Standardization

The objective of standardization is the compare the rates of a disease (or outcome) between population which differ in underlying characteristics (age, sex, race, etc) that may affect the overall rate of disease; In essence it's another way to account for confounding.

Standardization of Rates

The difference between crude rates and standardized rates is that the crude rates are calculated on the population under study, whereas standardized rates are based on particular characteristic(s) as standard. If the rates are calculated based on the specific characteristic(s), they are called specific rates (eg. age specific mortality rates, or sex specific mortality rates).

While the MH and Precision based tests assume uniformity of estimates across strata, standardized estimators do not. Thus, these summary measures ignore non-uniformity and mask interactions.

CI of Standardized Rate

Variance:

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

Standardized Rate is a Weighted Average:

$$s(r) = \frac{\sum w_i \hat{p}_i}{\sum w_i}$$

Variance for a Standardized Rate:

$$Var(s(r)) = \frac{\sum w_i^2 Var(\hat{p}_i)}{\left(\sum w_i\right)^2}$$

There are two approaches to Standardization of Rates:

Direct Standardization (Most Common)

$$Adjusted \;\; rate_D = \frac{\sum w_i r_i}{\sum w_i}$$

w_i are the frequency from "standard" population, r_i are rates from testing population. Very similar to the "weighted average" method from the last chapter.

- Stratum-specific rates (r_i) from population under study
- Strata distribution (w_i) from a "standard" population
- **My rates, standard weights**

There are many choices for the 'standard' population.

- Internal Standard - one of the groups under study, total population of the study
- External standard - outside, a larger group (ex. US population)

Direct Standardization Test-Based CI for sRR

Chi-square test for s(r1) = s(r2) is a variation on z test comparing two proportions:

$$\chi^2 = \frac{(s(r_1) - s(r_2))^2}{(Var(s(r_1)) + Var(s(r_2)))}$$

$$CI = sRR^{1 \pm Z_{critical}/\sqrt{\chi^2}}$$

Indirect Standardization

Used when rates in your population are not estimated well due to small sample sizes. We need :

1. Crude rate in population/sample under study
2. Strata distribution of popuation/sample under study

3. Stratum-specific rates (and crude rates) for "standard" population

$$Adjusted\ \ rate_I = \frac{\sum w_i r_i}{\sum w_i}$$

w_i are the frequency from testing population and r_i are rates from "standard" population.

- The stratum-specific rates in the study of interest are not reliable due to small sample size.
- Stratum-specific weights (wi) from population under study
- Strata distribution (ri) from "standard" population
- **My weights, standard rates**

Results are often presented as a **Standardized Mortality Ratio (SMR):**

$$SMR = \frac{O\ (\text{Observed number of deaths})}{E\ (\text{Expected number of deaths})}$$

We can also calculate the standardized rate based on the standard population rate and teh target cohort-specific SMR:

s(r) = (Population Rate) x SMR

Indirect Standardization Test-Based CI for sRR

CI for an observed count using the Poisson model (SE = sqrt(mean)):

$$X \pm z_{crit}\sqrt{X}$$

Standardized Estimators

All summary measures ignore non-uniformity and mask interaction, very similar to last chapter's tests for confounding with weights.

$$\hat{\theta} = \frac{\sum_{\text{strata}} w_g \hat{\theta}_g}{\sum_{\text{strata}} w_g}$$

$$\hat{\theta}_g = \text{RR or OR}$$

	RR	OR
general form of weight	$s\ b/n_0$	$s\ b/d$
Internal	$s = n_1$	$s = c$
External	$s = n_0$	$s = d$

> 's' can also be weights (frequencies) derived from a standard population.

For RR:

- Internally Standardized
  - s = n_1 = exposed subjects sRR
- Externally Standardized
  - s = n_0 = unexposed subjects s'RR

For OR:

- Internally Standardized
  - s = c = exposed non-cases sOR
  - The interpretation of the OR would be something like: "If study participants who were not exposed to X have the same distribution of a confounding variable as those exposed to X who did not develop Y, then the increased odds of Y is OR for study participants exposed to X."
- Externally Standardized
  - s = d = unexposed non-cases s'OR
  - The interpretation of the OR would be something like: "If study participants have the same distribution of a confounding variable as those not exposed to X who developed Y, then the increased odds of Y is OR for study participants exposed to X."

Equivalent Estimators to the Standardized Risk Ratio:

- Directly standardized risk ratio with unexposed used as standard population
- Externally standardized RR
- SMR with unexposed as standard population
- Internally standardized RR

Test for Linear Trend in Odds Ratios

Test with several exposures, 2 * k tables (for association), or a series of 2 * 2 tables (for effect modification/interaction).

Note: Odds ratios are sensitive to non-linear transformations, but non-linear transformations can alter results.

Mantel Extension Chi-Square for 2 * k Tables

Let X = an ordinal score for exposure

H_0: No linear (monotone) association between exposure and disease (proportion)

$$\chi^2 = \frac{\left[\sum_{strata}\sum_{exposures} a_i x_i - \sum_{strata}\left(\frac{m_1}{n}\sum_{exposures} n_i x_i\right)\right]}{\sum_{strata} Var\left(\sum_{exposures} a_i x_i\right)}$$ with df = 1

And each stratum has variance:

$$Var\left(\sum_{exposures} a_i x_i\right) = \frac{m_1 m_0}{n(n-1)}\left[\sum_{exposures} n_i x_i^2 - \frac{\left(\sum_{exposures} n_i x_i\right)^2}{n}\right]$$

Conclusion: We find a [de/in]creasing trend in incidence of [Outcome] with increasing levels of [independent] ($p < .05$), with ORs changing from OR_before to OR_after.

Test for Linear Trend in Odds Ratio for Several 2x2 Tables

The model we use is logarithmic: $\ln(OR) = b_0 + b_1 X$

We are testing H_0: b_0 = Odds Ratios for all tables OR All Odds Ratios are equal

Measure if there is an interaction present (ORs not all the same), or modelling interaction (ORs

$$\chi^2 = \frac{\left[\sum_{strata} X\left[a - E(a)\right]\right]^2}{\sum_{strata} X^2\, Var\,(a) - \frac{\left[\sum_{strata} X\, Var\,(a)\right]^2}{\sum_{strata} Var\,(a)}}$$ with df = 1

Where X is the score, and E(a) is the expected score under H_0. E(a) and Var(a) can be computed as in the Test for homogeneity of Odds Ratios (Breslow-Day):

$$a' = E(a) = \frac{t \pm \sqrt{t^2 - 4(a\hat{O}R - 1)(a\hat{O}R)m_1 n_1}}{2(a\hat{O}R - 1)}$$
$$\text{with } t = a\hat{O}R(m_1 + n_1) + m_0 - n_1$$

$$Var(a) = \left(\frac{1}{a'} + \frac{1}{b'} + \frac{1}{c'} + \frac{1}{d'}\right)^{-1}$$
$$= \left(\frac{1}{a'} + \frac{1}{m_1 - a'} + \frac{1}{n_1 - a'} + \frac{1}{m_0 - n_1 + a'}\right)^{-1}$$

Revision #6
Created 26 September 2022 14:06:50 by Elkip
Updated 26 September 2022 16:48:26 by Elkip