詳細解答

# STEP III 2011，第 12 題

概率與統計 • 以概率生成函數處理複合分佈

> **STEP III 2011，第 12 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/11-s3-q12

## 關鍵思路

關於一個分佈的平均值（mean）和方差（variance），我們所需要的一切都藏在它的概率母函數（probability generating function, pgf）在 $t = 1$ 處的首兩階導數之中。因此，我們用連鎖法則（chain rule）對複合函數 $G(H(t))$ 求導（兩次），代入 $t = 1$，並利用 $H(1) = 1$。至於擲硬幣的部分，$N$ 服從幾何分佈，而每個 $X_i$ 都是單次擲公平硬幣的指示變量（indicator），所以兩個概率母函數都很簡單，而 $G(H(t))$ 是一個有理函數（rational function），可以展開成幾何級數（geometric series）。

## 解答

### 一般結果

對於任何取非負整數值、概率母函數為 $P(t) = \mathrm{E}(t^Z)$ 的隨機變量 $Z$，我們有 $P(1) = 1$，

$$P'(1) = \mathrm{E}(Z), \qquad P''(1) = \mathrm{E}(Z(Z-1)),$$

因此

$$\mathrm{Var}(Z) = P''(1) + P'(1) - (P'(1))^2.$$

記 $\mu = \mathrm{E}(X_i) = H'(1)$、$\sigma^2 = \mathrm{Var}(X_i)$，並以 $K(t) = G(H(t))$ 表示 $Y$ 的概率母函數。由連鎖法則，$K'(t) = G'(H(t))\,H'(t)$。由於 $H(1) = 1$，

$$\mathrm{E}(Y) = K'(1) = G'(1)H'(1) = \mathrm{E}(N)\,\mathrm{E}(X_i),$$

這就是第一個結果。

再求導一次（先用乘積法則（product rule），再用連鎖法則），

$$K''(t) = G''(H(t))(H'(t))^2 + G'(H(t))\,H''(t),$$

所以，再次利用 $H(1) = 1$，

$$K''(1) = G''(1)\mu^2 + \mathrm{E}(N)\,H''(1).$$

把上面的方差公式分別應用於 $N$ 和 $X_i$，得

$$G''(1) = \mathrm{Var}(N) - \mathrm{E}(N) + (\mathrm{E}(N))^2, \qquad H''(1) = \sigma^2 - \mu + \mu^2.$$

? **想一想** 為甚麼不直接由 $K''(1)$ 計算 $\mathrm{Var}(Y)$，而要先改寫 $G''(1)$ 和 $H''(1)$？

因為我們想要的答案是以方差表示，而不是以階乘矩（factorial moments）表示；透過 $\mathrm{Var}(N)$ 和 $\sigma^2$ 來表達 $G''(1)$ 和 $H''(1)$，可以讓多餘的項互相抵消。記 $m = \mathrm{E}(N)$，

$$\begin{aligned}\mathrm{Var}(Y) &= K''(1) + K'(1) - (K'(1))^2 \\ &= (\mathrm{Var}(N) - m + m^2)\mu^2 + m(\sigma^2 - \mu + \mu^2) + m\mu - m^2\mu^2 \\ &= \mathrm{Var}(N)\mu^2 + m\sigma^2 + (-m\mu^2 + m\mu^2) + (m^2\mu^2 - m^2\mu^2) + (-m\mu + m\mu) \\ &= \mathrm{Var}(N)(\mathrm{E}(X_i))^2 + \mathrm{E}(N)\,\mathrm{Var}(X_i),\end{aligned}$$

即為所求。

**擲硬幣**

*$N$ 的概率母函數。*第一次擲出正面發生在第 $n$ 次，當且僅當首 $n-1$ 次都是反面而第 $n$ 次是正面，所以對 $n = 1,2,3,\ldots$，$\mathrm{P}(N = n) = (\frac{1}{2})^n$（這是幾何分佈（geometric distribution））。因此，當 $|t| < 2$ 時，

$$G(t) = \sum_{n=1}^{\infty} \left(\frac{t}{2}\right)^n = \frac{t/2}{1 - t/2} = \frac{t}{2-t}.$$

*每個 $X_i$ 的概率母函數。*若其後的第 $i$ 次擲出正面，令 $X_i = 1$，否則令 $X_i = 0$。於是 $Y = X_1 + \cdots + X_N$，而 $H(t) = \frac{1}{2} + \frac{1}{2}t = \frac{1}{2}(1+t)$。其後各次擲硬幣彼此獨立，亦與決定 $N$ 的各次擲硬幣獨立，所以一般結果的條件成立。

*$Y$ 的概率母函數。*

$$G(H(t)) = \frac{\frac{1}{2}(1+t)}{2 - \frac{1}{2}(1+t)} = \frac{1+t}{3-t}.$$

*平均值與方差。*對於 $p = \frac{1}{2}$ 的幾何分佈，$\mathrm{E}(N) = 1/p = 2$，$\mathrm{Var}(N) = (1-p)/p^2 = 2$（或者對 $G$ 求導：$G'(t) = 2/(2-t)^2$、$G''(t) = 4/(2-t)^3$，得 $G'(1) = 2$、$G''(1) = 4$，

以及 $\mathrm{Var}(N) = 4 + 2 - 4 = 2$）。每個 $X_i$ 都服從伯努利（Bernoulli）分佈 $(\frac{1}{2})$，所以 $\mathrm{E}(X_i) = \frac{1}{2}$，$\mathrm{Var}(X_i) = \frac{1}{4}$。因此

$$\mathrm{E}(Y) = 2 \cdot \tfrac{1}{2} = 1, \qquad \mathrm{Var}(Y) = 2 \cdot \tfrac{1}{4} + 2 \cdot \tfrac{1}{4} = 1.$$

$Y$ 的分佈。當 $|t| < 3$ 時，我們展開

$$\frac{1+t}{3-t} = \frac{1+t}{3} \cdot \frac{1}{1 - t/3} = \frac{1+t}{3} \sum_{k=0}^{\infty} \frac{t^k}{3^k} = \sum_{k=0}^{\infty} \frac{t^k}{3^{k+1}} + \sum_{k=0}^{\infty} \frac{t^{k+1}}{3^{k+1}}.$$

$t^0$ 的係數只來自第一個和。當 $r \geq 1$ 時，$t^r$ 的係數是 $\dfrac{1}{3^{r+1}} + \dfrac{1}{3^r} = \dfrac{4}{3^{r+1}}$。所以

$$\mathrm{P}(Y = 0) = \frac{1}{3}, \qquad \mathrm{P}(Y = r) = \frac{4}{3^{r+1}} \quad (r = 1, 2, 3, \ldots).$$

## 驗算

- 各概率之和為 1：$\frac{1}{3} + 4\sum_{r\geq 1} 3^{-(r+1)} = \frac{1}{3} + 4 \cdot \frac{1}{6} = 1$。
- 直接計算 $\mathrm{P}(Y = 0)$：$\sum_{n\geq 1} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^n = \sum_{n\geq 1} 4^{-n} = \frac{1}{3}$。同理，$\mathrm{P}(Y = 1) = \sum_{n\geq 1} n\, 4^{-n} = \frac{4}{9}$，與 $4/3^2$ 吻合。
- 對 $n \leq 200$ 精確求和 $\mathrm{P}(N = n)\binom{n}{r}2^{-n}$，可重現 $r \leq 7$ 時的 $\mathrm{P}(Y = r)$，而所得平均值和方差均為 1.000；以 200 000 次模擬，得到 0.996 和 0.994。

## 常見錯誤

- 把 $N$ 當作從 0 開始（概率母函數為 $1/(2-t)$）：這裏 $N$ 計算的是擲出正面的那一次，所以 $N \geq 1$。
- 忘記 $P''(1)$ 是 $\mathrm{E}(Z^2) - \mathrm{E}(Z)$ 而不是 $\mathrm{E}(Z^2)$，以致在方差公式中漏掉 $+P'(1)$。
- 對所有 $r \geq 0$ 都給出 $4/3^{r+1}$：$r = 0$ 的一項是不同的。

## STEP 之外

$Y$ 是一個隨機和（random sum）（即複合分佈（compound distribution）），而上述方差公式是全方差定律（law of total variance）$\mathrm{Var}(Y) = \mathrm{E}\big(\mathrm{Var}(Y \mid N)\big) + \mathrm{Var}\big(\mathrm{E}(Y \mid N)\big)$ 的一個特例。概率母函數的複合也是分支過程（branching process）的核心：在分支過程中，$G(G(\cdots G(t)))$ 給出第 $n$ 代的規模，而滅絕的概率是 $G$ 的一個不動點（fixed point）。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率導論》）正是在這個程度上講解母函數、隨機和及分支過程。

Beyond STEP 詳細解答：https://beyondstep.keithmcnulty.org/zh-hant/。解答文字採用 CC0 1.0；題目本身版權屬 Cambridge University Press & Assessment 所有（©），本文件不予轉載。由機器從英文翻譯而成；尚未經審閱。