Zadanie 4

Wiktoria Kuna

316418

Polecenie

Rozkład t-studenta z k stopniami swobody ma gęstość:

$$f(x) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi}\Gamma(\frac{k}{2})}\left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}}, \ x \in \mathbb{R}$$

Dla ustalonego $t \in \mathbb{R}$ obliczyć wartość całki:

$$G(t) = \int_{-\infty}^{t} f(x)\,dx$$

Parzystość gęstości

Zauważmy, że gęstość rozkładu t-studenta jest funkcją parzystą.

Dowód:

$$f(x) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi}\Gamma(\frac{k}{2})}\left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}} = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi}\Gamma(\frac{k}{2})}\left(1 + \frac{(-x)^2}{k}\right)^{-\frac{k+1}{2}} = f(-x)$$

Lemat: G(t) = 1 - G(-t)

Dowód:

$$G(t) = \int_{-\infty}^{t} f(x)\,dx \overset{(1)}{=} \int_{-\infty}^{t} f(-x)\,dx = \int_{-t}^{\infty} f(x)\,dx =$$
$$\int_{\mathbb{R}} f(x) - \int_{-\infty}^{-t} f(x)\,dx = 1 - G(-t)$$

(1) Korzystamy z parzystości gęstości rozkładu t-studenta.

Przekształcenia

Aby móc skorzystać ze złożonego wzoru trapezów i metody Romberga, całka którą liczymy musi być właściwa. Musimy zatem przeprowadzić odpowiednie przekształcenia.

$$G(t) \overset{Lemat}{=} 1 - G(-t) = 1 - \int_{-\infty}^{-t} f(x)\,dx =$$
$$1 - \int_{-\infty}^{-t} \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi}\Gamma(\frac{k}{2})} \Big(1 + \frac{x^2}{k}\Big)^{-\frac{k+1}{2}} dx =$$
$$1 - \int_{-\infty}^{-t} \frac{\Gamma(\frac{1}{2}(k+1))}{\sqrt{k\pi}\Gamma(\frac{1}{2}k)} \Big(1 + \frac{x^2}{k}\Big)^{-\frac{k+1}{2}} dx$$

Wiemy, że $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

$$1 - \int_{-\infty}^{-t} \frac{\Gamma(\frac{1}{2}(k+1))}{\sqrt{k}\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}k)} \Big(1 + \frac{x^2}{k}\Big)^{-\frac{k+1}{2}} dx \tag{1}$$

Na ćwiczeniach udowadnialiśmy następującą zależność: $\Gamma(p)\Gamma(q) = \Gamma(p+q)B(p,q)$.

Niech $p = \frac{1}{2}$ oraz $q = \frac{1}{2}k$. Z powyższej zależności mamy:

$$\Gamma\Big(\frac{1}{2}\Big)\Gamma\Big(\frac{1}{2}k\Big) = \Gamma\Big(\frac{1}{2}(k+1)\Big)B\Big(\frac{1}{2}, \frac{1}{2}k\Big)$$
$$\frac{\Gamma\Big(\frac{1}{2}(k+1)\Big)}{\Gamma\Big(\frac{1}{2}\Big)\Gamma\Big(\frac{1}{2}k\Big)} = \frac{1}{B\Big(\frac{1}{2}, \frac{1}{2}k\Big)}$$

Teraz możemy podstawić powyższą zależność do (1):

$$1 - \int_{-\infty}^{-t} \frac{1}{B\Big(\frac{1}{2}, \frac{1}{2}k\Big)\sqrt{k}} \Big(1 + \frac{x^2}{k}\Big)^{-\frac{k+1}{2}} dx =$$
$$1 - \frac{1}{B\Big(\frac{1}{2}, \frac{1}{2}k\Big)\sqrt{k}} \int_{-\infty}^{-t} \Big(1 + \frac{x^2}{k}\Big)^{-\frac{k+1}{2}} dx$$

Dokonajmy teraz podstawienia za x. Od teraz rozpatrujemy wyłącznie $t \geqslant 0$.

$$\left|\begin{array}{c} y = \frac{k}{x^2+k} \\ \\ |x| = \sqrt{\frac{k(1-y)}{y}} \implies x = -\sqrt{\frac{k(1-y)}{y}} \; (x \leqslant 0) \\ \\ dx = -\frac{1}{2}(\frac{k(1-y)}{y})^{-\frac{1}{2}} \cdot \frac{-ky-k(1-y)}{y^2} dy = \\ \\ -\frac{1}{2}(\frac{k(1-y)}{y})^{-\frac{1}{2}} \cdot \frac{-k}{y^2} dy \end{array}\right|$$

Wraz z podstawieniem musimy oczywiście zmienić granice całkowania. Skoro x znajdował się na przedziale $(-\infty, -t]$, a $y = \frac{k}{x^2+k}$ to y będziemy całkować na przedziale $[0, \frac{k}{t^2+k}]$.

$$1 - \frac{1}{B\left(\frac{1}{2}, \frac{1}{2}k\right)\sqrt{k}} \int_0^{\frac{k}{t^2+k}} \left(1 + \frac{\frac{k(1-y)}{y}}{k}\right)^{-\frac{k+1}{2}} \frac{-1}{2} \left(\frac{k(1-y)}{y}\right)^{-\frac{1}{2}} \cdot \frac{-k}{y^2} \, dy =$$

$$1 - \frac{1}{B\left(\frac{1}{2}, \frac{1}{2}k\right)\sqrt{k}} \frac{1}{2} \int_0^{\frac{k}{t^2+k}} \left(1 + \frac{(1-y)}{y}\right)^{-\frac{k+1}{2}} \left(\frac{k(1-y)}{y}\right)^{-\frac{1}{2}} \cdot \frac{k}{y^2} \, dy =$$

$$1 - \frac{1}{B\left(\frac{1}{2}, \frac{1}{2}k\right)\sqrt{k}} \frac{1}{2} \int_0^{\frac{k}{t^2+k}} \left(\frac{1}{y}\right)^{-\frac{k+1}{2}} \left(\frac{k(1-y)}{y}\right)^{-\frac{1}{2}} \cdot \frac{k}{y^2} \, dy =$$

$$1 - \frac{1}{B\left(\frac{1}{2}, \frac{1}{2}k\right)} \frac{1}{2} \int_0^{\frac{k}{t^2+k}} \left(y\right)^{\frac{k}{2}-1} \cdot y^{\frac{1}{2}} \cdot \left(\frac{y}{1-y}\right)^{\frac{1}{2}} \cdot \frac{1}{y} \, dy =$$

$$1 - \frac{1}{B\left(\frac{1}{2}, \frac{1}{2}k\right)} \frac{1}{2} \int_0^{\frac{k}{t^2+k}} \left(y\right)^{\frac{k}{2}-1} \cdot \left(\frac{1}{1-y}\right)^{\frac{1}{2}} \, dy =$$

$$1 - \frac{1}{B\left(\frac{1}{2}, \frac{1}{2}k\right)} \frac{1}{2} \int_0^{\frac{k}{t^2+k}} \left(y\right)^{\frac{k}{2}-1} \cdot \left(1-y\right)^{\frac{1}{2}-1} \, dy =$$

$$1 - \frac{1}{2} \frac{B\left(\frac{k}{t^2+k}; \frac{k}{2}, \frac{1}{2}\right)}{B\left(\frac{1}{2}k, \frac{1}{2}\right)}$$

W ten sposób do policzenia mamy funkcję Beta i niepełną funkcję Beta, czyli dwie całki właściwe.

Wyprowadziliśmy postać dla $t \geqslant 0$, jednak wcześniej pokazaliśmy, że rozkład t-studenta jest symetryczny, stąd mamy równocześnie wartości dla $t \leqslant 0$.

Mamy zatem:

$$G(t) = \begin{cases} 1 - \frac{1}{2} \frac{B\left(\frac{k}{t^2+k}; \frac{k}{2}, \frac{1}{2}\right)}{B\left(\frac{1}{2}k, \frac{1}{2}\right)} & t \geqslant 0 \\ \frac{1}{2} \frac{B\left(\frac{k}{t^2+k}; \frac{k}{2}, \frac{1}{2}\right)}{B\left(\frac{1}{2}k, \frac{1}{2}\right)} & t \leqslant 0 \end{cases}$$

Metoda Romberga

Metoda Romberga polega na rekurencyjnym przybliżaniu wartości funkcji zaczynając od aporksymacji z złożonej metody trapezów.

Wzory realizujące metodę:

$$h_n = \frac{b-a}{2^n}$$
$$R(0,0) = h_1 \cdot (f(b) + f(a))$$
$$R(n,0) = R(n-1,0) + h_n \cdot \sum_{k=1}^{2^{(n-1)}} f(a + (2 \cdot k - 1) \cdot h_n)$$
$$R(n,m) = \frac{1}{4^m - 1} \cdot (4^m \cdot R(n, m-1) + R(n-1, m-1))$$

Program

Do obliczenia całki wykorzystamy złożony wzór trapezów i metodę Romberga. Do otrzymania wartości rozkładu t-studenta potrzebna nam watrość funkcji Beta i niekompletnej funkcji Beta. Do pierwszej użyjemy funkcji bibliotecznej, natomiast drugą policzymy metodą Romberga.

```
import scipy.integrate as inte
import scipy.special as sc

def fun(a, b):
    def iner(y):
        if y == 0:
            return 0.
        if y == 1:
            return 0.
        return y**(a - 1) * (1-y)**(b-1)
    return iner

def Tstudent_beta(k):
    return sc.beta(k/2., .5)

def Tstudent_beta_incomplete(k, t):
    f = fun(k/2., 1/2.)
    return inte.romberg(f, 0., k/(t**2+k), divmax=50, tol=10**(-8), show=True)

def Tstudent_CDF(k,t):
    romb = 0.5 * Tstudent_beta_incomplete(k,t)/Tstudent_beta(k)
    if t < 0:
        return romb
    else:
        return 1 - romb
```

Wartość divmax określa maksymalną liczbę kroków, na którą pozwalamy metodzie Romberga. Osiągnięcie jednak dokładności do 8 miejsca po przecinku może zająć dużo czasu, gdyż wymaga wielu iteracji.

Dokładność

Python zapewnia, że biblioteczna funkcja Beta wyliczy się z dokładnością do 10^{-14}. Jeśli damy wystarczająco czasu na wykonanie się programu, procedura korzystająca z metody Romberga da nam dokładność 10^{-8} (w wypadku niecierpliwym 10^{-5}). Stąd jesteśmy w stanie wyliczyć wartość $G(t)$ z dokładnością do 8 miejsc po przecinku.