

El problema oculto de la replicación: las mediciones repetidas no son experimentos independientes

Juan Segurola

3Dresyns by Resyner Technologies S.L.

© 2026 Juan Segurola. Todos los derechos reservados.

Correspondencia: info@3Dresyns.com · www.3Dresyns.com

HEB-008 · Artículo técnico independiente · Versión v0.3

DOI: 10.5281/zenodo.22915191

Resumen

La replicación a menudo se cuenta en lugar de definirse. Cinco lecturas instrumentales, cinco posiciones sobre una misma probeta, cinco probetas de una misma fabricación y cinco lotes preparados de forma independiente pueden producir todos el número $n = 5$ mientras estiman componentes de varianza diferentes. Tratar estos niveles como intercambiables crea pseudoreplicación, exagera la cantidad de información independiente y puede hacer que la incertidumbre parezca menor de lo que permite el experimento. Este artículo metodológico desarrolla el problema oculto de la replicación como una cuestión de identidad de la unidad experimental y estructura de varianza. Separa lecturas repetidas, reposicionamiento, replicación de probetas, replicación de preparación, replicación de fabricación, efectos de día, efectos de operador y efectos de laboratorio, y muestra por qué el nivel de replicación debe corresponder al nivel de la afirmación científica. La proposición central es que la independencia es una propiedad del proceso generador de datos, no del número de filas de una hoja de cálculo.

Palabras clave: replicación; pseudoreplicación; repetibilidad; reproducibilidad; varianza jerárquica; independencia; unidad experimental; diseño anidado

1. Propósito y límite de evidencia

HEB-001 introdujo la distinción entre lecturas repetidas y replicas independientes [1]. HEB-008 desarrolla esa distinción en un marco general. ER-209 conserva el diseño de ejercicios interlaboratorio y su arquitectura estadística específica [4]. La proposición central es:

Cuente unidades experimentales, no mediciones.

2. La unidad experimental es la unidad asignada independientemente al tratamiento

Si un tratamiento se aplica una sola vez a un lote y se miden diez alícuotas de ese lote, el tratamiento no se ha replicado de forma independiente diez veces. Las diez mediciones pueden estimar heterogeneidad intralote o repetibilidad de medición, pero no estiman la variabilidad del tratamiento entre lotes. La unidad experimental es la unidad asignada independientemente al tratamiento, mientras que la unidad observacional es la unidad sobre la que se realiza una medición. Pueden ser distintas: diez alícuotas medidas de un lote tratado independientemente son diez unidades observacionales, pero una sola unidad experimental para el contraste de tratamiento [5].

3. Las lecturas repetidas responden una pregunta estrecha

Las lecturas repetidas sin perturbar la probeta sondan principalmente la variación instrumental y de análisis a corto plazo. Retirar y reinsertar la probeta añade posicionamiento. Preparar una nueva probeta añade variación de preparación. Repetir la fabricación añade variación de manufactura. Repetir en otro día añade componentes temporales y ambientales. La replicación se hace más amplia a medida que se regenera independientemente una parte mayor de la cadena causal. La repetibilidad y la reproducibilidad metrológicas también distinguen las condiciones bajo las que se evalúa la precisión de medición [2].

4. La pseudoreplicación crea falsa precisión

Si observaciones correlacionadas se analizan como si fueran independientes, el tamaño muestral nominal supera al tamaño muestral efectivo. Los errores estándar pueden reducirse artificialmente y la significación estadística puede parecer más fuerte de lo que permite el diseño. Promediar replicas técnicas antes de un análisis de nivel superior puede ser apropiado en algunos casos, pero el modelo correcto depende de la jerarquía y de la afirmación [5].

5. La variación anidada es información, no una molestia

Una jerarquía útil es laboratorio, con día anidado dentro de laboratorio, fabricación dentro de día, probeta dentro de fabricación, posición dentro de probeta y lectura dentro de posición. Preguntas diferentes viven en niveles distintos. Un estudio de cribado de materiales puede necesitar lotes independientes; un estudio de fabricación puede necesitar fabricaciones independientes; un estudio metrológico puede necesitar colocaciones y lecturas repetidas [3].

6. Mas mediciones en el nivel equivocado no solucionan la falta de replicación

Cien lecturas repetidas sobre una sola probeta pueden estimar con gran precisión la lectura media de esa probeta y dejar completamente desconocida la incertidumbre entre preparaciones. Aumentar la replicación técnica no puede sustituir la replicación independiente biológica, de material, de lote, de fabricación o de laboratorio cuando son esos niveles los que sustentan la afirmación.

7. Aleatorización y replicación resuelven problemas diferentes

La aleatorización reduce la alineación sistemática entre variación de nuisance y niveles de factor; la replicación estima variación. Ninguna sustituye a la otra. Un estudio perfectamente aleatorizado sin replicación independiente no puede estimar el error puro relevante, mientras que un estudio replicado con un orden de ejecución fuertemente confundido puede atribuir la deriva al tratamiento.

8. La replicación debe corresponder a la afirmación de transferencia

ER-209 trata el diseño interlaboratorio y distingue la repetibilidad intraprobeta de la variación entre fabricaciones y entre laboratorios; esa cualificación específica de aplicación permanece fuera de HEB-008 [4]. HEB-015 conserva por separado la cuestión más amplia del nivel de transferencia, mientras HEB-008 trata la estructura de independencia dentro del experimento.

9. Las medias pueden ocultar la jerarquía

Colapsar todas las observaciones en una sola media y desviación estándar puede mezclar variabilidad intraunidad y entre unidades. Dos estudios pueden informar el mismo n y la misma desviación estándar y representar estructuras de evidencia muy diferentes. Por ello, el informe debe declarar que se replicó y como estaban anidadas o cruzadas las observaciones.

10. Un mapa práctico de replicación

Antes de recoger datos, dibuje la cadena causal y marque cada nivel que puede variar: lote de materia prima, preparación, lote, fabricación, probeta, posición, lectura, día, operador, instrumento y laboratorio. Para cada afirmación científica, identifique el nivel más alto al que se pretende generalizar. Replique en ese nivel o estreche la afirmación.

11. Una jerarquía ilustrativa de replicación

Imaginemos un estudio de recubrimientos con tres lotes preparados de forma independiente. De cada lote se colan cuatro probetas; en cada probeta se miden cinco posiciones; en cada posición el instrumento registra tres lecturas repetidas. El conjunto de datos contiene 180 números, pero no contiene 180 lotes, probetas o posiciones independientes. Las tres lecturas informan principalmente de la repetibilidad de lectura a corto plazo; las cinco posiciones informan de la variación espacial dentro de la probeta; las cuatro probetas informan de la variación entre probetas condicionada al lote; y los tres lotes informan de la variación entre preparaciones. Si la afirmación científica se refiere a la preparación de formulación, el recuento independiente más fuerte se aproxima más al número de lotes que al número de lecturas brutas. Las mediciones de nivel inferior siguen siendo valiosas porque ayudan a descomponer la estructura del error. Solo pierden su valor cuando se etiquetan erróneamente como evidencia independiente para una afirmación de nivel superior.

12. Las estructuras anidadas y cruzadas responden preguntas de transferencia diferentes

Las jerarquías suelen estar anidadas: las posiciones pertenecen a probetas, las probetas a fabricaciones y las fabricaciones a lotes. Otros factores pueden estar cruzados: los mismos operadores pueden medir probetas de todos los lotes o los mismos instrumentos pueden ensayar material de varias preparaciones. Cruzar factores puede separar efectos de operador o instrumento de efectos de material de forma más eficiente que asignar un operador a un lote, lo que los confunde. El diseño debe reflejar, por tanto, los componentes de varianza que importan. Un estudio destinado a comparar operadores debe proporcionar a los operadores estados de material solapados; un estudio

destinado solo a estimar repetibilidad intraoperador puede mantener intencionadamente fijo al operador. La replicación no es una petición genérica de «más n». Es una arquitectura para hacer estimables fuentes concretas de variación.

13. Potencia estadística y generalización no son el mismo objetivo

Añadir muchas réplicas técnicas puede aumentar la precisión de una media estrecha y aportar poco a la generalización entre lotes, días, máquinas o laboratorios. A la inversa, distribuir recursos limitados entre más unidades independientes de alto nivel puede aumentar la información de transferencia a costa de menos lecturas por unidad. El diseño experimental implica, por tanto, un compromiso entre caracterizar con precisión cada unidad y muestrear la población a la que se aplicará la conclusión. La asignación óptima depende de la estructura de varianza y de la afirmación. Lo que no debe ocurrir es utilizar abundante replicación de bajo nivel para insinuar evidencia sobre una variabilidad de alto nivel que nunca fue muestreada.

14. Conclusiones

La replicación solo tiene significado respecto a una unidad experimental definida. Las mediciones repetidas pueden mejorar la precisión en un nivel mientras dejan intacta la incertidumbre de niveles superiores. La pseudorreplicación aparece cuando observaciones correlacionadas se cuentan como evidencia independiente. Un experimento robusto declara, por tanto, la jerarquía, conserva la identidad de las unidades y estima los componentes de varianza que corresponden a la afirmación prevista.

Referencias

- [1] Seguro, J. The Hidden Error Budget: Measurement-System Failure, False Precision and Experimental Design in Laboratory Science. HEB-001. 3Dresyns Independent Technical Articles (2026). DOI: 10.5281/zenodo.22862642.
- [2] Joint Committee for Guides in Metrology. International Vocabulary of Metrology - Basic and General Concepts and Associated Terms (VIM), 3rd ed., JCGM 200:2012. DOI: 10.59161/JCGM200-2012.
- [3] Heckert, N. A.; Filliben, J. J.; Croarkin, C. M.; Hembree, B.; Guthrie, W. F.; Tobias, P.; Prinz, J. Handbook 151: NIST/SEMATECH e-Handbook of Statistical Methods. National Institute of Standards and Technology, 2002.
- [4] Seguro, J. Interlaboratory Round-Robin Design for Vat-Photopolymer Materials. ER-209. 3Dresyns Engineering Publications (2026). Versión DOI: 10.5281/zenodo.22862020.
- [5] Hurlbert, S. H. Pseudoreplication and the Design of Ecological Field Experiments. Ecological Monographs 54(2) (1984) 187-211. DOI: 10.2307/1942661.

Límite de conocimiento

HEB-008 trata la distinción general entre observaciones repetidas y replicación experimental independiente. ER-209 conserva el diseño interlaboratorio, la ejecución de round-robin y el análisis de reproducibilidad específico de normas.

Limitaciones

Este artículo metodológico es una síntesis acotada de los corpus controlados HEB y Engineering Review. No informa de un nuevo conjunto de datos experimental, no establece umbrales numéricos universales y no convierte ejemplos de un dominio en especificaciones aplicables a todas las aplicaciones. Las afirmaciones que requieren cualificación específica de aplicación permanecen en la revisión técnica o estudio primario correspondiente.