Tablero de Galton (Para la clase 8)

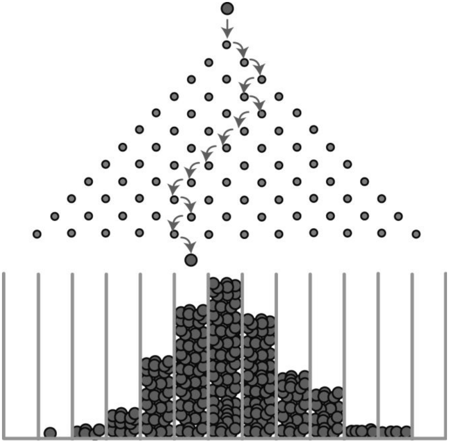

Figura: Verlo en acción en
https://www.youtube.com/watch?v=1DTRzPRfu6s

Probabilidad - Clase 8
Teorema local de De Moivre-Laplace

Ernesto Mordecki

Facultad de Ciencias, Universidad de la República. Montevideo, Uruguay

Curso de la Licenciatura en Matemática - 2020

Contenidos

Teorema local de De Moivre-Laplace

Teorema

- *Consideremos una serie de n experimentos independientes, con probabilidad de éxito en cada experimento igual a p $(0 < p < 1)$, $q = 1 - p$.*
- *Sean $P_n(m)$ la probabilidad de obtener m éxitos en n experimentos,*
- *Sea $x = x_{n,m} = \frac{m-np}{\sqrt{npq}}$.*

Entonces, tiene lugar la convergencia

$$\frac{\sqrt{npq}P_n(m)}{\frac{1}{\sqrt{2\pi}}e^{-x^2/2}} \to 1, \quad si \quad n \to \infty, \tag{1}$$

uniformemente en el conjunto de los valores de m tales que $|x_{n,m}| \leq C$, donde C es una constante arbitraria.

Enunciado alternativo

Una manera alternativa de escribir la convergencia que tiene lugar en (15), es

$$\sup \left| \frac{\sqrt{npq} P_n(m)}{e^{-x^2/2} / \sqrt{2\pi}} - 1 \right| \to 0, \quad \text{si} \quad n \to \infty,$$

donde el supremo se toma en el conjunto de valores de m tales que $|x_{n,m}| \leq C$.

Demostracion

La demostración se basa en la distribución binomial y en la fórmula de Stirling. En vista de la definición de x, tenemos

$$m = np + x\sqrt{npq}, \tag{2}$$
$$n - m = nq - x\sqrt{npq}. \tag{3}$$

Estas fórmulas, y la condición $|x_{n,m}| \leq C$, implican que $m \to \infty$ y $n - m \to \infty$ cuando $n \to \infty$.

De la fórmula binomial, se obtiene, que

$$P_n(m) = \binom{n}{n} p^m q^{n-m} = \frac{n!}{m!(n-m)!} p^m q^{n-m}.$$

Escribimos la fórmula de Stirling como

$$n! = n^n e^{-n} \sqrt{2\pi n}(1 + \alpha_n),$$

donde

$$\alpha_n \to 0 \ (n \to \infty),$$

Sustituyendo, obtenemos

$$\begin{aligned} P_n(m) = {} & \frac{n^n e^{-n}\sqrt{2\pi n}(1+\alpha_n)}{m^m e^{-m}\sqrt{2\pi m}(1+\alpha_m)} \\ & \times \frac{1}{(n-m)^{n-m} e^{-n+m}\sqrt{2\pi(n-m)}(1+\alpha_{n-m})} \\ & \times p^m q^{n-m}, \end{aligned}$$

Primero observamos que los e se cancelan. Agrupando los errores (los α)

Entonces

$$P_n(m) = \frac{n^n\sqrt{2\pi n}}{(n-m)^{n-m}\sqrt{2\pi(n-m)}m^m\sqrt{2\pi m}} \times p^m q^{n-m}\frac{(1+\alpha_n)}{(1+\alpha_m)(1+\alpha_{n-m})}$$

$$= \frac{n^{n+1/2}\, p^m q^{n-m}}{\sqrt{2\pi}(n-m)^{n-m+1/2}m^{m+1/2}}(1+\beta_{n,m})$$

donde

$$\beta_{n,m} = \frac{1+\alpha_n}{(1+\alpha_m)(1+\alpha_{n-m})} - 1.$$

Seguimos:

$$P_n(m) = \frac{n^{n+1/2}\, p^m q^{n-m}}{\sqrt{2\pi}(n-m)^{n-m+1/2} m^{m+1/2}}(1+\beta_{n,m})$$

$$= \frac{n^{m+1/2}\; n^{n-m+1/2}\; p^{m+1/2} q^{n-m+1/2}}{\sqrt{2\pi npq}\;(n-m)^{n-m+1/2}\; m^{m+1/2}}(1+\beta_{n,m})$$

$$= \frac{1}{\sqrt{2\pi npq}}\left(\frac{np}{m}\right)^{m+\frac{1}{2}}\left(\frac{nq}{n-m}\right)^{n-m+\frac{1}{2}}(1+\beta_{n,m}).$$

Recordemos el enunciado:

$$\frac{\sqrt{npq}P_n(m)}{\frac{1}{\sqrt{2\pi}}e^{-x^2/2}} \to 1, \quad \text{si} \quad n \to \infty,$$

Llegamos a

$$P_n(m) = \frac{1}{\sqrt{2\pi npq}}\Big(\frac{np}{m}\Big)^{m+\frac{1}{2}}\Big(\frac{nq}{n-m}\Big)^{n-m+\frac{1}{2}}(1+\beta_{n,m}).$$

Primero vemos que

$$\beta_{n,m} \to 0,$$

cuando $n \to \infty$, uniformemente, en el conjunto de los valores de m tales que $|x_{n,m}| \leq C$.

En efecto, teníamos

$$x = x_{n,m} = \frac{m - np}{\sqrt{npq}}, \qquad |x_{n,m}| \leq C.$$

De aquí

$$\begin{aligned} m &= np + x\sqrt{npq}, \\ n - m &= nq - x\sqrt{npq}, \end{aligned}$$

lo que obliga a $m \to \infty$ y $n - m \to \infty$ si $n \to \infty$. Como

$$\beta_{n,m} = \frac{1 + \alpha_n}{(1 + \alpha_m)(1 + \alpha_{n-m})} - 1.$$

obtenemos $\beta_{n,m} \to 0$ si $n \to \infty$.

Es conveniente reescribir la fórmula anterior, como

$$\sqrt{2\pi npq}P_n(m) = \left(\frac{m}{np}\right)^{-(m+1/2)}\left(\frac{n-m}{nq}\right)^{-(n-m+1/2)}(1+\beta_{n,m}).$$

De las expresiones (2) y (3) se obtiene, que

$$\frac{m}{np} = 1 + x\sqrt{\frac{q}{np}},$$

$$\frac{n-m}{nq} = 1 - x\sqrt{\frac{p}{nq}}.$$

Así aparece la x en la fórmula:

$$\sqrt{2\pi npq}P_n(m) = T_{n,m}(1 + \beta_{n,m}), \tag{4}$$

donde el término $T_{n,m}$ está dado por

$$T_{n,m} = \Big(1 + x\sqrt{\frac{q}{np}}\Big)^{-(m+1/2)}\Big(1 - x\sqrt{\frac{p}{nq}}\Big)^{-(n-m+1/2)}.$$

Resta tomar logaritmos naturales y aplicar (dos veces) el desarrollo de Taylor. se tiene

$$\ln T_{n,m} = -(m+1/2)\ln\Big(1+x\sqrt{\frac{q}{np}}\Big)-(n-m+1/2)\ln\Big(1-x\sqrt{\frac{p}{nq}}\Big). \tag{5}$$

Nos queda demostrar que

$$\ln T_{n,m} = -x^2/2 + r_{n,m}$$

y ahí aparece la exponencial: equivale a

$$T_{n,m} = e^{-x^2/2} e^{r_{n,m}}$$

donde $r_{n,m} \to 0$ ($n \to \infty$) uniformemente.

Recordemos una vez mas el enunciado:

$$\frac{\sqrt{npq} P_n(m)}{\frac{1}{\sqrt{2\pi}} e^{-x^2/2}} \to 1, \quad \text{si} \quad n \to \infty,$$

Utilizamos entonces el desarrollo de Taylor de la función $\ln(1+u)$, en el punto $u=0$, que establece que

$$\ln(1+u) = u - \frac{u^2}{2} + \theta u^3, \qquad \text{si } |u| < 1/2.$$

Poniendo $u = x\sqrt{\frac{q}{np}}$ en el primer sumando en (5), y $u = -x\sqrt{\frac{p}{nq}}$ en el segundo, y teniendo en cuenta (2) y (3), resulta

$$\begin{aligned}\ln T_{n,m} = &-(np + x\sqrt{npq} + 1/2)\Big(x\sqrt{\frac{q}{np}} - \frac{x^2 q}{2np} + \theta_1\Big(x\sqrt{\frac{q}{np}}\Big)^3\Big)\\ &- (nq - x\sqrt{npq} + 1/2)\Big(-x\sqrt{\frac{p}{nq}} - \frac{x^2 p}{2nq} + \theta_2\Big(x\sqrt{\frac{p}{nq}}\Big)^3\Big)\end{aligned}$$

donde $|\theta_1| < 3$, $|\theta_2| < 3$, para todo n suficientemente grande.

¡Tenemos 18 términos!

Lo importante es entender el tamaño de cada uno.

- Hay dos términos de orden $\sqrt{n}$, pero son opuestos: se cancelan

$$-np\,x\sqrt{\frac{q}{np}} \quad \text{y} \quad (-nq)\left(-x\sqrt{\frac{p}{nq}}\right)$$

- Hay cuatro términos de orden constante (se cancela el n):

$$-x^2q + \frac{x^2q}{2} - x^2p + \frac{x^2p}{2} = -\frac{x^2}{2}.$$

- Todos los demás términos tienden a cero.

En resumen

$$
\begin{aligned}
\ln T_{n,m} &= -x\sqrt{npq} + x\sqrt{npq} \\
&\quad - x^2 q + \frac{x^2 q}{2} - x^2 p + \frac{x^2 p}{2} \\
&\quad + r_{n,m} \\
&= -\frac{x^2}{2} + r_{n,m},
\end{aligned}
$$

donde $|r_{n,m}| \leq C_0/\sqrt{n}$, con C_0 una constante que depende únicamente de p, q y C.

Sustituyendo en

$$\sqrt{2\pi npq}P_n(m) = T_{n,m}(1 + \beta_{n,m}),$$

obtenemos

$$\sqrt{2\pi npq}P_n(m) = e^{-x^2/2}e^{r_{n,m}}(1 + \beta_{n,m}),$$

que escrita en forma similar al enunciado del teorema, es

$$\frac{\sqrt{npq}P_n(m)}{\frac{1}{\sqrt{2\pi}}e^{-x^2/2}} = e^{r_{n,m}}(1 + \beta_{n,m}),$$

Como $\beta_{m,n}$ y $r_{n,m}$ convergen uniformemente a cero, en el conjunto de los valores de m tales que $|x_{m,n}| \leq C$, la demostración está terminada (RH)[1].

---

[1] Respiramos hondo

Algunas consecuencias

En vista del teorema recién demostrado, se obtiene, que para n suficientemente grande, tiene lugar la identidad aproximada

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \varphi\left(\frac{m - np}{\sqrt{npq}}\right),$$

donde

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

La función $\varphi(x)$ se denomina *densidad normal*, o también *campana de gauss*.

Aplicación

Con esta aproximación, estamos ahora en condiciones de calcular (aproximadamente) la probabilidad

$$P_{10\ 000}(230) = \binom{10\ 000}{230} p^{230} q^{9770}$$

de un ejemplo anterior. Tenemos

- $p = 0{,}02$
- $q = 1 - p = 0{,}98$,
- El x es
$$x = \frac{m - np}{\sqrt{npq}} = \frac{230 - 200}{14} = 2{,}14.$$

Entonces

$$P_{10000}(230) \approx \frac{1}{14}\varphi(2.14) = \frac{1}{14} \times 0{,}0404 = 0{,}0029.$$

—¿Cómo se llama el nuevo profesor del Departamento de Matemáticas?
—No lo sé pero creo que le llaman «el épsilon».
—¿Por qué?
—Porque es pequeño y despreciable.

—Tú que eres matemático, ¿crees en Dios?
—Sí, salvo isomorfismos.

Figura: ¡Que pasen lindo!