Algorithms for Rates

Crude Rate

A crude rate is the number of cases per 100,000 in a given population.

$$cruderate = \frac{count}{population} \times 100{,}000$$

Age-Adjusted Rate

An age adjusted rate is a weighted average of crude rates, where the crude rates are calculated for different age groups and the weights are the proportions of persons in the corresponding age groups of a standard population. Several sets of standard populations are included in SEER*Stat. These include the total U.S. populations (1940, 1950, 1960, 1970, 1980, and 1990), an estimate of the U.S. 2000 population, 1991 Canadian population, and the world population. The age-adjusted rate for an age group comprised of the ages x through y is calculated using the following formula:

$$aarate_{x-y} = \sum_{i=x}^{y} \left[\left(\frac{count_i}{pop_i} \right) \times 100{,}000 \times \left(\frac{stdmil_i}{\sum_{j=x}^{y} stdmil_j} \right) \right]$$

where counti is the number of cases for the ith age group, popi is the relevant population for the same age group, and stdmili is the standard population for the same age group.

Standard Error for a Crude Rate

This calculation assumes that the cancer counts have Poisson distributions.

$$SE_{crude} = \frac{\sqrt{count}}{population} \times 100{,}000$$

Standard Error for an Age-Adjusted Rate

This calculation assumes that the cancer counts have Poisson distributions. Suppose that the age-adjusted rate is comprised of age groups x through y.

$$SE_{AArate} = \left[\sum_{i=x}^{y} \left(\frac{stdmil_i}{\sum_{j=x}^{y} stdmil_j} \right)^2 \times \left(\frac{count_i}{population_i^2} \right) \right]^{\frac{1}{2}} \times 100{,}000$$

Crude Rate Confidence Intervals

The endpoints of a p ? 100% confidence interval are calculated as:

$$CI_{low} = \frac{\left(\frac{1}{2} \left(ChiInv\left(\frac{p}{2}, 2 \times count \right) \right) \right)}{population} \times 100{,}000$$

$$CI_{high} = \frac{\left(\frac{1}{2} \left(ChiInv\left(1 - \frac{p}{2}, 2 \times (count + 1) \right) \right) \right)}{population} \times 100{,}000$$

where Chi Inv(p,n) is the inverse of the chi-squared distribution function evaluated at p and with n degrees of freedom, and we define Chi Inv (p,0) = 0.

Although the normal approximation may be used with the standard errors to obtain confidence intervals when the count is large, we use the above exact method that holds even with small counts [see Johnson and Kotz, 1969, or Fay and Feuer, 1997]. When the count is large the 2 methods produce similar results.

Age-Adjusted Rate Confidence Intervals

Suppose that the age-adjusted rate is comprised of age groups x through y, and let:

$$w_i = \frac{stdmil_i}{\left(pop_i \times \sum_{j=x}^{y} stdmil_j \right)}$$

$$w_m = \max\left(w_i \right)$$

$$v = \sum_{i=x}^{y} \left(w_i^2 \times count_i \right)$$

The endpoints of a p x 100% confidence interval are calculated as:

$$CI_{low} = \left(\frac{v}{2 \times rate} \right) \times \left(ChiInv\left(\frac{p}{2}, \frac{\left(2 \times rate^2 \right)}{v} \right) \right) \times 100,000$$

$$CI_{high} = \left(\frac{v + w_m^2}{2\left(rate + w_m \right)} \right) \times \left(ChiInv\left(1 - \frac{p}{2}, \frac{2\left(rate + w_m \right)^2}{\left(v + w_m^2 \right)} \right) \right) \times 100,000$$

This method for calculating the confidence interval was developed in Fay and Feuer (1997). The method produces similar confidence limits to the standard normal approximation when the counts are large and the population being studied is similar to the standard population. In other cases, the above method is more likely to ensure proper coverage.

Note

The rate used in the above formulas is not per 100,000 population.