Analysis of Gap times

1. Renewal process and related models
2. Extension of renewal models
3. Marginal gap time probabilities

UNC Charlotte

P.h.d student Peilin Chen

1. Gap time and renewal process

1.1 Gap times:

- Times between sequentially ordered events (gap times) are often of interest in biomedical studies.
- For example, in a cancer study, the gap times from incidence-to-remission and remission-to-recurrence may be examined.

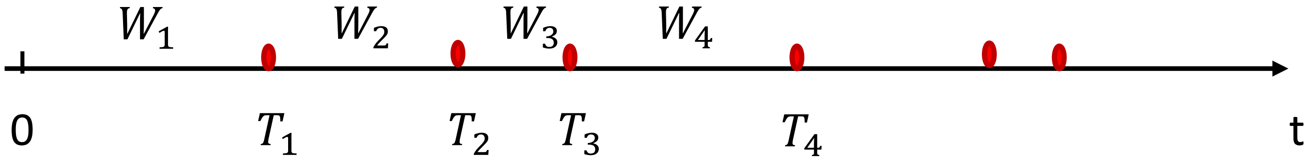
1.2

- Renewal processes are ones in which the gaps between successive events are independent and identically distributed.

Renewal process:

$T_1, T_2, T_3, \dots$ *are event times*;

$W_1, W_2,$ ...are the gaps between event times,which are assumed to be i.i.d



For renewal process

- The intensity is:

$$\lambda(t|H(t)) = h(B(t))\, t > 0\,, (1)$$

Where $B(t) = t - T_{N(t-)}$ is the time since the most recent event before t.

h(.) is the hazard functions for the variables W_j

- Hazard rate:

Wj have common density function f(w) and survivor function S(w) = P (W ≥ w), then :

$$h(w) = \frac{f(w)}{S(w)} = \lim_{\Delta w \downarrow 0} \frac{\Pr(W < w + \Delta w | W \geq w)}{\Delta w}\,.$$

1.3 Likelihood function for observed gap times W_{ij}

Assumption:

1. $t = 0$ is the start of the event process
2. individual i is observed over time [0, τ_i], there are m individuals
3. if n_i event is observed at times $0 < t_{i1} < t_{i2} <\ldots t_{in} \leq \tau_i$
4. allowing for fixed covariates x_i for gaps W_{ij}

then the likelihood function for m individuals are:

$$L = \prod_{i=1}^{m} \left\{ \prod_{j=1}^{ni} h\left(w_{ij} \middle| x_i\right) \exp\left(-H\left(w_{ij} \middle| x_i\right)\right) \right\} \exp\left(-H\left(w_{i,n_{i+1}} \middle| x_i\right)\right) \qquad (2)$$

Where,

$$\mathrm{H}(w|x_i) = \int_0^w h(u|x_i)du$$

is the cumulative hazard function for W_{ij}

Furthermore, let $f(w|x) = h(w|x)\exp(-H(w|x)$, $S(w) = \exp(-H(w|x)$ be the density and survival function for $W_{ij}|x$,

Then (2) could be rewrite as:

$$L = \prod_{i=1}^{m} \prod_{j=1}^{n_i} f\left(w_{ij}\middle|x_i\right) \cdot S\left(w_{i,n_{i+1}}\middle|x_i\right) \qquad (3)$$

- 1.4 common models

1. Parametric -AFT(accelerated failure time) model

$Y = \log W$ has a location-scale distribution of the form

$$Y = \beta_0 + x'\beta + \sigma\epsilon\,, \qquad (4)$$

where $x = (x1, \ldots, xk)'$ is a covariate vector, $\beta = (\beta1, \ldots\ \beta k)'$ is a vector of regression coefficients, $\sigma > 0$ is a scale parameter, ε is a random variable whose distribution is independent of x.

AFT model handles cases where the covariate values are fixed within gaps but vary across gaps:

$$Y_{ij} = \beta_0 + x'_{ij}\beta + \sigma\epsilon_{ij}\,, \qquad (5)$$

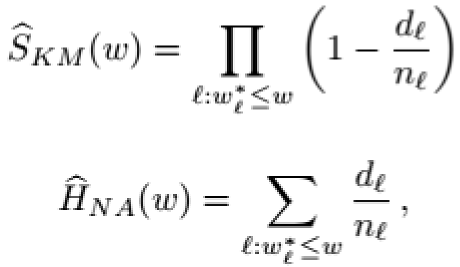
If covariates vary within gaps, cox model is preferred

AFT model a summary of the commonly used distribution of ϵ, and the corresponding distribution of $W_{ij} = \exp(\mathrm{Y}_{ij})$ is :

Distribution of ϵ	Distribution of W_{ij}	Syntax in $survreg$
extreme values(2 par.)	Weibull	dist=weibull
extreme values(1 par.)	exponential	dist= exponential
Log-gamma	gamma	dist=gamma
logistic	log-logistic	dist=llogistic
normal	Log-normal	dist=lognormal

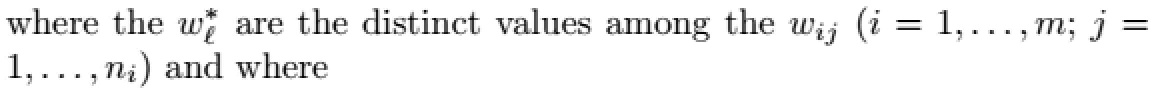
- K-M (Kaplan- Meier estimate) (Non-parametric)

The Kaplan–Meier nonparametric estimate of $S(w)$ and the Nelson–Aalen estimate of $H(w)$ can be used when there are no covariates.



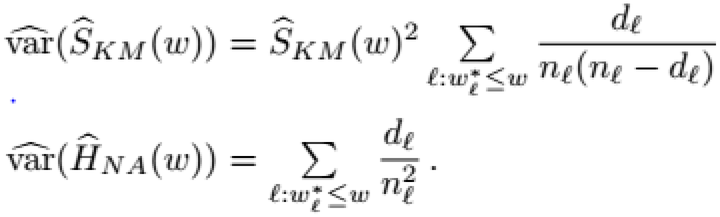
$$\widehat{S}_{KM}(w) = \prod_{\ell : w_\ell^* \le w} \left(1 - \frac{d_\ell}{n_\ell}\right) \tag{6}$$

$$\widehat{H}_{NA}(w) = \sum_{\ell : w_\ell^* \le w} \frac{d_\ell}{n_\ell}\,, \tag{7}$$



where the w_ℓ^* are the distinct values among the w_{ij} $(i = 1, \ldots, m;\ j = 1, \ldots, n_i)$ and where

Variance estimate is (6) and (7) are



$$\widehat{\mathrm{var}}(\widehat{S}_{KM}(w)) = \widehat{S}_{KM}(w)^2 \sum_{\ell : w_\ell^* \le w} \frac{d_\ell}{n_\ell(n_\ell - d_\ell)}$$

$$\widehat{\mathrm{var}}(\widehat{H}_{NA}(w)) = \sum_{\ell : w_\ell^* \le w} \frac{d_\ell}{n_\ell^2}\,.$$

- Semi-parametric Cox model

The Cox semiparametric multiplicative hazards model in which the hazard function for W_{ij} given x_{ij} is of the form

$$h(w|x_{ij}) = h_0(w)\exp(x'_{ij}\beta) \tag{8}$$

The partial likelihood function:

$$L(\beta) = \prod_{i=1}^{m}\prod_{j=1}^{n_i}\left\{\frac{\exp(x'_{ij}\beta)}{\sum_{l=1}^{m}\sum_{k=1}^{n_\ell+1} I(w_{lk} \geq w_{ij})\exp(x'_{lk}\beta)}\right\}. \tag{9}$$

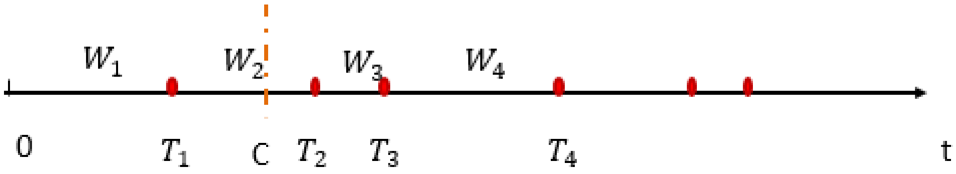
Estimate of $H_0\,(w) = \int_0^w h_0(u)du$

$$\widehat{H}_0(w) = \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left\{\frac{I(w_{ij} \leq w)}{\sum_{l=1}^{m}\sum_{k=1}^{n_l+1} I(w_{lk} \geq w_{ij})\exp(x'_{lk}\widehat{\beta})}\right\}. \tag{10}$$

2. Extension of renewal models

- Without covariates, the assumption that gap times W_{ij} are i.i.d are too strong.

For example, in cancer study, the data are usually right censored. The longer the a give individual's time until fist event, the greater the probability that the time between their first and second event is censored.



If the observation period for individual i is $[0, \tau_i]$, then the potential censoring time for W_{i2} is $\tau_i - \min(W_{i1}, \tau_i)$, which is not independent of W_{i1}. So the marginal analysis of W_{i2} without conditioning on W_{i1} involves dependent censoring.

- If we ignore the effect of dependent censoring, the second (W_{i2}) and subsequent gap times(W_{i2}, W_{i3}, $W_{i4,\ldots}$) would be underestimated.

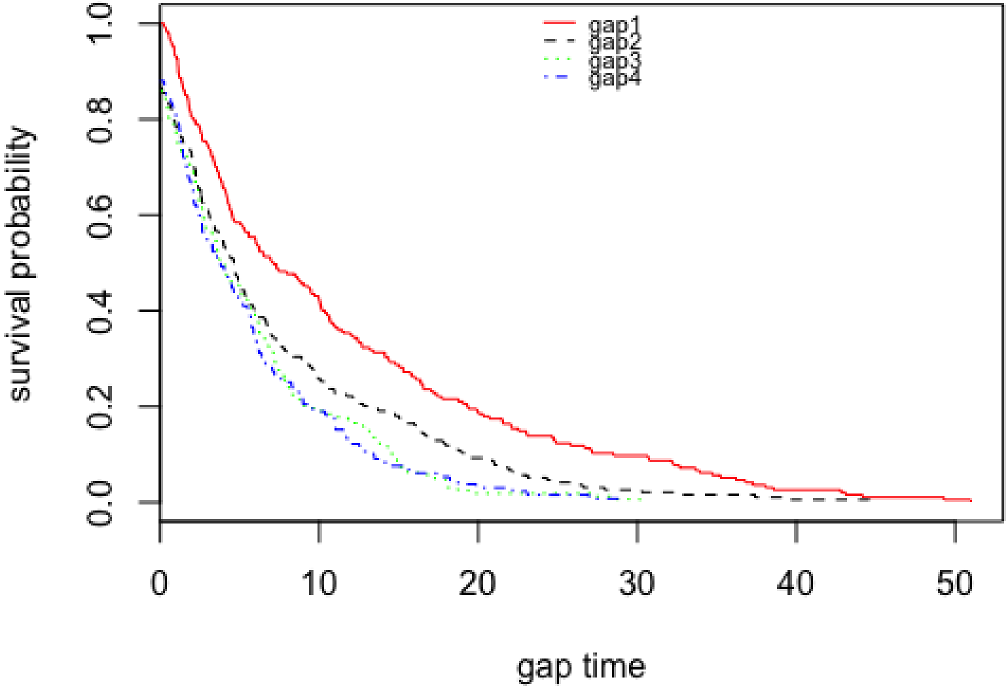
- simulation:

a random sample of 200 individuals(i.e. $i = 1,2, \dots 200$).

A common censoring time, $C_i = 52$

For individual i, log gap times Y_{ij} =log (W_{ij}) (j=1,2,...) follows $MN(2, \ \Sigma \quad)$,

$$\Sigma = \begin{pmatrix} 2 & 1 & \dots & 1 \\ 1 & 2 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & & 2 \end{pmatrix}$$



- AFT and Cox model are under the assumption that conditional on covariates $x_{ij}, W_{ij}|x_{ij}$ are independent.

The **independence checking** can be approached by:

1. informal checking: look at graphs and there should be no trend; empirical distributions for different gaps

2. fitting models that include renewal processes as a special case => extensions of renewal processes are needed

- Extension of renewal models

We need to extend the renewal models because the independent assumption of gap times is untenable in most case. Beside, under the framework log-likelihood ratio test, we can do model check of the independent assumption by comparing complicate models with simpler ones.

Here we introduce three types of common extension models:

1. Random effects models
2. Conditional models
3. Joint gap time distributions approach

- Random effect model:

2.1 Marginal approach:

Assume that given a random effect $u_i \sim G(u_i)$, W_{ij} $(j = 1,2,3,\ldots)$ for individual i are independent.

Then the likelihood function is of the form:

$$L = \prod_{i=1}^{m} \int_{-\infty}^{\infty} \prod_{j=1}^{n_i} f(w_{ij}|x_{ij}, u_i) \cdot S(w_{i,n_i+1}|x_{ij}, u_i) dG(u_i) \,, \qquad (11)$$

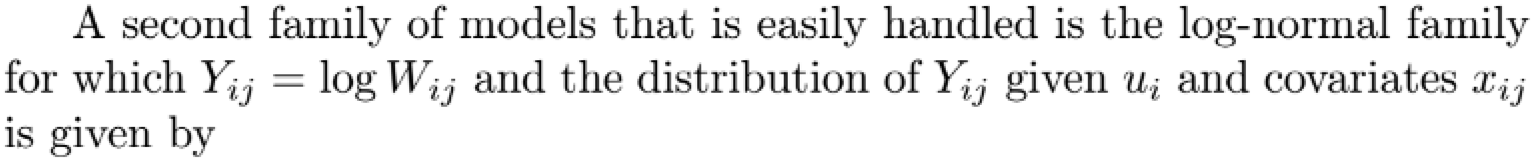
This model would be very useful when the gap times arising from a particular individual are independent, but unobservable factors create heterogeneity in the gap time distribution across individuals.

- 2.2 proportional hazard frailty models

However, obtaining an estimate of the marginal distribution for W_{ij} given x_{ij} is not easy in general. The resulting model will not, in general, have a multiplicative form for the covariate effects.

$$h_{ij}(w|u_i) = u_i h_{0j}(w) \exp(x'_{ij}\beta_j) \qquad j = 1, 2, \ldots . \tag{12}$$

- 2.3



A second family of models that is easily handled is the log-normal family for which $Y_{ij} = \log W_{ij}$ and the distribution of Y_{ij} given u_i and covariates x_{ij} is given by

$$Y_{ij} = \beta_{0j} + x'_{ij}\beta_j + u_i + \epsilon_{ij} \ , \tag{13}$$

where the $\epsilon_{ij} \sim i.i.d. N(0, \sigma^2)$,the $u_i \sim$i.i.d. $N(0, \sigma_{\mathrm{u}}^2)$.

For this model $var(Y_{ij}) = \sigma^2 + \sigma_{\mathrm{u}}^2$, $cov(Y_{ij}, Y_{ik}) = \frac{\sigma_{\mathrm{u}}^2}{\sigma^2 + \sigma_{\mathrm{u}}^2}$

- Conditional model (conditional on event history)

2.2 AFT model:

Notation:

$w_i^{(j-1)} = (w_{i1}, \dots, w_{i,j-1})'$

x_{ij} is a vector of covariates associated with the j^{th} gap time for individual i

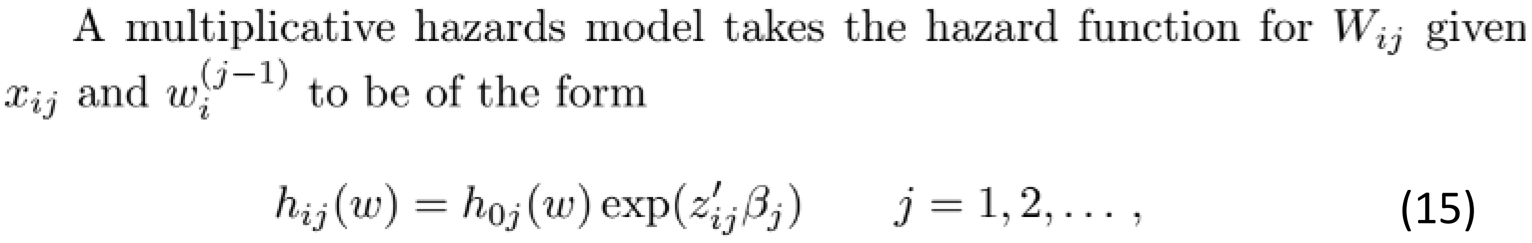
z_{ij} is a vector that models dependence of W_{ij} on x_{ij} and $w_i^{(j-1)}$

Analogous to (4), and define $Y_{ij} = \log W_{ij}$, THE AFT model is of the form:

$$Y_{ij} = \beta_{0j} + z'_{ij}\beta_j + \sigma_j \epsilon_{ij} \qquad j = 1, 2, \dots , \tag{14}$$

Where $\epsilon_{1j}, \epsilon_{2j, \dots,} \epsilon_{mj}$ are i. i. d random variables follow $\mathrm{G}_j(\epsilon)$
$\mathrm{G}_j(\epsilon)$ is fully specified, common used ones are standard normal, logistic, or extreme value distribution

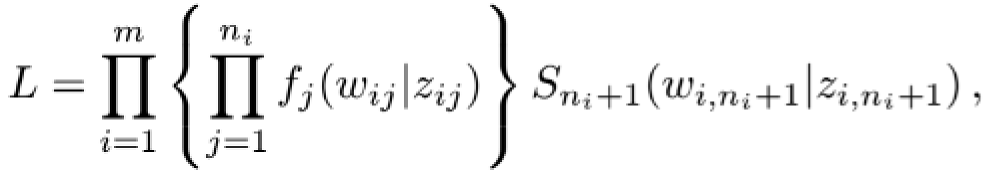
- 2.3 Multiplicative hazards model:



A multiplicative hazards model takes the hazard function for W_{ij} given x_{ij} and $w_i^{(j-1)}$ to be of the form

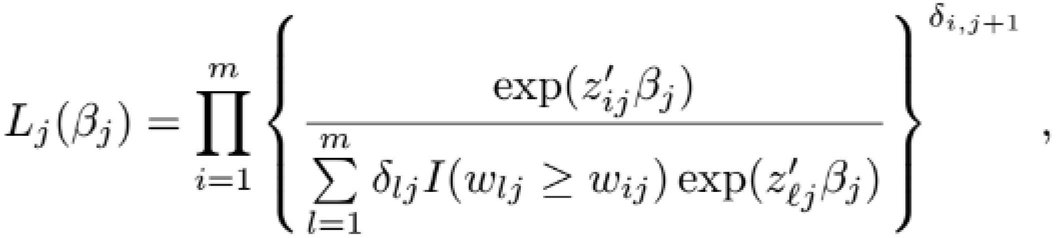
$$h_{ij}(w) = h_{0j}(w)\exp(z'_{ij}\beta_j) \qquad j = 1, 2, \ldots\ , \tag{15}$$

For parametric models the likelihood function from a set of m independent processes is an extension of (3):



$$L = \prod_{i=1}^{m}\left\{\prod_{j=1}^{n_i} f_j(w_{ij}|z_{ij})\right\} S_{n_i+1}(w_{i,n_i+1}|z_{i,n_i+1})\ ,$$

In case that the h_{0j}, β_j are distinct in (12) are distinct, $\widehat{\beta_j}$ is obtained by maximize:



$$L_j(\beta_j) = \prod_{i=1}^{m}\left\{\frac{\exp(z'_{ij}\beta_j)}{\sum_{l=1}^{m}\delta_{lj}I(w_{lj} \geq w_{ij})\exp(z'_{\ell j}\beta_j)}\right\}^{\delta_{i,j+1}}\ , \tag{16}$$

$\delta_{ij} = I(individual\ i\ experienced\ a\ (j-1)st\ event)$

- 3. Join gap time distributions

The random effect models might not obtain marginal distributions for gaps which are of a simple form.

A special case is the multivariate normal model,

The distribution of Y_{ij} given ${y_i}^{(j-1)} = (y_{i1}, \dots, y_{i,j-1})'$ is normal :

$$E(Y_{ij}|x_i, y_i^{(j-1)}) = x'_{ij}\beta_j + \Sigma_{j,j-1}\Sigma_{j-1}^{-1}(y_i^{(j-1)} - \mu_i^{(j-1)}) \qquad (17)$$

and variance $\sigma_j^2 - \Sigma_{j,j-1}\Sigma_{j-1}^{-1}\Sigma'_{j,j-1}$, where $\mu_i^{(j-1)} = (x'_{i1}\beta_1, \dots, x'_{i,j-1}\beta_{j-1})'$ and Σ_j is partitioned as

$$\Sigma_j = \begin{pmatrix} \Sigma_{j-1} & \Sigma_{j-1,j} \\ \Sigma_{j,j-1} & \sigma_j^2 \end{pmatrix} . \qquad (18)$$

- Example: Pulmonary Exacerbations and rhDNase Treatment

Therneau and Hamilton(1997) discussed data that arose in a clinical trail involving persons with cystic fibrosis. Subjects in the study were randomly assigned to receive either a daily dose of the experimental treatment rhDNase or a daily dose of a placebo. The study was double blind, and most subjects were followed for approximately 169 days.

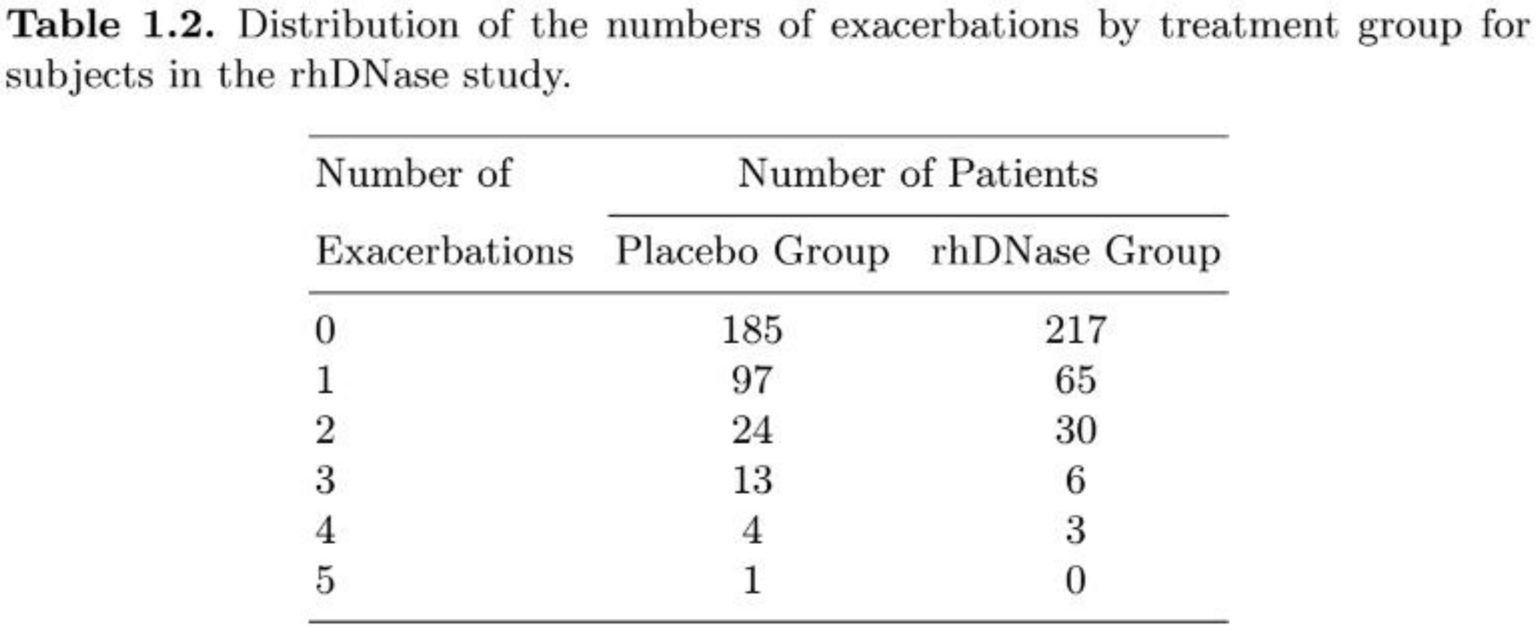


Table 1.2. Distribution of the numbers of exacerbations by treatment group for subjects in the rhDNase study.

Number of	Number of Patients	
Exacerbations	Placebo Group	rhDNase Group
0	185	217
1	97	65
2	24	30
3	13	6
4	4	3
5	1	0

- Data structure:

Id: the patient ID number

trt: 1 for groups receiving rhDNase; 0 for placebo group

fev: the forced respiratory volume

time1: start of a period (indicating at risk)

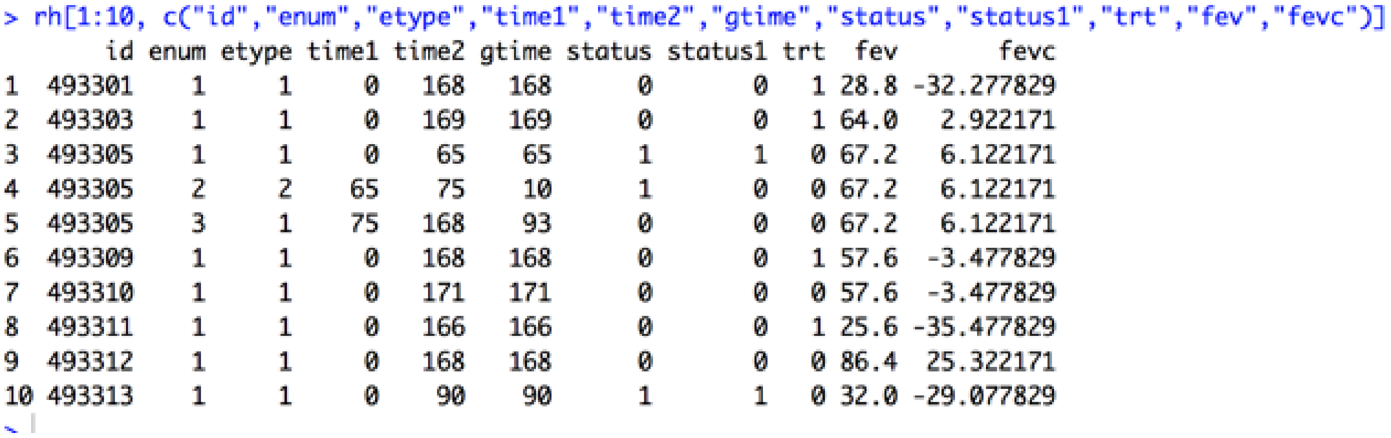
time2: event time or censoring time

status: if time2 is event time, status=1; if time2 is censoring time, status=0

enum: cumulative number of lines in the data frame for each individual

enum1: the cumulative number of exacerbation-free periods for each individual

Part of the data frame:



```
> rh[1:10, c("id","enum","etype","time1","time2","gtime","status","status1","trt","fev","fevc")]
        id enum etype time1 time2 gtime status status1 trt  fev       fevc
1  493301    1     1     0   168   168      0       0   1 28.8 -32.277829
2  493303    1     1     0   169   169      0       0   1 64.0   2.922171
3  493305    1     1     0    65    65      1       1   0 67.2   6.122171
4  493305    2     2    65    75    10      1       0   0 67.2   6.122171
5  493305    3     1    75   168    93      0       0   0 67.2   6.122171
6  493309    1     1     0   168   168      0       0   1 57.6  -3.477829
7  493310    1     1     0   171   171      0       0   0 57.6  -3.477829
8  493311    1     1     0   166   166      0       0   1 25.6 -35.477829
9  493312    1     1     0   168   168      0       0   0 86.4  25.322171
10 493313    1     1     0    90    90      1       1   0 32.0 -29.077829
> 
```

Because independent gap times are so strong assumption, it might not hold, extensions of renewal models are more preferred.

(i) Cox model in the form of (12):

$$h_{ij}(w|u_i) = u_i h_{0j}(w) \exp(x'_{ij}\beta_j) \qquad j = 1, 2, \ldots .$$

(i) Log-normal AFT model in the form of (14):

$$Y_{ij} = \beta_{0j} + z'_{ij}\beta_j + \sigma_j \epsilon_{ij} \qquad j = 1, 2, \ldots ,$$

(i) Joint distribution model in the form of (17):

$$E(Y_{ij}|x_i, y_i^{(j-1)}) = x'_{ij}\beta_j + \Sigma_{j,j-1}\Sigma_{j-1}^{-1}(y_i^{(j-1)} - \mu_i^{(j-1)})$$

Because relatively few persons experienced two or more exacerbations, we consider only the first two gap times.

Let

W_{i1}: the times to the first exacerbation

W_{i2}: the gap times between the first and second exacerbations

x_{i1}: $I(individual\ i\ received\ rhDNase)$

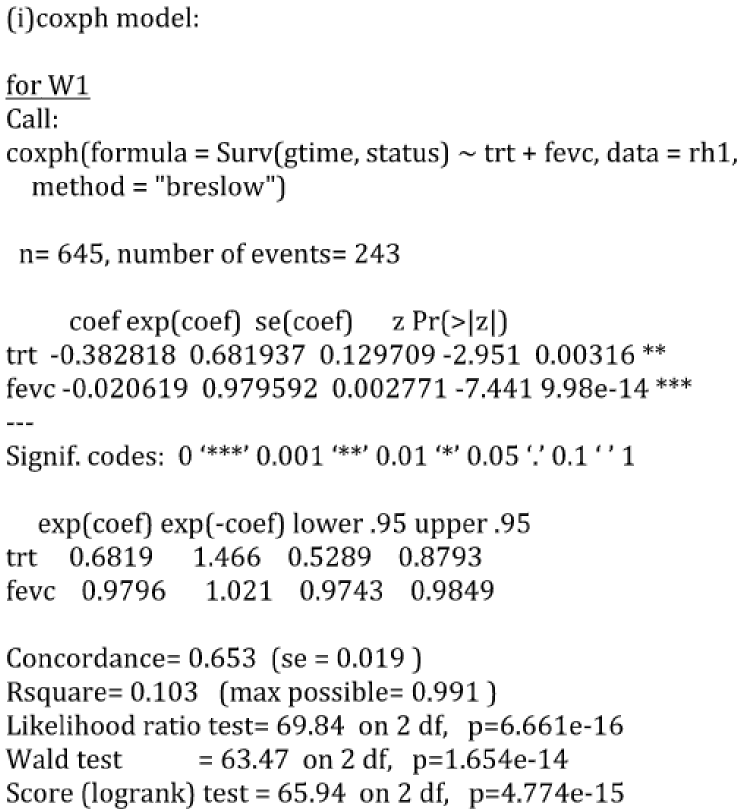
x_{i2}: individual $i's$ forced expiratory volume (fev), a measure of a person’s lung function, measured at the time of randomization. In the analysis below, x_{i2} is centered fev, obtained subtracting the mean fev value across all individuals.

The three types of model would be:

(i) Cox model:

$$h_{ij}(w) = h_{0j}(w)\exp\left(\beta_{j1}x_{i1} + \beta_{j2}x_{i2} + \beta_{j3}\mathrm{w}_{i1}I(\mathrm{j} = 2)\right) \qquad \mathrm{j} = 1{,}2$$

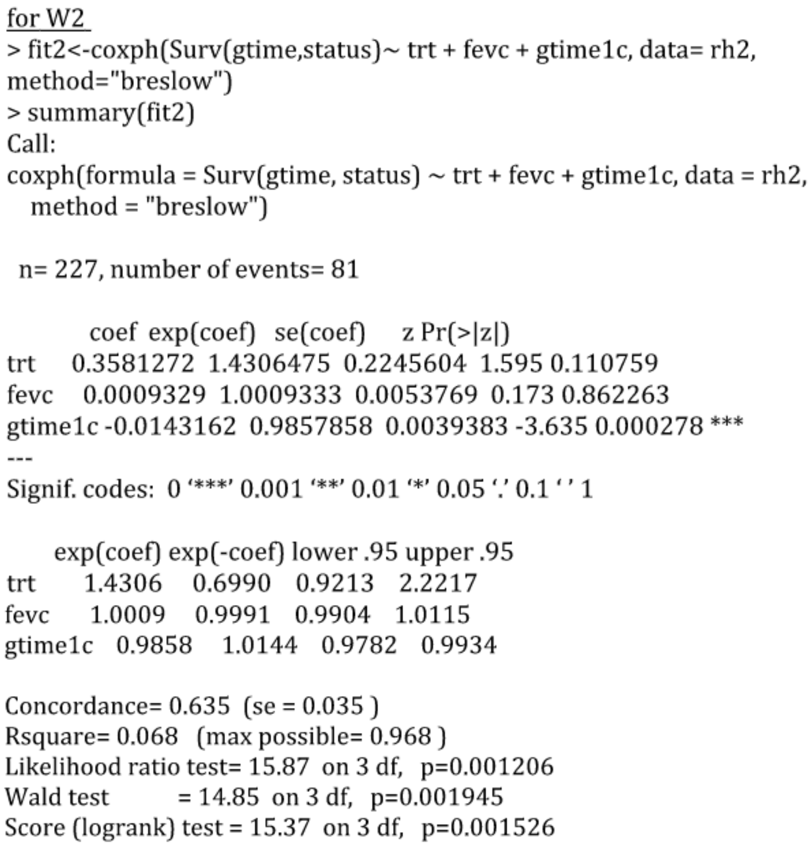
Results:



(i)coxph model:

for W1

```
Call:
coxph(formula = Surv(gtime, status) ~ trt + fevc, data = rh1,
   method = "breslow")

  n= 645, number of events= 243

        coef exp(coef)  se(coef)      z Pr(>|z|)
trt  -0.382818  0.681937  0.129709 -2.951  0.00316 **
fevc -0.020619  0.979592  0.002771 -7.441 9.98e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

    exp(coef) exp(-coef) lower .95 upper .95
trt    0.6819     1.466    0.5289    0.8793
fevc   0.9796     1.021    0.9743    0.9849

Concordance= 0.653  (se = 0.019 )
Rsquare= 0.103   (max possible= 0.991 )
Likelihood ratio test= 69.84  on 2 df,   p=6.661e-16
Wald test            = 63.47  on 2 df,   p=1.654e-14
Score (logrank) test = 65.94  on 2 df,   p=4.774e-15
```



for W2

```
> fit2<-coxph(Surv(gtime,status)~ trt + fevc + gtime1c, data= rh2,
method="breslow")
> summary(fit2)
Call:
coxph(formula = Surv(gtime, status) ~ trt + fevc + gtime1c, data = rh2,
   method = "breslow")

  n= 227, number of events= 81

           coef  exp(coef)   se(coef)      z Pr(>|z|)
trt      0.3581272  1.4306475  0.2245604  1.595 0.110759
fevc     0.0009329  1.0009333  0.0053769  0.173 0.862263
gtime1c -0.0143162  0.9857858  0.0039383 -3.635 0.000278 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

      exp(coef) exp(-coef) lower .95 upper .95
trt       1.4306     0.6990    0.9213    2.2217
fevc      1.0009     0.9991    0.9904    1.0115
gtime1c   0.9858     1.0144    0.9782    0.9934

Concordance= 0.635  (se = 0.035 )
Rsquare= 0.068   (max possible= 0.968 )
Likelihood ratio test= 15.87  on 3 df,   p=0.001206
Wald test            = 14.85  on 3 df,   p=0.001945
Score (logrank) test = 15.37  on 3 df,   p=0.001526
```

- (ii)log-normal AFT model:

$$Y_{ij} = \beta_{j0} + \beta_{j1}x_{i1} + \beta_{j2}x_{i2} + \beta_{j3}\log(w_{i1})\, I(j=2) + \sigma_j\epsilon_{ij}$$

Where $j = 1{,}2$ $\epsilon_{ij} \sim N(0{,}1)$

- (iii) joint distribution model

Considering only the first two gap times, a bivariate normal model for $(Y_{i1}, Y_{i2}) = (\log W_{i1}, log\, W_{i2})$, given x_{i1}(trt) and x_{i2}(fev) .

The model would be:

$$E(Y_{i1}|x_{i1}, x_{i2}) = \beta_{10} + \beta_{11}x_{i1} + \beta_{12}x_{i2}$$
$$E(Y_{i2}|x_{i1}, x_{i2}) = \gamma_{20} + \gamma_{21}x_{i1} + \gamma_{22}x_{i2}$$
$$Var(Y_{i1}|x_{i1}, x_{i2}) = \sigma_1^2,\ Var(Y_{i2}|x_{i1}, x_{i2}) = \sigma_{2m}^2,$$
$$\text{cov}(Y_{i1}, Y_{i2}|x_{i1}, x_{i2}) = \rho\sigma_1\sigma_{2m}$$

(ii)lognormal AFT model:

for W1

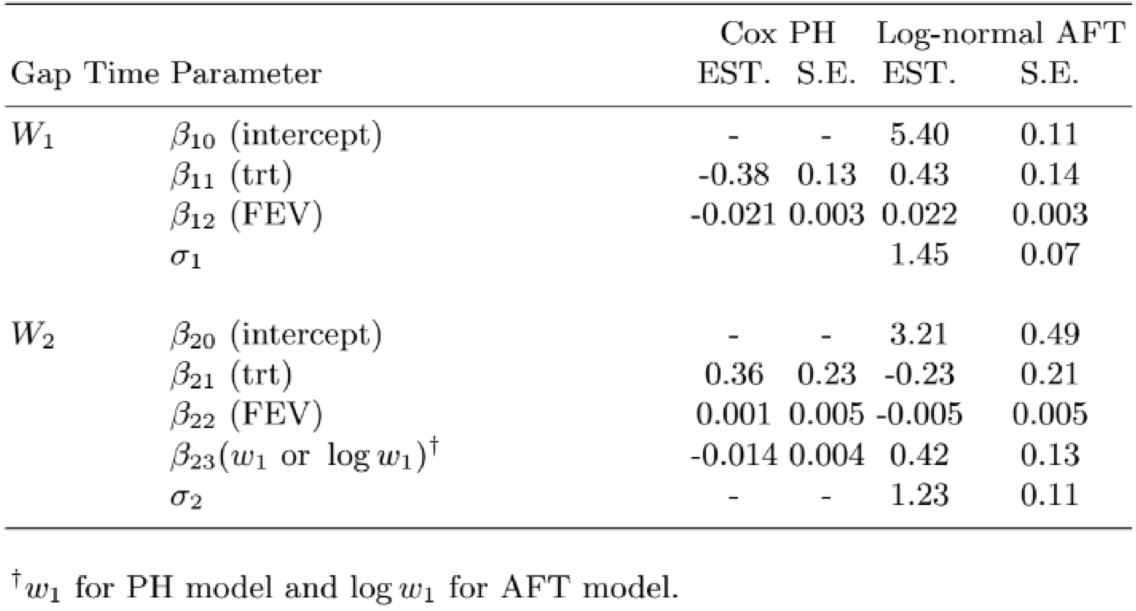
```
> fitlog1<- survreg(Surv(gtime,status)~ trt + fevc, data= rh1, dist="lognormal")
> summary(fitlog1)

Call:
survreg(formula = Surv(gtime, status) ~ trt + fevc, data = rh1,
    dist = "lognormal")
             Value Std. Error     z       p
(Intercept) 5.4030     0.1048 51.53 0.00e+00
trt         0.4302     0.1371  3.14 1.71e-03
fevc        0.0217     0.0029  7.47 8.03e-14
Log(scale)  0.3688     0.0512  7.21 5.68e-13

Scale= 1.45

Log Normal distribution
Loglik(model)= -1625   Loglik(intercept only)= -1660.8
        Chisq= 71.51 on 2 degrees of freedom, p= 3.3e-16
Number of Newton-Raphson Iterations: 4
n= 645
```

for W2 (use logw1)

```
> fitlog2<-survreg(Surv(gtime,status)~ trt + fevc + lgtime1, data= rh2, dist =
"lognormal")
> summary(fitlog2)

Call:
survreg(formula = Surv(gtime, status) ~ trt + fevc + lgtime1,
    data = rh2, dist = "lognormal")
              Value Std. Error      z        p
(Intercept)  3.20898     0.4867  6.593 4.30e-11
trt         -0.22657     0.2105 -1.077 2.82e-01
fevc        -0.00454     0.0047 -0.965 3.34e-01
lgtime1      0.41730     0.1322  3.157 1.59e-03
Log(scale)   0.20559     0.0859  2.392 1.67e-02

Scale= 1.23

Log Normal distribution
Loglik(model)= -484.9   Loglik(intercept only)= -490.9
        Chisq= 11.99 on 3 degrees of freedom, p= 0.0074
Number of Newton-Raphson Iterations: 4
n= 227
```

- Fitting summary:

Fitted models for W1 and for W2 given W1.



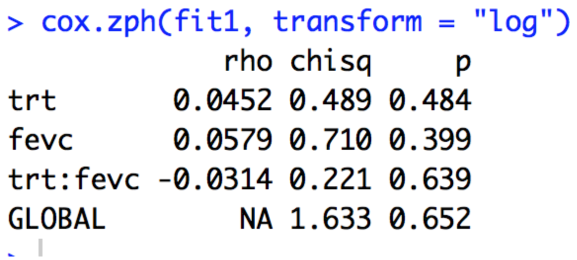
Gap Time	Parameter	Cox PH EST.	Cox PH S.E.	Log-normal AFT EST.	Log-normal AFT S.E.
W_1	β_{10} (intercept)	-	-	5.40	0.11
	β_{11} (trt)	-0.38	0.13	0.43	0.14
	β_{12} (FEV)	-0.021	0.003	0.022	0.003
	σ_1			1.45	0.07
W_2	β_{20} (intercept)	-	-	3.21	0.49
	β_{21} (trt)	0.36	0.23	-0.23	0.21
	β_{22} (FEV)	0.001	0.005	-0.005	0.005
	$\beta_{23}(w_1 \text{ or } \log w_1)^\dagger$	-0.014	0.004	0.42	0.13
	σ_2	-	-	1.23	0.11

$^\dagger w_1$ for PH model and $\log w_1$ for AFT model.

The results for W_1 indicate a strong positive treatment effect after adjustment for fev. $w_{i1} is$ highly significant in connection with W_2, which indicates a strong positive association between $W_{i1} and\ W_{i2}$ even after adjustment for treatment and fev.
It is hard to separate clearly the effect of w_1, treatment and fev on W_2 since W_2 is limitedly observed and has a strong association with W_{i1}. From the table we can see that neither treatment nor fev is significant for W_2, after adjustment for W_1

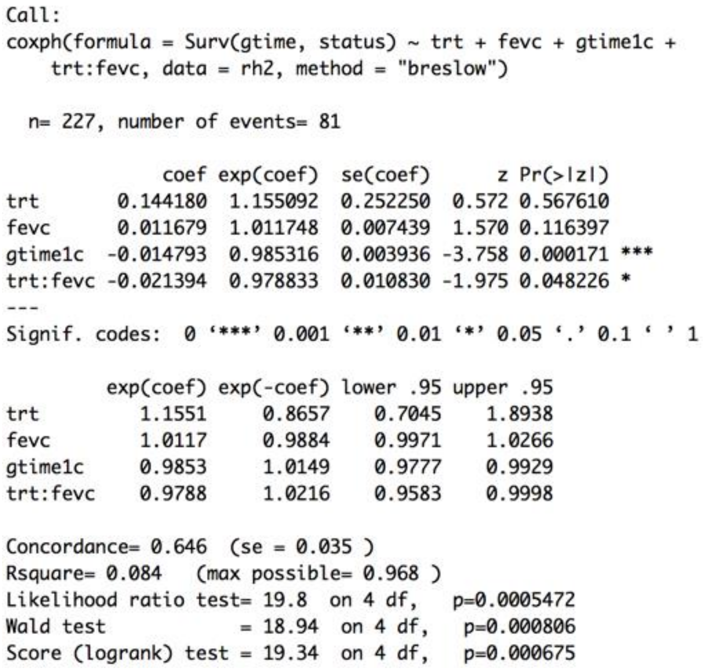
- Checks for trt by fev interactions did not provide evidence of an interaction.

for cox model, apply cox.zph(), which does not provide evidence of an interaction



```
> cox.zph(fit1, transform = "log")
             rho chisq     p
trt       0.0452 0.489 0.484
fevc      0.0579 0.710 0.399
trt:fevc -0.0314 0.221 0.639
GLOBAL        NA 1.633 0.652
```

for log-normal AFT model, no strong evidence of an interaction



```
Call:
coxph(formula = Surv(gtime, status) ~ trt + fevc + gtime1c +
    trt:fevc, data = rh2, method = "breslow")

  n= 227, number of events= 81

              coef exp(coef)  se(coef)      z Pr(>|z|)
trt       0.144180  1.155092  0.252250  0.572 0.567610
fevc      0.011679  1.011748  0.007439  1.570 0.116397
gtime1c  -0.014793  0.985316  0.003936 -3.758 0.000171 ***
trt:fevc -0.021394  0.978833  0.010830 -1.975 0.048226 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

         exp(coef) exp(-coef) lower .95 upper .95
trt         1.1551     0.8657    0.7045    1.8938
fevc        1.0117     0.9884    0.9971    1.0266
gtime1c     0.9853     1.0149    0.9777    0.9929
trt:fevc    0.9788     1.0216    0.9583    0.9998

Concordance= 0.646  (se = 0.035 )
Rsquare= 0.084   (max possible= 0.968 )
Likelihood ratio test= 19.8  on 4 df,   p=0.0005472
Wald test            = 18.94  on 4 df,   p=0.000806
Score (logrank) test = 19.34  on 4 df,   p=0.000675
```

- The analysis above showed a general problem:

in assessing the effects of fixed baseline covariates on gap times, if the gap times are not independent (after conditional on the baseline covariates), then the **effects of covariates on second and subsequent gap times are confounded with the effects of prior gap times.**

- We might want to get the marginal distribution of W_{i2} given $x_{i1}$and x_{i2}.

Another approach is considering the joint distributions for W_1 , W_2 , ….

From model (iii), the marginal distribution for Y_{i2} (parameters were estimated by mle)

	MLE	$s.e.$
$\hat{\gamma}_{20}$(intercept)	5.40	0.30
$\hat{\gamma}_{21}$ (trt)	-0.048	0.215
$\hat{\gamma}_{22}$ (fev)	0.0045	0.0053
$var(Y_{i2}\|x_{i1}, x_{i2})$	1.37	0.15

Note: neither trt nor fev is significant on W_2 , which is different than their effects on W_1 . A qualification of this is that less than half of the subjects in the study experienced even a first exacerbation, and it would be misleading to compute the marginal distribution of W_2 . Because to find the marginal distribution of W_2, It is implicitly assumed that everyone eventually experienced a first exacerbation.

- There are other models, for example the proportional hazards frailty model with a gamma random effect, and we would obtain similar result to the bivariate normal models.

Limitation of gap time analysis based on previous method

- The association between the gaps times for an individual makes it difficult to address questions concerning persistence of trends in treatment effects.
- Nonparametric or semiparametric estimator is also difficult with all approaches

Another approach to the marginal distribution:

Inspired by the idea inverse probability of censoring weights (IPCW), here we introduce the following approach, which is used by Lin et.al(1999) and developed further by Van der Laan et al(2002)

- By Lin et al(1999)

Consider function:

$$H(w_1, w_2) = \Pr(W_{i1} \leq w_1, W_{i2} > w_2)\,, \qquad (18)$$

$H(w_1, w_2)$ is estimable only for values in a dataset that satisfy $w_1 + w_2 \leq C_{max}$, where $C_{max} = \max(C_1, \dots, C_m)$

Assumption:

(a)Censoring times C_i is completely independent of event process $\{N_i(t), 0 \leq t\}$

(b)$G(c) = \Pr(C_i > c)$ i.e. survival function for C_i

(c) Let C_i represent the censoring time for individual i, so that what we may observe is $\tilde{T}_{i1} = \min(W_{i1}, C_i)$, $\tilde{T}_{i2} = \min(W_{i1} + W_{i2}, C_i)$, $\delta_{i1} = I(W_{i1} \leq C_i)$, and $\delta_{i2} = I(W_{i1} + W_{i2} \leq C_i)$, with $(t_{i1}, t_{i2}, \delta_{i1}, \delta_{i2})$, $i = 1, \dots, m$

The estimate for $H(w_1, w_2)$ of Lin et al. (1999) is based on the observation that

$$E\left\{\frac{I(W_{i1} \leq w_1, W_{i2} > w_2) I(C_i > W_{i1} + w_2)}{G(W_{i1} + w_2)}\right\} = H(w_1, w_2)\,, \qquad (19)$$

Because C_i is independent of $(W_{i1}, W_{i2}, \dots)$

$$E\{I(C_i > W_{i1} + w_2 | N_i^{(\infty)})\} = G(W_{i1} + w_2)$$

If we use an estimate $\widehat{G}(c)$ of $G(c)$, then we have:

$$\widehat{H}(w_1, w_2) = \frac{1}{m} \sum_{i=1}^{m} \frac{I(w_{i1} \leq w_1, w_{i2} > w_2, C_i > w_{i1} + w_2)}{\widehat{G}(w_{i1} + w_2)} . \qquad (20)$$

This can be expressed in the equivalent form

$$\widehat{H}(w_1, w_2) = \frac{1}{m} \sum_{i=1}^{m} \frac{I(\tilde{t}_{i1} \leq w_1, \tilde{t}_{i2} - \tilde{t}_{i1} > w_2)}{\widehat{G}(\tilde{t}_{i1} + w_2)} , \qquad (21)$$

where we note that for $w_1 > 0, w_2 > 0$, the condition $\tilde{t}_{i2} - \tilde{t}_{i1} > w_2$ implies that $\tilde{t}_{i1} = w_{i1}$ and $\delta_{i1} = 1$. Assuming that $\widehat{G}(c)$ is a consistent estimator of $G(c)$, (20) is a consistent estimator of $H(w_1, w_2)$.

The estimate $\widehat{G}(c)$ can be the empirical survivor function, if all of $C_1, \ldots,$ C_n are observed. or a Kaplan–Meier estimate

then the marginal distribution $F_1(w) = H(w_1, 0)$ is,

$$\widehat{F}_1(w_1) = \frac{1}{m} \sum_{i=1}^{m} \delta_{i1} \frac{I(\tilde{t}_{i1} \leq w_1)}{\widehat{G}(\tilde{t}_{i1})} . \qquad (22)$$

for W_2 we can consider estimated probabilities

$$\widehat{\Pr}(W_2 > w_2 | W_1 \le w_1) = \frac{\widehat{H}(w_1, w_2)}{\widehat{H}(w_1, 0)}\,, \qquad (23)$$

for (w_1, w_2), where $w_1 + w_2 < C_{\max}$.

with $\widehat{\text{var}}\{\widehat{\Pr}(W_2 > w_2 | W_1 \le w_1)\}$ given by

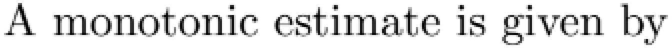
$$\frac{1}{m^2 \widehat{H}(w_1,0)^2} \sum_{i=1}^{m} \left\{ \delta_{i1} I(\tilde{t}_i \le w_1) \left[\frac{\widehat{H}(w_2|w_1)}{\widehat{G}(\tilde{t}_{i1})} - \frac{I(\tilde{t}_{i2} - \tilde{t}_{i1} > w_2)}{\widehat{G}(\tilde{t}_{i1} + w_2)} \right]^2 \right.$$

$$\left. - \frac{m^2 (1 - \delta_{i2}) \widehat{B}(w_1, w_2; \tilde{t}_{i2})^2}{\left[1 + \sum_{j=1}^{m} I(\tilde{t}_{j2} > \tilde{t}_{i2})\right] \sum_{j=1}^{m} I(\tilde{t}_{j2} > \tilde{t}_{i2})} \right\},$$

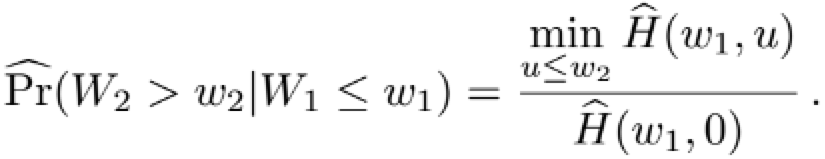
where $\widehat{H}(w_2|w_1) = \widehat{H}(w_1, w_2)/\widehat{H}(w_1, 0)$ and

$$\widehat{B}(w_1, w_2; u) = \widehat{H}(w_2|w_1)\{\widehat{H}(w_1, 0) - \widehat{H}(u, 0)\}^+ - \{\widehat{H}(w_1, w_2) - \widehat{H}(u - w_2, w_1)\}^+$$

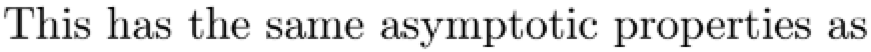
The estimator (23) is not necessarily strictly monotonic in w_2



A monotonic estimate is given by



$$\widehat{\Pr}(W_2 > w_2 | W_1 \leq w_1) = \frac{\min\limits_{u \leq w_2} \widehat{H}(w_1, u)}{\widehat{H}(w_1, 0)} . \qquad (24)$$



This has the same asymptotic properties as (23)

This model put forward by Lin et al. (1999) has a strict assumption that censoring times C_i is independent of event process $\{N_i(t), 0 \leq t\}$.
Van der Lann et al.(2002) relaxed the assumption by allowing censoring to depend on prior event history or on previously observed covariates.

- The frame work of Van der Laan et al.(2002):

Let :

$B_i = g(W_{i1}, \ldots, W_{ik})$ be a function of some number of gap times
$\theta = E(B_i)$ is a parameter of our interests

$\Delta_i = I(B_i \text{ is observed})$

$V_i = \min\{t : \Delta_i = 1\}$

(the earliest time in the process $\{N_i(t), 0 \le t\}$ at which B_i can be observed.)

e.g: if $B_i = I(W_{i1} \le w_1, W_{i2} > w_2)$ for given values $w_1 > 0, w_2 > 0$ then

$\theta = H(w_1, w_2)$

key idea is based on having a model for

$$E(\Delta_i | N_i^{(\infty)}) = \Pr(C_i > V_i | N_i^{(\infty)})\,, \qquad (25)$$

where $N_i^{(\infty)} = \{N_i(t), 0 \le t\}$

- Analogous to (19), we would have

$$E\left\{\frac{\Delta_i B_i}{E(\Delta_i|N_i^{(\infty)})}\right\} = E(B_i) = \theta\,. \qquad (26)$$

This motivates the estimator

$$\widehat{\theta} = \frac{1}{m}\sum_{i=1}^{m}\frac{\Delta_i B_i}{\widehat{E}(\Delta_i|N_i^{(\infty)})}\,. \qquad (27)$$

If censoring times C_i is completely independent of event process $\{N_i(t), 0 \leq t\}$, then $\hat{E}(\Delta_i|N^{\infty}) = \hat{E}(\Delta_i)$ could be get from observed data. This is the case by Lin(1999).

If censoring times C_i is not completely independent of event process $\{N_i(t), 0 \leq t\}$, but only depend on the past observations. That is , we can allow the hazard function for C_i at process time t to be:

$$\lambda_c(t|N^{(\infty)}) = \lambda_c(t|N^{(t^-)})\,, \qquad (28)$$

Then we could estimate (26) by (24),

$$E(\Delta_i|N_i^{(\infty)}) = \Pr(C_i > V_i|N_i^{(\infty)})\,,$$

- e.g.

Consider the hazard function of censoring times C_i is a cox model:

$$\lambda_c(t) = \lambda_0(t)\exp\{\beta I(N_i(t^-) > 0)\} \tag{29}$$

Then,

$$\begin{aligned} \Pr(C_i > V_i | N_i^{(\infty)}) &= \Pr(C_i > w_{i1} + w_2 | N_i^{(\infty)}) \\ &= \exp\{-\Lambda_0(w_{i1}) + e^{\beta}[\Lambda_0(w_{i1} + w_2) - \Lambda_0(w_{i1})]\}\,, \end{aligned} \tag{30}$$

where $\Lambda_0(t) = \int_0^t \lambda_0(u)du$

- The framework above can also be used for any subsequent gap time W_j:

Treating W_j as W_2, and $T_{j-1} = W_1$+…+W_{j-1} as W_1. To estimate $S_j(w) = \Pr(W_j > w)$

by estimating $H_j(t, w) = \Pr(T_{j-1} \le t, W_j > w)$

for pairs (t, w) that satisfy $t + w < C_{\max}$

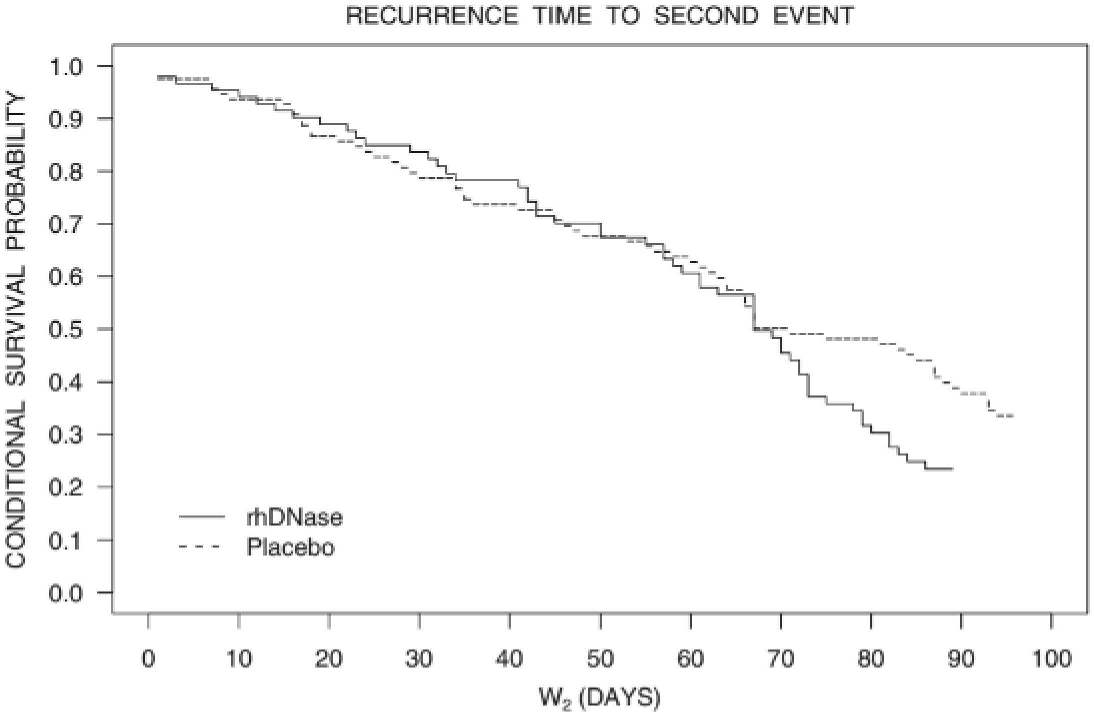
Note: The variance estimation of confidence interval would be preferred to be approached by bootstrap

- Exmaple: Pulmonary Exacerbations and rhDNase Treatment

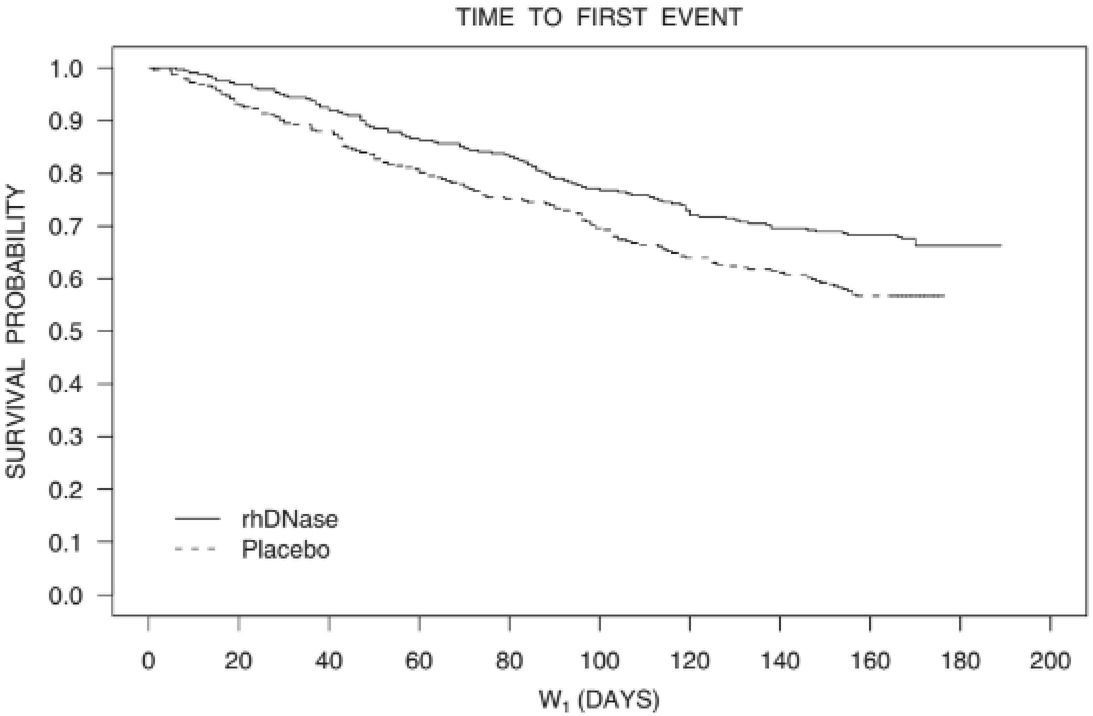
1. Exam the distribution of W_2, given $W_1 \leq 100$, without adjusting for fev

(condition on the first exacerbation occurs before 100 days is because in that way second exacerbations would have more time to be observed; fev is not adjusted since treatment is randomly assigned to subjects and is therefore independent of fev)

The plot shows no significant difference between the rhDNase and placebo groups. The estimates beyond $w_2 = 69$ is imprecise because $P(W_2 > w_2 | W_1 \leq 100)$ is estimable only for $W_2 \leq C_{max} - 100 = 169 - 100 = 69$. The divergence of the two curves does not imply a significant differenc



- The K-M estimator for $S_1(w_1)$ indicates a significant difference between the rhDNase and placebo groups.



- Because we have no observation for $w_1 + w_2 \geq 170\ days,$

  So when using conditional regression model to estimate marginal probabilities, we choose $W_1 \leq 100$

The conditional probability based on the log-normal model introduced before:

($Recall$: $Y_{ij} = \beta_{j0} + \beta_{j1} x_{i1} + \beta_{j2} x_{i2} + \beta_{j3} \log(w_{i1})\, I(j = 2) + \sigma_j \epsilon_{ij}$

Where $j = 1,2$ $\epsilon_{ij} \sim N(0,1)$)

Letting $y_1 = \log w_1$, $y_2 = \log w_2$, and $y_L = \log L$,

$$\Pr(W_2 > w_2 | W_1 \leq L,\ x) = \frac{\int_{-\infty}^{y_L} \bar{F}_N\left(\frac{y_2 - x'\beta_2 - \beta_{23} y_1}{\sigma_2}\right) \sigma_1^{-1} f_N\left(\frac{y_1 - x'\beta_1}{\sigma_1}\right) dy_1}{F_N\left(\frac{y_L - x'\beta_1}{\sigma_1}\right)}, \qquad (31)$$

where $x = (1, x_1, x_2)'$, and $f_N(\cdot)$, $F_N(\cdot)$, $\bar{F}_N(\cdot)$ are, the density, distribution, and survivor functions for the standard normal distribution.

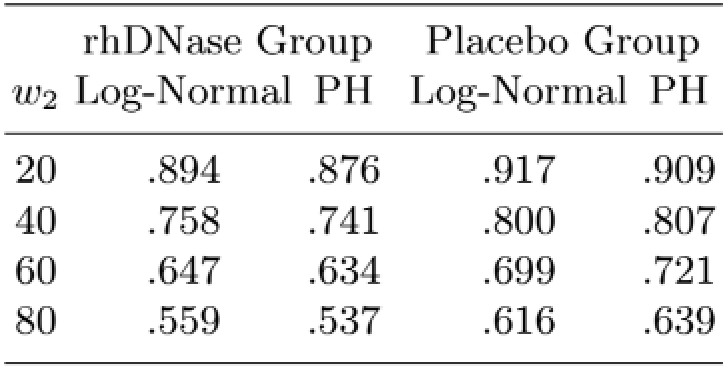
- Similarly, based on cox proportional hazards model, the conditional probability would be:

$$\Pr(W_2 > w_2 | W_1 \leq L, x) =$$

$$\frac{\int_0^L \exp\{-\Lambda_{20}(w_2)e^{x'\beta_2+\beta_{23}w_1}\}\exp\{-\Lambda_{10}(w_1)e^{x'\beta_1}\}e^{x'\beta_1}d\Lambda_{10}(w_1)}{1-\exp\{-\Lambda_{10}(L)e^{x'\beta_1}\}},$$

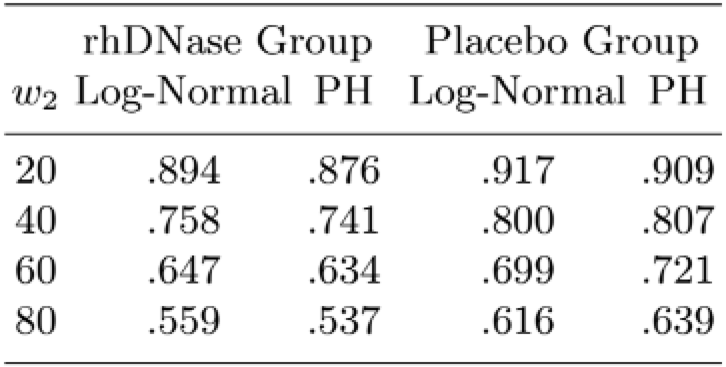
where $\Lambda_{10}(w)$ and $\Lambda_{20}(w)$ are the generalized Nelson–Aalen estimates for the baseline hazard functions

For $L = 100$, and selected values of w_2, for persons with $x_2 = 0$ (average fev), the estimated $\Pr(W_2 > w_2 | W_1 \leq 100, x)$ is shown below:



	rhDNase Group		Placebo Group	
w_2	Log-Normal	PH	Log-Normal	PH
20	.894	.876	.917	.909
40	.758	.741	.800	.807
60	.647	.634	.699	.721
80	.559	.537	.616	.639

- We observe that the estimated conditional survival probability under the two models are similar.
- The survival probabilities are slightly higher on the placebo group, but the difference is not significant.
- For a person with average fev (i.e. $x_{i2} = 0$), the conditional probability are similar to those shown in the conditional survival probability plot of W_2, without adjustment for x_{i2}



	rhDNase Group		Placebo Group	
w_2	Log-Normal	PH	Log-Normal	PH
20	.894	.876	.917	.909
40	.758	.741	.800	.807
60	.647	.634	.699	.721
80	.559	.537	.616	.639

- Further topic:

for observational studies it is the norm that the event process for an individual has started prior to her selection for a study. In this type of situation we frequently do not know the precise time at which the process started (left truncated).

We could treat the initial conditions differently:

- On way is to ignore the incomplete first gap and treat the individual's follow up as starting from t_1 (first event time) as far as estimation of $S(w)$ is censored.
- The other option might be to assume that the times of events occurring before t=0 follows a certain distribution.

- Reference:

1. Richard J.Cook, Jerald F. Lawless, *The Statistical Analysis of Recurrent Events*. Springer press.
2. http://www.math.uwaterloo.ca/~rjcook/book_code.html