

Estadísticas y probabilidades 2023-1

Lenin Chavez

Contents

Chapter 1

Page 2

1.1	Muestreo aleatorizado	2
	Debilidades del muestreo aleatorizado — 2	
1.2	Tabla de datos	3
1.3	Tipos de variables	3
1.4	Limpieza de datos	3
1.5	Descriptores	3

Chapter 2

Page 4

2.1	Experimento aleatorio	4
-----	-----------------------	---

Chapter 1

1.1 Muestreo aleatorizado

Definition 1.1.1

MUESTREO ALEATORIO : es aquel procedimiento de selección de la muestra en el que todos y cada uno de los elementos de la población tiene una cierta probabilidad de resultar elegidos. (uv.es)

Definition 1.1.2

MUESTREO ALEATORIO SIMPLE : Elegir una muestra representativa de la población en la que cada elemento de la población tiene la misma probabilidad de ser elegido y todas las muestras del mismo tamaño son igualmente probables.

Note:-

Cuidado con los "unicornios"

Definition 1.1.3

MUESTREO ALEATORIO ESTRATIFICADO : es aquel en el que la población se divide en subpoblaciones o estratos y se elige una muestra aleatoria simple de cada estrato, de manera que en nuestra muestra total la relación entre la población de los estratos y sus unidades muestrales sea de manera proporcional.

Note:-

Garantiza representatividad de los estratos. Por ejemplo, si un estrato representa el $1/6$ de la población, entonces $1/6$ de nuestra muestra debe ser de ese estrato.

Definition 1.1.4

MUESTREO ALEATORIO POR CONGLOMERADOS : Asume que todos los conglomerados son equivalentes y que cada uno de ellos tiene la misma probabilidad de ser elegido.

1.1.1 Debilidades del muestreo aleatorizado

- M.A.S.
 - Es necesario tener acceso explícito a toda la población
 - Es potencialmente costoso y demandante de tiempo
- M.E.
 - El conocimiento sobre los estratos y sus tamaños es imorescindible
 - Es necesario acceso a todos los estratos y sus U.M.

- M.C.
 - Es necesario que los conglomerados sean genuinamente parecidos
 - Hay conglomerados que no se muestrean

Note:-

El resultado típico de tomar una muestra es una tabla de datos.

1.2 Tabla de datos

Definition 1.1

Tabla de datos : es prolija si tiene una observación por fila y una variable por columna.

1.3 Tipos de variables

Definition 1.2

Variable : es una característica que puede tomar diferentes valores.

- **Variable numérica** : es una variable que puede tomar valores numéricos.
 - **Variable discreta** : es una variable que puede tomar valores numéricos enteros. Se puede contar, ordenar, aritmética.
 - **Variable continua** : es una variable que puede tomar valores numéricos reales. Se puede contar, ordenar, aritmética.
- **Variable categórica** : es una variable que puede tomar valores no numéricos.
 - **Variable nominal** : es una variable que puede tomar valores no numéricos que no tienen orden. Se puede contar.
 - **Variable ordinal** : es una variable que puede tomar valores no numéricos que tienen orden. Se puede contar, ordenar.

1.4 Limpieza de datos

- **Política de subsanación de errores**: Puedes eliminar todo lo que tenga errores, intentar arreglar el error (si haces esto la mejor opción es preguntar a la fuente original ya que de lo contrario estarías suponiendo y puede contaminar la muestra, reemplazarlos por N.A)
- **Política de manejo de errores** : Eliminar todo lo que tiene datos faltantes, intentar llenar los datos faltantes (Mejor opción es preguntar a la fuente original, la otra opción es técnicas de imputación), vivir con ellos lo cuál es peligroso para la muestra

1.5 Descriptores

- **Numéricos**: - Es un número que trata de resumir algo sobre los datos: Hay de posición (Moda, mediana, promedio), dispersión (rango, rango intercuartil, varianza, desviación estándar) e interacción (covarianza, correlación).
- **Gráficos**: - Un conjunto de números junto con una visualización tratan de resumir algo sobre los datos hay de posición (diagramas de barras, puntos, líneas, histogramas), dispersión (caja y bigotes) e interacción (diagramas disp., boxplots indexados, mosaicos) .
- **Tabulares**: - Listas de valores extraídos de la muestra (típicamente, tablas de frecuencia)

Chapter 2

2.1 Experimento aleatorio

Ω FP Modelo de un experimento aleatorio.

Ω : Espacio de todos los resultados de posibles interés.

$F \subseteq 2^\Omega \rightarrow F$: Álgebra de eventos, 2^Ω : conjunto de partes de Ω

F : Están todos los eventos de interés. F es un conjunto de conjuntos que debe cumplir condiciones técnicas. Las uniones, intersecciones, complementos, etc de eventos son eventos.

P es una función de conjuntos que a cada conjunto en F le asigna un número entre 0 y 1 conocido como su probabilidad. P también debe cumplir condiciones técnicas.

$$P = F \rightarrow [0, 1]$$

Las condiciones técnicas deben obedecer los siguientes axiomas.

Axioma 1: $\forall A \in F, P(A) \geq 0$

Axioma 2: $P(\Omega) = 1$

Axioma 3: $\forall A_1, A_2, \dots \in F$

$$\text{disjuntos } P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Propiedades de P :

- Propiedad 1: $P(\emptyset) = 0$
- Propiedad 2: $P(\bar{A}) = 1 - P(A)$
- Propiedad 3: Si $A, B \in F$ y $A \cap B = \emptyset$ entonces, $P(A \cup B) = P(A) + P(B)$
- Propiedad 4: Si $A, B \in F$ entonces $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Propiedad 5: Si $A, B \in F$ y $A \subseteq B$ entonces $P(A) \leq P(B)$

Note:-

La P nos da una forma alterna de "medir" el tamaño de un conjunto

Example 2.1.1 (Lanzamiento de una moneda justa)

$$\Omega_1 = \{\text{cara, sello}\}$$

$$F_1 = \{\emptyset, \{\text{cara}\}, \{\text{sello}\}, \{\text{cara, sello}\}\}$$

$$P_1 = \{\emptyset = 0, \{\text{cara}\} = 1/2, \{\text{sello}\} = 1/2, \{\text{cara, sello}\} = 1\}$$

Example 2.1.2 (Lanzamos un dado de seis caras, juusto y anotamos el resultado)

$$\Omega_2 = \{1,2,3,4,5,6\}$$

$$F_2 = \{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}, \{1,6\}, \{2,3\}, \{2,4\}, \{2,5\}, \{2,6\}, \dots\}$$

$$P_2(\text{sale par}) = 1/2$$

$$P_2(\text{sale impar}) = 1/2$$

$$P_2(\text{sale 1}) = 1/6$$

$$P_2(\text{no sale 1}) = 5/6$$

Example 2.1.3 (Lanzo un dado justo y anoto que sale pero si sale 4, 5 o 6 anoto No en vez)

$$\Omega_3 = \{1,2,3,\text{No}\}$$

$$F_3 = 2^{\Omega_3}$$

$$|F_3| = 16$$

Usando EA del siguiente ejemplo

$$P_3(\{i\}) = 1/6$$

$$P_3(\text{Sale 3 o No}) = 4/6$$

Example 2.1.4 (Experimento accesorio)

Lanzo un dado justo y anoto lo que sale.

$$P_A(\{4,5,6\}) = 1/2$$

$$P_3(\{1,2,3\}) = 1/2$$

$$P_3(\{\text{No}\}) = 1/2$$

Definition 2.1.1: EA: Experimento aleatorio

(ΩFP)(Con Ω finito) Es un espacio **Equiprobable** si todos sus eventos atómicos tienen la misma probabilidad. E1, E2, EA son espacios Equiprobable, E3 no lo es.

Cuando un espacio es equiprobable es fácil calcular probabilidades (ΩFP) es equiprobable entonces:

$$\forall x \in \Omega, P(\{x\}) = c$$

$$P(A) = \sum_{x \in A} P(\{x\}) = \sum_{x \in A} c = c|A| = \frac{|A|}{|\Omega|}$$

A es casos favorables y Ω es casos totales.

ESTO SOLO FUNCIONA EN ESPACIOS EQUIPROBABLES.

Example 2.1.5 (Tenemos una bolsa con 6 bolas numeradas del 1 al 6 sacamos 3 bllas a la vez y anotamos lo que salió)

$$\Omega_4 = \{(x,y,z) / x,y,z \in \{1,2,3,4,5,6\}\}$$

$$F_4 = 2^{\Omega_4}$$

$$P_4(\{(x,y,z) / x,y,z \in \{1,2,3,4,5,6\}\}) = 1/20$$

$$P_4(\{(x,y,z) / x,y,z \in \{1,2,3,4,5,6\} \text{ y } x,y,z \text{ son distintos}\}) = 6/20$$

$$P_4(\{(x,y,z) / x,y,z \in \{1,2,3,4,5,6\} \text{ y } x,y,z \text{ son iguales}\}) = 1/20$$