MATH 341/641 LEC 13

Last class we proved the following asymptotic result:

$$\frac{s(\theta; X_1, \ldots, X_n)}{\sqrt{I_n(\theta)}} \xrightarrow{d} N(0,1) \quad \text{by CLT}$$

If we are testing $H_a: \theta \neq \theta_0 \Longrightarrow H_0: \theta = \theta_0$

$$\Rightarrow \quad \hat{Z} \mid H_0 = \frac{s(\theta_0; X_1, \ldots, X_n)}{\sqrt{I_n(\theta_0)}}$$

which gives us an approximate test.

So what is the usefulness of this thm? The most useful is when its difficult to derive a $\hat{\theta}$! When is this the case?

Consider DGP: $X_1, \ldots X_n \overset{iid}{\sim} \text{Logistic}(\theta, 1) := \dfrac{e^{-(x-\theta)}}{\left(1+e^{-(x-\theta)}\right)^2}$

MM gives us $\hat{\theta}^{MM} = \bar{X}$ always. From there, there's a CLT. Can we find an alternate test? Let's get the MLE

$$\mathcal{L}(\theta; X_1 \ldots X_n) = \prod_{i=1}^{n} \frac{e^{-(X_i-\theta)}}{\left(1+e^{-(X_i-\theta)}\right)^2} = \frac{e^{-\sum(X_i-\theta)}}{\prod_{i=1}^{n}\left(1+e^{-(X_i-\theta)}\right)^2} = \frac{e^{-n\bar{X}+n\theta}}{\prod_{i=1}^{n}\left(1+e^{-(X_i-\theta)}\right)^2}$$

$$\ell(\theta; X_1 \ldots X_n) = -n\bar{X} + n\theta - 2\sum_{i=1}^{n} \ln\left(1+e^{\theta}e^{-X_i}\right)$$

$$\ell'(\theta; X_1 \ldots X_n) = n - 2\sum_{i=1}^{n} \frac{e^{\theta}e^{-X_i}}{1+e^{\theta}e^{-X_i}} \overset{set}{=} 0 \quad \text{Good Luck!}$$

$\Rightarrow$ MLE not available in closed form!

$$\ell''(\theta; X_1 \ldots X_n) = -2\sum \frac{d}{d\theta}\left[\frac{e^{\theta}e^{-X_i}}{1+e^{\theta}e^{-X_i}}\right] = -2\sum \frac{(1+e^{\theta}e^{-X_i})(e^{\theta}e^{-X_i}) - (e^{\theta}e^{-X_i})^2}{(1+e^{\theta}e^{-X_i})^2}$$

$$= -2\sum \frac{e^{\theta}e^{-X_i}}{(1+e^{\theta}e^{-X_i})^2}$$

$$I_n(\theta) = E\left[-\ell''(\theta; X_1 \ldots X_n)\right] = 2n\, E\left[\frac{e^{-(X-\theta)}}{\left(1+e^{-(X-\theta)}\right)^2}\right] = \int_{\mathbb{R}} \frac{e^{-(x-\theta)}}{\left(1+e^{-(x-\theta)}\right)^2} \frac{e^{-(x-\theta)}}{\left(1+e^{-(x-\theta)}\right)^2} dx$$

$$= \int_{\mathbb{R}} \frac{\left(e^{-(x-\theta)}\right)^2}{\left(1+e^{-(x-\theta)}\right)^4} dx$$

$$= \int_{\mathbb{R}} \left(\frac{1}{1+e^{-(x-\theta)}}\right)^2 \left(\frac{e^{-(x-\theta)}}{(1+e^{-(x-\theta)})^2}\right)^2$$

let $u = \frac{1}{1+e^{-(x-\theta)}}$, $1-u = 1 - \frac{1}{1+e^{-(x-\theta)}} = \frac{e^{-(x-\theta)}}{1+e^{-(x-\theta)}}$

$\frac{du}{dx} = \left(1+e^{-(x-\theta)}\right)^{-2} e^{-(x-\theta)} = \frac{1}{1+e^{-(x-\theta)}} \cdot \frac{e^{-(x-\theta)}}{1+e^{-(x-\theta)}} = u(1-u) \Rightarrow dx = \frac{1}{u(1-u)} du$

$x = -\infty \Rightarrow u = 0$, $x = \infty \Rightarrow u = 1$

$$= \int_0^1 u^2(1-u)^2 \frac{1}{u(1-u)} du = \int_0^1 u(1-u)\, du = \int_0^1 (u - u^2) du = \left[\frac{u^2}{2} - \frac{u^3}{3}\right]_0^1$$

$$= \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

$\Rightarrow J_n(\theta) = 2n\left(\frac{1}{6}\right) = \frac{n}{3}$ $\Rightarrow$ To test $H_0: \theta = \theta_0$, we ask

$$\Rightarrow \hat{z} = \frac{n - 2\sum \frac{e^{\theta_0} e^{-x_i}}{1+e^{\theta} e^{-x_i}}}{\sqrt{\frac{n}{3}}} \overset{?}{\in} [-2, 2]$$

e.g. $X = \langle 1.85, 4.18, 2.63, 3.41, 4.23, 4.24, 3.17, 3.64, 3.13, 1.04 \rangle$, $n = 10$

$H_0: \theta = 0$

$\Rightarrow \hat{z} = 4.77 \Rightarrow$ Reject H_0

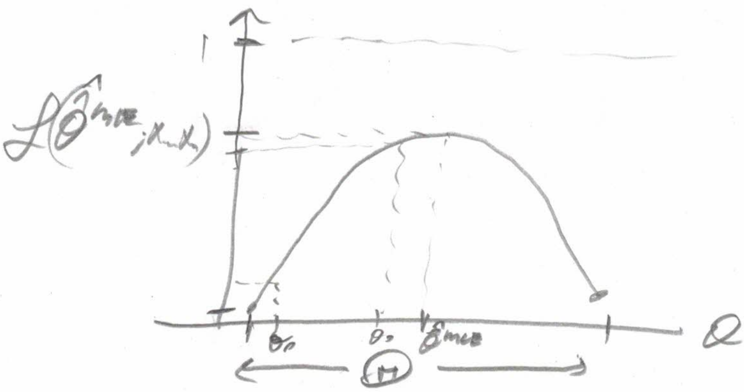
Again, no estimate of θ is necessary...

Let's explore another strategy to create a test for $H_a: \theta \neq \theta_0$

Consider the following test statistic, the likelihood ratio (LR):

$$\hat{LR} = \frac{\mathcal{L}(\hat{\theta}^{MLE}; X_1, \ldots, X_n)}{\mathcal{L}(\theta_0; X_1, \ldots, X_n)} = \prod_{i=1}^{n} \frac{\mathcal{L}(\hat{\theta}^{MLE}; X_i)}{\mathcal{L}(\theta_0; X_i)}$$

Note this ratio at a minimum is 1 since $\mathcal{L}(\hat{\theta}^{MLE}; X_1, \ldots, X_n)$ is the maximum the likelihood could ever become (by def)



Note: the denominator can never be zero since that would mean $\theta_0 \notin \Theta \Rightarrow \hat{LR} \in [1, \infty)$

What indicates a departure from H_0? If the ratio is large i.e. $\mathcal{L}(\hat{\theta}^{MLE}; X_1, \ldots, X_n) >> \mathcal{L}(\theta_0; X_1, \ldots, X_n)$

this means the data estimates θ far from $\theta_0 \Rightarrow$ Reject H_0!

The last thing we need to have a working test is the distr. of the test statistic. Consider instead:

$$\hat{\Lambda} := 2\ln(\hat{LR}) = 2\ln\left(\frac{\mathcal{L}(\hat{\theta}^{mle};X_1,\dots,X_n)}{\mathcal{L}(\theta_0;X_1,\dots,X_n)}\right) = 2\left(\ell(\hat{\theta}^{mle};X_1,\dots,X_n) - \ell(\theta_0;X_1,\dots,X_n)\right)$$

"capital Lambda"

$\hat{\Lambda}$ and $\hat{LR}$ are 1:1. If $\hat{LR}\uparrow \Leftrightarrow \hat{\Lambda}\uparrow$

So we reject if $\hat{\Lambda}$ is large. We now prove the asymptotic distr of $\hat{\Lambda}$ (see p. 489 of C & B)

Just like when proving the core MLE thm, we consider the 2nd order Taylor series for the log likelihood:

$$\ell(\theta_0;X_1,\dots,X_n) \hat{=} \ell(\hat{\theta}^{mle};X_1,\dots,X_n) + \underbrace{(\theta_0-\hat{\theta}^{mle})\,\ell'(\hat{\theta}^{mle};X_1,\dots,X_n)}_{=0} + \frac{1}{2}(\theta_0-\hat{\theta}^{mle})^2\,\ell''(\hat{\theta}^{mle};X_1,\dots,X_n)$$

drop this when going formal

$$\Rightarrow \ell(\theta_0;X_1,\dots,X_n) = \ell(\hat{\theta}^{mle};X_1,\dots,X_n) + \frac{1}{2}(\theta_0-\hat{\theta}^{mle})^2\,\ell''(\hat{\theta}^{mle};X_1,\dots,X_n)$$

$$\Rightarrow \ell(\hat{\theta}^{mle};X_1,\dots,X_n) - \ell(\theta_0;X_1,\dots,X_n) = -\frac{1}{2}(\theta_0-\hat{\theta}^{mle})^2\,\ell''(\hat{\theta}^{mle};X_1,\dots,X_n)$$

mult. by 2

$$\Rightarrow \hat{\Lambda} = -(\theta_0-\hat{\theta}^{mle})^2\,\ell''(\hat{\theta}^{mle};X_1,\dots,X_n)\cdot\frac{\ell''(\theta;X_1,\dots,X_n)}{\ell''(\theta;X_1,\dots,X_n)}$$

$$= -(\theta_0-\hat{\theta}^{mle})^2\,\ell''(\theta;X_1,\dots,X_n)\,\underbrace{\frac{\ell''(\hat{\theta}^{mle};X_1,\dots,X_n)}{\ell''(\theta;X_1,\dots,X_n)}}_{A_1}$$

$$= \frac{\left(\theta_0 - \hat{\theta}^{MLE}\right)^2}{\frac{1}{-\ell''(\theta; X_1..X_n)}} A_1 = \frac{\left(\theta_0 - \hat{\theta}^{MLE}\right)^2}{\frac{1}{\sum -\ell''(\theta; X_i) \cdot \frac{n}{n}}} \cdot \frac{\frac{1}{I(\theta)}}{I(\theta)} A_1$$

$$= \frac{\left(\theta_0 - \hat{\theta}^{MLE}\right)^2}{\frac{1}{n\, I(\theta)}} \cdot \underbrace{\frac{\frac{1}{n} \sum -\ell''(\theta; X_i)}{I(\theta)}}_{A_2} A_1$$

$$= \underbrace{\left(\frac{\theta_0 - \hat{\theta}^{MLE}}{\sqrt{\frac{I(\theta)^{-1}}{n}}}\right)^2}_{B} A_1 A_2$$

Facts $B \xrightarrow{d} N(0,1)$ by core MLE thm

$\Rightarrow B^2 \xrightarrow{d} \chi^2$ by CMT which we never proved but it's true...

$A_1 = \frac{\ell''(\hat{\theta}^{MLE}; X_1..X_n)}{\ell''(\theta; X_1..X_n)} \xrightarrow{p} 1$ by CMT and $\hat{\theta}^{MLE} \xrightarrow{p} \theta$

$A_2 \xrightarrow{p} 1$ by CMT and $\frac{1}{n} \sum -\ell''(\theta; X_i) \xrightarrow{p} E[-\ell''(\theta; X_i)] = I(\theta)$

$\Rightarrow A_1 A_2 \xrightarrow{p} 1$ by Slutsky's A

$\Rightarrow \hat{\Lambda} \xrightarrow{d} \chi^2_1$ by Slutsky's A

$\Rightarrow$ we have a test ole LR test (LRT) e.g. HP: $X_1,\ldots,X_n \overset{iid}{\sim} \text{Bern}(\theta)$, $\hat{\theta}^{mle} = \bar{X}$

By wald test AND score test, the test statistic was $\hat{z}|H_0 = \dfrac{\hat{\theta}^{mle} - \theta_0}{\sqrt{\dfrac{\theta_0(1-\theta_0)}{n}}}$

By the LRT,

$$\hat{LR} = \prod_{i=1}^{n} \frac{\mathcal{L}(\bar{X}; X_1 \ldots X_n)}{\mathcal{L}(\theta_0; X_1 \ldots X_n)} = \prod_{i=1}^{n} \frac{\bar{X}^{x_i}(1-\bar{X})^{1-x_i}}{\theta_0^{x_i}(1-\theta_0)^{1-x_i}} = \left(\frac{\bar{X}}{\theta_0}\right)^{\sum X_i} \left(\frac{1-\bar{X}}{1-\theta_0}\right)^{n-\sum X_i}$$

$$\hat{\Lambda} = 2\ln(\hat{LR}) = 2\left(\sum X_i \ln\left(\frac{\bar{X}}{\theta_0}\right) + \left(n - \sum X_i\right) \ln\left(\frac{1-\bar{X}}{1-\theta_0}\right)\right)$$

which is not the same test statistic!

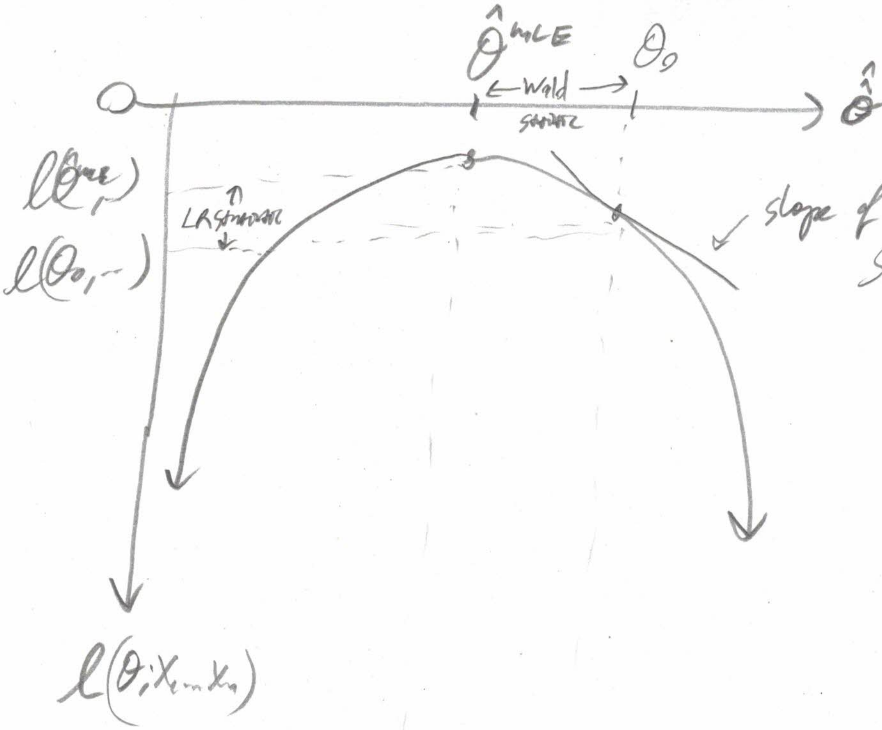
Calculate $\hat{\Lambda}$. The RET is always 1-sided because the difference in log-lik is always ≥ 0, $F_{\chi^2_1}(c) = .95 \Rightarrow c \approx 3.84$

$\Rightarrow RET = [0, 3.84]$

Flip a coin $n=100$ times, $\sum X_i = 61$. Test $H_a: \theta \neq .5$ at $\alpha = 5\%$.

$$\hat{\Lambda} = 2\left(61 \ln\left(\frac{.61}{.5}\right) + 39 \ln\left(\frac{.39}{.5}\right)\right) = 4.88 \notin RET \Rightarrow \text{Reject } H_0!$$

Visualizing the Wald, Score and LR tests... for $H_0: \theta = \theta_0$



slope of tangent line is score statistic

They all measure the same thing in slightly different ways. Sometimes those ways are exactly equivalent, sometimes not. Sometimes one way is more convenient. Sometimes one has higher power. All asymptotic...