



**MODELOS DE DISTRIBUCIONES
OBTENIDAS DE LA DISTRIBUCIÓN
NORMAL**

María F. Martínez de Paz



MODELOS DE DISTRIBUCIONES OBTENIDAS DE LA DISTRIBUCIÓN NORMAL

María F. Martínez de Paz

Memoria presentada como parte de los requisitos
para la obtención del título de Grado en Matemá-
ticas por la Universidad de Sevilla.

Tutorizada por

Inmaculada Barranco Chamorro

Índice general

Resumen	1
English Abstract	3
1. Herramientas básicas	5
1.1. Conceptos básicos en distribuciones de tiempos de vida	5
1.2. Inferencia en muestras completas	9
1.3. Inferencia en muestras censuradas	12
1.4. Distribución Normal	14
2. Distribución Log Normal	21
2.1. Propiedades básicas	22
2.2. Inferencia de la distribución Log Normal	25
2.3. Aplicaciones	29
3. Distribución Inversa Gaussiana	37
3.1. Propiedades básicas	39
3.2. Inferencia de la distribución Inversa Gaussiana	44
3.3. Aplicación	47

4. Distribución Birnbaum-Saunders	51
4.1. Introducción	51
4.2. Interpretación física	52
4.3. Propiedades básicas	53
4.3.1. Características de las distribuciones de tiempos de vida	56
4.4. Momentos y momentos inversos	58
4.5. Relación con la distribución Inversa Gaussiana	63
4.6. Inferencia	64
4.6.1. Método de los momentos	64
4.6.2. Método de los momentos modificado	65
4.6.3. Método de máxima verosimilitud	66
4.6.4. Máxima verosimilitud para muestras censuradas	68
4.7. Aplicaciones	69
Bibliografía	77

Resumen

El objetivo de este trabajo es el estudio de modelos continuos de tiempo de vida relacionados con la distribución Normal. Específicamente, nos centramos en las distribuciones Log-Normal, Inversa Gaussiana y Birnbaum-Saunders. Todos ellos son modelos biparamétricos en que se darán propiedades que permiten estudiar la forma de la distribución, obtener características de interés en teoría de fiabilidad y análisis de supervivencia, y realizar inferencia. Se ha estudiado la estimación de los parámetros del modelo, tanto para datos completos como para muestras censuradas. Cabe destacar que en la mayoría de las ocasiones, para el proceso de estimación habrá que recurrir a técnicas numéricas. El estudio teórico de cada modelo se ha completado con aplicaciones a datos reales (completos y censurados) utilizando el software RStudio.

La memoria se ha estructurado en cuatro capítulos. En el Capítulo 1 se introducen herramientas básicas en modelos de tiempo de vida, como son la función de supervivencia, función de riesgo o hazard, y cuantiles. Para la estimación de parámetros se propone el método de máxima verosimilitud, introduciendo las modificaciones adecuadas para tratar con muestras censuradas. El capítulo se completa con una recopilación de resultados en la distribución normal que se usarán posteriormente.

El Capítulo 2 se dedica a la distribución Log-Normal. Se obtienen su función de supervivencia, función de hazard y momentos, los cuáles nos serán útiles para estudiar la forma de la distribución y propiedades de los estimadores. Destacar que en la estimación para datos completos se tienen fórmulas cerradas, y para datos censurados se han de aplicar métodos numéricos (paquete survival de R).

En el Capítulo 3 se estudia la distribución Inversa Gaussiana siguiendo la parametrización propuesta por Tweedie. Se recogen propiedades, función generatriz de momentos y de cumulantes y resultados sobre la interpretación de los parámetros. Cabe destacar que al realizar inferencia para muestras completas se tiene un resultado análogo al Teorema de Fisher para poblaciones normales.

Finalmente, en el Capítulo 4 se realizará un estudio más extenso de la distribución Birnbaum-Saunders, debido a su importante base física entre otras propiedades interesantes. Veremos cómo surge la distribución a partir de un proceso de fatiga cíclica, y las propiedades fundamentales de la misma. Asimismo, se considerará la notable relación que guarda con la distribución Inversa Gaussiana, además de resultados de inferencia mediante diferentes métodos que se verán reflejados en las aplicaciones a datos reales, las cuales darán por concluido este trabajo.

English Abstract

The aim of this dissertation is to study continuous lifetime models related to the Normal distribution. Specifically, we will focus on the Log-Normal, Gaussian Inverse and Birnbaum Saunders distribution. All of them are two-parameter models in which properties will be given that allow us to study the shape of the distribution, to obtain summaries of interest in Reliability Theory and Survival Analysis, and to carry out Inference. The issue of estimation of parameters in the model is studied for complete data and for censored data. It is important to note that, quite often, numerical techniques are necessary in the process of estimation. Applications to real dataset (complete and censored data) ha been carried out by using the software RStudio.

The document is divided into four chapters. In Chapter 1, the basic tools in lifetime models are introduced. These are: survival function, hazard function and quantiles. Estimation of parameters is carried out by using maximum likelihood. Specific results for censored samples are given. To conclude the chapter, a section is included dealing with relevant results in the normal distribution, which will be used later.

Chapter 2 is devoted to Log-Normal distribution. Its survival function, hazard function and moments are given. Moments will be useful to study the shape of the distribution and to obtain properties of estimators. As for the estimation of parameters, we highlight that, in the case of complete data we have closed expressions. On the other hand, for censored data numerical techniques must be applied (R survival package).

In Chapter 3, the Gaussian Inverse distribution is studied with the parameterization proposed by Tweedie. Properties, moment and cumulant generating functions, and results, which allow us to interpret the parameters in this model, are given. As for the inference in this model for complete data, we highlight a result analogous to Fisher Theorem in normal populations.

Finally, in Chapter 4, the Birnbaum-Saunders model is studied in depth, due to its interesting physical interpretation amongst other important properties. The Birnbaum-Saunders distribution models fatigue life of a metal, subject to a cyclic stress, from which it has some special characteristics. It is also remarkable the relationship between Birnbaum-Saunders and Inverse Gaussian distribution. This work concludes with some inference results, where a numerical method is also given in order to estimate this model's parameters. These results will be reflected in the final applications to real data, differentiating between complete and censored samples.

1 | Herramientas básicas

Este trabajo es una introducción a la Teoría de Fiabilidad y Análisis de Supervivencia, en la que estudiaremos distribuciones de probabilidad relacionadas con este campo. Estas distribuciones se suelen utilizar para modelar tiempos de vida sobre muestras completas y sobre muestras censuradas, por lo que comenzaremos viendo las herramientas básicas para trabajar con este tipo de distribuciones. A lo largo de todo el trabajo, se utilizarán variables continuas no negativas T para representar los tiempos de vida.

En este primer capítulo, nos centraremos en conceptos y resultados necesarios para entender la utilidad de las distribuciones de probabilidad que aquí se estudian, además de recordar propiedades básicas de la distribución normal, sobre la cual se basarán las demás.

Existe un gran número de modelos de tiempo de vida: exponencial, gamma, Weibull, etc. En este trabajo nos centramos en tres de ellas: Log Normal, Inversa Gaussiana y Birnbaum-Saunders. En secciones posteriores, se verán propiedades básicas y resultados de interés sobre algunas de estas distribuciones y su relación con la normal, así como aplicaciones a datos reales tanto de muestras completas como censuradas. Dichas aplicaciones se han realizado utilizando el software RStudio.

1.1 Conceptos básicos en distribuciones de tiempos de vida

El término **fiabilidad** describe la probabilidad de que un elemento funcione correctamente durante un cierto período de tiempo.

Este trabajo se basa en el estudio de **tiempos de vida**, y bajo ese concepto pode-

mos dividir los objetos de estudio de la siguiente forma:

- Productos reparables: Aquellos en los que puede ocurrir más de un fallo, es decir, el producto se puede arreglar. *Ejemplos: ordenadores, electrodomésticos, etc.*
- Productos no reparables: Aquellos que se descartan una vez fallan, sin posibilidad de arreglo. *Ejemplos: aislante eléctrico, diodo, bombilla eléctrica, etc.*

Los productos que nos interesan para este trabajo son los no reparables. Las distribuciones que veremos más adelante estudian su tiempo de vida, usando datos independientes e idénticamente distribuidos (lo notaremos "i.i.d.").

Llamaremos T a la variable aleatoria (v.a.) continua y no negativa que describe el tiempo de vida del individuo en estudio. Por comodidad, denotamos la *función de densidad (fdd)* de T por $f(t)$ y la *Función de Distribución (FdD)* de T por $F(t)$. Sabemos que estas dos funciones cumplen

$$F(t) = P(T \leq t) = \int_0^t f(x) dx \quad \text{y} \quad F'(t) = f(t), \quad t > 0.$$

| Definición 1.1. Llamamos *función de fiabilidad o de supervivencia* a

$$S(t) = P(T \geq t).$$

Su valor indica la probabilidad de que el tiempo de vida del producto sea mayor que el instante t , es decir, la probabilidad de que el producto no falle antes de ese instante.

Por tanto, la función $S(t)$ se puede definir como

$$S(t) = 1 - F(t).$$

La siguiente proposición nos da algunas propiedades de la función de supervivencia, obtenidas a partir de su relación con la FdD.

| Proposición 1.1. Sea T una v.a. continua no negativa y sea $S(t)$ su función de fiabilidad. Entonces, se cumple:

1. $S(t)$ es continua y monótona decreciente.

2. $S(0) = 1$.
3. $S(+\infty) = 0$.

Demostración.

1. Tenemos que $S(t) = 1 - F(t)$, donde F es la FdD de T . Al ser T una v.a. continua, $F(t)$ debe ser una función continua y, por tanto, $S(t)$ también. Además, $F(t)$ es monótona creciente, luego $S(t)$ será monótona decreciente.
2. $S(0) = 1 - F(0) = 1 - \int_0^0 f(x) dx = 1 - 0 = 1$.
3. $S(+\infty) = \lim_{t \rightarrow +\infty} (1 - F(x)) = 1 - \lim_{t \rightarrow +\infty} (\int_0^t f(x) dx) = 1 - \int_0^{+\infty} f(x) dx = 1 - 1 = 0$.

Definición 1.2. Se le llama **función de riesgo, tasa de fallo o función de hazard**, a la función $h(t)$ que indica la probabilidad de que el individuo falle en el instante t . Se define como

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}. \quad (1.1)$$

De (1.1), se obtiene

$$h(t)\Delta t \simeq \lim_{\Delta t \rightarrow 0} P(t \leq T < t + \Delta t \mid T \geq t), \quad (1.2)$$

lo cual es una aproximación de la probabilidad de fallo en el intervalo $(t, t + \Delta t)$

La función $h(t)$ se puede definir también como

$$h(t) = \frac{f(t)}{S(t)}, \quad (1.3)$$

lo cual nos facilitará mucho su cálculo.

La función de riesgo $h(t)$ nos ayuda a estudiar la fiabilidad del individuo. Podemos dividir su comportamiento en tres tipos:

- **Comportamiento creciente:** Nos indica que es más probable que el producto falle conforme pasa el tiempo.
- **Comportamiento decreciente:** Indica que es más probable que haya fallos al principio de la vida útil del producto.

- **Comportamiento constante:** La probabilidad de fallo es la misma a lo largo de toda la vida útil del individuo.

Si en el comportamiento del individuo se observa al comienzo un comportamiento decreciente, luego constante y finalmente creciente, obtenemos la llamada **curva de bañera** o **bathtub curve**.

Definición 1.3. La **función de riesgo acumulada** se define como

$$H(t) = \int_0^t h(x) dx, \quad t > 0$$

Observación 1.1. Si $h(t)$ es constante, al representar $H(t)$ obtendremos una línea recta. Si $h(t)$ es creciente, obtendremos una función con un crecimiento más rápido que el de una recta, y si $h(t)$ es decreciente, una que crece más lento.

Proposición 1.2. Las funciones de riesgo y de fiabilidad cumplen

1. $f(t) = -S'(t), \quad \forall t > 0$
2. $h(t) = -\frac{d}{dt} \log S(t), \quad \forall t > 0$
3. $S(t) = \exp\{-H(t)\}, \quad \forall t > 0$

Demostración.

1. $S(t) = 1 - F(t) \Rightarrow S'(t) = -F'(t) = -f(t)$.
2. $\frac{d}{dt} \log S(t) = \frac{S'(t)}{S(t)} = -\frac{f(t)}{S(t)} = -h(t)$.
3. Integrando en la expresión anterior, y sabiendo que $S(0) = 1$,

$$\frac{d}{dt} \log S(t) = -h(t) \Rightarrow \int_0^x \frac{d}{dt} \log S(t) dt = \log S(t) \Big|_{t=0}^{t=x} = \log S(x) = - \int_0^x h(t) dt \Rightarrow$$

$$\Rightarrow \log S(x) = - \int_0^x h(t) dt = -H(x) \Rightarrow S(x) = e^{-H(x)}$$

|

Cuantiles. Los cuantiles son una medida interesante para el estudio de tiempos de vida. Concretamente, serán de utilidad la mediana y los cuantiles de orden muy bajo y muy alto, ya que nos pueden aportar detalles importantes para la interpretación de los datos. Para cualquier distribución, se definen como sigue

Definición 1.4. Sea T una variable aleatoria. Se define el **cuantil de orden p** , al valor t_p tal que

$$P(T \leq t_p) = p, \quad 0 < p < 1.$$

Y por tanto,

$$t_p = F_T^{-1}(p), \quad 0 < p < 1.$$

O sea, es el valor de t tal que un porcentaje $100p$ de la población total es menor o igual que dicho valor.

Los **percentiles** dividen la distribución en cien partes, luego tomando esta referencia, t_p es el percentil $100p$.

Observación 1.2. El cuantil de orden 0.5, $t_{0,5}$, se corresponde con la mediana.

En estudios de fiabilidad y análisis de supervivencia, además de la mediana, suelen ser de gran interés cuantiles de orden bajo como $t_{0,01}$, $t_{0,05}$, $t_{0,10}$ y de orden alto como $t_{0,90}$, $t_{0,95}$, $t_{0,99}$.

1.2 Inferencia en muestras completas

El método de estimación que se utilizará en este trabajo será el método de máxima verosimilitud. Vamos a ver de manera resumida propiedades generales de los estimadores máximo verosímiles (EMV).

El método de máxima verosimilitud es el más conocido de los métodos de estimación. El método original fue introducido por Fisher en 1922, y desde entonces ha recibido importantes contribuciones por parte de otros autores

Definición 1.5. Sea una m.a.s T_1, \dots, T_n de una v.a. T con función de densidad $f(t, \underline{\theta})$ donde $\underline{\theta}$ es un vector de parámetros desconocido que toma valores en el espacio paramétrico Θ . Entonces, se define la función de verosimilitud de las n variables como

$$L(t_1, \dots, t_n; \underline{\theta}) = L(t_1, \dots, t_n; \theta_1, \theta_2) = f(t_1, \dots, t_n; \underline{\theta}) = \prod_{i=1}^n f(t_i, \underline{\theta}), \quad \forall \underline{\theta} \in \Theta. \quad (1.4)$$

Método. El método para hallar el estimador de $\underline{\theta}$ consistirá en hallar aquel valor $\underline{\hat{\theta}} = \hat{\theta}(T_1, \dots, T_n)$ tal que maximiza la función de verosimilitud, es decir, que verifica

$$L(t_1, \dots, t_n; \underline{\hat{\theta}}) = \max_{\underline{\theta} \in \Theta} L(\underline{t}; \underline{\theta}).$$

Este valor recibirá el nombre de **estimador máximo verosímil** o **estimador de máxima verosimilitud** del parámetro $\underline{\theta}$.

Sin embargo, es común que la función de verosimilitud en (1.4) tenga una expresión complicada, dificultando los cálculos. Esta función es positiva y sus máximos se mantienen al aplicar un logaritmo neperiano, al ser esta función monótona creciente. Por ello, es frecuente considerar su logaritmo neperiano. Esta nueva función recibe el nombre de **función de log-verosimilitud**, cuya expresión es

$$l(\underline{\theta}) = \log L(\underline{t}; \underline{\theta}) = \sum_{i=1}^n \log f(t_i; \underline{\theta}), \quad \forall \underline{\theta} \in \mathbb{R}. \quad (1.5)$$

Generalmente, la ecuación o sistema de ecuaciones de verosimilitud se puede resolver sin grandes inconvenientes. Sin embargo, en ocasiones hay que recurrir a métodos numéricos iterativos para poder hallar el estimador.

Caso uniparamétrico. En el caso de que tengamos un solo parámetro $\theta \subseteq \Theta \subseteq \mathbb{R}$ y se verifiquen las siguientes condiciones de regularidad:

- El campo de variación de θ es un intervalo abierto del eje real.
- El campo de variación de la variable aleatoria poblacional no depende de θ .
- La fdd $f(t_i, \theta)$ es positiva y derivable respecto de θ .
- Se verifica la condición del máximo $\left. \frac{\partial^2 l}{\partial \theta^2} \right|_{\theta=\hat{\theta}} < 0$.

Entonces, se cumplen las siguientes propiedades.

1. Consistencia. Los EMV son consistentes, lo que significa que $\forall \epsilon > 0$, se verifica

$$\lim_{n \rightarrow \infty} P(|\underline{\hat{\theta}} - \underline{\theta}| < \epsilon) = 1, \quad \forall \underline{\theta} \in \Theta.$$

2. Insesgadez. En general, los EMV no serán insesgados. En caso de serlo, serán asintóticamente insesgados.

3. Normalidad asintótica. Al aumentar el tamaño muestral, la distribución de $\hat{\theta}$ puede aproximarse por una distribución normal.

Para un parámetro $\theta \in \Theta$, se verifica

$$\hat{\theta} \xrightarrow{\mathcal{L}} N(\theta, \sqrt{Var(\hat{\theta})}),$$

con $Var(\hat{\theta})$ la varianza asintótica de $\hat{\theta}$.

Caso multiparamétrico $k \geq 2$. En el caso de que la función de densidad dependa de k parámetros, los EMV se obtienen resolviendo el sistema de ecuaciones de verosimilitud en $\theta_1, \dots, \theta_k$. Las ecuaciones de este sistema vendrán dadas por

$$\frac{\partial l(t_1, \dots, t_n; \theta_1, \dots, \theta_k)}{\partial \theta_j} = \sum_{i=1}^n \frac{\partial \log f(t_i; \theta_1, \dots, \theta_k)}{\partial \theta_j} = 0, \forall j = 1, \dots, k.$$

En nuestro caso, será $k = 2$.

Bajo condiciones de regularidad los EMVs obtenidos tendrán buenas propiedades asintóticas, que son una generalización al caso multiparamétrico de las expuestas anteriormente. En particular, será de gran interés la matriz de información de Fisher, pues nos permite obtener la matriz de varianza asintótica de los EMVs.

Matriz de información de Fisher

En este trabajo consideramos modelos biparamétricos para T . Dada una muestra aleatoria de T , sus parámetros $\underline{\theta} = (\theta_1, \theta_2)$ se pueden estimar por máxima verosimilitud. En aquellas situaciones en que se pueda suponer la normalidad asintótica de los estimadores MV, nos interesa obtener la matriz de información de Fisher, pues a partir de ella se pueden proponer aproximaciones asintóticas para las varianzas de $\hat{\theta}_1$ y $\hat{\theta}_2$, así como para la covarianza $Cov(\hat{\theta}_1, \hat{\theta}_2)$.

Denotando por $I(\underline{\theta})$ a la matriz de información de Fisher, se tiene que el elemento (i, j) de esta matriz viene dado por

$$-\frac{\partial^2 l(\underline{\theta})}{\partial \theta_i \partial \theta_j}, \quad i, j = 1, 2,$$

donde $l(\underline{\theta})$ es la función de logverosimilitud de la muestra en cuestión, dada por (1.5). $I(\underline{\theta})$ depende de los valores de los parámetros que son desconocidos. Estos se suelen

estimar por MV, obteniendo así la matriz de información de Fisher observada; $\mathcal{I}(\hat{\theta})$. La matriz inversa de $\mathcal{I}(\hat{\theta})$, $\hat{V} = [\mathcal{I}(\hat{\theta})]^{-1}$, nos da estimaciones de la varianza y covarianza de los estimadores de los parámetros. En concreto, los elementos de la diagonal nos dan las varianzas, y los de fuera de la diagonal, la covarianza. Este resultado se utilizará para aproximar el error estándar de los EMV de los θ_i , $i = 1, 2$, y para obtener intervalos de confianza aproximados.

1.3 Inferencia en muestras censuradas

En el estudio de datos de supervivencia, es muy común encontrar que las muestras de datos puedan ser **censuradas**, esto es, nos dan información incompleta por algún motivo, el cual dependerá del tipo de censura. De esta manera, se necesita un mecanismo diferente para tratar con estos datos a los usados con las muestras completas. Hay diferentes tipos de censura, como podemos ver a continuación.

- **Censura de tipo I:** Se da cuando el investigador fija un tiempo máximo para la observación del tiempo de vida de los datos. Entonces, cuando un individuo observado no presenta una falla dentro de ese tiempo, no aporta la información necesaria al investigador y constituye una observación censurada.
- **Censura de tipo II:** Supongamos que se observa una muestra de n elementos y se fija un número r tal que $1 \leq r \leq n$. En este caso, el experimento termina cuando r de los individuos fallan, resultando por tanto r datos observados tales que $t_{1:n} \leq \dots \leq t_{r:n}$, y $n - r$ datos censurados, puesto que no han fallado al acabar la investigación. Por $t_{i:n}$ se ha denotado al estadístico de orden i en una muestra de tamaño n .
- **Censura aleatoria:** En este tipo, la censura no depende de ninguna decisión del investigador. Se da cuando cada individuo bajo estudio tiene un tiempo de vida T_i y un tiempo de censura C_i , siendo T_i y C_i variables aleatorias continuas e independientes entre ellas.

Para modelizar este tipo de datos, se utilizará la notación que sigue. Supongamos una muestra aleatoria de n individuos, cuyos tiempos de vida vienen dados por las variables aleatorias T_1, \dots, T_n , respectivamente. Llamamos t_i a lo que puede ser o bien el tiempo de vida observado o bien el tiempo de censura según lo que ocurra primero, esto es, $t_i = \min\{T_i, C_i\}$, siendo C_i el tiempo de censura en la observación i . Definimos

la variable δ_i , $i = 1, \dots, n$, tal que

$$\delta_i = \begin{cases} 1 & \text{si } t_i \text{ dato observado} \\ 0 & \text{si } t_i \text{ dato censurado} \end{cases}$$

De esta manera, las muestras censuradas se escriben como (t_i, δ_i) , $i = 1, \dots, n$.

Método de máxima verosimilitud

Para poder hacer inferencia sobre este tipo de observaciones, se suele utilizar el método de máxima verosimilitud, que se verá alterado por el carácter de estos datos. De forma general, denotaremos al vector con los parámetros del modelo bajo estudio como $\underline{\theta} = (\theta_1, \theta_2)$. La función de verosimilitud cuando partimos de una muestra de datos censurados (t_i, δ_i) , $i = 1, \dots, n$, viene dada por la siguiente expresión

$$L(\underline{\theta}) = \prod_{i=1}^n f(t_i)^{\delta_i} S(t_i)^{1-\delta_i}, \quad (1.6)$$

donde f es la fdd de la muestra y S denota la función de supervivencia correspondiente.

Esta forma de la función de verosimilitud es común a los tres tipos de datos censurados nombrados anteriormente (ver Lawless [11]).

También puede ser de utilidad la función de logverosimilitud, cuya expresión para este tipo de muestras es

$$l(\underline{\theta}) = \log L(\underline{\theta}) = \sum_{i=1}^n \delta_i \log f(t_i) + \sum_{i=1}^n (1 - \delta_i) \log S(t_i). \quad (1.7)$$

Matriz de información de Fisher

En el caso de muestras censuradas, sigue teniendo sentido considerar la matriz de información de Fisher. Al igual que en el caso de muestras completas, viene definida como la matriz donde cada elemento (i, j) viene dado por

$$\frac{-\partial l(\underline{\theta})}{\partial \theta_i \partial \theta_j}, \quad i, j = 1, 2.$$

En este caso, $l(\underline{\theta})$ representa la función de log-verosimilitud de la muestra censurada, dada por (1.7).

Esta matriz se utilizará para aproximar el error estándar y covarianza asintótica de los estimadores de máxima verosimilitud.

1.4 Distribución Normal

En esta sección se van a recoger algunos resultados importantes de la distribución normal, que nos serán útiles para abordar el estudio de modelos más complejos en capítulos posteriores.

| Definición 1.6. Sean $\mu \in \mathbb{R}$ y $\sigma > 0$, y sea X una variable aleatoria (v.a.). Decimos que X sigue una **distribución normal univariante de parámetros μ y σ** , y se denota por $X \sim N(\mu, \sigma^2)$, si su función de densidad (fdd) se expresa como sigue

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}, \quad -\infty < x < \infty. \quad (1.8)$$

Es conveniente remarcar el caso particular de la distribución normal donde $\mu = 0$ y $\sigma = 1$, debido a su gran utilidad en múltiples campos de la estadística y de este mismo trabajo.

| Definición 1.7. Sea una v.a. Z tal que $Z \sim N(0, 1)$. Entonces, se dice que Z sigue una **distribución normal estándar univariante**

La fdd de Z se denota por ϕ , y viene dada por

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty, \quad (1.9)$$

mientras que su FdD, denotada como Φ , es de la siguiente forma

$$\Phi(x) = \int_{-\infty}^x \phi(u) du. \quad (1.10)$$

La siguiente proposición nos da algunas de las propiedades de la distribución normal estándar, que serán importante para el posterior estudio de las propiedades tanto de la normal como de otras distribuciones.

| Proposición 1.3. Si $Z \sim N(0, 1)$, entonces se verifica

1. ϕ es una función par, i.e., $\phi(x) = \phi(-x)$, $\forall x \in \mathbb{R}$.
2. $\Phi(-x) = 1 - \Phi(x)$
3. Función generatriz de momentos de Z :

$$M_Z(t) = e^{\frac{1}{2}t^2}, \quad t \in \mathbb{R}. \quad (1.11)$$

4. $E(Z) = 0$; $Var(Z) = 1$.

Demostración.

1. Trivial a partir de (1.9).
2. Inmediato por la simetría de ϕ .
3. La FGM se define como $M_Z(t) = E(e^{tZ})$. Entonces

$$M_Z(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz} e^{-\frac{z^2}{2}} dz = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz - \frac{z^2}{2} - \frac{t^2}{2}} dz = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-t)^2} dz.$$

Tomamos el cambio de variable $r = z - t$, resultando así

$$M_Z(t) = e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}r^2} dr = e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \phi(r) dr = e^{\frac{t^2}{2}}, \quad t \in \mathbb{R}.$$

4. Recordemos que $E(Z^k) = M^{(k)}(t)|_{t=0}$. Entonces

$$\frac{dM_Z}{dt}(t) = te^{\frac{t^2}{2}} \Rightarrow E(Z) = 0.$$

Sabemos que $Var(Z) = E(Z^2) - E(Z)^2$. Acabamos de probar que $E(Z) = 0$, luego basta con hallar el valor de $E(Z^2)$.

$$E(Z^2) = \frac{d^2 M_Z}{dt^2}(t) = e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}} \Rightarrow Var(Z) = 1.$$

| Definición 1.8. Sea $Z \sim N(0, 1)$. Para $0 < p < 1$, se denotará por z_p al **cuantil de orden p** , es decir, al valor z_p tal que $\Phi(z_p) = p$.

Sea $X \sim N(\mu, \sigma^2)$. Denotamos por x_p al cuantil de orden p . Puesto que podemos expresar $X = \mu + \sigma Z$ con $Z \sim N(0, 1)$, se tiene la siguiente relación entre los cuantiles de ambas distribuciones

$$x_p = \mu + \sigma z_p.$$

La siguiente proposición es esencial en el estudio de la distribución normal, ya que nos da una relación entre esta y la distribución normal estándar. A esta relación se le llama estandarización.

Proposición 1.4 (Estandarización). Sea $X \sim N(\mu, \sigma^2)$ y sea $Z = \frac{X-\mu}{\sigma}$. Entonces, $Z \sim N(0, 1)$.

Demostración.

Sea $Z = \frac{X-\mu}{\sigma}$. Podemos calcular su FdD como sigue

$$F_Z(z) = P(Z \leq z) = P\left(\frac{X-\mu}{\sigma} \leq z\right) = P(X \leq \sigma z + \mu) = F_X(\sigma z + \mu).$$

A partir de esto, podemos obtener su fdd derivando F_Z y usando (1.8).

$$f_Z(z) = \frac{dF_X}{dz}(\sigma z + \mu) = \sigma f_X(\sigma z + \mu) = \sigma \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(\sigma z + \mu - \mu)^2}{2\sigma^2}\right\}.$$

Se tiene entonces $f_Z(z) = \phi(z) \quad \forall z \in \mathbb{R}$, luego $Z \sim N(0, 1)$ |

Estas propiedades nos dan las herramientas necesarias para ver distintas características de la distribución normal, empezando por el siguiente corolario, que nos da una caracterización de la FdD de la distribución.

Corolario 1.1. Dada $X \sim N(\mu, \sigma^2)$, esta puede expresarse como $X = \mu + \sigma Z$, y su FdD vendrá dada por

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right). \quad (1.12)$$

Proposición 1.5 (Propiedades). Si $X \sim N(\mu, \sigma^2)$, entonces se verifica

1. $f(x)$ es simétrica respecto de μ : $f(x) = f(2\mu - x)$, $\forall x \in \mathbb{R}$
2. $E(X) = \mu$, $Var(X) = \sigma^2$
3. Función generatriz de momentos de X :

$$M_X(t) = \exp\left\{t\mu + \frac{1}{2}\sigma^2 t^2\right\}, \quad t \in \mathbb{R}. \quad (1.13)$$

Demostración.

1. Tomando $x = 2\mu - x$ en (1.8), se observa

$$f(2\mu - x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(2\mu - x - \mu)^2}{2\sigma^2}\right\} = \phi(x), \quad x \in \mathbb{R}.$$

2. Sabemos que se cumple la relación $X = \mu + \sigma Z$, donde $Z \sim N(0, 1)$. Entonces, mediante propiedades conocidas de la esperanza y la varianza se tiene que

$$E(X) = E(\mu + \sigma Z) = \mu + \sigma E(Z) = \mu.$$

$$Var(X) = Var(\mu + \sigma Z) = \sigma^2 Var(Z) = \sigma^2.$$

3. Aplicando de nuevo la transformación lineal $X = \mu + \sigma Z$,

$$M_X(t) = E(e^{tX}) = E(e^{t(\mu + \sigma Z)}) = e^{t\mu} E(e^{t\sigma Z}).$$

Tomamos el cambio de variable $t\sigma = u$, resultando

$$M_X(t) = e^{t\mu} E(e^{uZ}) = e^{t\mu} e^{-\frac{1}{2}u^2} = e^{t\mu} e^{-\frac{1}{2}\sigma^2 t^2}.$$

Finalmente, deshaciendo el cambio,

$$M_X(t) = \exp \left\{ t\mu + \frac{1}{2}\sigma^2 t^2 \right\}, \quad t \in \mathbb{R}.$$

La función de densidad de $X \sim N(\mu, \sigma^2)$ tiene forma de campana (campana de Gauss). El parámetro μ es la media, mediana y moda de la distribución. Por otro lado, σ es la desviación típica, y también un parámetro de escala. En la figura (1.1) pueden observarse estos hechos junto con el efecto de variar los valores de estos parámetros.

| Proposición 1.6. Sea $X \sim N(\mu, \sigma^2)$. Entonces, $aX + b \sim N(a\mu + b, a^2\sigma^2)$ para $a, b \in \mathbb{R}$.

Demostración.

Sea $X \sim N(\mu, \sigma^2)$. Veamos que $Y = aX + b$ sigue una distribución normal, usando la FGM como caracterización de la misma.

$$M_Y(t) = E(e^{tY}) = e^{tb} M_X(at) = \exp \left\{ t(a\mu + b) + \frac{1}{2}(a\sigma)^2 t^2 \right\},$$

luego $Y \sim N(a\mu + b, a^2\sigma^2)$.

| Proposición 1.7 (Reproductividad). Sean $X_1 \sim N(\mu_1, \sigma_1^2)$ y $X_2 \sim N(\mu_2, \sigma_2^2)$ independientes entre sí. Entonces

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

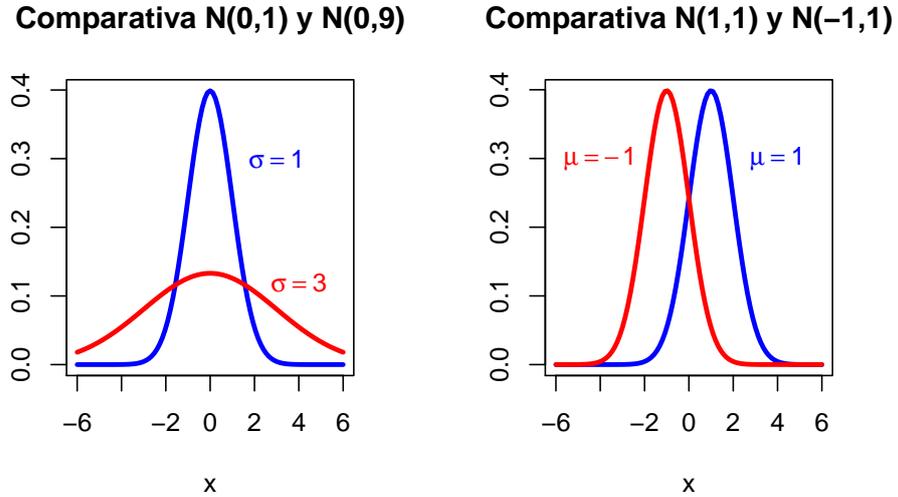


Figura 1.1: fdd de la distribución Normal con distintos parámetros

Demostración.

Tomamos las expresiones de las FGM de X_1 y X_2 vistas en (1.13). Debido a la independencia entre ambas variables, se cumple lo siguiente

$$M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t) = \exp \left\{ t(\mu_1 + \mu_2) + \frac{1}{2}t^2(\sigma_1^2 + \sigma_2^2) \right\},$$

que se corresponde con la expresión de la FGM de una distribución normal de media $\mu_1 + \mu_2$ y varianza $\sigma_1^2 + \sigma_2^2$.

Por tanto, $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. |

Esta propiedad se generaliza de forma inmediata a la suma de n v.a.s independientes. Destacamos el caso en que además de independientes, las v.a.s son idénticamente distribuidas.

Corolario 1.2. Sean X_1, X_2, \dots, X_n v.a.s independientes e idénticamente distribuidas (i.i.d.) según $N(\mu, \sigma^2)$. Entonces

1. $\sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2)$.
2. $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim N(\mu, \frac{\sigma^2}{n})$, donde \bar{X} es la media muestral.

La propiedad anterior es de gran utilidad al hacer inferencia en poblaciones normales, junto con los dos resultados que siguen. En el primero, destacamos que el mo-

delo normal pertenece a la familia exponencial de distribuciones, y en el Teorema de Fisher, se establece la independencia de la media muestral, \bar{X} , y la cuasivarianza muestral, $S_c^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / (n - 1)$, junto con la distribución de estos dos estadísticos en el caso normal.

| Proposición 1.8. Sea $X \sim N(\mu, \sigma^2)$. Nótese que podemos escribir su fdd como

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2\sigma^2} - \frac{\mu^2}{2\sigma^2} + \frac{2\mu x}{2\sigma^2} \right\} = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ \frac{\mu x}{\sigma^2} - \frac{x^2}{2\sigma^2} - \frac{\mu^2}{2\sigma^2} \right\},$$

donde $\mu \in \mathbb{R}, \sigma > 0$.

Pertenece a la familia exponencial biparamétrica de distribuciones. Dada una muestra aleatoria simple (variables aleatorias i.i.d.) $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$, entonces $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ es un estadístico suficiente y completo para (μ, σ^2) . Equivalentemente, (\bar{X}, S_c^2) también es un estadístico suficiente y completo para (μ, σ^2) .

| Teorema 1.1 (Fisher). Sean X_1, \dots, X_n variables aleatorias i.i.d. según $N(\mu, \sigma^2)$. Entonces

1. \bar{X} y S_c^2 son independientes, siendo \bar{X} es la media muestral y S_c^2 la cuasivarianza muestral.
2. $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.
3. $\frac{(n-1)}{\sigma^2} S_c^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi_{n-1}^2$.

Este resultado es la base para realizar inferencia en la distribución normal cuando tenemos muestras completas.

2 | Distribución Log Normal

Comenzamos el estudio de las distribuciones de tiempos de vida con la distribución log-normal. Esta distribución es la que guarda una relación más estrecha con la distribución normal de las que se estudiarán en este trabajo, lo que simplificará su estudio. Veremos propiedades notables de esta distribución, así como su aplicación a datos reales en los ejemplos del final del capítulo.

Definición 2.1. Sea T una v.a. positiva. Si se cumple que $Y = \log T \sim N(\mu, \sigma^2)$, se dice que T sigue una **distribución log-normal** de parámetros μ y σ , y se denota por $T \sim \text{Log}N(\mu, \sigma^2)$.

Sus funciones de densidad y distribución vienen dadas respectivamente por

$$f(t) = \frac{1}{\sqrt{2\pi\sigma t}} \exp \left\{ -\frac{1}{2} \left(\frac{\log t - \mu}{\sigma} \right)^2 \right\}, \quad t > 0. \quad (2.1)$$

$$F(t) = \Phi \left(\frac{\log t - \mu}{\sigma} \right), \quad t > 0, \quad (2.2)$$

siendo Φ la FdD de la distribución $N(0, 1)$ dada por (1.12).

Por tanto, su función de supervivencia $S(t)$ se puede expresar como

$$S(t) = 1 - F(t) = 1 - \Phi \left(\frac{\log t - \mu}{\sigma} \right), \quad t > 0. \quad (2.3)$$

Gráficos de la función de densidad $\text{Log}N(\mu, \sigma^2)$ para distintos valores de sus parámetros pueden verse en la Figuras 2.2 y 2.3.

2.1 Propiedades básicas

Esta sección recoge propiedades fundamentales de la distribución así como características de su forma que nos permitirán reconocerla a la hora de trabajar con datos reales de tiempos de vida.

Proposición 2.1 (Propiedades). Si $T \sim \text{Log}N(\mu, \sigma^2)$, se verifica

1. Momento no central de orden r :

$$E(T^r) = \exp \left\{ r\mu + \frac{1}{2}r^2\sigma^2 \right\} \quad (2.4)$$

2. Media:

$$E(T) = e^{\mu + \frac{1}{2}\sigma^2} \quad (2.5)$$

3. Varianza:

$$\text{Var}(T) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1) \quad (2.6)$$

4. Cuantil de orden p :

$$t_p = e^{\mu + \sigma z_p}, \quad 0 < p < 1. \quad (2.7)$$

En particular, se tiene que la mediana es $t_{0,50} = e^\mu$

Demostración.

1. Por definición de momento no central,

$$E(T^r) = \int_0^{+\infty} t^r f(t) dt = \int_0^{+\infty} t^r \frac{1}{\sqrt{2\pi\sigma t}} e^{-\frac{1}{2}\left(\frac{\log t - \mu}{\sigma}\right)^2} dt.$$

Tomando $y = \log t$, y sabiendo que $Y = \log T \sim N(\mu, \sigma^2)$, obtenemos

$$E(T^r) = \int_{-\infty}^{+\infty} e^{yr} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy = M_Y(r) = \exp \left\{ r\mu + \frac{1}{2}r^2\sigma^2 \right\}.$$

2. Tomando $r = 1$ en el momento no central definido en (2.4), obtenemos

$$E(T) = e^{\mu + \frac{1}{2}\sigma^2}$$

3. Tomando ahora el momento de orden $r = 2$, se obtiene

$$E(T^2) = e^{2\mu + \frac{1}{2}4\sigma^2},$$

y teniendo en cuenta que $E^2(T) = e^{2\mu + \sigma^2}$, se obtiene que

$$Var(T) = E(T^2) - E^2(T) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$$

4. Aplicando la definición de cuantil de orden p , t_p debe cumplir

$$\Phi\left(\frac{\log t_p - \mu}{\sigma}\right) = p \quad \Rightarrow \quad \frac{\log t_p - \mu}{\sigma} = z_p \quad \Rightarrow \quad \log t_p = \mu + \sigma z_p.$$

Se tiene entonces

$$t_p = e^{\mu + \sigma z_p},$$

siendo z_p el cuantil de orden p de la distribución $N(0, 1)$.

Puesto que $z_{0.5} = 0$, se tiene que la mediana es $t_{0.5} = e^\mu$.



La siguiente proposición define el comportamiento de la función de hazard de esta distribución, que se podrá comprobar gráficamente en la Figura 2.1.

Proposición 2.2. Dada $T \sim \text{Log}N(\mu, \sigma^2)$, su función de riesgo $h(t)$ tiende a cero cuando $t \rightarrow 0^+$ y aumenta hasta un máximo, tras lo cual decrece y tiende a cero de nuevo cuando $t \rightarrow +\infty$

Demostración. Sabemos que $h(t) = \frac{f(t)}{S(t)}$ y que, por propiedades conocidas del límite, $\lim \frac{f}{S} = \frac{\lim f}{\lim S}$.

1. Cuando $t \rightarrow 0^+$, obtenemos una indeterminación, luego usaremos L'Hôpital para calcular este límite.

a) Veamos primero $\lim_{t \rightarrow 0^+} f(t)$. Se cumple que

$$\lim_{t \rightarrow 0^+} \frac{1}{\sqrt{2\pi\sigma t}} \exp\left\{-\frac{1}{2}\left(\frac{\log t - \mu}{\sigma}\right)^2\right\} = 0,$$

ya que el comportamiento de este límite viene marcado por la función exponencial.

b) Estudiemos ahora el comportamiento de $\lim_{t \rightarrow 0^+} S(t)$.

$$\lim_{t \rightarrow 0^+} S(t) = \lim_{t \rightarrow 0^+} \left(1 - \Phi \left(\frac{\log t - \mu}{\sigma} \right) \right) = 1 - \lim_{t \rightarrow 0^+} \Phi \left(\frac{\log t - \mu}{\sigma} \right) = 1 - \Phi(-\infty) = 1.$$

Por tanto, por L'Hôpital se tiene

$$\lim_{t \rightarrow 0^+} h(t) = \frac{\lim_{t \rightarrow 0^+} f(t)}{\lim_{t \rightarrow 0^+} S(t)} = \frac{0}{1} = 0.$$

2. Cuando $t \rightarrow +\infty$, obtenemos una indeterminación del tipo $\frac{0}{0}$. Usaremos L'Hôpital para calcular este límite.

$$S'(t) = -f(t)$$

$$f'(t) = -f(t) \frac{1}{t} \left(1 + \left(\frac{\log t - \mu}{\sigma^2} \right) \right).$$

Entonces

$$\lim_{t \rightarrow +\infty} \frac{f'(t)}{S'(t)} = \frac{\lim_{t \rightarrow +\infty} f'(t)}{\lim_{t \rightarrow +\infty} S'(t)} = \lim_{t \rightarrow +\infty} \frac{1}{t} \left(1 + \left(\frac{\log t - \mu}{\sigma^2} \right) \right) = 0.$$

Conociendo este valor, podemos aplicar L'Hôpital como sigue

$$\lim_{t \rightarrow +\infty} h(t) = \lim_{t \rightarrow +\infty} \frac{f(t)}{S(t)} = \lim_{t \rightarrow +\infty} \frac{f'(t)}{S'(t)} = 0.$$

3. Es claro que al ser $h(t) > 0$, $\forall t > 0$, y cumplir que tiende a 0 cuando $t \rightarrow 0^+$ y cuando $t \rightarrow \infty$, deberá existir al menos un máximo de la función. Para más detalles sobre la unicidad de este, ver Lawless [13].

|

Observación 2.1. Se tienen los siguientes resultados sobre la forma de la fdd de la distribución, que se pueden ver reflejados en las Figuras 2.2 y 2.3.

- La fdd de la distribución Log Normal es siempre unimodal.
- La fdd de la Log Normal es asimétrica a la derecha.
- Por ello, la fdd de la Log Normal es muy útil para modelar datos asimétricos a la derecha.

Funciones de hazard Log–Normal

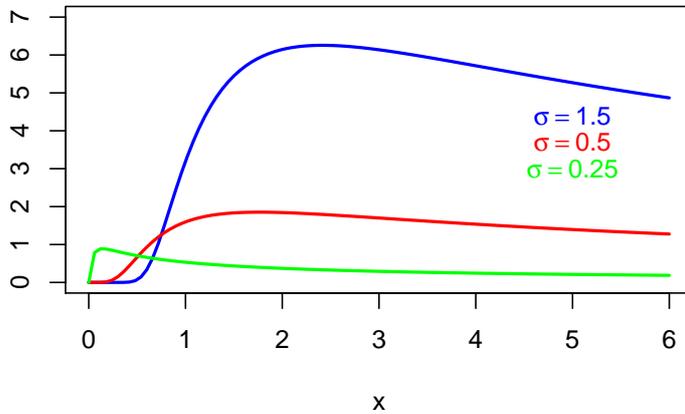


Figura 2.1: Función de hazard Log Normal con $\mu = 0$ y $\sigma = 0.25, 0.5$ y 1.5 .

Funciones de densidad Log–Normal

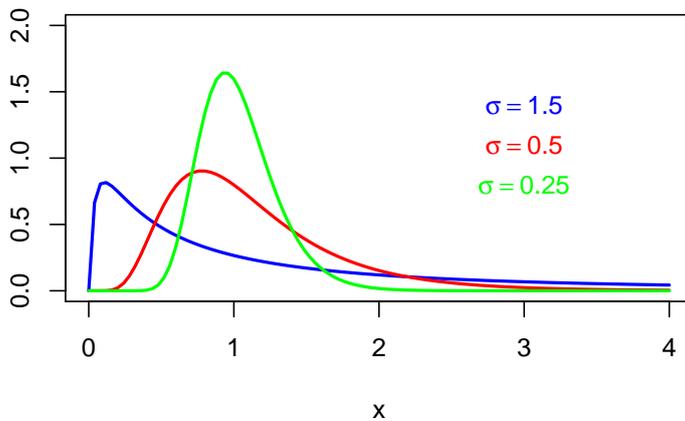


Figura 2.2: Función de densidad Log Normal con $\mu = 0$ y $\sigma = 0.25, 0.5$ y 1.5 .

2.2 Inferencia de la distribución Log Normal

Sea T una v.a. de tiempo de vida tal que $T \sim \text{LogN}(\mu, \sigma^2)$, y sea $Y = \log T$ tal que $Y \sim N(\mu, \sigma^2)$. El objetivo de este apartado es calcular $\hat{\mu}$ y $\hat{\sigma}$, los estimadores de

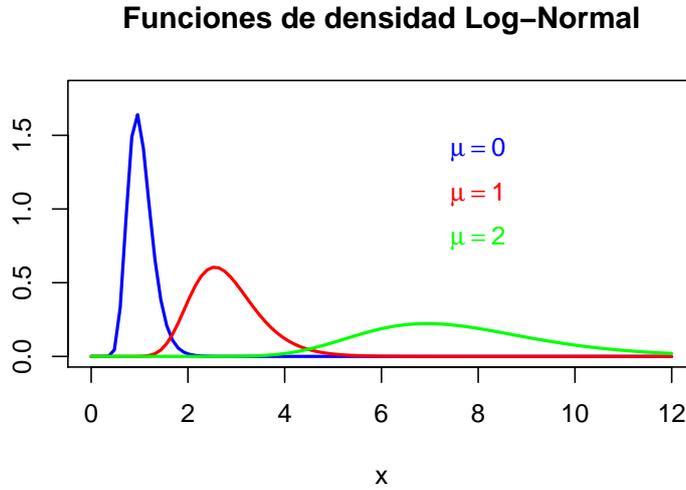


Figura 2.3: Función de densidad Log Normal con $\sigma = 0.25$ y $\mu = 1, 2$ y 3 .

máxima verosimilitud de los parámetros μ y σ de la distribución.

Muestras completas

Para muestras aleatorias completas, es fácil obtener $\hat{\mu}$ y $\hat{\sigma}$, ya que tomando la transformación $y_i = \log t_i$, reducimos la muestra a una que sigue una distribución normal. Usando (2.1), calculamos la función de verosimilitud y la función de log-verosimilitud.

$$L(\mu, \sigma) = \prod_{i=1}^n f(t_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma t_i}} \exp \left\{ -\frac{1}{2} \left(\frac{\log t_i - \mu}{\sigma} \right)^2 \right\}, \quad (2.8)$$

$$l(\mu, \sigma) = \log L(\mu, \sigma) = -\frac{n}{2} \log 2\pi - n \log \sigma - \sum_{i=1}^n \left(\log t_i + \frac{1}{2} \left(\frac{\log t_i - \mu}{\sigma} \right)^2 \right). \quad (2.9)$$

Calculamos las derivadas de (2.9) e igualamos a cero para obtener los estimadores $\hat{\mu}$ y $\hat{\sigma}$.

$$\frac{\partial l}{\partial \mu}(\mu, \sigma) = \sum_{i=1}^n \frac{\log t_i - \mu}{\sigma^2} = 0 \quad \Rightarrow \quad \sum_{i=1}^n \log t_i = n\mu \quad \Rightarrow \quad \hat{\mu} = \frac{1}{n} \sum_{i=1}^n \log t_i.$$

También puede expresarse como

$$\hat{\mu} = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad (2.10)$$

donde \bar{y} es la media muestral de $(y_1, \dots, y_n) = (\log t_1, \dots, \log t_n)$. Por otra parte,

$$\frac{\partial l}{\partial \sigma}(\mu, \sigma) = -\frac{n}{\sigma} + \sum_{i=1}^n \frac{(\log t_i - \mu)^2}{\sigma^3} = 0 \quad \Rightarrow \quad \hat{\sigma} = \left(\sum_{i=1}^n \frac{(\log t_i - \mu)^2}{n} \right)^{1/2}.$$

Calculando la matriz hessiana y evaluándola en $(\hat{\mu}, \hat{\sigma})$, se tiene que $(\hat{\mu}, \hat{\sigma})$ corresponde a un máximo.

Teniendo en cuenta la expresión alternativa de $\hat{\mu}$, la desviación típica de las y_i se puede escribir también como

$$\hat{\sigma} = \left(\sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n} \right)^{1/2} \quad (2.11)$$

o bien, como $\hat{\sigma} = [(n-1)s_i^2/n]^{1/2}$ donde $s_i^2 = \sum (y_i - \bar{y})^2 / (n-1)$ es la cuasivarianza muestral de las y_i .

Aplicando el Teorema de Fisher, se tiene que $\hat{\mu}$ y $\hat{\sigma}^2$ son independientes. Además

$$\hat{\mu} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \quad \frac{(n-1)\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-1}^2.$$

Por tanto, pueden aplicarse resultados para poblaciones normales al realizar inferencia sobre estos parámetros.

Muestras censuradas

Sea una muestra aleatoria censurada de tiempos de vida (t_i, δ_i) , $i = 1, \dots, n$, tal que sigue una distribución Log Normal. Para trabajar con estos datos, utilizaremos la fdd de la distribución, $f(t)$, ya vista en (2.1), y la función de supervivencia $S(t)$, cuya expresión se encuentra en (2.3).

Si denotamos

$$y_i = \log t_i, \quad z_i = \frac{y_i - \mu}{\sigma},$$

$$S_0(z) = 1 - \Phi(z), \quad f_0(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2},$$

entonces se tiene que

$$f(t_i) = \frac{1}{t_i} \frac{1}{\sigma} \phi\left(\frac{\log t_i - \mu}{\sigma}\right) = \frac{1}{t_i} \frac{1}{\sigma} f_0(z),$$

$$S(t_i) = 1 - \Phi\left(\frac{\log t_i - \mu}{\sigma}\right) = S_0(z_i).$$

Usando (1.6), la función de verosimilitud vendrá entonces dada por

$$L(\mu, \sigma) = \prod_{i=1}^n \left[\frac{1}{t_i} \right]^{d_i} \prod_{i=1}^n \left[\frac{1}{\sigma} f_0 \left(\frac{y_i - \mu}{\sigma} \right) \right]^{\delta_i} S_0 \left(\frac{y_i - \mu}{\sigma} \right)^{1 - \delta_i}. \quad (2.12)$$

Como $\prod_{i=1}^n \left[\frac{1}{t_i} \right]^{d_i}$ no depende de ninguno de los parámetros, para calcular la función de log-verosimilitud nos basta con tener en cuenta la siguiente relación

$$L(\mu, \sigma) \propto \prod_{i=1}^n \left[\frac{1}{\sigma} f_0 \left(\frac{y_i - \mu}{\sigma} \right) \right]^{\delta_i} S_0 \left(\frac{y_i - \mu}{\sigma} \right)^{1 - \delta_i}.$$

Si llamamos r al número de datos realmente observados, esto es, $r = \sum_{i=1}^n \delta_i$, entonces la función de log-verosimilitud para μ y σ viene dada por

$$l(\mu, \sigma) = -r \log \sigma + \sum_{i=1}^n \left[\delta_i \log f_0(z_i) + (1 - \delta_i) \log S_0(z_i) \right], \quad (2.13)$$

donde al sustituir f_0 , es lo mismo que

$$l(\mu, \sigma) = -r \log \sigma - \frac{1}{2} \sum_{i=1}^n \delta_i z_i^2 + \sum_{i=1}^n (1 - \delta_i) \log S_0(z_i). \quad (2.14)$$

Es claro que

$$\frac{\partial z_i}{\partial \mu} = -\frac{1}{\sigma}, \quad \frac{\partial z_i}{\partial \sigma} = \frac{-(y_i - \mu)}{\sigma^2} = \frac{-z_i}{\sigma}.$$

Usamos esto para calcular sus derivadas a partir de (2.14) y obtener así las **ecuaciones de verosimilitud**, formadas por las siguientes expresiones igualadas a cero.

$$\frac{\partial l}{\partial \mu} = \frac{1}{\sigma} \sum_{i=1}^n \left(\delta_i z_i + (1 - \delta_i) \frac{f_0(z_i)}{S_0(z_i)} \right), \quad (2.15)$$

$$\frac{\partial l}{\partial \sigma} = -\frac{r}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^n \left(\delta_i z_i^2 + (1 - \delta_i) z_i \frac{f_0(z_i)}{S_0(z_i)} \right). \quad (2.16)$$

Los estimadores $\hat{\mu}$ y $\hat{\sigma}$ no tienen una forma explícita para los datos censurados, siendo necesario un método numérico para resolver las ecuaciones de verosimilitud y obtener así los estimadores que maximizan (2.14).

A continuación, calculamos las derivadas segundas, ya que son necesarias para calcular la matriz de información observada. Estas vienen dadas por

$$\frac{\partial^2 l}{\partial \mu^2} = \frac{1}{\sigma^2} \sum_{i=1}^n \left(-\delta_i + (1 - \delta_i) \frac{f_0(z_i)}{S_0(z_i)} \left(z_i - \frac{f_0(z_i)}{S_0(z_i)} \right) \right), \quad (2.17)$$

$$\begin{aligned} \frac{\partial^2 l}{\partial \sigma^2} &= \frac{r}{\sigma^2} - \frac{2}{\sigma^2} \sum_{i=1}^n \left(\delta_i z_i^2 + (1 - \delta_i) z_i \frac{f_0(z_i)}{S_0(z_i)} \right) \\ &+ \frac{1}{\sigma^2} \sum_{i=1}^n \left(-\delta_i z_i + (1 - \delta_i) z_i^2 \frac{f_0(z_i)}{S_0(z_i)} \left(z_i - \frac{f_0(z_i)}{S_0(z_i)} \right) \right), \end{aligned} \quad (2.18)$$

$$\begin{aligned} \frac{\partial^2 l}{\partial \mu \partial \sigma} &= -\frac{1}{\sigma^2} \sum_{i=1}^n \left(\delta_i z_i^2 + (1 - \delta_i) z_i \frac{f_0(z_i)}{S_0(z_i)} \right) \\ &+ \frac{1}{\sigma^2} \sum_{i=1}^n \left(-\delta_i z_i^2 + (1 - \delta_i) z_i^2 \frac{f_0(z_i)}{S_0(z_i)} \left(z_i - \frac{f_0(z_i)}{S_0(z_i)} \right) \right). \end{aligned} \quad (2.19)$$

Entonces, para obtener la matriz de información observada de Fisher basta con sustituir en la siguiente expresión, donde todos los valores son conocidos.

$$I(\mu, \sigma) = \begin{pmatrix} -\partial^2 l / \partial \mu^2 & -\partial^2 l / \partial \mu \partial \sigma \\ -\partial^2 l / \partial \sigma \partial \mu & -\partial^2 l / \partial \sigma^2 \end{pmatrix}. \quad (2.20)$$

La matriz $I(\hat{\mu}, \hat{\sigma})$ nos da la matriz de covarianzas asintótica $\hat{V} = [I(\hat{\mu}, \hat{\sigma})]^{-1}$ de los estimadores $(\hat{\mu}, \hat{\sigma})$.

2.3 Aplicaciones

Aplicación 1. Muestra completa

Los datos utilizados a continuación han sido obtenidos del trabajo de Overduin [15]. Esta serie de datos es completa, y representa los tiempos de vida en meses de 184 pacientes fallecidos por cáncer de pulmón. Estos datos son los descritos en el Cuadro 2.1, y su representación gráfica puede observarse en el histograma de la Figura 2.4. Usaremos las expresiones calculadas a lo largo de este capítulo para estudiar e interpretar los datos dados.

Datos de cáncer de pulmón							
4,04	4,70	5,82	6,15	7,07	7,36	7,56	7,76
7,82	7,86	7,86	7,89	8,15	8,19	8,84	9,04
9,17	9,24	9,47	9,67	10,03	10,06	10,13	10,26
10,32	10,36	10,42	10,42	10,45	10,52	10,52	10,72
10,75	10,75	11,15	11,18	11,28	11,34	11,47	11,77
11,80	11,93	12,03	12,30	12,39	12,53	12,53	12,53
12,56	12,56	12,82	12,95	13,05	13,12	13,15	13,18
13,28	13,32	13,74	13,91	14,04	14,17	14,33	14,33
14,93	14,93	14,99	15,02	15,01	15,12	15,35	15,52
15,58	15,88	15,95	15,95	16,01	16,11	16,14	16,27
16,41	16,41	16,60	16,67	16,77	17,13	17,16	17,23
17,52	17,79	17,82	17,98	18,02	18,02	18,48	18,61
18,81	18,81	19,13	19,17	19,20	19,20	19,30	19,46
19,53	19,63	19,73	19,82	19,86	19,89	20,05	20,12
20,19	10,22	20,28	20,32	20,65	20,65	20,68	20,68
20,78	20,81	20,84	21,11	21,14	21,47	21,50	21,70
21,80	21,90	22,45	22,62	23,31	23,54	23,57	23,64
23,70	23,70	23,70	23,84	24,03	24,16	24,20	24,46
24,46	24,69	24,72	24,79	25,18	25,35	25,45	25,97
25,97	27,12	27,16	27,48	27,65	28,04	28,27	28,64
29,10	29,98	30,02	30,05	30,97	31,27	32,55	32,61
33,83	34,88	35,38	36,62	38,37	42,38	43,00	44,42
44,65	47,28	47,64	53,82	55,69	58,50	58,82	64,64

Cuadro 2.1: Tiempos de vida en meses de pacientes de cáncer de pulmón.

Tenemos entonces una muestra aleatoria (t_1, \dots, t_n) de $n = 184$ observaciones. El histograma de la muestra nos da una idea de que los datos siguen una distribución log-normal, ya que se puede ver una asimetría muy fuerte a la derecha, lo cual es muy característico de esta distribución.

Podemos comprobar esto mediante un test de hipótesis, ya que si llamamos $Y = \log T$ y tomamos $y_i = \log t_i$, $i = 1, \dots, n$, entonces la muestra (y_1, \dots, y_n) debe seguir una distribución normal. Se puede aplicar a la variable Y el test de normalidad de Shapiro-Wilk, que contrasta la hipótesis $H_0 : Y \sim Normal$ frente a la hipótesis alternativa $H_1 : Y \not\sim Normal$. Utilizando el programa R (ver [16]) para implementar este contraste, se obtiene un p-valor igual a 0,5595, por lo que podemos suponer la

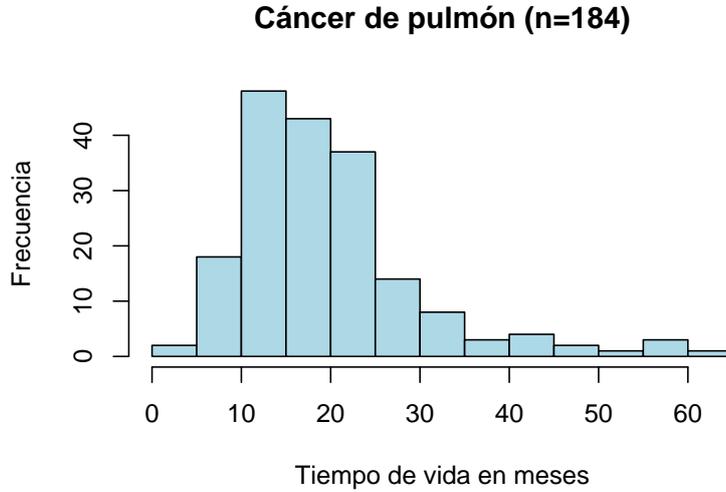


Figura 2.4: Histograma obtenido de los datos del Cuadro 2.1.

hipótesis H_0 . Esto implica de forma inmediata que, en efecto, la variable T sigue una distribución log-normal.

Para calcular los estimadores de los parámetros, partimos de los resultados de máxima verosimilitud dados por (2.10) y (2.11).

$$\hat{\mu} = \sum_{i=1}^{184} \frac{1}{184} \log t_i = 2,855664.$$

$$\hat{\sigma} = \left(\sum_{i=1}^{184} \frac{(\log t_i - \hat{\mu})^2}{184} \right)^{1/2} = 0,4965.$$

Podemos suponer de esta manera que la muestra aleatoria dada sigue una distribución $T \sim \text{LogN}(\hat{\mu}, \hat{\sigma}^2)$. En la Figura 2.5 se observa que el modelo propuesto con estos estimadores se ajusta correctamente a los datos.

Podemos estimar la media utilizando su expresión en (2.5), resultando

$$\widehat{E(T)} = 19,66692.$$

La información que nos da este resultado es que al diagnosticar a un paciente de cáncer de pulmón, se puede esperar que su tiempo de vida desde entonces sea de 19 meses aproximadamente.

También resulta interesante el estudio de algunos de los cuantiles de los datos. Para estimarlos, usamos (2.7) con los estimadores $\hat{\mu}$ y $\hat{\sigma}$ ya calculados.

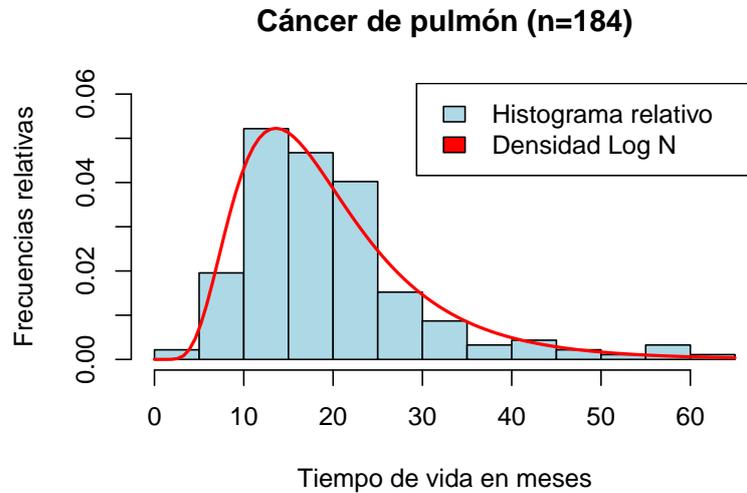


Figura 2.5: Histograma relativo de los datos y curva de la fdd $LogN(2'8556, 0'4965)$.

- Cuantil de orden 0,05

$$\widehat{t}_{0,05} = 7,682425$$

- Cuantil de orden 0,1

$$\widehat{t}_{0,1} = 9,201144$$

- Mediana

$$\widehat{t}_{0,5} = 17,38598$$

- Cuantil de orden 0,9

$$\widehat{t}_{0,9} = 32,85161$$

- Cuantil de orden 0,95

$$\widehat{t}_{0,95} = 39,34596$$

Las estimaciones obtenidas se corresponden con las conclusiones que podríamos sacar de la Figura 2.5. Podemos ver que la mitad de los individuos bajo estudio llegan a un tiempo de vida aproximado de 17 meses. Por otro lado, la información que nos aportan los demás cuantiles estimados es que se puede esperar que los pacientes vivan más de 7 y 9 meses tras el diagnóstico, ya que los casos en que el tiempo de vida es menor corresponde a apenas el 5 y el 10 % de los pacientes. Por otra parte, observamos que de un estudio de unos 60 meses, el 90 % de los individuos estudiados fallecen antes de los 32 meses, y el 95 %; antes de los 39 meses. Siendo así, podemos concluir que son muy excepcionales los casos en que un paciente sobrepase este límite.

Aplicación 2. Muestra censurada

El siguiente ejemplo se ha obtenido del trabajo de Lawless [11]. Cada dato de este conjunto recoge el número de miles de millas recorridas por una locomotora antes de que se produzca un fallo en sus controles. El estudio se realiza durante 135.000 millas y sobre una muestra de 96 locomotoras diferentes, de las cuales 37 presentan un fallo antes del fin del estudio. Se tienen entonces 37 muestras observadas y 59 muestras censuradas iguales a 135. En el Cuadro 2.2 podemos observar los datos, donde el asterisco (*) señala un dato censurado. Su representación gráfica en forma de histograma se encuentra en la Figura 2.6.

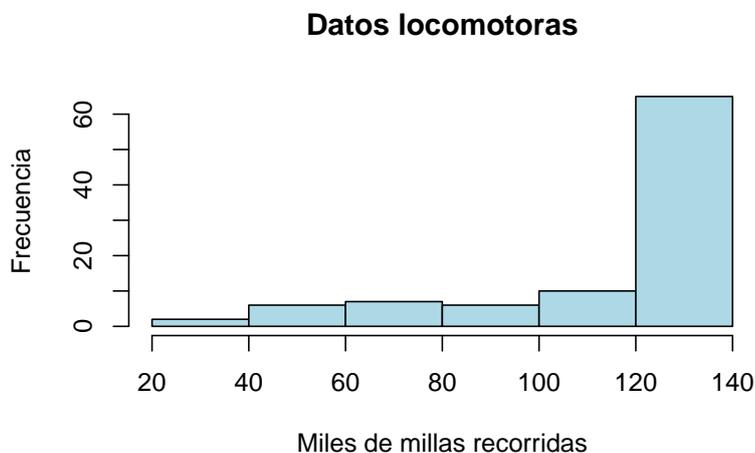


Figura 2.6: Histograma de datos Cuadro 2.2

Utilizamos el paquete "survival" de R (ver [18]) para calcular los estimadores, resultando

$$\hat{\mu} = 5,116925$$

$$\hat{\sigma} = 0,705494$$

En la Figura 2.7 se puede ver el ajuste resultante mediante estos estimadores. La densidad calculada se ajusta bien hasta los últimos valores, debido a que estos últimos datos se corresponden con los de censura. Podemos considerarlo, por tanto, un buen ajuste.

Millas recorridas por las locomotoras							
22,5	37,5	46,0	48,5	51,5	53,0	54,5	57,5
66,5	68,0	69,5	76,5	77,0	78,5	80,0	81,5
82,0	83,0	84,0	91,5	93,5	102,5	107,0	108,5
112,5	113,5	116,0	117,0	118,5	119,0	120,0	122,5
123,0	127,5	131,0	132,5	134,0	135,0*	135,0*	135,0*
135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*
135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*
135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*
135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*
135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*
135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*
135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*	135,0*

Cuadro 2.2: Miles de millas recorridas antes de un fallo en la locomotora.

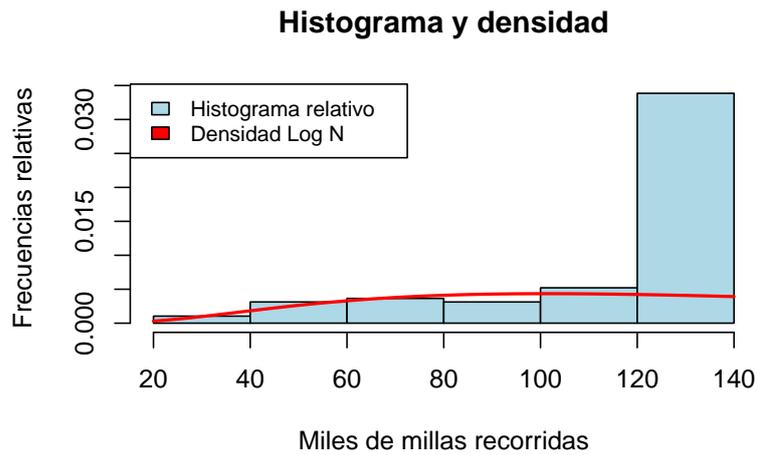


Figura 2.7: Histograma relativo de datos y curva de densidad $LN(\mu, \sigma)$.

Se puede obtener también la matriz de covarianzas asintóticas estimada de esta muestra utilizando R u otro software, resultando

$$\hat{V} = \begin{pmatrix} 0,01085 & 0,005729 \\ 0,005729 & 0,008686 \end{pmatrix}$$

Los valores de la diagonal de \hat{V} se corresponden con la varianza estimada de $\hat{\mu}$ y $\hat{\sigma}$, respectivamente, mientras que el valor de la diagonal inversa nos da la estimación

de la covarianza de dichos parámetros. Vemos que son valores muy próximos a cero, mostrando la poca variabilidad de estos valores y es coherente con la independencia de los estimadores, ya que su covarianza estimada es prácticamente cero.

3 | Distribución Inversa Gaussiana

La distribución Inversa Gaussiana es útil para modelar tiempos de vida en objetos donde el fallo ocurriría en cierto momento del proceso de deterioro. Concretamente, objetos cuya función de riesgo es **creciente** en sus primeros instantes, y luego decrece muy lentamente o bien se mantiene estable. Este modelo de dos parámetros tiene unas propiedades parecidas a la distribución Log-Normal vista anteriormente. Es un modelo bastante complejo, pero con unas propiedades muy interesantes, además de guardar una relación con la distribución Birnbaum-Saunders, que se especificará más adelante.

Veamos a continuación su relación con la distribución Normal.

Definición 3.1. *Un proceso estocástico continuo $\{X(t), t \geq 0\}$ se dice que es un **proceso de Wiener** con coeficiente de desplazamiento γ y coeficiente de dispersión σ^2 si*

1. $X(0) = 0$.
2. $\{X(t), t \geq 0\}$ tiene incrementos independientes y estacionarios.
3. Para cada $t \geq 0$, $X(t) \sim N(\gamma t, \sigma^2 t)$.

La distribución Inversa Gaussiana surge para estudiar el tiempo que tarda un proceso de Wiener continuo con coeficiente de desplazamiento γ y coeficiente de dispersión σ^2 en llegar a un nivel límite $d > 0$ dado. La v.a. $T = \inf\{t : X(t) = d\}$ cumpliría, por las propiedades del proceso de Wiener, $T \sim N(\gamma t, \sigma^2 t)$. Su fdd vendría dada entonces por

$$f(t; \gamma, \sigma) = \frac{d}{\sigma \sqrt{2\pi t^3}} \exp \left\{ -\frac{(d - \gamma t)^2}{2\sigma^2 t} \right\}, \quad t > 0. \quad (3.1)$$

Definición 3.2. *Se dice que una v.a. T sigue una **distribución Inversa Gaussiana** o distribución de Wald de parámetros μ y λ , denotado por $T \sim IG(\mu, \lambda)$, si su fdd es de*

la forma

$$f(t; \mu, \lambda) = \left(\frac{\lambda}{2\pi t^3} \right)^{1/2} \exp \left\{ - \frac{\lambda(t - \mu)^2}{2\mu^2 t} \right\}, \quad t > 0. \quad (3.2)$$

La distribución Inversa Gaussiana es una distribución positiva, asimétrica a la derecha y unimodal.

Observación 3.1. La fdd de la distribución Inversa Gaussiana (3.2) es una reparametrización de la expresión (3.1), tomando $\mu = d/\gamma$ y $\lambda = d^2/\sigma^2$. Hay distintas parametrizaciones posibles. A esta forma se le llama *parametrización de Tweedie*, y cumple que $\mu > 0$, $\lambda > 0$. Otras parametrizaciones pueden verse en [19].

Podemos ver la forma de la fdd de esta distribución en las Figuras 3.1 y 3.2, donde se puede observar el resultado de la variación de sus dos parámetros. La distribución es asimétrica y biparamétrica con densidad similar a la Gamma, aunque con una asimetría mayor y un pico más pronunciado.

Su FdD se puede expresar en términos de la FdD de la distribución normal estándar, $\Phi(\cdot)$, y viene dada por

$$F(t; \mu, \lambda) = \Phi \left[\sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} - 1 \right) \right] + e^{2\lambda/\mu} \Phi \left[- \sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} + 1 \right) \right], \quad t > 0. \quad (3.3)$$

Recordemos que la función de supervivencia se define como $S(t) = 1 - F(t)$, luego

$$S(t; \mu, \lambda) = \Phi \left[- \sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} - 1 \right) \right] - e^{2\lambda/\mu} \Phi \left[- \sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} + 1 \right) \right], \quad t > 0. \quad (3.4)$$

De (3.2) y (3.4), se tiene que la función de hazard de la distribución es de la siguiente forma, y su comportamiento podrá observarse en la Figura 3.3.

$$h(t) = \frac{\left(\frac{\lambda}{2\pi t^3} \right)^{1/2} \exp \left\{ - \frac{\lambda(t-\mu)^2}{2\mu^2 t} \right\}}{\Phi \left[- \sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} - 1 \right) \right] - e^{2\lambda/\mu} \Phi \left[- \sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} + 1 \right) \right]}, \quad t > 0. \quad (3.5)$$

Observación 3.2.

- La fdd de la distribución Inversa Gaussiana es unimodal.
- Es asimétrica a la derecha.
- Estas dos propiedades son comunes a la Log Normal.

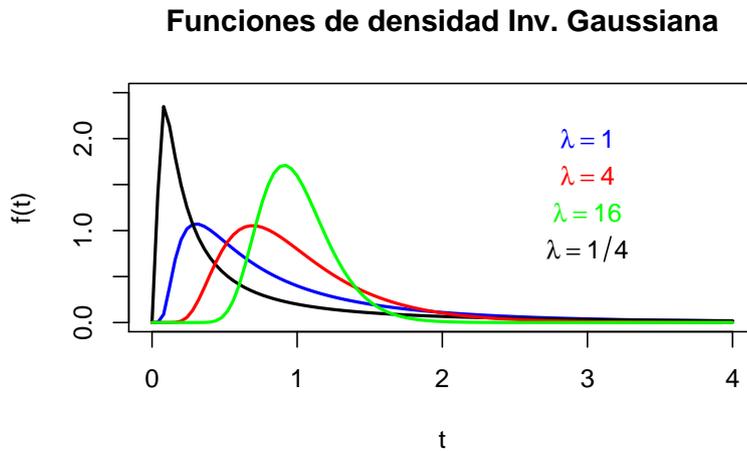


Figura 3.1: Funciones de densidad de la Inv. Gaussiana para $\mu = 1$ y $\lambda = 1, 4, 16$ y $1/4$.

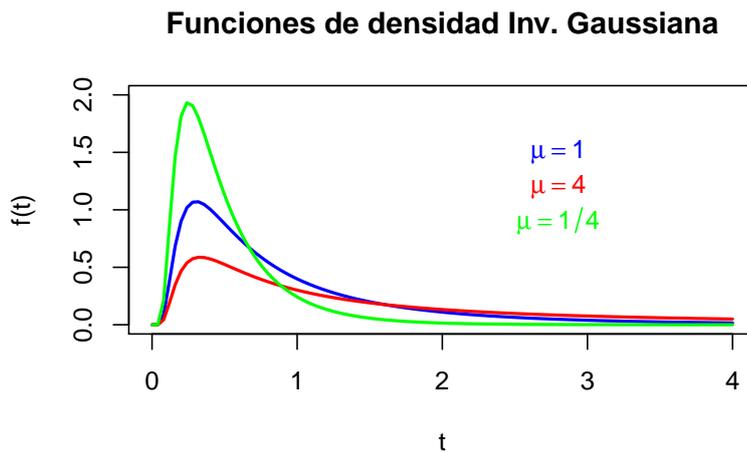


Figura 3.2: Funciones de densidad de la Inv. Gaussiana para $\lambda = 1$ y $\mu = 1, 4$ y $1/4$.

3.1 Propiedades básicas

En la distribución Inversa Gaussiana, el caso $\mu = 1$ se puede considerar como una forma estándar de la distribución. Sin embargo, en algunas aplicaciones físicas puede interesar también tomar $\lambda = 1$ y variar μ . Podemos ver ejemplos de ambos casos en las Figuras 3.1 y 3.2.

Funciones de hazard Inv. Gaussiana

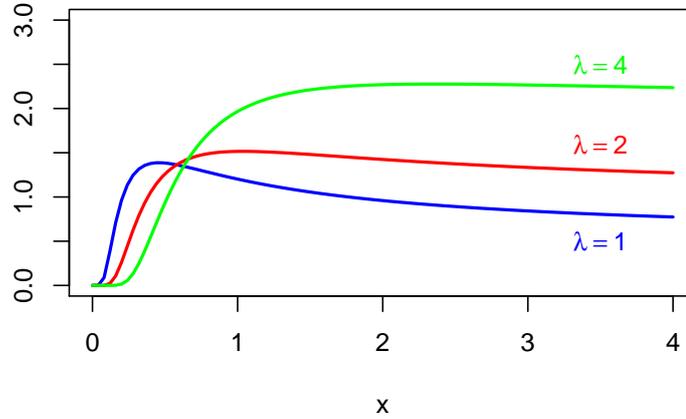


Figura 3.3: Función de hazard de la Inv. Gaussiana para $\mu = 1$ y $\lambda = 1, 2$ y 4

La distribución Inversa Gaussiana posee momentos positivos y negativos de todos los órdenes. Su función generatriz de momentos viene dada por

$$M_T(r) = \exp \left\{ \frac{\lambda}{\mu} \left(1 - \sqrt{1 - \frac{2\mu^2 r}{\lambda}} \right) \right\}, \quad r \in \mathbb{R}, \quad (3.6)$$

mientras que su función característica se expresa como

$$\varphi_T(r) = E[e^{irT}] = \exp \left\{ \frac{\lambda}{\mu} \left(1 - \sqrt{1 - \frac{2i\mu^2 r}{\lambda}} \right) \right\}, \quad r \in \mathbb{R}. \quad (3.7)$$

Estos resultados han sido obtenidos de [11].

A partir de la FGM, se tienen los siguientes resultados.

Proposición 3.1. Si $T \sim IG(\mu, \lambda)$, se verifica

1. Media

$$E(T) = \mu \quad (3.8)$$

2. Varianza

$$Var(T) = \mu^3 / \lambda \quad (3.9)$$

Demostración.

1. Derivando la FGM (3.6) respecto de r , se obtiene

$$\begin{aligned} \frac{dM_T}{dr}(r) &= \exp\left\{\frac{\lambda}{\mu}\left(1 - \sqrt{1 - \frac{2\mu^2 r}{\lambda}}\right)\right\} \frac{\lambda}{\mu} \frac{\frac{2\mu^2}{\lambda}}{2\sqrt{1 - \frac{2\mu^2 r}{\lambda}}} = \\ &= \exp\left\{\frac{\lambda}{\mu}\left(1 - \sqrt{1 - \frac{2\mu^2 r}{\lambda}}\right)\right\} \mu \left(1 - \frac{2\mu^2 r}{\lambda}\right)^{-1/2}. \end{aligned}$$

Evaluando en $r=0$, se obtiene $E(T) = \mu$.

2. Calculamos la derivada segunda de $M_T(r)$:

$$\begin{aligned} \frac{d^2 M_T}{dr^2}(r) &= \exp\left\{\frac{\lambda}{\mu}\left(1 - \sqrt{1 - \frac{2\mu^2 r}{\lambda}}\right)\right\} \mu^2 \left(1 - \frac{2\mu^2 r}{\lambda}\right)^{-1} \\ &+ \exp\left\{\frac{\lambda}{\mu}\left(1 - \sqrt{1 - \frac{2\mu^2 r}{\lambda}}\right)\right\} \mu \left(-\frac{1}{2}\right) \left(1 - \frac{2\mu^2 r}{\lambda}\right)^{-3/2} \left(-\frac{2\mu^2}{\lambda}\right) \end{aligned}$$

Evaluando en $r=0$, se obtiene

$$E(T^2) = \mu^2 + \mu \left(-\frac{1}{2}\right) \left(-\frac{2\mu^2}{\lambda}\right) = \mu^2 + \frac{\mu^3}{\lambda},$$

de donde

$$Var(T) = E(T^2) - E(T)^2 = \mu^2 + \frac{\mu^3}{\lambda} - \mu^2 = \frac{\mu^3}{\lambda}.$$

La siguiente proposición muestra una relación interesante entre los momentos y los momentos inversos de la distribución. Este resultado se ha obtenido de [14].

Proposición 3.2. Sea $T \sim IG(\mu, \lambda)$. Entonces, se cumple la siguiente relación.

$$E[T^{-r}] = \frac{E[T^{r+1}]}{\mu^{2r+1}}, \quad r \in \mathbb{R}. \quad (3.10)$$

La **moda** de la distribución viene dada por

$$t_{moda} = \frac{\mu}{2\lambda} \left(\sqrt{4\lambda^2 + 9\mu^2} - 3\mu \right). \quad (3.11)$$

La **función generatriz de cumulantes** se corresponde con el logaritmo neperiano de la función característica (3.7). De esta manera, su expresión es

$$\psi(r) = \log \varphi_T(r) = \frac{\lambda}{\mu} \left(1 - \sqrt{1 - \frac{2i\mu^2 r}{\lambda}} \right), \quad r \in \mathbb{R}. \quad (3.12)$$

La expresión de los cumulantes de orden t será por tanto la que sigue

$$\kappa_t = \frac{1}{i^t} \frac{d^t}{dr^t} (\log \varphi(r)) \Big|_{r=0} \quad (3.13)$$

Los cumulantes nos dan una forma más sencilla de calcular propiedades de interés de la distribución, como la media, la varianza y los coeficientes de asimetría y de kurtosis.

Proposición 3.3. Sea $T \sim IG(\mu, \lambda)$. Se verifica

1. $\kappa_1 = \mu$.
2. $\kappa_2 = \mu^3 / \lambda$.
3. En general, para $t \geq 2$, se tiene

$$\kappa_t = 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2t - 3) \mu^{2t-1} \lambda^{1-t}.$$

Demostración.

1. Usamos la función generatriz de cumulantes (3.12).

$$\frac{d\psi}{dr} = \frac{\lambda}{\mu} \frac{2\mu^2/\lambda}{2\sqrt{1 - \frac{2\mu^2 r}{\lambda}}}.$$

Evaluando en $r = 0$, se tiene $\kappa_1 = \mu$.

2. De igual manera,

$$\frac{d^2\psi}{dr^2} = -\frac{\mu}{2} \left(1 - \frac{2\mu^2 r}{\lambda} \right)^{-3/2} \left(-\frac{2\mu^2}{\lambda} \right).$$

Tomando $r = 0$, efectivamente $\kappa_2 = \mu^3 / \lambda$.

3. Ver [19].

Observación 3.3. Se tiene que

$$\frac{\kappa_2}{\kappa_1^3} = \frac{1}{\lambda}.$$

Por ello, se dice que λ es una medida inversa de precisión relativa, lo cual nos puede ser útil para posibles interpretaciones de λ .

| Proposición 3.4. Sea $\phi = \frac{\lambda}{\mu}$. Entonces, los coeficientes de asimetría y de kurtosis de la distribución vienen dados por

1. Coeficiente de asimetría

$$\gamma_1 = 3\phi^{-1/2}. \quad (3.14)$$

2. Coeficiente de kurtosis

$$\gamma_2 = 15\phi^{-1}. \quad (3.15)$$

Demostración.

Se tiene $\kappa_3 = 3\mu^5/\lambda^2$ y $\kappa_4 = 15\mu^2/\lambda^3$. Usando esto:

$$\gamma_1 = \kappa_3/\kappa_2^{3/2} = 3(\mu/\lambda)^{1/2}$$

$$\gamma_2 = \kappa_4/\kappa_2^{-2} = 15(\mu/\lambda)$$

Cuando ϕ aumenta, la distribución se asemeja a la normal. Es por tanto un parámetro de forma o parámetro de normalidad. El siguiente resultado especifica esta relación.

| Proposición 3.5. Si $T \sim IG(\mu, \lambda)$, se tiene que cuando $\phi \rightarrow \infty$, entonces

$$T \xrightarrow{\mathcal{L}} N\left(\mu, \frac{\mu^3}{\lambda}\right).$$

Un cambio de escala en la distribución conserva el modelo, salvo porque los parámetros se ven también multiplicados por el mismo factor que la variable. Lo vemos en la siguiente proposición.

| Proposición 3.6. Sea $T \sim IG(\mu, \lambda)$. Entonces, $aT \sim IG(a\mu, a\lambda)$.

Demostración. Llamamos $Y = aT$. Entonces, su FGM vendrá dada por

$$\begin{aligned} M_Y(r) &= E[e^{rY}] = M_T(ar) = \exp\left\{\frac{\lambda}{\mu}\left(1 - \sqrt{1 - \frac{2\mu^2 ar}{\lambda}}\right)\right\} \\ &= \exp\left\{\frac{a\lambda}{a\mu}\left(1 - \sqrt{1 - \frac{2\mu^2 a^2 r}{a\lambda}}\right)\right\}, \end{aligned}$$

de donde $Y \sim IG(a\mu, a\sigma)$

| Teorema 3.1. Si $T \sim IG(\mu, \lambda)$, su fdd $f(t)$ pertenece a la familia exponencial, y se puede expresar como

$$\begin{aligned} f(t, \mu, \lambda) &= \exp \left\{ -\frac{\lambda(t^2 - 2t\mu + \mu^2)}{2t\mu^2} + \frac{1}{2} \log(\lambda) - \frac{1}{2} \log(2\pi t^3) \right\} = \\ &= \exp \left\{ -\frac{\lambda}{2\mu^2} \left[t - 2\mu + \frac{\mu^2}{t} \right] + \frac{1}{2} \log(\lambda) - \frac{1}{2} \log(2\pi t^3) \right\}. \end{aligned}$$

3.2 Inferencia de la distribución Inversa Gaussiana

Sea una v.a. T que sigue una distribución $IG(\mu, \lambda)$, con μ y λ desconocidos. En este apartado vamos a calcular $\hat{\mu}$ y $\hat{\lambda}$, los estimadores de los parámetros, mediante el método de máxima verosimilitud, diferenciando entre los casos de muestras completas y censuradas.

Muestras completas

Sea la muestra aleatoria completa (t_1, \dots, t_n) . Tomando la fdd de la distribución dada por (3.2), calculamos la función de verosimilitud

$$L(\mu, \lambda) = \prod_{i=1}^n f_i(t) = \prod_{i=1}^n \left[\left(\frac{\lambda}{2\pi t_i^3} \right)^{1/2} \right] \exp \left\{ \sum_{i=1}^n \frac{-\lambda(t_i - \mu)^2}{2\mu^2 t_i} \right\}, \quad (3.16)$$

y a partir de ella, la función de log-verosimilitud

$$l(\mu, \lambda) = \frac{n}{2} \log \left(\frac{\lambda}{2\pi} \right) - \frac{3}{2} \sum_{i=1}^n \log(t_i) + \sum_{i=1}^n \frac{-\lambda(t_i - \mu)^2}{2\mu^2 t_i}. \quad (3.17)$$

Calculamos las derivadas de (3.17) respecto de μ y λ y las igualamos a cero para obtener $\hat{\mu}$ y $\hat{\lambda}$ respectivamente:

$$\frac{\partial l}{\partial \mu}(\mu, \lambda) = \sum_{i=1}^n \frac{\lambda \mu^2 (t_i - \mu) + \lambda \mu (t_i - \mu)^2}{\mu^4 t_i} = 0,$$

de donde

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n t_i = \bar{t}. \quad (3.18)$$

Por otro lado,

$$\frac{\partial l}{\partial \lambda} = \frac{n}{2\lambda} - \sum_{i=1}^n \frac{(t_i - \mu)^2}{2\mu^2 t_i} = 0 \quad \Rightarrow \quad \hat{\lambda} = n \left(\sum_{i=1}^n \frac{(t_i - \hat{\mu})^2}{\hat{\mu}^2 t_i} \right)^{-1},$$

donde usando que $\hat{\mu} = \bar{t}$,

$$\sum_{i=1}^n \frac{(t_i - \hat{\mu})^2}{\hat{\mu}^2 t_i} = \sum_{i=1}^n \frac{(t_i - \bar{t})^2}{\bar{t}^2 t_i} = \frac{1}{\bar{t}^2} \sum_{i=1}^n (t_i - 2\bar{t}) + \frac{\bar{t}}{\bar{t}^2} \sum_{i=1}^n \frac{\bar{t}^{-2}}{t_i} = -\sum_{i=1}^n \frac{1}{\bar{t}} + \sum_{i=1}^n \frac{1}{t_i}.$$

Por tanto, se tiene

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^n (t_i^{-1} - \bar{t}^{-1})}. \quad (3.19)$$

Los cálculos para obtener la matriz de información de Fisher son muy extensos. Se ha obtenido su expresión del trabajo de Henze [9], y es la siguiente

$$I(\mu, \lambda) = \begin{pmatrix} \lambda\mu^{-3} & 0 \\ 0 & (2\lambda^2)^{-1} \end{pmatrix} \quad (3.20)$$

Del trabajo de Tweedie [19], obtenemos algunos resultados interesantes relacionados con los estimadores de los parámetros, que enunciaremos a continuación. Estos resultados pueden ser útiles para hacer inferencia sobre esta distribución. Dada una muestra aleatoria (T_1, \dots, T_n) de una distribución $IG(\mu, \lambda)$ con $\mu > 0$, $\lambda > 0$, se tiene que el siguiente es un estadístico suficiente y completo para (μ, λ) .

$$W(\underline{T}) = \left(\sum_{i=1}^n T_i, \sum_{i=1}^n \frac{1}{T_i} \right).$$

| Teorema 3.2. Para los estimadores MV de los parámetros μ y λ , $\hat{\mu}$ y $\hat{\lambda}$, se verifica

1. Los estimadores $\hat{\mu}$ y $\hat{\lambda}$ son independientes.
2. $\hat{\lambda}$ es sesgado.
3. $\hat{\lambda}$ verifica que

$$\frac{n\lambda}{\hat{\lambda}} \sim \chi_{n-1}^2,$$

o lo que es lo mismo,

$$\frac{n}{\hat{\lambda}} \sim \frac{1}{\lambda} \chi_{n-1}^2.$$

4. El estimador $\hat{\mu}$ cumple

$$\hat{\mu} = \bar{t} \sim IG(\mu, n\lambda).$$

Muestras censuradas

Consideramos la muestra aleatoria censurada (t_i, δ_i) , $i = 1, \dots, n$, que sigue una distribución $IG(\mu, \lambda)$. Para calcular los estimadores de máxima verosimilitud $\hat{\mu}$ y $\hat{\lambda}$ utilizamos la función de verosimilitud dada por (1.6) y la fdd y función de supervivencia de la distribución, dadas por (3.2) y (3.4) respectivamente.

Sea entonces la función de logverosimilitud

$$l(\mu, \lambda) = \sum_{i=1}^n \delta_i \left\{ \frac{1}{2} \log \frac{\lambda}{2\pi} - \frac{3}{2} \log t_i - \lambda \frac{(t_i - \mu)^2}{2\mu^2 t_i} \right\} + \sum_{i=1}^n (1 - \delta_i) \log \left\{ \Phi(A_i) - \exp\left(\frac{2\lambda}{\mu} \Phi(B_i)\right) \right\}, \quad (3.21)$$

donde $A_i = -\sqrt{\frac{\lambda}{t_i}}\left(\frac{t_i}{\mu} - 1\right)$ y $B_i = -\sqrt{\frac{\lambda}{t_i}}\left(\frac{t_i}{\mu} + 1\right)$.

A partir de esto, podemos calcular sus ecuaciones de verosimilitud. La primera de ellas vendría dada por

$$\frac{\partial l}{\partial \mu}(\mu, \lambda) = \sum_{i=1}^n \delta_i \frac{\lambda(t_i - \mu)}{\mu^3} + \sum_{i=1}^n (1 - \delta_i) \frac{\partial \log S(t_i)}{\partial \mu}, \quad (3.22)$$

donde

$$\frac{\partial \log S(t)}{\partial \mu} = \frac{1}{\mu^2 S(t)} \left\{ \sqrt{\lambda t} \phi \left[\sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} - 1 \right) \right] + e^{2\lambda/\mu} \left[2\lambda \Phi \left(-\sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} + 1 \right) \right) - \sqrt{\lambda t} \phi \left(\sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} + 1 \right) \right) \right] \right\}.$$

Por otra parte, la segunda ecuación de verosimilitud es

$$\frac{\partial l}{\partial \lambda}(\mu, \lambda) = \sum_{i=1}^n \delta_i \left(\frac{1}{2\lambda} - \frac{(t_i - \mu)^2}{2\mu^2 t_i} \right) + \sum_{i=1}^n (1 - \delta_i) \frac{\partial \log S(t_i)}{\partial \lambda}, \quad (3.23)$$

donde

$$\frac{\partial \log S(t)}{\partial \lambda} = \frac{1}{S(t)} \left\{ \frac{-1}{2\sqrt{\lambda t}} \left(\frac{t}{\mu} - 1 \right) \phi \left(\sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} - 1 \right) \right) - e^{2\lambda/\mu} \left[\frac{2}{\mu} \Phi \left(-\sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} + 1 \right) \right) - \frac{1}{2\sqrt{\lambda t}} \left(\frac{t}{\mu} + 1 \right) \phi \left(\sqrt{\frac{\lambda}{t}} \left(\frac{t}{\mu} + 1 \right) \right) \right] \right\}.$$

Las ecuaciones de verosimilitud (3.22) y (3.23) nos darán la expresión de los estimadores de máxima verosimilitud $\hat{\mu}$ y $\hat{\lambda}$ utilizando algún método numérico, ya que no tienen una forma explícita.

3.3 Aplicación

Para aplicar los resultados de este capítulo, consideramos un conjunto de datos obtenido del paquete "invGauss" de R (ver [8]), que originalmente fueron cedidos por el *Radiation Therapy Oncology Group*, en Estados Unidos. En la muestra, se estudia a 192 pacientes de cáncer de orofaringe, y se recoge el tiempo de vida en años de los pacientes desde su diagnóstico. Estos datos se encuentran en el Cuadro 3.1, donde los asteriscos (*) representan los datos censurados. Para esta aplicación vamos a tomar solamente los datos observados. Lamentablemente, no hemos encontrado un paquete de R que nos permita calcular los estimadores para datos censurados con la parametrización de la distribución utilizada en este trabajo, por lo que no ha sido posible realizar una aplicación sobre este tipo de datos.

Aplicación. Muestra completa

Tomamos los datos no censurados del Cuadro 3.1, formando así una muestra completa de 139 observaciones que suponemos que es una m.a.s. de una distribución Inversa Gaussiana. El histograma de esta muestra, en la Figura 3.4, nos permite observar una asimetría muy grande a la derecha.

Procedemos a estimar los parámetros μ y λ por el método de máxima verosimilitud, cuya expresión viene dada en (3.18) y (3.19). Resulta entonces

$$\hat{\mu} = 1,067311$$

$$\hat{\lambda} = 0,888078$$

Por tanto, podemos suponer que la muestra sigue aproximadamente una distribución $IG(1'07, 0'89)$. Viendo su representación en la Figura 3.5, podemos tomar el

Años de vida tras el diagnóstico							
1,73	0,74	0,90	0,67	2,51	4,99*	1,75	0,64
0,70	0,50	2,92	1,13	0,59	0,89	1,32	0,67
4,29*	1,53	1,03	2,50	0,76	0,39	2,99	0,26
0,48	4,03*	1,44	0,47	1,58	0,61	0,46	4,29
0,70	0,37	1,11	4,10*	0,44	0,72	0,84	2,14
1,81	1,50	4,84*	1,02	4,08*	3,96*	0,20	4,41*
0,82	0,90	1,26	1,22	4,50*	1,35	0,76	2,51
0,62	0,35	4,31	1,54	1,01	2,21	0,53	0,75
3,77*	1,12	2,55	1,50	3,61*	3,61*	1,42	3,58*
0,63	2,09	0,47	3,99*	3,38*	1,49	2,19	4,00*
2,15	1,96	0,93	1,18	3,59*	0,96	0,56	3,34*
0,03	1,82	0,40	2,90*	1,31	2,90*	3,59*	1,91
0,31	0,84	0,04	0,36	0,81	0,80	1,49	2,98*
3,42*	0,40	1,99	0,85	1,64	2,73*	2,98*	1,05
2,55*	0,72	0,03	2,50*	0,24	1,44	1,46*	1,75
0,31	3,00*	0,47	2,58*	0,52	2,54*	2,52*	2,26*
0,27	0,27	2,56*	1,26	0,95	1,02	2,00*	0,65
1,62*	0,60	1,27	1,52	1,46	0,42	1,01	1,48
0,29	2,34*	2,25*	2,12	0,92	1,41	2,50*	2,07
2,18*	0,29	2,01*	1,64*	0,87	1,12	0,95	1,42
1,08	0,22	1,67	2,08*	0,94	0,89	0,70	2,06*
0,92	0,75	1,50*	0,31	0,50*	0,57	0,57	0,48
1,78*	1,84*	0,80	1,98*	1,36	0,76*	0,25*	0,58
0,10	0,35	1,22*	0,44	0,60	0,47	1,13*	0,75

Cuadro 3.1: Datos de pacientes de cáncer de orofaringe.

ajuste como bueno.

Podemos estimar la moda de la distribución utilizando (3.11), resultando

$$\hat{t}_{moda} = 0,2762016.$$

Lo que podemos concluir a partir de la moda es que usualmente, ese es el tiempo que tarda en fallecer un paciente de este tipo de cáncer, ya que es el valor más repetido.

Por otra parte, la varianza de la distribución se puede estimar usando (3.9), de donde obtenemos

$$\hat{\sigma}^2 = \widehat{Var}(T) = 1,369056.$$

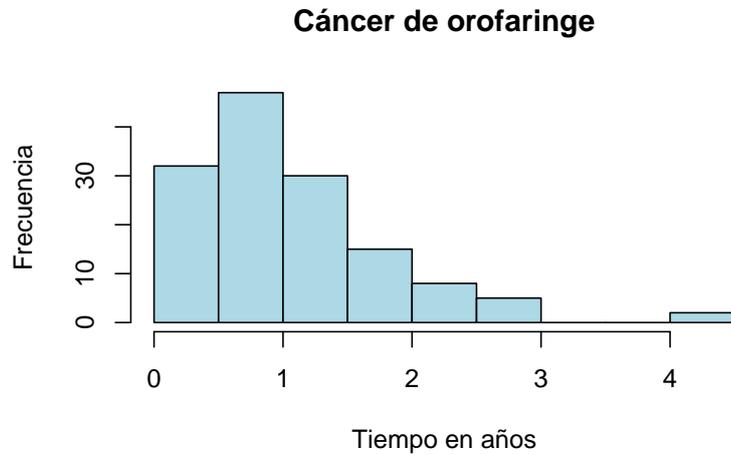


Figura 3.4: Histograma de datos no censurados del Cuadro 3.1.

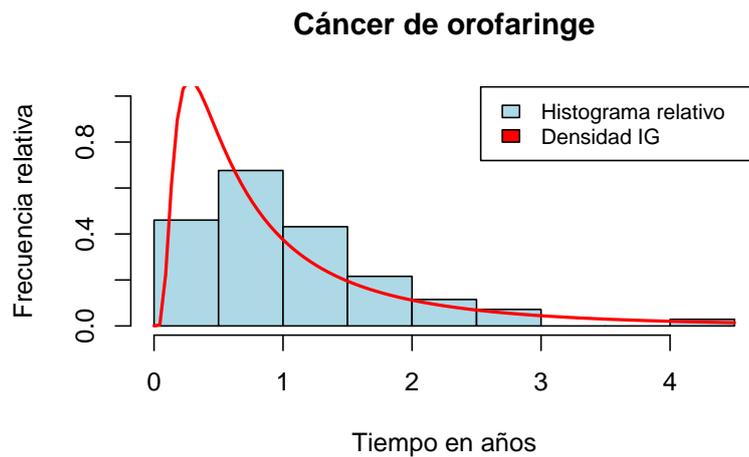


Figura 3.5: Histograma relativo de datos y densidad $IG(\hat{\mu}, \hat{\lambda})$.

La varianza tiene un valor bajo, ya que tal y como se ve en el histograma de la Figura 3.4, no hay mucha variabilidad de los datos. De (3.20), podemos calcular la matriz de información de Fisher estimada, que sería

$$\hat{I} = \begin{pmatrix} 0,7304302 & 0 \\ 0 & 0,6339686 \end{pmatrix}.$$

Calculando su inversa, se obtiene la matriz de covarianzas estimada, resultando

$$\hat{V} = \begin{pmatrix} 1,369056 & 0 \\ 0 & 1,577365 \end{pmatrix},$$

donde los elementos de la diagonal son las varianzas estimadas de $\hat{\mu}$ y $\hat{\lambda}$, respectivamente, lo cual nos da una idea de la variabilidad de estos estimadores. Es fácil estimar también el error estándar de cada estimador, tomando la raíz cuadrada de la diagonal de \hat{V} . Se tiene entonces

$$\sigma_{\mu} = 1,170067, \quad \sigma_{\lambda} = 1,255932.$$

Son también valores bajos, lo que nos da la idea de que las estimaciones no se alejan mucho del valor real de los parámetros.

4 | Distribución Birnbaum-Saunders

4.1 Introducción

La distribución Normal es la más conocida y utilizada de todas las distribuciones estadísticas. Ya hemos visto en capítulos anteriores dos ejemplos de distribuciones de tiempo de vida relacionadas con la Normal. Otra de ellas es la distribución de Birnbaum-Saunders, que abreviaremos como BS. Es una distribución de dos parámetros que puede expresarse como una transformación monótona de la distribución Normal Estándar. Aunque ya se conocía esta distribución, en 1969 fueron Birnbaum y Saunders quienes le dieron una justificación natural física: modelar el tiempo de vida de un metal sujeto a estrés periódico. Por este motivo, la distribución recibe su nombre. Desde entonces se ha trabajado mucho en esta distribución, hallando así más campos de aplicación y nuevas propiedades de esta, tal y como veremos en este capítulo. Entre estas nuevas conclusiones, se encuentra una importante relación entre la distribución Inversa Gaussiana (IG) y la BS, ya que esta última se puede obtener como una aproximación de la primera. También se puede ver la distribución BS como una mezcla entre la IG y su recíproca, lo cual provoca que la distribución BS herede ciertas propiedades ya conocidas de la IG y es útil para estudiar su inferencia.

La función de densidad y la función de hazard de esta distribución también son objetos interesantes de estudio. Se ha observado que la fdd de la BS es unimodal, al igual que su función de hazard. La función de hazard de las distribuciones IG y BS tienen comportamientos parecidos, ya que cuando $t \rightarrow \infty$, tras llegar a su máximo ambas decrecen hacia una constante, mientras que la de la distribución log-normal tiende a cero.

En este capítulo estudiaremos también los estimadores de máxima verosimilitud de los parámetros de forma y escala (que notaremos α y β respectivamente) basados en muestras completas y censuradas, y veremos que estos no pueden obtenerse de forma explícita, sino resolviendo una ecuación no lineal.

4.2 Interpretación física

Definición 4.1. Decimos que una variable aleatoria T sigue una **distribución Birnbaum-Saunders (BS) de dos parámetros** α y β , y se notará como $T \sim BS(\alpha, \beta)$, si su FdD puede escribirse como

$$F_T(t; \alpha, \beta) = \Phi \left[\frac{1}{\alpha} \left\{ \left(\frac{t}{\beta} \right)^{1/2} - \left(\frac{\beta}{t} \right)^{1/2} \right\} \right], \quad 0 < t < \infty, \quad \alpha, \beta > 0, \quad (4.1)$$

donde Φ es la FdD de la distribución Normal Estándar y los parámetros α y β son los parámetros de **forma** (ya que influye en la forma de la distribución) y **escala**, respectivamente.

Si llamamos

$$\xi(u) = u^{1/2} - u^{-1/2}, \quad u > 0, \quad (4.2)$$

cuya derivada viene dada por

$$\xi'(u) = \frac{1}{2} \left(u^{-1/2} + u^{-3/2} \right), \quad u > 0,$$

entonces podemos reescribir la FdD de la distribución como

$$F_T(t; \alpha, \beta) = \Phi \left[\frac{1}{\alpha} \xi \left(\frac{t}{\beta} \right) \right], \quad 0 < t < \infty, \quad \alpha, \beta > 0. \quad (4.3)$$

La correspondiente fdd de la distribución se define como

$$f_T(t; \alpha, \beta) = \frac{1}{2\sqrt{2\pi}\alpha\beta} \left[\left(\frac{\beta}{t} \right)^{1/2} + \left(\frac{\beta}{t} \right)^{3/2} \right] \exp \left[-\frac{1}{2\alpha^2} \left(\frac{t}{\beta} + \frac{\beta}{t} - 2 \right) \right], \quad t, \alpha, \beta > 0. \quad (4.4)$$

Utilizando $\xi(\cdot)$ vista en (4.2), podemos simplificar (4.4) como sigue

$$f_T(t; \alpha, \beta) = \phi \left[\frac{1}{\alpha} \xi \left(\frac{t}{\beta} \right) \right] \frac{1}{\alpha} \xi' \left(\frac{t}{\beta} \right), \quad 0 < t < \infty, \quad \alpha, \beta > 0. \quad (4.5)$$

Esta distribución de probabilidad en dos parámetros se basa en el proceso de rotura por fatiga de un material. Trata de modelar el número de ciclos necesarios para que se produzca dicha rotura bajo los siguientes supuestos:

1. El fallo por fatiga se produce debido a la aplicación repetida de un patrón de estrés (tensión) cíclico.
2. A causa de este estrés cíclico, se produce una grieta en el material que va creciendo hasta alcanzar un tamaño crítico ω , en el cual ocurre el fallo por fatiga.
3. La extensión de la rotura en cada ciclo es una variable aleatoria con las mismas media y varianza.
4. Las extensiones de la fisura en cada ciclo son estadísticamente independientes.
5. Por el Teorema Central del Límite (TCL), la distribución de la extensión total de la rotura tras un gran número de ciclos será aproximadamente normal.

En particular, en el j -ésimo ciclo suponemos que la extensión de la grieta, X_j , es una v.a. con media μ_0 y varianza σ_0^2 . Entonces, debido al TCL, se tiene que $\sum_{i=1}^n X_i$ sigue aproximadamente una distribución normal, de media $n\mu_0$ y varianza $n\sigma_0^2$. Por tanto, la probabilidad de que la fisura no sobrepase el umbral ω vendrá dada por

$$\Phi\left(\frac{\omega - n\mu_0}{\sigma_0\sqrt{n}}\right) = \Phi\left(\frac{\omega}{\sigma_0\sqrt{n}} - \frac{\mu_0\sqrt{n}}{\sigma_0}\right). \quad (4.6)$$

Si denotamos como T el tiempo de vida del material bajo estudio, es decir, el número de ciclos hasta la rotura, entonces la FdD de T será aproximadamente la siguiente

$$P(T \leq t) = F_T(t) \approx 1 - \Phi\left(\frac{\omega}{\sigma_0\sqrt{n}} - \frac{\mu_0\sqrt{n}}{\sigma_0}\right) = \Phi\left(\frac{\mu_0\sqrt{n}}{\sigma_0} - \frac{\omega}{\sigma_0\sqrt{n}}\right), \quad t > 0. \quad (4.7)$$

Si suponemos que (4.7) es el modelo exacto del tiempo de vida, es claro por (4.1) que T sigue una distribución *BS* donde

$$\beta = \frac{\omega}{\mu_0}, \quad \alpha = \frac{\sigma_0}{\sqrt{\omega\mu_0}}.$$

4.3 Propiedades básicas

La siguiente proposición nos da algunas propiedades importantes de la distribución relacionadas con el comportamiento de sus funciones de densidad y de distribución, que podemos ver gráficamente en la Figura 4.1.

| Proposición 4.1.

1. La fdd (4.4) tiende a 0 tanto cuando $t \rightarrow 0^+$ como cuando $t \rightarrow \infty$.

2. Para todos los valores de α y β , la fdd es unimodal, continua y asimétrica a la derecha.
3. β es la mediana de la distribución.
4. El parámetro de forma α varía la asimetría y la kurtosis.
5. Cuando α tiende a cero, la distribución tiende a ser simétrica respecto de β .

Demostración.

1. Ver Leiva [13].
2. Ver también Leiva [13].
3. La mediana es el valor $t_{0.5}$ tal que $F_T(t_{0.5}) = 0,5$. Entonces, utilizando la expresión de F_T dada por (4.3) se tiene

$$F_T(t) = 0.5 \Leftrightarrow \Phi \left[\frac{1}{\alpha} \xi \left(\frac{t}{\beta} \right) \right] = 0.5 \Leftrightarrow \xi \left(\frac{t}{\beta} \right) = 0 \Leftrightarrow t = \beta,$$

por la definición de $\xi(\cdot)$ en (4.2).

4. Se tiene de forma inmediata ya que, como se verá más adelante, tanto el coeficiente de asimetría (4.18) como el de kurtosis (4.19) dependen solo del parámetro α .
5. En secciones posteriores se probará que la media de esta distribución viene dada por $E(T) = \beta \left(1 + \frac{\alpha^2}{2} \right)$. Es fácil ver que cuando $\alpha \rightarrow 0$, la media es igual a β , siendo esta la mediana de la distribución, junto a los hechos de que el coeficiente de asimetría tiende a 0 cuando $\alpha \rightarrow 0$ y el coeficiente de kurtosis tiende a 3 cuando $\alpha \rightarrow 0$, por tanto, se tiene que la distribución es simétrica respecto de β cuando α tiende a cero.

■

Moda. La moda de la $BS(\alpha, \beta)$ no se puede obtener de forma explícita. Para hallarla, hay que calcular primero m_α , la moda de la $BS(\alpha, 1)$, para luego obtener el producto βm_α . La manera de calcular m_α es resolver la ecuación

$$t^3 + t^2(\alpha^2 + 1) + t(3\alpha^2 - 1) - 1 = 0,$$

que tiene solución única. Se cumple que m_α es una función creciente de α (ver [10]).

La siguiente propiedad de la distribución nos da una manera de relacionar la BS con la distribución normal, que utilizaremos para resultados posteriores.

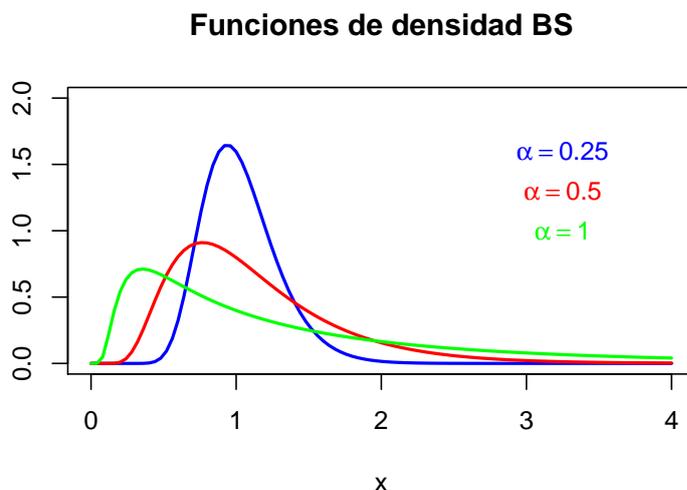


Figura 4.1: Función de densidad de la BS para $\beta = 1$ y $\alpha = 0.25, 0.5$ y 1 .

Proposición 4.2. Sea $T \sim BS(\alpha, \beta)$. Entonces, se puede expresar

$$T = h(Z) = \frac{\beta}{4} \left[\alpha Z + \sqrt{(\alpha Z)^2 + 4} \right]^2, \quad (4.8)$$

donde h es una función monótona creciente y $Z \sim N(0, 1)$.

Corolario 4.1. De la transformación (4.8) se deducen las siguientes igualdades:

1.

$$Z = \frac{1}{\alpha} \left[\left(\frac{T}{\beta} \right)^{1/2} - \left(\frac{T}{\beta} \right)^{-1/2} \right], \quad (4.9)$$

2.

$$W = Z^2 = \frac{1}{\alpha^2} \left[\frac{h(Z)}{\beta} + \frac{\beta}{h(Z)} - 2 \right] \sim \chi_1^2. \quad (4.10)$$

Con este resultado se observa que podemos generar variables aleatorias que sigan una distribución *BS* usando tanto (4.8) como (4.10). Para las pruebas de estos resultados, ver Leiva [13].

Corolario 4.2. El **cuantil de orden p** de una v.a. $T \sim BS(\alpha, \beta)$ viene dado por

$$t_p = \frac{\beta}{4} \left[\alpha z_p + \sqrt{(\alpha z_p)^2 + 4} \right]^2, \quad (4.11)$$

donde z_p es el cuantil de orden p de la v.a. normal estándar Z .

Demostración.

Por la definición de cuantil de $Z \sim N(0, 1)$, se cumple que $\Phi(z_p) = p$. Por ser h monótona creciente se tiene

$$P(Z \leq z_p) = p \Leftrightarrow P(h(Z) \leq h(z_p)) = p,$$

|

4.3.1 Características de las distribuciones de tiempos de vida

Las funciones de supervivencia y de riesgo son unos indicadores muy útiles en el estudio de tiempos de vida. En la siguiente proposición veremos las correspondientes expresiones de ambas funciones en la distribución que estamos estudiando.

Proposición 4.3. Sea $T \sim BS(\alpha, \beta)$. Entonces, las expresiones de la función de supervivencia $S(t)$ y la función de hazard $h_T(t)$ son

1.

$$S(t; \alpha, \beta) = \Phi \left[-\frac{1}{\alpha} \xi \left(\frac{t}{\beta} \right) \right], \quad t > 0. \quad (4.12)$$

2.

$$h_T(t; \alpha, \beta) = \frac{\phi \left[\frac{1}{\alpha} \xi \left(\frac{t}{\beta} \right) \right] \frac{1}{\alpha} \xi' \left(\frac{t}{\beta} \right)}{\Phi \left[-\frac{1}{\alpha} \xi \left(\frac{t}{\beta} \right) \right]}, \quad t > 0, \quad (4.13)$$

donde $\xi(\cdot)$ se definió en (4.2).

Demostración.

1. S se define como $S(t) = 1 - F(t)$. Por tanto,

$$S(t; \alpha, \beta) = 1 - \Phi \left[\frac{1}{\alpha} \xi \left(\frac{t}{\beta} \right) \right] = \Phi \left[-\frac{1}{\alpha} \xi \left(\frac{t}{\beta} \right) \right].$$

2. Por definición, $h(t) = \frac{f(t)}{S(t)}$. Entonces, por el apartado anterior y por (4.5), se tiene el resultado.

|

La forma de la función de hazard no depende del parámetro de escala β , luego podemos tomar $\beta = 1$ sin pérdida de generalidad.

Algunas de las propiedades de la función de hazard, que pueden observarse en la Figura 4.2, son

- La función de hazard h_T en (4.13) es unimodal para cualquier α .
- Se cumple que

$$\lim_{t \rightarrow \infty} \log(h_T(t; \alpha, 1)) = \frac{1}{2\alpha^2}.$$

- h_T tiende a ser creciente cuando α tiende a cero.

Para más detalles sobre estas propiedades, ver [4].

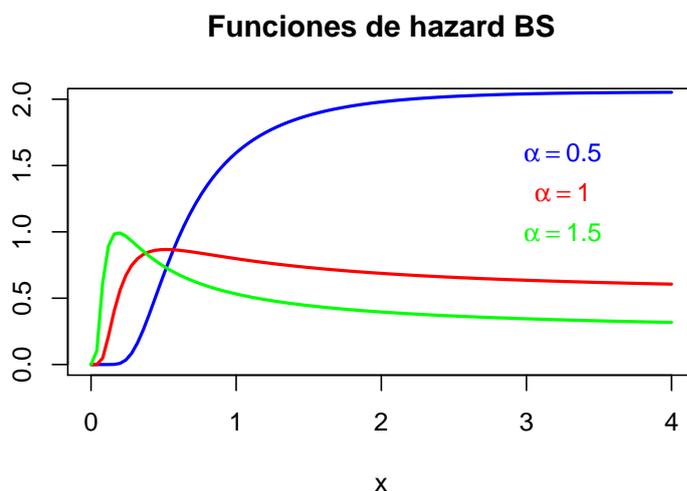


Figura 4.2: Función de hazard de la BS para $\beta = 1$ y $\alpha = 0.5, 1$ y 1.5 .

Definición 4.2. Dada una serie de datos, se llama **punto de cambio** o **change point** al punto a partir del cual se ha producido un cambio en la estructura del modelo de datos bajo estudio.

El punto de cambio de la función de hazard en (4.13), que notaremos como c_α , se puede obtener resolviendo la siguiente ecuación no lineal

$$\Phi\left(-\frac{1}{\alpha}\xi(t)\right)\left\{-\left(\xi'(t)\right)^2\xi(t) + \alpha^2\xi''(t)\right\} + \alpha\phi\left(-\frac{1}{\alpha}\xi(t)\right)\left(\xi'(t)\right)^2 = 0. \quad (4.14)$$

Como β es el parámetro de escala, el punto de cambio de la función de hazard en (4.13), que notaremos como $c_{\alpha,\beta}$ se puede calcular de forma inmediata como $c_{\alpha,\beta} = \beta c_\alpha$.

4.4 Momentos y momentos inversos

Proposición 4.4. *Sea T una v.a. tal que $T \sim BS(\alpha, \beta)$ y sea un entero r no negativo. Entonces, el momento no central de orden r vendrá dado por*

$$E(T^r) = \beta^r \sum_{j=0}^r \binom{2r}{2j} \sum_{k=0}^j \binom{j}{k} \frac{(2r-2j+2k)!}{2^{r-j+k}(r-j+k)!} \left(\frac{\alpha}{2}\right)^{2r-2j+2k}. \quad (4.15)$$

Demostración.

Partiendo de (4.8), obtenemos la expresión

$$E\left[\left(\frac{T}{\beta}\right)^r\right] = E\left[\left(\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1}\right)^{2r}\right].$$

Si aplicamos a esta expresión el binomio de Newton, esto es,

$$(a+b)^m = \sum_{k=0}^m \binom{m}{k} a^{m-k} b^k,$$

se tiene que

$$E\left[\left(\frac{T}{\beta}\right)^r\right] = \sum_{k=0}^{2r} \binom{2r}{k} E\left[\left(\left(\frac{\alpha Z}{2}\right)^2 + 1\right)^{k/2} \left(\frac{\alpha Z}{2}\right)^{2r-k}\right].$$

Observamos que

$$E\left\{\left[\left(\frac{\alpha Z}{2}\right)^2 + 1\right]^l \left(\frac{\alpha Z}{2}\right)^s\right\} = 0$$

si s es impar. Por tanto,

$$E\left[\left(\frac{T}{\beta}\right)^r\right] = \sum_{k=0}^r \binom{2r}{2k} E\left\{\left[\left(\frac{\alpha Z}{2}\right)^2 + 1\right]^k \left(\frac{\alpha Z}{2}\right)^{2(r-k)}\right\}.$$

Usando de nuevo el binomio de Newton:

$$\left[\left(\frac{\alpha Z}{2}\right)^2 + 1\right]^j = \sum_{k=0}^j \binom{j}{k} \left(\frac{\alpha Z}{2}\right)^{2k},$$

de donde

$$\begin{aligned} E\left[\left(\frac{T}{\beta}\right)^r\right] &= \sum_{j=0}^r \binom{2r}{2j} E\left[\sum_{k=0}^j \binom{j}{k} \left(\frac{\alpha Z}{2}\right)^{2k} \left(\frac{\alpha Z}{2}\right)^{2(r-j)}\right] \\ &= \sum_{j=0}^r \binom{2r}{2j} \sum_{k=0}^j \left(\frac{\alpha}{2}\right)^{2(r-j+k)} E\left[Z^{2(r-j+k)}\right]. \end{aligned}$$

Como $Z \sim N(0, 1)$, entonces

$$E\left(Z^{2(r-j+k)}\right) = \frac{(2(r-j+k))!}{2^{r-j+k}(r-j+k)!},$$

de donde obtenemos finalmente que

$$E\left[\left(\frac{T}{\beta}\right)^r\right] = \sum_{j=0}^r \binom{2r}{2j} \sum_{k=0}^j \binom{j}{k} \frac{(2r-2j+2k)!}{2^{r-j+k}(r-j+k)!} \left(\frac{\alpha}{2}\right)^{2r-2j+2k}.$$

La expresión de los momentos no centrales nos permite calcular fácilmente el valor de la esperanza y varianza de la distribución, así como los coeficientes de asimetría y kurtosis.

Proposición 4.5. Sea $T \sim BS(\alpha, \beta)$. Entonces,

1. Esperanza

$$E(T) = \beta \left(1 + \frac{\alpha^2}{2}\right) \quad (4.16)$$

2. Varianza

$$Var(T) = \beta^2 \alpha^2 \left(1 + \frac{5}{4} \alpha^2\right) \quad (4.17)$$

3. Coeficiente de asimetría

$$\gamma_1(T) = \frac{16\alpha^2(11\alpha^2 + 6)^2}{(5\alpha^2 + 4)^3} \quad (4.18)$$

4. Coeficiente de kurtosis

$$\beta_2(T) = 3 + \frac{6\alpha^2(93\alpha^2 + 40)}{(5\alpha^2 + 4)^2} \quad (4.19)$$

Demostración.

1. Tomamos $r = 1$ en (4.15):

$$\begin{aligned} E\left(\frac{T}{\beta}\right) &= \sum_{j=0}^1 \binom{2}{2j} \sum_{k=0}^j \binom{j}{k} \frac{(2-2j+2k)!}{2^{1-j+k}(1-j+k)!} \left(\frac{\alpha}{2}\right)^{2-2j+2k} \\ &= \binom{2}{0} \frac{2!}{1!2} \left(\frac{\alpha}{2}\right)^2 + \binom{2}{2} \sum_{k=0}^1 \binom{1}{k} \frac{(2k)!}{2^k k!} \left(\frac{\alpha}{2}\right)^{2k} = 1 + \frac{\alpha^2}{2}, \end{aligned}$$

de donde

$$E(T) = \beta \left(1 + \frac{\alpha^2}{2}\right).$$

2. Sabemos que $Var(T) = E(T^2) - E(T)^2$, luego calculamos $E(T^2)$ usando (4.15):

$$\begin{aligned} E\left[\left(\frac{T}{\beta}\right)^2\right] &= \sum_{j=0}^2 \binom{4}{2j} \sum_{k=0}^j \binom{j}{k} \frac{(4-2j+2k)!}{2^{2-j+k}(2-j+k)!} \left(\frac{\alpha}{2}\right)^{4-2j+2k} \\ &= 3\left(\frac{\alpha}{2}\right)^4 + 6\left(\frac{\alpha}{2}\right)^2 + 18\left(\frac{\alpha}{2}\right)^4 + 1 + 2\left(\frac{\alpha}{2}\right)^2 + 3\left(\frac{\alpha}{2}\right)^4 \\ &= \frac{3}{2}\alpha^4 + 2\alpha^2 + 1. \end{aligned}$$

Entonces,

$$E(T^2) = \beta^2 \left(\frac{3}{2}\alpha^4 + 2\alpha^2 + 1\right),$$

luego

$$Var(T) = \beta^2 \left(\frac{3}{2}\alpha^4 + 2\alpha^2 + 1\right) - \left(\beta \left(1 + \frac{\alpha^2}{2}\right)\right)^2 = \beta^2 \alpha^2 \left(1 + \frac{5}{4}\alpha^2\right).$$

3. El coeficiente de asimetría se define de forma general como

$$\gamma_1(T) = \frac{\mu_3}{(\mu_2)^{3/2}}, \quad (4.20)$$

donde $\mu_k = E[(T - E(T))^k]$ es el momento centrado de orden k . Entonces, solo necesitamos hallar μ_3 , ya que se tiene $\mu_2 = Var(T)$.

$$\begin{aligned} E\left[\left(\frac{T}{\beta}\right)^3\right] &= \sum_{j=0}^3 \binom{6}{2j} \sum_{k=0}^j \binom{j}{k} \frac{(4-2j+2k)!}{2^{2-j+k}(2-j+k)!} \left(\frac{\alpha}{2}\right)^{4-2j+2k} \\ &= 1 + \frac{9}{2}\alpha^2 + 9\alpha^4 + \frac{15}{2}\alpha^6, \end{aligned}$$

luego

$$E(T^3) = \frac{\beta^3}{2} (2 + 9\alpha^2 + 18\alpha^4 + 15\alpha^6).$$

μ_3 se puede expresar en función de momentos no centrales como sigue

$$\mu_3 = E[T^3 + 3TE(T)^2 - 3T^2E(T) - E(T)^3],$$

y aplicando propiedades conocidas de la esperanza, resulta

$$\mu_3 = E(T^3) + 2E(T)^3 - 3E(T)E(T^2).$$

Aplicando estos resultados a la definición de γ_1 en (4.20), se tiene el resultado.

4. El coeficiente de kurtosis se expresa como

$$\beta_2(T) = \frac{\mu_4}{\mu_2^2}. \quad (4.21)$$

Siguiendo el mismo proceso de antes, calculamos $E(T^4)$.

$$\begin{aligned} E\left[\left(\frac{T}{\beta}\right)^4\right] &= \sum_{j=0}^4 \binom{6}{2j} \sum_{k=0}^j \binom{j}{k} \frac{(4-2j+2k)!}{2^{2-j+k}(2-j+k)!} \left(\frac{\alpha}{2}\right)^{4-2j+2k} \\ &= 1 + 8\alpha^2 + 30\alpha^4 + 60\alpha^6 + \frac{105}{2}\alpha^8. \end{aligned}$$

Entonces,

$$E(T^4) = \frac{\beta^4}{2}(2 + 16\alpha^2 + 60\alpha^4 + 120\alpha^6 + 105\alpha^8).$$

Por otro lado, se cumple

$$\mu_4(T) = E[T^4 + 6T^2E(T)^2 - 4T^3E(T) - 4TE(T)^3 + E(T)^4],$$

de donde

$$\mu_4(T) = E(T^4) + 6E(T^2)E(T)^2 - 4E(T^3)E(T) - 3E(T)^4.$$

Entonces, sustituyendo los valores necesarios en (4.21), finaliza la prueba. |

Para obtener los **momentos inversos** de la BS, vamos a utilizar la propiedad descrita a continuación, que a su vez nos será útil posteriormente para la inferencia de la distribución.

| Proposición 4.6. *Sea T una v.a. tal que $T \sim BS(\alpha, \beta)$. Entonces, se cumple que tanto T/β como β/T siguen la misma distribución, esto es,*

$$\frac{T}{\beta} \sim BS(\alpha, 1), \quad \frac{\beta}{T} \sim BS(\alpha, 1).$$

Demostración.

Es claro que al ser β el parámetro de escala y ser $T \sim B(\alpha, \beta)$, entonces se tiene que $T/\beta \sim B(\alpha, 1)$.

Para la segunda parte de la prueba, hacemos el cambio de variable $R = \beta/T$. Entonces,

$$\begin{aligned} F_R(r) &= P\left(\frac{\beta}{T} \leq r\right) = P\left(\frac{\beta}{r} \leq T\right) = 1 - \Phi\left[\frac{1}{\alpha} \left\{ \left(\frac{\beta/r}{\beta}\right)^{1/2} - \left(\frac{\beta}{\beta/r}\right)^{1/2} \right\}\right] = \\ &= \Phi\left[\frac{1}{\alpha} \left\{ \left(\frac{r}{1}\right)^{1/2} - \left(\frac{1}{r}\right)^{1/2} \right\}\right], \end{aligned}$$

luego $R = \frac{\beta}{T} \sim B(\alpha, 1)$. |

A partir de esta última proposición, es fácil ver que

$$E(T^{-r}) = \frac{E(T^r)}{\beta^{2r}}, \quad (4.22)$$

relación que nos permitirá calcular de forma simple los momentos no centrados de la v.a. T^{-1} . En la siguiente proposición se darán la media y la varianza de la misma.

| Proposición 4.7. *Dada una v.a. $T \sim BS(\alpha, \beta)$, se tiene que la media y la varianza de T^{-1} vienen dadas por*

1. *Media*

$$E\left(\frac{1}{T}\right) = \frac{1}{\beta} \left(1 + \frac{\alpha^2}{2}\right) \quad (4.23)$$

2. *Varianza*

$$Var\left(\frac{1}{T}\right) = \frac{\alpha^2}{\beta^2} \left(\frac{5}{4}\alpha^2 + 1\right). \quad (4.24)$$

Demostración.

1. Por (4.22), se tiene que

$$E(T^{-1}) = \frac{E(T)}{\beta^2},$$

donde el valor de $E(T)$ viene dado por (4.16). Entonces,

$$E(T^{-1}) = \frac{1}{\beta^2} \beta \left(a + \frac{\alpha^2}{2}\right) = \frac{1}{\beta} \left(1 + \frac{\alpha^2}{2}\right).$$

2. De igual manera, por (4.22),

$$E(T^{-2}) = \frac{E(T^2)}{\beta^4} = \frac{1}{\beta^2} \left(\frac{3}{2} \alpha^4 + 2\alpha^2 + 1 \right).$$

Entonces, como $Var(T^{-1}) = E(T^{-2}) - E(T^{-1})^2$, se tiene finalmente

$$Var(T^{-1}) = \frac{1}{\beta^2} \left(\frac{3}{2} \alpha^4 + 2\alpha^2 + 1 \right) - \left[\frac{1}{\beta} \left(1 + \frac{\alpha^2}{2} \right) \right]^2,$$

de donde se tiene el resultado. |

4.5 Relación con la distribución Inversa Gaussiana

Consideramos dos variables aleatorias X_1 y X_2 tales que $X_1 \sim IG(\mu, \lambda)$ y $X_2^{-1} \sim IG(\mu^{-1}, \lambda\mu^2)$, donde $\mu > 0$ y $\lambda > 0$. Definimos la v.a. T como

$$T = \begin{cases} X_1 & \text{con probabilidad } 1/2 \\ X_2 & \text{con probabilidad } 1/2 \end{cases}$$

Entonces, la fdd de T vendrá dada por

$$f_T(t; \mu, \lambda) = \frac{1}{2} f_{X_1}(t; \mu, \lambda) + \frac{1}{2} f_{X_2}(t; \mu, \lambda), \quad t > 0, \quad (4.25)$$

donde f_{X_1} y f_{X_2} son las funciones de densidad de X_1 y X_2 , respectivamente. La primera viene dada por la expresión (3.2), mientras que la segunda será de la forma $f_{X_2}(t; \mu, \lambda) = t f_{X_1}(t; \mu, \lambda) / \mu$.

Tomando $\alpha = \sqrt{\mu/\lambda}$ y $\beta = \mu$, obtenemos que la v.a. T definida anteriormente sigue una distribución $BS(\alpha, \beta)$. Además, podemos obtener transformaciones de muchas de las propiedades de la IG en las mismas de la BS de forma inmediata, usando la relación vista en (4.25). Un ejemplo de esto es que podemos usarla para obtener la FGM de la distribución BS a partir de las FGMs de la IG . Veámoslo:

Si $T \sim BS(\alpha, \beta)$, su FGM será de la forma

$$M_T(r) = \frac{1}{2} [M_{X_1}(r) + M_{X_2}(r)],$$

siendo M_{X_1} y M_{X_2} son las FGM de X_1 y X_2 , respectivamente, las cuales tienen la siguiente expresión

$$M_{X_1}(r) = \exp \left\{ \frac{\lambda}{\mu} \left(1 - \left(1 - \frac{2\mu^2 r}{\lambda} \right)^{1/2} \right) \right\}, \quad r \in \mathbb{R},$$

$$M_{X_2}(r) = \exp \left\{ \frac{\lambda}{\mu} \left(1 - \left(1 - \frac{2\mu^2 r}{\lambda} \right)^{1/2} \right) \right\} \left(1 - \frac{2\mu^2 r}{\lambda} \right)^{-1/2}, \quad r \in \mathbb{R}.$$

De manera inmediata, se obtiene

$$M_T(r) = \frac{1}{2} \exp \left\{ \frac{\lambda}{\mu} \left(1 - \left(1 - \frac{2\mu^2 r}{\lambda} \right)^{1/2} \right) \right\} \left[1 + \left(1 - \frac{2\mu^2 r}{\lambda} \right)^{-1/2} \right].$$

4.6 Inferencia

En esta sección, estudiaremos distintos métodos de estimación de los parámetros α y β basados primero en una muestra aleatoria completa T_1, \dots, T_n de la v.a. $T \sim BS(\alpha, \beta)$ con fdd f_T y cuyas observaciones son $\{t_1, \dots, t_n\}$. Posteriormente, estudiaremos el cálculo de estos estimadores para el caso de muestras censuradas.

4.6.1 Método de los momentos

Los estimadores de α y β por el método de los momentos se obtienen igualando la media muestral y la varianza muestral a $E(T)$ y $Var(T)$, respectivamente, cuya expresión se vio en (4.16) y (4.17). Denotamos la media y la varianza muestrales como s y v , respectivamente. Entonces, los estimadores por el método de los momentos de α y β se obtienen resolviendo el siguiente sistema:

$$s = \frac{\beta}{2}(\alpha^2 + 2)$$

$$v = \frac{\beta^2}{4}(5\alpha^4 + 4\alpha^2)$$

De aquí, se tiene de forma inmediata que el estimador por el método de los momentos de α se puede obtener como raíz de la siguiente ecuación no lineal

$$\alpha^4(5s^2 - v) + 4\alpha^2(s^2 - 4) - 4v = 0.$$

Entonces, si el coeficiente de variación muestral, σ/\bar{t} , es menor que $\sqrt{5}$, el estimador de α tendrá la siguiente expresión

$$\hat{\alpha}_M = \left[\frac{-2(s^2 - 4) + 2\sqrt{(s^2 - 4)^2 + v(5s^2 - v)}}{5s^2 - v} \right]^{1/2}, \quad (4.26)$$

mientras que el de β será entonces

$$\hat{\beta}_M = \frac{2s}{\hat{\alpha}_M^2 + 2}. \quad (4.27)$$

Sin embargo, si el coeficiente de variación muestral v/s fuera mayor que $\sqrt{5}$, el estimador de α por el método de los momentos podría no existir, y por tanto, ocurriría lo mismo con el de β .

4.6.2 Método de los momentos modificado

Como veremos más tarde, los estimadores MV de α y β no pueden obtenerse de forma explícita, y a su vez, los estimadores por el método de los momentos no tienen por qué existir. Este es el motivo por el que se propone el método de los momentos modificado (MM) para calcular los estimadores de α y β . Para calcularlos, usamos el hecho de que si $T \sim BS(\alpha, \beta)$, entonces $T^{-1} \sim BS(\alpha, \beta^{-1})$ visto en la Proposición 4.6.

Igualamos la media muestral aritmética s y la media muestral armónica r a sus correspondientes expresiones poblacionales en (4.16) y (4.23), respectivamente, donde

$$s = \frac{1}{n} \sum_{i=1}^n t_i \quad \text{y} \quad r = \left[\frac{1}{n} \sum_{i=1}^n t_i^{-1} \right]^{-1}.$$

Resultan entonces las ecuaciones

$$s = \beta \left(1 + \frac{1}{2} \alpha^2 \right), \quad r^{-1} = \beta^{-1} \left(1 + \frac{1}{2} \alpha^2 \right).$$

Así, los estimadores MM de α y β poseen la siguiente expresión

$$\hat{\alpha}_{MM} = \left(2 \left[\left(\frac{s}{r} \right)^{1/2} - 1 \right] \right)^{1/2} \quad \text{y} \quad \hat{\beta}_{MM} = \sqrt{sr}. \quad (4.28)$$

Los estimadores MM siempre existen y tienen forma explícita. Además, utilizando el TCL en $\frac{1}{n} \sum_{i=1}^n T_i$ y $\frac{1}{n} \sum_{i=1}^n T_i^{-1}$, se puede probar que $\hat{\alpha}_{MM}$ y $\hat{\beta}_{MM}$ siguen una distribución normal bivalente (ver [6]), de manera que

$$\begin{pmatrix} \hat{\alpha}_{MM} \\ \hat{\beta}_{MM} \end{pmatrix} \sim N \left[\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \Sigma \right],$$

donde

$$\Sigma = \begin{pmatrix} \frac{\alpha^2}{2n} & 0 \\ 0 & \frac{\alpha\beta^2}{n} \left(\frac{1+\frac{3}{4}\alpha^2}{(1+\frac{1}{2}\alpha^2)^2} \right) \end{pmatrix}. \quad (4.29)$$

es la matriz de covarianzas asintótica.

Observación 4.1. Los estimadores MM son asintóticamente independientes.

4.6.3 Método de máxima verosimilitud

La función de log-verosimilitud de (α, β) basada en las observaciones t_1, \dots, t_n es la siguiente:

$$l(\alpha, \beta) = c_1 + \frac{n}{\alpha^2} - \frac{1}{2\alpha^2} \sum_{i=1}^n \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} \right) - n \log(\alpha) - \frac{n}{2} \log(\beta) + \sum_{i=1}^n \log(t_i + \beta), \quad (4.30)$$

donde c_1 es una constante que no depende de α ni de β .

Para maximizar (4.30), necesitamos calcular sus primeras derivadas respecto de α y de β , obteniendo así las siguientes expresiones:

$$\begin{aligned} \frac{\partial l(\alpha, \beta)}{\partial \alpha} &= -\frac{2n}{\alpha^3} + \frac{1}{\alpha^3} \sum_{i=1}^n \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} \right) - \frac{n}{\alpha}, \\ \frac{\partial l(\alpha, \beta)}{\partial \beta} &= \frac{1}{2\alpha^2} \sum_{i=1}^n \left(\frac{t_i}{\beta^2} - \frac{1}{t_i} \right) - \frac{n}{2\beta} + \sum_{i=1}^n \frac{1}{t_i + \beta}. \end{aligned}$$

A continuación, igualamos ambas expresiones a cero, y evaluamos en $\alpha = \hat{\alpha}_{MV}$ y $\beta = \hat{\beta}_{MV}$. Para el caso de α , obtenemos

$$\hat{\alpha}_{MV} = \left(\frac{1}{n} \sum_{i=1}^n \left(\frac{t_i}{\hat{\beta}_{MV}} + \frac{\hat{\beta}_{MV}}{t_i} \right) \right)^{1/2}, \quad (4.31)$$

que también podemos expresar como

$$\hat{\alpha}_{MV} = \left(\frac{s}{\hat{\beta}_{MV}} + \frac{\hat{\beta}_{MV}}{r} - 2 \right)^{1/2}, \quad (4.32)$$

donde s y r son las medias aritmética y armónica, respectivamente, de las observaciones t_1, \dots, t_n .

Por otro lado, al repetir el procedimiento con β , y sustituyendo $\hat{\alpha}_{MV}$ según la expresión (4.32), obtenemos la siguiente ecuación no lineal:

$$\beta^2 - \beta(2r + K(\beta)) + r(s + K(\beta)) = 0, \quad (4.33)$$

donde hemos llamado $K(x)$ a

$$K(x) = \left[\frac{1}{n} \sum_{i=1}^n (x + t_i)^{-1} \right]^{-1} \text{ para } x \geq 0.$$

El estimador de β , $\hat{\beta}_{MV}$, será la raíz positiva de (4.33). Birnbaum y Saunders [3] ya probaron que $\hat{\beta}$ es la única raíz positiva de (4.33), y además, $r < \hat{\beta} < s$. Para poder resolver la ecuación (4.33), se necesita un método numérico iterativo. Una vez calculado $\hat{\beta}$, se puede sustituir en (4.32), obteniendo así el estimador $\hat{\alpha}$.

El método iterativo propuesto por Leiva [13] se compone de las siguientes ecuaciones iterativas para $m = 0, 1, 2, \dots$

$$\hat{\alpha}^{(m+1)} = \frac{1}{\sqrt{n}} \left(\sum_{i=1}^n \left(\frac{t_i}{\hat{\beta}^{(m)}} + \frac{\hat{\beta}^{(m)}}{t_i} - 2 \right) \right)^{1/2}, \quad (4.34)$$

$$\hat{\beta}^{(m+1)} = \sqrt{\frac{\frac{1}{2(\hat{\alpha}^{(m)})^2} \sum_{i=1}^n t_i}{\frac{1}{2(\hat{\alpha}^{(m)})^2} \sum_{i=1}^n \frac{1}{t_i} - \sum_{i=1}^n \frac{1}{t_i + \hat{\beta}^{(m)}} + \frac{n}{2\hat{\beta}^{(m)}}}}. \quad (4.35)$$

Son necesarios unos valores iniciales $\hat{\alpha}^{(0)}$ y $\hat{\beta}^{(0)}$. Birnbaum y Saunders propusieron el valor inicial $\hat{\beta}^{(0)} = (sr)^{1/2}$, a partir del cual se puede obtener el valor inicial $\hat{\alpha}^{(0)}$ correspondiente mediante la expresión (4.32).

Los estimadores $\hat{\alpha}$ y $\hat{\beta}$ calculados por máxima verosimilitud (MV) son estimadores consistentes, para más detalles ver [6]. Además, estudios más recientes han demostrado que para un tamaño muestral $n = 1$, los estimadores MV de α y β no existen,

mientras que para $n > 1$, siempre existen y son únicos. La expresión de la matriz de información de Fisher esperada, vendrá dada por

$$I(\alpha, \beta) = \begin{pmatrix} \frac{2n}{\alpha^2} & 0 \\ 0 & \frac{n}{\beta^2} \left(\frac{1}{4} + \frac{1}{\alpha^2} + B(\alpha) \right) \end{pmatrix}, \quad (4.36)$$

donde

$$B(\alpha) = 2 \int_0^\infty \left(\frac{1}{1 + \frac{1}{\xi(\alpha z)}} - \frac{1}{2} \right)^2 \phi(z) dz,$$

con $\xi(\cdot)$ definida en (4.2) y $\phi(\cdot)$ la fdd normal estándar vista en (1.9). Para más información sobre los cálculos, ver Leiva [13].

Tomando $\theta = (\alpha, \beta)^t$ y $\mathcal{J}(\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} I(\theta)$, se tiene el siguiente resultado, descrito en el trabajo de Leiva [13].

| Proposición 4.8. *Sea $\hat{\theta}$ es el estimador MV de θ y sea $\Sigma_\theta = \mathcal{J}(\theta)^{-1}$. Entonces, cuando $n \rightarrow \infty$ se verifica*

1.

$$\sqrt{n}(\hat{\theta} - \theta) \rightarrow N(0_{2 \times 1}, \Sigma_\theta).$$

2.

$$(\theta - \hat{\theta})^t \Sigma_\theta^{-1} (\theta - \hat{\theta}) \rightarrow \chi_2^2$$

4.6.4 Máxima verosimilitud para muestras censuradas

En este apartado, trabajaremos con muestras censuradas de tipo II. Entonces, de una muestra $(t_i, \delta_i), i = 1, \dots, n$, se tiene que r de ellas son datos observados, con r conocido. Esto hace que el trato de los datos resulte más sencillo. La muestra sigue una $BS(\alpha, \beta)$, con α y β desconocidos. La función de logverosimilitud, dada por (1.7), es la misma en todos los tipos de censura, luego su expresión será

$$l(\alpha, \beta) = c_2 - \frac{r}{2\alpha^2} \sum_{i=1}^r \left(\frac{t_i}{\beta} - \frac{\beta}{t_i} \right) - r \log(\alpha) - \frac{r}{2} \log(\beta) + \sum_{i=1}^r \log(t_i + \beta) + (n - r) \log \left[\Phi \left(-\frac{1}{\alpha} \xi \left(\frac{t_{r:n}}{\beta} \right) \right) \right], \quad (4.37)$$

donde c_2 es una constante que no depende de (α, β) , $t_{r:n}$ es el valor del r -ésimo estadístico ordenado $T_{r:n}$, $\xi(\cdot)$ es la definida en (4.2) y $\Phi(\cdot)$ es la FdD normal estándar.

Para el cálculo de los estimadores MV procedemos de forma usual, derivando l respecto de cada parámetro para obtener las ecuaciones de verosimilitud. Las derivadas parciales de l son las siguientes

$$\frac{\partial l(\alpha, \beta)}{\partial \alpha} = -\frac{r}{\alpha} + \frac{1}{\alpha^3} \sum_{i=1}^r \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2 \right) + \frac{2(n-r)(t_{r:n} - \beta)h_T(t_{r:n})}{\alpha(t_{r:n} + \beta)t_{r:n}},$$

$$\frac{\partial l(\alpha, \beta)}{\partial \beta} = -\frac{r}{2\beta} + \sum_{i=1}^r \frac{1}{t_i + \beta} - \frac{1}{2\alpha^2} \sum_{i=1}^r \left(\frac{1}{t_i} - \frac{t_i}{\beta^2} \right) + \frac{(n-r)t_{r:n}h_T(t_{r:n})}{\beta},$$

donde $h_T(\cdot)$ es la función de hazard de la distribución, dada por (4.13). Igualando las derivadas a cero, se obtiene

$$\hat{\alpha} = \left[\frac{1}{r} \sum_{i=1}^r \left(\frac{t_i}{\hat{\beta}} + \frac{\hat{\beta}}{t_i} - 2 \right) + \frac{2(n-r)\hat{\alpha}^2(t_{r:n} - \hat{\beta})h_T(t_{r:n})}{r(t_{r:n}\hat{\beta})t_{r:n}} \right]^{1/2},$$

$$\hat{\beta} = \frac{\sum_{i=1}^r \frac{t_i}{2\hat{\alpha}^2}}{\sum_{i=1}^r \frac{1}{2\hat{\alpha}^2 t_i} - \sum_{i=1}^r \frac{1}{t_i + \hat{\beta}} + \frac{r}{2\hat{\beta}} - \frac{(n-r)}{\hat{\beta}} t_{r:n} h_T(t_{r:n})}.$$

De nuevo, los estimadores $\hat{\alpha}_{MV}$ y $\hat{\beta}_{MV}$ no tienen una forma explícita, debe usarse un método numérico para calcularlos. La matriz hessiana, necesaria para obtener la matriz de covarianzas asintótica, no tiene una forma sencilla. Para más información, se puede consultar el trabajo de Leiva [13].

4.7 Aplicaciones

Aplicación 1. Muestra completa

Este conjunto de datos ha sido obtenido del trabajo de Birnbaum y Saunders [3], aunque también se puede encontrar en el paquete "gbs" (ver [2]). En él, se recogen los datos de fatiga de 101 láminas de una aleación de aluminio llamada aluminio 6061-T, cortadas en paralelo a la dirección de laminado y que oscilan a 18 ciclos por segundo (cps). Cada observación mide los miles de ciclos que resiste cada lámina hasta que se produce el fallo. Los individuos se encuentran sometidos a un estrés por ciclo de 31.000 psi (libras por pulgada cuadrada), y sabemos que la muestra sigue una distribución Birnbaum-Saunders, debido a la forma de obtener los datos. Estos se pueden observar

en el Cuadro 4.1, y su representación en forma de histograma, en la Figura 4.3. En ella, se puede ver a simple vista que el histograma de los datos es muy simétrico, lo que nos lleva a pensar que el valor de α estará muy próximo a cero, por propiedades ya vistas de esta distribución.

Miles de ciclos antes del fallo							
70	90	96	97	99	100	103	104
104	105	107	108	108	108	109	109
112	112	113	114	114	114	116	119
120	120	120	121	121	123	124	124
124	124	124	128	128	129	129	130
130	130	131	131	131	131	131	132
132	132	133	134	134	134	134	134
136	136	137	138	138	138	139	139
141	141	142	142	142	142	142	142
144	144	145	146	148	148	149	151
151	152	155	156	157	157	157	157
158	159	162	163	163	164	166	166
168	170	174	196	212			

Cuadro 4.1: Datos de fatiga.

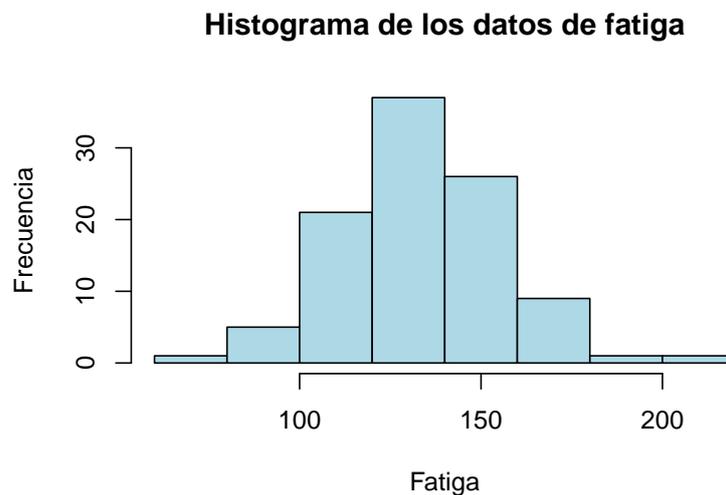


Figura 4.3: Histograma de datos de fatiga, obtenido de los datos del Cuadro 4.1.

Para obtener los estimadores por el método de los momentos, $\hat{\alpha}_M$ y $\hat{\beta}_M$, usamos (4.26) y (4.27), respectivamente, donde

$$s = \bar{t} = 133.7327 \quad y \quad v = \frac{1}{100} \sum_{i=1}^{101} (t_i - \bar{t})^2 = 499,7778$$

En este caso, sabemos que los estimadores por este método existen ya que se cumple que $v/s > \sqrt{5}$. Entonces, usando la expresión ya calculada de ambos parámetros, se tiene

$$\hat{\alpha}_M = 0,1644443, \quad \hat{\beta}_M = 131,9486.$$

Para su cálculo por el método de los momentos modificados, utilizamos (4.28), donde

$$r = \left[\frac{1}{101} \sum_{i=1}^{101} t_i^{-1} \right]^{-1} = 129,9332.$$

De esta manera, los estimadores resultantes son

$$\hat{\alpha}_{MM} = 0,170384, \quad \hat{\beta}_{MM} = 131,8193.$$

Los estimadores MV se pueden obtener fácilmente utilizando el programa R con el paquete "gbs" (ver [2]). Se tienen los siguientes resultados

$$\hat{\alpha}_{MV} = 0,1703847, \quad \hat{\beta}_{MV} = 131,8188.$$

Podemos ver que los estimadores mediante máxima verosimilitud y métodos modificado son prácticamente iguales.

Se obtienen unas densidades ajustadas muy parecidas con los tres métodos, cuya representación puede observarse en la Figura 4.4. Mediante los tres métodos, resulta una estimación de α muy cercana a cero, tal y como habíamos podido prever. Cuanto más cercano es α a cero, más se asemeja β a la media de la distribución, tal y como se puede intuir al ver el histograma de datos.

Para obtener más información sobre la muestra, podemos hallar algunos cuantiles de interés, utilizando los estimadores calculados. Usaremos para ello los estimadores MM, aunque son igualmente válidos los de los demás métodos, ya que son muy parecidos. Para el cálculo de los cuantiles, basta con sustituir los valores ya conocidos de $\hat{\alpha}_{MM}$ y $\hat{\beta}_{MM}$ en la expresión (4.11). Sabemos que la mediana de la distribución es β , luego

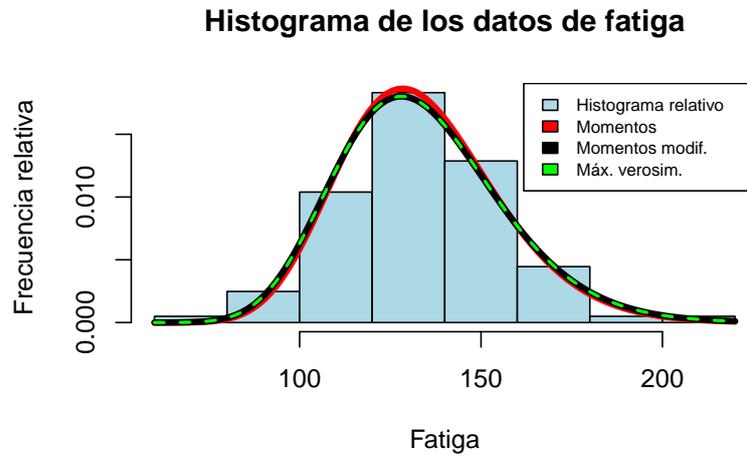


Figura 4.4: Histograma relativo de datos y funciones de densidad BS con parámetros $(\hat{\alpha}_M, \hat{\beta}_M)$, $(\hat{\alpha}_{MM}, \hat{\beta}_{MM})$ y $(\hat{\alpha}_{MV}, \hat{\beta}_{MV})$ respectivamente.

- Cuantil de orden 0,05

$$\hat{t}_{0,05} = 99,69176$$

- Cuantil de orden 0,10

$$\hat{t}_{0,1} = 106,00714$$

- Mediana

$$\hat{t}_{0,5} = \hat{\beta}_{MM} = 131,8193$$

- Cuantil de orden 0,90

$$\hat{t}_{0,9} = 163,91647$$

- Cuantil de orden 0,95

$$\hat{t}_{0,95} = 174,30042$$

A raíz de estos resultados, podríamos casi garantizar que los cupones de aluminio bajo estudio pueden soportar como mínimo 100.000 ciclos bajo estrés sin romperse. Por otro lado, también podemos asegurar que este tipo de láminas raramente van a resistir más de 174.000 ciclos. El valor de la mediana se interpreta como que la mitad de estos productos fallan antes de los 131.000 ciclos bajo presión. Podemos decir entonces que ese es el número esperado de ciclos que resistirá este producto, por ser β la media aproximada de la distribución en este caso.

Como hemos visto anteriormente, los estimadores MM de esta distribución siguen una distribución normal bivalente asintótica, cuya matriz de covarianzas asintótica es la dada por (4.29). Podemos calcular la matriz de covarianzas asintótica estimada a partir de los estimadores MM calculados. El resultado es

$$\hat{\Sigma} = \begin{pmatrix} 0,0001437175 & 0 \\ 0 & 29,10074 \end{pmatrix},$$

cuya diagonal nos da la varianza estimada de cada estimador. De la matriz $\hat{\Sigma}$ podemos obtener el error estándar estimado de $\hat{\alpha}_{MM}$ y $\hat{\beta}_{MM}$ tomando la raíz cuadrada de los elementos de la diagonal, de manera que

$$\hat{\sigma}_{\alpha} = 0,01198822, \quad \hat{\sigma}_{\beta} = 5,39451.$$

Aplicación 2. Muestra censurada

Para esta aplicación sobre datos censurados, partimos del mismo conjunto de datos de la aplicación anterior, ya definidos en el Cuadro 4.1. Para que estos datos se ajusten a los de una muestra censurada de tipo II, fijamos un número $r = 80$ de datos que vamos a tomar como observados, mientras que los 21 datos restantes, aproximadamente un 20 % de la muestra, los tomaremos como datos censurados. En otras palabras, hemos fijado un número de casos a partir del cual se termina el experimento, por lo que los casos sobrantes quedan como datos censurados. Esto no altera la representación gráfica de la muestra, luego sigue siendo válida para este caso la Figura 4.3.

Mediante el programa R con el paquete "gbs" (ver [2]), se pueden calcular de forma sencilla los estimadores MV de la muestra. El resultado es

$$\hat{\alpha} = 0,2071544$$

$$\hat{\beta} = 134,9966$$

El valor estimado de α vuelve a resultar próximo a cero, lo cual encaja con la simetría de los datos.

Observamos que los estimadores son diferentes con respecto a los de la muestra completa, ya que al tomar el 20 % de los datos como censurados, se pierde información. En la Figura 4.5 se puede ver el ajuste del modelo utilizando los estimadores calculados.

Comparando este resultado con las densidades ajustadas de la muestra completa en la Figura 4.3, vemos que el ajuste en este caso es peor, aunque podemos tomarlo igualmente como válido.

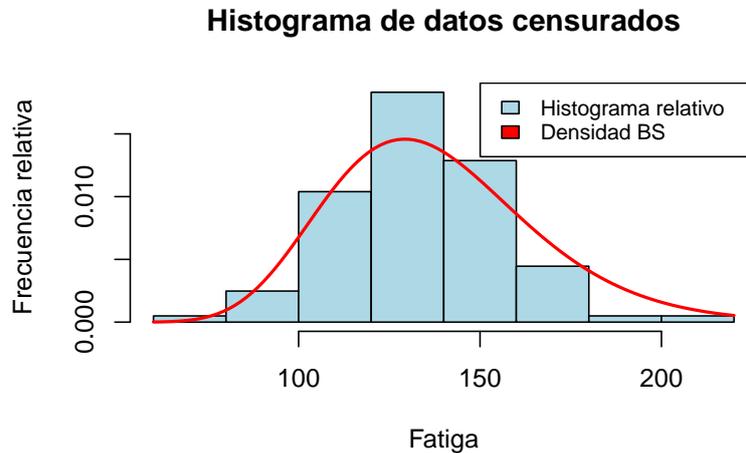


Figura 4.5: Histograma de datos de fatiga censurados y densidad $BS(\hat{\alpha}, \hat{\beta})$.

Procediendo como en la aplicación anterior, calculamos los cuantiles de interés de la distribución obtenida.

- Cuantil de orden 0,05

$$\hat{t}_{0,05} = 96,17202$$

- Cuantil de orden 0,10

$$\hat{t}_{0,1} = 103,60073$$

- Mediana

$$\hat{t}_{0,5} = \hat{\beta} = 134,99664$$

- Cuantil de orden 0,90

$$\hat{t}_{0,9} = 175,90700$$

- Cuantil de orden 0,95

$$\hat{t}_{0,95} = 189,49476$$

En vista de estos resultados, podemos esperar que las láminas de aluminio puedan resistir unos 135.000 ciclos bajo estrés antes de fallar. Podemos garantizar con un 5 % de significación que este tipo de producto puede soportar 96.000 ciclos sin romperse,

mientras que no podrán durar más de 190.000. Trabajar con datos censurados es más complejo que en el caso completo, por lo que no podemos extender mucho más el estudio de esta aplicación.

Bibliografía

- [1] BALAKRISHNAN, N., AND KUNDU, D. Birnbaum-saunders distribution: A review of models, analysis, and applications. *Appl Stochastic Models Bus Ind.* 35 (2019), 4–49.
- [2] BARROS, M., LEIVA, V., AND PAULA, G. *gbs: Package for Generalized Birnbaum-Saunders Distributions*, 2012. R package version 1.0.
- [3] BIRNBAUM, Z., AND SAUNDERS, S. Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability* 6, 2 (1969), 328–347. doi:10.2307/3212004.
- [4] CHANG, D., AND TANG, L. Reliability bounds and critical time for the birnbaum-saunders distribution. *IEEE Trans. Device Mater. Reliab.* 42 (1993), 464–469.
- [5] CHHIKARA, R., AND FOLKS, J. The inverse gaussian distribution as a lifetime model. *Technometrics* 19, 4 (1977), 461–468.
- [6] COX, D., AND HINKLEY, D. *Theoretical Statistics*. Chapman and Hall, 1974.
- [7] CRISTÓBAL, J. *Inferencia estadística*. Ed. Prensas de la Universidad de Zaragoza, 2000.
- [8] GJESSING, H. K. *invGauss: Threshold regression that fits the (randomized drift) inverse Gaussian distribution to survival data.*, 2014. R package version 1.1.
- [9] HENZE, N., AND KLAR, B. Goodness-of-fit tests for the inverse gaussian distribution based on the empirical laplace transform. *Annals of the Institute of Statistical Mathematics* 54 (2002), 425–444.
- [10] KUNDU, D., KANNAN, N., AND BALAKRISHNAN, N. On the hazard function of birnbaum-saunders distribution and associated inference. *Comput Stat Data Anal* 52, 5 (2008), 2692–2702.

- [11] LAWLESS, J. *Statistical Models and Methods for Lifetime Data*. Ed. Wiley- Interscience, 2003.
- [12] LEE, E., AND WANG, J. *Statistical Methods for Survival Data Analysis*. Ed. John Wiley Sons, 2003.
- [13] LEIVA, V. *The Birnbaum-Saunders Distribution*. Ed. Academic Press, 2016.
- [14] LEMESHKO, B., ET AL. *Inverse Gaussian Model and Its Applications in Reliability and Survival Analysis*. 10 2010, pp. 433–453.
- [15] OVERDUIN, S. Use of the lognormal distribution for survival data: Inference and robustness. Master of science, Simon Fraser University, 2004.
- [16] R CORE TEAM. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019.
- [17] ROHATGI, V., AND EHSANES SALEH, A. K. M. *An Introduction to Probability and Statistics*. Ed. John Wiley Sons, 2011.
- [18] THERNEAU, T. M. *A Package for Survival Analysis in S*, 2015. version 2.38.
- [19] TWEEDIE, M. C. K. Statistical properties of inverse gaussian distributions. i. *The Annals of Mathematical Statistics* 28, 2 (1957), 362–377.