G6P10

Estudie el grado de validez del teorema central del límite dibujando las distribuciones siguientes, y superponiendo sobre ellas la gaussiana con el μ y σ correspondiente.

(a) $B_k(5, 0{,}2), B_k(30, 0{,}4)$

(b) $P_n(4), P_n(10), P_n(40)$

a)

Sea x una variable aleatoria con distribución binomial. Su función generatriz se escribe como $G_x(z) = [p(z-1)+1]^n$. Defino otra variable aleatoria $y = \sum_{i=1}^{k} x_i$, donde las x_i son variables aleatorias independientes con distribución binomial. Entonces la función generatriz de y se escribe como

$$G_y(z) = \prod_{i=1}^{k} G_{x_i}(z) = [G_x(z)]^k = [p(z-1)+1]^{nk}$$

Obtuve que la distribución de y, que es suma de k variables con distribución binomial, también es binomial. Es más, no sólo es binomial, sino que su parámetro es suma de los parámetros de cada variable x, es decir $n_y = \sum_{i=1}^{k} n = nk$. Entonces no hace falta sumar variables aleatorias, basta con agrandar el parámetro.

En la Fig.1 se observan los gráficos obtenidos para las distribuciones binomiales $B_k(5, 0{,}2)$ y $B_k(30, 0{,}4)$.

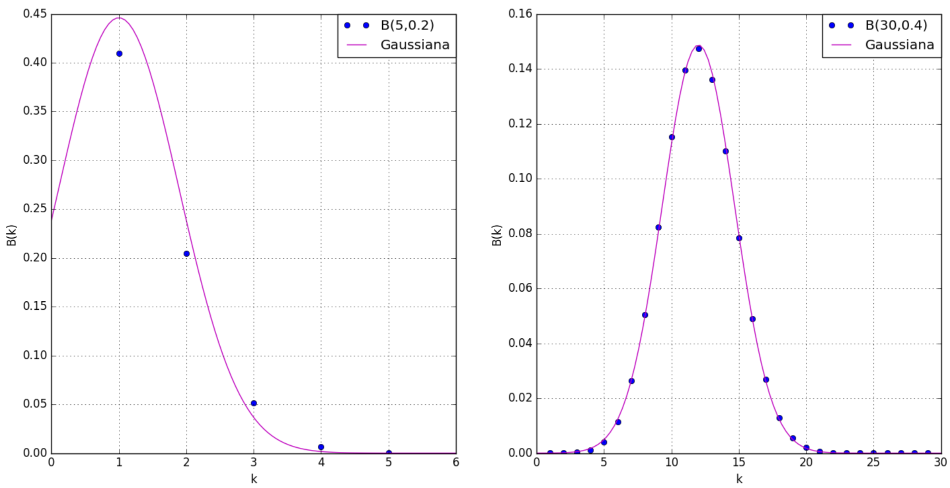

Figura 1: Gráficos de distribuciones binomiales con gaussianas superpuestas.

Se superpusieron a los gráficos las gaussianas N(1.00, 0.89) y N(12.00, 2.48), cuyos μ y σ fueron calculados como la esperanza y la raíz de la varianza para una distribución binomial:

$$\mu = np \qquad \sigma = \sqrt{np(1-p)}$$

En el primer gráfico, la gaussiana no representa una buena aproximación de los puntos. En cambio, en el segundo gráfico, donde se observa una binomial cuya variable aleatoria podría pensarse como $x^B = \sum_{i=1}^{6} x_i^A$, donde x^A son variables aleatorias con distribución $B_k(5, 0{,}2)$ (es decir, correspondientes al primer gráfico), la gaussiana sí representa una buena aproximación de los puntos.

b)

Sea x una variable aleatoria con distribución poissoniana. Su función generatriz se escribe como $G_x(z) = \exp[\mu(z-1)]$. Defino otra variable aleatoria $y = \sum_{i=1}^{k} x_i$, donde las x_i son variables aleatorias independientes con distribución poissoniana. Entonces la función generatriz de y se escribe como

$$G_y(z) = \prod_{i=1}^{k} G_{x_i}(z) = \prod_{i=1}^{k} \exp[\mu(z-1)] = \exp[\sum_{i=1}^{k} \mu(z-1)] = \exp[n\mu(z-1)]$$

Al igual que en ocurre con la binomial, obtuve que la distribución de y es la misma que la de x, con parámetro $n\mu$.
En la Fig.2 se observan los gráficos obtenidos para las distribuciones poissonianas $P_n(4)$, $P_n(10)$ y $P_n(40)$.

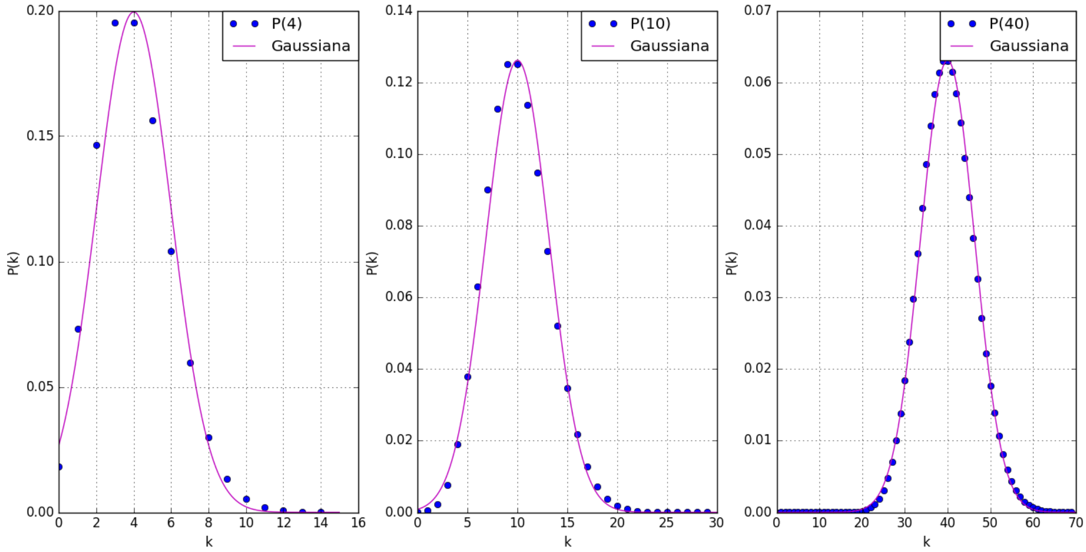

Figura 2: Gráficos de distribuciones poissonianas con gaussianas superpuestas.

Se superpusieron a los gráficos las gaussianas N(4, 2), N(10.00,3.16) y N(40.00, 6.32), cuyos μ y σ fueron calculados como la esperanza y la raíz de la varianza para una distribución poissoniana, siendo éstas μ y $\sqrt{\mu}$.

A medida que aumenta el μ, la gaussiana aproxima mejor los puntos. En este caso, aumentar μ es lo mismo que considerar una variable aleatoria suma de variables con distribución poissoniana. Por ejemplo, una variable con distribución $P_n(40)$ puede pensarse como la suma de 4 variables con distribución $P_n(10)$ o 10 variables con distribución $P_n(4)$.

Tanto en el caso binomial como en el poissoniano, se observa claramente que a medida que aumenta el parámetro, la gaussiana aproxima mejor los puntos. Estas dos distribuciones tienen la propiedad de que si se suman variables aleatorias con esa distribución, se obtiene otra variable aleatoria con la misma distribución, y de parámetro n $\times$ el parámetro original. Entonces aumentar el parámetro es lo mismo que sumar más variables aleatorias, y cuanto mayor sea el parámetro (más variables estén sumadas), por el teorema central del límite, más se parecerá la distribución de la nueva variable a una gaussiana.