CAPÍTULO 1

**1. DISTRIBUCIONES POBLACIONALES**

**1.1 Introducción**

En las secciones de este capítulo presentamos las principales definiciones a utilizar en el desarrollo del presente trabajo, teniendo así en la sección 1.2 la definición de parámetro poblacional, en la sección 1.3 la definición de estimador con las características principales deseadas los estimadores, en la sección 1.4 los momentos alrededor de la media y el origen poblacionales y muestrales y los coeficientes que se obtienen a partir de ellos como son el coeficiente de simetría, kurtosis, etc., en esta sección ilustramos con ejemplos lo expuesto anteriormente. En la sección 1.5 mostramos el método de la transformada inversa para la generación de números aleatorios, en la sección 1.6 se muestran e ilustran los métodos de estimación de máxima verosimilitud y de los momentos, en la sección 1.7 revisamos la definición de convergencia en distribución y en la sección 1.8 se hace referencia a las distribuciones discretas y continuas de las cuales presentamos su función de densidad y principales momentos poblacionales en el Anexo 1 y en el Anexo 2, respectivamente.

**1.2 Parámetro poblacional**

Sea X una variable aleatoria o población, discreta o continua; una constante , característica de esta población es denominada parámetro poblacional.

**1.3 Estimador**

Sea X1, X2,…, Xn una muestra aleatoria tomada de una población X, sea  un parámetro de dicha población; un estimador  de , es una función  definida en términos de los valores de la muestra y que en su definición no incluye al parámetro .



Un parámetro poblacional de  puede tener más de un estimador 

**1.3.1 Características de los estimadores**

Existen ciertas características de los estimadores, que son deseables al momento de realizar inferencias estadísticas acerca de los parámetros poblacionales como son las siguientes:

**Insesgadez**

Sea  un estimador de un parámetro poblacional .

Se dice que  es un estimador insesgado de si , caso contrario se dice que es sesgado.

**Sesgo de estimación:** El sesgo B de un estimador  de un parámetro poblacional  está dado por:

****

Si el sesgo de estimación es cero el estimador se denomina insesgado.

**Eficiencia relativa**

Sea  estimadores de un mismo parámetro poblacional  y  las varianzas de los estimadores respectivamente.

Entonces se define:



Si la eficiencia relativa es mayor que 1 entonces  es más eficiente que .

**Consistencia**

Sea  un estimador de un parámetro poblacional .

Se dice que  es consistente si:





Se puede probar que, si  es un estimador insesgado de un parámetro poblacional  tenemos que es consistente si se cumple:



**1.4 Momentos alrededor del origen y alrededor de la media.**

Para poblaciones podemos obtener los momentos poblacionales alrededor de la media y del origen que también constituyen parámetros poblacionales de la variable aleatoria, análogamente se pueden obtener los momentos muestrales alrededor de la media y del origen que constituyen estimadores de los parámetros poblacionales respectivos, en esta sección presentamos las definiciones mencionadas.

**1.4.1 Momento k-ésimo alrededor del origen:**

Sea X una variable aleatoria continua con densidad f, se define el k-ésimo momento poblacional alrededor del origen como:

****

Sea X una variable aleatoria discreta con distribución f, se define el k-ésimo momento poblacional alrededor del origen como:



Para k=1,  es la media poblacional.

Para estimar los momentos poblacionales alrededor del origen, se suelen utilizar los siguientes estimadores:

Sea X1, X2,…,Xn; una muestra aleatoria, entonces los momentos muestrales alrededor del origen se definen como:



* + 1. **Momento k-ésimo alrededor de la media**

Sea X una variable aleatoria continua, se define el k-ésimo momento poblacional alrededor de la media como:

****

Sea X una variable aleatoria discreta, se define el k-ésimo momento poblacional alrededor de la media como:



Para k=2, tenemos la varianza poblacional.

Para estimar los momentos poblacionales alrededor de la media, se suelen utilizar los siguientes estimadores:



Se puede probar que, entre momentos poblacionales con respecto a la media y momentos con respecto al origen se dan las siguientes relaciones:



Con los momentos poblacionales de la variable aleatoria X, podemos obtener algunos coeficientes que nos dan una idea de la forma de la función de distribución, como por ejemplo el coeficiente de sesgo y el coeficiente de Kurtosis.

**Sesgo de simetría:** Es el grado de simetría de una distribución.

****

Si este coeficiente es mayor a cero la distribución es asimétrica sesgada a la derecha, si es menor a cero la distribución es asimétrica sesgada a la izquierda y si es igual a cero la distribución es simétrica.

**Kurtosis:** Es el grado de apuntamiento de una distribución.

****

Si este coeficiente es mayor a tres se dice que la distribución es leptocúrtica, si es menor a tres se dice que la distribución es platicúrtica y si es igual a tres se dice que la distribución es mesocúrtica.

Función de verosimilitud:

Sea X1,X2,...,Xn variables aleatorias y x1, x2,..., xn n observaciones y , el conjunto de parámetros poblacionales.

Si X1,X2,...,Xn son variables aleatorias discretas entonces la función de verosimilitud  es la probabilidad conjunta de x1, x2,..., xn.

Si X1,X2,...,Xn son variables aleatorias continuas entonces la función de verosimilitud  es la función de densidad conjunta en x1, x2,..., xn.

**Ejemplos:**

A continuación se presentan las funciones de probabilidad de dos variables aleatorias discretas, una simétrica y la otra asimétrica, así como las distribuciones de los estimadores de la media y la mediana con sus respectivas medidas descriptivas, tratadas en las secciones anteriores. Se trabaja con un tamaño muestral (n=3).

* **Población Asimétrica**

Sea X una variable aleatoria discreta con función de probabilidad f(x)=P(X=x).



Para esta función de probabilidad, el estadístico de primer orden es 1, el estadístico de último orden es 5, la media poblacional es 3.33, la mediana poblacional es 3.5 y la varianza poblacional es 2.22.

**Gráfico 1.1**

# Estimación por el método Jacknife

**Gráfico de la Función de Probabilidad de la Variable Aleatoria X Asimétrica**

f(x)

X

**Elaboración:** *R. Plúa*

A partir de la función de probabilidad descrita presentamos la función de probabilidad para sus principales estimadores.

Sea  la media muestral, la mediana muestral,  el estimador insesgado de la varianza,  el estimador de máxima verosimilitud de la varianza,  el primer estadístico de orden yel n-ésimo estadístico de orden.





Las medidas descriptivas para cada uno de los estadísticos se muestran en la Tabla 1:

**Tabla I**

# Estimación por el Método Jacknife

Parámetros de las funciones de probabilidad de los principales estimadores para la población X Asimétrica.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Media muestral** | **Mediana Muestral** | **Primer estadístico de orden** | **Último estadístico de orden** | **Varianza muestral (Inses.)** | **Varianza muestral (máx. ver.)** |
| Media | 3.333 | 3.500 | 1.750 | 4.750 | 2.667 | 1.778 |
| Mediana | 3.333 | 3.500 | 1.500 | 5.000 | 2.333 | 1.556 |
| Moda | 3.333 | 3.000 | 1.000 | 5.000 | 2.333 | 1.556 |
| Varianza | 0.468 | 1.105 | 0.829 | 0.303 | 2.152 | 0.956 |
| Curtosis | -0.377 | -1.100 | 0.260 | 4.657 | -1.176 | -1.176 |
| Asimetría | 0.000 | 0.000 | 1.017 | -2.239 | 0.156 | 0.156 |
| Rango | 2.667 | 3.000 | 3.000 | 2.000 | 5.000 | 3.333 |
| Mínimo | 2.000 | 2.000 | 1.000 | 3.000 | 0.333 | 0.222 |
| Máximo | 4.667 | 5.000 | 4.000 | 5.000 | 5.333 | 3.556 |

**Elaboración:** *R. Plúa*

Como podemos el valor esperado para la función de probabilidad de la media muestral coincide con el valor esperado de la función de probabilidad para la variable aleatoria X, por lo que la media muestral es un estimador insesgado para la media poblacional, el estimador de la mediana muestral también es insesgado ya que el valor esperado del mismo coincide con el valor de la mediana poblacional. La función de probabilidad de estos estimadores es simétrica, como se puede constatar al observar el coeficiente de asimetría, dado en la Tabla I.

El sesgo de estimación para el estadístico de primer orden es , el sesgo de estimación para el último estadístico de orden es de , el sesgo de estimación para la varianza muestral , y el sesgo de estimación para el estimador de la varianza obtenido por el método de máxima verosimilitud es ; lo que indica que estos estimadores son sesgados. Las funciones de probabilidad para estos estimadores son asimétricas, como se muestra en la Tabla 1, el coeficiente de asimetría es diferente de cero.

* **Población Simétrica**

Sea X una variable aleatoria con función de probabilidad f(x).



Para esta función de probabilidad, el estadístico de primer orden es 1, el estadístico de último orden es 5, la media poblacional es 3, la mediana poblacional es 3 y la varianza poblacional es 1.4285.

A partir de la función de probabilidad descrita presentamos la función de probabilidad para sus principales estimadores:

**Gráfico 1.2**

# Estimación por el método Jacknife

**Gráfico de la Función de probabilidad f(X) Simétrica**

f(x)

X

**Elaboración:** *R. Plúa*





Las medidas descriptivas para cada uno de los estadísticos se muestran en la Tabla 2:

**Tabla II**

# Estimación por el Método Jacknife

Parámetros de las funciones de probabilidad de los principales estimadores para la población X Simétrica.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Media muestral** | **Mediana Muestral** | **Primer estadístico de orden** | **Último estadístico de orden** | **Varianza muestral (Inses.)** | **Varianza muestral (máx. ver.)** |
| Media | 3,000 | 3,000 | 1,857 | 4,143 | 1,667 | 1,111 |
| Mediana | 3,000 | 3,000 | 2,000 | 4,000 | 1,333 | 0,889 |
| Moda | 2,667 | 3,000 | 1,000 | 5,000 | 1,000 | 0,667 |
| Varianza | 0,317 | 0,294 | 0,714 | 0,714 | 1,536 | 0,683 |
| Curtosis | -0,736 | 0,773 | -1,553 | -1,553 | -0,009 | -0,009 |
| Asimetría | 0,000 | 0,000 | 0,285 | -0,285 | 0,929 | 0,929 |
| Rango | 2,000 | 2,000 | 2,000 | 2,000 | 4,333 | 2,889 |
| Mínimo | 2,000 | 2,000 | 1,000 | 3,000 | 0,000 | 0,000 |
| Máximo | 4,000 | 4,000 | 3,000 | 5,000 | 4,333 | 2,889 |

**Elaboración:** *R. Plúa*

Como podemos el valor esperado para la función de probabilidad de la media muestral coincide con el valor esperado de la función de probabilidad para la variable aleatoria X, por lo que la media muestral es un estimador insesgado para la media poblacional, el estimador de la mediana muestral es un estimador insesgado para la media poblacional porque su valor esperado también coincide con valor de la media poblacional. La función de probabilidad de estos estimadores es simétrica, como se puede constatar al observar el coeficiente de asimetría dado, en la Tabla 2.

El sesgo de estimación para el estadístico de primer orden es , el sesgo de estimación para el último estadístico de orden es de , el sesgo de estimación para la varianza muestral , y el sesgo de estimación para el estimador de la varianza obtenido por el método de máxima verosimilitud es ; lo que indica que estos estimadores son sesgados. Las funciones de probabilidad para estos estimadores son asimétricas, como se muestra en la Tabla II, el coeficiente de asimetría es diferente de cero.

**1.5 Generación de números aleatorios de distribuciones poblacionales por el método de la transformada inversa.**

Para trabajos experimentales como el que vamos a desarrollar, muchas veces se tiene que simular las observaciones que se obtendrían al realizar el experimento, si proviniesen de determinada población. Para esto, se suele utilizar el método de la transformada inversa.

Sea X una variable aleatoria que sigue una distribución U(0,1), esto es:



Sea Y una variable aleatoria descrita por la función de densidad g(y) y función de distribución acumulada G(y).

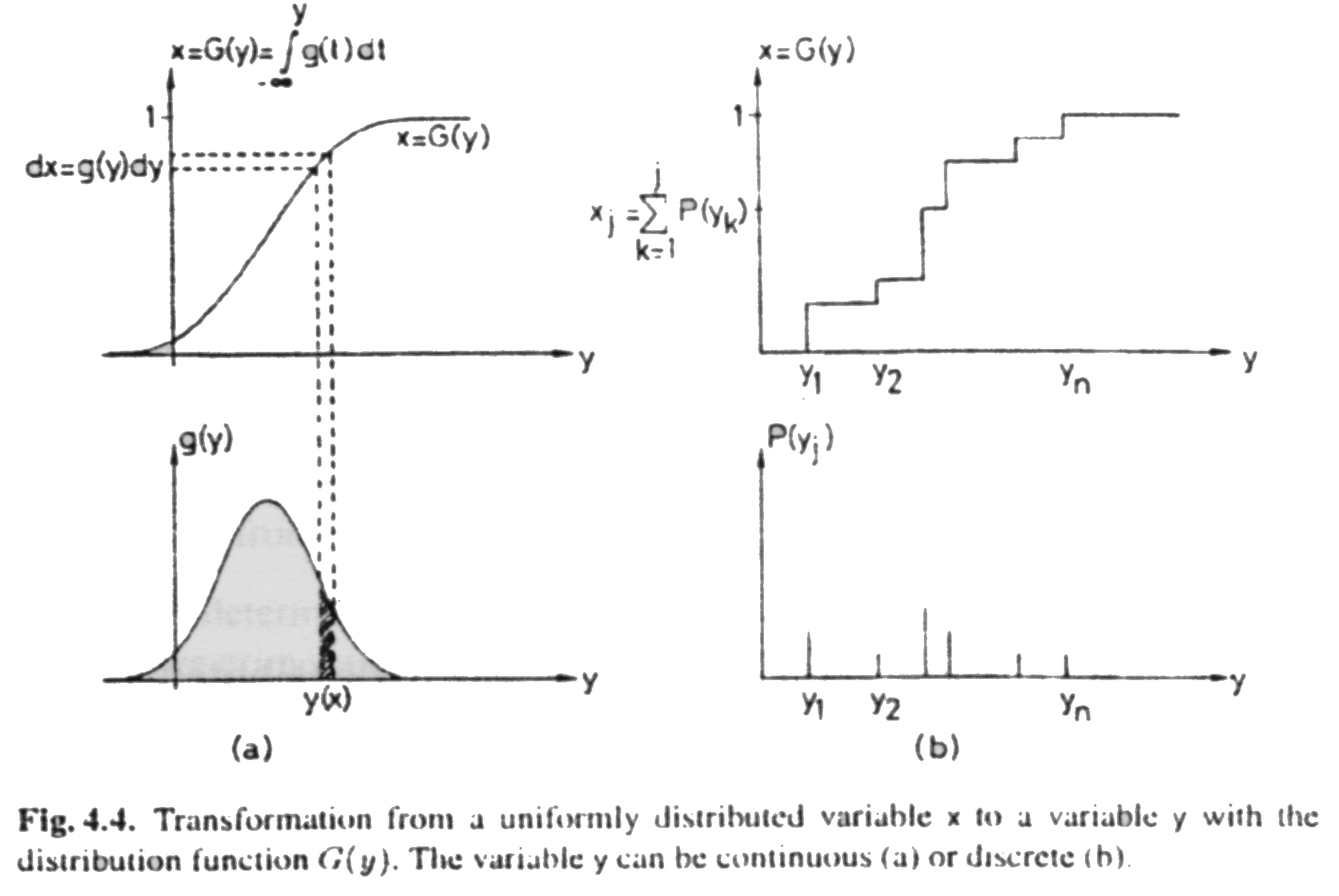


Tenemos un número aleatorio descrito por la función de densidad g(y).

**Gráfico 1.3**

# Estimación por el método Jacknife

**Transformación de una variable aleatoria uniforme X a una variable aleatoria Y con distribución G(y).**



## Fuente: *S. Brandt*

Como se puede observar en el gráfico 1.1 (b) la relación mencionada puede ser usada en el caso de distribuciones discretas.

En los siguientes ejemplos, estipulamos una variable aleatoria con función de densidad Y de la cual obtenemos un número aleatorio, por el método de la transformada inversa descrito anteriormente.

* Sea X una variable aleatoria U(0,1) y Y una variable aleatoria con función de densidad exponencial con parámetro *β*:



* Sea X una variable aleatoria U(0,1) y Y una variable aleatoria con función de probabilidad



La función de distribución para la función de densidad de Y es:



La números aleatorios obtenidos de la función de densidad g(y), se los obtiene con la siguiente regla de correspondencia:



* Sea Y una variable aleatoria con función de densidad N(0,1)



Se toma las coordenadas de un espacio de puntos uniformemente distribuido dentro del círculo unidad.

El punto (v1,v2)

Donde v1=2(u1)-1 y v2=2(u2)-1

Expresados en coordenadas polares;

v1=r Cos θ

v2=r Cos θ

donde 

Siendo 

El punto (y1,y2) en coordenadas polares es:



La densidad de probabilidad de r es:



La función de densidad conjunta de y1 y y2 es:



Por tanto, y1 y y2 provienen de distribuciones normales estándares independientes.

**1.6 Métodos de estimación**

Sabemos que un parámetro poblacional puede tener más de un estimador, por lo tanto se suelen utilizar métodos de estimación para la obtención de los mismos como lo son el método de los momentos y el método de máxima verosimilitud, los cuales se explican en esta sección.

**1.6.1 Método de Máxima Verosimilitud**

Sea X1,X2,…,Xn una muestra aleatoria tomada de una población X con p parámetros desconocidos .

El método consiste en escoger como estimadores de los parámetros poblacionales aquellos valores de los parámetros que maximizan la función de verosimilitud .

Se construye la función de verosimilitud , los valores de los parámetros que maximizan la función se los obtiene derivando la misma, pero algunas veces es dificultoso y tedioso la derivación en un producto, por lo que construimos la función  ya que por ser la función logaritmo natural monótona, creciente y no negativa se maximiza en el mismo punto. Posteriormente derivamos la función  con respecto a cada parámetro poblacional obteniendo p ecuaciones con p incógnitas; al resolver el sistema de ecuaciones tendremos los p estimadores de máxima verosimilitud para los p parámetros desconocidos.



En el siguiente ejemplo ilustraremos el método de máxima verosimilitud para la obtención de los estimadores de la media y varianza poblacional.

Sea X1,X2,…,Xn una muestra aleatoria tomada de una población Normal con media μ y varianza σ2.

 es la función de densidad conjunta de la muestra. Por lo tanto:



Tenemos un sistema de dos ecuaciones con dos incógnitas, de la ecuación (1) obtenemos:





Al sustituir  por  en la ecuación (2):



Demostraremos que  es un estimador insesgado de .



Por el contrario  es un estimador sesgado de .



El sesgo de estimación es .

**1.6.2 Método de los momentos**

Sea X1,X2,…,Xn una muestra aleatoria tomada de una población X con p parámetros desconocidos .

El método se basa en el supuesto de que los momentos alrededor del origen de la muestra deben proporcionar estimaciones apropiadas para los momentos alrededor del origen de la población. Por lo tanto podemos expresar:



Obteniendo un sistema de p ecuaciones con p incógnitas que son los parámetros desconocidos. Al resolver el sistema de ecuaciones obtendremos los estimadores de los parámetros poblacionales.

Presentaremos la estimación de los parámetros de la función de densidad Gamma utilizando el método de los momentos, la obtención de estos estimadores por medio del método de máxima verosimilitud resultaría muy dificultosa.

Sea X una variable aleatoria con función de densidad Gamma:



Su media y su varianza son respectivamente  y .

Por tanto, al igualar los momentos poblacionales alrededor del origen con los momentos muestrales alrededor del origen tenemos:



Reemplazando , tenemos:



Obteniendo finalmente:



**1.7 Convergencia en distribución (o en Ley)**

Dada una sucesión {Xn} de variables aleatorias independientes e idénticamente distribuidas X1, X2,…,Xn con funciones de distribución F1(X1), F2(X2),….,Fn(Xn), respectivamente, se dice que esa sucesión converge en distribución a la variable aleatoria X con función de distribución F(X), si:



Esto es Xn converge en ley a X, uno de los ejemplos más importantes de la convergencia en distribución es el Teorema del Límite Central.

**Teorema del límite Central:** Si X1, X2,…,Xn es una muestra aleatoria de tamaño n, tomada de una población X con media μ y varianza σ2<∞.

Definimos: .

Entonces la función de distribución de la sucesión *Un*converge en ley a una función de distribución normal estándar.

Es decir .

**1.8 Principales Distribuciones Discretas y Continuas**

En el Anexo 1 presentamos las principales distribuciones discretas y en el Anexo 2 las principales distribuciones continuas, donde podemos observar sus parámetros poblacionales así como los gráficos de las funciones de distribución para cada una de ellas con diferentes valores de los parámetros.