



- Métodos computacionales:
Alejandro Segura
- Estadística
 - a) Incluir el código Notebook (.ipynb).
 - b) Guardar la información en una carpeta llamada **Semana15_Nombre1_Nombre2**
 - c) **Hacer una sola entrega por grupo.**

Contents

1	Estadística	3
1.1	Bondad del ajuste para pseudo-generadores	4
1.1.1	Uniformidad en $d=2$	4
1.1.2	Uniformidad en $d=3$	4

List of Figures

1	Distribución χ^2_{99} , el valor observado es menor que el valor crítico de la distribución. $p -$ $value = 0.111$, se acepta H_0	5
2	Distribución χ^2_{999} , el valor observado es menor que el valor crítico de la distribución. $p -$ $value = 0.190$, se acepta H_0	5

1 Estadística

1.1 Bondad del ajuste para pseudo-generadores

¿Sabías que el generador de números de Numpy (Mersenne - Twister 1998), tiene uniformidad en $d=623$ dimensiones al 95% de nivel de confianza?

1. Para probar la uniformidad de los números aleatorios generados por pseudo-generadores es posible realizar una prueba de hipótesis. Por lo cuál, se tienen las siguientes hipótesis:
 - a) H_0 : La secuencia de números sigue una distribución uniforme en k -dimensiones.
 - b) H_1 : La secuencia de números no sigue una distribución uniforme en k -dimensiones.

1.1.1 Uniformidad en $d=2$

En el caso 2D, se divide el espacio entre $[0, 1]$ en k^d celdas de igual área, luego se genera una muestra aleatoria de tamaño $2n$ (note que el número de eventos es par), podemos construir n pares no solapados $(x_0, x_1), (x_2, x_3), \dots, (x_{2n-1}, x_{2n})$ y estimar la frecuencia observada de puntos en cada celda. Si suponemos que la hipótesis nula es verdadera, la frecuencia esperada de puntos en cada celda es:

$$f_e^i = \frac{n}{k^d} \quad (1)$$

El estadístico de prueba que mide la discrepancia entre la frecuencia esperada y observada está dado por el estadístico χ^2 :

$$\chi^2 = \sum_{i=1}^k \frac{(f_o^i - f_e^i)^2}{f_e^i} \quad (2)$$

Calcule el valor crítico de la distribución χ_{α, k^d-1}^2 , donde la significancia es $\alpha = 0.05$. k es el número de celdas en las que se divide el dominio. Note que los grados de libertad son $gl = k^d - 1$.

- (a) Genere una muestra de $2 * (1000)$ números aleatorios con el generador de Numpy.
- (b) Use una partición de $k = 10$ celdas.
- (c) Calcule las frecuencias observadas y estadístico de prueba. Guardar los conteos en objetos $M = \text{np.array}((k, k))$. En este caso, es posible usar histogramas 2D para contar los puntos en cada celda.
- (d) Calcule el p-value de la observación.
- (e) ¿Valida o rechaza la hipótesis al 95% de nivel de confianza de que los números tienen una distribución uniforme en $d = 2$ dimensiones?

1.1.2 Uniformidad en $d=3$

En el caso 3D, se divide el espacio entre $[0, 1]$ en k^d celdas de igual volumen, luego se genera una muestra aleatoria de tamaño $3n$ (note que el número de eventos es múltiplo de 3), podemos construir n pares no solapados $(x_0, x_1, x_2), (x_3, x_4, x_5), \dots, (x_{3n-2}, x_{3n-1}, x_{3n})$ y estimar la frecuencia observada de puntos en cada celda. Si suponemos que la hipótesis nula es verdadera, la frecuencia esperada de puntos en cada celda es:

$$f_e^i = \frac{n}{k^d} \quad (3)$$

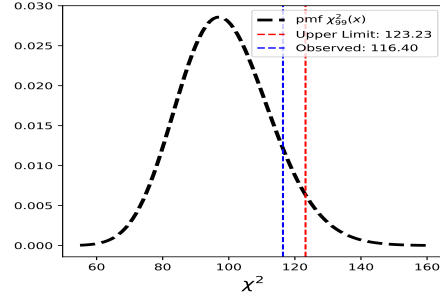


Figure 1: Distribución χ^2_{99} , el valor observado es menor que el valor crítico de la distribución. $p - value = 0.111$, se acepta H_0 .

El estadístico de prueba que mide la discrepancia entre la frecuencia esperada y observada está dado por el estadístico χ^2 :

$$\chi^2 = \sum_{i=1}^k \frac{(f_o^i - f_e^i)^2}{f_e^i} \quad (4)$$

Calcule el valor crítico de la distribución $\chi^2_{\alpha, k^d - 1}$, donde la significancia es $\alpha = 0.05$. k es el número de celdas en las que se divide el dominio. Note que los grados de libertad son $gl = k^d - 1$.

- Genere una muestra de $3 * (1000)$ números aleatorios con el generador de **Numpy**.
- Use una partición de $k = 10$ celdas.
- Calcule las frecuencias observadas y estadístico de prueba. Guardar los conteos en objetos **M** = `np.array((k,k,k))`.
- Calcule el p-value de la observación.
- ¿Valida o rechaza la hipótesis al 95% de nivel de confianza de que los números tienen una distribución uniforme en $d = 3$ dimensiones?

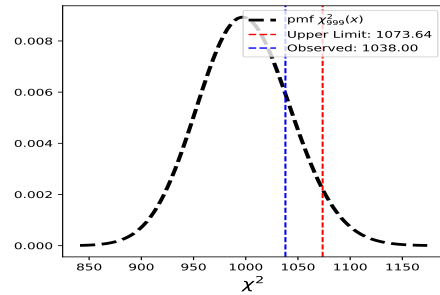


Figure 2: Distribución χ^2_{999} , el valor observado es menor que el valor crítico de la distribución. $p - value = 0.190$, se acepta H_0 .