



# The variance sum law and its implications for modelling

Pier-Olivier Caron <sup>a</sup> and Laura Lemardelet <sup>b</sup>

<sup>a</sup>Université TÉLUQ

<sup>b</sup>Université Laval

**Abstract** ■ Basic understanding of linear model is primordial in quantitative psychology and modelling. As an advanced topic for students in psychology, modelling data according to given sets of parameters can be challenging. In this paper, we address the variance sum law which states how to combine variance of variables into a new one. The paper uses a progressive pedagogical approach, presenting step by step the statistical prerequisites. An example with code in R is offered to support the presentation. Some matrix algebra is gently introduced to stimulate readers to more advanced topics. We hope this paper can stimulate research in data modelling and invite newcomers to the fields.

**Keywords** ■ Statistics, variance, modelling, computation. **Tools** ■ R.

✉ [Pier-Olivier.Caron@teluq.ca](mailto:Pier-Olivier.Caron@teluq.ca)

doi [10.20982/tqmp.17.2.p080](https://doi.org/10.20982/tqmp.17.2.p080)

**Acting Editor** ■ Denis Cousineau (Université d'Ottawa)

**Reviewers**  
■ No reviewer

## Introduction

Basic understanding of linear model and how to appropriately model data according to given sets of parameters of interest can be challenging. Sometimes, modellers fill free parameters (like a regression coefficient or fixing an error variance) while unknowingly neglecting other core parameters (like covariance and correlation). For instance, modellers might fix the regression coefficients and the same error variances for two comparable linear models, but change the variance of variables, thereby altering the magnitude of the regression coefficients. Another example, modeller might fix error variance and regression coefficients, but forget to compute covariance between variables, which decreases the variance of some variable (especially the outcome), thereby increasing the magnitude of the regression parameters. In these scenarios, parameters were fixed by design, but other properties were altered. Many more scenarios may happen (we do not point them out here) and may raise an eyebrow at times in the methodological literature.

In this paper, we address the variance sum law, an advanced topic for students in social and behavioral, and health sciences (and even researchers!), that has many implications in data modelling (for recent example, see Achim, 2020; Beribisky, Constance, & Cribbie, 2020; Caron & Valois, 2018; Caron, Valois, & Gellen-Kamel, 2020; Lorah,

2020, and there is many more). Basic understanding of linear regression is helpful. Some knowledge of matrix algebra is helpful but not mandatory. We hope this paper can be a gentle introduction to the usefulness of matrix algebra for behavioral sciences students. Code in R (R Core Team, 2021) is used to illustrate an example.

## Variance sum law

The variance sum law (Casella & Berger, 2002, p. 171) states how to sum the variance of random variables, regardless of their distribution. The simplest case being two variables, let say  $x_1$  and  $x_2$ , and how they add up to a third variable,  $y$ . Appendix 1 shows a complete derivation of the variance sum law for interested readers. Let's have the simple model being two independents  $x_i$  as:

$$y = x_1 + x_2, \tag{1}$$

then, the variance of their sum takes the following form:

$$\sigma_y^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 \tag{2}$$

where  $\sigma^2$  is the usual symbol for variance. The variance of the sum (or difference) of two independent random variables is their sum. In fact, the sum of  $p$  independent variables is the sum of their variances. In practice, it is rare that independent variables are uncorrelated. Developing the general case where the variables are correlated, if the covariance between  $x_1$  and  $x_2$  is  $\sigma_{x_1x_2}$ , then the variance



sum law is:

$$\sigma_y^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 + 2\sigma_{x_1x_2} \tag{3}$$

from which equation (2) is a special case when  $\sigma_{x_1x_2} = 0$ . Note that the correlation is obtained with  $\rho_{x_1,x_2} = \frac{\sigma_{x_1x_2}}{\sigma_{x_1}\sigma_{x_2}}$  and inversely the covariance is obtained by  $\rho_{x_1x_2}\sigma_{x_1}\sigma_{x_2} = \sigma_{x_1x_2}$  (Howell, 2012).

Some mathematically oriented readers might notice the relation between the variance sum law depicted in equation (2) and the Pythagoras theorem. Consequently, we can envision two uncorrelated variables to be represented by a right-angled triangle. Equation (3) is the cosine rule, or the generalisation of the Pythagoras theorem for non-right-angled triangles, so we can see the correlation as an angle (Rogers & Nicewander, 1988).

Another more intuitive fashion to compute the variance sum law is to see that the variance of the sum of  $p$  variables is the “grand sum” of their covariance matrix. The grand sum is an informal mathematical function who refers to the sum of all the element in a matrix. Let  $\Sigma$  be the variance-covariance matrix of two random variables  $x$  (remember that diagonal elements are variances), then

$$\begin{aligned} \sigma_y^2 &= \text{grand sum}(\Sigma) \\ &= \text{grand sum} \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1x_2} \\ \sigma_{x_1x_2} & \sigma_{x_2}^2 \end{bmatrix} \end{aligned} \tag{4}$$

which, if we sum it all up, lead to equation (3), and again, if  $\sigma_{x_1x_2} = 0$ , then equation (4) is equal to equation (2). This formulation has the advantage to show the origin of the two covariances in equation (3).

Even though the matrix notation might not seem appealing at first, it becomes much more interesting when the number of variables,  $p$ , increases, as there is  $p(p-1)/2$  off-diagonal elements to add to  $p$  variances.

In the  $p = 3$  example, equation (4) would be:

$$\begin{aligned} \sigma_y^2 &= \text{grand sum}(\Sigma) \\ &= \text{grand sum} \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1x_2} & \sigma_{x_1x_3} \\ \sigma_{x_1x_2} & \sigma_{x_2}^2 & \sigma_{x_2x_3} \\ \sigma_{x_1x_3} & \sigma_{x_2x_3} & \sigma_{x_3}^2 \end{bmatrix} \end{aligned} \tag{5}$$

In the form of equation (3), it is:

$$\sigma_y^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 + \sigma_{x_3}^2 + 2\sigma_{x_1x_2} + 2\sigma_{x_1x_3} + 2\sigma_{x_2x_3} \tag{6}$$

and getting longer with increasing  $p$ .

A matrix operation equivalent to the grand sum is  $\mathbf{1}'\Sigma\mathbf{1}$ , where  $\mathbf{1}$  is a vector of length  $p$  containing only 1. Doing this operation, yield the sum of all the elements in  $\Sigma$ . This will be useful to derive a more general case in the next section.

### Adding scaling constants $\beta$

Equation (3), (4), (5) and (6) are special cases of a more general law. They would not work if scaling constants were added (that will be later regression coefficients) or to compute the difference (a type of scaling also).

Consider the fact that the variable  $y$  is the product of a constant  $\beta$  and the variable  $x$  as:

$$y = \beta x \tag{7}$$

(a linear model without error). It might be useful to consider  $\beta$  the degree to which two variances are related but also as a pure standard deviation modifier (a scaling constant). The same way a random variable with mean of 0 and standard deviation of 1,  $x \sim \mathcal{N}(0, 1)$ , multiplied by value  $\delta$ , an arbitrary scaling factor, becomes distributed like  $\delta x \sim \mathcal{N}(0, \delta)$ . So,  $\delta$  modified (or scaled) the standard deviation of the distribution. A related and frequently encountered case of this is when data are standardized as  $z$ -score or unstandardized (divided or multiplied respectively by  $\sigma$ ). At this point, the contribution is a standard deviation, so  $\beta$  has to be squared to yield the variance of  $y$ , thus  $\beta^2$ . To compute the variance of  $y$ , the equation is

$$\sigma_y^2 = \beta^2 \sigma_x^2. \tag{8}$$

Now, let us go further with a model with two scaling constants. The linear model of the form:

$$y = \beta_1 x_1 + \beta_2 x_2 \tag{9}$$

which is the same model as in equation (1) where we added the scaling factor  $\beta_i$ . To consider scaling constants, the variance sum law becomes for two independent variables, based on equation (2):

$$\sigma_y^2 = \beta_1^2 \sigma_{x_1}^2 + \beta_2^2 \sigma_{x_2}^2 \tag{10}$$

and when they covary, like equation (3), it becomes:

$$\sigma_y^2 = \beta_1^2 \sigma_{x_1}^2 + \beta_2^2 \sigma_{x_2}^2 + 2\beta_1\beta_2\sigma_{x_1x_2} \tag{11}$$

In these equations, we see that  $\beta_i$  are scaling the variance and the covariance. Like previously, the matrix algebra form is simpler and more elegant:

$$\sigma_y^2 = \mathbf{B}'\Sigma\mathbf{B} \tag{12}$$

where  $\mathbf{B}$  is a vector containing all the regression coefficient  $\beta_i$  of length  $p$  predicting  $y$  and apply for any number of predictors  $x$ . The prime ' is the transpose symbol, an operator which flips a matrix over its diagonal, and in the case of vector, flip the columns into rows and vice-versa (see its implication in equation (13) for example). In equation (12), using the two vectors  $\mathbf{B}$  twice act the same as



squaring  $\beta_i$ . Notice that if  $\mathbf{B}$  are all 1, then equation (12) is the same as the definition of a grand sum. Finally, it is worthy to note that equation (12) is the same as the coefficients of determination,  $R^2$  (Cohen, Cohen, West, & Aiken, 2003), when all variables and regressions coefficients are standardized.

In the simple linear model depicted in equation (9), equation (11) becomes:

$$\sigma_y^2 = \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1x_2} \\ \sigma_{x_1x_2} & \sigma_{x_2}^2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \quad (13)$$

which is equivalent to equation (11). At this stage, some might have had the intuition that equations (1) to (6) were only a special case where all  $\beta = 1$ .

The variance of the difference of two variables is equivalent to stating that the subtracted variables'  $\beta$  is the reverse sign, i.e.,  $-\beta$ , which for the model lead to

$$y = \beta_1x_1 - \beta_2x_2 \quad (14)$$

or, equivalently,

$$y = (\beta_1)x_1 + (-\beta_2)x_2 \quad (15)$$

yield

$$\begin{aligned} \sigma_y^2 &= \begin{bmatrix} \beta_1 & -\beta_2 \end{bmatrix} \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1x_2} \\ \sigma_{x_1x_2} & \sigma_{x_2}^2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ -\beta_2 \end{bmatrix} \\ &= (\beta_1)^2\sigma_{x_1}^2 + (-\beta_2)^2\sigma_{x_2}^2 + 2(\beta_1)(-\beta_2)\sigma_{x_1x_2} \\ &= \beta_1^2\sigma_{x_1}^2 + \beta_2^2\sigma_{x_2}^2 - 2\beta_1\beta_2\sigma_{x_1x_2} \end{aligned} \quad (16)$$

so the variance of the difference of two random variables is the sum of their variances subtracting twice their covariance scale by  $\beta$ , as expected.

### Implications for modelling

The variance sum law has many implications in computational modelling, especially if data are shaped according to certain desirable characteristics, like in linear models. Linear regression is an approach to model additive effects (independent variables,  $x_i$ ) to predict a dependent variable ( $y$ ). Linearity refers to the property of a function of being compatible with addition and scaling. As such, there is a direct relation between equations above and the linear model:

$$y = \beta_1x_1 + \dots + \beta_px_p + \varepsilon \quad (17)$$

omitting the constant  $\beta_0$ , which plays no role in the variance of the dependent variables and adding the error, a particular independent variable which is assumed to be unrelated to  $x_i$ , have 0 mean, a standard deviation  $\sigma$ . Equation (17) is the general form of linear model in equation (1) and (8) in which the variance of  $y$  is a function of the variances-covariances of  $x$  weighted by  $\beta$ .

### Computing the error variance

Taking the most basic linear model:

$$y = x + \varepsilon \quad (18)$$

where  $\varepsilon$  is the residual error term. We have at this stage the very basic to build a bivariate model. It is the same bivariate sum as equation (1), but  $x_2$  is replaced and defined as independent of  $x_1$ . Based on this model, equation (2) leads to the following equation,

$$\sigma_y^2 = \sigma_x^2 + \sigma_\varepsilon^2 \quad (19)$$

In most cases in data modelling, the variances  $\sigma_x^2$  and  $\sigma_y^2$  are known rather than  $\sigma_\varepsilon$  so computing a priori the error variance instead of varying the parameters of interest is more relevant. Rearranging equation (18) to isolate  $\varepsilon$ , we get

$$\varepsilon = y - x \quad (20)$$

and in terms of variance:

$$\sigma_\varepsilon^2 = \sigma_y^2 + \sigma_x^2 - 2\sigma_{xy} \quad (21)$$

or the equation (16) for the form in matrix algebra:

$$\sigma_\varepsilon^2 = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} \sigma_y^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_x^2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \sigma_y^2 + \sigma_x^2 - 2\sigma_{xy}. \quad (22)$$

### Error variance with $\beta$

As we have seen, one concern during data modelling it to preserve the desired properties of the model in the data sets like variance of error, regression parameters, covariances, etc. After the variance of the error term, the last element to consider is the regression coefficients,  $\beta_i$ , in other words, the size of the relation between the independent and dependent variables. From equation (18), we add the slope the error:

$$y = \beta x + \varepsilon \quad (23)$$

The variance sum law for this model becomes

$$\sigma_y^2 = \beta^2\sigma_x^2 + \sigma_\varepsilon^2 \quad (24)$$

Remember the error is not correlated to the independent variable  $x_i$ . As before, the matrix algebra form is simpler and more elegant:

$$\sigma_y^2 = \mathbf{B}'\Sigma\mathbf{B} + \sigma_\varepsilon^2 \quad (25)$$

where  $\mathbf{B}$  is a vector containing all the regression coefficient  $\beta$  predicting  $y$  and apply for any number of predictors  $x$ . Taking advantage of independence of the error term, we can rearrange the equation as:

$$\sigma_\varepsilon^2 = \sigma_y^2 - \mathbf{B}'\Sigma\mathbf{B} \quad (26)$$



which gives the variance of the error term. In a linear form (for a single independent variable), we can also rearrange equation (24) to isolate  $\sigma_\varepsilon$

$$\sigma_\varepsilon^2 = \sigma_y^2 - \beta^2 \sigma_x^2 \tag{27}$$

Remember that  $\beta = \frac{\sigma_{xy}}{\sigma_x^2}$  and, as we mentioned earlier, the covariance  $\sigma_{xy}$ , is equal to  $\rho\sigma_x\sigma_y$ .

**The standardized case**

To conclude this section, we look at the standardize case, that is, when variances of variable are equal to 1. Specifically, the variance of every variable is fixed; only residual errors have to be adjusted to maintain these properties. In this scenario, the variance-covariance matrix,  $\Sigma$ , is a correlation matrix, which defines the residual error variances. To compute variables' error variance, we can use equation (26) and (21) replacing variance of variables by 1 as

$$\sigma_\varepsilon^2 = 1 - \mathbf{B}'\Sigma\mathbf{B} \tag{28}$$

or, for a single predictor,

$$\sigma_\varepsilon^2 = 1 - \beta^2 \sigma_x^2 \tag{29}$$

**Numerical examples in R**

Listing 1 illustrates some of the more relevant formula and confirmed what has been presented. The code begins by fixing an arbitrarily large sample size,  $n$ , and setting the seed for reproducibility. It then produce a variable  $x$  with standard deviation of  $\sigma_x = 3$ , which can be verified with the function `sd()` in R. Lines 13-18 depicts then equations (5), (8), for  $\beta = 4$ , i.e.,

$$\begin{aligned} \beta &= 4 \\ y &= \beta x \\ \sigma_y^2 &= \beta^2 \sigma_x^2 = (4)^2 (3)^2 = 144 \end{aligned}$$

After, lines 25-35 is used to generate a model with  $\rho_{xy} = .20$ , and as previously,  $\sigma_x^2 = 3^2$ ,  $\sigma_y^2 = 12^2$ . We can a priori compute the error variance by computing firstly the related  $\beta$  (an intuitive explanation is shown in Appendix 2) as

$$\beta = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\rho\sigma_x\sigma_y}{\sigma_x^2} = \frac{\rho\sigma_y}{\sigma_x} = \frac{.2 * 12}{3} = .80$$

and from equation (27) we get

$$\sigma_\varepsilon^2 = \sigma_y^2 - \beta^2 \sigma_x^2 = 12^2 - 0.8^2 \times 3^2 = 138.24$$

or in standard deviation, 11.75755. Lines 36-41 in Listing 1 then proceed to confirm the result by modelling,  $x$ ,  $\varepsilon$  and  $y$ .

Finally, lines 44-50 in Listing 1 shows a standardized scenario (where means are 0 and variance are 1) with

$\beta_{x_1,x_2} = 0.40$ ,  $\beta_{x_1,x_3} = 0.50$ , and  $\beta_{x_2,x_3} = 0.30$ , much like a mediation model, where error variances are computed with equation (27):

$$\sigma_{\varepsilon_{x_2}}^2 = \sigma_y^2 - \beta_{x_1x_2}^2 \sigma_{x_1}^2 = 1 - 0.40^2 \times 1 = 0.84$$

and

$$\begin{aligned} \sigma_{\varepsilon_{x_3}}^2 &= 1 - (\beta_{x_1x_3}^2 \sigma_{x_1}^2 + \beta_{x_2x_3}^2 \sigma_{x_2}^2 + 2\beta_{x_1x_3}\beta_{x_2x_3}\beta_{x_1x_2}) \\ &= 1 - (0.50^2 * 0.30^2 + 2 * 0.50 * 0.30 * 0.40) = 0.54, \end{aligned}$$

and in matrix algebra, equation (26):

$$\sigma_{\varepsilon_y}^2 = 1 - \begin{bmatrix} .50 & .30 \end{bmatrix} \begin{bmatrix} 1 & 0.40 \\ 0.40 & 1 \end{bmatrix} \begin{bmatrix} 0.50 \\ 0.30 \end{bmatrix} = 0.54$$

**Conclusion**

The purpose of the current paper is to describe the variance sum law and how to use it to model data according to some parameters defined by the modellers. As an advanced topic for most students in psychology, basic understanding on how to build these models can be challenging. We hope that by using a pedagogical approach, as well as supporting examples in R, more inclined toward students and researchers of behavioral, health and educational sciences, we can stimulate research in data modelling from a more applied-orientated perspective and invite newcomers to the fields.

**References**

Achim, A. (2020). Esprit et enjeux de l'analyse factorielle exploratoire. *The Quantitative Methods for Psychology*, 16(4), 213–247. doi:10.20982/tqmp.16.4.p213

Beribisky, N., Constance, A. M., & Cribbie, R. A. (2020). An equivalence testing approach for evaluating substantial mediation. *The Quantitative Methods for Psychology*, 16(4), 424–441. doi:10.20982/tqmp.16.4.p424

Caron, P.-O., & Valois, P. (2018). A computational description of simple mediation analysis. *The Quantitative Methods for Psychology*, 14(2), 147–158. doi:10.20982/tqmp.14.2.p147

Caron, P.-O., Valois, P., & Gellen-Kamel, A. (2020). Some computational descriptions of moderation analysis. *The Quantitative Methods for Psychology*, 16(1), 9–20. doi:10.20982/tqmp.16.1.p009

Casella, G., & Berger, R. L. (2002). *Statistical inference*. Pacific Grove: Duxbury.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences (iii ed.)* Mahwah: Erlbaum.

Howell, D. C. (2012). *Statistical methods for psychology*. Belmont: Wadsworth.



Lorah, J. A. (2020). Interpretation of main effects in the presence of non-significant interaction effects. *The Quantitative Methods for Psychology*, 16(1), 33–45. doi:10.20982/tqmp.16.1.p033

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: In R Foundation

for Statistical Computing. Retrieved from <http://www.Rproject.org/>

Rogers, J. L., & Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42(1), 59–66.

### Appendix 1

Here, we derive the variance sum law. We use a slightly different notation than the main article to simplify the presentation. If

$$z = ax + by$$

where  $x$  and  $y$  are the predictors,  $a$  and  $b$  are scaling constants, then

$$\sigma_z^2 = (n - 1)^{-1} \mathbb{E}((ax + by) - (a\mu_x + b\mu_y))^2$$

in which  $\mu_i$  is the mean of  $i$ . We continue as

$$\begin{aligned} \sigma_z^2 &= (n - 1)^{-1} \mathbb{E}(a(x - \mu_x) + b(y - \mu_y))^2 \\ &= (n - 1)^{-1} \mathbb{E}(a^2(x - \mu_x)^2 + b^2(y - \mu_y)^2 + 2ab(x - \mu_x)(y - \mu_y)) \\ &= a^2 \mathbb{E}(x - \mu_x)^2 (n - 1)^{-1} + b^2 \mathbb{E}(y - \mu_y)^2 (n - 1)^{-1} + 2ab \mathbb{E}(x - \mu_x)(y - \mu_y) (n - 1)^{-1} \end{aligned}$$

where we see that  $\mathbb{E}(x - \mu_x)^2 / (n - 1)$  and  $\mathbb{E}(y - \mu_y)^2 / (n - 1)$  are variances,  $\sigma_x^2$  and  $\sigma_y^2$  respectively, and  $\mathbb{E}(x - \mu_x)(y - \mu_y) / (n - 1) = \text{cov}(x, y) = \sigma_{xy}$ , then we end up with :

$$\sigma_z^2 = a^2 \sigma_x^2 + b^2 \sigma_y^2 + 2ab \sigma_{x,y}$$

### Appendix 2

Here we give an intuitive explanation on the case of a single regression coefficient,  $\beta$ . In the model,

$$y = \beta x$$

We can isolate  $\beta$

$$\beta = \frac{\bar{y}}{\bar{x}}$$

where the symbol  $\bar{\phantom{x}}$  denote the average of the variables. Multiplying  $\frac{\bar{x}}{\bar{x}} = 1$  on both side lead to

$$\beta = \frac{\overline{xy}}{\overline{x^2}}$$

If we center  $x$  and  $y$ , then we get

$$\beta = \frac{\sigma_{xy}}{\sigma_x^2}$$

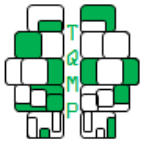
The above equations are the same as in the matrix notation to

$$\mathbf{B} = [\mathbf{X}'\mathbf{X}]^{-1} \mathbf{X}'\mathbf{y}$$

which computes regression coefficients as to minimize the error (omitted herein), the ordinary least square estimators.

**Listing 1. Illustration in R of the variance sum law.**

```
1 #We take an arbitrarily large sample size
2 #to ensure accuracy in estimates.
3 n = 100000000
4 #For reproducibility, we set a given seed.
5 set.seed(1234)
6 #rnorm generate n random values with mean mu (default = 0)
7 #and standard deviation sd = 3.
8 X = rnorm(n, sd=3)
9 #To compute variance, use the function var()
10 # or sd() for standard deviation.
11 var(X)
12 sd(X)
13 #We can check equation (5) (linear model) and equation (8)
14 # (variance sum law) for beta = 4
15 beta = 4
16 Y = beta * X
17 var(Y)
18 #which yield to 4^2*3^2=144 (or 12^2) as expected.
19
20 #Now, let say we use equation (23) to build a model with
21 #correlation (rho) of .2,
22 #var_x = 3^2 and var_y = 12^2. What would be the var_e of the error term?
23 #Let use equation (27) and excluding the last two terms,
24 #because we only have a single IV, we can compute
25 rho=.2
26 sd_x=3
27 sd_y=12
28 cov_xy=rho*sd_x*sd_y
29 beta = cov_xy/sd_x^2
30 beta
31 #equation (27):
32 var_e = sd_y^2 - beta^2*sd_x^2
33 #which we can verified with
34 e = rnorm(n, sd = sqrt(var_e))
35 Y = beta*X + e
36 var(X)
37 #9
38 var(Y)
39 #144
40 var(e)
41 #138.24
42 round(cov(X,Y),3)
43 round(cor(X,Y),3)
44 #Standardized case with b_x1x2 = .4; b_x2x3 = .3; b_x1x3 = .5.
45 b_x1x2 = .4; b_x2x3 = .3; b_x1x3 = .5;
46 x1 = rnorm(n)
47 x2 = b_x1x2 * x1 + rnorm(n, sd = sqrt(1-b_x1x2^2))
48 #Variance of the error in x3
49 var_ex3 = 1-(b_x1x3^2+b_x2x3^2+2*b_x1x3*b_x2x3*b_x1x2)
50 x3 = b_x1x3 * x1 + b_x2x3 * x2 + rnorm(n, sd = sqrt(var_ex3))
```



```
51
52 #Seeing the extension of equation (8) and (27), the
53 #covariance between x1 and x2 is their beta.
54 #We can also carry the above formula with matrix algebra
55 #The beta matrix:
56 B = matrix(c(b_x1x3, b_x2x3), 2, 1)
57 #The variance-covariance matrix
58 S = matrix(c(1, b_x1x2, v_x1x2, 1), 2, 2)
59
60 #Variance of the error in x3 in matrix algebra
61 #see equation (26)
62 var_ex3_m = 1 - (t(B) %*% S %*% B)
63 X3 = b_x1x3 * x1 + b_x2x3 * x2 + rnorm(n, sd = sqrt(var_ex3_m))
64 #To verify we can check
65 var_ex3
66 var_ex3_m
67 cor(cbind(x1, x2, x3))
68 lm(x2~x1)
69 lm(x3~x2+x1)
```

### Citation

Caron, P.-O., & Lemardelet, L. (2021). The variance sum law and its implications for modelling. *The Quantitative Methods for Psychology*, 17(2), 80–86. doi:[10.20982/tqmp.17.2.p080](https://doi.org/10.20982/tqmp.17.2.p080)

Copyright © 2021, Caron and Lemardelet. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 10/04/2021 ~ Accepted: 30/04/2021