



Medidas de Resumen PREGUNTAS

Jorge Galbiati Riesco

1. ¿Qué ventajas tiene la mediana sobre el promedio, como medida de centro? ¿Qué ventajas tiene el promedio sobre la mediana?
2. ¿Cuándo se prefiere usar la mediana en lugar del promedio, como medida de centro? ¿Cuándo se prefiere usar el promedio en lugar de la mediana)
3. ¿Qué tiene de malo el rango, como medida de dispersión?
4. ¿En qué unidad se expresan las siguientes medidas de dispersión, en relación a la unidad de medida de los datos originales?
Rango, desviación media, varianza, desviación estándar, desviación intercuartil.
5. Se tiene un conjunto de datos, y resulta que la mediana está a la izquierda del promedio. ¿Qué características presentan esos datos?
6. Se tiene un conjunto de datos. El valor más grande es 23, pero por un error de digitación, se registró como 203. ¿En qué afecta este error a la desviación intercuartil? ¿En qué afecta a la mediana?
7. Un conjunto de 40 datos presenta las siguientes medidas descriptivas:
El rango es 100, la desviación standard es 28.4, el rango inter cuartil es 19.3.
Explique qué forma tiene el conjunto de datos (puede ayudarse con un dibujo).
8. ¿Qué característica fundamental debe cumplir un plan de muestreo para que la muestra obtenida sea representativa de la población?
9. Describa los tipos de escalas de medida que conoce.
10. Un conjunto de 50 datos presenta las siguientes medidas descriptivas:
El cuartil 1 es 16, la mediama es 25, el cuartil 3 es 43, el promedio es 32.5.
Explique qué forma tiene el conjunto de datos (puede ayudarse con un dibujo).

11. Un conjunto de 30 datos presenta las siguientes medidas descriptivas:

El promedio es 23.6, el promedio recortado al 10% es 16.9.

Explique qué característica tiene el conjunto de datos.

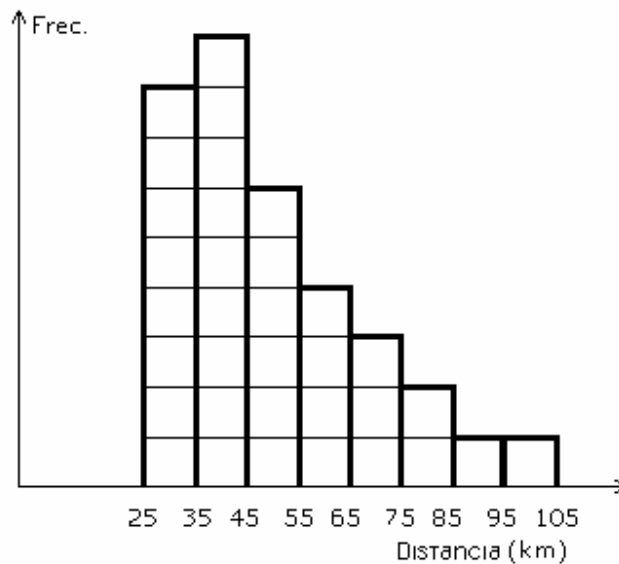
12. Explique qué miden los cuartiles.

13. Un conjunto de 40 datos contiene cinco valores mucho más grandes que todo el resto.

Indique cuáles de estas medidas de centro estarán afectadas por los cinco valores y cuáles no:

Promedio, mediana, promedio recortado al 5%, promedio recortado al 20%

14. En el siguiente histograma, ubique, en forma aproximada, el promedio y la mediana.



15. Explique las siguientes afirmaciones:

El promedio de notas de Juan está sobre el percentil 90.

La desviación intercuartil de las edades de los consumidores de este producto es 25 años.

16. Un conjunto de 40 datos presenta las siguientes medidas descriptivas:

El rango es 100, la desviación estándar es 28.4, el rango intercuartil es 19.3.

Explique qué forma tiene el conjunto de datos (puede ayudarse con un dibujo). Es decir, 3 personas respondieron "si" a la pregunta 10 y "si" a la pregunta 17, etc. Más allá de los resultados numéricos, ¿qué se puede decir, en general, respecto de la relación entre estas dos preguntas?

17. Discuta acerca de la apropiada elección de unidades muestrales y marco de referencia en las siguientes investigaciones:

- a) Un departamento de justicia de un estado, desea estimar la duración promedio de la detención preventiva de personas con cargos de ofensas criminales mayores.
- b) La división de marketing de una compañía farmacéutica desea estimar el porcentaje de hospitales en el país que usan una marca de solución desinfectante para esterilizar equipo quirúrgico.
- c) Un cuerpo gubernativo de una universidad esta interesado en estudiar opiniones de los estudiantes, en relación a un cambio en el calendario académico propuesto por una facultad.
- d) Un magistrado elegido desea determinar como un distrito electoral se siente sobre un trozo de legislación anti-obesidad propuesto.
- e) Una agencia estatal de salud pública esta desarrollando un proyecto para inquirir acerca de la calidad de la atención de salud en el estado. Parte del proyecto, esta destinado para estimar el gasto promedio anual en el cuidado dental de las familias.

18 Identifique las principales fuentes de sesgo en cada uno de los siguientes experimentos:

- a) Se conduce una encuesta para estudiar en que medida se utilizan los alimentos preparados en hogares de una comunidad. Se selecciona una muestra aleatoria de hogares y los datos se recogen por entrevistas telefónicas durante horas de trabajo (8 a.m. a 5 p.m.). Las no respuestas son ignoradas.
- b) Para estudiar la participación de residentes en una ciudad en particular en actividades deportivas al aire libre, se recogen datos por medio de entrevistadores que visitan los hogares muestreados, usualmente durante las tardes de fines de semana.
- c) Una agencia decide usar un conveniente y económico método de muestreo para conducir una encuesta de opinión pública sobre las marcas de cecinas. Entrevistadores ubicados en supermercados durante las tardes de los días de semana, recogen los datos de los compradores cada tres minutos a medida que entran al local.

19. Describa los siguientes tipos de muestreo:

- a) Muestreo aleatorio simple.
- b) Muestreo estratificado.
- c) Muestreo por conglomerados.
- d) Muestreo sistemático.

Dé un ejemplo de cada uno.

- 20.** a) ¿Cuál es la diferencia entre muestreo probabilística y muestreo no- probabilístico?
- b) De algunos ejemplos de este último.

21. De cada una de las 13 regiones del país se escogió, proporcionalmente al tamaño de cada región, un número de madres jóvenes (con edad menor o igual a 25 años) que en total sumaron

2000. Entre otras características, se observó el número de hijos por madre, las edades de cada uno de ellos y el ingreso total del grupo familiar.

- a) Defina la población en estudio.
- b) Defina y clasifique la(s) variable(s) en estudio.
- c) Describa la técnica de muestreo utilizada.

EJERCICIOS

1. Los siguientes son los tiempos que demoró una cierta reacción química en producirse, en 32 repeticiones, que serán considerados como nuestra población:

25	20	35	21	22	22	24	25	30	28
24	20	20	25	20	19	25	23	20	24
28	24	22	28	26	23	25	22	27	25
23	25								

- a) Utilizando números aleatorios, obtenga una muestra aleatoria simple de tamaño 10, de esta población.
- b) Calcule el promedio de los valores de la muestra y compárelo con el valor del promedio de todos los datos.
- c) Repita (a) y (b) varias veces; represente los valores en un gráfico apropiado y comente.

2. En una máquina automática de enlatado y etiquetado de productos del mar, se producen detenciones por latas que ingresan en mala posición y traban el sistema. Se registraron los números de detenciones que ocurrieron durante 70 días consecutivos, obteniéndose los resultados que se muestran más abajo. Encuentre :

- a) La media,
- b) la mediana,
- c) los cuartiles.
- d) el rango de la tabla. e) la varianza y la desviación estándar.

0	0	2	0	0	0	3	3	0	0
1	8	5	0	0	4	3	0	6	2
3	1	1	0	1	0	1	1	0	0
2	2	0	0	0	1	2	1	2	0
0	1	6	4	3	3	1	2	4	0
0	3	1	2	0	0	0	0	0	1
1	0	2	0	2	2	4	0	2	2

3. Haga un bosquejo de la forma que tendrá cada uno de los polígonos de frecuencia correspondientes a los siguientes casos:

a) $M_d=30$ $M_o=10$ $\bar{X} = 50$

b) $\bar{X} = 30$ $S=9$

c) $M_d=20$ $M_o=10$ $\bar{X} = 25$ $Q_2=30$

4. El departamento de ingeniería de una empresa realizó un control de tiempo de la elaboración de una pieza de dos secciones A y B. La tabla que se muestra a continuación entrega el número de operarios que fabrican la pieza y los tiempos empleados, en segundos :

Tiempo	40	41	42	43	44	45	46	47	48	49	50
N° Empleados											
Sección A	1	4	3	6	12	9	9	3	1	1	1
N° Empleados											
Sección B	0	0	3	2	6	10	12	7	4	2	3

- a) En base a los valores de las medias, decida cual sección tiene mejor rendimiento.
 b) Haga una representación gráfica que permita comparar los rendimientos.

5. Para estudiar la influencia de la vitamina B₁₂ en el aumento diario del peso de cerdos, se dispuso el siguiente experimento. A un grupo de 36 cerdos, cuyo aumento promedio era de 1,34 Kg. se les dio una dieta rica en la vitamina, se consiguió con esto que los aumentos de peso fuesen incrementados en un 40%. Otros 14 cerdos, cuyo aumento promedio diario era de 1,50 Kg. se usó como grupo control (se les dio dieta normal) y su incremento no se modificó. A un tercer grupo de 24 cerdos se les dio una dieta pobre en vitamina B₁₂, se observó que estos cerdos disminuyeron cada uno 0.3 Kg. su incremento de peso. El aumento promedio diario de todos los cerdos antes de ser sometidos a tratamiento era 1,7 Kg. ,

- a) ¿Cuál era el aumento de peso de los cerdos del tercer grupo antes del tratamiento?
 b) ¿Cuál es el aumento de peso promedio de todos los cerdos después del tratamiento?
 c) Si las desviaciones estándares de los tres grupos eran 0.31, 0.18, y 0.53 Kg. respectivamente, antes de realizar el experimento. Calcule los coeficientes de variación para cada grupo antes y después de someterlos a los tratamientos.

6. Se conocen los puntajes que obtuvo un grupo de postulantes en una prueba de aptitudes, pero no se conocen sus identificaciones. Uno de ellos, Andrés, no conoce su puntaje, pero sabe que es mayor que el promedio y menor que el tercer cuartil. Los puntajes son:

851 344 591 513 744 526 522 684
 491 618 750 739 527 765 590

- a) ¿Puede usted obtener los posibles puntajes que obtuvo Andrés?
 b) De estos valores calculados en (a), el verdadero puntaje de Andrés es aquel que al calcular la desviación standard de los 14 restantes, produce el mayor valor. ¿Cual es su puntaje?

7. Un grupo de 100 atletas viaja en 2 aviones. El primero lleva 40 atletas y el segundo los restantes. Se sabe que el peso medio de todos los atletas es 83,27 Kg. y que los del segundo avión tienen un peso medio de 45,35 Kg. menos que el de los atletas del primer avión. ¿cuál es el peso medio de cada grupo?

8. Suponga que el coeficiente de variación del peso inicial de 10 cerdos pequeños es de 20,6%. Después de someterlos a cierto tratamiento, se observó un aumento de aproximadamente 7 kilos por cerdo, de modo que el coeficiente de variación resultó igual a 18,2%.

- ¿Cuál es el peso medio de los cerdos después del tratamiento?
- ¿Cuál es la desviación estándar antes y después del tratamiento?

10. Un botánico, en un primer intento por comparar los efectos de dos tratamientos diferentes sobre la cantidad de agua que pierden las hojas de cierta planta (miligramos por centímetro cuadrado), calcula algunas medidas descriptivas. En su estudio toma 50 hojas al azar, de las cuales a 20 se les aplica el tratamiento "A" y al resto el tratamiento "B". Se obtiene la siguiente información:

	Tratamiento A	Tratamiento B
Media	65	82.5
Desviación estándar	13	16.5

- ¿Cual es la perdida media de agua de todas las hojas en tratamiento?
- ¿Cual de los dos tratamientos produce resultados más homogéneos?
- Posteriormente se descubre que los instrumentos utilizados para medir estaban descalibrados. Para corregir el error, se debe agregar 20 unidades a cada una de las observaciones del tratamiento "A", y se debe disminuir en un 10% cada una de las observaciones del tratamiento "B". ¿Cómo se ven afectados los resultados anteriores después de las modificaciones?

11. Se le pide a tres entrenadores información respecto a la distribución de los tiempos porcentuales de trabajo, obteniéndose la siguiente tabla:

Actividades		Entrenador		
		A	B	C
Teoría		2	14	7
Practica:	Pre calentamiento	8	5	10
	Entrenamiento	0	20	9
	Deporte	70	55	40
Sociabilidad		13	0	22

Analice descriptivamente esta información y haga un breve informe.

12. Para trabajo en grupo. Obtenga la siguiente información de un grupo de alumnos del curso: promedio de notas de enseñanza media, puntajes de la prueba de aptitud académica matemática y verbal, puntaje prueba específica en matemáticas, y promedio de notas de las asignaturas cursadas en la universidad, ponderándola por los créditos respectivos. Haga un análisis descriptivo de estos datos, que incluya tablas y gráficos apropiados, y correlaciones entre las variables. Haga un informe-resumen, incluyendo algunas conclusiones y comentarios.

13. En una universidad, los alumnos tienen dos posibilidades: pagar la matrícula semestral al contado o bien, obtener un crédito estatal por un porcentaje del valor de la matrícula, y en consecuencia, tener que pagar sólo la diferencia al contado. Al comienzo de un determinado semestre, se cuenta con los resultados de las solicitudes para el crédito estatal (para el semestre). Los datos de que se dispone son los siguientes:

Valor Arancel semestral.....	\$ 1.100.000
Total de alumnos regulares.....	2.300
Número de solicitantes de crédito.....	1.820
Monto total del crédito otorgado.....	\$ 107.000.000
Coeficiente de variación de los montos de crédito individuales	
asignados a los alumnos que solicitaron crédito.....	0,62

El cálculo del coeficiente de variación, no incluye a 260 alumnos que solicitaron y no obtuvieron crédito estatal (es decir, se les dio 0% de crédito). Se sabe además, que ningún alumno obtuvo el 100% de crédito.

- Calcule la desviación standard de los montos individuales de crédito, entre los alumnos que obtuvieron crédito estatal.
- Calcule la cantidad promedio de pago de matrícula - por alumno - al contado, sobre el total de alumnos.
- Calcule la desviación standard de los montos de crédito individual por alumno, sobre el total de alumnos.

14. Un vendedor ha perdido una boleta de entre las últimas 10 ventas realizadas en la última hora. El promedio de las 10 ventas (en dólares) es \$7.20 y las otras 9 boletas tienen los siguientes valores:

\$5.72 \$4.80 \$7.10 \$7.90 \$9.55 \$4.45 \$7.54 \$8.34 \$9.70

$\Sigma x = 65.1$, $\Sigma x^2 = 500.0806$.

- a) ¿Cuál es el valor de la boleta perdida?
- b) Calcule la desviación estándar de los valores de ventas de las 9 boletas restantes. Indique el significado de este valor.

15. En una industria, el coeficiente de variación (C.V.) de los sueldos (en U.F.) de 300 empleados es de 10%. Según la ley, todos los sueldos deben ser reajustados en un 15%. Además, debido a las utilidades que ha tenido la industria, la gerencia decide dar una bonificación de 2 U.F. por empleado, con lo cual el C.V. queda en un 9,2%.

- a) Determine el promedio y desviación estándar de los sueldos de los 300 empleados antes de las modificaciones salariales.
- b) Determine la cantidad de dinero que necesitará la industria para pagar los sueldos después de las modificaciones salariales.

16. Los datos siguientes corresponden al número de sobres de café defectuosos producidos por día durante el turno N° 1 en la máquina N°4:

10	11	15	16	18	15	10	13	14	10
13	12	14	15	16	16	14	15	16	11
13	18	19	16	18	12	16	17	15	13

- a) Calcule el número promedio de sobres de café defectuosos. Comente.
- b) Calcule la desviación estándar del número de sobres de café defectuosos. Comente.
- c) Calcule el Cuartil 3. Comente.
- d) Obtenga medidas de simetría y curtosis, e interpréte las.

17. Las cinco primeras desviaciones con respecto a la media muestral de un conjunto de seis observaciones de la medición de la concentración de cierta sustancia química en el agua potable son: -2; 3; 7; 4 y -1.

- a) ¿Cuál es el valor de la sexta desviación respecto de la media?
- b) Calcule la varianza de las seis observaciones.

18. Los datos siguientes corresponden a duración (en segundos) de 12 llamadas telefónicas. Calcule el promedio, la desviación estándar, la mediana, el coeficiente de simetría y de curtosis. Con ellas, dé una descripción de este conjunto de datos. ¿Cuál medida de tendencia central usaría? ¿Por qué?

34	27	25	47	32	27
20	23	26	20	53	22

19. Un Profesor de biología tomó una prueba a 8 alumnos, al momento de entregar las notas se dio cuenta que faltaba una de las pruebas. De antemano él sabe que el promedio es de 64.19 puntos. Los siguientes datos son los puntajes obtenidos por 7 alumnos:

25,6 55,6 85,2 75,3 35,9 89,2 96,2

$$\Sigma x = 513.5 \quad \Sigma x^2 = 37725.99$$

- a) ¿Cuál es el puntaje perdido por el profesor?
 b) Calcule la desviación estándar de los puntajes de los 8 alumnos. Indique el significado de este valor.

20. a) La siguiente pregunta y la correspondiente tabulación de las respuestas, aparecieron publicadas en un estudio sobre las intenciones de los consumidores.

"¿Planea comprar un automóvil dentro de los próximos 24 meses?"

Respuesta	Cantidad de personas	Porcentaje
Si	215	14
No	480	31
No sabe	864	55
Total	1559	100

Suponga que en lugar de la tabulación anterior, hubiese aparecido lo siguiente:

Respuesta	Cantidad de personas	Porcentaje
Si	215	31
No	480	69
Total	695	100

- a) ¿Hubiese usted extraído las mismas conclusiones a partir de las dos tablas? Explique.
 ¿Cuándo sería apropiada la segunda tabulación?
 b) Titule adecuadamente ambas tablas.
 c) Con el propósito de seleccionar 80 personas de un total de 400 postulantes se realiza una prueba escrita. Una persona obtiene 500 puntos. Si este puntaje corresponde al percentil 18 ¿Queda seleccionada esta persona ¿Por qué?

PROYECTOS PARA TRABAJO EN GRUPO

1. Organice una encuesta utilizando muestreo aleatorio simple. Escoja un problema que requiera estimar una proporción poblacional. Para su proyecto, discuta la factibilidad y las ventajas esperadas de estratificación versus otros tipos de muestreo.

2. Obtenga la siguiente información de un grupo de alumnos del curso: promedio de notas de enseñanza media, puntajes de la prueba de aptitud académica matemática y verbal, puntaje prueba específica en matemáticas, y promedio de notas de las asignaturas cursadas en la universidad, ponderándola por los créditos respectivos. Haga un análisis descriptivo de estos datos, que incluya tablas y gráficos apropiados, y correlaciones entre las variables. Haga un informe-resumen, incluyendo algunas conclusiones y comentarios.

Los datos siguientes corresponden a distancias de frenado (en pies) de autos guiados en un terreno húmedo. Treinta y dos personas guiaron el mismo auto a una misma velocidad. Elabore un informe que incluya una tabulación (6 intervalos), un gráfico, medidas de tendencia central y de dispersión con los comentarios correspondientes.

124	107	118	133	118	141	128	129
129	132	131	137	126	126	136	116
121	121	117	108	116	111	140	142
132	109	125	113	135	122	128	117

$$\sum_{i=1}^{32} distancia_i = 3988$$

$$\sum_{i=1}^{32} distancia_i^2 = 500024$$