Reforzamiento de Estadística en Machine Learning

1. El Teorema del Límite Central (TLC)

El Teorema del Límite Central (TLC) es un pilar fundamental en la teoría de la probabilidad y la estadística. Establece que, bajo ciertas condiciones, la distribución de la media (o la suma) de un número suficientemente grande de variables aleatorias independientes e idénticamente distribuidas (i.i.d.) se aproximará a una distribución normal, independientemente de la forma de la distribución original de las variables (Kwak & Kim, 2017; DataCamp, s.f.). Esta convergencia hacia la normalidad es lo que otorga al TLC su notable poder y aplicabilidad.

Condiciones para la Aplicación del TLC:

Aunque el TLC es robusto, su validez depende de ciertas condiciones:

1. **Independencia e Idéntica Distribución (i.i.d.):** Las muestras deben ser seleccionadas de manera que sean independientes entre sí y provengan de la misma distribución de población (DataCamp, s.f.).
2. **Tamaño de Muestra Suficientemente Grande:** Comúnmente, se considera que un tamaño de muestra (n \geq 30) es adecuado para que la aproximación normal sea razonable, aunque esto puede variar dependiendo de la simetría de la distribución original; poblaciones muy sesgadas pueden requerir muestras más grandes (Kwak & Kim, 2017; Penn State Eberly College of Science, s.f.-a).
3. **Varianza Finita:** La población de la cual se extraen las muestras debe tener una media (mu) y una varianza (sigma2) finitas y bien definidas (Statistics LibreTexts, 2023a).

Importancia e Implicaciones:

La principal importancia del TLC radica en que permite realizar inferencias estadísticas sobre parámetros poblacionales utilizando la distribución normal, incluso cuando se desconoce la distribución subyacente de la población (DataCamp, s.f.; Cornell University, s.f.). Esto es crucial para:

- **Pruebas de Hipótesis y Intervalos de Confianza:** Facilita la construcción de intervalos de confianza para la media poblacional y la realización de pruebas de hipótesis (e.g., t-tests, ANOVA) que asumen normalidad de las medias muestrales (Kwak & Kim, 2017).
- **Aplicaciones en Diversos Campos:** Se utiliza en control de calidad, finanzas, encuestas y, crucialmente, en Machine Learning para la validación de modelos, la estimación de errores y la comprensión del comportamiento de los estimadores (Cornell University, s.f.; DataCamp, s.f.). Por ejemplo, en el bootstrapping o en la evaluación de métricas de error mediante validación cruzada, el TLC puede ayudar a entender la distribución de dichas métricas.

Formulación:

Si X_1,X_2,ldots,X_n son n variables aleatorias i.i.d. con media mu y varianza sigma2, entonces la distribución de la media muestral barX=frac1nsum_i=1nX_i se aproxima a una distribución normal con media mu y varianza fracsigma2n (es decir, barXsimN(mu,fracsigma2n)) a medida que n aumenta (Statistics LibreTexts, 2023a). De manera similar, la distribución de la suma de estas variables, S_n=sum_i=1nX_i, se aproxima a una distribución normal con media nmu y varianza nsigma2 (Kwak & Kim, 2017).

En resumen, el TLC es esencial porque justifica el uso extendido de la distribución normal en la inferencia estadística, proporcionando una base teórica para muchas técnicas analíticas.

2. Sampling (Muestreo)

El muestreo es el proceso de seleccionar un subconjunto de individuos o elementos[1] de una población más grande con el objetivo de hacer inferencias sobre dicha población (Bazán, 2004; Lumen Learning, s.f.). Un muestreo adecuado es crucial para obtener resultados de investigación que sean representativos y generalizables.

Importancia del Muestreo:

- **Eficiencia:** Estudiar una muestra es generalmente más rápido, económico y práctico que estudiar una población entera (Bazán, 2004).
- **Representatividad:** Una muestra bien seleccionada puede reflejar con precisión las características de la población, permitiendo inferencias válidas (Newcastle University, s.f.).
- **Viabilidad:** En muchos casos, es imposible acceder a toda la población, haciendo del muestreo la única opción viable.

Tipos de Muestreo:

Existen dos categorías principales de muestreo: probabilístico y no probabilístico.

A. Muestreo Probabilístico:

En el muestreo probabilístico, cada miembro de la población tiene una probabilidad2 conocida y no nula de ser seleccionado. Esto permite la estimación del error de muestreo y la generalización estadística (Bazán, 2004).

1. **Muestreo Aleatorio Simple (MAS):**
    - **Definición:** Cada individuo y cada posible muestra de un tamaño dado tienen la misma probabilidad de ser seleccionados (Lumen Learning, s.f.). Se puede realizar mediante tablas de números aleatorios o generadores de números

aleatorios.
    - **Ventajas:** Fácil de entender e implementar si se dispone de un listado completo de la población; es fundamentalmente insesgado (Newcastle University, s.f.).
    - **Desventajas:** Puede no ser el más eficiente si la población es heterogénea y no garantiza la representación de subgrupos pequeños (Newcastle University, s.f.). Requiere un marco muestral completo.
2. **Muestreo Sistemático:**
    - **Definición:** Se selecciona un punto de partida aleatorio y luego se elige cada k-ésimo elemento de la lista de la población (Bazán, 2004; Lumen Learning, s.f.). El intervalo k se calcula dividiendo el tamaño de la población por el tamaño de la muestra deseado.
    - **Ventajas:** Más fácil de implementar que el MAS, especialmente con poblaciones grandes y ordenadas; asegura una dispersión uniforme de la muestra a lo largo de la población (Newcastle University, s.f.).
    - **Desventajas:** Puede introducir sesgo si hay un patrón o periodicidad en la lista que coincida con el intervalo de muestreo (Tille, s.f.).
3. **Muestreo Estratificado:**
    - **Definición:** La población se divide en subgrupos homogéneos mutuamente excluyentes llamados estratos (e.g., por edad, género, ubicación geográfica). Luego, se realiza un muestreo aleatorio simple o sistemático dentro de cada estrato (Bazán, 2004; Lumen Learning, s.f.). La selección de los estratos puede ser proporcional o no proporcional al tamaño del estrato en la población.
    - **Ventajas:** Asegura la representación de todos los subgrupos importantes, permite estimaciones más precisas para cada subgrupo y puede mejorar la precisión general en comparación con el MAS si los estratos son homogéneos internamente y heterogéneos entre sí (Newcastle University, s.f.).
    - **Desventajas:** Requiere conocimiento previo de las características de la población para formar los estratos; puede ser más complejo de implementar y analizar (Bazán, 2004).
4. **Muestreo por Conglomerados (Clusters):**
    - **Definición:** La población se divide en grupos o conglomerados (e.g., escuelas, ciudades, bloques de viviendas). Se selecciona aleatoriamente una muestra de conglomerados, y luego se incluyen en la muestra todos los individuos de los conglomerados seleccionados (muestreo de una etapa) o se realiza un muestreo adicional dentro de los conglomerados seleccionados (muestreo de múltiples etapas) (Lumen Learning, s.f.; Newcastle University, s.f.).
    - **Ventajas:** Útil cuando la población está geográficamente dispersa o cuando no existe un listado completo de individuos pero sí de conglomerados; puede ser más económico y rápido (Newcastle University, s.f.).
    - **Desventajas:** Menos preciso que el MAS o el estratificado si los

conglomerados son heterogéneos internamente y similares entre sí; el análisis es más complejo debido a la correlación intra-conglomerado (Bazán, 2004).

B. Muestreo No Probabilístico:

En el muestreo no probabilístico, la selección de los individuos no se basa en el azar, sino en el juicio del investigador, la conveniencia o cuotas preestablecidas. No permite la estimación del error de muestreo y la generalización a la población es limitada (Bazán, 2004).

1. **Muestreo por Conveniencia:** Se seleccionan los individuos que son más accesibles para el investigador (Lumen Learning, s.f.). Es rápido y económico, pero propenso a sesgos significativos.
2. **Muestreo Intencional o por Juicio:** El investigador selecciona a los individuos que considera más apropiados o representativos para el estudio (Bazán, 2004).
3. **Muestreo por Cuotas:** Se establecen cuotas para diferentes subgrupos de la población (e.g., 50% hombres, 50% mujeres) y se seleccionan individuos hasta llenar dichas cuotas, a menudo por conveniencia o juicio dentro de cada cuota (Tille, s.f.).
4. **Muestreo de Bola de Nieve:** Se comienza con unos pocos individuos que cumplen los criterios de inclusión y luego se les pide que refieran a otros individuos con características similares (Bazán, 2004). Útil para poblaciones difíciles de alcanzar.

La elección del método de muestreo depende de los objetivos de la investigación, los recursos disponibles, la naturaleza de la población y el grado de precisión requerido. En el contexto de Machine Learning, un muestreo adecuado es vital para la creación de conjuntos de entrenamiento y prueba que reflejen la distribución de los datos del mundo real, evitando así modelos sesgados o con pobre generalización (Tille, s.f.).

3. Diferencia entre Error Tipo I y Error Tipo II

En el contexto de la prueba de hipótesis estadísticas, se busca tomar una decisión sobre una afirmación acerca de una población (la hipótesis nula, H_0) basándose en la evidencia de una muestra. Debido a la incertidumbre inherente al muestreo, existe la posibilidad de cometer errores en esta decisión. Los dos tipos de errores que se pueden cometer son el Error Tipo I y el Error Tipo II (Nahm, 2010; Scribbr, 2023).

Definiciones:

- Error Tipo I (Falso Positivo):
  Ocurre cuando se rechaza la hipótesis nula (H_0) cuando esta es en realidad verdadera (Statistics LibreTexts, 2021). En otras palabras, se concluye que existe un efecto o diferencia cuando en realidad no lo hay.
  La probabilidad de cometer un Error Tipo I se denota con la letra griega alfa (alpha) y se conoce como el nivel de significancia de la prueba. El investigador establece este valor antes de realizar la prueba (comúnmente alpha=0.05, lo que implica un 5% de riesgo de cometer este error) (Nahm, 2010).

    - **Ejemplo:** Un estudio concluye que un nuevo medicamento es efectivo para tratar una enfermedad (rechaza H_0: el medicamento no es efectivo), cuando en realidad el medicamento no tiene ningún efecto (Scribbr, 2023). En diagnósticos médicos, un Error Tipo I es un resultado de prueba que indica incorrectamente la presencia de una enfermedad (Fortress Diagnostics, 2025).
- Error Tipo II (Falso Negativo):
  Ocurre cuando no se rechaza la hipótesis nula (H_0) cuando esta es en realidad falsa (Statistics LibreTexts, 2021). Es decir, se concluye que no hay un efecto o diferencia cuando en realidad sí existe.
  La probabilidad de cometer un Error Tipo II se denota con la letra griega beta (beta). El valor de beta no suele fijarse directamente por el investigador, pero está inversamente relacionado con la potencia de la prueba (Nahm, 2010).

    - **Ejemplo:** Un estudio concluye que un nuevo medicamento no es efectivo (no rechaza H_0: el medicamento no es efectivo), cuando en realidad el medicamento sí es efectivo (Scribbr, 2023). En diagnósticos médicos, un Error Tipo II es un resultado de prueba que indica incorrectamente la ausencia de una enfermedad cuando el paciente sí la padece (Fortress Diagnostics, 2025).

Tabla Resumen de Decisiones:

	H_0 es Verdadera	**H_0 es Falsa**
No se rechaza H_0	Decisión Correcta (Prob = 1–alpha)	Error Tipo II (Prob = beta)
Se rechaza H_0	Error Tipo I (Prob = alpha)	Decisión Correcta (Prob = 1–beta, Potencia)

(Basado en Statistics LibreTexts, 2021)

Relación entre alpha, beta y la Potencia de la Prueba:

- **Potencia de la Prueba (1–beta):** Es la probabilidad de rechazar correctamente una hipótesis nula falsa. Es la capacidad de la prueba para detectar un efecto real (Nahm, 2010; Scribbr, 2023). Se desea que la potencia sea alta.
- **Trade-off entre alpha y beta:** Existe una relación inversa entre alpha y beta. Si se disminuye la probabilidad de cometer un Error Tipo I (haciendo alpha más pequeño),

generalmente aumenta la probabilidad de cometer un Error Tipo II (beta), y viceversa, manteniendo constante el tamaño de la muestra y el efecto (Scribbr, 2023). La elección del nivel de alpha debe considerar las consecuencias relativas de cometer cada tipo de error.

Factores que Influyen en los Errores:

- **Nivel de Significancia (alpha):** Fijado por el investigador. Un alpha más bajo reduce el Error Tipo I pero aumenta el Error Tipo II.
- **Tamaño de la Muestra (n):** Aumentar el tamaño de la muestra generalmente aumenta la potencia de la prueba, lo que reduce beta para un alpha dado (Nahm, 2010).
- **Magnitud del Efecto:** Efectos más grandes son más fáciles de detectar, lo que aumenta la potencia y reduce beta.
- **Variabilidad de los Datos:** Una menor variabilidad en los datos conduce a una mayor potencia.

Consecuencias:

Las consecuencias de cometer cada tipo de error varían según el contexto:

- **Error Tipo I:** Puede llevar a la adopción de tratamientos ineficaces, la implementación de políticas basadas en hallazgos falsos o la inversión de recursos en direcciones incorrectas (Scribbr, 2023). En el diagnóstico médico, un falso positivo puede causar ansiedad innecesaria, tratamientos invasivos y costos adicionales (Fortress Diagnostics, 2025).
- **Error Tipo II:** Puede llevar a no reconocer tratamientos efectivos, perder oportunidades de innovación o no identificar problemas reales (Scribbr, 2023). En el diagnóstico médico, un falso negativo puede retrasar un tratamiento necesario, empeorando el pronóstico del paciente y permitiendo la propagación de enfermedades infecciosas (Fortress Diagnostics, 2025).

En Machine Learning, estos conceptos son análogos a los falsos positivos y falsos negativos en la evaluación de modelos de clasificación, donde las decisiones (e.g., si un correo es spam o no, si un paciente tiene una enfermedad o no) también conllevan riesgos y consecuencias asociadas a cada tipo de error.

4. Regresión Lineal y sus Métricas (p-value, coeficientes, R-cuadrado)

La regresión lineal es una técnica estadística fundamental utilizada para modelar la relación entre una variable dependiente (o respuesta, Y) y una o más variables independientes (o predictoras, X_1,X_2,ldots,X_k) asumiendo una relación lineal (IBM, 2021). Si hay una sola variable independiente, se denomina regresión lineal simple; si hay múltiples variables independientes, es una regresión lineal múltiple.

Propósito y Aplicaciones:

- **Predicción:** Estimar el valor de la variable dependiente para nuevos valores de las variables independientes (IBM, 2021).
- **Comprensión de Relaciones:** Cuantificar la fuerza y dirección de la relación entre las variables. Por ejemplo, determinar cómo un cambio en una variable predictora afecta a la variable respuesta (Statistics By Jim, s.f.-b).
- **Control de Variables:** Aislar el efecto de una variable predictora sobre la respuesta, controlando el efecto de otras variables en un modelo múltiple.

Es ampliamente utilizada en campos como la economía (predecir crecimiento), finanzas (evaluar riesgo), ciencias sociales (modelar comportamientos) y Machine Learning (como algoritmo base y para la ingeniería de características).

Ecuación del Modelo de Regresión Lineal Múltiple:

Y=beta_0+beta_1X_1+beta_2X_2+ldots+beta_kX_k+epsilon

Donde:

- Y es la variable dependiente.
- X_1,X_2,ldots,X_k son las variables independientes.
- beta_0 es el intercepto (el valor esperado de Y cuando todas las X_i son cero).
- beta_1,beta_2,ldots,beta_k son los coeficientes de regresión (pendientes), que representan el cambio esperado en Y por un cambio de una unidad en la X_i correspondiente, manteniendo constantes las demás variables (UCLA: Statistical Consulting Group, s.f.).
- epsilon es el término de error aleatorio, que representa la variabilidad no explicada por el modelo.

Los coeficientes (beta_i) se estiman a partir de los datos, comúnmente mediante el método de mínimos cuadrados ordinarios (MCO), que minimiza la suma de los cuadrados de las diferencias entre los valores observados y los predichos por el modelo (IBM, 2021).

Asunciones del Modelo de Regresión Lineal (Clásico):

Para que las estimaciones y las inferencias sean válidas, el modelo de regresión lineal se basa en varias asunciones sobre el término de error epsilon (IBM, 2021; Penn State Eberly College of Science, s.f.-b):

1. **Linealidad:** La relación entre las variables independientes y la variable dependiente es lineal en los parámetros.
2. **Independencia de los Errores:** Los errores epsilon_i son independientes entre sí (no hay autocorrelación).
3. **Homocedasticidad:** Los errores tienen una varianza constante (sigma2) para todos

los niveles de las variables independientes (varianza constante de los errores).

4. **Normalidad de los Errores:** Los errores epsilon_i se distribuyen normalmente con media cero y varianza sigma2. Esta asunción es particularmente importante para la inferencia (pruebas de hipótesis e intervalos de confianza).
5. **No Multicolinealidad Perfecta:** Las variables independientes no están perfectamente correlacionadas entre sí (en regresión múltiple).

Métricas Clave para la Evaluación del Modelo:

1. **Coeficientes de Regresión (beta_i):**
   - **Interpretación:** Como se mencionó, cada coeficiente beta_i (para igeq1) indica el cambio promedio en la variable dependiente Y asociado con un aumento de una unidad en la variable independiente X_i, manteniendo todas las demás variables independientes constantes (UCLA: Statistical Consulting Group, s.f.). El intercepto beta_0 es el valor esperado de Y cuando todas las X_i son cero.
   - **Significancia:** La importancia de cada coeficiente se evalúa mediante pruebas de hipótesis.
2. **p-value (para los coeficientes):**
   - **Definición:** En el contexto de los coeficientes de regresión, el p-value se asocia con la prueba de hipótesis de que un coeficiente particular es igual a cero (H_0:beta_i=0), lo que implicaría que la variable independiente X_i no tiene una relación lineal significativa con Y después de considerar las otras variables en el modelo (Universidad de Granada, s.f.).
   - **Interpretación:** Un p-value pequeño (típicamente $\< 0.05$) sugiere que se puede rechazar la hipótesis nula, lo que indica que el coeficiente es estadísticamente significativo y que la variable predictora correspondiente tiene un efecto lineal significativo sobre la variable dependiente. Un p-value grande sugiere que no hay suficiente evidencia para afirmar que el coeficiente es diferente de cero (UCLA: Statistical Consulting Group, s.f.).
3. **R-cuadrado (R2) o Coeficiente de Determinación:**
   - **Definición:** Mide la proporción de la varianza total en la variable dependiente (Y) que es explicada por el modelo de regresión (es decir, por las variables independientes incluidas) (Statistics By Jim, s.f.-b).
   - **Interpretación:** Varía entre 0 y 1 (o 0% y 100%). Un R2=0 indica que el modelo no explica nada de la variabilidad en Y, mientras que un R2=1 indica que el modelo explica toda la variabilidad. Por ejemplo, un R2=0.70 significa que el 70% de la varianza en Y es explicada por las variables predictoras del modelo (UCLA: Statistical Consulting Group, s.f.).
   - **Limitaciones:** R2 siempre aumenta o se mantiene igual cuando se añaden más variables predictoras al modelo, incluso si estas no son realmente útiles. No

indica si el modelo está bien ajustado o si las estimaciones de los coeficientes son insesgadas (Statistics By Jim, s.f.-b).

4. **R-cuadrado Ajustado (Adjusted R2):**

   - **Definición:** Es una modificación del R2 que penaliza la adición de variables predictoras irrelevantes al modelo. Tiene en cuenta el número de predictores en el modelo (Statistics By Jim, s.f.-b).
   - **Interpretación:** Generalmente es preferible al R2 para comparar modelos con diferente número de predictores. Puede disminuir si se añade una variable que no mejora el modelo lo suficiente. Proporciona una estimación más realista de la bondad de ajuste en la población (UCLA: Statistical Consulting Group, s.f.).

5. **p-value de la Prueba F (para el modelo global):**

   - **Definición:** La prueba F evalúa la hipótesis nula de que todos los coeficientes de regresión (excepto el intercepto) son simultáneamente iguales a cero (H_0:beta_1=beta_2=ldots=beta_k=0). Esto es, si el modelo en su conjunto tiene algún poder predictivo (UCLA: Statistical Consulting Group, s.f.).
   - **Interpretación:** Un p-value pequeño para la prueba F (típicamente $\< 0.05$) indica que se rechaza la hipótesis nula, sugiriendo que al menos una de las variables predictoras tiene una relación significativa con la variable dependiente y que el modelo global es estadísticamente significativo (Universidad de Granada, s.f.).

La evaluación de un modelo de regresión lineal implica analizar estas métricas en conjunto, así como verificar las asunciones del modelo mediante el análisis de residuos.

5. Estadística por Iteración (ej. bootstrapping)

La estadística por iteración, también conocida como métodos de remuestreo (resampling), se refiere a un conjunto de técnicas que utilizan los datos de la muestra observada para generar nuevas muestras simuladas, permitiendo así estimar propiedades de un estimador o la precisión de un modelo (Number Analytics, 2023). Estos métodos son a menudo computacionalmente intensivos pero ofrecen gran flexibilidad, especialmente cuando las asunciones de los métodos estadísticos tradicionales no se cumplen o son difíciles de verificar. El bootstrapping es uno de los ejemplos más prominentes y versátiles de estos enfoques.

Bootstrapping:

El bootstrapping es un método no paramétrico introducido por Bradley Efron en 1979. Su idea central es tratar la muestra original como si fuera la población y generar un gran número de "muestras bootstrap" extrayendo observaciones de la muestra original con reemplazo (Sustainability Methods Wiki, s.f.; Number Analytics, 2023). Cada muestra bootstrap tiene el mismo tamaño que la muestra original.

Propósito del Bootstrapping:

El principal propósito del bootstrapping es estimar la distribución muestral de un estadístico (e.g., la media, la mediana, un coeficiente de regresión, etc.) (FasterCapital, s.f.-a). A partir de esta distribución muestral empírica, se pueden:

1. **Estimar el Error Estándar (SE):** La desviación estándar de la distribución bootstrap del estadístico es una estimación de su error estándar (Number Analytics, 2023).
2. **Construir Intervalos de Confianza (IC):** Se pueden derivar intervalos de confianza para el parámetro de interés utilizando los percentiles de la distribución bootstrap del estadístico (e.g., el método de percentiles) u otros métodos más sofisticados como el BCa (bias-corrected and accelerated) (ResearchGate, 2024; Number Analytics, 2023).
3. **Realizar Pruebas de Hipótesis:** Aunque menos común que para SE e IC, se puede adaptar para pruebas de hipótesis, por ejemplo, evaluando si un intervalo de confianza bootstrap para un parámetro incluye el valor nulo (Sustainability Methods Wiki, s.f.).

Procedimiento Básico del Bootstrapping No Paramétrico (para estimar el SE y un IC para la media, por ejemplo):

El procedimiento, como lo describe Number Analytics (2023), se puede resumir en los siguientes pasos:

1. **Muestra Original:** Se tiene una muestra original de datos $x=(x_1,x_2,\ldots,x_n)$ de tamaño n.
2. **Remuestreo con Reemplazo:** Generar un gran número B de muestras bootstrap (e.g., B=1000 o más). Cada muestra bootstrap, x^{*b} (donde $b=1,\ldots,B$), se obtiene seleccionando n observaciones de la muestra original x con reemplazo. Esto significa que algunas observaciones originales pueden aparecer múltiples veces en una muestra bootstrap, mientras que otras pueden no aparecer en absoluto.
3. **Cálculo del Estadístico:** Para cada una de las B muestras bootstrap, calcular el estadístico de interés, $T(x^{*b})$. Por ejemplo, si se está interesado en la media, se calcularía la media de cada muestra bootstrap, $\bar{x}^{*b}$.
4. **Distribución Bootstrap:** La colección de los B valores del estadístico calculado, $(T(x^{*1}), T(x^{*2}), \ldots, T(x^{*B}))$, forma la distribución empírica bootstrap del estadístico T.
5. **Estimación del Error Estándar:** El error estándar del estadístico T se estima como la desviación estándar de los B valores $T(x^{*b})$.
6. **Construcción de Intervalos de Confianza:** Un intervalo de confianza del $(1-\alpha)100\%$ para el parámetro poblacional se puede construir utilizando los percentiles de la distribución bootstrap. Por ejemplo, un IC del 95% por el método de percentiles se obtendría tomando los valores en los percentiles 2.5 y 97.5 de la

distribución de los $T(x^{*b})$.

Ventajas del Bootstrapping:

- **Pocas Asunciones:** No requiere fuertes asunciones sobre la distribución de la población original (es no paramétrico) (Sustainability Methods Wiki, s.f.; ResearchGate, 2024).
- **Versatilidad:** Se puede aplicar a una amplia variedad de estadísticos, incluso aquellos con distribuciones muestrales complejas o desconocidas donde los métodos analíticos son difíciles (Number Analytics, 2023).
- **Precisión:** A menudo proporciona estimaciones más precisas del error estándar e intervalos de confianza, especialmente con muestras pequeñas o cuando las asunciones paramétricas no se cumplen (Sustainability Methods Wiki, s.f.).

Desventajas del Bootstrapping:

- **Intensidad Computacional:** Puede requerir una cantidad significativa de cálculos, especialmente con muestras grandes o estadísticos complejos (Sustainability Methods Wiki, s.f.).
- **Dependencia de la Muestra Original:** La calidad de los resultados del bootstrap depende de cuán representativa sea la muestra original de la población. Si la muestra original es sesgada o tiene poca información, el bootstrap reflejará estas deficiencias (Number Analytics, 2023).
- **Sensibilidad a Outliers:** Al igual que muchos métodos estadísticos, puede ser sensible a la presencia de outliers en la muestra original (Sustainability Methods Wiki, s.f.).
- **No Siempre Funciona:** Hay situaciones donde el bootstrapping puede fallar o dar resultados pobres (e.g., para estimar extremos de una distribución o con parámetros en el límite del espacio paramétrico).

Aplicaciones:

El bootstrapping se utiliza en una amplia gama de campos:

- **Estadística:** Estimar errores estándar e intervalos de confianza para medias, medianas, proporciones, coeficientes de correlación, coeficientes de regresión, etc.
- **Machine Learning:**
  - **Evaluación de Modelos:** Estimar la variabilidad del rendimiento de un modelo (e.g., precisión, error cuadrático medio) (FasterCapital, s.f.-a).
  - **Bagging (Bootstrap Aggregating):** Técnica de ensamble donde se entrenan múltiples modelos en diferentes muestras bootstrap de los datos de entrenamiento y luego se promedian sus predicciones (e.g., Random Forests) para mejorar la estabilidad y precisión (FasterCapital, s.f.-a).
  - **Estimación de Incertidumbre:** Cuantificar la incertidumbre en las

predicciones o en los parámetros del modelo (Number Analytics, 2023).
  - **Selección de Características:** Evaluar la estabilidad de las características seleccionadas por un algoritmo.

Otras técnicas de remuestreo incluyen la validación cruzada (principalmente para la evaluación de modelos), las pruebas de permutación (para pruebas de hipótesis no paramétricas) y el jackknife (un precursor del bootstrap).

6. Sesgo de Selección y Estrategias de Mitigación

El sesgo de selección es un tipo de error sistemático que ocurre en la investigación cuando los individuos o grupos seleccionados para un estudio difieren de la población objetivo de una manera que afecta los resultados del estudio (Choi & Pak, 2005; FasterCapital, s.f.-b). Esto lleva a que la muestra no sea representativa de la población que se pretende estudiar, lo que compromete la validez y generalizabilidad de las conclusiones.

Tipos y Fuentes Comunes de Sesgo de Selección:

Existen numerosas formas en las que puede surgir el sesgo de selección:

1. **Sesgo de Muestreo (Sampling Bias):** Ocurre cuando el método utilizado para seleccionar la muestra favorece a ciertos individuos o grupos sobre otros.
   - **Sesgo de Cobertura (Undercoverage Bias):** Ciertos grupos de la población no tienen la oportunidad de ser incluidos en la muestra (e.g., no incluir hogares sin teléfono en una encuesta telefónica) (Choi & Pak, 2005).
   - **Sesgo de Autoselección (Volunteer Bias):** Los individuos que se ofrecen voluntariamente para participar en un estudio pueden ser diferentes de aquellos que no se ofrecen (e.g., más saludables, más motivados) (PMC, 2010; FasterCapital, s.f.-b).
   - **Sesgo de No Respuesta (Non-response Bias):** Si las personas que no responden a una encuesta o estudio difieren significativamente de las que sí responden en variables relevantes para el estudio (FasterCapital, s.f.-b).
   - **Sesgo de Berkson (Berkson's Fallacy):** Ocurre en estudios hospitalarios cuando la muestra se toma de pacientes hospitalizados, lo que puede crear asociaciones espurias entre enfermedades (Choi & Pak, 2005).
   - **Sesgo de Neyman (Prevalence-Incidence Bias):** Surge cuando se estudian enfermedades crónicas y se seleccionan casos prevalentes en lugar de incidentes, lo que puede llevar a omitir casos fatales o de corta duración (Choi & Pak, 2005).
2. **Sesgo de Supervivencia (Survivorship Bias):** Se concentra en los individuos o casos que "sobrevivieron" a algún proceso de selección, ignorando a aquellos que no lo hicieron, lo que puede llevar a conclusiones optimistas o erróneas (FasterCapital, s.f.-b). Un ejemplo clásico es analizar solo empresas exitosas para determinar factores de éxito, ignorando las que fracasaron.

3. **Sesgo de Atrición (Attrition Bias / Loss to Follow-up Bias):** Ocurre en estudios longitudinales cuando los participantes que abandonan el estudio difieren sistemáticamente de los que permanecen, afectando la composición de la muestra a lo largo del tiempo (PMC, 2018; Choi & Pak, 2005).
4. **Sesgo de Detección (Detection Bias):** La exposición a un factor de riesgo puede influir en la probabilidad de que se detecte una enfermedad, llevando a una asociación incorrecta (Choi & Pak, 2005).
5. **Sesgo de Exclusión:** Ciertos participantes son excluidos del análisis por razones que están relacionadas con la exposición o el resultado (Formplus Blog, s.f.).

Impacto del Sesgo de Selección:

El sesgo de selección puede tener graves consecuencias para la investigación:

- **Validez Interna Comprometida:** Puede llevar a asociaciones o estimaciones de efectos incorrectas dentro de la muestra estudiada (PMC, 2010; Formplus Blog, s.f.).
- **Validez Externa (Generalizabilidad) Reducida:** Las conclusiones obtenidas de una muestra sesgada no pueden generalizarse de manera confiable a la población objetivo (PMC, 2018; Formplus Blog, s.f.).
- **Conclusiones Erróneas:** Puede llevar a conclusiones falsas sobre relaciones causales, efectividad de tratamientos o prevalencia de condiciones.
- **En Machine Learning:** Si los datos de entrenamiento están afectados por sesgo de selección, el modelo puede aprender patrones incorrectos o no generalizar bien a nuevos datos del mundo real, perpetuando o incluso amplificando desigualdades existentes (Hoyle et al., 2021). Por ejemplo, un modelo de diagnóstico médico entrenado principalmente con datos de un grupo demográfico puede tener un rendimiento inferior en otros grupos.

Estrategias de Mitigación:

Abordar el sesgo de selección requiere una cuidadosa planificación y ejecución del estudio, así como técnicas analíticas.

1. **Definición Clara de la Población Objetivo:** Especificar claramente la población a la que se quieren generalizar los resultados.
2. **Diseño de Muestreo Apropiado:**
    - **Muestreo Probabilístico:** Utilizar métodos como el muestreo aleatorio simple, estratificado o sistemático para asegurar que cada miembro de la población objetivo tenga una probabilidad conocida de ser seleccionado (FasterCapital, s.f.-b). La aleatorización es clave en los ensayos clínicos para prevenir el sesgo de selección en la asignación a grupos de tratamiento (Kahan et al., 2015).
    - **Muestreo Estratificado:** Asegurar la representación adecuada de subgrupos relevantes (Hoyle et al., 2021).
3. **Maximizar las Tasas de Respuesta y Minimizar la Atrición:**

    - En encuestas y estudios longitudinales, implementar estrategias para aumentar la participación y retención, como recordatorios, incentivos, múltiples modos de contacto y asegurar la confidencialidad (PMC, 2010).
4. **Blinding (Cegamiento):** En ensayos clínicos, el cegamiento de los reclutadores y de quienes asignan a los participantes a los grupos de tratamiento puede reducir el sesgo de selección (Kahan et al., 2015).
5. **Recolección de Datos Diversa y Representativa:** Esforzarse por recolectar datos de una amplia gama de fuentes y asegurar que la muestra refleje la diversidad de la población objetivo, especialmente en el contexto de Machine Learning (Hoyle et al., 2021).
6. **Documentación y Transparencia:** Especificar claramente los métodos de selección de la muestra, las tasas de participación y las características de los participantes versus los no participantes o los perdidos en el seguimiento (e.g., mediante diagramas CONSORT) (PMC, 2010).
7. **Técnicas Estadísticas y Analíticas:**
    - **Ajuste Estadístico:** En el análisis, se pueden utilizar técnicas como la ponderación para ajustar las diferencias conocidas entre la muestra y la población, o entre los que respondieron y los que no.
    - **Modelado del Sesgo de Selección:** Métodos como los modelos de selección de Heckman pueden usarse en econometría para corregir el sesgo de selección en ciertos contextos.
    - **Análisis de Sensibilidad:** Evaluar cómo los resultados podrían cambiar bajo diferentes supuestos sobre el sesgo (PMC, 2010).
    - **Propensity Score Matching (PSM):** En estudios observacionales o en Machine Learning, PSM puede ayudar a crear grupos comparables al equilibrar las covariables observadas entre los grupos, mitigando así el sesgo de selección y confusión (Hoyle et al., 2021; ResearchGate, 2024b).
8. **Protocolos Preespecificados:** En contextos como el uso de brazos de control externos (ECA) en ensayos clínicos, la selección a priori del grupo de control externo y la documentación de los enfoques en un protocolo preespecificado y un plan de análisis estadístico son cruciales para reducir el sesgo de selección (ResearchGate, 2024b).
9. **En Machine Learning (Estrategias Específicas):**
    - **Pre-procesamiento:** Técnicas como el sobremuestreo de grupos minoritarios (e.g., SMOTE), submuestreo de grupos mayoritarios o reponderación de instancias (Hoyle et al., 2021).
    - **In-procesamiento:** Modificar algoritmos de aprendizaje para que sean conscientes del sesgo, como la optimización con restricciones o el aprendizaje adversarial (Hoyle et al., 2021).
    - **Post-procesamiento:** Ajustar las predicciones del modelo, como la calibración o la selección de umbrales específicos para diferentes grupos (Hoyle et al., 2021).

Ninguna estrategia es perfecta, y a menudo se requiere una combinación de enfoques para minimizar el impacto del sesgo de selección.

7. Probabilidad Binomial

La distribución binomial es una distribución de probabilidad discreta fundamental que describe el número de éxitos en una secuencia fija de n ensayos de Bernoulli independientes entre sí, donde cada ensayo tiene solo dos resultados posibles (comúnmente denominados "éxito" y "fracaso") y la probabilidad de éxito, p, es constante para cada ensayo (SATHEE, s.f.; Wikipedia, 2024).

Condiciones para una Distribución Binomial:

Un experimento o situación se puede modelar con una distribución binomial si se cumplen las siguientes condiciones (Pickl.AI, s.f.; Khan Academy, s.f.):

1. **Número Fijo de Ensayos (n):** El experimento consta de un número fijo de repeticiones o ensayos, n.
2. **Ensayos Independientes:** El resultado de un ensayo no afecta el resultado de los otros ensayos.
3. **Dos Resultados Posibles:** Cada ensayo tiene solo dos resultados posibles, que pueden clasificarse como "éxito" o "fracaso".
4. **Probabilidad Constante de Éxito (p):** La probabilidad de "éxito", denotada por p, es la misma para cada ensayo. La probabilidad de "fracaso" es, por lo tanto, q=1−p.

Parámetros de la Distribución Binomial:

La distribución binomial está completamente definida por dos parámetros:

- n: el número de ensayos.
- p: la probabilidad de éxito en un solo ensayo.

Se denota como XsimB(n,p), donde X es la variable aleatoria que representa el número de éxitos en n ensayos.

Función de Masa de Probabilidad (FMP):

La probabilidad de obtener exactamente k éxitos en n ensayos está dada por la función de masa de probabilidad binomial (SATHEE, s.f.; Wikipedia, 2024):

P(X=k)=binomnkpk(1−p)n−k

Donde:

- k es el número de éxitos (puede tomar valores enteros desde 0 hasta n).
- binomnk=fracnk(n−k) es el coeficiente binomial, que representa el número de formas

diferentes de obtener k éxitos en n ensayos.
- pk es la probabilidad de obtener k éxitos.
- (1−p)n−k es la probabilidad de obtener n−k fracasos.

Media y Varianza de la Distribución Binomial:

- Media (Valor Esperado): El número esperado de éxitos en n ensayos es (SATHEE, s.f.; Wikipedia, 2024):
  E[X]=mu=np

- Varianza: La varianza del número de éxitos es (SATHEE, s.f.; Wikipedia, 2024):
  Var(X)=sigma2=np(1−p)=npq

- Desviación Estándar: La desviación estándar es la raíz cuadrada de la varianza:
  sigma=sqrtnp(1−p)=sqrtnpq

Propiedades de la Distribución Binomial:

- Es una distribución discreta.
- La suma de dos variables aleatorias binomiales independientes X_1simB(n_1,p) y X_2simB(n_2,p) (con la misma p) es también una variable binomial X_1+X_2simB(n_1+n_2,p) (SATHEE, s.f.).
- Cuando n es grande y p no está demasiado cerca de 0 o 1, la distribución binomial puede aproximarse bien por la distribución normal con media mu=np y varianza sigma2=np(1−p). Una regla común es que la aproximación es buena si npgeq5 y n(1−p)geq5 (Wikipedia, 2024).
- Si n=1, la distribución binomial es una distribución de Bernoulli (SATHEE, s.f.).

Aplicaciones Prácticas:

La distribución binomial se utiliza en una amplia variedad de campos para modelar situaciones con resultados binarios (Pickl.AI, s.f.; Alooba, s.f.):

1. **Control de Calidad:** Estimar la probabilidad de encontrar un cierto número de artículos defectuosos en un lote de producción, asumiendo que la probabilidad de que un artículo sea defectuoso es constante.
2. **Medicina y Salud Pública:** Evaluar la efectividad de un tratamiento (e.g., número de pacientes que se recuperan), la probabilidad de que un paciente experimente un efecto secundario, o modelar la propagación de enfermedades.
3. **Encuestas y Sondeos de Opinión:** Calcular la probabilidad de que un cierto número de encuestados apoye una política o candidato.
4. **Genética:** Modelar la herencia de ciertos rasgos genéticos.
5. **Finanzas y Seguros:** Evaluar el riesgo, como la probabilidad de un cierto número de

impagos de préstamos.

6. **Machine Learning:**
    - **Evaluación de Clasificadores Binarios:** Un modelo de clasificación binaria produce un resultado de "éxito" (clasificación correcta) o "fracaso" (clasificación incorrecta) para cada instancia. Si se evalúa el modelo en un conjunto de n instancias independientes, y se asume una probabilidad p constante de clasificación correcta (la exactitud del modelo), el número de clasificaciones correctas puede modelarse con una distribución binomial. Esto permite, por ejemplo, construir intervalos de confianza para la exactitud del modelo o probar si el rendimiento observado es significativamente mejor que el azar.
    - **Pruebas A/B:** Determinar si una versión de un producto (e.g., una página web) es significativamente mejor que otra en términos de una métrica binaria (e.g., tasa de conversión).

La comprensión de la distribución binomial es esencial para analizar y modelar datos donde los resultados son categóricos y de naturaleza binaria.

Referencias

Alooba. (s.f.). *Binomial Distribution: A Fundamental Concept in Statistics*. Recuperado el 21 de mayo de 2025, de https://www.alooba.com/blog/binomial-distribution-a-fundamental-concept-in-statistics

Bazán, J. L. (2004). Muestreo estadístico: Conceptos básicos y ejemplos de aplicación. *AJAYU Órgano de Difusión Científica del Departamento de Psicología UCBSP*, *2*(1), 73-112. http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S1815-02762004000100012

Choi, B. C. K., & Pak, A. W. P. (2005). Bias, overview. En P. Armitage & T. Colton (Eds.), *Encyclopedia of Biostatistics* (2nd ed.). Wiley. (Trabajo recuperado de https://rossisanusi.wordpress.com/wp-content/uploads/2020/10/choi_overview-of-bias.pdf)

Cornell University. (s.f.). *12.4 The Central Limit Theorem*. CS1380 Data Science for All. Recuperado el 20 de mayo de 2025, de https://www.cs.cornell.edu/courses/cs1380/2018sp/textbook/chapters/12/4/central-limit-theorem.html

DataCamp. (s.f.). *Teorema central del límite: Explicación de un concepto clave de la estadística*. Recuperado el 20 de mayo de 2025, de https://www.datacamp.com/es/tutorial/central-limit-theorem

FasterCapital. (s.f.-a). *Bootstrapping technique: The Power of Bootstrapping: Enhancing Model Stability*. Recuperado el 20 de mayo de 2025, de https://www.fastercapital.com/content/Bootstrapping-technique--The-Power-of-Bootstrappin g--Enhancing-Model-Stability.html

FasterCapital. (s.f.-b). *Volunteer Bias*. Recuperado el 20 de mayo de 2025, de https://fastercapital.com/keyword/volunteer-bias.html

Formplus Blog. (s.f.). *Selection Bias in Research: Types, Examples & Impact*. Recuperado el 21 de mayo de 2025, de https://www.formpl.us/blog/selection-bias

Fortress Diagnostics. (2025, 29 de abril). *The Critical Role of Quality Control in Diagnostic Testing*. Recuperado el 20 de mayo de 2025, de https://www.fortressdiagnostics.com/news/2025/april/the-critical-role-of-quality-control-in-diagnostic-testing

Hoyle, C., Meng, X., Seneviratne, S., & Yu, T. (2021). *Review of Data Bias in Healthcare AI Applications*. Chapman University Digital Commons. https://digitalcommons.chapman.edu/cgi/viewcontent.cgi?article=1226&context=pt_articles

IBM. (2021, 18 de agosto). *What Is Linear Regression?* IBM. Recuperado el 20 de mayo de 2025, de https://www.ibm.com/think/topics/linear-regression

Kahan, B. C., Rehal, S., & Cro, S. (2015). Risk of selection bias in randomised trials. *Trials*, *16*(1), 405. https://doi.org/10.1186/s13063-015-0921-9

Khan Academy. (s.f.). *Random variables and probability distributions*. [Unidad del curso de AP Statistics]. Recuperado el 21 de mayo de 2025, de https://www.khanacademy.org/math/ap-statistics/random-variables-ap[3] (Información específica sobre variables binomiales, condiciones, parámetros, PMF, CDF, valor esperado y varianza se encuentra dentro de esta unidad).

Kwak, S. G., & Kim, J. H. (2017). Central limit theorem: the cornerstone of modern statistics. *Korean Journal of Anesthesiology*, *70*(2), 144–156. https://doi.org/10.4097/kjae.2017.70.2.144

Lumen Learning. (s.f.). *Sampling and Data*. Introduction to Statistics. Recuperado el 20 de mayo de 2025, de https://courses.lumenlearning.com/introstats1/chapter/sampling-and-data/

Nahm, F. S. (2010). Hypothesis testing, type I and type II errors. *Korean Journal of Anesthesiology*, *58*(6), 585. (Error en la cita original, el artículo correcto con este tema y autor es de 2022 o el referenciado para el PMC es más adecuado) *Nota: Se utilizará la cita del artículo de PMC (2010) referenciado a continuación para este concepto, ya que la cita original parece tener un error de año y la información es consistente.*

Nahm, F. S. (2010). Understanding statistical inference:

hypothesis testing, type I, and type II errors. Anesthesia and Pain Medicine, 5(4), 289-292. (Referenciado como PMC2996198, Diciembre 2010). Recuperado de https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2996198/

Newcastle University. (s.f.). *Types of Sampling*. Maths Resources. Recuperado el 20 de mayo de 2025, de https://www.ncl.ac.uk/webtemplate/ask-assets/external/maths-resources/statistics/sampling/types-of-sampling.html

Number Analytics. (2023, 1 de diciembre). *Advanced Techniques for Bootstrap Confidence Interval in Modern Statistics*. Recuperado el 20 de mayo de 2025, de https://www.numberanalytics.com/blog/advanced-bootstrap-confidence-interval-technologies

Penn State Eberly College of Science. (s.f.-a). *7.4 - Central Limit Theorem*. STAT 200: Elementary Statistics. Recuperado el 20 de mayo de 2025, de https://online.stat.psu.edu/stat200/lesson/7/7.4

Penn State Eberly College of Science. (s.f.-b). *Lesson 9: Simple Linear Regression*. STAT 500: Applied Statistics. Recuperado el 20 de mayo de 2025, de https://online.stat.psu.edu/stat500/lesson/9 (Las asunciones se detallan en subsecciones como 9.2.3).

Pickl.AI. (s.f.). *Understanding Binomial Distribution in Statistics*. Recuperado el 21 de mayo de 2025, de https://www.pickl.ai/blog/binomial-distribution-in-statistics/

PMC. (2010). *Assessment of self-selection bias in a pediatric unilateral hearing loss study*. [PMC2975441]. Recuperado el 21 de mayo de 2025, de https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2975441/

PMC. (2018). *Modeling competence development in the presence of selection bias*. [PMC6267521]. Recuperado el 21 de mayo de 2025, de https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6267521/

ResearchGate. (2024a, Enero). *Nonparametric Bootstrap Likelihood Estimation to Investigate the Chance Set-up on Clustering Results*. [Publicación 389311462]. Recuperado el 20 de mayo de 2025, de https://www.researchgate.net/publication/389311462_Nonparametric_Bootstrap_Likelihood_Estimation_to_Investigate_the_Chance_Set-up_on_Clustering_Results

ResearchGate. (2024b, Febrero). *The Next Horizon of Drug Development: External Control Arms and Innovative Tools to Enrich Clinical Trial Data*. [Publicación 379279449]. Recuperado el 21 de mayo de 2025, de https://www.researchgate.net/publication/379279449_The_Next_Horizon_of_Drug_Developme nt_External_Control_Arms_and_Innovative_Tools_to_Enrich_Clinical_Trial_Data

SATHEE. (s.f.). *Maths Binomial Distribution*. Prutor.ai. Recuperado el 21 de mayo de 2025, de https://sathee.prutor.ai/article/maths/maths-binomial-distribution/

Scribbr. (2023, 22 de junio). *Type I & Type II Errors | Differences, Examples, Visualizations*. Recuperado el 20 de mayo de 2025, de https://www.scribbr.com/statistics/type-i-and-type-ii-errors/

Statistics By Jim. (s.f.-a). *Central Limit Theorem Explained*. Recuperado el 20 de mayo de 2025, de https://statisticsbyjim.com/basics/central-limit-theorem/ (Nota: Aunque este sitio es útil, se priorizaron fuentes académicas o ECR para el TLC; la información es consistente con las fuentes citadas).

Statistics By Jim. (s.f.-b). *How To Interpret R-squared in Regression Analysis*. Recuperado el 20 de mayo de 2025, de https://statisticsbyjim.com/regression/interpret-r-squared-regression/

Statistics LibreTexts. (2021, 11 de septiembre). *9.2: Type I and Type II Errors*. (CUNY, Introductory Statistics with Probability). Recuperado el 20 de mayo de 2025, de https://stats.libretexts.org/Courses/City_University_of_New_York/Introductory_Statistics_with_Probability_(CUNY)/09%3A_Hypothesis_Testing_for_a_Single_Variable_and_Population/9.02%3A_Type_I_and_Type_II_Errors

Statistics LibreTexts. (2023a, 28 de julio). *7: The Central Limit Theorem*. (Los Angeles City College, Introductory Statistics). Recuperado el 20 de mayo de 2025, de https://stats.libretexts.org/Courses/Los_Angeles_City_College/Introductory_Statistics/07%3A_The_Central_Limit_Theorem

Sustainability Methods Wiki. (s.f.). *Bootstrapping in Python*. Recuperado el 20 de mayo de 2025, de https://sustainabilitymethods.org/index.php/Bootstrapping_in_Python

Tille, A. (s.f.). *Introduction to Statistical Sampling and Resampling*. Towards Data Science. Recuperado el 20 de mayo de 2025, de https://towardsdatascience.com/introduction-to-statistical-sampling-and-resampling-1a6110965c3a

UCLA: Statistical Consulting Group. (s.f.). *Regression Analysis | SPSS Annotated Output*. Recuperado el 20 de mayo de 2025, de https://stats.oarc.ucla.edu/spss/output/regression-analysis/

Universidad de Granada. (s.f.). *Práctica 3 | Estadística*. Recuperado el 20 de mayo de 2025, de https://wpd.ugr.es/~bioestad/guia-spss/practica-3/

Wikipedia. (2024, 17 de mayo). *Binomial distribution*. Recuperado el 21 de mayo de 2025, de

https://en.wikipedia.org/wiki/Binomial_distribution