



UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICA

La Distribución de Gompertz y la Estimación de Máxima Verosimilitud de sus Parámetros.

Trabajo Especial de Grado presentado ante la ilustre Universidad Central de Venezuela por la **Br. Andrea Alfonso Silva** para optar al título de Licenciado en Matemática.

Tutor: Mercedes Arriojas.

Caracas, Venezuela

Mayo de 2014

Nosotros, los abajo firmantes, designados por la Universidad Central de Venezuela como integrantes del Jurado Examinador del Trabajo Especial de Grado titulado “**La Distribución de Gompertz y la Estimación de Máxima Verosimilitud de sus Parámetros**”, presentado por la **Br. Andrea Alfonzo Silva**, titular de la Cédula de Identidad **21.014.540**, certificamos que este trabajo cumple con los requisitos exigidos por nuestra Magna Casa de Estudios para optar al título de **Licenciado en Matemática**.

Mercedes Arriojas
Tutor

Mairene Colina
Jurado

Jose Benito Hernández
Jurado

--

DEDICATORIA

AGRADECIMIENTOS

ÍNDICE GENERAL

Introducción	1
Capítulo 1. Preliminares	3
1. Nociones Básicas de Probabilidades	3
2. Nociones Básicas de Estadística	16
3. Nociones Básicas de Teoría de la Medida	24
4. Transformada de Laplace	28
5. Integrales de Línea	37
Capítulo 2. Distribución de Gompertz	42
1. Introducción	42
2. Fuerza de Mortalidad de Gompertz.	44
3. Distribución de Gompertz	53
Capítulo 3. Método de Máxima Verosimilitud.	72
1. Estimación por Máxima Verosimilitud.	72
Capítulo 4. Curvas S.	80
1. Introducción.	80
2. Curvas S por Sectores.	84
Conclusiones	88
Bibliografía	90

En las últimas décadas, se han desarrollado a partir de distribuciones probabilísticas, modelos determinísticos de difusión que han sido aplicados con éxito para el ajuste de fenómenos de crecimiento en muchos campos científicos, tales como el estudio de la difusión de innovaciones técnicas o de nuevos productos comerciales. Por ejemplo, modelos de crecimiento determinístico como el Logístico o Gompertz, están basados sobre hipótesis de crecimiento que en principio son válidas para modelar la evolución de ciertos fenómenos reales. En particular, este modelo ha sido aplicado con éxito para describir el crecimiento de poblaciones animales y células, así como para estudiar niveles de elaboración de productos manufacturados.

Éste, es un tipo de modelo matemático para una serie de tiempo, donde el crecimiento es más lento al principio y al final de un período de tiempo.

La idea fundamental de éste modelo se basa en que la tasa instantánea de crecimiento de la población disminuye de forma exponencial con el tiempo, o lo que es lo mismo, la mortalidad crece de forma exponencial con la edad.

Teniendo en cuenta lo expuesto en el apartado anterior, dado que la tasa instantánea de crecimiento debe disminuir de forma exponencial con el tiempo, se propone una ecuación diferencial de la forma,

$$\frac{dy(x)}{y(x)} = k(x) = be^{-cx}$$

donde, $b > 0$.

Con la finalidad de comprender un poco más acerca de este tema, se desarrolló el artículo titulado *The Gompertz distribution and Maximum Likelihood Estimation of its parameters - a revision* de Adam Lenart [8] que presenta el estudio del método de Máxima Verosimilitud para estimar los parámetros de la distribución de Gompertz.

Este trabajo está estructurado en cuatro (4) capítulos. En el Capítulo 1 se presentan algunas definiciones y resultados de probabilidades y estadística que serán necesarios para la comprensión del presente trabajo, así como también definiciones básicas del análisis y teoría de la medida para facilitar el entendimiento del mismo. Las referencias son [1],[2],[4],[6],[10].

El Capítulo 2 está referido a la Distribución de Gompertz. Aquí, se exponen los antecedentes de dicha distribución, su definición formal, propiedades, cálculo de sus momentos, entre otros. [3],[5],[8],[9].

En el Capítulo 3 se presenta el método de Máxima Verosimilitud, tanto para muestras discretas como continuas, y a partir de este hecho, se procede a calcular los estimadores de los parámetros de la distribución de Gompertz. [7],[8],[11].

Por último, en el Capítulo 4 se estudia la relación entre la distribución de Gompertz y las Curvas S, que se caracterizan por tener una gráfica en forma de S. En este capítulo se evidencian las distintas aplicaciones de estas curvas en el modelaje de diversas situaciones cotidianas, abarcando distintos sectores, como lo pueden ser el estudio de crecimiento de poblaciones, el área de infraestructura, industria petrolera, entre otros.

CAPITULO 1

Preliminares

1. Nociones Básicas de Probabilidades

Definición 1.1. (*Espacio Muestral*)

El espacio muestral o espacio de muestreo (generalmente denotado por Ω) consiste en el conjunto de todos los posibles resultados individuales de un experimento aleatorio.

Por ejemplo, si el experimento consiste en lanzar dos monedas, el espacio muestral es el conjunto $\{(cara, cara), (cara, cruz), (cruz, cara) \text{ y } (cruz, cruz)\}$. Un **evento o suceso** es un subconjunto del espacio muestral, llamándose a los sucesos que contengan un único elemento sucesos elementales. En el ejemplo, el suceso "sacar cara en el primer lanzamiento", estaría formado por los sucesos elementales $\{(cara, cara) \text{ y } (cara, cruz)\}$.

Definición 1.2. (*Medida de Probabilidad*)

Sea Ω un espacio muestral y $\mathcal{A}$ una familia de eventos de Ω . Una medida de probabilidad o probabilidad sobre Ω , es una función P definida sobre $\mathcal{A}$ que satisface los siguientes axiomas:

- (i) P es no negativa, es decir, para todo evento $A \in \mathcal{A}$ se cumple que $P(A) \geq 0$.

- (ii) P es *contablemente aditiva* o σ -*aditiva*, esto es, si $A_n \subseteq \Lambda$, $n \geq 1$, son disjuntos, lo que significa que $A_i \cap A_j = \emptyset$ para cada $i \neq j$, entonces $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.
- (iii) Un evento es cierto o seguro si se realiza siempre. El evento cierto tiene probabilidad 1: $P(\Omega) = 1$.

La terna (Ω, Λ, P) se denomina un *espacio de probabilidad*, el valor $P(A)$ se denomina *probabilidad de A*

Definición 1.3. (*Variable Aleatoria*)

Sea (Ω, Λ, P) un espacio de probabilidad. Una variable aleatoria es una función

$$X : \Omega \rightarrow R$$

que satisface la propiedad de que para todo intervalo $I \subseteq R$ el conjunto

$$X^{-1}(I) = \{w \in \Omega : X(w) \in I\} \in \Lambda,$$

es decir, es un evento.

El conjunto $X^{-1}(I)$ se denomina preimágen de I por la función X ; por lo que se dice que la función X es una variable aleatoria si la preimágen de cualquier intervalo es un evento.

Definición 1.4. (*Función de Distribución*)

Si X es una variable aleatoria, se define la función de distribución de X como la función $F_x : R \rightarrow [0, 1]$ tal que para cada $y \in R$,

$$F_x(y) = P(w \in \Omega : X(w) \leq y) = P(X \leq y).$$

Observación.

Resumimos a continuación las propiedades más importantes de la función de distribución:

$$(1) \quad 0 \leq F_X \leq 1.$$

$$(2) \quad \lim_{x \rightarrow +\infty} F_X(x) = 1.$$

$$(3) \quad \lim_{x \rightarrow -\infty} F_X(x) = 0.$$

$$(4) \quad \text{Si } x_1 < x_2, \text{ entonces } F_X(x_1) \leq F_X(x_2).$$

$$(5) \quad \lim_{x \rightarrow a} F_X(x) = F_X(a) \quad \forall a \in \mathbb{R}. \text{ Función siempre continua por la derecha.}$$

Definición 1.5. (*Variable Aleatoria Discreta*)

Una variable aleatoria se denomina discreta, si su rango es un conjunto discreto (finito ó numerable). En este caso existe un conjunto $\{x_n\}_{n \geq 1}$ (conjunto de valores de X) tal que $\sum_{n \geq 1} P(X = x_n) = 1$.

Nota. Dada una variable aleatoria discreta X cuyo conjunto de valores es un conjunto $\{x_n\}_{n \geq 1}$, la función

$$\begin{aligned} p : \{x_n\}_{n \geq 1} &\rightarrow [0, 1] \\ x_n &\rightarrow p(n) = p_n = P(X = x_n) \end{aligned}$$

se denomina función de probabilidad de la variable aleatoria X ó función de masa de probabilidad de la variable aleatoria X .

Esta función satisface la condición $\sum_{n \geq 1} p_n = \sum_{n \geq 1} P(X = x_n) = 1$.

Si X es una variable aleatoria discreta tal que $X(\Omega) = \{x_1, x_2, \dots, x_n, \dots\}$, entonces se define $F_x = \sum_{x_n \leq x} p_n$, donde $p_n = P(X = x_n)$.

En particular, las variables aleatorias definidas sobre espacios de probabilidad discretos son discretas. Sin embargo, una variable aleatoria definida sobre un espacio de probabilidad

no discreto puede ser discreta.

Ejemplo 1.1.

Distribución Binomial.

Un experimento de Bernoulli se caracteriza por ser dicotómico, esto es, sólo son posibles dos resultados. A uno de estos se denomina éxito y tiene una probabilidad de ocurrencia p y al otro, fracaso, es decir,

✓ p es la probabilidad de éxito.

✓ $1-p$ es la probabilidad de fracaso.

La distribución binomial es la distribución de probabilidad discreta, de la variable aleatoria definida como el número de éxitos en una secuencia de n ensayos de Bernoulli independientes entre sí.

Si X es la variable aleatoria que representa el número de éxitos en n ensayos, el interés está entonces en determinar la probabilidad de obtener exactamente $X(w) = x$ éxitos durante los n ensayos.

La probabilidad de tener en n ensayos x éxitos consecutivos seguidos de $n-x$ fracasos consecutivos es:

$$\underbrace{p \cdot p \cdots p}_{x \text{ veces}} \cdot \underbrace{(1-p) \cdot (1-p) \cdots (1-p)}_{(n-x) \text{ veces}} = p^x (1-p)^{n-x}.$$

Para esta distribución se verifica que x puede tomar los valores

$$0, 1, 2, \dots, n.$$

La probabilidad de obtener exactamente x éxitos y $n-x$ fracasos en cualquier otro orden es la misma.

Tomando el número de combinaciones de n objetos tomando X a la vez se obtiene,

$$p_X(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \text{ para } k = 0, 1, 2, \dots, n.$$

Recordemos que la combinatoria está dada por:

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

La función de distribución Binomial es entonces,

$$F_X(x; n, p) = \sum_{i=0}^x \binom{n}{i} p^i (1 - p)^{n-i}, \quad \text{para } i = 0, 1, 2, \dots, n$$

Definición 1.6. (*Función de Densidad*)

Una función $f : R \rightarrow R$ es una función de densidad sobre R si y sólo si f satisface las condiciones siguientes:

- (i) Para todo $x \in R$, $f(x) \geq 0$.
- (ii) $\int_{-\infty}^{\infty} f(u) du = 1$.

Definición 1.7. (*Variable Aleatoria Continua*)

Una variable aleatoria X es continua ó absolutamente continua si existe una función de densidad f tal que para todo $a \in R$

$$F(a) = P(X \leq a) = \int_{-\infty}^a f(u) du.$$

Observación.

Cuando la variable es continua, no tiene sentido hacer una suma de las probabilidades de cada uno de los términos en el sentido anterior, ya que el conjunto de valores que puede tomar la variable es no numerable. En este caso, lo que generaliza de modo natural el concepto de suma (Σ) es el de integral ($\int$).

Observación.

Una propiedad de la función de distribución F , es que $\lim_{t \rightarrow a^-} F(t) = F(a) - P(X = a)$, la cual nos indica que para todo $a \in R$, $P(X = a)$ es igual al “salto” de la función de distribución F en el punto a . Por lo tanto $P(X = a) = 0$ en todos los puntos a de continuidad de F .

Así, para una función de distribución F , continua, para todo $x \in R$, se tiene que $P(X = x) = 0$.

Entonces, para variables continuas no tiene interés hablar de la probabilidad de que $X = x_n$, ya que ésta vale siempre 0.

Observación.

Tanto la función de distribución como la función de densidad para una variable aleatoria siempre verifican las siguientes propiedades:

$$(i) \quad F_X \geq 0.$$

$$(ii) \quad \int_{-\infty}^{\infty} f(x) dx = 1.$$

$$(iii) \quad P[a \leq x \leq b] = \int_a^b f(x) = F(b) - F(a) \text{ para } a < b.$$

$$(iv) \quad f(x) = F'_X(x). \text{ La función de densidad es la derivada de la función de distribución.}$$

Ejemplos 1.2.

- (i) 50 autos entraron a una carrera de 100 millas. Sea X el número de carros que terminaron la carrera. Aquí, X puede tomar los valores: 0, 1, 2, ..., 50.

- (ii) Sea el experimento de lanzar 3 veces una moneda y, el número de caras X es una variable aleatoria que toma los valores 0,1,2 ó 3; es decir, puede que ninguna vez, una sola, dos o tres veces se obtenga cara como resultado.

El espacio muestral de lanzar 3 veces una moneda es:

$$S = \{(c,c,c), (c,c,+), (c,+,+), (+,+,+), (+,+,c), (+,c,c), (c,+,c) \text{ y } (+,c,+) \}.$$

Si solamente nos interesan los elementos del espacio muestral donde aparece cara (c), entonces X toma las opciones :

$$E = \{(c,c,c), (c,c,+), (c,+,+), (+,+,c), (+,c,c), (c,+,c) \text{ y } (+,c,+) \}.$$

Por ejemplo, la probabilidad de que $X = 2$ es $\frac{4}{8}$, ya que de los ocho posibles resultados del espacio muestral hay cuatro en los cuales se encuentran dos caras.

- (iii) Una vez a la semana un estudiante compra un billete de lotería. Sea X el número de boletos que él comprará antes de que gane al menos 1000 dólares. Los valores posibles de X pueden ser 1, 2, 3, ... y la lista puede que nunca termine.

Definición 1.8. (*Esperanza*)

Caso Discreto.

Sea X una variable aleatoria discreta tal que su conjunto de valores es $\{x_n\}_{n \geq 1}$ y su función de probabilidad $\{p_n\}_{n \geq 1}$, donde $p_n = P(X_n = x_n)$. La esperanza matemática o valor esperado de X es el número denotado por $E[X]$ y definido como,

$$E[X] = \sum_{n=0}^{\infty} x_n P(X = x_n) = \sum_{n=0}^{\infty} x_n p_n,$$

siempre y cuando la serie anterior sea convergente. En este caso, se dice que existe la esperanza matemática de X .

Caso Continuo.

Sea X una variable aleatoria continua con función de densidad f . Se define la esperanza de X como,

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx;$$

siempre que la integral anterior sea finita. En este caso ($E[X] < \infty$) se dice que X tiene esperanza o también que es integrable con respecto a la medida de probabilidad dada.

Ejemplos 1.3.

(i) Sea X una variable aleatoria de Poisson con parámetro λ . Entonces $E[X] = \lambda$.

En efecto,

$$\begin{aligned} (1.1) \quad E[X] &= \sum_{k=1}^n \frac{k e^{-\lambda} \lambda^k}{k!} = \sum_{k=1}^n \frac{e^{-\lambda} \lambda^k}{(k-1)!} \\ &= \lambda \sum_{k=1}^n \frac{e^{-\lambda} \lambda^{k-1}}{(k-1)!} \\ &= \lambda \sum_{j=0}^n \frac{e^{-\lambda} \lambda^j}{j!}. \end{aligned}$$

De aquí, por propiedades de series, se obtiene que,

$$E[X] = \lambda.$$

(ii) Sea X una variable aleatoria exponencial con parámetro λ , entonces

$$E[X] = \int_0^{\infty} \lambda x e^{-\lambda x} dx = \lambda \left[\frac{-x}{\lambda} e^{-\lambda x} - \frac{1}{\lambda^2} e^{-\lambda x} \right]_0^{\infty} = \frac{1}{\lambda}.$$

Observación.

Si X es una variable aleatoria y a y b son constantes entonces

$$(1.2) \quad E[aX + b] = aE[X] + b.$$

siempre que las esperanzas existan.

Definición 1.9. (*Varianza*)

Si X es una variable aleatoria con media $E[X] = \mu$, entonces la varianza de X , denotada $Var[X]$, se define como,

$$(1.3) \quad Var[X] = E[(X - E[X])^2].$$

Observación.

$$(1.4) \quad Var[X] = E[X^2] - (E[X])^2.$$

En efecto,

$$\begin{aligned} Var[X] = E[(X - E[X])^2] &= E[X^2 - 2XE[X] + (E[X])^2] \\ &= E[X^2] - 2E[X]E[X] + (E[X])^2 \\ (1.5) \quad &= E[X^2] - (E[X])^2. \end{aligned}$$

Observación. (*Propiedad de la Esperanza*)

Caso Discreto.

$$E[X] = \sum_{n=0}^{\infty} x_n p_n.$$

Dada $g : R \rightarrow R$, la esperanza de $Y = g(X)$ viene dada por

$$E[Y] = E[g(X)] = \sum_{n=0}^{\infty} g(x_n) p_n.$$

Caso Continuo.

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx.$$

Dada $g : R \rightarrow R$, la esperanza de $Y = g(X)$ viene dada por

$$E[Y] = E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

Por ejemplo, consideremos $g(X) = X^2$ como una variable aleatoria discreta. Entonces,

$$E[Y] = E[X^2] = \sum_{n=0}^{\infty} x_n^2 p_n.$$

Igualmente, si tomamos $g(X) = X^2$ como una variable aleatoria continua, obtenemos

$$E[Y] = E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) dx.$$

Ejemplo 1.4.

Sea X una variable aleatoria de Poisson con parámetro λ . Entonces $Var[X] = \lambda$.

$$Var[X] = E[X^2] - (E[X])^2.$$

Es claro que $(E[X])^2 = \lambda^2$. Veamos que $E[X^2] = \lambda^2 + \lambda$.

$$\begin{aligned}
 E[X^2] &= \sum_{k=1}^n \frac{k^2 e^{-\lambda} \lambda^k}{k!} = \sum_{k=1}^n k(k-1) \frac{e^{-\lambda} \lambda^k}{k!} + E[X] \\
 &= \lambda^2 \sum_{j=0}^n \frac{e^{-\lambda} \lambda^j}{j!} + \lambda \\
 (1.6) \qquad &= \lambda^2 + \lambda.
 \end{aligned}$$

De aquí, $Var[X] = \lambda$.

1.1. Función Generadora de Momentos.

Sea X una variable aleatoria que toma sólo valores enteros no negativos y sea

$$a_j = P(X = j) \text{ para } j \geq 0,$$

su función de probabilidad.

Para una variable $z \in R$ consideremos la serie de potencias en z

$$g(z) = a_0 + a_1 z + a_2 z^2 \dots = \sum_{i=0}^{\infty} a_i z^i.$$

Como $\sum_{i=0}^{\infty} a_i = 1$ tenemos que

$$|g(z)| \leq \sum_{i=0}^{\infty} |a_i| |z|^i \quad \text{siempre que } |z| \leq 1;$$

por lo que la serie de potencias $g(z)$ converge siempre que $|z| \leq 1$. Esto implica que para $|z| \leq 1$ la serie $g(z)$ puede ser diferenciada “término a término”. Evaluando $g(z)$ y sus sucesivas derivadas en $z = 0$ se obtiene

$$\begin{aligned}
g(0) &= a'_0, \\
g'(z) &= a_1 + 2a_2z + 3a_3z^2 \dots \Rightarrow g'(0) = a_1, \\
g''(z) &= 2a_2 + 3 \times 2a_3z + 4 \times 3a_4z^2 \dots \Rightarrow g''(0) = 2a_2, \\
g'''(z) &= 3 \times 2a_3 + 4 \times 3a_4z + \dots \Rightarrow g'''(0) = 3 \times 2a_3, \\
&\vdots \\
g^n(z) &= n!a_n + n!(n+1)a_{n+1}z + n!(n+1)(n+1)z^2 \Rightarrow g^n(0) = n!a_n.
\end{aligned}$$

Por lo tanto, a partir de la función g se puede obtener la función de probabilidad de X evaluando en 0 las sucesivas derivadas de g ; esto significa que la función g en cierto modo caracteriza la función de distribución de X .

Definición 1.10. (*Función Generadora de Momentos*)

Dada una variable aleatoria X cuyos valores son enteros no negativos, se define la función generadora de momentos ó función generatriz de X como la serie de potencias

$$g(z) = \sum_{i=0}^{\infty} a_i z^i$$

donde, $a_i = P(X = i)$ para $i \geq 0$.

Sea $z = e^{-t}$ para $t \in (0, \infty)$ y sea $X > 0$ una variable aleatoria. Entonces, se expresa la función generadora de momentos de X como

$$\phi(t) = E[z^X] = E[e^{-tX}] = \begin{cases} \sum_{x:P(X=x)>0} e^{-xt} P(X=x) & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{\infty} e^{-tx} f(x) dx & \text{si } X \text{ es continua con densidad.} \end{cases}$$

Comentarios.

- (i) La función generadora de momentos existe siempre que X tome valores no negativos, en el caso continuo, esto significa que $f(u) = 0$ para todo $u < 0$. De lo contrario, la esperanza de z^x podría no existir.
- (ii) Con esta definición de la función generatriz no es tan sencillo obtener los valores de la función de probabilidad o la función de densidad; sin embargo, al igual que cuando $X > 0$ toma sólo valores enteros, la distribución de la variable aleatoria X está determinada en forma única por su función generatriz.

Definición 1.11. (*Momento de Orden K*)

Dada una variable aleatoria X y un entero $k > 0$, la esperanza $E[X^k]$ se denomina momento de orden k ó k -ésimo momento de X . Así mismo, la esperanza $E[(X - E[X])^k]$ se denomina momento centrado de orden k de X .

Si evaluamos las derivadas de $g(z)$ en $z = 1$ tenemos que,

$$\begin{aligned}
 g'(z) &= a_1 + 2a_2z + 3a_3z^2 + \dots \Rightarrow g'(1) = \sum_{i=1}^{\infty} ia_i = E[X], \\
 g''(1) &= \sum_{i=1}^{\infty} i(i-1)a_i = \sum_{i=1}^{\infty} i^2a_i - \sum_{i=1}^{\infty} i(i-1)a_i = E[X^2] - E[X], \\
 g'''(1) &= \sum_{i=1}^{\infty} i(i-1)(i-2)a_i, \\
 &\vdots
 \end{aligned}$$

Esto significa que también pueden obtenerse los momentos de cualquier orden evaluando las derivadas de la función generadora de momentos en $z = 1$.

2. Nociones Básicas de Estadística

Definición 2.1. (*Sesgo*)

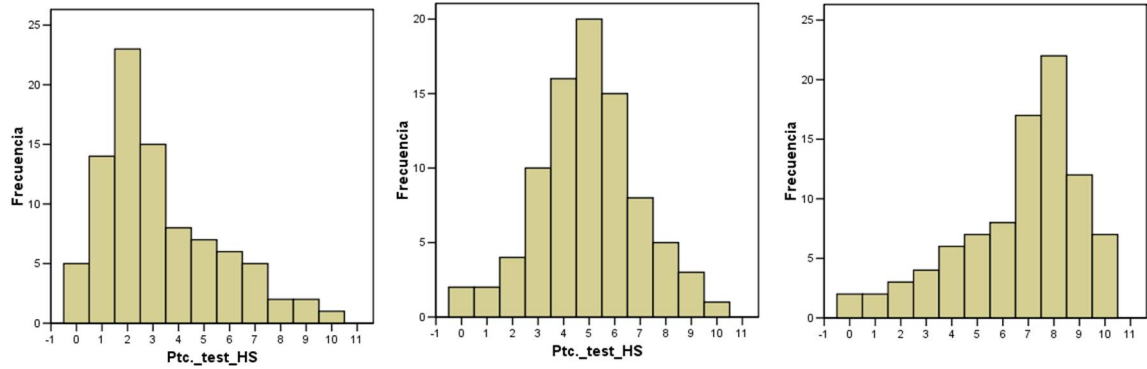
Sean μ y σ^2 tales que, denotan la esperanza y la varianza de una variable aleatoria X . Se define el sesgo como,

$$(1.7) \quad \kappa = \frac{E[(X - \mu)^3]}{\sigma^3} = \frac{\int (x - \mu)^3 dF(x)}{\left(\int (x - \mu)^2 dF(x) \right)^{\frac{3}{2}}}.$$

De aquí, el sesgo se define como el tercer momento centrado con respecto a la media sobre la raíz cuadrada de la varianza de X al cubo. El sesgo mide la asimetría de una distribución, es decir, el alargamiento de la distribución hacia la izquierda o hacia la derecha. Así:

- (i) Si el coeficiente de sesgo tiene un valor positivo se dice que la distribución es SESGADA a DERECHA o que tiene SESGO POSITIVO.
- (ii) Si el coeficiente de sesgo tiene un valor negativo se dice que la distribución es SESGADA a IZQUIERDA o que tiene SESGO NEGATIVO.
- (iii) Si el coeficiente de sesgo tiene un valor 0 se dice que la distribución es INSESGADA o que tiene SESGO 0.

A continuación, se muestra un ejemplo de las puntuaciones de un grupo de sujetos en un test de habilidades sociales antes, durante y después de recibir 6 sesiones de entrenamiento en habilidades sociales. En este ejemplo, la gráfica de la izquierda representa una distribución sesgada a izquierda (sesgo negativo), la gráfica central, una distribución insesgada, y por último, la gráfica del extremo derecho, representa una distribución con sesgo positivo.



Definición 2.2. (*Curtosis*)

La curtosis es una medida de la forma, es decir, medida que determina numéricamente algunas características, las cuales hacen referencia al grado de apuntamiento de los datos de una distribución, esto es, el coeficiente de curtosis mide cuán puntiaguda es una distribución respecto de un estándar. Este estándar es la forma acampanada denominada normal, y corresponde a una curva de gran importancia en estadística.

Así, la curtosis es el estudio que mide la mayor o menor cantidad de datos que se agrupan en torno a la moda. Así, las medidas de curtosis tratan de estudiar la proporción de la varianza que se explica por la combinación de datos extremos respecto a la media en contraposición con datos poco alejados de la misma. Una mayor curtosis implica una mayor concentración de datos muy cerca de la media de la distribución coexistiendo al mismo tiempo con una relativamente elevada frecuencia de datos muy alejados de la misma. En otras palabras, la curtosis mide el grado de agudeza o achatamiento de una distribución con relación a la distribución normal, es decir, mide cuán puntiaguda es la gráfica de la función de densidad de una distribución. Esta se define como,

$$(1.8) \quad \beta = \frac{E[(X - \mu)^4]}{(E[X - \mu]^2)^2} - 3.$$

con μ como el valor esperado de la variable aleatoria.

Ejemplos 2.1.

- (i) El coeficiente de Sesgo de una distribución de Poisson con parámetro λ es $\kappa = \frac{1}{\sqrt{\lambda}}$.

En efecto, usando el hecho de que $E[X] = \lambda$ y $Var[X] = \lambda$ en la definición de Sesgo, tenemos que,

$$\begin{aligned}
 \kappa &= \frac{E[(X - \mu)^3]}{\sigma^3} \\
 &= \frac{E[X^3 - 3X^2\mu + 3X\mu - \mu^3]}{\sigma^3} \\
 &= \frac{E[X^3] - 3E[X]Var[X] - E[X]^3}{Var[X]^{\frac{3}{2}}} \\
 &= \frac{E[X^3] - 3\lambda^2 - \lambda^3}{\lambda^{\frac{3}{2}}}.
 \end{aligned}$$

Faltaría calcular el valor de $E[X^3]$. Este es,

$$\begin{aligned}
 E[X^3] &= \sum_{k=1}^n \frac{k^3 e^{-\lambda} \lambda^k}{k!} \\
 &= \sum_{k=1}^n \frac{(k^3 + k^2 - k^2) e^{-\lambda} \lambda^k}{k!} \\
 &= \sum_{k=2}^n \frac{k e^{-\lambda} \lambda^k}{(k-2)!} + \lambda^2 + \lambda \\
 &= \lambda^2 \sum_{l=0}^n \frac{(l+2) e^{-\lambda} \lambda^l}{l!} + \lambda^2 + \lambda \\
 &= \lambda^3 + 3\lambda^2 + \lambda.
 \end{aligned}$$

Así,

$$(1.9) \quad E[X^3] = \lambda^3 + 3\lambda^2 + \lambda.$$

En conclusión,

$$\kappa = \frac{\lambda^3 + 3\lambda^2 + \lambda - 3\lambda^2 - \lambda^3}{\lambda^{\frac{3}{2}}} = \frac{\lambda}{\lambda^{\frac{3}{2}}} = \frac{1}{\sqrt{\lambda}}.$$

(ii) El coeficiente de Curtosis de una distribución de Poisson con parámetro λ es $\beta = \frac{1}{\lambda}$.

En efecto, usando el hecho de que $E[X] = \lambda$ y $Var[X] = \lambda$ en la definición de Curtosis, tenemos que,

$$\begin{aligned} \beta &= \frac{E[(X - \mu)^4]}{\sigma^4} - 3 \\ &= \frac{E[X^4 - 4X^3\mu + 6X^2\mu^2 - 4X\mu^3 + \mu^4]}{\sigma^4} - 3 \\ &= \frac{E[X^4] - \lambda^4 - 6\lambda^3 - 4\lambda^2}{\lambda^2} - 3. \end{aligned}$$

De aquí, procedemos a calcular el valor de $E[X^4]$.

$$\begin{aligned} E[X^4] &= \sum_{k=1}^n \frac{k^4 e^{-\lambda} \lambda^k}{k!} \\ &= \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda. \end{aligned}$$

Así,

$$(1.10) \quad E[X^4] = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda.$$

En conclusión,

$$\beta = \frac{\lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda - \lambda^4 - 6\lambda^3 - 4\lambda^2}{\lambda^2} - 3 = \frac{3\lambda^2 + \lambda}{\lambda^2} - 3 = \frac{1}{\lambda}.$$

2.1. Método de Máxima Verosimilitud.

El método más común para estimar parámetros en un modelo paramétrico es el método de máxima verosimilitud.

La verosimilitud consiste en otorgar a un estimador una determinada credibilidad, una mayor posibilidad de ser cierto el valor o el método para obtener dicho valor. Esto es, usamos la información disponible en la muestra para elegir el valor del parámetro para el cuál es más probable haber observado ese resultado muestral.

En términos probabilísticos podríamos hablar de que la verosimilitud es la probabilidad de que ocurra o se dé una determinada muestra si es cierta la estimación que hemos efectuado o el estimador que hemos planteado.

Evidentemente, la máxima verosimilitud, será aquel estimador que nos otorgue mayor credibilidad de ser cierto un valor.

Definición 2.3. (*Función de Verosimilitud*)

Supongamos que tenemos las variables aleatorias $X_1, X_2, \dots, X_n$ con común función de densidad $f(x; \theta)$. Aquí el parámetro desconocido es θ , el cual pertenece al espacio de parámetros Θ .

Dado que $X_i = x_i$, $i = 1, \dots, n$, la función de verosimilitud viene dada por,

$$(1.11) \quad L(\theta) = \prod_{i=1}^n f(x_i; \theta), \quad \theta \in \Theta.$$

Tanto x_i como θ pueden ser vectores. Se puede suponer que los valores observados $x_1, x_2, \dots, x_n$ son fijos mientras que θ puede variar libremente.

En muchos casos se trabaja con el logaritmo de la función de verosimilitud $l(\theta, y) = \log(L(\theta, y))$. Este es,

$$(1.12) \quad l(\theta) = \log(L(\theta)) = \sum_{i=1}^n \log(f(y_i), \theta).$$

La función de verosimilitud no es una función de densidad: en general, no se cumple que la integral de $L(\theta)$ sea 1.

Definición 2.4. (*Estimador de máxima verosimilitud*)

El estimador de máxima verosimilitud (MLE, por sus siglas en ingles), denotado por $\hat{\theta}$, es el valor de θ que maximiza L . Esto es,

$$(1.13) \quad L(\hat{\theta}) = \sup_{\theta \in \Theta} L(\theta).$$

El máximo de $l(\theta)$ ocurre en el mismo punto que el máximo de $L(\theta)$, así que, maximizar la función de verosimilitud logarítmica conlleva a la misma respuesta que maximizar la función de verosimilitud. Generalmente, es más fácil trabajar con la función logarítmica, puesto que al combinar observaciones independientes, hemos visto que en el cálculo de la función de verosimilitud interviene el producto de las probabilidades individuales, por lo que habitualmente interesa tomar logaritmos, ya que éstos transforman los productos en sumas y los cocientes en restas.

En el caso discreto, $L(\theta)$ significa $P_{\theta}\{X = x\}$, que es la probabilidad de observar el tipo de datos disponibles cuando $\hat{\theta}$ es el verdadero valor del parámetro desconocido. El estimador de máxima verosimilitud se interpreta como el valor de θ que maximiza la oportunidad de observar los datos particulares disponibles. En el caso continuo, una interpretación similar se da mediante la sustitución del hecho de probabilidad por una descripción análoga usando la función de densidad probabilística conjunta de X .

Para la definición de MLE, no hay ningún método matemático predilecto a seguir con el fin de localizar donde $L(\theta)$ alcanza su supremo. Si $L(\theta)$ es una función dos veces diferenciable de θ , entonces podemos aplicar las técnicas estándar del cálculo diferencial para encontrar $\hat{\theta}$. A veces tomamos el logaritmo natural de $L(\theta)$ primero y, a continuación, se maximiza dicho logaritmo para obtener $\hat{\theta}$.

Hay situaciones en las que se encuentra una solución única o situaciones donde encontramos más de una solución que maximice de manera global a $L(\theta)$. En algunas situaciones puede que no se logre maximizar a $L(\theta)$.

Observación.

Si multiplicamos $L(\theta)$ por una constante positiva c (no dependiente de θ) el estimador de máxima verosimilitud no se modificará.

Ejemplos 2.3.

- (i) Supongamos que $X_1, X_2, \dots, X_n \sim \text{Bernoulli}(p)$. La función de probabilidad viene dada por $f(x, p) = p^x(1-p)^{1-x}$ para $x = 0, 1$. El parámetro desconocido es p . Entonces,

$$L(p) = \prod_{i=1}^n f(X_i; p) = \prod_{i=1}^n p^{X_i}(1-p)^{1-X_i} = p^S(1-p)^{n-S}.$$

donde, $S = \sum_i X_i$. Por tanto,

$$l(p) = S \log(p) + (n-S) \log(1-p).$$

Tomando la derivada de $l(p)$, e igualandola a 0, obtenemos que el estimador es

$$\hat{p}_n = \frac{S}{n}.$$

- (ii) Sean $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$. El parámetro es $\theta = (\mu, \sigma)$ y la función de verosimilitud es,

$$L(\mu, \theta) = \prod_i \frac{1}{\sigma} e^{-\frac{(X_i - \mu)^2}{2\sigma^2}}$$

$$\begin{aligned}
&= \sigma^{-1} e^{-\frac{\sum_i (X_i - \mu)^2}{2\sigma^2}} \\
&= \sigma^{-1} e^{-\frac{nS^2}{2\sigma^2}} e^{-\frac{n(\bar{X} - \mu)^2}{2\sigma^2}},
\end{aligned}$$

donde, $\bar{X} = n^{-1}\sum_i X_i$ es la media muestral y $S^2 = n^{-1}\sum_i (X_i - \bar{X})^2$. La última igualdad de la ecuación anterior, se obtiene del hecho de que $\sum_i (X_i - \mu)^2 = nS^2 + n(\bar{X} - \mu)^2$, lo cual se puede verificar escribiendo $\sum_i (X_i - \mu)^2 = \sum (X_i + \bar{X} - \bar{X} - \mu)^2$. La función de verosimilitud logarítmica es,

$$l(\mu, \sigma) = -n \log(\sigma) - \frac{nS^2}{2\sigma^2} - \frac{n(\bar{X} - \mu)^2}{2\sigma^2}.$$

Resolviendo las ecuaciones,

$$\frac{\partial \ell(\mu, \sigma)}{\partial \mu} = 0 \text{ y } \frac{\partial \ell(\mu, \sigma)}{\partial \sigma} = 0,$$

se concluye que $\hat{\mu} = \bar{X}$ y $\hat{\sigma} = S$.

2.1.1. Propiedades de los Estimadores de Máxima Verosimilitud.

Una propiedad útil de los MLE es la de **invariancia** que dice que si g es una función con inversa g^{-1} , de manera que $\phi = g(\theta)$ implica que $\theta = g^{-1}(\phi)$, entonces el MLE de ϕ , $\hat{\phi}$, se calcula como

$$\hat{\phi} = g(\hat{\theta}),$$

siendo $\hat{\theta}$ el MLE de θ .

Como ya sabemos, podemos maximizar $L(\theta)$ o bien maximizar $l(\theta)$. En problemas regulares, el MLE puede hallarse igualando a 0 las derivadas primeras de $l(\theta)$ con respecto a θ .

La primera derivada de $l(\theta)$ respecto a θ se llama score. En el caso univariado tenemos:

$$l'(\theta) = \sum_{i=1}^n u_i(\theta),$$

donde

$$u_i(\theta) = \frac{\partial}{\partial \theta} \log(f(y_i, \theta)).$$

Si tenemos q parámetros, $\theta = (\theta_1, \theta_2, \dots, \theta_q)'$, el vector de scores es

$$l'(\theta) = \begin{bmatrix} \frac{\partial}{\partial \theta_1} \\ \frac{\partial}{\partial \theta_2} \\ \vdots \\ \frac{\partial}{\partial \theta_q} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n \frac{\partial}{\partial \theta_1} \log(f(y_i, \theta)) \\ \sum_{i=1}^n \frac{\partial}{\partial \theta_2} \log(f(y_i, \theta)) \\ \vdots \\ \sum_{i=1}^n \frac{\partial}{\partial \theta_q} \log(f(y_i, \theta)) \end{bmatrix}.$$

3. Nociones Básicas de Teoría de la Medida

Definición 3.1. (σ -álgebra)

Consideremos un conjunto cualquiera Ω . A la clase de todos los subconjuntos de Ω , partes de Ω , la denotamos por $\wp(\Omega)$. Se denomina σ -álgebra de subconjuntos de Ω a una familia $\mathcal{A} \subseteq \wp(\Omega)$ que cumple con las siguientes propiedades:

- (i) $\Omega \in \mathcal{A}$.
- (ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$, es decir, si A es un evento, entonces “ A no ocurre” también es un evento.
- (iii) $A_n \in \mathcal{A}$, para $n=1,2,\dots$ entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, es decir que la unión numerable de eventos es un evento.

Nótese que al realizar un experimento, la unión de eventos ocurre si el resultado pertenece a alguno de los eventos que integran la unión.

Dado un conjunto cualquiera Ω y una σ -álgebra $\mathcal{A}$ de subconjuntos de Ω , a la pareja $(\Omega, \mathcal{A})$ se le llama *espacio medible*, es decir, al conjunto Ω se le ha provisto de una estructura de medibilidad. A los elementos de $\mathcal{A}$ se les denomina conjuntos medibles.

Definición 3.2. (*σ -álgebra Generada por una Familia*)

Para toda $C \in \wp(\Omega)$, existe una σ -álgebra más pequeña que contiene a C . Ésta σ -álgebra se llama la σ -álgebra generada por C y se denota por $\sigma(C)$.

Ejemplo 3.1.

La colección B de conjuntos de Borel es la menor σ -álgebra que contiene todos los conjuntos abiertos en R .

Definición 3.3. (*Medida*)

Una medida η sobre un espacio medible $(\Omega, \mathcal{A})$ es una función

$$\eta : \mathcal{A} \rightarrow R^+$$

tal que

- (i) $\eta(\emptyset) = 0$.
- (ii) η es sigma-aditiva, es decir si $(A_n) \in \mathcal{A}$ es una sucesión de conjuntos disjuntos, entonces

$$\eta\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \eta(A_n).$$

A $\eta(\mathcal{A})$ se le llama **medida** de $\mathcal{A}$.

Las medidas más importantes en R^n son indudablemente aquellas que están definidas en la σ -álgebra de Borel de R^n y que son finitas sobre los subconjuntos acotados de R^n . Estas medidas se suelen conocer con el nombre de medidas de Lebesgue-Stieltjes y resultan estar asociadas a un tipo particular de funciones conocidas como funciones de distribución.

Definición 3.4. (*Medida de Lebesgue-Stieltjes*)

Dada una función de distribución F en R , existe una única medida η en $B(R)$ tal que $\eta(a, b] = F(b) - F(a)$, para cada $a < b$ de R .

Una medida definida de esta manera se llama **medida de Lebesgue-Stieltjes** asociada a F .

Supongamos, sea $f : R \rightarrow [0, \infty)$ una función integrable sobre cualquier intervalo acotado.

Definimos la función $F : R \rightarrow R$ en la siguiente forma

$$\begin{aligned} F(x) &= F(0) + \int_0^x f(t) dt \quad \text{si } x > 0, \\ F(x) &= F(0) - \int_x^0 f(t) dt \quad \text{si } x < 0. \end{aligned}$$

Siendo $F(0)$ un valor fijado de forma arbitraria.

De aquí, F origina una medida de Lebesgue-Stieltjes μ tal que

$$\mu((a, b]) = F(b) - F(a) = \int_a^b f(t) dt \quad a, b \in R, a < b.$$

En particular, si consideramos la función $f(x) = 1$, $x \in R$ y tomamos $F(0) = 0$, entonces $F(x) = x$, $x \in R$.

Se deduce como medida η asociada

$$\eta((a, b]) = b - a, \quad a, b \in R, a < b$$

es decir, la medida de Lebesgue sobre $B(R)$.

Definición 3.5. (*Espacio de Medida*)

Un espacio de medida es una terna (X, Ψ, η) formada por un espacio medible (X, Ψ) y una medida definida en él.

Ejemplo 3.2.

Si B es la clase de los conjuntos de Borel en R y m la medida de Lebesgue, (R, B, m) es un espacio de medida.

Definición 3.6. (*Función Medible*)

Diremos que una función entre espacios de medida

$$F : (\Omega_1, \Lambda_1) \rightarrow (\Omega_2, \Lambda_2),$$

es *medible* si para cada $B \in \Lambda_2$, $F^{-1}(B) \in \Lambda_1$.

Si $(\Omega_2, \Lambda_2) = (R, B(R))$ diremos que F es una función medible o función Borel medible.

En el contexto de la Teoría de Probabilidades las funciones medibles vienen a ser las llamadas *variables aleatorias*.

Teorema 3.1 (*Derivación bajo la integral*)

Sean I, J intervalos reales no triviales, con I compacto y J abierto. Sea $f : I \times J \rightarrow R$ una función tal que

- (1) $f(\cdot, \lambda)$ es integrable en I para todo $\lambda \in J$;
- (2) $f(x, \cdot)$ es derivable en J para todo $x \in I$;

Supongamos además que $\frac{\partial f}{\partial \lambda}$ es continua en $I \times J$. Entonces,

- $\frac{\partial f}{\partial \lambda}(\cdot, \lambda)$ es integrable para todo $\lambda \in J$.
- $\int_I f(x, \lambda) dx$ es derivable con derivada continua en J para todo $x \in I$.

Y se cumple la regla de derivación bajo la integral,

$$(1.14) \quad \frac{d}{d\lambda} \int_I f(x, \lambda) dx = \int_I \frac{df}{d\lambda}(x, \lambda) dx \quad \forall \lambda \in J.$$

4. Transformada de Laplace

Definición 4.1. (*Transformada de Laplace*)

Si $f(t)$ es una función de valores reales en el intervalo $[0, \infty)$, se define la transformada de Laplace de f , como la integral convergente para ciertos valores de s , tales que,

$$(1.15) \quad L[f(s)] = \int_0^\infty e^{-st} f(t) dt.$$

Esto es, el valor obtenido a partir de la integral es finito.

Observación.

Una condición que garantiza la existencia de la Transformada es que la función f sea suave a trozos en cada intervalo de la forma $[0, t_0]$, siendo $t_0 > 0$, es decir, si f es continua y tiene derivada continua $f'(t)$ excepto a lo sumo en un número finito de puntos de discontinuidad entonces el integrando será suave a trozos y la integral existirá.

Definición 4.2. (*Orden Exponencial*)

Se dice que una función f es de *orden exponencial* en $[0, \infty)$, si existen las constantes $\mathbf{C}$ y α , con $\mathbf{C} > 0$, tales que

$$|f(t)| \leq \mathbf{C}e^{\alpha t}.$$

para toda $t > 0$.

Observación.

La siguiente afirmación es una consecuencia inmediata de un teorema de comparación bien conocido del análisis; a saber: si f y g son integrables en cada intervalo de la forma $[a, b]$, donde a es fijo y $b > a$ es arbitrario, y si $|f(t)| \leq g(t)$ para todo $t \geq a$, entonces $\int_a^\infty f(t) dt$ existe siempre que $\int_a^\infty g(t) dt$ exista.

Teorema 4.1.

Si f es una función suave a trozos de orden exponencial existe un número real α tal que

$$(1.16) \quad \int_0^\infty e^{-st} f(t) dt$$

converge para todos los valores $s > \alpha$.

DEMOSTRACIÓN.

Aceptando la veracidad de este resultado, escójanse $\mathbf{C}$ y α tales que $|f(t)| \leq \mathbf{C}e^{\alpha t}$ para todo $t > 0$. Entonces,

$$(1.17) \quad \int_0^\infty e^{-st} (\mathbf{C}e^{\alpha t}) dt = \mathbf{C} \int_0^\infty e^{-(s-\alpha)t} dt = \frac{\mathbf{C}}{s-\alpha} \quad \text{si } s > \alpha,$$

y el teorema de comparación implica que $\int_0^\infty f(t) dt$ existe para todo $s > \alpha$.

□

Teorema 4.2.

Sea f continua en $(0, \infty)$, y supóngase que f' es suave a trozos y de orden exponencial en $[0, \infty)$. Entonces,

$$(1.18) \quad L[f] = sL[f] - f(0^+),$$

donde $f(0^+) = \lim_{t \rightarrow 0^+} f(t)$. En forma más general, si $f, f', \dots, f^{(n-1)}$ son continuas para todo $t > 0$ y si $f^{(n)}$ es suave a trozos y de orden exponencial en $[0, \infty]$, entonces,

$$(1.19) \quad L[f''] = s^2 L[f] - sf(0^+) - f'(0^+)$$

$$\vdots$$

$$(1.20) \quad L[f^{(n)}] = s^n L[f] - s^{n-1}f(0^+) - s^{n-2}f'(0^+) - \dots - f^{(n-1)}(0^+).$$

DEMOSTRACIÓN.

Para establecer (1.19) se utiliza la integración por partes para evaluar $L[f']$, como sigue:

$$\begin{aligned} L[f'] &= \int_0^\infty e^{-st} f'(t) dt \\ &= e^{-st} f(t) \Big|_0^\infty + s \int_0^\infty e^{-st} f(t) dt \\ &= sL[f] + e^{-st} f(t) \Big|_0^\infty, \end{aligned}$$

y la demostración se completará si se puede demostrar que $e^{-st} f(t) \Big|_0^\infty = -f(0^+)$.

Con esto se observa que ya que f es de orden exponencial, $e^{-st} f(t) \Big|_0^\infty \rightarrow 0$, cuando $t \rightarrow \infty$ siempre que s sea suficientemente grande. Así, $e^{-st} f(t) \Big|_0^\infty$ se anula en su límite superior y tomando en consideración el hecho de que f puede tener una discontinuidad por salto en el origen, se tiene

$$\begin{aligned}
e^{-st}f(t) \Big|_0^\infty &= \lim_{t \rightarrow \infty} e^{-st}f(t) - \lim_{t \rightarrow 0} e^{-st}f(t) \\
&= -\lim_{t \rightarrow 0} e^{-st}f(t) \\
&= -f(0^+)
\end{aligned}$$

como se requería.

□

Ejemplos 4.1.

(i) Sea $f(t) = -\frac{1}{a} \cos(at)$. Entonces $f'(t) = \sin(at)$ y así, usando (1.19),

$$\begin{aligned}
L[\sin(at)] &= -\frac{s}{a} L[\cos(at)] + \frac{1}{a} \\
&= -\frac{s}{a} \left(\frac{s}{s^2 + a^2} \right) + \frac{1}{a} \\
&= \frac{a}{s^2 + a^2}, \quad s > 0.
\end{aligned}$$

(ii) Ya que $D^n t^n = n!$, tenemos que,

$$L[D^n t^n] = L[n!] = n! L[1] = \frac{n!}{s}, \quad s > 0.$$

Por otra parte (1.20) da

$$\begin{aligned}
L[D^n t^n] &= s^n L[t^n] - s^{n-1} \cdot 0 - \dots - 0 \\
&= s^n L[t^n].
\end{aligned}$$

En consecuencia, $s^n L[t^n] = \frac{n!}{s}$ para todo entero n no negativo y

$$L[t^n] = \frac{n!}{s^{n+1}}, \quad s > 0.$$

4.1. Propiedades adicionales de la Transformada de Laplace.

Teorema 4.3.

Si $L[f] = \varphi(s)$, entonces

$$L[e^{at}f(t)] = \varphi(s - a).$$

DEMOSTRACIÓN.

Sabemos que,

$$L[f](s) = \int_0^{\infty} e^{-st} f(t) dt.$$

Entonces,

$$\begin{aligned} L[e^{at}f(t)] &= \int_0^{\infty} e^{-st} e^{at} f(t) dt && \text{(Por propiedades de la función exponencial)} \\ &= \int_0^{\infty} e^{-(s-a)t} f(t) dt \\ &= L[f(t)](s - a) \\ &= \varphi(s - a). \end{aligned}$$

□

Teorema 4.4

Si $L[f] = \varphi(s)$, entonces

$$L[t^n f(t)] = (-1)^n \frac{d^n}{ds^n} \varphi(s).$$

DEMOSTRACIÓN.

Este resultado se establece derivando ambos lados de la ecuación

$$\varphi(s) = \int_0^{\infty} e^{-st} f(t) dt$$

n veces con respecto a s . Así

Para $n=1$

$$\frac{d}{ds} \varphi(s) = \frac{d}{ds} \int_0^{\infty} e^{-st} f(t) dt.$$

Puesto que f es derivable, integrable y además tiene derivada continua, se cumple la derivación bajo la integral (1.15), por tanto,

$$\begin{aligned} \frac{d}{ds} \varphi(s) &= \int_0^{\infty} \frac{d}{ds} [e^{-st} f(t)] dt \\ &= - \int_0^{\infty} e^{-st} t f(t) dt \\ &= -L[tf(t)]. \end{aligned}$$

Para $n=2$

$$\frac{d^2}{ds^2} \varphi(s) = - \frac{d}{ds} \int_0^{\infty} e^{-st} t f(t) dt.$$

Nuevamente, a partir de la definición de f se cumple la derivación bajo la integral y así, a partir del Teorema 3.1,

$$\begin{aligned} \frac{d^2}{ds^2} \varphi(s) &= - \int_0^\infty \frac{d}{ds} [e^{-st} t f(t)] dt \\ &= \int_0^\infty e^{-st} t^2 f(t) dt \\ &= L[t^2 f(t)]. \end{aligned}$$

Suponiendo cierto el hecho de que para $n = k - 1$ se cumple que,

$$L[t^{k-1} f(t)] = (-1)^{k-1} \frac{d^{k-1}}{ds^{k-1}} \varphi(s).$$

Entonces, para $n=k$

$$\frac{d^k}{ds^k} \varphi(s) = \frac{d}{ds} \int_0^\infty (-1)^{k-1} t^{k-1} e^{-st} f(t) dt$$

Nuevamente, por (1.15),

$$\begin{aligned} \frac{d^k}{ds^k} \varphi(s) &= (-1)^{k-1} \int_0^\infty \frac{d}{ds} t^{k-1} e^{-st} f(t) dt \\ &= (-1)^k \int_0^\infty t^k e^{-st} f(t) dt \\ &= L[t^k f(t)]. \end{aligned}$$

□

Definición 4.3. (*Función Característica*)

Sea X una variable aleatoria y F_X su función de distribución. Definimos a la función característica de X por la función $\phi_X : R \rightarrow C$ asociada a F_X de la siguiente manera

$$\phi_X = E[e^{itX}] = E[\cos(tX)] + iE[\sin(tX)].$$

Observación.

Con la notación usual para la esperanza, considerando el caso particular en que f es densidad de una variable aleatoria no negativa, se tiene que

$$L[f](s) = E[e^{-sX}].$$

Ejemplos 4.2.

(i) Sea X tal que asume los valores $0, 1, 2, \dots$, con probabilidades $p_0, p_1, p_2, \dots$, entonces,

$$L[f](s) = \sum_{i=0}^n p_i e^{-is},$$

donde la función generatriz viene dada por $G(\lambda) = E[\lambda^X] = \sum_{i=0}^n p_i \lambda^i$.

Así, $L[f] = G(e^{-s})$ y la transformada de Laplace difiere de la función generatriz al cambiar la variable $\lambda = e^{-s}$. Esto explica la analogía entre las propiedades de las transformadas de Laplace y las funciones generatrices.

(ii) La distribución gamma con densidad $f_\alpha(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)} e^{-x}$ tiene la transformada

$$L[f] = \frac{1}{\Gamma(\alpha)} \int_0^\infty e^{-(s-1)x} x^{\alpha-1} dx,$$

$$L[f] = \frac{1}{(s+1)^\alpha}, \quad \alpha > 0.$$

Observación.

A partir del Teorema 4.4, usando la notación para la esperanza, obtenemos que,

$$L'[f] = -E[Xe^{-sX}].$$

De forma más general, inductivamente podemos afirmar que,

$$L^{(n)}[f] = (-1)^n E[X^n e^{-sX}].$$

4.2. Cálculo de Momentos.

Es posible usar la función generatriz para obtener momentos de cualquier orden. Se calcula la derivada de L y se evalúa en $t = 0$. En el caso discreto, se puede calcular la derivada de la serie correspondiente derivando término a término. De manera similar, en el caso continuo, la derivada de la integral se calcula como la integral de la derivada del integrando. Esto significa que,

$$(1.21) \quad \varphi(s) = \frac{d}{dt} E[e^{-sX}] = E \left[\frac{d}{dt} (e^{-sX}) \right] = E[-X e^{-sX}].$$

En el caso en que X es una variable aleatoria discreta se tiene que,

$$(1.22) \quad E[-X e^{-sX}] = \sum_{x: P(X=x) > 0} -x e^{-sx} P(X = x).$$

En el caso en que X es una variable aleatoria continua,

$$(1.23) \quad E[-X e^{-sX}] = \int_{-\infty}^{\infty} -x e^{-sx} f(x) dx.$$

Por lo tanto

$$\begin{aligned} \varphi(0) &= -E[X], \\ \varphi'(0) &= E[X^2], \\ &\vdots \\ \varphi^{(n)}(0) &= (-1)^n E[X^n]. \end{aligned}$$

5. Integrales de Línea

5.1. Integral de línea de un campo escalar.

Un camino γ en $A \subseteq R^n$ es una función continua de un intervalo $[a, b]$ de R en A . Se dice en este caso que el camino γ une los puntos $p = \gamma(a)$ y $q = \gamma(b)$.

El camino se llama *regular* si existe la derivada γ' y es continua en el intervalo (a, b) . El camino se llama *regular a trozos* si el intervalo $[a, b]$ puede descomponerse en un número finito de subintervalos en cada uno de los cuales el camino es regular.

Definición 5.1. (*Integral de Línea de un Campo Escalar*)

Sea $f : \Omega \rightarrow R$ un campo escalar continuo, con $\Omega \subseteq R^n$, y sea $\gamma : [a, b] \rightarrow \Omega$ un camino regular a trozos. La integral de línea de f a lo largo de γ es, por definición:

$$\int_{\gamma} f \, dl = \int_a^b f(\gamma(t)) \, \|\gamma'(t)\| \, dt$$

5.1.1. Existencia de la integral $\int_{\gamma} f \, dl$.

Está asegurada, ya que el integrando es una función acotada en $[a, b]$ y continua salvo, a lo sumo, en un número finito de puntos. De hecho, si hacemos una partición $a = t_0 < t_1 < \dots < t_n = b$ del intervalo $[a, b]$ de forma que, para $k = 1, 2, \dots, n$, la restricción de γ al subintervalo $[t_{k-1}, t_k]$ sea de clase C^1 , es decir, si existen todas las derivadas parciales de la función y además son continuas, entonces, podemos escribir

$$\int_{\gamma} f \, dl = \sum_{k=1}^n \int_{t_{k-1}}^{t_k} f(\gamma(t)) \, \|\gamma'(t)\| \, dt ,$$

obteniendo una suma finita de integrales de funciones continuas. Resaltamos que al campo escalar f sólo se le exige estar definido y ser continuo sobre la curva Γ recorrida por el camino de integración. Habitualmente f tendrá propiedades de regularidad mucho mejores, siendo por ejemplo diferenciable en un abierto Ω que contenga a la curva Γ .

Ejemplo 5.1.

Consideremos el campo escalar f definido en R^3 por $f(x) = \|x\|^2$ ($x \in R^3$) y el camino helicoidal γ dado por:

$$\gamma(t) = (\cos(t), \sin(t), t) \quad (0 \leq t \leq 4\pi).$$

En este caso tenemos claramente

$$f(\gamma(t)) = \cos^2(t) + \sin^2(t) + t^2 = 1 + t^2 \quad (0 \leq t \leq 4\pi).$$

Y también

$$\gamma'(t) = (-\sin(t), \cos(t), 1), \|\gamma'(t)\| = \sqrt{2}, \quad (0 \leq t \leq 4\pi).$$

Con lo cual,

$$\int_{\gamma} f \, dl = \int_0^{4\pi} (1 + t^2) \sqrt{2} \, dt = 4\pi \sqrt{2} \left(1 + \frac{16\pi^2}{3}\right).$$

5.2. Integral de línea de un campo vectorial.

Definición 5.2. (*Integral de Línea de un Campo Vectorial*)

Sea ahora $\mathbf{F} : \Omega \rightarrow R^n$ un campo vectorial continuo en un conjunto $\Omega \subseteq R^n$ y $\gamma : [a, b] \rightarrow \Omega$ un camino regular a trozos. La integral de línea de $\mathbf{F}$ a lo largo de γ es, por definición:

$$\int_{\gamma} \mathbf{F} \cdot dl = \int_a^b \langle \mathbf{F}(\gamma(t)), \gamma'(t) \rangle \, dt$$

La integral $\int_{\gamma} \mathbf{F} \cdot dl$ representa el trabajo realizado por el campo en ese recorrido.

La existencia de esta integral está asegurada por las mismas razones comentadas en el caso de un campo escalar.

Ejemplo 5.2.

Consideremos el campo vectorial definido en R^3 por $\mathbf{F}(x) = x$ ($x \in R^3$) y el camino helicoidal $x = \cos(t); y = \sin(t); z = t$ ($0 \leq t \leq 4$).

Tenemos entonces

$$\begin{aligned}\langle \mathbf{F}(\gamma(t)) | \gamma'(t) \rangle &= \langle \cos(t)\mathbf{i} + \sin(t)\mathbf{j} + t\mathbf{k} | -\sin(t)\mathbf{i} + \cos(t)\mathbf{j} + \mathbf{k} \rangle \\ &= -\sin(t)\cos(t) + \sin(t)\cos(t) + t = t \quad (0 \leq t \leq 4),\end{aligned}$$

con lo que,

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{l} = \int_0^{4\pi} t \, dt = 8\pi^2.$$

5.3. Relación entre las integrales de línea.

Si $\gamma : [a, b] \rightarrow R^n$ es un camino suave, es decir, es regular con $\gamma' \neq 0$ para todo $t \in [a, b]$, podemos definir

$$\mathbf{T}(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|} \quad (a \leq t \leq b),$$

que es un vector unitario tangente a la curva Γ recorrida por el camino γ en cada punto. Si ahora $\mathbf{F}$ es un campo vectorial continuo sobre dicha curva tendremos:

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{l} = \int_a^b \langle \mathbf{F}(\gamma(t)), \gamma'(t) \rangle \|\gamma'(t)\| \, dt,$$

expresión que recuerda la integral de línea de un campo escalar. Así pues, en ciertas condiciones, para la curva Γ recorrida por el camino γ existirá un campo escalar continuo $f : \Gamma \rightarrow R$ que verifica:

$$f(\gamma(t)) = \langle \mathbf{F}(\gamma(t)), \mathbf{T}(t) \rangle \quad (a \leq t \leq b) \text{ y, por tanto, } \int_{\gamma} \mathbf{F} \cdot d\mathbf{l} = \int_{\gamma} f \, dl.$$

En resumen, bajo ciertas condiciones sobre el camino, toda integral de línea de un campo vectorial coincide con la integral de línea de un campo escalar, y es lógico analizar la relación entre ambos campos. Observamos que para cada $t \in [a, b]$, el vector $f(\gamma(t))\mathbf{T}(t)$ es la proyección ortogonal de $\mathbf{F}(\gamma(t))$ sobre el vector $\mathbf{T}(t)$, es decir, la componente de $\mathbf{F}(\gamma(t))$ en la dirección tangencial a la curva Γ en el punto $\gamma(t)$. Esto se debe a que el vector $f(\gamma(t))\mathbf{T}(t)$ es perpendicular a la recta de proyección ($\mathbf{F}(\gamma(t))$), estableciéndose una relación entre todos los puntos del elemento proyectante con los proyectados. Por tanto, podemos ver el campo escalar f como la coordenada del campo vectorial F en dicha dirección tangencial a la curva Γ en cada punto de la misma.

En general, se interpreta a la integral de línea sobre γ de un campo vectorial $\mathbf{F}$ como la integral de línea del campo escalar que se obtiene al tomar la coordenada de $\mathbf{F}$ en la dirección tangencial al camino γ en cada punto.

5.4. Propiedades de las integrales de línea.

Linealidad. Las integrales dependen linealmente del campo que se integra. Más concretamente, se verifica que,

$$\int_{\gamma} (\alpha f + \beta g) dl = \alpha \int_{\gamma} f dl + \beta \int_{\gamma} g dl$$

Continuidad. Las integrales de línea también dependen de manera continua del campo que se integra; intuitivamente, pequeñas perturbaciones del campo dan lugar a pequeñas variaciones en la integral. Ello es consecuencia de las desigualdades que vamos a presentar.

Sea $\gamma : [a, b] \rightarrow R^n$ un camino regular a trozos que recorre una curva Γ , sea f un campo escalar continuo sobre Γ y supongamos que f está acotado en Γ por una constante k , es decir,

$$k \geq \max\{|f(x)| : x \in \Gamma\} = \max\{|f(\gamma(t))| : a \leq t \leq b\}.$$

Entonces si se tiene, claramente,

$$|\int_{\gamma} f dl| \leq kL(\gamma).$$

Se obtiene un resultado análogo al hacer el estudio para un campo vectorial F que sea continuo sobre Γ . De hecho, podemos considerar el campo escalar $\|F\|$, que también es continuo sobre Γ , y la desigualdad de Cauchy-Schwartz nos permite escribir:

$$|\int_{\gamma} \mathbf{F}.dl| \leq \int_{\gamma} \|F\| dl.$$

Naturalmente ahora, de la estimación

$$k \geq \max\{\|\mathbf{F}(x)\| : x \in \Gamma\} = \max\{\|\mathbf{F}(\gamma(t))\| : a \leq t \leq b\},$$

se deduce que,

$$|\int_{\gamma} \mathbf{F}.dl| \leq kL(\gamma).$$

Aditividad. Las integrales de línea son aditivas con respecto al camino de integración, en el sentido de que al recorrer consecutivamente dos caminos, las integrales se suman.

Más concretamente, sean $\gamma : [a, b] \rightarrow R^n$ y $\varsigma : [c, d] \rightarrow R^n$ caminos regulares a trozos consecutivos, esto es, verificando que $\gamma(b) = \varsigma(c)$, y consideremos el camino suma $\gamma \oplus \varsigma$. Si f y $\mathbf{F}$ son ,respectivamente, un campo escalar y un campo vectorial, ambos continuos sobre la unión de las curvas recorridas por γ y ς , se verifica que:

$$\int_{\gamma \oplus \varsigma} f dl = \int_{\gamma} f dl + \int_{\varsigma} f dl \text{ y } \int_{\gamma \oplus \varsigma} \mathbf{F}.dl = \int_{\gamma} \mathbf{F}.dl + \int_{\varsigma} \mathbf{F}.dl.$$

Para el camino opuesto, el comportamiento de ambas integrales no es el mismo, la de un campo escalar no se altera al cambiar el sentido del recorrido, mientras que la de un campo vectorial cambia de signo. Más concretamente, si f es un campo escalar y $\mathbf{F}$ un campo vectorial, ambos continuos sobre la curva recorrida por un camino regular a trozos γ , se tiene que,

$$\int_{\gamma_{op}} f dl = \int_{\gamma} f dl \text{ pero } \int_{\gamma_{op}} \mathbf{F}.dl = - \int_{\gamma} \mathbf{F}.dl.$$

CAPITULO 2

Distribución de Gompertz

1. Introducción

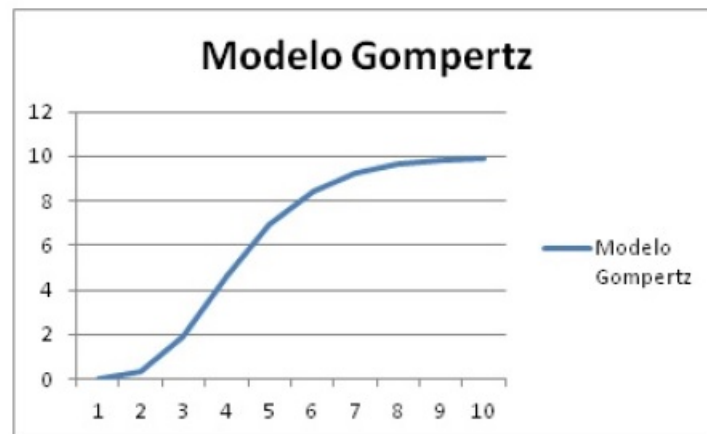
El nombre de la distribución de Gompertz, procede del matemático Benjamin Gompertz. En junio de 1825, Gompertz escribió al matemático Francis Baily ofreciéndole un artículo para la Real Sociedad de Matemáticos de Londres, que trataba sobre la naturaleza de la expresión de la función de la ley de mortalidad humana, y sobre una nueva forma de determinar el valor de lo que él denominaba “Life contingences” (y que se puede traducir por imprevistos de la vida).

Esta distribución se usa al considerar modelos asociados con curvas acotadas como las curvas de Gompertz. Estas curvas son conocidas por su forma en S o curvas sigmoidales, pues son monótonas y presentan un punto de inflexión donde la curva cambia su forma cóncava por la convexa. Estas curvas tienen un patrón de crecimiento, el cual reside en la localización del punto de inflexión; cuando se usan para modelar procesos de crecimiento, las curvas de Gompertz alcanzan este punto en la primera parte del ciclo de crecimiento (35 o 40 por ciento del crecimiento total).

El caso de la curva de Gompertz merece especial atención pues existen muchas expresiones de ella. De hecho, Gompertz introdujo la curva asociada al modelo de ley de mortalidad humana y la expresó como una doble exponencial. Desde entonces, la curva ha sido modificada

y reescrita en distintas formas para facilitar su estudio y entendimiento. Consecuentemente, hay una fórmula que engloba diferentes versiones de la curva, las cuales dependen de varios parámetros, con la doble exponencial como la característica más destacada.

La distribución de Gompertz tiene su importancia en la descripción de patrones de muerte en adultos. En adultos jóvenes se observa que, a menos que se encuentren factores externos, ajenos a la tradicional ley de vida, la tasa de mortalidad de este grupo de personas, o el riesgo de muerte, se mantiene a una velocidad prácticamente constante, a diferencia de los infantes y las personas de edad avanzada, las cuales son más propensas a padecer enfermedades o problemas de salud, lo cual propicia que la tasa de mortalidad aumente a una velocidad mayor que la de los adultos mencionados anteriormente. De aquí, se puede observar que al momento de realizar una gráfica que describa estos patrones, obtenemos que ésta representa la silueta de una S “chata” o “aplastada”. A continuación, se muestra una gráfica que describe este tipo de Curvas asociadas al modelo de Gompertz.



La desaceleración de la mortalidad representa un menor riesgo de muerte, como es el caso de los adultos jóvenes, como se mencionaba anteriormente. En un principio, la tasa de riesgo deja de aumentar con el tiempo, y luego procede a una velocidad constante, produciendo decaimiento exponencial, como en la desintegración radiactiva.

La distribución de Gompertz tiene densidad de probabilidad continua con un parámetro a , y un parámetro de forma b , llamado así, dado que indica la forma de la distribución.

La función de densidad de la distribución, con soporte en $(-\infty, \infty)$ viene dada por

$$(2.1) \quad f(x) = ae^{bx - \frac{a}{b}(e^{bx} - 1)},$$

En aplicaciones actuariales o demográficas, x usualmente denota edad, la cual no puede ser negativa, haciendo que el soporte sea en $(0, \infty)$. La función de distribución viene dada por,

$$(2.2) \quad F(x) = 1 - e^{-\frac{a}{b}(e^{bx} - 1)}.$$

2. Fuerza de Mortalidad de Gompertz.

Sea (x) un individuo, de edad x , donde $x \geq 0$. La muerte de (x) puede ocurrir a cualquier edad mayor a x , y se modela la vida futura de (x) a través de una variable aleatoria continua que denotaremos por T_x . Esto significa que $x + T_x$ es la variable aleatoria del momento de muerte de (x) .

Sea F_x la función de distribución de T_x tal que

$$F_x(t) = P(T_x \leq t).$$

Entonces, $F_x(t)$ representa la probabilidad de que (x) no sobreviva más allá de la edad $x + t$ y nos referimos a F_x como la **distribución de vida** a partir de la edad x . En muchos casos, se encuentra mayor interés en estudiar la probabilidad de supervivencia más que la probabilidad de muerte, por lo que definimos a S_x como

$$S_x(t) = 1 - F_x(t) = P(T_x > t).$$

Por tanto, $S_x(t)$ representa la probabilidad de que (x) sobreviva al menos t años, y $S_x(t)$ es conocida como la función de supervivencia.

Dada la interpretación de las variables aleatorias $\{T_x\}_{x>0}$ como la perspectiva de vida de un individuo después de alcanzar la edad “ x ”, necesitamos una conexión entre ellas. Para ello, consideremos T_0 y T_x para un individuo en particular, con edad x . Dicho individuo, pudo haber muerto antes de haber alcanzado la edad x , donde la probabilidad de este evento sería $P(T_0 < x)$. En caso de haber sobrevivido a la edad x , tal que $T_0 > x$, el tiempo restante de vida de este individuo, una vez que ha sobrevivido x años viene representado por T_x . Si el individuo muere en los siguientes t años, entonces $T_x < t$ y $T_0 < x + t$. Se requiere que los eventos $(T_x \leq t)$ y $(T_0 \leq x + t)$ sean equivalentes, dado que el individuo sobrevive a la edad x . Esto es posible haciendo la siguiente suposición para todo $x \leq 0$ y para todo $t \leq 0$

$$(2.3) \quad P(T_x \leq t) = P(T_0 \leq x + t | T_0 > x).$$

Ahora, para dos eventos A y B

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

De aquí,

$$P(T_x \leq t) = \frac{P(x < T_0 \leq x + t)}{P(T_0 > x)},$$

esto es,

$$(2.4) \quad F_x(t) = \frac{F_0(x + t) - F_0(x)}{S_0(x)}.$$

Además, usando $S_x(t) = 1 - F_x(t)$,

$$(2.5) \quad S_x(t) = \frac{S_0(x + t)}{S_0(x)}.$$

Cualquier función de supervivencia para una distribución de vida debe satisfacer las siguientes condiciones:

- (i) $S_x(0) = 1$, esto es, la probabilidad de que un individuo con edad x sobreviva 0 años es 1.
- (ii) $\lim_{t \rightarrow \infty} S_x(t) = 0$; esto es, todos los individuos eventualmente mueren.
- (iii) La función de supervivencia debe ser una función decreciente de t .

Ejemplo 2.1

Sea $F_0(t) = 1 - (1 - t/120)^{1/5}$ para $0 \leq t \leq 120$.

Calcule la probabilidad que

- (a) Un recién nacido sobreviva más allá de los 30.
- (b) Un individuo de 30 años muera antes de los 50.
- (c) Un individuo de 40 años sobreviva más de 65 años.

(a). $S_0(30) = 1 - F_0(30) = 1 - (1 - 30/120)^{1/5} = 0,9532$.

(b). $F_{30}(20) = \frac{F_0(50) - F_0(30)}{1 - F_0(30)} = 0,0489$.

(c). $S_{40}(25) = \frac{S_0(65)}{S_0(40)} = 0,9395$.

Cabe destacar, en el ejemplo anterior, que $S_0(120) = 0$, lo cual significa que en este modelo, supervivencia mayor a los 120 años no es posible. En este caso, 120 se refiere a la edad límite del modelo. En general, si hay una edad límite, se utiliza la letra ω para denotarla.

Notación Actuarial.

La notación usada anteriormente, $S_x(t)$, $F_x(t)$, $T_x(t)$ es estándar en estadística. Las ciencias actuariales han desarrollado su propia notación. La notación actuarial internacional, la cual encapsula las probabilidades y funciones de gran interés y uso para actuarios. La

notación actuarial para probabilidades de supervivencia y mortalidad son,

$$(2.6) \quad {}_t p_x = P(T_x > t) = S_x(t).$$

$$(2.7) \quad {}_t q_x = P(T_x \leq t) = 1 - S_x(t) = F_x(t).$$

$$(2.8) \quad {}_u | {}_t q_x = P(u < T_x \leq u + t) = S_x(u) - S_x(u + t).$$

- ${}_t p_x$ es la probabilidad de que (x) sobreviva por lo menos a la edad $x + t$.
- ${}_t q_x$ es la probabilidad de que (x) muera t años después de cumplir x años.
- ${}_u | {}_t q_x$ es la probabilidad de que (x) sobreviva u años, después de cumplir x años y muera en los siguientes t años. Esta es llamada la probabilidad de mortalidad diferida, puesto que es la probabilidad de que la muerte ocurra en un intervalo seguido de un periodo diferido.

En el caso mas usual, donde $t = 1$ y x es un número entero, la notación ${}_1 q_x$ se remplaza por q_x y ${}_1 p_x$ por p_x .

Las relaciones que se muestran a continuación son consecuencia de las definiciones y resultados mostrados anteriormente. Estas son:

$$\begin{aligned} {}_t p_x + {}_t q_x &= 1, \\ {}_u | {}_t q_x &= {}_u p_x - {}_{u+t} p_x. \end{aligned}$$

2.1. Tablas de Vida.

En el campo actuarial, un objetivo primario es estimar un patrón de mortalidad que exhiba un grupo de individuos. Un dispositivo básico para conseguir esto son las llamadas tablas de vida. (También conocidas como tablas de mortalidad).

Dado un modelo de supervivencia, con probabilidad de supervivencia ${}_t p_x$, se puede construir la tabla de vida para el modelo, a partir de una edad inicial x_0 y una edad máxima ω . Sea l_0 un número arbitrario, generalmente se toma una cifra redonda, como el número 100,000. Supongamos que empezamos con un grupo l_0 de recién nacidos. Nos gustaria predecir cuántos de estos individuos seguirán vivos en un momento dado en el futuro. Es claro que, no se pueden esperar unos cálculos exactos, pero se puede llegar a un estimado muy cercano al valor exacto a partir de buenas estadísticas.

Sea l_x el grupo de individuos con edad inicial 0 que seguirán vivos a la edad de x años, y sea d_x el grupo de individuos, nuevamente con edad inicial 0 que mueren entre las edades correspondientes a x y $x + 1$. La relación básica entre estas cantidades es

$$l_{x+1} = l_x - d_x.$$

Una tabla de vida es la tabulación de l_x y d_x , donde x es un entero positivo. Esta tabla terminará al alcanzar un valor máximo, generalmente denotado por ω , tal que $l_\omega = 0$. Este valor, denota la primera edad en la que todo el grupo original de individuos habrá muerto. El valor de ω varía dependiendo de cada tabla de vida en particular, pero generalmente se toma como un número mayor o igual a 110.

A pesar de que se puede predecir el valor de l_x , existe cierta aleatoriedad en estos modelos, dado que se desconoce si un individuo en particular se encontrará entre los sobrevivientes en algún momento dado. Para ello, retomamos la notación actuarial presentada anteriormente.

Para t y x enteros,

$$(2.9) \quad {}_t p_x = \frac{l_{x+t}}{l_x}.$$

Si nos preguntamos cuál es el significado de este término, basta con considerar l_x como los sobrevivientes con edad x , l_{x+t} los sobrevivientes para la edad $x+t$, y por consiguiente, el cociente en la ecuación anterior nos proporciona la probabilidad de que un individuo con edad x , sobreviva para la edad $x+t$.

Sea

$$(2.10) \quad {}_t q_x = \frac{l_x - l_{x+t}}{l_x}.$$

Esta ecuación representa la probabilidad de que (x) muera entre las edades x y $x+t$.

Para cualquier $x \geq x_0$ podemos interpretar a l_{x+t} como el valor esperado de sobrevivientes con edad $x+t$ a partir de l_x individuos independientes con edad x . Esta interpretación resulta más natural si consideramos a l_x como un entero, y se sigue puesto que el número de sobrevivientes a la edad $x+t$ es una variable aleatoria con distribución binomial y parámetros l_x y ${}_t p_x$. Esto es, suponiendo que se tienen l_x individuos independientes con edad x , cada uno con una probabilidad de supervivencia ${}_t p_x$ para la edad $x+t$.

En teoría, una tabla de vida es definida para todo x , desde la edad inicial x_0 hasta la edad máxima ω . En la práctica, es muy común que las tablas de vida se definan sólo para edades enteras.

Es usual en las tablas de vida, para edades enteras tabuladas, mostrar valores de d_x , donde

$$(2.11) \quad d_x = l_x - l_{x+1}.$$

Además, podemos interpretar a d_x como el valor esperado de muertes en el año desde la edad x hasta $x + 1$ para los l_x individuos con edad x , por tanto, usando la distribución binomial, donde se considera éxito la supervivencia y fracaso, el fallecimiento de un individuo en particular, se tiene

$$(2.12) \quad d_x = l_x \left(1 - \frac{l_{x+1}}{l_x} \right) = l_x(1 - p_x) = l_x q_x.$$

Una tabla de vida proporciona exactamente la misma información que una función de supervivencia S_{x0} . Sin embargo, una tabla de vida tabulada sólo para edades enteras no contiene toda la información correspondiente a su modelo de supervivencia, puesto que los valores enteros de l_x no son suficiente para calcular la probabilidad que involucra a los valores no enteros.

2.2. Fuerza de Mortalidad.

¿Cuál es la probabilidad de que (x) muera en algún momento en particular? Podríamos aproximar este resultado a partir de ${}_h q_x$, pero si asumimos continuidad y tomamos el límite cuando h tiende a 0, obtendríamos como resultado 0. Este resultado no es relevante ya que no nos da información alguna acerca de la mortalidad de (x) en ese momento. En lugar de esto, se puede calcular una tasa anual de mortalidad para (x) , dividiendo por h antes de tomar el límite. Esto nos lleva a la siguiente definición.

Definición 2.1. (*Fuerza de Mortalidad*).

La fuerza de mortalidad se define como el límite del cociente entre la proporción de los fallecidos en tiempo h , ${}_h q_x$, por unidad de tiempo, cuando h tiende a cero. La fuerza de mortalidad para (x) es entonces la cantidad

$$(2.13) \quad \mu(x) = \lim_{h \rightarrow 0} \frac{{}_h q_x}{h} = \lim_{h \rightarrow 0} \frac{l_x - l_{x+h}}{h l_x} = -\frac{\frac{d}{dx} l_x}{l_x} = \frac{d}{dx} \log(l_x).$$

La ecuación anterior representa un grupo de individuos cuyas vidas se encuentran en declive en función del tiempo debido a la mortalidad. La cantidad $\mu(x)$ nos da la tasa de declive relativa en este grupo de individuos para la edad x .

En muchos casos, este estudio se realiza para una edad fija x y tenemos una variable de tiempo t . De aquí, definimos,

$$\mu_x(t) = \mu(x + t).$$

Se puede considerar a $\mu_x(t)$ como la fuerza de mortalidad en el tiempo t para un individuo de edad x en el tiempo 0. A partir de la cuarta expresión en (2.13)

$$(2.14) \quad \mu_x(t) = -\frac{\frac{d}{dt} {}_t p_x}{{}_t p_x},$$

Teniendo en cuenta que la lógica nos impulsa a afirmar que la fuerza de mortalidad aumenta con la edad, ésta fuerza de mortalidad sólo tendría validez en periodos cortos. La función de supervivencia asociada a esta fuerza será:

$$(2.15) \quad S(x) = {}_t p_x = e^{-\int_0^t \mu_x(r) dr} = e^{\mu_x t}.$$

Ejemplo 2.2

Supongamos que la fuerza de mortalidad esta dada por

$$\mu(x) = \frac{1}{\omega - x}, \quad x < \omega.$$

Encontremos una expresión para ${}_t p_x$.

Para $t < \omega - x$,

$$-\int_0^t \frac{1}{\omega - x - r} dr = \log(\omega - x - t) - \log(\omega - x) = \log\left(\frac{\omega - x - t}{\omega - x}\right).$$

Sustituyendo en (2.13)

$${}_tp_x = \left(\frac{\omega - x - t}{\omega - x} \right) = 1 - \frac{t}{\omega - x}.$$

2.2.1. Fuerza de Mortalidad de Gompertz.

Recordemos de (2.2) que la función de distribución de Gompertz viene dada por

$$F_x(t) = 1 - e^{-\frac{a}{b}(e^{bt}-1)}.$$

Por otro lado, a partir de (2.6) y (2.7), se sigue que

$${}_tp_x = S_x(t) = 1 - F_x(t).$$

De aquí, es claro que

$$\begin{aligned}
 (2.16) \quad \frac{d}{dt} {}_tp_x &= \frac{d}{dt} (1 - F_x(t)) \\
 &= \frac{d}{dt} e^{-\frac{a}{b}(e^{bt}-1)} \\
 &= -ae^{bt-\frac{a}{b}(e^{bt}-1)}.
 \end{aligned}$$

Así, la fuerza de mortalidad de Gompertz o Ley de Gompertz viene dada por,

$$(2.17) \quad \mu(x) = -\frac{\frac{d}{dt} {}_tp_x}{{}_tp_x} = ae^{bx} \quad a, b > 0.$$

En esta ley, propuesta en 1825, después del estudio de la edad específica de las tasas de mortalidad en ese momento, también se argumentó que por razones fisiológicas, la intensidad de la mortalidad (en términos de Gompertz, el agotamiento promedio del poder del hombre para evitar la muerte) ganó proporciones iguales en intervalos iguales de edad, dando lugar a una fuerza de la mortalidad que crece exponencialmente con la edad. El parámetro b por

lo general tiene su valor en el entorno del 1,09. La ley de Gompertz es también muy conveniente para cálculos aproximados rápidos, incluso cuando la curva de la mortalidad no es estrictamente de la forma de Gompertz.

En virtud de esta ley de Gompertz, la función de supervivencia S_x , tal como se define en (2.15) se puede escribir como

$$(2.18) \quad S_x(t) = e^{axe^{bt}}.$$

3. Distribución de Gompertz

3.1. Notación.

A continuación, se presentan algunas expresiones usadas al momento de obtener ciertos resultados relacionados con la distribución de Gompertz, cuya finalidad recae en facilitar la expresión de los cálculos.

Integral Exponencial.

Con la finalidad de proveer métodos informáticos los cuales permitan evaluar funciones tabuladas en rangos completos de valores reales de sus parámetros, se evidencia el uso frecuente de funciones auxiliares para eliminar la parte infinita de la función original en sus singularidades y argumentos auxiliares para hacer frente a infinidad de rangos. Un ejemplo de este tipo de funciones, es la exponencial integral, identificada con el símbolo E_i , la cual se define como,

$$E_i(x) = \int_{-\infty}^x \frac{e^t}{t} dt$$

donde, para valores positivos de la parte real de z , se puede expresar como

$$E_n(z) = \int_1^{\infty} \frac{e^{-tz}}{t^n} dt \quad \operatorname{Re}(z) \geq 0$$

Esta definición puede ser utilizada para valores positivos de x , pero a causa de la singularidad del integrando en cero, la integral debe ser interpretada en términos del valor principal de Cauchy, el cual es un método que permite asignar valores a ciertas integrales impropias que sino resultarían indefinidas. Para valores complejos del argumento, esta definición es ambigua a causa de los puntos de ramificación en 0 y en ∞ . En general, se realiza un corte en el eje real negativo y E_i puede ser definida mediante una continuación analítica en el resto del plano complejo.

Para valores positivos de la parte real de z , se puede expresar como

$$E_n(z) = \int_1^{\infty} \frac{e^{-tz}}{t^n} dt, \quad \operatorname{Re}(z) \geq 0.$$

Función Integro - Exponencial Generalizada.

La función integro-exponencial generalizada se define en términos de la integral exponencial. Esta función juega un papel importante en la teoría de los procesos de transporte y el flujo de fluido, pero todavía no existe un resumen unificado de sus propiedades en la literatura.

Sean s y z variables continuas y n y j representan enteros no negativos. Definimos la “función integro - exponencial generalizada” $E_s^j(z)$ como

$$(2.19) \quad E_s^j(z) = \frac{(-1)^j}{j!} \frac{\partial^j}{\partial s^j} E_s(z),$$

donde, $E_s(z)$ representa a la integral exponencial para $n = s$.

La representación integral viene dada por

$$(2.20) \quad E_s^0(z) \equiv E_s(z) = \int_1^{\infty} t^{-s} e^{-zt} dt.$$

Por otra parte, una definición general de la G - función de Meijer está dada por la siguiente integral de línea en el plano complejo

$$\begin{aligned}
 G_{p,q}^{m,n} \left(z \left| \begin{smallmatrix} (a_i)_{1,p} \\ (b_1)_{1,q} \end{smallmatrix} \right. \right) &= G_{p,q}^{m,n} \left(z \left| \begin{smallmatrix} a_1, a_2, \dots, a_p \\ b_1, b_2, \dots, b_q \end{smallmatrix} \right. \right) \\
 &= \frac{1}{2\pi i} \int_L \frac{\prod_{j=1}^m \Gamma(b_j + s) \prod_{i=1}^n \Gamma(1 - a_i - s)}{\prod_{i=n+1}^p \Gamma(a_i + s) \prod_{j=m+1}^q \Gamma(1 - b_j - s)} z^{-s} ds,
 \end{aligned}$$

o lo que es igual,

$$\begin{aligned}
 G_{p,q}^{m,n} \left(z \left| \begin{smallmatrix} (a_i)_{1,p} \\ (b_1)_{1,q} \end{smallmatrix} \right. \right) &= G_{p,q}^{m,n} \left(z \left| \begin{smallmatrix} a_1, a_2, \dots, a_p \\ b_1, b_2, \dots, b_q \end{smallmatrix} \right. \right) \\
 &= \frac{1}{2\pi i} \int_L \frac{\prod_{j=1}^m \Gamma(b_j - s) \prod_{i=1}^n \Gamma(1 - a_i + s)}{\prod_{i=n+1}^p \Gamma(a_i - s) \prod_{j=m+1}^q \Gamma(1 - b_j + s)} z^s ds,
 \end{aligned}$$

donde Γ denota la función gamma alrededor del contorno L . La definición se mantiene bajo las siguientes suposiciones:

- (i) $0 \leq m \leq q$ y $0 \leq n \leq p$, donde m, n, p y q son números enteros.
- (ii) $z \neq 0$

Las funciones de este tipo fueron introducidas por primera vez por Meijer en 1936 y luego generalizadas por Fox en 1961. Se aplican ampliamente en la teoría de probabilidades, estadísticas, teoría de la difusión de anomalías y la teoría de ecuaciones diferenciales fraccionadas.

Para obtener una teoría unificada de esta función y sus propiedades, veremos la representación de e^{-z} en términos de la función G de Meijer

$$(2.21) \quad e^{-z} = G_{1,2}^{2,0} \left(z \left| \begin{smallmatrix} \cdot \\ 0 \end{smallmatrix} \right. \right) = \frac{1}{2\pi i} \int_{L_0} \Gamma(-t) z^t dt.$$

Donde, nuevamente Γ representa la función Gamma alrededor del contorno L_0 .

Ahora, sustituyendo (2.23) en (2.22) obtenemos,

$$(2.22) \quad E_s(z) = G_{1,2}^{2,0}(z \mid_{0,s-1}^s) = \frac{1}{2\pi i} \int_{L_0} \frac{\Gamma(-t)z^t}{s-1-t} dt,$$

donde L_0 es el contorno dirigido negativamente encerrando el eje t no negativo y el polo en $t = s - 1$. El tercer término de la igualdad anterior es la representación integral del contorno de la G-función. Así, se obtiene la representación integral de contorno,

$$(2.23) \quad \begin{aligned} E_s^j(z) &= \frac{1}{2\pi i} \int_{L_0} \frac{\Gamma(-t)z^t}{(s-1-t)^{j+1}} dt \\ &= G_{j+1,j+2}^{j+2,0}(z \mid_{0,s-1,\dots,s-1}^{s,\dots,s}). \end{aligned}$$

Función Hipergeométrica Generalizada.

Una serie hipergeométrica generalizada es una serie de potencias en que la relación de coeficientes sucesivos indexados por n es una función racional de n . La serie, si es convergente, define una función hipergeométrica generalizada, que puede entonces ser definida sobre un dominio más amplio del argumento por continuación analítica. Ésta, es una técnica para extender el dominio de definición de una función analítica dada. Una extensión analítica por lo general tiene éxito en definir valores adicionales de la función, por ejemplo en una región nueva en la que una representación mediante series infinitas con la que se había definido inicialmente a la función era divergente.

Una serie hipergeométrica se define formalmente como una serie de potencias,

$$\beta_0 + \beta_1 z + \beta_2 z^2 + \dots = \sum_{n \geq 0} \beta_n z^n,$$

tal que, la relación de coeficientes sucesivos es una función racional de n . Es decir,

$$\frac{\beta_{n+1}}{\beta_n} = \frac{A(n)}{B(n)},$$

donde $A(n)$ y $B(n)$ son polinomios en n .

Por ejemplo, en el caso de la serie para la función exponencial,

$$1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots$$

$\beta_n = n!^{-1}$ y $\beta_{n+1}/\beta_n = 1/(n+1)$. Así que esto satisface la definición de $A(n) = 1$ y $B(n) = n+1$.

Se acostumbra a factorizar el término principal, por lo que β_0 se asume como 1. Los polinomios se pueden factorizar por factores lineales de la forma $(a_j + n)$ y $(b_k + n)$ respectivamente, donde a_j y b_k son números complejos.

Por razones históricas, se asume que $(1+n)$ es un factor de B . Si este no es el caso, entonces ambos A y B pueden ser multiplicados por este factor, el cual se cancela, por lo que los términos no cambian y no hay pérdida de generalidad.

La relación entre los coeficientes consecutivos ahora tiene la forma

$$\frac{c(a_1 + n) \dots (a_p + n)}{d(b_1 + n) \dots (b_q + n)(1 + n)},$$

con c y d los coeficientes principales de A y B .

Entonces, la serie tiene la forma,

$$1 + \frac{a_1 \dots a_p}{b_1 \dots b_q} \frac{cz}{1 \cdot d} + \frac{a_1 \dots a_p}{b_1 \dots b_q} \frac{(a_1 + 1) \dots (a_p + 1)}{1 \cdot (b_1 + 1) \dots (b_q + 1) \cdot 2} \left(\frac{cz}{d}\right)^2 + \dots$$

o,

$$1 + \frac{a_1(a_1 + 1) \dots a_p(a_p + 1)}{b_1(b_1 + 1) \dots b_q(b_q + 1)} \frac{z^2}{2!} + \dots$$

con

$$(a)_0 = 1$$

$$(a)_n = a(a+1)(a+2) \dots (a+n-1), \quad n \geq 1.$$

Esto tiene la forma de una función generatriz exponencial. La notación estándar para esta serie es

$${}_pF_q \left[\begin{matrix} a_1 a_2 \dots a_p \\ b_1 b_2 \dots b_q \end{matrix}; z \right].$$

Por tanto,

$${}_pF_q \left[\begin{matrix} a_1 a_2 \cdots a_p \\ b_1 b_2 \cdots b_q \end{matrix}; z \right] = \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{z^n}{n!}.$$

Ahora, se presentan algunos resultados esenciales de la distribución de Gompertz.

3.2. Distribución de Gompertz.

La función generadora de momentos de una distribución de Gompertz, al igual que para cualquier otra distribución probabilística, viene dada por su transformada de Laplace, siempre que dicha transformada exista.

Proposición 2.1.

La transformada de Laplace de la distribución de Gompertz es

$$(2.24) \quad L[f(s)] = \frac{a}{b} e^{\frac{a}{b}} E_{\frac{s}{b}} \left(\frac{a}{b} \right) \quad a, b > 0,$$

$$\text{donde } E_n(z) = \int_1^{\infty} \frac{e^{-zt}}{t^n} dt.$$

DEMOSTRACIÓN.

La transformada de Laplace para la función de densidad de Gompertz dada en (2.1) es,

$$\begin{aligned}
 L[f(s)] &= \int_0^{\infty} e^{-sx} a e^{bx - \frac{a}{b}(e^{bx} - 1)} dx \\
 &= \int_0^{\infty} a e^{(bx - sx) - \frac{a}{b}(e^{bx} - 1)} dx \\
 &= a e^{\frac{a}{b}} \int_0^{\infty} e^{(bx - sx) - \frac{a}{b} e^{bx}} dx \\
 &= a e^{\frac{a}{b}} \int_0^{\infty} e^{x(b-s) - \frac{a}{b} e^{bx}} dx.
 \end{aligned}$$

Haciendo el cambio

$$q = e^{bx} \quad dq = b e^{bx} dx,$$

donde claramente, para $x \rightarrow 0$, $q \rightarrow 1$. Entonces,

$$L[f(s)] = \frac{a}{b} e^{\frac{a}{b}} \int_1^{\infty} \frac{e^{-\frac{a}{b}q}}{q^{\frac{s}{b}}} dq.$$

Nótese que la función integro - exponencial, $E_n(z)$, esta definida por,

$$E_n(z) = \int_1^{\infty} \frac{e^{-zt}}{t^n} dt \quad n > 0, \operatorname{Re}(z) > 0.$$

Por tanto,

$$L[f(s)] = \frac{a}{b} e^{\frac{a}{b}} E_{\frac{s}{b}} \left(\frac{a}{b} \right).$$

□

Los momentos de la Distribución de Gompertz se expresan como la derivada de su transformada de Laplace, la cual se puede expresar de forma general usando la función integro-exponencial generalizada.

Proposición 2.2.

El enésimo momento de una variable aleatoria X con distribución de Gompertz es,

$$(2.25) \quad E[X^n] = \frac{n!}{b^n} e^{\frac{a}{b}} E_1^{n-1} \left(\frac{a}{b} \right),$$

donde, $E_s^n(z) = \frac{1}{n!} \int_1^\infty (\ln(x))^n x^{-s} e^{-zx} dz$ es la función integro-exponencial generalizada.

DEMOSTRACIÓN.

Los momentos de una variable aleatoria X con distribución de Gompertz pueden ser calculados con la ayuda de la transformada de Laplace. A partir de (2.24),

$$(2.26) \quad L[f(s)] = \frac{a}{b} e^{\frac{a}{b}} E_{\frac{s}{b}} \left(\frac{a}{b} \right) \quad a, b > 0.$$

Para $n=1$

$$(2.27) \quad \begin{aligned} E[X] &= -\frac{d}{ds} L[f(s)] \\ &= -\frac{d}{ds} \left(\frac{a}{b} e^{\frac{a}{b}} \int_1^\infty \frac{e^{-\frac{a}{b}x}}{x^{\frac{s}{b}}} dx \right) \quad (\text{Por Teorema 3,1}) \\ &= -\frac{a}{b} e^{\frac{a}{b}} \int_1^\infty e^{-\frac{a}{b}x} \left(\frac{d}{ds} x^{-\frac{s}{b}} \right) dx. \end{aligned}$$

Para resolver ésta integral, consideremos

$$g(s) = x^{-\frac{s}{b}} \implies -\frac{s}{b} \ln(x) = \ln(g(s)).$$

Derivando a ambos lados de la igualdad con respecto a s ,

$$-\frac{1}{b} \ln(x) = \frac{1}{g(s)} g'(s).$$

Así, $g'(s) = -\frac{1}{b} x^{-\frac{s}{b}} \ln(x).$

De aquí,

$$(2.28) \quad E[X] = \frac{a}{b} e^{\frac{a}{b}} \int_1^{\infty} \frac{1}{b} e^{-\frac{a}{b}x} x^{-\frac{s}{b}} \ln(x) dx \Big|_{s=0} = \frac{a}{b} e^{\frac{a}{b}} \int_1^{\infty} \frac{1}{b} e^{-\frac{a}{b}x} \ln(x) dx.$$

Para $n=2$

$$(2.29) \quad \begin{aligned} E[X^2] &= -\frac{d^2}{ds^2} L[f(s)] \\ &= -\frac{d}{ds} \left(\frac{a}{b} e^{\frac{a}{b}} \int_1^{\infty} \frac{1}{b} e^{-\frac{a}{b}x} \ln(x) x^{-\frac{s}{b}} dx \right) \\ &= -\frac{a}{b} e^{\frac{a}{b}} \int_1^{\infty} \frac{1}{b} e^{-\frac{a}{b}x} \ln(x) \frac{d}{ds} (x^{-\frac{s}{b}}) dx. \end{aligned}$$

Nuevamente, para resolver ésta integral, consideremos

$$g(s) = x^{-\frac{s}{b}} \implies -\frac{s}{b} \ln(x) = \ln(g(s)).$$

Derivando a ambos lados de la igualdad con respecto a s ,

$$-\frac{1}{b} \ln(x) = \frac{1}{g(s)} g'(s).$$

Así, $g'(s) = -\frac{1}{b} x^{-\frac{s}{b}} \ln(x).$

De aquí, para (2.29)

$$(2.30) \quad E[X^2] = -\frac{a}{b} e^{\frac{a}{b}} \int_1^{\infty} \frac{1}{b^2} e^{-\frac{a}{b}x} x^{-\frac{s}{b}} (\ln(x))^2 dx \Big|_{s=0} = -\frac{a}{b} e^{\frac{a}{b}} \int_1^{\infty} \frac{1}{b^2} e^{-\frac{a}{b}x} (\ln(x))^2 dx.$$

A partir del teorema 4.4 de la sección de nociones básicas de Probabilidades, tenemos que para $n=k$,

$$(2.31) \quad E[X^k] = (-1)^{k+1} \frac{d^k}{ds^k} L[f(s)] = (-1)^{k+1} \frac{a}{b} e^{\frac{a}{b}} \int_1^{\infty} \frac{1}{b^k} e^{-\frac{a}{b}x} (\ln(x))^k dx.$$

Utilizando el método de integración por partes,

$$u = (\ln(x))^n \quad du = \frac{n}{x} (\ln(x))^{n-1} dx$$

$$dv = e^{-\frac{a}{b}x} dx \quad v = -\frac{b}{a} e^{-\frac{a}{b}x}$$

$$\frac{a}{b^{n+1}} e^{\frac{a}{b}} \int_1^{\infty} e^{-\frac{a}{b}x} (\ln(x))^n dx = \frac{a}{b^{n+1}} e^{\frac{a}{b}} \left[-\frac{b}{a} e^{-\frac{a}{b}x} (\ln(x))^n \Big|_1^{\infty} + \int_1^{\infty} \frac{b}{a} \frac{n}{x} e^{-\frac{a}{b}x} (\ln(x))^{n-1} dx \right].$$

$$\implies E[X^n] = e^{\frac{a}{b}} \int_1^{\infty} \frac{n}{b^n} x^{-1} e^{-\frac{a}{b}x} (\ln(x))^{n-1} dx.$$

Usando la función integro-exponencial generalizada, definida por,

$$E_s^j(z) = \frac{1}{\Gamma(j+1)} \int_1^{\infty} (\ln(x))^j x^{-s} e^{-zx} dx.$$

donde, $\Gamma(j) = (j-1)!$,

$$(2.32) \quad E_1^{n-1} \left(\frac{a}{b} \right) = \frac{1}{\Gamma(n)} \int_1^{\infty} (\ln(x))^{n-1} x^{-1} e^{-\frac{a}{b}x} dx = \frac{1}{(n-1)!} \int_1^{\infty} (\ln(x))^{n-1} x^{-1} e^{-\frac{a}{b}x} dx.$$

Por tanto,

$$(2.33) \quad E[X^n] = e^{\frac{a}{b}} \frac{n}{b^n} (n-1)! E_1^{n-1} \left(\frac{a}{b} \right) = \frac{n!}{b^n} e^{\frac{a}{b}} E_1^{n-1} \left(\frac{a}{b} \right).$$

Por otro lado, usando la función G de Meijer, definida por,

$$(2.34) \quad E_s^j(z) = G_{j+1, j+2}^{j+2, 0} \left(z \mid \begin{smallmatrix} s, \dots, s \\ 0, s-1, \dots, s-1 \end{smallmatrix} \right)$$

Se concluye que,

$$E[X^n] = \frac{n!}{b^n} e^{\frac{a}{b}} G_{n, n+1}^{n+2, 0} \left(\frac{a}{b} \mid \begin{smallmatrix} 1, \dots, 1 \\ 0, 0, \dots, 0 \end{smallmatrix} \right)$$

□

Observación.

La expansión en series de la función integro-exponencial ofrece una representación de los momentos de la distribución de Gompertz más simple, pero a su vez más larga. Tomando $s = 1$ y $z = \frac{a}{b}$,

$$(2.35) \quad E_1^{n-1} \left(\frac{a}{b} \right) = \left[\sum_{j=1}^{\infty} \frac{1}{(-j)^n} \frac{(-\frac{a}{b})^j}{j!} \right] + \frac{(-1)^n}{n!} \sum_{j=0}^n \binom{n}{j} \ln \left(\frac{a}{b} \right)^{n-j} \Psi_j,$$

donde,

$$(2.36) \quad \Psi_j = \lim_{t \rightarrow 0} \frac{d^j}{dt^j} \Gamma(1-t).$$

Para $j \in \{0, 1, 2, 3, 4\}$

$$\begin{aligned} \Psi_0 &= 1; \\ \Psi_1 &= \gamma; \\ \Psi_2 &= \gamma^2 + \frac{\Pi^2}{6}; \\ \Psi_3 &= \gamma^3 + \gamma \frac{\Pi^2}{2} + 2\zeta(3); \end{aligned}$$

$$\Psi_4 = \gamma^4 + \gamma^2 \Pi^2 + 8\gamma \zeta(3) + \frac{3}{20} \Pi^4;$$

donde, $\gamma \approx 0,57722$ es la constante de Euler - Mascheroni y $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ es la función zeta de Riemann. Nótese que $\frac{\Pi^2}{6} = \zeta(2)$ y $\frac{\Pi^4}{90} = \zeta(4)$. Además, $\zeta(3) \approx 1,2021$ es conocida como la constante de Apèry.

Proposición 2.3.

La esperanza de una variable aleatoria X con distribución de Gompertz es

$$(2.37) \quad E[X] = \frac{1}{b} e^{\frac{a}{b}} E_1\left(\frac{a}{b}\right) \approx \frac{1}{b} e^{\frac{a}{b}} \left(\frac{a}{b} - \ln\left(\frac{a}{b}\right) - \gamma\right).$$

Observación.

La precisión del valor resultante de la aproximación depende de la relación entre a y b . Para valores pequeños de $\frac{1}{b} e^{\frac{a}{b}} \frac{\left(\frac{a}{b}\right)^2}{4}$, la aproximación es precisa.

DEMOSTRACIÓN.

El valor esperado de una variable aleatoria X con distribución de Gompertz, o expectativa de vida, viene dado por,

$$\begin{aligned} E[X] &= \frac{1}{b} e^{\frac{a}{b}} \int_1^{\infty} x^{-1} e^{-\frac{a}{b}x} dx \\ &= \frac{1}{b} e^{\frac{a}{b}} E_1\left(\frac{a}{b}\right). \end{aligned}$$

Aproximación. La expansión en series del valor esperado es,

$$\begin{aligned} E[X] &= \frac{1}{b} e^{\frac{a}{b}} \left(- \sum_{j=1}^{\infty} \frac{\left(-\frac{a}{b}\right)^j}{j!} - \sum_{j=0}^1 \binom{1}{j} \ln\left(\frac{a}{b}\right)^{1-j} \Psi_j \right) \\ &= \frac{1}{b} e^{\frac{a}{b}} \left(- \sum_{j=1}^{\infty} \frac{\left(-\frac{a}{b}\right)^j}{j!} - \ln\left(\frac{a}{b}\right) - \gamma \right). \end{aligned}$$

Como a usualmente se acerca a 0, puesto que así está definido el modelo para ésta distribución en particular, $\frac{a}{b} \approx -\sum_{j=1}^{\infty} \frac{-\left(\frac{a}{b}\right)^j}{j!}$ y

$$E[X] \approx \frac{1}{b} e^{\frac{a}{b}} \left(\frac{a}{b} - \ln \left(\frac{a}{b} \right) - \gamma \right).$$

□

Nótese que como el signo de $-\sum_{j=1}^{\infty} \frac{\left(-\frac{a}{b}\right)^j}{j!}$ cambia, la ecuación anterior será ligeramente sobreestimada para valores de a relativamente grandes, pero mediante la adición de nuevos términos de la suma, la esperanza se puede aproximar por precisión arbitraria.

Proposición 2.4.

La varianza de una variable aleatoria X con distribución de Gompertz es,

$$(2.38) \quad Var[X] = \frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b^3} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] + \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln \left(\frac{a}{b} \right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right]^2$$

donde ${}_pF_q \left[\begin{matrix} a_1, a_2, \dots, a_p \\ b_1, b_2, \dots, b_q \end{matrix}; z \right]$ denota la hipergeometría generalizada.

DEMOSTRACIÓN.

Sabemos que la varianza viene dada por,

$$Var[X] = E[X^2] - E[X]^2.$$

De aquí,

$$Var[X] = e^{\frac{a}{b}} \int_1^{\infty} \frac{2}{b^2} x^{-1} e^{-\frac{a}{b}x} \ln(x) dx - \left[\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right]^2.$$

Concentrándonos en $E[X^2]$, si X tiene distribución de Gompertz, de (2.26)

$$(2.39) \quad E[X^2] = \frac{2}{b^2} e^{\frac{a}{b}} E_1^1 \left(\frac{a}{b} \right).$$

Una representación exacta de (2.40) se puede dar a partir de la función G de Meijer en (2.35):

$$E[X^2] = \frac{2}{b^2} e^{\frac{a}{b}} G_{2,3}^{3,0} \left(\frac{a}{b} \mid \begin{matrix} ; 1, 1 \\ 0, 0, 0; \end{matrix} \right).$$

De (2.36):

$$E_1^j(z) = \left[\sum_{l=1}^{\infty} \frac{1}{(-l)^{j+1}} \frac{(-z)^l}{l!} \right] + \frac{(-1)^{j+1}}{(j+1)!} \sum_{l=0}^{j+1} \binom{j+1}{l} \ln(z)^{1+j-l} \Psi_l,$$

donde,

$$\Psi_l = \lim_{t \rightarrow 0} \frac{d^l}{dt^l} \Gamma(1-t).$$

Nótese que,

$$\sum_{l=1}^{\infty} \frac{1}{(-l)^{j+1}} \frac{(-z)^l}{l!} = (-1)^j z_{j+2} F_{j+2} \left[\begin{matrix} 1_1, \dots, 1_{j+2} \\ 2_1, \dots, 2_{j+2}; \end{matrix} z \right],$$

donde ${}_pF_q \left[\begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q; \end{matrix} z \right] = \sum_{k=0}^{\infty} \frac{(a_1)_k \dots (a_p)_k}{(b_1)_k \dots (b_q)_k} \frac{z^k}{k!}$, denota la función hipergeométrica generalizada.

Por tanto,

$$\begin{aligned} E[X^2] &= \frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2; \end{matrix} \frac{a}{b} \right] + \frac{1}{2} \sum_{l=0}^2 \binom{2}{l} \ln \left(\frac{a}{b} \right)^{2-l} \psi_l \right\} \\ &= \frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2; \end{matrix} \frac{a}{b} \right] + \frac{1}{2} \left[\binom{2}{0} \ln \left(\frac{a}{b} \right)^2 \psi_0 + \binom{2}{1} \ln \left(\frac{a}{b} \right) \psi_1 + \binom{2}{2} \psi_2 \right] \right\} \end{aligned}$$

$$\begin{aligned}
&= \frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] + \frac{1}{2} \left[\ln \left(\frac{a}{b} \right)^2 + 2 \ln \left(\frac{a}{b} \right) \gamma + \gamma^2 + \frac{\pi^2}{6} \right] \right\} \\
&= \frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] + \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln \left(\frac{a}{b} \right) \right)^2 \right] \right\}
\end{aligned}$$

En conclusión,

$$Var[X] = \frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] + \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln \left(\frac{a}{b} \right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right]^2.$$

□

Proposición 2.5.

El sesgo κ de una variable aleatoria X con distribución de Gompertz es,

(2.40)

$$\kappa = \frac{\frac{6}{b^3} e^{\frac{a}{b}} G_{3,4}^{4,0} \left(\frac{a}{b} \mid \begin{matrix} 1, 1, 1 \\ 0, 0, 0, 0 \end{matrix} \right) - 3m_1 \frac{2}{b^2} e^{\frac{a}{b}} G_{2,3}^{3,0} \left(\frac{a}{b} \mid \begin{matrix} 1, 1 \\ 0, 0, 0 \end{matrix} \right) + 2(m_1)^3}{\left(\frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] + \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln \left(\frac{a}{b} \right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right]^2 \right)^{\frac{3}{2}}}$$

donde, m_1 denota el valor esperado de la distribución, es decir, la esperanza de la distribución de Gompertz.

DEMOSTRACIÓN.

Consideremos μ como el valor esperado de la variable aleatoria X y σ^2 su varianza. El sesgo viene dado por,

$$\begin{aligned}
\kappa &= \frac{E[(X - \mu)^3]}{\sigma^3} \\
&= \frac{E[X^3 - 3X^2\mu + 3X\mu^2 - \mu^3]}{\sigma^3}
\end{aligned}$$

$$\begin{aligned}
&= \frac{E[X^3] - 3E[\mu(X^2 - X\mu)] - \mu^3}{\sigma^3} \\
&= \frac{E[X^3] - 3E[X](E[X^2] - E[X]^2) - E[X]^3}{Var[X]^{\frac{3}{2}}}
\end{aligned}$$

Usando (2.12), (2.16) y el resultado obtenido en el Corolario 2.2 del capítulo de preliminares, obtenemos que,

$$\begin{aligned}
\kappa &= \frac{\frac{6}{b^3} e^{\frac{a}{b}} E_1^2\left(\frac{a}{b}\right) - \frac{3}{b} e^{\frac{a}{b}} E_1\left(\frac{a}{b}\right) \left[\frac{2}{b^2} e^{\frac{a}{b}} E_1^1\left(\frac{a}{b}\right) - \frac{1}{b^2} e^{2\frac{a}{b}} E_1\left(\frac{a}{b}\right)^2 \right] - \frac{1}{b^3} e^{3\frac{a}{b}} E_1\left(\frac{a}{b}\right)^3}{\left(\frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b^3} F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] + \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln\left(\frac{a}{b}\right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1\left(\frac{a}{b}\right) \right]^2 \right)^{\frac{3}{2}}} \\
&= \frac{\frac{6}{b^3} e^{\frac{a}{b}} E_1^2\left(\frac{a}{b}\right) - \frac{6}{b^3} e^{2\frac{a}{b}} E_1\left(\frac{a}{b}\right) E_1^1\left(\frac{a}{b}\right) + \frac{2}{b^3} e^{3\frac{a}{b}} E_1\left(\frac{a}{b}\right)^3}{\left(\frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b^3} F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] + \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln\left(\frac{a}{b}\right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1\left(\frac{a}{b}\right) \right]^2 \right)^{\frac{3}{2}}} \\
&= \frac{\frac{6}{b^3} e^{\frac{a}{b}} E_1^2\left(\frac{a}{b}\right) - 3 \left(\frac{1}{b} e^{\frac{a}{b}} E_1\left(\frac{a}{b}\right) \right) \left(\frac{2}{b^2} e^{\frac{a}{b}} E_1^1\left(\frac{a}{b}\right) \right) + 2 \left(\frac{1}{b} e^{\frac{a}{b}} E_1\left(\frac{a}{b}\right) \right)^3}{\left(\frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b^3} F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] + \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln\left(\frac{a}{b}\right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1\left(\frac{a}{b}\right) \right]^2 \right)^{\frac{3}{2}}} \\
&= \frac{\frac{6}{b^3} e^{\frac{a}{b}} E_1^2\left(\frac{a}{b}\right) - 3m_1 \frac{2}{b^2} e^{\frac{a}{b}} E_1^1\left(\frac{a}{b}\right) + 2(m_1)^3}{\left(\frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b^3} F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] + \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln\left(\frac{a}{b}\right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1\left(\frac{a}{b}\right) \right]^2 \right)^{\frac{3}{2}}}.
\end{aligned}$$

Por último, a partir de (2.35),

$$\kappa = \frac{\frac{6}{b^3} e^{\frac{a}{b}} G_{3,4}^{4,0} \left(\frac{a}{b} \mid \begin{matrix} ; 1, 1, 1 \\ 0, 0, 0, 0; \end{matrix} \right) - 3m_1 \frac{2}{b^2} e^{\frac{a}{b}} G_{2,3}^{3,0} \left(\frac{a}{b} \mid \begin{matrix} ; 1, 1 \\ 0, 0, 0; \end{matrix} \right) + 2(m_1)^3}{\left(\frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2; \end{matrix} \frac{a}{b} \right] + \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln \left(\frac{a}{b} \right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right]^2 \right)^{\frac{3}{2}}}$$

□

Proposición 2.6.

El coeficiente de Curtosis de una variable aleatoria X con distribución de Gompertz es,

$$\begin{aligned} \beta &= \left\{ \frac{24}{b^4} e^{\frac{a}{b}} G_{4,5}^{5,0} \left(\frac{a}{b} \mid \begin{matrix} ; 1, 1, 1, 1 \\ 0, 0, 0, 0, 0; \end{matrix} \right) - 4m_1 \frac{6}{b^3} e^{\frac{a}{b}} G_{3,4}^{4,0} \left(\frac{a}{b} \mid \begin{matrix} ; 1, 1, 1 \\ 0, 0, 0, 0; \end{matrix} \right) \right. \\ &+ 6(m_1)^2 \frac{2}{b^2} e^{\frac{a}{b}} G_{2,3}^{3,0} \left(\frac{a}{b} \mid \begin{matrix} ; 1, 1 \\ 0, 0, 0; \end{matrix} \right) + 3(m_1)^4 \left. \right\} \left\{ \frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2; \end{matrix} \frac{a}{b} \right] \right. \right. \\ &+ \left. \left. \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln \left(\frac{a}{b} \right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right]^2 \right\}^2 / -3. \end{aligned}$$

DEMOSTRACIÓN.

Sabemos que,

$$\beta = \frac{E[(X - \mu)^4]}{\sigma^4} - 3 = \frac{E[X^4] - 4E[X^3]E[X] + 6E[X^2]E[X]^2 - 4E[X]E[X]^3 + E[X]^4}{Var[X]^2} - 3$$

Luego, usando (2.36),

$$\begin{aligned}
\beta &= \left\{ \frac{24}{b^4} e^{\frac{a}{b}} E_1^3 \left(\frac{a}{b} \right) - 4 \left[\left(\frac{6}{b^3} e^{\frac{a}{b}} E_1^2 \left(\frac{a}{b} \right) \right) \left(\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right) \right] + 6 \left[\frac{2}{b^2} e^{\frac{a}{b}} E_1^1 \left(\frac{a}{b} \right) \left(\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right)^2 \right] \right. \\
&- 3 \left[\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right]^4 \left. \right\} / \left\{ \frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] \right. \right. \\
&+ \left. \left. \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln \left(\frac{a}{b} \right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right]^2 \right\}^2 - 3 \\
&= \frac{\frac{24}{b^4} e^{\frac{a}{b}} E_1^3 \left(\frac{a}{b} \right) - \frac{24}{b^4} e^{2\frac{a}{b}} E_1^2 \left(\frac{a}{b} \right) E_1 \left(\frac{a}{b} \right) + \frac{12}{b^4} e^{2\frac{a}{b}} E_1^1 \left(\frac{a}{b} \right) \left(E_1 \left(\frac{a}{b} \right) \right)^2 - \frac{3}{b^4} e^{4\frac{a}{b}} \left(E_1 \left(\frac{a}{b} \right) \right)^4}{\left\{ \frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] + \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln \left(\frac{a}{b} \right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right]^2 \right\}^2} - 3 \\
&= \frac{\frac{24}{b^4} e^{\frac{a}{b}} E_1^3 \left(\frac{a}{b} \right) - \frac{24}{b^3} e^{\frac{a}{b}} m_1 E_1^2 \left(\frac{a}{b} \right) + \frac{12}{b^2} e^{\frac{a}{b}} (m_1)^2 E_1^1 \left(\frac{a}{b} \right) - 3(m_1)^4}{\left\{ \frac{2}{b^2} e^{\frac{a}{b}} \left\{ -\frac{a}{b} {}_3F_3 \left[\begin{matrix} 1, 1, 1 \\ 2, 2, 2 \end{matrix}; \frac{a}{b} \right] + \frac{1}{2} \left[\frac{\pi^2}{6} + \left(\gamma + \ln \left(\frac{a}{b} \right) \right)^2 \right] \right\} - \left[\frac{1}{b} e^{\frac{a}{b}} E_1 \left(\frac{a}{b} \right) \right]^2 \right\}^2} - 3
\end{aligned}$$

Ahora, por (2.35),

$$E_1^3 \left(\frac{a}{b} \right) = G_{4,5}^{5,0} \left(\frac{a}{b} \mid ; 1, 1, 1, 1 \right. \\ \left. 0, 0, 0, 0, 0; \right),$$

$$E_1^2 \left(\frac{a}{b} \right) = G_{3,4}^{4,0} \left(\frac{a}{b} \mid ; 1, 1, 1 \right. \\ \left. 0, 0, 0, 0; \right),$$

$$E_1^1 \left(\frac{a}{b} \right) = G_{2,3}^{3,0} \left(\frac{a}{b} \mid ; 1, 1 \right. \\ \left. 0, 0, 0; \right).$$

y así, finalmente se obtiene el resultado deseado.

□

CAPITULO 3

Método de Máxima Verosimilitud.

1. Estimación por Máxima Verosimilitud.

1.1. Edad Continua.

En estadística aplicada, la censura es el fenómeno que ocurre cuando el valor de una muestra sólo se conoce parcialmente.

Por ejemplo, en el estudio del impacto de cierta medicina en la mortalidad que produce cierta enfermedad, puede saberse únicamente si cierto individuo está vivo a la edad de 75 años (que es cuando se le realizó el último seguimiento) pero no se sabe si sigue todavía vivo o no.

Decimos que una muestra se encuentra censurada hacia la derecha cuando dicha observación está por encima de cierto valor, pero se ignora cuánto.

Cuando se observa una muestra continua, de la distribución de muertes, disponible en su totalidad o censurada hacia la derecha, a partir de la distribución de Gompertz podemos

obtener su función de verosimilitud.

Sea X_i el tiempo de vida observado, ya sea en su totalidad o censurado hacia la derecha; consideremos $\delta_i = 1$ como el momento de muerte de un individuo i si éste logra ser estudiado, y $\delta_i = 0$ si se encuentra censurado a la derecha. En este caso, la densidad conjunta y la función de verosimilitud logarítmica de la distribución de Gompertz serán

$$p(X_i | a, b) = (ae^{bX_i})^{\delta_i} e^{-\frac{a}{b}} (e^{bX_i} - 1)$$

y

$$l(a, b | X) = \sum_i \delta_i (\ln(a) + bX_i) - \frac{a}{b} (e^{bX_i} - 1)$$

respectivamente.

Las funciones score en el caso continuo serán

$$\begin{pmatrix} S_a \\ S_b \end{pmatrix} = \begin{pmatrix} \sum_i \frac{\delta_i}{a} - \frac{e^{bX_i} - 1}{b} \\ \sum_i \frac{a}{b^2} (e^{bX_i} - 1) + \delta_i X_i - \frac{a}{b} e^{bX_i} X_i \end{pmatrix}$$

y por tanto,

$$\hat{a} = \frac{Db}{\frac{1}{n} \sum_i e^{bX_i} - 1}$$

donde, n representa el número de observaciones y $D = \frac{1}{n} \sum_i \delta_i$ es la proporción de muertes del total de las observaciones.

El estimador de máxima verosimilitud de b es la solución para

$$\sum_i \frac{De^{bX_i} X_i}{\frac{1}{n} \sum_i e^{bX_i} - 1} = \sum_i \left(\frac{D(e^{bX_i} - 1)}{\frac{b}{n} \sum_i e^{bX_i} - b} + \delta_i X_i \right).$$

Observaciones.

Si todos los X_i son momentos de muerte observados, denotando la fuerza de mortalidad promedio de la población como $\bar{\mu}$ y notando que $\bar{\mu} = \sum_i \frac{a}{n} e^{bX_i}$:

$$\bar{\mu} = \hat{a} + b.$$

Si $b = \hat{b}$ y la muestra viene dada por una distribución de Gompertz, la fuerza de mortalidad promedio observada es igual a $\hat{a} + \hat{b}$.

1.2. Edad Discreta.

Con frecuencia, sólo se encuentran disponibles muestras discretas y la fuerza de mortalidad de Gompertz no puede ser estimada de manera directa. Sin embargo, si el número de muertes y el número de años/personas expuestas a riesgo de muerte pueden ser observadas, se puede asumir que el número de muertes en un intervalo de edades dado sigue una distribución de Poisson.

Sea D el número de muertes y λ el parámetro de la distribución de Poisson

$$f(D; \lambda) = \frac{1}{D!} \lambda^D e^{-\lambda}.$$

El parámetro λ de la distribución de Poisson es consistente para $\mu(x)$, puesto que $\lambda, \mu(x) > 0$. Cuando se aplica la distribución de Poisson a muestras de mortalidad, la unidad de observación es el número de muertes en cada grupo de edad, D_x , y la ventana de observación es el número de personas que se encuentran expuestas a este riesgo de muerte. Claramente, el número de personas con riesgo a morir difiere de acuerdo al grupo de edad. Por tanto, λ tiene que sopesarse de acuerdo al número de personas expuestas a muerte, E_x , a cierta edad x . Luego, sustituyendo $\lambda = \lambda E_x$ en la ecuación anterior y tomando el logaritmo, además de eliminar las constantes aditivas, la función de verosimilitud logarítmica de

muertes D será

$$(3.1) \quad l(\lambda | D, E) \propto \sum_x D_x \ln(\lambda) - E_x \lambda.$$

Asumiendo que todos los grupos de edades comparten el mismo índice de mortalidad, λ , consideramos

$$E_x = \int_x^{x+1} N(t) dt.$$

donde, $N(x)$ representa el tamaño de la población para la edad x , denotada para casos discretos como N_x . El tiempo de exposición se asume como conocido o se estima a partir de N_x .

1.2.1. Estimación por Máxima Verosimilitud.

Tiempo de Exposición Conocido.

Si E_x se asume como conocido, lo denotaremos por $\hat{E}_x$. Si sustituimos $\mu(x)$ por λ , asumimos que tenemos n sorteos independientes (para cada edad) de una distribución de Poisson con un parámetro de tasa que arroja luz sobre la relación funcional previamente oculta entre la tasa λ y la edad x .

Para calcular los estimadores de Máxima Verosimilitud analíticamente, denotemos por S al score de la función de verosimilitud y la definimos por

$$S = \frac{\partial l(\theta | D)}{\partial \theta}.$$

Maximizar la función de verosimilitud implica que para los parámetros óptimos, $\hat{\lambda}$, el sistema de ecuaciones definido es homogéneo

$$S = 0,$$

y la matriz Hessiana

$$H = \frac{\partial^2 l(\hat{\lambda} | D)}{\partial \hat{\lambda}^2}$$

es definida negativa.

Las s funciones score, $i, \dots, s$, en el caso de la verosimilitud para Poisson aparece como

$$(3.2) \quad S_i = \sum_x D_x \frac{\frac{\partial f(\lambda_i, x)}{\partial \lambda_i}}{f(\lambda_i, x)} - \hat{E}_x \frac{\partial f(\lambda_i, x)}{\partial \lambda_i},$$

donde $f(\cdot)$ es alguna función de λ y x .

Sustituyendo $\mu(x) = ae^{bx}$ en la ecuación anterior, se tiene que

$$S = \begin{pmatrix} \sum_x \frac{D_x}{a} - \sum_x \hat{E}_x e^{bx} \\ \sum_x D_x x - \sum_x \hat{E}_x x a e^{bx} \end{pmatrix}.$$

Estableciendo

$$S = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

S_a puede ser resuelto para $\hat{a}$ como

$$(3.3) \quad \hat{a} = \frac{\sum_x D_x}{\sum_x \hat{E}_x e^{bx}}.$$

La ecuación (3.3) se puede utilizar inmediatamente para reducir la dimensión del problema de optimización de dos a uno.

Sustituyendo $\hat{a}$ en S_b se obtiene

$$S_b = \sum_x D_x x - \sum_x \hat{E}_x \frac{\sum_x D_x}{\sum_x \hat{E}_x e^{bx}} x e^{bx},$$

o

$$(3.4) \quad \frac{1}{\sum_x D_x} \sum_x D_x x = \frac{1}{\sum_x \hat{E}_x e^{\hat{b}x}} \sum_x \hat{E}_x e^{\hat{b}x} x.$$

Como $\hat{b}$ es la única incognita en (3.4), se puede resolver numéricamente y seguidamente, $\hat{b}$ se puede sustituir en $\hat{a}$ lo cual reduce el problema de optimización bidimensional a la resolución de una ecuación.

Ejemplo para resolver b .

El método Newton-Raphson es comúnmente usado para resolver sistemas de ecuaciones numéricamente. Este método utiliza un proceso iterativo para aproximar la raíz de una función. Si r denota la raíz y n el número de iteraciones,

$$r_{n+1} = r_n - \frac{f(r_n)}{f'(r_n)}.$$

En el caso del modelo de mortalidad de Gompertz para muestras discretas, $f(b)$ representa a (3.4) y

$$f'(b) = \left(\frac{\sum_x \hat{E}_x e^{bx} x}{\sum_x \hat{E}_x e^{bx}} \right)^2 - \frac{\sum_x \hat{E}_x e^{bx} x^2}{\sum_x \hat{E}_x e^{bx}}$$

Tiempo de Exposición Desconocido.

Si el tiempo de exposición es desconocido, éste se puede estimar a partir de la función de Hazard de manera continua o asumiendo en el intervalo $[x, x+1)$ trozos constantes. En particular, acá emplearemos el método de la verosimilitud binomial.

La función de verosimilitud para el modelo binomial, viene dado por

$$(3.5) \quad L(p|n, y) = \binom{n}{y} p^y (1-p)^{n-y}.$$

Esta función involucra un parámetro p , dada la muestra. La muestra discreta y el estadístico y son conocidos.

Debido a la ausencia de conocimiento del tiempo de exposición, pero con D_x y N_x conocidos, un enfoque alternativo para estimar los parámetros es

$$l_B = (q | D, N) \propto \sum_x D_x \ln(q) + (N_x - D_x) \ln(1 - q),$$

donde q denota la probabilidad de muerte en el intervalo $[x, x+1)$ y $q(x) = 1 - e^{-\int_x^{x+1} \mu(x) dx}$.

Similar al caso anterior, al sustituir $\mu(x)$ por q en la ecuación anterior, asumimos que tenemos n sorteos independientes a partir de una distribución binomial con probabilidad q , que es una función de x .

Observaciones.

Nótese que tomando ω_D y $\omega_{\hat{E}}$ como

$$\omega_D = \frac{D_x}{\sum_x D_x} y \quad \omega_{\hat{E}} = \frac{\hat{E}_x e^{\hat{b}x}}{\sum_x \hat{E}_x e^{\hat{b}x}}$$

y denotando sus respectivos vectores como $W_D, W_{\hat{E}}$ y $\mathbf{X}$, (3.4) se puede reescribir como,

$$\mathbf{W}_D^T \mathbf{X} = \mathbf{W}_{\hat{E}}^T \mathbf{X}.$$

En $b, \hat{b}$ óptimos, los dos promedios ponderados de las edades $\mathbf{X}$; uno medido por el número de muertes W_D y el otro, medido por las exposiciones $W_{\hat{E}}$ son iguales el uno al otro. La ponderación de las exposiciones son, a su vez, también ponderadas. Estas ponderaciones corresponden a la tasa de envejecimiento; cada edad consecutiva tiene $e^{\hat{b}}$ veces mayor importancia que la anterior.

Nota.

La estimación de los parámetros de la distribución de Gompertz se realiza usando el método de máxima verosimilitud puesto que es éste el que mejor se ajusta al análisis de crecimiento exponencial.

CAPITULO 4

Curvas S.

1. Introducción.

La curva logística o curva en forma de S es una función matemática que aparece en diversos modelos de crecimiento de poblaciones, propagación de enfermedades epidémicas, difusión en redes sociales, entre otros. Dicha función constituye un refinamiento del modelo exponencial para la evolución de una cierta cantidad o magnitud a lo largo del tiempo. La curva logística propone que, bajo ciertas circunstancias razonables muchas magnitudes en sistemas ecológicos y sociales evolucionan con el tiempo. Estas curvas reciben este nombre ya que, gráficamente presentan en la primera parte de su dominio un avance lento, seguido de un rápido crecimiento y nuevamente un lento crecimiento hacia el final hasta alcanzar su valor asintótico, obteniendo finalmente que la gráfica general luzca como una S .

Uno de los beneficios de estas curvas S , a partir de ciertas propiedades y la estimación de sus parámetros, los cuales varían dependiendo del objetivo a realizar, aunque siempre dentro de los estándares sobre los cuales están definidos, es la creación de distintos software para el control de proyectos, donde, basadas en la distribución de Gompertz, se modelan estas curvas a fin de poder pronosticar como se verá terminada cierta tarea con el paso del tiempo. Entre muchos, destacan los programas Primavera de Oracle y Microsoft Project, de la empresa que lleva el mismo nombre.

Si f es la función que genera la curva, se cumple que,

$$\lim_{x \rightarrow \infty} f(x) = a.$$

Ahora, redefinamos la función de Gompertz, cuya gráfica describe la curva S , la cual, al ser aplicada para describir el desarrollo desde el inicio de un proyecto hasta su momento de madurez, se define de la siguiente manera:

$$\begin{aligned} f : R^+ &\rightarrow R^+ \\ x &\rightarrow f(x), \end{aligned}$$

donde,

$$y = f(x) = \begin{cases} ae^{bx - \frac{a}{b}(e^{bx} - 1)} & \text{si } 0 \leq x \leq x_m \\ 0 & \text{en otro caso.} \end{cases}$$

En donde y corresponde a los valores obtenidos a partir del trabajo realizado en el tiempo x a medida que éste varía; x_m es el valor correspondiente al momento de madurez, mientras que a y b son constantes definidas así:

- $0 < a \in R$, corresponde al valor asintótico.
- $1 < b \in R$, es un parámetro de ajuste cuando $y \neq 0$ ó $x \neq 0$. Es decir, controla la diferencia entre los valores máximos y mínimos de y para algún x .
- $0 < \frac{a}{b} < 1$, es el índice de madurez; la tasa específica de crecimiento.

Teniendo en cuenta la información anterior se tiene que $0 < y \in R$. Además, tenemos que esta función es inyectiva, es decir, que para $x_1, x_2 \in R^+$ tales que $x_1 \neq x_2$, se cumple que, $f(x_1) \neq f(x_2)$.

La primera derivada de la función, que corresponde a la velocidad de crecimiento es:

$$\frac{dy}{dx} = ae^{bx - \frac{a}{b}(e^{bx} - 1)}(b - ae^{bx}).$$

Como $y = ae^{bx - \frac{a}{b}(e^{bx} - 1)}$ para $0 \leq x \leq x_m$, tenemos que,

$$\frac{dy}{dx} y(b - ae^{bx}).$$

Para encontrar el punto en donde cambia la concavidad de la función se obtiene la segunda derivada, la cual corresponde a la aceleración del crecimiento, esta se iguala a cero y se resuelve la ecuación resultante en términos de x para hallar la abscisa al punto de inflexión, luego se calcula la imagen de la misma para llegar a las coordenadas de dicho punto, esto es:

$$\frac{d^2y}{dx^2} = by(b - ae^{bx}) - aye^{bx}(2b - ae^{bx}).$$

Ahora, igualando esto a cero,

$$\frac{d^2y}{dx^2} = by(b - ae^{bx}) - aye^{bx}(2b - ae^{bx}) = 0.$$

Por otra parte, como $a, b, y, e^{bx} \neq 0$, entonces, se obtiene que

$$b - ae^{bx} = 0.$$

De aquí,

$$\begin{aligned} ae^{bx} &= b; \\ e^{bx} &= \frac{b}{a}; \\ x &= \frac{1}{b} \ln \left(\frac{b}{a} \right). \end{aligned}$$

Por consiguiente, el punto de inflexión es $\left(\frac{1}{b} \ln \left(\frac{a}{b} \right), be^{\frac{a}{b}-1} \right)$. La importancia de este punto de inflexión recae en el hecho de que a partir de él cambia la concavidad de la curva.

Algunos usos de este modelo en la vida laboral son:

- (1) Compra de teléfonos móvil, donde los costos eran inicialmente altos (de manera que la aceptación fue lenta), seguido por un periodo de crecimiento rápido, seguido, a su vez, de una disminución de la absorción cuando la saturación se alcanzó.
- (2) Población en un espacio cerrado, cuando aumenta la tasa de natalidad y luego ocurre un lento crecimiento cuando los límites de recursos se alcancen.
- (3) Modelado del crecimiento de tumores.

A continuación se muestra un ejemplo del uso de este modelo para describir el crecimiento corporal, y de los componentes químicos (proteína, lípido, ceniza y agua) de cerdos en condiciones comerciales de producción. La investigación fue realizada por el Departamento de Ciencias para la Producción Animal, Facultad de Medicina Veterinaria y de Zootecnia, Universidad Nacional de Colombia, Sede Bogotá.

Usando la función de Gompertz definida anteriormente, donde y es el peso del animal a un tiempo t , $a > 0$ es el peso adulto, el parámetro $\frac{a}{b} > 0$ describe el índice de madurez o la tasa específica de crecimiento, y el parámetro $b > 0$ controla la rapidez con que la curva alcanza la cota para un tiempo t .

El punto de inflexión es el punto en el tiempo donde $y = be^{\frac{a}{b}-1}$, este da un $t = \frac{1}{b} \ln(\frac{b}{a})$. La curva S o de Gompertz permite generar asimetría alrededor del punto de inflexión, esto es, que a partir de dicho punto, los extremos de la curva obtenida no deben ser necesariamente iguales, y en este caso, se alcanza este punto antes del 50 por ciento de peso adulto, donde se muestran crecimientos tempranos rápidos, pero unos valores más lentos en la medida que se aproxima a la asíntota, con un largo periodo de crecimiento lineal alrededor del punto de inflexión.

La parametrización del modelo de Gompertz respecto a la variable peso corporal se observa en la tabla a continuación.

Variable	Peso corporal	Agua	Lípidos	Proteína	Ceniza
A	222.7	91.68	72.65	28.14	5.47
B	1.6059	1.4576	2.1027	1.5975	1.5054
C	0.0109	0.0117	0.0114	0.0125	0.00977
Edad Punto de Inflexión, días	147.33	124.58	184.45	127.80	154.08
Peso al punto de inflexión, kg	81.92	33.72	26.72	10.35	2.01
dY/dt Max, kg/d	0.893	0.395	0.305	0.129	0.020
Peso al nacimiento, kg	1.52	1.24	0.020	0.201	0.060
Fase lag o estructuración, d	55.59	39.11	96.73	47.80	51.73
Tasa relativa de crecimiento a 100 kg, kg/d	0.0087	0.0069	0.0146	0.0074	0.0082
Tasa absoluta de crecimiento a 100 kg, kg/d	0.8727	0.3500	0.2947	0.1154	0.0194
50% de la madurez, d	180.96	155.91	216.60	157.12	191.60
98% de la madurez, d	505.31	458.08	526.72	439.96	553.43

En la siguiente gráfica, se presentan las curvas obtenidas a partir de los parámetros calculados para cada una de los componentes químicos y el peso corporal.

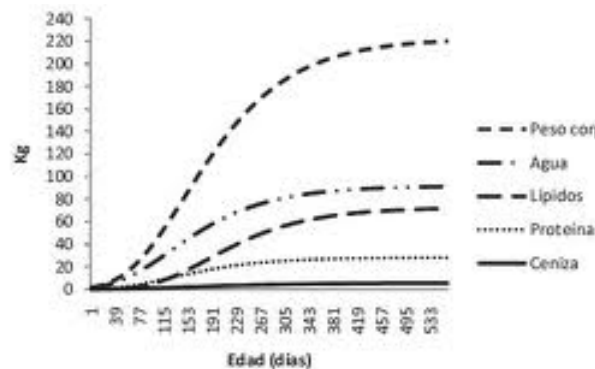


Figura 1. Curvas obtenidas de la parametrización del modelo Gompertz para el peso corporal y sus componentes químicos en cerdos castrados.

2. Curvas S por Sectores.

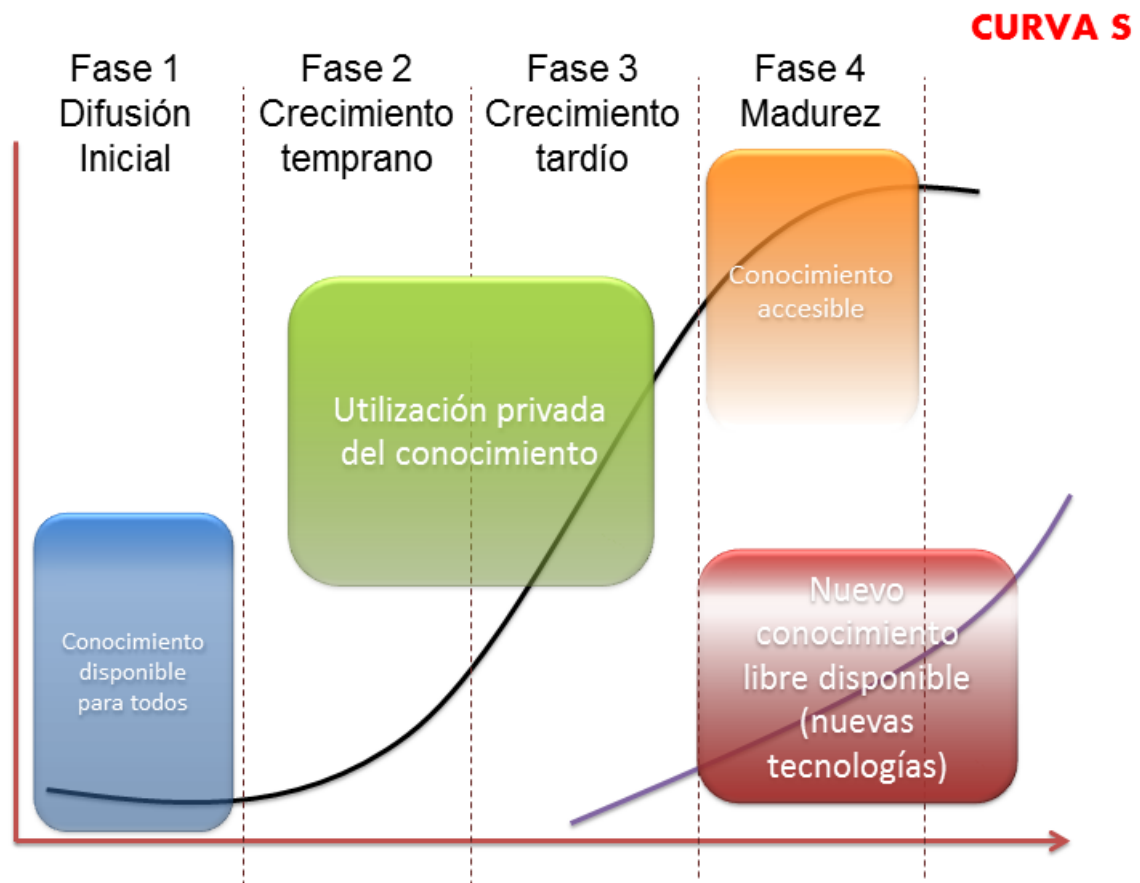
2.1. Curvas S en los negocios.

El ambiente de los negocios ha cambiado. Lo vemos de manera tangible en la forma de nuevos productos y servicios, que aparecen más rápido, se desarrollan más rápido, pero que también desaparecen más rápido. El 80 por ciento de los productos que usamos hoy, no existían hace diez años. Una memoria USB, un teléfono inteligente, Dropbox, Netflix, Waze, son algunos de los ejemplos de los productos o servicios que sustituyeron o están sustituyendo productos obsoletos. La cámara digital sustituyó a la cámara de rollo de película, y ni siquiera la cámara digital se salva, los propios teléfonos la sustituyeron.

Éste fenómeno de difusión lo explica la llamada “Curva S”. Un nuevo conocimiento, una nueva tecnología está a disposición de todos para ser desarrollada en nuevos productos y

servicios. Hay una etapa inicial de prueba donde son muy importantes los primeros usuarios. Cuando se logra un buen desempeño el producto o servicio es aceptado y adoptado por un público más masivo. Por último se llega a una etapa de madurez donde el producto ya no se desarrolla más. Una nueva tecnología terminará por sustituir al producto o servicio en etapa de madurez.

A continuación, se muestra una gráfica señalando lo explicado anteriormente.



La Curva S en los negocios se aplica de la misma manera que en el desarrollo personal, donde al comenzar a investigar en un nuevo dominio de conocimiento tenemos que lidiar con una curva de aprendizaje lineal de lenta penetración, aunque en la medida que seguimos practicando se obtiene mayor entendimiento y con ello nos acercamos a un círculo virtuoso donde finalmente alcanzamos el punto óptimo de competencia, aceleración y confianza.

Un ejemplo de lo comentado es Facebook, iniciativa de la que debemos suponer tenía una oportunidad de mercado de un billón. Llegar a la penetración del mercado del 10 por ciento tardó, en este caso, cuatro años para luego llegar a la masa crítica de un millón de usuarios donde comenzó el hipercrecimiento por el llamado “efecto de la red” y la “viralidad” de la iniciativa.

2.2. Curvas S en el área petrolera.

La función de Gompertz estándar describe el crecimiento de cierta cantidad física dentro de las fronteras de un sistema limitado. Aunque no existe una razón fundamental para creer que esta forma matemática particular, se prefiere en los demás como una manera de describir la dinámica de la exploración de petróleo o de producción, una base heurística deriva del hecho de que la extracción de combustibles fósiles creció considerablemente con el tiempo, lo que produjo mayor obtención de estos recursos finitos.

En resumen, el uso de curvas logísticas para la modelación de producción de crudo, al tiempo que proporciona una descripción satisfactoria, se debe hacer con la conciencia de que los resultados pueden ser menos precisos si una región se encuentra en un punto anterior al punto máximo de producción.

2.3. Curvas S para el control económico de obras.

Después de presupuestar una obra, se requiere controlar la obra, tratando de lograr que se cumpla el desarrollo de la misma, de acuerdo a lo programado, durante el plazo de duración de la obra. Es aquí donde trataremos de aplicar una metodología que nos permita determinar en cada etapa, si estamos en adelanto o en atraso, o si estamos ganando o perdiendo económicamente.

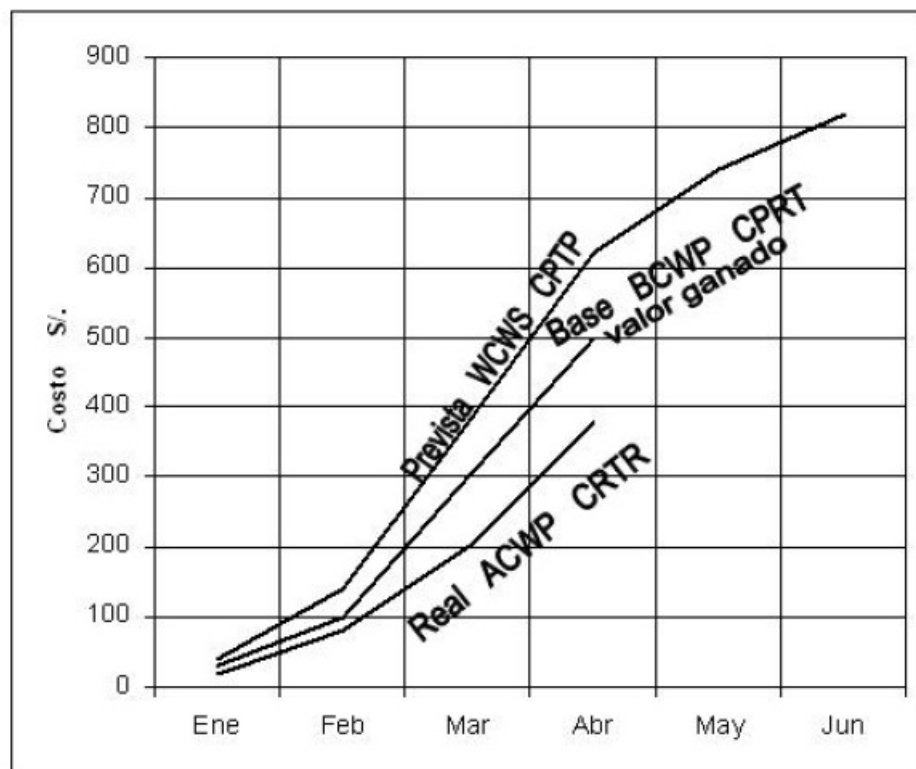
La metodología empleada para el control de estas obras es la de la Curva “S”, tanto en horas-hombre como la valorizada en montos. La curva en horas-hombre es la que se emplea

para programar y reprogramar la obra en el plazo previsto, mientras que la curva económica será aquella en la que se controlan los montos, de los materiales, la mano de obra, los equipos y los sub-contratos.

A medida que se va ejecutando la obra se irá obteniendo la curva real de horas-hombre, la cual indicará, si se está gastando más o menos horas de lo previsto.

Para poder controlar efectivamente la obra se requiere de una tercera curva, que se obtiene al considerar los avances físicos reales referidos a la programación base. Esto es cuanto se hubiese gastado realmente con los avances a una determinada fecha, en un presupuesto base.

Por lo tanto se tienen las 3 curvas, la prevista, la real y la valorizada.



CONCLUSIONES

En este trabajo, se realizó un estudio profundo de la distribución de Gompertz, la cual se origina para el estudio de patrones de vida en poblaciones a pesar que hoy en día es útil para modelar variadas situaciones de la vida. Previo a éste estudio, se debió indagar en el área de las ciencias actuariales, a fin de poder interpretar de manera más eficiente los conceptos básicos y fundamentales de esta ciencia para así, obtener una mejor comprensión de la utilidad de la función generada por la distribución de Gompertz en los estudios de crecimientos de poblacionales y patrones de muerte. Entre los conceptos estudiados, se encuentran la distribución de vida y su respectiva función de supervivencia, puesto que en ocasiones conviene más estudiar la probabilidad de sobrevivir que la de morir. Debido a que los conocimientos proporcionados por las ciencias actuariales son ajenos a los obtenidos en el estudio de las matemáticas, se hizo uso de una notación particular para actuarios.

A partir de aquí, se dio paso al estudio de las llamadas “tablas de vida”, las cuales consisten en la tabulación de los datos obtenidos a partir de la observación de cierta muestra con la finalidad de estimar un patrón de mortalidad que exhiba un grupo de individuos.

Todo lo estudiado hasta el momento sirvió de base para comprender el concepto de fuerza de mortalidad, la cual es la tasa de declive relativa de un grupo de individuos para cierta edad, que se rigen de acuerdo a un patrón obtenido a partir de las tablas de vida.

Con todo lo anterior, se realizó el cálculo de la fuerza de mortalidad para la distribución de Gompertz, obteniendo así, las respectivas funciones de densidad y distribución probabilística de esta distribución.

Luego, se realizó el estudio general de la distribución de Gompertz, en donde se presentaron sus propiedades elementales, como lo son el cálculo de sus momentos a partir de la transformada de Laplace, el estudio del valor esperado o esperanza, varianza, entre otros. También, se estimaron utilizando el método de máxima verosimilitud, los parámetros de esta distribución con la finalidad de maximizar la función de Gompertz y así, obtener una mayor credibilidad de ser ciertos los valores de estos parámetros al momento de ser aplicados a modelos matemáticos. Cabe destacar que, al momento de realizar la estimación de éstos parámetros, se estudiaron dos casos: el caso de una muestra discreta y el caso donde dicha muestra es continua.

Seguidamente, haciendo uso de estas propiedades básicas, se estudian las llamadas “Curvas S”, las cuales son creadas basándose en los distintos modelos matemáticos que resultan de modificaciones del modelo exponencial, con la doble exponencial como característica común, entre los cuales se encuentra el modelo de Gompertz para estudios de poblaciones.

A raíz de esto, se consideró el modelo que define estas curvas S, llamadas así por la forma de su gráfica. Específicamente, se calcularon la primera y segunda derivada de la función de Gompertz, y se obtuvo el punto de inflexión de la curva, el cual es, el punto donde la curva resultante cambia su concavidad. De aquí, se evidencia que éste modelo matemático, que arroja como resultado las curvas S, puede ser utilizado al momento de modelar distintas situaciones de la vida cotidiana, abarcando ciertos sectores laborales, pero siempre teniendo en cuenta que se sigue un patrón, el cuál describe un avance lento al inicio, seguido de un rápido crecimiento y nuevamente un lento crecimiento hacia el final hasta alcanzar su punto de madurez.

En la actualidad, se observa que existen diversos programas informáticos o software para control de proyectos usando las curvas S como referencia.

BIBLIOGRAFÍA

- [1] APOSTOL, T. (1967). One - Variable Calculus, with an Introduction to Linear Algebra. Jon Wiley and Sons, New York, 2nd edition.
- [2] ARRIOJAS, M. (2004). Teoría de las Probabilidades. Caracas.
- [3] ASKEY, R. A.; DAALHUIS, Adri B. Olde (2010), "Generalized hypergeometric function", in Olver, Frank W. J.; Lozier, Daniel M.; Boisvert, Ronald F. et al., NIST Handbook of Mathematical Functions, Cambridge University Press, ISBN 978-0521192255,MR2723248.
- [4] CHUNG, K. (2001). A Course in Probability Theory.
- [5] GUTIERREZ, R. (2005). Difusiones Estocásticas no Homogéneas Lognormales y Gompertz. Proceso de Rayleigh. Aplicaciones.
- [6] IRIBARREN, I. (2006). Intriducción a la Teoría de la Medida. Universidad Central de Venezuela, Caracas.
- [7] ISTONI, S. (2011). Estudio Comparativo de Procesos de Difusión Asociados con Curvas Gompertz.
- [8] LENART, A. (2012). The Gompertz distribution and Maximum Likelihood Estimation of its parameters - a revision. Max Planck Institute for Demographic Research, Germany.
- [9] MILGRAM, M. (1985). The generalized integro exponential function. Mathematics of computation.

- [10] MUKHOPADHYAY, N. Probability and Statistical Interference. University of Connecticut.
- [11] SLUD, E. (2001). Actuarial Mathematics and Life - Table Statistics. Mathematics Department, University of Maryland, College Park.