Stochastic Processes

PSTAT160A Fall 2025

John Inston

2025-10-22

Contents

1 Preliminary Material

1.1 Probability Spaces

Recall from our study of that a *probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ consists of a *sample space* Ω of all possible outcomes, a *σ-algebra* of measurable events and a *probability measure* $\mathbb{P}$ satisfying:

1. $\mathbb{P}(\emptyset) = 0 \leq \mathbb{P}(E) \leq 1 = \mathbb{P}(\Omega)$ for all $E \in \mathcal{F}$;
2. $\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} E_n\right) = \sum_n \mathbb{P}(E_n)$ for mutually exclusive $E_n \in \mathcal{F}$.

The *conditional probability* of E given F assuming $\mathbb{P}(F) > 0$ is

$$\mathbb{P}(E|F) := \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}.$$

Bayes formula states that for partition $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{F}$ and $E \in \mathcal{F}$ we have

$$\mathbb{P}(F_k|E) = \frac{\mathbb{P}(E|F_k)\mathbb{P}(F_k)}{\sum_n \mathbb{P}(E|F_n)\mathbb{P}(F_n)},$$

for all $k = 1, \dots, n$.

We say that two events E and F are *independent* if

$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F).$$

We say countable events $\{E_n\}_{n\in\mathbb{N}}$ are independent if for every $l \in \mathbb{N}$ and every subset $\mathbb{E}_{k_1}, \ldots, E_{k_l}$ of these events we have that

$$\mathbb{P}\left(\bigcap_{i=1}^{l} E_{k_i}\right) = \prod_{i=1}^{l} \mathbb{P}(E_{k_i}).$$

1.2 Random Variables

A *random variable* $X : \Omega \to \mathbb{R}$ is a measurable function mapping from the sample space to the reals. The *cumulative distribution function* of a random variable is defined by the nondecreasing, right-continuous function with left limits

$$F_X(x) := \mathbb{P}(X \leq x)$$

where this function naturally satisfies

$$\lim_{x\to\infty} F_X(x) = 1$$
$$\lim_{x\to-\infty} F_X(x) = 0.$$

If a random variable X takes countably many outcomes it is called *discrete* and is described by its *probability mass function*

$$p_X(x) = \mathbb{P}(X = x).$$

If a random variable X takes uncountably many outcomes it is called *continuous* in which case there exists a nonnegative function $f_X(\cdot)$ such that for $x \in \mathbb{R}$

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(u)\ du,$$

where $f_X(\cdot)$ is called the *probability density function*.

The expectation of a measurable function $g(\cdot)$ random variable is given by

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x)\ dF_X(x).$$

This can then be written as a weighted sum for discrete random variables

$$\mathbb{E}[g(X)] = \sum_{x} g(x) \cdot \mathbb{P}(X = x),$$

or as an integral with respect to the density function for continuous random variables

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) f_X(x)\ dx.$$

Similarly, when we have more than one random variables $X = (X_i)_{i=1}^n$ we can define the joint distribution as

$$F_X(x) := \mathbb{P}(X_i \leq x_i,\ \forall i \in \{1, \dots, n\}),$$

with similar results for joint densities and mass *functions*.

Example: Consider a Poisson random variable $X \sim \mathcal{P}(\lambda)$ which has probability mass function

$$p_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}; \quad k = 0, 1, 2, \dots .$$

The expectation of X can be computed as

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=0}^{\infty} k \cdot p_X(k) \\
&= \sum_{k=1}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!} \\
&= e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} \\
&= e^{-\lambda}\lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= e^{-\lambda}\lambda e^{\lambda} \\
&= \lambda.
\end{aligned}$$

1.3 Probability Generating Functions

Generating functions are power series expansions that allow us to quickly generate values of interest about particular functions.

> **Definition (Probability Generating Function):** For discrete random variable X we can define the *probability generating function* as
>
> $$P_X(t) := \mathbb{E}[t^X]$$
>
> when it exists.

Given the probability generating function we can compute the mass function for $k \in \mathbb{N}_0$ by taking the k-th derivative, evaluating at $t = 0$ and dividing by $k!$, or equivalently

$$\mathbb{P}(X = k) = \frac{1}{k!} \frac{d^k}{dt^k} P_X(t)\Big|_{t=0}.$$

1.4 Moment Generating Functions

n this case the *moments* of a random variable. Moments of a random variable describe the *shape* and *location* of the variable's probability distribution.

The k-th moment of a random variable X is given by $\mathbb{E}[X^k]$. Often we are also interested in the k-th central moments of a random variable X which are given by $\mathbb{E}[(X - \mathbb{E}[X])^k]$.

> **Definition (Moment Generating Function):** Given a random variable X we define the moment generating function
>
> $$\phi_X(t) := \mathbb{E}[e^{tX}]; \quad t \in \mathbb{R},$$
>
> when it is finite.

Given the moment generating function of a random variable we can compute the k-th moment by taking the k-th derivative with respect to t and evaluating at $t = 0$, that is

$$\mathbb{E}[X^k] = \phi_X^{(t)}(t)\Big|_{t=0}.$$

1.5 Integral Inequalities

In this section we provide proofs of some of the fundamental results used in probability theory. We begin by proving the which gives an upper bound on probabilities of non-negative, finite mean random variables.

> **Theorem (Markov Inequality):** For a non-negative, finite mean random variable X, for any $\epsilon > 0$ we have that
>
> $$\mathbb{P}(X \geq \epsilon) \leq \frac{1}{\epsilon}\mathbb{E}[X].$$

PROOF: We split the integration region $[0, \infty)$ to find that

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^\infty x \; dF_X(x) \\
&= \int_0^\epsilon x \; dF_X(x) + \int_\epsilon^\infty x \; dF_X(x) \\
&\geq \int_\epsilon^\infty x \; dF_X(x) \\
&\geq \int_\epsilon^\infty \epsilon \; dF_X(x) \\
&= \epsilon \int_\epsilon^\infty \; dF_X(x) \\
&= \epsilon \cdot \mathbb{P}(X \geq \epsilon).
\end{aligned}$$

□

> **Theorem (Chebyshev's Inequality):** For random variable X with mean μ and variance σ^2 then for any $\epsilon > 0$
>
> $$\mathbb{P}(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

PROOF: We apply the Markov inequality for the nonnegative random variable $|X-\mu|^2$

$$\begin{aligned}\mathbb{P}(|X-\mu| \geq \epsilon &= \mathbb{P}(|X-\mu|^2 \geq \epsilon^2 \\ &\leq \frac{\mathbb{E}[|X-\mu|^2]}{\epsilon^2} \\ &= \frac{\sigma^2}{\epsilon^2}.\end{aligned}$$

□

2 Markov Chains

2.1 Fundamentals of Markov Chains