FORMULARIO

Muestreo

POBLACIONES INFINITAS (n≥100,000)	POBLACIONES FINITAS (n<100,000 es decir 99,000)
$n_0 = \frac{Z^2 * P * Q}{E^2}$ **Si se dan proporciones** $n_0 = \frac{Z^2 * \sigma^2}{E^2}$ **Si se dan cantidades**	$n_0 = \frac{N * Z^2 * P * Q}{(N-1)E^2 + (Z^2 * P * Q)}$ **Si se dan proporciones** $n_0 = \frac{N * Z^2 * \sigma^2}{(N-1)E^2 + (Z^2 * \sigma^2)}$ **Si se dan cantidades**

Dónde:

n_0= Tamaño mínimo de muestra
N= Tamaño de la población
Z= Nivel de confianza en valores Z, p.ej. 95% ---- Z= 1.96
P = probabilidad de éxito o preferencia, en valor relativo, p.ej. 60% ---- 0.6
Q = probabilidad de fracaso o No preferencia, en valor relativo, p.ej. 40 % ---- 0.4
E= error muestral , en valor relativo, p.ej. 3% de error máximo ----- 0.03

Cuartiles y Percentiles

Datos Sueltos		Datos agrupados
Percentiles Exactos	Cálculo exacto	
$P_k = \frac{k\,(n+1)}{100}$ **Dónde:** P = Percentil n = tamaño de la muestra k = percentil deseado	$Q1 = \frac{(n+1)}{4}$ $Q3 = \frac{3(n+1)}{4}$	$Q_1 = L_i + \left(\frac{\frac{n}{4} - fa}{f}\right)A$ $Q_3 = L_i + \left(\frac{\frac{3n}{4} - fa}{f}\right)A$

Agrupación de datos

Rango (R) = Xmáx. – Xmín (No redondear)
Criterio de Stürges, No. Intervalos = 1 +(3.322 log n)

Amplitud (A ó k) = $\frac{\text{Rango}}{\text{No.intervalos(ya redondeado)}}$ .... (Redondear hacia arriba, según la complejidad del problema)

Por ejemplo, A=3.12255555, si los datos manejados por el problema son:
Enteros, el redondeo queda en 4 (A=3.~~12255555~~, el 3 se lleva a 4)
Si tienen 1 decimal, el redondeo queda en 3.2 (A=3.1~~2255555~~,
Si tienen 2 decimales, el redondeo queda en 3.13 (A=3.12~~255555~~,
Si tienen 3 decimales, el redondeo queda en 3.123 (A=3.122~~55555~~,

Nota: Todo esto se hace con el fin de ser práctico y que el valor máximo siempre sea contabilizado en el último intervalo y este no quede fuera, sin importar nos pasemos un poco.

FUNCIONES EN EXCEL						
PROMEDIO MEDIANA MODA.VARIOS	VAR.S DESVEST.M	ASIME				

Datos sueltos	***Datos agrupados***
Promedio o Media aritmética	
$\bar{x} = \frac{\sum x}{n}$	$\bar{x} = \frac{\sum xf}{n}$

Mediana

$\tilde{X}$ = **Por observación**, es el valor a la mitad de la distribución de datos una vez ordenados de menor a mayor. $Ubicación = \frac{n+1}{2}$ Si son datos pares (promedie bordes) Si son datos nones (saldrá directo)	$\tilde{x} = L_i + \left(\frac{\frac{n}{2} - fa}{f}\right) A$ $Ubicación = \frac{n}{2}$ Dónde: Li = Límite inferior **donde está la mediana** fa= frecuencia acumulada **ANTES** de llegar a la mediana A= Amplitud n = tamaño de la muestra f = frecuencia de la mediana *NOTA: La ubicación de la mediana se selecciona en base a fa*

Moda

$\hat{x}$ = **Por observación,** la frecuencia más alta. NOTA: Puede haber varias modas, tienen que tener el mismo valor númerico y ser las altas. AMODAL= Sin moda UNIMODAL= 1 moda BIMODAL= 2 MODAS MULTIMODAL= 3 ó más modas	$\hat{x} = L_i + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right) A$ Dónde: Li = Límite inferior donde está la moda A= Amplitud f modal= frecuencia de la moda f anterior= frecuencia antes de la moda f superior= frecuencia después de la moda $\Delta_1 = f_{modal} - f_{anterior}$ $\Delta_2 = f_{modal} - f_{superior}$ NOTA: El intervalo que contiene la moda(s) se ubica por tener la frecuencia “f” más alta. Puede haber varias modas, y habrá que repetir esta fórmula “n” veces las haya.

Varianza muestral

Fórmula Original $s^2 = \frac{\sum (x - \bar{x})^2}{n-1}$	Fórmula Alterna $s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$	Fórmula Original $s^2 = \frac{\sum (x - \bar{x})^2 f}{n-1}$	Fórmula alterna $s^2 = \frac{\sum x^2 f - \frac{(\sum xf)^2}{n}}{n-1}$

Desviación estándar muestral

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

Curtosis ó Apuntalamiento

Ap= $\left\{\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{x_j - \bar{x}}{s}\right)^4\right\} - \frac{3(n-1)^2}{(n-2)(n-3)}$	Ap >0 Leptocurtica Ap <0 Platicurtica Ap =0 Mesocurtica	$A_p = \frac{\sum (x - \bar{x})^4 f}{n\,\sigma^4}$ $\sigma^4 = \left(\frac{\sum (x - \bar{x})^2 f}{n}\right)^2$	Ap >3 Leptocurtica Ap <3 Platicurtica Ap =3 Mesocurtica

Asimetría

As= $\frac{n}{(n-1)(n-2)} \sum \left(\frac{x_j - \bar{x}}{s}\right)^3$	As >0 Asimetría **“ +”** As <0 Asimetría **“-“** As = 0 Simétrica	*Pearson* $A_s = \frac{3(\bar{x} - \tilde{x})}{s}$ (media, mediana, desv est muestral) *Bowley* $A_s = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$ *Fischer* $A_s = \frac{\sum (x - \bar{x})^3 f}{n\,\sigma^3}$	As >0 Asimetría **“ +”** As <0 Asimetría **“-“** As = 0 Simétrica

Coeficiente de Variabilidad

$c.v = \frac{s}{\bar{x}} \times 100$ Desviación Estándar, media	$c.v = \frac{s}{\bar{x}} \times 100$ Desviación Estándar, media	0≤ C.V <20 % Muy confiable 20≤ C.V<30 % Confiable 30≤ C.V<40% Relat. Confiable 40% ↑ No confiable