UCIrvine
University of California, Irvine

Lecture 4

Extensions of the Logrank Test

Statistics 255 - Survival Analysis

Presented January 21, 2016

Dan Gillen
Department of Statistics
University of California, Irvine

Weighted Logrank Tests

Logrank and Mantel-Haenzel Test

M-H Test Series of (independent) tables at different levels of a confounder C

- Data at level $C = k$:

	D	$\overline{\text{D}}$
E	a_k	b_k
$\overline{\text{E}}$	c_k	d_k

- M-H test compares $\Pr[\text{D}|\text{E}, C = k]$ and $\Pr[\text{D}|\overline{\text{E}}, C = k]$ and is designed (most powerful) for the case where the odds ratio, ψ_k is constant at all levels of C:

$$\psi_k = \frac{\Pr[\text{D}|\text{E}, C = k]/\Pr[\overline{\text{D}}|\text{E}, C = k]}{\Pr[\text{D}|\overline{\text{E}}, C = k]/\Pr[\overline{\text{D}}|\overline{\text{E}}, C = k]}$$

Weighted Logrank Tests

Logrank and Mantel-Haenzel Test

Logrank Test Series of (dependent) tables at different failure times

- Data at time t_k:

	D	$\overline{\text{D}}$	Risk Size
E	d_{0k}	$y_{0k} - d_{0k}$	y_{0k}
$\overline{\text{E}}$	d_{1k}	$y_{1k} - d_{1k}$	y_{1k}
Total	d_k	$y_k - d_k$	y_k

- We expect the logrank test to be most powerful when the "odds ratio" over infinitesimal time intervals are constant across time, ie $\psi_t = \psi$ for all t where

$$\psi_t = \frac{\text{Pr}[t \leq T < t + \Delta t | E, T \geq t] / \{1 - \text{Pr}[t \leq T < t + \Delta t | E, T \geq t]\}}{\text{Pr}[t \leq T < t + \Delta t | \overline{\text{E}}, T \geq t] / \{1 - \text{Pr}[t \leq T < t + \Delta t | \overline{\text{E}}, T \geq t]\}}$$

Proportional Hazards

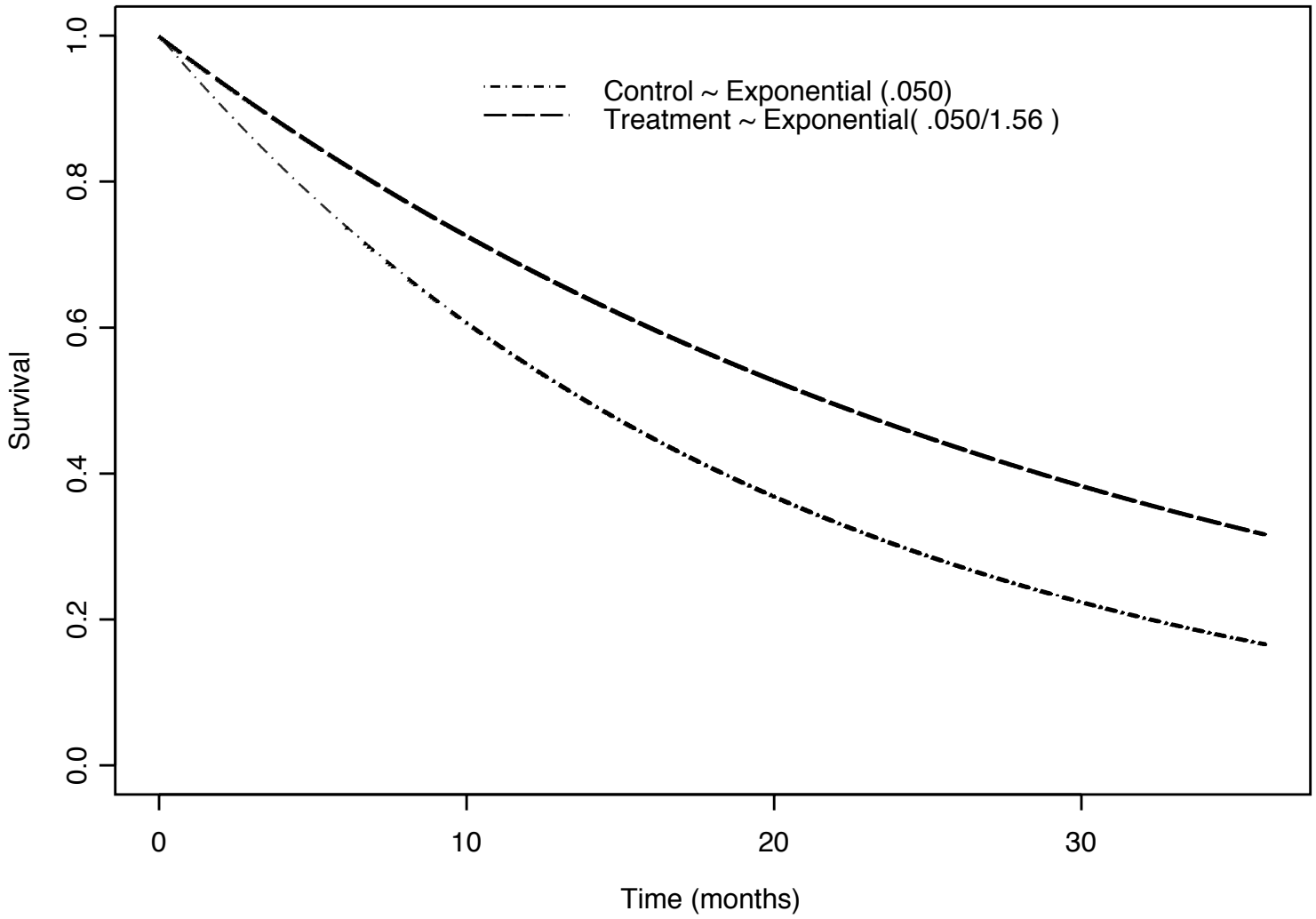
- But, as $\Delta t \downarrow 0$
  - $1 -$ Pr's $\uparrow 1$
  - Ratio of Pr's $\rightarrow$ ratio of hazards, ie

$$\psi_t \approx \frac{\lambda(t|\mathrm{E})}{\lambda(t|\overline{\mathrm{E}})}$$

- The logrank test will be most powerful for the case where the hazard ratio remains constant over time. This is called the **proportional hazards** case.

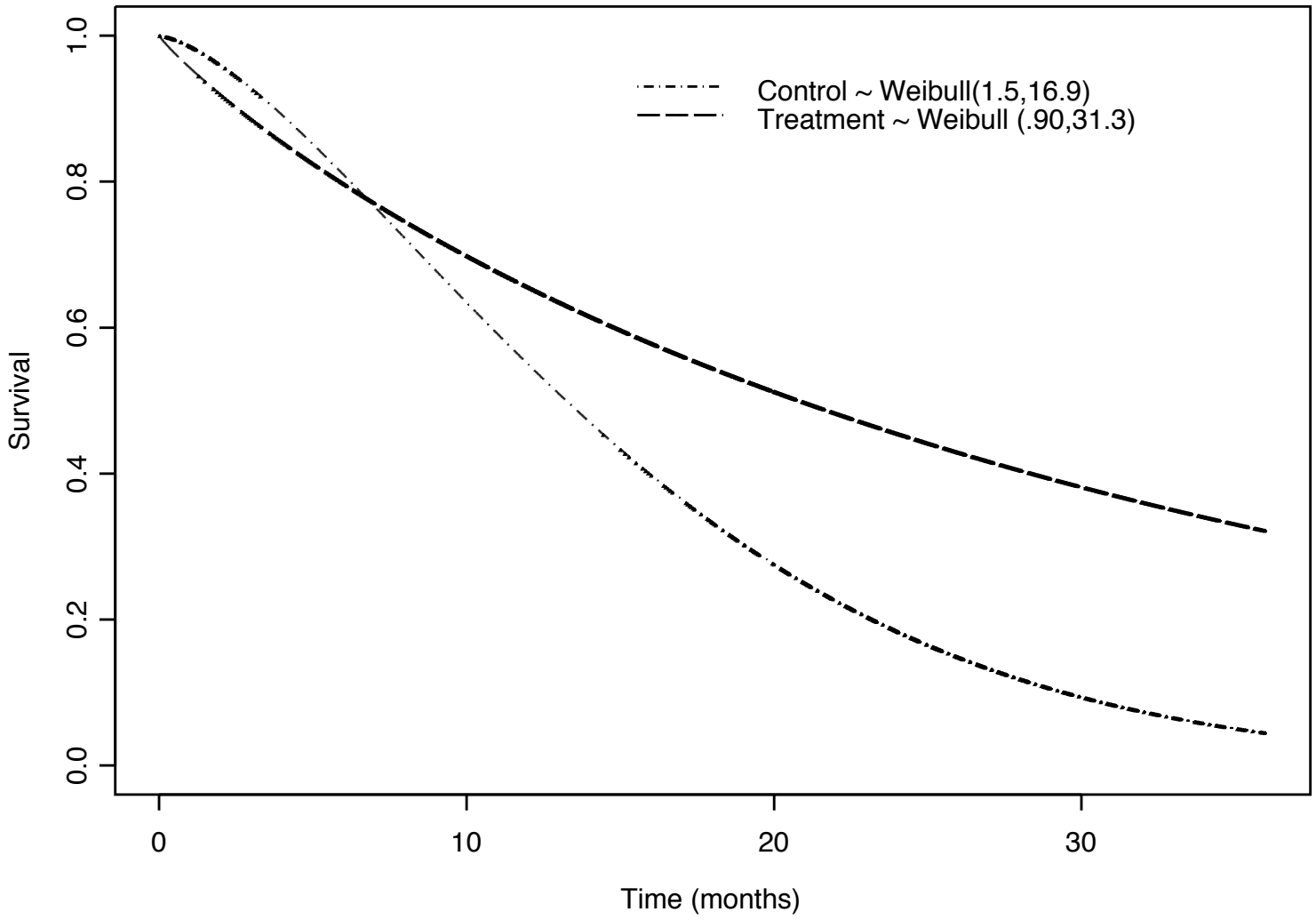
Weighted Logrank Tests

Ex. Proportional Hazards



Weighted Logrank Tests

Ex. Non-Proportional Hazards



Weighted Logrank Tests

Weighted Logrank Statistics

- Consider weighting $(Obs - Exp)$ differently over time
- This will enable us to inflate early or late differences

  $\rightarrow$ Potential for increased power under non-proportional hazards

$$T_W = \frac{\left[\sum_{k=1}^{D} w_k(O_k - E_k)\right]^2}{\sum_{k=1}^{D} w_k^2 V_k} = \frac{\left[\sum_{k=1}^{D} w_k U_k\right]^2}{\sum_{k=1}^{D} w_k^2 V_k}$$

Weighted Logrank Tests

Weighted Logrank Statistics

- Choices for w_k that have been proposed:

  1. $w_k = n_k$ gives the **Gehan-Breslow** test (weights equal to the total number of subjects at risk at each failure time). Applies greater weight to early failure times.

  2. $w_k = \hat{S}_{KM}(t_k-)$ gives the **generalized Wilcoxon** test (weights equal to the pooled estimate of survival just prior to time t_k). Applies greater weight to early failure times.

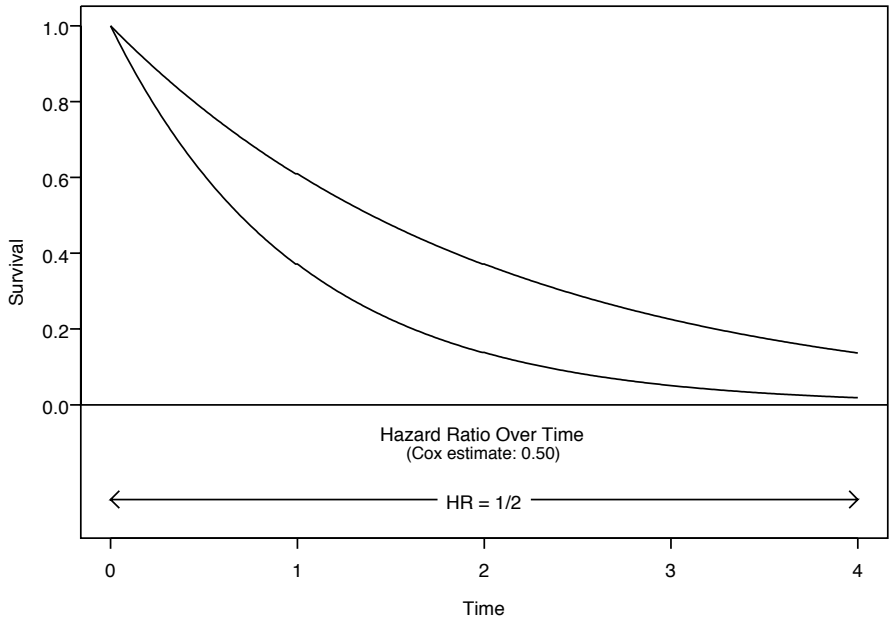
     - Equivalent to the Wilcoxon rank sum statistic when there is no censoring.

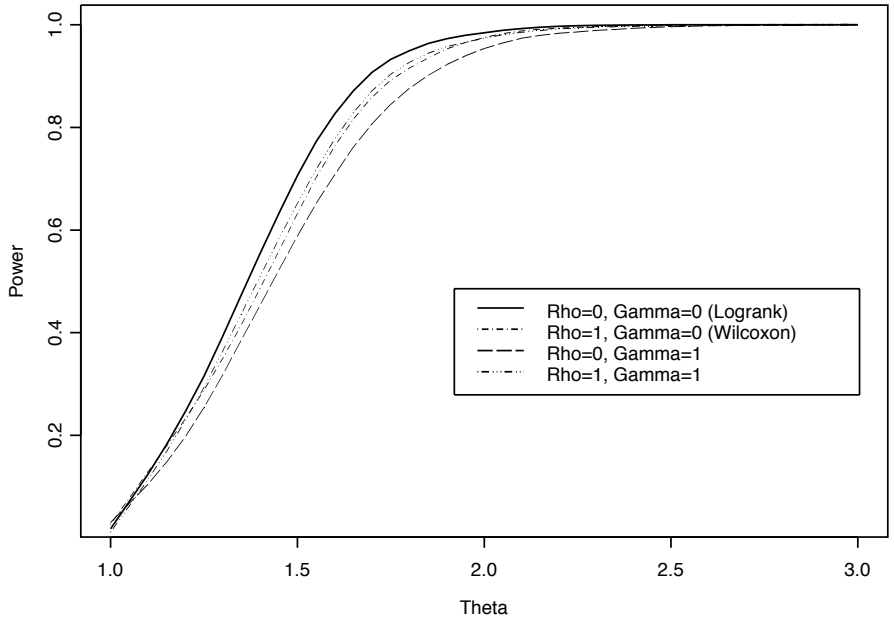
- The $G^{\rho,\gamma}$ family (Fleming and Harrington; 1991)

  - $w_k = \left[\hat{S}_{KM}(t_k-)\right]^{\rho}\left[1-\hat{S}_{KM}(t_k-)\right]^{\gamma}$
  - $\rho = \gamma = 0$ gives the usual logrank statistic
  - $\rho = 1$ and $\gamma = 0$ gives the generalized Wilcoxon test

Weighted Logrank Tests

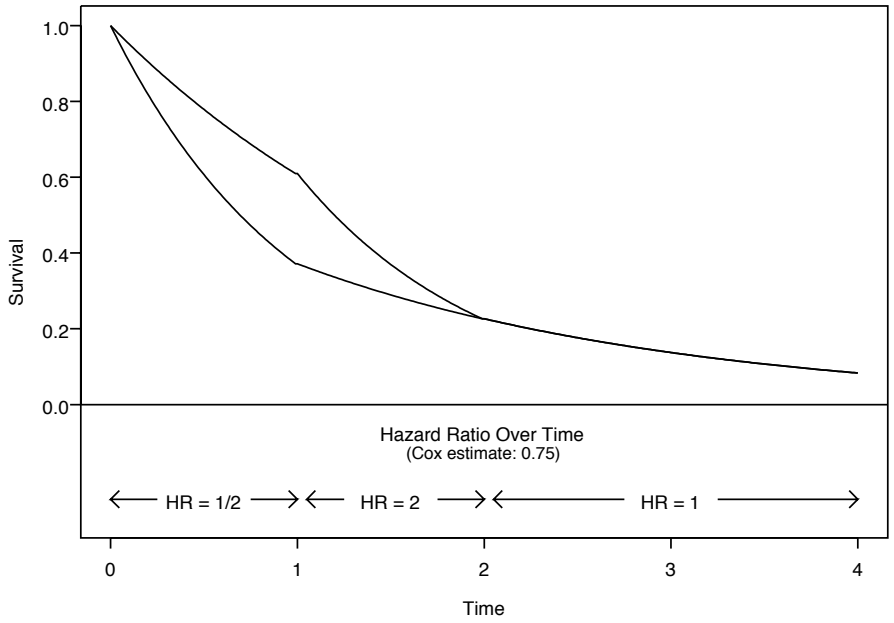
Power Comparisons - Proportional Hazards

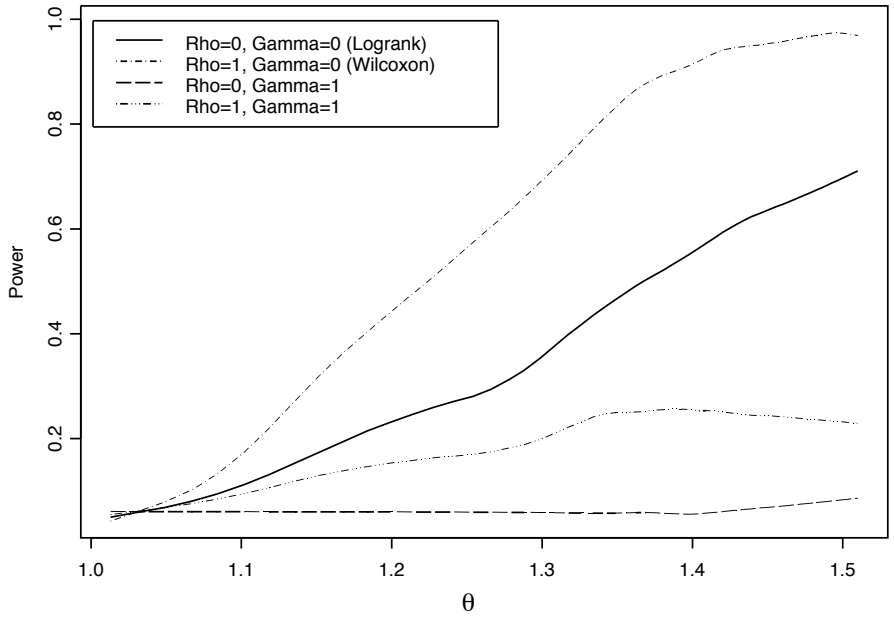




Weighted Logrank Tests

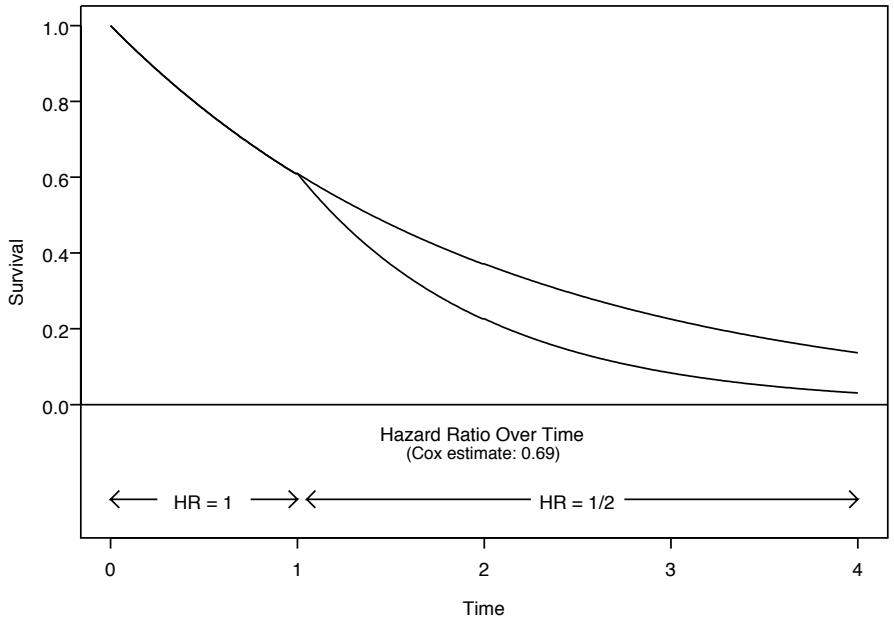
Power Comparisons - Early Diverging Hazards

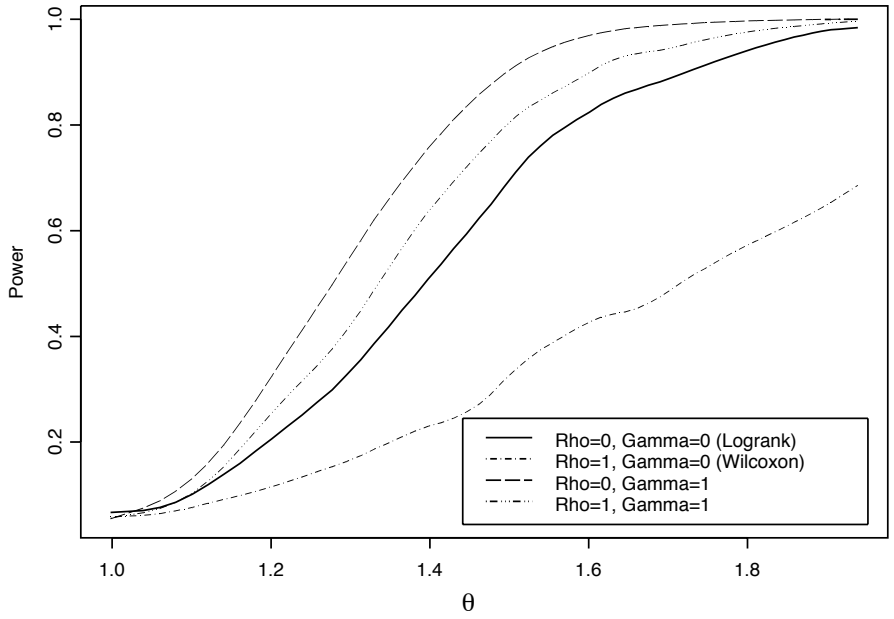




Weighted Logrank Tests

Power Comparisons - Late Diverging Hazards





Implementation in R - 6MP Example

- Know that the (unweighted) logrank statistic will be most powerful under proportional hazards
- How can we (informally) check the proportional hazards assumption?
  - If we have proportional hazards, then

$$\lambda_1(t) = \phi\lambda_0(t)$$

    so that

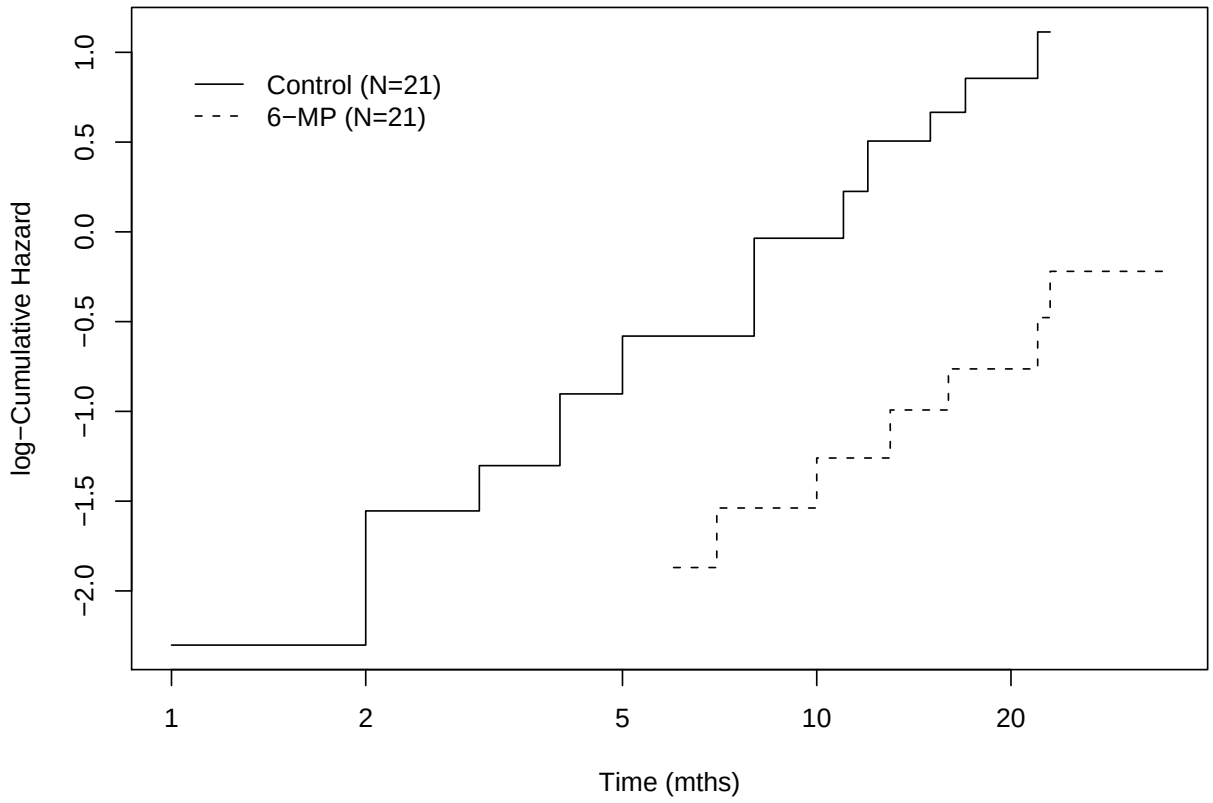
$$\log \Lambda_1(t) = \log(\phi) + \log \Lambda_0(t)$$

- So, if the log cumulative hazards are roughly parallel, the logrank test will tend to be most powerful

Weighted Logrank Tests

6MP log-Cumulative Hazards Plot

```
plot( survfit( Surv( time, irelapse ) ~ sixmp, data=sixmpLong ),
        fun="cloglog", lty=1:2, mark.time=FALSE,
        xlab="Time (mths)", ylab="log-Cumulative Hazard" )
legend( 1,1, lty=1:2, legend=c("Control (N=21)", "6-MP (N=21)"),
        bty="n" )
```



Weighted Logrank Tests

Implementation in R - 6MP Example

- Not too bad...We do not expect the generalized Wilcoxon test to be as powerful as the logrank test in this situation
- To compute the generalized Wilcoxon, specify the option `rho=1` in the `survdiff()` function
  - Note: The $G^{\rho,\gamma}$ statistic is not currently implemented in the `survival` package

```
> #####        Usual (unweight LR test)
> survdiff( Surv( time, irelapse ) ~ sixmp, rho=0, data=sixmpLong )
         N Observed Expected (O-E)^2/E (O-E)^2/V
sixmp=0 21       21     10.7      9.77      16.8
sixmp=1 21        9     19.3      5.46      16.8

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


> #####        Generalized Wilcoxon test
> survdiff( Surv( time, irelapse ) ~ sixmp, rho=1, data=sixmpLong )
         N Observed Expected (O-E)^2/E (O-E)^2/V
sixmp=0 21    14.55     7.68      6.16      14.5
sixmp=1 21     5.12    12.00      3.94      14.5

 Chisq= 14.5  on 1 degrees of freedom, p= 0.000143
```

Weighted Logrank Tests

How should weights be chosen?

- For scientific inference it is not reasonable to look at the survival curves first, then choose weights
- First, ask whether there is a reason to believe we will have non-proportional hazards
  - If not, go with the logrank test
  - If so, consider what survival differences are most meaningful (early vs late)

    $\rightarrow$ Childhood cancer (late differences)
    $\rightarrow$ Late stage lung cancer remission (early differences)

K-Sample Logrank Tests

- Suppose we have $K > 2$ groups and we wish to simultaneously compare them with respect to survival time distributions (or equivalently, hazards)

$$H_0: \quad \lambda_1(t) = \lambda_2(t) = \ldots \lambda_K(t), \text{ for all } t > 0$$

(i.e. the survival curves for the all groups are equal everywhere)

- We are particularly concerned with the alternatives

$$H_A: \quad \lambda_k(t) > \lambda_{k'}(t), \text{ for some } t > 0$$
$$\text{or}$$
$$\lambda_k(t) < \lambda_{k'}(t), \text{ for some } t > 0$$
$$\text{for at least some } k \neq k'$$

K-Sample Logrank Tests

K-Sample Logrank Tests

- Test statistic is a generalization of the two sample statistic that depends on the covariance between the $(O - E)$'s between each group
- Consider the data at the i^{th} observed event time t_i in the pooled sample:

1	2		k		K	Total
d_{1i} $y_{1i} - d_{1i}$	d_{2i} $y_{2i} - d_{2i}$	$\ldots$ $\ldots$	d_{ki} $y_{ki} - d_{ki}$	$\ldots$ $\ldots$	d_{Ki} $y_{Ki} - d_{Ki}$	d_i $y_i - d_i$

K-Sample Logrank Tests

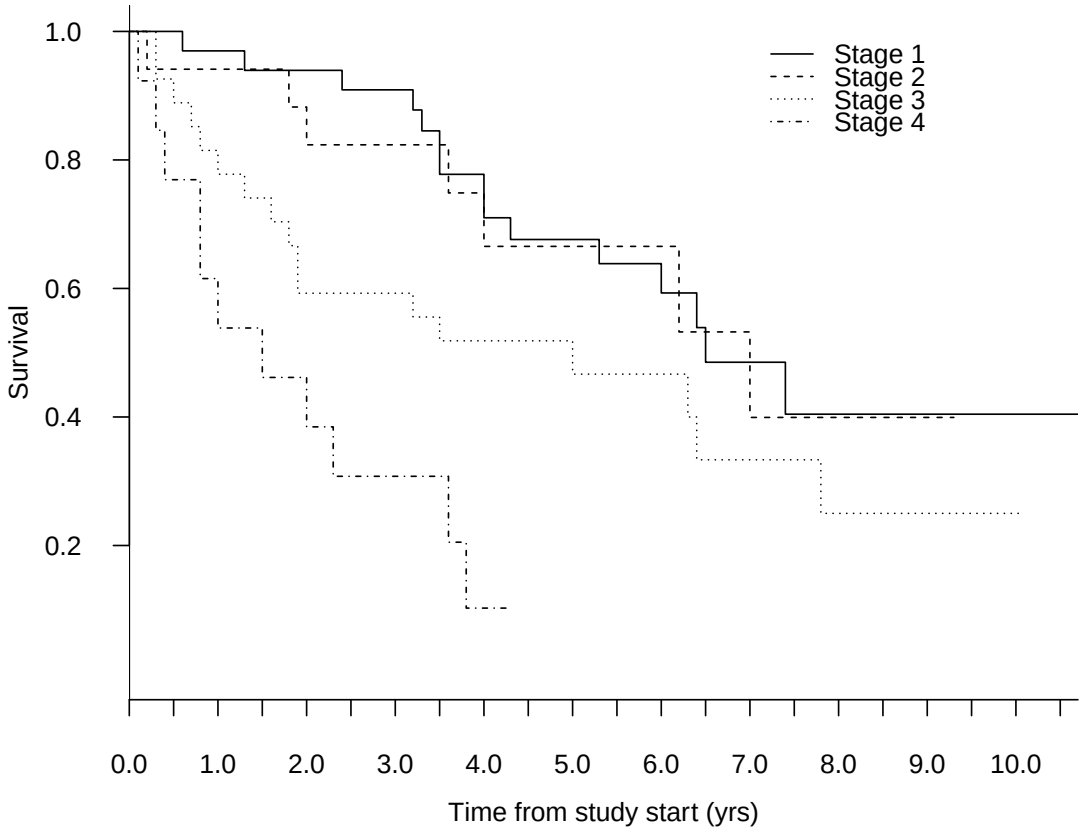
Ex: Survival in patients with cancer of the larynx (Sect 1.8 in K&M)

- Time origin: diagnosis with cancer
- Failure event: death
- Question of interest: How does survival time from diagnosis to death vary by *stage of disease* at presentation?

```
> larynx[1:10,]
   stage t2death age year death
1      1     0.6  77   76     1
2      1     1.3  53   71     1
3      1     2.4  45   71     1
4      1     2.5  57   78     0
5      1     3.2  58   74     1
6      1     3.2  51   77     0
7      1     3.3  76   74     1
8      1     3.3  63   77     0
9      1     3.5  43   71     1
10     1     3.5  60   73     1
```

K-Sample Logrank Tests

Ex: Survival in patients with cancer of the larynx (Sect 1.8 in K&M)



Stage 1	33 (0)	23 (7)	6 (14)	1 (15)
Stage 2	17 (0)	11 (3)	3 (7)	0 (7)
Stage 3	27 (0)	14 (13)	4 (16)	0 (17)
Stage 4	13 (0)	3 (9)	0 (11)	0 (11)
Total	90 (0)	51 (32)	13 (48)	1 (50)

K-Sample Logrank Tests

Ex: Survival in patients with cancer of the larynx (Sect 1.8 in K&M)

- `survfit()` can be used to test differences in K-samples as before

```
> survdiff( Surv(t2death,death) ~ stage, data=larynx )
Call:
survdiff(formula = Surv(t2death, death) ~ stage, data = larynx)

         N Observed Expected (O-E)^2/E (O-E)^2/V
stage=1 33       15    22.57     2.537     4.741
stage=2 17        7    10.01     0.906     1.152
stage=3 27       17    14.08     0.603     0.856
stage=4 13       11     3.34    17.590    19.827

 Chisq= 22.8  on 3 degrees of freedom, p= 4.53e-05
```

Conclusion: The hypothesis that all four survival curves are equal is clearly rejected. We conclude that at least one group is different with respect to survival

K-Sample Logrank Tests

Ex: Survival in patients with cancer of the larynx (Sect 1.8 in K&M)

- Additional Notes:
- The G^ρ family of weighted logrank statistics can be extended to K-samples by specifying the `rho` value in `survfit()`
- These tests say nothing about *how* the groups differ; which one is worst, best, etc. (though the sum of ranks gives a clue). That can be further explored with a trend test or regression modeling...
- For now, we could also think about testing for *trend* since stage is *ordinal*

K-Sample (Tarone) Test for Trend

Larynx Cancer Example

- Recall that there were 4 stages of disease recorded at baseline (the origin)
- The 4 stage of disease groups can be ordered in a meaningful way
- Suppose we wish to examine the hypothesis that the survival experience by stage of disease is **either** *progressively worse* **or** *progressively better* by stage of disease
- That is, we wish to take advantage of the *ordinal* nature of the stage of disease variable `stagedx`

K-Sample (Tarone) Test for Trend

Larynx Cancer Example

- Formally, for K ordered groups with dose vector $s_1, ..., s_K$ (could be $s_1 = 1, ..., s_K = K$), we want to test the hypothesis

$$H_0 : \quad \lambda_1(t) = \lambda_2(t) = \ldots = \lambda_K(t), \text{ for all } t > 0$$

(*Note: Same H_0 as the general K-sample problem!)

vs.

$$H_A : \quad \phi^{s_1}\lambda_1(t) = \phi^{s_2}\lambda_2(t) = \ldots = \phi^{s_K}\lambda_K(t),$$
$$\text{for } \phi \neq 1 \text{ and for all } t > 0$$

- What is the general form of this alternative in terms of survival curves?

K-Sample (Tarone) Test for Trend

Formulation of the test

- Recall the log-rank test is:

$$X^2 = \frac{U^2}{V} \dot{\sim} \chi^2_1$$

where

$$U \equiv U_1 \equiv \sum_j (\text{obs}_{(j)1} - \text{exp}_{(j)1}) = \sum_j U_{(j)1}$$

is the "observed" – "expected" discrepancy for group 1, and

$$V \equiv V_1 \equiv \text{Var}[U_1] = \sum_j V_{(j)1}$$

is the variance of U_1 under H_0

K-Sample (Tarone) Test for Trend

Formulation of the test

- Suppose group 1 (e.g. treatment) being compared to group 0 (e.g. placebo). Could similarly define

$$U_0 = \sum_j (\text{obs}_{(j)0} - \text{exp}_{(j)0}) = \sum_j U_{(j)0}$$

and write

$$U = 1 \times U_1 + 0 \times U_0 = U_1$$

assigning "scores" 1 and 0 to the two groups

K-Sample (Tarone) Test for Trend

Formulation of the test

- For K-samples, assign scores $s_1, s_2, \ldots, s_K$ and compute

$$U_T = s_1 \times U_1 + s_2 \times U_2 + \cdots + s_K \times U_K = \sum_{k=1}^{K} s_k (O_k - E_k)$$

- Then U_T will be *large* (positive or negative) if the U_ks increase (or decrease) with s_k

- V_T is computed using the variance-covariance matrix of $(U_1, U_2, \ldots, U_K)$:

$$V_T = \sum_{k=1}^{K} s_k^2 V_{kk} + 2 \sum_{k<k'} s_k s_{k'} V_{kk'}$$

K-Sample (Tarone) Test for Trend

Formulation of the test

- So our test statistic is

$$X_T^2 = \frac{U_T^2}{V_T} \dot{\sim} \chi_1^2$$

- Why 1-degree of freedom?

  $\rightarrow$ Test is really a regression of $\log \lambda_k(t)$ on s_k

K-Sample (Tarone) Test for Trend

Larynx Cancer Example

- There is no dedicated function for the trend test in R, but I have written the function `survtrend()` and posted it on the course webpage for this purpose

```
> survtrend( Surv(t2death,death) ~ stage, data=larynx )
         N Observed Expected
stage=1 33       15  22.5660
stage=2 17        7  10.0117
stage=3 27       17  14.0845
stage=4 13       11   3.3377

Logrank Test : Chi(3) = 22.763, p-value = 4.5252e-05
Tarone Test Trend : Chi(1) = 13.815, p-value = 0.00020169
```

Conclusion: Reject the hypothesis that all four survival curves are equal and conclude that stage is positively associated with the hazard for death

K-Sample (Tarone) Test for Trend

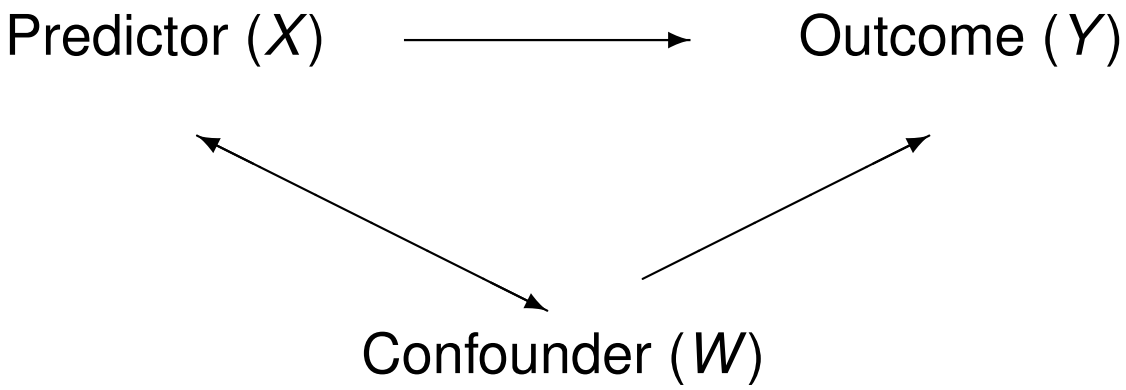
Comments

- The trend test depends on the order of the covariate being tested while the general K-sample test does not
- Why use a trend test (on 1 df) vs. a general K-sample test (on $K - 1$ df)?
  - If effects are monotonically ordered it will be more *sensitive*
    - The general K-sample test has less power because it does not take advantage of the ordinal nature of the data
  - It seeks to detect a more *specific* alternative
    - If survival curves differ, but differences are not ordered, trend test less likely to reject
- The trend test is essentially a regression of the hazard on the covariate of interest

Stratified Logrank Tests

Confounding

- One definition: A **confounder** is a variable that is associated with the predictor of interest (X) *and* causally related to the outcome of interest (Y).



Stratified Logrank Tests

Confounding

- Example: Weight may be a confounder in the relationship between diabetes and blood pressure:
  - Diabetics tend to be heavier than non-diabetics
  - Increased weight is associated with higher blood pressure

- Note: When considering potential confounders, need to carefully consider whether or not the potential confounder lies in the *causal pathway* of the association of interest

$$\text{Exposure} \longrightarrow Z \longrightarrow \text{Outcome}$$

Stratified Logrank Tests

Confounding

- How do we deal with confounding? Adjust for the confounder
- Adjustment involves the assumption that the effect of interest is *similar across all strata* of the potential confounder
- What if we want to test for differences in risk (ie survival data) after adjustment for a potential confounding factor?

  $\rightarrow$ One solution is to **stratify** the sample

Stratified Logrank Tests

Set-up and Notation

- Suppose the variable we wish to stratify on has J levels
- Consider testing the hypothesis

$$H_0: \quad \lambda_{j1}(t) = \lambda_{j0}(t), \text{ for } j = 1, ..., J \text{ and } t > 0$$
$$H_A: \quad \lambda_{j1}(t) = \phi\lambda_{j0}(t), \text{ for } j = 1, ..., J \text{ and } t > 0, \phi \neq 1$$

- Notes:

1. $\lambda_{j1}(t)$ can differ from $\lambda_{j'1}(t)$ for two strata j and j', as can $\lambda_{j0}(t)$ and $\lambda_{j'0}(t)$
2. Testing whether, on average across *strata* $j = 1, \ldots, J$ and across *time* t, the *within-stratum hazard* $\lambda_{j1}(t)$ greater (or less) than $\lambda_{j0}(t)$?
3. Testing for *similar (proportional hazards) effects* across time and strata

Stratified Logrank Tests

Set-up and Notation

- Suppose that, *for the jth stratum*

$n_{i(j)}$ = the number at risk at time $t_{i(j)}$

$d_{i(j)}$ = the number failing at time $t_{i(j)}$

- Define

$n_{i(j)1}$ = the number at risk *in group 1* and stratum j at time $t_{i(j)}$

$d_{i(j)1}$ = the number failing *in group 1* and stratum j at time $t_{i(j)}$

Stratified Logrank Tests

Set-up and Notation

- Recall: the **log-rank test** for the jth stratum only would compare "observed" to "expected":

$$\begin{aligned} U_j &= \sum_i (\text{obs}_{i(j)} - \text{exp}_{i(j)}) = \sum_i U_{i(j)} \\ &= \sum_i \left\{ d_{i(j)1} - n_{i(j)1} \left(\frac{d_{i(j)}}{n_{i(j)}} \right) \right\} \end{aligned}$$

using the variance

$$V_j = \text{Var}[U_j] = \sum_i v_{i(j)}$$

- If U_j is large (positive or negative), then the test will reject within the jth stratum

Stratified Logrank Tests

Set-up and Notation

- The **stratified log-rank test** sums (averages) over strata just as the log-rank test sums (averages) over times:

$$U_S = \sum_j U_j = \sum_j \sum_i U_{i(j)}$$

and

$$V_S = \text{Var}[U_S] = \sum_j V_j = \sum_j \sum_i v_{i(j)}$$

- Under H_0

$$X_S^2 = \frac{U_S^2}{V_S} \dot{\sim} \chi_1^2$$

- The **stratified log-rank test** statistic U_S is a weighted average of the within-stratum log-rank test statistics U_i

Stratified Logrank Tests

Back to the larynx cancer example...

- Let's consider adjustment for age as a potential confounding factor in the relationship between stage of disease and time to death
- Specifically, consider age discretized into 3 groups

```
> ##
> ##### Consider potential confounding by age
> ##
> summary(larynx$age)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   41.0    57.0    65.0    64.6    72.0    86.0
> round( quantile( larynx$age, c(0,.33,.66,1) ) )
  0%  33%  66% 100%
  41   60   70   86
> larynx$agegrp <- cut( larynx$age, c(41,60,70,86),
                       include.lowest=TRUE )
> summary( larynx$agegrp )
[41,60] (60,70] (70,86]
     30      32      28
```

Stratified Logrank Tests

Back to the larynx cancer example...

- Let's consider first consider whether or not age is likely to meet the definition of a confounder...

```
> ##
> ##### Does age meet the definition of a confounder? (not really...)
> ##
> chisq.test( table( larynx$stage, larynx$agegrp ) )

        Pearson's Chi-squared test

data:  table(larynx$stage, larynx$agegrp)
X-squared = 4.7134, df = 6, p-value = 0.5811



> survdiff( Surv(t2death,death) ~ agegrp, data=larynx )
Call:
survdiff(formula = Surv(t2death, death) ~ agegrp, data = larynx)

                N Observed Expected (O-E)^2/E (O-E)^2/V
agegrp=[41,60] 30       14     15.9     0.221     0.330
agegrp=(60,70] 32       16     20.8     1.103     1.938
agegrp=(70,86] 28       20     13.3     3.325     4.615

 Chisq= 4.7  on 2 degrees of freedom, p= 0.0937
```

Stratified Logrank Tests

Back to the larynx cancer example...

- From the above, age is not significantly associated with stage or with time to death (in the dataset)
- The implication of this is that adjustment for age is unlikely to have any impact on the conclusions of our analysis (we will lose some efficiency though...)
  - In a real setting, when testing a well-defined hypothesis we should decide upon adjustment for age before assessing the data in order to avoid data-driven inflation of the type 1 error rate!

Stratified Logrank Tests

Back to the larynx cancer example...

- Let's stratify by age here as an example...To do this, use the `strata()` function in the formula statement of `survdiff()`

```
> ##
> ##### LR test of association between stage and t2death,
> ##### stratified by agegrp
> ##
> survdiff(Surv(t2death,death) ~ stage + strata(agegrp), data=larynx)
Call:
survdiff(formula = Surv(t2death, death) ~ stage + strata(agegrp),
    data = larynx)

         N Observed Expected (O-E)^2/E (O-E)^2/V
stage=1 33       15    23.60     3.134     6.430
stage=2 17        7     9.38     0.602     0.763
stage=3 27       17    13.23     1.074     1.547
stage=4 13       11     3.79    13.686    16.182

 Chisq= 20.1  on 3 degrees of freedom, p= 0.00016
```

Stratified Logrank Tests

Conclusions

- Stage of disease at diagnosis is positively related to the risk of death
- This relationship still holds after adjusting for the effect of age
- The association is not due to any (positive or negative) association of age with stage of disease and / or age with risk of death
- The association is not due to any **confounding** effect of age

Matched Tests

Matching

- When explicit control for confounders will be difficult, comparative studies are sometime performed on samples of *matched pairs*:
  - one member of pair is exposed or treated and the other is not (or gets placebo)
  - matching on: age $\times$ sex, neighborhood, clinic, etc.
  - twins
  - matched pairs like many strata of size 2

Matched Tests

Matching

- To account for matching in the sampling scheme, we can:
  1. stratify on the matching set,
  2. compare outcomes within that strata, then
  3. combine the results across (independent) strata
- As an example, consider the 6-MP data where subjects were actually matched by remission status and hospital
  - One member randomized to 6-MP (vs. placebo) maintenance therapy

Matched Tests

Ex: 6-MP data

- A proper analysis should account for the correlation induced by matching...

```
> ##
> #####        Matched analysis of the 6-MP data
> ##
> sixmp <- read.table( "http://www.ics.uci.edu/~dgillen/
                        STAT255/Data/sixmp.txt" )
> sixmp[1:5,]
  pairid tpbo t6mp irelapse
1      1    1   10        1
2      2   22    7        1
3      3    3   32        0
4      4   12   23        1
5      5    8   22        1


> ##
> #####        Transform data to long format
> ##
> sixmpLong <- cbind( rep(sixmp$pairid, 2), c(sixmp$tpbo, sixmp$t6mp),
+                  rep(0:1, each=21), c( rep(1,21), sixmp$irelapse ) )
> sixmpLong <- as.data.frame( sixmpLong )
> names( sixmpLong ) <- c( "pairid", "time", "sixmp", "irelapse" )
```

Matched Tests

Ex: 6-MP data

```
> ##
> #####         Matched analysis stratifying by pairid
> ##
> survdiff(Surv(time, irelapse)~sixmp+strata(pairid), data=sixmpLong)
Call:
survdiff(formula = Surv(time, irelapse) ~ sixmp + strata(pairid),
    data = sixmpLong)

         N Observed Expected (O-E)^2/E (O-E)^2/V
sixmp=0 21       21     13.5      4.17      10.7
sixmp=1 21        9     16.5      3.41      10.7

 Chisq= 10.7  on 1 degrees of freedom, p= 0.00106
```

Conclusion: After accounting for correlation induced by matching on remission status and hospital, we conclude that there is a difference in time-to-relapse between the 6-MP arm and the control arm using a level .05 test

Stratified Analyses

Summary

1. Analyses by separate strata, stratified tests and adjustment are statistical activities, but . . .
   - . . . the identification of confounders and the decision to adjust for them are extra-statistical considerations
   - . . . they involve (1) the scientific question of interest and (2) possible chains of causality
2. If the *study design* is stratified or matched, always adjust
3. Stratified tests will have good power for alternatives that are in the same direction in each stratum
4. When effects are different by stratum (*interaction or effect modification*), analyses are better performed and reported separately on each stratum

Stratified Analyses

Summary

5. Weighted, K-sample and K-sample trend test versions of the stratified log-rank test exist

6. Strata can be quite small for adjustment (but not for within-stratum analyses)

7. When data are in the form of matched pairs (or small sets), think of them as many small strata

8. Stratified log-rank test on matched pairs is the censored data analogue of the the signed-rank test for paired data

Stratified Analyses

Analogous Methods For Binomial Data

	Proportions	Survival Data
1. Description	$\hat{p}, \hat{RR}$ $\hat{OR}$	$\hat{S}, \hat{\Lambda}$, $\hat{\lambda}, \hat{RR}$
2. Two-sample test	Z test/χ^2 test	Logrank test
3. Stratified two -sample test	Mantel-Haenzel test	Stratified logrank test
4. K-sample heterogeneity test	K-sample heterogeneity test	K-sample logrank test
5. K-sample trend test	Cochran-Armitage trend test	Tarone trend test
6. Regression models	Logistic regression	Cox regression