Probability and Statistics

Events, Random Variables and Distributions

XYZW Project

Contents

Preface **1**
Structure of the book 1
Notation 1

1 Events **2**
1.1 Sample spaces 2
1.2 Operations on events 2
1.3 Axioms of probability 2

2 Random Variables **4**
2.1 Definition 4
2.2 Expectation 4
2.2.1 Linearity of expectation 4
2.3 Variance 4
2.4 Simulation 5

Preface

Probability gives uncertainty a model you can compute with; statistics then uses that model to draw conclusions from data. This book covers the foundations of the former: events, probability and random variables.

Structure of the book

- Events: sample spaces, events and the axioms of probability.
- Random Variables: random variables, expectation and variance.

  The calculus you need

  The expectation of a continuous random variable is an integral. Limits and Derivatives from *Calculus* are enough.

Notation

$\mathbb{P}(A)$ is the probability of an event A; $\mathbb{E}[X]$ and $\mathrm{Var}(X)$ are the expectation and variance of a random variable X.

Chapter 1

Events

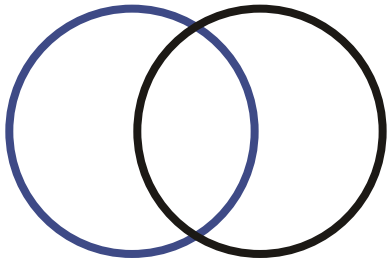
1.1 Sample spaces

The set of all possible outcomes of a random experiment is its **sample space**, written Ω. For one roll of a die, $\Omega = \{1, 2, 3, 4, 5, 6\}$.

An **event** is a subset of Ω. "Roll an even number" is the event $A = \{2, 4, 6\}$.

1.2 Operations on events

Events combine like sets. In the figure, the overlap of the two circles is $A \cap B$.



Notation	Read as	Meaning
$A \cup B$	A or B	at least one happens
$A \cap B$	A and B	both happen
A^c	not A	A does not happen
$A \cap B = \emptyset$	disjoint	they cannot both happen

1.3 Axioms of probability

Probability

A function $\mathbb{P}$ assigns every event a number in $[0, 1]$ such that

1. $\mathbb{P}(\Omega) = 1$;

2. for pairwise disjoint events $A_1, A_2, ...$, $\mathbb{P}(\bigcup_i A_i) = \sum_i \mathbb{P}(A_i)$.

Inclusion–exclusion

For any events A and B,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

The subtracted term is exactly the overlap in the figure: $\mathbb{P}(A)$ and $\mathbb{P}(B)$ each counted it once.

Chapter 2

Random Variables

2.1 Definition

A **random variable** is a function $X : \Omega \to \mathbb{R}$ from the sample space to the reals. It turns "what happened" into "how much we got".

2.2 Expectation

The **expectation** of a discrete random variable X is its probability-weighted average:

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

For a continuous random variable the sum becomes an integral, $\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\, \mathrm{d}x$, where f is the density.

2.2.1 Linearity of expectation

Linearity

For any random variables X, Y and constants a, b,

$$\mathbb{E}[aX + bY] = a\, \mathbb{E}[X] + b\, \mathbb{E}[Y].$$

No independence between X and Y is required. In the language of *Linear Algebra*, expectation preserves every linear combination.

Aside

"aside" is not one of Obsidian's built-in callout types. The publisher still renders it, with a neutral colour.

2.3 Variance

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

Variance measures how far X strays from its expectation. Its square root is the **standard deviation**.

2.4 Simulation

```python
import random

rolls = [random.randint(1, 6) for _ in range(100_000)]
print(sum(rolls) / len(rolls))  # close to 3.5
```

The expected value of a die roll is 3.5, and the simulated mean lands close to it. Event notation is in Events › Operations on events.