CS 312 HW8

Rohan Mukherjee

May 30, 2023

1. (a) We recall that

$$\mathcal{L}(x_1, \dots, x_n;\ \theta_B, \theta_C) = \prod_{i=1}^{n} \mathbb{P}(x_i;\ \theta_B, \theta_C)$$

If x_i is a Bulbasaur,

$$\mathbb{P}(x_i;\ \theta_B, \theta_C) = \theta_B$$

And similarly, if x_i is a Charmander, $\mathbb{P}(x_i;\ \theta_B, \theta_C) = \theta_C$, and finally if x_i is a Squirtle, $\mathbb{P}(x_I;\ \theta_B, \theta_C) = 1 - \theta_B - \theta_C$. Now, since precisely n_B of the x_i's are equal to Bulbasaur, n_C are equal to Charmander, and n_S are equal to Squirtle, we can simplify this product to

$$\mathcal{L}(x_1, \dots, x_n;\ \theta_B, \theta_C) = \theta_B^{n_B} \cdot \theta_C^{n_C} \cdot (1 - \theta_B - \theta_C)^{n_S}$$

Since multiplication is commutative.

(b) The log of my previous answer is just

$$\log(\mathcal{L}(x_1, \dots, x_n;\ \theta_B, \theta_C)) = n_B \log(\theta_B) + n_C \log(\theta_C) + n_S \log(1 - \theta_B - \theta_C)$$

(c) Taking partials gives

$$\frac{\partial \log(\mathcal{L}(-))}{\partial \theta_B} = \frac{n_B}{\theta_B} - \frac{n_S}{1 - \theta_B - \theta_C}$$

And similarly,

$$\frac{\partial \log(\mathcal{L}(-))}{\partial \theta_C} = \frac{n_C}{\theta_C} - \frac{n_S}{1 - \theta_B - \theta_C}$$

We want each of these to be equal to 0 to find the critical point. Setting the first to 0 gives

$$\begin{aligned} & \frac{n_S}{1 - \theta_B - \theta_C} = \frac{n_B}{\theta_B} \\ \iff & n_S\theta_B = n_B(1 - \theta_B - \theta_C) \\ \iff & (n_S + n_B)\theta_B + n_B\theta_C = n_B \end{aligned}$$

Doing the exact same thing in the other variable gives

$$\begin{aligned} & n_S\theta_C = n_C(1-\theta_B-\theta_C) \\ \iff \ & n_C\theta_B + (n_S+n_C)\theta_C = n_C \end{aligned}$$

(d) Plugging this into wolfram gives

$$\theta_B = \frac{n_B}{n_B+n_C+n_S} \text{ and } \theta_C = \frac{n_C}{n_B+n_C+n_S}$$

Exactly what you might expect.

2. (a) The graph is a flipped absolute value shifted up a little bit and stretched. If theta is large the graph is flatter and wider, and if theta is small the graph is taller. It also is continuous at $x = \pm\theta$ with value 0, and zero at $|x| \geq \theta$ (i.e., the support of f is just $[-\theta, \theta]$).

   (b) We recall from part c that the likelihood is only nonzero when all samples have absolute value strictly smaller than θ, i.e. $\max\{|x_1|, \ldots, |x_n|\} < \theta$. If this is not the case, i.e. if there is an x_i with $|x_i| \geq \theta$, then the likelihood is just 0. So, in the case where $\max\{|x_1|, \ldots, |x_n|\} < \theta$, the likelihood equals

$$\mathcal{L}(x_1, \ldots, x_n;\ \theta) = \prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} \left(-\frac{|x_i|}{\theta^2} + \frac{1}{\theta}\right)$$

   (c) The likelihood in this case would be

$$f_X(x_1) \cdot \prod_{i=2}^{n} f_X(x)$$

   If $|x_1| > \theta$, then the density gives 0, so the product (= likelihood) is 0. If $|x_1| = 1$, the density would give

$$-\frac{\theta}{\theta^2} + \frac{1}{\theta} = 0$$

   So again the product would be 0, and hence the likelihood is 0.

   (d) Under the assumption that we have done two draws, with $x_1 = -x_2$, if $|x_1| \geq \theta$ (equivalently, $|x_2| \geq \theta$), then we know that the likelihood is just equivalently 0, so we could take for example $\theta = 1$ to be a maximizer, since constant functions are maximized (and minimized) everywhere. In the more interesting case that $|x_1| < \theta$,

$$\mathcal{L}(x_1, -x_1;\ \theta) = \prod_{i=1}^{2} \left(-\frac{|x_i|}{\theta^2} + \frac{1}{\theta}\right) = \left(-\frac{|x_1|}{\theta^2} + \frac{1}{\theta}\right)^2$$

   Taking the log of this quantity gives

$$\log(\mathcal{L}(-)) = 2\log\left(\frac{-|x_1|}{\theta^2} + \frac{1}{\theta}\right) = 2\log\left(\frac{-|x_1| + \theta}{\theta^2}\right) = 2(\log(-|x_1| + \theta) - 2\log(\theta))$$

   Taking derivatives now gives

$$\frac{\mathrm{d}\log(\mathcal{L}(-))}{\mathrm{d}\theta} = 2\left(\frac{1}{-|x_1| + \theta} - \frac{2}{\theta}\right)$$

   Setting this to zero gives

$$\begin{aligned} \frac{2}{\theta} &= \frac{1}{-|x_1| + \theta} \\ \iff 2\theta - 2|x_1| &= \theta \\ \iff \theta &= 2|x_1| \end{aligned}$$

3. (a) We recall that

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

We must find $f_X(x)$. We recall that this equals

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$

For $x \notin [0,3]$, the density is always zero, so $f_X(x) = 0$ there. For $x \in [0,3]$, this integrand is only not zero when $x \leq y \leq 3$, so we get

$$f_X(x) = \int_x^3 \frac{4}{27}x^2 dy = \frac{4}{27}x^2 \cdot y\Big|_x^3 = \frac{4}{27}x^2(3-x)$$

We conclude that

$$f_{Y|X}(y \mid x) = \frac{\frac{4}{27}x^2}{\frac{4}{27}x^2(3-x)} = \frac{1}{3-x}$$

(b) For $x \notin [0,3]$, this conditional probability isn't even defined, so I will say it is 0. So assume that $x \in [0,3]$.

$$\mathbb{E}[Y \mid X = x] = \int_{-\infty}^{\infty} y f_{Y|X}(y \mid x)dy$$

Once again the integrand is only not zero when $x \leq y \leq 3$, so this integral simplifies to

$$\int_x^3 y\frac{1}{3-x}dy = \frac{1}{3-x} \cdot \frac{y^2}{2}\Big|_x^3 = \frac{1}{2} \cdot \frac{1}{3-x}(9-x^2) = \frac{1}{2}(3+x)$$

(c) By the law of total expectation,

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} \mathbb{E}[Y \mid X = x] f_X(x)dx$$

The support of X is just $[0,3]$, so this integral simplifies to

$$\int_0^3 \mathbb{E}[Y \mid X = x] f_X(x)dx$$

Now we may use the results from part (a) and (b) to conclude that this integral equals

$$\int_0^3 \frac{1}{2}(3+x)\frac{4}{27}x^2(3-x)dx = \frac{4}{54}\int_0^3 x^2(9-x^2)dx = \frac{4}{54}\int_0^3 9x^2 - x^4 dx$$
$$= \frac{4}{54}\left(3x^3 - \frac{1}{5}x^5\right)\Bigg|_0^3 = \frac{4}{54}\left(3 \cdot 27 - \frac{1}{5}243\right) = \frac{12}{5}$$

4. Since $X_i \sim \text{Uniform}(0, \theta)$, $\mathbb{E}[X_i] = \theta/2$ from the zoo. Hence,

$$\mathbb{E}\left[\hat{\theta}\right] = \mathbb{E}\left[\frac{2}{n}\sum_{i=1}^{n} X_i\right] = \frac{2}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \frac{2}{n}\cdot n\cdot\theta/2 = \theta$$

So yes, this estimator is unbiased.