Evaluating a Theory-Based Hypothesis Against Its Complement Using an AIC-Type Information Criterion With an Application to Facial Burn Injury

Leonard Vanbrabant
Ghent University and Municipal Health Service (GGD) West-Brabant, Noord-Brabant, Breda, the Netherlands

Nancy Van Loey
Association of Dutch Burn Centers, Noord-Holland, Beverwijk, the Netherlands, and Utrecht University

Rebecca M. Kuiper
Utrecht University

Abstract
An information criterion (IC) like the Akaike IC (AIC), can be used to select the best hypothesis from a set of competing theory-based hypotheses. An IC developed to evaluate theory-based order-restricted hypotheses is the Generalized Order-Restricted Information Criterion (GORIC). Like for any IC, the values themselves are not interpretable but only comparable. To improve the interpretation regarding the strength, GORIC weights and related evidence ratios can be computed. However, if the unconstrained hypothesis (the default) is used as competing hypothesis, the evidence ratio is not affected by sample-size nor effect-size in case the hypothesis of interest is (also) in agreement with the data. In practice, this means that in such a case strong support for the order-restricted hypothesis is not reflected by a high evidence ratio. Therefore, we introduce the evaluation of an order-restricted hypothesis against its complement using the GORIC (weights). We show how to compute the GORIC value for the complement, which cannot be achieved by current methods. In a small simulation study, we show that the evidence ratio for the order-restricted hypothesis versus the complement increases for larger samples and/or effect-sizes, while the evidence ratio for the order-restricted hypothesis versus the unconstrained hypothesis remains bounded. An empirical example about facial burn injury illustrates our method and shows that using the complement as competing hypothesis results in much more support for the hypothesis of interest than using the unconstrained hypothesis as competing hypothesis.

Translational Abstract
In an informative hypothesis, academic expertise (i.e., theory) about the population of interest is included in the hypothesis in terms of order-restrictions on the model parameters. As an example, in an ANOVA setting, we might expect that the group means follow a certain order (e.g., $H_1$: $\mu_1 < \mu_2 < \mu_3 < \mu_4$). As another example, in a linear regression model, we might expect that the (standardized) regression coefficients are subject to multiple one-sided restrictions (e.g., $H_2$: $\beta_1 > 0; \beta_2 > 0; \beta_3 > 0$). In the absence of a competing informative hypothesis, an order-restricted hypothesis is typically evaluated against the unconstrained hypothesis $H_u$ (all orderings are allowed). However, a problem arises if the order-restricted hypothesis $H_1$ is in agreement with the data: Then, the estimated parameters for $H_1$ are identical to the estimated parameters of $H_u$. Therefore, we introduce an AIC-type/information-theoretic method for evaluating an informative hypothesis $H_u$ against its complement, where the complement $H_c$ is defined as $H_c$: not $H_u$ (all orderings are allowed except the ordering in $H_1$). The method is illustrated using an empirical example about facial burn injury. In addition, the method is implemented in the user-friendly R package restrktor.

Keywords: complement, evidence ratio, level probabilities, model selection, order restrictions

The evaluation of a researcher’s theory is often key to research. For example, consider a study about facial burn injury. A burn event can have an avers impact on a person’s quality of life. The scars can affect physical appearance and may constitute a source of
ruminations acting as a reminder to the event (Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008). Based on previous research (Van Loey et al., 2014), it is expected that injury characteristics that may be perceived as distressing such as facial burn injury and larger burns may be triggers for the activation and prolongation of rumination. In addition, a gender effect is also expected because disfiguring scars from burns may be of greater importance to women as compared with men (Ghirwati et al., 2017). Therefore, we hypothesized that the means of rumination for men with and without facial burn injury and the mean of rumination for women without facial burn injury would be lower than the mean of rumination for women with facial burn injury. In symbols, this hypothesis can be stated as: $H_1$: \{ $\mu_{\text{men, no facial burns}} \leq \mu_{\text{men, facial burns}}$, $\mu_{\text{women, no facial burns}} \leq \mu_{\text{women, facial burns}}$ \} $\leq \{ \mu_{\text{women, facial burns}} \}$, where $\mu$ are the population means for rumination for the four groups determined by gender and facial burns, adjusted for the population effects of some covariates. Note that no particular order is assumed among the first three means. This form of theory-based hypothesis is known as an order-restricted (OR) hypothesis or informative hypothesis (Hu, 2012) because the order of the means is restricted based on theory and/or academic expertise.

To evaluate such OR hypothesis, three methods can be distinguished, that is, OR hypothesis testing (e.g., Kudo, 1963), model selection using OR information criteria (e.g., Anraku, 1999; Mulder, & Raftery, in press), and model selection using the Bayes factor (e.g., Mulder, Hoijtink, & Klugkist, 2010). In this current article, we focus on model selection using information criteria. Akaike’s IC (AIC; see, e.g., Akaike, 1973, 1998) is probably the most familiar and widely used information criterion employed in the social and behavioral sciences. Nevertheless, the AIC is not suitable when the model parameters (e.g., means or regression coefficients) are subject to order restrictions. A modification of the AIC that can deal with simple order restrictions (i.e., Order-Restricted Information Criterion [ORIC]) in the exponential family was proposed by Amraku (1999). Kuiper, Hoijtink, and Silvapulle (2011) generalized the ORIC (GORIC) to accommodate any linear inequality restrictions in multivariate normal models (except for range restrictions, which bounds a parameter to a specific interval, e.g., $-1 \leq \mu \leq 1$). Information criteria like the AIC, ORIC, and GORIC are calculated as minus two times the maximum log-likelihood (under the hypothesized model) plus twice a penalty term value. The main difference between the methods is in calculating the penalty term value, which is less straightforward to compute in case of order restrictions.

The evaluation of an OR hypothesis (e.g., $H_1$) requires at least one competing hypothesis. Sometimes researchers have another hypothesis of interest and want to know which hypothesis is best, for example, $H_1$ versus $H_2$: $\mu_{\text{men, no facial burns}} \leq \mu_{\text{men, facial burns}}$, $\mu_{\text{women, no facial burns}} \leq \mu_{\text{women, facial burns}}$. The difference between the two is that in $H_2$ the means are fully ordered. In practice, researchers often do not have such a specific competing hypothesis and only the unconstrained hypothesis $H_m$, where no restrictions are imposed on the model parameters, remains included as competing hypothesis in the set. Therefore, in this article, we focus solely on the set of hypotheses with one OR hypothesis $H_m$ and the unconstrained hypothesis $H_m$.

The hypothesis with the lowest GORIC value is the preferred one. The GORIC values themselves are not interpretable and only the differences between the values can be inspected. To improve the interpretation, so-called GORIC weights ($w_m$) can be computed, which are derived from the Akaike weights (Akaike, 1978; Burnham & Anderson, 2002) and are comparable with posterior model probabilities (see Burnham & Anderson, 2002, pp. 302–305). An IC weight $w_m$ represents the relative likelihood of hypothesis $m$ given the data and the set of $M$ hypotheses (Burnham & Anderson, 2002; Kuiper, 2011; Wagenmakers & Farrell, 2004). For example, if we compare hypothesis $H_1$ against hypothesis $H_m$, we can examine the ratio of the two corresponding weights, that is $w/m$. This evidence ratio is considered as the strength of evidence in favor of model $m$ (in this example, $m = 1$) of being the best model (Burnham & Anderson, 2002; Wagenmakers & Farrell, 2004).

The evidence ratio should increase for larger samples and/or effect-sizes. However, if the OR hypothesis of interest $H_m$ is in agreement with the data, increasing the sample-size and/or effect-size does not affect the evidence ratio if the unconstrained hypothesis is used as competing hypothesis (assuming that $H_m$ is true and remains in agreement with the data). In that case, both hypotheses $H_m$ and $H_r$ are in line with the data, since $H_r$ is always in line with the data, and consequently both hypotheses have the same maximized log-likelihood value. Then, the difference in GORIC values equals the difference in penalty term values, which are independent of sample-size and effect-size. The latter case is illustrated in Figure 1, where we generated 500 data sets according to an ANOVA model with four uncorrelated ordered means which are in agreement with $H_1$, with a sample-size of $n = 50$ per group and various effect-sizes $f$ (Cohen, 1988, pp. 274–275). The results (cf. see triangles in Figure 1) show that at first the mean evidence ratio (on a log scale) of $w/\mu$ increases slightly for increasing effect-sizes and that afterward it stabilizes at an upper-bound value of approximately $\exp(1.09) \approx 2.97$ on the original scale. It is at this point that the data are for each simulation run in agreement with $H_1$ and thus the maximized log-likelihood values of $H_1$ and $H_m$ are the same. The boundary value equates the exponential difference of the penalty term values between $H_m$ and $H_r$, that is, $\exp(5.00–3.91) = \exp(1.09) \approx 2.97$; as will become clear later on. Consequently, strong support for the OR parameters is not expressed in a high evidence ratio if compared to the unconstrained hypothesis and many research questions may be erroneously dismissed as irrelevant. It is important to note that the boundary issue is not specific to the order-restricted IC literature but can also be found in the Bayes factor literature (e.g., Mulder et al., 2009, 2010). They solve the boundary issue by comparing the order-restricted hypothesis against its complement.

The objective of this study is to show that this upper bound issue can also be solved by replacing the unconstrained hypothesis by the complement of the hypothesis of interest (cf. see circles in Figure 1). The complement is defined as $H_r = \neg H_m$, where $\neg$ denotes "not." For example, for the OR hypothesis $H_1$: \{ $\mu_{\text{men, no facial burns}} \leq \mu_{\text{men, facial burns}}$, $\mu_{\text{women, no facial burns}} \leq \mu_{\text{women, facial burns}}$ \} $\leq \{ \mu_{\text{women, facial burns}} \}$, there are four means there are four ways in which the four means can be ordered in such an ordering. Hypothesis $H_1$ consists of one of these four combinations, therefore the complement represents the $4 - 1 = 3$ remaining ways in which the four means can be ordered. In

\[ H_1 : \{ \mu_{\text{men, no facial burns}} \leq \mu_{\text{men, facial burns}} \leq \mu_{\text{women, no facial burns}} \leq \mu_{\text{women, facial burns}} \}\]
Before we introduce the GORIC for the complement, some technical background is inevitable. The results given in this part are for the linear regression model, where the regression coefficients are subject to linear inequality and/or linear equality restrictions. The method can readily be adapted to multivariate normal linear models. This is briefly discussed in the last section.

This "simple" case, it is easy to write out the complement but this is often not the case. In many cases, it is a cumbersome or even impossible task to write up all possible combinations that belong to the complement, because the number of combinations increases excessively with the number of parameters. For example, for the OR hypothesis \( H_2 \) with four means there are 24 ways (i.e., \( 4! = 4 \times 3 \times 2 \times 1 \)) in which the four means can be ordered in a simple order ordering. Moreover, the GORIC is often not defined when the complement comprises multiple combinations. This is because, the GORIC is only defined for restrictions that form a closed convex cone (ccc, e.g., \( H_1 \) and \( H_2 \)).

The novelty of this article is that we introduce the GORIC for the situation that the restrictions in the complement are not a ccc. We show (a) how to determine the log-likelihood for the complement and (b) how to determine the penalty term value for the complement.

The remainder of this article is organized as follows. First, we provide some technical background about the computation of the GORIC and the corresponding penalty term value for the unconstrained hypothesis and an OR hypothesis \( H_m \). Second, we introduce how the GORIC is computed for the complement of \( H_m \). Third, we illustrate our method with the empirical example introduced at the beginning of this section. Finally, we give some concluding remarks and recommendations.

**Technical Background**

Before we introduce the GORIC for the complement, some technical background is inevitable. The results given in this part are for the linear regression model, where the regression coefficients are subject to linear inequality and/or linear equality restrictions. The method can readily be adapted to multivariate normal linear models. This is briefly discussed in the last section.

**Linear Model and Order-Restricted Hypotheses**

Consider the standard linear regression model,

\[
y_i = x_i^T \theta + \epsilon_i, \quad i = 1, \ldots, n,
\]

where \( \theta = (\theta_1, \ldots, \theta_p)^T \) is the parameter vector of interest, \( x_i = (x_{i1}, \ldots, x_{ip})^T \) is a vector of predictor variables for person \( i \), and \( \epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{ip})^T \) is a vector of normally distributed random errors: \( \epsilon_i \sim N(0, \sigma^2) \). Let the (unconstrained) maximum likelihood estimates be denoted by \( \hat{\theta} \) and the order-restricted mle's denoted by \( \hat{\theta}_m \). The latter is the solution of maximizing the likelihood under the restrictions in \( H_m \), a well-studied restricted optimization problem in the statistical literature (Nocedal & Wright, 2006).

We consider three types of hypotheses, namely \( H_0 : \theta \in \mathbb{R}^p \), where \( \mathbb{R}^p \) is the p-dimensional Euclidean space, \( H_m : \theta \in C \), where \( C \) is also a space in \( \mathbb{R}^p \) and is a (reallocated) closed convex cone (Kuiper, Hooijink, & Silvapulle, 2012), and \( H_1 : \neg H_m \), which is not necessarily a (reallocated) closed convex cone. Because, most applications only involve linear restrictions, we only consider linear hypotheses.

**The GORIC**

The GORIC for the unconstrained hypothesis \( H_0 \) is defined as

\[
\text{GORIC}_u = -2 \times \text{LL}_u + 2 \times \text{PT}_u,
\]

where \( \text{LL}_u \) is the maximized log-likelihood value and the penalty term value is defined as \( \text{PT}_u = 1 + p \). Note that GORIC\(_u\) equals the AIC for \( H_u \).

The GORIC for the OR hypothesis \( H_m \) is defined as

\[
\text{GORIC}_m = -2 \times \text{LL}_m + 2 \times \text{PT}_m,
\]

where \( \text{LL}_m \) is the maximized log-likelihood value for the OR hypothesis \( H_m \) and \( \text{PT}_m \) is the penalty term value for \( H_m \). The penalty term value equals

\[
\text{PT}_m = 1 + \sum_{j=0}^{p} j \times \text{LP}(p, \Sigma, H_m).
\]

In an univariate regression model, \( \Sigma = (X^T X)^{-1} \) is the unscaled covariance matrix of the parameters with \( X = (x_{11}, \ldots, x_{n1})^T \) of order \( n \times p \) and \( \text{LP}(p, \Sigma, H_m) \) are the level probabilities (chi-bar-square weights). A level probability \( \text{LP}_p \) is the probability that the OR mle \( \hat{\theta}_m \) has \( j \) levels (under the null-hypothesis), where \( j = p - \text{“the number of active restrictions”}; and the \( \text{LP} \)'s sum to 1. We will clarify the computation of \( \text{PT}_m \) using two examples. Consider Figure 2a, where

Figure 1. Mean evidence ratio (on a log scale and based on 500 simulations) for hypothesis \( H_1 : [\mu_1, \mu_2, \mu_3] \leq \mu_4 \) compared with the unconstrained hypothesis (mean \( w_i / w_c \)) and compared with its complement (mean \( w_i / w_c \)) when \( H_1 \) is true, for \( n = 50 \), and various effect-sizes (f).
the unconstrained parameter space is determined by the two parameters \( \theta_1 \) and \( \theta_2 \) and is divided into four quadrants (Q1 to Q4). If we assume that \( \theta_1 \) and \( \theta_2 \) are independent of each other (i.e., \( \Sigma = I \), where \( I \) is an identity matrix), then each quadrant is assigned a level probability of 0.25 under \( H_0 : \theta_1 = \theta_2 = 0 \). The permissible gray shaded area is defined by the order restrictions \( H_3 : \theta_1 \geq 0, \theta_2 \geq 0 \). Then, the probability that \( j = 2 \) that is, none of the restrictions are active (i.e., \( j = p - 0 = 2 - 0 = 2 \), is 0.25 \((Q_2)\). The probability that \( j = 1 \), that is, that one restriction is active (i.e., \( j = 2 - 1 = 1 \), is 0.25 + 0.25 = 0.50 \((Q_2 \) and \( Q_3 \)). The probability that \( j = 0 \), that is, that both restrictions are active (i.e., \( j = 2 - 2 = 0 \), is 0.25 \((Q_4)\). Consequently, the penalty term value for the OR hypothesis \( H_3 \) can be computed as \( PT_3 = 1 + 0 \times 0.25 + 1 \times 0.50 + 2 \times 0.25 = 2 \). In addition, consider Figure 2b, where the parameter space is restricted by the order restrictions \( H_3 : \theta_1 \geq 0 \). Because the order restriction divides the unconstrained parameter space into two spaces, \( Q_1 \) and \( Q_2 \) are now two half-spaces. Again, assume that \( \Sigma = I \) and again we have two parameters (i.e., \( p = 2 \)), but now we only have one order restriction (i.e., \( q_1 = 1 \)). Consequently, there can be zero or at maximum one active restriction and thus the probability that \( j = 0 \), that is, that we have two active restrictions, is 0. This is because, if we impose one order restriction on two parameters, one parameter is allowed to vary freely, while the other parameter is restricted by the value of this free parameter. The probability that \( j = 1 \), that is, that the order restriction is active, is 0.5 \((Q_2)\). The probability that \( j = 2 \), that is, that the order restriction is not active, is 0.5 \((Q_1)\). Hence, the penalty term value for the OR hypothesis \( H_3 \) is computed by \( PT_3 = 1 + 0 \times 0 + 1 \times 0.5 + 2 \times 0.25 = 2.5 \).

Equipped with this knowledge, we rewrite the penalty term \( PT_m \). This becomes helpful for determining the penalty for the complement of \( H_m \), which is discussed later on. Let \( q_1 > 0 \) be the number of order restrictions and \( q_2 > 0 \) the number of equality restrictions. Then, \( p = q_1 + q_2 + F \), with \( F \) the number of remaining, free parameters. Then, the penalty for \( H_m \) can be rewritten as:

\[
PT_m = 1 + \sum_{j=0}^{p} j \times LP_j(p, \Sigma, H_m) = 1 + 0 \times LP_0 + 1 \times LP_1 + \ldots + p \times LP_p \\
= 1 + (F + 0) \times LP_{F+0} + \ldots + (p - q_2) \times LP_{p-q_2} \\
= 1 + (F + 0) \times LP_{F+0} + \ldots + (F + q_1) \times LP_{F+q_1} \\
= 1 + F + 0 \times LP_{F+0} + \ldots + q_1 \times LP_{F+q_1},
\]

using that \( LP_0 \) to \( LP_{F-1} \) are 0 (because \( F \) free parameters in \( H_m \) denote that there are always at least \( F \) “inactive” restrictions); and that \( LP_{F-q_2+1} \) to \( LP_F \) are 0 (because the \( q_2 \) equality restrictions in \( H_m \) are always active); and that the \( LP \)'s sum to 1.

From the penalty term value \( PT_m \), it follows that for \( q_1 \) order restrictions and \( q_2 \) equality restrictions, and thus \( F = p - q_2 \) free parameters, the penalty term value for a hypothesis with solely equality restrictions equals \( 1 + F = 1 + p - q_2 \), which equals the penalty term value of the AIC.

In case of order restrictions, the exact computation of the level probabilities when \( \Sigma \neq I \) and for \( q > 4 \) is a difficult task in general because the probabilities can no longer be expressed in a closed form. Fortunately, the probabilities can be approximated by using the multivariate normal probability distribution function with additional Monte Carlo steps (Grömping, 2010) or they can be computed easily and sufficiently precise by Monte Carlo simulation (Silvapulle & Sen, 2005; Wolak, 1987).

**Introduction of the GORIC for \( H_c \)**

Here, we introduce the GORIC for the complement of \( H_m \), which is defined as

\[
\text{GORIC}_c = -2 \times \text{LL}_C + 2 \times PT_c, \tag{5}
\]

where \( \text{LL}_C \) is the maximized log-likelihood value for the complement of \( H_m \) and \( PT_c \) is the penalty term value. Recall that for the computation of the GORIC value for \( H_m \) the order-restricted
hypothesis is required to be a ccc. However, the complement $H_c$ is in many cases not a ccc. For example, the complement of $H_3 : \theta_1 \geq 0, \theta_2 \geq 0$, that is, $H_c$, is constructed by the quadrants $Q_2, Q_3$, and $Q_4$ (see Figure 3a), and can be written as

$$\begin{align*}
\theta_1 &\leq 0 \land \theta_2 \geq 0 \quad (Q_2) \\
\theta_1 &\geq 0 \land \theta_2 \leq 0 \quad (Q_3) \\
\theta_1 &\geq 0 \land \theta_2 \geq 0 \quad (Q_4).
\end{align*}$$

Note that in this case, the complement can be written out easily but, for many hypotheses, it is a difficult or even impossible task to write up the complement. Because, $H_c$ is not a ccc, the LLc and the PTc values cannot be computed directly like the LLm and the PTm values.

### Computing the Log-Likelihood for $H_c$

To compute the LLc value, we first need to ascertain whether the restrictions in $H_c$ are in line with the data or not. If at least one inequality restriction is violated, then the data are automatically in line with the complement and the LLc value equals the LLm value. This is illustrated for $H_3 : \theta_1 \geq 0, \theta_2 \geq 0$ in Figure 3a, where the permissible area is $Q_1$, and the quadrants $Q_2, Q_3$, and $Q_4$ form the complement. Because the unconstrained mle’s $\hat{\theta}$ lie in $Q_1$, both restrictions are violated, the data are in line with the complement and the LLc value is equal to LLm. Note that the same applies if the mle’s lie in $Q_2$ or $Q_4$. On the other hand, if the data are in line with the restrictions in $H_3$, then we have to find the mle’s of $\theta$ that are closest to $\theta \in H_c$, given $\Sigma$, which is denoted by $\hat{\theta}_c$. Let us inspect a bivariate case with $\Sigma = I$ (solid circles), as depicted in Figure 3b. The solid circles of the contour plot indicate that the two parameters $\theta_1$ and $\theta_2$ are uncorrelated. As a reminder, the lines of the contour plot correspond to parameter values which have equal log-likelihood values and lines closer to $\hat{\theta}$ result in a higher log-likelihood value, because $\hat{\theta}$ is the value for which the log-likelihood is maximized (without imposing restrictions on the parameters). Clearly, the solution $\hat{\theta}_c$ is on the boundary of the restricted parameter space $H_m$. Because there are many boundary solutions (see thick black lines), we have to search for a solution that has the shortest distance between $\hat{\theta}$ and the two boundaries, given $\Sigma$. Fortunately, we do not have to investigate each point on the thick black lines but only the points $\hat{\theta}_{c1}$ and $\hat{\theta}_{c2}$. The point $\hat{\theta}_{c1}$ is computed by treating the inequality restriction for $\theta_1$ as equality restriction (i.e., $\theta_1 = 0, \theta_2 \geq 0$). Analogously, for the point $\hat{\theta}_{c2}$, where $\theta_2$ is treated as equality restriction (i.e., $\theta_1 \geq 0, \theta_2 = 0$). Thus, in general, there are in total $q_1$ possibilities to be investigated. Notably, in case of equality restrictions, all $q_2$ equalities are “freed.” The LL corresponding to that point that results in the highest log-likelihood value, given $\Sigma$, equals the LLc value (here, $\hat{\theta}_{c1}$).

As mentioned above, the solution of $\hat{\theta}_c$ is dependent on the covariance matrix $\Sigma$. To clarify this, again consider Figure 3b. The solid contour lines show the solution of $\hat{\theta}_c$ if $\Sigma$ is an identity matrix (i.e., $\hat{\theta}_{c1}$). It can easily been seen that, if the covariance matrix is not an identity matrix (e.g., see dot-dashed line) that the solution of $\hat{\theta}_c$ alters (here, it is $\hat{\theta}_{c2}$).

### Computing the Penalty Term Value for $H_c$

The penalty term value for $H_m$ can be seen as the expected number of “inactive” restrictions (i.e., $j = p - "the number of active restrictions"$ which implies that free parameters denote “inactive” restrictions) plus one for the variance term. The expected number of “inactive” restrictions is the sum of the expected number of “inactive” restrictions for $1 + p - q_2$ subspaces (where $0$ to $p - q_2$ restrictions are “inactive”). If we apply this

![Figure 3](image-url)
principle on the complement, then only two distinct subspaces can be distinguished (and not $1 + p - q_2$). The first subspace is the one not in agreement with the complement (cf. $Q_1$; i.e., the space fully in agreement with $H_m$). For this subspace, the number of “inactive” restrictions equals the number of free parameters in $H_m$, denoted by $F$. Note that the free parameters in $H_m$ ($F$) remain free in $H_c$ and the $q_2$ equality restrictions in $H_m$ are “freed” in $H_c$. Thus, $F = F + q_2 = p - q_1$. The probability of $F$ levels in $H_c$ ($LP_F$) equals the probability of having $p - q_1$ levels in $H_m$ ($LP_{p-q_1}$; cf. probability of ending up in $Q_1$, is $LP_{p-q_1} = LP_2 = 0.25$), that is, the probability finding the $mle$ in $H_m$ (under the null). The second subspace is the one fully in agreement with the complement (cf. not $Q_1$, that is, $Q_2, Q_3$, and $Q_4$) and has $p = F + q_1$ “inactive” restrictions (note that there are no equality restrictions in $H_c$). The corresponding level probability is denoted by $LP_{F+q_1}$ and equals the probability of not finding the $mle$ in $H_m$ but in $H_c$ which equals $(1 - LP_{p-q_1})$. Mimicking Equation 4, the penalty term value for $H_c$ is given by

$$PT_c = 1 + (F + 0) \times LP_{F+q_1} + (F + q_1) \times LP_{F+q_1} = 1 + (p - q_1) + q_1 \times (1 - LP_{p-q_1})$$

(7)

Another way of establishing $PT_c$ follows from the fact that the complement is the whole space minus the space fully in agreement with $H_m$ and, therefore, its penalty equals the penalty of the whole space (i.e., $1 + p$) minus the expected number of “inactive” restrictions in the space $H_m$. The expression for the latter equals the last part in Equation 4, that is, $q_1 \times LP_{F+q_1}$. Then, it follows that the expected number of “inactive” restrictions in $H_c$ equals

$$PT_c = 1 + p - q_1 \times LP_{F+q_1}$$

(8)

This expression is equal to Equation 7, because $F + q_1 = p - q_2$.

To illustrate, the penalty term value for the complement of $H_1$; $\theta_0 \geq 0, \theta_2 \geq 0$ is computed by $PT_c = 1 + 2 - 2 \times 0.25 = 2.5$ and the penalty term value for the complement of $H_c$; $\theta_0 \geq \theta_2$ is computed by $PT_c = 1 + 2 - 1 \times 0.5 = 2.5$. Because the complement of $H_c$ (i.e., $\theta_1 \leq \theta_2$) is a ccc, we can use the $PT_c$ formula as well, which also renders a penalty of 2.5. In Appendix A, we illustrate the computation of the $PT_m$ and the $PT_c$ values in case of three parameters.

In the next section, we show by means of a brief illustration and the results from a simulation study that the evidence ratio for the order-restricted hypothesis compared with its complement is boundless.

**Unbounded GORIC Weights in Case of the Complement**

Once the GORIC values for $H_m$ and its complement are known, the GORIC weights can be easily obtained as follows

$$w_s = \frac{\exp(-0.5(GORIC_m))}{\exp(-0.5(GORIC_m)) + \exp(-0.5(GORIC_c))}$$

(9)

where the subscript $s$ equals $m$ or $c$ for hypothesis $H_m$ and hypothesis $H_c$, respectively. From these weights, we can determine the evidence ratio for $H_m$ against its complement $w_m/w_c$. This ratio is interpreted as the strength of evidence for $H_m$ given the data and $H_c$ (Kuiper, 2011; Wagenmakers & Farrell, 2004). For example, for Figure 3b with $\Sigma = 1$ (e.g., dot-dashed lines), $n = 50$, and $f = 0.20$, the evidence ratio for $H_1$; $\theta_0 \geq 0, \theta_2 \geq 0$ compared with $H_c$ equals $w_m/w_c = 0.92/0.08 = 11.50$. This ratio tells us that hypothesis $H_3$ is the best out of the two (because it is larger than 1) and that $H_c$ is 11.50 times more supported than $H_1$. Notably, because we sampled from $H_c$, $H_1$ is true. To contrast, if we want to determine the evidence ratio for $H_3$ against the unconstrained hypothesis $H_0$, that is, $w_3/w_m$, we have to replace the GORIC, by the GORIC in Equation 9 and $s$ equals $m$ or $U$ for hypothesis $H_m$ and hypothesis $H_u$, respectively. Note that $w_3$ now not equals 0.92 from above, because the weights depend on the set of hypotheses. Therefore, if $H_u$ is replaced by $H_m$, the weights must be recomputed for the two hypotheses in the set. Then, the evidence ratio equals $w_m/w_u = 0.62/0.38 = 1.63$. This clearly shows the advantage of using the complement as competing hypothesis. Namely for the same data, we could obtain support for $H_3$ of 1.63 (the maximum support) or a support of 11.50 when comparing to $H_c$ (using $n = 50$).

In Appendix B, we present a simulation study in which we investigated the performance of the evidence ratio weights. The simulation results show the benefits of evaluating an OR hypothesis against its complement. While, for small effect-sizes and/or sample-sizes, the difference between the evidence ratio for the true $H_m$ when using the complement or unconstrained as competing hypothesis is minimal, the difference increases rapidly and profoundly for larger effect-sizes and/or sample-sizes. More importantly, the evidence ratio for the true $H_m$ against its complement is boundless for increasing effect-sizes and/sample-size and might therefore be more compelling, whereas the evidence ratio for the unconstrained hypothesis as competing hypothesis has an upper bound. Therefore, we recommend to replace the unconstrained hypothesis by the complement of the hypothesis of interest as competing hypothesis.

**Bums Example**

To illustrate the method, we analyze the empirical example introduced in the introduction in which we sought to determine possible risk factors for ruminating thoughts after a burn injury. The data are based on a cohort study consisting of 245 individuals with burns, aged 18- to 74-years-old. The response variable is rumination. Moreover, for the current illustration, we included gender ($0 =$ men, $1 =$ women) and facial burns ($0 =$ no, $1 =$ yes) together with its interaction as predictor variables and Hospital Anxiety and Depression Scale (HADS; $M = 3.85, SD = 3.66$), age ($M = 41.06, SD = 13.94$), and the number of surgical operations, which is a measure of severity of the burns ($SO; M = 1.14, SD = 1.76$) as covariates.

Reconsider, the hypothesis of interest $H_1: \{p_{\text{men, no facial burns}} - p_{\text{men, facial burns}}, p_{\text{women, no facial burns}} - p_{\text{women, facial burns}} \leq 0\}$. A natural choice to evaluate the OR hypothesis $H_1$ would be an order-restricted $2 \times 2$ ANCOVA model. Because an ANCOVA is just a special case of the linear regression model, the model can be written as a linear function. To obtain adjusted means for a person with an average score on the covariates, the covariates HADS, age and SO are centered at their average and are denoted by $Z_{\text{HADS}}, Z_{\text{age}}$ and $Z_{\text{SO}}$, respectively. Then, the model can be written as follows:
Rumination\(_i\) = \(\theta_1 + \theta_2\) \(\text{facialBurns}_i + \theta_3\) \(\text{gender}_i\)  
\(\times\) \(\text{facialBurns}_i + \theta_5\) \(\hat{Z}_i\) \(\text{HADS}_i\)  
\(\hat{\theta}_6\) \(\hat{Z}_i\) \(\text{age}_i + \theta_7\) \(\hat{Z}_i\) \(\text{SO}_i + \epsilon_i\),  
for \(i = 1, \ldots, 245\).

On the left-hand side of the = operator, we have the response variable rumination and on the right-hand side we have the factors facial burns and gender and its interaction, and the centered covariates \(\hat{Z}_\text{HADS}, \hat{Z}_\text{age}\) and \(\hat{Z}_\text{SO}\). The interaction between gender and facial burns is included using the \(\times\) operator. Then, the four adjusted means with average scores on the covariates are computed as:

\[
\begin{align*}
\mu_{\text{men}, \text{no facial burns}} &= \theta_1 \\
\mu_{\text{men}, \text{facial burns}} &= \theta_1 + \theta_2 \\
\mu_{\text{women}, \text{no facial burns}} &= \theta_1 + \theta_3 \\
\mu_{\text{women}, \text{facial burns}} &= \theta_1 + \theta_3 + \theta_4.
\end{align*}
\]

The R (R Core Team, 2019) code and the output from the analyses can be found in Appendix C.

The results show that the OR hypothesis \(H_1\) is 0.891/0.109 \(\approx\) 8.198 times more supported by the data than its complement. For comparison, the results for the un constrained hypothesis show that hypothesis \(H_1\) is 0.734/0.266 \(\approx\) 2.754 times more supported by the data, which is its maximum support. Assuming \(H_1\) is true, using the complement of \(H_m\) instead of \(H_m\), we have now more compelling evidence.

**Summary and Discussion**

In this article, we introduced the evaluation of an order-restricted (OR) hypothesis against its complement using the GORIC (weights). The GORIC is an information criterion that can be used to evaluate competing hypotheses in univariate and multivariate normal linear models, where the regression parameters \(\theta\) are subject to linear (in)equality restrictions. The interpretation can be improved by computing GORIC weights and related evidence ratios reflecting the strength of evidence for one hypothesis versus another.

We advise that one should evaluate their theory against its complement \(H_1\) instead of the unconstrained hypothesis \(H_m\). The advantage of our method is that the evidence ratio for an OR hypothesis \(H_m\) compared with its complement is bounded. The evidence ratio for \(H_m\) compared to \(H_m\) is neither increased by a larger sample-size nor by a larger effect-size, if the data are in agreement with the hypothesis of interest (i.e., theory). Consequently, using the complement as competing hypothesis leads to much more support for the hypothesis of interest assuming it is true, compared with using the unconstrained hypothesis as competing hypothesis. In case the complement is not true, then the results are comparable to evaluating the order-restricted hypothesis against the unconstrained hypothesis. Furthermore, in case that \(H_m\) is almost true, then there is less support for \(H_m\) when compared with its complement than when compared with the unconstrained hypothesis. This is because the log-likelihood values are almost identical and the difference between the penalty term values is larger between \(H_m\) and \(H_m\) than between \(H_m\) and \(H_m\). Consequently, in such cases, it is also better to evaluate \(H_m\) against its complement as the support for \(H_m\) against \(H_m\) might be overstated because \(H_m\) is not true. Besides that the complement has practical benefits, it is also more substantive: When looking at the interpretation of the complement and the unconstrained hypothesis, the latter can be seen as all possible theories including the target hypothesis \(H_m\), whereas the former denotes the possible theories without the target hypothesis \(H_m\). We believe that this makes more sense if you would like to know whether your theory is better than all other theories (and how much). Comparing your theory against all possible theories including the targeted one makes less sense to us.

The method was illustrated using an empirical example about facial burn injury. In six easy steps (shown in Appendix C), we showed how to compute the evidence ratio of the researchers theory against its complement using the R package *restrktor* (see http://www.restrktor.org).

We assumed that researchers often do not have specific competing hypotheses. While, this is probably often the case, it is conceivable that the set of hypotheses contains more than one competing hypothesis. In these cases, the problem that the evidence ratio for \(H_m\) against \(H_m\) is not affected by increasing sample-size and/or effect-size after a specific value can still occur. For example, consider the set with three hypotheses: \(H_2, H_3 : \mu_1 \leq \mu_2 \leq \mu_3 = \mu_4\) (which is a subset of \(H_2\)) and the unconstrained hypothesis \(H_m\). If \(H_m\) is true, then all three hypotheses are true and all evidence ratios are bounded (Kuiper et al., 2011, p. 107). However, the evaluation of a set of multiple OR hypotheses against its complement is less straightforward because determining the complement for multiple hypotheses might not always be trivial (especially for software).

The results presented in this article are for the univariate linear regression model but fortunately they can easily be adapted for the multivariate normal linear model. One should keep in mind that, unlike in the univariate setting, where \(\hat{\theta}\) does not depend on the order-restricted covariance matrix, denoted by \(\Sigma\), in the multivariate normal linear model \(\hat{\theta}\) does depend on \(\Sigma\) and \(\Sigma\) on \(\hat{\theta}\) (Kuiper et al., 2012). Hence, an iterative procedure is needed to calculate them. The procedure is implemented in *restrktor*.

**References**


### Appendix A

**Example of Computing the PT<sub>c</sub> in Case of Three Parameters**

Consider Figure A1, the unconstrained parameter space is determined by the three parameters \( \theta_1, \theta_2, \) and \( \theta_3 \) (and is of course the whole space and not just the one depicted in Figure A1). The gray shaded area is a closed convex cone and is defined by the order restrictions \( H_{A1} : \theta_1 \leq \theta_2, \theta_1 \leq \theta_3 \). The level probabilities corresponding to \( H_{A1} \) (assuming \( \Sigma = I \)) equal \( LP_0 = 0, LP_1 = \frac{1}{6}, LP_2 = \frac{1}{2} \) and \( LP_3 = \frac{3}{4} \). Next, we will elaborate on the values for \( LP_0 \) and \( LP_3 \). The level probability \( LP_3 = \frac{3}{4} \) is the probability that the vector with OR estimates \( \hat{\theta} \) is identical to the unconstrained estimates (i.e., \( \hat{\theta} = \theta \) and \( j = 3 \)). This probability equals the proportion of the gray shaded area compared with the whole cube and of course also of \( H_{A1} \) versus the whole space. A level probability of \( LP_0 = 0 \), that is, the probability that there are \( j = 0 \) levels, which means that it is impossible that the vector with OR estimates \( \hat{\theta} \) has three active restrictions, which makes sense with two order restrictions. Hence, the penalty term value for \( H_{A1} \) equals \( PT_{A1} = 1 + 0 \times 0 + \frac{1}{6} \times 1 + \frac{1}{2} \times 2 + \frac{1}{3} \times 3 = \frac{31}{6} \).

Using the expression in Equation 8, the penalty term value for \( H_c \) equals \( PT_c = 1 + p - q_1 \times LP_{p-1} - 1 = 1 + 3 - 2 \times \frac{1}{3} = \frac{31}{4} \). For comparison, \( PT_w = 1 + 3 = 4 \).

![Figure A1](https://example.com/figure.png)

*Figure A1.* The permissible, gray area is defined by \( H_{A1} : \theta_1 \leq \theta_2, \theta_1 \leq \theta_3 \), depicted for \( \theta_1, \theta_2, \) and \( \theta_3 \), between \(-1 \) and \( 1 \).

(Appendices continue)
**Appendix B**

**Simulation Study**

**Design**

We generated 500 samples according to the ANOVA model\(^5\)

\[
y_i = \mu_1 x_{i1} + \ldots + \mu_k x_{ik} + \epsilon_i, \quad i = 1, \ldots, n
\]

where we assume that the residuals are standard normally distributed. We considered the OR hypothesis \(H_1 : \mu_1 = \mu_2 = \mu_3 = \mu_4\), its complement \(H_c : \mu_1\), \(\mu_2\), \(\mu_3\), \(\mu_4\). Note that \(H_c\) does not equal \(\mu_1 = \mu_2 = \mu_3 \neq \mu_4\); it does contain this but also the other (22) orderings of simple ordering combinations of \(\mu_1\) to \(\mu_4\) (excluding the one ordering in \(H_m\)). Data were generated under hypothesis \(H_1\) with four uncorrelated independent means of size \(n = 30, 50, 100, 200, 500\) per group and for a variety of differences among the population means, using effect-size \(f = 0.10, 0.20, \ldots, 1\) (Cohen, 1988, pp. 274–275). Notably, \(f = 0\) corresponds to sampling from the boundary of both \(H_m\) and \(H_c\). If we sample values from a \(H_1\) population with increasing effect-size, this will evidently lead to more and more support for \(H_1\). Let the differences between the means, \(d\), be equally spaced, where \(d\) is defined as

\[
d = \frac{2\sqrt{\frac{\sigma^2}{2(I-1)p^2}}}{I-1}
\]

under the restriction that \(\sum_{i=1}^{p} \mu_i = 0\) and \(\sigma = 1\). Then, the \(p\) ordered means can be computed as

\[
\mu_1 = \frac{-\left(\frac{p-1}{I\sigma}\right)}{2} + (i-1)d.
\]

Appendix D shows the computed population means for the various effect-sizes (\(f\)).

**Simulation Results**

All results are obtained using the R package restriktor (see [http://www.restriktor.org](http://www.restriktor.org)) employing the GORIC function. The results of the simulation study are presented in Figure B1a, B2, and B3, and are obtained by computing the mean value of the relative evidences in each of the 500 simulations. Furthermore, to improve visibility, we took the natural logarithm values of the means and we used a varying range of sample-sizes and effect-sizes.

The results clearly illustrate the benefits of evaluating \(H_m\) versus its complement: The mean evidence ratio for \(H_1\) versus \(H_m\) (mean \(w_i/w_m\)) increases rapidly for larger effect-sizes (see Figure B1a) and sample-sizes (see Figure B1b), while the mean evidence ratio using the unconstrained hypothesis as competing hypothesis (mean \(w_i/w_{cu}\)) is clearly bounded after a certain value (see Figures B1c and B1d, respectively). To illustrate, consider for example Figure B1a, where the mean evidence ratio for \(H_1\) versus \(H_m\) (mean \(w_i/w_m\)) for a medium effect-size (\(f = 0.30\)) and \(n = 100\) is \(\exp(2.63) \approx 13.87\) (on the original scale), while the mean evidence ratio for \(H_1\) versus \(H_m\) (mean \(w_i/w_{cu}\)) is bounded at \(\exp(1.92) \approx 6.82\). Note that the value 1.92 equals the difference in penalty term values; with \(PT_u - PT_1 = (1.00 + 4.00) - (1.00 + 2.08) \approx 1.92\), which equals the difference in GORIC values, because the log-likelihood values are here the same (i.e., \(LL_m = LL_1\)).

For small effect-sizes and small samples, the mean evidence ratio for \(H_1\) using \(H_m\) is slightly lower than when using \(H_{cu}\). For example, for \(f = 0.10\) and \(n = 30\) the mean evidence ratio for \(w_i/w_m\) is \(\exp(1.50) \approx 4.48\) and for \(w_i/w_{cu}\) the mean evidence ratio is \(\exp(1.61) \approx 5.00\). In this case, using the complement is a bit more conservative; although the conclusion is not different of course. Furthermore, the evidence ratio for small effect-sizes (\(f \leq 0.20\)) does not increase very rapidly (see Figure B1b), independent of sample-size. This is because, when examining small effects, the complement is often true (even though the data were generated under \(H_1\)). This is illustrated in Figure B2. For example, if \(f = 0\), the mle’s are (except from some sampling variation) in 23/24 (approximately 95.8%) of the time not in agreement with \(H_1\) and thus in agreement with \(H_c\). Thus, both hypotheses \(H_1\) and \(H_m\) have the same maximized log-likelihood value with a probability of \(\text{prob}_{w_m} = 23/24\). When \(f\) increases, the data/mle’s will be more and more in agreement with \(H_1\), and thus not with its complement \(H_c\) and hence the proportion of equal maximized log-likelihood values of \(H_m\) and \(H_c\) (and thus \(\text{prob}_{w_m}\)) decreases. Logically, the proportion of equal maximized log-likelihood values of \(H_1\) and \(H_{cu}\), that is \(1 - \text{prob}_{w_{cu}}\), then increases.

The presented results so far are for the scenario that \(H_1\) is true, but we are also interested in the performance if \(H_1\) is not true (i.e., \(H_c\) is true). Figure B3 shows the results for the situation that the complement is true. Data were generated under the complement of \(H_1\), for which we choose \(H_c : \mu_1 = \mu_2 = \mu_3 \neq \mu_4\). The means are given in Appendix D and are now in reversed order compared with the previous simulation. Again, we considered the OR hypothesis \(H_1\), its complement \(H_c\), and the unconstrained hypothesis \(H_{cu}\). The results in Figure B3a show that the mean evidence ratio for \(H_1\) versus \(H_c\) (mean \(w_i/w_c\)) and for \(H_1\) versus \(H_{cu}\) (mean \(w_i/w_{cu}\)) decreases rapidly for larger \(f\). This is because, when the effect-size and/or the sample-size increases, the data/mle’s will be more and more in agreement with the complement \(H_c\) and of course also with the unconstrained hypothesis \(H_{cu}\). The results shown in Figure B3b are based on the same numerical results shown in Figure B3a but now for \(H_c\) versus \(H_1\) (mean \(w_i/w_1\)) and for \(H_{cu}\) versus \(H_1\) (mean \(w_i/w_1\)). They clearly show the nice property that if the complement (and also \(H_c\)) is true, both evidence ratios \(w_i/w_1\) and \(w_i/w_1\) show more support for larger effect-sizes and samples sizes. Stated otherwise, both are boundless.

---

5 Note that the ANOVA model is a special case of the multiple regression model discussed in the main text.
Figure B1. Mean of the evidence ratio (on a log scale) for the situation that the OR hypothesis $H_1$ is true (based on 500 simulations). (a) Hypothesis $H_1$ is compared with its complement $H_c$ (mean $w_1/w_c$), for various effect-sizes ($f$) and for $n = 30, 50, 100$ and $200$. (b) Hypothesis $H_1$ is compared with its complement $H_c$ (mean $w_1/w_c$), for various sample-sizes ($n$) and for $f = 0.10, 0.20, 0.30$ and $0.40$. (c, d) same as (a, b) but now for hypothesis $H_1$ versus the unconstrained hypothesis $H_u$ (mean $w_1/w_u$).

(Appendices continue)
Figure B2. Proportion of data sets (based on 500 simulations) that result in equal log-likelihood values for the complement of $H_1$: $\mu_1 \leq \mu_2 \leq \mu_3 \leq \mu_4$, that is, $H_c$, and the unconstrained hypothesis $H_u$ (i.e., $LL_c = LL_u$), for various effect-sizes ($f$) and $n = 30, 50, 100, 200$.

Figure B3. Mean of the evidence ratio (based on 500 simulations) for the situation that the complement $H_c$ of the OR hypothesis $H_1$ is true, for various effect-sizes ($f$) and for $n = 30$ and 200. (a) The closed circles denote $H_1$ versus $H_c$ (mean $w_1/w_2$), for $n = 30$ (solid line) and $n = 200$ (dashed line). The closed triangles denote $H_1$ versus the unconstrained hypothesis $H_u$ (mean $w_1/w_u$). (b) Same as (a) but now for $H_c$ versus $H_1$ (mean $w_c/w_1$) and $H_u$ versus $H_1$ (mean $w_u/w_1$), for $n = 30$ and $n = 200$. Note that Figure B3b is on a log scale.

(Appendices continue)
Appendix C

R Code to Run the Burns Example

In what follows, we describe all steps to compute the evidence ratio for hypothesis $H_1$ compared with its complement and $H_1$ compared with the unconstrained hypothesis using the R package \texttt{restriktor}.

Step 1: Load your data set into R.

```r
burns <- read.csv("burns.csv", header = TRUE, sep = " ")
```

More information about how to load your data into R, can be found online at http://restriktor.org/tutorial/importdata.html.

Step 2: Center the covariates HADS, age, and SO at their average. This can be done in R as follows:

```r
burns$Z_HADS <- burns$HADS - mean(burns$HADS, na.rm = TRUE)
burns$Z_age <- burns$age - mean(burns$age, na.rm = TRUE)
burns$Z_SO <- burns$SO - mean(burns$SO, na.rm = TRUE)
```

Step 3: Fit the unconstrained linear regression model using the \texttt{lm()} function.

```r
fit.lm <- lm(Rumination ~ 1 + gender + facialBurns + gender:facialBurns + Z_HADS + Z_age + Z_SO, data = burns)
```

For clarity reasons, we explicitly added an intercept term by specifying the value 1. The interaction between gender and facial burns is included using the \texttt{:} operator.

Step 4: Create the restriction syntax for \texttt{restriktor}. Now that the model is defined in R, we are left with specifying the order restrictions. This is done in \texttt{restriktor} by specifying a so-called restriction syntax. Order restrictions are defined by means of inequality restrictions ($<$ or $>$) or by equality restrictions ($=$). In addition, a convenient feature of the \texttt{restriktor} restriction syntax is the option to define new parameters that are linear in the original model parameters. This can be done using the := operator.

In this way, we can compute the four adjusted means and impose order restrictions among these means. The restriction syntax is enclosed within single quotes. Then, for hypothesis $H_3$, the restriction syntax might looks as follows:

```r
myRestrictions <- 'm1 := .Intercept.
m2 := .Intercept. + facialBurns
m3 := .Intercept. + gender
m4 := .Intercept. + facialBurns + gender + gender.facialBurns
m1 < m4
m2 < m4
m3 < m4 '
```

It is important to note that variable/factor names of the interaction effects in objects of class \texttt{lm} contain a semicolon (:) between the variable names (e.g., \texttt{gender:facialBurns}). To use these parameters in the restriction syntax, the semicolon must be replaced by a dot (\texttt{.}); e.g., \texttt{gender.facialBurns}). In addition, the intercept of a fitted objects of class \texttt{lm} is denoted in the output as (\texttt{Intercept}) and not as 1 anymore. To use the intercept in the restriction syntax, the parentheses must also be replaced by a dot (\texttt{i.e., .Intercept.}). More information about the restriction syntax can be found online at http://restriktor.org/tutorial/syntax.html.
Step 5: Compute the GORIC weights and the evidence ratio using the `goric()` function from the `restriktor` package.

```r
out.c <- goric(fit.lm,
               constraints = myRestrictions,
               comparison = "complement")
```

The first argument to the `goric()` function is the unconstrained fitted object of class `lm`. The second argument is the restriction syntax specified in the previous step. To compare $H_3$ with its complement $H_c$, the argument `comparison` has to be set to "complement" (by default it is set to "unconstrained").

Step 6: Interpret the results. A brief overview can be requested using the `print()` function. A more detailed overview can requested by the `summary()` function as follows:

```r
summary(out.c, brief = TRUE)
```

Restriktor: generalized order-restricted information criterion (GORIC):

Results:

<table>
<thead>
<tr>
<th>model</th>
<th>loglik</th>
<th>penalty</th>
<th>goric</th>
<th>goric.weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 myRestrictions</td>
<td>-660.024</td>
<td>6.987</td>
<td>1334.022</td>
<td>0.891</td>
</tr>
<tr>
<td>2 Complement</td>
<td>-661.942</td>
<td>7.173</td>
<td>1338.230</td>
<td>0.109</td>
</tr>
</tbody>
</table>

Relative GORIC-weights:

<table>
<thead>
<tr>
<th></th>
<th>vs myRestrictions</th>
<th>vs complement</th>
</tr>
</thead>
<tbody>
<tr>
<td>myRestrictions</td>
<td>1.000</td>
<td>8.198</td>
</tr>
<tr>
<td>complement</td>
<td>0.122</td>
<td>1.000</td>
</tr>
</tbody>
</table>

The order-restricted hypothesis myRestrictions has 8.198 times more support than its complement.

Below are the results for $H_1$ compared with the unconstrained hypothesis.

```r
out.u <- goric(fit.lm,
                constraints = myRestrictions,
                comparison = "unconstrained")
summary(out.u, brief = TRUE)
```

Restriktor: Generalized order-restricted information criterion (GORIC):

Results:

<table>
<thead>
<tr>
<th>model</th>
<th>loglik</th>
<th>penalty</th>
<th>goric</th>
<th>goric.weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 myRestrictions</td>
<td>-660.024</td>
<td>6.987</td>
<td>1334.022</td>
<td>0.734</td>
</tr>
<tr>
<td>2 unconstrained</td>
<td>-660.024</td>
<td>8.000</td>
<td>1336.048</td>
<td>0.266</td>
</tr>
</tbody>
</table>

Relative GORIC-weights:

<table>
<thead>
<tr>
<th></th>
<th>vs myRestrictions</th>
<th>vs unconstrained</th>
</tr>
</thead>
<tbody>
<tr>
<td>myRestrictions</td>
<td>1.000</td>
<td>2.754</td>
</tr>
<tr>
<td>unconstrained</td>
<td>0.363</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: In case of equal log-likelihood (loglik) values, the relative weights are solely based on the difference in penalty values.

(Appendices continue)
Appendix D
Population Means for First Simulation Study

<table>
<thead>
<tr>
<th>Effect-size</th>
<th>Population means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>μ₁</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>.1</td>
<td>-1.34</td>
</tr>
<tr>
<td>.2</td>
<td>-0.268</td>
</tr>
<tr>
<td>.3</td>
<td>-0.402</td>
</tr>
<tr>
<td>.4</td>
<td>-1.341</td>
</tr>
</tbody>
</table>

Note. In the second simulation, we used the reverse ordering of these means.