

Inferencia Estadística

Daniel Franzani
Universidad Católica Silva Henríquez

01 - 03 - 2022

Contents

Acerca del curso	5
1 Variables aleatorias continuas	7
1.1 Definición v.a continua	11
1.2 Función de densidad y distribución	11
1.3 Momentos de una v.a continua	14
1.4 Función generadora de momentos	16
1.5 Modelos de v.a continua	16
1.6 Vectores de v.a continuas	17
1.7 Probabilidad condicionada	17
2 Muestras y Distribuciones muestrales	19
2.1 Conceptos básicos de estadística inferencial	19
2.2 Distribución de parámetros	19
2.3 Distribuciones	19
3 Estimación de parámetros	21
3.1 Estimación puntual	21
3.2 Propiedades de los estimadores puntuales	21
3.3 Estimación por intervalos	21
4 Prueba de hipótesis	23
4.1 Hipótesis estadísticas	23
4.2 Tipos de errores	23
4.3 Lema de Neyman y Pearson	23
4.4 Pruebas de dos parámetros	23
4.5 Pruebas de bondad y ajuste	23

Acerca del curso

El siguiente documento abarca los contenidos correspondientes al curso de Inferencia Estadística. Las temáticas que se abordan, son las siguientes:

- Unidad 1: Variables aleatorias continuas.
- Unidad 2: Muestras y Distribuciones muestrales.
- Unidad 3: Estimación de parámetros.
- Unidad 4: Prueba de hipótesis.

Para mayor detalle respecto a los temas, evaluaciones, resultados e indicadores de aprendizaje, consultar el programa del curso.

Además, pueden encontrar el documento en un versión PDF descargable aquí.

Chapter 1

Variables aleatorias continuas

Al escuchar de variable aleatoria lo asociamos a los valores en un experimento azaroso, lo cual no está realmente tan alejado de la definición formal. Una variable aleatoria es una función $X(\omega)$ con recorrido en los reales (asigna valores reales a los resultados de un experimento), definida sobre un espacio de probabilidad (Ω, F, P) , tal que cumple la siguiente condición:

$$\{\omega \in \Omega : X(\omega) \leq x\} \in F, \forall x \in R. \quad (1.1)$$

Sin embargo, ¿qué significa está condición?

Para entenderlo, consideremos el experimento de tirar 1 vez un dado equilibrado de 6 caras, como bien sabemos, los resultados posibles de tirar el dado son: $\Omega = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}$ (conjunto que contiene todos los posibles resultados, cada uno de ellos es un conjunto de un elemento).

Ahora, elijamos un función con recorrido en los reales, por ejemplo $X(\omega) = 3\omega$ (pueden probar con la función que deseen). Luego, para verificar si la función propuesta es una variable aleatoria, debemos verificar la condición de (1.1). La clave de la condición, es probar que se cumple para todo x en los reales. Para llegar a esto seguiremos los siguientes pasos:

1. Determinar el recorrido de $X(\omega)$, es decir, aplicaremos la función X a los posibles resultados del experimento:

```
# Valores del experimento
valores = c(1,2,3,4,5,6)

funcion_X = function(evaluar){
```

```

    return(3*evaluar) # Devuelve los valores del experimento multiplicados por 3
}

# Valores al aplicar la función
funcion_X(valores)

## [1]  3  6  9 12 15 18

```

2. Luego debemos ver que valores del experimento (1, 2, 3, 4, 5 y/o 6) cumple que, al reemplazarlos en la función son menores o igual a un número real. Como habrán notado, no podemos verificar esta condición para cada número real por separado (porque básicamente son infinitos, y no tenemos tiempo para ello), por ende, debemos particionar los reales en intervalos (pero ¿cómo?, 0.0).

La clave de la partición de los intervalos, es acotarlos solamente por la derecha, por los elementos obtenidos al aplicar la función X : $(-\infty, 3]$, $(-\infty, 6]$, $(-\infty, 9]$, $(-\infty, 12]$, $(-\infty, 15]$, $(-\infty, 18]$, además, debemos agregar los siguientes dos intervalos: $(-\infty, 3)$ y $(18, \infty)$.

Para entender esta partición, consideremos lo siguiente:

- La unión de todos ellos es igual a los reales.
- El intervalo $(-\infty, 3)$ está considerado intencionadamente, para que no contenga a ningún valor del recorrido de X (por la izquierda).
- El intervalo $(18, \infty)$ está considerado intencionadamente, para que no contenga a ningún valor del recorrido de X (por la derecha).
- Los 6 intervalos acotados por los resultados del experimento, contienen al número de la cota y aquellos que son menores a él, que también pertenecen al experimento, por ejemplo, el intervalo $(-\infty, 12]$ contiene al 12, incluyendo a 3, 6 y 9.

Naturalmente, está no es la única forma de particionar los reales, pero es posible considerar esto como una técnica en caso de que deseen replicarlo en otros ejercicios.

3. Por último, debemos visualizar la desigualdad presenten en la condición. Para ello, utilicemos el intervalo $(-\infty, 6]$.

Consideremos un valor cualquiera de x en $(-\infty, 6]$, por ejemplo $x = 5$. Luego, debemos ver que valores del recorrido son menores o igual a 5, en este caso la respuesta es solo el 3, y luego, veamos cual (o cuales) es el valor en Ω que me lleva a 3, la respuesta es el 1. Sin embargo, si consideramos $x = 6$, los valores del experimento que son menores o iguales son 3 y 6, y los valores de Ω que me llevan a esos valores son 1 y 2 respectivamente. Entonces, ¿qué valor de x debemos tomar? La respuesta es, el “extremo” derecho del intervalo del intervalo, porque abarcará la mayor cantidad de opciones (esto no es una explicación formal, pero en el fondo tratamos de abarcar todas las opciones posibles de Ω).

Lo anterior implica que, para los intervalos $(-\infty, 3]$, $(-\infty, 6]$, $(-\infty, 9]$, $(-\infty, 12]$, $(-\infty, 15]$, $(-\infty, 18]$ se considera x igual a 3, 6, 9, 12, 15 y 18 respectivamente. En el caso del intervalo $(-\infty, 3)$ consideramos el “extremo” de $x < 3$ abierta (“es decir 2.999999999...”), y en el caso $(18, \infty)$ puede ser cualquier número que sea mayor a cualquier elemento del recorrido (por ejemplo $x = 19$).

La siguiente tabla muestra un resumen.

Intervalo	Valor de x	Valores del recorrido $\leq x$	Valores de Ω correspondientes
$(-\infty, 3)$	2.99999...	$\{\emptyset\}$	$\{\emptyset\}$ (ninguno)
$(-\infty, 3]$	3	$\{3\}$	$\{1\}$
$(-\infty, 6]$	6	$\{3, 6\}$	$\{1, 2\}$
$(-\infty, 9]$	9	$\{3, 6, 9\}$	$\{1, 2, 3\}$
$(-\infty, 12]$	12	$\{3, 6, 9, 12\}$	$\{1, 2, 3, 4\}$
$(-\infty, 15]$	15	$\{3, 6, 9, 12, 15\}$	$\{1, 2, 3, 4, 5\}$
$(-\infty, 18]$	18	$\{3, 6, 9, 12, 15, 18\}$	$\{1, 2, 3, 4, 5, 6\}$
$(18, \infty)$	19	$\{3, 6, 9, 12, 15, 18\}$	$\{1, 2, 3, 4, 5, 6\} = \Omega$

4. El paso final, es verificar si los conjuntos que se muestran en la cuarta columna de la tabla pertenecen a F (σ - álgebra). Usualmente, F se da en los problemas, aunque podría pedirse al estudiante que la construya (para efectos de este curso, F siempre será dada en este contexto).

A modo ilustrativo, consideremos un F como el conjunto (conjunto de conjuntos) $\{\{\emptyset\}, \{1, 2, 3\}, \{1, 3\}, \{1, 2, 6\}\}$ (este ejemplo, realmente no cumple con las condiciones de una σ - álgebra). Como podemos observar, muchos de los conjuntos de la cuarta columna **no se encuentran** en F , por lo que la función X **no sería una variable aleatoria** (en caso de que todos los conjuntos de la cuarta columna se encontraran en F , entonces X sí sería un variable aleatoria).

5. **Nota:** Independiente de la forma en la que se eligen los intervalos, la conclusión siempre es la misma (aunque los resultados de la cuarta columna pueden cambiar). Esto es debido a propiedades de F , las cuales comentaremos más adelante.

Llegando a este punto, se preguntarán de que sirve la función X . Esta función da origen a otra, llamada función de distribución, la se define de la siguiente manera.

Definición: Dado un espacio de probabilidad y una variable aleatoria $X(\omega)$. Se denomina función de distribución de la variable aleatoria X , a la función $F_X(x)$ definida para todos los valores reales de x mediante la fórmula:

$$F_X(x) = P_X(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x\}) \quad (1.2)$$

Como se puede apreciar, se está interesado en calcular las probabilidades de los conjuntos que aparecen en la cuarta columna de la tabla anteriormente vista. Cabe mencionar, que los conjuntos que están en la tabla no son los únicos que se pueden formar con los elementos de Ω , por ejemplo falta el conjunto $\{1, 3\}$.

Veamos el siguiente ejemplo, que relaciona el espacio muestral Ω , la variable aleatoria X y la probabilidad P .

Ejemplo: Un experimento consiste en tirar 2 monedas, y se está interesado en determinar la cantidad de sellos obtenidos. Determine la probabilidad de que ocurra cada una de las opciones posibles.

Como se puede apreciar, el experimento es discreto, de tal manera que en el diagrama de más abajo en la columna de Ω se observan los distintos resultados posibles del experimento (cara = c, sello = s).

Luego, la columna de la variable aleatoria X refleja la cantidad de sellos evidenciadas en cada uno de los posibles resultados del experimento.

El último paso, es determinar la probabilidad de cada uno de las opciones posibles en X , la cual arroja solo tres resultados distintos: 0, 1 y 2. Para calcular la probabilidad de cada caso, debemos contar la cantidad de elementos de Ω que nos envían a cada uno de los resultados en X y dividirlo por la cantidad total de elementos de Ω (*definición clásica de probabilidad*).

En el caso de $X = 1$, se observa que hay dos elementos del espacio muestral que obtiene ese valor a aplicar la función X , los cuales son cs y sc , por lo que son 2 elementos, mientras que la cantidad total de elementos de Ω es de 4 (tal como se ve en el diagrama), por lo tanto, la probabilidad de $X = 1$ es de $2/4 = 0.5$.

Ω	\rightarrow	$X(\omega)$	\rightarrow	$P(X = x)$
cc		0		$P(X = 0) = 1/4 = 0.25$
cs		1		$P(X = 1) = 2/4 = 0.5$
sc		1		
ss		2		$P(X = 2) = 1/4 = 0.25$

Podemos apreciar que existe un leve cambio entre la definición y el ejemplo, y es que calculamos la expresión $P(X = x)$ en vez de $P(X \leq x)$. Sin embargo, como ya saben en el caso del ejemplo (caso discreto) se tiene que

$$P(X \leq x) = \sum_{i=0}^x P(X = i) = F_X(x). \quad (1.3)$$

De una manera más formal y genérica, podemos escribirlo de la siguiente forma

$$P(X \leq x) = \sum_{i=-\infty}^x P(X = x_i) = F_X(x). \quad (1.4)$$

Por ejemplo, $F_X(1) = P(X \leq 1) = P(X = 0) + P(X = 1) = 0.25 + 0.5 = 0.75$.

Por último, cabe mencionar, que es posible verificar si una función F es de distribución (sin asociar a la variable aleatoria). Para ello, se deben satisfacer 3 condiciones:

1. Monótona: $x < y \Rightarrow F(x) \leq F(y)$
2. Continuidad por la derecha
3. $\lim_{x \rightarrow \infty} F(x) = 1$ y $\lim_{x \rightarrow -\infty} F(x) = 0$

A partir de este punto, trabajaremos directamente con X , es decir, la variable aleatoria, dejaremos de lado el trabajo con el espacio muestral Ω y F . Para entender de otra perspectiva el concepto de variable aleatoria pueden mirar este vídeo.

1.1 Definición v.a continua

Los tipos de variables aleatorias (v.a) se diferencian por el tipo de recorrido que tiene la función $X(\omega)$, en este sentido se identifican dos:

1. **Variables aleatorias discretas:** los valores del recorrido de la función X son numerables. Por ejemplo, contar la cantidad de sellos que se obtienen al tirar dos monedas (valores discretos entre un mínimo y un máximo, no es estrictamente necesario la presencia de un mínimo y/o máximo).
2. **Variables aleatorias continuas:** los valores del recorrido de la función X son no numerables. Por ejemplo, determinar la estatura de las personas de una determinada ciudad (valores reales entre una estatura mínima y máxima, no es estrictamente necesario la presencia de un mínimo y/o máximo).

1.2 Función de densidad y distribución

Al igual que una v.a discreta, una del tipo continua posee su propia expresión para definir una función de distribución y densidad. Las expresiones son las siguientes:

1. Dado un espacio de probabilidad y una v.a continua X . Decimos que X tiene distribución continua si la **función de distribución** $F_X(x)$ puede representarse de la forma

$$F_X(x) = \int_{-\infty}^x f_X(u) du, \quad (1.5)$$

donde $f_X(x)$ es una función no negativa y Lebesgue integrable (para este curso nos bastará utilizar Riemann integrable).

2. La función $f_X(x)$ se denomina **función de densidad** de la variable aleatoria X . Algunas propiedades de esta función son:

- a. $\int_{-\infty}^{\infty} f_X(u) du = 1$.
- b. Si $f_X(u)$ es continua en todos los puntos, entonces $F'_X(x) = f_X(x)$ para todo x (también cierto en caso puntual).

Ejemplo: Consideremos la siguiente función de densidad de una variable aleatoria X

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in (a, b) \\ 0 & \text{si } x \notin (a, b) \end{cases}$$

Primero, verifiquemos si la integral en todo el dominio es igual a 1.

$$\int_a^b \frac{1}{b-a} du = \frac{1}{b-a} \int_a^b du = \frac{1}{b-a} (u|_a^b) = \frac{b-a}{b-a} = 1$$

Esto implica que la función de densidad está “bien definida”.

Ahora, determinemos la función de distribución $F_X(x)$.

$$F_X(x) = \int_a^x \frac{1}{b-a} du = \frac{1}{b-a} \int_a^x du = \frac{1}{b-a} (u|_a^x) = \frac{x-a}{b-a}$$

La figura 1.1 muestra la relación que existe entre ambas funciones en el intervalo $(1, 10)$. En particular, el área de bajo la curva de densidad desde el extremo izquierdo hasta 5, es igual a evaluar el 5 en la función de distribución. Ambas funciones son admisibles para el cálculo de probabilidades, ya que

$$P_X(X \leq x) = F_X(x) = \int_{-\infty}^x f_X(u) du, \quad (1.6)$$

lo cual, en el caso del ejemplo es

$$P_X(X \leq 5) = F_X(5) = \int_1^5 f_X(u) du = 0.44,$$

Como bien se puede apreciar, el ejemplo calcula expresamente el área de bajo la curva de la función densidad desde el extremo izquierdo del dominio hasta el 5. Sin embargo, es posible calcular áreas de entre puntos que no sean necesariamente los extremos, por ejemplo, entre 3 y 6.

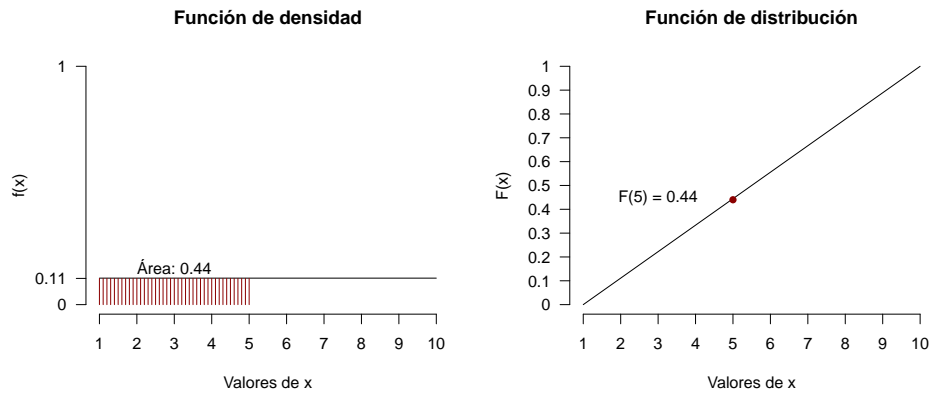


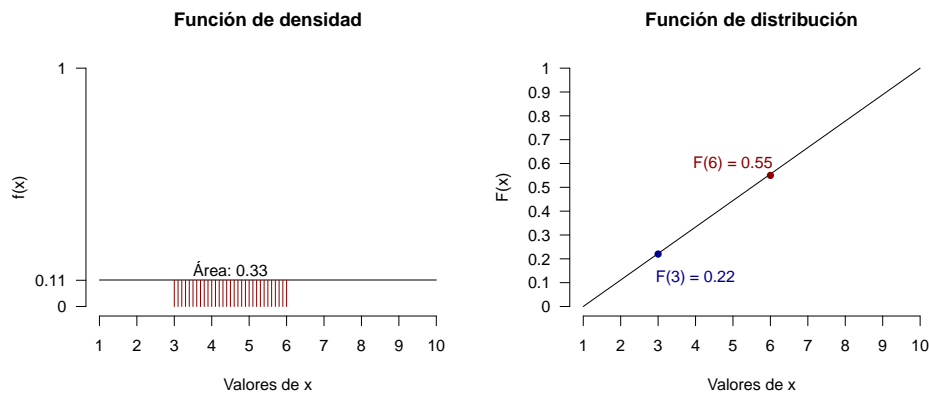
Figure 1.1: Relación entre la función de densidad y distribución

La diferencia está en como determinar el área calculada en la función de densidad desde la función de distribución. Para ello contamos con la siguiente propiedad:

$$P_X(a \leq X \leq b) = \int_a^b f_X(u) du = F_X(b) - F_X(a), \quad (1.7)$$

la cual, en el caso del ejemplo se traduce en:

$$P_X(3 \leq X \leq 6) = \int_3^6 f_X(u) du = F_X(6) - F_X(3) = 0.33,$$



Nota: En la propiedad (1.7) los signos de desigualdad pueden ser estrictos o no, ya que no afecta al cálculo de integración ni a la propiedad en si.

Otra propiedad a tener en cuenta, es el hecho de que la probabilidad puntual de un v.a continua es cero, es decir,

$$P_X(X = x) = 0, \forall x \in R.$$

Esta propiedad es verificable utilizando lo mostrado en (1.7).

1.3 Momentos de una v.a continua

1.3.1 Esperanza

EL concepto de esperanza se entiende como “el valor esperado en un *experimento*”. Sin embargo, podemos dotar de exactitud a este concepto mediante la siguiente definición.

Definición: Dado un espacio de probabilidad y una función $g(X(\omega))$ sobre la variable aleatoria continua X . La esperanza o media de $g(X)$ esta dada por

$$E_X\{g(X)\} = \int_{-\infty}^{\infty} g(x)f_X(x)dx \quad (1.8)$$

Hoy en día, esta definición es aquella que engloba el concepto de promedio, ese tan usado a la hora de determinar la nota final de un curso (en el caso discreto).

Ejemplo: La función $f_X(x)$ expresa el tiempo que transcurre hasta que llega la primera llamada a mi celular, determinar la $E_X(X)$, donde x corresponde al tiempo que transcurre y $1/\lambda$ indica la tasa de fallo (podemos entenderla como, uno menos la evidencia que existe respecto a la cantidad de llamadas que entran en ese periodo de tiempo en un celular, ejemplo: 10 llamadas por hora).

$$f_X(x) = \frac{1}{\lambda}e^{-x/\lambda}, 0 \leq x < \infty$$

Para determinar la esperanza, seguiremos la definición dada en (1.8). Primeramente podemos ver que $g(X(\omega)) = X(\omega)$, es decir, se utiliza la función identidad. Luego,

$$\begin{aligned}
E_X\{g(X)\} &= \int_{-\infty}^{\infty} g(x)f_X(x)dx \\
&= \int_0^{\infty} x \frac{1}{\lambda} e^{-x/\lambda} dx \\
&= \frac{1}{\lambda} \int_0^{\infty} x e^{-x/\lambda} dx \\
&= \frac{1}{\lambda} \left(-x\lambda e^{-x/\lambda} \Big|_0^{\infty} - \int_0^{\infty} -\lambda e^{-x/\lambda} dx \right) \\
&= \frac{1}{\lambda} \left(0 - \lambda^2 e^{-x/\lambda} \Big|_0^{\infty} \right) \\
&= \frac{1}{\lambda} (0 - 0 + \lambda^2) \\
&= \lambda
\end{aligned}$$

En este caso, y tomando como ejemplo el enunciado, si $1/\lambda = 10$, entonces se espera que la cantidad de tiempo medio hasta que llega la primera llamada sea de 0.1 horas.

En la figura 1.2, se puede apreciar comportamiento de la función de densidad y la esperanza.

Función densidad y valor esperado

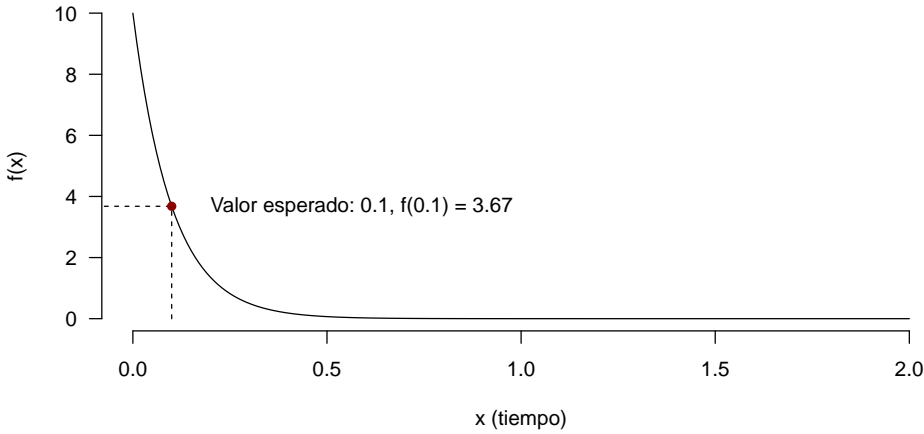


Figure 1.2: Función densidad y valor esperado.

Nota: el valor esperado nunca debe depender de la variable aleatoria, pero si puede depender de constantes, tal como sucede en el ejemplo.

Como se pudo apreciar, el valor de esperanza está influenciado por la expresión de la función g . En este sentido, cuando g corresponde a la función identidad,

estaremos hablando el promedio “usual”, para otros valores de g es necesario saber interpretar, por ejemplo si $g = X^2$, estaremos obteniendo el promedio de los valores al cuadrado.

Por último, algunas propiedades de la esperanza son:

1. $E_X(a \cdot g(X) + b) = aE_X(g(X)) + b$.
2. $E_X(g_1(X) + g_2(X)) = E_X(g_1(X)) + E_X(g_2(X))$.
3. Si $g(X) \geq 0$ para todo x , entonces $E_X(g(X)) \geq 0$.

1.3.2 Momentos

Los momentos están definidos en base a la esperanza. En este sentido, para cada entero n , el n - ésimo momento de una variable aleatoria X es:

$$\mu'_n = E_X(X^n). \quad (1.9)$$

Esto quiere decir, que los momentos son potencias la variable aleatoria, lo cual solo modifica la función $g(X) = X^n$.

Una definición similar es la de **momento centrados**, la cual es una variación de (1.9), donde un momento central está definido como:

$$\mu_n = E_X\{(X - \mu)^n\}, \quad (1.10)$$

donde $\mu = \mu'_1 = E_X(X)$.

En este caso, nuevamente nos encontramos con una variación de la función $g(X)$. Además, dos de la propiedades más relevantes de estos momentos son las siguientes:

1. Si $E_X(X^n) < \infty$, entonces $E_X\{(X - \mu)^n\} < \infty$, para todo entero $n > 0$.
2. Si $E_X(X^n) < \infty$, entonces $E_X(X^m) < \infty$, para todos enteros $k > m > 0$.

1.3.3 Varianza y Desviación Estándar

1.3.4 Momentos normalizados

1.4 Función generadora de momentos

1.5 Modelos de v.a continua

Uniforme, exponencial y normal

1.6 Vectores de v.a continuas

1.7 Probabilidad condicionada

Chapter 2

Muestras y Distribuciones muestrales

2.1 Conceptos básicos de estadística inferencial

2.2 Distribución de parámetros

Media,, varianza, proporción, diferencia de medias con varianza conocida, diferencia de proporciones, cociente de varianzas.

2.3 Distribuciones

Normal, T-student, Chi-cuadrado, F

Chapter 3

Estimación de parámetros

3.1 Estimación puntual

3.2 Propiedades de los estimadores puntuales

3.3 Estimación por intervalos

Chapter 4

Prueba de hipótesis

- 4.1 Hipótesis estadísticas
- 4.2 Tipos de errores
- 4.3 Lema de Neyman y Pearson
- 4.4 Pruebas de dos parámetros
- 4.5 Pruebas de bondad y ajuste