Review class of two-sample test:

Two groups, $j = 1, 2$, with hazards $\alpha_j(t)$, counting processes $N_j(t)$, no. at risk $Y_j(t)$ and Nelson-Aalen estimators $\hat{A}_j(t) = \int_0^t dN_j(s)/Y_j(s)$.

Let $L(t)$ be a general predictable weight function and define a general two-sample test by

$$Z_1 = \int_0^{t_0} L(s)[d\hat{A}_1(s) - d\hat{A}_1(s)] = \int_0^{t_0} L(s)[\frac{dN_1(s)}{Y_1(s)} - \frac{dN_2(s)}{Y_2(s)}]$$

Under $\mathrm{H}_0 : \alpha_1(t) = \alpha_2(t) = \alpha(t)$ Z_1 has expectation 0 and variance

$$\mathrm{E}[\int_0^{t_0} L(s)^2 (\frac{1}{Y_1(s)} + \frac{1}{Y_2(s)}) \alpha(s) ds]$$

which may be estimated

$$V_{11} = \int_0^{t_0} \frac{L(s)^2}{Y_1(s)Y_2(s)} dN_\bullet(s)$$

Some weight functions

Let $L(t) = K(t)\frac{Y_1(s)Y_2(s)}{Y_\bullet(s)}$

Choices of $K(t)$ up or down-weights early and late events:

- $K(t) = 1$ gives log-rank
- $K(t) = Y(t)$ Gehan's generalisation of the Wilcoxon-test
- $K(t) = \hat{S}(t)$ Peto and Prentice generalization of Wilcoxon
- $K(t) = \hat{S}(t)^p(1 - \hat{S}(t))^q$ Fleming and Harrington generalization of Prentice' test.
- $K(t) = \hat{S}(t)^p$ implementered in R

Ex. 3.9 Logrank Relapse Leukemia

```
> tpl<-c(1,1,2,2,3,4,4,5,5,8,8,8,8,11,11,12,12,15,17,22,2
> dpl<-rep(1,length(tpl))
> t6MP<-c(6,6,6,6,7,9,10,10,11,13,16,17,19,20,22,23,
          25,32,32,34,35)
> d6MP<-c(1,1,1,0,1,0,1,0,0,1,1,0,0,0,1,1,0,0,0,0,0)
> times<-c(tpl,t6MP)
> relapse<-c(dpl,d6MP)
> group<-c(rep(1,length(tpl)),rep(2,length(t6MP)))
>
> #Log-rank
> survdiff(Surv(times,relapse)~group)

         N Observed Expected (O-E)^2/E (O-E)^2/V
group=1 21       21     10.7      9.77      16.8
group=2 21        9     19.3      5.46      16.8

 Chisq= 16.8  on 1 degrees of freedom, p= 4.17e-05
```

Ex. 3.9 Relapse Leukemia - two other tests

```
> #Wilcoxon type,
> survdiff(Surv(times,relapse)~group,rho=1)

         N Observed Expected (O-E)^2/E (O-E)^2/V
group=1 21    14.55     7.68      6.16      14.5
group=2 21     5.12    12.00      3.94      14.5

 Chisq= 14.5  on 1 degrees of freedom, p= 0.000143
>
> #In between
> survdiff(Surv(times,relapse)~group,rho=0.5)

         N Observed Expected (O-E)^2/E (O-E)^2/V
group=1 21    17.22     8.98      7.57      15.7
group=2 21     6.68    14.92      4.56      15.7

 Chisq= 15.7  on 1 degrees of freedom, p= 7.4e-05
```

Ex. 3.10 Unbiased variance estimator

We will show that, see slide 2, $\mathrm{E}[V_{11}] = \mathrm{Var}(Z_1)$.

Since $N_\bullet(t) = N_1(t) + N_2(t)$ under H_0 has intensity process $Y_\bullet(t)\alpha(t) = Y_1(t)\alpha(t) + Y_2(t)\alpha(t)$, i.e. the sum of the intensity processes of the $N_j(t)$ we have

$$\begin{aligned} V_{11} &= \int_0^{t_0} \frac{L(s)^2}{Y_1(s)Y_2(s)} dN_\bullet(s) \\ &= \int_0^{t_0} \frac{L(s)^2}{Y_1(s)Y_2(s)} Y_\bullet(s)\alpha(s)ds + \int_0^{t_0} \frac{L(s)^2}{Y_1(s)Y_2(s)} dM_\bullet(s) \end{aligned}$$

where $M_\bullet(t) = N_\bullet(t) - \int_0^t Y_\bullet(s)\alpha(s)ds$ is a martingale, thus the last term is an integral wrt martingale and has exp. 0. Thus

$$\begin{aligned} \mathrm{E}[V_{11}] &= \mathrm{E}[\int_0^{t_0} \frac{L(s)^2}{Y_1(s)Y_2(s)} Y_\bullet(s)\alpha(s)ds] \\ &= \mathrm{E}[\int_0^{t_0} L(s)^2 (\frac{1}{Y_1(s)} + \frac{1}{Y_2(s)})\alpha(s)ds] = \mathrm{Var}(Z_1) \end{aligned}$$

Ex. 3.11

Want to show that under the null where $\alpha_1(t) = \alpha_2(t) = \alpha(t)$ we have $\mathrm{E}[N_1(t_0)] = \mathrm{E}[E_1(t_0)]$ where from eq. (3.60), ABG

$$E_1(t_0) = \int_0^{t_0} Y_1(s) \frac{dN_\bullet(s)}{Y_\bullet(s)} = \int_0^{t_0} Y_1(s)\alpha(s)ds + \int_0^{t_0} Y_1(s) \frac{dM_\bullet(s)}{Y_\bullet(s)}$$

where the last equality is added in here. The last term is a martingale with expectation 0 and we get

$$\mathrm{E}[E_1(t_0)] = \mathrm{E}[\int_0^{t_0} Y_1(s)\alpha(s)ds].$$

But we also have

$$\mathrm{E}[N_1(t_0)] = \mathrm{E}[\int_0^{t_0} Y_1(s)\alpha(s)ds + M_1(t_0)] = \mathrm{E}[\int_0^{t_0} Y_1(s)\alpha(s)ds],$$

thus equality (under the null).

About variance of Z_1

Under the null we may write, $M_j(t) = N_j(t) - \int Y_j(s)\alpha(s)ds$,

$$Z_1 = \int_0^{t_0} L(s)\frac{dM_1(s)}{Y_1(s)} - \int_0^{t_0} L(s)\frac{dM_2(s)}{Y_2(s)}$$

as the difference between two martingales and the variance result follows directly when they are uncorrelated.

In general $\mathrm{Cov}(\sum_i U_i, \sum_j V_j) = \sum_i \sum_j \mathrm{Cov}(U_i, V_j)$ and this will extend to integrals, thus

$$\mathrm{Cov}(\int_0^{t_0} L(s)\frac{dM_1(s)}{Y_1(s)}, \int_0^{t_0} L(s)\frac{dM_2(s)}{Y_2(s)})$$

$$= \int_0^{t_0}\int_0^{t_0} \mathrm{Cov}(L(s_1)\frac{dM_1(s_1)}{Y_1(s_1)}, L(s_2)\frac{dM_2(s_2)}{Y_2(s_2)})$$

About variance of Z_1, contd.

But the latter term simplifies to

$$\int_0^{t_0} \mathrm{Cov}(L(s)\frac{dM_1(s)}{Y_1(s)}, L(s)\frac{dM_2(s)}{Y_2(s)})$$

since for $s_1 \neq s_2$, say $s_1 < s_2$,

$$\mathrm{Cov}(L(s_1)\frac{dM_1(s_1)}{Y_1(s_1)}, L(s_2)\frac{dM_2(s_2)}{Y_2(s_2)}) =$$

$$\mathrm{E}[\mathrm{E}(L(s_1)\frac{dM_1(s_1)}{Y_1(s_1)}L(s_2)\frac{dM_2(s_2)}{Y_2(s_2)})|\mathcal{F}_{s_2-}]$$

$$\mathrm{E}[L(s_1)\frac{dM_1(s_1)}{Y_1(s_1)}\mathrm{E}(L(s_2)\frac{dM_2(s_2)}{Y_2(s_2)}|\mathcal{F}_{s_2-})] = 0$$

About variance of Z_1, III

Furthermore

$$\int_0^{t_0} \mathrm{Cov}(L(s)\frac{dM_1(s)}{Y_1(s)}, L(s)\frac{dM_2(s)}{Y_2(s)}) = 0$$

since

$$\mathrm{Cov}(L(s)\frac{dM_1(s)}{Y_1(s)}, L(s)\frac{dM_2(s)}{Y_2(s)}) = \mathrm{E}[\frac{L(s)^2}{Y_1(s)Y_2(s)}\mathrm{E}[dM_1(s)dM_2(s)|\mathcal{F}_{s-}]$$

which equal zero since $N_1(s) + N_2(s)$ has intensity process $(Y_1(s) + Y_2(s))\alpha(s)ds$, thus

$$\begin{aligned} \mathrm{Var}(dM_1(s) + dM_2(s)|\mathcal{F}_{s-}) &= (Y_1(s) + Y_2(s))\alpha(s)ds \\ &= \mathrm{Var}(dM_1(s)|\mathcal{F}_{s-}) + \mathrm{Var}(dM_2(s)|\mathcal{F}_{s-}) \end{aligned}$$

giving $\mathrm{E}[dM_1(s)dM_2(s)|\mathcal{F}_{s-}] = \mathrm{Cov}[dM_1(s), dM_2(s)|\mathcal{F}_{s-}] = 0$

About variance of Z_1, IV

A couple of the arguments on the previous page may need some more explanation: $N_j(t)$ has intensity process $Y_j(t)\alpha(t)$ and so for the martingale $M_j(t) = N_j(t) - \int_0^t Y_j(s)\alpha(s)ds$ we have

$$\text{Var}(dM_j(t)|\mathcal{F}_{t-}) = Y_j(t)\alpha(t)dt$$

Furthermore $N_1(t) + N_2(t)$ has intensity process $(Y_1(t) + Y_2(t))\alpha(t)$ and martingale

$$M_\bullet(t) = (N_1(t)+N_2(t)) - \int_0^t (Y_1(s)+Y_2(s))\alpha(s)ds = M_1(t)+M_2(t)$$

and so $\text{Var}(dM_\bullet(t)|\mathcal{F}_{t-}) = (Y_1(t) + Y_2(t))\alpha(t)dt$
$= \text{Var}(dM_2(t)|\mathcal{F}_{t-}) + \text{Var}(dM_2(t)|\mathcal{F}_{t-})$. Thus $M_1(t)$ and $M_2(t)$ have uncorrelated increments.