



Estadística Sesgo y Pruebas Paramétricas

Dr. Joel De La Cruz Oller
2-2025-2026

Importante



Al final de esta presentación hay un enlace a una prueba corta de *Microsoft Forms*. Debe obtener al menos un 80% en la prueba para aprobarla.

Le exhortamos a que estudie el módulo con detenimiento antes de tomar la prueba.

Objetivos



Después de estudiar este módulo, los estudiantes podrán:

- Comprender y diferenciar los conceptos de sesgo (bias) y factor de confusión, su impacto en la validez del estudio y ejemplos prácticos.
- Aprender métodos para detectar y corregir sesgos y confusores mediante diseño (aleatorización, emparejamiento, restricción) y análisis (estratificación, ajuste multivariado)

Objetivos



Después de estudiar este módulo, los estudiantes podrán:

- Entender el valor p y el nivel de significancia (α) como criterio para determinar significancia estadística .
- Diferenciar error tipo I (falso positivo) y tipo II (falso negativo), comprender β , potencia ($1-\beta$) y su relación con el tamaño muestral.

Objetivos



Después de estudiar este módulo, los estudiantes podrán:

- Aplicar el procedimiento para realizar pruebas de hipótesis: identificar y clasificar variables dependientes e independientes y seleccionar pruebas paramétricas adecuadas.

LA DISTRIBUCIÓN NORMAL

- Diferencia entre sesgo (Bias) y factor de confusión (Cofounder)
- El sesgo (bias) es un error sistemático que conduce a una medición falsa de la variable dependiente y, por lo tanto, a una conclusión errónea.
- Un factor de confusión es una variable que se relaciona con la variable dependiente e independiente, no se encuentra en la vía causal (que se llamaría una variable intermedia) y conduce al supuesto de una relación causal entre la variable dependiente e independiente, que, sin embargo, se sobreestima o induce a error.

Sesgo



- El sesgo en estadística es un error sistemático que provoca que las estimaciones o mediciones se alejen de los valores verdaderos de la población.
- A diferencia del error aleatorio, que fluctúa alrededor del valor real y se reduce con muestras más grandes, el sesgo persiste independientemente del tamaño muestral y puede conducir a conclusiones incorrectas.
- Las fuentes comunes de sesgo incluyen un muestreo no representativo, errores en la recolección de datos, instrumentos mal calibrados y el sesgo del experimentador o del encuestado.
- Detectarlo requiere comparar métodos, revisar el diseño del estudio y examinar patrones consistentes en los residuos o en las diferencias entre subgrupos.
- Corregirlo implica rediseñar el muestreo, ajustar por variables relevantes (p. ej., mediante estratificación o modelado) o mejorar la calidad de la medición.

Sesgo



- Desde la perspectiva práctica, el sesgo compromete la validez interna y externa de un estudio: puede llevar a inferencias equivocadas sobre relaciones causales o a estimaciones sistemáticamente sobre- o subestimadas de efectos.
- En epidemiología y ciencias sociales, identificar y controlar confusores es esencial para reducir sesgos de asociación; sin embargo, algunos sesgos, como el de información o el de selección, requieren intervenciones específicas en la recolección y manejo de datos.
- La evaluación crítica del diseño del estudio, el uso de técnicas de ajuste estadísticas y la transparencia en la metodología ayudan a mitigar su impacto.
- Los análisis de sensibilidad y la replicación en diferentes muestras o condiciones también permiten evaluar cuánto puede haber influido el sesgo en los resultados.
- Finalmente, reconocer la posibilidad de sesgo y reportarla adecuadamente fortalece la interpretación responsable de los hallazgos.

Sesgo



- Ejemplo de sesgo:
- Si una encuesta telefónica sobre ingresos se realiza solo por teléfonos fijos, se excluirá sistemáticamente a adultos jóvenes que usan solo móviles, produciendo una sobreestimación del ingreso medio (sesgo de selección).

Factor de confusión



- Un factor de confusión es una variable que está asociada tanto con la exposición (o variable independiente) como con el resultado (variable dependiente) y que no se encuentra en la cadena causal entre ambas.
- Su presencia puede generar una asociación espuria o exagerada entre exposición y resultado porque parte o todo el efecto atribuido a la exposición proviene en realidad del confusor.
- Para ser variable de confusión debe relacionarse con la exposición, con el resultado, y no ser un efecto intermedio de la exposición sobre el resultado. Identificar confusores exige conocimiento del contexto y examinar asociaciones entre variables antes del ajuste.
- Controlar confusores puede lograrse mediante diseño (aleatorización, emparejamiento, restricción) o análisis (estratificación, ajuste multivariado, modelado).

Factor de confusión

- Ejemplo práctico: suponga que queremos estudiar si el consumo de café se asocia con mayor riesgo de infarto; el tabaquismo podría ser un factor de confusión porque las personas que fuman tienden a beber más café y fumar aumenta el riesgo de infarto.
- Si no se controla el tabaquismo, podríamos atribuir a la cafeína un efecto que en realidad corresponde al tabaco.
- Para abordar esto se puede ajustar por tabaquismo en un modelo multivariado, realizar análisis estratificados por estado de fumador o usar emparejamiento por tabaquismo en el diseño.
- Informar y probar la robustez del resultado con análisis de sensibilidad ayuda a evaluar cuánto influye el confusor en la estimación.

Pruebas Paramétricas

- **Valor p**
- El valor p proporciona una estimación de la probabilidad del evento y, si es lo suficientemente pequeño, el resultado se denomina estadísticamente significativo.
- Definición del valor p: La probabilidad del resultado observado o algo más extremo bajo la hipótesis nula.

Pruebas Paramétricas

- **Valor p**
- Este umbral se llama alfa (α), el nivel de significancia.
- Es ampliamente aceptado considerar un valor p igual o menor que un α de 0.05 lo suficientemente pequeño (que corresponde a una probabilidad de 1 sobre 20).

Pruebas Paramétricas

No se puede rechazar H_0

Rechazar H_0

H_0 es verdadero correcto

Error de tipo I

**H_0 es falso
(potencia)**

Error de tipo II

Correcto

Pruebas Paramétricas

- **Error tipo I (falso positivo):** Rechazar la hipótesis nula aunque la nula sea verdadera (en otras palabras, alegar una diferencia significativa cuando en realidad no hay diferencia).
- Nivel de significancia Alfa = α — probabilidad de cometer un error de tipo I

Pruebas Paramétricas

- **Error tipo I (falso positivo):**
- Un α de 0.05 significa que está aceptando un 5% de probabilidad máxima de rechazar incorrectamente la hipótesis nula (H_0).
- Cuanto menor sea el Alfa, menor será esta posibilidad "permitida".

Pruebas Paramétricas

- **Error tipo II (falso NEGATIVO):** No rechazar la hipótesis nula cuando la nula es falsa (en otras palabras, alegando que no hay diferencia significativa cuando en realidad sí hay diferencia);
- Esto sucede cuando el experimento tiene poca potencia).
- $BETA = \beta$ — probabilidad de cometer un error de tipo II = Relacionado directamente con la potencia ($Potencia = 1 - \beta$).

Pruebas Paramétricas

- **Error tipo II (falso NEGATIVO):**
- La mayoría de los estudios establecen un β de 0,2.
- Esto significa que su poder será 0.8 (80%), y que está aceptando un 20% de probabilidad de no rechazar la hipótesis nula (H_0) cuando en realidad es falsa.

Pruebas Paramétricas

- **CÓMO REALIZAR Pruebas de hipótesis**
- Primer paso: comprender las variables; Primer paso importante: clasificar las variables de estudio.
- El investigador debe determinar las variables dependientes (o el resultado) e independientes.

Referencias



- Banack, H. R., Hayes-Larson, E., & Mayeda, E. R. (2022). Monte Carlo simulation approaches for quantitative bias analysis: A tutorial. *Epidemiologic Reviews*, 43(1), 106–117. <https://doi.org/10.1093/epirev/mxab012>
- Christley, R. M. (2008). Statistical significance, power and sample size—What does it all mean? *Journal of Small Animal Practice*, 49(6), 263. <https://doi.org/10.1111/j.1748-5827.2008.00612.x>
- Delgado-Rodríguez, M., & Llorca, J. (2004). Bias. *Journal of Epidemiology and Community Health*, 58(8), 635–641. <https://doi.org/10.1136/jech.2003.008466>
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486–489. <https://doi.org/10.5812/ijem.3505>

Referencias



- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, *P* values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology*, 31(4), 337–350.
<https://doi.org/10.1007/s10654-016-0149-3>
- Infante-Rivard, C. (2018). Reflection on modern methods: Selection bias—A review of recent developments. *International Journal of Epidemiology*, 47(5), 1714–1722.
<https://doi.org/10.1093/ije/dyy138>
- Mascha, E. J., & Vetter, T. R. (2018). Significance, errors, power, and sample size: The blocking and tackling of statistics. *Anesthesia & Analgesia*, 126(2), 691–698.
<https://doi.org/10.1213/ANE.0000000000002741>

Referencias



Tripepi, G., Jager, K. J., Dekker, F. W., & Zoccali, C. (2010). Selection bias and information bias in clinical research. *Nephron Clinical Practice*, 115(2), c94–c99.

<https://doi.org/10.1159/000312871>

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p -values: Context, process, and purpose. *The American Statistician*, 70(2), 129–133.

<https://doi.org/10.1080/00031305.2016.1154108>

Para finalizar, haga clic sobre el enlace para tomar la prueba y enviar el informe a su profesor(a):

Estadística: Sesgo y Pruebas Paramétricas – Fill out form

¡Gracias por utilizar los recursos del  !

Conozca más sobre nuestros servicios:

<http://generales.uprrp.edu/competencias-linguisticas/>

Para revisiones: <https://forms.office.com/r/f2pYYSQnHp>