



桑慧敏 著

醍醐「統計學」



2

機率基本概念

本章大綱

- 2.1 前言
 - 2.1.1 醍醐灌頂
 - 2.1.2 本章架構
 - 2.2 樣本空間, 事件空間, 與機率函數
 - 2.3 機率三公設及基本機率定理
 - 2.3.1 機率三公設
 - 2.3.2 八個基本機率定理
 - 2.4 條件機率及條件機率定理
 - 2.4.1 條件機率
 - 2.4.2 基本條件機率定理
 - 2.4.3 乘法定理與總機率定理
 - 2.4.4 貝氏定理
 - 2.4.5 獨立事件
 - 2.5 重點整理
練習題
-

2.1 前言

2.1.1 醍醐灌頂

進入本章前, 請先想想以下問題:

1. 文獻中第一個機率問題是賭金分配, 是個世紀難題, 難倒了許多數學家, 包括巴斯卡與費馬。今日為何能輕易解此題?
2. 在微觀物質中, 粒子的位置與動量是服從「機率模式」嗎? 兩位物理學家愛因斯坦與波爾看法迥異。愛因斯坦錯了嗎?

愛因斯坦 (1926) 說:
「上帝是不擲骰子的」。
請看本書第一章說明。

2.1.2 本章架構

賭金分配問題見本章
習題第3題。

「三八加一線」是指三
公設, 八定理, 及條件
機率。

文獻中記載的第一個賭金分配的機率問題 (Pacioli, 1494) 經過許多數學家 (包括巴斯卡與費馬) 的討論, 150 多年後才得解。不得解的原因是「機率」沒有定義, 而且還未制定「機率三公設」。

本章首先說明使用機率模式的動機, 然後定義「機率」與制定「機率三公設」。接著就能證明出有用的八個機率定理。最後貝氏 (Bayes) 先生定義了條件機率 (以「一直線」符號表示), 補足了機率三公設與八定理無法分析的條件事件。有了機率定義, 在加上「三八加一線」, 整個機率論已完整; 能夠計算出任何事件的機率。

作者將本章內容濃縮成四句歌詞, 猜猜此四句歌詞的涵意。

此四句歌詞, 可以三輪
車調吟唱。熟讀本章
後, 再回想此歌詞涵意,
看看此歌詞是否掌握
該章的精神。

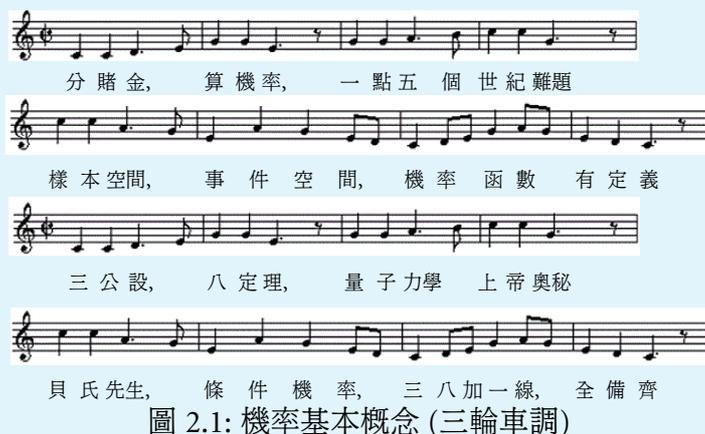


圖 2.1: 機率基本概念 (三輪車調)

3

一元隨機變數基本概念

本章大綱

- 3.1 前言
 - 3.1.1 醍醐灌頂
 - 3.1.2 本章架構
 - 3.2 隨機變數
 - 3.2.1 以隨機變數表達事件
 - 3.2.2 隨機變數分類
 - 3.3 隨機變數的證件
 - 3.3.1 累積分配函數
 - 3.3.2 分配函數
 - 3.3.3 動差母函數
 - 3.4 分配函數的屬性
 - 3.4.1 描述集中趨勢之屬性
 - 3.4.2 描述分散趨勢之屬性
 - 3.4.3 尾端事件機率
 - 3.4.4 四分位數與百分位數
 - 3.4.5 動差
 - 3.5 重點整理
 - 練習題
-

3.1 前言

3.1.1 醍醐灌頂

投擲一個錢幣 3 次,「出現一次正面的事件」可寫成 $\{(正, 反, 反), (反, 正, 反), (反, 反, 正)\}$, 這種寫法很繁瑣。如果是投擲一個錢幣 100 次, 幾乎是無法一一列出「出現一次正面的事件」中所有樣本點。有何簡便的方法表達此事件?

3.1.2 本章架構

請不要被「隨機變數」的名字誤導, 名字中雖有變數, 但它的性質卻不是變數; 它是函數。

就像身分證或護照是人的 ID, ID 能讓我們辨認出某人的名字。「隨機變數」的 ID, 能讓我們辨認出「隨機變數」的名字, 進而知道所有與其相關的屬性。

回答醍醐灌頂的課題, 很自然的一個想法是建立一個函數, 將隨機實驗中樣本點轉換成實數; 此函數稱為「隨機變數」。隨機變數的好處: (1) 量化樣本點, (2) 簡化「機率事件的表示方式」, (3) 其後續發展出的“累積分配函數 (cdf) 與分配函數 (pdf)”使得我們能與「微積分」這個有力的數學工具連上線, 藉此簡化分析, (4) 更進一步產生了所謂的機率模式, 如常態機率模式。

本章有三大課題: (i) 定義隨機變數, (ii) 定義能分辨隨機變數的證件 (ID), 包括累積分配函數, 分配函數, 與動差母函數; (iii) 定義隨機變數的屬性, 包括描述集中趨勢, 分散趨勢, 尾端事件之機率, 四分位數, 百分位數, 與動差等屬性。

作者將本章內容濃縮成四句歌詞, 猜猜此四句歌詞的涵意。

此四句歌詞, 可以三輪車調吟唱。熟讀本章後, 再回想此歌詞涵意, 看看此歌詞是否掌握該章的精神。

樣本點, 實數值, 隨機變數結連理

c d f, p d f, 身份證件辨唯一

期望值, 變異數, 又名一二階動差

馬可夫, 柴比雪夫, 神龍現尾, 真稀奇

4

常用的機率模式

本章大綱

- 4.1 前言
 - 4.1.1 醍醐灌頂
 - 4.1.2 本章架構
 - 4.2 常用的離散形機率模式
 - 4.2.1 均勻 (Uniform)
 - 4.2.2 伯努力 (Bernoulli)
 - 4.2.3 二項 (Binomial)
 - 4.2.4 幾何 (Geometric)
 - 4.2.5 負二項 (Negative Binomial)
 - 4.2.6 超幾何 (Hypergeometric)
 - 4.2.7 卜瓦松 (Poisson)
 - 4.3 常用的連續型機率模式
 - 4.3.1 常態與標準常態 (Normal and Standard Normal)
 - 4.3.2 矩形 (Rectangular)
 - 4.3.3 指數 (Exponential)
 - 4.3.4 伽瑪 (Gamma) 與 Erlang
 - 4.3.5 韋伯 (Weibull) 與 Rayleigh
 - 4.3.6 對數常態 (Lognormal)
 - 4.3.7 貝他 (Beta)
 - 4.3.8 卡方 (Chi-Squared)
 - 4.3.9 T
 - 4.3.10 F
 - 4.4 重點整理
練習題
-

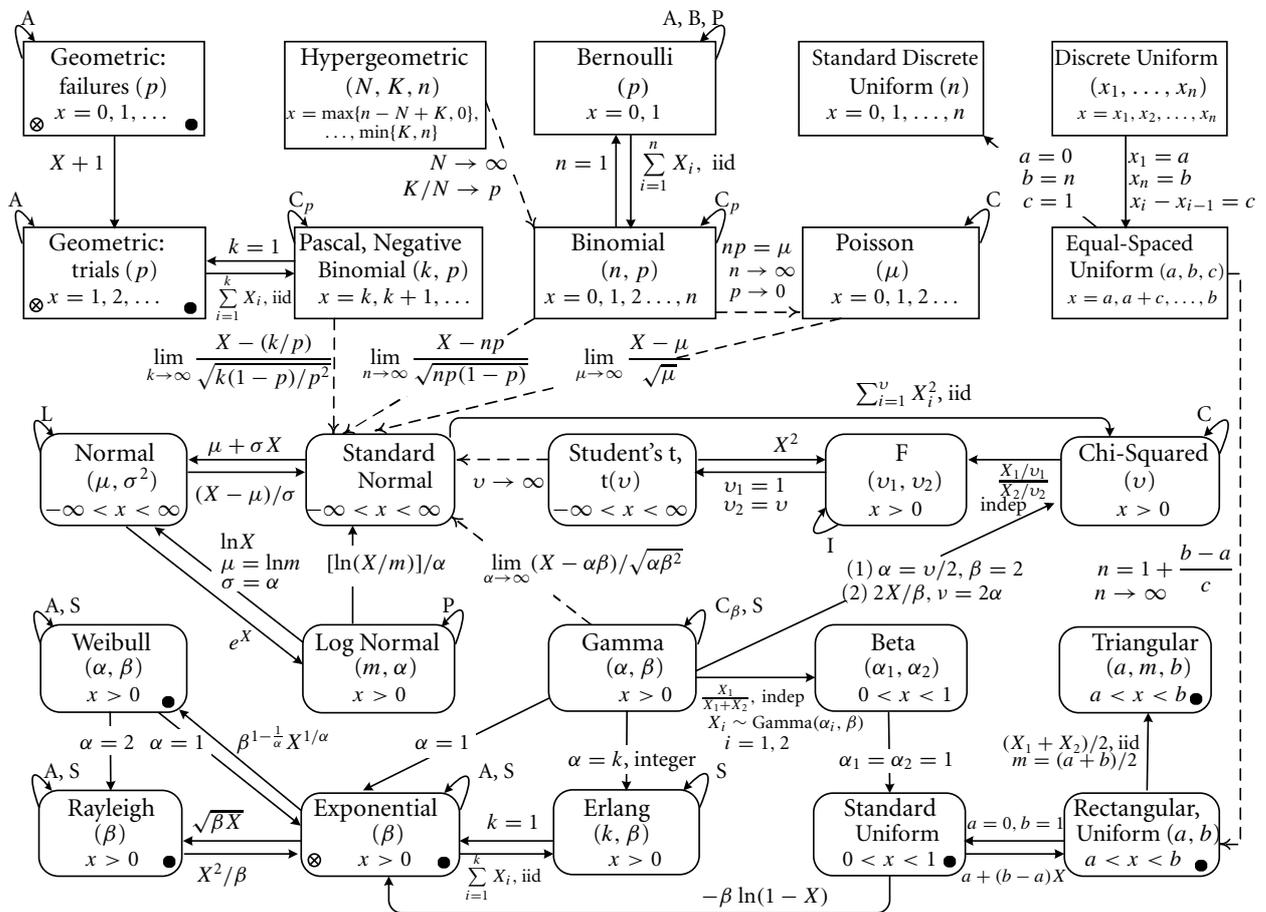
4.1 前言

4.1.1 醍醐灌頂

你同意這世界是「亂中有序」的嗎? 如果同意, 如何證明。

4.1.2 本章架構

常用的機率模式彙整於圖 4.1。先認識這些機率模式 (名字, 參數, 圖形, 相互關係), 再學習利用這些模式來解釋「亂中有序」的世界。



- 註 1. 圖中以實線 (→) 表示兩隨機變數間的轉換, 其中 X 為箭尾的隨機變數; (---) 虛線連接表示極限的關係。
- 註 2. 圖中自我關聯箭號 (↻): A: Minimum, B: Maximum, C: Convolution, I: Inverse, L: Linear Combination, P: Product, S: Scaling. (Source random variables: A and B require iid; C, L, and P require independent; C_β, C_p require “variational iid”).
- 註 3. \bullet 表示 cdf 為 “隱函數 (closed-form inverse cdf)”。 \otimes 表示 “無記憶性 (memoryless property)”。

圖 4.1: 25 個常用機率模式的關聯圖