VARIANCE

STAT 510

DEFN

THE VARIANCE OF X, WITH $\mu = \mathbb{E}[X]$ IS

$$\sigma^2 = \mathbb{E}\left[(X-\mu)^2\right] = \begin{cases} \sum_x (x-\mu)^2 f(x) & X \text{ DISC} \\ \int x f(x)\, dx & X \text{ CONT} \end{cases}$$

$$= \sigma_X^2 = \mathbb{V}[X]$$

THE STANDARD DEVIATION IS $\sigma = \sqrt{\sigma^2}$

$$\sigma = \sigma_X = SD[X]$$

THM THE VARIANCE HAS THREE PROPERTIES:

1. $\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

2. GIVEN CONSTANTS a, b: $\mathbb{V}[aX+b] = a^2\mathbb{V}[X]$

3. $X_1, \ldots, X_n$ INDEPENDENT, $a_1, \ldots, a_n$ CONSTANTS THEN

$$\mathbb{V}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 \mathbb{V}[X_i]$$

Example

$X \sim \text{binom}(n, p)$

$$\mathbb{V}[X] = \sum_{i=1}^{n} (x - np)^2 \binom{n}{x} p^x (1-p)^{n-x} = \ldots\ldots$$

BERNOULLI

Again $X = \sum_{i=1}^{n} X_i$ $\quad X_i$ IND $\quad$ Recall $\mathbb{E}[X_i] = p$

$$\mathbb{E}[X_i^2] = 0 \cdot (1-p) + 1^2 \cdot p = p$$

$$\mathbb{V}[X_i] = \mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2 = p - p^2 = p(1-p)$$

$$\mathbb{V}[X] = \mathbb{V}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{V}[X_i] = \sum_{i=1}^{n} p(1-p) = np(1-p) \checkmark$$

THM $X_1, X_2, \ldots X_n$ IID WITH $\mathbb{E}[X_i] = \mu$ AND $\mathbb{V}[X_i] = \sigma^2$

1. $\mathbb{E}[\overline{X}_n] = \mu$ WHERE $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

   SAMPLE MEAN

2. $\mathbb{V}[\overline{X}_n] = \sigma^2/n$

3. $\mathbb{E}[S_n^2] = \sigma^2$ WHERE $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$

   SAMPLE VARIANCE

DEFN

X WITH MEAN μ_X AND VARIANCE σ_X^2

Y WITH MEAN μ_Y AND VARIANCE σ_Y^2

THE COVARIANCE OF X AND Y IS

$$\mathrm{Cov}[X,Y] = \mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right]$$

THE CORRELATION OF X AND Y IS

$$\rho = \rho_{X,Y} = \rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\sigma_X^2\,\sigma_Y^2}}$$

THM

1. $\mathrm{Cov}[X,Y] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$

2. $-1 \leq \rho(X,Y) \leq 1$

3. IF X, Y IND THEN $\mathrm{Cov}[X,Y] = \rho = 0$.

   THE REVERSE IS <u>NOT</u> TRUE IN GENERAL.

THM

$$\mathbb{V}[X+Y] = \mathbb{V}[X] + \mathbb{V}[Y] + 2\,\mathrm{Cov}[X,Y]$$

OR, IN GENERAL

$$\mathbb{V}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2\, \mathbb{V}[X_i] + 2 \sum_{i<j}\sum a_i a_j\, \mathrm{Cov}[X_i, X_j]$$

Variance - Covariance Matrix

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix} \qquad \mu = \mathbb{E}[X] = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix} = \begin{pmatrix} \mathbb{E}[X_1] \\ \vdots \\ \mathbb{E}[X_k] \end{pmatrix}$$

$$\Sigma = \mathbb{V}[X] = \begin{bmatrix} \mathbb{V}(X_1) & \mathrm{Cov}(X_1, Y_2) & \cdots & \mathrm{Cov}(X_1, X_k) \\ \mathrm{Cov}(X_2, X_1) & \mathbb{V}(X_1) & \cdots & \mathrm{Cov}(X_2, X_k) \\ \vdots & & \ddots & \vdots \\ \mathrm{Cov}(X_k, X_1) & \cdots & & \mathbb{V}(X_k) \end{bmatrix}$$