Universidad Autónoma Metropolitana
Unidad Iztapalapa
División de Ciencias Básicas e Ingeniería

Casa abierta al tiempo

"Modelado de regresión basado en cópula para eventos secuenciales de Supervivencia"

Tesis que presenta:

Itzel Moctezuma Barona

Matrícula 2192802610

para obtener el grado de:

Maestra en Ciencias
(Matemáticas Aplicadas e Industriales)

Director:
Dr. Gabriel Escarela Pérez

Jurados:
Dra. Lizbeth Naranjo Albarrán
Dr. Alberto Castillo Morales
Dr. Gabriel Núñez Antonio

Iztapalapa, Ciudad de México, Diciembre 2021

Índice general

Modelado de regresión basado en cópula para eventos secuenciales de Supervivencia.

Agradecimientos

En especial quiero agradeder a mis padres, por el apoyo incondicional que he recibido en todo momento, por la fortaleza y sacrificio que han hecho para darme siempre lo mejor y hacer de mí una mejor persona.

Al Dr. Gabriel Escarela, quien me dio la oportunidad de descubrir y aprender un poco más acerca de la estadística, por su paciencia y consejos brindados durante la elaboración de esta tesis.

A mis sinodales: Dra. Lizbeth Naranjo, Dr. Alberto Castillo y Dr. Gabriel Núñez, por el tiempo dedicado a leer este trabajo, por compartirme sus ideas y hacer de esta tesis un mejor trabajo.

Al CONACYT, por otorgarme la beca durante mi estancia, sin este apoyo hubiese sido complicado culminar mi posgrado.

A mis amigos y a todas aquellas personas que me apoyaron y formaron parte de esta trayectoria.

Introducción

El objetivo principal de un análisis de supervivencia es modelar y analizar los datos del tiempo transcurrido hasta la ocurrencia del evento de interés, a estos tiempos generalmente se les conoce como tiempos de vida, de falla o de supervivencia, y son de gran importancia en diversas áreas como la biomedicina, la ingeniería o incluso en las ciencias sociales, algunas aplicaciones están basados en el tiempo hasta la recurrencia de una enfermedad después de la aplicación de un tratamiento, tiempo hasta la muerte, tiempo hasta la reincidencia delictiva, entre otros.

Es posible utilizar diferentes métodos estadísticos estándar cuando todos los tiempos de supervivencia son observados, sin embargo, una de las características principales que poseen estos datos es la presencia de censura. Decimos que los tiempos de supervivencia están censurados cuando el evento de interés no ha sido observado, así pues solo una proporción de individuos han experimentado dicho evento. La censura puede ocurrir en diferentes formas: el individuo decide no participar más tiempo en el estudio y lo abandona, o el individuo no experimenta el resultado relevante (una recaída o la muerte) durante el período de seguimiento. Así, la única información brindada por este tipo datos es que el tiempo de supervivencia de cada participante es mayor al último tiempo observado de cada individuo.

Cuando hablamos de los tiempos de supervivencia observados secuencialmente nos referimos a que los tiempos son observados uno después del otro, es decir, un individuo en el estudio puede experimentar más de un tiempo de supervivencia. Como mencionamos anteriormente, la presencia de censura induce dificultades en el análisis de estos datos. En el caso de datos bivariados ocurre algo particular, si los tiempos para un individuo no son independientes y el primer tiempo de supervivencia está censurado, entonces surge la censura inducida para el segundo tiempo de supervivencia, es decir, los segundos

tiempos de supervivencia son observados solo si los primeros tiempos no están censurados.

La motivación de esta investigación involucra a pacientes tratados con cáncer de colon para datos bivariados, donde el cáncer se ha diseminado al menos a un ganglio linfático cercano al intestino (etapa C de Dukes). Existen diferentes metodologías que aportaron investigadores para este tipo de datos, Lawless y Yilmaz modelaron la distribución conjunta de los tiempos de supervivencia mediante el uso de cópulas y proporcionaron procedimientos de estimación semiparamétricos (Lawless y Yilmaz 2011), Lin se basó en el enfoque de riesgo marginal, donde las distribuciones marginales de los tiempos de falla satisfacen el modelo de riesgos proporcionales de Cox y la estructura de dependencia para fallas relacionadas no está especificada (D. Y. Lin 1994).

En esta tesis desarrollamos un método totalmente paramétrico para la inferencia de datos de supervivencia bivariados: estimaremos la asociación entre los tiempos de supervivencia, así como la distribución marginal del segundo tiempo. El método se basa en caracterizar las distribuciones marginales mediante cuatro funciones de distribución: Burr, Log-logística, Weibull y Exponencial, donde cada función de riesgo será formulada utilizando el modelo de riesgos proporcionales de Cox, además como suponemos que los tiempos de supervivencia bivariados no son independientes modelamos la distribución conjunta mediante las cópulas paramétricas: Gaussiana, Gumbel y Clayton.
La metodología será aplicada a un ensayo clínico aleatorizado sobre la efectividad de tratamientos (Placebo, Levamisol con y sin fluorouracilo) a 929 pacientes con cáncer de colon obtenidos del software `R`, donde los tiempos de supervivencia secuenciales están definidos de la siguiente manera: T_1 tiempo de recurrencia del cáncer y T_2 tiempo de muerte debida al cáncer, en ambos eventos tenemos presencia de censura, además T_2 es observado solo si es observado T_1, es decir, consideramos que nadie en el estudio murió de cáncer sin haber presentado recurrencia. Asimismo supondremos que en la población existe una proporción p de pacientes susceptibles a recurrencias, este parámetro será modelado por medio de regresión logística, las variables explicativas que consideramos causarán efectos sobre p son tratamiento, género, cantidad de nodos y edad.

Nuestra meta será elegir el modelo más adecuado que ajuste mejor a los datos, por consiguiente obtendremos la medida de asociación entre estas variables que permitirá inferir la existencia o no de dependencia entre estas variables, también analizaremos la efectivi-

dad del modelo de riesgos proporcionales de Cox e igualmente deduciremos cuáles fueron las covariables que tuvieron efectos significativos sobre p. Para lograr dicho objetivo la investigación se divide en los siguientes capítulos: el primer capítulo está dedicado a un apartado de preliminares, describimos aspectos básicos sobre el análisis de supervivencia como: función de supervivencia, de riesgo, y datos censurados, definimos también algunos modelos paramétricos que utilizamos en nuestra aplicación. En el segundo capítulo explicamos la construcción del estimador de Kaplan-Meier, introduciremos el modelo de riesgos proporcionales, el cual es ampliamente utilizado en el análisis de datos de supervivencia. Exponemos el método de máxima verosimilitud y damos un ejemplo sobre la construcción de la función de verosimilitud cuando existen tiempos de fallas censurados. Finalmente, agregamos una sección destinado a definir y dar algunos resultados importantes sobre la teoría de cópulas que construyen la base para puntos específicos en la investigación.
En el tercer capítulo damos una pequeña introducción acerca del cáncer de colon, desarrollamos la metodología propuesta aplicada a la muestra, mostramos los resultados y realizamos un estudio de simulación para analizar el desempeño del método propuesto.

Capítulo 1

Preliminares

En esta sección definimos conceptos básicos del análisis de supervivencia univariado, mostraremos algunos resultados y ejemplos importantes, describimos el concepto de datos censurados específicamente la censura tipo III, además mencionamos algunas funciones de distribución utilizadas en el análisis de supervivencia.

1.1. Funciones de Supervivencia y Riesgo

Función de Supervivencia.

Cuando hablamos sobre el análisis de supervivencia nos referimos al tiempo en que ocurre algún evento de interés desde un punto de partida específico como el tiempo de vida o tiempo de supervivencia. El evento de interés puede ser la muerte o recurrencia de una determinada enfermedad, donde el tiempo de supervivencia es el tiempo de vida real o la fecha en que se observó la recurrencia de la enfermedad.
Sea T la variable aleatoria asociada al tiempo de supervivencia que toma valores no negativos y sea $f(t)$ la función de densidad de probabilidad de T, entonces la función

$$F(t) = \Pr(T < t) = \int_0^t f(u)du,$$

representa la probabilidad de que el tiempo de supervivencia sea menor que un tiempo t. Por otro lado podemos pensar en la probabilidad de que el tiempo de supervivencia de un individuo sea mayor o igual que un tiempo t, esta función se define como la *función de*

supervivencia y la denotamos por $S(t)$, notemos que esta probabilidad está determinada por el complemento de la ecuación anterior

$$S(t) = \Pr(T \geq t) = \int_t^{\infty} f(u)du = 1 - F(t).$$

Observe que $S(t)$ es una función continua monótona decreciente con $S(0) = 1$ y $S(\infty) = \lim_{t\to\infty} S(t) = 0$.

Función de Riesgo.

La *función de riesgo*

$$h(t) = \lim_{\delta t\to 0} \frac{\Pr(t \leq T \leq t + \delta t \mid T \geq t)}{\delta t},$$

describe la tasa instantánea de falla de un individuo en el tiempo t, dado a que haya sobrevivido hasta ese momento t; $h(t)\delta t$ es la probabilidad aproximada de muerte en el intervalo $(t, t + \delta t)$, dada la supervivencia hasta t.

Existe una relación importante entre las funciones de supervivencia (S) y de riesgo (h), a continuación se muestra el siguiente resultado

$$\begin{aligned} h(t) &= \lim_{\delta t\to 0} \frac{\Pr(t \leq T \leq t + \delta t \mid T \geq t)}{\delta t} = \lim_{\delta t\to 0} \frac{\Pr(t \leq T \leq t + \delta t)/\Pr(T \geq t)}{\delta t} \\ &= \lim_{\delta t\to 0} \frac{F(t + \delta t) - F(t)}{\delta t} \cdot \frac{1}{S(t)} = \frac{F'(t)}{S(t)} = \frac{f(t)}{S(t)} \end{aligned} \tag{1.1}$$

una implicación de este resultado es

$$H(t) = -\log[S(t)] \tag{1.2}$$

pues $h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt}\log[S(t)]$. La función $H(t) = \int_0^t h(u)du$ se define como el *riesgo acumulativo.*

Frecuentemente se desea conocer la estructura de la función de riesgo; una función creciente indica que los elementos tienen mayor probabilidad de presentar el evento con el paso del tiempo, mientras que una función decreciente representaría lo opuesto.

Ejemplo 1.1.1. *Sea la función de riesgo*

$$h(t) = \frac{\alpha}{\lambda}\left(\frac{t}{\lambda}\right)^{\alpha-1}, \qquad 0 \leq t < \infty$$

con parámetros α, λ mayores que cero. Una característica importante que posee esta función es la monotonía, para distintos valores de α la función es creciente o decreciente.

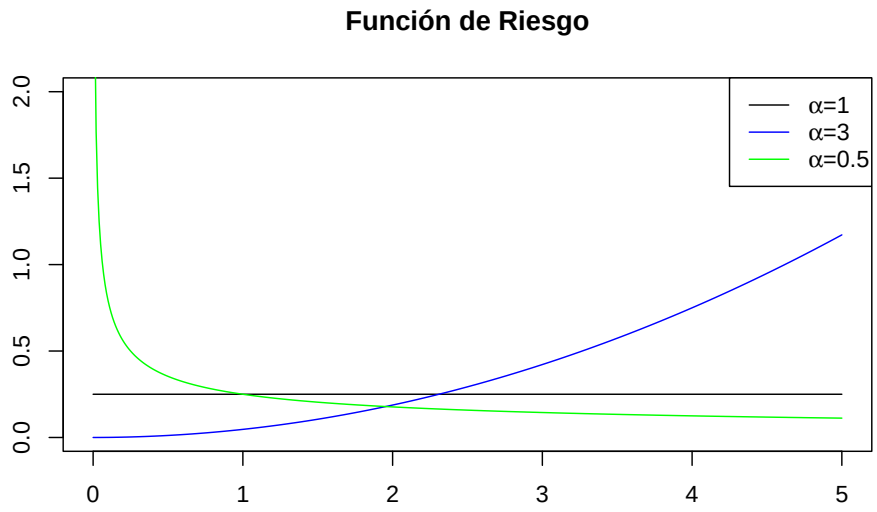


Figura 1.1: Forma de la función de riesgo Weibull, cuando $\alpha = 1$, 3, 0.5 y $\lambda = 4$.

El modelo más sencillo ocurre cuando $\alpha = 1$, pues la función de riesgo es constante sobre el tiempo y el tiempo de supervivencia sigue una distribución exponencial, $T \sim \exp(\lambda)$.
A continuación calculamos la función de supervivencia utilizando los resultados anteriores

$$S(t) = \exp\left[-\int_0^t \frac{\alpha}{\lambda}\left(\frac{u}{\lambda}\right)^{\alpha-1} \mathrm{du}\right] = \exp\left[-\left(\frac{t}{\lambda}\right)^{\alpha}\right].$$

La correspondiente función de densidad de probabilidad es

$$f(t) = h(t) \cdot S(t) = \frac{\alpha}{\lambda}\left(\frac{t}{\lambda}\right)^{\alpha-1} \exp\left[-\left(\frac{t}{\lambda}\right)^{\alpha}\right], \qquad 0 \leq t < \infty$$

que es la densidad de una variable aleatoria que tiene una distribución Weibull con parámetro de escala λ y parámetro de forma α.

1.2. Datos censurados

Una característica importante de los tiempos de supervivencia es que pueden estar sujetos a censura, básicamente, una observación censurada contiene información parcial sobre la variable de interés. La censura ocurre cuando los tiempos de supervivencia no se observan para una proporción de individuos en el estudio, en consecuencia, el tamaño de muestra efectivo del estudio se reduce.
Se dice que las observaciones están censuradas por la derecha cuando solo conocemos un límite inferior en el tiempo de supervivencia. Comúnmente la censura por la derecha surge porque el individuo no presenta el evento cuando finaliza el estudio, en otros casos sucede porque los pacientes han decidido retirarse del estudio debido a un pronóstico que empeora o mejora su estado de salud.

En esta investigación estamos interesados en la *censura tipo III*, en ella, los pacientes son reclutados en el estudio en diferentes momentos durante un período establecido, algunos pueden desarrollar el evento de interés antes del punto final del estudio, en consecuencia, proporcionan tiempos de supervivencia exactos. Por otro lado, las personas pueden retirarse durante el período de estudio, otros nunca desarrollan el resultado de interés (sus tiempos de supervivencia son al menos desde el inicio hasta el final del estudio), cuando esto ocurre decimos que los tiempos de seguimiento están censurados.
Bajo la censura tipo III, los tiempos de censura son diferentes para cada individuo que lo presenta y se comporta como una variable aleatoria. En nuestro estudio supondremos que la *censura es no informativa*, es decir, la distribución de censura no contiene información sobre los parámetros de interés.

1.3. Funciones de distribución

En nuestra investigación estaremos interesados en cuatro funciones de distribución que generalmente son utilizadas en el análisis de supervivencia. Esta sección está dedicada a definir las funciones: Burr, Weibull, Log-logística y Exponencial, y mostraremos cuáles son las relaciones que existen entre ellas.

La función de distribución Burr es el duodécimo ejemplo de soluciones de una ecuación diferencial que define el sistema de distribuciones Burr, introducido por Irving Burr en 1942. Dicha distribución fue redescubierta de forma independiente en diversas áreas, por tal motivo se conoce con una variedad de nombres: Burr XII, Burr, Singh-Maddala, Log-logística generalizada, entre otras. A continuación definimos dicha función.

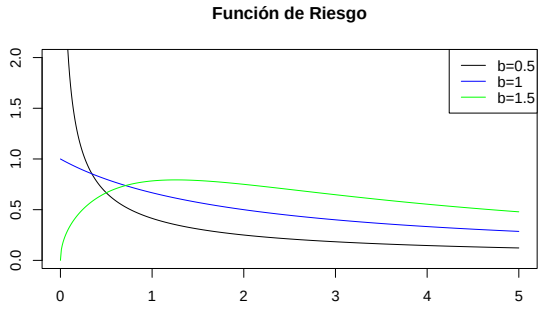
Definición. *La distribución Burr* tiene función de densidad

$$f(x;a,b,s) = \frac{abx^{b-1}}{s^b\left[1+\left(\frac{x}{s}\right)^b\right]^{1+a}}, \qquad a,b,s,x>0 \tag{1.3}$$

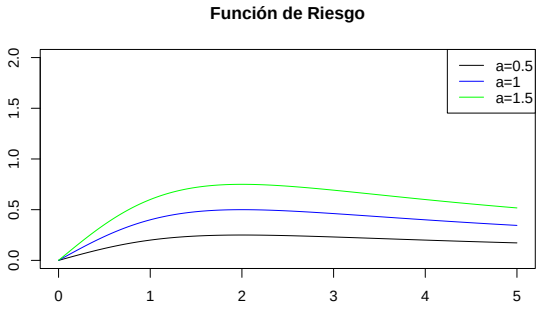
donde s es un parámetro de escala y a, b parámetros de forma; a solo afecta a la cola derecha, mientras que b afecta ambas colas.

La distribución Burr es una distribución de probabilidad continua que toma valores no negativos, a diferencia de las demás distribuciones del sistema Burr, la Burr XII ha mostrado ser más flexible, pues permite a su correspondiente función de riesgo tener una forma decreciente o en forma de bañera invertida (creciente en determinado tiempo y luego decreciente). Conforme a lo anterior, la distribución Burr puede ser utilizada cuando:

- La tasa de recurrencia disminuye.
- La tasa de recurrencia aumenta drásticamente durante un período de tiempo, y luego disminuye lentamente.



(a) Valores fijos $a = s = 2$ y varios valores de b.



(b) Valores fijos $b = s = 2$ y varios valores de a.

Figura 1.2: Forma de la función de riesgo Burr.

La función de distribución está definida por

$$F(x) = \int_0^x f(t)\,\mathrm{dt} = \int_0^x \frac{abt^{b-1}}{s^b\left[1+\left(\frac{t}{s}\right)^b\right]^{1+a}}\,\mathrm{dt} = 1 - \frac{1}{\left[1+\left(\frac{x}{s}\right)^b\right]^a},$$

obtenido por el cambio de variable $u = 1 + \left(\frac{t}{s}\right)^b$. Como vemos, está determinada por una expresión simple en forma cerrada, por lo que también obtenemos una expresión simple para la función cuantil

$$F^{-1}(u) = s[(1-u)^{-1/a} - 1]^{1/b}, \qquad \text{para} \quad 0 < u < 1.$$

Otras funciones importantes a considerar son

$$S(t) = \frac{1}{\left[1+\left(\frac{t}{s}\right)^b\right]^a} \quad \text{y} \quad h(t) = \frac{abt^{b-1}}{s^b\left[1+\left(\frac{t}{s}\right)^b\right]},$$

la función de supervivencia y de riesgo, respectivamente.

La función de *distribución log-logística* es un caso particular de la función de distribución Burr cuando el parámetro de forma a de la ecuación (1.3) es igual a 1,

$$f(x;b,s) = \frac{bx^{b-1}}{s^b\left[1+\left(\frac{x}{s}\right)^b\right]^2} = \frac{\left(\frac{b}{s}\right)\left(\frac{x}{s}\right)^{b-1}}{\left[1+\left(\frac{x}{s}\right)^b\right]^2}, \qquad \text{con} \quad b,s,x > 0 \tag{1.4}$$

por lo que se deducen fácilmente las funciones de supervivencia y riesgo.

Recordemos que una variable aleatoria tiene distribución Weibull si su distribución está dada por

$$F(x;\alpha,\lambda) = 1 - \exp\left[-\left(\frac{x}{\lambda}\right)^{\alpha}\right], \qquad \alpha,\lambda,x > 0 \tag{1.5}$$

donde α es el parámetro de forma y λ el de escala, con función de supervivencia y riesgo

$$S(t) = \exp\left[-\left(\frac{t}{\lambda}\right)^{\alpha}\right] \quad \text{y} \quad h(t) = \frac{\alpha}{\lambda}\left(\frac{t}{\lambda}\right)^{\alpha-1},$$

respectivamente. Como mencionamos anteriormente la forma de la función de riesgo depende de los valores de α, ya que puede ser monótona creciente o decreciente.

De la ecuación (1.5) cuando $\alpha = 1$, tenemos como caso particular a la función de distribución exponencial

$$F(x;\lambda) = 1 - \exp\left[-\left(\frac{x}{\lambda}\right)\right], \qquad \lambda,x > 0. \tag{1.6}$$

Un resultado importante es, la familia Weibull puede verse como un caso límite (cuando el parámetro de forma $a \to \infty$) de la familia Burr, esto se deduce del siguiente resultado. Sea X una variable aleatoria que se distribuye Burr, entonces

$$\begin{aligned}
P\left(X \le \left(\frac{1}{a}\right)^{1/b} y\right) &= 1 - \left[1 + \left(\frac{\left(\frac{1}{a}\right)^{1/b} y}{s}\right)^{b}\right]^{-a} \\
&= 1 - \left[1 + \frac{(y/s)^b}{a}\right]^{-a} \\
&= 1 - \exp\left[-a\ln\left(1 + \frac{(y/s)^b}{a}\right)\right] \\
&= 1 - \exp\left[-a\left(\frac{(y/s)^b}{a} - \frac{1}{2}\left[\frac{(y/s)^b}{a}\right]^2 + \cdots\right)\right], \quad \text{con} \quad \left|\frac{(y/s)^b}{a}\right| < 1 \\
&= 1 - \exp\left[-\left(\frac{y}{s}\right)^{b}\right], \quad \text{cuando} \quad a \to \infty.
\end{aligned}$$

que es la función de distribución de una variable aleatoria Weibull.

Capítulo 2

Métodos Estadísticos

2.1. Estimador de Kaplan-Meier (KM)

Herramientas como histogramas, funciones de distribución empírica y los gráficos de densidad son importantes en la descripción y el análisis de datos, sin embargo, para los datos de supervivencia la presencia de censura modifica los métodos estándar, por lo cual se opta por un método no paramétrico llamado así debido a que no se necesita asumir una distribución específica para los tiempos de supervivencia.
Existen varios métodos no paramétricos para estimar la función $S(t)$, en esta sección nos enfocaremos en el método de Kaplan-Meier (1958) que es especialmente desarrollado para datos censurados por la derecha, la construcción es la siguiente.

Sean n individuos con sus respectivos tiempos de supervivencia observados $t_1, t_2, \ldots, t_n$, estas observaciones pueden ser censuradas a la derecha. Supondremos que entre esos individuos hay r tiempos de muerte observados, con $r \leq n$ y los ordenamos de forma ascendente $t_{(1)} < t_{(2)} < \cdots < t_{(r)}$, donde el $j-$ésimo tiempo observado se denota por $t_{(j)}$ con $j = 1, \ldots, r$. Por otro lado, el número de individuos que están vivos antes del tiempo $t_{(j)}$ incluyendo los que están por morir en ese tiempo, se denotan por η_j para $j = 1, \ldots, r$, y denotemos por d_j como el número de individuos que mueren en este tiempo. El intervalo $[t_{(j)} - \delta, t_{(j)}]$ incluye un tiempo de muerte, donde δ es un tiempo infinitesimal. Como hay η_j individuos que están vivos poco antes de $t_{(j)}$, y d_j muertes en $t_{(j)}$, la probabilidad de que un individuo muera durante el intervalo $[t_{(j)} - \delta, t_{(j)}]$ es estimado por d_j/η_j, por tanto,

la probabilidad de supervivencia estimada es

$$1 - \frac{d_j}{\eta_j} = \frac{\eta_j - d_j}{\eta_j}.$$

El tiempo inmediatamente anterior al siguiente tiempo de muerte $(t_{(j)}, t_{(j+1)} - \delta)$ no contiene muertes, por lo tanto, la probabilidad de sobrevivir en el intervalo $(t_{(j)}, t_{(j+1)} - \delta)$ es uno, así la probabilidad conjunta de sobrevivir en $[t_{(j)} - \delta, t_{(j)}]$ y $(t_{(j)}, t_{(j+1)} - \delta)$ es estimada por $(\eta_j - d_j)/\eta_j$. En el límite, cuando δ tiende a cero, $(\eta_j - d_j)/\eta_j$ se convierte en un estimador de la probabilidad de supervivencia en $[t_{(j)}, t_{(j+1)})$.

Suponiendo que las muertes de los individuos en la misma muestra son independientes unas de otras, entonces el estimador de la función de supervivencia en cualquier tiempo t, en el $k-$ésimo intervalo de tiempo construido $[t_{(k)}, t_{(k+1)})$ con $k = 1, 2, \ldots, r$, será la probabilidad estimada de sobrevivir más allá de $t_{(k)}$. Esta es la probabilidad de sobrevivir en el intervalo $[t_{(k)}, t_{(k+1)})$ y todos los intervalos anteriores, que conduce al *estimador de Kaplan-Meier* de la función de supervivencia, el cual se calcula de la siguiente manera:

$$\widehat{S}(t) = \prod_{j=1}^{k} \left(\frac{\eta_j - d_j}{\eta_j} \right),$$

para $t_{(k)} \leq t < t_{(k+1)}$, $k = 1, 2, \ldots, r$, con $\widehat{S}(t) = 1$ cuando $t < t_{(1)}$ (Collet 2003).
Note que si la observación más grande es un tiempo de supervivencia censurado t^*, entonces $\widehat{S}(t)$ está indefinido para $t > t^*$. Por otro lado, si la observación más grande es un tiempo sin censura $t_{(r)}$, entonces $\eta_r = d_r$, así $\widehat{S}(t) = 0$ cuando $t \geq t_{(r)}$.

Si no hay observaciones censuradas, la función de supervivencia empírica en el tiempo t es la proporción de sobrevivientes en el tiempo t y el tamaño de la muestra es igual a n. Esta función tiene la propiedad de ser escalonada, disminuyendo en $1/n$ después de cada falla observada.

Ejemplo 2.1.1. *Consideramos el conjunto de datos* $\{10, 13^*, 18^*, 19, 23^*, 30, 36, 38^*, 54^*, 56^*, 59, 75, 93, 97, 104^*, 107, 107^*, 107^*\}$ *que representan los tiempos de supervivencia de* 18 *mujeres, definidos como el número de semanas desde la colocación del DIU (implante de un dispositivo intrauterino) hasta la interrupción (WHO,* 1987*). Denotamos con un asterisco los tiempos que son censurados.*

En la siguiente tabla presentamos el desarrollo de Kaplan-Meier, donde la función $\widehat{S}(t)$ *es la función de supervivencia estimada utilizando dicho método.*

Intervalo	η_j	d_j	$(\eta_j - d_j)/\eta_j$	$\widehat{S}(t)$
$0-$	18	0	1.0000	1.0000
$(0, 10]$	18	1	0.9444	0.9444
$(10, 19]$	15	1	0.9333	0.8815
$(19, 30]$	13	1	0.9231	0.8137
$(30, 36]$	12	1	0.9167	0.7459
$(36, 59]$	8	1	0.8750	0.6526
$(59, 75]$	7	1	0.8571	0.5594
$(75, 93]$	6	1	0.8333	0.4662
$(93, 97]$	5	1	0.8000	0.3729
$(97, 107]$	3	1	0.6667	0.2486

Cuadro 2.1: Estimador de Kaplan-Meier para la función S

Observamos a continuación que la gráfica de $\widehat{S}(t)$ *es precisamente una función escalonada, además como la observación más grande,* 107*, está censurada, entonces* $\widehat{S}(t)$ *no está definido cuando* $t > 107$*.*

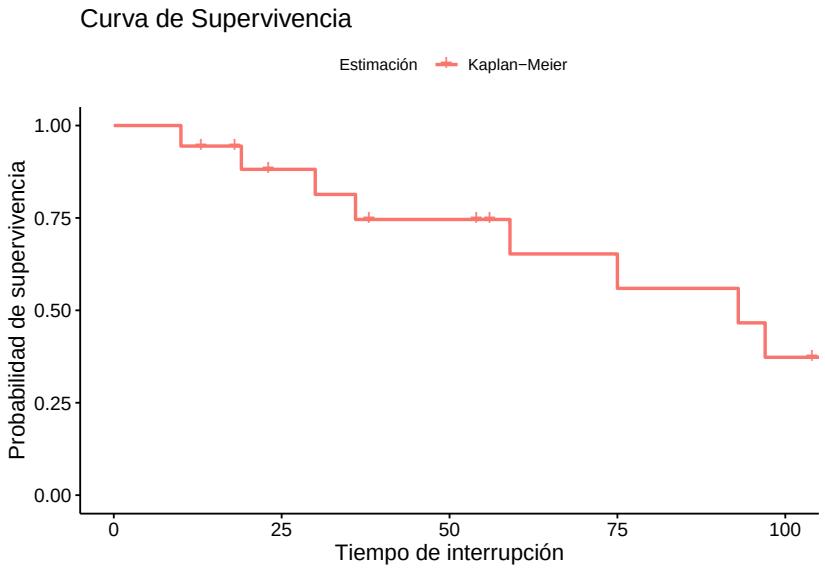


Figura 2.1: Estimador de Kaplan-Meier de la funcion S.

2.2. Modelo de Riesgos Proporcionales

En algunas aplicaciones prácticas, la población en estudio no es homogénea, en otras palabras, el análisis de datos pueden diferir de factores como: edad, género, educación, alimentación, etc. Estos factores son conocidas como variables explicativas, predictoras o covariables, y son de gran importancia pues en particular pueden determinar el efecto de un tratamiento en un ensayo clínico.
El modelo de riesgos proporcionales introducido por Cox (1972) es ampliamente utilizado en el análisis de datos de supervivencia y, posiblemente es el más aplicado y popular. Este modelo de regresión se especifica mediante la función de riesgo, donde la variable dependiente será el tiempo del evento. A continuación presentamos dicho modelo.

Sea $h(t|X)$ el riesgo de un individuo en el tiempo t con vector de covariables X^T (vector transpuesto de X) y β el vector de parámetros de regresión, donde $X^T\beta = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$. El *modelo de riesgos proporcionales* especifica que

$$h(t|X) = h_0(t)\exp(X^T\beta),$$

donde $h_0(t)$ es la función de riesgo base que especifica cómo cambia la función de riesgo en términos del tiempo cuando no existen efectos de las variables explicativas. El modelo supone que todos los individuos de la población tienen la misma función de riesgo base y a su vez están relacionadas por una constante de proporcionalidad que no depende de t, esto significa que la proporción de riesgos para cualesquiera dos individuos es constante sobre el tiempo cuando Y y Y^* son fijos:

$$\frac{h(t|Y)}{h(t|Y^*)} = \frac{h_0(t)\exp(Y^T\beta)}{h_0(t)\exp((Y^*)^T\beta)} = \frac{\exp(Y^T\beta)}{\exp((Y^*)^T\beta)} = \exp[(Y^T - (Y^*)^T)\beta],$$

a este valor se le conoce como *proporción de riesgos* (hazards ratio HR).

Al igual que con el modelado de regresión lineal, un objetivo estadístico de un análisis de supervivencia es obtener alguna medida de efecto que describa la relación entre una variable predictora de interés y el tiempo de falla. En el modelado de regresión lineal, la medida del efecto es el coeficiente de regresión, en análisis de supervivencia, la medida del efecto es la proporción de riesgos (HR), por ejemplo, supongamos que tenemos una

variable explicativa binaria X_1 que consta de los tipos de tratamiento que puede recibir un individuo: $X_1 = 0$ si un individuo recibe placebo y $X_1 = 1$ si recibe tratamiento, luego el modelo de riesgos proporcionales se escribe como

$$\begin{aligned} h(t|X_1 = 0) &= h_0(t) \\ h(t|X_1 = 1) &= h_0(t) \cdot \exp(\beta_1), \end{aligned}$$

donde β_1 denota el coeficiente del grupo covariable, entonces la proporción de riesgo está dada por

$$HR = \frac{h_0(t) \cdot \exp(\beta_1)}{h_0(t)} = \exp(\beta_1),$$

si $HR = 1$, entonces no hay ningún efecto, supongamos que se obtuvo un $HR = 5$, entonces el grupo tratamiento tiene cinco veces el riesgo del grupo placebo, análogamente un $HR = 1/5$ implica que el grupo tratamiento tiene una quinta parte del riesgo del grupo placebo.

El modelo de riesgos proporcionales también puede escribirse en términos de la función de supervivencia, esto se obtiene a partir del resultado $S(t) = \exp(-H(t))$, donde $H(t)$ es la función de riesgo acumulativa, entonces

$$\begin{aligned} S(t|X) &= \exp\left[-\int_0^t h(u|X)\,\mathrm{du}\right] = \exp\left[-\int_0^t h_0(u)\exp(X^T\beta)\,\mathrm{du}\right] \\ &= \exp\left[-\exp(X^T\beta)\int_0^t h_0(u)\,\mathrm{du}\right] = \exp\left[-\int_0^t h_0(u)\,\mathrm{du}\right]^{\exp(X^T\beta)} \\ &= \exp[-H_0(t)]^{\exp(X^T\beta)} = S_0(t)^{\exp(X^T\beta)}, \end{aligned}$$

donde $S_0(t)$ es la distribución de supervivencia base.

Para evaluar si los riesgos son proporcionales para los diferentes niveles de X, gráficamente se debe observar que las curvas de riesgo para cada nivel son paralelas. Mostraremos un método alternativo para verificar el supuesto de proporcionalidad.
Consideremos el modelo

$$S(t|X) = S_0(t)^{\exp(X^T\beta)},$$

aplicando logaritmos obtenemos,

$$\log[-\log[S(t|X)]] = \log[-\log[S_0(t)^{\exp(X^T\beta)}]] = X^T\beta + \log[-\log[S_0(t)]],$$

como $h(t|X) = h_0(t)\exp(X^T\beta)$ es equivalente a $H(t|X) = H_0(t)\exp(X^T\beta)$, y además $H(t) = -\log[S(t)]$, luego por los resultados anteriores se deduce que al graficar $-\log[S(t|X)]$ contra t, o bien $\log[-\log[S(t|X)]]$ contra $\log(t)$ las curvas obtenidas para varios niveles de X deben de ser paralelas.

2.2.1. Residuales de Cox-Snell

Los residuales son cantidades utilizadas en la verificación de modelos, cuando la función de supervivencia ha sido estimada, *los residuales de Cox-Snell* (1968) son calculados

$$rc_i = \widehat{H}(t_i) = -\log(\widehat{S}(t_i))$$

donde $\widehat{H}$ y $\widehat{S}$ son los valores estimados de la función de riesgo acumulativa y de supervivencia en el tiempo t_i.
A continuación demostraremos a partir del Teorema de Transformación que si T es la variable aleatoria asociada con el tiempo de supervivencia de un individuo, y $S(t)$ es la correspondiente función de supervivencia, entonces la variable aleatoria $Y = -\log[S(t)]$ tiene una distribución exponencial con media igual a uno.

Teorema 2.2.1. ***Teorema de Transformación****. Sea X una variable aleatoria continua con función de densidad de probabilidad $f_X(x)$ y soporte en $\mathcal{A}$, y sea $Y = g(X)$ variable aleatoria, donde g es una función diferenciable uno a uno en $\mathcal{A}$. Denotamos la inversa de g por g^{-1}.*
Entonces la función de densidad de la variable aleatoria Y está dada por

$$f_Y(y) = f_X[g^{-1}(y)] \Big/ \left|\frac{dy}{dx}\right|, \quad y \in \mathcal{B}$$

donde $\mathcal{A} = \{x : \ f_X(x) > 0\}$ y $\mathcal{B} = \{y = g(x) : \ x \in \mathcal{A}\}$.

De acuerdo al teorema anterior obtenemos que la función de densidad de probabilidad de $Y = -\log[S(t)]$ es

$$\begin{aligned} f_Y(y) &= f_T[g^{-1}(y)] \Big/ \left|\frac{dy}{dt}\right| \\ &= f_T[S^{-1}[\exp(-y)]] \Big/ \left|\frac{dy}{dt}\right|, \end{aligned}$$

donde $f_T(t)$ es la función de probabilidad de T. Por otro lado, calculamos

$$\begin{aligned} \frac{dy}{dt} &= \frac{d[-\log[S(t)]]}{dt} = -\frac{1}{S(t)} \cdot [1 - F(t)]' \\ &= \frac{f_T(t)}{S(t)} = \frac{f_T[S^{-1}[\exp(-y)]]}{S[S^{-1}[\exp(-y)]]} = \frac{f_T[S^{-1}[\exp(-y)]]}{\exp(-y)}, \end{aligned}$$

así, la función de densidad de $Y = -\log[S(t)]$ es

$$\begin{aligned} f_Y(y) &= f_T[S^{-1}[\exp(-y)]] \Big/ \left|\frac{dy}{dt}\right| \\ &= \frac{f_T[S^{-1}[\exp(-y)]] \cdot \exp(-y)}{f_T[S^{-1}[\exp(-y)]]} = \exp(-y) \end{aligned}$$

función de densidad de una variable aleatoria exponencial con media igual a uno (Collet 2003).
Este resultado nos proporciona una forma de verificar si una función de supervivencia es adecuada, si los valores $-\log(S(t_i)) \sim \exp(1)$, entonces las estimaciones, $-\log(\widehat{S}(t_i))$, se comportarán como observaciones de una distribución exponencial con media igual a uno. Por otro lado, la presencia de censura en los datos de supervivencia es posible, los residuos para este tipo de observaciones no pueden ser consideradas de la misma forma que los residuos derivados de observaciones sin censura, lo cual implica modificar los residuales de Cox-Snell para datos censurados.

Sea t_i^* el $i-$ésimo tiempo de supervivencia censurado y supongamos que t_i es el tiempo real de supervivencia pero desconocido, entonces $t_i > t_i^*$. El residual para este individuo evaluado en el tiempo de supervivencia censurada es

$$rc_i = \widehat{H}(t_i^*) = -\log(\widehat{S}(t_i^*)).$$

Observe que la función H es creciente pues la función de supervivencia S es decreciente al igual que la función $-\log$, luego $H(t_i) > H(t_i^*)$, es decir, cuanto mayor es el tiempo de supervivencia, mayor es el valor del residual de Cox-Snell. Considerando esta observación, los residuos de Cox-Snell se modifican mediante la adición de una constante positiva Δ, llamado *exceso residual*.
Los residuos de Cox-Snell modificados se definen por

$$rc_i = \begin{cases} rc_i, & \text{observaciones sin censura.} \\ rc_i + \Delta, & \text{observaciones censuradas.} \end{cases}$$

A continuación vamos a calcular un valor adecuado para Δ, para ello, primero demostramos la propiedad de falta de memoria de la distribución exponencial. Supongamos que la variable T tiene una distribución exponencial con media $1/\lambda$, y consideremos la probabilidad de que T excede $t_0 + t_1$, con $t_1 > 0$, condicionado a que T sea al menos igual a t_0, luego calculamos

$$\begin{aligned} \Pr(T \geq t_0 + t_1 | T \geq t_0) &= \frac{\Pr(T \geq t_0 + t_1, T \geq t_0)}{\Pr(T \geq t_0)} \\ &= \frac{\Pr(T \geq t_0 + t_1)}{\Pr(T \geq t_0)} = \frac{S(t_0 + t_1)}{S(t_0)}, \end{aligned}$$

como $T \sim \exp(\lambda)$, entonces $S(t) = 1 - F(t) = 1 - (1 - \exp(-\lambda t)) = \exp(-\lambda t)$, sustituyendo

$$\Pr(T \geq t_0 + t_1 | T \geq t_0) = \frac{\exp(-\lambda(t_0 + t_1))}{\exp(-\lambda t_0)} = \exp(-\lambda t_1),$$

esto significa que condicionada a la función de supervivencia al tiempo t_0, el exceso de tiempo de supervivencia más allá de t_0 también tiene una distribución exponencial con media $1/\lambda$, de este modo la probabilidad de supervivencia más allá del tiempo no se ve afectada por el conocimiento de que los individuos ya han sobrevivido en el tiempo t_0.
Por lo tanto, como $rc_i \sim \exp(1)$, entonces Δ también tendrá una distribución exponencial unitaria, esto implica que el valor esperado de Δ es igual a 1, así $i-$ésimo residual modificado Cox-Snell está definido por

$$rc_i = \begin{cases} rc_i, & \text{observaciones sin censura.} \\ rc_i + 1, & \text{observaciones censuradas.} \end{cases}$$

Una forma equivalente de evaluar el buen ajuste del modelo de supervivencia es utilizando el resultado anterior, sabemos que $rc_i \sim \exp(1)$, entonces su función de densidad es $f(rc_i) = \exp(-rc_i)$, calculando la función de riesgo acumulativa obtenemos

$$H(rc_i) = -\log(S(rc_i)) = -\log(\exp(-rc_i)) = rc_i.$$

Por lo tanto, al graficar rc_i contra la tasa de riesgo acumulada $\widehat{H}(rc_i)$ se espera una línea recta con intersección cero y pendiente igual a uno.

2.3. Método de Máxima Verosimilitud

Una de las técnicas más populares para obtener estimadores es el método de máxima verosimilitud pues elige como estimador del parámetro aquel valor que máximiza la probabilidad de obtener una muestra como la que se ha obtenido y en la que se ha basado la estimación. En esta sección definimos el concepto de función de verosimilitud y mencionaremos algunos argumentos que respaldan este método, también describiremos brevemente la prueba de Wald y el estadístico de prueba de razón de verosimilitud.

Definición. Sean $X_1, X_2, \ldots, X_n$, una muestra aleatoria con función de densidad conjunta $f(X_1, X_2, \ldots, X_n; \theta)$. Para cada muestra particular $(x_1, x_2, \ldots, x_n)$ definimos la *función de verosimilitud* como

$$L(\theta; \mathbf{x}) = f(x_1, x_2, \ldots, x_n; \theta).$$

En particular, si los $X_1, X_2, \ldots, X_n$ son independientes e idénticamente distribuidas (iid) con función de densidad común $f(x; \theta)$, la función de verosimilitud se reduce a

$$L(\theta; \mathbf{x}) = \prod_{i=1}^{n} f(x_i; \theta).$$

Debido a que la función logarítmica es monótona, maximizar la función de verosimilitud es lo mismo que maximizar la función log-verosimilitud

$$\log L(\theta) = \sum_{i=1}^{n} \log f(x_i; \theta),$$

esta última suele ser más conveniente de usar.

Definición. Sea X una variable aleatoria con función de distribución acumulada $F(x, \theta)$. Sea $x_1, x_2, \ldots, x_n$, una muestra de la distribución de X y sea T_n un estadístico. Decimos que T_n es un *estimador consistente* de θ si

$$\lim_{n\to\infty} \Pr(|T_n - \theta| < \epsilon) = 1 \quad \text{para todo} \;\; \epsilon > 0$$

T_n converge en probabilidad al verdadero valor del parámetro.

La justificación teórica para considerar el estimador de máxima verosimilitud es porque cumple las siguientes propiedades:

1. Sea θ el verdadero valor del parámetro, bajo condiciones de regularidad se tiene que

   $$\lim_{n\to\infty} \Pr_{\theta}(L(\theta, x) > L(\theta_0, x)) = 1 \quad \text{para toda} \;\; \theta \neq \theta_0.$$

   La interpretación de esta propiedad es, de forma asintótica la función de verosimilitud se maximiza en el verdadero valor θ, por esta razón al considerar estimaciones de θ parece natural elegir el valor de θ_0 que maximiza la probabilidad.

2. El estimador de máxima verosimilitud $\widehat{\theta_n}$ es consistente del verdadero valor θ.
   Esta propiedad se cumple si se consideran las siguientes condiciones: la muestra $X_1, X_2, \ldots, X_n$, satisface ciertas condiciones de regularidad, $f(x; \theta)$ es diferenciable respecto de θ y la función de verosimilitud tiene una única solución $\widehat{\theta_n}$.

3. Los estimadores de máxima verosimilitud siguen asintóticamente una distribución normal.
   Bajo condiciones de regularidad y suponiendo que la información de Fisher satisface $0 < I(\theta) < \infty$, la aproximación asintótica de la distribución muestral del estimador de máxima verosimilitud es normal multivariada con media θ (verdadero valor desconocido del parámetro) y varianza $I(\theta)^{-1}$.
   En la práctica, este resultado no es útil porque no conocemos el verdadero valor del parámetro, si lo conociéramos entonces no lo estaríamos estimando. Sin embargo, podemos aplicar el principio de plug-in y caracterizar la varianza de la distribución asintótica usando $I(\widehat{\theta})^{-1}$ o $[- \bigtriangledown^2 \log L(\widehat{\theta})]^{-1}$, donde $\bigtriangledown^2 f$ representa la matriz de segundas derivadas parciales de una función escalar f.

Construcción de la función de verosimilitud para datos censurados.

Como se indicó en el capítulo anterior, los tiempos de supervivencia pueden involucrar censura y estos deben considerarse cuidadosamente al construir funciones de probabilidad. Una suposición fundamental es que los tiempos de supervivencia y de censura sean independientes. Para construir la función de verosimilitud cuando hay datos censurados debemos considerar qué información nos brinda cada observación, por ejemplo, una observación correspondiente al tiempo exacto en que ocurrió el evento proporciona información sobre la probabilidad de que el evento ocurra en ese tiempo, que es igual a la función de densidad f de X en ese momento, por otro lado, para una observación censurada por la derecha sabemos que el tiempo en que se presentará el evento será mayor al censurado.
Presentamos a continuación el esquema de los casos anteriores: sea (T, δ) una variable aleatoria bivariada, donde δ indica si el tiempo de supervivencia X es observado ($\delta = 1$) o no ($\delta = 0$), con $T = X$ si el tiempo de vida es observado y $T = C$ si está censurada a la derecha, para este último caso los tiempos son considerados desde el inicio hasta el final del estudio, o último tiempo de contacto con el individuo, así $T = \text{mín}(X, C)$.
Para la construcción de la función de verosimilitud hacemos lo siguiente: denotamos las funciones de densidad de X y C, respectivamente, por $f(x)$ y $g(c)$. Suponemos que los tiempos de supervivencia X y los tiempos de censura C son independientes, luego calculamos la función de distribución acumulada de los datos no-censurados y censurados

1. Cuando $\delta = 1$, obtenemos

$$\begin{aligned} F(t,1) &= \Pr(T \leq t, \delta = 1) = \Pr(X \leq t, X \leq C) \\ &= \int_{x \leq t} \int_{x \leq c} f(x) \cdot g(c)\ dx\ dc \\ &= \int_{-\infty}^{t} f(x) \left(\int_{x}^{\infty} g(c)\ dc \right)\ dx \end{aligned}$$

   En consecuencia la función de densidad es $f(x)\left(\int_x^\infty g(c)\ dc\right)$, puesto que la censura a la derecha es no informativa se deduce que la contribución a la función de verosimilitud es $f(x)$.

2. Cuando $\delta = 0$,

$$\begin{aligned} F(t,0) &= \Pr(T \leq t, \delta = 0) = \Pr(C \leq t, X > C) \\ &= \int_{c \leq t} \int_{x > c} f(x) \cdot g(c)\ dx\ dc \\ &= \int_{-\infty}^{t} \left(\int_{c}^{\infty} f(x) \cdot g(c)\ dx \right)\ dc \\ &= \int_{-\infty}^{t} g(c) \cdot \left(\int_{c}^{\infty} f(x)\ dx \right)\ dc \\ &= \int_{-\infty}^{t} g(c) \cdot (S(c))\ dc \end{aligned} \tag{2.1}$$

Por lo tanto la función de densidad es $g(t) \cdot (S(c))$, como la censura es no informativa entonces se sigue que la contribución a la verosimilitud es $S(c)$.
Dada una muestra aleatoria (T_i, δ_i), con $i = 1, 2, \ldots, n$, la función de verosimilitud será el producto de las funciones f y S evaluados en sus respectivos tiempos

$$L = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} \cdot [S(t_i)]^{1-\delta_i}.$$

Prueba de Wald.

En nuestro análisis estaremos interesados en examinar si existe una relación significativa entre la variable dependiente y las variables independientes contenidas en el modelo logístico, la *prueba de Wald* es una de varias formas de probar si los parámetros asociados con un grupo de variables son cero. La distribución del estadístico de Wald es una herramienta que permite aceptar o rechazar la hipótesis nula establecida sobre el $i-$ ésimo parámetro (β_i):

$$H_0 : \beta_i = 0 \qquad \text{vs} \qquad H_A : \beta_i \neq 0.$$

Para calcular la prueba de Wald debemos calcular el estadístico de Wald para la variable en cuestión, dicho estadístico se obtiene dividiendo la estimación de máxima verosimilitud

del parámetro, $\widehat{\beta_n}$, por el estimador de su error éstandar:

$$Z = \frac{|\widehat{\beta_n}|}{se(\widehat{\beta_n})} = \frac{|\widehat{\beta_n}|}{\sqrt{V(\widehat{\beta_n})}},$$

bajo la hipótesis nula este estadístico se distribuye normal con media 0 y varianza igual a 1, $Z \sim N(0,1)$.
Sea $z_{\alpha/2}$ único punto tal que $\Pr(|Z| > z_{\alpha/2}) = \alpha$, entonces la prueba de Wald de tamaño α (nivel de significancia) rechaza la hipótesis nula cuando $|Z| > z_{\alpha/2}$.

Estadístico de prueba de razón de verosimilitud.

Una forma de evaluar la adecuación de un modelo es compararlo con un modelo más general con el número máximo de parámetros que se pueden estimar, *prueba de razón de verosimilitud*, ésta es una prueba de hipótesis utilizada para comparar la bondad de ajuste de dos modelos anidados: un modelo nulo contra un modelo alternativo. Dicha prueba se basa en la razón de verosimilitudes que especifica cuántas veces es más probable que los datos estén bajo un modelo que de otro, esta razón o de forma equivalente su logaritmo es una herramienta para decidir si se rechaza o no el modelo nulo.

Antes de definir el estadístico de prueba de razón de verosimilitud, escribimos el espacio de parámetro Ω como una partición disjunta, $\Omega = \Omega_0 \cap \Omega_0^c$, luego la hipótesis que queremos constrastar es $H_0 : \theta \in \Omega_0$.

Definición. *El estadístico de prueba de razón de verosimilitud* para probar $H_0 : \theta \in \Omega_0$ (modelo reducido) contra H_A (modelo completo) está dado por

$$r(x) = \frac{\sup_{\Omega_0} L(\theta)}{\sup_{\Omega} L(\theta)}$$

donde $L(\theta)$ es la función de verosimilitud.
$r(x)$ es un valor que toma valores entre cero y uno, un valor de $r(x)$ cercano a uno indica que H_0 es aceptable, en caso contrario rechazamos la hipótesis de H_0, equivalentemente

rechazamos H_0 para valores grandes de

$$r^*(x) = -2\log[r(x)].$$

Un resultado importante acerca del estadístico de prueba r^* es el siguiente: bajo $H_0 : \theta \in \Omega_0$, la distribución de r^* converge a una chi-cuadrada $\chi^2_{(df)}$ cuando $n \to \infty$, donde los grados de libertad df = núm. de parámetros libres en Ω−núm. de parámetros libres en Ω_0. Por lo tanto, una prueba de tamaño α aproximada es rechazar H_0 si

$$r^* = -2\log r(x) \geq \chi^2_\alpha,$$

donde χ^2_α es el α−ésimo cuantil superior de una chi-cuadrada con df grados de libertad.

2.4. Cópulas

La noción de cópula fue introducido por Abe Sklar en 1959 como respuesta al problema planteado por Fréchet sobre la relación entre una función de distribución de probabilidades multivariada y sus distribuciones marginales de menor dimensión. Puede consultar los detalles de la historia en (Burney y Bin 2020).
La idea intuitiva sobre las cópulas es juntar n funciones de distribución unidimensionales F_i, con el fin de construir una función de distribución conjunta n−dimensional con marginales F_i. La ventaja del uso de las cópulas es que pueden extraer la estructura de dependencia de la función de distribución conjunta de un vector de variables aleatorias.
Dedicamos esta sección al concepto de cópulas, a su vez mostramos el teorema de Sklar, el cual establece que cualquier función de distribución bivariada puede ser descrita en términos de dos distribuciones marginales y una cópula.

Definición. Una *cópula bidimensional* $\mathcal{C}$ es una función de distribución bivariada definida sobre el cuadrado unitario $[0,1]\times[0,1]$ con distribuciones marginales uniformes $U(0,1)$.

El principal motivo del uso de las cópulas proviene del ***teorema de Sklar***, en dicho teorema se menciona lo siguiente:

Teorema de Sklar. *Dadas las variables aleatorias Y_1, Y_2, con marginales continuas F_1, F_2, respectivamente, y función de distribución conjunta F, entonces existe una única cópula $\mathcal{C}$ tal que para todo y_1, $y_2 \geq 0$ satisface*

$$F(y_1, y_2) = \mathcal{C}[F_1(y_1), F_2(y_2)] \tag{2.2}$$

Inversamente, si $\mathcal{C}$ es una cópula y, F_1 y F_2 son funciones de distribución, entonces F definida en la ecuación (2.2) *es una función de distribución conjunta con marginales F_1 y F_2.*

La idea de la demostración de este teorema se basa en probar la igualdad de la ecuación (2.2) en un dominio reestringido de la cópula, luego se extiende este resultado para los datos faltantes, la demostración puede ser consultada en (Nelsen 2006, sección 2.3).

El teorema de Sklar también nos brinda una posibilidad de relación entre el parámetro de dependencia de la cópula y la dependencia de las variables Y_1 y Y_2. Para obtener y entender dicha relación primeramente demostramos que las cópulas son invariantes bajo transformaciones estrictamente crecientes, esto significa que las propiedades y medidas no se modifican cuando se realizan dichas transformaciones. A continuación demostramos dicho teorema.

Teorema 2.4.1. *Sea X y Y variables aleatorias continuas con cópula $\mathcal{C}_{XY}$. Si α y β son funciones estrictamente crecientes en el rango de X y rango de Y, respectivamente, entonces $\mathcal{C}_{\alpha(X)\beta(Y)} = \mathcal{C}_{XY}$. Así $\mathcal{C}_{XY}$ es invariante bajo transformaciones estrictamente crecientes de X y Y.*

Demostración. Sean F_1, F_2, G_1 y G_2 las funciones de distribución de X, Y, $\alpha(X)$ y $\beta(Y)$, respectivamente. Como α y β son estrictamente crecientes, entonces se cumple

$$\begin{aligned} G_1(x) &= \Pr(\alpha(X) \leq x) = \Pr(X \leq \alpha^{-1}(x)) = F_1(\alpha^{-1}(x)) \\ G_2(y) &= \Pr(\beta(Y) \leq y) = \Pr(Y \leq \beta^{-1}(y)) = F_2(\beta^{-1}(y)), \end{aligned}$$

sustituyendo en el siguiente resultado

$$\begin{aligned}\mathcal{C}_{\alpha(X)\beta(Y)}(G_1(x), G_2(y)) &= F_{\alpha(X)\beta(Y)}(x, y) = \Pr(\alpha(X) \leq x, \beta(Y) \leq y) \\ &= \Pr(X \leq \alpha^{-1}(x), Y \leq \beta^{-1}(y)) \\ &= F_{XY}(\alpha^{-1}(x), \beta^{-1}(y)) \\ &= C_{XY}(F_1(\alpha^{-1}(x)), F_2(\beta^{-1}(y))) \\ &= C_{XY}(G_1(x), G_2(y)).\end{aligned}$$

Puesto que X y Y son continuas, entonces el rango de F_1 y F_2 es igual a $\mathbf{I}$, por lo tanto $\mathcal{C}_{\alpha(X)\beta(Y)} = \mathcal{C}_{XY}$ en $\mathbf{I}^2$. □

En nuestro estudio, la función de distribución conjunta será modelada por la cópula Gaussiana o por cópulas contenidas en la familia de cópulas arquimedianas, estas últimas definidas por Genest y MacKay utilizando una función generatriz (Genest y MacKay 1986). Las cópulas arquimedianas tienen propiedades deseables pues se construyen utilizando una función generatriz que facilita la mayoría de los cálculos.
A continuación definimos estas cópulas y presentamos algunos ejemplos que utilizaremos para la estimación paramétrica.

Definición. Sea $\Psi : [0, 1] \to [0, \infty]$ una función continua y estrictamente decreciente, tal que $\Psi(1) = 0$. Definimos la pseudo inversa de Ψ como $\Psi^{(-1)} : [0, \infty] \to [0, 1]$ que satisface

$$\Psi^{(-1)}(u) = \begin{cases} \Psi^{-1}(u), & 0 \leq u \leq \Psi(0), \\ 0, & \text{en otro caso.} \end{cases}$$

La función $\mathcal{C} : \mathbf{I} \times \mathbf{I} \to \mathbf{I}$ definida por

$$\mathcal{C}(u, v) = \Psi^{(-1)}(\Psi(u) + \Psi(v)) \tag{2.3}$$

es una cópula si y sólo si Ψ es convexa, satisfaciendo Ψ las propiedades de la definición anterior, la función Ψ es llamada un *generador* de la cópula. Las cópulas generadas por la ecuación (2.3) se denominan *cópulas arquimedianas*, algunos ejemplos son los siguientes:

Ejemplo 2.4.1. ***Cópula Gumbel****. La función generatriz de esta cópula está dada por*

$$\Psi(u) = (-\ln u)^\theta, \quad \theta \geq 1,$$

con inversa $\Psi^{-1}(x) = \exp(-x^{1/\theta})$*, luego por* (2.3) *la cópula se define*

$$\mathcal{C}(u, v; \theta) = \Psi^{[-1]}(\Psi(u) + \Psi(v)) = \exp(-[(-\ln u)^{\theta} + (-\ln v)^{\theta}]^{1/\theta}). \tag{2.4}$$

Ejemplo 2.4.2. ***Cópula Clayton****. Se caracteriza por tener como función generatriz*

$$\Psi(u) = \frac{1}{\theta}(u^{-\theta} - 1), \quad \theta \in (0, \infty)$$

con función inversa $\Psi^{-1}(x) = (1 + \theta \cdot x)^{-1/\theta}$*, así la cópula está definida por*

$$\mathcal{C}(u, v; \theta) = \left(1 + \theta \left(\frac{1}{\theta}(u^{-\theta} - 1 + v^{-\theta} - 1)\right)\right)^{-1/\theta} = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta} \tag{2.5}$$

donde el parámetro de dependencia θ *toma valores en el intervalo* $(0, \infty)$*.*

Como mencionamos anteriormente, trabajaremos también con la cópula Gaussiana.

Ejemplo 2.4.3. ***Cópula Gaussiana****. La cópula Gaussiana bivariada con parámetro* $\rho \in (-1, 1)$ *está definida por*

$$\mathcal{C}(u, v; \rho) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v)) \tag{2.6}$$

donde dadas las funciones de densidad normal estándar

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad y \quad \varphi_2(x, y; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right)$$

para el caso univariado y bivariado, respectivamente, definimos la función de distribución de la normal estándar

$$\Phi(h) = \int_{-\infty}^{h} \varphi(x)dx$$

y la función de distribución normal estándar bivariada con parámetro de correlación $\rho \in (-1, 1)$*.*

$$\Phi_2(h, k; \rho) = \int_{-\infty}^{h} \int_{-\infty}^{k} \varphi_2(x, y; \rho)dy\ dx.$$

Puesto que $\mathcal{C}$ es una función de distribución, es usual pensar sobre la función de densidad de la cópula, seguidamente presentamos dicha función.

Para calcular la función de densidad de la cópula suponemos que $F_1(y_1)$, $F_2(y_2)$ y la cópula $\mathcal{C}(u, v)$ deben ser diferenciables, así derivando la ecuación (2.2) con respecto a ambas variables tenemos la siguiente expresión

$$\frac{\partial^2 F(y_1, y_2)}{\partial y_1 \partial y_2} = \frac{\partial^2 \mathcal{C}[F_1(y_1), F_2(y_2)]}{\partial F_1(y_1) \partial F_2(y_2)} \cdot \frac{\partial F_2(y_2)}{\partial y_2} \cdot \frac{\partial F_1(y_1)}{\partial y_1}$$
$$f(y_1, y_2) = c[F_1(y_1), F_2(y_2)] \cdot f_2(y_2) \cdot f_1(y_1)$$

donde $c[F_1(y_1), F_2(y_2)] = \dfrac{\partial^2 \mathcal{C}[F_1(y_1), F_2(y_2)]}{\partial F_1(y_1) \partial F_2(y_2)}$ es la *función de densidad de la cópula* $\mathcal{C}$.

Por lo tanto, dado la densidad de la cópula y las funciones de densidad marginales, si existen, es posible calcular la función de densidad conjunta a partir de la ecuación anterior.

Ejemplo 2.4.4. ***Función de densidad de la cópula Gaussiana.*** *Utilizando el resultado anterior tenemos que*

$$c[F_1(y_1), F_2(y_2)] = \frac{f(y_1, y_2)}{f_1(y_1) \cdot f_2(y_2)}$$

Sustituyendo los respectivos valores obtenemos la función de densidad de la cópula Gaussiana

$$\begin{aligned} c(u, v; \rho) &= \frac{\varphi_2(\Phi^{-1}(u), \Phi^{-1}(v); \rho)}{\varphi(\Phi^{-1}(u)) \cdot \varphi(\Phi^{-1}(v))} = \frac{\varphi_2(\varepsilon_1, \varepsilon_2; \rho)}{\varphi(\varepsilon_1) \cdot \varphi(\varepsilon_2)} \\ &= \left[\frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{\varepsilon_1^2 - 2\rho\varepsilon_1\varepsilon_2 + \varepsilon_2^2}{2(1-\rho^2)} \right) \right] \Big/ \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\varepsilon_1^2}{2} \right) \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\varepsilon_2^2}{2} \right) \right] \\ &= \frac{1}{\sqrt{1-\rho^2}} \exp\left(\frac{2\rho\varepsilon_1\varepsilon_2 - \rho^2(\varepsilon_1^2 + \varepsilon_2^2)}{2(1-\rho^2)} \right) \end{aligned}$$

donde $\varepsilon_1 = \Phi^{-1}(u)$ *y* $\varepsilon_2 = \Phi^{-1}(v)$.

En el estudio utilizaremos la derivada parcial de cada una de las cópulas definidas anteriormente, en particular, mostraremos este resultado para la cópula Gaussiana.

$$\frac{\partial}{\partial u} C(u, v; \rho) = \Phi\left[\frac{\Phi^{-1}(v) - \rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}} \right],$$

donde $\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx$ y ρ es el parámetro de dependencia.
Utilizando la regla de Leibniz

$$\frac{\partial}{\partial u} C(u, v; \rho) = \frac{d}{du}\Phi^{-1}(u) \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{\Phi^{-1}(u)^2 - 2\rho\Phi^{-1}(u)y + y^2}{2(1-\rho^2)}\right) dy$$

Resolviendo solo la integral

$$\begin{aligned}
&\int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2} \cdot \frac{[y - \rho\Phi^{-1}(u)]^2 + (1-\rho^2)\Phi^{-1}(u)^2}{1-\rho^2}\right] dy \\
&\qquad = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{\Phi^{-1}(u)^2}{2}\right] \int_{-\infty}^{\Phi^{-1}(v)} \exp\left[-\frac{1}{2} \cdot \frac{[y - \rho\Phi^{-1}(u)]^2}{1-\rho^2}\right] dy \\
&\qquad = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{\Phi^{-1}(u)^2}{2}\right] \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\Phi^{-1}(v)} \exp\left[-\frac{\left(\frac{y-\rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}\right)^2}{2}\right] \frac{1}{\sqrt{1-\rho^2}} dy
\end{aligned}$$

Haciendo cambio de variable en la integral, $w = \frac{y-\rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}$ implica $dw = \frac{1}{\sqrt{1-\rho^2}} dy$, si y tiende a $-\infty$, entonces w tiende a $-\infty$, además si y tiende a $-\Phi^{-1}(v)$, entonces w tiende a $\frac{\Phi^{-1}(v)-\rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}$, así

$$\begin{aligned}
\frac{\partial}{\partial u} C(u, v; \rho) &= \frac{d}{du}\Phi^{-1}(u) \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{\Phi^{-1}(u)^2}{2}\right] \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\Phi^{-1}(v)-\rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}} \exp\left[-\frac{w^2}{2}\right] dw \\
&= \frac{d}{du}\Phi^{-1}(u) \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{\Phi^{-1}(u)^2}{2}\right] \Phi\left(\frac{\Phi^{-1}(v) - \rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}\right)
\end{aligned}$$

Por otro lado calculamos

$$\frac{d}{du}\Phi^{-1}(u) = \frac{1}{\Phi'(\Phi^{-1}(u))} = \frac{1}{\frac{1}{\sqrt{2\pi}} \cdot \exp\left[-\frac{\Phi^{-1}(u)^2}{2}\right]}.$$

Sustituyendo se deduce que

$$\frac{\partial}{\partial u} C(u, v; \rho) = \Phi\left[\frac{\Phi^{-1}(v) - \rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}\right].$$

Medidas de Concordancia.

La relación de dependencia estadística entre variables aleatorias es uno de los temas más estudiados, como veremos en esta sección, una forma natural de medir la dependencia es mediante el uso de la cópula.
Como consecuencia directa del Teorema 2.4.1, las cópulas tienen la propiedad de ser invariantes bajo transformaciones estrictamente crecientes de las variables aleatorias, esto induce la idea de relacionar medidas de dependencia invariantes por escala con las cópulas, como la concordancia. Dos medidas utilizadas en este tipo de análisis y basados en la concordancia son: la tau de Kendall y la rho de Spearman, en esta investigación sólo nos enfocamos en estudiar la relación entre la tau de Kendall y la cópula, comenzaremos a definir el concepto de concordante.

Una forma en la que se puede estudiar la asociación de dos variables aleatorias es mediante una medida de concordancia (tipo de correlación no paramétrica), decimos que dos observaciones (x_i, y_i) y (x_j, y_j) de un vector (X, Y) de variables aleatorias continuas son *concordantes* si

$$x_i < x_j \text{ y } y_i < y_j, \text{ o } x_i > x_j \text{ y } y_i > y_j \Leftrightarrow (x_i - x_j)(y_i - y_j) > 0.$$

Similarmente decimos que la pareja es *discordante* si

$$x_i < x_j \text{ y } y_i > y_j, \text{ o } x_i > x_j \text{ y } y_i < y_j \Leftrightarrow (x_i - x_j)(y_i - y_j) < 0.$$

En otras palabras, un vector aleatorio es concordante cuando a valores grandes (pequeños) de una variable están asociadas a grandes (pequeños) valores de la otra variable aleatoria. De la misma manera se obtiene algo análogo al concepto discordante.
Una medida de asociación usual basada en la concordancia y discordancia es la tau de Kendall. La versión poblacional de la *Tau de Kendall* para un vector (X, Y) de variables aleatorias continuas con función de distribución F se define de la siguiente forma.

Definición. Sean (X_1, Y_1) y (X_2, Y_2) vectores aleatorios continuos independientes e idénticamente distribuidos con función de distribución conjunta F. Definimos la *Tau de Kendall* (τ) para el vector aleatorio (X, Y) como la probabilidad de concordancia menos la pro-

babilidad de discordancia

$$\tau = \Pr[(X_1 - X_2)(Y_1 - Y_2) > 0] - \Pr[(X_1 - X_2)(Y_1 - Y_2) < 0].$$

En el siguiente teorema se muestra la importancia que tienen las cópulas en la concordancia.

Teorema 2.4.2. *Sean (X_1, Y_1) y (X_2, Y_2) dos vectores aleatorios continuos independientes e idénticamente distribuidos cuyas marginales son F y G, y cuya cópula asociada es $\mathcal{C}$, entonces*

$$\tau = 4 \int \int_{\mathbf{I}^2} \mathcal{C}(u, v) d\mathcal{C}(u, v) - 1.$$

Demostración. Puesto que los vectores aleatorios son continuos, entonces se deduce la siguiente igualdad

$$\Pr[(X_1 - X_2)(Y_1 - Y_2) < 0] = 1 - \Pr[(X_1 - X_2)(Y_1 - Y_2) > 0],$$

sustituyendo este resultado en la definición poblacional de la Tau de Kendall obtenemos

$$\tau = 2 \Pr[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1$$

Observe que

$$\Pr[(X_1 - X_2)(Y_1 - Y_2) > 0] = \Pr[X_1 < X_2, Y_1 < Y_2] + \Pr[X_1 > X_2, Y_1 > Y_2]$$

Dado este resultado calculamos

$$\Pr[X_1 < X_2, Y_1 < Y_2] = \int \int_{\mathbb{R}} \mathcal{C}(F(x), G(y)) d\mathcal{C}(F(x), G(y))$$

como F y G son funciones de distribución se sigue que el rango de dichas funciones toma valores en el intervalo $(0, 1)$, haciendo la transformación $u = F(x)$ y $v = F(y)$ tenemos

$$\Pr[X_1 < X_2, Y_1 < Y_2] = \int \int_{\mathbf{I}^2} \mathcal{C}(u, v) d\mathcal{C}(u, v)$$

Concluimos que

$$\begin{aligned}\tau &= 2\Pr[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1 \\ &= 2[\Pr(X_1 < X_2, Y_1 < Y_2) + \Pr(X_1 > X_2, Y_1 > Y_2)] - 1 \\ &= 2\left(2\int\int_{\mathbf{I}^2} \mathcal{C}(u,v)d\mathcal{C}(u,v)\right) - 1 \\ &= 4\int\int_{\mathbf{I}^2} \mathcal{C}(u,v)d\mathcal{C}(u,v) - 1.\end{aligned}$$

□

Así la τ de Kendall está completamente determinada por la cópula y no por las distribuciones marginales. Por otro lado, una de las propiedades que tiene la τ de Kendall es que toma valores en el intervalo $[-1, 1]$, si este valor es aproximadamente cero la relación de las variables será débil. Un resultado importante es que si dos variables aleatorias X y Y son independientes, entonces $\tau = 0$ (Nelsen 2006, sección 5.1.1).
Para las cópulas Arquimedianas la τ de Kendall se expresa de una forma más sencilla utilizando la función generatriz Ψ (Genest y MacKay 1986),

$$\tau = 1 + 4\int_0^1 \frac{\Psi(t)}{\Psi'(t)}dt,$$

en particular, la τ para la cópula Gumbel es

$$\begin{aligned}\tau &= 1 + 4\int_0^1 \frac{t\cdot\ln(t)}{\theta}\,\mathrm{dt} = 1 + \frac{4}{\theta}\left(\left[\frac{t^2\cdot\ln(t)}{2}\right|_0^1 - \int_0^1 \frac{t}{2}\,\mathrm{dt}\right) \\ &= 1 + \frac{4}{\theta}\left(0 - \frac{1}{4}\right) = 1 - \frac{1}{\theta}\end{aligned}$$

Análogamente obtenemos la τ de Kendall de la cópula Clayton:

$$\begin{aligned}\tau &= 1 + 4\int_0^1 \frac{t^{\theta+1} - t}{\theta}\,\mathrm{dt} = 1 + \frac{4}{\theta}\left(\frac{1}{\theta+2} - \frac{1}{2}\right) \\ &= 1 - \frac{2}{\theta+2} = \frac{\theta}{\theta+2}\end{aligned}$$

Capítulo 3

Datos de cáncer de colon

Acorde a la página oficial de la OMS (Organización Mundial de la Salud), el cáncer es una de las causas principales de muerte en todo el mundo, en el año 2020 se reportaron aproximadamente 10 millones de muertes causadas por dicha enfermedad. El cáncer colorrectal es el tercer cáncer más común y la segunda causa más importante de muerte relacionada con el cáncer. Investigadores han encontrado factores de riesgo que podrían aumentar la probabilidad de que una persona desarrolle cáncer de colon y están relacionados con los malos hábitos alimenticios, obesidad, tabaquismo, enfermedad inflamatoria intestinal, pólipos, consumo de alcohol, factores genéticos y envejecimiento.
En este capítulo proporcionamos información esencial sobre el cáncer de colon, además presentamos la aplicación de la teoría antes vista a una base de datos de pacientes diagnósticados con cáncer de colon en etapa C de Dukes (estadificación que indica la diseminación del cáncer a al menos un ganglio linfático).

3.1. Cáncer de Colon

El cuerpo humano está compuesto por billones de células que normalmente crecen y se multiplican (proceso llamado división celular) formando células que el cuerpo necesita, cuando las células son anormales, mueren y nuevas células ocupan su lugar. Cuando algo sale mal en este proceso las células crecen sin control, dichas células forman masas de tejido conocidas como tumores, las cuales se clasifican en dos tipos: cancerosos (malignos) o no cancerosos (benignos).
Los tumores cancerosos pueden invadir los tejidos cercanos, con frecuencia, pueden desplazarse a otras partes del cuerpo para formar nuevos tumores (proceso de metástasis).

El cáncer es una enfermedad caracterizada por la división incontrolada de células anormales, cuando este proceso ocurre en el colon o recto se denomina *cáncer colorrectal*, también llamado *cáncer de colon* o *cáncer de recto*, estos comúnmente se agrupan porque tienen características comunes.
El colon y el recto forman el intestino grueso, que es parte del sistema digestivo, la función del colon es absorber agua y sal de la materia alimentaria luego la materia de desecho restante va directo al recto.

Pólipos.

La mayoría de los cánceres de colon comienzan como un pólipo (ver figura 3.1), un tumor benigno que se desarrolla en la capa mucosa (revestimento interno) del colon. Los pólipos son comunes, algunos de ellos pueden convertirse en cáncer con el tiempo, la probabilidad de que un pólipo se convierta en cáncer depende del tipo de pólipo, algunos de ellos son los siguientes:

a) **Pólipos Adenomatosos (adenomas).** Estos pólipos a veces se transforman en cáncer por lo cual se denominan condición precancerosa. Los tres tipos de adenomas son tubulares, vellosos y tubulo-vellosos.

b) **Pólipos hiperplásticos e inflamatorios.** Estos pólipos son cómunes, en general no son precancerosos (no existe riesgo de provocar cáncer).

c) **Pólipos serrados sésiles y adenomas serrados tradicionales.** Estos pólipos son menos comunes pero tienen un mayor riesgo de convertirse en cáncer colorrectal.

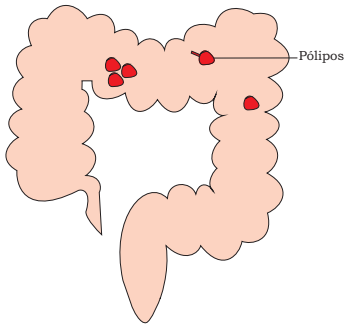


Figura 3.1: Pólipos en el colon.

Etapas del cáncer de colon.

Una vez que el pólipo se convierte en cáncer puede crecer hacia la pared del colon o el recto donde pueden invadir los vasos sanguíneos o linfáticos, generalmente se diseminan primero a los ganglios linfáticos cercanos (estructura en forma de frijol que forma parte del sistema inmunitario, estos ayudan a combatir infecciones y enfermedades), también pueden desplazarse a otros organos como: el hígado y los pulmones a través de los vasos sanguíneos. El grado de diseminación del cáncer en el momento del diagnóstico se le conoce como proceso de estadificación, la estadificación es esencial para determinar la gravedad del cáncer y las opciones de tratamiento. Uno de las sistemas utilizados para medir la estadificación del cáncer del colon es la escala A, B, C, D de Dukes propuesto en 1932. Consultar la referencia electrónica (Cancer research UK. 2018) para más detalles.

a) **A de Dukes.** El cáncer se encuentra en el revestimiento interno (capa mucosa) del intestino o está creciendo lentamente hacia la capa muscular (Cancer research UK. 2018).

b) **B de Dukes.** El cáncer ha crecido a través de la capa muscular del intestino (Cancer research UK. 2018).

c) **C de Dukes.** El cáncer se ha diseminado al menos a un ganglio linfático cercano al intestino (Cancer research UK. 2018).

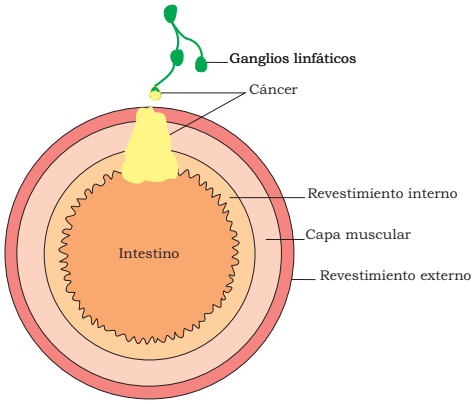


Figura 3.2: Etapa C del sistema Dukes.

d) **D de Dukes.** También llamado cáncer de intestino avanzado y es la etapa donde el cáncer se ha diseminado a otras partes del cuerpo, como el hígado, pulmón, la pared abdominal o los huesos (Cancer research UK. 2018).

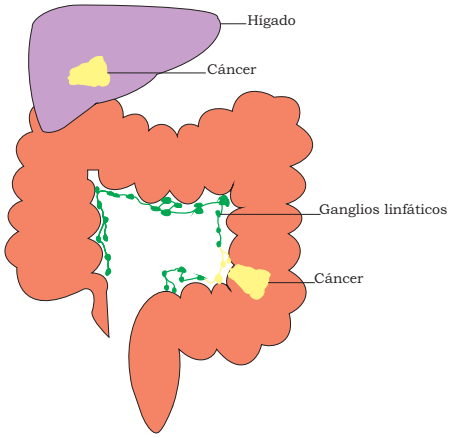


Figura 3.3: Etapa D del sistema Dukes.

Tratamiento.

El tratamiento para el cáncer de colon se basa principalmente en la estadificación (describe la cantidad de cáncer que hay en el cuerpo y cuánto se ha propagado), además de otros factores como el tipo de células en las que comenzó el cáncer y las condiciones de salud de cada individuo. Personas con cáncer de colon que aún no se ha propagado a sitios distantes generalmente se someten a cirugía como tratamiento principal o inicial.
Los principales tratamientos para el cáncer de colon son la *cirugía* y *quimioterapia*.

En la mayoría de los casos es poco probable que una cirugía cure los cánceres en etapa avanzada, pero puede disminuir el riesgo de muerte cuando hay áreas pequeñas de propagación del cáncer en el hígado o pulmones, por lo general, se recomienda quimioterapia después de la cirugía (llamado *tratamiento adjuvante*) cuando el cáncer tiene mayor riesgo de reaparición.
La quimioterapia usa medicamentos contra el cáncer, estos circulan por todo el cuerpo a través del torrente sanguíneo que permite destruir las células cancerosas. Para los cánceres de colon avanzando que no se pueden extirpar totalmente mediante la cirugía,

se recomienda la *quimioterapia neoadjuvante* (cirugía para trata de reducir el tamaño del cáncer) administrada junto con radiación para reducir el tamaño del cáncer y a su vez extiparlo con cirugía. Si el cáncer se ha extendido demasiado, es decir, es muy grande o hay demasiados, se puede administrar quimioterapia antes de la cirugía, si mediante este proceso los tumores se encogen entonces se puede intentar nuevamente una cirugía para extirparlos. Para una explicación explícita de los tipos de tratamientos, ver la referencia electrónica (Colorectal Cancer 2021).

3.2. Metodología Estadística

A continuación describiremos una metodología totalmente paramétrica que modela la distribución conjunta de los tiempos de supervivencia secuenciales T_1 y T_2, donde T_2 es observada si T_1 no está censurada. Para modelar dicha distribución utilizamos distintas funciones cópula y así poder descartar el supuesto de independencia entre dichas variables.

Sea (T_{1i}, T_{2i}) los tiempos de supervivencia secuenciales, donde T_2 *es observada sólo si es observada* T_1, y sean C_{1i}, C_{2i}, los tiempos de censura correspondientes a cada uno de los tiempos de supervivencia. El tiempo C_{1i} se obtiene de forma análoga al caso univariado, por otro lado, C_{2i} se obtiene restando el primer tiempo de supervivencia (T_{1i}) del último tiempo observado para cada individuo i.
Sean $t_{ki} = \text{mín}(T_{ki}, C_{ki})$ los tiempos de supervivencia observados y sus indicadores de censura $\delta_{ki} = I(T_{ki} \leq C_{ki})$ para $k = 1, 2$.
Las distribuciones marginales de T_1, T_2 están dadas por $F_1(t_1) = F(t_1, \infty)$ y $F_2(t_2) = F(\infty, t_2)$, por otro lado, las funciones de supervivencia marginal son $S_1(t_1) = S(t_1, 0)$ y $S_2(t_2) = S(0, t_2)$, respectivamente, con $t_1, t_2 \geq 0$.
Para construir la función de verosimilitud consideramos cada uno de los casos que se presentan en el problema, además supondremos que los tiempos de supervivencia son independientes a los tiempos de censura. Para ver los detalles de los cálculos consulte el apéndice A.

Caso 1. Si T_1 y T_2 son observadas, la contribución a la función de verosimilitud es

$$f(t_{1i}, t_{2i}) = \frac{\partial^2 F(t_{1i}, t_{2i})}{\partial t_{1i} \partial t_{2i}}.$$

Caso 2. Si T_1 es observada y T_2 es censurada, entonces la contribución a la función de verosimilitud es

$$-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}}$$

por el resultado $S(t_{1i}, t_{2i}) = 1 - F_1(t_{1i}) - F_2(t_{2i}) + F(t_{1i}, t_{2i})$, obtenemos

$$-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}} = f_1(t_{1i}) - \frac{\partial F(t_{1i}, t_{2i})}{\partial t_{1i}}.$$

Caso 3. Si T_1 está censurado, entonces la contribución es

$$S_1(C_{1i}) = 1 - F(C_{1i}).$$

Por lo tanto, la función de verosimilitud toma la forma

$$L = \prod_{i=1}^{n} \left[\frac{\partial^2 F(t_{1i}, t_{2i})}{\partial t_{1i} \partial t_{2i}}\right]^{\delta_{i1}\delta_{2i}} \left[f_1(t_{1i}) - \frac{\partial F(t_{1i}, t_{2i})}{\partial t_{1i}}\right]^{\delta_{1i}(1-\delta_{2i})} [1 - F_1(t_{1i})]^{1-\delta_{1i}} \tag{3.1}$$

Describiremos un mecanismo probabilístico el cual permitirá que los tiempos de supervivencia T_{1i} sean infinitos, esto es útil porque nos permite introducir una función de distribución acumulada propia para describir los tiempos de vida de las personas no inmunes (susceptibles). Supongamos que cada individuo i está asociado con una variable aleatoria Bernoulli B_i, diremos que el individuo es susceptible a recurrencia si toma el valor de 1 con probabilidad p, mientras que será inmune cuando $B_i = 0$ con probabilidad $1 - p$, luego las personas susceptibles a recurrencias tienen una distribución acumulativa de supervivencia $F_0(t_1) = \Pr[T_1 \leq t_1 \mid B_i = 1]$ con $t_1 \geq 0$, que cumple con las propiedades $F_0(0) = 0$ y $F_0(\infty) = 1$, en otras palabras, es la función de distribución condicional de T_1 dado $T_1 < \infty$, por otro lado deducimos que $\Pr[T_1 \leq t_1 \mid B_i = 0] = 0$.

Por el Teorema de la probabilidad total obtenemos que la función de distribución para T_1 es

$$\begin{aligned} F_1(t_1) &= \Pr(T_1 \leq t_1) \\ &= \Pr(T_1 \leq t_1 \mid B_i = 1) \cdot \Pr(B_i = 1) + \Pr(T_1 \leq t_1 \mid B_i = 0) \cdot \Pr(B_i = 0) \\ &= p \cdot F_0(t) + 0 = p \cdot F_0(t_1) \end{aligned}$$

con $t_1 < \infty$. Como T_2 solo se puede observar cuando el primer evento es observado que

es $T_1 < \infty$, la función de distribución de T_2 es

$$F_2(t_2) = \Pr(T_2 \leq t_2 \mid T_1 < \infty).$$

Suponiendo que ambas variables T_1 y T_2 son continuas, especificamos un modelo de cópula para (T_1, T_2), dado que $T_1 < \infty$,

$$\Pr(T_1 \leq t_1, T_2 \leq t_2 \mid T_1 < \infty) = \mathcal{C}(F_0(t_1), F_2(t_2)),$$

donde $\mathcal{C}$ es una función cópula.
Utilizando este resultado y por el teorema de probabilidad total deducimos

$$\begin{aligned}
\Pr(T_1 \leq t_1, T_2 \leq t_2) &= \Pr(T_1 \leq t_1, T_2 \leq t_2 \mid B_1 = 1) \cdot \Pr(B_1 = 1) \\
&\quad + \Pr(T_1 \leq t_1, T_2 \leq t_2 \mid B_1 = 0) \cdot \Pr(B_1 = 0) \\
&= \Pr(T_1 \leq t_1, T_2 \leq t_2 \mid B_1 = 1) \cdot p + 0 \\
&= p \cdot \mathcal{C}(F_0(t_1), F_2(t_2))
\end{aligned}$$

Reescribiendo la función de verosimilitud (3.1) obtenemos

$$\begin{aligned}
L &= \prod_{i=1}^{n} \left[p \cdot \frac{\partial^2 \mathcal{C}(F_0(t_{1i}), F_2(t_{2i}))}{\partial t_{1i} \partial t_{2i}} \right]^{\delta_{1i}\delta_{2i}} \left[p \cdot f_0(t_{1i}) - p \cdot \frac{\partial \mathcal{C}(F_0(t_{1i}), F_2(t_{2i}))}{\partial t_{1i}} \right]^{\delta_{1i}(1-\delta_{2i})} \\
&\qquad \cdot [1 - pF_0(t_{1i})]^{1-\delta_{1i}} \\
&= \prod_{i=1}^{n} p^{\delta_{1i}\delta_{2i} + \delta_{1i}(1-\delta_{2i})} \left[f(t_{1i}, t_{2i}) \right]^{\delta_{1i}\delta_{2i}} \left[f_0(t_{1i}) - \frac{\partial \mathcal{C}(F_0(t_{1i}), F_2(t_{2i}))}{\partial t_{1i}} \right]^{\delta_{1i}(1-\delta_{2i})} \\
&\qquad \cdot [1 - pF_0(t_{1i})]^{1-\delta_{1i}} \\
&= \prod_{i=1}^{n} p^{\delta_{1i}} \left[f(t_{1i}, t_{2i}) \right]^{\delta_{1i}\delta_{2i}} \left[f_0(t_{1i}) - \frac{\partial \mathcal{C}(F_0(t_{1i}), F_2(t_{2i}))}{\partial t_{1i}} \right]^{\delta_{1i}(1-\delta_{2i})} [1 - pF_0(t_{1i})]^{1-\delta_{1i}}
\end{aligned} \tag{3.2}$$

Dicho anteriormente, consideramos modelos en los que las funciones de cópula $\mathcal{C}$ se especifican paramétricamente como $\mathcal{C}_\theta(u_1, u_2)$, donde el parámetro θ determina la asociación entre T_1 y T_2 a través de la τ de Kendall. Para la construcción utilizaremos como herramienta las cópulas Gumbel (2.4), Clayton (2.5) y Gaussiana (2.6), donde cada una de las distribuciones marginales de los tiempos de vida serán ajustadas por cuatro distri-

buciones: Burr (1.3), Weibull (1.5), Log-logística (1.4) y Exponencial (1.6), además cada función de riesgo será formulada utilizando el modelo de riesgos proporcionales de Cox.

La población de pacientes con cáncer se divide en una proporción p que presentará recurrencia en algún momento, este paramétro será modelado mediante el método de regresión logística

$$\log\left(\frac{p}{1-p}\right) = \gamma^T \mathbf{x}.$$

Expresión de la supervivencia.

Examinaremos la efectividad de los modelos usando los residuales de Cox-Snell, para ello calcularemos la expresión de la supervivencia de T_2 dado que T_1 ha sido observado.

$$\begin{aligned}\Pr[T_2 > t_2 | T_1 = t_1, T_1 < \infty] = S_{T_2|T_1}[t_2|T_1 = t_1] &= \frac{\Pr[T_2 > t_2, T_1 = t_1 | T_1 < \infty]}{\Pr[T_1 = t_1 | T_1 < \infty]} \\ &= \frac{f_0(t_1)\left[1 - \frac{\partial \mathcal{C}(F_0(t_1), F_2(t_2))}{\partial F_0(t_1)}\right]}{f_0(t_1)} = 1 - \frac{\partial \mathcal{C}(F_0(t_1), F_2(t_2))}{\partial F_0(t_1)}\end{aligned} \tag{3.3}$$

donde F_0 es la función de distribución condicional de T_1 dado $T_1 < \infty$ y $F_2(t_2) = \Pr(T_2 \leq t_2 | T_1 < \infty)$, así la gráfica $rc = -\log(S_{T_2|T_1})$ vs $\widehat{H}(rc)$ debe parecerse a una línea recta si el modelo ajusta adecuadamente los datos.
Calculamos este resultado para cada cópula

a) Cópula Gumbel.

$$\begin{aligned}S_{T_2|T_1}(t_2|T_1 = t_1) = 1 - \exp[-[(-\log u)^\theta + (-\log v)^\theta]^{1/\theta}] \cdot \\ \frac{(-\log u)^{\theta-1}}{u} \cdot [(-\log u)^\theta + (-\log v)^\theta]^{1/\theta - 1}.\end{aligned} \tag{3.4}$$

b) Cópula Gaussiana.

$$S_{T_2|T_1}(t_2|T_1 = t_1) = 1 - \Phi\left[\frac{\Phi^{-1}(v) - \rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}\right]. \tag{3.5}$$

c) Cópula Clayton.

$$S_{T_2|T_1}(t_2|T_1 = t_1) = 1 - (u^{-\theta} + v^{-\theta} - 1)^{-1-1/\theta} \cdot (u^{-\theta-1}). \tag{3.6}$$

3.3. Simulación

Comenzaremos enunciando algunas definiciones importantes que utilizaremos en el estudio de simulación.

Sesgo y Error cuadrático medio.

Definición. El *sesgo* de un estimador $\widehat{\theta}$ de θ es el error de estimación esperado

$$\text{sesgo}(\widehat{\theta}, \theta) = E[\text{error}(\widehat{\theta}, \theta)] = E[\widehat{\theta}] - \theta,$$

donde $\text{error}(\widehat{\theta}, \theta) = \widehat{\theta} - \theta$.

El sesgo indica cuál es la ubicación de $\widehat{\theta}$ en relación con el valor real θ.
Esperaríamos que la estimación encontrada sea la mejor suposición del valor verdadero (desconocido), para ello debe cumplir con la propiedad de precisión. Una estimación precisa es aquella en la que la variabilidad del error de estimación es pequeña, esta propiedad se captura mediante el error cuadrático medio.

Definición. *El error cuadrático medio* de un estimador $\widehat{\theta}$ de θ está dado por

$$ECM(\widehat{\theta}, \theta) = E[(\widehat{\theta} - \theta)^2] = E[\text{error}(\widehat{\theta}, \theta)^2].$$

El error cuadrático medio mide la desviación cuadrática esperada de $\widehat{\theta}$ de θ, si este valor es pequeño, entonces sabremos que $\widehat{\theta}$ estará casi siempre cerca de θ, de lo contrario es posible ver muestras para las que $\widehat{\theta}$ está demasiado alejado del verdadero valor. Desarrollando la

fórmula anterior obtenemos el siguiente resultado.

$$\begin{aligned} ECM(\widehat{\theta},\theta) &= E[(\widehat{\theta}-\theta)^2] = E[\widehat{\theta}^2 - 2\widehat{\theta}\theta + \theta^2] \\ &= E[\widehat{\theta}^2] - 2\theta E[\widehat{\theta}] + \theta^2 + E^2[\widehat{\theta}] - E^2[\widehat{\theta}] \\ &= (E[\widehat{\theta}^2] - E^2[\widehat{\theta}]) + (E^2[\widehat{\theta}] - 2\theta E[\widehat{\theta}] + \theta^2) \\ &= \mathrm{Var}[\widehat{\theta}] + (E[\widehat{\theta}] - \theta)^2 \\ &= \mathrm{Var}[\widehat{\theta}] + (\mathrm{sesgo}(\widehat{\theta},\theta))^2. \end{aligned}$$

Simulación.

Realizamos un estudio de simulación para validar el desempeño de la estimación basado en el método de máxima verosimilitud. Generamos 500 muestras aleatorias de tiempo de supervivencia bivariadas con un tamaño de 700 y 1200 observaciones del modelo de cópula Gumbel con parámetro de dependencia $\theta = 1/3,\ 2$ correspondientes respectivamente a una τ de Kendall igual a 0.25 y 0.5.
Modelamos el estatus del tiempo T_1 mediante el modelo de regresión logística en función de dos predictores: tratamiento y nodos, cada una generada a partir de la distribución binomial, para la primera variable consideramos la siguiente representación: 0 si un individuo recibe placebo y 1 si recibe tratamiento con probabilidad de éxito igual a 0.4835924, para la variable nodos consideramos 0 si un individuo tiene una cantidad de nodos menor o igual a dos y 1 si tiene una cantidad de nodos mayor a dos con probabilidad de éxito igual a 0.4887737. Los valores de β (vector de parámetros) se muestran en el siguiente componente lineal:

$$\mathbf{x}\beta = 0.1126445 - 0.8800859 \cdot \text{Tratamiento} + 1.2091202 \cdot \text{Nodos}$$

la proporción de pacientes susceptibles a recurrencias se obtiene con

$$p = \frac{\exp(\mathbf{x}\beta)}{1+\exp(\mathbf{x}\beta)},$$

deducida a partir de las covariables que posee cada individuo. Usando este valor p generamos el estatus de T_1 utilizando una distribución binomial.

Las distribuciones marginales de T_1 y T_2 se tomaron Burr con parámetros $a_1 = 4$, $b_1 = 1$ y $s_1 = 3$, y $a_2 = 5$, $b_2 = 1$ y $s_2 = 4$, respectivamente. Los tiempos de censura para T_1 se generaron a partir de una distribución uniforme sobre el intervalo $(0, 20)$, los tiempos de censura de T_2 se generaron utilizando una distribución uniforme sobre $(0, 5)$ de modo que se censuraron aproximadamente el $20\,\%$ de los tiempos de T_2. Cuando T_1 está censurada entonces T_2 también lo está, luego el estatus de T_2 para este caso es 0.

Para propósitos de análisis, nos enfocaremos en ciertas características de un estimador como el valor esperado y la varianza. En el proceso de simulación diferentes realizaciones de variables aleatorias producirán diferentes valores de probabilidades condicionales $\Pr(T_2 > t_2|T_1 < t_1)$. Intuitivamente, un buen estimador es uno que en promedio es correcto (insesgado) y no se aleja demasiado del valor verdadero (varianza pequeña).

En la siguiente tabla presentamos las medias empíricas y las desviaciones de los estimadores paramétricos de probabilidades condicionales $\Pr(T_2 > t_2|T_1 < t_1)$, la notación v.r. significa valor real. Mostramos los resultados para los cuantiles 0.8, 0.5, 0.2 de la distribución de los tiempos de supervivencia T_1 y los cuantiles 0.6, 0.2 de la distribución de T_2.

			$\tau = 0.25$		$\tau = 0.5$	
			t_2		t_2	
n	t_1		$t_{0.6}$	$t_{0.2}$	$t_{0.6}$	$t_{0.2}$
700	$t_{0.8}$	v.r.	0.417970	0.082282	0.215380	0.015278
		m. empírica	0.412034	0.081888	0.211238	0.015021
		d. estándar	0.034790	0.013622	0.024597	0.003614
	$t_{0.5}$	v.r.	0.485503	0.105774	0.366049	0.034426
		m. empírica	0.483971	0.106675	0.366091	0.035383
		d. estándar	0.025735	0.014575	0.023332	0.006778
	$t_{0.2}$	v.r.	0.548774	0.141130	0.513212	0.088286
		m. empírica	0.551948	0.144227	0.516273	0.090822
		d. estándar	0.022543	0.015739	0.022194	0.011293
1200	$t_{0.8}$	v.r.	0.417970	0.082282	0.215380	0.015278
		m. empírica	0.412972	0.081806	0.209733	0.015116
		d. estándar	0.026498	0.010326	0.019526	0.002820
	$t_{0.5}$	v.r.	0.485503	0.105774	0.366049	0.034426
		m. empírica	0.482942	0.106332	0.364081	0.035713
		d. estándar	0.019412	0.011546	0.019101	0.004790
	$t_{0.2}$	v.r.	0.548774	0.141130	0.513212	0.088286
		m. empírica	0.547921	0.143551	0.516845	0.091505
		d. estándar	0.018667	0.011882	0.016880	0.007940

Cuadro 3.1: Valores verdaderos, medias empíricas y desviaciones estándar de los estimadores paramétricos de $\Pr(T_2 > t_2 | T_1 < t_1)$.

De la tabla observamos que cuando el modelo de cópula asumido es correcto, los estimadores paramétricos tienen poco sesgo y su varianza es pequeña. Cuando hay una asociación moderada ($\tau = 0.5$) y los tiempos t_i son grandes tenemos que el estimador presenta un poco de sesgo pero mostrando una menor varianza comparada con los otros casos, en general, las simulaciones muestran que los estimadores son muy cercanos a los valores reales.

3.4. Aplicación y resultados

Nuestro interés está en estimar la asociación entre los tiempos de supervivencia secuenciales recolectados del conjunto de datos "`colon`" del software `R`, así como estimar la distribución marginal para el segundo tiempo. Dicha muestra es un estudio de 929 pacientes con cáncer de colon en etapa C de Dukes que fueron asignados al azar a tres grupos de tratamientos: Observación (*Obs*) con la cantidad de 315 individuos, también llamado grupo Placebo, Levamisol (*Lev*) con 310 individuos y Levamisol combinado con fluorouracilo (*Lev + 5-FU*) con un total de 304 individuos; el tiempo máximo de seguimiento fue aproximadamente nueve años (Lawless y Yilmaz 2011).
Los tiempos de supervivencia secuenciales se consideran de la siguiente forma: T_{1i} se define como el tiempo desde el ingreso al estudio hasta que el cáncer fue detectado (después de la aplicación de un tratamiento) para el $i-$ésimo paciente y T_{2i} el tiempo desde la presencia de recurrencia hasta la muerte, en ambos eventos tenemos presencia de censura C_{1i} y C_{2i}, respectivamente.
Las variables explicativas que utilizaremos en el modelo logístico y de Cox son: tipo de tratamiento (*rx*), género (*sex*), edad (*age*) y nodos (*nodes*), este último dato lo dividimos en dos subgrupos: el primer grupo consta de individuos que tienen una cantidad de nodos mayor a dos y el segundo grupo estará conformada por el complemento. En la siguiente tabla se presenta la clasificación de algunas variables respecto al tipo de tratamiento.

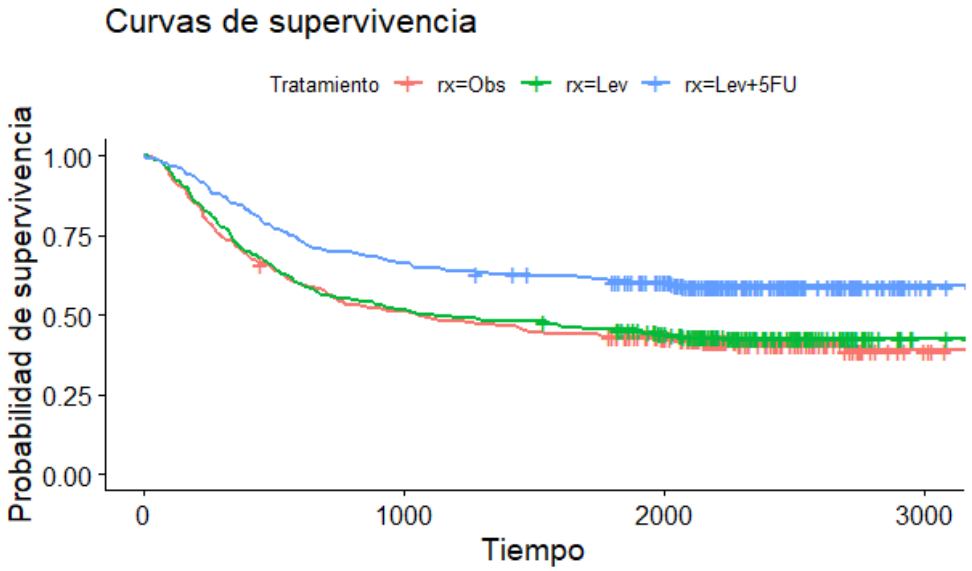
Tipo de	Sexo		Total		Nodos	
Tratamiento	Hombre	Mujer	Recurrencia	Muerte	Grupo 1	Grupo 2
Observación	166	149	177	168	218	94
Levamisol	177	133	172	161	218	86
Levamisol+5FU	141	163	119	123	199	96

Para estimar los parámetros utilizamos el método de máxima verosimilidad aplicado a la ecuación (3.2), este fue implementado en el software `R` donde hicimos un cambio de coordenadas parametrales, pues suele ser difícil modelar una variable que tiene rango restringido, por ejemplo, el paramétro de dependencia de la cópula gaussiana está definida en el intervalo $(-1, 1)$, así utilizar la función tangente hipérbolica es buena opción, ya que su rango es dicho intervalo. La propiedad de invarianza del estimador de máxima verosimilitud (EMV) respalda este cambio parametral siempre que la función sea derivable distinta

de cero, así calcular el EMV en las coordenadas originales es equivalente a encontrar el EMV en las coordenadas restringidas y transformadas.

Resultados.

Graficamos las curvas de supervivencia de los pacientes clasificados por el tipo de tratamiento a partir de los estimadores de Kaplan-Meier. De acuerdo a los resultados decidimos centrar el estudio en los grupos de observación (placebo) y levamisol más $5-FU$, pues la curva de supervivencia del tercer grupo (Levamisol) es similar a la del grupo Placebo como puede observar a continuación.



El primer análisis fue aplicar el modelo de Cox a ambas funciones marginales, como resultado obtuvimos que p era aproximadamente uno, en otras palabras, toda la población es susceptible a recurrencias, lo cual dado nuestra base de datos eso no es posible de suceder. Intuitivamente al modelo le resulta difícil distinguir entre los efectos de aumentar p y de aumentar los valores de los parámetros (aquellos que presentarán recurrencia en algún momento en el futuro lo harán más rápidamente). Intentamos resolver este problema reestringiendo los componentes que no son de intersección, pero obtuvimos el mismo resultado, esto nos condujo considerar solo algunas variables explicativas, sin embargo el modelo de riesgos proporcionales no tuvo efectos significativos en los resultados, por tal motivo decidimos no aplicarlo.

Se calcularon las estimaciones de parámetros mediante el método de máxima verosimilitud, utilizando la prueba de *Wald* mostraremos cuáles fueron las variables significativas en el modelo logístico, algunos casos son los siguientes:

a) T_1, $T_2 \sim$ *Burr y la función de densidad conjunta es modelada por la cópula Gaussiana.* Los resultados de la prueba *Wald* con un nivel de significancia del 5 % se resumen a continuación.

Covariables	Estimadores (γ)	Estadístico de Wald	p-valor
Tratamiento	-0.702068755	3.72625303	$9.717366e-05$
Edad	-0.001900498	0.24155409	0.4045628
Género	-0.130165728	0.69539282	0.2434046
Nodos	1.064549898	5.58056514	$1.198692e-08$

Observemos que las únicas variables individualmente significativas fueron Tratamiento y Nodos.
Los *odds* de éxito se definen como la proporción de la probabilidad de éxito sobre la probabilidad de fracaso, para la variable tratamiento, la razón de los odds para pacientes en el grupo placebo a pacientes en el grupo Lev+5FU es

$$\frac{\exp(0.116228925)}{\exp(0.116228925 - 0.702068755)} = 2.017923,$$

así las probabilidades de que un individuo sea susceptible dado a que pertenece al grupo placebo son 2.017923 veces superiores a las de un paciente en el grupo Lev+5FU, por otro lado, para la variable nodos la razón de los odds es:

$$\frac{\exp(0.116228925 + 1.064549898)}{\exp(0.116228925)} = 2.899534,$$

las probabilidades de que un paciente sea susceptible dado que tiene más de dos nodos son más altas que la de un paciente con una cantidad de nodos menor o igual a 2.

b) T_1, $T_2 \sim$ *Weibull y la función de densidad conjunta es modelada por la cópula Gumbel.* Los resultados de la prueba *Wald* con un nivel de significancia del 5 % se muestran en la siguiente tabla.

Covariables	Estimadores (γ)	Estadístico de Wald	p-valor
Tratamiento	-0.846481215	4.6884188	$1.376621e-06$
Edad	-0.002997042	0.3985972	0.345095
Género	-0.203851829	1.1365122	0.1278711
Nodos	1.086932152	6.0266093	$8.371755e-10$

c) T_1, $T_2 \sim$ *Burr y la función de densidad conjunta es modelada por la cópula Gumbel.* Los resultados de la prueba *Wald* con un nivel de significancia del 5 % se presentan en la siguiente tabla.

Covariables	Estimadores (γ)	Estadístico de Wald	p-valor
Tratamiento	-0.802467021	4.15539259	$1.623647e-05$
Edad	-0.002380948	0.29719019	0.3831607
Género	-0.149423721	0.78597447	0.2159412
Nodos	1.215691422	6.14827561	$3.916492e-10$

Nuevamente las únicas covariables con efectos significativas sobre p son tratamiento y nodos, en general este resultado fue el mismo para cada estudio.

Utilizamos el método de eliminación hacia atrás destinado a encontrar el mejor subconjunto de variables predictoras, se aplicó al modelo completo (variables explicativas: tratamiento, nodos, género y edad) de cada análisis y gradualmente eliminamos las variables edad y género en el modelo de regresión utilizando el criterio de Wald.

Para determinar cuál de estos modelos ofrece un mejor ajuste para los datos utilizamos la prueba de razón de verosimilitud que compara la bondad de ajuste de dos modelos anidados, donde los modelos no restringidos serán la distribución Burr y Weibull frente a los modelos restringidos Log-logística y Exponencial, respectivamente. A continuación se muestran algunos de los resultados.

(a) $H_0 : a = 1$ (Marginales Log-logística) vs $H_A : a \neq 1$ (Marginales Burr), con función de densidad conjunta modelada por la cópula Gaussiana.
El estadístico

$$r^*(x) = -2\log[r(x)] = 8.939103 \overset{\text{a}}{\sim} \chi^2_2$$

el p-valor es igual $\Pr(r^*(x) \geq 8.939103) \approx 0.01145245$, por lo que rechazamos la hipótesis nula H_0 con un nivel de significancia de 0.05.

(b) $H_0 : \alpha = 1$ (Marginales Exponencial) vs $H_A : \alpha \neq 1$ (Marginales Weibull), con función de densidad conjunta modelada por la cópula Gumbel.

$$r^*(x) = -2\log[r(x)] = 9.831332 \overset{\text{a}}{\sim} \chi^2_2$$

el p-valor es igual $\Pr(r^*(x) \geq 9.831332) \approx 0.007330834$, luego rechazamos la hipótesis nula H_0 con un nivel de significancia de 0.05.

(c) $H_0 : \alpha = 1$ (Marginales Exponencial) vs $H_A : \alpha \neq 1$ (Marginales Weibull), con función de densidad conjunta modelada por la cópula Clayton.

$$r^*(x) = -2\log[r(x)] = 10.80312 \overset{\text{a}}{\sim} \chi^2_2$$

el p-valor es igual $\Pr(r^*(x) \geq 10.80312) \approx 0.004509536$, así rechazamos la hipótesis nula H_0 con un nivel de significancia de 0.05.

El ánalisis de la prueba de razón de verosimilitud aplicado a cada uno de los casos induce a rechazar la hipótesis nula, en consecuencia nuestro estudio se reduce a trabajar con los modelos alternativos, las marginales se distribuyen Burr y Weibull.
Seguidamente presentamos en las siguientes tablas los resultados que obtuvimos cuando suponemos que las distribuciones marginales se distribuyen Burr/Weibull donde la función de densidad conjunta es modelada por las cópulas Gaussiana, Gumbel y Clayton.

a) T_1, $T_2 \sim$ Burr/Weibull, donde la función de densidad conjunta es modelada por la cópula Gaussiana.

Cópula Gaussiana						
	Burr			*Weibull*		
	Estimadores	Límite inferior	Límite superior	Estimadores	Límite inferior	Límite superior
τ	0.2026951	0.1213331	0.2807725	0.1957946	0.1139423	0.2744394
	Forma 1 (a)	Forma 2 (b)	Escala (s)	Forma (α)	Escala (λ)	
Marginal T_1	0.7952676	1.7205752	369.8061991	1.15777	622.7063	
Marginal T_2	2.38192	1.16445	1018.40841	0.9478959	595.1806	
p	Intercepción	Tratamiento	Nodos	Intercepción	Tratamiento	Nodos
	0.03583432	−0.84642292	1.15863321	−0.11460976	−0.78260672	1.06682202

Cuadro 3.2: Estimadores del modelo de dependencia y de los parámetros de F_0 y F_2.

b) $T_1, T_2 \sim$ Burr/Weibull y función de densidad conjunta modelada por la cópula Gumbel.

Cópula Gumbel						
	Burr			*Weibull*		
	Estimadores	Límite inferior	Límite superior	Estimadores	Límite inferior	Límite superior
τ	0.2953097	0.2038838	0.4067849	0.2140832	0.1357904	0.3207625
	a	b	s	α	λ	
Marginal T_1	0.6085487	1.7926178	306.2511700	1.144735	634.456411	
Marginal T_2	1.765056	1.163831	796.715484	0.928161	613.823394	
p	Intercepción	Tratamiento	Nodos	Intercepción	Tratamiento	Nodos
	0.1049103	−0.8735638	1.2105682	−0.10954324	−0.78532998	1.06894493

Cuadro 3.3: Estimadores del modelo de dependencia y de los parámetros de F_0 y F_2.

Cuando las marginales se distribuyen Burr y la distribución conjunta es modelada por la cópula Gumbel obtenemos que los odds de éxito para la variable tratamiento son

$$\frac{\exp(0.1049103)}{\exp(0.1049103 - 0.8735638)} = 2.395433,$$

las probabilidades de que un individuo sea susceptible dado que recibe Placebo son 2.395433 veces superiores a las de un paciente con tratamiento Lev+5FU. Para la

variable nodos la razón de los odds son

$$\frac{\exp(0.1049103 + 1.2105682)}{\exp(0.1049103)} = 3.355391$$

así las probabilidades de que un individuo presente recurrencia dado a que tiene una cantidad de nodos mayor de dos son 3.355391 veces más altas que la de un paciente con una cantidad de nodos menor o igual a 2.

A continuación proporcionamos la tasa de curación de los individuos pertenecientes a cada subgrupo, donde Nodos0 representa el grupo con una cantidad de nodos menor o igual a 2 y Nodos1 el complemento.

Covariables	Tasa de curación $1 - \widehat{p}$
Placebo y Nodos0	0.4737965
Placebo y Nodos1	0.2115715
Tratamiento y Nodos0	0.6832295
Tratamiento y Nodos1	0.3912848

Cuadro 3.4: Tasa de curación $1 - \widehat{p}$.

c) T_1, $T_2 \sim$ Burr/Weibull, donde la función de densidad conjunta es modelada por la cópula Clayton.

Cópula Clayton						
	Burr			*Weibull*		
	Estimadores	Límite inferior	Límite superior	Estimadores	Límite inferior	Límite superior
τ	0.1262324	0.07159372	0.21300308	0.1889509	0.1233552	0.2783524
	a	b	s	α	λ	
Marginal T_1	0.8150686	1.7059650	380.1780901	1.1564939	622.1000425	
Marginal T_2	2.904167	1.147443	1226.467600	0.9429919	585.5817340	
p	Intercepción	Tratamiento	Nodos	Intercepción	Tratamiento	Nodos
	0.04021045	−0.84642622	1.14922311	−0.11901183	−0.77857071	1.07129696

Cuadro 3.5: Estimadores del modelo de dependencia y de los parámetros de F_0 y F_2.

De acuerdo a los resultados, notamos que hay diferencias significativas entre las τ de Kendall, obtener una τ como la del modelo Clayton nos proporciona evidencia estadística de falta de correlación entre las variables T_1 y T_2, esto nos conduciría a una posibilidad de independencia entre las variables y limitar el uso de un modelo de cópula.
Para examinar la efectividad de los modelos consideramos los residuales de Cox-Snell, para ello utilizamos la expresión de supervivencia dadas en las ecuaciones (3.4), (3.5) y (3.6), así la gráfica $rc = -\log(S_{T_2|T_1})$ vs $\widehat{H}(rc)$ debe parecerse a una línea recta si el modelo ajusta adecuadamente los datos.
En las siguientes figuras mostramos los residuales para cada modelo cópula.

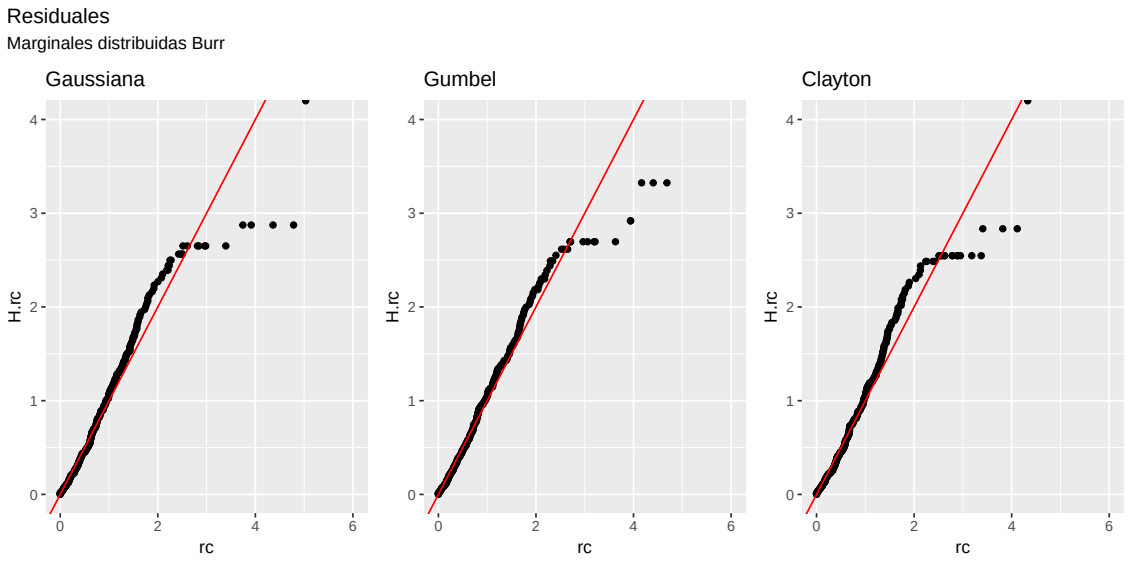


Figura 3.4: Gráfica de diagnóstico (marginales distribuidas Burr).

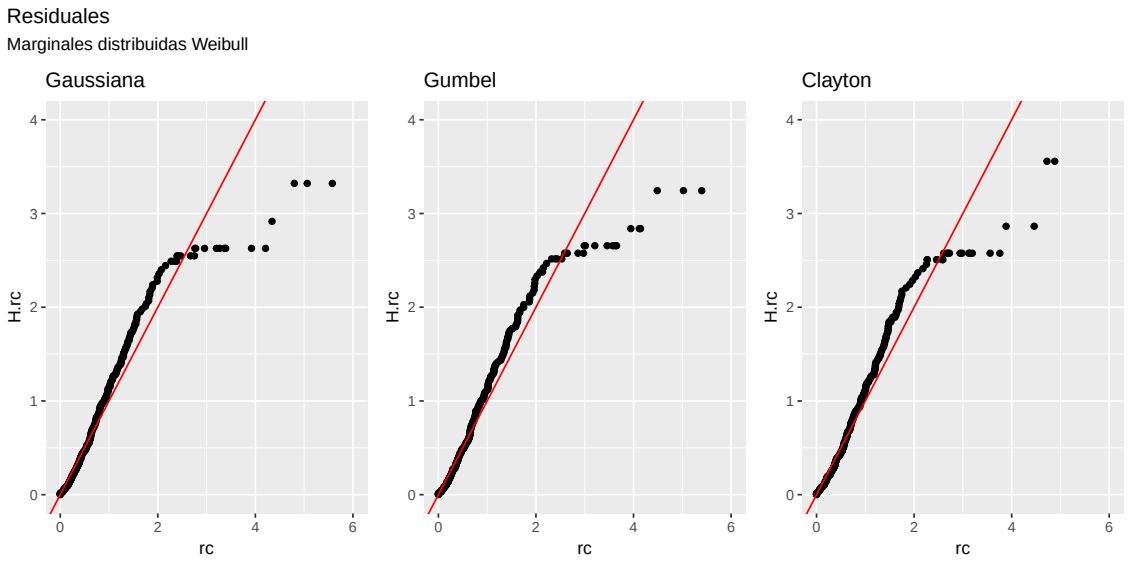


Figura 3.5: Gráfica de diagnóstico (marginales distribuidas Weibull).

Observamos que la gráfica que mejor ajusta a una línea recta es cuando las marginales se distribuyen Burr y la función de distribución conjunta es modelada por la cópula Gumbel, en base a esto decidimos optar este modelo y así construir una banda de confianza para las funciones de riesgo.

Una vez que hemos calculado el estimador de máxima verosimilitud y el error estándar, una banda de confianza puede ser construida de la siguiente manera: generamos $10,000$ observaciones de una distribución Normal con media el estimador de máxima verosimilitud $\widehat{\theta}$ y varianza $I(\widehat{\theta})^{-1}$ (inversa de la información de Fisher). Para cada t_i en el dominio evaluamos la función de riesgo utilizando los respectivos parámetros generados y calculamos los percentiles 25 y 97.5, la curva que une estos puntos será la banda de confianza al $95\,\%$ de la función de riesgo. A continuación se muestran las funciones de riesgo y sus correspondientes bandas de confianza.

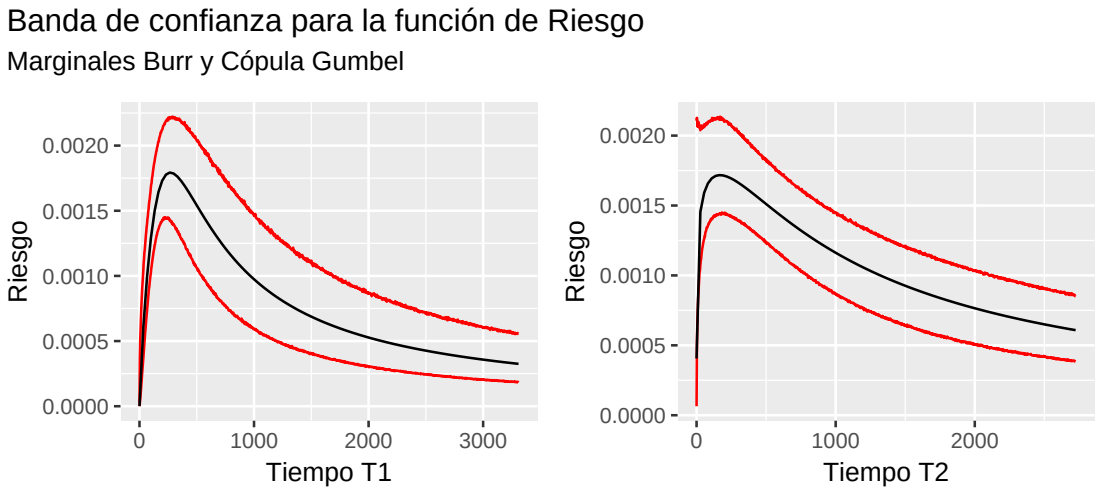


Figura 3.6: Bandas de confianza, donde T_1, $T_2 \sim$ Burr y la cópula utilizada es la Gumbel.

Observe que ambas funciones de riesgo no son monótonas, la interpretación de este resultado se debe a que las tasas de riesgo aumentan drásticamente durante un período de tiempo debido a la aplicación de un tratamiento/recurrencia y luego disminuye de acuerdo a la efectividad del tratamiento o por la cantidad de nodos que tiene cada paciente.

Conclusión

De acuerdo a las gráficas de diagnóstico 3.4 y 3.5, los resultados sugieren que el modelo de cópula Gumbel se prefiere al resto de las cópulas cuando las marginales se distribuyen Burr, pues es la que mejor se aproxima a una línea recta, por lo cual decidimos considerar este modelo para el ajuste de nuestros datos. La asociación entre las variables T_1 y T_2 de acuerdo a este modelo es $\tau = 0.2953097$, dicha medida de asociación es pequeña pero nos muestra evidencia estadística de dependencia entre las variables de supervivencia en comparación con la cópula Clayton, $\tau = 0.1262324$.

Las covariables que tuvieron efectos significativos sobre la proporción de pacientes susceptibles a recurrencia fueron tratamiento y cantidad de nodos, los resultados sugieren que la sustancia Lev + $5FU$ es un tratamiento eficaz para reducir la probabilidad de recurrencia del cáncer mientras que tener una cantidad de nodos mayor a dos aumenta la probabilidad de presentarla. Por la tabla 3.4 observamos que la tasa de curación $1 - \widehat{p} = 1 - 0.3167705 = 0.6832295$ es mayor que las probabilidades obtenidas para el resto de los grupos: pacientes que ingieren placebo y tienen una cantidad de nodos menor o igual a 2, pacientes que ingieren placebo y tienen una cantidad de nodos mayor que 2, y pacientes que reciben tratamiento con una cantidad de nodos mayor que 2. Esto significa que aquellos pacientes que se les administra Lev + $5FU$ y tienen una cantidad de nodos menor o igual que dos, tienen mayor probabilidad de convertirse en personas inmunes.

Por los resultados, observamos que la metodología elegida dificultó el uso de modelos de riesgos proporcionales de Cox, pues las estimaciones de máxima verosimilitud de los componentes del modelo logístico y de Cox estaban correlacionadas, en consecuencia no obtuvimos algunos resultados deseados. Los factores que pudieron haber provocado este problema fue que la cantidad de parámetros a estimar eran demasiadas, además las variables explicativas que sugieren mayor probabilidad de recurrencia serán también las

que sugieren mayor tasa de recurrencia. Como alternativa, para obtener un modelo más óptimo, sugerimos modelar la función de riesgo utilizando el modelo de tiempo de falla acelerado, o modificar el método paramétrico por uno semi-paramétrico o no paramétrico.

Observamos también que la función de distribución Burr obtuvo un mejor comportamiento que las otras funciones, por este motivo dicha distribución podría ser útil en la metodología paramétrica para el análisis de tiempos de supervivencia.

Bibliografía

[1] American Cancer Society (2020); *Colorectal Cancer Facts & Figures* 2020 − 2022. Atlanta.

[2] American Cancer Society (2020); *What is Cancer?*, publicado electrónicamente en `https://www.cancer.org/content/dam/CRC/PDF/Public/6041.00.pdf`.

[3] Burney S. M. y Bin Ajaz Osama (2020); *Copulas: A Historical Literature Review and Major developments*.

[4] *Cancer research UK* (2018). Dukes' staging system, publicado electrónicamente en `https://www.cancerresearchuk.org/`.

[5] Collet David (2003); *Modelling Survival Data in Medical Research*. Chapman & Hall. Second Edition.

[6] *Colon Cancer Treatment* (2020). National Cancer Institute, publicado electrónicamente en `https://www.cancer.gov/types/colorectal/patient/colon-treatment-pdq`.

[7] *Colorectal Cancer* (2021). American Cancer Society, publicado electrónicamente en `https://www.cancer.org/cancer/colon-rectal-cancer.html`.

[8] D. Y. Lin (1994); *Cox regression analysis of multivariate failure time data: the marginal approach*. Statistics in Medicine.

[9] Dobson Annette J. y Barnett Adrian G. (2008); *An Introduction to Generalized Linear Models*. Chapman & Hall. Third Edition.

[10] Escarela Gabriel y Hernández Angélica (2009); *Modelado de parejas aleatorias usando cópulas*. Revista Colombiana de Estadística.

[11] Genest Christian y MacKay R. J (1986); *Copules archimédiennes et families de lois bidimensionnelles dont les marges sont données*. The Canadian Journal of Statistics.

[12] Harrell Frank E., Jr (2016); *Regression Modeling Strategies. With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Springer. Second Edition.

[13] Hogg Robert V.; McKean Joseph W.; Craig Allen T. (2005); *Introduction to Mathematical Statistics*. Pearson. Seventh Edition.

[14] Klein John P.; Moeschberger Melvin L. (2003); *Survival Analysis. Techniques for Censored and Truncated and Truncated Data*. New York, Springer.

[15] Lawless Jerald F. (2003); *Statistical Models and Methods for Lifetime Data*. Wiley, Hoboken, New Jersey, Second Edition.

[16] Lawless Jerald F. y Yilmaz Yildiz E. (2011); *Semiparametric estimation in copula models for bivariate sequential survival times*. Biometrical Journal 53.

[17] Nelsen Roger B. (2006); *An Introduction to Copulas*. Springer Series in Statistics. Second Edition.

[18] Tableman Mara y Sung Kim Jong (2012); *Survival Analysis Using S/R^**. Fariborz Maseeh Department of Mathematics & Statistics. Portland State University.

[19] Weaam Alhadlaq y Abdulhamid Alzaid (2020); *Distribution Function, Probability Generating Function and Archimedean Generator*. Symmetry.

[20] *What is Cancer?* (2021). National Cancer Institute, publicado electrónicamente en `https://www.cancer.gov/about-cancer/understanding/what-is-cancer`.

[21] WHO (1987); Special Programme of Research, Development and Research Training in Human Production. *Vaginal bleeding patterns — the problem and an example data set*. Applied Stochastic Models and Data Analysis.

[22] Wienke Andreas (2011); *Frailty Models in Survival Analysis*. Chapman & Hall/CRC Biostatistics Series.

Apéndice A

Modelos de tiempos de supervivencia bivariados.

En este apartado proporcionamos los resultados utilizados en la construcción de la función de verosimilitud en la sección 3.2.
Estamos interesados en obtener la contribución a la función de verosimilitud cuando tenemos presencia de datos censurados a la derecha, para este propósito definimos lo siguiente: sean $(X_1, X_2, \delta_1, \delta_2)$ las observaciones censuradas bivariadas con $X_i = \text{mín}(T_i, C_i)$ y $\delta_i = I(T_i \leq C_i)$ para $i = 1, 2$. Denotamos la densidad de (T_1, T_2) y (C_1, C_2), respectivamente, por $f(t_1, t_2)$ y $g(c_1, c_2)$, y suponemos que los tiempos de supervivencia (T_1, T_2) son independientes a los tiempos censurados (C_1, C_2).
A continuación calculamos la distribución acumulada para cada caso.

1. Ambos tiempos de supervivencia son observados.

$$\begin{aligned} H(x_1, x_2, 1, 1) &= \Pr(X_1 > x_1, X_2 > x_2, \delta = 1, \delta = 1) \\ &= \Pr(T_1 > x_1, T_2 > x_2, T_1 \leq C_1, T_2 \leq C_2) \\ &= \int_{t_1 > x_1} \int_{t_2 > x_2} \int_{t_1 \leq c_1} \int_{t_2 \leq c_2} f(t_1, t_2) \cdot g(c_1, c_2)\ dt_1\ dt_2\ dc_1\ dc_2 \\ &= \int_{x_2}^{\infty} \int_{x_1}^{\infty} \int_{t_1}^{\infty} \int_{t_2}^{\infty} f(t_1, t_2) \cdot g(c_1, c_2)\ dc_2\ dc_1\ dt_1\ dt_2 \\ &= \int_{x_2}^{\infty} \int_{x_1}^{\infty} \left(f(t_1, t_2) \int_{t_1}^{\infty} \int_{t_2}^{\infty} g(c_1, c_2)\ dc_2\ dc_1 \right)\ dt_1\ dt_2 \end{aligned}$$

En consecuencia, la función de densidad para pares de tiempos de supervivencia no

censurados es

$$f(x_1, x_2) \int_{t_1}^{\infty} \int_{t_2}^{\infty} g(c_1, c_2)\ dc_2\ dc_1,$$

puesto que la censura es no informativa se reduce a la siguiente expresión $f(x_1, x_2)$.

2. Para el segundo caso cuando T_1 es observada y T_2 censurada obtenemos

$$\begin{aligned} H(x_1, x_2, 1, 0) &= \Pr(X_1 > x_1, X_2 > x_2, \delta = 1, \delta = 0) \\ &= \Pr(T_1 > x_1, C_2 > x_2, T_1 \leq C_1, T_2 > C_2) \\ &= \int_{t_1 > x_1} \int_{c_2 > x_2} \int_{t_1 \leq c_1} \int_{t_2 > c_2} f(t_1, t_2) \cdot g(c_1, c_2)\ dt_1\ dt_2\ dc_1\ dc_2 \\ &= \int_{x_1}^{\infty} \int_{x_2}^{\infty} \int_{t_1}^{\infty} \int_{c_2}^{\infty} f(t_1, t_2) \cdot g(c_1, c_2)\ dt_1\ dt_2\ dc_1\ dc_2 \\ &= \int_{x_2}^{\infty} \int_{x_1}^{\infty} \left(\int_{c_2}^{\infty} f(t_1, t_2)\ dt_2 \int_{t_1}^{\infty} g(c_1, c_2)\ dc_1 \right)\ dt_1\ dc_2 \end{aligned}$$

Así la función de densidad es

$$\int_{c_2}^{\infty} f(t_1, t_2)\ dt_2 \int_{t_1}^{\infty} g(c_1, c_2)\ dc_1,$$

puesto que la censura es no informativa se reduce a $\int_{c_2}^{\infty} f(x_1, t_2)\ dt_2$, el cual es equivalente al negativo de la derivada parcial de la función de supervivencia bivariada, $-\frac{\partial S(x_1, x_2)}{dx_1}$.

3. Cuando T_1 está censurada este cálculo es más sencillo pues no consideramos la segunda variable T_2, este resultado es análogo al construido en la ecuación (2.1).

Apéndice B

Código

```
F.BurrGumbell<- function(par, t.rec, t.mue, cens.rec, cens.mue, matriz.P){
  #num. de variables en el vector gama
  n<- dim(matriz.P)[2]
  #coeficientes del vector gama
  gama<- par[1:n]

  theta<- exp(par[n+1])+1
  forma1.1<- exp(par[n+2])
  forma1.2<- exp(par[n+3])
  scala1<- exp(par[n+4])
  forma2.1<- exp(par[n+5])
  forma2.2<- exp(par[n+6])
  scala2<- exp(par[n+7])

  #Matriz diseño para el modelo logistico
  x.matP<- matriz.P
  x.gama <- c(x.matP %*% gama)
  p<- exp(x.gama)/(1+exp(x.gama))

  #funciones de densidad Burr
  f.T1<- dburr(t.rec, shape1 = forma1.1, shape2 = forma1.2, scale = scala1)
  f.T2<- dburr(t.mue, shape1 = forma2.1, shape2 = forma2.2, scale = scala2)
```

```
  #funciones de distribucion T1 y T2
  F.T1<- pburr(t.rec, shape1 = forma1.1, shape2 = forma1.2, scale = scala1)
  F.T2<- pburr(t.mue, shape1 = forma2.1, shape2 = forma2.2, scale = scala2)

  #funcion de densidad conjunta de las variables T1 y T2
  #f.T1T2=c*f.T1*f.T2
  gumbel.cop<- gumbelCopula(theta)
  f.T1T2<- f.T1*f.T2*dCopula(cbind(F.T1, F.T2), gumbel.cop)

  log.Like<- (sum((cens.rec*log(p))+cens.rec*cens.mue*log(f.T1T2)))
    +sum(cens.rec*(1-cens.mue)*log(f.T1*(1-(exp(-(((-log(F.T1))^(theta))
    +((-log(F.T2))^(theta)))^(1/theta))*(((-log(F.T1))^(theta-1))/(F.T1))
    *((((-log(F.T1))^(theta))+((-log(F.T2))^(theta)))^(-1+(1/theta)))))))
    +sum((1-cens.rec)*(log(1-p*F.T1)))
  #Minimizar
  -1*log.Like
}
```

Casa abierta al tiempo
UNIVERSIDAD AUTÓNOMA METROPOLITANA

ACTA DE EXAMEN DE GRADO

No. 00223
Matrícula: 2192802610

Modelado de regresión basado en cópula para eventos secuenciales de supervivencia.

Con base en la Legislación de la Universidad Autónoma Metropolitana, en la Ciudad de México se presentaron a las 14:00 horas del día 9 del mes de diciembre del año 2021 POR VÍA REMOTA ELECTRÓNICA, los suscritos miembros del jurado designado por la Comisión del Posgrado:

DR. ALBERTO CASTILLO MORALES
DRA. LIZBETH NARANJO ALBARRAN
DR. GABRIEL NUÑEZ ANTONIO

Bajo la Presidencia del primero y con carácter de Secretario el último, se reunieron para proceder al Examen de Grado cuya denominación aparece al margen, para la obtención del grado de:

MAESTRA EN CIENCIAS (MATEMÁTICAS APLICADAS E INDUSTRIALES)

DE: ITZEL MOCTEZUMA BARONA

y de acuerdo con el artículo 78 fracción III del Reglamento de Estudios Superiores de la Universidad Autónoma Metropolitana, los miembros del jurado resolvieron:

A P R O B A R

Acto continuo, el presidente del jurado comunicó a la interesada el resultado de la evaluación y, en caso aprobatorio, le fue tomada la protesta.

UNIVERSIDAD AUTONOMA METROPOLITANA
Casa abierta al tiempo
DIRECCION DE SISTEMAS ESCOLARES

ITZEL MOCTEZUMA BARONA
ALUMNA

REVISÓ

MTRA. ROSALIA SERRANO DE LA PAZ
DIRECTORA DE SISTEMAS ESCOLARES

DIRECTOR DE LA DIVISIÓN DE CBI

DR. JESUS ALBERTO OCHOA TAPIA

PRESIDENTE

DR. ALBERTO CASTILLO MORALES

VOCAL

DRA. LIZBETH NARANJO ALBARRAN

SECRETARIO

DR. GABRIEL NUÑEZ ANTONIO

El presente documento cuenta con la firma –autógrafa, escaneada o digital, según corresponda- del funcionario universitario competente, que certifica que las firmas que aparecen en esta acta – Temporal, digital o dictamen- son auténticas y las mismas que usan los c.c. profesores mencionados en ella