4. Nonparametric Hypothesis Testing

- This lecture's topics: weighted logrank tests
  - deriving test statistic
  - large sample null distribution
  - various forms of weights
- Text: FH Chapters 3 and 7

Nonparametric Setting

- Objective: compare two groups of patients with respect to survival; e.g.,
  - treatment: $j = 1$
  - placebo: $j = 0$
- Set-up:
  - covariate data are not available
  - assume that covariate distributions are equal across the groups being compared
- Want to make no parametric assumptions for the survival functions being compared

Notation: Two Sample Setting

- Observed data: (X_i, Δ_i, Z_i) for $i = 1, \ldots, n$
  - $X_i = T_i \wedge C_i$
  - $\Delta_i = I(T_i < C_i)$
  - $Z_i = I_i(\text{treated}) = 1 - I_i(\text{placebo})$
- Group-specific notation:
  - $S_j(t) = P(T_i > t | Z_i = j)$
  - $\Lambda_j(t) = \Lambda(t | Z_i = j)$
  - $G_j(t) = P(C_i > t | Z_i = j)$
  - $n_1 = \sum_{i=1}^n Z_i$, $n_0 = \sum_{i=1}^n (1 - Z_i)$,
  - $\tau_j = \max\{X_i : Z_i = j\}$
- Filtration: $\mathcal{F}(t) = \sigma\{N_i(s), Y_i(s+), Z_i;\ s \in (0, t],\ i = 1, \ldots, n\}$
- Counting processes and risk sets,
  - $Y_{\bullet j}(t) = \sum_{i=1}^n I(Z_i = j) Y_i(t)$
  - $N_{\bullet j}(t) = \sum_{i=1}^n I(Z_i = j) N_i(t)$
  - $\rho_j = P(Z_i = j)$
- Hypotheses:
  - $H_0 : S_0(t) = S_1(t)$, $\forall\ t$
  - $H_1 : S_0(t) \neq S_1(t)$, for some t

Weighted Logrank Test

- One can study the logrank test through the hypergeometric distribution
- We employ a counting process approach
- Consider the observed data at time t,

	$Z_i = 0$	$Z_i = 1$	total
deaths	$dN_{\bullet 0}(t)$	$dN_{\bullet 1}(t)$	$dN(t)$
survivors	$Y_{\bullet 0}(t) - dN_{\bullet 0}(t)$	$Y_{\bullet 1}(t) - dN_{\bullet 1}(t)$	$Y(t) - dN(t)$
total	$Y_{\bullet 0}(t)$	$Y_{\bullet 1}(t)$	$Y(t)$

- We arbitrarily select group 1 to compare to the average
- Under H_0, $\Lambda_0 = \Lambda_1 = \Lambda$, with a consistent estimator given by $\widehat{\Lambda} = \int Y^{-1} dN$
- Therefore, under H_0, the following quantity has mean 0,

$$\int_0^\infty \{dN_{\bullet 1}(t) - Y_{\bullet 1}(t) d\Lambda(t)\}$$

- Replace Λ with $\widehat{\Lambda}$,

$$U = \int_0^\infty \{dN_{\bullet 1}(t) - Y_{\bullet 1}(t) d\widehat{\Lambda}(t)\}$$

- Note that U jumps only at the observed death times
- Kernel of U has the spirit of *observed* - *expected* $(O - E)$, at time t
  - observed (O)= $dN_{\bullet 1}(t)$
  - expected (E)= $Y_{\bullet 1}(t) d\Lambda(t)$ = at risk $\times$ hazard
  - **Note**: expected = $dN(t) \times Y_{\bullet 1}(t)/Y(t)$ =deaths $\times$ fraction of sample from group 1

- We can re-write the integrand as follows,

$$
\begin{aligned}
& dN_{\bullet 1}(t) - Y_{\bullet 1}(t) d\widehat{\Lambda}(t) \\
= \; & dN_{\bullet 1}(t) - \frac{Y_{\bullet 1}(t)\{dN_{\bullet 0}(t) + dN_{\bullet 1}(t)\}}{Y_{\bullet 0}(t) + Y_{\bullet 1}(t)} \\
= \; & \frac{\{Y_{\bullet 0}(t) + Y_{\bullet 1}(t)\} dN_{\bullet 1}(t) - Y_{\bullet 1}(t)\{dN_{\bullet 0}(t) + dN_{\bullet 1}(t)\}}{\{Y_{\bullet 0}(t) + Y_{\bullet 1}(t)\}} \\
= \; & Y(t)^{-1}\{Y_{\bullet 0}(t) dN_{\bullet 1}(t) - Y_{\bullet 1}(t) dN_{\bullet 0}(t)\}
\end{aligned}
$$

- Define $\tau_{01} = \tau_0 \wedge \tau_1$

- The *unweighted* logrank test statistic is then given by,

$$
\begin{aligned}
U &= \int_0^{\tau_{01}} Y(t)^{-1}\{Y_{\bullet 0}(t) dN_{\bullet 1}(t) - Y_{\bullet 1}(t) dN_{\bullet 0}(t)\} \\
&= \int_0^{\tau_{01}} \frac{Y_{\bullet 0}(t) Y_{\bullet 1}(t)}{Y(t)} \{\frac{dN_{\bullet 1}(t)}{Y_{\bullet 1}(t)} - \frac{dN_{\bullet 0}(t)}{Y_{\bullet 0}(t)}\}
\end{aligned}
$$

- *Weighted* logrank statistic,

$$
U_W = \int_0^{\tau_{01}} W(t) Y(t)^{-1}\{Y_{\bullet 0}(t) dN_{\bullet 1}(t) - Y_{\bullet 1}(t) dN_{\bullet 0}(t)\}
$$

where, typically, the weight function satisfies

1. $W(t) = W(t^-)$ (left continuity)
2. $\sup_t |W(t) - w(t)| \xrightarrow{p} 0$ with $W(t)$ uniformly bounded

Setting Up Martingale Structure

- Set $M_{\bullet j}(t) = N_{\bullet j}(t) - \int_0^t Y_{\bullet j}(t) d\Lambda_j(t)$, for $j = 0, 1$
- $M_{\bullet j}(t)$ is a martingale with respect to the filtration $\mathcal{F}(t) = \sigma\{\mathcal{F}_{\bullet 0}(t) \cup \mathcal{F}_{\bullet 1}(t)\}$, where

$$\mathcal{F}_{\bullet j}(t) = \sigma\{Y_i(s+), N_i(s) : Z_i = j,\ s \in (0, t];\ i = 1, \ldots, n\}$$

- Since the individuals are independent, the treatment groups are independent, and $M_{\bullet 0}(t) \perp M_{\bullet 1}(t)$
- We now write the logrank test statistic in terms of martingales
  - note: $dN_{\bullet j}(t) = dM_{\bullet j}(t) + Y_{\bullet j}(t) d\Lambda_j(t)$

$$\begin{aligned}
U_W &= \int_0^{\tau_{01}} W(t) Y(t)^{-1} \{Y_{\bullet 0}(t) dN_{\bullet 1}(t) - Y_{\bullet 1}(t) dN_{\bullet 0}(t)\} \\
&= \int_0^{\tau_{01}} W(t) Y(t)^{-1} Y_{\bullet 0}(t) \{dM_{\bullet 1}(t) + Y_{\bullet 1}(t) d\Lambda_1(t)\} \\
&\quad - \int_0^{\tau_{01}} W(t) Y(t)^{-1} Y_{\bullet 1}(t) \{dM_{\bullet 0}(t) + Y_{\bullet 0}(t) d\Lambda_0(t)\} \\
&= \int_0^{\tau_{01}} W(t) Y(t)^{-1} \{Y_{\bullet 0}(t) dM_{\bullet 1}(t) - Y_{\bullet 1}(t) dM_{\bullet 0}(t)\} \\
&\quad + \int_0^{\tau_{01}} W(t) Y(t)^{-1} Y_{\bullet 0}(t) Y_{\bullet 1}(t) d\{\Lambda_1(t) - \Lambda_0(t)\}
\end{aligned}$$

Mean of Null Distribution

- Now, under H_0, $\Lambda_1(t) = \Lambda_0(t)$, such that the second term vanishes, leaving

$$U_W = \int_0^{\tau_{01}} W(t)Y(t)^{-1}\{Y_{\bullet 0}(t)dM_{\bullet 1}(t) - Y_{\bullet 1}(t)dM_{\bullet 0}(t)\}$$

- A linear combination of $\mathcal{F}$ martingales will also be an $\mathcal{F}$ martingale
  - therefore, $E[U_W] = 0$

- Having derived the mean under the null distribution of U_W, we now focus on the variance

Variance of Null Distribution

- We now set $U_W = U_{W1} - U_{W2}$, where

$$\begin{aligned} U_{W1} &= \int_0^{\tau_{01}} W(t)Y(t)^{-1}Y_{\bullet 0}(t)dM_{\bullet 1}(t) \\ U_{W2} &= \int_0^{\tau_{01}} W(t)Y(t)^{-1}Y_{\bullet 1}(t)dM_{\bullet 0}(t) \end{aligned}$$

- Since U_{W1} and U_{W2} are uncorrelated,

$$V(U_W) = V(U_{W1}) + V(U_{W2})$$

- From the basic properties of martingales, $V(U_{Wk}) = E[\langle U_{Wk}\rangle]$, with the predictable variation process computed as

$$
\begin{aligned}
\langle U_W \rangle &= \langle U_{W1} \rangle + \langle U_{W2} \rangle \\
&= \left\langle \int_0^{\tau_{01}} W(t)Y(t)^{-1}Y_{\bullet 0}(t)dM_{\bullet 1}(t) \right\rangle \\
&\quad + \left\langle \int_0^{\tau_{01}} W(t)Y(t)^{-1}Y_{\bullet 1}(t)dM_{\bullet 0}(t) \right\rangle \\
&= \int_0^{\tau_{01}} W^2(t)Y(t)^{-2}Y_{\bullet 0}(t)^2 dA_{\bullet 1}(t) \\
&\quad + \int_0^{\tau_{01}} W^2(t)Y(t)^{-2}Y_{\bullet 1}(t)^2 dA_{\bullet 0}(t) \\
&= \int_0^{\tau_{01}} W^2(t)Y(t)^{-2}Y_{\bullet 0}(t)^2 Y_{\bullet 1}(t) d\Lambda_1(t) \\
&\quad + \int_0^{\tau_{01}} W^2(t)Y(t)^{-2}Y_{\bullet 1}(t)^2 Y_{\bullet 0}(t) d\Lambda_0(t)
\end{aligned}
$$

- Therefore, under H_0,

$$
\begin{aligned}
\langle U_W \rangle &= \int_0^{\tau_{01}} W^2(t)Y(t)^{-2}Y_{\bullet 0}(t)Y_{\bullet 1}(t)\{Y_{\bullet 0}(t) + Y_{\bullet 1}(t)\} d\Lambda(t) \\
&= \int_0^{\tau_{01}} W^2(t)Y(t)^{-1}Y_{\bullet 0}(t)Y_{\bullet 1}(t) d\Lambda(t)
\end{aligned}
$$

such that the variance under H_0 is given by

$$
\begin{aligned}
V(U_W) &= E[\langle U_W \rangle] \\
&= E\left[\int_0^{\tau_{01}} W^2(t)Y(t)^{-1}Y_{\bullet 0}(t)Y_{\bullet 1}(t) d\Lambda(t) \right]
\end{aligned}
$$

- We estimate the variance by substituting $\widehat{\Lambda}$ for Λ,

$$\begin{aligned}\widehat{V}(U_W) &= \int_0^{\tau_{01}} W^2(t)Y(t)^{-1}Y_{\bullet 0}(t)Y_{\bullet 1}(t)d\widehat{\Lambda}(t) \\ &= \int_0^{\tau_{01}} W^2(t)Y(t)^{-2}Y_{\bullet 0}(t)Y_{\bullet 1}(t)dN(t)\end{aligned}$$

- This is an unbiased estimator of $V(U_W)$ since

$$\int_0^{\tau_{01}} W^2(t)Y(t)^{-2}Y_{\bullet 0}(t)Y_{\bullet 1}(t)dM(t)$$

  is an $\mathcal{F}$ martingale and hence has mean 0; i.e.,

$$\begin{aligned}& E\left[\int_0^{\tau_{01}} W^2(t)Y(t)^{-2}Y_{\bullet 0}(t)Y_{\bullet 1}(t)dN(t)\right] \\ =\ & E\left[\int_0^{\tau_{01}} W^2(t)Y(t)^{-1}Y_{\bullet 0}(t)Y_{\bullet 1}(t)d\Lambda(t)\right]\end{aligned}$$

  where the second expression equals $V(U_W)$

- Asymptotic null distribution of Logrank statistic:

$$\text{Under } H_0,\ n^{-1/2}U_W \xrightarrow{D} N(0, \sigma_U^2)$$

  Note:

  – By MCLT

  – $\langle n^{-1/2}U_W\rangle \xrightarrow{p} \sigma_U^2$

  – Two conditions required for a use of the MCLT can be verified...

Logrank Test: Summary

- To summarize, for the weighted logrank test, under H_0 we have:
  - test statistic, numerator:

$$U_W = \int_0^{\tau_{01}} W(t)Y(t)^{-1}\{Y_{\bullet 0}(t)dN_{\bullet 1}(t) - Y_{\bullet 1}(t)dN_{\bullet 0}(t)\}$$

  - mean, asymptotic variance:

$$E[U_W] = 0 \text{ and } V(U_W) = E\left[\int_0^{\tau_{01}} W^2(t)Y(t)^{-1}Y_{\bullet 0}(t)Y_{\bullet 1}(t)d\Lambda(t)\right]$$

  - asymptotic distribution:

$$n^{-1/2}U_W \xrightarrow{D} N(0, \sigma_U^2)$$

  - unbiased, consistent variance estimator:

$$\widehat{\sigma}_U^2 = n^{-1}\int_0^{\tau_{01}} W^2(t)Y(t)^{-2}Y_{\bullet 0}(t)Y_{\bullet 1}(t)dN(t)$$

  - test:

$$\frac{n^{-1/2}U_W}{\widehat{\sigma}_U} \xrightarrow{D} N(0,1)$$

Weighted Logrank Test Under Local Alternatives

- So far, we have considered the behavior of the logrank test under the null
- We now evaluate the distribution of U_W when $S_0(t) \neq S_1(t)$
- In particular, we consider local alternatives of the form $\lambda_{1n}(t) = \lambda_0(t)\exp\{\beta_n\theta(t)\}$, where

- $\theta(t)$ is a known function of t
- $\beta_n \to 0$ at rate $n^{-1/2}$

  $n^{1/2}\beta_n \to \tau$

- Note: $\theta(t)$ is proportional to the log hazard ratio,

$$\log\left\{\frac{\lambda_{1n}(t)}{\lambda_0(t)}\right\} = \beta_n\theta(t)$$

- Weighted logrank statistic (recall), normalized:

$$n^{-1/2}U_W = n^{-1/2}\int_0^{\tau_{01}} \frac{W(t)}{Y(t)}\{Y_{\bullet 0}(t)dN_{\bullet 1}(t) - Y_{\bullet 1}(t)dN_{\bullet 0}(t)\}$$

- This can be decomposed as

$$\begin{aligned} n^{-1/2}U_W &= n^{-1/2}\int_0^{\tau_{01}} \frac{W(t)}{Y(t)}\{Y_{\bullet 0}(t)dM_{\bullet 1}(t) - Y_{\bullet 1}(t)dM_{\bullet 0}(t)\} \\ &\quad + n^{-1/2}\int_0^{\tau_{01}} \frac{W(t)}{Y(t)}Y_{\bullet 0}(t)Y_{\bullet 1}(t)\{\lambda_{1n}(t) - \lambda_0(t)\}dt \\ &= n^{-1/2}U_{W1}^1 + n^{-1/2}U_{W2}^1 \end{aligned}$$

- Under H_0, the second term disappears; under H_{1n} it can be written as $n^{-1/2}U_{W2}^1 =$

$$\begin{aligned} & n^{-1/2}\int_0^{\tau_{01}} \frac{W(t)}{Y(t)}Y_{\bullet 0}(t)Y_{\bullet 1}(t)[\lambda_0(t)\exp\{\beta_n\theta(t)\} - \lambda_0(t)]dt \\ = \; & n^{-1/2}\int_0^{\tau_{01}} \frac{W(t)}{Y(t)}Y_{\bullet 0}(t)Y_{\bullet 1}(t)[\exp\{\beta_n\theta(t)\} - 1]\lambda_0(t)dt \end{aligned}$$

- Using a first-order Taylor series expansion,

$$\exp\{\beta_n\theta(t)\} \approx 1 + \beta_n\theta(t)$$

  and recall that $n^{1/2}\beta_n \to \tau$, such that $n^{-1/2}U_{W2}^1 =$

$$\begin{aligned} & n^{-1}\int_0^{\tau_{01}} W(t)Y(t)^{-1}Y_{\bullet 0}(t)Y_{\bullet 1}(t)n^{1/2}\beta_n\theta(t)\lambda_0(t)dt \\ = \; & n^{-1}\int_0^{\tau_{01}} W(t)Y(t)^{-1}Y_{\bullet 0}(t)Y_{\bullet 1}(t)\tau\theta(t)\lambda_0(t)dt \end{aligned}$$

- Applying the WLLN and continuity,

$$\begin{aligned} n^{-1}\frac{W(t)}{Y(t)}Y_{\bullet 0}(t)Y_{\bullet 1}(t) &= \frac{W(t)}{\widehat{\pi}(t)}\frac{Y_{\bullet 0}(t)}{n_0}\frac{Y_{\bullet 1}(t)}{n_1}\frac{n_0}{n}\frac{n_1}{n} \\ &\xrightarrow{p} \frac{w(t)\rho_0\rho_1 G_0(t)S_0(t)G_1(t)S_{1n}(t)}{\pi(t)} \\ &= \frac{w(t)\rho_0\rho_1 G_0(t)S_0(t)G_1(t)S_{1n}(t)}{\rho_0 G_0(t)S_0(t) + \rho_1 G_1(t)S_{1n}(t)} \end{aligned}$$

  – Under H_{1n},

$$S_{1n}(t) = \exp\{-\Lambda_{1n}(t)\} = \exp\left\{-\int_0^t \lambda_0(s)\exp\{\beta_n\theta(s)\}ds\right\}$$

  – Now, since $\beta_n \to 0$ as $n \to \infty$,

$$S_{1n}(t) - S_0(t) \ \to \ 0$$

- Incorporating this result,

$$n^{-1}\frac{W(t)}{Y(t)}Y_{\bullet 0}(t)Y_{\bullet 1}(t) \ \xrightarrow{p} \ \frac{w(t)\rho_0\rho_1 G_0(t)S_0(t)G_1(t)}{\rho_0 G_0(t) + \rho_1 G_1(t)}$$

- Combining these ideas, we have

$$n^{-1/2}U^1_{W2} \ \xrightarrow{p} \ \int_0^{\tau_{01}} \tau\theta(t)w(t)\frac{\rho_0\rho_1 G_0(t)S_0(t)G_1(t)}{\rho_0 G_0(t) + \rho_1 G_1(t)}\lambda_0(t)dt$$

  which is a constant

- We now consider the first term, $n^{-1/2}U^1_{W1}$, under H_{1n}

$$n^{-1/2}U^1_{W1} \ = \ n^{-1/2}\int_0^{\tau_{01}} \frac{W(t)}{Y(t)}\{Y_{\bullet 0}(t)dM_{\bullet 1}(t) - Y_{\bullet 1}(t)dM_{\bullet 0}(t)\}$$

  – this term remains a martingale
  – still $\xrightarrow{W}$ to a zero-mean Gaussian process

– variance of limiting distribution is given by limit of $\langle n^{-1/2}U^1_{W1}\rangle$, which equals

$$\begin{aligned} & n^{-1}\int_0^{\tau_{01}} \frac{W^2(t)}{Y(t)^2}\{Y_{\bullet 0}(t)^2 dA_{1n}(t) + Y_{\bullet 1}(t)^2 dA_0(t)\} \\ = \; & n^{-1}\int_0^{\tau_{01}} \frac{W^2(t)}{Y(t)^2} \\ & \{Y_{\bullet 0}(t)^2 Y_{\bullet 1}(t)\lambda_{1n}(t) + Y_{\bullet 1}(t)^2 Y_{\bullet 0}(t)\lambda_0(t)\}dt \end{aligned}$$

– since $\lambda_{1n}(t) = \lambda_0(t)\exp\{\beta_n\theta(t)\}$, with $\beta_n \to 0$ as $n \to \infty$, $\lambda_{1n}(t) \to \lambda_0(t)$, such that

$$\begin{aligned} \langle n^{-1/2}U^1_{W1}\rangle &\xrightarrow{p} \rho_0\rho_1\int_0^{\tau_{01}} \frac{w^2(t)G_0(t)G_1(t)S(t)}{\rho_0 G_0(t)+\rho_1 G_1(t)}\lambda_0(t)dt \\ &\equiv \sigma_U^2 \end{aligned}$$

Summary: Logrank under Local Alternatives

- Under H_{1n}, $n^{-1/2}U^1_W = n^{-1/2}U^1_{W1} + n^{-1/2}U^1_{W2}$, where

$$\begin{aligned} n^{-1/2}U^1_{W1} &= n^{-1/2}\int_0^{\tau_{01}} \frac{W(t)}{Y(t)}\{Y_{\bullet 0}(t)dM_{\bullet 1}(t) - Y_{\bullet 1}(t)dM_{\bullet 0}(t)\} \\ n^{-1/2}U^1_{W2} &= n^{-1/2}\int_0^{\tau_{01}} \frac{W(t)}{Y(t)}Y_{\bullet 0}(t)Y_{\bullet 1}(t)\{\lambda_{1n}(t)-\lambda_0(t)\}dt \end{aligned}$$

- Under local alternatives of the form $\lambda_{1n}(t) = \lambda_0(t)\exp\{\beta_n\theta(t)\}$,

$$\begin{aligned} n^{-1/2}U^1_{W1} &\xrightarrow{D} N(0,\sigma_U^2) \\ n^{-1/2}U^1_{W2} &\xrightarrow{p} \int_0^{\tau_{01}} \tau\theta(t)w(t)\frac{\rho_0\rho_1 G_0(t)S_0(t)G_1(t)}{\rho_0 G_0(t)+\rho_1 G_1(t)}\lambda_0(t)dt \\ &\equiv \mu_U \end{aligned}$$

- Therefore, $n^{-1/2}U^1_W \xrightarrow{D} N(\mu_U, \sigma_U^2)$, where:

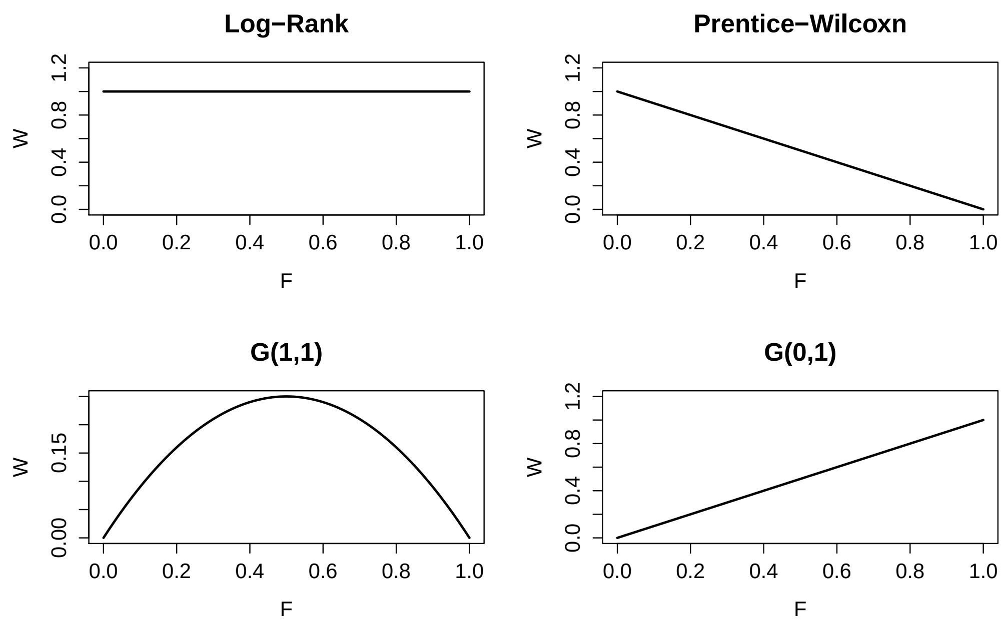
$$\begin{aligned} \mu_U &= \int_0^{\tau_{01}} \tau\theta(t)w(t)\frac{\rho_0\rho_1 G_0(t)S_0(t)G_1(t)}{\rho_0 G_0(t)+\rho_1 G_1(t)}\lambda_0(t)dt \\ \sigma_U^2 &= \rho_0\rho_1\int_0^{\tau_{01}} \frac{w^2(t)G_0(t)G_1(t)S(t)}{\rho_0 G_0(t)+\rho_1 G_1(t)}\lambda_0(t)dt \end{aligned}$$

Choice of Weight Function

- The following table gives examples of some common weight functions

Statistic	$W(t)$
Logrank	1
Prentice-Wilcoxon	$\widehat{S}(t^-)$
Harrington-Fleming G^ρ	$\widehat{S}(t^-)^\rho$
Gehan-Wilcoxon	$\widehat{\pi}(t^-)$
Tarone-Ware	$\widehat{\pi}(t^-)^\rho$
$G^{\rho,\gamma}$	$\widehat{S}(t^-)^\rho(1-\widehat{S}(t^-))^\gamma$

- Most common choices: logrank, Gehan-Wilcoxon



- *Logrank* applies equal weight to all t
  - most powerful under proportional hazards alternatives
  - a good choice intuitively, if investigator derives equal information from early and late deaths

- Wilcoxon over-weights early failure times

- **Question**: what is the optimal weighting function $w(u)$?

- Optimal weight function depends on distribution of deaths in both $j = 0$ and $j = 1$ groups

- Power increases as the noncentrality parameter μ_U/σ_U increases.

- The function $w(u)$ which maximazies μ_U/σ_U will yield the optimal solution.

- Cauchy-Schwartz inequality $\Rightarrow$

$$\int_0^\infty w(u)\theta(u)h(u)du \leq \left\{\int_0^\infty w^2(u)h(u)du \int_0^\infty \theta^2(u)h(u)du\right\}^{1/2},$$

  with equality iff $w(\cdot) \propto \theta(\cdot)$

- Therefore,

$$\frac{\int_0^\infty w(u)\theta(u)h(u)du}{\left\{\int_0^\infty w^2(u)h(u)du\right\}^{1/2}} \leq \left\{\int_0^\infty \theta^2(u)h(u)du\right\}^{1/2}$$

  where $h(u) = \frac{\rho_0\rho_1 G_0(t)S_0(t)G_1(t)}{\rho_0 G_0(t)+\rho_1 G_1(t)}\lambda_0(t)$,
  with equality iff $w(\cdot) \propto \theta(\cdot)$

- Optimal weight: $w(t) \propto \theta(t)$, $\forall\, t$
  - recall: $\theta(t) \propto \log\{\lambda_1(t)/\lambda_0(t)\}$
  - e.g., if $\theta(t) = \theta$, then $w(t) = 1$ is optimal

- The greatest power for detecting alternatives where the log hazard ratio is given by $\theta(u)$, among wegithed logrank tests, is obtained by choosing $W(u)$ so that $W(u) \xrightarrow{p} \theta(u)$.

Commonly Encounted Weights

1. Gehan/Breslow Generalized Wilcoxon test

- $w_i = Y(t_i)$, #at risk, or, $Y(t_i)/n = \hat{\pi}(t_i)$, where t_i are distinct failure times.
- Under H_0,

$$X_G^2 = \left(\frac{n^{-1/2}U_G}{\widehat{\sigma}_G}\right)^2 \xrightarrow[n\to\infty]{D} \chi_1^2$$

- Better than the logrank test at detecting early differences; worse at detecting later differences.

Note:

- a whole family of tests was proposed by Tarone & Ware:

$$w_i = Y(t_i)^\alpha \qquad \underset{\text{logrank}}{\overset{\uparrow}{0}} \ \le\ \alpha\ \le\ \underset{\text{Gehan/Breslow}}{\overset{\uparrow}{1}}$$

- Potential problem with X_G^2: If there is heavy censoring between t_{i-1} and $t_i, Y(t_i)$ is much smaller than $Y_{t_{i-1}}$. Any information after t_{i-1} would be given very little weight, and not because the $\lambda(t)'$s were small, but because of the lack of follow-up.

2. Peto/Prentice Generalized Wilcoxon Test

- $$w_i = \widetilde{S}(t_i) = \prod_{j=1}^{i} \frac{Y(t_j)}{Y(t_j)+D_j} = \prod_{j=1}^{i} \frac{S_j + D_j}{Y(t_j)+D_j}$$
- Under H_0,

$$X_P^2 = \left(\frac{n^{-1/2}U_P}{\widehat{\sigma}_P}\right)^2 \xrightarrow[n\to\infty]{D} \chi_1^2$$

- Like $w_i = Y(t_i), w_i = \widetilde{S}(t_i)$ decreases as time passes, but it doesn't jump as wildly due to censoring
- A whole family was proposed by Fleming & Harrington

$$W_i = (\widetilde{S}(t_i))^\alpha \quad \begin{array}{lll} \alpha = 0 & \Rightarrow & \text{logrank} \\ \alpha = 1 & \Rightarrow & X_P^2 \end{array}$$

- Both X_G^2 and X_P^2 are generalized Wilcoxon tests because when there is no censoring they are both identical to the Wilcoxon rank sum tests.

- All three of X_L^2,X_G^2 and X_P^2 are rank tests. They do not depend on the actual survival times, just the order they come in (the ranks).

Log-rank Test for Three or More Groups

- Test for differences in $\lambda(t)$ among $K > 2$ groups

$$\begin{array}{ll} & H_0 : \lambda_0(t) \equiv \lambda_1(t) \equiv \cdots \equiv \lambda_{K-1}(t) \text{ for all } t \\ \text{vs.} & \quad H_1 : \neq \text{somewhere (heterogeneity)} \end{array}$$

- Consider the observed data at time t,

	$Z_i = 0$	$Z_i = 1$	$\cdots$	Z_{K-1}	total
deaths	$dN_{\bullet 0}(t)$	$dN_{\bullet 1}(t)$	$\cdots$	$dN_{\bullet K-1}(t)$	$dN(t)$
survivors	$Y_{\bullet 0}(t) - dN_{\bullet 0}(t)$	$Y_{\bullet 1}(t) - dN_{\bullet 1}(t)$	$\cdots$	$Y_{\bullet K-1}(t) - dN_{\bullet K-1}(t)$	$Y(t) - dN(t)$
total	$Y_{\bullet 0}(t)$	$Y_{\bullet 1}(t)$	$\cdots$	$Y_{\bullet 1}(t)$	$Y(t)$

- Under H_0, $\Lambda_0 = \Lambda_1 = \cdots \Lambda_{K-1} = \Lambda$, with a consistent estimator given by $\widehat{\Lambda} = \int Y^{-1} dN$

- The test statistic is a generalization of the 2-sample statistic. Define the processes

$$U_h = \int_0^\infty W(t) \left\{ dN_j(t) - Y_{\bullet j}(t) d\hat{\Lambda}(t) \right\}$$

for $j = 0, 1, \ldots, K-1$.

- With $K > 2$, it now depends on the covariance between the U_j's for each group. A consistent estimator for the covariance between U_j and $U_{j'}$ is

$$V_{hj} = \int_0^\infty W^2(t)\frac{Y_{\bullet h}(t)}{Y(t)}\left(\delta_{hj} - \frac{Y_j(t)}{Y(t)}\right) dN(t)$$

where δ_{hj} is a Kronecker delta (= 1 when $h = j$ and = 0 otherwise).

- Note that $\sum_{j=0}^{K-1} U_j = 0$ (Check!). So, we consider only the first $K-1$ U_j's in constructing our test statistic.

- Let $\boldsymbol{U} = (Z_0, Z_i, \cdots, Z_{K-2})^\top$ is a $(K-1)\times 1$ vector and $\boldsymbol{V}$ is a $(K-1)\times(K-1)$ matrix whose entry (h, j) is equal to V_{hj} for $h, j = 0, 1, \ldots, K-2$. Then, the test statistic is

$$X_K^2 = \boldsymbol{Z}^\top \boldsymbol{V}^{-1}\boldsymbol{Z} \xrightarrow[n\to\infty]{D} \chi^2_{K-1}.$$

- An alternative expression a using different notation is given as follows:
  - Define $t_i (i = 1, \ldots, I)$ to be ordered observed failure times from the pooled (combined) sample $(t_1 < t_2 < \cdots < t_I)$.
  - For Group k $(k = 1, \ldots, K)$, define d_{ki}, n_{ki} and S_{ki} to be the numbers of failures at t_i, at risk at t_i, and survived past t_i (including those who are censored at t_i), respectively. D_i, S_i and $n + i$ are their respective sums across the groups.
  - Data at t_i in the pooled sample:

at t_i:		1	2		k		K	
	d	d_{1i}	d_{2i}	$\cdots$	d_{ki}	$\cdots$	d_{Ki}	D_i
	S	S_{1i}	S_{2i}		S_{ki}		S_{Ki}	S_i
		n_{1i}	n_{2i}		n_{ki}		n_{Ki}	n_i

–

$$O_{ki} = d_{ki}(k = 1, \cdots, K-1), \quad E_{ki} = \frac{n_{ki} D_i}{n_i}$$
$$V_{kki} = \frac{n_{ki}(n_i - n_{ki}) D_i S_i}{n_i^2 (n_i - 1)}, V_{kk'i} = -\frac{n_{ki} n_{k'i} D_i S_i}{n_i^2 (n_i - 1)}$$
$$O_k - E_k = \sum_{i=1}^{I} (O_{ki} - E_{ki}), \quad V_{kk} = \sum_{i=1}^{I} V_{kki}, \quad V_{kk'} = \sum_{i=1}^{I} V_{kk'i}$$

where $k = 1, \cdots, K-1$:

$$V = \begin{pmatrix} V_{11} & V_{12} & \cdots & V_{1,K-1} \\ V_{21} & V_{22} & & \cdot \\ & & & \cdot \\ V_{1,K-1} & \cdots & \cdot & V_{K-1,K-1} \end{pmatrix}$$

$$X_K^2 = \begin{pmatrix} O_1 - E_1 \\ O_2 - E_2 \\ \vdots \\ O_{K-1} - E_{K-1} \end{pmatrix}^\top V^{-1} \begin{pmatrix} O_1 - E_1 \\ O_2 - E_2 \\ \vdots \\ O_{K-1} - E_{K-1} \end{pmatrix} \underset{H_0}{\overset{\text{large } n}{\sim}} \chi^2_{K-1}$$

Stratified Tests

- What if we want to test for differences in risk, adjusted for some confounding factor?

One solution: stratify

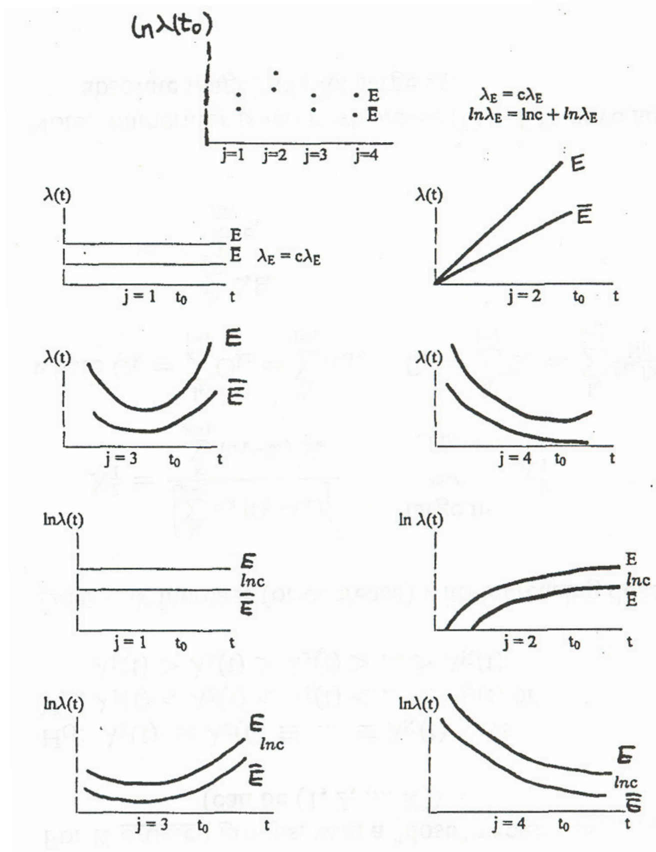
- We use the alternative expression.
- Suppose the stratification factor has J levels: $j = 1, \cdots, J$

For the logrank test, we would test

$$\begin{aligned} & H_0 : \lambda_j(t|E) = \lambda_j(t|\overline{E}) \quad \{\text{where } j = 1, \cdots, J\} \ (c = 1) \\ \text{vs. } & H_1 : \lambda_j(t|E) = c\, \lambda_j(t|\overline{E}) \quad \{\text{where } j = 1, \cdots, J;\ c \neq 1\} \end{aligned}$$

where E and $\overline{E}$ denote Groups 0 and 1.

- See pictures and note: $H_1 \Leftrightarrow \log\ \lambda_j(t|E) = \log\ c + \log\ \lambda_j(t|\overline{E}) \quad j = 1, \cdots, J$



- Stratified versions of X_L^2, X_G^2 and X_P^2:

$$X^2 = \frac{\left[\sum_{i1=1}^{I_1} w_{i1}(O_{i1}-E_{i1}) + \sum_{i2=1}^{I_2} w_{i2}(O_{i2}-E_{i2}) + \cdots + \sum_{iJ=1}^{I_J} \overbrace{w_{iJ}(O_{iJ}-E_{iJ})}^{\text{Stratum } J}\right]^2}{\underbrace{\sum_{i1=1}^{I_1} w_{i1}^2 V_{i1}}_{\text{Stratum } 1} + \underbrace{\sum_{i2=1}^{I_2} w_{i2}^2 V_{i2}}_{\text{Stratum } 2} + \cdots + \underbrace{\sum_{iJ=1}^{I_J} w_{iJ}^2 V_{iJ}}_{\text{Stratum } J}}$$

where O_{ij}, E_{ij}, V_{ij} and w_{ij} are calculated solely from subjects in the j^{th} stratum.

- Under H_0 and for large $n, X^2 \sim \chi^2_1$.

Log-rank test for Trend

- For K ordered groups, with a "dose" vector $(\omega_1, \cdots, \omega_K)$ (can be $(1, 2, \cdots, K)$)

$$\begin{aligned} H_0\text{:} &\quad \lambda_1(t) \equiv \lambda_2(t) \equiv \cdots \equiv \lambda_K(t) \text{ for all } t \\ \text{vs. } H_a\text{:} &\quad \lambda_1(t) < \lambda_2(t) < \cdots < \lambda_K(t) \text{ or} \\ &\quad \lambda_1(t) > \lambda_2(t) > \cdots > \lambda_K(t) \end{aligned}$$

- Does risk increase (or decrease) with increasing dose?

- We use the alternative expression:

$$X_T^2 = \frac{[\sum\limits_{k=1}^{K} \omega_k(O_k - E_k)]^2}{\sum\limits_{k=1}^{K} (\omega_k - \overline{\omega})^2 E_k} \quad \underset{H_0}{\overset{\text{large } n}{\sim}} \chi_1^2$$

$$\text{where } O_k = \sum_{i=1}^{I_k} O_{ki} = \sum_{i=1}^{I_k} d_{ki}, \quad E_k = \sum_{i=1}^{I_k} E_{ki} = \sum_{i=1}^{I_k} \frac{n_{ki} D_i}{N_i}$$

$$\overline{\omega} = \frac{\sum\limits_{i=1}^{K} \omega_k E_k}{\sum\limits_{k=1}^{K} E_k}$$

Sample Size Computation

- We have derived the distribution of $n^{-1/2}U_W$ under $H_0 : \lambda_1(t) = \lambda_0(t)$ and under $H_{1n} : \lambda_{1n}(t) = \lambda_0(t)\exp\{\beta_n\theta(t)\}$
  - in each case, the derived results apply as $n \to \infty$
- In planning a study, we assume a fixed treatment effect, and determine the minimum sample size to reject H_0 with a given pre-specified power at an acceptable Type I error probability
- Notation:

$$\begin{aligned}
&n_* = \text{minimum required sample size}\\
&\beta = \text{Type II error rate}\\
&\alpha = \text{Type I error rate}\\
&H_1 : \lambda_1(t) = \lambda_0(t)\exp\{\beta_0\theta(t)\}
\end{aligned}$$

- Test statistic (recall):

$$\begin{aligned}
T_W &\equiv \frac{n^{-1/2}U_W}{\widehat{\sigma}_U}\\
&\sim N(0,1), \text{under } H_0\\
&\sim N\left(\frac{\mu_U}{\sigma_U}, 1\right), \text{under } H_1
\end{aligned}$$

- Let $\Phi(z) = P\{N(0,1) \le z\}$, and let z_q be the qth quantile of the $N(0,1)$
- Desired power:

$$\begin{aligned}
(1-\beta) &= P(T_W > z_{1-\alpha}|H_1)\\
&= P\left(T_W - \frac{\mu_U}{\sigma_U} > z_{1-\alpha} - \frac{\mu_U}{\sigma_U}\right)\\
&= 1 - \Phi\left(z_{1-\alpha} - \frac{\mu_U}{\sigma_U}\right)
\end{aligned}$$

- Therefore, we have

$$\begin{aligned} \beta &= \Phi\left(z_{1-\alpha} - \frac{\mu_U}{\sigma_U}\right) \\ \Phi^{-1}(\beta) &= z_{1-\alpha} - \frac{\mu_U}{\sigma_U} \\ z_\beta &= z_{1-\alpha} - \frac{\mu_U}{\sigma_U} \\ z_{1-\beta} + z_{1-\alpha} &= \frac{\mu_U}{\sigma_U} \end{aligned}$$

- Now, $\mu_U = \tau\sigma_U^2$, based on the expression for μ_U and σ_U^2 assuming that $w(t) = \theta(t)$, with $\tau = n^{1/2}\beta_0$, which gives

$$\begin{aligned} z_{1-\beta} + z_{1-\alpha} &= \frac{\mu_U}{\sigma_U} = \tau\sigma_U \\ &= n^{1/2}\beta_0\sigma_U \end{aligned}$$

which yields the required sample size,

$$n_* = \frac{(z_{1-\beta} + z_{1-\alpha})^2}{\beta_0^2\sigma_U^2}$$

- Unlike the familiar sample size formulas, σ_U^2 is not easy to describe
  - greatly limits the applicability of the above formula

- Various other simplifications are typically employed; e.g.,

$$\begin{aligned} n_* &= \frac{(z_{1-\beta} + z_{1-\alpha})^2}{\beta_0^2\sigma_U^2} \\ &= \frac{(z_{1-\beta} + z_{1-\alpha})^2}{\beta_0^2\rho_0\rho_1 E[\Delta_i]}, \quad \text{if } w(t) = 1, G_0 = G_1 \\ &= \frac{4(z_{1-\beta} + z_{1-\alpha})^2}{\beta_0^2 E[\Delta_i]}, \quad \text{if } \rho_0 = \rho_1 = 0.5 \end{aligned}$$

- Therefore, the expected number of deaths is given by

$$n_* \times E[\Delta_i] = \frac{4(z_{1-\beta} + z_{1-\alpha})^2}{\beta_0^2}$$

Expected Number of Deaths

- To calculate the expected number of failures, one needs to specify the failure and censoring distribution
- Usually, the expected number of failures is computed separately for each treatment group, under H_1
- Often, it is assumed that there is no loss to follow-up
  - all censoring is administrative
  - in which case n_* can be determined by

$$n_*\{\rho_0 E[\Delta_i|Z_i = 0] + \rho_1 E[\Delta_i|Z_i = 1]\} = \frac{(z_{1-\beta} + z_{1-\alpha})^2}{\beta_0^2 \rho_0 \rho_1}$$