Stratified v.s. Unstratified Analysis

Chao Cheng

September 2, 2025

Contents

1 Introduction

In this note we will talk about the Cox's proportional hazards (Cox's PH) model. And more specifically, what will happen when unstratified analysis is used for a data where stratified analysis is the true model.

2 A simple parametric model

Consider the Weibull distribution, denote $T \sim W(p, \lambda)$. Then

$$\begin{aligned}
&f(t) = p\lambda^p t^{p-1}\exp\left(-(\lambda t)^p\right) \\
&F(t) = 1 - \exp\left(-(\lambda t)^p\right) \qquad S(t) = \exp\left(-(\lambda t)^p\right) \\
&h(t) = p\lambda^p t^{p-1} \\
&H(t) = (\lambda t)^p \\
&\mathrm{E}(T) = \frac{1}{\lambda}\cdot\Gamma\left(1+\frac{1}{p}\right) \qquad \mathrm{Var}(T) = \frac{1}{\lambda^2}\left(\Gamma\left(1+\frac{2}{p}\right) - \Gamma\left(1+\frac{1}{p}\right)\right) \\
&\mathrm{E}(T^m) = \frac{1}{\lambda^m}\Gamma\left(1+\frac{m}{p}\right)
\end{aligned}$$

Then the likelihood is

$$\begin{aligned}
L(t_1,\cdots,t_n|p,\lambda_1,\cdots,\lambda_n) &= \prod_{i=1}^n \left(f(t_i)^{\delta_i}(1-F(t_i))^{1-\delta_i}\right) = \prod_{i=1}^n \left(h(t_i)^{\delta_i} S(t_i)\right) \\
&= \prod_{i=1}^n p\lambda_i^p t_i^{p-1}\exp\left(-(\lambda_i t_i)^p\right)
\end{aligned} \tag{1}$$

Note that in (1), we assume all subjects share the same p in the Weibull distribution, but their λs can be different.

3 Cox model

For a Cox model, the key assumption is constant hazard ratio, that is

$$h\left(t|Z\right) = h_0\left(t\right) \cdot e^{\beta Z}$$

And if we plug-in the Weibull distribution, that is $h_0\left(t\right) = p\lambda_0^p t^{p-1}$. Then

$$h\left(t|Z\right) = h_0\left(t\right) \cdot e^{\beta Z} = p\lambda_0^p t^{p-1} \cdot e^{\beta Z} = p\left(\lambda_0 e^{\beta Z/p}\right)^p t^{p-1}$$

References