Instructions

- The homework is due on **Friday 3/17 at 5pm ET**.
- No extension will be provided, unless for serious documented reasons.
- **Despite having two weeks for this HW, better start early than late!**
- Study the material taught in class, and feel free to do so in small groups, but the solutions should be a product of your own work.
- This is not a multiple choice homework; reasoning, and mathematical proofs are required before giving your final answer.

1 MLE and MoM [30 points]

1. (5pts) Let $X_1, \ldots, X_n$ be iid Bernoulli(p) samples. In class we sketched the proof that the maximum likelihood estimator of p is $p_{MLE} = \frac{\sum_{i=1}^n x_i}{n}$. Write a complete proof.
2. (5pts) Assume you have a prior p that is a $beta(\alpha, \beta)$ and let $Y = \sum_{i=1}^n X_i$. Write down the joint distribution of Y, p.
3. (5pts) Let $X_1, \ldots, X_n$ be iid $\mathcal{N}(\mu, \sigma^2)$ where both μ, σ are unknown. What are the MLEs for μ, σ^2?
4. (5pts) Let $X_1, \ldots, X_n$ be iid $Exponential(\lambda)$. Find the method of moments estimator for λ.
5. (5pts+5pts) Let $X_1, \ldots, X_n$ be iid $\beta(\theta, 1)$. Find (a) the MLE and the (b) MoM estimator for θ.

   Hint: You may use the fact that the expected value of a $beta(\alpha, \beta)$ is equal to $\frac{\alpha}{\alpha+\beta}$ without proof.

2 To Handshake or Not? [20 points]

Suppose n people walk into a party. Due to covid-19, each pair $\{i, j\}$ shakes hands with probability only $\frac{1}{10}$. Prove that with probability that tends to 1 as $n \to +\infty$ **every** person from that party shook hands in the range $[0.95\frac{n}{10}, 1.05\frac{n}{10}]$.

3 Mixture of Gaussians [25 points]

Let X, Y be two independent normal RVs, with means $\mu_x = 100, \mu_y = 300$ and standard deviations $\sigma_x = \sigma_y = 10$. Consider the RV U defined by

$$U = \frac{1}{2}(X + Y).$$

Alternatively, consider the RV Z that is generated as follows:

(a) With probability $\frac{1}{2}$ we sample Z from $N(\mu = 100, \sigma^2 = 100)$.

(b) With probability $\frac{1}{2}$ we sample Z from $N(\mu = 300, \sigma^2 = 100)$.

1. **[5 points]** Simulate the sampling, and produce two histograms (one for U and one for Z) over 10 000 samples for each U, Z.
2. **[10 points]** Compute the expected values of U, Z.
3. **[10 points]** Compute the variances of U, Z.

4 Coding EM for Mixture of Gaussians [25 points]

Check the Jupyter notebook on our Git repo.