Expected Value Theory

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

P312, 2013

Expected Value Theory

Introduction

Many textbooks assume and require a knowledge of the basic theoretical results on expected values. Some introductory courses teach this theory, but some sidestep it in a misguided attempt to be user-friendly. Expected value *notation* is a bit cluttered, visually, but the underlying ideas are pretty straightforward.

In this module, we start by reviewing the basic concepts of expected value algebra, and then generalize to matrix expected values.

We hope to give you enough background so that you can negotiate most discussions in standard textbooks on regression and multivariate analysis.

Expected Value of a Random Variable

- The *expected value* of a random variable X, denoted by $\mathrm{E}(X)$, is the long run average (or mean) of the values taken on by that variable.
- As you might expect, one calculates $\mathrm{E}(X)$ differently for discrete and continuous random variables.
- However, in either case, since $\mathrm{E}(X)$ is a mean, it must follow the laws of linear transformation and linear combination!

Algebra of Expected Values

Given constants a,b and random variables X and Y,

1. $\mathrm{E}(a) = a$
2. $\mathrm{E}(aX) = a\,\mathrm{E}(X)$
3. $\mathrm{E}(aX + bY) = a\,\mathrm{E}(X) + b\,\mathrm{E}(Y)$

From the preceding rules, one may directly state other rules, such as

- $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$
- $\mathrm{E}(X - Y) = \mathrm{E}(X) - \mathrm{E}(Y)$
- $\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$
- $$\mathrm{E}\left(\sum_i a_i X_i\right) = \sum_i a_i\,\mathrm{E}(X_i)$$

Variance of a Random Variable

- A random variable X is in deviation score form if and only if $\mathrm{E}(X) = 0$.
- If X is a random variable and has a finite expectation, then $X - E(X)$ is a random variable with an expected value of zero. (*Proof.* C.P. !!)
- The variance of random variable X, denoted $\mathrm{Var}(X)$ or σ_x^2, is the long run average of its squared deviation scores, i.e.

$$\mathrm{Var}(X) = \mathrm{E}\left(X - E(X)\right)^2 \tag{1}$$

- A well-known and useful identity (to be proven by C.P.) is

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \left(\mathrm{E}(X)\right)^2 \tag{2}$$

Covariance of Two Random Variables

For two random variables X and Y, the covariance, denoted $\mathrm{Cov}(X, Y)$ or σ_{xy}, is defined as

$$\mathrm{Cov}(X, Y) = \mathrm{E}\left(X - \mathrm{E}(X)\right)\left(Y - \mathrm{E}(Y)\right) \tag{3}$$

The covariance may be computed via the identity

$$\mathrm{Cov}(X, Y) = \mathrm{E}\left(XY\right) - \mathrm{E}(X)\,\mathrm{E}(Y) \tag{4}$$

Correlation of Two Random Variables

The correlation between two random variables, denoted $\mathrm{Cor}(X, Y)$ or ρ_{xy}, is the expected value of the product of their Z-scores, i.e.

$$\begin{aligned} \mathrm{Cor}(X, Y) &= \mathrm{E}(Z_x Z_y) && (5) \\ &= \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} && (6) \\ &= \frac{\sigma_{xy}}{\sigma_x \sigma_y} && (7) \end{aligned}$$

Joint Distributions

Consider two random variables X and Y. Their joint distribution reflects the probability (or probability density) for a pair of values. For example, if X and Y are discrete random variables, then $f(x, y) = \Pr(X = x \cap Y = y)$.

If X and Y are independent, then $\Pr(X = x \cap Y = y) = \Pr(X = x)\Pr(Y = y)$, and so independence implies $f(x, y) = f(x)f(y)$. Moreover, if X and Y have a joint distribution, random variables like XY generally exist, and also have an expected value. In general, if X and Y are independent, then $E(XY) = E(X)E(Y)$.

Conditional Expectation and Variance

When X and Y are not independent, things are not so simple! In either case, we can talk about the *conditional expectation* $\mathrm{E}(Y|X=x)$, the expected value of the conditional distribution of Y on those occasions when $X=x$.

We can also define the *conditional variance* of $Y|X=x$, i.e., the variance of Y on those occasions when $X=x$.

Introduction

In order to decipher many discussions in multivariate texts, you need to be able to think about the algebra of variances and covariances in the context of random vectors and random matrices.

In this section, we extend our results on linear combinations of variables to *random vector* notation. The generalization is straightforward, and requires only a few adjustments to transfer our previous results.

Random Vectors

- A *random vector* $\boldsymbol{\xi}$ is a vector whose elements are random variables.
- One (informal) way of thinking of a random variable is that it is a process that generates numbers according to some law. An analogous way of thinking of a random vector is that it produces a vector of numbers according to some law.
- In a similar vein, a *random matrix* is a matrix whose elements are random variables.

Expected Value of a Random Vector or Matrix

- The expected value of a random vector (or matrix) is a vector (or matrix) whose elements are the expected values of the individual random variables that are the elements of the random vector.

Example (Expected Value of a Random Vector)

Suppose, for example, we have two random variables x and y, and their expected values are 0 and 2, respectively. If we put these variables into a vector $\boldsymbol{\xi}$, it follows that

$$\mathrm{E}\left(\boldsymbol{\xi}\right) = \mathrm{E}\left(\left[\begin{array}{c} x \\ y \end{array}\right]\right) = \left[\begin{array}{c} \mathrm{E}(x) \\ \mathrm{E}(y) \end{array}\right] = \left[\begin{array}{c} 0 \\ 2 \end{array}\right]$$

Variance-Covariance Matrix of a Random Vector

Given a random vector $\boldsymbol{\xi}$ with expected value $\boldsymbol{\mu}$, the variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\xi}}$ is defined as

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\xi}} &= \mathrm{E}(\boldsymbol{\xi} - \boldsymbol{\mu})(\boldsymbol{\xi} - \boldsymbol{\mu})' \qquad (8) \\ &= \mathrm{E}(\boldsymbol{\xi}\boldsymbol{\xi}') - \boldsymbol{\mu}\boldsymbol{\mu}' \qquad (9) \end{aligned}$$

If $\boldsymbol{\xi}$ is a deviation score random vector, then $\boldsymbol{\mu} = \mathbf{0}$, and

$$\boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\xi}} = \mathrm{E}(\boldsymbol{\xi}\boldsymbol{\xi}')$$

Comment

Let's "concretize" the preceding result a bit by giving an example with just two variables.

Example (Variance-Covariance Matrix)

Suppose

$$\boldsymbol{\xi} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

Note that $\boldsymbol{\xi}$ contains random variables, while $\boldsymbol{\mu}$ contains constants. Computing $\mathrm{E}(\boldsymbol{\xi}\boldsymbol{\xi}')$, we find

$$\begin{aligned} \mathrm{E}\left(\boldsymbol{\xi}\boldsymbol{\xi}'\right) &= \mathrm{E}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\begin{bmatrix} x_1 & x_2 \end{bmatrix}\right) \\ &= \mathrm{E}\left(\begin{bmatrix} x_1^2 & x_1x_2 \\ x_2x_1 & x_2^2 \end{bmatrix}\right) \\ &= \begin{bmatrix} \mathrm{E}(x_1^2) & \mathrm{E}(x_1x_2) \\ \mathrm{E}(x_2x_1) & \mathrm{E}(x_2^2) \end{bmatrix} \qquad (10) \end{aligned}$$

Comment

Example (Variance-Covariance Matrix [ctd.)

] In a similar vein, we find that

$$\begin{aligned}\boldsymbol{\mu}\boldsymbol{\mu}' &= \begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix}\begin{bmatrix}\mu_1 & \mu_2\end{bmatrix} \\ &= \begin{bmatrix}\mu_1^2 & \mu_1\mu_2 \\ \mu_2\mu_1 & \mu_2^2\end{bmatrix}\end{aligned} \tag{11}$$

Subtracting Equation 11 from Equation 10, and recalling that $\mathrm{Cov}(x_i, x_j) = \mathrm{E}(x_i x_j) - \mathrm{E}(x_i)\,\mathrm{E}(x_j)$, we find

$$\begin{aligned}\mathrm{E}(\boldsymbol{\xi}\boldsymbol{\xi}') - \boldsymbol{\mu}\boldsymbol{\mu}' &= \begin{bmatrix}\mathrm{E}(x_1^2) - \mu_1^2 & \mathrm{E}(x_1x_2) - \mu_1\mu_2 \\ \mathrm{E}(x_2x_1) - \mu_2\mu_1 & \mathrm{E}(x_2^2) - \mu_2^2\end{bmatrix} \\ &= \begin{bmatrix}\sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2\end{bmatrix}\end{aligned}$$

Covariance Matrix of Two Random Vectors

Given two random vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, their covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\eta}}$ is defined as

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\eta}} &= \mathrm{E}\left(\boldsymbol{\xi}\boldsymbol{\eta}'\right) - \mathrm{E}(\boldsymbol{\xi})\,\mathrm{E}(\boldsymbol{\eta}') && (12) \\ &= \mathrm{E}\left(\boldsymbol{\xi}\boldsymbol{\eta}'\right) - \mathrm{E}(\boldsymbol{\xi})\,\mathrm{E}(\boldsymbol{\eta})' && (13) \end{aligned}$$

Laws of Matrix Expected Value

Linear combinations on a random vector

- Earlier, we learned how to compute linear combinations of rows or columns of a matrix.
- Since data files usually organize variables in columns, we usually express linear combinations in the form $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B}$.
- When variables are in a random vector, they are in the rows of the vector (i.e., they are the elements of a column vector), so one linear combination is written $y = \boldsymbol{b}'\boldsymbol{x}$, and a set of linear combinations is written $\boldsymbol{y} = \boldsymbol{B}'\boldsymbol{x}$.

Laws of Matrix Expected Value

Expected Value of a Linear Combination

We now present some key results involving the "expected value algebra" of random matrices and vectors.

As a generalization of results we presented in scalar algebra, we find that, for a matrix of constants $\boldsymbol{B}$, and a random vector $\boldsymbol{x}$,

$$\mathrm{E}\left(\boldsymbol{B}'\boldsymbol{x}\right) = \boldsymbol{B}'\,\mathrm{E}(\boldsymbol{x}) = \boldsymbol{B}'\boldsymbol{\mu}$$

For random vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, we find

$$\mathrm{E}\left(\boldsymbol{x} + \boldsymbol{y}\right) = \mathrm{E}(\boldsymbol{x}) + \mathrm{E}(\boldsymbol{y})$$

Comment. The result obviously generalizes to the expected value of the difference of random vectors.

Laws of Matrix Expected Value

Matrix Expected Value Algebra

Some key implications of the preceding two results, which are especially useful for reducing matrix algebra expressions, are the following:

1. The expected value operator distributes over addition and/or subtraction of random vectors and matrices.
2. The parentheses of an expected value operator can be "moved through" multiplied matrices or vectors of constants from both the left and right of any term, until the first random vector or matrix is encountered.
3. $\mathrm{E}(\boldsymbol{x}') = (\mathrm{E}(\boldsymbol{x}))'$
4. For any vector of constants $\boldsymbol{a}$, $\mathrm{E}(\boldsymbol{a}) = (a)$. Of course, the result generalizes to matrices.

An Example

Example (Expected Value Algebra)

As an example of expected value algebra for matrices, we reduce the following expression. Suppose the Greek letters are random vectors with zero expected value, and the matrices contain constants. Then

$$\begin{aligned} \mathrm{E}\left(\boldsymbol{A}'\boldsymbol{B}'\boldsymbol{\eta}\boldsymbol{\xi}'\boldsymbol{C}\right) &= \boldsymbol{A}'\boldsymbol{B}'\,\mathrm{E}\left(\boldsymbol{\eta}\boldsymbol{\xi}'\right)\boldsymbol{C} \\ &= \boldsymbol{A}'\boldsymbol{B}'\boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\xi}}\boldsymbol{C} \end{aligned}$$

Variances and Covariances for Linear Combinations

As a simple generalization of results we proved for sets of scores, we have the following very important results:

Given $\boldsymbol{x}$, a random vector with p variables, having variance-covariance matrix $\boldsymbol{\Sigma_{xx}}$. The variance-covariance matrix of any set of linear combinations $\boldsymbol{y} = \boldsymbol{B}'\boldsymbol{x}$ may be computed as

$$\boldsymbol{\Sigma_{yy}} = \boldsymbol{B}'\boldsymbol{\Sigma_{xx}}\boldsymbol{B} \tag{14}$$

In a similar manner, we may prove the following:

Given $\boldsymbol{x}$ and $\boldsymbol{y}$, two random vectors with p and q variables having covariance matrix $\boldsymbol{\Sigma_{xy}}$. The covariance matrix of any two sets of linear combinations $\boldsymbol{w} = \boldsymbol{B}'\boldsymbol{x}$ and $\boldsymbol{m} = \boldsymbol{C}'\boldsymbol{y}$ may be computed as

$$\boldsymbol{\Sigma_{wm}} = \boldsymbol{B}'\boldsymbol{\Sigma_{xy}}\boldsymbol{C} \tag{15}$$