Pre-session tasks: Intro. to Markov Chains

University of Oxford

Balliol College Fibonacci Programme 2025

Table of Contents

1. Recap of Probability (Reading + Exercises)

2. Intro. to Stochastic Processes (Reading)

Table of Contents

1. Recap of Probability (Reading + Exercises)

2. Intro. to Stochastic Processes (Reading)

Conditional probability

Let $A, B \in \mathcal{F}$ be events with $P(B) > 0$. Supposing we know that B occurred, how likely is A given that information? That is, what is the conditional probability $P(A|B)$?

We define

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

We need a more sophisticated definition if we want to define the conditional probability $P(A|B)$ when $P(B) = 0$.

Example:

Tickets are drawn consecutively and *without replacement* from a box of tickets numbered $1 - 10$. What is the chance the second ticket is even numbered given the first is

- even?
- labelled 3?

Bayes' formula

Let $B_1, B_2, \cdots, B_n$ be mutually exclusive events with $A \subset \bigcup_{j=1}^{n} B_j$, then

$$P(A) = \sum_{j=1}^{n} P(A|B_j)P(B_j).$$

With the same assumptions as for the Law of Total Probability,

$$P(B_j|A) = \frac{P(B_j \cap A)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{k=1}^{n} P(A|B_k)P(B_k)}.$$

Example:

A disease affects 1/1000 newborns and shortly after birth a baby is screened for this disease using a cheap test that has a 2% false positive rate (the test has no false negatives). If the baby tests positive, what is the chance it has the disease?

Independent events

Events A and B are said to be independent if

$$P(A \cap B) = P(A)P(B).$$

If $\mathbb{P}(B) \neq 0$ or $\mathbb{P}(A) \neq 0$ then this is the same as $P(A|B) = P(A)$ and $P(B|A) = P(B)$.

Events $A_1, \cdots, A_n$ are independent if, for any subset $\{i_1, ..., i_k\}$ of $\{1, ..., n\}$,

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \times \cdots \times P(A_{i_k}).$$

Random variables

A random variable (rv) on a probability space $(\Omega, \mathcal{F}, P)$ is a function $X : \Omega \to \mathbb{R}$.

Usually, we want to talk about the probabilities that the values of random variables lie in sets of the form $(a, b) = \{x : a < x < b\}$. The smallest sigma-algebra of subsets of $\mathbb{R}$ that contains these sets is called the set $\mathcal{B}(\mathbb{R})$ of Borel sets.

The probability that $X \in (a, b)$ is the probability of the subset $\{\omega : X(\omega) \in (a, b)\}$. In order for this to make sense, we need this set to be in $\mathcal{F}$ for all $a < b$. When this holds, we say the function X is *measurable* with respect to $\mathcal{F}$.

So X is measurable with respect to $\mathcal{F}$ if $\{\omega : X(\omega) \in B\} \in \mathcal{F}$ for all Borel sets $B \subset \mathbb{R}$.

Distribution Functions

The function $F_X(t) = P(X \leq t) = P(\{\omega : X(\omega) \in (-\infty, t]\}$ that maps R to $[0, 1]$ is called the distribution function of the random variable X.

Any distribution function F

F1. is non-decreasing,

F2. is such that $F(x) \to 0$ as $x \to -\infty$ and $F(x) \to 1$ as $x \to \infty$, and

F3. is 'right-continuous', that is $\lim_{h \to 0^+} F(t + h) = F(t)$ for all t.

Distribution Functions

We say that

- the random variable X is discrete if it can take only countably-many values, with $P(X = x_i) = p_i > 0$ and $\sum_i p_i = 1$. Its distribution function $F_X(t)$ is commonly a step function.
- the random variable X is continuous if $F_X(t)$ is absolutely continuous, that is if there exists a function $f_X(t)$ that maps R to R_+ such that $F_X(t) = \int_{-\infty}^{t} f_X(u)du$.

A mixed random variable has some points that have positive probability and also some continuous parts.

Examples of distributions

- Examples of discrete random variables: binomial, Poisson, geometric, negative binomial, discrete uniform
  `http://en.wikipedia.org/wiki/Category:Discrete_distributions`
- Examples of continuous random variables: normal, exponential, gamma, beta, uniform on an interval (a, b)
  `http://en.wikipedia.org/wiki/Category:Continuous_distributions`

Expectation of X

For a discrete, continuous or mixed random variable X that takes on values in the set S_X, the expectation of X is

$$E(X) = \int_{S_X} x dF_X(x)$$

The integral on the right hand side is a Lebesgue-Stieltjes integral.

It can be evaluated as

$$E(X) = \begin{cases} \sum_{i=1}^{\infty} x_i P(X = x_i), & \text{if } X \text{ is discrete,} \\ \int_{-\infty}^{\infty} x f_X(x) dx, & \text{if } X \text{ is absolutely continuous,} \end{cases}$$

assuming these integrals/sums are absolutely convergent.

Expectation of $g(X)$

For a measurable function g that maps S_X to some other set S_Y, $Y = g(X)$ is a random variable taking values in S_Y and

$$E(Y) = E(g(X)) = \int_{S_X} g(x) dF_X(x).$$

We can also evaluate $E(Y)$ by calculating its distribution function $F_Y(y)$ and then using the expression

$$E(Y) = \int_{S_Y} y dF_Y(y).$$

Properties of Expectation

- $E(aX + bY) = aE(X) + bE(Y)$.
- If $X \leq Y$, then $E(X) \leq E(Y)$.
- If $X \equiv c$, then $E(X) = c$.

Exercise

Let $\Omega = \{a, b, c, d\}$, $P(\{a\}) = \frac{1}{2}$, $P(\{b\}) = P(\{c\}) = \frac{1}{8}$ and $P(\{d\}) = \frac{1}{4}$.
Define random variables,

$$Y(\omega) = \begin{cases} 1, & \omega = a \text{ or } b, \\ 0, & \omega = c \text{ or } d, \end{cases}$$

$$X(\omega) = \begin{cases} 2, & \omega = a \text{ or } c, \\ 5, & \omega = b \text{ or } d. \end{cases}$$

Compute $E(X)$, $E(X|Y)$ and $E(E(X|Y))$.

Conditional Probability

The conditional probability of event A given X is a random variable $P(A|X) = \eta(X)$, where

- if $P(X = x) > 0$, then

$$\eta(x) = P(A|X = x) = \frac{P(A \cap \{X = x\}}{P(\{X = x\}},$$

- if $P(X = x) = 0$, then

$$\eta(x) = P(A|X = x) = \lim_{\epsilon \to 0^+} \frac{P(A \cap \{X \in (x - \epsilon, x + \epsilon)\})}{P(\{X \in (x - \epsilon, x + \epsilon)\})}.$$

Conditional Distribution

- The conditional distribution function $F_{Y|X}(y|X)$ of Y evaluated at the real number y is given by $P(\{Y \leq y\}|X)$ as defined on the previous slide.
- If (X, Y) is absolutely continuous, then the conditional density of Y given that $X = x$ is

$$f_{Y|X}(y|x) = \frac{f_{(X,Y)}(x, y)}{f_X(x)}.$$

Limit Theorems

The Law of Large Numbers (LLN) states that if X_1, X_2, $\cdots$ are independent and identically-distributed with mean μ, then

$$\overline{X_n} \equiv \frac{1}{n}\sum_{j=1}^{n} X_j \to \mu$$

as $n \to \infty$, where we interpret the right hand side as a random variable that is equal to μ with probability one.
In the strong form, this is true almost surely, which means that it is true on a set A of sequences $x_1, x_2, \ldots$ that has probability one.
In the weak form, this is true in probability which means that, for all $\epsilon > 0$,

$$P(|\overline{X_n} - \mu| > \epsilon) \to 0$$

as $n \to \infty$.

Limit Theorems

The Central Limit Theorem (CLT) states that if X_1, X_2, $\cdots$ are independent and identically-distributed with mean μ and variance σ^2, then for any x,

$$P\left(\frac{\overline{X_n}-\mu}{\sigma/\sqrt{n}} < x\right) \to \Phi(x) \equiv \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

as $n \to \infty$. That is, a suitably-scaled variation from the mean approaches a standard normal distribution as $n \to \infty$.

Example

A biased coin that comes up heads with probability 1/3 is tossed 1000 times.

- What is the expected number of heads in the sequence?
- What is the probability that the number of heads that appears in the sequence is less than 300?

Table of Contents

1. Recap of Probability (Reading + Exercises)

2. Intro. to Stochastic Processes (Reading)

Bibliography


Ghahramani, Saeed (2018). *Fundamentals of probability: with stochastic processes*. Chapman and Hall/CRC.


Taylor, Peter (2023). *MAST30001 Stochastic Modelling*. Lecture slides.