

How to Calculate Standard Deviation of Frequency Distribution in Excel?

Authored by
stats writer

November 18, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Calculate Standard Deviation of Frequency Distribution in Excel?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=96108>

This comprehensive guide provides an expert walkthrough on calculating the Standard Deviation of a Frequency Distribution specifically within the environment of Microsoft Excel. We detail the foundational concepts of statistical dispersion, explain the necessity of using midpoints for grouped data, and provide precise, step-by-step instructions for manipulating the formulas in Excel to arrive at an accurate estimate of variance.

Mastering this technique is vital for anyone analyzing summarized datasets where individual data points are aggregated into class intervals. By the conclusion of this tutorial, you will possess the requisite knowledge to reliably calculate the standard deviation for any frequency distribution, thereby enhancing your ability to interpret and work effectively with statistical data.

Calculating the standard deviation for grouped data is a frequent requirement in statistical applications, especially when dealing with large datasets summarized into class intervals. This specialized approach ensures that statistical measures of spread remain accurate even when the raw data is unavailable.

To illustrate the methodology, we will use a concrete example of a frequency distribution. The following steps demonstrate exactly how to leverage Excel's capabilities to perform this complex calculation, focusing on clarity and accuracy at every stage.

For example, suppose you have the following frequency distribution:

	A	B	C	D
1	Lower Class Limit	Upper Class Limit	Frequency	
2	1	10	2	
3	11	20	7	
4	21	30	10	
5	31	40	3	
6	41	50	1	
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				

The following step-by-step example shows precisely how to calculate the standard deviation of this frequency distribution in Excel, preparing us for detailed analysis of the steps involved in estimating statistical measures for grouped information.

Setting Up the Data in Microsoft Excel

The initial and most crucial step involves accurately inputting the data from the frequency distribution table into the Microsoft Excel spreadsheet. We must designate separate columns for the class boundaries--the lower limit, the upper limit--and the corresponding frequencies. This structured approach facilitates easier formula application and review later in the process.

We begin by entering the class limits and the frequency values for our distribution into three adjacent columns. For this example, let's assume the lower limits are in Column A, the upper limits in Column B, and the frequency counts (f_i) are in Column C, starting from row 2. This setup directly translates the grouped data structure into a computational framework.

Ensuring the accuracy of this data entry is paramount, as subsequent statistical estimates are entirely dependent on these initial values. This step transforms the static data representation into a dynamic dataset ready for calculation of the midpoint and the weighted products necessary for determining the mean and variance.

First, we'll enter the class limits and frequency values for our frequency distribution:

	A	B	C	D
1	Lower Class Limit	Upper Class Limit	Frequency	
2	1	10	2	
3	11	20	7	
4	21	30	10	
5	31	40	3	
6	41	50	1	
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				

Theoretical Basis for Estimating the Mean

Before we can calculate the dispersion, we must establish the estimated Mean (μ) of the distribution. Since we are dealing with grouped data, we cannot calculate the true mean of the original observations. Instead, we estimate it by assuming all observations within a class interval are concentrated at the class midpoint.

The midpoint (m_i) is calculated as the average of the upper and lower limits of the class. This value is then weighted by its frequency (n_i). The estimated mean is the summation of these weighted products ($\sum m_i n_i$) divided by the total number of observations (N , which is the sum of all frequencies). This weighted average approach ensures that classes with higher frequencies contribute proportionally more to the central tendency estimate.

The formula used to estimate the mean of our frequency distribution provides the statistical foundation for all subsequent standard deviation calculations, as the mean serves as the crucial reference point from which deviation is measured.

The formula for the estimated mean is expressed as:

$$\text{Mean: } \frac{\sum m_i n_i}{N}$$

where:

m_i : The midpoint of the i^{th} group.

n_i : The frequency of the i^{th} group.

N : The total sample size (sum of all frequencies).

Practical Calculation of the Mean in Excel

To implement the mean calculation in Excel, we introduce two new columns: one for the midpoint (m_i) and one for the weighted product ($m_i n_i$). Using the data entered previously (A and B for limits, C for frequency), we start by calculating the midpoint in Column D.

In cell **D2**, we input the formula `=AVERAGE(A2:B2)` to calculate the midpoint of the first class. This formula is then copied down the column (D3, D4, etc.). Next, in cell **E2**, we calculate the weighted product by multiplying the midpoint by the frequency: `=D2*C2`. This formula is also dragged down, yielding the total weighted value for each class interval.

Finally, the estimated mean is calculated in a single cell, such as **F2**. This cell combines the sum of Column E (numerator) and the sum of Column C (denominator, N). The formula `=SUM(E$2:E$6)/SUM(C$2:C$6)` calculates the weighted average. The result in F2 is the estimated mean (μ) and will be referenced in the standard deviation calculation.

To apply this formula in Excel, we will type the following formulas into cells **D2**, **E2**, and **F2**:

D2: =AVERAGE(A2:B2)

E2: =D2*C2

F2: =SUM(\$E\$2:\$E\$6)/SUM(\$C\$2:\$C\$6)

We will then click and drag these formulas down to each remaining cell in each column (D and E):

	A	B	C	D	E	F
1	Lower Class Limit	Upper Class Limit	Frequency (n_i)	m_i	m_i * n_i	μ
2	1	10	2	5.5	11	22.8913
3	11	20	7	15.5	108.5	22.8913
4	21	30	10	25.5	255	22.8913
5	31	40	3	35.5	106.5	22.8913
6	41	50	1	45.5	45.5	22.8913
7						
8						
9						
10						
11						

The Standard Deviation Formula for Grouped Data

The Standard Deviation (s) for a frequency distribution (often referred to as the standard deviation for Grouped Data) is derived from the variance, which measures the average squared deviation from the mean. Since this calculation often relies on sample data to infer population characteristics, we use the sample formula, requiring division by the degrees of freedom ($N-1$).

The core of the standard deviation calculation involves determining the squared difference between each class midpoint (m_i) and the estimated Mean (μ). This squared difference ensures that negative deviations are treated equally to positive ones and mathematically inflates larger deviations, emphasizing outliers or high variability.

Crucially, this squared difference must then be weighted by the frequency (n_i) of the respective class before summation. This step correctly accounts for the fact that classes with more observations should have a greater influence on the overall measure of dispersion. The final result is the square root of the variance, returning the measure of spread to the original units of measurement.

The formula for the sample standard deviation of a frequency distribution is:

Standard Deviation: $\sqrt{\frac{\sum n_i(m_i - \mu)^2}{N-1}}$

where:

n_i : The frequency of the i^{th} group.

m_i : The midpoint of the i^{th} group.

μ : The estimated mean.

N : The total sample size.

Executing the Standard Deviation Calculation in Excel

To manage the complex components of the standard deviation formula, we introduce three new computational columns in our Excel sheet: Column G for Deviation, Column H for Squared Deviation, and Column I for Weighted Squared Deviation. This breakdown isolates each part of the numerator calculation, making debugging and verification simpler.

In Column G, we calculate the deviation: $=D2-F2$ (where D2 is the midpoint and F2 holds the fixed mean value). In Column H, we square this deviation: $=G2^2$. This result represents the variance contribution of a single, unweighted midpoint. Both formulas are copied down the relevant rows.

The final calculation for the numerator is performed in Column I (Weighted Squared Deviation): $=C2*H2$. This multiplies the squared deviation by the class frequency (C2). Once this formula is dragged down the column, the sum of Column I will yield $\sum n_i(m_i - \mu)^2$, the sum of squares required for the variance.

To apply this formula in Excel, we will type the following formulas into cells **G2**, **H2**, and **I2**:

G2: $=D2-F2$

H2: $=G2^2$

I2: $=C2*H2$

We will then click and drag these formulas down to each remaining cell in each column:

	A	B	C	D	E	F	G	H	I
1	Lower Class Limit	Upper Class Limit	Frequency (n_i)	m_i	m_i * n_i	μ	m_i - μ	(m_i - μ)²	n_i(m_i - μ)²
2	1	10	2	5.5	11	22.8913	-17.3913	302.4575	604.9149
3	11	20	7	15.5	108.5	22.8913	-7.3913	54.63138	382.4197
4	21	30	10	25.5	255	22.8913	2.608696	6.805293	68.05293
5	31	40	3	35.5	106.5	22.8913	12.6087	158.9792	476.9376
6	41	50	1	45.5	45.5	22.8913	22.6087	511.1531	511.1531
7									
8									
9									
10									
11									
12									

Finalizing the Result and Interpretation

The culmination of our work involves combining the sums from the previous steps into the final standard deviation formula, using Excel's built-in mathematical functions. The numerator is the sum of Column I (I2:I6), and the denominator is the sum of Column C (C2:C6, representing $\sum n_i$) minus 1.

The entire calculation is contained within the `SQRT` function, which takes the variance (Numerator / Denominator) and returns the standard deviation. We place this final formula into a convenient, separate cell (e.g., B8) to clearly display the result of the complex calculation performed on the Frequency Distribution.

Lastly, we can type the following formula into cell **B8** to calculate the standard deviation of this frequency distribution:

=SQRT(SUM(I2:I6)/(SUM(C2:C6)-1))

The following screenshot shows how to use this formula in practice:

B8 : f_x =SQRT(SUM(I2:I6)/(SUM(C2:C6)-1))									
	A	B	C	D	E	F	G	H	I
1	Lower Class Limit	Upper Class Limit	Frequency (n_i)	m_i	$m_i * n_i$	μ	$m_i - \mu$	$(m_i - \mu)^2$	$n_i(m_i - \mu)^2$
2	1	10	2	5.5	11	22.8913	-17.3913	302.4575	604.9149
3	11	20	7	15.5	108.5	22.8913	-7.3913	54.63138	382.4197
4	21	30	10	25.5	255	22.8913	2.608696	6.805293	68.05293
5	31	40	3	35.5	106.5	22.8913	12.6087	158.9792	476.9376
6	41	50	1	45.5	45.5	22.8913	22.6087	511.1531	511.1531
7									
8	Std. Dev	9.637705925							
9									
10									
11									
12									
13									

The final calculated Standard Deviation of this frequency distribution turns out to be **9.6377**. This result signifies that the data points, represented by their class midpoints, deviate from the estimated mean by an average of 9.64 units. This manual, detailed process ensures a deep understanding of the statistical estimation required when analyzing grouped data in Microsoft Excel.