

生存分析：竞争风险模型（Fine & Gray 模型）

当研究对象面临多种相互排斥事件的风险时即存在竞争风险（Gichangi & Vach 2005），并且其中一个事件的发生将阻止任何其他事件发生。例如来自不同原因的死亡。

当存在竞争风险时，在分析某一种事件的风险的时候，传统的将其它事件的发生视作丢失（ *censor*）计算出来的 Hazard Ratio（危害比）虽然可以解释，但其相应的生存函数则不能被描述为该事件的生存概率。这种方法假定其他事件导致的丢失是独立的，而这一假设没有办法检验的。

相互竞争的风险理论使我们能够计算出患者不仅有所关注的事件发生的风险而且还有任何其他事件发生的“真实世界”概率，提供一种分解事件发生（如不同死因死亡）概率的方法，使研究对象更清楚地了解他们每次做出决定时所面临的风险。

基本概念：

- 1) 特定事件风险函数 $h_k(t)$ ：在时间 t 时仍然存活条件下，发生特定事件 k 的危险。Cause-specific hazard, $h_k(t)$, is the instantaneous risk of dying from a particular cause k given that the subject is still alive at time t . Prentice et al. (1978)。
- 2) 累积发病率(CIF)函数 $C_k(t)$ ：考虑到观察对象可能发生其它事件的情况下，计算的在观察时间 t 累积发生事件 K 的比例（Cumulative incidence function, $C_k(t)$, the proportion of patients at time t who have died from cause k accounting for the fact that patients can die from other causes）。某事件的 CIF 不仅取决于该事件的风险 $h_k(t)$ ，还取决于其它事件的风险，因此该事件的 $h_k(t)$ 与 $C_k(t)$ 之间没有直接关系。危险因素对 CIF 的作用可能与它们对 $h_k(t)$ 作用不再一致。Fine & Gray (1999) 提出直接拟合累积发生函数 CIF 的模型。
- 3) 子事件风险 $h_{ks}(t)$ ：没有发生某事件的研究对象在时间 t 发生该事件的风险（Sub-distribution hazard, $h_{ks}(t)$, is the instantaneous risk of dying from a particular cause k given that the subject has not died from cause k ）。这里没有发生某事件的对象包括已经发生其它事件的对象，不同于特定事件风险 $h_k(t)$ 的定义里的仍然存活的对象。子事件风险和 CIF 之间有直接联系，因此可以用来评估危险因素对 CIF 的效应。

关于 Fine & Gray 模型，请参考文献：“A Proportional Hazards Model for the Subdistribution of a Competing Risk. Jason P. Fine and Robert J. Gray, *Journal of the American Statistical Association* Vol. 94, No. 446 (Jun., 1999), pp. 496-509”。

本模块采用 Fine & Gray 模型拟合竞争风险，模型中可以包括非参数项与参数项，非参数项的效应可以随时间变化，不是固定的；参数项（或称固定项）其效应不随时间变化，是固定的。非参数项输出：该项效应是否显著的检验结果，该项效应是否随时间变化的检验结果。参数项输出该项的效应大小及 95% 可信区间。

本模块输出结果包含 3 个模型，模型 1 是所有变量均为非参数项模型，模型 2 是所有变量均为参数项模型，模型 3 是按输入设置定义的含参数与非参数项模型（如果没有定义非参数项或所有变量都定义为参数项，则不运行模型 III）。

例：下载练习数据 [bmt.xls](http://r.empowerstats.cn/empowerStats/exdata/bmt.xls) (<http://r.empowerstats.cn/empowerStats/exdata/bmt.xls>)，该数据有 408 条记录 5 个变量，其中 cause 表示随访生存状态，0=删失、1=死亡、2=复发；time 表示随访时间；platelet=1 表示血小板计数 $>100 \times 10^9/L$ ，platelet=0 表示血小板计数 $\leq 100 \times 10^9/L$ ；tcell 表示是否 T-cell 耗尽，1=是、0=否；age 表示经标化的病人年龄 $(age-35)/15$ 。拟合 platelet、age、tcell 与 cause 的关系，输入界面如下：

在自变量框中，如选择某自变量（S）表示为非参数项，如本例中的 AGE。

输出结果及解读：

结局变量：Survival status

时间变量：TIME

模型 I：Event (TIME, CAUSE) ~ AGE+TCELL+PLATELET

模型中所有变量的效应均随时间变化 (allowing all covariates to have time-varying effects)

模型 I 分析结局状态 Survival status == Death

检验：	是否有显著效应	效应是否随时间变化	效应是否随时间变化
H0:	B(t)=0 (无效应)	效应不随时间变化 (固定效应)	效应不随时间变化 (固定效应)
Intercept	<0.0001	<0.0001	<0.0001
AGE	<0.0001	<0.0001	<0.0001

T-cell depleted	<0.0001	0.1020	0.0620
Platelet	0.0150	0.0460	0.1110

模型 I 分析结局状态 Survival status == Relapse

检验:	是否有显著效应	效应是否随时间变化	效应是否随时间变化
H0:	B(t)=0(无效应)	效应不随时间变化(固定效应)	效应不随时间变化(固定效应)
Intercept	<0.0001	<0.0001	<0.0001
AGE	0.5160	0.3170	0.4210
T-cell depleted	0.3070	0.0350	0.0500
Platelet	0.0020	0.2450	0.3710

模型 II: Event (TIME, CAUSE) ~ const (AGE)+const (TCELL)+const (PLATELET)

模型中所有变量的效应均随不时间变化(allowing all covariates to have constant effects).

模型 II 分析结局状态 Survival status == Death

固定效应	Sub-distribution HR	95%CI low	95%CI upp	P value	Coef.	Se	Z
AGE	1.3832	1.1554	1.6560	0.0004	0.32442	0.09182	3.53341
T-cell depleted	0.4994	0.2760	0.9035	0.0217	-0.69445	0.30253	-2.29547
Platelet	0.5708	0.3825	0.8518	0.0060	-0.56072	0.20425	-2.74531

模型 II 分析结局状态 Survival status == Relapse

固定效应	Sub-distribution HR	95%CI low	95%CI upp	P value	Coef.	Se	Z
AGE	1.0298	0.8042	1.3187	0.8159	0.02938	0.12617	0.23288
T-cell depleted	1.5168	0.8491	2.7097	0.1593	0.41661	0.29603	1.40735
Platelet	0.9115	0.5681	1.4627	0.7010	-0.09263	0.24127	-0.38392

模型 III: Event (TIME, CAUSE) ~ AGE+const (TCELL)+const (PLATELET)

模型中选择的变量效应随时间变化,其它变量效应不随时间变化(allowing selected covariates to have time-varying effects and others constant effects).

模型 III 分析结局状态 Survival status == Death

效应随时间变化的变量:

检验:	是否有显著效应	效应是否随时间变化 Kolmogorov-Smirnov test	效应是否随时间变化 Cramer von Mises test
H0:	B(t)=0(无效应)	效应不随时间变化(固定效应)	效应不随时间变化(固定效应)

Intercept	<0.0001	<0.0001	<0.0001
AGE	<0.0001	<0.0001	<0.0001

效应不随时间变化的变量:

固定效应	Sub-distribution HR	95%CI low	95%CI upp	P value	Coef.	Se	Z
T-cell depleted	0.4958	0.2766	0.8890	0.0185	-0.70148	0.29788	-2.35488
Platelet	0.5781	0.3886	0.8598	0.0068	-0.54805	0.20257	-2.70548

模型 III 分析结局状态 Survival status == Relapse

效应随时间变化的变量:

检验:	是否有显著效应	效应是否随时间变化 Kolmogorov-Smirnov test	效应是否随时间变化 Cramer von Mises test
H0:	B(t)=0(无效应)	效应不随时间变化(固定效应)	效应不随时间变化(固定效应)
Intercept	<0.0001	<0.0001	<0.0001
AGE	0.6770	0.4430	0.5490

效应不随时间变化的变量:

固定效应	Sub-distribution HR	95%CI low	95%CI upp	P value	Coef.	Se	Z
T-cell depleted	1.5059	0.8428	2.6906	0.1668	0.40937	0.29612	1.38245
Platelet	0.9132	0.5702	1.4623	0.7053	-0.09084	0.24024	-0.37811

输出的.txt 文件详细解读

```
Competing risk regression model (using timereg package: Fine & Gray model)
Outcome: Survival status
Time: TIME
```

模型 I 对每个自变量都采用非参数方法拟合，这是模型表达式

```
Model I: Event(TIME, CAUSE) ~ AGE+TCELL+PLATELET
```

首先拟合结局状态：死亡

```
Model for Survival status == Death
```

下面是拟合结果:

```
Competing risks Model
```

```
Test for nonparametric terms
```

```
Test for non-significant effects
```

```
Supremum-test of significance p-value H_0: B(t)=0
(Intercept) 11.00 0.000
AGE 5.63 0.000
TCELL 4.99 0.000
```

PLATELET

3.23

0.015

上面检验的是基线风险(intercept)是否为0, p 值为0, 非常显著。AGE、TCELL、PLATELET 是否与风险有关, AGE、TCELL、PLATELET 三自变量的 P 值分别为 0.000、0.000、0.015, 均显著。下面是对每个自变量的效应是否随时间有变化进行检验, 分别采用了两种检验方法: Kolmogorov-Smirnov 检验与 Cramer von Mises 检验。结果显示基线风险随时间变化显著, AGE 的效应随时间有变化显著。

```

Test for time invariant effects
      Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                0.454                0.000
AGE                        0.138                0.000
TCELL                      0.189                0.091
PLATELET                   0.153                0.046
      Cramer von Mises test p-value H_0:constant effect
(Intercept)                11.800               0.000
AGE                        0.955                0.000
TCELL                      1.900                0.056
PLATELET                   0.640                0.122

```

下面拟合结局状态: 复发, 结果解释同上

Model for Survival status == Relapse
Competing risks Model

Test for nonparametric terms

```

Test for non-significant effects
      Supremum-test of significance p-value H_0: B(t)=0
(Intercept)                6.97                0.000
AGE                        1.48                0.522
TCELL                      1.91                0.317
PLATELET                   3.47                0.004

```

检验基线风险(intercept)是否为0, p 值非常显著。AGE、TCELL、PLATELET 是否与风险有关, AGE、TCELL、PLATELET 三自变量的 P 值分别为 0.522、0.317、0.005, 表示 AGE、TCELL 与复发的风险无显著关系, PLATELET 则与复发风险关系显著。下面是对每个自变量的效应是否随时间有变化进行检验, 分别采用了两种检验方法: Kolmogorov-Smirnov 检验与 Cramer von Mises 检验。结果显示基线风险随时间变化显著, TCELL 的效应随时间有变化显著。

```

Test for time invariant effects
      Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                0.1450               0.000
AGE                        0.0341               0.301
TCELL                      0.2290               0.028
PLATELET                   0.0760               0.270
      Cramer von Mises test p-value H_0:constant effect
(Intercept)                0.7140               0.000
AGE                        0.0218               0.422
TCELL                      1.9700               0.038
PLATELET                   0.1170               0.374

```

下面采用模型 II (所有变量都采用参数拟合), 模型表达式如下:

Model II: Event(TIME, CAUSE) ~ const(AGE) + const(TCELL) + const(PLATELET)

首先拟合结局状态“死亡”

Model for Survival status == Death
Competing risks Model

Test for nonparametric terms

```

Test for non-significant effects
      Supremum-test of significance p-value H_0: B(t)=0
(Intercept)                10.3                0

```

上面检验的是基线风险(intercept)是否为0, p 值为0, 非常显著。

```

Test for time invariant effects
      Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                0.492                0

```

```

Cramer von Mises test p-value H_0:constant effect
(Intercept)                13.3                0

```

上面检验的是基线风险(intercept)是否随时间变化, p 值为 0, 非常显著。下面是参数拟合的结果, 各列数据分别是: 回归系数、标准误、稳健的标准误、Z (检验) 值、P 值、回归系数的 95%可信区间下限与上限。

```

Parametric terms :
      Coef.    SE Robust SE      z P-val lower2.5% upper97.5%
const (AGE)    0.324 0.092    0.092  3.533 0.000    0.144    0.504
const (TCELL) -0.694 0.303    0.303 -2.295 0.022   -1.288   -0.100
const (PLATELET) -0.561 0.204    0.204 -2.745 0.006   -0.961   -0.161

```

下面拟合结局状态“复发”, 结果解释同前。

```

Model for Survival status == Relapse
Competing risks Model

```

```

Test for nonparametric terms

```

```

Test for non-significant effects
Supremum-test of significance p-value H_0: B(t)=0
(Intercept)                7.23                0

```

```

Test for time invariant effects
      Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                0.174                0
      Cramer von Mises test p-value H_0:constant effect
(Intercept)                1.23                0

```

```

Parametric terms :
      Coef.    SE Robust SE      z P-val lower2.5% upper97.5%
const (AGE)    0.029 0.126    0.126  0.233 0.816   -0.218    0.276
const (TCELL)  0.417 0.296    0.296  1.407 0.159   -0.163    0.997
const (PLATELET) -0.093 0.241    0.241 -0.384 0.701   -0.565    0.379

```

下面采用模型 III, 按输入设置, age 采用非参数拟合, 其它变量采用参数拟合, 模型表达式如下:

```

Model III: Event (TIME, CAUSE) ~ AGE+const (TCELL)+const (PLATELET)

```

首先拟合结局状态“死亡”

```

Model for Survival status == Death
Competing risks Model

```

```

Test for nonparametric terms

```

```

Test for non-significant effects
Supremum-test of significance p-value H_0: B(t)=0
(Intercept)                11.70               0
AGE                        6.13                0

```

上面检验的是基线风险(intercept)是否为 0, AGE 是否与风险有关, p 值为 0, 非常显著。

```

Test for time invariant effects
      Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                0.512                0
AGE                        0.154                0
      Cramer von Mises test p-value H_0:constant effect
(Intercept)                14.30               0
AGE                        1.19                0

```

上面检验的是基线风险(intercept)与 AGE 是否随时间变化, p 值为 0, 非常显著。下面是参数拟合的结果, 各列数据分别是: 回归系数、标准误、稳健的标准误、Z (检验) 值、P 值、回归系数的 95%可信区间下限与上限。

```

Parametric terms :
      Coef.    SE Robust SE      z P-val lower2.5% upper97.5%
const (TCELL)  -0.701 0.298    0.298 -2.355 0.019   -1.285   -0.117
const (PLATELET) -0.548 0.203    0.203 -2.705 0.007   -0.946   -0.150

```

下面拟合结局状态“复发”，结果解释同前。

Model for Survival status == Relapse
Competing risks Model

Test for nonparametric terms

Test for non-significant effects

Supremum-test of significance p-value H_0: B(t)=0

(Intercept)	7.24	0.000
AGE	1.26	0.683

Test for time invariant effects

Kolmogorov-Smirnov test p-value H_0:constant effect

(Intercept)	0.1750	0.000
AGE	0.0291	0.457

Cramer von Mises test p-value H_0:constant effect

(Intercept)	1.2600	0.000
AGE	0.0153	0.532

Parametric terms :

	Coef.	SE	Robust SE	z	P-val	lower2.5%	upper97.5%
const(TCELL)	0.409	0.296	0.296	1.382	0.167	-0.171	0.989
const(PLATELET)	-0.091	0.240	0.240	-0.378	0.705	-0.561	0.379