Parameter collapsibility and noncollapsibility in statistical models: what you need to know

Tom Palmer

18th January 2021

Outline

What is collapsibility?

- Parameter/measure: risk difference, risk ratio, odds ratio, hazard ratio, etc.
- Greenland and Pearl (2011) define collapsibility as

  When an adjustment does not alter a measure, the measure is said to be collapsible over C or invariant with respect to the adjustment. Conversely, if an adjustment alters a measure, the measure is said to be non-collapsible over C.

- Originally used in the context of collapsing over a third variable in contingency tables
- Sometimes referred to as Simpson's paradox (Simpson 1951)
  - A dramatic example with different signs
- First use of the term by Bishop et al. (1975)

Example 1

- From the Introduction of Daniel et al. (2020)
  - X binary exposure
  - Y binary outcome
  - C third (in this case continuous) variable
  - OR between Y and X given C set at 10

Example 1

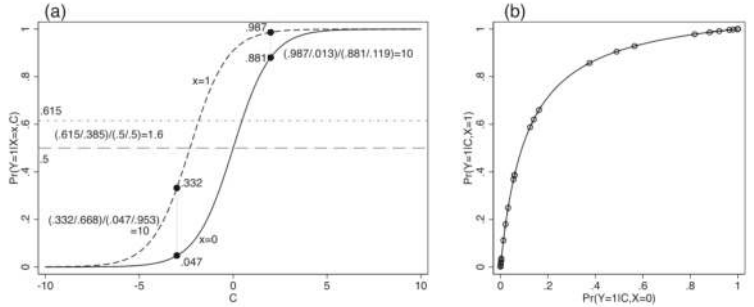
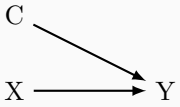


FIGURE 1 (a) A simple logistic regression model, and (b) the relationship between the conditional probability of $Y = 1$ given C between the two treatment groups as implied by the model in (a). Thirty randomly chosen values of $C \sim U[-10, 10]$ give rise to the superimposed scatter plot in (b)

- Take-away: the marginal OR can be very different to the conditional OR

Example 1

- The plot on the RHS is non-linear
- Data simulated as



- which will turn out to be important because neither $Y \perp\!\!\!\perp C|X$ nor $X \perp\!\!\!\perp C|Y$ hold

Example 2

- An example with a binary third variable *Z* comparing RDs, RRs, ORs from Greenland et al. (1999)

TABLE 1
Examples of collapsibility and noncollapsibility in a three-way distribution

	$Z=1$		$Z=0$		Marginal	
	$X=1$	$X=0$	$X=1$	$X=0$	$X=1$	$X=0$
$Y=1$	0.20	0.15	0.10	0.05	0.30	0.20
$Y=0$	0.05	0.10	0.15	0.20	0.20	0.30
Risks[a]	0.80	0.60	0.40	0.20	0.60	0.40
Risk differences	0.20		0.20		0.20	
Risk ratios	1.33		2.00		1.50	
Odds ratios	2.67		2.67		2.25	

[a]Probabilities of $Y=1$.

- In this example the two conditional RDs and marginal RD are all equal

Example 2

- This is an example of strict/strong collapsibility

- In this example a weighted average of the conditional RRs obtain the marginal RR
  - $(0.25 \times 2) + (0.75 \times 1.333) = 1.5$
- Here the RR is said to be collapsible across/over Z
  - The measure is not constant across strata, but a particular summary of the conditional measures equals the marginal measure

- But for the OR there is no positive set of weights for which $w_1 2.67 + w_2 2.67 = 2.25$
  - hence the OR is noncollapsible

What causes non-collapsibility?

- Quick refresher
  - GLMs: where $f(p)$ is the link function and η the linear predictor, $f(\mathbb{E}[Y]) = \eta$
  - e.g. logistic regression uses the logit link; $\log\left(\frac{p}{1-p}\right) = \eta$
- Alot of people guess (not unreasonably) it's the nonlinearity or non-separability of the link function
- Neuhaus and Jewell (1993) explained that where v is the conditional association between X and Y (on the scale of the linear predictor)
- Characteristic collapsibility function (CCF) $g_v(p)$, named by Daniel et al. (2020), maps $P(Y|X = 0, C)$ to $P(Y|X = 1, C)$, e.g. part (b) of first figure

What causes non-collapsibility?

$$\text{CCF: } g_v(p) = f^{-1}\{f(p) + v\}$$

- Take away: **A parameter is said to be non-collapsible when its CCF is non-linear**

What causes non-collapsibility?

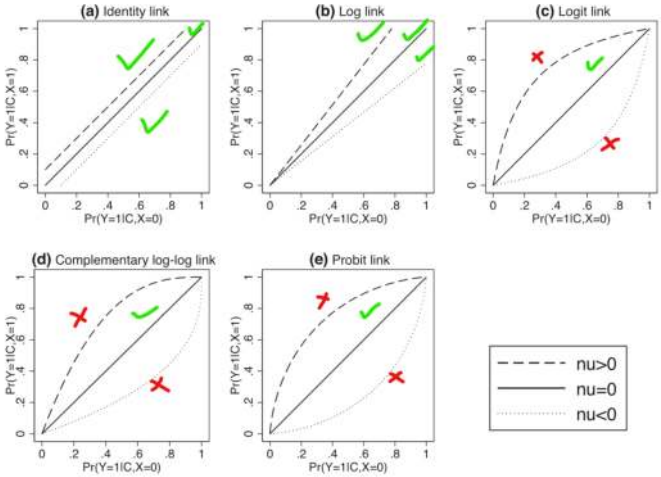


FIGURE 2 The (non)collapsibility of common effect measures for binary outcomes as determined by the concavity, convexity, or linearity of the characteristic collapsibility function (CCF) $g_\nu(\cdot) = f^{-1}\{f(\cdot) + \nu\}$, where f is the link function and ν is the conditional effect measure. $f(p) = \log[p/(1-p)]$ for the logit link, $f(p) = \log(-\log(1-p))$ for the complementary log-log link, and $f(p) = \Phi^{-1}(p)$, where $\Phi(\cdot)$ is the CDF of the standard normal distribution, for the probit link

Collapsibility over *C*

- Greenland and Pearl (2011) discussed that

(a) If *C* is separated from *Y* given *X* then *Y*-*X* will be (completely) collapsible over *C*, i.e. effectively $Y \perp\!\!\!\perp C|X$

(b) If *C* is separated from *X* unconditionally, the population standardized measured of *Y*-*X* will be collapsible over *C*

Collapsibility over C

(c) If C is separated from X conditional on Y then the OR_{YX} will be collapsible over C, effectively $X \perp\!\!\!\perp C|Y$

$$X \rightarrow Y \rightarrow C$$

- Didelez et al. (2010) additionally discussed that if we have an (outcome dependent) selection variable S
- $OR_{YX}(C, S)$ is collapsible over S iff $Y \perp\!\!\!\perp S|(C, X)$ or $X \perp\!\!\!\perp S|(C, Y)$
- A consequence of this is in a matched case-control study, OR_{YX} is collapsible if we condition on the matching variables

Collapsibility over *C*

FIG. 8. *DAG for a matched (on B) case-control study* (left) *and moral graph* (right).

- i.e. $OR_{YX}(B, S) = OR_{YX}(B)$
- So we can only ever recover OR_{YX} conditional on the matching variable (in the above DAG the matching variable is *B*)

Parameter collapsibility summary (for models not involving time)

- Collapsible parameters
  - *All parameters under the null*
  - Risk difference and models using an identity link
  - Risk ratio and models using a log link
- Non-collapsible parameters
  - Odds ratio and models using a logit link
  - Parameters from other GLM link functions: Probit, Complementary log-log, Cauchit
- but what about time-to-event models ...

Parameter collapsibility summary in time-to-event models

- Collapsible parameters
  - *All parameters under the null*
  - Rate difference in continuous time models
  - Hazard difference in continuous time models (Aalen model)
- Non-collapsible parameters
  - Rate difference in discrete time models
  - Hazard difference in discrete time models
  - Rate ratio (in both discrete and continuous time models)
  - Hazard ratio (in both discrete and continuous time models)
- Hey, what's going on?

Time-to-event models – what's the complication?

- Explanation from Daniel et al. (2020), see Appendix A.2
- Link functions for mapping rates/hazards to linear predictors in time-to-event models are the log link (e.g., in the Cox PH model, which ensures the hazard remains positive) and the identity link (e.g., Aalen additive hazards model, which does not ensure the hazard remains positive, Didelez and Stensrud (forthcoming))
- It might be tempting to think, therefore, that noncollapsibility is not an issue for rate/hazard differences/ratios
- Sjölander et al. (2016) explain why this reasoning is faulty
- Rates and hazards are based on conditional probabilities (conditional on prior survival), and these cannot be averaged over C as for non-time-to-event models

Time-to-event models – what's the complication?

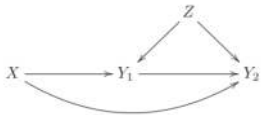


FIGURE 2. A randomized trial where survival is measured at two discrete follow-up times.

- E.g. at $t \geq 2$ Y_1 collides X and Z, introducing bias in models conditioning on Y_1
  - N.b. the numerator and denominator of a hazard ratio are based on different populations – technically a causal contrast should compare the same popn under different treatments
- An additional step that removes this conditioning on survival converts a rate/hazard model into a risk model (the probability of an event before time t)
  - (interestingly) this step alters the link function

Time-to-event models – what's the complication?

- For rates (arising from survival models in discrete time), the corresponding risk model link functions imply a nonlinear CCF even when the rate model link function is either the identity or log link
  - thus rate differences and rate ratios are non-collapsible
- As time is subdivided into more intervals, and the probability of the event in any given interval decreases (and the rate becomes a hazard)
  - the risk model link function corresponding to an additive hazards model is the complementary log link – collapsible
  - the risk model link function corresponding to a Cox PH model is the complementary log–log link – noncollapsible
  - Thus, hazard differences (unlike rate differences) are collapsible, but hazard ratios are not

Time-to-event models – what's the complication?

- For rates and hazard ratios non-collapsibility occurs for all link functions (apart from for the additive hazards model)

Time-to-event models – what's the complication?

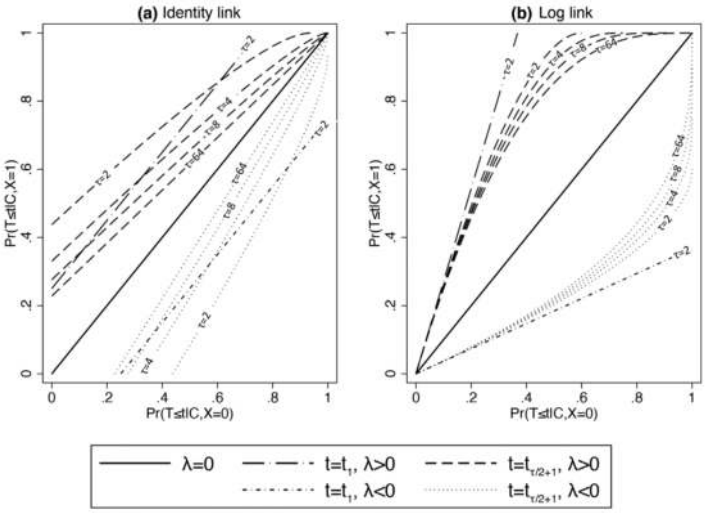


FIGURE A.2 The CCF implied by discrete-time rate models with (a) an identity link and (b) a log link, for both the first time-interval, and a subsequent time-interval, as well as for different values of τ, the total number of time intervals, and for different treatment effect values (λ) on the scale of the linear predictor. As $\tau \to \infty$, the discrete-time rate model becomes a continuous-time hazard model

Summary

- Neuhaus and Jewell (1993) described the solution to non-collapsibility, the now named, CCF
- Sjölander et al. (2016) argued a different explanation holds for time-to-event models
- Daniel et al. (2020) showed the CCF explanation holds for time-to-event models
- TLDR; thread by Tim Morris https://twitter.com/tmorris_mrc/status/1308330100400807940
- Additionally:
    - Didelez and Stensrud (forthcoming) on Liu et al. (2020) in clinical trials
    - Zhang (2008) proposed a method to marginalise conditional ORs after logistic regression with covariates
        - Daniel et al. (2020) adapted this for Cox PH model

Summary

- Issue of: marginal/conditional estimand and unadjusted/adjusted analysis
  - i.e. if you are totally clear what estimand you are targetting you theoretically can't get confused ...

References

Bishop, Y., S. Fienberg, and P. Holland. 1975. *Discrete Multivariate Analysis: Theory and Practice.* MIT Press.

Daniel, R., J. Zhang, and D. Farewell. 2020. “Making apples from oranges: Comparing noncollapsible effect estimators and their standard errors after adjustment for different covariate sets.” *Biometrical Journal* n/a (n/a).

Didelez, V., S. Kreiner, and N. Keiding. 2010. “On the use of graphical models for inference under outcome dependent sampling.” 25:368–387.

Didelez, V., and M. J. Stensrud. Forthcoming. “On the logic of collapsibility for causal effect measures.” *Biometrical Journal.*

Greenland, S., and J. Pearl. 2011. “Adjustments and their Consequences – Collapsibility Analysis using Graphical Models.” *International Statistical Review* 79 (3): 401–426.

Greenland, S., J. M. Robins, and J. Pearl. 1999. “Confounding and Collapsibility in Causal Inference.” *Statistical Science* 14, no. 1 (February): 29–46.

References

Liu, Y., B. Wang, M. Yang, J. Hui, H. Xu, S. Kil, and J. Hsu. 2020. “Correct and logical causal inference for binary and time-to-event outcomes in randomized controlled trials.” *Biometrical Journal.*

Neuhaus, J. M., and N. P. Jewell. 1993. “A geometric approach to assess bias due to omitted covariates in generalized linear models.” *Biometrika* 80 (4): 807–815.

Simpson, E. 1951. “The interpretation of interaction in contigency tables.” *Journal of the Royal Statistical Society (Series B)* 13 (2): 238–241.

Sjölander, A., E. Dahlqwist, and J. A. Zetterqvist. 2016. “Note on the Noncollapsibility of Rate Differences and Rate Ratios.” *Epidemiology* 27 (3): 356–359.

Zhang, Z. 2008. “Estimating a marginal causal odds ratio subject to confounding.” *Communications in Statistics—Theory and Methods* 38:309–321.