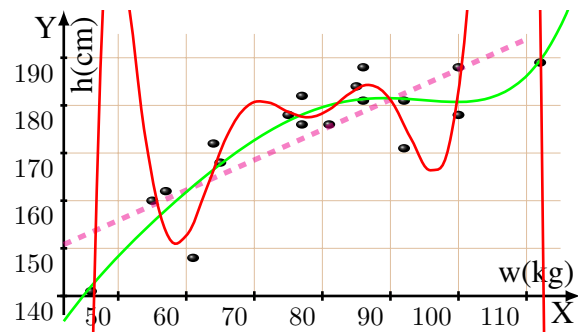


**Problem:**

We discussed estimates of the relationship between height and weight of people based on empirical data and various fits to the data:

H	148	171	168	160	178	162
W	61	92	65	55	75	57
H	188	141	182	184	189	176
W	100	46	77	85	112	77
H	188	172	178	181	181	176
W	86	64	100	92	86	81



The plots show the *generalised linear models* using the following bases and using the maximum likelihood estimation:

- linear  $\Phi = [1, x]$ ;
- cubic  $\Phi = [1, x, x^2, x^3]$ ;
- ninth order polynomial  $\Phi = [1, x, \dots, x^9]$ ;

Using the maximum likelihood estimation, **assess** the robustness of the general polynomial families. Consider the family of polynomials having order  $K$  as:  $f(x) = \sum_{k=1}^K \theta_k (x - 80)^k$ .

Implement the following procedure:

- choose a random subset of size 12 from the data.
- find the M.L. estimates for the subset.
- measure the sum of squares quadratic errors on the subset **not in the training set** – *i.e.* the complementary or left-out set.

Repeat the above procedure 100 times and for each repetition measure the error on (1) the training data and also on (2) the test data.

Report the results above experiments – as bar-plots, or in the format you will explain – for the  $K = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$ . The graph should contain **both** the training as well as the test errors.