3.3.1 in ABG: Two-sample tests
For these slides, see also Chapter 4 of ASAUR

Consider two counting processes $N_1(t)$ and $N_2(t)$ with intensity processes of the multiplicative form

$$\lambda_h(t) = Y_h(t)\alpha_h(t); \quad h = 1, 2$$

We want to test the null hypothesis

$$H_0: \ \alpha_1(t) = \alpha_2(t) \ \text{ for } 0 \leq t \leq t_0$$

Usually we will choose $t_0 = \tau$, the upper time limit of study.

The common (but unknown) value of the $\alpha_h(t)$ under H_0 will be called $\alpha(t)$.

A general two-sample test based on the $\hat{A}_h(t)$

Recall the Nelson-Aalen estimators

$$\hat{A}_h(t) = \int_0^t \frac{1}{Y_h(u)} dN_h(u) = \sum_{T_j \leq t} \frac{1}{Y_h(T_j)}$$

and consider the test statistic

$$Z_1(t_0) = \int_0^{t_0} L(t)\{d\hat{A}_1(t) - d\hat{A}_2(t)\}$$

Here $L(t)$ is a non-negative predictable weight process that is zero whenever at least one of the $Y_h(t)$ are zero.

The choice

$$L(t) = Y_1(t)Y_2(t)/Y_\bullet(t)$$

with $Y_\bullet(t) = Y_1(t) + Y_2(t)$ gives the **log-rank test**, to be considered later.

Two-sample tests (cont.)

The standardized test statistic

$$U(t_0) = \frac{Z_1(t_0)}{\sqrt{V_{11}(t_0)}}$$

is approximately standard normal under H_0 (can be shown by martingale central limit theorem).

Alternatively we may use the test statistic

$$X^2(t_0) = \frac{Z_1(t_0)^2}{V_{11}(t_0)}$$

which is approximately chi-square distributed with 1 df under H_0

The log-rank test

For $K(t) = I\{Y_\bullet(t) > 0\}$ we get

$$\begin{aligned} Z_1(t_0) &\equiv \int_0^{t_0} K(t)dN_1(t) - \int_0^{t_0} K(t)\frac{Y_1(t)}{Y_\bullet(t)}dN_\bullet(t) \\ &= N_1(t_0) - \int_0^{t_0} \frac{Y_1(t)}{Y_\bullet(t)}dN_\bullet(t) \\ &= N_1(t_0) - E_1(t_0) \equiv O_1 - E_1 \\ &= \textit{observed - expected} \quad \text{in sample 1} \end{aligned}$$

Thus the standardized log-rank test statistic can be written

$$\frac{Z_1}{\sqrt{V_{11}}} = \frac{O_1 - E_1}{\sqrt{V_{11}}} \sim_{H_0} N(0,1) \;\text{ or }\; \left(\frac{Z_1}{\sqrt{V_{11}}}\right)^2 = \frac{(O_1 - E_1)^2}{V_{11}} \sim_{H_0} \chi_1^2$$

Hand-calculation of log-rank test

$$O_1 - E_1 = N_1(t_0) - \int_0^{t_0} \frac{Y_1(t)}{Y_\bullet(t)} dN_\bullet(t), \qquad V_{11} = \int_0^{t_0} \frac{Y_1(t)Y_2(t)}{Y_\bullet(t)^2} dN_\bullet(t)$$

Go through all *failure times* $T_1, \cdots, T_r$:

	Group 1	Group 2	Total at T_j
# at risk at T_j	Y_{1j}	Y_{2j}	Y_j
Observed # fail at T_j	O_{1j}	O_{2j}	O_j
Est prob of fail under H_0			$\frac{O_j}{Y_j}$
Estim expect # failures	$E_{1j} = Y_{1j} \cdot \frac{O_j}{Y_j}$	$E_{2j} = Y_{2j} \cdot \frac{O_j}{Y_j}$	
Estimated variance			$V_j = \frac{Y_{1j}Y_{2j}O_j}{Y_j^2}$

Then sum over all failure times $T_1, \cdots, T_r$:

$$O_h = \sum_{j=1}^r O_{hj}, \quad E_h = \sum_{j=1}^r E_{hj} \text{ for } h = 1, 2, \quad \text{and } V_{11} = \sum_{j=1}^r V_j$$

Test statistics are then

$$\frac{(O_1 - E_1)^2}{V_{11}} \text{ or the conservative } \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}$$

Example Log-rank: Kidney transplantation

```
eldre<-(age>49)
survdiff(Surv(time,delta)~eldre)
```

```
                N Observed Expected (O-E)^2/E (O-E)^2/V
eldre=FALSE 574       73    100.3      7.44      26.5
 eldre=TRUE 289       67     39.7     18.81      26.5

 Chisq= 26.5  on 1 degrees of freedom, p= 2.64e-07
```

Calculate also

$$\frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} = 7.44 + 18.81 = 26.25 (< 26.5)$$

k-sample tests

Consider now k counting processes $N_1(t), N_2(t), \ldots, N_k(t)$ with intensity processes of the multiplicative form

$$\lambda_h(t) = Y_h(t)\alpha_h(t); \quad h = 1, 2, \ldots, k$$

We want to test the null hypothesis

$$H_0 : \alpha_1(t) = \cdots = \alpha_k(t) \text{ for } 0 \leq t \leq t_0$$

We introduce (where δ_{hj} is a Kronecker delta)

$$\begin{aligned} Z_h(t_0) &= \int_0^{t_0} K(t)dN_h(t) - \int_0^{t_0} K(t)\frac{Y_h(t)}{Y_\bullet(t)}dN_\bullet(t) \\ V_{hj}(t_0) &= \int_0^{t_0} K^2(t)\frac{Y_h(t)}{Y_\bullet(t)}\left(\delta_{hj} - \frac{Y_j(t)}{Y_\bullet(t)}\right)dN_\bullet(t) \end{aligned}$$

k-sample tests (cont.)

Then the test statistic takes the form

$$X^2(t_0) = \mathbf{Z}(t_0)^T \mathbf{V}(t_0)^{-1} \mathbf{Z}(t_0)$$

The statistic is chi-square distributed with $k-1$ d.f. when the null hypothesis is true.

For the log-rank test one may show that

$$\sum_{h=1}^{k} \frac{(N_h(t_0) - E_h(t_0))^2}{E_h(t_0)} \leq X^2(t_0) \quad (*)$$

where $E_h(t_0) = \int_0^{t_0} \{Y_h(t)/Y_\bullet(t)\} dN_\bullet(t)$

Thus the left-hand side of $(*)$ provides a *conservative* version of the log-rank test (see also the case $k=2$).

Example Log-rank: Kidney transplantation

Ex: Kidney transpl.

```
> agegr<-trunc(age/20)
> table(agegr)
  0   1   2   3
 29 304 429 101
> survdiff(Surv(time,delta)~agegr)
Call:
survdiff(formula = Surv(time, death) ~ agegr)

          N Observed Expected (O-E)^2/E (O-E)^2/V
agegr=0  29        1     5.65      3.82      3.99
agegr=1 304       21    56.76     22.53     38.17
agegr=2 429       88    65.45      7.77     14.63
agegr=3 101       30    12.15     26.24     28.97

 Chisq= 61.2  on 3 degrees of freedom, p= 3.26e-13
```

Stratified tests (cont.)

For each stratum s we define similar quantities as above:

$$\begin{aligned} Z_{hs}(t_0) &= \int_0^{t_0} K_s(t) dN_{hs}(t) - \int_0^{t_0} K_s(t) \frac{Y_{hs}(t)}{Y_\bullet(t)} dN_{\bullet s}(t) \\ V_{hjs}(t_0) &= \int_0^{t_0} K_s^2(t) \frac{Y_{hs}(t)}{Y_\bullet(t)} \left(\delta_{hj} - \frac{Y_{js}(t)}{Y_\bullet(t)} \right) dN_{\bullet s}(t) \end{aligned}$$

Further we define the $k-1$ dimensional vectors

$$\mathbf{Z}_s(t_0) = (Z_{1s}(t_0), \ldots, Z_{k-1,s}(t_0))^T$$

and the $(k-1) \times (k-1)$ dimensional matrices

$$\mathbf{V}_s(t_0) = \{V_{hjs}(t_0)\}_{h,j=1,\ldots,k-1}$$