



[Cod: CM274-A,B Curso: Introducción a la Estadística y Probabilidad]
[Profesores: Cesar Lara A., Fernando Zamudio P.]

PRIMERA PRÁCTICA DIRIGIDA

1. Intentando determinar la eficacia de un medicamento nuevo, un investigador médico ha comenzado con el contraste del medicamento frente a un placebo. Para asegurarse de que los dos grupos de pacientes voluntarios: aquellos que reciben el medicamento y los que reciben el placebo son lo más parecido posible, el investigador ha decidido no basarse en el azar sino que, por el contrario, ha analizado detalladamente a los voluntarios y luego él mismo ha elegido los grupos. ¿Es aconsejable este procedimiento? ¿Por qué sí? o ¿por que no?.
2. Explica por qué es importante que un investigador que intenta estudiar la utilidad de un nuevo medicamento no conozca qué pacientes son tratados con el medicamento y cuáles están recibiendo el placebo.
3. Los siguientes datos representan los tamaños de 30 familias que residen en una pequeña ciudad de Guatemala.

5, 13, 9, 12, 7, 4, 8, 6, 6, 10, 7, 11, 10, 8, 15, 8, 6

9, 12, 10, 7, 11, 10, 8, 12, 9, 7, 10, 7, 8

- (a) Construye una tabla de frecuencias para estos datos.
 - (b) Representa estos datos mediante un gráfico de líneas.
 - (c) Representa graficamente los datos mediante un polígono de frecuencia.
 - (d) Dibuja una tabla de frecuencia relativa y dibuja después un gráfico de barras.
4. En 1662 el comerciante inglés John Graunt, publicó un libro titulado *Observaciones naturales y políticas hechas a partir de los datos de la mortalidad*. En la siguiente tabla se incluye el número de total de fallecimientos en Inglaterra y el número de ellos causados por la peste en cinco diferentes años de epidemia, sacada de ese libro.

Grant utilizó los datos de la mortalidad de Londres para estimar la población de la ciudad. Por ejemplo en 1600 en la población de Londres se descubrió que se producían 3 defunciones por cada 88 habitantes. Dividiendo entre 3, observó que en media había una muerte por cada $88/3$ habitantes. Puesto que las cifras de mortalidad de Londres recogían 13200, Graunt estimó que la población de Londres era aproximadamente

$$13\ 200 \cdot \frac{88}{3} = 387\ 200$$

Graunt también usó este indicador para determinar el número de habitantes de Inglaterra, además usó las cifras de mortalidad de Londres, así como algunas inteligentes intuiciones sobre qué enfermedades mataban a qué personas y a qué edades, para inferir las tasas de defunción por edad. Como indica la siguiente tabla:

- (a) Comenta el método de Graunt para estimar la población de Londres. ¿Qué hipótesis implícita se está asumiendo?.

Año	Entierros	Muerte por peste
1592	25 886	11 503
1593	17 844	10 662
1603	37 294	30 561
1625	51 758	35 417
1636	23 359	10 400

Edad de muerte	Muerte por cada 100 nacimientos
0-6	36
6- 16	24
16-26	15
26-36	9
36-46	6
46-56	4
56-66	3
66-76	2
≥ 76	1

- (b) Las cifras de mortalidad de Londres registraban 12246 muertes en 1658. Suponiendo que una encuesta sobre las parroquias de Londres mostró que, grosso modo, un 2% de la población había fallecido en dicho año, utilice el método de Graunt para estimar la población de Londres en 1658.
- (c) Si eres un vendedor de planes de pensiones en 1662, año en el que se publicó el libro de Graunt. Explica cómo habría usado los datos sobre las edades en la que se producían los fallecimientos.
5. Los siguientes conjuntos de datos representan las puntuaciones obtenidas por 40 estudiantes de un curso de probabilidad en un test IQ en una determinada escuela matemática:

144, 122, 103, 118, 99, 105, 134, 125, 117, 106
109, 104, 111, 127, 133, 111, 117, 103, 120, 98
100, 130, 141, 119, 128, 106, 109, 115, 113, 121
100, 130, 125, 117, 119, 113, 104, 108, 110, 112

- (a) Presenta este conjunto de datos en un histograma de frecuencias.
- (b) ¿Qué intervalo de clase contiene el mayor número de valores de datos?.
- (c) ¿ Existe, el mismo número de datos en cada uno de los intervalos de clase?.
- (d) ¿ El histograma parece aproximadamente simétrico?
6. Un conjunto de datos puntuales se dividió en 8 clases, todas de tamaño 3 (en las unidades de los datos). Después se determinaron las frecuencias de cada clase y se contruyó una tabla de frecuencias. Sin embargo, ciertas entradas de esta tabla se perdieron. Supongamos que la parte de la tabla de frecuencia que se conservó es la siguiente
- Completa los valores perdidos de la tabla y dibuja un histograma de frecuencias relativas.
7. Para determinar la relación entre la temperatura que hay al mediodía (medida de grados Celsius) y el número de piezas defectuosas producidas dicho día, una compañía registro los datos siguientes correspondientes a 22 días laborales.

Intervalo de clase	Frecuencia	Frecuencia relativa
		0.05
	14	
	18	
15-18	38	
		0.10
	42	
	11	

Temperatura	Número de piezas defectuosas	Temperatura	Número de piezas defectuosas
24.2	25	24.8	23
22.7	31	20.6	20
30.5	36	25.1	25
28.6	33	21.4	25
25.5	19	23.7	23
32.0	24	23.9	27
28.6	27	25.2	30
26.5	25	27.4	33
25.3	16	28.3	32
26.0	14	28,8	35
24.4	22	26.6	24

(a) Dibuja un diagrama de dispersión.

(b) ¿Qué se puede concluir a partir del diagrama anterior?

(c) Si la temperatura al mediodía fuese de $24^{\circ}C$. ¿qué se podría conjeturar sobre el número de piezas defectuosas que se vayan a producir al día siguiente?.

8. Encuentre el percentil muestral de orden 90% del siguiente conjunto de datos,

75, 33, 51, 21, 46, 98, 103, 88, 35, 22, 29, 73, 37, 101, 121, 144, 133, 52, 54, 63, 21, 7

9. Los cuartiles de un extenso conjunto de datos son los siguientes:

Primer cuartil = 35

Segundo cuartil = 47

Tercer cuartil = 66

- Indica un intervalo que contenga aproximadamente un 50% de los datos.
- Determina un valor que aproximadamente sea mayor que un 50% de los datos.
- Determina un valor para el que aproximadamente un 25% de los datos sean mayores que él.

10. Suponga que la media muestral de un conjunto de 10 datos puntuales es $\bar{x} = 20$.

(a) Si se descubre que se ha leído incorrectamente un dato con valor 15 y que se le ha dado el valor 13. ¿Cuál será el valor revisado de la media muestral?.

- (b) Si existiera un dato adicional con valor 22, ¿aumentaría o disminuiría el valor de \bar{x} ?
- (c) Con los datos originales. ¿Cuál sería el nuevo valor de \bar{x} del apartado anterior.
11. Supongamos que se disponen de dos muestras distintas, de tamaños n_1 y n_2 . Si la media muestral de la primera muestra es \bar{x}_1 y la segunda muestra es \bar{x}_2 , ¿cuál es la media de la muestra conjunta, de tamaño $n_1 + n_2$.
12. La mediana de un conjunto de datos simétricos es igual a 40 y su tercer cuartil es igual a 55. ¿Cuál es el valor del primer cuartil?
13. Suponga que deseas descubrir el salario del vicepresidente de un banco, al que acabas de conocer. Si pretendes tener la mayor probabilidad de acertar a menos de 1000 soles, ¿le gustaría conocer la media muestral, la mediana muestral o la moda muestral de los salarios de los vicepresidentes de bancos?
14. Varios corredores utilizan una pista de atletismo de un cuarto de milla de longitud. En una muestra de 17 de corredores, 1 corrió 2 vueltas, 4 corrieron 4 vueltas, 5 corrieron 6 vueltas, 6 corrieron 8 vueltas y 1 corrió 12 vueltas.
- ¿Cuál es la moda muestral del número de vueltas que han hecho estos corredores?
 - ¿Cuál es la moda muestral de las distancias en millas recorridas por los corredores?
15. Un individuo que necesitaba asegurar su coche, preguntó cuales eran las cuotas para idénticas coberturas en 10 compañías de seguro. Obtuvo los siguientes valores (correspondientes a las cuotas anuales, en dolares)
- 720, 880, 630, 590, 1140, 908, 677, ,720, 1260, 800

Encuentre :

- la media muestral
 - la mediana muestral
 - la desviación típica muestral
16. Si s es la desviación típica muestral de los datos $x_i, i = 1, \dots, n$. ¿cuál es la desviación típica muestral de $ax_i + b, i = 1, \dots, n$?. En este problema a y b son constantes dadas.
17. En la siguiente situaciones, ¿cuál de los estadísticos media muestral o mediana muestral piensas que es más informativo?
- (a) Para analizar si se debe cerrar una línea de autobús entre Lima y Chiclayo, un ejecutivo ha recopilado el número de viajeros en una muestra de días.
- (b) Para comparar a los estudiantes universitarios actuales con los de los años anteriores, se consultan muestras de las calificaciones obtenidas en los exámenes de acceso a la universidad durante varios años.
- (c) El abogado defensor de un proceso judicial con jurado popular está analizando las puntuaciones de un test de inteligencia (IQ) obtenida por los miembros del jurado.
- (d) Tu has comprado su casa hace 6 años en una pequeña comunidad por un precio de 105000 dólares, que coincidía con el precio medio y mediano de todas las casas que se vendieron aquel año en dicha comunidad. Sin embargo, en los dos últimos años, se han construido varias casas nuevas mucho más caras que las anteriores. Para obtener una idea de valor actual de su casa, decides analizar los precios de venta de las casas vendidas recientemente en su comunidad.

18. La siguiente tabla muestra el número de médicos y dentistas que había en Japón en los años pares comprendidos entre 1984 y 2000.

-	Médicos	Dentistas
1984	173 452	61 283
1986	183 129	64 904
1988	193 682	68 692
1990	203 797	72 087
1992	211 498	75 628
1994	220 853	79 091
1996	230 297	83 403
1998	236 933	85 669
2000	243 201	88 410

- Determina la varianza muestral del número de médicos en los años citados.
- Determina la varianza muestral del número de dentistas en dichos años.

19. Considere los dos siguientes conjuntos de datos

$A : 4, 5, 0, 5, 1, 5, 0, 10, 5, 2$

$B : 0, 4, 0, 1, 9, 0, 10, 9, 5$

- Determine el rango de cada conjunto de datos.
- Calcule la desviación típica muestral de cada conjunto de datos.
- Determine el rango intercuartílico de cada conjunto de datos.

20. Explica por qué el coeficiente de correlación muestral de los pares de datos

$(121, 360), (242, 362), (363, 364)$

es el mismo que el de los pares

$(1, 0), (2, 2), (3, 4)$

el cual, a su vez, coincide con el de los pares

$(1, 0), (2, 1), (3, 2)$

Además, calcula el coeficiente de correlación muestral para los pares de datos anteriores.

Los Profesores
UNI, August 29, 2018