Mikołaj Drozd 339139

Zadanie 2 (egzaminacyjne)

Treścią zadania było obliczenie dystrybuanty funkcji o rozkładzie $\chi^2(k)$ mającej postać $f(x)=1/(2^{k/2}\,\Gamma(k/2))*x^{k/2-1}*e^{-x/2}$ dla $k \in N$, $x \in [0, +\infty)$. Rozwiązanie należało uzyskać używając jak najmniej wbudowanych funkcji matematycznych.

Wiemy, że obliczenie wartości dystrybuanty z funkcji gęstości polega na obliczeniu całki z tej funkcji na przedziale od 0 do x. Działanie to było jednak utrudnione przez bardzo złożony wzór funkcji. Możemy jednak rozwikłać ten problem poprzez zastosowanie jednej z metod całkowania numerycznego poznanej w poprzednim semestrze, np. złożonego wzoru trapezów.

Takie podejście do zadania zdecydowanie upraszcza nam problem, ponieważ wiemy, że całka oznaczona równa się polu pod wykresem funkcji. Oznacza to, że wystarczy policzyć jej wartość w n równoodległych punktach oraz stworzyć trapezy, na podstawie których będziemy obliczać pole powierzchni. Wiemy, że dystrybuanta na przedziale od $-\infty$ do 0 wynosi 0, zaś na przedziale od x do $+\infty$ będzie ona wynosić 1, interesuje nas zatem jedynie przedział od 0 do x. Należy jeszcze zaznaczyć, że ponieważ korzystamy ze złożonego wzoru trapezów, to podczas obliczania wartości funkcji każdą z nich z wyjątkiem tej w 0 i w x obliczymy dwa razy, jako koniec jednego z przedziałów i jako początek drugiego. Skrajne wartości zaś wystarczy obliczyć tylko raz.

Pisząc rozwiązanie w programie Octave, podzieliłem je na kilka podproblemów. Jedna z funkcji oblicza wszystkie składniki funkcji gęstości, czyli kolejno $2^{k/2}$, $e^{-x/2}$ oraz $x^{k/2-1}$.

Druga z nich oblicza wartość funkcji gamma w sposób rekurencyjny. Wiemy bowiem, że wartość tej funkcji w punkcie 1 jest równa 1 (z definicji funkcji) oraz wartość w ½ wynosi $\sqrt{\pi}$ (zadanie 3 z listy 5). Znamy również następującą własność: $\Gamma(x+1)$ = x*Γ(x) (zadanie 3 z listy 1). Jest ona prawdziwa dla wszystkich k, ponieważ wiemy, że należą one do zbioru liczb naturalnych, zatem będziemy obsługiwać jedynie liczby naturalne lub ich połowy, a co za tym idzie, możemy używać rekurencyjnego wzoru i sprowadzać każdą rozwiązanie dla każdych danych do bazowych przypadków.

Ostatnia część rozwiązania to zastosowanie samego wzoru trapezów. Iterując się przez kolejne przedziały, uzyskujemy wartości funkcji w tych przedziałach, które sumujemy oraz mnożymy razy wysokość równą $\frac{x-0}{n}$, czyli długość przedziału podzieloną przez n (ilość równoodległych punktów). Ze wzoru na pole trapezu musimy otrzymany wynik podzielić jeszcze przez 2. Chciałbym jeszcze zaznaczyć, że w mojej implementacji tego rozwiązania n nie jest stałą, lecz daną możliwą do modyfikacji w wybrany przez użytkownika sposób. Umożliwia to obliczanie wartości dystrybuanty z podziałem na różną liczbę trapezów, co pozwala na liczenie z różnymi dokładnościami i porównywanie uzyskanych wyników. W pliku z rozwiązaniem zamieściłem kilka testów, których wyniki były zweryfikowane przez następujący kalkulator danego w treści zadania rozkładu funkcji chi-kwadrat: https://homepage.divms.uiowa.edu/~mbognar/applets/chisq.html