Ordinal Scale

Authored by

mohammad looti

October 2, 2025

Ordinal Scale

Primary Disciplinary Field(s): Statistics, Psychometrics, Social Sciences, Research Methodology

1. Core Definition

An **ordinal scale** is a fundamental level of measurement used in statistics and research to categorize and rank data according to a meaningful order. This scale allows for the qualitative ordering of items or categories, establishing a hierarchy where one category is "greater than" or "preferred over" another, but without quantifying the precise difference or distance between these categories. The primary characteristic of an ordinal scale is its ability to convey relative position rather than absolute magnitude or equal intervals between successive ranks. It signifies a sequential relationship, indicating which element comes before or after another in a given attribute or quality.

Unlike more advanced measurement scales, an ordinal scale does not possess properties that permit arithmetic operations such as addition or subtraction to determine the exact magnitude of differences. For instance, if individuals rank their preference for ice cream flavors as 1) chocolate, 2) vanilla, and 3) strawberry, the ordinal scale clearly indicates that chocolate is preferred over vanilla, and vanilla over strawberry. However, it does not tell us by how much chocolate is preferred over vanilla, nor does it imply that the subjective preference gap between chocolate and vanilla is identical to the gap between vanilla and strawberry. The numerical assignment (1, 2, 3) merely serves as a label for rank, not a measure of interval distance.

The inherent limitation regarding the absence of equal intervals distinguishes ordinal scales from interval and ratio scales, which possess this crucial property. This means that while we know the order, we cannot assume uniform psychological or physical distances between adjacent categories. This characteristic has profound implications for the types of statistical analyses that are appropriate for data collected using an ordinal scale, often necessitating the use of non-parametric methods that do not rely on assumptions of equal intervals or a true zero point.

2. Etymology and Historical Development

The concept of measurement scales, including the ordinal scale, was formally introduced and systematized by the American psychologist Stanley Smith Stevens in his seminal 1946 article, "On the Theory of Scales of Measurement," published in the journal Science. Prior to Stevens' work, the understanding and classification of different types of measurement in psychology and other social sciences were less rigorous, leading to inconsistencies in data analysis and interpretation. Stevens' contribution provided a foundational framework that categorized measurement into four distinct levels: nominal, ordinal, interval, and ratio. This hierarchy helped researchers understand the mathematical properties inherent in different types of data and, consequently, the appropriate statistical techniques for their analysis.

Stevens' motivation for developing these scales stemmed from the need to bring greater scientific rigor to the emerging field of psychometrics. He argued that the type of scale used to measure a variable dictates what statistical operations are permissible and meaningful. For example, applying statistical methods appropriate for ratio data to nominal or ordinal data could lead to invalid conclusions. The ordinal scale, positioned above the nominal scale (which only categorizes without order) and below interval and ratio scales (which incorporate properties of equal intervals and a true zero, respectively), represented a crucial step in understanding data that has a natural order but lacks precise quantitative differences between its categories.

Since its introduction, Stevens' classification of measurement scales has become a cornerstone of statistical education and research methodology across various disciplines. While there have been debates and alternative classifications proposed over the decades, the nominal, ordinal, interval, ratio framework remains the most widely recognized and applied model for understanding the properties of data and guiding the selection of appropriate statistical tests. The ordinal scale, in particular, proved invaluable for fields like psychology, sociology, and marketing, where many phenomena (e.g., attitudes, preferences, social status) can be ranked but not always precisely quantified.

3. Key Characteristics

The **ordinal scale** is defined by several key characteristics that differentiate it from other levels of measurement. Foremost among these is the property of **order** or rank. Data measured on an ordinal scale can be meaningfully arranged in a sequential manner, indicating a higher or lower degree of an attribute. For instance, customer satisfaction ratings (e.g., "very dissatisfied," "dissatisfied," "neutral," "satisfied," "very satisfied") clearly convey an increasing level of satisfaction. This ordering is critical for understanding relative positions and making comparative judgments.

A second defining characteristic is the **absence of equal intervals** between adjacent categories. While the order is established, the "distance" or difference in magnitude between any two consecutive ranks is not assumed to be uniform or measurable. Using the example of Olympic medals (Gold, Silver, Bronze), we know that Gold is better than Silver, and Silver is better than Bronze. However, the performance gap between the gold medalist and the silver medalist might be minuscule (e.g., a fraction of a second), whereas the gap between the silver medalist and the bronze medalist could be much larger. The ordinal scale captures the rank but not the quantifiable difference in achievement. This lack of equidistant intervals means that arithmetic operations like addition or subtraction, which are common for interval and ratio data, are generally inappropriate and can lead to misleading interpretations if applied to ordinal data.

Furthermore, ordinal scales typically **lack a true zero point** that signifies the complete absence of the measured attribute. For example, while a student's grade (A, B, C) indicates their academic performance relative to others, there isn't a "zero performance" point in the same way that zero weight indicates no weight. Even in scales where '0' might be used (e.g., a pain scale from 0-10 where 0 means "no pain"), the numbers still primarily function as ranks, and the distance between 0 and 1 is not necessarily the same as between 9 and 10. Consequently, ratios are also not meaningful with ordinal data; for example, one cannot assert that a "very satisfied" customer is twice as satisfied as a "satisfied" customer. These characteristics guide researchers in selecting appropriate descriptive statistics, such as the mode and median, which are suitable for ordinal data, in contrast to the mean, which assumes equal intervals.

4. Examples and Applications

The **ordinal scale** finds extensive application across a multitude of disciplines due to its intuitive nature for ranking and ordering qualitative attributes. One of the most frequently cited examples is the ranking of athletes in competitions, such as the Olympic Games, where medals (**Gold**, **Silver**, **Bronze**) denote a clear hierarchy of achievement. A gold medalist has performed better than a silver medalist, who in turn has performed better than a bronze medalist. However, the specific performance difference between gold and silver, or silver and bronze, is not uniformly implied by these ranks, underscoring the core characteristic of unequal intervals.

In the social sciences, another pervasive example is the categorization of **socioeconomic status** (SES), often classified into categories such as **upper class**, **middle class**, and **lower class**. These categories represent a clear ordinal progression in terms of wealth, education, and occupational prestige. An individual in the upper class typically possesses more resources and opportunities than someone in the middle class, and similarly for the middle versus lower class. Yet, the social and economic "distance" between these strata is not necessarily equal; the disparity between the lower and middle class might be different from that between the middle and upper class, reflecting complex societal structures.

Moreover, Likert scales, widely used in surveys and questionnaires, are prime examples of ordinal measurement. These scales typically ask respondents to indicate their level of agreement, frequency, or satisfaction using categories like "**Strongly Disagree**," "**Disagree**," "**Neutral**," "**Agree**," and "**Strongly Agree**." While these responses can be assigned numerical values (e.g., 1 to 5) for ease of data entry and analysis, these numbers represent an ordered preference or intensity, not quantifiable intervals. The psychological gap between "Strongly Disagree" and "Disagree" may not be perceived as identical to the gap between "Neutral" and "Agree" by respondents. Other applications include educational grades (A, B, C), product quality ratings (Excellent, Good, Fair, Poor), military ranks, and disease severity classifications (Mild, Moderate, Severe), all of which establish a clear order without assuming equal step sizes between categories.

5. Distinction from Other Measurement Scales

To fully grasp the nature of the **ordinal scale**, it is crucial to understand its distinctions from the other three levels of measurement proposed by S. S. Stevens: nominal, interval, and ratio scales. These four levels form a hierarchy, each possessing the properties of the preceding scale while adding new ones. The most basic is the **nominal scale**, which merely categorizes data without any intrinsic order or numerical value. Examples include gender (male, female), hair color (blonde, brown, black), or religious affiliation. While ordinal scales also categorize, they add the crucial property of order, allowing for ranking.

Moving up the hierarchy, the **interval scale** possesses all the properties of an ordinal scale (categorization and order) but adds the characteristic of **equal intervals** between successive values. This means that the difference between any two adjacent points on an interval scale is consistent and meaningful. For instance, temperature measured in Celsius or Fahrenheit is an interval scale. The difference between 20°C and 30°C is the same as the difference between 30°C and 40°C (10 degrees). However, interval scales lack a true absolute zero point, meaning that zero does not signify the complete absence of the measured attribute (0°C does not mean no temperature). This lack of a true zero is the primary distinction between interval and ratio scales, and it means that ratios are not meaningful (e.g., 40°C is not twice as hot as 20°C).

Finally, the **ratio scale** is the most sophisticated level of measurement, encompassing all the properties of nominal, ordinal, and interval scales, plus the critical addition of a **true absolute zero point**. This true zero indicates the complete absence of the quantity being measured, making ratios meaningful. Examples include height, weight, age, income, and reaction time. A person who weighs 100 kg is indeed twice as heavy as a person who weighs 50 kg, and zero weight means no weight at all. Because ratio scales possess all properties, they allow for the widest range of statistical analyses, including all arithmetic operations. The ordinal scale, by contrast, sacrifices the precision of equal intervals and a true zero for the ability to establish a simple, yet powerful, order of preference or magnitude in qualitative data.

6. Statistical Analysis with Ordinal Data

The unique properties of the **ordinal scale** dictate the types of statistical analyses that can be appropriately and meaningfully applied to data collected at this level. Because ordinal data conveys order but not precise, equal intervals between ranks, parametric statistical tests that assume normally distributed data and equal variances (like the mean, standard deviation, t-tests, or ANOVA) are generally considered inappropriate. Applying these tests to ordinal data can lead to misleading results, as they treat the numerical labels of ranks as if they represent continuous, interval-level measurements.

For descriptive statistics, measures that rely on order rather than interval distance are suitable. The **mode**, which identifies the most frequently occurring category, is always appropriate for ordinal data, as it is for nominal data. The **median**, representing the middle value when data are ordered, is also highly suitable for ordinal data because it depends solely on the rank of values, not the distances between them. **Percentiles** and quartiles, which divide the ordered data into specific proportions, are similarly appropriate. However, the **mean** (average) and **standard deviation**, which require addition and subtraction and assume equal intervals, are typically not recommended for ordinal data.

For inferential statistics, **non-parametric tests** are the preferred choice for ordinal data, as they do not make assumptions about the underlying distribution of the data or the equality of intervals. Key non-parametric tests include the **Mann-Whitney U Test** (for comparing two independent groups), the **Wilcoxon Signed-Rank Test** (for comparing two related groups), the **Kruskal-Wallis H Test** (for comparing more than two independent groups), and the **Friedman Test** (for comparing more than two related groups). For examining associations between two ordinal variables, **Spearman's Rank Correlation Coefficient (rho)** and **Kendall's Tau** are widely used. These tests operate on the ranks of the data rather than their raw values, thus respecting the ordinal nature of the measurement and providing robust statistical inferences without violating fundamental assumptions.

7. Advantages and Disadvantages

The **ordinal scale** offers several distinct advantages that make it a valuable tool in various research contexts, particularly in the social sciences, marketing, and psychology. One of its primary benefits is its **simplicity and intuitive nature**. It allows researchers to easily collect and categorize data that involves subjective judgments, preferences, or attitudes, where precise quantification might be difficult or impossible. For instance, it's easier for respondents to rank their agreement with a statement or their satisfaction level than to assign an exact numerical value on a ratio scale. This ease of implementation makes it highly practical for surveys and questionnaires, contributing to higher response rates and less respondent burden.

Furthermore, ordinal scales are highly effective for capturing **qualitative distinctions and hierarchies**. They provide meaningful insights into the relative standing of individuals or items along a particular dimension. For example, knowing that one student achieved an 'A' while another received a 'B' offers a clear understanding of their relative academic performance, even without knowing the exact numerical difference in their scores. This ability to establish a clear order is crucial in fields where phenomena are inherently rankable, such as consumer preferences, disease severity, or educational attainment, where the "more than" or "less than" relationship is significant.

However, the ordinal scale also comes with notable **disadvantages and limitations**. The most significant drawback is the **limited scope for statistical analysis**. As discussed, parametric tests that assume equal intervals and a true zero are generally inappropriate, restricting researchers to non-parametric methods. While non-parametric tests are robust, they can sometimes be less powerful than their parametric counterparts, potentially requiring larger sample sizes to detect significant effects. This limitation can lead to a loss of information, as the magnitude of differences between ranks is ignored, even if those differences hold some underlying significance. For instance, if a researcher converts precise numerical scores into ordinal grades (e.g., 90-100 to A, 80-89 to B), the nuance of a score of 99 versus 91 is lost, reducing both to an 'A'.

8. Debates and Criticisms

Despite its widespread use and clear theoretical foundation, the application and interpretation of the **ordinal scale** have been subjects of considerable debate and criticism within the statistical and research communities, particularly concerning the treatment of ordinal data in quantitative analysis. The central point of contention often revolves around whether it is permissible, under certain circumstances, to treat ordinal data as if it were interval data, thereby allowing the use of more powerful parametric statistical tests.

This debate is especially prominent in disciplines heavily reliant on Likert scales and similar rating scales (e.g., 1 to 5, or 1 to 7) in psychology and social research. Proponents of treating ordinal data as interval argue that when an ordinal scale has a sufficient number of categories (e.g., five or more) and the distribution of responses approximates normality, the errors introduced by using parametric tests like means, t-tests, or ANOVA are negligible and do not significantly impact the validity of conclusions. They contend that the robustness of these parametric tests, coupled with their greater statistical power and ease of interpretation, outweighs the theoretical purism of strictly adhering to non-parametric methods. Some studies have even empirically demonstrated that parametric tests often yield similar conclusions to non-parametric tests when applied to ordinal data, especially with larger sample sizes.

Conversely, strict adherents to Stevens' original framework strongly maintain that violating the assumption of equal intervals fundamentally undermines the validity of parametric tests. They argue that applying arithmetic operations (like calculating a mean) to data that only represent ranks can produce nonsensical or misleading results, as the average of ranks does not necessarily reflect the true average of the underlying construct. For instance, calculating the mean of "Strongly Disagree" (1) and "Strongly Agree" (5) might yield "Neutral" (3), but this average is based on arbitrary numerical assignments and does not reflect a true midpoint if the psychological distance between 1 and 2 is far greater than between 2 and 3. This perspective emphasizes that the choice of statistical test should always align with the measurement properties of the data to ensure accurate and defensible research findings, advocating for the consistent use of non-parametric methods for all ordinal data.

9. Relationship to Ordinal Variable

The concept of an **ordinal scale** is intrinsically linked to the definition of an **ordinal variable**. In essence, an ordinal scale is the mechanism or method used to measure a phenomenon, and the outcome of that measurement is an ordinal variable. An ordinal variable is a type of categorical variable where the categories have a natural, ordered relationship.

When data is collected using an ordinal scale, the resulting data points are values of an ordinal variable. For example, if researchers use a Likert scale (an ordinal scale) to measure agreement, the responses ("Strongly Agree," "Agree," etc.) constitute the values of an ordinal variable. Similarly, when students receive grades (A, B, C), these grades form an ordinal variable because they represent ordered categories of academic performance. The scale provides the framework, and the variable is the data generated from that framework.

Further Reading

Level of measurement - Wikipedia
Ordinal data - Wikipedia
Stanley Smith Stevens - Wikipedia
Psychometrics - Wikipedia
Social sciences - Wikipedia
Research methods - Wikipedia
Likert scale - Wikipedia
Mann-Whitney U Test - Wikipedia
Kruskal-Wallis one-way analysis of variance - Wikipedia
Spearman's rank correlation coefficient - Wikipedia