

What is a Marginal Distribution?

Authored by
Mohammed looti

November 6, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *What is a Marginal Distribution?*. PSYCHOLOGICAL STATISTICS.
Retrieved from <https://statistics.arabpsychology.com/?p=11487>

Understanding the Two-Way Frequency Table

In statistical analysis, organizing data efficiently is the first step toward drawing meaningful conclusions. A [two-way frequency table](#), often referred to as a contingency table, is a powerful tool designed to display the relationship between two distinct [categorical variables](#). This table systematically presents the frequencies, or counts, of how often specific combinations of these two variables occur simultaneously.

To illustrate this concept, consider a foundational survey involving 100 participants asked to identify their favorite sport among baseball, basketball, or football. This scenario involves two primary variables: the respondent's **Gender** and their preferred **Sport**. The data collected forms the basis for our table.

The structure of the two-way table organizes the data such that the rows represent one variable (Gender) and the columns represent the outcomes of the second variable (Sport). The intersection of a row and a column provides the **joint frequency** for that specific combination, showing the distribution of both variables together.

	Baseball	Basketball	Football	Total
Male	13	15	20	48
Female	23	16	13	52
Total	36	31	33	100

Defining the Marginal Distribution

While joint frequencies focus on interactions, a [marginal distribution](#) provides a crucial simplified perspective. It is defined as the [distribution](#) of each individual variable, entirely disregarding the influence or category of the other variable in the table. In essence, it answers the question: what is the overall frequency of outcomes for just one characteristic?

The terminology "marginal" is directly related to where these totals appear. In a two-way table, the marginal distributions are consistently displayed in the **margins**--the total row and the total column--of the table. These figures are calculated by summing all the counts across the rows or down the columns, respectively.

Using the sports survey example, the marginal distributions provide the overall counts for Sport preference (regardless of gender) and Gender (regardless of sport preference). The following

visual highlights these essential marginal totals:

	Baseball	Basketball	Football	Total
Male	13	15	20	48
Female	23	16	13	52
Total	36	31	33	100

Marginal distribution of gender

Marginal distribution of sports

From the totals calculated in the bottom margin, we derive the specific marginal distribution of sports preferences across the entire sample:

Baseball: 36 respondents

Basketball: 31 respondents

Football: 33 respondents

Calculating Marginal Distributions in Percentages

Presenting marginal distributions using raw counts is useful, but converting these totals into percentages often enhances interpretation, as it standardizes the data relative to the total sample size (N). To calculate the percentage for a category, we divide its marginal count by the grand total (the count in the very bottom-right corner) and multiply by 100.

For the marginal distribution of Sports, based on the grand total of 100 respondents, the corresponding percentage breakdown is calculated as follows:

Baseball: $36 / 100 = 36\%$

Basketball: $31 / 100 = 31\%$

Football: $33 / 100 = 33\%$

We apply the same calculation method to find the marginal distribution for the variable "Gender":

Male: $48 / 100 = 48\%$

Female: $52 / 100 = 52\%$

A key statistical property is that the components of any marginal [distribution](#), when summed, must always equal the grand total sample size or 100% when expressed as proportions. This inherent self-checking mechanism ensures the accuracy of the calculations.

Why Use Marginal Distributions?

The primary utility of marginal distributions lies in their ability to isolate the analysis of individual variables within a multivariate context. While the core of a [two-way frequency table](#) is often to explore conditional relationships (e.g., "What percentage of males prefer basketball?"), marginal distributions address research questions focused solely on overall prevalence.

For instance, a market researcher may only be interested in the total number of people who chose Football, regardless of gender, to gauge its overall popularity. The marginal total provides this exact, concise figure, eliminating the need to examine every cross-tabulation.

By extracting these distributions, analysts can establish the fundamental composition and characteristics of their sample population. This univariate perspective is often necessary for strategic planning, resource allocation, or as an initial step before proceeding to complex statistical tests of independence.

These techniques are essential practice for gaining a solid, intuitive understanding of how distributions operate within complex datasets.

Example 1: Marginal Distributions for Percentages

The following two-way table displays the results of a survey that asked 238 people which movie genre they liked best, categorized by gender:

	Fantasy	Drama	Action	Total
Male	22	30	70	122
Female	25	58	33	116
Total	47	88	103	238

Question A: What is the marginal [distribution](#) for movie genre (in percentages)?

Answer A: We calculate the total count for each genre category and divide it by the grand total (N=238):

Fantasy: $47 / 238 \approx 19.7\%$

Drama: $88 / 238 \approx 37.0\%$

Action: $103 / 238 \approx 43.3\%$

Question B: What is the marginal distribution for gender (in percentages)?

Answer B: We calculate the total count for each gender category and divide it by the grand total (N=238):

Male: $122 / 238 \approx 51.3\%$

Female: $116 / 238 \approx 48.7\%$

Example 2: Marginal Distributions for Counts

The following two-way table tracks the exam scores of 64 students in a class, cross-tabulated based on the number of hours they spent studying prior to the exam:

	71-80	81-90	91-100	Total
1 Hour	7	5	2	14
2 Hours	8	3	3	14
3 Hours	3	5	8	16
4 Hours	4	9	7	20
Total	22	22	20	64

Question A: What is the marginal [distribution](#) for exam scores (in counts)?

Answer A: We read the totals directly from the bottom row (the margin for exam scores):

71-80: 22 students

81-90: 22 students

91-100: 20 students

Question B: What is the marginal distribution for hours studied (in counts)?

Answer B: We read the totals directly from the right column (the margin for hours studied):

1 Hour: 14 students

2 Hours: 14 students

3 Hours: 16 students

4 Hours: 20 students

As expected, the sum of the exam score counts ($22 + 22 + 20$) and the sum of the hours studied counts ($14 + 14 + 16 + 20$) both align perfectly with the grand total of 64 students.

Additional Resources