



The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation

Variance

Contributors: Dorothy J. Musselwhite & Brian C. Wesolowski

Edited by: Bruce B. Frey

Book Title: The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation

Chapter Title: "Variance"

Pub. Date: 2018

Access Date: February 26, 2018

Publishing Company: SAGE Publications, Inc.

City: Thousand Oaks,

Print ISBN: 9781506326153

Online ISBN: 9781506326139

DOI: <http://dx.doi.org/10.4135/9781506326139.n737>

Print pages: 1790-1793

©2018 SAGE Publications, Inc.. All Rights Reserved.

This PDF has been generated from SAGE Knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

Variability is a numerical description that refers to the spread values within a given distribution. Variability can be described using four common measures: range, the mean deviation score, standard deviation, and variance. Variance is most useful in determining the overall spread of a data set. Range simply takes two data points, the lowest and highest values, and measures the space between them. The interquartile range also interprets the amount of spread within a set of data by using only the middle 50% of scores. Unlike range, variance takes into account every data point within the set and measures each distance from the mean. Variance is calculated using a number of values from data and the associated distribution. Specifically, variance requires raw data scores and the sample size of the data.

The sum of squares (or sum of squared errors) is also used as a method to determine total spread or dispersion. The problem with sum of squares is this value cannot be compared across samples that differ in size. Variance deals with average spread, which is comparable across groups that may change sizes.

Variances are most impactful when comparing multiple distributions. Furthermore, this statistic becomes the basis for various statistical comparisons, including an analysis of variance (ANOVA). In experiments, people or groups undergo treatments to potentially elicit various responses. If all scores, or responses, are different, variability will be large. If all scores are exactly the same, variability will be zero.

Variance is used in many different experimental designs. After discussing a brief history, variance is described not only by its formulae but also by its conceptual properties. Variance is then applied to statistical testing, including ANOVA and multiple regression.

History

Although credit for the concept of variance is given to Ronald Fisher, it is apparent that the concept of variance existed long before Fisher, in the work of Carl Friedrich Gauss and his endeavor to estimate the locations of stars. Within his search, Gauss encountered a probability distribution with deviations that may have given way to the current concept of variance.

Fisher introduced the concept of variance in his 1918 paper "The Correlation Between Relatives on the Supposition of Mendelian Inheritance." Fisher was the first to introduce the test now known as ANOVA. Most of his later work involved significance and hypothesis testing.

Formulae

When research designs are considered, it is the goal of the researcher to collect as many real-world observations as possible, knowing that the collected observations will not account for all possible observations that exist. As mean and variance are calculated from these select observations, it can be assumed that the values will not perfectly match the values that would have been obtained using every possible observation. Therefore, the researcher must estimate the mean and variance using an equation to account for observational bias. Two formulae exist to calculate variance, one describing the population variance and the other describing the sample variance.

Population variance is calculated using the mean of the squared deviations from the

distribution mean. Deviation is the difference between a specific value of data and the distribution's mean. The formula for population variance can be defined as

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N},$$

where σ^2 represents the population variance, $X - \mu$ represents the deviation between each score and the population mean, and N is the population size. This formula is conceptual and represents a hypothetical population. A population refers to the entire group being defined. For example, if the experiment involves 20-year-old females, the population would be every 20-year-old female on the planet. Because the population in question is not always readily available, the researcher must obtain a sample or a proportion of the population. This sample is meant to provide a snapshot of what the population actually looks like.

If a population is extremely large, specifically when it is not possible to count every observation available, a sample of the population must be used to compute variance. This formula for sample variance is used most frequently and can be defined as:

$$S^2 = \frac{\sum (X - \bar{X})^2}{n - 1},$$

where S^2 is the variance, $X - \bar{X}$ represents the deviation between each score and the sample mean, and n is the sample size. The numerator is often referred to as the sum of squared errors, or *SS*. Using $n - 1$ as the denominator for sample variance is called Bessel's correction. The result is an unbiased estimator of the population variance called the corrected sample variance or unbiased sample variance. By taking the sum of the deviations and dividing by Bessel's correction, what results is an average of all deviations. The sum of the deviations of any data set will always be zero; therefore, the deviations must be squared before being summed to account for negative values. If the distribution mean is not easily attainable, another formula may be used while employing raw data:

$$S^2 = \frac{\sum X^2}{N} - \frac{(\sum X)^2}{N^2}.$$

Both formulas yield identical answers but are dependent on the data and statistics available.

Once variance is calculated, variability can be further explored through another statistic, standard deviation. Standard deviation is another method of determining spread, specifically focusing on the data and its location from the mean. Variance is more about the average distance away from the mean and is reported in square units. Standard deviation can be calculated by taking the square root of the variance:

$$S = \sqrt{S^2}.$$

By taking the square root of the variance, the units are also restored to their original form. Standard deviation is more appropriate when units do not make sense to be squared (e.g., people).

Properties of Variance

Variance will always yield a nonnegative value because squares are positives or zero. If a constant is applied to all values of a variable through addition, variance will remain unchanged. In other words, variance is invariant as it relates to the location parameter.

However, if a constant is applied through multiplication, or scaling, variance will adjust by the square of that constant.

Variance is often the preferred method of determining the spread of a given data set due to the fact that the variance of the sum of the uncorrelated (independent) random variables can be calculated using the sum of the variances. The Bienaymé formula is a derivation of the aforementioned concept. This derived formula was discovered in 1853 and is defined as:

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n},$$

where \bar{X} refers to the sample mean. The Bienaymé formula can be interpreted to indicate that the mean in every experiment with a sample size of n will fall within $\pm\sigma/n$ of the true population 68% of the time. The 68% refers to the area under the normal curve within ± 1 standard deviation from the mean. This formula also shows that the variance of the mean will decrease when the sample size, n , increases when the variables have the same variance.

Correlated variables are treated differently in terms of the variance of their sum. Instead of adding the variances as in uncorrelated variables, the variance of the sum of correlated variables is equal to the sum of their covariances. Because the variables are now related, the covariance must be used to determine how the variables change together. Variance of the mean is calculated differently when the variables are correlated and have equal variance:

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n} + \frac{n-1}{n}\rho\sigma^2,$$

where ρ refers to the average correlation of distinct variables.

The variance sum law states that the variance of a difference between two independent variables can be calculated using the sum of their respective variances. For example, a test is given to fifth graders to determine level of mathematics ability based on gender. Therefore, boys and girls are split into two subgroups for testing. The two populations are fifth grade boys and fifth grade girls. For each group, the variance is calculated separately. When the two variables are independent, the variance can be added together to show the variance of the sum. The standard error of that variance can be calculated by taking the square root of the aforementioned sum.

Variance should always be reported in units squared. For example, if data points are measured in centimeters (cm), variance will be reported as cm^2 . This reporting can be problematic, depending on the units. It may be reasonable to report cm^2 or kg^2 , but, as noted earlier, when the units are people, it does not make sense to report people. When this is the case, standard deviation is often reported instead, by taking the square root of the variance and defining by the same units as the original variable.

Between Groups and Within Groups Variance

Within one set of data exists two types of variance: between-groups variance and within-groups variance. These types of variance are most relevant when conducting an ANOVA, specifically when multiple groups are being compared. Between-groups variance is frequently referred to as explained or systematic variance. Systematic variance is derived from the independent variable or among the groups in the study. Explained variance is accompanied by a direct manipulation from the researcher. Within-groups variance refers specifically to the

error variance and comes from within each group. It does not come from the independent variable. This unsystematic variance is due to variations in performance, by factors outside of the manipulation directly from the researcher. This type of variance is not consistent and may fluctuate from one testing administration to the next. Comparing examinees in the same group would result in within-groups variance. Comparing examinees in two different groups would result in between-groups variance. Both of these variances will impact the validity and reliability of tests.

Homogeneity of Variance

The assumption of homogeneity of variance, or homoscedasticity, is used when several groups of examinees are tested for differences in group means. In such a design, it is assumed that all samples come from populations of equal variance. This assumption may also be applicable in a correlational design, where the variance of the dependent variable should be equal at all levels of the independent variable. Any fluctuation of the variance would be a violation of the assumption of homogeneity of variance.

If the spread around the scores are roughly the same from each sample taken, the assumption has been satisfied. If the spread fluctuates greatly from sample to sample, this assumption has been violated and is called heterogeneity of variance or heteroscedasticity. The violation of this assumption is most important when group sizes are unequal.

Homogeneity of variance is most important when using the method of least squares to estimate parameters. Unequal variances will create a bias and an inconsistent estimate of the standard error. Inconsistent standard error will directly affect confidence intervals and significance tests. Levene's test will test the null hypothesis, stating different groups have equal variances. This test is executed through a one-way ANOVA, where a significant p value indicates the variances among groups are significantly different. This significant p value from Levene's test, therefore, indicates a violation of the assumption of homogeneity of variance. However, a large sample size may result in an incorrect reporting of significance. Small differences among group variances when the sample size is large will generate a significant Levene's test. It is important to note that Levene's test works best with equal group sizes.

Another way to compare groups is to use the variance ratio or Hartley's F_{\max} . Hartley's F_{\max} is equal to the ratio of the largest group variance to the smallest group variance. This value is then compared to critical values that are dependent on the group size and the number of variances being evaluated. The variance ratio is expected to be smaller than the critical values specified by Hartley's table to be nonsignificant.

Homogeneity of variance should be evaluated by both Levene's test and the variance ratio, as false significance may be reported by one test.

Variance in Multiple Regression

In a multiple regression design, fit of the regression line is measured by correlation (r) and r^2 . In relation to variance, however, R^2 is often reported specifically in reference to explained variance. R^2 , often called the coefficient of determination, may be reported as a decimal ($R^2 = .71$) or a percent. When reported as a percentage, it is often stated as the percentage of variance in the independent variable as explained by the model. This value does not make the

implication that causation occurs. Using the output from statistical analysis software, or from hand calculations, R^2 can be calculated using the following formula:

$$R^2 = \frac{SS_M}{SS_T},$$

where SS_M refers to the model sum of squares and SS_T refers to the total sum of squares. The square root of this value yields Pearson's correlation coefficient, r . Adjusted R^2 will often be reported in addition to R^2 , which removes the bias associated with the latter by diminishing its value.

See also [Analysis of Variance](#); [Correlation](#); [Multiple Linear Regression](#); [Normal Distribution](#); [Standard Deviation](#); [Standard Error of Measurement](#)

Dorothy J. Musselwhite Brian C. Wesolowski
<http://dx.doi.org/10.4135/9781506326139.n737>
10.4135/9781506326139.n737

Further Readings

- Boyle, J. D., & Radocy, R. E. (1987). *Measurement and evaluation of musical experiences*. New York, NY: Schirmer Books.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London, UK: Sage.
- Gravetter, F. J., & Wallnau, L. B. (2011). *Essentials of statistics for the behavioral sciences*. Belmont, CA: Wadsworth.
- Kubiszyn, T., & Borich, G. (2003). *Educational testing and measurement: Classroom application and practice*. New York, NY: Wiley.
- Payne, D. A. (2003). *Applied educational assessment*. Toronto, Canada: Wadsworth.
- Reid, H. M. (2014). *Introduction to statistics: Fundamental concepts and procedures of data analysis*. Thousand Oaks, CA: Sage.