

Tarea 4. Análisis de Supervivencia.

Antonio, H. F.

2022-12-09

Problema 1.

(Modelos de Fragilidad) Kidney data: Considera el conjunto de datos dados en **MCGilchrist and Aisbett, Biometrics 47, 461-466, 1991**. Datos sobre los tiempos de recurrencia de infección, en el punto de inserción del catéter, para pacientes renales que utilizan equipos portátiles de diálisis. Los catéteres pueden retirarse por razones distintas a la infección, en cuyo caso se censura la observación. Cada paciente tiene exactamente 2 observaciones. Varios autores han utilizado el conjunto de datos para ilustrar modelos de efectos aleatorios (“**fragilidad**”) para datos de supervivencia. Sin embargo, cualquier estimación distinta de cero del efecto aleatorio se debe casi en su totalidad a un valor atípico, el sujeto 21.

- **id**: Número de paciente para los 38 pacientes.
- **time**: Tiempo hasta la infección desde la inserción del catéter.
- **age**: Edad de paciente, en años.
- **sex**: Sexo de cada paciente (**1 = hombre, 2 = hombre**).
- **disease**: Tipo de enfermedad (**GN, AN, PKD, Other**).
- **frail**: Fragilidad estimada del artículo original.

id	time	status	age	sex	disease	frail	id	time	status	age	sex	disease	frail
1	8	1	28	1	3	2.3	12	141	1	34	2	3	1.2
1	16	1	28	1	3	2.3	12	8	0	34	2	3	1.2
2	23	1	48	2	0	1.9	13	96	1	35	2	1	1.4
2	13	0	48	2	0	1.9	13	38	1	35	2	1	1.4
3	22	1	32	1	3	1.2	14	149	0	42	2	1	0.4
3	28	1	32	1	3	1.2	14	70	0	42	2	1	0.4
4	447	1	31	2	3	0.5	15	536	1	17	2	3	0.4
4	318	1	32	2	3	0.5	15	25	0	17	2	3	0.4
5	30	1	10	1	3	1.5	16	17	1	60	1	1	1.1
5	12	1	10	1	3	1.5	16	4	0	60	1	1	1.1
6	24	1	16	2	3	1.1	17	185	1	60	2	3	0.8
6	245	1	17	2	3	1.1	17	177	1	60	2	3	0.8
7	7	1	51	1	0	3	18	292	1	43	2	3	0.8
7	9	1	51	1	0	3	18	114	1	44	2	3	0.8
8	511	1	55	2	0	0.5	19	22	0	53	2	0	0.5
8	30	1	56	2	0	0.5	19	159	0	53	2	0	0.5
9	53	1	69	2	1	0.7	20	15	1	44	2	3	1.3
9	196	1	69	2	1	0.7	20	108	0	44	2	3	1.3
10	15	1	51	1	0	0.4	21	152	1	46	1	2	0.2
10	154	1	52	1	0	0.4	21	562	1	47	1	2	0.2
11	7	1	44	2	1	0.6	22	402	1	30	2	3	0.6
11	333	1	44	2	1	0.6	22	24	0	30	2	3	0.6

id	time	status	age	sex	disease	frail	id	time	status	age	sex	disease	frail
23	13	1	62	2	1	1.7	31	27	1	56	2	1	1.7
23	66	1	63	2	1	1.7	31	58	1	56	2	1	1.7
24	39	1	42	2	1	1	32	5	0	50	2	1	1.3
24	46	0	43	2	1	1	32	43	1	51	2	1	1.3
25	12	1	43	1	1	0.7	33	152	1	57	2	2	2.9
25	40	1	43	1	1	0.7	33	30	1	57	2	2	2.9
26	113	0	57	2	1	0.5	34	190	1	44	2	0	0.7
26	201	1	58	2	1	0.5	34	5	0	45	2	0	0.7
27	132	1	10	2	0	1.1	35	119	1	22	2	3	2.2
27	156	1	10	2	0	1.1	35	8	1	22	2	3	2.2
28	34	1	52	2	1	1.8	36	54	0	42	2	3	0.7
28	30	1	52	2	1	1.8	36	16	0	42	2	3	0.7
29	2	1	53	1	0	1.5	37	6	0	52	2	2	2.1
29	25	1	53	1	0	1.5	37	78	1	52	2	2	2.1
30	130	1	54	2	0	1.5	38	63	1	60	1	2	1.2
30	26	1	54	2	0	1.5	38	8	0	60	1	2	1.2

- Ajuste el modelo de riesgos proporcionales Cox sin usar la variable **frail** (es solo referencia).
- Ajuste el modelo de Cox suponiendo fragilidad ya que existe recurrencia y compare.
- Suponga un modelo paramétrico con fragilidad gama y tiempos de vida Weibull, estime parámetros y compare.
- Usando un enfoque bayesiano proponga una solución numérica para el modelo usado en c). (Puede usar INLA o BRMS de R).

Solución a)

Problema 2.

Análisis de riesgos en competencia. El trasplante de médula ósea (TMO) es un tratamiento estándar para la leucemia aguda. Klein y Moeschberger (1997) presentan un conjunto de datos de TMO para 137 pacientes, agrupados en tres categorías de riesgo según su estado en el momento del trasplante: leucemia linfoblástica aguda (LLA), leucemia mielocítica aguda (LMA) de bajo riesgo y LMA alto riesgo. Durante el período de seguimiento, algunos pacientes pueden recaer o algunos pacientes pueden morir mientras están en remisión. Considere la recaída como el evento de interés.

La muerte es un riesgo competitivo porque la muerte impide la aparición de una recaída de la leucemia.

Se utiliza el modelo de Fine y Gray (1999) para comparar las categorías de riesgo sobre la supervivencia libre de enfermedad. La variable Enfermedad representa el grupo de riesgo de un paciente, que es LLA, LMA de bajo riesgo o LMA de alto riesgo. La variable T representa la supervivencia libre de enfermedad en días, que es el tiempo hasta la recaída, el tiempo hasta la muerte o censurado. La variable Estado tiene tres valores: 0 para observaciones censuradas, 1 para pacientes con recaída y 2 para pacientes que mueren antes de experimentar una recaída.

Disease	Time	Status	Disease	Time	Status	Disease	Time	Status	Disease	Time	Status
1	2081	0	1	110	1	2	414	2	3	422	1
1	1377	0	2	2218	0	2	641	2	3	47	1
1	1111	0	2	1363	0	2	748	1	3	32	1
1	383	1	2	1870	0	2	381	1	3	105	2
1	487	2	2	1527	0	2	248	1	3	80	2
1	122	2	2	848	0	3	2640	0	3	16	2
1	466	2	2	1384	0	3	1238	0	3	63	2
1	107	2	2	105	2	3	845	0	1	1433	0
2	2409	0	2	79	2	3	2	2	1	1199	0
2	1470	0	2	1074	2	3	318	2	1	418	2
2	2246	0	2	35	2	3	183	2	1	172	2
2	1568	0	2	606	1	3	120	1	1	526	2
2	847	0	3	2133	0	3	74	2	1	86	2
2	1447	0	3	1136	0	3	273	1	1	1	2
2	481	2	3	100	1	1	1462	0	2	2506	0
2	421	1	3	268	1	1	226	0	2	1562	0
2	272	1	3	390	1	1	1167	0	2	1258	0
2	80	2	3	93	1	1	609	1	2	1674	0
2	219	1	3	168	2	1	230	1	2	932	0
3	2140	0	3	48	1	1	122	1	2	1535	0
3	1345	0	3	363	2	1	55	1	2	1063	2
3	84	1	1	1496	0	2	2569	0	2	288	2
3	456	1	1	996	0	2	1829	0	2	48	2
3	47	1	1	1182	0	2	860	0	2	53	2
3	164	2	1	104	1	2	1709	0	2	211	1
3	64	1	1	194	2	2	957	0	3	2252	0
3	625	1	1	74	1	2	1843	0	3	2024	0
3	113	1	1	109	1	2	2204	2	3	162	2
1	1602	0	1	332	2	2	390	2	3	242	1
1	1330	0	2	1857	0	2	486	1	3	467	1
1	530	0	2	1030	0	2	10	2	3	115	1
1	276	2	2	1799	0	2	704	2	3	677	2
1	662	1	2	1324	0	3	2430	0	3	157	1
1	129	1	2	1850	0	3	1631	0	3	76	1

- Estime y grafique la incidencia acumulada de los riesgos competitivos.
- Utilice un modelo de Cox causa específica y estime e interprete.
- Especifique y utilice un modelo de regresión de Fine-Gray y compare.
- Proponga una modelo paramétrico de causa específica y compare.

Solución a)

Se exploran los datos:

```
data <- read.csv("Disease.csv")
summary(data)
```

```
##      Disease      Time      Status
##  Min.   :1.000   Min.    : 1   Min.    :0.0000
## 1st Qu.:1.000   1st Qu.: 122   1st Qu.:0.0000
## Median :2.000   Median : 467   Median :1.0000
## Mean   :2.051   Mean    : 782   Mean    :0.9051
## 3rd Qu.:3.000   3rd Qu.:1363   3rd Qu.:2.0000
```

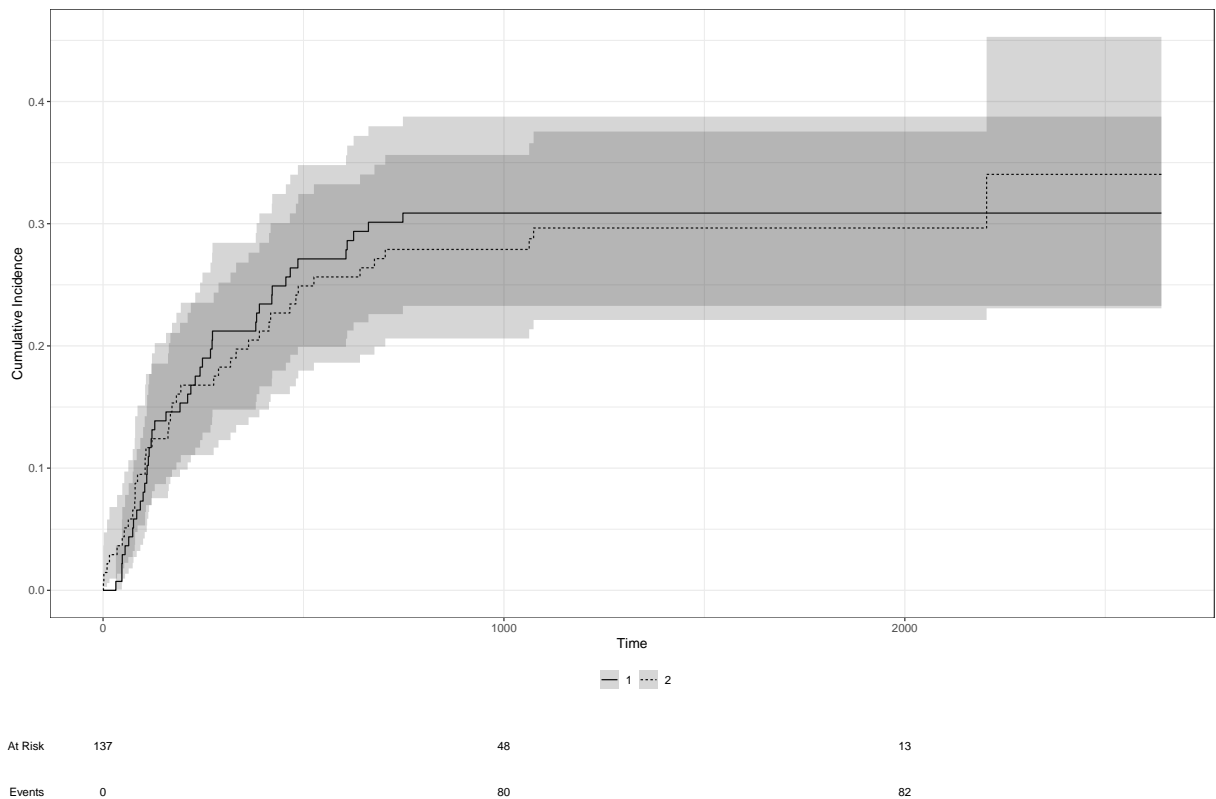
```
## Max. :3.000 Max. :2640 Max. :2.0000
```

Se convierten a factores las variables categóricas.

```
data$Disease <- as.factor(data$Disease)
data$Status <- as.factor(data$Status)
```

Se estima la incidencia acumulada de los riesgos en competencia.

```
oSup <- Surv(Time, Status) ~ 1
cuminc(oSup, data = data) %>%
ggcuminc(outcome = c("1", "2")) +
  add_confidence_interval() +
  add_risktable()
```



Solución b)

Se ajusta un modelo de Cox para cada Status.

```
oSup1 <- Surv(Time, Status == 1) ~ 1
oSup2 <- Surv(Time, Status == 2) ~ 1
mCox1 <- coxph(oSup1, data=data)
mCox2 <- coxph(oSup2, data=data)
```

Revisamos si los modelos son iguales o distintos.

```
LM1 <- mCox1$loglik
LM2 <- mCox2$loglik
pVal <- 1 - pchisq(LM2-LM1, df=1)
pVal
```

```
## [1] 0.007549638
```

Se concluyen que son diferentes con un $\alpha = 0.005$.

Solución c)

Se ajusta el modelo de Fine-Gray para cada **Status**

```
oSup3 <- Surv(Time,Status) ~ .  
crr(oSup3,data=data)
```

```
##
```

```
## -- crr() -----
```

```
## * Call Surv(Time, Status) ~ .
```

```
## * Failure type of interest "1"
```

```
##
```

## Variable	Coef	SE	HR	95% CI	p-value
## Disease2	-0.803	0.428	0.45	0.19, 1.04	0.061
## Disease3	0.508	0.366	1.66	0.81, 3.41	0.16

```
crr(oSup3, failcode = "2", data=data)
```

```
##
```

```
## -- crr() -----
```

```
## * Call Surv(Time, Status) ~ .
```

```
## * Failure type of interest "2"
```

```
##
```

## Variable	Coef	SE	HR	95% CI	p-value
## Disease2	-0.115	0.371	0.89	0.43, 1.84	0.76
## Disease3	-0.092	0.402	0.91	0.42, 2.01	0.82

Se observa que solamente para **Status** = 1 el grupo de riesgo es significativo.

Solución c)

Problema 3

Tamaño de muestra. En un estudio con pacientes con melanoma múltiple, el interés es investigar el efecto del logaritmo del nitrógeno diuretico en sangre 1BUN en la supervivencia de los pacientes. Estime el tamaño de muestra requerido para lograr un poder del 80% para detectar un cambio de 0.5 en el riesgo logarítmico asociado a un cambio de una unidad en el 1BUN usando una prueba de Wald unilateral con un nivel de significancia del 5%, después de ajustar por otros factores. A partir de estudios previos se ha estimado que la desviación estándar de 1BUN es 0.3126.

Asume:

- 1BUN es independiente de las otras covariables (experimento aleatorio)
- 1BUN está correlacionado con las otras covariables ($R^2 = 0.1837$)
- Nivel de censura (sobre la tasa de muerte = 0.8)

Solución para independencia de 1BUN

```
sEst <- 0.3126
nSig <- 0.05
pReq <- 0.8
cReq <- 0.5
cEst <- 0.8
r2 <- 0
ssizeEpi.default(power = pReq, theta = cReq, p = 0.5,
                  psi = cEst,
                  rho2 = r2,
                  alpha = nSig)
```

```
## [1] 82
```

Solución 1BUN correlacionado y nivel de censura 0.8

```
sEst <- 0.3126
nSig <- 0.05
pReq <- 0.8
cReq <- 0.5
cEst <- 0.8
r2 <- 0.3126
ssizeEpi.default(power = pReq, theta = cReq, p = 0.5,
                  psi = cEst,
                  rho2 = r2,
                  alpha = nSig)
```

```
## [1] 119
```