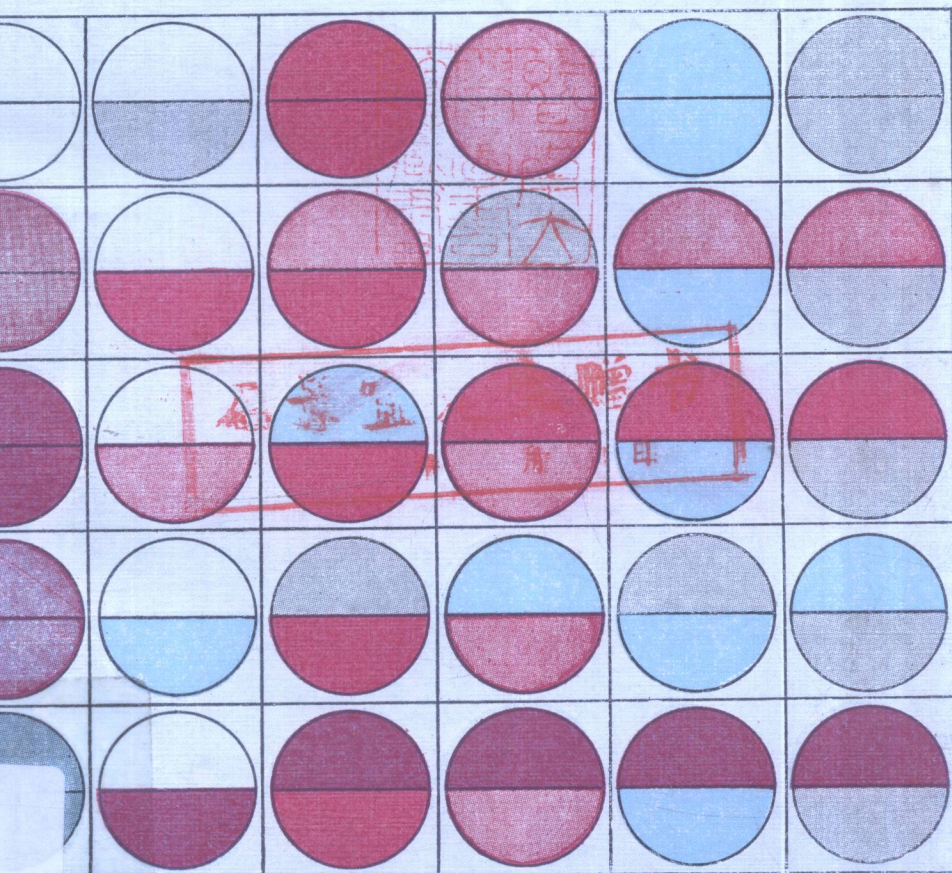


統計學

下 冊

編 著 者

廖 繼 敏



東華書局印行

887
2

018005



S9004183

統計學

下 冊

廖 繼 敏 編 著



石景宜先生贈書

年 月 日

東華書局印行



版權所有・翻印必究

中華民國六十二年十月初版

中華民國七十四年二月五版

大學
用書

統計學 (全二冊)

下冊 定價 新台幣九十元整

(外埠酌加運費滙費)

編著者 廖 繼 敏

發行人 卓 鑫 森

出版者 臺灣東華書局股份有限公司
臺北市博愛路一〇五號
電話：3819470 郵撥：6481

印刷者 中 臺 印 刷 廠
臺中市公園路三十七號

行政院新聞局登記證 局版臺業字第零柒貳伍號
(62033)

統計學

下冊目次

章節	頁碼
第九章 間斷機率分配.....	1-26
導論。等機率分配。貝努力分配。二項分配。 超幾何分配。波以生分配。負二項分配。	
第十章 常態分配.....	27-48
常態曲線。常態曲線下的面積。 二項分配的常態近似值。	
第十一章 抽樣原理.....	49-92
導論。隨機抽樣。抽樣設計。平均數的抽樣分配。 兩平均數差之抽樣分配。比率之抽樣分配。 大數法則及中央極限定理。 t 分配。 χ^2 分配。 F 分配。	
第十二章 估計原理.....	93-135
決策原理。傳統的估計方法。最大概似法。 平均數的區間估計。兩平均數差的估計。 比率之估計。變異數之估計。兩變異數比之估計。	
第十三章 假設的檢定.....	136-197
統計的假設。第一過誤與第二過誤。 一尾及兩尾檢定。平均數的檢定。 兩平均數差的檢定。比率之檢定。檢定力函數。 變異數之檢定。兩變異數相等之檢定。 配合度之檢定。獨立性之檢定。	

第十四章	迴歸與相關.....	198-229
	直線迴歸。母體迴歸線之點估計。	
	母數之區間估計及假設檢定。	
	母體迴歸線的信賴區間及預測。	
	迴歸之直線性的檢定。指數迴歸。複迴歸。	
	相關。	
第十五章	變異數分析.....	230-258
	導論。一因子變異數分析。	
	兩因子變異數分析，每方格一觀察值。	
	兩因子變異數分析，每方格有數觀察值。	
附 錄		259-270

第九章

間斷機率分配

9.1 導論 (Introduction)

本章將涉及間斷機率分配 (Discrete probability distributions)，前於第三章已就間斷機率分配用表、圖或公式加以表示，許多隨機變數具有間斷的性質，其隨機變數之可能數值之間有一固定數值的差異。換言之，間斷隨機變數的樣本空間為由有限的或無限但可數的數值或樣本點構成。例如，一硬幣試行 n 次，其出現正面只有 n 種可能的結果，而沒有其他分數單位的數值，若每次試行出現正面的機率皆相同，代表此試驗出現正面數之所有母全體皆有相同之形式，因而可用一簡單公式表示之。若拋擲一硬幣，以出現正面代表成功，則拋擲五次中出現正面的分配公式，其結構與投擲一骰子七次有四次成功的分配公式相同。

選擇一機率分配以正確說明由實驗所導出的結果，應予審慎。本章將探討用以敘述常用之母全體的幾個重要間斷機率分配。

9.2 等機率分配 (Uniform Distribution)

間斷機率分配中之最簡單者為隨機變數之所有值皆有相等的機率，此一機率分配稱為等機率分配。

定義： 若隨機變數 X 的樣本空間為 $S = \{x_1, x_2, \dots, x_k\}$ ，則間斷等機率分配為：

$$f(x; k) = \frac{1}{k}, x = x_1, x_2, \dots, x_k.$$

上式中以 $f(x; k)$ 代替 $f(x)$ 以表示等機率分配隨母數 k 而不同。

例 9.1 投擲一粒骰子，樣本空間 $S = \{1, 2, 3, 4, 5, 6\}$ 中，每一元素出現的機率皆為 $\frac{1}{6}$ ，故可得一等機率分配如下：

$$f(x; 6) = \frac{1}{6}, x = 1, 2, 3, 4, 5, 6.$$

例 9.2 自一副牌中隨機抽出一張，每一張被抽出的機率皆為 $\frac{1}{52}$ ，因此，其等機率分配如下：

$$f(x; 52) = \frac{1}{52}, x = 1, 2, \dots, 52.$$

此處將每一張牌予以編號，使每一張牌各以不同的號碼代表之。

等機率分配可以直方圖表示，隨機變數之每一數值，皆具有相同高度的長方圖，例 9.1 的直方圖以圖 9.1 表示之。

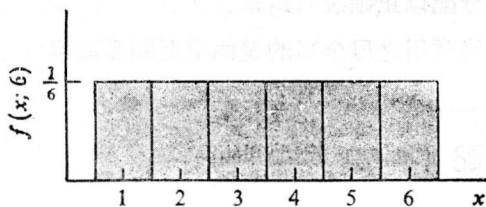


圖 9.1 投擲一骰子之直方圖

自第四章的隨機變數的定義可知，自一有限母全體所抽出所有可

能樣本的分配爲等機率。如自一含有 A, B, C 及 D 等四個牧師的母全體中隨機抽出樣本爲兩人，其可能樣本的總數爲 $\binom{4}{2} = 6$ ，列示如下： AB, AC, AD, BC, BD, CD 。因每一組被抽出的機率皆相同，故此樣本的分配爲等機率分配，即

$$f(x;6) = \frac{1}{6}, \quad x=1, 2, \dots, 6.$$

因此

$$Pr(B \text{ 及 } D \text{ 皆被選出}) = Pr(x=5) = f(5;6) = \frac{1}{6}.$$

等機率分配中的 K 爲自一有限母全體 N 中隨機抽出樣本 n ，即

$$\binom{N}{n}.$$

例 9.3 在一年中抽取含有三個月的隨機樣本，求其等機率分配。

解：因一年有十二個月，隨機抽取三個月有 $\binom{12}{3} = 220$ 種方法，以號碼

1 至 220 表各方法，其機率分配爲

$$f(x;220) = \frac{1}{220}, \quad x=1, 2, \dots, 220.$$

故選擇號碼爲 5 之樣本的機率爲

$$f(5;220) = \frac{1}{220}.$$

9.3 貝努力分配 (Bernoulli Distribution)

許多的隨機現象具有貝努力程序，此爲一甚簡單的機率分配——僅包含兩個互相排斥及周延的事件。例如拋擲一硬幣，只有兩種可能結果，即正面或反面。自一大學中選出一學生，被選出者必爲男生或

女生，爲已結婚或未結婚。將這些名詞標準化可得許多相同的程序，可歸納爲“成功”及“失敗”等兩種可能結果，就此兩種事件數量化而以“1”代表成功，以“0”代表失敗，並令 Y 爲貝努力隨機變數，其樣本空間爲 $S=\{0,1\}$ ，以 p 及 q 各代表成功及失敗的機率，則隨機變數 Y 的機率分配如表 9.1 所示，此一分配稱爲貝努力機率分配。

表 9.1 貝努力隨機變數 Y 的機率分配

y	$P(y)$
0	$q=1-p$
1	p

與其他分配相似，貝努力分配亦有平均數及變異數，表 9.2 說明計算平均數及變異數的程序， Y 的平均數爲：

表 9.2 計算貝努力隨機變數 Y 的平均數及變異數的程序

y	$p(y)$	y^2	$y^2 p(y)$
0	$1-p$	0	0
1	p	1	p
合計			p

$$\begin{aligned}
 E(Y) &= 0(1-p) + 1(p) \\
 &= 0 + p \\
 &= p
 \end{aligned}$$

Y 的變異數及標準差以 σ_y^2 及 σ_y 表示，可自下式求得：

$$\sigma_y^2 = E(Y^2) - [E(Y)]^2$$

$$\begin{aligned}
 &= 0^2(1-p) + 1^2(p) - p^2 \\
 &= p - p^2 \\
 &= p(1-p) \\
 &= pq
 \end{aligned}$$

及

$$\sigma_y = \sqrt{pq}$$

例如，在投擲一硬幣中，獲得正面（表示成功）的機率為 0.5，得到反面（表示失敗）的機率為 0.5，則此一分配的平均數為 0.5，變異數為 $0.5 \times 0.5 = 0.25$ ，及標準差為 $\sqrt{0.25} = 0.50$ 。

9.4 二項分配 (Binomial Distribution)

若自貝努利分配中隨機抽出 n 個樣本，令此樣本為 $(Y_1, Y_2, \dots, Y_n)$ ，若該試行的結果為成功，則以值 1 表示之，而以 0 表示此結果為失敗，在 n 次試行中出現 1 的數目為： $(Y_1 + Y_2 + \dots + Y_n)$ 或 $\sum_{i=1}^n Y_i$ ，今以一新的隨機變數 X 表示二項隨機變數，則在 n 個樣本中成功的數目為：

$$X = \sum_{i=1}^n Y_i$$

此一實驗 (experiment) 為重複獨立試行若干次，每次皆有成功及失敗等兩種可能的結果，如投擲硬幣五次，每次試行的可能結果為正面或反面，此種實驗稱為二項實驗。

二項實驗具有下列特性：

1. 每一實驗包括 n 個重複試行。
2. 每次試行的結果可分類為成功或失敗。

3. 以 p 表示成功的機率，每一試行的 p 皆相同。

4. 重覆試行皆具獨立性。

就投擲硬幣三次而以正面表示成功為例，隨機變數 X 表成功的數目，其值為自 0 至 3，其八種可能的結果及其對應的 X 值為

結果	TTT THT TTH HTT THH HTH HHT HHH							
x	0	1	1	1	2	2	2	3

因試行皆獨立，且每次成功的機率皆為 $\frac{1}{2}$ ，如 $Pr(HTH) = Pr(H)Pr(T)Pr(H) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{8}$ ，其他任一可能結果的機率亦為 $\frac{1}{8}$ ， X 的機率分配如下：

x	0	1	2	3
$Pr(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

此時可以下列公式表此機率分配：

$$f(x) = \frac{\binom{3}{x}}{8}, \quad x=0, 1, 2, 3.$$

定義：在試行 n 次的二項實驗中，成功的數目 X 稱為二項隨機變數。

二項變數 X 的機率分配稱為二項分配，以 $b(x; n, p)$ 表示，因其

值決定於試行次數及每次試行成功的機率，故在投擲三硬幣中，其機率分配為：

$$b\left(x; 3, \frac{1}{2}\right) = \frac{\binom{3}{x}}{8}, \quad x=0, 1, 2, 3.$$

今可將上面說明化為一般公式 $b(x; n, p)$ ，表示在二項實驗中， n 次試行得 x 次成功的機率，因每次皆為獨立試行，故可將每次不同結果的機率相乘，每次成功的機率為 p ，失敗的機率為 $q = 1 - p$ ，因此，按一特定次序的機率為 $p^x q^{n-x}$ ，今可決定在一次實驗中得 x 次成功及 $n-x$ 次失敗的樣本點之總數，亦即將 n 個結果分成 x 及 $n-x$ 兩組而加以組合，可得 $\binom{n}{x}$ 之數目，因每一分割皆為互斥，故可將所有不同分割之機率相加或以 $\binom{n}{x}$ 乘 $p^x q^{n-x}$ 即得一般公式。

二項分配：若二項試行成功的機率為 p ，失敗的機率為 $q = 1 - p$ ，則在 n 次獨立試行中，以二項隨機變數 X 表示的成功次數的機率分配為：

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, \quad x=0, 1, 2, \dots, n.$$

當 $n=3$, $p=\frac{1}{2}$ 時，

$$b\left(x; 3, \frac{1}{2}\right) = \binom{3}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{3-x} = \frac{\binom{3}{x}}{8},$$

例 9.4 投擲一骰五次，求恰得三次兩點之機率。

解：在五次獨立試行中，每次成功的機率為 $\frac{1}{6}$ ，失敗的機率為 $\frac{5}{6}$ ，出現兩點視為成功，則：

$$\begin{aligned} b\left(2; 4, \frac{3}{4}\right) &= \binom{4}{2} \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^2 \\ &= \frac{4!}{2!2!} \cdot \frac{3^2}{4^4} \\ &= 0.211 \end{aligned}$$

二項分配源於 $(q+p)^n$ 的展開式，計有 $(n+1)$ 項，其各項相當於 $b(x; n, p)$ 在 $x=0, 1, 2, \dots, n$ 之各值。即

$$\begin{aligned} (q+p)^n &= \binom{n}{0} q^n + \binom{n}{1} p q^{n-1} + \binom{n}{2} p^2 q^{n-2} + \dots + \binom{n}{n} p^n \\ &= b(0; n, p) + b(1; n, p) + b(2; n, p) + \dots \\ &\quad + b(n; n, p)。 \end{aligned}$$

因 $p+q=1$ ，故 $\sum_{x=0}^n b(x; n, p)=1$ ，此符合機率分配的條件。

若欲求 $Pr(X < r)$ 或 $Pr(a \leq X \leq b)$ ，可由附錄表 A·2 之二項總和 $\sum_{x=0}^r b(x; n, p)$ 求得。

例 9·5 生壞血病者復原之機率為 0.4，若十五人已患此病，求 (a) 至少有十人復原，(b) 三至八人復原之機率。

解：(a) 以 X 表復原之人數，則

$$\begin{aligned} Pr(X \geq 10) &= 1 - Pr(x < 10) \\ &= 1 - \sum_{x=0}^9 b(x; 15, 0.4) \\ &= 1 - 0.9662 \\ &= 0.0338。 \end{aligned}$$

$$\begin{aligned}
 (b) \Pr(3 \leq X \leq 8) &= \sum_{x=3}^8 b(x; 15, 0.4) \\
 &= \sum_{x=0}^8 b(x; 15, 0.4) - \sum_{x=0}^2 b(x; 15, 0.4) \\
 &= 0.9050 - 0.0271 \\
 &= 0.8779.
 \end{aligned}$$

證明：令第 j 次試行的結果以隨機變數 I_j 表示，其值為 0 及 1，各以機率 q 及 p 表示。此稱為指示變數 (indicator Variable)，因 $I_j = 0$ 指示失敗及 $I_j = 1$ 指示成功。二項實驗之成功數可以 n 個獨立指示變數之和表示，即：

$$X = I_1 + I_2 + \cdots + I_n.$$

I_j 的平均數為 $E(I_j) = 0 \cdot q + 1 \cdot p = p$ ，應用定理 3.6，得二項分配的平均數為

$$\begin{aligned}
 \mu &= E(X) = E(I_1) + E(I_2) + \cdots + E(I_n) \\
 &= \underbrace{p + p + \cdots + p}_{n \text{ 項}} \\
 &= np.
 \end{aligned}$$

任一 I_j 的變異數為

$$\begin{aligned}
 \sigma_{I_j}^2 &= E[(I_j - p)^2] = E(I_j^2) - p^2 = (0)^2 q + (1)^2 \cdot p - p^2 \\
 &= p(1 - p) = pq.
 \end{aligned}$$

因此，依定理 4.4，二項分配的變異數為

$$\begin{aligned}
 \sigma_x^2 &= \sigma_{I_1}^2 + \sigma_{I_2}^2 + \cdots + \sigma_{I_n}^2 \\
 &= \underbrace{pq + pq + \cdots + pq}_{n \text{ 項}} \\
 &= npq.
 \end{aligned}$$

定理 9.1 二項分配 $b(x; n, p)$ 的平均數及變異數為
及 $\mu = np$ 及 $\sigma^2 = npq$.

例 9.6 應用經驗法則求例 9.5 之區間 $\mu \pm 2\sigma$ 並加解釋。

解：例 9.5 爲 $n=15$ 及 $p=0.4$ 之二項分配，由定理 9.1 得

$$\mu = (15)(0.4) = 6 \text{ 及 } \sigma^2 = (15)(0.4)(0.6) = 3.6.$$

求 3.6 的平方根得 $\sigma = 1.897$ ，故所求之區間爲 $6 \pm (2)(1.897)$ ，或自 2.206 至 9.794，當 p 趨近於 0.5 時，二項分配的直方圖趨近於鐘形分配，可得如下結論：在任一以十五人爲一組的病患中，其復原人數在 2.206 至 9.794 之間者有 95% 的機率。

若每一試行之結果有二種以上的可能結果，則二項實驗變爲多項實驗 (multinomial experiment)，如拋擲二粒骰子，其可能的點數和爲 7 或 11，或其他任何情況，此構成多項實驗。自一副牌中以歸還法取牌，按四色牌記載，亦爲多項實驗。

若一試行的結果爲 k 種可能結果 $E_1, E_2, \dots, E_k$ 之任一種，其機率各爲 $p_1, p_2, \dots, p_k$ ，則多項分配爲在 n 次獨立試行中， E_1 出現 x_1 次， E_2 出現 x_2 次， $\dots$ ， E_k 出現 x_k 次的機率，其 $x_1 + x_2 + \dots + x_k = n$ ，以 $f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n)$ 表聯合機率分配，其 $p_1 + p_2 + \dots + p_k = 1$ ，因每一試行的結果爲 k 種可能結果的任何一種。

今可導出一般公式，因每次試行皆獨立，任一獲得 x_1 次 E_1 ， x_2 次 E_2 ， $\dots$ ， x_k 次 E_k 的特定次序的機率爲 $p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$ ，在 n 次試行獲得相同結果的各種次序的總數爲將 n 項分割爲第一組含 x_1 項，第二組含 x_2 項， $\dots$ ，第 k 組含 x_k 項的 k 組組合數目，即：

$$\binom{n}{x_1, x_2, \dots, x_k} = \frac{n!}{x_1! x_2! \dots x_k!} \text{ 種方法}$$

因此分割爲互相排斥且發生之機率相等，將特定次序之機率乘其分割總數即得多項分配 (Multibinomial Distribution)。

多項分配: 若一試行有機率為 $p_1, p_2, \dots, p_k$ 的 k 種結果 $E_1, E_2, \dots, E_k$, 則以隨機變數 $X_1, X_2, \dots, X_k$ 代表 $E_1, E_2, \dots, E_k$ 出現的數目, 各隨機變數的聯合機率分配為:

$$f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n) \\ = \binom{n}{x_1, x_2, \dots, x_k} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k},$$

$$\text{其中 } \sum_{i=1}^k x_i = n \text{ 與 } \sum_{i=1}^k p_i = 1.$$

多項分配為由 $(p_1 + p_2 + \dots + p_k)^n$ 的多項展開式, 其各項相當於 $f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n)$ 之所有可能數值之一。

例 9.7 若兩骰拋擲六次, 求獲得兩次其和為 7 或 11, 一次兩骰點數相同, 三次各種其他組合的機率。

解: 各可能事件列示如下:

E_1 : 總和為 7 或 11 的事件,

E_2 : 兩骰點數相同的事件,

E_3 : 非兩骰點數相同, 亦非總和為 7 或 11 的事件。

在每一試行中, E_1, E_2 及 E_3 的相對機率為 $p_1 = \frac{2}{9}$, $p_2 = \frac{1}{6}$, 與 $p_3 = \frac{11}{18}$, 各機率在六次試行中皆相同, 以 $x_1 = 2, x_2 = 1$, 及 $x_3 = 3$, 求多項分配的機率為:

$$f\left(2, 1, 3; \frac{2}{9}, \frac{1}{6}, \frac{11}{18}, 6\right) = \binom{6}{2, 1, 3} \left(\frac{2}{9}\right)^2 \left(\frac{1}{6}\right)^1 \left(\frac{11}{18}\right)^3 \\ = \frac{6!}{2!1!3!} \cdot \frac{2^2}{9^2} \cdot \frac{1}{6} \cdot \frac{11^3}{18^3} \\ = 0.1127.$$

9.5 超幾何分配 (Hypergeometric Distribution)

二項分配是建立在母全體為無限與隨機樣本為歸還抽樣的假設上，故其每次的觀察值皆為獨立，亦即 n 次試行皆為獨立試行的實驗。當母全體為有限，且隨機抽樣為不歸還，(Without replacement)，則每一觀察值的機率皆不同，此情況之機率分配稱為超幾何分配。

若自一副牌中隨機抽取五張，而欲求在26張紅牌中抽出3張及26張黑牌中抽出2張的機率，抽出3張紅牌有 $\binom{26}{3}$ 種方法，抽出2張黑牌有 $\binom{26}{2}$ 種方法，因此，在抽出5張牌中為3張紅牌2張黑牌之取法總數為 $\binom{26}{3}\binom{26}{2}$ 。自一副牌抽取5張的方法為 $\binom{52}{5}$ 種，故在不歸還抽樣中，抽取5牌而獲得3紅牌及2黑牌的機率為：

$$\frac{\binom{26}{3}\binom{26}{2}}{\binom{52}{5}} = \frac{26!}{3!23!} \cdot \frac{26!}{2!24!} = 0.3251,$$

$$\frac{52!}{5!47!}$$

自含有 N 個之有限母全體中，取 n 個為一隨機樣本，而欲求自 K 項表成功之中抽取 x 次成功及自 $N-k$ 項表失敗中抽取 $n-x$ 次失敗的機率，此稱為超幾何實驗。

超幾何實驗具有如下兩性質：

1. 隨機樣本數 n 為自母全體為 N 項中所抽出。
2. N 項中的 k 項屬於成功之一類， $N-k$ 項屬於失敗之一類。

定義： 在超幾何實驗中，成功之個數 X 稱為超幾何隨機變數。