

**Uso del Método “Combinación de Riesgos”
para estimar la función de supervivencia en
presencia de riesgos competitivos
dependientes.**

Osnamir Elias Bru Cordero

Universidad Nacional de Colombia
Facultad de Ciencias, Escuela de Estadística
Medellín, Colombia
2015

Uso del Método “Combinación de Riesgos” para estimar la función de supervivencia en presencia de riesgos competitivos dependientes.

Osnamir Elias Bru Cordero

Tesis de grado presentada como requisito parcial para optar al título de:
Magister en Ciencias Estadística

Director:

Mario César Jaramillo Elorza, Ph.D. en Estadística

Líneas de Investigación:

Estadística industrial

Grupos de Investigación:

Estadística industrial

Universidad Nacional de Colombia
Facultad de Ciencias, Escuela de Estadística
Medellín, Colombia

2015

A Dios, a mis padres Gilberto (Q.E.P.D) y Tulia Ester por darme la vida y educarme, a todos mis familiares que de una u otra forma me apoyaron anímicamente y/o económicamente en este logro, en especial a mi esposa Erika María y a nuestro hijo Said Elias. A mi hermano uterino Samir, que siempre ha sido y será un modelo a seguir.

Agradecimientos

Los autores agradecen de manera especial el apoyo financiero de la convocatoria del programa nacional de apoyo a estudiantes de posgrado para el fortalecimiento de la investigación, creación e innovación de la Universidad Nacional de Colombia 2013 – 2015, como parte del proyecto de investigación “Uso del Método Combinación de Riesgo para estimar la función de supervivencia en presencia de riesgos competitivos dependientes” Código: 22352.

De manera especial, le agradezco al profesor Mario Cesar Jaramillo por sus aportes y estar siempre dispuesto a sacar adelante este proyecto. Además, al profesor Sergio Yañez por ser un apoyo incondicional durante mi maestría y a mis compañeros Cristian Correa y Elisabeth Estrada.

Resumen

En este trabajo se comparan distintas estructuras de dependencia para los riesgos que compiten en un modelo de riesgos competitivos dependientes trivariado, haciendo uso de las técnicas C-Vines y D-Vines cópulas, vía simulación estadística. Los *Vines* permiten obtener flexibilidad multivariada y son capaces de capturar todo el rango posible de dependencias entre los riesgos competitivos, las cuales son de gran interés en los mercados financieros, problemas sociales, genéticos, entre otros. Seguidamente, se estima la función de supervivencia para el tiempo mínimo, tanto para el caso independiente, por medio del estimador Kaplan Meier, como para el caso dependiente, en el que usaremos el método combinación de riesgos propuesto por Lo & Wilke (2010) el cual es una extensión del estimador cópula gráfico propuesto por Zheng & Klein (1995). Los C-D *Vines* cópulas, trabajan con una variedad de cópulas bivariadas, las cuales se pueden seleccionar de manera independiente y permiten tener una amplia gama de posibilidades para la caracterización de la dependencia de los riesgos que compiten, se estudian casos particulares donde dos de los tres riesgos tienen igual dependencia y el riesgo restante es independiente a los anteriores. También se estudia el caso donde dos riesgos presentan igual dependencia y el otro con dependencia alta. Además se analiza un caso particular donde los tres riesgos presentan distinta dependencia. En todos los casos estudiados, el método de combinación de riesgos es una buena alternativa para estimar las funciones de distribución marginal y la función de supervivencia cuando se tiene dependencia entre los riesgos de un modelo de riesgos competitivos dependientes.

Palabras clave: C-Vines, D-Vines, Combinación de Riesgos, Cópula Gráfico.

Abstract

This paper compares different structures of dependence for the risks that compete in a trivariate competitive risk dependent model, making use of the technical C-Vines and D-vines copula, via statistical simulation. The vines can obtain multivariate flexibility and are able to capture all the possible range of dependencies between the competitive risks which are of great interest in financial markets, social, genetic among others problems. Then, estimated survival function for the minimum time, both for the case independent, through the Kaplan Meier estimator, to the dependent case, in which we will use the risk pooling method proposed by Lo & Wilke (2010), which is an extension of the estimator intercourse chart proposed by Zheng & Klein (1995).

The C-DVines copulations work with a cascade of bivariate copulations, which can be selected independently and allow a wide range of possibilities for the characterization of the dependence of the risks that compete, we study particular cases where two of the three risks have equal dependence and the remaining risk is independent to the previous ones. It

also explores the case where two risks have equal dependence and the other with high dependency. In addition, a particular case where the three risks have different dependence is analyzed. In all the cases studied, the risk pooling method is a good alternative to estimate the marginal distribution functions and the survival function when you have dependency between the risks of a dependent competitive risks model.

Keywords: C-Vines, D-Vines, Risk pooling method, Copula Graphic.

Contenido

Agradecimientos	vii
Resumen	ix
1. Introducción	1
2. Riesgos Competitivos	5
2.1. Conceptos Fundamentales	5
2.2. Problema de Identificabilidad	6
2.3. Estimación de la Función de Supervivencia bajo Independencia	7
3. Cópulas	9
3.1. Cópulas Arquimedianas	9
3.2. Una clase de cópulas simétricas	10
3.3. Propiedades de $C(x, y)$	11
3.3.1. La cópula Frank	13
3.3.2. τ de Kendall	13
3.3.3. La cópula Frank trivariada	14
4. Estimador Cópula Gráfico y Extensión	15
4.1. Estimador Cópula Gráfico	15
4.2. Construcción del Estimador Cópula Gráfico	16
4.3. Extensión del Estimador Cópula Gráfico: Método de Combinación de Riesgos	18
5. La dependencia modelada con C-Vine y D-Vine cópulas	23
5.1. Metodología de los vines	24
5.2. Simulación de un C-Vine o un D-Vine	26
6. Estudio de simulación	27
6.1. Metodología y Escenarios de simulación	27
6.2. Resultados	29
6.3. Escenario 1 de dependencia	29
6.4. Escenario 2 de dependencia	40
6.5. Escenario 3 de dependencia	52

7. Conclusiones y Trabajos Futuros	63
A. Programa en R del Método Combinación de Riesgo y C-DVines	65

Lista de Tablas

6-1. Estructuras de dependencia del modelo de riesgos competitivos dependientes trivariado	27
6-2. Parámetros de la distribución Weibull y tamaños de muestra	28
6-3. Porcentaje de falla para el modelo Weibull con tamaños de muestra $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$ y C-Vines.	29
6-4. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$	30
6-5. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$	31
6-6. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$	32
6-7. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$	33
6-8. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$	35
6-9. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$	36
6-10. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$	37
6-11. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$	38

6-12. Eficiencia relativa para el modelo Weibull con tamaños muestrales $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$ y C-Vines	38
6-13. Porcentaje de falla para el modelo Weibull con tamaños de muestra $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.9$ y C-Vines.	40
6-14. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.9$	41
6-15. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.9$	42
6-16. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.9$	42
6-17. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.9$	42
6-18. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.9$	46
6-19. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.9$	46
6-20. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.9$	47
6-21. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.9$	48
6-22. Eficiencia relativa para el modelo Weibull con tamaños muestrales $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.9$ y C-Vines	49
6-23. Porcentaje de falla para el modelo Weibull con tamaños de muestra $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.2, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.8$ y C-Vines.	52
6-24. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.2, \tau_{13} = 0.5, \tau_{23 1} = 0.8$	53

6-25. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.2, \tau_{13} = 0.5, \tau_{23 1} = 0.8$	54
6-26. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.2, \tau_{13} = 0.5, \tau_{23 1} = 0.8$	54
6-27. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.2, \tau_{13} = 0.5, \tau_{23 1} = 0.8$	56
6-28. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.2, \tau_{13} = 0.5, \tau_{23 1} = 0.8$	57
6-29. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.2, \tau_{13} = 0.5, \tau_{23 1} = 0.8$	58
6-30. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.2, \tau_{13} = 0.5, \tau_{23 1} = 0.8$	59
6-31. Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.2, \tau_{13} = 0.5, \tau_{23 1} = 0.8$	60
6-32. Eficiencia relativa para el modelo Weibull con tamaños muestrales $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.2, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.8$ y C-Vines	61

Lista de Gráficas

5-1.	Representación gráfica de los C-DVines: a) C-Vines b) D-Vines	26
6-1.	Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.5, \tau_{13} = 0.5, \tau_{23 1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	31
6-2.	Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	32
6-3.	Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	33
6-4.	Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	34
6-5.	Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	35
6-6.	Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	36
6-7.	Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	37
6-8.	Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 100$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	39
6-9.	Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, n = 50$ y $\tau_{12} = 0.5, \tau_{13} = 0.5$ y $\tau_{23 1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	41

- 6-10.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 43
- 6-11.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 44
- 6-12.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 45
- 6-13.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$. 47
- 6-14.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$. 48
- 6-15.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$. 49
- 6-16.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 50
- 6-17.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 53
- 6-18.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 55
- 6-19.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 55
- 6-20.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 56
- 6-21.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 57
- 6-22.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 58

-
- 6-23.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 59
- 6-24.**Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ 60

1. Introducción

Un sistema con tres modos de falla puede ser modelado usando un modelo de riesgos competitivos, el sistema fallará cuando se presente cualquiera de los tres modos de falla, es decir, el tiempo tiempo de falla observado del sistema es el mínimo de los posibles tiempos de falla de los respectivos modos de falla, además un modo de falla censura a los otros dos modos. La metodología tradicional asume independencia entre los tiempos de falla para la estimación de la función de supervivencia marginal, pero comúnmente este supuesto no se cumple. Para riesgos competitivos la estructura de los datos obtenidos no permite estimar la función de supervivencia ni las densidades marginales, esto es el llamado problema de identificabilidad. En el caso de dos riesgos competitivos, el estimador cópula gráfico permite estimar de la función de supervivencia marginal bajo una cópula conocida y la dependencia entre los dos riesgos que compiten. Cuando se tienen modelos de riesgos competitivos con más de dos riesgos, se considera el método denominado *combinación de riesgos*, el cual es una extensión del estimador cópula gráfico y es usado para la estimación de la función de supervivencia. En este trabajo se analiza la función de supervivencia para un modelo con tres riesgos bajo distintas dependencias entre los modos de falla que compiten, por medio del método antes mencionado. Para encontrar la estimación de la función de sobrevivencia se supone independencia entre los modos de falla. En teoría de probabilidad las funciones llamadas cópulas son en realidad funciones de distribución que representan las relaciones de dependencia entre las variables aleatorias. Las funciones cópula bidimensionales son funciones bivariadas que unen o bien copulan dos funciones de distribución univariadas para construir funciones de distribución bivariadas. La cópula representa una forma paramétrica conveniente para modelar la estructura de dependencia en distribuciones conjuntas de variables aleatorias. Este concepto de cópula permite construir modelos que van más allá de los estándares en el análisis de dependencia entre variables estocásticas. La metodología de cópula es capaz de capturar relaciones no lineales y en particular permite relacionar eventos extremos que ocurren en la naturaleza, Nelsen (2006).

En el caso no paramétrico, la función de sobrevivencia conjunta en un problema con múltiples modos de falla, no pueden ser completamente identificada en la situación usual cuando sólo se conocen los tiempos mínimos de falla, Tsiatis (1975). Incluso en el caso paramétrico los datos pueden contener poca información sobre el coeficiente de asociación entre las variables y resulta necesario hacer algunos supuestos al respecto, Meeker et al. (2009). Para el caso no paramétrico el estimador cópula gráfico de Zheng & Klein (1995), resuelve el problema bajo el supuesto del conocimiento de la forma funcional de la cópula

que proporciona la estructura de asociación entre los riesgos competitivos. Es de anotar que bajo independencia el problema de identificabilidad no existe. Para la estimación de la función de supervivencia Lo & Wilke (2010) proponen el método denominado combinación de riesgos para un modelo con más de dos riesgos competitivos cuando la cópula es Arquimediana. En el método la variable combinada es sólo creada conceptualmente para verificar lo relacionado con la identificación. Schweizer & Sklar (1983) realizaron estudio de dependencia entre los riesgos, el cual es un procedimiento para una iteración serial de cópulas Arquimedianas las cuales tienen la misma estructura de dependencia, además estas cópulas que utilizaron para ese estudio satisfacen las propiedades de simetría y asociatividad, las cuales son requeridas para que el método de combinación de riesgos funcione en forma adecuada. Lo & Wilke (2010) realizaron una aplicación del método de combinación de riesgos en un estudio de desempleo en Alemania, donde consideraron tres riesgos competitivos con igual dependencia, en este trabajo se estudia el caso en el cual los tres riesgos que compiten presentan diferentes dependencias. Paz-Sabogal. et al. (2014) estudiaron una extensión del estimador cópula gráfico para un modelo con más de dos riesgos competitivos e ilustraron la metodología con datos de la confiabilidad de interruptores tipo *FL245* en Interconexión Eléctrica S.A. E.S.P. (ISA) y mostraron la utilidad de la temática en confiabilidad industrial.

En este trabajo, para estudiar distintas dependencias entre los riesgos, se usa la técnica de los C-D Vines cópulas, los cuales trabajan con una variedad de cópulas bivariadas, dichas cópulas se pueden seleccionar de manera independiente y permiten tener una amplia gama de posibilidades para la caracterización de la dependencia de los riesgos que compiten. Se estudiaron tres casos de dependencia. En el primer caso dos de los tres riesgos tienen igual dependencia y el riesgo restante es independiente a los anteriores. En el segundo caso se tomaron dos riesgos con igual dependencia y el otro riesgo con una dependencia alta. En el tercer caso se tomaron los tres riesgos con distintas dependencias.

A continuación presentamos la forma como se desarrollo este trabajo:

En el capítulo 2 se definen conceptos fundamentales usados en la teoría de riesgos competitivos, el problema de identificabilidad y para finalizar se ilustra la forma como se estima la función de sobrevivencia bajo el supuesto de independencia. Seguidamente en el capítulo 3 se trata todo lo referente al concepto de cópula, se muestran cópulas simétricas, algunas propiedades de las cópulas, se presenta la cópula Frank, la cuales se usan en este trabajo y finalmente se ilustran los grados de asociación conocido como τ de Kendall.

En el capítulo 4 se muestra el estimador cópula gráfico y la extensión llamada método de combinación de riesgos, por medio del cual se estima la función de sobrevivencia para el tiempo mínimo asumiendo independencia entre los modos de falla del modelo de riesgos competitivos trivariado dependientes. En el capítulo 5, se presentan todo lo referente a la dependencia modelada con C-Vines y D-Vines cópulas.

En el capítulo 6 se presentan los resultados obtenidos de la función de sobrevivencia para un modelo con distintas estructuras de dependencia, las cuales se generan usando simulación

Monte Carlos y la técnica C-Vines y D-Vines. Además las marginales del modelo están distribuidas Weibull. Se presentan gráficos asociados a los intervalos de confianza empíricos en los percentiles t_p , con $p = 0.05, 0.25, 0.50, 0.75, 0.95$, aproximados del 95% para los estimadores de la función de sobrevivencia para el tiempo mínimo asumiendo dependencia e independencia entre los riesgos que compiten, como un ingrediente adicional, se muestran tablas de eficiencia relativa, estas tablas permiten comparar puntualmente el desempeño de los dos estimadores presentados en este estudio.

Finalmente en el capítulo 7 se muestran las conclusiones del estudio y trabajos futuros.

2. Riesgos Competitivos

El tiempo de falla de un sistema con múltiples modos de falla puede ser modelado como un sistema en serie o un modelo de riesgos competitivos. El tiempo de falla observado es el mínimo de esos tiempos potenciales individuales, donde cada unidad tiene un tiempo potencial de falla, Pintilie (2006).

Los datos para un modelo en riesgos competitivos trivariado se presentan como una variable aleatoria bivariada de la forma (T, δ) , con T el tiempo mínimo de falla $T = \min(X, Y, Z)$, con tiempos de falla potenciales X, Y, Z y donde δ es una variable discreta, la cual toma valores de 0 si la observación es una censura, 1 cuando la falla se presenta por el primer modo de falla, 2 cuando la falla se presenta por el segundo modo de falla y 3 cuando la falla se presenta por el tercer modo de falla.

2.1. Conceptos Fundamentales

Pintilie (2006) presenta algunos conceptos fundamentales para el desarrollo del presente trabajo.

La función de *subdistribución* o función de incidencia acumulada (FIA) para el modo de falla i , $i = 1, 2, 3$, viene dada por:

$$G_i = P(T \leq t, \delta = i)$$

La función de distribución total es la probabilidad que un evento de cualquier tipo ocurra en o antes del tiempo t , es decir, es la suma de las FIA para todos los tipos de eventos;

$$G(t) = P(T \leq t) = \sum_{i=1}^3 P(T \leq t, \delta = i) = \sum_{i=1}^3 G_i(t).$$

En ausencia de riesgos competitivos la función de distribución total se encuentra en el intervalo $[0, 1]$. Luego en presencia de riesgos competitivos la función de subdistribución sólo puede tomar valores hasta $P(\delta = i)$ puesto que:

$$\lim_{t \rightarrow \infty} G_i(t) = P(\delta = i)$$

Por lo tanto, la función $G_i(t)$ no es una función de distribución propia. La función de *subsobrevivencia* es la probabilidad de que el modo de falla i no ocurra antes del tiempo t

$$\overline{G}_i = P(T > t, \delta = i)$$

La función de subdensidad para el modo de falla i y T continuo, se define de la siguiente manera

$$g_i(t) = \frac{dG_i(t)}{dt} = -\frac{d\overline{G}_i(t)}{dt}$$

2.2. Problema de Identificabilidad

Para el caso no paramétrico, la distribución conjunta en un problema con múltiples modos de falla no puede ser completamente identificada en la situación usual cuando sólo se conocen los tiempos mínimos de falla, para un problema de este tipo, la aproximación usada para especificar el modelo de riesgos competitivos es vía tiempos latentes. Para el caso trivariado se tienen tiempos de falla potenciales X , Y , y Z asociados a cada modo de falla, donde si el sistema falla a causa del primer modo de falla, el segundo o el tercero no puede ser observado. Así $T = \min(X, Y, Z)$ determina el tiempo de falla del sistema completo.

Dada la tripleta (X, Y, Z) con distribución multivariada:

$$F(x, y, z) = P(X \leq x, Y \leq y, Z \leq z)$$

y funciones de distribución marginales para cada modo de falla, los cuales tienen la siguiente forma $F_1(t) = P(X \leq t)$, $F_2(t) = P(Y \leq t)$ y $F_3(t) = P(Z \leq z)$. Uno de los intereses en riesgos competitivos es identificar las distribuciones marginales de las variables a partir de los datos de riesgos competitivos. En el caso donde los tiempos de falla son independientes, las distribuciones marginales son identificables, y corresponden a las funciones de distribución. Si este supuesto no se tiene, entonces a partir únicamente de los datos, en riesgos competitivos, no es posible identificar las distribuciones marginales y la distribución conjunta de los tiempos, existen muchas funciones de distribución conjunta diferentes que comparten las mismas funciones de subdistribución, lo cual es llamado el problema de identificabilidad, Tsiatis (1975).

Para este problema en particular existen varias soluciones:

1. Asumir independencia.
2. Asumir una estructura de dependencia conocida usando una cópula particular para la distribución conjunta de los modos de falla Zheng & Klein (1995) proponen el algoritmo cópula gráfico.
3. Especificar un modelo paramétrico.

2.3. Estimación de la Función de Supervivencia bajo Independencia

Para un sistema en serie con tres componentes independientes, la función de distribución para el tiempo mínimo T está dada por:

$$\begin{aligned}
 F(t) &= P(T \leq t) \\
 &= 1 - P(T > t) \\
 &= 1 - P(X > t)P(Y > t)P(Z > t) \\
 &= 1 - [1 - F_1(t)][1 - F_2(t)][1 - F_3(t)]
 \end{aligned}$$

donde (X, Y, Z) son los tiempos de vida asociados a cada una de las tres componentes, $T = \min(X, Y, Z)$ es el tiempo observado y $F_i, i = 1, 2, 3$, las respectivas funciones de distribución para los componentes. La función de supervivencia $S(t) = 1 - F(t)$, para el caso de modos de falla independientes es:

$$\begin{aligned}
 S(t) &= P(T > t) \\
 &= P(X > t, Y > t, Z > t) \\
 &= P(X > t)P(Y > t)P(Z > t) \\
 &= S_1(t)S_2(t)S_3(t).
 \end{aligned}$$

donde $S_i(t), i = 1, 2, 3$ son las funciones de supervivencia de las tres componentes. Para un sistema con tres componentes en serie y tiempos de falla dependientes, $F(t)$ tiene la siguiente forma:

$$F(t) = P(T \leq t) = 1 - P(T > t) = 1 - P(X > t, Y > t, Z > t)$$

Para este caso, la evaluación se debe hacer con respecto a la distribución conjunta de X, Y y Z , en la cual se debe incluir el parámetro de dependencia. La función de supervivencia para este caso es:

$$S(t) = 1 - F(t) = 1 - [1 - P(T > t)] = P(T > t) = P(X > t, Y > t, Z > t).$$

El estimador de la supervivencia del tiempo mínimo de falla de un sistema con tres modos de falla independientes que compiten $S(t)$ se denotará $\widehat{S}^*(t)$, Para obtener este estimador basta multiplicar las funciones de supervivencia estimadas bajo una distribución específica, donde para cada función marginal se estiman los parámetros, considerando los tiempos del otro modo de falla como tiempos de censura o censurados.

$$\widehat{S}^*(t) = \widehat{S}_1(t)\widehat{S}_2(t)\widehat{S}_3(t).$$

donde $\widehat{S}_1(t), \widehat{S}_2(t)$ y $\widehat{S}_3(t)$, corresponden a las funciones de supervivencia marginales estimadas para cada modo de falla.

3. Cópulas

Kurowicka & Cooke (2006) definen una cópula como una distribución multivariada con distribuciones marginales uniformes en $[0, 1]$. Las cópulas permiten representar de manera separada la estructura de dependencia y las distribuciones marginales de una distribución multivariada. La metodología de cópulas se ha convertido en una herramienta importante a la hora de modelar o generar datos multivariados.

Nelsen (2006) contiene una gran cantidad de información acerca del modelado de dependencias con cópulas. Combinando varias distribuciones marginales con diferentes cópulas, es posible modelar distribuciones multivariadas con una amplia gama de marginales y tipos de dependencia. A continuación se muestra el teorema fundamental para cópulas conocido como el teorema de Sklar.

Teorema 1 (Teorema de Sklar). *Sean Y_1, Y_2 variables aleatorias con función de distribución conjunta F , con marginales F_1 y F_2 respectivamente. Entonces existe una cópula C tal que satisface:*

$$F(y_1, y_2) = C(F_1(y_1), F_2(y_2))$$

Si F_1 y F_2 son continuas, entonces C es única. Inversamente, si C es una cópula y F_1, F_2 son funciones de distribución, entonces F es una función de distribución conjunta con marginales F_1 y F_2 .

El teorema de Sklar se puede extender al caso multivariado, para nuestro caso se utilizan cópulas bivariadas y trivariadas.

3.1. Cópulas Arquimedianas

Un importante caso particular de cópulas son las cópulas Arquimedianas, se caracterizan por la facilidad con que pueden ser construidas y por la gran variedad de estructuras de dependencia que pueden reproducir. En este trabajo se utilizó la cópula Frank debido a que en la implementación del algoritmo de Lo & Wilke (2010), ellos la utilizaron en un estudio de desempleo en Alemania y los resultados fueron óptimos, puesto que en el fondo se llega a una extensión del estimador cópula gráfico, propuesto por Zheng & Klein (1995).

3.2. Una clase de cópulas simétricas

Genest & MacKay (1986) considera una clase Φ de funciones $\phi : [0, 1] \rightarrow [0, \infty]$, la cual tiene dos derivadas continuas en $(0, 1)$ y satisface:

$$\phi(1) = 0, \quad \phi'(t) < 0, \quad \phi''(t) > 0 \quad (3.1)$$

para todo $0 < t < 1$. Las condiciones dadas en (3.1) son suficientes para garantizar que ϕ tiene inversa, es decir ϕ^{-1} y que también existen las dos primeras derivadas.

Por notación $\phi(0) = \infty$ si $\lim_{t \rightarrow 0^+} \phi(t) = \infty$. Miembros típicos de esta clase Φ incluyen:

1. $\phi(t) = -\log(t)$
2. $\phi(t) = (1-t)^\alpha$, con $\alpha > 1$
3. $\phi(t) = t^{-\alpha} - 1$, con $\alpha > 1$

Una caracterización de las cópulas arquimedianas, es que cada miembro ϕ de la clase Φ genera una función de distribución bivariada para el par (X, Y) de la siguiente manera:

$$\begin{aligned} C(x, y) &= \phi^{-1}[\phi(x) + \phi(y)], \quad \text{si } \phi(x) + \phi(y) \leq \phi(0) \\ &= 0, \quad \text{en otro caso} \end{aligned} \quad (3.2)$$

Si $\phi(0) = \infty$, entonces $C(x, y)$ es estrictamente positiva excepto cuando $x = 0$ o $y = 0$. Para encontrar la densidad $c(x, y)$ asociada (3.2), sea $\phi(C) = \phi(x) + \phi(y)$, diferenciando $C(x, y)$ con respecto a x y luego con respecto a y , se sigue que:

$$\begin{aligned} C(x, y) &= \phi^{-1}(\phi(x) + \phi(y)) \\ \phi(C(x, y)) &= \phi(x) + \phi(y) \end{aligned}$$

derivando con respecto a x , se tiene:

$$\phi'(C) \frac{\partial C}{\partial x} = \phi'(x)$$

y luego derivando con respecto a y , se sigue que:

$$\phi''(C) \frac{\partial C}{\partial x} \frac{\partial C}{\partial y} + \phi'(C) \frac{\partial^2 C}{\partial x \partial y} = 0 \quad (3.3)$$

Luego

$$\begin{aligned}
\phi'(C) \frac{\partial^2 C}{\partial x \partial y} &= -\phi''(C) \frac{\partial C}{\partial x} \frac{\partial C}{\partial y} \\
\phi'(C) c(x, y) &= -\phi''(C) \frac{\partial C}{\partial x} \frac{\partial C}{\partial y} \\
c(x, y) &= \frac{-\phi''(C) \frac{\partial C}{\partial x} \frac{\partial C}{\partial y}}{\phi'(C)} \\
c(x, y) &= \frac{-\phi''(C) \frac{\phi'(x)}{\phi'(C)} \frac{\phi'(y)}{\phi'(C)}}{\phi'(C)} \\
c(x, y) &= \frac{-\phi''(C) \phi'(x) \phi'(y)}{[\phi'(C)]^3}
\end{aligned}$$

De las condiciones de ϕ dadas en (3.1), se tiene que $c(x, y) > 0$ para todo (x, y) tal que $\phi(x) + \phi(y) \leq \phi(0)$. En general, las derivadas no existen en la frontera $\phi(x) + \phi(y) = \phi(0)$.

3.3. Propiedades de $C(x, y)$

a) La distribución es simétrica en x y y , esto es $C(x, y) = C(y, x)$.

En efecto:

$$\begin{aligned}
C(x, y) &= \phi^{-1}[\phi(x) + \phi(y)] \\
&= \phi^{-1}[\phi(y) + \phi(x)] \\
&= C(y, x)
\end{aligned}$$

b) La cópula C es asociativa, es decir; $C(C(x, y), z) = C(x, C(y, z))$.

$$\begin{aligned}
C(C(x, y), z) &= \phi^{-1}\{\phi[C(x, y)] + \phi(z)\} \\
&= \phi^{-1}\{\phi[\phi^{-1}(\phi(x) + \phi(y))] + \phi(z)\} \\
&= \phi^{-1}\{\phi(x) + \phi(y) + \phi(z)\} \\
&= \phi^{-1}\{\phi(x) + \phi[C(y, z)]\} \\
&= C(x, C(y, z))
\end{aligned}$$

c) Las distribuciones marginales de X e Y son uniformes en el intervalo $[0, 1]$. Por ejemplo $C(x, 1) = x$ para todo $0 \leq x \leq 1$. En otras palabras, las distribuciones de la forma (3.2) son cópulas. Por definición se tiene que:

$$C(x, 1) = \phi^{-1}[\phi(x) + \phi(1)]$$

pero como $\phi(1) = 0$, entonces tenemos:

$$\begin{aligned}
C(x, 1) &= \phi^{-1}[\phi(x) + 0] \\
&= \phi^{-1}[\phi(x)] \\
&= x \\
\therefore C(x, 1) &= x
\end{aligned}$$

Análogamente;

$$\begin{aligned}
C(1, y) &= \phi^{-1}[\phi(1) + \phi(y)] \\
&= \phi^{-1}[0 + \phi(y)] \\
&= \phi^{-1}[\phi(y)] \\
&= y \\
\therefore C(1, y) &= y
\end{aligned}$$

Luego, las marginales son uniformes en el intervalo $[0, 1]$.

d) El soporte de la distribución es $\{(x, y) : \phi(x) + \phi(y) \leq \phi(0)\}$, el cual es un cuadrado unitario si $\phi(0) = \infty$. Por definición, el soporte de una función es donde la función es estrictamente positiva, así si $\phi(0) = \infty$, entonces $C(x, y)$ es estrictamente positiva.

e) $P\{X > x, Y > y\} = C(x, y) - x - y + 1$

Tenemos que:

$$\begin{aligned}
P\{X > x, Y > y\} &= \int_x^1 \int_y^1 c(t, w) dt dw \\
&= \int_x^1 \left[1 - \int_0^y c(t, w) dt \right] dw \\
&= \int_x^1 dw - \int_x^1 \int_0^y c(t, w) dt dw \\
&= \int_x^1 dw - \int_0^y \left[\int_x^1 c(t, w) dw \right] dt \\
&= 1 - x - \int_0^y \left[1 - \int_0^x c(t, w) dw \right] dt \\
&= 1 - x - \int_0^y dt + \int_0^y \int_0^x c(t, w) dw dt \\
&= 1 - x - y + C(x, y).
\end{aligned}$$

f) Para 3 riesgos competitivos, es decir $N = 3$, la función de sobrevivencia conjunta está definida por:

$$\bar{C}(u_1, u_2, u_3) = 1 - u_1 - u_2 - u_3 + C(u_1, u_2) + C(u_1, u_3) + C(u_2, u_3) - C(u_1, u_2, u_3)$$

g) $\max\{0, x + y - 1\} \leq C(x, y) \leq \min\{x, y\}$

Sea (x, y) un punto arbitrario en el dominio de $C(x, y)$. Como $C(x, y) \leq C(x, 1) = x$ y $C(x, y) \leq C(1, y) = y$, entonces

$$C(x, y) = \min\{x, y\}$$

Ahora, como la función de sobrevivencia es siempre positiva o cero y además por la propiedad e) , se tiene:

$$C(x, y) - x - y + 1 \geq 0$$

$$C(x, y) \geq x + y - 1$$

$$C(x, y) \geq \begin{cases} x + y - 1 \\ 0 \end{cases}$$

entonces

$$C(x, y) \geq \max\{x + y - 1, 0\}$$

Los límites inferior y superior en la desigualdad presentada son usualmente llamados límites de Fréchet.

3.3.1. La cópula Frank

$$C_{\theta}(x, y) = -\frac{1}{\theta} \ln \left[1 + \frac{\{exp(-\theta x) - 1\} \{exp(-\theta y) - 1\}}{\{exp(-\theta) - 1\}} \right], \theta \neq 0 \quad (3-1)$$

La cópula Frank tiene grandes propiedades, captura el rango completo de dependencias, incluye las cópulas de cota superior de Fréchet cuando $\theta \rightarrow -\infty$, de cota inferior de Fréchet cuando $\theta \rightarrow \infty$ y de independencia cuando $\theta \rightarrow 0$. La cópula Frank proporciona cantidades cerradas y permite mayor facilidad al momento de programar (Escarela & Carriere (2003)).

3.3.2. τ de Kendall

El τ de Kendall, mide el grado de asociación entre dos variables aleatorias, en particular para las marginales dos a dos de un modelo de riesgos trivariado dependiente generado en nuestro caso. Está dado como sigue:

$$\tau_{\theta} = 1 - \frac{4}{\theta} [1 - D_1(\theta)]$$

donde $D_k(x)$ es la función Debye, la cual está definida para cualquier entero k por:

$$D_k(x) = \frac{k}{x^k} \int_0^x \frac{t^k}{exp(t)-1} dt$$

El τ de Kendall, cumple con la siguientes propiedades $\lim_{\theta \rightarrow -\infty} \tau_{\theta} = -1$, $\lim_{\theta \rightarrow \infty} \tau_{\theta} = 1$ y $\lim_{\theta \rightarrow 0} \tau_{\theta} = 0$ (independiente).

3.3.3. La cópula Frank trivariada

$$C_{\theta_{123}}^3(\mathbf{X}) = -\frac{1}{\theta_1} \left[1 + \frac{\{exp(-\theta_1 x_1) - 1\} \{exp(-\theta_2 x_2) - 1\} \{exp(-\theta_3 x_3) - 1\}}{\{exp(-\theta_1) - 1\}^2} \right]$$

donde θ_1 es el parámetro de la cópula C_{12} , θ_2 es el parámetro de la cópula C_{13} y θ_3 es el parámetro de la cópula $C_{23|1}$, donde la cópula $C_{23|1}$ es una cópula condicional.

4. Estimador Cópula Gráfico y Extensión

En este capítulo ilustraremos la manera como surge el estimador cópula gráfico y cual fue el problema que Zheng & Klein (1995) solucionaron. Se presentarán detalles del estimador y la extensión para la construcción del método combinación de riesgos de Lo & Wilke (2010) y Paz-Sabogal. et al. (2014) el cual es uno de los objetivos principales de este trabajo, es decir, cuando los riesgos presentan distintas dependencias entre los modos de falla que participan en un modelo de riesgos competitivos dependientes trivariado.

4.1. Estimador Cópula Gráfico

Zheng & Klein (1995) aplican cópulas Arquimedianas en modelos de riesgos en competencia con datos censurados dependientes, proponiendo un estimador no paramétrico para las distribuciones marginales llamado estimador cópula gráfico, el cual da solución al problema de identificabilidad, dicho problema es resuelto con datos de riesgos competitivos y bajo el supuesto de que la cópula y su parámetro de asociación son conocidos de antemano.

En el marco de riesgos competitivos se observa $T_j = \min(X_j, Y_j)$ y $\delta_j = I_{(X_j \leq Y_j)}$ ($j = 1, \dots, n$). Suponiendo que $P(X_j = Y_j) = 0$. Haciendo uso de esta información se puede estimar directamente:

$$k(t) = P(X > t, Y > t), \quad p_1(t) = P(X \leq t, X < Y),$$

$$p_2(t) = P(Y \leq t, Y < X), \quad \text{para } 0 \leq t < \infty$$

donde $p_1(t)$ y $p_2(t)$ corresponden a las funciones de subdistribución.

Zheng & Klein (1995) afirman que bajo el supuesto de independencia entre X y Y , las distribuciones marginales son determinadas de manera única por esas probabilidades. Muestran en general, que si la cópula de (X, Y) es conocida, entonces las distribuciones marginales son determinadas de manera única por los datos de riesgos competitivos.

4.2. Construcción del Estimador Cópula Gráfico

El estimador cópula gráfico está basado en el algoritmo de búsquedas de raíces, el cual trabaja dividiendo el intervalo a la mitad y seleccionando el subintervalo que contiene la raíz de la función de variable real. Este método es sencillo y fácil para resolver ecuaciones, se basa en el teorema del valor intermedio, el cual establece que toda función continua f en el intervalo $[a, b]$ y $f(a)$ y $f(b)$ tienen signos opuestos, entonces f tiene al menos una raíz en el intervalo (a, b) .

Seguidamente se ilustran los detalles de la construcción del estimador cópula gráfico. Para riesgos competitivos, supongamos una muestra de tamaño n de tiempos de vida, con $T_j = \min(X_j, Y_j)$. Si $F_1(t)$ y $F_2(t)$ son las distribuciones acumuladas marginales de X y Y , para todo t , entonces para cualquier cópula se tiene:

$$\mu_c(A_t) = P(X > t, Y > t) = k(t) \quad (4.1)$$

$$\mu_c(B_t) = P(X \leq t, X < Y) = p_1(t) \quad (4.2)$$

donde

$$A_t = \{(x, y) : F_1(t) < x \leq 1, F_2(t) < y \leq 1\},$$

$$B_t = \{(x, y) : 0 < x \leq F_1(t), F_2 F_1^{-1}(t) < y \leq 1\}.$$

Zheng & Klein (1995) demuestran que las relaciones (4.1) y (4.2) determinan de manera única a $F_1(t)$ y $F_2(t)$. Por lo tanto se pueden encontrar estimadores para $\widehat{F}_1(t)$ y $\widehat{F}_2(t)$ los cuales conservan las propiedades sobre una rejilla seleccionada de m puntos distintos de falla o censura $0 < t_1 < t_2 < \dots < \max\{T_j, j = 1, \dots, n\}$.

A continuación se definen los siguientes conjuntos

$$\widehat{A}_t = \{(x, y) : \widehat{F}_1(t) < x \leq 1, \widehat{F}_2(t) < y \leq 1\}$$

$$\widehat{B}_t = \{(x, y) : 0 < x \leq \widehat{F}_1(t), \widehat{F}_2 \widehat{F}_1^{-1}(t) < y \leq 1\}$$

Adicionalmente, las estimaciones empíricas de $P(X > t, Y > t)$ y $P(X \leq t, X < Y)$, vienen dadas de la siguiente manera

$$\widehat{P}(X > t, Y > t) = \frac{1}{n} \sum_{i=1}^n I_{\{T_i > t\}}$$

$$\widehat{P}(X \leq t, X < Y) = \frac{1}{n} \sum_{i=1}^n I_{\{T_i \leq t, \delta=1\}}$$

Se encuentra a $\widehat{F}_1(t_i)$ como la raíz de (4.4), basado en la definición de $\widehat{F}_2(t_i)$, como una función de $\widehat{F}_1(t_i)$ solucionando la ecuación (4.3). Sean \widehat{F}_1 y \widehat{F}_2 funciones escalonadas en cada intervalo $(t_i, t_i + 1)$.

$$\mu_c(\widehat{A}_t) - \widehat{P}(X > t_i, Y > t_i) = 0 \quad (4.3)$$

$$\mu_c(\widehat{B}_t) - \widehat{P}(X \leq t_i, X < Y) = 0 \quad (4.4)$$

Sean t_1, t_2, \dots, t_m tiempos distintos de falla o censura, tomando estos tiempos como la rejilla y usando la función escalonada en cada intervalo (t_{i-1}, t_i) el estimador es más fácil de calcular. En particular, si $\delta_i = 1$, entonces $\widehat{F}_2(t_i) = \widehat{F}_2(t_{i-1})$, es decir, la función calculada en t_i es igual a la función calculada en t_{i-1} debido a que la falla se presenta por la causa X, pero si $\delta_i = 2$, entonces $\widehat{F}_1(t_i) = \widehat{F}_1(t_{i-1})$. Como punto inicial, sea $t_0 = 0$ y $\widehat{F}_1(t_0) = \widehat{F}_2(t_0) = 0$

$$\delta_i = \begin{cases} 1, & \text{si } X < Y; \text{ es decir falla por la causa } X \\ 2, & \text{si } Y < X; \text{ es decir falla por la causa } Y \\ 0, & \text{no falla por ninguna de las dos causas} \end{cases}$$

Para $\delta_i = 1$

$$\mu_c(\widehat{A}_{t_i}) = 1 - \widehat{F}_1(t_i) - \widehat{F}_2(t_{i-1}) + C\{\widehat{F}_1(t_i), \widehat{F}_2(t_{i-1})\} = \widehat{P}(X > t_i, Y > t_i) \quad (4.5)$$

Para $\delta_i = 2$

$$\mu_c(\widehat{A}_{t_i}) = 1 - \widehat{F}_1(t_{i-1}) - \widehat{F}_2(t_i) + C\{\widehat{F}_1(t_{i-1}), \widehat{F}_2(t_i)\} = \widehat{P}(X > t_i, Y > t_i) \quad (4.6)$$

donde $C\{\widehat{F}_1, \widehat{F}_2\}$ es la cópula que relaciona las marginales estimadas \widehat{F}_1 y \widehat{F}_2 . Las ecuaciones (4.5) y (4.6) definen la sobrevivencia conjunta bivariada presentada por Genest & MacKay (1986).

$$P(X > x, Y > y) = C(x, y) - x - y + 1$$

Pasos para la construcción del Estimador Cópula Gráfico

1. Para $i = 1$, dado un punto inicial $\widehat{F}_1(t_1)$, el cual denotaremos $\widehat{F}_1^{(1)}(t_1)$, luego se encontrará $\widehat{F}_2^{(1)}(t_1)$ resolviendo (4.3).
2. Si el punto $(\widehat{F}_1^{(1)}(t_1), \widehat{F}_2^{(1)}(t_1))$ satisface (4.4), ir al paso 4. En caso contrario use (4.4) para decidir si el siguiente valor, $\widehat{F}_1^{(2)}(t_1)$ es mayor o menor que $\widehat{F}_1^{(1)}(t_1)$. Use el punto medio del intervalo $(\widehat{F}_1^{(1)}(t_1), 1)$ ó $(0, \widehat{F}_1^{(1)}(t_1))$ como el valor de $\widehat{F}_1^{(2)}(t_1)$, respectivamente.

Observación: En la ecuación (4.4) se tiene que $\mu_c(\widehat{B}_t) - \widehat{P}(X \leq t_i, X < Y) = 0$, donde

$$\begin{aligned} \mu_c(\widehat{B}_{t_i}) &= \widehat{F}_1(t_i) - \int_{t_i}^{\infty} C_{F_1}(\widehat{F}_1, \widehat{F}_2) d\widehat{F}_1 dt. \\ &= \widehat{F}_1(t_i) - \int_{t_i}^{\infty} C_{F_1}(\widehat{F}_1, \widehat{F}_2) f_1 dt. \end{aligned}$$

donde $C_{F_1}(\widehat{F}_1, \widehat{F}_2)$ es la derivada parcial de la cópula respecto a F_1 .

3. Repita los pasos 1 y 2 usando la estimación actual de $\widehat{F}_1(t_1)$ y $\widehat{F}_2(t_1)$. Después de j pasos el nuevo valor de $\widehat{F}_1(t_1)$ es el punto medio del intervalo $(a_i, \widehat{F}_1^{(i)}(t_1))$ o el punto medio del intervalo $(\widehat{F}_1^{(i)}(t_1), b_i)$, donde a_i y b_i estan dados por:

$$a_i = \max \left\{ \widehat{F}_1^{(k)}(t_1) : k < i, \widehat{F}_1^{(k)}(t_1) < \widehat{F}_1^{(i)}(t_1) \right\}$$

$$b_i = \min \left\{ \widehat{F}_1^{(k)}(t_1) : k < i, \widehat{F}_1^{(k)}(t_1) > \widehat{F}_1^{(i)}(t_1) \right\}$$

El proceso continúa hasta encontrar $\widehat{F}_1(t_i)$ y $\widehat{F}_2(t_i)$ que satisfacen (4.3) y (4.4). La convergencia del algoritmo es garantizada.

4. Repita los pasos 1 – 3 para $i = 2, \dots, m$. Para $t > \max T_j$, se define $\widehat{F}_1(t)$ y $\widehat{F}_2(t)$ que serán $\widehat{F}_1(t_m)$ y $\widehat{F}_2(t_m)$.

4.3. Extensión del Estimador Cópula Gráfico: Método de Combinación de Riesgos

Lo & Wilke (2010), presentaron una extensión del estimador cópula gráfico propuesto por Zheng & Klein (1995), la propuesta es llamada el método de combinación de riesgos,

la cual aplica para modelos con más de dos riesgos competitivos cuando la cópula es Arquimediana. A continuación se presentan detalles del método cuando se tienen tres riesgos en competencia.

Sean $(T_1, T_2, T_3) \in R_+^3$ tiempos de falla latentes $j = 1, 2, 3$, de un modelo de riesgos competitivos 3 – *dimensional*. En el caso trivariado se tienen tres tiempos potenciales X , Y y Z , asociados a tres modos de falla, de manera que cuando la falla se presenta por el primer modo, los otros dos modos de falla no pueden ser observados. En este sentido, se dice que los tiempos son latentes, es decir, pueden ser dependientes uno de otro.

Sea $S_j : R_+ \rightarrow [0, 1]$ una función de sobrevivencia marginal desconocida, continua y estrictamente decreciente de T_j

$$S_j = P(T_j > t_j) = r_j, j = 1, 2, \dots, J \tag{4.7}$$

donde r_j es considerada como la posición relativa o rango de t_j , y es una realización de una variable aleatoria R_j que se distribuye uniforme $[0, 1]$.

Se asume que la estructura básica de dependencia de T_j es generada por una cópula conocida, la cual es una distribución conjunta de los rangos de la variable de tiempo de falla. La J -cópula , $C^J : [0, 1]^J \rightarrow [0, 1]$, está definida de la siguiente manera:

$$C^J(r_1, \dots, r_J) = P(R_1 \leq r_1, \dots, R_J \leq r_J) \tag{4.8}$$

donde la cópula determina la estructura básica de dependencia entre las variables T_j .

La función de sobrevivencia conjunta $S(t_1, \dots, t_J) = P(T_1 > t_1, \dots, T_J > t_J)$ es determinada sustituyendo la ecuación (4.7) en la ecuación (4.8).

$$\begin{aligned} C^J \{S_1(t_1), \dots, S_J(t_J)\} &= P \{S_1(T_1) \leq S_1(t_1), \dots, S_J(T_J) \leq S_J(t_J)\} \\ &= S(t_1, \dots, t_J). \end{aligned}$$

La función de sobrevivencia conjunta $S(t) = S(t, \dots, t)$ y la curva de incidencia acumulada de causa específica $Q_j(t_j)$, están dadas de la siguiente manera:

$$\begin{aligned} S(t) &= P(T_1 > t, \dots, T_J > t) \\ &= \int_0^{r_J(t)} \dots \int_0^{r_1(t)} dC^J(r_1, \dots, r_J) \end{aligned} \tag{4.9}$$

$$\begin{aligned} Q_j(t_j) &= P(T_j \leq t_j, T_j < \min_{i \neq j} \{T_i\}) \\ &= \int_0^{\zeta_J(r_j)} \dots \int_0^{\zeta_{j+1}(r_j)} \int_{r_j(t_j)}^1 \int_0^{\zeta_{j-1}(r_j)} \int_0^{\zeta_1(r_j)} dC^J(r_1, \dots, r_J) \end{aligned} \tag{4.10}$$

donde $\zeta_k(r_j) = S_k \{S_j^{-1}\}$ para todo $k \neq j$, es de anotar que la inversa existe puesto que S_j es continua y estrictamente decreciente.

El enfoque de la propuesta de Lo & Wilke (2010) busca determinar las funciones de supervivencia marginales $\{S_1(t_1), \dots, S_J(t_J)\}$ usando $\{S(t), Q_1(t_1), \dots, Q_J(t_J), C^J\}$. La estrategia es estimar $S(t)$ y $Q_j(t_j)$, cuando la función de la cópula verdadera C^J es conocida apriori. Si T_i es independiente de T_j para todo $i \neq j$, cualquier cópula se reduce al producto cópula

$$C^J(r_1, \dots, r_J) = \prod_{j=1}^J G_j(U_j) \quad (4.11)$$

El estimador Kaplan - Meier requiere independencia.

La propuesta numérica de Zheng & Klein (1995) para un modelo con dos riesgos puede ser aplicada directamente para dar solución a las ecuaciones (4.9) y (4.10). Lo & Wilke (2010) presentan una extensión para el caso en que más de dos riesgos compiten ($J > 2$), donde $S_j(t)$ puede ser calculada combinando todos los riesgos con ($k \neq j$), dicha extensión es llamada **método de combinación de riesgos**. Seguidamente se ilustra para el caso de 3 riesgos en competencia ($J = 3$). El método consiste en combinar las variables T_2 y T_3 para formar una nueva variable $T_{23} = \min(T_2, T_3)$. La función de supervivencia marginal de T_{23} se define como $S_{23}(t) = P(T_{23} > t)$.

Si existe una cópula de supervivencia entre las variables T_1 y T_{23}

$$C_1^2 \{S_1(t_1), S_{23}(t_{23})\} = S(t) \quad (4.12)$$

la aproximación de Zheng & Klein (1995) puede aplicarse directamente para calcular $S_1(t_1)$ y $S_{23}(t_{23})$. Lo anterior se consigue resolviendo las dos ecuaciones

$$S(t) = \int_0^{r_{23}(t)} \int_0^{r_1(t)} dC_1^2(S_1, S_{23}) \quad (4.13)$$

La ecuación anterior se cumple, dado que:

$$S(t) = P(T_1 > t, T_{23} > t)$$

por la propiedad de las cópulas definidas en el ítem e) de la sección 3.3, se tiene que

$$S(t) = 1 - F_1(t) - F_{23}(t) + \widehat{C}_1^2(F_1(t), F_{23}(t))$$

Ahora, por la relación de la función de supervivencia y la función de distribución acumulada, se tiene

$$S(t) = S_1(t) + S_{23}(t) - 1 + \widehat{C}_1^2(1 - S_1(t), 1 - S_{23}(t))$$

entonces,

$$\begin{aligned}
S(t) &= C_1^2(S_1(t), S_{23}(t)) \\
&= P[S_1(T_1) < S_1(t), S_{23}(T_{23}) < S_{23}(t)] \\
&= \int_0^{S_{23}(t)} \int_0^{S_1(t)} dC_1^2(S_1, S_{23})
\end{aligned}$$

donde \widehat{C} hace referencia a la cópula evaluada en las distribuciones acumuladas y C corresponde a la cópula de sobrevivencia.

La ecuación (4.13) se cumple puesto que

$$\begin{aligned}
\int_0^{r_{23}(t)} \int_0^{r_1(t)} dC_1^2(S_1, S_{23}) &= P\{S_1 \leq r_1(t), S_{23} \leq r_{23}(t)\} \\
&= P(T_1 > t, T_{23} > t) \\
&= P(T_1 > t, T_2 > t, T_3 > t) \\
&= S(t).
\end{aligned}$$

note que $T_{23} = \min(T_2, T_3)$. Se tiene que $T_{23} > t$ si y sólo si $T_2 > t$ y $T_3 > t$.

$$Q_1 = \int_0^{\zeta_{23}(r_1)} \int_{r_1(t)}^1 dC_1^2(S_1, S_{23}) \quad (4.14)$$

donde $r_{23}(t) = S_{23}(t)$ y $\zeta_{23}(r_1) = S_{23}\{S_1^{-1}\}$, en efecto:

$$\begin{aligned}
\int_0^{\zeta_{23}(r_1)} \int_{r_1(t)}^1 dC_1^2(S_1, S_{23}) &= P(S_1(t) > r_1(t), S_{23} \leq \zeta_{23}(r_1)) \\
&= P(T_1 \leq t, T_{23} > T_1) \\
&= P(T_1 \leq t, T_2 > T_1, T_3 > T_1) \\
&= Q_1(t).
\end{aligned}$$

Por otra parte, no es necesario crear la variable combinada $T_{23} = \min(T_2, T_3)$ para la implementación del método de combinación de riesgo, puesto que $S_1(t_1)$ y $S_{23}(t_{23})$ pueden ser identificadas a partir de $S(t)$, $Q_1(t)$ y $C_1^2\{S_1(t), S_{23}(t_{23})\}$ sin la construcción de $Q_{23}(t)$. La variable combinada es sólo creada conceptualmente para probar la identificabilidad. $S_2(t)$ y $S_3(t)$ pueden ser obtenidas a partir de $S(t)$, $Q_2(t)$, $Q_3(t)$, $C_2^2 = \{S_2(T_2), S_{13}(T_{13})\}$ y $C_3^2 = \{S_3(T_3), S_{12}(T_{12})\}$. El problema inicial se redujo de tener 3 variables a tener 2 variables, luego se puede utilizar el estimador cópula gráfico.

Para que el método descrito funcione se debe garantizar que: para una 3-cópula conocida $C^3(r_1, r_2, r_3)$, existan todas las 2-cópula, es decir, $C_{ij}(r_i, r_j) = S_{ij}(t)$ tal que

$$C^3(r_1, r_2, r_3) = C_1^2 \{r_1, C_{23}(r_2, r_3)\} = C_2^2 \{r_2, C_{13}(r_1, r_3)\} = C_3^2 \{r_3, C_{12}(r_1, r_2)\}$$

Nelsen (2006) muestra que en este caso C_{ij} es compatible con C^3 . Hay que tener en cuenta que algunas veces una 2-cópula es generalmente inconsistente con la 3-cópula no agrupada de T_1 , T_2 y T_3 .

Genest et al. (1995) muestra que la cópula $C_1^2 \{S_1(t_1), S_{23}(t_{23})\}$ en la ecuación (4.12) puede no existir. Por tanto, para la implementación de la metodología se utilizan cópulas Arquimedianas que satisfacen la propiedad de simetría y asociatividad, las cuales se mostraron en la sección (3.3). Teniendo en cuenta la restricción a esta clase de funciones cópula se deduce también que no importa el orden en que se realice el agrupamiento de los riesgos. A manera de ejemplo, se toma una 3-cópula de la clase Arquimediana como se indica en la ecuación (4.12). En efecto:

$$\begin{aligned} S(t) &= C_1^2(S_1(t_1), S_{23}(t_{23})) = C_2^2(S_2(t_2), S_{13}(t_{13})) = C_3^2(S_3(t_3), S_{12}(t_{12})) \\ &= C_1^2(S_1(t_1), S_{32}(t_{32})) = C_2^2(S_2(t_2), S_{31}(t_{31})) = C_3^2(S_3(t_3), S_{21}(t_{21})) \end{aligned}$$

De esta manera, a partir de $C_1^2(S_1(t_1), S_{23}(t_{23}))$ (o $C_1^2 \{S_1(t_1), S_{32}(t_{32})\}$), $S(t)$ y $Q_1(t)$ se identifica $S_1(t)$, a partir de $C_2^2(S_2(t_2), S_{13}(t_{13}))$ (o $C_2^2 \{S_2(t_2), S_{31}(t_{31})\}$), $S(t)$ y $Q_2(t)$ se identifica $S_2(t)$, a partir de $C_3^2(S_3(t_3), S_{12}(t_{12}))$ (o $C_3^2 \{S_3(t_3), S_{21}(t_{21})\}$), $S(t)$ y $Q_3(t)$ se identifica $S_3(t)$. Luego los resultados son independientes del orden de las variables. Por lo anterior, se puede utilizar la estrategia de Zheng & Klein (1995) a un modelo con más de dos riesgos, siempre que la cópula de la variable combinada sea compatible.

5. La dependencia modelada con C-Vine y D-Vine cópulas

El objetivo de éste capítulo es trabajar distintas dependencias entre los riesgos de un modelo de riesgos competitivos dependiente trivariado. Este tipo de dependencias se ilustran con la técnica basada en C-Vines y D-Vines cópulas.

Un tipo especial de *vines* son los C-Vines (*canonical vines*) y los D-Vines (*drawable vines*). Desde un punto de vista gráfico, tanto un C-Vine como un D-Vine están formados por una cascada de árboles y a cada uno de ellos le corresponde una manera específica de descomponer la distribución.

Brechman & Shepsmeier (2013); proponen los C-Vines y D-Vines para simular distribuciones multivariadas flexibles las cuales son necesarias en distintas áreas. La distribución Gaussiana es muy restrictiva y no puede detectar características como asimetría y colas pesadas, las cuales en los análisis financieros son muy comunes. El uso de las cópulas es un reto en dimensiones superiores donde las cópulas multivariadas estándar tienen o poseen estructuras inflexibles. Los vines cópulas superan estas limitaciones y son capaces de modelar patrones de dependencia complejos al tener una gran variedad de cópulas bivariadas como componentes básicos. Los vines son modelos gráficos flexibles para describir cópulas multivariadas usando una cascada de cópulas multivariadas llamadas pair-cópulas. Tales construcciones, descomponen una densidad de probabilidad multivariada en cópulas bivariadas, donde cada pair-cópula puede ser elegido de manera independiente de los otros. Esto permite una enorme flexibilidad en el modelamiento de la dependencia. En particular; las asimetrías y la dependencia de las colas pueden ser tenidas en cuenta, así como la independencia condicional para construir modelos más parsimoniosos. Los vines combinan las ventajas de modelización con cópulas multivariadas, es decir, separan del modelo marginal, la dependencia y la flexibilidad de la cópula bivariada. El enfoque basado en cópulas multivariadas presenta inconvenientes. El número de cópulas disponibles para modelar la estructura de dependencia entre más de dos variables es limitado. De hecho, la gran mayoría de las cópulas paramétricas disponibles son bivariadas. Además, el enfoque basado en una cópula multivariada puede no resultar apropiado cuando los pares de variables presentan estructuras de dependencia diferentes. Finalmente, en muchos casos en que sí se cuenta con una extensión multivariada de la cópula, se utiliza un solo parámetro para describir la estructura de dependencia entre todos los pares de variables, lo cual es una limitación importante si las relaciones entre pares de variables son diferentes. Las contrucciones con

cópulas bivariadas y los vines constituyen una alternativa al enfoque basado en cópula multivariada. Estos modelos permiten extender las cópulas bivariadas a dimensiones superiores utilizando solamente cópulas bivariadas y densidades univariadas como bloques constructivos. El creciente interés en estos modelos se debe a la gran flexibilidad que brindan para modelar una amplia variedad de dependencias combinando cópulas bivariadas de familias diferentes.

5.1. Metodología de los vines

Los vines son una representación gráfica para especificar las llamadas construcciones pair-cópulas (PCCs) introducidas por Aas et al. (2009). En particular, las PCC en tres dimensiones, estarían dadas como sigue:

Sea $X = (x_1, x_2, x_3)^T \sim F$ con funciones de distribución marginales F_1 , F_2 y F_3 y la respectivas funciones de densidad f_1 , f_2 y f_3 , luego la función de densidad trivariada se define como:

$$f(x_1, x_2, x_3) = f_1(x_1)f(x_2|x_1)f(x_3|x_1, x_2)$$

Por el teorema de Sklar, sabemos que

$$f(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)}$$

como

$$c_{12}(x_1, x_2) = \frac{\partial^2 C_{12}(x_1, x_2)}{\partial x_1 \partial x_2}$$

Así

$$f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2)$$

$$\begin{aligned} f(x_2|x_1) &= \frac{f(x_1, x_2)}{f_1(x_1)} \\ &= \frac{c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2)}{f_1(x_1)} \\ &= c_{12}(F_1(x_1), F_2(x_2)) \cdot f_2(x_2) \\ f(x_2|x_1) &= c_{12}(F_1(x_1), F_2(x_2)) \cdot f_2(x_2) \end{aligned}$$

Nota: c_{12} es la densidad de la cópula.

Vale la pena destacar que:

$$\begin{aligned} F(x_1, x_2) &= C_{12}(F_1(x_1), F_2(x_2)) \\ \frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2} &= \frac{\partial^2 C_{12}(F_1(x_1), F_2(x_2))}{\partial x_1 \partial x_2} \\ f(x_1, x_2) &= c_{12}(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2) \end{aligned}$$

y

$$\begin{aligned}
f(x_3|x_1, x_2) &= \frac{f(x_2, x_3|x_1)}{f(x_2|x_1)} \\
&= \frac{c_{2,3|1}(F(x_2|x_1), F(x_3|x_1))f(x_2|x_1)f(x_3|x_1)}{f(x_2|x_1)} \\
&= c_{2,3|1}(F(x_2|x_1), F(x_3|x_1))f(x_3|x_1) \\
&= c_{2,3|1}(F(x_2|x_1), F(x_3|x_1))c_{1,3}(F_1(x_1), F_3(x_3))f_3(x_3)
\end{aligned}$$

La densidad conjunta trivariada, se representa en términos de las cópulas bivariadas C_{12} , C_{13} y $C_{23|1}$ con densidades c_{12} , c_{13} y $c_{23|1}$ llamadas pair-cópulas, las cuales pueden ser seleccionadas en forma independiente la una de la otra, para lograr una amplia gama de diferentes estructuras de dependencia. Normalmente se asume que la cópula condicional $C_{2,3|1}$ es independiente de las variables que condicionan, x_1 , para facilitar la inferencia Kurowicka & Cooke (2006) y Kurowicka & Joe (2011).

Los vines organizan $d(d-1)/2$ pair-cópulas de un d -dimensional en $d-1$ árboles enlazados (gráficos acíclicos conectados con los nodos y los bordes). En el primer árbol C-Vines, la dependencia es con respecto a una variable en particular, el primer nodo raíz se modela usando cópulas bivariadas para cada par. Condicionando a esta variable la dependencia de parejas con respecto a la segunda variable y así se modela el segundo nodo raíz. En general, se elige un nodo raíz en cada árbol y todas las dependencias entre parejas con respecto a este nodo se modelan condicionadas en todos los nodos de las raíces anteriores, es decir, los árboles C-Vines tienen una estructura en estrella.

Esto da a la siguiente descomposición de una densidad multivariada sin pérdida de generalidad la densidad C-Vine con nodos de raíz $1, \dots, d$. (de lo contrario los nodos se pueden reetiquetar).

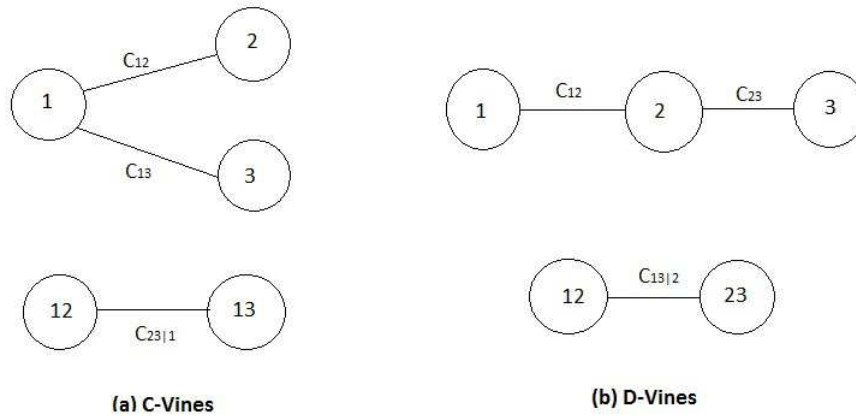
$$f(x) = \prod_{k=1}^d f_k(x_k) \cdot \prod_{i=1}^{d-1} \prod_{j=1}^{d-i} c_{i,i+j|1:(i-1)}(F(x_i|x_1, \dots, x_{i-1}), F(x_{i+j}|x_1, \dots, x_{i-1})|\theta_{i,i+j|1:(i-1)})$$

donde f_k , $k = 1, \dots, d$ denota las densidades marginales y $C_{i, i+j|1:(i-1)}$ las densidades de las cópulas bivariadas con parámetros $\theta_{i, i+j|1:(i-1)}$ (en general $i_k : i_m$ mediante i_k, \dots, i_m).

Aquí, el producto exterior se ejecuta sobre $d-i$ árboles y nodos de raíz i , mientras que el producto interno se refiere a la $d-i$ par-cópulas en cada árbol $i = 1, \dots, d-1$. Nuestro ejemplo tridimensional puede ser interpretado como un C-vine y su probabilidad se puede encontrar en Aas et al. (2009) y en Czado et al. (2012).

Del mismo modo, D-vines también se construyen mediante la elección de un orden específico de variables. Luego, en el primer árbol, la dependencia de la primera y la segunda variable, de la segunda y la tercera variable, de la tercera y de la cuarta variable y así sucesivamente se modela utilizando pares de cópulas, es decir, asumiendo el orden $1, \dots, d$, vamos a modelar los pares $(1, 2)$, $(2, 3)$, $(3, 4)$, ... y así sucesivamente.

Para realizar la selección de la estructura de los C-Vines y D-Vines; se deben seleccionar qué pares de variables serán modeladas explícitamente con cópulas. Un C-Vine resulta



Gráfica 5-1.: Representación gráfica de los C-DVines: a) C-Vines b) D-Vines

apropiado si una variable domina las interacciones entre las variables. Un D-Vine permite una selección más flexible de dependencias. En la gráfica 5-1, se ilustran los vines; C-Vines en la parte izquierda y los D-Vines en la parte derecha. En este trabajo se usaron los C-Vines, puesto que permiten una estructura de dependencia la cual es determinada tanto por el primer árbol, como por el segundo, pero para los D-Vines solo el primer árbol determina la estructura de dependencia.

5.2. Simulación de un C-Vine o un D-Vine

La simulación de un C-Vine o un D-Vine con n variables se inicia con la generación de igual número de observaciones uniformes independientes $w_i \in (0, 1)$. A partir de estas observaciones, se obtienen u_1, u_2, \dots, u_n con la estructura de dependencia representada por el vine de acuerdo a:

$$\begin{aligned}
 u_1 &= w_1 \\
 u_2 &= F_{2|1}^{-1}(w_2|u_1) \\
 &\vdots \\
 u_n &= F_{n|1,2,\dots,n-1}^{-1}(w_n|u_1, \dots, u_{n-1})
 \end{aligned}$$

Las variables de la nueva población se obtienen a partir de la muestra de variables uniformes, u_1, u_2, \dots, u_n , utilizando el método de inversión, Devroye (1986).

6. Estudio de simulación

En este capítulo se presentan los resultados obtenidos al comparar la función de sobrevivencia verdadera del tiempo mínimo de falla $S(t)$, con la función de sobrevivencia estimada asumiendo independencia entre los modos de falla ($S_I(t)$) y la función de sobrevivencia bajo dependencia ($S_D(t)$) del tiempo mínimo de falla, la cual fue estimada haciendo uso del método de combinación de riesgos, para un modelo de riesgos competitivos trivariado. Se analizan distintas dependencias entre los riesgos, dichas estructuras son generadas mediante la técnica C-DVines.

6.1. Metodología y Escenarios de simulación

En esta sección se describen los escenarios y la metodología de simulación utilizados para el análisis de la función de sobrevivencia en un modelo de riesgos competitivos dependientes trivariados bajo distintas estructuras de dependencias entre los modos de falla.

Inicialmente se generán muestras de tamaño $n = 50$ y $n = 100$ de tiempos de falla trivariados con la dependencia deseada entre los riesgos del modelo en riesgos competitivos dependientes y marginales Weibull. Se tomaron estos tamaños de muestra, puesto que en confiabilidad siempre es de gran interés estudiar problemas con tamaños de muestra pequeños, por los costos y no es común estudiar problemas con tamaños de muestra demasiado grande. Las dependencias estudiadas entre los modos de falla se relacionan en la siguiente tabla:

Dependencia	τ_{12}	τ_{13}	$\tau_{23 1}$
Escenario 1	0.5	0.5	0.05
Escenario 2	0.5	0.5	0.9
Escenario 3	0.2	0.5	0.8

Tabla 6-1.: Estructuras de dependencia del modelo de riesgos competitivos dependientes trivariado

Las dependencias propuestas se generaron mediante las técnicas: C-Vines; la cual es apropiada para dependencias donde una de las variables domina toda la estructura de dependencia y D-Vines; permiten una selección más flexible de las dependencias. En el caso 1 se tomaron dos riesgos con igual dependencia y el otro riesgo independiente a los otros

dos. En el caso 2 se tomaron dos riesgos con igual dependencia y el otro riesgo con una dependencia alta. En el caso 3 se tomaron los tres riesgos con distintas dependencias.

Seguidamente se encuentra el tiempo mínimo, es decir, $T = \min(T_1, T_2, T_3)$. Se rotulan con 1 si la falla ocurre por el modo de falla 1; 2 si la falla ocurre por el modo 2 y 3 si la falla se presenta por el modo 3.

Para obtener los tiempos de falla Weibull, se hace uso de lo propuesto por Lu & Bhattacharyya (1990) en el caso bivariado. Sean T_1 y T_2 tiempos de falla Weibull, una función de sobrevivencia conjunta para la Weibull bivariada se define como sigue:

$$S(t_1, t_2) = \exp \left\{ - \left[\left(\frac{t_1}{\eta_1} \right)^{\beta_1/(1-\tau)} + \left(\frac{t_2}{\eta_2} \right)^{\beta_2/(1-\tau)} \right]^{(1-\tau)} \right\}$$

donde β_1, β_2 son los parámetros de forma y η_1, η_2 son los parámetros de escala asociados a T_1 y T_2 respectivamente, además $0 \leq \tau \leq 1$, es el parámetro de dependencia entre T_1 y T_2 . En los escenarios Weibull estudiados se consideraron diferentes formas de la función *Hazard* de los tiempos de falla. Cuando el parámetro $\beta < 1$, la tasa de falla es decreciente, cuando $\beta > 1$, la tasa es creciente y cuando $\beta = 1$, la tasa de falla es constante.

Parametros	Tamaño	β_1	β_2	β_3	η_1	η_2	η_3
Caso 1	50	2	2	2	1	1	1
Caso 2		1.5	1	1	1	1	1
Caso 3		0.5	1	1	1	1	1
Caso 4		1.2	1	0.5	1	1	1
Caso 5	100	2	2	2	1	1	1
Caso 6		1.5	1	1	1	1	1
Caso 7		0.5	1	1	1	1	1
Caso 8		1.2	1	0.5	1	1	1

Tabla 6-2.: Parámetros de la distribución Weibull y tamaños de muestra

Seguidamente se estimaron los parámetros para las distribuciones del tiempo de falla para cada riesgo, con estos estimadores se obtienen las funciones de sobrevivencia del tiempo mínimo de falla asumiendo independencia. Luego se estiman las distribuciones marginales para cada modo de falla mediante el método de combinación de riesgos, el cual nos permite estimar la función de sobrevivencia bajo dependencia entre los riesgos del modelo trivariado. Mediante simulación de Monte Carlo se construyen intervalos de confianza empíricos aproximados al 95 % para los percentiles t_p , con $p = 0.05, 0.25, 0.50, 0.75, 0.95$, tanto para función de sobrevivencia estimada bajo independencia y bajo dependencia.

Finalmente se comparó los dos estimadores en cuestión calculando la eficiencia relativa ER_p de $(\hat{S}_I(t))$ con referencia a $(\hat{S}_D(t))$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$

$$ER_p = \frac{ECM(\widehat{S}_D(t_p))}{ECM(\widehat{S}_I(t_p))} = \frac{E[(\widehat{S}_D(t_p) - S(t_p))^2]}{E[(\widehat{S}_I(t_p) - S(t_p))^2]}$$

Si $ER_p < 1$ entonces $S_I(t_p)$ es menos eficiente en relación a $S_D(t_p)$ para la estimación de la función de sobrevivencia verdadera en los percentiles estudiados.

Para efectos de tener mayor claridad a la hora de leer los resultados, *DLI* y *DLS* hacen referencia a las estimación del limite inferior y superior del intervalo de confianza para la función de sobrevivencia bajo dependencia entre los modos de falla del modelo en riesgos competitivos trivario. *INLI* y *INLS* quiere decir que es la estimación de el limite inferior y el limite superior del intervalo de confianza para la función de sobrevivencia cuando se tiene independencia entre los modos de falla del modelo en cuestión. $S(t)$ corresponde a la función de sobrevivencia verdadera.

6.2. Resultados

A continuación se muestran los resultados obtenidos en todos los escenarios de simulación analizados:

6.3. Escenario 1 de dependencia

En este escenario se simularon dos riesgos con igual dependencia y el otro riesgo se tomó independiente a los anteriores con tamaños de muestrales de 50 y 100. Los riesgos tienen dependencia $\tau_{12} = 0.5$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.05$ y cuyo parámetro de cada cópula son respectivamente; $\theta_{12} = 5.7$, $\theta_{13} = 5.7$ y $\theta_{23|1} = 0.45$ y se usó el método de C-Vines, el cual genera la dependencia teniendo en cuenta el primer riesgo.

n	β_1	β_2	β_3	Modo 1	Modo 2	Modo 3
50	2	2	2	46 %	28 %	26 %
	1.5	1	1	46 %	24 %	30 %
	0.5	1	1	56 %	26 %	18 %
	1.2	1	0.5	28 %	24 %	48 %
100	2	2	2	42 %	33 %	25 %
	1.5	1	1	39 %	31 %	30 %
	0.5	1	1	52 %	21 %	27 %
	1.2	1	0.5	25 %	29 %	46 %

Tabla 6-3.: Porcentaje de falla para el modelo Weibull con tamaños de muestra $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$ y C-Vines.

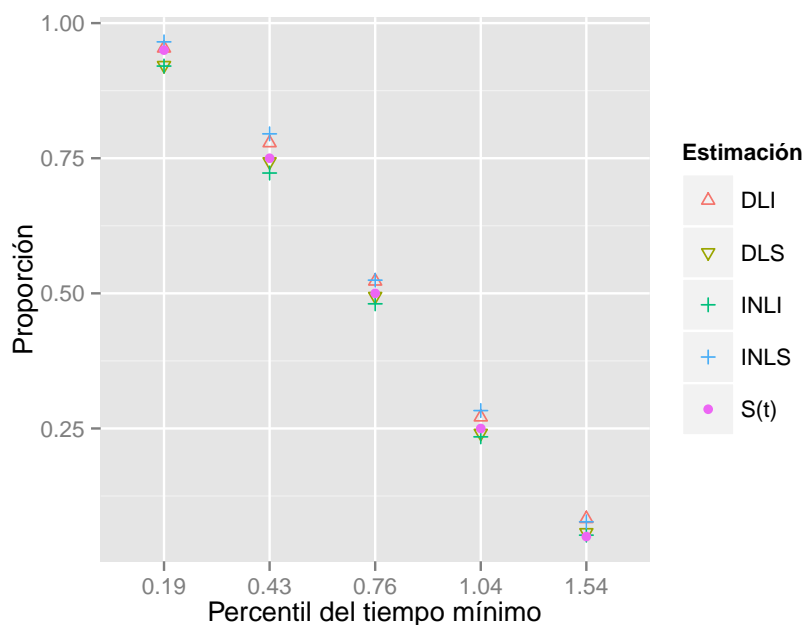
En la Tabla **6-3**, se observa que el porcentaje de falla para cada uno de los modos de falla son mayores o iguales al 15 %, lo que nos garantiza tener buenas estimaciones de la función de

sobrevivencia para el tiempo mínimo cuando se realizan las estimaciones bajo dependencia e independencia entre los modos de fallas del modelo en riesgos competitivos trivariado dependiente. Los porcentajes de falla en la gran mayoría de los casos son distintos, por tanto garantiza que los modelos estudiados no se dejan influenciar por las dependencias. Lo anterior permite ilustrar claramente el problema de identificabilidad y es claro que no se están trabajando con funciones de distribución marginales sino con funciones de subdistribución Paz-Sabogal. et al. (2014).

En la Tablas **6-4**, **6-5**, **6-6**, **6-7** se muestran los límites inferior y superior de la función de sobrevivencia estimada bajo dependencia $LI.S_D(t)$ y $LS.S_D(t)$ respectivamente. Los límites inferior y superior de la función de sobrevivencia estimada bajo independencia denotados por $LI.S_I(t)$, $LS.S_I(t)$, y la precisión de cada intervalo de confianza empírico aproximado para la sobrevivencia verdadera en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$ y tama no de muestral 50. Se evidencia que los intervalos de confianza bajo dependencia en los modos de falla resultan ser más precisos que los intervalos de confianza bajo independencia entre los modos de falla del modelo. En las gráficas **6-1**, **6-2**, **6-3** y **6-4** se observa que los intervalos de confianza bajo dependencia entre los modos de falla del modo trivariado siempre se encuentran contenidos en los intervalos de confianza bajo independencia en los modo de falla del modelo. En algunos casos el limite inferior o superior se encuentran cercanos al valor verdadero pero lo importante es que contengan al valor de la sobrevivencia verdadera.

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.92194	0.95360	0.03166	0.92051	0.96526	0.04475
0.25	0.74373	0.77831	0.03458	0.72263	0.79506	0.07243
0.50	0.49458	0.52236	0.02778	0.48063	0.52442	0.04379
0.75	0.24129	0.27104	0.02975	0.23449	0.28312	0.04864
0.95	0.05746	0.08293	0.02547	0.05262	0.08731	0.03466

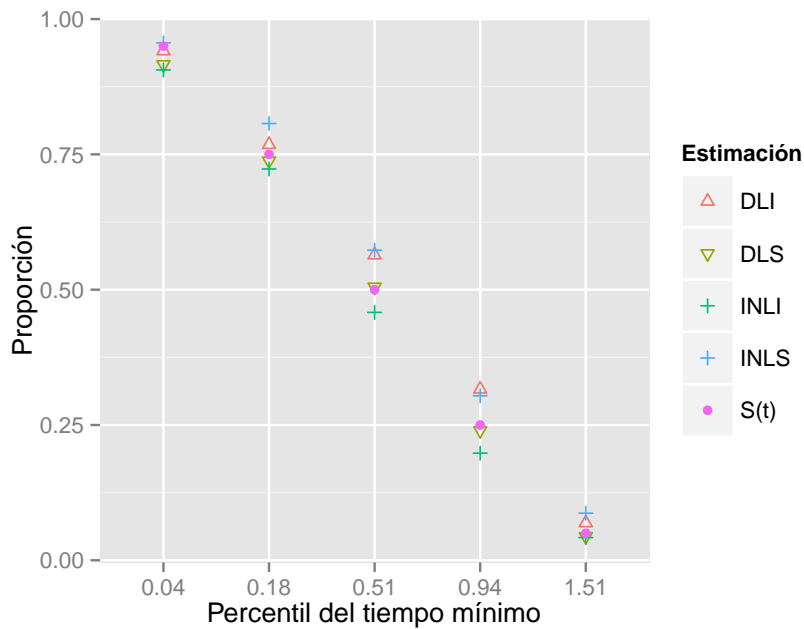
Tabla 6-4.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$



Gráfica 6-1.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.91630	0.95091	0.03461	0.90648	0.95629	0.04981
0.25	0.73801	0.76825	0.03025	0.72281	0.80682	0.08401
0.50	0.50586	0.56409	0.05823	0.45817	0.57274	0.11457
0.75	0.24050	0.31619	0.07570	0.19825	0.30360	0.10534
0.95	0.04427	0.06895	0.02468	0.04236	0.08662	0.04426

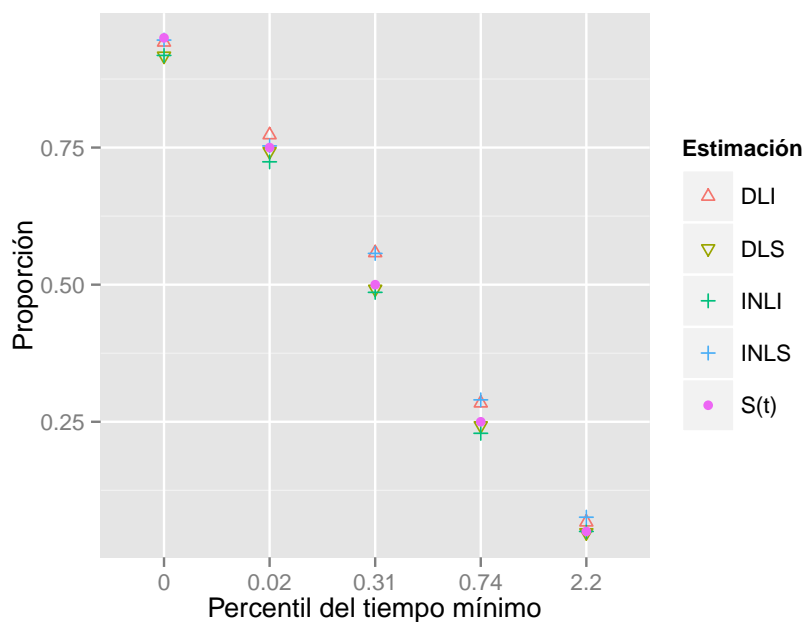
Tabla 6-5.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$



Gráfica 6-2.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.91749	0.95199	0.03450	0.91824	0.95595	0.03771
0.25	0.74329	0.77265	0.02936	0.72363	0.75330	0.02967
0.50	0.49172	0.55792	0.06620	0.48593	0.55701	0.07108
0.75	0.24294	0.28412	0.04118	0.22911	0.28969	0.06057
0.95	0.04805	0.06686	0.01881	0.04975	0.07602	0.02627

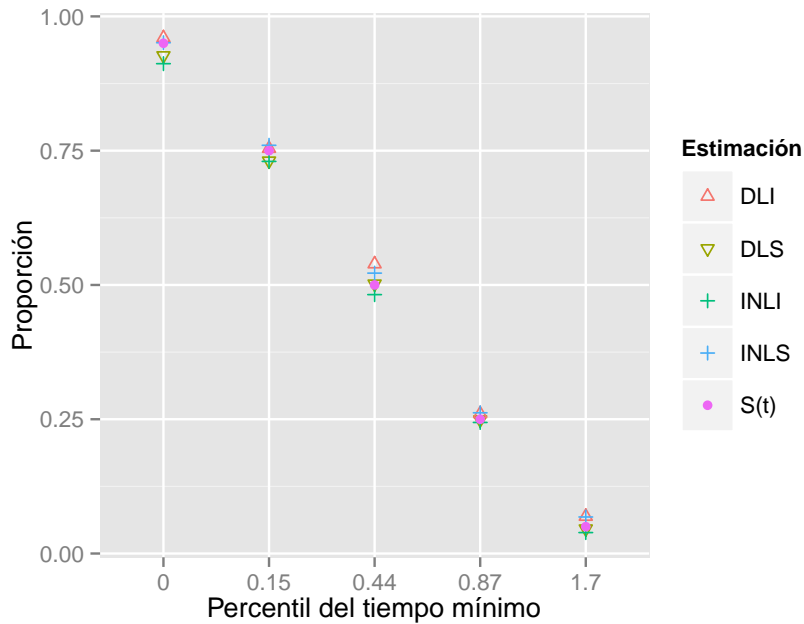
Tabla 6-6.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$



Gráfica 6-3.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.92666	0.95999	0.03333	0.91185	0.95107	0.03922
0.25	0.73058	0.75438	0.02380	0.73049	0.76006	0.02957
0.50	0.50191	0.53937	0.03746	0.48191	0.52224	0.04033
0.75	0.24899	0.26009	0.01110	0.24449	0.26162	0.01714
0.95	0.04619	0.06875	0.02256	0.03906	0.06842	0.02936

Tabla 6-7.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$

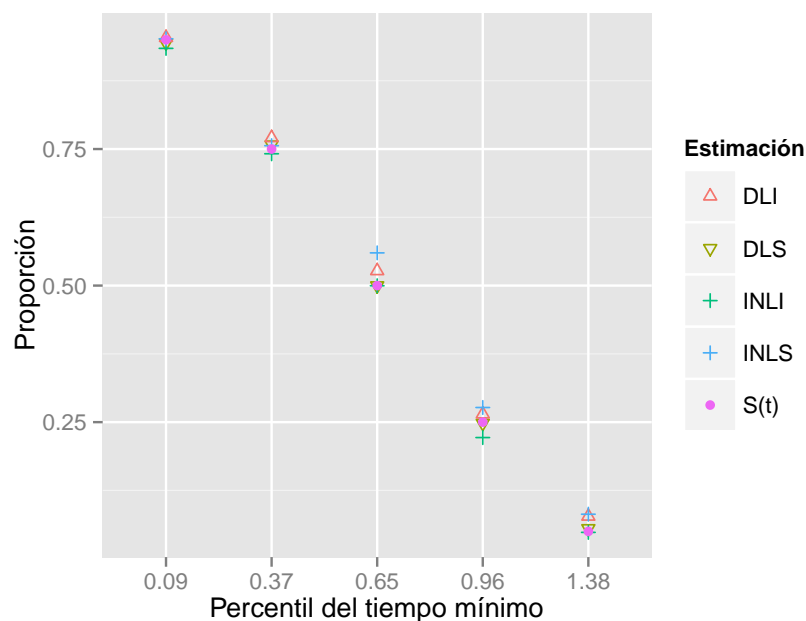


Gráfica 6-4.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Para las tablas **6-8**, **6-9**, **6-10** y **6-11** se muestran todos los intervalos de confianza para la estimación de la función de sobrevivencia para el caso donde el tama no muestral es de 100, evidenciandose que los intervalos para la función de sobrevivencia cuando se tienen dependencia entre los modos de falla del modelo trivariado son más precisos en cada uno de los percentiles estudiados. En las gráficas **6-5**, **6-6**, **6-7** y **6-8** todos los intervalos de confianza contienen el valor verdadero de la función de sobrevivencia tanto para los casos donde los modos de falla del modelo en riesgos competitivos trivariado son independientes o dependientes. Además, no siempre los intervalos para la función de sobrevivencia están enteramente contenidos en los intervalos para función de sobrevivencia cuando se asume independencia entre los modos de falla del modelo trivariado dependientes. En la gráfica **6-6** para el percentil $p = 0.25$ los intervalos no estan encajonados pero en la intercepción contiene al valor verdadero de la función de sobrevivencia.

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.94427	0.95384	0.00957	0.93426	0.95176	0.01749
0.25	0.75763	0.77055	0.01291	0.74134	0.75608	0.01474
0.50	0.49985	0.52695	0.02710	0.49968	0.55991	0.06023
0.75	0.24802	0.26471	0.01669	0.22177	0.27679	0.05502
0.95	0.05529	0.07762	0.02233	0.04823	0.08146	0.03322

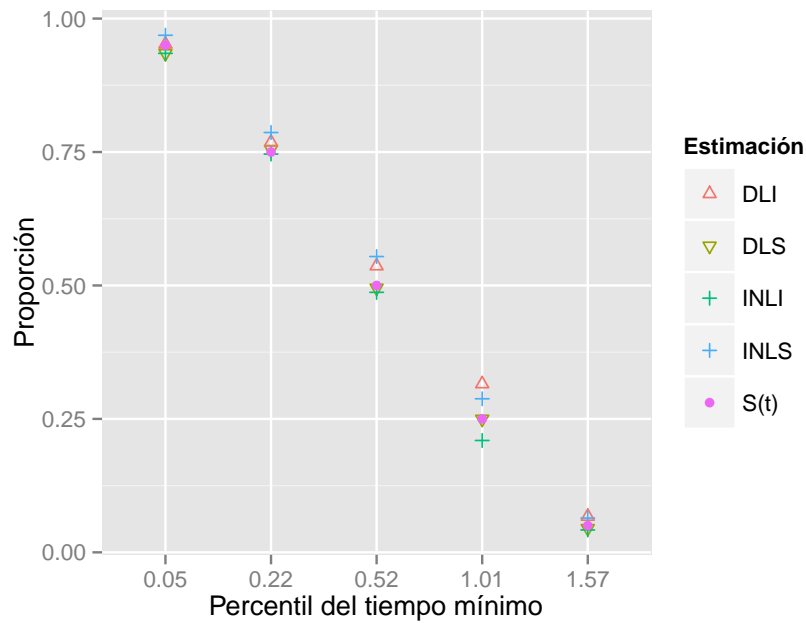
Tabla 6-8.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$



Gráfica 6-5.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.93470	0.95197	0.01726	0.93481	0.96872	0.03392
0.25	0.75510	0.76838	0.01327	0.74617	0.78655	0.04038
0.50	0.49532	0.53644	0.04112	0.48713	0.55419	0.06706
0.75	0.24975	0.31549	0.06574	0.20957	0.28769	0.07812
0.95	0.04496	0.06718	0.02222	0.04194	0.06421	0.02227

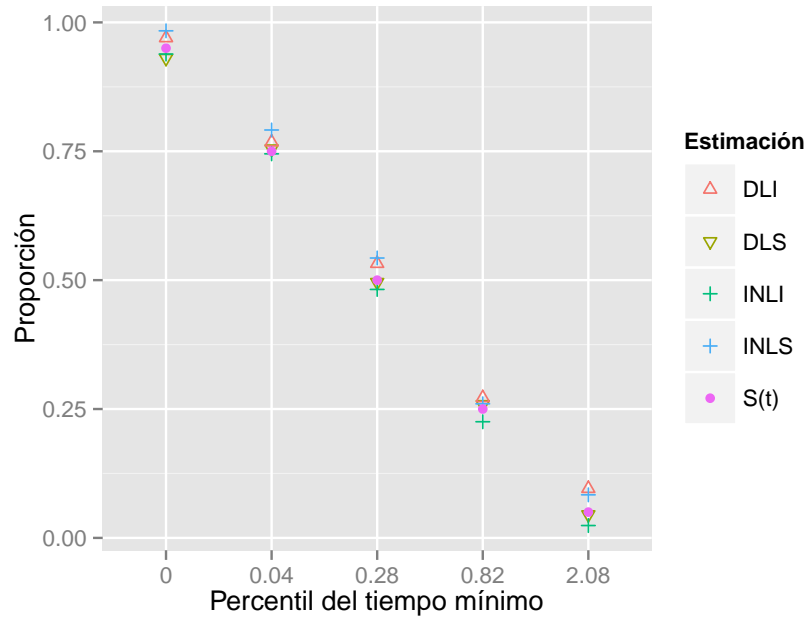
Tabla 6-9.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$



Gráfica 6-6.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.93051	0.96997	0.03946	0.93857	0.98365	0.04508
0.25	0.75484	0.76768	0.01284	0.74508	0.79127	0.04619
0.50	0.49577	0.53219	0.03642	0.48201	0.54294	0.06094
0.75	0.25882	0.27172	0.01290	0.22533	0.26012	0.03479
0.95	0.04511	0.09564	0.05054	0.02401	0.08355	0.05954

Tabla 6-10.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$



Gráfica 6-7.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

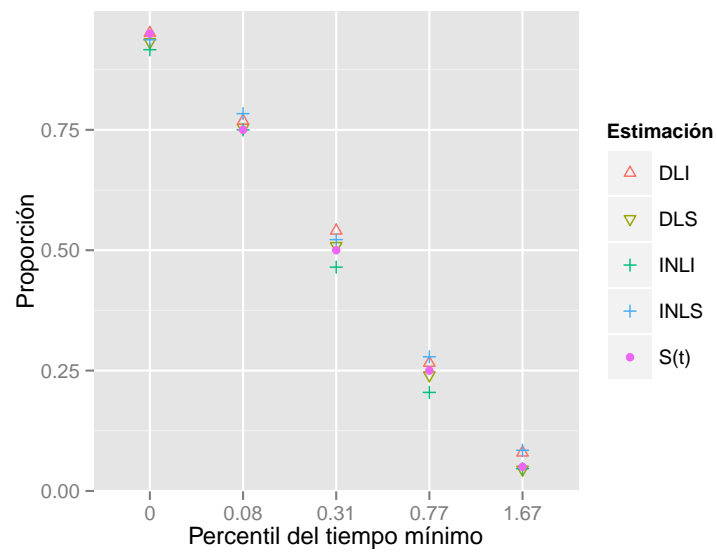
En la Tabla **6-12**, se muestra la eficiencia relativa ER_p para la función de sobrevivencia conjunta estimada bajo dependencia $S_D(t)$ en relación a la función de sobrevivencia conjunta estimada bajo independencia $S_I(t)$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$. Para todos los casos donde las marginales están distribuidas Weibull con parámetros estipulados en la Tabla **6-2** y tamaños de muestra $n = 50, 100$, con dependencias $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$. Se observa que el estimador bajo independencia es menos eficiente que el estimador bajo dependencia puesto que los valores calculados son menores que 1, vale la pena resaltar que en algunos percentiles el valor calculado es cercano a 1 y general no se evidenció ningún patrón en particular.

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.93242	0.95103	0.01861	0.91650	0.93706	0.02056
0.25	0.75464	0.76846	0.01382	0.74990	0.78368	0.03377
0.50	0.50898	0.54041	0.03142	0.46489	0.52205	0.05716
0.75	0.24051	0.26588	0.02536	0.20471	0.27866	0.07395
0.95	0.04448	0.07952	0.03504	0.04612	0.08453	0.03840

Tabla 6-11.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$

Parámetros				Percentiles				
n	β_1	β_2	β_3	0.05	0.25	0.50	0.75	0.95
50	2	2	2	0.8506	0.7185	0.8513	0.4378	0.9981
	1.5	1	1	0.8512	0.7819	0.7325	0.4053	0.9917
	0.5	1	1	0.9595	0.9493	0.4341	0.1043	0.2312
	1.2	1	0.5	0.9949	0.8011	0.4204	0.3413	0.4169
100	2	2	2	0.0860	0.0493	0.2444	0.7289	0.1211
	1.5	1	1	0.8512	0.7819	0.7325	0.4053	0.9917
	0.5	1	1	0.9595	0.9493	0.4341	0.1043	0.2312
	1.2	1	0.5	0.9949	0.8011	0.4204	0.3413	0.4169

Tabla 6-12.: Eficiencia relativa para el modelo Weibull con tamaños muestrales $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$ y C-Vines



Gráfica 6-8.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.05$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

6.4. Escenario 2 de dependencia

En este escenario se simularon dos riesgos con igual dependencia y el otro riesgo se tomo con una dependencia alta en relación a los otros dos para tamaños de muestrales de 50 y 100. Los riesgos tienen dependencia $\tau_{12} = 0.5$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.9$, cuyo parámetro de cada cópula son respectivamente; $\theta_{12} = 5.7$, $\theta_{13} = 5.7$ y $\theta_{23|1} = 38.28$ y se usó el método de C-Vines el cual genera la dependencia teniendo en cuenta el primer riesgo.

n	β_1	β_2	β_3	Modo 1	Modo 2	Modo 3
50	2	2	2	48 %	20 %	32 %
	1.5	1	1	42 %	36 %	22 %
	0.5	1	1	58 %	26 %	16 %
	1.2	1	0.5	32 %	16 %	52 %
100	2	2	2	48 %	26 %	26 %
	1.5	1	1	41 %	39 %	20 %
	0.5	1	1	50 %	22 %	28 %
	1.2	1	0.5	28 %	15 %	57 %

Tabla 6-13.: Porcentaje de falla para el modelo Weibull con tamaños de muestra $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ y C-Vines.

En la Tabla **6-13**, se observa que el porcentaje de falla para cada uno de los modos de falla son mayores al 15 %, lo que nos garantiza tener buenas estimaciones de la función de sobrevivencia para el tiempo mínimo cuando se realizan las estimaciones bajo dependencia e independencia entre los modos de fallas del modelo en riesgos competitivos trivariado dependiente.

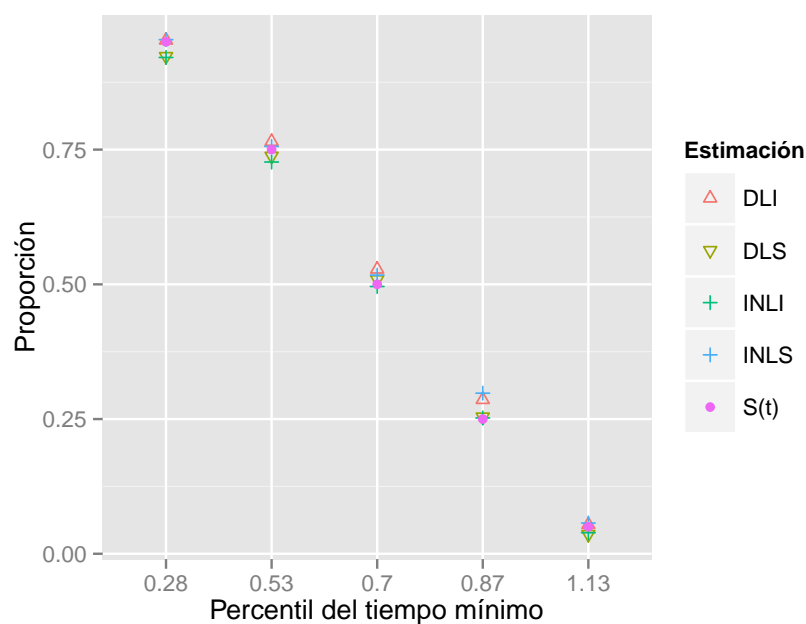
A continuación se muestran los límites inferior y superior de la función de sobrevivencia estimada bajo dependencia $LI.S_D(t)$ y $LS.S_D(t)$ respectivamente, los límites inferior y superior de la función de sobrevivencia estimada bajo independencia denotados por $LI.S_I(t)$ y $LS.S_I(t)$. Además, se incluye la precisión de cada intervalo de confianza empírico aproximado para la sobrevivencia verdadera, en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En este caso las tablas **6-14**, **6-15**, **6-16** y **6-17** con tamaño de muestra de $n = 50$ muestran las estimaciones de los intervalos de confianza empíricos aproximados para la función de sobrevivencia. Los resultados encontrados reflejan que los intervalos de confianza para la función de sobrevivencia cuando se asume dependencia son más precisos que los intervalos de confianza para la función de sobrevivencia bajo independencia entre los modos de falla del modelo trivariado. Las gráficas **6-9**, **6-10**, **6-11** y **6-12** todos los intervalos contienen la función de sobrevivencia verdadera y los intervalos de confianza para dicha función bajo dependencia entre los modos de falla en su gran mayoría están contenidos en los intervalos de confianza bajo independencia entre los modos de falla del modelo de riesgos competitivos.

En general estos intervalos estimados para este caso de dependencia y tama no muestral 50 son más estrechos que los estimados en el primer escenario de dependencia.

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.92272	0.95281	0.03008	0.92062	0.95411	0.03349
0.25	0.73819	0.76539	0.02720	0.72724	0.75640	0.02917
0.50	0.50900	0.52790	0.01889	0.49568	0.51740	0.02172
0.75	0.25423	0.28645	0.03222	0.25212	0.29802	0.04589
0.95	0.03624	0.05374	0.01750	0.03869	0.05678	0.01809

Tabla 6-14.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$



Gráfica 6-9.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.94504	0.95696	0.01192	0.94439	0.96403	0.01964
0.25	0.75334	0.77694	0.02360	0.74499	0.77048	0.02549
0.50	0.50718	0.53566	0.02849	0.50916	0.54302	0.03386
0.75	0.22217	0.25028	0.02811	0.21303	0.26775	0.05472
0.95	0.04917	0.08376	0.03458	0.05475	0.09368	0.03893

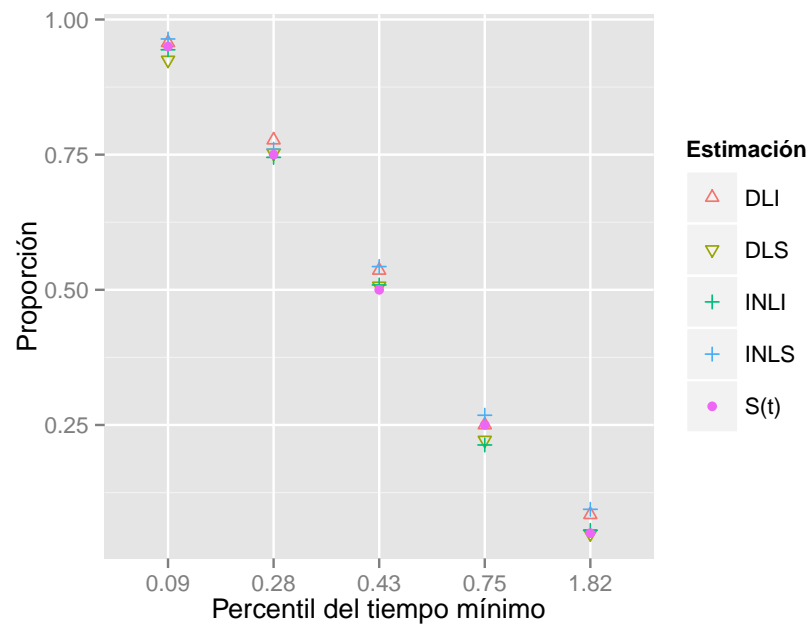
Tabla 6-15.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.91732	0.95994	0.04262	0.91414	0.96390	0.04976
0.25	0.71900	0.75911	0.04010	0.73771	0.79254	0.05483
0.50	0.48626	0.51775	0.03149	0.47618	0.51895	0.04277
0.75	0.23148	0.26697	0.03548	0.20111	0.25081	0.04971
0.95	0.04131	0.06537	0.02406	0.04286	0.07004	0.02718

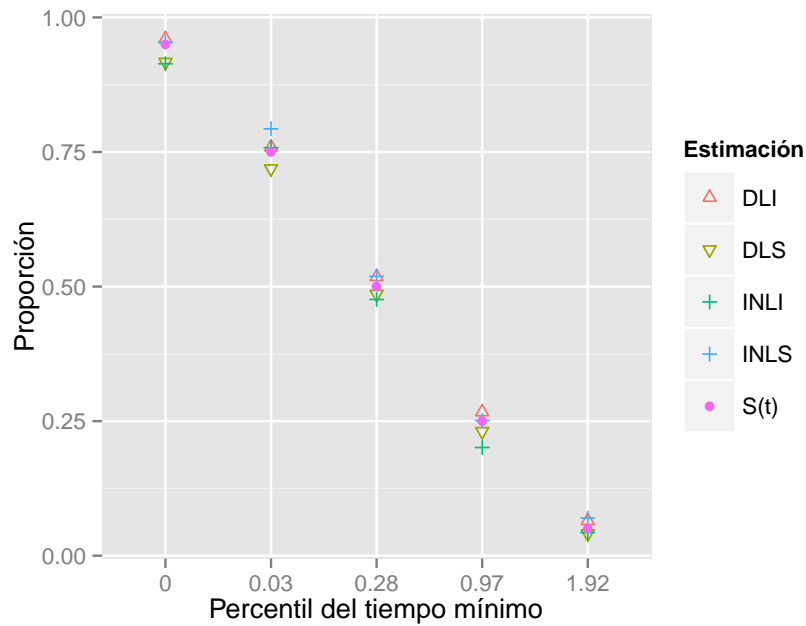
Tabla 6-16.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.94748	0.95194	0.00446	0.93488	0.95377	0.01889
0.25	0.72176	0.75150	0.02973	0.74397	0.78309	0.03912
0.50	0.47544	0.50562	0.03018	0.48419	0.52724	0.04305
0.75	0.21452	0.25935	0.04483	0.20748	0.25763	0.05015
0.95	0.03691	0.05743	0.02052	0.05693	0.08290	0.02597

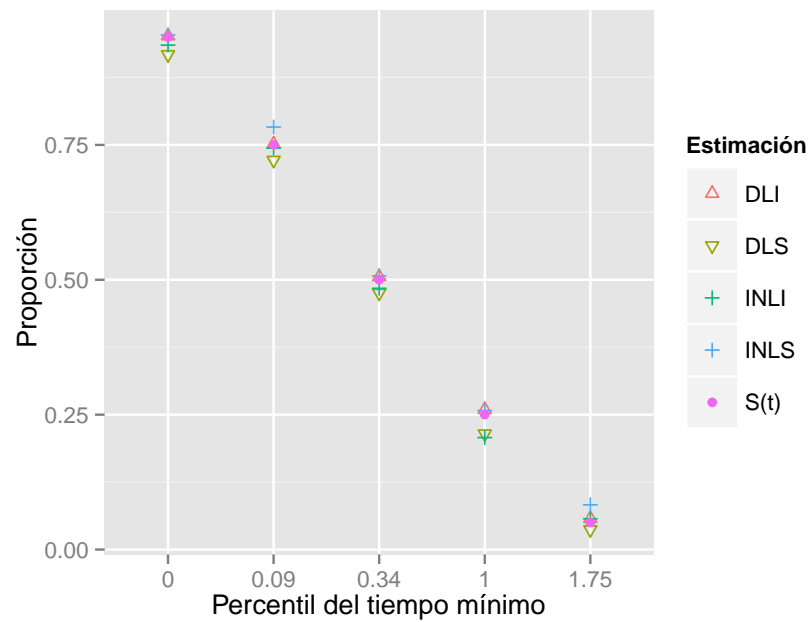
Tabla 6-17.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$



Gráfica 6-10.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$



Gráfica 6-11.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$



Gráfica 6-12.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

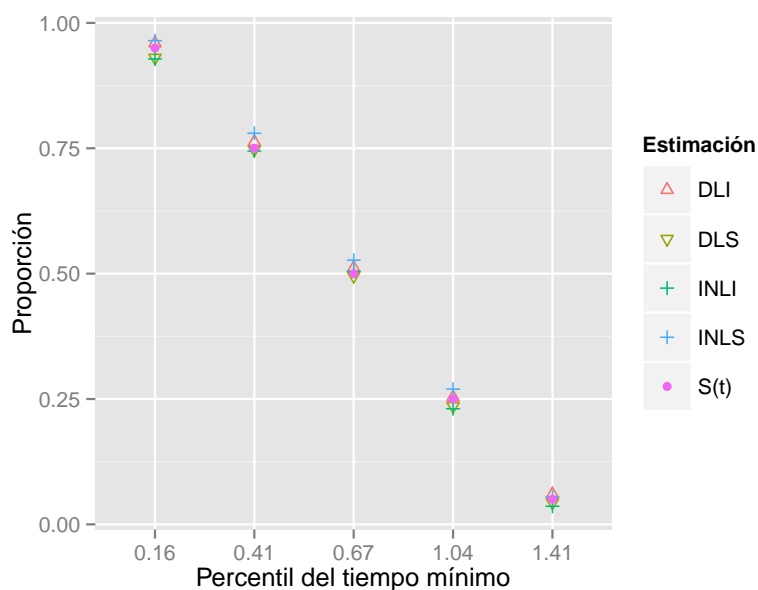
En las tablas **6-18**, **6-19**, **6-20** y **6-21** se puede evidenciar que todas las estimaciones de los límites de confianza para la función de sobrevivencia verdadera resultan ser más estrechos, es decir, tienen mayor precisión. En este escenario de dependencia y tamaño muestral de $n = 100$ el valor verdadero no lo toman los límites inferior o el límite superior, lo que quiere decir que los contienen enteramente al valor verdadero de la función de sobrevivencia. Lo anterior es ratificado por el comportamiento mostrado en las gráficas **6-15**, **6-16**, **6-17** y **6-18** donde siempre los intervalos de confianza empíricos para $S(t_p)$ cuando los modos de falla son dependientes, siempre contienen el valor verdadero que la función de sobrevivencia y cuando los modos de falla son independientes contienen la sobrevivencia verdadera pero son más amplios en relación a los primeros.

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.93038	0.95947	0.02910	0.93806	0.97461	0.03655
0.25	0.74767	0.76262	0.01495	0.74460	0.77992	0.03533
0.50	0.49542	0.51142	0.01600	0.50434	0.52686	0.02252
0.75	0.23583	0.25352	0.01769	0.23045	0.26963	0.03919
0.95	0.04372	0.06001	0.01629	0.03615	0.05508	0.01892

Tabla 6-18.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.94338	0.95813	0.01475	0.92520	0.94754	0.02234
0.25	0.76171	0.77416	0.01245	0.75085	0.76594	0.01510
0.50	0.49272	0.50791	0.01520	0.48289	0.51634	0.03345
0.75	0.22746	0.25213	0.02467	0.20841	0.26038	0.05197
0.95	0.04922	0.06480	0.01558	0.04989	0.08009	0.03020

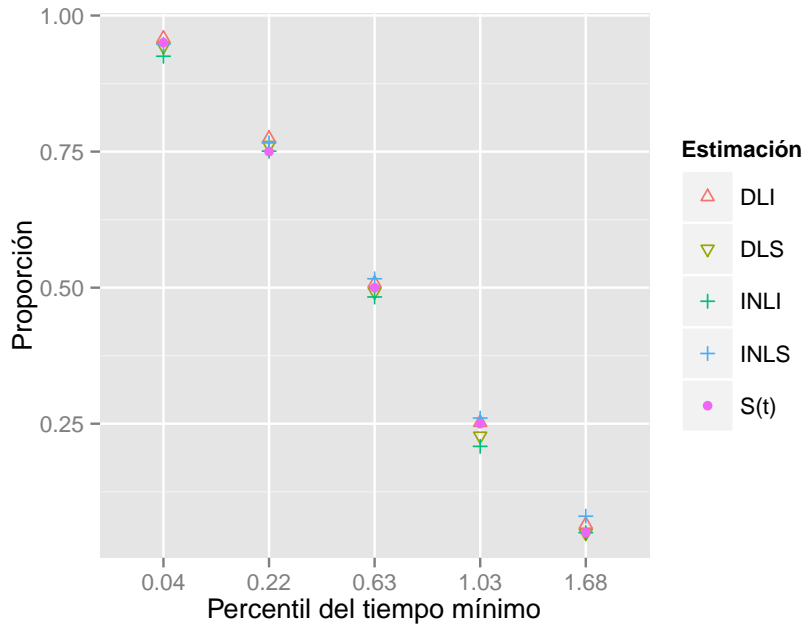
Tabla 6-19.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$



Gráfica 6-13.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.93239	0.95014	0.01775	0.92185	0.96804	0.04619
0.25	0.74199	0.75502	0.01303	0.74283	0.80754	0.06470
0.50	0.49502	0.51263	0.01761	0.48145	0.52798	0.04652
0.75	0.21668	0.25062	0.03394	0.18805	0.26183	0.07377
0.95	0.03677	0.06351	0.02674	0.02877	0.05849	0.02971

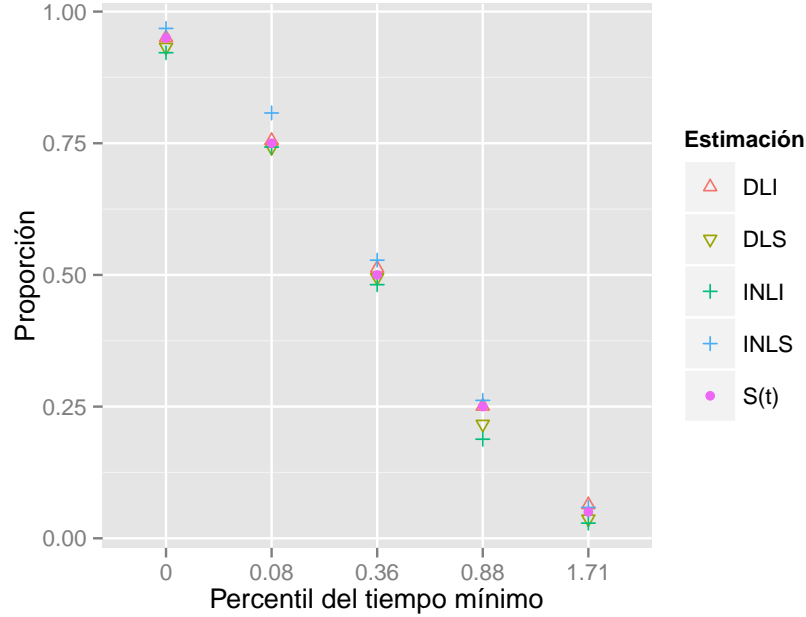
Tabla 6-20.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$



Gráfica 6-14.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.93148	0.95015	0.01867	0.90653	0.95664	0.05011
0.25	0.73413	0.76897	0.03484	0.74239	0.78563	0.04323
0.50	0.49693	0.50193	0.00500	0.48172	0.50100	0.01928
0.75	0.22753	0.25838	0.03085	0.22759	0.26802	0.04043
0.95	0.05674	0.07393	0.01719	0.03920	0.06076	0.02156

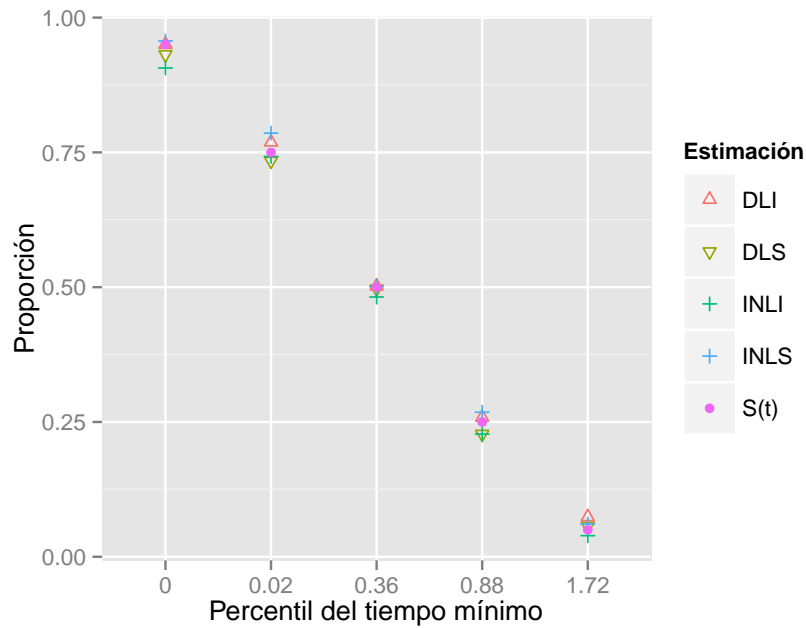
Tabla 6-21.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$



Gráfica 6-15.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Parámetros				Percentiles				
n	β_1	β_2	β_3	0.05	0.25	0.50	0.75	0.95
50	2	2	2	0.6735	0.2097	0.0796	0.2521	0.9891
	1.5	1	1	0.9596	0.5930	0.1995	0.1510	0.0287
	0.5	1	1	0.1647	0.7936	0.1066	0.1869	0.2803
	1.2	1	0.5	0.9897	0.9914	0.6969	0.5765	0.9801
100	2	2	2	0.3490	0.7165	0.1002	0.8826	0.1961
	1.5	1	1	0.0654	0.1232	0.2904	0.2281	0.1297
	0.5	1	1	0.3114	0.4045	0.8647	0.5225	0.2447
	1.2	1	0.5	0.1343	0.6804	0.2618	0.1386	0.2943

Tabla 6-22.: Eficiencia relativa para el modelo Weibull con tamaños muestrales $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ y C-Vines



Gráfica 6-16.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

En la Tabla **6-22**, se muestra la eficiencia relativa ER_p , para la función de sobrevivencia conjunta estimada bajo dependencia $S_D(t)$, con relación a la función de sobrevivencia conjunta estimada bajo independencia $S_I(t)$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$ para todos los casos donde las marginales están distribuidas Weibull con parámetros estipulados en la Tabla **6-2** y tamaños de muestra $n = 50, 100$, con dependencias $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$. Se evidencia que el estimador bajo independencia es menos eficiente que el estimador bajo dependencia, puesto que los valores calculados son menores que 1, vale la pena resaltar que en algunos percentiles el valor calculado es cercano a 1 y general no se evidencia ninguna tendencia en particular.

6.5. Escenario 3 de dependencia

En este escenario se simularon los tres riesgos con distinta dependencia. Los riesgos tienen dependencia $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ cuyo parámetro de cada cópula son respectivamente $\theta_{12} = 1.86$, $\theta_{13} = 5.7$ y $\theta_{23|1} = 18.29$ y se usó el método de C-Vines, el cual genera la dependencia teniendo en cuenta el primer riesgo y el parámetro de la cópula condicional tiene una dependencia alta.

n	β_1	β_2	β_3	Modo 1	Modo 2	Modo 3
50	2	2	2	34 %	36 %	30 %
	1.5	1	1	24 %	50 %	26 %
	0.5	1	1	38 %	22 %	40 %
	1.2	1	0.5	26 %	16 %	58 %
100	2	2	2	56 %	22 %	22 %
	1.5	1	1	50 %	25 %	25 %
	0.5	1	1	50 %	35 %	15 %
	1.2	1	0.5	28 %	15 %	57 %

Tabla 6-23.: Porcentaje de falla para el modelo Weibull con tamaños de muestra $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.2$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.8$ y C-Vines.

En la Tabla **6-23**, se observan los porcentaje de falla para cada uno de los modos de falla, los cuales resultan ser mayores o igual al 15 %, lo que nos garantiza tener buenas estimaciones de la función de sobrevivencia para el tiempo mínimo cuando se realizan las estimaciones bajo dependencia e independencia entre los modos de fallas del modelo en riesgos competitivos trivariado dependiente.

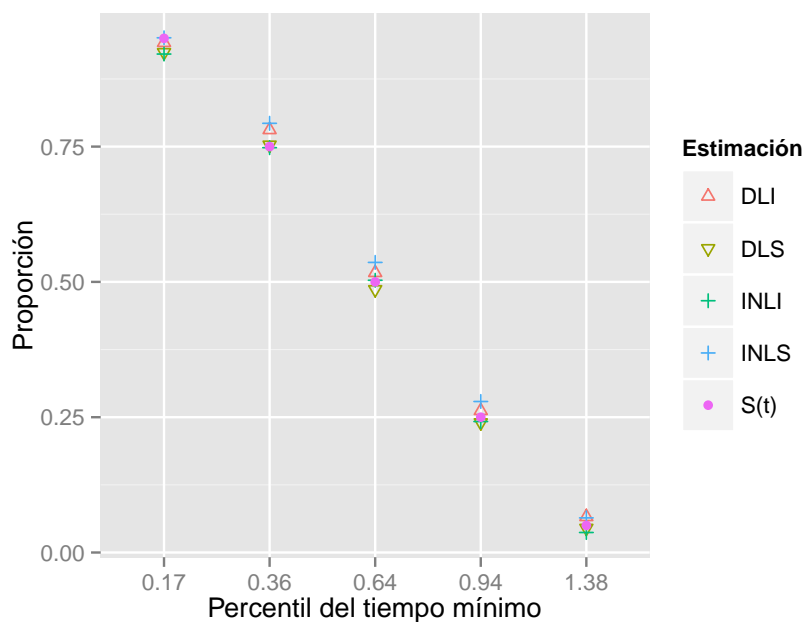
A continuación se muestran los límites inferior y superior de la función de sobrevivencia estimada bajo dependencia $LI.S_D(t)$ y $LS.S_D(t)$, respectivamente, los límites inferior y superior de la función de sobrevivencia estimada bajo independencia denotados por $LI.S_I(t)$ y $LS.S_I(t)$. Además, se incluye la precisión de cada intervalo de confianza empírico aproximado para la sobrevivencia verdadera, en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$. Para todos los casos estudiados del modelo Weibul.

En las tablas **6-24**, **6-25**, **6-26** y **6-27** se muestran los resultados para el treces escenario de dependencia y para un tama no muestral de $n = 50$ los resultados obtenidos para los intervalos de confianza empíricos para $S(t)$ bajo dependencia resultan ser más precisos y siempre el valor verdadero de la sobrevivencia se encuentra en dichos intervalos. Los intervalos bajo independencia son más amplios en algunos casos, lo cual es ratifican en las gráficas **6-17**, **6-18**, **6-19** y **6-20** donde los intervalos de confianza contienen el valor verdadero de la función de sobrevivencia tanto para el caso donde los modos de falla son independientes o dependientes y a su vez los intervalos bajo dependencia son de menor

longitud.

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.92378	0.94229	0.01850	0.92133	0.95123	0.02990
0.25	0.75273	0.78072	0.02800	0.74812	0.79253	0.04441
0.50	0.48568	0.51721	0.03153	0.50316	0.53614	0.03298
0.75	0.24021	0.26170	0.02149	0.24238	0.27906	0.03668
0.95	0.04529	0.06633	0.02104	0.03736	0.06434	0.02698

Tabla 6-24.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$



Gráfica 6-17.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

En las tablas **6-28**, **6-29**, **6-30** y **6-31** se muestran los resultados para el treces escenario de dependencia y para un tama no muestral de $n = 100$ los resultados obtenidos para los intervalos de confianza empíricos para $S(t)$ y en las gráficas **6-17**, **6-18**, **6-19** y **6-20** donde los intervalos de confianza contienen el valor verdadero de la función de sobrevivencia tanto para el caso donde los modos de falla son independientes o dependientes y a su vez los intervalos bajo dependencia tienen mayor precisión.

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.92796	0.95561	0.02765	0.92757	0.96510	0.03754
0.25	0.75821	0.77287	0.01466	0.72781	0.77767	0.04985
0.50	0.50033	0.56471	0.06438	0.45417	0.52642	0.07225
0.75	0.24409	0.29694	0.05285	0.23536	0.31051	0.07515
0.95	0.04189	0.07062	0.02873	0.05837	0.08238	0.02400

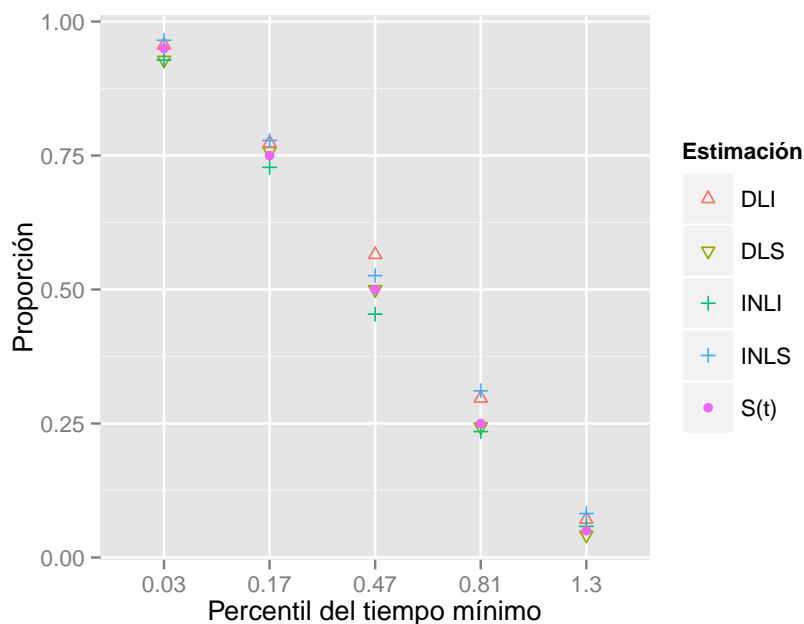
Tabla 6-25.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.91748	0.95199	0.03451	0.92478	0.97412	0.04935
0.25	0.73636	0.76129	0.02493	0.70098	0.75578	0.05480
0.50	0.50534	0.54178	0.03644	0.48546	0.53078	0.04532
0.75	0.24297	0.27717	0.03420	0.23460	0.27205	0.03745
0.95	0.03900	0.06826	0.02926	0.04699	0.07917	0.03218

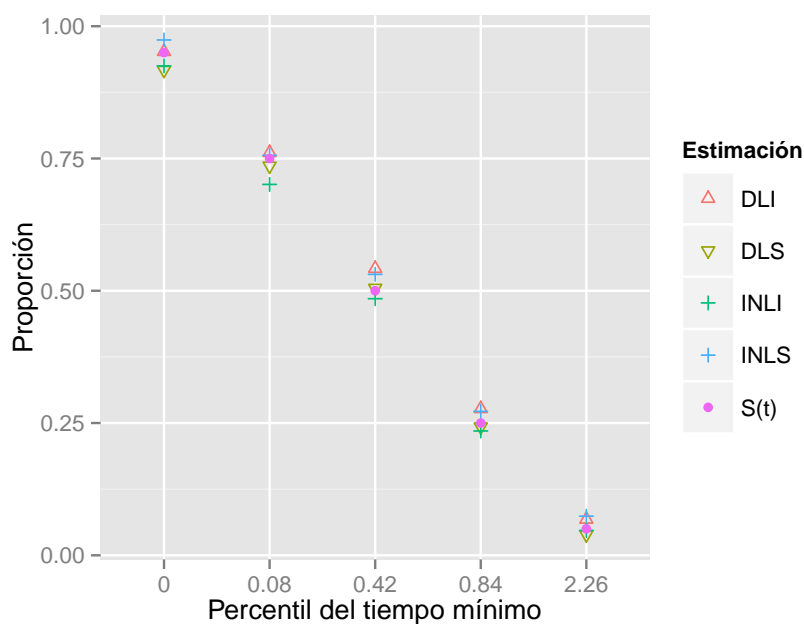
Tabla 6-26.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$

En la Tabla **6-32**, se muestra la eficiencia relativa ER_p , para la función de sobrevivencia conjunta estimada bajo dependencia $S_D(t)$, con relación a la función de sobrevivencia conjunta estimada bajo independencia $S_I(t)$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$ para todos los casos donde las marginales están distribuidas Weibull con parámetros estipulados en la Tabla **6-2** y tamaños de muestra $n = 50, 100$, con dependencias $\tau_{12} = 0.5$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.9$. Se evidencia que el estimador bajo independencia es menos eficiente que el estimador bajo dependencia, puesto que los valores calculados son menores que 1, vale la pena resaltar que en algunos percentiles el valor calculado es cercano a 1 y en general no se evidencia ningún patrón en particular.

En este trabajo el esquema de simulación, se realizó en el software R.



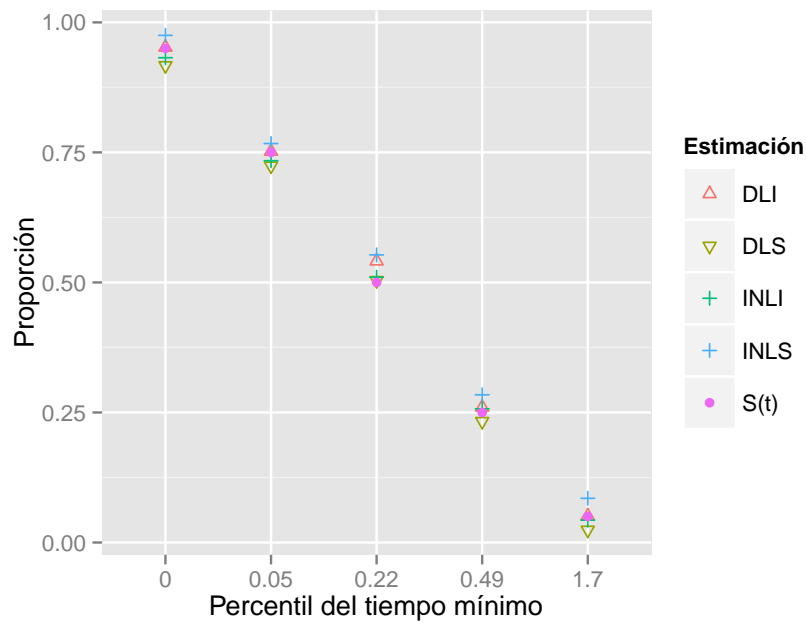
Gráfica 6-18.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$



Gráfica 6-19.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.91749	0.95200	0.03450	0.93164	0.97548	0.04384
0.25	0.72351	0.75179	0.02828	0.73436	0.76686	0.03250
0.50	0.50357	0.54075	0.03718	0.50965	0.55254	0.04289
0.75	0.23263	0.25971	0.02709	0.25670	0.28392	0.02722
0.95	0.02447	0.05092	0.02645	0.04340	0.08509	0.04169

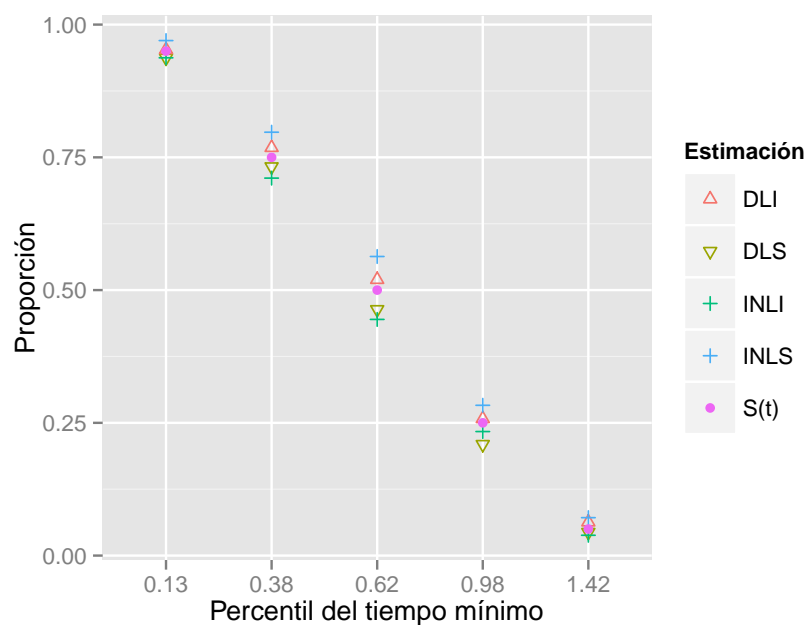
Tabla 6-27.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 50$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$



Gráfica 6-20.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.93716	0.95267	0.01550	0.93748	0.96990	0.03242
0.25	0.73239	0.76833	0.03594	0.71087	0.79731	0.08644
0.50	0.46347	0.51978	0.05631	0.44475	0.56325	0.11850
0.75	0.20947	0.25764	0.04817	0.23360	0.28303	0.04943
0.95	0.04392	0.06311	0.01919	0.03828	0.07141	0.03314

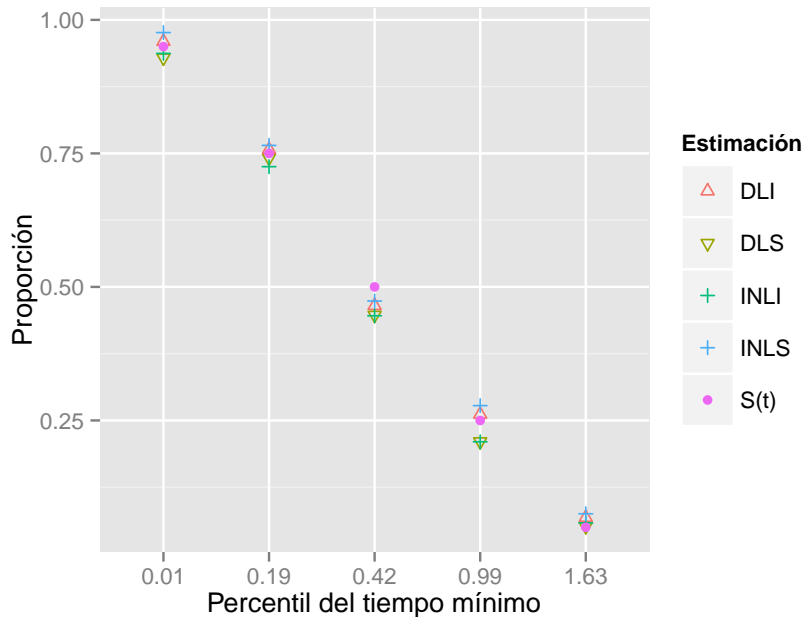
Tabla 6-28.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$



Gráfica 6-21.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 2$, $\beta_2 = 2$, $\beta_3 = 2$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.92894	0.95956	0,03062	0.93721	0.97623	0.03901
0.25	0.74156	0.75710	0.01554	0.72515	0.76498	0.03983
0.50	0.44727	0.46491	0.01763	0.44599	0.47366	0.02767
0.75	0.21042	0.26144	0.05102	0.21001	0.27774	0.06772
0.95	0.05172	0.06822	0.01649	0.05815	0.07524	0.01708

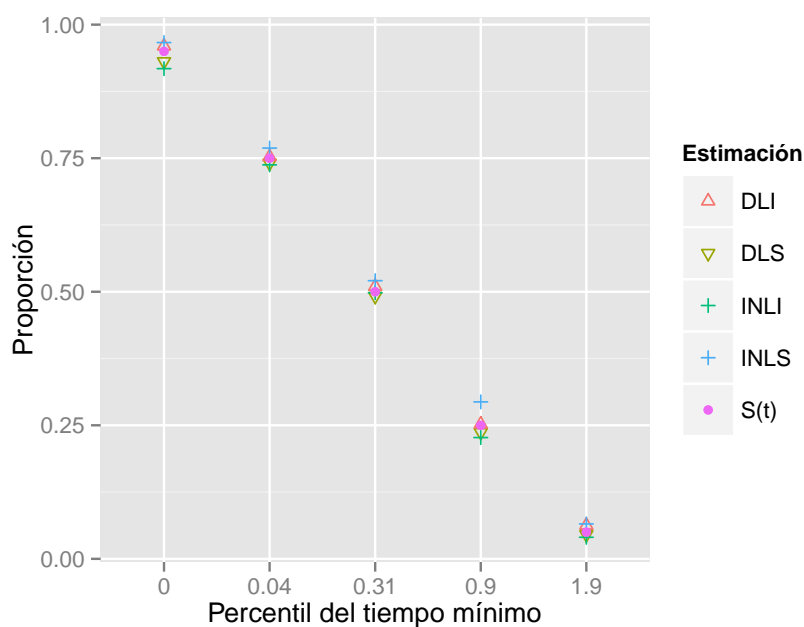
Tabla 6-29.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$



Gráfica 6-22.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.93085	0.95993	0.02908	0.91764	0.96643	0.04878
0.25	0.74100	0.75251	0.01151	0.73750	0.76893	0.03143
0.50	0.49161	0.50990	0.01829	0.49811	0.52073	0.02262
0.75	0.23683	0.25216	0.01533	0.22705	0.29384	0.06679
0.95	0.04545	0.06278	0.01733	0.04042	0.06534	0.02491

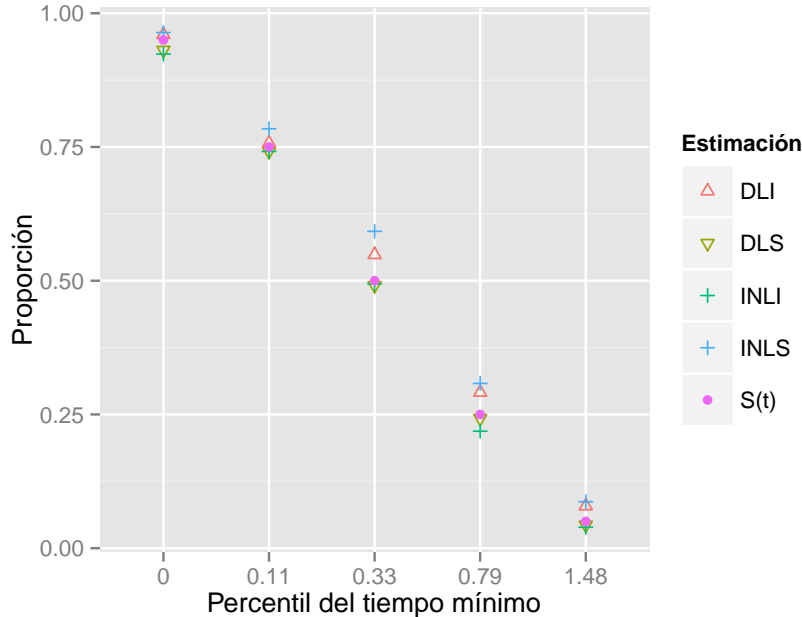
Tabla 6-30.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$



Gráfica 6-23.: Intervalos de confianzas empíricos aproximados del 95 % para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 0.5$, $\beta_2 = 1$, $\beta_3 = 1$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Percentil	$LI.S_D(t)$	$LS.S_D(t)$	Precisión	$LI.S_I(t)$	$LS.S_I(t)$	Precisión
0.05	0.93098	0.95998	0.02900	0.92364	0.96384	0.04020
0.25	0.74187	0.75708	0.01521	0.74166	0.78400	0.04235
0.50	0.49126	0.54856	0.05730	0.49390	0.59233	0.09843
0.75	0.24286	0.29120	0.04834	0.21875	0.30804	0.08929
0.95	0.04343	0.07888	0.03545	0.03915	0.08687	0.04772

Tabla 6-31.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$



Gráfica 6-24.: Intervalos de confianzas empíricos aproximados del 95% para $S(t_p)$. Modelo Weibull con parámetros $\beta_1 = 1.2$, $\beta_2 = 1$, $\beta_3 = 0.5$, $\eta_1 = 1$, $\eta_2 = 1$, $\eta_3 = 1$, $n = 100$ y $\tau_{12} = 0.2$, $\tau_{13} = 0.5$, $\tau_{23|1} = 0.8$ con $p = 0.05, 0.25, 0.50, 0.75, 0.95$

Parámetros				Percentiles				
n	β_1	β_2	β_3	0.05	0.25	0.50	0.75	0.95
50	2	2	2	0.4005	0.0876	0.7865	0.5555	0.0342
	1.5	1	1	0.0229	0.0212	0.6027	0.9786	0.9992
	0.5	1	1	0.9943	0.9807	0.1808	0.9768	0.6471
	1.2	1	0.5	0.8530	0.4541	0.7515	0.5051	0.9221
100	2	2	2	0.4905	0.8336	0.6317	0.4990	0.1947
	1.5	1	1	0.9521	0.7679	0.5640	0.4827	0.2651
	0.5	1	1	0.4867	0.2184	0.2907	0.4663	0.4624
	1.2	1	0.5	0.4973	0.6100	0.1044	0.3102	0.3036

Tabla 6-32.: Eficiencia relativa para el modelo Weibull con tamaños muestrales $n = 50, 100$ y parámetros de dependencia; $\tau_{12} = 0.2$, $\tau_{13} = 0.5$ y $\tau_{23|1} = 0.8$ y C-Vines

7. Conclusiones y Trabajos Futuros

- En gran mayoría de los casos estudiados el estimador bajo dependencia $S_D(t)$ presentó mayor eficiencia frente al estimador bajo independencia $S_I(t)$ de $S(t)$. Por lo tanto se recomienda utilizar el estimador $S_D(t)$ cuando no se tiene certeza de que los tiempos de falla sean independientes.
- Los intervalos de confianza para la función de sobrevivencia cuando se asume dependencia entre los modos de falla son más estrechos que los intervalos de confianza para la función de sobrevivencia cuando se asume independencia entre los modos de falla del modelo de riesgos trivariados con marginales Weibull.
- Para un modelo de riesgos competitivos trivariados, el método de combinación riesgos (extensión del estimador cópula gráfico) de es una buena técnica para estimar las funciones de distribución marginal y la función de sobrevivencia del tiempo mínimo en un modelo en riesgos competitivos cuando se tiene dos riesgos con igual dependencia y el otro independiente a los anteriores. También cuando se tienen dos riesgos con igual dependencia y el otro riesgo presenta una alta dependencia y por último, cuando los riesgos para un caso particular, presentan distintas dependencias (generados por medio de C-DVines).
- Posibles trabajos futuros, es analizar todo lo referente al parámetro de dependencia de la cópula por medio de elicitación (enfoque bayesiano). También puede ser de interés trabajar toda la teoría de combinación de riesgos para cópulas de dos parámetros. En el caso de los intervalos de confianza empíricos se trabajo mediante bootstrap, lo cual se pueden estimar por medio de *Jackknife* delete I o delete II.

A. Programa en R del Método Combinación de Riesgo y C-DVines

A continuación se presenta el código implementado para obtener todas las simulaciones de los escenarios que se trabajaron en este trabajo. Cabe resaltar que el código fue adaptado del propuesto por Lo & Wilke (2010) y Paz-Sabogal. et al. (2014).

```
# Programa adaptado para generar riesgos con distintas  
# dependencias para un modelo de riesgos competitivos  
# modelo trivariado
```

```
library(stats)  
library(utils)  
library(base)  
library(pracma)  
library(polylog)  
library(psi)  
library(sinc)  
  
library(mvtnorm)  
library(mnormt)  
library(scatterplot3d)  
library(pspline)  
library(sn)  
library(ADGofTest)  
library(stabledist)  
library(copula)  
library(splines)  
library(survival)  
library(cmprsk)  
library(mstate)  
library(CDVine)  
library(MASS)
```

```
library(igraph)

# Declare constants

BETA1 <-1.2
BETA2 <- 1
BETA3 <- 0.5
ETA<- 1
alpha<-2
ALPHA<-2
ALPHA1 <-1.86
ALPHA2<-5.7
ALPHA3<-18.29

lockBinding("BETA1", globalenv())
lockBinding("BETA2", globalenv())
lockBinding("BETA3", globalenv())
lockBinding("ETA", globalenv())
lockBinding("ALPHA", globalenv())
#lockBinding("TAO", globalenv())
lockBinding("ALPHA1", globalenv())
lockBinding("ALPHA2", globalenv())
lockBinding("ALPHA3", globalenv())

# Implementacion en R de la estimacion
# "Un Modelo Copula de riesgos competitivos Dependientes"
# propuesto po Simon Lo and Ralf Wilke 2010.

MCopGE = function(type, alpha, S, Q, kf)
{
  type = tolower(type)
  if (type == 'frank')
  {
    if ((missing(type)) || (missing(alpha)) || (missing(S))
|| (missing(Q)) || (missing(kf)))
    {
      stop("Requires exactly five input arguments:
```

```
MCopGE(TYPE, tau, S, Q, kf) with Q should be a matrix of CIC")
}

k = length(t(Q))
ns = length(S)
nq = length(Q[,1])

if (is.numeric(alpha) != 1)
{
  stop("Alpha debe ser un escalar")
}

if (alpha == 0)
{
  stop("El valor alpha debe ser diferente de cero")
}
else if (alpha < -200)
{
  stop("El valor alpha debe ser mayor que -200")
}
else if (alpha > 35)
{
  stop("El valor alfa debe ser inferior a 35")
}

if (is.numeric(kf) != 1)
{
  stop("kf debe ser un escalar: n??mero de marginales a estimar")
}

if (kf > k)
{
  stop("kf debe ser inferior a k: numero de marginales que
se estiman debe ser menor o igual al numero de CIC")
}

if (k < 1)
{
  stop("matrix de CIC necesita tener dimensiones mayores que 1")
}
```

```
}

if (ns != nq)
{
  stop("Número de filas de la matriz CIC y Supervivencias no coinciden")
}

# estimar las tres marginales de manera separadas

F = matrix(0, ns, kf)

for (nt in 1:kf)
{
  F[,nt] = suppressWarnings(frankest3redo(alpha, S, Q[,nt]))
}

#S12 is the 21-dimensions marginal survivor function of risk 1,2
}
return (F)
}
else
{
  stop("Only frank estimator is available")
}
}

frankest3redo = function(a, S, Q1)
{

  #given joint survival, CIC, and the
#copulas parameter, the marginal
#survival functions are identified

  tol = 1e-5

  floor = 0
  F1 = matrix(0, length(S), 1) # estimator of H1
  F2 = matrix(0, length(S), 1) # estimator of H2
```

```

I = matrix(0, length(S), 1) # integrated value at each it
lower = floor
upper = 1
E = matrix(0, 100, 1)
i = 1
w = 1

# j = time index,
# i = iteration index

# step one
# initial guess of h1(t1)

h1=S[1] #max as initial guess of H1 at t1.

# use iteration method to solve H2 given H1 and S. EQT 4.3
h2 = solve43(h1, S[1], a)

# w = difference of the estimated H1 in iteration
while (abs(w) > tol)
{
  myFun = function (x,y) ((y > as.numeric(h2/h1) * x) *
(-a*exp(-a*x)*exp(-a*y)*(exp(-a)-1)/(exp(-a)-1
+(exp(-a*x)-1)*(exp(-a*y)-1))^2))
  I[1] = quad2d(myFun, floor, h1, 0, 1)

  # it is the integrated value of the area between u \in [floor,h1] and v
  # \in[0,1] given that v > h2/h1*u
  # at t1, floor =0.

  #compare the estimated area with the CIC at t1. If the estimated value of
  #s1 and s2 are correct, they should be the same
  width = I[1] - Q1[1]

  #the first estimated value of h1 is recorded in E[1]
  E[i]= h1;

  if (i > 1)
  {
    J = which(E < h1 & E > 0)

```

```
if (length(J) > 1)
{
  lower = max(E[J])
}
else if (length(J) == 1)
{
  lower = E[J]
}

J = which(E > h1 & E < 1)

if (length(J) > 1)
{
  upper = min(E[J])
}
else if (length(J) == 1)
{
  upper = E[J]
}
}

rm(J)

# ==== use mid point iteration STEPS 2 & 3

if ((abs(w) > tol) && (width > 0)) # it means h1 too large, use a lower estimated val
{
  h1 = (h1 + lower)*0.5
  h2 = solve43(h1, S[1], a)
}
else if ((abs(w) > tol) && (width < 0)) # h1 too small
{
  h1 = (h1 + upper)*0.5
  h2 = solve43(h1, S[1], a)
}
```

```
    if (i > 1)
    {
        w = E[i]-E[i-1];
    }

    i = i + 1

}

F1[1]=h1 # estimated H1 at t1
F2[1]=h2 # estimated H2 at t1
floor= h1 # h1 at t1 becomes the lower value (floor) of h1 at t2
h1c=h1
h2c=as.numeric(h2)

# loop for higher j == step4

for (j in 2:length(S))
{
    i = 1
    lower = floor
    upper = 1
    E = matrix(0,100,1)
    w = 1 # initialize the estimation
    h1 = S[j]

    h2 = solve43(h1, S[j], a)

    while (abs(w) > tol)
    {

        myFun2 = function (u,v) (v > as.numeric(h2c) + (as.numeric(h2)-as.numeric(h2c))
/(as.numeric(h1) - as.numeric(h1c)) * (u - as.numeric(h1c))) *
(-a*exp(-a*u)*exp(-a*v)*(exp(-a)-1)/(exp(-a)-1+(exp(-a*u)-1)*(exp(-a*v)-1))^2)
        I[j] = quad2d(myFun2, floor, h1, 0, 1)
```

```
if (is.nan(I[j]) || is.na(I[j]))
{
  I[j] = 0
}

II = I[j] + I[j - 1]
width = II - Q1[j] # If estimation is correct, these two should be very closed
E[i] = h1

if (i > 1)
{
  J = which(E < h1 & E > 0)

  if (length(J) > 1)
  {
    lower = max(E[J])
  }
  else if (length(J) == 1)
  {
    lower = E[J]
  }

  J = which(E > h1 & E < 1)

  if (length(J) > 1)
  {
    upper = min(E[J])
  }
  else if (length(J) == 1)
  {
    upper = E[J]
  }

  rm(J)
}

if ((abs(w) > tol) && (width > 0)) # s1 too big
```

```

    {
      h1 = (h1 + lower)* 0.5
      h2 = solve43(h1, S[j], a)
    }
    else if ((abs(w) > tol) && (width < 0)) #s1 too small
    {
      h1 = (h1 + upper)* 0.5
      h2 = solve43(h1, S[j], a)
    }

    if (i > 1)
    {
      w = E[i] - E [i - 1]
    }

    i = i + 1

    # here i set the length of E as the maximun number of iterations
    if (i > length(E))
    {
      stop("not converging")
    }

    #cat("i:",i,"; h1:", h1, "; h2: ",h2, "; I[j]:",I[j],"; II ", II ,"
; width: ", width, "; E[i]: ", E[i], "; upper: ", upper, "; lower: "
, lower, "; w: ", w, " \n ")

  }

F1[j] = h1
F2[j] = h2
floor = h1
h1c = h1
h2c = h2

I[j] = I[j] + I [j - 1]

```

```
}

return (F1)

}

e43 = function(h1,h2,a)
{
  #Survival given u,v, and copula
  s = 1-h1-h2-1/a*log(1+(exp(-a*h1)-1)*(exp(-a*h2)-1)/(exp(-a)-1))
  return (s)
  list(s)
}

solve43 = function (h1,s,a)
{

  lower = matrix(0, length(s), length(s))
  upper = matrix(1, length(s), length(s))
  width = 1
  tol = 1e-12

  while (width > tol)
  {
    h2 = .5*(lower + upper)
    lo = e43(h1,h2,a) > s
    lower[lo] = h2[lo]
    upper[!lo] = h2[!lo]
    width = 0.5 * width
  }
  return (h2)
}

calculate_CIC_S = function(n)
{
```

```
CIC = matrix(0, 1, 1)
flag = TRUE;
datos1 = 0
datos = 0
y = NA

while (flag == TRUE)
{

  while (length(which(is.na(y))) != 0)
  {
    d = 3      # Generacion de las riesgos con distintas dependencias
    dd = d*(d-1)/2
    fam1 = rep(5,dd)
    par1 = c(BiCopTau2Par(family=5,tau=0.20),BiCopTau2Par(family=5,tau=0.50)
,BiCopTau2Par(family=5,tau=0.8))
    N = 100
    yo<- CDVineSim(N,fam1,par1,type=1)
    y1=(-log(1-yo[,1]))^(1/BETA1)
    y2=(-log(1-yo[,2]))^(1/BETA2)
    y3=(-log(1-yo[,3]))^(1/BETA3)
    y=cbind(y1,y2,y3)

  }

  minimo<-apply(y,1,min)
  indicadora<-apply(y,1,which.min)

  datos1=data.frame(minimo,indicadora)
  datos=datos1[order(datos1[,1]),]
  minmar1<-datos$minimo[datos$indicadora==1]
  posmar1<-which(datos$indicadora == 1)
  minmar2 <- datos$minimo[datos$indicadora==2]
  posmar2<-which(datos$indicadora == 2)
  minmar3<-datos$minimo[datos$indicadora==3]
  posmar3<-which(datos$indicadora == 3)
  conteo=as.matrix(table(datos1$indicadora))
```

```

print(conteo)

  if ((nrow(conteo) > 2) && (length(posmar1) > 1)
&& (length(posmar2) > 1) && (length(posmar3) > 1))
  {
    flag = FALSE
  }

  y = NA
}

fit=Cuminc(datos[,1],datos[,2])
CIC= cbind(fit$CI.1[-1],fit$CI.2[-1],fit$CI.3[-1])
S=fit$Surv[-1]
#print(datos[,1])

value = data.frame(CIC = CIC, S = S, minT=datos[,1],ind=datos[,2])
return (value)
#print(value)
}

#n Tama~{n}o de muestra
principal = function (n, kf)
{

  #if (is.numeric(alpha) != 1)
  #{
    # stop("Alpha must be a scalar")
  #}

# if (alpha == 0)
# {
  # stop("The alpha value must be different from zero")
  #}

```

```
#else if (alpha < -20)
#{
# stop("The alpha value must be greater than -20")
#}
#else if (alpha > 35)
#{
# stop("The alpha value must be less than 35")
#}

if (is.numeric(kf) != 1)
{
  stop("kf must be a scalar: Numero de marginales estimadas")
}

#print('***** Calculo de CIC y S *****')

resultado = suppressWarnings(calculate_CIC_S(n))
CIC = resultado[,1:3]
S = resultado[,4]
minT=resultado[,5]
ind=resultado[,6]

mcpoge = MCopGE('frank', alpha, S, CIC, kf)

value = data.frame(mcpoge = mcpoge, minT = minT, ind=ind)

return(value)
}

MDF=principal(100,3)
```


Bibliografía

- Aas, K., Czado, C., Frigessi, A. & Bakken, H. (2009), 'Pair-copula constructions of multiple dependence', *Mathematics and Economics* **44**(2), 182–198.
- Brechman, E. C. & Shepsmeier, U. (2013), 'Modeling dependence with c- and d-vine copula: The r package cdvine', *Journal of statistical software* **52**.
- Czado, C., Schepsmeier, U. & Min, A. (2012), 'Maximum likelihood estimation of mixed c-vines with application to exchange rates', *Statistical Modelling* **12**(3), 229–255.
- Devroye, L. (1986), *Non uniform random variate generation*, Springer Verlag, New York.
- Escarela, G. & Carriere, J. (2003), 'Fitting competing risks with an assumed copula', *Statistical Methods in Medical Research* **12**, 33–349.
- Genest, C., Ghousi, K. & Rivest, L. (1995), 'A semiparametric estimation procedure of dependence parameters in multivariate families of distributions', *Biometrika* **82**(3), 543–552.
- Genest, C. & MacKay, J. (1986), 'The joy of copulas: Bivariate distributions with uniform marginals', *The American Statistician* **40**(4), 280–283.
- Kurowicka, D. & Cooke, R. (2006), *Uncertainty analysis with high dimensional dependence modelling*, wiley series in probability and statistics edn, John Wiley & Sons, New York.
- Kurowicka, D. & Joe, H. (2011), *Dependence Modelling: Vine Copula Handbook*, World Scientific Publishing Co., Singapore.
- Lo, S. & Wilke, R. A. (2010), 'A copula model for dependent competing risks', *Journal of the Royal Statistical Society* **59**(2), 359–376.
- Lu, J. & Bhattacharyya, G. (1990), 'Some new constructions of bivariate weibull models', *Annals of the Institute of Statistical Mathematics* **42**(3), 543–559.
- Meeker, W. Q., Escobar, L. A. & Hong, Y. (2009), 'Using accelerated life tests results to predict product field reliability', *Technometrics* **51**(2), 146–161.
- Nelsen, R. B. (2006), *An introduction to copulas*, second edn, Springer, New York.

- Paz-Sabogal., M. C., Lopera-Gómez., C. M. & Yáñez-Canal., S. (2014), 'Extensión del estimador cópula gráfico para un modelo con más de dos riesgos competitivos dependientes', *Revista Ingeniería y Competitividad-Univalle* **16**(1).
- Pintilie, M. (2006), *Competing Risks: Apractical Perspective*, Jhon Wiley and Sons.
- Schweizer, B. & Sklar, A. (1983), *Probabilistic Metric Spaces*, Dover Publications, New York.
- Tsiatis, A. (1975), 'A nonidentifiability aspect of the problem of competing risks', *Proc. Natl. Acad. Sci. USA* **72**(1), 20–22.
- Zheng, M. & Klein, J. P. (1995), 'Estimates of marginal survival for dependent competing risks based on an assumed copula', *Biometrika* **82**(1), 127–138.