



UNIVERSIDAD NACIONAL DE COLOMBIA

Extensión del Estimador Cópula Gráfico para un Modelo con más de dos Riesgos Competitivos Dependientes

María Carolina Paz Sabogal

Universidad Nacional de Colombia
Facultad de Ciencias
Escuela de Estadística
Medellín, Colombia
2012

Extensión del Estimador Cópula Gráfico para un Modelo con más de dos Riesgos Competitivos Dependientes

María Carolina Paz Sabogal

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

Director:

Sergio Yáñez Canal

M.Sc Mathematics New York University

M.Sc Statistics Loughboroygh University

Línea de Investigación:

Confiabilidad Industrial

Universidad Nacional de Colombia

Facultad de Ciencias

Escuela de Estadística

Medellín, Colombia

2012

A mi familia

Agradecimientos

El presente trabajo es un esfuerzo en el cual, directa o indirectamente, participaron varias personas como profesores, compañeros y demás integrantes de la Escuela de Estadística que tuvieron su aporte leyendo, opinando, corrigiendo para el desarrollo de esta investigación.

Debo agradecer de manera especial y sincera al Profesor Sergio Yáñez Canal por aceptarme para realizar esta tesis de Maestría bajo su dirección. Las ideas propias, siempre enmarcadas en su orientación y rigurosidad, han sido la clave del buen trabajo que hemos realizado juntos, el cual no se puede concebir sin su siempre oportuna participación.

Resumen

En un modelo de riesgos competitivos dependientes es imposible determinar las distribuciones marginales y la distribución conjunta a partir solamente de los datos de riesgos competitivos. Esta situación se conoce como el problema de identificabilidad. Zheng and Klein (1995) proponen el estimador cópula gráfico como solución al problema de identificabilidad para dos riesgos competitivos. Para ello asumen una estructura de dependencia usando una cópula para la distribución conjunta entre los tiempos de falla y su parámetro de dependencia conocido. En el caso de un modelo con más de dos riesgos competitivos, Lo and Wilke (2010) proponen el método de combinación de riesgo (“ risk pooling method ”) como una extensión del estimador cópula gráfico cuando la cópula es Arquimediana. En este trabajo para el caso trivariado, se compara la función de sobrevivencia conjunta verdadera, con la función de sobrevivencia conjunta estimada asumiendo independencia entre los tiempos de falla y la función de sobrevivencia estimada mediante el método de combinación de riesgos. Estas comparaciones se realizan vía simulación teniendo en cuenta tiempos de falla asociados a una distribución Weibull y lognormal multivariada y diferentes niveles de dependencia entre los tiempos de falla. Se concluye que el estimador asumiendo independencia es menos eficiente que el estimador de la función de sobrevivencia utilizando el método de combinación de riesgos.

Palabras clave: Método de combinación de riesgos, Riesgos competitivos, Cópula Arquimediana, Identificabilidad.

Abstract

In a dependent competing risks model is impossible to determine the marginal distributions and the joint distribution from the competing risks data alone. This is known as the identifiability problem. Zheng and Klein (1995) propose the copula graphic estimator as a solution to the identifiability problem for two competing risks. For that, they assume a dependence structure using a copula for the joint distribution of failure times and its dependence parameter known. Lo and Wilke (2010) propose the risk pooling method as an extension of the copula graphic estimator when the copula is Archimedean. This research for the trivariate case, is compared the true joint survival function, with joint survival function estimated assuming independence among failure times and the survival function estimated by the risk pooling method. These comparisons are performed via simulation considering failure times associated with multivariate Weibull and lognormal distributions and different levels of dependence between failure times. We conclude that the estimator assuming independence is less efficient than the estimator of the survival function using the risk pooling method.

Key words: Risk pooling method, Competing risks, Archimedean copula, Identifiability.

Contenido

Agradecimientos	vii
Resumen	ix
1 Introducción	3
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
3.1.1 Propiedades de $C(x, y)$	11
3.1.2 La cópula Frank	14
3.2 Medidas de Concordancia	15
3.2.1 τ de Kendall	15
4 Extensión del Estimador Cópula Gráfico: Método de Combinación de Riesgos	23
4.1 Estimador Cópula Gráfico	23
4.1.1 Construcción del Estimador Cópula Gráfico	24
4.1.2 Algoritmo Cópula Gráfico	26
4.2 Extensión del Estimador Cópula Gráfico	28
4.2.1 Chequeo Empírico del Método de Combinación de Riesgos	33
5 Generación de Tiempos de Falla Weibull y Lognormales	43
5.1 Obtención de Tiempos de Falla Lognormales	43
5.1.1 Generación de Tiempos de Falla Lognormales Dependientes	45
5.2 Obtención de Tiempos de Falla Weibull	47
5.2.1 Generación de Tiempos de Falla Weibull Dependientes	48
5.3 Esquema de Simulación	48
6 Análisis de Resultados	58
6.1 Casos Weibull	58

6.1.1	Caso Weibull: $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	58
6.1.2	Caso Weibull: $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	63
6.1.3	Caso Weibull: $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	68
6.1.4	Caso Weibull: $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	72
6.2	Casos Lognormal	77
6.2.1	Caso Lognormal: $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8$	77
6.2.2	Caso Lognormal: $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1$	82
6.2.3	Caso Lognormal: $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1$	87
7	Conclusiones y Trabajo Futuro	93
A	Anexo: Funcionamiento del Estimador Cópula Gráfico	94
A.1	Comparación entre las funciones de distribución marginales verdaderas y las estimadas por el algoritmo cópula gráfico	94
A.1.1	Función de Distribución Marginal de X	95
A.1.2	Función de Distribución Marginal de Y	102
A.2	Comparación entre la Función de Supervivencia Verdadera y la Estimada por el Algoritmo Cópula Gráfico	106
B	Anexo: Algoritmo y Programa en R del Método de Combinación de Riesgos	116
	Bibliografía	125

Lista de Figuras

2-1	Sistema con dos componentes en serie.	5
4-1	Relación entre $F_1(t)$ y $F_2(t)$ en el cuadrado unitario.	24
4-2	Transformación para la cópula.	25
4-3	Representación gráfica del modelo de riesgos competitivos.	31
4-4	Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 1, $\beta_1 = 2.4, \eta_1 = 1, \tau = 0.2$, para $n = 50$ y $n = 200$	35
4-5	Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 2, $\beta_2 = 2.6, \eta_2 = 1, \tau = 0.2$, para $n = 50$ y $n = 200$	35
4-6	Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 3, $\beta_3 = 2.8, \eta_3 = 1, \tau = 0.2$, para $n = 50$ y $n = 200$	36
4-7	Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 1, $\beta_1 = 2.4, \eta_1 = 1, \tau = 0.5$, para $n = 50$ y $n = 200$	37
4-8	Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 2, $\beta_2 = 2.6, \eta_2 = 1, \tau = 0.5$, para $n = 50$ y $n = 200$	38
4-9	Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 3, $\beta_3 = 2.8, \eta_3 = 1, \tau = 0.5$, para $n = 50$ y $n = 200$	38
4-10	Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 1, $\beta_1 = 2.4, \eta_1 = 1, \tau = 0.8$, para $n = 50$ y $n = 200$	40
4-11	Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 2, $\beta_2 = 2.6, \eta_2 = 1, \tau = 0.8$, para $n = 50$ y $n = 200$	40
4-12	Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 3, $\beta_3 = 2.8, \eta_3 = 1, \tau = 0.8$, para $n = 50$ y $n = 200$	41
5-1	Relación entre el Coeficiente de Correlación Lineal ρ y el ρ_L , $\sigma_1 = 2, \sigma_2 = 1$	44
5-2	Comportamiento del coeficiente de correlación para diferentes valores de los parámetros de la distribución Lognormal.	45
5-3	Gráfico de probabilidad Weibull para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	50
5-4	Gráfico de probabilidad Weibull para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	51
5-5	Gráfico de probabilidad Weibull para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	52

5-6	Gráfico de probabilidad Weibull para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	53
5-7	Gráfico de probabilidad Weibull para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = \sigma_2 = \sigma_3 = 1$	54
5-8	Gráfico de probabilidad Weibull para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8$	55
5-9	Gráfico de probabilidad Weibull para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8$	56
6-1	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	59
6-2	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	61
6-3	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	62
6-4	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	64
6-5	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	65
6-6	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	67
6-7	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	69
6-8	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	70
6-9	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	72
6-10	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	74

6-11	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	75
6-12	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	77
6-13	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0, \sigma_1 = \sigma_2 = \sigma_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	79
6-14	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0, \sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	80
6-15	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0, \sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	82
6-16	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0, \sigma_1 = \sigma_2 = \sigma_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	84
6-17	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0, \sigma_1 = \sigma_2 = \sigma_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	85
6-18	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0, \sigma_1 = \sigma_2 = \sigma_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	87
6-19	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	89
6-20	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	90
6-21	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	92
A-1	Gráfico de probabilidad Weibull para la función de distribución marginal de $X, \beta_1 = 2, \eta_1 = 1, \tau = 0.2$, para $n = 50$ y $n = 200$	96
A-2	Gráfico de probabilidad Weibull para la función de distribución marginal de $X, \beta_1 = 2, \eta_1 = 1, \tau = 0.5$, para $n = 50$ y $n = 200$	97
A-3	Gráfico de probabilidad Weibull para la función de distribución marginal de $X, \beta_1 = 2, \eta_1 = 1, \tau = 0.8$, para $n = 50$ y $n = 200$	97

A-4	Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2, \eta_1 = 1, \tau = 0.9$, para $n = 50$ y $n = 200$	98
A-5	Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2.5, \eta_1 = 1, \tau = 0.2$, para $n = 50$ y $n = 200$	99
A-6	Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2.5, \eta_1 = 1, \tau = 0.5$, para $n = 50$ y $n = 200$	100
A-7	Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2.5, \eta_1 = 1, \tau = 0.8$, para $n = 50$ y $n = 200$	100
A-8	Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2.5, \eta_1 = 1, \tau = 0.9$, para $n = 50$ y $n = 200$	101
A-9	Gráfico de probabilidad Weibull para la función de distribución marginal de Y , $\beta_2 = 2, \eta_2 = 1, \tau = 0.2$, para $n = 50$ y $n = 200$	103
A-10	Gráfico de probabilidad Weibull para la función de distribución marginal de Y , $\beta_2 = 2, \eta_2 = 1, \tau = 0.5$, para $n = 50$ y $n = 200$	104
A-11	Gráfico de probabilidad Weibull para la función de distribución marginal de Y , $\beta_2 = 2, \eta_2 = 1, \tau = 0.8$, para $n = 50$ y $n = 200$	104
A-12	Gráfico de probabilidad Weibull para la función de distribución marginal de Y , $\beta_2 = 2, \eta_2 = 1, \tau = 0.9$, para $n = 50$ y $n = 200$	105
A-13	Función de sobrevivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de sobrevivencia estimada $\widehat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.2$, para $n = 50$	108
A-14	Función de sobrevivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de sobrevivencia estimada $\widehat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.5$, para $n = 50$	109
A-15	Función de sobrevivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de sobrevivencia estimada $\widehat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.8$, para $n = 50$	109
A-16	Función de sobrevivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de sobrevivencia estimada $\widehat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.9$, para $n = 50$	110
A-17	Función de sobrevivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de sobrevivencia estimada $\widehat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.2$, para $n = 200$	111
A-18	Función de sobrevivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de sobrevivencia estimada $\widehat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.5$, para $n = 200$	112
A-19	Función de sobrevivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de sobrevivencia estimada $\widehat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.8$, para $n = 200$	113

A-20 Función de sobrevivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de sobrevivencia estimada $\widehat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.9$, para $n = 200$	114
---	-----

Lista de Tablas

3-1	Familia de cópulas Arquimedianas	13
4-1	Diseño de simulación.	34
4-2	Error cuadrático medio para las funciones de distribución marginales estimadas en los percentiles t_p , con $p = 0.25, 0.50, 0.95$ para $\tau = 0.2$	37
4-3	Error cuadrático medio para las funciones de distribución marginales estimadas en los percentiles t_p , con $p = 0.25, 0.50, 0.95$ para $\tau = 0.5$	39
4-4	Error cuadrático medio para las funciones de distribución marginales estimadas en los percentiles t_p , con $p = 0.25, 0.50, 0.95$ para $\tau = 0.8$	41
5-1	Porcentaje de falla para el modelo Weibull $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	50
5-2	Porcentaje de falla para el modelo Weibull $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	51
5-3	Porcentaje de falla para el modelo Weibull $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	52
5-4	Porcentaje de falla para el modelo Weibull $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$	53
5-5	Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1$	54
5-6	Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8$	55
5-7	Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1$	56
6-1	Porcentaje de falla para el modelo Weibull $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$	58
6-2	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	59
6-3	Porcentaje de falla para el modelo Weibull $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$	60

6-4	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	60
6-5	Porcentaje de falla para el modelo Weibull $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$	61
6-6	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	62
6-7	Eficiencia relativa.	63
6-8	Porcentaje de falla para el modelo Weibull $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$	63
6-9	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	63
6-10	Porcentaje de falla para el modelo Weibull $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$	64
6-11	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	65
6-12	Porcentaje de falla para el modelo Weibull $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$	66
6-13	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	66
6-14	Eficiencia relativa.	68
6-15	Porcentaje de falla para el modelo Weibull $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$	68
6-16	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	68
6-17	Porcentaje de falla para el modelo Weibull $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$	69
6-18	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	70
6-19	Porcentaje de falla para el modelo Weibull $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$	71
6-20	Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	71

6-21 Eficiencia relativa.	72
6-22 Porcentaje de falla para el modelo Weibull $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$	73
6-23 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	73
6-24 Porcentaje de falla para el modelo Weibull $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 1.2, \eta_1 = 1, \eta_2 = 0.5, \eta_3 = 1, \tau = 0.5$	74
6-25 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	75
6-26 Porcentaje de falla para el modelo Weibull $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$	76
6-27 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	76
6-28 Eficiencia relativa.	77
6-29 Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.2$	78
6-30 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	78
6-31 Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.5$	79
6-32 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	80
6-33 Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.8$	81
6-34 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	81
6-35 Eficiencia relativa.	82
6-36 Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.2$	83
6-37 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	83
6-38 Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.5$	84

6-39 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	85
6-40 Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.8$	86
6-41 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	86
6-42 Eficiencia relativa.	87
6-43 Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.2$	88
6-44 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	88
6-45 Porcentaje de falla para el modelo Weibull $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.5$	89
6-46 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	90
6-47 Porcentaje de falla para el modelo Weibull $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.8$	91
6-48 Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$	91
6-49 Eficiencia relativa.	92
A-1 Error cuadrático medio para la función de distribución marginal de X estimada en los percentiles t_p , con $p = 0.25, 0.5, 0.95$	102
A-2 Error cuadrático medio para la función de distribución marginal estimada de Y en los percentiles t_p , con $p = 0.25, 0.50, 0.95$	106
A-3 Error cuadrático medio para la función de sobrevivencia conjunta de la Weibull bivariada estimada en los percentiles t_p , con $p = 0.05, 0.25, 0.50, 0.75, 0.95$, para $n = 50$	115
A-4 Error cuadrático medio para la función de sobrevivencia conjunta de la Weibull bivariada estimada en los percentiles t_p , con $p = 0.05, 0.25, 0.50, 0.75, 0.95$, para $n = 200$	115

1 Introducción

El tiempo de falla de un sistema con varios modos de falla puede ser modelado como un sistema en serie o un modelo de riesgos competitivos. Cada unidad tiene un tiempo potencial de falla. El tiempo de falla observado es el mínimo de esos tiempos potenciales individuales.

En riesgos competitivos bajo el supuesto de independencia entre los tiempos de falla, los datos observados proporcionan información suficiente para determinar de manera única las funciones de supervivencia marginal. Sin embargo, el supuesto de independencia no siempre se cumple. Cuando existe dependencia entre los tiempos de falla, queda imposible identificar las distribuciones marginales y la distribución conjunta a partir solamente de los datos de riesgos competitivos, este se conoce como el problema de identificabilidad.

En 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 (Tsiatis, 1975). Inclusive en el caso paramétrico, los datos pueden contener poca información sobre el coeficiente de asociación entre las variables y es necesario hacer algunos supuestos al respecto (Meeker, Escobar, and Hong, 2009). Para el caso no paramétrico Zheng and Klein (1995) proponen el estimador cópula gráfico como una solución al problema de identificabilidad. Para ello utilizan cópulas como una función no paramétrica que captura la dependencia entre dos variables aleatorias. El problema de identificabilidad es resuelto con los datos de riesgos competitivos y bajo el supuesto de que la cópula y su parámetro de asociación son conocidos.

Yáñez, Manotas, Lopera, and Jaramillo (2008) estudiaron para tiempos de falla con distribución Weibull y lognormal, el efecto de la estimación de la función de supervivencia para el tiempo mínimo asumiendo independencia entre los tiempos de falla que realmente son dependientes. Yáñez, Brango, Lopera, and Jaramillo (2011) estimaron la función de supervivencia para el tiempo mínimo asumiendo dependencia entre los tiempos de falla, para un modelo con dos riesgos competitivos, mediante el estimador cópula gráfico propuesto por Zheng and Klein (1995) y realizaron un estudio de simulación, para compararlo con el estimador propuesto por Yáñez, Manotas, Lopera, and Jaramillo (2008).

Lo and Wilke (2010) proponen el método de combinación de riesgos como una extensión del estimador cópula gráfico cuando la cópula es Arquimediana para uno modelo con más de

dos riesgos competitivos. La propuesta de Lo and Wilke (2010) es computacionalmente más eficiente que las propuestas por Carriere (1995) y Rivest and Wells (2001). Lo and Wilke (2010) proponen esta metodología basada en cópulas a la flexibilidad que dichos modelos permiten especificar la estructura de dependencia entre variables aleatorias competitivas.

En el capítulo 2 se definen algunos conceptos fundamentales de la teoría de riesgos competitivos, el problema de identificabilidad y como se estima la función de sobrevivencia del tiempo mínimo asumiendo independencia entre los tiempos de falla. En el capítulo 3 se presenta todo lo relacionado con la teoría de las cópulas y como se emplean para modelar la estructura de dependencia para los tiempos de falla trivariados.

En el capítulo 4 y 5 se describe el método de combinación de riesgos con el que se estima la función de sobrevivencia para el tiempo mínimo asumiendo dependencia y un chequeo empírico de este método, y la metodología utilizada en este estudio para la obtención de los resultados.

En el capítulo 6 se muestran los resultados obtenidos asociados a cada uno de los casos Weibull y lognormal. Se exhiben gráficos correspondientes a intervalos de confianza empíricos en los percentiles t_p , con $p = 0.05, 0.25, 0.50, 0.75, 0.95$, basados en la simulación Monte Carlo aproximados del 95% para los estimadores de la función de sobrevivencia para el tiempo mínimo asumiendo independencia y dependencia. Además, se presentan tablas correspondientes a la eficiencia relativa para comparar puntualmente los dos estimadores.

En el capítulo 7 se presentan las conclusiones del estudio y finalmente en los anexos se muestra un chequeo empírico del estimador cópula gráfico, el algoritmo y el programa implementado en el software estadístico R del método de combinación de riesgos.

2 Riesgos Competitivos

El tiempo de falla de un sistema con dos modos de falla puede ser modelado como un sistema en serie o un modelo de riesgos competitivos, como se ilustra en la Figura 2-1. Cada unidad tiene un tiempo potencial de falla. El tiempo de falla observado es el mínimo de esos tiempos potenciales individuales.

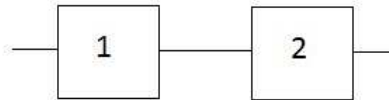


Figura 2-1: Sistema con dos componentes en serie.

Los datos en riesgos competitivos son presentados como una variable aleatoria bivariada de la forma (T, δ) , donde T es el tiempo mínimo de falla $T = \min(X, Y)$ y δ es una variable discreta, δ toma el valor de 0 cuando la observación es censurada, 1 cuando la falla se presenta por el primer modo de falla y 2 cuando la falla se presenta por el segundo modo de falla.

Para estudiar el modelo de riesgos competitivos es necesario definir algunos conceptos fundamentales.

2.1 Conceptos Fundamentales

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

$$G_i(t) = P(T \leq t, \delta = i). \quad (2-1)$$

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

$$F(t) = P(T \leq t) = \sum_{i=1}^2 P(T \leq t, \delta = i) = \sum_{i=1}^2 G_i(t). \quad (2-2)$$

Nótese que cuando no hay riesgos competitivos la función de distribución total se encuentra en el intervalo $[0,1]$. En contraste, en presencia de riesgos competitivos la función de subdistribución sólo puede tomar valores hasta $P(\delta = i)$ debido a que

$$\lim_{t \rightarrow \infty} G_i(t) = P(\delta = i). \quad (2-3)$$

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

$$\bar{G}_i(t) = P(T > t, \delta = i). \quad (2-4)$$

La función de subdensidad para el modo de falla i y T continuo, se define como

$$g_i(t) = \frac{dG_i(t)}{dt} = -\frac{d\bar{G}_i(t)}{dt}. \quad (2-5)$$

Para mayores detalles ver Pintilie (2006).

2.2 Problema de Identificabilidad

La aproximación tradicional para especificar el modelo de riesgos competitivos es vía tiempos de falla latentes. Se dice que los tiempos de falla son latentes debido a que, por ejemplo, para el caso bivariado se tienen los tiempos de falla potenciales X, Y asociados a los dos modos de falla, de manera que si el primero ocurre, el segundo no puede ser observado. Así $T = \min(X, Y)$ determina el tiempo de falla del sistema completo y una vez que el sistema haya fallado, el otro tiempo de falla no puede ser observado. Se supone que la pareja (X, Y) tiene una distribución multivariada $F(x, y) = P(X \leq x, Y \leq y)$ y las funciones de distribución marginales para los dos modos de falla son de la forma $F_1(t) = P(X \leq t)$ y $F_2(t) = P(Y \leq t)$, respectivamente.

Un problema de interés en riesgos competitivos es identificar las distribuciones marginales de las variables a partir de los datos de riesgos competitivos. Cuando los tiempos de falla potenciales son independientes las distribuciones marginales son identificables y corresponden a las funciones de distribución. Si este supuesto no se cumple, no es posible identificar las distribuciones marginales y la distribución conjunta de los tiempos potenciales a partir

únicamente de los datos obtenidos en riesgos competitivos, es decir, existen muchas funciones de distribución conjuntas diferentes que comparten las mismas funciones de subdistribución, esto es el problema de identificabilidad.

Existen distintas soluciones al problema de identificabilidad cuando se tienen datos de la forma (T, δ) y se hace necesario identificar el modelo conjunto $F(x, y)$, entre ellas se encuentran:

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

2.3 Estimación de la Función de Supervivencia bajo Independencia

Una estructura en serie con s componentes funciona sí y sólo sí, todas las componentes funcionan. Para un sistema en serie con tres componentes independientes, la función de distribución para el tiempo mínimo T es

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

donde (X, Y, Z) son los tiempos de vida asociados a las tres componentes y $F_i, i = 1, 2, 3$ son respectivamente las funciones de distribución para los componentes y $T = \min\{X, Y, Z\}$ es el tiempo de falla observado del sistema.

La función de supervivencia $S(t) = 1 - F(t)$, para el caso de independencia es

$$\begin{aligned} S(t) &= 1 - \{[1 - [1 - F_1(t)][1 - F_2(t)][1 - F_3(t)]]\} \\ &= [1 - F_1(t)][1 - F_2(t)][1 - F_3(t)] \\ &= S_1(t)S_2(t)S_3(t), \end{aligned}$$

donde $S_i(t), i = 1, 2, 3$ son las funciones de sobrevivencia de las tres componentes.

Para un sistema de tres componentes en serie y tiempos de falla dependientes

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

En este caso, la evaluación debe hacerse con respecto a la distribución conjunta de X, Y y Z , que incluye el parámetro de dependencia.

La función de sobrevivencia para este caso es

$$S(t) = P(T > t) = P(X > t, Y > t, Z > t).$$

El estimador de la función de sobrevivencia asumiendo independencia $S^*(t)$ se denotará $\widehat{S}^*(t)$, que es un estimador de la sobrevivencia del tiempo mínimo de falla de un sistema con tres modos de falla que compiten. Este se obtiene como el producto de las funciones de sobrevivencia 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

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

donde $\widehat{S}_1(t), \widehat{S}_2(t)$ y $\widehat{S}_3(t)$ se suponen son las funciones de sobrevivencia marginales estimadas para cada modo de falla.

3 Cópulas

Las cópulas se han convertido en una herramienta popular de modelamiento multivariado en muchos campos donde la dependencia multivariada es de gran interés y la usual normalidad multivariada está en cuestión. En finanzas, las cópulas son usadas en la asignación de activos, evaluación de crédito, el modelamiento de incumplimiento y la gestión de riesgos. En estudios biomédicos, las cópulas son utilizadas en el modelamiento de la correlación de tiempos de eventos correlacionados y riesgos en competencia. En ingeniería, las cópulas son usadas en procesos de control multivariados y en el modelamiento hidrológico (Yan, 2007).

Definición: Sea X y Y variables aleatorias con función de distribución conjunta F , con distribuciones marginales F_1 y F_2 respectivamente. X y Y son dos variables uniformes si F_1 y F_2 son continuas. De esta manera, si las marginales F_1 y F_2 de la distribución bivariada F son continuas, entonces existe una cópula única, la cual es una función de distribución acumulada, con sus marginales uniformes (Drouet Mari and Kotz, 2001). Formalmente una función $C : [0, 1]^2 \rightarrow [0, 1]$ tal que

$$F(x, y) = C(F_1(x), F_2(y)),$$

es una cópula.

Las cópulas bidimensionales son funciones bivariadas que unen dos funciones de distribución univariadas para construir distribuciones bivariadas continuas. La cópula es una función que permite modelar la estructura de dependencia de dos variables aleatorias.

3.1 Cópulas Arquimedianas

Genest and Mackay (1986) hacen referencia a una clase de funciones Φ , $\phi : [0, 1] \rightarrow [0, \infty]$ la cual tiene dos derivadas continuas en $(0, 1)$ y satisface:

$$\phi(1) = 0, \phi'(t) < 0, \phi''(t) < 0. \tag{3-1}$$

Para todo $0 < t < 1$. Las condiciones (3-1) son suficientes para garantizar que ϕ tiene inversa ϕ^{-1} y que también tiene dos derivadas. Miembros típicos de esta clase Φ incluyen:

- $\phi(t) = -\log(t)$.

- $\phi(t) = (1 - t)^\alpha$ donde $\alpha > 1$.
- $\phi(t) = t^{-\alpha} - 1$ donde $\alpha > 1$.

Cada miembro ϕ de la clase Φ genera una función de distribución bivariada para el par (X, Y) como se muestra a continuación:

$$\begin{aligned} C(x, y) &= \phi^{-1}[\phi(x) + \phi(y)], \quad \text{si } \phi(x) + \phi(y) \leq \phi(0) \\ &= 0 \quad \text{en otro caso.} \end{aligned} \tag{3-2}$$

Para encontrar la densidad $c(x, y)$ asociada con (3-2), sea $\phi(C) = \phi(x) + \phi(y)$ al diferenciar $C(x, y)$ con respecto a x y luego con respecto a y , se obtiene

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

y

$$\phi''(C) \frac{\partial C}{\partial x} \frac{\partial C}{\partial y} + \phi'(C) \frac{\partial^2 C}{\partial y \partial x} = 0. \tag{3-3}$$

De ahí que:

$$\phi'(C) \frac{\partial^2 C}{\partial x \partial y} = -\phi''(C) \frac{\partial C}{\partial x} \frac{\partial C}{\partial y}$$

$$\phi'(C) c(x, y) = -\phi''(C) \frac{\partial C}{\partial x} \frac{\partial C}{\partial y}$$

$$c(x, y) = \frac{-\phi''(C) \frac{\partial C}{\partial x} \frac{\partial C}{\partial y}}{\phi'(C)} = \frac{-\phi''(C) \frac{\phi'(x)}{\phi'(C)} \frac{\phi'(y)}{\phi'(C)}}{\phi'(C)}$$

$$c(x, y) = \frac{-\phi''(C) \phi'(x) \phi'(y)}{[\phi'(C)]^3}.$$

A partir de las propiedades dadas en (3-1), es claro que $c(x, y) > 0$ para todo (x, y) tal que $\phi(x) + \phi(y) < \phi(0)$. En general, las derivadas no existen en la frontera $\phi(x) + \phi(y) = \phi(0)$.

3.1.1 Propiedades de $C(x, y)$

- a. C es simétrica; esto es, $C(x, y) = C(y, x)$.

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

- b. C es asociativa; esto es, $C(C(x, y), z) = C(x, C(y, z))$.

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

- 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 tenemos que:

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

como $\phi(1) = 0$, entonces $\phi^{-1}[\phi(x)] = x$.

Análogamente,

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

De ahí que 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. Si $m > 0$ es cualquier constante, ϕ y $m\phi$ genera la misma cópula.

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

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

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

En general, la función de sobrevivencia conjunta está dada por

$$\bar{\mathbf{C}}(u_1, \dots, u_n, \dots, u_N) = \sum_{n=0}^N \left[(-1)^n \sum_{\mathbf{v}(u_1, \dots, u_n, \dots, u_N) \in Z(N-n, N, 1)} \mathbf{C}(v_1, \dots, v_n, \dots, v_N) \right],$$

donde $Z(M, N, \epsilon)$ denota el conjunto $\left\{ \mathbf{v} \in [0, 1]^N \mid v_n \in \{u_n, \epsilon\}, \sum_{n=1}^N \chi_\epsilon(v_n) = M \right\}$ (ver Georges, Lamy, Nicolas, Quibel, and Roncalli, 2001).

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

$$\overline{\mathbf{C}}(u_1, u_2, u_3) = 1 - u_1 - u_2 - u_3 + \mathbf{C}(u_1, u_2) + \mathbf{C}(u_1, u_3) + \mathbf{C}(u_2, u_3) - \mathbf{C}(u_1, u_2, u_3). \quad (3-4)$$

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

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

$$C(x, y) \leq \min(x, y).$$

Ahora, como la función de sobrevivencia es siempre positiva o cero y además por la propiedad (f), tenemos $C(x, y) - x - y + 1 \geq 0 \implies C(x, y) \geq x + y - 1$.

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

entonces,

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

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

Existen distintas clases de cópulas que describen diferentes estructuras de dependencia básicas de las variables. Nelsen (2006) ha suministrado una visión general sobre las diferentes familias. La tabla **3-1** lista algunas familias importantes de cópulas Arquimedianas con un solo parámetro, que se denotaran por C_θ .

Tabla 3-1: Familia de cópulas Arquimedianas

Familia	$C_\theta(x, y)$	$\phi^{-1}(t)$	Transformada de Laplace
Clayton	$\max(x^{-\theta} + y^{-\theta} - 1, 0)^{-1/\theta}$	$\left(\frac{1}{1 + \theta t}\right)^{1/\theta}$	Si
Ali-Mikhail-Haq	$\frac{xy}{1-\theta}(1-x)(1-y)$	$\left(\frac{1-\theta}{\exp(t)-\theta}\right)$	Si
Frank	$-\frac{1}{\theta} \ln \left[1 + \frac{\{\exp(-\theta x) - 1\}\{\exp(-\theta y) - 1\}}{\exp(-\theta) - 1} \right]$	$-\frac{1}{\theta} \ln[1 + \exp(-t)\exp(-\theta) - 1]$	Si

3.1.2 La cópula Frank

En este trabajo se hace énfasis en la cópula Frank debido a que ésta se utiliza para la implementación del algoritmo de Lo and Wilke (2010) para la extensión del estimador cópula gráfico.

La cópula Frank, cuyo generador es $\phi(t) = -\ln[(\exp(-\theta t) - 1)/(\exp(-\theta) - 1)]$, está dada por

$$C_{\theta}(x, y) = -\frac{1}{\theta} \ln \left[1 + \frac{\{\exp(-\theta x) - 1\}\{\exp(-\theta y) - 1\}}{\exp(-\theta) - 1} \right], \theta \neq 0. \quad (3-5)$$

El uso de la cópula Frank es atractivo ya que se puede capturar el rango completo de dependencia; esto es, al igual que la cópula Gaussiana, la cópula Frank 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 $\theta \rightarrow 0$. De hecho, cuando se trata de inferencia, algunos estadísticos prefieren usar la cópula Frank a la Gaussiana, ya que mientras ambas cópulas tienen propiedades similares, la cópula Frank proporciona cantidades cerradas y por lo tanto, más fáciles de programar (Escarela and Carriere, 2003).

Para evaluar el grado de asociación entre las marginales en el modelo generado por la cópula Frank, el τ de Kendall correspondiente está dado por

$$\tau_{\theta} = 1 - \frac{4}{\theta} [1 - D_1(\theta)], \quad (3-6)$$

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

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

(Genest, 1987). El τ de Kendall de la cópula Frank toma valores en el rango completo de concordancia. Observando los casos especiales de la cópula Frank, $\lim_{\theta \rightarrow -\infty} \tau_{\theta} = -1$, $\lim_{\theta \rightarrow \infty} \tau_{\theta} = 1$ y $\lim_{\theta \rightarrow 0} \tau_{\theta} = 0$.

La familia Frank se puede generalizar de 2-cópulas a la familia Frank de n -cópulas para $\theta > 0$ y cualquier $n \geq 2$

$$C_{\theta}^n(\mathbf{x}) = -\frac{1}{\theta} \ln \left[1 + \frac{\{\exp(-\theta x_1) - 1\}\{\exp(-\theta x_2) - 1\} \dots \{\exp(-\theta x_n) - 1\}}{\{\exp(-\theta) - 1\}^{n-1}} \right]. \quad (3-7)$$

3.2 Medidas de Concordancia

Cuando se considera una pareja de variables aleatorias, es útil saber qué tanto tienden a estar asociados valores grandes de una de las variables aleatorias con valores grandes de la otra, y que tanto están asociados valores pequeños de una con valores pequeños de la otra. Una formalización de la idea intuitiva de este grado de asociación fue propuesta por Yanagimoto and Okamoto (1969), quienes proponen el uso del orden de concordancia de distribuciones bivariadas con marginales univariadas dadas de acuerdo con la fuerza de su asociación positiva.

3.2.1 τ de Kendall

Suponga que (X, Y) y (X^*, Y^*) son dos observaciones independientes de una distribución conjunta. Entonces τ es la diferencia entre la probabilidad de concordancia y la probabilidad de discordancia de esas dos observaciones, esto es,

$$\tau = P\{(X^* - X)(Y^* - Y) \geq 0\} - P\{(X^* - X)(Y^* - Y) < 0\}. \quad (3-8)$$

$$\begin{aligned} \tau &= P\{(X^* - X)(Y^* - Y) \geq 0\} - P\{(X^* - X)(Y^* - Y) < 0\} \\ &= P\{(X^* - X)(Y^* - Y) \geq 0\} - \{1 - P\{(X^* - X)(Y^* - Y) \geq 0\}\} \\ &= 2P\{(X^* - X)(Y^* - Y) \geq 0\} - 1. \end{aligned} \quad (3-9)$$

Propiedades del τ de Kendall

Se asumen que las distribuciones marginales son continuas.

- i. $-1 \leq \tau \leq 1$.

Si todas las parejas son concordantes:

$$\begin{aligned} \tau &= 2P\{(X^* - X)(Y^* - Y) \geq 0\} - 1 \\ &= 2(1) - 1 = 1. \end{aligned}$$

Si todas las parejas son discordantes:

$$\begin{aligned} \tau &= 2P\{(X^* - X)(Y^* - Y) \geq 0\} - 1 \\ &= 2(0) - 1 = -1. \end{aligned}$$

Sabemos que,

$$\tau = 2P\{(X^* - X)(Y^* - Y) \geq 0\} - 1$$

Por definición de probabilidad

$$\begin{aligned} 0 &\leq P\{(X^* - X)(Y^* - Y) \geq 0\} \leq 1, \\ 0 &\leq 2P\{(X^* - X)(Y^* - Y) \geq 0\} \leq 2, \\ -1 &\leq 2P\{(X^* - X)(Y^* - Y) \geq 0\} - 1 \leq 1 \end{aligned}$$

Así, $-1 \leq \tau \leq 1$.

- ii. $\tau = 1 (= -1)$ si y sólo si $Y = f(X)$ para alguna función f monótona creciente (decreciente).

Restringiendo la atención a distribuciones sobre el cuadrado unitario con marginales uniformes, se mostrará que $Y = X$ con probabilidad 1 cuando $\tau = 1$. En este caso, se sabe que:

$$P\{(X^* - X)(Y^* - Y) \geq 0\} = 1.$$

Condicionando sobre un valor posible de (X, Y) , es decir, (x, y) , y usando la independencia de las dos observaciones (X, Y) y (X^*, Y^*) , se obtiene:

$$\begin{aligned} P\{(X^* - x)(Y^* - y) \geq 0\} &= 1, \\ P(X^* > x, Y^* > y) + P(X^* < x, Y^* < y) &= 1, \\ C(x, y) - x - y + 1 + C(x, y) &= 1. \end{aligned}$$

Por lo tanto:

$$C(x, y) = \frac{x + y}{2} \quad \forall \quad 0 < x, y < 1.$$

Por propiedades de las cópulas $C(x, y) \leq \min(x, y)$, entonces $(x + y)/2 \leq \min(x, y)$

- Si $\min(x, y) = x \longrightarrow \frac{x+y}{2} \leq x$. Por lo tanto $x \geq y$.
- Si $\min(x, y) = y \longrightarrow \frac{x+y}{2} \leq y$. Por lo tanto $x \leq y$.

De ahí que: $x = y$.

Ahora si $Y = F(X)$

Partiendo de la ecuación (3-9) y condicionando sobre un valor de (X, Y) , es decir (x, y)

$$\begin{aligned}\tau &= 2P\{(X^* - x)(Y^* - y) \geq 0\} - 1 \\ &= 2[P(X^* > x, Y^* > y) + P(X^* < x, Y^* < y)] - 1 \\ &= 2[1] - 1 = 1.\end{aligned}$$

iii. $\tau = 0$ si X e Y son independientes (pero no inversamente).

$$\begin{aligned}\tau &= 2P[(X^* - X)(Y^* - Y) \geq 0] - 1 \\ &= 2[P(X^* > X, Y^* > Y) + P(X^* < X, Y^* < Y)] - 1 \\ &= 2[P(X^* > X)P(Y^* > Y) + P(X^* < X)P(Y^* < Y)] - 1.\end{aligned}$$

Observación: Veamos

$$\begin{aligned}P(X^* > X) &= \int_{-\infty}^{\infty} \int_x^{\infty} f_X(x)f_{X^*}(x^*)dx^*dx \\ &= \int_{-\infty}^{\infty} f_X(x)[1 - F_X(x)]dx \\ &= - \int_1^0 udu = \left[\frac{-u^2}{2} \right]_1^0 \\ u &= [1 - F_X(x)] = \frac{1}{2}\end{aligned}$$

Por ser X y Y independientes

$$\tau = 2 \left[\left(\frac{1}{2} \right) \left(\frac{1}{2} \right) + \left(\frac{1}{2} \right) \left(\frac{1}{2} \right) \right] - 1 = 2 \left[\left(\frac{1}{4} \right) + \left(\frac{1}{4} \right) \right] - 1 = 2 \left[\frac{1}{2} \right] - 1 = 0.$$

iv. $\tau = 4E[C(X, Y)] - 1$

$$\begin{aligned}
\tau &= 2P\{(X^* - X)(Y^* - Y) \geq 0\} - 1 \\
&= 2E\{E[P\{(X^* - X)(Y^* - Y) \geq 0 | X = x, Y = y\}]\} - 1 \\
&= 2E\{P[(X^* - x)(Y^* - y) \geq 0]\} - 1 \\
&= 2E[P(X^* > x, Y^* > y) + P(X^* < x, Y^* < y)] - 1 \\
&= 2\{E[P(X^* > x, Y^* > y)] + E[P(X^* < x, Y^* < y)]\} - 1 \\
&= 2 \left[\iint_{\mathbb{R}^2} P(X^* > x, Y^* > y) dC(x, y) + \iint_{\mathbb{R}^2} P(X^* < x, Y^* < y) dC(x, y) \right] - 1.
\end{aligned}$$

Ahora,

$$\iint_{\mathbb{R}^2} P(X^* < x, Y^* < y) dC(x, y) = \iint_{I^2} C(x, y) dC(x, y).$$

$$\begin{aligned}
\iint_{\mathbb{R}^2} P(X^* > x, Y^* > y) dC(x, y) &= \iint_{I^2} [1 - x - y + C(x, y)] dC(x, y) \\
&= \iint_{I^2} 1 dC(x, y) - \iint_{I^2} x dC(x, y) - \iint_{I^2} y dC(x, y) \\
&= + \iint_{I^2} C(x, y) dC(x, y).
\end{aligned}$$

Debido a que $C(x, y)$ es la función de distribución conjunta de una pareja (X, Y) de variables aleatorias, con funciones de distribución marginales $F(X)$ y $F(Y)$ uniformes $[0, 1]$, $E(X) = 1/2$, $E(Y) = 1/2$. Entonces:

$$\begin{aligned}
P(X^* < X, Y^* < Y) &= 1 - \frac{1}{2} - \frac{1}{2} + \iint_{I^2} C(x, y) dC(x, y) \\
&= \iint_{I^2} C(x, y) dC(x, y).
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
\tau &= 2 \left[\iint_{R^2} P(X^* > x, Y^* > y) dC(x, y) + \iint_{R^2} P(X^* < x, Y^* < y) dC(x, y) \right] - 1 \\
&= 2 \left[\iint_{I^2} C(x, y) dC(x, y) + \iint_{I^2} C(x, y) dC(x, y) \right] - 1 \\
&= 2 \left[2 \iint_{I^2} C(x, y) dC(x, y) \right] - 1 \\
&= 4E[C(X, Y)] - 1.
\end{aligned}$$

Para mayores detalles ver Nelsen (2006).

Teorema 3.1 τ es invariante bajo transformaciones monótonas. Esto es, si f y g son funciones monótonas crecientes o decrecientes, entonces $\tau(f(X), g(Y)) = \tau(X, Y)$

Prueba 3.1 Por la monotonía tenemos:

$$\{(X_1 - X_2)(Y_1 - Y_2) > 0\} \iff \{(f(X_1) - f(X_2))(g(Y_1) - g(Y_2)) > 0\},$$

y

$$\{(X_1 - X_2)(Y_1 - Y_2) < 0\} \iff \{(f(X_1) - f(X_2))(g(Y_1) - g(Y_2)) < 0\}.$$

Ahora bien, usando (3-8) se tiene

$$\begin{aligned}
\tau &= P\{(f(X_1) - f(X_2))(g(Y_1) - g(Y_2)) \geq 0\} - P\{(f(X_1) - f(X_2))(g(Y_1) - g(Y_2)) < 0\} \\
&= P\{(X_1 - X_2)(Y_1 - Y_2) \geq 0\} - P\{(X_1 - X_2)(Y_1 - Y_2) < 0\}.
\end{aligned}$$

Debido a que son funciones monótonas crecientes. Si $X_1 \leq X_2$ entonces $f(X_1) \leq f(X_2)$ y $X_1 \geq X_2$ entonces $f(X_1) \geq f(X_2)$. ■

Cabe notar que si f y g son las funciones de distribución marginales de X y Y , respectivamente, entonces $f(X)$ y $g(Y)$ son uniformes.

Sea $Y = F(X)$

$$\begin{aligned}
F_Y(y) &= P(Y \leq y) = P(F(X) \leq y) = P(X \leq F^{-1}(y)) \\
&= F(F^{-1}(y)) = y.
\end{aligned}$$

Para calcular el τ de Kendall, es conveniente transformar la distribución conjunta de (X, Y) al cuadrado unitario.

Existe una fórmula simple para calcular el τ de Kendall cuando un cópula pertenece a la familia definida en (3-2).

Teorema 3.2 *Sea (X, Y) una pareja de variables aleatorias cuya distribución C es de la forma (3-2) para algún ϕ en Φ . Entonces,*

$$\tau(X, Y) = 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt + 1.$$

Prueba 3.2 *Primero note que $C(x, y) = 0$ para todo (x, y) tal que $\phi(x) + \phi(y) = 0$. Por lo tanto para calcular τ se integra C sobre la región en la cual hay una densidad. Esto es,*

$$\tau = \iint_{\phi(x)+\phi(y)<\phi(0)} C(x, y) \frac{-\phi''(C)\phi'(x)\phi'(y)}{[\phi'(C)]^3} dx dy.$$

Sea

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

$$u < x \quad \text{y} \quad \phi(0) = \infty \Rightarrow 0 < x < 1,$$

$$C(x, y) \leq \min(x, y) \leq x,$$

$$\text{Si } \phi(0) = \infty \Rightarrow 0 < x = v < 1,$$

y

$$J = - \left| \frac{\phi'(u)}{\phi'(y)} \right| = \frac{\phi'(u)}{\phi'(y)}.$$

Entonces:

$$\begin{aligned}
 \tau &= \int_0^1 \int_u^1 u \frac{-\phi''(u)\phi'(v)\phi'(y)\phi'(u)}{[\phi'(u)]^3} \frac{\phi'(u)}{\phi'(y)} dv du \\
 &= - \int_0^1 \int_u^1 u \frac{\phi''(u)\phi'(v)}{[\phi'(u)]^2} dv du \\
 &= - \int_0^1 u \frac{\phi''(u)}{[\phi'(u)]^2} \left[\int_u^1 \phi'(v) dv \right] du \\
 &= - \int_0^1 u \frac{\phi''(u)}{[\phi'(u)]^2} [-\phi(u)] du \\
 &= \int_0^1 u \frac{\phi''(u)}{[\phi'(u)]^2} \phi(u) du.
 \end{aligned}$$

Ahora

$$\begin{aligned}
 \frac{d}{du} \left[u \frac{\phi(u)}{\phi'(u)} \right] &= \frac{\phi(u)\phi'(u) + u[\phi'(u)]^2 - u\phi(u)\phi''(u)}{[\phi'(u)]^2} \\
 &= \frac{\phi(u)}{\phi'(u)} + u - \frac{u\phi(u)\phi''(u)}{[\phi'(u)]^2},
 \end{aligned}$$

$$\int_0^1 \frac{d}{du} \left[u \frac{\phi(u)}{\phi'(u)} \right] du = \int_0^1 \frac{\phi(u)}{\phi'(u)} du + \int_0^1 u du - \int_0^1 \frac{u\phi(u)\phi''(u)}{[\phi'(u)]^2} du$$

$$\left[u \frac{\phi(u)}{\phi'(u)} \right]_0^1 = \int_0^1 \frac{\phi(u)}{\phi'(u)} du + 1 - \int_0^1 \frac{u\phi(u)\phi''(u)}{[\phi'(u)]^2} du$$

$$\int_0^1 \frac{u\phi(u)\phi''(u)}{[\phi'(u)]^2} du = \int_0^1 \frac{\phi(u)}{\phi'(u)} du + 1 - \left[u \frac{\phi(u)}{\phi'(u)} \right]_0^1.$$

$$\int_0^1 \frac{u\phi(u)\phi''(u)}{[\phi'(u)]^2} du = \int_0^1 \frac{\phi(u)}{\phi'(u)} du + 1 - \left[\frac{\phi(1)}{\phi'(1)} \right].$$

Por lo tanto

$$\tau = \int_0^1 \frac{\phi(u)}{\phi'(u)} du + 1.$$

■

La relación entre el coeficiente de correlación ρ_N y el τ de Kendall para una distribución normal bivariada, está dada por

$$\tau = \frac{2}{\pi} \sin^{-1}(\rho_N). \quad (3-10)$$

4 Extensión del Estimador Cópula Gráfico: Método de Combinación de Riesgos

En este capítulo se presenta el método de combinación de riesgos propuesto por Lo and Wilke (2010) para la extensión del estimador cópula gráfico (Zheng and Klein, 1995). Además se presentan los detalles del algoritmo cópula gráfico, el cual es la base fundamental para la construcción del método de combinación de riesgos.

4.1 Estimador Cópula Gráfico

Zheng and Klein (1995) proponen un estimador no paramétrico para las distribuciones marginales, al cual denominan estimador cópula gráfico como una solución al problema de identificabilidad descrito en la sección 2.2. Para ello utilizan cópulas que capturan la dependencia entre dos variables aleatorias. El problema de identificabilidad es resuelto con los datos de riesgos competitivos y bajo el supuesto de que la cópula entre X y Y y su parámetro de asociación son conocidos.

En el marco de riesgos competitivos se observa $T_j = \min(X_j, Y_j)$ y $\delta_j = I_{\{X_j \leq Y_j\}}$ ($j = 1, \dots, n$). Asumiendo que $P(X_j = Y_j) = 0$, utilizando estos datos se pueden 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.

Se sabe bien que, bajo el supuesto de independencia de X y Y , la distribución marginal de ambas son determinadas de manera única por esas probabilidades. Zheng and Klein (1995) muestran en general, que si la cópula de (X, Y) es conocida, entonces las distribuciones marginales de X y Y son determinadas de manera única por los datos de riesgos competitivos. Dicho resultado es la base del estimador cópula gráfico.

4.1.1 Construcción del Estimador Cópula Gráfico

A continuación se presentan los detalles de la construcción del estimador cópula gráfico. Suponga que se tiene, en el contexto de riesgos competitivos, una muestra de tamaño n de tiempos de vida, donde $T_j = \min(X_j, Y_j)$. Si $F_1(t)$ y $F_2(t)$ son las distribuciones acumuladas marginales de X y Y respectivamente, entonces para cualquier cópula se tiene (ver, Figura 4-1)

$$\mu_c(A_t) = P(X > t, Y > t), \quad (4-1)$$

$$\mu_c(B_t) = P(X \leq t, X < Y), \quad (4-2)$$

donde

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

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

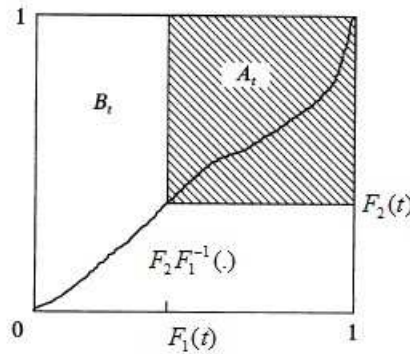


Figura 4-1: Relación entre $F_1(t)$ y $F_2(t)$ en el cuadrado unitario.

A continuación se explica la transformación realizada para que la relación entre $F_1(t)$ y $F_2(t)$ se encuentre en el cuadrado unitario. Observe que en la parte izquierda de la figura 4-2 se tienen las dos variables aleatorias X y Y originales y la línea recta corresponde a los valores donde $x = y$. En la parte derecha de la figura 4-2 se observa las nuevas variables U y V y la línea corresponde a la imagen de la diagonal en el cuadrado unitario bajo la transformación $(x, y) \rightarrow (u, v) = (F_1(X), F_2(Y))$.

Zheng and Klein (1995) prueban que estas dos relaciones (4-1) y (4-2) determinan de manera única a $F_1(t)$ y $F_2(t)$. Por lo tanto se encuentran estimadores $\hat{F}_1(t)$ y $\hat{F}_2(t)$ de $F_1(t)$ y $F_2(t)$,

los cuales conservan las propiedades sobre una rejilla seleccionada de m puntos distintos de falla o censura $0 < t_1 < t_2 < \dots < t_m < \max\{T_j, j = 1, \dots, n\}$.

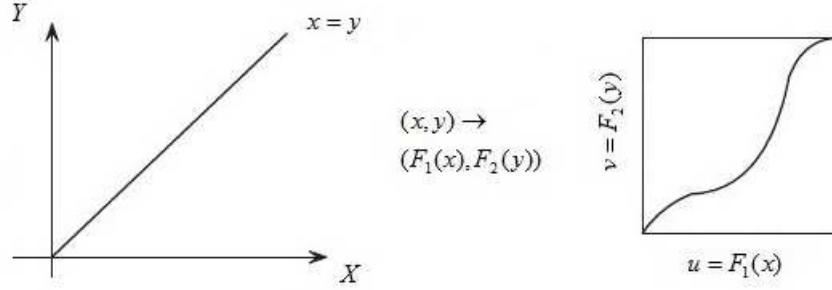


Figura 4-2: Transformación para la cópula.

Para la construcción del estimador cópula gráfico se define

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

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

Adicionalmente, sean

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

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

las estimaciones empíricas de $P(X > t, Y > t)$ y $P(X \leq t, X < Y)$ respectivamente.

Se encuentra a $\hat{F}_1(t_i)$ como la raíz de (4-4), sujeto a la definición de $\hat{F}_2(t_i)$ como una función de $\hat{F}_1(t_i)$ resolviendo la ecuación (4-3). Sean \hat{F}_1 y \hat{F}_2 funciones escalonadas en cada intervalo $(t_i, t_i + 1)$.

$$\mu_c(\hat{A}_t) - \hat{P}(X > t_i, Y > t_i) = 0, \quad (4-3)$$

$$\mu_c(\hat{B}_t) - \hat{P}(X \leq t_i, X < Y) = 0. \quad (4-4)$$

Una manera de escoger la rejilla en la cual el estimador se va a calcular, es tomar t_1, t_2, \dots, t_m tiempos distintos de falla o censura. Con esta rejilla y usando la función escalonada en cada intervalo t_{i-1}, t_i el estimador es más fácil de calcular. Para este estimador, 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 , mientras que si $\delta_i = 2$ entonces $\widehat{F}_1(t_i) = \widehat{F}_1(t_{i-1})$. Sea $t_0 = 0$ y $\widehat{F}_1(t_0) = \widehat{F}_2(t_0) = 0$.

Donde:

$$\delta_i = \begin{cases} 1, X < Y, \text{ es decir falla por la causa } X. \\ 2, Y < X, \text{ es decir falla por la causa } Y. \\ 0, \text{ Censura a derecha, esto es al terminar el estudio no falla por ninguna de las dos.} \end{cases}$$

Para $\delta_i = 1$

$$\mu_c(\widehat{A}_{t_i}) = 1 - \widehat{F}_1(t_i) - \widehat{F}_2(t_{i-1}) + C \left\{ \widehat{F}_1(t_i), \widehat{F}_2(t_{i-1}) \right\} = \widehat{P}(X > t_i, Y > t_i). \quad (4-5)$$

Para $\delta_i = 2$

$$\mu_c(\widehat{A}_{t_i}) = 1 - \widehat{F}_1(t_{i-1}) - \widehat{F}_2(t_i) + C \left\{ \widehat{F}_1(t_{i-1}), \widehat{F}_2(t_i) \right\} = \widehat{P}(X > t_i, Y > t_i), \quad (4-6)$$

donde $C \left\{ \widehat{F}_1, \widehat{F}_2 \right\}$ es la cópula que relaciona las marginales estimadas \widehat{F}_1 y \widehat{F}_2 . Las relaciones (4-5) y (4-6) definen la función de sobrevivencia conjunta bivariada presentada en Genest and Mackay (1986)

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

4.1.2 Algoritmo Cópula Gráfico

El algoritmo para construir \widehat{F}_1 y \widehat{F}_2 está basado en el algoritmo de bisección de búsqueda de raíces. El método de bisección es un método de búsqueda de raíces que repetidamente divide un intervalo y luego selecciona un subintervalo, en el que una raíz debe ser encontrada para su posterior procesamiento. Es un método muy simple y robusto.

El método se aplica cuando se desea resolver la ecuación $f(x) = 0$ para la variable real x , donde f es una función continua definida en el intervalo $[a, b]$ y $f(a)$ y $f(b)$ tienen signos

opuestos. Por el teorema del valor intermedio, f debe tener al menos una raíz en el intervalo (a, b) .

En cada paso el método divide el intervalo en dos calculando el punto medio $c = (a + b)/2$ del intervalo y el valor de la función $f(c)$ en ese punto. El método selecciona el subintervalo como un nuevo intervalo que se usará en el próximo paso. El proceso continua hasta que el intervalo sea suficientemente pequeño.

Explícitamente, si $f(a)$ y $f(c)$ tienen signos opuestos, entonces el método establece c como el nuevo valor para b , y si $f(b)$ y $f(c)$ tienen signos opuestos entonces el método establece c como el nuevo valor para a . Si $f(c) = 0$ entonces c puede ser tomada como la solución y el proceso para. En ambos casos, si los nuevos $f(a)$ y $f(b)$ tienen signos opuestos, el método es aplicable al intervalo más pequeño. El método garantiza que converge a una raíz de f , si f es una función continua en el intervalo $[a, b]$ y $f(a)$ y $f(b)$ tiene signos opuestos.

Pasos para la Construcción del Estimador Cópula Gráfico

1. Para $i = 1$, dado un punto inicial $\widehat{F}_1(t_1)$, es decir $\widehat{F}_1^{(1)}(t_1)$, encontrar $\widehat{F}_2^{(1)}(t_1)$ resolviendo (4-3).
2. Si el punto $(\widehat{F}_1^{(1)}(t_1), \widehat{F}_2^{(1)}(t_1))$ satisface (4-4), vaya al paso 4. En otro caso 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

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

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

Este proceso continúa hasta encontrar $\widehat{F}_1(t_1)$ y $\widehat{F}_2(t_1)$ que satisfacen (4-3) y (4-4). Es claro que 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 $\widehat{F}_2(t_m)$ como corresponde.

Observación: En el Anexo A se dan los detalles del funcionamiento del estimador cópula gráfico.

4.2 Extensión del Estimador Cópula Gráfico

Lo and Wilke (2010) realizan la extensión del estimador cópula gráfico propuesto por Zheng and Klein (1995) utilizando el método de combinación de riesgos para un modelo con más de dos riesgos competitivos cuando la cópula es Arquimediana. A continuación se dan los detalles de este método.

Sean $(T_1, \dots, T_J) \in \mathbf{R}_+^J$ tiempos latentes de falla $j = 1, \dots, J$ en un modelo de riesgos competitivos \mathbf{J} -dimensional. Se observa solamente T_j si $T_j < T_i$ para todo $i \neq j$. Todos los demás T_i no son observados, es decir, en el caso trivariado se tienen 3 tiempos de falla potenciales X , Y y Z , asociados a tres modos de falla, de manera que si ocurre el primero, los otros dos modos de falla no pueden ser observados. En este sentido, se dice que los tiempos son latentes. (T_1, \dots, T_J) pueden ser dependientes uno de otro.

Sea $S_j : \mathbf{R}_+ \rightarrow [0, 1]$ la función de sobrevivencia marginal desconocida, continua y estrictamente decreciente de T_j

$$S_j = P(T_j > t_j) = r_j, \quad (4-7)$$

donde r_j se considera 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 variables de tiempos de falla. La \mathbf{J} -cópula, $C^J : [0, 1]^J \rightarrow [0, 1]$, está definida por

$$C^J(r_1, \dots, r_J) = P(R_1 \leq r_1, \dots, R_J \leq r_J). \quad (4-8)$$

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

Si las funciones de supervivencia marginales están dadas por (4-7), la función de supervivencia conjunta $S(t_1, \dots, t_J) = P(T_1 > t_1, \dots, T_J > t_J)$ es determinada únicamente 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 supervivencia conjunta $S(t) = S(t, \dots, t)$ y la curva de incidencia acumulada de causa específica $Q_j(t_j)$, definida en la sección 2.1, están dadas por

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

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

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

Lo and Wilke (2010) pretenden determinar las funciones de supervivencia marginales $\{S_1(t_1), \dots, S_J(t_J)\}$ usando $\{S(t), Q_1(t_1), \dots, Q_J(t_J), C^J\}$. $S(t)$ y $Q_j(t_j)$ son estimadas y la función de la cópula verdadera C^J debe ser conocida o asumida.

Lo and Wilke (2010) sugieren la extensión de la implementación que fue propuesta por Zheng and Klein (1995). Para el caso $J > 2$, $S_j(t)$ puede ser calculada combinando todos los otros riesgos con $k \neq j$. El algoritmo numérico que fue propuesto por Zheng and Klein para un modelo de dos riesgos puede ser directamente aplicado para determinar $S_j(t)$. A continuación se ilustra el **método de combinación de riesgos** para el caso de $J = 3$. La idea es 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} es definida como $S_{23}(t) = P(T_{23} > t)$. Si existe una cópula de supervivencia entre las variables T_1 y T_{23} ,

$$C_1^2\{S_1(t_1), S_{23}(t_{23})\} = S(t), \quad (4-9)$$

la aproximación de Zheng and Klein puede aplicarse directamente para calcular $S_1(t_1)$ y $S_{23}(t_{23})$. Esto se hace resolviendo las dos ecuaciones

$$S(t) = \int_0^{r_{23}(t)} \int_0^{r_1(t)} dC_1^2(S_1, S_{23}). \quad (4-10)$$

Observación:

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

por la propiedad de las cópulas definida en el ítem (f) de la Sección 3.1.1

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

además, por definición (Nelsen, 2006, Página 32)

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

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 y C corresponde a la cópula de sobrevivencia.

$$Q_1(t) = \int_0^{\zeta_{23}(r_1)} \int_{r_1(t)}^1 dC_1^2(S_1, S_{23}), \quad (4-11)$$

donde $r_{23}(t) = S_{23}(t)$ y $\zeta_{23}(r_1) = S_{23}\{S_1^{-1}\}$.

Observación:

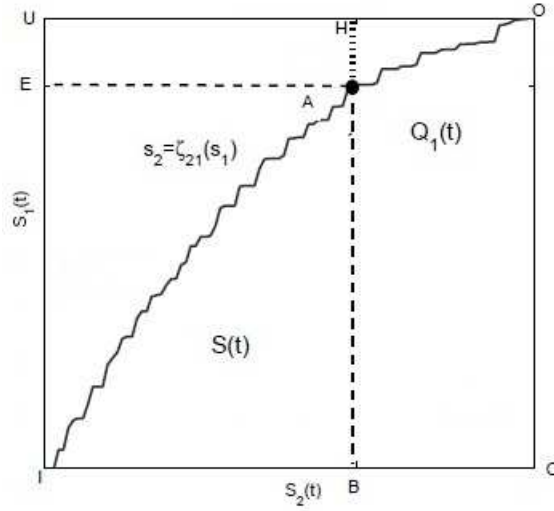


Figura 4-3: Representación gráfica del modelo de riesgos competitivos.

$Q_1(t)$ es la probabilidad asociada con el área OABC.

La ecuación (4-10) se cumple porque

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

note que $T_{23} = \min(T_2, T_3)$. Se tiene que $T_{23} > t$ sí y sólo sí $T_2 > t$ y $T_3 > t$. Similarmente, la ecuación (4-11) se cumple porque

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

No es necesario crear la variable combinada $T_{23} = \min(T_2, T_3)$ para la implementación del método de combinación de riesgo. La razón es que $S_1(t_1)$ y $S_{23}(t_{23})$ pueden ser identificadas a partir de $S(t)$ y $Q_1(t)$ junto con $C_1^2\{S_1(t_1), S_{23}(t_{23})\}$ sin la construcción de $Q_{23}(t)$. La variable combinada es sólo creada conceptualmente para probar el resultado de identificación. Para los otros casos, se procede análogamente utilizando la ecuación (4-9). $S_2(t)$ y

$S_3(t)$ pueden ser obtenidas a partir de $S(t)$, $Q_2(t)$ y $Q_3(t)$ junto con $C_2^2 = \{S_2(t_2), S_{13}(t_{13})\}$ y $C_3^2 = \{S_3(t_3), S_{12}(t_{12})\}$. Obsérvese, como se redujo el problema de tener inicialmente 3 variables a tener finalmente 2 variables y en consecuencia poder utilizar el estimador cópula gráfico.

Para que la metodología descrita anteriormente funcione, se debe garantizar la siguiente condición: para una 3-cópula conocida $C^3(r_1, r_2, r_3)$, existen todas las 2-cópulas $C_{ij}(r_i, r_j) = S_{ij}(t)$ tal que

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

En este caso C_{ij} es compatible con C^3 (Nelsen, 2006). Lo anterior debido a que desafortunadamente, una 2-cópula agrupada es generalmente inconsistente con la 3-cópula no agrupada de T_1, T_2 y T_3 . Esto significa que la cópula $C_1^2\{S_1(t_1), S_{23}(t_{23})\}$ en la ecuación (4-9) puede no existir (Genest, Ghousi, and Rivest, 1995). Por lo tanto, para la implementación de la metodología se utilizan cópulas Arquimedianas que satisfacen la propiedad de simetría y asociativa requeridas, como se describieron en la Sección 3.1.1. Como un resultado C^J puede ser construida paso a paso a partir de una 2-cópula

$$\begin{aligned} C^J(r_1, \dots, r_J) &= C^2\{C^{J-1}(r_1, \dots, r_{J-1}), r_J\} \\ &= C^2\{C^2[C^{J-2}(r_1, \dots, r_{J-2}), r_{J-1}], r_J\} \\ &= \dots \\ &= C^2\{C^2[\dots C^2(C^2(r_1, r_2), r_3), \dots], r_J\}. \end{aligned}$$

Schweizer and Sklar (1983) denotaron este procedimiento como una iteración serial de la 2-cópula Arquimediana la cual implica que la estructura de dependencia entre todos los riesgos r_j es la misma.

Teniendo en cuenta la restricción a esta clase de funciones cópula se deduce también que no importa el orden en que realicemos el agrupamiento de riesgo. Como un ejemplo, tome una 3-cópula de la clase Arquimediana como se indica en la ecuación (4-9). En este caso se tiene

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

De esta manera, a partir de $C_1^2\{S_1(t_1), S_{23}(t_{23})\}$ (o $C_1^2\{S_1(t), S_{32}(t_{32})\}$), $S(t)$ y $Q_1(t)$ se identifica $S_1(t)$, a partir de $C_2^2\{S_2(t_2), S_{13}(t_{13})\}$ (o $C_2^2\{S_2(t), S_{31}(t_{31})\}$), $S(t)$ y $Q_2(t)$ se identifica $S_2(t)$ y a partir de $C_3^2\{S_3(t_3), S_{12}(t_{12})\}$ (o $C_3^2\{S_3(t), S_{21}(t_{21})\}$) $S(t)$ y $Q_3(t)$ se identifica $S_3(t)$.

Los resultados son por lo tanto independientes del orden de las variables.

Se puede entonces utilizar la estrategia de Zheng and Klein (1995) a un modelo con más de dos riesgos: las distribuciones marginales S_j para $j = 1, \dots, J$ pueden ser identificadas por el método de combinación como se describió anteriormente, si la cópula de la variable combinada es compatible.

Nota: La propiedad de consistencia fuerte del estimador cópula gráfico se puede extender para el caso multivariado. Los detalles se pueden ver en Lo and Wilke (2010).

4.2.1 Chequeo Empírico del Método de Combinación de Riesgos

A continuación se realiza un estudio de simulación para la comparación de las funciones de distribución marginales estimadas por el método de combinación de riesgos, y las funciones de distribución marginales verdaderas, para tres riesgos competitivos dependientes, con el objetivo de chequear empíricamente el correcto funcionamiento del método de combinación de riesgos.

Esquema de Simulación

1. Se generan datos para tres riesgos competitivos dependientes a partir de la cópula Frank y el parámetro de la cópula es elegido tal que la correlación entre todos los riesgos sea la misma. Se consideran diferentes valores del nivel de dependencia, $\tau = 0.2, 0.5, 0.8$, donde τ es el Tau de Kendall.
2. Cada uno de los tres riesgos considerados en el estudio siguen una distribución Weibull como se presenta en la tabla **6-3**. Los parámetros de escala se fijan en $\eta_i = 1, i = 1, 2, 3$, ya que el parámetro de dependencia, no depende del parámetro de escala (Lu and Bhattacharyya, 1990). Este modelo Weibull se selecciona sólo con el objetivo de chequear empíricamente el método de combinación de riesgos. En el análisis de resultados se estudian otros modelos Weibull.

Se repiten las simulaciones para dos tamaños de muestra diferentes $n = 50$ y $n = 200$. Por lo tanto se consideran 6 escenarios posibles.

3. Se analizan gráficos de probabilidad Weibull para las funciones de distribución marginales con el fin de determinar la bondad de ajuste a la distribución Weibull, para los distintos tamaños de muestra.
4. Se calcula el error cuadrático medio (ECM) entre la función de distribución marginal verdadera en los percentiles t_p , con $p = 0.25, 0.50, 0.95$, $F_i(t_p), i = 1, 2, 3$, y la función de distribución marginal estimada por método de combinación de riesgos $\hat{F}_i(t_p), i = 1, 2, 3$.

$$ECM[\widehat{F}_i(t_p)] = \frac{1}{N_i} \sum_{j=1}^N \left(\widehat{F}_i(t_p) - F(t_p) \right)^2, i = 1, 2, 3,$$

donde N_i corresponde al número de veces que $\widehat{F}_i(t_p)$ es estimada de las 1000 simulaciones realizadas.

Tabla 4-1: Diseño de simulación.

Riesgo	Distribución	Parámetros
1	Weibull	$(\beta, \eta) = (2.4, 1)$
2	Weibull	$(\beta, \eta) = (2.6, 1)$
3	Weibull	$(\beta, \eta) = (2.8, 1)$

Las Figuras 4-4 - 4-6 comparan las funciones de distribución marginales de los riesgos 1, 2 y 3 estimadas por el método de combinación de riesgos y las funciones de distribución marginales verdaderas para los dos tamaños de muestra, $n = 50$ y $n = 200$ y el nivel de dependencia $\tau = 0.2$ correspondiente a $\theta = 2$, mediante gráficos de probabilidad Weibull. La línea recta corresponde a la distribución marginal verdadera, denotada en el gráfico por F-Ver, y los puntos corresponden a los valores de la función de distribución marginal estimada. En las Figuras se observa el buen ajuste que tienen las funciones de distribución marginales estimadas por el método de combinación de riesgos a las distribuciones marginal reales, mejorando el ajuste cuando el tamaño de muestra es mayor, indicando que este método estima correctamente las marginales de cada uno de los riesgos para este caso de estudio, en el cual las distribuciones son muy similares y bajo el supuesto de que la estructura de dependencia entre los riesgos es la misma.

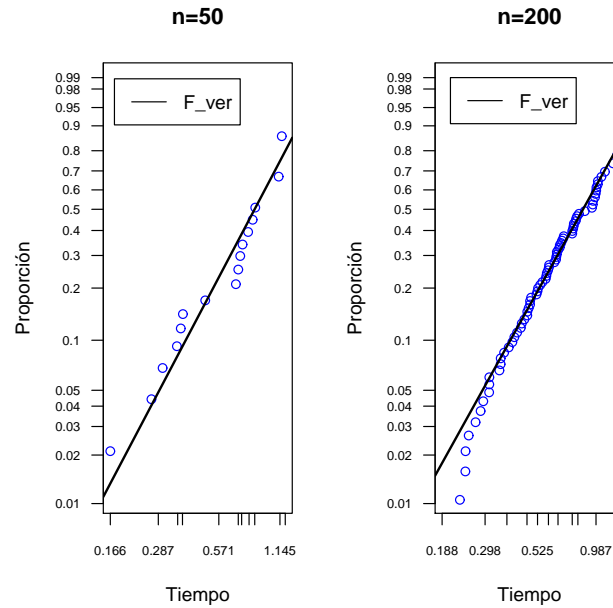


Figura 4-4: Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 1, $\beta_1 = 2.4, \eta_1 = 1, \tau = 0.2$, para $n = 50$ y $n = 200$.

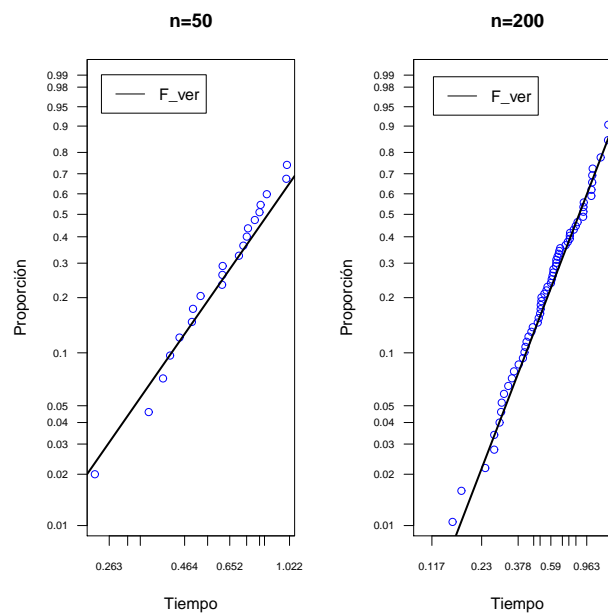


Figura 4-5: Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 2, $\beta_2 = 2.6, \eta_2 = 1, \tau = 0.2$, para $n = 50$ y $n = 200$.

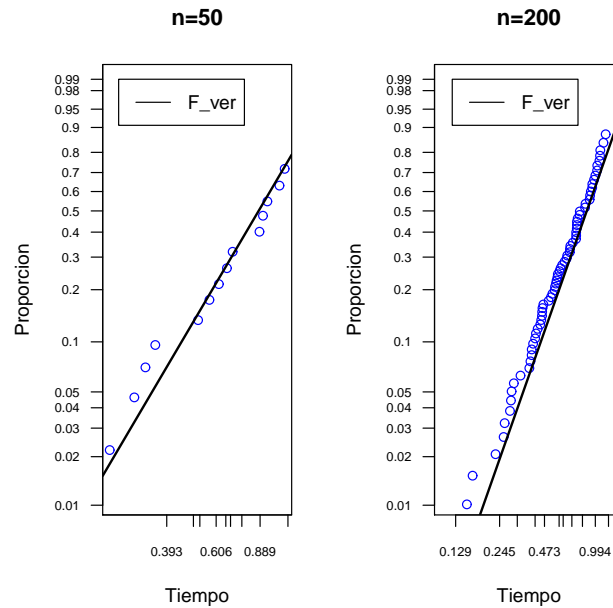


Figura 4-6: Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 3, $\beta_3 = 2.8$, $\eta_3 = 1$, $\tau = 0.2$, para $n = 50$ y $n = 200$.

La Tabla 4-2 muestra el error cuadrático medio para las funciones de distribución marginales estimadas en los percentiles t_p con $p = 0.25, 0.50, 0.95$, para un nivel de dependencia $\tau = 0.2$. Se observa que los valores del ECM son pequeños indicando que la funciones de distribución marginales se aproximan muy bien a la función de distribución marginal verdadera para los dos tamaños de muestra considerados $n = 50$ y $n = 200$. Por ejemplo, para $p = 0.50$ con $n = 50$ para el riesgo 1 el $ECM = 0.00440$ y con $n = 200$ un $ECM = 0.00089$. Es decir las estimaciones son buenas.

Tabla 4-2: Error cuadrático medio para las funciones de distribución marginales estimadas en los percentiles t_p , con $p = 0.25, 0.50, 0.95$ para $\tau = 0.2$.

p	$n = 50$			$n = 200$		
	Riesgo 1	Riesgo 2	Riesgo 3	Riesgo 1	Riesgo 2	Riesgo 3
0.25	0.00205	0.00209	0.00182	0.00059	0.00083	0.00030
0.50	0.00440	0.00412	0.00545	0.00089	0.00139	0.00170
0.95	0.01042	0.01013	0.00880	0.00441	0.00363	0.00191

Las Figuras 4-7 - 4-9 comparan las funciones de distribución marginales de los riesgos 1, 2 y 3 estimadas por el método de combinación de riesgos y las funciones de distribución marginales verdaderas para los dos tamaños de muestra, $n = 50$ y $n = 200$ y el nivel de dependencia $\tau = 0.5$ correspondiente a $\theta = 5.8$, mediante gráficos de probabilidad Weibull. En las Figuras se observa que las distribuciones marginales estimadas por el método de combinación de riesgos se encuentran muy cercanas a las distribuciones marginales reales indicando que las estimaciones son buenas. Cabe resaltar que en la Figuras 4-7 y 4-8 para $n = 200$ las funciones de distribución marginales estimadas en la cola inferior no presentan un buen ajuste. Sin embargo la existencia de mayor variabilidad en las colas es un comportamiento de esperar en estos gráficos de probabilidad.

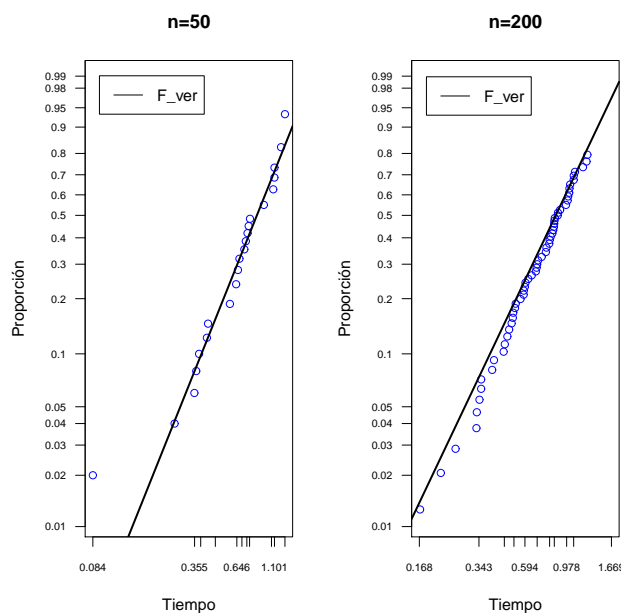


Figura 4-7: Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 1, $\beta_1 = 2.4, \eta_1 = 1, \tau = 0.5$, para $n = 50$ y $n = 200$.

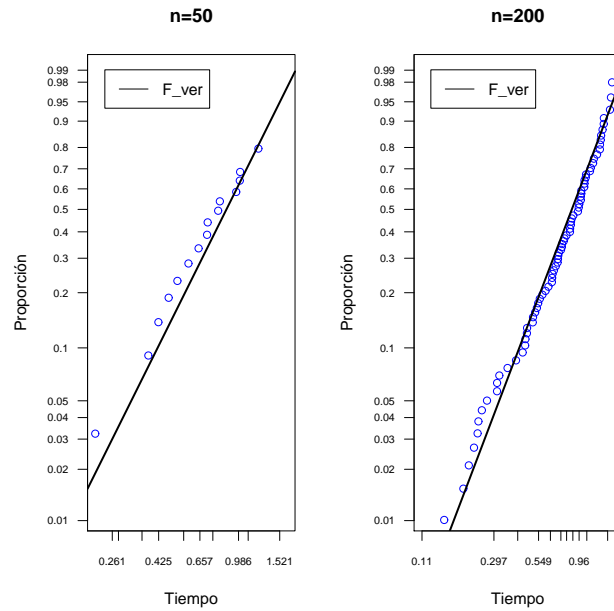


Figura 4-8: Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 2, $\beta_2 = 2.6$, $\eta_2 = 1$, $\tau = 0.5$, para $n = 50$ y $n = 200$.

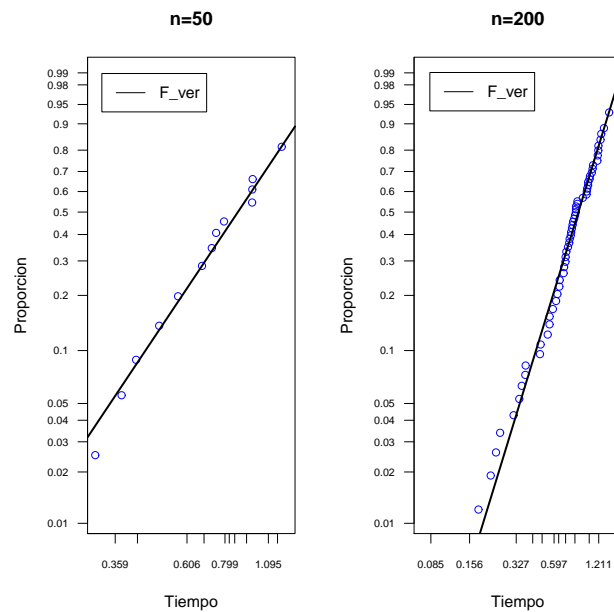


Figura 4-9: Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 3, $\beta_3 = 2.8$, $\eta_3 = 1$, $\tau = 0.5$, para $n = 50$ y $n = 200$.

El error cuadrático medio para las funciones de distribución marginales estimadas en los percentiles t_p , con $p = 0.25, 0.5, 0.95$ se muestra en la Tabla **4-3** para un nivel de dependencia $\tau = 0.5$. En este caso, se observa que los valores del ECM son pequeños indicando que las estimaciones son buenas, la funciones de distribución marginales estimadas están muy cercanas a la función de distribución marginal real, para los dos tamaños de muestra considerados $n = 50$ y $n = 200$.

Tabla 4-3: Error cuadrático medio para las funciones de distribución marginales estimadas en los percentiles t_p , con $p = 0.25, 0.50, 0.95$ para $\tau = 0.5$.

p	$n = 50$			$n = 200$		
	Riesgo 1	Riesgo 2	Riesgo 3	Riesgo 1	Riesgo 2	Riesgo 3
0.25	0.00244	0.00360	0.00264	0.00128	0.00107	0.00057
0.50	0.00555	0.00591	0.00765	0.00187	0.00153	0.00189
0.95	0.00628	0.00619	0.00420	0.00087	0.00094	0.00101

Las Figuras **4-10** - **4-12** comparan las funciones de distribución marginales de los riesgos 1, 2 y 3 estimadas por el método de combinación de riesgos y las funciones de distribución marginales verdaderas para los dos tamaños de muestra, $n = 50$ y $n = 200$ y el nivel de dependencia $\tau = 0.8$ correspondiente a $\theta = 20$, mediante gráficos de probabilidad Weibull. En las Figuras se observa que las distribuciones marginales estimadas por el método de combinación de riesgos se encuentran muy cercanas a las distribuciones marginales reales indicando que las estimaciones son buenas. En la Figura **4-12** para $n = 50$ no se observa un buen ajuste de la función de distribución marginal estimada a la función de distribución marginal verdadera.

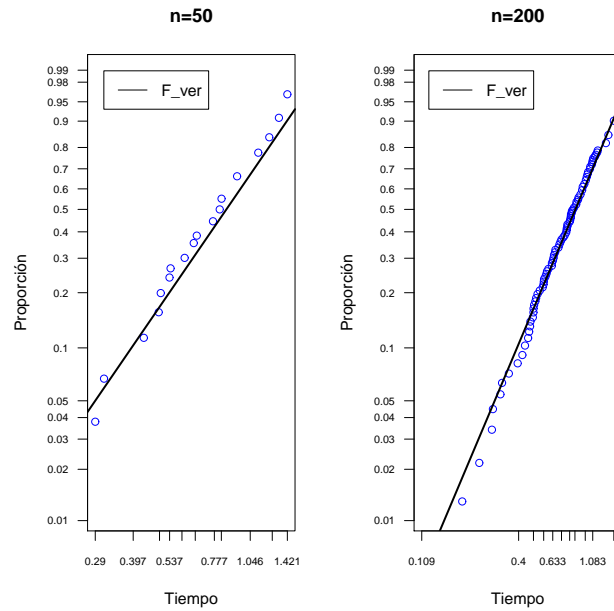


Figura 4-10: Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 1, $\beta_1 = 2.4$, $\eta_1 = 1$, $\tau = 0.8$, para $n = 50$ y $n = 200$.

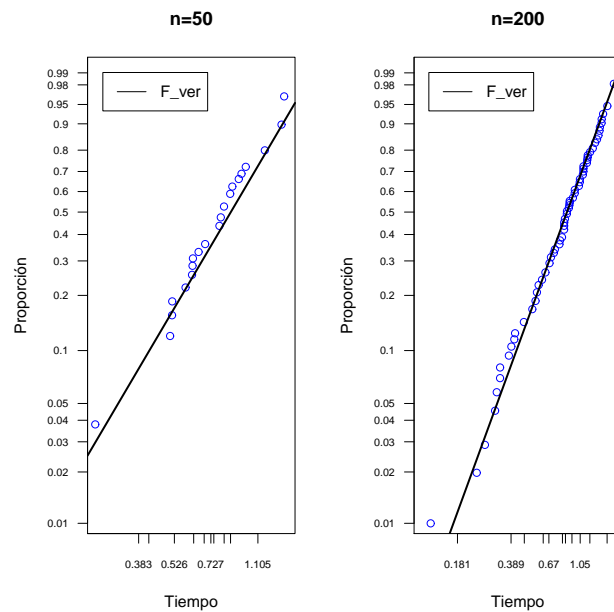


Figura 4-11: Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 2, $\beta_2 = 2.6$, $\eta_2 = 1$, $\tau = 0.8$, para $n = 50$ y $n = 200$.

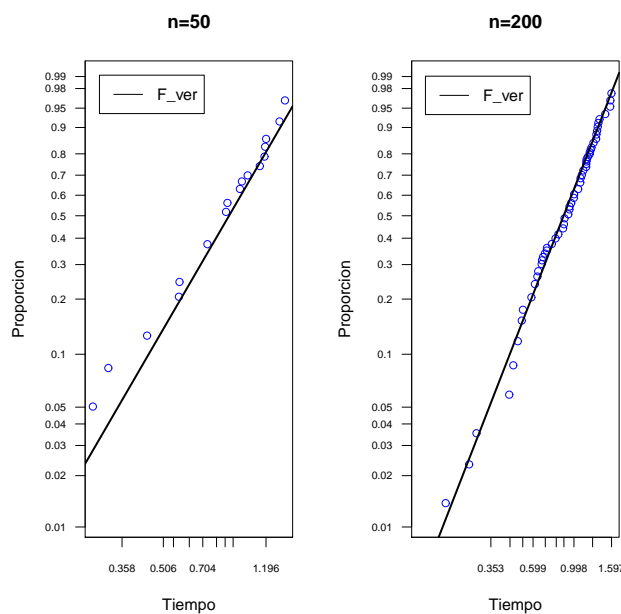


Figura 4-12: Gráfico de probabilidad Weibull para la función de distribución marginal del riesgo 3, $\beta_3 = 2.8$, $\eta_3 = 1$, $\tau = 0.8$, para $n = 50$ y $n = 200$.

La Tabla 4-4 muestra el error cuadrático medio para las funciones de distribución marginales estimadas en los percentiles t_p , con $p = 0.25, 0.5, 0.95$. Los resultados obtenidos del ECM indican que las funciones de distribución marginales se aproximan muy bien a la función de distribución marginal verdadera para los dos tamaños de muestra considerados $n = 50$ y $n = 200$ debido a que estos valores son pequeños. Lo anterior permite afirmar que las estimaciones son buenas.

Tabla 4-4: Error cuadrático medio para las funciones de distribución marginales estimadas en los percentiles t_p , con $p = 0.25, 0.50, 0.95$ para $\tau = 0.8$.

p	$n = 50$			$n = 200$		
	Riesgo 1	Riesgo 2	Riesgo 3	Riesgo 1	Riesgo 2	Riesgo 3
0.25	0.00379	0.00377	0.00491	0.00084	0.00126	0.00177
0.50	0.00516	0.00509	0.00767	0.00124	0.00075	0.00170
0.95	0.00448	0.00302	0.00218	0.00066	0.00048	0.00060

5 Generación de Tiempos de Falla Weibull y Lognormales

En este capítulo se presentan los algoritmos para la generación de tiempos de falla Weibull y lognormales trivariados dependientes. Además de ello, se proponen dos algoritmos para la generación de tiempos de falla lognormales bivariados utilizando como medida de dependencia el τ de Kendall.

5.1 Obtención de Tiempos de Falla Lognormales

Para obtener tiempos de falla lognormales trivariados, se define la distribución lognormal multivariada. Sea $\mathbf{X} = [X_1, X_2, \dots, X_p]$ un vector aleatorio de p componentes que tiene una distribución normal multivariada con vector de medias $\boldsymbol{\xi}$ y matriz de varianzas y covarianzas \mathbf{V} . Ahora use la transformación $Y_i = \exp(X_i)$ y defina $\mathbf{Y} = [Y_1, Y_2, \dots, Y_p]$. Entonces \mathbf{Y} tiene una distribución lognormal multivariada (Kotz, Balakrishnan, and Johnson, 2000).

Si $Y = [\ln(y_1), \dots, \ln(y_p)]'$

$$g(y) = \frac{1}{\prod_{i=1}^p y_i (2\pi)^{(p/2)} |\Sigma|^{1/2}} \exp \left[-\frac{1}{2} (y - \mu)' \Sigma^{-1} (y - \mu) \right], \quad y \in R^{P+}$$

En el caso bivariado, se exponencian los tiempos de falla generados a partir de variables aleatorias normales bivariadas (X, Y) con coeficiente de correlación ρ , es decir, $(U, V) = [\exp(X), \exp(Y)]$, para generar tiempos de falla lognormales bivariados (U, V) .

Se sabe que el coeficiente de correlación ρ_L de la lognormal bivariada se expresa así en terminos de ρ

$$\rho_L = \frac{\exp(\rho_N \sigma_1 \sigma_2) - 1}{\{[\exp(\sigma_1^2) - 1][\exp(\sigma_2^2) - 1]\}^{1/2}}$$

Como el interés es la medida de asociación entre las variables aleatorias lognormales, ρ_L no resulta ser una buena medida de dependencia, puesto que el coeficiente de correlación ρ_N no es invariante bajo transformaciones monótonas. Para ilustrar lo anterior, en la figura **5-1**,

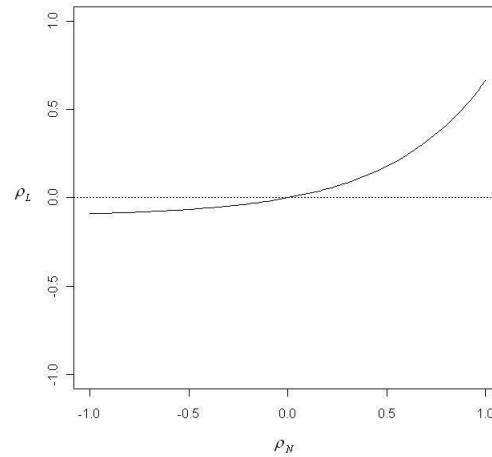


Figura 5-1: Relación entre el Coeficiente de Correlación Lineal ρ y el ρ_L , $\sigma_1 = 2$, $\sigma_2 = 1$.

en el eje x está el coeficiente de correlación de las variables aleatorias normales ρ_N y en el eje y el coeficiente de correlación de las variables aleatorias lognormales ρ_L . Se observa que por ejemplo para un valor de $\rho_N = 0.8$ el valor de $\rho_L = 0.5$. Además es claro que los valores de ρ_L no se encuentran entre -1 y 1.

En la figura **5-2** se evidencia claramente la sensibilidad que tiene el coeficiente de correlación cuando se modifican los parámetros de la distribución lognormal bivariada. Obsérvese que cuando $\rho = 0$, el resultado esperado se presenta en la gráfica de la derecha para parámetros $\sigma_1 = 0.05$, $\sigma_2 = 0.05$, donde se muestra que no hay dependencia entre las variables. Sin embargo para el mismo valor de $\rho = 0$, en la gráfica de la izquierda cuando se modifican los parámetros a $\sigma_1 = 0.5$, $\sigma_2 = 0.1$, se evidencia que existe cierta dependencia entre las variables. De ahí la dificultad que se presenta de interpretar a ρ como medida de dependencia en el caso lognormal. El coeficiente de correlación de Pearson, es idealmente adecuado para la distribución normal bivariada.

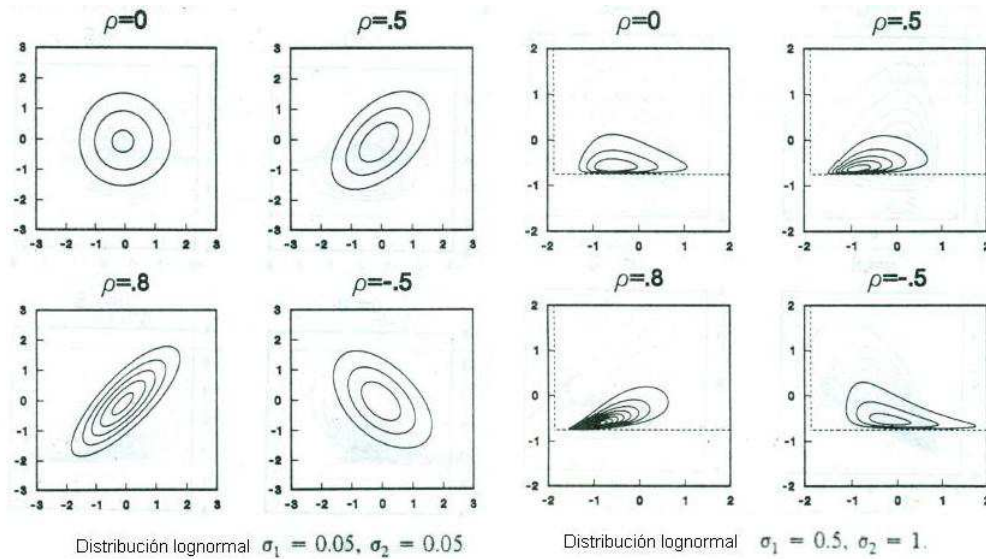


Figura 5-2: Comportamiento del coeficiente de correlación para diferentes valores de los parámetros de la distribución Lognormal.

Teniendo en cuenta todo lo anterior, se propone dos algoritmos para la generación de tiempos de falla lognormales dependientes.

5.1.1 Generación de Tiempos de Falla Lognormales Dependientes

Algoritmo 1

A continuación se presenta el algoritmo para la generación de tiempos de falla lognormales bivariados dependientes utilizando como medida de asociación el τ de Kendall que permite controlar la dependencia entre las variables aleatorias lognormales.

1. Escoger un valor del τ de Kendall.
2. Calcular el coeficiente de correlación ρ_N a partir de la ecuación (3-10), es decir $\rho_N = \sin\left(\frac{\tau * \pi}{2}\right)$
3. Generar tiempos de falla normales bivariados (X, Y) con coeficiente de correlación ρ_N obtenido en el paso 2.
4. Generar tiempos de falla lognormales bivariados (U, V) exponenciando los tiempos de falla generados a partir (X, Y) , es decir, $(U, V) = [\exp(X), \exp(Y)]$.
5. Así se obtiene lognormales bivariadas dependientes con parámetros $\mu_1, \mu_2, \sigma_1, \sigma_2$ y τ .

Una metodología más general, es la generación de distribuciones multivariadas dependientes con marginales lognormales utilizando cópulas. En el siguiente algoritmo se utiliza cópulas que pertenecen a la familia de las cópulas arquimedianas.

Algoritmo 2

Generar tiempos de falla con marginales lognormales usando cópulas, a partir del paquete cópula implementado por Yan (2007), en el software estadístico R. A continuación se presenta el algoritmo utilizando la cópula Frank y la cópula Gumbel-Hougaard, sin embargo para el desarrollo de este trabajo se utilizó únicamente la cópula Frank.

1. Para el caso trivariado:

- La cópula Frank

$$C_{\theta}^3(\mathbf{x}) = -\frac{1}{\theta} \ln \left[1 + \frac{\{\exp(-\theta x_1) - 1\}\{\exp(-\theta x_2) - 1\}\{\exp(-\theta x_3) - 1\}}{\{\exp(-\theta) - 1\}^2} \right], \theta > 0.$$

- La cópula Gumbel-Hougaard

$$C_{\theta}^3(\mathbf{x}) = \exp \left\{ -\left[(-\ln x_1)^{\theta} + (-\ln x_2)^{\theta} + (-\ln x_3)^{\theta} \right]^{1/\theta} \right\}, \theta \geq 1,$$

donde θ es el parámetro de dependencia.

2. Para el caso bivariado:

- La cópula Frank definida en

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

Se calcula el τ de Kendall a partir del parámetro de dependencia θ , como se definió en (3-6)

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

- La cópula Gumbel-Hougaard está definida por

$$C_\theta(x, y) = \exp \left\{ -[(-\ln x)^\theta + (-\ln y)^\theta]^{1/\theta} \right\}, \theta \geq 1.$$

Se calcula el τ de Kendall a partir del parámetro de dependencia θ .

$$\tau = 1 - \frac{1}{\theta}$$

Liu (2010) muestra que la distribución lognormal bivariada también juega un papel muy importante en el área de ingeniería electrónica en el procesamiento de señales y comunicaciones.

5.2 Obtención de Tiempos de Falla Weibull

Para simular tiempos de falla Weibull, se presenta una clase de distribución multivariada Weibull: Considere variables aleatorias X_1, \dots, X_n que tienen una distribución conjunta, cada $X_i, i = 1, \dots, n$ tiene una distribución Weibull de la forma $\bar{F}_i(t) = \exp(-\lambda_i t^{\alpha_i})$, con $\lambda_i > 0, i = 1, \dots, n$ (Lee, 1979).

Para el caso bivariado, sean T_1 y T_2 tiempos de falla Weibull, una función de confiabilidad conjunta para la Weibull bivariada (Lu and Bhattacharyya, 1990), es:

$$\bar{F}(t_1, t_2) = \exp \left\{ - \left[\left(\frac{t_1}{\eta_1} \right)^{\beta_1/(1-\tau)} + \left(\frac{t_2}{\eta_2} \right)^{\beta_2/(1-\tau)} \right]^{(1-\tau)} \right\}, 0 \leq \tau < 1, t_i \geq 0, \beta_i \geq 0, \eta_i \geq 0,$$

donde $\beta_1, \eta_1, \beta_2, \eta_2$, son los parámetros de forma y escala asociados a T_1 y T_2 respectivamente, $0 \leq \tau < 1$, es el parámetro de dependencia entre T_1 y T_2 , cuyas funciones de confiabilidad marginales son:

$$\bar{F}_k(t) = \exp \left\{ - \left(\frac{t}{\theta_k} \right)^{\beta_k} \right\}, k = 1, 2.$$

5.2.1 Generación de Tiempos de Falla Weibull Dependientes

Algoritmo

Generar tiempos de falla con marginales Weibull usando cópulas, a partir del paquete cópula implementado por Yan (2007), en el software estadístico R. A continuación se presenta el algoritmo utilizando la cópula Frank.

1. Para el caso trivariado:

$$C_{\theta}^3(\mathbf{x}) = -\frac{1}{\theta} \ln \left[1 + \frac{\{\exp(-\theta x_1) - 1\}\{\exp(-\theta x_2) - 1\}\{\exp(-\theta x_3) - 1\}}{\{\exp(-\theta) - 1\}^2} \right], \theta > 0.$$

2. Para el caso bivariado:

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

donde θ es un parámetro de dependencia.

Para el caso bivariado, se calcula el τ de Kendall a partir del parámetro de dependencia θ , como se definió en (3-6)

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

Cabe resaltar que para este estudio se trabaja bajo el supuesto que todos los riesgos están igualmente correlacionados, es decir, en el caso de tiempos de falla lognormal se asume que para el caso trivariado $cov(t_1, t_2) = cov(t_1, t_3) = cov(t_2, t_3)$. La aplicación de lo anterior se encuentra en áreas como la biología (ver, Johnson and Wichern, 2007, Página 469) y la economía. Lo and Wilke (2010) también utilizan esta clase de matrices en un estudio del desempleo en Alemania, considerando tres riesgos competitivos. Sin embargo, se puede contemplar el caso en el cual los riesgos presenten diferentes correlaciones y se trabaje con cópulas de la familia Arquimediana de dos parámetros (ver, Nelsen, 2006).

5.3 Esquema de Simulación

A continuación se presenta el esquema general para realizar la comparación entre la función de sobrevivencia verdadera del tiempo mínimo de falla $S(t)$ con la función de sobrevivencia

estimada asumiendo independencia ($\widehat{S}^*(t)$) y la función de sobrevivencia del tiempo mínimo de falla utilizando el método de combinación de riesgos, teniendo en cuenta la dependencia entre los modos de falla ($\widehat{S}(t)_{MCR}$).

1. Las funciones de sobrevivencia verdaderas del tiempo mínimo de falla de un sistema corresponden a diferentes escenarios para tiempos de falla con marginales lognormales y tiempos de falla con marginales Weibull.

a. Casos Weibull

Los escenarios Weibull a estudiar consideran 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 de falla es creciente y cuando $\beta = 1$, la tasa de falla es constante. Los parámetros de escala de la distribución Weibull se fijan en $\eta_1 = \eta_2 = \eta_3 = 1$, ya que el parámetro de dependencia, no depende del parámetro de escala (Lu and Bhattacharyya, 1990). Además, se escogen los modelos Weibull de tal manera que se garantice que existan fallas para cada modo, $i = 1, 2, 3$. Para ello, se establece que como mínimo haya un 15% de fallas para cada uno de los modos, de tal forma que no se presente un alto porcentaje de censura que dificulta la estimación de la función del tiempo mínimo asumiendo independencia ($\widehat{S}^*(t)$) y dependencia ($\widehat{S}(t)_{MCR}$). Por ejemplo, para un tamaño de muestra de $n = 50$, si $\min\{T_1, T_2, T_3\} = T_1$ un 15% equivale a tener aproximadamente 7 fallas y 43 censuras, que corresponden a 43 observaciones pérdidas (missing) lo que hace muy inestable la estimación.

A continuación se describen los modelos Weibull a estudiar y se presentan los gráficos de probabilidad Weibull correspondientes, con el fin de ilustrar gráficamente la escogencia de cada uno de ellos.

$$- \beta_1 = 2, \beta_2 = 2, \beta_3 = 2 \text{ y } \eta_1 = \eta_2 = \eta_3 = 1$$

En la Figura 5-3 se observa que para este modelo Weibull el porcentaje de falla es similar para cada uno de los modos, puesto que las tres funciones de distribución marginales son iguales.

Tabla 5-1: Porcentaje de falla para el modelo Weibull $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$.

Modo	Porcentaje
1	33.31%
2	33.20%
3	33.48%

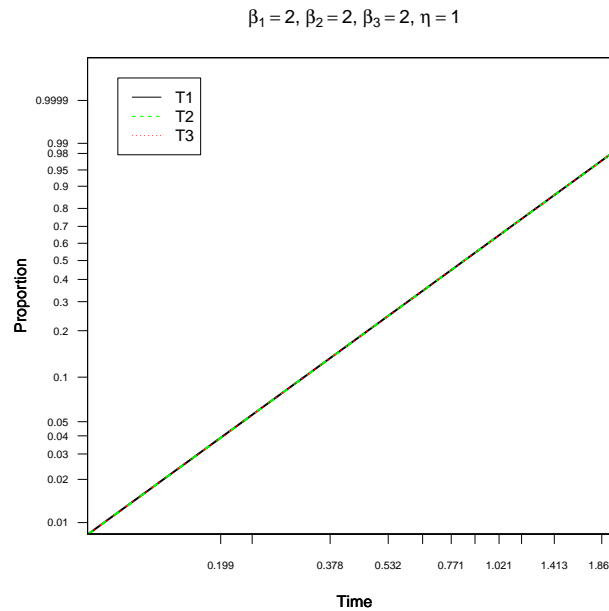


Figura 5-3: Gráfico de probabilidad Weibull para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$.

– $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5$ y $\eta_1 = \eta_2 = \eta_3 = 1$

En la Figura 5-4 se observa que para este modelo Weibull las funciones de distribución son diferentes. El modo de falla 3 es el que tiene mayor ocurrencia que los otros dos modos de falla, sin embargo se garantiza que haya como mínimo un 15% de fallas para cada uno de los modos.

Es importante anotar que en la Weibull con $\eta = 1$, las tres rectas se cortan en el cuantil 0.632 (En general η es el cuantil 0.632 para una distribución Weibull).

Tabla 5-2: Porcentaje de falla para el modelo Weibull $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$.

Modo	Porcentaje
1	26.01%
2	28.98%
3	45.00%

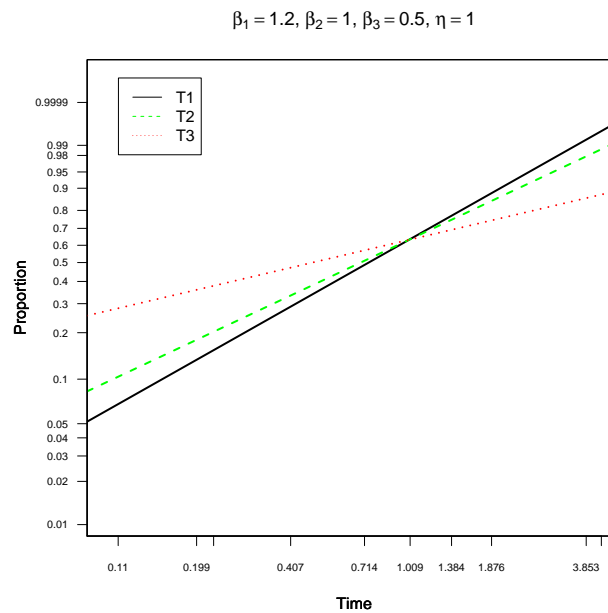


Figura 5-4: Gráfico de probabilidad Weibull para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$.

– $\beta_1 = 1.5, \beta_2 = 1.0, \beta_3 = 1.0$ y $\eta_1 = \eta_2 = \eta_3 = 1$

En este modelo Weibull se observa como se presenta en la Figura 5-5 que los modos de falla 2 y 3 tienen mayor ocurrencia que el modo de falla 1. De igual manera que en el caso anterior se garantiza que haya como mínimo un 15% de fallas para cada uno de los modos.

Tabla 5-3: Porcentaje de falla para el modelo Weibull $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$.

Modo	Porcentaje
1	27.43%
2	36.49%
3	36.08%

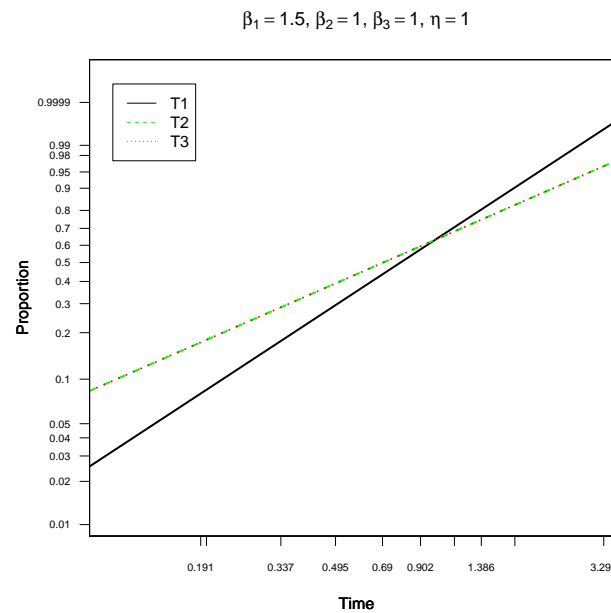


Figura 5-5: Gráfico de probabilidad Weibull para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$.

– $\beta_1 = 0.5, \beta_2 = 1.0, \beta_3 = 1.0$ y $\eta_1 = \eta_2 = \eta_3 = 1$

En la Figura 5-6 se observa que el modo de falla 1 tiene mayor ocurrencia que los otros dos modos de falla. Se garantiza que haya como mínimo un 15% de fallas para cada uno de los modos.

Tabla 5-4: Porcentaje de falla para el modelo Weibull $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$.

Modo	Porcentaje
1	43.70%
2	28.19%
3	28.11%

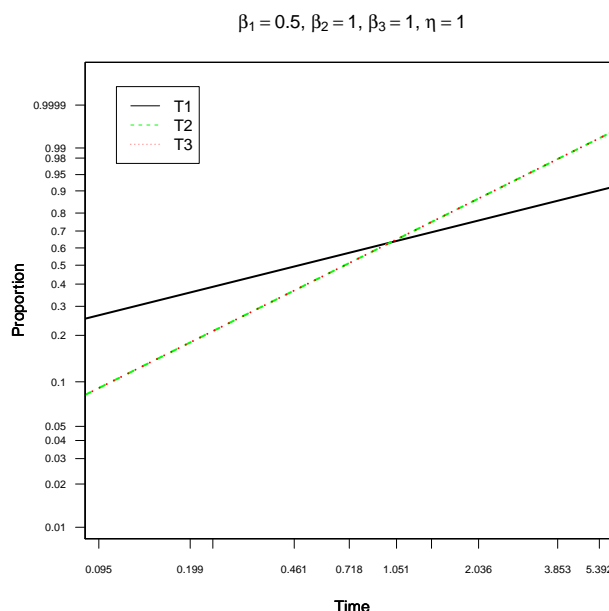


Figura 5-6: Gráfico de probabilidad Weibull para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$.

b. Casos Lognormal

Los escenarios de los casos lognormales se escogen de tal manera que se garantice que ocurran los tres modos de falla. Cabe resaltar que en la distribución lognormal μ corresponde a un parámetro de escala y σ a un parámetro de forma. Además, la media y la varianza de una distribución lognormal son, $E(T) = \exp(\mu + 0.5\sigma^2)$ y $Var(T) = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1]$, respectivamente.

En los modelos seleccionados se fija el parámetro de escala $\mu = 0$ y se varía el parámetro de forma σ , ya que éste último permite considerar diferentes pendientes y de esta manera modificar la ocurrencia de cada modo de falla.

$$- \mu_1 = \mu_2 = \mu_3 = 0 \text{ y } \sigma_1 = \sigma_2 = \sigma_3 = 1$$

En la Figura 5-7 se observa que para este modelo lognormal las tres funciones de distribución marginales son iguales de tal manera que el porcentaje de falla de los modos es muy similar.

Tabla 5-5: Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1$.

Modo	Porcentaje
1	33.29%
2	33.30%
3	33.41%

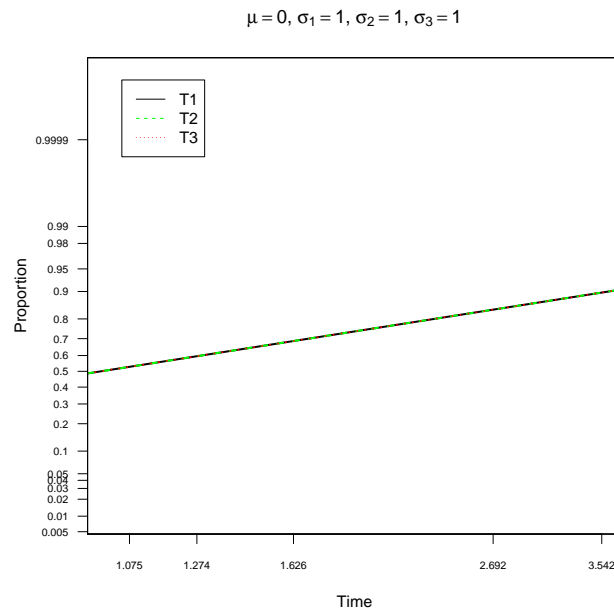


Figura 5-7: Gráfico de probabilidad Weibull para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = \sigma_2 = \sigma_3 = 1$.

– $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8$

En la Figura 5-8 se observa que el modo de falla 1 tiene mayor ocurrencia que los otros dos modos de falla. Se garantiza que haya como mínimo un 15% de fallas para cada uno de los modos.

Tabla 5-6: Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8$.

Modo	Porcentaje
1	27.91%
2	32.65%
3	39.44%

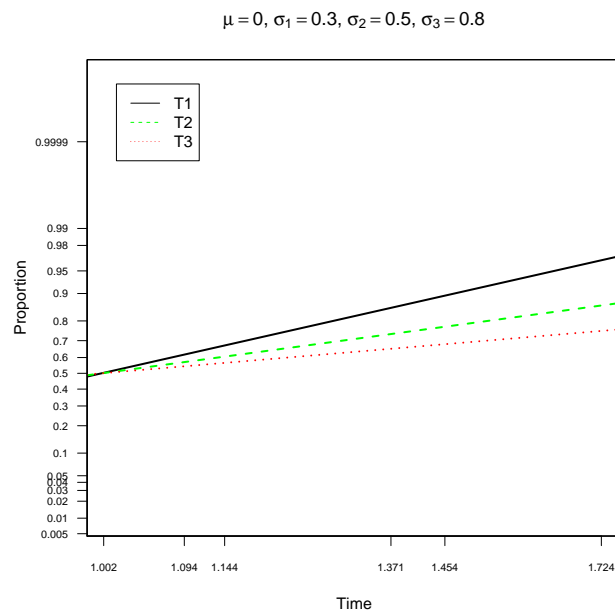


Figura 5-8: Gráfico de probabilidad Weibull para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8$.

– $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1$

En la Figura 5-9 se observa que el modo de falla 3 tiene mayor ocurrencia que los otros dos modos de falla. Se garantiza que haya como mínimo un 15% de fallas para cada uno de los modos.

Tabla 5-7: Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1$.

Modo	Porcentaje
1	39.74%
2	33.07%
3	27.19%

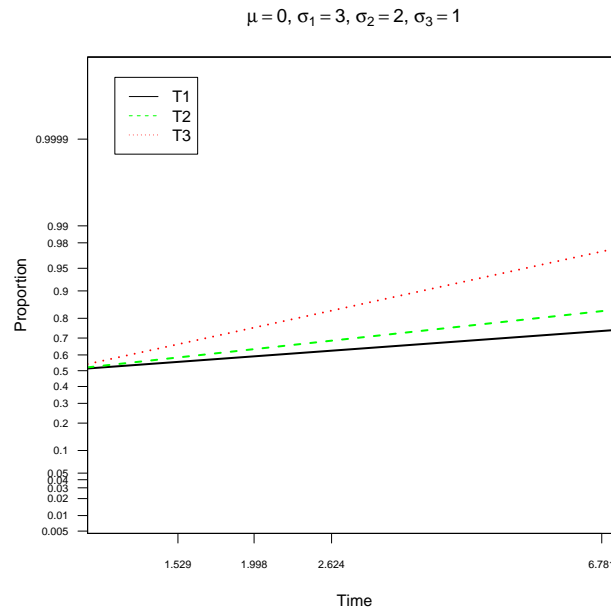


Figura 5-9: Gráfico de probabilidad Weibull para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8$.

Asociados a cada uno de los casos mencionados anteriormente, se consideran tres situaciones diferentes asociadas al parámetro de dependencia entre los tiempos de falla, $\tau = 0.2, 0.5, 0.8$ que corresponden a $\theta = 2, 5.8, 20$, respectivamente. Es decir, para cada uno de los escenarios hay 3 situaciones distintas que abarcan baja, mediana y alta dependencia. De esta manera se consideran 12 casos de simulación para los modelos Weibull y 9 casos de simulación para los modelos lognormales.

2. Se estiman las funciones de supervivencia verdaderas $S(t)$, asumiendo independencia entre los tiempos de falla utilizando el estimador $\widehat{S}^*(t)$ y teniendo en cuenta la dependencia utilizando el estimador $\widehat{S}(t)_{MCR}$. Para realizar estas estimaciones se lleva a cabo la siguiente metodología
 - Se generan muestras de tamaño $n = 50$ de tiempo de falla trivariados, tanto con marginales lognormales como con marginales Weibull. Se elige este tamaño de muestra ya que en general en experimentos de confiabilidad se manejan bases

datos no muy grandes, estos estudios son costosos y en algunas ocasiones los datos de falla son difíciles de obtener.

- En cada una de las muestras de tiempos de falla obtenidas en los casos Weibull y Lognormal, se encuentra $T = \min\{T_1, T_2, T_3\}$ y se etiqueta estos tiempos mínimos con 1, 2 ó 3, indicando si el tiempo proviene del modo de falla 1, 2 ó 3.
 - Se estiman los parámetros asociados a las distribuciones de los tiempos de falla de cada modo de falla, con estos estimadores se obtienen las funciones de sobrevivencia del tiempo mínimo de falla asumiendo independencia.
 - Se estiman las distribuciones marginales asociadas a cada modo de falla mediante el método de combinación de riesgos descrito en 4.2 y con estas funciones de distribución marginales se estima la función de sobrevivencia bajo dependencia utilizando la ecuación (3-4).
3. Se construyen intervalos de confianza empíricos en los percentiles t_p , con $p = 0.05, 0.25, 0.50, 0.75, 0.95$, basados en la simulación Monte Carlo aproximados del 95% para $\widehat{S}^*(t_p)$ y $\widehat{S}(t_p)_{MCR}$, con el fin de observar si estos intervalos contienen el verdadero valor de $S(t_p)$. Estos intervalos de confianza sirven como región de aceptación para realizar pruebas de hipótesis en cada percentil. Se tiene entonces, dos hipótesis nulas en cada punto, en las que se prueban la igualdad estadística de las funciones $\widehat{S}^*(t_p)$ vs. $S(t_p)$ y $\widehat{S}(t_p)_{MCR}$ vs. $S(t_p)$.
 4. Se comparan puntualmente los dos estimadores calculando la eficiencia relativa ER_p de $\widehat{S}^*(t_p)$ con relación $\widehat{S}(t_p)_{MCR}$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

$$ER_p = \frac{ECM(\widehat{S}(t_p)_{MCR})}{ECM(\widehat{S}^*(t_p))} = \frac{E \left[\left(\widehat{S}(t_p)_{MCR} - S(t_p) \right)^2 \right]}{E \left[\left(\widehat{S}^*(t_p) - S(t_p) \right)^2 \right]}$$

donde $ECM(\widehat{S}(t_p)_{MCR})$ y $ECM(\widehat{S}^*(t_p))$ corresponden a los errores cuadráticos medios de los estimadores $\widehat{S}(t_p)_{MCR}$ y $\widehat{S}^*(t_p)$, respectivamente.

Si $ER_p < 1$ entonces $ECM(\widehat{S}^*(t_p)) > ECM(\widehat{S}(t_p)_{MCR})$ lo cual indica que $\widehat{S}^*(t_p)$ es menos eficiente frente a $\widehat{S}(t_p)_{MCR}$ para la estimación de la función de sobrevivencia verdadera en t_p .

5. Se utilizó el software estadístico **R** 2.15.0 para la programación del método de combinación de riesgos. Los detalles del algoritmo se presentan en el Anexo B.

6 Análisis de Resultados

En este capítulo se presentan los resultados obtenidos del estudio, en el que se realizaron comparaciones entre la función de sobrevivencia verdadera del tiempo mínimo de falla $S(t)$ con la función de sobrevivencia estimada asumiendo independencia ($\widehat{S}^*(t)$) y la función de sobrevivencia del tiempo mínimo de falla utilizando el método de combinación de riesgos, teniendo en cuenta la dependencia entre los modos de falla ($\widehat{S}(t)_{MCR}$). Estas comparaciones se realizan para tiempos de falla lognormales y Weibull.

6.1 Casos Weibull

6.1.1 Caso Weibull: $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$

En la Tabla **6-1** se observa que el porcentaje de falla para cada uno de los modos de falla es mayor al 15%, que permite la estimación de la función del tiempo mínimo asumiendo independencia ($\widehat{S}^*(t)$) y dependencia ($\widehat{S}(t)_{MCR}$).

Cabe resaltar que si se compara la Tabla **6-1** con la Tabla **5-1** presentada en la sección 5.3 del capítulo anterior, los porcentajes de ocurrencia para cada modo de falla son diferentes, debido a que este porcentaje se ve influenciado por la dependencia considerada en los modelos, lo cual permite ilustrar claramente el problema de identificabilidad y enfatizar que no se están trabajando con funciones de distribución marginales sino con funciones de subdistribución. El método de combinación de riesgos permite resolver el problema de identificabilidad y estimar las funciones de distribución marginales.

Tabla 6-1: Porcentaje de falla para el modelo Weibull $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$.

Modo	Porcentaje
1	40.00%
2	32.00%
3	28.00%

En las gráficas **6-1**, **6-2** y **6-3** se muestran los límites de confianza para este caso Weibull $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$ para los diferentes niveles de dependencia $\tau = 0.2, 0.5, 0.8$ ($\theta = 2, 5.8, 20$). En estas Figuras se presenta el valor verdadero de la

función de supervivencia conjunta denotado en el gráfico por $S(t)$, los límites inferior y superior de la función de supervivencia estimada bajo dependencia denotados por $LI.S(t)_{MCR}$ y $LS.S(t)_{MCR}$ respectivamente, y los límites inferior y superior de la función de supervivencia estimada bajo independencia denotados por $LI.S^*(t)$ y $LS.S^*(t)$, respectivamente. Estos intervalos de confianza muestran que en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$, el estimador asumiendo dependencia contiene en todos los casos considerados, el verdadero valor de la supervivencia conjunta $S(t_p)$. Lo anterior se puede observar en las Tablas **6-2**, **6-4** y **6-6**.

Tabla 6-2: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0461	0.0744	0.0284	0.0309	0.0811	0.0501
0.25	0.2264	0.2858	0.0594	0.2198	0.3091	0.0893
0.50	0.4793	0.5284	0.0491	0.4669	0.5358	0.0688
0.75	0.7308	0.7731	0.0423	0.7249	0.7813	0.0564
0.95	0.9171	0.9617	0.0446	0.9125	0.9713	0.0588

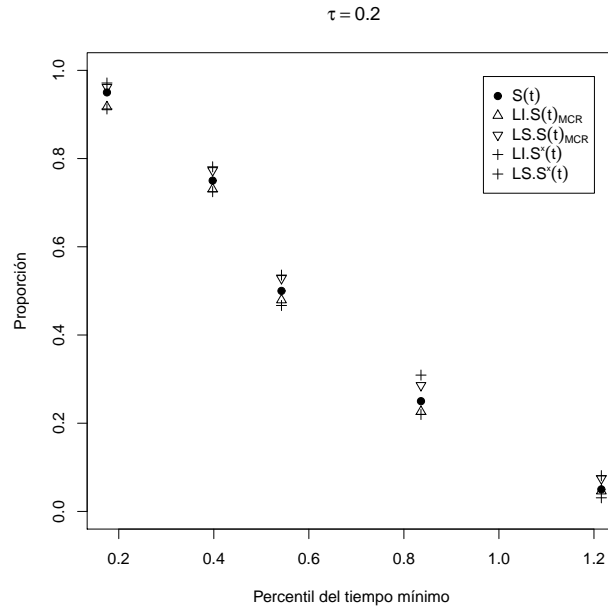


Figura 6-1: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla **6-3** se observa que el porcentaje de falla para cada uno de los modos es mayor al 15% y el modo 3 presenta mayor ocurrencia para este nivel de dependencia $\tau = 0.5$ que corresponde a $\theta = 5.8$.

Tabla 6-3: Porcentaje de falla para el modelo Weibull $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$.

Modo	Porcentaje
1	24.00%
2	30.00%
3	46.00%

Tabla 6-4: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0439	0.0718	0.0279	0.0267	0.0794	0.0527
0.25	0.2203	0.2783	0.0580	0.2412	0.3205	0.0643
0.50	0.4894	0.5328	0.0434	0.4893	0.5515	0.0622
0.75	0.7249	0.7639	0.0390	0.7362	0.7991	0.0629
0.95	0.9264	0.9591	0.0328	0.9250	0.9819	0.0569

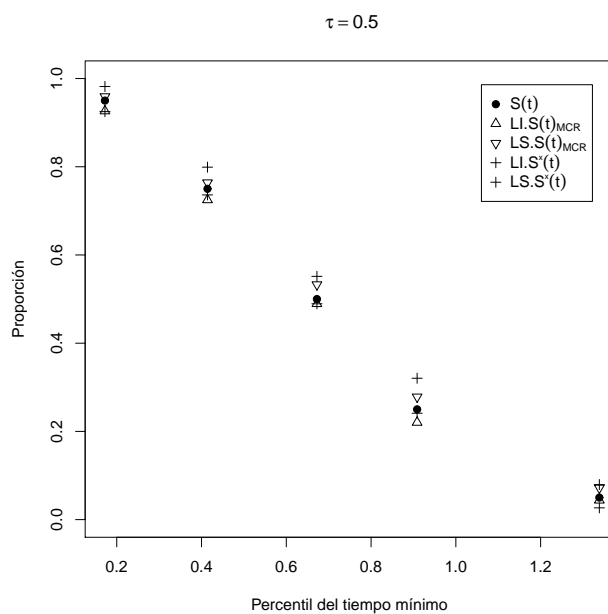


Figura 6-2: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla 6-5 se observa que el porcentaje de falla para cada uno de los modos es mayor al 15% y el modo 2 es el que presenta menor ocurrencia para este nivel de dependencia $\tau = 0.8$ que corresponde a $\theta = 20$.

Tabla 6-5: Porcentaje de falla para el modelo Weibull $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$.

Modo	Porcentaje
1	36.00%
2	26.00%
3	38.00%

Tabla 6-6: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0395	0.0642	0.0247	0.0303	0.0850	0.0547
0.25	0.2327	0.2647	0.0320	0.2050	0.2714	0.0664
0.50	0.4810	0.5191	0.0381	0.4609	0.5284	0.0675
0.75	0.7202	0.7509	0.0307	0.7410	0.8280	0.0684
0.95	0.9459	0.9611	0.0152	0.9272	0.9895	0.0623

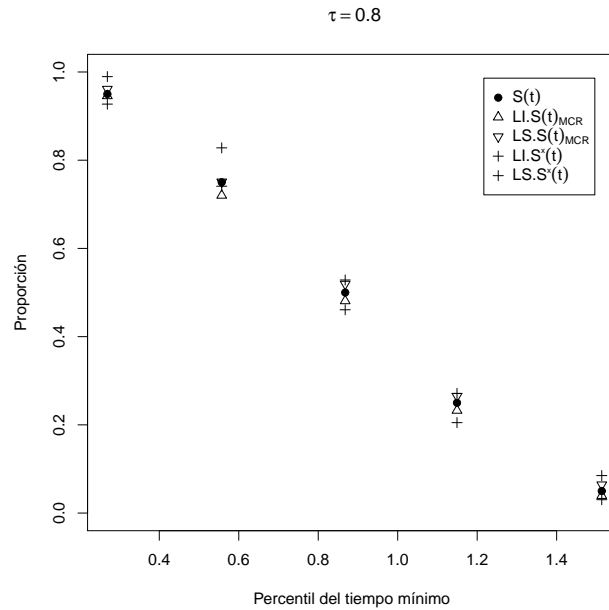


Figura 6-3: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 2, \beta_2 = 2, \beta_3 = 2, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla 6-7 se muestra la eficiencia relativa ER_p de la función de supervivencia conjunta estimada bajo independencia $\widehat{S}^*(t_p)$ con relación a la función de supervivencia conjunta estimada bajo dependencia $\widehat{S}(t_p)_{MCR}$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$, para los diferentes niveles de dependencia $\tau = 0.2, 0.5, 0.8$, la cual indica que el estimador $\widehat{S}^*(t_p)$ es menos eficiente que $\widehat{S}(t_p)_{MCR}$ para estimar $S(t_p)$, debido a que los valores encontrados de ER_p son todos menores que 1. Se observa que a medida que aumenta la dependencia el valor de la eficiencia relativa disminuye.

Tabla 6-7: Eficiencia relativa.

τ	$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
0.2	0.2445	0.1238	0.1682	0.2088	0.1893
0.5	0.2026	0.1082	0.1502	0.1450	0.1759
0.8	0.1355	0.0713	0.0897	0.1046	0.1178

6.1.2 Caso Weibull: $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$

En la Tabla 6-8 se observa que el porcentaje de falla para cada uno de los modos es mayor al 15% y que el modo 3 presenta mayor ocurrencia con respecto a los otros dos modos de falla.

Tabla 6-8: Porcentaje de falla para el modelo Weibull $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$.

Modo	Porcentaje
1	20.00%
2	32.00%
3	48.00%

Los límites de confianza para este caso Weibull se muestran en las graficas 6-4, 6-5 y 6-6, para los diferentes niveles de dependencia $\tau = 0.2, 0.5, 0.8$. Estas Figuras muestran que en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$, el estimador asumiendo dependencia contiene el verdadero valor de la sobrevivencia conjunta $S(t_p)$. Los límites para $\widehat{S}^*(t_p)$ son más amplios que los encontrados para $\widehat{S}(t_p)_{MCR}$, indicando con ello que el estimador $\widehat{S}(t_p)_{MCR}$ tiene mayor precisión para estimar $S(t_p)$ (ver Tablas 6-9, 6-11 y 6-13).

Tabla 6-9: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0465	0.0845	0.0381	0.0444	0.0985	0.0541
0.25	0.2386	0.2802	0.0417	0.2230	0.2891	0.0661
0.50	0.4878	0.5361	0.0483	0.4787	0.5394	0.0607
0.75	0.7199	0.7703	0.0504	0.6774	0.7554	0.0780
0.95	0.9293	0.9709	0.0415	0.9195	0.9759	0.0564

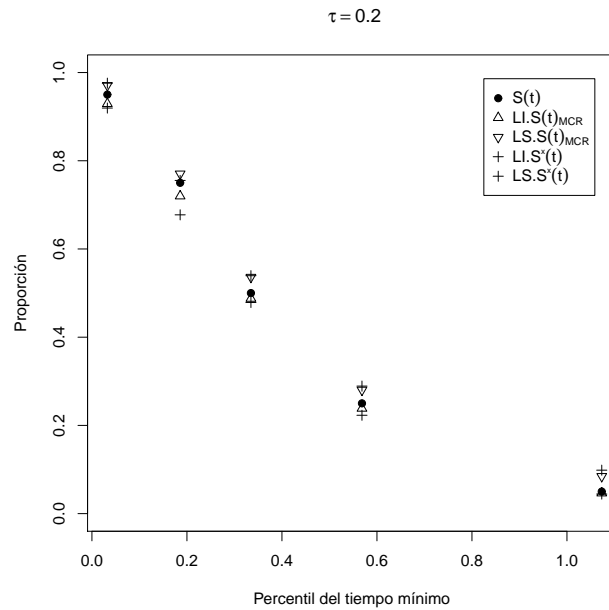


Figura 6-4: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla **6-10** se observa que el porcentaje de falla para cada uno de los modos es mayor al 15%.

Tabla 6-10: Porcentaje de falla para el modelo Weibull $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$.

Modo	Porcentaje
1	40.00%
2	26.00%
3	34.00%

Tabla 6-11: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0396	0.0723	0.0327	0.0309	0.0843	0.0534
0.25	0.2424	0.2798	0.0374	0.2394	0.3013	0.0619
0.50	0.4860	0.5261	0.0401	0.4715	0.5392	0.0677
0.75	0.7251	0.7699	0.0447	0.7452	0.8208	0.0757
0.95	0.9271	0.9618	0.0348	0.9214	0.9748	0.0535

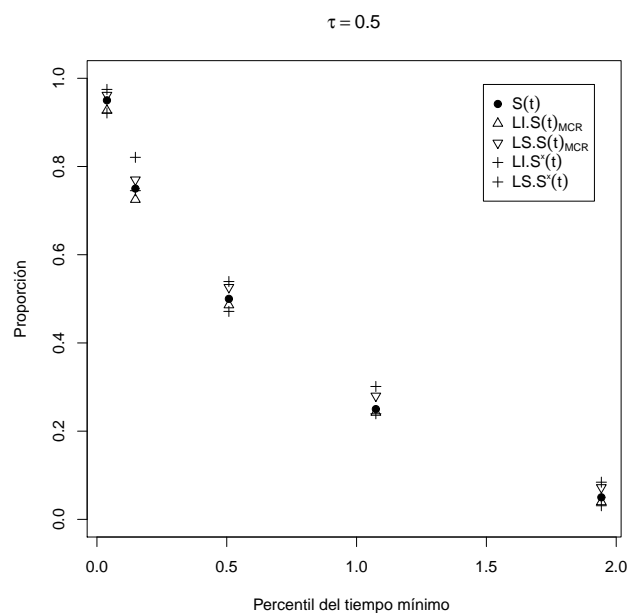


Figura 6-5: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla **6-12** se observa que el porcentaje de falla para cada uno de los modos es mayor al 15% y el modo de falla 3 es el que presenta mayor ocurrencia para este nivel de dependencia $\tau = 0.8$ que corresponde a $\theta = 20$.

Tabla 6-12: Porcentaje de falla para el modelo Weibull $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$.

Modo	Porcentaje
1	28.00%
2	28.00%
3	44.00%

Tabla 6-13: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0384	0.0633	0.0249	0.0390	0.1019	0.0629
0.25	0.2330	0.2634	0.0304	0.2148	0.2788	0.0640
0.50	0.4737	0.5105	0.0368	0.4903	0.5701	0.0609
0.75	0.7224	0.7622	0.0398	0.6999	0.7784	0.0785
0.95	0.9374	0.9594	0.0220	0.9154	0.9696	0.0542

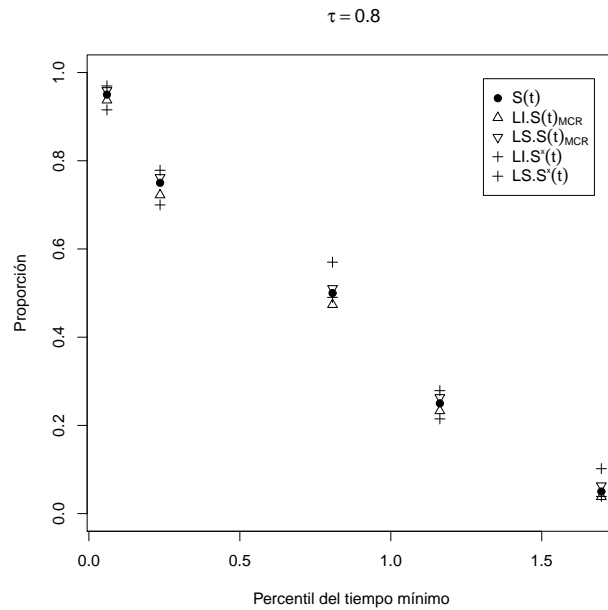


Figura 6-6: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla **6-14** se muestra la eficiencia relativa ER_p de la función de supervivencia conjunta estimada bajo independencia $\widehat{S}^*(t_p)$ con relación a la función de supervivencia conjunta estimada bajo dependencia $\widehat{S}(t_p)_{MCR}$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$, para el caso Weibull $\beta_1 = 1.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$, para los diferentes niveles de dependencia $\tau = 0.2, 0.5, 0.8$. Se observa también en este caso de estudio que el estimador $\widehat{S}^*(t_p)$ es menos eficiente que $\widehat{S}(t_p)_{MCR}$ para estimar $S(t_p)$, debido a que los valores encontrados de ER_p son todos menores que 1 y a medida que aumenta la dependencia el valor de la eficiencia relativa disminuye.

Tabla 6-14: Eficiencia relativa.

τ	$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
0.2	0.3004	0.2203	0.2301	0.1814	0.2751
0.5	0.2116	0.1876	0.1500	0.1055	0.2176
0.8	0.1204	0.1073	0.1375	0.0618	0.1869

6.1.3 Caso Weibull: $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$

El modo 1 presenta mayor ocurrencia que los otros dos modos de falla como se puede observar en la Tabla 6-15. Además el porcentaje de falla para cada uno de los modos es mayor al 15%.

Tabla 6-15: Porcentaje de falla para el modelo Weibull $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$.

Modo	Porcentaje
1	50.00%
2	28.00%
3	22.00%

Los límites de confianza se muestran en las Figuras 6-7, 6-8 y 6-9. Los intervalos de confianza para el estimador $\widehat{S}(t_p)_{MCR}$ contiene el verdadero valor de $S(t_p)$. Lo anterior se puede observar en las Tablas 6-16, 6-18 y 6-21.

Tabla 6-16: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0472	0.0759	0.0287	0.0308	0.0797	0.0489
0.25	0.2310	0.2701	0.0391	0.2165	0.2828	0.0663
0.50	0.4752	0.5131	0.0379	0.4818	0.5463	0.0645
0.75	0.7228	0.7653	0.0425	0.7241	0.7930	0.0688
0.95	0.9275	0.9620	0.0345	0.9133	0.9711	0.0578

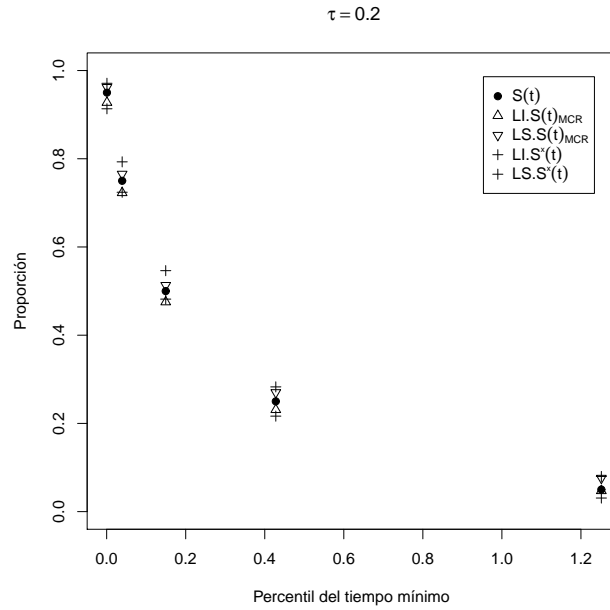


Figura 6-7: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla 6-17 se observa que el porcentaje de falla para cada uno de los modos es mayor al 15%. El modo 2 es el que presenta menor ocurrencia para este nivel de dependencia $\tau = 0.5$ que corresponde a $\theta = 5.8$.

Tabla 6-17: Porcentaje de falla para el modelo Weibull $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$.

Modo	Porcentaje
1	58.00%
2	17.00%
3	25.00%

Tabla 6-18: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0404	0.0663	0.0259	0.0429	0.0929	0.0500
0.25	0.2476	0.2827	0.0351	0.1958	0.2605	0.0647
0.50	0.4909	0.5251	0.0343	0.4830	0.5475	0.0645
0.75	0.7232	0.7620	0.0388	0.7450	0.8536	0.0685
0.95	0.9275	0.9602	0.0327	0.9193	0.9841	0.0648

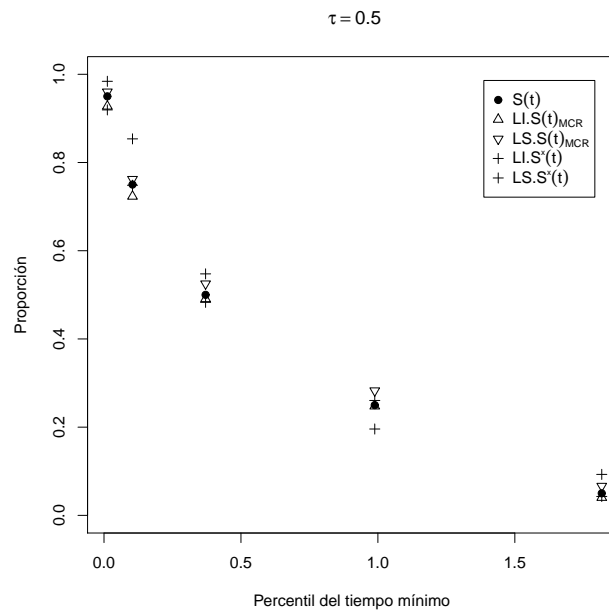


Figura 6-8: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla **6-19** se observa que el modo 2 tiene la menor ocurencia con respecto a los otros dos modos, equivalente a 16%, sin embargo se garantiza que el porcentaje de falla para cada uno de los modos sea mayor al 15%.

Tabla 6-19: Porcentaje de falla para el modelo Weibull $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$.

Modo	Porcentaje
1	63.00%
2	16.00%
3	21.00%

Tabla 6-20: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0411	0.0610	0.0198	0.0235	0.0750	0.0515
0.25	0.2304	0.2607	0.0303	0.2122	0.2896	0.0774
0.50	0.4832	0.5108	0.0276	0.4954	0.5822	0.0668
0.75	0.7300	0.7520	0.0220	0.7135	0.7769	0.0634
0.95	0.9318	0.9590	0.0273	0.9164	0.9799	0.0635

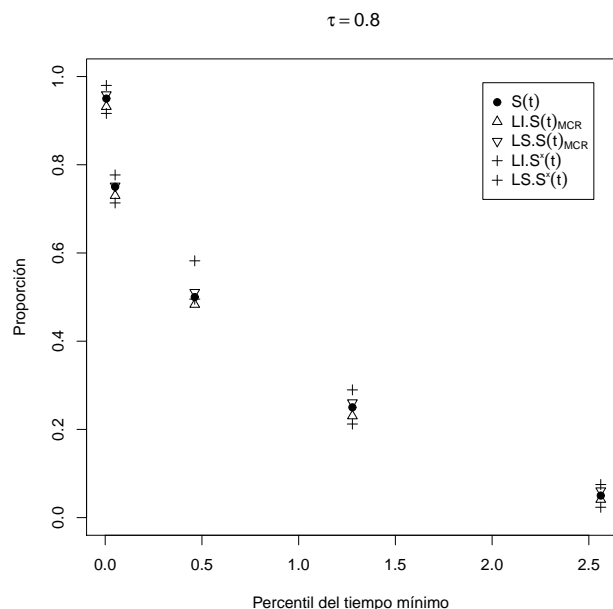


Figura 6-9: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla **6-21** se muestra la eficiencia relativa ER_p de la función de sobrevivencia conjunta estimada bajo independencia $\widehat{S}^*(t_p)$ con relación a la función de sobrevivencia conjunta estimada bajo dependencia $\widehat{S}(t_p)_{MCR}$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$, para el caso Weibull $\beta_1 = 0.5, \beta_2 = 1, \beta_3 = 1, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$, para los diferentes niveles de dependencia $\tau = 0.2, 0.5, 0.8$. Se observa también en este caso de estudio que el estimador $\widehat{S}^*(t_p)$ es menos eficiente que $\widehat{S}(t_p)_{MCR}$ para estimar $S(t_p)$, debido a que los valores encontrados de ER_p son todos menores que 1 y a medida que aumenta la dependencia el valor de la eficiencia relativa disminuye.

Tabla 6-21: Eficiencia relativa.

τ	$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
0.2	0.2584	0.1749	0.1988	0.1426	0.2283
0.5	0.1580	0.1263	0.1320	0.0986	0.1194
0.8	0.1127	0.0440	0.0767	0.0829	0.0817

6.1.4 Caso Weibull: $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1$

En la Tabla **6-22** se observa que el porcentaje de falla para cada uno de los modos es mayor al 15%. El modo 3 presenta mayor ocurrencia para este nivel de dependencia $\tau = 0.2$ equivalente a $\theta = 2$.

Tabla 6-22: Porcentaje de falla para el modelo Weibull $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$.

Modo	Porcentaje
1	26.00%
2	30.00%
3	44.00%

En las Figuras **6-10**, **6-11** y **6-12** se muestran los límites de confianza. Los intervalos de confianza para el estimador $\hat{S}(t_p)_{MCR}$ contiene el verdadero valor de $S(t_p)$. Lo anterior se puede observar en las Tablas **6-23**, **6-25** y **6-27**.

Tabla 6-23: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0331	0.0705	0.0374	0.0353	0.0772	0.0419
0.25	0.2309	0.2864	0.0554	0.2258	0.2824	0.0566
0.50	0.4824	0.5341	0.0517	0.4843	0.5419	0.0576
0.75	0.7235	0.7729	0.0494	0.7144	0.7740	0.0596
0.95	0.9274	0.9719	0.0445	0.9190	0.9729	0.0539

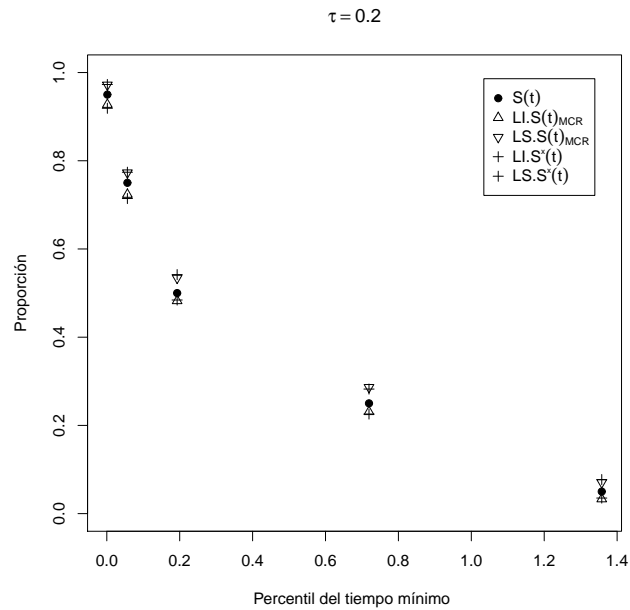


Figura 6-10: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla **6-24** se observa que el porcentaje de falla para cada uno de los modos es mayor al 15%.

Tabla 6-24: Porcentaje de falla para el modelo Weibull $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 1.2, \eta_1 = 1, \eta_2 = 0.5, \eta_3 = 1, \tau = 0.5$.

Modo	Porcentaje
1	32.00%
2	22.00%
3	46.00%

Tabla 6-25: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0413	0.0718	0.0305	0.0316	0.0805	0.0489
0.25	0.2223	0.2716	0.0493	0.2429	0.3040	0.0511
0.50	0.4758	0.5251	0.0493	0.4881	0.5437	0.0556
0.75	0.7191	0.7593	0.0402	0.7079	0.7698	0.0619
0.95	0.9175	0.9600	0.0425	0.9281	0.9835	0.0554

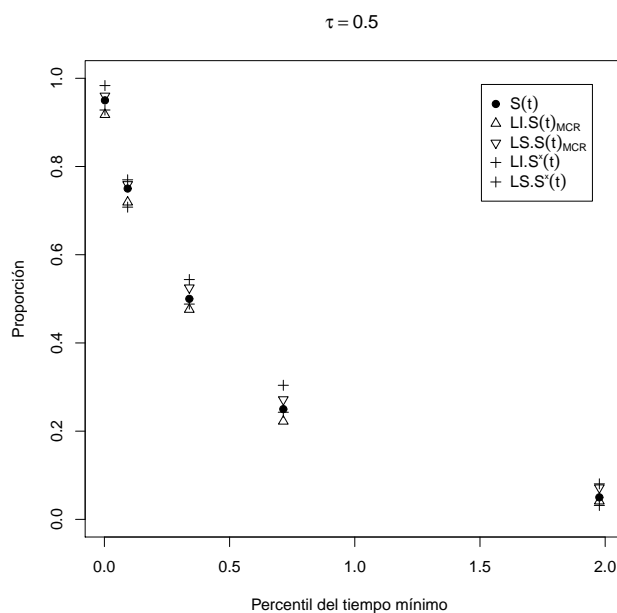


Figura 6-11: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

El modo 3 presenta una alta ocurrencia con respecto a los otros dos modos, como se puede observar en la Tabla 6-26. Se garantiza que el porcentaje de falla para cada uno de los modos sea mayor al 15%.

Tabla 6-26: Porcentaje de falla para el modelo Weibull $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$.

Modo	Porcentaje
1	22.00%
2	16.00%
3	62.00%

Tabla 6-27: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0423	0.0669	0.0246	0.0444	0.0972	0.0527
0.25	0.2362	0.2661	0.0299	0.2298	0.2858	0.0560
0.50	0.4749	0.5048	0.0299	0.4602	0.5244	0.0642
0.75	0.7188	0.7483	0.0296	0.7416	0.8310	0.0663
0.95	0.9375	0.9620	0.0245	0.9099	0.9627	0.0528

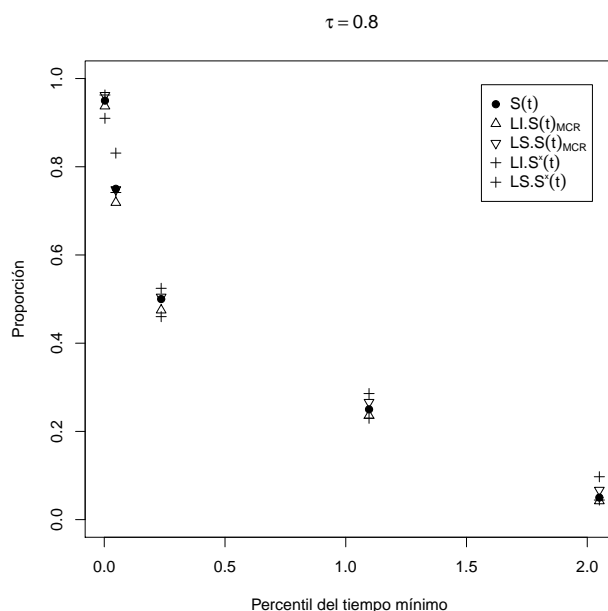


Figura 6-12: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\beta_1 = 1.2, \beta_2 = 1, \beta_3 = 0.5, \eta_1 = 1, \eta_2 = 1, \eta_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

La eficiencia relativa ER_p de la función de sobrevivencia conjunta estimada bajo independencia $\widehat{S}^*(t_p)$ con relación a la función de sobrevivencia conjunta estimada bajo dependencia $\widehat{S}(t_p)_{MCR}$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$ se presenta en la Tabla **6-28**. Se observa también en este caso de estudio que el estimador $\widehat{S}^*(t_p)$ es menos eficiente que $\widehat{S}(t_p)_{MCR}$ para estimar $S(t_p)$, debido a que los valores encontrados de ER_p son todos menores que 1 y a medida que aumenta la dependencia el valor de la eficiencia relativa disminuye.

Tabla 6-28: Eficiencia relativa.

τ	$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
0.2	0.2691	0.2117	0.2017	0.1885	0.2380
0.5	0.1768	0.1486	0.1233	0.0906	0.1279
0.8	0.1190	0.1091	0.0776	0.0549	0.1104

6.2 Casos Lognormal

6.2.1 Caso Lognormal: $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8$

En la Tabla **6-29** se observa que el porcentaje de falla para cada uno de los modos es mayor al 15%. El modo 1 es el que presenta menor ocurrencia para este nivel de dependencia $\tau = 0.2$.

Tabla 6-29: Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.2$.

Modo	Porcentaje
1	20.00%
2	40.00%
3	40.00%

En las Figuras **6-13**, **6-14** y **6-15** se muestran los límites de confianza para este caso lognormal para los diferentes niveles de dependencia $\tau = 0.2, 0.5, 0.8$, correspondientes a $\theta = 2, 5.8, 20$. En estos gráficos se presenta el valor verdadero de la función de sobrevivencia conjunta denotado en el gráfico por $S(t)$, los límites inferior y superior de la función de sobrevivencia estimada bajo dependencia denotados por $LI.S(t)_{MCR}$ y $LS.S(t)_{MCR}$ respectivamente, y los límites inferior y superior de la función de sobrevivencia estimada bajo independencia denotados por $LI.S^*(t)$ y $LS.S^*(t)$, respectivamente. Estos intervalos de confianza muestran que en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$, tanto el estimador asumiendo independencia como el que asume dependencia contienen el verdadero valor de la sobrevivencia conjunta $S(t_p)$. Sin embargo los límites para $\widehat{S}^*(t_p)$ son mucho más amplios que los encontrados para $\widehat{S}(t_p)_{MCR}$, indicando con ello que el estimador $\widehat{S}(t_p)_{MCR}$ tiene mayor precisión para estimar $S(t_p)$. Lo anterior se puede observar en las Tablas **6-30**, **6-32** y **6-34**.

Tabla 6-30: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0314	0.0729	0.0414	0.0262	0.1370	0.1108
0.25	0.2277	0.2978	0.0701	0.1380	0.3709	0.2328
0.50	0.4822	0.5412	0.0590	0.3902	0.6198	0.2296
0.75	0.7226	0.7823	0.0597	0.6033	0.8418	0.2385
0.95	0.9287	0.9720	0.0432	0.8512	0.9698	0.1186

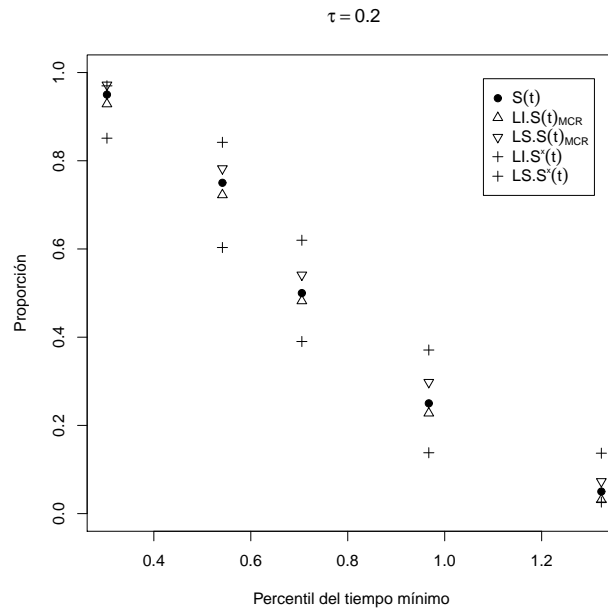


Figura 6-13: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$, $\sigma_1 = \sigma_2 = \sigma_3 = 1$, $\tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla **6-31** se observa que el porcentaje de falla para cada uno de los modos es mayor al 15%.

Tabla 6-31: Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.5$.

Modo	Porcentaje
1	32.00%
2	36.00%
3	32.00%

Tabla 6-32: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0434	0.0776	0.0341	0.0135	0.1479	0.1343
0.25	0.2298	0.2823	0.0525	0.1493	0.3849	0.2356
0.50	0.4909	0.5386	0.0476	0.4196	0.6586	0.2391
0.75	0.7225	0.7705	0.0480	0.6342	0.8618	0.2276
0.95	0.9174	0.9592	0.0418	0.8579	0.9753	0.1173

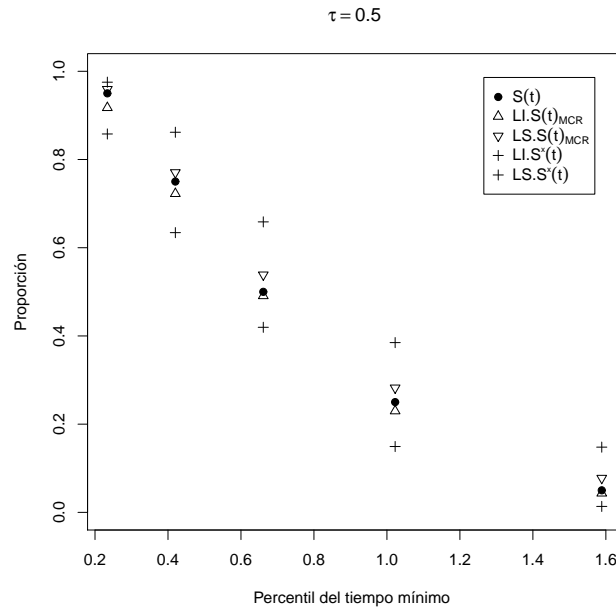


Figura 6-14: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0, \sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

El modo 2 es el que menos ocurre para este nivel de dependencia $\tau = 0.8$ correspondiente a $\theta = 20$, con un porcentaje de falla igual a 16% como se puede observar en la Tabla **6-33**. Sin embargo se garantiza que el porcentaje de falla para cada uno de los modos sea mayor al 15%.

Tabla 6-33: Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.8$.

Modo	Porcentaje
1	38.00%
2	16.00%
3	46.00%

Tabla 6-34: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0382	0.0626	0.0244	0.0037	0.1796	0.1758
0.25	0.2328	0.2651	0.0323	0.1810	0.4346	0.2536
0.50	0.4840	0.5155	0.0315	0.4144	0.6674	0.2530
0.75	0.7301	0.7581	0.0281	0.6450	0.8626	0.2176
0.95	0.9379	0.9620	0.0241	0.8626	0.9626	0.1000

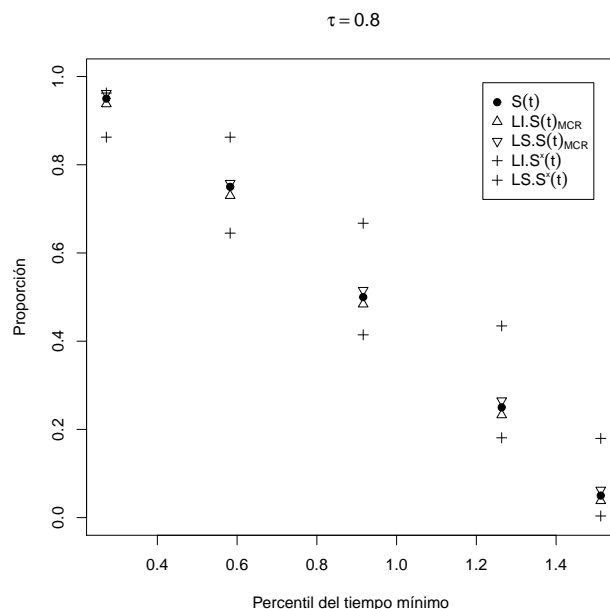


Figura 6-15: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$, $\sigma_1 = 0.3$, $\sigma_2 = 0.5$, $\sigma_3 = 0.8$, $\tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla **6-35** se muestra la eficiencia relativa ER_p de la función de sobrevivencia conjunta estimada bajo independencia $\widehat{S}^*(t_p)$ con relación a la función de sobrevivencia conjunta estimada bajo dependencia $\widehat{S}(t_p)_{MCR}$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$, para el caso lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 0.3, \sigma_2 = 0.5, \sigma_3 = 0.8$, la cual indica que el estimador $\widehat{S}^*(t_p)$ es menos eficiente que $\widehat{S}(t_p)_{MCR}$ para estimar $S(t_p)$, debido a que los valores encontrados de ER_p son todos menores que 1. Se observa además, que a medida que aumenta la dependencia el valor de la eficiencia relativa disminuye.

Tabla 6-35: Eficiencia relativa.

τ	$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
0.2	0.1505	0.0693	0.0869	0.0619	0.1436
0.5	0.1091	0.0417	0.0372	0.0574	0.1383
0.8	0.0745	0.0092	0.0109	0.0347	0.0977

6.2.2 Caso Lognormal: $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1$

En la Tabla **6-36** se observa que el porcentaje de falla para cada uno de los modos es mayor al 15%.

Tabla 6-36: Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.2$.

Modo	Porcentaje
1	32.00%
2	38.00%
3	30.00%

Los límites de confianza se muestran en las Figuras **6-16**, **6-17** y **6-18** . El estimador que asume dependencia $\widehat{S}(t_p)_{MCR}$ contiene el verdadero valor de $S(t_p)$ en todos los casos analizados. Además, los límites para $\widehat{S}^*(t_p)$ son mucho más amplios que los encontrados para $\widehat{S}(t_p)_{MCR}$, por lo tanto el estimador $\widehat{S}(t_p)_{MCR}$ tiene mayor precisión para estimar $S(t_p)$. Lo anterior se puede observar en las Tablas **6-37**, **6-39** y **6-41**.

Tabla 6-37: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0258	0.0833	0.0575	0.0167	0.1122	0.0954
0.25	0.2336	0.2927	0.0591	0.2127	0.4590	0.2462
0.50	0.4730	0.5322	0.0592	0.4447	0.6890	0.2443
0.75	0.7208	0.7808	0.0600	0.6702	0.8629	0.1926
0.95	0.9174	0.9719	0.0545	0.8584	0.9552	0.0568

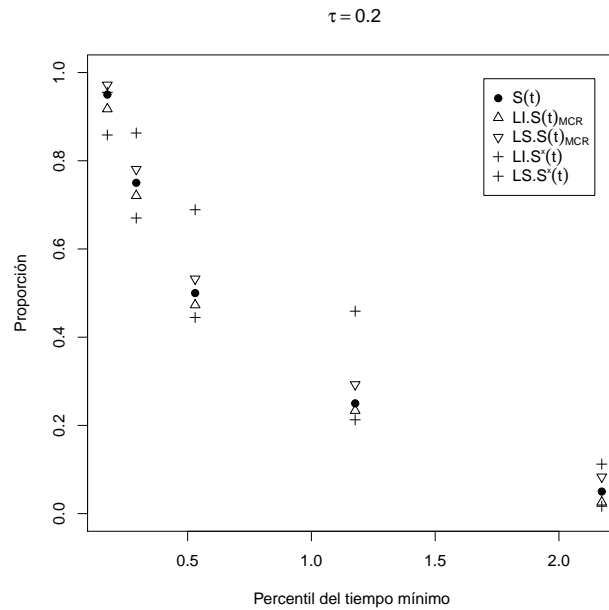


Figura 6-16: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$, $\sigma_1 = \sigma_2 = \sigma_3 = 1$, $\tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla **6-38** se observa que el porcentaje de falla para cada uno de los modos es mayor al 15%. Además es modo 1 presenta mayor ocurrencia con respecto a los otros dos modos de falla.

Tabla 6-38: Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.5$.

Modo	Porcentaje
1	42.00%
2	28.00%
3	30.00%

Tabla 6-39: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0379	0.0728	0.0349	0.0218	0.1327	0.1109
0.25	0.2455	0.2875	0.0420	0.1270	0.3355	0.2084
0.50	0.4873	0.5285	0.0412	0.3934	0.5886	0.1952
0.75	0.7187	0.7663	0.0476	0.6610	0.8566	0.1956
0.95	0.9216	0.9698	0.0482	0.8599	0.9629	0.1030

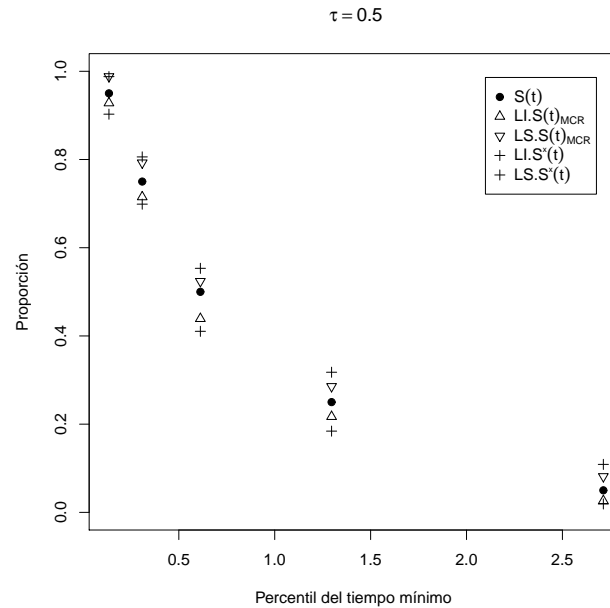


Figura 6-17: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0, \sigma_1 = \sigma_2 = \sigma_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

El modo 3 presenta menor ocurrencia para este nivel de dependencia $\tau = 0.8$ como se observa en la Tabla 6-40. Además se garantiza que el porcentaje de falla para cada uno de los modos sea mayor al 15%.

Tabla 6-40: Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.8$.

Modo	Porcentaje
1	30.00%
2	50.00%
3	20.00%

Tabla 6-41: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 1, \sigma_2 = 1, \sigma_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0329	0.0667	0.0338	0.0400	0.1517	0.1117
0.25	0.2310	0.2600	0.0290	0.1346	0.3885	0.2539
0.50	0.4792	0.5196	0.0404	0.3184	0.6087	0.2903
0.75	0.7238	0.7527	0.0289	0.5667	0.8680	0.3013
0.95	0.9193	0.9611	0.0419	0.8718	0.9869	0.1152

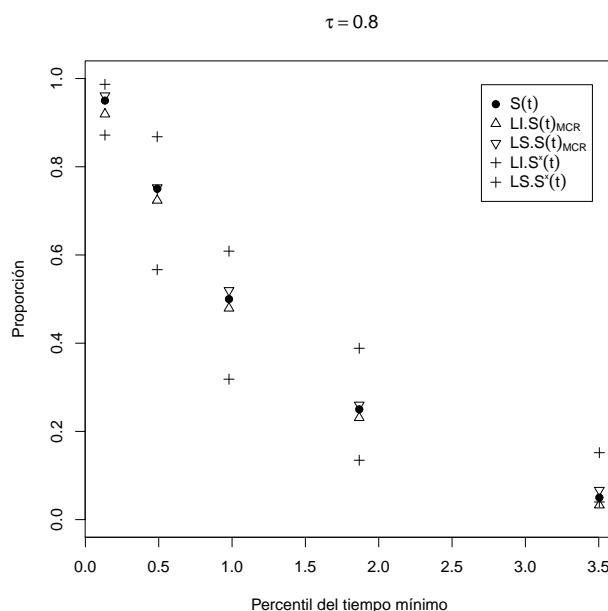


Figura 6-18: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$, $\sigma_1 = \sigma_2 = \sigma_3 = 1$, $\tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

La eficiencia relativa ER_p de la función de sobrevivencia conjunta estimada bajo independencia $\widehat{S}^*(t_p)$ con relación a la función de sobrevivencia conjunta estimada bajo dependencia $\widehat{S}(t_p)_{MCR}$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$, para el caso lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = \sigma_2 = \sigma_3 = 1$ se presenta en la Tabla 6-42, en donde se observa que el estimador $\widehat{S}^*(t_p)$ es menos eficiente que $\widehat{S}(t_p)_{MCR}$ para estimar $S(t_p)$, debido a que los valores encontrados de ER_p son todos menores que 1. Además de ello cuando se aumenta la dependencia el valor de la eficiencia relativa disminuye.

Tabla 6-42: Eficiencia relativa.

τ	$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
0.2	0.2573	0.0712	0.0875	0.1780	0.3020
0.5	0.1840	0.0227	0.0501	0.0595	0.2003
0.8	0.1359	0.0909	0.0209	0.0091	0.1215

6.2.3 Caso Lognormal: $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1$

En la Tabla 6-43 se observa que el porcentaje de falla para cada uno de los modos es mayor al 15%.

Tabla 6-43: Porcentaje de falla para el modelo lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.2$.

Modo	Porcentaje
1	36.00%
2	26.00%
3	38.00%

En las Figuras **6-19**, **6-20** y **6-21** se muestran los límites de confianza para este caso lognormal para los diferentes niveles de dependencia $\tau = 0.2, 0.5, 0.8$, que corresponden a valores de $\theta = 2, 5, 8, 20$. Estos intervalos de confianza muestran que en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$, tanto el estimador asumiendo independencia como el que asume dependencia contienen el verdadero valor de la sobrevivencia conjunta $S(t_p)$. Cabe resaltar que los límites para $\widehat{S}^*(t_p)$ son mucho más amplios que los encontrados para $\widehat{S}(t_p)_{MCR}$, indicando con ello que el estimador $\widehat{S}(t_p)_{MCR}$ tiene mayor precisión para estimar $S(t_p)$. Lo anterior se puede observar en las Tablas **6-44**, **6-46** y **6-48**.

Tabla 6-44: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0336	0.0831	0.0495	0.0154	0.1457	0.1303
0.25	0.2163	0.2762	0.0599	0.1429	0.3620	0.2192
0.50	0.4810	0.5260	0.0450	0.3501	0.5892	0.2391
0.75	0.7005	0.7606	0.0601	0.6683	0.8621	0.1938
0.95	0.9276	0.9800	0.0524	0.8622	0.9742	0.1120

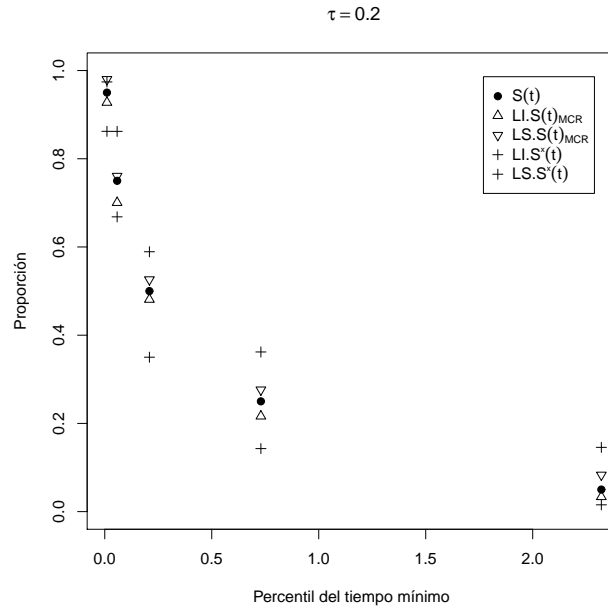


Figura 6-19: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.2$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

Para este nivel de dependencia $\tau = 0.5$ que corresponde a $\theta = 5.8$, el modo 1 es el que presenta mayor ocurrencia con respecto a los otros dos modos, lo cual se observa en la Tabla 6-45. Además se garantiza que el porcentaje de falla para cada uno de los modos sea mayor al 15%.

Tabla 6-45: Porcentaje de falla para el modelo Weibull $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.5$.

Modo	Porcentaje
1	42.00%
2	24.00%
3	34.00%

Tabla 6-46: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0392	0.0736	0.0345	0.0318	0.1672	0.1354
0.25	0.2438	0.2925	0.0487	0.1358	0.3553	0.2195
0.50	0.4893	0.5204	0.0311	0.4029	0.6250	0.2221
0.75	0.7208	0.7690	0.0482	0.6166	0.8262	0.2096
0.95	0.9175	0.9632	0.0457	0.8662	0.9732	0.1070

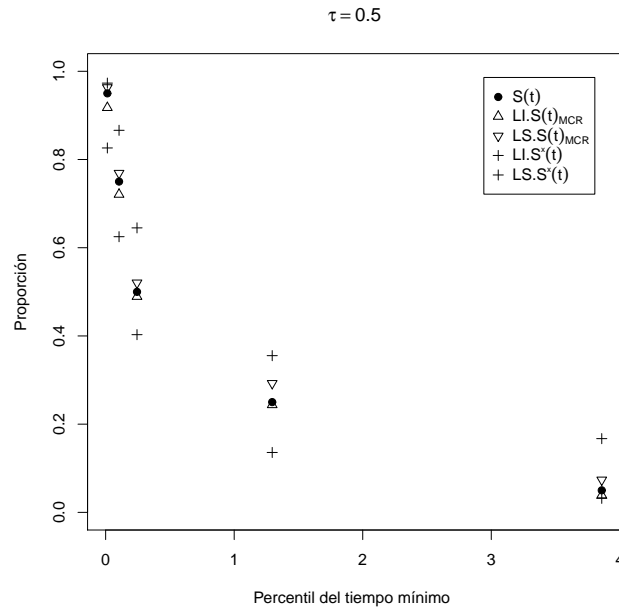


Figura 6-20: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.5$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla **6-47** se observa que el porcentaje de falla para cada uno de los modos es mayor al 15% y que el modo 2 es el que menos ocurre para este nivel de dependencia $\tau = 0.8$ que corresponde a $\theta = 20$.

Tabla 6-47: Porcentaje de falla para el modelo Weibull $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.8$.

Modo	Porcentaje
1	38.00%
2	16.00%
3	46.00%

Tabla 6-48: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

p	$LI.S(t)_{MCR}$	$LS.S(t)_{MCR}$	Precisión	$LI.S^*(t)$	$LS.S^*(t)$	Precisión
0.05	0.0412	0.0637	0.0225	0.0186	0.1430	0.1244
0.25	0.2300	0.2601	0.0301	0.1817	0.4336	0.2519
0.50	0.4833	0.5128	0.0296	0.4076	0.6471	0.2395
0.75	0.7188	0.7487	0.0298	0.6276	0.8539	0.2263
0.95	0.9367	0.9612	0.0245	0.8661	0.9769	0.1107

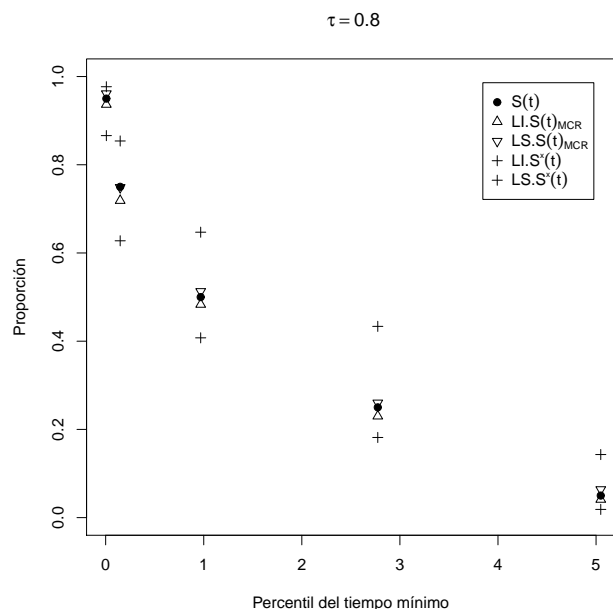


Figura 6-21: Intervalos de confianza empíricos aproximados del 95% para $S(t_p)$ para $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1, \tau = 0.8$ en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.

En la Tabla 6-49 se muestra la eficiencia relativa ER_p de la función de sobrevivencia conjunta estimada bajo independencia $\widehat{S}^*(t_p)$ con relación a la función de sobrevivencia conjunta estimada bajo dependencia $\widehat{S}(t_p)_{MCR}$ en los percentiles $p = 0.05, 0.25, 0.50, 0.75, 0.95$, para el caso lognormal $\mu_1 = \mu_2 = \mu_3 = 0$ y $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1$, en donde se observa que el estimador $\widehat{S}^*(t_p)$ es menos eficiente que $\widehat{S}(t_p)_{MCR}$ para estimar $S(t_p)$, debido a que los valores encontrados de ER_p son todos menores que 1. Además de ello cuando se aumenta la dependencia el valor de la eficiencia relativa disminuye.

Tabla 6-49: Eficiencia relativa.

τ	$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
0.2	0.2378	0.1493	0.1166	0.1933	0.2674
0.5	0.1346	0.0852	0.0783	0.1033	0.1696
0.8	0.1200	0.0358	0.0583	0.8222	0.1527

7 Conclusiones y Trabajo Futuro

- El coeficiente de correlación ρ_L no resulta ser una buena medida de dependencia entre las variables aleatorias lognormales, puesto que este coeficiente de correlación no es invariante bajo transformaciones monótonas. Por lo tanto, se propone utilizar el τ de Kendall que permite controlar la dependencia entre las variables aleatorias lognormales.
- El método de combinación de riesgos es una buena herramienta para estimar las funciones de distribución marginal y la función de sobrevivencia del tiempo mínimo cuando se tiene dependencia entre los riesgos competitivos.
- En todos los casos analizados el estimador $\widehat{S}^*(t_p)$ resultó ineficiente frente al estimador $\widehat{S}(t_p)_{MCR}$. En los casos lognormales los intervalos de confianza para $\widehat{S}^*(t_p)$ resultaron mucho más amplios que los encontrados en los casos Weibull.
- En los casos estudiados el estimador $\widehat{S}(t_p)_{MCR}$ es mejor estimador que $\widehat{S}^*(t_p)$ para la estimación de $S(t_p)$. Por lo tanto se recomienda utilizar el estimador $\widehat{S}(t_p)_{MCR}$, cuando no se tiene certeza que los tiempos de falla son independientes.
- Yáñez, Manotas, Lopera, and Jaramillo (2008) encontraron que el supuesto de independencia en el contexto de riesgos competitivos no genera errores significativos en la estimación de la sobrevivencia. Sin embargo, Yáñez, Brango, Lopera, and Jaramillo (2011) y Brango (2011) encontraron que la estimación de la sobrevivencia bajo este supuesto de independencia puede subestimar o sobreestimar la sobrevivencia verdadera. Para los casos analizados en este trabajo, se encontró que de igual manera, el asumir independencia puede subestimar o sobreestimar la sobrevivencia verdadera, debido a que los intervalos de confianza empíricos son mucho más amplios en comparación a los encontrados para el método de combinación de riesgos.
- En el caso en que los riesgos competitivos no estén igualmente correlacionados, el método de combinación de riesgos implementado en este trabajo no funciona. Este sería un tópico importante para el trabajo futuro tratando de adecuar el método de combinación de riesgos para tales efectos, ya que en muchas disciplinas los riesgos competitivos no siempre se encuentran igualmente correlacionados (Nelsen, 2006).

A Anexo: Funcionamiento del Estimador Cópula Gráfico

Se realiza una comparación entre las funciones de distribución marginales de X y Y reales con las funciones de distribución marginales estimadas mediante el algoritmo cópula gráfico. También se compara la función de sobrevivencia conjunta de la Weibull bivariada con la estimada por medio de este estimador. Las comparaciones se realizan vía simulación utilizando el gráfico de probabilidad Weibull y el error cuadrático medio.

A.1 Comparación entre las funciones de distribución marginales verdaderas y las estimadas por el algoritmo cópula gráfico

La comparación se realiza vía simulación de la siguiente manera:

1. La función de sobrevivencia conjunta de la Weibull bivariada esta dada por

$$S(t_1, t_2) = \exp \left\{ - \left[\left(\frac{t_1}{\eta_1} \right)^{\beta_1/(1-\tau)} + \left(\frac{t_2}{\eta_2} \right)^{\beta_2/(1-\tau)} \right]^{(1-\tau)} \right\}, 0 \leq \tau < 1, \quad (\text{A-1})$$

donde (β_i, η_i) , $i = 1, 2$ corresponden al parámetro de forma y de escala respectivamente de cada una de las funciones de distribución marginales X y Y , τ es el parámetro de dependencia y las distribuciones marginales son Weibull.

Esta forma corresponde a la cópula de Gumbel-Hougaard de manera que para simular se hace a partir de ella utilizando el paquete cópula de R.

2. Los casos de tiempos Weibull a estudiar son:

- Tiempos de falla con parámetros de forma β_i iguales: $\beta_1 = 2$ y $\beta_2 = 2$.
- Tiempos de falla con parámetros de forma β_i diferentes: $\beta_1 = 2.5$ y $\beta_2 = 2$.

Los parámetros de escala de la distribución Weibull bivariada se fijan $\eta_1 = \eta_2 = 1$, ya que el parámetro de dependencia, no depende del parámetro de escala (Lu and Bhattacharyya, 1990). Estos modelos Weibull se seleccionan sólo con el objetivo de chequear empíricamente el estimador cópula gráfico. Para otros casos de estudio ver (Branco, 2011)

3. Asociados a cada uno de los dos escenarios de tiempos de falla Weibull expuestos anteriormente, hay 4 situaciones diferentes de acuerdo al parámetro de dependencia τ de Kendall, que toma los siguientes valores $\tau = 0.2, 0.5, 0.8, 0.9$. Se considera además dos tamaños de muestra $n = 50$ y $n = 200$ de tiempos bivariados Weibull.
4. El chequeo procede de la siguiente manera:
 - Se analizan gráficos de probabilidad Weibull para la función de distribución marginal de X y Y , para determinar la bondad de ajuste a la distribución Weibull. Si el estimador cópula gráfico es correcto, es de esperar que se ajuste a una Weibull ya que las marginales de un distribución Weibull bivariada son Weibull univariadas con parámetros $(\beta_i, \eta_i), i = 1, 2$.
 - Se calcula el error cuadrático medio (ECM) entre la función de distribución marginal verdadera en los percentiles t_p , con $p = 0.25, 0.50, 0.95$, $F_i(t_p), i = 1, 2$, y la función de distribución marginal estimada por el algoritmo cópula gráfico $\hat{F}_i(t_p), i = 1, 2$.

$$ECM(\hat{F}_i(t_p)) = \frac{1}{N} \sum_{j=1}^N \left(\hat{F}_i(t_p) - F(t_p) \right)^2, i = 1, 2,$$

donde N es el número de muestras independientes para cada uno de los escenarios de simulación planteados. Se considera $N = 1000$.

A.1.1 Función de Distribución Marginal de X

A continuación se presenta un análisis de bondad de ajuste para la función de distribución marginal del X , para los diferentes niveles de τ de Kendall como se definieron en el numeral 3 la subsección A.1.

Las Figuras **A-1** - **A-4** comparan la función de distribución marginal de X estimada por el algoritmo cópula gráfico y la función de distribución marginal verdadera para los dos tamaños de muestra, $n = 50$ y $n = 200$, para el caso $\beta_1 = 2, \eta_1 = 1$, mediante gráficos de probabilidad Weibull. La línea recta corresponde a la distribución marginal verdadera,

denotada en el gráfico por F-Ver, y los puntos corresponden a los valores de la función de distribución marginal estimada. En las figuras se observa el buen ajuste que tiene la función de distribución marginal estimada por el algoritmo cópula gráfico a la distribución marginal real para todos los niveles de dependencia considerados $\tau = 0.2, 0.5, 0.8, 0.9$, mejorando el ajuste cuando el tamaño de muestra es mayor. Cabe resaltar que en las Figuras A-2 y A-4 el ajuste en la cola inferior para el tamaño de muestra $n = 200$ no es tan bueno. Sin embargo la existencia de mayor variabilidad en las colas es un comportamiento de esperar en estos gráficos de probabilidad.

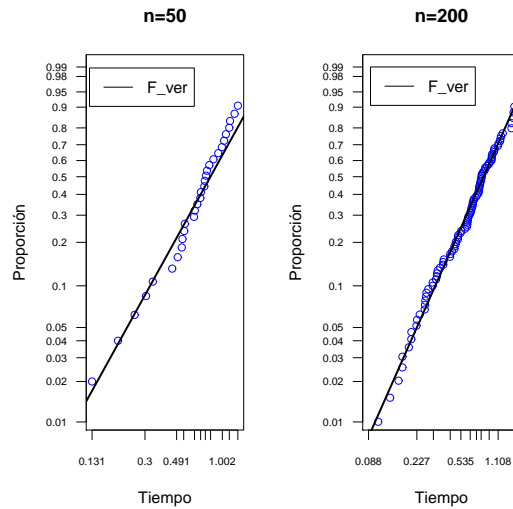


Figura A-1: Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2, \eta_1 = 1, \tau = 0.2$, para $n = 50$ y $n = 200$.

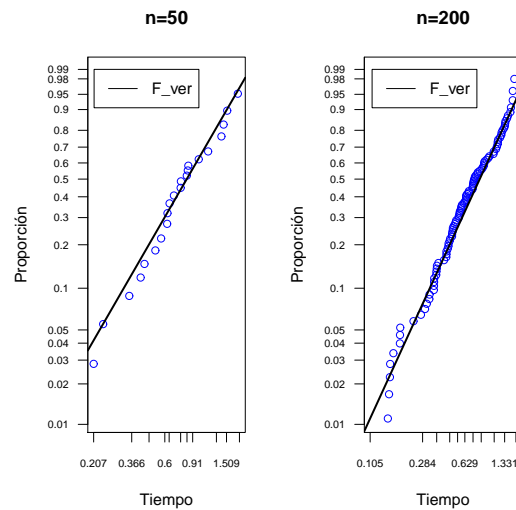


Figura A-2: Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2, \eta_1 = 1, \tau = 0.5$, para $n = 50$ y $n = 200$.

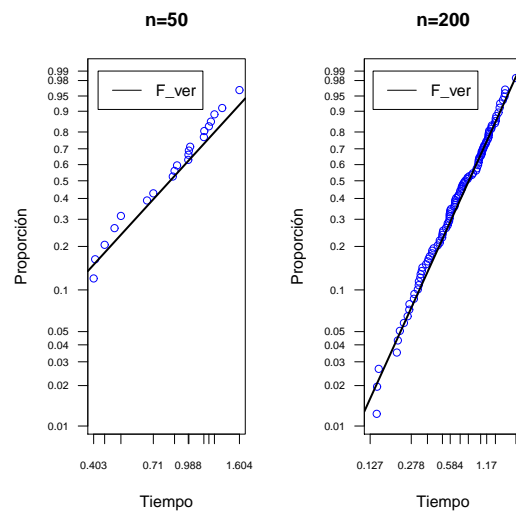


Figura A-3: Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2, \eta_1 = 1, \tau = 0.8$, para $n = 50$ y $n = 200$.

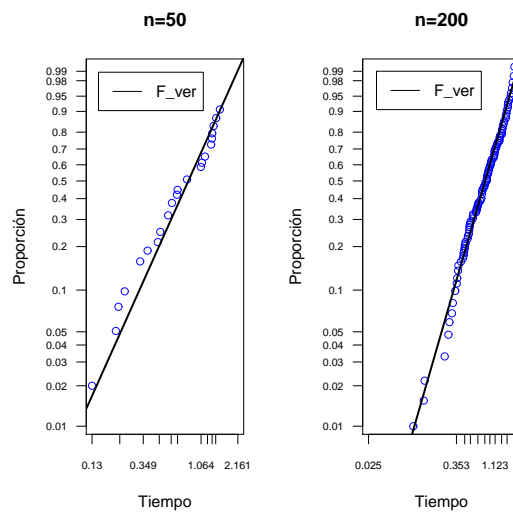


Figura A-4: Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2, \eta_1 = 1, \tau = 0.9$, para $n = 50$ y $n = 200$.

Las Figuras **A-5** - **A-8** comparan la función de distribución marginal de X estimada por el algoritmo cópula gráfico y la función de distribución marginal verdadera para los dos tamaños de muestra, $n = 50$ y $n = 200$, para el caso $\beta_1 = 2.5, \eta_1 = 1$, mediante gráficos de probabilidad Weibull. En las figuras se observa el buen ajuste que tiene la función de distribución marginal estimada por el algoritmo cópula gráfico a la distribución marginal real, mejorando el ajuste cuando el tamaño de muestra es mayor, indicando que el algoritmo cópula gráfico estima correctamente la marginal X en los dos casos considerados. En la Figura **A-6** se observa que en la cola inferior no se presenta un buen ajuste de la función de distribución marginal estimada por el estimador cópula gráfico a la función de distribución marginal real.

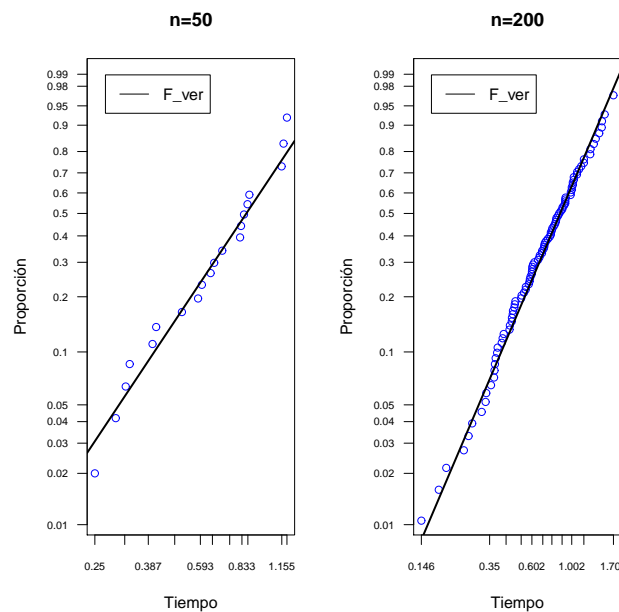


Figura A-5: Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2.5, \eta_1 = 1, \tau = 0.2$, para $n = 50$ y $n = 200$.

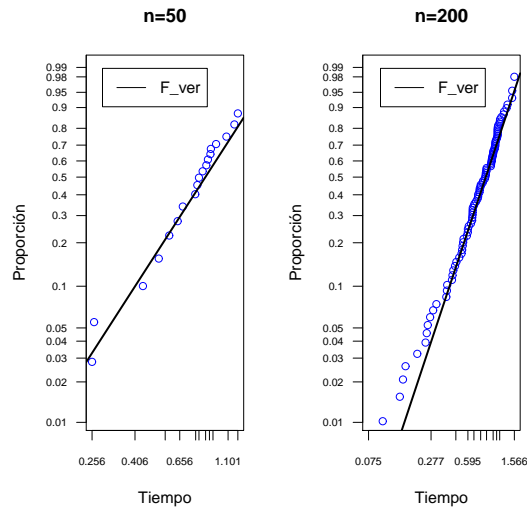


Figura A-6: Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2.5, \eta_1 = 1, \tau = 0.5$, para $n = 50$ y $n = 200$.

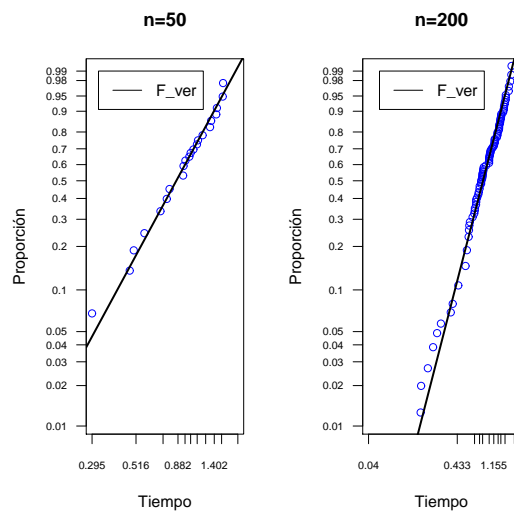


Figura A-7: Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2.5, \eta_1 = 1, \tau = 0.8$, para $n = 50$ y $n = 200$.

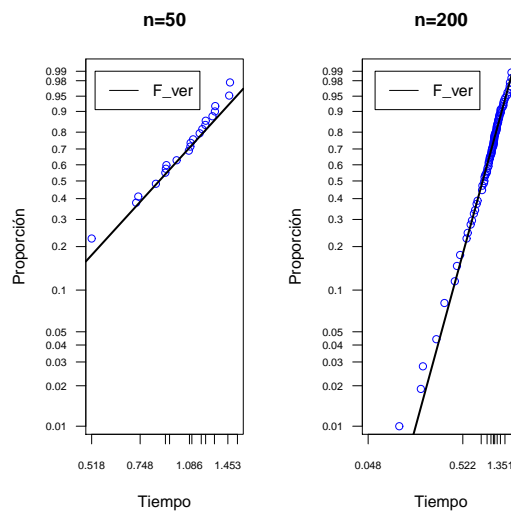


Figura A-8: Gráfico de probabilidad Weibull para la función de distribución marginal de X , $\beta_1 = 2.5$, $\eta_1 = 1$, $\tau = 0.9$, para $n = 50$ y $n = 200$.

La Tabla **A-1** muestra el error cuadrático medio para la función de distribución marginal de X estimada en los percentiles t_p , con $p = 0.25, 0.5, 0.95$ para los diferentes casos de tiempos Weibull descritos en la sección A.1 para los distintos niveles de dependencia $\tau = 0.2, 0.5, 0.8, 0.9$. Se observa que los valores del ECM son pequeños indicando que la función de distribución marginal de X estimada por el algoritmo cópula gráfico se aproxima muy bien a la función de distribución marginal verdadera para los dos tamaños de muestra considerados $n = 50$ y $n = 200$. Es decir las estimaciones son buenas.

Tabla A-1: Error cuadrático medio para la función de distribución marginal de X estimada en los percentiles t_p , con $p = 0.25, 0.5, 0.95$.

β_1	η_1	τ	$n = 50$			$n = 200$		
			$p = 0.25$	$p = 0.50$	$p = 0.95$	$p = 0.25$	$p = 0.5$	$p = 0.95$
2	1	0.2	0.0008	0.0029	0.0035	2.4675×10^{-4}	0.0009	0.0010
2	1	0.5	0.0007	0.0034	0.0019	1.9941×10^{-4}	0.0006	0.0007
2	1	0.8	0.0004	0.0023	0.0015	2.1930×10^{-4}	0.0005	0.0002
2	1	0.9	0.0003	0.0016	0.0012	1.3641×10^{-4}	0.0003	0.0010
2.5	1	0.2	0.0003	0.0025	0.0048	6.7164×10^{-5}	0.0009	0.0010
2.5	1	0.5	0.0001	0.0030	0.0029	9.4249×10^{-5}	0.0030	0.0006
2.5	1	0.8	0.0003	0.0020	0.0018	6.3703×10^{-5}	0.0010	0.0002
2.5	1	0.9	0.0004	0.0011	0.0018	2.4133×10^{-5}	0.0005	0.0003

A.1.2 Función de Distribución Marginal de Y

Las Figuras **A-9** - **A-12** comparan la función de distribución marginal real y la función de distribución marginal de Y estimada por el algoritmo cópula gráfico para los dos tamaños de muestra, $n = 50$ y $n = 200$, mediante gráficos de probabilidad Weibull. La línea recta corresponde a la distribución marginal verdadera, denotada en el gráfico por F-Ver, y los puntos corresponden a los valores de la función de distribución marginal estimada. En las Figuras se observa en general el buen ajuste que tiene la función de distribución marginal estimada por el algoritmo cópula gráfico a la distribución marginal verdadera para todos los niveles de dependencia considerados $\tau = 0.2, 0.5, 0.8, 0.9$, mejorando el ajuste cuando el tamaño de muestra es mayor indicando que las estimaciones son buenas. En la Figura **A-10** se observa que para $n = 200$ en la cola inferior no hay un buen ajuste de la función de distribución marginal estimada por el estimador cópula gráfico a la función de distribución marginal verdadera. Sin embargo como se ha aclarado anteriormente la existencia de mayor variabilidad en las colas es un comportamiento de esperar en estos gráficos de probabilidad.

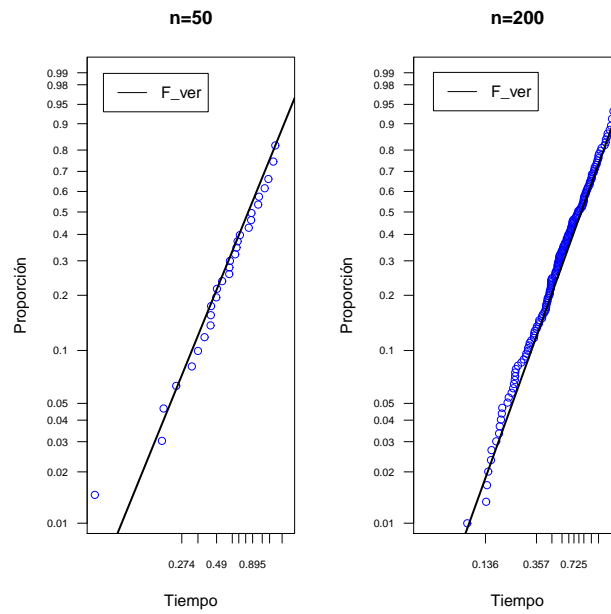


Figura A-9: Gráfico de probabilidad Weibull para la función de distribución marginal de Y , $\beta_2 = 2, \eta_2 = 1, \tau = 0.2$, para $n = 50$ y $n = 200$.

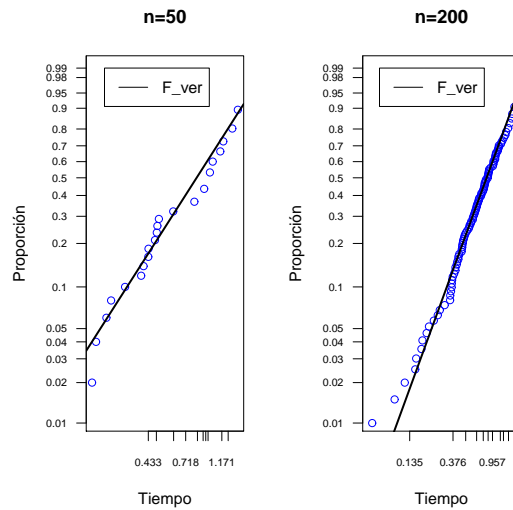


Figura A-10: Gráfico de probabilidad Weibull para la función de distribución marginal de Y , $\beta_2 = 2, \eta_2 = 1, \tau = 0.5$, para $n = 50$ y $n = 200$.

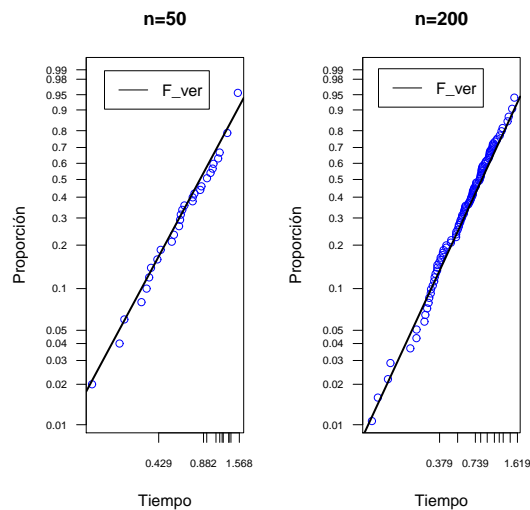


Figura A-11: Gráfico de probabilidad Weibull para la función de distribución marginal de Y , $\beta_2 = 2, \eta_2 = 1, \tau = 0.8$, para $n = 50$ y $n = 200$.

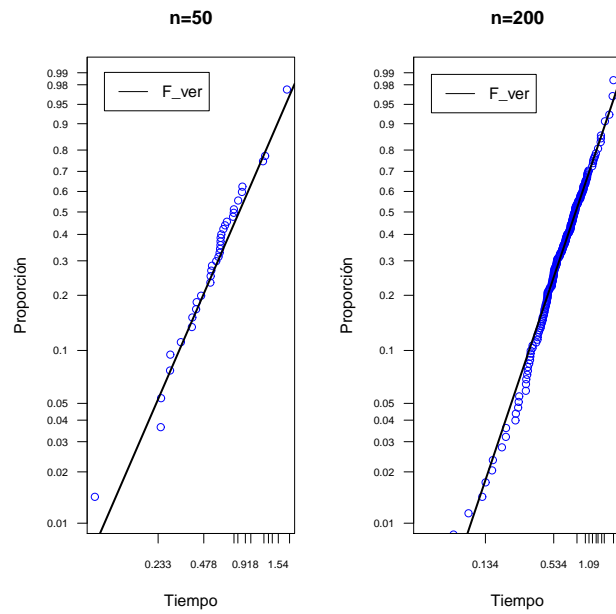


Figura A-12: Gráfico de probabilidad Weibull para la función de distribución marginal de Y , $\beta_2 = 2, \eta_2 = 1, \tau = 0.9$, para $n = 50$ y $n = 200$.

El error cuadrático medio para la función de distribución marginal de Y estimada en los percentiles t_p , con $p = 0.25, 0.5, 0.95$ se muestra en la Tabla **A-2**. En este caso, se observa al igual que los resultados obtenidos para la función de distribución marginal de X , que los valores del ECM son pequeños indicando que las estimaciones son buenas, la función de distribución marginal de Y estimada por el algoritmo cópula gráfico se aproxima muy bien a la función de distribución marginal verdadera para los dos tamaños de muestra considerados $n = 50$ y $n = 200$.

Tabla A-2: Error cuadrático medio para la función de distribución marginal estimada de Y en los percentiles t_p , con $p = 0.25, 0.50, 0.95$.

β_1	η_1	τ	$n = 50$			$n = 200$		
			$p = 0.25$	$p = 0.50$	$p = 0.95$	$p = 0.25$	$p = 0.5$	$p = 0.95$
2	1	0.2	0.0021	0.0065	0.0095	0.0005	0.0017	0.0017
2	1	0.5	0.0009	0.0051	0.0058	0.0002	0.0016	0.0014
2	1	0.8	0.0016	0.0040	0.0030	0.0003	0.0013	0.0006
2	1	0.9	0.0009	0.0049	0.0024	0.0002	0.0010	0.0004

A.2 Comparación entre la Función de Supervivencia Verdadera y la Estimada por el Algoritmo Cópula Gráfico

Se comparan los percentiles t_p , con $p = 0.05, 0.25, 0.50, 0.75, 0.95$, de la función de supervivencia conjunta de la Weibull bivariada verdadera definida en (A-1) denotada por $S(t_p)$, con la función de supervivencia conjunta estimada mediante el algoritmo cópula gráfico $\widehat{S}(t_p)_{copula}$.

A continuación se presenta un esquema general que permite realizar esta comparación vía simulación.

1. Para cada uno de los escenarios descritos en la subsección A.1 se calcula la función de supervivencia conjunta de la Weibull bivariada verdadera $S(t_p)$ en los percentiles t_p , con $p = 0.05, 0.25, 0.50, 0.75, 0.95$.
2. En los percentiles t_p se encuentra la función de supervivencia conjunta estimada mediante el algoritmo cópula gráfico. Esta función está definida como

$$\widehat{S}(t_p)_{copula} = \left\{ \left(\frac{1}{1 - \widehat{F}_1} \right)^{(\alpha-1)} + \left(\frac{1}{1 - \widehat{F}_2} \right)^{(\alpha-1)} - 1 \right\}^{-1/(\alpha-1)}. \quad (\text{A-2})$$

donde \widehat{F}_1 y \widehat{F}_2 son las estimaciones de las distribuciones marginales de X y Y , encontradas a partir del estimador cópula gráfico y α es el parámetro de dependencia de la cópula que mide el grado de asociación entre los tiempos de falla X y Y (Yáñez, Brango, Lopera, and Jaramillo, 2011). Este parámetro se relaciona con el coeficiente τ de Kendall, de la siguiente manera

$$\tau = \frac{1}{\alpha - 1}.$$

Para encontrar los valores de $\widehat{S}(t_p)_{copula}$, se calcula inicialmente el valor del tiempo mínimo correspondiente a cada cuantil con $p = 0.05, 0.25, 0.50, 0.75, 0.95$ y finalmente se encuentra el valor aproximado de la función de la supervivencia estimada en cada uno de los cuantiles.

Para realizar la comparación de la función de supervivencia estimada por el algoritmo cópula gráfico con la función de supervivencia verdadera, se generan $N = 1000$ muestras independientes para cada uno de los escenarios de simulación planteados.

Para realizar esta estimación se lleva a cabo la siguiente metodología

- Se generan muestras de tamaño $n = 50$ y $n = 200$ de tiempos bivariados Weibull, como se explicó en la subsección A.1 .
- En cada una de las muestras de tiempos bivariados Weibull obtenidos se encuentra el $T = \min \{T_1, T_2\}$.
- Se estiman las marginales \widehat{F}_1 y \widehat{F}_2 por medio del estimador cópula gráfico.
- Se estima la función de supervivencia mediante la ecuación (A-2) para cada uno de los casos de interés.

3. El chequeo procede de la siguiente manera

- a. Se realiza un gráfico de probabilidad Weibull para comparar la función de supervivencia conjunta de la Weibull bivariada verdadera y la función de supervivencia estimada mediante el algoritmo cópula gráfico, lo que permite tener una idea visual de cercanía entre estas dos funciones.
- b. Se calcula el error cuadrático medio (ECM) entre la función de supervivencia verdadera $S(t_p)$ y la función de supervivencia estimada $\widehat{S}(t_p)_{copula}$ en cada uno de los percentiles mencionados anteriormente.

$$ECM(\widehat{S}(t_p)_{copula}) = \frac{1}{N} \sum_{j=1}^N \left(\widehat{S}(t_p)_{copula} - S(t_p) \right)^2.$$

Las Figuras **A-13** - **A-16** comparan mediante gráficos de probabilidad Weibull la función de sobrevivencia conjunta Weibull bivariada verdadera y la función de sobrevivencia estimada mediante el algoritmo cópula gráfico para el tamaño de muestra $n = 50$. La línea recta corresponde a la función de sobrevivencia real $S(t_p)$, denotada en el gráfico por S-Ver y los puntos corresponden a la función de sobrevivencia estimada $\widehat{S}(t_p)_{copula}$. Estos gráficos ilustran en general la buena aproximación que tiene la función de sobrevivencia estimada a la función de sobrevivencia verdadera para este tamaño de muestra. Cabe resaltar que para todos los niveles de dependencia considerados existe una alta variabilidad en la cola inferior, lo cual es un comportamiento de esperar en estos gráficos de probabilidad.

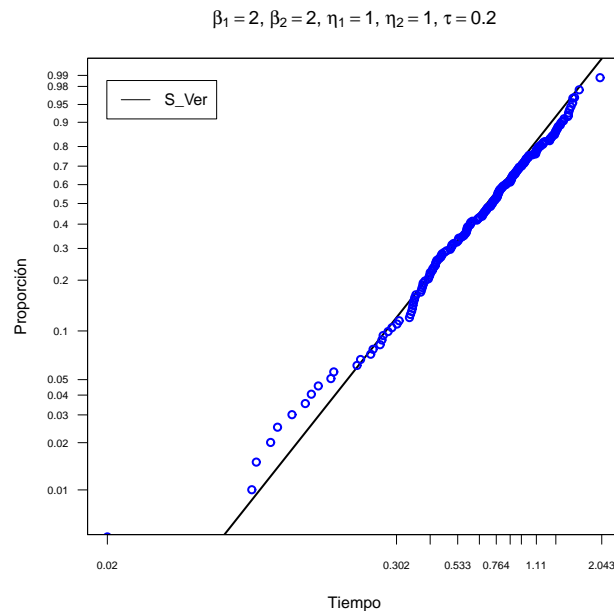


Figura A-13: Función de sobrevivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de sobrevivencia estimada $\widehat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.2$, para $n = 50$.

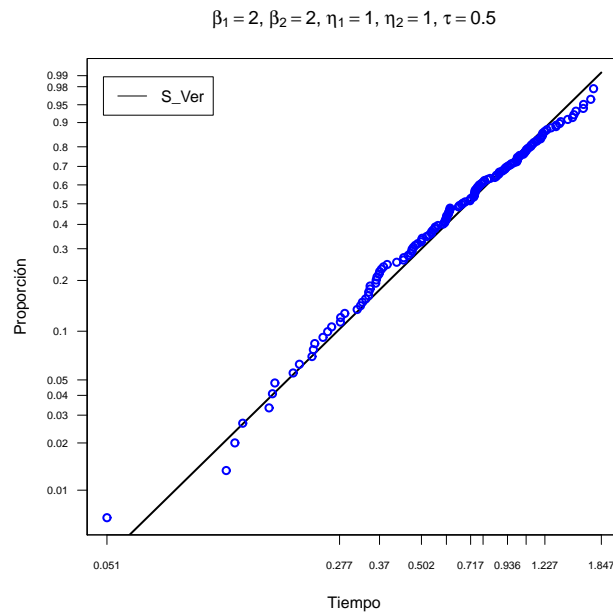


Figura A-14: Función de supervivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de supervivencia estimada $\hat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.5$, para $n = 50$.

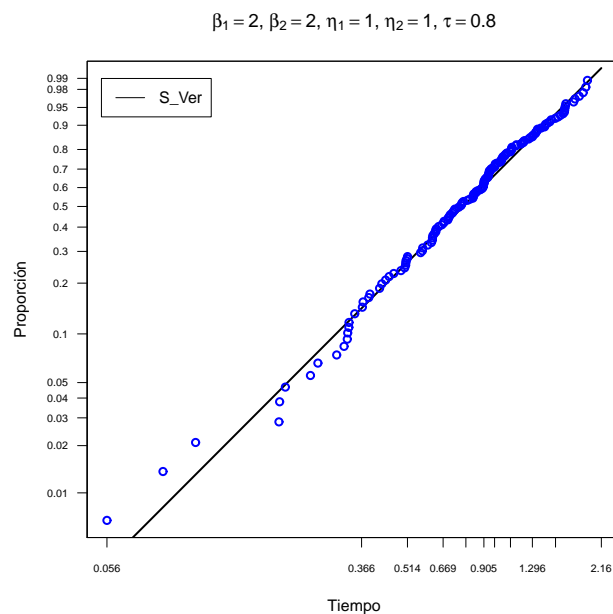


Figura A-15: Función de supervivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de supervivencia estimada $\hat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.8$, para $n = 50$.

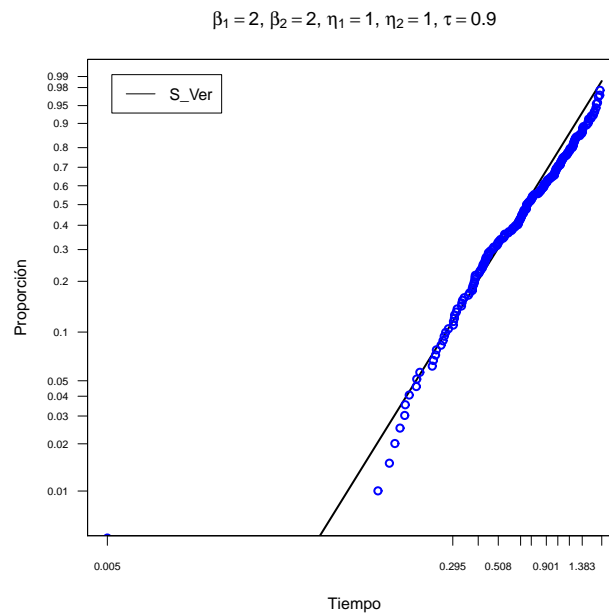


Figura A-16: Función de supervivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de supervivencia estimada $\hat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.9$, para $n = 50$.

Las Figuras **A-17** - **A-20** comparan mediante gráficos de probabilidad Weibull la función de supervivencia conjunta Weibull bivariada verdadera y la función de supervivencia estimada mediante el algoritmo cópula gráfico para el tamaño de muestra $n = 200$, observando que para este tamaño de muestra también la función de supervivencia conjunta estimada $\widehat{S}(t_p)_{copula}$ tiene una buena aproximación a la función de supervivencia conjunta real. Sin embargo, como se puede observar en la Figura **A-17**, correspondiente al nivel de dependencia $\tau = 0.2$, la función de supervivencia conjunta estimada por el algoritmo cópula gráfico no presenta una muy buena aproximación a la función de supervivencia conjunta verdadera, para los demás casos de estudio se observa un buen ajuste, resaltando la alta variabilidad en la cola inferior.

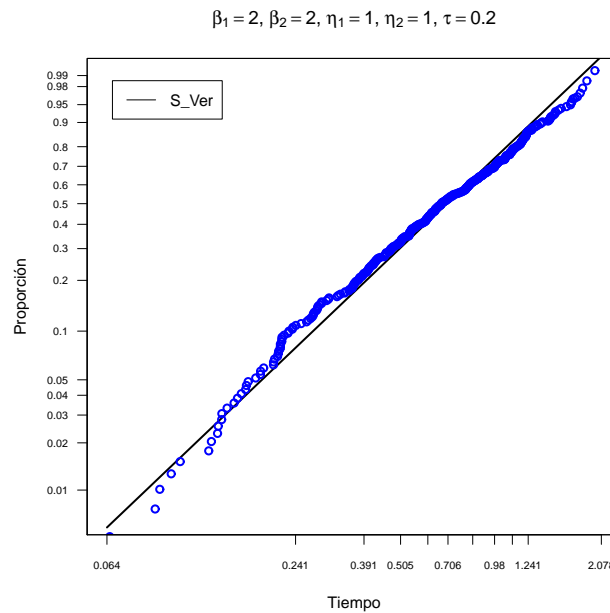


Figura A-17: Función de supervivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de supervivencia estimada $\widehat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.2$, para $n = 200$.

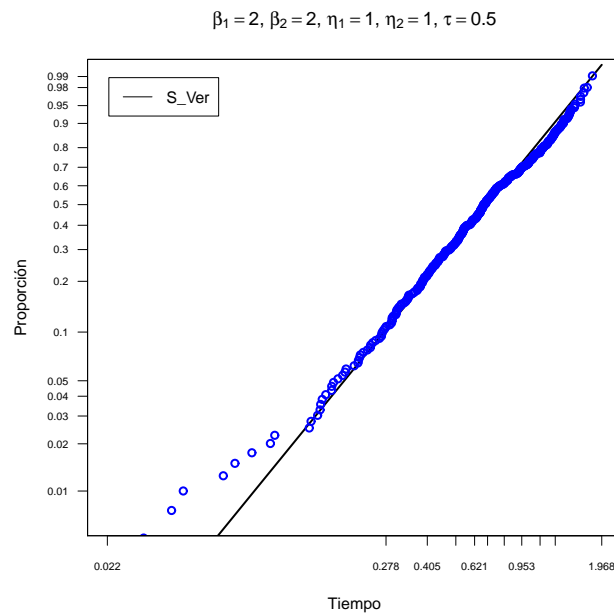


Figura A-18: Función de supervivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de supervivencia estimada $\hat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.5$, para $n = 200$.

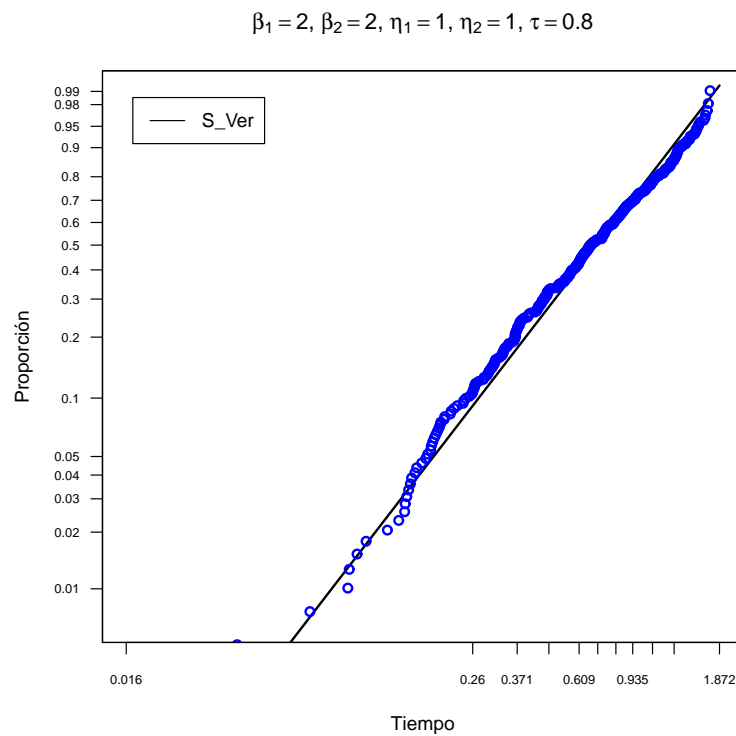


Figura A-19: Función de supervivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de supervivencia estimada $\hat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.8$, para $n = 200$.

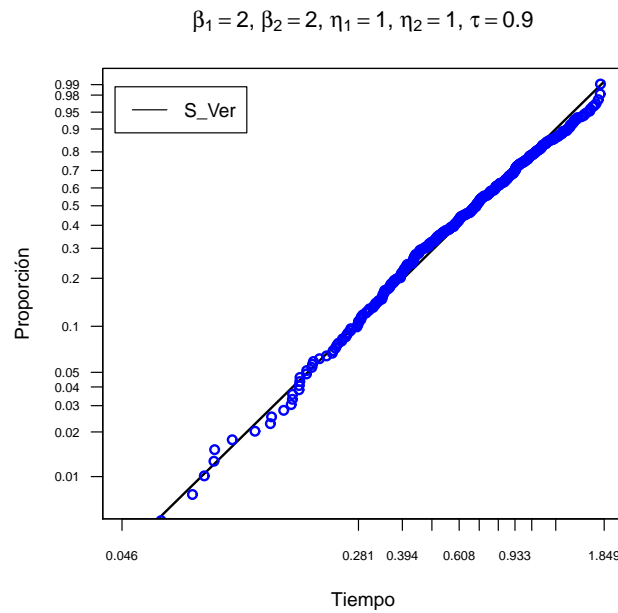


Figura A-20: Función de supervivencia conjunta Weibull bivariada verdadera $S(t_p)$ y función de supervivencia estimada $\hat{S}(t_p)_{copula}$, con $\beta_1 = 2, \beta_2 = 2, \eta_1 = 1, \eta_2 = 1$ y $\tau = 0.9$, para $n = 200$.

La Tabla **A-3** y **A-4** muestra el error cuadrático medio para la función de supervivencia conjunta de la Weibull bivariada estimada en cada uno de los percentiles t_p para los tamaños de muestra $n = 50$ y $n = 200$ respectivamente, encontrándose que el valor de la función de supervivencia estimada mediante el algoritmo cópula gráfico está muy cercano al valor de la función de supervivencia conjunta verdadera, lo cual indica que las estimaciones son buenas. Se observa que el error cuadrático medio para el tamaño de muestra $n = 200$ no es significativamente más pequeño al encontrado para el tamaño de muestra $n = 50$, aunque cabe resaltar que en algunos casos si resulta mucho menor.

Tabla A-3: Error cuadrático medio para la función de supervivencia conjunta de la Weibull bivariada estimada en los percentiles t_p , con $p = 0.05, 0.25, 0.50, 0.75, 0.95$, para $n = 50$.

β_1	β_2	η_1	η_2	τ	$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
2	2	1	1	0.2	0.019116	0.017026	0.008821	0.002379	0.016225
2	2	1	1	0.5	0.017444	0.016960	0.004629	0.001820	0.014120
2	2	1	1	0.8	0.015582	0.012939	0.001369	0.001487	0.004780
2	2	1	1	0.9	0.015405	0.011931	0.001191	0.001411	0.003591
2.5	2	1	1	0.2	0.024286	0.020919	0.004491	0.008910	0.018125
2.5	2	1	1	0.5	0.016149	0.014983	0.001197	0.002724	0.023373
2.5	2	1	1	0.8	0.014387	0.011803	0.001132	0.001857	0.007489
2.5	2	1	1	0.9	0.012955	0.008638	0.001016	0.001880	0.006339

Tabla A-4: Error cuadrático medio para la función de supervivencia conjunta de la Weibull bivariada estimada en los percentiles t_p , con $p = 0.05, 0.25, 0.50, 0.75, 0.95$, para $n = 200$.

β_1	β_2	η_1	η_2	τ	$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
2	2	1	1	0.2	0.014776	0.012260	0.001005	0.002053	0.010589
2	2	1	1	0.5	0.014115	0.012043	0.000629	0.000714	0.006236
2	2	1	1	0.8	0.013322	0.009034	0.000784	0.000465	0.004737
2	2	1	1	0.9	0.012265	0.008394	0.000642	0.000282	0.003927
2.5	2	1	1	0.2	0.013583	0.009868	0.013759	0.010460	0.008265
2.5	2	1	1	0.5	0.013247	0.007565	0.000779	0.007039	0.006420
2.5	2	1	1	0.8	0.012688	0.005060	0.000525	0.004534	0.002470
2.5	2	1	1	0.9	0.000352	0.000350	0.000306	0.000294	0.000736

El estudio de simulación presentado anteriormente, mostró que al realizar la comparación entre las marginales verdaderas y las estimadas por el algoritmo cópula gráfico, tanto las distribuciones marginales de X y Y para los tamaños de muestra $n = 50$ y $n = 200$ tienen un buen ajuste a una distribución Weibull. También se encontró que la función de supervivencia estimada por el algoritmo cópula gráfico tiene una buena aproximación a la función de supervivencia verdadera. El error cuadrático medio calculado para la supervivencia conjunta de la Weibull bivariada estimada en los percentiles t_p con $p = 0.05, 0.25, 0.50, 0.75, 0.95$, resultó ser pequeño indicando que las estimaciones son buenas para los dos tamaños de muestra considerados. Lo anterior afirma que el estimador cópula gráfico estima muy bien las funciones de distribución marginales y la función de supervivencia conjunta.

B Anexo: Algoritmo y Programa en R del Método de Combinación de Riesgos

Algoritmo del método de combinación de riesgos

1. Calcule la función de supervivencia conjunta $S(t)$, las funciones de subdistribución para cada uno de los riesgos (CIC) y defina la cópula y el parámetro de la cópula.
2. Establezca un valor inicial de $\hat{F}_1^1(t_1)$ y encuentre $\hat{F}_{23}^1(t_1)$.
3. Si el valor $\hat{F}_1^i(t_i)$ es suficientemente cercano al valor de $\hat{F}_1^{i-1}(t_{i-1})$, vaya al paso 4. En caso contrario encuentre el nuevo valor $\hat{F}_1^i(t_i)$ como el punto medio del intervalo $(a_i, \hat{F}_1^{(i)}(t_1))$ o $(\hat{F}_1^{(i)}(t_1), b_i)$, donde

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

y finalmente encuentre $\hat{F}_{23}^i(t_i)$.

4. Una vez encontrado $\hat{F}_1^i(t_1)$, repita los pasos 2 y 3 para encontrar $\hat{F}_1^j(t_j)$ $j = 2, \dots, m$.
5. Repita los pasos (2 – 4) para encontrar $\hat{F}_2(t)$ y $\hat{F}_3(t)$.

Código R B.1

```
#Esta es la implementación en el software estadístico R del método de
#combinación de riesgos sugerido en el Artículo "A Copula Model for
#Dependent Competing Risks" por Simon Lo and Ralf Wilke.
MCopGE = function(type, alpha, S, Q, kf)
library(stats)
library(utils)
library(base)
library(pracma)
library(mvtnorm)
```

```
library(mnormt)
library(scatterplot3d)
library(apspline)
library(sn)
library(ADGofTest)
library(stabledist)
library(copula)
library(splines)
library(survival)
library(cmprsk)
library(mstate)
#Declaración de constantes
BETA1 <- 2.4
BETA2 <- 2.6
BETA3 <- 2.8
ETA<- 1
ALPHA <- 20
TAO<- 0.8
lockBinding("BETA1", globalenv())
lockBinding("BETA2", globalenv())
lockBinding("BETA3", globalenv())
lockBinding("ETA", globalenv())
lockBinding("ALPHA", globalenv())
lockBinding("TAO", globalenv())
{
  type = tolower(type)
  if (type == 'frank')
  {
    if ((missing(type)) || (missing(alpha)) || (missing(S)) || (missing(Q)) ||
    {
      stop("Requires exactly five input arguments: MCopGE(TYPE, tau, S, Q, kf)
    }

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

    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 < -200)
    {
      stop("The alpha value must be greater than -200")
    }
  }
}
```

```

    }
    else if (alpha > 35)
    {
      stop("The alpha value must be less than 35")
    }
    if (is.numeric(kf) != 1)
    {
      stop("kf must be a scalar: number of marginals to be estimated")
    }
    if (kf > k)
    {
      stop("kf must be less than k: number of marginals to be estimated should
          be smaller or equal to number of CIC")
    }
    if (k < 1)
    {
      stop("matrix of CIC needs to have dimensions greater than 1")
    }
    if (ns != nq)
    {
      stop("number of rows in the matrix of CIC and Survivals do not agree")
    }
  }

F = matrix(0, ns, kf)
  for (nt in 1:kf)
  {
    F[,nt] = suppressWarnings(frankest3redo(alpha, S, Q[,nt]))
  }
  return (F)
}
else
{
  stop("Only frank estimator is available")
}
}

frankest3redo = function(a, S, Q1)
{
#Dada la sobrevivencia conjunta, las funciones de subdistribución (CIC), la
cópula y el parámetro de la cópula, las funciones de distribución marginales
son identificadas.
  tol = 1e-5
  floor = 0
  F1 = matrix(0, length(S), 1) # estimador de H1 (H1 función de distribución)
  F2 = matrix(0, length(S), 1) # estimador de H2
  I = matrix(0, length(S), 1) # valor integrado en cada t
  lower = floor
  upper = 1

```

```

E = matrix(0, 100, 1)
i = 1
w = 1
# j = índice del tiempo,
# i = índice de cada iteración

h1=S[1] # valor inicial de H1 en t1
h2 = solve43(h1, S[1], a) #función solve43 definida Ecuación 2 al final del programa
while (abs(w) > tol)
{
#función de densidad de la cópula Frank
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)
width = I[1] - Q1[1]
#el primer valor de h1 es almacenado en 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)

if ((abs(w) > tol) && (width > 0)) # esto significa h1 es muy grande,
use un valor estimado menor.
{
h1 = (h1 + lower)*0.5
h2 = solve43(h1, S[1], a)
}
else if ((abs(w) > tol) && (width < 0)) # esto significa que h1 es muy pequeño.

```

```

{
  h1 = (h1 + upper)*0.5
  h2 = solve43(h1, S[1], a)
}
if (i > 1)
{
  w = E[i]-E[i-1] # condición para finalizar la iteración,
  si el valor estimado en la iteración i está suficientemente
  cercano al valor en la iteración i-1.
}
i = i + 1
}
F1[1]=h1 # valor estimado de H1 en t1
F2[1]=h2 # valor estimado de H2 en t1
floor= h1 # h1 en t1 se convierte en el valor inferior (floor) de h1 en t2
h1c=h1
h2c=as.numeric(h2)

for (j in 2:length(S))
{
  i = 1
  lower = floor
  upper = 1
  E = matrix(0,100,1)
  w = 1 # inicializar la estimación
  h1 = S[j]
  h2 = solve43(h1, S[j], a)
  while (abs(w) > tol)
  {
    # NOTA: el valor integrado I(t2) sólo cubre el área entre t1 y t2,
    # por lo tanto se necesita adicionar el área entre t1 and
    # t0, I(t1), el valor estimado de la integración en el último periodo.
    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]
    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

    if (i > length(E))
    {
        stop("not converging")
    }
}
F1[j] = h1
F2[j] = h2
floor = h1
h1c = h1
h2c = h2
I[j] = I[j] + I [j - 1]
```

```

}
return (F1)
}

#Ecuación 1
e43 = function(h1,h2,a)
{
  #sobrevivencia dado u,v, y la cópula
  s = 1-h1-h2-1/a*log(1+(exp(-a*h1)-1)*(exp(-a*h2)-1)/(exp(-a)-1))
  return (s)
}

#Ecuación 2
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)
}

#Cálculo de la función de sobrevivencia conjunta y las funciones
de subdistribución.
calculate_CIC_S = function(n)
{
  CIC = matrix(0, 1, 1)
  flag = TRUE;
  datos1 = 0
  datos = 0
  y = NA
  #Calculo de la CIC, riesgo 1, riesgo 2 y 3 generados por
  una distribución Weibull, con la copula frank
  while (flag == TRUE)
  {
    while (length(which(is.na(y))) != 0)
    {
      myCop.frank<- archmCopula(family= "frank",dim=3, param=ALPHA)
      myMvd<-mvdc(copula= myCop.frank,margins=c("weibull","weibull","weibull"),
      paramMargins=list(list(shape=BETA1,scale=1),
      list(shape=BETA2,scale=1),list(shape=BETA3,scale=1)))
    }
  }
}

```

```

        y<-rmvdc(myMvd,n)
    }
    minimo<-apply(y,1,min) #obtención del mínimo
    indicadora<-apply(y,1,which.min) #obtención de la indicadora
    datos1=data.frame(minimo,indicadora)
    datos=datos1[order(datos1[,1]),]
    conteo=as.matrix(table(datos1$indicadora))
    if (nrow(conteo) > 2)
    {
        flag = FALSE
    }
    y = NA
}
fit=Cuminc(datos[,1],datos[,2])
CIC= cbind(fit$CI.1[-1],fit$CI.2[-1],fit$CI.3[-1]) #Calculo de las funciones
de subdistribución
S=fit$Surv[-1] #Calculo de la función de sobrevivencia
conjunta del tiempo mínimo t.
value = data.frame(CIC = CIC, S = S, minT=datos[,1],ind=datos[,2])
return (value)
}

#n Tamaño de muestra
#parámetro de la cópula
#número de marginales a ser estimadas
principal = function (n, alpha, 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: number of marginals to be estimated")
    }
    resultado = suppressWarnings(calculate_CIC_S(n))
}

```

```
CIC = resultado[,1:3]
S = resultado[,4]
minT=resultado[,5]
ind=resultado[,6]
mcpge = MCopGE('frank', alpha, S, CIC, kf)
value = data.frame(mcpge = mcpge)
return(value)
}
MDF=principal(50,ALPHA,3)
```

Bibliografía

- Brago, H. A. (2011). Estimación de la función de sobrevivencia con datos de riesgos en competencia bivariados dependientes. Master's thesis, Universidad Nacional de Colombia.
- Carriere, J. F. (1995). Removing cancer when it is correlated with other causes of death. *Biometrical* 37(3), 339–350.
- Drouet Mari, D. and S. Kotz (2001). *Correlation and Dependence*. Imperial College Press.
- Escarela, G. and J. Carriere (2003). Fitting competing risks with an assumed copula. *Statistical Methods in Medical Research* 12, 33–349.
- Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika* 74(3), 549–555.
- Genest, C., K. Ghousi, and L. Rivest (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* 82(3), 543–552.
- Genest, C. and J. Mackay (1986). The joy of copulas: Bivariate distributions with uniform marginals. *The American Statistician* 40(4), 280–283.
- Georges, P., A. Lamy, E. Nicolas, G. Quibel, and T. Roncalli (2001). Multivariate survival modelling: a unified approach with copulas. Technical report, Groupe de Recherche Opérationnelle Crédit Lyonnais.
- Johnson, R. A. and D. W. Wichern (2007). *Applied Multivariate Statistical Analysis* (Sixth edition ed.). New Jersey: Pearson Prentice Hall.
- Kotz, S., N. Balakrishnan, and N. L. Johnson (2000). *Continuous Multivariate Distributions*. John Wiley & Sons.
- Lee, L. (1979). Multivariate distributions having weibull properties. *Journal of Multivariate Analysis* 9, 267–277.
- Liu, X. (2010). Copulas of bivariate rayleigh and lognormal distributions. *Electronics Letters* 46(25), 1–2.
- Lo, S. and R. A. Wilke (2010). A copula model for dependent competing risks. *Journal of the Royal Statistical Society* 59(2), 359–376.

- Lu, J. and G. Bhattacharyya (1990). Some new constructions of bivariate weibull models. *Annals of the Institute of Statistical Mathematics* 42(3), 543–559.
- Meeker, W., L. A. Escobar, and Y. Hong (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 edition ed.). New York: Springer.
- Pintilie, M. (2006). *Competing Risks: A Practical Perspective*. John Wiley & Sons.
- Rivest, L. and M. Wells (2001). A martingale approach to the copula graphic estimator for the survival function under dependent censoring. *J. Multiva. Anal* 79, 138–155.
- Schweizer, B. and A. Sklar (1983). *Probabilistic Metric Spaces*. New York: North-Holland.
- Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proc. Natl. Acad. Sci. USA* 72(1), 20–22.
- Yan, J. (2007). Enjoy the joy of copulas. *Journal of Statistical Software* 21(2), 1–21.
- Yanagimoto, T. and M. Okamoto (1969). Partial orderings of permutations and monotonicity of a rank correlation statistic. *Annals of the Institute of Statistical Mathematics* 21, 489–506.
- Yáñez, S., H. A. Brango, C. M. Lopera, and M. C. Jaramillo (2011). Comparación entre riesgos competitivos vía estimador cópula gráfico. *Revista Colombiana de Estadística* 34(2), 231–248.
- Yáñez, S., E. Manotas, C. M. Lopera, and M. C. Jaramillo (2008). Estudio del efecto de la dependencia en la estimación de la confiabilidad de un sistema con dos modos de falla concurrentes. *Dyna* 75(154), 29–38.
- Zheng, M. and J. P. Klein (1995). Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika* 82(1), 12–38.