# *CSL003P1M : Probability and Statistics Lecture 23 (Some Problems on Variances & Covariances and Correlation Coefficient)*

Sumit Kumar Pandey

November 08, 2021

# *Exercise*-1

Let $X_1, \ldots, X_n$ be independent and identically distributed random variables having expected value $\mu$ and variance $\sigma^2$. Let $\bar{X} = \sum_{i=1}^{n} X_i/n$ be the sample mean. The quantities $X_i - \bar{X}$, $i = 1, \ldots, n$, are called deviations, as they equal the differences between the individual data and the sample mean. The random variable

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

is called the sample variance. Find (a) $Var(\bar{X})$ and (b) $E[S^2]$.

$$
\begin{aligned}
Var(\bar{X}) &= \left(\frac{1}{n}\right)^2 Var\left(\sum_{i=1}^{n} X_i\right) \\
&= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} Var(X_i) \qquad \text{by independence} \\
&= \frac{\sigma^2}{n}.
\end{aligned}
$$

# $Exercise$-1

$$
\begin{aligned}
(n-1)S^2 &= \sum_{i=1}^{n}(X_i-\mu+\mu-\bar{X})^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2+\sum_{i=1}^{n}(\bar{X}-\mu)^2-2(\bar{X}-\mu)\sum_{i=1}^{n}(X_i-\mu) \\
&= \sum_{i=1}^{n}(X_i-\mu)^2+n(\bar{X}-\mu)^2-2(\bar{X}-\mu)n(\bar{X}-\mu) \\
&= \sum_{i=1}^{n}(X_i-\mu)^2-n(\bar{X}-\mu)^2
\end{aligned}
$$

Taking expectations both sides, we obtain

$$
\begin{aligned}
(n-1)E[S^2] &= \sum_{i=1}^{n}E[(X_i-\mu)^2]-nE[(\bar{X}-\mu)^2] \\
&= n\sigma^2-nVar(\bar{X})=(n-1)\sigma^2
\end{aligned}
$$

# *Exercise*-2

Compute the variance of a binomial random variable $X$ with parameters $n$ and $p$.

Solution:

- Let

  $$X = X_1 + X_2 + \cdots + X_n$$

  where the $X_i$ are independent Bernoulli random variables such that

  $$X_i = \begin{cases} 1 & \text{if the } i\text{th trial is a success} \\ 0 & \text{otherwise} \end{cases}$$

- Hence,

  $$Var(X) = Var(X_1) + \cdots + Var(X_n)$$

# *Exercise*-2

But,

$$\begin{aligned} Var(X_i) &= E[X_i^2] - (E[X_i])^2 \\ &= E[X_i] - (E[X_i])^2 \qquad \text{since } X_i^2 = X_i \\ &= p - p^2. \end{aligned}$$

Thus,

$$Var(X) = np(1-p)$$

# Correlation Coefficient

# *Correlation Coefficient*

Let $X$ and $Y$ be two random variables having finite nonzero variances. One measure of the degree of dependence between the two random variables is the correlation coefficient $\rho(X, Y)$ defined by

$$\rho = \rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{(Var(X)Var(Y))}}$$

- These random variables are said to be uncorrelated if $\rho = 0$.
- If $X$ and $Y$ are independent, $Cov(X, Y) = 0$ and we see at once that random variables are uncorrelated.
- The correlation coefficient $\rho$ is always between $-1$ and $1$.

Thank You