第七章 離散型機率分配

本章內容

7-1 一致分配(uniformly distribution)

7-2 二項分配(binomial distribution)

7-3 幾何分配 (geometric distribution)

7-4 超幾何分配 (hypergeometric distribution)

7-5 卜瓦松分配 (Poisson distribution)

本章介紹幾種常用的離散型機率分配：

(1)一致分配 (uniformly distribution)

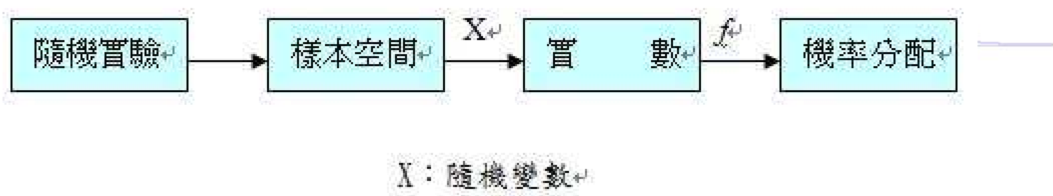
(2)白努利分配 (Bernoulli distribution)

(3)二項分配 (binomial distribution)

(4)超幾何分配 (hyper-geometric distribution)

(5)幾何分配 (geometric distribution)

(6)卜瓦松分配 (Poisson distribution)



7-1 一致分配

我們考慮一個如下的隨機實驗，假設有一個個數為 N 的母體，成員為 $x_1, x_2, \cdots, x_N$ ，今從中隨機抽出任意一個，令隨機變數 X 代表抽出的結果。這個實驗中，我們假設母體中每一個成員被抽出的機率相同，即隨機變數 X 的機率分配函數為

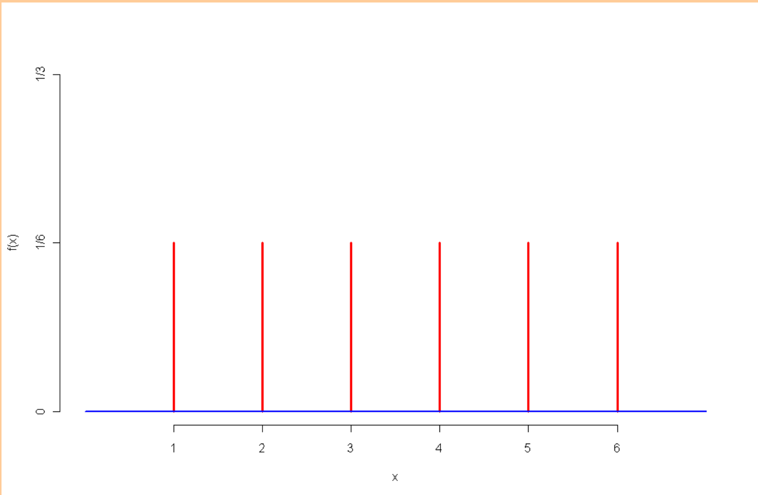
$$f(x_i) = P(X = x_i) = \frac{1}{N}, \quad \text{i} = 1, 2, \ldots, N$$

我們稱上述機率分配為一致分配。

例 7- 1

擲一個公平的骰子，令 X 代表出現的點數，則機率分配函數為

$$f(i)=P(X=i)=\frac{1}{6}\,,\qquad \mathrm{i}=1,2,\ldots,6$$



一致分配的期望值與變異數

假設隨機變數 X 的所有可能結果為 $x_1,x_2,\cdots,x_N$，且其具有一致分配。根據期望值與變異數的定義，我們得

期望值 $\mu=E(X)=\sum_{i=1}^{N}x_i f(x_i)=\sum_{i=1}^{N}x_i\frac{1}{N}=\sum_{i=1}^{N}\frac{x_i}{N}$ 式 7- 1

變異數 $\sigma^2=Var(X)=E\left(X-u\right)^2=\sum_{i=1}^{N}(x_i-\mu)^2 f(x_i)=\sum_{i=1}^{N}\frac{(x_i-\mu)^2}{N}$ 式 7- 2

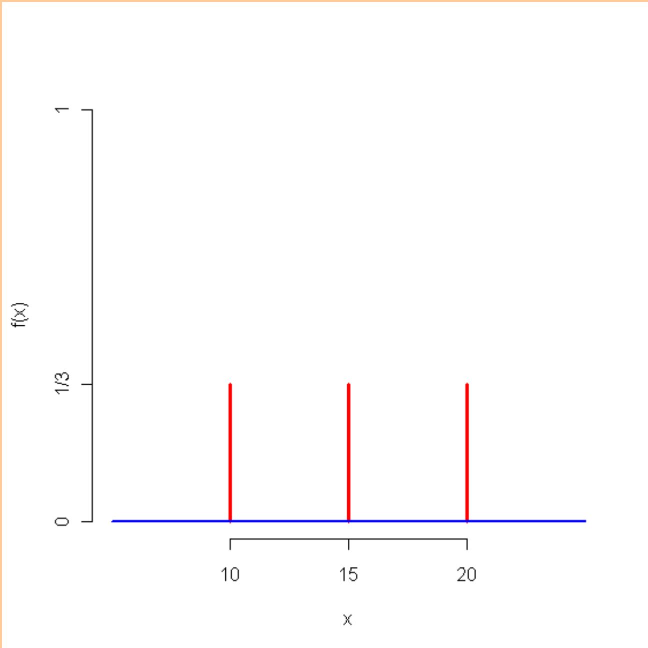
例 7- 2

某便利商店舉行折價活動，凡購買抽過 100 元的顧客可在摸彩箱摸出一顆彩球。假設摸彩箱內有三顆彩球，黃、紅和藍，黃色可折價 20 元，紅色彩球可折價 15 元，藍色彩球可折價 10 元。令隨機變數 X 代表顧客折價金額，請問 X 的期望值與變異數。

解：

根據題意，X 的機率分配為

$$f(x)=\frac{1}{3}\text{，}\ x=10,15,20$$



即 $x_1=10, x_2=15, x_3=20$ ，所以

$$\mu=E(X)=\sum_{i=1}^{3}x_i(\frac{1}{3})=\frac{10+15+20}{3}=15$$

$$\sigma^2=\sum_{i=1}^{N}\frac{(x_i-\mu)^2}{N}=\frac{(10-15)^2+(15-15)^2+(20-15)^2}{3}=\frac{50}{3}$$

當 $x_i=i$ ，則

$$E(X)=\frac{\sum_{i=1}^{N}i}{N}=\frac{1+2+...+N}{N}=\frac{N(N+1)}{2}\Big/N=\frac{N+1}{2}$$

$$Var(X)=\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}=\frac{\sum_{i=1}^{N}x_i^2-N\mu^2}{N}$$

$$=\frac{\frac{N(N+1)(2N+1)}{6}-\frac{N(N+1)^2}{4}}{N}$$

$$=\frac{(N+1)(4N+2-3N-3)}{12}=\frac{N^2-1}{12}$$

7-2 二項分配

在介紹二項分配(binomial distribution)前，我們先介紹一個最為基礎的白

努利試驗（Bernoulli trial）。假若一個隨機實驗的可能結果僅有兩種，我們稱此隨機實驗為白努利試驗。例如：

（a）擲一個銅板的隨機實驗，這個實驗只會出現正面 H 和反面 T，其樣本空間為 $S=\{H,T\}$。

（b）一個學生修統計學是否會被當，其結果也只有兩種？ S={Yes,No}

（c）一個問卷調查中受訪者性別？ S={Male,Female}

令 X 代表一個白努利試驗結果，通常記成

$$X=\begin{cases}1,\ \text{Head, Yes, Male}\\0,\ \text{Tail, No, Female}\end{cases}$$

它的機率分配函數為

$$f(x)=P(X=x)=p^{x}(1-p)^{1-x},\quad x=0,1 \qquad \text{式 7- 3}$$

其中 $p=P(X=1)$ 代表此一試驗成功的機率，所以失敗的機率為 $q=1-p=P(X=0)$。

$$\begin{aligned}P(X=0)&=f(0)=p^{0}(1-p)^{1-0}=1-p\\P(X=1)&=f(1)=p^{1}(1-p)^{1-1}=p\end{aligned}$$

進一步可計算白努利隨機變數的期望值與變異數為

$$E(X)=p,Var(X)=p(1-p)$$

證明：

$$E(X)=\sum_{X=0}^{1}xf(x)=0f(0)+1\times f(1)=p$$

$$E(X^2)=\sum_{X=0}^{1}x^2f(x)=f(0)+1f(1)=p$$

$$Var(X)=E(X^2)-(E(X))^2=p-p^2=p(1-p)$$

現在，我們重複執行相同的白努利試驗，假設這些試驗的結果為 $X_1,X_2,\dots,X_n$，它們是 n 個獨立且有相同的成功率 p 的白努利試驗，令 X 代表這些白努利試驗成功次數的和，即

$$X=X_1+X_2+\cdots+X_n$$

我們稱為 X 稱二項隨機變數。

例 7- 3

(1) 令 X_i 代表第 i 次擲銅板出現的結果，這是一個白努利試驗。

$$X_i = \begin{cases} 1, & \text{Head} \\ 0, & \text{Tail} \end{cases}$$

令 X 代表投擲 20 銅板出現正面的次數，即

$$X = X_1 + \ldots + X_{20}$$

(2) 一個問卷調查中，我們隨機訪問 10 位同學，令 X_i 第 i 個受訪者的性別。

$$X_i = \begin{cases} 1, \text{Male} \\ 0, \text{Female} \end{cases}$$

讓 $X = X_1 + X_2 + \cdots + X_{10}$ 代表這此訪問中受訪男生總數。

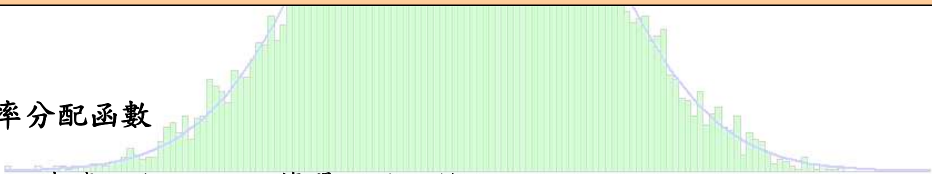
假設問卷結果如下

$$X_1 = 1, X_2 = 1, X_3 = 0, X_4 = 1, X_5 = 0, X_6 = 1, X_7 = 0, X_8 = 0, X_9 = 1, X_{10} = 0$$

則

$$X = 1+1+0+1+0+1+0+0+1+0 = 5$$

即有 5 位男生受訪，相對地也有 5 位女生受訪。



二項機率分配函數

首先，我們考慮一個 $n = 2$ 的簡單例子，換言之，

$$X = X_1 + X_2$$

隨機變數 X 的所有可能結果為 0, 1, 2。下表列出各種可能發生的機率，我們可以將隨機變數 X 的機率分配函數表如下

$$f(x) = P(X = x) = C_x^2 p^x (1-p)^{2-x}, x = 0,1,2$$

結果 X	(X_1, X_2) 的組合	組合數	機率
0	(0, 0)	$C_0^2 = 1$	$(1-p)^2$
1	(0, 1), (1, 0)	$C_1^2 = 2$	$2p(1-p)$
2	(1, 1)	$C_2^2 = 1$	p^2

我們作一簡單的推廣，考慮

$$X = X_1 + X_2 + X_3$$

則 X 的所有可能結果為 0, 1, 2, 3。同樣地，我們下表列出各種可能發生的機率，我們可以將隨機變數 X 的機率分配函數表如下

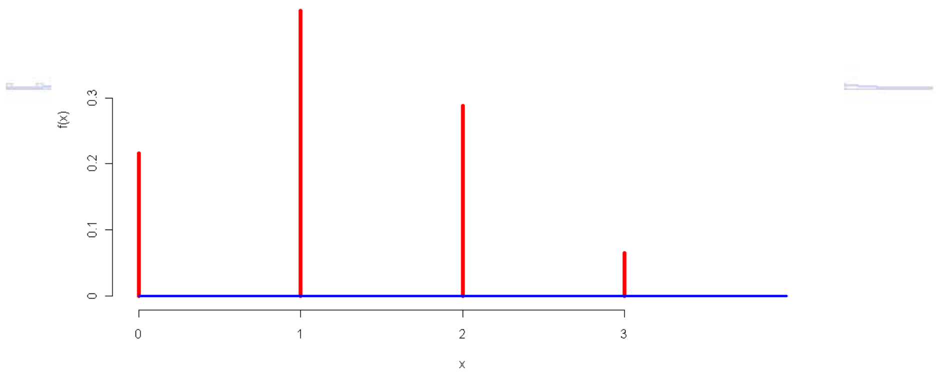
$$f(x) = P(X = x) = C_x^3 p^x (1-p)^{3-x}, x = 0,1,2,3$$

上式中 C_x^3 代表 3 次白努利試驗中 x 次成功的可能的組合數，而 $p^x(1-p)^{3-x}$ 代表每一個 x 次成功的機率（x 次成功必定有 $n-x$ 次失敗）。

結果 X	(X_1, X_2, X_3) 的組合	組合數	機率
0	(0, 0, 0)	$C_0^3 = 1$	$(1-p)^3$
1	(0, 0, 1), (0, 1, 0), (0, 0, 1)	$C_1^3 = 3$	$3p(1-p)^2$
2	(1, 1, 0), (1, 0, 1), (0, 1, 1)	$C_2^3 = 3$	$3p^2(1-p)$
3	(1, 1, 1)	$C_3^3 = 1$	p^3

當 $p = 0.4$ 時，

x	0	1	2	3
$f(x)$	0.216	0.432	0.288	0.064



例 7- 4

品管人員隨機檢驗 3 個產品，p 代表該製程產品不良的機率，X_i =代表第 i 次檢驗的結果，其中 i=1,2,3。記錄為

$$X_i = \begin{cases} 1, & Bad \\ 0, & Good \end{cases}$$

令 $X = X_1 + X_2 + X_3$ 代表 3 個受檢產品的總不良品個數。所以，可能的結果和發生

的機率為								
實驗結果	GGG	GGB	GBG	BGG	GBB	BGB	BBG	BBB
機率值	$(1-p)^3$	$p(1-p)^2$	$p(1-p)^2$	$p(1-p)^2$	$p^2(1-p)$	$p^2(1-p)$	$p^2(1-p)$	p^3

二項隨機變數 X ，表示成 $X \sim bin(n,p)$ 其的機率分配函數為

$$f(x)=P(X=x)=C_x^n p^x(1-p)^{n-x},$$
$$x=0,1,2,\ldots,n$$

例 7- 5

假設 $X \sim bin(2,0.4)$ ，則 X 的機率分配函數為 $f(x)=C_x^2 0.4^x 0.6^{2-x}$, $x=0,1,2,$	
x	$f(x)$
0	$C_0^2 0.4^0 0.6^2=0.6^2=0.36$
1	$C_1^2 0.4^1 0.6^1=2\times 0.24=0.48$
2	$C_2^2 0.4^2 0.6^0=0.16$

例 7- 6

假設產品的不良率為 0.1, 今檢驗 10 的產品，請問

(1) 恰好有 3 個不良品的機率,

(2) 不良品的個數不超過 5 的機率,

(3) 不良品個數介於 2 到 6 的機率。

解：

令 X 代表此檢驗結果的不良品個數，則 $X \sim bin(10,0.1)$ ，機率分配函數為

$$f(x)=C_x^{10}0.1^x 0.9^{10-x}, \quad x=0,1,2,\ldots,10$$

(1) $P(X=3)=C_3^{10}0.1^3 0.9^7=P(X\le 3)-P(X\le 2)$
$=0.9872-0.9298=0.0574$

(2) $P(X\le 5)=\sum_{x=0}^{5}C_x^{10}0.1^x 0.9^{10-x}=0.9999$

(3) $P(2\le X\le 6)=P(X\le 6)-P(X\le 1)=1.0\text{-}0.7361\text{=}0.2639$

例 7- 7

假設一個袋中有 5 顆紅球(red)和 15 顆白球(white)。每一次從袋中抽出一球後放回，令 X 代表 4 次抽樣紅球出現的總次數，因為每次抽中紅球的機率為 $p=1/4$，所以

$$X \sim bin(4,1/4)$$

試問：

(1) 紅球出現次數恰好 1 次的機率？

(2) 紅球出現次數最多 2 次的機率？

解：

(1)

$$P(X=1)=f(1)=C_1^4(\frac{1}{4})^1(\frac{3}{4})^3=4\times\frac{1}{4}\times\frac{27}{64}=\frac{27}{64}$$

(2)

$$P(0\le X\le 2)=f(0)+f(1)+f(2)=C_0^4(\frac{1}{4})^0(\frac{3}{4})^4+C_1^4(\frac{1}{4})^1(\frac{3}{4})^3+C_2^4(\frac{1}{4})^2(\frac{3}{4})^2$$
$$=\frac{3^4+4\times 3^3+6\times 3^2}{4^4}=\frac{81+108+54}{256}$$
$$=\frac{243}{256}=0.9492$$

二項分配的期望值與變異數

因為　$X = X_1 + \cdots + X_n$，所以

$$E(X) = E(X_1 + ... + X_n) = p + p + ... + p = np$$
$$Var(X) = Var(X_1 + ... + X_n) = p(1-p) + \cdots + p(1-p) = np(1-p)$$

例 7- 8

假設 $X \sim bin(4, 0.3)$，請問 X 的期望值與變異數。

解：

$$E(X) = 4 \times 0.3 = 1.2,$$
$$Var(X) = 4 \times 0.3 \times 0.7 = 0.84$$

練習：

(a) 假設 $X \sim bin(5, 0.2)$，請問 X 的期望值與變異數。

解：

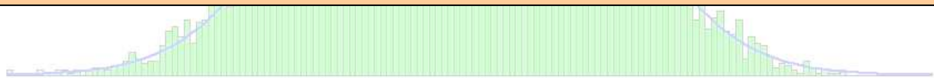
$$E(X) = 5 \times 0.2 = 1, Var(X) = 5 \times 0.2 \times 0.8 = 0.8$$

(b) 已知 $X \sim bin(n, p)$ 且 $E(X) = 0.4,\quad Var(X) = 1.92$，請問 n, p 分別為何？

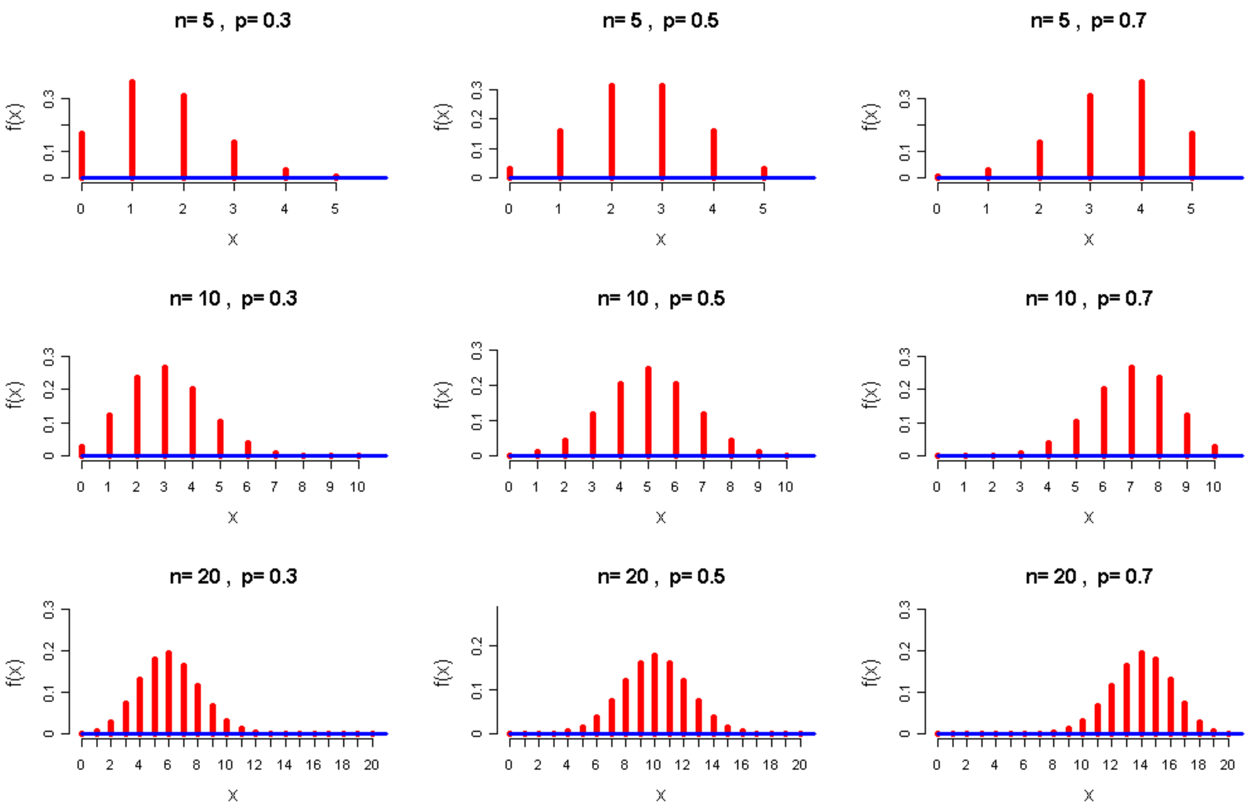
解：

$$\begin{cases} n \times p = 4.8 \\ n \times p \times (1-p) = 1.92 \end{cases} \Rightarrow \frac{Var(X)}{E(X)} = 0.4 = 1 - p$$

所以，$p = 0.6$，$n = 8$



成功機率 p 與機率分配圖形的關係

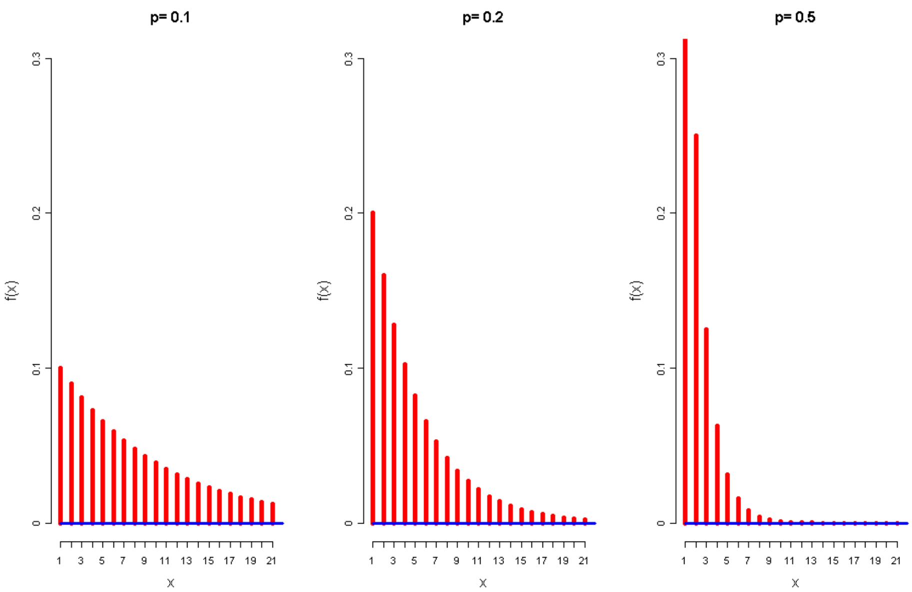


7-3 幾何分配

重複執行相同白努利試驗直到成功的事件發生才停止試驗，此試驗稱為幾何隨機實驗。例如擲一個銅板直到出現正面為止。我們令隨機變數 X 代表實驗的次數，稱為幾何隨機變數，其機率質量函數為

$$f(x)=q^{x-1}p, x=1,2,\ldots$$

其中 p,q 為白努力試驗成功與失敗的機率，滿足 $p+q=1$。



期望值與變異數

假設 X 幾何隨機變數，成功的機率為 p ，則

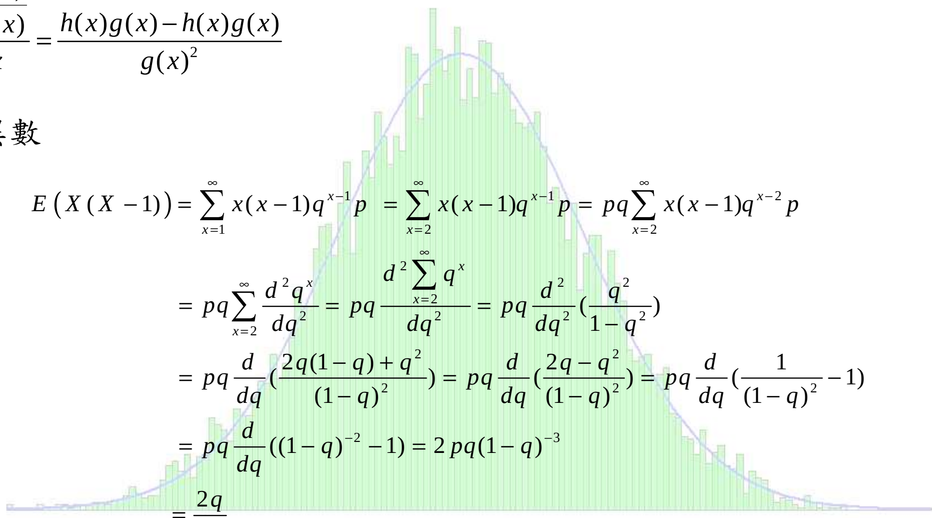
$E(X)=\frac{1}{p}$ ，$Var(X)=\frac{1-p}{p^2}$

證明：

(1) 期望值

$$E(X)=\sum_{x=1}^{\infty}xf(x)=\sum_{x=1}^{\infty}xq^{x-1}p=p\sum_{x=1}^{\infty}xq^{x-1}=p\sum_{x=1}^{\infty}\frac{dq^x}{dq}$$
$$=p\frac{d}{dq}(\frac{q}{1-q})=p\frac{(1-q)+q}{(1-q)^2}=\frac{p}{p^2}=\frac{1}{p}$$

註：$\frac{d\frac{h(x)}{g(x)}}{dx}=\frac{h(x)g(x)-h(x)g(x)}{g(x)^2}$



(2) 變異數

$$E\left(X(X-1)\right)=\sum_{x=1}^{\infty}x(x-1)q^{x-1}p=\sum_{x=2}^{\infty}x(x-1)q^{x-1}p=pq\sum_{x=2}^{\infty}x(x-1)q^{x-2}p$$
$$=pq\sum_{x=2}^{\infty}\frac{d^2q^x}{dq^2}=pq\frac{d^2\sum_{x=2}^{\infty}q^x}{dq^2}=pq\frac{d^2}{dq^2}(\frac{q^2}{1-q^2})$$
$$=pq\frac{d}{dq}(\frac{2q(1-q)+q^2}{(1-q)^2})=pq\frac{d}{dq}(\frac{2q-q^2}{(1-q)^2})=pq\frac{d}{dq}(\frac{1}{(1-q)^2}-1)$$
$$=pq\frac{d}{dq}((1-q)^{-2}-1)=2pq(1-q)^{-3}$$
$$=\frac{2q}{p^2}$$

因為

$$E\left(X(X-1)\right)=E(X^2-X)=EX^2-EX$$

所以， $EX^2=E\left(X(X-1)\right)+E(X)=\frac{2g}{p^2}+\frac{1}{p}=\frac{2g+p}{p^2}=\frac{2-p}{p^2}$

因此，幾何隨機變數的變異數為

$$Var(X)=EX^2-(EX)^2=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}$$

例 7- 9

假設某張證照通過率為 $p=0.1$ ，某人考第 11 次取得證照的機率為何？

FFFFFFFFFFS

前面 10 次失敗(Fail)，最後一次成功(Success)

$f(11)=0.9^{10}\times 0.1=0.03487$ $f(11)=0.9^{10}\times 0.1=0.03487$

例 7- 10

接續上題，請問取得該證照的期望考照次數與變異數。

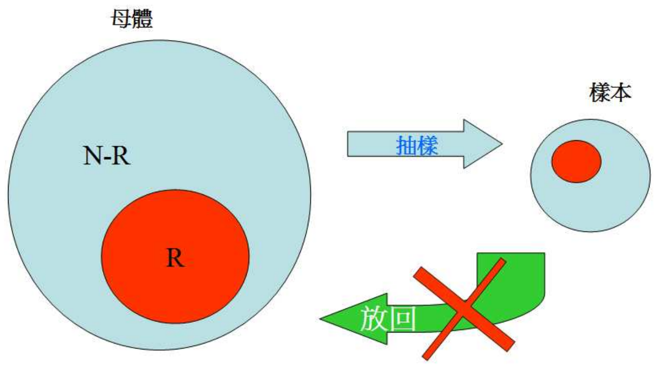
解：

$$E(X)=\frac{1}{p}=\frac{1}{0.1}=10$$

$$Var(X)=\frac{q}{p^2}=\frac{0.9}{0.1^2}=90$$

7-4 超幾何分配

假設一個個數為 N 的有限母體，母體中的成員被分類成兩種，第一類為成功有 $N-R$ 個，第二類為失敗有 R 個。今從母體以不放回的抽樣方式抽出 n 個成員，記錄各類別出現的次數，此稱為超幾何隨機實驗(hypergeometric distribution)。令 X 代表成功類別被抽中的個數，我們稱 X 為超幾何隨機變數。其機率質量分配函數稱為超幾何機率分配。



例 7- 11

假設一個袋中 $N=10$ 球，編號 $1,2,\cdots,10$，今隨機抽出兩個，請問有幾種抽法。

解：

根據組合數公式，我們可得

$$C_2^{10}=\frac{10\times 9}{2\times 1}=45$$

共有 45 種可能結果。

例 7- 12

接續上題，假設 10 球當中有 8 紅球，2 個白球。請問恰有 1 個紅球的組合數？

解：

假設前 8 號為紅色，其餘為白色

①②③④⑤⑥⑦⑧⑨⑩

恰有 1 個紅球的組合有

①⑨ ②⑨ ③⑨ ④⑨ ⑤⑨ ⑥⑨ ⑦⑨ ⑧⑨

①⑩ ②⑩ ③⑩ ④⑩ ⑤⑩ ⑥⑩ ⑦⑩ ⑧⑩

組合數=$C_1^8C_1^2=8\times 2=16$

例 7- 13

接續上題，請問恰有 1 個紅球的機率？

解：

令 R_1 代表恰有 1 個紅球事件，根據古典機率測度，假設每一種組合出現的機會相同，所以

$$P(R_1)=\frac{C_1^8C_1^2}{C_2^{10}}=\frac{16}{45}$$

假設 E 代表超幾何隨機實驗的結果有 x 成功，另外有 $n-x$ 失敗。根據組合公式，我們可以知道事件 E 的組合數有 $C_x^RC_{n-x}^{N-R}$，所以，事件 E 發生的機率，也就是隨機變數 $X=x$ 的機率為

$$f(x)=P(X=x)=\frac{\#(E)}{\#(S)}=\frac{C_x^RC_{n-x}^{N-R}}{C_n^N}$$

其中 $x=\max(0,n-(N-R)),\cdots,\min(n,R)$ 。上式為超幾何隨機變數的機率質量函數，若一個隨機變數 X 具有超幾何機率分配，記為

$$X \sim H(N,R,n)$$

隨機變數 X 成功的次數不可能超過抽樣個數 n ，也不可能超過母體中成功的個數 R ，所以 x 的上界為 $\min(n,R)$ 。隨機變數 X 成功的次數不可能低於 0，也不低於 $n-(N-R)$ ，其意為當抽完所有失敗者，剩下被抽中的都屬成功者。

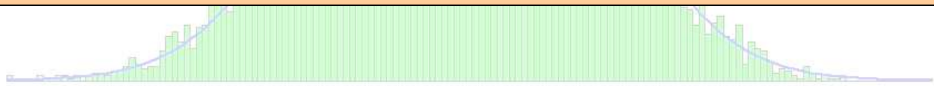
例 7- 14

假設 10 球當中有 8 紅球，2 個白球，今抽三球不放回。令 X 代表白球的個數，Y 代表紅球的個數，請問 X ，Y 的範圍分別為何？

解：

(1) 對隨機變數 X 而言，母體個數 $N=10$，成功的個數(白) $R=2$，失敗的個數(紅) $N-R=8$ 。雖抽三球，但最多只有 2 個白球， $\min(n,R)=2$ 。

(2) 對隨機變數 Y 而言，母體個數 $N=10$，成功的個數(紅) $R=8$，失敗的個數(白) $N-R=2$ 。因為只有 2 個白球，所以最少抽中一個紅球， $\max(0,n-(N-R))=\max(0,3-2)=1$ 。



例 7- 15

某公司進貨一批 50 個，假設其中有 5 個不良品。今作進貨檢驗，抽驗 5 個貨品，令 X 代表不良品的個數。請問：

(1) 隨機變數 X 的機率分配？

(2) 請問沒有檢驗出不良品的機率為何？

解：

(1) 隨機變數 $X \sim H(50,5,5)$ 的機率質量函數為

$$f(x)=\frac{C_x^5 C_{5-x}^{45}}{C_5^{50}},\ x=0,\cdots,5$$

(2) 請問沒有不良品的機率？

$$f(0)=P(X=0)=\frac{C_0^{50}C_5^{45}}{C_5^{50}}=0.577$$

例 7- 16

假設本系新生有 100 人，其中有 30 個女生。今系學會隨機抽出 10 人參加愛校 DVD 拍攝，令 X 代表女生被抽中的人數。請問：

(1) 隨機變數 X 的機率分配？

(2) 請問恰抽中 2 個女生的機率？

解：

(1) 隨機變數 $X\sim H(100,30,10)$ 的機率質量函數為

$$f(x)=\frac{C_x^{30}C_{10-x}^{70}}{C_{10}^{100}},\ x=0,\cdots,5$$

(2)

$$f(2)=\frac{C_2^{30}C_8^{70}}{C_{10}^{100}}$$

超幾何隨機變數的期望值與變異數

假設隨機變數 $X\sim H(N,R,n)$，成功的比例=$\frac{R}{N}$，其期望值與變異數為

$E(X)=n\cdot\frac{R}{N}$ ($=n\times$成功的比例)

$Var(X)=n\cdot\frac{R}{N}\cdot\frac{N-R}{N}\cdot\frac{N-n}{N-1}$ ($=n\times$成功的比例$\times$失敗的比例$\times$有限母體的修正因子)

其中係數 $\frac{N-n}{N-1}$ 稱為有限母體的修正因子，當 $N>>n$ 時， $\frac{N-n}{N-1}\approx 1$，超幾何隨機變數的期望值與變異數近似二項分配隨機變數。

註：二項隨機變數的期望值與變異數

$$E(X)=np, Var(X)=np(1-p)$$

超幾何分配與二項分配的關係

在一個有限母體中抽樣，若每次抽出一個後放回再抽，則母體總數會固定不

變，每一次抽樣可以是為相同的白努利試驗，所以，n 抽樣結果的和為一個二項分配。但是若是不放回，則母體的個數會逐次減少，換言之，每次抽樣的條件都在改變，雖然每次抽樣都依然是白努利試驗，但成功的機率並非不變的，而且與已經發生的抽樣結果有關，這 n 抽樣結果的和是一個超幾何分配。例如，我們從一副牌中隨機抽出一張，第一次抽中紅心的機率為 1/4，若把抽出的排再放回去，經過適當洗牌後再抽一張，抽中紅心的機率仍為 1/4。但是，若我們不將排放回去，則第二次抽中紅心的機率就與第一次不同，而且與第一次的結果有關。令 A_1 代表第一次抽中紅心的事件，A_2 代表第二次抽中紅心的事件。假設第一次抽中紅心，則第二次在抽中紅心的機率為

$$P(A_2 \mid A_1) = \frac{12}{51} = \frac{4}{17} < 1/4$$

但是若第一次沒有抽中紅心，即 A_1 的餘事件 A_1^c 發生，剩下 51 張牌中還是有 13 紅心，所以，

$$P(A_2 \mid A_1^c) = \frac{13}{51} > 1/4。$$

我們注意到超幾何隨機變數的變異數為 $n\frac{R}{N}\frac{N-R}{N}\frac{N-n}{N-1}$ 小於 $np(1-p)$（$p=\frac{R}{N}$）。也就是說，抽放回的隨機試驗的變數小於抽不放回的試驗。真實上，不管是民意調查或是工廠的產品檢驗都屬抽不放回的檢驗，我們不可能對相同的個體做一次以上訪問或是檢驗。當 $n=N$ 時，也就是普查而非抽樣，因此沒有抽樣的不確定性，所以超幾何隨機變數的變異數為零，它的意義是再次抽樣的結果會是相同的。但是，就抽有放回的二項隨機變數，它的變異數卻是會隨著試驗的次數增加而越來越大。

但是，當抽樣的個數遠小於母體總數 $N \gg n$，有無放回的影響性就很小了，就像在大海中舀水，有無倒回對大海的影響幾乎可略去不計。一般民意調查也是如此，為數僅千人的受訪者相對於上千萬的母體可說微乎其微。我們再次注意到超幾何隨機變數的變異數為 $n\frac{R}{N}\frac{N-R}{N}\frac{N-n}{N-1}$

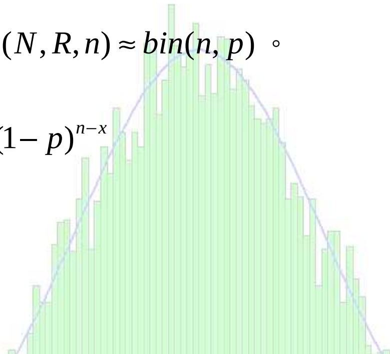
當 $N \gg n$(通常指 $\frac{n}{N} \le 0.1$)時，末項為有限母體的修正因子 $\frac{N-n}{N-1}$ 就很接近 1，換言之，超幾何分配的變異數就接近二項分配的變異數。當在此種條件下，兩個分配的集中趨勢(期望值)相同，且分散趨勢(變異數)相近，機率分配圖形上兩者差異很小，因此，我們可以利用二項分配計算超幾何分配的機率值。為何要用近似值呢？主要是計算上的效益，計算超幾何分配的機率值用很大的數字階乘，不僅計算費時而且數字級數過大，容易超出計算機程式的負荷，造成很大的計算誤差。

使用二項分配的近似流程：

(1) 判斷 $\frac{n}{N} \le 0.1$ 是否成立？當這個條件成立後方可使用近似解。

(2) 令 $p = \frac{R}{N}$ ，假設 $X \sim H(N, R, n) \approx bin(n, p)$ 。

(3) $f(x) = \frac{C_x^R C_{n-x}^{N-R}}{C_n^N} \approx C_x^n p^x (1-p)^{n-x}$



例 7- 17

假設 $X \sim H(20,10,5)$ ，請計算 X 的期望值與變異數。

解：

成功的比例 $= \frac{R}{N} = \frac{10}{20} = 0.5$

$E(X) = 5 \times 0.5 = 2.5$

$Var(X) = 5 \times 0.5 \times 0.5 \times \frac{20-5}{20-1} = 1.25 \times \frac{15}{19} = \frac{18.75}{19}$

例 7- 18

接續例 7- 16， $X \sim H(100,30,10)$ ，請利用二項分配近似求得恰有 2 個女生的機率值。

解：

因為 $N=100$，$n=10$，所以成功的機率為 $p=\dfrac{n}{N}=0.1$。又 $n/N=0.1$，隨機變數 X 近似 $bin(10,0.3)$，因此，

$$f(2)=\frac{C_2^{30}C_8^{70}}{C_{10}^{100}}\approx C_2^{10}0.3^2 0.7^8=0.3823-0.1493=0.2330$$

7-5 卜瓦松分配

我們在一個固定時間間隔，或固定範圍內，觀察某一特定事件發生的次數，會發生幾次是一個隨機變數。例如：

（1）一片 12 吋晶圓片上的雜點數目。

（2）桃園縣一天中發生的車禍次數。

（3）一片汽車烤漆的氣泡數目。

（4）中午 12:00 到 13:00 到便利商店買餐的人數。

這些隨機行為變數具有以下介紹一些性質，稱為卜瓦松隨機變數：

(a) 某事件發生一次的機率與時間的長度或是區域大小成正比。

(b) 在很極短的時間或極小的區域內，某事件發生兩次的機率幾乎為零。也就是說，任何兩事件發生，就時間而言可以區分前後，就發生地點而言可以分別彼此。

(c) 任兩個不重疊的時間或區域，某事件發生的次數彼此間相互獨立。

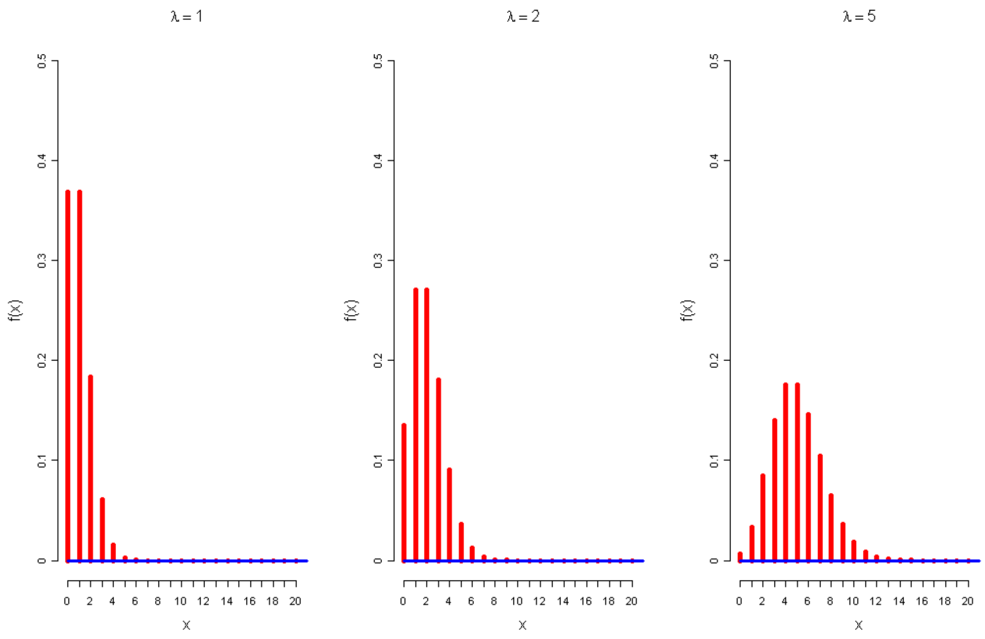
卜瓦松分配

假設隨機變數 X 的機率質量函數為

$$f(x)=\frac{e^{-\lambda}\lambda^x}{x!}\quad , \quad x=0,1,2\ldots$$

我們稱 X 為具有參數 λ 的卜瓦松隨機變數，表示成 $X\sim P(\lambda)$。

期望值 λ 與機率分配的關係



卜瓦松分配的期望值與變異數

假設 X 為具有參數 λ 的卜瓦松隨機變數，則其期望值與變異數為

$$E(X)=\lambda$$
$$Var(X)=\lambda$$
。

註：卜瓦松隨機變數的期望值與變異數相同。

例 7- 19

假設通過中山高速公路中壢收費站的汽車數為一個卜瓦松隨機變數，已知每 5 秒鐘 2 部車通過，請問：

(1)5 秒內恰有一車通過的機率？

(2)10 秒內恰有一車通過的機率？

(3)10 秒內恰沒有車通過的機率？

(4) 一分鐘平均有幾輛車通過？

解：

(1) 因為 λ=2 ，所以

$$f(1)=\frac{e^{-2}2^{1}}{1!}=2e^{-2}=0.271$$
$$=P(X\le 1)-P(X\le 0)=0.406-0.135=0.271$$

(2)每 10 秒平均有 4 輛車通過，即 λ=4 ，所以

$$f(1)=\frac{e^{-4}4^{1}}{1!}=4e^{-4}=0.092-0.018=0.074$$

(3) $f(0)=\frac{e^{-4}4^{0}}{0!}=e^{-4}=0.018$

(4) 因為每 5 秒鐘 2 部車通過，可以推論每分鐘有 24 輛車通過。

例 7- 20

假設麥當勞中午用餐時間 11:00-14:00 平均到客量為 180 人，請問：

(1) 中午用餐時間到客人數的期望值和變異數。

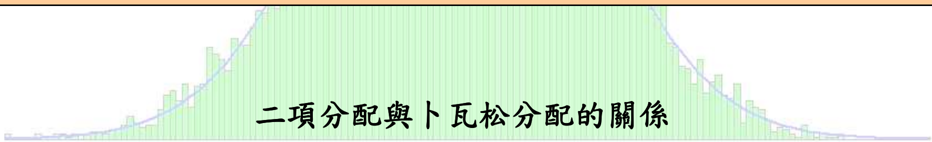
(2) 一分鐘內有兩個顧客的機率為何？

解：

(1) $E(X)=180$ ，$Var(X)=180$ 。

(2) 三小時平均 180 人，所以每分鐘的平均值為 1 人，即 λ=1。有兩個顧客的機率為

$$f(2)=\frac{e^{-1}1^{2}}{2!}=0.184 \text{ 。}$$



二項分配與卜瓦松分配的關係

在二項隨機試驗中，當 n 很大而 p 很小時，我們可以用卜瓦松分配求得二項分配的近似機率值(Simeon D. Poisson(1837), 1781~1840)。我們想像這樣一個白努利隨機試驗，每此投擲銅板的時間間隔是相同的 h ，所以，當 n 很大時，這個投擲實驗像是在一個綿密的持續時間為 nh 的區間內，做觀察某一個特定事件(成功)出現的次數。二項隨機實驗是這 n 白努利試驗的和，具有(1) 每一次試驗彼此相互獨立，又(2)成功的次數與持續時間 nh 成正比，這兩點性質滿足前面提到卜瓦松分配三個性質的(a)與(c)。當 p 很小時，這意謂接連兩次正面的機率非常非常地小，也就滿足(b)的性質。以上說明，當 n 很大而 p 很小時，二項隨機實驗的觀察機制非常類似卜瓦松隨機實驗，這也就是為什麼 Poisson(1837)得出可以用卜瓦松分配求得二項分配的近似機率值。

在這 nh 的持續時間內，特定事件(成功)出現的次數的期望值為 np ，所以，取 λ=np 近似之。

例 7- 21

假設產品的不良率為 $p=0.01$，今檢驗 $n=100$ 個，試問：

(1) 不良品個數不超過 2 的機率?

(2) 請問恰有 3 個不良品的機率?

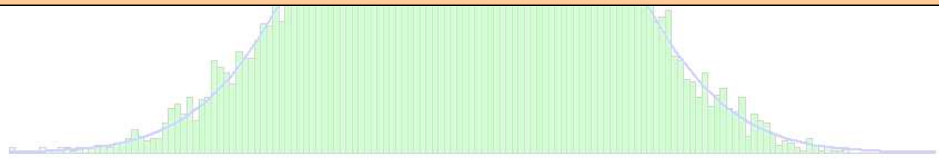
解：令隨機變數 X 代表不良品個數，因為 $p=0.01$很小，$n=100$ 很大，我們可以利用卜瓦松分配求得二項分配的近似機率值

$$\lambda = np = 100 \times 0.01 = 1$$

（1）

$$\begin{aligned}P(X \le 2) &= f(0) + f(1) + f(2) \\ &= \frac{e^{-1}1^0}{0!} + \frac{e^{-1}1^1}{1!} + \frac{e^{-1}1^2}{2!} = 2.5e^{-1} \\ &=0.9197\end{aligned}$$

（2） $f(3) = \dfrac{e^{-1}1^3}{3!} = \dfrac{1}{6}e^{-1} = 0.0613$



習　題

1. 你被要求作一抽樣計劃從供應商所提供的 300 單位的量中，抽取 40 件，此供應商過去品質約有 2%的不良品，利用卜瓦松分配近似發現 2 件或超過 2 件不良品的機率。
2. 從一連續製程所產出的產品中，隨機抽樣 100 項，此製程平均約 0.7%的不良項，則可發現無不良項存在的近似機率為何？
3. 從 1000 坪的地毯，隨機抽樣 3 坪的檢驗計劃，若在 3 坪的地毯沒有發現任何缺點，則接受此地毯，今一捲的地毯中平均每坪，具有一個缺點。在此計畫中被拒絕的機率為何？
4. 意外事件發生的機率常用何種分配？
5. 自家門前約 5 分鐘經過 1 輛車，求 10 分鐘內無任何一輛車經過之機率？
6. 某機械之故障率為 0.004，運轉該機器 50 次，其故障至少 2 次之機率？
7. 設某重機器之故障率為 0.045，茲運轉該機器 20 次，問其間發生故障次數超過 1 次之機率為多少？
8. 一個製程，生產物料有 30%不良品，隨時選出 5 件作檢驗，則在樣本中剛好發現 2 件良品的機率為何？

9. 令 X 表投擲一個銅板出現正面的次數；Y 表投擲 n 個銅板出現正面的次數，$n>1$。請問：

    (a)隨機變數 X 具有何種分配？

    (b)　隨機變數 Y 具有何種分配？

10. 隨機變數 $X \sim bin(5,\ 0.2)$，請問：

    (a)　$P(X=2)=$？______________

    (b)　$P(X<2)=$？______________

    (c)　$P(X>2)=$？______________

11. 假設大華公司有一批產品 100 件欲交貨給顧客，而其交貨合約書上規定，若產品中有超過 1 件不良品時，需退貨。今已知此批產品之不良率為 0.02，請

問此批產品退貨的機率為何？

12. 假設本班有男生 20 人，女生 40 人。今隨機抽取 5 位同學作調查（不得重複調查），請問，

(a)全部為男生的機率為何？

(b) 全部為女生的機率為何？

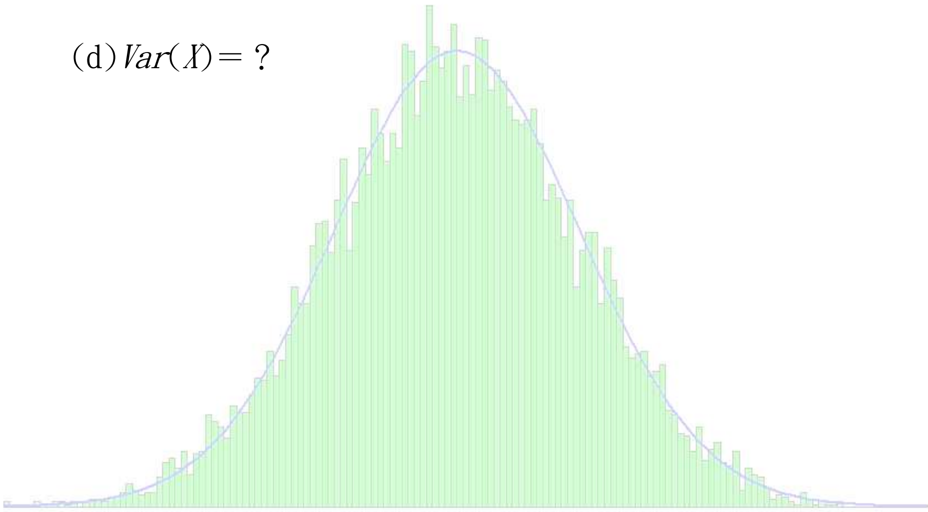
13. 假設一離散型隨機變數 X 具有參數 $N=100$、$R=30$、$n=3$ 的超幾何分配。請問，

(a) $P(X=1)=$？

(b) $P(X\leq 2)=$？

(c) $E(X)=$？

(d) $Var(X)=$？



重點摘錄