2019 年 4 月 17 日補足–2 項変数の分散

$m = E[X]$ として定義される X の分散

$$V[X] := E[(X-m)^2] = \sum_{k=0}^{n}(k-m)^2 P(X=k) = \sum_{k=0}^{n}(k-m)^2 {}_nC_k p^k q^{n-k}$$

を計算しましょう．

$$\begin{aligned}\sum_{k=0}^{n}(k-m)^2 P(X=k) &= \sum_{k=0}^{n}(k^2-2mk+m^2)^2 P(X=k) \\ &= \sum_{k=0}^{n} k^2 P(X=k) - 2m\sum_{k=0}^{n} kP(X=k) + m^2 \sum_{k=0}^{n} P(X=k) \\ &= \sum_{k=0}^{n} k^2 P(X=k) - \left(\sum_{k=0}^{n} kP(X=k)\right)^2\end{aligned}$$

から一般的に成立する公式

$$\boxed{V[X] = E[X^2] - (E[X])^2 = \sum_{k=0}^{n} k^2 P(X=k) - \left(\sum_{k=0}^{n} kP(X=k)\right)^2}$$

が導かれます．この最右辺辺の第 1 項を計算します．

$$\begin{aligned}\sum_{k=0}^{n} k^2 {}_nC_k p^k q^{n-k} &= \sum_{k=0}^{n}\left\{(k^2-k)+k\right\}{}_nC_k p^k q^{n-k} \\ &= \sum_{k=0}^{n} k(k-1){}_nC_k p^k q^{n-k} + \sum_{k=0}^{n} k{}_nC_k p^k q^{n-k} \\ &= \sum_{k=2}^{n} k(k-1){}_nC_k p^k q^{n-k} + np = \sum_{k=2}^{n} n(n-1){}_{n-2}C_{k-2} p^k q^{n-k} + np\end{aligned}$$

$t = k-2$ とする．

$$\begin{aligned}&= n(n-1)\sum_{t=0}^{n-2} {}_{n-2}C_t p^{t+2} q^{n-2-t} + np = n(n-1)p^2 \sum_{t=0}^{n-2} {}_{n-2}C_t p^t q^{n-2-t} + np \\ &= p^2 n(n-1)(p+q)^{n-2} + np = p^2 n(n-1) + np\end{aligned}$$

と計算されます．これから

$$\begin{aligned}V[X] &= p^2 n(n-1) + np - n^2p^2 \\ &= np\left(p(n-1)+1-np\right) = np(1-p) = npq\end{aligned}$$

から

$$\boxed{V[X] = npq}$$

が従います．途中で

$$k(k-1){}_nC_k = k(k-1){}_{n-2}C_{k-2}$$

を用いました．