

Eventos recurrentes

Giancarlo Sal y Rosas

Departamento de Ciencias
Pontificia Universidad Católica del Perú

June 7, 2017

Outline

- 1 Motivación
- 2 Proceso de conteo
- 3 Proceso de estratificación

Motivación

- Cáncer:
 - La persona se puede curar a su primer episodio (proceso de remisión)
 - Después un tiempo el mal reaparece (a veces mas fuerte)
 - ¿Cómo estudio los factores asociados con desarrollar cáncer?
- Faltas de tránsito:
 - En el Perú, se tiene un sub registro de faltas cometidas por conductores
 - ¿Cómo estudio los factores asociados con el tiempo a recibir la primera (segunda, tercera, ...) falta de tránsito?

Tumor de vejiga

- Estudio aleatorizado en pacientes con tumores de vejiga superficial
- El objetivo era entender el efecto de un tratamiento (thiotepa) en la frecuencia de ocurrencia de tumores.
- 38 pacientes fueron asignados al tratamiento nuevo (thiotepa) y 48 a placebo.
- La posible dependencia de la tasa de recurrencia de el número y tamaño de tumores pre-aleatorización

Estudio de tumor de vejiga

● Estructura 1

```
> head(bladder)
  id rx number size stop event  enum
1  1  1      1    3     1      0     1
2  1  1      1    3     1      0     2
3  1  1      1    3     1      0     3
4  1  1      1    3     1      0     4
5  2  1      2    1     4      0     1
6  2  1      2    1     4      0     2
```

● Estructura 2

```
> head(bladder2)
  id rx number size start stop event  enum
1  1  1      1    3     0     1      0     1
2  2  1      2    1     0     4      0     1
3  3  1      1    1     0     7      0     1
4  4  1      5    1     0    10      0     1
5  5  1      4    1     0     6      1     1
6  5  1      4    1     6    10      0     2
```

● Datos

- **treatment:** placebo, pyridoxine y thiotepa
- **number:** Número inicial de tumores
- **size:** Tamaño (cm) del tumor inicial mas grande

Modelos

- Eventos recurrentes
 - Tiempo a eventos (modelo de proceso)
 - Tiempo entre eventos (Modelos de gap)
- Eventos ordenados
 - HIV - AIDS - Muerte
 - Estadio de enfermedad I - II - III - IV
- Eventos multiples sin orden

Modelos

- Tiempo a eventos
 - Interes en la tasa de ocurrencia de recurrencias en el tiempo
 - El tiempo es medido desde el *origen* hasta los eventos
 - El origen puede ser un tiempo calendario, o un evento biologico o inicio de tratamiento
- Modelos de tiempo entre eventos
 - Variables de interes son los tiempos entre eventos
 - Son de interes cuando existen patrones ciclicos: Ciclos menstruales de una mujer

Proceso de intensidad

- Sea T_1, T_2, \dots denota el tiempo hasta el primer, segundo y demas eventos
- Para eventos recurrentes, todas las observaciones son censuradas eventualmente
- Sea $N(t)$ el número de eventos que un individuo experimenta hasta el tiempo t , entonces *el proceso de intensidad* (que puede depender de covariables) es

$$\lambda(t) = \lim_{h \rightarrow 0} \frac{P[N(t+h) - N(t) = 1]}{h}$$

Proceso de conteo

- Modelo

$$\lambda_{ik}(t) = \lambda_0(t) e^{Z_i(t)\beta}$$

donde

- $\lambda_{ik}(\cdot)$ es la función de riesgo para el k -ésimo evento del i -ésimo individuo
- $\lambda_0(\cdot)$ función de riesgo basal para todos lo eventos en el tiempo
- $Z(\cdot)$ el valor de la covariable para el sujeto i -ésimo en el tiempo t

Proceso de conteo

- Paciente 14

```
> bladder2[bladder2$id==14,]  
  id rx number size start stop event  enum  
25 14  1      2   3      0   7      1     1  
26 14  1      2   3      7  10      1     2  
27 14  1      2   3     10  16      1     3  
28 14  1      2   3     16  24      1     4
```

- El paciente con *id* 14 esta en observación para cualquier evento entre el tiempo 0 y el tiempo 24.
 - Entra en 0 y sale en 3: $t_1 = 3$
 - Entra en 3 y sale en 9: $t_2 = 6$
 - Entra en 9 y sale en 21: $t_3 = 12$
 - Entra en 21 y sale en 24: $t_4 = 3$

Proceso de conteo

• Código de R

```
> model1 <- coxph(Surv(start,stop,event) ~ factor(rx) + number + size ,
+                  data=bladder2)
```

```
n= 178, number of events= 112
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
factor(rx) 2	-0.46469	0.62833	0.19973	-2.327	0.019989 *
number	0.17496	1.19120	0.04707	3.717	0.000202 ***
size	-0.04366	0.95728	0.06905	-0.632	0.527196

	exp(coef)	exp(-coef)	lower .95	upper .95
factor(rx) 2	0.6283	1.5915	0.4248	0.9294
number	1.1912	0.8395	1.0862	1.3063
size	0.9573	1.0446	0.8361	1.0960

- *Interpretación:* Existe evidencia que el tratamiento con thiotepa reduce el riesgo de recurrencia de tumores en un 40% (HR=0.63, 95%IC: 0.42-0.93, p=0.02)

Proceso de conteo

- Lin and Wei (1989) desarrollo el *estimador robusto de la varianza*:

$$\hat{J}_n^{-1}(\hat{\theta}) \hat{V}_n(\hat{\theta}) \hat{J}_n^{-1}(\hat{\theta})$$

donde

$$\hat{V}_n(\theta) = \sum_{i=1}^n \nabla l_n(\theta) (\nabla l_n(\theta))^t \quad , \quad \hat{J}_n(\theta) = \sum_{i=1}^n \nabla^2 l_n(\theta)$$

son estimadores de

$$V_n(\theta) = \text{Var}(\nabla(l_n(\theta))) \quad , \quad J_n(\theta) = -E[\nabla^2 l_n(\theta)]$$

respectivamente.

- En general se cumple

$$\sqrt{n}(\hat{\theta}_n - \theta^*) \rightarrow N(0, J_1(\theta^*) V_1(\theta) J_1(\theta)^{-1})$$

Proceso de conteo: Varianza robusta

● Código de R

```
> model2 <- coxph(Surv(start,stop,event) ~ factor(rx) + number + size ,
+               data=bladder2, robust = TRUE)
> summary(model2)
```

n= 178, number of events= 112

	coef	exp(coef)	se(coef)	robust se	z	Pr(> z)	
factor(rx) 2	-0.46469	0.62833	0.19973	0.23517	-1.976	0.0482	*
number	0.17496	1.19120	0.04707	0.06669	2.624	0.0087	**
size	-0.04366	0.95728	0.06905	0.07521	-0.581	0.5616	

	exp(coef)	exp(-coef)	lower .95	upper .95
factor(rx) 2	0.6283	1.5915	0.3963	0.9963
number	1.1912	0.8395	1.0452	1.3575
size	0.9573	1.0446	0.8261	1.1093

- *Interpretación:* Existe evidencia que el tratamiento con thiotepa reduce el riesgo de recurrencia de tumores en un 40% (HR=0.63, 95%IC: 0.40-0.99, p=0.048)

Proceso de estratificación

- No siempre podemos asumir independencia entre los eventos recurrentes:
 - EL orden de los eventos es importante
 - Las categorías de el evento son diferentes
- En este caso, una forma de analizar la data es estratificando por el número de eventos.
 - Es decir suponer una función de riesgo basal diferente para una persona que experimento uno, dos, tres o k eventos.
- Consideremos dos diferentes formas de estratificar

Proceso de estratificación

- **Caso 1:** La estructura de la data es igual al proceso de conteo.

```
> bladder2[bladder2$id==14,]
  id rx number size start stop event enum
25 14  1      2   3     0   7      1    1
26 14  1      2   3     7  10      1    2
27 14  1      2   3    10  16      1    3
28 14  1      2   3    16  24      1    4
```

- **Caso 2:** El tiempo de inicio es siempre cero (0) terminado el proceso anterior

```
> bladderB      <- bladder2
> bladderB$start <- 0
> bladderB$stop  <- bladder2$stop - bladder2$start
> bladderB[bladderB$id==14,]
  id rx number size start stop event enum
25 14  1      2   3     0   7      1    1
26 14  1      2   3     0   3      1    2
27 14  1      2   3     0   6      1    3
28 14  1      2   3     0   8      1    4
```

Proceso de estratificación

- El modelo de Cox en este caso tiene la forma

$$\lambda_k(t) = \lambda_{0,g}(t)e^{Z_i\beta}$$

donde

- λ_{0k} representa la función de riesgo basal específica para el evento k -ésimo en el tiempo
- *Supuesto*: Un sujeto no está en riesgo del evento k -ésimo si el evento $k - 1$ no ha ocurrido

Modelo: Condicional caso 1

```
> model4 <- coxph(Surv(start,stop,event) ~ factor(rx) + number + size + strata(enum),
+                 data=bladder2)
> summary(model4)
```

n= 178, number of events= 112

	coef	exp(coef)	se(coef)	z	Pr(> z)
factor(rx)2	-0.43228	0.64903	0.22082	-1.958	0.0503 .
number	0.11557	1.12252	0.05373	2.151	0.0315 *
size	-0.01566	0.98446	0.07297	-0.215	0.8300

	exp(coef)	exp(-coef)	lower .95	upper .95
factor(rx)2	0.6490	1.5408	0.4210	1.001
number	1.1225	0.8909	1.0103	1.247
size	0.9845	1.0158	0.8533	1.136

- *Interpretación:* Condicionado en la recurrencia previa, el tratamiento reduce en 35% el riesgo de contraer la subsiguiente (CRI: 0.65, 95%IC: 0.42-1.00)

Modelo: Condicional 1

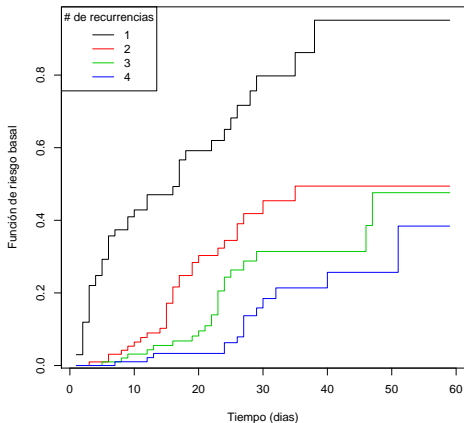


Figure : Funciones basales de riesgo

Modelo: Condicional caso 2

```
> model5 <- coxph(Surv(stop,event) ~ factor(rx) + number + size + strata(enum),
+                 data=bladderB, robust = TRUE)
> summary(model5)
```

```
n= 178, number of events= 112
```

		coef	exp(coef)	se(coef)	robust se	z	Pr(> z)
factor(rx)	2	-0.279005	0.756536	0.207348	0.217334	-1.284	0.19923
number		0.158046	1.171220	0.051942	0.055976	2.823	0.00475 **

		exp(coef)	exp(-coef)	lower .95	upper .95
factor(rx)	2	0.7565	1.3218	0.4941	1.158
number		1.1712	0.8538	1.0495	1.307
size		1.0074	0.9926	0.8792	1.154

- *Interpretación:* el tratamiento no reduce el riesgo a la ocurrencia de un tumor ($p = 0.199$)

Modelo: Condicional 2

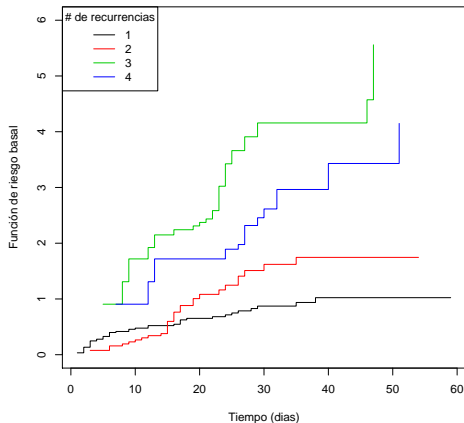


Figure : Funciones basales de riesgo

Conclusiones

- ¿Qué modelo es mejor?
 - Si el orden de los eventos no es importante
 - Modelo de conteo
 - Si el orden es importante
 - Condicional A: El tiempo de interes es el tiempo desde la entrada al estudio
 - Condicional B: El tiempo de interes es el tiempo entre dos eventos

Referencias

Lin, D. and Wei, L. (1989). The robust inference for the Cox proportional hazards model. *J Am Stat Assoc*, 82(1):1075–1078.