Distinguishing monophonies from polyphonies using Weibull bivariate distributions

Hélène Lachambre, Régine André-Obrecht, Julien Pinquier
IRIT – Université de Toulouse
118 route de Narbonne – 31062 Toulouse Cedex 9 – France
{lachambre, obrecht, pinquier}@irit.fr

Abstract—In the context of music indexation, it would be useful to have a precise information about the number of sources performing: a source is a solo voice or an isolated instrument which produces a single note at any time. This article discusses the automatic distinction between monophonic music excerpts, where only one source is present, and polyphonic ones. Our method is based on the analysis of a “confidence indicator”, which gives the confidence (in fact its inverse) on the current estimated fundamental frequency (pitch). In a monophony, the confidence indicator is low. In a polyphony, the confidence indicator is higher and varies more. This leads us to compute the short term mean and variance of this indicator, take this 2-dimension vector as the observation vector and model its conditional distribution with Weibull bivariate models. This probability density function is characterized by five parameters. A method to perform their estimation is developed (in theory and practice). The decision is taken considering the maximum likelihood, computed over one second. The best configuration gives a global error rate of 6.3 %, performed on a balanced corpus (18 minutes in total).

Index Terms—Monophony, Polyphony, Weibull bivariate distribution

I. INTRODUCTION

In the context of music indexation or annotation, much research has been conducted to determine the rhythm of a melody [1], [2], its tonality [3], the performing instruments [4], [5], [2], or the (potentially multiple) pitch(es) of the instrument(s). In the studies intending to extract multiple pitches, the number of sources is either known or not. When it is unknown [6], [7], the number of notes is found a posteriori considering the number of estimated pitches. Even in [8], in which this problem – polyphony inference – is studied for itself, the number of sources is found a posteriori. The existing algorithms have to estimate at the same time the sources and their number. We think that a priori knowing the number of sources could allow to build more efficient algorithms.

The question explored here is the automatic distinction between monophonic musical excerpts – a solo voice or an isolated instrument which plays a single note at any given time, and polyphonic excerpts.

A few papers have been published recently on related subjects. Tsai et al. [9] suggest to distinguish solo singing voices from duos, as a preprocessing for singer identification. This can be viewed as a sub-problem of the one addressed here. Modeling the repartition of MFCC with GMM, they achieve an accuracy of 96 %.

Smit and Ellis [10] present a method to distinguish solo singing voice from polyphonic music. Their algorithm is based on the following idea: considering a frame, is it possible to model it with a single periodic function? If it is, the frame is considered monophonic; otherwise it is polyphonic. With this method, they achieve a precision and recall, both around 70 %. They compare their system with a base-line: MFCC, modeled by GMM. They conclude that, for this larger task (larger than the one addressed by Tsai et al., the base-line is unsuccessful: for a precision of 70 %, the recall is only about 50 %.

Our method is based on the fact that in polyphonic segments, the presence of several notes at the same time makes the pitch extraction more difficult. We have been inspired by a “confidence indicator” proposed by de Cheveigné (YIN [11]), which measures the “inverse” of the confidence on the current estimated pitch. In the case of a monophony, where the pitch is relatively easy to estimate, the “confidence indicator” is low. In a polyphonic context, where the pitch is more difficult to determine, multiple pitches may appear, not only the “confidence indicator” is higher, but its variation is larger. These remarks lead us to compute the short term mean and variance of the indicator over an analysis window of 50 ms, and to model their conditional distributions with Weibull bivariate distributions. A reference model is estimated for each class (monophony and polyphony), and the decision is taken by computing the maximum likelihood. The global scheme is presented on figure 1.

In part II, we present our method. Part III is dedicated to the description of the experimental corpus and the assessment of the proposed system, including comparisons with classical classifiers. The new method for the estimation of the parameters of the Weibull bivariate distribution is presented in appendix.

Fig. 1. Global scheme of the system.

II. STATISTICAL BEHAVIOR OF THE CUMULATIVE MEAN NORMALIZED DIFFERENCE

A. The YIN estimator

In [11], de Cheveigné and Kawahara present a method, named YIN, to estimate the pitch.
The algorithm is based on the search of the minimum of a function, the cumulative mean normalized difference. First, the following difference function is computed:

\[ d_t(\tau) = \sum_{k=1}^{N} (x_{t+k} - x_{t+k+\tau})^2 \]  

with \( N \) the window sample size and \( x_t \) the sample at frame \( t \).

Obviously, in the case of a periodic function of period \( T \), \( d_t(T) = 0 \), and reciprocally. However, in practice, \( d_t(T) \) is small, but not equal to zero, due to imperfect periodicity. Some low minima can also appear for small values of \( \tau \), due for example to the presence of high frequency noise. So finding the minimum of this function in order to determine the frequency would fail too often. To solve these problems, the authors propose to use the Cumulative Mean Normalized Difference, which essentially modifies the first values of \( d_t(\tau) \):

\[ d_t'(\tau) = \begin{cases} 
1 & \text{if } \tau = 0 \\
\frac{d_t(\tau)}{\tau} \left( \frac{1}{\tau} \sum_{k=1}^{n} d_t(k) \right) & \text{otherwise}
\end{cases} \]  

If \( T = \arg \min \), \( d_t'(\tau) \), then \( T^{-1} \) is an estimate of the fundamental frequency, and \( 1/d_t'(T) \) can be considered as a “confidence indicator”: the smaller \( d_t'(T) \) is, the more periodic the signal is, the more confident the estimation of \( T \) is. In our study, at each frame \( t \), we will study the value of \( d_t'(T) \), which will be named \( C(t) \). \( C(t) \) is computed every 10 ms.

**B. Definition of the observation 2-vector**

Since \( C(t) \) is the inverse of a confidence indicator, it should be low in the case of a single note (instrument or singer), and high in presence of multiple notes. We may also presume that in a polyphony, if an instrument is louder than the others, \( C(t) \) can lower a bit: the variance of \( C(t) \) should be higher for polyphonies. Figure 2 shows the behavior of the \( C(t) \) function for a monophonic and a polyphonic excerpt (5 seconds each). This observation confirms our assumptions, and thus leads us to study the joint behavior of the short term mean \( C_m(t) \) and variance \( C_v(t) \) of \( C(t) \). They are both computed on a sliding 50 ms window (5 frames), every 10 ms.

\[ C(t) \] differs from the polyphonic one.

**C. Weibull bivariate distribution**

Considering these histograms, we model the bivariate distribution of the observation vectors with a Weibull bivariate distribution. We choose the one described by Hougaard [12], and extended by Lu and Bhattacharya [13].

Its cumulative distribution function is given by:

\[ F(x, y) = 1 - \exp \left( - \left[ \frac{x}{\theta_1} \right]^\beta_1 + \left( \frac{y}{\theta_2} \right)^\beta_2 \right)^\delta \]  

for \((x, y) \in \mathbb{R}^+ \times \mathbb{R}^+\), with \((\theta_1, \theta_2) \in \mathbb{R}^+ \times \mathbb{R}^+\) the scale parameters, \((\beta_1, \beta_2) \in \mathbb{R}^+ \times \mathbb{R}^+\) the shape parameters and \(\delta \in [0, 1]\) the correlation parameter.

The choice of this distribution is motivated by the fact that two very different shapes have to be modelled, a goal that Weibull distributions are well known to achieve.

This choice is validated by the Kolmogorov test: compared to gaussian and exponential distributions, Weibull is the only one that passes the test.

Figure 4 shows the different Weibull bivariate distributions estimated for each class, obtained on the training corpus which is described in the next section. As for the distribution histograms, the monophonic shape strongly differs from the polyphonic one.
III. EXPERIMENTS AND RESULTS

Our final task is to distinguish between two classes: monophony and polyphony. In order to assess the two alternative decision systems (see sections III-C1 and III-C3, it is necessary to have musical excerpts labelled into 5 subclasses: single instrument or single singer (monophonies), several instruments, several singers, or instrument(s) AND singer(s) (polyphonies).

A. Corpus

For the purpose of our experiments, a home made corpus is used, that contains all classes and subclasses considered in our study. The extracts come from commercial recordings. These recordings were made in live or in studio. Our corpus is balanced: both classes, as well as all subclasses, contain sufficient and roughly equal amount of data. The repartition of the corpus between classes and subclasses is described in tab. I.

Our corpus contains music of various styles for each class: rock, popular music, classical music, jazz, country, ... In the “multiple instruments” class, the size of the orchestra are very different, from duos, trios or small groups (rock, jazz), to symphonic orchestra or big bands. The training set contains jazz, Pop, modern, renaissance and baroque music. In the test set, we also have classical, rock and country.

The “multiple singers” class contains, we also have duos, trios, small and huge choirs. These ensembles are composed either men and women (train and test), of men only, or of women only (test).

The “instruments and singers” class contains music in which at least one instrument, and at least one singer perform at each instant. It contains jazz, Pop, modern music (train and test), opera, country, Rap and rock&roll (test).

The “single instrument” class contains instruments such as violin, recorder, double bass (train and test), oboe, flute, guitar, clarinet, triangle, piano, trumpet and cello (test).

Finally, the “single singer” class contains extracts from twelve different singers, males and females, professionals and amateurs.

B. Training phase

Two classes are considered, which are subdivided into five subclasses. So seven Weibull bivariate models are trained, one for each class and one for each subclass.

For each subclass, the model is trained with 25 s of signal (5 s of 5 different musical excerpts). As the mean \( \mu(t) \) and the variance \( \sigma^2(t) \) are computed every 10 ms, 2500 statistical samples are used to estimate the five parameters of the Weibull bivariate distribution associated to each model.

For each class, the models are trained with all the data of the corresponding subclasses: 50 s from 10 different excerpts for the monophonic class model, 75 s from 15 different excerpts for the polyphonic class model, corresponding to 5000 and 7500 statistical samples.

No musical excerpt is common between the training corpus and the test corpus.

C. Results

The performances are measured using the overall error rate: \( \text{err} = \frac{\text{Number of seconds misclassified}}{\text{Total number of seconds}} \). For each experiment, the confusion matrix between both classes is also given.

1) Primary system – Class approach: In this primary system, two models are learned, one for each class. The confusion matrix is presented in tab. II, the overall error rate is 8.5%.

We note that there are more classification errors in the monophonic class. These errors are mainly due to single singers or single instruments performing on a fast tempo. Indeed, the YIN estimator needs a minimum duration to
estimate a note. If a note is too short, the estimation is made considering either two consecutive notes, or the note and some noise, which skews the result. The estimator performs as if there are two simultaneous notes: $C(t)$ behaves as in polyphonic context.

The few errors made in the other way (polyphonies recognized as monophonies) are due to very harmonic chords, such as major triads.

2) **Comparison with classical systems:**

   a) **Comparison with a baseline system:** Our method is first compared with the classical approach, which we name the “baseline” classifier. During a sequence of experiments, two parameters configurations are tested:

   - Energy and 12 MFCC.
   - Energy, $\Delta$Energy, 12 MFCC and $\Delta$MFCC.

   The distribution of these parameters (in both cases) is modeled with GMM. We make the number of GMM component vary (1, 2, 4, 8, 16, 32, 64, 128, 256), with diagonal covariance matrices. For the 1-component GMM, we also test the full covariance matrix configuration.

   The best configuration appears to be: 26 coefficients, modeled with a 16 component GMM, with diagonal covariance matrix.

   The overall error rate is **19.2 %**. The confusion matrix is presented in tab. III.

### TABLE II

**Confusion matrix - 2 Weibull bivariate models.**

<table>
<thead>
<tr>
<th></th>
<th>Monophony</th>
<th>Polyphony</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monophony</td>
<td>82.1 %</td>
<td>17.9 %</td>
<td>100 %</td>
</tr>
<tr>
<td>Polyphony</td>
<td>1.3 %</td>
<td>98.7 %</td>
<td>100 %</td>
</tr>
</tbody>
</table>

This experiment first proves the suitability of our parameters, if compared with the previous experiment. With the same modeling (Gaussian), the global error rate is almost divided by two.

It secondly shows that Weibull bivariate distributions are better than Gaussian ones: the global error rate falls from 11.4 % to 8.5 %. This is because the Gaussian distributions are not able to capture the main difference between the two histograms: not only their mean and variance are specific, but also their shapes. Weibull bivariate distributions, are able to capture this shape difference.

   c) **Our parameters with independent Weibull models:** To assess the use of bivariate models, the Weibull bivariate models are replaced by two independent Weibull univariate models, one modeling $C_m$, the other one modeling $C_v$. The confusion matrix is presented in tab. V, the overall error rate is **15.5 %**.

### TABLE III

**Confusion matrix - Baseline classifier.**

<table>
<thead>
<tr>
<th></th>
<th>Monophony</th>
<th>Polyphony</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monophony</td>
<td>88.3 %</td>
<td>11.7 %</td>
</tr>
<tr>
<td>Polyphony</td>
<td>25.4 %</td>
<td>74.6 %</td>
</tr>
</tbody>
</table>

The following observations can be done:

- With only the Energy and 12 MFCC (without $\Delta$), the global error rate is between 30 and 35 %, for a number of GMM components between 4 and 256.
- With the derivatives, the optimal number of GMM components is 16, but numbers between 1 and 64 do not change much the results: the error rate is always around 20 %.

The results clearly show that these parameters and modeling are not appropriate for this task. The MFCC do not represent the difference between monophonic sounds and polyphonic sounds. The distributions of the MFCC are therefore not enough different for the two classes. This conclusion is consistent with the study of Smit and Ellis mentioned in introduction.

b) **Our parameters with a Gaussian bivariate model:** In this experiment, the use of our parameters as well as the use of Weibull bivariate models are assessed. In the proposed scheme, the Weibull bivariate models are replaced by Gaussian bivariate distributions (with full covariance matrix). The confusion matrix is presented in tab. IV, the overall error rate is **11.4 %**.

### TABLE IV

**Confusion matrix - 2 Gaussian bivariate models.**

<table>
<thead>
<tr>
<th></th>
<th>Monophony</th>
<th>Polyphony</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monophony</td>
<td>89.1 %</td>
<td>10.9 %</td>
</tr>
<tr>
<td>Polyphony</td>
<td>11.7 %</td>
<td>88.3 %</td>
</tr>
</tbody>
</table>

With this experiment, we show that the correlation between the short term mean and variance is informative. Indeed, the correlation is higher in the monophonic case than in the polyphonic case, as we can see on figure 2.

d) **Comparison with a SVM:** Finally, the probabilistic approach is assessed, by comparing our method with a SVM classifier, based on the same parameters. To specify correctly this classifier, we test several kernels (Gaussian, polynomial, sigmoid), with a variable selection for each one. The best results are obtained with the Gaussian kernel.

The following classical procedure is used:

- each 2D vector (frame) is classified with the trained SVM (naive approach),
- each second (100 2D vectors) is labelled by majority vote on the frame labels.

The overall error rate is **22.5 %**. The confusion matrix is presented in tab. VI.

### TABLE V

**Confusion matrix - 2 Classes - Independent Weibull univariate models.**

<table>
<thead>
<tr>
<th></th>
<th>Monophony</th>
<th>Polyphony</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monophony</td>
<td>85.3 %</td>
<td>14.7 %</td>
</tr>
<tr>
<td>Polyphony</td>
<td>16.2 %</td>
<td>83.8 %</td>
</tr>
</tbody>
</table>

One interesting result is the huge number of vectors selected as support: in each experiment, for each kernel, more than 7000 (over 12500) vectors are selected. More
precisely, all the training vectors from the monophonic class are selected, which explains why the classifiers act randomly on the monophonic class. This proves the complexity of the classes we consider in this problem.

In order to consider a 1-second classification, we add a postprocessing with a majority vote. However, we still obtain poor results. This is probably due to the fact that the “frame level” classifier has so bad performances.

3) An improvement – Subclass approach: In this experiment, five models are considered, one for each subclass, in the maximum likelihood step. The merging process is simple: a single instrument or a single singer detection is a monophonic decision, while the three other detections are polyphonic decisions.

The confusion matrix is presented in tab. VII, the overall error rate is 6.3 %.

### TABLE VII

<table>
<thead>
<tr>
<th></th>
<th>Monophony</th>
<th>Polyphony</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monophony</td>
<td>88.7 %</td>
<td>11.3 %</td>
</tr>
<tr>
<td>Polyphony</td>
<td>4.8 %</td>
<td>95.2 %</td>
</tr>
</tbody>
</table>

We note that considering five models in the classification step and merging the five subclasses into two groups improve notably the results. There are a bit more errors in the classification of the polyphonic class, but there are much less errors in the monophonic class, leading to a fall of the global error rate of more than 2 %. A close look at the results for each sub-class shows us that:

- Both monophonic subclasses are better classified, with a very significant improvement for the “single singer” subclass.
- A small number of errors are made in all polyphonic subclasses.

We therefore conclude that the subclass approach mostly allow us to better model the “single singer” subclass, with only little degradation for the modeling of each polyphonic subclass.

### IV. CONCLUSION AND PERSPECTIVES

In this article, a new method to distinguish between mono- and polyphonies was presented. This method is based on an observation vector composed of the short term mean and variance of a confidence indicator on the pitch. These mean and variance are computed over 50 ms and modeled by Weibull bivariate distributions. The classification is made by computing the joint likelihood of a sequence of 100 vectors over 1 second. Our method gave very good results: the global error rate is 6.3 %.

We assessed each step of our method by comparing it with classical methods (baseline, Gaussian distributions, Weibull univariate distributions and SVM).

The simplicity of this method implies that some improvement may rapidly be obtained: a simple post post-processing may lift some aberrant decisions. The most complex problem of short notes must be more explored as the problem of the non stable sounds. Our main interest will be to merge this approach with our study [14] on singing voice detection, in which it will be used as a pre-processing.

### APPENDIX

#### Parameter estimation of Weibull bivariate distribution

We estimate the five variables of the Weibull bivariate distribution by the moment method; the moments are given by Lu and Bhattacharyya in [13]. If we call $(X_1, X_2)$ a random vector, and $\Gamma$ the Gamma function:

$$E[X_{i}] = \theta_{i} \Gamma\left(\frac{1}{\beta_{i}} + 1\right), \quad i = 1, 2$$

$$\text{Var}(X_i) = \theta_{i}^{2} \left( \Gamma\left(\frac{2}{\beta_{i}} + 1\right) - \Gamma^{2}\left(\frac{1}{\beta_{i}} + 1\right) \right)$$

$$\text{Cov}(X_1, X_2) = \theta_{1} \theta_{2} \frac{\delta_{1} \delta_{2}}{\text{Var}(X_1)} \left[ \Gamma\left(\frac{\delta_{1}}{\beta_{1}} + 1\right) \Gamma\left(\frac{\delta_{2}}{\beta_{2}} + 1\right) \Gamma\left(\frac{1}{\beta_{1}} + \frac{1}{\beta_{2}} + 1\right) \right]$$

$$- \Gamma\left(\frac{1}{\beta_{1}} + 1\right) \Gamma\left(\frac{1}{\beta_{2}} + 1\right)$$

$\theta_{1}, \theta_{2}, \beta_{1} \text{ and } \beta_{2}$ are the parameters of the marginal Weibull distributions. Since this distribution has been studied for more than 40 years, several methods have been given to estimate the shape and scale parameters, which are available on the web, and summarized in [15].

With the moment method, the method is the following one: first, compute $\frac{\Gamma\left(\frac{\delta_{1}}{\beta_{1}} + 1\right) \Gamma\left(\frac{\delta_{2}}{\beta_{2}} + 1\right)}{\text{Var}(X_1)}$. Relying on tables for the values of $\Gamma\left(\frac{1}{\beta_{1}} + 1\right)$ and $\Gamma\left(\frac{2}{\beta_{2}} + 1\right) - \Gamma^{2}\left(\frac{1}{\beta_{1}} + 1\right)$, we get $\delta_{1}$. Finally, we get $\theta_{1}$ by inverting equation 4: $\theta_{1} = \frac{\Gamma\left(\frac{1}{\beta_{1}} + 1\right)}{\text{Cov}(X_1, X_2)}$. Variables $\theta_{1}, \theta_{2}, \beta_{1} \text{ and } \beta_{2}$ being easily deduced from equations 4 and 5, we therefore assume in the following that they are known.

To estimate variable $\delta$ from (6), we will show that:

- this equation can be written as $f(\delta) = C$, with $C$ a constant depending from $\theta_{1}, \theta_{2}, \beta_{1}$ and $\beta_{2}$, the function $f$ is a strictly decreasing function of $\delta$, its derivative being strictly negative for all triplet $(\beta_{1}, \beta_{2}, \delta) \in \mathbb{R}^{+} \times \mathbb{R}^{+} \times [0, 1]$.

So finding the unique zero of $f(\delta) - C$ (by dichotomy in our case) will give us $\delta$.

#### A. Transformation of equation 6

We first show that $(6) \Leftrightarrow \delta B\left(\frac{\delta_{1}}{\beta_{1}}, \frac{\delta_{2}}{\beta_{2}}\right) = C$, with $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ the Beta function. We have:

$$\frac{\text{Cov}(X_1, X_2)}{\theta_{1} \theta_{2}} = \frac{\Gamma\left(\frac{\delta_{1}}{\beta_{1}} + 1\right) \Gamma\left(\frac{\delta_{2}}{\beta_{2}} + 1\right) \Gamma\left(\frac{1}{\beta_{1}} + \frac{1}{\beta_{2}} + 1\right)}{\Gamma\left(\frac{\delta_{1}}{\beta_{1}} + \frac{\delta_{2}}{\beta_{2}} + 1\right) \Gamma\left(\frac{1}{\beta_{1}} + 1\right) \Gamma\left(\frac{1}{\beta_{2}} + 1\right)}$$

$$- \Gamma\left(\frac{1}{\beta_{1}} + 1\right) \Gamma\left(\frac{1}{\beta_{2}} + 1\right)$$

(7)
We put \( \Gamma \left( \frac{1}{\beta_1} + 1 \right) \Gamma \left( \frac{1}{\beta_2} + 1 \right) = C_1 \); it does not depend on \( \delta \).

Using the following property of the Gamma function: 
\[
\Gamma(a + 1) = a\Gamma(a),
\]
equation (7) becomes:
\[
\frac{\text{Cov}(X_1, X_2)}{\theta_1\theta_2} + C_1
= \frac{\delta}{\beta_1} \frac{\delta}{\beta_2} \frac{\Gamma \left( \frac{\delta}{\beta_1} \right) \Gamma \left( \frac{\delta}{\beta_2} \right)}{\Gamma \left( \frac{1}{\beta_1} + \frac{1}{\beta_2} + 1 \right)}
= \frac{\beta_1 + \beta_2}{\beta_1 \beta_2} \left( \frac{\Gamma \left( \frac{1}{\beta_1} + \frac{1}{\beta_2} + 1 \right)}{\Gamma \left( \frac{\delta}{\beta_1} + \frac{\delta}{\beta_2} \right)} \right)
= \frac{\Gamma \left( \frac{1}{\beta_1} + \frac{1}{\beta_2} + 1 \right)}{\beta_1 + \beta_2} \left( \frac{\Gamma \left( \frac{\delta}{\beta_1} + \frac{\delta}{\beta_2} \right)}{\Gamma \left( \frac{1}{\beta_1} + \frac{1}{\beta_2} + 1 \right)} \right)
\equiv \frac{C_1}{C_2}
\]
\( \delta B \left( \frac{\delta}{\beta_1}, \frac{\delta}{\beta_2} \right) = \frac{\text{Cov}(X_1, X_2)}{\theta_1\theta_2} + C_1
\]
where \( C \) is independent of \( \delta \). We note \( f(\delta) = \delta B \left( \frac{\delta}{\beta_1}, \frac{\delta}{\beta_2} \right) \).

**B. Derivative of the function \( f \) and its sign**

The derivative \( f' \) of \( f \) is:
\[
f'(\delta) = B \left( \frac{\delta}{\beta_1}, \frac{\delta}{\beta_2} \right)
\]
\[
\left[ 1 + \frac{1}{\beta_1} \left( \psi_0 \left( \frac{\delta}{\beta_1} \right) - \psi_0 \left( \frac{\delta}{\beta_1} + \frac{\delta}{\beta_2} \right) \right)
+ \frac{1}{\beta_2} \left( \psi_0 \left( \frac{\delta}{\beta_2} \right) - \psi_0 \left( \frac{\delta}{\beta_1} + \frac{\delta}{\beta_2} \right) \right) \right]
\]
with \( \psi_0(x) = \frac{d}{dx} \psi(x) \) the digamma function.

We know that \( B \left( \frac{\delta}{\beta_1}, \frac{\delta}{\beta_2} \right) \) is independent of \( \delta \).

We integrate (13) between \( a \) and \( a+b \) (with \( a+b > a \)):
\[
\psi_1(x) > \frac{1}{x^2}, \quad x > 0
\]
\[
\psi_1(x) > \frac{1}{x^2} \Rightarrow \int_a^{a+b} \psi_1(x) dx > \int_a^{a+b} \frac{1}{x^2} dx
\]
\[
\psi_0(a+b) - \psi_0(a) > -\frac{1}{a+b} + \frac{1}{a}
\]
\[
a(\psi_0(a+b) - \psi_0(a)) > \frac{b}{a+b}
\]

Symmetrically, \( b(\psi_0(b) - \psi_0(a+b)) \leq -\frac{a}{a+b} \), which finally leads to:
\[
\left[ a(\psi_0(a) - \psi_0(a+b)) + b(\psi_0(b) - \psi_0(a+b)) \right] < -1
\]

**References**


