Probabilidad

Dr. Julio César Galindo López

Ley de los grandes números y el teorema central del límite

1. La ley de los grandes números

Definición 1. (**Convergencia en Probabilidad**) Decimos que $X_1, \ldots, X_n, \ldots$ una sucesión de variables aleatorias converge, *en probabilidad*, a $\mathbb{E}[X_i] = \mu$ si,

$$\lim_{n\to\infty} \mathbb{P}\left(|X_n - \mu| > \varepsilon\right) = 0$$

para todo $\varepsilon > 0$

Teorema 2. (**La ley débil de los grandes números**) Sea $X_1, \ldots, X_n, \ldots$ una sucesión de variables aleatorias independientes e idénticamente distribuídas tales que $\mathbb{E}[X_i] = \mu < +\infty$. Entonces,

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to\infty]{} \mu$$

en probabilidad.

Usaremos la desigualdad de Chebyshev para probar el teorema 2 **bajo la hipótesis (innecesaria)** de que $\mathrm{Var}(X_i) = \sigma^2 < \infty$.

1.1. **Desigualdad de Chebyshev.** Primero comencemos con la desigualdad de Markov.

Proposición 3. (**Desigualdad de Markov**) Sea X una variable aleatoria positiva con $\mathbb{E}[X] < \infty$, entonces

$$\mathbb{P}(X > a) \leq \frac{\mathbb{E}[X]}{a},$$

para todo $a > 0$.

Demostración. Consideremos a la variable aleatoria $\mathbf{1}_{\{X>a\}}$. Sabemos que $\mathbb{E}[\mathbf{1}_{\{X>a\}}] = \mathbb{P}(X > a)$. Ahora bien,

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X\mathbf{1}_{\{X>a\}}] + \mathbb{E}[X\mathbf{1}_{\{X\leq a\}}] \\
&\geq a\mathbb{E}[\mathbf{1}_{\{X>a\}}] + \mathbb{E}[X\mathbf{1}_{\{X\leq a\}}] \\
&\geq a\mathbb{P}(X > a)
\end{aligned}$$

en donde la desigualdad se sigue debido a que X es positiva (y por lo tanto $\mathbb{E}[X\mathbf{1}_{\{X\leq a\}}] > 0$).

Corolario 4. (**Desigualdad de Chebyshev**) Sea X una variable aleatoria con $\mathbb{E}[X] < \infty$ y $\mathrm{Var}(X) < \infty$. Entonces,

$$\mathbb{P}\left(|X - \mathbb{E}[X]| > a\right) \leq \frac{\mathrm{Var}(X)}{a^2},$$

para toda $a > 0$.

Demostración. Se tiene la siguiente igualdad de eventos

$$\{|X - \mathbb{E}[X]| > a\} = \{(X - \mathbb{E}[X])^2 > a^2\}.$$

Así, de la desigualdad de Markov,

$$\begin{aligned}
\mathbb{P}(|X - \mathbb{E}[X]| > a) &= \mathbb{P}((X - \mathbb{E}[X])^2 > a^2) \\
&\leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2} \\
&= \frac{\mathrm{Var}(X)}{a^2}.
\end{aligned}$$

1.2. **Demostración de la ley débil de los grandes números.** Consideremos a la sucesión de variables aleatorias

$$\left\{\overline{X}_n := \frac{X_1 + \cdots + X_n}{n}\right\}_{n\geq 1}.$$

Se tiene que:

$$\mathbb{E}[\overline{X}_n] = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu$$

y además

$$\mathrm{Var}[\overline{X}_n] = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n}.$$

Por la desigualdad de Chebyshev 4,

$$\mathbb{P}\left(|\overline{X}_n - \mu| > \varepsilon\right) \leq \frac{\mathrm{Var}(\overline{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

Finalmente,

$$\lim_{n\to\infty}\mathbb{P}\left(|\overline{X}_n - \mu| > \varepsilon\right) = 0,$$

lo que prueba el resultado.

Corolario 5. (Método Monte-Carlo) Sean $U_1, \ldots, U_n$ variables aleatorias independientes e idénticamente distribuídas sobre el intervalo $(0,1)$. Sea $f : [0,1] \to \mathbb{R}$ una función continua y acotada (por lo tanto integrable). Entonces, las variables aleatorias $f(U)_1, \ldots, f(U)_n$ son independientes e idénticamente distribuídas y

$$\frac{f(U)_1 + \cdots + f(U)_n}{n} \xrightarrow[n\to\infty]{} \int_0^1 f(t)dt.$$

1.3. **Ejemplos.**

Ejemplo 6. Suponga que tomamos $X_1, \ldots, X_n, \ldots$ variables aleatorias independientes uniformes sobre $(0,1)$. Entonces, $\mathbb{E}[X_i] = \frac{1}{2}$, $\mathrm{Var}(X_i) = \frac{1}{12}$. Entonces,

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to\infty]{} \frac{1}{2}$$

en probabilidad.

Ejemplo 7. Consideremos n lanzamientos independientes de un dado. Sea X_i el resultado del i-ésimo lazamiento con $\mathbb{E}[X_i] = 7/2$. Entonces, $S_n = X_1 + \cdots + X_n$ es la suma de los primeros n lanzamientos. La ley de los grandes números nos dice que

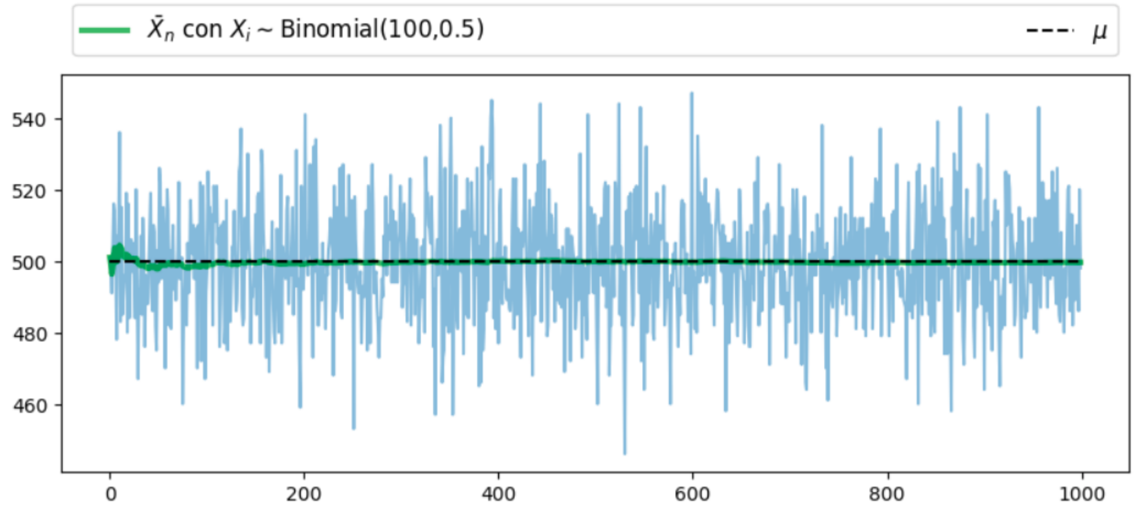
$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to\infty]{} \frac{7}{2}$$

Prueba el siguiente código:

```python
#Librerías
import matplotlib.pyplot as plt
plt.rcParams["figure.figsize"] = (11, 5)
import random
import numpy as np
from scipy.stats import t, beta, lognorm, expon, gamma, uniform
from scipy.stats import gaussian_kde, poisson, binom, norm, chi2
from mpl_toolkits.mplot3d import Axes3D
from matplotlib.collections import PolyCollection
from scipy.linalg import inv, sqrtm

```

```
# tamaño
n = 100

# Colección de distribuciones
distributions = {"Binomial(100,0.5)": binom(1000, 0.5),
"poisson(4)": poisson(4),
"uniforme": beta(2, 2),
"beta(5, 1/2)": gamma(5, scale=2),
"exponencial con media 1": expon(1)}


num_plots = 3
fig, axes = plt.subplots(num_plots, 1, figsize=(10, 20))

bbox = (0., 1.02, 1., .102)
legend_args = {’ncol’: 2,
      ’bbox_to_anchor’: bbox,
      ’loc’: 3,
      ’mode’: ’expand’}
plt.subplots_adjust(hspace=0.5)

for ax in axes:
# Se eligen distribuciones
      dist = random.choice(list(distributions.keys()))
      distribution = distributions.pop(dist)

# Generamos las muestras
     data = distribution.rvs(n)

# Medias muestrales (promedios)
    sample_mean = np.empty(n)
   for i in range(n):
        sample_mean[i] = np.mean(data[:i+1])

        ax.plot(list(range(n)), data, ’o’, color=’grey’, alpha=0.5)
        axlabel = ’$\\bar{X}_n$ for $X_i \sim$’ + dist
        ax.plot(list(range(n)), sample_mean, ’g-’, lw=3, alpha=0.6, label=axlabel)
        m = distribution.mean()
        ax.plot(list(range(n)), [m] * n, ’k--’, lw=1.5, label=’$\mu$’)
        ax.vlines(list(range(n)), m, data, lw=0.2)
        ax.legend(**legend_args, fontsize=12)

plt.show()
```

Ejemplo 8. (El estimador de Kolmogorov-Smirnov es un estimador Monte.Carlo) Sea $X_1, \ldots, X_n, \ldots$ una sucesión de variables aleatorias independientes e idénticamente distribuídas con función de distribución acumulada F. Sea $x \in \mathbb{R}$ un real fijo y defínase $Z_i = \mathbf{1}_{\{X_i \leq x\}}$ (note que las variables aleatorias Z_i son



independientes e idénticamente distribuídas). Así, por la ley de los grandes números,

$$F_n(x) \xrightarrow[n\to\infty]{} \mathbb{E}[Z_i] = F,$$

en probabilidad. Además, si $a < b$ son números reales y definimos

$$W_n(x) := \sum_{i=1}^{n} Z_i = \operatorname{card}\left(\{i \in \{1, \dots, n\} : X_i \leq x\}\right),$$

$$\frac{W_n(b) - W_n(a)}{n} = \frac{\operatorname{card}\left(\{i \in \{1, \dots, n\} : a < X_i \leq x\}\right)}{den} \xrightarrow[n\to\infty]{} F(b) - F(a),$$

en probabilidad. Si X_1 admite densidad $f(x)$, por el teorema fundamental del cálculo,

$$\frac{W_n(b) - W_n(a)}{n} = \frac{\operatorname{card}\left(\{i \in \{1, \dots, n\} : a < X_i \leq b\}\right)}{n} \xrightarrow[n\to\infty]{} \int_a^b f(x)dx.$$

2. El Teorema Central del Límite

Teorema 9. Sea $X_1, \dots, X_n, \dots$ una sucesión de variables aleatorias independientes e idénticamente distribuídas tales que $\mathbb{E}[X_i] = \mu < +\infty$ y $\operatorname{Var}(X_i) = \sigma^2 < \infty$. Definamos a la siguiente sucesión de variables aleatorias $S_n = X_1 + \cdots + X_n$ y tomemos su estandarización:

$$Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}.$$

Entonces, para todo cualesquiera reales $a < b$,

$$\lim_{n\to\infty} \mathbb{P}\left(a \leq Z_n \leq b\right) = \Phi(b) - \Phi(a),$$

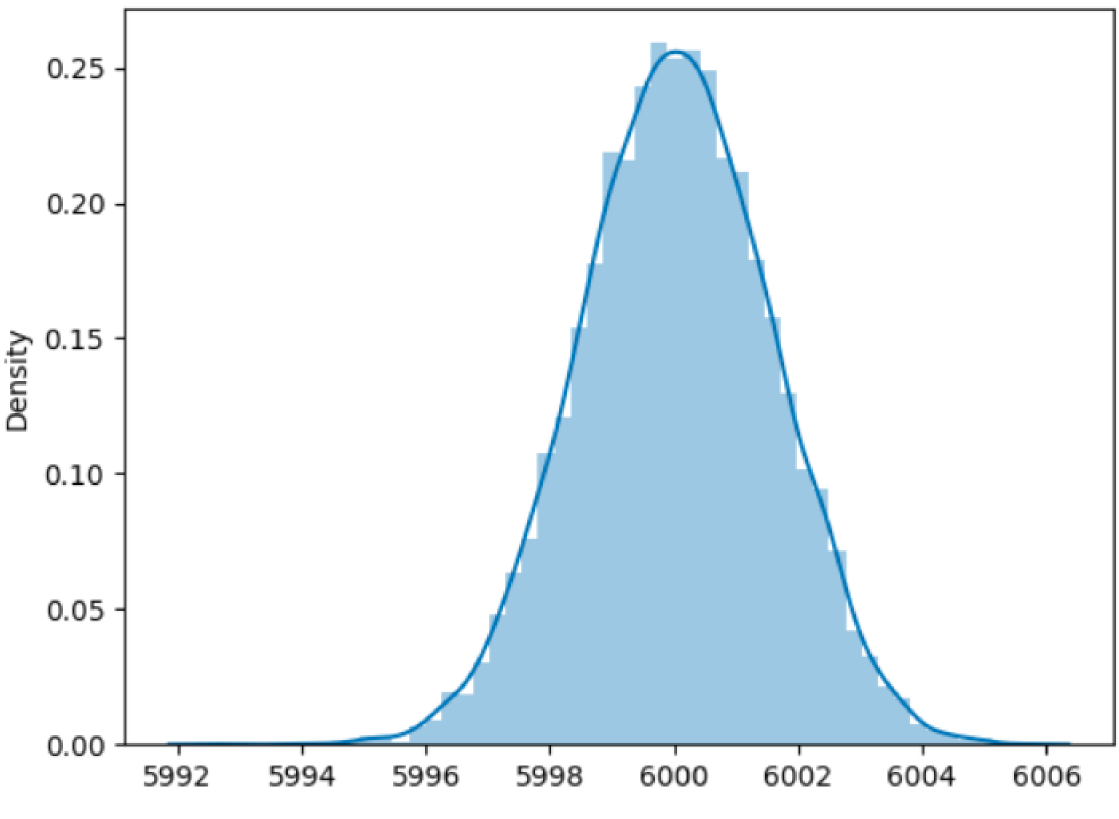
en donde

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{2\pi} e^{-x^2/2} dx.$$

El principio del TCL es el siguiente:

- Obtener muestras aleatorias.
- Necesitamos remuestreo.
- Obtener la muestra de las medias.
- Realizar el histograma.

```python
#Librerías
import numpy as np
import pandas as pd
import seaborn as sns


# Vamos a crear distintas muestras (100) y los acomodaremos en un panel
np.random.seed(1)

df = pd.DataFrame()

for i in range(1, 10001):
    muestras_binomiales = np.random.binomial(10000, 0.6, 1000)
    columnas = f"muestra {i}"
    df[columnas]=muestras_binomiales

muestra_medias_binom = pd.DataFrame(df.mean(), columns = ["Muestra_medias_binomial"])
sns.distplot(muestra_medias_binom)
```



2.1. **Ejemplos.**

Ejemplo 10. El análisis de los viajes diarios muestra que el número de pasajeros por automóvil, X, es una variable aleatoria discreta con distribuciones idénticas e independientes, tales que $\mathbb{E}[X] = 1.2$ y $\mathrm{Var}(X) = 1$. Estima la probabilidad de que, en una muestra de $n = 100$ automóviles, el número total de pasajeros es de 140 o menos.

Por lo tanto,

$$\mathbb{P}(S_n \leq 140) \approx \Phi(2) = 0.9772.$$

Ejemplo 11. Hay 100 personas en un avión. Sea X_i el peso (en kg) de la i-ésima persona en el avión. Supongamos que las variables aleatorias $X_1, \ldots, X_n$ son independientes e idénticamente distribuídas y $\mathbb{E}[X_i] = \mu = 170$ and $\mathrm{Var}(X_i) = \sigma^2 = 30$. Estimemos la probabilidad de que el peso total de las personas en el avión exceda los 18,000 kg. Si S_{100} es la variable aleatoria que denota el peso total, entonces $S_{100} = X_1 + X_2 + \cdots + X_{100}$. Tenemos $\mathbb{E}[S_{100}] = (100)(170) = 17000$. y $\mathrm{Var}(S_{100}) = (100)(30)^2 = 90000$. Por el TCL

$$\begin{aligned}
\mathbb{P}(S_{100} > 18000) &= \mathbb{P}\left(\frac{S_{100} - 17000}{300} > \frac{18000 - 17000}{300}\right) \\
&= \mathbb{P}\left(\frac{S_{100} - 17000}{300} > \frac{10}{3}\right) \\
&= 1 - \Phi\left(\frac{10}{3}\right) \\
&\approx 4.3 \times 10^{-4}.
\end{aligned}$$

Ejemplo 12. Sean $X_1, X_2, \cdots, X_{25}$ variables aleatorias discretas independientes e idénticamente distribuídas con función de masa de probabilidades dada por

$$P_X(k) = \begin{cases} 0.6 & k = 1 \\ 0.4 & k = -1 \\ 0 & \text{otherwise} \end{cases}$$

Sea $S_n = X_1 + \cdots + X_n$. Usando el TCL, estimemos $\mathbb{P}(4 \leq S_n \leq 6)$. Se tiene $\mathbb{E}[X_i] = \frac{1}{5}$ y $\mathbb{E}[X_i^2] = 1$. Por lo tanto, $\mathrm{Var}(X_i) = 1 - \frac{1}{25} = \frac{24}{25}$. De esta forma, $\mathbb{E}[S_n] = 5$ y $\mathrm{Var}(S_n) = 24$. Así,

$$\begin{aligned}
\mathbb{P}(4 \leq S_n \leq 6) &= \mathbb{P}\left(\frac{4-5}{2\sqrt{6}} \leq \frac{S_n - 5}{2\sqrt{6}} \leq \frac{6-5}{2\sqrt{6}}\right) \\
&= \mathbb{P}\left(-\frac{1}{2\sqrt{6}} \leq Z_n \leq \frac{1}{2\sqrt{6}}\right) \\
&\approx \Phi(0.2041) - \Phi(-0.2041) \\
&= 2\Phi(0.2041) - 1 \\
&\approx 0.16172
\end{aligned}$$

Ejemplo 13. Sean $X_1, X_2, \cdots, X_n$ variables aleatorias independientes e idénticamente distribuídas exponencialmente con $\lambda = 1$. Sea

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

Veamos qué tan grande debe ser n para asegurar que

$$\mathbb{P}(0.9 \leq \overline{X} \leq 1.1) \geq 0.95$$

Sea $S_n = X_1 + \cdots + X_n$, so $\bar{X} = \frac{Y}{n}$. Tenemos que,

$$\mathbb{E}[X_i] = 1, \quad \mathrm{Var}(X_i) = 1$$

De aquí que, $\mathbb{E}[S_n] = nE[X_i] = n$ y $\mathrm{Var}(X_i) = n$.

Por lo tanto,

$$
\begin{aligned}
\mathbb{P}(0.9 \leq \overline{X}_n \leq 1.1) &= \mathbb{P}\left(0.9 \leq \frac{S_n}{n} \leq 1.1\right) \\
&= \mathbb{P}(0.9n \leq S_n \leq 1.1n) \\
&= \mathbb{P}\left(\frac{0.9n - n}{\sqrt{n}} \leq \frac{S_n - n}{\sqrt{n}} \leq \frac{1.1n - n}{\sqrt{n}}\right) \\
&= \mathbb{P}\left(-0.1\sqrt{n} \leq Z_n \leq 0.1\sqrt{n}\right)
\end{aligned}
$$

Por el TCL

$$
\begin{aligned}
\mathbb{P}(0.9 \leq \bar{X} \leq 1.1) &\approx \Phi(0.1\sqrt{n}) - \Phi(-0.1\sqrt{n}) \\
&= 2\Phi(0.1\sqrt{n}) - 1
\end{aligned}
$$

Necesitamos que $2\Phi(0.1\sqrt{n}) - 1 \geq 0.95$ o, equivalentemente, $\Phi(0.1\sqrt{n}) \geq 0.975$
Por lo tanto,

$$
\begin{aligned}
0.1\sqrt{n} &\geq \Phi^{-1}(0.975) = 1.96 \\
\sqrt{n} &\geq 19.6 \\
n &\geq 384.16
\end{aligned}
$$

De esta manera, podemos tomar $n \geq 385$.