

Análisis de Supervivencia Aplicado a las Operaciones Empresariales a Inicios de la Pandemia Covid 19

Elmer David Zelaya
Universidad Nacional Autónoma de Honduras
e-mail: edzelaya@unah.hn

ÍNDICE

I.	Introducción	2
II.	Preliminares y Notación	3
II-A.	Análisis de Supervivencia	3
II-B.	Método de Kaplan Meier	3
II-C.	Test Log-Rank	3
II-D.	Cociente de Riesgos	3
III.	Resultados	4
III-A.	Primer resultado	4
III-B.	Segundo resultado	5
III-C.	Tercer resultado	7
III-D.	Resultado Central	8
IV.	Análisis e Interpretación	8
V.	Conclusiones	9
VI.	Trabajo Futuro	9
VII.	Agradecimientos	9
VIII.	Disponibilidad de Datos	9
	Referencias	9

ÍNDICE DE FIGURAS

1.	Gráfica de la frecuencia absoluta de las observaciones y censuras en toda la muestra.	4
2.	Gráfica de frecuencias de las observaciones y censuras por actividades económicas.	4
3.	Gráfica de frecuencias de las observaciones y censuras por Tipo de Financiamiento.	4
4.	Gráfica de frecuencias de las observaciones y censuras por Tamaño de la empresa.	5
5.	Gráfica de frecuencias relativas de los tiempos de supervivencia.	5
6.	Curva de Supervivencia Global.	5
7.	Curva de Supervivencia (Financiamiento Bancario).	6
8.	Curva de Supervivencia (Financiamiento Familiar).	6
9.	Curvas de Supervivencia (Micro y Pequeña empresa).	7
10.	Curvas de Supervivencias de los Grupos de Actividades Económicas	7
11.	Curvas de Supervivencias de los Grupos de los Tipos de Financiamiento	7
12.	Curvas de Supervivencias de los Grupos del Tamaño de la empresa	8

ÍNDICE DE CUADROS

I.	Resumen de frecuencias del número de empresas, censuras y observaciones según Actividad Económica.	4
II.	Resumen de frecuencias del número de empresas, censuras y observaciones según el Tipo de Financiamiento . .	4
III.	Resumen de frecuencias del número de empresas, censuras y observaciones según el Tamaño de la empresa . . .	5
IV.	Estimación de la Función de Supervivencia Global, error estándar e I.C. del 95. %	5
V.	Estimación de la Función de Supervivencia (Actividades Profesionales), error estándar e I.C. del 95 %	6
VI.	Estimación de la Función de Supervivencia (Servicios de comida), error estándar e I.C. del 95 %	6
VII.	p-valores de comparaciones en pareja (Actividades Económicas)	7
VIII.	p-valores de comparaciones en pareja (Tipos de Financiamientos)	7
IX.	Número de eventos observados y esperados, Cociente de riesgos (Alojamiento - Transporte) e IC del 95 % . . .	8
X.	Número de eventos observados y esperados, Cociente de riesgos (Comida - Transporte) e IC del 95 %	8
XI.	Número de eventos observados y esperados, Cociente de riesgos (Comercio - Comida) e IC del 95 %	8
XII.	Número de eventos observados y esperados, Cociente de riesgos (Banco - Familiar) e IC del 95 %	8

Análisis de Supervivencia Aplicado a las Operaciones Empresariales a Inicios de la Pandemia Covid 19

Resumen—En este proyecto se estimó la función de supervivencia de un conjunto de empresas de la ciudad de Choluteca, a las que se les hizo un seguimiento correspondiente a las primeras 42 semanas una vez declarado el Covid-19 como pandemia, esta función nos da la probabilidad de que estas empresas no cierren en un determinado momento, siendo este nuestro evento de interés, es decir, el cierre de operaciones. Dado que no todas las empresas cerraron, estas no proveen una información completa al estudio, dicha información no fue desechada ya que el método de Kaplan-Meier con el cual estimamos la función de supervivencia la incluye para realizar las estimaciones. Realizamos comparaciones de la supervivencia en parejas para los grupos de los diferentes rubros económicos con los que contamos en nuestra base de datos de igual manera con los tipos de financiamientos empleados para desarrollar sus actividades y el tamaño empresarial, mediante la prueba Log Rank y determinamos cuales de los grupos de empresas por cada uno de los factores antes mencionado fue significativamente distinto y luego establecimos quienes tuvieron mayor o menor riesgo de cerrar sus operaciones utilizando el cociente de riesgos.

Los resultados para los rubros económicos nos mostraron que las empresas del rubro transportista fueron significativamente distintas en su supervivencia con las que ofrecen servicios de comida y alojamiento, donde estas últimas dos tenían casi el doble riesgo de cerrar en comparación con las que ofrecen servicios de transporte, de igual forma hubo diferencia en la supervivencia entre las empresas del rubro del comercio y servicios de comida en donde las dedicadas al comercio tenían un 31% menos de riesgo de cerrar.

En cuanto a los tipos de financiamientos solo dos grupos presentaron diferencia significativa en su supervivencia, fueron los que se financian por medio de instituciones bancarias y aquellas que poseen financiamiento familiar las cuales tenían más riesgo de cerrar en comparación a las empresas financiadas por medio de bancos.

Solamente para los grupos de las pequeñas y micro empresas, es decir, para los tamaños empresariales los datos no proveyeron evidencia suficiente para determinar diferencias significativas.

I. INTRODUCCIÓN

La pandemia del Covid19 fue declarada como tal por la O.M.S. el 11 de marzo del 2020 y con estos los gobiernos se vieron en la necesidad de tomar medidas para prevenir la propagación de la mortal enfermedad, Honduras no fue la excepción y a mediados de marzo se comenzaron a aplicar restricciones a toda clase de negocios, restricciones que afectaron a los propietarios obligándolos a cerrar operaciones o buscar alternativas para evitar esto.

Para este estudio se considerarán datos recolectados mediante encuestas aplicadas a negocios de la ciudad de Choluteca que nos brindan información sobre su situación operacional, así como el tipo de financiamiento para sus operaciones, el tamaño de la empresa y el tipo de servicio que ofrecen desde la

segunda semana de marzo hasta la última semana de diciembre del 2020.

Nuestro objetivo para este trabajo está en estimar la función de supervivencia la cual nos da la probabilidad de que una empresa sobreviva, es decir, que no cierre su negocio, hasta un momento determinado y la comparación de esta por grupos para luego determinar (en caso de ser estadísticamente diferente) cuales de estas empresas tienen mayor o menor riesgo de cerrar.

Haremos uso del análisis de supervivencia el cual es un conjunto de técnicas donde la variable de interés de resultado es el tiempo hasta que ocurre un evento, para nuestro estudio el evento de interés será el cierre de operaciones. Dado que no todos los negocios suspendieron sus labores estos no presentan el evento de interés y por tanto estos proporcionan una información parcial al estudio, gracias al análisis de supervivencia podemos incorporar estos datos para estimar la función de supervivencia en el sentido del tiempo (semanas) que transcurre hasta que se presenta el evento, por medio del método de Kaplan Meier y además compararemos las curvas de supervivencia utilizando la prueba Log-Rank de grupos de cada factor (Tipo de financiamiento, tamaño de la empresa, tipo de servicio) cuya hipótesis nula es que no hay diferencia en la experiencia de supervivencia versus que si hay diferencias en la experiencia, en caso de existir diferencia significativa entre los grupos determinaremos quienes tienen mayor riesgo de cerrar su negocio mediante el cociente de riesgos.

La estructura de este trabajo es la siguiente: en la sección II que corresponde a los preliminares y notación se presentan las definiciones necesarias para la comprensión de los resultados de este proyecto, los cuales encontramos en la sección III, iniciando con un resumen descriptivo del conjunto de datos disponibles en [6], para luego aplicar el método de Kaplan Meier para la estimación de la función de supervivencia global y por cada uno de los grupos de entre los tipos de actividad económica, tipo de financiamiento y tamaño empresarial, una vez estimada esta función procedemos a realizar la comparación de esta por medio de la prueba Log-Rank y posteriormente hacemos uso del cociente de riesgos para establecer que grupos tienen mayor o menor riesgo al ser comparada con otro grupo (esto en caso de haber resultado estadísticamente diferentes en su supervivencia). En la sección IV presentamos un análisis e interpretación de nuestros resultados, en la sección V se exponen las conclusiones de este trabajo seguido de una sección de trabajo futuro donde planteamos interrogantes que permitan una mejora y mayor análisis a este estudio, la sección VII y VIII corresponden a

los agradecimientos y disponibilidad de los datos y códigos para la obtención de resultados respectivamente.

II. PRELIMINARES Y NOTACIÓN

II-A. Análisis de Supervivencia

El análisis de supervivencia es una colección de procedimientos estadísticos para el análisis de datos para los cuales la variable de resultado de interés es el tiempo hasta que ocurre un evento de interés (muerte, remisión, cierre de operaciones, etc.).

Uno de los problemas que se pueden presentar es que ciertos individuos no presenten el evento de interés, con el análisis de supervivencia podemos incorporar esta información parcial a nuestro estudio.

Definición 1. Sea T una variable aleatoria no negativa que representa la vida de los individuos en alguna población. La probabilidad de que un individuo sobreviva hasta el tiempo t viene dada por la función de supervivencia [1]

$$S(T) = P(T \geq t) \quad (1)$$

Definición 2. La función de riesgo se define como la probabilidad instantánea de muerte o fallo en el instante [3] t .

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad (2)$$

Definición 3. Se dice que una observación en un estudio está censurada cuando el individuo no presenta el evento de interés durante el tiempo duración del estudio. [4]

II-B. Método de Kaplan Meier

Sea t_1, t_2, \dots, t_k los k tiempos de supervivencia distintos y ordenados ascendentemente.

Sean n_j el número de sujetos que están vivos justo antes del tiempo t_j y d_j denota el número de “muertos” en el momento t_j , dada la función de supervivencia $S(t)$ el método de Kaplan Meier es un estimador no paramétrico de la función S definido como [5]:

$$\hat{S}(t) = \prod_{j: t_j < t} \left(\frac{n_j - d_j}{n_j} \right) \quad (3)$$

El estimador de la de varianza de (3) lo obtenemos a través de la fórmula de *Greenwood* [2]:

$$\widehat{Var}(\hat{S}(t)) = \hat{S}^2(t) \sum_{t_j \leq t} \frac{d_j}{n_j(n_j - d_j)} \quad (4)$$

Los Intervalos de confianza con $(1 - \alpha) \times 100\%$ para el momento t de $\hat{S}(t)$, asumiendo una distribución normal para las estimaciones vienen dados por [2]:

$$\hat{S}(t) \pm z_{\frac{\alpha}{2}} \times e.e.(\hat{S}(t)) \quad (5)$$

donde $e.e.(\hat{S}(t))$ es el error estándar del estimador de la función de supervivencia que se obtiene de la raíz cuadrada de (4).

II-C. Test Log-Rank

El test Log-Rank es una prueba de chi-cuadrado, en el cual se compara la experiencia de supervivencia de dos o más grupos, en esta prueba la hipótesis nula establece que la experiencia de supervivencia de los grupos es la misma y la hipótesis alternativa nos dice que al menos uno de los grupos posee una supervivencia distinta.

Al igual que muchas otras estadísticas utilizadas en otros tipos de pruebas de chi-cuadrado, para realizar la prueba es necesario calcular el número de “muertes” observadas (O_j) y esperadas (E_j). El estadístico de este test es [1]

$$X^2 = \sum_{j=1}^g \frac{(O_j - E_j)^2}{E_j} \quad (6)$$

donde $X^2 \sim \chi^2$ con $g-1$ grados de libertad (g es el número de grupos que se comparan)

II-D. Cociente de Riesgos

La prueba Log-Rank la usamos para comparar la supervivencia de dos o más grupos, pero esta al ser una prueba de hipótesis no provee información directa de como estos se diferencian.

Una forma de medir la experiencia de supervivencia relativa entre dos grupos es por medio del cociente de riesgos (HR), el cual da una estimación de las tasas relativas de riesgos de los dos grupos, con esto determinamos cuál de los grupos comparados tiene mayor o menor riesgo de cerrar su negocio.

Consideremos un grupo A, sean O_A y E_A el número de eventos observados y esperados del grupo A, la razón O_A/E_A nos da la tasa relativa de riesgo del grupo, análogamente para un grupo B, combinando estas dos tasas en un solo estadístico calculamos el cociente de riesgos (HR), el cual se define como el cociente de estas dos tasas:

$$HR = \frac{O_A/E_A}{O_B/E_B} \quad (7)$$

Su intervalo de confianza del 95 % viene dado por :

$$\begin{aligned} \text{limite inferior} &: \exp(\log(HR) - 1,96 \times e.e.(\log(HR))) \\ \text{limite superior} &: \exp(\log(HR) + 1,96 \times e.e.(\log(HR))) \end{aligned}$$

$$\text{Donde } e.e.(\log(HR)) = \sqrt{\left(\frac{1}{E_A} + \frac{1}{E_B} \right)} \quad (8)$$

Considerando HR como en (7).

- Si $HR > 1$, el grupo A tiene mayor riesgo de ocurrencia del evento en comparación con el grupo B.
- Si $HR < 1$, el riesgo de ocurrencia del evento es mayor en el grupo B en comparación con el grupo A.
- Si $HR = 1$, el riesgo en ambos grupos es el mismo. [7], [8], [9].

III. RESULTADOS

III-A. Primer resultado

Para el desarrollo de este proyecto utilizaremos un conjunto de datos que están disponibles en [6], estos nos proporcionan información como ser el tipo de servicio que es ofrecido al público por parte de los negocios, entre ellos tenemos servicios de actividades profesionales, alojamiento, comercio, comida y de transporte, así como el tipo de financiamiento empleado para desarrollar sus actividades en las cuales encontramos financiamientos por medio de instituciones bancarias y no bancarias, cooperativas y financiamientos de tipo familiar.

Los datos proveen información sobre el tamaño empresarial en los cuales se consideran las micro y pequeñas empresas y también el estado en que se encuentra el negocio, es decir, si cerraron sus establecimientos o si continuaron sus operaciones.

En este estudio se considera el tiempo en semanas contabilizadas desde la segunda semana de marzo que es donde comienza el periodo de observación hasta que se presenta el evento de interés (cierre de operaciones), aquellas empresas en las que el periodo de observación concluyó (última semana de diciembre) y no pararon sus actividades se consideran datos censurados.

A continuación, presentaremos un resumen descriptivo del conjunto de datos disponibles en [6].

La base de datos consta con información de 474 empresas en las cuales se observaron 233 eventos que representan un 49,16% del total de los negocios y el restante 50,84% continuó laborando cuando culminó el periodo de observación.

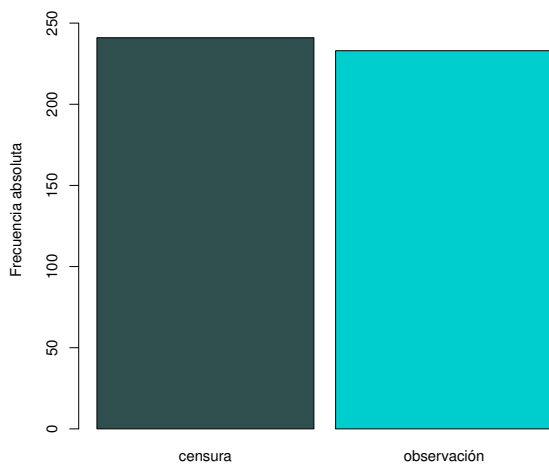


Figura 1: Gráfica de la frecuencia absoluta de las observaciones y censuras en toda la muestra.

En cuanto al tipo de servicio que ofrecen los negocios, encontramos que aquellos que brindan servicios de transporte, el 64% continuó laborando (no presentó el evento de interés) y el 56% de las que ofrecen servicio de Alojamiento cerraron operaciones.

Actividades económicas	número de empresas	Censuras	observaciones
Actividades profesionales	60	32	28
Alojamiento	69	31	38
Comercio	105	59	46
Servicios de comida	174	77	97
Transporte	66	42	24

Cuadro I: Resumen de frecuencias del número de empresas, censuras y observaciones según Actividad Económica.

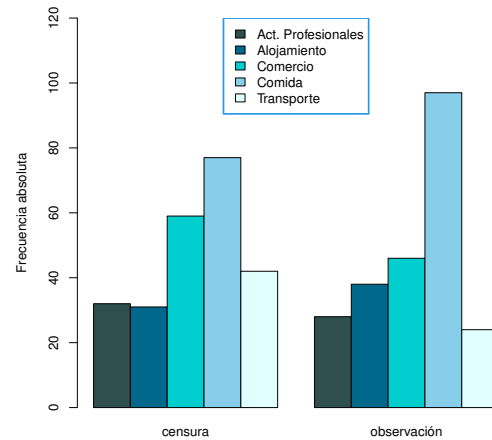


Figura 2: Gráfica de frecuencias de las observaciones y censuras por actividades económicas.

Por otro lado, 76 negocios cerraron sus establecimientos, estos representan el 16,03% con respecto a los tipos de financiamiento, en estos el financiamiento era familiar, mientras que 87 negocios que representan un 18,35% siguieron trabajando, estos se financiaban por medio de bancos.

Tipo de Financiamiento	número de empresas	Censuras	observaciones
Bancos	155	87	68
Cooperativas	130	67	63
Inst. No Bancarias	59	33	26
Familiar	130	54	76

Cuadro II: Resumen de frecuencias del número de empresas, censuras y observaciones según el Tipo de Financiamiento

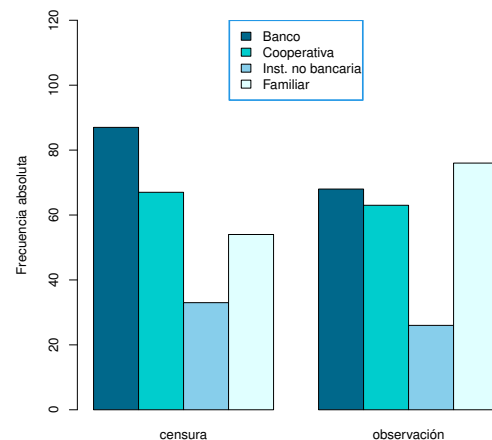


Figura 3: Gráfica de frecuencias de las observaciones y censuras por Tipo de Financiamiento.

Con respecto al tamaño de la empresa se encontró que 75,32 % son micro empresas y el 24,68 % corresponde a las pequeñas empresas, donde el 50,14 % y 46,15 % de las micro y pequeñas empresas respectivamente cerraron.

Tamaño de la empresa	número de empresas	Censuras	observaciones
Micro	357	178	179
Pequeña	117	63	54

Cuadro III: Resumen de frecuencias del número de empresas, censuras y observaciones según el Tamaño de la empresa

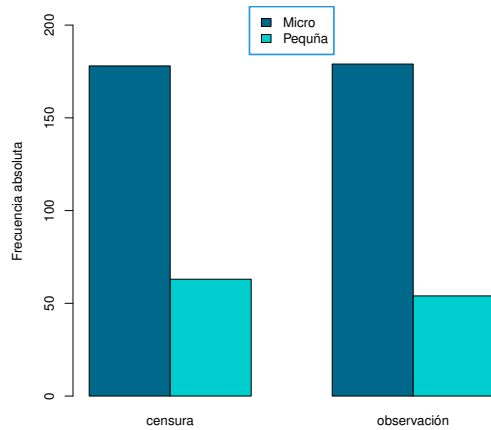


Figura 4: Gráfica de frecuencias de las observaciones y censuras por Tamaño de la empresa.

En los tiempos de supervivencia podemos apreciar que entre el 15 % y 16 % de los negocios cerraron en la segunda y séptima semana, notemos que es en la semana 20 (última semana de Julio) donde se registró una mayor cantidad eventos.

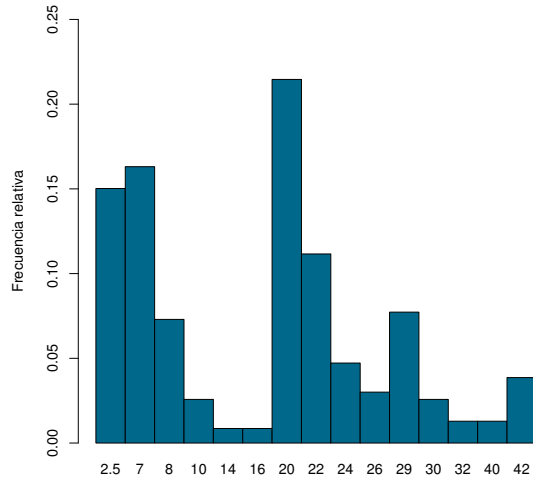


Figura 5: Gráfica de frecuencias relativas de los tiempos de supervivencia.

III-B. Segundo resultado

En estudios de análisis de supervivencia nos interesa el tiempo que transcurre desde la fecha de inicio de un periodo de observación, hasta que se presenta un evento que es de nuestro interés, para este proyecto la fecha de inicio de observación es la segunda semana de marzo del 2020 y el evento de interés es el cierre de operaciones empresariales.

Por medio del método de Kaplan-Meier (K-M) podemos estimar la función de supervivencia(1), la cual nos da la probabilidad de que una empresa sobreviva, es decir, que no cierre su negocio, hasta un momento determinado en el tiempo.

En esta sección comenzaremos con la presentación de los resultados obtenidos de la estimación de la función de supervivencia (global), así como el cálculo del error estándar y los intervalos de confianza del 95 % para esta.

t_j	n_j	d_j	$\hat{S}(t)$	$e.e$	LI	LS
2.5	474	35	0.926	0.0120	0.903	0.950
7.0	439	38	0.846	0.0166	0.813	0.878
8.0	401	17	0.810	0.0180	0.775	0.845
10.0	384	6	0.797	0.0185	0.761	0.834
14.0	378	2	0.793	0.0186	0.757	0.830
16.0	376	2	0.789	0.0187	0.752	0.826
20.0	374	50	0.684	0.0214	0.642	0.725
22.0	324	26	0.629	0.0222	0.585	0.672
24.0	298	11	0.605	0.0224	0.561	0.649
26.0	287	7	0.591	0.0226	0.546	0.635
29.0	280	18	0.553	0.0228	0.508	0.598
30.0	262	6	0.540	0.0229	0.495	0.585
32.0	256	3	0.534	0.0229	0.489	0.579
40.0	253	3	0.527	0.0229	0.482	0.572
42.0	250	9	0.508	0.0230	0.463	0.553

Cuadro IV: Estimación de la Función de Supervivencia Global, error estándar e I.C. del 95. %

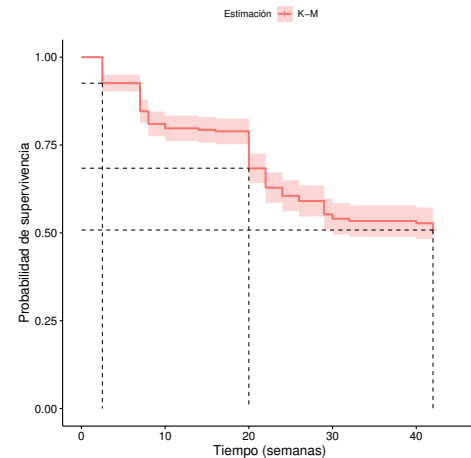


Figura 6: Curva de Supervivencia Global.

La Supervivencia global estimada con K-M mostró que para el primer tiempo de supervivencia $t_1 = 2.5$, las empresas tenían una probabilidad del 92.6 % de no cerrar sus operaciones.

En la semana 20 (última semana de julio) habían 374 empresas abiertas y en esa semana registraron 50 eventos, siendo esta la semana donde más empresas cerraron sus establecimientos, en ese momento la probabilidad de supervivencia fue de 64.8 %.

Al finalizar el periodo de observación (última semana de diciembre) se tuvo un 50.8 % de probabilidad que las empresas no pararan sus labores.

Notemos que a medida transcurre el tiempo la probabilidad de sobrevivir disminuye, pero esta no llega a 0 dado que existe una fuerte presencia de empresas que no cerraron

(datos censurados) al finalizar el estudio.

Ahora presentaremos la estimación de la función de supervivencia por grupos en donde consideramos aquellos que presentaron una mayor probabilidad de no cerrar sus empresas en los primeros tiempos de observación y aquellos en los que al finalizar el estudio tuvieron una probabilidad de supervivencia menor.

Comenzaremos con el tipo de actividades económicas, en este factor el grupo que tuvo una mejor supervivencia fue el de las actividades profesionales, en el siguiente cuadro apreciamos la estimación de esta.

t_j	n_j	d_j	$\hat{S}(t)$	$e.e$	LI	LS
2.5	60	2	0.967	0.0232	0.921	1.000
7.0	58	7	0.850	0.0461	0.760	0.940
8.0	51	1	0.833	0.0481	0.739	0.928
20.0	50	8	0.700	0.0592	0.584	0.816
22.0	42	3	0.650	0.0616	0.529	0.771
29.0	39	4	0.583	0.0636	0.459	0.708
40.0	35	1	0.567	0.0640	0.441	0.692
42.0	34	2	0.533	0.0644	0.407	0.660

Cuadro V: Estimación de la Función de Supervivencia (Actividades Profesionales), error estándar e I.C. del 95 %

En el primer tiempo de observación se estimó una supervivencia del 96.7%, luego en la séptima semana se encontró que habían 158 empresas abiertas, cuya probabilidad de no cerrar oscilaba entre 77.1 % y 88.4 %. En la semana 20 es en la que se registró un mayor cantidad de eventos para esta actividad económica, aquí la probabilidad de supervivencia se redujo a 70 %.

En cuanto al grupo de las actividades económicas que tuvo una supervivencia menor al finalizar el periodo de observación tenemos al grupo de las empresas de ofrecen servicios de comida.

t_j	n_j	d_j	$\hat{S}(t)$	$e.e$	LI	LS
2.5	174	16	0.908	0.0219	0.865	0.951
7.0	158	14	0.828	0.0286	0.771	0.884
8.0	144	9	0.776	0.0316	0.714	0.838
10.0	135	5	0.747	0.0330	0.683	0.812
20.0	130	19	0.638	0.0364	0.567	0.709
22.0	111	11	0.575	0.0375	0.501	0.648
24.0	100	8	0.529	0.0378	0.455	0.603
26.0	92	4	0.506	0.0379	0.431	0.580
29.0	88	2	0.494	0.0379	0.420	0.569
30.0	86	4	0.471	0.0378	0.397	0.545
32.0	82	2	0.460	0.0378	0.386	0.534
42.0	80	3	0.443	0.0377	0.369	0.516

Cuadro VI: Estimación de la Función de Supervivencia (Servicios de comida), error estándar e I.C. del 95 %

Hasta la semana 20 ya habían ocurrido 44 eventos dejando a 130 en riesgo con una probabilidad de sobrevivir del 63.8 %, al final de periodo de observación la supervivencia se

redujo al 44.3 %.

En cuanto al tipo de financiamiento, 155 empresas se financian por medio de bancos, en el comienzo del estudio tenían un probabilidad de supervivencia de 93.5 %, aproximadamente el 75 % tenían una probabilidad de sobrevivir las primeras 22 semanas del 72.9 %. En el último tiempo de supervivencia ($t = 42$) se registraron 2 eventos y estos tenían una probabilidad de continuar operando de entre 48.3 % y 63.9 %.

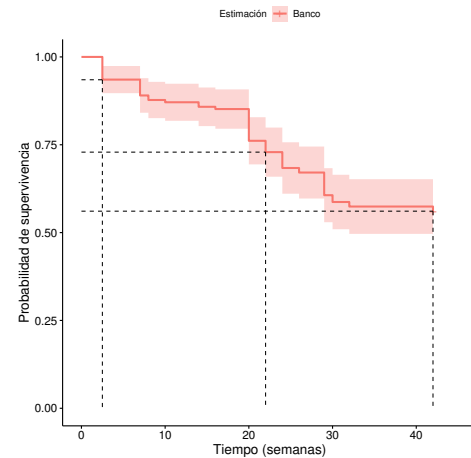


Figura 7: Curva de Supervivencia (Financiamiento Bancario).

Por otro lado los negocios que disponen de un financiamiento familiar, con el estimador de Kaplan-Meier se obtuvo una supervivencia en el primer tiempo de observación $t_1 = 2.5$ del 89.2 %. Aproximadamente el 50 % de estas empresas sobrevivieron a la semana 26 con una probabilidad de lograrlo de 48.5 %. Al finalizar el estudio tenían una probabilidad de no detener sus negocios del 41.5 % (I.C. 95 %, 33.1 % a 50 %) siendo la más baja en cuanto a los tipos de financiamiento.

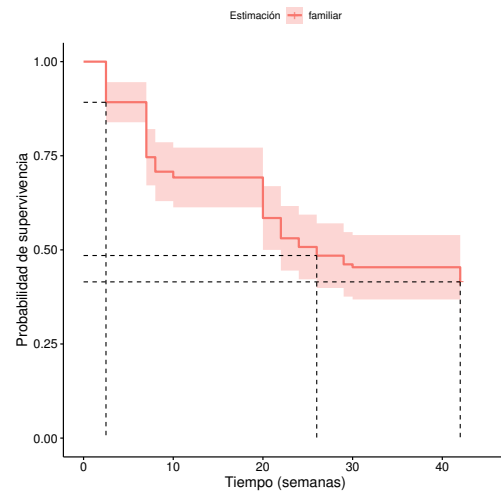


Figura 8: Curva de Supervivencia (Financiamiento Familiar).

Para el tamaño de las empresas en la semana 7 se registraron 26 y 12 eventos en las micro y pequeñas empresas respecti-

vamente, para esta semana la probabilidad de sobrevivir es de 84.9 % y 83.8 %, siendo esta dentro de las primeras 16 semanas donde se reportaron más eventos, luego de esto, fue hasta en la semana 20 donde hubo un crecimiento en número de empresas cerradas, aquí la probabilidad fue del 67.5 % y 70.9 %. Se finalizó el estudio con 7 y 2 eventos en las micro y pequeñas empresas en la cual que para esta última semana se tenía una probabilidad de de no parar sus labores económicas del 49.9 % y 53.8 % respectivamente.

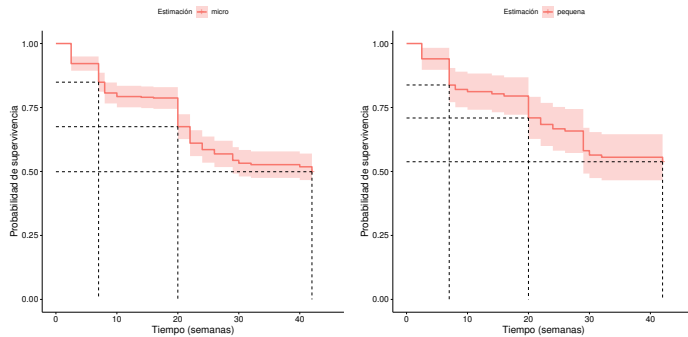


Figura 9: Curvas de Supervivencia (Micro y Pequeña empresa).

III-C. Tercer resultado

Luego de realizar la estimación de la función de supervivencia para cada uno de los grupos que hemos definido, nos preguntamos si existe diferencia en la experiencia de supervivencia, es decir, si las probabilidades de cerrar una empresa varían o son las mismas según al rubro que esta se dedica, de forma análoga con el tipo de financiamiento utilizado y el tamaño empresarial.

Con las curvas de supervivencia las cuales son una representación gráfica del comportamiento de la supervivencia a lo largo del tiempo de seguimiento del estudio, podemos darnos una idea intuitiva de esto, para determinar si hay diferencia significativa empleamos la prueba Log-Rank con la cual obtuvimos los siguientes resultados:

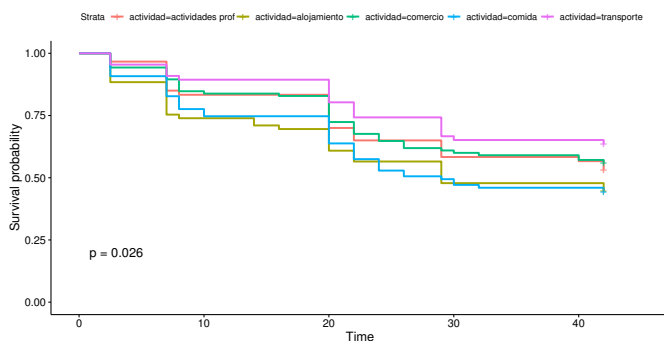


Figura 10: Curvas de Supervivencias de los Grupos de Actividades Económicas

En los grupos de las actividades económicas se obtuvo un valor p de $0.026 < 0.05$, con esto rechazamos la hipótesis

nula y por tanto al menos una de las supervivencias es significativamente diferente.

Realizando una comparación en parejas para determinar cuáles de estas curvas fueron significativamente distintas se obtuvo que el grupo de las empresas que ofrecen servicios de transporte presentaron diferencia significativa al ser comparadas con aquellas que ofrecen servicios de alojamiento y aquellos que brindan servicios de comida de igual forma existió diferencia en la supervivencia entre las empresas del rubro de comida y comercio.

	Actividades Prof.	Alojamiento	Comercio	Comida
Alojamiento	0.2310	-	-	-
Comercio	0.7769	0.0991	-	-
Comida	0.1818	0.9525	0.0453	-
Transporte	0.2253	0.0158	0.2895	0.0058

Cuadro VII: p-valores de comparaciones en pareja (Actividades Económicas)

Aplicando la prueba Log-Rank con un nivel de significancia del 5 %, con un p-valor de 0.024 concluimos que al menos una de las curvas de supervivencia es significativamente diferente.

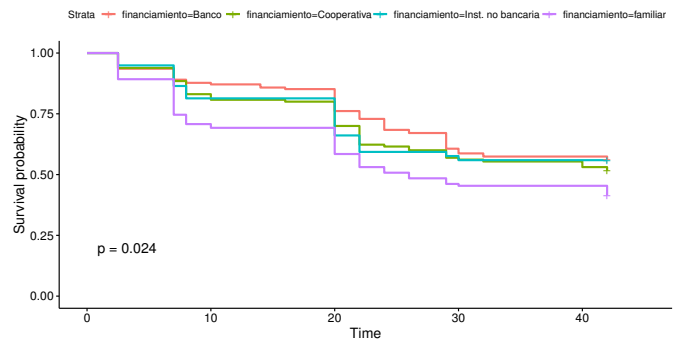


Figura 11: Curvas de Supervivencias de los Grupos de los Tipos de Financiamiento

Nuevamente realizamos comparaciones en pareja para los diferentes tipos de financiamientos, los grupos que presentaron una diferencia significativa fueron aquellos que emplean financiamientos por medio instituciones bancarias y los que lo hacen mediante un financiamiento familiar.

	Banco	Cooperativa	Inst. no bancaria
Cooperativa	0.3551	-	-
Inst. no bancaria	0.7628	0.6933	-
familiar	0.0039	0.0583	0.0669

Cuadro VIII: p-valores de comparaciones en pareja (Tipos de Financiamientos)

Solamente para los grupos del tamaño empresarial, es decir los grupos de las micro y pequeñas empresas, los datos no dieron evidencia suficiente para determinar que las curvas de supervivencia son distintas, el p-valor para esta comparación

fue de 0.42 con lo cual no es posible rechazar la hipótesis nula de la prueba.

En la figura 12 podemos observar cierta similitud en el comportamiento de la supervivencia entre estos grupos.

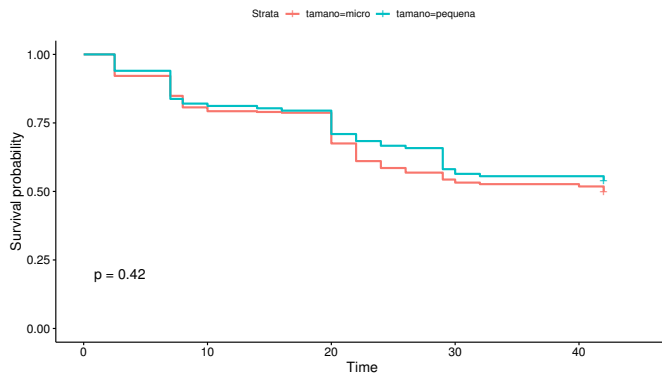


Figura 12: Curvas de Supervivencias de los Grupos del Tamaño de la empresa

III-D. Resultado Central

Ahora que hemos determinado quienes de estos grupos son estadísticamente diferentes en su experiencia de supervivencia, utilizaremos el cociente de riesgos (HR) para establecer quienes de estos tienen mayor o menor riesgo de cerrar su negocio al ser comparado con el grupo con el cual que se presentó la diferencia significativa.

En los grupos de actividades económicas, las empresas que ofrecen servicios de transporte y las que ofrecen servicios de alojamiento fueron estadísticamente diferentes.

	Observados	Esperados	HR	Li	Ls
Alojamiento(1)	38	28.9191	$\frac{O_1/E_1}{O_2/E_2} = 1.8112$	1.2582	2.6072
Transporte(2)	24	33.0809			

Cuadro IX: Número de eventos observados y esperados, Cociente de riesgos (Alojamiento - Transporte) e IC del 95 %

Al calcular el cociente de riesgo obtuvimos un valor de 1.81 con lo cual las empresas del rubro del hospedaje tienen mayor riesgo de cerrar comparado con los del rubro de Transporte, estas empresas tienen casi el doble de riesgo.

El grupo de las empresas que ofrecen servicios de comida también tuvo una experiencia de supervivencia distinta con las del grupo de las empresas del rubro transportista.

	Observados	Esperados	HR	Li	Ls
Comida(1)	97	83.5120	$\frac{O_1/E_1}{O_2/E_2} = 1.8143$	1.4641	2.2482
Transporte(2)	24	37.4880			

Cuadro X: Número de eventos observados y esperados, Cociente de riesgos (Comida - Transporte) e IC del 95 %

El riesgo de cerrar sus negocios entre aquellas empresas del rubro de las comidas fue entre 1.46 y 2.25 veces más alto que el riesgo de cerrar sus operaciones entre los que ofrecen servicios de transporte.

Otros de los grupos que al ser comparados fueron significativamente diferentes fueron los del rubro del comercio y comidas.

	Observados	Esperados	HR	Li	Ls
Comercio(1)	46	57.2870	$\frac{O_1/E_1}{O_2/E_2} = 0.7095$	0.5477	0.9192
Comidas(2)	97	85.7130			

Cuadro XI: Número de eventos observados y esperados, Cociente de riesgos (Comercio - Comida) e IC del 95 %

En este caso el cociente de riesgo fue de 0.71 (IC 95 %, 0.55 a 0.92) y con esto los negocios dedicados al comercio tienen un 31 % menos riesgo de parar sus actividades en comparación con los negocios dedicados al servicio de comida.

Dentro de los grupos de los tipos de financiamientos que utilizan los negocios para desarrollar sus actividades económicas, dos grupos presentaron diferencia significativa en su supervivencia, estos fueron los que se financian por medio de instituciones bancarias y los que lo hacen con un financiamiento familiar.

	Observados	Esperados	HR	Li	Ls
Banco(1)	68	84.3916	$\frac{O_1/E_1}{O_2/E_2} = 0.6320$	0.5106	0.7823
Familiar(2)	76	59.6084			

Cuadro XII: Número de eventos observados y esperados, Cociente de riesgos (Banco - Familiar) e IC del 95 %

Calculando el HR se obtuvo que este es de $0.63 < 1$, con lo cual los negocios con un financiamiento bancario tienen un menor riesgo (37 % menos riesgo) de cerrar en comparación con los negocios financiados familiarmente.

IV. ANÁLISIS E INTERPRETACIÓN

Nuestro interés en este estudio fue el de estimar la función de supervivencia de un conjunto de empresas de la ciudad de Choluteca y a partir de esta dar a conocer con qué probabilidad estos negocios podrían mantenerse en funcionamiento una vez declarada la actual crisis sanitaria hasta la última semana de diciembre del 2020, ante una considerable cantidad de datos censurados, es decir, aquellos establecimientos que no presentaron el evento de interés durante el periodo de observación, utilizamos el método de Kaplan Meier para realizar estas estimaciones dado que este método incluye información parcial y completa en sus cálculos. Una vez que estimamos la función de supervivencia global y por cada grupo dentro de los factores que con base en la información proporcionada consideramos, aplicamos la prueba Log Rank para determinar si existía de diferencias significativas en la supervivencia entre

estos grupos dando como resultado que entre los grupos de los rubros económicos tres de estas actividades económicas fueran estadísticamente distintas en su supervivencia, las empresas que presentaron diferencias fueron las dedicadas al hospedaje y las dedicadas al transporte de igual forma las que ofrecen servicios de comida fue estadísticamente diferente en comparación a las empresas de transporte, las empresas del rubro del comercio también presentó diferencias con las del rubro de comidas, una vez determinado esto, por medio del cociente de riesgos pudimos establecer quienes de estas empresas tiene mayor o menor riesgo de cerrar en comparación con la otra empresa con la que fue estadísticamente diferente en su experiencia de supervivencia, donde las empresas del transporte tuvieron menor riesgo de cerrar comparado con las dedicadas al alojamiento y comida, de igual forma las empresas del rubro de comercio tuvo un menor riesgo en comparación con las que ofrecen servicios de comida. Realizando un análisis similar dentro de los tipos de financiamientos, logramos establecer que con base en la información disponible las empresas con un financiamiento bancario fueron estadísticamente diferentes en su supervivencia con las empresas familiares en las cuales el riesgo de cerrar era casi del doble, finalmente en cuanto a los tamaños empresariales que parar nuestro estudio fueron considerados únicamente las micro y pequeñas empresas los datos no proporcionaron evidencia suficiente para establecer diferencias, con esto logramos cumplir con nuestros objetivos de estimar, comparar y determinar quiénes de las empresas entre los factores considerados (Tipo de actividad económica, tipo de financiamiento, tamaño empresarial) tienen más posibilidades de mantener o no sus operaciones.

V. CONCLUSIONES

En el presente proyecto se ha evaluado la supervivencia de un conjunto de datos que provee información de 474 empresas de la ciudad de Cholul de Cholul a las que se le realizó un seguimiento de 42 semanas iniciando desde la segunda semana de marzo hasta la última de diciembre del 2020, casi el 50 % de estas empresas de cerraron sus establecimientos, al inicio contaban con una probabilidad de sobrevivir de 92,6 % y la cual se redujo un 41,8 % llegando a ser de 50,8 % al finalizar el estudio, donde el grupo menos afectado dentro del tipo de rubro económico fue el de las dedicadas al transporte con una supervivencia final de 63,6 %, esta al ser comparada con los otros rubros fue significativamente diferente en su supervivencia con los negocios que ofrecen servicios de comida y alojamiento, estas últimas dos tenían casi el doble del riesgo de cerrar en comparación a las que ofrecen servicio de transporte.

En cuanto a los tipos de financiamientos empleados para desarrollar sus actividades económicas, los negocios cuyo financiamiento era familiar fueron las más afectadas ya que casi el 60 % de estos cerraron, por otro lado, los negocios financiados por medio de instituciones bancarias fueron los que mejor resistieron ya que al finalizar del estudio tenían una probabilidad de no parar sus labores de 56,10 % siendo esta superior a los demás tipos de financiamientos.

Con base en datos que contamos, no logramos determinar diferencias significativas en cuanto a los tamaños empresariales.

VI. TRABAJO FUTURO

Al realizar los cálculos para obtener los resultados surgieron unas preguntas como: ¿ Será que existe diferencia significativa al comparar los tipos de financiamientos de las empresas cuyo rubro es el comercio? tambien nos preguntamos si las micro empresas cuyo rubro es comercio y son financiadas a través de cooperativas tienen mayor riesgo que aquellas cuyo financiamiento es familiar.

Podríamos contruir un modelo multivariante que nos permita dar respuesta a estas preguntas y variantes de estas.

VII. AGRADECIMIENTOS

Gracias a Dios, mi padre Elmer Zelaya Granera y madre Susy Cruz por su apoyo incondicional y a cada miembro de mi familia y amigos que me han dado ánimos para seguir adelante, a mis profesores en especial al PhD Fredy Vides por cada una de sus observaciones y consejos para el desarrollo de este proyecto y formación profesional.

VIII. DISPONIBILIDAD DE DATOS

El conjunto de datos y scripts que se desarrollan para la obtención de resultados para este proyecto está disponible en [6].

REFERENCIAS

- [1] David G. Kleinbaum and Mitchel Klein, *Survival Analysis A Self-Learning Text*, 3rd ed, Springer.
- [2] Mara Tableman & Jong Sung Kim with a contribution from Stephen Portnoy, *Survival analysis using S: analysis of time-to-event data*, Chapman & Hall/CRC.
- [3] Kris Bogaerts, Arnošt Komárek, *Survival Analysis With interval-censored DATA*, Chapman & Hall/CRC.
- [4] Lawless, J.F., *Statistical Models and Methods for Lifetime Data*, 2nd ed, John Wiley & Sons.
- [5] Annette J. Dobson, *An introduction to generalized linear models*, 2nd ed, Chapman & Hall/CRC.
- [6] <https://github.com/ElmerZelaya/MM-700-Seminario-de-Investigacion>.
- [7] David Machin, Yin Bun Cheung, Mahesh K.B. Parmar, *Survival Analysis: A Practical Approach*, 2nd ed, John Wiley & Sons.
- [8] Douglas G. Altman, *Practical statistics for medical research*, Chapman & Hall.
- [9] T.D.V. Swinscow and M.J. Campbell, *Statistics at Square One*, 10th ed, BMJ Books.