A hazards approach to the biometric analysis of infant mortality

Göran Broström[1] Tommy Bengtsson[2]

[2]CED, Lund University (tommy.bengtsson@ekh.lu.se)

[1]CEDAR, Umeå University (gb@ehar.se, https://github.com/goranbrostrom)

Outline

- Jean Bourgeois-Pichat (1951): Biometric analysis of infant mortality (background).
- Hazard-based alternative.
- Theoretical considerations
- Real-world examples.

Bourgeois-Pichat and causes of death

- Two categories:
  - Endogenous: *inherited*, *delivery*, etc.
  - Exogenous: *accidental*, *infectious diseases*, etc.
- How to differentiate between the two categories without information of causes of death?

Bourgeois-Pichat's biometric model.

The biometric model

Two postulates:

1. Endogenous deaths only occur during the neonatal period (0–28 days).
2. On a specific time scale, exogenous infant mortality is uniformly distributed.
   - The log-cube time scale.

The log-cube transform

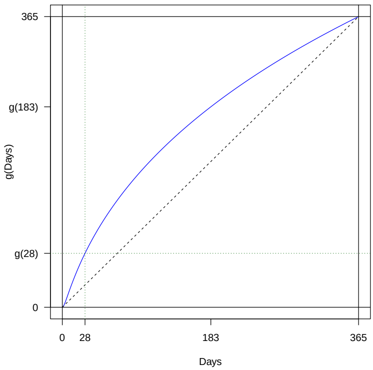
$$g(t) = C \log^3(t+1),\ 0 < t \leq 365.$$

where C is a normalizing constant:

$$C = \frac{365}{\log^3(365+1)}$$

t is age in days.

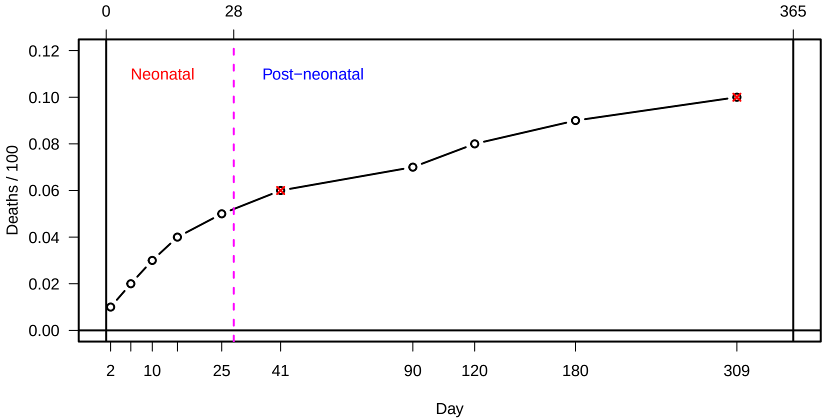
Note: C is not part of B-P's original definition.



Demonstration of the B-P plot

Assume 100 births, of which 90 survives infancy, ten death ages:

2, 6, 10, 15, 25, 41, 90, 120, 180, 309.

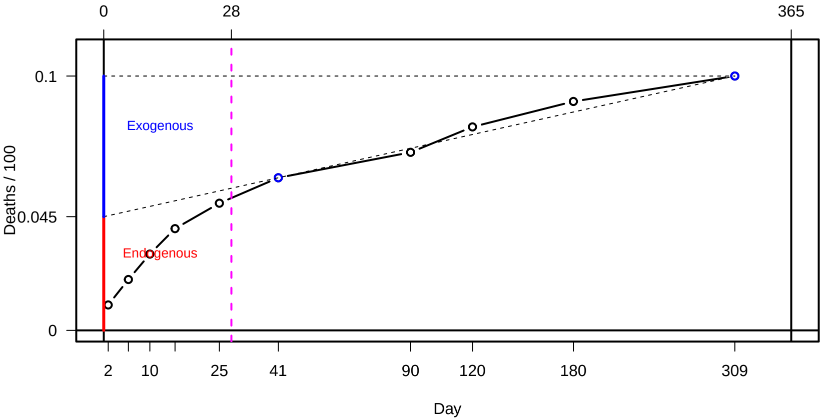


Calculation (no ties): $\frac{1}{100}, \frac{2}{100}, \frac{3}{100}, \ldots, \frac{10}{100}$.

Demonstration of the B-P plot II

Assume 100 births, of which 90 survives infancy, ten death ages:

2, 6, 10, 15, 25, 41, 90, 120, 180, 309.

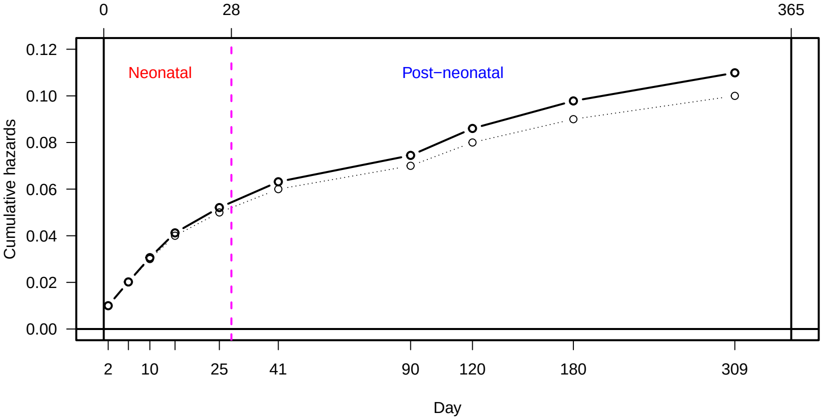


Calculation (no ties): $\frac{1}{100}, \frac{2}{100}, \frac{3}{100}, \ldots, \frac{10}{100}$.

The hazards plot

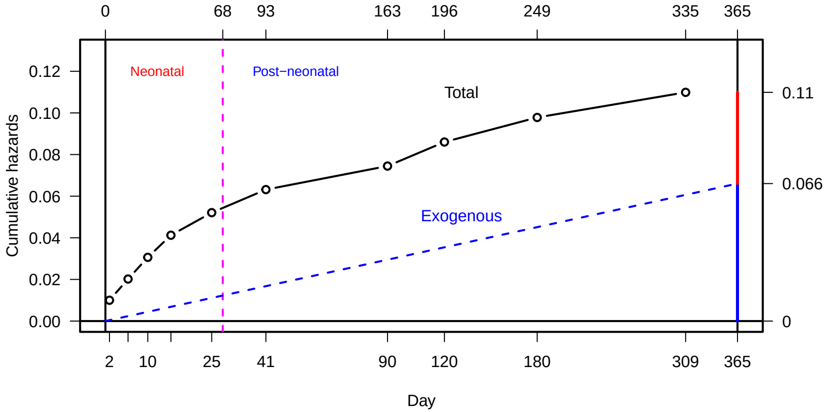
Assume 100 births, of which 90 survives infancy, ten death ages:

2, 6, 10, 15, 25, 41, 90, 120, 180, 309 (same as before).



Calculation (no ties): $\frac{1}{100}, \frac{2}{99}, \frac{3}{98}, \ldots, \frac{10}{91}$.

The hazards plot II

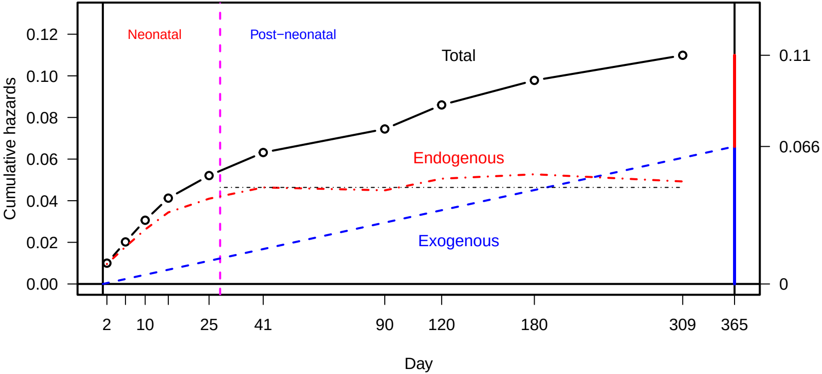


$$\text{Exogenous slope} = \frac{\#(\text{post deaths})}{\text{post exposure}} = \frac{5}{27626} \approx 0.00018$$

Post exposure $= 90(365 - 68) + (93 - 68) + (163 - 68) + \cdots + (335 - 68) \approx 27626$

Adding the endogenous mortality

By subtraction: Endogenous = Total − Exogenous

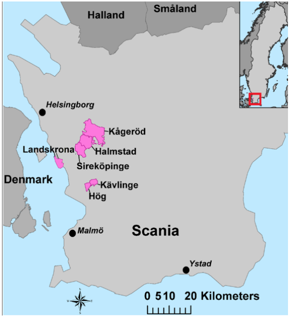


Earlier work

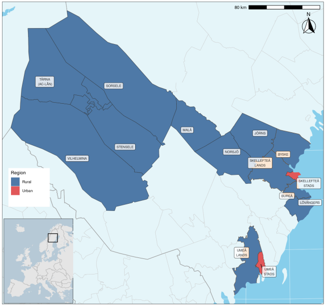
- Knodel & Kintner (1977).
  - Discusses effects of breastfeeding habits.
- Wrigley (1977).
  - Problems with data recording.
- Lynch, Greenhouse & Brändström (1998).
  - Mentions "constant hazard".
  - Suggest linear regression methods.
- Bengtsson (1999).
  - Data quality.
- Manfredini (2004).
  - Deals with effect of climate variation.

Study areas

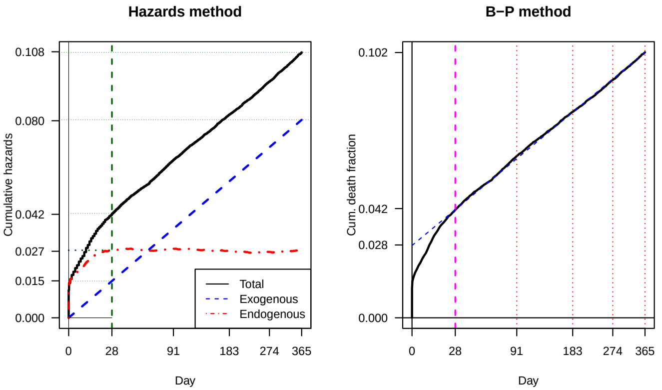
Skåne 1711–1800



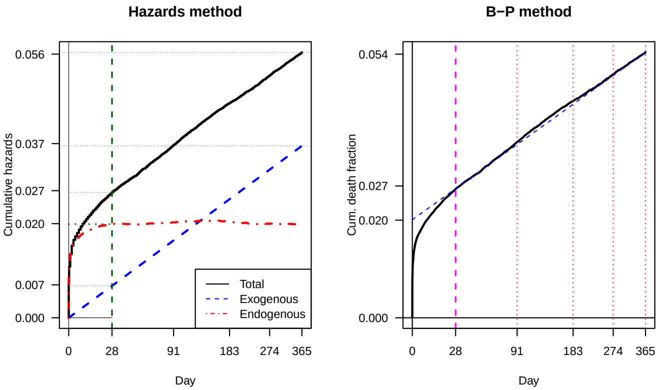
Västerbotten 1801–1950



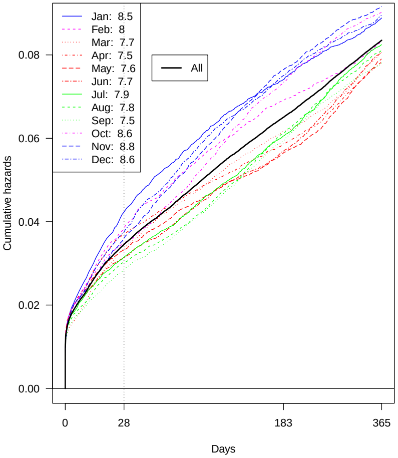
Västerbotten 1861-1890



Västerbotten 1921–1950



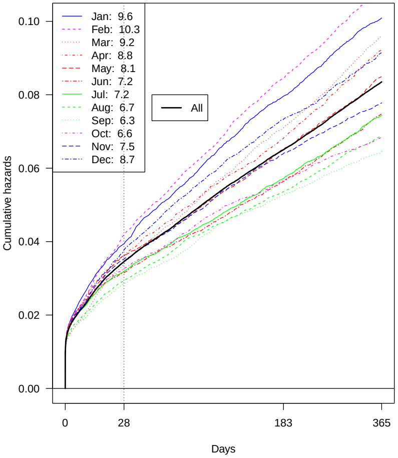
Cohort effect of birth month, Västerbotten



Conclusion:

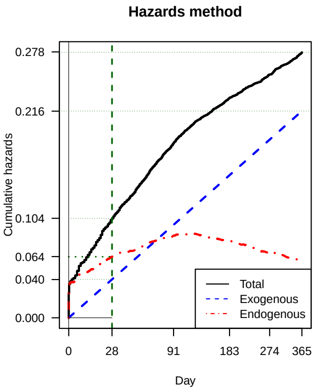
Post-neonatal linearity is a population property, not an individual.

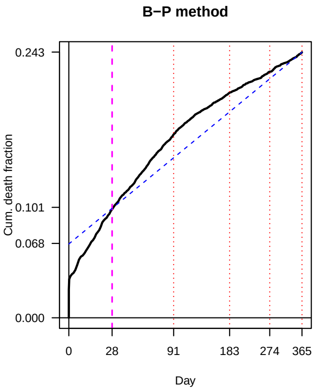
Period effect of month, Västerbotten



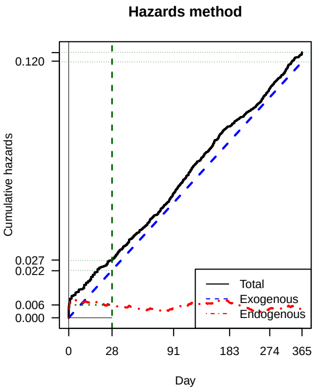
Looks quite proportional!

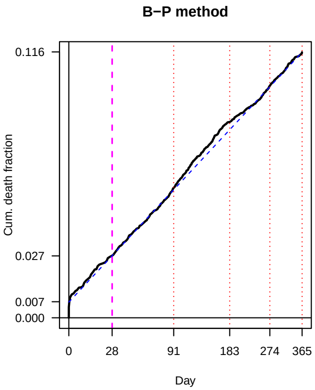
Scania 1711–1800





Skellefteå 1801–1820





Biometric analysis on the natural time scale

Given that the exogenous mortality follows an Exponential distribution on the log-cube scale, it follows that the cumulative hazards function H on the natural time scale is

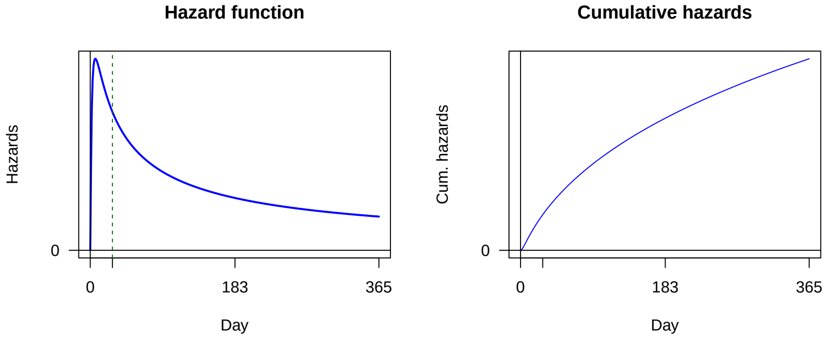
$$H(t; \lambda) = \lambda \log^3(t + 1), \; t \geq 0,$$

and it follows that the hazard function h is given by

$$h(t; \lambda) = \lambda \frac{3 \log^2(t + 1)}{t + 1}, \; t \geq 0,$$

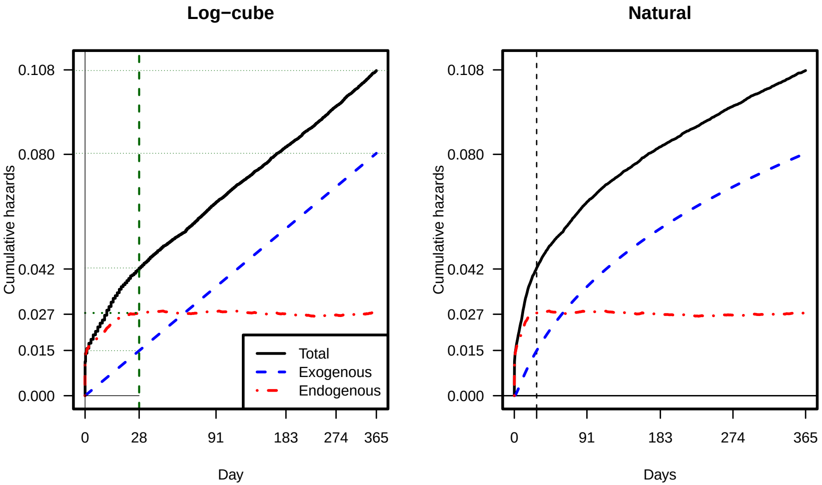
As λ varies, this constitutes a proportional hazards family of distributions.

Exogenous mortality on the normal time scale



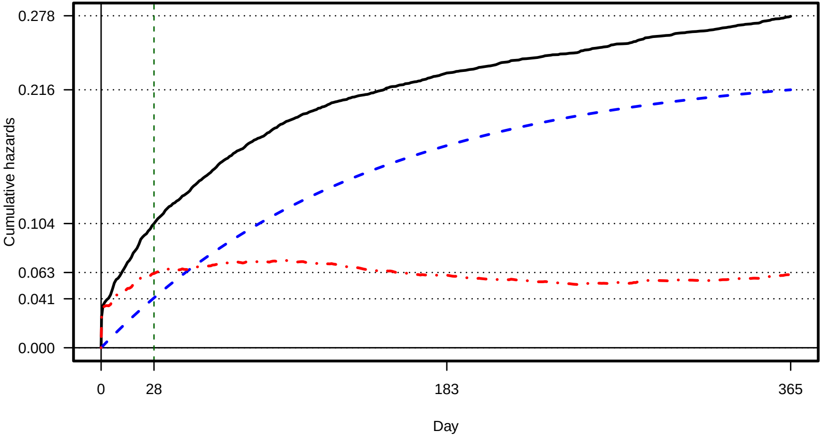
Natural vs. log-cube time scales

Västerbotten 1861–1890.



Natural time scale and Gompertz exogenous distribution

Skåne 1711–1800



Conclusions

- Advantages of the hazards method.
  - Fits naturally into modern survival analysis, allows censored and truncated data, and
  - The exogenous component forms a class of proportional hazards distributions.
  - So, theoretically more appropriate.
- Restrictions for both models.
  - Models for populations, not for individuals.
  - Less useful for traditional survival analysis.
  - The log-cube transform is not universal.

Suggestions for future work

- Work on the natural time scale, if
  - access to a computer, and
  - suitable software.
- Try standard distributions for exogenous mortality, eg.
  - Weibull,
  - Gompertz.
- Remember that we are dealing with competing risks!
  - Requires special care.

https://github.com/goranbrostrom/ESSHC_23

R code for this paper and talk.