Statistics with Recitation

Editor: Hsin Lin [*] (林鑫君)

November 28, 2021

[*]The Teaching Assistant at Department of Economics, National Taiwan University.

8. 抽樣分配與漸近理論

Property 1 前情提要

本章節我們將首先介紹幾個前一章節未提到的某些分配，這些分配將時常運用在往後統計推論等章節，如：區間估計、假設檢定、迴歸分析等等。再來我們將介紹統計學裡的大樣本性質，其中包含弱大數法則、中央極限定理以及各種收斂性質。以下有幾點須特別注意：

1. 本章介紹的分配中，pdf 多不重要，重點在分配形式以及與其他分配的關聯
2. 本章將延續應用先前某些連續型分配，比如說常態、Gamma 等分配
3. 本章也會大量運用許多單雙變量隨機變數章節的基本性質

≪ *notes* ≫

Property 2 卡方分配 (Chi-square distribution)

- 若 $X \sim \chi^2(k)$，則

$$f(x) = \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}} \,, \; x > 0$$

$$E(X) = k$$

$$Var(X) = 2k$$

1. 卡方是 Gamma 分配的特例，其中 $\alpha = \frac{k}{2}$，$\beta = 2$，k 稱為卡方分配的自由度
2. 卡方分配也具有可加性 (參數為自由度之和)、可乘性
3. 若 $Z \sim N(0,1)$，則 $Z^2 \sim \chi^2(1)$
4. 若 $\{Z_i\}_{i=1}^{n} \overset{i.i.d.}{\sim} N(0,1)$，則 $\sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$

$\ll notes \gg$

Property 3 卡方分配重要性質

- 若 $X \overset{i.i.d.}{\sim} N(\mu,\sigma^2)$，則

$$\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(1)$$

$$\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2(1)$$

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$$

1. 這三個傢伙以後很常用，務必熟練怎麼來的

$\ll notes \gg$

Property 4　Cochran 定理 (Cochran's theorem)

- 給定 $Z \overset{i.i.d.}{\sim} N(0,1)$，並令

$$C = \sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$$

若 C 可以拆成多個二次型 (quadratic forms) 的和，意即

$$C = \sum_{i=1}^{n} Z_i^2 = \underbrace{\sum_{i=1}^{n} W_{1i}^2}_{\text{C1}} + \underbrace{\sum_{i=1}^{n} W_{2i}^2}_{\text{C2}} + \ldots + \underbrace{\sum_{i=1}^{n} W_{ki}^2}_{\text{Ck}}$$

則 C_i 為卡方分配，其自由度之和為 n，且 C_i 之間彼此獨立 (這點超重要)

1. 這個定理真的太無情了，可以拿來推卡方分配，也可以拿來證某些變數之間彼此獨立，在往後章節也會不時看到
2. 講人話一點，若已知 $\sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$，$\sum_{i=1}^{n} B_i^2 \sim \chi^2(1)$，且

$$\underbrace{\sum_{i=1}^{n} Z_i^2} = \sum_{i=1}^{n} A_i^2 + \underbrace{\sum_{i=1}^{n} B_{2i}^2}$$

   則根據 Cochran 定理，$\sum_{i=1}^{n} A_i^2 \sim \chi^2(n-1)$，且 $\sum_{i=1}^{n} A_i^2 \perp\!\!\!\perp \sum_{i=1}^{n} B_i^2$

$\ll notes \gg$

Property 5 一個超重要的卡方分配性質

- 若 $\{X_i\}_{i=1}^{n} \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$，則

$$\frac{(n-1)s^2}{\sigma^2} = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$$

1. 此推導非常重要，除了 Cochran 定理，也可以順便複習分配性質

$\ll notes \gg$

Property 6 t 分配 (t distribution)

- 給定 $Z \sim N(0, 1)$，$W \sim \chi^2(k)$，且 $Z \perp\!\!\!\perp W$，則

$$X = \frac{Z}{\sqrt{W/k}} \sim t(k)$$

1. t 分配是從常態分配推來的，要用他之前請先確定母體分配為常態
2. 當 t 分配的自由度越大，會越接近標準常態分配
3. k 為 t 分配的自由度，當 $k = 1$，$t(1)$ 為標準柯西分配

$\ll notes \gg$

Property 7 一個超重要的 t 分配性質

- 若 $\{X_i\}_{i=1}^{n} \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$，則

$$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t(n-1)$$

1. 此推導非常重要，需用到 Cochran 定理
2. 非常重要的一點，資料須來自常態母體，才可以得到此分配性質
3. 若資料來自常態母體，且 σ 已知，則 $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
4. 若資料來自常態母體，且 σ 未知，則可用樣本標準差取代，$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t(n-1)$
5. 3、4 兩點非常重要，往後在統計推論章節會時常用到

$\ll notes \gg$

Property 8 F 分配 (F distribution)

- 給定 $Y_1 \sim \chi^2(k_1)$，$Y_2 \sim \chi^2(k_2)$，且 $Y_1 \perp\!\!\!\perp Y_2$，則

$$X = \frac{Y_1/k_1}{Y_2/k_2} = \frac{\chi^2\,(k_1)\,/k_1}{\chi^2\,(k_2)\,/k_2} \sim F(k_1, k_2)$$

1. F 分配的 pdf、期望值、變異數都不重要，F 分配主要用於統計推論
2. 若 $T \sim t(k)$，則 $T^2 \overset{d}{=} F(1, k)$

$\ll notes \gg$

Property 9 F 分配的機率性值 (查表會用到)

- 若 $X \sim F(k_1, k_2)$，$Y \sim F(k_2, k_1)$，若 $\alpha = P(X \geq c)$，則

$$1 - \alpha = P(X \leq c) = P\left(Y > \frac{1}{c}\right)$$

$$F_{1-\alpha}\,(k_1, k_2) = \frac{1}{F_{\alpha}\,(k_2, k_1)}$$

$$F_{\alpha}\,(k_1, k_2) = \frac{1}{F_{1-\alpha}\,(k_2, k_1)}$$

1. F 分配表大多只提供單尾機率，所以會常搭配以上性質轉換

$\ll notes \gg$

Property 10　實數序列的極限

- 給定一實數序列 $\{Ci\}_{i=1}^{n}$，對任何 $\epsilon > 0$ 來說，存在實數 C 使得

$$|C_n - C| < \epsilon$$

此時我們稱 C 為實數序列 $\{Ci\}_{i=1}^{n}$ 的極限，亦可表示為:

$$\lim_{n\to\infty} C_n = C$$

$\ll notes \gg$

Property 11 Markov 不等式

- 給定 X 為非負隨機變數，對於任意實數 $m > 0$ 來說：

$$P(X \geq m) \leq \frac{E(X)}{m}$$

1. 推導建議要會，可由不自覺統計學家法則出發

$$E(X) = \int_0^\infty xf(x)dx = \int_0^m xf(x)dx + \int_m^\infty xf(x)dx$$

$$\geq \int_m^\infty xf(x)dx \geq \int_m^\infty mf(x)dx = m\int_m^\infty f(x)dx = mP(X \geq m)$$

$\ll notes \gg$

Property 12 Chebshev 不等式
- 給定一隨機變數 Y，對於任意 $\epsilon > 0$ 而言:

$$P\left(|Y-\mu_Y| \geq \epsilon\right) \leq \frac{\sigma_Y^2}{\epsilon^2}$$

$$P\left(|Y-\mu_Y| \geq k\sigma_Y\right) \leq \frac{1}{k^2}$$

1. Chebshev 不等式可由 Markov 不等式推得，所以你可以挑一個記

給定 $X = (Y-\mu_Y)^2$，則 $E(X) = E[(Y-\mu_y)^2] = \sigma_Y^2$，根據 Markov 不等式:

$$P\left(|Y-\mu_Y| \geq \epsilon\right) = P\left(\left(Y-\mu_Y\right)^2 \geq \epsilon^2\right) = P\left(X \geq \epsilon^2\right) \leq \frac{\sigma_Y^2}{\epsilon^2}$$

令 $\epsilon = k\sigma_Y$，則可得:

$$P\left(|Y-\mu_Y| \geq k\sigma_Y\right) \leq \frac{\sigma_Y^2}{\left(k\sigma_Y\right)^2} = \frac{1}{k^2}$$

$\ll notes \gg$

Property 13 機率收斂 (converge in probability)

- 給定一隨機變數序列 $\{Y_n\}$，對於任何 $\epsilon > 0$，若

$$\lim_{n\to\infty} P\left(|Y_n - c| > \varepsilon\right) = 0$$

則稱 Y_n 機率收斂到實數 c，以 $Y_n \xrightarrow{p} c$ 或 $plim_{n\to\infty} Y_n = c$ 表示

1. 講白話一點，機率收斂代表隨著 $n \to \infty$，隨機變數的實現值會退化成一個常數 c，意即隨機變數只有一個實現值 c
2. 我們稱 c 為 Y_n 的機率極限 (probability limit)

$\ll notes \gg$

Property 14 分配收斂 (converge in distribution)

- 給定一隨機變數序列 $\{Y_n\}$，其相對應的 CDF 為 $F_n(y)$，而另一隨機變數 Y 的 CDF 為 $F_Y(y)$，若

$$\lim_{n\to\infty} F_n(y) = F_Y(y)$$

則稱 Y_n 分配收斂到隨機變數 Y，以 $Y_n \xrightarrow{d} Y$ 表示

1. 講白話一點，隨著 $n \to \infty$，Y_n 和 Y 有著相同分配
2. 我們稱 Y 是 Y_n 的極限分配 (limiting distribution)
3. 既然是說明隨著 $n \to \infty$，Y_n 和 Y 有著相同分配，我們亦可利用 MGF 證明，意即 $\lim_{n\to\infty} M_{Y_n}(t) = \lim_{n\to\infty} M_Y(t)$, 則$Y_n \xrightarrow{d} Y$

$\ll notes \gg$

Property 15 機率收斂隱含分配收斂

1. 如題，若 $Y_n \xrightarrow{p} Y$，則 $Y_n \xrightarrow{d} Y$，反過來不一定成立
2. 給定 Y 唯一常數 c，則若 $Y_n \xrightarrow{p} c$，則 $Y_n \xrightarrow{d} c$
3. 分配收斂是較弱的收斂，以上一點為例，若 Y_n 機率收斂到常數，則隱含 Y_n 的極限分配是一個常數 c

$\ll notes \gg$

Property 16 均方收斂 (converge in mean square)

- 給定一隨機變數序列 $\{Y_n\}$，若 $\{Y_n\}$ 滿足

$$\lim_{n\to\infty} E\left[(Y_n - c)^2\right] = 0$$

則稱 Y_n 均方收斂到實數 c，可用 $Y_n \xrightarrow{m.s} c$ 表示

1. $Y_n \xrightarrow{m.s} c$ 隱含 $Y_n \xrightarrow{p} c$ 隱含 $Y_n \xrightarrow{d} c$，反之則不一定
2. 均方收斂是三種收斂中最強的收斂，其個別收斂速度由慢到快如下:
   (a) 分配收斂為隨機變數 Y_n 之實現值收斂至一個特定區間 (Y 的值域)
   (b) 機率收斂為隨機變數 Y_n 之實現值收斂至一常數
   (c) 均方收斂也為 Y_n 之實現值收斂至一常數，但收斂速度比機率收斂快

≪ *notes* ≫

Property 17 均方收斂的充分且必要條件

$$\lim_{n\to\infty} E\left(Y_n\right) = c$$

$$\lim_{n\to\infty} Var\left(Y_n\right) = 0$$

1. 推導可從均方收斂定義出發

$$E\left[\left(Y_n - c\right)^2\right] = E\left[\left(Y_n - E\left(Y_n\right) + E\left(Y_n\right) - c\right)^2\right]$$

$$= E\left[\left(Y_n - E\left(Y_n\right)\right)^2\right] + \left[E\left(Y_n\right) - c\right]^2 + 2\underbrace{E\left[\left(Y_n - E\left(Y_n\right)\right)\left(E\left(Y_n\right) - c\right)\right]}_{=(E(Y_n)-c)E(Y_n - E(Y_n))=0}$$

$$= E\left[\left(Y_n - E\left(Y_n\right)\right)^2\right] + \left[E\left(Y_n\right) - c\right]^2 = \underbrace{\mathrm{Var}\left(Y_n\right)}_{\geq 0} + \underbrace{\left[E\left(Y_n\right) - c\right]^2}_{\geq 0}$$

$\ll$ *notes* $\gg$

Property 18 弱大數法則 (Weak Law of Large Numbers, WLLN)
- 給定隨機樣本 $\{X_i\}_{i=1}^n$，其中 $E(X) < \infty$，令 $\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n}$

$$X_n \xrightarrow{p} E(X)$$

1. 簡單來說，弱大數法則表示隨著樣本數 (n) 增加，樣本平均會離母體平均越來越接近
2. 一般化弱大數法則 (計量常用):

$$\frac{\sum_{i=1}^n X_i^r}{n} \xrightarrow{p} E(X^r)$$

$$\frac{\sum_{i=1}^n (X_i - \mu)^2}{n} \xrightarrow{p} E\left[(X-\mu)^2\right] = \mathrm{Var}(X)$$

$$\frac{\sum_{i=1}^n (X_i - \mu_X)(Y_i - \mu_Y)}{n} \xrightarrow{p} E\left[(X-\mu_X)(Y-\mu_Y)\right] = \mathrm{Cov}(X, Y)$$

≪ *notes* ≫

Property 19 中央極限定理 (Central limit theorem, CLT)

- 給定隨機樣本 $\{X_i\}_{i=1}^{n}$，$E(X)=\mu<\infty$，$Var(X)=\sigma^2<\infty$，令 $\bar{X}_n=\frac{\sum_{i=1}^{n}X_i}{n}$，則

$$Z_n=\frac{\bar{X}_n-E\left(X_n\right)}{\sqrt{\operatorname{Var}\left(\bar{X}_n\right)}}=\sqrt{n}\left(\frac{\bar{X}_n-\mu}{\sigma}\right)\xrightarrow{d}N(0,1)$$

$$\sqrt{n}\left(\bar{X}_n-\mu\right)\xrightarrow{d}N(0,\sigma^2)$$

1. 簡單來說，中央極限定理是說當樣本數夠大，樣本平均所形成的分配會很接近常態分配，意即給定 n 足夠大時，

$$\bar{X}_n\overset{A}{\sim}N\left(\mu,\frac{\sigma^2}{n}\right)$$

2. 基本上只要 $n>30$，我們就可認定他為” 足夠大”，但樣本數還是越多越好

$\ll notes \gg$

Property 20 漸近分配 (Asymptotic) 與極限分配 (Limiting) 的差異

$$\sqrt{n}\left(\bar{X}_n - \mu\right) \xrightarrow{d} N\left(0, \sigma^2\right)$$

$$\bar{X}_n \overset{A}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$$

1. $\sqrt{n}\left(\bar{X}_n - \mu\right)$ 的極限分配為 $N\left(0, \sigma^2\right)$
2. $\bar{X}_n$ 的漸進分配為 $N\left(\mu, \frac{\sigma^2}{n}\right)$
3. 隨著 $n \longrightarrow \infty$，$Var(\bar{X}_n) \longrightarrow 0$，此時 $\bar{X}_n$ 收斂成常數，故 $\bar{X}_n$ 的極限分配為常數
4. 漸進分配可以想成 n 足夠大，但不到無限大，開始接近常態但還沒收斂到常數 μ 的近似分配

$\ll$ *notes* $\gg$

Property 21 連續映射定理 (Continuous Mapping Theorem, CMT)
- 給定 $\bar{X}_n \xrightarrow{p} X$，且 g 為連續函數

$$g\left(X_n\right) \xrightarrow{p} g(X)$$

1. 此定理應用非常廣泛，在迴歸分析常拿來證明一致性及搭配 WLLN 使用
2. 舉例來說，若 $\bar{X}_n \xrightarrow{p} \mu$，則 $\bar{X}_n^2 \xrightarrow{p} \mu^2$

$\ll notes \gg$

Property 22 Slutsky 定理 (Slutsky's theorem)

- 給定 $\bar{X}_n \xrightarrow{d} X$，$\bar{Y}_n \xrightarrow{p} b$

$$X_n + Y_n \xrightarrow{d} X + b$$

$$X_n Y_n \xrightarrow{d} bX$$

$$\frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{b},\ b \neq 0$$

1. 此定理需延用前述有間機率收斂、分配收斂的相對強弱來判定
2. Slutsky 定理主要用來找隨機變數的極限分配與漸進分配，以及推導樞紐量

$\ll notes \gg$

Property 23 The delta method

- 給定

$$\sqrt{n}\left(\bar{X}_n-\mu_X\right)\xrightarrow{d}N\left(0,\sigma_X^2\right)$$

若 g 為連續函數且 $g'(\mu_X)\neq 0$ 存在，則

$$\sqrt{n}\left[g\left(X_n\right)-g\left(\mu_X\right)\right]\xrightarrow{d}N\left(0,\left[g'\left(\mu_X\right)\right]^2\sigma_X^2\right)$$

$$g\left(X_n\right)\overset{A}{\sim}N\left(g\left(\mu_X\right),\left[g'\left(\mu_X\right)\right]^2\frac{\sigma_X^2}{n}\right)$$

1. delta method 主要用來找 $\bar{X}_n$ 之函數的漸進分配，多從中央極限定理開始
2. 證明可從 (X_n) 在 μ_X 做一階泰勒展開出發
3. 我們常搭配 Slutsky 定理推導許多漸進分配

$\ll notes \gg$