



Estadística Descriptiva

Índice

1	Introducción	3
2	Conceptos Básicos	3
3	Variables Estadística.....	4
4	Escalas de Medición	4
5	Organización de la Información.....	5
6	Tablas de Frecuencia.....	8
7	Ejemplos	9
8	Resumen	14
9	Referencias Bibliográficas	14

Red SUMMA ©

Objetivos

- Objetivo: Conocer y comprender los fundamentos básicos en los que se compone la estadística.

1 Introducción

Esta unidad de aprendizaje está encaminada a orientar al alumno a través de los conceptos básicos y las herramientas descriptivas que ofrece la estadística como disciplina ligada a las ciencias básicas, es de notar que los conceptos aquí consignados permitirán al estudiante obtener el conocimiento necesario para la comprensión de fenómenos presentes en cualquier tipo de estudios científicos, esto a través de ejemplos léxico-gráficos, pero sin hacer énfasis en los detalles y rigurosidad matemáticas propias de la disciplina.

2 Conceptos Básicos

Estadística: Es la disciplina que se ocupa de los **métodos y procedimientos para recoger, clasificar, resumir, hallar regularidades y analizar los datos**, siempre y cuando la variabilidad e incertidumbre sea una causa intrínseca de los mismos. Así como de realizar inferencias a partir de ellos, con la finalidad de ayudar a la toma de decisiones y en su caso formular predicciones.

Estadística Descriptiva: Comprende los **procesos de consolidación, resumen y descripción de los datos recopilados**. Tablas, gráficos o índices que permiten un análisis referido, exclusivamente, a los datos coleccionados. No se generaliza ni se extrapola a la totalidad de los elementos.

Estadística Inferencial: Incluye procedimientos que permiten la **extrapolación y generalización sobre características que tipifican a todos los elementos de la población**. Puede decirse que, es el proceso de hacer afirmaciones o predicciones sobre toda la población, tomando como base, sólo a la información recolectada a través de una muestra.

Población: Conjunto de elementos que son de interés en un estudio (poseen características comunes acerca de los cuales se desea tener información). Usualmente a dichos elementos se les **denomina individuos, observaciones o mediciones**. La población puede ser **finita o infinita**.

“**Población:** individuos, observaciones, mediciones.

Muestra: parte de la población.

Censo: información de la población

Muestreo: selección elementos de una población”

Muestra: Es una **parte de la población**. Para estudios estadísticos, se requieren muestras que nos den información real de la población.

Censo: Información de toda la población.

Muestreo: Proceso mediante el cual **se seleccionan los elementos de una población**.

3 Variables Estadística

Son **características de los elementos poblacionales que se atribuyen al azar**. Se clasifican en:

Cualitativas: Aquellas que **describen cualidades de los elementos**. **Ejemplo:** Tipo de sangre, cuyas modalidades o categorías son: O, A, B, AB. **Otros ejemplos son:** Estado civil, documento de identificación, filiación política, tipo de religión, etc.

Cuantitativas: Aquellas que generalmente **resultan de un proceso de medición**. Pueden ser:

- **Discretas:** Resultan de conteos, y el **resultado es un número entero**. **Ejemplo:** Número de hermanos, cuyas modalidades o categorías son: 0,1,..., N. **Otros ejemplos son:** Número de pacientes que llegan a un centro de salud en un intervalo de tiempo dado, número de sillas en un salón de clases, número de horas que un estudiante dedica semanalmente a sus asignaturas.
- **Continuas:** El resultado es un **subconjunto de los números reales**. **Ejemplo:** Tiempo de espera en una parada de bus, velocidad de un vehículo en una autopista, ingreso económico del jefe de hogar en una familia.

4 Escalas de Medición

La medición hace referencia a la **asignación de números a las características objeto de estudio**.

Escala nominal: Es la **más baja de las escalas** de medición. Identifica las **categorías de la variable de interés**, y se pueden diferenciar las categorías una de la otra haciendo **uso de dígitos**. **Ejemplo:** Estado civil (soltero, casado, viudo,

“Cualitativas: describen cualidades de los elementos.”

“Cuantitativas: resultan de un proceso de medición”

“Tipos de escala: nominal, ordinal, de intervalos y, de razón”

unión libre, separado), cédula de ciudadanía, género musical, tipo de sangre, estado del paciente.

Escala ordinal: Identifica las **categorías de la variable y pueden ser clasificadas por grados de acuerdo a algún criterio**. La **función** de los dígitos asignados a datos ordinales es la de **ordenar**. **Ejemplo:** Grado de escolaridad (ninguno, primaria, secundaria, profesional, postgrado), rangos militares, grados de desnutrición, tipo de quemadura.

Escala de intervalos: La escala de intervalos es una escala más especializada que la nominal y la ordinal, en el sentido de que no solo es posible **ordenar las mediciones**, sino que también se **conoce la distancia entre las observaciones** cualesquiera. Aquí **no** hay un **punto cero único**. **Ejemplo:** La escala en la que se mide la temperatura. No es posible decir que 30° es doble de frío que 60° , ya que depende de la escala (grados

Celsius o Fahrenheit). **Otros ejemplos son:** Pérdida auditiva en decibeles y coeficiente intelectual en puntaje.

Escala de razón: Es el nivel más alto de las escalas de mediciones y se **caracteriza** por el hecho de que se puede **determinar tanto la igualdad de razones como la de intervalos**. Existe un **punto cero único**. **Por ejemplo:** Altura, peso, longitud, velocidad, área, volumen.

5 Organización de la Información

Existen **dos formas** básicas para la representación de la información recolectada, a través de tablas o cuadros estadísticos, y a través de un gráfico.

Una **tabla o cuadro estadístico** es una **representación en forma ordenada de la variación de un fenómeno, clasificado bajo uno o más variables**. Puede ser **simple** (clasificación bajo **una variable**) o **compuesto** (clasificación bajo **dos o más variables**).

Términos relacionados para tablas estadísticas:

- **Frecuencia absoluta (n_i):** Sea X una variable estadística cuyos valores son X_1, X_2, \dots, X_k ; de una muestra de tamaño n , con ($k \leq n$). La frecuencia absoluta corresponde al **número de veces que se repite cada valor de la variable**.

“Términos: frecuencia absoluta, relativa, absoluta acumulada, relativa acumulada”

Propiedades:

1. $0 \leq n_i \leq n$, para $i = 1, \dots, k$
2. $\sum_{i=1}^k n_i = n_1 + n_2 + \dots + n_k = n$

- **Frecuencia relativa (h_i):** Sea X una variable estadística cuyos valores son X_1, X_2, \dots, X_k ; de una muestra de tamaño n , con

($k \leq n$). La frecuencia relativa corresponde al **cociente entre la frecuencia absoluta y el tamaño de muestra.**

$$h_i = \frac{n_i}{n}$$

Propiedades:

1. $0 \leq h_i \leq 1$, para $i = 1, \dots, k$
2. $\sum_{i=1}^k h_i = h_1 + h_2 + \dots + h_k = 1$

- **Frecuencia absoluta acumulada (N_i):** Es la **suma sucesiva de frecuencias absolutas.**

Propiedades:

1. $N_i = n_1 + n_2 + \dots + n_i = \sum_{j=1}^i n_j$, para $i = 1, \dots, k$
2. $N_1 = n_1$
3. $n_i = N_i - N_{i-1}$

- **Frecuencia Relativa acumulada (H_i):** Es la **suma sucesiva de frecuencias relativas.**

Propiedades:

1. $H_i = h_1 + h_2 + \dots + h_i = \sum_{j=1}^i h_j$, para $i = 1, \dots, k$
2. $H_1 = h_1$
3. $h_i = H_i - H_{i-1}$

Para algunos tipos de tablas, es necesario definir algunos términos adicionales.

- **Clases:** Intervalos de números reales que tienen las siguientes características:
 - Son disjuntos, es decir, **no comparten** elementos.
 - Son exhaustivas, es decir, **no dejan** elementos por **fuera**.
 - El elemento de **menor magnitud** cae en la **primera clase** y el de **mayor magnitud** cae en la **última clase**.

Observaciones:

- Se recomienda construir entre 5 y 25 clases según el número de datos.
- Las clases que se construyen deben ser reales, esto es, sea $[L_i; L_{i+1}]$ la clase i -ésima ($i = 1, \dots, k$), donde L_i es el **límite inferior** de la clase i , y L_{i+1} es el **límite superior** de la clase i .

De acuerdo con la información contenida en la muestra, los límites reales de clase toman la siguiente forma.

Información	L_i	L_{i+1}
Enteros	-0.5	+0.5
Décimas	-0.05	+0.05
Centésimas	-0.05	+0.05
⋮	⋮	⋮

Tabla 1: Límites reales

- **Marca de clase (X_i):** Es el **punto medio de cada clase**.

$$X_i = \frac{L_i + L_{i+1}}{2}$$

Comúnmente representa a todos los elementos en la respectiva clase.

- **Número de clases (K):** Es un indicador para obtener el número de clases. Regla de *Stegers*.

$$K = 1 + 3.33 \log n$$

Donde n : es el tamaño de la muestra. Otro indicador muy usado es:

$$k \approx \sqrt{n}$$

- **Rango (R): Diferencia entre la observación de mayor magnitud y la de menor magnitud.**

$$R = X_{(max)} - X_{(min)}$$

Se dice que el rango es la **longitud total de los datos**.

- **Amplitud de Clase (A_i): Es la diferencia entre el límite superior y el límite inferior de clase.**

$$A_i = L_{i+1} - L_i ; i = 1, 2, \dots, k$$

Para efectos de interpretación, se recomienda que todas las clases tengan la misma amplitud. Esto se obtiene por medio de la siguiente expresión.

$$A_i = \frac{R}{K} ; i = 1, 2, \dots, k$$

6 Tablas de Frecuencia

La forma general de una tabla de frecuencias se presenta en el siguiente cuadro.

Modalidad	Frec. Abs.	Frec. Rel.	Frec. Abs. Acu.	Frec. Rel. Acu.
<i>C</i>	<i>n_i</i>	<i>h_i</i>	<i>N_i</i>	<i>H_i</i>
<i>c₁</i>	<i>n₁</i>	$h_1 = \frac{n_1}{n}$	<i>N₁ = n₁</i>	<i>H₁ = h₁</i>
...
<i>c_j</i>	<i>n_j</i>	$h_j = \frac{n_j}{n}$	<i>N_j = n₁ + ... + n_j</i>	$H_j = \frac{N_j}{n} = h_1 + \dots + h_j$
...
<i>c_k</i>	<i>n_k</i>	$h_k = \frac{n_k}{n}$	<i>N_k = n_k</i>	<i>H_k = 1</i>

Tabla 2: Tabla de frecuencias explícita.

Modalidad	Frec. Abs.	Frec. Rel.	Frec. Abs. Acu.	Frec. Rel. Acu.
C	n_i	h_i	N_i	H_i
c_1	n_1	$h_1 = \frac{n_1}{n}$	N_1	H_1
...
c_j	n_j	$h_j = \frac{n_j}{n}$	N_j	H_j
...
c_k	n_k	$h_k = \frac{n_k}{n}$	N_k	H_k

Tabla 3: Tabla de frecuencias.

7 Ejemplos

Ejemplo 1:

El número de vehículos que llegan a un taller automotor en un día dado, es una variable de tipo estadístico que se observó durante un período de 25 días y se obtuvieron los siguientes datos:

8	6	7	9	8
7	8	10	4	10
8	7	9	8	7
6	6	10	7	8
5	5	8	10	11

Se puede describir la variable X como: X : número de vehículos que llegan al taller. Tipo: **cuantitativa discreta**. Escala: **razón**.

La distribución de frecuencias para el número de vehículos que llegan a un taller automotor en un día dado es:

Valor de X_i	Frec. Abs. n_i	Frec. Rel. h_i	Frec. Abs. Acu. N_i	Frec. Rel. Acu. H_i
4	1	0.04	1	0.04
5	2	0.08	3	0.12
6	3	0.12	6	0.24
7	5	0.20	11	0.44
8	7	0.28	18	0.72
9	2	0.08	20	0.80
10	4	0.16	24	0.96
11	1	0.04	25	1
Total	25	1	-	-

Tabla 4: Ejemplo 1

Ejemplo 2:

Una encuesta realizada a 30 fumadores para determinar el número de cigarrillos que encienden (*fuman*) en un día corriente arrojó los siguientes resultados:

3	7	5	10	8	4
5	8	10	8	8	4
5	3	10	5	7	10
8	5	5	12	8	4
4	3	5	8	12	10

Se puede describir la variable X como: X : Número de cigarros que enciende un fumador. Tipo: Cuantitativa discreta. Escala: razón.

La distribución de frecuencias para el número de cigarrillos que enciende un fumador en un día dado es:

Valor de X_i	Frec. Abs. n_i	Frec. Rel. h_i	Frec. Abs. Acu. N_i	Frec. Rel. Acu. H_i
3	3	0.100	3	0.100
4	4	0.133	7	0.233
5	7	0.233	14	0.467
7	2	0.067	16	0.533
8	7	0.233	23	0.767
10	5	0.167	28	0.933
12	2	0.067	30	1
Total	30	1	-	-

Tabla 5: Ejemplo 2

Ejemplo 3:

A continuación se presentan los datos sobre el octanaje del combustible para motores de varias marcas de gasolina:

88.5	89.8	89.9	90.6	93.4	90.7	90.1
94.7	91.6	<u>98.8</u>	92.2	96.1	88.6	89.3
84.3	90.3	83.3	87.7	89.6	88.3	91.1
90.1	90.0	90.4	91.1	90.4	94.2	<u>83.4</u>
89.0	91.5	91.2	86.7	91.6	85.3	93.2

Se puede describir la variable X como: X : Octanaje del combustible. Tipo: Cuantitativa continua. Escala: razón.

Datos necesarios:

Número de clases: $K = 1 + 3.33 \log 35 = 6.095 \approx 6$

Rango: $R = X_{(max)} - X_{(min)} = 98.8 - 83.4 = 15.4$

$$\text{Amplitud: } A = \frac{R}{K} = \frac{15.4}{6} = 2.5666 \approx 2.6$$

La distribución de frecuencias para el Octanaje del combustible es:

L_i	L_s	Frec. Abs. n_i	Frec. Rel. h_i	Frec. Abs. Acu. N_i	Frec. Rel. Acu. H_i	Marca de clase X_i
83.4	86	5	0.143	5	0.143	84.7
	88.6	4		9		
86	91.2	15	0.114	24	0.257	87.3
	93.8	7	0.429	31	0.685	89.9
88.6	96.4	3		34	0.885	
	99	1	0.200	35	0.971	92.5
91.2			0.086		1	95.1
93.8			0.029			97.7
96.4						
Total		35	1	-	-	

Tabla 6: Ejemplo 3

Ejemplo 4 Resuelto:

Se ha preguntado a 40 personas el número de personas que forman el hogar familiar obteniéndose los siguientes resultados:

Número de personas en el hogar	2	3	4	5	6	7
Frecuencia	4	11	11	6	6	2

- a. Calcula la media, la mediana, la moda y la desviación típica.

Para calcular la media, moda, mediana y varianza se tendrá que realizar la siguiente tabla:

Tabla para calcular la media y desviación típica				
personas x_i	frecuencia f_i	F_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
2	4	4	8	16
3	11	15	33	99
4	11	26	44	176
5	6	32	30	150
6	6	38	36	216
7	2	40	14	98
Σ	40		165	755

Por lo que, ya tenemos todo lo suficiente para poder hacer las fórmulas.

Media: $\bar{x} = \frac{\sum x_i \cdot f_i}{N} = \frac{165}{40} = 4,125$

Moda M_o : hay dos valores con la mayor frecuencia 11, son los valores 3 y 4 es bimodal.

Mediana: $M = \frac{40}{2} = 20$ columna F_i el valor >20 , es el 26 se corresponde con 4. $M = 4$.

Desviación típica: $\sigma = \sqrt{\frac{\sum (x_i)^2 \cdot f_i}{N} - (\bar{x})^2} = \sqrt{\frac{755}{40} - (4,125)^2} = 1,364$

8 Resumen

- Estadística: Es la disciplina que se ocupa de los métodos y procedimientos para recoger, clasificar, resumir, hallar regularidades y analizar los datos, siempre y cuando la variabilidad e incertidumbre sea una causa intrínseca de los mismos.
- Cualitativas: Aquellas que describen cualidades de los elementos.
- Cuantitativas: Aquellas que generalmente resultan de un proceso de medición.
- Tipos de escala: Nominal, ordinal, de intervalos y, de razón.

9 Referencias Bibliográficas

- Montgomery, D. C y Runger, R. (2008). Probabilidad y Estadística Aplicadas a la Ingeniería. 2da. Edición. Limusa Wiley. Mexico.
- Rincón. L. (2010). Curso Elemental de Probabilidad y Estadística. 1ra Edición. Circuito Exterior de CU. México D.F.
- Walpole, R. E., Myers, R. H., Myers, S. L. y Ye, K. (2007). Probabilidad y Estadística para Ingeniería y Ciencias. 8va. Edición. Pearson. Mexico
- I. Espejo Miranda, F. Fernández Palacín, M. A. López Sánchez, M. Muñoz Márquez, A. M. Rodríguez Chía, A. Sánchez Navas, C. Valero Franco. (2006). Estadística Descriptiva y Probabilidad. 3ra. Edición. Servicio de Publicaciones de la Universidad de Cádiz.
- Devore, J. (2008). Probabilidad y Estadística para Ingeniería y Ciencias. 7ma. Edición. Cengage Learning Editores. México
- Canavos, G. Probabilidad y Estadística. Métodos y Aplicaciones.
- Daniel, W. (2009). Bioestadística. Base para el análisis de las ciencias de la salud. Cuarta edición. Limusa Wiley. México.
- Mendenhall, W., Wackerly, D. y Scheaffer, R. (1994). Estadística Matemática con Aplicaciones. Grupo Editorial Iberoamericana. México.