Universidad de Santander
UDES

VIGILADA MINEDUCACIÓN | SNIES 2832

Estadística

Estadística descriptiva

Daniel Martínez Bello

Universidad de Santander
Maestría en Biotecnología

Octubre de 2025

Que es estadística

Estadística se refiere a los métodos científicos para recolectar, organizar, resumir, presentar y analizar datos, asi como obtener conclusiones validas y hacer decisiones razonables sobre la base de tales análisis.

Para que sirve la estadística

Sin técnicas estadísticas las conclusiones de estudios experimentales y observacionales no tienen validez

El estudio experimental y el estudio observacional

La diferencia principalmente es que el estudio experimental ofrece la posibilidad de determinar la relación causa y efecto entre una o mas variables predictoras o explicatorias o independientes y una variable respuesta. Con el estudio observacional podemos establecer una asociación entre las variables explicatorias o predictoras y la variable respuesta, pero no una causación.

Que es un experimento?

- "Un experimento es una operación conducida bajo condiciones controladas para descubrir un efecto previamente desconocido, para probar o establecer una hipótesis o para demostrar un ley conocida".
- "El objetivo de un experimento es aclarar la relación entre las condiciones controlables y el resultado de un experimento."
- "Un análisis experimental se realiza sobre observaciones que son afectadas no solamente por las condiciones controladas, sino también por condiciones no controladas y errores de medida."
- (The Concise Encyclopedia of Statistics, 2008).

Objetivo de la experimentación

Obtener conclusiones acerca de la pregunta científica planteada por un investigador, basado en un método científico que implica controlar al máximo la variabilidad de los datos generados por un experimento

Población

- La colección completa de objetos, personas, animales o entidades individuales mínimas de las cuales el investigador desea obtener información.
- Ejemplos
- El investigador desea conocer algun atributo específico de la población.
- Por ejemplo
- Podemos definir como un “modelo” de la población, la colección de números para cada individuo en la población real.

Parámetro

- Un investigador requiere conocer la distribución de los numeros de la población.
- El investigador esta interesado en los parametros de una población.
- Los parámetros son números asociados con la distribución de la población, son números que resumen la distribución de una población.
- Ejemplos: Media, mediana, desviación estándar.

Parámetro

- En muchos estudios no es factible obtener información acerca de todas la unidades de una población.
- La población puede tener características que impiden obtener datos para la población entera.
- Ejemplo: la población puede ser muy grande, o muy pequeña , o estar muy dispersa, o puede ser muy costoso recoger los datos para la población completa.

Muestra

- Una muestra es una parte de la población.
- Una muestra se toma para conocer alguna característica de una población.
- Ejemplo: La población de estudiantes de la Universidad.
- La población de niños de tercer año de primaria en la ciudad de Bucaramanga.

Estadístico

- Un estadístico de una muestra es una característica numérica, es un número, que resume la distribución de una muestra.
- La media de la muestra, la mediana de la muestra y la desviación standard de la muestra.
- Un estadístico tiene una relación similar a una muestra, como un parámetro es a una población.
- La proporción de mujeres fumadoras en la facultad de veterinaria.
- El promedio de gastos en servicios psicológicos de una muestra de IPS en Santander.

Inferencia estadística

- Inferencia estadística es realizar una declaración, una afirmación un enunciado acerca de un parámetro poblacional basado en un estadístico de una muestra.
- Para que la inferencia estadística sea valida la muestra debe ser representativa de la población
- La representatividad de la muestra esta dada por la similitud de la distribución de la muestra con respecto a la distribución de la población de la cual proviene.
- Sesgo por muestreo es una tendencia a colectar una muestra no representativa, generando inferencias no validas.

Letras griegas

α alfa	β beta	γ gama	δ delta
τ tau	ϵ epsilon	σ sigma	ϕ fi
χ chi	λ lambda	μ mu	ω omega
θ teta	ρ ro	ν nu	η eta
κ kappa	υ upsilon	ξ xi	ψ psi
ι iota	ζ zeta		

Notación

- Una población se denota con una letra mayúscula.
- Los parámetros de una población se denotan con una letra griega.
- Una muestra se denota con una letra mayúscula.
- Un estadístico se denota con una letra mayúscula o minúscula.

Variables

- Es un atributo de una población o una muestra.
- Es la representación de una característica de una población o una muestra que toma diferentes valores.
- Una variable se denota con una letra mayúscula o minúscula.
- Ejemplos: Peso, edad, color de ojos, tiempo al momento en que una intervención psicológica es efectiva, proporcion de uso de drogas blandas en una población de universitarios, dinero gastado en servicios veterinarios por una comunidad, cantidad de productos lácteos producidos en una región.

Tipos de variables

- Variables continuas.
- Es una variable cuyos valores se encuentran en un intervalo de valores.
- Peso, edad, nivel de una sustancia en la sangre, resultados de una prueba psicológica.
- Variables discretas (conocidas tambien como variables categóricas)
- Es una variable cuyos valores se definen en categorías.
- Presencia o ausencia de una trastorno de personalidad, sexo, estrato socio-económico.

Toma de datos

Formularios, registros, encuestas, estudios, pruebas psicológicas, observación .

Distribución de frecuencias

- Para resumir una cantidad grande de datos se recurre a organizarlos en clases o categorias y determinar el número de individuos en cada clase o sea la frecuencia de clase.
- Un arreglo en forma de tabla de los datos de acuerdo a sus frecuencias se conoce como distribución de frecuencias o tabla de frecuencias.

Ejemplo

449, 400, 459, 428, 359,
328, 330, 355, 432, 330,
324, 354, 385, 445, 330,
420, 350, 415, 375, 435

Distribución de frecuencias

- Se ordenan los datos de menor a mayor

324	$y_{[1]}$	385	$y_{[11]}$
328	$y_{[2]}$	400	$y_{[12]}$
330	$y_{[3]}$	415	$y_{[13]}$
330	$y_{[4]}$	420	$y_{[14]}$
330	$y_{[5]}$	428	$y_{[15]}$
350	$y_{[6]}$	432	$y_{[16]}$
354	$y_{[7]}$	435	$y_{[17]}$
355	$y_{[8]}$	445	$y_{[18]}$
359	$y_{[9]}$	449	$y_{[19]}$
375	$y_{[10]}$	459	$y_{[20]}$

Distribución de frecuencias

- Se definen intervalos de clase, intervalos iguales entre el valor superior y el valor inferior de todos los datos ordenados.
- Se seleccionan límites de clase, valores que representan límite superior e inferior para cada intervalo de clase.

	Límite inferior	Límite superior
Intervalo de clase uno	320	339
Intervalo de clase dos	340	359
Intervalo de clase tres	360	379
Intervalo de clase cuatro	380	399
Intervalo de clase cinco	400	419
Intervalo de clase seis	420	439
Intervalo de clase siete	440	459

Distribución de frecuencias

- El tamaño o anchura de un intervalo de clase se determinan de acuerdo al nivel de precisión requerido, aunque pueden ser de 5 a 20, pero no es una norma rígida.
- Despues de definidos los intervalos de clase, se procede a contar las observaciones que caen en cada intervalo de clase.
- Cada conteo en cada intervalo de clase se conoce como la frecuencia de observaciones en cada intervalo.
- La suma acumulada de las frecuencias de acuerdo a cada intervalo de clase configura la frecuencia acumulada de observaciones en cada intervalo.
- El valor medio entre el l'imite superior y el límite inferior para cada intervalo de clase se conoce como Marca de Clase.

Distribución de frecuencias

- Una forma de establecer el número de intervalos es establecer el valor máximo y mínimo del conjunto de datos, y subdividir en intervalos iguales el espacio entre estos valores máximo y mínimo.
- Por ejemplo, si el valor mínimo es 12 y el valor máximo es 28, entonces se pueden hacer intervalos de tres en tres iniciando en 10 y terminando en 30. De tal forma quedan los intervalos $[10-13), [13-16), [16-19), [19-22), [22-25), [25-28), [28-31)$
- Otra forma para este mismo ejemplo podría ser, iniciar en 12 y terminar en 28 o en número ligeramente superior, por ejemplo,y con intervalos de 3 unidades, esto es, $[12-15), [15-18), [18-21), [21-24), [24-27), [27-30)$

Distribución de frecuencias

Resultados de pesos de bovinos hembras

Tabla de frecuencias.

Límite inferior	Límite superior	Marca de Clase
320	339	329.5
340	359	349.5
360	379	369.5
380	399	389.5
400	419	409.5
420	439	429.5
440	459	449.5

Distribución de frecuencias

- Se ordenan los datos de menor a mayor

324	$y_{[1]}$	385	$y_{[11]}$
328	$y_{[2]}$	400	$y_{[12]}$
330	$y_{[3]}$	415	$y_{[13]}$
330	$y_{[4]}$	420	$y_{[14]}$
330	$y_{[5]}$	428	$y_{[15]}$
350	$y_{[6]}$	432	$y_{[16]}$
354	$y_{[7]}$	435	$y_{[17]}$
355	$y_{[8]}$	445	$y_{[18]}$
359	$y_{[9]}$	449	$y_{[19]}$
375	$y_{[10]}$	459	$y_{[20]}$

Distribución de frecuencias
Tabla de frecuencias.

Límite inferior	Límite superior	Frecuencia
320	339	5
340	359	4
360	379	1
380	399	1
400	419	2
420	439	4
440	459	3

Distribución de frecuencias

Tabla de frecuencias y frecuencias acumuladas

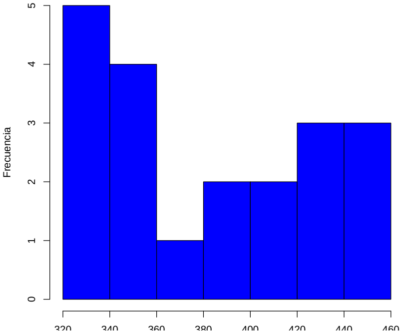
Límite inferior	Límite superior	Frecuencia	Frecuencia Acumulada
320	339	5	5
340	359	4	9
360	379	1	10
380	399	1	11
400	419	2	13
420	439	4	17
440	459	3	20

Distribución de frecuencias

- Se puede obtener una impresión gráfica de la distribución de frecuencias a traves de los histogramas, polígonos de frecuencias.
- Un histograma consiste en un gráfico de barras en la cual cada barra representa la frecuencia de observaciones en cada intervalo de clase.
- Un polígono de frecuencias consiste en unir todos las marcas de clase en los techos de los histogramas
- Un ojiva o gráfica de frecuencias acumuladas se produce utilizando las frecuencias acumuladas.

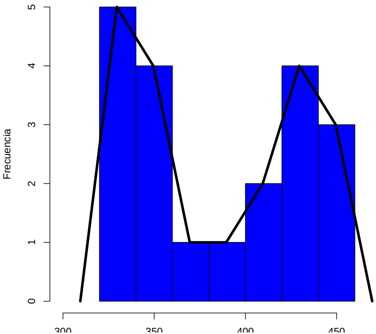
Distribución de frecuencias

El histograma es una forma de representar las frecuencias de conteo en una forma gráfica.

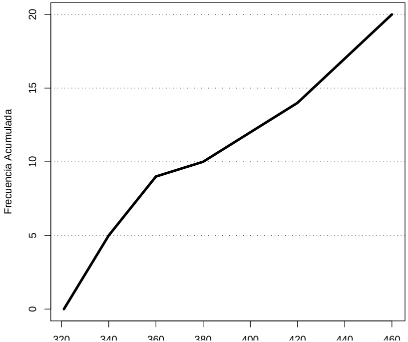


Distribución de frecuencias

El polígono de frecuencias. El polígono se prolonga en los extremos inferior y superior hasta las marcas de clase inferior y superior inmediatas.



Distribución de frecuencias
Gráfica de frecuencias acumuladas.



Medidas de tendencia central

- La media de la población
- La media de la muestra.
- La mediana de la población
- La mediana de la muestra
- La proporción de la población
- La proporción de la muestra

La media de la población

- La media de la población, los datos los tenemos de acuerdo a como los recolectamos, y N es el número total de datos de la población.

$$y_1, y_2, \cdots, y_N$$

- Donde el subíndice indica el orden en que los datos fueron recogidos
- Sumamos los datos y los dividimos por el número total de datos

$$\mu = \frac{\sum_{i=1}^{N} Y}{N}$$

Media de la Muestra

- La media de la muestra, los datos los tenemos de acuerdo a como los recolectamos

$$y_1, y_2, \cdots, y_n$$

- Donde el subíndice indica el orden en que los datos fueron recogidos
- Sumamos los datos y los dividimos por el numero total de datos

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y}{n} = \frac{7703}{20} = 385{,}15$$

La proporción de la población

- Para la proporción de la población los datos consisten en una coleccion de 0 y 1 donde el valor 1 corresponde a un individuo de la población que presenta la característica de interes y el valor 0 corresponde a un individuo de la población que no tiene la característica de interes

$$1, 0, 1, 0, 1, 0, 0, 1, 1, 1$$

- Sumamos los datos y los dividimos por el número total de datos de la población N.

$$\pi = \frac{\sum_{i=1}^{N} Y_i}{N} = \frac{6}{10} = 0{,}6$$

- La proporción toma valores entre 0 y 1, o puede ser expresado en porcentaje multiplicando el valor por 100.

La proporción de la muestra

- Para la proporción de la muestra los datos consisten en una colección de 0 y 1 donde el valor 1 corresponde a un individuo de la población que presenta la característica de interes y el valor 0 corresponde a un individuo de la población que no tiene la característica de interes

$$1, 0, 1, 0, 1, 0, 0$$

- Sumamos los datos y los dividimos por el número total de datos de la muestra n.

$$P = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{3}{7} = 0{,}43$$

- La proporción toma valores entre 0 y 1, o puede ser expresado en porcentaje multiplicando el valor por 100.

Media para datos agrupados

- Se utilizan las marcas de clase de cada intervalo de clase de una tabla de frecuencias se denotan Y_i donde $i = 1, \cdots, n$, y n es el numero de intervalos.
- Se selecciona la marca de clase central, o la marca de clase con frecuencia mas alta, este valor se denota como A.
- Se calculan las desviaciones de todas las marcas de clase con respecto a la marca de clase central, y se denotan como d_i, entonces $d_i = Y_i - A$ donde Y_i es el valor de cada marca de clase.

Media para datos agrupados

- Se multiplica cada frecuencia f_i por cada desviación d_i y se denota como fd_i.
- La media para datos agrupados es igual a

$$\bar{Y} = A + \frac{\sum_{i=1}^{n} fd_i}{N}$$

- Donde N es el número total de observaciones

Media para datos agrupados
Tabla de frecuencias.

	Límite inferior	Límite superior	Marca de Clase
	320	339	329.5
	340	359	349.5
	360	379	369.5
A $\rightarrow$	380	399	389.5
	400	419	409.5
	420	439	429.5
	440	459	449.5

Media para datos agrupados

	Marca de Clase	$d_i = Y_i - A$	f_i	fd_i
	329.5	-60	5	-300
	349.5	-40	4	-160
	369.5	-20	1	-20
A $\rightarrow$	389.5	0	1	0
	409.5	20	2	40
	429.5	40	4	160
	449.5	60	3	180

$$\bar{Y} = A + \frac{\sum_{i=1}^{n} fd_i}{N} = 389{,}5 + \frac{-100}{20} = 384{,}5$$

Mediana de la población

- Los datos los tenemos de acuerdo a como los recolectamos

$$y_1, y_2, \cdots, y_n$$

- Donde el subíndice indica el orden en que los datos fueron recogidos
- Primero ordenamos los datos de menor a mayor

$$y_{[1]}, y_{[2]}, \cdots, y_{[n]}$$

- Donde el subíndice indica el orden de los datos de menor a mayor asi [1] es el menor, [2] es el siguiente , hasta$[n]$ que es el numero mayor

Mediana de la población

- La mediana se determina como el valor medio que se encuentra en los datos ordenados.

$$mediana = y_{\left[\frac{n+1}{2}\right]}$$

Si el subíndice no es entero se toma el promedio entre los dos datos con subíndices mas cercanos al subíndice que corresponde a la mediana.

Se ordenan los datos de menor a mayor

324	$y_{[1]}$	385	$y_{[11]}$
328	$y_{[2]}$	400	$y_{[12]}$
330	$y_{[3]}$	415	$y_{[13]}$
330	$y_{[4]}$	420	$y_{[14]}$
330	$y_{[5]}$	428	$y_{[15]}$
350	$y_{[6]}$	432	$y_{[16]}$
354	$y_{[7]}$	435	$y_{[17]}$
355	$y_{[8]}$	445	$y_{[18]}$
359	$y_{[9]}$	449	$y_{[19]}$
375	$y_{[10]}$	459	$y_{[20]}$

Ejemplo: mediana de la población

- Para estimar la mediana

$$mediana = y_{[\frac{20+1}{2}]} = y_{[10,5]}$$

- Como el subíndice no es entero se toman los datos con subíndice $y_{[10]}$ correspondiente a 375 y $y_{[11]}$ correspondiente a 385, y se toma un valor promedio

$$mediana = \frac{(375 + 385)}{2} = 380$$

- La mediana de la población es de 380 kilogramos

Mediana de la Muestra: el mismo procedimiento para la mediana de la población.

Moda

- La moda de una serie de numeros es aquel valor que se presenta con la mayor frecuencia, es decir, es el valor mas comun. La moda puede no existir, o puede no ser unica.
- 3, 5, 6, 4, 5, 6, 5, 4, 3, 5, 5, 6, 8
- La moda de los datos es 5.

Cuartiles, deciles y percentiles

- En una serie de datos ordenados, de menor a mayor, el valor medio (que divide los datos en dos partes iguales es la mediana)
- Se puede pensar en otros valores que dividen los datos ordenados en cuatro partes iguales. Estos valores representados por las letras Q_1, Q_2 y Q_3 se denominan primer, segundo y tercer quartiles.
- Q_2 es igual a la mediana.
- Los valores que dividen los datos en 10 partes iguales se conocen como deciles $(D_1 \cdots D_9)$, y cuando los datos se dividen en cien partes iguales se conocen como percentiles $(P_1 \cdots P_9)$

Medidas de dispersión

- Rango
- Rango intercuartílico.
- Varianza
- Desviación estándar

Rango

- Primero ordenamos los datos de menor a mayor

$$y_{[1]}, y_{[2]}, \cdots, y_{[n]}$$

- Donde el subindice indica el orden de los datos de menor a mayor asi [1] es el menor, [2] es el siguiente , hasta$[n]$ que es el numero mayor
- Se selecciona el valor menor $y_{[1]}$ y el valor mayor $y_{[n]}$.
- Se resta el valor mayor $y_{[n]}$ del menor valor $y_{[1]}$, y se obtiene el rango

Ejemplo: rango

Se ordenan los datos de menor a mayor

324	$y_{[1]}$	385	$y_{[11]}$
328	$y_{[2]}$	400	$y_{[12]}$
330	$y_{[3]}$	415	$y_{[13]}$
330	$y_{[4]}$	420	$y_{[14]}$
330	$y_{[5]}$	428	$y_{[15]}$
350	$y_{[6]}$	432	$y_{[16]}$
354	$y_{[7]}$	435	$y_{[17]}$
355	$y_{[8]}$	445	$y_{[18]}$
359	$y_{[9]}$	449	$y_{[19]}$
375	$y_{[10]}$	459	$y_{[20]}$

El valor menor es 324 y el valor mayor es 459. El rango es 135.

Rango intercuartílico

- Primero ordenamos los datos de menor a mayor

$$y_{[1]}, y_{[2]}, \cdots, y_{[n]}$$

- Donde el subindice indica el orden de los datos de menor a mayor asi [1] es el menor, [2] es el siguiente , hasta $[n]$ que es el numero mayor
- Se seleccionan los valores que corresponden al 25 % inferior y al 75 % superior

$$Cuartil_{25} = y_{\left[\frac{n+1}{4}\right]}, Cuartil_{75} = y_{\left[\frac{3(n+1)}{4}\right]}$$

Ejemplo: Rango intercuartílico

- Primero ordenamos los datos de menor a mayor $y_{[1]}, y_{[2]}, \cdots, y_{[n]}$
- Se seleccionan los valores que corresponden al 25 % inferior y al 75 % superior

$$Cuartil_{25} = y_{\left[\frac{20+1}{4}\right]} = y_{[5,5]}$$

$$Cuartil_{75} = y_{\left[\frac{3(20+1)}{4}\right]} = y_{[15,75]}$$

- Los valores mas cercanos a $y_{[5,5]}$ son $y_{[5]}$ (330) y $y_{[6]}$ (350), el valor promedio entre los dos es 340, este corresponde al cuartil 25 %. Los valores mas cercanos a $y_{[15,75]}$ son $y_{[15]}$ (428) y $y_{[16]}$ (432), el valor promedio entre los dos es 430, este corresponde al cuartil 75 %.

Varianza de la población

- Los datos los tenemos de acuerdo a como los recolectamos

$$y_1, y_2, \cdots, y_n$$

- Donde el subíndice indica el orden en que los datos fueron recogidos
- La varianza nos informa acerca de la dispersión o variabilidad de los datos, se denota con la letra griega σ^2.

$$\sigma^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n}$$

Ejemplo: varianza de la población

Y	$\bar{Y}$	$(Y - \bar{Y}_i)$	$(Y_i - \bar{Y})^2$
324	385,2	-61,2	3739,3
328	385,2	-57,2	3266,1
330	385,2	-55,27	3041,5
330	385,2	-55,2	3041,5
330	385,2	-55,2	3041,5
350	385,2	-35,2	1235,5
354	385,2	-31,2	970,3
355	385,2	-30,2	909
359	385,2	-26,2	683,8
375	385,2	-10,2	103

Ejemplo: varianza de la población

Y	$\bar{Y}$	$(Y - \bar{Y}_i)$	$(Y_i - \bar{Y})^2$
385	385,2	-0,1	0
400	385,2	14,9	220,5
415	385,2	29,9	891
420	385,2	34,9	1214,5
428	385,2	42,9	1836,1
432	385,2	46,9	2194,9
435	385,2	49,9	2485
445	385,2	59,9	3582
449	385,2	63,9	4076,8
459	385,2	73,9	5453,8

Ejemplo: varianza de la población

$$\sigma^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n} = \frac{41986,6}{20} = 2099,3 \tag{1}$$

Desviación estándar de la población

La desviación standard tiene la misma interpretación que la varianza, nos indica la variabilidad de los datos (la variabilidad de una población o de una muestra), se denota con la letra griega σ

$$\sigma = \sqrt{\sigma^2}$$

Ejemplo:Desviación estándar de la población

$$\sigma = \sqrt{\sigma^2} = \sqrt{2099{,}3} = 45,82 \text{ Kilogramos}$$

Varianza de la muestra

- Los datos los tenemos de acuerdo a como los recolectamos

$$y_1, y_2, \cdots, y_n$$

- Donde el subíndice indica el orden en que los datos fueron recogidos
- La varianza de la muestra indica la variabilidad de los datos alrededor de la media de la muestra.

$$S^2 = \frac{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}{n-1}$$

Ejemplo: varianza de la muestra

$$S^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1} = \frac{41986,6}{19} = 2209,8$$

Desviación estándar de la muestra

La desviación standard es la raíz cuadrada de la varianza.

$$S = \sqrt{S^2}$$

Desviación estándar de la muestra

$$S = \sqrt{S^2} = \sqrt{2209{,}8} = 47,01 \text{ Kilogramos}$$

Coeficiente de variación

El coeficiente de variación es una medida absoluta de dispersión y sirve para comparar la dispersión entre grupos.

$$CV = \frac{s}{\bar{y}}$$

$$CV = \frac{47,01}{385,15} = 0{,}1212\,\%$$

Varianza de la proporción

- Para la proporción de la población los datos consisten en una colección de 0 y 1 donde el valor 1 corresponde a un individuo de la poblacion que presenta la caracteristica de interes y el valor 0 corresponde a un individuo de la poblacion que no tiene la caracteristica de interes

$$1, 0, 1, 0, 1, 0, 0, 1, 1, 1$$

$$\sigma_{\pi}^{2} = \pi(1 - \pi) = 0{,}6 * 0{,}4 = 0{,}24$$

- La varianza de la proporción toma valores entre 0 y 1, o puede ser expresado en porcentaje multiplicando el valor por 100.

Varianza de la proporción

- Para la proporción de la muestra los datos consisten en una coleccion de 0 y 1 donde el valor 1 corresponde a un individuo de la poblacion que presenta la caracteristica de interes y el valor 0 corresponde a un individuo de la poblacion que no tiene la caracteristica de interes

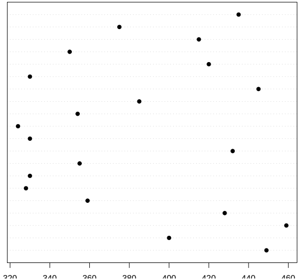
$$1, 0, 1, 0, 1, 0, 0, 1, 1, 1$$

$$S_P^2 = P(1-P) = 0{,}6 * 0{,}4 = 0{,}24$$

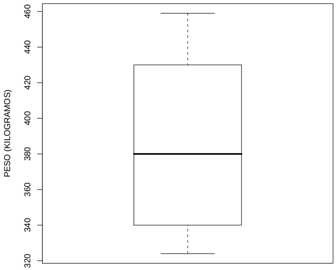
- La varianza de la proporción de la muestra toma valores entre 0 y 1, o puede ser expresado en porcentaje multiplicando el valor por 100.

- Gráfico de puntos (Dotplot)
- Gráfico de cajas y bigotes (Box and Whiskers plot, Box plot)
- Gráfico (Stem and Leaf plot)
- Tabla de frecuencias
- Histograma
- Polígono de frecuencias acumuladas

Gráfico de puntos



Cajas y bigotes



Raíz y hojas (stem and leaf plot)

Gráfico de raices y hojas . El punto decimal queda 2 digitos a la derecha de la barra vertical.

3	23333
3	556689
4	022334
4	556

Proporciones

Proporciones en la población.

La proporción en la población se refiere al numero de elementos en la población entera perteneciendo a una categoria de interes, dividido por el numero total de elementos en la población. La proporción de la población se denota con la letra griega π.

- Ejemplos
- La proporción de personas con depresión en una región.
- La proporción de personas con habilidades superiories de lecto-escritura pero con baja habilidad matemática en una institución de educación superior.

Proporciones en la población.

Como se presentan los datos para mirar una proporción.

- La proporción de la población se estima a partir de la proporción de la muestra.

$$P = \frac{\sum_{i=1}^{n} X_i}{n}$$