**Parametric and non-parametric statistics**

Parametric and non-parametric statistics are two different approaches to analyzing data in statistics. They differ primarily in the assumptions they make about the underlying distribution of the data and the types of tests they use.

Parametric statistics:

1. Assumptions: Parametric statistics assume that the data follow a specific distribution, usually a normal (Gaussian) distribution. This assumption requires that the data have certain characteristics, such as being continuous, having a mean and a standard deviation, and being symmetric around the mean.
2. Types of tests: Parametric tests are designed based on these assumptions, and they are more powerful when the assumptions hold true. Some common parametric tests include t-tests, analysis of variance (ANOVA), linear regression, and Pearson's correlation coefficient.
3. Sample size: Parametric tests generally require larger sample sizes to achieve reliable results. However, they are more efficient and can provide more accurate estimates when the assumptions are met.

Non-parametric statistics:

1. Assumptions: Non-parametric statistics do not assume any specific distribution for the data. They are distribution-free methods that make fewer assumptions about the underlying population. This makes non-parametric statistics more robust and flexible when dealing with data that may not meet the requirements of parametric statistics.
2. Types of tests: Non-parametric tests are designed for data that do not meet the assumptions of parametric tests. Some common non-parametric tests include the Wilcoxon signed-rank test, the Mann-Whitney U test, the Kruskal-Wallis test, and Spearman's rank correlation coefficient.
3. Sample size: Non-parametric tests can be used with smaller sample sizes and are more appropriate for data that are not normally distributed, have outliers, or are ordinal or ranked in nature.

In summary, parametric statistics make strong assumptions about the distribution of data and are more powerful when those assumptions hold true. Non-parametric statistics make fewer assumptions and are more robust but generally less powerful when the assumptions of parametric statistics are met. The choice between these two approaches depends on the nature of the data and the objectives of the analysis.

**Biased and unbiased estimation**

In statistics, estimation is the process of using sample data to make inferences about a population parameter. Biased and unbiased estimators are two types of estimators used to draw conclusions about population parameters.

Biased Estimator:

1. Definition: A biased estimator is an estimator whose expected value (average) over many samples does not equal the true population parameter. In other words, the estimator systematically overestimates or underestimates the true value of the parameter.
2. Example: A common example of a biased estimator is the sample variance calculated using the formula (1/n) \* Σ(xi - x̄)^2, where n is the sample size, xi are the individual data points, and x̄ is the sample mean. This formula systematically underestimates the true population variance, particularly when the sample size is small.

Unbiased Estimator:

1. Definition: An unbiased estimator is an estimator whose expected value (average) over many samples is equal to the true population parameter. This means that, on average, an unbiased estimator will provide an accurate estimate of the population parameter, although individual estimates may still be off.
2. Example: A common example of an unbiased estimator is the sample variance calculated using the formula (1/(n-1)) \* Σ(xi - x̄)^2, where n is the sample size, xi are the individual data points, and x̄ is the sample mean. This formula, known as Bessel's correction, corrects the bias found in the previous formula and provides an unbiased estimate of the population variance.

In practice, statisticians prefer using unbiased estimators because they provide accurate estimates of population parameters on average. However, it's important to remember that unbiasedness is only one desirable property of an estimator. Other properties, such as efficiency (the estimator's variance) and consistency (the estimator's behavior as the sample size increases), should also be considered when selecting an estimator for a specific problem.

**The standard error of an estimator**

The standard error of an estimator is a measure of the precision or accuracy of that estimator. In other words, it quantifies the variability or dispersion of the estimator when it is repeatedly applied to different samples drawn from the same population. A smaller standard error indicates a more precise estimator, while a larger standard error indicates a less precise one.

For a given estimator, the standard error is the square root of its variance. The variance of an estimator is a measure of how much the estimator's values would vary if you were to draw many different samples from the same population and compute the estimator for each sample.

For example, let's consider the sample mean (x̄) as an estimator of the population mean (μ). The standard error of the sample mean (SE) is given by:

SE = σ / √n

where σ is the population standard deviation, and n is the sample size. This formula assumes that the data follows a normal distribution or that the sample size is large enough for the Central Limit Theorem to apply.

The standard error is an important concept in inferential statistics because it plays a key role in constructing confidence intervals and conducting hypothesis tests. Confidence intervals provide a range of values within which the population parameter is likely to fall, and the standard error helps determine the width of this interval. In hypothesis tests, the standard error is used to calculate test statistics, such as the t-statistic or z-statistic, which help determine the significance of the results.

In general, the standard error decreases as the sample size increases, indicating that larger samples provide more precise estimates of population parameters. However, the rate of decrease diminishes as the sample size gets larger, so there is a trade-off between the precision gained and the additional effort required to collect more data.

**UMVUE "Uniformly Minimum Variance Unbiased Estimator."**

UMVUE is a concept in statistical estimation theory. An estimator is called UMVUE if it satisfies the following two conditions:

1. Unbiasedness: The estimator is unbiased, meaning that its expected value (average) over many samples equals the true population parameter. In other words, the estimator doesn't systematically overestimate or underestimate the true value.
2. Minimum Variance: Among all unbiased estimators, the UMVUE has the smallest variance. The variance of an estimator is a measure of how much its values would vary if you were to draw many different samples from the same population and compute the estimator for each sample. A smaller variance indicates a more precise estimator.

The UMVUE is an optimal estimator because it provides the most accurate estimates (in terms of unbiasedness) with the highest precision (lowest variance) among all possible unbiased estimators. However, finding the UMVUE can be challenging, and it may not always exist for a given problem.

One common method for finding the UMVUE is the Lehmann-Scheffé theorem, which states that if a sufficient and complete statistic exists for a population parameter, the conditional expectation of the complete statistic given the sufficient statistic is the UMVUE of the parameter.

In practice, statisticians strive to use estimators that are unbiased and have low variance, although the UMVUE may not always be available or easy to compute. Other properties, such as consistency (the estimator's behavior as the sample size increases) and efficiency, should also be considered when selecting an estimator for a specific problem.

**Methods of moments**

The method of moments is a technique used in statistics for estimating population parameters based on the moments of a sample. Moments are summary statistics that describe the shape and characteristics of a distribution. The method of moments involves equating the sample moments to their corresponding population moments and solving for the population parameters.

There are two types of moments:

1. Raw moments: These are the expected values of the data raised to various powers. The nth raw moment is given by E[X^n], where X is a random variable, and n is a positive integer.
2. Central moments: These are the expected values of the deviations of the data from the mean, raised to various powers. The nth central moment is given by E[(X - μ)^n], where μ is the mean of the random variable X, and n is a positive integer.

The first moment corresponds to the mean, and the second central moment corresponds to the variance.

To apply the method of moments, follow these steps:

1. Choose the number of moments: Decide on the number of moments you want to use for estimation. This number should be equal to the number of unknown parameters in the distribution.
2. Calculate sample moments: Compute the sample moments based on the available data. For example, the sample mean (first moment) is calculated as x̄ = (Σx\_i) / n, and the sample variance (second central moment) is calculated as s^2 = (Σ(x\_i - x̄)^2) / (n-1).
3. Equate sample moments to population moments: Set the sample moments equal to their corresponding population moments. For instance, equate the sample mean to the population mean and the sample variance to the population variance.
4. Solve for population parameters: Solve the resulting equations for the unknown population parameters.

The method of moments is relatively easy to apply and computationally simple, but it may not always yield the most efficient or accurate estimators. Other estimation techniques, such as maximum likelihood estimation, may provide better results in some cases. However, the method of moments remains an important tool in the toolbox of a statistician, particularly for its simplicity and ease of use.