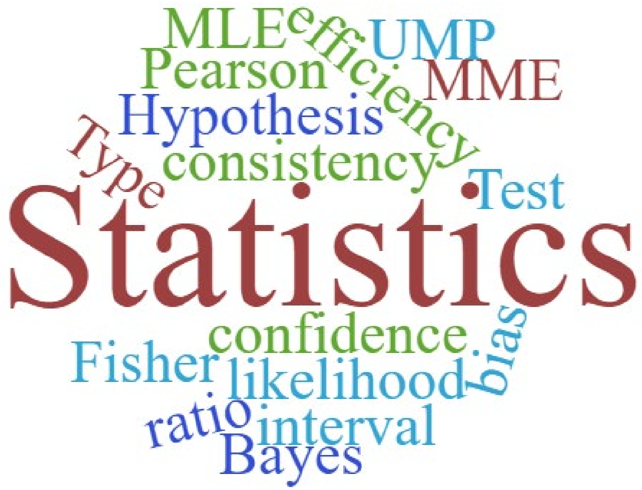
Math 181A: Mathematical Statistics



David Quarfoot
Associate Teaching Professor
Email: dquarfoot@ucsd.edu
Office: APM 5220

Learning Goals:

- Learn to deal with MLEs involving 1) complicated likelihoods (requiring indicator functions), 2) more than one variable, and 3) expressions for the estimate that do not have a closed form solution

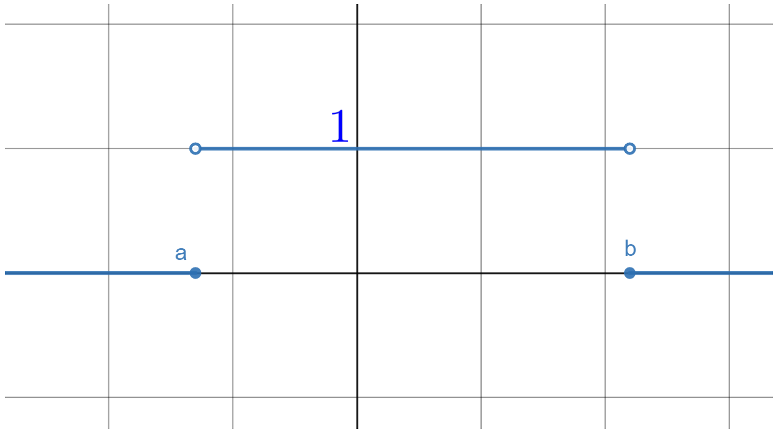
MLEs, Day 2: Harder Situations

Ways to step up the difficulty: 1) Difficult likelihood function, 2) More than one parameter, 3) No closed form solution.

Helpful notation for today:

“indicator function”

$$\text{Let } I_{(a,b)} = I_{(a,b)}(x) = \begin{cases} 1, & x \in (a,b) \\ 0, & \text{else} \end{cases}$$



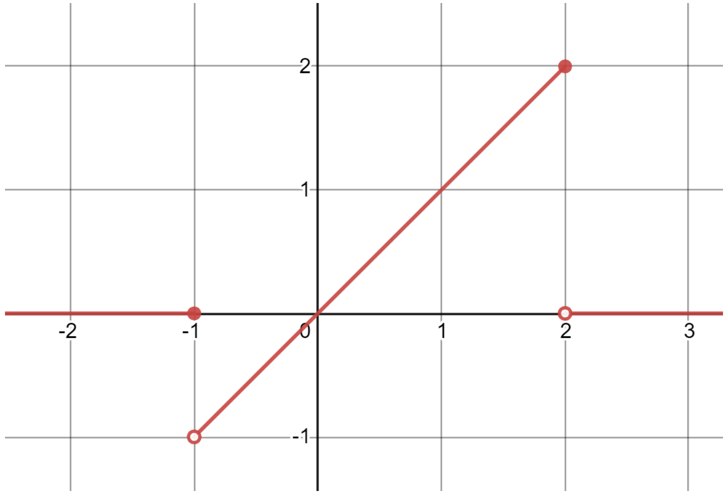
This notation can simplify the presentation of piecewise functions:

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \iff f(x;\lambda) = \lambda e^{-\lambda x} \cdot I_{[0,\infty)}(x)$$

Your Turn! Indicator Extravaganza

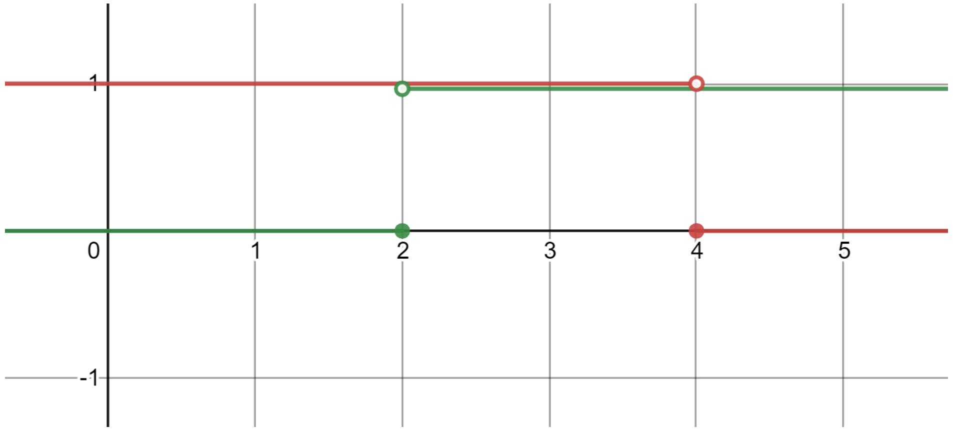
All answers must involve indicator functions!

1. Write a simple formula for this graph:



$$\boxed{f(x) = x \cdot I_{(-1,2]}}$$

2. Simplify $I_{(-\infty,4)} \cdot I_{(2,\infty)}$



$$= \boxed{I_{(2,4)}}$$

In general, $I_{\text{set } A} \cdot I_{\text{set } B} = I_{A \cap B}$

3. Simplify $I_{[1,\infty)} \cdot I_{[3,\infty)} \cdot I_{[5,\infty)}$

$$= I_{[1,\infty)\cap[3,\infty)\cap[5,\infty)} = \boxed{I_{[5,\infty)}} = \boxed{I_{[\max(1,3,5),\infty)}}$$

A Star Is Born!

In Legend of Zelda: Breath of the Wild, star fragments fall once per night at some moment between 9 PM and an unknown stopping time (on the in-game timer) with equal likelihood.

For four random nights, you record the times the star appears: $x_1 = 10$ PM, $x_2 = 1{:}15$ AM, $x_3 = 2$ AM, and $x_4 = 11{:}45$ PM. Find the MLE for the time at which fragments no longer spawn.

Model this situation via $X \sim Unif(0, \theta)$ where $f_X(x; \theta) = \dfrac{1}{\theta}$ where $0 \leq x \leq \theta$ where 9 PM is 0, θ is the end time, and $x_1 = 1, x_2 = 4.25, x_3 = 5, x_4 = 2.75$.

PT 13: Usually, it is easiest to do the problem without the specific data (1, 4.25, 5, 2.75) and instead use the n data points $x_1, x_2, \ldots, x_n$. Then, plug in the real data at the end. This also creates an answer that works beyond your data.

We have $L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} = \frac{1}{\theta^n}$. So $\ell(\theta) = \ln\left(\theta^{-n}\right) = -n \ln \theta$.

$$\implies \ell'(\theta) = -\frac{n}{\theta}, \text{ and so } 0 = -\frac{n}{\widehat{\theta}} \implies 0 = -n!$$

Two signs the usual $L \to \ell \to \ell' = 0$ pipeline won't work:

1. The algebra leads to nonsense!
2. The support of the pdf depends on the parameter (here: $0 \le x \le \theta$).

The "Unusual" Approach: $f \to f$ with indicator on $x \to L$ with indicator on θ

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}, & 0 \le x \le \theta \\ 0, & \text{else} \end{cases} \longrightarrow f(x;\theta) = \frac{1}{\theta} \cdot I_{[0,\theta]}(x)$$

Note: $0 \le x \le \theta < \infty$

$$\longrightarrow L(\theta) = \frac{1}{\theta} \cdot I_{[x,\infty)}(\theta)$$

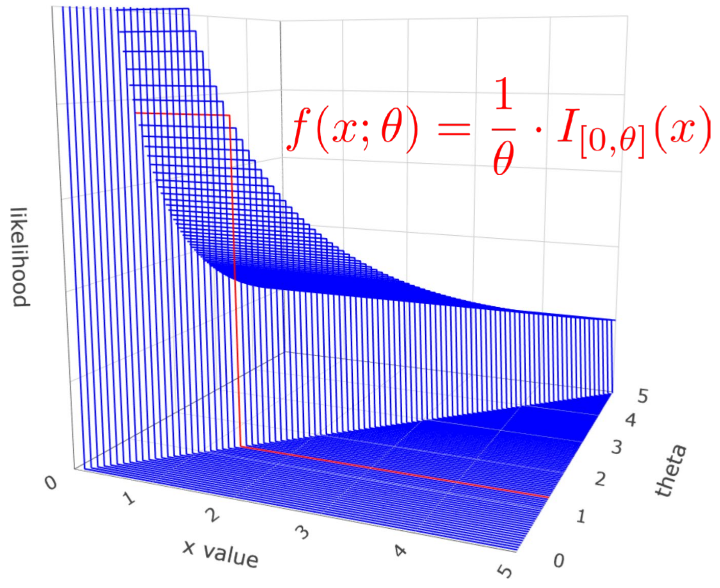
Note: $0 \le x \le \theta < \infty$

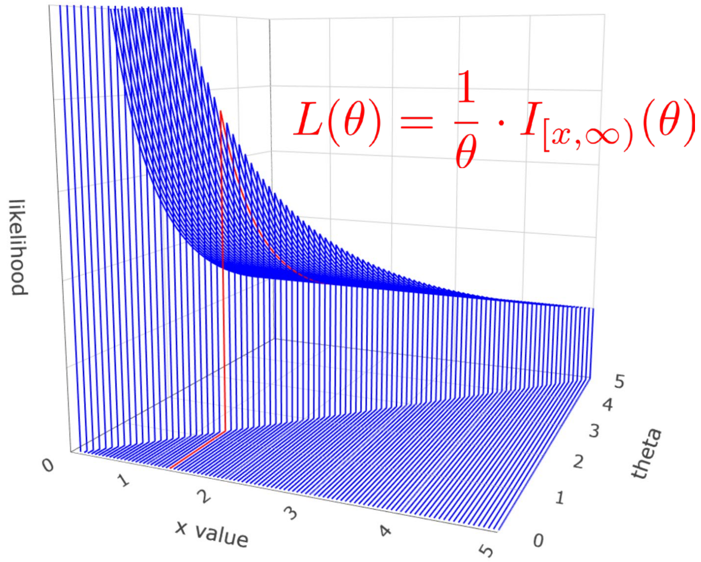
Thus, $L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} I_{[x_i,\infty)} = \frac{1}{\theta^n} \cdot \prod_{i=1}^{n} I_{[x_i,\infty)} = \frac{1}{\theta^n} \cdot I_{[\max x_i,\infty)}(\theta)$

Because of the $\frac{1}{\theta^n}$ term, $L(\theta)$ is maximal when θ is as small as possible.

The indicator function forces $\max x_i \leq \theta < \infty$ (unless we want $L = 0$), so $\boxed{\widehat{\theta}_{\text{MLE}} = \max x_i}$. For our data, $\widehat{\theta}_{\text{MLE}} = 5 = \boxed{\text{2 AM}}$.

Thinking





Fighting Two Enemies at Once

Some models depend on more than one parameter: $N(\mu, \sigma^2)$ or $Gamma(r, \lambda)$. In such cases, the likelihood will be multivariate!

Plan of attack:

$$L(\theta_1, \theta_2) = \text{some expression}$$

$$\ell(\theta_1, \theta_2) = \ln(\text{expression})$$

$$\left.\begin{aligned} \frac{\partial \ell}{\partial \theta_1} &\overset{\text{set}}{=} 0 \\ \frac{\partial \ell}{\partial \theta_2} &\overset{\text{set}}{=} 0 \end{aligned}\right\}$$

System of two equations and two unknowns (θ_1 and θ_2). Solve!

Class Convention!

With one parameter, I forced you to find ℓ'' and show $\ell''(\widehat{\theta}_{\text{MLE}}) < 0$. Proving you actually found a maximum when *multiple* parameters are present is **not** required.

G(r)ammatically Correct

Find the MLEs for r and λ in the Gamma distribution using the sample $x_1, \ldots, x_n$.

Recall if $X \sim Gamma(r, \lambda)$, then $f_X(x; r, \lambda) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}$ for $x > 0$

Thus, $L(r, \lambda) = \prod_{i=1}^{n} \frac{\lambda^r}{\Gamma(r)} x_i^{r-1} e^{-\lambda x_i}$

So, $\ell(r, \lambda) = \sum_{i=1}^{n} [r \ln \lambda - \ln \Gamma(r) + (r-1) \ln x_i - \lambda x_i]$.

$$= nr \ln \lambda - n \ln \Gamma(r) + (r-1) \sum \ln x_i - \lambda \sum x_i$$

If you can't get to this line in three steps or less, you are spending too much time showing small steps that need to be internalized.

Your Turn: Partial Pressure!

$$\ell(r,\lambda) = nr\ln\lambda - n\ln\Gamma(r) + (r-1)\sum\ln x_i - \lambda\sum x_i$$

$$\frac{\partial \ell}{\partial \lambda} = \frac{nr}{\lambda} + 0 + 0 - \sum x_i$$

Solve $0 = \dfrac{n\widehat{r}}{\widehat{\lambda}} - \sum x_i$

So, $\boxed{\widehat{\lambda} = \dfrac{n\widehat{r}}{\sum x_i} = \dfrac{\widehat{r}}{\frac{\sum x_i}{n}} = \dfrac{\widehat{r}}{(\overline{x})}}$

$$\frac{\partial \ell}{\partial r} = n\ln\lambda - n\frac{\Gamma'(r)}{\Gamma(r)} + \sum\ln x_i - 0$$

Solve $0 = n\ln\widehat{\lambda} - n\dfrac{\Gamma'(\widehat{r})}{\Gamma(\widehat{r})} + \sum\ln x_i$

Substituting $\widehat{\lambda} = \dfrac{\widehat{r}}{(\overline{x})}$ gives:

$$\boxed{0 = n\ln\left(\frac{\widehat{r}}{(\overline{x})}\right) - n\frac{\Gamma'(\widehat{r})}{\Gamma(\widehat{r})} + \sum\ln x_i}$$

The right equation has no closed-form solution. In such cases, the answer must be numerically approximated using iterative approaches such as the Newton-Raphson method. When this is the case, the problem will ask for "an equation the MLE satisfies". This phenomenon is one downside of MLE.

Additional Practice

Swipe Left? Swipe Right? Tinder "Success" Percentages

On Tinder, if a person swipes right on your picture, they are interested. The percentage of all people that swipe right on your profile is your "success" percentage. A researcher believes that percentages for different people might be modeled by $X \sim N(\mu, \sigma^2)$ so they collect data $(x_1, \ldots, x_n)$ on a random sample of users. Find the MLE for μ and σ^2.

Recall that $$f_X(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

exp(a) is a way of writing e^a that is useful when the exponent is complicated.

Thus, $$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

So, $$\ell(\mu, \sigma^2) = \sum_{i=1}^{n} \left[-\frac{1}{2}\ln(2\pi\sigma^2) - \frac{(x_i-\mu)^2}{2\sigma^2}\right] = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\text{So, } \ell(\mu,\sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{1}{2\sigma^2}\sum 2(x_i-\mu)(-1)$$

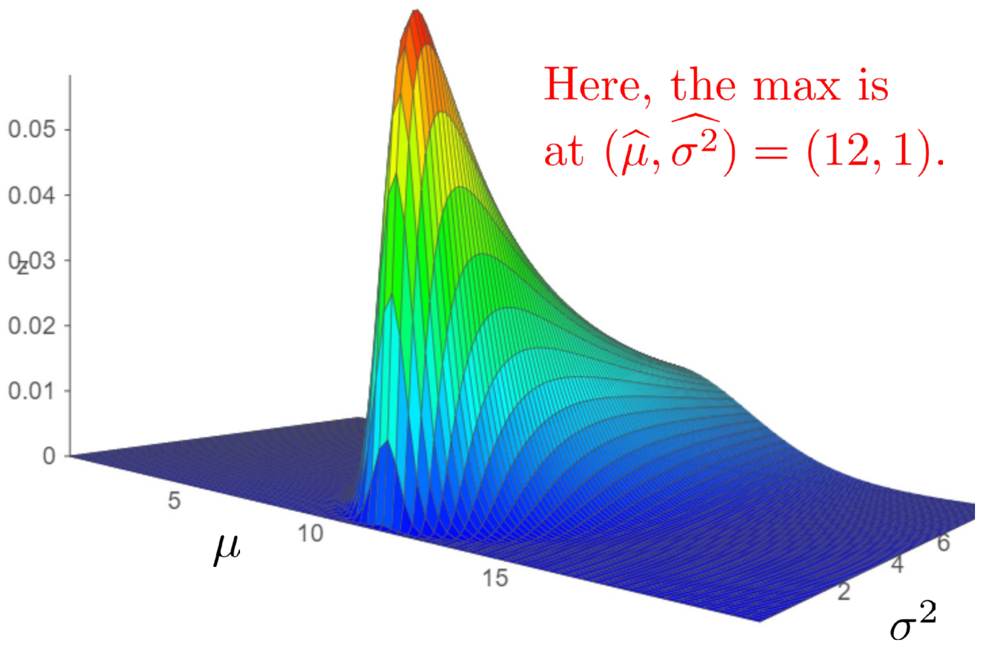
$$0 = \sum(x_i - \widehat{\mu}) = \sum(x_i) - n\widehat{\mu}$$

$$\text{So, } \boxed{\widehat{\mu} = \frac{\sum x_i}{n} = \overline{x}}$$

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2}\frac{2\pi}{2\pi\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum(x_i-\mu)^2$$

$$0 = -\frac{n}{2\widehat{\sigma^2}} + \frac{1}{2\left(\widehat{\sigma^2}\right)^2}\sum(x_i-\widehat{\mu})^2 \implies \boxed{\widehat{\sigma^2} = \frac{1}{n}\sum(x_i-\overline{x})^2}$$

$L(\mu,\sigma^2)$ for $x_1 = 11, x_2 = 13$



Get an intuitive feel for things here. Click on the circles next to 1, 6, and 18. Move the sliders.

Your Turn: No Soup For You!

Each week you volunteer to make pots of soup for the local soup kitchen. You always make at least one pot, but some weeks require more pots because additional people show up. Your pot counts for six random weeks are $x_1 = 1, x_2 = 1, x_3 = 2, x_4 = 1, x_5 = 4, x_6 = 2$. You think Poisson might help model this, but you'll never make 0 pots, so you use the "truncated Poisson distribution":

$$f_X(x; \lambda) = \frac{\lambda^x}{x!(e^\lambda - 1)} \text{ where } x = 1, 2, 3, \ldots$$

Find an equation that the MLE for λ satisfies given these data.

We have $L(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!(e^\lambda - 1)}$ where $\lambda > 0$.

$$\text{Thus, } \ell(\lambda) = \sum_{i=1}^{n} \left[x_i \ln \lambda - \ln(x_i!) - \ln\left(e^{\lambda} - 1\right)\right]$$

$$= (\ln \lambda) \sum x_i - \sum \ln(x_i!) - n \ln\left(e^{\lambda} - 1\right)$$

$$\ell'(\lambda) = \frac{\sum x_i}{\lambda} - 0 - \frac{n}{e^{\lambda} - 1}(e^{\lambda})$$

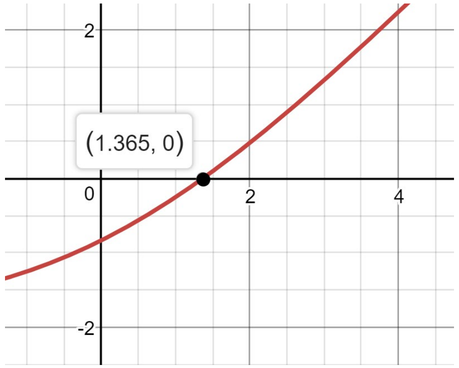
$$\text{Solve } 0 = \frac{\sum x_i}{\widehat{\lambda}} - \frac{n e^{\widehat{\lambda}}}{e^{\widehat{\lambda}} - 1}$$

$$\text{This simplifies to: } \boxed{\frac{\widehat{\lambda} e^{\widehat{\lambda}}}{e^{\widehat{\lambda}} - 1} = \overline{x}}$$

We skip the analysis of ℓ'' in this problem.

For our data, $\overline{x} = \dfrac{11}{6}$, so we must solve:

$$\frac{y e^{y}}{e^{y} - 1} = \frac{11}{6}, \text{ or } \frac{y e^{y}}{e^{y} - 1} - \frac{11}{6} = 0$$



Desmos gives $\boxed{\widehat{\lambda} \approx 1.365 \text{ pots of soup/week}}$.