Rachunek prawdopodobieństwa i statystyka

Lista zadań nr 14. 5 i 8 czerwca 2017

[Do zadań 1–2] Niezależne obserwacje x_{ij} pochodzą z tego samego rozkładu $N\left(\mu,\sigma^2\right)$, $(i = 1,\ldots,I;\ j = 1,\ldots,J)$.

1. **(2 pkt)**. Udowodnić, że prawdziwe jest równanie:
$\sum_{i,j}(x_{ij}-\bar{x})^2 = J\sum_i (x_{i\bullet}-\bar{x})^2 + I\sum_j (x_{\bullet j}-\bar{x})^2 + \sum_{ij}(x_{ij}-x_{i\bullet}-x_{\bullet j}-\bar{x})^2.$
2. Jakie rozkłady można wskazać w powyższym równaniu?

[Do zadań 3–4] Zmienna losowa X podlega rozkładowi normalnemu z parametrami jak poniżej:

$$N \sim \left(\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 16 & -3 \\ -3 & 9 \end{bmatrix}\right).$$

3. Niech $Y_1 = 3X_1 + 2X_2,\ Y_2 = 3X_1 + 4X_2$. Znaleźć rozkład zmiennej Y.
4. Niech $Z_1 = X_1 + 4,\ Z_2 = X_2 - 3$. Znaleźć rozkład zmiennej Z.
5. Zmienna losowa (X,Y) ma gęstość $f(x,y)$. Zmienne X oraz Y są niezależne. Wykazać, że $\mathrm{E}(X\cdot Y) = \mathrm{E}(X)\cdot\mathrm{E}(Y)$.
6. **(2 pkt)**. Gęstość zmiennej (X,Y) to $f(x,y) = 6xy$, dla $0 < y < 2-2x < 1$. Znaleźć postać dystrybuanty $F(x,y)$.

Plik temperatura.csv zawiera kolejno: średnie temperatury miesięczne, średnią temperaturę roczną, amplitudę roczną, szerokość i długość geograficzną, wysokość (npm), miesięczne opady, roczna suma opadów.

7. **E1** Jako zmienne niezależne (X_i) wybieramy szerokość i długość geograficzną oraz wysokość. Która ze zmiennych: amplituda, suma opadów jest lepiej objaśniana przez wybrane trzy zmienne?
8. **E1** ANOVA 1-czynnikowa. Czy prawdą jest, że: (a) opady nie zależą od miesiąca. (b) opady w okresie VI–IX nie zależą od miesiąca.
9. **E2** Skalujemy dane i znajdujemy 2 kierunki główne. Wymienić 4 zmienne i 4 miasta które mają największy wpływ na wspomniane 2 kierunki główne.

Witold Karczewski