STA261 (Summer 2022) - Assignment 2

These problems are meant to test your understanding of the concepts in Module 2. They are *not* to be handed in. Some of these have been modified (or in some cases taken directly) from questions in the *Additional Resources* listed in the course syllabus, and no claims of originality are made.

---

1. Suppose $\mathcal{X} = \{1, 2, 3, 4\}$ and $\Theta = \{a, b\}$. Two mass functions on $\mathcal{X}$ – one for each value of $\theta \in \Theta$ – are specified in the following table:

	$x = 1$	$x = 2$	$x = 3$	$x = 4$
$p_a(x)$	$1/2$	$1/6$	$1/6$	$1/6$
$p_b(x)$	$1/3$	$1/3$	$1/3$	0

   Suppose $X \sim p_\theta$. Find the MLE $\hat{\theta}(X)$, and then calculate $\mathbb{E}_\theta\left[\hat{\theta}(X)\right]$ for each $\theta \in \Theta$.

2. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Geom}(p)$, where $p \in (0, 1]$. Find the MOM estimator and the MLE of p.

3. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Gamma}(\alpha, \beta)$ with $\beta > 0$ and α known. Find the MOM estimator and the MLE of β.

4. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\theta, 1)$, where $\theta \in \mathbb{R}$. Find the MOM estimator and the MLE of $\tau(\theta) = \mathbb{P}_\theta(X_i \leq 1)$.

5. Let $X_1, X_2, \ldots, X_n$ be a random sample from an inverse Gaussian distribution, which has density

$$f_{\mu,\lambda}(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right), \quad x \in \mathbb{R}, \quad \lambda > 0, \quad \mu \in \mathbb{R}.$$

   Find the MOM estimator of (μ, λ), and then assume λ is known and find the MLE of μ. For the MOM, you can freely use the facts that $\mathbb{E}[X] = \mu$ and $\text{Var}(X) = \mu^3/\lambda$ (or you can show these yourself, if you're really into integration by parts). You can also find the MLE of (μ, λ) itself, but it's pretty tedious.

6. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Lognormal(μ, σ^2) distribution, which has density

$$f_{\mu,\sigma}(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log(x) - \mu)^2}{2\sigma^2}\right), \quad x > 0, \quad \mu \in \mathbb{R}, \quad \sigma > 0.$$

   Find the MLE of σ^2 assuming μ is known, and the MLE of μ assuming σ^2 is known. If both parameters are unknown, can you guess the MLE of (μ, σ^2) just from inspection?

7. For the same Lognormal(μ, σ^2) distribution as above, find the MOM estimator of (μ, σ^2). There's a really nifty trick for easily finding $\mathbb{E}\left[X^t\right]$ for any $t > 0$: check that $Y = \log(X) \sim \mathcal{N}\left(\mu, \sigma^2\right)$, and think about Normal mgfs.

8. Example 2.17 was a particular case of *linear regression*, which is probably the single most widely used statistical method out there. To keep things simple, we'll stick to *simple linear regression* in this exercise. The formulation is reasonable: for each $i = 1, \ldots, n$, we imagine that some real-life quantity z_i is linearly related to some other real-life quantity x_i, in the sense that

$$z_i = \alpha + \beta x_i, \quad i = 1, \ldots, n,$$

for some fixed $\alpha, \beta \in \mathbb{R}$. For example, if z_i is the height of person i and x_i is their weight, this is saying that everyone's height is the same linear function of their weight. Now, when we collect our data on people's heights and weights, our measurements aren't perfect – we assume the data is noisy and there's some kind of random measurement error. Thus, we don't actually observe z_i. Instead, we independently observe $Y_i = y_i$ for $i = 1, \ldots, n$, where

$$Y_i = \alpha + \beta x_i + \epsilon_i, \quad i = 1, \ldots, n,$$

where $\epsilon_1, \epsilon_2, \ldots, \epsilon_n \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right)$. This is the simple linear regression model. In statistics jargon, the y_i's are the **response variable**, and each x_i is a **covariate**. α is the **intercept**, β is the **slope**, and α and β are collectively called the **regression coefficients**, while ϵ_i is called a **random error**.[1] Our goal here is to say something interesting about the MLEs of α and β.

(a) Explain why $Y_i \sim \mathcal{N}\left(\alpha + \beta x_i, \sigma^2\right)$ and the Y_i's are independent.

(b) Show that finding the MLE of (α, β) is equivalent to finding the (α, β) that minimizes $g(\alpha, \beta) := \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$, and interpret this quantity geometrically.

(c) First fix β, and show that the value of α that minimizes $g(\alpha, \beta)$ is $\tilde{\alpha} = \bar{y} - \beta \bar{x}$.

(d) Now let β vary, and show that the value of β that minimizes $g(\tilde{\alpha}, \beta)$ is

$$\hat{\beta} = \frac{\sum_{i=1}^n (\bar{y} - y_i)(\bar{x} - x_i)}{\sum_{i=1}^n (\bar{x} - x_i)^2}.$$

It follows that if we let $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$, then $g(\alpha, \beta)$ is minimized at $(\hat{\alpha}, \hat{\beta})$. This is the MLE!

(e) Show that $\hat{\alpha}(\mathbf{Y})$ is unbiased for α, and that $\hat{\beta}(\mathbf{Y})$ is unbiased for β.

(f) Think about exponential families and show that

$$T(\mathbf{Y}) = \left(\sum_{i=1}^n Y_i^2, \sum_{i=1}^n Y_i, \sum_{i=1}^n Y_i x_i \right)$$

is a complete sufficient statistic for $(\alpha, \beta, \sigma^2)$.

(g) Show that both $\hat{\alpha}(\mathbf{Y})$ and $\hat{\beta}(\mathbf{Y})$ are functions of $T(\mathbf{Y})$.

(h) Explain why we can conclude that $\hat{\alpha}(\mathbf{Y})$ and $\hat{\beta}(\mathbf{Y})$ are the UMVUEs of α and β, respectively.

9. As we saw in Example 2.16, it's not always possible to find the MLE by differentiating the log-likelihood. That doesn't mean the MLE doesn't exist, however; it just means that calculus won't help us find it, so we need to resort to other methods.

(a) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Unif}\,(\theta_1, \theta_2)$, where $-\infty < \theta_1 < \theta_2 < \infty$. Find the MLE of (θ_1, θ_2).

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from a Laplace(μ) distribution, which has density

$$f_\mu(x) = \frac{1}{2} e^{-|x-\mu|}, \quad x \in \mathbb{R}, \qquad \mu \in \mathbb{R}.$$

Find <u>an</u> MLE of μ. This one is tough. Here are two hints from two different textbooks, plus a third from me. *Hint 1*: Maximize the log-likelihood in each of the intervals $(-\infty, x_{(1)})$, $[x_{(1)}, x_{(2)})$, etc. *Hint 2:* Consider the case of even n separate from that of odd n, and find the MLE in terms of the order statistics. *Hint 3:* $\sum_i |x_i - \mu| = \sum_i |x_{(i)} - \mu|$.

[1] These terms each have about 10 different commonly-used synonyms. No one can agree on what to call anything.

10. Suppose that $T_1(\mathbf{X})$ and $T_2(\mathbf{X})$ are two unbiased estimators of $\tau(\theta) \in \mathbb{R}$, with $\operatorname{Var}_\theta(T_1(\mathbf{X})) = \sigma_1^2$ and $\operatorname{Var}_\theta(T_2(\mathbf{X})) = \sigma_2^2$. Let $\alpha \in [0,1]$.

    (a) Show that $T(\mathbf{X}) := \alpha T_1(\mathbf{X}) + (1-\alpha)T_2(\mathbf{X})$ is unbiased for $\tau(\theta)$.

    (b) Assuming $T_1(\mathbf{X})$ and $T_2(\mathbf{X})$ are independent, find the α that minimizes $\operatorname{Var}_\theta(T(\mathbf{X}))$.

    (c) Do the same thing, but this time without assuming independence. If you want, you can call $\rho = \operatorname{Corr}_\theta(T_1(\mathbf{X}), T_2(\mathbf{X}))$.

11. Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with mean μ and variance σ^2. Let $a_1, a_2, \ldots, a_n \in \mathbb{R}$.

    (a) Suppose that $\sum_{i=1}^n a_i = 1$. Show that $T(\mathbf{X}) = \sum_{i=1}^n a_i X_i$ is unbiased for μ.

    (b) Show that finding the a_i's that minimize $\operatorname{Var}(T(\mathbf{X}))$ is the same as finding the a_i's that minimize $\sum_{i=1}^n a_i^2$ (all subject to $\sum_{i=1}^n a_i = 1$).

    (c) Show that $a_i = \frac{1}{n}$ for all i does the trick. There are plenty of ways to do this; you can bring out the Lagrange multipliers if you want, but it's much easier to just see what happens to $\sum_{i=1}^n a_i^2$ when you write $a_i^2 = ([a_i - \frac{1}{n}] + \frac{1}{n})^2$ and expand the square.

12. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$ where $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. Find the UMVUEs of $\mu^2 + \sigma^2$ and $\mu + \sigma^2$.

13. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$ where $\theta \in \mathbb{R}$ and σ^2 is known. Fix $a, b \in \mathbb{R}$. Find some $\tau(\theta)$ such that $a\bar{X}_n + b$ is the UMVUE of $\tau(\theta)$.

14. For each of the following densities, find some $\tau(\theta)$ such that the UMVUE of $\tau(\theta)$ exists.

    (a)
    $$f_\theta(x) = \theta x^{\theta-1}, \quad x \in (0,1), \quad \theta > 0.$$

    (b)
    $$f_\theta(x) = \frac{\log(\theta)}{\theta - 1}\theta^x, \quad x \in (0,1), \quad \theta > 1.$$

15. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \operatorname{Unif}(0,\theta)$ where $\theta > 0$. I'll give you that $T(\mathbf{X}) = X_{(n)}$ is a complete sufficient statistic.

    (a) Show that $T(\mathbf{X})$ is biased for θ, and correct that bias to find the UMVUE of θ, which you can call $U(\mathbf{X})$.

    (b) Show that $V(\mathbf{X}) = 2X_1$ is unbiased for θ, and explain why it can't be the UMVUE of θ.

    (c) Explain why $\mathbb{E}_\theta[V(\mathbf{X}) \mid T(\mathbf{X})]$ must be the exact same thing as $U(\mathbf{X})$, and explicitly show this. You can save yourself practically all the calculations by using Basu's theorem, similar to what we did in Example 1.41. Hint: why is $X_1/X_{(n)}$ independent of $X_{(n)}$, and why does that imply $\mathbb{E}_\theta\left[X_1/X_{(n)}\right] = \mathbb{E}_\theta[X_1]/\mathbb{E}_\theta\left[X_{(n)}\right]$?

16. Let $X_1, X_2, \ldots, X_n$ be a random sample from a continuous one-parameter exponential family of the form $f_\theta(x) = h(x)g(\theta)\exp(T(x)\cdot w(\theta))$, where $g(\cdot)$ and $w(\cdot)$ are differentiable on Θ. Let's find the UMVUE of $\tau(\theta) = -\frac{g'(\theta)}{w'(\theta)\cdot g(\theta)}$.

    (a) Explain why
    $$0 = \frac{\mathrm{d}}{\mathrm{d}\theta}\left(g(\theta)\int_{\mathcal{X}} h(x)e^{T(x)\cdot w(\theta)}\,\mathrm{d}x\right).$$

(b) Carry out the differentiation on the right, using product rule and the fact that $g'(\theta) = \frac{g'(\theta)}{g(\theta)} g(\theta)$ to find that

$$0 = w'(\theta) \int_{\mathcal{X}} T(x) f_\theta(x) \, \mathrm{d}x + \frac{g'(\theta)}{g(\theta)}.$$

(c) Find an unbiased estimator $U(\mathbf{X})$ of $\tau(\theta)$.

(d) Assuming the "open-interval" condition holds, explain why $U(\mathbf{X})$ is the UMVUE of $\tau(\theta)$.

17. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Exp}(\lambda)$, where $\lambda > 0$. Find an unbiased estimator of $1/\lambda$ based on $X_{(1)}$, and show that there's a better unbiased estimator of $1/\lambda$ out there with lower variance by computing the variances of each.

18. Prove Theorem 2.10.

19. Finish off the proof of Theorem 2.11.