(1) Learning probability distributions
    (a) Approximating distributions from data
        (i) empirical distributions
        (ii) Kernel smoothing
    (b) Learning parameters of distributions
        (i) Estimators. Vaiance and bias.
        (ii) Maximum likelihood estimation
        (iii) Bayesian perspective