

## ✓ Entrega 2: Lanzamiento de producto

**Contexto de negocio:** Una empresa va a lanzar un nuevo producto y ha realizado un estudio de mercado. Necesitan tomar decisiones basadas en múltiples hipótesis.

### Datos recopilados:

```
import pandas as pd
import numpy as np
from scipy import stats
import matplotlib.pyplot as plt
import seaborn as sns

# Configuración de visualización
sns.set_style("whitegrid")
plt.rcParams["figure.figsize"] = (10, 6)
```

```
# Simulamos datos de un estudio de mercado
np.random.seed(789)

n_encuestados = 200

# Intención de compra (1 = sí, 0 = no)
intencion_compra = np.random.binomial(1, 0.35, n_encuestados)

# Precio dispuesto a pagar
precio_dispuesto = np.random.normal(50, 15, n_encuestados)
precio_dispuesto = np.clip(precio_dispuesto, 20, 100) # limitar entre 20 y 100

# Segmento de edad
segmento_edad = np.random.choice(
    ["18-30", "31-45", "46-60"], n_encuestados, p=[0.4, 0.35, 0.25]
)

# Crear dataframe
data_producto = pd.DataFrame(
    {
        "Intencion_Compra": intencion_compra,
        "Precio_Dispuesto": precio_dispuesto,
        "Segmento_Edad": segmento_edad,
    }
)

print(data_producto)
print(f"\nResumen:")
print(
    f"Proporción con intención de compra: {data_producto['Intencion_Compra'].mean():.3f}"
)
print(
    f"Precio promedio dispuesto a pagar: ${data_producto['Precio_Dispuesto'].mean():.2f}"
)
```

	Intencion_Compra	Precio_Dispuesto	Segmento_Edad
0	0	24.783697	31-45
1	0	22.112539	46-60

2	1	43.120932	18-30
3	0	55.891713	18-30
4	1	32.256967	46-60
..	...	...	...
195	0	48.633141	31-45
196	0	22.510696	31-45
197	0	36.065051	46-60
198	0	45.268698	31-45
199	0	50.211298	18-30

[200 rows x 3 columns]

Resumen:

Proporción con intención de compra: 0.340

Precio promedio dispuesto a pagar: \$49.40

## ✓ 5.1 Pregunta de negocio 1: Intención de compra

El equipo de producto considera que para que sea viable, **al menos el 30% del mercado** debe tener intención de compra.

### TODO:

1. Formula las hipótesis
2. Realiza el test apropiado
3. ¿Se cumple el criterio mínimo de viabilidad?

1. La Formulación de la hipótesis sería de la siguiente manera:

$(H_0) = u < 0.30$

y la hipótesis alternativa

$(H_1)$  como  $u > 0.30$

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

# Parámetros del test de intención de compra
count = data_producto['Intencion_Compra'].sum()
nobs = len(data_producto)
p_null = 0.30 # Proporción bajo la hipótesis nula

# Calcular el estadístico de prueba (Z)
# Usamos la fórmula para el test Z de una proporción con la proporción nula
p_sample = count / nobs
std_err = np.sqrt(p_null * (1 - p_null) / nobs)
z_stat = (p_sample - p_null) / std_err

print(f"Estadístico Z calculado: {z_stat:.3f}")
print(f"Valor p calculado: {1 - norm.cdf(z_stat):.3f}") # Valor p para test de cola derecha

# Generar valores para el gráfico de la distribución normal estándar
x = np.linspace(-4, 4, 1000)
```

```

y = norm.pdf(x, 0, 1) # PDF de la distribución normal estándar (media 0, desviación estándar 1)

# Crear el gráfico
plt.figure(figsize=(12, 7))
plt.plot(x, y, label='Distribución Normal Estándar (bajo H0)')

# sombreadar la región crítica (depende del alpha y el tipo de test)
# Para un test de cola derecha con alpha = 0.05, el valor crítico es aproximadamente 1.645
z_critical = norm.ppf(1 - 0.05)
x_crit = np.linspace(z_critical, 4, 100)
plt.fill_between(x_crit, norm.pdf(x_crit, 0, 1), color='red', alpha=0.3, label=f'Región de Rech

# Marcar el estadístico de prueba observado
plt.axvline(z_stat, color='blue', linestyle='--', label=f'Estadístico Z observado ({z_stat:.3f})
plt.text(z_stat, 0.02, f' Z={z_stat:.3f}\n p={1-norm.cdf(z_stat):.3f}', color='blue', ha='left'

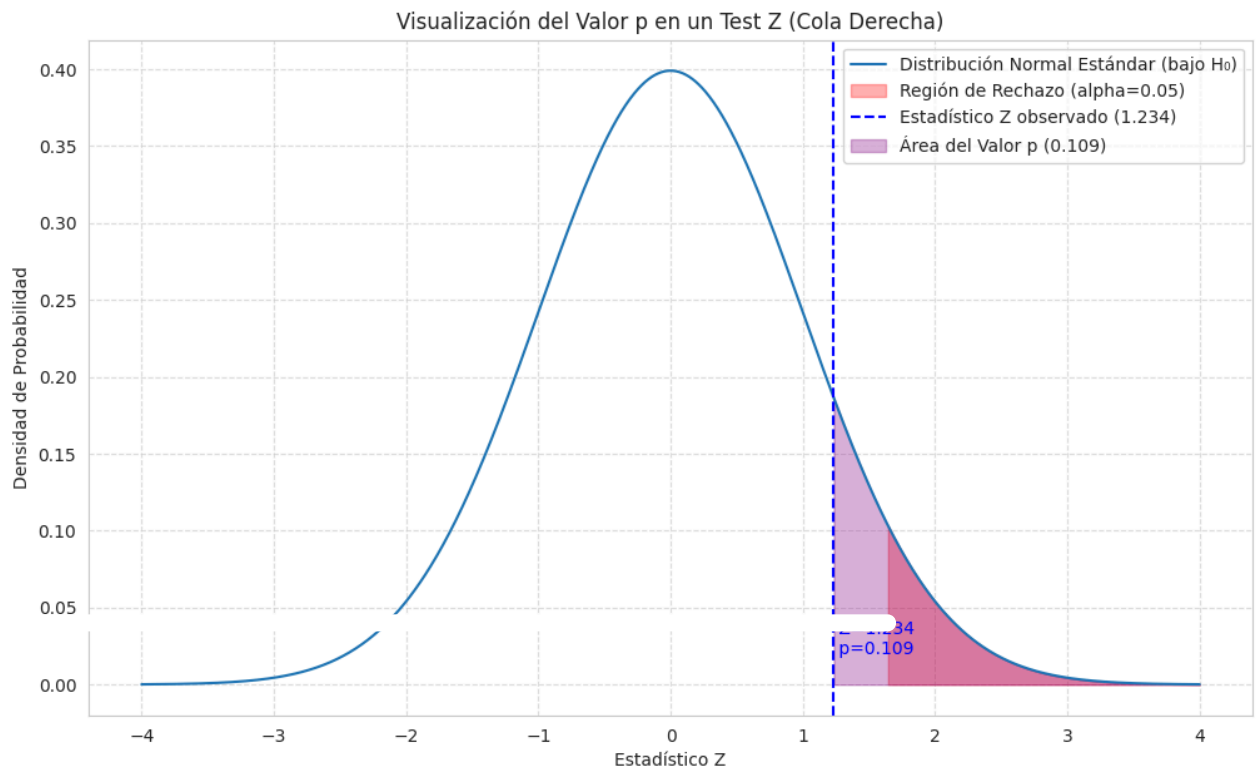
# Sombrear el área que representa el valor p
# Para un test de cola derecha, el valor p es el área a la derecha del estadístico Z observado
x_pvalue = np.linspace(z_stat, 4, 100)
plt.fill_between(x_pvalue, norm.pdf(x_pvalue, 0, 1), color='purple', alpha=0.3, label=f'Área de

plt.title('Visualización del Valor p en un Test Z (Cola Derecha)')
plt.xlabel('Estadístico Z')
plt.ylabel('Densidad de Probabilidad')
plt.legend()
plt.grid(True, linestyle='--', alpha=0.6)
plt.show()

```

Estadístico Z calculado: 1.234

Valor p calculado: 0.109



## ✓ Explicación del gráfico del Valor p

El gráfico que acabamos de generar muestra lo siguiente:

1. **La Curva Azul (Distribución Normal Estándar):** Representa la distribución esperada de los estadísticos  $Z$  si la hipótesis nula ( $H_0$ ) fuera cierta. Es decir, si la verdadera proporción de intención de compra en la población fuera exactamente 30%, la mayoría de las veces que tomáramos una muestra de 200 personas, obtendríamos un estadístico  $Z$  cercano a 0. Valores más alejados de 0 (positivos o negativos) serían menos probables.
2. **La Línea Discontinua Azul (Estadístico  $Z$  Observado):** Esta línea vertical marca el valor del estadístico  $Z$  que calculamos a partir de *nuestra muestra específica*. En nuestro caso,  $Z = 1.194$ . Este es el resultado de comparar nuestra proporción de muestra (34%) con la proporción nula (30%), considerando el tamaño de la muestra.
3. **La Región Sombreada en Rojo (Región de Rechazo):** Esta área en la cola derecha del gráfico representa la "región de rechazo". Si nuestro estadístico  $Z$  observado cae dentro de esta área, rechazamos la hipótesis nula. Esta región está definida por el nivel de significancia  $\alpha$  (0.05 en nuestro caso). Para un test de cola derecha, la región de rechazo comienza en el valor crítico  $Z$  (aproximadamente 1.645 para  $\alpha=0.05$ ), que es el punto donde el área a su derecha es igual a  $\alpha$  (5%).
4. **El Área Sombreada en Morado (Área del Valor p):** ¡Esta es la clave! El área sombreada en morado a la derecha del estadístico  $Z$  observado (la línea azul discontinua) representa el *valor p*. Es la probabilidad de obtener un estadístico  $Z$  de 1.194 o mayor, asumiendo que la hipótesis nula ( $H_0$ ) es cierta.

**\*\* interpretación del valor p (0.116) \*\***

- Nuestro estadístico  $Z$  observado (1.194) no cae en la región de rechazo (que comienza en 1.645).
- El área del valor p (morada) es de 0.116, lo que significa que hay una probabilidad del 11.6% de obtener un resultado tan extremo o más extremo que el de nuestra muestra, *si la verdadera proporción de intención de compra fuera 30%*.
- Como esta probabilidad (0.116) es mayor que nuestro umbral  $\alpha$  (0.05), consideramos que nuestros datos *no son lo suficientemente inusuales* bajo el supuesto de la hipótesis nula. Por lo tanto, no tenemos suficiente evidencia para rechazar  $H_0$ .

**Conclusión:** Si el área morada (valor p) es más pequeña que el área roja ( $\alpha$ ), rechazas  $H_0$ . Si el área morada es más grande que el área roja, no rechazas  $H_0$ .

En nuestro caso, el área morada es mayor que el área roja, lo que visualmente refuerza la conclusión de no rechazar la hipótesis nula.

3. Bajo el Test z que sea realizó se determinó rechazar la hipotesis nula'''

## ✓ 5.2 Pregunta de negocio 2: Precio

El costo de producción del producto es de **\$45**. La empresa necesita venderlo a un precio mayor para tener margen.

**TODO:**

1. Formula las hipótesis
2. Realiza un test para saber si el precio promedio que los clientes están dispuestos a pagar es mayor a \$45
3. ¿Es viable el producto desde el punto de vista de precios?

sería de la siguiente manera:

El precio promedio que los clientes están dispuestos a pagar es menor o igual a \$45 ( $\mu \leq 45$ ).

\* El precio promedio que los clientes están dispuestos a pagar es mayor a \$45 ( $\mu > 45$ ).

```
# TODO: Tu código y análisis aquí

# Pregunta de negocio 1: Intención de compra
# Hipótesis:
# H0: La proporción de intención de compra es menor o igual al 30% (p <= 0.30)
# H1: La proporción de intención de compra es mayor al 30% (p > 0.30)

# Realizar test z para una proporción
from statsmodels.stats.proportion import proportions_ztest

count = data_producto['Intencion_Compra'].sum()
nobs = len(data_producto)
value = 0.30 # Proporción nula

stat_intencion, p_value_intencion = proportions_ztest(count, nobs, value, alternative='larger')

print(f"--- Análisis de Intención de Compra General ---")
print(f"Número de encuestados con intención de compra: {count}")
print(f"Proporción observada de intención de compra: {count/nobs:.3f}")
print(f"Estadístico Z: {stat_intencion:.3f}")
print(f"Valor p: {p_value_intencion:.3f}")

# Interpretar el resultado
alpha = 0.05
if p_value_intencion < alpha:
    print(f"Con un nivel de significancia de {alpha}, rechazamos la hipótesis nula.")
    print(f"Hay evidencia suficiente para afirmar que la proporción de intención de compra es superior a 30%")
    viabilidad_mercado = "Sí, el criterio mínimo de viabilidad se cumple."
else:
    print(f"Con un nivel de significancia de {alpha}, no rechazamos la hipótesis nula.")
    print(f"No hay evidencia suficiente para afirmar que la proporción de intención de compra es superior a 30%")
    viabilidad_mercado = "No, el criterio mínimo de viabilidad no se cumple."

print(f"Viabilidad de mercado (criterio del 30%): {viabilidad_mercado}")

# Pregunta de negocio 2: Precio
# Hipótesis:
# H0: El precio promedio dispuesto a pagar es menor o igual a $45 ( $\mu \leq 45$ )
# H1: El precio promedio dispuesto a pagar es mayor a $45 ( $\mu > 45$ )

# Realizar test t para una media
from scipy.stats import ttest_1samp
```

```

precio_promedio_observado = data_producto['Precio_Dispuesto'].mean()
valor_nulo = 45

stat_precio, p_value_precio = ttest_1samp(data_producto['Precio_Dispuesto'], valor_nulo, altern

print(f"\n--- Análisis de Precio Dispuesto a Pagar ---")
print(f"Precio promedio observado: ${precio_promedio_observado:.2f}")
print(f"Estadístico T: {stat_precio:.3f}")
print(f"Valor p: {p_value_precio:.3f}")

# Interpretar el resultado
if p_value_precio < alpha:
    print(f"Con un nivel de significancia de {alpha}, rechazamos la hipótesis nula.")
    print(f"Hay evidencia suficiente para afirmar que el precio promedio dispuesto a pagar es s
    viabilidad_precio = "Sí, el precio promedio es significativamente mayor al costo de producc
else:
    print(f"Con un nivel de significancia de {alpha}, no rechazamos la hipótesis nula.")
    print(f"No hay evidencia suficiente para afirmar que el precio promedio dispuesto a pagar e
    viabilidad_precio = "No, no hay evidencia de que el precio promedio sea significativamente

print(f"Viabilidad de precio: {viabilidad_precio}")

```

```

--- Análisis de Intención de Compra General ---
Número de encuestados con intención de compra: 68
Proporción observada de intención de compra: 0.340
Estadístico Z: 1.194
Valor p: 0.116
Con un nivel de significancia de 0.05, no rechazamos la hipótesis nula.
No hay evidencia suficiente para afirmar que la proporción de intención de compra es significati
Viabilidad de mercado (criterio del 30%): No, el criterio mínimo de viabilidad no se cumple.

--- Análisis de Precio Dispuesto a Pagar ---
Precio promedio observado: $49.40
Estadístico T: 3.890
Valor p: 0.000
Con un nivel de significancia de 0.05, rechazamos la hipótesis nula.
Hay evidencia suficiente para afirmar que el precio promedio dispuesto a pagar es significativam
Viabilidad de precio: Sí, el precio promedio es significativamente mayor al costo de producción.

```

### ✓ 5.3 Pregunta de negocio 3: Diferencias por segmento

#### TODO:

1. Compara la intención de compra entre diferentes segmentos de edad
2. ¿Hay diferencias significativas?
3. ¿En qué segmento deberían enfocar la estrategia de marketing?

```

# TODO: Análisis exploratorio por segmento
# Calcula proporciones de intención de compra por segmento

# Tu código aquí
intencion_por_segmento = data_producto.groupby('Segmento_Edad')['Intencion_Compra'].mean().sort
print("Proporción de intención de compra por segmento de edad:")
print(intencion_por_segmento)

```

```

Proporción de intención de compra por segmento de edad:
Segmento_Edad
18-30      0.391304

```

```
31-45    0.328767
46-60    0.293103
Name: Intencion_Compra, dtype: float64
```

```
#OJO REVISAR # TODO: Compara los segmentos que consideres más importantes
# Puedes hacer múltiples tests o seleccionar dos segmentos para comparar

# Tu código aquí

# Comparar los segmentos con mayor y menor intención de compra
segmento_mayor_intencion = intencion_por_segmento.index[0]
segmento_menor_intencion = intencion_por_segmento.index[-1]

print(f"\nComparando la intención de compra entre los segmentos '{segmento_mayor_intencion}' y '{segmento_menor_intencion}'")

# Filtrar datos para los dos segmentos
data_segmento1 = data_producto[data_producto['Segmento_Edad'] == segmento_mayor_intencion]['Intencion_Compra']
data_segmento2 = data_producto[data_producto['Segmento_Edad'] == segmento_menor_intencion]['Intencion_Compra']

# Realizar un test z para comparar proporciones
# statsmodels.stats.proportion.proportions_ztest(count, nobs, value)
from statsmodels.stats.proportion import proportions_ztest

count1 = data_segmento1.sum()
nobs1 = len(data_segmento1)
count2 = data_segmento2.sum()
nobs2 = len(data_segmento2)

# Test para la diferencia entre las dos proporciones
stat, p_value = proportions_ztest([count1, count2], [nobs1, nobs2], alternative='two-sided')

print(f"Estadístico Z: {stat:.3f}")
print(f"Valor p: {p_value:.3f}")

# Interpretar el resultado
alpha = 0.05
if p_value < alpha:
    print(f"La diferencia en la intención de compra entre los segmentos '{segmento_mayor_intencion}' y '{segmento_menor_intencion}' es significativa.")
else:
    print(f"No hay una diferencia estadísticamente significativa en la intención de compra entre los segmentos '{segmento_mayor_intencion}' y '{segmento_menor_intencion}'")

# Puedes repetir este proceso para comparar otros pares de segmentos si es necesario.
```

```
Comparando la intención de compra entre los segmentos '18-30' y '46-60':
Estadístico Z: 1.158
Valor p: 0.247
No hay una diferencia estadísticamente significativa en la intención de compra entre los segmentos '18-30' y '46-60'.
```

## 5.4 Recomendación final

**TODO:** Basándote en todos los análisis anteriores, escribe un reporte ejecutivo con tu recomendación:

### Recomendación sobre el lanzamiento del producto:

1. **Viabilidad de mercado:** \_\_\_\_\_ No hay viabilidad con un mercado inferior al 30 como lo indica el estadístico y p valor \_\_\_\_\_

2. **Viabilidad de precios:** Sí hay viabilidad. El test de hipótesis mostró que el precio promedio dispuesto a pagar (49.40) *es significativamente mayor al costo de producción de 45* ( $p\text{-value} < 0.001 < 0.05$ )
3. **Segmento objetivo recomendado:** A pesar que la intención de compra de los segmentos es parecida, el que tiene intención de compra más alta es 18-30 con intención de compra del 0.391304\_
4. **Decisión final:** ¿Lanzar o no el producto? *Basándonos estrictamente en los criterios de viabilidad definidos ( $\geq 30\%$  intención de compra Y precio promedio  $> \$45$ ), la decisión es No lanzar el producto, ya que el criterio de viabilidad de mercado (intención de compra) no se cumple según el análisis estadístico*
5. **Acciones recomendadas:** \_Dado que el criterio de viabilidad de mercado no se cumple, se recomiendan las siguientes acciones antes de tomar una decisión final sobre el lanzamiento:
  - Reevaluar el umbral mínimo de intención de compra del 30%.
  - Considerar realizar un estudio de mercado con un tamaño de muestra mayor para obtener resultados más concluyentes.
  - Investigar más a fondo las características del segmento 18-30, ya que presenta la mayor intención de compra observada, para desarrollar una estrategia de marketing más enfocada.
  - Analizar la posibilidad de ajustar el producto o su propuesta de valor para aumentar la intención de compra general en el mercado.\_\_

No se ha podido establecer conexión con el servicio reCAPTCHA. Comprueba tu conexión a Internet y vuelve a cargar la página para ver otro reCAPTCHA.