Wartość oczekiwana zawężonego rozkładu t-Studenta

Zmienna losowa R ma niestandardowy rozkład t-Studenta zawężony do przedziału $(\alpha;\beta)$, co zapisujemy $R \sim Tt_{(\alpha;\beta)}(\mu,\sigma^2;v)$, gdy jej gęstość prawdopodobieństwa jest opisana wzorem:

$$g(x) = \sigma^{-1} f_v\left(\frac{x-\mu}{\sigma}\right)\left(F_v\left(\frac{\beta-\mu}{\sigma}\right) - F_v\left(\frac{\alpha-\mu}{\sigma}\right)\right)^{-1}, \quad \alpha < x < \beta. \tag{1}$$

We wzorze (1) $f_v(\cdot)$ i $F_v(\cdot)$ to odpowiednio gęstość prawdopodobieństwa i dystrybuanta standardowego rozkładu t-Studenta $t(0,1;v)$ z v stopniami swobody. Gdy $\alpha = -\infty$ i $\beta = +\infty$, wzór (1) opisuje gęstość niestandardowego rozkładu t-Studenta na niezawężonym nośniku.

Wartość oczekiwana zmiennej losowej R może być zapisana jako

$$\mathbb{E}(R) = \int_\alpha^\beta x g(x)\, dx \quad \text{dla } v > 1$$

lub równoważnie

$$\mathbb{E}(R) = \mu + \sigma \frac{\Gamma((v-1)/2)((v+a^2)^{-(v-1)/2} - (v+b^2)^{-(v-1)/2})v^{v/2}}{2(F_v(b) - F_v(a))\Gamma(v/2)\Gamma(1/2)} \quad \text{dla } v > 1. \tag{2}$$

We wzorze (2) $\Gamma(\cdot)$ to funkcja gamma Eulera, $a = (\alpha-\mu)/\sigma$, $b = (\beta-\mu)/\sigma$. Funkcje $\Gamma(\cdot)$ i $F_v(\cdot)$ są dostępne w pakiecie R jako `gamma` i `pt`, w MATLAB-ie jako `gamma` i `tcdf`.

Przykład. Wektor losowy $\mathbf{R}$ opisuje 3-wymiarowy rozkład t-Studenta z 4 stopniami swobody, którego wartości składowych zostały zawężone do przedziału $[20;50]$. Parametry $\boldsymbol{\mu}$ oraz $\boldsymbol{\Sigma}$ niezawężonego rozkładu t-Studenta są następujące:

$$\boldsymbol{\mu} = \begin{pmatrix} 45 \\ 35 \\ 40 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & -2 & -1 \\ -2 & 36 & -8 \\ -1 & -8 & 9 \end{pmatrix}.$$

Należy wyznaczyć wartości oczekiwane zawężonych rozkładów brzegowych.

Rozwiązanie. Rozkład t-Studenta jest ciągły, więc wartość oczekiwana na przedziale domkniętym jest taka sama jak na przedziale otwartym. Z warunków zadania mamy $R_1 \sim Tt_{(20;50)}(45,1;4)$, $R_2 \sim Tt_{(20;50)}(35,36;4)$ i $R_3 \sim Tt_{(20;50)}(40,9;4)$. Po podstawieniu danych do (2) dostajemy:

$$\mathbb{E}(R_1) = 45 + 1 \cdot \frac{\Gamma(3/2)((4+(-25)^2)^{-3/2} - (4+5^2)^{-3/2}) \cdot 4^2}{2(F_4(5) - F_4(-25))\Gamma(2)\Gamma(1/2)} \approx 44{,}97,$$

$$\mathbb{E}(R_2) = 35 + 6 \cdot \frac{\Gamma(3/2)((4+(-5/2)^2)^{-3/2} - (4+(5/2)^2)^{-3/2}) \cdot 4^2}{2(F_4(5/2) - F_4(-5/2))\Gamma(2)\Gamma(1/2)} = 35,$$

$$\mathbb{E}(R_3) = 40 + 3 \cdot \frac{\Gamma(3/2)((4+(-20/3)^2)^{-3/2} - (4+(10/3)^2)^{-3/2}) \cdot 4^2}{2(F_4(10/3) - F_4(-20/3))\Gamma(2)\Gamma(1/2)} \approx 39{,}83.$$