# Mantel-Haenszel Type Estimators and Cohort Sampling Schemes

**Larry Goldstein and Bryan Langholz**

University of Southern California

# Classical Mantel-Haenszel Estimator

Group $j \in \{0, 1\}$ has $n_j(t)$ individuals at time $t$, exposure $Z = j$, failure rate $\lambda_0(t)\phi_0^Z$ and failure times $\mathcal{T}_j$. Let

$$R_{jk} = \sum_{t \in \mathcal{T}_j} \frac{n_k(t)}{n_0(t) + n_1(t)},$$

the classical Mantel-Haenszel estimator of $\phi_0$ is *explicitly* given by

$$\hat{\phi} = \frac{R_{10}}{R_{01}}.$$

# Features

1. Self Evident Model. Baseline is unspecified, and the exposed group has additional relative risk $\phi_0 \in (0, \infty)$ over those unexposed.
2. Simple Calculation. Estimator $\hat{\phi}$ can be computed easily, a 'back of the envelope calculation'.

$J_2 = \{.09, 1.8, 3.07\}$ $J_0 = \{1.7, 2.8\}$

$n_0(1.7) = 5$, $n_1(1.7) = 3$

$n_0(2.8) = 2$, $n_1(2.8) = 1$

$n_0(.09) = 7$, $n_1(.09) = 6$, $n_0(1.8) = 3$

$n_0(1.8) = 5$, $n_0(3.07) = 2$, $n_1(3.07) = 2$

$$R_{01} = \frac{3}{5+3} + \frac{1}{2+1} = \frac{3}{8} + \frac{1}{3} = \frac{17}{24}$$

$$R_{10} = \frac{7}{7+6} + \frac{3}{3+5} + \frac{2}{2+2}$$

$$= \frac{7}{13} + \frac{3}{8} + \frac{2}{4} = \frac{56+39+52}{104}$$

$$= \frac{147}{104} \therefore \phi = \frac{R_{10}}{R_{01}} = \frac{147/104}{17/24}$$

$$\approx (1.4)^2 \cong 2$$

# Other Features

At the null, the Mantel-Haenszel estimator is as efficient as the Maximum Partial Likelihood Estimator (MPLE), which is more complicated to calculate.

Less efficient away from the null.

# For Better or For Worse

1. Medline search of papers in the years 2000-2005 gives a total of 420 references where Mantel-Haenszel is cited in the abstract as the method applied.
2. From page 156 of Kahn and Sempos, “when a method is as simple and free of assumptions as the Mantel-Haenszel procedure, it deserves a strong recommendation, and we do not hesitate to give it.”

The estimator is well used; can we contribute something so that it is better used?

# Using Mantel-Haenszel with Complex Event History Data and Complex Sampling

Zhang, Fujii, and Yanagawa (2000), and Zhang (2000), consider some sampling designs for the Mantel-Haenszel estimator.

Can we generalize enough to accommodate

1. Histories more complex than binary exposure?
2. General framework for complex sampling?

Then we need to look to see if the allowances for all this additional complexity has made the estimator too complex.

# More Complex Exposure

Consider the case when $Z \in \{\alpha_0, \ldots, \alpha_\eta\}$ with

$$\alpha_0 < \ldots < \alpha_\eta,$$

and by absorbing a factor into $\lambda_0(t)$ if necessary, without loss of generality, $\alpha_0 = 0$.

Can make the connection to the exponential relative risk model for a one-dimensional covariate $Z$ by letting $\beta_0 = \log \phi_0$, since then

$$e^{\beta_0 Z} = \phi_0^Z.$$

# Extension of Mantel-Haenszel estimator to multiple exposure levels

In the classical two level case, view

$$\hat{\phi} = \frac{R_{10}}{R_{01}}$$

as the solution to the (linear! martingale?) estimating equation

$$\phi R_{01} - R_{10} = 0,$$

or as the minimizer of $G_{01}^2$, where with $j < k$, we let

$$G_{jk} = \phi^{\alpha_k} R_{jk} - \phi^{\alpha_j} R_{kj}.$$

# One Proposal, Multiple Exposure Levels

Define the estimator as a value $\hat{\phi}_n$ which minimizes a weighted sum of squares of the form

$$n^{-1} \sum_{j<k} c_{jk} G_{jk}^2(\phi).$$

Weights $c_{jk}$ can be chosen to take into account factors such as varying group sizes, or even, with enough information (asymptotically), to minimize the variance among all estimators of this form.

But what about when the full cohort is too large for data collection, and must be sampled? What are the extensions (even for the classical two level case)?

# Main Interest: More Complex Sampling

From Borgan, Goldstein, Langholz (1995), consider sampling designs as particular specifications of

$$\pi_t(\mathbf{r}|i),$$

the probability of choosing $\mathbf{r}$ as the sampled risk set should $i$ fail at time $t$.

BGL developed a theory for such designs in general, and studied particular examples, for the partial likelihood estimator and the model

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp(\beta_0 \mathbf{Z}_i(t)).$$

# Simple Random Sampling Design

Take a sampled risk set of size $m$, including the failure, from $\mathcal{R}(t)$, those at risk at time $t$, uniformly over all such sets. So, with $n(t) = |\mathcal{R}(t)|$,

$$\pi_t(\mathbf{r}|i) = \binom{n(t)-1}{m-1}^{-1} \mathbf{1}(\mathbf{r} \ni i, \mathbf{r} \subset \mathcal{R}(t), |\mathbf{r}| = m)$$

# Matching Design

When $\mathcal{R}(t)$ is the disjoint union of strata $\bigcup_l \mathcal{C}_l(t)$, we take a specified number of controls from the case's strata (controls for confounder).

If $C_i(t)$ is strata of $i$ at time $t$, $c_i(t) = |\mathcal{C}_i(t)|$, and we specify $\mathbf{m} = (m_l)_{l \in \mathcal{C}}$, the matching design is given by

$$\pi_t(\mathbf{r}|i) = \binom{c_{C_i(t)}(t) - 1}{m_{C_i(t)} - 1}^{-1} \mathbf{1}(\mathbf{r} \subset \mathcal{C}_{C_i(t)}(t), \mathbf{r} \ni i, |\mathbf{r}| = m_{C_i(t)})$$

# Counter Matching Design

Want to have the sampled risk set be 'representative' in the strata variables, so take $\mathbf{r}$ uniformly over $\mathbf{r} \ni i, \mathbf{r} \subset \mathcal{R}(t), |\mathbf{r} \cap \mathcal{R}_l(t)| = m_l$; $l \in \mathcal{C}$, that is, with probability

$$\pi_t(\mathbf{r}|i) = \left[ \prod_{l \in \mathcal{C}} \binom{n_l(t)}{m_l} \right]^{-1} \frac{n_{C_i(t)}(t)}{m_{C_i(t)}}.$$

... *and others* ...

# Counting Process Framework

Following the pioneering work of Andersen and Gill (1982), by placing the problem in a counting process framework, in particular, by considering

$$N_i(t) \quad \text{and} \quad N_{i,\mathbf{r}}(t),$$

their intensities, and the associated martingales, BGL was able to give general conditions under which MPLE was consistent and asymptotically normal.

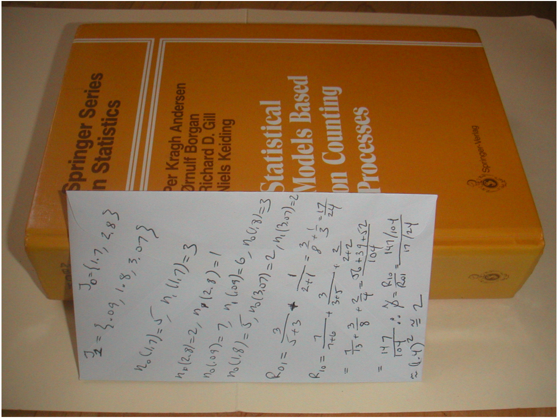
And the same for an estimate of the integrated baseline hazard.

# Mantel-Haenszel Estimator

Can we do the same for the Mantel-Haenszel estimator?

We first need to see if we can support the Mantel-Haenszel calculations on the same solid foundation.

# Yes, the calculations can rest on the theory



# Mantel-Haenszel in Counting Process Framework

Exposure groups $\mathcal{R}_k(t), k = 0, \ldots, \eta$ at time $t$, sampling scheme $\pi_t(\mathbf{r}|i)$. Define

$$A^k_{\mathbf{r}}(t) = \sum_{i \in \mathcal{R}_k(t)} \pi_t(\mathbf{r}|i) \quad k = 0, \ldots, \eta,$$

and for a given continuous function $a : \mathbf{R}^{\eta+1} \to [0, \infty)$, define

$$a_{\mathbf{r}}(t) = a(A^0_{\mathbf{r}}(t), \ldots, A^\eta_{\mathbf{r}}(t));$$

the choice which extends the classical case is

$$a(u_0, \ldots, u_\eta) = (u_0 + \cdots + u_\eta)^{-1}.$$

# General (in terms of both exposure and sampling) Estimator

With

$$R_{jk}(t) = \int_0^t \sum_{\mathbf{r}\subset\mathcal{R}} a_{\mathbf{r}}(s) A_{\mathbf{r}}^k(s) dN_{\mathbf{r}}^j(s),$$

generalization of classical estimator is one which minimizes the magnitude (e.g. sum of squares) at $t = \tau$ of the local square integrable martingales

$$G_{jk}(t) = \phi^{\alpha_k} R_{jk}(t) - \phi^{\alpha_j} R_{kj}(t).$$

By resting on existing theory, can estimate integrated baseline hazard as in BGL.

# Analysis and Results

For $\mathbf{v}$ a multi-subset of $\{0, \ldots, \eta\}$, e.g. $\mathbf{v} = \{0, 0, 1\}$, let

$$H_{\mathbf{v}}(t) = \sum_{\mathbf{r} \subset \mathcal{R}} a_{\mathbf{r}}^{|\mathbf{v}|-1}(t) \prod_{k \in \mathbf{v}} A_{\mathbf{r}}^{k}(t).$$

Assume

$$\frac{1}{n} H_{\mathbf{v}}(t) \to h_{\mathbf{v}}(t), \quad \text{and let}$$

$$I_{\mathbf{v}}(t) = \int_0^t h_{\mathbf{v}}(s) \lambda_0(s) ds \text{ and } \beta_{jk} = (\alpha_k - \alpha_j) \phi_0^{\alpha_k + \alpha_j - 1} I_{jk}(\tau)$$

All are consistently estimable

# Asymptotic

**Theorem 1**

$$\sqrt{n}\left(\hat{\phi}_n - \phi_0\right) \rightarrow_d \mathcal{N}(0, \sigma^2) \quad \textit{where} \quad \sigma^2 = v^2/\gamma^2,$$

*with*

$$v^2 = \sum_{j<k, p<q} c_{jk}\beta_{jk} < g_{jk}, g_{pq} >_\tau \beta_{pq} c_{pq},$$

*and*

$$\gamma = \sum_{j<k} c_{jk}\beta_{jk}^2,$$

*and* $g_{jk}(t)$ *are scaled limits of* $G_{jk}(t)$*.*

# Optimization

$B$ with diagonal entries $\beta_{jk}$, $\Gamma$ with entries $< g_{jk}, g_{pq} >_\tau$ positive definite, and $B\Gamma B = M'M$.

**Proposition 1** *Let* $\mathbf{1}$ *be the vector all of whose entries are 1 and*

$$X = (M^{-1})'B^2\mathbf{1}\mathbf{1}'B^2M^{-1}.$$

*Then taking*

$$\mathbf{c} = M^{-1}\mathbf{d}$$

*where* $\mathbf{d}$ *is any eigenvector corresponding to the largest eigenvalue* $\lambda$ *of* $X$ *minimizes the asymptotic variance at the value* $\sigma^2 = \lambda^{-1}$.

In accommodating this much complexity, the estimator is now a bit more complex.

What about the two exposure group setting?

# Two Exposure Groups, General Complex Sampling

Even under complex sampling, the estimator, so generalized, retains its simple form,

$$\hat{\phi}_n = \frac{R_{10}}{R_{01}} = \frac{n^{-1}R_{10}}{n^{-1}R_{01}} \rightarrow_p \phi_0, \quad \text{and}$$

$$\sqrt{n}\left(\hat{\phi}_n - \phi_0\right) \rightarrow_d \mathcal{N}(0, \sigma^2) \quad \text{as } n \rightarrow \infty,$$

where

$$\sigma^2 = \frac{\int_0^\tau \left(\phi_0^2 h_{011}(t) + \phi_0 h_{100}(t)\right)\lambda_0(t)dt}{\left(\int_0^\tau h_{01}(t)\lambda_0(t)dt\right)^2}.$$

# Consequences

The Mantel-Haenszel Estimator is expressed as before in a simple closed form, even for complex sampling, in the classical case of two exposure groups.

Confine ourselves to this case henceforth.

Efficiencies for the various designs?

# Simple Random Sampling

Under the null, we have

$$\sigma^2 = \left(\frac{m}{m-1}\right) \frac{1}{\int_0^\tau p(t) f_0(t) f_1(t) \lambda_0(t) dt},$$

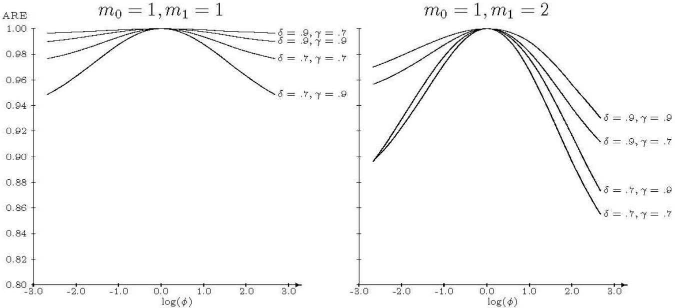
giving an asymptotic relative efficiency of $(m-1)/m$ with respect to the full cohort variance, same as the partial likelihood.

# Counter Matching

When the design matches one control with binary 'surrogate exposure' $C(t)$ of the binary exposure $Z(t)$ to the value opposite that of the case, the asymptotic variance is equal to that for this same counter matching design when using the maximum partial likelihood estimator.

What about away from the null?

# ARE relative to MPLE




Sensitivity $\delta = P(Z(t) = 1 | C(t) = 1, Y(t) = 1)$

Specificity $\gamma = P(Z(t) = 0 | C(t) = 0, Y(t) = 1)$

# Conclusion

The Mantel-Haenszel estimator can be generalized, both in terms of exposure variable, and to accommodate general sampling schemes. In the classical case of two exposure variables, it has the same efficiency, under complex sampling, as the MPLE at the null, and nearly so around the null.

# Conclusion

Due to its simplicity, the Mantel-Haenszel estimator will, in all ‘likelihood’, continue to be used. We hope this contribution allows the use of this estimator to be used in conjunction with sampling schemes that can help produce more accurate estimates than it would have otherwise.

The calculation remains simple, but we have added some power to its use:

# The Mantel-Haenszel Estimator with greater power