ZMUM: laboratorium 6

1. Generate 5 datasets using the following schemes.
   (a) Example 1:
      - Sample size $n = 1000$, number of features $p = 50$,
      - $X_1 \sim U(0, 4)$, $X_2, \ldots, X_{50} \sim N(0, 1)$, $\epsilon \sim N(0, 0.1)$,
      - Target variable $Y = \sqrt{X_1} + \epsilon$.

   (b) Example 2:
      - Sample size $n = 1000$, number of features $p = 50$,
      - $X_1 \sim U(0, 4)$, $X_2, \ldots, X_{50} \sim N(0, 1)$, $\epsilon \sim N(0, 0.1)$,
      - Target variable $Y = X_1^2 + \epsilon$.

   (c) Example 3:
      - Sample size $n = 1000$, number of features $p = 50$,
      - $X_1, \ldots, X_{50} \sim N(0, 1)$, $\epsilon \sim N(0, 0.1)$,
      - Target variable $Y = (X_1 - 0)_+ + (X_1 - 1)_+ + \epsilon$.

   (d) Example 4:
      - Sample size $n = 1000$, number of features $p = 50$,
      - $X_1 \sim U(0, 4)$, $X_2, \ldots, X_{50} \sim N(0, 1)$, $\epsilon \sim N(0, 0.1)$,
      - Target variable $Y = \sin(X_1) + \epsilon$.

   (e) Example 5:
      - Sample size $n = 1000$, number of features $p = 50$,
      - $X_1, \ldots, X_{50} \sim N(0, 1)$, $\epsilon \sim N(0, 0.1)$,
      - Target variable $Y = I(X_1 < 0)$.

2. Run MARS (Multivariate Adaptive Regression Splines) method.

3. Prepare a scatter plot showing the dependence between X_1 and Y. Then add to the scatter plot a curve showing predicted values from MARS method.