

Universidad Nacional Autónoma de
México

Facultad de Ciencias

Modelos de Supervivencia
y
de Series de Tiempo

Proyecto: Análisis de Supervivencia

Integrantes:

- Galindo Torres Bernardo Antonio
- Martínez Guerrero Jazmín Alejandra
- Rivera Hernández León Diego
- Sánchez Morales Sandra Itzel

Dra. Lizbeth Naranjo Albarrán

Act. Yadira Rivas Godoy

Grupo: 9249

Semestre: 2020-2

Índice

1. Introducción	1
2. Análisis descriptivo	2
2.1. Descripción de las variables	2
2.2. Distinción por edades	3
2.3. Estado marital de los pacientes	4
2.4. Raza de las pacientes	5
2.5. Uso de condón e infección inicial	5
2.6. Número de parejas sexuales	6
2.7. Años de escolaridad.	6
3. Análisis y ajuste del modelo de Kaplan - Meier	8
3.1. Ajuste del modelo de Kaplan - Meier	8
3.2. Identificación de variables que afectan al tiempo de supervivencia	8
3.3. Contraste de supervivencia en poblaciones que tuvieron sexo oral en los últimos 30 días contra las que no tuvieron	10
3.4. Otras características	11
4. Análisis y ajuste del modelo Cox-PH	13
4.1. Ajuste de modelo Cox-PH	13
4.2. Verificación de riesgos proporcionales y estimación de $\underline{\beta}$	13
4.3. Interpretación de los coeficientes	14
4.4. Análisis de residuales	15
4.4.1. Residuos Cox - Snell	15
4.4.2. Residuos de Devianza	16
4.4.3. Residuos DfBeta	17
5. Conclusión y Referencias	19
6. Anexo	20
6.1. Gráficas de estimadores Kaplan-Meier	20
6.1.1. Comparación entre poblaciones.	26
6.2. Gráficas del test de Schoenfeld	28

1. Introducción

Las Enfermedades de Transmisión Sexual (ETS) son una preocupación de la salud pública a nivel mundial, debido a que éstas tienden a propagarse rápidamente. Es importante estudiar las ETS ya que algunas han sido vinculadas a un mayor riesgo de adquirir ciertos tipos de cánceres e infecciones. Una de las principales desventajas es que muchos portadores no presentan síntomas, sin embargo, siguen siendo un foco de infección.

Las ETS afectan tanto a los hombres como a las mujeres, y casi la mitad de todas las ETS ocurren en personas de menores a 25 años, esto podría deberse a la falta de información o falta de responsabilidad en los individuos. La exposición a una ETS puede ocurrir en cualquier momento que tiene contacto sexual con alguien, que involucra los genitales, la boca (oral) o el recto (anal). La exposición es más probable si tiene más de una pareja sexual o no se tiene el hábito del uso de condón.

Aunque la enfermedad de transmisión sexual más conocida es el SIDA, otras enfermedades sexuales aún tienen un impacto significativo en la mortalidad de la población, la base de datos se centra en las siguientes dos: gonorrea y clamidia.

Éstas enfermedades son de especial interés debido a que son asintomáticas cuando se presentan en mujeres y si no se tratan pueden llevar a complicaciones, incluyendo la esterilidad. Ambas enfermedades pueden ser fácilmente prevenidas y efectivamente tratadas. Por lo tanto, es un misterio por qué la incidencia de estas enfermedades siguen siendo alta en varias subpoblaciones.

El propósito del estudio del cual se tomó la base de datos es identificar esos factores que están relacionados con el tiempo de reinfección de gonorrea o clamidia, teniendo como infección inicial gonorrea o clamidia. Si los factores relacionados a un mayor riesgo de reinfección pueden ser identificados, entonces, se podrían hacer intervenciones en aquellos individuos que tengan un mayor riesgo de reinfección y así se podría reducir el número de contagiados.

2. Análisis descriptivo

2.1. Descripción de las variables

La base consta de 877 observaciones (pacientes mujeres) y 24 variables, se describen a continuación:

- **Obs:** Número de observación.
- **rinfct** Reinfección (1: sí, 0: no).
- **time** Tiempo de la reinfección.

Variables demograficas:

- **Race:** Raza (W: blanca, B: negra).
- **Marital status:** Estado marital (D: divorciada/separada, M: casada, S: soltera).
- **Age:** Edad de la paciente cuando tuvo la infección inicial.
- **yschool:** Años de escolaridad.

Factores de comportamiento registrados en el examen, cuando se diagnostico la infección inicial:

- **npartner:** Número de parejas sexuales cuando la infección sexual fue diagnosticada, en los últimos 30 días.
- **os12m:** Sexo oral dentro de los 12 meses (1: sí, 0: no).
- **os30d:** Sexo oral dentro de los últimos 30 días(1: sí, 0: no).
- **rs12m:** Sexo anal en los 12 meses pasados (1: sí, 0: no).
- **rs30d:** Sexo anal dentro de los últimos 30 días (1: sí, 0: no).
- **condom:** Uso de condón (1:siempre, 2:a veces, 3:nunca).

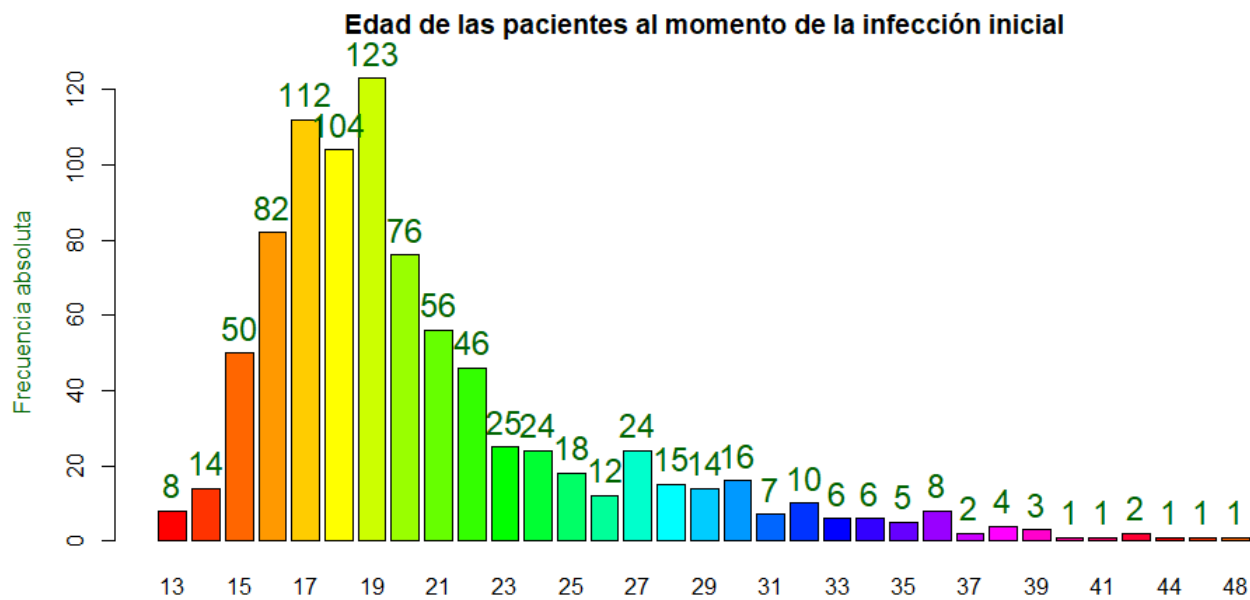
Síntomas notados al tiempo de la infección inicial:

- **iinfct:** Infección inicial (1: gonorrea, 2: clamidia, 3: ambos).
- **abdpain:** Presencia de dolor abdominal (1: sí, 0: no).
- **discharge:** Signo de secreción (1: sí, 0: no).

- **dysuria**: Signo de disuria (1:sí, 0:no).
- **itch**: Signo de picazón (1: sí, 0: no).
- **lesion**: Signo de lesión (1: sí, 0: no).
- **rash**: Signo de erupción (1: sí, 0: no).
- **lymph** Signo de linfa (1: sí, 0: no).
- **vagina** Examen vaginal (1: sí, 0: no).
- **dchexam** Secreción en el examen (1: sí, 0: no).
- **abnode** Ganglio anormal en el examen (1: sí, 0: no).

2.2. Distinción por edades

Si bien la edad promedio para el inicio de la vida sexual de las personas varía dependiendo de factores que van desde la región en que habitan hasta la religión que se profesa, se tiene que en la base de datos a mujeres desde los 13 años y hasta los 48 años:

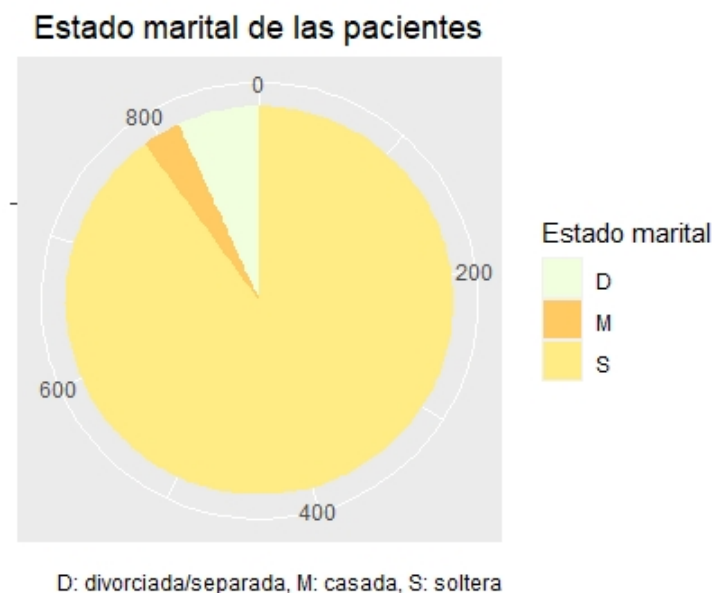


En promedio, la edad sexual activa es de 20 años. Notemos además que el conjunto de personas que tienen entre 15 y 21 años son una población importante para nuestro estudio, de tal manera que se podría pensar en la falta de responsabilidad ante las enfermedades sexuales en este grupo de pacientes.

Además se encontró que el **58 % de las pacientes menores a 16 años presentan reinfección**, mientras que sólo el **38 % de las pacientes mayores o iguales a 16 años presentan reinfección**. Esto señala que las pacientes menores son más propensas a presentar reinfección.

2.3. Estado marital de los pacientes

Habitualmente se concibe el matrimonio como una especie de fin del libertinaje que tienen las personas al estar solteras, es por ello que es relevante analizar la situación sentimental de los pacientes para ver de que manera influye esto en el estudio:



D: 60 pacientes, M: 28 pacientes, S: 789 pacientes

Es posible observar que el número de pacientes solteras es considerablemente mayor a las pacientes casadas y separadas, esto hace pensar en un mayor libertinaje en el celibato que podría desencadenarse en enfermedades de transmisión sexual.

En la siguiente tabla se muestra el porcentaje que tuvo o no reinfección de acuerdo al estado marital:

	No Reinfección	Reinfeccion
Soltera	59.19 %	40.81 %
Casada	67.86 %	32.14 %
Divorciada	73.33 %	26.67 %

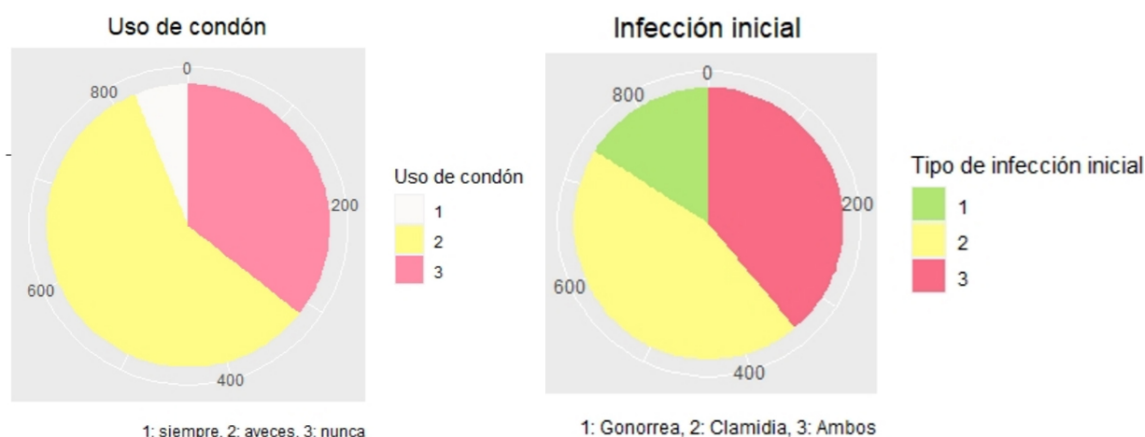
Cuadro 1: Porcentajes de no reinfecciones y reinfecciones en cada población dependiendo el estado marital de la paciente.

Se puede notar que las reinfecciones son menos propensas en las pacientes divorciadas y casadas que en las pacientes solteras.

2.4. Raza de las pacientes

Si bien esto no nos da información concluyente, se podría pensar en que las mujeres de raza negra son más propensas a las ETS, ya que al menos para el estudio éstas son aproximadamente dos terceras partes del total de los pacientes. Con un total de **585 pacientes negras** y **292 blancas**.

2.5. Uso de condón e infección inicial



En la gráfica de la izquierda se puede notar que el un mayor número de pacientes son las que hacen uso ocasional del condón, mientras que la minoría usan siempre condón.

Por otro lado, en la gráfica de la derecha muestra que **la mayoría de pacientes tuvo clamidia como infección inicial**. Cabe destacar que el **52.14 %** de pacientes con gono-

rea presentaron **reinfección**, mientras que solo el **34.09 %** de pacientes diagnosticadas con **clamidia** presentaron **reinfección**.

2.6. Número de parejas sexuales

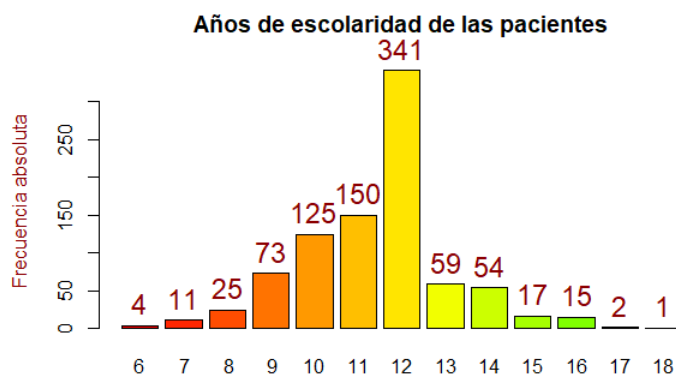
Número parejas sexuales	Cantidad de pacientes
0	70
1	607
2	146
3	39
4	5
5 o más	10

Cuadro 2: CANTIDAD DE PACIENTES CLASIFICADAS POR EL NÚMERO DE PAREJAS SEXUALES REGISTRADAS AL MOMENTO DE LA INFECCIÓN INICIAL.

El número de parejas sexuales cuando la infección sexual fue diagnosticada, en los últimos 30 días va desde 0 hasta 19, con un promedio de 1.27 parejas. Obsérvese que el número de pacientes con sólo una pareja es considerablemente mayor a las demás, resultando un 69 % de la población total del estudio.

2.7. Años de escolaridad.

Los años de escolaridad de las pacientes van desde 6 hasta 18 años, con un promedio de 12 años. Se puede ver que la mayoría de las pacientes cuenta con entre 10 y 12 años de escolaridad.



Enfermedades Sexuales

5 de junio de 2020

	No Reinfección	Reinfeccion
< 8 años	33.33 %	66.67 %
7 < años < 16	60.42 %	39.57 %
> 15 años	83.33 %	16.67 %

Cuadro 3: Porcentajes de no reinfecciones y reinfecciones en cada población dependiendo la cantidad de años de estudios de las pacientes.

Se puede notar que de las pacientes con menos de 8 años de escolaridad, 2/3 de ellas tuvieron reinfección, mientras que de las pacientes con más de 15 años de escolaridad, sólo el 16.67 % tuvo reinfección.

3. Análisis y ajuste del modelo de Kaplan - Meier

3.1. Ajuste del modelo de Kaplan - Meier

A continuación, se presenta la gráfica de la estimación de Kaplan - Meier hecha al modelo y sus respectivas bandas con un nivel de confianza de 95 %. El evento de interés en este caso es la presencia de reinfección en las pacientes registradas.

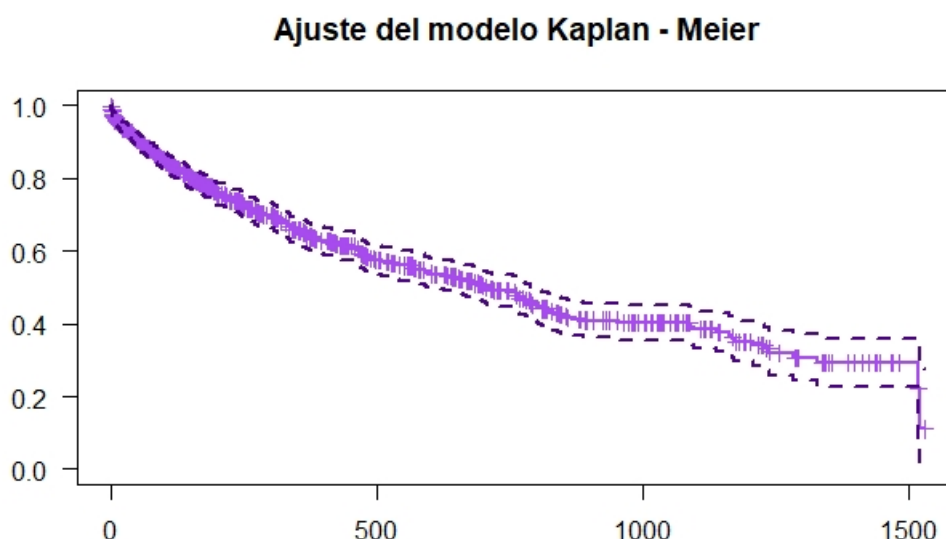


Figura 1: Estimación de de Kaplan - Meier de la función de supervivencia para la base de datos `std`.

3.2. Identificación de variables que afectan al tiempo de supervivencia

Para identificar qué variables afectan al tiempo de supervivencia, la población en la base de datos fue dividida tomando diversos factores significantes, tales como el uso del condón, la práctica de sexo oral en los últimos 30 días, la práctica de sexo rectal en los últimos 30 días, el estado marital de cada paciente y el tipo de infección inicial, entre otros.

Esto con el fin de comparar las diversas funciones de supervivencia, es decir, contrastar la hipótesis:

$$H_0 : S_A(t) = S_B(t) \quad \forall t > 0 \quad vs \quad H_0 : S_A \neq S_B(t) \text{ p.a. } t > 0$$

Para cada factor se compararon las funciones de supervivencia de cada grupo, usando la prueba de *Log-Rank* de *Mantel-Haenszel* y usando la prueba de *Log-Rank* de *Peto-Peto*, las cuales prueban la hipótesis de que las funciones de supervivencia son iguales para cada categoría de la variable. En el caso de las variables con tres categorías se realizó un análisis dos a dos para cada nivel.¹

Por lo tanto, podemos concluir que las variables siguientes tienen una distribución distinta entre sus respectivas categorías.

- race
- rinfct
- os12m
- os30d
- abdpain
- vagina
- dchexam
- condom (a veces *vs.* nunca)
- marital (soltera *vs.* divorciada)
- iinfct (gonorrea *vs.* clamidia)
- iinfct (gonorrea *vs.* ambas)

¹Los resultados de las pruebas se muestran en el Anexo, cuadro 8.

3.3. Contraste de supervivencia en poblaciones que tuvieron sexo oral en los últimos 30 días contra las que no tuvieron

Se realizó una partición de la base teniendo en cuenta las categorías de la variable `os30d`, las cuales identifican si se practicó el sexo oral o no en los últimos 30 días.

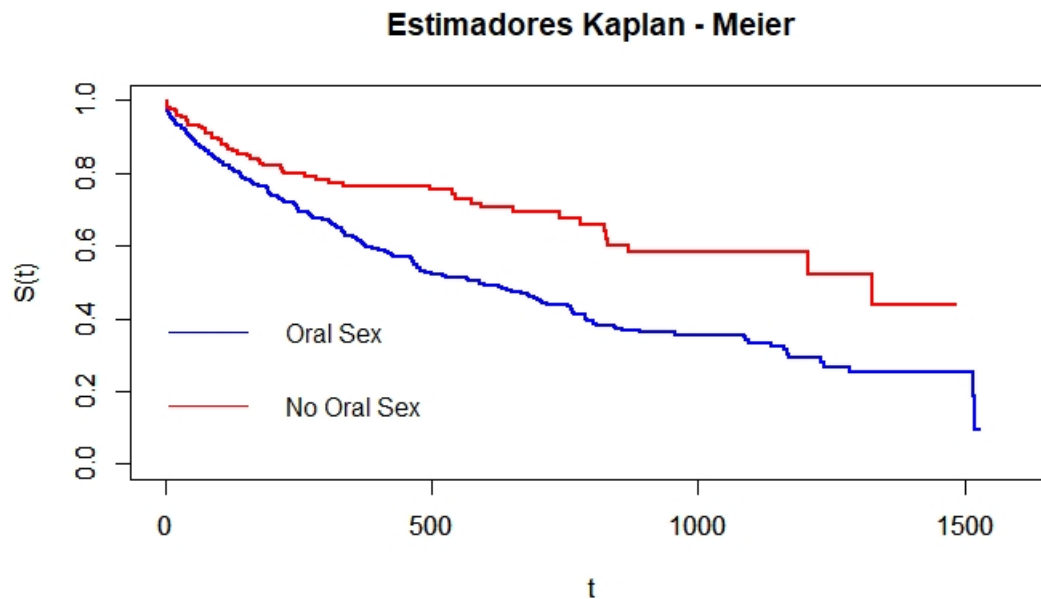


Figura 2: Se comparó las distintas distribuciones para contrastar la supervivencia en cada caso. Teniendo como resultados la afectación que tiene practicar el sexo oral para presentar una reinfección.

3.4. Otras características

A continuación se presentan dos gráficas, las cuales contienen la función de supervivencia de la estimación ajustada, sin embargo, la diferencia que radica en ellas es el número de variables tomadas en cuenta. En la primera se toman en cuenta todas las variables, y en la segunda la variable `os30d`. Además, en ambas gráficas se señala una aproximación al valor de la mediana, señalando para cada caso, el tiempo en el que se realizó la observación.

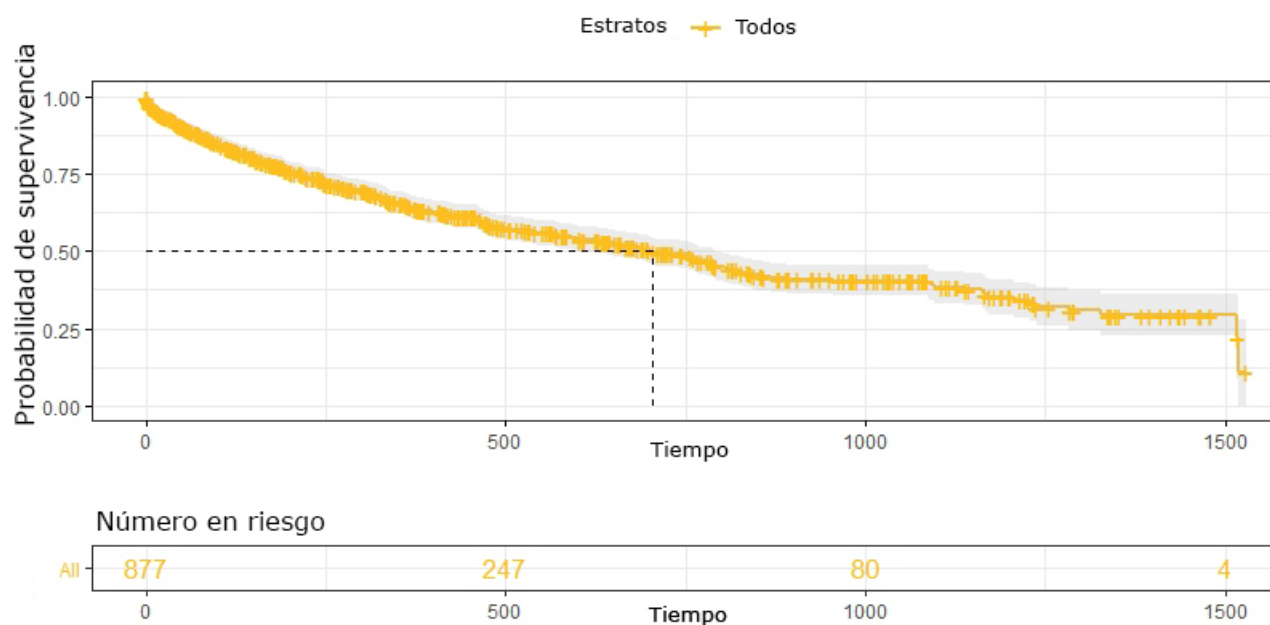


Figura 3: Es posible observar una mayor concentración de sujetos en riesgo para los primeros tiempos de estudio, y una reducción considerable para los últimos.

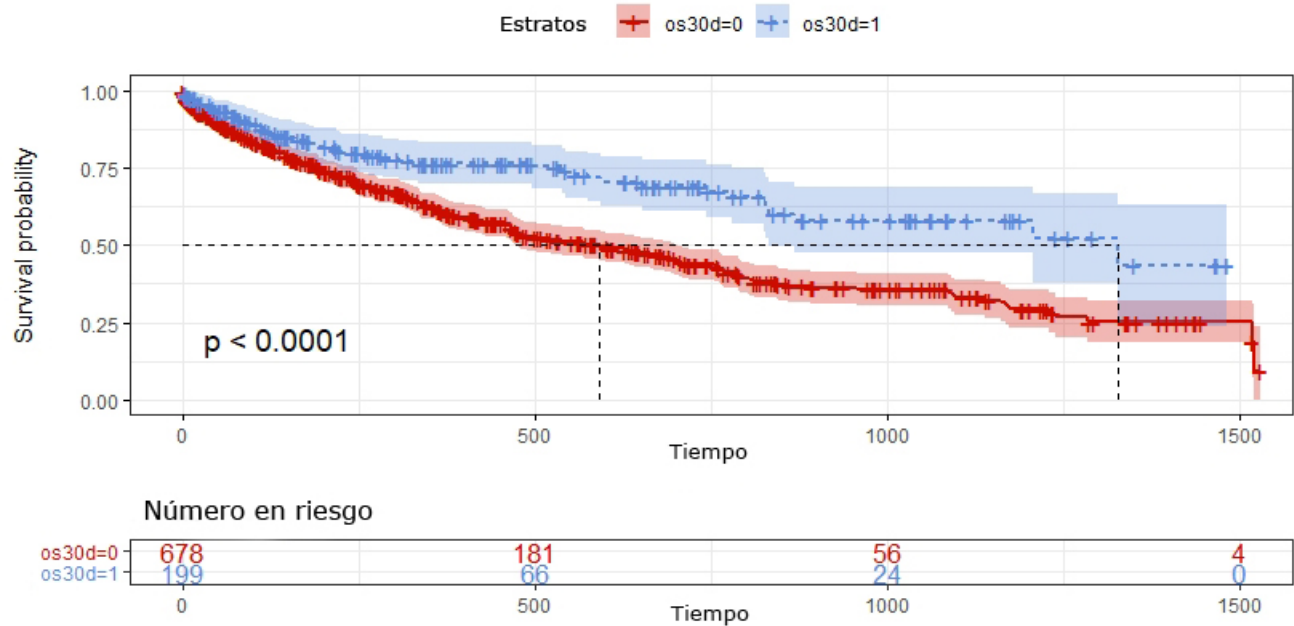


Figura 4: En esta gráfica se puede apreciar que las funciones de supervivencia no intersectan y por ende, sí hay una afectación en el tiempo de supervivencia dependiendo de si se practicó o no sexo oral en los últimos 30 días, además de afirmar esta decisión con ayuda del *p-value*, pues éste rechaza la hipótesis de que ambas distribuciones sean iguales. También se puede ver que la mediana para ambos casos está acorde con cada una de las distribuciones.

4. Análisis y ajuste del modelo Cox-PH

4.1. Ajuste de modelo Cox-PH

Se ajustó un modelo Cox-PH a los datos, como hay 21 variables ² se realizó una selección de variables *stepwise*³. Dicha selección dio como resultado al modelo con las variables `yschool`, `iinfct2`⁴, `os30d`, `abdpain`, `vagina` y `dchexam` como las únicas variables significativas (véase el cuadro 4). Obsérvese además que, en todas las variables dicotómicas seleccionadas, se observó una diferencia de las funciones de supervivencia entre sus respectivas categorías en la sección 3.2.

	z	Pr(> z)	VIF
vagina	2.311	0.020830	1.025655
yschool	-4.203	2.63e-05	1.017647
dchexam	-1.900	0.057446	1.013283
iinfct2	-1.709	0.087448	1.010733
abdpain	1.872	0.061232	1.020770
os30d	-3.782	0.000155	1.019209
AIC del modelo= 4103.907 (edf = 6)			

Cuadro 4: Significancia de las variables seleccionadas para el modelo, además del valor VIF para cada una. Se observa que todas las variables tienen un VIF bajo. El valor 'z' es la estadística de la prueba de Wald para $\beta's = 0$.

4.2. Verificación de riesgos proporcionales y estimación de $\underline{\beta}$

Ya que se determinaron las variables significativas con el modelo de Cox-PH, se verificó el supuesto de riesgos proporcionales con la prueba de Schoenfeld, ésta prueba nos proporciona información tanto de forma individual como globalmente para las variables del modelo, los resultados obtenidos se muestran en el cuadro 5.

²Sin contar el tiempo y el estatus de los sujetos del estudio, pues forman la variable de tipo 'survival'.

³La selección se realizó considerando un Nivel de Entrada y Nivel de Permanencia iguales a 0.15 (*p-values* de la prueba de Wald para $\beta_i = 0$) y un umbral de VIF igual a 10 para las variables continuas y 2.5 para variables categóricas.

⁴La variable `iinfct2` es de tipo *dummy*; indica si un individuo tiene valor '2' o no en la variable original `iinfct`, cabe destacar que '1' indica una infección inicial de gonorrea y '3' indica una infección inicial de gonorrea y clamidia, `iinfct2` representa los individuos que **solo** se infectaron inicialmente de clamidia.

	chisq	df	p
vagina	0.881	1	0.348
yschool	0.019	1	0.890
dchexam	0.347	1	0.556
iinfct2	3.725	1	0.054
abdpain	0.267	1	0.605
os30d	0.882	1	0.348
GLOBAL	6.221	6	0.339

Cuadro 5: Resultados de la prueba de Schoenfeld para verificar riesgos proporcionales; tanto para las variables individuales como para el modelo global.

Dado que todos los $p - value > 0.05$, se falla en rechazar la hipótesis nula de riesgos proporcionales (en la sección 6.2 se pueden encontrar las gráficas de dicha prueba); más aún, no se desecha ninguna de las variables seleccionadas como significativas.

Como no es posible desechar ninguna de las variables significativas, se procede a hacer los cálculos de las respectivas β_i puntualmente y por intervalos de confianza (al 95 %). Los resultados de tales computaciones se muestran en el cuadro 6.

Variable i	β_i			$\exp(\beta_i)$		
	Puntual	lower .95	upper .95	Puntual	lower .95	upper .95
vagina	0.39034	0.05930	0.72137	1.47748	1.0611	2.0573
yschool	-0.14318	-0.20994	-0.07641	0.86660	0.8106	0.9264
dchexam	-0.42217	-0.85769	0.01335	0.65562	0.4241	1.0134
iinfct2	-0.18951	-0.40684	0.02783	0.82737	0.6657	1.0282
abdpain	0.27676	-0.01303	0.56656	1.31885	0.9871	1.7622
os30d	-0.57523	-0.87330	-0.27716	0.56258	0.4176	0.7579

Cuadro 6: Estimaciones para los coeficientes β_i .

Por lo tanto, la ecuación ajustada al modelo Cox-PH esta dada por (1).

$$\lambda(t|\mathbf{Z}) = \lambda_0(t) * e^{0,39x_{\text{vagina}} - 0,14x_{\text{yschool}} - 0,42x_{\text{dchexam}} - 0,19x_{\text{iinfct2}} + 0,28x_{\text{abdpain}} - 0,58x_{\text{os30d}}} \quad (1)$$

4.3. Interpretación de los coeficientes

Los términos e^{β_i} presentes en la tabla 6 se interpretan para las variables dicotómicas como: A cualquier momento dado, alguien que da positivo a dicha variable ('1') es e^{β_i} veces más propenso a presentar la falla. En el caso de las variables cuyos valores $e^{\beta_i} < 1$ el riesgo, entonces, disminuye; por ejemplo, para la variable **abdpain**, alguien que presenta dolor abdominal ($x_{\text{abdpain}} = '1'$) es apróx. 1.32 (o sea, un 32 % **más**) veces propenso de presentar una

reinfección en comparación con alguien que no presenta dolor abdominal ($x_{\text{abdpain}} = '0'$). En cambio, una persona cuya infección inicial fue solo clamidia ($x_{\text{iinfct2}} = '1'$) es 0.83 (en otras palabras, un 17 % **menos**) veces propensa de presentar una reinfección en comparación con alguien cuya infección no fue (solo) clamidia ($x_{\text{iinfct2}} = '0'$). Para la variable **vagina** el riesgo aumenta en un 48 %; para **dchexam** el riesgo disminuye en un 34 %; para **abdpain** el riesgo aumenta en un 32 % y para la variable **os30d** el riesgo disminuye en un 44 %.

La variable **yschool** se tomó como continua, pero ocurre algo similar: Por cada unidad que aumenta esta variable, un individuo aumenta en un 0.87 su riesgo de presentar la falla; i.e. dicho riesgo disminuye en un $(1 - 0.87) * 100 = 13$ % por año escolar (en comparación con el riesgo que se tiene con un año escolar menos).

4.4. Análisis de residuales

Los residuos en supervivencia son un poco distintos a los de otro tipo de modelos por la existencia de la censura. Es por ello que una vez que se ha estimado el modelo es necesario revisar qué tan bien se ajusta a nuestros datos.

4.4.1. Residuos Cox - Snell

Al aplicar la prueba de Residuos de Cox-Snell se tiene que los valores no se ajustan alrededor de los intervalos de confianza, esto principalmente se debe a que la mayoría de las variables son binarias, por lo cual se van a los extremos, así que al igual que con los Residuos de Martingalas no son los mejores postores para el tipo de variables que se trabajan en el estudio.

Algo que se puede rescatar de esta prueba es el gráfico obtenido para la variable **yschool** ya que esta al ser una variable continua se ve la siguiente manera:

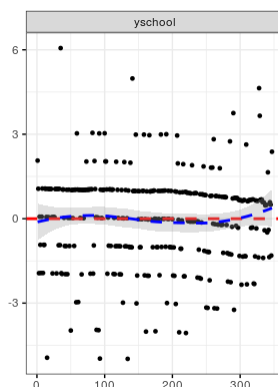


Figura 5: Residuos Cox-Snell de la variable yschool

Notemos que aunque sea una variable continua sigue sin agruparse alrededor de las bandas

de confianza, sin embargo, muestra algunos datos atípicos en la variable que se alejan por mucho del centro.

4.4.2. Residuos de Devianza

Recordemos que la devianza de un modelo de regresión es el estadístico que se utiliza para cuantificar hasta que punto el modelo actual que hemos estimado se aleja de un modelo teórico que se ajustase perfectamente a nuestros datos. Estos residuos tienen características similares a los residuos de regresión lineal, es decir, tienen media cero y varianza uno. Son negativos para observaciones con tiempos de supervivencia menores que el esperado.

De manera gráfica obtenemos lo siguiente:

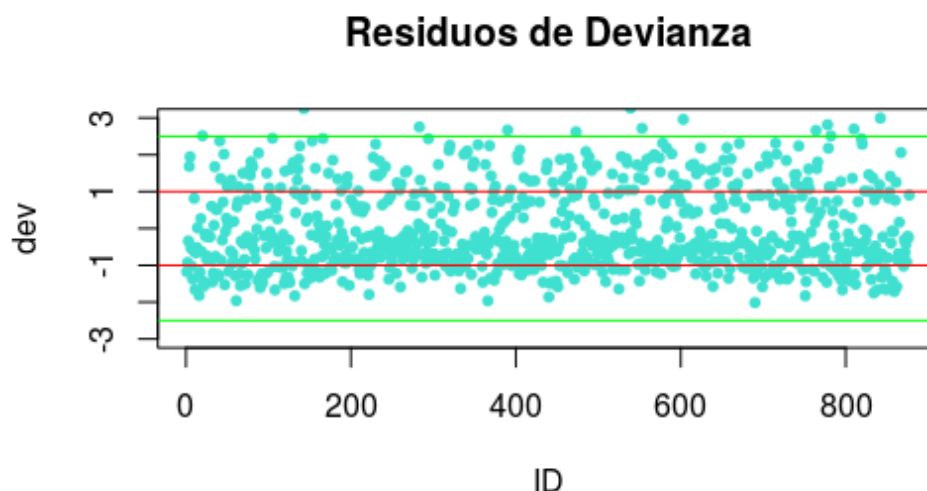


Figura 6: Residuos de Devianza

Notemos que tenemos muchas observaciones fuera del intervalo $(-1,1)$, esto principalmente puede verse afectado por la variación entre las variables, ya que de manera numérica obtenemos los resultados del cuadro 7 y vemos que justo son datos muy extremos entre sí.

Mín	Mediana	Máx.
-2.013	-0.394	3.267

Cuadro 7: Residuos de Devianzas

4.4.3. Residuos DfBeta

La especificación del argumento **type = dfbeta** produce una matriz de cambios estimados en los coeficientes de regresión al eliminar cada observación a su vez; asimismo, **type = dfbetas** produce los cambios estimados en los coeficientes divididos por sus errores estándar. Este análisis se hace para cada una de las covariables electas en el modelo.

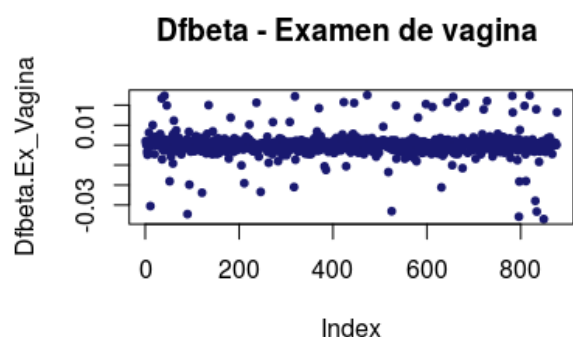


Figura 7: Dfbeta para la variable 'vagina'.

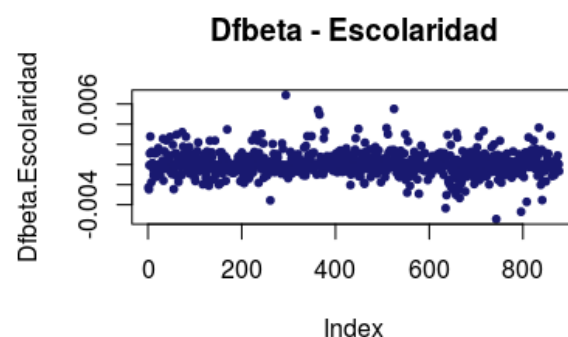


Figura 8: Dfbeta para la variable 'yschool'.

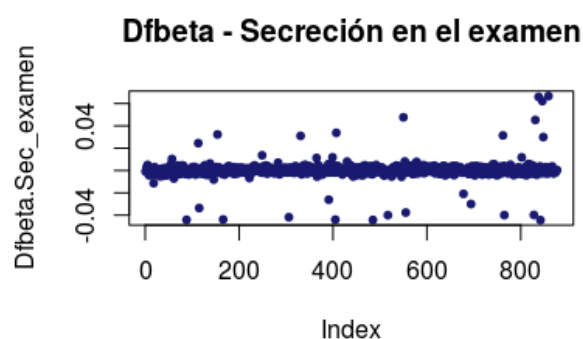


Figura 9: Dfbeta para la variable 'dchexam'.

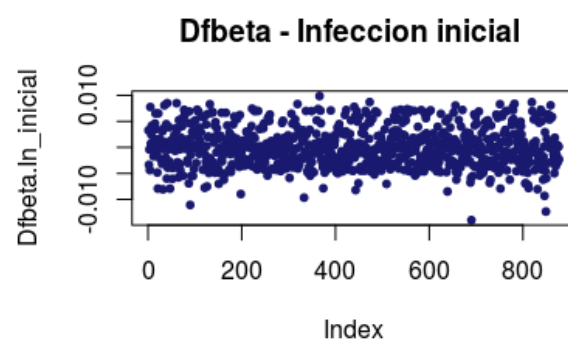


Figura 10: Dfbeta para la variable 'iinfct2'.

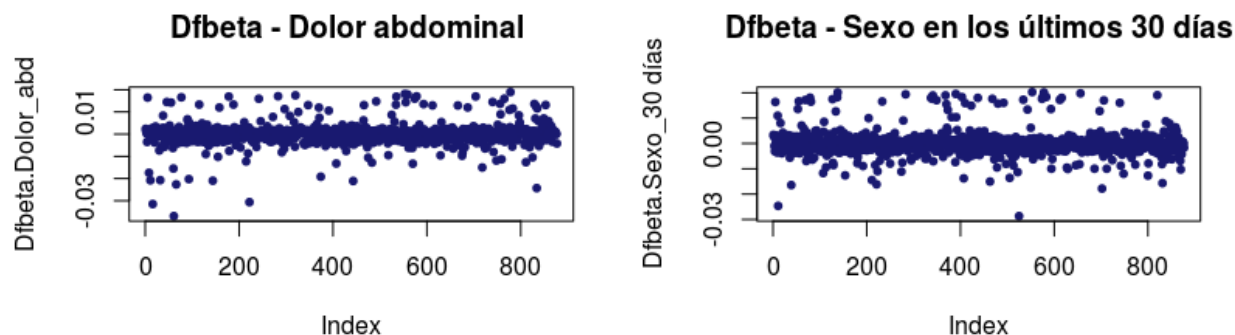


Figura 11: Dfbeta para la variable 'abdpain'. Figura 12: Dfbeta para la variable 'os30d'.

En particular podríamos decir que para esta prueba de residuales los datos parecen estar mas agrupados, sin embargo, aún hay variables como *“infct2”* que se muestra un poco dispersa en los datos, esto como ya lo habíamos mencionado antes puede deberse principalmente al tipo de variable y que la mayoría de las pacientes incluidas en el estudio tenían clamidia como infección inicial.

5. Conclusión y Referencias

Si bien hay muchos factores que contribuyen a la reaparición de una enfermedad, hay algunos que propician o aumentan el riesgo de que exista alguna reinfección.

La posibilidad de reinfección por clamidia o gonorrea en las pacientes si está relacionada con factores tales como: el grado de estudios, el chequeo periódico, las preexistencias de alguna enfermedad, entre otras.

En conclusión, sí importa la raza, que se haya practicada sexo oral en los últimos 12 meses ó 30 días, la presencia de dolor abdominal, la presencia de exámenes vaginales, las secreciones en el examen, el uso de condón ocasional, estar soltera, o presentar gonorrea inicialmente, y los años de escolaridad para que exista la posibilidad de presentar una reinfección.

- *Klein and Moeschberger (1997) Survival Analysis Techniques for Censored and truncated data, Springer.*
- *Enfermedades de transmisión sexual. (s. f.).Recuperado 25 de mayo de 2020, de <https://medlineplus.gov/spanish/sexuallytransmitteddiseases.html>.*

6. Anexo

6.1. Gráficas de estimadores Kaplan-Meier

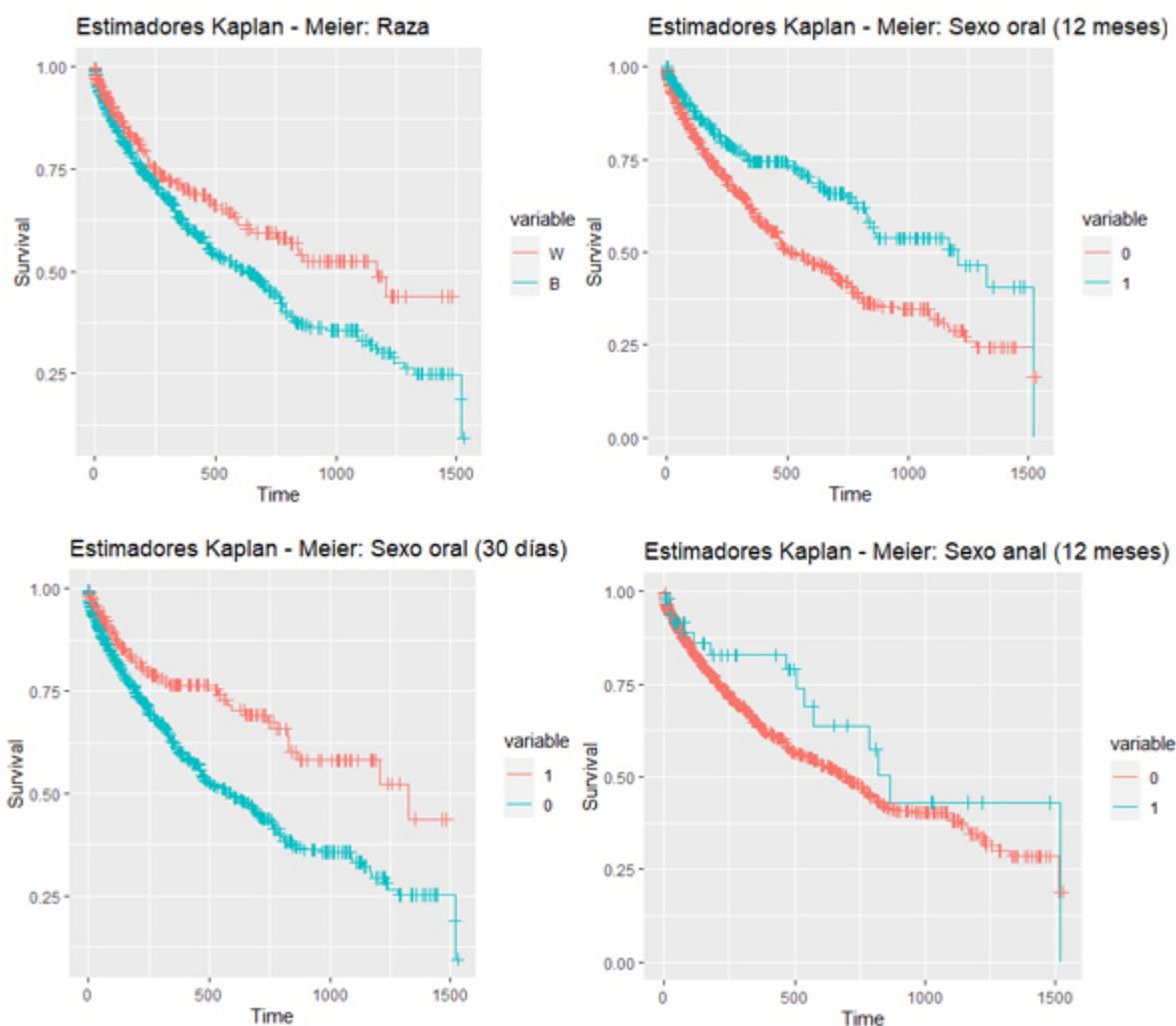


Figura 13: Gráficas que muestran las estimaciones de las funciones de supervivencia para las variables: raza, sexo oral (12 meses y 30 días) y sexo anal (12 meses). Las tres primeras gráficas parecen mostrar diferencias entre cada grupo.

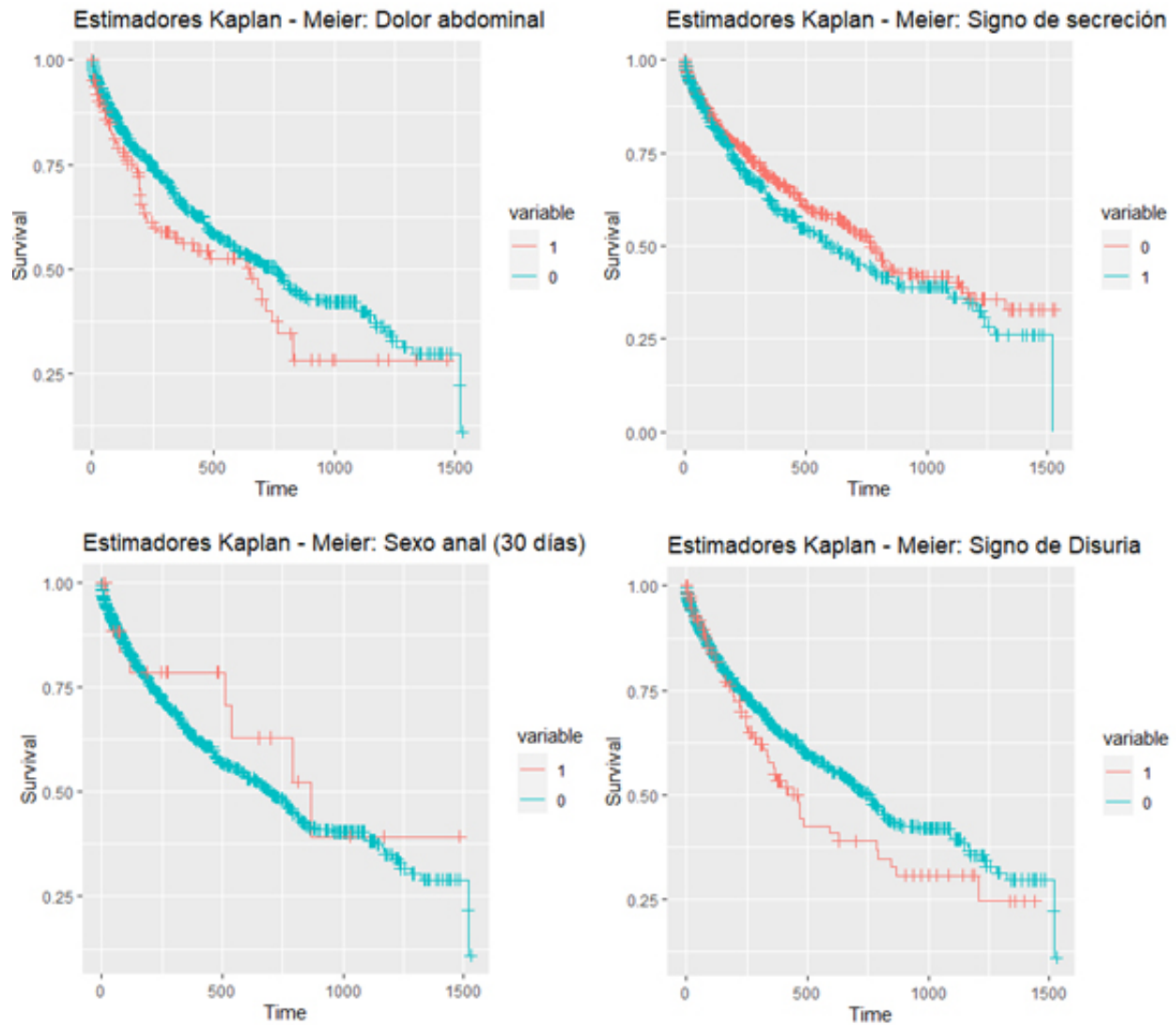


Figura 14: Gráficas que muestran las estimaciones de las funciones de supervivencia para las variables: Dolor abdominal, signo de secreción, sexo anal (30 días) y signo de disuria. Ninguna de las variables parece tener diferencias entre sus grupos.

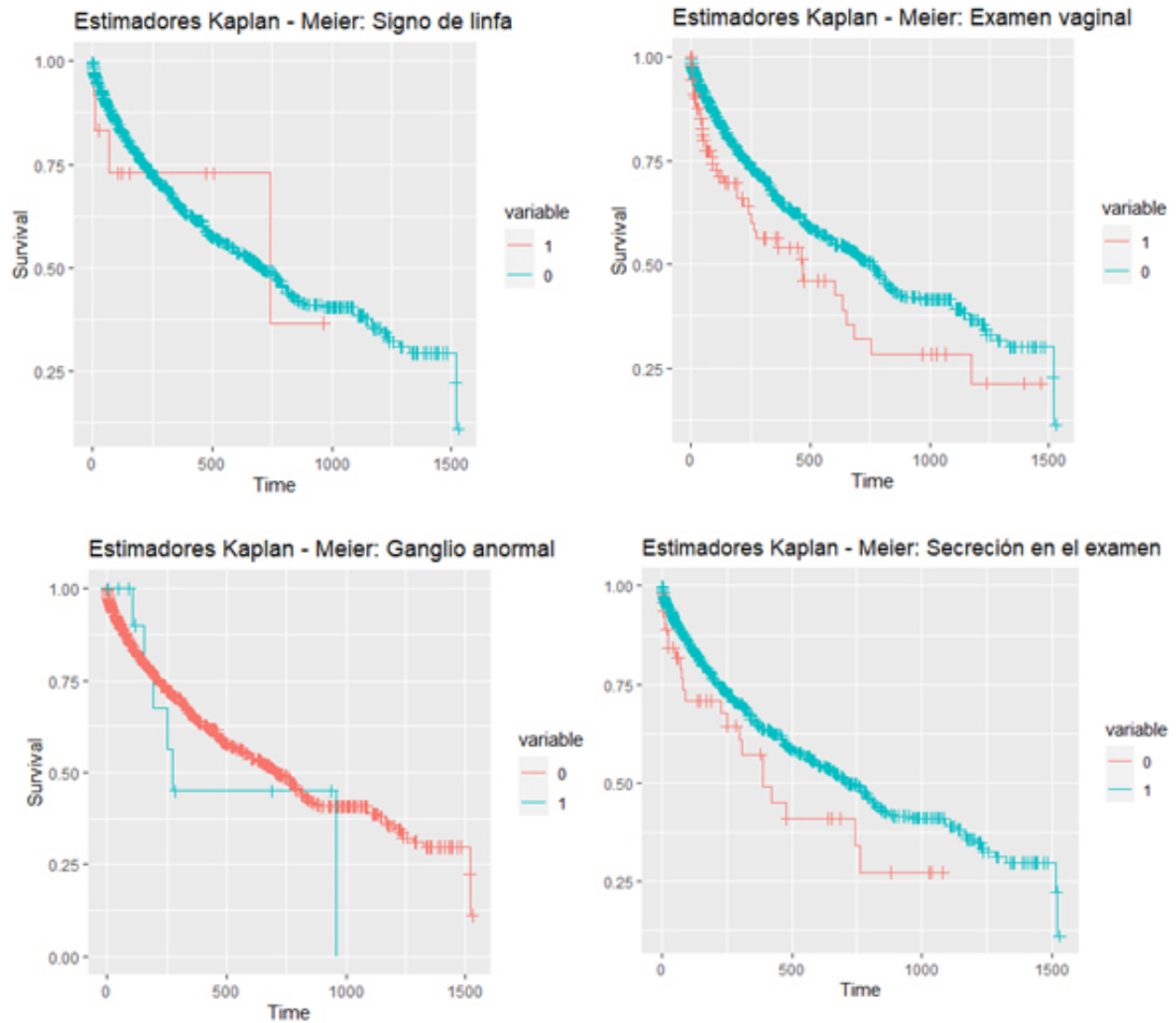


Figura 15: Gráficas que muestran las estimaciones de las funciones de supervivencia para las variables: Signo de linfa, examen vaginal, ganglio anormal y secreción en el examen. La primer y última variable parecen tener diferencias entre las supervivencia de sus poblaciones.

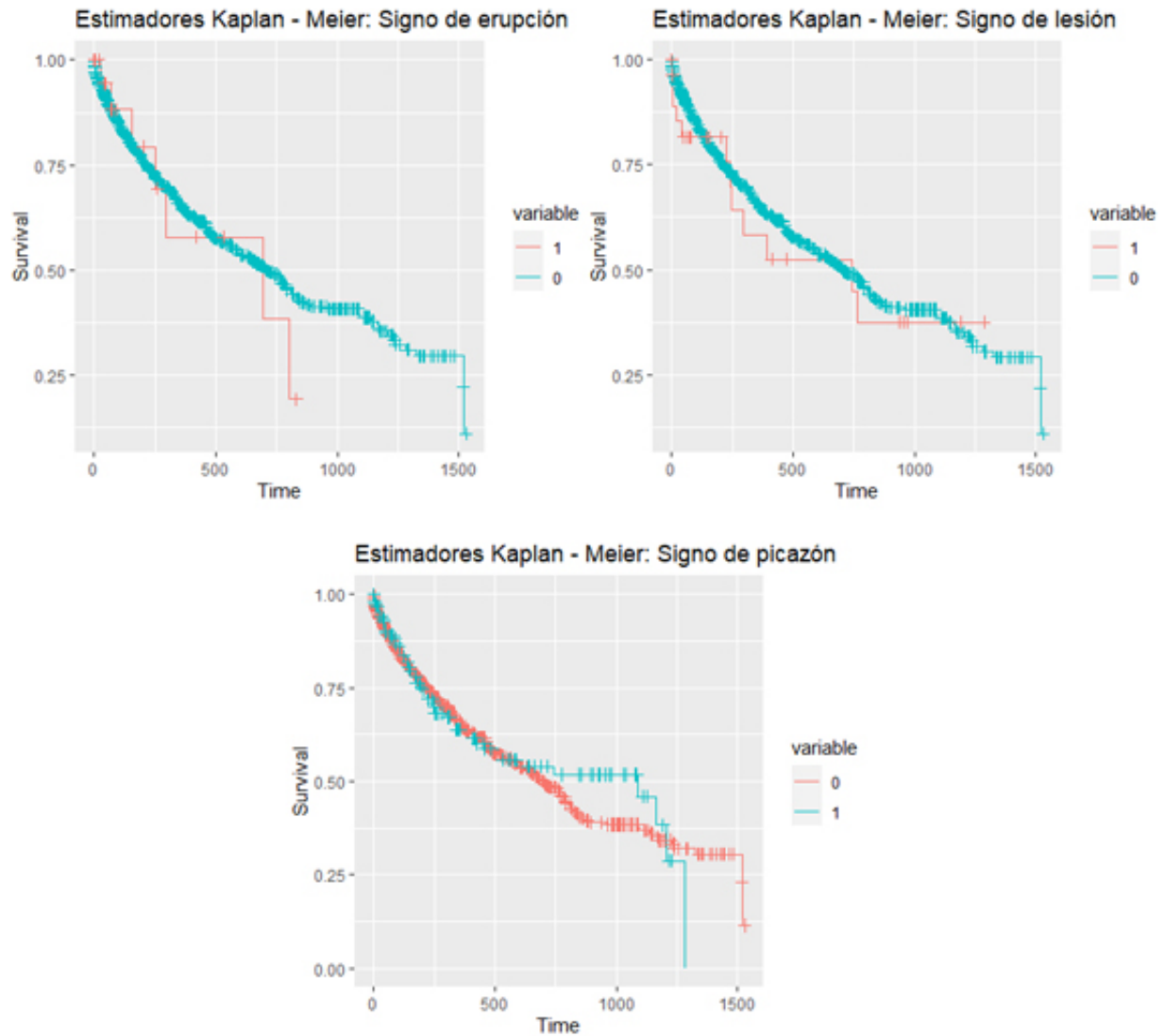


Figura 16: Gráficas que muestran las estimaciones de las funciones de supervivencia para las variables: Signo de erupción, signo de lesión y signo de picazón. Ninguna variable parece tener una diferencia significativa entre sus grupos.

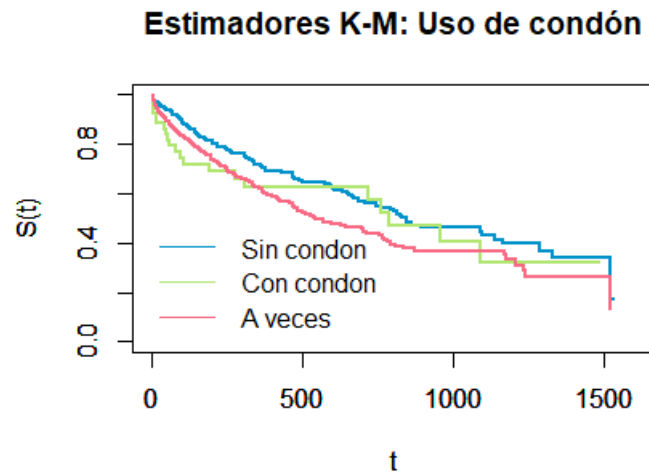


Figura 17: Es posible observar la intersección dos a dos de cada función, indicando la igualdad entre la función de supervivencia entre usar condón siempre, y a veces, y siempre y nunca usarlo.

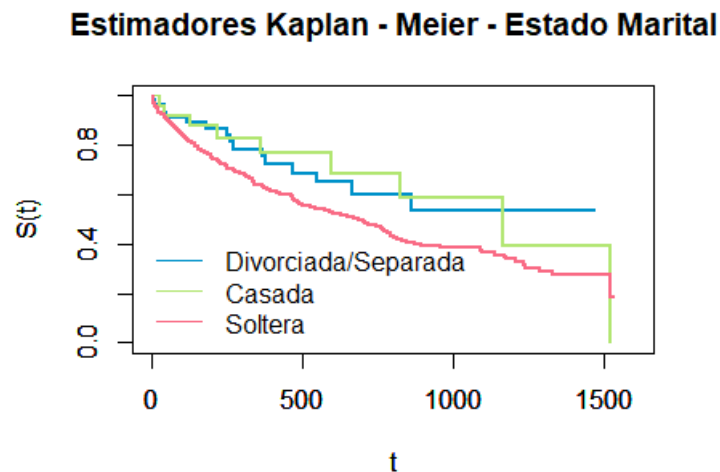


Figura 18: La gráfica señala la intersección de las funciones de supervivencia de las pacientes casadas y divorciadas/separadas, lo que significa una distribución igual, y diferentes a la supervivencia de las pacientes solteras.

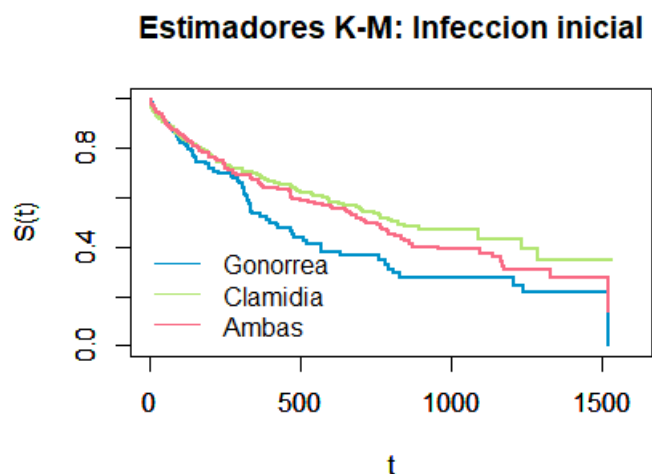


Figura 19: En la presente figura se observa cómo las gráficas de las distintas distribuciones en los primeros tiempos son similares hasta ser distintos al final del estudio, sin embargo, se aprecia la similitud entre la distribución de las pacientes que presentaron clamidia en la primera infección y las que presentaron ambas.

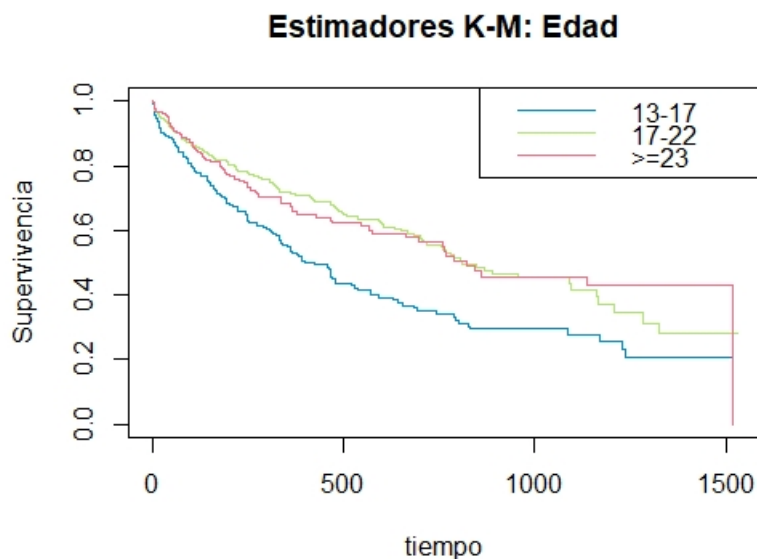


Figura 20: En la figura se puede observar que hay diferencia entre la función de supervivencia de las pacientes de más de 36 años con la de las pacientes de 13 a 22 años.

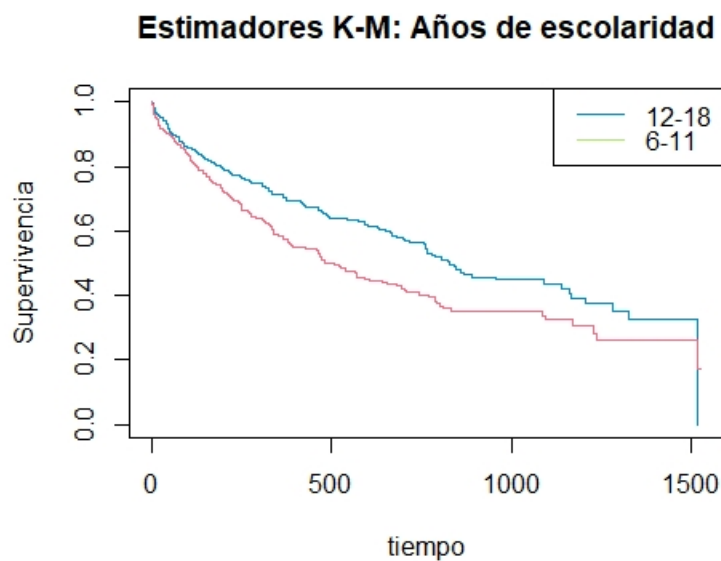


Figura 21: En la figura se puede observar que hay diferencia entre la función de supervivencia dependiendo de los años de escolaridad de las pacientes.

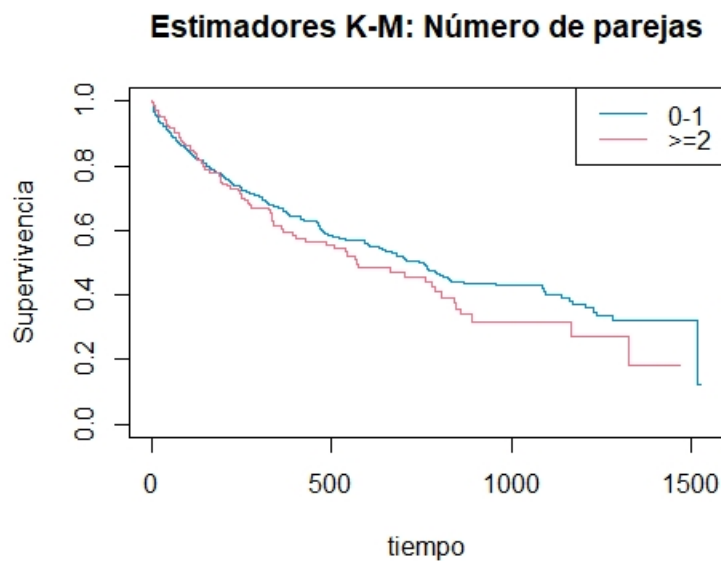


Figura 22: En la figura se puede observar que la supervivencia entre las pacientes que tienen entre 0 y 1 pareja difiere un poco de las que tienen 2 o más parejas.

6.1.1. Comparación entre poblaciones.

En la siguiente tabla se muestran los resultados de los p-values obtenidos en la comparación de poblaciones para todas las variables categóricas, en el caso de las variables continuas se

Enfermedades Sexuales

5 de junio de 2020

realizó una partición para categorizarlas.

Variable	Categorías	M-H _p -value	P-P _p -value	Conclusión
race	W/B	0.0026	0.0072	Rechazo
rinfct	0/1	0.0145	0.0411	Rechazo
os12m	0/1	4.743e-06	8.934e-06	Rechazo
os30d	0/1	2.034e-05	6.37e-05	Rechazo
rs12m	0/1	0.1565	0.1235	Acepto
rs30d	0/1	0.4285	0.4269	Acepto
abdpain	0/1	0.0394	0.0313	Rechazo
discharge	0/1	0.0917	0.1044	Acepto
itch	0/1	0.7606	0.8168	Acepto
lesion	0/1	0.6838	0.5504	Acepto
rash	0/1	0.6426	0.8689	Acepto
lymph	0/1	0.8289	0.7044	Acepto
vagina	0/1	0.0068	0.0031	Rechazo
dchexam	0/1	0.0558	0.0430	Rechazo
abnode	0/1	0.5121	0.6757	Acepto
condom	1/2	0.8972	0.8227	Acepto
condom	1/3	0.2311	0.1172	Acepto
condom	2/3	0.0031	0.0023	Rechazo
marital	S/M	0.1230	0.0916	Acepto
marital	S/D	0.0496	0.0559	Rechazo
marital	M/D	0.8890	0.8395	Acepto
iinfct	1/2	0.0031	0.0119	Rechazo
iinfct	1/3	0.0415	0.05138	Rechazo
iinfct	2/3	0.3326	0.5116	Acepto
age	13-11/>23	5e-04	0.001	Rechazo
age	13-11/17-22	2e-05	1e-05	Rechazo
age	17-22/>23	2e-05	2e-05	Rechazo
yschool	6-11/12-18	0.001	9e-04	Rechazo
npartner	0-1/2-19	0.2	0.4	Acepto

Cuadro 8: COMPARACIÓN ENTRE LAS POBLACIONES DE CADA VARIABLE, DONDE **M-H** CORRESPONDE A LA PRUEBA DE MANTEL-HAENSZEL Y **P-P** A LA PRUEBA DE PETO-PETO

6.2. Gráficas del test de Schoenfeld

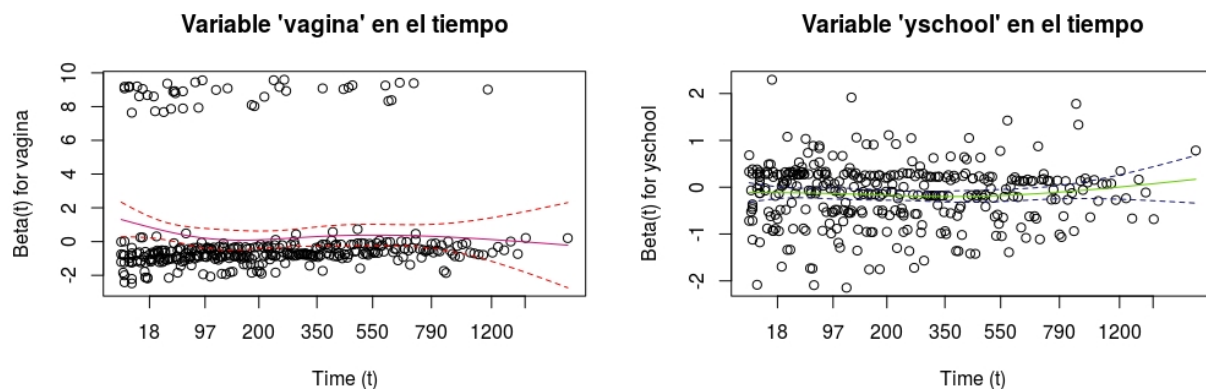


Figura 23: Coeficiente de la variable 'vagina' a lo largo del tiempo. Figura 24: Coeficiente de la variable 'yschool' a lo largo del tiempo.

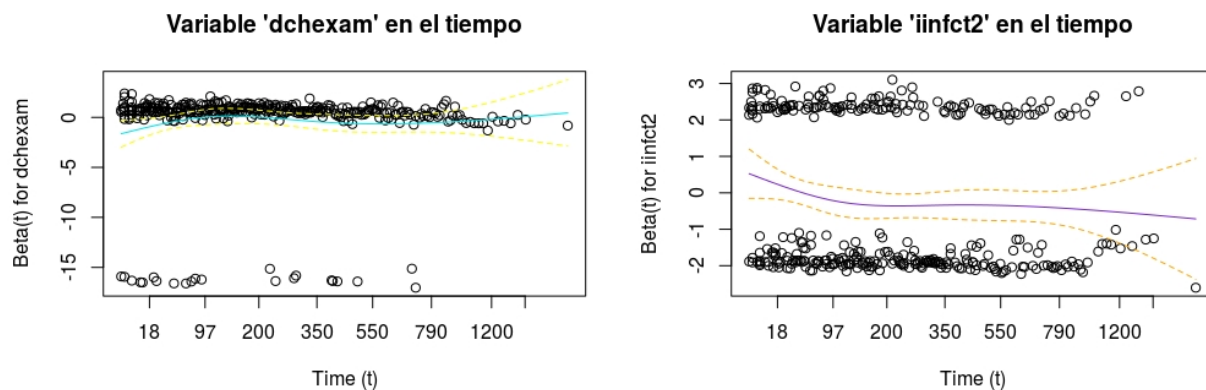


Figura 25: Coeficiente de la variable 'dhexam' a lo largo del tiempo. Figura 26: Coeficiente de la variable 'iinfct2' a lo largo del tiempo.

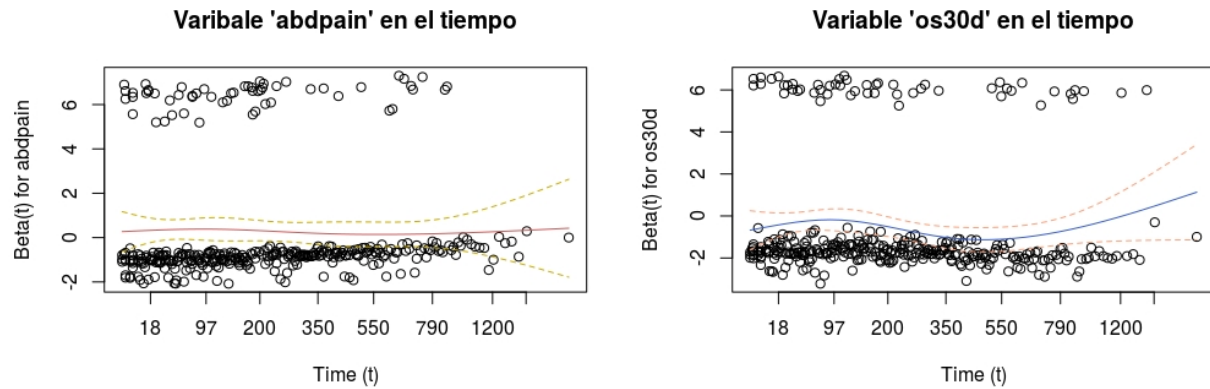


Figura 27: Coeficiente de la variable 'abdpain' a lo largo del tiempo. Figura 28: Coeficiente de la variable 'os30d' a lo largo del tiempo.