

CAPÍTULO I

CONCEPTOS BÁSICOS DE PROBABILIDAD Y ESTADÍSTICA

El campo de la estadística tiene que ver con la recopilación, presentación, análisis y uso de datos para tomar decisiones y resolver problemas. Montgomery (1996). Es por esta razón que es tan importante la estadística en la vida cotidiana, ya que las técnicas estadísticas se utilizan en casi todos los aspectos. Se diseñan encuestas para recabar información previa a un día de elecciones, se seleccionan al azar consumidores para obtener información con el fin de predecir la preferencia de un producto, el ingeniero muestrea las características de calidad de un producto, el economista considera varios índices de la situación económica durante cierto periodo y utiliza la información para predecir la situación económica futura. Es por eso que las técnicas estadísticas desempeñan una función importante en el logro del objetivo de cada uno de estos problemas prácticos. Es por eso que en este Capítulo se hace una recopilación de los temas más importantes dentro de la probabilidad y estadística, para dar un panorama más amplio del tema en esta tesis.

1.1 Probabilidad

La probabilidad se desarrolló en un principio en relación con los juegos de azar y lleva una relación entre las posibilidades y las probabilidades. Este concepto clásico solo se aplica cuando los eventos son igualmente probables. Freund (1994).

Es muy útil saber y cuantificar la posibilidad de que se presente algún resultado de un experimento aleatorio. Esta cuantificación se hace asignando un número entre el intervalo $[0,1]$. Un hecho tan sencillo como puede ser conocer la posibilidad de lluvia de un determinado día.

1.1.1 Evento y Espacios Muestrales

Un evento se encuentra asociado con el espacio de muestreo de un experimento Hines (1986). Este espacio muestral es el conjunto de los posibles resultados de un experimento aleatorio y se denota con la letra S . Generalmente, un espacio muestral se clasifica de acuerdo al número de elementos con que cuenta. Éstos pueden ser discretos cuando contienen un número finito o numerablemente infinito de elementos, mientras que los continuos contienen un número infinito no contable de elementos.

Existen dos tipos de experimentos que son: determinísticos, los cuales se efectúan bajo condiciones que determinan el resultado del mismo y los no determinísticos o aleatorios, llamados así por que no se puede predecir su resultado atendiendo las condiciones bajo las cuales se lleva a cabo y además proporcionan diferentes resultados aún cuando se repita siempre de la misma manera. En esta tesis se describirán los segundos.

1.1.2 Axiomas de Probabilidad

Estos axiomas deben de satisfacer las posibilidades de cualquier experimento aleatorio. Los axiomas y sus consecuencias restringen las asignaciones de probabilidad de una manera que permite interpretar éstas como frecuencias relativas sin inconsistencias. Montgomery (1996). Los axiomas de probabilidad son:

1. $P(S) = 1$
2. $0 \leq P(E) \leq 1$
3. $P(E_1 \cup E_2) = P(E_1) + P(E_2)$, cuando ambos eventos son mutuamente excluyentes.

Donde: S es el espacio muestral.

E es un evento.

$P(E)$ la probabilidad del evento E .

A partir de estos axiomas se implican algunos otros resultados como son:

- $P(\emptyset) = 0$, es decir, la probabilidad de un conjunto vacío es cero.
- $P(E') = 1 - P(E)$, de igual forma, la probabilidad del complemento de E (E') es 1 menos la probabilidad del evento E.

1.1.3 Regla de la Adición

Esta regla se aplica a dos eventos mutuamente excluyentes, pero se generaliza a que se aplique a más de dos eventos mutuamente excluyentes, es decir que ningún elemento de un conjunto pertenezca a otro.

Si k eventos son mutuamente excluyentes, la probabilidad de que uno de éstos ocurra es igual a la suma de las probabilidades de cada uno de ellos.

$$P(A_1 \cup A_2 \cup \dots \cup A_k) = P(A_1) + P(A_2) + \dots + P(A_k)$$

Esta regla también se puede generalizar de la siguiente manera según Freund (1994).

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

1.1.4 Regla de la Multiplicación

La probabilidad de que dos eventos independientes ocurran es simplemente el producto de las probabilidades respectivas.

$$P(A \cap B) = P(A) * P(B)$$

En ocasiones esta regla es usada para verificar si dos eventos son independientes. Esta regla será de gran utilidad, ya que posteriormente se usará para calcular la Función de Máxima Verosimilitud de la distribución Hockey Stick.

Dos eventos son independientes cuando la ocurrencia del evento A no se afecta por la ocurrencia del evento B . Cuando esto sucede, se puede afirmar que el evento A es independiente del evento B , y esta relación se puede expresar por la siguiente definición:

$$P(A/B) = P(A) \quad \text{o} \quad P(B/A) = P(B)$$

1.2 Función de Densidad

En el estudio de las variables aleatorias, generalmente lo más interesante es la probabilidad que toman los diversos valores dentro de su amplitud, los cuales son descritos por la Función de Densidad.

Las variables aleatorias se clasifican de dos formas: variables aleatorias discretas las cuales toman un número finito o contablemente infinito, tantos valores como números enteros existan y variables aleatorias continuas, que son las que se usan cuando se manejan cantidades de medida en una escala continua, por ejemplo: el tiempo, el peso, la distancia, etc.

Una distribución de probabilidad es una correspondencia que asigna probabilidades a una variable aleatoria. Freund (1994). Siempre que es posible, se trata de expresar las distribuciones de probabilidad por medio de fórmulas que permitan calcular las probabilidades asociadas con los diversos valores de una variable aleatoria.

Los valores de una distribución de probabilidad deben ser números que se encuentren en un intervalo de $[0,1]$. Además, la suma de todos los valores de una distribución de probabilidad debe ser equivalente a 1.

1.3 Función de Distribución Acumulativa

La Función de Distribución Acumulativa sirve para calcular la probabilidad dentro de un intervalo y se define como:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(x) dx$$

La Función de Distribución cuenta con las siguientes propiedades:

1. $\lim_{x \rightarrow -\infty} F_X(x) = 0$
2. $\lim_{x \rightarrow \infty} F_X(x) = 1$
3. $P(a < x < b) = \int_a^b f_X(x) dx = F(b) - F(a)$, donde $f_X(x)$ es la Función de Densidad correspondiente a la Función de Distribución $F_X(x)$.
4. $f_X(x) = \frac{dF_X(x)}{dx}$, donde $f_X(x)$ es la Función de Densidad correspondiente a la Función de Distribución $F_X(x)$.

1.4 Función de Máxima Verosimilitud

La Función de Máxima Verosimilitud normalmente se ve representada por la letra L. Uno de los métodos para obtener un estimador puntual de un parámetro es por medio del Estimador de Máxima Verosimilitud. Un estimador es una función numérica de un dato.

Existen muchos caminos para especificar la forma de un estimador para un parámetro particular de una distribución dada. Law y Kelton (2000). Un estimador de un parámetro de la población consiste en un solo valor de un estadístico y se le conoce como estimador puntual del parámetro.

Dada una muestra aleatoria de una $f(x)$ con distribución de probabilidad $f(x, \theta)$, donde θ es un parámetro desconocido. Sean x_1, x_2, \dots, x_n los valores observados en una muestra aleatoria de tamaño n . La Función de Máxima Verosimilitud de la muestra es:

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta)$$

(1.1)

Donde:

$f(x_i; \theta)$ es la Función de Densidad.

θ es un parámetro o un conjunto de parámetros para la distribución dada.

Esto es, $L(\theta)$ es la verosimilitud de la muestra que es la "Probabilidad" de observar la muestra. Por lo tanto el Estimador de Máxima Verosimilitud es un estimador que maximiza la probabilidad de ocurrencia de los valores muestrales. Ver Montgomery (1996).

Para el caso de una variable aleatoria continua, la cual se describirá en esta tesis. Se tiene que, la Función de Máxima Verosimilitud esta dada por:

$$L(\theta) = f_{\theta}(X_1) f_{\theta}(X_2) \dots f_{\theta}(X_n)$$

Se dice que $\hat{\theta}$ de θ es definido como el valor de θ que maximice $L(\theta)$ sobre todos los valores permisibles de θ . Law y Kelton (2000).

Por lo tanto la función que maximiza la Función de Máxima Verosimilitud esta dada por:

$$L(\hat{\theta}) = \text{Max } L(\theta) \quad \forall \theta$$

Dado que el logaritmo natural de una función es una función continua creciente que se maximiza en el mismo punto que la función original, es posible utilizar la siguiente igualdad para facilidad de cálculo:

$$l = \ln L$$

(1.2)

Sustituyendo en la ecuación 1.2 la ecuación de la Función de Máxima Verosimilitud, la ecuación 1.1, se obtiene la siguiente igualdad:

$$l = \ln \prod_{i=1}^n f(x_i)$$

Una de las formas para calcular los estimadores de algunas distribuciones es, una vez obtenida la expresión l , se deriva con respecto a cada uno de los estimadores correspondientes a cada distribución. Posteriormente igualando a cero las derivadas y por ultimo despejando el estimador deseado.

La probabilidad de un punto x_i en ese mismo punto es cero y formalmente se obtiene integrando la Función de Densidad del punto x_i a ese mismo punto y como se observa en la siguiente igualdad:

$$P [X_1 = x_1] = 0 = \int_{x_1}^{x_1} f(x) dx$$

En la Figura 1.1 se aprecia de manera gráfica lo anteriormente dicho ya que se esta integrando un diferencial que es una parte tan pequeña y que al ser integrado el mismo punto x_1 esta probabilidad es cero.

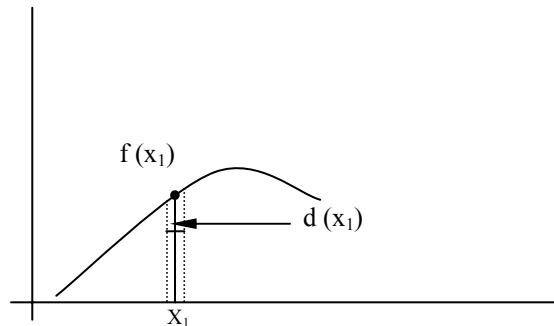


Figura 1.1 Diferencial de un punto x_1 .

Sin embargo, cabe mencionar que se ha buscado una interpretación a esta probabilidad y a continuación se muestra la manera de como se hace Bladt (1993):

$$P [X_1 \leq x_1 \leq X_1 + dx] = f(x_1) dx$$

Al encontrar esta interpretación de la $f(x)$ se encuentran dos ventajas:

1. La definición de probabilidad es sumamente clara de la siguiente forma:

$$P [a < x_1 < b] = \int_a^b f(x_1) dx$$

2. Por otro lado, permite definir claramente el principio de verosimilitud, por que se puede interpretar claramente bajo cualquier distribución, ya sea continua o discreta, que la función de verosimilitud es la "probabilidad" de observar la muestra

y los Estimadores de Máxima Verosimilitud son aquellos valores de los parámetros que maximicen dicha probabilidad.

La técnica del Estimador de Máxima Verosimilitud es por lo tanto equivalente a maximizar una ecuación de algunas variables. En general, los cálculos que más se acercan es obtener la derivada parcial con respecto a cada uno de los parámetros desconocidos, esto produce un sistema de ecuaciones que tienen los estimadores de máxima verosimilitud como solución. Tobias (1986). Si una distribución tiene más de un parámetro, se puede definir el Estimador de Máxima Verosimilitud para cada uno de los parámetros en un camino natural.

Ejemplo 1.1 Encuentre el Estimador de Máxima Verosimilitud de λ de la distribución exponencial. Si se tiene una muestra aleatoria X_1, X_2, \dots, X_n , de una distribución exponencial con una tasa de llegada de λ .

Solución:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & d.o.m. \end{cases}$$

Reemplazando la Función de Densidad de la distribución Exponencial en la ecuación 1.1 la cual es la Función de Máxima Verosimilitud, es posible escribir la siguiente ecuación:

$$L(\lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i}$$

Aplicando la ecuación 1.2 para facilitar los cálculos se tiene lo siguiente:

$$l = \ln L(\lambda)$$

$$\begin{aligned}
&= \ln \left[\prod_{i=1}^n \lambda e^{-\lambda x_i} \right] \\
&= \sum_{i=1}^n \ln \left[\lambda e^{-\lambda x_i} \right] \\
&= \sum_{i=1}^n [\ln \lambda - \lambda x_i] \\
&= n \ln \lambda - \lambda \sum_{i=1}^n x_i
\end{aligned}$$

Como siguiente paso se procede a derivar la ecuación anterior con respecto a λ ya que es el parámetro de la distribución Exponencial y es el cual se quiere estimar:

$$\frac{dl}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i$$

Luego la derivada obtenida anteriormente es igualada a cero para maximizar la función y se procede como último paso a despejar λ el cual es el parámetro que se desea estimar.

$$\frac{n}{\lambda} - \sum_{i=1}^n x_i = 0$$

$$\lambda = \frac{n}{\sum_{i=1}^n x_i}$$

Finalizando se usa una muestra aleatoria para estimar el parámetro de la población a la cual se hacer referencia con la muestra que se obtuvo.

Ejemplo 1.2. La distribución Gama tiene dos parámetros (α y β), y la Función de Máxima Verosimilitud es definida como:

$$L(\alpha, \beta) = \frac{\beta^{-n\alpha} \left(\prod_{i=1}^n X_i\right)^{\alpha-1} \exp\left[-\frac{1}{\beta} \sum_{i=1}^n X_i\right]}{[\Gamma(\alpha)]^n}$$

Los Estimadores de Máxima Verosimilitud $\hat{\alpha}$ y $\hat{\beta}$ de los valores desconocidos de α y β están definidos para ser los valores de α y β que juntamente maximizan $L(\alpha, \beta)$. Para encontrar $\hat{\alpha}$ y $\hat{\beta}$ se procede obteniendo $l(\alpha, \beta) = \ln L(\alpha, \beta)$ y posteriormente se resuelve un sistema de ecuaciones $dl/d\alpha = 0$ y $dl/d\beta = 0$ simultáneamente para α y β .

1.5 La Distribución Exponencial

La distribución Exponencial obtiene su nombre de la función exponencial que aparece en la Función de Densidad. Esta distribución tiene como parámetro λ el cual es de escala y sus unidades generalmente son número de ocurrencias entre unidad de tiempo. La Figura 1.2 muestra la gráfica de la distribución Exponencial para distintos valores de λ .

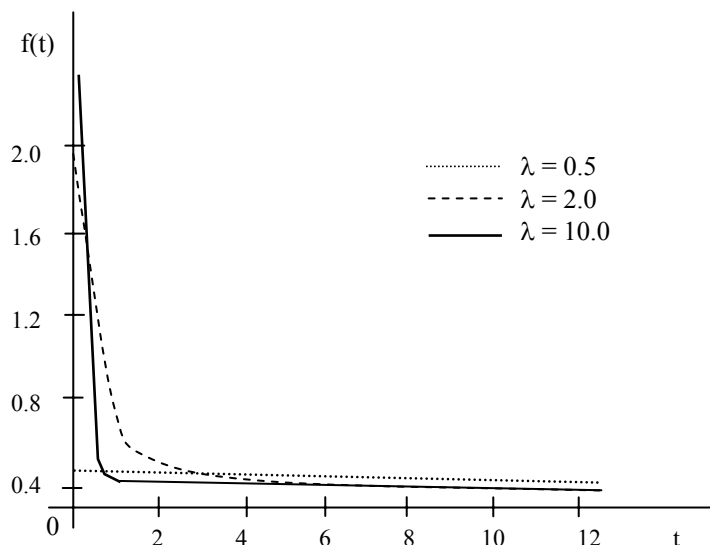


Figura 1.2 Función de densidad de una variable aleatoria exponencial para diferentes valores de λ .

La distribución Exponencial es ampliamente usada en el campo de la Ingeniería de confiabilidad como un modelo del tiempo de falla de un componente o sistema. En estas aplicaciones, el parámetro λ es llamado tasa de fallas del sistema y la media de la distribución $1/\lambda$ el cual es llamado tiempo medio de falla. Montgomery (2001).

La distribución Exponencial (λ) es un caso especial de la distribución Weibull ya que $t \sim Weibull(1,c) = Exp\left(\frac{1}{c}\right)$ y de la distribución Gamma por que $t \sim \Gamma(1,\beta) = Exp\left(\frac{1}{\beta}\right)$ con parámetros de forma y de escala respectivamente de $m = 1$ y $\alpha = 1$ y λ en ambos casos. Law y Kelton, (2000). Y se escribe de la siguiente manera:

$$t \sim Exp(\lambda)$$

Supóngase que X es una variable aleatoria continua con Función de Densidad $f_x(x), -\infty < x < \infty$. La media o el valor esperado de una variable aleatoria continúa esta denotada por μ_x o por $E(X)$ y es:

$$\mu_x = E(X) = \int_{-\infty}^{\infty} x f_x(x) dx$$

Aplicando la ecuación anterior y sustituyendo la Función de Densidad de la distribución Exponencial se tiene que la media es:

$$\mu = \frac{1}{\lambda}$$

La varianza de una variable aleatoria se calcula ponderando el cuadrado de cada desviación con respecto a la media, con la probabilidad asociada con la desviación. La varianza de una variable aleatoria esta denotada por σ_x^2 o $V(X)$ y es:

$$\sigma_x^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu_x)^2 f_x(x) dx$$

Aplicando la ecuación anterior y sustituyendo la Función de Densidad de la distribución Exponencial se tiene que la varianza es:

$$\sigma^2 = \frac{1}{\lambda^2}$$

La Función de Densidad es la que trata de expresar las distribuciones de probabilidad por medio de fórmulas que estas a su vez permiten calcular las probabilidades asociadas con las variables aleatorias. La Función de Densidad para la distribución Exponencial esta dada por:

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & d.o.m. \end{cases}$$

Donde: λ es el coeficiente de intensidad o número de fallas esperadas por unidad de tiempo.

t es la variable aleatoria la cual mide el tiempo de falla y debe ser mayor o igual a cero.

La Función de Distribución Acumulativa por lo tanto se calcula integrando la Función de Densidad de la distribución de la siguiente manera:

$$\begin{aligned} F(t) &= P\{x \leq t\} = 0, \quad t \leq 0 \\ &= \int_0^t \lambda e^{-\lambda t} dt \\ &= 1 - e^{-\lambda t} \quad t \geq 0 \end{aligned}$$

La Figura 1.3 muestra la Función de Distribución Acumulativa de la distribución Exponencial.

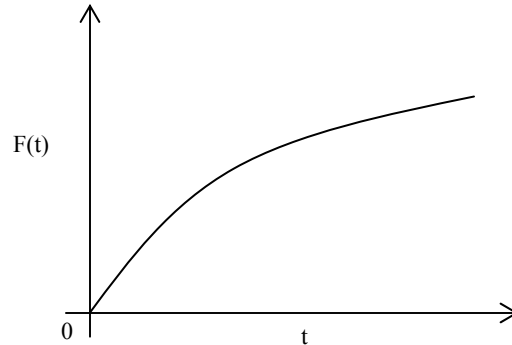


Figura 1.3 Gráfica de la función de distribución exponencial acumulativa.

Dada la Función de Distribución Acumulativa y con la igualdad $R(t) = 1 - F(t)$ que se verá con más detalle en el Capítulo II, es fácil decir que la función de confiabilidad esta dada por:

$$R(t) = \begin{cases} e^{-\lambda t}, & t > 0 \\ 1, & t \leq 0 \end{cases}$$

Para esta distribución la tasa de fallas esta definida como el cociente de la Función de Densidad y la Función de Confiabilidad. La tasa de fallas tiene como unidades el número de fallas por unidad de tiempo. No es una probabilidad y puede tener valores mayores que uno, sin embargo siempre será positiva. Se dará más información en el Capítulo II. La tasa de fallas para la distribución Exponencial es:

$$\begin{aligned} h(t) &= \frac{f(t)}{R(t)} \\ &= \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} \\ &= \lambda \end{aligned}$$

Existe una propiedad característica de la distribución exponencial; es la única distribución con una tasa de fallas constante dado que la tasa de fallas de la distribución se reduce solo a λ para todo el tiempo, y la Figura 1.4 lo muestra.



Figura 1.4 Tasa de fallas $h(t)$ de la distribución exponencial.

La distribución Exponencial tiene una propiedad muy importante que es la de la pérdida de memoria. Esta propiedad consiste en que un componente o sistema no recuerda que tanto ha estado operando y entonces la probabilidad condicional de que falle en la siguiente hora es la misma que si fuera nuevo.

1.6 La Distribución Weibull

La familia Weibull es una de las distribuciones de vida que provee buenos modelos en muchos casos empíricos. Esta distribución puede ser usada como un modelo de fallas causadas por procesos de degradación como fatiga, corrosión, difusión y abrasión mecánica como fallas de cojinetes y material de potencia. La Figura 1.5 muestra una gráfica de la Función de Densidad de la distribución para diferentes valores del parámetro de forma.

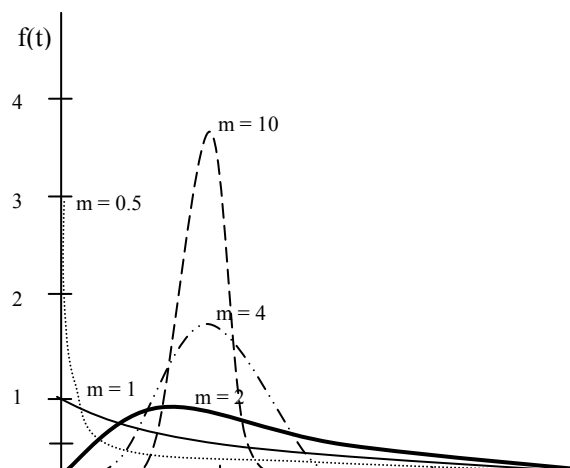


Figura 1.5 Función de densidad de la distribución Weibull para distintos valores de m .

La distribución Weibull cuenta con dos parámetros, los cuales son: m el cual que es conocido como el parámetro de forma y c como un parámetro de escala, también llamado característica de vida. Ambos deben ser mayores que cero y esta distribución es una distribución de vida definida solo por tiempos positivos t . Tobias (1986). Y se escribe de la siguiente forma:

$$t \sim Weibull(m, c)$$

Al igual que en la distribución Exponencial X es una variable aleatoria continua con Función de Densidad $f_x(x), -\infty < x < \infty$. La media o el valor esperado de una variable aleatoria continúa se denota por μ_x o por $E(X)$ y es:

$$\mu_x = E(X) = \int_{-\infty}^{\infty} xf_x(x)dx$$

Si se aplica la ecuación anterior y se sustituye la Función de Densidad de la distribución Weibull se tiene que la media es:

$$\mu = c\Gamma\left(1 + \frac{1}{m}\right)$$

De igual forma como se mencionó en la distribución Exponencial, la varianza de una variable aleatoria se calcula ponderando el cuadrado de cada desviación con respecto a la media, con la probabilidad asociada con la desviación. La varianza de una variable aleatoria se denota por σ_x^2 o $V(X)$ y es:

$$\sigma_x^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu_x)^2 f_x(x) dx$$

Aplicando la ecuación anterior y sustituyendo la Función de Densidad de la distribución Weibull se tiene que la varianza es:

$$\sigma^2 = c^2 \Gamma\left(1 + \frac{2}{m}\right) - \left[c \Gamma\left(1 + \frac{1}{m}\right) \right]^2$$

La Función de Densidad es la que trata de expresar las distribuciones de probabilidad por medio de fórmulas que éstas a su vez permiten calcular las probabilidades asociadas con las variables aleatorias. La Función de Densidad para distribución Weibull esta dada por:

$$f(t) = \begin{cases} \frac{mt^{m-1}}{c^m} e^{-\left(\frac{t}{c}\right)^m} & t > 0 \\ 0 & d.o.m. \end{cases}$$

Donde: m es el parámetro de forma.

c es el parámetro de escala.

t es la variable aleatoria la cual mide el tiempo de falla y debe ser mayor que uno.

La Función de Distribución Acumulativa de la distribución Weibull por lo tanto se puede calcular integrando la función de densidad obteniendo la siguiente ecuación:

$$F(t) = \begin{cases} 1 - e^{-\left(\frac{t}{c}\right)^m} & t > 0 \\ 0 & d.o.m. \end{cases}$$

Dada la Función de Distribución Acumulativa y conociendo la igualdad $R(t) = 1 - F(t)$ que se verá con más detalle en el Capítulo II, es fácil obtener la función de confiabilidad $R(t)$:

$$R(t) = \begin{cases} e^{-\left(\frac{t}{c}\right)^m}, & t > 0 \\ 0 & , d.o.m. \end{cases}$$

La tasa de fallas esta definida como el resultado de la Función de Densidad entre la Función de Confiabilidad. La tasa de fallas tiene como unidades el número de fallas por unidad de tiempo. No es una probabilidad y puede tener valores mayores que uno, sin embargo siempre será positiva. Se dará más información sobre esta en el Capítulo II. La tasa de fallas para la distribución Weibull es:

$$\begin{aligned} h(t) &= \frac{f(t)}{R(t)} \\ &= \frac{\frac{mt^{m-1}}{c^m} e^{-\left(\frac{t}{c}\right)^m}}{e^{-\left(\frac{t}{c}\right)^m}} \\ &= \frac{mt^{m-1}}{c^m} \end{aligned}$$

La distribución Weibull se emplea a menudo para modelar el tiempo hasta presentarse una falla en muchos sistemas físicos diferentes, ya que los parámetros de esta distribución proporcionan mucha flexibilidad para modelar sistemas en los que el número de fallas aumenta con el tiempo, que disminuyen con el tiempo o que permanece constante. La Figura 1.6 muestra una gráfica de la tasa de fallas de la distribución Weibull para distintos valores del parámetro m .

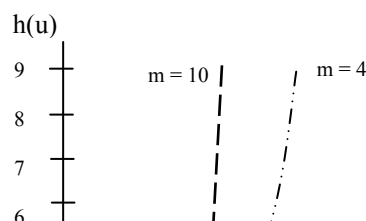


Figura 1.6 Tasa de fallas de la distribución Weibull para distintos valores de m .

1.7 Intervalos de confianza

El intervalo de confianza es un método que utiliza las mediciones de la muestra para calcular dos números que forman los extremos del intervalo. El intervalo de confianza en su caso ideal debe de tener dos propiedades: que contenga al parámetro objetivo y además que sea lo más estrecho posible. El intervalo se encuentra en función de las mediciones de la muestra y es por esa razón que varía de manera aleatoria en uno o en ambos de sus extremos. Mendenhall (1992).

Una estimación por intervalos de un parámetro desconocido θ es un intervalo de la forma $l \leq \theta \leq u$, donde los puntos extremos l y u dependen del valor numérico del estadístico $\hat{\theta}$ para una muestra en particular, y de la distribución de muestreo de $\hat{\theta}$. Puesto que muestras diferentes producen valores distintos de $\hat{\theta}$ y, en consecuencia, valores diferentes de los puntos extremos l y u , estos puntos son valores de variables aleatorias, por ejemplo L y U , respectivamente. De la distribución de muestreo de $\hat{\theta}$ es posible

determinar los valores aleatorios de L y U tales que la siguiente proposición de probabilidad es verdadera:

$$P(L \leq \theta \leq U) = 1 - \alpha$$

Donde: $0 < \alpha < 1$ significa que se tiene una probabilidad de $1 - \alpha$ de seleccionar una muestra que produzca un intervalo que contiene el valor verdadero de θ . Ver Holland y Sielken (1993).

Si x_1, x_2, \dots, x_n es una realización de la muestra aleatoria X_1, X_2, \dots, X_n y se calcula l como una realización de L y u una realización de U , entonces, el intervalo resultante $l \leq \theta \leq u$ se conoce como intervalo de confianza del $100(1 - \alpha)$ por ciento para el parámetro desconocido θ . Las cantidades l y u reciben el nombre de límites de confianza inferior y superior, respectivamente, y $1 - \alpha$ es el coeficiente de confianza. La interpretación de un intervalo de confianza es que, si se recopila un número infinito de muestras aleatorias y se calcula un intervalo de confianza del $100(1 - \alpha)$ por ciento para θ , para cada una de las muestras, entonces el $100(1 - \alpha)$ por ciento de esos intervalos contienen el valor verdadero de θ .

1.8 Método Bootstrap

Bradley Efron (1979) desarrolló el método Bootstrap. Desde entonces, se ha convertido rápidamente en una popular y poderosa herramienta estadística usada para problemas difíciles en el análisis estadístico. Este método es computacionalmente intenso, sin embargo, las modernas computadoras son más que suficientes para los requerimientos computacionales requeridos para este método.

El método Bootstrap es una forma de hacer inferencia de los datos originales mediante el uso de una gran cantidad de métodos, los cuales son llamados procedimientos de remuestreo. El procedimiento Bootstrap se puede definir como B muestras Bootstrap generadas de un conjunto de datos originales, cada muestra Bootstrap tiene n elementos generados con muestreo con reemplazo n veces. Las replicas Bootstrap son obtenidas para calcular el valor del estimador de la replica. Ver Devore (1987). Las estadísticas son con frecuencia las salidas del análisis de los datos. La estimación Bootstrap del error estándar no requiere de cálculos teóricos. Un pequeño número de réplicas $B = 25$ da solamente resultados informativos, replicas de $B = 50$ es a menudo suficiente para dar como resultado un buen estimador del error estándar, sin embargo se requieren una B de mayor tamaño para los intervalos de confianza Bootstrap, una $B = 1000$ es necesaria para un intervalo de confianza estable. Ver Thomas (2002).

Existen varios métodos de muestreo para hacer inferencia a los parámetros de una población y en esta tesis se usará para este propósito el Bootstrap paramétrico. El método de Bootstrap paramétrico se usa cuando se es conocida la forma funcional de la población bajo estudio, es decir, su distribución estadística, pero al menos uno de los parámetros definidos es desconocido. Este método se aplica a las siguientes situaciones:

- Cuando es conocida la distribución de la población la cual se esta estudiando.
- Cuando se tiene una muestra disponible de la población.
- Cuando los parámetros desconocidos para completar la distribución, se pueden estimar de la sola muestra actual.

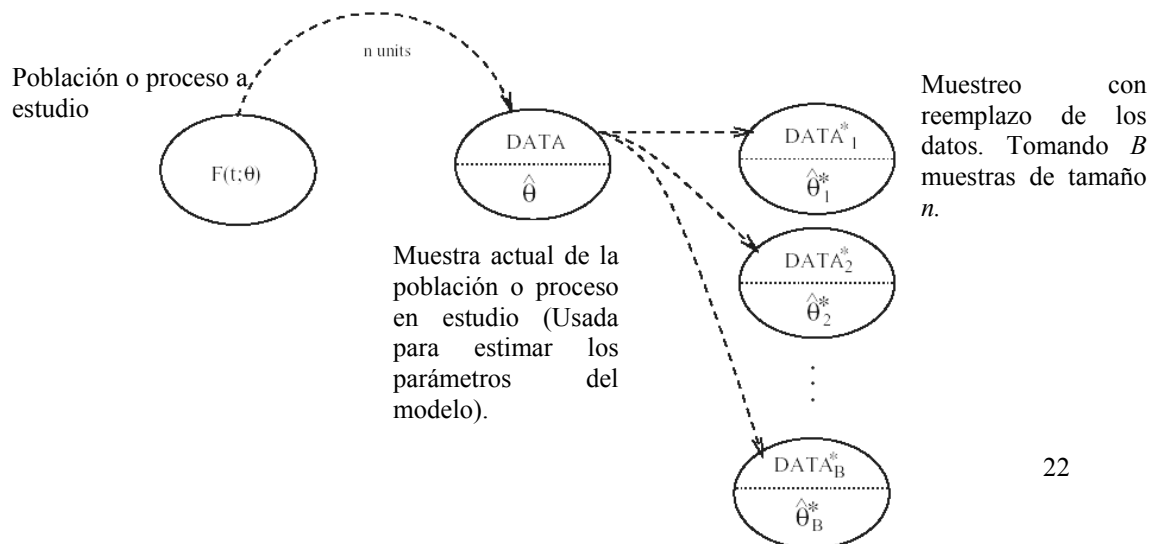


Figura 1.7 Diagrama del método de Bootstrap paramétrico.

Suponga que se tiene una muestra X_1, X_2, \dots, X_n , de una distribución, con una distribución de densidad $f(x/\theta)$, donde θ es un vector de parámetros. Se puede estimar θ con $\hat{\theta}$, que es el Estimador de Máxima Verosimilitud del muestreo.

$$X_1^*, X_2^*, \dots, X_n^* \sim f(X/\hat{\theta})$$

Si se toman B muestras, se puede estimar la varianza de θ^* usando:

$$Var^*_B(\hat{\theta}) = \frac{1}{B-1} \sum_{i=1}^B \left(\hat{\theta}_i^* - \bar{\hat{\theta}}^* \right)^2$$

Donde: $\hat{\theta}_i^*$ es el estimador de la observación i -ésima y $\bar{\hat{\theta}}^*$ es el promedio del estimador de la muestra tomada. Ver Cassella y Berger (2002).