UNDERSTANDING A LIKELIHOOD FLAT PROBLEM: INFERENCEs ON THE RATIO OF REGRESSION COEFFICIENTS IN LINEAR MODELS

ENTENDIENDO UN PROBLEMA DE VEROSIMILITUD PLANA: INFERENCIAS SOBRE EL COCIENTE DE COEFICIENTES DE REGRESIÓN EN MODELOS LINEALES

JORGE ESPÍNDOLA ZEPEDA, JOSE A. MONTOYA

Recibido 13-08-2021, aceptado 6-04-2022, versión final 13-06-2022
Research Paper

ABSTRACT: In this paper, we analyze a flat likelihood function shape that arises when performing inferences on the ratio of two regression coefficients in a linear regression model, parameter of interest in various applications. Due to this shape, infinite length likelihood-confidence intervals can be obtained. In the cases discussed here these likelihood-confidence intervals are related to the nested models problem, which is analyzed in detail through three illustrative simulated cases. It is essential to understand the shapes of the likelihood function in order to legitimately criticize likelihood inferences. This is of particular importance since the likelihood function is a key ingredient used in many inference methods.

KEYWORDS: Shape of the likelihood function; nested models; linear regression model; profile likelihood function.

RESUMEN: En este artículo se analiza una forma plana de la función de verosimilitud que surge cuando se realizan inferencias sobre la razón de dos coeficientes de regresión en un modelo lineal, parámetro de interés en diversas aplicaciones. Debido a esta forma pueden obtenerse intervalos de verosimilitud-confianza de longitud infinita. En los casos que se discuten aquí, estos intervalos de verosimilitud-confianza están relacionados con el problema de modelos anidados, que es analizado a detalle a través de la simulación de tres casos ilustrativos. Es fundamental comprender las formas de la función de verosimilitud para criticar de manera legítima las inferencias por verosimilitud. Esto es de particular importancia ya que la función de verosimilitud es un ingrediente clave utilizado en muchos métodos inferenciales.

PALABRAS CLAVE: Forma de la función de verosimilitud; modelos anidados; modelo de regresión lineal; función de verosimilitud perfil.


* Corresponding author: arturo.montoya@unison.mx
1. INTRODUCTION

Academic literature has consistently illustrated that statistical methods based on the likelihood function can lead to misleading, strange, unstable, preposterous or ridiculous estimates, (Harter & Moore, 1966; Breusch et al., 1997; Berger et al., 1999; Martins & Stedinger, 2000; Pewsey, 2000; Martins & Stedinger, 2001; El Adlouni et al., 2007; Tumlinson, 2015). The shape of the likelihood function is often cited as an underlying cause of these strange or unintuitive results. The criticisms of maximum likelihood estimates, associated with strange likelihood functions that grow rapidly to infinite and caused by the use of density functions that have singularities, are invalid. This issue is well known; see Montoya et al. (2009), Liu et al. (2015), and the references cited therein. However, there are also criticisms of the maximum likelihood estimation that occur quite often and that arise when the shape of the likelihood function becomes flat. In fact, flat likelihoods are used to promote integrated likelihoods (Berger et al., 1999; Ghosh et al., 2006), Bayesian posteriors (Tsionas, 2001), penalized methods (Li & Sudjianto, 2005; Cole et al., 2013; Lima & Cribari-Neto, 2019) and also new optimization methods (Frery et al., 2004).

Some authors have found that flat likelihoods can be related with an overparameterization of the model (Catchpole & Morgan, 1997), inappropriate reparametrizations (Farcomeni, & Tardella, 2012), embedded models (Cheng & Iles, 1990), limited amount and quality of experimental data (Raue et al., 2009; Kreutz et al., 2013), and even a poor model adjustment to the data, when the sample size is large (Sundberg, 2010). Although these studies are based on relevant statistical models (exponential family of distributions, three-parameters distributions, capture-recapture models and dynamical models), and all helped to clarify what is behind the problem of these flat likelihoods, additional work is needed to enhance our understanding.

This article is focused on the linear regression model. We consider the profile likelihood function to analyze the ratio of two regression coefficients, our parameter of interest, which naturally appears in many statistical applications that usually arise in biopharmaceutical and economic research, among others. This parameter can mean: a measure of the relative potency of a test drug versus a reference drug (Fay et al., 2009); the location of a turning point in a quadratic specification where the marginal impact of a regressor changes sign (Rosenblad, 2020); the interpretation of the marginal effect of one regressor when interacted with another (Hirschberg & Lye, 2010), among many others meanings.

Profile likelihood function for the ratio of two regression coefficients was given by Ghosh et al. (2003) and reviewed again in Ghosh et al. (2006), where they showed that inferences about this parameter, based on profile likelihood-confidence intervals, may result the entire real line. This fact occurs because we are dealing with an essentially flat likelihood function that has a horizontal asymptote, located at some distance above the global minimum.

In this paper, we analyze conditions under which a confidence interval for the interest parameter, based on
the profile likelihood function, has an infinite length and even becomes the entire real line. These conditions are related to the nested models problem. In the cases discussed here, infinite length confidence intervals arise when a hypothesis test, under an $F$ statistic, shows that full and reduced regression models fit the data equally well. When this situation is considered, the shape of the likelihood function is very informative and explains this model criticism problem.

The paper is structured as follows. Section 2 includes some results used to perform hypothesis testing about the coefficients of the classical linear regression model. In addition, some connections between these tests and a pivotal quantity (or pivot) for the ratio of two regression coefficients, our parameter of interest, are presented. Section 3 shows how the shape of the profile likelihood function of the parameter of interest is determined by the pivotal quantity. Additionally, connections between the shape of the likelihood and hypotheses tests about the coefficients of the linear regression model are shown. In Section 4 we include three illustrative simulated cases to exemplify theoretical results shown in previous sections. Finally, some concluding remarks are presented in Section 5.

2. **F-TESTS AND NESTED REGRESSION MODEL**

Cox (2006, p. 3) described that a statistical inference process can be basically divided into two stages: choice of the family of models and model criticism. The family that is chosen in the first stage is often fully specified, except for a limited number of unknown parameters. In our case, we address flat likelihoods and estimation issues, within the multiple linear regression framework. For simplicity, and without loss of generality, we have selected a linear regression model with two predictor variables. In the second stage, we must ask ourselves whether the selected model is consistent with the data, if some changes on it may be needed, or even if the whole model should be modified. A particular problem related to this stage arises when comparing two nested models. Two models are called nested if one contains all the predictors of the other one and also some additional predictors. In linear regression, this is known as reduced models (Searle & Gruber, 1971).

Consider the regression model

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + e_i,$$  \hspace{1cm} (1)

for $i = 1, \ldots, n$, $n > 2$, where $e_i$'s are independent random errors, normally distributed with mean 0 and common variance $\sigma^2$. Here, regression coefficients, $\beta_1$ and $\beta_2$, are considered real unknown parameters. Model presented in (1) can be written in matrix form

$$y = X\beta + e,$$  \hspace{1cm} (2)

where $y^T = (y_1, \ldots, y_n)$, $X^T = (x_{11}, \ldots, x_{1n}, x_{21}, \ldots, x_{2n})^T$, $i = 1, \ldots, n$, $\beta^T = (\beta_1, \beta_2)$, $e^T = (e_1, \ldots, e_n)$, and rank of $X$ equals 2.
The nested modelling problem associated with the aim of this article occurs when \( y_i = e_i \) or \( y_i = \beta_1 x_i + e_i \) fits the data evenly well. The typical statistical hypotheses, related to these nested problems are \( H_{01} : \beta_1 = \beta_2 = 0 \) and \( H_{02} : \beta_2 = 0 \), respectively. For the nested model \( y_i = e_i \), called the white noise problem, we test the hypothesis \( H_{01} : \beta_1 = \beta_2 = 0 \), using the standard test statistic
\[
F_1 = \frac{n\hat{\beta}^T C \hat{\beta}}{2MSE},
\]
where \( C = (c_{ij}) = n^{-1} X^T X \) is a positive matrix, \( i, j = 1, 2 \), \( \hat{\beta}^T = (\hat{\beta}_1, \hat{\beta}_2) = (nC)^{-1} X^T y \), and \( MSE = (y - X \hat{\beta})^T (y - X \hat{\beta}) / (n - 2) \). It is well known that \( F_1 \) has an \( F \) distribution with 2 and \( n - 2 \) degrees of freedom. For \( 0 < \alpha < 1 \), the hypothesis \( H_{01} : \beta_1 = \beta_2 = 0 \) is rejected at \( \alpha \) significance level when \( F_1 \) is greater than the \((1 - \alpha)\) quantile of the \( F \) distribution with 2 and \( n - 2 \) degrees freedom, quantile denoted here as \( F_{2,n-2,1-\alpha} \).

On the other hand, for the nested model \( y_i = \beta_1 x_{i1} + e_i \), we can test the null hypothesis \( H_{02} : \beta_2 = 0 \) using the typical test statistic
\[
F_2 = \frac{n|C| \sqrt{\beta_2^2 / c_{11}}}{MSE}.
\]
This statistic has an \( F \) distribution with 1 and \( n - 2 \) degrees of freedom. Thus, when \( F_2 \) is greater than \( F_{1,n-2,1-\alpha} \), the \((1 - \alpha)\) quantile of the \( F \) distribution with 1 and \( n - 2 \) degrees freedom, the hypothesis \( H_{02} : \beta_2 = 0 \) is rejected at \( \alpha \) significance level.

Let us now consider the problem of making inferences on \( \theta_1 = \beta_1 / \beta_2 \), the ratio of the regression coefficients \( \beta_1 \) and \( \beta_2 \); when \( \beta_1 = \beta_2 = 0 \) or \( \beta_2 = 0 \), it can cause indeterminacy and lead to some paradoxes. For such reason, inferences about \( \beta^T = (\beta_1, \beta_2) \) play an important role in the determination of plausible values of \( \theta_1 \), and this must be fully understood. Particularly, those inferences related to hypotheses \( H_{01} \) and \( H_{02} \).

Ghosh et al. (2006) showed that
\[
F_1(\theta_1) = \frac{n|C|(\hat{\beta}_2 \theta_1 - \bar{\beta}_1)^2 / Q(\theta_1)}{MSE}.
\]
has an \( F \) distribution with 1 and \( n - 2 \) degrees of freedom, where \( Q(\theta_1) = c_{11} \theta_1^2 + 2c_{12} \theta_1 + c_{22} \), being \( F_1(\theta_1) \) a pivotal quantity (Sprott, 2008, p. 63) related to the hypotheses tests \( H_{01} \) and \( H_{02} \). They also found the following relationship between \( F_1(\theta_1) \) and the statistic \( F_1 \):
\[
\sup_{\theta_1} F_1(\theta_1) = 2F_1.
\]

Now, considering (3) and (6), the hypothesis \( H_{01} : \beta_1 = \beta_2 = 0 \) is rejected at \( \alpha \) significance level if
\[
\sup_{\theta_1} F_1(\theta_1) > 2F_{2,n-2,1-\alpha}.
\]

On the other hand, there is also a relationship between \( F_1(\theta_1) \) and \( F_2 \) that is as follows:
\[
\lim_{|\theta_1| \to \infty} F_1(\theta_1) = F_2.
\]
since \( \lim_{|\theta_1| \to \infty} (\hat{\beta}_2 \theta_1 - \hat{\beta}_1)^2 / Q(\theta_1) = \hat{\beta}_2^2 / c_{11} \). Hence, observing (4) and (8), the hypothesis \( H_{02} : \beta_2 = 0 \) is rejected at \( \alpha \) significance level if
\[
\lim_{|\theta_1| \to \infty} F_1(\theta_1) > F_{1,n-2,1-\alpha}.
\] (9)

The profile likelihood function of \( \theta_1 \) is presented in the next section, where it is used to get a closed mathematical expression for the likelihood ratio test statistic, associated to the pivotal quantity \( F_1(\theta_1) \). In addition, we show some connections between relative profile likelihood function and confidence intervals for \( \theta_1 \). But, even more important than that is to show how profile likelihood function shape is associated with the nature of the nested models problem. Thus, the characteristics of this problem are naturally inherited by \( \theta_1 = \beta_1/\beta_2 \) and their confidence intervals.

3. INFEERENCE ABOUT PARAMETER \( \theta_1 \)

The likelihood approach is one of the most popular techniques for deriving statistics, as explained in Casella & Berger (2002, p. 315), and this is performed via the likelihood function. The resulting statistics play an important role in statistical inference since, besides the fact that they can be used to obtain point estimates, it is possible to construct confidence intervals or perform hypothesis tests based on them. In our case, the likelihood function for model (2) is given by
\[
L(\beta_1, \beta_2, \sigma) \propto \sigma^{-n} \exp \left[ -\frac{1}{2\sigma^2} (y - X\hat{\beta})^T (y - X\hat{\beta}) \right],
\]

and since \((y - X\hat{\beta})^T (y - X\hat{\beta}) = (y - X\hat{\beta} + X\hat{\beta} - X\hat{\beta})^T (y - X\hat{\beta} + X\hat{\beta} - X\hat{\beta}) \) and \((X\hat{\beta} - X\hat{\beta})^T (y - X\hat{\beta}) = 0 \), the likelihood function can be also written as
\[
L(\beta_1, \beta_2, \sigma) \propto \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \left[ SSE + (\beta - \hat{\beta})^T X^T X (\beta - \hat{\beta}) \right] \right\},
\] (10)

where \( SSE = (y - X\hat{\beta})^T (y - X\hat{\beta}) \).

To make inferences regarding a parameter of interest, in the presence of nuisance parameters, the likelihood approach relies on the profile likelihood function (Sprott, 2008, p. 66). This is obtained by fixing values for the parameter of interest and maximizing with respect to nuisance parameters. Here, \( \theta_1 = \beta_1/\beta_2 \) is the parameter of interest and \( \beta_2 \) and \( \sigma \) are considered nuisance parameters. Hence, in order to obtain the profile likelihood function for \( \theta_1 \), the likelihood function (10) is reparametrized in terms of \( (\theta_1, \beta_2, \sigma) \).

\[
L(\theta_1, \beta_2, \sigma) \propto \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \left[ SSE + n(\beta_2 \phi - \hat{\beta})^T C(\beta_2 \phi - \hat{\beta}) \right] \right\},
\] (11)

where \( \phi^T = (\theta_1, 1) \). Then, maximizing with respect to \( \sigma \) and \( \beta_2 \), the resulting profile likelihood function of \( \theta_1 \) is
\[
L_p(\theta_1) \propto \left[ SSE + n|C|(\beta_2 \theta_1 - \hat{\beta}_1)^2 / Q(\theta_1) \right]^{-n/2},
\] (12)
where \( Q(\theta_1) = \phi^T C \phi = c_{11} \theta_1^2 + 2c_{12} \theta_1 + c_{22} > 0 \).

By the likelihood invariance property, the MLE of \( \theta_1 \) is \( \hat{\theta}_1 = \hat{\beta}_1/\hat{\beta}_2 \). Now, the relative profile likelihood function of \( \theta_1 \) is a standardized version that takes a value of one at the maximum of the profile likelihood function of \( \theta \) given in (12),

\[
R(\theta_1) = \frac{L_P(\theta_1)}{L_P(\hat{\theta}_1)} = \left[ 1 + \frac{n|C(\hat{\beta}_2 \theta_1 - \hat{\beta}_1)^2/Q(\theta_1)|}{(n-2)MSE} \right]^{-n/2}.
\]

(13)

A level \( c \) profile likelihood region for \( \theta_1 \) is given by

\[
\{ \theta_1 : R(\theta_1) \geq c \},
\]

where \( 0 \leq c \leq 1 \). Since \( -2\ln[R(\theta_1)] \) follows an asymptotically \( \chi^2 \) distribution (Spratt, 2008; Kalbfleisch, 1985), approximate likelihood-confidence intervals can be easily obtained. \( R(\theta_1) \) is a transformation of the pivotal quantity \( F_{1}(\theta_1) \) defined in (5),

\[
R(\theta_1) = \left[ 1 + \frac{F_{1}(\theta_1)}{n-2} \right]^{-n/2},
\]

(14)

then, a \( 100(1-\alpha) \% \) likelihood-confidence interval for \( \theta_1 \) is

\[
LCI(\gamma) = \{ \theta_1 : R(\theta_1) > c_{1,n-2,1-\alpha} \},
\]

(15)

with

\[
c_{1,n-2,1-\alpha} = \left[ 1 + \frac{F_{1,n-2,1-\alpha}}{n-2} \right]^{-n/2},
\]

where \( F_{1,n-2,1-\alpha} \) is the \( (1-\alpha) \) quantile of the \( F \) distribution with 1 and \( n-2 \) degrees freedom, and \( 0 < \alpha < 1 \).

Equation (15) implies that, in a \( \theta_1 \) versus \( R(\theta_1) \) plot, the likelihood-confidence interval will consist of all \( \theta_1 \) values where relative profile likelihood function exceeds the horizontal cutoff \( c_{1,n-2,1-\alpha} \).

On the other hand, since \( F_{1}(\theta_1) \) properties are inherited by \( R(\theta_1) \), then, in order not only to understand the shape of the profile likelihood function of \( \theta_1 \), but also to interpret it and to carefully state the results from statistical inferences for \( \theta_1 = \beta_1/\beta_2 \) and their corresponding hypotheses tests for \( (\beta_1, \beta_2) \), it is necessary to make a link between the characteristics of the pivotal quantity \( F_{1}(\theta_1) \) and \( R(\theta_1) \).

In particular, since the quantity \( F_{1}(\theta_1) \) attains its maximum and minimum and has a horizontal asymptote, then \( R(\theta_1) \) will also reach its maximum and minimum and will have a horizontal asymptote. Relative profile likelihood function \( R(\theta_1) \) crosses the horizontal asymptote, as illustrated in Figure 1. Note that, by (8) and (14), \( F_{1}(\theta_1) \) inherits the asymptote to \( R(\theta_1) \) function, when \(|\theta_1| \) becomes larger, so that \( R(\theta_1) \) approaches to the asymptote \( (1 + F_2/(n-2))^{-n/2} \). Moreover, by varying \( c_{1,n-2,1-\alpha} \) level, \( 0 < \alpha < 1 \), the length of the likelihood-confidence intervals for \( \theta_1 \) can be finite or infinite. Something far more interesting is the fact that
likelihood inferences for \( \theta_1 \) are related to the classical hypotheses tests on the regression coefficients \( \beta_1 \) and \( \beta_2 \), as is explained below.

A case of likelihood-confidence intervals of infinite length for \( \theta_1 \) occurs when

\[
\inf_{\theta_1 \in \mathbb{R}} R(\theta_1) > c_{1,\nu-2,1-\alpha}.
\]

(16)

This fact leads to the typical criticism of absurd likelihood-confidence intervals. That is, there is always a confidence level \( \alpha < 1 \) for which the associated likelihood-confidence interval is the entire real line. This issue should not be taken lightly, because it conceals information about the plausibility of the hypothesis \( H_{01} : \beta_1 = \beta_2 = 0 \). Since (16) can equivalently be written as

\[
2F_1 < F_{1,\nu-2,1-\alpha},
\]

and \( F_{1,\nu-2,1-\alpha} < 2F_{2,\nu-2,1-\alpha} \), see Casella & Berger (2002, p. 258) and Casella et al. (2001, p. 5-7), then the hypothesis \( H_{01} : \beta_1 = \beta_2 = 0 \) is not rejected at \( \alpha \) significance level. Therefore, intervals of this type are related to a non-rejection of \( H_{01} \). On the other hand, when \( H_{01} \) is rejected, it is usual to analyze if any of the regression coefficients is zero. The connection between \( H_{02} : \beta_2 = 0 \) and the infinite length likelihood-confidence intervals is shown below.

Another possibility of likelihood-confidence intervals of infinite length for \( \theta_1 \) occurs when

\[
\lim_{|\theta_1| \to \infty} R(\theta_1) \geq c_{1,\nu-2,1-\alpha}.
\]

(17)

In this case, likelihood-confidence interval may result the entire real line or the union of two infinite intervals, and that occurs when \( c_{1,\nu-2,1-\alpha} \) is less than the minimum value of \( R(\theta_1) \) or when it lies between the
horizontal asymptote and the minimum value of $R(\theta_1)$, respectively. Note that an infinite length interval is also obtained when the value of $c_{1,n-2,1-\alpha}$ coincides exactly with the asymptote. However, in any case, the hypothesis $H_{02} : \beta_2 = 0$ is not rejected at $\alpha$ significance level, because (17) is equivalent to

$$F_2 < F_{1,n-2,1-\alpha}.$$ 

Overall, likelihood-confidence intervals of infinite length suggest moving back to the model criticism stage and testing for model adequacy, in terms of the significance of the regression model coefficients.

4. ILLUSTRATIVE EXAMPLES

This section exemplifies the relationship between parameter $\theta_1$ inferences and hypotheses tests regarding nested models involving $\beta_1$ and $\beta_2$. In that sense, the three following cases are considered:

- Example 1: $H_{01} : \beta_1 = \beta_2 = 0$ is not rejected,
- Example 2: $H_{02} : \beta_2 = 0$ is not rejected but $H_{01} : \beta_1 = \beta_2 = 0$ is rejected,
- Example 3: $H_{01} : \beta_1 = \beta_2 = 0$ and $H_{02} : \beta_2 = 0$ are both rejected.

In all these examples sample size is set at $n = 25$ and significance level is fixed at $\alpha = 0.05$, so $c_{1,n-2,1-\alpha}$ shown in (15) takes a value of 0.11, which will be used in all these examples. In the same way, $\sigma = 5$ remains fixed, as well as the covariable $(x_1, x_2)$ values obtained from Rawlings et al. (1998, p. 177). In view of all the above, the procedure results as follows:

- In all the examples, $\beta_1$ and $\beta_2$ values are provided.
- Given $\beta_1$ and $\beta_2$ values, $y_i$ is then simulated from a normal distribution with mean $\beta_1 x_{i1} + \beta_2 x_{i2}$ and a standard deviation $\sigma$.
- Once $F_1$ statistic is computed, then $H_{01} : \beta_1 = \beta_2 = 0$ is tested. The critical value corresponding to this hypothesis test is $F_{2,n-2,1-\alpha} = 3.42$.
- In case of $H_{01} : \beta_1 = \beta_2 = 0$ rejection, then $H_{02} : \beta_2 = 0$ is tested by computing $F_2$ statistic and comparing it with $F_{1,n-2,1-\alpha} = 4.27$ critical value.
- Finally, a plot of the relative profile likelihood function for $\theta_1$ is constructed and a $100(1 - \alpha)\%$ likelihood-confidence interval for this parameter is computed.

4.1. Example 1: $H_{01} : \beta_1 = \beta_2 = 0$ is not rejected

In this example $\beta_1 = 0.2$ and $\beta_2 = 0.02$ are considered, simulating then $y_i$ values from a normal distribution with mean $0.2x_{i1} + 0.02x_{i2}$ and standard deviation $\sigma$. Once the sample for this regression model is generated, the hypothesis $H_{01} : \beta_1 = \beta_2 = 0$ is tested at $\alpha$ significance level. In this case, the test statistic value
results $F_1 = 1.02$. Now, since $F_1 < F_{2,n-2,1-\alpha}$, there is no statistical evidence to reject the hypothesis $H_{01} : \beta_1 = \beta_2 = 0$; so that these covariables could be excluded from the model. However, if this information is not considered and the inference process regarding $\theta_1$ continues, a relative profile likelihood function of $\theta_1$, like the one shown in Figure 2, is obtained. This type of graph clearly shows a function that is always above the horizontal cutoff $c_{1,n-2,1-\alpha}$; that is, the whole 100$(1 - \alpha)$% likelihood-confidence interval for $\theta_1$ results in the real line. Moreover, as explained in Equation (16), when the infimum of the relative profile likelihood function for $\theta_1$ is greater than $c_{1,n-2,1-\alpha}$, then, there is no statistical evidence to reject $H_{01} : \beta_1 = \beta_2 = 0$, at the proposed $\alpha$ significance level.

4.2. Example 2: $H_{02} : \beta_2 = 0$ is not rejected but $H_{01} : \beta_1 = \beta_2 = 0$ is rejected

In this example $\beta_1 = 0.5$ and $\beta_2 = 0.02$, so $y_i$ values are simulated from a normal distribution with mean $0.5x_{i1} + 0.02x_{i2}$ and standard deviation $\sigma$. Under this scenario $F_1 = 20.6$, so it follows that $F_1 > F_{2,n-2,1-\alpha}$, and at $\alpha$ significance level there is statistical evidence to reject the hypothesis $H_{01} : \beta_1 = \beta_2 = 0$. According to the procedure previously stated, hypothesis $H_{02} : \beta_2 = 0$ is now tested. In this case $F_2 = 1.63$, so $F_2 < F_{1,n-2,1-\alpha}$; that is, there is no statistical evidence to reject the hypothesis $H_{02} : \beta_2 = 0$, at $\alpha$ significance level.
level, so $x_2$ variable could be excluded from the regression model. However, if this variable remains in the model and the inference process about parameter $\theta_1$ continues, a $100(1 - \alpha)\%$ likelihood-confidence interval of infinite length for $\theta_1$ is also obtained. That occurs because the asymptote of the relative profile likelihood function for $\theta_1$ is above the cutoff $c_{1,n - 2,1 - \alpha}$, as shown in Figure 3. Note that, in this case, the infimum of the relative profile likelihood is not above this cutoff, like in previous example where $H_{01} : \beta_1 = \beta_2 = 0$ is not rejected, but the asymptote of the relative profile likelihood function for $\theta_1$, as explained in Equation 17, is above the cutoff $c_{1,n - 2,1 - \alpha}$, so there is no statistical evidence to reject $H_{02} : \beta_2 = 0$, at $\alpha$ significance level.

4.3. **Example 3:** $H_{01} : \beta_1 = \beta_2 = 0$ and $H_{02} : \beta_2 = 0$ are both rejected

To exemplify this scenario $\beta_1 = 0.5$ and $\beta_2 = 0.6$ are considered, so $y_i$ values are generated based on a normal distribution with mean $0.5x_{i1} + 0.6x_{i2}$ and standard deviation $\sigma$. Now, since $F_1 = 27.44$, then $F_1 > F_{2,n - 2,1 - \alpha}$; therefore, the hypothesis $H_{01} : \beta_1 = \beta_2 = 0$ is rejected at $\alpha$ significance level. Similarly, as $F_2 = 4.64$, then $F_2 > F_{1,n - 2,1 - \alpha}$ and $H_{02} : \beta_2 = 0$ is also rejected at $\alpha$ significance level. When continuing with the inference process regarding $\theta_1$, under this situation, a finite $100(1 - \alpha)\%$ likelihood-confidence interval for $\theta_1$ is
obtained. In this case the asymptote of the relative profile likelihood function for $\theta_1$ is below the cutoff $c_{1,\pi-2,1-\alpha}$, as shown in Figure 4.

As can be observed in the examples included here, when an infinite $100(1-\alpha)\%$ likelihood-confidence interval for $\theta_1$ is obtained, it implies that there will be no statistical evidence to reject any of the hypotheses $H_{01} : \beta_1 = \beta_2 = 0$ or $H_{02} : \beta_2 = 0$, at $\alpha$ significance level. Note that, as mentioned in previous section, when hypothesis $H_{02} : \beta_2 = 0$ is not rejected, at $\alpha$ significance level, then a $100(1-\alpha)\%$ likelihood-confidence interval of infinite length, for parameter $\theta_1 = \beta_1/\beta_2$, is obtained.

5. CONCLUSIONS

The purpose of this article is to emphasize the importance of deeply understanding and analyzing the shape of the likelihood function and their corresponding inferences. In the cases discussed here, the infinite length of likelihood-confidence intervals are due to the fact of considering full models when reduced models are not rejected by the data. Occurrence of flat likelihood shapes, like the ones presented here, are very informative regarding model criticisms and should not be the reason for seeking or promoting alternative estimation.
methods such as integrated likelihoods or posterior Bayesian distributions.

**Acknowledgements**

We thank Dr. Gudelia Figueroa Preciado for her help in the development of this paper and for her insightful discussions. We also thank the anonymous referees for their helpful suggestions.

**Funding**

This article was supported by Consejo Nacional de Ciencia y Tecnología. Doctoral fellowship of the first author.

**References**


