University of São Paulo
“Luiz de Queiroz” College of Agriculture

**Statistical modelling of data from performance of
broiler chickens**

Reginaldo Francisco Hilário

Thesis presented to obtain the degree of Doctor in Science.
Area: Statistics and Agricultural Experimentation

Piracicaba
2018

Reginaldo Francisco Hilário
Degree in Mathematics

Statistical modelling of data from performance of broiler chickens
versão revisada de acordo com a resolução CoPGr 6018 de 2011.

Advisor:
Prof^ª Dr^ª **CLARICE GARCIA BORGES DEMÉTRIO**

Thesis presented to obtain the degree of Doctor in Science.
Area: Statistics and Agricultural Experimentation

Piracicaba
2018

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP

Hilário, Reginaldo Francisco
Statistical modelling of data from performance of broiler chickens/
Reginaldo Francisco Hilário. – – versão revisada de acordo com a resolução
CoPGr 6018 de 2011. – – Piracicaba, 2018.
160 p.

Tese (Doutorado) – – USP / Escola Superior de Agricultura “Luiz de
Queiroz”.

1. Frango de corte 2. Poder de teste 3. Tamanho amostral 4. Modelos de
mistura . I. Título.

DEDICATION

*I dedicate this work in memory of
my parents and my brother.*

ACKNOWLEDGMENTS

I would like to thank my family, for the immense love for me, for the patience and all affection, without you my life would not make sense.

To my adviser, Prof. Dr. Clarice Garcia Borges Demétrio, for guidance, for her confidence in me, for her motivation, patience, dedication and shared wisdom, thank you so much.

To Prof. Dr. José Fernando Machado Menten, for the time dedicated, for the attention and clarifications.

To Professors Dr. Geert Molenberghes and Dr. Geert Verbeke, for their valuable guidance, enthusiasm and motivation, I am immensely grateful.

I am also very grateful to Martine Machiels, who helped to arrange a great stay for my family and school for my children in Belgium.

I would like to thank Prof. Dr. Silvio Sandoval Zocchi, for the contribution that helped me to enrich the work.

To Professor Dr. Cristian Marcelo Villegas Lobos, for the attention and good will to help me.

To the Professors of the Department of Exact Sciences at ESALQ/USP, who were present at this time of course, for their shared experiences that collaborated to build my knowledge.

To colleagues and employees of the Department of Exact Sciences at ESALQ/USP, for the friendship and companionship.

Special thanks to CNPq for the financial support in Brazil and CAPES for the financial support in Belgium, I am very grateful.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

EPIGRAPH

*“He who has no love has no knowledge of God,
because God is love.”*

1 John 4:8

*“Experience is not what happens to a man;
it is what a man does with what happens to him.”*

Aldous Huxley

CONTENTS

RESUMO	8
ABSTRACT	9
LIST OF FIGURES	10
LIST OF TABLES	12
1 INTRODUCTION	19
References	20
2 STATISTICAL TEST POWER ANALYSIS ON BROILER CHICKEN DATA . . .	21
Abstract	21
2.1 Introduction	21
2.2 Case-study	22
2.3 Modelling	27
2.3.1 Type I and Type II errors	27
2.3.2 Power of a Statistical test	28
2.3.3 Non-central χ^2 , F and t distributions	28
2.3.3.1 Non-central χ^2 distribution	28
2.3.3.2 Non-central F -distribution	30
2.3.3.3 Non-central t -distribution	32
2.3.4 Power of the F -Test	34
2.3.5 Mixed Models	38
2.3.6 Selection of models	39
2.3.6.1 Likelihood Ratio Test	39
2.3.6.2 Information criteria	40
2.3.7 Tests for the fixed effects	41
2.3.7.1 Approximate Wald Tests	41
2.3.7.2 Approximate t-Tests and F-Tests	41
2.3.8 Diagnostics	42
2.4 Results	43
2.5 Discussion	47
References	47
3 MIXTURE MODELS FOR THE ANALYSIS OF CHICKENS WEIGHT	53
Abstract	53
3.1 Introduction	53
3.2 Case-study	54
3.3 Modelling	54
3.3.1 Mixture Models	54
3.3.2 Mixtures of normal distributions	56
3.3.3 Methods of estimation	60

3.3.4 EM Algorithm	61
3.3.5 Mixture model for the sum of chicken weights: Cross-sectional case	68
3.3.5.1 Gender-specific mean and variance	68
3.3.5.2 Different means and common variance	69
3.3.6 Methods of estimation	70
3.3.7 Bayesian approach	70
3.3.8 Simulated case study using the classical approach	71
3.3.9 Simulated case study using the Bayesian approach	73
3.4 Results	74
3.4.1 Analysis of individual weights	74
3.4.1.1 Analysis of individual weights using the classical approach	74
3.4.1.2 Analysis of the individual weights using the Bayesian approach	78
3.4.2 Analysis of the sum of the weights of chickens	82
3.4.2.1 Simulated case study of the sum of chicken weights using the classical approach	82
3.4.2.2 Analysis of the sum of chicken weights of the real data using the classical approach	87
3.4.2.3 Simulated case study of the sum of chicken weights using the Bayesian Approach	89
3.4.2.4 Analysis of the sum of chicken weights of the real data using the Bayesian approach	91
3.5 Discussion	96
References	97
APPENDIX	101

RESUMO

Modelagem estatística de dados de desempenho de frangos de corte

Experimentos com frangos de corte são comuns atualmente, pois devido à grande demanda de mercado da carne de frango surgiu a necessidade de melhorar os fatores ligados à produção do frango de corte. Muitos estudos têm sido feitos para aprimorar as técnicas de manejo. Nesses estudos os métodos e técnicas estatísticas de análise são empregados. Em estudos com comparações entre tratamentos, não é incomum observar falta de efeito significativo mesmo quando existem evidências que apontam a significância dos efeitos. Para evitar tais eventualidades é fundamental realizar um bom planejamento antes da condução do experimento. Nesse contexto, foi feito um estudo do poder do teste F enfatizando as relações entre o poder do teste, tamanho da amostra, diferença média a ser detectada e variância para dados de pesos de frangos. Na análise de dados provenientes de experimentos com frangos de corte com ambos os sexos e que a unidade experimental é o boxe, geralmente os modelos utilizados não levam em conta a variabilidade entre os sexos das aves, isso afeta a precisão da inferência sobre a população de interesse. Foi proposto um modelo para o peso total por boxe que leva em conta a informação do sexo dos frangos.

Palavras-chave: Frango de corte; Poder do teste F ; Tamanho amostral; Modelos de mistura

ABSTRACT

Statistical modelling of data from performance of broiler chickens

Experiments with broiler chickens are common today, because due to the great market demand for chicken meat, the need to improve the factors related to the production of broiler chicken has arisen. Many studies have been done to improve handling techniques. In these studies statistical analysis methods and techniques are employed. In studies with comparisons between treatments, it is not uncommon to observe a lack of significant effect even when there is evidence to indicate the significance of the effects. In order to avoid such eventualities it is fundamental to carry out a good planning before conducting the experiment. In this context, a study of the power of the F test was made emphasizing the relationships between test power, sample size, mean difference to be detected and variance for chicken weights data. In the analysis of data from experiments with broilers with mixed sexes and that the experimental unit is the box, generally the models used do not take into account the variability between the sexes of the birds, this affects the precision of the inference on the population of interest . We propose a model for the total weight per box that takes into account the sex information of the broiler chickens.

Keywords: Broiler chickens; Power of the F test; Sample size; Mixture models

LIST OF FIGURES

Figure 2.1 - Graph of profiles over time of total weight in kilograms per box for each treatment	23
Figure 2.2 - Histogram for the weight in grams at 7 days for each treatment	24
Figure 2.3 - Histogram for the weight in kilograms at 42 days for each treatment	25
Figure 2.4 - Histogram of residuals for the model considering the weight at 42 days	26
Figure 2.5 - Graph of the distribution χ^2 with $\nu = 4$ degrees of freedom and some values for the parameter of non-centrality	29
Figure 2.6 - Graph of the central and non-central χ^2 distributions with $\nu = 4$ degrees of freedom, type II error rate (β), power of the test ($1 - \beta$) for a significance level $\alpha = 0,05$	30
Figure 2.7 - Graph of the central and non-central F -distribution with $\nu_1 = 6$ and $\nu_2 = 12$ degrees of freedom and some values for the non-centrality parameter λ	31
Figure 2.8 - Graph of the central and non-central F -distributions with $\nu_1 = 6$ e $\nu_2 = 12$ degrees of freedom, type II error rate (β), power of the test ($1 - \beta$) for a significance level $\alpha = 0,05$	31
Figure 2.9 - Graph of the central and non-central t distribution with $\nu = 5$ degrees of freedom and some values for the non-centrality parameter δ	33
Figure 2.10 - Graph of the central and non-central t distribution with $\nu = 5$ degrees of freedom, type II error rate (β), test power ($1 - \beta$) for a significance level $\alpha = 0,05$	34
Figure 2.11 - Power of test as a function of the effect size	36
Figure 2.12 - Power of test as a function of the significance level	36
Figure 2.13 - Power of test as a function of the sample size	36
Figure 2.14 - Sample size as a function of the effect size	37
Figure 2.15 - Sample size as a function of the test power	37
Figure 2.16 - Sample size as a function of the significance level	38
Figure 2.17 - Power of the F test as a function of the mean difference for the experiment with chicken weight data at 42 days with different variances	44
Figure 2.18 - Power of the F test as a function of sample size (number of replicates per treatment) for the experiment with chicken weight data at 42 days with different variances	45
Figure 2.19 - Sample size (number of replicates per treatment) as a function of the mean difference for the experiment with chicken weight data at 42 days with different variances	45
Figure 2.20 - Power of the F test as a function of σ for the experiment with chicken weights at 42 days considering $\Delta = 50g$, 5 replicates and $\alpha = 0.05$	46

Figure 2.21 - Sample size as a function of σ for the experiment with chicken weights
 at 42 days considering $\Delta = 50\text{g}$, $(1 - \beta) \approx 0.8$ and $\alpha = 0.05$ 46

LIST OF TABLES

Table 2.1 -	Number of individuals per box at 7 days and 42 days for experiment with chickens	22
Table 2.2 -	Table of analysis of variances with subsamples considering the weight of chickens at 42 days of age	26
Table 2.3 -	Estimates of the components of variance and the ratio between $\hat{\sigma}_\epsilon^2$ and $\hat{\sigma}_e^2$ for the chicken weight data at 42 days	27
Table 2.4 -	Possible scenarios for a hypothesis test	27
Table 2.5 -	Number of replicates (r) required to detect the mean difference in grams (Δ) with probability 0.8, at each time of the experimental period for live weight data of chickens	43
Table 3.1 -	Estimates of the parameters for the models with a normal distribution and the mixture of two normal distributions with likelihood ratio test considering the weight of chickens at 42 days of age	74
Table 3.2 -	Bayesian estimates of the mixture model parameters with Gaussian components of individual weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$). Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and credibility interval with 95% probability for each model parameter	79
Table 3.3 -	Bayesian estimates of the mixture model parameters with Gaussian components of individual weights of chickens with heterogeneous variances (σ_1 and σ_2). Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and credibility interval with 95% probability for each model parameter	80
Table 3.4 -	Estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values	83
Table 3.5 -	Estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values	83
Table 3.6 -	Estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values	84

Table 3.7 -	Estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ were varied and kept the initial values of the other parameters fixed in the true values	84
Table 3.8 -	Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values	85
Table 3.9 -	Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values	85
Table 3.10 -	Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values	86
Table 3.11 -	Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_1 were varied and kept the initial values of the other parameters fixed in the true values	86
Table 3.12 -	Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_2 were varied and kept the initial values of the other parameters fixed in the true values	87
Table 3.13 -	Estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the weights data of chickens at 42 days by treatment. The parameters estimates by the Nelder-Mead algorithm, as well as the respective standard errors (SE) are presented	88
Table 3.14 -	Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the weights data of chickens at 42 days by treatment. The parameters estimates by the Nelder-Mead algorithm, as well as the respective standard errors (SE) are presented	88

Table 3.15 - Bayesian estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and quantiles of the posterior distribution for each model parameter	89
Table 3.16 - Bayesian estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and quantiles of the posterior distribution for each model parameter	90
Table 3.17 - Bayesian estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) with informative priors for the simulated data considering 5 boxes and 46 individuals per box. Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and quantiles of the posterior distribution for each model parameter	90
Table 3.18 - Bayesian estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) with informative priors for the simulated data considering 5 boxes and 46 individuals per box. Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and quantiles of the posterior distribution for each model parameter	90
Table 3.19 - Bayesian estimates of the model parameters of sum of the chickens weights with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$). Also shown are standard deviation (SD), Monte Carlo standard error (MCSE) and the credibility interval of 95% for each parameter of the model	92
Table 3.20 - Bayesian estimates of the model parameters of sum of the chickens weights with heterogeneous variances (σ_1 and σ_2). Also shown are standard deviation (SD), Monte Carlo standard error (MCSE) and the credibility interval of 95% for each parameter of the model	93
Table 3.21 - Bayesian estimates of the model parameters of sum of the chickens weights with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) with informative priors. Also shown are standard deviation (SD), Monte Carlo standard error (MCSE) and the credibility interval of 95% for each parameter of the model	94

Table 3.22 - Bayesian estimates of the model parameters of sum of the chickens weights with heterogeneous variances (σ_1 and σ_2) with informative priors. Also shown are standard deviation (SD), Monte Carlo standard error (MCSE) and the credibility interval of 95% for each parameter of the model	95
Table 3.23 - Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values	107
Table 3.24 - Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values	107
Table 3.25 - Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values	108
Table 3.26 - Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ were varied and kept the initial values of the other parameters fixed in the true values	108
Table 3.27 - Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values	109
Table 3.28 - Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values	109

- Table 3.29 - Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values 110
- Table 3.30 - Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_1 were varied and kept the initial values of the other parameters fixed in the true values 110
- Table 3.31 - Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_2 were varied and kept the initial values of the other parameters fixed in the true values 111
- Table 3.32 - Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values 111
- Table 3.33 - Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values 112
- Table 3.34 - Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values 112
- Table 3.35 - Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ were varied and kept the initial values of the other parameters fixed in the true values 113

- Table 3.36 - Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values113
- Table 3.37 - Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values114
- Table 3.38 - Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values114
- Table 3.39 - Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_1 were varied and kept the initial values of the other parameters fixed in the true values115
- Table 3.40 - Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_2 were varied and kept the initial values of the other parameters fixed in the true values115

1 INTRODUCTION

The world production of chicken meat reached about 88.72 million tons in 2016. Brazil remained as the largest exporter and second largest producer of chicken meat with 12.90 million tons behind only the United States, playing a leading role in the global poultry industry scenario as projections of the United States Department of Agriculture - USDA - ABPA (Associação Brasileira de Proteína Animal).

Advances in genetic improvement, in management conditions, sanitary control and nutrition favor the growing increase in world poultry production. Among these existing aspects of poultry production, nutrition plays an important role as it represents about 70% of production costs (RIZZO, 2008). In this sense, there is great interest among researchers in exploring scientifically this aspect in order to reduce costs and increase productivity. Many studies have been performed, but those that are distinguished as promising are the ones that have features often overlooked. The question that becomes the main obstacle to research is to know what is really important, or essential, to take into account. Taking all views is impractical, so it is up to the researcher to select the necessary and feasible aspects.

In poultry science, a common practice is to compare new treatments with control or to make comparisons between treatments in planned experiments (DEMÉTRIO et al., 2013). Often the results of the experiments do not show what the researcher expected, effects of non-significant treatments may occur even though there is considerable evidence pointing to the contrary in similar studies. In order to minimize this type of eventuality, it is essential to pay special attention to the planning of the experiment. With an adequate statistical planning it is possible to extract the maximum of useful information that leads the answer of the research question. In this sense, it is necessary to understand the factors that directly influence the results. Among them we can mention sample size, variability between experimental units and effect size. In addition, there are uncontrollable variables (experimental error), which tend to mask the effects of the treatments. From the previous knowledge of these aspects about an experiment that one wishes to conduct, we are able to elaborate a good planning and achieve more accurate analysis and reach more reliable inferences.

In experiments with broiler chickens we can have batches separated by sex or mixed sexes. There are different handling specifications for each batch type. Generally mixed sexes experiments are performed according to the recommended specifications, but statistical analyzes are done as if there was no mixing of the sexes, since the models used do not take into account the variability between the sexes.

In chapter 2 of this work we did a study of the power of the F test emphasizing the relationships between test power, sample size, mean difference to be detected and variance for chicken weight data. In chapter 3 we propose a model that takes into account the sex information of the birds when the observation is the total weight per box.

References

ASSOCIAÇÃO BRASILEIRA DE PROTEÍNA ANIMAL - ABPA. Disponível em: <<http://www.abpa-br.org>>. Acesso em: 08 jun. 2018.

DEMÉTRIO, C.G.B.; MENTEN, J.F.M.; LEANDRO, R.A.; BRIEN, C. Experimental power considerations - justifying replication for animal care and use committees. **Poultry Science**, Savoy, v. 92, p. 2490-2497, 2013.

RIZZO, P. **Misturas de extratos vegetais como alternativas ao uso de antibióticos melhoradores do desempenho nas dietas de frangos de corte**. 2008. 69 p. Dissertação (Mestrado em Ciências Animal e Pastagens) - Escola Superior de Agricultura “Luiz de Queiroz” - Universidade de São Paulo, Piracicaba, 2008.

2 STATISTICAL TEST POWER ANALYSIS ON BROILER CHICKEN DATA

Abstract

In experimental designs, one of the aims is to study statistical differences between treatments. However, it is not uncommon to observe the lack of significant differences even when many evidences point to the existence of differences. The good planning of the experiment has a determinant role in the inference about the parameters involved in the study to obtain reliable inferences. For this, it is fundamental to have some prior knowledge about the subject to be studied. Such knowledge can be obtained from a pilot study, or from some systematic investigation of similar studies already performed. In this context, prior knowledge of the power of the test, sample size and effect size related to the study in question are indispensable in every planning.

Keywords: Test power; Sample size; Effect size; Experimental design

2.1 Introduction

Poultry production in Brazil has international recognition and provides the country with excellent positions among the world's largest producers.

The technological advances in genetics, management and ambience have provided the great development of the country in the sector, in this way Brazil has been intensively increasing the production of chickens. Due to increased production, alternatives to reduce costs have been explored.

Feeding for poultry represents about two thirds of the cost of producing broilers (RIZZO, 2008). Thus, many efforts have been made to improve the efficiency of poultry diets. In this context, it is common to use experiments with the objective of comparing different diets. It is not difficult to find studies in which the results point to the non-significance of the effects even the researcher knowing evidence in favor of their significance. In order to minimize such eventualities, it is essential to carry out good planning before conducting the experiment. With proper planning, more reliable inferences about the study population will be obtained. For this, it is necessary to understand the factors that directly influence the results, among them we can mention the sample size, the size of the desired effect, the type I and type II error rates and the natural variability present in the type of data to be studied.

In this chapter a study of the power of the F test was made emphasizing the relationships between test power, sample size, mean difference to be detected and variance for chicken weight data.

2.2 Case-study

In order to assess the effect of physical form of pre-starter diet on performance of broiler chickens born from eggs hatched from Ross breeders of different ages, Traldi (2009) conducted in the experimental aviary of Department of Animal Science - Sector Non Ruminants of the College of Agriculture “Luiz de Queiroz”, a completely randomized design with six treatments (factorial 2×3) and five replicates.

In this experiment, there were 30 plots with 46 birds (23 male and 23 females from Ross breeders) for each of them. The treatments were a combination of three physical forms of diets for two ages of breeders.

For the pre-initial phase with unique formula feed and also to each phase (initial, growth and final) the nutritional recommendations of Rostagno (2005) were observed. The diets were produced at the feed mill in the Department of Animal Science of ESALQ. Mortality and culling were observed throughout the trial period.

The response variable live weight of birds in grams was observed in two different ways, which are described below. In one of the ways the value of the individual weight of each bird was observed in two occasions, at 7 and 42 days. Once the experiment was conducted with 30 plots and 46 birds within each, in a balanced case we would have 1380 observations on each of the two experimental days (at 7 and at 42 days). It is worth noting that it is not common to have the individual observations for an experiment of this size with 1380 birds. The other way consisted of the total weight in each box, that was observed at 21 and 35 days. Therefore, in this case, there were 30 observations recorded, which are related to the number of boxes in the experiment (30 boxes). Generally for this type of experiment we have the average of the weights per box and not the information of the individual weights of the birds.

The Table (2.1) shows the unbalance in the number of individuals per box due to culled and/or mortality at 7 and 42 days respectively. Note that for this experiment there was little unbalance of individuals within the boxes.

Table 2.1 – Number of individuals per box at 7 days and 42 days for experiment with chickens

Treatment	Repetitions (7 days)						Repetitions (42 days)					
	1	2	3	4	5	Total	1	2	3	4	5	Total
T1	46	46	46	46	46	230	45	46	46	45	45	227
T2	46	46	46	46	46	230	45	44	45	46	46	226
T3	45	46	46	46	46	229	45	43	44	45	43	220
T4	46	46	45	46	46	229	45	45	45	46	46	227
T5	46	46	45	46	46	229	44	46	45	45	45	225
T6	46	46	46	46	45	229	45	43	45	43	42	218

An important information is that in each box, at the beginning of the experiment, 23 female birds and 23 male birds were placed, without identifying them during the experiment, i.e. the observations were recorded in the case of the individual measurements, without knowing if the individual was male or female. Since during the experiment there was mortality of the birds, the number of males and/or females can be considered as being a random variable whose value changes during the conduction of the experiment. Figure (2.1) shows the graph of profiles over time of the total weight in kilograms per box for each treatment.

Figure 2.1 – Graph of profiles over time of total weight in kilograms per box for each treatment

The histograms with the empirical density of the weight in grams at 7 days for each treatment are shown in Figure (2.2). One observes in some of the graphs a bimodal structure and in others a negative skewness. In this age of birds, weights are similar between

males and females, and that is why the apparent bimodal structure of the data is not so obvious, but it is still possible to visualize a slight bimodality.

Figure 2.2 – Histogram for the weight in grams at 7 days for each treatment

The Figure (2.3) shows the histogram with the empirical density for the weight at 42 days for each treatment. It is possible to observe the bimodality present in the histograms for each of the treatments, in that age of the birds, the weights between males and females distanced themselves and the distinction between the subpopulations of males and females presented in the histogram is easily discernible.

Figure 2.3 – Histogram for the weight in kilograms at 42 days for each treatment

Considering the cross-sectional case, where the analysis is made for each time point independent of the others, a commonly used model for this case, taking into account the individual measures is as follows

$$y_{ijk} = \mu + \tau_i + e_{ij} + \varepsilon_{ijk} \quad (2.1)$$

where, $i = 1, \dots, t$; $j = 1, \dots, r$; $k = 1, \dots, s$; μ is the general mean inherent for all observations; τ_i is the treatment effect; e_{ij} is the experimental error (error between) and ε_{ijk} is the sample error (error within).

The common assumptions for this model are as follows, the treatment effect τ_i is fixed, i.e., $E(\tau_i) = \tau_i$, the experimental error e_{ij} is random following the normal

distribution with mean equal to zero and variance equal to σ_e^2 independent and identically distributed, and the sample error ε_{ijk} is random following the normal distribution with mean equal to zero and variance equal to σ_e^2 independent and identically distributed.

Considering the analysis using the model (2.1) we have the Table (2.2) of analysis of variance with subsamples for the data of chicken weight at 42 days. Figure (2.4) shows the histograms of the residuals for the experimental error and sample error of the model (2.1). The histograms of the residuals (2.4) suggest a nonnormal distribution of the errors and present bimodality.

Table 2.2 – Table of analysis of variances with subsamples considering the weight of chickens at 42 days of age

Source of variation	D.f.	SS	MS
Treatment	5	3308225	661645.1
Residuals	24	3325311	138554.6
(Plots)	29	—	—
Residuals (Within)	1313	148283198	112934.7
Total	1342

Figure 2.4 – Histogram of residuals for the model considering the weight at 42 days; (a) Residuals for the experimental error (between plots); (b) Residuals for the sample error (within)

Estimates for the variance components using the method of moments (MM), the method of moments with the harmonic mean (MMH) (RAMALHO et al., 2005), restricted maximum likelihood method (REML) and the ratio between $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_e^2$ are found in Table (2.3). Note that the variation between individuals within the plot is approximately 200 times the variance of the error. In addition, $\hat{\sigma}_\varepsilon^2$ represents approximately 81% of the contribution to the residual mean square, which is very likely due to differences between males and females, in addition to natural variability.

Table 2.3 – Estimates of the components of variance and the ratio between $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_e^2$ for the chicken weight data at 42 days

	MM	MMH	REML
$\hat{\sigma}_e^2$	572.34	543.96	588.58
$\hat{\sigma}_\varepsilon^2$	112934.7	113045.1	112920.88
$\hat{\sigma}_\varepsilon^2/\hat{\sigma}_e^2$	197.32	207.82	191.85

Note that the model presented above does not take into account the information about the sex of the animals, this omission entails an increase in the residual and less precision in the inference about the population of interest.

2.3 Modelling

According to Aaron and Hays (2004), “an understanding of some basic statistical concepts and the essence of classical hypothesis testing is a necessary precursor to a discussion of statistical power”. In this section, we review some general concepts and definitions on the issue of statistical power, as well as notions about the statistical models that will be used in this work.

2.3.1 Type I and Type II errors

When testing a hypothesis, the researcher aims to decide which of two complementary hypothesis is true, taking into account a sample of the population (CASELLA; BERGER, 2002). These two complementary hypotheses are called null hypothesis and alternative hypothesis, denoted respectively by H_0 and H_1 . When doing a decision on a statistical test four scenarios can occur as shown in Table 2.4. If the null hypothesis H_0 is accepted if true or rejected when it is false, no error has been committed. However, there are two possibilities of error, a type I error, which occurs when a true null hypothesis is rejected and Type II error, when a false null hypothesis is accepted.

Table 2.4 – Possible scenarios for a hypothesis test

Decision	H_0	
	True	False
Accept H_0	Correct decision	Type II Error
Reject H_0	Type I Error	Correct decision

The probability of accepting H_0 given that H_0 is true is $(1 - \alpha)$ and the probability of type I error is $P(\text{Reject } H_0 | H_0 \text{ is true}) = \alpha$. The probability of rejecting H_0 given that H_0 is false is $(1 - \beta)$ and the probability of type II error is $P(\text{Accept } H_0 | H_0 \text{ is false}) = \beta$.

When determining the significance level α of the test, the researcher is only controlling the probability of type I error, not the type II error. When the sample size is

fixed, it is not feasible to make both types of error arbitrarily small. In practice, when the researcher set the type I error rate at a very low value, this will result in high values for the type II error rate, so it is necessary to maintain a balance between these two error rates. It is important to note that when researchers observe in their experiments absence of treatment effect should incorporate in their interpretations the possibility of type II error (COUSENS; MARSHALL, 1987; COHEN, 1977).

2.3.2 Power of a Statistical test

The sensitivity or power of the test is the probability $(1 - \beta)$ to reject the null hypothesis H_0 when it is false, in which β is the probability of type II error. According to Berndtson (1991), the power of a statistical test is the probability that a treatment effect does not go unnoticed, if there is an effect. In this context, power is the ability to detect real differences, if they exist, in an experiment with significance level α stipulated by the researcher.

Generally the power depends on the magnitude of the difference to be detected, the significance level α and the size of the experimental error. When the experimental error is reduced with removing irrelevant sources of variability, the probability of detecting minor differences increases and the power of the test also. The increase in sample size decreases the experimental error, and therefore increases the power of the test. A very common question among researchers who want to define a protocol for an experiment is how many replicates per treatment are necessary, a discussion on this subject can be found in Demétrio et al. (2013).

2.3.3 Non-central χ^2 , F and t distributions

For a better understanding of the power of the test, a notion is needed about the parameter of non-centrality admitted in the distribution referring to the statistical test used. The following are presented probability non-central distributions χ^2 , F and t with the respective definitions of non-centrality parameters. Such definitions can be found in Appendix IV of Scheffé (1959).

2.3.3.1 Non-central χ^2 distribution

If a random variable X has normal distribution with mean ξ and variance σ^2 , we denote $X \sim N(\xi, \sigma^2)$.

Definition 2.1 *If $X_1, X_2, \dots, X_\nu$ are independently distributed and $U = \sum_{i=1}^{\nu} X_i^2$ has distribution χ^2 with ν degrees of freedom and non-centrality parameter $\delta = \sum_{i=1}^{\nu} \xi_i^2$.*

The probability density function of a random variable with distribution χ^2 with ν degrees of freedom and non-centrality parameter $\lambda = \delta^2$ is given by

$$f(x; \nu, \lambda) = e^{-\frac{\lambda}{2}} \sum_{r=0}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^r}{r!} f(x; \nu + 2r) = \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{1}{2}(x+\lambda)} \sum_{r=0}^{\infty} \frac{(\lambda x)^r \Gamma\left(\frac{1}{2} + r\right)}{(2r)! \Gamma\left(\frac{\nu}{2} + r\right)} \quad (2.2)$$

where $x > 0$, $\nu > 0$ and $f(x; \nu + 2r)$ is the density function of ordinary χ^2 with $\nu + 2r$ degrees of freedom (JOHNSON et al., 1995).

An ordinary or central χ^2 distribution is said to be a special case of the non-central distribution when the non-centrality parameter is zero, $\delta = 0$ (SCHEFFÉ, 1959). By convention it will be called the χ^2 distribution without mention of the non-centrality parameter as being the ordinary or central χ^2 . In the literature, some authors use the non-centrality parameter as $\lambda = \delta^2$, others use $\lambda = \frac{1}{2}\delta^2$. To denote that a random variable X follows a χ^2 distribution with ν degrees of freedom and non-centrality parameter λ , we usually use the notation $X \sim \chi^2_{\nu}(\lambda)$. According to Kendall and Stuart (1961), the distribution (2.2) was introduced by Fisher (1928) and studied further by Wishart (1932) and Patnaik (1949).

The Figure 2.5 shows the graphical representation of the probabilistic density function of χ^2 with $\nu = 4$ degrees of freedom and some values for the non-centrality parameter. A hypothetical situation of a hypothesis test is represented in Figure 2.6, where the density curve of the central χ^2 is shown under the null hypothesis with $\nu = 4$ degrees of freedom and also the density of the non-central χ^2 under the alternative hypothesis. Additionally, this figure illustrates the power of the hypothesis test and the type II error rate.

Figure 2.5 – Graph of the distribution χ^2 with $\nu = 4$ degrees of freedom and some values for the parameter of non-centrality

Figure 2.6 – Graph of the central and non-central χ^2 distributions with $\nu = 4$ degrees of freedom, type II error rate (β), power of the test ($1 - \beta$) for a significance level $\alpha = 0,05$

2.3.3.2 Non-central F -distribution

The non-central F distribution was first studied by Fisher (1928), in a special context by Wishart (1932) and later by Tang (1938) and Patnaik (1949) (KENDALL; STUART, 1961).

Definition 2.2 If U_1 and U_2 are independent random variables and $U_1 \sim \chi_{\nu_1}^2(\lambda)$, $U_2 \sim \chi_{\nu_2}^2(\lambda)$, the distribution of the ratio

$$F = \frac{U_1/\nu_1}{U_2/\nu_2}$$

is called non-central F distribution with ν_1 and ν_2 degrees of freedom and non-centrality parameter λ .

The probability density function of the non-central distribution F can be written as

$$f(F; \nu_1, \nu_2, \lambda) = \sum_{r=0}^{\infty} \frac{e^{-\lambda/2} (\lambda/2)^r}{B\left(\frac{\nu_2}{2}, \frac{\nu_1}{2} + r\right) r!} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2} + r} \left(\frac{\nu_2}{\nu_2 + \nu_1 F}\right)^{\frac{\nu_1 + \nu_2}{2} + r} (F)^{\frac{\nu_1}{2} - 1 + r} \quad (2.3)$$

where $F \geq 0$, the number of degrees of freedom of the numerator and denominator are positive and the parameter of non-centrality λ is non-negative. The term $B(a, b)$ corresponds to the beta function, where

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

The central F distribution is a special case of the non-central F with non-centrality parameter equal to zero, $\lambda = 0$. We will use the notation $F_{\nu_1, \nu_2}(\lambda)$ for the non-central F distribution with ν_1 and ν_2 degrees of freedom for the numerator and denominator, respectively, with non-centrality parameter λ . The Figure 2.7 shows the graphical representation of the probability density function of the distribution F with $\nu_1 = 6$ and $\nu_2 = 12$ degrees of freedom and some values for the non-centrality parameter. A hypothetical situation of a hypothesis test is represented in Figure 2.8, which shows the density curve of the central F distribution under the null hypothesis with $\nu_1 = 6$ and $\nu_2 = 12$ degrees of freedom, and also the density of the non-central F distribution under the alternative hypothesis. Additionally, the power of the hypothesis test and the type II error rate for this situation are illustrated.

Figure 2.7 – Graph of the central and non-central F -distribution with $\nu_1 = 6$ and $\nu_2 = 12$ degrees of freedom and some values for the non-centrality parameter λ

Figure 2.8 – Graph of the central and non-central F -distributions with $\nu_1 = 6$ e $\nu_2 = 12$ degrees of freedom, type II error rate (β), power of the test ($1 - \beta$) for a significance level $\alpha = 0,05$

2.3.3.3 Non-central t -distribution

Considering the probability density function of the non-central F distribution, we can obtain the non-central t distribution. For this, in the expression (2.3), making $\nu_1 = 1$ we have the non-central t^2 distribution with non-centrality parameter $\delta^2 = \lambda$ and ν_2 degrees of freedom, and, by applying a transformation from t^2 to t , we obtain the non-central t distribution. The notation $t_{\nu,\delta}$ to designate the non-central t distribution with non-centrality parameter δ and ν degrees of freedom will be used.

Definition 2.3 *If X and U are independent random variables and $X \sim N(\delta, 1)$, $U \sim \chi_\nu^2$, the distribution of the ratio*

$$T = \frac{X}{\sqrt{U/\nu}}$$

is called of non-central t distribution with ν degrees of freedom and non-centrality parameter δ .

The probability density function of the non-central t distribution with ν degrees of freedom and non-centrality parameter δ can be expressed by

$$f(T; \nu, \delta) = \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \sum_{r=0}^{\infty} \frac{(T\delta)^r}{r!\nu^{\frac{r}{2}}} \left(1 + \frac{(T)^2}{\nu}\right)^{-\frac{n+r+1}{2}} 2^{\frac{r}{2}} \Gamma\left(\frac{n+r+1}{2}\right). \quad (2.4)$$

One can also write

$$T = \frac{Z + \delta}{\sqrt{W/\nu}}$$

where $Z \sim N(0, 1)$ and $W \sim \chi_\nu^2$.

The Figure 2.9 shows the graphic representation of the probability density function of the t distribution with $\nu = 5$ degrees of freedom and some values for the non-centrality parameter. The non-central t distribution is a generalization of the t distribution. It can be shown that the estimator

$$T = \frac{\bar{X} - \mu}{S/\sqrt{\nu}}, \quad (2.5)$$

where $\bar{X}$ is the sample mean and S is the sample standard deviation of a random sample of size ν from a normal population with mean μ . If the population mean is μ_a , then $T \sim t_{\nu-1,\delta}$ where

$$\delta = \frac{\mu_a - \mu}{\sigma/\sqrt{\nu}}. \quad (2.6)$$

Figure 2.9 – Graph of the central and non-central t distribution with $\nu = 5$ degrees of freedom and some values for the non-centrality parameter δ

The non-centrality parameter is a normalized difference between μ_a and μ . The t distribution provides the probability of a t test reject correctly a false null hypothesis of the mean μ when the population mean is actually μ_a . This probability is called power of the t test. The increase in the $\mu_a - \mu$ difference, as well as the increase in the sample size ν , increases the test power.

Consider the hypotheses

$$H_0 : \mu \leq \mu_0 \quad \text{versus} \quad H_1 : \mu > \mu_0.$$

For a given level of significance α , the power of the t test is the probability of rejecting the null hypothesis when in fact the true mean μ is greater than μ_0 , given by

$$P(t > t_{\nu-1, 1-\alpha} | H_1) = P(t_{\nu-1, \delta} > t_{\nu-1, 1-\alpha}), \quad (2.7)$$

where t is given by (2.5), $t_{\nu-1, 1-\alpha}$ denotes the $(1-\alpha)$ th quantile of the t distribution with $\nu-1$ degrees of freedom, and $t_{\nu-1, \delta}$ denotes the random variable T with $\nu-1$ degrees of freedom and non-centrality parameter given by (2.6). The test power for bilateral hypotheses is calculated in a similar way.

A hypothetical situation of a bilateral hypothesis test is represented in Figure 2.10, which shows the density curve of the central t distribution under the null hypothesis with $\nu = 5$ degrees of freedom and also the density of the non-central t distribution under the alternative hypothesis. Additionally, the power of the hypothesis test and the type II error rate for this situation are illustrated.

Figure 2.10 – Graph of the central and non-central t distribution with $\nu = 5$ degrees of freedom, type II error rate (β), test power ($1 - \beta$) for a significance level $\alpha = 0,05$

An approximation can be made to a standard normal using

$$Z = \frac{T \left(1 - \frac{1}{4\nu}\right) - \delta}{\sqrt{1 + \frac{(T)^2}{2\nu}}}$$

where Z is distributed asymptotically as a standard normal variable.

2.3.4 Power of the F -Test

The power or sensitivity of the F test depends of the level of significance α , the numbers of degrees of freedom of the numerator and denominator of the statistic F and of the parameter of non-centrality given by

$$\lambda = \frac{r \sum_{i=1}^k (\mu_i - \mu)^2}{\sigma^2} \quad (2.8)$$

where μ is the average of μ_i , $i = 1, \dots, k$. Since (2.8) is obtained, the non-central F distribution can be used to calculate power. However, there is need for values of μ_i that are unknown. One way to reverse this is to stipulate a difference Δ between the means of the treatments tested; so the non-centrality parameter becomes:

$$\lambda = \frac{rm}{2} \left(\frac{\Delta}{\sigma} \right)^2 \quad (2.9)$$

where m is the multiplier of r which gives the number of observations (rm) used to calculate the averages to be compared. For σ^2 , we usually use the estimate given by the mean square residuals of some experiment performed.

The non-centrality parameter given by (2.8) resembles the F statistic in its structure, thus, replacing its constituents by values from the sample has an estimate given

by

$$\hat{\lambda} = \frac{r \sum_{i=1}^k (\bar{Y}_i - \bar{Y})^2}{\hat{\sigma}^2} = (k-1)F. \quad (2.10)$$

This estimate in terms of the F statistic is the product of the value of the statistic by the number of degrees of freedom of the numerator. According to Helms (1992), in the approximate F tests for mixed effects models, this same idea, to approximate the parameter of non-centrality by the product of the statistic by the number of degrees of freedom, was considered very favorable for small samples, based in simulation studies. In the same context of approximate F tests, Verbeke and Lesaffre (1999), Stroup (2002), Tempelman (2005), Rosa et al. (2005), among other authors, also used this approximation to calculate the test power for fixed effects in mixed effects models.

In a study of the power of a statistical test, the criterion of significance, sample size, effect size, and power are related to each other so that each of them is a function of the other three (COHEN, 1988; NAKAGAWA; FOSTER, 2004). This relationship makes possible four types of statistical power analysis (COHEN, 1965; NAKAGAWA; CUTHILL, 2007). To exemplify these types of analysis, three values were considered for the variance ($\sigma^2 = 2000$, $\sigma^2 = 3000$ and $\sigma^2 = 4000$):

(i) Power as a function of the significance level, effect size and sample size

This type of power analysis is useful to the researcher as part of the research planning, which can change the experiment settings in view of the test power result. Consider, for example, the planning of a performance experiment with broiler chickens in the completely randomized design with the following preliminary configuration: 4 treatments, 5 replicates and significance level $\alpha = 0.05$. It is possible to evaluate the power as a function of the effect size with significance level and sample size fixed, according to Figure 2.11. It is noted that the increase in effect size provides greater test power, characterizing a non-decreasing relationship. In addition, it is observed that the lower is the value of the variance, more accelerated is the growth of the test power. By setting the effect size to 50g, one can evaluate the power according to the sample size or the significance level, according to Figures 2.12 and 2.13, respectively. It is observed that these last two relations are also non-decreasing.

Figure 2.11 – Power of test as a function of the effect size

Figure 2.12 – Power of test as a function of the significance level

Figure 2.13 – Power of test as a function of the sample size

(ii) Sample size as a function of the effect size, significance level and power

The investigator specifies the effect size he wants to detect, the level of significance, the expected power, and determines the required sample size to meet those specifications. This type of analysis should be at the center of the planning in any research on

the sample size decision (COHEN, 1965). As an example, consider planning a performance experiment with broiler chickens in the completely randomized design with 4 treatments, significance level 0.05 and test power 0.80. It is possible to evaluate the sample size in function of the effect size, according to Figure 2.14. Note that as the size of the effect is increased the number of repetitions decreases while keeping the other parameters fixed. By setting the size of the effect to 50g, one can evaluate the sample size in function of the power or level of significance, according to Figures 2.15 and 2.16, respectively. It is observed that the increase in the test power requires a greater number of repetitions and in opposition to this, the increase in the significance level decreases the number of repetitions required.

Figure 2.14 – Sample size as a function of the effect size

Figure 2.15 – Sample size as a function of the test power

Figure 2.16 – Sample size as a function of the significance level

(iii) Effect size as a function of the significance level, sample size and power

This type of power analysis is generally less used than the first two. A researcher can know the size of the detectable effect for a particular experiment by specifying the significance level, sample size, and test power by considering an estimate of the variance of a pilot study.

(iv) Significance level as a function of the sample size, test power and effect size

This type of analysis is rare due to strong convention adopted by most researchers as to the significance level.

Four types of test power analysis have been described, but as mentioned, the first two are generally more interesting to the researcher.

2.3.5 Mixed Models

A mixed model is a statistical model that contains fixed effect factors and random effects factors simultaneously.

Described in Laird and Ware (1982) and Harville (1977) the mixed model for each vector $\mathbf{y}_i$ of observations is denoted by:

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\alpha} + \mathbf{Z}_i \mathbf{b}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \dots, N, \quad (2.11)$$

where $\mathbf{y}_i$ is a vector ($n_i \times 1$) of response of the i th experimental unit, $\boldsymbol{\alpha}$ is a vector ($p \times 1$) of the unknown fixed effects, $\mathbf{X}_i$ is a known design matrix ($n_i \times p$) of fixed effects linking $\boldsymbol{\alpha}$ to $\mathbf{y}_i$; $\mathbf{b}_i$ is an unknown vector ($k \times 1$) of random effects, $\mathbf{Z}_i$ is a known design matrix ($n_i \times k$) of random effects linking $\mathbf{b}_i$ to $\mathbf{y}_i$; N is the number of observations, p is the number of parameters of fixed effects, k is the number of random effects.

It is assumed that $\boldsymbol{\epsilon}_i$ is normally distributed with mean $\mathbf{0}$ and matrix of variance and covariance $\mathbf{R}_i$. The variance-covariance matrix $\mathbf{R}_i$ has dimension $(n_i \times n_i)$ and by definition is positive-definite, its size depends on i , but not the parameters in $\mathbf{R}_i$. The vector of random effects $\mathbf{b}_i$ is distributed as normal with mean $\mathbf{0}$ and matrix of variance and covariance $\mathbf{G}$, by hypothesis it is positive definite of dimension $(k \times k)$ and $\mathbf{b}_i$ are independently of each other and of the $\boldsymbol{\epsilon}_i$. Then,

$$E(\mathbf{Y}_i) = \mathbf{X}_i\boldsymbol{\alpha} \quad \text{and} \quad \text{Var}(\mathbf{Y}_i) = \mathbf{V}_i = \mathbf{R}_i + \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^T.$$

If all variance-covariance parameters are known, then, an estimator for $\boldsymbol{\alpha}$ is given by

$$\hat{\boldsymbol{\alpha}} = \left(\sum_1^m \mathbf{X}_i^T \mathbf{W}_i \mathbf{X}_i \right)^{-1} \sum_1^m \mathbf{X}_i^T \mathbf{W}_i \mathbf{y}_i \quad (2.12)$$

and a predictor for $\mathbf{b}_i$ is

$$\hat{\mathbf{b}}_i = \mathbf{G}\mathbf{Z}_i^T \mathbf{W}_i (\mathbf{y}_i - \mathbf{X}_i \hat{\boldsymbol{\alpha}}), \quad (2.13)$$

where $\mathbf{W}_i = \mathbf{V}_i^{-1}$.

If the variance-covariance matrix parameters are unknown, but estimates of $\mathbf{R}_i$ and $\mathbf{G}$ are available, then $\hat{\mathbf{V}}_i = \hat{\mathbf{R}}_i + \mathbf{Z}_i \hat{\mathbf{G}} \mathbf{Z}_i^T = \hat{\mathbf{W}}_i^{-1}$, $\boldsymbol{\alpha}$ is estimated and predictions are obtained for $\mathbf{b}_i$ using the equations (2.12) and (2.13) replacing $\mathbf{W}_i$ by $\hat{\mathbf{W}}_i$.

2.3.6 Selection of models

The selection of the appropriate model is an important step in the analysis of the data set, it is sought to choose the model that explains well the behavior of the response variable and that contains the minimum of possible parameters to be estimated. Model selection is used when there is no particular clear choice among the many possible different models. In the literature there are several discussions on this subject, some of them can be found in Jennrich e Schluchter (1986), Diggle (1988), Lindsey (1993), Pinheiro and Bates (2000), Verbeke and Molenberghs (2000), Weiss (2005), among others. Several criteria for model selection are presented, including the likelihood ratio test (LRT), the Akaike information criterion - AIC (AKAIKE, 1974; SAKAMOTO et al., 1986) and the Bayesian information criterion - BIC (SCHWARZ, 1978).

2.3.6.1 Likelihood Ratio Test

The likelihood ratio test (LRT) can be used to compare nested models, that is, when one model represents a special case of the other, fitted by maximum likelihood

or restricted maximum likelihood. The alternative hypothesis (H_1) presents the general model with more parameters, this being the reference model, while the null hypothesis (H_0) presents the restricted model with fewer parameters. The statistic used for the test is given by:

$$\Lambda = 2\log\left(\frac{L_2}{L_1}\right) = 2[\log(L_2) - \log(L_1)]$$

where L_2 is the likelihood of the general model and L_1 is the likelihood of the restricted model. Wilks (1938) has shown that if l_k is the number of parameters to be estimated in the k model, then the asymptotic distribution of the LRT statistic under the null hypothesis, which is suitable for the restricted model, follows a χ^2 distribution with $l_2 - l_1$ degrees of freedom. Thus, to test H_0 versus H_1 , with significance level α , we compare Λ to a χ_k^2 with $k = l_2 - l_1$ degrees freedom. When $\Lambda \geq \chi_{(k,\alpha)}^2$ we reject H_0 in favor of H_1 .

In selecting the random effects structure, different nested models are usually fitted, the random effects structure is altered, and the likelihood ratio test is applied to evaluate the terms significance. According to Stram and Lee (1994), tests on the structure of random effects using LRT may be conservative, that is, the p value calculated from the $\chi_{l_2-l_1}^2$ distribution may be greater than it should actually be.

2.3.6.2 Information criteria

An alternative to the likelihood ratio test when comparing non-nested models are the information criteria, which can also be used in comparisons of nested models. The two most popular criteria for selecting models are Akaike's information criterion (AIC) and Bayesian information criterion (BIC) (WEISS, 2005). These criteria use a penalty term applied to the likelihood function. For the calculation of AIC and BIC the following expressions are used:

$$AIC = -2l(\hat{\beta}, \hat{\theta}, \hat{\sigma}) + 2k$$

$$BIC = -2l(\hat{\beta}, \hat{\theta}, \hat{\sigma}) + k\log(n)$$

where θ is the vector of parameters of variance components, $l(\hat{\beta}, \hat{\theta}, \hat{\sigma})$ is the value of the logarithm of the likelihood function of the calculated model with the estimates obtained in the maximization process, k represents the total number of model parameters and n is the number of observations used in the estimation of the model under study. AIC or BIC is used to compare two or more models for the same data; the model with the lowest AIC or BIC value is selected as the most appropriate.

Guerin e Stroup (2000) compared the AIC and BIC information criteria for the ability to select the "correct model" and the impact of choosing the "wrong model" based on the type I error rate. They confirmed that the AIC tends to select more complex models than the BIC, and also, the choice of a very simple model affects the control of the type I error rate of negative way. Thus, when the priority is the control of the type I error rate,

AIC is recommended, however, if power loss is relatively more severe, BIC is preferable (LITTELL et al., 2006).

2.3.7 Tests for the fixed effects

The main objective of a statistical analysis is not simply the fitting of a model; the primary interest is in making inferences about its parameters in order to generalize the results to the population from a specific sample. The fixed effects vector is estimated by (2.12) and since the variance components associated with the matrix W_i are unknown, there is a need to replace them with their estimates of maximum likelihood (ML) or restricted maximum likelihood (REML).

2.3.7.1 Approximate Wald Tests

The approximate Wald test, also called the Z test, for each parameter α_j in $\boldsymbol{\alpha}$, $j = 1, \dots, p$, as well as a confidence interval is obtained from an approximation of the distribution of $(\hat{\alpha}_j - \alpha_j)/s.e(\hat{\alpha}_j)$ by a standard univariate normal distribution, where $s.e$ is the associated standard deviation. Generally, for any known $\mathbf{L}$ matrix, a test for the hypothesis

$$H_0 : \mathbf{L}\boldsymbol{\alpha} = \mathbf{0} \quad \text{versus} \quad H_A : \mathbf{L}\boldsymbol{\alpha} \neq \mathbf{0} \quad (2.14)$$

from the fact that the distribution of

$$(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})' \mathbf{L}' \left[\mathbf{L} \left(\sum_1^m \mathbf{X}'_i \mathbf{V}_i^{-1}(\hat{\boldsymbol{\theta}}) \mathbf{X}_i \right)^{-1} \mathbf{L}' \right]^{-1} \mathbf{L}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \quad (2.15)$$

asymptotically follows a χ^2 distribution with number of degrees of freedom given by $\text{rank}(\mathbf{L})$, where $\boldsymbol{\theta}$ is the vector of variance components.

2.3.7.2 Approximate t-Tests and F-Tests

The Wald test statistics underestimate the true variability of $\hat{\boldsymbol{\alpha}}$ because they do not take into account the variability introduced by the $\boldsymbol{\theta}$ estimate as discussed by Dempster, Rubin e Tsutakawa (1981). Due to this limitation of the Wald test, for the tests concerning the fixed parameters, Verbeke e Molenberghs (2000) advise the use of the approximate t and F tests.

An approximate t test, as well as a confidence interval for each parameter α_j in $\boldsymbol{\alpha}$, $j = 1, \dots, p$, can be obtained by approximating the distribution of $(\hat{\alpha}_j - \alpha_j)/s.e(\hat{\alpha}_j)$ by an appropriate t distribution. The approximate F test to test hypotheses as presented

in (2.14) is based on the approximation of the F distribution whose statistics is as follows:

$$F = \frac{(\hat{\alpha} - \alpha)' \mathbf{L}' \left[\mathbf{L} \left(\sum_1^m \mathbf{X}'_i \mathbf{V}_i^{-1} (\hat{\theta}) \mathbf{X}_i \right)^{-1} \mathbf{L}' \right]^{-1} \mathbf{L} (\hat{\alpha} - \alpha)}{\text{rank}(\mathbf{L})} \quad (2.16)$$

with the number of degrees of freedom of the numerator given by $\text{rank}(\mathbf{L})$. Several methods can be used for the appropriate calculation of the number of degrees of freedom of the denominator of the F test and the number of degrees of freedom associated with the t test. Among the methods, we can mention: the Residual method, the Containment method, which is the SAS software standard method (SAS INSTITUTE, 2004), the method of Satterthwaite (1941, 1946) and the method of Kenward-Roger (KENWARD; ROGER, 1997).

2.3.8 Diagnostics

Model diagnostics are important for the construction of a model, because with them the assumptions of distribution for the residuals and the sensitivity of the model for the unusual observations are verified. The diagnostic tools for classical linear models are well established in the literature, for example, details of development and applications can be seen in Cook (1977), Hoglin and Welsch (1978), Welsch and Kuh (1977), Belsley et al. (1980), Atkinson (1985) and others. For mixed models, the volume of work in this area is relatively smaller because of complexity and because it has been formulated later in relation to the classic models. In general, mixed models require iterative optimization, have more components, may have more types of residuals, have conditional and marginal distributions, and are most often applied to data with clustered structures (LITTELL et al., 2006).

Nobre and Singer (2007) and Hilden-Minton (1995) defined three types of residuals in mixed linear models

- (i) Marginal residual: $\hat{\xi} = \mathbf{y} - \mathbf{X}\hat{\beta}$;
- (ii) Conditional residual: $\hat{\epsilon} = \mathbf{y} - \mathbf{X}\hat{\beta} - \mathbf{Z}\hat{\mathbf{b}}$;
- (iii) EBLUP: $\mathbf{Z}\hat{\mathbf{b}}$, which predicts the random effects $\mathbf{Z}\mathbf{b} = \text{E}[\mathbf{Y}|\mathbf{b}] - \text{E}[\mathbf{Y}]$.

The authors make recommendations regarding the use of each type of residual to evaluate some kind of model assumption (2.11). For example, Hilden-Minton (1995) suggests the use of the marginal residual ($\hat{\xi}$) to evaluate the linearity assumption of the relationship between $\text{E}[\mathbf{Y}]$ and the covariates $\mathbf{X}$, in addition to their use in the evaluating of the validity of covariance structure. Pinheiro and Bates (2000) suggest the use of the conditional residual to verify the hypothesis of normality and homoscedasticity of conditional error. This type of residual can also be used to identify discrepant observations.

The EBLUP can be used to detect possible discrepant experimental units, to evaluate the normality assumption of random effects, as well as to verify its variance and covariance structure as suggested by Pinheiro and Bates (2000).

Available computational tools aid in the diagnosis of mixed linear models. For more details see a description of existing methods for the SAS software in Schabenberger (2004). For the R software, recently, the `HLMdiag` library has been developed and the details of its use can be observed in Loy e Hofmann (2014) and in the documentation itself of the library.

2.4 Results

The power of the F test, represented by $(1 - \beta)$, was calculated for the live weight data of chicken at each time of the experimental period considering a mean difference (Δ) of approximately 2% of the average weight of chickens in each of the times (7, 21, 35 and 42) with a significance level $\alpha = 0.05$. In addition, was calculated the number of repetitions (r) required to detect the mean difference (Δ) with an approximate probability of 0.8. Table (2.5) shows these results together with the mean weight in grams per treatment at each time and the mean square of the residual (MSE). Note that at 7 days, to detect a difference between averages of 5 grams with a probability of 0.8 are required 24 replicates. At 21 days, 21 replicates for a difference of 25 grams, at 35 days, 45 replicates for a difference of 40 grams and at 42 days, 33 replicates for a difference of 50 grams. Note that these amounts of replicates required to obtain a test power of approximately 0.8 are difficult in practice, for example at 42 days, we would have 33 replicates (boxes) with 46 birds each for one of the 6 treatments, totaling 9108 birds distributed in 198 boxes.

Table 2.5 – Number of replicates (r) required to detect the mean difference in grams (Δ) with probability 0.8, at each time of the experimental period for live weight data of chickens

Time (Days)	MSE	r	Δ	$1 - \beta$	Treatment mean (g)					
					T1	T2	T3	T4	T5	T6
7	22.2	24	5	0.17	129	149	160	144	175	188
21	482.0	21	25	0.20	807	839	856	838	869	895
35	2727.7	45	40	0.11	1945	1983	2024	2014	2018	2102
42	3070.6	33	50	0.14	2658	2685	2726	2713	2738	2818

r is the number of replicates (boxes per treatment) to detect Δ with a probability of approximately 0.8
 Δ is the difference to detect in grams

Note that the power of the F test was low to detect the mean differences (Δ) at each time. Knowing that the power of test F , besides depending of the level of significance, of the number of degrees of freedom of the numerator and of the denominator of the F statistic, also depends of the non-centrality parameter given by expression (2.8). The greater the variance (σ^2), the lower the non-centrality parameter of the F distribution under the

alternative hypothesis, consequently, the lower the power of the test, that is, the variance is one of the factors that influence the power of the F test.

Based on the chicken weight data at 42 days, graphs were used to evaluate the power of the F test, the mean difference between the treatments and the sample size with the significance level fixed at 0.05. It was considered the estimate obtained with the mean square residual as the value of variance σ^2 and percentages of 75% and 50% of the variance for the construction of curves in each graph.

Figure (2.17) shows graph with the curves of power of the F test as a function of the mean difference. Note the curve with $\sigma^2 = 3071$, the probability of 0.8 is reached with mean difference around 140 grams. By reducing the value of the variance by 50%, the probability of approximately 0.8 is reached with mean difference around 70 grams. There is considerable gain in detecting smaller mean differences when variance is reduced.

Figure 2.17 – Power of the F test as a function of the mean difference for the experiment with chicken weight data at 42 days with different variances

Figure (2.18) shows the graph of the power of the F test as a function of the sample size in the detection of $\Delta = 50$ grams. As the size of the sample increases, the power of the test also increases and the reduction in the value of the variance requires a smaller number of boxes per treatment. We can also note that for the experiment with 5 replicates per treatment, the power of the test is less than 0.2 in the detection of 50 grams.

Figure 2.18 – Power of the F test as a function of sample size (number of replicates per treatment) for the experiment with chicken weight data at 42 days with different variances

The sample size as a function of the mean difference to be detected with an approximate probability of 0.8 can be visualized in Figure (2.19). Note that the larger the mean difference to be detected, the less replicates are required. Note also that reductions in the value of variance considerably decrease the required sample size when the mean difference to be detected is 50 grams. As we increase the difference to be detected, reductions in the value of the variance do not cause large changes in sample size.

Figure 2.19 – Sample size (number of replicates per treatment) as a function of the mean difference for the experiment with chicken weight data at 42 days with different variances

Figures (2.20) and (2.21) show the graphs of the power of the F test and the sample size as functions of the σ . In the experiment under study, at 42 days, the estimate for σ was equal to 55.47. Note in Figure (2.20) that as we increase the value of σ , the power of the test decreases and the relation between the power of the test and σ is not linear. To obtain a test power of approximately 0.8, we would need a reduction of approximately 65% of σ to detect a difference of 50 grams. In Figure (2.21), note that the greater the value of σ , the greater the number of replicates required to detect a difference of 50 grams, also we note a nonlinear relationship between sample size and σ .

Figure 2.20 – Power of the F test as a function of σ for the experiment with chicken weights at 42 days considering $\Delta = 50\text{g}$, 5 replicates and $\alpha = 0.05$

Figure 2.21 – Sample size as a function of σ for the experiment with chicken weights at 42 days considering $\Delta = 50\text{g}$, $(1 - \beta) \approx 0.8$ and $\alpha = 0.05$

In Appendix A of this work is the code in R (R Core Team, 2017) used for the preparation of charts in this section.

2.5 Discussion

In this chapter we work with power analysis of the F test and some of the factors that influence it for chicken weight data. We emphasize the relationships between test power, sample size, mean difference to be detected and variance.

We observed that the larger the mean difference to be detected, the greater the power of the test, while maintaining fixed the sample size, the level of significance and the variance. The power of the test also increases as we increase the sample size by keeping the other factors involved fixed. We also noticed that the sample size depends on the size of the mean difference that the researcher wants to detect by the statistical test, the smaller the difference, the larger the sample size required.

We note that the variance has a strong influence on the power of the F test, the lower the variance, the higher the power of the test, the smaller the sample size needed for the experiment and can be detected the smaller differences between the treatment means.

The data of chicken weights worked in this chapter presents great variability within the plot due to the presence of male and female birds inside the same box. The models commonly used do not take into account the sex of the birds because there is no such identification. The variability between males and females contributes to the increase of the mean square of the residual, which reflects in the loss of test power and the need to increase the sample size.

With the intention of reducing the mean square of the residual, we propose in the next chapter a model that takes into account the information about the sex of the birds.

References

- AARON, D.K.; HAYS, V.W. How many pigs? Statistical power considerations in swine nutrition experiments. **Journal of Animal Science**, Champaign, v. 82, p. E245-E254, 2004.
- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, New York, v. 19, n. 6, p. 716-723, Dec. 1974.
- ATKINSON, C.A. **Plots, Transformations and Regression: An Introduction to graphical methods of diagnostic regression analisys**. Oxford: Oxford University Press, 1985. 282 p.
- BELSLEY, D.A., KUH, E., WELSCH, R.E. **Regression Diagnostics: Identifying influential data and sources of collinearity**. New York: John Wiley & Sons, 1980. 292 p.
- BERNDTSON, W.E. A simple, rapid and reliable method for selecting or assessing the number of replicates for animal experiments. **Journal of animal science**, Champaign, v. 69, p. 67-76, 1991.

CASELLA, G.; BERGER, R.L. **Statistical Inference**. Pacific Grove: Thomson Learning, 2002. 686 p.

COHEN, J. **Some statistical issues in psychological research**. In B.B. Wolman (Ed.). *Handbook of clinical psychology*, New York: McGraw-Hill, 1965. 121 p.

_____. **Statistical Power Analysis for the Behavioral Sciences**. New York: Academic Press, INC. 1977. 474 p.

_____. **Statistical Power Analysis for the Behavioral Sciences**. New Jersey: Hillsdale, 1988. 567 p.

COOK, R.D. Detection of Influential Observation in Linear Regression. **Technometrics**, v. 19, n. 1, p. 15-18, 1977.

COUSENS, R.; MARSHALL, C. Dangers in testing statistical hypotheses. **Annals of Applied Biology**, Hoboken, v. 111, p. 469-476, 1987.

DEMÉTRIO, C.G.B.; MENTEN, J.F.M.; LEANDRO, R.A.; BRIEN, C. Experimental power considerations - justifying replication for animal care and use committees. **Poultry Science**, Savoy, v. 92, p. 2490-2497, 2013.

DEMPSTER, A.P.; RUBIN, D.B.; TSUTAKAWA, R.K. Estimation in Covariance Components Models. **Journal of the American Statistical Association**, Alexandria, v. 76, p. 341-353, 1981.

DIGGLE, P.J. An Approach to the Analysis of Repeated Measures. **Biometrics**, Arlington, v. 44, p. 959-971, 1988.

FISHER, R.A. Moments and product-moments of sampling distributions. **Proc. Lond. Math. Soc.**, London, p. 199-238, 1928.

GUERIN, L.; STROUP W. A Simulation Study to Evaluate PROC MIXED Analysis of Repeated Measures Data. In **Proceedings of the Twelfth Annual Conference on Applied Statistics in Agriculture**, Manhattan: Kansas State University, 2000.

HARVILLE, D.A. Maximum likelihood approaches to variance component estimation and to related problems. **Journal of the American Statistical Association**, Alexandria, v. 72, p. 320-340, 1977.

HELMS, R.W. Intentionally incomplete longitudinal designs: Methodology and comparison of some full span designs. **Statistics in Medicine**, Chichester, v. 11, p. 1889-1913, 1992.

HILDEN-MINTON, J.A. **Multilevel diagnostics for mixed and hierarchical linear models**. 1995. 110p. Tese (Doutorado) - University of California, Los Angeles, 1995.

HOGLIN, D.C., WELSCH, R.E. The Hat Matrix in Regression and ANOVA. **The American Statistician**, v. 32, n. 1, p. 17-22, 1978.

JENNRICH, R.I.; SCHLUCHTER, M.D. Unbalanced Repeated-Measures Models with Structured Covariance Matrices. **Biometrics**, Arlington, v. 42, p. 805-820, 1986.

JOHNSON, N.L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions**. New York: Wiley, 1995. 717 p.

KENDALL, M.G.; STUART, A. **The Advanced Theory of statistics**. London: Charles Griffin & Company Limited, 1961. 457 p.

KENWARD, M.G.; ROGER, J.H. Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. **Biometrics**, Hoboken, v. 53, n. 3, p. 983-997, 1997.

LAIRD, N.M.; WARE, J.H. Random-effects models for longitudinal data. **Biometrics**, Hoboken, v. 38, p. 963-974, 1982.

LINDSEY, J.K. **Models for Repeated Measurements**. Oxford: Clarendon Press, 1993. 536 p.

LITTELL, R.C.; MILLIKEN, G.A.; STROUP, W.W.; WOLFINGER, R.D.; SCHABENBERGER, O. **SAS® for Mixed Models**. 2.ed. Cary, NC: SAS Institute Inc. 2006. 814 p.

LOY, A., HOFMANN, H. HLMdiag: A Suite of Diagnostics for Hierarchical Linear Models in R. **Journal of Statistical Software**, Insbruck, v. 56, n. 5, 2014.

NAKAGAWA, S.; FOSTER, T.M. The case against retrospective statistical power analyses with an introduction to power analysis. **Acta Ethologica**, v.7, p. 103-108, 2004.

NAKAGAWA, S.; CUTHILL, I.C. Effect size, confidence interval and statistical significance: a practical guide for biologists. **Biol. Rev.**, v.82, p. 591-605, 2007.

NOBRE, J.S., SINGER, J.M. Residual analysis for linear mixed models. **Biometrical Journal**, Berlin, v. 49, n. 6, p. 863-875, 2007.

PATNAIK, P.B. The non-central χ^2 and F-distributions and their applications. **Biometrika**, v. 36, 202-232, 1949.

PINHEIRO, J.C.; BATES, D.M. **Mixed-Effects Models in S and S-PLUS**. New York: Springer-Verlag, 2000. 528 p.

R: A Language and Environment for Statistical Computing, R Development Core Team, R Foundation for Statistical Computing, Vienna, 2017. Disponível em: <<http://www.R-project.org>>. Acesso em: 08 jun. 2018.

RAMALHO, M.A.; FERREIRA, D.F.; OLIVEIRA, A.C. **Experimentação em Genética e Melhoramento de Plantas**. Lavras: UFLA, 2005. 300 p.

RIZZO, P. **Misturas de extratos vegetais como alternativas ao uso de antibióticos melhoradores do desempenho nas dietas de frangos de corte**. 2008. 69 p. Dissertação (Mestrado em Ciências Animal e Pastagens) - Escola Superior de Agricultura "Luiz de Queiroz" - Universidade de São Paulo, Piracicaba, 2008.

ROSA, G.J.M.; STEIBEL, J.P.; TEMPELMAN, R.J. Reassessing Design and Analysis of Two-Colour Microarray Experiments Using Mixed Effects Models. **Comparative and Functional Genomics**, Oxford, v. 6, n. 3, p. 123-131, 2005.

ROSTAGNO, H.S. **Tabelas Brasileiras para aves e suínos: composição de Alimentos e exigências nutricionais**. Viçosa: UFV, Departamento de Zootecnia, 2005. 186 p.

SAKAMOTO, Y.; ISHIGURO, M.; KITAGAWA, G. **Akaike Information Criterion Statistics**. Dordrecht: Reidel, 1986. 290 p.

SAS INSTITUTE INC. **SAS OnlineDoc® 9.1.3**. Cary, 2004. Disponível em: <<http://support.sas.com/onlinedoc/913/docMainpage.jsp>>. Acesso em: 21 jan. 2018.

SATTERTHWAITE, F.E. Synthesis of Variance. **Psychometrika**, New York, v. 6, n. 5, 1941, p. 309-316.

———. An approximate distribution of estimates of variance components. **Biometrics**, Hoboken, v. 2, n. 6, p. 110-114, 1946.

SCHABENBERGER, O. **Mixed Model Influence Diagnostics, in Proceedings of the Twenty-Ninth Annual SAS Users Group International Conference**. Paper 189-29, Cary, NC: SAS Institute Inc. 2004. 17p.

SCHEFFÉ, H. **The Analysis of Variance**. New York: John Wiley & Sons, 1959. 479 p.

SCHWARZ, G. Estimating the dimension of a model. **Annals of Statistics**, Hayward, v. 6, n. 2, p. 461-464, 1978.

STRAM, D.O.; LEE, J.W. Variance Components in the Longitudinal Mixed Effects Model. **Biometrics**, Arlington, v. 50, p. 1171-1177, 1994.

STROUP, W.W. Power Analysis Based on Spatial Effects Mixed Models: A Tool for Comparing Design and Analysis Strategies in the Presence of Spatial Variability. **Journal of Agricultural, Biological, and Environmental Statistics**, London, v. 7, n. 4, p. 491-511, 2002.

TANG, P.C. The power function of the analysis of variance tests with tables and illustrations of their use. **Statist. Res. Mem.**, v.2, n.126, 1938.

TEMPELMAN, R.J. Assessing Statistical Precision, Power, and Robustness of Alternative Experimental Designs for Two Colour Microarray Platforms Based on Mixed Effects Models. **Veterinary Immunology and Immunopathology**, Amsterdam, v. 105, p. 175-186, 2005.

TRALDI, A.B. **Influência da idade da matriz e do peso do ovo incubado nas respostas de pintos de corte alimentados com rações pré-iniciais farelada, triturada ou micro-peletizada**. 2009. 205p. Tese (Doutorado em Ciências) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2009.

VERBEKE, G.; LESAFFRE, E. The Effect of Drop-Out on the Efficiency of Longitudinal Experiments. **Journal of the Royal Statistical Society**, Hoboken, v. 48, n. 3, p. 363-375, 1999.

VERBEKE, G.; MOLENBERGHS, G. **Linear Mixed Models for Longitudinal Data**. New York: Springer-Verlag, 2000. 568 p.

WEISS, R. **Modeling longitudinal data**. New York: Springer-Verlag, 2005. 432 p.

WELSCH, R.E., KUH, E. Linear Regression Diagnostics. **Working Paper**, Massachusetts: National Bureau of Economic Research, n. 173, 1977.

WILKS, S.S. The large-sample distribution of the likelihood ratio for testing composite hypothesis. **The Annals of Mathematical Statistics**, New York, v. 9, n. 1, p. 60-62, 1938.

WISHART, J. A note on the distribution of the correlation ratio. **Biometrika**, v.24 ,n. 3/4, pp. 441-456, 1932.

3 MIXTURE MODELS FOR THE ANALYSIS OF CHICKENS WEIGHT

Abstract

In experiments with broilers in which the sexing of the birds is not done and the experimental unit is the box, the statistical analyzes are usually made based on models that do not take into account the mixture of sexes. This omission implies the increase of the residual and less accuracy of the inference on the population of interest. In this chapter we propose a model for the total weight of the birds per box that takes into account the sex information of the chickens. We used the classical and Bayesian approach to estimate the parameters of the proposed model. We observed that, in relation to the classical approach, Bayesian methods provided better estimates of the parameters of the proposed model for the total weight of the birds.

Keywords: Mixture model; Gaussian components; Heterogeneous groups; Heterogeneous components

3.1 Introduction

Brazilian poultry has a prominent position in the international market, Brazil is among the world's three largest producers of broilers. The production of broilers shows great dynamism, because in less than 50 days the chickens are ready for slaughter.

The broilers can be kept in batches separated by sex or mixed, for each batch there are different specifications for handling. Generally, mixed-batch experiments are performed according to the recommended specifications, but the statistical analyzes are done as if there was no mixing of the sexes, since the models used do not take into account the variability between the sexes, see for example Kalmendal et al. (2011), Korver et al. (2004), Rodrigues et al. (2008), Teixeira et al. (2007) and Qureshi (1996).

In mixed sex experiments with a large number of birds, the variables of interest are usually not observed by birds and there is also no gender identification. In these experiments the birds are distributed by box and the variable of interest is observed by box. In this way, it is a great challenge to model this type of data taking into account the variability between the sexes.

In this chapter we propose a model that takes into account the sex information of the birds when the observation is the total weight per box.

3.2 Case-study

The data studied in this chapter come from the experiment conducted by Traldi (2009) in the experimental aviary of the Department of Animal Science - Non-Ruminant Sector of the “Luiz de Queiroz” School of Agriculture. The description of the data can be found in section 2.2 of chapter 2 of this work.

3.3 Modelling

3.3.1 Mixture Models

The mixture models are used to analyze data from populations composed by subpopulations that are present in different proportions. Pearson (1894) developed one of the first studies involving mixing distributions, which fitted a mixture of two probability density functions of normals with different means, μ_1 and μ_2 , and different variances, σ_1^2 and σ_2^2 . To estimate the parameters Pearson used the moments method to equate the first five moments with their sample values, then reduced the five equations obtained to a ninth degree polynomial with at least one real root. Each real root provided a set of different estimates for the parameters. To select the most plausible set of estimates Pearson choose the set of estimates that resulted in greater agreement between the sixth central sample moment and the corresponding moment of the fitted mixture distribution. Because of the large number of calculations involved in the solution of the ninth degree polynomial, this approach was considered intimidating by potential users of this methodology at the time. Moreover, the procedure used by Pearson provides multiple solutions, requiring criteria to select the set of estimates more plausible for the problem. Martin (1936) used the method applied by Pearson and emphasized that such procedure could be extremely sensitive to experimental errors in the distribution and could not be relied upon to give reasonable results. For five decades after the work of Pearson (1894) the related works the mixture of the distributions involved, mainly, the development of techniques designed to simplify the approach used by Pearson. Rao (1948) was the first to suggest the estimation by maximum likelihood for the problem of mixture of normals with equal variances. In the data set used by Pearson, the maximum likelihood method presented similar estimates as those obtained by Pearson. However, the maximum likelihood method is more efficient (TAN; CHANG, 1972). Wolfe (1965, 1967, 1969, 1970) in a series of articles worked with maximum likelihood estimation for situations involving mixture of multivariate normals, and was also the pioneer in creating computer routines to deal with such models. According to Everitt (1996), the procedure suggested by Wolfe to solve the maximum likelihood equations is one of the first examples of the algorithm E.M., formulated later in more general terms by Dempster, Laird and Rubin (1977).

Due to its flexibility for data analysis from heterogeneous populations, the use of mixture models was broadcast in various areas of the knowledge. The areas in which these models have been successfully used include astronomy, biology, genetics, medicine, economics, engineering, marketing beyond many fields in biology, physics and social sciences (MCLACHLAN; PEEL, 2000).

Let $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ be a random sample of size n , wherein $\mathbf{Y}_i$ is a random vector with probability density function $f(\mathbf{y}_i)$ in $\mathbb{R}^p$. Suppose that the density function $f(\mathbf{y}_i)$ of $\mathbf{Y}_i$ can be written as

$$f(\mathbf{y}_i) = \sum_{j=1}^g \pi_j f_j(\mathbf{y}_i), \quad (3.1)$$

where $f_j(\mathbf{y}_i)$ are densities and π_j are non-negative quantities whose sum is one, i.e.,

$$0 \leq \pi_j \leq 1 \quad (j = 1, \dots, g) \quad \text{and} \quad \sum_{j=1}^g \pi_j = 1.$$

The quantities $\pi_1, \dots, \pi_g$ are called proportions or mixing weights. Through the property of convexity, given that each $f_j(\mathbf{y}_i)$ defines a probability density function, then $f(\mathbf{y}_i)$ also define a probability density function. The functions $f_j(\mathbf{y}_i)$ are called component densities of the mixture. The aim of finite mixture modeling is to estimate the parameters of each component density, $f_j(\mathbf{y}_i)$, as well as the mixing parameters, π_j .

There are different methods for the estimation of parameters that may be mentioned, for example, the moment method, maximum likelihood method, and Bayesian methods. Using the maximum likelihood method, since the equations have no analytical solution, numerical methods are used. Among the numerical methods, for example, the Newton-Raphson method or Expectation-Maximization algorithm (EM) shown in Dempster et al. (1977) can be used.

Under the Bayesian perspective, a posteriori distribution is given by

$$p(\Psi|\mathbf{y}) \propto L(\Psi)p(\Psi), \quad (3.2)$$

where

$$L(\Psi) = \prod_{i=1}^n \sum_{j=1}^g \pi_j f_j(\mathbf{y}_i) \quad (3.3)$$

is the likelihood function and $p(\Psi)$ is a prior distribution for the parameter vector Ψ .

Due to the expansion of the likelihood function in g^n terms it is difficult to obtain the maximum likelihood estimator or the Bayes estimator analytically. Some useful methods of approximations for normal mixtures in a Bayesian context can be observed in Roeder and Wasserman (1997), while in Diebolt and Robert (1994) estimation methods of the parameters are shown for missing data structures.

There is an extensive literature on mixture models and for better understanding of the subject some important references are Everitt and Hand (1981), Titterton et al. (1985), McLachlan and Basford (1988), Everitt (1996), McLachlan and Peel (2000) and Wirjanto and Xu (2009).

3.3.2 Mixtures of normal distributions

Gaussian mixture models are probabilistic models used to analyze data from populations composed of normally distributed subpopulations. In general, they are models belonging to a class of mixture models in which the component densities are densities of Gaussian probabilities. One of the first applications of such models back from the seminal work of Pearson (1894), who used the model with two components on crab data, Pearson got a model fit and raised evidence of the presence of two subspecies of crabs (MCLACHLAN; PEEL, 2000). Since then the Gaussian mixture models have been studied and applied in several fields of study, for example, in Psychology, Bock and Kolakowski (1973) used the mixture of two normal distributions with equal variances to relate spatial capacity of the human with a sex-linked recessive gene, in Medicine, Bogardus et al. (1989) used the mixture of three normal distributions with equal variances to fit a model with the fasting insulin levels among the young adult nondiabetic Pima Indians, in Archeology, Dong (1997) used the mixture of two normal distributions ones with equal variances to discriminate the dimorphism sexual in archaeological materials and samples of prehistoric human populations, and in Finance, Wirjanto and Xu (2009) provided a review of the developments and applications of mixture of normal distribution models in empirical finance.

Consider probability density functions for a vector of random variables $\mathbf{y}$, of dimension n , with the following form:

$$f(\mathbf{y}; \boldsymbol{\pi}, \boldsymbol{\Sigma}, \boldsymbol{\mu}) = \sum_{j=1}^g \pi_j f_j(\mathbf{y}; \boldsymbol{\Sigma}_j, \boldsymbol{\mu}_j)$$

where $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_{g-1})$ are the $g - 1$ independent mixing proportions of the mixture, with

$$0 < \pi_j < 1 \quad \text{and} \quad \pi_g = 1 - \sum_{j=1}^{g-1} \pi_j.$$

$\Sigma = \{\Sigma_1, \Sigma_2, \dots, \Sigma_g\}$ is a $(n \times gn)$ matrix, and $\boldsymbol{\mu} = \{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \dots, \boldsymbol{\mu}_g\}$ is a $(n \times g)$ matrix, in which $\boldsymbol{\mu}_j$ and Σ_j are the mean vector and variance-covariance matrix, respectively, of the j^{th} component multivariate normal density, $f_j(\mathbf{y}; \Sigma_j, \boldsymbol{\mu}_j)$, given by

$$f_j(\mathbf{y}; \Sigma_j, \boldsymbol{\mu}_j) = (2\pi)^{-\frac{n}{2}} |\Sigma_j|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \boldsymbol{\mu}_j)' \Sigma_j^{-1} (\mathbf{y} - \boldsymbol{\mu}_j) \right\}.$$

Let X indicate from which population an observation has been sampled, so $X = j$ means that the observation belongs to the j^{th} population. The distribution of X is discrete with support $1, 2, \dots, g$ and corresponding probabilities $\pi_1, \pi_2, \dots, \pi_g$, so we have

$$X \sim \begin{pmatrix} 1 & 2 & \dots & g \\ \pi_1 & \pi_2 & \dots & \pi_g \end{pmatrix}.$$

We can use the representation of the latent variable X to show properties of expectation and variance as follows

$$E(Y) = E[E(Y|X)] \quad \text{and} \quad \text{Var}(Y) = \text{Var}[E(Y|X)] + E[\text{Var}(Y|X)]. \quad (3.4)$$

Conditional expectation and variance follows directly from the component density f_j .

For the univariate case of the finite normal mixture distributions, we can consider three situations:

- (i) *Normals With Common Variance;*
- (ii) *Normals With Common Mean;*
- (iii) *Normals With General Mean and Variance.*

The notation for situation (i) considering g components of the mixture can be expressed by $Y|\mu \sim N(\mu, \sigma^2)$ with discrete distribution for μ given by

$$\mu \sim \begin{pmatrix} \mu_1 & \mu_2 & \dots & \mu_g \\ \pi_1 & \pi_2 & \dots & \pi_g \end{pmatrix}.$$

Using the properties of expectation and conditional variance given by (3.4) we obtain the following expressions for expectation and variance:

$$E(Y) = E(\mu) = \sum_j \pi_j \mu_j,$$

$$\text{Var}(Y) = \text{Var}(\mu) + E(\sigma^2) = \text{Var}(\mu) + \sigma^2 = \sum_j \pi_j \mu_j^2 - \left(\sum_j \pi_j \mu_j \right)^2 + \sigma^2.$$

To illustrate the flexibility of the class of models presented in (i) we will consider the case where we have two components. Thus, consider a mixture of two normal populations as follows

$$f(y) = \pi_1 \phi(y; \mu_1, \sigma^2) + \pi_2 \phi(y; \mu_2, \sigma^2) \quad (3.5)$$

where

$$\phi(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} (y - \mu)^2 \right\}. \quad (3.6)$$

Figure (3.1) presents different shapes of the mixture density of two normal distributions with homogeneous variances. Observe that asymmetry and multimodality can be modeled by this class of models by changing the proportions of the component densities of the mixture as well as the distance between the means of the same. Let Δ as the Mahalanobis distance between the homoscedastic components of the gaussian mixture density, i.e., $\Delta = |\mu_1 - \mu_2|/\sigma$. If $\Delta \leq 2$ then the mixture is unimodal for all mixture proportions π , this is illustrated in Figure (3.1) items (a), (b) and (c). If we have $\Delta > 2$ then the modality of the mixture depends on π , Figure (3.1) items (d), (e) and (f) exemplify this. We have that $f(y)$ defined in (3.5) is symmetric if $\pi_1 = \pi_2$ and $\sigma_1 = \sigma_2$, or if $\mu_1 = \mu_2$, this is illustrated in Figure (3.1) items (a), (c) and (d) (EVERITT; HAND, 1981).

Figure 3.1 – The dense line represents the graph of the density of the mixture of two normal distributions with $\sigma = 1$ and different values for the other parameters; The dashed lines represent the density components of the mixture: (a) $\mu_1 = \mu_2 = 0$, $\pi_1 = 0.6$ and $\pi_2 = 0.4$; (b) $\mu_1 = 0$, $\mu_2 = 2$, $\pi_1 = 0.55$ and $\pi_2 = 0.45$; (c) $\mu_1 = 0$, $\mu_2 = 2$ and $\pi_1 = \pi_2 = 0.5$; (d) $\mu_1 = 0$, $\mu_2 = 2.5$ and $\pi_1 = \pi_2 = 0.5$; (e) $\mu_1 = 0$, $\mu_2 = 2.5$, $\pi_1 = 0.55$ and $\pi_2 = 0.45$; (f) $\mu_1 = 0$, $\mu_2 = 2.5$, $\pi_1 = 0.85$ and $\pi_2 = 0.15$

The model in (ii) can be represented by $Y|\sigma^2 \sim N(\mu, \sigma^2)$ with discrete distribution for σ^2 given by

$$\sigma^2 \sim \begin{pmatrix} \sigma_1^2 & \sigma_2^2 & \dots & \sigma_g^2 \\ \pi_1 & \pi_2 & \dots & \pi_g \end{pmatrix}.$$

The expressions of expectation and variance in this case are given by

$$E(Y) = E(\mu) = \mu,$$

$$\text{Var}(Y) = \text{Var}(\mu) + \text{E}(\sigma^2) = \text{E}(\sigma^2) = \sum_j \pi_j \sigma_j^2.$$

For the case in (iii) we have the representation $Y|(\mu, \sigma^2) \sim N(\mu, \sigma^2)$ with discrete distribution for (μ, σ^2) given by

$$(\mu, \sigma^2) \sim \begin{pmatrix} (\mu_1, \sigma_1^2) & (\mu_2, \sigma_2^2) & \dots & (\mu_g, \sigma_g^2) \\ \pi_1 & \pi_2 & \dots & \pi_g \end{pmatrix}.$$

For the mixture model of normals with general mean and variance the expressions for expectation and variance are given by

$$\text{E}(Y) = \text{E}(\mu) = \sum_j \pi_j \mu_j,$$

$$\text{Var}(Y) = \text{Var}(\mu) + \text{E}(\sigma^2) = \text{Var}(\mu) + \sum_j \pi_j \sigma_j^2 = \sum_j \pi_j \mu_j^2 - \left(\sum_j \pi_j \mu_j \right)^2 + \sum_j \pi_j \sigma_j^2.$$

3.3.3 Methods of estimation

Estimating parameters of a finite mixture of normal distributions has been considered by several authors, see Cooper (1964) and Cohen (1967) who worked with moment estimators, Wolfe (1965, 1967, 1969, 1970) in a series of papers considers maximum likelihood estimators, Xu (2010) worked with discrete mixtures of normal family based on the continuous empirical characteristic function method and Lourens (2013) compared under various settings the performance of the maximum likelihood estimation method via the EM algorithm and the continuous empirical characteristic function method, among others. A general problem of likelihood estimation for mixture models arises when the likelihood is unlimited, so the maximum likelihood estimate will no longer exist. According to Day (1969), this situation happens when we have a mixture of two normals with different means and variances. The author points out that in such situation, each sample point generates a singularity in the likelihood function, as well as any pair of sample points, triplets, quadruplets, and so forth, which are close enough. Kiefer and Wolfowitz (1956) also mention the inconsistency of the maximum likelihood estimator in this type of model. An illustration of the problem with the maximum likelihood estimate in this context can be seen in Wirjanto and Xu (2009).

3.3.4 EM Algorithm

The EM algorithm consists of an iterative procedure for the calculation of the maximum likelihood estimator in problems involving incomplete data. The general formulation with its basic properties and examples of applications was first presented by Dempster et al. (1977). There is an extensive bibliography with many examples, applications and extensions of the EM algorithm. Meng and Pedlow (1992) list numerous items related to the algorithm and considerable content about this subject can be found in Little and Rubin (1987), Tanner (1991, 1993), Schafer (1996), McLachlan and Krishnan (2008), among others. One of the applications of the EM algorithm is involving mixture models.

Let us assume that the population of interest P is composed of g subpopulations $P_1, P_2, \dots, P_g$.

$$P = \{P_1, P_2, \dots, P_g\}$$

The proportion of each population P_j is represented by π_j , with the total population $\sum_{j=1}^g \pi_j = 1$.

Let X indicate from which population an observation has been sampled, $X = j$ means that the observation belongs to population P_j . The distribution of X is discrete with support $1, 2, \dots, g$ and corresponding probabilities $\pi_1, \pi_2, \dots, \pi_g$, thus

$$X \sim \begin{pmatrix} 1 & 2 & \dots & g \\ \pi_1 & \pi_2 & \dots & \pi_g \end{pmatrix}.$$

Note that X is not observed, therefore it is a latent variable.

The density of Y in the total population P is then equal to

$$f(y) = \sum_{j=1}^g f(y|X = j)P(X = j) = \sum_{j=1}^g \pi_j f_j(y) \quad (3.7)$$

The distribution of Y is called a (finite) mixture with g components. The densities $f_1(y), \dots, f_g(y)$ often depend on vectors of (un-)known parameters $\theta_1, \dots, \theta_g$, the densities can be continuous, discrete, or a mixture of both types.

Considering the number of components g fixed (known) we will study how the parameters in a finite mixture distribution can be estimated using maximum likelihood.

Let $Y_1, \dots, Y_N$ be distributed as

$$\begin{aligned} Y_i &\sim \pi_1 f_{i1}(y_i) + \pi_2 f_{i2}(y_i) + \dots + \pi_g f_{ig}(y_i) \\ &= \sum_{j=1}^g \pi_j f_{ij}(y_i) \end{aligned} \quad (3.8)$$

where $f_{i1}(y_i), f_{i2}(y_i), \dots, f_{ig}(y_i)$ are the density functions of Y_i in the g components of the mixture. Often, we have that $f_{ij}(y_i) = f_j(y_i)$, for all j assuming that all Y_i follow the same distribution. Now, the index i allows the Y_i to have different distributions.

As mentioned above, the densities $f_{ij}(y_i)$ depend on the unknown parameters $\theta_1, \dots, \theta_g$, which are combined in the vector $\boldsymbol{\theta}$, this is often denoted as $f_{ij}(y_i|\boldsymbol{\theta})$. Moreover, let $\boldsymbol{\pi}$ be the vector of component probabilities $\boldsymbol{\pi}' = (\pi_1, \dots, \pi_g)$. The vector $\boldsymbol{\psi}$ contain all unknown parameters in the model, that is $\boldsymbol{\psi}' = (\boldsymbol{\pi}', \boldsymbol{\theta}')$.

Assuming that the observations are independent of each other, the likelihood function is equal to

$$L(\boldsymbol{\psi}|\mathbf{y}) = \prod_{i=1}^N \left\{ \sum_{j=1}^g \pi_j f_{ij}(y_i|\boldsymbol{\theta}) \right\}, \quad (3.9)$$

where $\mathbf{y}' = (y_1, \dots, y_N)$ is the vector of independent observed response values.

The corresponding log-likelihood function is given by

$$\ell(\boldsymbol{\psi}|\mathbf{y}) = \sum_{i=1}^N \ln \left\{ \sum_{j=1}^g \pi_j f_{ij}(y_i|\boldsymbol{\theta}) \right\}. \quad (3.10)$$

In general, maximizing $\ell(\boldsymbol{\psi}|\mathbf{y})$ with respect to $\boldsymbol{\psi}$ requires iterative numerical procedures. However, the maximization of this log-likelihood is far from straightforward because procedures based on the calculation of the first and second order derivatives can become numerically complex due to the sum of these densities in the denominator of the first order derivative. The EM algorithm, initially designed to handle problems with incomplete data, is a convenient alternative procedure to deal with problems involving mixture of distributions. In this context, the component associated with the mixture, here designated as the latent variable X , will be considered missing.

Define the indicators Z_{ij} , $i = 1, \dots, N$, $j = 1, \dots, g$, such that

$$Z_{ij} = \begin{cases} 1 & \text{if observation } i \text{ belongs to component } j \\ 0 & \text{otherwise.} \end{cases}$$

Then, we have that

$$P(Z_{ij} = 1) = \pi_j.$$

The joint density of Y_i and all component indicators Z_{ij} equals

$$\begin{aligned} f_i(y_i, Z_{i1} = z_{i1}, \dots, Z_{ig} = z_{ig}) \\ &= f_i(y_i | Z_{i1} = z_{i1}, \dots, Z_{ig} = z_{ig}) \times P(Z_{i1} = z_{i1}, \dots, Z_{ig} = z_{ig}) \\ &= \left\{ \prod_{j=1}^g [f_{ij}(y_i | \boldsymbol{\theta})]^{z_{ij}} \right\} \times \left\{ \prod_{j=1}^g \pi_j^{z_{ij}} \right\} = \prod_{j=1}^g [\pi_j f_{ij}(y_i | \boldsymbol{\theta})]^{z_{ij}}. \end{aligned} \quad (3.11)$$

The expression obtained in (3.11) is the result of the product of two factors, the first is the known density of our response Y and the second is the density of the multinomial distribution associated with the probability of the observation to belong to component j of the mixture.

Assuming again that distinct observations are independent of each other, the joint likelihood function for the observed measurements $\mathbf{y}$ and unobserved measurements $\mathbf{Z}$ is given by the following expression called complete data likelihood

$$L(\boldsymbol{\psi} | \mathbf{y}, \mathbf{Z}) = \prod_{i=1}^N \prod_{j=1}^g [\pi_j f_{ij}(y_i | \boldsymbol{\theta})]^{z_{ij}}. \quad (3.12)$$

The product of expression (3.12) refers to independent contribution and the second is the contribution of density to each observation in the data set.

The corresponding complete data log-likelihood function is given by

$$\ell(\boldsymbol{\psi} | \mathbf{y}, \mathbf{Z}) = \sum_{i=1}^N \sum_{j=1}^g z_{ij} \{ \ln \pi_j + \ln f_{ij}(y_i | \boldsymbol{\theta}) \}. \quad (3.13)$$

We now have a double summation and note that the $\boldsymbol{\pi}$ parameters are separate from the $\boldsymbol{\theta}$ parameters. This makes expression (3.13) computationally convenient.

Maximizing the complete data log-likelihood is simpler than maximizing the log-likelihood of the observed data, however the maximization of $\ell(\boldsymbol{\psi} | \mathbf{y}, \mathbf{Z})$ depends not only on the observed data $\mathbf{y}$, but also on the unobserved indicators z_{ij} . Therefore, this can now be used in practice to estimate the unknown parameter vector $\boldsymbol{\psi}$.

The maximization procedure of $\ell(\boldsymbol{\psi}|\mathbf{y}, \mathbf{Z})$ consists first of replacing the missing observations z_{ij} by their expected values, that is, we maximize the expected value with respect to the observations unobserved z_{ij} given the observations $\mathbf{y}$, therefore we maximize

$$E[\ell(\boldsymbol{\psi}|\mathbf{y}, \mathbf{Z})|\mathbf{y}].$$

While $\boldsymbol{\psi}$ is not known the calculation of conditional expectation is not possible, so the EM algorithm acts iteratively when we give the initial value $\boldsymbol{\psi}^{(1)}$ for $\boldsymbol{\psi}$. A series of estimates $\boldsymbol{\psi}^{(t)}$ is constructed, which eventually converge to the maximum likelihood estimate $\hat{\boldsymbol{\psi}}$ of $\boldsymbol{\psi}$. Once we have $\boldsymbol{\psi}^{(t)}$, the updated estimate $\boldsymbol{\psi}^{(t+1)}$ is obtained through a step E and a step M.

Step E is then based on the calculation of the following function

$$Q(\boldsymbol{\psi}|\boldsymbol{\psi}^{(t)}) = E[\ell(\boldsymbol{\psi}|\mathbf{y}, \mathbf{Z})|\mathbf{y}, \boldsymbol{\psi}^{(t)}]. \quad (3.14)$$

Given $\boldsymbol{\psi}^{(t)}$ the step E will be the calculation of the conditional expectation given $\mathbf{y}$, this will be used to calculate the expectation. The function $Q(\boldsymbol{\psi}|\boldsymbol{\psi}^{(t)})$ can be maximized with respect to the unknown vector $\boldsymbol{\psi}$ of parameters.

Step M consists of maximizing $Q(\boldsymbol{\psi}|\boldsymbol{\psi}^{(t)})$ with respect to $\boldsymbol{\psi}$ to obtain the updated estimate $\boldsymbol{\psi}^{(t+1)}$.

The procedure continues to iterate between step E and step M until convergence is achieved. The convergence is monitored in the log-likelihood between success steps when a change in log-likelihood in the iterative process is smaller than a pre-specified value ε , that is, until

$$|\ell(\boldsymbol{\psi}^{(t+1)}|\mathbf{y}) - \ell(\boldsymbol{\psi}^{(t)}|\mathbf{y})| < \varepsilon$$

for some small $\varepsilon > 0$. Once this criterion is satisfied the procedure is stopped and then the value that is obtained for $\boldsymbol{\psi}$ will be used as the maximum likelihood estimate for $\boldsymbol{\psi}$.

E step

The expected value of the complete log-likelihood of the data needs to be calculated conditionally to the observed values $\mathbf{y}$. Using the expression of the complete log-likelihood of the data we obtain for $Q(\boldsymbol{\psi}|\boldsymbol{\psi}^{(t)})$ the following expression

$$\begin{aligned}
 Q(\boldsymbol{\psi}|\boldsymbol{\psi}^{(t)}) &= E[\ell(\boldsymbol{\psi}|\mathbf{y}, \mathbf{Z})|\mathbf{y}, \boldsymbol{\psi}^{(t)}] \\
 &= E \left[\left\{ \sum_{i=1}^N \sum_{j=1}^g Z_{ij} [\ln \pi_j + \ln f_{ij}(y_i|\boldsymbol{\theta})] \right\} | \mathbf{y}, \boldsymbol{\psi}^{(t)} \right] \\
 &= \sum_{i=1}^N \sum_{j=1}^g E [Z_{ij} | \mathbf{y}, \boldsymbol{\psi}^{(t)}] [\ln \pi_j + \ln f_{ij}(y_i|\boldsymbol{\theta})] \tag{3.15}
 \end{aligned}$$

Since $\boldsymbol{\theta}$, $\boldsymbol{\pi}$ and $\mathbf{y}$ were considered fixed in the calculation of

$$E[\ell(\boldsymbol{\psi}|\mathbf{y}, \mathbf{Z})|\mathbf{y}, \boldsymbol{\psi}^{(t)}]$$

all we need to do is to calculate the expected value of Z_{ij} conditional to $\mathbf{y}$ and $\boldsymbol{\psi}$, where $\boldsymbol{\psi}$ is equal to $\boldsymbol{\psi}^{(t)}$. Therefore, step E requires only the calculation of

$$E [Z_{ij} | \mathbf{y}, \boldsymbol{\psi}^{(t)}] .$$

Since $\mathbf{Z}$ is an indicator variable then this conditional expectation is the probability that Z_{ij} is equal to 1 given y_i and the current value for $\boldsymbol{\psi}$. Therefore, using the Bayes rule we have that this probability equals the density of y_i given that the indicator $Z_{ij} = 1$ multiplied by the probability of the indicator is equal to 1 and divided by the marginal density of y_i . This calculation will be given below and denoted by the function $\pi_{ij}(\boldsymbol{\psi}^{(t)})$.

$$\begin{aligned}
 E [Z_{ij} | y_i, \boldsymbol{\psi}^{(t)}] &= P(Z_{ij} = 1 | y_i, \boldsymbol{\psi}^{(t)}) = \frac{f_i(y_i | Z_{ij} = 1) P(Z_{ij} = 1)}{f_i(y_i | \boldsymbol{\theta})} \Bigg|_{\boldsymbol{\psi}^{(t)}} \\
 &= \frac{\pi_j f_{ij}(y_i | \boldsymbol{\theta})}{\sum_j \pi_j f_{ij}(y_i | \boldsymbol{\theta})} \Bigg|_{\boldsymbol{\psi}^{(t)}} = \pi_{ij}(\boldsymbol{\psi}^{(t)}) \tag{3.16}
 \end{aligned}$$

The function $\pi_{ij}(\boldsymbol{\psi}^{(t)})$ is the posterior probability, that is, it is the probability of the observation i sampled from the j component of the mixture given the response value y_i for the observation i . The probability π_j will be called the prior probability for the observation i belonging to the j^{th} component of the mixture, it is the probability that an observation is sampled from the component j of the mixture without taking into account

the response value for this observation.

The conclusion we get here is that step E is simple and it is basically limited to the calculation of all posterior probabilities π_{ij} for all values i and j in our data set.

M step

The updated estimate $\boldsymbol{\psi}^{(t+1)}$ is obtained by maximizing

$$\begin{aligned}
 Q(\boldsymbol{\psi}|\boldsymbol{\psi}^{(t)}) &= \sum_{i=1}^N \sum_{j=1}^g \pi_{ij}(\boldsymbol{\psi}^{(t)}) [\ln \pi_j + \ln f_{ij}(y_i|\boldsymbol{\theta})] \\
 &= \sum_{i=1}^N \sum_{j=1}^g \pi_{ij}(\boldsymbol{\psi}^{(t)}) \ln \pi_j + \sum_{i=1}^N \sum_{j=1}^g \pi_{ij}(\boldsymbol{\psi}^{(t)}) \ln f_{ij}(y_i|\boldsymbol{\theta}) \\
 &= Q_1(\boldsymbol{\psi}|\boldsymbol{\psi}^{(t)}) + Q_2(\boldsymbol{\psi}|\boldsymbol{\psi}^{(t)})
 \end{aligned} \tag{3.17}$$

with respect to $\boldsymbol{\psi}' = (\boldsymbol{\pi}', \boldsymbol{\theta}')$.

We need to maximize the calculated Q function with respect to all elements in the unknown vector $\boldsymbol{\psi}$. Note that Q is a sum of two components, where the first component depends only on the π parameters and the second depends only on the θ parameters, so we can maximize Q by maximizing these two components separately.

The maximization of the first component with respect to π requires the maximization under the restriction $\sum_{j=1}^g \pi_j = 1$. In order to construct this restriction in the maximization procedure we replace the last probability π_g by 1 minus the sum of all other probabilities $(\pi_1, \dots, \pi_{g-1})$, so we need to maximize the following expression

$$\begin{aligned}
 Q_1(\boldsymbol{\psi}|\boldsymbol{\psi}^{(t)}) &= \sum_{i=1}^N \sum_{j=1}^g \pi_{ij}(\boldsymbol{\psi}^{(t)}) \ln \pi_j = \sum_{i=1}^N \sum_{j=1}^{g-1} \pi_{ij}(\boldsymbol{\psi}^{(t)}) \ln \pi_j + \\
 &\quad + \sum_{i=1}^N \pi_{ig}(\boldsymbol{\psi}^{(t)}) \ln \left[1 - \sum_{j=1}^{g-1} \pi_j \right]
 \end{aligned} \tag{3.18}$$

with respect to $\pi_1, \dots, \pi_{g-1}$.

Equating the first order derivatives to zero we have

$$\begin{aligned}
\frac{\partial Q_1(\boldsymbol{\psi}|\boldsymbol{\psi}^{(t)})}{\partial \pi_j} = 0 &\Leftrightarrow \sum_{i=1}^N \frac{\pi_{ij}(\boldsymbol{\psi}^{(t)})}{\pi_j^{(t+1)}} = \frac{\pi_{ig}(\boldsymbol{\psi}^{(t)})}{\pi_g^{(t+1)}} \Leftrightarrow \\
&\Leftrightarrow \frac{\pi_j^{(t+1)}}{\pi_g^{(t+1)}} = \frac{\sum_{i=1}^N \pi_{ij}(\boldsymbol{\psi}^{(t)})}{\sum_{i=1}^N \pi_{ig}(\boldsymbol{\psi}^{(t)})}.
\end{aligned} \tag{3.19}$$

From (3.19) we obtain the following

$$\begin{aligned}
1 = \sum_{j=1}^g \pi_j^{(t+1)} &= \sum_{j=1}^g \frac{\pi_g^{(t+1)} \sum_{i=1}^N \pi_{ij}(\boldsymbol{\psi}^{(t)})}{\sum_{i=1}^N \pi_{ig}(\boldsymbol{\psi}^{(t)})} \\
&= \frac{\pi_g^{(t+1)} \sum_{i=1}^N \sum_{j=1}^g \pi_{ij}(\boldsymbol{\psi}^{(t)})}{\sum_{i=1}^N \pi_{ig}(\boldsymbol{\psi}^{(t)})} = \frac{N \pi_g^{(t+1)}}{\sum_{i=1}^N \pi_{ig}(\boldsymbol{\psi}^{(t)})}.
\end{aligned} \tag{3.20}$$

Therefore, $\pi_g^{(t+1)}$ is given by

$$\pi_g^{(t+1)} = \frac{\sum_{i=1}^N \pi_{ig}(\boldsymbol{\psi}^{(t)})}{N}. \tag{3.21}$$

It follows also that all $\pi_j^{(t+1)}$ are given by

$$\pi_j^{(t+1)} = \frac{\sum_{i=1}^N \pi_{ij}(\boldsymbol{\psi}^{(t)})}{N}. \tag{3.22}$$

The expression (3.22) says that the prior probability to belong to g components is the average of all posterior probabilities to belong to that same components. So far the maximization of the first part of the sum in (3.17) has been completed.

The other part of step M is the maximization of the second part of the sum

in (3.17), that is, the maximization with respect to the elements in $\boldsymbol{\theta}$ of the following expression

$$Q_2(\boldsymbol{\psi}|\boldsymbol{\psi}^{(t)}) = \sum_{i=1}^N \sum_{j=1}^g \pi_{ij}(\boldsymbol{\psi}^{(t)}) \ln f_{ij}(y_i|\boldsymbol{\theta}). \quad (3.23)$$

In some simple cases this maximization can be done analytically but in most of the cases this is not possible and numerical maximization as Newton-Raphson is required. When this is necessary the EM algorithm is double iterative and can have serious consequences in computational time.

3.3.5 Mixture model for the sum of chicken weights: Cross-sectional case

Considering how the data of chicken weights were obtained, we can think of the following mechanism without considering the effect of the treatment preliminarily. A number of individuals formed by samples from two subpopulations of chickens, males and females, were allocated to the box. Since there was no gender identification of the individuals within box, one can not consider the sex factor to be included in a model, however, the information that the data come from two subgroups should not be left aside in modeling. In addition, it is not common to have the data of individual weights in an experiment of this size, the usual in this case is to have the average or sum of the individual weights of the birds for each box. Thus, we can think of a model for the sum of the weights of the birds that takes into account the mixture of the different distributions of males and females. A model constructed for the sum of chick weights inside box will be presented, which takes into account the mixture of the distributions of weights of males and females.

The model will be presented in two situations, the first being a more general case with mean and variance specific for each sex, the second being a simplified version with different means for each sex and the same variance.

3.3.5.1 Gender-specific mean and variance

Let Y_{ij} be the weight of chicken j in box i . Without loss of generality, we assume that all boxes contain the same number of chicken. Let n be the total number of chickens in each box. It is assumed that Y_{ij} is normally distributed with gender-specific mean and variance:

$$Y_{ij} \sim \begin{cases} N(\mu_1, \sigma_1^2) & \text{if female} \\ N(\mu_2, \sigma_2^2) & \text{if male} \end{cases}.$$

Furthermore, let p be the proportion of female chickens in the total population, and let N_i be the number of female chickens in box i . The variable N_i then follows a

Binomial distribution with n trials and success probability p , that is

$$N_i \sim \text{Binomial}(n, p).$$

Assuming all chickens to be independent, we have that the distribution of the total weight $Y_i = \sum_{j=1}^n Y_{ij}$ of all chickens in a box i with N_i female chickens (and $n - N_i$ male chickens) equals

$$Y_i | N_i \sim N[N_i \mu_1 + (n - N_i) \mu_2, N_i \sigma_1^2 + (n - N_i) \sigma_2^2].$$

The marginal density of Y_i then equals

$$\begin{aligned} f(y_i) &= \sum_{x=1}^n P(N_i = x) f(y_i | N_i = x) \\ &= \sum_{x=1}^n \frac{\binom{n}{x} p^x (1-p)^{n-x} \exp \left\{ -\frac{[y_i - x\mu_1 - (n-x)\mu_2]^2}{2[x\sigma_1^2 + (n-x)\sigma_2^2]} \right\}}{\sqrt{2\pi[x\sigma_1^2 + (n-x)\sigma_2^2]}} \end{aligned} \quad (3.24)$$

3.3.5.2 Different means and common variance

In this situation the weight of chicken j in box i , Y_{ij} , is assumed to be normally distributed with different means and common variance:

$$Y_{ij} \sim \begin{cases} N(\mu_1, \sigma^2) & \text{if female} \\ N(\mu_2, \sigma^2) & \text{if male} \end{cases}.$$

The distribution of the total weight $Y_i = \sum_{j=1}^n Y_{ij}$ of all the chickens in the box i with N_i female chickens (and $n - N_i$ male chickens) equals

$$Y_i | N_i \sim N[N_i \mu_1 + (n - N_i) \mu_2, N \sigma^2]$$

The marginal density of Y_i then equals

$$\begin{aligned} f(y_i) &= \sum_{x=1}^n P(N_i = x) f(y_i | N_i = x) \\ &= \sum_{x=1}^n \frac{\binom{n}{x} p^x (1-p)^{n-x} \exp \left\{ -\frac{[y_i - x\mu_1 - (n-x)\mu_2]^2}{2n\sigma^2} \right\}}{\sqrt{2\pi n\sigma^2}} \end{aligned} \quad (3.25)$$

3.3.6 Methods of estimation

All parameters in the models (3.24) and (3.25) can be estimated using maximum likelihood estimation, i.e., maximizing the likelihood for a data set where observations for N boxes are available:

$$L(\mu_1, \mu_2, \sigma_1, \sigma_2, p) = \prod_{i=1}^N f(y_i)$$

3.3.7 Bayesian approach

In Bayesian inference the different degrees of uncertainty about quantities of interest are represented by probabilistic models. Considering the quantity of interest θ , which takes values in a space Θ , we have that the true value of θ is unknown and the interest is to obtain information about θ .

We can summarize some information about θ using a prior distribution called $p(\theta)$. Using a prior distribution together with a sample, y , of a random quantity related to θ expressed by its likelihood function $p(y|\theta)$, we apply the theorem Bayes and we obtain the posteriori density

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{p(y)} \quad (3.26)$$

where $p(y) = \sum_{\theta} p(\theta)p(y|\theta)$, the sum is over all possible values of θ , in the case of continuous θ we have $p(y) = \int p(\theta)p(y|\theta)d\theta$.

An equivalent form of Bayes' Theorem is given by

$$p(\theta|y) \propto p(\theta)p(y|\theta). \quad (3.27)$$

Note that the factor $p(y)$ of the expression (3.26), which does not depend of θ , was omitted in (3.27) and functions as a normalizing constant of $p(\theta|y)$.

The technical essence of Bayesian methods is in the expressions (3.26) and (3.27), the main task of Bayesian Inference is to fit a model of probability to a set of data and summarize the results by an appropriate probability distribution $p(\theta|y)$.

In many applications the obtaining of the posterior distribution requires the integration of high-dimensional functions and their calculation can be difficult or even impossible. There are some algorithms used to approximate the posterior distribution computationally and thus allow the realization of inferences. A review of computational methods can be found, for example, in Smith (1991) and Tanner (1991).

The method widely used to approximate a target distribution, which in the Bayesian context is the posterior distribution $p(\theta|y)$, is the Markov chain Monte Carlo Method (MCMC).

A sequence $\{\theta^{(i)}\}$ is called Markov Chain (MC) if

$$p(\theta^{(i+1)}|\theta^{(1)}, \dots, \theta^{(i)}) = p(\theta^{(i+1)}|\theta^{(i)}).$$

For a given actual value of chain $\theta^{(i)}$, the next value $\theta^{(i+1)}$ depends only of $\theta^{(i)}$. Under certain regularity conditions the MC converges to a stationary distribution and the probability of realization of θ depends of the current value of the chain. A common practice used to remove the effects of the initial values sampled is known as burn-in, consisting in discarding a certain amount of these initial values. One strategy to reduce autocorrelation is to decrease the production and store only sufficiently spaced points after the burn-in period. This strategy is called thinning of the sequence and in addition to reducing the autocorrelation between the samples also alleviates the computational cost in the construction of the chain.

Considering only the sufficiently spaced values we can obtain samples of uncorrelated values of this stationary distribution. From these samples we can obtain any measure regarding this stationary distribution, such as estimating the probability density, calculating percentiles, means, variances, etc.

In practice, more than one chain with different initial values between them are constructed to approximate some target distribution that one wants. The convergence monitoring is done by means of methods proposed for this purpose. Among the authors who proposed these methods are Fosdick (1959), Ripley (1987), Tanner and Wong (1987), Gelfand and Smith (1990), Gelman and Rubin (1992), among others. The method proposed by Gelman and Rubin (1992) consists in determining, after the burn-in period, the variance of the simulations of each chain, to calculate the mean of these variances and to compare it with the variance of all the chains together. The variance considering all the chains is divided by the mean variance within the chain, then the square root of that ratio is calculated. This statistic is called the “potential scale reduction factor”. According to Congdon (2003) values of this statistic below 1.2 indicate convergence, while Brooks et al. (2011) recommend stopping the iterations when the “potential scale reduction factor” is smaller than 1.1 and advise the use of other convergence diagnoses together with this statistic.

3.3.8 Simulated case study using the classical approach

In order to evaluate the estimates of the model parameters designed for the sum of chicken weights, data were generated and used the maximum likelihood method to estimate the parameters in two different cases, gender specific mean and variance and different means with common variance. The values of the parameters were chosen based on observations of the individual weights of chickens at 42 days. The E.M. algorithm was used to estimate the parameters of the mixture of two Gaussian components with homogeneous

variances for each treatment. Then, the average of the estimates of each parameter was calculated and this result was considered as the value of the parameter in the generation of the simulated data. In the case of heterogeneous variances were considered the average of the estimates obtained by the Bayesian approach. The sample size considered in the simulated data generation was based on the balanced sample size of the real experiment, 5 boxes with 46 individuals per box. The number of individuals per box was fixed to avoid interference of the unbalance effect in this assessment of the estimates.

Next, the mechanism of data generation in the different cases will be described.

Let Y_{ij} be the weight of chicken j in box i , in the case of the homogeneous variances the data of individual weights were generated considering Y_{ij} as

$$Y_{ij} \sim \begin{cases} N(\mu_1 = 2.481, \sigma = 0.1935) & \text{if female} \\ N(\mu_2 = 3.0358, \sigma = 0.1935) & \text{if male.} \end{cases}$$

After the generation of the individual data Y_{ij} , the data were assigned to the boxes considering the number of female chickens N_i in box i as

$$N_i \sim \text{Binomial}(n, p = 0.56),$$

where n represents the number of individuals within box i .

Then the sum of the individual data generated, Y_{ij} , was calculated within each box i . This sum represents the total weight $Y_i = \sum_{j=1}^n Y_{ij}$ of all chickens in a box i with N_i female chickens and $(n - N_i)$ male chickens.

In the case with the heterogeneous variances the data of individual weights were generated considering Y_{ij} as

$$Y_{ij} \sim \begin{cases} N(\mu_1 = 2.4633, \sigma_1 = 0.1897) & \text{if female} \\ N(\mu_2 = 2.9933, \sigma_2 = 0.2247) & \text{if male.} \end{cases}$$

The number of female chickens N_i in box i , in this case, were generated following the distribution

$$N_i \sim \text{Binomial}(n, p = 0.5).$$

Different sets of initial values were used in the optimization algorithm based on the robust method of Nelder and Mead (1965). The procedure consisted of varying different initial values for each model parameter in the two different cases, homogeneous and heterogeneous variances, while maintaining the other initial values of the other parameters fixed in the true values. The obtained estimates were compared to the values of the parameters, also the value of the likelihood of each set of estimates was calculated and the respective standard error was calculated.

3.3.9 Simulated case study using the Bayesian approach

In order to evaluate the estimates of the parameters of the model designed for the sum of the chicken weights, data were generated and the Bayesian approach was used to estimate the parameters in two different cases, one case with different means and the same variances and the another with gender-specific means and variances. The values of the parameters were chosen based on the estimates of individual weights of chickens at 42 days. In the case of the homogeneous variances, we considered the average of the maximum likelihood estimates of the mixture model of two homogeneous Gaussian components obtained through the EM algorithm. In the case of heterogeneous variances, the averages of the Bayesian estimates of the mixture model of two heterogeneous Gaussian components were considered. The sample size considered in the simulated data generation was based on the real data, 5 boxes with 46 individuals per box.

The data generating mechanism was the same used for the simulated case study in the Classical approach cited in the previous section (section 3.3.8). Based on the marginal densities (3.24) and (3.25) of the total weight distribution Y_i of all chickens in a box i with N_i female chickens (and $n - N_i$ male chickens) the Bayesian approach was used first considering non-informative prior (uniform) for all the parameters of the models and after we considered informative priors for the parameters p , μ_1 and μ_2 . In each of the cases studied were used 3 chains with different initial values each. To avoid the divergence of the chains the following parameter restrictions were used, $0 \leq p \leq 1$; $0 \leq \mu_1 \leq 4.0$; $0 \leq \mu_2 \leq 4.0$; $0 \leq \sigma_1 \leq 1.0$ and $0 \leq \sigma_2 \leq 1.0$. Such restrictions were based on the formulation of the model of the sum of the weights and in the estimates of the mixture model for the individual weights data.

According to recommendations of Brooks et al. (2011) the first half of the simulated sequences were discarded to avoid the dependence of the simulations with the initial values. In order to decrease the computational cost, was considered a thin equal to 10 and the convergence monitoring was performed comparing between and within of the chains by the statistic “potential scale reduction factor”.

Appendix C of this work is the code used in R (R Core Team, 2017) to the results of this chapter.

3.4 Results

The results are organized in two main sections, the first one referring to the analyzes for the individual data of chicken weights and the second one referring to the analyzes for the sum of chicken weights. In each of these sections there is the division between the classical and Bayesian approach.

3.4.1 Analysis of individual weights

We used the individual weights data of broilers at 42 days on the scale in kilograms, since the individual measurements were available only at 7 and 42 days, we chose to work with the data at 42 days because of the greater distinction between the weights of males and females during the experimental period. The models used for the data of individual weights refer to the mixture model of two Gaussian components, being considered the case of homogeneous variances in the classical and Bayesian approach. Additionally, the parameters of the mixture model with heterogeneous variances were estimated using the Bayesian approach.

3.4.1.1 Analysis of individual weights using the classical approach

In order to verify the adequacy of the Gaussian mixture model with two components to the data in relation to the adjustment of only one Gaussian component, in the case of the homogeneous variances, the parameters in both models were estimated by the maximum likelihood method using the algorithm EM. Subsequently, was performed the likelihood ratio test between the models adjusted for each treatment. The results are presented in Table (3.1), where we can observe that the likelihood ratio test favors the model with two Gaussian components in all treatments.

Table 3.1 – Estimates of the parameters for the models with a normal distribution and the mixture of two normal distributions with likelihood ratio test considering the weight of chickens at 42 days of age

Treatment	$N(\mu, \sigma^2)$			$\pi N(\mu_1, \sigma^2) + (1 - \pi)N(\mu_2, \sigma^2)$					χ^2 (p -value)
	$\hat{\mu}$	$\hat{\sigma}$	-2Log L1	$\hat{\pi}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	-2Log L2	
T1	2.66	0.33	140.2	0.53	2.41	2.94	0.20	123.5	16.7 (< 0.01)
T2	2.69	0.34	150.4	0.53	2.42	2.98	0.19	122.9	27.5 (< 0.01)
T3	2.73	0.34	147.2	0.58	2.49	3.06	0.19	120.8	26.4 (< 0.01)
T4	2.71	0.33	141.5	0.60	2.49	3.05	0.19	113.1	28.4 (< 0.01)
T5	2.74	0.34	158.6	0.53	2.47	3.04	0.20	135.2	23.4 (< 0.01)
T6	2.82	0.34	143.2	0.59	2.59	3.15	0.20	122.3	20.9 (< 0.01)

Note that estimates of the proportion parameter, π , were close to 0.5 for treatments T1, T2 and T5, while for the other treatments were close to 0.6. The results of these estimates agree with the information about the proportion of males and females from the data, since at the beginning of the experiment were allocated 50% of each sex to the boxes and there was low mortality.

Calculating the Mahalanobis distance for each treatment we have $\Delta_{T1} = 2.65$, $\Delta_{T2} = 2.94$, $\Delta_{T3} = 3.0$, $\Delta_{T4} = 2.95$, $\Delta_{T5} = 2.85$ and $\Delta_{T6} = 2.80$. Note that the calculated distances are close to each other and the modality of the mixture depends of π , since the Mahalanobis distance is greater than 2 for every treatment.

Figures (3.2) and (3.3) show the histograms of the data with the graphical representations of the estimated densities, in the cases of one and two Gaussian components, respectively. Observe in these figures that the fit of the mixture model with two homogeneous Gaussian components was more appropriate to the data than the model with one Gaussian component.

The inferences to the mixture model with two heterogeneous Gaussian components do not follow the standard likelihood theory. As pointed out by Day (1969), such model generates a singularity in the likelihood function and leads to the inconsistency of the maximum likelihood estimator. Sometimes the estimates can be obtained for this type of model, but convergence leads to a local maximum of the likelihood.

In the next section will be shown the results of the estimates for the mixture model with two heterogeneous Gaussian components using the Bayesian approach.

Figure 3.2 – Chicken weights histograms at 42 days of age with the density plot of the normal distribution adjusted for each treatment separately

Figure 3.3 – Histograms of chick weights at 42 days of age with the graph of the density of the two gaussian mixture adjusted for each treatment separately

3.4.1.2 Analysis of the individual weights using the Bayesian approach

In this section we present the modeling results using the Bayesian approach for of the individual weights data of the chickens at 42 days. For each mixture model with homogeneous and heterogeneous Gaussian components were considered 3 chains with different initial values for each chain. We used the standard practice of burn-in recommended by Brooks et. al (2011), so the first half of the simulated sequence was discarded in both models to decrease the influence of the initial values on the target distribution.

We attributed restrictions for the parameters according as the parametric space of each of them. We used Improper Uniform Prior for all parameters to avoid interference of external information to data on posterior distribution.

Convergence was monitored by observing the “potential scale reduction factor” (GELMAN; RUBIN, 1992) and the Monte Carlo standard errors for all parameters in both models. We used 6000 iterations for convergence with thin equals to 10 to decrease computational time.

Tables (3.2) and (3.3) present the estimates of the parameters of the mixture models with two homogeneous and heterogeneous Gaussian components, respectively. In these tables we find the Monte Carlo standard error (MCSE), the standard deviation (SD), the median and the credibility interval with 95% probability (95% CrI) for each estimated parameter. In both models, the number of iterations was satisfactory, because in all treatments the MCSE was reasonably small and the “potential scale reduction factor” ranged from 1 to 1.03.

Note in Table (3.2) that the values of the estimates were similar to those obtained using the classical approach (Table 3.1). In general, the low standard deviation and the median close to value of the mean estimate give evidence of the absence of atypical observations in each treatment, that is, the sample did not present observations that would cause distortions in the parameter estimates. The narrow credibility interval for each model parameter helps confirm the good fit of the mixture model with homogeneous Gaussian components to the data.

When adjusting the mixture model with heterogeneous Gaussian components, where the results are found in Table (3.3), in general, the standard deviations and the credibility intervals were slightly higher than those obtained in the fit of the homogeneous mixture model. Additionally, the values of the medians remained close to values of the mean estimates of the parameters.

Note also, the estimates for the proportion parameter were moderately different. In cases where the values of the estimates of σ_1 and σ_2 were more distant from each other we observed a greater difference between the estimated values of the proportion in relation to the estimates of the parameters of proportion of the model with a single σ . This leads us to suppose that there may be a relation between the proportion parameter and the variances.

The estimates of the means, μ_1 and μ_2 , were close to the estimates of the homogeneous mixture model.

Table 3.2 – Bayesian estimates of the mixture model parameters with Gaussian components of individual weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$). Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and credibility interval with 95% probability for each model parameter

Treatment		$\hat{\pi}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$
T1	mean	0.5377	2.4209	2.9305	0.2114
	MCSE	0.0024	0.0018	0.0025	0.0014
	SD	0.0711	0.0510	0.0464	0.0243
	median	0.5400	2.4200	2.9400	0.2100
	95% CrI	(0.43 – 0.66)	(2.36 – 2.5)	(2.84 – 2.99)	(0.18 – 0.28)
T2	mean	0.5293	2.4253	2.9791	0.1963
	MCSE	0.0017	0.0009	0.0009	0.0005
	SD	0.0438	0.0250	0.0270	0.0136
	median	0.5300	2.4200	2.9800	0.2000
	95% CrI	(0.45 – 0.62)	(2.38 – 2.48)	(2.93 – 3.03)	(0.17 – 0.22)
T3	mean	0.5870	2.4930	3.0575	0.1953
	MCSE	0.0014	0.0008	0.0010	0.0005
	SD	0.0418	0.0231	0.0282	0.0133
	median	0.5900	2.4900	3.0600	0.1900
	95% CrI	(0.5 – 0.67)	(2.45 – 2.54)	(3 – 3.11)	(0.17 – 0.22)
T4	mean	0.6046	2.4948	3.0480	0.1930
	MCSE	0.0015	0.0009	0.0009	0.0005
	SD	0.0434	0.0218	0.0282	0.0134
	median	0.6100	2.4900	3.0500	0.1900
	95% CrI	(0.52 – 0.69)	(2.45 – 2.54)	(2.99 – 3.1)	(0.17 – 0.22)
T5	mean	0.5349	2.4781	3.0383	0.2027
	MCSE	0.0015	0.0009	0.0010	0.0006
	SD	0.0448	0.0274	0.0287	0.0161
	median	0.5400	2.4800	3.0400	0.2000
	95% CrI	(0.46 – 0.61)	(2.43 – 2.53)	(2.99 – 3.09)	(0.18 – 0.24)
T6	mean	0.5945	2.5946	3.1437	0.2038
	MCSE	0.0015	0.0010	0.0010	0.0005
	SD	0.0461	0.0259	0.0311	0.0156
	median	0.5900	2.5900	3.1400	0.2000
	95% CrI	(0.51 – 0.69)	(2.55 – 2.65)	(3.08 – 3.21)	(0.18 – 0.24)

Table 3.3 – Bayesian estimates of the mixture model parameters with Gaussian components of individual weights of chickens with heterogeneous variances (σ_1 and σ_2). Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and credibility interval with 95% probability for each model parameter

Treat.		$\hat{\pi}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$
T1	mean	0.2858	2.3250	2.7917	0.1292	0.2874
	MCSE	0.0038	0.0015	0.0022	0.0012	0.0014
	SD	0.1090	0.0430	0.0615	0.0351	0.0341
	median	0.2700	2.3200	2.7800	0.1200	0.2900
	95% CrI	(0.13 – 0.53)	(2.26 – 2.42)	(2.7 – 2.93)	(0.07 – 0.21)	(0.2 – 0.34)
T2	mean	0.4288	2.3814	2.9164	0.1689	0.2355
	MCSE	0.0049	0.0021	0.0027	0.0015	0.0014
	SD	0.1305	0.0585	0.0759	0.0392	0.0432
	median	0.4400	2.3800	2.9200	0.1700	0.2300
	95% CrI	(0.18 – 0.66)	(2.28 – 2.5)	(2.77 – 3.04)	(0.09 – 0.25)	(0.16 – 0.32)
T3	mean	0.6711	2.5403	3.1071	0.2293	0.1567
	MCSE	0.0025	0.0014	0.0015	0.0010	0.0011
	SD	0.0737	0.0413	0.0448	0.0292	0.0312
	median	0.6700	2.5400	3.1100	0.2300	0.1500
	95% CrI	(0.54 – 0.82)	(2.46 – 2.63)	(3.02 – 3.18)	(0.18 – 0.3)	(0.1 – 0.22)
T4	mean	0.5806	2.4899	3.0290	0.1903	0.2046
	MCSE	0.0030	0.0012	0.0022	0.0008	0.0012
	SD	0.0896	0.0368	0.0628	0.0234	0.0357
	median	0.5900	2.4900	3.0400	0.1900	0.2000
	95% CrI	(0.39 – 0.73)	(2.42 – 2.56)	(2.89 – 3.13)	(0.15 – 0.24)	(0.15 – 0.29)
T5	mean	0.4914	2.4643	3.0084	0.1944	0.2229
	MCSE	0.0043	0.0017	0.0027	0.0013	0.0017
	SD	0.1163	0.0485	0.0711	0.0346	0.0444
	median	0.5100	2.4600	3.0200	0.1900	0.2100
	95% CrI	(0.19 – 0.68)	(2.37 – 2.56)	(2.82 – 3.11)	(0.12 – 0.26)	(0.16 – 0.34)
T6	mean	0.5292	2.5861	3.0968	0.1952	0.2289
	MCSE	0.0072	0.0020	0.0049	0.0016	0.0027
	SD	0.1668	0.0526	0.1095	0.0412	0.0615
	median	0.5800	2.5900	3.1400	0.2000	0.2100
	95% CrI	(0.17 – 0.76)	(2.49 – 2.68)	(2.87 – 3.24)	(0.11 – 0.27)	(0.15 – 0.35)

Note in Figure (3.4) that the fit of the mixture model with two heterogeneous Gaussian components was quite reasonable.

Distribution and Estimated Density for Treatment 1

(a)

Distribution and Estimated Density for Treatment 2

(b)

Distribution and Estimated Density for Treatment 3

(c)

Distribution and Estimated Density for Treatment 4

(d)

Distribution and Estimated Density for Treatment 5

(e)

Distribution and Estimated Density for Treatment 6

(f)

Figure 3.4 – Histograms of chick weights at 42 days of age with the graph of the density of the two gaussian mixture adjusted for each treatment separately

3.4.2 Analysis of the sum of the weights of chickens

In this section we present the results of the application of the mixture model of the sum of chicken weights. First, the model was used in simulated data and then it was applied to the real data, both in the classical and Bayesian approaches. The simulated data were generated based on information from the real data considering 5 boxes and 46 individuals per box. The data generation mechanism is described in section 3.3.8 for the classical approach and section 3.3.9 for the Bayesian approach. In the Bayesian approach, we first use non-informative (uniform) priors for all the parameters of the models (3.24) and (3.25), after based on prior knowledge about the data we use informative priors for the parameters p , μ_1 and μ_2 . Since p is the Binomial distribution parameter, we use a Beta distribution as prior for p , since its support is in agreement with the parametric space of p . With the intention of concentrating 95% of its mass in the interval $(0.3, 0.7)$, we consider a distribution $\text{Beta}(12, 12)$. For the parameters μ_1 and μ_2 we used the priors

$$\mu_1 \sim N(2.5, 0.25) \quad \text{for female}$$

and

$$\mu_2 \sim N(3.0, 0.25) \quad \text{for male,}$$

where the prior for μ_1 concentrates 95% of its mass in the interval $(2.0, 3.0)$, whereas a prior for μ_2 concentrates 95% of its mass in the interval $(2.5, 3.5)$.

3.4.2.1 Simulated case study of the sum of chicken weights using the classical approach

In this section we present the estimates of the model parameters of the sum of the chicken weights for the simulated data. The likelihood value and the standard error estimate are also presented. We used the optimization algorithm based on the robust method of Nelder and Mead (1965) to obtain the estimates. In order for the search region of the optimization algorithm to be the real line, we use a reparameterization in the model for the parameters of the proportion (p) and scale (σ) given by

$$p = \frac{1}{1 + e^{-\theta}} \quad \text{and} \quad \sigma = e^{\theta^*}.$$

The set of the true values of the parameters used in the data simulation are highlighted in each table. Different initial values were assigned to the proportion parameter (p), while the initial values of the other parameters were fixed at the true values, the results are presented in Table (3.4). We follow the same idea for the parameters μ_1 , μ_2 and σ , where the results are presented in Tables (3.5), (3.6) and (3.7), respectively, for the model with a single σ . For the model with two σ 's different the results for the parameters p , μ_1 ,

μ_2 , σ_1 and σ_2 are given in Tables (3.8), (3.9), (3.10), (3.11) and (3.12), respectively.

Table 3.4 – Estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_p
1	initial values	0.40	2.4810	3.0358	0.1935		
2	at convergence	0.67	2.5071	3.1735	0.0269	-10.2041	0.1357
3	initial values	0.50	2.4810	3.0358	0.1935		
4	at convergence	0.78	2.5824	3.2386	0.0210	-10.01	0.1588
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.63	2.4826	3.1386	0.0210	-10.0338	0.1363
7	initial values	0.60	2.4810	3.0358	0.1935		
8	at convergence	0.71	2.5259	3.2197	0.0224	-10.0432	0.1452
9	initial values	0.70	2.4810	3.0358	0.1935		
10	at convergence	0.98	2.6914	4.7934	0.0575	-8.6439	0.5047

Table 3.5 – Estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_{μ_1}
1	initial values	0.56	1.5000	3.0358	0.1935		
2	at convergence	0.82	2.6022	3.2925	0.0300	-10.2691	0.0105
3	initial values	0.56	2.0000	3.0358	0.1935		
4	at convergence	0.90	2.6300	3.5684	0.0215	-8.3472	0.0034
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.63	2.4826	3.1386	0.0210	-10.0338	0.0092
7	initial values	0.56	2.8000	3.0358	0.1935		
8	at convergence	0.77	2.5710	3.2655	0.0223	-10.0075	0.0066
9	initial values	0.56	2.8500	3.0358	0.1935		
10	at convergence	0.99	2.6934	4.2763	0.0411	-8.9117	0.0037

Table 3.6 – Estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_{μ_2}
1	initial values	0.56	2.4810	2.5800	0.1935		
2	at convergence	0.23	3.2650	2.5711	0.0224	-10.0039	0.0067
3	initial values	0.56	2.4810	2.8000	0.1935		
4	at convergence	0.73	2.5406	3.2376	0.0364	-10.4994	0.0431
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.63	2.4826	3.1386	0.0210	-10.0338	0.0144
7	initial values	0.56	2.4810	3.5000	0.1935		
8	at convergence	0.65	2.4969	3.1527	0.0209	-10.0342	0.0149
9	initial values	0.56	2.4810	3.7000	0.1935		
10	at convergence	0.85	2.5891	3.5284	0.0219	-8.4771	0.0255

Table 3.7 – Estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_{σ}
1	initial values	0.56	2.4810	3.0358	0.0800		
2	at convergence	0.60	2.4971	3.0679	0.0147	-9.1396	0.3205
3	initial values	0.56	2.4810	3.0358	0.1000		
4	at convergence	0.60	2.4971	3.0679	0.0147	-9.1396	0.3205
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.63	2.4826	3.1386	0.0210	-10.0338	0.3631
7	initial values	0.56	2.4810	3.0358	0.2200		
8	at convergence	0.62	2.5095	3.0803	0.0147	-9.1414	0.3205
9	initial values	0.56	2.4810	3.0358	0.2400		
10	at convergence	0.67	2.4957	3.1895	0.0224	-10.0799	0.3617

Table 3.8 – Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_p
1	initial values	0.42	2.4633	2.9933	0.18970	0.2247		
2	at convergence	0.72	2.5568	3.1778	0.00248	0.0180	-5.6505	0.1464
3	initial values	0.45	2.4633	2.9933	0.18970	0.2247		
4	at convergence	0.86	2.6468	3.2653	0.00430	0.3693	-9.7471	1.76
5	initial values	0.50	2.4633	2.9933	0.18970	0.2247		
6	at convergence	0.92	2.6569	3.6198	0.00000	0.0665	-6.208	0.2455
7	initial values	0.60	2.4633	2.9933	0.18970	0.2247		
8	at convergence	0.88	2.6126	3.5965	0.02212	0.0599	-7.9028	0.2048
9	initial values	0.65	2.4633	2.9933	0.18970	0.2247		
10	at convergence	0.92	2.6560	3.6319	0.01593	0.0716	-7.0445	0.2477

Table 3.9 – Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{μ_1}
1	initial values	0.50	2.3000	2.9933	0.18970	0.22470		
2	at convergence	0.90	2.6423	3.5050	0.00008	0.05380	-6.7107	0.0023
3	initial values	0.50	2.4000	2.9933	0.18970	0.22470		
4	at convergence	0.89	2.6615	3.2811	0.04659	0.46339	-9.7462	0.0195
5	initial values	0.50	2.4633	2.9933	0.18970	0.22470		
6	at convergence	0.92	2.6569	3.6198	0.00000	0.06653	-6.208	0.0018
7	initial values	0.50	2.4700	2.9933	0.18970	0.22470		
8	at convergence	0.86	2.6622	3.1445	0.02204	0.49251	-9.7562	0.0709
9	initial values	0.50	2.4750	2.9933	0.18970	0.22470		
10	at convergence	0.92	2.6569	3.6198	0.00003	0.06653	-6.208	0.0018

Table 3.10 – Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{μ_2}
1	initial values	0.50	2.4633	2.9200	0.18970	0.22470		
2	at convergence	0.86	2.6353	3.3152	0.03943	0.21682	-9.7569	0.7545
3	initial values	0.50	2.4633	2.9500	0.18970	0.22470		
4	at convergence	0.92	2.6569	3.6198	0.00000	0.06653	-6.208	0.0264
5	initial values	0.50	2.4633	2.9933	0.18970	0.22470		
6	at convergence	0.92	2.6569	3.6198	0.00000	0.06653	-6.208	0.0264
7	initial values	0.50	2.4633	3.0000	0.18970	0.22470		
8	at convergence	0.92	2.6568	3.6221	0.00587	0.06723	-6.3639	0.029
9	initial values	0.50	2.4633	3.1000	0.18970	0.22470		
10	at convergence	0.87	2.6620	3.2006	0.02414	0.49071	-9.7482	1.7998

Table 3.11 – Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_1 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{σ_1}
1	initial values	0.50	2.4633	2.9933	0.1500	0.22470		
2	at convergence	0.92	2.6604	3.5309	0.0000	0.06186	-6.6117	862.0739
3	initial values	0.50	2.4633	2.9933	0.1700	0.22470		
4	at convergence	0.92	2.6569	3.6198	0.0000	0.06651	-6.208	847.0398
5	initial values	0.50	2.4633	2.9933	0.1897	0.22470		
6	at convergence	0.92	2.6569	3.6198	0.0000	0.06653	-6.208	1457.1678
7	initial values	0.50	2.4633	2.9933	0.2000	0.22470		
8	at convergence	0.92	2.6569	3.6198	0.0000	0.06653	-6.208	2122.1686
9	initial values	0.50	2.4633	2.9933	0.2100	0.22470		
10	at convergence	0.49	2.5670	2.8892	0.0487	0.26659	-9.7797	14.8991

Table 3.12 – Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_2 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{σ_2}
1	initial values	0.50	2.4633	2.9933	0.18970	0.19000		
2	at convergence	0.96	2.6433	4.3467	0.01701	0.12776	-5.8129	0.6613
3	initial values	0.50	2.4633	2.9933	0.18970	0.20000		
4	at convergence	0.92	2.6569	3.6199	0.00167	0.06663	-6.222	0.3196
5	initial values	0.50	2.4633	2.9933	0.18970	0.22470		
6	at convergence	0.92	2.6569	3.6198	0.00000	0.06653	-6.208	0.3174
7	initial values	0.50	2.4633	2.9933	0.18970	0.22500		
8	at convergence	0.92	2.6569	3.6198	0.00017	0.06652	-6.2081	0.3174
9	initial values	0.50	2.4633	2.9933	0.18970	0.22700		
10	at convergence	0.92	2.6569	3.6198	0.00000	0.06653	-6.208	0.3174

In general, we note in these tables that the estimates did not converge to a single solution, for each initial value we obtained different values for the parameter estimates. For the highest values of the likelihood function, we not observed closeness of the estimates in relation to the true values. We observed that the largest discrepancies of the estimates in relation to the true values occurred for the parameters of proportion (p) and scale (σ). Note also that the standard error estimates were larger for the parameters p , σ , σ_1 and σ_2 .

In Appendix B of this work are the tables with model estimates for the sum of the weights using the other two optimization methods (BFGS and SANN).

3.4.2.2 Analysis of the sum of chicken weights of the real data using the classical approach

The mixture model was used for the sum of the chicken weights to the real data considering the weights at 42 days for each treatment. We used the algorithm of Nelder and Mead (1965) to obtain the estimates of the parameters of the reparameterized model. The initial values used for the algorithm were the maximum likelihood estimates of the individual data in the case of the model with a single σ and the Bayesian estimates in the case of the model with two σ 's.

Based on the estimations of the mixture model of the individual data presented in Tables (3.13) and (3.14) we observe discrepancies in the values of the estimates, mainly in the parameters of proportion (p) and scale (σ , σ_1 and σ_2), which also presented higher estimates for the standard error.

Table 3.13 – Estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the weights data of chickens at 42 days by treatment. The parameters estimates by the Nelder-Mead algorithm, as well as the respective standard errors (SE) are presented

Treatment	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$
T1	0.3200 (0.1426)	2.1311 (0.0094)	2.9017 (0.0045)	0.0146 (0.3162)
T2	0.5400 (0.1336)	2.3555 (0.0046)	3.0792 (0.0055)	0.0115 (0.3163)
T3	0.2500 (0.1548)	2.2861 (0.0144)	2.8762 (0.0051)	0.0176 (0.3273)
T4	0.0800 (0.2456)	1.4886 (0.0430)	2.8190 (0.0041)	0.0252 (0.3163)
T5	0.6500 (0.1401)	2.4874 (0.0060)	3.2110 (0.0112)	0.0196 (0.3191)
T6	0.2200 (0.1634)	2.3573 (0.0095)	2.9470 (0.0028)	0.0131 (0.3163)

Table 3.14 – Estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the weights data of chickens at 42 days by treatment. The parameters estimates by the Nelder-Mead algorithm, as well as the respective standard errors (SE) are presented

Treatment	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$
T1	0.0500 (0.2966)	1.0913 (0.0681)	2.7447 (0.0046)	0.0002 (319.2766)	0.0368 (0.3163)
T2	0.6800 (0.1428)	2.3948 (0.0065)	3.3066 (0.0134)	0.0217 (0.3204)	0.0023 (11.842)
T3	0.1000 (1.3575)	2.1518 (0.9737)	2.7867 (0.0348)	0.5882 (0.8106)	0.0128 (8.8187)
T4	0.1000 (1.1403)	2.1128 (0.8618)	2.7791 (0.0449)	0.7295 (0.7568)	0.0217 (6.7906)
T5	0.9800 (0.4522)	2.6872 (0.0034)	4.9831 (0.0986)	0.0377 (0.3162)	0.0006 (169.1829)
T6	0.9600 (0.3276)	2.6363 (0.0062)	4.9301 (0.1195)	0.0424 (0.3633)	0.0776 (1.4974)

3.4.2.3 Simulated case study of the sum of chicken weights using the Bayesian Approach

We present in this section the modeling results for the simulated data of the sum of chicken weights using the Bayesian approach. We used 3 chains with different initial values and the convergence was monitored by observing the “potential scale reduction factor” (GELMAN; RUBIN, 1992) and the Monte Carlo standard error for all parameters of the models. We used 6000 iterations with a burn-in of 3000 and the thinning of the sequence equal to 10. For all the parameters the Monte Carlo standard error estimate was reasonably small and the statistic “potential scale reduction factor” was equal to 1, confirming that the number of iterations was satisfactory for the convergence.

Tables (3.15), (3.16), (3.17) and (3.18) present the estimates of the parameters, Monte Carlo standard error (MCSE), standard deviation (SD) and the quantiles of the posterior distribution for each parameter considering the model with a single σ and two σ 's.

The results of modeling of the simulated data with non-informative priors are in Tables (3.15) and (3.16). We noticed that, in relation to the estimates obtained using the classical approach, the estimates from the non-informative prior model were closer to the true values, however, the credibility intervals were not narrow.

In Tables (3.17) and (3.18) the results of the modeling with informative priors are presented. We observed that the results were sensitive to the informative priors, in comparison to the modeling with non-informative priors, there were changes in the values of the estimates, in general, there was a reduction in the value of the standard deviation and the intervals of credibility were narrower.

Table 3.15 – Bayesian estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and quantiles of the posterior distribution for each model parameter

	truth	mean	MCSE	SD	2.5%	25%	50%	75%	97.5%
$\hat{p}$	0.56	0.53	0.00	0.29	0.06	0.27	0.53	0.78	0.98
$\hat{\mu}_1$	2.48	2.66	0.01	0.59	1.09	2.44	2.73	2.97	3.75
$\hat{\mu}_2$	3.04	2.63	0.01	0.66	0.87	2.38	2.73	3.00	3.74
$\hat{\sigma}$	0.19	0.38	0.00	0.22	0.05	0.21	0.34	0.51	0.90

Table 3.16 – Bayesian estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and quantiles of the posterior distribution for each model parameter

	truth	mean	MCSE	SD	2.5%	25%	50%	75%	97.5%
$\hat{p}$	0.50	0.53	0.00	0.30	0.06	0.25	0.54	0.80	0.98
$\hat{\mu}_1$	2.46	2.67	0.01	0.54	1.27	2.51	2.73	2.90	3.70
$\hat{\mu}_2$	2.99	2.65	0.01	0.61	0.96	2.47	2.73	2.96	3.70
$\hat{\sigma}_1$	0.19	0.38	0.00	0.26	0.02	0.17	0.34	0.56	0.93
$\hat{\sigma}_2$	0.22	0.40	0.00	0.27	0.02	0.18	0.36	0.60	0.95

Table 3.17 – Bayesian estimates of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) with informative priors for the simulated data considering 5 boxes and 46 individuals per box. Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and quantiles of the posterior distribution for each model parameter

	truth	mean	MCSE	SD	2.5%	25%	50%	75%	97.5%
$\hat{p}$	0.56	0.50	0.00	0.09	0.31	0.43	0.49	0.56	0.68
$\hat{\mu}_1$	2.48	2.49	0.01	0.17	2.13	2.39	2.49	2.60	2.82
$\hat{\mu}_2$	3.04	2.96	0.01	0.17	2.65	2.85	2.96	3.07	3.29
$\hat{\sigma}$	0.19	0.36	0.01	0.22	0.04	0.20	0.33	0.50	0.85

Table 3.18 – Bayesian estimates of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) with informative priors for the simulated data considering 5 boxes and 46 individuals per box. Also shown are the standard deviation (SD), Monte Carlo standard error (MCSE) and quantiles of the posterior distribution for each model parameter

	truth	mean	MCSE	SD	2.5%	25%	50%	75%	97.5%
$\hat{p}$	0.50	0.49	0.00	0.10	0.30	0.42	0.50	0.57	0.69
$\hat{\mu}_1$	2.46	2.53	0.01	0.17	2.18	2.43	2.54	2.65	2.84
$\hat{\mu}_2$	2.99	2.93	0.01	0.17	2.62	2.82	2.93	3.04	3.28
$\hat{\sigma}_1$	0.19	0.38	0.01	0.25	0.02	0.17	0.34	0.57	0.93
$\hat{\sigma}_2$	0.22	0.35	0.01	0.25	0.02	0.15	0.31	0.50	0.91

3.4.2.4 Analysis of the sum of chicken weights of the real data using the Bayesian approach

In the Bayesian approach we used 5 chains with distinct initial values, 20000 iterations with discard of the first half of the iterations, and thin equal to 10. The number of iterations was satisfactory for convergence, since for all parameters in both models (a single σ and two σ 's) the Monte Carlo standard error estimate was reasonably small and the “potential scale reduction factor” statistic was less than 1.01.

Tables (3.19), (3.20), (3.21) and (3.22) present the estimates of the parameters for the mixture model considering the sum of the weights, for the cases with a single σ and two σ 's. In these tables we also find the Monte Carlo standard error (MCSE), the standard deviation (SD), the median and the 95% credibility interval (95% CrI) for each estimated parameter.

Tables (3.19) and (3.20) present the modeling results with the real data using non-informative priors. We note in these tables that, in relation to the estimates obtained in the classical approach, the estimates of the parameters p , μ_1 e μ_2 were closer to the estimates of the parameters of the mixture model for the individual data presented in Tables (3.2) and (3.3).

As for the scale parameters, σ , σ_1 and σ_2 , the estimates were overestimated in relation to the estimates for the individual weights. Overall, 95% credibility intervals were not narrow and the estimates deviation values were greater than deviations from the mixture model estimates for the individual weights.

The results of modeling with informative priors are in Tables (3.21) and (3.22). We observed that the results were sensitive to informative priors. The estimates for μ_1 and μ_2 were closer to the estimates of the parameters of the mixture model for the individual weights than the model with non-informative priors. For the parameter p the estimates, practically, did not vary among themselves, they were all close to 0.5. As for the scale parameters, σ , σ_1 and σ_2 , the estimates were overestimated in relation to the estimates of individual weights. We also noticed that, in relation to modeling with non-informative priors, there was in general a reduction in the values of the standard deviation and the intervals of credibility were narrowed.

Table 3.19 – Bayesian estimates of the model parameters of sum of the chickens weights with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$). Also shown are standard deviation (SD), Monte Carlo standard error (MCSE) and the credibility interval of 95% for each parameter of the model

Treatment		$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$
T1	mean	0.5245	2.5809	2.5345	0.4210
	MCSE	0.0045	0.0105	0.0110	0.0036
	SD	0.2872	0.6539	0.7201	0.2368
	median	0.5200	2.6600	2.6400	0.4000
	95% CrI	(0.06 – 0.98)	(0.92 – 3.75)	(0.65 – 3.77)	(0.05 – 0.93)
T2	mean	0.5199	2.5912	2.5905	0.4214
	MCSE	0.0042	0.0106	0.0106	0.0036
	SD	0.2845	0.6635	0.7005	0.2327
	median	0.5200	2.6900	2.6800	0.3900
	95% CrI	(0.06 – 0.98)	(0.85 – 3.75)	(0.74 – 3.8)	(0.05 – 0.92)
T3	mean	0.5399	2.6638	2.5867	0.3229
	MCSE	0.0057	0.0142	0.0172	0.0038
	SD	0.2976	0.5597	0.6887	0.2092
	median	0.5500	2.7300	2.7100	0.2800
	95% CrI	(0.06 – 0.98)	(1.11 – 3.67)	(0.58 – 3.7)	(0.04 – 0.85)
T4	mean	0.5285	2.6322	2.5983	0.3878
	MCSE	0.0042	0.0093	0.0101	0.0032
	SD	0.2881	0.6061	0.6786	0.2215
	median	0.5400	2.7200	2.7100	0.3500
	95% CrI	(0.06 – 0.98)	(1.04 – 3.75)	(0.74 – 3.76)	(0.05 – 0.89)
T5	mean	0.5158	2.6810	2.6168	0.4200
	MCSE	0.0052	0.0183	0.0109	0.0042
	SD	0.2850	0.6707	0.7012	0.2359
	median	0.5200	2.7500	2.7100	0.4000
	95% CrI	(0.07 – 0.98)	(0.88 – 3.9)	(0.75 – 3.79)	(0.04 – 0.92)
T6	mean	0.5401	2.7511	2.6644	0.3548
	MCSE	0.0043	0.0095	0.0164	0.0052
	SD	0.2919	0.5650	0.6833	0.2191
	median	0.5500	2.8300	2.7900	0.3200
	95% CrI	(0.06 – 0.98)	(1.29 – 3.72)	(0.83 – 3.81)	(0.04 – 0.88)

Table 3.20 – Bayesian estimates of the model parameters of sum of the chickens weights with heterogeneous variances (σ_1 and σ_2). Also shown are standard deviation (SD), Monte Carlo standard error (MCSE) and the credibility interval of 95% for each parameter of the model

Treat.		$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$
T1	mean	0.5357	2.5772	2.5316	0.4379	0.4489
	MCSE	0.0041	0.0089	0.0101	0.0038	0.0038
	SD	0.2867	0.6284	0.7114	0.2654	0.2708
	median	0.5500	2.6700	2.6400	0.4100	0.4200
	95% CrI	(0.06 – 0.98)	(0.89 – 3.71)	(0.64 – 3.77)	(0.02 – 0.96)	(0.03 – 0.96)
T2	mean	0.5241	2.6001	2.5883	0.4425	0.4620
	MCSE	0.0040	0.0091	0.0105	0.0038	0.0039
	SD	0.2801	0.6358	0.6924	0.2689	0.2715
	median	0.5200	2.6900	2.6900	0.4200	0.4500
	95% CrI	(0.06 – 0.97)	(0.9 – 3.73)	(0.68 – 3.78)	(0.03 – 0.95)	(0.02 – 0.96)
T3	mean	0.5429	2.6666	2.6122	0.3821	0.3974
	MCSE	0.0045	0.0082	0.0097	0.0037	0.0037
	SD	0.2982	0.5391	0.6512	0.2596	0.2629
	median	0.5600	2.7400	2.7100	0.3400	0.3600
	95% CrI	(0.06 – 0.98)	(1.16 – 3.66)	(0.67 – 3.73)	(0.02 – 0.94)	(0.02 – 0.94)
T4	mean	0.5312	2.6488	2.5847	0.4252	0.4335
	MCSE	0.0041	0.0084	0.0097	0.0037	0.0040
	SD	0.2907	0.5843	0.6715	0.2620	0.2730
	median	0.5400	2.7200	2.7000	0.4000	0.4100
	95% CrI	(0.06 – 0.98)	(1.1 – 3.72)	(0.69 – 3.72)	(0.02 – 0.95)	(0.02 – 0.96)
T5	mean	0.5256	2.6781	2.6301	0.4482	0.4617
	MCSE	0.0040	0.0089	0.0104	0.0039	0.0040
	SD	0.2823	0.6281	0.7147	0.2735	0.2731
	median	0.5200	2.7500	2.7300	0.4300	0.4500
	95% CrI	(0.07 – 0.97)	(0.97 – 3.8)	(0.62 – 3.83)	(0.03 – 0.96)	(0.02 – 0.96)
T6	mean	0.5377	2.7498	2.6751	0.4013	0.4055
	MCSE	0.0041	0.0078	0.0092	0.0038	0.0038
	SD	0.2922	0.5470	0.6530	0.2625	0.2657
	median	0.5500	2.8300	2.7900	0.3600	0.3700
	95% CrI	(0.06 – 0.98)	(1.31 – 3.7)	(0.88 – 3.78)	(0.02 – 0.94)	(0.02 – 0.94)

Table 3.21 – Bayesian estimates of the model parameters of sum of the chickens weights with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) with informative priors. Also shown are standard deviation (SD), Monte Carlo standard error (MCSE) and the credibility interval of 95% for each parameter of the model

Treatment		$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$
T1	mean	0.5090	2.4196	2.9162	0.4260
	MCSE	0.0014	0.0025	0.0026	0.0034
	SD	0.0954	0.1710	0.1810	0.2249
	median	0.5100	2.4200	2.9200	0.4000
	95% CrI	(0.32 – 0.69)	(2.09 – 2.76)	(2.56 – 3.27)	(0.05 – 0.91)
T2	mean	0.5097	2.4480	2.9406	0.4341
	MCSE	0.0013	0.0025	0.0026	0.0032
	SD	0.0952	0.1744	0.1842	0.2240
	median	0.5100	2.4500	2.9400	0.4100
	95% CrI	(0.32 – 0.7)	(2.09 – 2.8)	(2.59 – 3.31)	(0.06 – 0.91)
T3	mean	0.5040	2.5094	2.9496	0.3173
	MCSE	0.0014	0.0024	0.0025	0.0030
	SD	0.0957	0.1627	0.1676	0.2065
	median	0.5000	2.5100	2.9500	0.2800
	95% CrI	(0.32 – 0.69)	(2.18 – 2.84)	(2.62 – 3.29)	(0.03 – 0.84)
T4	mean	0.5066	2.4731	2.9638	0.3755
	MCSE	0.0014	0.0030	0.0034	0.0040
	SD	0.0968	0.1725	0.1790	0.2227
	median	0.5100	2.4700	2.9600	0.3500
	95% CrI	(0.32 – 0.69)	(2.13 – 2.82)	(2.62 – 3.31)	(0.02 – 0.88)
T5	mean	0.5047	2.4902	2.9935	0.4419
	MCSE	0.0013	0.0025	0.0026	0.0034
	SD	0.0943	0.1774	0.1806	0.2252
	median	0.5100	2.4900	2.9900	0.4200
	95% CrI	(0.32 – 0.69)	(2.13 – 2.84)	(2.64 – 3.35)	(0.06 – 0.92)
T6	mean	0.4904	2.5752	3.0428	0.3324
	MCSE	0.0030	0.0024	0.0023	0.0085
	SD	0.0966	0.1685	0.1593	0.2166
	median	0.4900	2.5800	3.0500	0.3000
	95% CrI	(0.31 – 0.68)	(2.23 – 2.92)	(2.72 – 3.36)	(0.01 – 0.85)

Table 3.22 – Bayesian estimates of the model parameters of sum of the chickens weights with heterogeneous variances (σ_1 and σ_2) with informative priors. Also shown are standard deviation (SD), Monte Carlo standard error (MCSE) and the credibility interval of 95% for each parameter of the model

Treat.		$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$
T1	mean	0.5089	2.4397	2.8984	0.4404	0.4326
	MCSE	0.0014	0.0024	0.0029	0.0045	0.0040
	SD	0.0964	0.1694	0.1822	0.2655	0.2691
	median	0.5100	2.4400	2.8900	0.4200	0.4100
	95% CrI	(0.32 – 0.69)	(2.11 – 2.77)	(2.56 – 3.28)	(0.02 – 0.95)	(0.02 – 0.95)
T2	mean	0.5083	2.4639	2.9237	0.4447	0.4493
	MCSE	0.0014	0.0025	0.0026	0.0038	0.0038
	SD	0.0956	0.1735	0.1809	0.2683	0.2661
	median	0.5100	2.4700	2.9200	0.4200	0.4400
	95% CrI	(0.32 – 0.7)	(2.11 – 2.8)	(2.58 – 3.29)	(0.02 – 0.95)	(0.02 – 0.95)
T3	mean	0.5009	2.5267	2.9306	0.3628	0.3636
	MCSE	0.0014	0.0023	0.0024	0.0035	0.0036
	SD	0.0990	0.1651	0.1670	0.2431	0.2487
	median	0.5000	2.5300	2.9300	0.3300	0.3300
	95% CrI	(0.31 – 0.7)	(2.18 – 2.85)	(2.6 – 3.28)	(0.02 – 0.92)	(0.02 – 0.92)
T4	mean	0.5060	2.4998	2.9386	0.4139	0.4116
	MCSE	0.0014	0.0024	0.0026	0.0038	0.0038
	SD	0.0981	0.1679	0.1731	0.2589	0.2613
	median	0.5100	2.5100	2.9400	0.3800	0.3800
	95% CrI	(0.31 – 0.7)	(2.16 – 2.82)	(2.6 – 3.29)	(0.02 – 0.94)	(0.02 – 0.95)
T5	mean	0.5030	2.5133	2.9692	0.4469	0.4552
	MCSE	0.0014	0.0025	0.0025	0.0039	0.0039
	SD	0.0981	0.1753	0.1783	0.2677	0.2615
	median	0.5000	2.5200	2.9700	0.4400	0.4400
	95% CrI	(0.31 – 0.7)	(2.16 – 2.86)	(2.62 – 3.34)	(0.02 – 0.95)	(0.02 – 0.95)
T6	mean	0.4904	2.6050	3.0143	0.3974	0.3804
	MCSE	0.0014	0.0025	0.0023	0.0036	0.0036
	SD	0.0969	0.1723	0.1626	0.2563	0.2552
	median	0.4900	2.6100	3.0100	0.3700	0.3500
	95% CrI	(0.3 – 0.68)	(2.24 – 2.93)	(2.7 – 3.35)	(0.02 – 0.93)	(0.02 – 0.93)

3.5 Discussion

In this chapter, we propose a model for total weights data of broilers from a mixture of the sexes of birds. This proposed model takes into account the mixture of distributions of the sexes, information that is not taken into consideration in models commonly used for the data type under study. Since individual weights data were available for 7 and 42 days, which is not common to have in an experiment with 1380 birds, modeling for individual weights was done using the mixture model with two Gaussian components in order to establish relations between the estimates for this model and the estimates for the proposed model. To obtain the estimates of both models, the classical and Bayesian approaches were used.

Bayesian method provided a greater agreement between the estimates of the mixture model for the individual weights and the estimates of the model for the total weights. While in the classical approach, the numeric estimation algorithm used was very sensitive to the initial values with lower agreement between the estimates of both models.

We noticed that the largest discrepancies between the estimates occurred with the parameters of proportion (p) and scale (σ 's).

Although the results obtained here are not conclusive, we started an investigation motivated by a problem presented when using a usual model to the data studied. As future proposals we see the need to review the estimation methods and improve the proposed model in relation to the parameters of proportion and scale and to include covariables and longitudinal factor.

References

- BOCK, R. Darrell; KOLAKOWSKI, Donald. Further evidence of sex-linked major-gene influence on human spatial visualizing ability. **American Journal of Human Genetics**, v. 25, n. 1, p. 1, 1973.
- BOGARDUS, Clifton et al. Distribution of in vivo insulin action in Pima Indians as mixture of three normal distributions. **Diabetes**, v. 38, n. 11, p. 1423-1432, 1989.
- BROOKS, Steve et al. (Ed.). **Handbook of markov chain monte carlo**. CRC press, 2011. 592 p.
- CONGDON, Peter. **Applied Bayesian Modelling**. Chichester: John & Sons, 2003. 472 p.
- COHEN, A. Clifford. Estimation in mixtures of two normal distributions. **Technometrics**, v. 9, n. 1, p. 15-28, 1967.
- COOPER, D.B.; COOPER, P.W. Nonsupervised adaptive signal detection and pattern recognition. **Information and Control**, v. 7, n. 3, p. 416-444, 1964.
- DAY, Neil E. Estimating the components of a mixture of normal distributions. **Biometrika**, v. 56, n. 3, p. 463-474, 1969.
- DEMPSTER, A.P.; LAIRD, N.M.; RUBIN, D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm. **Journal of the Royal Statistical Society**, v. 39, n. 1, p. 1-38, 1977.
- DIEBOLT, J.; ROBERT, C.P. Estimation of Finite Mixture Distributions through Bayesian Sampling. **Journal of the Royal Statistical Society**, v. 56, n. 2, p. 363-375, 1994.
- DONG, Zhuan. Mixture analysis and its preliminary application in archaeology. **Journal of Archaeological Science**, v. 24, n. 2, p. 141-162, 1997.
- EVERITT, B.S. An introduction to finite mixture distributions. 1996.
- EVERITT, B.S.; HAND, D.J. **Finite Mixture Distributions**. New York: Chapman and Hall, 1981. 143 p.
- FOSDICK, Lloyd D. Calculation of order parameters in a binary alloy by the Monte Carlo method. **Physical Review**, v. 116, n. 3, 1959. 565 p.
- GELFAND, Alan E.; SMITH, Adrian FM. Sampling-based approaches to calculating marginal densities. **Journal of the American statistical association**, v. 85, n. 410, p. 398-409, 1990.
- GELMAN, Andrew; RUBIN, Donald B. Inference from iterative simulation using multiple sequences. **Statistical science**, p. 457-472, 1992.
- KALMENDAL, Robin et al. High-fibre sunflower cake affects small intestinal digestion and health in broiler chickens. **British poultry science**, v. 52, n. 1, p. 86-96, 2011.

KIEFER, Jack; WOLFOWITZ, Jacob. Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. **The Annals of Mathematical Statistics**, p. 887-906, 1956.

KORVER, D. R.; ZUIDHOF, M. J.; LAWES, K. R. Performance characteristics and economic comparison of broiler chickens fed wheat-and triticale-based diets. **Poultry Science**, v. 83, n. 5, p. 716-725, 2004.

LOURENS, Spencer et al. Bias in estimation of a mixture of normal distributions. **Journal of biometrics & biostatistics**, v. 4, 2013.

LITTLE, R.A.; RUBIN, D.B. **Statistical Analysis with Missing Data**. New York: Wiley, 1987. 408 p.

MARTIN, E. S. A study of an Egyptian series of mandibles, with special reference to mathematical methods of sexing. **Biometrika**, v. 28, n. 1/2, p. 149-178, 1936.

MCLACHLAN, G.J.; BASFORD, K.E. **Mixture models: inference and applications to clustering**. New York: Marcel Dekker, 1988. 253 p.

MCLACHLAN, G.; KRISHNAN, T. **The EM Algorithm and Extensions**. Hoboken: John Wiley & Sons. 2008. 400 p.

MCLACHLAN, G.; PEEL, D. **Finite Mixture Models**. New York: John Wiley & Sons, 2000. 291 p.

MENG, X.-L.; PEDLOW, S. EM: a bibliographic review with missing articles. **American Statistical Association Proceedings of the Statistical Computing Section**, Alexandria, VA: American Statistical Association, p. 24-27, 1992.

NELDER, John A.; MEAD, Roger. A simplex method for function minimization. **The computer journal**, v. 7, n. 4, p. 308-313, 1965.

PEARSON, K. Contributions to the theory of mathematical evolution. **The Royal Society**, London, v. 185, p. 71-110, 1894.

QURESHI, M. A.; GARLICH, J. D.; KIDD, M. T. Dietary Spirulina platensis enhances humoral and cell-mediated immune functions in chickens. **Immunopharmacology and Immunotoxicology**, v. 18, n. 3, p. 465-476, 1996.

R: A Language and Environmental for Statistical Computing, R Development Core Team, R Foundation for Statistical Computing, Vienna, 2017. Disponível em: <<http://www.R-project.org>>. Acesso em: 08 jun. 2018.

RAO, C. Radhakrishna. Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. In: Mathematical Proceedings of the Cambridge Philosophical Society. **Cambridge University Press**, p. 50-57, 1948.

RIPLEY, R. D. **Stochastic simulation**. New York: John Wiley & Sons, 1987. 237 p.

RODRIGUES, Kênia Ferreira et al. Relação lisina digestível: proteína bruta em dietas para frangos de corte no período de 1 a 21 dias de idade. Desempenho e metabolismo. **Revista Brasileira de Zootecnia**, Viçosa, v. 37, n. 3, p. 450-457, 2008.

- ROEDER, K.; WASSERMAN, L. Practical Bayesian Density Estimation Using Mixture of Normals. **Journal of the American Statistical Association**, Alexandria, v. 92, n. 439, p. 894-902, 1997.
- SCHAFER, J. **Analysis by Simulation of Incomplete Multivariate Data**. London: Chapman & Hall, 1996. 432 p.
- SMITH, Adrian FM. Bayesian computational methods. **Phil. Trans. R. Soc. Lond. A**, v. 337, n. 1647, p. 369-386, 1991.
- TAN, W. Y.; CHANG, W. C. Some comparisons of the method of moments and the method of maximum likelihood in estimating parameters of a mixture of two normal densities. **Journal of the American Statistical Association**, v. 67, n. 339, p. 702-708, 1972.
- TANNER, M.A. **Tools for Statistical Inference: Observed Data and Data Augmentation Methods**, Vol. 67. New York: Springer-Verlag, 1991. 110 p.
- TANNER, M.A. **Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions**. New York: Springer- Verlag, 1993. 208 p.
- TANNER, Martin A.; WONG, Wing Hung. The calculation of posterior distributions by data augmentation. **Journal of the American statistical Association**, v. 82, n. 398, p. 528-540, 1987.
- TEIXEIRA, Elisanie Neiva Magalhães et al. Inclusão do ovo desidratado em rações peletizadas ou fareladas para frangos de corte. **Revista Brasileira de Zootecnia**, p. 1372-1381, 2007.
- TITTERINGTON, D.M.; SMITH, A.F.M.; MAKOV, U.E. **Statistical analysis of finite mixture distributions**. New York: John Wiley & Sons, 1985. 243 p.
- TRALDI, A.B. **Influência da idade da matriz e do peso do ovo incubado nas respostas de pintos de corte alimentados com rações pré-iniciais farelada, triturada ou micro-peletizada**. 2009. 205p. Tese (Doutorado em Ciências) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2009.
- XU, Dinghai; KNIGHT, John. Continuous empirical characteristic function estimation of mixtures of normal parameters. **Econometric Reviews**, v. 30, n. 1, p. 25-50, 2010.
- WIRJANTO, Tony S.; XU, Dinghai. The applications of mixtures of normal distributions in empirical finance: A selected survey. Department of Statistics, University of Waterloo, Canada, 2009.
- WOLFE, J.H. **A computer program for the maximum likelihood analysis of types**. Naval Personnel Research Activity San Diego CA, 1965.
- WOLFE, J.H. **NORMIX: Computational methods for estimating the parameters of multivariate normal mixtures of distributions**. Naval Personnel Research Activity San Diego CA, 1967.

WOLFE, J.H. **Pattern clustering by multivariate mixture analysis**. Naval Personnel Research Activity San Diego CA, 1969.

WOLFE, J.H. Pattern clustering by multivariate mixture analysis. **Multivariate Behavioral Research**, v. 5, n. 3, p. 329-350, 1970.

APPENDIX

Appendix A: code related to the results of Chapter 2

```

rm(list=ls())
I <- 6 # treat
J <- 5 # rep.
soma.inv.T1 <- sum(1/45,1/46,1/46,1/45,1/45)
soma.inv.T2 <- sum(1/45,1/44,1/45,1/46,1/46)
soma.inv.T3 <- sum(1/45,1/43,1/44,1/45,1/43)
soma.inv.T4 <- sum(1/45,1/45,1/45,1/46,1/46)
soma.inv.T5 <- sum(1/44,1/46,1/45,1/45,1/45)
soma.inv.T6 <- sum(1/45,1/43,1/45,1/43,1/42)
mh <- (I*J)/sum(soma.inv.T1,soma.inv.T2,soma.inv.T3,soma.inv.T4,
soma.inv.T5,soma.inv.T6)
rm <- J*mh
#QMRes at 42 days: 138554.6
#sigma <- sqrt(138554.6)
sigma <- sqrt(3071)

#1. delta

#
require(dae)
#1. poder como função de delta
y1 <- vector("numeric",20)
y2 <- vector("numeric",20)
y3 <- vector("numeric",20)
for(i in 1:20){
y1[i] <- power.exp(rm=5, df.num=I-1, df.denom=24, delta=10*i,
sigma=sigma,alpha=0.05, print=FALSE)
y2[i] <- power.exp(rm=5, df.num=I-1, df.denom=24, delta=10*i,
sigma=sigma*0.75,alpha=0.05, print=FALSE)

y3[i] <- power.exp(rm=5, df.num=I-1, df.denom=24, delta=10*i,
sigma=sigma*0.50,alpha=0.05, print=FALSE)
}

x <- seq(10,200,10)

plot(x,y1,type="b", xaxt="n",

```



```

ylab=expression("Power"(1-beta)), xlab="Mean Difference",
cex.axis=0.8, las=2, pch=15)
abline(h=c(0.2,0.4,0.6),v=seq(10,200,10),lty=3,col="gray")
lines(x,y1,type="b",pch=15)
lines(x,y2,type="b",pch=16)
lines(x,y3,type="b",pch=17)
axis(1,x,x,cex.axis=0.8)
abline(h=0.8,lty=2)
legend(130,0.5,legend=c(expression(sigma^2 == 3071),
expression('75% of' ~ sigma^2),
expression('50% of' ~ sigma^2)),
pch=c(15,16,17),bty="n",lty=1)

```

#2. Poder como função do tamanho da amostra

```

y1 <- vector("numeric",40)
y2 <- vector("numeric",40)
y3 <- vector("numeric",40)

for(i in 2:41){
y1[i-1] <- power.exp(rm=i, df.num=I-1, df.denom=(i*I-1)-(I-1), delta=50,
sigma=sigma,alpha=0.05, print=FALSE)
y2[i-1] <- power.exp(rm=i, df.num=I-1, df.denom=(i*I-1)-(I-1), delta=50,
sigma=sigma*0.75,alpha=0.05, print=FALSE)

y3[i-1] <- power.exp(rm=i, df.num=I-1, df.denom=(i*I-1)-(I-1), delta=50,
sigma=sigma*0.50,alpha=0.05, print=FALSE)
}

```

```

x <- seq(2,41,1)

```

```

plot(x,y1,type="b", xaxt="n",
ylab=expression("Power"(1-beta)), xlab="Sample Size",
cex.axis=0.8,pch=15,ylim=c(0.05,0.99),las=2)
abline(h=c(0.2,0.4,0.6),v=seq(2,41,1),lty=3,col="gray")
lines(x,y1,type="b",pch=15)
lines(x,y2,type="b",pch=16)
lines(x,y3,type="b",pch=17)

axis(1,x,x,cex.axis=0.8)

```

```

legend(27,0.6,legend=c(expression(sigma^2 == 3071),
expression('75% of' ~ sigma^2),
expression('50% of' ~ sigma^2)),
pch=c(15,16,17),bty="n",lty=1)
abline(h=0.8,lty=2)

```

#3. Tamanho da amostra em função do tamanho do efeito; mantendo
o poder em 0.8

```

poder <- 0.8
y1 <- vector("numeric",20)
y2 <- vector("numeric",20)
y3 <- vector("numeric",20)

delta <- seq(50,240,10)

for(i in 1:20){
y1[i] <- no.reps(multiple=1, df.num=I-1,
df.denom=expression(29-df.num), delta=delta[i],
sigma=sigma, alpha=0.05, power=poder, tol=0.1,
print=FALSE)$nreps

y2[i] <- no.reps(multiple=1, df.num=I-1,
df.denom=expression(29-df.num), delta=delta[i],
sigma=sigma*0.75, alpha=0.05, power=poder, tol=0.1,
print=FALSE)$nreps

y3[i] <- no.reps(multiple=1, df.num=I-1,
df.denom=expression(29-df.num), delta=delta[i],
sigma=sigma*0.50, alpha=0.05, power=poder, tol=0.1,
print=FALSE)$nreps

}

x <- seq(50,240,10)

plot(x,y1,type="b", xaxt="n", yaxt="n", ylab="Sample size",
xlab="Mean Difference",cex.axis=0.8, pch=15,ylim=c(0,50),
lty=1,las=2)

```

```
abline(h=seq(0,50,5),v=seq(50,240,10),lty=3,col="gray")
lines(x,y1,type="b",pch=15)
lines(x,y2,type="b",pch=16)
lines(x,y3,type="b",pch=17)
axis(1,x,x,cex.axis=0.8)
axis(2,seq(0,50,5),
seq(0,50,5),cex.axis=0.8,las=2)

legend(130,46,legend=c(expression(sigma^2 == 3071),
expression('75% of' ~ sigma^2),
expression('50% of' ~ sigma^2)),
lty=1,pch=c(15,16,17),bty="n")
```

Appendix B: estimates with other optimization methods (BFGS and SANN) referring to Chapter 3

Table 3.23 – Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_p
1	initial values	0.40	2.4810	3.0358	0.1935		
2	at convergence	0.43	2.6994	2.7490	0.3078	-10.7924	
3	initial values	0.50	2.4810	3.0358	0.1935		
4	at convergence	0.38	2.7127	2.7371	0.3089	-10.7924	
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.84	2.6075	3.3776	0.0280	-10.3536	0.1819
7	initial values	0.60	2.4810	3.0358	0.1935		
8	at convergence	0.87	2.5129	4.1612	0.0302	-8.6527	0.1958
9	initial values	0.70	2.4810	3.0358	0.1935		
10	at convergence	0.84	2.6075	3.3777	0.0280	-10.3536	0.1819

Table 3.24 – Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_{μ_1}
1	initial values	0.56	1.5000	3.0358	0.1935		
2	at convergence	0.58	0.9556	5.1580	0.0572	-13.079	0.1494
3	initial values	0.56	2.0000	3.0358	0.1935		
4	at convergence	0.87	2.6096	3.5480	0.0215	-8.3979	0.004
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.84	2.6075	3.3776	0.0280	-10.3536	0.006
7	initial values	0.56	2.8000	3.0358	0.1935		
8	at convergence	0.87	2.6096	3.5480	0.0215	-8.3979	0.004
9	initial values	0.56	2.8500	3.0358	0.1935		
10	at convergence	0.13	-26.1558	7.0605	0.3089	-19.5879	

Table 3.25 – Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_{μ_2}
1	initial values	0.56	2.4810	2.5000	0.1935		
2	at convergence	0.35	29.2656	-11.4255	0.3089	-21.2808	
3	initial values	0.56	2.4810	2.9000	0.1935		
4	at convergence	0.94	2.6708	3.6092	0.0215	-8.5452	0.0275
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.84	2.6075	3.3776	0.0280	-10.3536	0.0301
7	initial values	0.56	2.4810	3.5000	0.1935		
8	at convergence	0.78	2.5823	3.2389	0.0210	-10.0102	0.0188
9	initial values	0.56	2.4810	3.7000	0.1935		
10	at convergence	0.87	2.5129	4.1612	0.0302	-8.6527	0.0629

Table 3.26 – Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_{σ}
1	initial values	0.56	2.4810	3.0358	0.0800		
2	at convergence	0.95	2.6653	3.8646	0.0189	-6.3267	0.3162
3	initial values	0.56	2.4810	3.0358	0.1000		
4	at convergence	0.86	2.5611	3.7604	0.0189	-7.012	0.3162
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.84	2.6075	3.3776	0.0280	-10.3536	0.4435
7	initial values	0.56	2.4810	3.0358	0.2200		
8	at convergence	0.85	2.5892	3.5276	0.0215	-8.4736	0.3163
9	initial values	0.56	2.4810	3.0358	0.2400		
10	at convergence	0.85	2.5892	3.5276	0.0215	-8.4736	0.3163

Table 3.27 – Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_p
1	initial values	0.42	2.4633	2.9933	0.18970	0.2247		
2	at convergence	0.95	2.6423	4.3677	0.00042	0.0760	-3.9245	0.2961
3	initial values	0.45	2.4633	2.9933	0.18970	0.2247		
4	at convergence	0.97	2.6782	4.4585	0.00023	0.1388	-0.071	0.3682
5	initial values	0.50	2.4633	2.9933	0.18970	0.2247		
6	at convergence	0.92	2.6569	3.6198	0.00030	0.0665	-6.2084	0.2455
7	initial values	0.60	2.4633	2.9933	0.18970	0.2247		
8	at convergence	0.90	2.6356	3.6022	0.00028	0.0605	-6.6435	0.2198
9	initial values	0.65	2.4633	2.9933	0.18970	0.2247		
10	at convergence	0.92	2.6604	3.5309	0.00024	0.0619	-6.6119	0.2395

Table 3.28 – Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{μ_1}
1	initial values	0.50	2.3000	2.9933	0.18970	0.22470		
2	at convergence	0.92	2.6569	3.6198	0.00030	0.06653	-6.2084	0.0018
3	initial values	0.50	2.4000	2.9933	0.18970	0.22470		
4	at convergence	0.93	2.6782	3.6238	0.00036	0.08574	-3.0659	4e-04
5	initial values	0.50	2.4633	2.9933	0.18970	0.22470		
6	at convergence	0.92	2.6569	3.6198	0.00030	0.06652	-6.2084	0.0018
7	initial values	0.50	2.4700	2.9933	0.18970	0.22470		
8	at convergence	0.92	2.6569	3.6198	0.00040	0.06653	-6.2088	0.0018
9	initial values	0.50	2.4750	2.9933	0.18970	0.22470		
10	at convergence	0.90	2.6356	3.6022	0.00035	0.06051	-6.6437	0.0028

Table 3.29 – Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{μ_2}
1	initial values	0.50	2.4633	2.9200	0.18970	0.22470		
2	at convergence	0.88	2.6143	3.5833	0.00015	0.05593	-6.9688	0.0287
3	initial values	0.50	2.4633	2.9500	0.18970	0.22470		
4	at convergence	0.90	2.6356	3.6022	0.00082	0.06051	-6.6452	0.0281
5	initial values	0.50	2.4633	2.9933	0.18970	0.22470		
6	at convergence	0.92	2.6569	3.6198	0.00030	0.06652	-6.2084	0.0264
7	initial values	0.50	2.4633	3.0000	0.18970	0.22470		
8	at convergence	0.77	2.6239	3.1024	0.02745	0.31994	-9.7621	
9	initial values	0.50	2.4633	3.1000	0.18970	0.22470		
10	at convergence	0.90	2.6356	3.6022	0.00038	0.06054	-6.6437	0.0281

Table 3.30 – Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_1 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{σ_1}
1	initial values	0.50	2.4633	2.9933	0.15000	0.22470		
2	at convergence	0.98	2.6781	4.6800	0.00113	0.60199	-5.5271	0.9958
3	initial values	0.50	2.4633	2.9933	0.17000	0.22470		
4	at convergence	0.92	2.6569	3.6198	0.00020	0.06653	-6.2082	34.3309
5	initial values	0.50	2.4633	2.9933	0.18970	0.22470		
6	at convergence	0.92	2.6569	3.6198	0.00030	0.06652	-6.2084	23.4468
7	initial values	0.50	2.4633	2.9933	0.20000	0.22470		
8	at convergence	0.92	2.6569	3.6198	0.00016	0.06653	-6.2081	43.828
9	initial values	0.50	2.4633	2.9933	0.21000	0.22470		
10	at convergence	0.97	2.6782	4.4423	0.00031	0.09714	-0.2808	2.8981

Table 3.31 – Estimates by the BFGS optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_2 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{σ_2}
1	initial values	0.50	2.4633	2.9933	0.18970	0.19000		
2	at convergence	0.95	2.6423	4.3677	0.00032	0.07603	-3.9244	0.3169
3	initial values	0.50	2.4633	2.9933	0.18970	0.20000		
4	at convergence	0.92	2.6569	3.6198	0.00037	0.06651	-6.2087	0.3174
5	initial values	0.50	2.4633	2.9933	0.18970	0.22470		
6	at convergence	0.92	2.6569	3.6198	0.00030	0.06652	-6.2084	0.3174
7	initial values	0.50	2.4633	2.9933	0.18970	0.22500		
8	at convergence	0.95	2.6782	4.4515	0.00020	0.10861	-0.8282	0.3613
9	initial values	0.50	2.4633	2.9933	0.18970	0.22700		
10	at convergence	0.97	2.6782	4.5688	0.00067	0.30574	-3.2782	

Table 3.32 – Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_p
1	initial values	0.40	2.4810	3.0358	0.1935		
2	at convergence	0.99	2.7168	4.0580	0.2811	-10.7731	4.4071
3	initial values	0.50	2.4810	3.0358	0.1935		
4	at convergence	1.00	2.6994	6.3014	0.1089	-8.5605	1.0546
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.97	2.6041	6.7483	0.0730	-8.7676	0.4105
7	initial values	0.60	2.4810	3.0358	0.1935		
8	at convergence	0.81	2.8112	2.3400	0.2395	-10.8801	0.4552
9	initial values	0.70	2.4810	3.0358	0.1935		
10	at convergence	0.99	2.7014	6.3005	0.1228	-8.5544	0.6655

Table 3.33 – Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_{μ_1}
1	initial values	0.56	1.5000	3.0358	0.1935		
2	at convergence	0.92	2.6093	3.8472	0.0279	-7.8309	0.0066
3	initial values	0.56	2.0000	3.0358	0.1935		
4	at convergence	0.92	2.6526	3.6359	0.1626	-10.6526	
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.98	2.6924	5.8146	0.1260	-9.6842	0.0108
7	initial values	0.56	2.8000	3.0358	0.1935		
8	at convergence	0.99	2.6901	7.0485	0.0537	-6.3884	0.0046
9	initial values	0.56	2.8500	3.0358	0.1935		
10	at convergence	0.97	2.6911	4.3434	0.0330	-6.8784	0.0029

Table 3.34 – Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_{μ_2}
1	initial values	0.56	2.4810	2.5000	0.1935		
2	at convergence	0.96	2.8031	1.0999	0.0550	-8.9679	0.1627
3	initial values	0.56	2.4810	2.9000	0.1935		
4	at convergence	0.99	2.7842	-1.4592	0.0805	-7.3671	0.4933
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.08	4.2440	2.6247	0.0502	-9.163	0.0111
7	initial values	0.56	2.4810	3.5000	0.1935		
8	at convergence	0.99	2.7884	-1.6255	0.0853	-7.7743	0.5642
9	initial values	0.56	2.4810	3.7000	0.1935		
10	at convergence	0.98	2.7842	-1.3440	0.0497	-7.4184	0.3066

Table 3.35 – Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with homogeneous variances ($\sigma = \sigma_1 = \sigma_2$) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}$	l	SE_{σ}
1	initial values	0.56	2.4810	3.0358	0.0800		
2	at convergence	0.91	2.6165	3.7728	0.0227	-7.5785	0.5053
3	initial values	0.56	2.4810	3.0358	0.1000		
4	at convergence	0.04	-1.4703	2.8790	0.0671	-8.985	0.3779
5	initial values	0.56	2.4810	3.0358	0.1935		
6	at convergence	0.93	2.6366	3.8715	0.0364	-8.0567	0.7249
7	initial values	0.56	2.4810	3.0358	0.2200		
8	at convergence	0.97	2.5868	7.3513	0.1247	-10.6836	1.3919
9	initial values	0.56	2.4810	3.0358	0.2400		
10	at convergence	0.97	2.7818	0.7763	0.0558	-9.642	0.3308

Table 3.36 – Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of p were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_p
1	initial values	0.42	2.4633	2.9933	0.18970	0.2247		
2	at convergence	0.07	4.3860	2.6428	0.03868	0.0561	-8.2261	0.2662
3	initial values	0.45	2.4633	2.9933	0.18970	0.2247		
4	at convergence	0.10	1.8315	2.8309	0.09881	0.0153	-9.2991	0.2235
5	initial values	0.50	2.4633	2.9933	0.18970	0.2247		
6	at convergence	1.00	2.7305	16.7990	0.25252	31198.7440	-9.8014	
7	initial values	0.60	2.4633	2.9933	0.18970	0.2247		
8	at convergence	0.31	2.4752	2.8393	0.00002	0.3559	-10.1866	
9	initial values	0.65	2.4633	2.9933	0.18970	0.2247		
10	at convergence	0.02	-4.8480	2.9020	1.73308	0.0204	-14.6323	0.4563

Table 3.37 – Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_1 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{μ_1}
1	initial values	0.50	2.3000	2.9933	0.18970	0.22470		
2	at convergence	0.02	18.8762	2.3734	0.23414	0.24866	-14.7476	
3	initial values	0.50	2.4000	2.9933	0.18970	0.22470		
4	at convergence	0.50	2.4000	2.9933	0.18970	0.22470	-11.4309	
5	initial values	0.50	2.4633	2.9933	0.18970	0.22470		
6	at convergence	0.48	3.0762	2.4439	0.01895	0.00025	-6.8439	0.0097
7	initial values	0.50	2.4700	2.9933	0.18970	0.22470		
8	at convergence	0.98	2.7645	0.8556	0.10249	0.01889	-9.5586	0.0061
9	initial values	0.50	2.4750	2.9933	0.18970	0.22470		
10	at convergence	1.00	2.7334	-14.4711	0.25325	0.00174	-9.7987	

Table 3.38 – Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of μ_2 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{μ_2}
1	initial values	0.50	2.4633	2.9200	0.18970	0.22470		
2	at convergence	0.75	3.1274	1.5649	0.03880	0.00118	-9.0187	
3	initial values	0.50	2.4633	2.9500	0.18970	0.22470		
4	at convergence	1.00	2.7332	8.8281	0.25213	0.00703	-9.7969	0
5	initial values	0.50	2.4633	2.9933	0.18970	0.22470		
6	at convergence	0.93	2.8618	1.1969	0.00611	0.08510	-5.4678	0.0643
7	initial values	0.50	2.4633	3.0000	0.18970	0.22470		
8	at convergence	0.10	4.1716	2.5786	0.04839	0.02851	-6.962	0.0249
9	initial values	0.50	2.4633	3.1000	0.18970	0.22470		
10	at convergence	0.06	4.3895	2.6400	0.00001	0.05237	-7.5427	0.0172

Table 3.39 – Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_1 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{σ_1}
1	initial values	0.50	2.4633	2.9933	0.15000	0.22470		
2	at convergence	0.97	2.6804	4.3973	0.02695	0.00009	-4.3324	0.5399
3	initial values	0.50	2.4633	2.9933	0.17000	0.22470		
4	at convergence	0.07	4.4006	2.6003	0.00004	0.03611	-7.1178	2284.4124
5	initial values	0.50	2.4633	2.9933	0.18970	0.22470		
6	at convergence	0.93	2.8238	1.2326	0.00003	0.09907	-5.431	408.7373
7	initial values	0.50	2.4633	2.9933	0.20000	0.22470		
8	at convergence	1.00	2.7362	15.1911	0.25611	0.00740	-9.8951	0.3218
9	initial values	0.50	2.4633	2.9933	0.21000	0.22470		
10	at convergence	1.00	2.7318	8.6481	0.25427	140773344.86524	-9.796	

Table 3.40 – Estimates by the Simulated-annealing (SANN) optimization method of the model parameters of the sum of the weights of chickens with heterogeneous variances (σ_1 and σ_2) for the simulated data considering 5 boxes and 46 individuals per box. The initial values of σ_2 were varied and kept the initial values of the other parameters fixed in the true values

	Setting	$\hat{p}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	l	SE_{σ_2}
1	initial values	0.50	2.4633	2.9933	0.18970	0.19000		
2	at convergence	1.00	2.7330	10.9644	0.25075	0.00134	-9.8136	
3	initial values	0.50	2.4633	2.9933	0.18970	0.20000		
4	at convergence	0.21	3.2023	2.5987	0.18430	0.00001	-9.839	27397.0788
5	initial values	0.50	2.4633	2.9933	0.18970	0.22470		
6	at convergence	0.27	3.9567	2.2796	0.00000	0.03511	-7.8947	0.5781
7	initial values	0.50	2.4633	2.9933	0.18970	0.22500		
8	at convergence	0.12	1.2413	2.8992	0.05491	0.00226	-6.339	
9	initial values	0.50	2.4633	2.9933	0.18970	0.22700		
10	at convergence	1.00	2.7328	-14.3949	0.25322	9713.53055	-9.7959	

Appendix C: code related to Chapter 3

```
// Stan code
data {
  int<lower=1> K; // number of mixture components
  int<lower=1> N; // number of data points
  real y[N]; // observations
}

parameters {
  simplex[K] theta; // mixing proportions
  ordered[K] mu; // locations of mixture components
  real sigma; // scales of mixture components
}

model {
  vector[K] log_theta = log(theta); // cache log calculation
  for (n in 1:N) {
    vector[K] lps = log_theta;
    for (k in 1:K)
      lps[k] = lps[k] + normal_lpdf(y[n] | mu[k], sigma);
    target += log_sum_exp(lps);
  }
}

data {
  int<lower=1> K; // number of mixture components
  int<lower=1> N; // number of data points
  real y[N]; // observations
}

parameters {
  simplex[K] theta; // mixing proportions
  ordered[K] mu; // locations of mixture components
  vector<lower=0>[K] sigma; // scales of mixture components
}

model {

  vector[K] log_theta = log(theta); // cache log calculation
  // priors
  theta ~ beta(10,10);
```

```

for (n in 1:N) {
  vector[K] lps = log_theta;
  for (k in 1:K)
    lps[k] = lps[k] + normal_lpdf(y[n] | mu[k], sigma[k]);
  target += log_sum_exp(lps);
}

//Stan code
functions{
  real newmix_log(vector y, vector n, int[] Nx,
    real p,
    real mu1,
    real mu2,
    real sigma){
    vector[num_elements(y)] ys;
    real lprob;

    for(i in 1:num_elements(y)){
      vector[Nx[i]] x;
      vector[Nx[i]] each;
      for(k in 1:Nx[i]){
        x[k] = k;
      }

      for(j in 1:num_elements(x)){
        each[j] = ( (tgamma(n[i]+1)/(tgamma(x[j]+1)*(tgamma(n[i]-x[j]+1))))
          *(pow(p,x[j]))*pow((1-p),(n[i]-x[j]))
          *(exp( -pow(y[i]-x[j]*mu1-(n[i]-x[j])*mu2,2)/ (2*(n[i])*pow(sigma,2)) ) )
          / (sqrt(2*pi()*(n[i])*pow(sigma,2))) ) );
      }

      ys[i] = sum(each);
    }

    lprob = sum(log(ys));
    return lprob;
  }
}

```

118

```
}  
}
```

```
data {  
  int<lower=0> N;  
  int Nx[5];  
  vector[N] y;  
  vector[N] n;  
}  
parameters {  
  real<lower=0,upper=1> p_est;  
  real<lower=0,upper=4.0> mu1_est;  
  real<lower=0,upper=4.0> mu2_est;  
  real<lower=0,upper=1.0> sigma_est;  
}
```

```
model{  
  y ~ newmix(n,Nx,p_est,mu1_est,mu2_est,sigma_est);  
}
```

```
//Stan code  
functions{  
  real newmix_log(vector y, vector n, int[] Nx,  
  real p,  
  real mu1,  
  real mu2,  
  real sigma1,  
  real sigma2){  
    vector[num_elements(y)] ys;  
    real lprob;
```

```
    for(i in 1:num_elements(y)){  
      vector[Nx[i]] x;  
      vector[Nx[i]] each;  
      for(k in 1:Nx[i]){  
        x[k] = k;  
      }  
    }
```

```
    for(j in 1:num_elements(x)){
```

```

each[j] = ( ((tgamma(n[i]+1)/(tgamma(x[j]+1)*(tgamma(n[i]-x[j]+1))))
*(pow(p,x[j]))*pow((1-p),(n[i]-x[j]))
*(exp( -pow(y[i]-x[j]*mu1-(n[i]-x[j])*mu2,2)/
(2*(x[j]*pow(sigma1,2)+(n[i]-x[j])*pow(sigma2,2)))) ) ) )
/ (sqrt(2*pi()*(x[j]*pow(sigma1,2)+(n[i]-x[j])*pow(sigma2,2)))) );
}

ys[i] = sum(each);

}

lprob = sum(log(ys));
return lprob;
}
}

data {
int<lower=0> N;
int Nx[5];
vector[N] y;
vector[N] n;
}

parameters {
real<lower=0,upper=1> p_est;
real<lower=0,upper=4.0> mu1_est;
real<lower=0,upper=4.0> mu2_est;
real<lower=0,upper=1.0> sigma1_est;
real<lower=0,upper=1.0> sigma2_est;
}

model{
y ~ newmix(n,Nx,p_est,mu1_est,mu2_est,sigma1_est,sigma2_est);
}

//Stan code
functions{
real newmix_log(vector y, vector n, int[] Nx,
real p,
real mu1,

```


120

```
real mu2,
real sigma){
vector[num_elements(y)] ys;
real lprob;

for(i in 1:num_elements(y)){
vector[Nx[i]] x;
vector[Nx[i]] each;
for(k in 1:Nx[i]){
x[k] = k;
}

for(j in 1:num_elements(x)){
each[j] = ( ((tgamma(n[i]+1)/(tgamma(x[j]+1)*(tgamma(n[i]-x[j]+1))))
*(pow(p,x[j]))*pow((1-p),(n[i]-x[j])))
*(exp( -pow(y[i]-x[j]*mu1-(n[i]-x[j])*mu2,2)/ (2*(n[i])*pow(sigma,2)) ) ))
/ (sqrt(2*pi()*(n[i])*pow(sigma,2))) );
}

ys[i] = sum(each);

}

lprob = sum(log(ys));
return lprob;
}
}

data {
int<lower=0> N;
int Nx[5];
vector[N] y;
vector[N] n;
}
parameters {
real<lower=0,upper=1> p_est;
real<lower=0,upper=4.0> mu1_est;
real<lower=0,upper=4.0> mu2_est;
real<lower=0,upper=1.0> sigma_est;
```

```

}

model{
  p_est ~ beta(12,12);
  mu1_est ~ normal(2.5,0.25);
  mu2_est ~ normal(3.0,0.25);

  y ~ newmix(n,Nx,p_est,mu1_est,mu2_est,sigma_est);
}

//Stan code
functions{
  real newmix_log(vector y, vector n, int[] Nx,
  real p,
  real mu1,
  real mu2,
  real sigma1,
  real sigma2){
    vector[num_elements(y)] ys;
    real lprob;

    for(i in 1:num_elements(y)){
      vector[Nx[i]] x;
      vector[Nx[i]] each;
      for(k in 1:Nx[i]){
        x[k] = k;
      }

      for(j in 1:num_elements(x)){
        each[j] = ( ((tgamma(n[i]+1)/(tgamma(x[j]+1)*(tgamma(n[i]-x[j]+1))))
        *(pow(p,x[j]))*pow((1-p),(n[i]-x[j])))
        *(exp( -pow(y[i]-x[j]*mu1-(n[i]-x[j])*mu2,2)/
        (2*(x[j]*pow(sigma1,2)+(n[i]-x[j])*pow(sigma2,2))) ) ) ) /
        (sqrt(2*pi()*(x[j]*pow(sigma1,2)+(n[i]-x[j])*pow(sigma2,2)))) ) );
      }

      ys[i] = sum(each);
    }
  }
}

```

```
}
```

```
lprob = sum(log(ys));
return lprob;
}
}
```

```
data {
  int<lower=0> N;
  int Nx[5];
  vector[N] y;
  vector[N] n;
}
parameters {
  real<lower=0,upper=1> p_est;
  real<lower=0,upper=4.0> mu1_est;
  real<lower=0,upper=4.0> mu2_est;
  real<lower=0,upper=1.0> sigma1_est;
  real<lower=0,upper=1.0> sigma2_est;
}
```

```
model{
  p_est ~ beta(12,12);
  mu1_est ~ normal(2.5,0.25);
  mu2_est ~ normal(3.0,0.25);

  y ~ newmix(n,Nx,p_est,mu1_est,mu2_est,sigma1_est,sigma2_est);
}
// Stan code end-----
```

```
### R code-----
```

```
#-----
```

```
# rmixsum() gera números aleatórios da soma por boxe de uma mistura de duas normais
rmixsum <- function(n.boxes,n,p,mu1,mu2,sigma1,sigma2,seed){
  set.seed(seed)
  Ni <- vector("numeric",n.boxes)
  for(i in 1:n.boxes){
    Ni[i] <- sum(rbinom(n=n,size=1,p=p))
```

```

}
n.male <- rnorm(sum(Ni),mean=mu1,sd=sigma1)
n.female <- rnorm((n.boxes*n)-sum(Ni),mean=mu2,sd=sigma2)
x <- c(n.male,n.female)
y <- sample(x, length(x), replace = FALSE, prob = NULL)
boxes <- matrix(0,nr=n.boxes,nc=n)
for(i in 1:n.boxes){
boxes[i,] <- c(sample(n.male,Ni[i]),sample(n.female,n-Ni[i]))
}
ysum <- vector("numeric",n.boxes)
for(i in 1:n.boxes){
ysum[i] <- sum(boxes[i,])
}
return(list("y"=y,"ysum"=ysum))
}
#-----
# a log-verossimilhança com sinal invertido
#uma variância (sigma)
#Reparametrizando
# p <- (1/(1+exp(-theta)))
# sigma <- exp(alpha)

ll.f1 <- function(par,y,n){
theta <- par[1]
mu1 <- par[2]
mu2 <- par[3]
alpha <- par[4]
ys <- vector("numeric",length(y))
for(i in 1:length(y)){
x <- 1:n[i]
a <- choose(n[i],x)*(1/(1+exp(-theta)))^x*(1-(1/(1+exp(-theta))))^(n[i]-x)
b <- exp( - (y[i]-x*mu1-(n[i]-x)*mu2)^2/
(2*(x*exp(alpha)^2+(n[i]-x)*exp(alpha)^2))
)
c <- sqrt(2*pi*(x*exp(alpha)^2+(n[i]-x)*exp(alpha)^2))
sum((a*b)/c)

ys[i] <- sum((a*b)/c)
}

```

124

```
ll <- sum(log(ys))
return(-ll)
}

# p <- (1/(1+exp(-theta)))
# sigma1 <- exp(alpha1)
# sigma2 <- exp(alpha2)
#duas variâncias (sigma1 e sigma2)
ll.f2 <- function(par,y,n){
  theta <- par[1]
  mu1 <- par[2]
  mu2 <- par[3]
  alpha1 <- par[4]
  alpha2 <- par[5]
  ys <- vector("numeric",length(y))
  for(i in 1:length(y)){
    x <- 1:n[i]
    a <- choose(n[i],x)*(1/(1+exp(-theta)))^x*(1-(1/(1+exp(-theta))))^(n[i]-x)
    b <- exp( - (y[i]-x*mu1-(n[i]-x)*mu2)^2/
      (2*(x*exp(alpha1)^2+(n[i]-x)*exp(alpha2)^2))
    )
    c <- sqrt(2*pi*(x*exp(alpha1)^2+(n[i]-x)*exp(alpha2)^2))
    sum((a*b)/c)

    ys[i] <- sum((a*b)/c)
  }
  ll <- sum(log(ys))
  return(-ll)
}

# Considerando 1 sigma (sigma)

#Estimativas EM
p <- mean(c(0.53,0.53,0.58,0.60,0.53,0.59))#0.56
mu1 <- round(mean(c(2.4126,2.4246,2.4915,2.4934,2.4731,2.5905)),4)#2.481
mu2 <- round(mean(c(2.9375,2.9818,3.0593,3.0499,3.0408,3.1458)),4)#3.0358
sigma.ests <- sqrt(c(0.04001,0.03661,0.03616,0.03517,0.03828,0.03846))
#round(sigma.ests,3)
```

```

sigma <- round(mean(sigma.ests),4)#0.1935

y.sum <- rmixsum(n.bboxes=5,n=46,p=0.56,mu1=2.481,mu2=3.0358,
sigma1=0.1935,sigma2=0.1935,seed=100)
names(y.sum)
y <- y.sum$ysum
n <- rep(46,5)
#par <- c(p=0.5,mu1=2.4633,mu2=2.9933,
#        sigma1=0.0367,sigma2=0.0517)
#optim(par,ll.f2,gr=NULL,y,n=n,method = "Nelder-Mead",
#      control = list(maxit = 20000, temp = 20),hessian = TRUE)

#-----
#----- p -----
par <- c(p=0.56,mu1=2.481,mu2=3.0358,
sigma=0.1935) #true

#Variação de p:
m <- matrix(c(log(0.4/(1-0.4)),2.481,3.0358,log(0.1935),
log(0.5/(1-0.5)),2.481,3.0358,log(0.1935),
log(0.56/(1-0.56)),2.481,3.0358,log(0.1935),
log(0.6/(1-0.6)),2.481,3.0358,log(0.1935),
log(0.7/(1-0.7)),2.481,3.0358,log(0.1935)),byrow=T,nc=4)

m.ests <- matrix(0,nc=4,nr=5)
ll.value <- vector("numeric",5)
st.err <- vector("numeric",5)
for(i in 1:length(ll.value)){
par<-m[i,]
results <- optim(par,ll.f1,gr=NULL,y,n=n,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE)
ests <- results$par
m.ests[i,] <- ests
ll.value[i] <- -results$value
st.err[i] <- sqrt(diag(solve(results$hessian))[1])
}

m.total <- matrix(0,nc=4,10)

```

126

```
for(i in 1:5){
m.total[(2*i-1),] <- m[i,]
m.total[2*i,] <- m.ests[i,]
}
colnames(m.total) <- c("p","mu1","mu2","sigma")

Setting <- rep(c("initial values","at convergence"),5)

ll.value <- round(ll.value,4)
l <- vector("character",10)
for(i in 1:5){
l[(2*i-1)] <- "          "
l[2*i] <- ll.value[i]
}

st.err <- round(st.err,4)
s <- vector("character",10)
for(i in 1:5){
s[(2*i-1)] <- "          "
s[2*i] <- st.err[i]
}

m.total[,1] <- 1/(1+exp(-m.total[,1]))
m.total[,4] <- exp(m.total[,4])

m.total <- round(m.total,4)
m.total[,1] <- round(m.total[,1],2)

d.p <- data.frame(Setting,m.total,"l"=l,"SE.p"=s)

knitr::kable(d.p)
xtable(d.p,auto=TRUE)
#-----
#----- mu1 -----
par <- c(p=0.56,mu1=2.481,mu2=3.0358,
sigma=0.1935) #true

#p <- c(0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,0.95)
#theta <- round(log(p/(1-p)),1)
```

```

#Variação de p:
m <- matrix(c(log(0.56/(1-0.56)),1.5,3.0358,log(0.1935),
log(0.56/(1-0.56)),2.0,3.0358,log(0.1935),
log(0.56/(1-0.56)),2.481,3.0358,log(0.1935),
log(0.56/(1-0.56)),2.8,3.0358,log(0.1935),
log(0.56/(1-0.56)),2.85,3.0358,log(0.1935)),byrow=T,nc=4)

m.ests <- matrix(0,nc=4,nr=5)
ll.value <- vector("numeric",5)
st.err <- vector("numeric",5)
for(i in 1:length(ll.value)){
  par<-m[i,]
  results <- optim(par,ll.f1,gr=NULL,y=n,n=n,method = "Nelder-Mead",
  control = list(maxit = 20000, temp = 20),
  hessian = TRUE)
  ests <- results$par
  m.ests[i,] <- ests
  ll.value[i] <- -results$value
  st.err[i] <- sqrt(diag(solve(results$hessian))[2])
}

m.total <- matrix(0,nc=4,10)
for(i in 1:5){
  m.total[(2*i-1),] <- m[i,]
  m.total[2*i,] <- m.ests[i,]
}
colnames(m.total) <- c("p","mu1","mu2","sigma")

Setting <- rep(c("initial values","at convergence"),5)

ll.value <- round(ll.value,4)
l <- vector("character",10)
for(i in 1:5){
  l[(2*i-1)] <- "          "
  l[2*i] <- ll.value[i]
}

st.err <- round(st.err,4)

```


128

```
s <- vector("character",10)
for(i in 1:5){
  s[(2*i-1)] <- "          "
  s[2*i] <- st.err[i]
}

m.total[,1] <- 1/(1+exp(-m.total[,1]))
m.total[,4] <- exp(m.total[,4])

m.total <- round(m.total,4)
m.total[,1] <- round(m.total[,1],2)

d.mu1 <- data.frame(Setting,m.total,"l"=1,"SE.mu1"=s)

knitr::kable(d.mu1)
xtable(d.mu1,auto=TRUE)
#----- mu2 -----
par <- c(p=0.56,mu1=2.481,mu2=3.0358,
sigma=0.1935) #true

m <- matrix(c(log(0.56/(1-0.56)),2.481,2.58,log(0.1935),
log(0.56/(1-0.56)),2.481,2.8,log(0.1935),
log(0.56/(1-0.56)),2.481,3.0358,log(0.1935),
log(0.56/(1-0.56)),2.481,3.5,log(0.1935),
log(0.56/(1-0.56)),2.481,3.7,log(0.1935)),byrow=T,nc=4)

m.ests <- matrix(0,nc=4,nr=5)
ll.value <- vector("numeric",5)
st.err <- vector("numeric",5)
for(i in 1:length(ll.value)){
  par<-m[i,]
  results <- optim(par,ll.f1,gr=NULL,y=n,n,method = "Nelder-Mead",
  control = list(maxit = 20000, temp = 20),
  hessian = TRUE)
  ests <- results$par
  m.ests[i,] <- ests
  ll.value[i] <- -results$value
  st.err[i] <- sqrt(diag(solve(results$hessian))[3])
}
```

```

m.total <- matrix(0,nc=4,10)
for(i in 1:5){
  m.total[(2*i-1),] <- m[i,]
  m.total[2*i,] <- m.ests[i,]
}
colnames(m.total) <- c("p","mu1","mu2","sigma")

Setting <- rep(c("initial values","at convergence"),5)

ll.value <- round(ll.value,4)
l <- vector("character",10)
for(i in 1:5){
  l[(2*i-1)] <- "          "
  l[2*i] <- ll.value[i]
}

st.err <- round(st.err,4)
s <- vector("character",10)
for(i in 1:5){
  s[(2*i-1)] <- "          "
  s[2*i] <- st.err[i]
}

m.total[,1] <- 1/(1+exp(-m.total[,1]))
m.total[,4] <- exp(m.total[,4])

m.total <- round(m.total,4)
m.total[,1] <- round(m.total[,1],2)

d.mu2 <- data.frame(Setting,m.total,"l"=1,"SE.mu2"=s)

knitr::kable(d.mu2)
xtable(d.mu2,auto=TRUE)
#----- sigma -----
par <- c(p=0.56,mu1=2.481,mu2=3.0358,
sigma=0.1935) #true

m <- matrix(c(log(0.56/(1-0.56)),2.481,3.0358,log(0.08),

```

130

```
log(0.56/(1-0.56)),2.481,3.0358,log(0.1),
log(0.56/(1-0.56)),2.481,3.0358,log(0.1935),
log(0.56/(1-0.56)),2.481,3.0358,log(0.22),
log(0.56/(1-0.56)),2.481,3.0358,log(0.24)),byrow=T,nc=4)

m.ests <- matrix(0,nc=4,nr=5)
ll.value <- vector("numeric",5)
st.err <- vector("numeric",5)
for(i in 1:length(ll.value)){
  par<-m[i,]
  results <- optim(par,ll.f1,gr=NULL,y=n,n=method = "Nelder-Mead",
  control = list(maxit = 20000, temp = 20),
  hessian = TRUE)
  ests <- results$par
  m.ests[i,] <- ests
  ll.value[i] <- -results$value
  st.err[i] <- sqrt(diag(solve(results$hessian))[4])
}

m.total <- matrix(0,nc=4,10)
for(i in 1:5){
  m.total[(2*i-1),] <- m[i,]
  m.total[2*i,] <- m.ests[i,]
}
colnames(m.total) <- c("p","mu1","mu2","sigma")

Setting <- rep(c("initial values","at convergence"),5)

ll.value <- round(ll.value,4)
l <- vector("character",10)
for(i in 1:5){
  l[(2*i-1)] <- "
  l[2*i] <- ll.value[i]
}

st.err <- round(st.err,4)
s <- vector("character",10)
for(i in 1:5){
  s[(2*i-1)] <- "
  s[2*i] <- st.err[i]
```

```

s[2*i] <- st.err[i]
}

m.total[,1] <- 1/(1+exp(-m.total[,1]))
m.total[,4] <- exp(m.total[,4])

m.total <- round(m.total,4)
m.total[,1] <- round(m.total[,1],2)

d.sig <- data.frame(Setting,m.total,"l"=1,"SE.sig"=s)

knitr::kable(d.sig)
xtable(d.sig,auto=TRUE)

#-----
#-----
#-----
#   2 sigmas
#-----
#

#Estimativas Bayesianas
p <- mean(c(0.28,0.43,0.67,0.57,0.50,0.55))#0.5
mu1 <- round(mean(c(2.32,2.38,2.54,2.48,2.47,2.59)),4)#2.4633
mu2 <- round(mean(c(2.79,2.91,3.11,3.02,3.02,3.11)),4)#2.9933
sigma1.est <- sqrt(c(0.02,0.03,0.05,0.04,0.04,0.04))
sigma2.est <- sqrt(c(0.08,0.06,0.03,0.04,0.05,0.05))
round(sigma1.est,2)
round(sigma2.est,2)

sigma1 <- round(mean(sigma1.est),4)#0.1897
sigma2 <- round(mean(sigma2.est),4)#0.2247

y.sum <- rmixsum(n.bboxes=5,n=46,p=0.5,mu1=2.4633,mu2=2.9933,
sigma1=0.1897,sigma2=0.2247,seed=100)
names(y.sum)

```

132

```
y <- y.sum$ysum
n <- rep(46,5)
#par <- c(p=0.5,mu1=2.4633,mu2=2.9933,
#          sigma1=0.0367,sigma2=0.0517)
#optim(par,ll.f2,gr=NULL,y,n=n,method = "Nelder-Mead",
#       control = list(maxit = 20000, temp = 20),hessian = TRUE)

#-----
#----- p -----
par <- c(p=0.5,mu1=2.4633,mu2=2.9933,
sigma1=0.1897,sigma2=0.2247) #true

#p <- c(0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,0.95)
#theta <- round(log(p/(1-p)),1)

#Variação de p:

m <- matrix(c(log(0.42/(1-0.42)),2.4633,2.9933,log(0.1897),log(0.2247),
log(0.45/(1-0.45)),2.4633,2.9933,log(0.1897),log(0.2247),
log(0.5/(1-0.5)),2.4633,2.9933,log(0.1897),log(0.2247),
log(0.6/(1-0.6)),2.4633,2.9933,log(0.1897),log(0.2247),
log(0.65/(1-0.65)),2.4633,2.9933,log(0.1897),log(0.2247)),
byrow=T,nc=5)

m.ests <- matrix(0,nc=5,nr=5)
ll.value <- vector("numeric",5)
st.err <- vector("numeric",5)
for(i in 1:length(ll.value)){
  par<-m[i,]
  results <- optim(par,ll.f2,gr=NULL,y,n=n,method = "Nelder-Mead",
  control = list(maxit = 20000, temp = 20),
  hessian = TRUE)
  ests <- results$par
  m.ests[i,] <- ests
  ll.value[i] <- -results$value
  st.err[i] <- sqrt(diag(solve(results$hessian))[1])
}
```

```

m.total <- matrix(0,nc=5,10)
for(i in 1:5){
  m.total[(2*i-1),] <- m[i,]
  m.total[2*i,] <- m.ests[i,]
}
colnames(m.total) <- c("p","mu1","mu2","sigma1","sigma2")

```

```

Setting <- rep(c("initial values","at convergence"),5)

```

```

ll.value <- round(ll.value,4)
l <- vector("character",10)
for(i in 1:5){
  l[(2*i-1)] <- "          "
  l[2*i] <- ll.value[i]
}

```

```

st.err <- round(st.err,4)
s <- vector("character",10)
for(i in 1:5){
  s[(2*i-1)] <- "          "
  s[2*i] <- st.err[i]
}

```

```

m.total[,1] <- 1/(1+exp(-m.total[,1]))
m.total[,4] <- exp(m.total[,4])
m.total[,5] <- exp(m.total[,5])

```

```

m.total[,c(2,3,5)] <- round(m.total[,c(2,3,5)],4)
m.total[,4] <- round(m.total[,4],5)
m.total[,1] <- round(m.total[,1],2)

```

```

d.p <- data.frame(Setting,m.total,"l"=1,"SE.p"=s)

```

```

knitr::kable(d.p)
xtable(d.p,auto=TRUE)

```

```

#-----

```

```

#-----

```

```

#----- mu1 -----
par <- c(p=0.5,mu1=2.4633,mu2=2.9933,
sigma1=0.1897,sigma2=0.2247) #true

#p <- c(0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,0.95)
#theta <- round(log(p/(1-p)),1)

#Variação de mu1:

m <- matrix(c(log(0.5/(1-0.5)),2.3,2.9933,log(0.1897),log(0.2247),
log(0.5/(1-0.5)),2.4,2.9933,log(0.1897),log(0.2247),
log(0.5/(1-0.5)),2.4633,2.9933,log(0.1897),log(0.2247),
log(0.5/(1-0.5)),2.47,2.9933,log(0.1897),log(0.2247),
log(0.5/(1-0.5)),2.475,2.9933,log(0.1897),log(0.2247)),
byrow=T,nc=5)

m.ests <- matrix(0,nc=5,nr=5)
ll.value <- vector("numeric",5)
st.err <- vector("numeric",5)
for(i in 1:length(ll.value)){
par<-m[i,]
results <- optim(par,ll.f2,gr=NULL,y=n,n=method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE)
ests <- results$par
m.ests[i,] <- ests
ll.value[i] <- -results$value
st.err[i] <- sqrt(diag(solve(results$hessian))[2])
}

m.total <- matrix(0,nc=5,10)
for(i in 1:5){
m.total[(2*i-1),] <- m[i,]
m.total[2*i,] <- m.ests[i,]
}
colnames(m.total) <- c("p","mu1","mu2","sigma1","sigma2")

Setting <- rep(c("initial values","at convergence"),5)

```

```

ll.value <- round(ll.value,4)
l <- vector("character",10)
for(i in 1:5){
  l[(2*i-1)] <- "          "
  l[2*i] <- ll.value[i]
}

st.err <- round(st.err,4)
s <- vector("character",10)
for(i in 1:5){
  s[(2*i-1)] <- "          "
  s[2*i] <- st.err[i]
}

m.total[,1] <- 1/(1+exp(-m.total[,1]))
m.total[,4] <- exp(m.total[,4])
m.total[,5] <- exp(m.total[,5])

m.total[,c(2,3)] <- round(m.total[,c(2,3)],4)
m.total[,c(4,5)] <- round(m.total[,c(4,5)],5)
m.total[,1] <- round(m.total[,1],2)

d.mu1 <- data.frame(Setting,m.total,"l"=l,"SE.mu1"=s)

knitr::kable(d.mu1)
xtable(d.mu1,auto=TRUE)

#----- mu2 -----
par <- c(p=0.5,mu1=2.4633,mu2=2.9933,
sigma1=0.1897,sigma2=0.2247) #true

#p <- c(0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,0.95)
#theta <- round(log(p/(1-p)),1)

#Variação de mu2:

```



```

m <- matrix(c(log(0.5/(1-0.5)),2.4633,2.92,log(0.1897),log(0.2247),
log(0.5/(1-0.5)),2.4633,2.95,log(0.1897),log(0.2247),
log(0.5/(1-0.5)),2.4633,2.9933,log(0.1897),log(0.2247),
log(0.5/(1-0.5)),2.4633,3.0,log(0.1897),log(0.2247),
log(0.5/(1-0.5)),2.4633,3.1,log(0.1897),log(0.2247)),
byrow=T,nc=5)

```

```

m.ests <- matrix(0,nc=5,nr=5)
ll.value <- vector("numeric",5)
st.err <- vector("numeric",5)
for(i in 1:length(ll.value)){
  par<-m[i,]
  results <- optim(par,ll.f2,gr=NULL,y=n,n=method = "Nelder-Mead",
  control = list(maxit = 20000, temp = 20),
  hessian = TRUE)
  ests <- results$par
  m.ests[i,] <- ests
  ll.value[i] <- -results$value
  st.err[i] <- sqrt(diag(solve(results$hessian))[3])
}

```

```

m.total <- matrix(0,nc=5,10)
for(i in 1:5){
  m.total[(2*i-1),] <- m[i,]
  m.total[2*i,] <- m.ests[i,]
}
colnames(m.total) <- c("p","mu1","mu2","sigma1","sigma2")

```

```

Setting <- rep(c("initial values","at convergence"),5)

```

```

ll.value <- round(ll.value,4)
l <- vector("character",10)
for(i in 1:5){
  l[(2*i-1)] <- "          "
  l[2*i] <- ll.value[i]
}

```

```

st.err <- round(st.err,4)

```

```

s <- vector("character",10)
for(i in 1:5){
  s[(2*i-1)] <- "          "
  s[2*i] <- st.err[i]
}

m.total[,1] <- 1/(1+exp(-m.total[,1]))
m.total[,4] <- exp(m.total[,4])
m.total[,5] <- exp(m.total[,5])

m.total[,c(2,3)] <- round(m.total[,c(2,3)],4)
m.total[,c(4,5)] <- round(m.total[,c(4,5)],5)
m.total[,1] <- round(m.total[,1],2)

d.mu2 <- data.frame(Setting,m.total,"l"=1,"SE.mu2"=s)

knitr::kable(d.mu2)
xtable(d.mu2,auto=TRUE)

#-----
#----- sigma1 -----
par <- c(p=0.5,mu1=2.4633,mu2=2.9933,
sigma1=0.1897,sigma2=0.2247) #true

#p <- c(0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,0.95)
#theta <- round(log(p/(1-p)),1)

#Variação de sigma1:

m <- matrix(c(log(0.5/(1-0.5)),2.4633,2.9933,log(0.15),log(0.2247),
log(0.5/(1-0.5)),2.4633,2.9933,log(0.17),log(0.2247),
log(0.5/(1-0.5)),2.4633,2.9933,log(0.1897),log(0.2247),
log(0.5/(1-0.5)),2.4633,2.9933,log(0.2),log(0.2247),
log(0.5/(1-0.5)),2.4633,2.9933,log(0.21),log(0.2247)),
byrow=T,nc=5)

```

138

```
m.ests <- matrix(0,nc=5,nr=5)
ll.value <- vector("numeric",5)
st.err <- vector("numeric",5)
for(i in 1:length(ll.value)){
  par<-m[i,]
  results <- optim(par,ll.f2,gr=NULL,y=n,n=n,method = "Nelder-Mead",
    control = list(maxit = 20000, temp = 20),
    hessian = TRUE)
  ests <- results$par
  m.ests[i,] <- ests
  ll.value[i] <- -results$value
  st.err[i] <- sqrt(diag(solve(results$hessian))[4])
}
```

```
m.total <- matrix(0,nc=5,10)
for(i in 1:5){
  m.total[(2*i-1),] <- m[i,]
  m.total[2*i,] <- m.ests[i,]
}
colnames(m.total) <- c("p","mu1","mu2","sigma1","sigma2")
```

```
Setting <- rep(c("initial values","at convergence"),5)
```

```
ll.value <- round(ll.value,4)
l <- vector("character",10)
for(i in 1:5){
  l[(2*i-1)] <- "          "
  l[2*i] <- ll.value[i]
}
```

```
st.err <- round(st.err,4)
s <- vector("character",10)
for(i in 1:5){
  s[(2*i-1)] <- "          "
  s[2*i] <- st.err[i]
}
```

```
m.total[,1] <- 1/(1+exp(-m.total[,1]))
m.total[,4] <- exp(m.total[,4])
```

```

m.total[,5] <- exp(m.total[,5])

m.total[,c(2,3)] <- round(m.total[,c(2,3)],4)
m.total[,c(4,5)] <- round(m.total[,c(4,5)],5)
m.total[,1] <- round(m.total[,1],2)

d.sig1 <- data.frame(Setting,m.total,"l"=1,"SE.sig1"=s)

knitr::kable(d.sig1)
xtable(d.sig1,auto=TRUE)

#-----
#----- sigma2 -----
par <- c(p=0.5,mu1=2.4633,mu2=2.9933,
sigma1=0.1897,sigma2=0.2247) #true

#p <- c(0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,0.95)
#theta <- round(log(p/(1-p)),1)

#Variação de sigma2:

m <- matrix(c(log(0.5/(1-0.5)),2.4633,2.9933,log(0.1897),log(0.19),
log(0.5/(1-0.5)),2.4633,2.9933,log(0.1897),log(0.20),
log(0.5/(1-0.5)),2.4633,2.9933,log(0.1897),log(0.2247),
log(0.5/(1-0.5)),2.4633,2.9933,log(0.1897),log(0.225),
log(0.5/(1-0.5)),2.4633,2.9933,log(0.1897),log(0.227)),
byrow=T,nc=5)

m.ests <- matrix(0,nc=5,nr=5)
ll.value <- vector("numeric",5)
st.err <- vector("numeric",5)
for(i in 1:length(ll.value)){
  par<-m[i,]
  results <- optim(par,ll.f2,gr=NULL,y=n,n=n,method = "Nelder-Mead",
  control = list(maxit = 20000, temp = 20),
  hessian = TRUE)

```

140

```
ests <- results$par
m.ests[i,] <- ests
ll.value[i] <- -results$value
st.err[i] <- sqrt(diag(solve(results$hessian))[5])
}

m.total <- matrix(0,nc=5,10)
for(i in 1:5){
m.total[(2*i-1),] <- m[i,]
m.total[2*i,] <- m.ests[i,]
}
colnames(m.total) <- c("p","mu1","mu2","sigma1","sigma2")

Setting <- rep(c("initial values","at convergence"),5)

ll.value <- round(ll.value,4)
l <- vector("character",10)
for(i in 1:5){
l[(2*i-1)] <- "          "
l[2*i] <- ll.value[i]
}

st.err <- round(st.err,4)
s <- vector("character",10)
for(i in 1:5){
s[(2*i-1)] <- "          "
s[2*i] <- st.err[i]
}

m.total[,1] <- 1/(1+exp(-m.total[,1]))
m.total[,4] <- exp(m.total[,4])
m.total[,5] <- exp(m.total[,5])

m.total[,c(2,3)] <- round(m.total[,c(2,3)],4)
m.total[,c(4,5)] <- round(m.total[,c(4,5)],5)
m.total[,1] <- round(m.total[,1],2)

d.sig2 <- data.frame(Setting,m.total,"l"=1,"SE.sig2"=s)
```

```

knitr::kable(d.sig2)
xtable(d.sig2,auto=TRUE)

#-----
##-----Estimando para cada tratamento
# Reparametrização
# p <- (1/(1+exp(-theta)))
# sigma <- exp(alpha)
#-----1 variância
#-----T1
y1
n1
#Valores iniciais (E.M. estimates)
# Problema na convergência impossibilitando o cálculo do st.err
p <- 0.53
mu1 <- 2.41
mu2 <- 2.94
sigma <- 0.20
#---
# Novos valores iniciais próximos das estimativas E.M.
p <- 0.53
mu1 <- 2.41
mu2 <- 2.94
sigma <- 0.23

par <- c(log(p/(1-p)),mu1,mu2,log(sigma)) #E.M. estimates

ests <- optim(par,ll.f1,gr=NULL,y=y1,n=n1,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE);ests
#-----
theta <- ests$par[1]
mu1 <- ests$par[2]
mu2 <- ests$par[3]
alpha <- ests$par[4]
#-----
p <- (1/(1+exp(-theta)))

```

142

```
sigma <- exp(alpha)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma"=sigma)
SE <- sqrt(diag(solve(ests$hessian)))

T1.est <- rbind("T1"=estimativas,"se.T1"=SE)

#-----T2
y2
n2
#Valores iniciais (E.M. estimates)
p <- 0.53
mu1 <- 2.42
mu2 <- 2.98
sigma <- 0.19
#---

par <- c(log(p/(1-p)),mu1,mu2,log(sigma)) #E.M. estimates

ests <- optim(par,ll.f1,gr=NULL,y=y2,n=n2,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE);ests
#-----
theta <- ests$par[1]
mu1 <- ests$par[2]
mu2 <- ests$par[3]
alpha <- ests$par[4]
#-----
p <- (1/(1+exp(-theta)))
sigma <- exp(alpha)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma"=sigma)
SE <- sqrt(diag(solve(ests$hessian)))

T2.est <- rbind("T2"=estimativas,"se.T2"=SE)

#-----T3
y3
```

```

n3
#Valores iniciais (E.M. estimates)
p <- 0.58
mu1 <- 2.49
mu2 <- 3.06
sigma <- 0.19
#---

par <- c(log(p/(1-p)),mu1,mu2,log(sigma)) #E.M. estimates

ests <- optim(par,ll.f1,gr=NULL,y=y3,n=n3,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE);ests
#-----
theta <- ests$par[1]
mu1 <- ests$par[2]
mu2 <- ests$par[3]
alpha <- ests$par[4]
#-----
p <- (1/(1+exp(-theta)))
sigma <- exp(alpha)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma"=sigma)
SE <- sqrt(diag(solve(ests$hessian)))

T3.est <- rbind("T3"=estimativas,"se.T3"=SE)

#-----T4
y4
n4
#Valores iniciais (E.M. estimates)
p <- 0.6
mu1 <- 2.49
mu2 <- 3.05
sigma <- 0.19
#---

par <- c(log(p/(1-p)),mu1,mu2,log(sigma)) #E.M. estimates

```



```

ests <- optim(par,ll.f1,gr=NULL,y=y4,n=n4,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE);ests
#-----
theta <- ests$par[1]
mu1 <- ests$par[2]
mu2 <- ests$par[3]
alpha <- ests$par[4]
#-----
p <- (1/(1+exp(-theta)))
sigma <- exp(alpha)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma"=sigma)
SE <- sqrt(diag(solve(ests$hessian)))

T4.est <- rbind("T4"=estimativas,"se.T4"=SE)

#-----T5
y5
n5
#Valores iniciais (E.M. estimates)
p <- 0.53
mu1 <- 2.47
mu2 <- 3.04
sigma <- 0.20
#---

par <- c(log(p/(1-p)),mu1,mu2,log(sigma)) #E.M. estimates

ests <- optim(par,ll.f1,gr=NULL,y=y5,n=n5,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE);ests
#-----
theta <- ests$par[1]
mu1 <- ests$par[2]
mu2 <- ests$par[3]
alpha <- ests$par[4]
#-----

```

```

p <- (1/(1+exp(-theta)))
sigma <- exp(alpha)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma"=sigma)
SE <- sqrt(diag(solve(ests$hessian)))

T5.est <- rbind("T5"=estimativas,"se.T5"=SE)

#-----T6
y6
n6
#Valores iniciais (E.M. estimates)
p <- 0.59
mu1 <- 2.59
mu2 <- 3.15
sigma <- 0.20
#---
p <- 0.59
mu1 <- 2.59
mu2 <- 3.15
sigma <- 0.23

par <- c(log(p/(1-p)),mu1,mu2,log(sigma)) #E.M. estimates

ests <- optim(par,ll.f1,gr=NULL,y=y6,n=n6,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE);ests
#-----
theta <- ests$par[1]
mu1 <- ests$par[2]
mu2 <- ests$par[3]
alpha <- ests$par[4]
#-----
p <- (1/(1+exp(-theta)))
sigma <- exp(alpha)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma"=sigma)
SE <- sqrt(diag(solve(ests$hessian)))

```

```

T6.est <- rbind("T6"=estimativas,"se.T6"=SE)

#-----
# Todos os tratamentos em uma tabela
#-----

T.trats <- rbind(T1.est,T2.est,T3.est,T4.est,
T5.est,T6.est)

T.trats <- round(T.trats,4)

for(i in 1:6){
T.trats[2*i-1,1] <- round(T.trats[2*i-1,1],2)
}

T.trats

require(xtable)
d <- xtable(T.trats,auto=TRUE)

##-----
##-----
#---- 2 sigmas
#-----

##-----Estimando para cada tratamento
# Reparametrização
# p <- (1/(1+exp(-theta)))
# sigma <- exp(alpha)
#-----2 variâncias
#-----T1
y1
n1
#Valores iniciais (Bayesian estimates - pesos individuais)
#
p <- 0.28
mu1 <- 2.32
mu2 <- 2.79

```

```

sigma1 <- 0.14
sigma2 <- 0.28
#--

par <- c(log(p/(1-p)),mu1,mu2,log(sigma1),log(sigma2)) #Bayesian estimates

ests <- optim(par,ll.f2,gr=NULL,y=y1,n=n1,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE);ests
#-----
theta <- ests$par[1]
mu1 <- ests$par[2]
mu2 <- ests$par[3]
alpha1 <- ests$par[4]
alpha2 <- ests$par[5]
#-----
p <- (1/(1+exp(-theta)))
sigma1 <- exp(alpha1)
sigma2 <- exp(alpha2)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma1"=sigma1,"sigma2"=sigma2)
SE <- sqrt(diag(solve(ests$hessian)))

T1.est.2sig <- rbind("T1"=estimativas,"se.T1"=SE)

#-----2 variâncias
#-----T2
y2
n2
#Valores iniciais (Bayesian estimates - pesos individuais)
# Problema na convergência impossibilitando o cálculo do st.err
p <- 0.5
mu1 <- 2.38
mu2 <- 2.91
sigma1 <- 0.25
sigma2 <- 0.3

par <- c(log(p/(1-p)),mu1,mu2,log(sigma1),log(sigma2)) #Bayesian estimates

```

```

ests <- optim(par,ll.f2,gr=NULL,y=y2,n=n2,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE);ests
#-----
theta <- ests$par[1]
mu1 <- ests$par[2]
mu2 <- ests$par[3]
alpha1 <- ests$par[4]
alpha2 <- ests$par[5]
#-----
p <- (1/(1+exp(-theta)))
sigma1 <- exp(alpha1)
sigma2 <- exp(alpha2)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma1"=sigma1,"sigma2"=sigma2)
SE <- sqrt(diag(solve(ests$hessian)))

T2.est.2sig <- rbind("T2"=estimativas,"se.T2"=SE)

#-----2 variâncias
#-----T3
y3
n3
#Valores iniciais (Bayesian estimates - pesos individuais)
# Problema na convergência impossibilitando o cálculo do st.err
#
p <- 0.6
mu1 <- 2.54
mu2 <- 3.11
sigma1 <- 0.22
sigma2 <- 0.4

par <- c(log(p/(1-p)),mu1,mu2,log(sigma1),log(sigma2)) #Bayesian estimates

ests <- optim(par,ll.f2,gr=NULL,y=y3,n=n3,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),

```

```

hessian = TRUE);ests
#-----
theta <- ests$par[1]
mu1 <- ests$par[2]
mu2 <- ests$par[3]
alpha1 <- ests$par[4]
alpha2 <- ests$par[5]
#-----
p <- (1/(1+exp(-theta)))
sigma1 <- exp(alpha1)
sigma2 <- exp(alpha2)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma1"=sigma1,"sigma2"=sigma2)
SE <- sqrt(diag(solve(ests$hessian)))

T3.est.2sig <- rbind("T3"=estimativas,"se.T3"=SE)

#-----2 variâncias
#-----T4
y4
n4
#Valores iniciais (Bayesian estimates - pesos individuais)
#
p <- 0.57
mu1 <- 2.48
mu2 <- 3.02
sigma1 <- 0.21
sigma2 <- 0.2

par <- c(log(p/(1-p)),mu1,mu2,log(sigma1),log(sigma2)) #Bayesian estimates

ests <- optim(par,ll.f2,gr=NULL,y=y4,n=n4,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE);ests
#-----
theta <- ests$par[1]
mu1 <- ests$par[2]

```

150

```
mu2 <- ests$par[3]
alpha1 <- ests$par[4]
alpha2 <- ests$par[5]
#-----
p <- (1/(1+exp(-theta)))
sigma1 <- exp(alpha1)
sigma2 <- exp(alpha2)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma1"=sigma1,"sigma2"=sigma2)
SE <- sqrt(diag(solve(ests$hessian)))

T4.est.2sig <- rbind("T4"=estimativas,"se.T4"=SE)

#-----2 variâncias
#-----T5
y5
n5
#Valores iniciais (Bayesian estimates - pesos individuais)
#
p <- 0.50
mu1 <- 2.47
mu2 <- 3.02
sigma1 <- 0.2
sigma2 <- 0.22
#
p <- 0.8
mu1 <- 2.47
mu2 <- 3.02
sigma1 <- 0.2
sigma2 <- 0.22

par <- c(log(p/(1-p)),mu1,mu2,log(sigma1),log(sigma2)) #Bayesian estimates

ests <- optim(par,ll.f2,gr=NULL,y=y5,n=n5,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),
hessian = TRUE);ests
#-----
theta <- ests$par[1]
```

```

mu1 <- ests$par[2]
mu2 <- ests$par[3]
alpha1 <- ests$par[4]
alpha2 <- ests$par[5]
#-----
p <- (1/(1+exp(-theta)))
sigma1 <- exp(alpha1)
sigma2 <- exp(alpha2)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma1"=sigma1,"sigma2"=sigma2)
SE <- sqrt(diag(solve(ests$hessian)))

T5.est.2sig <- rbind("T5"=estimativas,"se.T5"=SE)

#-----2 variâncias
#-----T6
y6
n6
#Valores iniciais (Bayesian estimates - pesos individuais)
#
p <- 0.55
mu1 <- 2.59
mu2 <- 3.11
sigma1 <- 0.2
sigma2 <- 0.22
#
p <- 0.80
mu1 <- 2.59
mu2 <- 3.11
sigma1 <- 0.2
sigma2 <- 0.22

par <- c(log(p/(1-p)),mu1,mu2,log(sigma1),log(sigma2)) #Bayesian estimates

ests <- optim(par,ll.f2,gr=NULL,y=y5,n=n5,method = "Nelder-Mead",
control = list(maxit = 20000, temp = 20),

```


152

```
hessian = TRUE);ests
#-----
theta <- ests$par[1]
mu1 <- ests$par[2]
mu2 <- ests$par[3]
alpha1 <- ests$par[4]
alpha2 <- ests$par[5]
#-----
p <- (1/(1+exp(-theta)))
sigma1 <- exp(alpha1)
sigma2 <- exp(alpha2)
#-----
estimativas <- c("p"=p,"mu1"=mu1,"mu2"=mu2,"sigma1"=sigma1,"sigma2"=sigma2)
SE <- sqrt(diag(solve(ests$hessian)))

T6.est.2sig <- rbind("T6"=estimativas,"se.T6"=SE)

Treat.table <- rbind(T1.est.2sig,T2.est.2sig,T3.est.2sig,
T4.est.2sig,T5.est.2sig,T6.est.2sig)

Treat.table <- round(Treat.table,4)

for(i in 1:6){
Treat.table[2*i-1,1] <- round(Treat.table[2*i-1,1],2)
}

Treat.table

require(xtable)
d <- xtable(Treat.table,auto=TRUE)

#-----1 variância (sigma)

#-----T1
y <- y1
n <- n1
```

```

N <- length(y)
Nx <- n1

fit1 <- stan(file = "newmix_1var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma_est=sqrt(0.04)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma_est=sqrt(0.02)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma_est=sqrt(0.01))),
warmup = 2000,
iter = 4000,
chains = 3,thin=10,
control = list(adapt_delta = 0.99, max_treedepth = 15))

print(fit1)
save(fit1, file = "fit.T1.1sig.rda")

fit.summary <- summary(fit1)
xtable(fit.summary$summary[1:4,1:8])

#-----Análise das estimativas
#posterior predictions
mcmc_chain = as.matrix(fit1)
#autocorrelation plot
par(mfrow=c(1,1))
acf(mcmc_chain[, "p_est"])
#par(mfrow=c(2,2))
acf(mcmc_chain[, "mu1_est"])
acf(mcmc_chain[, "mu2_est"])
acf(mcmc_chain[, "sigma_est"])
traceplot(fit1)
plot(fit1)
plot(fit1, show_density = TRUE, ci_level = 0.5, fill_color = "purple")
plot(fit1, plotfun = "hist", include = TRUE)
plot(fit1, plotfun = "trace", pars =
  c("p_est", "mu1_est", "mu2_est", "sigma_est"), inc_warmup = TRUE)
#-----
#-----T2
y <- y2
n <- n2

```

154

```
N <- length(y)
Nx <- n2

fit2 <- stan(file = "newmix_1var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma_est=sqrt(0.04)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma_est=sqrt(0.02)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma_est=sqrt(0.01))),
warmup = 2000,
iter = 4000,
chains = 3,thin=10,
control = list(adapt_delta = 0.99, max_treedepth = 15))

print(fit2)
save(fit2, file = "fit.T2.1sig.rda")
fit.summary <- summary(fit2)
xtable(fit.summary$summary[1:4,1:8])
#-----T3
y <- y3
n <- n3
N <- length(y)
Nx <- n3
fit3 <- stan(file = "newmix_1var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma_est=sqrt(0.04)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma_est=sqrt(0.02)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma_est=sqrt(0.01))),
warmup = 2000,
iter = 4000,
chains = 3,thin=10,
control = list(adapt_delta = 0.99, max_treedepth = 15))

print(fit3)
save(fit3, file = "fit.T3.1sig.rda")

fit.summary <- summary(fit3)
xtable(fit.summary$summary[1:4,1:8])
#-----T4
y <- y4
```

```

n <- n4
N <- length(y)
Nx <- n4
fit4 <- stan(file = "newmix_1var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma_est=sqrt(0.04)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma_est=sqrt(0.02)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma_est=sqrt(0.01))),
warmup = 2000,
iter = 4000,
chains = 3,thin=10,
control = list(adapt_delta = 0.99, max_treedepth = 15))

print(fit4)
save(fit4, file = "fit.T4.1sig.rda")

fit.summary <- summary(fit4)
xtable(fit.summary$summary[1:4,1:8])
#-----T5
y <- y5
n <- n5
N <- length(y)
Nx <- n5

fit5 <- stan(file = "newmix_1var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma_est=sqrt(0.04)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma_est=sqrt(0.02)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma_est=sqrt(0.01))),
warmup = 2000,
iter = 4000,
chains = 3,thin=10,
control = list(adapt_delta = 0.99, max_treedepth = 15))
print(fit5)
save(fit5, file = "fit.T5.1sig.rda")

fit.summary <- summary(fit5)
xtable(fit.summary$summary[1:4,1:8])
#-----T6

```

156

```
y <- y6
n <- n6
N <- length(y)
Nx <- n6

fit6 <- stan(file = "newmix_1var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma_est=sqrt(0.04)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma_est=sqrt(0.02)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma_est=sqrt(0.01))),
warmup = 2000,
iter = 4000,
chains = 3,thin=10,
control = list(adapt_delta = 0.99, max_treedepth = 15))

print(fit6)
save(fit6, file = "fit.T6.1sig.rda")

fit.summary <- summary(fit6)
xtable(fit.summary$summary[1:4,1:8])

#-----2 variâncias (sigma1 e sigma2)
#-----T1
y <- y1
n <- n1
N <- length(y)
Nx <- n1

fit1.2var <- stan(file = "newmix_2var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma1_est=sqrt(0.04),
sigma2_est=sqrt(0.02)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma1_est=sqrt(0.02),
sigma2_est=sqrt(0.01)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma1_est=sqrt(0.01),
sigma2_est=sqrt(0.001))),
warmup = 2000,
iter = 4000,
chains = 3,thin=10,
```

```

control = list(adapt_delta = 0.99, max_treedepth = 15))

print(fit1.2var)
save(fit1.2var, file = "fit1.T1.2sig.rda")

fit.summary <- summary(fit1.2var)
xtable(fit.summary$summary[1:5,1:8])

#-----Análise das estimativas
#posterior predictions
mcmc_chain = as.matrix(fit1.2var)
#autocorrelation plot
par(mfrow=c(1,1))
acf(mcmc_chain[, "p_est"])
#par(mfrow=c(2,2))
acf(mcmc_chain[, "mu1_est"])
acf(mcmc_chain[, "mu2_est"])
acf(mcmc_chain[, "sigma1_est"])
acf(mcmc_chain[, "sigma2_est"])
traceplot(fit1.2var)
plot(fit1.2var)
plot(fit1.2var, show_density = TRUE, ci_level = 0.5, fill_color = "purple")
plot(fit1.2var, plotfun = "hist", include = TRUE)
plot(fit1.2var, plotfun = "trace", pars = c("p_est", "mu1_est", "mu2_est",
      "sigma1_est", "sigma2_est"), inc_warmup = TRUE)
#-----
#-----T2
y <- y2
n <- n2
N <- length(y)
Nx <- n2

fit2.2var <- stan(file = "newmix_2var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma1_est=sqrt(0.04),
sigma2_est=sqrt(0.02)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma1_est=sqrt(0.02),
sigma2_est=sqrt(0.01)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma1_est=sqrt(0.01),

```

158

```
sigma2_est=sqrt(0.001))),
warmup = 2000,
iter = 4000,
chains = 3,thin=10,
control = list(adapt_delta = 0.99, max_treedepth = 15))

print(fit2.2var)
save(fit2.2var, file = "fit2.T2.2sig.rda")

fit.summary <- summary(fit2.2var)
xtable(fit.summary$summary[1:5,1:8])

#-----T3
y <- y3
n <- n3
N <- length(y)
Nx <- n3

fit3.2var <- stan(file = "newmix_2var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma1_est=sqrt(0.04),
sigma2_est=sqrt(0.02)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma1_est=sqrt(0.02),
sigma2_est=sqrt(0.01)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma1_est=sqrt(0.01),
sigma2_est=sqrt(0.001))),
warmup = 2000,
iter = 4000,
chains = 3,thin=10,
control = list(adapt_delta = 0.99, max_treedepth = 15))

print(fit3.2var)
save(fit3.2var, file = "fit3.T3.2sig.rda")

fit.summary <- summary(fit3.2var)
xtable(fit.summary$summary[1:5,1:8])

#-----T4
```

```

y <- y4
n <- n4
N <- length(y)
Nx <- n4

fit4.2var <- stan(file = "newmix_2var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma1_est=sqrt(0.04),
sigma2_est=sqrt(0.02)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma1_est=sqrt(0.02),
sigma2_est=sqrt(0.01)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma1_est=sqrt(0.01),
sigma2_est=sqrt(0.001))),
warmup = 2000,
iter = 4000,
chains = 3,thin=10,
control = list(adapt_delta = 0.99, max_treedepth = 15))

print(fit4.2var)
save(fit4.2var, file = "fit4.T4.2sig.rda")

fit.summary <- summary(fit4.2var)
xtable(fit.summary$summary[1:5,1:8])

#-----T5
y <- y5
n <- n5
N <- length(y)
Nx <- n5

fit5.2var <- stan(file = "newmix_2var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma1_est=sqrt(0.04),
sigma2_est=sqrt(0.02)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma1_est=sqrt(0.02),
sigma2_est=sqrt(0.01)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma1_est=sqrt(0.01),
sigma2_est=sqrt(0.001))),
warmup = 2000,

```


160

```
iter = 4000,
chains = 3,thin=10,
control = list(adapt_delta = 0.99, max_treedepth = 15))

print(fit5.2var)
save(fit5.2var, file = "fit5.T5.2sig.rda")

fit.summary <- summary(fit5.2var)
xtable(fit.summary$summary[1:5,1:8])

#-----T6
y <- y6
n <- n6
N <- length(y)
Nx <- n6

fit6.2var <- stan(file = "newmix_2var.stan",
data = list(y=y,n=n,N=N,Nx=Nx),
init=list(list(p_est=.3,mu1_est=2.5,mu2_est=2.0,sigma1_est=sqrt(0.04),
sigma2_est=sqrt(0.02)),
list(p_est=.4,mu1_est=2.8,mu2_est=2.2,sigma1_est=sqrt(0.02),
sigma2_est=sqrt(0.01)),
list(p_est=.5,mu1_est=3.0,mu2_est=2.4,sigma1_est=sqrt(0.01),
sigma2_est=sqrt(0.001))),
warmup = 2000,
iter = 4000,
chains = 3,thin=10,
control = list(adapt_delta = 0.99, max_treedepth = 15))

print(fit6.2var)
save(fit6.2var, file = "fit6.T6.2sig.rda")

fit.summary <- summary(fit6.2var)
xtable(fit.summary$summary[1:5,1:8])
```