This problem set covers material from Week 4, dates 3/05- 3/08. Unless otherwise noted, all problems are taken from the textbook. Problems can be found at the end of the corresponding subsection. "AP" stands for additional problems not found in the book.

Instructions: Write or type complete solutions to the following problems and submit answers to the corresponding Canvas assignment. Your solutions should be neatly-written, show all work and computations, include figures or graphs where appropriate, and include some written explanation of your method or process (enough that I can understand your reasoning without having to guess or make assumptions). A general rubric for homework problems appears on the final page of this assignment.

Tuesday 3/05

1. Let $X_1, \ldots, X_n | \boldsymbol{\theta} \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$ where $\boldsymbol{\theta} = (\mu, \sigma^2)$. Find the method of moments estimator for $\boldsymbol{\theta}$ (simplify as much as possible), and compare it to $\hat{\boldsymbol{\theta}}_{MLE}$.
   In your comparison, it might be worth remembering the following identity you showed on HW 2: if $X_1, \ldots, X_n iid$, then $\sum_{i=1}^n (X_i - a)^2 = \sum_{i=1}^n (X_i - \bar{X})^2 + n(\bar{X} - a)^2$ for any $a \in \mathbb{R}$.

2. Let $X_1, \ldots, X_n | \theta \overset{\text{iid}}{\sim} f(x|\theta) = \frac{1}{2\theta} e^{-\frac{|x|}{\theta}}$ for $x \in \mathbb{R}$ and unknown parameter $\theta > 0$. Find a method of moments estimator for θ. It may be useful to remember the following facts:
   - If ϕ is a continuous even function, then $\int_{-a}^{a} \phi(x)dx = 2\int_0^a \phi(x)dx$
   - If ϕ is a continuous odd function, then $\int_{-a}^{a} \phi(x)dx = 0$

Thursday 3/07

3. Let $X_1, \ldots, X_n | \theta \overset{\text{iid}}{\sim} \text{Exp}(\theta)$. We found that that the MLE and MoM estimators for θ are the same: $\hat{\theta} = \frac{1}{\bar{X}} = \frac{n}{\sum X_i}$. Let us try to say something about the bias of this estimator for θ. The WRONG thing to do is to say $\mathbb{E}[\hat{\theta}] = \frac{1}{\mathbb{E}[\bar{X}]}$; remember the function of the expectation and the expectation of the function are rarely equal. Instead, let's try to use the following important theorem:

   Jensen's inequality for probability theory states that if X is a random variable and $g()$ a convex function, then $g(\mathbb{E}[X]) \leq \mathbb{E}[g(X)]$.

   Remember that a convex function is one where a linear segment between any two distinct points on the graph of the function always lies above the graph between the two points. We can demonstrate convexity of a univariate function $g(x)$ on an interval I if the second derivative $g''(x) > 0$ for all $x \in I$.

   Using Jensen's inequality, what can you say about the bias of $\hat{\theta}_{MLE}$ for θ in this problem? Does the MLE tend to overestimate or underestimate θ, or is it unbiased?

4. For $X_1, \ldots, X_n|\theta \overset{\text{iid}}{\sim} \text{Unif}[0, \theta]$, we found the estimators of θ under the method of moments and the method of maximum likelihood to be:

$$\hat{\theta}_{MM} = 2\bar{X} \qquad \text{and} \qquad \hat{\theta}_{MLE} = \max\{X_i\}$$

   (a) Find the bias of $\hat{\theta}_{MM}$ for θ.

   (b) Find the bias of $\hat{\theta}_{MLE}$ for θ.

   (c) For the estimators that you found to biased in (a) and/or (b), propose a new *unbiased* estimator $\tilde{\theta}_a$ and/or $\tilde{\theta}_b$ that is a modification of the estimator from (a) and/or (b), respectively.

5. Suppose we have a single $X|\theta \sim \text{Poisson}(\theta)$, and we are interested in estimating $g(\theta) = P(X = 0)^2 = e^{-2\theta}$.

   (a) Find an unbiased estimator of $g(\theta)$. *Hint: LoTUS and* $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.

   (b) Comment on at least two reasons why the unbiased estimator in (a) is not very useful.

General rubric

Points	Criteria
5	The solution is correct *and* well-written. The author leaves no doubt as to why the solution is valid.
4.5	The solution is well-written, and is correct except for some minor arithmetic or calculation mistake.
4	The solution is technically correct, but author has omitted some key justification for why the solution is valid. Alternatively, the solution is well-written, but is missing a small, but essential component.
3	The solution is well-written, but either overlooks a significant component of the problem or makes a significant mistake. Alternatively, in a multi-part problem, a majority of the solutions are correct and well-written, but one part is missing or is significantly incorrect.
2	The solution is either correct but not adequately written, or it is adequately written but overlooks a significant component of the problem or makes a significant mistake.
1	The solution is rudimentary, but contains some relevant ideas. Alternatively, the solution briefly indicates the correct answer, but provides no further justification.
0	Either the solution is missing entirely, or the author makes no non-trivial progress toward a solution (i.e. just writes the statement of the problem and/or restates given information).
Notes:	For problems with multiple parts, the score represents a holistic review of the entire problem. Additionally, half-points may be used if the solution falls between two point values above.
Notes:	For problems with code, well-written means only having lines of code that are necessary to solving the problem, as well as presenting the solution for the reader to easily see. It might also be worth adding comments to your code.