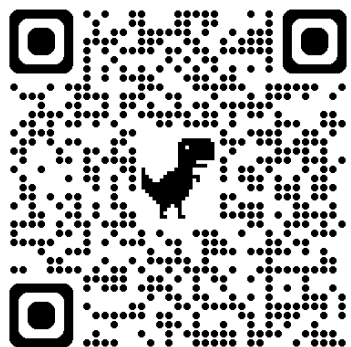


Inferencia Estadística con Datos Reales

De Instagram al Aula Escolar

Un taller práctico sobre metodologías activas y enseñanza contextualizada



Dr. Pedro Vidal Szabó



La Estadística Vive en el Mundo Digital

Nuestros estudiantes interactúan constantemente con datos:

- Métricas de redes sociales
- Estadísticas de videojuegos
- Tendencias digitales
- Análisis de contenido

La clave: aprovechar esta realidad para enseñar estadística de manera relevante y motivadora.



Objetivos del Taller



Vivenciar Investigación

Ciclo completo: desde recolección hasta interpretación de resultados



Aplicar Inferencia

Estimar parámetros usando intervalos de confianza



Usar Herramientas

Excel y GeoGebra para análisis y visualización



Reflexionar

Proyecciones didácticas y ética en el tratamiento de datos

Fundamentos: Parámetro vs. Estadístico

Parámetro

Valor **fijo y único** que describe la **población completa**

- La "verdad" que queremos conocer
- Ejemplos: μ (media poblacional), p (proporción poblacional)
- Casi siempre **desconocido**

Estadístico

Valor **variable** calculado desde una **muestra**

- Se usa para **estimar** el parámetro
- Ejemplos: \bar{x} (media muestral), \hat{p} (proporción muestral)
- Cambia de muestra a muestra



Variabilidad Muestral y el TCL

¿Por qué muestras diferentes dan resultados diferentes?

Variabilidad muestral: fenómeno natural donde cada muestra produce un estadístico ligeramente distinto.

Una sola muestra no da la "respuesta exacta", sino una **estimación**.

- ❏ **Teorema Central del Límite (TCL):** Con $n \geq 30$ (habitualmente en teoría, no es necesariamente así en la realidad), la distribución de estadísticos muestrales se comporta de manera predecible y aproximadamente normal. Esto nos da confianza para inferir sobre la población.

⚠ Nota importante:

Si la muestra es **pequeña** ($n < 30$) y no se conoce σ , se debe usar la **distribución t de Student**:

$$IC_{95\%} = \bar{x} \pm t_{\alpha/2, df} \cdot \frac{s}{\sqrt{n}}$$

donde $t_{\alpha/2, df}$ es el valor crítico t con $df = n - 1$ grados de libertad.

La Lógica del Intervalo de Confianza

Un intervalo de confianza es un **rango de valores plausibles** para un parámetro poblacional desconocido.

1

No es exacto

Captura la incertidumbre del muestreo

2

IC 95% significa

95% de los intervalos contendrían μ si repitiéramos el proceso (**IMPORTANTE**)

3

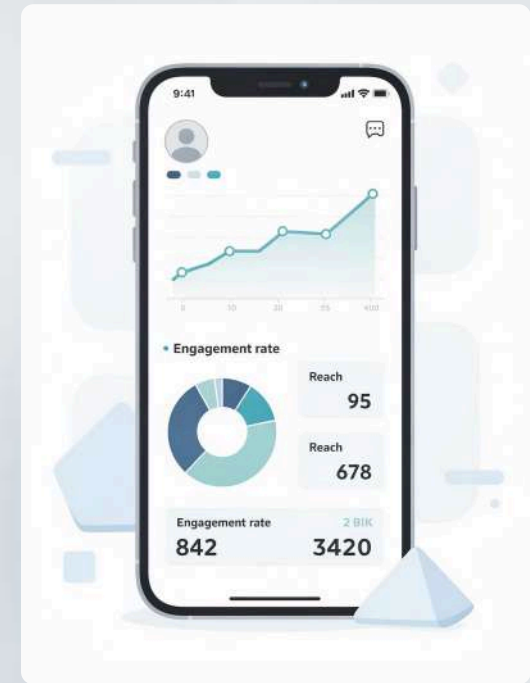
Interpretación procesal

No "probabilidad de que μ esté aquí"

Actividad Central

Análisis de Perfiles de Instagram

Aplicamos conceptos de parámetro, estadístico, variabilidad y confianza en un contexto cotidiano para los estudiantes.



Diseño del Estudio

01

Objetivo

Estimar media de publicaciones (μ : media poblacional) y proporción de cuentas "influyentes" (p : proporción poblacional)

02

Población

(Caso particular de una pareja de estudiantes de pregrado)

N = 306 cuentas seguidas (Individuo A: 102, Individuo B: 204)

03

Muestreo

Aleatorio proporcional: n = 41 (14 de A, 27 de B)

04

Ética

Anonimato total: cuentas codificadas, solo datos públicos numéricos



Análisis 1: Media de Publicaciones

Estadísticos muestrales

- $\bar{x} = 162$ publicaciones
- $s = 534.806$
- $n = 41$

Fórmula IC 95%:

$$IC_{95\%} = \bar{x} \pm 1.96 \cdot \frac{s}{\sqrt{n}}$$



Distribución	Estadísticas
Intervalo Z de una media	
Nivel de confianza 0.95	
Muestra	
Media	162
σ	534.806
n	41
Resultado	
Intervalo Z de una media	
Media	162
σ	534.806
ES	83.5227
n	41
Extremo inferior	-1.7014
Extremo superior	325.7014
Intervalo	162 \pm 163.7014

Resultado

IC 95% ajustado: [0, 326] publicaciones

Interpretación: Enorme amplitud refleja **alta variabilidad** en tipos de cuentas: desde perfiles personales hasta grandes creadores.

Análisis 2: Cuentas "Influyentes"

1

Definición

Cuenta "influyente"

si **Ratio** ≥ 2

$$R = \frac{\text{Seguidores}}{\text{Seguidos}}$$

2

Datos

10 de 41 cuentas

$\hat{p} = 24.39\%$

Fórmula IC 95%:

$$IC_{95\%} = \hat{p} \pm 1.96 \cdot \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

3

Resultado IC 95%

11% - 38%

El feed no está dominado por influencers

¿Cuándo se usa esta fórmula?

Se utiliza para estimar un intervalo de confianza para la proporción poblacional p , cuando:

- \hat{p} es la proporción muestral.
- n es el tamaño de la muestra.
- Se usa el valor crítico 1.96, que corresponde al nivel de confianza del 95 %.
- La muestra es suficientemente grande, tal que:

$$n\hat{p} \geq 5 \quad \text{y} \quad n(1 - \hat{p}) \geq 5$$

Esto asegura que la distribución binomial pueda aproximarse con la normal (por el Teorema Central del Límite).

Desglose de la fórmula:

- \hat{p} : Proporción muestral (por ejemplo, si 40 de 100 personas dicen "sí", entonces $\hat{p} = 0.4$)
- $\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$: Error estándar de la proporción
- 1.96: Valor crítico z para un nivel de confianza del 95 %

Distribución	Estadísticas
Intervalo Z de una proporción	
Nivel de confianza 0.95	
Muestra	
Éxitos	10
n	41
Resultado	
Intervalo Z de una proporción	
Éxitos	10
n	41
ES	0.0671
Extremo inferior	0.1125
Extremo superior	0.3754
Intervalo	0.2439 ± 0.1314

Visualización Dinámica con GeoGebra



Beneficios pedagógicos

- **Transnumeración:** ver datos desde diferentes representaciones
- **Pensamiento crítico:** comprender impacto del tamaño muestral
- **Interactividad:** simular "qué pasaría si..."

Permite a estudiantes experimentar con n y observar cómo cambia la amplitud del IC.

Contextos Múltiples

Actividades Análogas para el Aula

Aplicar los mismos principios inferenciales en contextos variados ayuda a los estudiantes a **generalizar el método** y apreciar su universalidad.



Learn Probability and Statistics – GeoGebra Math Resources

Make math come alive with our collection of over 74 free and interactive GeoGebra resources for statistics and probability for grades 4–8. Enhance students' understanding of statistics and probability with our engaging and interactive resources. Improve math proficiency an...



Caso 1: ¿Cuántos Gajos tiene una Mandarina?



Objetivo

Estimar μ (gajos promedio)

Intervalo T de una media

Nivel de confianza 0.95

Muestra

Media 10.7714

s 1.0791

n 70

Resultado

Intervalo T de una media

Media	10.7714
s	1.0791
ES	0.129
n	70
glib	69
Extremo inferior	10.5141
Extremo superior	11.0287
Intervalo	10.7714 ± 0.2573

IC 95% para μ : [10, 11] |

Intervalos estrechos → alta homogeneidad y calidad consistente del producto.

Reflexiones Didácticas: Mandarinas

Nivel curricular

7° Básico a II° Medio

Metodología

Experimentación directa por estudiantes

1

2

3

4

Conexión

Descriptiva → Inferencia

Herramientas

Simuladores GeoGebra para visualizar
repetición



Caso 2: ¿Quién Ganará la Carrera?

Objetivo

Estimar p (probabilidad victoria Auto 7)

Datos (n=30 carreras)

- $\hat{p} = 40\%$ victorias Auto 7

CONCLUSIONES

Existe un 95% de confianza en que la probabilidad de que el auto 7 gane cualquier carrera se encuentra aproximadamente entre 23% y 58%.

Notar que la gran amplitud del intervalo es el resultado de tener un tamaño de muestra relativamente pequeño ($n = 30$).

Intervalo Z de una proporción

Nivel de confianza 0.95

Muestra

Éxitos 12

n 30

Resultado

Intervalo Z de una proporción

Éxitos 12

n 30

ES 0.0894

Extremo inferior 0.2247

Extremo superior 0.5753

Intervalo 0.4 ± 0.1753



Implementación Gradual por Nivel



7° Básico

Probabilidad experimental y exploración de variable dicotómica (victoria)



8° Básico - 1° Medio

Estadística descriptiva: tendencia central, dispersión. Inferencia informal.



II° Medio

Estudio inferencial completo: TCL, interpretación formal de IC

Estrategias para el Aula



Partir de lo Descriptivo

Base sólida en descripción y visualización antes de generalizar



Fomentar Experimentación

Que estudiantes recolecten datos para vivenciar variabilidad



Utilizar Simulación

GeoGebra para visualizar TCL y construcción de IC



Enfatizar Interpretación

"¿Qué significa este resultado?" debe ser constante

Errores Conceptuales Frecuentes

Error	Descripción
Confusión de parámetros	Mezclar \bar{x} (estimación variable) con μ (parámetro fijo)
Interpretación del IC	Creer que IC 95% = "95% de probabilidad que μ esté aquí"
Omisión de contexto	Calcular sin conectar conclusiones al problema real
Generalización apresurada	Ignorar la incertidumbre inherente al proceso
Causalidad vs. aleatoriedad	Concluir causa cuando puede ser variabilidad aleatoria
Falta de justificación	No justificar decisiones metodológicas
Errores de procedimiento	Usar n en lugar de \sqrt{n} en el error estándar



Mensaje Central del Taller

La inferencia estadística trasciende el mero cálculo

Aplicar un mismo método riguroso a contextos diversos —**Instagram**, **mandarinas**, **carreras de autos**— transforma la estadística en una **herramienta universal** para el pensamiento crítico.

Equipar a estudiantes con estas habilidades

Fomentar ciudadanía informada y reflexiva

Promover decisiones basadas en evidencia

Referencias y Recursos



Lecturas recomendadas

- Batanero, C., Burrill, G., & Reading, C. (2016). *Teaching statistics in school mathematics*. Springer.
- Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning*. Springer.
- Konold, C., & Miller, C. (2005). *TinkerPlots: Dynamic data exploration*. Key Curriculum Press.

Herramientas digitales: Excel, GeoGebra, TinkerPlots

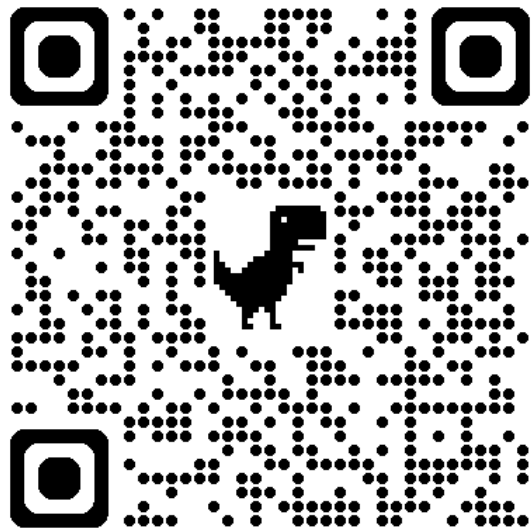


De_Instagram_al_Aula.mp4

Inferencia Estadística con Datos Reales

De Instagram al Aula Escolar

Un taller práctico sobre metodologías activas y enseñanza contextualizada



Instagram



Dr. Vidal-Szabó (@drpedrovidalszabo) • In...

295 Followers, 741 Following, 8 Posts – See Instagram photos and videos from Dr. Vidal-...



PREGUNTAS - COMENTARIOS

Dr. Pedro Vidal Szabó

pedro.vidalszabo@ulagos.cl