

單元 18: 負二項機率分布

(課本 §3.6)

獨立地重複相同的試驗. 若 $P(\text{成功}) = p$, 且令

$Y =$ 出現第 r 次成功時, 所需的試驗次數

則 Y 的可能值為 $r, r+1, r+2, \dots$ 且對於 $y = r, r+1, r+2, \dots$,

$$\begin{aligned} P(Y = y) &= P(\text{前面 } (y-1) \text{ 次試驗中有 } (r-1) \text{ 次成功} \\ &\quad \text{且第 } y \text{ 次成功}) \\ &= \binom{y-1}{r-1} p^{r-1} q^{y-r} \cdot p \quad (\text{由獨立性與乘法交換律}) \\ &= \binom{y-1}{r-1} p^r q^{y-r}, \quad q = 1 - p \end{aligned}$$

並稱隨機變數 Y 有負二項機率分布 (negative binomial probability distribution). 正式的定義如下述:

定義 1. $Y \sim \text{negative } b(p, r)$ 若且為若

$$\begin{aligned} p(y) &= P(Y = y) \\ &= \binom{y-1}{r-1} p^r q^{y-r} \end{aligned}$$

其中 $y = r, r + 1, r + 2, \dots, 0 \leq p \leq 1$.

定理 3.9. 設 $Y \sim \text{negative } b(p, r)$. 則

$$\mu = E(Y) = \frac{r}{p}$$

且

$$\sigma^2 = \text{Var}(Y) = \frac{r(1-p)}{p^2}$$

<證> (略) 第五章會有一簡易的證法.

註. $\text{negative } b(p, r)$ 的期望值與變異數均為 $\text{geometric}(p)$ 的 r 倍 (與直觀相符, 因為出現第 r 次成功時, 所需的試驗次數應與出現第 1 次成功時, 所需的試驗次數的 r 倍有關聯).

例 1. 用過的大量泵 (浦) (pump, 抽水機, 唧筒) 中, 有 20% 需要修理. 維護工帶著三套材料, 以不放回的方式隨機抽取泵浦檢查, 並以一套材料修理不合格的. 若需時 10 分鐘檢查好的泵浦且需 30 分鐘檢查並修理不合格的, 試求用完三套材料所需時間的期望值與標準差.

<解> 因為有大量的泵浦，故可假設

$$Y \stackrel{\text{def}}{=} \text{找到第三個不合格泵浦所需的檢查次數} \\ \sim \text{negative } b(0.2, 3)$$

因此，用完三套材料所需的時間

$$T = 10Y + 3(20) = 10Y + 60$$

其中 $10Y$ 為檢查時間， $3(20)$ 為修理 3 個不合格泵浦的時間。所以，所需時間的期望值

$$\begin{aligned} E(T) &= 10E(Y) + 60 \\ &= 10(3) \left(\frac{1}{0.2} \right) + 60 \\ &= 150 + 60 = 210 \end{aligned}$$

且變異數

$$\begin{aligned} \text{Var}(T) &= \text{Var}(10Y + 60) \\ &= 10^2 \text{Var}(Y) \\ &\quad \text{因為 } \text{Var}(aY + b) = a^2 \text{Var}(Y) \\ &= 10^2(3) \left(\frac{1 - 0.2}{(0.2)^2} \right) \\ &= 100(3)(20) = 6000 \end{aligned}$$

以及標準差

$$\sigma = \sqrt{6000} = 77.46$$