

Análisis bivariado de confiabilidad basado en cópulas

Reliability Bivariate Analysis Based on Copulas

ENRIQUE RAÚL VILLA-DIHARCE^{1,a}, PEDRO ENRIQUE MONJARDIN^{1,b}

¹CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS A. C., GUANAJUATO, MÉXICO

Resumen

En la industria se requieren estudios de confiabilidad para asegurar el nivel de calidad de los productos que se fabrican. Algunas veces tales productos tienen varios modos de falla que deben considerarse en el análisis de confiabilidad. Este análisis se complica cuando se incluye el patrón de dependencia de los modos de falla. En este artículo se muestra el análisis de confiabilidad de componentes que presentan dos modos de falla dependientes, expresando el patrón de dependencia de ambos modos de falla por medio de una función de cópula. Esta representación es adecuada, debido a que las distribuciones de los tiempos de falla son diferentes, Lognormal y Weibull. Se presenta un ejemplo donde se ilustra el análisis de confiabilidad de componentes con dos modos de falla.

Palabras clave: cópula, distribución bivariada, estimación en dos etapas, función de confiabilidad, función de verosimilitud.

Abstract

In industry, reliability studies are done to assure the quality level of products manufactured. Sometimes these products have several failure modes that must be considered in reliability analysis. This analysis is complicated when the dependence pattern is included. In this article a reliability analysis of components that have two dependent failure modes is proposed. The dependence pattern is expressed through a copula function, this representation is appropriate, because the marginal distributions of failure times are from different families, Lognormal and Weibull. We provide an example which illustrates the reliability analysis of components with two failure modes.

Key words: Bivariate distribution, Copula, Likelihood function, Reliability function, Two-stage estimation.

^aProfesor e investigador. E-mail: villadi@cimat.mx

^bEstudiante de doctorado. E-mail: pemo@cimat.mx

1. Introducción

En la industria existe la preocupación de mejorar los procesos de producción y aumentar la calidad de sus productos, lo que se busca es disminuir costos e incrementar la confiabilidad de los equipos o componentes que producen. Para el segundo objetivo se realizan estudios de confiabilidad para verificar si el producto satisface los requerimientos del cliente. En este trabajo comentamos una forma de analizar la confiabilidad de un producto que tiene dos modos de falla. Usualmente los equipos están formados por varios componentes, y cuando falla uno de los componentes, falla el equipo (o el desempeño del mismo disminuye), a esto se le puede llamar un sistema con varios modos de falla, de hecho un componente también puede tener varios modos de falla. Como los componentes de un sistema o producto operan bajo las mismas condiciones, se espera que sus tiempos a la falla sean dependientes. Para el análisis de este tipo de información se pueden utilizar técnicas de riesgos en competencia. En estas técnicas, los diferentes modos de falla están en competencia, esto es, solo se observa el primero que ocurre, ya que una vez que la unidad falla por alguno de los modos de falla, deja de ser observada. El tiempo a la falla de una unidad que tiene varios modos de falla, será el mínimo de los tiempos de falla potenciales asociados a cada modo de falla. En algunos estudios de confiabilidad, sobre todo en los que se realizan en laboratorio, es posible corregir los modos de falla que se van presentando, esto permite continuar con el estudio hasta que se observen todos los modos de falla. En este caso, los datos son multivariados y se pueden usar modelos multivariados para su análisis, los datos que analizamos en este trabajo son de este tipo.

En este artículo se presenta un estudio de confiabilidad de encendedores de horno de estufa a gas, que tienen dos modos de falla, que se describen posteriormente en la sección 4. En este estudio, se observan ambos modos de falla y se registran sus tiempos representados por las variables aleatorias T_1 y T_2 , respectivamente. Como la falla de un componente no censura la ocurrencia de una falla en el otro componente, los datos en este estudio son bivariados.

El objetivo en este estudio consiste en conocer diferentes puntos de interés como son: el patrón de dependencia de los dos tiempos a la falla, la confiabilidad en dos tiempos de referencia importantes, uno y diez años, así como los cuantiles para diferentes probabilidades.

Las distribuciones identificadas para ambas variables aleatorias, son de diferentes familias: Weibull y Lognormal. Esto nos lleva a usar el enfoque de cópulas para modelar la dependencia entre ambas variables. Las cópulas separan la estructura de dependencia de los efectos de las marginales, por esta razón, es posible considerar marginales de familias de diferentes distribuciones, como ocurre en el presente trabajo.

En la sección 2, se discute la relación entre función de supervivencia y cópulas, también se presentan algunas de las cópulas arquimedeanas comunes. Se presenta además una técnica para identificar una cópula que modele el patrón de dependencia de los datos. En la sección 3, se describe la manera de hacer inferencia, cuando tenemos funciones de supervivencia basadas en cópulas. En la sección 4, se

discuten los detalles del análisis de los datos de tiempos de falla bivariados, para los dos modos de falla de los encendedores de horno. En la sección 5 se presentan las conclusiones del presente trabajo.

2. Función de confiabilidad bivariada y cópulas

El objetivo de un estudio de confiabilidad es estimar cantidades de interés como cuantiles para diferentes probabilidades y niveles de confiabilidad en diferentes tiempos. Con frecuencia, en la industria la confiabilidad bivariada se estima asumiendo independencia entre los tiempos a la falla T_1 y T_2 de los modos de falla. Por supuesto que esto no es del todo correcto, ya que al trabajar bajo las mismas condiciones, los modos de falla generalmente son dependientes. Así, es necesario considerar modelos que tomen en cuenta la estructura de dependencia que pueden tener los datos. Además de esto, la censura puede estar presente. De manera más precisa, si hay censura por la derecha y si C_1 y C_2 son las variables de censura para T_1 y T_2 , respectivamente, los datos que se observan son parejas (X, Y) e indicadores de censura (δ_1, δ_2) , con $X = \min(T_1, C_1)$, $Y = \min(T_2, C_2)$, $\delta_1 = \mathbb{1}_{[X=T_1]}$ y $\delta_2 = \mathbb{1}_{[Y=T_2]}$, donde $\mathbb{1}_A$ es la función indicadora del evento A .

Si las distribuciones marginales de T_1 y T_2 son continuas, podemos unir las con una función para formar una función de confiabilidad bivariada como sigue (*e. g.* Wang & Wells 2000).

$$S(t_1, t_2) = C_\theta[S_1(t_1), S_2(t_2)] \quad (1)$$

donde $S_j(\cdot)$, $j = 1, 2$, son las funciones de confiabilidad marginales de T_1 y T_2 , $C_\theta(\cdot, \cdot)$ es una función que une a las funciones de confiabilidad marginales, llamada función cópula, y $\theta \in \mathbb{R}$ es un parámetro de asociación desconocido. Algunos ejemplos de estas funciones de confiabilidad bivariada se encuentran en Clayton (1978), Frank (1979), Joe (1993) y Hougaard (1986). Note que la función de cópula $C_\theta(\cdot, \cdot)$ es una función de supervivencia en $[0, 1] \times [0, 1]$. En la literatura, existen una gran cantidad de trabajos relacionados con las propiedades de las funciones cópulas y el modelado de variables aleatorias usando éstas (*e.g.* Roger (2006), Joe (1997), Escarela & Hernández (2009), Jaworski, Durante, Hädel & Rychlik (2010)). Una característica especial de las funciones cópula es que la estructura de dependencia es separada de los efectos de las marginales, así que la relación de dependencia se puede estudiar sin especificar las distribuciones marginales.

La función de densidad conjunta correspondiente a la función de confiabilidad dada en la ecuación (1) está dada por

$$f(t_1, t_2) = f_1(t_1)f_2(t_2) \times c_\theta[S_1(t_1), S_2(t_2)]$$

donde $f_j(z) = -\frac{d}{dz}S_j(z)$, $j = 1, 2$, son las densidades de los tiempos a la falla y $c_\theta(u, v) = \frac{\partial^2}{\partial u \partial v}C_\theta(u, v)$.

El parámetro θ se puede ver como un parámetro de asociación global, y está relacionado con la τ de Kendall a través de la ecuación (ver Escarela & Hernández 2009)

$$\tau = 4 \int_0^1 \int_0^1 C_\theta(u, v) dC_\theta(u, v) - 1 \equiv \Lambda(\theta) \quad (2)$$

Dado un nivel de asociación medido por τ o θ , $C_\theta(u, v)$ determina el grado de dependencia local en $(u, v) \in [0, 1] \times [0, 1]$. Cuando la asociación se aproxima a los niveles extremos, todos los modelos toman la misma forma: bajo independencia ($\tau = 0$), $C_\theta(u, v) = uv \equiv C^\perp(u, v)$ y bajo la máxima dependencia positiva ($\tau = 1$), $C_\theta(u, v) = \min(u, v) \equiv C^+(u, v)$.

Existe una clase de cópulas que engloba a la mayoría de las cópulas bivariadas más comunes, su característica principal es que es indexada por una función univariada, por lo que sus propiedades matemáticas y computacionales son fáciles, esta es la llamada clase de cópulas arquimedeanas. Una función cópula es arquimedea si puede expresarse como

$$C_\theta(u, v) = \varphi_\theta^{-1} [\varphi_\theta(u) + \varphi_\theta(v)], \quad (u, v) \in [0, 1] \times [0, 1]$$

donde $\varphi_\theta(\cdot)$ es una función convexa definida en $[0, 1]$ tal que $\varphi_\theta(1) = 0$ y $\varphi_\theta^{-1}(\cdot)$ denota la pseudoinversa de $\varphi_\theta(\cdot)$. A φ_θ se le llama la función generadora de la cópula.

Algunas de las familias de cópulas más comunes que se pueden expresar en una representación de cópulas arquimedeanas se muestran en la tabla 1.

TABLA 1: Ejemplos de cópulas arquimedeanas.

Familia	$C_\theta(u, v)$	$\varphi_\theta(t)$
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$	$\frac{1}{\theta}(t^{-\theta} - 1)$
Frank	$-\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right)$	$-\ln \left(\frac{e^{-\theta t} - 1}{e^{-\theta} - 1} \right)$
Gumbel	$\exp \left\{ - [(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta} \right\}$	$(-\ln t)^\theta$
Joe	$1 - [(1-u)^\theta + (1-v)^\theta - (1-u)^\theta(1-v)^\theta]$	$-\ln[1 - (1-t)^\theta]$

Una función de confiabilidad bivariada modelada con una función cópula arquimedea toma la forma

$$\mathbf{S}(t_1, t_2) = \varphi_\theta^{-1} \{ \varphi_\theta[S_1(t_1)] + \varphi_\theta[S_2(t_2)] \}$$

Para poder usar el modelo dado en (1) es necesario elegir una cópula. Genest & Riverst (1993) proponen un método no-paramétrico que sin tomar en cuenta las marginales, estima una función que identifica a una cópula arquimedea. Ellos prueban, cuando no hay datos censurados, que la función generadora $\varphi_\theta(\cdot)$ de una cópula arquimedea se puede identificar usando la función $K(z) = \Pr[\mathbf{S}(T_1, T_2) \leq z]$ a través de la ecuación diferencial

$$\lambda(z) = z - K(z) = \frac{\varphi_\theta(z)}{\varphi_\theta'(z)}$$

donde $\varphi'_\theta(z) = \frac{d}{dz}\varphi_\theta(z)$. Este método se usa en Arcila, Jaramillo & Lopera (2009) para analizar datos de seguros.

Esta propuesta no considera la censura en los datos, lo cual se tiene con gran frecuencia en estudios de confiabilidad. Wang & Wells (2000) proponen una técnica similar a la de Genest & Riverst (1993), pero para el caso en que hay datos censurados por la derecha, el método es el siguiente: Sean $x_{(1)} \leq \dots \leq x_{(n)}$ y $y_{(1)} \leq \dots \leq y_{(n)}$ las observaciones ordenadas de los datos $\{(X_i, Y_i) : i = 1, \dots, n\}$.

1. Calcular el estimador $\tilde{K}(z)$ de $K(z) = 1 - \Pr[\mathbf{S}(T_1, T_2) > z]$ como

$$\tilde{K}(z) = 1 - \sum_{i=1}^n \sum_{j=1}^n \mathbb{1}_{[\tilde{\mathbf{S}}(x_{(i)}, y_{(j)}) > z]} \tilde{\mathbf{S}}(\Delta x_{(i)}, \Delta y_{(j)})$$

donde $\tilde{\mathbf{S}}$ es un estimador no-paramétrico de \mathbf{S} y

$$\begin{aligned} \tilde{\mathbf{S}}(\Delta x_{(i)}, \Delta y_{(j)}) &= \tilde{\mathbf{S}}(x_{(i)}, y_{(j)}) - \tilde{\mathbf{S}}(x_{(i-1)}, y_{(j)}) \\ &\quad - \tilde{\mathbf{S}}(x_{(i)}, y_{(j-1)}) + \tilde{\mathbf{S}}(x_{(i-1)}, y_{(j-1)}) \end{aligned}$$

es la masa estimada en $(x_{(i)}, y_{(j)})$. Existen varias propuestas de estimadores no paramétricos de $\mathbf{S}(t_1, t_2)$ cuando los datos son censurados, por ejemplo Dabrowska (1988), Lin & Ying (1993), Prentice & Cai (1992), Wang & Wells (1997) y Batún-Cutz (2005). Luego calcular $\tilde{\lambda}(z) = z - \tilde{K}(z)$ y graficar $\tilde{\lambda}(z_i)$ para $z_i = \tilde{\mathbf{S}}(x_i, y_i)$, $i = 1, \dots, n$.

2. Sean $C_{\theta_l}(\cdot, \cdot)$, $l = 1, \dots, m$ las m funciones cópulas bajo consideración. Obtener un valor θ_l , digamos $\tilde{\theta}_l$, tal que $\hat{\tau}_0 = \Lambda(\tilde{\theta}_l)$ donde $\hat{\tau}_0$ es un estimador de τ . Si $\tau = \Lambda(\theta)$ es una función uno a uno, $\tilde{\theta}_l = \Lambda^{-1}(\hat{\tau}_0)$. Podemos usar el estimador de τ propuesto por Wang y Wells,

$$\hat{\tau}_0 = 4 \sum_{i=1}^n \sum_{j=1}^n \tilde{\mathbf{S}}(x_{(i)}, y_{(j)}) \tilde{\mathbf{S}}(\Delta x_{(i)}, \Delta y_{(j)}) - 1 \tag{3}$$

Si $x_{(n)}$ y $y_{(n)}$ son no censuradas, $\hat{\tau}_0$ es un estimador consistente de τ . Para cada $\tilde{\theta}_l$, obtener el estimador paramétrico $\lambda_{\tilde{\theta}_l}(z)$.

3. Comparar la gráfica de $(z_i, \tilde{\lambda}(z_i))$ con las gráficas teóricas $\lambda_{\tilde{\theta}_l}(z)$ para los modelos bajo consideración. Elegir el modelo que mejor ajuste los puntos $(z_i, \tilde{\lambda}(z_i))$, por ejemplo el que tenga menor suma de cuadrados

$$\sum_{i=1}^n \left[\tilde{\lambda}(z_i) - \lambda_{\tilde{\theta}_l}(z_i) \right]^2$$

3. Inferencia

Lo primero que hay que hacer para ajustar un modelo basado en cópulas es elegir la función cópula que usaremos en el modelo (1); para esto se puede usar

el método propuesto por Wang y Wells, descrito anteriormente. El siguiente paso es determinar las distribuciones marginales apropiadas, una buena referencia para el ajuste de modelos paramétricos univariadas es el libro de Meeker & Escobar (1998).

Para estimar los parámetros θ de la confiabilidad bivariada \mathbf{S} usaremos el método de máxima verosimilitud. Con el modelo basado en cópulas

$$\mathbf{S}(t_1, t_2; \theta) = C_\theta \left(S_1(t_1; \theta^{(1)}), S_2(t_2; \theta^{(2)}) \right)$$

donde $\theta = (\theta^{(1)}, \theta^{(2)}, \theta)$, siendo, $\theta^{(1)}$ y $\theta^{(2)}$ los parámetros correspondientes a las confiabilidades marginales, y θ es el parámetro de la función cópula.

Sea $\mathbf{d} = \{(x_i, \delta_{i1}, y_i, \delta_{i2}) : i = 1, \dots, n\}$ el conjunto de datos. Para especificar la función de verosimilitud es necesario distinguir entre las observaciones que son tiempos de falla y las que son censuradas.

Si ambas observaciones (x_i, y_i) son tiempos de falla ($\delta_{i1} = 1, \delta_{i2} = 1$), la contribución de esta pareja a la función de verosimilitud es

$$c_\theta \left(S_1(x_i; \theta^{(1)}), S_2(y_i; \theta^{(2)}) \right) f_1 \left(x_i; \theta^{(1)} \right) f_2 \left(y_i; \theta^{(2)} \right)$$

donde $f_j(z; \theta^{(j)}) = -\frac{\partial}{\partial z} S_j(z; \theta^{(j)})$, $j = 1, 2$, son las densidades de los tiempos de falla y $c_\theta(u, v) = \frac{\partial^2}{\partial u \partial v} C_\theta(u, v)$.

Observar que x_i es tiempo de falla y y_i es censurada ($\delta_{i1} = 1, \delta_{i2} = 0$), es equivalente a observar el evento $\{x_i < T_1 \leq x_i + h, T_2 > y_i\}$, para h pequeña. Pero

$$\begin{aligned} \Pr(x_i < T_1 \leq x_i + h, T_2 > y_i) &= \mathbf{S}(x_i, y_i; \theta) - \mathbf{S}(x_i + h, y_i; \theta) \\ &\approx h \cdot \left[-\frac{\partial}{\partial x_i} \mathbf{S}(x_i, y_i; \theta) \right] \end{aligned}$$

y la contribución a la verosimilitud por esta pareja es

$$-\frac{\partial}{\partial x_i} \mathbf{S}(x_i, y_i; \theta) = \frac{\partial}{\partial S_1(x_i; \theta^{(1)})} C_\theta \left(S_1(x_i; \theta^{(1)}), S_2(y_i; \theta^{(2)}) \right) f_1(x_i; \theta^{(1)})$$

De la misma forma, si x_i es una censura y y_i es tiempo de falla, la contribución de esta pareja a la verosimilitud es

$$-\frac{\partial}{\partial y} \mathbf{S}(x_i, y_i; \theta) = \frac{\partial}{\partial S_2(y_i; \theta^{(2)})} C_\theta \left(S_1(x_i; \theta^{(1)}), S_2(y_i; \theta^{(2)}) \right) f_2(y_i; \theta^{(2)})$$

Por último, si x_i y y_i son censuras, su contribución a la verosimilitud es

$$\Pr(T_1 > x_i, T_2 > y_i) = \mathbf{S}(t_1, t_2; \theta) = C_\theta \left(S_1(x_i; \theta^{(1)}), S_2(y_i; \theta^{(2)}) \right)$$

Por lo tanto, la función de verosimilitud es

$$\begin{aligned}
 L(\boldsymbol{\theta}; \mathbf{d}) &= \prod_{i=1}^n \left\{ c_{\theta} \left(S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right) f_1(x_i; \boldsymbol{\theta}^{(1)}) f_2(y_i; \boldsymbol{\theta}^{(2)}) \right\}^{\delta_{i1} \delta_{i2}} \\
 &\times \left\{ \frac{\partial C_{\theta} \left(S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right)}{\partial S_1(x_i; \boldsymbol{\theta}^{(1)})} f_1(x_i; \boldsymbol{\theta}^{(1)}) \right\}^{\delta_{i1}(1-\delta_{i2})} \\
 &\times \left\{ \frac{\partial C_{\theta} \left(S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right)}{\partial S_2(x_i; \boldsymbol{\theta}^{(2)})} f_2(y_i; \boldsymbol{\theta}^{(2)}) \right\}^{(1-\delta_{i1})\delta_{i2}} \\
 &\times C_{\theta} \left(S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right)^{(1-\delta_{i1})(1-\delta_{i2})}
 \end{aligned} \tag{4}$$

El estimador máximo verosimil para $\boldsymbol{\theta}$ es

$$\widehat{\boldsymbol{\theta}}_{\text{MV}} = \arg \max_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta}; \mathbf{d})$$

con Θ el espacio paramétrico. Equivalentemente

$$\widehat{\boldsymbol{\theta}}_{\text{MV}} = \arg \max_{\boldsymbol{\theta} \in \Theta} \ell(\boldsymbol{\theta}; \mathbf{d}) \tag{5}$$

donde $\ell(\boldsymbol{\theta}; \mathbf{d}) = \ln L(\boldsymbol{\theta}; \mathbf{d})$ es la función log-verosimilitud. Si la función log-verosimilitud alcanza su máximo en un punto interior de Θ , el estimador máximo verosimil es la solución de la ecuación

$$\frac{\partial}{\partial \boldsymbol{\theta}} \ell(\boldsymbol{\theta}; \mathbf{d}) = \mathbf{0}.$$

Sea $\boldsymbol{\theta}_0$ el vector de valores reales. Bajo ciertas condiciones de regularidad,

$$n^{1/2}(\widehat{\boldsymbol{\theta}}_{\text{MV}} - \boldsymbol{\theta}_0) \underset{\text{as.}}{\sim} N(\mathbf{0}, \mathcal{J}^{-1}(\boldsymbol{\theta}_0))$$

donde $\mathcal{J}(\boldsymbol{\theta}_0)$ es la matriz información.

Con este método (la estimación máxima verosimilitud completa), los parámetros $\boldsymbol{\theta}$ se estiman simultáneamente. Sin embargo, el enfoque con cópulas sugiere hacer la estimación en dos pasos:

1. primero estimar $\boldsymbol{\theta}^{(1)}$ y $\boldsymbol{\theta}^{(2)}$ separadamente maximizando las verosimilitudes

$$\begin{aligned}
 L_1(\boldsymbol{\theta}^{(1)}; \mathbf{d}^{(1)}) &= \prod_{i=1}^n f_1(x_i; \boldsymbol{\theta}^{(1)})^{\delta_{i1}} S_1(x_i; \boldsymbol{\theta}^{(1)})^{(1-\delta_{i1})} \\
 L_2(\boldsymbol{\theta}^{(2)}; \mathbf{d}^{(2)}) &= \prod_{i=1}^n f_2(y_i; \boldsymbol{\theta}^{(2)})^{\delta_{i2}} S_2(x_i; \boldsymbol{\theta}^{(2)})^{(1-\delta_{i2})}
 \end{aligned}$$

de los datos univariados. Equivalentemente, se puede maximizar $\ell_j(\boldsymbol{\theta}^{(j)}; \mathbf{d}^{(j)}) = \ln L_j(\boldsymbol{\theta}^{(j)}; \mathbf{d}^{(j)})$, $j = 1, 2$.

2. luego, estimar θ maximizando $L(\theta; \hat{\boldsymbol{\theta}}^{(1)}, \hat{\boldsymbol{\theta}}^{(2)}; \mathbf{d})$ dadas las estimaciones $\hat{\boldsymbol{\theta}}^{(1)}$ y $\hat{\boldsymbol{\theta}}^{(2)}$

Este método es llamado inferencia por funciones marginales (IFM) por Joe & Xu (1996). Los estimadores $\hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)}$ y $\hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)}$ están dados por

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)} &= \arg \text{máx } \ell_1(\boldsymbol{\theta}^{(1)}; \mathbf{d}^{(1)}) \\ \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)} &= \arg \text{máx } \ell_2(\boldsymbol{\theta}^{(2)}; \mathbf{d}^{(2)})\end{aligned}\tag{6}$$

y se sigue que el estimador para el parámetro de la cópula es

$$\hat{\theta}_{\text{IFM}} = \arg \text{máx } \ell(\theta; \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)}, \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)}, \mathbf{d})\tag{7}$$

Este método es computacionalmente más simple que la estimación simultánea de todos los parámetros. Una optimización numérica con muchos parámetros es más costosa en tiempo de computo comparada con varias optimizaciones con pocos parámetros (Joe 1997, pp. 300).

Quizás el parámetro de mayor interés es el parámetro θ , ya que refleja la dependencia que hay en los datos. La verosimilitud relativa del parámetro θ se define como

$$R(\theta) = \frac{L(\theta; \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)}, \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)}, \mathbf{d})}{L(\hat{\theta}_{\text{IFM}}; \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)}, \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)}, \mathbf{d})}$$

y asintóticamente $r^*(\theta) = -2 \ln[R(\theta)]$ tiene una distribución ji-cuadrada con un grado de libertad, cuando se evalúa en el valor real de θ . Al hacer una prueba de razón de verosimilitudes, se rechaza la hipótesis nula $\theta = \theta_0$, al nivel de significancia α , si

$$r^*(\theta) = -2 \ln[R(\theta)] > \chi_{(1-\alpha, 1)}^2$$

con $\chi_{(1-\alpha, 1)}^2$ el cuantil $1-\alpha$ de la distribución ji-cuadrada con un grado de libertad. También es posible, usando esta aproximación asintótica, construir intervalos de confianza aproximada del $100(1-\alpha)\%$.

La principal ventaja del método IFM es la simplificación de los cálculos numéricos. Sin embargo los estimadores IFM tienen otras propiedades deseables. Al igual que los estimadores de máxima verosimilitud, los estimadores IFM también son asintóticamente eficientes, y se tiene que

$$n^{1/2}(\hat{\boldsymbol{\theta}}_{\text{IFM}} - \boldsymbol{\theta}_0) \underset{\text{as.}}{\sim} N(\mathbf{0}, \mathfrak{B}(\boldsymbol{\theta}_0))$$

donde $\mathfrak{B}(\boldsymbol{\theta}_0)$ es llamada inversa de la matriz de información de Godambe. Si se define

$$\mathbf{g}(\boldsymbol{\theta}; \mathbf{d}) = \begin{bmatrix} \frac{\partial}{\partial \boldsymbol{\theta}^{(1)}} \ell_1(\boldsymbol{\theta}^{(1)}; \mathbf{d}^{(1)})^T \\ \frac{\partial}{\partial \boldsymbol{\theta}^{(2)}} \ell_2(\boldsymbol{\theta}^{(2)}; \mathbf{d}^{(2)})^T \\ \frac{\partial}{\partial \theta} \ell(\theta; \boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \mathbf{d}) \end{bmatrix}$$

entonces

$$\mathfrak{B}(\boldsymbol{\theta}_0) = \mathbf{D}^{-1} \mathbf{M} (\mathbf{D}^{-1})^T$$

donde $\mathbf{D} = \text{E} \left[\frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{g}(\boldsymbol{\theta}; \mathbf{d})^T \right]$ y $\mathbf{M} = \text{E} \left[\mathbf{g}(\boldsymbol{\theta}; \mathbf{d}) \mathbf{g}(\boldsymbol{\theta}; \mathbf{d})^T \right]$.

Para estimar la matriz de covarianzas asintóticas $n^{-1} \mathfrak{B}(\boldsymbol{\theta}_0)$ de $\hat{\boldsymbol{\theta}}_{\text{IFM}}$ es necesario calcular $\text{E} \left[\frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{g}(\boldsymbol{\theta}; \mathbf{d})^T \right]$, lo cual puede ser tedioso y complicado. Para evitar esto se pueden usar métodos de remuestreo como jackknife y bootstrap, en este trabajo decidimos usar jackknife, ya que, de acuerdo con Shao & Tu (1995), el procedimiento bootstrap no da mejores estimadores de la varianza que el procedimiento jackknife y además este último es computacionalmente más sencillo. Sea $\hat{\boldsymbol{\theta}}_{(-i)\text{IFM}}$ el estimador de $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \boldsymbol{\theta})$ con la i -ésima observación d_i excluida de los datos, $i = 1, \dots, n$. Si $\hat{\boldsymbol{\theta}}_{(-i)\text{IFM}}$ y $\hat{\boldsymbol{\theta}}_{\text{IFM}}$ son vectores renglones, el estimador jackknife de $n^{-1} \mathfrak{B}(\boldsymbol{\theta}_0)$ es (ver Joe 1997)

$$\sum_{i=1}^n \left(\hat{\boldsymbol{\theta}}_{(-i)\text{IFM}} - \hat{\boldsymbol{\theta}}_{\text{IFM}} \right)^T \left(\hat{\boldsymbol{\theta}}_{(-i)\text{IFM}} - \hat{\boldsymbol{\theta}}_{\text{IFM}} \right)$$

El método jackknife también se puede usar para calcular los errores estándar de funciones de los parámetros. El método delta requiere las derivadas de las funciones respecto a los parámetros, el método jackknife evita la necesidad de estas. Sea $h(\boldsymbol{\theta})$ la cantidad de interés (de valores reales), el estimador del error estándar de $h(\hat{\boldsymbol{\theta}}_{\text{IFM}})$ es

$$\left\{ \sum_{i=1}^n \left[h(\hat{\boldsymbol{\theta}}_{(-i)\text{IFM}}) - h(\hat{\boldsymbol{\theta}}_{\text{IFM}}) \right]^2 \right\}^{1/2} \quad (8)$$

Ya que la unidad falla cuando ocurre por primera vez alguno de los modos de falla, la función de confiabilidad de la unidad está dada por

$$\begin{aligned} S(t) &= \Pr(T_1 > t, T_2 > t) \\ &= \mathbf{S}(t, t) = C_{\theta}(S_1(t), S_2(t)) \end{aligned}$$

si se usa el modelo (1). Entonces, un estimador de la función de confiabilidad de la unidad es

$$\hat{S}(t) = C_{\hat{\boldsymbol{\theta}}_{\text{IFM}}} \left[S_1(t; \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)}), S_2(t; \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)}) \right]$$

Usando el método jackknife, específicamente la ecuación (8), podemos encontrar intervalos de confianza para la función de confiabilidad estimada de la unidad $\hat{S}(t)$.

Dado $p \in [0, 1]$, si resolvemos la ecuación $p = \hat{S}(t_p)$ para t_p se obtiene una estimación del tiempo en el que la unidad tiene una confiabilidad del $100p\%$. Ya que la función de confiabilidad es una función monótona, un intervalo de confianza aproximado del $100(1 - \alpha)\%$ para t_p es el conjunto $\{t : A(p) \leq \hat{S}(t) \leq B(p)\}$, donde $(A(p), B(p))$ es un el intervalo de confianza aproximado del $100(1 - \alpha)\%$ para $\hat{S}(t_p)$.

4. Aplicación a un estudio de encendedores

Se considera un estudio de confiabilidad de encendedores de hornos de estufas a gas, que tienen dos modos de falla. El motivo del estudio fue conocer la confiabilidad en dos tiempos de referencia, un año y diez años. Estos tiempos son de interés, debido a que las unidades donde se instalan los encendedores, tienen un tiempo de garantía de un año y un tiempo de vida de diez años. El fabricante del electrodoméstico donde se instalan los encendedores, planea cambiar los encendedores con el fin de reducir el número de reclamos de los usuarios, por fallas de los electrodomésticos debidas a la falla del encendedor. El fabricante planea introducir un nuevo encendedor que satisfaga las confiabilidades objetivo a uno y diez años, de 0.992 y 0.974, respectivamente. Un proveedor asegura que su producto cumple con estas exigencias, así que se llevó a cabo un estudio de confiabilidad de los encendedores que oferta el proveedor.

Los datos constan de 27 parejas de tiempos a la falla (x_i, y_i) , donde x_i corresponde a una falla debida a que el encendedor no alcanza el amperaje necesario para que se active el control electrónico del horno y y_i corresponde a la falla debida a que el encendedor no alcanza en el tiempo reglamentario (20 segundos) la temperatura mínima de operación. Los encendedores que se estudian, son resistencias de grafito en las que se aplica un voltaje determinado y su temperatura va creciendo hasta alcanzar un nivel estable determinado. Consideramos aquí dos modos de falla. Un modo de falla se presenta cuando la corriente que circula por el encendedor en operación es menor que el nivel mínimo de la corriente necesaria para activar la tarjeta de control electrónico. El otro modo de falla se presenta cuando la temperatura del encendedor a los 20 segundos de operación es menor a la temperatura mínima de operación, de 1050 °C.

Con el envejecimiento de los encendedores, ocurre que al aplicar un nivel determinado de voltaje, el crecimiento de su temperatura es cada vez más lento, y la corriente que circula por él, es cada vez menor. En este conjunto de datos tenemos varias censuras por la derecha, debido a que el proceso de medición de la corriente o la temperatura se interrumpió. Para acelerar las fallas en este estudio, no se llevó a cabo un estudio de vida acelerada, sino que se incrementó la tasa de uso de los componentes. Después de hacer un estudio de campo, se determinó el patrón de operación y el número promedio de horas de uso del encendedor al año. Posteriormente, en el laboratorio las unidades se mantuvieron en operación, con una frecuencia de uso mayor a la real y los tiempos a la falla registrados se expresaron en años de operación real. La gráfica de los datos se muestra en la figura 1. En ella se observa que hay una dependencia moderada entre ambos tiempos de falla.

Lo primero que hacemos es seleccionar la función cópula que se usará en el modelo (1). Recurrimos al método propuesto por Wang & Wells (2000) descrito en la sección 2 para elegir una de las funciones cópula mostradas en la tabla 1, para lo cual es necesario tener una estimación no-paramétrica de la función de confiabilidad bivariada $\mathcal{S}(\cdot, \cdot)$. El estimador de Dabrowska (1988) es el más común, aun cuando en algunas ocasiones el resultado obtenido no es una confiabilidad propia. Esto es lo que ocurre para los datos de encendedores, por lo que optamos por usar el estimador de $\mathcal{S}(\cdot, \cdot)$ propuesto por Batún-Cutz (2005) en su tesis doctoral.

Este estimador de la confiabilidad bivariada es una modificación del estimador de Dabrowska (1988) que corrige este problema.

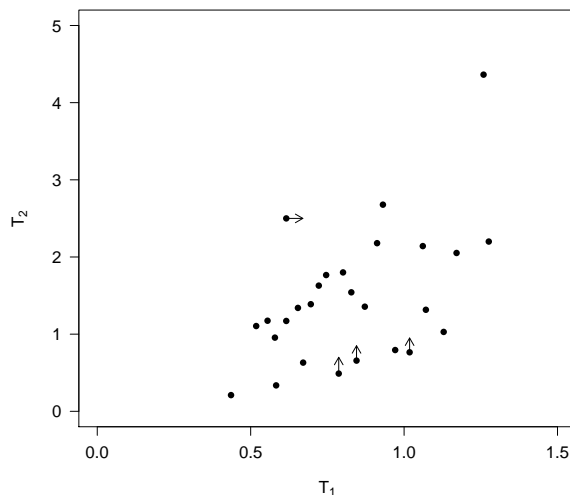


FIGURA 1: Tiempos de falla observados para los encendedores. La flecha indica el tiempo censurado.

Con el estimador $\tilde{S}(t_1, t_2)$ de Batún-Cutz (2005) calculamos $\tilde{\lambda}(z_i)$ para $z_i = \tilde{S}(x_i, y_i)$, $i = 1 \dots, 27$. El estimador de la τ de Kendall dado por la ecuación (3) es $\tilde{\tau} = 0.3939$, con éste obtenemos los estimadores $\tilde{\theta}$'s de los parámetros de las cópulas en la tabla 1 como se describe en el punto 2 del método para elegir una función cópula. Luego construimos los estimadores $\lambda_{\tilde{\theta}}(z)$ de las funciones $\lambda(z)$. La gráfica de las curvas $\lambda_{\tilde{\theta}}(z)$ y de los puntos $(z_i, \tilde{\lambda}(z_i))$ se muestran en la Figura 2 y en la tabla 2 se dan las distancias de los puntos a las curvas. Como los puntos $(z_i, \tilde{\lambda}(z_i))$ están más cerca de la curva $\lambda_{\tilde{\theta}}(z)$ que corresponde a la función cópula de Gumbel, elegimos esta función cópula para hacer la inferencia.

TABLA 2: Suma de las diferencias al cuadrado, $\sum_{i=1}^n [\tilde{\lambda}(z_i) - \lambda_{\tilde{\theta}}(z_i)]^2$

Cópula	Distancia $\times 10^3$
Frank	34.5867
Joe	28.0921
Clayton	96.7035
Gumbel	25.8142

Cuando se analizan los modos de falla por separado, se encuentra que las distribuciones más adecuadas para los tiempos de falla T_1 y T_2 son la Lognormal y la Weibull, respectivamente. Luego, las funciones de confiabilidad marginales para

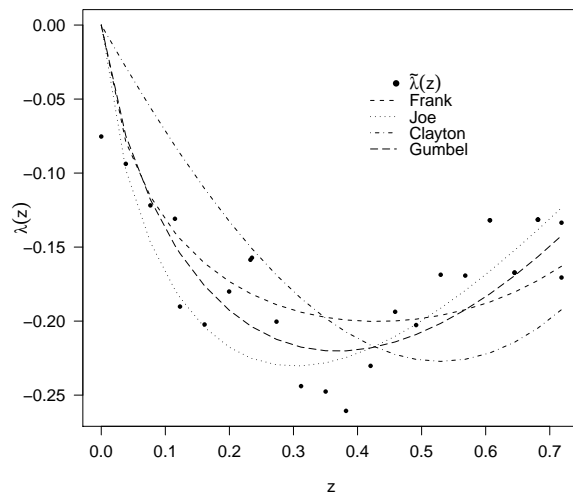


FIGURA 2: Gráfica de $\tilde{\lambda}(z)$ y $\lambda_{\tilde{\delta}}(z)$ para las cópulas consideradas.

los tiempos de falla serán

$$\begin{aligned}
 S_1(t_1; \mu, \sigma) &= 1 - \Phi\left(\frac{\ln t_1 - \mu}{\sigma}\right) \\
 S_2(t_2; \beta, \eta) &= \exp\left[-\left(\frac{t_2}{\eta}\right)^\beta\right]
 \end{aligned} \tag{9}$$

donde $\Phi(\cdot)$ es la función de distribución normal estándar. En la notación de la sección anterior, $\boldsymbol{\theta}^{(1)} = (\mu, \sigma)$ y $\boldsymbol{\theta}^{(2)} = (\beta, \eta)$ son los parámetros de las marginales.

La calidad del ajuste de las distribuciones marginales puede examinarse con una comparación gráfica de las confiabilidades ajustadas y de sus versiones empíricas como se muestra en la figura 3. Debido a la censura usamos el estimador de Kaplan-Meier para las confiabilidades empíricas. Note que la confiabilidad del primer modo de falla es menor, es decir, el primer modo de falla es el que determina el comportamiento de la unidad.

Entonces la función de confiabilidad bivariada de la unidad toma la forma

$$\mathcal{S}(t_1, t_2; \boldsymbol{\theta}) = \exp\left\{-\left[\left(-\ln S_1(t_1; \boldsymbol{\theta}^{(1)})\right)^\theta + \left(-\ln S_2(t_2; \boldsymbol{\theta}^{(2)})\right)^\theta\right]^{1/\theta}\right\}$$

donde $S_j(t_j; \boldsymbol{\theta}^{(j)})$, $j = 1, 2$ están dadas por las ecuaciones (9).

Los valores de los estimadores IFM de los parámetros $\boldsymbol{\theta} = (\mu, \sigma, \beta, \eta, \theta)$ se dan en la tabla 3, y la gráfica de contornos de la confiabilidad bivariada estimada se muestra en la figura 4. De la gráfica vemos que hay indicios de una dependencia entre los tiempos a la falla.

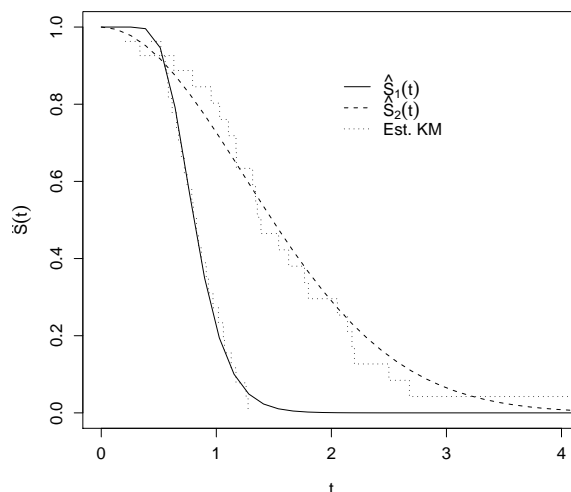


FIGURA 3: Funciones de confiabilidad univariadas estimadas de los modos de falla. La línea punteada corresponde a los estimadores Kaplan-Meier.

TABLA 3: Estimadores IFM, errores estándar e intervalos de confianza aproximados del 95 % para los datos de encendedores.

Parámetro	Estimación	Intervalo de	
		Error Estándar	confianza 95 %
μ	-0.2160	0.0566	(-0.327, -0.105)
σ	0.2806	0.0334	(0.2152, 0.3461)
β	1.9561	0.5591	(0.8603, 3.0519)
η	1.7932	0.2126	(1.3765, 2.2100)
θ	1.4732	0.2862	(0.9123, 2.0341)

Usamos la técnica jackknife para calcular la matriz de covarianzas asintótica de Godambe de las estimaciones IFM de los parámetros, la matriz de correlaciones correspondiente se da en la tabla 4. También es posible calcular intervalos de confianza aproximados del 95 % para las estimaciones de los parámetros, que se muestran en la tabla 3.

Como ya comentamos, el método IFM facilita los cálculos numéricos para estimar los parámetros. Esto es cierto sobre todo si el número de parámetros es grande. En el ejemplo aquí tratado sólo es necesario estimar cinco parámetros, y es posible estimarlos conjuntamente sin mucho esfuerzo. Los estimadores máximo verosímiles dados al resolver la ecuación (5) se dan en la tabla 5, observe que básicamente no hay diferencias en las estimaciones obtenidas por los dos métodos.

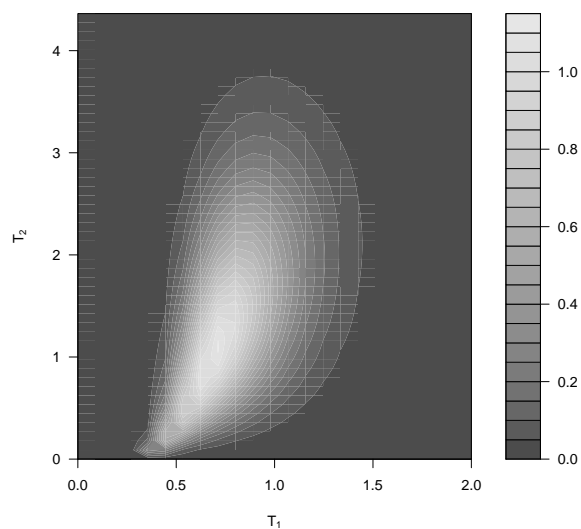


FIGURA 4: Gráfica de contornos de la densidad conjunta estimada de los tiempos a la falla de los encendedores.

TABLA 4: Matriz de correlaciones de las estimaciones IFM de los parámetros

	μ	σ	β	η	θ
μ	1.0000	-0.1682	-0.0451	0.6445	-0.2240
σ	-0.1682	1.0000	-0.4964	-0.0648	0.6323
β	-0.0451	-0.4964	1.0000	-0.3953	-0.7322
η	0.6445	-0.0648	-0.3953	1.0000	0.1373
θ	-0.2240	0.6323	-0.7322	0.1373	1.0000

TABLA 5: Estimadores máximo verosímiles, errores estándar e intervalos de verosimilitud-confianza aproximados del 95 % para los datos de encendedores.

Parámetro	Estimación	Intervalo del 95 % de confianza
μ	-0.2078	(-0.3165, -0.0939)
σ	0.2802	(0.2216, 0.3720)
β	1.9613	(1.4248, 2.5539)
η	1.8033	(1.4351, 2.2473)
θ	1.4687	(1.0429, 2.3307)

En la cópula de Gumbel tenemos el caso de independenciam cuando $\theta = 1$, por lo que podemos probar la hipótesis de modos de falla independientes, probando la

hipótesis $\theta = 1$. El estadístico de la prueba de la razón de verosimilitudes $r^*(\theta)$ toma el valor $r^*(1) = 9.6808$ que es mayor que $\chi^2_{(0.95,1)} = 3.84$, por lo que debemos rechazar la hipótesis nula $\theta = 1$, ya que tenemos una fuerte evidencia en contra de la independencia de los modos de falla.

Por la ecuación (2), la relación entre la τ de Kendall y el parámetro de la función cópula de Gumbel es $\tau = 1 - \theta^{-1}$. Entonces, es posible calcular una estimación de τ como $\hat{\tau} = 1 - (\hat{\theta}_{\text{IFM}})^{-1}$. Para los datos de ejemplo obtenemos $\hat{\tau} = 0.3212$, lo que indica que existe una dependencia moderada entre los modos de falla, como ya lo vimos en la prueba de hipótesis anterior.

En la figura 5 se muestran las estimaciones de la función de confiabilidad $\hat{S}(t) = \mathcal{S}(t, t; \hat{\theta}^{(1)}, \hat{\theta}^{(2)}, \theta)$ considerando el modelo con la cópula de Gumbel. Si se asume que los modos de falla son independientes, la función de confiabilidad de los encendedores toma la forma $\hat{S}^\perp(t) = S_1(t; \theta_{\text{IFM}}^{(1)})S_2(t; \theta_{\text{IFM}}^{(2)})$. Su gráfica también se muestra en la figura 5. Note que $\hat{S}^\perp(t) \leq \hat{S}(t)$, lo cual indica que al asumir independencia se subestima la función de confiabilidad del encendedor.

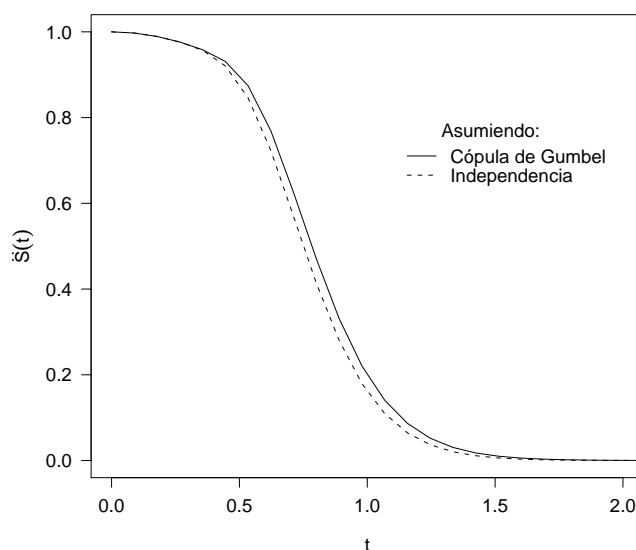


FIGURA 5: Confiabilidades estimadas de encendedores, se muestra la función de confiabilidad estimada con la cópula de Gumbel y la confiabilidad asumiendo independencia.

Con el fin de verificar las afirmaciones del proveedor, usamos la función de confiabilidad estimada $\hat{S}(t)$ para calcular estimaciones de la confiabilidad del encendedor en los tiempos de uno y diez años. Las estimaciones de la confiabilidad para estos tiempos, además de un mes y dos años, se dan en la tabla 6. Vemos que las confiabilidades estimadas están muy por debajo de las confiabilidades ob-

jetivos. Al estimar el intervalo de confianza de la confiabilidad a los 10 años con el método jackknife, encontramos que las estimaciones de la confiabilidad a los 10 años prácticamente no cambian, al excluir cualquiera de las observaciones. En consecuencia la desviación estándar calculada con este método es cero. Así, el intervalo de confianza aproximado del 95 % para la confiabilidad a 10 años consta sólo del punto cero.

TABLA 6: Confiabilidades estimadas para tiempos dados.

t_p	Confiabilidad	Desviación	Intervalo
	Estimada	Estándar	Confianza 95 %
Un mes	0.9975	0.0027	(0.9923, 1.0000)
Un año	0.1992	0.0580	(0.0856, 0.3129)
Dos años	0.0004	0.0006	(0.0000, 0.0154)
10 años	0.0000	–	–

Para ver los tiempos en que se consiguen las confiabilidades objetivo que plantea el cliente, resolvemos la ecuación $c = \widehat{S}(t_c)$ donde c es una confiabilidad dada. Así mismo, podemos dar un intervalo de confianza aproximado del 95 % para el tiempo t_c como se describe al final de la sección anterior. Los resultados para estas, y algunas otras, se muestran en la tabla 7. En ella observamos que los tiempos a los que se consiguen las confiabilidades objetivos son muy bajos. Por ejemplo, la confiabilidad de 0.992 se alcanza en $t_c = 0.1522$; esto significa que se esperan ocho fallas, de cada 1000 unidades, a los 55 días, aproximadamente. Más grave aún es lo que indica el último renglón en la tabla, el 20 % de las unidades fallarán antes de 7.3 meses.

TABLA 7: Cuantiles obtenidos para confiabilidades dadas.

Confiabilidad	Intervalo	
dada	t_c	Confianza 95 %
0.992	0.1522	(0.0000, 0.1818)
0.974	0.2794	(0.0000, 0.2931)
0.950	0.3887	(0.0000, 0.3576)
0.900	0.5014	(0.1408, 0.4012)
0.800	0.6005	(0.3159, 0.6811)

El análisis de confiabilidad de estos encendedores, muestra que éstos no cumplen con los niveles requeridos por el fabricante de electrodomésticos que planea hacer el cambio. Los resultados del estudio se enviaron al proveedor con la intención de que modificara el diseño del encendedor, para que su confiabilidad satisfaga los requerimientos del cliente.

5. Conclusiones

Los fabricantes siempre han estado preocupados por incrementar la confiabilidad de sus productos. Para lo cual se ven en la necesidad de hacer estudios de confiabilidad de los artículos que producen, así como de exigir estudios que avalen la confiabilidad de aquellos que adquieren. En este trabajo se ha mostrado el uso de la función cópula para modelar la confiabilidad de un artículo con dos modos de falla. Este enfoque es muy útil en la industria ya que es muy flexible con respecto a la selección de las distribuciones marginales.

Hemos utilizado funciones de cópula para modelar la estructura de dependencia para dos distribuciones marginales provenientes de diferentes familias, lognormal y Weibull. Lo cual no es posible con los procedimientos tradicionales que suponen marginales de la misma familia. Además, utilizamos el procedimiento de Wang y Wells para seleccionar la función de cópula entre un conjunto de cópulas propuestas. Esta selección se basa en las discrepancias entre la cópula empírica y las cópulas propuestas.

Hemos estimado los parámetros con dos métodos, IFM y máxima verosimilitud, y hemos obtenido resultados similares. Esto nos lleva a proponer el uso del método IFM debido a que es computacionalmente más simple que el de máxima verosimilitud y tiene propiedades similares a este como lo han demostrado Joe & Xu (1996).

Comparamos gráficamente las confiabilidades estimadas suponiendo independencia y dependencia modelada con la cópula de Gumbel. En esta comparación observamos que sistemáticamente es menor la confiabilidad en el caso de independencia.

Hemos ilustrado el método propuesto con un conjunto de datos reales, provenientes de la industria de electrodomésticos.

Como trabajo futuro consideramos hacer una comparación más exhaustiva entre los modelos que asumen dependencia e independencia. Esta comparación es interesante debido a que el análisis bajo el modelo de independencia es más simple que cuando suponemos un determinado patrón de dependencia, y es conveniente conocer cuanto se pierde al usar el modelo de independencia.

Agradecimientos

Agradecemos al Dr. Luis A. Escobar R. (Louisiana State University) sus comentarios y sugerencias, ya que fueron de gran utilidad para mejorar nuestra presentación; así mismo, queremos expresar nuestra gratitud al Editor invitado y a los árbitros por su lectura cuidadosa y recomendaciones que contribuyeron a mejorar el trabajo.

[Recibido: agosto de 2010 — Aceptado: febrero de 2011]

Referencias

- Arcila, L. D., Jaramillo, M. C. & Lopera, C. M. (2009), 'Selección de un modelo cópula para el ajuste de datos bivariados dependientes', *Dyna* **76**(158), 253–263.
- Batún-Cutz, J. L. (2005), Algunos problemas en inferencia no paramétrica: estimación de distribuciones ordenadas por picudez y supervivencia bivariada con censura, Tesis de doctorado, Centro de Investigación en Matemáticas, A. C., Guanajuato, México.
- Clayton, D. G. (1978), 'A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence', *Biometrika* **65**, 141–151.
- Dabrowska, D. M. (1988), 'Kaplan-Meier estimates on the plane', *The Annals of Statistics* **16**, 1475–1489.
- Escarela, G. & Hernández, A. (2009), 'Modelado de parejas aleatorias usando cópulas', *Revista Colombiana de Estadística* **32**(1), 33–58.
- Frank, M. J. (1979), 'On the simultaneous associativity of $f(x, y)$ and $1 - x + y - f(x, y)$ ', *Aequationes Mathematicae* **19**, 194–226.
- Genest, C. & Rivest, L. P. (1993), 'Statistical inference procedures for bivariate Archimedean copulas', *Journal of the American Statistical Association* **88**, 1034–1043.
- Hougaard, P. (1986), 'A class of multivariate failure time distributions', *Biometrika* **73**, 671–678.
- Jaworski, P., Durante, F., Hädel, W. & Rychlik, T. (2010), *Copula Theory and its Applications*, Vol. 1, Springer, Germany.
- Joe, H. (1993), 'Parametric families of multivariate distributions with given margins', *Journal of Multivariate Analysis* **46**, 573–581.
- Joe, H. (1997), *Multivariate Models and Dependence Concepts*, Chapman & Hall.
- Joe, H. & Xu, J. J. (1996), The estimation method of inference functions for margins for multivariate models, Technical Report 166, Department of Statistics, University of British Columbia.
- Lin, D. Y. & Ying, Z. (1993), 'A simple nonparametric estimator of the bivariate survival function under univariate censoring', *Biometrika* **80**, 573–581.
- Meeker, W. L. & Escobar, L. A. (1998), *Statistical Methods for Reliability Data*, John Wiley & Sons.
- Prentice, R. L. & Cai, J. (1992), 'Covariance and survivor function estimation using censored multivariate failure time data', *Biometrika* **79**, 495–512.

- Roger, B. N. (2006), *An Introduction to Copulas*, second edn, Springer.
- Shao, J. & Tu, D. (1995), *The Jackknife and Bootstrap*, Springer-Verlag.
- Wang, W. & Wells, M. T. (1997), 'Nonparametric estimation of the bivariate survival function under simplified censoring conditions', *Biometrika* **84**, 863–880.
- Wang, W. & Wells, M. T. (2000), 'Model selection and semiparametric inference for bivariate failure-time data', *Journal of the American Statistical Association, Theory and Methods* **95**(449), 62–72.