Math 3280 20-11-16

Review

- $E[X_1 + \cdots + X_n] = \sum_{i=1}^{n} E[X_i]$

(linearity of expectation)

- For two r.v.'s X and Y,

$$\mathrm{Cov}(X, Y) = E[(X - E[X]) \cdot (Y - E[Y])]$$

In particular, $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.

Remark: $\mathrm{Cov}(X, Y) = E[XY] - E[X] \cdot E[Y]$.

Pf. Write $\mu = E[X]$, $\nu = E[Y]$.

Then
$$\begin{aligned}
\mathrm{Cov}(X, Y) &= E[(X-\mu)(Y-\nu)] \\
&= E[XY - \mu Y - \nu X + \mu\nu] \\
&= E[XY] - \mu E[Y] - \nu E[X] + \mu\nu \\
&= E[XY] - \mu\nu.
\end{aligned}$$

$\boxtimes$

- If X and Y are independent, then

$$\mathrm{Cov}(X, Y) = 0.$$

Prop 1. (1) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$.

(2) $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.

(3) $\mathrm{Cov}(aX, Y) = a\,\mathrm{Cov}(X, Y)$, $a \in \mathbb{R}$.

(4) $$\mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Cov}(X_i, Y_j).$$

(1), (3), (4) imply that $\mathrm{Cov}(\cdot, \cdot)$ is bi-linear.

Pf. Let us prove (4) only. Write $\mu_i = E[X_i]$,
$\nu_i = E[Y_j]$.

Then by definition,

$$\mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = E\left[\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i\right) \cdot \left(\sum_{j=1}^{m} Y_j - \sum_{j=1}^{m} \nu_j\right)\right]$$

$$= E\left[\sum_{i=1}^{n} (X_i - \mu_i) \sum_{j=1}^{m} (Y_j - \nu_j) \right]$$

$$= E\left[\sum_{i=1}^{n} \sum_{j=1}^{m} (X_i - \mu_i)(Y_j - \nu_j) \right]$$

By the linearity of E

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} E\left[(X_i - \mu_i)(Y_j - \nu_j) \right]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} \mathrm{Cov}(X_i, Y_j).$$

$\square$

Corollary 2. $\mathrm{Var}\left(\sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2 \cdot \sum_{1 \le i < j \le n} \mathrm{Cov}(X_i, X_j)$.

Moreover if $X_1, \cdots, X_n$ are pairwise independent, then

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \mathrm{Var}(X_i).$$

Example 1. Let $X_1, \cdots, X_n$ be independent and identical distributed, having expected value μ and variance σ^2.

Let $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ (Sample mean)

$$S^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1} \quad \text{(sample variance)}$$

Find out ① $Var(\bar{X})$.

② $E[S^2]$.

Solution: $Var(\bar{X}) = Var\left(\frac{X_1 + \cdots + X_n}{n}\right)$

$$= \frac{1}{n^2} Var(X_1 + \cdots + X_n)$$

(notice $X_1, \cdots, X_n$ are independent)

$$= \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i)$$

$$= \frac{n \cdot \sigma^2}{n^2} = \sigma^2/n.$$

Recall that $S^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}$.

So

$$(n-1)S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2$$

$$= \sum_{i=1}^{n} \left((X_i - \mu) - (\bar{X} - \mu)\right)^2$$

$$= \sum_{i=1}^{n} \left[(X_i - \mu)^2 - 2(X_i - \mu)(\bar{X} - \mu) + (\bar{X} - \mu)^2\right]$$

$$= \sum_{i=1}^{n} (X_i - \mu)^2 - 2(\bar{X} - \mu) \cdot \sum_{i=1}^{n} (X_i - \mu) + n \cdot (\bar{X} - \mu)^2$$

$$= \sum_{i=1}^{n} (X_i - \mu)^2 - 2n(\bar{X} - \mu)^2 + n(\bar{X} - \mu)^2$$

$$= \sum_{i=1}^{n} (X_i - \mu)^2 - n(\bar{X} - \mu)^2.$$

Hence

$$E\left[(n-1)S^2\right] = \sum_{i=1}^{n} E\left[(X_i - \mu)^2\right] - n\, E\left[(\bar{X} - \mu)^2\right]$$

$$= n \cdot \sigma^2 - n \cdot \frac{\sigma^2}{n}$$

$$= (n-1)\sigma^2.$$

Thus $E[S^2] = \sigma^2$. $\boxed{/\!/}$

Example 2. Define $\bar{X}$ as in Example 1.

Show that $\text{Cov}(X_i - \bar{X}, \bar{X}) = 0$

Pf. $\text{Cov}(X_i - \bar{X}, \bar{X})$

$$= \text{Cov}(X_i, \bar{X}) - \text{Cov}(\bar{X}, \bar{X})$$

$$= \text{Cov}\left(X_i, \frac{X_1 + \cdots + X_n}{n}\right) - \frac{\sigma^2}{n}$$

$$= \frac{1}{n} \sum_{j=1}^{n} \text{Cov}(X_i, X_j) - \frac{\sigma^2}{n}$$

$$= \frac{1}{n}\left[\text{Cov}(X_i, X_i) + \sum_{j \neq i} \text{Cov}(X_i, X_j)\right] - \sigma^2/n$$

$$= \frac{1}{n} \text{Var}(X_i) - \sigma^2/n = \frac{\sigma^2}{n} - \frac{\sigma^2}{n} = 0.$$