2021年 01月 13日

第14回 数理統計 レポート

小森 一輝

定義6.6 標本積率

標本を $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ とする．このとき，**原点まわりの r 次の標本積率** M_r' (r th sample moment about the origin) を，

$$M_r'(\boldsymbol{X}) = \frac{1}{n}\sum_{i=1}^{n} X_i^r \tag{6.9}$$

で定義する．また，$c\,(\in \mathbb{R})$ **まわりの r 次の標本積率** M_r (r th sample moment about c) を

$$M_r(\boldsymbol{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - c)^r \tag{6.10}$$

で定義する．特に，原点まわりの 1 次の標本積率を**標本平均** (sample mean) とよび，$\bar{X}_n$ で記す．すなわち，

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{6.11}$$

である．また，$\bar{X}_n$ まわりの 2 次の標本積率を**標本分散** (sample variance) とよび，$S_n^2(\boldsymbol{X})$ で記す．すなわち，

$$S_n^2(\boldsymbol{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \tag{6.12}$$

である．さらに，

$$U_n^2(\boldsymbol{X}) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \tag{6.13}$$

を**不偏標本分散** (unbiased sample variance) とよぶ．

- 原点まわりの r次の積率

$$ ¥mu_r^{¥prime} = E(X^n) $$

- 平均まわりの r次の積率

$$ ¥mu_r = E((X-¥mu)^n)¥ ¥mu = E(X) = ¥mu_{1}^{¥prime} $$

定義6.7 標本共分散, 標本相関係数

2 つの変数に関して対応する標本 $(\boldsymbol{X}, \boldsymbol{Y}) = ((X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n))$ が与えられているとき，**標本共分散** (sample covariance) $C_n(\boldsymbol{X}, \boldsymbol{Y})$，**標本相関係数** (sample correlation coefficient) $R_n(\boldsymbol{X}, \boldsymbol{Y})$ はそれぞれ以下で定義される．

$$C_n(\boldsymbol{X}, \boldsymbol{Y}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)(Y_i - \bar{Y}_n), \quad R_n(\boldsymbol{X}, \boldsymbol{Y}) = \frac{C_n(\boldsymbol{X}, \boldsymbol{Y})}{\sqrt{S_n^2(\boldsymbol{X})S_n^2(\boldsymbol{Y})}} \tag{6.14}$$

定理 6.6. 標本共分散の性質
標本共分散について，以下が成り立つ．

$$C_n(\boldsymbol{X}, \boldsymbol{X}) = S_n^2(\boldsymbol{X}), \quad C_n(\boldsymbol{X}, \boldsymbol{Y}) = C_n(\boldsymbol{Y}, \boldsymbol{X}) \tag{6.15}$$

$$C_n(\boldsymbol{X}, \boldsymbol{Y} + \boldsymbol{Z}) = C_n(\boldsymbol{X}, \boldsymbol{Y}) + C_n(\boldsymbol{X}, \boldsymbol{Z}) \tag{6.16}$$

$$C_n(a\boldsymbol{X}, \boldsymbol{Y}) = aC_n(\boldsymbol{X}, \boldsymbol{Y}) \ (a: \text{定数}) \tag{6.17}$$

ここで，$\boldsymbol{Y} + \boldsymbol{Z}$ および $a\boldsymbol{X}$ はそれぞれ標本 $Y_1 + Z_1, Y_2 + Z_2, \ldots, Y_n + Z_n$ および標本 $aX_1, aX_2, \ldots, aX_n$ を表す．

- データ：共分散 $$ S_{xy} = ¥frac{1}{n} ¥sum_{i=1}^{n}(x_i - ¥overline{x})(y_i - ¥overline{y}) $$
- 標本共分散

$$ C(¥textbf{X}, ¥textbf{Y}) = ¥frac{1}{n}¥sum(X_i - ¥overline{X})(Y_i - ¥overline{Y}) $$

定義6.7 標本積率の期待値と分散

母集団確率変数 X の r 次の積率を $E(X^r) = \mu_r'$ $(r = 1, 2, \ldots)$ とし，無作為標本を $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ とする．このとき，標本積率 M_r' について，以下が成り立つ．

$$E(M_r') = \mu_r', \quad V(M_r') = \frac{1}{n}(\mu_{2r}' - \mu_r'^2) \tag{6.18}$$

証明

$$ ¥begin{aligned} E(M_r^{¥prime}) &= E(¥frac{1}{n}¥sum_{i=1}^{n}X_i^{r})¥ &= ¥frac{1}{n}¥sum_{i=1}^{n}E(X_i^{r})¥ &= ¥frac{1}{n}n¥mu_r^{¥prime} = ¥mu_r^{¥prime}¥ ¥end{aligned} $$

母集団	標本
X	$X_1, X_2 ¥dots$
$E(X) = ¥mu$	$E(X_i) = E(X) = ¥mu$
$V(X) = ¥sigma^2$	$V(X_i) = V(X) = ¥sigma^2$
$E(X^n) = ¥mu_r^{¥prime}$	$E(X_i^{r}) = E(X^r) = ¥mu_r^{¥prime}$

原点まわりの r次の標本積率の期待値は, 母集団確率変数の原点まわりの r次の積率と一致する

$$ ¥begin{aligned} V(M_r^{¥prime}) &= V(¥frac{1}{n}¥sum_{i=1}^{n}X_i^{r})¥ &= ¥frac{1}{n^2}¥sum_{i=1}^{n}V(X_i^{r})¥ &= ¥frac{1}{n^2}¥sum_{i=1}^{n}{E(X_i^{2r}) - (E(X_i^r))^2 }¥ &= ¥frac{1}{n^2}¥sum_{i=1}^{n}(¥mu_{2r}^{¥prime} - (¥mu_r^{¥prime})^2) = ¥frac{1}{n}(¥mu_{2r}^{¥prime} - (¥mu_r^{¥prime})^2)¥ ¥end{aligned} $$

定義6.8 標本平均・標本分散の期待値と分散

母集団確率変数Xに関して，$E(X)=\mu$, $V(X)=\sigma^2$, $E((X-\mu)^r)=\mu_r\ (r=1,2,\ldots)$ とし，無作為標本を $\boldsymbol{X}=(X_1,X_2,\ldots,X_n)$ とする　このとき，標本平均$\bar{X}_n$，標本分散S_n^2について，以下が成り立つ．

$$\text{i)}\ E(\bar{X}_n)=\mu, \qquad \text{ii)}\ V(\bar{X}_n)=\frac{\sigma^2}{n} \qquad (6.19)$$

$$\text{iii)}\ E(S_n^2)=\frac{n-1}{n}\sigma^2, \qquad \text{iv)}\ V(S_n^2)=\frac{(n-1)^2}{n^3}\left(\mu_4-\frac{n-3}{n-1}\sigma^4\right) \qquad (6.20)$$

$$\text{v)}\ E(U_n^2)=\sigma^2, \qquad \text{vi)}\ V(U_n^2)=\frac{1}{n}\left(\mu_4-\frac{n-3}{n-1}\sigma^4\right) \qquad (6.21)$$

ただし，$n>1$とする．

証明

i) $$ E(¥overline{X}) = E(¥frac{1}{n}¥sum_{i=1}^{n}X_i) = ¥frac{1}{n}¥sum_{i=1}^{n}E(X_i) = ¥frac{1}{n}n¥mu = ¥mu $$

ii) $$ V(¥overline{X}) = V(¥frac{1}{n}¥sum_{i=1}^{n}X_i) = ¥frac{1}{n}¥sum_{i=1}^{n}V(X_i) = ¥frac{1}{n^2}n¥sigma^2 = ¥frac{¥sigma^2}{n} $$

iii)

Tips

$E(X) = E(X_i) = E(¥overline{X}) = ¥mu$

$$ ¥begin{aligned} E(S^2) &= E(¥frac{1}{n}¥sum_{i=1}^{n}(X_i - ¥overline{X})^2)¥ &= E(¥frac{1}{n}¥sigma{(X_i - ¥mu) - (¥overline{X} - ¥mu)}^2)¥ &= ¥frac{1}{n}¥sum E[(X_i - ¥mu)^2] - ¥frac{2}{n}¥sum E[(X_i - ¥mu)(¥overline{X} - ¥mu)] + ¥frac{1}{n}¥sum E[(¥overline{X} - ¥mu)^2]¥ ¥end{aligned}¥ $$ $$ ¥mu = E(X_i) より, E[(X_i - ¥mu) = V(X_i)¥ ¥mu = E(¥overline{X}) より, E[(¥overline{X} - ¥mu) = V(¥overline{X})¥ $$

ここで, $$ ¥begin{aligned} E[(X_i - ¥mu)(¥overline{X} - ¥mu)] &= E[(X_i - ¥mu)(¥frac{1}{n}¥sum_{j=1}^{n}X_j - ¥mu)]¥ &= E[(X_i - ¥mu)(¥frac{1}{n}(¥sum_{j=1}^{n}(X_j - ¥mu)))]¥ &= ¥frac{1}{n}¥sum_{j=1}^{n}E((X_i - ¥mu)(X_j - ¥mu))¥ &= ¥frac{1}{n}¥sum_{j=1}^{n}Cov(X_i, X_j)¥ ¥end{aligned} $$

なので,

$$ ¥begin{aligned} &= ¥frac{1}{n}¥sum E[(X_i - ¥mu)^2] - ¥frac{2}{n}¥sum E[(X_i - ¥mu)(¥overline{X} - ¥mu)] + ¥frac{1}{n}¥sum E[(¥overline{X} - ¥mu)^2]¥ &= ¥frac{1}{n}¥sum V(X_i) - ¥frac{2}{n^2}¥sum_{i=1}^{n}¥sum_{j=1}^{n} Cov(X_i, X_j) + ¥frac{1}{n}¥sum V(¥overline{X})¥ &= ¥frac{1}{n}n¥sigma^2 - ¥frac{2}{n^2}n¥sigma^2 + ¥frac{1}{n}n¥frac{¥sigma^2}{n}¥ &= ¥sigma^2 - ¥frac{2}{n}¥sigma^2 + ¥frac{1}{n}¥sigma^2¥ &= ¥frac{n-1}{n}¥sigma^2 ¥end{aligned} $$ Tips

$¥sum_{i¥neq j}¥sum_{j¥neq i}Cov(X_i, X_j) = 0, ¥sum_{i=j}¥sum_{j=i}V(X_i) = ¥sigma^2$

定理 6.9. 母集団確率変数 X に関して，$E(X)=\mu$, $E\left((X-\mu)^3\right)=\mu_3$ とし，無作為標本を $\boldsymbol{X}=(X_1, X_2, \ldots, X_n)$ とする．このとき，以下が成り立つ．

$$E\left((\bar{X}_n-\mu)^3\right)=\frac{\mu_3}{n^2}, \quad \mathrm{Cov}(\bar{X}_n, S_n^2)=\frac{n-1}{n^2}\mu_3 \tag{6.22}$$

証明

$$ ¥begin{aligned} E((¥overline{X}-¥mu)^3) &= E((¥frac{1}{n}¥sum_{i=1}^{n}X_i - ¥mu)^3)¥ &= E[(¥frac{1}{n}¥sum (X_i - ¥mu))^3]¥ &= ¥frac{1}{n^3}E[(¥sum(X_i - ¥mu))^3]¥ &= ¥frac{1}{n^3}E[(X_i - ¥mu)^3]¥ &= ¥frac{1}{n^3}n¥mu_3 = ¥frac{¥mu^3}{n^2}¥ ¥end{aligned} $$

Tips $$ ¥begin{aligned} &E(¥sum_{i=1}¥sum_{j=1}¥sum_{k=1}(X_i - ¥mu)(X_j - ¥mu)(X_k - ¥mu))¥ &= ¥sum_{i=1}¥sum_{j=1}¥sum_{k=1}E[(X_i - ¥mu)(X_j - ¥mu)(X_k - ¥mu)]¥ &= ¥sum_{i=1}^{n}E(X_i - ¥mu)^3 ¥end{aligned} $$

$$ ¥begin{aligned} &Cov(¥overline{X}, S^2) = Cov(¥frac{1}{n}¥sum(X_i-¥mu), ¥frac{1}{n}¥sum (X_i - ¥mu)^2 - (¥overline{X} - ¥mu)^2)¥ &= Cov(¥frac{1}{n}¥sum(X_i - ¥mu), ¥frac{1}{n}¥sum(X_i - ¥mu)^2) - Cov(¥frac{1}{n}¥sum(X_i - ¥mu),(¥overline{X} - ¥mu)^2)¥ &= ¥frac{1}{n^2}¥sum (Cov(X_i-¥mu, (X_i-¥mu)^2) - Cov(¥overline{X} - ¥mu, (¥overline{X} - ¥mu)^2))¥ &= ¥frac{1}{n^2}(E(X_i - ¥mu)^3 - E((¥overline{X}-¥mu)^3))¥ &= ¥frac{1}{n^2}n¥mu_3 - ¥frac{¥mu_3}{n^2} = ¥frac{n-1}{n^2}¥mu_3 ¥end{aligned} $$

Tips $$ ¥begin{aligned} &¥frac{1}{n}¥sum (X_i - ¥overline{X})^2¥ &= ¥frac{1}{n}¥sum ((X_i - ¥mu) - (¥overline{X} - ¥mu))^2¥ &= ¥frac{1}{n}¥sum {(X_i - ¥mu)^2 + (¥overline{X} - ¥mu)^2 -2(X_i - ¥mu)(¥overline{X} - ¥mu)}¥ ¥end{aligned} $$

標本平均と標本分散の Cov は $¥mu_3$ で表される．(3次の積率)

※正規分布 気数字の積率は0

命題 6.10. 母集団確率変数 X に関して，$E(X)=\mu$，$V(X)=\sigma^2$ とし，無作為標本を $\boldsymbol{X}=(X_1, X_2, \ldots, X_n)$ とする．このとき $Y_i=\sum_{j=1}^{i} X_j\ (i=1, 2, \ldots, n)$ とすれば，以下が成り立つ．

$$\mathrm{Cov}(Y_{n-1}, Y_n)=(n-1)\sigma^2, \quad \rho(Y_{n-1}, Y_n)=\sqrt{\frac{n-1}{n}} \tag{6.23}$$