Probability Review

John Norstad

j.norstad@mac.com
http://homepage.mac.com/j.norstad

September 11, 2002
Updated: February 10, 2005

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 *$Var(X)$ = the* variance *of X = $E([X - E(X)]^2)$.*

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

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

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

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

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

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

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

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

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

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

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

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

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

2 Proofs of the Propositions

Proposition 1 $Var(X) = E(X^2) - E(X)^2$

Proof:

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

Proposition 2 $Var(aX+b) = a^2\, Var(X)$

Proof:

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

Proposition 3 $Stdev(aX+b) = |a|Stdev(X)$

Proof:

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

Proposition 4 $Cov(X,Y) = E(XY) - E(X)E(Y)$

Proof:

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

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

Proof:

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

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

Proof:

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

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

Proof:

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

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

Proof:

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

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

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

where:

w = column vector of the values w_i for $i = 1 \ldots n$
w' = the transpose of w, a row vector
$V = n \times n$ matrix of the covariances $\mathrm{Cov}(X_i, X_j)$

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:

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

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

Proof:

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

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

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

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

Proof:

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

Proposition 10 $|Cor(X_1, X_2)| \leq 1$

Proof:[1]

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

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

Similarly,

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

[1] 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.