6 常用的離散型機率分配

挑戰問題

請同學先用上一個單元所教的概念算一下這一題，然後我們再說明為什麼要學「常用的離散型機率分配」。

例題 6.1

擲一個公正的骰子一次，令隨機變數 X 表示出現的點數，試求

(1) X 之機率分配。

(2) $E(X),V(X)$。

解：(1) $P(X=x)=f(x)=\dfrac{1}{6} \qquad x=1,2,3,4,5,6$

一個骰子有六面，每面之點數出現的機率均等為 1/6，故為離散均等分配。

(2) ① 關於期望值的計算：

(a) 代入離散型均等分配的期望值公式：$E(X)=\dfrac{n+1}{2}=\dfrac{6+1}{2}=3.5$

(b) 從期望值的定義計算：

$$E(X)=1\left(\frac{1}{6}\right)+2\left(\frac{1}{6}\right)+3\left(\frac{1}{6}\right)+4\left(\frac{1}{6}\right)+5\left(\frac{1}{6}\right)+6\left(\frac{1}{6}\right)=3.5$$

② 關於變異數的計算：

(a) 代入離散型均等分配的變異數公式：

$$V(X)=\frac{n^2-1}{12}=\frac{36-1}{12}=2.92$$

(b) 從變異數的定義計算：

$$V(X)=(1-3.5)^2\frac{1}{6}+(2-3.5)^2\frac{1}{6}+(3-3.5)^2\frac{1}{6}+(4-3.5)^2\frac{1}{6}$$
$$+(5-3.5)^2\frac{1}{6}+(6-3.5)^2\frac{1}{6}=2.92$$

END

統計學是利用統計方法將資料轉成資訊的一門學問。有用的資訊不外集中趨勢（平均數）和離散趨勢（標準差），以及衡量不確定性下可能性大小的機率。如果有樣本資料，那麼透過敘述統計（統計圖、統計表、統計量）與機率論（事件機率）就可以直接算出來樣本的相關資訊了。同樣地，如果我們有母體的資料，依樣畫葫蘆即可「算機率」、「求特性」了。

可是可是可是，通常我們沒有母體的資料。統計學家想到利用「隨機實驗」的觀念，像上面擲骰子的例子一樣，透過古典機率的定義方式，就可以透過統計表、統計圖與數學式來表示母體的機率分配，不但可以「算機率」，也可以「求特性」了（計算期望值和變異數）。由於數學式子是最方便代表母體機率分配的表示方式，如果有一個數學式是能代表母體特性的機率分配，像上面擲骰子的例子一樣，我們就可以帶「原始定義」的公式來「算機率」、「求特性」了。

只是只是只是，數學式很抽象很難，我們用數學式把常用的機率分配表示出來有什麼好處呢？(1)可以用更簡單的「特殊」公式來「求特性」，就像上面例題，同學可以用「離散均等分配」的公式來算期望值和變異數，不需用「原始定義」的公式了；(2) 更好的是統計

學家已經替我們「算機率」了，只要知道是哪一種特殊的機率分配，就可以直接查表使用了。如果同學學一下統計軟體，那「算機率」、「求特性」的工作就可以讓電腦做了。

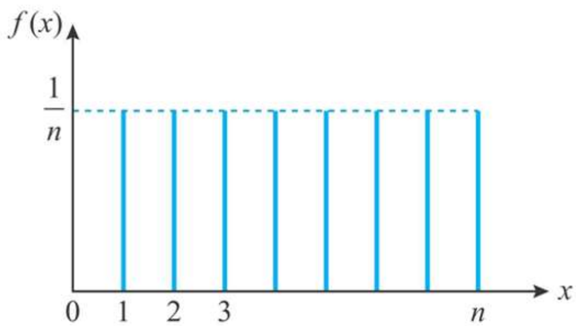
6-1 離散均等分配

1. 若離散型隨機變數 X 可能出現的值有 n 種可能性， 且每一個值出現的機率皆相同，即發生的機率為$1/n$，則此機率函數稱**離散均等分配 (discrete uniform distribution)**。

數學式： $X \sim uniform(n)$，有一個參數n

$$P(X=x)=f(x)=\frac{1}{n}, x=1,2,...,n$$

圖形：一條水平線



2. 求特性：期望值與變異數

- 期望值：$E(X)=\frac{n+1}{2}$
- 變異數：$V(X)=\frac{n^2-1}{12}$

Proof:

$$E(X)=1*\frac{1}{n}+2*\frac{1}{n}+...+n*\frac{1}{n}=\frac{1}{n}(1+2+...+n)=\frac{1}{n}*\frac{n(n+1)}{2}=\frac{n+1}{2}$$

$$V(X)=E(X^2)-[E(X)]^2=\frac{1}{n}(1^2+2^2+...+n^2)-[\frac{n+1}{2}]^2=\frac{1}{n}*\frac{n(n+1)(2n+1)}{6}-[\frac{n+1}{2}]^2=\frac{n^2-1}{12}$$

例題 5.1

一個箱子中有 5 個球，在球上編 1 至 5 號，並在箱子中隨機抽出一個球。設隨機變數 X 表示抽出球的編號，試求：

(1) X 的機率分配。

(2) $E(X), V(X)$。

解 (1) 隨機變數 X 表示抽出球的編號，可能出現 1 至 5 號中的一個數字，且每個球被抽出的機率皆相同，故為 $n=5$ 的離散均等分配。

$$P(X=x)=f(x)=\frac{1}{5} \qquad x=1,2,3,4,5$$

(2) 關於期望值的計算：

(a) 從期望值的定義計算：

$$E(X)=1\left(\frac{1}{5}\right)+2\left(\frac{1}{5}\right)+3\left(\frac{1}{5}\right)+4\left(\frac{1}{5}\right)+5\left(\frac{1}{5}\right)=3$$

(b) 代入離散均等分配的期望值公式:

$$E(X)=\frac{n+1}{2}=\frac{5+1}{2}=3$$

關於變異數的計算：

(a) 從變異數的定義計算：

$$V(X)=(1-3)^2\frac{1}{5}+(2-3)^2\frac{1}{5}+(3-3)^2\frac{1}{5}+(4-3)^2\frac{1}{5}+(5-3)^2\frac{1}{5}=2$$

(b) 代入離散均等分配的變異數公式:

$$V(X)=\frac{n^2-1}{12}=\frac{25-1}{12}=2$$

機率分配
- 離散型機率分配
  - 1. 離散均等分配
  - 2. 伯努力分配
  - 3. 二項分配
  - 4. 超幾何分配
  - 5. 卜瓦松分配
- 連續型機率分配
  - 1. 均勻分配
  - 2. 常態分配
  - 3. 卡方分配
  - 4. t 分配
  - 5. F 分配

學習重點：

- 常見的特殊離散型機率分配
  - 母體分配特性（產生樣本的方法，使用數學式）
  - 求算機率分配的特性（期望值與變異數）
  - 求算機率分配的機率 （手算與查表）

分配	隨機變數 X	符號	機率分配 $f(x)$	$E(x)$	$V(x)$	$M(t)$
百努利	不是成功就是失敗	$Ber(p)$	$p^x q^{1-x}$ ； $x=0,1$	p	pq	pe^t+q
二項	n 次獨立百努利實驗中成功次數	$B(n,p)$	$C^n_x p^x q^{n-x}$ ； $x=1,2,\dots,n$	np	npq	$(pe^t+q)^n$
超幾何	n 次不獨立百努利實驗中成功次數	$HG(N,K,n)$	$\frac{C^K_x C^{N-K}_{n-x}}{C^N_n}$ ；$x=1,2,\dots,n$	$n(\frac{K}{N})$	$n(\frac{K}{N})(1-\frac{K}{N})(\frac{N-n}{N-1})$	×
幾何	實驗至第一次成功所需實驗次數	$G(p)$	pq^{x-1} ； $x=1,2,\dots.$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pe^t}{1-qe^t}$
負二項	實驗至第 r 次成功所需實驗次數	$Nb(r,p)$	$C^{x-1}_{r-1} p^r q^{x-r}$ ；$x=r,r+1,\dots$	$\frac{r}{p}$	$\frac{rq}{p^2}$	$(\frac{pe^t}{1-qe^t})^r$
Poission	一段時間內發生的次數	$Poi(\lambda)$	$\frac{e^{-\lambda}\lambda^x}{x!}$ ；$x=0,1,2,\dots$	λ	λ	$e^{\lambda(e^t-1)}$

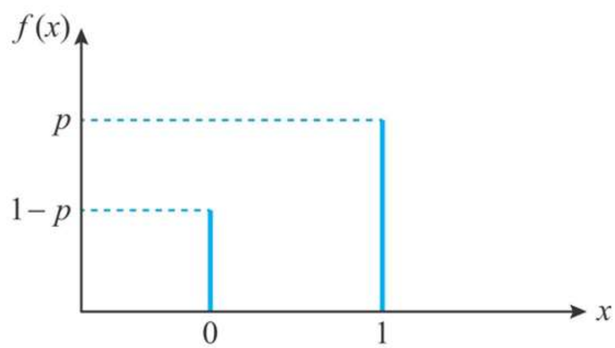
6-2 伯努力分配

1. 若離散型隨機變數 X 可能出現的結果只有兩種：成功或失敗；成功的發生機率為 p，失敗的發生機率為 $1-p$ 。試行一次，出現的結果的機率分配稱為伯努力分配 (Bernoulli distribution)，又名點二項分配。可用 $X\sim B(1,p)$來表示，其中 p為此機率分配的參數，$0\le p\le 1$ 。

數學式： $X\sim B(1,p)$，有一個參數p為成功機率。

$$P(X=x)=f(x)=p^x(1-p)^{1-x}, x=0,1$$

圖形：



2. 求特性： 期望值與變異數

- 期望值：$E(X)=p$
- 變異數：$V(X)=p(1-p)$

proof:

$$E(X)=0*p^0(1-p)^1+1*p^1(1-p)^0=p$$

$$V(X)=E(X^2)-[E(X)]^2=0^2*p^0(1-p)^1+1^2*p^1(1-p)^0-p^2=p-p^2=p(1-p)$$

3. **伯努力（隨機）實驗 (Bernoulli trial)**

- 一個隨機實驗出現的結果共有兩種可能性。
  - 一種稱為成功事件 ，另一種則稱為失敗事件 。
    - 出現成功就不可能失敗
    - 出現失敗就不可能成功
- 成功事件發生的機率為$ p$，失敗事件發生的機率為$1-p$ 。

例題 5.2

一個箱子中有 5 個色球，其中三個紅球、兩個白球，在箱子中隨機抽出一個球。設隨機變數 X 表示抽出白球的數量，試求：

(1) X 之機率分配。

(2) $E(X), V(X)$。

解 (1) 此隨機實驗可能的結果為紅球或白球（兩種），並且抽一個色球一次（試行一次），故此隨機實驗為伯努力實驗，機率分配為伯努力分配。隨機變數 X 表示抽出白球的數量，用白球數量占箱中球數的比例 2/5，訂出抽中白球的機率為 $p = 2/5 = 0.4$，而抽中紅球（非白球）的機率為 $1-p = 3/5 = 0.6$。

$$P(X=x) = f(x) = p^x(1-p)^{1-x} = (0.4)^x(0.6)^{1-x} \qquad x = 0, 1$$

(2) $E(X) = p = 0.4$

$V(X) = p(1-p) = 0.4 \times 0.6 = 0.24$

6-3 二項分配

6-3-1 基本觀念

1. 二項分配(Binominal Distribution)是二項實驗的結果，也就是描述試行 *n* 次伯努力實驗現象的機率分配。

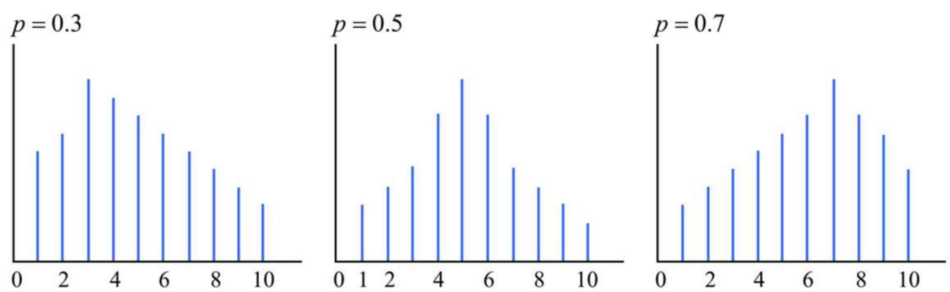
隨機變數X是在n次試驗中「成功」次數，隨機變數X的機率分配用 $X \sim B(n,p)$ 來表示，其中試行的次數 n和成功的機率 p 為此機率分配的參數，$0 \leq p \leq 1$ 。

數學式：$X \sim B(n,p)$，有試行次數 n和成功機率 p兩個參數

$$P(X=x) = f(x) = C_x^n p^x (1-p)^{1-x}, x = 0, 1, ..., n \text{，} 0 \leq p \leq 1 \text{，} C_x^n = \frac{n!}{x!(n-x)!}$$

圖形： 當$p = \frac{1}{2}$為對稱分配； \$p < 為右偏分配； \$$p > \frac{1}{2}$ 為左偏分配。

$n = 10$時之二項分配



2. 求特性：期望值與變異數

- 期望值：$E(X) = np$

$$E(X) = E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n) = p + p + ... + p = np$$

- 變異數：$V(X) = np(1-p)$

$$V(X) = V(X_1 + X_2 + ... + X_n) = V(X_1) + V(X_2) + ... + V(X_n) = p(1-p) + p(1-p) + ... + p(1-p) = np(1-p)$$

3. **二項實驗(binomial trial)**

隨機實驗若滿足下列四個條件，則此隨機實驗稱為二項實驗。

- 重複進行 n 次完全相同的試驗。
- 每一次試驗皆僅有兩種可能結果(outcome)；其一稱為「成功」(S)，另一則為「失敗」(F)。
- 每一次試驗中，出現成功結果的機率固定皆為p，出現失敗結果的機率固定為皆為$q(=1-p)$
- 每一次試驗之間皆互為獨立。

4. **算機率**

- 方法一：可用手算，直接帶機率分配的公式計算
- 方法二：也可以查表(Optional)
- 方法三：在實務上，常用常態分配來計算二項分配的機率值(後面會教)

例題 6.6

張先生新婚，計畫婚後生育三個小孩（假設均為單胞胎），令 X 為生男孩的個數，試求：

(1) X 之機率分配。

(2) $E(X), V(X)$。

(3) 僅有一個男孩的機率。

(4) 恰有二個男孩的機率。

(5) 全部女孩的機率。

解： (1) 因為生育不是男孩就是女孩（機率各半），而且第一次生育性別與第二次生育性別無關，張先生打算有三個小孩 $n=3$，所以符合二項實驗的條件，故可以使用二項分配來描述其機率分配狀況。

① 令 X 為生男孩的個數，

$$P(X=x)=f(x)=C_x^n p^x q^{n-x}=C_x^3\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{3-x} \qquad x=0,1,2,3$$

$$P(X=3)=f(3)=C_3^3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^0=\frac{1}{8}$$

$$P(X=2)=f(2)=C_2^3\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1=3\times\frac{1}{8}=\frac{3}{8}$$

$$P(X=1)=f(1)=C_1^3\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^2=3\times\frac{1}{8}=\frac{3}{8}$$

$$P(X=0)=f(0)=C_0^3\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^3=\frac{1}{8}$$

② 以此例，說明二項係數出現的原因。

將其所生三個小孩的性別狀況全部列出，如下表：

所有的可能性 （第一個，第二個，第三個）	男孩的人數 (X)	對應的機率
男孩，男孩，男孩	3	$\left(\frac{1}{2}\right)^3=\frac{1}{8}$
男孩，男孩，女孩	2	$\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1=\frac{1}{8}$
男孩，女孩，男孩	2	$\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1=\frac{1}{8}$
女孩，男孩，男孩	2	$\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1=\frac{1}{8}$
女孩，女孩，男孩	1	$\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^2=\frac{1}{8}$
女孩，男孩，女孩	1	$\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^2=\frac{1}{8}$
男孩，女孩，女孩	1	$\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^2=\frac{1}{8}$
女孩，女孩，女孩	0	$\left(\frac{1}{2}\right)^3=\frac{1}{8}$

由上表很清楚地可以看到，除了三個小孩全部是男孩、三個小孩全部是女孩的狀況外，生兩男一女或一男兩女時，都有男孩女孩出生的順序問題；例如：生兩男一女的情況，就可以分成三種狀況：(1) 頭兩胎是男孩，最後生女孩；(2) 男孩、女孩、男孩；(3) 老大是女孩，帶兩個弟弟。因此同樣都是兩男一女，但是因為女孩出生的順序，又可以細分為上表中所顯示出來的三種可能性。因此張先生生出兩男一女的機率，應是上述三種狀況的總和，也就是 $\frac{1}{8}+\frac{1}{8}+\frac{1}{8}=3\times\frac{1}{8}$，其中的 3 就是二項係數。

整理上表所列出的狀況可得：

$$P(\text{三男})=P(X=3)=\frac{1}{8}$$
$$P(\text{二男一女})=P(X=2)=\frac{3}{8}$$
$$P(\text{一男二女})=P(X=1)=\frac{3}{8}$$
$$P(\text{三女})=P(X=0)=\frac{1}{8}$$

此機率分配即為所求。

例題 6.9

設 X 為二項隨機變數，期望值與變異數分別為 $E(X)=6, V(X)=3$，試求該機率分配的參數 n 與 p。

解：

$$\begin{cases} E(X)=np=6 \\ V(X)=npq=3 \end{cases}$$

解方程式，可得

$$\frac{E(X)}{V(X)}=\frac{np}{npq}=\frac{6}{3}\Rightarrow 6q=3\Rightarrow q=\frac{1}{2}$$

因此 $p=\frac{1}{2}$，代回原式可得

$$n\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)=3$$

$$n=12$$

參數 $n=12, p=\frac{1}{2}$，該隨機變數 X 之機率分配為 $X\sim B\left(12,\frac{1}{2}\right)$ END

6-3-2 補充說明

1. 查表算機率

- 查表前要先確定n與p的值，通常$n<25$。
- 然後再由二項機率表中，所提供的累積機率$P(X\le k)$，找出對應的k值。
- 機率的換算（點機率不為0）
  - $P(X<k)=P(X\le[k-1])$
  - $P(X=k)=P(X\le k)–P(X<k)$ = $P(X\le k)–P(X\le[k–1])$
  - $P(X>k)=1–P(X\le k)$
  - $P(X\ge k)=1–P(X<k)=1–P(X\le[k–1])$

表 1　二項分配

x 值	機率 $f(x)$
0	$f(0)$
1	$f(1)$
2	$f(2)$
$\vdots$	$\vdots$
$\vdots$	$\vdots$
n	$f(n)$
總計	1

表 2　二項分配值

C	表中數據的意義
0	$f(0)$
1	$f(0)+f(1)$
2	$f(0)+f(1)+f(2)$
$\vdots$	$\vdots$
$\vdots$	$\vdots$
n	$f(0)+f(1)+\cdots+f(n)=1$

例題 5.4

利用附表 1 二項分配累積機率表，求算下列機率：

(1) $n=12$, $p=0.5$, $P(X \le 8)$

(2) $n=12$, $p=0.5$, $P(2 \le X \le 8)$

(3) $n=10$, $p=0.2$, $P(X=3)$

(4) $n=25$, $p=0.7$, $P(X>12)$

解 (1) $n=12$, $p=0.5$, $P(X \le 8)$

首先，在最左方欄位找出 $n=12$ 的附表，再找 c 為「8」，往右查；再在最上方欄位 p 為「0.5」，往下查，交叉得到數值「0.927」，即為

$$P(X \le 8)=0.927$$

(2) $n=12$, $p=0.5$, $P(2 \le X \le 8)$

首先，在最左方欄位找出 $n=12$ 的附表，再找 c 為「1」與「8」，分別往右查；再在最上方欄位 p 為「0.5」，往下查，交叉得到數值「0.003」與「0.927」，即為

$$P(X \le 1)=0.003$$
$$P(X \le 8)=0.927$$
$$P(2 \le X \le 8)=P(X \le 8)-P(X \le 1)=0.927-0.003=0.924$$

(3) $n=10$, $p=0.2$, $P(X=3)$

首先，在最左方欄位找出 $n=10$ 的附表，再找 c 為「2」與「3」，分別往右查；再在最上方欄位 p 為「0.2」，往下查，交叉得到數值「0.678」與「0.879」，即為

$$P(X \le 2)=0.678$$
$$P(X \le 3)=0.879$$
$$P(X=3)=P(X \le 3)-P(X \le 2)=0.879-0.678=0.201$$

(4) $n=25$, $p=0.7$, $P(X>12)$

首先，在最左方欄位找出 $n=25$ 的附表，再找 c 為「12」，往右查；再在最上方欄位 p 為「0.7」，往下查，交叉得到數值「0.017」，即為

$$P(X \le 12)=0.017$$
$$P(X>12)=1-P(X \le 12)=1-0.017=0.983$$

【查表應用題】：已知機率分配（手算/查表）求機率

例題 5.6

假設研究指出有 20% 的寵物主人出國旅遊會將寵物寄養於寵物醫院，而不是託親友照顧。今隨機詢問 25 位準備出國旅遊的寵物主人，試求最多有 5 位寵物主人會將寵物寄養於寵物醫院的機率。

解　因寵物主人出國旅遊會將寵物寄養於寵物醫院（成功事件）或是託親友照顧（失敗事件），只會出現兩種結果。研究指出有 20% 的寵物主人出國旅遊會將寵物寄養於寵物醫院。隨機詢問 25 位準備出國旅遊的寵物主人，每位相互獨立。因此，此隨機實驗為一試行次數 $n=25$，成功機率 $p=0.2$ 的二項實驗，其機率分配可以用二項分配描述。

令隨機變數 X 表示隨機詢問 25 位準備出國旅遊的寵物主人會將寵物寄養於寵物醫院的人數，則 X 的機率分配為 $n=25, p=0.2$ 的二項分配，可簡記為 $X \sim B(25, 0.2)$，其機率分配為

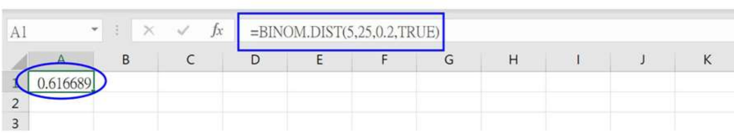
$$P(X=x)=f(x)=C_x^{25}(0.2)^x(0.8)^{25-x} \qquad x=0,1,2,\cdots,25$$

最多有 5 位寵物主人會將寵物寄養於寵物醫院的機率：

$$P(X \le 5)=\sum_{x=0}^{5} C_x^{25}(0.2)^x(0.8)^{25-x}$$

查附表 1 二項分配累積機率表，可得 $P(X \le 5)=0.617$。

利用 Excel 2016 操作，可得



Excel 2016 函式：= BINOM.DIST (5, 25, 0.2, TRUE)

2. 二項分配的加法性

假設$X \sim B(n_1, p), Y \sim B(n_2, p)$，則$X+Y \sim B\ (\ n_1+n_2, p)$

3. 二項分配可以延伸到多項機率分配(Optional)

Multi-nominal distribution: $f(x_1, x_2, \ldots, x_k)=\frac{n!}{x1!\ldots xk!}p_1^{x1}P_2^{x2}\ldots P_k^{xk}, x_i=0,1,\ldots n, \sum p_i=1$