

- `exp(coef)` is the hazard ratio $\frac{\lambda_T(t \mid x+1)}{\lambda_T(t \mid x)} = \frac{\lambda_T(t \mid \text{gene expression})}{\lambda_T(t \mid \text{no gene expression})} [= \exp(\beta)]$, where λ_T is our hazard function.
x is the treatment parameter. E.g. in this example x is given by GENE1, being 1 for samples that express the gene and 0 for samples that do not express the gene.
- `exp(-coef)` is therefore the (inverse) hazard ratio $\frac{\lambda_T(t \mid \text{no gene expression})}{\lambda_T(t \mid \text{gene expression})}$
- `coef` is this estimated coefficient $\hat{\beta}$ from the model (see below).
- `se(coef)` is the standard error $\sqrt{\text{Var}(\hat{\beta})}$
- `z` is the z-score $\frac{\text{coef}}{\text{se(coef)}}$ (how many standard errors is $\hat{\beta}$ away from 0)
- `Pr(>|z|)` the probability that the estimated $\hat{\beta}$ could be 0.
- `lower .95` and `upper .95` are the 95%-confidence interval for the estimated hazard ratio `exp(coef)`
- Then there are different test scores, which I'm unfortunately not versed enough on.

Some details on the model

The cox model is a linear transformation model of the form

$$\mathbb{P}(T \leq t \mid x) = \exp(-\exp(g(t) + \tilde{x}^T \beta))$$

where $g(t)$ is an unspecified linear transformation function.

The cool thing is that this unknown $g(t)$ goes into a baseline hazard $\lambda_0(t)$ which is independent of β . This allows us to estimate the optimal parameter $\hat{\beta}$ independent of the baseline hazard. (Like we're only interested in the hazard ratio but not in the absolute values)

Leaving out calculations, the hazard function has the form: $\lambda_T(t) = \lambda_0(t) \cdot \exp(\tilde{x}^T \beta)$ and in order to estimate $\hat{\beta}$ we take λ_0 as piecewise constant (changes only when an event happens) and minimize the log-likelihood.