Hypothesis Testing $\hat{\theta}$: unrestricted mle; $H_0: \theta = \theta_0$ vs. $H_1: \theta \neq \theta_0$.
Wald Test - MLE $\hat{\theta}$ under null $\theta = \theta_0$,
$W = \frac{(\hat{\theta}-\theta_0)^2}{1/I(\hat{\theta})} \sim \chi_{(1)}$; $W = \frac{\hat{\theta}-\theta_0}{\text{SE}(\hat{\theta})} \sim N(0,1)$[I($\hat{\theta}$) obs info matrix]

Score Test - Null distribution θ_0 by CLT $S(\theta) \sim N(0, I_n(\theta))$ under null $\frac{S(\theta_0)^2}{I(\theta_0)} \sim \chi_{(1)}$; $\frac{S(\theta_0)}{\sqrt{I(\theta_0)}} \sim N(0,1)$ $E[S(\theta)] = 0$ and $V(S(\theta)) = I_n(\theta)$
Res. Deviance : pf. $l(\theta_0)$ expand around $\hat{\theta}$. Power Function: type I: P(reject null| null); power: p(accept null| h1); type II: 1- power

---

Survival Analysis

Basics $S(t) = P(\tilde{T} > t) = 1 - F(t)$ $f(t) = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt}$; hazard fun. (rate, instantaneous occurrence) $\lambda(t) = \lim_{\delta \to 0} \frac{P(t \le \tilde{T} < t+\delta | \tilde{T} > t) = \lambda\delta}{\delta}$
$\lim \frac{F(t+\delta)-F(t)}{S(t)\delta} = \frac{f(t)}{S(t)} = -\frac{dS(t)}{S(t)dt} = -\frac{d}{dt}\ln S(t) = -\frac{d}{dt}\ln 1 - F(t)$;
$\int_0^t \lambda(u)du = \int_0^t \frac{f(u)}{1-F(u)}du = -\ln(1-F(u))\Big|_0^t \to \Lambda(t) = -\ln S(t)$
$\to S(t) = \exp(-\Lambda(t)) = \exp(-\int_0^t \lambda(t)dt)$
likelihood(cont.) $L_n(\theta; t_i, X_i) \propto \prod_{i=1}^n [\lambda_i(t_i|x_i,\theta)^{e_i} S_i(t_i|x_i,\theta)]$
(1)uncensored $E_i = 1$ at t $\Leftrightarrow P(t <= T_i < t + dt; E_i = 1)$
$\overset{A1}{=} P(C_i > t|x_i,\theta)P(t \le \tilde{T}_i < t+dt|\tilde{T}_i > t)P(\tilde{T}_i > t|x_i,\theta) \overset{A2}{=}$
$\propto \lambda_i(t|x_i,\theta)dtS_i(t|x_i,\theta)$
(2) censored($E_i = 0$) at t $\Leftrightarrow P(t \le T_i < t + dt; E_i = 0) \overset{A1}{=}$
$P(t \le C_i < t+dt|x_i,\theta)$ $P(\tilde{T}_i > t|x_i,\theta) \overset{A2}{=} \propto S_i(t_i|x_i,\theta)$
A1: indep censoring $C_i \perp \tilde{T}_i|X_i$ A2: noninfo. censor C_i not involve θ
$\ell(\theta) \propto \sum e_i \log(\lambda_i(t_i|x_i,\theta)) + \sum \log S_i(t_i|x_i,\theta) / \sum - \int_0^t \lambda(t;\theta)dt$
pois likelihood $D \sim$ pois(λ) (1) one sample:
$\ell(\lambda) = D\log(\lambda) - \lambda\sum_{i=1}^n t_i/\lambda Y$, Y: total person-year; $\hat{\lambda}) = \frac{D}{Y}$,
se($\hat{\lambda}$) = $\frac{\sqrt{D}}{Y}$, reparam $\alpha = \log(\lambda)$ [more symmetric] ,se $\log(\hat{\lambda}) = \sqrt{\frac{1}{D}}$
[FI](2) two sample $\ell(\lambda_0, \lambda_1) = d_0\log(\lambda_0) - \lambda_0\sum_{i=1}^n t_i^{\lambda_0 Y_0} +$
$d_1\log(\lambda_1) - \lambda_1\sum_{i=1}^n t_i^{\lambda_1 Y_1}$ mle $\hat{\lambda}_i = \frac{d_i}{Y_i}$ reparam: $\log(\lambda_0) = \alpha$,
$\log(\lambda_1) = \alpha + \beta$; mle $\hat{\alpha} = \log(\frac{d_0}{Y_0})$, $\hat{\beta} = \log(\frac{d_0}{Y_0}/\frac{d_1}{Y_1})$
bern likelihood N bins each iid Bern(π), length h($Y = Nh$). $\mu = \lambda Y$
$= N\pi \to \lambda^{constant} = \pi\frac{N}{Y} = \frac{\pi}{h}/\frac{\mu}{Y}$[bern/pois]. surv one bin=1- risk
=1-π = (1-λh). T-year surv:$(1-\lambda h)^N = \exp(\log(1-\lambda h)^N)$
$= \exp(N \times (-\lambda h)) \approx \exp(-\lambda T)$ [T = Nh],trick $e^x \approx 1 + x, \log(1+x) \approx x$
exp likelihood $\lambda\exp(-\lambda t)$, then $L(\lambda) = e^{-\lambda U_{n+1}}\prod_{j=1}^D \lambda\exp(-\lambda u_j) =$
$\lambda^D \exp^{-\lambda T}$, [aggre data,each interval U_j is exp dist, U_{j+1} censor at T]
(1) one sample mle $\frac{n}{\sum y_i}$ $\sum y_i \sim \Gamma \to \chi$ cal CI (2) two sample: HR ω
(OI), ρ hazard in group0, LRT=2($\ell(\hat{\omega},\hat{\rho})$ - $\ell(\omega_0,\hat{\rho}_{\omega_0})$)
exp model repa $\tilde{T} \sim \exp(\lambda = \lambda_0 e^{\beta X} = e^{\alpha+\beta X})$, model $\lambda = \lambda_0 e^{\beta X}$.
$S_i(t;\theta) = e^{-\int_0^t \lambda_0 e^{\beta x_i}dt} = e^{-t\lambda_0 e^{\beta x_i}} = S_0(t;\theta)e^{\beta x_i}$ [log(-logS)]
AFT $\log\tilde{T}_i = \alpha + \gamma x_i + \sigma\epsilon$, $S_i(te^{\gamma x_i}) = S_0(t)$ or $S_i(t) = S_0(te^{-\gamma x_i})$
$\lambda_0 = e^{-\alpha}$, $\beta = -\gamma$ or $(-\gamma/\sigma)$; $\sigma = 1$ or $(1/\kappa)$
Weibull $\lambda_i(t;\theta) = \lambda\kappa(\lambda t)^{\kappa-1}e^{\beta x_i}$; $S_i(t;\theta) = e^{-(\lambda t)^\kappa e^{\beta x_i}}$ shape param $\kappa > 1 \uparrow$ hazard
Discrete case $S(t_j) = P(\tilde{T} > t_j) = \frac{P(\tilde{T}>t_j)}{P(\tilde{T}>t_{j-1})}P(\tilde{T} > t_{j-1}) =$
$\frac{P(\tilde{T}>t_j)}{P(\tilde{T}>t_{j-1})}\frac{P(\tilde{T}>t_{j-1})}{P(\tilde{T}>t_{j-2})}\ldots = \prod\frac{S(t_j)}{S(t_{j-1})} = \prod(1 - h(t_j))$
Discrete hazard is a condit. prob $h(t_j) = P(\tilde{T} = t_j|\tilde{T} \ge t_j)$
$= \frac{S(t_{j-1})-S(t_j)}{S(t_{j-1})} = 1 - \frac{S(t_j)}{S(t_{j-1})} \to S(t_j) = \prod\frac{S(t_j)}{S(t_{j-1})}$
$= \prod_{t_j \le t}(1 - h(t_j))$ [S(t) multiplication of previous survivals]
Discrete likelihood $\ell(\vec{h}) = \sum_{j=1}^T d_j\log(h_j) + \sum_{j=1}^T(r_j - d_j)\log(1 - h_j)$
Kaplan-Meier from liklhd, $\hat{h}_j = \frac{d_j}{r_j}$, so $\hat{S}(t) = \prod_{j:t_j \le t}\left[1 - \frac{d_j}{r_j}\right]$
$= \prod_{u \le t}\left[1 - \frac{\mathbf{1}\{Y(u)>0\}}{Y(u)}dN(u)\right]$ term2:$(1 - \lambda(t)dt)$ [d_j death at j; r_j risk at j]; , var($\hat{h}_j$) = var($1-\hat{h}_j$) = $\frac{\hat{h}_j(1-\hat{h}_j)}{r_j}$[OFI inverse]. curve [5,14)
Greenwood $V(\hat{S}(t))$ $V(log\hat{S}(t)) = \sum_{j:t_j \le t} V(log(1 - \frac{d_j}{r_j}))$
$\overset{\text{g=log(1-h)}}{\approx} \sum_{j:t_j \le t}[1/(1-d_j/n_j)^2]V(\frac{d_j}{r_j}) = \sum\frac{d_j}{r_j(r_j-d_j)}$;
$V(\hat{S}(t)) = \hat{S}(t)^2\sum_{j:t_j \le t}\frac{d_j}{n_j(n_j-d_j)} = \hat{S}(t)^2\int_0^t\frac{dN(u)}{Y(u)[Y(u)-dN(u)]}$
[delta2:$g = exp(x)$ $x = logS(t)$]; twice delta method. Problem: CI not bounded in(0,1).
Nelson-Aalen unbiased cum. hazard $\hat{\Lambda}_{\text{NA}}(t) = \sum_{j:t_j \le t}\frac{d_j}{r_j}$,
$= \int_0^t\frac{\mathbf{1}\{Y(u)>0\}}{Y(u)}dN(u)$ $\hat{S}_{\text{KM}}(t) \approx \exp[-\hat{\Lambda}_{\text{NA}}(t)]$,
$CI = \Lambda(t) \pm z_{1-\frac{\alpha}{2}} \times S.E.[\hat{\Lambda}(t)]$,var($\hat{\Lambda}(t)$) =var(log($S(t)$)) as
$\hat{\Lambda}(t) = -\log(S(t))$ see greenwood; KM var : $V(\hat{H}(t)) = \sum_{j:t_j \le t}\frac{d_j}{r_j^2}$

COX reg partial likelihood model $\lambda_i(t; Z_i) = \lambda_0(t)\exp(\beta' Z_i(t))$
$z_i(t) = 1_{\{v_i<t\}}$, $\prod_{i=1}^n\left(\frac{\exp(\beta' z_i(t_i))}{\sum_{l=1}^n Y_l(t_i)\exp(\beta' z_l(t_i))}\right)^{e_i}$, i and j with $v_i < t$ and $v_j > t$ $\frac{\lambda_i(t)}{\lambda_j(t)} = \exp(\beta)$ [time-varying cov/immortal]
log-rank $d_{kj} \sim Bin(r_{kj}, \lambda_j h)$ & $d_{kj}|d_j \sim hyp(r_{kj}, r_j, d_j)$, r_{kj} is the number at risk for group k at time t_j, d_{kj} is event count for group k at time t_j, n_j is total count. $E_j(0) = \frac{d_j r_{1j}}{r_j}$ $V_j(0) = \frac{d_j r_{0j} r_{1j}(r_j-d_j)}{r_j^2(r_j-1)}$
$U(0)^2/I(0) = \left(\sum_{j=1}^g d_{1j} - \sum_{j=1}^g E_j\right)^2 / \sum_{j=1}^g V_j \sim \chi^2(1)$
Breslow Estimator $\hat{\Lambda}_0(t)$ $S_i(t) = e^{-\Lambda_0(t)\exp(\hat{\beta}' X)}$ [Model-based survival probability], $\hat{\Lambda}_0(t_i) = \sum_{j:t_j \le t}\left(\frac{d_j}{\sum_l^n Y_l\exp(\beta X_l)}\right)$ where $Y_i(t) = 1_{T_i \ge t}$ at-risk. $\hat{\Lambda}(t) = \hat{\Lambda}_0(t)e^{\hat{\beta}' X)}$
s-year risk $\pi_i(s) = 1 - S_i(s) = 1 - e^{-\Lambda_0(s)\exp(\hat{\beta}' X)}$
Diagnostic (1) martingale resi fun form cont var (2) cum hazard vs cox-snell resi (model fit linear) (3)score res outlier/influential (4) schoenfeld res PH assump (5) dfbeta(s) influential obs $\frac{2}{\sqrt{n}}$ threshold
Assess PH assmp. (1) Schoenfeld residuals, (2) log(-log(S(t)) plots - parallel lines, (3) covariate-time interaction.

---

homogeneous pois process: (1) $N(t) - N(s) \sim poi(\lambda(t - s))$ (2) interarrival time U_j idep $exp(1/\lambda)$ (3) num event in non-overlap tiem intervals $\perp$
Counting process: (1)$\tilde{N}_i(t) = \mathbf{1}_{\{\tilde{\mathbf{T}}_\mathbf{i} \le \mathbf{t}_\mathbf{i}\}}$; At-risk $Y_i(t) = 1_{\{\tilde{T}_i \ge t_i\}}$;
(2)jump: $d\tilde{N}_i(t) = \tilde{N}_i(t^- + dt) - \tilde{N}_i(t^-)$; (3) $P(d\tilde{N}_i(t) = 1|\ \tilde{N}_i(t) = 0)$ $= \lambda(t)dt$. (4) $N(t) = \sum N_i(t)$ (cum case); dN(t): jump/cases at t; $Y(t) = \sum Y_i(t)$ (5) $P(dN_i(t) = 1|F_{t-}) = E(dN_i(t)|F_{t-}) = Y_i(t)\lambda_i(t)dt$
intensity process $Y_i(t)\lambda_i(t)$ define diff $dM_i(t) = dN_i(t) - Y_i(t)\lambda_i(t)dt$ (O-E), take $\mathbf{E}$, $E[dM_i(t) \mid \mathcal{F}_{t-}] = E[dN_i(t) \mid \mathcal{F}_{t-}]^{=5}$ $-E[Y_i(t)\lambda_i(t)\,dt \mid \mathcal{F}_{t-}] = 0$
martingale ϵ $M_i(t) = N_i(t) - \int_0^t Y_i(u)\lambda_i(u)du$; $M_i = \delta_i - e^{\hat{\beta}^T z_i}\hat{H}_0(t_i)$; $E[M_i(t) \mid \mathcal{F}_s] = M_i(s)$, $E[M(t) - M(s) \mid \mathcal{F}_s] = 0$

---

competing risk $1 - p$(transplant) : die before trans or surv not trans.
cause-speci hazard $\lambda_j(t) \equiv \lim \text{P}(t \le \tilde{T}_i < t + h, \tilde{E}_i = j \mid \tilde{T}_i \ge t)/h$;
$f_j(t)dt = \lambda_j(t)\,dt\exp\left\{-\int_0^t\sum_{k=1}^J\lambda_k(u)\,du\right\}$ no event occ, $\sum f_j = 1$;
cox-type:$\lambda_{ij}(t) = \lambda_{0j}(t)\exp\{\beta_j' x_i\}$; cum incidence: $\hat{\pi}_{ij}(s) =$
$\sum_k d\hat{\Lambda}_{ij}(t_k)\exp\left\{-\sum_{j=1}^J\hat{\Lambda}_{ij}(t_{k-1})\right\}$, $\Lambda_{ij}(t) = \Lambda_{0j}(t)\exp\{\beta_j' x_i\}$ cum base from breslow; non-param: nelson+KM, AalenJohansen
$\hat{\pi}_j(s) = \sum_{k:t_k \le s}\frac{d_{jk}}{r_k}\hat{S}_{KM}(t_{k-1})$, $\hat{S} = \prod_{k:t_k \le t}\left(1 - \frac{\sum_{j=1}^J d_{jk}}{r_k}\right)$
subdist hazard at risk or devlp other event $\psi_{ij}(t) = \psi_{0j}(t)\exp\{\gamma_j' x_i\}$
,$\psi_j(t) \equiv \lim P\left(t \le \tilde{T}_i < t + h, \tilde{E}_i = j \mid \tilde{T}_i \ge t \text{ or } (\tilde{T}_i < t \text{ and } \tilde{E}_i \neq j)\right)$,
$\hat{\pi}_{ij}(s) = 1 - \exp\left\{-\hat{\Psi}_{0j}(s)\exp\left\{\hat{\gamma}_j' x_i\right\}\right\}$ connect cox-type-finegray
$\pi_j(t) = 1 - \exp\{-\Psi_j(t)\}$
Piecewise hazard: $\prod_{i=1}^n\prod_{k=1}^k\prod_{l=1}^l\lambda_{kl}^{d_{ikl}}exp(-y_{ikl}\lambda_{kl}dt)$ where $\sum_i^n d_{ikl} \sim pois(\lambda_{kl}, \sum_i^n y_{ikl})$, model $log(\lambda_{ikl}) = \alpha_k + \beta_l + \gamma' X_i$
$L(\beta) = \prod_{i=1}^n\prod_{k=1}^K[(\lambda_{0k}exp(\beta' x_i))^{d_{ik}}exp(-y_{ik}\lambda_{0k}exp(\beta' x_i))$,
$\hat{\lambda}_{0k}(\beta) = d_k/\{\sum_i^n y_{ik}exp(\beta' x_i)\}$
Diagnostics (1)Sen/TP = $P(\hat{\pi}_i(s) > \pi^*|\tilde{N}_i(s) = 1)$ and $PPV = P(\tilde{N}_i(s) = 1|\hat{\pi}_i(s) > \pi^*)$. (2) Spe/TN= $P(\hat{\pi}_i(s) < \pi^*|N_i(s) = 0)$ and NPV=$P(N_i(s) = 0|\hat{\pi}_i(s) < \pi^*)$
AUC sen vs 1-spe $P(\hat{\pi}_i(s) > \hat{\pi}_j(s)|N_i(s) = 1, N_j(s) = 0)$, w/o censoring, estimate as prop. of concordant pairs:$P(\hat{\pi}_i(s) > \hat{\pi}_j(s)|\tilde{T}_i < \tilde{T}_j)$ when censor, use C-index: $(T_{\text{censor}} > T_{\text{non-censor}})$.
Estimate sen/spe w/ censoring: use bayes & KM est.
Sen=$\frac{(1-P(N_i(s)=0|\hat{\pi}_i(s)\ge\pi^*)(1-P(\hat{\pi}_i(s)<\pi^*)}{1-P(N_i(s)=0)}$.
Spe= $\frac{P(N_i(s)=0|\hat{\pi}_i(s)<\pi^*)P(\hat{\pi}_i(s)<\pi^*)\text{[empirical prop]}}{P(N_i(s)=0)^{\text{s-year survival KM}}}$,
Calibration Hosmer-Lemeshow $\sum_{k=1}^K\frac{(O_k-E_k)^2}{N_k\bar{\pi}_k(1-\bar{\pi}_k)} \sim \chi_{k-2}^2$, $E_k = N_k\bar{\pi}_k$; $\bar{\pi}_k$ avg risk in grp k, w/ censor $O_k = N_k(1 - S_k(s))$ from KM. Brier score $(1/n)\sum_{i=1}^n\left(\tilde{N}_i(s) - \hat{\pi}_i(s)\right)^2$

Eliminate nuisance param: (1) data $y = (v, w)$, $P(v, w|\theta, \psi) = P(w|v, \theta)P(v|\theta, \psi)$ if $v3$ is ancillary. (2) Profile $\prod_{i=1}^n\prod_{k=1}^K[\lambda_{0k}\exp(\beta' x)]^{d_{ik}}\exp(-y_{ik}\lambda_{ik}\exp(\beta' x_i))$ (person i, int k). MLE BaseHaz $\hat{\lambda}_{0k} = \frac{d_k}{\sum_{i=1}^n y_{ik}exp(\beta x_i)}$ to get
$\prod_{i=1}^n\prod_{k=1}^k\left(\frac{exp(\beta x_i)}{\sum_{l=1}^n y_l(t_i)exp(\beta x_l)}\right)^{d_{ik}}$ where i =individual, k =interval.
Only i individuals with events contribute, Cox PH simplifies to
$\prod_{i=1}^n\left(\frac{exp(\beta' x_i)}{\sum_l^n Y_l(t_i)dtexp(\beta' x_l)}\right)^{e_i}$ where $Y_l(t_i) = 1_{\{T_l \ge t_i\}}$ and $e_i = 1_{\{T_i \le C_i\}}$. Partial lkhd for 1-1 match :
$P(D_{i1} = 1|D_{i1} + D_{i2} = 1, z_{i1}, x_{i1}, z_{i2}, x_{i2}, ;\theta) =$
$P(dN_i(t) = 1|dN_{i1}(t) + dN_{i2}(t) = 1, F_{t_i-}, \theta) =$
$\frac{exp(\alpha_i+\beta' Z_{i1}+\gamma' X_{i1})}{exp(\alpha_i+\beta' Z_{i1}+\gamma' X_{i1})+exp(\beta' Z_{i2}+\gamma' X_{i2})}$. Partial LKHD becomes
$\prod_{i=1}^d\frac{exp(\beta' Z_{i1}+\gamma' X_{i1})}{exp(\beta' Z_{i1}+\gamma' X_{i1})+exp(\beta' Z_{i2}+\gamma' X_{i2})}$ where $\theta = (\alpha, \beta, \gamma)$.

Counting Process: (1) $N(0) = 0$, (2) $N(t) \in \{\ldots, -1, 0, 1, \ldots\}$, (3 & 4) $s < t \Rightarrow N(s) < N(t)$ and $N(t) - N(s) = \#$ events in interval $(s, t]$.
$N(s + t) - N(s) \sim \text{Poisson}(\lambda t)$ Properties: (1) Memoryless $P(X > s + t|X > t) = P(X > s)\ \forall\ s, t \ge 0$. depends only on the length (2) Independent non-overlapping increments: $N(s + t) - N(s) \perp N(t)$. (3) Poisson-Gamma Duality: $N(t) \sim \text{Poiss}(\lambda t), S_n \sim \text{Gamma}(n, \lambda) \Rightarrow$ $P(N(t) \ge n) = P(S_n \le t)$, $N(t) = \sum N_i(t) \sim \text{Pois}(\sum\lambda_k)$ Interarrival Times $\{T_n, n \ge 1\} \sim \exp(\lambda)$ = time between $(n-1)^{\text{th}}$ and n^{th} event.
$P(T_1 > t) = P(N(t) = 0)$. $P(T_2 > t) = E[P(T_2 > t \mid T_1] =$
$\int_s P(N(t + s) - N(s) = 0 \mid N(s) = 1)ds \overset{\perp}{=} e^{-\lambda t}$
Waiting Times $S_n = \sum_{i=1}^n T_i \sim \text{Gamma}(n, \lambda)$,
Order stat: $P(T_1 < s_1, T_2 < s_1 + s_2, \ldots T_n < s_1 + \ldots + s_n|N(t) = n) =$ $n!\prod_{i=1}^n\frac{s_i}{t^n}\mathbb{1}(0 < s_{(1)} < \ldots < s_{(n)} < t) \Rightarrow U(0, t)$ Ran. Samp. Order S.
Cond. Dist: $T_1|N(t) = 1 \sim U(0, t)$, $N(s)|N(t) = n \sim Bin(n, \frac{s}{t})$
$T_i \sim exp(\lambda_i) \Rightarrow P(X_1 < X_2) = \int_0^\infty P(X_1 < X_2|X_1 = x)P(X_1 = x)dx =$
$\int_0^\infty P(x < X_2)\lambda_1\exp(-\lambda_1 x)dx = \frac{\lambda_1}{\lambda_1+\lambda_2}$
MC: at time n, dist of $X_n = \vec{S}_{1\times M}$ (m states) $\to$ dist of $X_{n+1} = \vec{S}Q$ as: $p(x_{n+1} = j) = \sum_i P(x_{n+1} = j|x_n = i)^{p_{ij}}P(x_n = i)^{s_i} = \sum_i s_i p_{ij}$