データ分布と予測
確率変数・確率分布についてのノート

堀田敬介

Revised 2008 年 5 月, 2003 年 10 月

1　確率分布の期待値・分散

1.1　期待値・分散

$$
\begin{aligned}
E(X) &= \sum_x f(x), \\
V(X) &= E((x-\mu)^2) \\
&= \sum_x (x-\mu)^2 f(x) \quad (\mu := E(X)) \\
&= \sum_x x^2 f(x) - 2\mu \sum_x x f(x) + \mu^2 \sum_x f(x) \\
&= E(X^2) - E(X)^2
\end{aligned}
$$

また,

$$
\begin{aligned}
E(aX) &= \sum_x a x f(x) \\
&= a \sum_x x f(x) \ = \ aE(X), \\
V(aX) &= E(a^2X^2) - E(aX)^2 \\
&= a^2E(X^2) - a^2E(X)^2 \ = \ a^2V(X)
\end{aligned}
$$

1.2　2 つの確率変数について期待値・分散の和・積

2 つの確率変数 X, Y について，同時確率分布 $f(x,y)$ と周辺確率分布 $g(x), h(y)$ を考える.

$$
f(x,y) \ = \ P(X=x, Y=y), \quad f(x,y) \geq 0,\ \sum_x \sum_y f(x,y) = 1,
$$

$$
\begin{aligned}
g(x) &= P(X=x) = \sum_y f(x,y), \\
h(y) &= P(Y=y) = \sum_x f(x,y).
\end{aligned}
$$

$Y \backslash X$	x_1	x_2	$\cdots$	x_n	$h(y)$
y_1	$P(X=x_1, Y=y_1)$	$P(X=x_2, Y=y_1)$	$\cdots$	$P(X=x_n, Y=y_1)$	$P(Y=y_1)$
y_2	$P(X=x_1, Y=y_2)$	$P(X=x_2, Y=y_2)$	$\cdots$	$P(X=x_n, Y=y_2)$	$P(Y=y_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
y_m	$P(X=x_1, Y=y_m)$	$P(X=x_2, Y=y_m)$	$\cdots$	$P(X=x_n, Y=y_m)$	$P(Y=y_m)$
$g(x)$	$P(X=x_1)$	$P(X=x_2)$	$\cdots$	$P(X=x_n)$	1

このとき，

$$
\begin{aligned}
E(X+Y) &= \sum_x \sum_y (x+y) f(x,y) \\
&= \sum_x \{x \sum_y y f(x,y)\} + \sum_y \{y \sum_x x f(x,y)\} \\
&= \sum_x \{x g(x)\} + \sum_y \{y h(y)\} = E(X) + E(Y), \\
V(X+Y) &= E(\{(X+Y) - E(X+Y)\}^2) \\
&= E(\{(X - E(X)) + (Y - E(Y))\}^2) \\
&= E\{(X-E(X))^2 + 2(X-E(X))(Y-E(Y)) + (Y-E(Y))^2\} \\
&= E\{(X-E(X))^2\} + 2E\{(X-E(X))(Y-E(Y))\} + E\{(Y-E(Y)^2)\} \\
&= V(X) + 2Cov(X,Y) + V(Y), \\
Cov(X,Y) &= E\{(X-E(X))(Y-E(Y))\} \\
&= E(XY) - E(XE(Y)) - E(YE(X)) + E(E(X)E(Y)) \\
&= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\
&= E(XY) - E(X)E(Y).
\end{aligned}
$$

また，確率変数 X, Y が互いに**独立** *independent*[1] の時，即ち，同時確率分布において，

$$
\forall x, y \quad f(x,y) = g(x)h(y)
$$

が成り立つとき，

$$
\begin{aligned}
E(XY) &= \sum_x \sum_y xy \cdot f(x,y) \\
&= \sum_x \sum_y xy \cdot g(x)h(y) = \sum_x xg(x) \sum_y yh(y) = E(X)E(Y), \\
Cov(X,Y) &= 0, \\
V(X+Y) &= V(X) + V(Y).
\end{aligned}
$$

[1]相関係数 $\rho = 0$ ($Cov(X,Y) = 0$) のとき，**無相関** *uncorrelated* という．独立ならば無相関であるが，逆は必ずしも成り立たない．

1.3 Coffee Break!

$$\begin{aligned}\sum_{k=1}^{n} k &= 1+2+\cdots+n \\ &= \frac{1}{2}\{(1+2+\cdots+n)+(n+(n-1)+\cdots+1)\} \\ &= \frac{1}{2}(n+1)\cdot n = \frac{1}{2}n(n+1)\end{aligned}$$

$$\begin{aligned}\sum_{k=1}^{n}\{(k+1)^3-k^3\} &= \{2^3-1^3\}+\{3^3-2^3\}+\cdots+\{(n+1)^3-n^3\} \\ &= (n+1)^3-1\end{aligned}$$

である．また，

$$\begin{aligned}(\text{左辺}) &= \sum_{k=1}^{n}(3k^2+3k+1) \\ &= 3\sum_{k=1}^{n}k^2+3\sum_{k=1}^{n}k+n \\ &= 3\sum_{k=1}^{n}k^2+\frac{3}{2}n(n+1)+n\end{aligned}$$

である．よって，

$$\begin{aligned}\sum_{k=1}^{n}k^2 &= \frac{1}{3}(n+1)^3-\frac{1}{3}-\frac{1}{2}n(n+1)-\frac{1}{3}n \\ &= \frac{1}{6}(n+1)\{2(n+1)^2-3n-2\} = \frac{1}{6}n(n+1)(2n+1)\end{aligned}$$

2 モーメント・モーメント母関数

2.1 モーメント

$$\mu_r := E(X^r)$$

を X の (原点まわりの)r 次の**積率 (モーメント, moment)** という．

$$\mu_r' := E(X-\mu)^r$$

を X の期待値のまわりの r 次のモーメントという．[2]

$$\alpha_r \ := \ E\{(X-\mu)/\sigma\}^r$$

は X の r 次の標準化モーメントという．

$$\mu_1 = E(X),\ \mu_2' = V(X)$$

$$\mu_0 \equiv 1,\ \mu_1' \equiv 0$$

2.2 モーメント母関数 moment generating function

$$M_X(t) \ := \ E(e^{tX})$$

をモーメント母関数という．モーメント母関数は任意の次数のモーメントを生成する (計算する) ために使う．

$$\begin{array}{ll} \text{離散型分布} & M_X(t) = \displaystyle\sum_x e^{tx} f(x), \\ \text{連続型分布} & M_X(t) = \displaystyle\int_{-\infty}^{\infty} e^{tx} f(x) dx \end{array}$$

[3]

$$\begin{aligned} M_X'(0) &= \mu_1 \quad (= E(X)), \\ M_X''(0) &= \mu_2 \quad (= E(X^2)), \\ M_X'''(0) &= \mu_3 \quad (= E(X^3)), \\ &\vdots \\ M_X^{(r)}(0) &= \mu_r \quad (= E(X^r)) \end{aligned}$$

何故こうなるか?

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \\ \rightarrow \quad e^{tX} &= 1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots \end{aligned}$$

[2] モーメント：力学のモーメント（積率，能率) と数学的に似ているかららしい
[3] この定義で無限和・積分が存在しないこともある．

より，両辺の期待値をとると，

$$\begin{aligned} E(e^{tX}) &= E(1)+E(tX)+E(\frac{(tX)^2}{2!})+E(\frac{(tX)^3}{3!})+\cdots \\ \rightleftharpoons \quad M_X(t) &= 1+tE(X)+\frac{1}{2!}E(X^2)t^2+\frac{1}{3!}E(X^3)t^3+\cdots \\ &= 1+\mu_1 t+\frac{\mu_2}{2!}t^2+\frac{\mu_3}{3!}t^3+\cdots \end{aligned}$$

である．即ち，モーメント母関数 $M_X(t)$ は各項の係数にモーメントを含んでいる！従って，微分すれば低次の項は消え，さらに $t=0$ とおくことで高次の項も消えるのである！すばらしい！

2.3 (離散型) 確率分布

2.4 (離散) 一様分布

$$f(x) = \frac{1}{n} \quad (x=1,2,\ldots,n)$$

$$\begin{aligned} E(X) &= \sum_x x\frac{1}{n} \\ &= \frac{1}{n}\sum_x x = \frac{1}{n}\frac{1}{2}n(n+1) = \frac{n+1}{2} \\ V(X) &= E(X^2)-E(X)^2 \\ &= \sum_x x^2\frac{1}{n}-\left(\frac{n+1}{2}\right)^2 \\ &= \frac{1}{n}\cdot\frac{1}{6}n(n+1)(2n+1)-\frac{(n+1)^2}{4} \\ &= \frac{1}{12}(n+1)\{2(2n+1)-3(n+1)\} = \frac{1}{12}(n-1)^2 \end{aligned}$$

2.5 ベルヌーイ分布

$$f(x) = \begin{cases} p & (x=0) \\ 1-p & (x=1) \end{cases}$$

$$\begin{aligned} E(X) &= 0p+1(1-p) = 1-p, \\ V(X) &= \{0^2p+1^2(1-p)\}-(1-p)^2 = p(1-p) \end{aligned}$$

2.6　二項分布 $Bi(n,p)$

$$f(x) \;=\; {}_nC_x p^x (1-p)^{n-x} \quad (x = 0, 1, \cdots, n)$$

$$\begin{aligned}
E(X) &= \sum_x x \cdot_n C_x p^x q^{n-x} \quad (q := 1-p) \\
&= \sum_x x \cdot \frac{n}{x} \cdot \frac{(n-1)!}{(n-x)!(x-1)!} p \cdot p^{x-1} q^{n-x} \\
&= \sum_x n \cdot_{n-1} C_{x-1} p \cdot p^{x-1} q^{n-x} \\
&= np \sum_x {}_{n-1}C_{x-1} p^{x-1} q^{(n-1)-(x-1)} \\
&= np \sum_y {}_mC_y p^y q^{m-y} \quad (m := n-1, y := x-1) \\
&= np \sum_y f(y) \;=\; np, \\
E(X^2) &= \sum_x x^2 \cdot_n C_x p^x q^{n-x} \quad (q := 1-p) \\
&= np \sum_x x_{n-1}C_{x-1} p^{x-1} q^{n-x} \\
&= np \sum_x (x-1+1)_{n-1}C_{x-1} p^{x-1} q^{n-x} \\
&= np \left\{ \sum_x (x-1)_{n-1}C_{x-1} p^{x-1} q^{n-x} + \sum_x {}_{n-1}C_{x-1} p^{x-1} q^{n-x} \right\} \\
& \left(\begin{array}{ll} = & np \left\{ \sum_y y_m C_y p^y q^{m-y} + \sum_y {}_mC_y p^y q^{m-y} \right\} \quad (m := n-1, y := x-1) \\ = & np(mp+1) \end{array} \right) \\
&= np\{(n-1)p + 1\} \;=\; n(n-1)p^2 + np, \\
V(X) &= E(X^2) - E(X)^2 \;=\; n(n-1)p^2 + np - (np)^2 \;=\; np(1-p)
\end{aligned}$$

2.6.1　二項分布を正規分布で近似

二項分布 $Bi(n,p)$（平均 np, 分散 npq）は $n \to \infty$ の時正規分布に近づく．二項分布 $f(x) =_n C_x p^x q^{n-x}$ について $n \to \infty$ の極限を考える．$f(x)$ が最大となる点 $x = \bar{x}$ でテーラー展開する．

自然対数をとって，スターリング近似[4]を利用すると，

$$\begin{aligned}
\ln f(x) &= \ln n! - \ln x! - \ln(n-x)! + x \ln p + (n-x) \ln q \\
&= (n \ln n - n) - (x \ln x - x) - ((n-x)\ln(n-x) - (n-x)) + x \ln p + (n-x) \ln q \\
&= (\ln p - \ln q)x - x \ln x - (n-x)\ln(n-x) + n \ln n + n \ln q
\end{aligned}$$

[4] $\ln y! \approx y \ln y - y$

$x = \bar{x} + \Delta x$ としてテーラー展開すると，

$$\ln f(\bar{x} + \Delta x) \ = \ \ln f(\bar{x}) + B_1 \Delta x + \frac{1}{2} B_2 (\Delta x)^2 + \cdots$$

ここで

$$B_1 \ = \ \frac{d \ln f(x)}{dx} \ = \ \frac{f'(\bar{x})}{f(\bar{x})} \ = \ 0 \quad (\bar{x} \text{ で極大})$$

また，

$$\begin{aligned} \frac{d \ln f(x)}{dx} \ &= \ (\ln p - \ln q) - \ln x - x \frac{1}{x} + \ln(n - x) + (n - x) \frac{1}{n - x} \\ &= \ (\ln p - \ln q) - \ln x + \ln(n - x) \end{aligned}$$

より，

$$\begin{aligned} & B_1 \ = \ \frac{d \ln f(\bar{x})}{dx} \ = \ (\ln p - \ln q) - \ln \bar{x} + \ln(n - \bar{x}) \ = \ 0 \\ & \rightleftharpoons \ln \left\{ \frac{p(n - \bar{x})}{q \bar{x}} \right\} \ = \ 0 \\ & \rightleftharpoons \frac{p(n - \bar{x})}{q \bar{x}} \ = \ 1 \ \rightleftharpoons \ np = (p + q) \bar{x} = \bar{x} \end{aligned}$$

即ち，二項分布の平均 $(np = \bar{x})$ が $f(x)$ の極大値を与える．(平均が極大を与える $\to$ 正規分布と同じ性質)

　また，

$$\begin{aligned} B_2 \ = \ \frac{d^2 \ln f(x)}{dx^2} \ &= \ -\frac{1}{\bar{x}} - \frac{1}{n - \bar{x}} \\ &= \ -\frac{1}{np} - \frac{1}{n - np} \ = \ -\frac{1}{n} \left(\frac{1}{p} + \frac{1}{q} \right) \ = \ -\frac{1}{npq} \end{aligned}$$

n が十分大きい場合を考えるので 3 次以降の項は無視すると，

$$\ln f(\bar{x} + \Delta x) \ = \ \ln f(\bar{x}) + \frac{1}{2} B_2 (\Delta x)^2 \ = \ \ln f(\bar{x}) - \frac{1}{2npq} (\Delta x)^2$$

npq は二項分布の分散なので $\sigma^2 := npq$ とおき，一般の変数 $x = \bar{x} + \Delta x$ について考えると，

$$f(x) \ = \ f(\bar{x}) \exp \left\{ -\frac{1}{2\sigma^2} (x - \bar{x})^2 \right\}$$

これが $n \to \infty$ としたときの二項分布の確率密度関数！

$$\int_{-\infty}^{\infty} f(x) dx = 1 \quad \text{より，} \quad f(\bar{x}) \ = \ \frac{1}{\sigma \sqrt{2\pi}}$$

であるので，

$$f(x) \ = \ \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(x - \bar{x})^2}{2\sigma^2} \right\}$$

となる．これはまさに平均 $\bar{x}$，分散 σ^2 の正規分布！

　即ち，n が十分大きい場合は，二項分布は，平均 $\bar{x} := np$，分散 $\sigma^2 := npq$ の正規分布で近似できる！

2.7 ポアソン分布

$$f(x) \;=\; \exp(-\lambda)\frac{\lambda^x}{x!} \quad (x = 0, 1, 2, \cdots)$$

$$\begin{aligned}
E(X) \;=\; \sum_x xf(x) \;&=\; \sum_{x=0}^{\infty} e^{-\lambda}\frac{\lambda^x}{(x-1)!} \\
&=\; \sum_{x=1}^{\infty} e^{-\lambda}\frac{\lambda^x}{(x-1)!} \\
&=\; e^{-\lambda}\lambda\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!} \\
&=\; e^{-\lambda}\lambda\sum_{t=0}^{\infty}\frac{\lambda^t}{t!} \;=\; e^{-\lambda}\lambda e^{\lambda} \;=\; \lambda, \\
E(X(X-1)) \;&=\; \sum_{x=0}^{\infty} x(x-1)e^{-\lambda}\frac{\lambda^x}{x!} \\
&=\; e^{-\lambda}\sum_{x=2}^{\infty} x(x-1)\frac{\lambda^x}{x!} \\
&=\; e^{-\lambda}\sum_{x=2}^{\infty}\frac{\lambda^x}{(x-2)!} \\
&=\; e^{-\lambda}\lambda^2\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!} \;=\; e^{-\lambda}\lambda^2 e^{\lambda} \;=\; \lambda^2 \\
V(X) \;=\; E(X^2) - E(X)^2 \;&=\; E(X^2) - E(X) - E(X)^2 + E(X) \\
&=\; E(X^2 - X) - E(X)^2 + E(X) \\
&=\; E(X(X-1)) - E(X)^2 + E(X) \;=\; \lambda^2 - \lambda^2 + \lambda \;=\; \lambda
\end{aligned}$$

2.7.1 二項分布からポアソン分布へ

二項分布について $n \to \infty$, $p \to 0$ とする．ただし，二項分布の期待値が一定 $np = \lambda$ のもとで.

$$\begin{aligned}
f(x) \;&=\; {}_nC_x p^x(1-p)^{n-x} \\
&=\; \frac{n!}{(n-x)!x!}p^x(1-p)^{n-x} \\
&=\; \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!}p^x(1-p)^{n-x} \\
&=\; \frac{n^x}{x!}\left\{1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\right\}p^x(1-p)^{n-x}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{n^x}{x!}\left\{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\right\}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} \\
&= \frac{\lambda^x}{x!}\left\{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\right\}\left(1-\frac{\lambda}{n}\right)^{n-x} \\
&= \frac{\lambda^x}{x!}\left\{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\right\}\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x} \\
&= \frac{\lambda^x}{x!}\left\{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\right\}\left\{\left(1+\frac{1}{m}\right)^{m}\right\}^{-\lambda}\left(1-\frac{\lambda}{n}\right)^{-x}
\end{aligned}
$$

最後の等式では，$-\frac{\lambda}{n}=\frac{1}{m}$ とした．ここで $n\to\infty$ (このとき $p\to 0$, $m\to-\infty$) とすると，

$$
\lim_{m\to\pm\infty}\left(1+\frac{1}{m}\right)^m \ = \ e
$$

より，

$$
f(x) \ = \ \frac{1}{x!}\lambda^x\exp(-\lambda)
$$

2.8 幾何分布

$$
f(x) \ = \ p(1-p)^{x-1} \quad (x=1,2,\cdots)
$$

$$
\begin{aligned}
E(X) \ = \ \sum_{x=1}^{\infty}xp(1-p)^{x-1} &= \sum_{x=1}^{\infty}(x-1+1)pq(1-p)^{x-2} \\
&= (1-p)\sum_{x=1}^{\infty}(x-1)p(1-p)^{x-2}+\sum_{x=1}^{\infty}p(1-p)^{x-1} \\
&= (1-p)\sum_{x=2}^{\infty}(x-1)p(1-p)^{x-2}+1 \\
&= (1-p)\sum_{x=1}^{\infty}xp(1-p)^{x-1}+1 \ = \ (1-p)E(X)+1 \\
\rightleftharpoons \qquad & E(X) \ = \ \frac{1}{p}, \\
E(X^2) \ = \ \sum_{x=1}^{\infty}x^2p(1-p)^{x-1} &= \sum_{x=1}^{\infty}(x-1+1)^2p(1-p)^{x-1} \\
&= \sum_{x=1}^{\infty}(x-1)^2p(1-p)^{x-1}+2\sum_{x=1}^{\infty}xp(1-p)^{x-1}-\sum_{x=1}^{\infty}p(1-p)^{x-1} \\
&= (1-p)\sum_{x=2}^{\infty}(x-1)^2p(1-p)^{x-2}+\frac{2}{p}-1
\end{aligned}
$$

$$
\begin{aligned}
&= (1-p)\sum_{x=1}^{\infty} xp(1-p)^{x-1} + \frac{2-p}{p} \\
&= (1-p)E(X^2) + \frac{2-p}{p} \\
\rightleftharpoons \quad & E(X^2) \ = \ \frac{2-p}{p^2}, \\
V(X) \ = \ E(X^2) - E(X)^2 \ &= \ \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 \ = \ \frac{1-p}{p^2}
\end{aligned}
$$

2.9 負の二項分布

$$
f(x) \ = \ {}_{k+x-1}C_x p^k (1-p)^x, \quad (x = 0, 1, 2, \cdots)^5
$$

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x \cdot {}_{k+x-1}C_x p^k (1-p)^x \\
&= \sum_{x=1}^{\infty} x \cdot \frac{k+x-1}{x} \cdot \frac{(k+x-2)!}{(x-1)!(k-1)!} p^k (1-p)^x \\
&= \sum_{x=1}^{\infty} \{k + (x-1)\} \cdot {}_{k+x-2}C_{x-1} p^k (1-p)^x \\
&= k(1-p)\sum_{x=1}^{\infty} {}_{k+x-2}C_{x-1} p^k (1-p)^{x-1} + (1-p)\sum_{x=1}^{\infty} (x-1) \cdot {}_{k+x-2}C_{x-1} p^k (1-p)^{x-1} \\
&= k(1-p)\sum_{y=0}^{\infty} {}_{k+y-1}C_y p^k (1-p)^y + (1-p)\sum_{y=0}^{\infty} y \cdot {}_{k+y-1}C_y p^k (1-p)^y \\
&= k(1-p) + (1-p)E(X) \\
\rightleftharpoons \quad & E(X) \ = \ \frac{k(1-p)}{p}, \\
E(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot {}_{k+x-1}C_x p^k (1-p)^x \\
&= \sum_{x=1}^{\infty} (x-1+1)^2 \cdot {}_{k+x-1}C_x p^k (1-p)^x \\
&= \sum_{x=1}^{\infty} (x-1)^2 \cdot {}_{k+x-1}C_x p^k (1-p)^x + 2\sum_{x=1}^{\infty} x \cdot {}_{k+x-1}C_x p^k (1-p)^x - \sum_{x=1}^{\infty} {}_{k+x-1}C_x p^k (1-p)^x \\
&= \sum_{x=1}^{\infty} (x-1)^2 \cdot {}_{k+x-1}C_x p^k (1-p)^x + 2E(X) - 1
\end{aligned}
$$

[5] $k = 1$ の時は幾何分布

$$
\begin{aligned}
&= \cdots \\
&= \frac{k(1-p)(1+k(1-p))}{p^2}, \\
V(X) &= E(X^2)-E(X)^2 = \frac{k(1-p)(1+k(1-p))}{p^2} - \frac{k^2(1-p)^2}{p^2} = \frac{k(1-p)}{p^2}
\end{aligned}
$$

2.10 超幾何分布

$$
f(x) = \frac{{}_M C_x \cdot {}_{N-M} C_{n-x}}{{}_N C_n}, \quad (x = \max\{0, n-(N-M)\}, \cdots, \min\{n, M\})
$$

$0 \geq n-(N-M)$ かつ $n \leq M$ の場合のみ示す.

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{n} x \cdot \frac{{}_M C_x \cdot {}_{N-M} C_{n-x}}{{}_N C_n} \\
&= \sum_{x=1}^{n} x \cdot \frac{{}_M C_x \cdot {}_{N-M} C_{n-x}}{{}_N C_n} \\
&= \sum_{x=1}^{n} x \cdot \frac{M!}{x!(M-x)!} \cdot \frac{n!(N-n)!}{N!} \cdot {}_{N-M} C_{n-x} \\
&= \sum_{x=1}^{n} \frac{M \cdot (M-1)!}{(x-1)!(M-x)!} \cdot \frac{n \cdot (n-1)!(N-n)!}{N \cdot (N-1)!} \cdot {}_{N-M} C_{n-x} \\
&= \frac{nM}{N} \sum_{x=1}^{n} \frac{{}_{M-1} C_{x-1} \cdot {}_{N-M} C_{n-x}}{{}_{N-1} C_{n-1}} \\
&= \frac{nM}{N} \sum_{x=0}^{n-1} \frac{{}_{M-1} C_x \cdot {}_{(N-1)-(M-1)} C_{(n-1)-x}}{{}_{N-1} C_{n-1}} \\
&= \frac{nM}{N} \sum_{x=0}^{n'} \frac{{}_{M'} C_x \cdot {}_{N'-M'} C_{n'-x}}{{}_{N'} C_{n'}} = \frac{nM}{N}, \\
E(X^2) &= \sum_{x=0}^{n} x^2 \cdot \frac{{}_M C_x \cdot {}_{N-M} C_{n-x}}{{}_N C_n} \\
&= \frac{nM}{N} \sum_{x=1}^{n} x \cdot \frac{{}_{M-1} C_{x-1} \cdot {}_{N-M} C_{n-x}}{{}_{N-1} C_{n-1}} \qquad (E(X) \text{ の証明より}) \\
&= \frac{nM}{N} \sum_{x=1}^{n} x \cdot \frac{{}_{N-M} C_{n-x}}{{}_{N-1} C_{n-1}} \cdot \{{}_M C_x - {}_{M-1} C_x\} \qquad (\text{脚注}^{6}\text{参照}) \\
&= \cdots \\
&= \frac{nM}{N} \cdot \frac{N-M+n(M-1)}{N-1} \\
V(X) &= E(X^2) - E(X)^2
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{nM}{N}\cdot\frac{N-M+n(M-1)}{N-1}-\frac{n^2M^2}{N^2} \\
&= \frac{nM}{N}\cdot\frac{N(N-M)+nN(M-1)-nM(N-1)}{N(N-1)} \\
&= \frac{nM}{N}\cdot\frac{N(N-M)-n(N-M)}{N(N-1)} \\
&= \frac{nM}{N}\cdot\frac{N-M}{N}\cdot\frac{N-n}{N-1}
\end{aligned}
$$

[6] ${}_MC_x - {}_{M-1}C_x = \frac{M}{x}\cdot\frac{(M-1)!}{(x-1)!(M-x)!} - \frac{M-x}{x}\frac{(M-1)!}{(x-1)!(M-x)!} = \frac{M-(M-x)}{x}{}_{M-1}C_{x-1} = {}_{M-1}C_{x-1}$