Stat 563
5-26-26

We were here:

(1)

$$\sqrt{n}\,(\hat{\theta}-\theta) = \frac{\frac{1}{\sqrt{n}}\,\ell'(\theta)}{-\frac{1}{n}\ell''(\theta) - \frac{1}{n}\ell'''(\theta^*)\,\frac{\hat{\theta}-\theta}{2}}$$

Numerator: $L(\theta) = \prod_{i=1}^{n} f(x_i|\theta)$

$$\ell(\theta) = \sum_{i=1}^{n} \ln f(x_i|\theta)$$

$$\ell'(\theta) = \sum_{i=1}^{n} \underbrace{\frac{\partial}{\partial\theta}\ln f(x_i|\theta)}_{Z_i}$$

(2)

Denominator:

1st term: $-\frac{1}{n}\ell''(\theta) = -\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta^2}\ln f(x_i|\theta)$

$$= \frac{1}{n}\sum_{i=1}^{n} \underbrace{-\frac{\partial^2}{\partial\theta^2}\ln f(x_i|\theta)}_{W_i}$$

The WLLN says $\frac{1}{n}\sum_{i=1}^{n} W_i \xrightarrow{P} E[W_i] = I(\theta)$

③

2^{nd} term:

$$-\frac{1}{n} \ell'''(\theta^*) \frac{\hat{\theta}-\theta}{2}$$

Claim: This converges in probability to 0

If we can show $\hat{\theta} \xrightarrow{P} \theta$, then this claim is justified.

That is, we still need to show that, under our regularity conditions, MLEs are consistent.

④

Recall from March that Z_i is a score function.

So $E[Z_i] = 0$ and $V[Z_i] = I(\theta)$

By the Central Limit Theorem,

$$\frac{\ell'(\theta) - 0}{\sqrt{n}\sqrt{I(\theta)}} \xrightarrow{\mathcal{D}} N(0,1)$$

(5)

We had:

$$\sqrt{n}(\hat{\theta}-\theta) = \frac{\frac{1}{\sqrt{n}} l'(\theta) \xrightarrow{D} N(0, I(\theta))}{\underbrace{-\frac{1}{n} l''(\theta)}_{\xrightarrow{P} I(\theta)} - \underbrace{\frac{1}{n} l'''(\theta^*) \frac{\hat{\theta}-\theta}{2}}_{\xrightarrow{P} 0}}$$

$$\xrightarrow{D} N\left(0, \frac{1}{I(\theta)}\right)$$

by Slutsky's Theorem //

(6)

$\therefore$ For large n,

$$\hat{\theta} \approx N\left(\theta, \frac{1}{n I(\theta)}\right)$$

<u>Sketch of proof of consistency</u>

$$\hat{\theta} \text{ maximizes } L(\theta) = \prod_{i=1}^{n} f(x_i \mid \theta)$$

$$\hat{\theta} \text{ maximizes } \underbrace{\ln L(\theta)}_{l(\theta)} = \sum_{i=1}^{n} \ln f(x_i \mid \theta)$$

(7)

So $\hat{\theta}$ maximizes $\frac{1}{n}l(\theta) = \frac{1}{n}\sum_{i=1}^{n} \ln f(X_i|\theta)$

By WLLN, $\frac{1}{n}\sum_{i=1}^{n} \ln f(X_i|\theta) \xrightarrow{P} E[\ln f(X_i|\theta)]$

What value maximizes $E[\ln f(X_i|\theta)]$?

Let θ' be any value other than the true value of θ.

$E[\ln f(X|\theta')] - E[\ln f(X|\theta)]$

(8)

$= E\left[\ln \frac{f(X|\theta')}{f(X|\theta)}\right]$

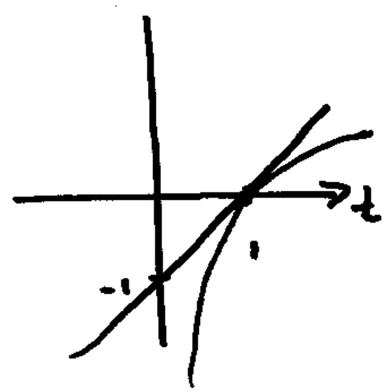
$\le E\left[\frac{f(X|\theta')}{f(X|\theta)} - 1\right]$

$= \int_{-\infty}^{\infty} \left[\frac{f(x|\theta')}{f(x|\theta)} - 1\right] f(x|\theta)\,dx$

$= \int_{-\infty}^{\infty} f(x|\theta')\,dx - \int_{-\infty}^{\infty} f(x|\theta)\,dx$

$= 1 - 1 = 0$

$\ln t \le t - 1$



(9)

So $E[\ln f(X_i|\theta)]$ is maximized at the true value of θ.

[Big gap]

$$\frac{1}{n}\sum_{i=1}^{n} \ln f(x_i|\theta) \xrightarrow{P} E[\ln f(X_i|\theta)]$$

$\hat{\theta}$ maximizes this

θ maximizes

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

(10)

<u>Theorem</u>: Assuming the same regularity conditions as before,

Under H_0, $\quad -2\ln\Lambda \xrightarrow{D} \chi^2_1$

<u>Proof</u>: Write the first order Taylor series, with remainder, for $l(\hat{\theta})$, expanded around θ_0.

$$l(\hat{\theta}) = l(\theta_0) + l'(\theta_0)(\hat{\theta} - \theta_0) + l''(\theta^*)\frac{(\hat{\theta}-\theta_0)^2}{2}$$

where θ^* lies between $\hat{\theta}$ and θ_0

(11)

Under H_0, $\hat{\theta} \xrightarrow{P} \theta_0$

Since θ^* lies between $\hat{\theta}$ and θ_0,

$$\theta^* \xrightarrow{P} \theta_0$$

Then $\quad l''(\theta^*) \xrightarrow{P} l''(\theta_0)$ (By a regularity condition)

Write $\quad l''(\theta^*) = l''(\theta_0) + \underbrace{R_0}_{\downarrow P \; 0}$

(12)

$$-\frac{1}{n} l''(\theta^*) = -\frac{1}{n} l''(\theta_0) - \frac{1}{n} R_0$$

$$= \underbrace{-\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta_0^2} \ln f(x_i \mid \theta_0)}_{\frac{1}{n}\sum_{i=1}^{n} -\frac{\partial^2}{\partial \theta_0^2} \ln f(x_i \mid \theta_0)} - \frac{1}{n} R_0$$

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n} -\frac{\partial^2}{\partial \theta_0^2} \ln f(x_i \mid \theta_0)}_{\downarrow P}$$

$$E\left[-\frac{\partial^2}{\partial \theta_0^2} \ln f(X \mid \theta_0)\right] = I(\theta)$$

(13)

So write

$$-\frac{1}{n} l''(\theta^*) = I(\theta) + \underset{\substack{\downarrow P \\ 0}}{R_1} - \frac{1}{n} R_0$$

$$l''(\theta^*) = -n I(\theta) - n R_1 - R_0 \quad \star$$

From previous theorem, under H_0

$$\sqrt{n}(\hat{\theta} - \theta_0) = \frac{\frac{1}{\sqrt{n}} l'(\theta_0)}{-\frac{1}{n} l''(\theta_0) - \frac{1}{n} l'''(\theta^*) \frac{\hat{\theta} - \theta_0}{2}}$$

(14)

So

$$\frac{1}{\sqrt{n}} l'(\theta_0) = \overbrace{-\frac{1}{n} l''(\theta_0)}^{\xrightarrow{P} I(\theta_0)} \sqrt{n}(\hat{\theta} - \theta_0)$$

$$- \frac{1}{n} l'''(\theta^*) \frac{\hat{\theta} - \theta_0}{2} \sqrt{n} \underbrace{(\hat{\theta} - \theta_0)}_{\substack{\downarrow P \\ 0}}$$

$$= I(\theta_0) \sqrt{n}(\hat{\theta} - \theta_0) + \underset{\substack{\downarrow P \\ 0}}{R_2} \quad \star\star$$

Taylor series:

$$\frac{1}{\sqrt{n}} [l(\hat{\theta}) - l(\theta_0)] = \frac{1}{\sqrt{n}} \left[\underline{l'(\theta_0)} (\hat{\theta} - \theta_0) + \underline{l''(\theta^*)} \frac{(\hat{\theta} - \theta_0)^2}{2} \right]$$

Using ★ and ★★ (15)

$$= I(\theta_0)\sqrt{n}\frac{(\hat{\theta} - \theta_0)^2}{2} + \underset{\substack{\downarrow P \\ 0}}{R_3} \quad \text{(collects } R_0, R_1, R_2)$$

$$\frac{1}{\sqrt{n}}[l(\hat{\theta}) - l(\theta_0)]$$

$$\parallel$$

$$\frac{1}{\sqrt{n}}[\ln L(\hat{\theta}) - \ln L(\theta_0)] = \frac{1}{\sqrt{n}} \ln \frac{L(\hat{\theta})}{L(\theta_0)}$$

$$= -\frac{1}{\sqrt{n}} \ln \frac{L(\theta_0)}{L(\hat{\theta})}$$

$$= -\frac{1}{\sqrt{n}} \ln \Lambda$$

(16)

So $-\frac{1}{\sqrt{n}} \ln \Lambda = I(\theta_0)\sqrt{n}\frac{(\hat{\theta} - \theta_0)^2}{2} + R_3$

$$-2 \ln \Lambda = n I(\theta_0)(\hat{\theta} - \theta_0)^2 + R_4$$

$$= \underbrace{\left[\frac{\hat{\theta} - \theta_0}{\sqrt{\frac{1}{n I(\theta_0)}}}\right]^2}_{\substack{\downarrow \mathcal{D} \\ \chi^2_1}} + \underset{\substack{\downarrow P \\ 0}}{R_4}$$

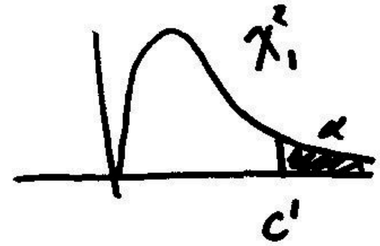
By Slutsky's Theorem, then

$$-2 \ln \Lambda \xrightarrow{\mathcal{D}} \chi^2_1$$

for large n, Rejecting H_0 when $\Lambda \le c$

is now equivalent to rejecting H_0

when $-2 \ln \Lambda \ge c'$



10.19 Another way in which underlying assumptions can be violated is if there is correlation in the sampling, which can seriously affect the properties of the sample mean. Suppose we introduce correlation in the case discussed in Exercise 10.2.1; that is, we observe $X_1, \dots, X_n$, where $X_i \sim \mathrm{n}(\theta, \sigma^2)$, but the X_is are no longer independent.

(a) For the equicorrelated case, that is, $\mathrm{Corr}(X_i, X_j) = \rho$ for $i \neq j$, show that

$$\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n} + \frac{n-1}{n}\rho\sigma^2,$$

so $\mathrm{Var}(\bar{X}) \not\to 0$ as $n \to \infty$.

(b) If the X_is are observed through time (or distance), it is sometimes assumed that the correlation decreases with time (or distance), with one specific model being $\mathrm{Corr}(X_i, X_j) = \rho^{|i-j|}$. Show that in this case

$$\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n} + \frac{2\sigma^2}{n^2}\frac{\rho}{1-\rho}\left(n + \frac{1-\rho^n}{1-\rho}\right),$$

so $\mathrm{Var}(\bar{X}) \to 0$ as $n \to \infty$. (See Miscellanea 5.8.2 for another effect of correlation.)

(c) The correlation structure in part (b) arises in an *autoregressive AR(1) model*, where we assume that $X_{i+1} = \rho X_i + \delta_i$, with δ_i iid $\mathrm{n}(0, 1)$. If $|\rho| < 1$ and we define $\sigma^2 = 1/(1-\rho^2)$, show that $\mathrm{Corr}(X_1, X_i) = \rho^{i-1}$.