HealthStats NSW

Confidence intervals

General information

Usage

Confidence intervals, also known as error margins, are used to illustrate the noise or error present in data. The standard margin of error used in surveys like opinion polls is 95%, so the statement "the poll found 46% support with a 3% margin of error" is equivalent to saying the poll had a 95% confidence interval of 43-49%. HealthStats NSW uses this standard.

Roughly speaking, a 95% confidence interval can be interpreted to say that there is a 95% probability that the true population value lies within the interval provided. For example, consider Figure 1, which comes from a survey of NSW residences. The sample's estimate of high blood pressure in 16-24-year old females is 5.4%, and there is a 95% chance that if we measured all 16-24-year old females in NSW, the percentage who had high blood pressure would lie between 3.2% and 7.6%. Note that while this interpretation can be used for general purposes, the technical definition of a confidence interval differs slightly.

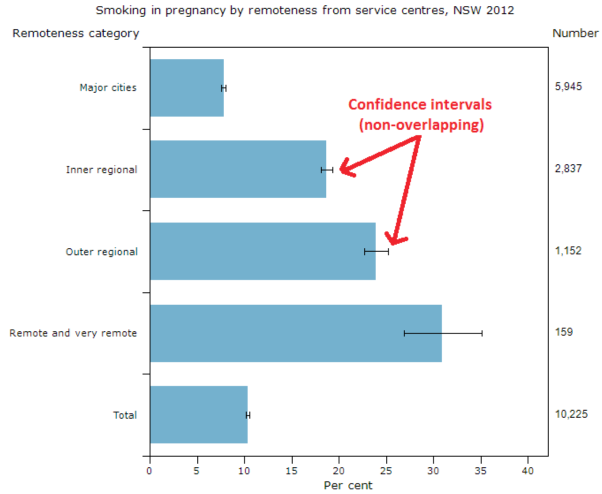
Figure 1. Confidence intervals in data tables

High blood pressure by age and sex, persons aged 16 years and over, NSW 2013

Age (years)	Sex	Number of Respondents	Actual estimate (Per cent)	LL 95% CI	UL 95% CI
16-24	Males	530	9.0	5.5	12.5
	Females	510	5.4	3.2	7.6
	Persons	1,040	7.2	5.1	9.3
25-34	Males	592	11.2	8.4	14.1
	Females	785	12.0	9.2	14.9
	Persons	1,377	11.7	9.6	13.7
35-44	Males	638	19.9	16.0	23.9
	Females	881	19.9	16.6	23.1
	Persons	1,519	19.9	17.3	22.5

The more data we have the more certain we can be about our indicator value; this means that for smaller sample sizes there will be larger confidence intervals. Figure 2 illustrates the varying widths of confidence intervals for remoteness categories within the Smoking in Pregnancy indicator. Given that there are far fewer pregnancies in remote and very remote areas, the confidence interval for this category is much wider than for other categories.

Figure 2. Confidence intervals in figures



Presented visually, confidence intervals not only give an indication of the uncertainty present in each category, but also offer a simple (but inexact) means of assessing potential differences between categories. When confidence intervals for different categories overlap to a large degree, differences between the categories cannot be concluded. If the confidence intervals are wholly separated from one another (as they are in Figure 2), one may infer a difference between the categories. Please note, this should be considered a guide only and a formal test would be required to arrive at statistically credible conclusions.

Sampling and administrative datasets

The size of a confidence interval covers only the statistical uncertainty (error) owing to the randomness inherent in sampling. In surveys such as the NSW Population Health Survey, notions of the sample and estimated population are clear. However, it is less clear in administrative datasets, such as the Perinatal Data Collection, which provide an exhaustive record of all relevant events in NSW. Despite this, these indicators are viewed as an example of the kinds of values we would expect from years and areas similar to the one we are considering.

It is also true that where administrative databases are concerned, there may be other sources of uncertainty. These could stem from diagnostic, patient-characteristic or coding ambiguities. Every effort has been made in HealthStats NSW to reduce these possible sources of additional uncertainty.

Technical information

Binomial data

(Example: www.healthstats.nsw.gov.au/indicator/mab_smo_cat)

Where estimates are provided as percentages derived from binomial data, HealthStats NSW uses the Clopper-Pearson (CP) method for generating confidence intervals.

The CP interval (also called the "Exact" interval) is considered a conservative alternative to the more common Wald interval. Several papers have suggested that the CP method can be too conservative in certain circumstances, however a convenient feature of the CP interval is that no matter what combination of sample p and n, the coverage of a 95% CP interval would be at least 95% for any underlying true population percentage[1,2,3]. It is also noted that the CP method is considerably more conservative for very small rates or proportions.[4] Several such small rates currently exist on HealthStats NSW, including illicit drug-attributable hospitalisations (approx. 30 per 100,000 population).

HealthStats NSW considers that a conservative approach to the generation of confidence intervals is appropriate in a public health setting, and particularly appropriate in cases of small proportions, which are more affected by random variation. Administrative datasets are considered to impose some additional variation onto estimates, further justifying the use of a conservative approach to the generation of the interval.

Lower Bound

$$\pi_l = \frac{x}{x + (n - x + 1)F}$$

Where F is the $F^{v_1}_{v_2 1-\frac{\alpha}{2}}$ quantile
with $v_1 = 2n - 2x + 2$ and
$v_2 = 2x$ degrees of freedom

Upper Bound

$$\pi_u = \frac{(x+1)F}{n - x + (x+1)F}$$

Where F is the $F^{v_1}_{v_2 1-\frac{\alpha}{2}}$ quantile
with $v_1 = 2x + 2$ and
$v_2 = 2n - 2x$ degrees of freedom

Directly standardised rates

(Example: www.healthstats.nsw.gov.au/indicator/beh_bmiafdth)

Where estimates are provided as directly standardised rates (e.g. hospitalisations per 100,000 population), HealthStats NSW uses the method prescribed in Dobson et al.[5] These standardised rates are derived from weighted sums of age-specific rates for males and females respectively. The numbers of events observed in each age group are considered to be independent and have Poisson distributions. The lower and upper bounds of the rates are given by:

$$\text{Lower Bound: } T_l = r + \sqrt{\frac{\sum_{i=1}^{k} w_i^2 X_i}{X}} \times (X_l - X)$$

$$\text{Upper Bound: } T_u = r + \sqrt{\frac{\sum_{i=1}^{k} w_i^2 X_i}{X}} \times (X_u - X)$$

Where:
X is the total count of events and
X_l, X_u are the exact upper and lower confidence intervals for Poisson count X
r is the direct standardised rate and
w_i are their weights used in direct standardisation

Survey indicators

(Example: www.healthstats.nsw.gov.au/indicator/beh_alc_age)

Where data is sourced from NSW Population Health Surveys, a Taylor series linearisation is performed to generate the variance of the estimates. This method obtains a first-order linear approximation for the rate estimator and then uses the variance estimate for this approximation to estimate the variance of the estimate itself.[6,7] This is conducted through the PROC STDRATE function in SAS[8] and is provided as the default variance estimation method for the output confidence intervals.

References

1. Vollset S. Confidence intervals for a binomial proportion. *Stat Med* 1993; 12(9): 809-24.
2. Brown DL, Cai TT, DasGupta A. Interval estimation for a binomial proportion. *Stat Sci* 2001; 16(2): 101-33.
3. Agresti A, Coull B. Approximate is better than 'exact' for interval estimation of binomial proportions. *The American Statistician* 1998; 52(2): 119-26.
4. Dunnigan K. Confidence interval calculation for binomial proportions. Statking Consulting Inc, P08-2008.
5. Dobson AJ, Kuulasmaa K, Eberle E, Scherer J. Confidence intervals for weighted sums of Poisson parameters. *Stat Med* 1991; 10(3): 457-62.
6. Fuller WA. Regression analysis for sample surveys. Sankhya 1975; C37: 117-32.
7. Woodruff R. A simple method for approximating the variance of a complicated estimate. *J Am Stat Assoc* 1971; 66: 411-4.
8. SAS Institute. SAS v9.3. Cary, North Carolina: SAS Institute, 2012.