

Nuevas cartas de control basadas en la distribución Birnbaum-Saunders y su implementación

New Control Charts Based on the Birnbaum-Saunders Distribution and their Implementation

VÍCTOR LEIVA^{1,a}, GERSON SOTO^{2,b}, ENRIQUE CABRERA^{1,3,c},
GUILLERMO CABRERA^{4,d}

¹DEPARTAMENTO DE ESTADÍSTICA, UNIVERSIDAD DE VALPARAÍSO, VALPARAÍSO, CHILE

²INSTITUTO NACIONAL DE ESTADÍSTICAS, SANTIAGO, CHILE

³INSTITUTO DE ESTADÍSTICA, PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO,
VALPARAÍSO, CHILE

⁴ESCUELA DE INGENIERÍA INFORMÁTICA, PONTIFICIA UNIVERSIDAD CATÓLICA DE
VALPARAÍSO, VALPARAÍSO, CHILE

Resumen

El modelo Birnbaum-Saunders (BS) es una distribución de vida que tiene propiedades interesantes y aplicaciones en varias áreas. Esto la ha convertido en un foco de investigación importante en el último tiempo. Sin embargo, la suma de variables aleatorias independientes BS (BSsum) no sigue una distribución BS. A través de la distribución BSsum, se pueden monitorear los tiempos de vida de productos expuestos a fallas mediante una carta de control de calidad. Los procedimientos clásicos de cartas de control suponen normalidad en la distribución de los datos. No obstante, una de las características principales de los tiempos de vida es que éstos generalmente siguen distribuciones asimétricas. Por tanto, si se quiere monitorear estos tiempos, se deben considerar cartas de control para distribuciones asimétricas, como es el caso de la distribución BS. El monitoreo de los tiempos de vida se realiza generalmente mediante el tiempo acumulado o el tiempo promedio hasta la ocurrencia de cierto número de fallas. Entonces, usando la distribución BSsum, desarrollamos, implementamos y aplicamos una nueva metodología para cartas de control basada en la distribución BS.

Palabras clave: lenguaje de computación R, métodos de verosimilitud, tiempos de vida.

^aProfesor titular. E-mail: victor.leiva@uv.cl

^bIngeniero estadístico. E-mail: gerson.soto@ine.cl

^cProfesor adjunto. E-mail: enrique.cabrera@uv.cl

^dProfesor permanente. E-mail: guillermo.cabrera@ucv.cl

Abstract

The Birnbaum-Saunders (BS) model is a life distribution with interesting properties and applications in several fields. This has transformed the BS model in an important research focus in recent decades. However, the sum of BS (BSsum) independent random variables does not follow a BS distribution. By means of the BSsum distribution, we can monitor the lifetime of products subject to failures using a quality control chart. Classic procedures for control charts assume normality in the distribution of the data. Nevertheless, one of the main characteristics of the lifetimes is that they generally follow asymmetric distributions. Therefore, if we want to monitor these lifetimes, we must consider control charts for asymmetric distributions, such as it is the case of the BS distribution. The monitoring of the lifetimes is carried out generally by the accumulated lifetime or the lifetime average until than a number of failures occurs. Thus, by using the BSsum distribution, we develop, implement and apply a new methodology for control charts based on the BS distribution.

Key words: Lifetime data, Likelihood methods, R computer language.

1. Introducción

Motivados por un problema de vibración en aviones comerciales, lo que produce fatiga de materiales, y basados en la Ley de Miner o de daño acumulativo, Birnbaum & Saunders (1969*b*) desarrollaron un modelo probabilístico asimétrico de dos parámetros que describe los tiempos de falla (o de vida) de especímenes de materiales expuestos a cargas cíclicas bajo estrés y que producen fatiga; ver Johnson, Kotz & Balakrishnan (1995, pp. 651-663) y Sanhueza, Leiva & Balakrishnan (2008). La distribución Birnbaum-Saunders (BS) está relacionada a la distribución normal y describe el tiempo de falla que transcurre hasta que cierta clase de daño acumulativo excede un umbral de resistencia máximo, provocando la falla del material. Desmond (1985) proporcionó una derivación más general de esta distribución basada en un modelo biológico y consolidó una justificación física para el uso de esta distribución relajando los supuestos hechos por Birnbaum & Saunders (1969*b*) indicando, por ejemplo, que la distribución BS puede también obtenerse desde modelos distintos al normal. Basados en este principio y usando argumentos estadísticos, Díaz-García & Leiva (2005) generalizaron la distribución BS obteniendo su densidad, algunas propiedades y varios casos particulares de la distribución Birnbaum-Saunders generalizada (BSG). Esta generalización permite obtener el modelo BS desde una clase de distribuciones simétricas en la recta de los números reales produciendo una gran flexibilidad y una estimación de parámetros robusta.

La distribución BS fue implementada por Leiva, Hernández & Riquelme (2006) en el software estadístico R (<http://www.R-project.org>) en un paquete llamado `bs`; ver R Development Core Team (2009). En este paquete, se pueden encontrar funciones para calcular probabilidades, estimar parámetros, generar números aleatorios y hacer estudios de bondad de ajuste y análisis de confiabilidad. Por otro lado, Barros, Paula & Leiva (2009) desarrollaron un paquete R llamado `gbs` para la

distribución BSG. Los tres generadores de números aleatorios BS y BSG existentes pueden revisarse en Leiva, Sanhueza, Sen & Paula (2008).

La distribución BS tiene características interesantes, pero no posee la propiedad reproductiva. Así, la suma de variables aleatorias (v.a.) BS (que llamaremos BSsum) no sigue una distribución BS. Raaijmakers (1980, 1981) halló la distribución del tiempo de vida útil de un sistema “standby” cuyas componentes fallan independientemente, de acuerdo con una distribución BS. La vida útil de tal sistema corresponde a la suma de los tiempos de falla de cada componente de este sistema siguiendo una distribución BSsum. Raaijmakers (1980) se basó en la transformada de Laplace para hallar la distribución de la suma de v.a. independientes (convolución). Este resultado permite relacionar la función de densidad de probabilidades (f.d.p.) de la vida útil de cada componente a la f.d.p. de la vida útil de todo el sistema.

Debido a que algunos procesos industriales evidencian la presencia de observaciones con un comportamiento asimétrico, distribuciones que tengan este patrón son las adecuadas para analizar características de calidad de este tipo de procesos. Como se mencionó, una de estas distribuciones es el modelo BS, la que es principalmente útil para modelar tiempos de vida de productos expuestos a fallas. Este tipo de tiempos, debería monitorearse mediante cartas de control basadas en distribuciones asimétricas. A la fecha, existen pocos trabajos en esta dirección, menos aún basados en la distribución BS, salvo el estudio desarrollado por Lio & Park (2008) para monitorear percentiles de esta distribución.

Conocer la distribución BSsum es particularmente importante en estadística industrial, por ejemplo, para implementar cartas de control para el tiempo de fallas de productos siguiendo una distribución BS. Al conocer la distribución BSsum, se pueden obtener de forma exacta los límites de control inferior (LCI) y superior (LCS) de la carta de control. Aunque el modelo BS se ha usado ampliamente como distribución de vida en ingeniería, trabajos recientes han aplicado este modelo a otras áreas, permitiendo considerarlo como una distribución probabilística general más que restringirla solamente al modelamiento de datos de tiempos de vida. Así, aunque una metodología de cartas de control basada en la distribución BS puede ser más apropiada para tiempos de falla, esta metodología puede usarse para cualquier v.a. positiva. Para más detalles acerca de nuevas aplicaciones del modelo BS, ver Leiva, Sanhueza & Saunders (2009). Para una revisión de distribuciones de vida, ver Marshall & Olkin (2007) y Saunders (2007). Para aplicaciones de la distribución BS en áreas diferentes a la ingeniería, ver Leiva, Barros, Paula & Galea (2007), Podlaski (2008), Barros, Paula & Leiva (2008), Leiva, Barros, Paula & Sanhueza (2008), Leiva, Sanhueza & Angulo (2009), Bhatti (2010), Vilca, Sanhueza, Leiva & Christakos (2010) y Leiva, Vilca, Balakrishnan & Sanhueza (2010).

El supuesto básico para implementar cartas de control clásicas para la media de un proceso es que la v.a. a controlar debe seguir una distribución normal; ver Duncan (1996). Este supuesto no siempre se cumple; ver Schoonhoven & Does (2010). La falta de normalidad es particularmente frecuente cuando se estudian tiempos de vida de productos sujetos a fallas. Como la información que aportan

estos tiempos es crucial para la estabilidad de un proceso productivo, se deben considerar cartas de control para distribuciones asimétricas, las que, como se mencionó, no han sido ampliamente desarrolladas. Algunas cartas de este tipo fueron propuestas por Cheng & Xie (2000) y Surucu & Sazak (2009) para las distribuciones lognormal y Weibull, respectivamente; ver también Vargas & Montañó (2005). Como se mencionó, una carta de control para la distribución BS fue propuesta por Lio & Park (2008), quienes crearon una metodología para el monitoreo de los percentiles de la distribución BS. Cartas de control para la media y el tiempo acumulado de procesos de producción gobernados por una distribución BS no han sido propuestas.

El objetivo principal de este artículo es desarrollar una metodología de cartas de control para la media y el tiempo acumulado de fallas de un proceso basadas en la distribución BS. Específicamente, en este artículo: i) introducimos la distribución BSsum; ii) desarrollamos cartas de control basadas en esta distribución; iii) implementamos la distribución BSsum en el software R, así como cartas de control basadas en la distribución BS; iv) llevamos a cabo un estudio de simulación que detecta la sensibilidad de la metodología propuesta a salidas de control del proceso productivo, y v) aplicamos los resultados obtenidos a datos industriales reales. Cabe destacar que en este artículo discutimos también los aspectos computacionales de este trabajo. Específicamente, en dos apéndices localizados después de las conclusiones, discutimos una nueva versión del paquete `bs` llamada `bs 2.0` y un paquete para la distribución BSsum llamado `bssum` que contiene las cartas de control BS. El paquete `bs 2.0` es más completo que el `bs 1.0`, ya que incorpora, por ejemplo, funciones para datos censurados. Este tipo de datos no son analizados aquí, ya que una metodología como la desarrollada en este trabajo no puede aplicarse directamente cuando existen datos censurados. Esto nos propone un desafío para un futuro trabajo. El lector interesado en este tipo de métodos, puede revisar Steiner & Mackay (2000), Zhang & Chen (2004) y Vargas & Montañó (2005). El paquete `bs 1.0` puede descargarse desde CRAN (<http://CRAN.R-project.org>), mientras que los paquetes `bs 2.0` y `bssum` puede descargarse desde <http://staff.deuv.cl/leiva/archivos>.

El resto de este artículo está organizado como sigue. En la sección 2 proporcionamos algunos elementos preliminares de la distribución BS útiles para desarrollar la metodología propuesta. En la sección 3, introducimos la distribución BSsum. En la sección 4 discutimos una carta de control para percentiles de la distribución BS y desarrollamos una nueva metodología para cartas de control basadas en esta distribución. En esta sección, llevamos a cabo también un estudio de simulación y dos ejemplos con datos industriales reales. En la sección 5, bosquejamos algunas conclusiones de este trabajo. En la parte final, en dos apéndices analizamos algunas características de los paquetes `bs 1.0`, `bs 2.0` y `bssum`.

2. Preliminares

En esta sección proporcionamos algunas propiedades y características de la distribución BS.

2.1. Distribución Birnbaum-Saunders

Una v.a. T con distribución BS tiene dos parámetros, uno de forma (α) y otro de escala (β), con β siendo además un parámetro de posición, pues corresponde a la mediana. Esto se denota por $T \sim BS(\alpha, \beta)$. Las variables aleatorias BS y normal estándar, denotadas respectivamente por T y Z , están relacionadas mediante

$$T = \beta \left[\frac{\alpha Z}{2} + \sqrt{\left\{ \frac{\alpha Z}{2} \right\}^2 + 1} \right]^2 \quad \text{y} \quad Z = \frac{1}{\alpha} \left[\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right] \quad (1)$$

Esto nos permite llegar a

$$W = \frac{1}{\alpha^2} \left[\frac{T}{\beta} + \frac{\beta}{T} - 2 \right] \sim \chi^2(1) \quad (2)$$

cuyo resultado es útil para bondad de ajuste y detectar datos atípicos mediante la distancia de Mahalanobis. Si $T \sim BS(\alpha, \beta)$, entonces las siguientes características se cumplen. La f.d.p. de T es

$$f_T(t) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2\alpha^2} \left[\frac{t}{\beta} + \frac{\beta}{t} - 2 \right] \right) \frac{t^{-3/2}[t + \beta]}{2\alpha\sqrt{\beta}}, \quad t > 0, \alpha > 0, \beta > 0 \quad (3)$$

La función de distribución acumulativa (f.d.a.) de T es

$$F_T(t) = \mathbb{P}(T \leq t) = \Phi \left(\frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right] \right) \quad t > 0 \quad (4)$$

donde $\Phi(\cdot)$ es la f.d.a. de $Z \sim N(0, 1)$. La función cuantil (f.q.) de T (cuantil q -ésimo) es

$$t(q) = F_T^{-1}(q) = \frac{\beta}{4} \left[\alpha z(q) + \sqrt{\{\alpha z(q)\}^2 + 4} \right]^2, \quad 0 < q < 1 \quad (5)$$

donde $F_T^{-1}(\cdot)$ es la función inversa de $F_T(\cdot)$ y $z(q)$ es el cuantil q -ésimo de $Z \sim N(0, 1)$. Por tanto, desde (5), $t(0.5) = \beta$ y así β es la mediana del modelo BS, tal como fue mencionado. Algunas propiedades de $T \sim BS(\alpha, \beta)$ son: (i) $cT \sim BS(\alpha, c\beta)$, con $c > 0$, y (ii) $1/T \sim BS(\alpha, 1/\beta)$. Estas propiedades indican que la distribución BS pertenece a las familias de transformaciones de escala y cerradas bajo recíproco. El momento k -ésimo de T está dado por

$$E [T^k] = \beta^k \sum_{j=0}^k \binom{2k}{2j} \sum_{i=0}^j \binom{j}{i} \frac{(2k - 2j + 2i)!}{2^{k-j+i}(k - j + i)!} \left[\frac{\alpha}{2} \right]^{2[k-j+i]}, \quad k = 1, 2, \dots \quad (6)$$

La media, la varianza y los coeficientes de variación (CV), sesgo (CS) y curtosis (CC) de T están dados respectivamente por

$$E[T] = \frac{\beta}{2} [2 + \alpha^2], \quad V[T] = \frac{\beta^2}{4} [5\alpha^4 + 4\alpha^2], \quad CV[T] = \frac{\sqrt{5\alpha^4 + 4\alpha^2}}{\alpha^2 + 2} \quad (7)$$

$$CS[T] = \frac{44\alpha^3 + 24\alpha}{[5\alpha^2 + 4]^{3/2}} \quad \text{y} \quad CC[T] = 3 + \frac{558\alpha^4 + 240\alpha^2}{[5\alpha^2 + 4]^2} \quad (8)$$

Entonces, desde (8), $CS[T] \rightarrow 0$ y $CC[T] \rightarrow 3$, cuando $\alpha \rightarrow 0$, es decir, cuando α es pequeño, el sesgo y la curtosis de la distribución BS se acercan al sesgo y la curtosis de la distribución normal degenerando en el parámetro β . Los CV, CS y CC son invariantes bajo escala, es decir, estos coeficientes son funciones independientes del parámetro de escala β . Por otra parte, si T tiene una distribución BS con parámetros α y β y ya que $1/T$ tiene también una distribución BS con los correspondientes parámetros α y $1/\beta$, respectivamente, entonces se tiene que

$$E\left[\frac{1}{T}\right] = \frac{1}{2\beta} [\alpha^2 + 2] \quad \text{y} \quad V\left[\frac{1}{T}\right] = \frac{1}{4\beta^2} [5\alpha^4 + 4\alpha^2] \quad (9)$$

La figura 1 muestra el comportamiento de la f.d.p. de la distribución BS para algunos valores del parámetro de forma α . Note que a medida que α decrece, la f.d.p. es aproximadamente simétrica. Gráficos para diferentes valores de β no se han considerado, porque al ser éste un parámetro de escala y de posición, la forma de la f.d.p. no cambia al variar este parámetro. Así, sin pérdida de generalidad, hemos considerado $\beta = 1$ en esta figura.

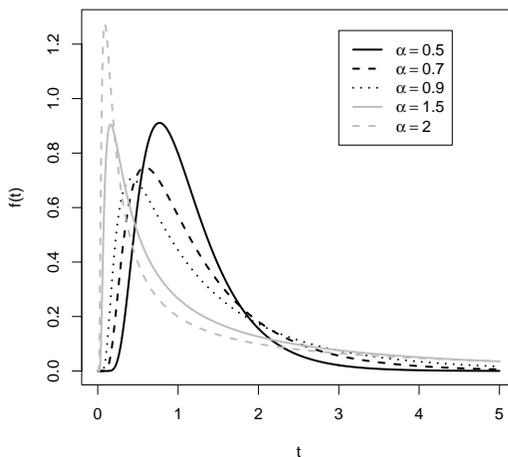


FIGURA 1: f.d.p. de $T \sim BS(\alpha, \beta = 1.0)$.

2.2. Métodos de estimación para la distribución BS

Se han planteado varios estimadores para los parámetros de forma α y de escala β de la distribución BS. A continuación se presentan tres métodos para estimar estos parámetros. Estos métodos son el de verosimilitud máxima (VM), uno gráfico que permite la estimación de α y β por mínimos cuadrados y otro de momentos modificado; ver Birnbaum & Saunders (1969a), Pérez & Correa (2008), Chang & Tang (1994) y Ng, Kundu & Balakrishnan (2003). Para cada uno de estos

métodos, considere que T_1, \dots, T_n es una muestra aleatoria de tamaño n , donde $T_i \sim BS(\alpha, \beta)$, para $i = 1, \dots, n$.

2.2.1. Método de verosimilitud máxima

La función de log-verosimilitud para α y β está dada por

$$\ell(\alpha, \beta) \propto \frac{n}{\alpha^2} - n \log(\alpha) - \frac{n}{2} \log(\beta) + \sum_{i=1}^n \left\{ \log(t_i + \beta) - \frac{1}{2\alpha^2} \left[\frac{t_i}{\beta} + \frac{\beta}{t_i} \right] \right\} \quad (10)$$

Derivando (10) con respecto a los parámetros α y β , se obtiene, respectivamente,

$$\frac{\partial \ell(\alpha, \beta)}{\partial \alpha} = -\frac{2n}{\alpha^3} - \frac{n}{\alpha} + \sum_{i=1}^n \left\{ \frac{1}{\alpha^3} \left[\frac{t_i}{\beta} + \frac{\beta}{t_i} \right] \right\} \quad y \quad (11)$$

$$\frac{\partial \ell(\alpha, \beta)}{\partial \beta} = -\frac{n}{2\beta} + \sum_{i=1}^n \left\{ \frac{1}{t_i + \beta} \right\} + \frac{1}{2\alpha^2} \sum_{i=1}^n \left\{ \frac{t_i}{\beta^2} - \frac{1}{t_i} \right\} \quad (12)$$

Al igualar ecuaciones de log-verosimilitud dadas en (11) y (12) a cero, se obtiene

$$\hat{\alpha}^2 = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{t_i}{\hat{\beta}} + \frac{\hat{\beta}}{t_i} \right\} - 2 \quad y \quad (13)$$

$$\hat{\alpha}^2 = \frac{2\hat{\alpha}^2 \hat{\beta}}{n} \sum_{i=1}^n \left\{ \frac{1}{t_i + \hat{\beta}} \right\} + \frac{1}{n} \sum_{i=1}^n \left\{ \frac{t_i}{\hat{\beta}} - \frac{\hat{\beta}}{t_i} \right\} \quad (14)$$

Considere la media aritmética y la media armónica de un conjunto de números positivos, digamos t_1, \dots, t_n , respectivamente, dadas por

$$s = \frac{1}{n} \sum_{i=1}^n t_i \quad y \quad r = \frac{1}{\left[\frac{1}{n} \sum_{i=1}^n \frac{1}{t_i} \right]} \quad (15)$$

Considere además una función media armónica $K(x) = n / \sum_{i=1}^n \{1/(x + t_i)\}$, para $x > 0$. Sustituyendo s y r y $K(x)$ en (13) y (14), se obtiene

$$\hat{\alpha}^2 = \frac{s}{\hat{\beta}} - \frac{\hat{\beta}}{r} + \frac{2\hat{\alpha}^2 \hat{\beta}}{K(\hat{\beta})} \quad y \quad (16)$$

$$\frac{\hat{\beta}}{r} = 1 + \frac{\hat{\alpha}^2 \hat{\beta}}{K(\hat{\beta})} \quad (17)$$

Si se sustituye (16) en (17), se obtiene $\hat{\beta}K(\hat{\beta})/r = K(\hat{\beta}) + \hat{\beta}[s/\hat{\beta} + \hat{\beta}/r - 2]$, de modo que $\hat{\beta}^2 - \hat{\beta}[2r - K(\hat{\beta})] + r[s + K(\hat{\beta})] = 0$. Ahora, tomando $g(x) = x^2 - x[2r - K(x)] + r[s + K(x)]$, se tiene que el estimador de VM de β , digamos $\hat{\beta}$, es la solución de $g(x) = 0$, la cual es única ya que $0 < x < \infty$ y $s > \hat{\beta} > r$; ver

Birnbaum & Saunders (1969a). Aś, el estimador de VM de α , digamos $\hat{\alpha}$, queda expresado en t́rminos de $\hat{\beta}$ mediante

$$\hat{\alpha} = \left[\frac{s}{\hat{\beta}} + \frac{\hat{\beta}}{r} - 2 \right]^{1/2} \quad (18)$$

mientras que $\hat{\beta}$ debe obtenerse usando un ḿtodo numérico iterativo por lo que se puede usar la mediana muestral como valor de partida o la estimación media-media para β propuesta por Birnbaum & Saunders (1969a) y dada por $\bar{\beta} = [sr]^{1/2}$.

2.2.2. Ḿtodo de ḿnimos cuadrados

Desde (4) se tiene que

$$t = \beta + \alpha \sqrt{\beta} \sqrt{t} \Phi^{-1}(F_T(t)) \quad (19)$$

Note que (19) no es una funci3n lineal en t , lo que es esencial para gráfcos de probabilidad de bondad de ajuste. Se puede eludir este problema reemplazando $p = \sqrt{t} \Phi^{-1}(F_T(t))$ en (19), dando como resultado la forma lineal $y = a + bx$, donde la ordenada es $y = t$, el intercepto $a = \beta$, la abscisa es $x = p$ y la pendiente es $b = \alpha \sqrt{\beta}$, es decir, $t = \beta + \alpha \sqrt{\beta} p$. Luego, se debe graficar t_i versus p_i , para $i = 1, \dots, n$, donde $p_i = \sqrt{t_i} \Phi^{-1}(F(t_i))$. Existen varias opciones para obtener $F_n(t_i)$. Aqú se utiliza el rango medio, es decir, $F_n(t_i) = [i - 0.3]/[n + 0.4]$. Aś, este ḿtodo puede usarse como una herramienta de bondad de ajuste para detectar si un conjunto de datos podŕa o no provenir de una distribuci3n BS. En particular, si el gráfcico de t_i versus p_i , para $i = 1, \dots, n$, sigue aproximadamente una ĺnea recta, entonces esto es una indicaci3n de que la muestra aleatoria podŕa provenir de una distribuci3n BS. Chang & Tang (1994) obtuvieron estimaciones del intercepto a y de la pendiente b mediante el ḿtodo de cuadrados ḿnimos, lo que es denotado por \tilde{a} y \tilde{b} , respectivamente. Ellos indicaron que una vez obtenidas estas estimaciones, uno puede determinar los parámtros de la distribuci3n BS mediante las expresiones

$$\tilde{\beta} = \tilde{a} \quad \text{y} \quad \tilde{\alpha} = \frac{\tilde{b}}{\sqrt{\tilde{a}}}$$

2.2.3. Ḿtodo de momentos modificados

En el caso de distribuciones de dos parámtros, sus estimadores de momentos se obtienen igualando los primeros dos momentos poblacionales a sus momentos muestrales correspondientes. Por tanto, los estimadores de momentos de α y β de la distribuci3n BS se pueden obtener como las soluciones de α y β de estas ecuaciones de momentos. Si el CV muestral es mayor que $\sqrt{5}$, entonces el estimador de momentos de α no existe. Por el contrario, si el CV es menor que $\sqrt{5}$, el estimador de momentos sí existe, pero el estimador de β no es único. En lugar de utilizar expresiones para los momentos de la distribuci3n BS dadas en (7), Ng et al. (2003) propusieron utilizar las expresiones del lado izquierdo de las ecuaciones (7)

y (9) e igualarlas a sus momentos muestrales correspondientes. En este caso, se tienen las ecuaciones de momentos

$$s = \beta \left[1 + \frac{1}{2} \alpha^2 \right] \quad \text{y} \quad \frac{1}{r} = \frac{1}{\beta} \left[1 + \frac{1}{2} \alpha^2 \right] \quad (20)$$

Resolviendo las ecuaciones dadas en (20) para α y β , se obtienen las estimaciones de momentos modificados para α y β denotadas por $\bar{\alpha}$ y $\bar{\beta}$ como

$$\bar{\alpha} = \left[2 \left\{ \left(\frac{s}{r} \right)^{1/2} - 1 \right\} \right]^{1/2} \quad \text{y} \quad \bar{\beta} = [sr]^{1/2} \quad (21)$$

2.3. El modelo BS como una distribución de vida

Una de las principales herramientas del análisis de confiabilidad es la función tasa de riesgo (f.r.) o tasa de fallas. En la ciencia actuarial, por ejemplo, la f.r. es la probabilidad por año que una persona de una edad determinada muera en el instante siguiente, expresada como una tasa de mortalidad por año. Una propiedad de la f.r. es que hace posible caracterizar el comportamiento de las distribuciones de vida. Este tipo de distribuciones describen los tiempos de vida de unidades hasta la ocurrencia de una falla, pudiendo estas unidades ser, por ejemplo, sistemas, componentes, órganos, personas o productos; para más detalles acerca de distribuciones de vida, ver Marshall & Olkin (2007) y Saunders (2007). Por ejemplo, modelos de probabilidad con densidades con formas similares en algunos casos tienen f.r. distintas. Éste es el caso de las distribuciones BS, Gaussiana inversa (GI), lognormal, gamma y Weibull; ver Balakrishnan, Leiva & López (2007). Para revisar algunas relaciones entre las distribuciones BS y GI, ver Balakrishnan, Leiva, Sanhueza & Cabrera (2009). Por tanto, es muy importante tener en cuenta la f.r. para seleccionar o descartar distribuciones, incluso cuando no se estén analizando datos de tiempos de vida. Por ejemplo, si la f.r. es creciente, la probabilidad de sobrevivencia disminuye con el tiempo. Este tipo de f.r. puede considerarse para modelar tiempos de vida cuando el desgaste o envejecimiento está presente. Sin embargo, una f.r. creciente puede ser inadecuada cuando se modela mortalidad humana, ya que, en este caso, el riesgo está disminuyendo, se estabiliza durante un período y luego éste es cada vez mayor, de manera que la forma que se debe suponer para la f.r. es aquella conocida como gráfico de bañera (en inglés *bathtub*). Otro tipo de f.r. tiene forma de bañera invertida, la que aparece con frecuencia cuando se analizan datos de fatiga y de contaminación ambiental. Una f.r. decreciente puede ser adecuada para modelar la sobrevivencia después de una cirugía exitosa, cuando hay un riesgo inicial alto causado, por ejemplo, por una infección o hemorragia, pero este riesgo comienza luego a disminuir a medida que el paciente se va recuperando. Por último, existen casos de f.r. constante, la que se observa, por ejemplo, cuando se estudia la duración de chips de computadores, los que no envejecen con el tiempo.

Distribuciones de vida que modelan los diferentes tipos de riesgo mencionados arriba pueden escogerse mediante la tasa de falla. Entonces, establecer la forma de la f.r. que un conjunto de datos tiene es de vital importancia en análisis de

confiabilidad para determinar la distribución que estos datos pueden tener. La confiabilidad o sobrevivencia se define como la probabilidad de que una entidad sobreviva o funcione adecuadamente en un período de tiempo t ; ver Lawless (1982). Así, si T es el tiempo de vida de una unidad, entonces la función de confiabilidad (o sobrevivencia), lo cual abreviamos como f.c., al tiempo t está dada por $R_T(t) = \mathbb{P}(T \geq t) = 1 - F_T(t) = \int_t^\infty f_T(s) ds$, para $t > 0$, donde $f_T(\cdot)$ y $F_T(\cdot)$ son la f.d.p. y la f.d.a. de T , respectivamente. La f.r. es un indicador muy importante en el análisis de tiempos de vida y se define como la tasa instantánea de fallas en el tiempo $t + \Delta t$, dado que la unidad ha sobrevivido a t . Entonces, la f.r. de una v.a. T es

$$h_T(t) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{P}(t < T < \Delta t | T > t)}{\Delta t} = \frac{f_T(t)}{R_T(t)} = -\frac{d \log(R_T(t))}{dt}, \quad t > 0 \quad (22)$$

donde $0 < R_T(t) < 1$. La f.d.p. puede estimarse fácilmente en forma empírica a través de un histograma. Sin embargo, una estimación como ésa para la f.r. no es fácil. Una herramienta que permite determinar la forma de la f.r. es el gráfico TTT (en inglés, *total time on test*). Específicamente, si T es una v.a. positiva, sus funciones TTT y TTT escalada son $H_T^{-1}(u) = \int_0^{F_T^{-1}(u)} R_T(y) dy$ y $W_T(u) = H_T^{-1}(u)/H_T^{-1}(1)$, respectivamente, para $0 < u < 1$. La función TTT escalada $W_T(\cdot)$ puede aproximarse en forma empírica mediante

$$W_n(k/n) = \frac{\sum_{i=1}^k T_{(i)} + [n-k] T_{(k)}}{\sum_{i=1}^n T_{(i)}}, \quad k = 1, \dots, n \quad (23)$$

donde $T_{(i)}$ es el estadístico de orden i -ésimo. Esto permite construir el gráfico TTT empírico mediante los puntos $[k/n, W_n(k/n)]$. Así, por medio de este gráfico se puede detectar el tipo de f.r. que los datos tienen. En la figura 8 (izquierda) es posible observar diferentes formas teóricas para el gráfico TTT, las cuales están asociadas con la f.r. respectiva. Si este gráfico arroja una curva cóncava, la f.r. es creciente. Por el contrario, si esta curva es convexa, la f.r. es decreciente. Ahora bien, si el gráfico TTT produce una curva que es primero cóncava y luego convexa, la f.r. tiene forma de bañera invertida. Si esta curva es convexa y luego cóncava, la f.r. tiene forma de bañera. Por último, si el gráfico TTT produce una línea recta, la f.r. es constante, como es el caso de la distribución exponencial. Así, el gráfico TTT nos puede dar una indicación del tipo de distribución asociada con los datos; ver la figura 2.

El modelo BS ha sido usado ampliamente como distribución de vida debido a la justificación física que lo originó; ver Birnbaum & Saunders (1969b) y Johnson et al. (1995, pp. 651-652). Si $T \sim \text{BS}(\alpha, \beta)$, entonces la f.r. de T es

$$h_T(t) = \frac{\phi([1/\alpha][\sqrt{t/\beta} - \sqrt{\beta/t}] t^{-3/2}[t + \beta])}{\Phi([1/\alpha][\sqrt{t/\beta} - \sqrt{\beta/t}]) 2\alpha\sqrt{\beta}} \quad t > 0 \quad (24)$$

donde $\phi(\cdot)$ es la f.d.p. $N(0,1)$. La figura 3 (izquierda) muestra que la f.c. BS decrece a medida que α aumenta. La figura 3 (derecha) indica que la f.r. BS tiende a ser creciente a medida que α decrece.

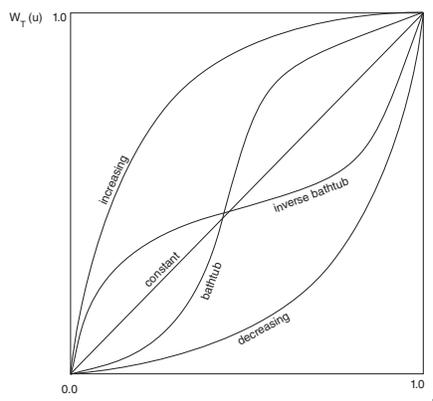


FIGURA 2: Curvas TTT para distribuciones con la f.r. indicada.

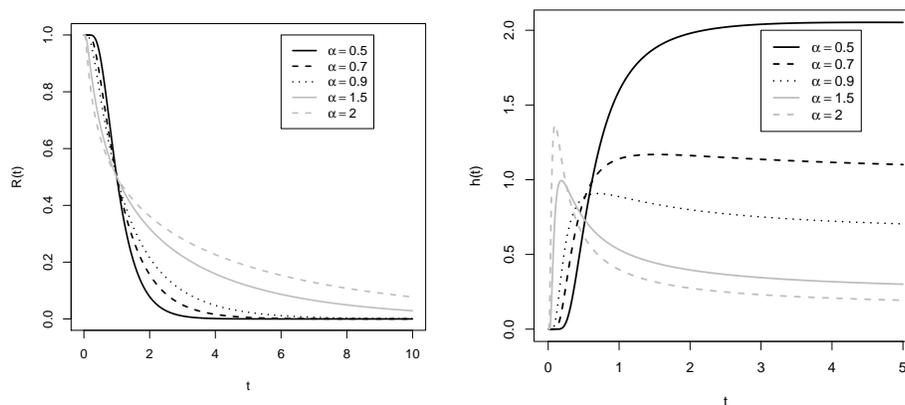


FIGURA 3: f.c. (izquierda) y f.r. (derecha) BS para diferentes valores de α .

3. Distribución BSsum

En esta sección introducimos algunas propiedades y características de la distribución BSsum.

3.1. Vida útil de un sistema de k componentes independientes

Un sistema se muestra consistente cuando la primera unidad falla y la segunda unidad, sin demora, se hace cargo de la operación. Este proceso continúa hasta que todas las k unidades que componen el sistema han fallado. Esto se conoce como sistema “standby”. Considere T_i como la vida útil de la unidad i -ésima en el sistema, para $i = 1, \dots, k$, y f como la f.d.p. para todos los T_i , que se suponen independientes. Considere además que Y_k es la vida útil del sistema de k unidades

independientes y f_k es la f.d.p. de Y_k , para $k \in \mathbb{N}$. Claramente, se puede concluir que $Y_k = \sum_{i=1}^k T_i$. Es posible establecer una relación entre las funciones f y f_k por medio de sus transformadas de Laplace. Específicamente, para cualquier función g definida entre $(0, \infty)$, digamos Lg , cuando existe, es la transformada de Laplace definida como $Lg(s) = \int_0^\infty \exp(-su) g(u) du$, para $s > 0$. La relación existente entre las funciones f y f_k es $Lf_k = [Lf]^k$. Por tanto, para obtener la distribución BSSum se debe calcular la transformada de Laplace de la f.d.p. BS, elevarla a la potencia k (donde k es el número de términos de la suma) y después calcular la transformada de Laplace inversa de esta potencia.

3.2. Distribución BSSum

Considere $f_{\alpha,\beta}$ como la f.d.p. BS con parámetros α y β definida en (3). Sin pérdida de generalidad, se puede tomar el parámetro de escala $\beta = 1$ y aplicar la transformada de Laplace a esta f.d.p. Así, la transformada de Laplace de $f_{\alpha,1}$ es

$$Lf_{\alpha,1}(s) = \frac{\exp\left(\frac{1}{\alpha^2}\right)}{2\alpha\sqrt{2\pi}} \left[\sqrt{\frac{\pi}{u}} + \sqrt{\frac{\pi}{a}} \right] \exp(-2\sqrt{ua}) \quad (25)$$

donde $u = s + a$ y $a = 1/2\alpha^2$. Se define la función $q(\cdot)$ como

$$q(s) = \frac{1}{2} \left[1 + \frac{1}{\sqrt{s}} \right] \exp\left(\frac{[1 - \sqrt{s}]}{\alpha^2}\right) \quad (26)$$

Después de algunos cálculos se desprende que $Lf_{\alpha,1}(s) = r[1+2\alpha^2s]$. Este resultado se debe elevar a la cantidad de términos que tiene la suma. Para esto se define $Q(s) = q(s)^k$ y $g = k/\alpha^2$. Entonces, se obtiene

$$H(s) = \frac{\exp(g)}{2^k} \left[1 + \frac{1}{\sqrt{s}} \right]^k \exp(-g\sqrt{s}) = \frac{\exp(g)}{2^k} \sum_{i=0}^k \binom{k}{i} \frac{1}{s^{i/2}} \exp(-g\sqrt{s}) \quad (27)$$

Luego, considere $\mu_i(t)$ con

$$L\mu_i(s) = \frac{1}{s^{i/2}} \exp(-\sqrt{s}) \quad (28)$$

y una función $z(t)$, tal que $Lz(s) = H(s)$. Sustituyendo las ecuaciones (27) y (28) en $Lz(s)$, se tiene que

$$Lz(s) = \frac{\exp(g)}{2^k} \sum_{i=0}^k \binom{k}{i} g^i L\mu_i(g^2s) \quad (29)$$

lo cual implica

$$z(t) = \frac{\exp(g)}{2^k} \sum_{i=0}^k \binom{k}{i} g^{i-2} \mu_i\left(\frac{t}{g^2}\right)$$

Así, después del uso de algunos teoremas de la transformada de Laplace, la f.d.p. BSsum es

$$f_k(t) = \frac{a}{2^k} \exp(g - at) \sum_{i=0}^k \binom{k}{i} g^{i-2} \mu_i \left(\frac{at}{g^2} \right) \quad (30)$$

donde $a = 1/2\alpha^2$ y $g = k/\alpha^2$. La f.d.a. de Y_k , que es la suma de v.a. BS, está dada por $F_k(t) = \mathbb{P}(Y_k \leq t) = \int_0^t f_k(u) du$, para $t > 0$, donde $f_k(t)$ está dada en (30), esto es,

$$F_k(t) = \frac{1}{2^k} \exp(g - at) \sum_{i=2}^k \left[l_i g^{i-2} \mu_i \left(\frac{at}{g^2} \right) \right] + \Phi(\alpha^{-1} \varphi_k(t)), \quad t > 0 \quad (31)$$

donde $l_i = l_{i+2} - \binom{k}{i}$, con $l_{k+2} = l_{k+1} = 0$ y $\varphi_m(t) \approx \sqrt{t} - m/\sqrt{t}$, para $m = 1, 2, \dots$, y $i = 2, \dots, k$.

3.3. Análisis de forma

La figura 4 (izquierda) muestra gráficas para la f.d.p. de la suma de cinco v.a. BS para diferentes valores del parámetro de forma α . Se ve que a medida que el parámetro de forma α crece, la forma de la distribución es más sesgada hacia la derecha. La figura 4 (derecha) muestra gráficas de la f.d.a. BSsum con parámetro de forma $\alpha = 1.5$ y escala $\beta = 1$. Se ve que a medida que se suman más v.a. la forma de la f.d.p. tiende a ser más simétrica.

4. Cartas de control para la distribución BS

En esta sección introducimos dos cartas de control basadas en la distribución BS. La primera de ellas es una carta de control bootstrap para percentiles BS disponible en la literatura y propuesta por Lio & Park (2008). La segunda es una nueva carta propuesta en este artículo.

4.1. Carta de control bootstrap para percentiles BS

Como se mencionó, la carta de control \bar{T} clásica de Shewhart supone que los datos observados en el proceso provienen de una distribución normal. No obstante, cuando la distribución es desconocida o no normal, la distribución muestral del estimador del parámetro de interés para controlar el proceso puede no estar disponible en teoría y así los límites de control no pueden obtenerse. El método bootstrap se puede utilizar para construir los límites para el seguimiento de un determinado percentil de la distribución BS. Si bien el método funciona para cualquier percentil, la preocupación por parte de los ingenieros está especialmente en los percentiles inferiores, ya que un cambio a la baja en un percentil inferior de la distribución puede ser algo más grave que en otro tipo de percentil.

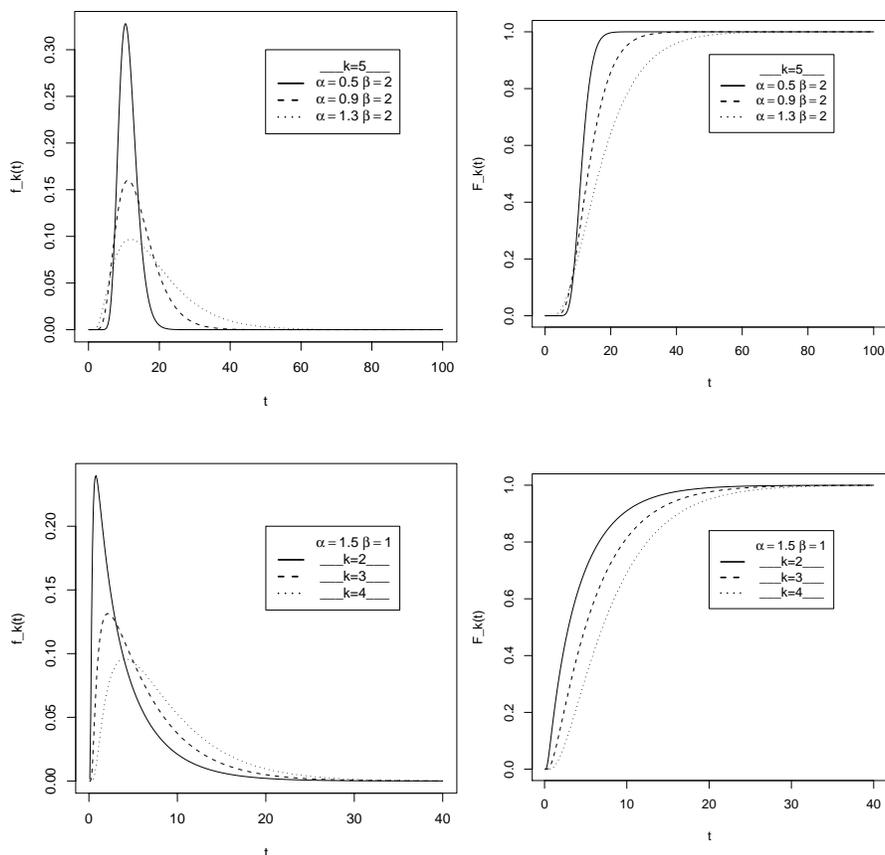


FIGURA 4: f.d.p. (izquierda) y f.d.a. (derecha) de la distribución BSsum para diferentes valores de α y k .

4.1.1. Construcción de la carta de control bootstrap

Para construir una carta de control para percentiles BS asumiendo que el proceso es estable y está bajo control se debe seguir un algoritmo que contiene los siguientes pasos de acuerdo a lo mencionado en Lio & Park (2008):

- (B1) Obtener k muestras aleatorias independientes de tamaño n_j , para $j = 1, \dots, k$, asumiendo que provienen de una distribución BS con parámetros α y β desconocidos, donde las observaciones de la muestra j -ésima se denotan como t_{ij} , para $i = 1, \dots, n_j$.
- (B2) Encontrar las estimaciones de VM de α y β con la muestra combinada de tamaño $N = \sum_{j=1}^k n_j$.
- (B3) Generar observaciones $t_1^*, t_2^*, \dots, t_m^*$ mediante el método bootstrap de tamaño m desde una distribución BS y utilizar las estimaciones obtenidas en el

paso (B2). Aquí, m es el tamaño de muestra, el cual se usa en el futuro como tamaño de los subgrupos.

- (B4) Encontrar las estimaciones de VM de α y β basadas en la muestra generada en el paso (B3) y denotarlas como $\hat{\alpha}^*$ y $\hat{\beta}^*$.
- (B5) Calcular las observaciones bootstrap $\hat{t}_p^* = [\hat{\beta}^*/4][\hat{\alpha}^* z(p) + [\hat{\alpha}^{*2} z(p)^2 + 4]^{1/2}]^2$, donde p representa el percentil que se quiere monitorear y $z(p)$ es el cuantil p -ésimo de la distribución $N(0, 1)$, para la muestra bootstrap del paso (B3) y con las estimaciones obtenidas en el paso (B4).
- (B6) Repetir los pasos (B3), (B4) y (B5) una gran cantidad de veces, obteniendo B (por ejemplo, $B = 10000$) muestras bootstrap de \hat{t}_p , denotada por $\hat{t}_{p1}^*, \dots, \hat{t}_{pB}^*$.
- (B7) Hallar los percentiles $(\gamma/2) \times 100\%$ y $(1 - \gamma/2) \times 100\%$ usando la muestra bootstrap obtenida en paso (B6). Estos percentiles son los LCI y LCS, respectivamente.

Ejemplo 1. Para ilustrar el procedimiento de la carta de control BS bootstrap, considere el monitoreo de la resistencia a la ruptura de un material de aluminio. Para ello se han simulado datos de un proceso productivo de hojas de aluminio 6061-T6. Así, se simularon veinte grupos de tamaño cinco en forma independiente desde un proceso bajo control con distribución BS de parámetros $\alpha = 0.280$ y $\beta = 1.358$. Estos veinte subgrupos se presentan en la tabla 1. A partir de estos datos simulados se obtienen los límites de control bootstrap para el monitoreo del primer percentil de la distribución BS, realizando $B = 10000$ veces los pasos (B2)-(B7). Así, el LCI y el LCS de la carta de control para monitorear el primer percentil de la distribución BS con razón de falsa alarma $\delta = 0.003$ son LCI = 0.920 y LCS = 1.874, respectivamente; ver el libro de Duncan (1996) para más detalles del concepto de falsa alarma. Se supone ahora que el proceso se ha convertido en uno fuera de control. Entonces, se simularon veinte subgrupos de tamaño cinco correspondientes a un modelo BS cuyo parámetro de forma ha cambiado a $\alpha = 0.878$. Estos datos se presentan también en la tabla 1. La figura 5 muestra la carta de control generada mediante el método bootstrap para el monitoreo de la resistencia a la ruptura a través del primer percentil de la distribución BS. Los límites se obtienen a partir de la muestra bajo control. Se grafican los 20 subgrupos que se asumen como fuera de control. Esto es para evidenciar si la carta puede detectar rápidamente una señal fuera de control o no. El límite central (LC) también está incluido en la carta de control. Esto es para mostrar que la distribución muestral del primer percentil de la distribución BS es asimétrica. Se ve claramente que el proceso inmediatamente da señales de estar fuera de control, ya que sólo seis puntos de los veinte subgrupos están dentro de los límites y sólo uno de éstos se encuentra por sobre el límite central.

4.2. Cartas de control \bar{T} para la distribución BS

Como se mencionó, una gran cantidad de datos generados desde procesos productivos siguen distribuciones asimétricas, tal como ocurre con los tiempos de vida. Cartas de control para estas distribuciones son de gran importancia para el

control adecuado de los procesos en cuestión. En la subsección 4.1.1, se mostró la única carta de control que se ha desarrollado para la distribución BS. Los pocos desarrollos de cartas de control para esta distribución se deben básicamente a que esta distribución no posee la propiedad reproductiva. Por tanto, cartas de control para el monitoreo del tiempo acumulado o promedio hasta la ocurrencia de m fallas para esta distribución no han sido implementadas, ya que éstas necesitan de la distribución BSsum.

TABLA 1: datos simulados de tiempos de falla por fractura debido a estrés de piezas de aluminio 6061-T6.

Subgrupo	Datos de un proceso bajo control					Datos de un proceso fuera de control				
1	1.140	1.430	1.076	2.113	1.489	0.280	8.532	0.627	0.767	0.414
2	1.080	1.556	1.669	1.595	1.247	0.519	1.290	1.689	2.051	1.592
3	2.066	1.514	1.142	0.738	1.857	2.344	0.717	1.384	0.807	4.217
4	1.341	1.352	1.767	1.708	1.603	0.416	3.147	0.565	1.083	0.753
5	1.754	1.689	1.387	0.784	1.615	1.511	1.688	0.369	2.564	0.536
6	1.337	1.300	0.902	1.188	1.526	0.219	0.772	1.173	0.807	1.649
7	1.981	1.320	1.513	1.338	0.926	1.436	3.991	0.957	1.344	1.907
8	1.209	1.216	1.336	1.845	1.680	1.726	0.771	5.854	1.631	3.047
9	1.297	1.265	1.650	1.586	1.121	0.988	0.854	7.217	3.632	1.221
10	1.115	1.504	1.683	1.316	1.736	4.562	1.186	4.448	1.575	0.595
11	1.518	1.145	1.494	0.992	2.022	4.474	0.436	2.346	0.444	0.568
12	2.346	1.226	1.015	1.592	1.308	0.546	1.198	2.190	3.836	2.028
13	2.625	1.343	1.646	1.369	1.104	0.435	1.126	5.999	1.145	1.155
14	1.432	0.824	2.040	1.417	2.470	1.779	0.703	1.221	3.087	0.524
15	1.551	1.114	1.610	1.047	0.958	1.946	0.532	2.614	1.399	5.211
16	1.473	1.200	1.358	1.387	1.152	2.201	0.624	4.550	0.886	1.387
17	1.159	1.308	1.885	0.890	1.603	0.565	3.595	1.409	0.303	0.767
18	1.490	1.826	1.247	1.506	1.463	1.531	1.391	1.639	3.103	2.940
19	1.167	1.900	1.876	1.651	2.108	0.532	0.774	0.817	0.537	5.010
20	1.587	0.952	1.157	0.966	1.190	1.436	1.117	1.044	1.038	0.816

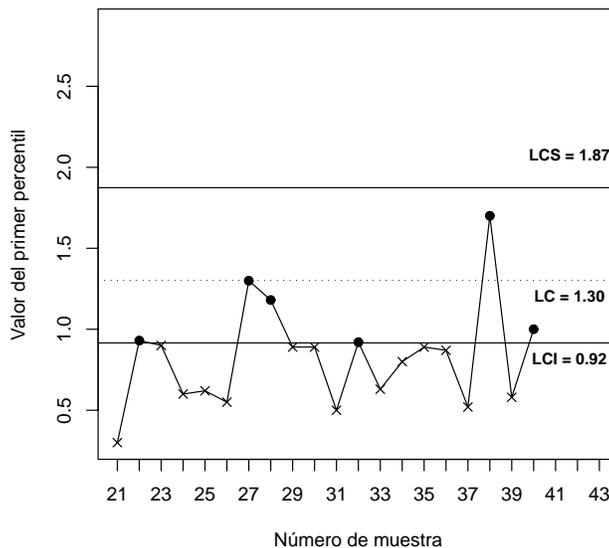


FIGURA 5: Carta de control *bootstrap* BS.

4.2.1. Construcción de la carta de control \bar{T}

Siguiendo la propuesta de Surucu & Sazak (2009), implementamos cartas de control para monitorear la ocurrencia de fallas de un proceso que genera datos provenientes de una distribución BS. La ventaja de la metodología desarrollada aquí frente a la carta de control creada (Surucu & Sazak 2009), donde se tuvo que aproximar la distribución de la suma de v.a. Weibull con tres parámetros a la distribución normal y así aproximar los límites de control, es que al tener implementada la distribución BSSum en forma exacta, los límites de control para la distribución BS también lo son. Lo primero que se debe decidir cuando se quiere monitorear el tiempo de ocurrencia de fallas en un proceso productivo es el número de fallas que se van a monitorear. Este número se denota por m . Una vez tomada esta decisión, para construir una carta de control para la media \bar{T} basada en la distribución BS, asumiendo que el proceso es estable y está bajo control, proponemos un algoritmo con los siguientes pasos:

- (C1) Registrar los tiempos de fallas, denotados por T_i , para $i = 1, 2, \dots, k$, los cuales son v.a. i.i.d. con distribución BS. Estos tiempos deben tomarse desde un proceso que esté bajo control, ya que son ocupados para determinar los límites de la carta de control.
- (C2) Calcular el tiempo acumulado hasta la ocurrencia de la falla m -ésima, generando n muestras de tamaño m , el cual se obtiene mediante la v.a. $Y_i = \sum_{j=m[i-1]+1}^{im} X_j$, para $i = 1, 2, \dots, n$.
- (C3) Estimar los parámetros de la distribución de la suma de m variables aleatorias BS para la muestra obtenida en el paso (C2). Esto se puede realizar mediante la función `estbssum()` del paquete `bssum`; ver detalles en Apéndice 2. Estas estimaciones se denotan por $\hat{\alpha}_{\text{bssum}}$ y $\hat{\beta}_{\text{bssum}}$.
- (C4) Calcular el LCI y el LCS de la carta de control, los que están determinados por los percentiles $(\gamma/2) \times 100\%$ y $(1 - \gamma/2) \times 100\%$ de la distribución BSSum, respectivamente. Estos percentiles se pueden obtener con la función `qbsum()` del paquete `bssum`. Se debe mencionar que en esta carta de control no tiene mucho sentido incorporar el LC de control, ya que éste es más apropiado para datos que presentan simetría. Sin embargo, con el objetivo de detectar si existe aleatoriedad en los datos, se ha construido igualmente un LC basado en la mediana.
- (C5) Finalmente, determinados los límites de control de la carta, se pueden graficar los tiempos acumulados de fallas, para así corroborar que estos tiempos de fallas están bajo control.

Ahora, estamos en condiciones de monitorear cualquier conjunto de datos que siga una distribución BS para los que se desee saber su estado de control o fuera de control en relación al tiempo acumulado hasta la ocurrencia de m fallas. Si los tiempos graficados están dentro de los límites de control y no presentan un comportamiento anómalo, se juzga que el proceso está bajo control. Por el contrario, si los tiempos están por debajo del LCI, se concluye que las fallas del proceso están

ocurriendo con demasiada frecuencia y se juzga que el proceso está fuera de control. Ahora bien, si los tiempos están por sobre el LCS, se concluye que las fallas están ocurriendo lentamente, lo cual es un buen indicador del proceso productivo.

4.3. Implementación

Aquí ilustramos mediante un ejemplo la metodología de construcción de cartas de control desarrolladas para la distribución BS. Específicamente, presentamos una carta de control para el monitoreo de la ocurrencia de fallas de un proceso que genera datos provenientes desde una distribución BS. Esta carta de control está implementada en el paquete `bssum` e incorpora todos los pasos de la metodología de construcción en una única función.

Ejemplo 2. Para ilustrar el procedimiento de construcción de la carta de control para la ocurrencia de fallas de un proceso que genera datos provenientes desde una distribución BS, se usan los datos presentados en Cheng & Xie (2000) empleados allí para ilustrar cartas de control para la distribución lognormal. Estos datos se muestran en la tabla 2, representan los tiempos de falla de válvulas en un proceso industrial y se encuentran implementados en el paquete `bssum` con el nombre de `valve`.

TABLA 2: Tiempos de falla (en horas) de la válvula i -ésima en la muestra j -ésima indicada.

j	t_{ij}					j	t_{ij}				
1	4.55	4.99	3.62	3.52	3.77	16	2.53	4.16	3.78	3.77	1.72
2	1.93	3.95	4.10	4.16	1.61	17	3.41	3.10	6.02	1.09	2.92
3	2.22	1.73	5.10	4.52	4.06	18	2.85	4.46	3.17	2.50	3.91
4	2.71	2.45	4.60	2.09	1.90	19	3.16	3.70	2.61	2.65	3.42
5	2.91	5.68	4.33	3.51	3.24	20	2.54	4.77	1.63	2.64	3.59
6	2.20	5.66	3.71	3.35	1.61	21	3.61	2.13	5.08	2.01	1.92
7	2.82	5.22	3.75	3.50	3.31	22	3.16	4.20	2.32	2.44	1.62
8	2.76	4.40	3.13	1.55	3.70	23	2.96	6.09	3.78	2.29	4.16
9	4.98	4.05	4.00	7.20	3.18	24	2.47	3.49	3.38	4.45	2.61
10	4.88	2.71	3.51	3.15	4.81	25	3.55	3.35	3.18	4.75	8.72
11	4.50	1.95	3.41	2.87	1.90	26	1.35	2.50	2.51	4.20	3.50
12	3.07	4.02	4.17	4.33	4.06	27	2.30	2.26	2.22	1.60	9.70
13	2.39	2.91	3.09	3.15	2.52	28	3.71	3.06	1.53	2.45	6.40
14	2.92	4.25	3.02	2.26	5.72	29	9.48	1.72	4.20	3.37	5.58
15	2.56	4.38	1.24	2.62	1.92	30	1.90	2.56	4.28	3.18	1.94

Para implementar cartas de control para la distribución BS para el proceso de producción de válvulas, se debe comprobar que los datos de la tabla 2 provienen de una distribución BS. Para esto se utiliza la función `ksbs()` de los paquetes `bs` 1.0 y `bs` 2.0; ver detalles en Apéndice 1. A continuación se presentan los resultados de la

aplicación de la instrucción `data(valve)` en el paquete `bssum` y de la instrucción `ksbs(valve, graph = FALSE, alternative = two.sided)` en el paquete `bs`:

```
One-sample Kolmogorov-Smirnov test
Data: valve
D = 0.1038
p-value = 0.8702
Alternative hypothesis: two-sided
```

Para apoyar la decisión obtenida mediante la prueba de Kolmogorov-Smirnov (KS) que indica que prácticamente no hay evidencia ($\text{valor-}p = 0.8702$) como para indicar que los datos no siguen una distribución BS, producimos la figura 6, la que respalda gráficamente lo indicado por la prueba KS.

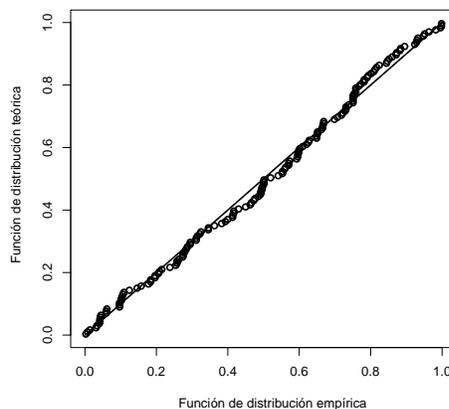


FIGURA 6: PP para los tiempos de fallas de las válvulas.

Con esta buena relación entre la distribución BS y los datos `valve`, podemos implementar una carta de control BS que permite monitorear la ocurrencia de fallas del proceso productivo de válvulas. El monitoreo de la producción de válvulas se realiza hasta que se acumulan cinco fallas. Una vez determinado esto, se construye la variable Y usando el paso (C2). En la tabla 3 se muestran estos tiempos.

TABLA 3: Tiempo acumulado hasta la quinta falla, Y_i , para $i = 1, 2, \dots, 30$.

1	2	3	4	5	6	7	8	9	10
20.45	15.75	17.63	13.75	19.67	16.53	18.60	15.54	23.41	19.06
11	12	13	14	15	16	17	18	19	20
14.63	19.65	14.06	18.17	12.72	15.96	16.54	16.59	15.54	15.17
21	22	23	24	25	26	27	28	29	30
14.75	13.74	19.28	16.40	23.55	14.06	18.08	17.15	24.35	13.86

A continuación, para la muestra generada de tiempos acumulados Y , se estiman los parámetros de la distribución de la suma de m v.a. independientes BS. Esto se realiza con la función `estbssum()` del paquete `bssum`. La instrucción

`estbssum(5, valve)` arroja las estimaciones de α y β de la suma de m v.a. BS independientes, que son $\hat{\alpha} = 0.358$ y $\hat{\beta} = 3.306$. Una vez estimados los parámetros, se deben calcular los límites de la carta de control que monitorea la ocurrencia de fallas en un proceso productivo. Como se menciona en el paso (C4), los límites están determinados por los percentiles $(\gamma/2) \times 100\%$ y $(1 - \gamma/2) \times 100\%$ de la distribución `BSSum`, respectivamente. A continuación se muestra la instrucción que calcula estos percentiles en el paquete `bssum`, que en definitiva son respectivamente el LCI y el LCS de la carta BS:

```
> qbssum(q = 0.0027/2, k = 5, alpha = 0.358, beta = 3.306,
+       lower.tail = TRUE)
[1] 10.720
> qbssum(q = 1 - 0.0027/2, k = 5, alpha = 0.358, beta = 3.306,
+       lower.tail = TRUE)
[1] 28.024
```

El conjunto de datos de la tabla 3 se usa para construir los límites de control de la carta, los que son graficados para comprobar que el proceso está bajo control en relación a los tiempos de fallas. Sin embargo, se necesita un conjunto de datos simulados con un cambio en el proceso con la idea de ver si la carta de control es capaz de detectar dicho cambio en los tiempos de fallas. Para esto, se simularon cincuenta (50) observaciones que representan el tiempo de falla de cincuenta (50) válvulas desde un proceso fuera de control. Estos datos simulados están cargados en el paquete `bssum` con el nombre de `valve1`. A continuación se muestran los tiempos acumulados hasta la ocurrencia de cinco (5) fallas en el proceso productivo, a partir de estos datos simulados. Estos datos se generan usando paso (C2) y se muestran en la tabla 4.

TABLA 4: Tiempo acumulado hasta la quinta falla, y_i , para $i = 1, \dots, 10$.

1	2	3	4	5	6	7	8	9	10
11.35	13.13	13.77	11.85	11.98	11.56	12.29	12.27	9.99	12.46

Con la idea de construir la carta de control para la distribución BS, se implementó en el paquete `bssum` la instrucción `chartControlT()`, la que incorpora los pasos (C1)-(C5). Se debe tener en cuenta la cantidad de fallas que son monitoreadas, los tiempos de fallas que se utilizan para obtener los límites de control (`valve`), los datos a monitorear (`valve 1`) y la razón de falsa alarma. La instrucción `chartControlT(5, valve, valve1, 0.0027)` genera la carta de control graficada en la figura 7 (izquierda). Las primeras 30 muestras corresponden a los datos con los cuales se obtuvieron los límites de control. Las últimas 10 muestras corresponden a los datos simulados con un cambio en el proceso. Se ve claramente que la carta de control para el monitoreo de tiempos de falla detecta rápidamente el cambio en el proceso.

A medida que el valor del parámetro α disminuye, la forma de la distribución BS tiende a ser simétrica. En estos casos sería de gran ayuda tener también implementada una carta de control basada en la distribución BS para \bar{T} . Así,

primero se debe verificar que los tiempos que se quieren monitorear se distribuyan BS y que tengan un comportamiento tendiente a la simetría. En el paquete `bssum`, se implementó también una carta de control basada en la distribución BS para \bar{T} . Su uso es similar al de la función `chartControlT()`. La instrucción `chartControlMean(5, valve, valve1, 0.0027)` genera la carta de control de la figura 7 (derecha).

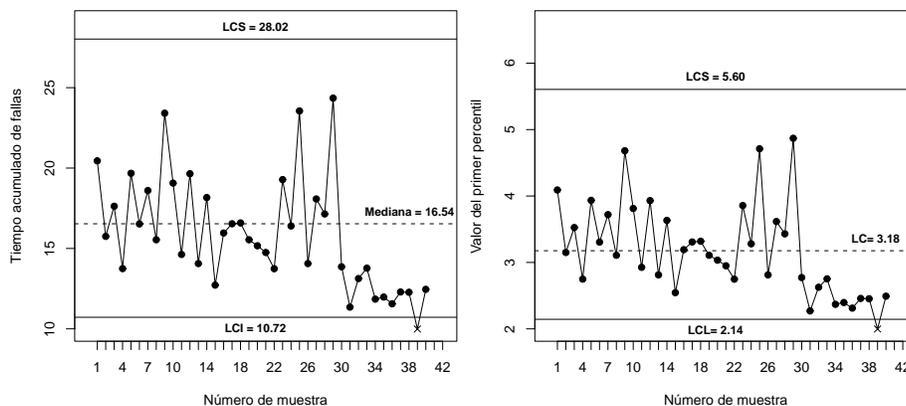


FIGURA 7: Cartas de control para tiempos de fallas acumulado (izquierda) y \bar{T} (derecha) de la distribución BS.

5. Conclusiones

Debido a que muchos procesos industriales de productos expuestos a fallas presentan datos con un comportamiento asimétrico, cartas de control basadas en la Birnbaum-Saunders deberían estar disponibles. A la fecha existe poca literatura disponible en esta dirección, salvo el estudio desarrollado por Lio & Park (2008) para monitorear percentiles de esta distribución. En el presente trabajo, hemos desarrollado una nueva metodología para cartas de control basadas en la distribución Birnbaum-Saunders que permite monitorear la media de procesos de producción de unidades expuestas a fallas y el tiempo acumulado de estas fallas. Ya que últimamente el modelo Birnbaum-Saunders se ha considerado más ampliamente como una distribución de probabilidades y no sólo como una distribución de vida, la metodología de cartas de control desarrollada en este trabajo puede ser también aplicada a otros tipos de procesos y de fenómenos. Para producir esta metodología, introducimos e implementamos la distribución de la suma de variables aleatorias independientes Birnbaum-Saunders. Esta suma no sigue una distribución Birnbaum-Saunders, lo que fue probado en un estudio de confiabilidad conducido por Raaijmakers (1980) en donde se obtuvo la distribución de vida de un sistema compuesto por unidades en standby. Cada una de estas unidades tiene tiempos de vida independientes que siguen una distribución Birnbaum-Saunders. Por tanto, la distribución de la vida útil del sistema corresponde a la distribución de la suma de variables aleatorias independientes Birnbaum-Saunders. La imple-

mentación de esta nueva distribución se desarrolló en el software R en el paquete `bssum`. Este paquete contiene funciones de probabilidad, un generador de números aleatorios y un método de estimación de parámetros para la distribución de la suma de variables aleatorias independientes Birnbaum-Saunders. Además, agregamos en este paquete dos funciones gráficas que generan las cartas de control para datos modelados mediante la distribución Birnbaum-Saunders. Estas cartas permiten monitorear el tiempo acumulado o el tiempo promedio hasta la ocurrencia de cierta cantidad de fallas. El paquete `bssum` es un primer acercamiento a la implementación de esta distribución en un paquete computacional estadístico. Éste puede usarse como base para una implementación más completa de esta distribución tal como ocurre con los paquetes `bs` y `gbs`. Otro desafío sobre el que los autores están trabajando es una metodología como la desarrollada en este artículo para datos censurados.

Agradecimientos

Los autores agradecen a los editores Dr. Leonardo Trujillo y Dra. Piedad Urdinola y a dos referees anónimos por sus valiosos comentarios que permitieron mejorar la versión preliminar de este artículo. Este estudio contó con el apoyo del proyecto FONDECYT 1080326 del gobierno de Chile.

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

Referencias

- Balakrishnan, N., Leiva, V. & López, J. (2007), ‘Acceptance sampling plans from truncated life tests based on the generalized Birnbaum-Saunders distribution’, *Communications in Statistics: Simulation and Computation* **36**, 643–656.
- Balakrishnan, N., Leiva, V., Sanhueza, A. & Cabrera, E. (2009), ‘Mixture inverse Gaussian distribution and its transformations, moments and applications’, *Statistics* **43**, 91–104.
- Barros, M., Paula, G. A. & Leiva, V. (2008), ‘A new class of survival regression models with heavy-tailed errors: robustness and diagnostics’, *Lifetime Data Analysis* **14**, 316–332.
- Barros, M., Paula, G. A. & Leiva, V. (2009), ‘An R implementation for generalized Birnbaum-Saunders distributions’, *Computational Statistics and Data Analysis* **53**, 1511–1528.
- Bhatti, C. R. (2010), ‘The Birnbaum-Saunders autoregressive conditional duration model’, *Mathematics and Computers in Simulation* **80**, 2062–2078.
- Birnbaum, Z. W. & Saunders, S. C. (1969a), ‘Estimation for a family of life distributions with applications to fatigue’, *Journal of Applied Probability* **6**, 328–347.

- Birnbaum, Z. W. & Saunders, S. C. (1969b), 'A new family of life distributions', *Journal of Applied Probability* **6**, 319–327.
- Chang, D. S. & Tang, L. C. (1994), 'Graphical analysis for Birnbaum-Saunders distribution', *Microelectronics and Reliability* **34**, 17–22.
- Cheng, S. W. & Xie, H. (2000), 'Control charts for lognormal data', *Tamkang Journal of Science and Engineering* **3**, 131–137.
- Desmond, A. (1985), 'Stochastic models of failure in random environments', *Canadian Journal of Statistics* **13**, 171–183.
- Díaz-García, J. A. & Leiva, V. (2005), 'A new family of life distributions based on elliptically contoured distributions', *Journal of Statistical Planning and Inference* **128**, 445–457.
- Duncan, A. (1996), *Control de Calidad y Estadística Industrial*, Alfaomega Grupo Editor, México.
- Johnson, N. L., Kotz, S. & Balakrishnan, N. (1995), *Continuous Univariate Distributions*, Vol. 2, second edn, John Wiley & Sons, New York.
- Lawless, J. F. (1982), *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, New York.
- Leiva, V., Barros, M., Paula, G. A. & Galea, M. (2007), 'Influence diagnostics in log-Birnbaum-Saunders regression models with censored data', *Computational Statistics and Data Analysis* **51**, 5694–5707.
- Leiva, V., Barros, M., Paula, G. & Sanhueza, D. (2008), 'Generalized Birnbaum-Saunders distributions applied to air pollutant concentration', *Environmetrics* **19**, 235–249.
- Leiva, V., Hernández, H. & Riquelme, M. (2006), 'A new package for the birnbaum-saunders distribution', *R Journal* **6**, 35–40.
*http://www.R-project.org/doc/Rnews/Rnews_2006-4.pdf
- Leiva, V., Sanhueza, A. & Angulo, J. M. (2009), 'A length-biased version of the Birnbaum-Saunders distribution with application in water quality', *Stochastic Environmental Research and Risk Assessment* **23**, 299–307.
- Leiva, V., Sanhueza, A. & Saunders, S. C. (2009), New developments and applications on life distributions under cumulative damage, Technical Report 4, CIMFAV.
*<http://www.cimfav.cl/reports.html#2009>
- Leiva, V., Sanhueza, A., Sen, P. K. & Paula, G. A. (2008), 'Random number generators for the generalized Birnbaum-Saunders distribution', *Journal of Statistical Computation and Simulation* **78**, 1105–1118.

- Leiva, V., Vilca, F., Balakrishnan, N. & Sanhueza, A. (2010), 'A skewed sinh-normal distribution and its properties and application to air pollution', *Communications in Statistics: Theory and Methods* **39**, 426–443.
- Lio, Y. L. & Park, C. (2008), 'A bootstrap control chart for Birnbaum-Saunders percentiles', *Quality and Reliability Engineering International* **24**, 85–600.
- Marshall, A. W. & Olkin, I. (2007), *Life Distributions*, Springer, New York.
- Ng, H. K. T., Kundu, D. & Balakrishnan, N. (2003), 'Modified moment estimation for the two-parameter Birnbaum-Saunders distribution', *Computational Statistics and Data Analysis* **43**, 283–298.
- Pérez, R. A. & Correa, J. C. (2008), 'Intervalos de confianza vía verosimilitud relativa de los parámetros de la distribución Birnbaum-Saunders', *Revista Colombiana de Estadística* **37**, 645–670.
- Podlaski, R. (2008), 'Characterization of diameter distribution data in near-natural forests using the Birnbaum-Saunders distribution', *Canadian Journal of Forest Research* **18**, 518–526.
- R Development Core Team (2009), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria.
*<http://www.R-project.org>
- Raaijmakers, F. J. M. (1980), 'The lifetime of a standby system of units having the Birnbaum-Saunders distribution', *Journal of Applied Probability* **17**, 490–497.
- Raaijmakers, F. J. M. (1981), 'Reliability of standby system for units with the Birnbaum-Saunders distribution', *IEEE Transactions on Reliability* **30**, 198–199.
- Sanhueza, A., Leiva, V. & Balakrishnan, N. (2008), 'The generalized Birnbaum-Saunders distribution and its theory, methodology and application', *Communications in Statistics: Theory and Methods* **37**, 645–670.
- Saunders, S. C. (2007), *Reliability Life Testing and Prediction of Service Lives*, Springer, New York.
- Schoonhoven, M. & Does, R. J. M. M. (2010), 'The control chart under non-normality', *Quality and Reliability Engineering International* **26**, 167–176.
- Steiner, H. S. & Mackay, J. R. (2000), 'Monitoring processes with highly censored data', *Journal of Quality Technology* **32**, 199–208.
- Surucu, B. & Sazak, H. (2009), 'Monitoring reliability for a three-parameter Weibull distribution', *Reliability Engineering and System Safety* **94**, 503–508.
- Vargas, J. & Montañó, T. (2005), 'Carta de control CEV \bar{X} para distribuciones Weibull con datos censurados', *Revista Colombiana de Estadística* **28**, 125–139.

Vilca, F., Sanhueza, A., Leiva, V. & Christakos, G. (2010), 'An extended Birnbaum-Saunders model and its application in the study of environmental quality in Santiago, Chile', *Stochastic Environmental Research and Risk Assessment* **24**, 771–782.

Zhang, L. & Chen, G. (2004), 'EWMA charts for monitoring the mean of censored Weibull lifetimes', *Journal of Quality Technology* **36**, 321–328.

Apéndice A.

Implementación de la distribución BS

Aquí se discuten dos implementaciones para la distribución BS en el software R. La primera de éstas se encuentra en el paquete `bs` 1.0 elaborado por Leiva et al. (2006), el que incorpora funciones de probabilidad, de análisis de confiabilidad, de estimación de parámetros y de bondad de ajuste. El segundo es el paquete `bs` 2.0, una implementación elaborada recientemente que es una versión más completa del paquete `bs` 1.0. A continuación se detallan algunas de las funciones principales de ambos paquetes.

Funciones básicas

Para calcular la f.d.p., la f.d.a. y la f.q. de la distribución BS con parámetros α y β se usan las funciones `dfs()`, `pbs()` y `qbs()`, respectivamente. En la tabla 5 se encuentran las instrucciones que ilustran estas funciones.

TABLA 5: Funciones de probabilidad básicas BS.

Función	Instrucción	Resultado
f.d.p.	<code>dfs(3.0, alpha = 0.5, beta = 1.0, log = FALSE)</code>	0.021
f.d.a.	<code>pbs(1.0, alpha = 0.5, beta = 1.0, log = FALSE)</code>	0.500
f.q.	<code>qbs(0.5, alpha = 0.5, beta = 1.0, log = FALSE)</code>	1.000

Para generar números aleatorios desde el modelo BS se puede usar la f.q. dada en (5), la que requiere de un generador de números desde la distribución normal estándar. El paquete `bs` 1.0 tiene incorporado tres métodos de generación de números aleatorios BS y una función que selecciona automáticamente el método más apropiado (entre los tres existentes). El paquete `bs` 2.0 tiene disponible sólo uno de los tres generadores de la versión 1.0, mediante la función `rbs()`, el que es más eficiente y está basado en la f.q. BS. La siguiente instrucción ilustra este comando:

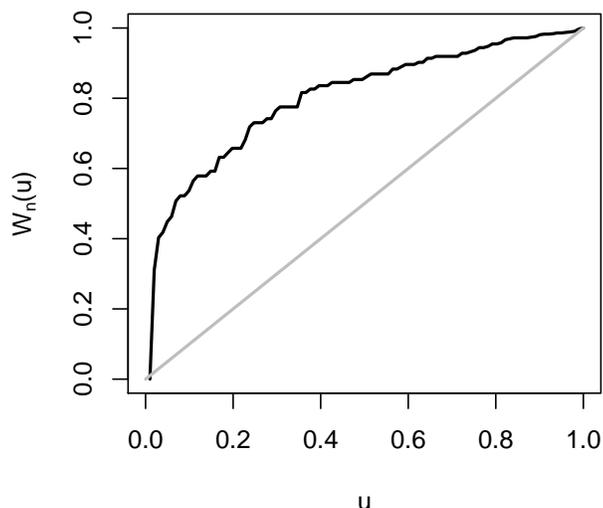
```
> rbs(n = 6, alpha = 0.5, beta = 2.5)
[1] 1.876529 2.785264 2.388317 2.898078 1.903702 1.366595
```

Para un análisis de confiabilidad basado en la distribución BS se ha implementado, entre otras funciones, la f.c. y la f.r. Las funciones mencionadas se muestran en la tabla 6.

TABLA 6: Funciones de confiabilidad BS.

Función	Instrucción	Resultado
f.r.	<code>frbs(3.0, alpha = 0.5, beta = 1.0)</code>	2.040
f.c.	<code>rfbs(1.0, alpha = 0.5, beta = 1.0)</code>	0.500

El paquete `bs` 2.0 tiene implementado el gráfico TTT mediante la función `TTT()`. Adicionalmente, esta función permite agregar en el gráfico una línea recta que representa a la distribución exponencial, tal como ocurre con los gráficos de probabilidad, usando la instrucción `explLine = TRUE`. Un ejemplo del uso de esta función para el conjunto de datos `psi31`, los cuales quedan disponible cuando se aplica la instrucción `data(psi31)`, es `TTT(psi31, explLine = TRUE)`, lo cual arroja el gráfico de la figura 8.

FIGURA 8: TTT para `psi31`.

Estimación de parámetros

Con el fin de estimar los parámetros de forma y de escala de la distribución BS, α y β , respectivamente, se han implementado los tres métodos descritos en la sección 2.2, los que están disponibles en ambos paquetes, tanto `bs` 1.0 como `bs` 2.0. A continuación se presenta un ejemplo relacionado al uso de estas funciones.

Ejemplo 3. Birnbaum & Saunders (1969a) introdujeron un conjunto de datos de fatiga de materiales de $n = 101$ hojas de aluminio del tipo 6061-T6. Estas hojas fueron cortadas en un ángulo paralelo al sentido de rotación y fueron expuestas a una presión con un máximo de tensión de 31.000 psi (del inglés “pounds per square inch” correspondiente a libras por pulgada cuadrada). Todas las hojas fueron probadas hasta que fallaran, así que no existen datos censurados. La instrucción

`est1bs(psi31)` calcula las estimaciones de VM de α y β , utilizando el estimador “media-media” como valor inicial para el procedimiento numérico, obteniendo

```
$beta.start
[1] 131.8193
$alpha
[1] 0.1703
$beta
[1] 131.8188
$converge
[1] "TRUE"
$iteration
[1] 2
```

Note que para ocupar la función `est1bs(psi31)` es necesario ejecutar previamente el comando `data(psi31)`. Las estimaciones de α y β pueden guardarse en una variable R como `estimate <- est1bs(psi31)`, de modo que las estimaciones de α y β pueden obtenerse mediante las instrucciones `alpha <- estimate$alpha` y `beta <- estimate$beta`. Además, por la propiedad de invarianza de los estimadores de VM, se pueden obtener estimaciones de la media, la varianza y los CV, CS y CC de la distribución BS mediante la instrucción `indicatorsbs(psi31)` obteniendo

```
The MLE's are:
Alpha = 0.1703
Beta = 131.8188
Mean = 143.0487
Variance = 522.7530
Coefficient of variation = 0.1710
Coefficient of skewness = 0.5103
Coefficient of kurtosis = 3.4329
```

Funciones de bondad de ajuste

Para determinar si un conjunto de datos proviene desde una distribución BS, se pueden usar diferentes métodos de bondad de ajuste. La distribución de los datos se puede juzgar mediante un histograma. Sin embargo, la información que nos arroja el histograma no es concluyente. Métodos de bondad de ajuste más precisos que están implementados en los paquetes `bs 1.0` y `bs 2.0` son los gráficos de probabilidad versus probabilidad (PP) y cuantil versus cuantil (QQ) y la prueba KS. Por tanto, un análisis que determine si un conjunto de datos proviene desde una distribución BS puede realizarse usando las funciones mencionadas anteriormente, las cuales son el histograma, la prueba KS y los gráficos PP y QQ.

Para ilustrar las funciones de bondad de ajuste implementadas en los paquetes `bs 1.0` y `bs 2.0`, se utilizan los mismos datos del ejemplo 3, es decir, los datos `psi31`. La función `histbs()` genera un histograma de los datos. Esta función, si es que el usuario lo determinara necesario, puede generar simultáneamente un gráfico de caja (box-plot) que permite determinar la simetría de los datos y si alguno de

éstos es atípico. Además, usando nuevamente la propiedad de invarianza de los estimadores de VM, se puede agregar a este histograma la f.d.p. estimada de la distribución BS para este conjunto de datos usando las estimaciones de VM de α y β . La instrucción `histbs(psi31, boxPlot = TRUE, densityLine = TRUE)` genera la figura 9 (izquierda). Continuando con el análisis distribucional de los datos, la instrucción

```
ksbs(psi31, graph = FALSE, alternative = "two.sided")
```

realiza la prueba KS para la distribución BS basado en `psi31` obteniendo

```
One-sample Kolmogorov-Smirnov test
Data: psi31
D = 0.085
p-value = 0.4594
Alternative hypothesis: two-sided
```

Además, la instrucción `ppbs(psi31, line = TRUE)` proporciona el gráfico PP basado en el modelo BS mostrado en la figura 9 (centro). Finalmente, se presenta una función que sólo está en la versión avanzada del paquete `bs` 1.0, esto es, la versión `bs` 2.0. La función `MdBS()` puede determinar si un dato es atípico para la distribución BS basada en la transformación dada en (1). Esta transformación permite aproximar la distancia de Mahalanobis usando las estimaciones de los parámetros de la distribución BS. Así, se puede calcular esta distancia para cada uno de los datos, graficarla y estableciendo un punto de corte, que es el percentil 95 de la distribución $\chi^2(1)$, permitiendo establecer si el dato es atípico o no. Para los datos `psi31`, la instrucción `MdBS(psi31)` permite graficar su distancia de Mahalanobis dada en la figura 9 (derecha) identificando cuatro datos atípicos.

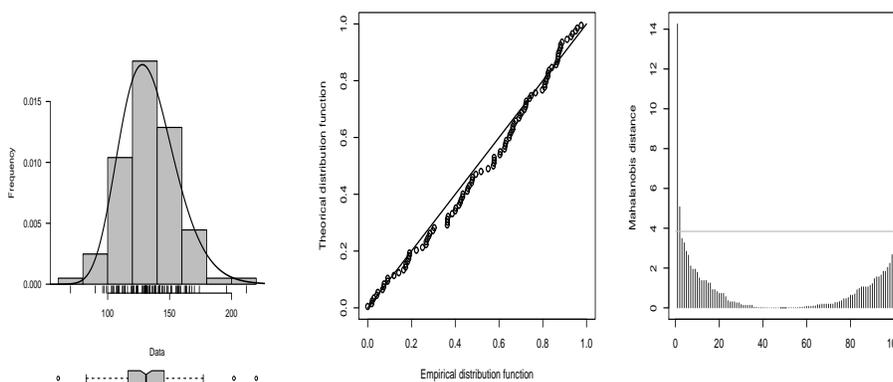


FIGURA 9: Histograma, boxplot y f.d.p. (izquierda), gráfico PP (centro) y distancia de Mahalanobis (derecha) para `psi31`.

Apéndice B.

Implementación de la distribución BSsum

Aquí, discutimos la implementación del paquete R `bssum`. Se han implementado funciones de probabilidad, un generador de números aleatorios y un método de estimación de parámetros para esta distribución. El paquete `bssum` se ha desarrollado con la finalidad de implementar una metodología de cartas de control para la media basadas en la distribución BS. La implementación está desarrollada para k variables aleatorias BS desde $k = 1, \dots, 24$, ya que más allá de este valor, se podría aplicar el teorema del límite central pudiendo obtener cartas de control usando la aproximación normal.

Funciones básicas

Para calcular la f.d.p., la f.d.a. y la f.q. de la distribución BSsum con parámetros α y β se usan las funciones `dbssum()`, `pbssum()` y `qbssum()`, respectivamente. En la tabla 7 se encuentran las instrucciones que ilustran estas funciones.

TABLA 7: Funciones de probabilidad básicas BSsum.

Función	Instrucción	Resultado
f.d.p.	<code>dbssum(2.0, k = 4, alpha = 1.0, beta = 1.0, log = F)</code>	0.062
f.d.a.	<code>pbssum(7.0, k = 4, alpha = 1.0, beta = 1.0, log = F)</code>	0.699
f.q.	<code>qbssum(0.6, k = 6, alpha = 1.0, beta = 1.0, log.q = F)</code>	9.305

Un punto importante en la implementación de la distribución BSsum es la generación de números aleatorios, ya que esto ayuda a analizar diferentes formas de la distribución y es también útil para estudios de simulación. A continuación se ilustra la función `rbssum()`:

```
> rbssum(n = 6, k = 8, alpha = 1.0, beta = 1.0)
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
[1,] 11.36413 8.061949 8.519679 8.749674 16.54859 8.264141
```

La figura 10 muestra el histograma de cuatro conjuntos de números aleatorios. Se ve que al aumentar el número de v.a. que se suman en la distribución, los valores de t generados aumentan y la forma de la distribución es más simétrica.

Estimación de parámetros

Los parámetros de la distribución BSsum son k , α y β . Esta distribución se considera una distribución en sí misma, olvidando que viene de una suma de k v.a., tal como ocurre con la suma de v.a. independientes exponenciales que producen una distribución gamma con parámetro de forma entero (también conocida como distribución Erlang). Por este motivo, el parámetro k debe estimarse. Sin embargo, como la idea es usar esta distribución para implementar cartas de control para la distribución BS, el valor de k se asume conocido. Esto se debe a que cuando se

monitorea el tiempo de ocurrencia de fallas, la cantidad de fallas se fija de antemano. Por tanto, los par metros que se deben estimar en este caso son s lo α y β . El m todo de estimaci n ocupado es el de VM. La funci n `optim()` permite optimizar la funci n de verosimilitud encontrando los valores de α y β que maximizan esta funci n. Como se mencion  en la subsecci n 4.2.1, el comando implementado para esta estimaci n se denomina `estbssum()` y arroja las estimaciones de VM de los par metros de la distribuci n de la suma de k v.a. BS. A continuaci n se proporciona un ejemplo para esta funci n usando los datos `valve` que se llaman mediante la instrucci n `data(valve)`. El n mero de t rminos que se suman en la distribuci n es $k = 5$ y las estimaciones son:

```
> estbssum(k = 5, valve)
$alpha
[1] 0.3581214
$beta
[1] 3.306289
```

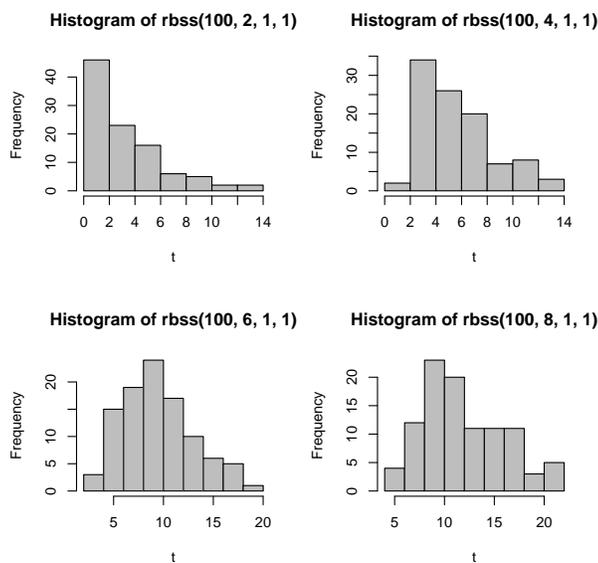


FIGURA 10: Histogramas de cuatro conjuntos de n meros aleatorios BSsum.