

Pruebas A/B en entorno profesional

Objetivo de la Sesión

El objetivo de esta sesión es comprender el proceso de creación, planificación y análisis de **pruebas A/B** para apoyar la toma de decisiones empresariales, especialmente en la **fijación de precios**. Aprenderemos cómo diseñar un experimento controlado para evaluar cambios de precio, analizar los resultados y obtener conclusiones valiosas que puedan guiar a la empresa en la mejora de sus indicadores clave de rendimiento (KPIs).

¿Qué es una Prueba A/B?

Las **pruebas A/B** son un método de experimentación utilizado para comparar dos versiones de una variable (A y B) para determinar cuál de ellas tiene mejor desempeño en función de un KPI o métrica. Este tipo de prueba es muy común en el marketing, implementación de modelos productivos de machine learning y también en la fijación de precios de productos y servicios.

En el contexto de fijación de precios, una prueba A/B puede ser útil para comparar el impacto de dos precios diferentes sobre el comportamiento del consumidor, como la tasa de conversión, ingresos o el valor del ciclo de vida del cliente.

Caso Práctico: Pruebas A/B en la Empresa ABC

Imaginemos que la empresa **ABC**, que ofrece un servicio de suscripción, está evaluando ajustar el precio de uno de sus planes. Actualmente, tienen un plan mensual a **10 USD** y quieren probar si aumentar el precio a **12 USD** afectaría sus tasas de conversión y retención de clientes. Para tomar esta decisión, realizarán una **prueba A/B**.

Contexto del Caso:

- **Problema:** La empresa está considerando aumentar el precio de su plan de suscripción mensual, pero necesita datos para respaldar la decisión.
- **Hipótesis:** El aumento de precio a **12 USD** podría reducir las conversiones, pero también podría aumentar los ingresos por cliente si el impacto negativo en la conversión es limitado.
- **Objetivo:** Evaluar el efecto de un aumento de precio sobre dos métricas principales:
 1. **Tasa de conversión:** Proporción de usuarios que se suscriben después de ver el nuevo precio.
 2. **Valor de por vida del cliente (LTV):** El valor esperado que un cliente generará a lo largo de su relación con la empresa.

Proceso para Ejecutar una Prueba A/B

Paso 1: Definir la Hipótesis y los KPIs

El primer paso para realizar una prueba A/B es formular la hipótesis que deseas probar. En este caso, la hipótesis es que el cambio de precio impactará tanto la **tasa de conversión** como el **LTV** de los clientes. Los KPIs principales que mediremos son:

- **Tasa de conversión:** Proporción de usuarios que completan la suscripción.
- **Ingresos:** La cantidad total de ingresos generados por el grupo de prueba.
- **LTV:** El valor total que un cliente promedio generará para la empresa.

$$H_0 = \Delta TC = 0\$$$

$$H_i = \Delta TC > 0\$$$

Paso 2: Selección de la Muestra y Segmentación

- **Cambios grandes es mejor usar muestras pequeñas**
- **Cambios pequeños es mejor usar muestras grandes**

Para llevar a cabo la prueba A/B, es crucial dividir a los usuarios en dos grupos aleatorios y comparables:

- **Grupo A:** Usuarios que ven el precio original de 10 USD.
- **Grupo B:** Usuarios que ven el nuevo precio de 12 USD.

La segmentación debe realizarse de manera aleatoria para garantizar que ambos grupos sean representativos y no haya sesgo en los resultados.

Antes de dividir a los usuarios en dos grupos, es crucial calcular el **tamaño de la muestra** necesario para asegurar que los resultados de la prueba A/B sean **estadísticamente significativos**. El tamaño de la muestra dependerá de varios factores, como la **tasa de conversión base**, el **nivel de confianza** y el **poder estadístico** que se espera obtener con la prueba.

Factores clave para calcular el tamaño de la muestra:

1. **Tasa de conversión base:** La tasa de conversión actual del grupo de control (por ejemplo, si el 5% de los usuarios actualmente se suscriben al plan de 10 USD).
2. **Magnitud del cambio esperado:** La diferencia mínima en la tasa de conversión que se espera detectar entre los dos grupos (por ejemplo, esperar un aumento de la tasa de conversión del 5% al 6% con el precio de 12 USD).
3. **Nivel de confianza:** El nivel de confianza, comúnmente del 95%, es la probabilidad de que el resultado observado no sea debido al azar.
4. **Poder estadístico:** El poder estadístico, normalmente del 80%, es la probabilidad de detectar un verdadero efecto si este existe.

Fórmula simplificada para el cálculo para proporciones:

Se puede usar una calculadora de tamaño de muestra en línea o herramientas estadísticas como Python (`statsmodels`) para hacer este cálculo. Una fórmula simplificada para el tamaño de la muestra es:

$$n = \frac{2 \times (Z_{\alpha/2} + Z_{\beta})^2 \times p(1-p)}{(p_1 - p_2)^2}$$

Donde:

- $Z_{\alpha/2}$ es el valor Z correspondiente al nivel de confianza (1.96 para el 95%).
- Z_{β} es el valor Z correspondiente al poder estadístico (0.84 para el 80%).
- p es la tasa de conversión base.
- p_1 y p_2 son las tasas de conversión esperadas para los grupos A y B.

Ejemplo con tasa de conversión:

- **Tasa de conversión base:** 5% (0.05).
- **Diferencia esperada en la tasa de conversión:** del 5% al 6% (0.06 - 0.05).
- **Nivel de confianza:** 95% (valor Z = 1.96).
- **Poder estadístico:** 80% (valor Z = 0.84).

Con estos valores, podríamos calcular el tamaño de la muestra usando la fórmula o una herramienta en línea.

¿Cual es el tamaño mínimo de la muestra necesaria en cada grupo?

Fórmula para el cálculo del tamaño de la muestra con valores:

$$n = \frac{2 \times (Z_{\alpha/2} + Z_{\beta})^2 \times \sigma^2}{\Delta^2}$$

\$7448\$

Donde:

- $Z_{\alpha/2}$: El valor Z para el nivel de confianza (1.96 para 95% de confianza).
- Z_{β} : El valor Z para el poder estadístico (0.84 para 80% de poder).
- σ : Desviación estándar del LTV.
- Δ : Diferencia mínima en el LTV que queremos detectar entre los dos grupos.

Ejemplo con tasa de conversión:

Imaginemos que:

- El **LTV actual** (μ_1) es de 100 USD.
- El **LTV esperado** (μ_2) con el nuevo precio es de 120 USD.
- La **desviación estándar** (σ) es de 30 USD.
- La **diferencia esperada en el LTV** ($\Delta = \mu_2 - \mu_1$) es de 20 USD.
- Queremos un **nivel de confianza** del 95% y un **poder estadístico** del 80%.

¿Cuál es el tamaño mínimo de la muestra necesaria en cada grupo?

¿Qué creen que pase entre mas chica es la diferencia que queremos comprobar?

Paso 3: Ejecutar la Prueba

La empresa ABC debe ejecutar la prueba durante un período definido (por ejemplo, 4 semanas), asegurándose de que la prueba cubra suficientes usuarios para obtener resultados estadísticamente significativos.

Durante la prueba, se monitorearán los KPIs en ambos grupos para observar cómo se comportan con los dos precios diferentes.

Paso 4: Análisis de Resultados

Una vez finalizada la prueba, se analizarán los resultados mediante técnicas estadísticas para determinar si el cambio de precio tiene un impacto significativo en los KPIs definidos.

Los puntos clave del análisis incluyen:

1. **Comparar tasas de conversión:** ¿El grupo B (12 USD) tuvo una tasa de conversión significativamente menor que el grupo A (10 USD)?
2. **Ingresos totales:** ¿El grupo B generó mayores ingresos por cliente a pesar de una posible caída en las conversiones?
3. **Valor del Tiempo de Vida (LTV):** ¿El LTV promedio del grupo B fue superior o inferior al del grupo A?

Una vez finalizada la prueba, es fundamental analizar los resultados para determinar si el cambio de precio ha tenido un impacto **estadísticamente significativo** en los KPIs, particularmente en el **Valor del Tiempo de Vida (LTV)**.

En lugar de simplemente observar las diferencias en las medias, utilizaremos un **t-test** para comprobar si la diferencia en los KPIs entre los grupos es significativa desde un punto de vista estadístico.

Pasos Clave en el Análisis:

1. **Comparar LTV promedio entre Grupo A (10 USD) y Grupo B (12 USD).**
2. **Realizar una prueba t de dos muestras** para comparar las medias de LTV en ambos grupos:
 - **Hipótesis nula (\$H_0\$):** No hay diferencia significativa en el LTV entre los dos grupos.
 - **Hipótesis alternativa (\$H_1\$):** Existe una diferencia significativa en el LTV entre los dos grupos.
3. **Calcular el valor p** para determinar si podemos rechazar la hipótesis nula:

- Si $p \leq 0.05$, rechazamos H_0 y concluimos que la diferencia en LTV es estadísticamente significativa.
- Si $p > 0.05$, no hay evidencia suficiente para rechazar H_0 , por lo tanto, no podemos afirmar que haya una diferencia significativa.

```
In [3]: # Mismo número de observaciones en ambas muestras
# Compruebo si son distintos los LTV

import numpy as np
import scipy.stats as stats

# LTV del Grupo A y Grupo B
ltv_grupo_a = [100, 110, 105, 98, 102, 95, 107] # LTV con precio de 10 USD
ltv_grupo_b = [120, 125, 118, 130, 115, 122, 119] # LTV con precio de 12 US

# Realizamos un t-test de dos muestras
t_stat, p_value = stats.ttest_ind(ltv_grupo_a, ltv_grupo_b)

# Mostramos los resultados
print(f"T-statistic: {t_stat}, P-value: {p_value}")
```

T-statistic: -6.906043053510043, P-value: 1.6370006736192083e-05

```
In [4]: # Experimento controlado, distinto número de observaciones
# Compruebo si son distintos los LTV
ltv_grupo_a = [100, 16, 105, 98, 102, 92, 107, 105, 112, 101, 106] # Más ob
ltv_grupo_b = [120, 125, 118, 130, 115, 122, 119] # Menos observaciones

# Realizamos un Welch's t-test, que es robusto a tamaños de muestras desigua
t_stat, p_value = stats.ttest_ind(ltv_grupo_a, ltv_grupo_b, equal_var=False)

# Mostramos los resultados
print(f"T-statistic: {t_stat}, P-value: {p_value}")
```

T-statistic: -3.1924461222722558, P-value: 0.008514555865328299

¿Qué decisión tomarían?

Sin embargo el objetivo era aumentar el LTV 20 USD, ¿esto se cumplió?

```
In [7]: # Objetivo: tener diferencia de 20 dolares en el LTV
objective_difference = 20

# Promedio y desviación estandar por grupo
mean_a = np.mean(ltv_grupo_a)
mean_b = np.mean(ltv_grupo_b)

std_a = np.std(ltv_grupo_a, ddof=1)
std_b = np.std(ltv_grupo_b, ddof=1)

n_a = len(ltv_grupo_a)
n_b = len(ltv_grupo_b)

# Realizar un t-test de una cola (mayor o igual que el objetivo)
# Null hypothesis H0: (mean_b - mean_a) <= objective_difference
```

```
# Alternative hypothesis H1: (mean_b - mean_a) > objective_difference

# Calculate the t-statistic and p-value
pooled_std = np.sqrt(((std_a**2 / n_a) + (std_b**2 / n_b)))
t_stat = (mean_b - mean_a - objective_difference) / pooled_std
p_value = 1 - stats.t.cdf(t_stat, df=min(n_a, n_b) - 1)

# Output the results
print(f"Mean LTV of Group A: {mean_a:.2f}")
print(f"Mean LTV of Group B: {mean_b:.2f}")
print(f"Objective Difference: {objective_difference}")
print(f"T-statistic: {t_stat:.4f}")
print(f"P-value: {p_value:.4f}")
```

Mean LTV of Group A: 94.91
 Mean LTV of Group B: 121.29
 Objective Difference: 20
 T-statistic: 0.7718
 P-value: 0.2348

¿Qué decisión tomarían?

```
In [9]: # Comparar conversiones
import numpy as np
import statsmodels.api as sm

conversions_a = 48
total_a = 1000
conversions_b = 30
total_b = 900

# Calcular conversion rate
rate_a = conversions_a / total_a
rate_b = conversions_b / total_b

# Hacer un z-test de proporciones
count = np.array([conversions_a, conversions_b])
nobs = np.array([total_a, total_b])
z_stat, p_value = sm.stats.proportions_ztest(count, nobs)

print(f"Z-statistic: {z_stat}, P-value: {p_value}")
```

Z-statistic: 1.6088219735707359, P-value: 0.10765527467052376

¿Qué decisión tomarían?

Paso 5: Toma de Decisiones

Con base en los resultados del análisis, la empresa ABC tomará una decisión fundamentada sobre si incrementar el precio a 12 USD. Si el incremento en el precio compensa la disminución en las conversiones y resulta en un mayor LTV e ingresos, podría ser una buena estrategia.

Conclusión

Las pruebas A/B son una herramienta poderosa en la toma de decisiones empresariales, especialmente en la fijación de precios. Al realizar experimentos controlados y analizar los datos de manera rigurosa, las empresas pueden minimizar el riesgo asociado con los cambios de precio, garantizando que sus decisiones estén basadas en datos concretos.

Actividad de la Sesión

1. **Discusión en clase:** Analizaremos ejemplos de pruebas A/B reales y discutiremos cómo aplicar esta metodología en diferentes áreas de negocio, como marketing, cambios en ambiente productivo, diseño de productos y fijación de precios.
2. **Trabajo en equipo:** Diseñen una prueba A/B para evaluar una hipótesis de cambio de precio.

Trabajo de equipo

CompuTech, una empresa de servicios de consulta de AI a través de API, está evaluando cambiar el precio de uno de sus servicios más populares. Tienen como objetivo incrementar 20 USD el ingreso promedio por usuario.

El servicio tiene actualmente un precio de 1 USD por consulta, y la empresa está considerando aumentar el precio a 1.25 USD. Antes de implementar este cambio, desean evaluar cómo este aumento afectará su KPI ventas por usuario. Se llevará a cabo una prueba A/B para respaldar esta decisión.

Objetivo del Caso

El objetivo de la prueba es determinar si el cambio de precio tiene un impacto significativo en los siguientes KPIs:

Ingreso promedio por usuario: Ingreso promedio por usuario. Desean que por lo menos aumente un 20 USD.

Tasa de Conversión: Proporción de usuarios que completan la suscripción.

Hipótesis

- Hipótesis Nula (\$H_0\$): El aumento de precio a 1.25 USD no tiene un impacto significativo en la tasa de conversión ni en el ingreso promedio por usuario.
- Hipótesis Alternativa (\$H_1\$): El aumento de precio a 1.25 USD tiene un impacto significativo en la tasa de conversión y el ingreso promedio por usuario.

Reporte de Laboratorio

```
In [13]: import pandas as pd
datos = pd.read_csv('data/historical_arpu.csv')
datos.head()
```

Out [13]:

	User_ID	ARPU
0	1	107.450712
1	2	97.926035
2	3	109.715328
3	4	122.845448
4	5	96.487699

El B es el de precio modificado

- $p \leq 0.05$: Se rechaza H_0
- $p > 0.05$: No se Rechaza H_0

```
In [15]: conversion_A = pd.read_csv('Data/group_A_conversion.csv')
conversion_B = pd.read_csv('Data/group_B_conversion.csv')

revenue_A = pd.read_csv('Data/group_A_revenue.csv')
revenue_B = pd.read_csv('Data/group_B_revenue.csv')
```

```
In [16]: # Objetivo: tener diferencia de 20 dolares en el LTV
objective_difference = 20

# Promedio y desviación estandar por grupo
mean_a = np.mean(revenue_A['ARPU'])
mean_b = np.mean(revenue_B['ARPU'])

std_a = np.std(revenue_A['ARPU'], ddof=1)
std_b = np.std(revenue_B['ARPU'], ddof=1)

n_a = len(revenue_A['ARPU'])
n_b = len(revenue_B['ARPU'])

# Realizar un t-test de una cola (mayor o igual que el objetivo)
# Null hypothesis H0: (mean_b - mean_a) <= objective_difference
# Alternative hypothesis H1: (mean_b - mean_a) > objective_difference

# Calculate the t-statistic and p-value
pooled_std = np.sqrt(((std_a**2 / n_a) + (std_b**2 / n_b)))
t_stat = (mean_b - mean_a - objective_difference) / pooled_std
p_value = 1 - stats.t.cdf(t_stat, df=min(n_a, n_b) - 1)

# Output the results
print(f"Mean LTV of Group A: {mean_a:.2f}")
print(f"Mean LTV of Group B: {mean_b:.2f}")
print(f"Objective Difference: {objective_difference}")
print(f"T-statistic: {t_stat:.4f}")
print(f"P-value: {p_value:.4f}")

if p_value <= 0.05:
    print('Se rechaza la hipótesis nula (H0)')
```



```
else:
    print('No se rechaza la hipótesis nula (H0)')
```

Mean LTV of Group A: 100.06
 Mean LTV of Group B: 120.75
 Objective Difference: 20
 T-statistic: 0.4893
 P-value: 0.3126
 No se rechaza la hipótesis nula (H0)

Hipótesis:

$H_0: \Delta \text{ARPU} \leq 20\$$

$H_i: \Delta \text{ARPU} > 20\$$

El resultado de la prueba es que el ARPU no aumenta 20 USD al subir el precio de la consulta.

Calcular tamaño de muestra:

$$n = \frac{2 \times (Z_{\alpha/2} + Z_{\beta})^2 \times \sigma^2}{\Delta^2}$$

```
In [20]: var = np.var(datos['ARPU'])
n = (2*(1.96+0.84)**2*var)/(20**2)
print(f'Tamaño de muestra para la prueba: {n}')
```

Tamaño de muestra para la prueba: 8.44872200943199

```
In [21]: # Objetivo: tener diferencia de 20 dolares en el LTV
objective_difference = 18

# Promedio y desviación estandar por grupo
mean_a = np.mean(revenue_A['ARPU'])
mean_b = np.mean(revenue_B['ARPU'])

std_a = np.std(revenue_A['ARPU'], ddof=1)
std_b = np.std(revenue_B['ARPU'], ddof=1)

n_a = len(revenue_A['ARPU'])
n_b = len(revenue_B['ARPU'])

# Realizar un t-test de una cola (mayor o igual que el objetivo)
# Null hypothesis H0: (mean_b - mean_a) <= objective_difference
# Alternative hypothesis H1: (mean_b - mean_a) > objective_difference

# Calculate the t-statistic and p-value
pooled_std = np.sqrt(((std_a**2 / n_a) + (std_b**2 / n_b)))
t_stat = (mean_b - mean_a - objective_difference) / pooled_std
p_value = 1 - stats.t.cdf(t_stat, df=min(n_a, n_b) - 1)

# Output the results
print(f"Mean LTV of Group A: {mean_a:.2f}")
print(f"Mean LTV of Group B: {mean_b:.2f}")
print(f"Objective Difference: {objective_difference}")
```

```
print(f"T-statistic: {t_stat:.4f}")
print(f"P-value: {p_value:.4f}")

if p_value <= 0.05:
    print('Se rechaza la hipótesis nula (H0)')
else:
    print('No se rechaza la hipótesis nula (H0)')
```

Mean LTV of Group A: 100.06

Mean LTV of Group B: 120.75

Objective Difference: 18

T-statistic: 1.9010

P-value: 0.0294

Se rechaza la hipótesis nula (H0)

Hipótesis:

$H_0: \Delta \text{ARPU} \leq 18\$$

$H_i: \Delta \text{ARPU} > 18\$$

```
In [23]: conversions_a = conversion_A['Converted'].sum()
total_a = len(conversion_A)
conversions_b = conversion_B['Converted'].sum()
total_b = len(conversion_B)

# Calcular conversion rate
rate_a = conversions_a / total_a
rate_b = conversions_b / total_b

# Hacer un z-test de proporciones
count = np.array([conversions_a, conversions_b])
nobs = np.array([total_a, total_b])
z_stat, p_value = sm.stats.proportions_ztest(count, nobs)

print(f"Z-statistic: {z_stat}, P-value: {p_value}")

if p_value <= 0.05:
    print('Se rechaza la hipótesis nula (H0)')
else:
    print('No se rechaza la hipótesis nula (H0)')
```

Z-statistic: 0.33634590917760526, P-value: 0.7366100338880623

No se rechaza la hipótesis nula (H0)

$H_0: \Delta \text{Conversión} = 0\$$

$H_i: \Delta \text{Conversión} \neq 0\$$

La conversión es igual entre ambos grupos.

Calcular tamaño de muestra:

$$n = \frac{2 \times (Z_{\alpha/2} + Z_{\beta})^2 \times p(1-p)}{(p_1 - p_2)^2}$$

```
In [27]: n_p = (2*(1.96+0.84)**2*rate_a*(1-rate_a))/(rate_a-rate_b)**2  
print(f'Tamaño de muestra para la prueba: {n_p}')
```

Tamaño de muestra para la prueba: 71694.186666666684

Conclusión

Al incrementar el precio de la consulta de 1 USD a 1.25 USD, el ARPU no logra el incremento 20 USD, pero la tasa de conversión entre ambos grupos se mantiene constante a pesar del aumento de precio. Además, al ajustar la prueba estadística para el ARPU, observamos un incremento de al menos 18 USD. Dado que la conversión no cambia entre los grupos y el ARPU aumenta, subir el precio de la consulta resulta favorable para la empresa, ya que genera mayores ganancias.

Created with Jupyter by Luis Márquez, Ana Sofía Hinojosa, and Ivanna Herrera