DOI: https://doi.org/10.52206/jsmc.2023.13.4.886
OPEN ACCESS

Concept and Interpretation of Confidence Interval: How to Overcome Limitations of P-value

Umar Hussain,[1] Farhad Ali[2]

This article may be cited as: Hussain U, Ali F. Concept and Interpretation of Confidence Interval: How to Overcome Limitations of P-value. J Saidu Med Coll Swat 2024;14(1):66-67. DOI: https://doi.org/10.52206/jsmc.2023.13.4.886

ABSTRACT

Knowing the estimate (mean or proportion) for the whole population is not feasible due to a lack of resources. Therefore, we rely on a sample estimate to draw conclusions about the population. The range of values in which the true population estimate lies is called a confidence interval (CI) which is more reliable than single estimate. For descriptive studies, CIs should be reported along with the point estimate (mean or proportion). For effect sizes in inferential statistics (mean difference, odds ratio, risk ratio, etc.), the p-value should be supplemented with a CI for strength and direction of the association.

INTRODUCTION

In the previous paper, we discussed what a p-value is, how to interpret it, and its limitations.[1] In this paper, we introduce a related but more important concept: the Confidence Interval (CI). For reporting research results, it is highly recommended to present CIs along with p-values.

In statistics, there are two concepts: point estimate and interval estimate. The point estimate is a single value, such as the mean or proportion, while the interval estimate is the range of those estimates from a lower to an upper border. An example of an interval estimate is the Confidence Interval (CI).

CI for Descriptive Statistics

Let's consider a simple scenario: a hematologist of Saidu Medical College wants to determine the mean Hemoglobin (Hb) level of MBBS students. He randomly selected 60 students out of 800, and their calculated results indicate a mean Hb of 14.71 mg/dl. If he were to randomly select another 60 students, their Hb levels would not be exactly the same, possibly a little higher or lower than 14.71 mg/dl. This difference between the sample estimate (mean) and the population is termed sampling error. Standard error (SE) quantifies sampling errors. Mathematically, for the mean, SE = SD/vn, where SD is the standard deviation, and n is the sample size. However, the formula for SE varies with the parameter (Table 01). We don't need to remember the formulas; computer programs like Excel, SPSS, or R, etc., can easily compute SE and CI.

1. Saidu Medical College, Saidu Sharif Swat, Pakistan.
2. Biostatistician at Khyber College of Dentistry, Peshawar.

Correspondence: Dr. Umar Hussain
Associate Editor & Biostatistician at J Saidu Med Coll, Swat
Lecturer, Saidu College of Dentistry
Email: drumarhusain@gmail.com

Received: January 07th, 2024 - Accepted: February 11th, 2024

Concentrating on the above example again, the hematologist will be interested in determining the real mean Hb level of the population (800 students), which is not possible without including the whole population in data collection, a process called a census. Normally, due to a lack of resources, the researcher estimate the population from a sample, but the sample should be randomly selected. If it is not possible to determine the real mean value (Hb level), the hematologist must have an interest in knowing the possible range of Hb levels, referred to as the period estimate or confidence interval (CI). Typically, the level of CI is set to be 95%. Mathematically, CI = ± 1.96SE, where is the mean value and SE is the standard error. If the hematologist calculates the standard deviation (let's say 3.15) or SE for 60 students and computes the CI, the range will be 14.49 to 14.92. This will indicate that we are 95% confident that the real mean Hb of the entire population (800 students) falls between 14.49 and 14.92 mg/dl.

In published research, authors report a point estimate with a confidence interval (CI), such as the prevalence of body dysmorphic disorders as 13.5% (CI: 11.2%, 15.4%). This CI indicates that the real prevalence can be as low as 11.2% and as high as 15.4%.

How CI shows the statistical significance

One other important use of confidence intervals (CI) is to overcome the limitation of p-values, as discussed in a previous paper.[1] The p-value is all or none; it only shows whether the difference is significant (leading to the rejection of the null hypothesis) or not (resulting in the failure to reject the null hypothesis). However, it does not indicate the strength of significance. Confidence intervals consist of two boundaries: the lower and upper bounds. When the null value (zero for differences such as mean difference, risk difference, and one for ratios like odds ratio, risk ratio) falls between these boundaries, the result is not statistically significant. A narrower width of the confidence interval indicates a higher strength of association.[2]

CI for Effect Sizes in the Inferential Statistics

Various effect sizes are employed in medical and dental research to quantify the impact of clinical interventions, such as mean difference, odds ratio, risk ratio, hazard ratio, and difference in median survival time. CONSORT guidelines require the reporting of study outcomes for each group, including summary statistics (mean, standard deviation, and proportion), and effect sizes (odds ratio, risk ratio, hazard ratio, and means, etc.). It is recommended to report confidence intervals for differences rather than separately for each group outcome.[3]

Let's take an example where we assess two surgical procedures, A and B, for the outcome of mortality. We used Cox regression to calculate the hazard ratio (HR). The HR for procedure A was 1.2 (p = 0.1), with a 95% CI of 0.9 to 1.45. The HR suggests a 20% higher hazard of mortality for A compared to B, but the non-significant p-value (0.1) and the CI including 1 indicate no statistically significant difference at the 0.05 level. The wide CI (0.9 to 1.45) suggests uncertainty, showing mortality could decrease by 10% or increase by 45%, suggesting potential benefits in some cases.

Table 1: Standard errors and confidence intervals of various parameter

Parameter	SE Formula	CI Formula	Parameter Descriptions
Odds Ratio (OR)	$SE_{\log(OR)} = \sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$	CI_{OR} = OR x exp(±1.96 x $SE_{\log(OR)}$)	*a,b,c,d*: Elements of 2x2 contingency table
Risk Ratio (RR)	$SE_{\log(RR)} = \sqrt{\frac{1}{a}-\frac{1}{a+b}+\frac{1}{c}-\frac{1}{c+d}}$	CI_{RR} = RR x exp(±1.96 x $SE_{\log(RR)}$)	*a,b,c,d*: Elements of 2x2 contingency table
t-test (sample mean)	$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$	$CI_{\bar{x}} = \bar{x} \pm 1.96 \times SE_{\bar{x}}$	s: Sample standard deviation, *n*: Sample size
ANOVA (one-way)	$SE_{mean} = \sqrt{\frac{MS_{error}}{n_i}}$	CI_{mean} = mean ± 1.96 x SE_{mean}	MS_{error}: Mean square error, n_i: Sample size in *i*-th group
Proportion	$SE_{\hat{p}} = \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$	$CI_{\hat{p}} = \hat{p} \pm 1.96 \times SE_{\hat{p}}$	$\hat{p}$: Sample proportion, *n*: Sample size
Mean (sample)	$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$	$CI_{\bar{x}} = \bar{x} \pm 1.96 \times SE_{\bar{x}}$	s: Sample standard deviation, *n*: Sample size
Mean Difference	$SE_{diff} = \sqrt{SE_{\bar{x}_1}^2 + SE_{\bar{x}_2}^2}$	CI_{diff} = diff ± 1.96 x SE_{diff}	$SE_{\bar{x}_1}$, $SE_{\bar{x}_2}$: Standard errors of the means, diff: Mean difference

REFERENCES

1. Hussain U, Ali F. Meaning of P-value. J Saidu Med Coll. 2023 Dec 19;13(4):200-1.
2. Pandis N. Confidence intervals rather than P values. Am J Orthod Dentofacial Orthop. 2013 Feb 1;143(2):293-4.
3. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomized trials. Br Med J 2010;340:c869.