Inferencia Estadística

Daniel Franzani
Universidad Católica Silva Henriquez

01 - 03 - 2022

Contents

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 que toma una determinada variable 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, definida sobre un espacio de probabilidad (Ω, F, P), tal que cumple la siguiente condición:

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

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){
```

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

# 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:

   a. La unión de todos ellos es igual a los reales.
   b. El intervalo $(-\infty, 3)$ está considerado intecionadamente, para que no contenga a ningún valor del recorrido de X (por la izquierda).
   c. El intervalo $(18, \infty)$ está considerado intecionadamente, para que no contenga a ningún valor del recorrido de X (por la derecha).
   d. 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 quie 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 (conjuto 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 si 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:

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

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 solo3 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		$P(X = 2) = 1/4 = 0.25$

1.1 Definición variable aleatoria continua

1.2 Función de densidad y funciónde distribución

1.3 Momentos de una v.a continua: esperanza, varianza, curtosis, asimetría

1.4 Función generadora de momentos

1.5 Modelos devaribales aletatoria continua

Uniforme, exponencial y normal

1.6 Vectores de variables aleatorias 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 estadisticas

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