

Informe Tarea 1

Probabilidad con distribución χ^2

Nombre: Diego Román
RUT: 20.299.495-4
Profesor: Valentino González
Auxiliares: Vicente Donaire
Benjamín Navarrete
Fecha de entrega: 26 de septiembre

1. Introducción

El problema a resolver es encontrar numéricamente el valor de a para la siguiente ecuación:

$$0.95 = \mathbb{P}(x < a) = \int_0^a \chi^2(x) dx$$

Donde la distribución $\chi^2(x) = \frac{1}{2^{k/2}\Gamma(k/2)}x^{k/2-1}e^{-x/2}$, y $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$.

Se comenzará con la implementación de la función $\Gamma(z)$. Para su implementación en código se regularizará la integral por medio de un cambio de variable apropiado. La integración se hará usando el método de Simpson, teniendo como parámetro la convergencia relativa deseada. Debido a la singularidad, la integración irá acompañada por el método del punto medio para obtener una aproximación de la integral en valores cercanos a $x = 0$.

Luego, usando el mismo método de Simpson se calcularán valores de $\mathbb{P}(x < a)$ para $5 < a < 15$, que irán acercándose a $\mathbb{P}(x < a) = 0.95$ con la tolerancia absoluta que se requiera gracias al método de la bisección.

2. Desarrollo

Tomando $u = e^{-x}$ la integral queda:

$$\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx = \int_0^1 \left(\ln\left(\frac{1}{u}\right)\right)^{z-1} du$$

La integración numérica de $\Gamma(z)$ requirió del uso del método del punto medio para el caso de la singularidad en $z = 0$, donde se calculó la integral entre 0 y un $\Delta x = 1 \times 10^{-6}$ como $I_0 = f(\frac{\Delta x}{2})\Delta x$.

Desde $z = \Delta x$ hasta $z = 1$ se usó Simpson vía método de los trapecios. Se calculó $S_{2N} = \frac{1}{3}(4T_{2N} - T_N)$, donde S_{2N} es la integral usando Simpson, mientras que T_N , T_{2N} son las integrales usando trapecios.

Las integrales T_N se calculan como $T_N = \left(\frac{f(x_0) + f(x_N)}{2} + \sum_{i=1}^{N-1} f(x_i) \right) \Delta x$.

Se ha definido un parámetro llamado `tol_rel` que permite determinar el cumplimiento de la siguiente desigualdad:

$$\left| \frac{S_{2N} - T_N}{T_N} \right| > tol_rel \quad (1)$$

En caso contrario, no será necesario volver a llamar recursivamente la función considerando el doble de puntos. Se determinó el valor a usar de la tolerancia relativa gracias a la figura 1. Se usaron valores conocidos de $\Gamma(z)$ (si $n \in \mathbb{N}$, $\Gamma(z) = (z-1)!$)

Se aprecia una línea casi recta para un espacio logarítmico (log-log). El promedio se calculó tomando la media aritmética del valor absoluto de la diferencia entre el valor real de $\Gamma(z)$ y la aproximación numérica para $z \in \{1, 2, 3, 4\}$.

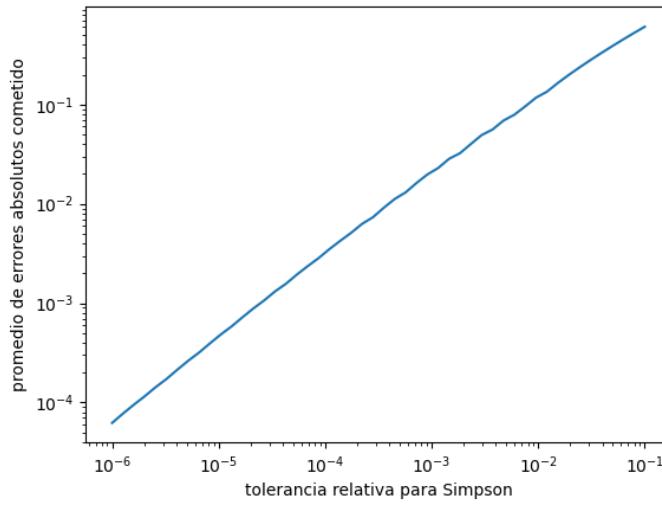


Figura 1: Error absoluto promedio en función de la tolerancia relativa otorgada como parámetro al método de Simpson.

La función $\chi^2(x)$ tiene su dominio en $\mathbb{R}^+ \cup \{0\}$, por lo que para el cálculo de $\mathbb{P}(x < a)$ no se requirió un cambio de variables que solucionara el problema del extremo inferior $x = -\infty$ (se calcula desde $x = 0$).

Finalmente para encontrar el valor de a se ha usado el método de la bisección otorgándole un parámetro de tolerancia absoluta de 1×10^{-12} . Este último se obtuvo gracias a la figura 2, que indica una mejora de precisión a menor valor de tolerancia.

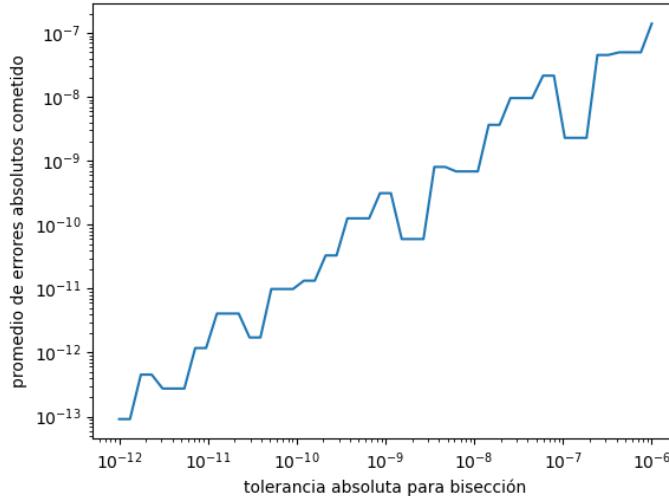


Figura 2: Error absoluto para distintas tolerancias absolutas.

Visualmente, la resolución del problema consistía en encontrar el a que hiciera el área azul bajo la curva $y = \chi^2(x)$ igual a 0.95, cuestión ilustrada en la figura 3

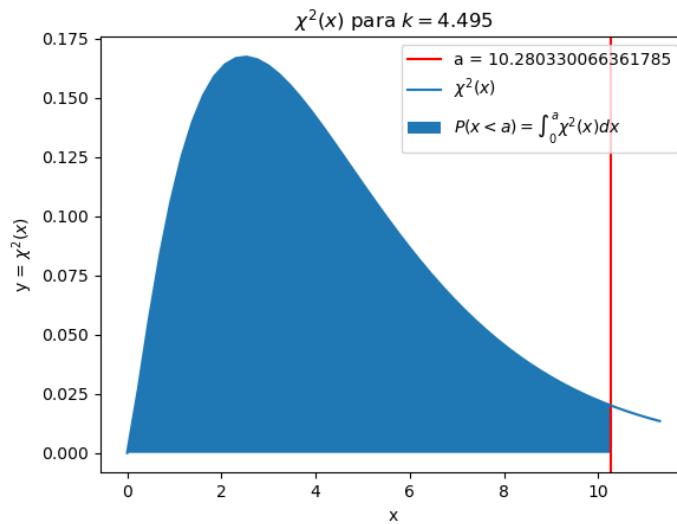


Figura 3: Área bajo la curva $y = \chi^2(x)$ desde $x = 0$ hasta el valor encontrado $x = a$.

3. Resultados

4. Discusión y Conclusiones