Agenda For Today

1. MLE Theory Review.
   a) Poisson Example
2. Constructing CI's with MLE
3. Interludes
   a) t dist
   b) Sample Variance vs. Gaussian MLE
4. Laplace Dist
5. MoM
6. δ-method
7. More ex. if time.

1. The Asymptotic Normality of the MLE.

Take some p_θ, some mild regularity conditions (θ can be multivariate)

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

Let θ denote true parameter.

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

The Fisher information

equiv.

$$\frac{\hat{\theta}_{MLE} - \theta}{\sqrt{I_1(\theta)^{-1}/n}} \dot{\sim} N(0,1)$$

" $\overset{\cdot}{\sim}$ " denotes approx dist.

1. Asymptotic unbiasedness

<u>Def.</u> An unbiased estimator

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

2. Consistency.

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

Q: Can we have an unbiased estimator that is not consistent?

Ex. $X_i \overset{iid}{\sim} N(\mu, 1)$.

want to estimate μ.

$\hat{\mu} = X_1$, $E(\hat{\mu}) = \mu$

X_1 is not consistent for anything.

Q: The reverse ?

Take $X_i \overset{iid}{\sim} N(0, \sigma^2)$

Want to estimate σ^2.

The MLE for σ^2 is

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

This is not unbiased.

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

Other nice properties talked about later

3. Efficiency

4. Sufficiency.

(Don't worry rn)

what is $I_1(\theta)$?

Denote $l_n(\theta; \underline{x}) = \log \prod_{i=1}^{n} f_\theta(x_i)$

$$l_1(\theta; x_1) = \log f_\theta(x_1),$$

$$I_1(\theta) = -E\left(\frac{\partial^2}{\partial\theta^2} l_1(\theta; x_1)\right)$$

$$= E\left(\left(\frac{\partial}{\partial\theta} l_1(\theta; x_1)\right)^2\right)$$

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

b/c $E\left(\frac{\partial}{\partial\theta} \ell_1(\theta; x_1)\right) = 0.$

Ex. Poisson

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

$E(X_i) = \lambda \quad \mathrm{Var}(X_i) = \lambda$

$\hat{\lambda}_{MLE} = \bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$

$$E(\hat{\lambda}_{MLE}) = E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} EX_i$$

$$= \frac{1}{n}\sum_{i=1}^{n} \lambda = \frac{1}{n}\cdot n\lambda$$

$$= \lambda$$

$$\operatorname{var}(\hat{\lambda}_{MLE}) = \frac{\lambda}{n}$$

$$\ell_1(\lambda; x_1) = \log\left(\frac{\lambda^{x_1} e^{-\lambda}}{x_1!}\right)$$

$$= x_1 \log\lambda - \lambda - \log x_1!$$

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

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

$$-E(\ell''(\lambda; X_1)) = \frac{1}{\lambda} = I_1(\lambda).$$

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

What if I use

$$\ell_n(\theta; x)$$

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

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

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

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

$$= \frac{n}{\lambda}$$

Using asymptotic dist to construct Confidence Interval (CI)

So we know this fact about our estimator

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

Def. A $(1-\alpha)\cdot 100\%$ confidence interval is a random interval such that for true parameter θ

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

$$\alpha \in [0,1]$$

We cannot make probabilistic statements about θ.

$$\frac{\sqrt{n}\,(\hat{\theta}-\theta)}{\sqrt{I_1(\theta)^{-1}}} \overset{\cdot}{\sim} N(0,1)$$

This is called a pivot. It's a function of the true parameter that is a known distribution.

Another useful fact is that for MLE.

$$\frac{\sqrt{n}\,(\hat{\theta}-\theta)}{\sqrt{I_1(\hat{\theta})^{-1}}} \overset{\cdot}{\sim} N(0,1)$$

(under regularity conditions $\ell \in C^3$)

Constructing the CI

We want

$$P\left(c_1 \le \frac{\hat{\theta}-\theta}{\sqrt{I_1(\hat{\theta})^{-1}}/\sqrt{n}} \le c_2\right) \approx 1-\alpha.$$

$$c_2 = z\left(1-\frac{\alpha}{2}\right) \quad c_1 = z\left(\frac{\alpha}{2}\right).$$

These are the standard normal quantiles.

Def. ~~If~~ $z(\alpha)$ is αth quantile and F is distribution function $F(z(\alpha)) = \alpha$.

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

$$= 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha.$$

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

$$P\left(z\left(\frac{\alpha}{2}\right) \cdot \frac{\sqrt{I(\hat{\theta})^{-1}}}{\sqrt{n}} \le \hat{\theta} - \theta \le z\left(1 - \frac{\alpha}{2}\right) \frac{\sqrt{I(\hat{\theta})^{-1}}}{\sqrt{n}}\right)$$

$$P\left(z\left(\frac{\alpha}{2}\right) \frac{\sqrt{I(\hat{\theta})^{-1}}}{\sqrt{n}} - \hat{\theta} \le -\theta \le z\left(1 - \frac{\alpha}{2}\right) \frac{\sqrt{I(\hat{\theta})^{-1}}}{\sqrt{n}} - \hat{\theta}\right)$$

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

$$= 1 - \alpha.$$

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

A brief Interlude

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

we often see

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

$X_i \sim N(\mu, \sigma^2)$.

S^2 is unbiased while second is not, so usually S^2.

$$\sqrt{I_n(\hat{\theta})} \cdot (\hat{\theta} - \theta) \overset{\cdot}{\sim} N(0,1)$$

Laplace Dist

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

I claim MLE is sample median.

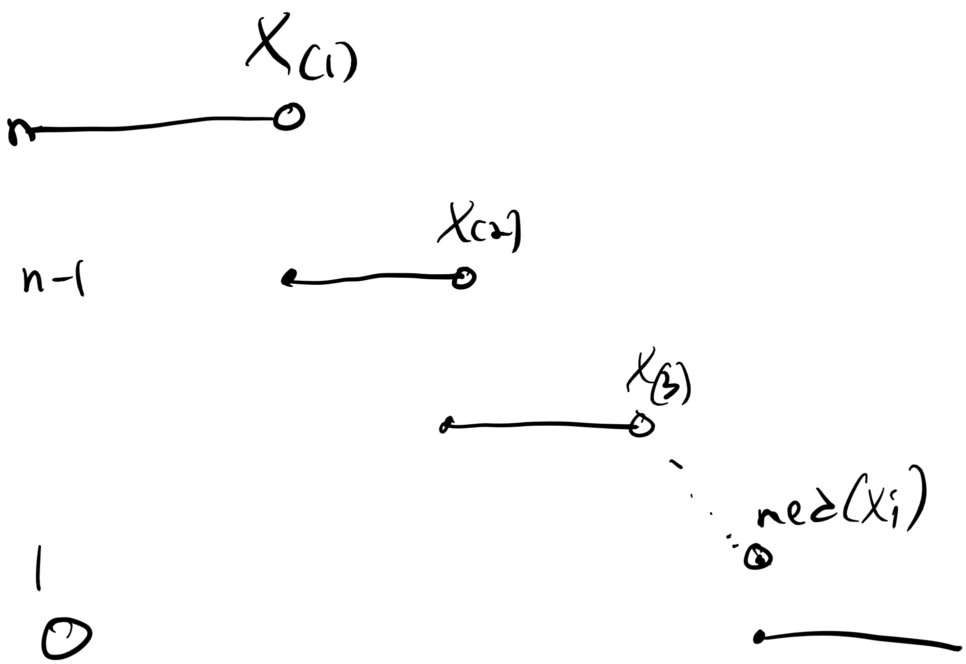
$$\ell_n(\theta; x) = \log \prod_{i=1}^{n} \frac{1}{2}\exp(-|x_i - \theta|)$$

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

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

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

When does the derivative change signs. (suppose n is odd).



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

"sample median X_i"

$$l_1'(\theta; x_1) = \text{sign}(x_1 - \theta).$$

Can use variance def'n of Fisher information

Symmetric distribution takes on values ± 1.

$$P(\ell_1' = 1) = \frac{1}{2}$$

$$P(\ell_1' = -1) = \frac{1}{2}.$$

$$E(\ell_1') = 1 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2} = 0.$$

$$I_1(\theta) E((\ell_1')^2) = 1^2 \cdot \frac{1}{2} + (-1)^2 \cdot \frac{1}{2} = 1.$$

$$\sqrt{n}\,(\hat{\theta}_{MLE} - \theta) \xrightarrow{d} N(0, 1).$$

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

Take the other estimator

$\bar{X}$ (Method of Moments)

$E(\bar{X}) = \theta$

$Var(\bar{X}) = \frac{2}{n}$, $Var(X_i) = 2$.

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

$$\sqrt{n}(\bar{X} - \theta) \overset{\cdot}{\sim} N(0,2)$$

Method of Moments

We want to use LLN

to get a consistent estimator.

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

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

μ_0 known.

Want to estimate σ^2

$$E(X_i) = \mu_0$$

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

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

we can estimate any moment with LLN

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

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

Consistent by construction.

MoM

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

2. Solve for param in terms of moments

3. Replace moments with sample moments

You now have consistent estimator.

δ-method

Suppose we know.

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

I want to know the distribution

of $f(\hat{\theta}_n)$

Suppose $f'(m)$ exists

$$f'(m) \neq 0.$$

then

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

equiv.

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

Example: Exponential Dist

$$g_\lambda(x) = \lambda e^{-\lambda x} \quad x > 0$$

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

Use MoM

1. Find moments

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

2. solve param in terms of moments

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

3. Replace with sample moment

(approx)

How do I find distribution of $\frac{1}{\overline{X}}$.

$$\overline{X} \sim N\left(\frac{1}{\lambda}, \frac{1}{n\lambda^2}\right) \text{ by CLT}$$

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

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

by δ-method

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

$$\overset{d}{\approx} N\left(\lambda, \frac{1}{n\lambda^2} \lambda^4\right)$$

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

$$\frac{1}{n} \sum_{i=1}^{n} f(X_i) \xrightarrow{P} E(f(X_i))$$

fcts

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

both μ, τ are unknown.

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

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

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

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

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

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

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

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

replaced μ with $\hat{\mu}$

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

$$< 0.$$

$$\frac{\partial^2 \ell_n}{\partial \tau^2} = \frac{n}{2\tau^2} - \frac{1}{\tau^3} \sum_{i=1}^{n} (x_i - \mu)^2.$$

$$< 0.$$

$$-E\left(\frac{\partial^2 \ell_n}{\partial \tau^2}\right) = \frac{-n}{2\tau^2} + \frac{n}{\tau^2}$$

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

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

$$= 0.$$

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

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

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

Look in AM

file for more

problems!