

# Informe Tarea 1

Probabilidad con distribución  $\chi^2$

Nombre: Diego Román  
RUT: 20.299.495-4  
Github: [@DiegoRomanCortes](#)  
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  $0 < a < 20$ , que irán acercándose a  $\mathbb{P}(x < a) = 0.95$  con la tolerancia absoluta que se requiera gracias al método de Newton, otorgándole un valor inicial dado por el método de la bisección.

# 2. Desarrollo

El integrando original describe un área que se puede apreciar en la figura 1a 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  $z = 0$  que se puede apreciar en la figura 1b, donde se calculó la integral entre 0 y un  $\Delta x > 0$  como  $I_0 = f\left(\frac{\Delta x}{2}\right)\Delta x$ .

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 y  $N$  es el número de puntos considerados.

Las integrales usando método de trapecios se calculan como  $T_N = \left( \frac{f(a) + f(b)}{2} + \sum_{i=1}^{N-1} f(x_0 + i\Delta x) \right) \Delta x$ , con  $\Delta x = \frac{b-a}{N}$  (se ha usado el mismo valor de  $\Delta x$  para el punto medio y para Simpson en el cálculo de  $\Gamma(k/2)$ ).

Se ha definido un parámetro llamado **rel\_tol** que permite determinar el cumplimiento de la siguiente desigualdad:

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

En caso contrario, no será necesario volver a llamar recursivamente la función considerando el doble de puntos.

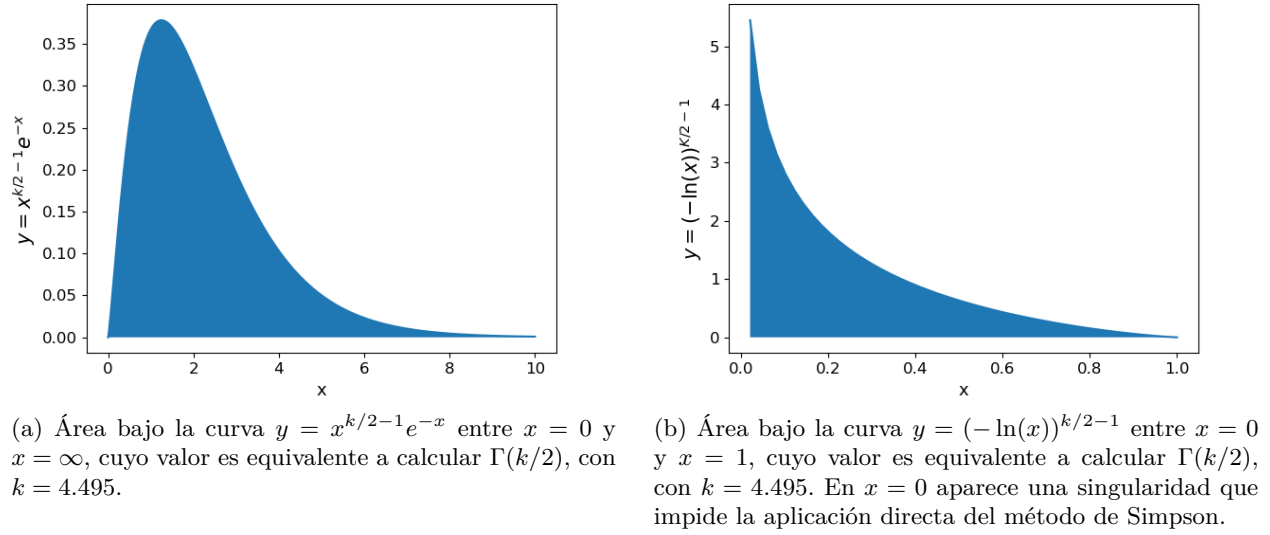


Figura 1: Área bajo los integrandos para el cálculo de  $\Gamma(k/2)$ .

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 Newton-Rhapon. Se definió en el código la función **prob\_menos\_95** para que el problema final sea calcular el cero de la función  $f(a) = \int_0^a \chi^2(x)dx - 0.95$ . Para otorgarle un punto inicial cercano al cero a encontrar, se utilizó el método de bisección para  $0.1 < a < 20$ . El método de la bisección toma dos puntos  $x_1$  y  $x_2$  tales que  $f(x_1)f(x_2) < 0$ . Se escoge  $p = \frac{x_1+x_2}{2}$  para luego determinar a qué lado del cero está  $f(p)$ . Este proceso se itera tantas veces como precisión se quiera.

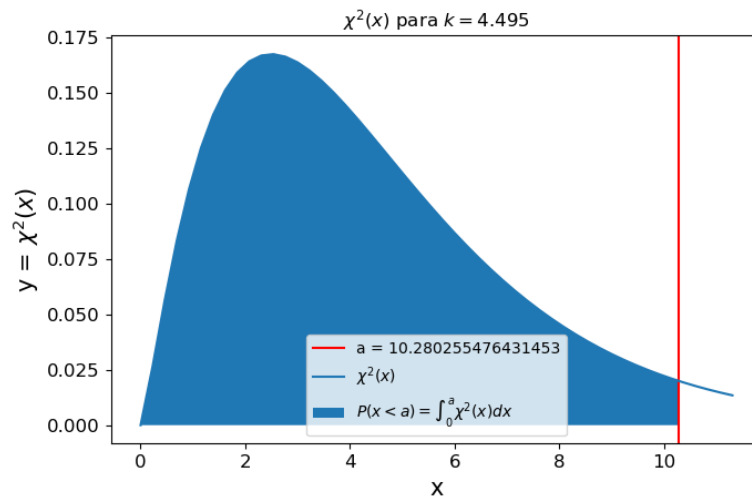


Figura 2: Área bajo la curva  $y = \chi^2(x)$  desde  $x = 0$  hasta el valor encontrado  $x = a$ .

Se usó el método de Newton-Rhapson además como comparación, consistente en iterar  $x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$ . La derivada de  $f(x)$  es  $f'(x) = \chi^2(x)$  por teorema fundamental del cálculo.

### 3. Resultados

El cambio de variable  $u = e^{-x}$  permitió obtener el valor de  $\Gamma(k/2) = 1.1313837482276219$  usando una tolerancia relativa **rel\_tol** =  $1 \times 10^{-6}$ . El método de Newton entrega el valor  $a = 10.280255476431453$  para una tolerancia absoluta **tol\_abs** =  $1 \times 10^{-12}$  en Newton y una relativa de  $1 \times 10^{-8}$  en Simpson (se usan ambos para determinar el valor de  $a$ ).

$\Delta x = \text{rel\_tol}$	Valor de $\Gamma(k/2)$	Tiempo con %timeit
$1 \times 10^{-5}$	<b>1.1313772563462283</b>	577 ms $\pm$ 16.1 ms
$1 \times 10^{-6}$	<b>1.1313837482276219</b>	4.46 s $\pm$ 137 ms
$1 \times 10^{-7}$	<b>1.131384407180971</b>	1min 12s $\pm$ 1.98 s
<b>scipy.special.gamma(k/2)</b>	<b>1.131384485494286</b>	-

Tabla 1: Tiempo de ejecución y valor de  $\Gamma(k/2)$  para tres valores de  $dx$ . Se aprecia un aumento significativo en el tiempo de ejecución a cambio de poca mejoría en la precisión. Se adjunta en la última fila el valor otorgado por la librería **scipy**, la cual se asume como infinitamente preciso para efectos de comparación.

Método	Tolerancia Absoluta	Número de pasos	Valor de $a$
Bisección	$1 \times 10^{-5}$	22	<b>10.280251502990723</b>
Newton	$1 \times 10^{-5}$	6	<b>10.280255476431392</b>
	$1 \times 10^{-12}$	3	<b>10.280255476431453</b>

Tabla 2: Comparación entre el método de la bisección y el de Newton para encontrar  $a$ . La tercera fila es en realidad la combinación del método de la bisección y el de Newton. Al último se le ha otorgado como valor inicial el resultado del primero.

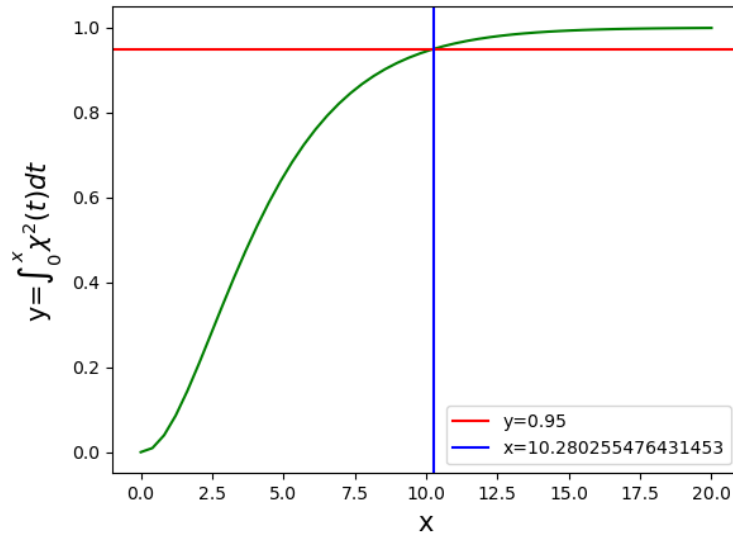


Figura 3: Valor de  $a$  encontrado con el método de Newton. Se muestra además la función probabilidad computada con el método de Simpson.

## 4. Discusión y Conclusiones

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 2. El valor de  $a$  que se retorna depende fuertemente de los valores de las tolerancias entregadas.

La resolución del problema requería del uso de un cambio de variable que ocasionó un problema adicional (la singularidad). Una mejor elección de esto último pudo haber evitado su aparición. El costo de tener que calcular aparte la región cercana a la singularidad se aprecia en la tabla 1, donde por razones de tiempo no es posible calcular más allá del séptimo decimal. Esta es la razón de la elección de  $\Delta x = 1 \times 10^{-6}$  para el cálculo de  $\Gamma(k/2)$ . Como contraste, el método `scipy.specials.gamma`, que es un análogo del  $\Gamma(z)$  de la librería `scipy`, otorga un valor que sólo coincide en seis cifras significativas con el usado en este informe.

Se eligió el método de Newton debido a su rapidez; aunque como no está asegurada su convergencia se usó como intermediario al método de la bisección, el cual le otorgó a su par un valor inicial cercano al cero buscado. Es interesante destacar la notoriedad de la convergencia cuadrática del método de Newton en la tabla 2 asumiendo precisión exacta en la función  $\Gamma(z)$  programada, pues con solo 6 pasos obtiene 7 nuevos decimales de precisión (contrastado con los 22 pasos y 5 decimales del método de bisección, se percibe la mejora).