

# Probability Review

John Norstad

[j-norstad@northwestern.edu](mailto:j-norstad@northwestern.edu)

<http://www.norstad.org>

September 11, 2002

Updated: November 3, 2011

## **Abstract**

We define and review the basic notions of variance, standard deviation, covariance, and correlation coefficients for random variables. We give proofs of their basic properties.

## 1 Definitions and Summary of the Propositions

**Definition 1**  $E(X)$  = the expected value of a random variable  $X$  is the mean or average value of  $X$ .

**Definition 2**  $\text{Var}(X)$  = the variance of  $X = E([X - E(X)]^2)$ .

**Definition 3**  $\text{Stdev}(X)$  = the standard deviation of  $X = \sqrt{\text{Var}(X)}$ .

**Definition 4**  $\text{Cov}(X, Y)$  = the covariance of  $X$  and  $Y = E([X - E(X)][Y - E(Y)])$ .

**Definition 5**  $\text{Cor}(X, Y)$  = the correlation coefficient of  $X$  and  $Y = \frac{\text{Cov}(X, Y)}{\text{Stdev}(X)\text{Stdev}(Y)}$ .

**Proposition 1:**  $\text{Var}(X) = E(X^2) - E(X)^2$

**Proposition 2:**  $\text{Var}(aX + b) = a^2\text{Var}(X)$

**Proposition 3:**  $\text{Stdev}(aX + b) = |a|\text{Stdev}(X)$

**Proposition 4:**  $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$

**Proposition 5:**  $\text{Cov}(aX + b, cY + d) = ac\text{Cov}(X, Y)$

**Proposition 6:**  $\text{Cov}(X, X) = \text{Var}(X)$

**Proposition 7:**  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

**Proposition 8:**  $\text{Var}\left(\sum_{i=1}^n w_i X_i\right) = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \text{Cov}(X_i, X_j)$

**Proposition 9:**  $\text{Cov}\left(\sum_{i=1}^n w_i X_i, Y\right) = \sum_{i=1}^n w_i \text{Cov}(X_i, Y)$

**Proposition 10:**  $|\text{Cor}(X_1, X_2)| \leq 1$

## 2 Proofs of the Propositions

**Proposition 1**  $\text{Var}(X) = E(X^2) - E(X)^2$

**Proof:**

$$\begin{aligned}\text{Var}(X) &= E([X - E(X)]^2) \\ &= E(X^2 - 2XE(X) + E(X)^2) \\ &= E(X^2) - 2E(X)E(X) + E(X)^2 \\ &= E(X^2) - E(X)^2\end{aligned}$$

**Proposition 2**  $\text{Var}(aX + b) = a^2 \text{Var}(X)$

**Proof:**

$$\begin{aligned}\text{Var}(aX + b) &= E([(aX + b) - E(aX + b)]^2) \\ &= E([aX + b - aE(X) - b]^2) \\ &= E([a(X - E(X))]^2) \\ &= E(a^2[X - E(X)]^2) \\ &= a^2E([X - E(X)]^2) \\ &= a^2\text{Var}(X)\end{aligned}$$

**Proposition 3**  $\text{Stdev}(aX + b) = |a|\text{Stdev}(X)$

**Proof:**

$$\begin{aligned}\text{Stdev}(aX + b) &= \sqrt{\text{Var}(aX + b)} \\ &= \sqrt{a^2\text{Var}(X)} \quad (\text{by Proposition 2}) \\ &= |a|\sqrt{\text{Var}(X)} \\ &= |a|\text{Stdev}(X)\end{aligned}$$

**Proposition 4**  $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$

**Proof:**

$$\begin{aligned}\text{Cov}(X, Y) &= E([X - E(X)][Y - E(Y)]) \\ &= E(XY - XE(Y) - E(X)Y + E(X)E(Y)) \\ &= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y)\end{aligned}$$

**Proposition 5**  $Cov(aX + b, cY + d) = acCov(X, Y)$

**Proof:**

$$\begin{aligned}
 Cov(aX + b, cY + d) &= E([aX + b - E(aX + b)][cY + d - E(cY + d)]) \\
 &= E([aX + b - aE(X) - b][cY + d - cE(Y) - d]) \\
 &= E(a[X - E(X)]c[Y - E(Y)]) \\
 &= acE([X - E(X)][Y - E(Y)]) \\
 &= acCov(X, Y)
 \end{aligned}$$

**Proposition 6**  $Cov(X, X) = Var(X)$

**Proof:**

$$\begin{aligned}
 Cov(X, X) &= E([X - E(X)][X - E(X)]) \\
 &= E([X - E(X)]^2) \\
 &= Var(X)
 \end{aligned}$$

**Proposition 7**  $Cov(X, Y) = Cov(Y, X)$

**Proof:**

$$\begin{aligned}
 Cov(X, Y) &= E([X - E(X)][Y - E(Y)]) \\
 &= E([Y - E(Y)][X - E(X)]) \\
 &= Cov(Y, X)
 \end{aligned}$$

$$\text{Proposition 8} \quad \text{Var} \left( \sum_{i=1}^n w_i X_i \right) = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \text{Cov}(X_i, X_j)$$

**Proof:**

$$\begin{aligned} \text{Var} \left( \sum_{i=1}^n w_i X_i \right) &= \text{E} \left( \left[ \sum_{i=1}^n w_i X_i \right]^2 \right) - \text{E} \left( \sum_{i=1}^n w_i X_i \right)^2 \\ &\quad \text{(by Proposition 1)} \\ &= \text{E} \left( \sum_{i=1}^n \sum_{j=1}^n w_i w_j X_i X_j \right) - \left( \sum_{i=1}^n w_i \text{E}(X_i) \right)^2 \\ &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j \text{E}(X_i X_j) - \sum_{i=1}^n \sum_{j=1}^n w_i w_j \text{E}(X_i) \text{E}(X_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j [\text{E}(X_i X_j) - \text{E}(X_i) \text{E}(X_j)] \\ &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j \text{Cov}(X_i, X_j) \quad \text{(by Proposition 4)} \end{aligned}$$

Proposition 8 can be stated in terms of matrix algebra as follows:

$$\text{Var} \left( \sum_{i=1}^n w_i X_i \right) = w' V w$$

where:

$w$  = column vector of the values  $w_i$  for  $i = 1 \dots n$

$w'$  = the transpose of  $w$ , a row vector

$V$  =  $n \times n$  matrix of the covariances  $\text{Cov}(X_i, X_j)$

**Proposition 9**  $Cov\left(\sum_{i=1}^n w_i X_i, Y\right) = \sum_{i=1}^n w_i Cov(X_i, Y)$

**Proof:**

$$\begin{aligned}
 Cov\left(\sum_{i=1}^n w_i X_i, Y\right) &= E\left(\sum_{i=1}^n w_i X_i Y\right) - E\left(\sum_{i=1}^n w_i X_i\right) E(Y) \\
 &\quad \text{(by Proposition 4)} \\
 &= \sum_{i=1}^n w_i E(X_i Y) - \sum_{i=1}^n w_i E(X_i) E(Y) \\
 &= \sum_{i=1}^n w_i [E(X_i Y) - E(X_i) E(Y)] \\
 &= \sum_{i=1}^n w_i Cov(X_i, Y) \quad \text{(by Proposition 4)}
 \end{aligned}$$

**Lemma 1** Let  $\mu = E(X)$  and  $\sigma = Stdev(X)$ . Define  $\hat{X} = \frac{X - \mu}{\sigma}$ . Then  $E(\hat{X}) = 0$  and  $Var(\hat{X}) = Stdev(\hat{X}) = 1$ .

**Proof:**

$$\begin{aligned}
 E(\hat{X}) &= E\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma}(E(X) - \mu) = \frac{1}{\sigma}(0) = 0 \\
 Var(\hat{X}) &= E\left(\left[\frac{X - \mu}{\sigma} - E\left(\frac{X - \mu}{\sigma}\right)\right]^2\right) \\
 &= E\left(\left[\frac{1}{\sigma}(X - E(X))\right]^2\right) \\
 &= \frac{1}{\sigma^2} E([X - E(X)]^2) \\
 &= \frac{1}{\sigma^2} Var(X) \\
 &= 1
 \end{aligned}$$

**Lemma 2** Let  $\mu_1 = E(X_1)$ ,  $\sigma_1 = \text{Stdev}(X_1)$ ,  $\mu_2 = E(X_2)$ , and  $\sigma_2 = \text{Stdev}(X_2)$

Define  $\hat{X}_1 = \frac{X_1 - \mu_1}{\sigma_1}$  and  $\hat{X}_2 = \frac{X_2 - \mu_2}{\sigma_2}$

Then  $\text{Cor}(X_1, X_2) = \text{Cov}(\hat{X}_1, \hat{X}_2)$ .

**Proof:**

$$\begin{aligned} \text{Cov}(\hat{X}_1, \hat{X}_2) &= \text{Cov}\left(\frac{X_1 - \mu_1}{\sigma_1}, \frac{X_2 - \mu_2}{\sigma_2}\right) \\ &= \frac{1}{\sigma_1 \sigma_2} \text{Cov}(X_1, X_2) \quad (\text{by Proposition 5}) \\ &= \text{Cor}(X_1, X_2) \end{aligned}$$

**Proposition 10**  $|\text{Cor}(X_1, X_2)| \leq 1$

**Proof:**<sup>1</sup>

Define  $\hat{X}_1$  and  $\hat{X}_2$  as in [2].

$$\begin{aligned} 0 &\leq \text{Var}(\hat{X}_1 + \hat{X}_2) \\ &= \text{Var}(\hat{X}_1) + 2\text{Cov}(\hat{X}_1, \hat{X}_2) + \text{Var}(\hat{X}_2) \\ &\quad (\text{by Propositions 6, 7 and 8}) \\ &= 2(1 + \text{Cor}(X_1, X_2)) \quad (\text{by Lemmas 1 and 2}) \\ 0 &\leq 1 + \text{Cor}(X_1, X_2) \\ -1 &\leq \text{Cor}(X_1, X_2) \end{aligned}$$

Similarly,

$$\begin{aligned} 0 &\leq \text{Var}(\hat{X}_1 - \hat{X}_2) \\ &= \text{Var}(\hat{X}_1) - 2\text{Cov}(\hat{X}_1, \hat{X}_2) + \text{Var}(\hat{X}_2) \\ &\quad (\text{by Propositions 5, 6, 7 and 8}) \\ &= 2(1 - \text{Cor}(X_1, X_2)) \quad (\text{by Lemmas 1 and 2}) \\ 0 &\leq 1 - \text{Cor}(X_1, X_2) \\ 1 &\geq \text{Cor}(X_1, X_2) \end{aligned}$$

---

<sup>1</sup>This nifty proof is from Feller [1]

## References

- [1] William Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. John Wiley & Sons, third edition, 1968.