

# MAT-206: Test basados en la verosimilitud

**Felipe Osorio**

`fosorios.mat.utfsm.cl`

Departamento de Matemática, UTFSM



## Problema:

Deseamos probar hipótesis (no lineales) de la forma:

$$H_0 : \mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}, \quad \text{versus} \quad H_1 : \mathbf{g}(\boldsymbol{\theta}) \neq \mathbf{0},$$

donde  $\mathbf{g} : \mathbb{R}^p \rightarrow \mathbb{R}^q$ , tal que  $\mathbf{G}(\boldsymbol{\theta}) = \partial \mathbf{g}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}^\top$  es una matriz  $q \times p$  con rango  $q$ . En otras palabras, bajo  $H_0$  deseamos resolver el problema restringido:

$$\max_{\boldsymbol{\theta} \in \Theta} \ell_n(\boldsymbol{\theta}), \quad \text{sujeto a: } \mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}.$$



## Observación:

Hipótesis del tipo

$$H_0 : g(\theta) = \mathbf{0}, \quad \text{vs.} \quad H_1 : g(\theta) \neq \mathbf{0},$$

se dicen en **forma implícita**.

Mientras que,

$$H_0 : \theta = h(\gamma), \quad \text{vs.} \quad H_1 : \theta \neq h(\gamma),$$

con espacio paramétrico nulo,

$$\Theta_0 = \{\theta : \theta = h(\gamma), \gamma \in \Gamma\}, \quad h(\gamma) = (h_1(\gamma), \dots, h_p(\gamma))^\top,$$

está en **forma explícita**. Finalmente, la hipótesis nula definida como:

$$H_0 : \{\theta : \exists a \in \mathcal{A} \subset \mathbb{R}^k, g(\theta, a) = \mathbf{0}\},$$

está escrita en **forma mixta**.<sup>1</sup>

---

<sup>1</sup>Gourieroux y Monfort (1989). *Econometric Theory* 5, 63-82.



## Resultado 1 (test de Wald):<sup>2</sup>

El **test de Wald** para probar  $H_0 : \mathbf{g}(\boldsymbol{\theta}) = \mathbf{0}$ , es definido por la región crítica,

$$\{W_n \geq \chi_{1-\alpha}^2(q)\},$$

donde

$$W_n = n\mathbf{g}^\top(\hat{\boldsymbol{\theta}}_n)[\mathbf{G}(\hat{\boldsymbol{\theta}}_n)\mathcal{F}^{-1}(\hat{\boldsymbol{\theta}}_n)\mathbf{G}^\top(\hat{\boldsymbol{\theta}}_n)]^{-1}\mathbf{g}(\hat{\boldsymbol{\theta}}_n),$$

y bajo  $H_0$  es asintóticamente de tamaño  $\alpha$ .

---

<sup>2</sup>Wald (1943). Transactions of the American Mathematical Society **54**, 426-482.



## *Demostración:*

Sabemos que

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N_p(\mathbf{0}, \mathcal{F}^{-1}(\theta_0)).$$

De ahí que

$$\sqrt{n}(g(\hat{\theta}_n) - g(\theta_0)) \xrightarrow{D} N_q(\mathbf{0}, G(\theta_0)\mathcal{F}^{-1}(\theta_0)G^\top(\theta_0)).$$

Bajo  $H_0$ , tenemos  $g(\theta_0) = \mathbf{0}$ . Luego,

$$\sqrt{n}g(\hat{\theta}_n) \xrightarrow{D} N_q(\mathbf{0}, G(\theta_0)\mathcal{F}^{-1}(\theta_0)G^\top(\theta_0)). \quad (\text{W.1})$$

Como  $G(\theta_0)$  es de rango fila completo, sigue que

$$\sqrt{n}[G(\theta_0)\mathcal{F}^{-1}(\theta_0)G^\top(\theta_0)]^{-1/2}g(\hat{\theta}_n) \xrightarrow{D} N_q(\mathbf{0}, I).$$



Note que  $\hat{\theta}_n$  es un estimador consistente de  $\theta_0$ . Haciendo,

$$\mathbf{Z}_n = \sqrt{n}[\mathbf{G}(\hat{\theta}_n)\mathcal{F}^{-1}(\hat{\theta}_n)\mathbf{G}^\top(\hat{\theta}_n)]^{-1/2}\mathbf{g}(\hat{\theta}_n) \xrightarrow{D} \mathbf{N}_q(\mathbf{0}, \mathbf{I}).$$

De ahí que, bajo  $H_0$

$$\begin{aligned} W_n &= n\mathbf{g}^\top(\hat{\theta}_n)[\mathbf{G}(\hat{\theta}_n)\mathcal{F}^{-1}(\hat{\theta}_n)\mathbf{G}^\top(\hat{\theta}_n)]^{-1}\mathbf{g}(\hat{\theta}_n) \\ &= \mathbf{Z}_n^\top \mathbf{Z}_n \xrightarrow{D} \chi^2(q). \end{aligned}$$



## Observación:

Para hipótesis de la forma  $H_0 : \theta_1 = \theta_1^0$ , donde  $\theta = (\theta_1^\top, \theta_2^\top)^\top$ . Es este caso el estadístico de Wald asume la forma:

$$W_n = n(\hat{\theta}_{1n} - \theta_1)^\top K_{11}^{-1}(\hat{\theta}_n)(\hat{\theta}_{1n} - \theta_1),$$

con

$$K(\theta) = \begin{pmatrix} K_{11}(\theta) & K_{12}(\theta) \\ K_{21}(\theta) & K_{22}(\theta) \end{pmatrix} = \mathcal{F}^{-1}(\theta).$$

Note que

$$K_{11}(\theta) = (\mathcal{F}_{11}(\theta) - \mathcal{F}_{12}(\theta)\mathcal{F}_{22}^{-1}(\theta)\mathcal{F}_{21}(\theta))^{-1},$$

con

$$\mathcal{F}(\theta) = \begin{pmatrix} \mathcal{F}_{11}(\theta) & \mathcal{F}_{12}(\theta) \\ \mathcal{F}_{21}(\theta) & \mathcal{F}_{22}(\theta) \end{pmatrix}.$$



Sea  $\tilde{\theta}_n$  el MLE de  $\theta$  sujeto a la restricción  $g(\theta) = 0$ . La **función Langrangiana** asociada con el problema restringido es

$$\ell_n(\theta) - g^\top(\theta)\lambda,$$

y las **condiciones de primer orden** son

$$\begin{aligned}\frac{\partial \ell_n(\tilde{\theta}_n)}{\partial \theta} - G^\top(\tilde{\theta}_n)\tilde{\lambda}_n &= 0, \\ g(\tilde{\theta}_n) &= 0,\end{aligned}$$

donde  $\tilde{\lambda}_n$  es un vector de **multiplicadores de Lagrange**.





## Resultado 2 (test score o de multiplicadores de Lagrange):<sup>3</sup>

El estadístico score o de multiplicadores de Lagrange para probar la hipótesis  $H_0 : g(\theta_0) = 0$ , es dado por

$$\begin{aligned} R_n &= \frac{1}{n} \left( \frac{\partial \ell_n(\tilde{\theta}_n)}{\partial \theta} \right)^\top \mathcal{F}^{-1}(\tilde{\theta}_n) \left( \frac{\partial \ell_n(\tilde{\theta}_n)}{\partial \theta} \right) \\ &= \frac{1}{n} \tilde{\lambda}_n^\top G(\tilde{\theta}_n) \mathcal{F}^{-1}(\tilde{\theta}_n) G^\top(\tilde{\theta}_n) \tilde{\lambda}_n, \end{aligned}$$

y bajo  $H_0$  tiene distribución asintótica chi-cuadrado con  $q$  grados de libertad.

---

<sup>3</sup>Rao (1948). Proceedings of the Cambridge Philosophical Society **44**, 50-57.



## Demostración:

Considere una expansión de Taylor en torno de  $\theta_0$ ,<sup>4</sup> como

$$g(\hat{\theta}_n) \stackrel{a}{=} g(\theta_0) + \frac{\partial g(\theta_0)}{\partial \theta^\top} (\hat{\theta}_n - \theta_0),$$

y análogamente para  $g(\tilde{\theta}_n)$ . De ahí que

$$\sqrt{n}g(\hat{\theta}_n) \stackrel{a}{=} G(\theta_0)\sqrt{n}(\hat{\theta}_n - \theta_0),$$

$$\sqrt{n}g(\tilde{\theta}_n) \stackrel{a}{=} G(\theta_0)\sqrt{n}(\tilde{\theta}_n - \theta_0),$$

tomando diferencias, obtenemos

$$\begin{aligned}\sqrt{n}g(\hat{\theta}_n) - \sqrt{n}g(\tilde{\theta}_n) &\stackrel{a}{=} G(\theta_0)\sqrt{n}(\hat{\theta}_n - \theta_0) - G(\theta_0)\sqrt{n}(\tilde{\theta}_n - \theta_0) \\ &= G(\theta_0)\sqrt{n}(\hat{\theta}_n - \tilde{\theta}_n),\end{aligned}$$

como  $\tilde{\theta}_n$  es el MLE restringido, tenemos  $g(\tilde{\theta}_n) = 0$ . De este modo,

$$\sqrt{n}g(\hat{\theta}_n) \stackrel{a}{=} G(\theta_0)\sqrt{n}(\hat{\theta}_n - \tilde{\theta}_n).$$

(R.1)

---

<sup>4</sup>  $X \stackrel{a}{=} Y$  indica que  $X - Y = o_p(1)$



## Test basados en la verosimilitud

Por otro lado,

$$\frac{1}{\sqrt{n}} \frac{\partial \ell_n(\tilde{\theta}_n)}{\partial \theta} \stackrel{a}{=} \frac{1}{\sqrt{n}} \frac{\partial \ell_n(\theta_0)}{\partial \theta} - \mathcal{F}(\theta_0) \sqrt{n}(\tilde{\theta}_n - \theta_0),$$

y

$$0 = \frac{1}{\sqrt{n}} \frac{\partial \ell_n(\hat{\theta}_n)}{\partial \theta} \stackrel{a}{=} \frac{1}{\sqrt{n}} \frac{\partial \ell_n(\theta_0)}{\partial \theta} - \mathcal{F}(\theta_0) \sqrt{n}(\hat{\theta}_n - \theta_0). \quad (\text{R.2})$$

Tomando diferencias, obtenemos

$$\begin{aligned} \frac{1}{\sqrt{n}} \frac{\partial \ell_n(\tilde{\theta}_n)}{\partial \theta} &\stackrel{a}{=} \frac{1}{\sqrt{n}} \frac{\partial \ell_n(\theta_0)}{\partial \theta} - \frac{1}{\sqrt{n}} \frac{\partial \ell_n(\theta_0)}{\partial \theta} \\ &\quad - \mathcal{F}(\theta_0) \sqrt{n}(\tilde{\theta}_n - \theta_0) + \mathcal{F}(\theta_0) \sqrt{n}(\hat{\theta}_n - \theta_0) \\ &\stackrel{a}{=} \mathcal{F}(\theta_0) \sqrt{n}(\hat{\theta}_n - \tilde{\theta}_n). \end{aligned}$$

Por tanto,

$$\sqrt{n}(\hat{\theta}_n - \tilde{\theta}_n) \stackrel{a}{=} \mathcal{F}^{-1}(\theta_0) \frac{1}{\sqrt{n}} \frac{\partial \ell_n(\tilde{\theta}_n)}{\partial \theta}. \quad (\text{R.3})$$



# Test basados en la verosimilitud

Desde Ecuación (R.1), obtenemos

$$\sqrt{n}g(\hat{\theta}_n) \stackrel{a}{=} G(\theta_0)\sqrt{n}(\hat{\theta}_n - \tilde{\theta}_n) \stackrel{a}{=} G(\theta_0)\mathcal{F}^{-1}(\theta_0)\frac{1}{\sqrt{n}}\frac{\partial\ell_n(\tilde{\theta}_n)}{\partial\theta}.$$

Por la condición de primer orden, podemos escribir

$$\frac{\partial\ell_n(\tilde{\theta}_n)}{\partial\theta} = G^\top(\tilde{\theta}_n)\tilde{\lambda}_n. \quad (\text{R.4})$$

De este modo,

$$\begin{aligned}\sqrt{n}g(\hat{\theta}_n) &\stackrel{a}{=} G(\theta_0)\mathcal{F}^{-1}(\theta_0)G^\top(\tilde{\theta}_n)\frac{\tilde{\lambda}_n}{\sqrt{n}} \\ &\stackrel{a}{=} G(\theta_0)\mathcal{F}^{-1}(\theta_0)G^\top(\theta_0)\frac{\tilde{\lambda}_n}{\sqrt{n}}.\end{aligned}$$

Es decir,

$$\frac{\tilde{\lambda}_n}{\sqrt{n}} \stackrel{a}{=} [G(\theta_0)\mathcal{F}^{-1}(\theta_0)G^\top(\theta_0)]^{-1}\sqrt{n}g(\hat{\theta}_n).$$



Por (W.1), sigue que

$$\frac{\tilde{\boldsymbol{\lambda}}_n}{\sqrt{n}} \xrightarrow{D} N_q(\mathbf{0}, [\mathbf{G}(\boldsymbol{\theta}_0) \boldsymbol{\mathcal{F}}^{-1}(\boldsymbol{\theta}_0) \mathbf{G}^\top(\boldsymbol{\theta}_0)]^{-1}).$$

De ahí que la forma cuadrática

$$R_n = \frac{\tilde{\boldsymbol{\lambda}}_n^\top}{\sqrt{n}} \mathbf{G}(\tilde{\boldsymbol{\theta}}_n) \boldsymbol{\mathcal{F}}^{-1}(\tilde{\boldsymbol{\theta}}_n) \mathbf{G}^\top(\tilde{\boldsymbol{\theta}}_n) \frac{\tilde{\boldsymbol{\lambda}}_n}{\sqrt{n}} \xrightarrow{D} \chi^2(q).$$

Nuevamente por la condición en (R.4), podemos escribir:

$$R_n = \frac{1}{n} \left( \frac{\partial \ell_n(\tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}} \right)^\top \boldsymbol{\mathcal{F}}^{-1}(\tilde{\boldsymbol{\theta}}_n) \frac{\partial \ell_n(\tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}}.$$



## Resultado 3 (test de razón de verosimilitudes):<sup>5</sup>

El test de razón de verosimilitudes es definido por el estadístico

$$LR_n = 2(\ell_n(\hat{\theta}_n) - \ell_n(\tilde{\theta}_n)),$$

y bajo  $H_0$  tiene región crítica asintótica de tamaño  $\alpha$ , dada por

$$\{LR_n \geq \chi_{1-\alpha}^2(q)\}.$$

---

<sup>5</sup>Wilks (1938). The Annals of Mathematical Statistics 9, 60-62.



## *Demostración:*

Considere las expansiones de Taylor de  $\ell_n(\hat{\theta}_n)$  y  $\ell_n(\tilde{\theta}_n)$  en torno de  $\theta_0$ . Bajo  $H_0 : \mathbf{g}(\theta_0) = \mathbf{0}$ , estas expresiones son:

$$\begin{aligned}\ell_n(\hat{\theta}_n) &\stackrel{a}{=} \ell_n(\theta_0) + \left( \frac{\partial \ell_n(\theta_0)}{\partial \theta} \right)^\top (\hat{\theta}_n - \theta_0) - \frac{n}{2} (\hat{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0) (\hat{\theta}_n - \theta_0) \\ \ell_n(\tilde{\theta}_n) &\stackrel{a}{=} \ell_n(\theta_0) + \left( \frac{\partial \ell_n(\theta_0)}{\partial \theta} \right)^\top (\tilde{\theta}_n - \theta_0) - \frac{n}{2} (\tilde{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0) (\tilde{\theta}_n - \theta_0).\end{aligned}$$

Tomando diferencias, obtenemos

$$\begin{aligned}\ell_n(\hat{\theta}_n) - \ell_n(\tilde{\theta}_n) &\stackrel{a}{=} \ell_n(\theta_0) - \ell_n(\theta_0) + \left( \frac{\partial \ell_n(\theta_0)}{\partial \theta} \right)^\top (\hat{\theta}_n - \tilde{\theta}_n) \\ &\quad - \frac{n}{2} (\hat{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0) (\hat{\theta}_n - \theta_0) + \frac{n}{2} (\tilde{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0) (\tilde{\theta}_n - \theta_0)\end{aligned}$$



Es decir,

$$\begin{aligned} LR_n &= 2(\ell_n(\hat{\theta}_n) - \ell_n(\tilde{\theta}_n)) \\ &\stackrel{a}{=} 2\left(\frac{\partial \ell_n(\theta_0)}{\partial \theta}\right)^\top (\hat{\theta}_n - \tilde{\theta}_n) - n(\hat{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0)(\hat{\theta}_n - \theta_0) \\ &\quad + n(\tilde{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0)(\tilde{\theta}_n - \theta_0) \end{aligned} \tag{L.1}$$

Usando (R.3), tenemos

$$\frac{1}{\sqrt{n}} \frac{\partial \ell_n(\theta_0)}{\partial \theta} \stackrel{a}{=} \mathcal{F}(\theta_0) \sqrt{n}(\hat{\theta}_n - \theta_0).$$

Substituyendo en (L.1), sigue que

$$\begin{aligned} LR_n &\stackrel{a}{=} 2n(\hat{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0)(\hat{\theta}_n - \tilde{\theta}_n) - (\hat{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0)(\hat{\theta}_n - \theta_0) \\ &\quad + n(\tilde{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0)(\tilde{\theta}_n - \theta_0) \\ &\stackrel{a}{=} 2n(\hat{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0)(\hat{\theta}_n - \tilde{\theta}_n) - (\hat{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0)(\hat{\theta}_n - \theta_0) \\ &\quad + n(\tilde{\theta}_n - \hat{\theta}_n + \hat{\theta}_n - \theta_0)^\top \mathcal{F}(\theta_0)(\tilde{\theta}_n - \hat{\theta}_n + \hat{\theta}_n - \theta_0). \end{aligned}$$





Después de simple álgebra, obtenemos

$$LR_n \stackrel{a}{=} n(\hat{\boldsymbol{\theta}}_n - \tilde{\boldsymbol{\theta}}_n)^\top \mathcal{F}(\boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}_n - \tilde{\boldsymbol{\theta}}_n)$$

*Observación:*

Por Ecuación (R.3), sigue que

$$LR_n \stackrel{a}{=} \frac{1}{n} \left( \frac{\partial \ell_n(\tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}} \right)^\top \mathcal{F}^{-1}(\boldsymbol{\theta}_0) \frac{\partial \ell_n(\tilde{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}} = R_n.$$

Es decir,  $LR_n$  es asintóticamente equivalente a  $R_n$ . De ahí que, bajo  $H_0$  tienen la misma distribución asintótica chi-cuadrado.



## Definición 1 (test gradiente):<sup>6</sup>

El estadístico gradiente  $T_n$ , para probar la hipótesis nula  $H_0 : \theta = \theta_0$  versus  $H_1 : \theta \neq \theta_0$  es dado por

$$T_n = \mathbf{U}_n^\top(\theta_0)(\hat{\theta}_n - \theta_0),$$

y es asintóticamente chi-cuadrado con  $p$  grados de libertad.

### Demostración:

Sigue de notar que

$$\frac{1}{\sqrt{n}} \mathbf{U}_n(\theta_0) \xrightarrow{D} N_p(\mathbf{0}, \mathcal{F}(\theta_0)), \quad \sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N_p(\mathbf{0}, \mathcal{F}^{-1}(\theta_0)),$$

de ahí que

$$\left\{ \frac{1}{\sqrt{n}} \mathcal{F}^{-1/2}(\theta_0) \mathbf{U}_n(\theta_0) \right\}^\top \sqrt{n} \mathcal{F}^{1/2}(\theta_0) (\hat{\theta}_n - \theta_0) \xrightarrow{D} \chi^2(p).$$

---

<sup>6</sup>Terrell (2002). Computing Sciences and Statistics 34, 206-215.

## Resultado 4 (test de forma bilineal):<sup>7</sup>

El estadístico de forma bilineal  $BF_n$ , para probar la hipótesis  $H_0 : \mathbf{g}(\boldsymbol{\theta}_0) = \mathbf{0}$ , adopta la forma

$$BF_n = \tilde{\boldsymbol{\lambda}}_n^\top \mathbf{g}(\hat{\boldsymbol{\theta}}_n),$$

y bajo  $H_0$ ,  $BF_n$  tiene una distribución asintótica chi-cuadrado con  $q$  grados de libertad.

### *Demostración:*

Sabemos que

$$\begin{aligned} \frac{\tilde{\boldsymbol{\lambda}}_n}{\sqrt{n}} &\xrightarrow{D} N_q(\mathbf{0}, [\mathbf{G}(\boldsymbol{\theta}_0) \mathcal{F}^{-1}(\boldsymbol{\theta}_0) \mathbf{G}^\top(\boldsymbol{\theta}_0)]^{-1}), \\ \sqrt{n} \mathbf{g}(\hat{\boldsymbol{\theta}}_n) &\xrightarrow{D} N_q(\mathbf{0}, \mathbf{G}(\boldsymbol{\theta}_0) \mathcal{F}^{-1}(\boldsymbol{\theta}_0) \mathbf{G}^\top(\boldsymbol{\theta}_0)), \end{aligned}$$

y el resultado sigue.

---

<sup>7</sup>Crudu y Osorio (2020). Economics Letters **187**, 108885.

## Observación:

Podemos escribir el estadístico  $BF_n$  de forma equivalente como

$$BF_n = \mathbf{U}_n^\top (\tilde{\boldsymbol{\lambda}}_n) \mathbf{G}^+ \mathbf{g}(\hat{\boldsymbol{\theta}}_n),$$

donde  $\mathbf{G}^+ = \mathbf{G}^\top (\mathbf{G}\mathbf{G}^\top)^{-1}$  denota la inversa Moore-Penrose de  $\mathbf{G} = \mathbf{G}(\tilde{\boldsymbol{\theta}}_n)$ . Por (R.1), tenemos que

$$BF_n \stackrel{a}{=} \mathbf{U}_n^\top (\tilde{\boldsymbol{\theta}}_n) \mathbf{G}^+ \mathbf{G}(\hat{\boldsymbol{\theta}}_n - \tilde{\boldsymbol{\theta}}_n).$$

Además,

$$BF_n \stackrel{a}{=} \frac{1}{n} \tilde{\boldsymbol{\lambda}}_n^\top \mathbf{G}(\tilde{\boldsymbol{\theta}}_n) \mathcal{F}^{-1}(\tilde{\boldsymbol{\theta}}_n) \mathbf{G}^\top(\tilde{\boldsymbol{\theta}}_n) \tilde{\boldsymbol{\lambda}}_n = R_n.$$



# Referencias bibliográficas



Crudu, F., Osorio, F. (2020).

Bilinear form test statistics for extremum estimation.

*Economics Letters* **187**, 108885.



Rao, C.R. (1948).

Large sample tests of Statistical hypotheses concerning several parameters with applications to problems of estimation.

*Proceedings of the Cambridge Philosophical Society* **44**, 50-57.



Terrell, G.R. (2002).

The gradient statistic.

*Computing Sciences and Statistics* **34**, 206-215.



Wald, A. (1943).

Test of Statistical hypothesis concerning several parameters when the number of observations is large.

*Transactions of the American Mathematical Society* **54**, 426-482.



Wilks, S.S. (1938).

The large-sample distribution of the likelihood ratio for testing composite hypothesis.

*The Annals of Mathematical Statistics* **9**, 60-62.

