Assessed Coursework for MA 413 – Statistics for Data Science

This sheet will take the place of the midterm exam. It is due for submission in moodle on October 21st (before midnight Central European Time). The handed in coursework must be your own work and not copied from a fellow student. You must include a signed statement stating "This is all my own work".

1. Assume you observe $Y_1, \ldots, Y_n$ independently from the distribution (for $c_1 \in \mathbb{R}^+$ and $0 < \theta < 1$)

$$f_Y(y) = c_1 \begin{cases} -\theta(y-0.5)^2 + 1 & \text{if} \quad 0 < y < 1 \\ 0 & \text{o/w} \end{cases} . \tag{1}$$

   (a) Determine the value of c_1.

   (b) Form the likelihood of the random sample $Y_1, \ldots, Y_n$.

   (c) Determine the maximum likelihood estimator of θ.

   (d) Evaluate the maximum likelihood estimate of θ if we observe $y_1 = 0.0114$, $y_2 = 0.0134$, $y_3 = 0.0254$, $y_4 = 0.5371$, $y_5 = 0.8721$, $y_6 = 0.9123$ and $y_7 = 0.9912$.

2. Consider the sets A_1 and A_2 that are all subsets of Ω, and that they are independent. Assume that $\Pr(A_1) = 0.05$ and $\Pr(A_2) = 0.1$. Compute $\Pr(A_1^c)$ and $\Pr(A_2^c)$. What property do A_1^c and A_2^c have?

3. Consider the set $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Define the subsets $A = \{1, 3, 5, 7, 9\}$, $B = \{2, 4, 6, 8\}$, $C = \{3, 5\}$, $D = \{2, 4, 6, 8\}$.

   (a) Determine $A \cup B$.

   (b) Determine $A \cap B$.

   (c) Determine A^c.

   (d) Determine $B \cap B$.

   (e) Determine $B \cap C$.

4. Let X and Y be generated from the joint distribution of

$$f_{X,Y}(x, y) = c_2 \begin{cases} xy & \text{if} \quad 0 < x < y < 1 \\ 0 & \text{o/w} \end{cases} . \tag{2}$$

   (a) Determine the value of c_2.

   (b) Determine the marginal density of X. Determine the marginal density of Y.

   (c) Determine the conditional density of $X|Y$ and of $Y|X$.

   (d) Are X and Y independent? Justify your answer.

5. Assume that $X_1 \sim N(\mu_1, \sigma_1^2)$ and that $X_2 \sim N(\mu_2, \sigma_2^2)$, and that the two random variables are independent.

   (a) Using MGFs or otherwise determine the distribution of $X_1 + X_2$.

   (b) Using MGFs or otherwise determine the distribution of $X_1 - X_2$.

   (c) Determine the distribution of $(X_1 + X_2)/(X_1 - X_2)$.

6. Assume you observe $Y_1, \ldots, Y_n$ all independent exponential random variables with mean unity (1).

   (a) Determine the distribution of $U = \max_i Y_i$ and $V = \min_i Y_i$ from first principles.

   (b) Determine the distribution of the mean of Y_i.

   (c) Assume we observe $X_i = \lambda Y_i$. Use your answer to (a) and (b) to propose an estimator of λ.

7. Assume you observe $Y_1, \ldots, Y_n$ all independent uniform random variables, uniform on $(0, 1)$.

   (a) Using the law of large numbers, or otherwise, determine the approximate distribution of $\overline{Y} = (1/n)(Y_1 + \cdots + Y_n)$ as $n >> 1$.

   (b) Determine the approximate distribution of $h(\overline{Y})$ for $h(\cdot)$ a continuous function.