Agenda For Today

(completely subject to change)

1. MLE asymptotic property review
2. Construction of Confidence intervals (CI's)
3. MLE examples
   - Poisson
   - Laplace
   - Geometric
   - Normal
4. Delta method and MoM

(5.) If time permits theory of MLE asymptotic distribution.

1. The asymptotic normality of the MLE

Under relatively mild regularity conditions conditions on P_θ

θ can be multivariate

For $\hat{\theta}_{MLE}$ from $\{X_i\}_{i=1}^{n} \overset{iid}{\sim} P_\theta$

$$\sqrt{n}\left(\hat{\theta}_{MLE} - \theta\right) \xrightarrow{d} N\left(0, I_1(\theta)^{-1}\right)$$

θ is the true value,

$\hat{\theta}_{MLE}$ is our maximum likelihood estimator for θ.

but also

$$\frac{(\hat{\theta} - \theta)}{\sqrt{I_1(\hat{\theta})^{-1}/n}} \xrightarrow{d} N(0,1)$$

"$\Rightarrow$"

1. Asymptotic unbiasedness.

Def: An unbiased estimator

$$E(\hat{\theta}) = \theta$$

2. $\hat{\theta} \xrightarrow{p} \theta$ consistency

Can we have consistency and not unbiased?

Take $X_i \overset{iid}{\sim} N(\mu, 1)$ $i = 1, \ldots, n$

Take X_1 as my estimator.

$$E(X_1) = \mu$$ Unbiased but not consistent.

$$X_1 \not\to \text{anything.}$$

Take $X_i \overset{iid}{\sim} N(0, \sigma^2)$ $i = 1, \ldots, n$

The MLE for σ^2

$$\frac{1}{n} \sum_{i=1}^{n} X_i^2 \xrightarrow{p} \sigma^2 \quad (\text{by LLN})$$

This is not unbiased!

3. Asymptotic Efficiency (More later)

4. Sufficiency (Don't worry rn)

Now the actual Fisher information

Denote $\ell_n(\theta; \underline{x}) = \log f_\theta(\underline{x})$

$$\ell_1(\theta; x_1) = \log f_\theta(x_1)$$

$$I_1(\theta) = -E\left(\frac{\partial^2}{\partial\theta^2}\ell_1(\theta; x_1)\right)$$

$$= E\left(\left(\frac{\partial}{\partial\theta}\ell_1(\theta; x_1)\right)^2\right)$$

$$= \mathrm{var}\left(\frac{\partial}{\partial\theta}\ell_1(\theta; x_1)\right)$$

Example: Poisson

$X_i \overset{iid}{\sim} \text{Pois}(\lambda) \quad P(X_i = x_i) = \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$

$\hat{\lambda}_{MLE} = \bar{X} \quad E(\bar{X}) = \lambda$

$\text{Var}(X_i) = \lambda \ , \ \text{Var}(\bar{X}) = \frac{\lambda}{n}$

$$l_1(\lambda; x_1) = \log\left(\frac{\lambda^{x_i} e^{-\lambda}}{x_i!}\right)$$

$$= x_i \log(\lambda) - \lambda - \log(x_i!)$$

$$l_1'(\lambda; x_1) = \frac{x_i}{\lambda} - 1$$

$$l_1''(\lambda; x_1) = \frac{-x_i}{\lambda^2}$$

$$-E\left(\ell_1''(\lambda; X_1)\right) = E\left(\frac{X_i}{\lambda^2}\right)$$

$$= \frac{1}{\lambda^2} E(X_i) = \frac{1}{\lambda^2} \cdot \lambda = \frac{1}{\lambda}$$

$$\sqrt{n}\left(\hat{\lambda} - \lambda\right) \xrightarrow{d} N(0, \lambda)$$

what if I use

$$\ell_n(\theta; \underline{x})$$

$$= \log \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

$$= \sum_{i=1}^{n} \log\left(\frac{\lambda^{x_i} e^{-\lambda}}{x_i!}\right)$$

$$= \log\lambda \sum_{i=1}^{n} x_i \quad - n\lambda \quad - \sum_{i=1}^{n} \log(x_i!)$$

$$l_n' = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n$$

$$l_n'' = \frac{-\sum_{i=1}^{n} x_i}{\lambda^2} \Rightarrow -E(l_n'') = \frac{n}{\lambda}$$

Using asymptotic dist
to construct CI

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{d} N(0, \lambda)$$

$$\frac{\sqrt{n}(\hat{\lambda} - \lambda)}{\sqrt{\lambda}} \sim N(0,1)$$

we don't want to deal with λ in variance, so replace it with estimate $\hat{\lambda}$.

So we also have

$$\sqrt{n}\,\frac{(\hat{\lambda}-\lambda)}{\sqrt{\hat{\lambda}}} \overset{\cdot}{\sim} N(0,1)$$

This is called a pivot

it's a function of our parameter with a known distribution

<u>Def.</u> A $(1-\alpha)100\%$ confidence interval is a random interval such that for true parameter

$\theta \qquad P([a,b] \ni \theta) = 1-\alpha$

θ is either in or not in the CI (confidence interval)

We cannot make probalistic statements about θ.

Back to Poisson

$$\sqrt{n}\,\frac{(\hat{\lambda}-\lambda)}{\sqrt{\hat{\lambda}}} \sim N(0,1)$$

$$P\left(a \le \frac{\hat{\lambda} - \lambda}{\sqrt{\hat{\lambda}}/\sqrt{n}} \le b\right) \simeq 1-\alpha.$$

$$b = z\left(1 - \frac{\alpha}{2}\right)$$

$$a = z\left(\frac{\alpha}{2}\right)$$

These are quantiles from standard normal.

$$F\left(z(\beta)\right) = \beta$$

$$P\left(z\left(\frac{\alpha}{2}\right) \le \frac{\hat{\lambda} - \lambda}{\sqrt{\hat{\lambda}}/\sqrt{n}} \le z\left(1 - \frac{\alpha}{2}\right)\right)$$

$$= P\left(z\left(\frac{\alpha}{2}\right)\frac{\sqrt{\hat{\lambda}}}{\sqrt{n}} \le \hat{\lambda} - \lambda \le z\left(1 - \frac{\alpha}{2}\right)\frac{\sqrt{\hat{\lambda}}}{\sqrt{n}}\right)$$

$$= P\left(z\left(\frac{\alpha}{2}\right)\frac{\sqrt{\hat{\lambda}}}{\sqrt{n}} - \hat{\lambda} \le -\lambda \le z\left(1-\frac{\alpha}{2}\right)\frac{\sqrt{\hat{\lambda}}}{\sqrt{n}} - \hat{\lambda} \right)$$

$$= P\left(\hat{\lambda} - z\left(1-\frac{\alpha}{2}\right)\frac{\sqrt{\hat{\lambda}}}{\sqrt{n}} \le \lambda \le \hat{\lambda} - z\left(\frac{\alpha}{2}\right)\frac{\sqrt{\hat{\lambda}}}{\sqrt{n}} \right)$$

$$= 1 - \alpha$$

$$\hat{\lambda} \pm z\left(1-\frac{\alpha}{2}\right)\frac{\sqrt{\hat{\lambda}}}{\sqrt{n}}$$

<u>Any MLE (from iid)</u>

$$\hat{\theta}_{MLE} \pm z\left(1-\frac{\alpha}{2}\right)\frac{\sqrt{I(\hat{\theta})^{-1}}}{\sqrt{n}}.$$

A brief interlude

$N(\mu, \sigma^2)$ (HOLDS FOR FINITE SAMPLE)

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

if σ known

Suppose σ unknown

we estimate

$$\sigma^2 \text{ w/ } s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$\frac{\bar{X} - \mu}{s / \sqrt{n}} \sim t_{n-1}$$

FOR MLE CAN STILL USE NORMAL B/C OF ASYMPTOTIC

When to use Sample Variance

$$\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 = S^2$$

$$\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2 \text{ is}$$

MLE of σ^2

$X_i \sim N(\mu_0, \sigma^2)$

Laplace Dist

$$f_\theta(x) = \frac{1}{2}\exp\left(-|X_i-\theta|\right)$$

I claim MLE is $\text{med}(X_i)$

$$\ell_n(\theta;\underline{x}) = \log \prod_{i=1}^{n} \frac{1}{2} \exp\left(-|x_i-\theta|\right)$$

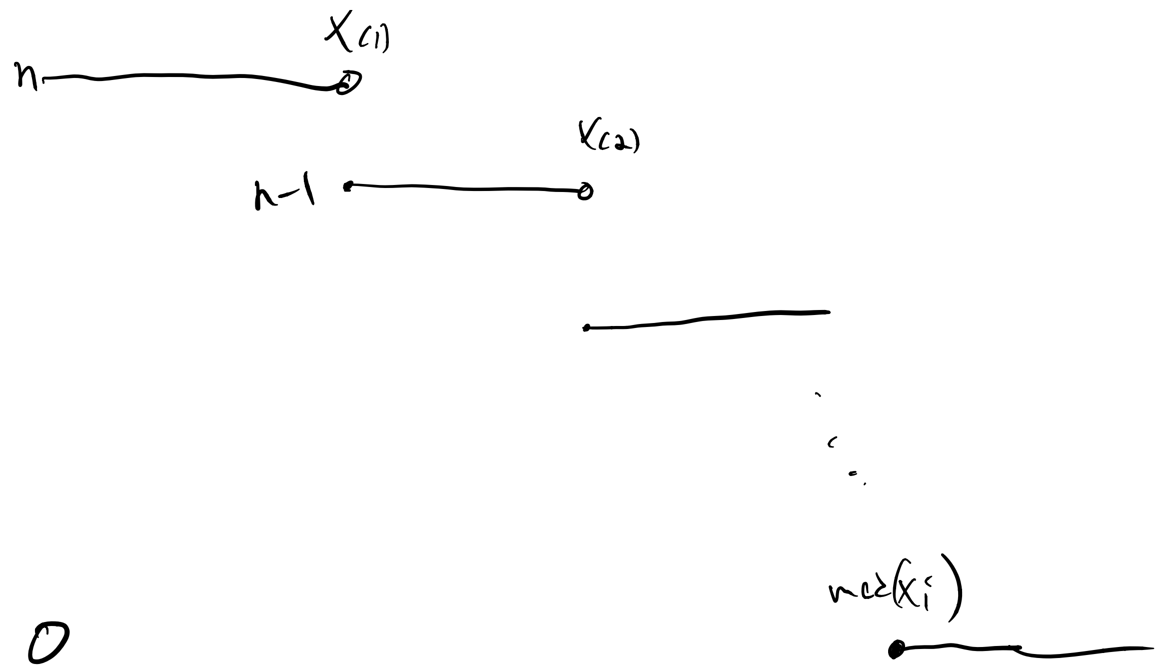
$$= \sum_{i=1}^{n} \log\left(\frac{1}{2} \exp\left(-|x_i-\theta|\right)\right)$$

$$= n\log\left(\frac{1}{2}\right) - \sum_{i=1}^{n} |x_i - \theta|$$

This is differentiable everywhere except x_i's

$$\ell_n'(\theta;x) = \sum_{i=1}^{n} \operatorname{sign}(x_i - \theta)$$

Where the derivative changes signs? (Suppose n is odd).



$$\hat{\theta}_{MLE} = \text{med}(X_i).$$

$$l_1'(\theta; X_1) = \text{sign}(X_1 - \theta).$$

$$P(\text{sign}(X_1 - \theta) = 1) = \frac{1}{2}$$

$$P(\text{sign}(X_1 - \theta) = -1) = \frac{1}{2}.$$

"Rademacher dist".

$$\mathrm{Var}\left(l_1'(\theta; x_1)\right) = 1 = I(\theta).$$

$$\sqrt{n}\left(\hat{\theta} - \theta\right) \xrightarrow{d} N(0, 1)$$

$$E\left(\mathrm{sign}(x_i - \theta)\right) = 1 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2}$$
$$= 0.$$

$$E\left(\left(\mathrm{sign}(x_i - \theta)\right)^2\right) = 1^2 \cdot \frac{1}{2} + (-1)^2 \cdot \frac{1}{2}$$
$$= 1.$$

$$\frac{\hat{\theta} - \theta}{1/\sqrt{n}} \dot\sim N(0, 1)$$

note $\dfrac{\bar{X} - \theta}{1/\sqrt{n}} \dot\sim N(0, 2)$

twice as large var!

Method of Moments

We want to use LLN
to get a consistent estimator

"consistency" $\hat{\theta} \xrightarrow{P} \theta$

$$P\left(|\hat{\theta}-\theta| > \varepsilon\right) \longrightarrow 0, \quad \forall \varepsilon > 0$$

$X_i \overset{iid}{\sim} N(\mu_0, \sigma^2)$ μ_0 known

want to estimate σ^2.

$$E(X_i) = \mu_0$$

$$E\left(\left(X_i^2\right)\right) = Var(X_i) + \left(E(X_i)\right)^2$$

$$= \sigma^2 + \mu_0^2.$$

We can estimate any moment with LLN

$$\frac{1}{n}\sum_{i=1}^{n}(X_i)^r \xrightarrow{p} E(X^r).$$

$$E(X_i^2) - \mu_0^2 = \sigma^2.$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \mu_0^2.$$

$$E(\hat{\sigma}^2) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^2) - \mu_0^2$$

$$= \frac{1}{n} n(\sigma^2 + \mu_0^2) - \mu_0^2$$

$$= \sigma^2 + \mu_0^2 - \mu_0^2$$

$$= \sigma^2.$$

MoM

1. Calculate moments in increasing order until you can solve the system of equations for each unknown parameter
2. Solve for parameters in terms of moments.

3. Replace moments with sample moments.

You now have a consistent estimator.

δ-method

$$\hat{\theta}_n \xrightarrow{d} N\left(m, \frac{\sigma^2}{n}\right)$$

I want to know distribution of

$f(\hat{\theta}_n)$, suppose that $f'(m)$ exists and $f'(m) \neq 0$.

then

$$f(\hat{\theta}_n) \xrightarrow{d} N\left(f(m), \frac{\sigma^2}{n} \cdot (f'(m))^2\right)$$

$$\frac{\sqrt{n}\,\left(f(\hat{\theta}_n) - f(m)\right)}{|f'(m)|} \mathrel{\dot\sim} N(0,1)$$

<u>Ex.</u> <u>Expon. Dist</u>

$$g_\lambda(x) = \lambda e^{-\lambda x}, \; x \geq 0.$$

$$E(X) = \frac{1}{\lambda}$$

Use MoM

Find moment

$$\lambda = \frac{1}{E(X)}$$

Solve for param

$$\hat{\lambda} = \frac{1}{\bar{X}}$$

replace with sample moment

How do I find distribution of this?

$$\bar{X} \sim N\left(\frac{1}{\lambda}, \frac{1}{n\lambda^2}\right)$$

B/c $var(X) = \frac{1}{\lambda^2}$

$$f(a) = \frac{1}{a}$$

$$f'(a) = \frac{-1}{a^2}$$

$$\frac{1}{\bar{X}} \sim N\left(\lambda, \frac{1}{n\lambda^2} \cdot \left(f'\left(\frac{1}{\lambda}\right)\right)^2\right)$$

$$\dot{\sim} N\left(\lambda, \frac{1}{n\lambda^2} \cdot \lambda^4\right)$$

$$= N\left(\lambda, \frac{\lambda^2}{n}\right).$$

Examples I didn't get to

Geometric Dist

On HW2 we showed MLE is $\hat{p} = \frac{1}{\bar{x}}$

Now for asymptotic

distribution.

$$P_p(X = x_i) = (1-p)^{x_i - 1} p$$

$$l_1(p; x_1) = \log p + (x_i - 1) \log(1-p)$$

$$l_1'(p; x_1) = \frac{1}{p} - \frac{x_i - 1}{1-p}$$

$$l_1''(p; x_1) = -\frac{1}{p^2} - \frac{x_i - 1}{(1-p)^2}$$

$$-E(l_1''(p; x_1)) = \left(\frac{1}{p^2} + \frac{\frac{1}{p} - 1}{(1-p)^2} \right)$$

$$= \frac{1}{p^2(1-p)}$$

$$\Rightarrow \sqrt{n}\left(\frac{1}{\bar{X}} - p\right) \xrightarrow{d} N\left(0, p^2(1-p)^2\right)$$

Normal with

Unknown mean and variance.

So $X_i \overset{iid}{\sim} N(\mu, \tau)$

both μ, τ are unknown

we first solve

$$\ell_n(\mu, \tau; \underline{x})$$

$$= \log \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\tau}} \exp\left(\frac{-(x_i-\mu)^2}{2\tau}\right)$$

$$= \sum_{i=1}^{n} \log\left(\frac{1}{\sqrt{2\pi\tau}} \exp\left(\frac{-(x_i-\mu)^2}{2\tau}\right)\right)$$

$$= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\tau) - \frac{1}{2\tau}\sum_{i=1}^{n}(x_i-\mu)^2$$

we solve $\nabla \ell_n(\mu, \tau; \underline{x}) = \underline{0}$

$$\frac{\partial \ell}{\partial \mu} = \frac{1}{\tau}\sum_{i=1}^{n}(x_i - \mu) = 0$$

$$\Rightarrow \hat{\mu} = \bar{x}$$

$$\frac{\partial \ell}{\partial \tau} = \frac{-n}{2\tau} + \frac{1}{2\tau^2} \sum_{i=1}^{n} (x_i - \mu)^2 = 0$$

$$\hat{\tau} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2$$

So the MLE estimate is

$$\begin{pmatrix} \hat{\mu} \\ \hat{\tau} \end{pmatrix} = \begin{pmatrix} \bar{x} \\ \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \end{pmatrix}$$

note I replaced μ with it's MLE $\bar{x}$.

Now we want the Fisher information which will be a matrix in this case.

$$\frac{\partial^2 \ell_n}{\partial \mu^2} = \frac{-n}{\gamma} \Rightarrow -E\left(\frac{\partial^2 \ell}{\partial \mu^2}\right) = \frac{n}{\gamma}$$

$$\frac{\partial^2 \ell_n}{\partial \gamma^2} = \frac{n}{2\gamma^2} - \frac{1}{\gamma^3}\sum_{i=1}^{n}(X_i - \mu)^2$$

$$-E\left(\frac{\partial^2 \ell_n}{\partial \gamma^2}\right) = -\frac{n}{2\gamma^2} + \frac{1}{\gamma^3}\sum_{i=1}^{n} E\left((X_i - \mu)^2\right)$$

$$= -\frac{n}{2\gamma^2} + \frac{1}{\gamma^3} \cdot n\gamma$$

$$= \frac{-n}{2\tau^2} + \frac{n}{\tau^2} = \frac{n}{2\tau^2}$$

Note also

$$-E\left(\frac{\partial l_n}{\partial \tau \partial \mu}\right) = \frac{+2}{n} \sum_{i=1}^{n} E(X_i - \mu)$$

$$= 0$$

$$= -E\left(\frac{\partial l_n}{\partial \mu \partial \tau}\right)$$

Therefore our Fisher information matrix is

$$I_n(\mu,\sigma) = n \begin{pmatrix} \frac{1}{\tau} & 0 \\ 0 & \frac{1}{2\tau^2} \end{pmatrix}$$

Therefore we have that

$$\sqrt{n}\left(\begin{pmatrix} \hat{\mu} \\ \hat{\tau} \end{pmatrix} - \begin{pmatrix} \mu \\ \tau \end{pmatrix} \right) \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, I_1(\mu,\sigma)^{-1} \right)$$

$$I_1(\mu,\sigma)^{-1} = \begin{pmatrix} \frac{1}{\tau} & 0 \\ 0 & \frac{1}{2\tau^2} \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} T & 0 \\ 0 & 2T^2 \end{pmatrix}$$

(Diagonal matrix inverse is inverse of diagonal elements).