SYMMETRY OF ODDS RATIO

Nazrul Islam*

ABSTRACT

Unlike the case with normally distributed continuous variable, the Confidence Interval (CIs) of Odds Ratio (OR), the often-used measure of association in medical literature, is said NOT to be symmetrical about the point estimates. The inference is based on the conventional additive construct, which is why it is shown that the CIs of natural logarithm (*ln*) of OR is symmetrical about the *ln*(OR), and NOT the OR itself. There is consensus that the log-scale is not quite intuitive, and hence the CI of *ln*(OR) is later expressed (taking anti-log) in natural scale. However, after the transformation from *ln*(OR) to OR, the CI of OR is blamed not to have symmetry about the point estimate, and hence the construct of CI of OR remains unintuitive. However, we show mathematically that CI of OR is also symmetrical about the point estimate of OR, only that it is on multiplicative construct. The same is applicable for Risk Ratios (RR) too.

Keywords: Confidence Interval (CI), Multiplicative, Odds Ratio (OR), Risk Ratio (RR), Symmetry/Symmetrical

* FRSPH, MSc, MPH, MBBS; School of Population and Public Health, Faculty of Medicine, University of British Columbia, Vancouver, Canada

The Confidence Interval (CI) of the Point Estimate (e.g., mean) of a normally-distributed continuous variable intuitively tells us that the range within which we are (1-α)% confident in our estimation is symmetric about the 'point estimate', e.g., sample mean ($\bar{x}$). The lower bound of a (1-α)% confidence interval, $(\mathbf{CI}_{\alpha/2}) = \bar{\mathbf{x}} - \mathbf{Z}_{\alpha/2} * \mathbf{SE}(\bar{\mathbf{x}})$, and upper bound, $(\mathbf{CI}_{1-\alpha/2}) = \bar{\mathbf{x}} + \mathbf{Z}_{\alpha/2} * \mathbf{SE}(\bar{\mathbf{x}})$; where, SE=Standard Error. This is surely symmetrical, as $\mathbf{point\ estimate} - \mathbf{lower\ bound\ CI} = \mathbf{upper\ bound\ CI} - \mathbf{point\ estimate}$; $[\bar{\mathbf{x}} - \mathbf{CI}_{\alpha/2} = \mathbf{Z}_{\alpha/2} * \mathbf{SE}(\bar{\mathbf{x}}) = \mathbf{CI}_{1-\alpha/2} - \bar{\mathbf{x}}]$; meaning the CI spreads, on either side (in two-sided CIs), by $\mathbf{Z}_{\alpha/2} * \mathbf{SE}(\bar{\mathbf{x}})$

However, the CI of Odds Ratio (OR), the often-used measure of association in the medical literature, is blamed NOT to have the symmetry about the point estimate ($\widehat{\mathbf{OR}}$),[1] i.e., $\widehat{\mathbf{OR}} - \mathbf{CI}_{\alpha/2} \neq \mathbf{CI}_{1-\alpha/2} - \widehat{\mathbf{OR}}$. This is because of the fact the OR is skewed to the right (being ranged between 0 and ∞), and hence it is log-transformed [taking the natural logarithm of OR, **ln(OR)**] to make the distribution normal. This is done to construct the confidence interval and perform hypothesis testing. The CI of ln(OR) is symmetric about ln(OR), however we seldom report the CI of OR in log-scale as this is quite unintuitive. Hence, CI is presented taking the anti-log (exponent), which certainly does not appear symmetric about the point estimate. Therefore, we do not look at the CIs of OR to *intuitively* understand the spread of the CI, as we do in cases of sample means for example. However, that is NOT the end of the story. While OR has thus far been so famous to the statisticians and epidemiologists thanks to its amazing mathematical properties, it is no less magical in CI either! And, here we say, CI of OR is also symmetrical, not in conventional additive construct, rather in multiplicative construct. Let us proceed to the mathematical notation: The CI for $\mathbf{ln}(\widehat{\mathbf{OR}}) = \mathbf{ln}\,(\widehat{\mathbf{OR}}) \pm \mathbf{Z}_{\alpha/2} * \mathbf{SE}(\mathbf{ln}(\widehat{\mathbf{OR}}))$. Taking the exponent, lower bound of a (1-α)% confidence interval of the estimated OR, $\mathbf{CI}_{\alpha/2} = \mathbf{e}^{[\mathbf{ln}(\widehat{\mathbf{OR}}) - \mathbf{Z}_{\alpha/2} * \mathbf{SE}(\mathbf{ln}(\widehat{\mathbf{OR}}))]}$ and upper bound, $\mathbf{CI}_{1-\alpha/2} = \mathbf{e}^{[\mathbf{ln}(\widehat{\mathbf{OR}}) + \mathbf{Z}_{\alpha/2} * \mathbf{SE}(\mathbf{ln}(\widehat{\mathbf{OR}}))]}$.

Lower bound CI can be simplified, following the rules of exponent, as follows: $\mathbf{e}^{[\mathbf{ln}(\widehat{\mathbf{OR}}) - \mathbf{Z}_{\alpha/2} * \mathbf{SE}(\mathbf{ln}(\widehat{\mathbf{OR}}))]} = \mathbf{e}^{\mathbf{ln}(\widehat{\mathbf{OR}})} \div \mathbf{e}^{[\mathbf{Z}_{\alpha/2} * \mathbf{SE}(\mathbf{ln}(\widehat{\mathbf{OR}}))]}$. Similarly, the upper bound CI, $\mathbf{e}^{[\mathbf{ln}(\widehat{\mathbf{OR}}) + \mathbf{Z}_{\alpha/2} * \mathbf{SE}(\mathbf{ln}(\widehat{\mathbf{OR}}))]} = \mathbf{e}^{\mathbf{ln}(\widehat{\mathbf{OR}})} * \mathbf{e}^{[\mathbf{Z}_{\alpha/2} * \mathbf{SE}(\mathbf{ln}(\widehat{\mathbf{OR}}))]}$

There we go! We see a perfectly symmetrical relationship among the $\widehat{\mathbf{OR}}$, $\mathbf{CI}_{\alpha/2}$ and $\mathbf{CI}_{1-\alpha/2}$, this time taking a little different shape: $\mathbf{point\ estimate} \div \mathbf{lower\ bound\ CI} = \mathbf{upper\ bound\ CI} \div \mathbf{point\ estimate}$

$\mathbf{\widehat{OR} \div CI_{\alpha/2} = \left[e^{\ln(\widehat{OR})}\right] \div \left[e^{\ln(\widehat{OR})} \div e^{[Z_{\alpha/2}*SE(\ln(\widehat{OR}))]}\right] = e^{Z_{\alpha/2}*SE(\ln(\widehat{OR}))}}$, and

$\mathbf{CI_{1-\alpha/2} \div \widehat{OR} = \left[e^{\ln(\widehat{OR})} * e^{[Z_{\alpha/2}*SE(\ln(\widehat{OR}))]}\right] \div \left[e^{\ln(\widehat{OR})}\right] = e^{Z_{\alpha/2}*SE(\ln(\widehat{OR}))}}$, meaning the Confidence Intervals of Odds Ratio spread, on either side (in two-sided CIs), by $\mathbf{e^{[Z_{\alpha/2}*SE(\ln(\widehat{OR}))]}}$, **on a multiplicative construct**. Also the product of the upper and lower confidence bounds is equal to the square of the point estimate of the OR (see below).

$$\mathbf{\left[e^{\ln(\widehat{OR})} \div e^{[Z_{\alpha/2}*SE(\ln(\widehat{OR}))]}\right] * \left[e^{\ln(\widehat{OR})} * e^{[Z_{\alpha/2}*SE(\ln(\widehat{OR}))]}\right] = \left[e^{\ln(\widehat{OR})}\right]^2 = [\widehat{OR}]^2}$$

Let's take a quick look on a real example: In a recent paper by Shindel and Vittinghoff,[2] the adjusted OR of every 10-year increase in age for Erectile Dysfunction was reported as 1.495 (95% CI: 1.353–1.653, p<0.001). Now, $1.495 \div 1.353 = 1.11 = 1.653 \div 1.495$, and $1.353*1.653=2.24=1.495^2$

I hope this helps us appreciate the fact that the construct of CI of OR is not as unintuitive as we thought it to be! Note, all the explanations about the symmetry of OR are also applicable for Risk Ratio (RR)!

Acknowledgement: I am grateful to **Prof. Eric Vittinghoff** at the Division of Biostatistics, Department of Epidemiology and Biostatistics, University of California, San Francisco, for his review and encouragement in preparing this manuscript.

Declaration of Interest:

The authors report no declarations of interest.

References:

1. Bland JM, Altman DG. Statistics notes: the odds ratio. *BMJ.* 2000;320(7247):1468.
2. Shindel AW, Vittinghoff E, Breyer BN. Erectile Dysfunction and Premature Ejaculation in Men Who Have Sex with Men. *J Sex Med.* 2012;9(2):576-584.