



Pontificia Universidad
JAVERIANA
Bogotá

MAE STRÍA EN 
EPIDEMIOLOGÍA
CLÍNICA

BIOESTADÍSTICA AVANZADA

MÓDULO III

Semana 12

Modelo de Riesgos Proporcionales de Cox

Material de contenido y aplicación

Martín Alonso Rondón Sepúlveda

Introducción

En muchos estudios médicos, epidemiológicos e incluso en otras áreas que no necesariamente pertenecen al campo de la salud, el interés principal consiste en analizar el tiempo hasta que ocurre un evento de interés (por ejemplo: muerte, recaída de una enfermedad, falla de un dispositivo, aparición de una complicación).

Como se revisó previamente, un análisis inicial puede realizarse mediante técnicas de análisis de supervivencia utilizando métodos descriptivos, como la estimación de la función de supervivencia mediante el método de Kaplan–Meier o la comparación entre grupos mediante la prueba de Log-rank. Sin embargo, estos métodos presentan limitaciones importantes, ya que generalmente permiten evaluar el efecto de una variable explicativa a la vez y no permiten ajustar simultáneamente por múltiples covariables. En consecuencia, su capacidad para controlar posibles factores de confusión o evaluar interacciones entre variables es limitada.

Para superar estas limitaciones, David Cox desarrolló en 1972 el modelo de riesgos proporcionales, conocido como modelo de Cox. A diferencia de la regresión logística convencional, que se limita a evaluar la presencia o ausencia de un evento, el modelo de Cox incorpora explícitamente la dimensión temporal. Esto permite no solo estimar la posibilidad de que ocurra el evento, sino también la velocidad o tasa a la cual este ocurre a través del tiempo de seguimiento. Además, el modelo permite evaluar simultáneamente el efecto de múltiples covariables y manejar adecuadamente datos censurados, una característica frecuente en estudios de supervivencia.

El modelo de Cox se clasifica como un modelo semiparamétrico, ya que permite estimar el efecto de las covariables sobre el riesgo instantáneo (hazard) sin asumir una forma específica para la distribución del tiempo de supervivencia. Esta característica le otorga gran flexibilidad y lo convierte en una herramienta ampliamente utilizada en epidemiología y en investigación clínica, especialmente cuando el seguimiento de los individuos es incompleto. De esta manera, el modelo aprovecha la información disponible de cada participante, incluso de aquellos que no han experimentado el evento al final del estudio.

Para comprender adecuadamente este modelo es necesario familiarizarse con algunos conceptos fundamentales del análisis de supervivencia, en particular la función de riesgo instantáneo (hazard), la función de riesgo acumulado (cumulative hazard) y su relación con la función de supervivencia. Estos conceptos constituyen la base teórica para la formulación del modelo de Cox y para la correcta interpretación de sus resultados.

Funciones de riesgo

Función de riesgo instantáneo (hazard o peligro)

Se define como la tasa instantánea de ocurrencia del evento en un tiempo específico, dado que el individuo ha sobrevivido hasta ese momento. En otras palabras, describe la intensidad con la que el evento ocurre en un instante determinado entre los individuos que permanecen en riesgo. Está dada por:

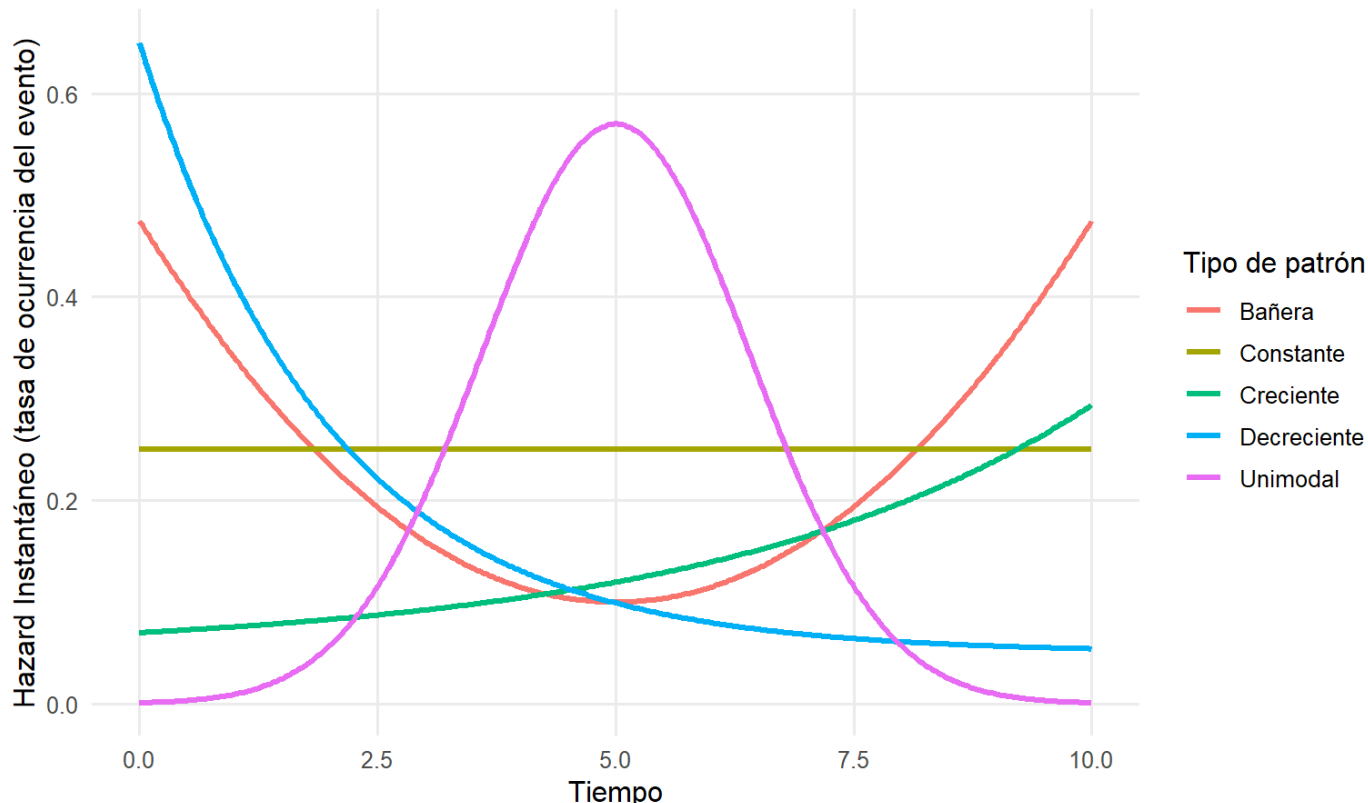
$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}$$

donde:

- T es la variable aleatoria que representa el tiempo hasta la ocurrencia del evento
- t constituye un tiempo específico dentro del periodo de seguimiento
- Δt es un intervalo de tiempo muy pequeño que tiende a cero.
- $P(t \leq T < t + \Delta t \mid T \geq t)$ es la probabilidad de que el evento ocurra en el intervalo $[t, t + \Delta t)$ dado que el individuo ha sobrevivido hasta el tiempo t
- $h(t)$ es la función de riesgo instantáneo en el tiempo t

El gráfico 1, muestra ejemplos de posibles escenarios o patrones que puede adoptar la función de hazard a lo largo del tiempo, lo que refleja distintas dinámicas temporales del riesgo del evento entre los individuos que permanecen en riesgo. La flexibilidad del modelo de Cox permite adaptarse a estas diversas formas de hazard sin necesidad de especificar una distribución paramétrica para el tiempo de supervivencia, lo que es una de sus principales fortalezas.

Gráfico 1. Ejemplo de patrones frecuentes de la función de hazard



Hazard creciente: La tasa instantánea de ocurrencia del evento aumenta a medida que transcurre el tiempo entre los individuos que permanecen en riesgo. Esto indica que el evento ocurre con mayor frecuencia en etapas tardías del seguimiento. Ejemplos típicos de este patrón se observa en enfermedades degenerativas como la Enfermedad de Alzheimer o la Enfermedad de Parkinson, donde el deterioro es progresivo y se incrementa progresivamente el riesgo de presentar el evento.

Hazard decreciente: La tasa instantánea de ocurrencia del evento es mayor al inicio del seguimiento y disminuye con el tiempo entre los individuos que continúan en riesgo. Esto ocurre cuando existe un periodo inicial de mayor vulnerabilidad. Un ejemplo común es el periodo posterior a una cirugía mayor o tras el inicio de un tratamiento intensivo, donde el riesgo inicial de complicaciones es alto y luego disminuye conforme el paciente se estabiliza.

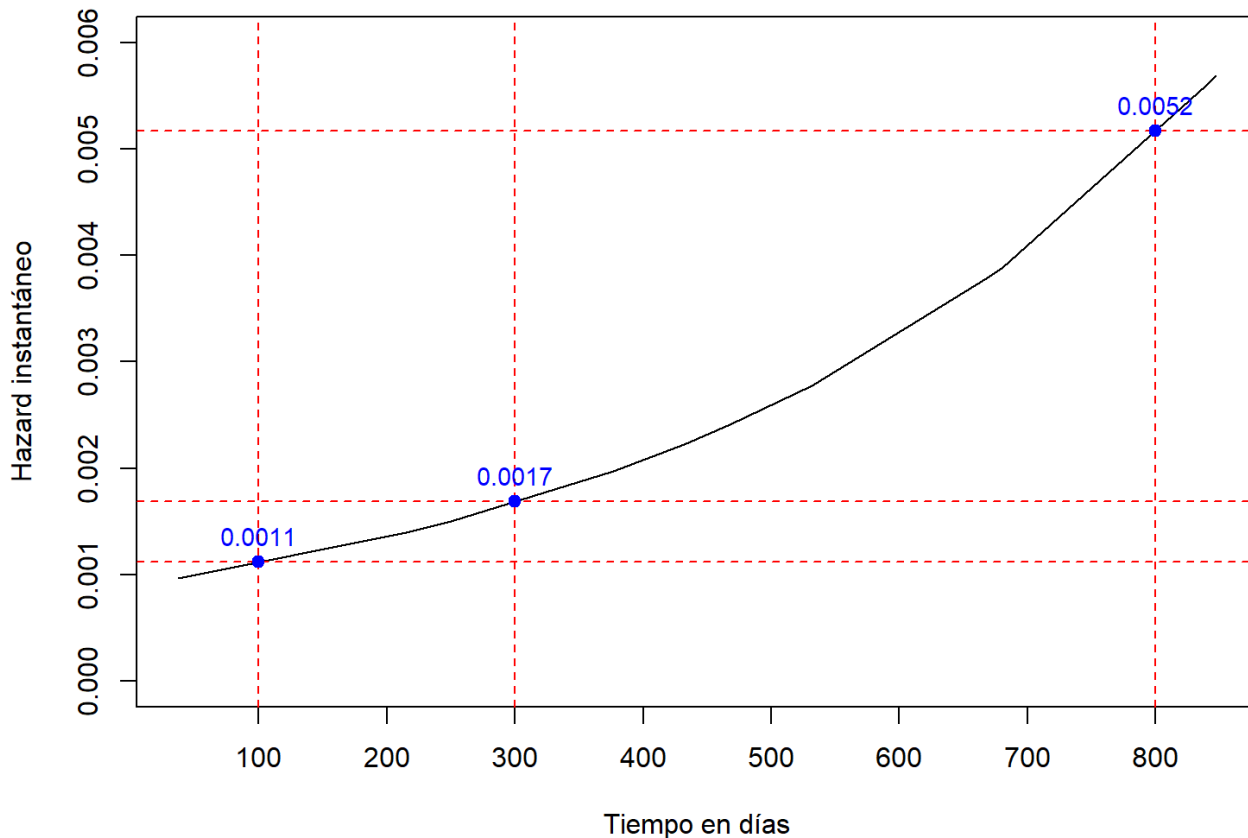
Hazard aproximadamente constante: La tasa instantánea de ocurrencia del evento se mantiene relativamente estable a lo largo del tiempo. Esto implica que el evento puede ocurrir con una intensidad similar en cualquier momento del seguimiento. Este patrón puede observarse, por ejemplo, en algunos estudios clínicos donde se evalúan periodos de seguimiento relativamente cortos, en los cuales factores como la edad o ciertas condiciones basales cambian muy poco durante el tiempo observado, por lo que el riesgo del evento se mantiene aproximadamente constante a lo largo del seguimiento.

Hazard en forma de bañera (bathtub hazard): El hazard es inicialmente alto, luego disminuye durante un periodo intermedio y finalmente vuelve a aumentar con el tiempo. Un ejemplo puede observarse en pacientes sometidos a un trasplante de riñón, donde el riesgo de muerte o complicaciones inmediatamente después de la cirugía es elevado debido a problemas postoperatorios o rechazo temprano del órgano. Cuando el paciente supera esta etapa inicial, el riesgo tiende a estabilizarse mientras el injerto funciona adecuadamente, sin embargo, con el paso de los años, el riesgo vuelve a incrementarse debido al deterioro progresivo del órgano trasplantado o al rechazo crónico.

Hazard unimodal: En este caso el hazard aumenta hasta alcanzar un máximo y posteriormente disminuye. Este patrón puede observarse en algunas enfermedades infecciosas o en complicaciones posteriores a un procedimiento médico, donde el riesgo crece durante un periodo crítico y luego disminuye a medida que los pacientes se recuperan, como por ejemplo con la neumonía, donde el riesgo de complicaciones graves o incluso la probabilidad de morir aumenta durante los primeros días de la infección, alcanzando un pico alrededor del tercer o cuarto día, y luego disminuye a medida que el paciente responde al tratamiento y se recupera.

En el Gráfico 2 se presenta la función de riesgo correspondiente al ejemplo sobre la supervivencia de pacientes después de haber sido diagnosticados con cáncer, analizado previamente. Se observa claramente un **patrón creciente**, lo que indica que el hazard instantáneo aumenta a medida que transcurre el tiempo de seguimiento. Al inicio del estudio (alrededor de los 100 días), la tasa de riesgo instantáneo es aproximadamente 0.0011; posteriormente aumenta a cerca de 0.0017 alrededor del día 300, y alcanza valores cercanos a 0.0052 hacia el día 800.

Gráfico 2. Ejemplo real de la función de hazard instantáneo



Estas tasas son relativamente pequeñas porque el hazard representa una tasa instantánea de ocurrencia del evento en un momento específico del tiempo, es decir, describe la intensidad con la que el evento ocurre en un instante determinado entre los individuos que permanecen en riesgo. Por lo tanto, esta medida no refleja la ocurrencia acumulada del evento a lo largo del seguimiento.

Por esta razón, en muchos análisis resulta útil examinar cómo ese riesgo se acumula a lo largo del tiempo. Para ello se utiliza la función de riesgo acumulado, que permite describir cómo el riesgo total del evento se va incrementando progresivamente a medida que avanza el seguimiento.

Función de riesgo acumulado (cumulative hazard)

Esta función muestra cómo se va acumulando el riesgo de que ocurra el evento a medida que transcurre el tiempo, desde el inicio del seguimiento hasta un tiempo t .

Se define como:

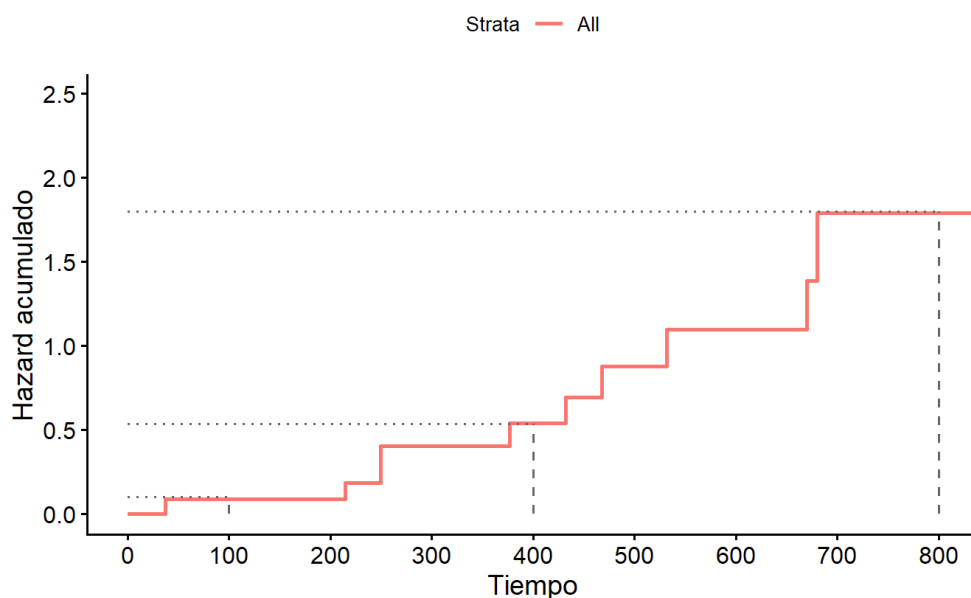
$$H(t) = \int_0^t h(u) du$$

donde:

- $h(u)$ es la función de riesgo instantáneo en el tiempo

Entre mayor sea el valor de $H(t)$, mayor es la acumulación de riesgo. En estudios con varios grupos, si la curva de un grupo crece más lentamente que la de otro, significa que el riesgo se está acumulando más despacio, lo que puede sugerir un efecto beneficioso o protector de una intervención o tratamiento.

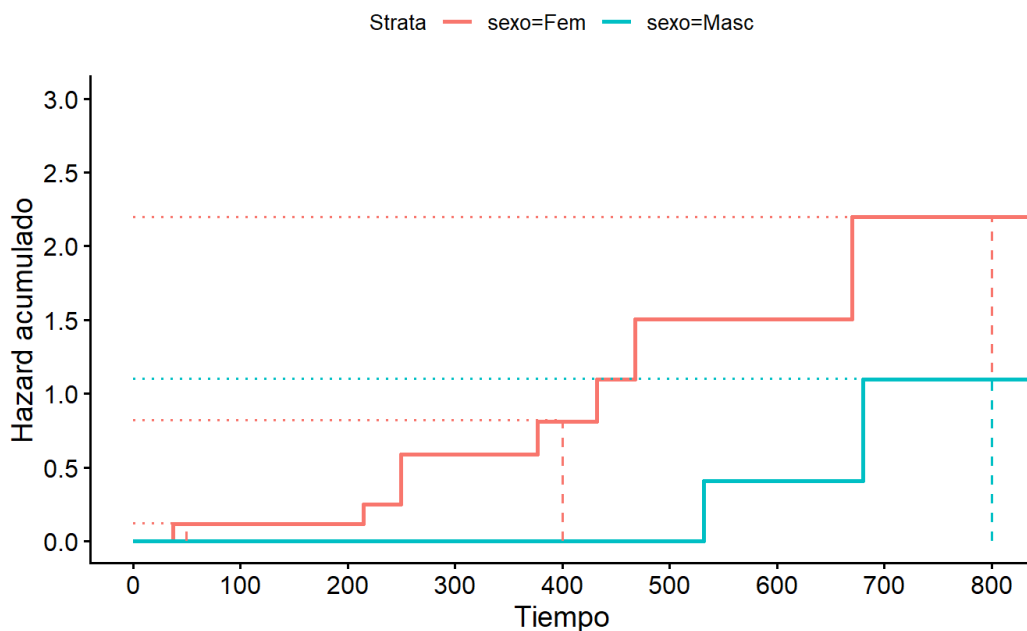
Gráfico 3. Ejemplo de función de Hazard acumulado ($H(t)$) para población total



El Gráfico 3 muestra cómo se acumula el riesgo de que ocurra el evento a lo largo del tiempo (en días), para la población total. Al inicio del seguimiento el aumento del hazard acumulado es

pequeño; por ejemplo, alrededor de los 100 días el valor es cercano a 0.1, lo que indica que hasta ese momento se han presentado pocos eventos. A medida que transcurre el tiempo, el riesgo acumulado aumenta progresivamente: alrededor de los 400 días alcanza valores alrededor de 0.52, reflejando que comienzan a ocurrir más eventos durante el seguimiento. Hacia el final del periodo de observación (alrededor de 800 días o más), el hazard acumulado ya supera valores de 1.8, lo cual indica que una gran parte de los individuos ya ha experimentado el evento hacia ese momento del estudio.

Gráfico 4. Ejemplo de función de Hazard acumulado (H(t)) por sexo

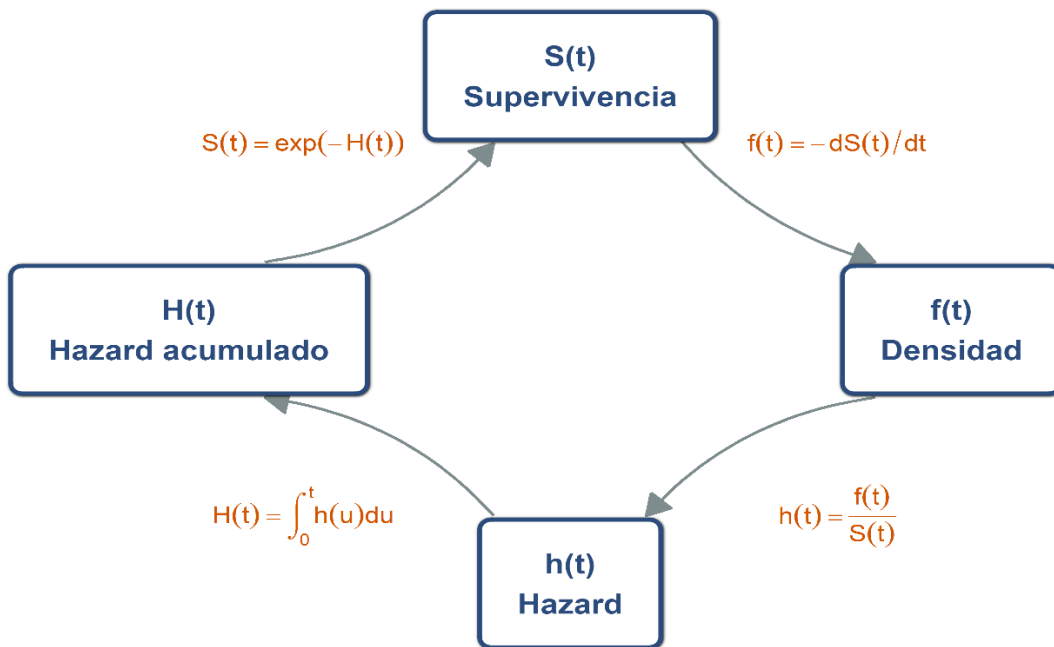


Al analizar el hazard acumulado por sexo (Gráfico 4) se observa que ambos grupos presentan un patrón creciente, lo que indica que el riesgo del evento se incrementa progresivamente a medida que avanza el seguimiento. Sin embargo, durante la mayor parte del periodo observado la curva correspondiente a las mujeres se mantiene por encima de la de los hombres, lo que sugiere una mayor acumulación del riesgo en este grupo. Por ejemplo, alrededor de los 400 días el hazard acumulado en mujeres se aproxima a valores cercanos a 0.82, mientras que en hombres permanece cercano a 0, indicando una ocurrencia más tardía de eventos en este grupo. Hacia el día 700, el riesgo acumulado en hombres aumenta de manera más marcada, alcanzando valores cercanos a 1.1, aunque aún se mantiene por debajo del observado en mujeres que para este tiempo era casi 2.2.

Relación entre la función de riesgo instantáneo, la función de riesgo acumulado y la función de supervivencia

Estas tres funciones están matemáticamente relacionadas entre sí, por lo tanto, es posible derivar las demás a partir de cualquiera de ellas. La Figura 1 muestra cómo la función de supervivencia, la función de riesgo instantáneo y la función de riesgo acumulado están estrechamente relacionadas entre sí, sino diferentes formas de describir la misma distribución del tiempo hasta la ocurrencia del evento. Por ejemplo, lo que nos indican las flechas es que a partir del hazard se puede obtener la función de riesgo acumulado mediante integración y, a partir de esta, se puede obtener la función de supervivencia mediante exponenciación.

Figura 1. Relación entre las funciones $S(t)$, $f(t)$, $h(t)$ y $H(t)$



- $S(t)$ es la función de supervivencia
- $f(t)$ es la función de densidad del tiempo hasta el evento
- $h(t)$ es la función de riesgo instantáneo (hazard)
- $H(t)$ es la función de riesgo acumulado

A manera de ejemplo, veamos a cuanto corresponde la supervivencia en el caso de tener un hazard acumulado de 2.5:

$$S(t) = e^{-H(t)} = e^{-2.5} \approx 0.082$$

Como se mencionó previamente, el hazard acumulado representa la cantidad total de riesgo que se ha acumulado hasta un tiempo t . En este caso, un valor de 2.5 indica que el riesgo total del evento ha sido relativamente alto durante el seguimiento. Al calcular la función de supervivencia a partir de este hazard acumulado, obtenemos un valor de aproximadamente 0.082, lo que significa que la probabilidad de sobrevivir más allá del tiempo t es de aproximadamente el 8.2%.

Forma general del modelo de Cox

El modelo de riesgos proporcionales de Cox describe cómo un conjunto de covariables influye sobre la función de riesgo instantáneo de un evento en el tiempo.

La forma general del modelo es:

$$h(t|X) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)$$

donde:

- $h(t|X)$: función de riesgo para un individuo con covariables X
- $h_0(t)$: función de riesgo basal
- X_1, \dots, X_p : covariables
- β_1, \dots, β_p : coeficientes de regresión

El modelo se compone de dos elementos fundamentales. El primero es la función de riesgo basal, que describe cómo evoluciona el riesgo instantáneo del evento a lo largo del tiempo para un individuo de referencia (por ejemplo, con covariables iguales a cero o en sus categorías basales). El segundo corresponde al componente asociado a las covariables, el cual permite evaluar cómo diferentes características de los individuos modifican de forma multiplicativa dicho riesgo.

Bajo el supuesto de riesgos proporcionales, las covariables actúan de forma multiplicativa sobre el riesgo basal. En consecuencia, los factores incluidos en el modelo modifican la magnitud del riesgo instantáneo, pero no cambian la forma de la función de riesgo en el tiempo, lo cual constituye la esencia del supuesto de riesgos proporcionales.

Para ilustrar esta idea, supóngase que se desea estudiar el riesgo de sufrir un infarto durante un período de diez años. El riesgo basal representaría la evolución del riesgo instantáneo en la población de referencia a lo largo del tiempo, mientras que variables como el tabaquismo, la obesidad o la hipertensión arterial permiten evaluar cómo estos factores modifican el riesgo relativo de ocurrencia del evento.

Naturaleza semiparamétrica del modelo

Como se mencionó anteriormente, el modelo de Cox se clasifica como un modelo semiparamétrico, ya que combina componentes paramétricos y no paramétricos en su formulación. Esta característica permite estimar el efecto de las covariables sobre el riesgo instantáneo del evento sin necesidad de especificar una forma funcional para el cambio en el tiempo del riesgo basal.

El *componente paramétrico* corresponde a los coeficientes asociados a las covariables $\beta_1, \beta_2, \dots, \beta_p$, los cuales cuantifican el efecto de cada variable explicativa sobre el riesgo instantáneo del evento.

Por otra parte, el *componente no paramétrico* corresponde a la función de riesgo basal $h_0(t)$, la cual describe cómo evoluciona el riesgo instantáneo del evento a lo largo del tiempo en la población de referencia. Como se mencionó previamente, este modelo no impone una forma funcional específica para esta función, lo que le da una gran flexibilidad para analizar diferentes patrones de riesgo observados en los datos, por lo tanto, este modelo permite estimar el efecto de las covariables sin asumir una distribución específica para los tiempos de supervivencia.

Conjunto de riesgo (Risk Set)

Como se vio en la semana 1, en el análisis de supervivencia, el conjunto de riesgo corresponde al grupo de individuos que continúan bajo seguimiento y que aún no han experimentado el evento en un tiempo determinado. Para cada tiempo de evento t_i , se define el conjunto de riesgo como $R(t_i)$, el cual incluye a todos los individuos que permanecen en observación y en riesgo de experimentar el evento justo antes del tiempo t_i . A partir de estos conjuntos de riesgo se construye la verosimilitud parcial, en la cual, para cada tiempo de evento, se compara al individuo que experimenta el evento con todos los individuos que se encuentran en su conjunto de riesgo.

Verosimilitud parcial y estimación del modelo de Cox

Como se mencionó previamente, el modelo de Cox permite estimar los coeficientes de regresión sin requerir la especificación de la función de riesgo basal $h_0(t)$, gracias al uso de la verosimilitud parcial propuesta por Cox (1972). Este enfoque se basa en considerar exclusivamente los tiempos de ocurrencia de los eventos y , en cada uno de ellos, comparar al individuo que experimenta el evento con el conjunto de individuos que permanecen en riesgo en ese instante.

Sea t_i el tiempo en el cual ocurre un evento en el individuo i . La contribución de este evento a la verosimilitud parcial se expresa como:

$$\frac{\exp(\beta^T X_i)}{\sum_{j \in R(t_i)} \exp(\beta^T X_j)}$$

Si en el estudio se observan k eventos, la verosimilitud parcial total se define como:

$$L(\beta) = \prod_{i=1}^k \frac{\exp(\beta^T X_i)}{\sum_{j \in R(t_i)} \exp(\beta^T X_j)}$$

En la práctica, se trabaja con la log-verosimilitud parcial:

$$\ell(\beta) = \sum_{i=1}^k \left[\beta^T X_i - \log \left(\sum_{j \in R(t_i)} \exp(\beta^T X_j) \right) \right]$$

Los coeficientes se estiman maximizando esta función mediante métodos numéricos. La función de riesgo basal no necesita especificarse, ya que se cancela algebraicamente al comparar individuos dentro del mismo conjunto de riesgo. Esto permite obtener estimaciones de los efectos de las covariables sin asumir una forma funcional para el riesgo basal, lo que constituye la principal ventaja del modelo de Cox.

Supuesto de riesgos proporcionales

Este supuesto es fundamental para este modelo, ya que establece que la razón de riesgos entre dos individuos permanece constante a lo largo del tiempo. En otras palabras, si se comparan dos individuos con diferentes valores en las covariables, la relación entre sus riesgos instantáneos de experimentar el evento permanece proporcional durante todo el período de seguimiento.

Si se consideran dos individuos con vectores de covariables X_i y X_j , la razón entre sus funciones de riesgo se expresa como

$$\frac{h(t|X_i)}{h(t|X_j)} = \frac{h_0(t)\exp(\beta X_i)}{h_0(t)\exp(\beta X_j)} = \exp(\beta(X_i - X_j))$$

Es importante notar que la función basal $h_0(t)$ es la misma para ambos individuos y, por lo tanto, se cancela. Como consecuencia, esta razón no depende del tiempo, lo que implica que el efecto de las covariables es multiplicativo y constante durante todo el período de seguimiento.

Basados en lo anterior, como los coeficientes aparecen dentro de una función exponencial, su interpretación se realiza generalmente mediante el **hazard ratio**.

Hazard Ratio (HR)

El **hazard ratio** ó **Razón de riesgos instantáneos** se obtiene exponenciando el coeficiente estimado en el modelo de Cox:

$$HR = e^{\beta}$$

Esta medida de efecto indica cuántas veces se multiplica el riesgo instantáneo de ocurrencia del evento cuando la covariable aumenta en una unidad o cuando se comparan dos grupos.

La interpretación habitual es:

- $HR = 1$ → no existe asociación entre la covariable y el riesgo instantáneo de ocurrencia del evento.
- $HR > 1$ → la covariable se asocia con un aumento en el riesgo instantáneo del evento
- $HR < 1$ → la covariable se asocia con una disminución en el riesgo instantáneo del evento

Por ejemplo, en un modelo que evalúa la mortalidad y cuya principal variable de exposición es el hábito de fumar, si el coeficiente estimado para el grupo de fumadores es $\hat{\beta} = 0.69$, entonces:

$$HR = e^{0.69} \approx 2$$

lo que indica que los pacientes fumadores presentan un riesgo instantáneo de muerte aproximadamente dos veces mayor que los no fumadores, manteniendo constantes las demás variables del modelo. En este ejemplo, es importante notar que un valor de 2 implica un incremento del 100% en el riesgo instantáneo.

Una vez establecidos los fundamentos teóricos del modelo, es útil ilustrar su aplicación en un contexto real. A continuación, se presenta un ejemplo basado en datos clínicos, que permite mostrar cómo se estiman los parámetros del modelo y cómo se interpretan los resultados en términos de hazard ratios.

Tabla 1. Modelos de Cox para supervivencia

Predictors	Modelo crudo			Modelo ajustado			Modelo con interacción		
	HR	95% CI	p-value	HR	95% CI	p-value	HR	95% CI	p-value
Placebo (vs Tratamiento)	4.8 2	2.15 – 10.8 1	<0.00 1	4.0 0	1.74 – 9.20	0.001	10.7 5	0.38 – 304.1 6	0.164
log(WBC)				5.4 2	2.81 – 10.4 8	<0.00 1	6.50	2.69 – 15.75	<0.00 1
Placebo × log(WBC)							0.73	0.26 – 2.04	0.546
Observations	42			42			42		
R ² Nagelkerk	0.326			0.679			0.682		
AIC	172.017			143.656			145.297		

Interpretación del Hazard Ratio (HR)

En la Tabla 1 se presentan tres modelos de riesgos proporcionales de Cox aplicados a datos de un estudio clínico en pacientes con leucemia (tomado de Kleinbaum y Klein, 2012), cuyo objetivo fue evaluar el efecto de un tratamiento sobre el tiempo hasta la recaída de la enfermedad en comparación con un grupo placebo. Dado que el conteo basal de leucocitos se reconoce como un factor pronóstico relevante, potencialmente asociado tanto con la exposición como con el desenlace, este se incorporó al análisis mediante su transformación logarítmica (logWBC).

En el modelo crudo, los pacientes asignados al grupo placebo presentan un riesgo instantáneo de recaída aproximadamente 4.82 veces mayor que aquellos en el grupo de tratamiento. Tras ajustar por el logaritmo del conteo de leucocitos (logWBC), este efecto se atenúa ligeramente, pero con un riesgo cuatro veces mayor en el grupo placebo. Por su parte, el logWBC indica que valores más elevados del conteo de leucocitos se asocian con un incremento significativo en el riesgo de recaída.

Finalmente, en el modelo que incorpora un término de interacción, el hazard ratio asociado al tratamiento debe interpretarse de forma condicional al valor de logWBC, y no como un efecto promedio global.

Inferencia en el modelo de Cox

Intervalo de confianza (IC) para el Hazard Ratio

Aunque en el modelo de Cox el parámetro de interés se expresa como el Hazard Ratio (HR), el intervalo de confianza de $(1 - \alpha)\%$ se realiza a través del coeficiente $\hat{\beta}$ y se basa en una aproximación asintótica normal. Lo podemos calcular de la siguiente forma:

$$\hat{\beta} \pm Z_{1-\alpha/2} \cdot SE(\hat{\beta})$$

donde:

$SE(\hat{\beta})$ es el error estándar del estimador y $Z_{1-\alpha/2}$ es el valor crítico de la distribución normal estándar.

Sea L y U los límites inferior y superior del intervalo de confianza para $\hat{\beta}$, respectivamente.

$$L = \hat{\beta} + Z_{1-\alpha/2} \cdot SE(\hat{\beta})$$

$$U = \hat{\beta} - Z_{1-\alpha/2} \cdot SE(\hat{\beta})$$

Para obtener los límites del intervalo en términos del HR, aplicamos la función exponencial así:

$$IC(1 - \alpha)\%_{HR} = (e^L, e^U)$$

Ejemplo

Considerando nuevamente los resultados de la Tabla 1, en el modelo crudo el efecto del tratamiento presenta un IC95% de 2.15 a 10.81. Dado que este intervalo se encuentra completamente por encima de 1, existe evidencia de una asociación positiva estadísticamente significativa al nivel del 5%. No obstante, la amplitud del intervalo sugiere una considerable variabilidad en la estimación, por lo que la magnitud del efecto debe interpretarse con cautela, especialmente en términos de su precisión.

Intervalos de confianza para el HR en presencia de interacción

Cuando se incorpora un término de interacción entre dos covariables X_1 y X_2 , el modelo se especifica como:

$$\eta = \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_1 X_2).$$

Esta formulación implica que el efecto de X_1 sobre el logaritmo del hazard depende del valor de X_2 . Para hacerlo explícito, se puede reescribir el predictor lineal agrupando los términos asociados a X_1 :

$$\eta = (\beta_1 + \beta_3 X_2) X_1 + \beta_2 X_2.$$

De esta expresión se deduce que el efecto de un incremento de una unidad en X_1 , manteniendo X_2 constante, está dado por la combinación lineal:

$$\beta_1 + \beta_3 X_2.$$

En consecuencia, el HR asociado a X_1 condicionado en un valor específico de X_2 es:

$$HR_{X_1|X_2} = \exp(\beta_1 + \beta_3 X_2).$$

La inferencia debe realizarse sobre dicha combinación, por lo que la varianza se obtiene de la siguiente forma:

$$Var(\beta_1 + \beta_3 X_2) = Var(\beta_1) + X_2^2 Var(\beta_3) + 2X_2 Cov(\beta_1, \beta_3),$$

y el error estándar correspondiente es:

$$SE(\beta_1 + \beta_3 X_2) = \sqrt{Var(\beta_1 + \beta_3 X_2)}.$$

El intervalo de confianza en escala logarítmica se construye como:

$$(\beta_1 + \beta_3 X_2) \pm Z_{1 - \alpha/2} \cdot SE(\beta_1 + \beta_3 X_2),$$

y al aplicar la función exponencial a los límites se obtiene el intervalo de confianza para el HR.

Es importante ver que estos intervalos son condicionales al valor de X_2 , por lo que no existe un único efecto de X_1 . En la práctica, el efecto de X_1 se evalúa para valores específicos de X_2 . En términos generales, la inclusión de términos de interacción tiende a incrementar la incertidumbre de las estimaciones. Esto se traduce en intervalos de confianza más amplios, como resultado de la propagación de la varianza asociada a múltiples parámetros y de la posible correlación entre ellos, reflejada en los términos de covarianza.

En el modelo con interacción, algunos intervalos de confianza son más amplios, lo cual es consistente con el incremento en la variabilidad del modelo. En este escenario, es fundamental reconocer que los efectos son condicionales: los HR y sus intervalos de confianza deben interpretarse para combinaciones específicas de covariables. Por tanto, la interpretación no debe centrarse en los coeficientes individuales, sino en los efectos derivados de las combinaciones relevantes, lo que requiere un análisis más cuidadoso y contextualizado.

Nota: Es importante indicar que lo acá presentado es en el caso de una sola interacción. En el caso más general, se recomienda revisar el capítulo XX del libro de Kleinbaum y Klein (2012).

Ejemplo

Si para el modelo con interacción tenemos la siguiente matriz de varianzas y covarianzas:

	Placebo	logWBC	Placebo x logWBC
Placebo	2.9086116	0.5266902	-0.8687027
logWBC	0.5266902	0.2037604	-0.1606420
Placebo x logWBC	-0.8687027	-0.1606420	0.2764537

Podemos estimar el intervalo de confianza del 95% para el log(WBC) en el modelo con interacción para un valor específico de la siguiente forma:

$$HR_{Placebo} = 10.75 \Rightarrow \beta_1 = \log(10.75)$$

$$HR_{interacción} = 0.73 \Rightarrow \beta_3 = \log(0.73)$$

$$\beta_1 \approx 2.375, \beta_3 \approx -0.315$$

Para un valor específico $X_2 = \log(\text{WBC})$:

$$\theta = \beta_1 + \beta_3 X_2$$

- $Var(\beta_1) = 2.9086116$
- $Var(\beta_3) = 0.2764537$
- $Cov(\beta_1, \beta_3) = -0.8687027$

Entonces:

$$Var(\theta) = 2.9086116 + X_2^2(0.2764537) + 2X_2(-0.8687027)$$

Supongamos:

$$\text{Caso: } \log(\text{WBC}) = 1$$

a) Efecto (log-HR)

$$\theta = 2.375 + (-0.315)(1) = 2.060$$

b) Varianza

$$\begin{aligned} Var(\theta) &= 2.9086116 + (1)^2(0.2764537) + 2(1)(-0.8687027) \\ &= 2.9086116 + 0.2764537 - 1.7374054 = 1.4476599 \end{aligned}$$

c) Error estándar

$$SE = \sqrt{1.4476599} \approx 1.203$$

d) Intervalo en log-escala

$$2.060 \pm 1.96(1.203)$$

$$2.060 \pm 2.358 \Rightarrow (-0.298, 4.418)$$

e) HR e Intervalo de confianza

$$HR = \exp(2.060) \approx 7.85$$

$$IC_{95\%} = (\exp(-0.298), \exp(4.418)) \approx (0.74, 82.9)$$

Para $\log(WBC) = 1$, el efecto del tratamiento se estimó como $HR = 7.85$ con $(IC_{95\%}: 0.74-83.4)$, calculado a partir de la combinación lineal de coeficientes y la matriz de varianzas-covarianzas del modelo. La amplitud del intervalo refleja una alta incertidumbre en la estimación.

Evaluación de confusión (Confounding)

La confusión ocurre cuando una variable está asociada tanto con la exposición (tratamiento) como con el desenlace (riesgo), distorsionando la estimación de la asociación entre ambos. Para evaluarla, se compara el efecto estimado en el modelo crudo con el obtenido en un modelo ajustado que incluye variables relevantes. Si la inclusión de una variable en el modelo ajustado produce un cambio significativo en la magnitud del hazard ratio, esto sugiere que dicha variable actúa como un factor de confusión.

En este caso, el modelo crudo muestra un HR de 4.82 para el tratamiento, mientras que el modelo ajustado por $\log(WBC)$ presenta un HR de 4.00. La reducción en la magnitud del estimador tras el ajuste sugiere que parte de la asociación observada inicialmente estaba influenciada por la variable $\log(WBC)$, indicando su papel como potencial factor de confusión.

Suele considerarse que existe confusión relevante cuando el cambio relativo en el estimador supera el 10%. En este caso, la variación relativa es aproximadamente del 17.0%, lo que indicaría la presencia de posible confusión. Sin embargo, este criterio debe interpretarse con cautela, ya que la magnitud del cambio también depende de la variabilidad de los datos, el tamaño de muestra y la precisión de las estimaciones. Por ello, la evaluación de la confusión no debe basarse exclusivamente en un umbral numérico, sino que debe integrarse con la plausibilidad clínica y el conocimiento sustantivo del fenómeno en estudio.

Ejemplo

En el modelo ajustado (Tabla 1), el IC95% (1.74 a 9.20) conserva tanto la dirección como un orden de magnitud similar al observado en el modelo crudo. Esto indica que, tras el ajuste por covariables, el efecto del tratamiento se mantiene relativamente estable, lo que sugiere la ausencia de confusión sustancial o, al menos, que el ajuste no modifica de manera relevante la estimación del efecto.

Pruebas de hipótesis

En el modelo de riesgos proporcionales de Cox, las pruebas de hipótesis permiten evaluar si las covariables están asociadas con el riesgo del evento. Para cada covariable X_j , se contrasta:

$$H_0: \beta_j = 0 \text{ vs } H_1: \beta_j \neq 0$$

Bajo H_0 , no hay asociación ($HR = e^{\beta_j} = 1$).

Las tres pruebas principales son:

- Wald
- Score
- Razón de verosimilitud (LR)

Aunque son asintóticamente equivalentes, difieren en su formulación y desempeño en muestras finitas.

Prueba de Wald

Se basa en el estimador y su error estándar:

$$Z = \frac{\hat{\beta}}{SE(\hat{\beta})}, W = Z^2 \sim \chi_1^2$$

Es fácil de calcular e interpretar, pero puede ser inestable con muestras pequeñas o estimaciones imprecisas.

Prueba Score

Evalúa la pendiente de la log-verosimilitud en $\beta = 0$:

$$U^T I^{-1} U \sim \chi^2$$

No requiere estimar el modelo completo, por lo que es útil para evaluar la inclusión de variables.

Prueba de razón de verosimilitud

Compara modelos anidados:

$$LR = -2(\ell_0 - \ell_1) \sim \chi^2$$

Es generalmente la más robusta, especialmente en muestras pequeñas o moderadas.

Interpretación en el ejemplo

El tratamiento es significativo en el modelo crudo ($p < 0.001$) y ajustado ($p = 0.001$). La variable logWBC también es significativa ($p < 0.001$). En contraste, la interacción no es significativa ($p = 0.546$), lo que sugiere que no hay evidencia de modificación de efecto.

Comparación entre pruebas

Prueba	Idea clave	Ventaja	Limitación
Wald	Usa $\hat{\beta}$ y su SE	Simple	Sensible a varianza
Score	Evalúa en $\beta=0$	No requiere modelo completo	Menos intuitiva
LR	Compara modelos	Más robusta	Requiere 2 modelos

En la práctica, se recomienda priorizar la **prueba de razón de verosimilitud**. En modelos con una sola covariable categórica, la prueba Score es equivalente a la **prueba Log-Rank**.

Ejemplo:

En el caso de la tabla 1, Los valores p reportados en la Tabla 1 corresponden, por defecto, a la prueba de Wald, por lo que permiten evaluar directamente la hipótesis nula $H_0: \beta_j = 0$ para cada covariable incluida en el modelo.

En el modelo crudo, el efecto del tratamiento (Placebo vs Tratamiento) es altamente significativo ($p < 0.001$), lo que indica evidencia estadística suficiente para rechazar la hipótesis nula y concluir que el tratamiento está asociado con el riesgo de recaída.

En el modelo ajustado, el efecto del tratamiento se mantiene significativo ($p = 0.001$), lo que sugiere que la asociación persiste incluso después de controlar por $\log(\text{WBC})$. Por su parte, el $\log(\text{WBC})$ también muestra una asociación estadísticamente significativa con el riesgo ($p < 0.001$), indicando su papel como factor pronóstico relevante.

En el modelo con interacción, el término de interacción (Placebo \times $\log(\text{WBC})$) no resulta estadísticamente significativo ($p = 0.546$), lo que sugiere que no hay evidencia suficiente para afirmar que el efecto del tratamiento varíe según los niveles de $\log(\text{WBC})$. En este escenario, se recomienda eliminar la interacción del modelo, por lo que el modelo más adecuado sería el modelo ajustado.

Comparación de modelos

En el análisis de supervivencia, es frecuente comparar diferentes modelos con el fin de identificar cuál describe mejor la relación entre las covariables y el riesgo del evento. Una estrategia común consiste en comparar **modelos anidados**, es decir, modelos en los que uno de ellos incluye todas las covariables del otro más algunas adicionales.

Por ejemplo:

- Modelo 1: Tratamiento
- Modelo 2: Tratamiento + $\log(\text{WBC})$

La comparación entre estos modelos puede realizarse mediante la **prueba de razón de verosimilitud**, evaluando si la inclusión de nuevas variables mejora significativamente el ajuste del modelo. Además de las pruebas estadísticas, también es posible utilizar criterios de información como:

- **AIC (Akaike Information Criterion)**
- **BIC (Bayesian Information Criterion)**

Estos criterios penalizan la complejidad del modelo, favoreciendo modelos que logren un buen ajuste utilizando un menor número de parámetros. Valores más pequeños de AIC o BIC indican un mejor equilibrio entre ajuste del modelo y complejidad.

Ejemplo

En la Tabla 1, el modelo ajustado presenta un AIC menor que el modelo crudo (143.66 vs 172.02), lo que sugiere que el ajuste por $\log(\text{WBC})$ mejora la capacidad explicativa del modelo. Por otro lado, el modelo con interacción tiene un AIC más alto (145.30), lo que indica que la inclusión del término de interacción no mejora el ajuste del modelo en comparación con el modelo ajustado sin interacción.

Curvas de supervivencia

En el contexto del modelo de Cox, la función de supervivencia está relacionada con el **riesgo acumulado** mediante

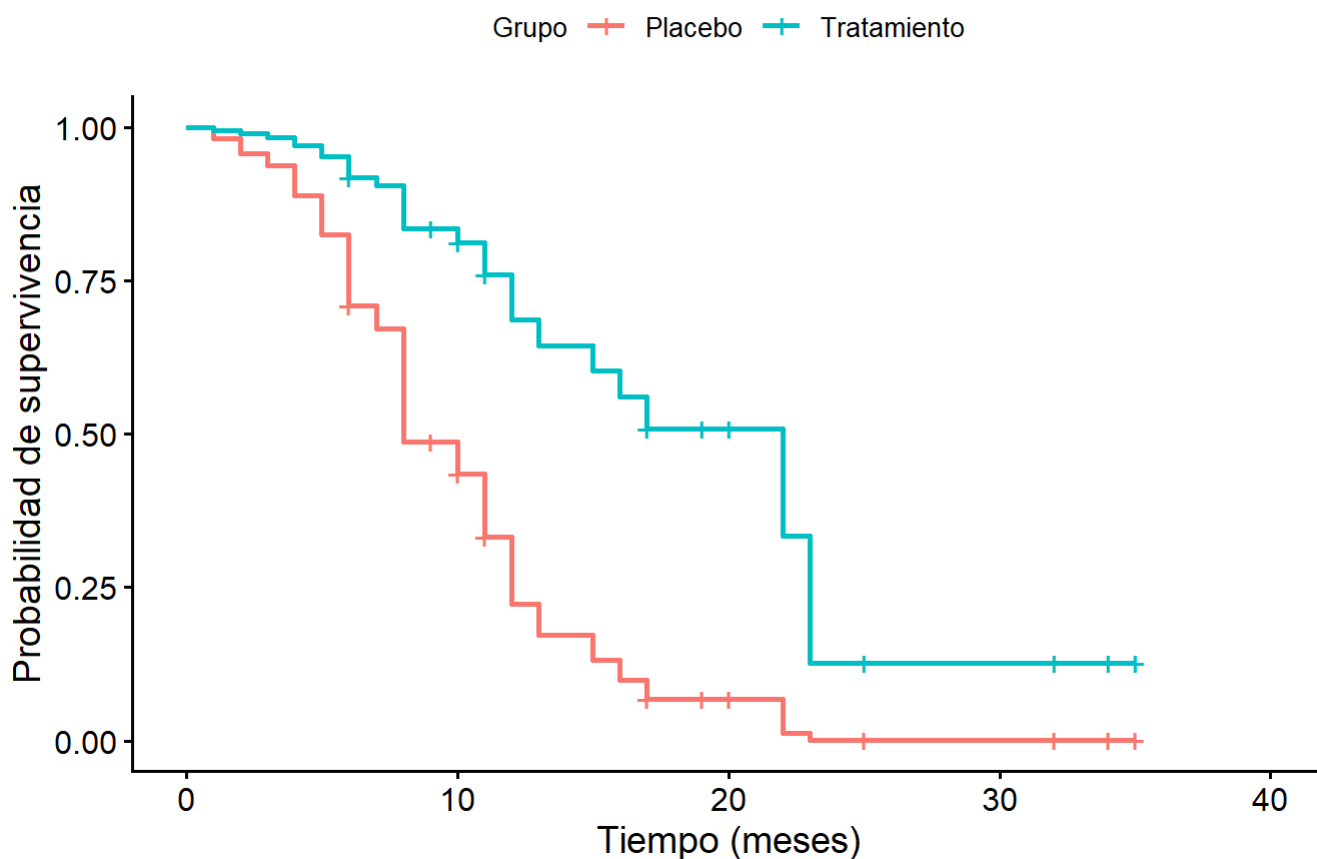
$$S(t) = \exp(-H(t))$$

donde $(H(t))$ corresponde a la función de riesgo acumulado.

Una de las principales aplicaciones de las curvas de supervivencia es **comparar la evolución del riesgo entre diferentes grupos**, como por ejemplo pacientes sometidos a distintos tratamientos.

A partir del modelo de Cox es posible estimar **curvas de supervivencia ajustadas**, las cuales incorporan el efecto de las covariables incluidas en el modelo. Estas curvas permiten visualizar cómo cambian las probabilidades de supervivencia cuando varían las características de los individuos.

Gráfica 5. Curvas de supervivencia ajustadas por log(WBC)



Como podemos observar en la gráfica 5, el grupo de tratamiento presenta una mayor probabilidad de supervivencia a lo largo del tiempo en comparación con el grupo placebo, lo que es consistente con los resultados obtenidos en los modelos de Cox. Además, al ajustar por log(WBC), estas curvas reflejan el efecto del conteo de leucocitos en la supervivencia, proporcionando una visión más precisa de la relación entre el tratamiento y el riesgo de recaída.

Ejemplo de Aplicación

Evaluación de un nuevo tratamiento para falla cardíaca

Un grupo de investigadores en cardiología ha desarrollado un nuevo tratamiento para pacientes con falla cardíaca. En análisis preliminares se observó que los pacientes que recibieron este tratamiento parecían presentar una mayor supervivencia en comparación con aquellos que recibieron el tratamiento estándar. Sin embargo, esta diferencia podría no reflejar necesariamente un efecto real del tratamiento, ya que podría estar influenciada por otras características de los pacientes. Por ejemplo, es posible que quienes recibieron el nuevo tratamiento fueran más jóvenes o tuvieran menos comorbilidades, lo que indicaría una mejor condición de salud al inicio del seguimiento.

Con el fin de evaluar esta posibilidad, se realizó un estudio en el que se seleccionó una muestra de pacientes de la institución y se efectuó un seguimiento a lo largo del tiempo, registrando la ocurrencia del evento de interés, definido como la muerte (falla_cardiaca.xlsx). El objetivo del análisis es determinar si existen diferencias en el riesgo de muerte entre los grupos de tratamiento, ajustando por posibles factores de confusión, mediante un análisis de supervivencia utilizando el modelo de riesgos proporcionales de Cox.

Las variables evaluadas en el estudio fueron las siguientes:

- **tiempo:** tiempo de seguimiento (en meses) hasta la ocurrencia del evento o la censura.
- **evento:** estado vital del paciente al final del seguimiento (1 = fallece, 0 = no fallece (censurado))
- **tratamiento:** tipo de tratamiento recibido (0 = tratamiento estándar, 1 = nuevo tratamiento)
- **edad:** edad del paciente al inicio del seguimiento (en años cumplidos).
- **diabetes:** diagnóstico de diabetes mellitus al inicio del seguimiento (0 = no, 1 = sí)

Base de datos

Nota: La base de datos se encuentra en la plataforma, por si desean replicar el ejercicio.

Abriendo la base datos

Crearemos un objeto llamado `base` a partir del archivo “falla_cardiaca.xlsx” utilizando la función `read_excel()` del paquete `readxl`.

```
library(readxl)
base <- read_excel("falla_cardiaca.xlsx")
```

Visualizando la información de la base de datos

Podemos revisar que contiene la base de datos utilizando la opción `paged_table()` del paquete `rmarkdown`, que nos permite visualizar la información de manera ordenada y paginada, facilitando la revisión de los datos sin sobrecargar la pantalla. En caso de hacerlo en la consola, se recomienda utilizar la función `head()` para mostrar solo las primeras filas de la base de datos, lo que permite obtener una vista general sin imprimir toda la información.

```
library(rmarkdown)
paged_table(
  base,
  options = list(rows.print = 20)
)
```

tiempo <dbl>	evento <dbl>	tratamiento <chr>	edad <dbl>	diabetes <chr>
1.6	1	Estándar	67	Si
31.9	1	Estándar	65	No
13.4	1	Estándar	65	Si
12.0	0	Nuevo	76	Si
9.6	1	Estándar	63	No
23.9	1	Nuevo	77	No
18.8	0	Nuevo	53	No
33.4	0	Nuevo	70	No
24.1	0	Estándar	66	No
5.3	1	Estándar	67	No
24.4	0	Nuevo	68	No
7.8	1	Nuevo	61	No
25.6	0	Nuevo	62	Si
6.7	1	Estándar	57	No
38.2	0	Nuevo	56	No
28.1	1	Estándar	67	Si
8.8	1	Nuevo	69	Si
5.6	0	Estándar	65	No
1.5	1	Estándar	72	No
22.2	0	Estándar	81	Si

1-20 of 100 rows

Análisis descriptivo preliminar

Inicialmente, es importante realizar un análisis descriptivo de las variables incluidas en la base de datos para comprender su distribución y características. Esto nos permitirá identificar posibles patrones, valores atípicos o inconsistencias que podrían afectar el análisis posterior.

Distribución del evento de interés (muerte, 1=fallece, 0=no fallece)

Para esta variable, es fundamental conocer la proporción de pacientes que experimentaron el evento (fallecimiento) en comparación con aquellos que no lo hicieron (censurados), ya que, en caso de tener muy pocos eventos, el análisis de supervivencia podría verse afectado por la falta de información suficiente para estimar adecuadamente los riesgos.

```
table(base$evento)
```

```
##  
##  0  1  
## 48 52
```

Podemos ver que contamos con 52 eventos de fallecimiento y 48 pacientes censurados, lo que indica una proporción de eventos adecuada para realizar un análisis de supervivencia.

Distribución del tiempo de seguimiento según grupo de tratamiento

Otra variable relevante es el tiempo de seguimiento, ya que nos permite evaluar la duración del seguimiento para cada paciente y cómo esta variable se distribuye según el grupo de tratamiento y la presencia o ausencia de diabetes. Esto es importante para identificar posibles diferencias en el tiempo de seguimiento entre los grupos, lo que podría influir en la estimación del riesgo y la interpretación de los resultados.

Inicialmente revisemos la distribución del tiempo de seguimiento según el grupo de tratamiento (0=Tratamiento estándar, 1=Nuevo tratamiento). Inicialmente, podemos calcular la media del tiempo de seguimiento para cada grupo utilizando la función `tapply()`, que nos permite aplicar una función (en este caso, la media) a un vector (tiempo) agrupado por otra variable (tratamiento).

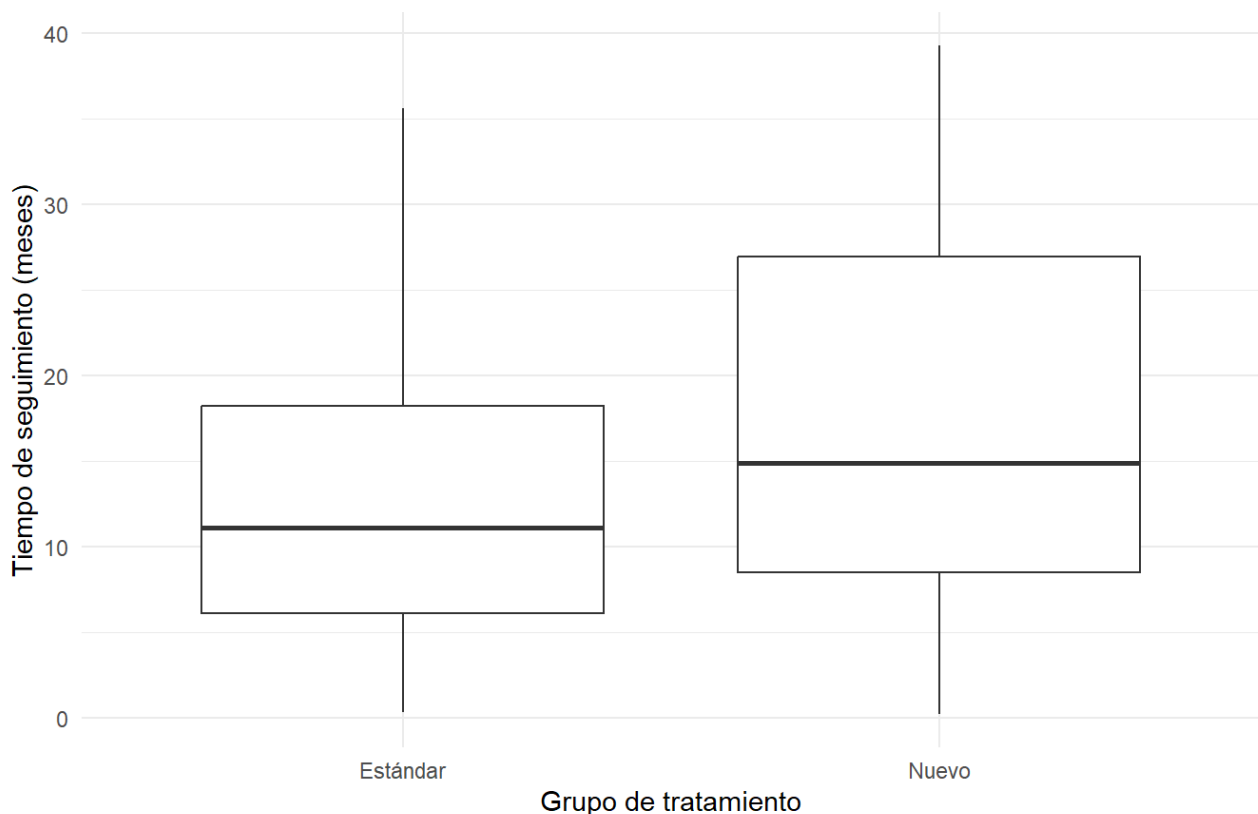
```
tapply(base$tiempo, base$tratamiento, mean)
```

```
## Estándar      Nuevo
## 13.49825 18.17209
```

Otra alternativa es a través de un gráfico de boxplot, el cual nos permite visualizar y comparar la distribución del tiempo de seguimiento entre los grupos de tratamiento. Su principal ventaja es que podemos ver directamente la mediana, los cuartiles y los posibles valores atípicos.

```
library(ggplot2)
ggplot(base, aes(x = tratamiento, y = tiempo)) + geom_boxplot() +
  labs(title = "Gráfica 6. Distribución del tiempo de seguimiento por grupo
de tratamiento", x = "Grupo de tratamiento", y = "Tiempo de seguimiento (meses)") +
  theme_minimal()
```

Gráfica 6. Distribución del tiempo de seguimiento por grupo de tratamiento



El grupo que recibe el tratamiento nuevo presenta tiempos de seguimiento más prolongados en comparación con el grupo estándar, aunque con mayor heterogeneidad. En particular, la mediana

de seguimiento es superior en el grupo nuevo (≈ 15 meses frente a ≈ 11 meses), acompañada de una mayor dispersión en los datos.

En ambos grupos se observa asimetría positiva, lo que indica una mayor concentración de observaciones en tiempos bajos y una cola hacia valores más altos. Estos hallazgos sugieren una mayor acumulación de tiempo observado en el grupo con tratamiento nuevo; sin embargo, esta diferencia debe interpretarse con cautela, dado que el gráfico no incorpora información sobre la censura ni la ocurrencia del evento, por lo que resulta necesario complementarlo con métodos de análisis de supervivencia.

Distribución del tiempo de seguimiento según presencia o no de diabetes (0=No 1=Si)

También vale la pena revisar la distribución del tiempo de seguimiento según la presencia o ausencia de diabetes, ya que esta condición podría estar asociada con un mayor riesgo de eventos adversos y, por lo tanto, influir en el tiempo hasta el evento. Al analizar esta variable, podemos identificar si los pacientes con diabetes tienen tiempos de seguimiento más cortos en comparación con aquellos sin diabetes, lo que podría indicar una mayor vulnerabilidad y un impacto negativo en la supervivencia.

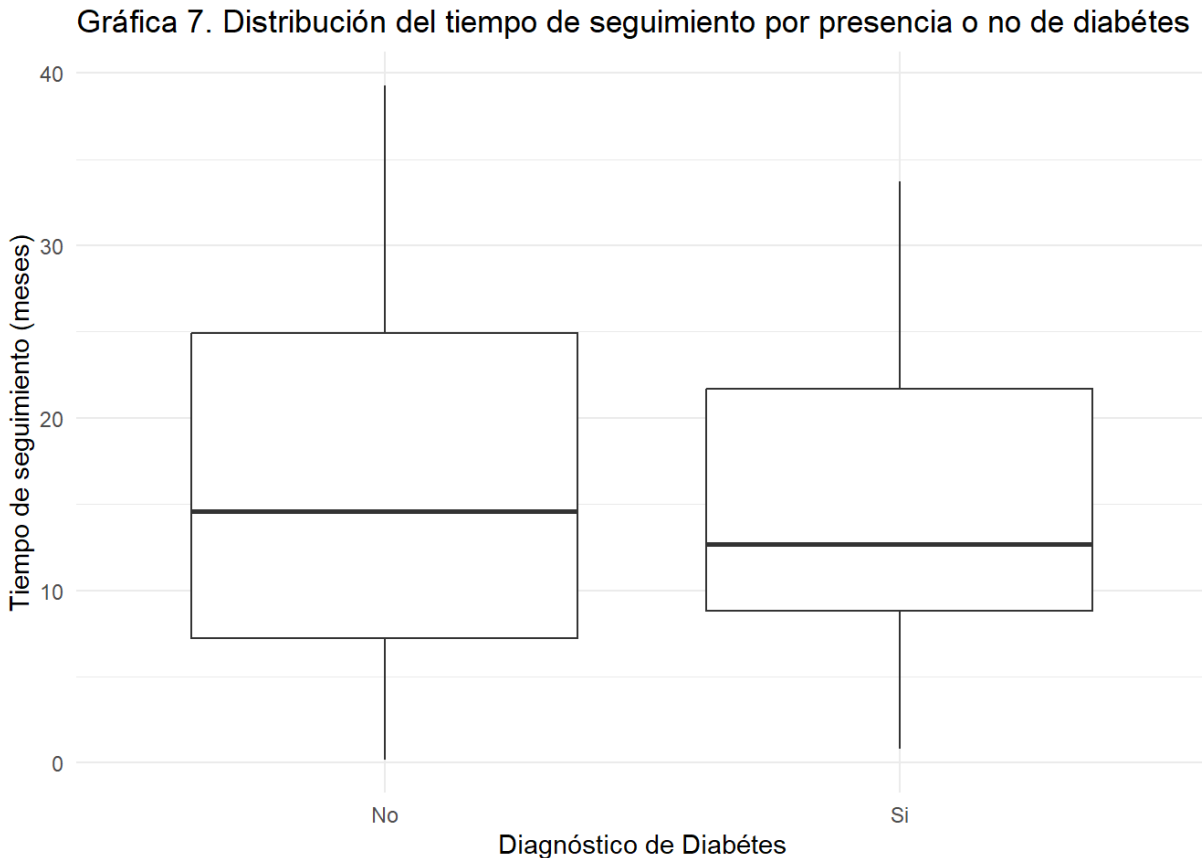
```
tapply(base$tiempo, base$diabetes, mean)
##      No      Si
## 15.99851 14.51212
```

Al calcular la media del tiempo de seguimiento para cada grupo observamos que los pacientes con diabetes presentan un tiempo de seguimiento promedio más corto en comparación con aquellos sin diabetes, lo que podría indicar una mayor vulnerabilidad y un impacto negativo en la supervivencia.

Veámoslo gráficamente:

```
library(ggplot2)
```

```
ggplot(base, aes(x = diabetes, y = tiempo)) + geom_boxplot() +
  labs(title = "Gráfica 7. Distribución del tiempo de seguimiento por presencia o no
de diabétes", x = "Diagnóstico de Diabétes", y = "Tiempo de seguimiento (meses)") +
  theme_minimal()
```



Podemos ver que los individuos sin diagnóstico de diabetes presentan un tiempo mediano de seguimiento ligeramente mayor (≈ 15 meses, en comparación con ≈ 13 meses), aunque ambas distribuciones evidencian asimetría positiva, con presencia de valores altos de seguimiento, aunque el grupo sin diabetes alcanza máximos más elevados. En conjunto, no se observan diferencias marcadas en la distribución del tiempo de seguimiento entre los grupos, por lo que cualquier posible efecto de la diabetes debería evaluarse mediante métodos de análisis de supervivencia que consideren la censura y la ocurrencia del evento.

Evaluación de la supervivencia

Definiendo el objeto de supervivencia

Antes de ajustar un modelo de supervivencia, definimos la variable respuesta utilizando la función `Surv()`, que combina el tiempo de seguimiento y la variable evento.

```
library(survival)
Surv(base$tiempo, base$evento)
## [1] 1.6 31.9 13.4 12.0+ 9.6 23.9 18.8+ 33.4+ 24.1+ 5.3 24.4+ 7.8
## [13] 25.6+ 6.7 38.2+ 28.1 8.8 5.6+ 1.5 22.2+ 33.7 33.4+ 3.3 9.9
## [25] 35.6+ 9.7 14.9+ 28.3+ 8.4 1.3 23.4+ 7.1+ 10.1 9.2+ 8.1+ 3.9
## [37] 10.6 16.8+ 17.0 6.1+ 12.7 29.0+ 12.9+ 16.1+ 11.1+ 26.8 4.0 2.0
## [49] 9.5 18.1+ 27.0+ 8.4+ 0.9 15.7 25.4+ 29.6 25.7 31.4+ 39.3+ 6.5+
## [61] 2.1 0.2 8.9 18.2+ 6.1 17.8+ 14.6+ 18.3 17.7+ 15.7+ 31.6+ 34.6+
## [73] 0.8 9.5 0.3 12.7+ 7.4 15.4 33.7+ 12.6 25.8+ 14.9 1.8 18.5+
## [85] 8.2+ 26.9+ 23.8 8.0+ 11.5 2.0 12.9 37.1+ 21.7+ 24.8+ 7.3 10.3
## [97] 11.9 1.8 0.3 17.3+
```

Esta función identifica los tiempos de seguimiento y los diferencia entre aquellos que corresponden a eventos (fallecimientos) y aquellos que corresponden a censura (pacientes que no fallecieron durante el seguimiento). El resultado es un objeto de clase `Surv`, que encapsula esta información y se utiliza como variable respuesta en el modelo de Cox.

Estimado el tiempo de Supervivencia utilizando el método de Kaplan-Meier

A continuación procederemos a calcular el tiempo de supervivencia utilizando la función `survfit()`, que estima la función de supervivencia a partir de los datos de tiempo y evento, sin estratificar por ninguna covariable.

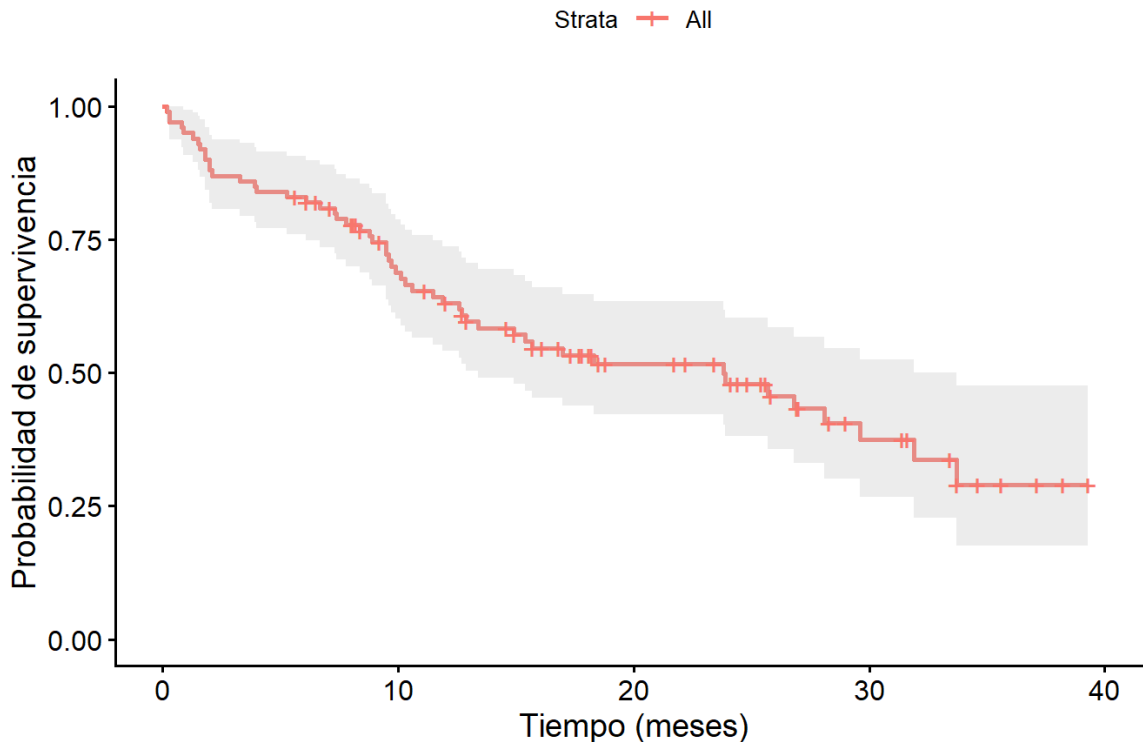
```
ajuste_km <- survfit(Surv(tiempo, evento) ~ 1, data = base)
```

Graficamente tenemos:

```
library(survminer)
ggsurvplot(
```

```
ajuste_km,
data = base,
conf.int = TRUE,
xlab = "Tiempo (meses)",
ylab = "Probabilidad de supervivencia",
title = "Gráfica 8. Curva de supervivencia global (Kaplan-Meier)"
)
```

Gráfica 8. Curva de supervivencia global (Kaplan-Meier)



Encontramos una disminución progresiva de la supervivencia conforme avanzan los meses de seguimiento. Aproximadamente, la supervivencia se sitúa alrededor de 0.75 hacia los 10 meses, desciende a cerca de 0.50 entre los 18 y 20 meses, y continúa disminuyendo hasta valores cercanos a 0.30 hacia los 35–40 meses.

Las marcas sobre la curva indican observaciones censuradas, es decir, individuos que no presentaron el evento durante su tiempo de seguimiento. El intervalo del 95% de confianza se ensancha a medida que aumenta el tiempo, reflejando una mayor incertidumbre en las estimaciones debido a la reducción del número de individuos en riesgo.

Veamos si hay diferencia según el grupo de tratamiento

```
ajuste_trat <- survfit(Surv(tiempo, evento) ~ tratamiento, data = base)
```

Calculemos la mediana y los intervalos del 95% de confianza para cada grupo de tratamiento:

```
summary(ajuste_trat)$table
```

##	records	n.max	n.start	events	rmean	se(rmean)	median
## tratamiento=Estándar	57	57	57	36	17.99708	1.933933	15.4
## tratamiento=Nuevo	43	43	43	16	26.94864	2.400604	NA
##	0.95LCL	0.95UCL					
## tratamiento=Estándar	10.1	28.1					
## tratamiento=Nuevo	23.9	NA					

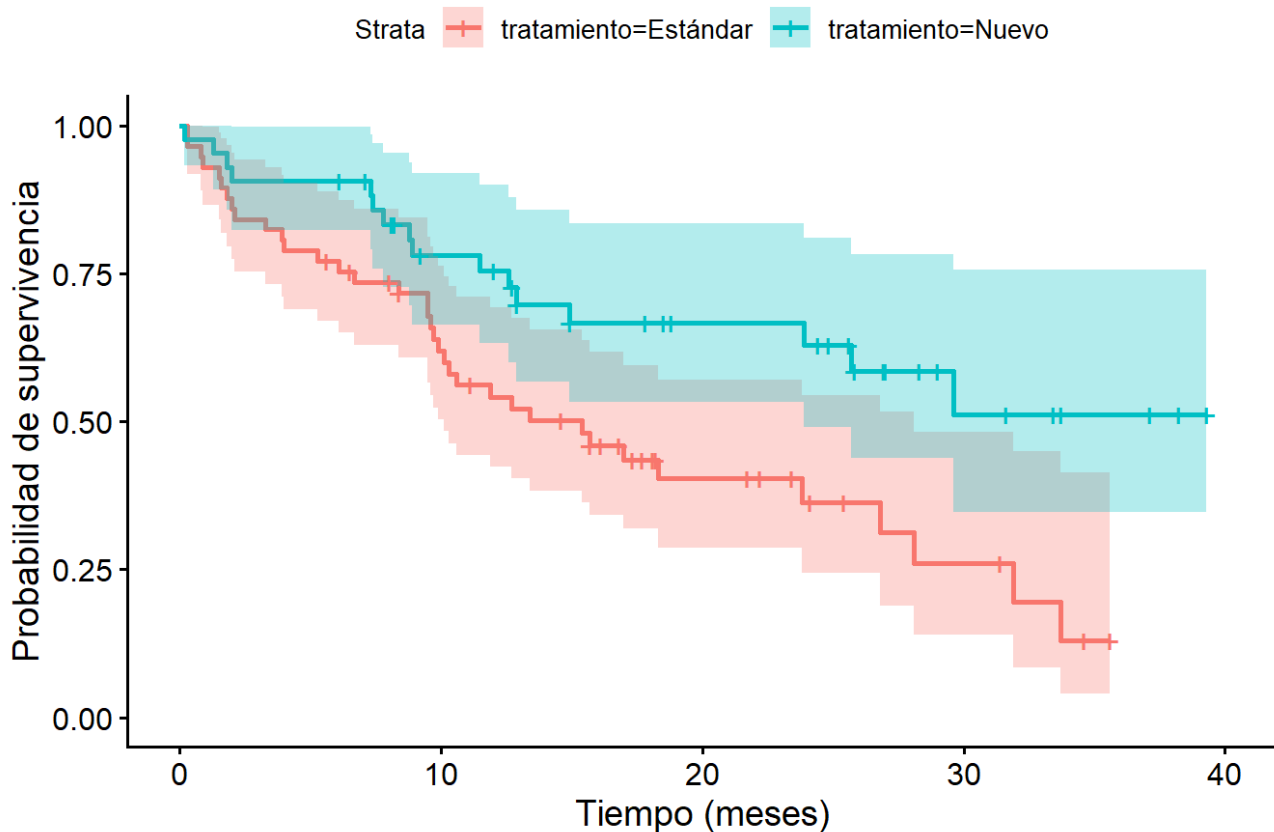
Vemos que hay diferencias relevantes entre las curvas de supervivencia de los dos grupos de tratamiento. El grupo que recibió el tratamiento estándar presentó una mediana de supervivencia de 15.4 meses (IC 95%: 10.1–28.1), lo que indica que el 50% de los pacientes experimentó el evento antes de este tiempo, mientras que en el grupo con el tratamiento nuevo no fue posible estimar la mediana de supervivencia, debido a que la función de supervivencia no descendió por debajo del 50% durante el periodo de seguimiento. Este hallazgo sugiere una menor ocurrencia del evento y, en consecuencia, una mayor supervivencia en comparación con el grupo estándar.

Veamos la gráfica

```
ggsurvplot(
  ajuste_trat,
  data = base,
  conf.int = TRUE,
  xlab = "Tiempo (meses)",
```

```
ylab = "Probabilidad de supervivencia",
title = "Gráfica 9. Curva de supervivencia global (Kaplan-Meier)"
)
```

Gráfica 9. Curva de supervivencia global (Kaplan-Meier)



Podemos observar una clara separación entre los grupos de tratamiento; el grupo con tratamiento nuevo mantiene de forma consistente una mayor probabilidad de supervivencia en comparación con el grupo estándar. A lo largo del seguimiento, el grupo estándar presenta un descenso más pronunciado, indicando una mayor ocurrencia del evento, mientras que el grupo del nuevo tratamiento conserva niveles de supervivencia más elevados; específicamente, alrededor de los 20 meses la supervivencia es cercana a 0.40–0.45 en el grupo estándar frente a aproximadamente 0.65–0.70 en el grupo nuevo. Aunque las bandas de confianza presentan cierta superposición, especialmente en etapas tempranas, la diferencia entre las curvas se mantiene en el tiempo, lo que

sugiere un posible efecto beneficioso del tratamiento nuevo, el cual debe confirmarse mediante pruebas estadísticas formales como el test de log-rank o el modelo de Cox.

Estimado el efecto del tratamiento a través de un modelo de Cox

Para evaluar el efecto del tratamiento, haremos el modelo de Cox incluyendo el tratamiento como variable independiente, así:

```
mod_cox <- coxph(Surv(tiempo, evento) ~ tratamiento, data = base)
summary(mod_cox)

## Call:
## coxph(formula = Surv(tiempo, evento) ~ tratamiento, data = base)
##
## n= 100, number of events= 52
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## tratamientoNuevo -0.7870    0.4552  0.3031 -2.597  0.00941 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## tratamientoNuevo    0.4552      2.197   0.2513   0.8245
##
## Concordance= 0.583 (se = 0.036 )
## Likelihood ratio test= 7.29 on 1 df,  p=0.007
## Wald test              = 6.74 on 1 df,  p=0.009
## Score (logrank) test = 7.08 on 1 df,  p=0.008
```

Una alternativa para presentar los resultados es usando el comando `tab_model` de la librería `sjPlot`, que permite graficar los coeficientes del modelo de Cox y sus intervalos de confianza de una forma más visual e intuitiva.

```
library(sjPlot)
```

```
tab_model(mod_cox, show.ci = 0.95, show.se = FALSE, show.p = TRUE, show.aic = TRUE)
```

Surv(tiempo,evento)			
<i>Predictors</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
tratamiento [Nuevo]	0.46	0.25 – 0.82	0.009
Observations	100		
R ² Nagelkerke	0.071		
AIC	412.949		

En el resultado se muestra el hazard ratio (HR) para el tratamiento, junto con su intervalo de confianza al 95% y el valor p asociado a la prueba de hipótesis. En este caso, el HR es de 0.46 con un IC95% de 0.25 a 0.82, lo que indica que el riesgo de muerte en el grupo de tratamiento nuevo es aproximadamente un 54% menor que en el grupo de tratamiento estándar. Además, el intervalo de confianza no incluye el valor 1, lo que sugiere que esta asociación es estadísticamente significativa al nivel del 5% y que se confirma con el valor p de 0.009.

Este resultado es consistente con lo observado en las curvas de supervivencia, sin embargo, una de las ventajas del modelo de Cox es que permite ajustar por otras variables lo que es fundamental para evaluar si realmente el efecto se debe al tratamiento.

A continuación, evaluaremos si el efecto del tratamiento se mantiene ajustando por edad y diabetes.

Modelo de Cox ajustando por covariables (Modelo Multivariado)

Vamos a correr el modelo incluyendo las covariables edad, presencia o no de diabetes y la interacción entre esta última y tratamiento.

```
mod_mult <- coxph( Surv(tiempo,evento) ~ tratamiento*diabetes + edad, data = base )
```

```
library(sjPlot)
```

```
tab_model(mod_cox, mod_mult, dv.labels = c("Modelo Crudo", "Modelo con Interacción"),
  show.ci = 0.95, show.se = FALSE, show.p = TRUE)
```

<i>Predictors</i>	Modelo Crudo			Modelo con Interacción		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
tratamiento [Nuevo]	0.46	0.25 – 0.82	0.009	0.26	0.11 – 0.62	0.002
diabetes [Si]				0.78	0.35 – 1.72	0.537
edad				1.06	1.02 – 1.11	0.002
tratamiento [Nuevo] × diabetes [Si]				2.77	0.77 – 10.01	0.120
Observations	100			100		
R ² Nagelkerke	0.071			0.176		

¿Hay efecto de interacción entre el tratamiento y la diabetes?

Inicialmente evaluamos si hay efecto de interacción entre el tratamiento y la diabetes usando el valor p asociado a este término en el modelo multivariado. Como podemos ver, el valor p para la interacción entre tratamiento y diabetes es de 0.120 (>0.05), lo que indica que no hay evidencia estadística suficiente para afirmar que el efecto del tratamiento varíe según la presencia o ausencia de diabetes, es decir, no hay efecto de interacción y podemos eliminarla del modelo.

```
mod_red <- coxph( Surv(tiempo,evento) ~ tratamiento + diabetes + edad, data = base)

tab_model(mod_cox, mod_mult, mod_red, dv.labels = c("Modelo Crudo", "Modelo con Interacción", "Modelo Reducido"),
  show.ci = 0.95, show.se = FALSE, show.p = TRUE)
```

Predictors	Modelo Crudo			Modelo con Interacción			Modelo Reducido		
	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p
tratamiento [Nuevo]	0.46	0.25 – 0.82	0.009	0.26	0.11 – 0.62	0.002	0.41	0.22 – 0.74	0.003
diabetes [Si]				0.78	0.35 – 1.72	0.537	1.13	0.63 – 2.05	0.677
edad				1.06	1.02 – 1.11	0.002	1.06	1.02 – 1.10	0.004
tratamiento [Nuevo] × diabetes [Si]				2.77	0.77 – 10.01	0.120			
Observations	100			100			100		
R ² Nagelkerke	0.071			0.176			0.155		

Al evaluar este modelo ajustado, se observa que el efecto del tratamiento se mantiene estadísticamente significativo (HR = 0.41; IC95%: 0.22–0.74; p = 0.003), lo que indica que el tratamiento nuevo reduce el riesgo de muerte incluso después de controlar por edad y diabetes. Asimismo, la edad presenta una asociación significativa con el riesgo (HR = 1.03; IC95%: 1.01–1.06; p = 0.015), sugiriendo que cada año adicional incrementa el riesgo de muerte en aproximadamente un 3%. En contraste, la diabetes no muestra una asociación estadísticamente significativa (HR = 1.40; IC95%: 0.80–2.45; p = 0.240), lo que indica que, en este modelo, no se evidencia un efecto independiente de esta variable sobre el riesgo de muerte tras el ajuste por las demás covariables.

Sin embargo, antes de considerar la eliminación de esta variable del modelo, evaluaremos si actúa como un factor de confusión en la asociación entre el tratamiento y el riesgo de muerte.

Postestimación en modelos de Cox con interacción

Cuando el término de interacción es estadísticamente significativo, la interpretación de los coeficientes individuales deja de ser marginal. En otras palabras, el efecto de una variable depende del nivel de la otra, por lo que la postestimación debe centrarse en combinaciones lineales de parámetros para obtener efectos específicos.

En este contexto, la postestimación consiste en:

- Definir el contraste de interés (combinación lineal de coeficientes).
- Estimar su valor, error estándar e inferencia.
- Transformar a escala interpretable (por ejemplo, hazard ratio mediante exponenciación).

En caso de que la interacción hubiera dado significativa, podríamos utilizar la función `emmeans()` del paquete `emmeans` para obtener estimaciones marginales de los efectos del tratamiento en cada nivel de diabetes, lo que nos permitiría interpretar cómo varía el efecto del tratamiento según la presencia o ausencia de diabetes.

Revisamos el nombre generado para las variables en el modelo

```
names(mod_mult$coefficients)
## [1] "tratamientoNuevo"          "diabetesSi"
## [3] "edad"                       "tratamientoNuevo:diabetesSi"
```

Calculando el valor para efecto del nuevo tratamiento en comparación con el tratamiento estándar considerando la interacción con diabetes

```
library(multcomp)
res <- glht(mod_mult, linfct = c("tratamientoNuevo + tratamientoNuevo:diabetesSi = 0"))

s <- summary(res)
ci <- exp(confint(res)$confint)

HR <- ci[, "Estimate"]
LI <- ci[, "lwr"]
LS <- ci[, "upr"]
p <- s$test$pvalues

data.frame(
```

```
HR_CI = paste0(round(HR,3), " (", round(LI,3), "-", round(LS,3), ")"),
p_value = round(p, 4)
)
## HR_CI p_value
## 1 0.727 (0.286-1.847) 0.5026
```

Como podemos ver, el efecto del tratamiento nuevo en comparación con el tratamiento estándar ajustando por la interacción es de 0.73 (IC95%: 0.29–1.85; $p = 0.5026$), lo que ratifica que realmente tener diabetes no afecta esta relación.

¿Es diabétes un factor de confusión?

Predictors	Modelo Crudo			Modelo Reducido con diabetes			Modelo Reducido sin diabetes		
	Estimates	CI	<i>p</i>	Estimates	CI	<i>p</i>	Estimates	CI	<i>p</i>
tratamiento [Nuevo]	0.46	0.25 – 0.82	0.009	0.41	0.22 – 0.74	0.003	0.41	0.22 – 0.75	0.004
diabetes [Si]				1.13	0.63 – 2.05	0.677			
edad				1.06	1.02 – 1.10	0.004	1.06	1.02 – 1.10	0.002
Observations	100			100			100		
R ² Nagelkerke	0.071			0.155			0.154		

Al comparar el modelo ajustado que incluye diabetes con el modelo reducido que la excluye, vemos que el hazard ratio para el tratamiento se mantiene igual (HR = 0.41), lo que sugiere que la inclusión de la variable diabetes no produce un cambio significativo en la estimación del efecto del tratamiento. Además, el valor p para el tratamiento sigue siendo significativo en ambos modelos ($p = 0.003$ con diabetes y $p = 0.004$ sin diabetes), lo que indica que la asociación entre el tratamiento y el riesgo de muerte se mantiene robusta independientemente de la inclusión de la variable diabetes.

En conclusión, el mejor modelo para este ejemplo es el modelo reducido sin diabetes.

Interpretación de los resultados del modelo

El tratamiento nuevo se asocia con una reducción del riesgo de muerte en comparación con el tratamiento estándar, ajustando por la edad (HR = 0.41; IC95%: 0.22–0.74; p = 0.003). En términos del modelo, los pacientes que reciben el tratamiento nuevo presentan un 59% menor riesgo instantáneo de morir en comparación con aquellos que reciben el tratamiento estándar, manteniendo constante la edad.

Por su parte, la edad muestra una asociación positiva con el riesgo de muerte (HR = 1.06; IC95%: 1.02;1.10; p = 0.004). Específicamente, cada año adicional de edad se asocia con un incremento aproximado del 6% en el riesgo instantáneo de muerte, manteniendo constante el tratamiento.

Comparación de los modelos

Esta sección se centra en la comparación de los modelos ajustados, utilizando diferentes pruebas de hipótesis y criterios de información para evaluar cuál modelo ofrece un mejor ajuste a los datos.

Prueba de Razón de Verosimilitud

Con esta prueba vamos a comparar el modelo reducido (sin interacción) con el modelo completo (con interacción) para evaluar si la inclusión del término de interacción mejora significativamente el ajuste del modelo.

```
## Analysis of Deviance Table
## Cox model: response is Surv(tiempo, evento)
## Model 1: ~ tratamiento + diabetes + edad
## Model 2: ~ tratamiento * diabetes + edad
##      loglik  Chisq Df Pr(>|Chi|)
## 1 -200.84
## 2 -199.59 2.4828 1 0.1151
```

El resultado de la prueba de razón de verosimilitud indica que no hay una mejora significativa en el ajuste del modelo al incluir el término de interacción entre tratamiento y diabetes (p = 0.115). Esto confirma la decisión obtenida anteriormente, de que la interacción no aporta información adicional relevante al modelo.

Si comparamos el modelo crudo con el modelo ajustado por edad tenemos:

```
anova(mod_cox, mod_red2)
```

```
## Analysis of Deviance Table
## Cox model: response is Surv(tiempo, evento)
## Model 1: ~ tratamiento
## Model 2: ~ tratamiento + edad
##      loglik  Chisq Df Pr(>|Chi|)
## 1 -205.47
## 2 -200.92 9.1067 1 0.002547 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Acá claramente si hay diferencias estadísticamente significativas entre ambos modelos ($p < 0.001$), lo que indica que el modelo ajustado por edad ofrece un mejor ajuste a los datos en comparación con el modelo crudo, lo que refuerza la importancia de incluir la edad como covariable en el análisis para obtener estimaciones más precisas del efecto del tratamiento sobre el riesgo de muerte.

También podemos comparar simultáneamente los tres modelos (crudo, ajustado por edad y diabetes, y el modelo ajustado por edad) utilizando la función `anova()` de la siguiente manera:

```
anova(mod_cox, mod_red, mod_red2)
```

```
## Analysis of Deviance Table
## Cox model: response is Surv(tiempo, evento)
## Model 1: ~ tratamiento
## Model 2: ~ tratamiento + diabetes + edad
## Model 3: ~ tratamiento + edad
##      loglik  Chisq Df Pr(>|Chi|)
## 1 -205.47
## 2 -200.84 9.2777 2 0.009669 **
## 3 -200.92 0.1711 1 0.679175
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Donde vemos que hay diferencias estadísticamente significativas entre los modelos ($p < 0.001$), lo que indica que el modelo ajustado por edad (mod_red2) ofrece un mejor ajuste a los datos en comparación con el modelo crudo y el modelo ajustado por edad y diabetes, lo que refuerza la importancia de incluir la edad como covariable en el análisis para obtener estimaciones más precisas del efecto del tratamiento sobre el riesgo de muerte, mientras que la inclusión de la diabetes no aporta una mejora adicional relevante para el modelo.

Prueba Score (Log-Rank)

Es una prueba alternativa para evaluar la asociación global entre las covariables y el riesgo.

En el caso de la comparación de los dos grupos de tratamiento, recordemos que la prueba Score es equivalente a la prueba Log-Rank, que evalúa si hay diferencias significativas en las curvas de supervivencia entre los grupos.

```
survdif(Surv(tiempo, evento) ~ tratamiento , data = base)
```

```
## Call:
## survdiff(formula = Surv(tiempo, evento) ~ tratamiento, data = base)
##
##           N Observed Expected (O-E)^2/E (O-E)^2/V
## tratamiento=Estándar 57      36      26.5      3.41      7.08
## tratamiento=Nuevo   43      16      25.5      3.54      7.08
##
## Chisq= 7.1 on 1 degrees of freedom, p= 0.008
```

Encontramos que si hay diferencias estadísticamente significativas entre las curvas de supervivencia de los grupos de tratamiento ($p = 0.008$), por lo que el tratamiento nuevo tiene un efecto beneficioso sobre la supervivencia en comparación con el tratamiento estándar.

Al ajustar por edad, la prueba de log-rank ya no es directamente aplicable, por lo que la decisión la debemos tomar a partir de la significancia del tratamiento en el modelo de Cox ajustado por edad, donde concluimos que esta si era relevante para el modelo.

Valores de AIC y BIC

Estas medidas penalizan la complejidad del modelo, favoreciendo modelos que logren un buen ajuste utilizando un menor número de parámetros. Valores más pequeños de AIC o BIC indican un mejor equilibrio entre ajuste del modelo y complejidad.

Valores de AIC

```
AIC(mod_cox, mod_mult, mod_red,mod_red2)
##           df           AIC
## mod_cox    1 412.9487
## mod_mult    4 407.1882
## mod_red     3 407.6710
## mod_red2    2 405.8421
```

Valores de BIC

```
BIC(mod_cox, mod_mult, mod_red,mod_red2)
```

```
##           df           BIC
## mod_cox    1 414.9000
## mod_mult    4 414.9932
## mod_red     3 413.5247
## mod_red2    2 409.7445
```

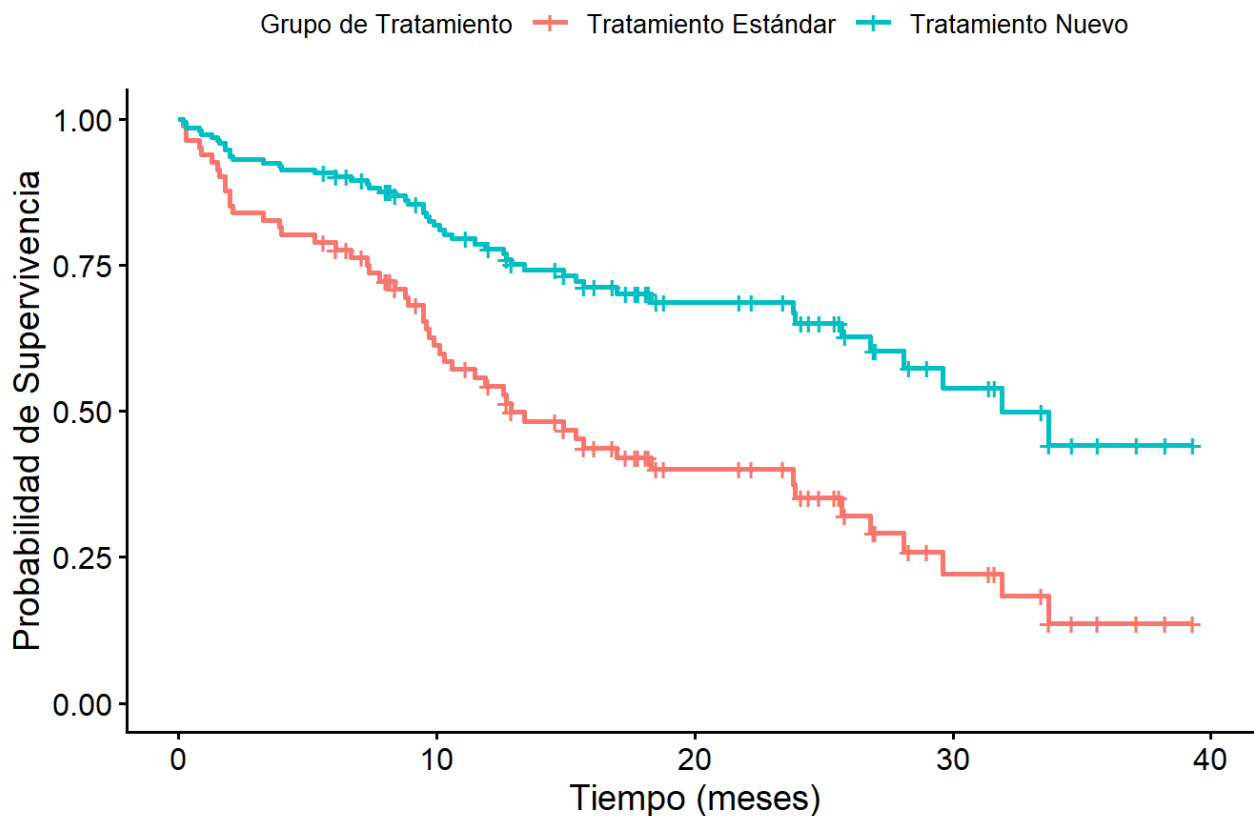
Podemos ver que el mejor modelo bajo las dos medidas es el modelo ajustado por edad (mod_red2), confirmando ue la inclusión de la edad mejora significativamente el ajuste del modelo en comparación con el modelo crudo y losque modelos que incluyen la diabetes o la interacción entre tratamiento y diabetes, no aportan una mejora adicional relevante para el modelo.

Estimación y comparación de curvas de supervivencia ajustadas

Estas curvas las vamos a estimar a partir del modelo de Cox ajustado por edad, lo que nos permitirá visualizar la supervivencia esperada para cada grupo de tratamiento, manteniendo constante la edad.

```
ggsurvplot(  
  survfit(mod_red2, newdata = data.frame(  
    tratamiento = c("Estándar", "Nuevo"),  
    edad = mean(base$edad, na.rm = TRUE)  
  )),  
  data = base,  
  conf.int = FALSE,  
  legend.labs = c("Tratamiento Estándar", "Tratamiento Nuevo"),  
  legend.title = "Grupo de Tratamiento",  
  xlab = "Tiempo (meses)",  
  ylab = "Probabilidad de Supervivencia",  
  title = "Gráfica 10. Curvas de supervivencia ajustadas por edad"  
)
```

Gráfica 10. Curvas de supervivencia ajustadas por edad



En este gráfico podemos observar que, como ajustando por edad, el grupo que recibió el tratamiento nuevo mantiene una mayor probabilidad de supervivencia a lo largo del tiempo en comparación con el grupo estándar. Esto refuerza la conclusión de que el tratamiento nuevo tiene un efecto beneficioso sobre la supervivencia, independientemente de la edad de los pacientes.

Específicamente, se observa una separación entre las curvas, donde el grupo estándar presenta una disminución más rápida de la supervivencia, lo que indica una mayor ocurrencia del evento. Por ejemplo, hacia los 20 meses, la supervivencia en el grupo estándar es cercana a 0.40, mientras que en el grupo nuevo se sitúa alrededor de 0.70, diferencia que se mantiene en periodos posteriores.

Lecturas complementarias.

- **Survival analysis: A self-learning text.** David G. Kleinbaum and Mitchel Klein. Springer, New York, 2012.
- **Función de riesgo instantáneo:** Cox, D. R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological), 34(2), 187-202.
- **Función de riesgo acumulado:** Nelson, W. (1972). Theory and applications of hazard plotting for censored failure data. Technometrics, 14(4), 945-966.
- **R:** Therneau, T. M. (2026). A package for survival analysis in R. R package version 3.8-6, <https://CRAN.R-project.org/package=survival>.
- klein, J. P., & Moeschberger, M. L. (2003). Survival analysis: techniques for censored and truncated data. Springer Science & Business Media.