

## ***Conceptos fundamentales en estadística.***

**Agrupación.** Es la acción y efecto de agrupar (reunir en grupo, apiñar).

**Agrupación de datos.** Son aquellos que se han clasificado en categorías o clases, tomando como criterio su frecuencia. Esto se hace con la finalidad de simplificar el manejo de grandes cantidades de datos y establecer sus tendencias.

**Aleatorio.** Es todo dato obtenido al azar, es decir, que todo número tenga la misma probabilidad de ser elegido y que la elección de uno no dependa de la elección del otro.

**Amplitud.** La diferencia entre el valor máximo y mínimo de los valores de una variable se encuentran comprendidos el 100% de los valores muestrales.

**Análisis.** Un fenómeno aleatorio es aquel que, bajo el mismo conjunto aparente de condiciones iniciales, puede presentar resultados diferentes, es decir, no se puede predecir el resultado exacto de cada experiencia particular. (Ej: Lanzamiento de un dado).

**Aplastamiento.** Sinónimo de curtosis.

**Asimetría.** Se refiere a si la curva que forman los valores de la serie presenta la misma forma a la izquierda y derecha de un valor central (media aritmética). Es positiva si hay tendencia a puntuaciones bajas y negativa si la hay a puntuaciones altas.

**Bernoulli.** Una distribución de Bernoulli se compone de dos sucesos contrarios A y B, a los que se les suele llamar éxito y fracaso, con probabilidades p y q respectivamente.

**Bidimensional.** Una variable estadística bidimensional es el conjunto (X,Y) de valores que pueden tomar dos caracteres diferentes X e Y medidos sobre cada uno de los individuos de una población o muestra, una de las dos variables es la variable independiente y la otra variable dependiente.

**Binomial.** Es una distribución de probabilidad discreta que cuenta el número de éxitos en una secuencia de n ensayos de Bernoulli independientes entre sí, con una probabilidad fija p de ocurrencia del éxito entre los ensayos.

**Campana de Gauss.** Curva de forma acampanada correspondiente a la función de densidad de probabilidad de una distribución normal. Nombre asignado coloquialmente a la representación gráfica de la distribución normal.

**Característica.** es una propiedad de los individuos de una población. Uno de los objetivos de la investigación psicológica es determinar las propiedades de una determinada población, propiedades que estarán presentes en todos y cada uno de los individuos que la conforman.

**Censo.** Es una investigación estadística que consiste en el recuento de la totalidad de los elementos que componen la población por investigar. Es necesario que se especifique el espacio y el tiempo al que se refiere el recuento. Medición o examen de cada elemento de la población.

**Coeficiente.** Cuando se desea hacer referencia a la relación entre el tamaño de la media y la variabilidad de la variable, se utiliza el coeficiente de estadística de variación.

**Confiabilidad.** Probabilidad de que la estimación efectuada se ajuste a la realidad. Cualquier información que queremos recoger está distribuida según una ley de probabilidad (Gauss o Student), así llamamos nivel de confianza a la probabilidad de que el intervalo construido en torno a un estadístico capte el verdadero valor del parámetro.

**Constante.** Llamaremos constante a una característica que sólo admite una modalidad, por ejemplo, la constante de gravitación universal.

**Continua.** Una variable se llama continua si entre cada dos valores suyos pueden existir infinitos otros, como el peso, la estatura, etc.

**Contraste.** Particularmente en análisis de varianza y regresión lineal, un contraste es una combinación lineal de variables cuyos coeficientes suman cero, permitiendo la comparación de diferentes tratamientos

**Correlación.** Es una medida de la relación entre dos o más variables. La correlación puede tomar valores entre  $-1$  y  $+1$ . El valor de  $-1$  representa una correlación negativa perfecta mientras un valor de  $+1$  representa una correlación perfecta positiva.

**Covarianza.** Relación sistemática entre dos variables, en la cual el cambio en una implica un cambio correspondiente en la otra.

**Cualitativo.** Se aplica a la variable (o dato, o medida) que sólo admite una medida nominal.

**Cuasi-varianza.** Es un índice de variabilidad que se obtiene sumando los cuadrados de las diferencias de todas las puntuaciones respecto a la media y dividiendo entre  $(n-1)$ .

**Cuantil.** Diremos que un número es el cuantil de orden  $p$  en una distribución de frecuencias si el porcentaje de datos inferiores a él es igual a  $p$  (y los superiores  $100-p$ ). Por ejemplo, el cuantil  $C_{85}$  será un punto que cumple que el  $85\%$  de los datos es inferior a él.

**Cuartil.** Los cuartiles son las puntuaciones que dividen la distribución en cuatro partes iguales con el  $25\%$  de los datos en cada una de ellas.

**Curtosis.** Grado de apuntamiento de la distribución de frecuencias. Si es alto, se denomina distribución leptocúrtica y si es bajo platicúrtica.

**Chi-cuadrado.** Es la distribución teórica que representa la distribución muestral de la suma de cuadrados de los datos dividida entre la varianza de la población.

**Dato.** Conocido también como información, es el valor de la variable asociada a un elemento de una población o una muestra.

**Decil.** Percentil cuyo valor que indica su proporción es un múltiplo de diez. Percentil 10 es el primer decil, percentil el segundo decil, etc.

**Descriptiva.** Parte de la estadística que resume la información de la muestra. La información recogida y resumida en los estadísticos se usa para la estimación de parámetros poblacionales.

**Desviación.** Es una medida de la diferencia entre el valor observado de una variable y algún otro valor, a menudo la media de esa variable. El signo de la desviación informa la dirección de esa diferencia

**Determinación.** Es una medida estadística de la bondad del ajuste o fiabilidad del modelo estimado a los datos.

**Dicotómico.** Adjetivo que se aplica a toda medida o proceso que sólo puede presentar dos valores, como SÍ/NO, Hombre/Mujer, Encendido/Apagado.

**Discreta.** Si una variable solo admite un número finito de valores entre cada dos, recibirá el nombre de discreta (edades medidas en años, número de hermanos, etc.).

**Distribución.** La distribución de probabilidad de una variable aleatoria es una función que asigna a cada suceso definido sobre la variable la probabilidad de que dicho suceso ocurra.

**Error.** Es una medida del error que se comete al tomar la media calculada en una muestra como estimación de la media de la población total.

**Escala.** La distribución de datos puede recogerse en distintas escalas: nominal, dicotómica, discreta o continua.

**Esperanza.** (valor esperado o media) de una variable aleatoria discreta es la suma de los productos de sus valores por sus probabilidades asociadas.

**Estadística.** Es la ciencia que se ocupa de la ordenación y análisis de datos procedentes de muestras y de la realización de inferencias sobre las poblaciones de las que éstas proceden.

**Estimación.** Un Estimador es un estadístico cuyos valores se consideran próximos a un parámetro que, por ser generalmente desconocido, se desea estimar.

**Estimador.** Un Estimador es un estadístico cuyos valores se consideran próximos a un parámetro que, por ser generalmente desconocido, se desea estimar.

**Extremo inferior.** Es el valor mínimo que puede tener un valor incluido en un intervalo de datos agrupados.

**Extremo superior.** Es el valor máximo posible en un intervalo de datos agrupados. Se considera no alcanzable. Así si un intervalo comprende desde 5 hasta 10, incluiremos en el mismo los valores comprendidos entre estos dos, incluyendo el 5 y sin incluir el 10.

**Frecuencia.** Número a veces que se repite, en una población o muestra, cada una de las categorías de una variable cualitativa o de los valores de una variable cuantitativa.

**Hipótesis.** Enunciado o proposición no probados acerca de un factor o fenómeno de interés para el investigador. Una hipótesis estadística a un enunciado respecto a una población y usualmente es un enunciado respecto a uno a más parámetros de la población.

**Histograma.** Gráfico de la representación de distribuciones de frecuencias, en el que se emplean rectángulos dentro de unas coordenadas.

**Indirecta.** Es un método de recolección de datos sobre las características y propiedades de un individuo, fenómeno o situación particular.

**Inferencia.** Es una parte de la estadística cuya finalidad es obtener conclusiones respecto a la población a partir de datos observados en muestras. Es el proceso por medio del cual se hacen aseveraciones o estimaciones de un todo, a partir de sus partes o elementos.

**Investigación.** Es una actividad que apela a diversas técnicas con el propósito de llegar a la esencia de la realidad.

**Marco.** Muestral es una lista de elementos que componen el universo que queremos estudiar y de la cual se extrae la muestra.

**Marca de clase.** Es el punto medio de una clase y se obtiene sumando los límites inferiores y superiores de una clase y dividiendo el resultado entre dos.

**Media.** Suma de los datos dividida entre el número de datos.

**Mediana.** Medida de tendencia central que indica la puntuación que es superada por la mitad de las observaciones, por tanto, no por la otra mitad. Puede utilizarse con variables cuantitativas y ordinales.

**Medida.** Es la operación de asignar un número a cada una de las modalidades de una característica, convirtiendo algunas relaciones entre modalidades en sus correspondientes relaciones entre los números que representan su medida.

**Moda.** Es el valor con una mayor frecuencia en una distribución de datos.

**Modalidad.** La modalidad, cada una de las maneras en que se presenta una característica. Por ejemplo, en el caso de que la característica fuese el género, varón y mujer. Véase Característica.

**Muestra.** es una selección de datos extraídos de una población o universo de posibles observaciones.

**Muestreo.** Procedimiento seguido para seleccionar un subconjunto de los elementos de una población.

**Nivel de confianza.** Probabilidad que el intervalo contenga el parámetro. Por ejemplo, podemos afirmar con un nivel de confianza del 89% que si lanzamos una moneda no trucada obtendremos entre 3 y 7 caras.

**Nivel de significación.** La probabilidad de que unos valores caigan en la región de rechazo n un contraste de hipótesis, a pesar de que  $H_0$  sea verdadera, se conoce con el nombre de nivel de significación  $\alpha$ ,

**Nominal.** Una escala **nominal** es una escala de medición en la cual los números sirven como “etiquetas” solamente para identificar o clasificar un objeto.

**Ordenada en el origen.** Es la ordenada del punto donde una recta corta al eje y.

**Parámetro.** Es cualquier valor característico de la población. Ejemplo: la media de la población, la desviación típica de la población.

**Pendiente.** Es la razón de cambio de la variación de la variable dependiente, en función de la variación de la variable independiente.

**Percentil.** Es el valor de la variable que deja por debajo un porcentaje determinado de observaciones.

**Periodo de referencia.** Intervalo entre dos momentos determinados, a los que se hace alusión en la información estadística solicitada. [INEGI. 2012. Planeación y control administrativo en proyectos estadísticos.]

**Población.** Es cualquier conjunto de unidades o elementos claramente definido, en el espacio y el tiempo, donde los elementos pueden ser personas, granjas, hogares, y cualquier otro. Las poblaciones pueden ser finitas e infinitas.

**Procesamiento.** Serie de actividades mediante las cuales se ordenan, almacenan y preparan los archivos con la información captada, asegurando su congruencia a fin de proceder a su explotación para la presentación de resultados estadísticos.

**Poisson.** Esta distribución, llamada de los sucesos raros, es el caso límite de la binomial, con las siguientes condiciones: a) El número de intentos  $n$  debe tender a infinito. b) La propiedad  $p$  debe ser muy pequeña (de ahí el nombre de suceso raro) c) El producto de  $n \cdot p$  ha de ser constante, y al que llamaremos  $m$ .

**Porcentaje.** Proporción de una cantidad respecto de otra referida al valor cien como unidad.

**P-valor.** Proporción de una cantidad respecto de otra referida al valor cien como unidad.

**Prueba de F.** Prueba estadística que sirve para comparar varianzas. El estadístico  $F$  experimental es el estadístico de contraste en el ANOVA y otras pruebas de comparación de varianzas.

**Rango.** Al intervalo entre el valor máximo y el valor mínimo; por ello, comparte unidades con los datos. Permite obtener una idea de la dispersión de los datos, cuanto mayor es el rango, más dispersos están los datos de un conjunto.

**Regresión.** Proceso de estimación de un valor de una variable a partir del valor de otra. Si el proceso se realiza gráficamente a través de una curva, a ésta se le denomina “Curva de

Regresión”, y si se realiza a través de una ecuación, a esta se le denomina “Ecuación de Regresión”.

**Resultado.** Una salida de la ejecución de procesos y actividades de dirección de proyectos.

**Relación.** Una Relación entre un Conjunto A y un Conjunto B es cualquier conjunto de parejas ordenadas de  $A \times B$ . Se llama dominio de la relación al conjunto de las primeras componentes de las parejas y rango o imagen de la relación al conjunto de las segundas componentes.

**Sumas de cuadrados.** Es la suma de las diferencias al cuadrado entre los datos experimentales y su media.

**Supuesto.** Es una afirmación que se hace de una población en la Estadística Inferencial: si es simétrica, normal, continua... y sobre la muestra, si es aleatoria simple, es de tamaño mayor que 30....

**T de Student.** Distribución que sigue la estimación de la desviación típica.

**Teorema.** Es una proposición teórica, enunciado o fórmula que incorpora una verdad, axioma o postulado que es comprobada por otros conjuntos de teorías o fórmulas

**Test.** Son el instrumento para validar o rechazar las hipótesis de modelación probabilistas.

**Tipificación.** Es la operación de convertir una medida en típica restándole la media y dividiendo entre la desviación típica.

**Uniforme.** es el modelo (absolutamente) continuo más simple. Corresponde al caso de una variable aleatoria que sólo puede tomar valores comprendidos entre dos extremos a y b.

**Variable.** Es una característica de la población o de la muestra cuya medida puede cambiar de valor. Se representa simbólicamente mediante las letras del alfabeto. Según su naturaleza puede ser cualitativa y cuantitativa.

**Varianza.** Medida de variabilidad que consiste en la distancia media de las distancias de las puntuaciones en una variable hasta su media aritmética al cuadrado.

***Para más conceptos de términos estadístico visiten.***

<https://glosarios.servidor-alicante.com/terminos-estadistica>