भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probability and Stochastic Processes

Expectations over Different Spaces, Absolute Continuity of Measures, Radon–Nikodym Theorem, Expectations of Continuous Random Variables, Variance, Covariance, Uncorrelatedness and Independence

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

14/17 October 2024

Expectation Over Different Spaces

Expectation Over Different Spaces

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a discrete random variable w.r.t. $\mathscr{F}$.
Let $g : \mathbb{R} \to \mathbb{R}$ be Borel-measurable.
Let $Y = g(X)$.

Theorem (Expectation Over Different Spaces)

We have

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

Proof of Theorem – 1

Suppose g is simple, and $\mathrm{Range}(g) = \{y_1, \ldots, y_n\}$.

- $Y = g(X)$ is a simple random variable taking values $y_1, \ldots, y_n \geq 0$
- We then have

$$\int_{\mathbb{R}} y \, \mathrm{d}\mathbb{P}_Y = \sum_{i=1}^{n} y_i \, \mathbb{P}_Y(\{y_i\})$$

$$= \sum_{i=1}^{n} y_i \, \mathbb{P}(\{\omega \in \Omega : Y(\omega) = y_i\})$$

$$\int_{\Omega} g(X) \, \mathrm{d}\mathbb{P} = \sum_{i=1}^{n} y_i \, \mathbb{P}(\{\omega \in \Omega : g(X(\omega)) = y_i\})$$

$$\int_{\mathbb{R}} g \, \mathrm{d}\mathbb{P}_X = \sum_{i=1}^{n} y_i \, \mathbb{P}_X(\{x \in \mathbb{R} : g(x) = y_i\})$$

$$= \sum_{i=1}^{n} y_i \, \mathbb{P}_X(g^{-1}(\{y_i\}))$$

$$= \sum_{i=1}^{n} y_i \, \mathbb{P}(\{\omega \in \Omega : X(\omega) \in g^{-1}(\{y_i\})\})$$

$$= \sum_{i=1}^{n} y_i \, \mathbb{P}(\{\omega \in \Omega : g(X(\omega)) = y_i\})$$

Proof of Theorem – 2

Suppose g is non-negative

- $Y = g(X)$ is a non-negative random variable
- There exists a sequence of simple functions $\{g_n\}_{n=1}^{\infty}$ such that $g_n \uparrow g$ pointwise
- $Y_n = g_n(X) \uparrow Y$ pointwise, Y_n simple for all n
- We have

$$\begin{aligned}
\int_\Omega Y \,\mathrm{d}\mathbb{P} &\overset{\text{MCT}}{=} \lim_{n\to\infty} \int_\Omega Y_n \,\mathrm{d}\mathbb{P} \\
&= \lim_{n\to\infty} \int_\Omega g_n(X) \,\mathrm{d}\mathbb{P} \\
&= \lim_{n\to\infty} \int_\mathbb{R} g_n \,\mathrm{d}\mathbb{P}_X \\
&\overset{\text{MCT}}{=} \int_\mathbb{R} g \,\mathrm{d}\mathbb{P}
\end{aligned}$$

Expectations of Continuous Random Variables

Absolute Continuity of Measures

Consider a measurable space $(\Omega, \mathscr{F})$.
Let $\mu : \mathscr{F} \to [0, +\infty]$ and $\nu : \mathscr{F} \to [0, +\infty]$ be two measures.

Definition (Absolute Continuity of Measures)

We say ν is absolutely continuous with respect to μ if

$$\mu(A) = 0 \implies \nu(A) = 0.$$

Notation: $\nu \ll \mu$.

Remark: The above definition applies to probability measures also

Indian Institute of Technology Hyderabad

Radon–Nikodym Theorem

Consider a measurable space $(\Omega, \mathscr{F})$.
Let $\mu : \mathscr{F} \to [0, +\infty]$ and $\nu : \mathscr{F} \to [0, +\infty]$ be two measures.

Theorem (Radon–Nikodym Theorem)

Suppose that $\nu \ll \mu$.
Then, there exists a non-negative, measurable function $f : \Omega \to [0, +\infty]$ such that

$$\nu(A) = \int_A f \, \mathrm{d}\mu = \int_\Omega f \, \mathbf{1}_A \, \mathrm{d}\mu, \qquad \forall A \in \mathscr{F}.$$

Notation: $f = \dfrac{\mathrm{d}\nu}{\mathrm{d}\mu}$.

Continuous Random Variable – New Definition

Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let $X : \Omega \to \mathbb{R}$ be a random variable with respect to $\mathscr{F}$.

$$
\begin{tikzcd}
(\Omega, \mathscr{F}, \mathbb{P}) \arrow[r, "X"] & (\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathbb{P}_X) \arrow[d] \\
X^{-1}(B) \arrow[u] & B \arrow[l, "\text{pre-image}"']
\end{tikzcd}
$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Continuous Random Variable – New Definition

$$\begin{tikzcd}
(\Omega, \mathscr{F}, \mathbb{P}) \arrow[rr, "X"] & & (\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathbb{P}_X) \arrow[d] \\
X^{-1}(B) \arrow[u] & & B \arrow[ll, "\text{pre-image}"']
\end{tikzcd}$$

Definition (Continuous Random Variable)

A random variable X is said to be continuous if $\mathbb{P}_X \ll \lambda$, where λ is the Lebesgue measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.
Then, by the Radon–Nikodym theorem, there exists a non-negative, measurable function, say $f : \mathbb{R} \to [0, +\infty]$, such that

$$\mathbb{P}_X(A) = \int_A f \, \mathrm{d}\lambda, \qquad A \in \mathscr{B}(\mathbb{R}).$$

The function f is called the probability density function (PDF) of X.

Expectation of a Continuous Random Variable

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let $X : \Omega \to \mathbb{R}$ be a continuous random variable w.r.t. $\mathscr{F}$, with PDF f_X.

Theorem (Expectation for Continuous Random Variables)

Suppose that $g : \mathbb{R} \to \mathbb{R}$ is measurable. Then,

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g \, f_X \, \mathrm{d}\lambda.$$

In particular,

$$\mathbb{E}[X] = \int_{\mathbb{R}} x \, f_X \, \mathrm{d}\lambda$$

Proof of Theorem – 1

Assume that g is simple and takes values $y_1, \ldots, y_n \geq 0$.

- We have

$$\begin{aligned}
\mathbb{E}[g(X)] &= \int_{\mathbb{R}} g \, \mathrm{d}\mathbb{P}_X \\
&= \sum_{i=1}^{n} y_i \, \mathbb{P}_X(\underbrace{\{x \in \mathbb{R} : g(x) = y_i\}}_{B_i}) \\
&= \sum_{i=1}^{n} y_i \, \mathbb{P}_X(B_i) \\
&\overset{\text{R.N.Thm}}{=} \sum_{i=1}^{n} y_i \int_{B_i} f_X \, \mathrm{d}\lambda \qquad = \sum_{i=1}^{n} \int_{\mathbb{R}} y_i \mathbf{1}_{B_i} f_X \, \mathrm{d}\lambda = \int_{\mathbb{R}} \sum_{i=1}^{n} y_i \mathbf{1}_{B_i} f_X \, \mathrm{d}\lambda = \int_{\mathbb{R}} g f_X \, \mathrm{d}\lambda.
\end{aligned}$$

Proof of Theorem – 2

Assume that g is non-negative.

- There exists a sequence of simple functions $g_n \uparrow g$ pointwise
- We have

$$
\begin{aligned}
\mathbb{E}[g(X)] &\overset{\text{MCT}}{=} \lim_{n\to\infty} \mathbb{E}[g_n(X)] \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} g_n f_X \, \mathrm{d}\lambda \\
&\overset{\text{MCT}}{=} \int_{\mathbb{R}} \lim_{n\to\infty} g_n f_X \, \mathrm{d}\lambda \qquad (\text{because } g_n f_X \uparrow g f_X \text{ pointwise, as } f_X \geq 0) \\
&= \int_{\mathbb{R}} g f_X \, \mathrm{d}\lambda.
\end{aligned}
$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Examples

- Suppose $X \sim \text{Exponential}(\mu)$. Compute $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$.
- Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$. Compute $\mathbb{E}[X]$, $\mathbb{E}[X^2]$, and $\mathbb{E}[(X-\mu)^3]$.
- Suppose $f_X(x) = \dfrac{1}{\pi} \cdot \dfrac{1}{1+x^2}, \quad x \in \mathbb{R}$.
  Compute $\mathbb{E}[X]$.

Exercises

Compute $\mathbb{E}[X], \mathbb{E}[X^2]$ for each of the following cases:

- $X \sim \mathrm{Ber}(p)$.
- $X \sim \mathrm{Poisson}(\lambda)$?
- $X \sim \mathrm{Unif}([a,b])$?

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Variance, Covariance, and Correlation

Variance

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be random variable with respect to $\mathscr{F}$.
Let $\mathbb{E}[X]$ be well defined (i.e., not of the form $\infty - \infty$).

Definition (Variance)

The variance of X is defined as

$$\operatorname{Var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2].$$

Remarks:

- $\operatorname{Var}(X) \geq 0$.
- The quantity $\sigma_X = \sqrt{\operatorname{Var}(X)}$ is called the standard deviation of X.

A Result on Zero Variance

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be random variable with respect to $\mathscr{F}$.
Let $\mathbb{E}[X]$ be well defined (i.e., not of the form $\infty - \infty$).

Lemma (Zero Variance)

The variance of X is zero if and only

$$\mathbb{P}(\{X = c\}) = 1 \qquad \text{for some constant } c.$$

An Alternative Expression for Variance

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be random variable with respect to $\mathscr{F}$.

Alternative Expression for Variance

Let $\mathbb{E}[X]$ be well defined (i.e., not of the form $\infty - \infty$).

1. If $\left|\mathbb{E}[X]\right| = +\infty$, then $\mathrm{Var}(X) = +\infty$.
2. If $\left|\mathbb{E}[X]\right| < +\infty$, then

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \big(\mathbb{E}[X]\big)^2.$$

Remark: Because $\mathrm{Var}(X) \geq 0$, we always have $\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Examples

- Compute the variance of $X \sim \mathrm{Ber}(p)$.
- What is the variance of $X \sim \mathrm{Poisson}(\lambda)$?
- What is the variance of $X \sim \mathrm{Unif}([a, b])$?
- What is the variance of $X \sim \mathrm{Exponential}(\mu)$?
- What is the variance of $X \sim \mathcal{N}(\mu, \sigma^2)$?
- Give an example of a random variable X for which $\left|\mathbb{E}[X]\right| < +\infty$, but $\mathrm{Var}(X) = +\infty$.

Covariance

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X, Y be random variables with respect to $\mathscr{F}$.
Let $\mathbb{E}[X], \mathbb{E}[Y]$ be well defined (i.e., not of the form $\infty - \infty$).

Definition (Covariance)

The covariance of X and Y is defined as

$$\mathrm{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])],$$

provided the expectation on the right-hand side is well defined (i.e., not $\infty - \infty$).
Furthermore,

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y],$$

provided the right-hand side is not of the form $\infty - \infty$.

Uncorrelated Random Variables

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X, Y be random variables with respect to $\mathscr{F}$.
Let $\mathbb{E}[X], \mathbb{E}[Y]$ be well defined (i.e., not of the form $\infty - \infty$).

Definition (Uncorrelated Random Variables)

X and Y are said to be uncorrelated if

$$\mathrm{Cov}(X, Y) = 0.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Uncorrelatedness and Independence

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X, Y be random variables with respect to $\mathscr{F}$.
Let $\mathbb{E}[X], \mathbb{E}[Y]$ be well defined (i.e., not of the form $\infty - \infty$).

Theorem (Uncorrelatedness and Independence)

If $X \perp\!\!\!\perp Y$, then
$$\mathrm{Cov}(X, Y) = 0.$$

The converse is not true in general.