MAXIMUM LIKELIHOOD ESTIMATORS & HYPOTHESIS TESTING

PRELIMINARIES

CONSIDER A RANDOM VARIABLE y AND A SAMPLE OF n OBSERVATIONS:

$$Y = \{y_1, y_2, \dots, y_n\}$$

SUPPOSE THAT THE OBSERVATIONS ARE INDEPENDENTLY DRAWN FROM SOME PROBABILITY DISTRIBUTION WITH PARAMETER θ. FOR EXAMPLE THE DISTRIBUTION MAY BE THE NORMAL DISTRIBUTION WITH MEAN μ & VARIANCE σ^2. THEN THE JOINT PDF OF THE SAMPLE IS

$$f(y_1, y_2, \dots, y_n; \mu, \sigma^2) = \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) e^{-\frac{(y_i-\mu)^2}{2\sigma^2}}$$

$$= (2\pi)^{-n/2} (\sigma^2)^{-n/2} e^{-\sum_i \frac{(y_i-\mu)^2}{2\sigma^2}}$$

CLEARLY, THIS JOINT PDF IS A FUNCTION OF THE SAMPLE Y & VARIES FROM SAMPLE TO SAMPLE.

NOW CONSIDER THE FOLLOWING QUESTION. WE KNOW THAT THE SAMPLE OBSERVATIONS CAME FROM SOME NORMAL DISTRIBUTION. OUR OBJECTIVE ~~IS~~ IS TO DETERMINE WHICH PARTICULAR DISTRIBUTION (i.e. WHICH PAIR OF μ & σ^2) IS MOST LIKELY TO HAVE GENERATED THIS SAMPLE DATA SET.

FOR THIS WE WRITE THE SAME JOINT PDF AS A FUNCTION OF THE PARAMETERS μ & σ^2 :

$$\mathcal{L}(\mu, \sigma^2; y_1, y_2, \cdots, y_n)$$

THIS IS KNOWN AS THE LIKELIHOOD FUNCTION.

THE MAXIMUM LIKELIHOOD PROCEDURE MAXIMIZES THE LIKELIHOOD FUNCTION $\mathcal{L}$ WITH RESPECT TO THE ~~LIKELI~~ PARAMETERS (μ, σ^2).

THE 2- VARIABLE REGRESSION

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

$$u_i \sim N(0, \sigma^2)$$

$$u_i = Y_i - \beta_1 - \beta_2 X_i$$

RECALL $$p(y_i) = p(u_i) \cdot \left| \frac{\partial u_i}{\partial y_i} \right|$$

$$\frac{\partial u_i}{\partial y_i} = 1$$

$$p(y_i) = p(Y_i - \beta_1 - \beta_2 X_i)$$

$$p(u_1, u_2, \ldots u_n) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\sum_i \frac{u_i^2}{2\sigma^2}}$$

$$p(y_1, y_2, \ldots y_n) = (2\pi)^{-n/2} (\sigma^2)^{-n/2} e^{-\frac{\sum_i (Y_i - \beta_1 - \beta_2 X_i)^2}{2\sigma^2}}$$

$$= \mathcal{L}(\beta_1, \beta_2, \sigma^2; y_1, y_2, \ldots, y_n).$$

NEXT CONSIDER THE LOG LIKELIHOOD FUNCTION

$$\ln \mathcal{L} = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum(Y_i - \beta_1 - \beta_2 X_i)^2$$

MAXIMIZING $\ln \mathcal{L}$ IS EQUIVALENT TO MAXIMIZING $\mathcal{L}$.

FOCS:

$$\frac{\partial \ln \mathcal{L}}{\partial \beta_1} = \frac{1}{2\sigma^2} 2 \sum(Y_i - \beta_1 - \beta_2 X_i) = 0$$

$$\Rightarrow \boxed{\frac{\sum(Y_i - \beta_1 - \beta_2 X_i)}{\sigma^2} = 0} \quad (A)$$

$$\frac{\partial \ln \mathcal{L}}{\partial \beta_2} = \boxed{\frac{\sum X_i(Y_i - \beta_1 - \beta_2 X_i)}{\sigma^2} = 0} \quad (B)$$

$$\frac{\partial \ln \mathcal{L}}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum(Y_i - \beta_1 - \beta_2 X_i)^2 = 0$$

$$\Rightarrow \boxed{\frac{\sum(Y_i - \beta_1 - \beta_2 X_i)^2}{\sigma^2} = n} \quad (C)$$

NOTE

$$\boxed{\begin{aligned} (A) &\Rightarrow \sum(Y_i - \beta_1 - \beta_2 X_i) = 0 \\ (B) &\Rightarrow \sum X_i(Y_i - \beta_1 - \beta_2 X_i) = 0 \end{aligned}}$$

THESE ARE EXACTLY THE SAME AS THE FOCS FOR OLS. THUS

SOLUTION OF A + B YIELDS

$$\boxed{\begin{aligned} \beta_2^* &= \frac{\Sigma (x_i-\bar{x})(y_i-\bar{y})}{\Sigma (x_i-\bar{x})^2} \quad \& \\ \beta_1^* &= \bar{y} - \beta_2^* \bar{x} \end{aligned}}$$

THAT IS β_1^* & β_2^* ARE THE SAME AS $(\hat{\beta}_1, \hat{\beta}_2)$. HENCE

$$y_i - \beta_1^* - \beta_2^* x_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i = e_i$$

(THE OLS RESIDUAL).

HENCE, BY VARTUE OF A & B,

$$C \Rightarrow \frac{\Sigma e_i^2}{\sigma_*^2} = n \quad \& \quad \boxed{\sigma_*^2 = \frac{\Sigma e_i^2}{n}}$$

NOTE THAT THE MLE OF σ^2 IS DIFFERENT FROM THE OLS BASED ESTIMATOR $\hat{\sigma}^2$.

THUS, $\ln \mathcal{L}^* = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\left(\frac{SSE}{n}\right) - \frac{1}{2}\frac{(SSE)}{\left(\frac{SSE}{n}\right)}$

$$= -\frac{n}{2}\ln 2\pi - \frac{n}{2} - \frac{n}{2}\ln\left(\frac{SSE}{n}\right) = -\frac{n}{2}\ln 2\pi - \frac{n}{2} + \frac{n}{2}\ln n - \frac{n}{2}\ln S$$

DEFINE $\frac{SSE}{\cancel{\sigma}} = Q$ & $-\frac{n}{2}\ln 2\pi - \frac{n}{2} + n\ln A = c$

THEN $\ln \mathcal{L}^* = c - \frac{n}{2}\ln\frac{SSE}{n} = \boxed{c - \frac{n}{2}\ln Q}$

THE LIKELIHOOD RATIO TEST

LET $ESS_R = Q_R$ BE THE ERROR SUM OF SQUARES FROM THE RESTRICTED MODEL

& $ESS_U = Q_U$ BE THE ERROR SUM OF SQUARES FROM THE UNRESTRICTED MODEL

THEN

$$\ln \mathcal{L}^*_R = c - \frac{n}{2} \ln Q_R$$

$$\ln \mathcal{L}^*_U = c - \frac{n}{2} \ln Q_u$$

$$\ln \mathcal{L}^*_R - \ln \mathcal{L}^*_u$$

$$= -\frac{n}{2}\left[\ln Q_R - \ln Q_u\right]$$

$$= -\frac{n}{2} \ln \left(\frac{Q_R}{Q_u}\right)$$

THUS THE LIKELIHOOD RATIO IS

$$\frac{\mathcal{L}^*_R}{\mathcal{L}^*_u} = \left(\frac{Q_R}{Q_u}\right)^{-n/2} = \lambda$$

$$-2 \ln \lambda = n\left(\ln Q_R - \ln Q_u\right) \sim \chi^2_p$$

WHERE p = # OF RESTRICTIONS.

CONSIDER THE RESTRICTION $\beta_2 = 0$
IN THE MODEL

$$Y_i = \beta_1 + \beta_2 X_i + u_i .$$

UNDER THE RESTRICTION,

$$Y_i = \beta_1 + u_i$$

THEN $\hat{\beta}_1 = \bar{Y}$ &

$$Q_R = (Y_i - \bar{Y})^2 \equiv S_{yy} .$$

ALSO IN THE UNRESTRICTED MODEL

$$1 - R^2 = \frac{SSE}{SST} = \frac{Q_U}{Q_R}$$

THUS

$$\frac{Q_R}{Q_U} \equiv \frac{1}{1-R^2}$$

$$\& \quad -2\ln\lambda = -n\ln(1-R^2) = n\ln\left(\frac{1}{1-R^2}\right) \sim \chi^2_1 .$$

THE WALD TEST

EXCEPT IN SIMPLE CASE LIKE THE ONE CONSIDERED ABOVE THE LR TEST REQUIRES ONE TO ESTIMATE THE RESTRICTED & THE UNRESTRICTED MODELS TO GET $\mathcal{L}^*_U$ & $\mathcal{L}^*_R$ FOR COMPUTING THE TEST STATISTIC. THE WALD TEST IS BASED ON THE PROPERTY THAT WHEN $u \sim N(0, \Sigma_i)$ UNDER THE RESTRICTION $R\beta = q$,

$$(R\hat{\beta} - q)' \, V(R\hat{\beta} - q)(R\hat{\beta} - q) \sim \chi^2_p \quad \text{WHERE}$$

p = # OF RESTRICTIONS.

IN THIS SIMPLE CASE, UNDER $H_0: \beta_2 = 0$

$$W = \frac{\hat{\beta}_2^{\,2}}{\operatorname{Var} \hat{\beta}_2} \sim \chi^2_1$$

THUS $$W = \frac{\hat{\beta}_2^{\,2}}{\dfrac{\sigma^2}{\sum (x_i - \bar{x})^2}} = \frac{\hat{\beta}_2^{\,2} \sum (x_i - \bar{x})^2}{\sigma^2} \sim \chi^2_1$$

IF WE USE $\hat{\sigma}^2_* = \dfrac{\sum e^2}{n}$ [THE MLE OF σ^2]

$$W = \left[\frac{\hat{\beta}_2^{\,2} \sum (x_i - \bar{x})^2}{\sum e_i^2} \right] n$$

$$= \left[\frac{SSR}{SSE} \right] n = \left[\frac{SSR/SST}{SSE/SST} \right] n$$

$$W = \left[\frac{R^2}{1 - R^2} \right] n \sim \chi^2_1$$

THE LAGRANGE MULTIPLIER TEST

IF WE USE THE MLE OF σ^2 FROM THE RESTRICTED MODEL,

$$LM = \frac{\hat{\beta}_2^{\,2} \sum (x_i - \bar{x})^2}{\sigma^2_{*R}} = \frac{\hat{\beta}_2^{\,2} \sum (x_i - \bar{x})^2}{\sum (Y_i - \bar{Y})^2 / n}$$

$$= \frac{SSR}{SST} \cdot n = \boxed{n R^2 \sim \chi^2_1}.$$

NOTE THAT

$$\boxed{\begin{aligned} LR &= n \ln\left(\frac{1}{1-R^2}\right) \\ W &= nR^2/(1-R^2) \\ LM &= nR^2 \end{aligned}}$$

CLEARLY, BECAUSE $0 \leq R^2 \leq 1$,

$$0 \leq 1-R^2 \leq 1 \quad \Rightarrow \quad \frac{1}{1-R^2} \geq 1$$

THUS, $W \geq LM$

FURTHER, GIVE THAT

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \quad \text{, FOR } x \geq 0$$

$$e^x \geq 1 + x \quad \Rightarrow$$

$$x \geq \ln(1+x).$$

$$\frac{LR}{n} = \ln\left(\frac{1}{1-R^2}\right)$$

$$\frac{W}{n} = \frac{R^2}{1-R^2}$$

$$\frac{W}{n} + 1 = \frac{1}{1-R^2}$$

$$\ln\left(\frac{W}{n}+1\right) = \ln \frac{1}{1-R^2}$$

THUS

$$\frac{W}{n} \geq \ln\left(\frac{W}{n}+1\right)$$

$$\Rightarrow \quad \frac{W}{n} \geq \frac{LR}{n} \quad \Rightarrow \quad \boxed{W \geq LR.}$$

FINALLY

$$\ln(1+x) \geq \frac{x}{1+x}$$

DEFINE

$$\frac{LM}{n} = \frac{W/n}{1+W/n}$$

$$\frac{LR}{n} = \ln\left(1+\frac{W}{n}\right)$$

HENCE

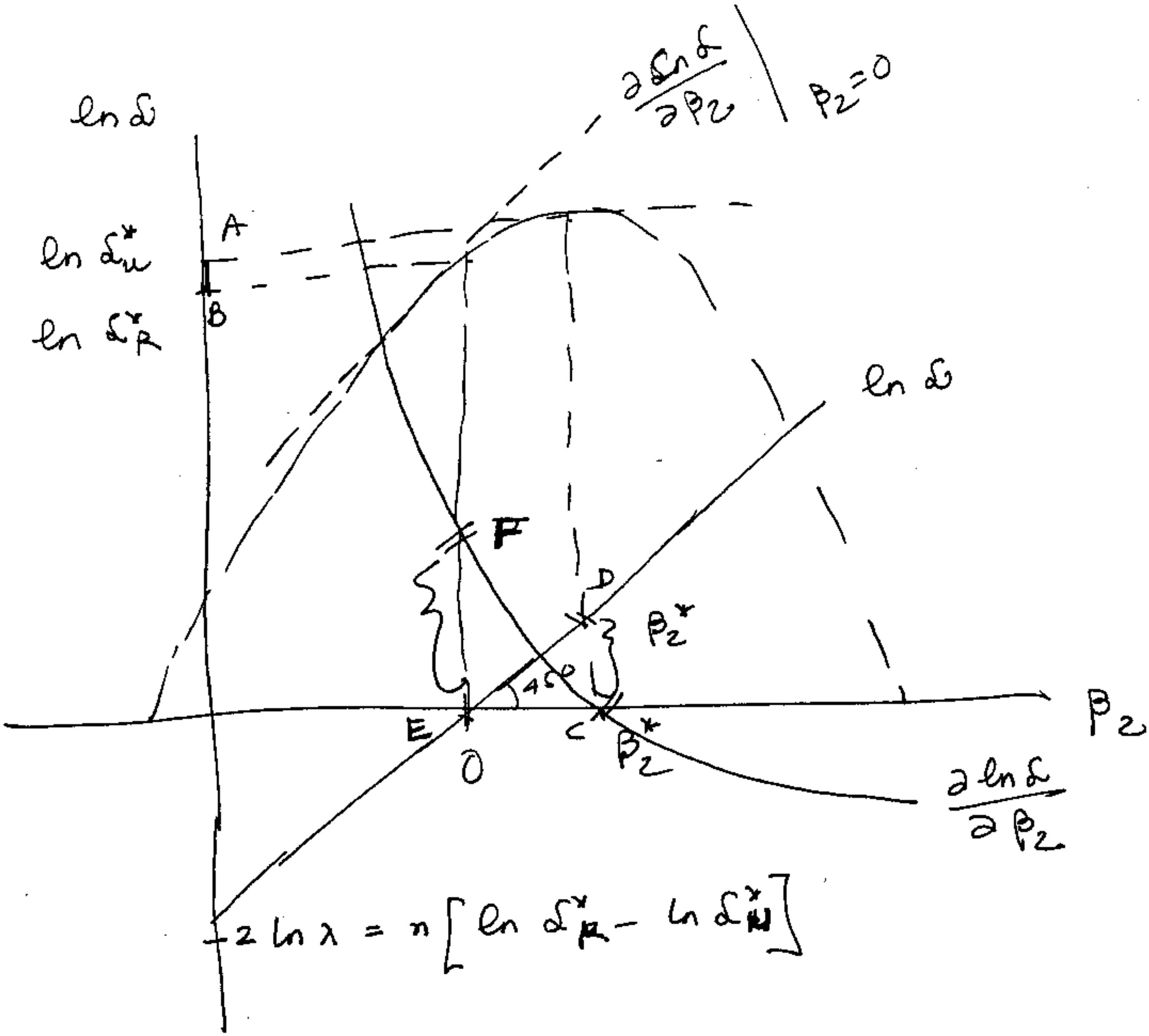
$$\frac{LR}{n} \geq \frac{LM}{n}.$$

THUS

$$\frac{W}{n} \geq \frac{LR}{n} \geq \frac{LM}{n}$$

$$\Rightarrow \quad \boxed{W \geq LR \geq LM}$$

BECAUSE EACH OF THE TEST STATISTICS IS $\sim \chi^2_1$,
IF $LM > \chi^2_{1\alpha}$, BOTH LR & W
ARE ALSO SIGNIFICANT.



LR : IS AB SIGNIFICANTLY DIFFERENT FROM 0 ?

W : IS CD SIGNIFICANTLY DIFFERENT FROM 0 ?

LM : IS EF SIGNIFICANTLY DIFFERENT FROM 0 ?