

機率與統計簡介

機率與統計的關係(Relation between probability and statistic)

sample由某個機率分佈產生，而model由sample找出來

- Sample -----> Model: 統計(Statistic)
- Sample <----- Model: 機率(Probability)

基礎定理(Probability rules)

1. 互斥事件的加法(Addition rule for mutually exclusive events) :
$$P(A \text{ or } B) = P(A) + P(B)$$
2. 獨立事件的乘法(Multiplication rule for independent events) :
$$P(A \text{ and } B) = P(A) \cdot P(B)$$
3. 通用加法(General addition rule) :
$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$
4. 通用乘法(General multiplication rule) :
$$P(A \text{ and } B) = P(A)P(B|A) \text{ or } P(B)P(A|B)$$

貝氏理論(Bayes theory)

- Let $A_{i=1\sim n}$ be mutually exclusive and exhaustive event.
- Then for any other event B(Law of total probability):

$$P(A_k|B) = \frac{P(A_k \text{ and } B)}{P(B)} = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^n P(B|A_i)P(A_i)}$$

期望值與變異數計算(Expected value and Variance)

Item	Formula
期望值(Expected value)	$E(X) = \sum_{i=1}^k x_i p(x_i)$
變異數(Variance)	$V(X) = E(X^2) - E(X)^2 = \sum [x_i^2 p(x_i)] - \mu^2$
共變異數(Covariance)	$Cov(X, Y) = E(XY) - E(X)E(Y)$
相關係數(Correlation coefficient)	$Corr(X, Y) = \frac{Cov(X, Y)}{S.D(X) \cdot S.D(Y)}$

常見的機率分佈(Probability distributions)

機率分佈大分類可分為

1. 離散型(Discrete)
2. 連續型(Continuous)

離散型(Discrete)

1. Bernoulli

結果只有兩種(0 or 1，正面或反面)可能的實驗，例如投擲硬幣。

2. Geometric

在伯努利實驗中，第一次成功所需要的實驗次數。

3. Binomial

進行伯努利實驗n次，成功x次的機率分佈。

4. Poisson

當Binomial的參數n很大，但p很小時的機率分佈。

連續型(Continuous)

1. Uniform

在指定的範圍(a,b)間，每個變數x的機率密度都是相同的機率分佈模型。

2. Normal

以平均數和標準差為參數的對稱機率分佈模型。

3. Exponential

以 λ 為參數的連續機率分佈，通常發生在描述獨立事件變數發生的時間間隔。

4. Gamma

假設 X_1, X_2, \dots, X_n 為連續發生事件的等候時間，且這n次等候時間為獨立的，那麼這n次等候時間之和Y ($Y=X_1+X_2+\dots+X_n$) 服從伽瑪分布，即 $Y \sim \text{Gamma}(\alpha, \beta)$ ，亦可記作 $Y \sim \text{Gamma}(\alpha, \lambda)$ ，其中 $\alpha = n$ ，而 β 與 λ 互為倒數關係， λ 表單位時間內事件的發生率。指數分布為 $\alpha = 1$ 的伽瑪分布。

5. Chi-Square(將 (α, β) 設成 $(\frac{k}{2}, 2)$)

若k個隨機變數 Z_1, Z_2, \dots, Z_k 是相互獨立，符合標準常態分布的隨機變數（數學期望為0、變異數為1），則隨機變數Z的平方和 $X = \sum_{i=1}^k Z_i^2$ 被稱為服從自由度為 k 的卡方分布，記作 $X \sim \chi^2(k)$ 。

機率分佈圖

離散型

圖片來源:[維基百科](#)

- Geometric distribution
- Binomial distribution
- Poisson distribution

連續型

圖片來源:[維基百科](#)

- Uniform distribution
- Normal distribution
- Exponential distribution
- Gamma distribution
- chi-square distribution

