}),$$

  where $B_1, \ldots, B_n$ partition $\mathbb{R}$
- Then, we have

$$\int_{\mathbb{R}} f(x) \, \mathrm{d}\mathbb{P}_X(x) = \sum_{i=1}^{n} y_i \, \mathbb{P}_X(B_i) = \sum_{i=1}^{n} y_i \, \mathbb{P}(\{X \in B_i\}) = \sum_{i=1}^{n} y_i \, \mathbb{P}(X \in f^{-1}(\{y_i\})) = \sum_{i=1}^{n} y_i \, \mathbb{P}(\{f(X) = y_i\}). \tag{2}$$

- Trivially, the function $y \mapsto \sum_{i=1}^{n} y_i \, \mathbf{1}_{\{y_i\}}(y)$ is a simple function from $\mathbb{R} \to \mathbb{R}$, and its expectation with respect to $\mathbb{P}_Y$ is given by

$$\sum_{i=1}^{n} y_i \, \mathbb{P}_Y(\{y_i\}) = \sum_{i=1}^{n} y_i \, \mathbb{P}(\{Y = y_i\}) = \sum_{i=1}^{n} y_i \, \mathbb{P}(\{f(X) = y_i\}). \tag{3}$$