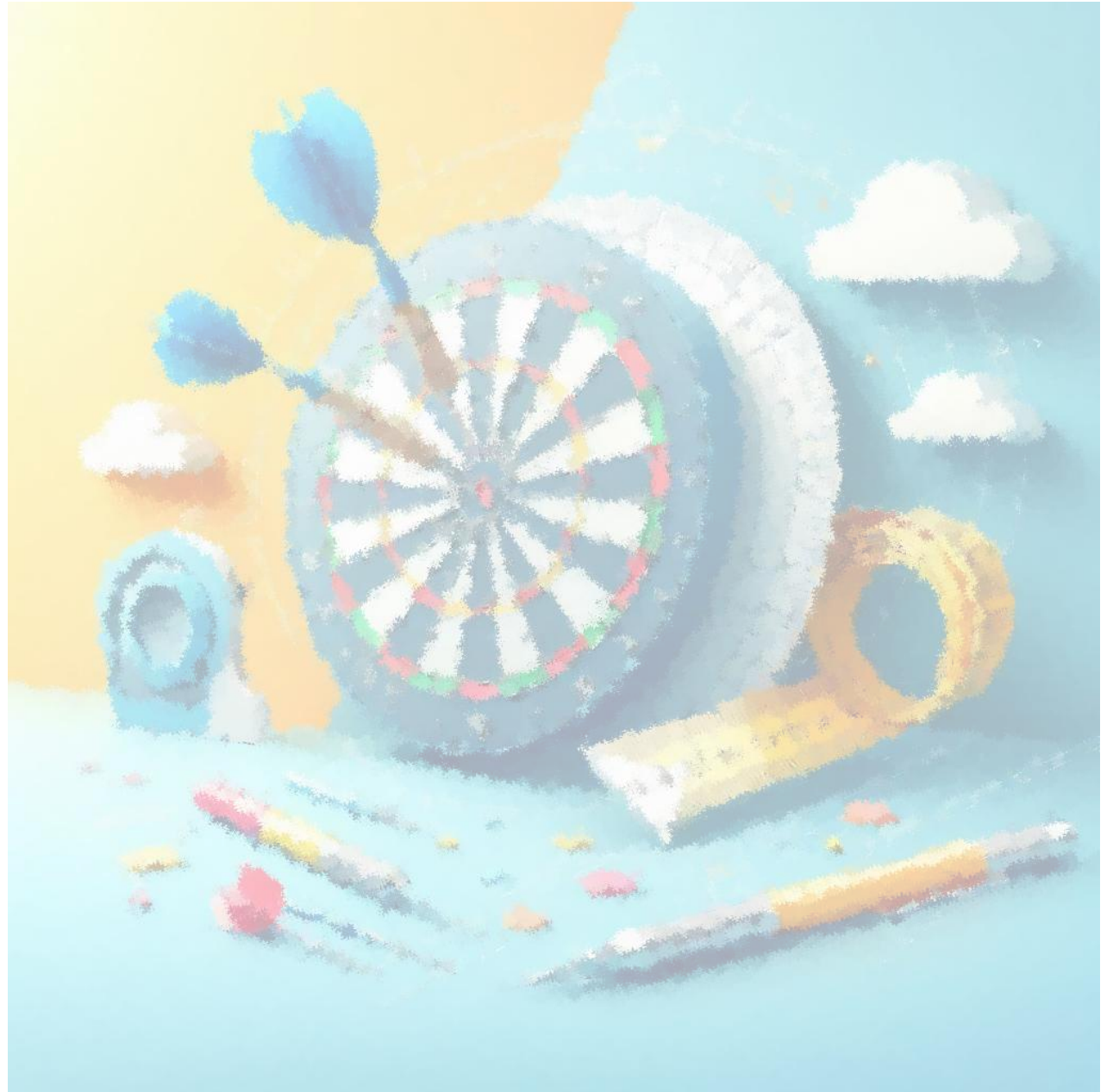


Precisión y exactitud

Ignacio Díaz Oreiro

CI0131 Diseño de Experimentos



Precisión y exactitud

- En el contexto del diseño experimental, precisión y exactitud son conceptos clave para evaluar la calidad de las mediciones y los resultados.
- Un diseño debe buscar ambas, tanto la precisión como la exactitud.

Precisión y exactitud

- Aunque a menudo se usan como sinónimos, tienen significados distintos:
- Exactitud: *"Proximidad al valor verdadero o teórico."*
- Precisión: *"Consistencia en mediciones repetidas."*

Exactitud

- Definición: Grado en que el resultado de un experimento se acerca al valor *verdadero* o teóricamente esperado.
- Enfoque: Está relacionada con la ausencia de sesgos sistemáticos (errores que afectan consistentemente las mediciones en una dirección).

Precisión

- Definición: Grado de repetibilidad o consistencia de las mediciones bajo las mismas condiciones.
- Enfoque: Refleja la variabilidad aleatoria (dispersión de los datos alrededor de su media).

Precisión y exactitud



Preciso
Exacto



Preciso
No exacto



No preciso
Exacto



No preciso
No exacto

Ejemplo

- En un diseño experimental la variable de respuesta mide el tiempo de ejecución de un algoritmo en milisegundos.
- Exactitud: hace referencia a qué tan cerca están las mediciones del valor real o verdadero del tiempo de ejecución.
- En este caso, el valor "verdadero" sería el tiempo exacto que tarda el algoritmo en ejecutarse, en condiciones ideales y sin errores en la medición.

Ejemplo - Exactitud

- En algoritmos, el valor "verdadero" rara vez se conoce (a menos que sea un algoritmo con tiempo teórico calculable, como en complejidad computacional).
- Si no hay referencia, la exactitud no puede medirse directamente, solo estimarse indirectamente (ej., comparando con otras mediciones).

Ejemplo - Exactitud

- La exactitud se puede evaluar mediante:
 - *Consistencia con modelos teóricos* (ej., si el algoritmo es $O(n)$, verificar que el tiempo crece linealmente con el tamaño de entrada).
 - Si no se tiene un valor teórico, se puede obtener una estimación aproximada con base en múltiples ejecuciones y ver si las mediciones promedio se acercan a un valor central.

Ejemplo

- Precisión: La precisión se refiere a la consistencia o repetibilidad de las mediciones.
- Es decir, si se ejecuta el algoritmo varias veces bajo las mismas condiciones, una medición precisa debería dar valores similares en cada repetición, independientemente de si el valor es exacto o no.

Ejemplo - Precisión

- La precisión depende de factores como:
 - Estabilidad del entorno (ej.: fluctuaciones en la carga del CPU).
 - Resolución del instrumento de medición (ej.: método de medición utilizado).

Ejemplo – Precisión

- La precisión se puede evaluar mediante:
 - *La realización de múltiples mediciones del tiempo de ejecución bajo las mismas condiciones (mismas entradas, mismo entorno).*
 - *Un método común es calcular la desviación estándar o varianza de las mediciones: cuanto menor sea esta, mayor será la precisión.*

Ejemplo

- Para definir la exactitud y precisión en este caso, ya que aún no se han tomado las mediciones, se puede:
- Exactitud estimada: Indicar cómo se planea reducir el error en la medición del tiempo (asegurando que no haya otros procesos que interfieran, utilizando un cronómetro confiable).
- Si no se tiene un valor de referencia, se puede especificar que el valor medido se espera que esté dentro de un rango razonable.

Ejemplo

- Precisión estimada: Definir cuántas veces se planea repetir la medición y cómo se va a calcular la variabilidad entre mediciones.
- Por ejemplo, se podría indicar que se va a ejecutar el algoritmo 30 veces y que se espera que la desviación estándar sea baja, indicando alta precisión.

Otro ejemplo

- Evaluar la experiencia de usuario de un software, donde la variable de respuesta es un resultado entre -3 y 3 que se obtiene de aplicar un cuestionario estandarizado con 6 preguntas.
- En este caso, los conceptos de exactitud y precisión se enfocan más en la calidad de las mediciones y el diseño del cuestionario.

Otro ejemplo - Exactitud

- Exactitud: En este contexto se refiere a qué tan bien las respuestas capturan el verdadero sentimiento o percepción de los usuarios respecto de su experiencia.
- En otras palabras, ¿el cuestionario refleja de manera precisa lo que los usuarios realmente piensan sobre el software?

Otro ejemplo - Exactitud

- Definición: La exactitud dependerá de qué tan bien está validado el cuestionario.
- Si es un cuestionario estandarizado, ya debe haber pasado por un proceso de validación donde se asegura que las preguntas miden lo que se proponen medir.

Otro ejemplo - Exactitud

- Medición: Se podrían comparar los resultados del cuestionario con otras mediciones o métodos considerados más precisos o con un "gold standard" en la evaluación de experiencia de usuario (si existiera).
- Otra forma sería hacer una validación de contenido o hacer pruebas piloto para ver si los usuarios interpretan las preguntas correctamente y si las respuestas corresponden con lo esperado.

Otro ejemplo - Precisión

- La precisión en este caso hace referencia a la consistencia de los resultados obtenidos del cuestionario.
- ¿Qué tan reproducibles son los resultados si aplicas el cuestionario en condiciones similares a los mismos usuarios o a usuarios con características similares?

Otro ejemplo - Precisión

- Definición: La precisión se refiere a la estabilidad y consistencia de las respuestas en el cuestionario.
- Si el mismo usuario u otros con experiencias similares responden el cuestionario en diferentes momentos, ¿obtendrán puntuaciones similares?

Otro ejemplo - Precisión

- Medición: Utilizar medidas de consistencia interna como el coeficiente alfa de Cronbach, que evalúa qué tan coherentes son las respuestas entre las diferentes preguntas del cuestionario.
- También se podría medir la fiabilidad *test-retest*, aplicando el mismo cuestionario a los mismos usuarios en diferentes momentos y comparando los resultados.

Otro ejemplo - Resumen

¿Cómo expresarlo en el diseño?

- Exactitud: “El cuestionario utilizado para la evaluación de UX ha sido validado en estudios previos, asegurando que mide de manera precisa la percepción de los usuarios sobre su experiencia con el software.”

Otro ejemplo - Resumen

- Precisión: “La consistencia interna del cuestionario será evaluada utilizando el coeficiente alfa de Cronbach, y se espera un valor de $\alpha \geq 0.7$ para garantizar que las preguntas sean consistentes entre sí. Adicionalmente, si es necesario, se realizará una prueba *test-retest* con un subgrupo de usuarios para verificar la estabilidad de las respuestas a lo largo del tiempo.”