Section 13

Lecture 7

Relation between Nelson-Aalen and Kaplan-Meier

Partition an interval $[0, t]$ into a number of intervals $0 < t_0, t_1, \cdots t_K = t$, define $\Delta t_k = t_k - t_{k-1}$. Then

$$S(t_k) - S(t_{k-1}) \approx -S(t_{k-1})dH(t_k) = -S(t_{k-1})(H(t_k) - H(t_{k-1})),$$

and thus $S(t_k \mid t_{k-1}) \approx 1 - (H(t_k) - H(t_{k-1}))$. Consider the product

$$\begin{aligned} S(t) &= \prod_{k=1}^{K} S(t_k \mid t_{k-1}), \\ &\approx \prod_{k=1}^{K} \{1 - (H(t_k) - H(t_{k-1}))\} \end{aligned}$$

Let K increase, then $S(t)$ is equal to the so-called product integral

$$S(t) = \lim_{\max \Delta t_k \to 0} \prod_{k=1}^{K} \{1 - (H(t_k) - H(t_{k-1}))\} = \prod_{u \leq t} \{1 - dH(u)\}.$$

Product integrals and certain relations

Suppose we considered discrete (survival) distributions. Then, the product integral would simply be

$$\prod_{u \leq t} \{1 - dH(u)\} = \prod_{u \leq t} \{1 - dH(u)\}.$$

Suppose now that H is absolutely continuous and thus $dH(u) = \alpha(u)du$. Use the (Taylor expansion) approximation $e^{-\alpha(u)du} \approx 1 - \alpha(u)du$ we have for small du (informally) that

$$\prod_{u \leq t} \{1 - dH(u)\} = \prod_{u \leq t} \{1 - \alpha(u)du\} = e^{-\int_{u \leq t} \alpha(u)du} = e^{-H(t)}.$$

Relation between Nelson-Aalen and Kaplan-Meier

Remember that $\sum_{T_j \leq t} \Delta \hat{H}(T_j)$, is the Nelson-Aalen estimator.
Now, the Kaplan-Meier estimator can be expressed as

$$\hat{S}(t) = \prod_{u \leq t} \{1 - d\hat{H}(u)\} = \prod_{T_j \leq t} \{1 - d\hat{H}(T_j)\}.$$

Properties of the Kaplan-Meier estimator

Let S_1 and S_2 be two cadlag survival functions. Then, Duhamel's equation gives

$$\frac{S_1(t)}{S_2(t)} = 1 + \int_0^t \frac{S_1(s-)}{S_2(s)} d(H_2 - H_1)(s).$$

As a heuristic motivation for Duhamel's equation, note that using the usual formula for differention, we would expect that

$$d(S_1/S_2) = \frac{(dS_1)S_2 - S_1 dS_2}{S_2^2} = S_1 d(H_2 - H_1)/S_2,$$

then integrate on both sides.

Properties of the Kaplan-Meier estimator

$$S^*(t) = \prod_{u \leq t} \{1 - dH^*(u)\},$$

where $H^*(t) = \int_0^t J(u) dH(u) du$.
We use Duhamel's equation

$$\frac{\hat{S}(t)}{S^*(t)} - 1 = -\int_0^t \frac{\hat{S}(u-)}{S^*(u)} d(\hat{H} - H^*)(u).$$

However, $\hat{H} - H^*$ is a martingale, and the right hand side of the equation above is a stochastic integral,that is, we can define $G(t) = \frac{\hat{S}(t-)}{S^*(t)}$. Thus, $\mathbb{E}\{\frac{\hat{S}(t)}{S^*(t)} - 1\} = 0$ and $\mathbb{E}\{\hat{S}(t)/S^*(t)\} = 1$. This shows that the Kaplan-Meier estimator consistently estimates $S^*(t)$.

Large sample properties (informal argument)

To see that the Kaplan-Meier estimator is uniformly consistent, note that when n grows, we have $\hat{S}(u-)/S^*(u) \approx 1$, $S^*(t) \approx S(t)$ and $H^*(u) \approx H(u)$. $\hat{S}(u-) \approx S^*(u)$. Using the representation by Duhamel's equation when n grows,

$$\frac{\hat{S}(t)}{S(t)} - 1 \approx -\int_0^t d(\hat{H} - H)(u),$$

and thus

$$\hat{S}(t) - S(t) \approx -S(t)\big(\hat{H}(t) - H(t)\big),$$

and $\text{Var}(\hat{S}(t)) \approx S(t)^2 \text{Var}(\hat{H}(t))$,
and thus we can use the properties of the Nelson-Aalen estimator.

Estimation of the variance

Thus, we can estimate the variance of the Kaplan-Meier estimator by

$$\hat{\tau}^2(t) = \hat{S}(t)^2 \hat{\sigma}^2(t)$$

where $\hat{\sigma}^2(t) = \int_0^t W(s)^2 dN(s)$ is the variance of the Nelson-Aalen estimator.

In practice, the Greenwood estimator is often used, which is very similar:

$$\tilde{\tau}^2(t) = \hat{S}(t)^2 \sum_{T_j \leq t} \frac{1}{Z(T_j)(Z(T_j) - 1)}.$$

Restricted mean survival

- The mean survival $\int_0^\infty S(t)dt$ depends heavily on the tails.
  And we only observed individuals in a time interval $[0, \tau]$, so we cannot say much about $S(u)$, $u > \tau$ unless we make strong assumptions.
- However we can define the t restricted mean survival $\mu_t = \int_0^t S(u)du$.
- μ_t can be interpreted as a the expected number of years alive up to time $t \leq \tau$ for a group of individuals.
- We estimate μ_t by using the Kaplan-Meier estimator,

$$\hat{\mu}_t = \int_0^t \hat{S}(u)du.$$

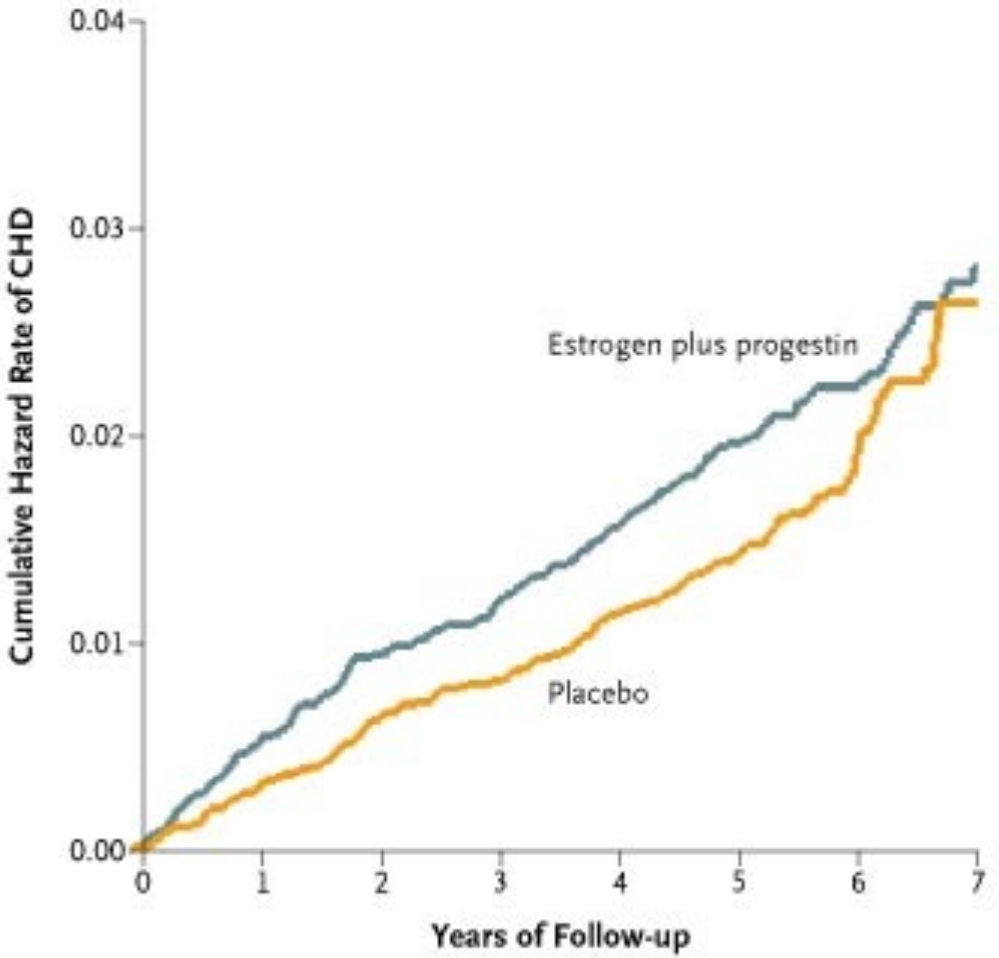
- We can estimate the variance of $\hat{\mu}_t$ by

$$\hat{\text{var}}(\hat{\mu}_t) = \sum_{T_j \leq t} (\hat{\mu}_t - \hat{\mu}_{T_j})^2 W(T_j)^2.$$

Section 14

Hypothesis testing

Harmful medication



No. at Risk								
Estrogen plus progestin	8506	8375	8281	8196	7971	5794	3062	1339
Placebo	8102	8007	7920	7835	7636	5481	2725	988

Testing whether two hazards are equal

- Consider two counting processes with intensities of the multiplicative form
  - N_1 with intensity $\lambda_1(t) = \alpha_1(t)Z_1(t)$,
  - N_2 with intensity $\lambda_2(t) = \alpha_2(t)Z_2(t)$.
- Suppose we want to test the null hypothesis

$$\alpha_1(t) = \alpha_2(t), \quad \text{for all } t \in [0, t_0].$$

- In the survival setting, there is a one-to-one correspondence between $S(t)$ and the cumulative hazard $H(t)$, and thus this hypothesis test is also a test of

$$S_1(t) = S_2(t), \quad \text{for all } t \in [0, t_0].$$

Testing based on increments of Nelson-Aalen

- Remember that the Nelson-Aalen estimator for $H_h(t)$, $h \in \{1, 2\}$, is $\hat{H}_h(t) = \int_0^t W_h(s) dN_h(s)$, that is, $\int_0^t J_h(s)/Z_h(s) dN_h(s)$ when $Z_h(t) > 0$.
- Let $L(t)$ be a non-negative predictable weight process, which is zero whenever $J_1(s)J_2(s) = 0$.
- Define

$$Q_1(t_0) = \int_0^{t_0} L(t)(d\hat{H}_1(t) - d\hat{H}_2(t)),$$

  which accumulates weighted differences in increments of two Nelson-Aalen estimators.

We will use $Q_1(t_0)$ to define test statistics.

Test statistic

Theorem (The rank test statistic)

A test for the null hypothesis, $\mathbf{H}_0$*, that* $\alpha_1(t) = \alpha_2(t), \quad$ *for all* $t \in [0, t_0]$ *can be based on the statistic*

$$U_1(t_0) = \frac{Q_1(t_0)}{\sqrt{V_{11}(t_0)}}$$

where $V_{11}(t_0) = \int_0^{t_0} \frac{L^2(t)}{Z_1(t)Z_2(t)} dN_\bullet(t)$*,* $N_\bullet(t) = N_1(t) + N_2(t)$*, and* $U(t_0)$ *is approximately standard normally distributed under the null hypothesis.*

We will show this in the next slides.

The two next slides give a sketch of proof

$$
\begin{aligned}
Q_1(t_0) &= \int_0^{t_0} L(t)(d\hat{H}_1(t) - d\hat{H}_2(t)) \\
&= \int_0^{t_0} \frac{L(t)}{Z_1(t)} dN_1(t) - \int_0^{t_0} \frac{L(t)}{Z_2(t)} dN_2(t) \text{ (bc } J_1 L = J_2 L = L) \\
&= \int_0^{t_0} \frac{L(t)}{Z_1(t)} dM_1(t) - \int_0^{t_0} \frac{L(t)}{Z_2(t)} dM_2(t) \text{ (under } \mathbf{H}_0\text{, bc } \alpha_1(t) = \alpha_2(t)
\end{aligned}
$$

Remember that $dN_h(t) = \alpha(t) Z_h(t) dt + dM_h(t)$ under the null hypothesis. $Q_1(t_0)$ is the difference between two stochastic integrals of counting process martingales, and therefore Q_1 is a mean zero martingale.

The variance of Q_1

We use the definition of a predictable variation process of a stochastic integral of a counting process martingale (See slide 155) to find that

$$\begin{aligned}\langle Q_1 \rangle(t_0) &= \int_0^{t_0} \left(\frac{L(t)}{Z_1(t)} \right)^2 \alpha(t) Z_1(t) dt + \int_0^{t_0} \left(\frac{L(t)}{Z_2(t)} \right)^2 \alpha(t) Z_2(t) dt \\ &= \int_0^{t_0} \left(\frac{L^2(t) Z_\bullet(t)}{Z_1(t) Z_2(t)} \right) \alpha(t) dt.\end{aligned}$$

The estimator is defined by replacing $\alpha(t)dt$ with $d\hat{H}(t) = \frac{dN_\bullet(t)}{Z_\bullet(t)}$.
We will rely on the large-sample behaviour of this statistic. Thus, we should check that the conditions for the Martingale central limit theorem hold. You can do this by yourself for the log-rank test. Hint, note that $\frac{1}{\sqrt{n}} Q_1(t_0)$ is a martingale (as a process in t_0), and check that the 2 sufficient conditions for the martingale CLT hold.

How to select an appropriate $L(t)$ function?

- Now we have a general way of defining test statistics.
- The log-rank test, which is used in the applied literature *very* frequently defines $L(t) = \frac{Z_1(t)Z_2(t)}{Z_\bullet(t)}$ such that the statistic is

$$Q_{1,logrank}(t_0) = \int_0^{t_0} \frac{Z_1(t)Z_2(t)}{Z_\bullet(t)}(d\hat{H}_1(t) - d\hat{H}_2(t)),$$

$$\langle Q_{1,logrank} \rangle(t_0) = \int_0^{t_0} L(t)\alpha(t)dt.$$

- Curiosity: note that $2L(t)$ is a harmonic mean, $2L(t) = 2\frac{1}{\frac{Z_\bullet(t)}{Z_1(t)Z_2(t)}} = \frac{2}{\frac{Z_1(t)+Z_2(t)}{Z_1(t)Z_2(t)}} = \frac{2}{\frac{1}{Z_1(t)}+\frac{1}{Z_2(t)}}$, which is also used to study rates in other settings.

How does log-rank perform

- How does this perform?
- What if the interest is survival at a given t?

What is a good null hypothesis

Instead of assessing hazards, let us study tests of parameters X_1 and X_2 in groups 1 and 2 at a prespecified time t_0. The null hypothesis is

$$\mathbf{H}_0^X\colon\ X_1(t_0) = X_2(t_0), \tag{13}$$

where X could be the survival function ($S_h(t_0)$, $h = 1, 2$), the restricted mean survival function ($\mu_h(t_0)$, $h = 1, 2$) etc.

Example survival

Instead of the log-rank test, consider for example a test of survival at time t_0 for two independent groups $h = 1, 2$, defined by the statistic

$$Q_S(t_0) = (\widehat{S}_1(t_0) - \widehat{S}_2(t_0))^2 / \widehat{V}(t_0).$$

where

$$\widehat{V}(t_0) = var[\widehat{S}_1(t_0) - \widehat{S}_2(t_0)] = var[\widehat{S}_1(t_0)] + var[\widehat{S}_2(t_0)].$$

Note that $Q_S(t_0)$ is approximately χ^2 with 1 degree of freedom. Hence, we can plug-in the variance estimator from the Kaplan-Meier estimator and get a test of the survival functions.

PS: Alternatively we could use the statistic $(\widehat{S}_1(t_0) - \widehat{S}_2(t_0))/\sqrt{\widehat{V}(t_0)}$, which would be standard normal.

When do we expect treatments to have identical survival curves at all $t \in [0, t_0]$?

- Perhaps justified when a treatment has no effect whatsoever.
- More dubious when comparing two different treatment, for example a new treatment vs. an old treatment.
- The time-dependent profile before t_0 is not our primary interest in many scenarios. In medicine, for example, we are often interested in comparing different treatment regimes, such as radiation and surgery for a particular cancer. Then, time to treatment failure is expected to differ in the shorter term due to the fundamental difference between the treatment regimes, but the study objective is to assess longer-term treatment differences.

Hazards are building blocks

We must be careful about assigning causal interpretations to hazards. However, hazards are key elements in the modelling of other parameters that are easier to interpret, serving as building blocks.