SDS 383C: Statistical Modeling I, Fall 2022
Homework 2, Due Oct 05, 12:00 Noon

Instructor: Abhra Sarkar (abhra.sarkar@utexas.edu)
Teaching Assistant: Preston Biro (prestonbiro@utexas.edu)
Department of Statistics and Data Sciences
The University of Texas at Austin
2317 Speedway D9800, Austin, TX 78712-1823, USA

All homework must be submitted typed-in as a single pdf file. Name the file "firstname-lastname-SDS383C-HW-2.pdf" Submit this file without compression such as zip or rar. Figures accompanying the solutions must be presented close to the actual solution. Computer codes will be rarely evaluated but must still be submitted separately from the main file. Codes must be commented properly and should run easily on other machines. Precise, concise, clear, innovative solutions may be rewarded with bonus points. Explain your answer with logic reasoning and/or mathematical proofs. Organize your solutions in the same order as they were presented. If you can solve a problem using multiple techniques, present only your best solution.

1. (10 points) For $y_1, \ldots, y_n \overset{iid}{\sim}$ Normal(μ, σ^2), show that $\mathbf{s} = (\overline{y}, s^2)$ are sufficient. Also, find the observed and the expected information matrices.

2. (10 points) From AC Davison's Statistical Models, do problem 4.4.6.

   [Hints: For the last part, use $z_i = y_i/\mu \sim$ Normal$(1, c)$.]

3. (30 points) Consider a Cauchy(θ, γ) distribution with unknown location θ but known scale $\gamma = 1$. Recall that a Cauchy(θ, γ) distribution has pdf

$$f(y \mid \theta, \gamma) = \frac{1}{\pi\gamma} \frac{\gamma^2}{\{\gamma^2 + (y-\theta)^2\}}, \qquad y \in \mathbb{R}, \theta \in \mathbb{R}, \gamma \in \mathbb{R}^+.$$

   The distribution is symmetric around θ but, due to heavy tails, the moments $\mathbb{E}\,|x|^r$ do not exist for any $r \geq 1$. So WLLN does not apply and $\overline{y}_n \not\to \theta$ in probability. Also, CLT does not apply. In fact, $\overline{y}_n$ also follows Cauchy$(\theta, 1)$ for all n.

   (a) Draw a random sample of size $n = 20$ from a Cauchy$(0, 1)$ distribution. Compute $\overline{y}_n$. Repeat this $B = 1000$ times. Draw a histogram of the sampled $\overline{y}_n$ values, superimposed over a Cauchy$(0, 1)$ density.

   Although we cannot directly apply MoM to estimate θ, we can still use maximum likelihood to estimate θ. Unfortunately, the maximum likelihood equation does not result in a closed form solution for θ, so we need to use numerical approximation to find out $\widehat{\theta}_{MLE}$. Consider now a random sample, drawn from a Cauchy$(\theta, 1)$ distribution, given below.

7.52	9.92	9.52	21.97	8.39	8.09	9.22	9.37	7.33	15.32
1.08	8.51	17.73	11.20	8.33	10.83	12.40	14.49	9.44	3.67

   Using this data points, compute $\widehat{\theta}_{MLE}$ using

(b) step-wise gradient ascent,
(c) Newton-Raphson, and
(d) stochastic gradient ascent.

Supplement your analysis with appropriate plots.

4. (10 points) Let $y_{i,j} \stackrel{ind}{\sim} \text{Normal}(\mu_i, \sigma^2)$, $i = 1, \ldots, n$ and $j = 1, \ldots, m$ with $n \to \infty$ but m fixed. Find out the MLEs of μ_i and σ^2. Show that $\widehat{\sigma}^2_{MLE}$ is inconsistent for σ^2. Adjust this estimator to come up with a consistent estimator of σ^2.

5. (10 points) For $y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Uniform}(0, \theta)$, show that $z_n = \frac{n(\theta - \widehat{\theta}_{MLE})}{\theta} \xrightarrow{D} E$ where E is exponential. Design a simulation study illustrating this nice result.

6. (30 points) Consider the spring failure data from Davison, providing the failure times at stress 950N/mm^2 in 10^3 cycles:

$$225, 171, 198, 189, 189, 135, 162, 135, 117, 162.$$

Using the method of maximum likelihood, fit a Weibull distribution to these data [use the parametrization considered in class]. Also find an approximate 90% joint confidence region for the parameters (k, λ). Supplement your analysis with appropriate plots.

7. (20 points) (computationally expensive bonus problem) For a Weibull(k, λ) distribution [use the parametrization considered in class with $\boldsymbol{\theta} = (k, \lambda)^{\mathrm{T}}$], design a simulation study to show that $(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^{\mathrm{T}} J(\widehat{\boldsymbol{\theta}})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} \chi^2_2$ and $W(\boldsymbol{\theta}) \xrightarrow{D} \chi^2_2$.

8. Read pages 127-131 (Section 4.5.2: Profile log likelihood) from AC Davison's Statistical Models.

9. Read pages 150-155 (Section 4.7: Model Selection) from AC Davison's Statistical Models, paying special attention to Example 4.47.