



## from PhD Course Advanced Biostatistics

# **Survival Analysis**

### **Peter Nazarov**

petr.nazarov@lih.lu

10-2020

L5. Data analysis in transcriptomics





#### **Survival Data**

#### **Survival analysis**

is a branch of statistics which deals with analysis of time to events, such as death in biological organisms and failure in mechanical systems (i.e. **reliability theory** in engineering).

Survival analysis attempts to answer questions such as:

- What is the proportion of a population which will survive past a certain time?
- Of those that survive, at what rate will they die or fail?
- Can multiple causes of death or failure be taken into account?
- How do particular circumstances or characteristics increase or decrease the probability of survival?



http://www.partek.com/webinars/survival-analysis-partek-genomics-suite-software





#### **Cox's Survival Model**

$$f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}}$$
 for  $x \ge 0, \mu > 0$ 



$$h(t) = h_0(t)e^{\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n}$$

h(t) – hazard function

 $h_0(t)$  – basic hazard

x<sub>1</sub>...x<sub>n</sub> – variables (expression of gene X) and covariates (age, smoking,) which are time-independent!

 $\beta_1 ... \ \beta_n$  – fitted parameters

$$HR = \frac{h_i(t)}{h_j(t)} = \exp\{\beta_1(x_{i1} - x_{j1}) + \dots + \beta_n(x_{in} - x_{jn})\}$$

To identify significantly involved covariate: partial likelihood is calculated



## L5.2. Survival Analysis



