Varianza y desviación estándar

- Indica que tan dispersos están los datos respecto a la media.
- La desviación estándar es la raíz cuadrada de la varianza.

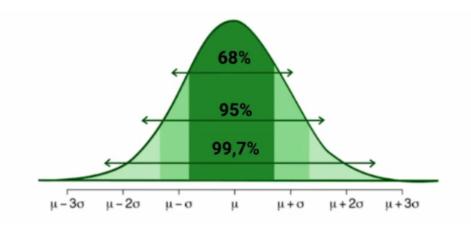
Ejemplo: edades de una población

Cálculo de la varianza y desviación estándar

	Varianza	Desviación estándar	Media		
Población	$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}$	$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}}$	$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$		
Muestra	$s^{2} = \frac{\sum_{i=1}^{n} (x_{i} - \bar{x})^{2}}{n-1}$	$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$		

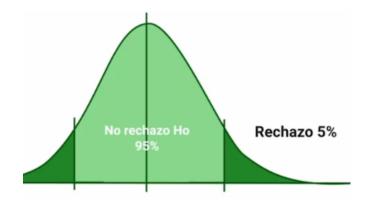
Intervalos de confianza

Son un par o varios pares de números entre los cuales se estima que estará cierto valor desconocido respecto de un parámetro poblacional con un determinado nivel de confianza. Son simétricos respecto a la media.



Nivel de significación

- El nivel de significación o alfa es el nivel límite para juzgar si un resultado es o no es estadísticamente significativo.
- Si el valor de significación es menor que el nivel de significación, el resultado es estadísticamente significativo.



Interpretación de un resultado

Intervalo de confianza del 95%:

Sabemos que con un 95% de certeza las edades de las personas que esquían están entre dos valores.

Ejemplo:

Si tenemos tablas donde tenemos diferentes estudiantes, podemos comparar unos estudiantes que estudian 20 horas, y otros estudiantes que estudian 5 horas. Lo que esperaremos es que la calificación final sea diferente entre ambas poblaciones; en el caso de los que estudian más, esperaremos una calificación más alta, por lo que tendrá un intervalo más estrecho cuando veamos estos índices de confianza, frente a los que estudian menos que podría obtener una calificación más baja, que obtendrán un intervalo más estrecho.

Ejercicio 1: nivel de significación



No conocemos información de nuestra población. Lo que vamos a calcular es el valor mínimo y máximo es que se encuentra la población. Tiene un valor del 95% de incertidumbre.

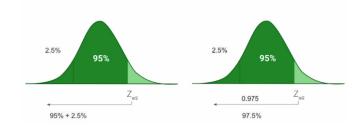
En el lado izquierda, vamos a tener el valor mínimo y en el lado derecho, vamos a tener el valor máximo.

Si lo estamos haciendo sobre el 95%, significa que la concentración tiene el 95% de lo valores y el 5% restante, va a estar distribuido entre el máximo y el mínimo.

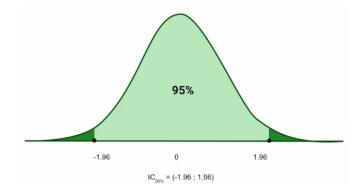
Consideraciones al buscar en la tabla



Al aplicar a nuestro ejercicio, nos quedaría de la siguiente forma:



Resultado



Ejercicio 2:

La duración en días de un cepillo de dientes se ajusta a la distribución normal (28,4). ¿Cuál es el intervalo de confianza al 80%?

$$\mu = 28$$

$$\sigma = 4$$

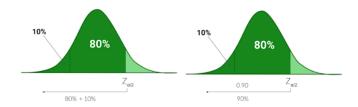
$$80\%$$
de confianza
$$Z_{\alpha/2}$$

Explicando el enunciado, los 28 se refiere al número de días que suele durar un cepillo de dientes, pero tiene una desviación de 4 días. Todo esto lo aplicaríamos al 80% de confianza.

Nivel de significación



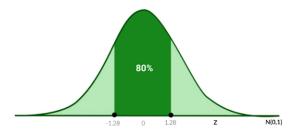
Consideraciones al buscar en la tabla



Búsqueda en la tabla

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
+1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
+1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91308	.91466	.91621	.91774
+1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
+1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
+1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
+1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
+1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
+1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
+2	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
+2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
+2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899

Resultado



Este es el caso de media 0, ahora tenemos que convertirlo a la media 28 del ejercicio con la fórmula: $Z=rac{x-\mu}{\sigma}$

Conversión

