Inference re Epidemiologic Parameter: Rate or Incidence Density

Theoretical: Rate or ID or \( \lambda \)

Empirical: \( c \) cases in PT population-time units: \( \hat{ID} = \hat{\lambda} = c / PT \)

Model: \( C \sim \text{Poisson}(\mu), \mu = \lambda \times PT \); \( c \) : realization of \( C \)

P-value:

\[ H_0: ID = ID_0, \lambda = \lambda_0; \rightarrow \mu = \lambda_0 \times PT \equiv \mu_0; \]

Exact
\( P[C \leq c & C \geq c | \mu_{null}] \) (lower & upper-tails)

Approx.
\( \hat{N} \) 1 Approx. to distr’n of \( C \) or transform, \( t(C) \), of \( C \)

CI 100(1-\( \alpha \))%:

Exact
\( P[C \geq c|\mu_L] = \alpha/2; \ P[C \leq c|\mu_U] = \alpha/2; \rightarrow \{\mu_L, \mu_U\} / PT \)

Approx.
\( \cdot \) reverse transform of \( ci = \{t(c) \equiv z_{\alpha/2} \times SE[t(c)]\}; ci \div PT \)

\begin{align*}
\text{transform} & \quad t(c) & SE[t(c)] = Var^{1/2} & \quad CI = \text{reverse of } ci \\
\text{identity} & \quad c & c^{1/2} & n/a \\
\log & \quad \log[c] & (1/c)^{1/2} & \exp(ci) \text{ or } \exp\{ci\} \\
\sqrt{c} & \quad \sqrt{c} & (0.25)^{1/2} = 0.5 & \{c \div 0.5z_{\alpha/2}\}^2 \\
\end{align*}

Notes:

[1] “Rate” in the incidence density sense.

[2-4] helps to separate obs’d & exp’d numerator, \( \mu \) & \( c \), from rate (\( ID, \lambda, \hat{\lambda} \)); \( \mu = \lambda \times PT \)

[5] Interested in \( \lambda \), not \( \mu \), but it is \( C \) which has the Poisson distribution!

[6] ppois(c,\( \mu \)) & 1-ppois(c-1,\( \mu \)) in R; Poisson(\( \mu \),\( X \)) in Excel; See c634 Resources.

[11] via- Tables; cii PT c, poisson in Stata; poiss.exact in epitools in R; etc.

[11] via trial & error using ppois or Poisson; or exactly using Poisson \( \Rightarrow \) Chi-sq link.

[13] Note that \( CI \) for \( \lambda \) or \( ID \) is of the form \( \{CI for \mu\} / PT \).

[14] Using ‘\( t() \)’ as shorthand for a generic ‘transform’ \& ‘function of’.

[15] These transforms will be called ‘\( \text{links} \)’ when we come to generalized linear models.

[16] The variance of a Poisson random variable is a function only of the mean \( \mu \).

[17] Rothman2002 (‘BabyRothman’) p132 uses identity link, i.e. untransformed version.

[18] Log link typically used for rate ratio; makes more sense than identity link for 1 rate.


[1] Comparison of ID’s or Rates in index (\( 1 \)) vs. reference (\( 0 \)) category

[2] Theoretical: \( \lambda_1 \& \lambda_0 \rightarrow \lambda_1 - \lambda_0 \) (IDD); \( \lambda_1 \div \lambda_0 \) (IDR)

[3] Empirical: \( c_1/PT_1 \& c_0/PT_0 \rightarrow \lambda_1 - \lambda_0; \ \lambda_1 \div \lambda_0 \)

[4] Model: \( c_i \sim \text{Poisson}(\mu_i = \lambda_i \times PT_i), i = 0, 1; c_1 \) independent of \( c_0 \).

[5] P-value: \( H_0: ID = 0; IDR = 0 \)

Exact
\( \cdot \) 1 \( c_1(c_1 + c_0) \sim \text{Binomial}(\pi'' = c_1 + c_0, \pi = PT_1/(PT_1 + PT_0)) \)

Approx. \( \cdot \) using the \( \hat{\lambda} \)'s, or transforms, \( t(\hat{\lambda}) \), of them

\( z = \sqrt{X^2} = \{t(\hat{\lambda}) - t(\hat{\lambda}_0)\}/\{Var_{H_0}[t(\hat{\lambda}_1)] + Var_{H_0}[t(\hat{\lambda}_0)]\}^{1/2} \)

[10] CI:

Exact – IDR only
\( \cdot \) IDR: \( P[C \geq c_1(IDR_L) = \alpha/2; \ IDR_U : - similarly \)

Approx. – both IDD and IDR: ci on t scale \( \rightarrow CI \) on desired scale
\( \cdot \) ci: \( \{t(\hat{\lambda}_1) - t(\hat{\lambda}_0) \pm z_{\alpha/2}(Var[t(\hat{\lambda}_1)] + Var[t(\hat{\lambda}_0)])^{1/2} \} \rightarrow CI \)

[15] Measure \quad \text{transform} \quad t(\hat{\lambda}) \quad ci \rightarrow CI \quad \text{test-based* CI}

ID Diff. \quad identity \quad \hat{\lambda} \quad n/a \quad \hat{IDD} \times (1 \pm z_{\alpha/2}/X)

ID Ratio \quad log \quad log[\hat{\lambda}] \quad e^{ci} \quad IDR^{(1 \pm z_{\alpha/2}/X)}

Notes:


[7] [11] Fixing \( c_1 + c_0 \) eliminates 1 nuisance parameter leaving just the ratio IDR = \( \lambda_1/\lambda_0 \).


... Don’t force \( c_1, c_0, PT_1, PT_0 \) into a \( 2 \times 1 \) table. See depiction as a \( 2 \times 1 \) table (jh Ch9).

[11] Use Binomial distr’n; \( \pi \) is determined by (is function of) the IDR & ratio of the PT’s.

... Use def’n. of \( \mu \)’s to show: \( \pi = c_1(\mu_1 + \mu_0) = IDR/(IDR + PT_0/PT_1) \)

... Lower limit \( \pi_1 \) for \( \pi \rightarrow \) lower limit for IDR: IDR_L = \( \{\pi_1/(1 - \pi_1)\} \div \{PT_1/PT_0\} \) etc.

... Can use same Excel spreadsheet (jh c607 ch8 resources) for exact test and exact CI.

... Via Stata immediate command iri c1 c0 PT; PT0 or rateratio.* in epitools

[14] Can also use a test-based CI for a risk difference/ratio or an odds ratio.

... Test-based CI’s use the Variance under the Null, used when testing the null value.

[17] Var[log(IDR)] had just 2 terms, \( 1/c_1 + 1/c_0 \). Since \( PT \) is just a constant,

\begin{align*}
\text{Variance under the Null, used when testing the null value.} \\
\text{Var[log(IDR)]} & \text{ had just 2 terms, } 1/c_1 + 1/c_0. \text{ Since } PT \text{ is just a constant,} \\
\text{Variance-stabilizing transformation.} \\
\end{align*}
Comparison of ID’s (Rates, λ’s) in index₁ vs. ref(₀) categories - **stratified data**

Empirical:  
\( c₁,ₙ \) cases in \( PT₁,ₙ \) and \( c₀,ₙ \) cases in \( PT₀,ₙ \) stratum ‘s’ \((s = 1, \ldots, S)\)  

ID **Ratio** (IDR): aliases: Rate Ratio; Incidence Ratio (‘IR’) - Rothman’s term

(1)  
\[
\text{Antilog of weighted average (}Wⁿ\text{Ave.}) \text{ of stratum-specific } log(I\overline{DR})'s
\]

Weights \( \{w₁, w₂, \ldots, wₙ, \ldots, w_S\} \) are precision-based  
\[
w_s = \frac{1/V_s}{1/V_1 + \ldots + 1/V_S}; \quad V_s = \text{Var}[log(I\overline{DR}_s)] = 1/c₁,ₙ + 1/c₀,ₙ
\]

**Point Estimate:** \( \hat{I\overline{DR}} = \exp [\sum_s w_s \times \log(I\overline{DR}_s)] = \exp [Wⁿ\text{Ave.}] \)

Variance of \( Wⁿ\text{Ave.} \):  
\[
\text{Var} = 1/\{\sum_s 1/V_s\}; \quad \text{SE} = \text{Var}^{1/2}
\]

**CI:**  
\[
\text{exp}[Wⁿ\text{Ave} \mp zα/2 \times \text{SE}[Wⁿ\text{Ave}]] = \hat{I\overline{DR}} \mp \exp[zα/2 \times \text{SE}]
\]

(2)  
**Mantel-Haenszel Summary IDR** See Rothman2002Ch8

No. cases, PT in stratum s: \( c_s = c₁,ₙ + c₀,ₙ \); \( PT_s = PT₁,ₙ + PT₀,ₙ \)

**Point Estimate:** \( \hat{I\overline{DR}}_{MH} = \frac{\sum_s (c₁,ₙ \times PT₀,ₙ) / PT₁,ₙ}{\sum_s (c₀,ₙ \times PT₁,ₙ) / PT₀,ₙ} = \frac{\text{Num}_{MH}}{\text{Den}_{MH}} \)

Variance of \( \log(I\overline{DR}_{MH}) \):  
\[
\text{Var} = \frac{\sum_s (c₁,ₙ \times PT₀,ₙ) / PT₁,ₙ \times PT₀,ₙ}{\sum_s (c₀,ₙ \times PT₁,ₙ) / PT₀,ₙ \times PT₀,ₙ} \times \frac{\text{Num}_{MH} \times \text{Den}_{MH}}{\text{Num}_{MH} \times \text{Den}_{MH}}
\]

**CI:**  
\[
\text{exp} \left[ \log(I\overline{DR}_{MH}) \mp zα/2 \times \text{Var}^{1/2} \right] = \hat{I\overline{DR}}_{MH} \mp \times \text{M.E.}’
\]

ID Diff. (IDD): aliases: Rate Diff.; Incidence Difference (‘ID’) - Rothman’s term

Precision-weighted average of stratum-specific \( \hat{IDD}’s \) See Rothman2002Ch8

\[
w_s = \frac{Q_s}{Q₁ + \ldots + Qₙ}; \quad Q_s = 1/(1/PT₁,ₙ + 1/PT₀,ₙ); \quad V_s = \frac{c₁,ₙ}{PT₁,ₙ} + \frac{c₀,ₙ}{PT₀,ₙ}
\]

**Point Estimate:** \( \hat{IDD} = \sum_s w_s \times \hat{IDD}_s \)

**CI:**  
\[
\hat{ID} \mp zα/2 \times \text{SE} \times \text{SE} = \text{Var}^{1/2}; \quad \text{Var} = \sum_s w_s^2 \times V_s.
\]