Nelson-Aalen & Kaplan-Meier

Nelson-Aalen Non-parametric estimator of the cumulative hazard rate. $A(t) = \int_0^t \alpha(u)du$ where $\alpha(t)$ is the hazard rate at time t.

Formulas $\hat{A}(t) = \sum_{T_j \le t} \Delta\hat{A}(T_j)$
No ties $\Delta\hat{A}(T_j) = \frac{1}{Y(T_j)}$ Rounded Ties $\Delta\hat{A}(T_j) = \sum_{k=1}^{d_j-1} \frac{1}{Y(T_j)-k}$ True Ties $\Delta\hat{A}(T_j) = \frac{d_j}{Y(T_j)}$
$\hat{\sigma}^2_{\text{N-A}}(t) = \sum_{T_j \le t} \Delta\hat{\sigma}^2(T_j)$
No ties $\Delta\hat{\sigma}^2(T_j) = \frac{1}{Y(T_j)^2}$ Rounded Ties $\Delta\hat{\sigma}^2(T_j) = \sum_{k=1}^{d_j-1} \frac{1}{(Y(T_j)-k)^2}$ True Ties $\Delta\hat{\sigma}^2(T_j) = \frac{(Y(T_j)-d_j)d_j}{Y(T_j)^3}$

Derivation of the Nelson-Aalen Estimator. $M(t)$ is a mean-zero m.g. $H(t) = \frac{J(t)}{Y(t)}$ is predictable where $J(t) = \mathbf{1}(Y(t) > 0)$. We then get $\hat{A}(t) = \int_0^t H(s)dN(s) = \int_0^t \alpha(s)J(s)ds + \int_0^t \frac{J(s)}{Y(s)}dM(s)$. As $M(t)$ is mean-zero m.g. we then get that the Nelson-Aalen estimator is a unbiased estimator of $A^*(t) = \int_0^t \alpha(s)J(s)ds$ since $\mathbb{E}\left[\hat{A}(t) - A^*(t)\right] = 0$. $\hat{A}(t)$ is however a biased estimator of $A(t)$ since $\mathbb{E}[J(s)] = \mathbb{P}(Y(s) > 0)$. This bias is however very small.

Derivation of variance of Nelson-Aalen Estimator. Recall that $\left[\int HdM\right](t) = \int H(s)^2 dN(s)$ where we have $H(s) = \frac{J(s)}{Y(s)} \implies \left[\int HdM\right](t) = \int \left(\frac{J(s)}{Y(s)}\right)^2 dN(s) = [\hat{A} - A^*](t) \implies \hat{\sigma}^2_{\text{N-A}} = \sum_{T_j \le t} \frac{1}{Y(T_j)}$

Delta Method CI Nelson-Aalen We have that $\hat{A}(t) \overset{\text{approx}}{\sim} N(A(t), \hat{\sigma}^2(t))$

$$g(\hat{A}(t)) \approx g(A(t)) + g'(A(t)) \underbrace{(\hat{A}(t) - A(t))}_{\mathbb{E}[\dots]=0} \implies \mathbb{E}[g(\hat{A}(t))] \approx g(A(t)) \text{ and } \mathbb{E}[(g(\hat{A}(t)) - g(A(t))^2] \approx g'(\hat{A}(t))^2 \underbrace{\mathbb{E}[(\hat{A}(t) - A(t))^2]}_{\hat{\sigma}^2}$$

$$\implies g(\hat{A}(t)) \overset{\text{approx}}{\sim} N(g(A(t), |g'(\hat{A}(t)|\hat{\sigma})$$

Let $g(x) = \log(x)$ which then gives us $g^{-1}(x) = e^x$ and $g'(x) = \frac{1}{x}$. The interval then becomes as follows.
$g^{-1}(CI) = \exp\left\{\log(\hat{A}(t)) \pm z_{1-\alpha/2}\frac{\hat{\sigma}}{\hat{A}(t)}\right\} = \hat{A}(t)\exp\left\{\pm z_{1-\alpha/2}\frac{\hat{\sigma}}{\hat{A}(t)}\right\}$

Kaplan-Meier Non-parametric estimator of the survival function. $S(t) = e^{-A(t)}$ where $A(t)$ is the cumulative hazard rate at time t.

Formulas $\hat{S}(t) = \prod_{T_j \le t}\left(1 - \frac{1}{Y(T_j)}\right) = \prod_{T_j \le t}(1 - \Delta\hat{A}(T_j))$
$\hat{\tau}^2(t) = \hat{S}(t)^2 \sum_{T_j \le t} \frac{1}{Y(T_j)^2} = \hat{S}(t)^2\hat{\sigma}^2_{\text{N-A}}$ $\hat{\tau}^2(t) = \hat{S}(t)^2 \sum_{T_j \le t} \frac{d_j}{Y(T_j)(Y(T_j)-dj)}$ (Greenwood)

Derivation of the Kaplan-Meier Estimator Recall $\mathbb{P}(T > t) \implies S(t_k|t_{k-1}) = \mathbb{P}(T > t_k | T > t_{k-1}) = \frac{S(t_k)}{S(t_{k-1})}$. Let $0 = t_0 < t_1 < \ldots < t_n$ and note that $\mathbb{P}(T > t_0) = 1$ which gives us $S(t_n) = \prod_{k=1}^n \frac{S(t_k)}{S(t_{k-1})}$. We formally we define the survival function as $S(t) = \prod_{u \le t}(1 - dA(u))$ since $\frac{S(t_k)}{S(t_{k-1})} = dA(t_k)$ when $t_k - t_{k-1} << 1$. This gives us the estimator $\hat{S}(t) = \prod_{T_j \le t}(1 - \Delta\hat{A}(T_J))$ as $\Delta\hat{A}(t)$ serves as an estimator for $dA(t)$.

Kaplan Meier CI $\hat{S}(t) \pm z_{1-\alpha/2}\hat{\tau}(t)$. Log-transforms (using same method as for Nelson-Aalen above) etc.

Derivation of variance of Kaplan-Meier Estimator Let $S^*(t) = \prod_{u \le t}(1 - dA^*(t))$ where $A^*(t) = \int_0^t J(u)dA(u)$. If $\mathbb{P}(J(s) = 0) << 1$ then S^* and S are close. We measure this closeness by $\frac{\hat{S}(t)}{S^*(t)} - 1 = -\int_0^t \frac{\hat{S}(u-)}{S^*(u)} d(\hat{A} - A^*)(u)$. We then have that $\mathbb{E}\left[\frac{\hat{S}(t)}{S^*(t)}\right] = 1$. We can then repat the arguments as we do for the variance of the Nelson-Aalen estimator above. $\left[\frac{\hat{S}}{S^*} - 1\right] = \left[\int \frac{-\hat{S}}{S^*} \underbrace{d(\hat{A} - A^*)}_{dM}\right] = \{\text{Theorem}\} = \int \left(\frac{\hat{S}}{S^*}\right)^2 d[M]$. Note that M is the same mean-zero m.g. as in the Nelson-Aalen case which gives us $d[M](t) = \frac{J(t)}{Y(t)}dN(t)$. This does in turn give us that $\left[\frac{\hat{S}}{S^*} - 1\right] = \int \left(\frac{\hat{S}}{S^*}\right)^2 \frac{J}{Y^2}dN$. Now by assuming $S^* = S$ and $\hat{S}(u) \approx \hat{S}(u-)$ we get that $\text{Var}\left(\frac{\hat{S}(t)}{S(t)} - 1\right) = \hat{\sigma}^2_{\hat{S}/S-1}(t) = \int_0^t \frac{J}{Y^2}dN = \hat{\sigma}^2_{\text{N-A}}(t) \implies \hat{\sigma}^2_{\hat{S}}(t) \approx \hat{S}^2(t)\int_0^t \frac{J}{Y^2}dN = \hat{S}^2(t)\hat{\sigma}^2_{\text{N-A}}(t)$

Martingales

Definition of Martingales: M is a Martingale if $E[M_t|\mathcal{F}_s] = M_s, \quad t \geq s$ and $E[|M_t|] < \infty$.

Formulas

$$(H \bullet M)_n = H_0M_0 + H_1(M_1 - M_0) + \ldots + H_n(M_n - M_{n-1})$$

$$\langle H \bullet M \rangle_n = \sum_{i=1}^{n} H_i^2 \Delta\langle M \rangle_i \text{ where } \Delta\langle M \rangle_i = [(M_i - M_{i-1})^2|\mathcal{F}_{i-1}]$$

$$\langle H \bullet M \rangle = H^2 \bullet \langle M \rangle$$

$$\langle M \rangle_n = \sum_{i=1}^{n} \mathrm{Var}(\Delta M_i|\mathcal{F}_{i-1}) = \sum_{i=1}^{n} \mathbb{E}[(M_i - M_{i-1})^2|\mathcal{F}_{i-1}]$$

$$\left\langle \int H dM \right\rangle = \int H^2(s) d\langle M \rangle(s) = \int H^2(s)\lambda(s)ds$$

$$\Delta M_t = M_t - M_{t-1} \text{ called m.g. difference}$$

$$[H \bullet M]_n = \sum_{i=1}^{n} H_i^2 \Delta[M]_i \text{ where } \Delta[M]_i = (M_i - M_{i-1})^2$$

$$[H \bullet M] = H^2 \bullet [M]$$

$$[M]_n = \sum_{i=1}^{n} (\Delta M_i)^2 = \sum_{i=1}^{n} (M_i - M_{i-1})^2$$

$$\left[\int H dM\right] = \int H^2(s) d[M]s) = \int H^2(s) dN(s)$$

$$Cov(M_s, M_t - M_s) = 0$$

$M^2 - \langle M \rangle$ and $M^2 - [M]$ are zero mean m.g.s.

$$Var(M(t)) = E\Big(M(t)\Big)^2 = E\langle M \rangle(t) = E[M](t)$$

$$E[M_t - M_s|\mathcal{F}_s] = E[M_t|\mathcal{F}_s] - M_s = 0$$

$$E[\Delta M_t|\mathcal{F}_{t-1}] = 0$$

Doob decomposition: Let X with X_0 be a general discrete time proc. and let M be def. by $M_0 = X_0 = 0$ and $\underbrace{M_n - M_{n-1}}_{\Delta M_n} = X_n - E[X_n|\mathcal{F}_{n-1}] \implies E[\Delta M_n|\mathcal{F}_{n-1}] = 0 \implies X_n = E[X_n|\mathcal{F}_{n-1}] + \Delta M_n = \underbrace{E[X_n|\mathcal{F}_{n-1}]}_{pred.} + \underbrace{(X_n - E[X_n|\mathcal{F}_{n-1}])}_{noise}$

Frailty

Hazard ratio: e.g. $\frac{\alpha(t|\mathbf{x}=(1,0))}{\alpha(t|\mathbf{x}=(0,1))} = \frac{r(\beta,(1,0))}{r(\beta,(0,1))}$

Laplace: $\mu(t) = \frac{-S'(t)}{S(t)} = \alpha(t)\frac{\mathcal{L}'(A(t))}{\mathcal{L}(A(t))}$, Frailty model: $\alpha(t|Z) = \alpha(t)Z \implies S(t) = E[e^{-A(t)Z}] = \mathcal{L}_Z(A(t))$
$S(t) = E[S(t|Z)]$

Testing

Breslow-estimator: $\hat{A}_0(t) = \hat{A}_0(t;\hat{\beta})$ where $\hat{A}_0(t;\hat{\beta}) = \int_0^t \frac{dN_\bullet(u)}{\sum_{l=1}^n Y_l(u) r(\beta, \mathbf{x}_l)}$,
Cox prop. haz. model $\implies \alpha(t|\mathbf{x}) = \alpha_0(t) r(\beta \mathbf{x}) \implies A(t|\mathbf{x}) = A_0(t) r(\beta \mathbf{x})$
Gehan-Breslow: $U(t_0) = \frac{Z_1(t_0)}{\sqrt{V_{11}}} \sim N(0,1)$ where $Z_1(t_0) = \int_0^{t_0} Y_2(t) dN_1(t) - \int_0^{t_0} Y_1(t) dN_2(t)$ and $V_{11}(t_0) = \int_0^{t_0} Y_1(t) Y_2(t) dN_\bullet(t)$
Cox regression: Multiplicative model: $\alpha(t|x_1, ..., x_p) = \alpha_0(t) \exp(\beta_1 x_1 + ... + \beta_p x_p)$.
Additive model: $\alpha(t|x_1, ..., x_p) = \underbrace{\beta_0(t)}_{baselinehaz.} + \beta_1 x_1 + ... + \beta_p x_p$.
Cox partial likelihood: $L(\beta) = \prod_{T_j} \frac{r(\beta, x_j)}{\sum_{i \in \mathcal{R}_j} r(\beta, x_j)}$
Example: $\alpha(t;x) = \alpha_0(t) e^{\beta x}$ and $\{(T_i, \delta_i, x_i)\}_{i=1...5} = \{(1,1,1), (3,1,0), (4,0,1), (7,0,0), (10,1,1)\}$ then
$L(\beta) = \frac{e^\beta}{e^\beta + 1 + e^\beta + 1 + e^\beta} \cdot \frac{1}{1 + e^\beta + 1 + e^\beta} \cdot \frac{e^\beta}{e^\beta} = ...$
LR test: $\chi^2_{LR} = 2(\log(L(\hat{\beta}) - \log(\beta_0)) \sim \chi^2(1)$

Misc

Accelerated failure time models: $\log U_i = \beta^T \mathbf{x}_i + \varepsilon_i$ where $E[\varepsilon_i] = 0$ iid. $\implies S_{U_i}(u) = P(U_i > u) = P(e^{\beta^T \mathbf{x}_i + \varepsilon_i} > u) = P(\underbrace{\varepsilon_\mathbf{i}}_{w_i} > ue^{-\beta^T \mathbf{x}_i} = S_{w_i}(ue^{-\beta^T \mathbf{x}_i})$, (change of time), $\implies S'_{U_i}(u) = S'_{w_i}(ue^{-\beta^T \mathbf{x}_i})e^{-\beta^T \mathbf{x}_i} \implies \alpha_{U_i}(u) = \frac{-S'_{U_i}(u)}{S_{U_i}(u)} = \alpha_{w_i}(ue^{-\beta^T \mathbf{x}_i})ue^{-\beta^T \mathbf{x}_i}$

Basic stuff
$S(t) = \exp(-\int_0^t \alpha(s)ds)$
$A(t) = \int_0^t \alpha(s)ds$
$-S'(t) = \alpha(t)S(t)$
$f(t)?\alpha(t)S(t)$
$P(T > x|T > y) = \frac{P(T>x)}{P(T>y)} = \frac{S(x)}{S(y)}$

Problems, solutions and examples

Problem Show $\hat{A}(t) = \int_0^t \frac{I(Y(s)>0)}{Y(s)} dN(s)$ is an unbiased estimator of $A^*(t) = \int_0^t I(Y(s) > 0)\alpha(s)ds$.
Solution $\hat{A}(t) - A^*(t) = \int_0^t \frac{I(Y(s)>0)}{Y(s)} dN(s) - \int_0^t I(Y(s) > 0)\alpha(s)ds = \int_0^t \frac{I(Y(s)>0)}{Y(s)} dN(s) - \int_0^t \frac{I(Y(s)>0)}{Y(s)} Y(s)\alpha(s)ds = \int_0^t \frac{I(Y(s)>0)}{Y(s)} (dN(s) - Y(s)\alpha(s)ds) = \int_0^t \underbrace{\frac{I(Y(s) > 0)}{Y(s)}}_{\text{pred.}} \underbrace{dM(s)}_{\text{mean zero m.g}} = 0.$
Problem Show $A^*(t)$ is a biased estimator of $A(t) = \int_0^t \alpha(s)ds$.
Solution $E[A^*(t)] = E[\int_0^t I(Y(s) > 0)\alpha(s)ds] \le E[\int_0^t 1 \cdot \alpha(s)ds] = A(t)$.
Problem Calculate the optional variation of $\hat{A}(t)A^*(t)$, i.e. $[\hat{A}(t)A^*(t)]$ and write this expression as a sum.
Solution $[\hat{A}(t)A^*(t)] = [\int \frac{1}{Y} dM](t) = \int_0^t \left(\frac{I(Y(s)>0)}{Y(s)}\right)^2 dN(s) = \int_0^t \frac{I(Y(s)>0)}{Y(s)^2} dN(s)$.
We also have that $Var(\hat{A}(t) - A^*(t)) = E\Big([\hat{A}(t) - A^*(t)](t)\Big)$ and thus $\hat{A}(t) - A^*(t)](t)$ is an unbiased est. of the variance.
Morover, assuming no ties $[\hat{A}(t) - A^*(t)](t) = \int_0^t \frac{I(Y(s)>0)}{Y(s)^2} dN(s) = \sum_{T_j \le t} \frac{1}{Y(T_j)}$, which is N-A estimator.
Problem Let X_n be discrete time m.g. Show that $E[X_n^2]$ is non-decreasing in n.
Solution First show $M_{n+1} = X_n(X_{n+1} - X_n)$ has zero mean. $E[X_n(X_{n+1} - X_n)|\mathcal{F}_n] = X_n E[X_{n+1} - X_n|\mathcal{F}_n] = 0$. Now note that $(X_{n+1} - X_n)^2 = (X_{n+1} - X_n)(X_{n+1} - X_n) = X_{n+1}(X_n + 1 - X_n) - M_n$. We get that $E[(X_{n+1} - X_n)^2] = E[X_{n+1}(X_n + 1 - X_n) - M_n] = E[X_{n+1}(X_n + 1 - X_n)] = E[X_{n+1}^2] - E[X_{n+1}X_n] = E[X_{n+1}^2] - E[E[X_{n+1}X_n|\mathcal{F}_n]] = E[X_{n+1}^2] - E[X_n^2] \ge 0$
Table example

t	$Y(t)$	$d(t)$	$\Delta\hat{A}_1(t)$	$\Delta\hat{\sigma}_1^2(t)$	$\Delta\hat{A}_2(t)$	$\Delta\hat{\sigma}_2^2(t)$
0.2	16	1	$\frac{1}{16}$	$\frac{(16-1)\cdot 1}{16^3}$	$\frac{1}{16}$	$\frac{1}{16^2}$
0.5	15	3	$\frac{3}{15}$	$\frac{(15-3)\cdot 3}{15^3}$	$\frac{1}{15} + \frac{1}{14} + \frac{1}{13}$	$\frac{1}{15^2} + \frac{1}{14^2} + \frac{1}{13^2}$
0.7	12	1	$\frac{1}{12}$	$\frac{(12-1)\cdot 1}{12^3}$	$\frac{1}{12}$	$\frac{1}{12^2}$
1.1	11	1	$\frac{1}{11}$	$\frac{(11-1)\cdot 1}{11^3}$	$\frac{1}{11}$	$\frac{1}{11^2}$