

Notas de un curso elemental de Estadística

- *Nombre Archivo:* 2020:04:02:Dia_12:41:30:Hora
 - *Palabra clave:* Meme
 - *Descripción:* Perfecta para hacer memes de **El estadista responde.**



- *Nombre Archivo:* 2020:04:02:Dia_12:53:58:Hora
 - *Palabra clave:* censo , encuesta
 - *Descripción:* **Censo:** Aplicación del instrumento estadístico para obtener información al total de la población. **Encuesta:** Aplicación del instrumento estadístico para obtener información a una muestra de la población.



- *Nombre Archivo:* 2020:04:02:Dia_12:58:46:Hora
 - *Palabra clave:* SuperEdición
 - *Descripción:* Imagen que muestra la altura de las diferentes versiones de *Gozilla* con sus respectivas siluetas en un eje que permite, a simple vista, distinguir las alturas. Gran trabajo de diseño.



- *Nombre Archivo:* 2020:04:02:Dia_12:59:04:Hora
 - *Palabra clave:* lista
 - *Descripción:* Lista de todas las versiones de *Gozilla*: Altura, largo y peso

1. Altura 50 metros, largo 122 metros, peso 20,000 toneladas.
2. Altura 50 metros, largo 100 metros, peso 20,000 toneladas.
3. Altura 50 metros, largo 100 metros, peso 20,000 toneladas.
4. Altura 50 metros, largo 100 metros, peso 20,000 toneladas.
5. Altura 50 metros, largo 100 metros, peso 20,000 toneladas.
6. Altura 80 metros, largo 190 metros, peso 50,000 toneladas.
7. Altura 80 metros, largo 190 metros, peso 50,000 toneladas.
8. Altura 100 metros, largo 200 metros, peso 60,000 toneladas.
9. Altura 100 metros, largo 200 metros, peso 60,000 toneladas.
10. Altura 60 metros, largo 120 metros, peso 500 toneladas.
11. Altura 55 metros, largo 122.5 metros, peso 25,000 toneladas.
12. Altura 60 metros, largo 120 metros, peso 30,000 toneladas.
13. Altura 55 metros, largo 122.5 metros, peso 25,000 toneladas.
14. Altura 100 metros, largo 200 metros, peso 55,000 toneladas.
15. Altura 108 metros, largo 274 metros, peso 55,000 toneladas.
16. Altura 118.5 metros, largo 333 metros, peso 92,000 toneladas.
17. Altura 50 metros, peso 10,000 toneladas.
18. Altura 318 metros, peso 100,000 toneladas.

- *Nombre Archivo:* 2020:04:02:Dia_13:00:41:Hora
 - *Palabra clave:* tabla, representaciones tabulares
 - *Descripción:* Tabla en formato lista con los datos de las alturas de *Gozilla*.



GODZILLA	
Cases (18 cases)	
Index	Altura (m)
1	50
2	50
3	50
4	50
5	50
6	50
7	55
8	55
9	60
10	60
11	80
12	80
13	100
14	100
15	100
16	108
17	118.5
18	318

- *Nombre Archivo:* 2020:04:02:Dia_13:01:09:Hora
 - *Palabra clave:* tabla, representaciones tabulares
 - *Descripción:* Tabla en formato lista con los datos de las alturas de *Gozilla*, con los datos repetidos agrupados.



GODZILLA		
Cases (8 cases)		
Index	Altura (m)	Frecuencia
1	50	6
2	55	2
3	60	2
4	80	2
5	100	3
6	108	1
7	118.5	1
8	318	1

- *Nombre Archivo:* 2020:04:02:Dia_13:10:16:Hora
 - *Palabra clave:* tabla, representaciones tabulares
 - *Descripción:* Tabla con los datos agrupados y grafico con las siluetas de los *Gozillas* de diferentes películas.



- *Nombre Archivo:* 2020:04:02:Dia_17:23:31:Hora
 - *Palabra clave:* frecuencias absolutas acumuladas
 - *Descripción:* Es el primer item.

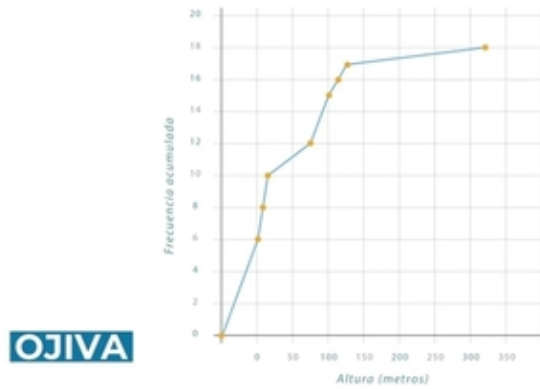
¿Qué nos muestra la columna de frecuencias absolutas acumuladas?

☐ El número de observaciones que son iguales o menores a un determinado valor de variable.

☐ El número de veces que se presenta un valor de variable.

OMITIR ENVIAR

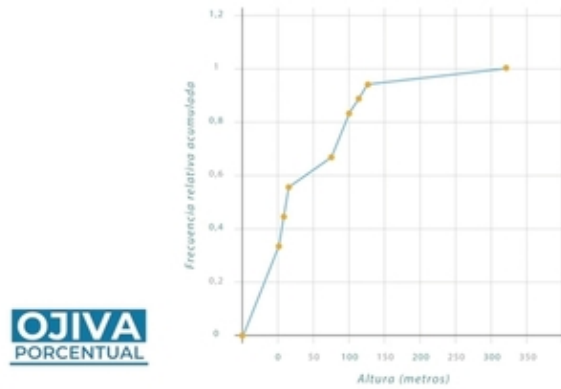
- *Nombre Archivo:* 2020:04:02:Dia_17:24:02:Hora
 - *Palabra clave:* ojiva, frecuencia acumulada
 - *Descripción:* Grafica que muestra como se ha realizado el grafico de una **ojiva** con la ayuda de las **frecuencia absolutas acumuladas**



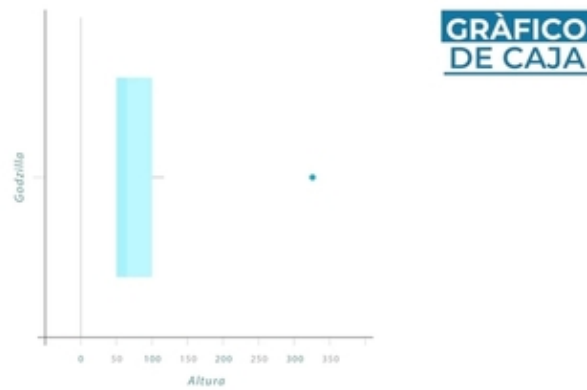
- *Nombre Archivo:* 2020:04:02:Dia_17:25:30:Hora
 - *Palabra clave:* tabla, frecuencia relativa, frecuencia absoluta, frecuencia relativa acumulada
 - *Descripción:* Tabla mejorada que muestra las frecuencia: relativa, absoluta y absoluta acumulada.

GODZILLA					
Versiones (8 cases)					
Index	Altura (m)	Frecuencia	F. Acumulada	F. Relativa	F. R. A.
1	50	6	6	0.33	0.33
2	55	2	8	0.11	0.44
3	60	2	10	0.11	0.56
4	80	2	12	0.11	0.67
5	100	3	15	0.17	0.83
6	108	1	16	0.06	0.89
7	118.5	1	17	0.06	0.94
8	318	1	18	0.06	1

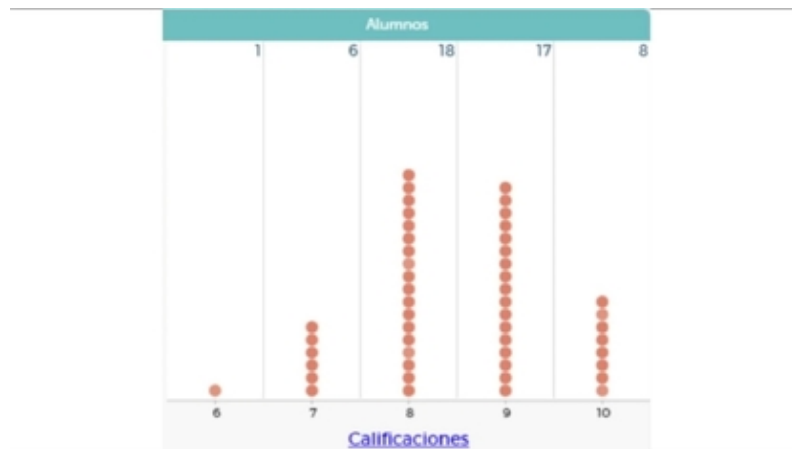
- *Nombre Archivo:* 2020:04:02:Dia_17:26:41:Hora
 - *Palabra clave:* ojiva porcentual
 - *Descripción:* Muestra la **ojiva porcentual** de la tabla de datos anterior.



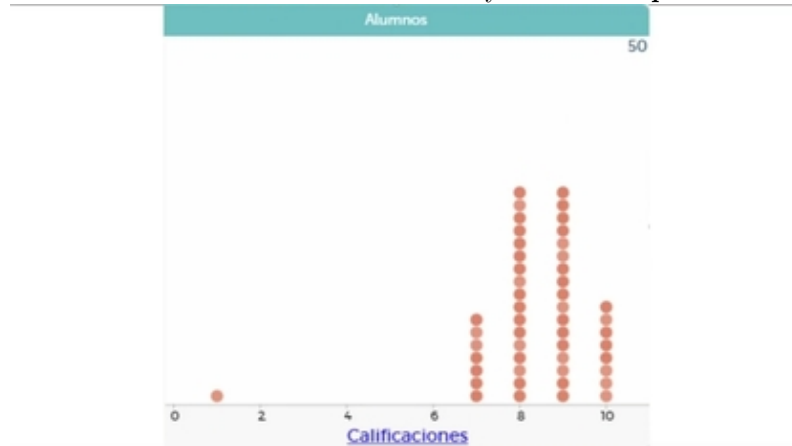
- *Nombre Archivo:* 2020:04:02:Dia_17:27:05:Hora
 - *Palabra clave:* gráfico de caja, datos atípicos
 - *Descripción:* Imagen que muestra el **grafico de caja** de los datos se puede ver con el punto en azul más fuerte, la presencia de *datos atípicos*



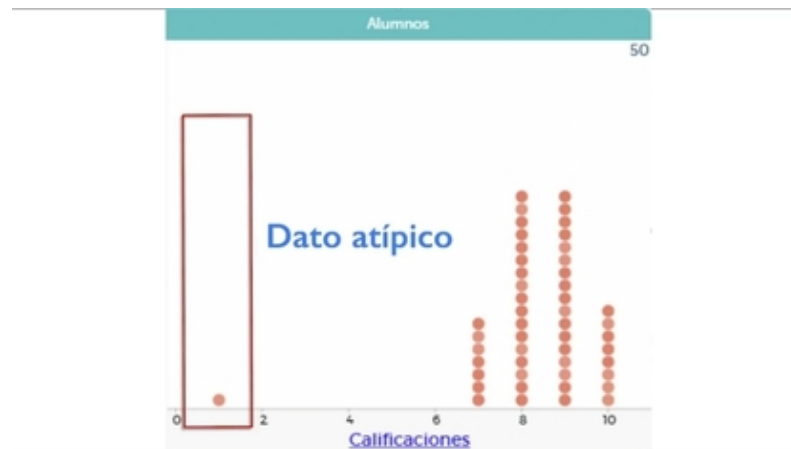
- *Nombre Archivo:* 2020:04:02:Dia_17:28:37:Hora
 - *Palabra clave:* histograma
 - *Descripción:* **Histograma** de alguna de las actividades propuestas, observamos una **distribución normal asimétrica**.



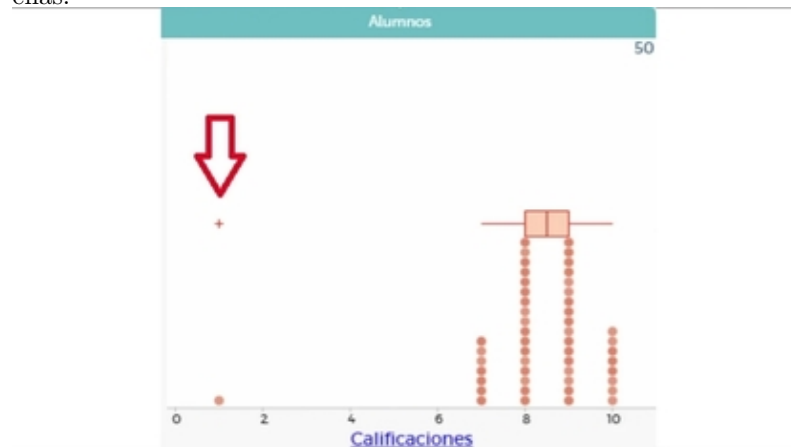
- *Nombre Archivo:* 2020:04:02:Dia_17:29:01:Hora
 - *Palabra clave:* histograma, datos atípicos
 - *Descripción:* **Histograma** de alguna de las actividades propuestas, observamos una **distribución normal** y un **dato atípico** en 1.



- *Nombre Archivo:* 2020:04:02:Dia_17:29:18:Hora
 - *Palabra clave:* histograma, datos atípicos
 - *Descripción:* **Histograma** de alguna de las actividades propuestas, observamos una **distribución normal** y un **dato atípico** en 1. Ahora se remarca el dato atípico con un recuadro rojo.



- *Nombre Archivo:* 2020:04:02:Dia_17:29:24:Hora
 - *Palabra clave:* histograma, datos atípicos, grafico de caja
 - *Descripción:* Datos anteriores, ahora sale junto a el **histograma** el **grafico de cajas** observamos una **distribución normal** y un **dato atípico** senyalado con una flecha roja. Observamos que las dos representaciones de los datos ofrecen una información coherente entre ellas.



- *Nombre Archivo:* 2020:04:02:Dia_17:50:24:Hora
 - *Palabra clave:* tabla, tabla variables cualitativas
 - *Descripción:* Tabla que muestra la representación resumida, en una tabla que podria responder a las pregunta, ¿Siglo XX? Sí o no.



GODZILLA		
Siglo	XX	11
	XXI	7
Column Summary		18

`S1 = count ()`

- *Nombre Archivo:* 2020:04:02:Dia_17:59:32:Hora
 - *Palabra clave:* probabilidad
 - *Descripción:* **Probabilidad:** Ciencia que se encarga de estudiar los fenómenos aleatorios y que sirve como soporte para la estadística.



PROBABILIDAD

Ciencia que se encarga de estudiar los fenómenos aleatorios y que sirve como soporte para la Estadística.

- *Nombre Archivo:* 2020:04:02:Dia_17:59:41:Hora
 - *Palabra clave:* probabilidad, fenómeno aleatorio
 - *Descripción:* **Fenómeno aleatorio:** Fenómeno que tiene más de un posible resultado, que no se puede predecir y que depende del azar.



PROBABILIDAD

Ciencia que se encarga de estudiar los fenómenos aleatorios y que sirve como soporte para la Estadística.

FENÓMENO ALEATORIO

Fenómeno que tiene más de un posible resultado, que no se puede predecir y que depende del azar.

- *Nombre Archivo:* 2020:04:02:Dia_18:00:09:Hora
 - *Palabra clave:* estadística
 - *Descripción:* **Estadística:** Ciencia que se encarga de recabar, organizar, presentar y analizar datos numéricos para poder predecir situaciones futuras.



ESTADÍSTICA

Ciencia que se encarga de recabar, organizar, presentar y analizar datos numéricos para poder predecir situaciones futuras.

- *Nombre Archivo:* 2020:04:02:Dia_18:02:25:Hora
 - *Palabra clave:* muestra
 - *Descripción:* **Muestra:** Sector de la población dentro del cual se pretende replicar el comportamiento de la variable de interés.



MUESTRA

Sector de la población dentro del cual se pretende replicar el comportamiento de la variable de interés.

- *Nombre Archivo:* 2020:04:03:Dia_02:14:11:Hora
 - *Palabra clave:* moda
 - *Descripción:* **Moda:** Observación que tiene la frecuencia absoluta mayor.



MODA

GODZILLA		
Cases (8 cases)		
Index	Altura (m)	Frecuencia
1	50	6
2	55	2
3	60	2
4	80	2
5	100	3
6	108	1
7	118.5	1
8	318	1

Observación que tiene la frecuencia absoluta mayor.

- *Nombre Archivo:* 2020:04:03:Dia_16:00:44:Hora
 - *Palabra clave:* mediana, hallar mediana
 - *Descripción:* Indicación visual de donde se encuentra la **mediana**, divide las frecuencias absolutas en dos mitades.



- *Nombre Archivo:* 2020:04:03:Dia_16:03:37:Hora
 - *Palabra clave:* media aritmética
 - *Descripción:* Resultado de la media aritmética de los valores de la tabla. En este caso $\bar{X} = 85.25$



- *Nombre Archivo:* 2020:04:03:Dia_16:05:06:Hora
 - *Palabra clave:* media, moda, mediana
 - *Descripción:* Comparación de la (50) **moda** \hat{X} , (60) **mediana** \tilde{X} y (85.25) la **media** \bar{X} , de los valores de la tabla del inicio.



MODA

$\hat{X} = 50$

MEDIANA

$\hat{X} = 60$

MEDIA

$\bar{X} = 85.25$

- *Nombre Archivo:* 2020:04:03:Dia_16:24:57:Hora
 - *Palabra clave:* frecuencia relativa acumulada, percentil, calculo de un percentil, semejanza de triángulos
 - *Descripción:* calculo del valor asociado a un percentil mediante semejanza de triángulos



- *Nombre Archivo:* 2020:04:03:Dia_16:27:43:Hora
 - *Palabra clave:* cuantiles, cuartiles
 - *Descripción:* 25%, 50%, 75%, 100% son cuantiles, concretamente, cuartiles.



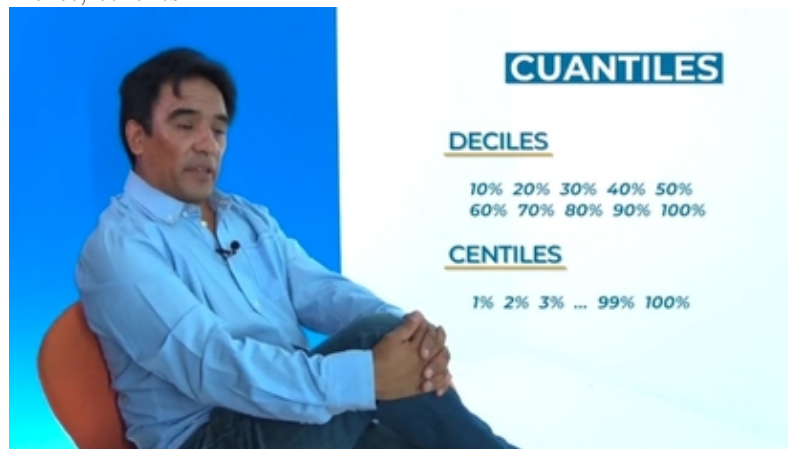
- *Nombre Archivo:* 2020:04:03:Dia_16:28:52:Hora
 - *Palabra clave:* cuantiles, cuartiles, quintiles
 - *Descripción:* 20%, 40%, 60%, 80%, 100% son cuantiles, concretamente, quintiles.



- *Nombre Archivo:* 2020:04:03:Dia_16:29:18:Hora
 - *Palabra clave:* cuantiles, deciles
 - *Descripción:* 10%, 20%, 30%, 40%, ..., 90%, 100% son cuantiles, concretamente, deciles.



- *Nombre Archivo:* 2020:04:03:Dia_16:29:53:Hora
 - *Palabra clave:* cuantiles, deciles, centiles
 - *Descripción:* 1%, 2%, 3%, ..., 99%, 100% son cuantiles, concretamente, centiles.



- *Nombre Archivo:* 2020:04:03:Dia_22:32:23:Hora
 - *Palabra clave:* error
 - *Descripción:* **Error:** Diferencia de algún valor observado de una variable con respecto de algún valor fijo.



- *Nombre Archivo:* 2020:04:03:Dia_22:35:19:Hora
 - *Palabra clave:* tabla con error
 - *Descripción:* Tabla con una columna para los errores.

GODZILLA			
Versiones (8 casos)			
Index	Altura (m)	Frec.	Error (m)
1	50	6	-35.25
2	55	2	-20.25
3	60	2	-25.25
4	80	2	-5.25
5	100	3	14.75
6	108	1	22.75
7	118.5	1	33.25
8	318	1	232.75

- *Nombre Archivo:* 2020:04:03:Dia_22:40:10:Hora
 - *Palabra clave:* varianza discreta, error, error cuadrado, error cuadrado por frecuencia relativa
 - *Descripción:* Tabla con error, error cuadrado, error cuadrado por frecuencia relativa añadidos.

VARIANZA

GODZILLA					
Versiones (8 cses)					
Index	Altura (m)	Frec.	Error (m)	Error Cuadrado	Error C. por Frec. (m)
1	50	6	-35.25	1242.5625	7455.375
2	55	2	-20.25	915.0625	1830.125
3	60	2	-25.25	637.5625	1275.125
4	80	2	-5.25	27.5625	55.125
5	100	3	14.75	217.5625	652.6875
6	108	1	22.75	517.5625	517.5625
7	118.5	1	33.25	1105.5625	1105.5625
8	318	1	232.75	54172.5625	54172.5625

67064.125/17=3944.9485

- Nombre Archivo: 2020:04:03:Dia_22:40:32:Hora
 - Palabra clave: varianza discreta, error, error cuadrado
 - Descripción: Fórmulas error y error cuadrado.

VARIANZA

1. Calcular los errores de cada observación respecto de la media.

$$x_i - \bar{x}$$

2. Elevar al cuadrado estos errores.

$$(x_i - \bar{x})^2$$

- Nombre Archivo: 2020:04:03:Dia_22:40:50:Hora
 - Palabra clave: calculo rebuscado varianza discreta, varianza
 - Descripción: Tercer paso del calculo de la varianza discreta n-1.

VARIANZA

3. Multiplicar los cuadrados de los errores por sus respectivas frecuencias absolutas

$$f(x_i - \bar{x})^2$$

- *Nombre Archivo:* 2020:04:03:Dia_22:41:01:Hora
 - *Palabra clave:* calculo rebuscado varianza, varianza
 - *Descripción:* Cuarto paso del calculo de la **varianza discreta n-1**.

VARIANZA

5. Multiplicar los cuadrados de los errores por sus respectivas frecuencias absolutas

$$f(x_i - \bar{x})^2$$

4. Sumar todos estos productos.

$$\sum_{i=1}^n f(x_i - \bar{x})^2$$

- *Nombre Archivo:* 2020:04:03:Dia_22:41:13:Hora
 - *Palabra clave:* calculo rebuscado varianza, varianza
 - *Descripción:* Quinto paso del calculo de la **varianza discreta n-1**.

VARIANZA

5. Dividir entre $n - 1$

$$\frac{\sum_{i=1}^n f(x_i - \bar{x})^2}{n - 1}$$

- *Nombre Archivo:* 2020:04:03:Dia_22:42:34:Hora
 - *Palabra clave:* amplitud
 - *Descripción:* **Amplitud:** Diferencia entre el mayor valor esperado y el menor valor esperado. ($A=M-m$)



- *Nombre Archivo:* 2020:04:03:Dia_22:42:41:Hora
 - *Palabra clave:* amplitud
 - *Descripción:* Ejemplo calculo amplitud.



AMPLITUD

Amplitud = Mayor valor de variable - Menor valor de variable

318-50=268

- *Nombre Archivo:* 2020:04:03:Dia_22:55:28:Hora
 - *Palabra clave:* coeficiente de variación
 - *Descripción:* Relación que tiene la desvinculación estándar con respecto de la media.



COEFICIENTE DE VARIACIÓN

Relación que tiene la desviación estándar con respecto de la media.

$$CV = \frac{S}{\bar{X}}$$

- *Nombre Archivo:* 2020:04:03:Dia_22:55:33:Hora
 - *Palabra clave:* coeficiente de variación
 - *Descripción:* **coeficiente de variación:** cociente de la raíz cuadrada de la varianza entre la media aritmética.



- *Nombre Archivo:* 2020:04:03:Dia_22:56:00:Hora
 - *Palabra clave:* coeficiente de variación
 - *Descripción:* **Comentario Coeficiente:** Para poderlo utilizar necesitamos cumplir con dos condiciones. **1.** La primera es que las dos poblaciones tengan más o menos la misma distribución para la variable con la que se está trabajando. La segunda **2.** condición es que en ambos casos se utilice la misma variable, es decir, podremos comparar variabilidad entre los pesos de los tigres y los pesos de los gatos, pero no la variabilidad entre las longitudes de los tigres y los pesos de los gatos a través del coeficiente de variación.



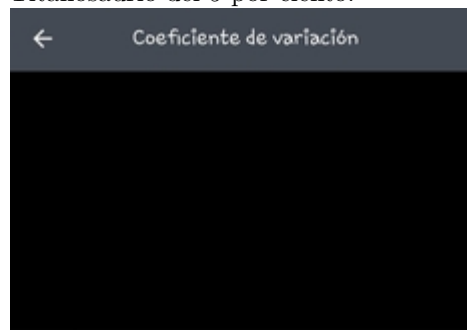
Su cálculo es, pues, el cociente de la desviación estándar entre la media. Para poderlo utilizar necesitamos cumplir con dos condiciones. La primera es que las dos poblaciones tengan más o menos la misma distribución para la variable con la que se está trabajando.

La segunda condición es que en ambos casos se utilice la misma variable, es decir, podremos comparar variabilidad entre los pesos de los tigres y los pesos de los gatos, pero no la variabilidad entre las longitudes de los tigres y los pesos de los gatos a través del coeficiente de variación.

- *Nombre Archivo:* 2020:04:03:Dia_23:02:28:Hora
 - *Palabra clave:* coeficiente de variación, coeficiente de variación ejemplo
 - *Descripción:* En la imagen se puede ver como *cv* para el Paravicursor es de $0.13/0.39=0.333$, mientras que para el Titanosaurio, el *cv* es de $2/40=0.05$.



- *Nombre Archivo:* 2020:04:03:Dia_23:03:19:Hora
 - *Palabra clave:* coeficiente de variación, coeficiente de variación ejemplo
 - *Descripción:* **Interpretación de los resultados del cv para el ejemplo:** Observando ambos valores, encontramos que el Parvicursor tienen un coeficiente de variación mayor que el del Titanosaurio, por lo que podemos decir que, para la variable altura, el Parvicursor presenta **mayor variabilidad** que el Titanosaurio. También es posible que presentemos el **coeficiente de variación como un porcentaje**. En el caso Del Paravicursor será del 33 por ciento y en el caso del Titanosaurio del 5 por ciento.



coeficiente de variación será 2 sobre 40
que es igual a 0,5.

Observando ambos valores, encontramos
que el parvicursor tiene un coeficiente de
variación mayor que el del titanosaurio, por
lo que podemos decir que, para la variable
altura, el parvicursor presenta mayor
variabilidad que el titanosaurio.

También es posible que presentemos
el coeficiente de variación como un
porcentaje. En el caso del parvicursor
será del 33 por ciento y en el caso del
titanosaurio del 5 por ciento.

- *Nombre Archivo:* 2020:04:04:Dia_23:18:38:Hora
 - *Palabra clave:* regla empírica
 - *Descripción:* Wikipedia: *En estadística, la regla 68-95-99.7, también conocida como regla empírica, es una abreviatura utilizada para recordar el porcentaje de valores que se encuentran dentro de una banda alrededor de la media en una distribución normal con un ancho de dos, cuatro y seis veces la desviación típica, respectivamente. Más exactamente, el 68.27%, el 95.45% y el 99.73% de los valores se encuentran dentro de bandas con semiancho de una, dos y tres veces*

la desviación típica respecto a la media.. En la imagen se muestra: media más menos una desviación estándar ~68% y media más menos dos desviaciones estándar ~95%.

1. Si a la media le restamos y le sumamos una vez el valor de la desviación estándar para crear un intervalo, este intervalo contendrá a aproximadamente el 68% de todas las observaciones:

$$\begin{array}{ll} 171.2 - 15.4 = 155.8 & [155.8, 186.6] \\ 171.2 + 15.4 = 186.6 & \end{array}$$

2. Si a la media le restamos y le sumamos dos veces la desviación estándar para crear un intervalo, este intervalo contendrá a aproximadamente el 95% de las observaciones:

$$\begin{array}{ll} 171.2 - (2 \times 15.4) = 140.4 & [140.4, 202] \\ 171.2 + (2 \times 15.4) = 202 & \end{array}$$

- *Nombre Archivo:* 2020:04:04:Dia_23:19:45:Hora
 - *Palabra clave:* regla empírica
 - *Descripción:* En la imagen se muestra: media más menos *tres* desviación estándar ~99% y media más menos *cuatro* desviaciones estándar ~0.999936657516334%.

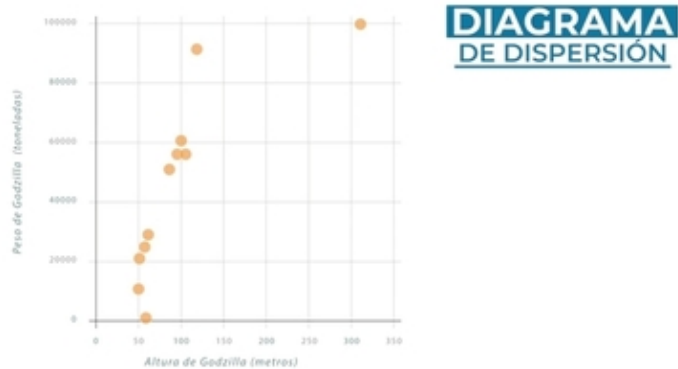
3. Si a la media le restamos y le sumamos tres veces la desviación estándar para crear un intervalo, este intervalo contendrá a aproximadamente el 99% de las observaciones:

$$\begin{array}{ll} 171.2 - (3 \times 15.4) = 125 & [125, 217.4] \\ 171.2 + (3 \times 15.4) = 217.4 & \end{array}$$

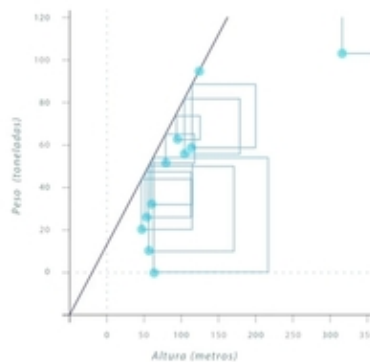
4. Si a la media le restamos y le sumamos tres veces y media la desviación estándar, en el intervalo que se construya se tendrá el total de las observaciones:

$$\begin{array}{ll} 171.2 - (3.5 \times 15.4) = 117.3 & [117.3, 225.1] \\ 171.2 + (3.5 \times 15.4) = 225.1 & \end{array}$$

- *Nombre Archivo:* 2020:04:05:Dia_12:32:42:Hora
 - *Palabra clave:* Diagrama de dispersión, Nube de puntos
 - *Descripción:* Wikipedia: *Un diagrama de dispersión o gráfica de dispersión o gráfico de burbujas es un tipo de diagrama matemático que utiliza las coordenadas cartesianas para mostrar los valores de dos variables para un conjunto de datos.*



- *Nombre Archivo:* 2020:04:05:Dia_12:37:25:Hora
 - *Palabra clave:* Diagrama de dispersión, Recta de regresión, Recta de mejor ajuste, recta de ajuste, método de mínimos cuadrados
 - *Descripción:* En la imagen se muestran los puntos en color azul claro del diagrama de dispersión en el plano, también se muestra una recta en negro (que será la recta de regresión), finalmente en ese fragmento del primer cuadrante del plano con los ejes podemos observar cuadrados con perímetro azul oscuro que representan el cuadrado de la distancia mínima de cada punto a la recta de regresión. Una forma de encontrar la recta que mejor se ajusta es el *método de mínimos cuadrados*



- *Nombre Archivo:* 2020:04:05:Dia_13:04:10:Hora
 - *Descripción:*
 - *Descripción:* Diferentes valores necesarios para calcular la recta de ajuste o de regression. Suma de todas las x_i , suma de todas las y_i , suma del producto de $x_i \cdot y_i$, suma de los cuadrados de x_i , suma de los cuadrados de y_i .

GODZILLA					
Indice	X	Y	X*Y	X.cuadrado	Y.cuadrado
1	50	20000	1000000	2500	400000000
2	50	20000	1000000	2500	400000000
3	50	20000	1000000	2500	400000000
4	50	20000	1000000	2500	400000000
5	50	20000	1000000	2500	400000000
6	50	50000	5000000	2500	1000000000
7	55	50000	1375000	3025	625000000
8	55	40000	1375000	3025	625000000
9	60	60000	3600000	3600	2100000000
10	60	500	1800000	3600	900000000
11	80	25000	4000000	6400	2500000000
12	80	30000	4000000	6400	2100000000
13	100	25000	6000000	10000	3600000000
14	100	55000	6000000	10000	3600000000
15	100	55000	5500000	10000	3025000000
16	108	92000	9940000	11664	3025000000
17	118.5	10000	10952000	14042.25	8444000000
18	318	100000	31800000	101124	10000000000

$$\sum_{i=1}^{18} x_i = 1534.50$$

$$\sum_{i=1}^{18} y_i = 712500$$

$$\sum_{i=1}^{18} x_i y_i = 84222000$$

$$\sum_{i=1}^{18} x_i^2 = 197880.25$$

$$\sum_{i=1}^{18} y_i^2 = 40964250000$$

197880.25 y para los cuadrados de "y", 40,964,250,000.

- **Nombre Archivo:** 2020:04:05:Dia_13:13:46:Hora
 - **Palabra clave:** coeficiente de correlación, coeficiente de correlación lineal, coeficiente de Pearson
 - **Descripción:** Mide el grado de asociación lineal entre las variables sin verse afectado por las unidades.



- **Nombre Archivo:** 2020:04:06:Dia_19:49:05:Hora
 - **Palabra clave:** estimación, regression, no lineal, limitaciones de los modelos lineales, correlación causalidad, dependencia
 - **Descripción:** Por ejemplo dos variables podrían comportarse de manera conjunta como una función cuadrática, como una logarítmica o como una función exponencial. Es importante resaltar que correlación no quiere decir ni *causalidad* ni *dependencia*. Que hay *correlación* entre dos variables, únicamente significa que ambas varían más o menos de la misma forma. No significa que la variación en una de estas variables, provoque la variación en la otra y tampoco significa que los valores de una de estas variables, dependan de los valores que tome la otra.



Por ejemplo dos variables podrían comportarse de manera conjunta como una función cuadrática, como una logarítmica o como una función exponencial. Es importante resaltar que correlación no quiere decir ni causalidad ni dependencia.

Que haya correlación entre dos variables, únicamente significa que ambas varían más o menos de la misma forma.

No significa que la variación en una de estas variables, provoque la variación en la otra y tampoco significa que los valores de una de estas variables, dependan de los valores que tome la otra. [MÚSICA]

- *Nombre Archivo:* 2020:04:06:Dia_19:52:15:Hora
 - *Palabra clave:* Enunciado problema Profesor Lewis Dodgson
 - *Descripción:* Lewis Dodgson tenía 18 amigas. Nueve eran rubias y nueve morenas, once tenían los ojos oscuros y siete tenían los ojos azules. ¡Ah, y seis eran rubias y de ojos oscuros! ¿Cómo eran todas las amigas del professor Dodgson?. ***Observamos** que hay dos variables: color de los ojos y color del cabello.



"Lewis Dodgson tenía 18 amigas.
Nueve eran rubias y nueve morenas,
once tenían los ojos oscuros y siete
tenían los ojos azules.
¡Ah, y seis eran rubias y de ojos oscuros!
¿Cómo eran todas las amigas del profesor
Dodgson?"

tenemos involucradas dos variables,
el color de ojos y el color de cabello.

- *Nombre Archivo:* 2020:04:06:Dia_19:52:42:Hora
 - *Palabra clave:* Tabla cruzada, Tabla de contingencia
 - *Descripción:* Tabla de contingencia del problema de las amigas del Profesor Dodgson.

TABLA CRUZADA

AMIGAS		Cabello		
		Morena	Rubia	Row Summary
Ojos	Azules	4	3	7
	Oscuros	5	6	11
Column Summary		9	9	18

S1 = count ()

tabla de contingencia o tabla cruzada.

- *Nombre Archivo:* 2020:04:06:Dia_19:55:07:Hora
 - *Palabra clave:* Tabla cruzada, Tabla de contingencia, Otras Tablas
 - *Descripción:* Tabla de contingencia del problema de las amigas del Profesor Dodgson esta vez en formato _simil histograma.



- *Nombre Archivo:* 2020:04:06:Dia_20:03:44:Hora
 - *Palabra clave:* Tabla de contingencia, Ejercicio propuesto, Otro acertijo Lewis Carroll
 - *Descripción:* El señor Charles Lutwidge Carroll coleccionaba marionetas y tenía veinte. Ocho eran de animales y doce de personas, seis eran grandes y ocho eran de personas y pequeñas. ¿Cómo eran todas las marionetas del señor Carroll?

Para el siguiente acertijo de Lewis Carroll, construye la tabla de doble entrada con frecuencias absolutas.

"El señor Charles Lutwidge Carroll coleccionaba marionetas y tenía veinte. Ocho eran de animales y doce de personas, seis eran grandes y ocho eran de personas y pequeñas. ¿Cómo eran todas las marionetas del señor Carroll?"

	Animal	Persona	TOTAL
Chica			
Grande			
TOTAL			

Escribir el número que va en la casilla que se encuentra iluminada:

	Animal	Persona	TOTAL
Chica			
Grande			
TOTAL			

NOTA: La respuesta es un número sin espacio ni

- *Nombre Archivo:* 2020:04:07:Dia_16:55:27:Hora
 - *Palabra clave:* Enfoque subjetivo
 - *Descripción:* Lo usual bajo el enfoque subjetivo es asignar la medida de probabilidad, no en términos numéricos, sino simplemente en

términos de más probable, menos probable, altamente probable o poco probable.



**ENFOQUE
SUBJETIVO**

- *Nombre Archivo: 2020:04:07:Dia_16:56:10:Hora*
 - *Palabra clave: Enfoque frecuencial*
 - *Descripción: El siguiente enfoque para calcular medidas de probabilidad se conoce como enfoque frecuencial. En este caso, para poder asignar una medida de probabilidad a algún resultado en particular del fenómeno aleatorio lo que se hace es contabilizar observaciones sobre repeticiones sucesivas del fenómeno.*



**ENFOQUE
FRECUENCIAL**

- *Nombre Archivo: 2020:04:07:Dia_16:57:05:Hora*
 - *Palabra clave: Enfoque clásico*
 - *Descripción: El tercer enfoque es el enfoque clásico de la probabilidad. En este caso, calcularemos el cociente del total de formas en las que puede darse un resultado en específico del fenómeno aleatorio sobre el total de formas en que puede ocurrir el fenómeno aleatorio.*



**ENFOQUE
CLÁSICO**

- *Nombre Archivo:* 2020:04:07:Dia_17:01:52:Hora
 - *Palabra clave:* espacio muestral
 - *Descripción:* Wikipedia: el conjunto de todos los posibles resultados de un experimento aleatorio, junto con una estructura sobre el mismo. Pensar en el ejemplo de los posibles escenarios al tirar un dado. A veces también se denota el conjunto con una letra omega majúscula.



**ESPACIO
MUESTRAL**

$S = \{1, 2, 3, 4, 5, 6\}$

- *Nombre Archivo:* 2020:04:07:Dia_17:02:45:Hora
 - *Palabra clave:* evento
 - *Descripción:* En el ejemplo de tirar el dado nos referimos al **evento** de que salga 5. Podemos pensar un evento, como un subconjunto del espacio muestral.



EVENTOS

$A=\{5\}$

- *Nombre Archivo:* 2020:04:07:Dia_17:04:52:Hora
 - *Palabra clave:* eventos
 - *Descripción:* Más ejemplos de eventos del *ejemplo de lanzar un dado*, el evento A se puede leer cómo: *que salga un número par*. El evento B se puede leer cómo: “que salga un número menor que 5”.



EVENTOS

$A=\{2, 4, 6\}$

$B=\{1, 2, 3, 4\}$

- *Nombre Archivo:* 2020:04:07:Dia_17:07:36:Hora
 - *Palabra clave:* Eventos compuestos, Unión de eventos, Conjunción de eventos, Complemento de eventos
 - *Descripción:* La **unión** de dos eventos A y B, se puede interpretar cómo *pasa A o pasa B*, la **intersección o conjunción** de dos eventos A y B se puede leer cómo: *sucede A y también sucede B*. Finalmente el **complemento** de un evento A se interpreta cómo: *No pasa A*



UNIÓN DE EVENTOS

$$(A \cup B) = \{1, 2, 3, 4, 6\}$$

CONJUNCIÓN

$$(A \cap B) = \{2, 4\}$$

COMPLEMENTO

$$\{1, 3, 5\}$$

$$A', A^c, \sim A \text{ o } \bar{A}$$

- *Nombre Archivo:* 2020:04:07:Dia_17:14:04:Hora
 - *Palabra clave:* Tabla de contingencia de probabilidades, problema Profesor Lewis Dodgson
 - *Descripción:* Ejemplo de tabla de contingencia como probabilidades del ejemplo de *Profesor Dodgson*.



CONJUNTAS

AMIGAS	Moreno	Rubio	Total
Azules	0.22	0.17	0.39
Oscuros	0.28	0.33	0.61
Total	0.50	0.50	1

- *Nombre Archivo:* 2020:04:07:Dia_17:29:22:Hora
 - *Palabra clave:* Tabla de contingencia de probabilidades, problema Profesor Lewis Dodgson
 - *Descripción:* Tabla de contingencia de probabilidades del ejemplo de *Profesor Dodgson*, con el calculo explicito de *La probabilidad de tener cabello rubio u ojos claros*, nótese que es una unión de eventos.

AMIGAS		Cabello		
Ojos		Moreno	Rubio	Row Summary
		Azules	Oscuros	
	Azules	4	3	7
	Oscuros	5	6	11
Column Summary		9	9	18

S1 = count ()

$$P(\text{Cabello rubio u Ojos oscuros}) = \frac{11+9-6}{18} = \frac{14}{18}$$

- Nombre Archivo: 2020:04:07:Dia_17:29:52:Hora
 - Palabra clave: Tabla de contingencia de probabilidades, problema Profesor Lewis Dodgson, Grassmann
 - Descripción: Símil de fórmula de Grassmann para calcular la probabilidad de que en escoger una amiga al azar del professor Dodgson y que cumpla que *cabello rubio u ojos oscuros*.

$$P(\text{Cabello rubio u Ojos oscuros}) = P(\text{Cabello rubio}) + P(\text{Ojos oscuros}) - P(\text{Cabello rubio y ojos oscuros})$$

AMIGAS		Cabello		
Ojos		Moreno	Rubio	Row Summary
		Azules	Oscuros	
	Azules	4	3	7
	Oscuros	5	6	11
Column Summary		9	9	18

S1 = count ()

- Nombre Archivo: 2020:04:07:Dia_17:35:00:Hora
 - Palabra clave: evento o sucesos imposible, evento o sucesos seguro
 - Descripción: En el *ejemplo del dado* un **evento imposible** podría ser *obtener un 7* (recordemos que el espacio muestral es de 1 a 6). En el mismo ejemplo un **evento seguro** es obtener un número menor que 7.



IMPOSIBLE

$S=\{1, 2, 3, 4, 5, 6\}$
 $P(\text{Obtener } 7)=0/6=0$

SEGURO

$S=\{1, 2, 3, 4, 5, 6\}$
 $P(\text{Número menor a } 7)=6/6=1$

- *Nombre Archivo:* 2020:04:07:Dia_17:41:39:Hora
 - *Palabra clave:* axiomas de Kolomogorov, probabilidades marginales
 - *Descripción:* **1.** Probabilidad valor entre 0 y 1. **2.** Probabilidad de todos el espacio muestral igual a 1. **3.** La probabilidad de la unión de una serie de eventos del mismo espació muestral es igual a la suma de las probabilidades marginales de cada uno de los eventos, siempre que todos los eventos sean mutuamente excluyentes y del mismo espació muestral.

AXIOMAS DE KOLMOGOROV

1. Un valor de probabilidad siempre estará entre 0 y 1.

$$0 \leq P(E) \leq 1$$

2. La probabilidad de todo el espacio muestral siempre es igual a 1.

$$P(S)=1$$

3. La probabilidad de la unión de una serie de eventos del mismo espacio muestral es igual a la suma de las probabilidades marginales de cada uno de los eventos, siempre que todos los eventos sean mutuamente excluyentes y del mismo espacio muestral.

$$P(E_1 \cup E_2 \cup \dots \cup E_n) = P(E_1) + P(E_2) + \dots + P(E_n)$$

- *Nombre Archivo:* 2020:04:09:Dia_12:39:19:Hora
 - *Palabra clave:* Tabla de contingencia de probabilidades, problema Profesor Lewis Dodgson, probabilidad condicional
 - *Descripción:* Con la misma tabla de contingencia, se enuncia que significado se dan a los valores A y B que sirvan para el posterior calculo de la probabilidad condicionada.



AMIGAS	Moreno	Rubio	Total
Azules	0.22	0.17	0.39
Oscuros	0.28	0.33	0.61
Total	0.50	0.50	1

A= rubia
B= ojos azules

- *Nombre Archivo:* 2020:04:09:Dia_12:39:45:Hora
 - *Palabra clave:* Tabla de contingencia de probabilidades, problema Profesor Lewis Dodgson, probabilidad condicional
 - *Descripción:* Misma Imagen que antes pero con titulo identificativo.



PROBABILIDAD CONDICIONAL

AMIGAS	Moreno	Rubio	Total
Azules	0.22	0.17	0.39
Oscuros	0.28	0.33	0.61
Total	0.50	0.50	1

- *Nombre Archivo:* 2020:04:09:Dia_12:48:43:Hora
 - *Palabra clave:* Probabilidad condicional
 - *Descripción:* Tabla con cuatro formulas para la probabilidad condicional de eventos *dependientes* y *independientes*.

PROBABILIDAD CONDICIONAL

$$P(A|B) = P(A \text{ y } B) / P(B)$$

$$P(A \text{ y } B) = P(A|B)P(B)$$

$$P(A|B) = P(A)$$

$$P(A \text{ y } B) = P(A)P(B)$$