

Assignment 4: Writing functions

Due Sunday, November 18, 2019 @ 11:59pm

Public Health 280.346, Fall 2019

Your assignment for this module is to write a series of functions for calculating 95% confidence intervals for different population parameters. To help with this, relevant entries from the formula sheet in Public Health Biostatistics are reproduced at the end of this assignment.

To complete your assignment, do the following items. You can use the `Assignment4_template.Rmd` file to do your work. When completed, please submit your assignment through Github.

- (1) Write a function called `mean_ci()` that takes the arguments listed below and returns either a 95% confidence interval for a single mean or a 95% confidence interval for the difference in means depending on those arguments. I've given default values for these arguments as well.

- `x` (no default value)
- `y` (default value is `NULL` which corresponds to `one.mean` option for the `type` argument)
- `type` (default value is `one.mean`, other value is `two.means`)
- ... (to be able to pass the `na.rm` option through to any functions as needed)

Consider including a way to check whether the vectors of `x` and `y` are numeric and thus such a confidence interval is appropriate.

- (2) Write a function called `prop_ci()` that takes the arguments listed below and returns either a 95% confidence interval for a single proportion or a 95% confidence interval for the difference in proportions depending on those arguments. I've given default values for these arguments as well.

- `x` (no default value)
- `y` (default value is `NULL` which corresponds to `one.mean` option for the `type` argument)
- `type` (default value is `one.prop`, other value is `two.props`)
- ... (to be able to pass the `na.rm` option through to any functions as needed)

Consider including a way to check whether the vectors of `x` and `y` are 0/1 vectors and thus such a confidence interval is appropriate.

- (3) Extra credit: Write a function called `ratio_ci()` that takes the arguments listed below and returns either a 95% confidence interval for the odds ratio or a 95% confidence interval for the relative risk between two groups depending on those arguments. I've given default values for these arguments as well.

- `p1` the sample proportion in group 1 (no default value)
- `n1` the sample size in group 1 (no default value)
- `p2` the proportion in group 2 (no default value)
- `n2` the sample size in group 2 (no default value)
- `type` (default value is `odds.ratio`, other value is `risk.ratio`)

Consider including a way to check whether the values of `p1` and `p2` that are provided are proportions between 0 and 1.

The information below from the formula sheet for Public Health Biostatistics may be helpful as you complete this assignment!

Standard Errors

Mean

$$se_{\bar{x}} = \sqrt{\frac{s^2}{n}}$$

Proportion

$$se_{\hat{p}} = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

Difference of two Means

$$se_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Difference of two Proportions

$$se_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$$

Log Relative Risk

$$se_{LRR} = \sqrt{\left(\frac{1}{\hat{p}_1 n_1} + \frac{1}{\hat{p}_2 n_2}\right) - \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Log Odds Ratio

$$se_{LOR} = \sqrt{\frac{1}{\hat{p}_1 n_1} + \frac{1}{(1 - \hat{p}_1) n_1} + \frac{1}{\hat{p}_2 n_2} + \frac{1}{(1 - \hat{p}_2) n_2}}$$