Syllabus

1. Introduction
   - Survival data
   - Censoring mechanism
   - Application in medical field
2. Concepts and definitions
   - Survival function
   - Hazard function
3. Non-parametric approach
   - Life table
   - Kaplan-Meier survival estimate
   - Hazard function
   - Median and percentile survival time
4. Hypothesis testing
   - Overview – hypothesis, test statistics, p-values
   - Log-rank
   - Wilcoxon
   - Gehan test
5. Study design and sample size estimation
   - Overview
   - Survival sample size estimation
   - Accrual time and Study duration
6. Semiparametric model – proportional hazard model
   - Partial likelihood
   - Inference
   - Time varying covariates
   - Stratification
7. Model checking in the PH model
   - Model checking
   - Residuals
8. Parametric model
   - Parametric proportional hazard model
   - Accelerate failure model
9. Other topics
   - Competing risk
   - **Recurrent events**
   - Non-proportional hazard ratio
   - Interval censoring

Multivariate Survival Data

- Different types of events occur in one subject
  - Time to vision loss in each eye
  - Time to stroke and MI
- Recurrent events
  - Opportunistic infections in patients with HIV
  - Repeated exacerbations in COPD patients
- Recurrent events with a terminating event
  - Death can be a terminating event
  - May have a heavier weight than a recurrent event in evaluating treatment effect
  - May informatively censor the future recurrent event

Topics

- Marginal Approach – Wei-Lin-Weissfeld (WLW) method
- Conditional approach – Prentice, Williams and Peterson (PWP) method
- Intensity Approach– Andersen-Gill (AG) method

- The rate or mean approaches
  - Andersen-Gill approach
  - Poisson regression
  - Negative binormal

- Frailty model

- Recurrent events with a terminating event

Robust Sandwich Estimator for Variance

- Model a set of data Y_i for $i = 1, \dots, n$ with $f_i(Y_i|\theta)$
- For statistical inference of θ, use the log likelihood function

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} \log f_i(Y_i|\theta)$$

- Take the first partial derivatives of $\mathcal{L}(\theta)$

$$\mathcal{L}'(\theta) = \frac{\partial}{\partial\theta} \sum_{i=1}^{n} \log f_i(Y_i|\theta)$$

  - $U = \mathcal{L}'(\theta)$ is the score function
  - Let θ_0 be the truth, can be estimated by $\mathcal{L}'(\theta) = 0$

- Take the second partial derivatives of $\mathcal{L}(\theta)$

$$\mathcal{L}''(\theta) = \frac{\partial^2}{\partial\theta^2} \sum_{i=1}^{n} \log f_i(Y_i|\theta)$$

  - $I = \mathcal{L}''(\theta_0)$ is the information matrix

Robust Sandwich Estimator for Variance

- Taking Taylor expansion $\mathcal{L}(\theta)$ at $\theta = \theta_0$

$$\mathcal{L}(\theta) \approx \mathcal{L}(\theta_0) + \mathcal{L}'(\theta_0)(\theta - \theta_0) + \frac{1}{2}(\theta - \theta_0)'\mathcal{L}''(\theta_0)\,(\theta - \theta_0)$$

- Finding the maximum by solving $\mathcal{L}'(\hat{\theta}) = 0$
- Expanding $\mathcal{L}'(\hat{\theta})$ at $\hat{\theta} = \theta_0$, we obtain

$\mathcal{L}'(\hat{\theta}) \approx \mathcal{L}'(\theta_0) + \mathcal{L}''(\theta_0)'(\hat{\theta} - \theta_0) = 0$ when maximum

$$\mathcal{L}'(\theta_0) + (\hat{\theta} - \theta_0)'\mathcal{L}''(\theta_0) \approx 0$$
$$(\hat{\theta} - \theta_0) \approx \left(-\mathcal{L}''(\theta_0)\right)^{-1}\left(\mathcal{L}'(\theta_0)\right)'$$

- Then

$$cov\hat{\theta} = \left(-\mathcal{L}''(\theta_0)\right)^{-1}\left(cov\mathcal{L}'(\theta_0)\right)\ \left(-\mathcal{L}''(\theta_0)\right)^{-1}$$
$$= I^{-1}cov(U)I^{-1}$$

- Why use robust sandwich estimator?
  - Model is mis-specified, such as data is correlated, clustered, etc.

Wei-Lin-Weissfeld (WLW) Method

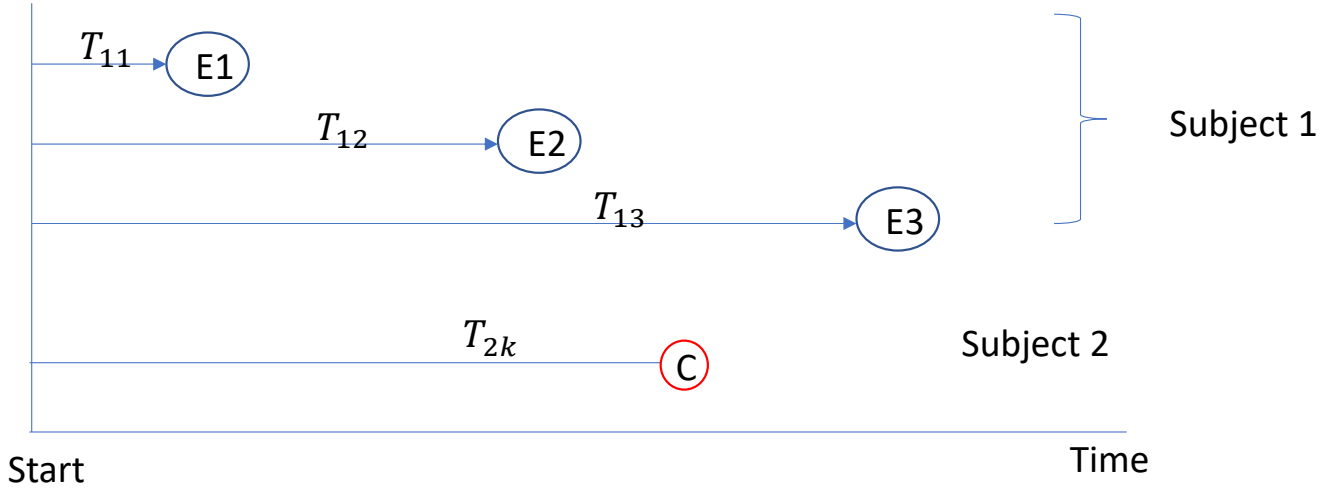
- Events are ordered or classified into different types
- The order/type is used as a stratification factor
  - Correlation among strata due to cluster effect - different types of events occurred in the same subjects
- Apply stratified Cox proportional hazard regression model
  - Time scale - Time from the study entry or randomization to the different types of events
  - Risk set – a marginal approach
    - Each stratum has its own risk set
    - All subjects are eligible in the risk set for the analysis of all stratum at the begining
  - Regression
    - Covariates can be stratum-specific
    - Effect parameter can be
      - Strata-specific
      - Common effect

- References

Wei, L. J., Lin, D. Y. and Weissfeld, L. 'Regression analysis of multivariate incomplete failure time data by modeling marginal distributions', ASA, 84, 1065—1073 (1989).

Lin, D. Y. (1994). "Cox Regression Analysis of Multivariate Failure Time Data: The Marginal Approach." *Statistics in Medicine* 13:2233–2247.

WLW Approach



- The recurrent event time is analyzed from start

Subject 2 in the risk set of E_1 & E_2 & E_3

WLW Method

- Let X_{ki} be the survival time for the k^{th} type of event of the i^{th} subject,
  - $k = 1, \dots, K$
  - $i = 1, \dots, n$
  - Time from randomization or study entry to the k^{th} type of event

- The observed survival time$(T_{ki}, \Delta_{ki}, Z_{ki})$,
  - $T_{ki} = \min(X_{ki}, C_{ki})$
  - $\Delta_{ki} = 0$ represents independent censoring
  - Z_{ki} is a vector of p dimensional

WLW Method

- For the k^{th} type of event

$$h_k(t, Z_k) = h_{k0}(t) e^{\beta_k' Z_k}$$

- The partial likelihood for the k^{th} type of event

$$L_k(\beta_k) = \prod_{i=1}^{n} \left\{ \frac{e^{\beta_k' z_{ki}}}{\sum_{l \in R_{ki}(t_{ki})} e^{\beta_k' z_{kl}}} \right\}^{\Delta_{ki}}$$

$$L(\beta) = \prod_{k=1}^{K} L_k(\beta_k)$$

  - Possible to fit a set of common parameters across the K strata

if you have common parameters, you must "glue" the event-specific likelihood together.

Inference

- Taking the first derivative of $\log L(\beta)$, we obtain MLEs $\hat{\beta}$

$$\widehat{\beta}' = \left(\hat{\beta}_1', \ldots, \hat{\beta}_K'\right)$$

  - $\hat{\beta}_k, k = 1, \ldots, K$ are correlated

$$\sqrt{n}\left(\hat{\beta}_1', \ldots, \hat{\beta}_K'\right)_{1\times pK} \sim N\left((\beta_1', \ldots, \beta_K'), \Sigma\right)$$

  Σ is $pK \times pK$ covariance matrix,

- Σ can be estimated with the robust sandwich estimator

Inference

- The covariates of interest
  - Let Z_1 be 0, 1 – a treatment indicator in the vector of p covariates

$$\left(\hat{\beta}_{11}, \dots, \hat{\beta}_{1K}\right)_{1\times K} \sim N\left((\beta_{11}, \dots, \beta_{1K}), \Sigma_{K\times K}\right)$$

  - To test H_0: $\beta_{11} = \cdots = \beta_{1K} = 0$
  - Form the test statistics to jointly test

$$\left(\hat{\beta}_{11}, \dots, \hat{\beta}_{1K}\right)\hat{\Sigma}_{K\times K}\left(\hat{\beta}_{11}, \dots, \hat{\beta}_{1K}\right)' \sim \chi^2_K$$

Inference

- Assume $\beta_{11}, = \cdots = \beta_{1K} = \beta_0$
- β_0 can be estimated by

$$\hat{\beta}_0 = \sum_{k=1}^{K} c_k \hat{\beta}_{1k} \text{ and } \sum_{k=1}^{K} c_k = 1$$

- $c = (c_1, \ldots, c_K)'$ can be estimated by

$$[e'(\Sigma_{K\times K})^{-1}e]^{-1}(\Sigma_{K\times K})^{-1}e$$
where $e = (1, \ldots, 1)'$

  - WLW shows that the estimated c results in the optimal weights – leads to the smallest asymptotic variance

WLW Method Data Structure

- Structure data set for K types of events
  - Form K strata
- Observations are in counting process style (start, stop)
  - Start – always study entry or randomization
  - Stop – the k^{th} event time or censored
  - Id – cluster indicator
  - Strata – indicator for the K types of events
  - Censoring indicator
- Each subject has K observations no matter the subjects has how many events

Bladder Data

- Variables
  - obs – Subject ID
  - Trt – Treatment 1=placebo 2=thiotepa
  - Time – follow-up time
  - Number – Initial number of tumor
  - Size – size (cm) of largest initial tumour
  - T1 – 1st event time
  - T2 – 2nd event time
  - T3 – 3rd event time
  - T4 – 4th event time

- Convert data set for the WLW method
  - ID – Subject ID
  - Visit – stratification indicator (with value k for the kth potential tumor recurrence)
  - Tstart – time of the (k – 1) recurrence for Visit=k, or the entry time 0 if VISIT=1, or the follow-up time if the (k – 1) recurrence does not occur
  - TStop, time of the kth recurrence if Visit=k or follow-up time if the kth recurrence does not occur
  - Status, censoring indicator of TStop (1=recurrence and 0=censored)
  - Other covariates

Obs	Trt	Time	Number	Size	T1	T2	T3	T4
1	1	0	1	1	.	.	.	.
2	1	1	1	3	.	.	.	.
3	1	4	2	1	.	.	.	.
4	1	7	1	1	.	.	.	.
5	1	10	5	1	.	.	.	.
6	1	10	4	1	6	.	.	.
7	1	14	1	1	.	.	.	.
8	1	18	1	1	.	.	.	.
9	1	18	1	3	5	.	.	.
10	1	18	1	1	12	16	.	.
11	1	23	3	3	.	.	.	.
12	1	23	1	3	10	15	.	.
13	1	23	1	1	3	16	23	.
14	1	23	3	1	3	9	21	.
15	1	24	2	3	7	10	16	24
16	1	25	1	1	3	15	25	.
17	1	26	1	2	.	.	.	.
18	1	26	8	1	1	.	.	.
19	1	26	1	4	2	26	.	.
20	1	28	1	2	25	.	.	.
21	1	29	1	4	.	.	.	.
22	1	29	1	2	.	.	.	.

Bladder Data – Formatted for WLW Method

Obs	ID	TStart	Trt	Number	Size	Visit	TStop	Status
1	1	0	1	1	1	1	0	0
2	1	0	1	1	1	2	0	0
3	1	0	1	1	1	3	0	0
4	1	0	1	1	1	4	0	0
5	2	0	1	1	3	1	1	0
6	2	1	1	1	3	2	1	0
7	2	1	1	1	3	3	1	0
8	2	1	1	1	3	4	1	0
9	3	0	1	2	1	1	4	0
10	3	4	1	2	1	2	4	0
11	3	4	1	2	1	3	4	0
12	3	4	1	2	1	4	4	0
13	4	0	1	1	1	1	7	0
14	4	7	1	1	1	2	7	0
15	4	7	1	1	1	3	7	0
16	4	7	1	1	1	4	7	0
17	5	0	1	5	1	1	10	0

45	12	0	1	1	3	1	10	1
46	12	10	1	1	3	2	15	1
47	12	15	1	1	3	3	23	0
48	12	23	1	1	3	4	23	0
49	13	0	1	1	1	1	3	1
50	13	3	1	1	1	2	16	1
51	13	16	1	1	1	3	23	1
52	13	23	1	1	1	4	23	0
53	14	0	1	3	1	1	3	1
54	14	3	1	3	1	2	9	1
55	14	9	1	3	1	3	21	1
56	14	21	1	3	1	4	23	0
57	15	0	1	2	3	1	7	1
58	15	7	1	2	3	2	10	1
59	15	10	1	2	3	3	16	1
60	15	16	1	2	3	4	24	1
61	16	0	1	1	1	1	3	1
62	16	3	1	1	1	2	15	1
63	16	15	1	1	1	3	25	1
64	16	25	1	1	1	4	25	0

ID: 14 doesn't have event T4.

Example – Bladder Data for WLW Method

```
title 'Wei-Lin-Weissfeld Model';
proc phreg data=Bladder covs(aggregate);
   model TStop*Status(0)=Trt1-Trt4 Number1-Number4 Size1-Size4;
   Trt1= Trt * (Visit=1);
   Trt2= Trt * (Visit=2);
   Trt3= Trt * (Visit=3);
   Trt4= Trt * (Visit=4);
   Number1= Number * (Visit=1);
   Number2= Number * (Visit=2);
   Number3= Number * (Visit=3);
   Number4= Number * (Visit=4);
   Size1= Size * (Visit=1);
   Size2= Size * (Visit=2);
   Size3= Size * (Visit=3);
   Size4= Size * (Visit=4);
   strata Visit;
   id ID;
   TREATMENT: test trt1,trt2,trt3,trt4/average e;
run;
```

for robust sandwich cov estimate .

Example – Bladder Data for WLW Method

The PHREG Procedure

Summary of the Number of Event and Censored Values

Stratum	Visit	Total	Event	Censored	Percent Censored
1	1	86	47	39	45.35
2	2	86	29	57	66.28
3	3	86	22	64	74.42
4	4	86	14	72	83.72
Total		344	112	232	67.44

Linear Coefficients for Test TREATMENT

Parameter	Row 1	Row 2	Row 3	Row 4	Average Effect
Trt1	1	0	0	0	0.67684
Trt2	0	1	0	0	0.25723
Trt3	0	0	1	0	-0.07547
Trt4	0	0	0	1	0.14140
Number1	0	0	0	0	0.00000
Number2	0	0	0	0	0.00000
Number3	0	0	0	0	0.00000

Test TREATMENT Results

	Wald Chi-Square	DF	Pr > ChiSq
	3.9668	4	0.4105

Average Effect for Test TREATMENT

Estimate	Standard Error	z-Score	Pr > |z|
-0.5489	0.2853	-1.9240	0.0543

Example – Bladder Data for WLW Method

```
title 'Wei-Lin-Weissfeld Model for
Common Effect';

proc phreg data=Bladder
covs(aggregate);

   model TStop*Status(0)=Trt Number
Size;

   strata Visit;

   id ID;

run;
```

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	24.7124	3	<.0001
Score (Model-Based)	27.8873	3	<.0001
Score (Sandwich)	11.7522	3	0.0083
Wald (Model-Based)	26.9033	3	<.0001
Wald (Sandwich)	15.5639	3	0.0014

Analysis of Maximum Likelihood Estimates							
Parameter	DF	Parameter Estimate	Standard Error	StdErr Ratio	Chi-Square	Pr > ChiSq	Hazard Ratio
Trt	1	-0.57984	0.30344	1.508	3.6516	0.0560	0.560
Number	1	0.20852	0.06568	1.400	10.0811	0.0015	1.232
Size	1	-0.05093	0.09304	1.335	0.2997	0.5841	0.950

Andersen & Gill (AG) Model

- AG (1982) model assumes independence between all observed event times
  - Within subjects
  - Between subjects
- The risk of the next event at t from the previous event or since randomization or study entry is the same, no matter what happened before t
  - Assume Poisson process
  - Proportional intensity model
- Lin and Wei (1989) proposed a robust variance estimator to account subject heterogeneity

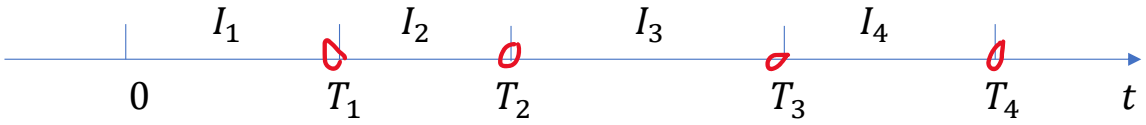
- Reference

Andersen PK, Gill RD. Cox's regression model for counting processes: A large sample study. Ann Stat. 1982;10(4):1100–20

Lin DY, Wei LJ. The robust inference for the cox proportional hazards model. J Am Stat Assoc. 1989;84(408):1074–8.

AG Model

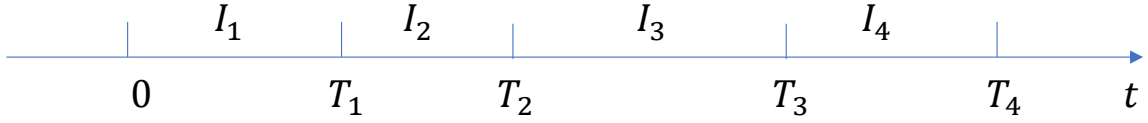
- Let $N(t)$ be a counting process that has $N(t)$ events occurred in the time interval $[0, t]$
  - $N(0) = 0$
  - $N(t) \in \{0,1,2, \dots\}, for\ all\ t \in [0, \infty)$
  - For $0 \leq s < t, N(t) - N(s)$ is the number of events occurred in the time interval $(s, t]$
- Let
  - T_i be the time from randomization or study entry to the i^{th} event
  - I_i be the gap time from the $(i-1)^{th}$ event to the i^{th} event
  - $i = 1, 2, \dots$



-

AG Model

- AG method assumes recurrent events follow Poisson process
- Definition
  - If $N(t)$ is a Poisson process with rate λ, then the interval times I_i are independent and $I_i \sim \exp(\lambda)$
    - The rate λ is the intensity
    - $N(0) = 0$
    - $N(t)$ consists independent increments with $\exp(\lambda)$
    - The number of events $k, k = 0,1,2, \ldots$ in any interval of length t has Poisson (λt) distribution
      - $P(k = m) = \frac{e^{-\lambda t}\lambda t^m}{m!}$



AG Model

- The observed survival time in the i^{th} subject, (T_{ri}, Δ_i, Z_i),
  - $r = 1, \ldots, R_i$ represents the r^{th} recurrent event in the i^{th} subject

$$\lambda(t, Z_i(t)) = \lambda_0(t) e^{\beta' Z_i(t)}$$

if not tim varying, should be a constant.

- The partial likelihood for all events

$$L(\beta) = \prod_{i=1}^{n} \prod_{r=1}^{R_i} \left\{ \frac{e^{\beta' z_i(t_{ri})}}{\sum_{l \in R_{ri}(t_{ri})} e^{\beta' z_l(t_{ri})}} \right\}^{\Delta_{ri}}$$

- Notice $R_{ri}(t_{ri})$ includes all subjects, unless
  - death
  - censored

in WLW, their exist stratas.

AG Model

- Inference

$$\mathcal{L}(\beta) = \log L(\beta)$$

$$\frac{\partial}{\partial \beta}\mathcal{L}(\hat{\beta}) = 0$$

- The robust variance estimator Σ Lin and Wei (1989)

$$\sqrt{n}(\hat{\beta} - \beta_0) \sim N(0, \Sigma)$$

AG Method – Data Structure

- If a subject has k events
  - The subject has $k+1$ observations if follow-up is beyond the last event
    - The first k observations are for events
    - The last observation is censored
  - The subject has k observations if follow-up is stopped – all observations are for events see ID 13
- Survival time is in counting process style (start, stop)
  - Start
    - The previous event time
    - The entry of study for the first event
  - Stop
    - The event time
    - The censoring time if the last observation

```
data bladder_AG;
  set bladder;
  if tstart<tstop
  run;
```

Obs	ID	TStart	Trt	Number	Size	Visit	TStop	Status
1	2	0	1	1	3	1	1	0
2	3	0	1	2	1	1	4	0
3	4	0	1	1	1	1	7	0
4	5	0	1	5	1	1	10	0
5	6	0	1	4	1	1	6	1
6	6	6	1	4	1	2	10	0
7	7	0	1	1	1	1	14	0
8	8	0	1	1	1	1	18	0
9	9	0	1	1	3	1	5	1
10	9	5	1	1	3	2	18	0
11	10	0	1	1	1	1	12	1
12	10	12	1	1	1	2	16	1
13	10	16	1	1	1	3	18	0
14	11	0	1	3	3	1	23	0
15	12	0	1	1	3	1	10	1
16	12	10	1	1	3	2	15	1
17	12	15	1	1	3	3	23	0
18	13	0	1	1	1	1	3	1
19	13	3	1	1	1	2	16	1
20	13	16	1	1	1	3	23	1
21	14	0	1	3	1	1	3	1

Example – Bladder Data for AG Method

```
title 'Intensity Model and
Proportional Means Model';

proc phreg data=Bladder covm
covs(aggregate);

   model (TStart, TStop) * Status(0)
= Trt Number Size;

   id id;

run;
```

Analysis of Maximum Likelihood Estimates with Model-Based Variance Estimate						
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
Trt	1	-0.45979	0.19996	5.2873	0.0215	0.631
Number	1	0.17165	0.04733	13.1541	0.0003	1.187
Size	1	-0.04256	0.06903	0.3801	0.5375	0.958

Analysis of Maximum Likelihood Estimates with Sandwich Variance Estimate							
Parameter	DF	Parameter Estimate	Standard Error	StdErr Ratio	Chi-Square	Pr > ChiSq	Hazard Ratio
Trt	1	-0.45979	0.25801	1.290	3.1757	0.0747	0.631
Number	1	0.17165	0.06131	1.296	7.8373	0.0051	1.187
Size	1	-0.04256	0.07555	1.094	0.3174	0.5732	0.958

Prentice, Williams and Peterson (PWP) Method conditional approach.

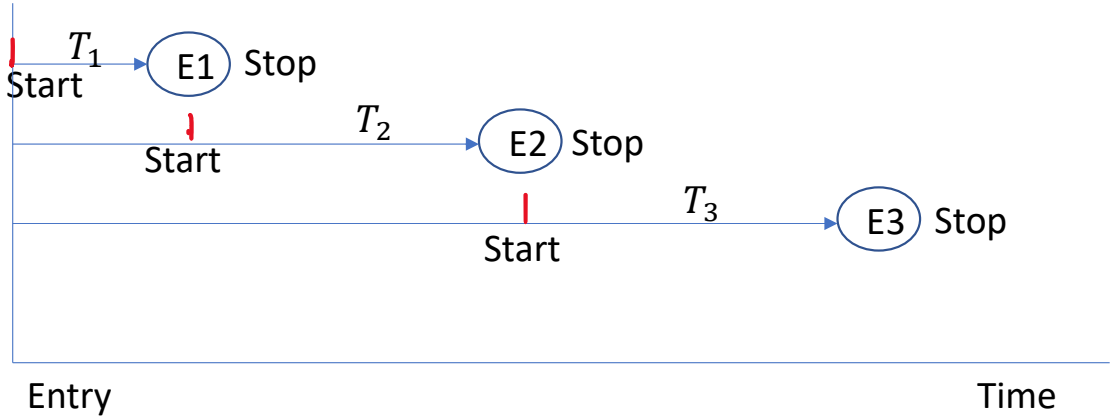
- Events are ordered
- The order is used as a stratification factor
- Apply Cox proportional hazard regression model
  - Time scale
    - Gap time – Time from the previous event to the next event
    - Total time – Time from randomization or the study entry to events
  - Risk set – a conditional approach
    - Only subjects who had the previous event are eligible for the analysis of the next event
  - Regression
    - Covariates can be stratum-specific
    - Effect parameter can be
      - Strata-specific
      - Common effect
- Reference

  Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika. 1981;68(2):373–9.

PWP Method

- Let X_{ki} be the survival time for the k^{th} type of event of the i^{th} subject,
  - $k = 1, \dots, K$
  - $i = 1, \dots, n$
  - X_{ki} can be gap time or time from the study entry
- The observed survival time$(T_{ki}, \Delta_{ki}, Z_{ki})$,
  - $T_{ki} = \min(X_{ki}, C_{ki})$
  - $\Delta_{ki} = 0$ represents independent censoring
  - Z_{ki} is a vector of p dimensional

PWP Method



- The recurrent event time can be analyzed
  - From start to event time (stop)
  - From study entry to event time
- If a subject experienced only one event, the subject will be in the risk set for the rest of the events

PWP Method

- For the k^{th} event

$$h_k(t, Z_k) = h_{k0}(t) e^{\beta_k' Z_k}$$

- The partial likelihood for the k^{th} type of event

$$L_k(\beta_k) = \prod_{i=1}^{n} \left\{ \frac{e^{\beta_k' z_{ki}}}{\sum_{l \in R_{ki}^{PWP}(t_{ki})} e^{\beta_k' z_{kl}}} \right\}^{\Delta_{ki}}$$

- The risk set includes only those who had the $(k-1)^{th}$ event
- Take the first derivative of $\log L_k(\beta_k)$, we obtain MLEs $\hat{\beta}_k$
- Use robust sandwich variance estimator for inference

PWP Method

- Structure data set for K events
  - For K ordered strata
- If a subject has k events
  - The subject has $k+1$ observations if follow-up is beyond the last event
    - The first k observations are for events
    - The last observation is censored at the end of follow-up
  - The subject has k observations if follow-up is stopped at the last event
- All subjects should be in the first strata for the first event

1 obs if no event

```
data Bladder2(drop=LastStatus);
  retain LastStatus;
  set Bladder;
  by ID;
  if first.id then LastStatus=1;
  if (Status=0 and LastStatus=0)
     then delete;
  LastStatus=Status;
  Gaptime=Tstop-Tstart;
run;
```

- If a subject has k events
  - The subject has $k+1$ observations if follow-up is beyond the last event
    - The first k observations are for events
    - The last observation is censored
  - The subject has k observations if follow-up is stopped – all observations are for events see ID=13
- Survival time is in counting process style (start, stop)
  - Start
    - The previous event time
    - The entry of study for the first event
  - Stop
    - The event time
    - The censoring time if the last observation

Obs	ID	TStart	Trt	Number	Size	Visit	TStop	Status	Gaptime
1	1	0	1	1	1	1	0	0	0
2	2	0	1	1	3	1	1	0	1
3	3	0	1	2	1	1	4	0	4
4	4	0	1	1	1	1	7	0	7
5	5	0	1	5	1	1	10	0	10
6	6	0	1	4	1	1	6	1	6
7	6	6	1	4	1	2	10	0	4
8	7	0	1	1	1	1	14	0	14
9	8	0	1	1	1	1	18	0	18
10	9	0	1	1	3	1	5	1	5
11	9	5	1	1	3	2	18	0	13
12	10	0	1	1	1	1	12	1	12
13	10	12	1	1	1	2	16	1	4
14	10	16	1	1	1	3	18	0	2
15	11	0	1	3	3	1	23	0	23
16	12	0	1	1	3	1	10	1	10
17	12	10	1	1	3	2	15	1	5
18	12	15	1	1	3	3	23	0	8
19	13	0	1	1	1	1	3	1	3
20	13	3	1	1	1	2	16	1	13
21	13	16	1	1	1	3	23	1	7
22	13	23	1	1	1	4	23	0	0

Example – Bladder Data for PWP Method

```
title 'PWP Total Time Model with Noncommon Effects';
proc phreg data=Bladder2 covs(aggregate);
   model Tstop * Status(0) = Trt1-Trt4 Number1-Number4
                             Size1-Size4;
   Trt1= Trt * (Visit=1);
   Trt2= Trt * (Visit=2);
   Trt3= Trt * (Visit=3);
   Trt4= Trt * (Visit=4);
   Number1= Number * (Visit=1);
   Number2= Number * (Visit=2);
   Number3= Number * (Visit=3);
   Number4= Number * (Visit=4);
   Size1= Size * (Visit=1);
   Size2= Size * (Visit=2);
   Size3= Size * (Visit=3);
   Size4= Size * (Visit=4);
   strata Visit;
run;
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	17.4211	12	0.1344
Score (Model-Based)	18.5546	12	0.0999
Score (Sandwich)	20.2426	12	0.0626
Wald (Model-Based)	17.7388	12	0.1239
Wald (Sandwich)	20.8363	12	0.0528

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	StdErr Ratio	Chi-Square	Pr > ChiSq	Hazard Ratio
Trt1	1	-0.51757	0.30750	0.974	2.8330	0.0923	0.596
Trt2	1	-0.42584	0.37389	0.929	1.2972	0.2547	0.653
Trt3	1	-0.89894	0.51378	0.952	3.0613	0.0802	0.407
Trt4	1	-0.23739	0.52974	0.776	0.2008	0.6541	0.789
Number1	1	0.23605	0.07208	0.948	10.7243	0.0011	1.266
Number2	1	0.00117	0.08249	0.880	0.0002	0.9886	1.001
Number3	1	0.01468	0.10423	0.786	0.0198	0.8880	1.015
Number4	1	0.29306	0.16433	0.745	3.1805	0.0745	1.341
Size1	1	0.06790	0.08529	0.842	0.6338	0.4260	1.070
Size2	1	-0.12515	0.10367	0.885	1.4574	0.2273	0.882
Size3	1	-0.21520	0.17895	1.005	1.4462	0.2291	0.806
Size4	1	0.25135	0.33916	1.166	0.5492	0.4586	1.286

Example – Bladder Data for PWP Method

```
title 'PWP Gap Time Model with Noncommon Effects';
proc phreg data=Bladder2 covs(aggregate);
   model gaptime * Status(0) = Trt1-Trt4 Number1-Number4
                               Size1-Size4;
   Trt1= Trt * (Visit=1);
   Trt2= Trt * (Visit=2);
   Trt3= Trt * (Visit=3);
   Trt4= Trt * (Visit=4);
   Number1= Number * (Visit=1);
   Number2= Number * (Visit=2);
   Number3= Number * (Visit=3);
   Number4= Number * (Visit=4);
   Size1= Size * (Visit=1);
   Size2= Size * (Visit=2);
   Size3= Size * (Visit=3);
   Size4= Size * (Visit=4);
   strata Visit;
run;
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	17.8089	12	0.1216
Score (Model-Based)	19.6097	12	0.0748
Score (Sandwich)	20.4570	12	0.0589
Wald (Model-Based)	18.4759	12	0.1020
Wald (Sandwich)	24.9959	12	0.0148

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	StdErr Ratio	Chi-Square	Pr > ChiSq	Hazard Ratio
Trt1	1	-0.51757	0.30750	0.974	2.8330	0.0923	0.596
Trt2	1	-0.25911	0.40206	0.992	0.4153	0.5193	0.772
Trt3	1	0.22105	0.61953	1.128	0.1273	0.7212	1.247
Trt4	1	-0.19498	0.62892	0.980	0.0961	0.7565	0.823
Number1	1	0.23605	0.07208	0.948	10.7243	0.0011	1.266
Number2	1	-0.00571	0.10040	1.039	0.0032	0.9547	0.994
Number3	1	0.12935	0.16189	1.014	0.6384	0.4243	1.138
Number4	1	0.42079	0.17125	0.864	6.0373	0.0140	1.523
Size1	1	0.06790	0.08529	0.842	0.6338	0.4260	1.070
Size2	1	-0.11636	0.11748	0.985	0.9810	0.3219	0.890
Size3	1	0.24995	0.16876	0.730	2.1935	0.1386	1.284
Size4	1	0.03557	0.28427	0.979	0.0157	0.9004	1.036

Example – Bladder Data for PWP Method

```
title 'PWP Total Time Model with Common
Effects';
proc phreg data=Bladder2 covs(aggregate);
   model tstop * Status(0) = Trt Number Size;
strata Visit;
run;
```

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	8.7512	3	0.0328
Score (Model-Based)	8.8795	3	0.0309
Score (Sandwich)	10.2693	3	0.0164
Wald (Model-Based)	8.7957	3	0.0321
Wald (Sandwich)	9.3663	3	0.0248

Analysis of Maximum Likelihood Estimates							
Parameter	DF	Parameter Estimate	Standard Error	StdErr Ratio	Chi-Square	Pr > ChiSq	Hazard Ratio
Trt	1	-0.48972	0.19738	0.943	6.1559	0.0131	0.613
Number	1	0.11027	0.05000	0.979	4.8645	0.0274	1.117
Size	1	-0.03773	0.06512	0.964	0.3357	0.5623	0.963

Example – Bladder Data for PWP Method

```
title 'PWP Gap Time Model with Common
Effects';

proc phreg data=Bladder2 covs(aggregate);

   model gaptime * Status(0) = Trt Number
Size;

strata Visit;

run;
```

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	8.7559	3	0.0327
Score (Model-Based)	9.5977	3	0.0223
Score (Sandwich)	9.4752	3	0.0236
Wald (Model-Based)	9.4570	3	0.0238
Wald (Sandwich)	8.2253	3	0.0416

Analysis of Maximum Likelihood Estimates							
Parameter	DF	Parameter Estimate	Standard Error	StdErr Ratio	Chi-Square	Pr > ChiSq	Hazard Ratio
Trt	1	-0.26952	0.21023	1.012	1.6436	0.1998	0.764
Number	1	0.15353	0.05414	1.039	8.0414	0.0046	1.166
Size	1	0.00684	0.06708	0.958	0.0104	0.9188	1.007

Comparisons of the Three Models

- AG
  - Make strongest assumptions
    - Assuming constant intensity
    - Independency between events
  - Interest in the overall effect on the intensity of all events
  - Same as Poisson regression if no time-varying covariates
- PWP
  - Risk sets for events other than the first no longer preserve the randomization principle
  - Need extra care in interpretation
- WLW
  - Easy to interpret the results for different types of events
  - Estimating marginal hazard ratios

Count Data

- Treating recurrent events as count data Y
  - Evaluate the rate of recurrent events λ, intensity
  - The counts over a follow-up period t
- Assume the count data follow
  - Poisson distribution
    - $Y \sim Poi(\mu)$
    - $E(Y) = \mu$ and $\text{var}(Y) = \mu$
  - Quasi-Poisson distribution
    - $Y \sim qPoi(\mu, \theta)$
    - $E(Y) = \mu$ and $\text{var}(Y) = \theta\mu,\ \theta > 1$
  - Negative binomial distribution
    - $Y \sim NB(\mu, \kappa)$
    - $E(Y) = \mu$ and $\text{var}(Y) = \mu + \kappa\mu^2, \kappa > 0$

Mean and Rate Functions of the Recurrent Events

- Using the counting process framework
- $E\{dN(t)\} = \lambda_0(t)dt$
- $E\{dN(t)|Z(s)\} = \exp\left(\beta' Z(s)\right)\lambda_0(t)dt$
- Let $d\mu_z(t) = E\{dN(t)|Z(s)\}$, and $\mu_z(t) = E\{N(t)|Z(s)\}$

$$d\mu_z(t) = \exp\left(\beta' Z(s)\right) d\mu_0(t)$$
$$\mu_z(t) = \exp\left(\beta' Z(s)\right) \mu_0(t)$$

- Note, $\mu_z(t)$ and $\mu_0(t)$ pertain to the mean functions of the recurrent events,
- When AG model has no time-varying covariates, AG model is the same as the mean and rate model

- Reference
  - Lin DY, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. J R Stat Soc B 2000;62:711–30.

Count Data Regression Model

- Let λ be the intensity, the mean count is

$$E(Y|t) = \mu = \lambda t$$

$$\log\{E(Y_i|t_i)\} = \log t_i + \log \lambda$$

- Let Z be a covariate vector and $\lambda_i(Z_i) = \lambda_0 e^{\beta' Z_i}$

$$\log\{E(Y_i|Z_i)\} = \log t_i + \log \lambda_i(Z_i) = \log t_i + \beta' Z_i + \log \lambda_0$$

- Log follow-up time $\log t_i$ as an offset for subject i

- For over-dispersed count data
  - $E(Y) > \text{var}(Y)$
  - Negative binomial distribution often used to account over-dispersion
    - A more flexible model than Poisson distribution

Count Data

```
proc genmod data = bladder_nb;
  class Trt;
  model count = Trt number size / dist = Poisson offset = ltime link=log;
  *estimate 'trt=1' Intercept 1 Trt 0 1 / exp;
  *estimate 'trt=2' Intercept 1 Trt 1 0 / exp;
run;
```

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.7582	0.2708	-4.2890	-3.2274	192.58	<.0001
Trt	1	1	0.4333	0.2001	0.0411	0.8255	4.69	0.0304
Trt	2	0	0.0000	0.0000	0.0000	0.0000	.	.
Number		1	0.1726	0.0490	0.0766	0.2685	12.42	0.0004
Size		1	-0.0462	0.0705	-0.1844	0.0920	0.43	0.5120
Scale		0	1.0000	0.0000	1.0000	1.0000		

```
proc genmod data = bladder_nb;
  class Trt;
  model count = Trt number size / dist = poisson offset = ltime link=log
dscale;
  *estimate 'trt=1' Intercept 1 Trt 0 1 / exp;
  *estimate 'trt=2
run;
```

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.7582	0.3518	-4.4478	-3.0686	114.10	<.0001
Trt	1	1	0.4333	0.2600	-0.0763	0.9429	2.78	0.0956
Trt	2	0	0.0000	0.0000	0.0000	0.0000	.	.
Number		1	0.1726	0.0636	0.0479	0.2972	7.36	0.0067
Size		1	-0.0462	0.0916	-0.2258	0.1333	0.25	0.6138
Scale		0	1.2991	0.0000	1.2991	1.2991		

Note: The scale parameter was estimated by the square root of DEVIANCE/DOF.

Count Data

```
proc genmod data = bladder_nb;
  class Trt;
  model count = Trt number size / dist = NB offset = ltime link=log;
  *estimate 'trt=1' Intercept 1 Trt 0 1 / exp;
  *estimate 'trt=2' Intercept 1 Trt 1 0 / exp;
run;
```

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.8047	0.3589	-4.5082	-3.1013	112.37	<.0001
Trt	1	1	0.4560	0.2719	-0.0769	0.9889	2.81	0.0935
Trt	2	0	0.0000	0.0000	0.0000	0.0000	.	.
Number		1	0.1874	0.0738	0.0426	0.3321	6.44	0.0112
Size		1	-0.0324	0.0924	-0.2136	0.1488	0.12	0.7261
Dispersion		1	0.5085	0.2555	0.1899	1.3612		

Frailty Models

- Model random effect from
  - Clustered survival data
    - Patients are clustered by hospitals or villages
    - Patients may share similar effect if from the same hospital
    - Hospitals and villages can be modeled as random effect when the number of hospitals or villages is large
  - Recurrent events from the same subjects
    - Events from the same subjects may share similar risk
    - Subjects are considered as random effect

- Let R represent the random factor
  - $R\ contains\ levels$: $r_1, r_2, \dots, r_n$ for n groups
    - n subjects, hospitals, …
  - Values $r_1, r_2, \dots, r_n$ represent the frailty of developing events
  - Incorporate the values in the PH models
    - Obtain the shared frailty model

same cluster, same frailty.

Frailty Models

- Let r_j represent the j^{th} level of frailty, $j = 1,2, \dots, n$
- The shared frailty models for the i^{th} events in the j^{th} group

$$h_{ij}(t, Z_{ij}) = h_0(t)e^{\beta' Z_{ij} + r_j}$$

- Let $\xi_j = e^{r_j}$

$$h_{ij}(t, Z_{ij}) = \xi_j h_0(t)e^{\beta' Z_{ij}}$$

- The distribution of ξ_j can be
  - Lognormal
    - r_j is normally distributed with mean=0
  - Gamma distribution

Example – Bladder Data for Frailty Models

```
title 'Frailty Model for Common Effect';
proc phreg data=Bladder ;
  class id;
  model TStop*Status(0)=Trt Number Size;
  random ID;
  hazardratio 'Frailty Model Analysis'
trt;
run;
*** WLW data set format
Default frailty lognormal distribution;
*** Fit total time;
```

Type 3 Tests

Effect	Wald Chi-Square	DF	Pr > ChiSq	Adjusted DF	Adjusted Pr > ChiSq
Trt	3.4591	1	0.0629	0.3777	0.0173
Number	7.0505	1	0.0079	0.3161	0.0014
Size	0.0423	1	0.8370	0.3504	0.4516
ID	133.7122	.	.	46.3647	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
Trt	1	-0.70722	0.38025	3.4591	0.0629	0.493
Number	1	0.27083	0.10200	7.0505	0.0079	1.311
Size	1	-0.02622	0.12741	0.0423	0.8370	0.974

Frailty Model Analysis: Hazard Ratios for Trt

Description	Point Estimate	95% Wald Confidence Limits	
Trt Unit=1	0.493	0.234	1.039

Example – Bladder Data for Frailty Models

```
title 'Frailty Model for Common Effect';
proc phreg data=Bladder_ag ;
  class id;
  model (Tstart, Tstop)*Status(0)=Trt
Number Size;
  random ID;
  hazardratio 'Frailty Model Analysis'
trt;
run;
*** AG data set format
Default frailty lognormal distribution;
```

Type 3 Tests

Effect	Wald Chi-Square	DF	Pr > ChiSq	Adjusted DF	Adjusted Pr > ChiSq
Trt	3.3889	1	0.0656	0.4774	0.0243
Number	6.4099	1	0.0113	0.4138	0.0030
Size	0.0475	1	0.8274	0.4580	0.5359
ID	75.1231	.	.	41.9277	0.0012

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
Trt	1	-0.58384	0.31715	3.3889	0.0656	0.558
Number	1	0.22493	0.08884	6.4099	0.0113	1.252
Size	1	-0.02335	0.10708	0.0475	0.8274	0.977

Frailty Model Analysis: Hazard Ratios for Trt

Description	Point Estimate	95% Wald Confidence Limits	
Trt Unit=1	0.558	0.300	1.038

Example – Bladder Data for Frailty Models

```
title 'Frailty Model for Common Effect';
proc phreg data=Bladder2;
  class id;
  model (Tstart, Tstop)*Status(0)=Trt
Number Size;
  random ID;
  hazardratio 'Frailty Model Analysis'
trt;
run;
*** PWP data set format,
Default frailty lognormal distribution;
```

Type 3 Tests

Effect	Wald Chi-Square	DF	Pr > ChiSq	Adjusted DF	Adjusted Pr > ChiSq
Trt	3.2594	1	0.0710	0.5167	0.0293
Number	7.2978	1	0.0069	0.4530	0.0020
Size	0.0454	1	0.8313	0.5007	0.5743
ID	72.6970	.	.	35.3180	0.0002

Analysis of Maximum Likelihood Estimates

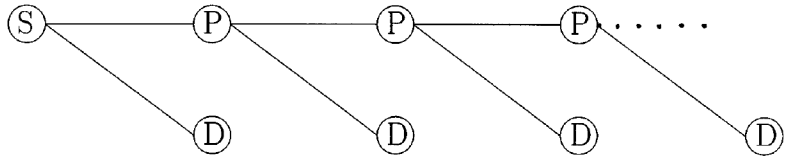
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
Trt	1	-0.54591	0.30238	3.2594	0.0710	0.579
Number	1	0.21574	0.07986	7.2978	0.0069	1.241
Size	1	-0.02164	0.10157	0.0454	0.8313	0.979

Frailty Model Analysis: Hazard Ratios for Trt

Description	Point Estimate	95% Wald Confidence Limits	
Trt Unit=1	0.579	0.320	1.048

Recurrent Events with a Terminating Event

- Death may not have the same weight as the recurrent events
- Death may informatively censor the recurrent events
- Analyzing the observed events only may bias the interpretation of treatment effect
- Methods
  - Re-configure the events – use WLW methods
  - Joint modeling recurrent and terminating events



Re-configure the Events

Observations	Scenario I				Scenario II				Scenario III				Scenario IV			
	$k=1$	$k=2$	$k=3$	$k=4$	$k=1$	$k=2$	$k=3$	$k=4$	$k=1$	$k=2$	$k=3$	$k=4$	$k=1$	$k=2$	$k=3$	$k=4$
$P_1P_2P_3D$	P_1	P_2	P_3	D	P_1	P_2	P_3	D	P_1	P_2	P_3	D	P_1	P_2	P_3	D
P_1P_2D	P_1	P_2	D^+	D	P_1	P_2	D	D	P_1	P_2	D	D^+	P_1	P_2	0^+	D
P_1D	P_1	D^+	D^+	D	P_1	D	D	D	P_1	D	D^+	D^+	P_1	0^+	0^+	D
D	D^+	D^+	D^+	D	D	D	D	D	D	D^+	D^+	D^+	0^+	0^+	0^+	D
$P_1P_2P_3C$	P_1	P_2	P_3	C^+	P_1	P_2	P_3	C^+	P_1	P_2	P_3	C^+	P_1	P_2	P_3	C^+
P_1P_2C	P_1	P_2	C^+	C^+	P_1	P_2	C^+	C^+	P_1	P_2	C^+	C^+	P_1	P_2	C^+	C^+
P_1C	P_1	C^+	C^+	C^+	P_1	C^+	C^+	C^+	P_1	C^+	C^+	C^+	P_1	C^+	C^+	C^+
C	C^+	C^+	C^+	C^+	C^+	C^+	C^+	C^+	C^+	C^+	C^+	C^+	C^+	C^+	C^+	C^+

removed from the risk set

Joint Modeling

- Model recurrent events

$$h_{ik}^{R}(t|Z_{ik},\mu_i,\omega_{ik}) = h_{0ik}^{R}(t)\exp(\beta_R' Z_{ik} + \mu_i + \omega_{ik})$$

  ω_{ik} - frailty for the recurrent events

  μ_i - frailty for between events correlation, recurrent events vs. terminating events

- Model terminating events

$$h_i^D(t|Z_i,\mu_i) = h_0^D(t)\exp(\beta_S' Z_{ik} + \phi\mu_i)$$

Homework 11

- 1. Use the Diabetic dataset from R survival package to perform analyses
  - Identify two methods for multi-variate survival analyses and lay out the rationale
  - Perform the analyses and interpreting the results