

Estadística para Químicos

Podemos decir, informalmente, que la estadística es el estudio de la **variabilidad**. Sus técnicas permiten describirla y plantear hipótesis con el objetivo de comprenderla.

Si todos opináramos igual respecto a la elección de un determinado candidato, si todos tuviésemos el mismo peso y altura, si todos reaccionáramos de idéntica forma ante un medicamento, o si los productos resultantes de un proceso industrial fuesen idénticos no serían necesarios los estudios estadísticos.

Punto de partida una *población de interés* y una o varias *variables* que se desea analizar en la población.

Una *población* es la totalidad de objetos o individuos de interés en una investigación.

Todos los niños sanos con edad entre 0 y 5 años.

Una **variable** es una característica que varía de individuo en individuo.

edad, peso, altura, sexo concentración de colesterol en sangre, etc.

Los datos son los **valores** de la variable en estudio.

Los datos disponibles se obtienen a partir de una *muestra* de la población de interés, como los valores observados de la o las variables de interés.

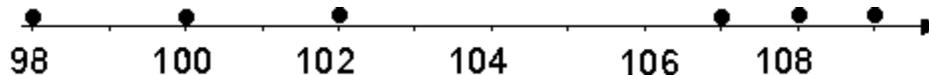
Una *muestra* es un subconjunto de la población.

Errores en el Proceso de Medición.

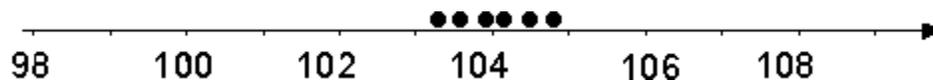
- Errores aleatorios -----> precisión
- Errores sistemáticos (sesgo) -----> exactitud

Un químico sintetiza un reactivo creyendo que es completamente nuevo. Lo estudia utilizando un método espectrométrico y obtiene un valor promedio de 104 (en unidades adecuadas). En los libros de referencia encuentra que todos los compuestos estudiados con el "mismo" método tienen un valor inferior a 100. ¿Descubrió nuestro químico un nuevo compuesto? La respuesta depende del grado de confianza que se pueda asignar al valor experimental 104. Si los valores muestrales fueron 98, 100, 102, 104, 107, 108 y 109

Diagrama de punto



el 100 aparece como un valor tan razonable como el 104. Se deberían realizar más mediciones. Si en cambio fueron 103.3 103.5 103.9 104.1 104.3 104.5 104.7



pensamos que el químico puede haber descubierto un nuevo compuesto o que ha cometido alguna equivocación que lo conduce a medir de más.

Los ERRORES ALEATORIOS NUNCA SE PUEDEN ELIMINAR.
Se pueden MINIMIZAR.

Además realizando medidas repetidas INDEPENDIENTES
se pueden EVALUAR por
procedimientos estadísticos.

Los ERRORES SISTEMÁTICOS no se pueden evaluar
realizando medidas repetidas.

Debe emplearse una forma de comparación alternativa.

Por ejemplo se pueden utilizar dos (o más) métodos físicos o químicos no relacionados para efectuar un análisis y si conducen en forma consistente a resultados que sólo muestran diferencias aleatorias es razonable suponer que no hay errores sistemáticos significativos.

MUESTREO--necesidad

- No es posible analizar la totalidad de un tanque lleno de leche (contenido de grasa)
- No es posible analizar toda el agua de un río (contaminante)
- Muchas técnicas de análisis son destructivas y por lo tanto no pueden aplicarse a la totalidad de un objeto.

EJEMPLO. Se desea obtener el peso promedio de las pastillas de un

gran lote. Elegimos 10 y pesamos cada una de ellas.

Población: todas las pastillas del lote

Variable: peso de una pastilla cualquiera del lote

Valores poblacionales de la **variable peso:** peso de cada una de las pastillas del lote.

Datos: el peso de cada una de las 10 las pastillas

¿Cómo elegimos las 10 pastillas del lote?

Mediante un procedimiento de **MUESTREO ALEATORIO**.

Decimos que nuestros datos son una **muestra aleatoria**.

MUESTREO ALEATORIO: procedimiento de selección por el cual **todos los miembros de la población tienen la misma posibilidad de ser elegidos**.

Más simple: **muestreo SISTEMÁTICO** o pseudo aleatorio. Se elige la unidad k-ésima a partir de un comienzo aleatorio entre los primeros k.

En el caso de las pastillas: se selecciona una de cada cien de una línea de producción.

Inconveniente: por ejemplo un cambio o variación periódica en el peso de las pastillas coincidente con el período de selección de la muestra (k).

MUESTREO A GRANEL: falta de unidades identificables para el muestreo.

EJEMPLO. Pureza del cloruro sódico contenido en un barril.

Se tomaron cinco “porciones” , A, B, C, D, E, de diferentes partes del barril elegidas al azar.

Se realizaron cuatro análisis repetidos sobre cada porción para evaluar la concentración de cloruro sódico.

Hay dos fuentes de variación posibles:

- Error aleatorio en la medición de la pureza
- Variaciones reales de la pureza del cloruro sódico en diferentes puntos.

Los resultados se analizan mediante el procedimiento ANOVA.

¿Qué veremos en este curso?

- **Estadística Descriptiva:** Métodos descriptivos visuales.
Medidas resumen.
- **Probabilidades:** lo mínimo necesario para realizar inferencias.
- **Inferencia Estadística:** Estimación. Intervalos de confianza. Tests de hipótesis. Análisis de la Varianza. Regresión lineal simple.