where $F_{u,v}$ denotes an F distribution with the degrees of freedom u for the numerator and v for the denominator. This F distribution is widely used for inference about the general linear hypotheses.

When $\mathbf{x}_i$ are concurrently sampled with y_i, we use the asymptotic distribution of the MLE for inference about the general linear hypothesis. The two most popular approaches for testing the null of linear hypothesis are the *Wald* and *likelihood ratio* tests.

If $\widehat{\boldsymbol{\theta}} \sim AN\left(\boldsymbol{\theta}, \frac{1}{n}\Sigma_\theta\right)$, then it follows from the properties of multivariate normal distribution that $K\widehat{\boldsymbol{\theta}} \sim N\left(K\boldsymbol{\theta}, \frac{1}{n}K\Sigma_\theta K^\top\right)$. Under the null, $K\widehat{\boldsymbol{\theta}} \sim AN\left(\mathbf{b}, \frac{1}{n}K\Sigma_\theta K^\top\right)$. By a slight abuse of notation, let Q_w^2 denote the *Wald* statistic obtained from (2.37) by replacing Σ_β with $\widehat{\Sigma}_\beta$. It follows from Slutsky's theorem that Q_w^2 has an asymptotic χ_l^2 distribution for large n. Alternatively, we can use the *likelihood ratio test* (LRT) to examine the general linear hypothesis.

Let $L_n(\boldsymbol{\theta})$ denote the likelihood function and let $\widehat{\boldsymbol{\theta}}$ the MLE of $\boldsymbol{\theta}$. Let $\widehat{\boldsymbol{\theta}}_0$ be the MLE of $\boldsymbol{\theta}$ under the null H_0 in (2.36). Then, the *likelihood ratio* based statistic, $-2\ln R\left(\widehat{\boldsymbol{\theta}}_0\right) = -2\ln\left(\frac{L_n(\widehat{\boldsymbol{\theta}}_0)}{L_n(\widehat{\boldsymbol{\theta}})}\right)$, has an asymptotic χ_l^2 distribution under H_0 as asserted by the following theorem.

Theorem 2. Under mild regularity assumptions,

$$-2\ln R\left(\widehat{\boldsymbol{\theta}}_0\right) = -2\ln\left(\frac{L_n\left(\widehat{\boldsymbol{\theta}}_0\right)}{L_n\left(\widehat{\boldsymbol{\theta}}\right)}\right) \to_d \chi_l^2, \quad \text{as } n \to \infty \tag{2.38}$$

Proof. Let $\mathbf{u}_n(\boldsymbol{\theta}) = \frac{1}{n}\frac{\partial l_n}{\partial \boldsymbol{\theta}}$. We prove (2.38) in three steps.

Case 1. First, assume $K = I_p$. In this case, the null in (2.36) reduces to $H_0 : \boldsymbol{\theta} = \mathbf{b}$. For notational brevity, we continue to use $\boldsymbol{\theta}$, but with the understanding that $\boldsymbol{\theta}$ is a known vector under H_0.

By (2.23), we have

$$\mathbf{0} = \mathbf{u}_n\left(\widehat{\boldsymbol{\theta}}\right) = \mathbf{u}_n(\boldsymbol{\theta}) + \frac{\partial}{\partial \boldsymbol{\theta}^\top}\mathbf{u}_n(\boldsymbol{\theta})\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) + o_p\left(n^{-\frac{1}{2}}\right)$$

It follows that

$$\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^\top \mathbf{u}_n(\boldsymbol{\theta}) = -\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^\top \frac{\partial}{\partial \boldsymbol{\theta}^\top}\mathbf{u}_n(\boldsymbol{\theta})\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) + \left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^\top o_p\left(n^{-\frac{1}{2}}\right)$$