Retos Integradores de Estadística: Distribuciones Muestrales

Contents

1 Introducción

Los siguientes retos integradores están diseñados para aplicar los conceptos de distribución muestral, Ley de los Grandes Números, Teorema del Límite Central, cálculo de probabilidades con la distribución muestral, y distribución muestral de proporciones. Cada reto plantea un escenario práctico que requiere simulación, cálculo de probabilidades, y visualización en Python, fomentando el trabajo en equipo.

2 Reto 1: Análisis de Consumo de Energía en Hogares

Contexto: Una compañía eléctrica estudia el consumo mensual de energía (en kWh) de los hogares en una ciudad. La población tiene un consumo promedio de 300 kWh con una desviación estándar de 50 kWh, y se sabe que el 40% de los hogares consumen más de 320 kWh. La distribución poblacional no es normal (por ejemplo, sesgada a la derecha).

Tareas:

1. **Ley de los Grandes Números**: Simular muestras de tamaños $n = 5, 20, 100$ (1000 muestras por tamaño) y mostrar cómo la media muestral del consumo converge al valor poblacional. Usar un gráfico para ilustrar la convergencia.
2. **Teorema del Límite Central**: Para $n = 40$, simular la distribución muestral de la media (1000 muestras) y comparar con una distribución normal teórica. £Es aproximadamente normal?
3. **Cálculo de Probabilidades**: Calcular la probabilidad de que la media muestral de una muestra de $n = 40$ esté por debajo de 290 kWh. Visualizar la región correspondiente.
4. **Distribución Muestral de Proporciones**: Simular la distribución muestral de la proporción de hogares con consumo mayor a 320 kWh para $n = 50$ (1000 muestras). Calcular la probabilidad de que la proporción muestral sea menor a 0.35. Visualizar la distribución.
5. **Análisis en Equipo**: Discutir cómo el tamaño de la muestra afecta los resultados y la validez de las aproximaciones normales.

Entregable: Un script en Python que genere los cálculos y gráficos, y un informe breve explicando los resultados.

3 Reto 2: Calidad de Producción de Tornillos

Contexto: Una fábrica produce tornillos con una longitud promedio de 5 cm y una desviación estándar de 0.8 cm. La distribución de las longitudes es uniforme. Además, el 25% de los tornillos tienen una longitud mayor a 5.5 cm.

Tareas:

1. **Ley de los Grandes Números**: Simular muestras de tamaños $n = 15, 60, 300$ (1000 muestras por tamaño) y mostrar cómo la media muestral converge a 5 cm. Usar un diagrama de caja para visualizar.
2. **Teorema del Límite Central**: Para $n = 60$, simular la distribución muestral de la media (1000 muestras) y verificar si se asemeja a una distribución normal. Comparar con la curva teórica.
3. **Cálculo de Probabilidades**: Calcular la probabilidad de que la media muestral de una muestra de $n = 60$ esté entre 4.9 y 5.1 cm. Visualizar la región de probabilidad.
4. **Distribución Muestral de Proporciones**: Simular la distribución muestral de la proporción de tornillos con longitud mayor a 5.5 cm para $n = 80$ (1000 muestras). Calcular la probabilidad de que la proporción muestral sea mayor a 0.3. Visualizar la distribución.
5. **Análisis en Equipo**: Analizar cómo la distribución uniforme afecta la convergencia y la normalidad de la distribución muestral.

Entregable: Un script en Python con cálculos y gráficos, y un resumen de las conclusiones del equipo.

4 Reto 3: Tiempos de Entrega de Paquetes

Contexto: Una empresa de mensajería registra tiempos de entrega con un promedio de 48 horas y una desviación estándar de 12 horas. La distribución es sesgada (por ejemplo, exponencial). El 65% de los paquetes se entregan en menos de 50 horas.

Tareas:

1. **Ley de los Grandes Números**: Simular muestras de tamaños $n = 10, 50, 500$ (1000 muestras por tamaño) y mostrar la convergencia de la media muestral a 48 horas. Usar un gráfico apropiado.
2. **Teorema del Límite Central**: Para $n = 50$, simular la distribución muestral de la media (1000 muestras) y compararla con la distribución normal teórica.
3. **Cálculo de Probabilidades**: Calcular la probabilidad de que la media muestral de

una muestra de $n = 50$ sea mayor a 52 horas. Visualizar la región correspondiente.

4. **Distribución Muestral de Proporciones**: Simular la distribución muestral de la proporción de paquetes entregados en menos de 50 horas para $n = 100$ (1000 muestras). Calcular la probabilidad de que la proporción muestral esté entre 0.6 y 0.7. Visualizar la distribución.

5. **Análisis en Equipo**: Discutir cómo el tamaño de la muestra influye en la precisión de las estimaciones y la normalidad de las distribuciones.

Entregable: Un script en Python con cálculos y visualizaciones, y un análisis grupal de los resultados.

5 Reto 4: Satisfacción en un Restaurante

Contexto: Un restaurante sabe que el 80% de sus clientes están satisfechos con el servicio, y el tiempo promedio de espera por mesa es de 15 minutos con una desviación estándar de 4 minutos. La distribución del tiempo de espera es sesgada.

Tareas:

1. **Ley de los Grandes Números**: Simular muestras de tamaños $n = 20, 100, 400$ (1000 muestras por tamaño) para los tiempos de espera y mostrar la convergencia a 15 minutos. Usar un gráfico para ilustrar.

2. **Teorema del Límite Central**: Para $n = 30$, simular la distribución muestral de la media de los tiempos de espera (1000 muestras) y verificar su normalidad. Comparar con la distribución normal teórica.

3. **Cálculo de Probabilidades**: Calcular la probabilidad de que la media muestral de una muestra de $n = 30$ esté por debajo de 14 minutos. Visualizar la región de probabilidad.

4. **Distribución Muestral de Proporciones**: Simular la distribución muestral de la proporción de clientes satisfechos para $n = 60$ (1000 muestras). Calcular la probabilidad de que la proporción muestral sea menor a 0.75. Visualizar la distribución.

5. **Análisis en Equipo**: Analizar cómo los resultados reflejan los conceptos estadísticos y discutir la importancia de muestras grandes.

Entregable: Un script en Python con cálculos y gráficos, y un informe breve con conclusiones del equipo.

6 Reto 5: Producción de Jugos en Botellas

Contexto: Una fábrica de jugos llena botellas con un volumen promedio de 500 ml y una desviación estándar de 20 ml. La distribución es no normal. El 90% de las botellas tienen un volumen mayor a 480 ml.

Tareas:

1. **Ley de los Grandes Números**: Simular muestras de tamaños $n = 8, 40, 200$ (1000 muestras por tamaño) y mostrar cómo la media muestral converge a 500 ml. Usar un gráfico para visualizar.
2. **Teorema del Límite Central**: Para $n = 40$, simular la distribución muestral de la media (1000 muestras) y compararla con la distribución normal teórica.
3. **Cálculo de Probabilidades**: Calcular la probabilidad de que la media muestral de una muestra de $n = 40$ esté entre 495 y 505 ml. Visualizar la región correspondiente.
4. **Distribución Muestral de Proporciones**: Simular la distribución muestral de la proporción de botellas con volumen mayor a 480 ml para $n = 100$ (1000 muestras). Calcular la probabilidad de que la proporción muestral sea mayor a 0.92. Visualizar la distribución.
5. **Análisis en Equipo**: Discutir cómo la distribución no normal de la población afecta los resultados y la validez del Teorema del Límite Central.

Entregable: Un script en Python con cálculos y visualizaciones, y un análisis grupal de los resultados.

7 Reto 6: Rendimiento Académico de Estudiantes

Contexto: Una universidad analiza el promedio de calificaciones de sus estudiantes, con una media de 75 puntos y una desviación estándar de 10 puntos (distribución no especificada). El 55% de los estudiantes obtienen más de 80 puntos.

Tareas:

1. **Ley de los Grandes Números**: Simular muestras de tamaños $n = 12, 60, 300$ (1000 muestras por tamaño) y mostrar la convergencia de la media muestral a 75 puntos. Usar un gráfico apropiado.
2. **Teorema del Límite Central**: Para $n = 60$, simular la distribución muestral de la media (1000 muestras) y verificar si es aproximadamente normal. Comparar con la curva teórica.
3. **Cálculo de Probabilidades**: Calcular la probabilidad de que la media muestral de

una muestra de $n = 60$ sea mayor a 78 puntos. Visualizar la región de probabilidad.

4. **Distribución Muestral de Proporciones**: Simular la distribución muestral de la proporción de estudiantes con más de 80 puntos para $n = 80$ (1000 muestras). Calcular la probabilidad de que la proporción muestral esté entre 0.5 y 0.6. Visualizar la distribución.

5. **Análisis en Equipo**: Analizar cómo los resultados ilustran los conceptos estadísticos y discutir las limitaciones de las aproximaciones normales.

Entregable: Un script en Python con cálculos y gráficas, y un resumen de las conclusiones del equipo.

8 Instrucciones Generales

- **Implementación**: Utilizar Python con bibliotecas como `numpy`, `scipy.stats.norm`, y `matplotlib` o `seaborn`. Instalar con:

```
pip install numpy scipy matplotlib seaborn
```

- **Distribuciones**: Asumir distribuciones no normales (por ejemplo, exponencial, uniforme) para enfatizar el Teorema del Límite Central. Los equipos pueden elegir la distribución específica.
- **Condiciones**: Verificar que $n \cdot p \geq 5$ y $n \cdot (1 - p) \geq 5$ para la distribución de proporciones, y que $n \geq 30$ para el Teorema del Límite Central.
- **Trabajo en Equipo**: Dividir tareas (simulación, cálculos, visualización, análisis) entre los miembros del equipo.
- **Entregables**: Un script en Python por reto y un informe breve que explique los resultados, destacando la aplicación de los conceptos estadísticos.