

MATEMÁTICA: CASUÍSTICA

ESTADÍSTICA

RESUMEN TEÓRICO

Estadística descriptiva: obtención, organización y representación de datos

La **estadística descriptiva** es la rama de la estadística que se ocupa de describir conjuntos de mediciones, tanto de **muestras** como **poblaciones**. Una vez que hayamos recolectado un conjunto de **mediciones** o **datos**, ¿cómo se puede mostrar este conjunto en una forma clara, entendible y fácil de leer e interpretar? Primero, se debe conocer la definición de **medición** o **"datos"** y clasificar los tipos de datos que probablemente se encuentre en la vida real, luego organizarlo en tablas y gráficos estadísticos para su lectura análisis e interpretación.

Variable es una característica que cambia o varía con el tiempo y/o para diferentes personas u objetos bajo consideración.

Unidad experimental es el individuo u objeto en el que se mide una variable. Resulta una sola **medición** o dato cuando una variable se mide en realidad en una unidad experimental.

Población es el conjunto completo de elementos, con alguna característica común, que es el objeto de estudio. Una población puede ser finita o infinita.

Muestra es un subconjunto de mediciones seleccionado de la población de interés.

Dato es el resultado de la medición de cierta característica variable.

VARIABLES Y DATOS

VARIABLES		
CUANTITATIVA		CUALITATIVA
DISCRETA	CONTINUA	
<p>Una variable cuantitativa discreta toma valores numéricos exactos, es decir a los números naturales: 0; 1; 2; 3; 4; ...</p> <p>Ejemplos:</p> <ul style="list-style-type: none"> • La cantidad de televisores LCD que se vende en la tienda CARSA. • La cantidad de hijos por familia en la provincia de Virú. • El número de electrones de un átomo. 	<p>Una variable cuantitativa continua puede tomar cualquier valor correspondiente a los puntos en un intervalo de recta real. (Considera fracciones y decimales)</p> <p>Por ejemplo, son variables continuas la longitud, el peso, el tiempo, la tasa de interés que paga o cobra un banco, las ventas diarias en un supermercado, la velocidad de un móvil, etc.</p>	<p>Las variables cualitativas miden una cualidad o característica en cada unidad experimental. Las variables cualitativas producen datos que se pueden clasificar de acuerdo a similitudes o diferencias en clase.</p> <p>Algunos ejemplos:</p> <ul style="list-style-type: none"> • Raza: mestizo, blanco, indígena, mulato, negro, asiático. • Tipo de logro: excelente, bueno, regular, malo. • Estado civil: soltero, casado, divorciado, viudo.

Datos cuantitativos	Datos cualitativos
<p>Los datos cuantitativos describen información que puede ser obtenida mediante conteo o medición. Algunas preguntas de las que surgen datos cuantitativos son:</p> <p>¿Cuántos lapiceros tienes?</p> <p>¿Cuánto tiempo demoras en llegar de tu habitación a tu aula?</p> <p>¿Cuántos estudiantes hay en el aula?</p> <p>¿Cuál es la velocidad del automóvil?</p>	<p>Los datos cualitativos determinan categorías y se les denomina datos categóricos. Algunas preguntas de las que surgen datos cualitativos son:</p> <p>¿Cuál es tu color favorito?</p> <p>¿Qué medio utiliza para trasladarse de su casa al colegio?</p> <p>¿Cuál es la marca de tu computadora?</p> <p>¿Cuál es el nivel de aprendizaje logrado?</p>

¿Cuál es la diferencia entre una población y una muestra?

Cuando pensamos en el término población, usualmente pensamos en gente de nuestra ciudad, región o país.

En estadística, el término **población** incluye **todos los elementos** o miembros de un grupo definido que estamos estudiando para tomar decisiones basadas en datos obtenidos.

La **muestra** es una **parte de la población**, es decir es un **subconjunto de la población**, una selección de individuos de la población.

Las muestras aleatorias deben tener dos características:

1. Cada individuo tiene igual oportunidad de selección.
2. La muestra tiene esencialmente las mismas características que la población.

Algunas precisiones sobre las medidas de tendencia central.

	Ventajas	Desventajas
Moda La moda puede utilizarse para datos cualitativos o cuando se pide elegir el elemento más frecuente	<ul style="list-style-type: none"> • Los valores extremos no afectan el valor de la moda. 	<ul style="list-style-type: none"> • No utiliza a todos los elementos del conjunto de datos. • No es necesariamente única, puede haber más de una respuesta. • Cuando no hay valores repetidos en el conjunto de datos. no existe la moda. • Cuando hay más de una moda, es difícil de interpretar
Media La media describe el centro de un conjunto de datos.	<ul style="list-style-type: none"> • Es la medida más ampliamente utilizada en áreas como negocios, ingeniería e informática. • Se utiliza a todos los elementos del conjunto de datos. • Es única y tiene sólo una respuesta. • Es útil en la comparación de conjuntos de datos. 	<ul style="list-style-type: none"> • Los valores extremos afectan la media.
Mediana La mediana describe el centro de un conjunto de datos.	<ul style="list-style-type: none"> • Los valores extremos no la afectan tanto como a la media. • Es útil en la comparación de conjunto de datos. • Es única y tiene solo una respuesta. • Por tratarse del valor del medio, deja al 50% de los datos a cada lado. 	<ul style="list-style-type: none"> • No es ampliamente utilizada como la media. • Es menos utilizada en cálculos avanzados.

Situación de aprendizaje

¿Es justo o injusto?

En el Proyecto Especial Chavimochic, el gerente general tiene un sueldo de S/ 12 000; cada subgerente S/ 7000, los jefes de oficina S/ 4000 cada uno, otros empleados S/ 3000 cada uno y los obreros, los nombrados antiguos ganan S/ 950 cada uno, los nombrados nuevos S/ 900 y los contratados el sueldo mínimo de S/ 850 cada uno. En dicho proyecto hay un gerente, 6 subgerentes, 4 jefes, 14 empleados, 38 obreros nombrados antiguos, 30 obreros nombrados nuevos y 32 obreros con contrato temporal.

Los obreros del proyecto a través de su sindicato hicieron un reclamo al Gobierno Regional, pidiendo un aumento de sueldo a un mínimo de S/ 1200 y un máximo de S/ 1400. El gerente general en su informe indica que el sueldo promedio es de S/ 1618,4; tras realizar el siguiente cálculo:

CARGO	SUELDO MENSUAL	CANTIDAD DE TRABAJADORES (f)	F	PLANILLA
Obreros contratados	850	32	32	27200
Obreros nuevos	900	30	62	27000
Obreros antiguos	950	38	100	36100
Empleados de oficina	3000	14	114	42000
Jefes	4000	4	118	16000
Subgerente	7000	6	124	42000
Gerente	12000	1	125	12000
TOTAL			125	202300

$$\mu = \frac{32 \times 850 + 30 \times 900 + 38 \times 950 + 14 \times 3000 + 4 \times 4000 + 6 \times 7000 + 1 \times 12000}{32 + 30 + 38 + 14 + 4 + 6 + 1}$$

$$\mu = \frac{202300}{125} = 1618,4$$

El presidente regional al observar que el sueldo promedio era superior a la petición sindical, determina no realizar el aumento.

- ¿Es correcta la decisión tomada por el presidente regional? ¿por qué?
- Si el sueldo de la mayoría oscila entre 850 y 950 soles, ¿por qué el sueldo promedio es S/ 1618,4?
- ¿Es la media aritmética una medida confiable?
- ¿Qué medida estadística debió calcular el gerente para adjuntar al informe de la realidad económica de los obreros para obtener el aumento respectivo?
- ¿Qué gráficos estadísticos nos proporcionan una mejor idea sobre la realidad de los sueldos de todos los trabajadores del PECH?

PROBLEMAS PROPUESTOS:

1. Un docente tiene como propósito que sus estudiantes seleccionen la medida de tendencia central apropiada para analizar un conjunto de datos al resolver problemas.
¿Cuál de los siguientes problemas favorece el logro del propósito planteado?
 - a) Un grupo de 7 directores de colegios, procedentes de diferentes instituciones educativas de la región Lima, han viajado para reunirse en la ciudad de Lima. Arturo viajó 40 km; Benjamín, 120 km; Cristina, 73 km; Doris, 60 km; Ernesto, 75 km; Federico, 85 km; y Gabriela, 60 km. Calcule la media, la mediana y la moda de las distancias que han viajado estos amigos. ¿Cuál de estas medidas de tendencia central es la de mayor valor numérico?
 - b) Se desea realizar un estudio comparativo de las masas corporales de un grupo de estudiantes mujeres de una sección de segundo grado del colegio Santa María. Se registró las masas de 10 estudiantes mujeres por sección. Las masas expresadas en kilogramos de una de las muestras son: 56, 63, 61, 57, 58, 60, 62, 63, 57 y 59 kg. ¿Cuál es el valor de la media, mediana y moda de estos datos?
 - c) En un aula del colegio San Juan, se realizó una encuesta a los estudiantes sobre el tiempo que necesitan para llegar a su colegio caminando. Un 15% de los encuestados necesita 10 minutos; 20%, 25 minutos; 25%, 20 minutos; y el resto necesita 15 minutos. ¿Qué medida de tendencia central describe mejor el tiempo más frecuente que necesitan los estudiantes encuestados para llegar a su colegio caminando?
2. Un docente hizo una encuesta a sus estudiantes sobre la cantidad panes que consumen en su desayuno cada día. Luego, les presentó estos datos y les pidió que obtuvieran el promedio.
2; 4; 0; 1; 0; 5; 4; 2; 0; 4; 2; 0; 3; 1; 2; 1; 0; 4; 0; 3; 2; 4; 1; 3; 0
Uno de los estudiantes presentó su solución:

$$\bar{x} = \frac{5(2) + 5(4) + 7(0) + 4(1) + 1(5) + 3(3)}{25} = \frac{48}{25} = 1,92$$

Y dijo: “En promedio, tenemos un consumo 1,92 panes por estudiante, ¿1,92 panes?”.

¿Cuál de las siguientes acciones pedagógicas es pertinente para favorecer en este estudiante la interpretación del promedio?

- a) Pedirle que explique por qué le parece extraño el promedio obtenido y si cree que es incorrecto. Luego, preguntarle: “Si reunimos la cantidad de panes obtenidos y tuviéramos que distribuir entre todos, de forma equitativa, ¿es posible repartir 1 pan a cada estudiante? ¿Será posible repartir 2 panes a cada uno? ¿Por qué? ¿Cómo se entiende que el promedio sea 1,92?”.
- b) Presentarle un caso similar, por ejemplo, el de considerar las estaturas de sus compañeros y pedirle que calcule el promedio de estas estaturas. Luego, preguntarle: “¿Cuánto es la suma de las estaturas? ¿Cuántas estaturas fueron registradas? ¿Cuál es el promedio de estas estaturas? ¿Cómo se entiende este promedio?”.

- c) Decirle que no siempre el promedio de un conjunto de datos es un número entero y que es válido que pueda resultar un número decimal. Luego, preguntarle: “¿Cuántos panes consumen en total? ¿Cuántos estudiantes fueron registrados? ¿Cuál es el promedio de panes que consume cada estudiante?”.

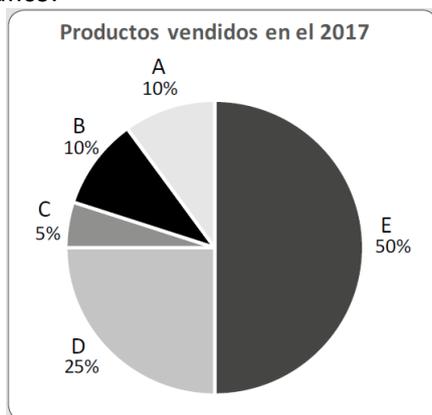
3. ¿Cuál de las siguientes tareas es de mayor demanda cognitiva?

- a) La media de la edad de 5 personas es 18 años y la mediana de sus estaturas es 1,65 m. Si se sabe que 4 de estas personas tienen 14, 18, 19 y 20 años y sus estaturas son 1,60 m; 1,62 m; 1,70 m y 1,73 m, ¿cuál es la edad y estatura que tiene la persona restante, de modo que cumplan, respectivamente, con la media y mediana dadas? Explica tu procedimiento.
- b) David desea saber qué valor representa la asistencia de aficionados al estadio de su comunidad durante dos meses. Para ello, solicitó la asistencia durante dicho periodo: 1800, 2000, 1600, 1800, 2400, 2200, 2800 y 8000. ¿Cuál de las siguientes medidas es pertinente para encontrar ese valor: la media, mediana o moda? Explica tu respuesta.
- c) El equipo de básquet femenino está conformado por Andrea (1,80 m), Blanca (1,65 m), Cinthya (1,60 m), Doris (1,70 m) y Elena (1,58 m). El día de hoy, Fernanda (1,60 m) se incorpora al equipo. ¿Su inclusión aumentará o disminuirá la media de las estaturas del equipo? Explica tus razones.

Lea la siguiente situación y responda las preguntas 4 y 5.

Una docente tiene como propósito evaluar el logro del siguiente desempeño: “Interpreta información de un gráfico circular estableciendo relaciones entre porcentajes y fracciones”; para ello, selecciona la siguiente actividad.

Durante una reunión sobre el balance anual del 2017, el gerente de una empresa está brindando información a sus colaboradores. Para representar la cantidad de productos vendidos (A, B, C, D y E), ha presentado el siguiente gráfico:



Luego, la docente indica que observen el gráfico y elaboren conclusiones acerca de la venta anual de los productos ofrecidos por la empresa estableciendo relaciones entre porcentajes y fracciones. Para evaluar la respuesta de los estudiantes, la docente ha elaborado una rúbrica con las descripciones de los niveles “En inicio”, “En proceso” y “Logrado”.

En inicio	En proceso	Logrado	Destacado
Menciona información explícita del gráfico.	Elabora conclusiones que implican comparaciones entre los sectores y el total utilizando solo porcentajes.	Elabora conclusiones que implican comparaciones entre los sectores y el total utilizando porcentajes y fracciones.	

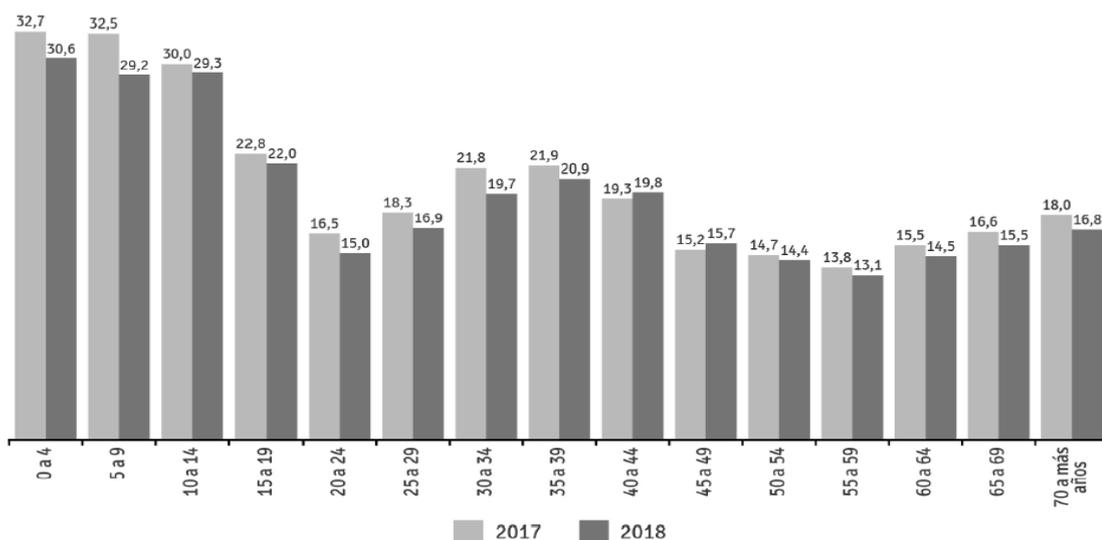
4. La docente está interesada en plantear la descripción del nivel “Destacado”. ¿Cuál de las siguientes descripciones es la que corresponde al nivel “Destacado”?
- Elabora de cinco a más conclusiones que implican comparaciones entre los sectores y el total utilizando porcentajes y fracciones.
 - Elabora conclusiones que implican comparaciones entre los sectores y el total, y entre sectores entre sí utilizando porcentajes y fracciones.
 - Elabora conclusiones que implican comparaciones entre los sectores y el total utilizando porcentajes y fracciones. Además, encuentra el ángulo que representa cada sector del gráfico circular.
5. Al revisar las conclusiones de los estudiantes, la docente encuentra el siguiente registro de un estudiante:

La cantidad vendida del producto E representa el 50% de la venta anual y equivale a la mitad de dicha venta, y es mayor que la cantidad vendida del producto D que representa el 25% de la venta anual y equivale a la cuarta parte de dicha venta.

Considerando la rúbrica presentada, ¿cuál es el nivel de logro alcanzado por el estudiante?

- En proceso
 - Logrado
 - Destacado
6. Una docente presentó el siguiente gráfico a los estudiantes:

PERÚ: INCIDENCIA DE LA POBREZA POR GRUPOS DE EDAD, 2017 - 2018
(Porcentaje respecto de cada grupo de edad)



Adaptado de INEI (2019). *Evolución de la pobreza monetaria. 2017-2018. Informe Técnico.*

¿Cuál de los siguientes grupos de preguntas es **más** pertinente para ayudar a los estudiantes a realizar inferencias a partir del gráfico?

- ¿Qué fenómeno representa el gráfico de barras? ¿En cuál de los ejes se expresa cada una de las variables?, ¿con qué unidades? ¿A qué años corresponde el estudio de incidencia? ¿Cómo evoluciona la incidencia de la pobreza en los grupos de edad de personas mayores o iguales que 50 años?
- ¿Cuáles son las variables que intervienen? ¿Es adecuado representarlas mediante un gráfico de barras? ¿Con qué escala se ha representado el gráfico? ¿Cuál fue la secuencia de la construcción

del gráfico? ¿En qué grupo de edad la incidencia de la pobreza fue mayor en el año 2017?, ¿y en el 2018?

- c) ¿Cuál es la tendencia general en la evolución porcentual de la pobreza en el 2017 y 2018? ¿Hay algún grupo de edad en el cual la proporción de incidencia de la pobreza se ha incrementado? ¿A qué grupo de edad debería otorgarle prioridad un programa destinado a reducir la incidencia de la pobreza?, ¿por qué?

7. Una docente ha propuesto el siguiente problema a los estudiantes:

El entrenador de un equipo de básquet pretende incorporar a un jugador que sea muy hábil encestando. Se presentan dos candidatos: Juan y Pablo. Para tomar su decisión final, el entrenador considerará los puntos que hagan al encestar en los partidos que enfrenten a los equipos A, B, C y D. En la siguiente tabla, se indican los puntos que hicieron al enfrentar a dichos equipos.

Equipos	Puntos de Juan	Puntos de Pablo
Equipo A	26	30
Equipo B	23	20
Equipo C	24	20
Equipo D	25	28

El asistente técnico recomendó al entrenador que incorpore a Juan, con el argumento de que es el jugador con mayor regularidad en los puntos realizados en cada partido. Explica cuál es la posible medición que consideró el asistente para hacer dicha recomendación.

¿Cuál es el propósito que busca lograr la docente con el problema propuesto?

- a) Que los estudiantes justifiquen el uso de una de las medidas de dispersión en la resolución de un problema.
- b) Que los estudiantes fundamenten la pertinencia de la medida de tendencia central elegida en la resolución de un problema.
- c) Que los estudiantes expliquen el procedimiento más eficaz para el cálculo de una medida de dispersión como parte de la resolución de un problema.

8. Raúl, Susana y Teresa trabajan en una misma empresa. Los salarios de los trabajadores de esta empresa se calculan según el cargo y las funciones que desempeñan. Además, se sabe que el primer cuartil de los salarios es 2000 soles, la mediana es 2300 soles y el tercer cuartil es 3000 soles.

Teniendo en cuenta esta información, ¿cuál de las siguientes afirmaciones es verdadera?

- a) Raúl gana 1650 soles; por lo tanto, su salario está por encima del 25 % de los salarios de los trabajadores que menos ganan en la empresa.
- b) Teresa gana 2800 soles; por lo tanto, su salario está dentro del 25 % de los salarios de los trabajadores que más ganan en la empresa.
- c) Susana gana 2500 soles; por lo tanto, como mínimo, el 25 % de los trabajadores de la empresa ganan más que ella.

9. Un docente plantea a los estudiantes la siguiente situación:

La tabla que se presenta a continuación muestra la cantidad de niños que atendió la posta médica de un centro poblado en cierto día.

Edad (años)	f_i
[0; 2[8
[2; 4[4
[4; 6[8
[6; 8[10
[8; 10[7
[10; 12]	13
Total	50

En este contexto, el docente propone diversas tareas. ¿Cuál de las siguientes tareas propuestas es de **mayor** demanda cognitiva?

- Se requiere repetir los análisis de los pacientes cuya edad coincide con la mediana de las edades de todos los pacientes. El mismo procedimiento se debe realizar con los pacientes cuya edad coincide con el segundo cuartil de todas las edades. ¿Qué relación existe entre estos pacientes?
- Se hará un taller sobre factores de riesgo de enfermedades infecciosas con los padres de los pacientes que se encuentran por debajo del segundo decil. ¿Cuál es la edad que corresponde al segundo decil de las edades de los pacientes?
- Se debe aplicar una encuesta de satisfacción a los padres o apoderados de los pacientes que corresponden al primer, segundo y tercer quintil, según su edad. ¿Cuáles son las edades de estos pacientes?

10. Para integrar el elenco de actores de una obra de teatro, se inscribieron 9 estudiantes. Según el orden en que fueron inscritos, sus edades son 15, 13, 11, 9, 10, 11, 11, 13 y 15 años.

En relación con esta información, ¿cuál de las siguientes proposiciones es verdadera?

- La distribución de edades es bimodal; las modas son 13 y 15 años.
- La media aritmética de las edades es 12 años.
- La mediana de las edades es 10 años.

11. Según una nota periodística, publicada el 28 de marzo del presente año en la página web de una emisora radial, los peruanos anualmente gastan en promedio S/ 40 en la adquisición de bloqueadores solares.

¿Cuál de las siguientes proposiciones puede concluirse **a partir de tal afirmación**?

- El precio más común de los bloqueadores solares ofrecidos a los peruanos es S/ 40.
- Por lo menos la mitad de los peruanos anualmente gastan S/ 40 en la adquisición de bloqueadores solares.
- Los peruanos anualmente gastan por debajo de los S/ 40, igual o por encima de dicho monto en adquirir bloqueadores solares.

12. Una empresa fabricante de medicamentos desea conocer cuán efectivo puede ser un nuevo medicamento para controlar cierta enfermedad en personas adultas. Para ello, realizó un estudio en 3600 personas adultas que tenían esta enfermedad. Así, encontró que el 75 % de las personas que recibieron el medicamento pudieron controlar la enfermedad.

¿Cuál de las siguientes alternativas presenta a la **muestra** considerada para realizar el estudio mencionado?

- Conjunto de todas las personas adultas que tienen la enfermedad.
- Conjunto de las 3600 personas sometidas al estudio que tienen la enfermedad.
- Conjunto del 75 % de personas que pudo controlar la enfermedad con el medicamento.