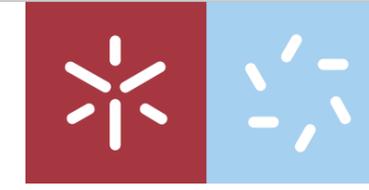


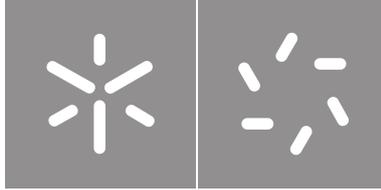


José Miguel Torres Maia

**Simulação de Monte Carlo de
sistemas óticos de alta qualidade,
utilizando distribuições de dados
reais para processos de produção**

Universidade do Minho
Escola de Ciências





Universidade do Minho

Escola de Ciências

José Miguel Torres Maia

**Simulação de Monte Carlo de
sistemas óticos de alta qualidade,
utilizando distribuições de dados
reais para processos de produção**

Dissertação de Mestrado

Mestrado em Estatística para Ciência de Dados

Trabalho efetuado sob a orientação de:

Professora Doutora Ana Paula Costa da

Conceição Amorim

Dr. João Tiago Costa Silva

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



Atribuição-NãoComercial-SemDerivações CC BY-NC-ND

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Agradecimentos

Em primeiro lugar, gostaria de expressar a minha gratidão aos meus pais, às minhas irmãs e ao meu irmão, que sempre me apoiaram e incentivaram ao longo do meu percurso académico. Foram, sem dúvida, uma inspiração ao longo destes anos.

Um enorme agradecimento aos meus orientadores, nomeadamente à Professora Doutora Ana Paula Amorim e ao Dr. João Tiago Silva, por toda a ajuda, paciência e partilha de conhecimentos.

Um agradecimento à Leica, em especial à equipa do Cálculo Ótico e aos restantes colaboradores do departamento de Tecnologia & Inovação, pela disponibilidade e pela forma como me integraram. Um obrigado aos restantes colaboradores que se cruzaram comigo e que, de alguma forma, me ajudaram durante esta etapa.

Por último, um agradecimento especial a todos os Professores, por contribuírem para a minha formação e desenvolvimento pessoal, e aos meus colegas, pela entreaajuda, companheirismo e por todos os bons momentos.

“O mundo não pode ser compreendido sem números.

E não pode ser compreendido só com números”

Hans Rosling

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

Resumo

Durante o desenvolvimento de um sistema ótico, na fase do *Design* para a produção, variam-se os parâmetros (características) de uma lente, definidos de acordo com as tolerâncias assumidas para cada parâmetro no processo de produção da lente, e verifica-se o desempenho do produto "*as-built*".

A atribuição de tolerâncias tem como objetivo determinar para que valores os parâmetros da lente minimizam os custos de produção e mantêm o desempenho pretendido do sistema ótico. As distribuições de probabilidade são úteis para a escolha de tolerâncias adequadas e para simular o desempenho esperado do sistema ótico real, por intermédio das simulações Monte Carlo.

O objetivo principal deste estudo é ajustar distribuições de probabilidade às características das lentes e desenvolver modelos para a previsão da espessura da lente na produção em série. Para isso, efetuaram-se medições do diâmetro, do erro de centragem e da espessura de vários tipos de lentes. No final, compararam-se as previsões efetuadas com as simulações de Monte Carlo para perceber que metodologia estava mais próxima das medições realizadas.

Para o ajuste das distribuições de probabilidade foram usados quatro métodos de estimação pontual dos parâmetros: o método de máxima verossimilhança (MMV), o *maximum goodness-of-fit* (MGOF), o método dos momentos (MM) e o método da correspondência dos quantis (MCQ). Para as distribuições com melhor desempenho, foram geradas amostras por *bootstrap*, com a mesma dimensão que as originais, para estimar os parâmetros e a respetiva incerteza associada.

Em seguida, para perceber que variáveis apresentavam maior relevância para a previsão, analisaram-se as correlações entre as características da lente e os parâmetros da distribuição, através do coeficiente de *Pearson*, para eleger as variáveis a incluir nos modelos. O método de reamostragem, *Cross-validation*, foi usado para escolher o melhor modelo para a previsão.

Por fim, concluiu-se que as distribuições de probabilidade associadas às medições das diferentes características da lente diferiram entre si. Para o caso da espessura, variaram de acordo com a tolerância associada à lente. Para esta característica, a distribuição apresentava-se, maioritariamente, à direita do valor nominal da lente.

Assim sendo, constatou-se que as simulações de Monte Carlo implementadas no *software* utilizado pela empresa conduziram a uma perceção distorcida dos valores da espessura, uma vez que os dados gerados por este método eram provenientes de uma distribuição Normal truncada, centrada no valor nominal e dentro do intervalo da tolerância.

Palavras-chave: Distribuições de probabilidade, *Bootstrap*, Previsões, Simulações de Monte Carlo

Abstract

During the development of an optical system, in the Design for production phase, the parameters of a lens are varied, defined according to the tolerances assumed for each one parameter the lens production process, and the performance of the as-built product is checked.

The purpose of assigning tolerances is to determine the values of lens parameters that minimise production costs and maintain the desired performance of the optical system. Probability distributions are useful for selecting appropriate tolerances and for simulating the expected performance of the real optical system using Monte Carlo simulations.

The main objective of this study is to fit probability distributions to the characteristics of lenses and to develop models for predicting lens thickness in serial production. To this end, measurements were taken of the diameter, centering error and thickness of different types of lenses. Then, the predictions made were compared with Monte Carlo simulations to see which was closest to the measured values obtained.

Four estimation methods were used to fit the probability distributions: maximum likelihood, maximum goodness-of-fit, moments matching and quantile matching estimation. For the distributions with the best performance, bootstrap samples of the same size as the originals were generated to estimate the parameters and their associated uncertainty.

Next, to understand which variables were most relevant to the prediction, the correlations between the lens characteristics and the distribution parameters were analysed using the Pearson coefficient, which was used to identify the variables to be included in the models. In addition, resampling methods such as cross-validation, were used to select the best model for prediction.

Finally, it can be concluded that the probability distributions associated with the measurements of the different lens characteristics differed from each other. As far as the thickness of the lens is concerned, they varied according to the tolerance assigned to the lens. For this characteristic, the distribution was mostly to the right of the nominal value of the lens.

Therefore, the results of the Monte Carlo simulations show that the software used by the company would give a biased perception of the thickness values, as the data generated by this method would be from a truncated normal distribution, centred on the nominal value and within tolerance.

Keywords: Probability distributions, *Bootstrap*, Predictions, Monte Carlo simulations

Índice

1	Introdução	1
1.1	Motivação	1
1.2	Leica - Aparelhos Ópticos de Precisão S.A	1
1.2.1	História	2
1.2.2	Áreas de produção	4
1.2.3	Tecnologia & Inovação	5
1.3	Estado da Arte	6
1.4	Objetivos	8
1.5	Estrutura da Dissertação	9
2	Ótica	11
2.1	Sistemas óticos	11
2.2	Qualidade de imagem	11
2.2.1	<i>Modulation Transfer Function</i>	12
2.2.2	Aberrações óticas	13
2.3	Interação da luz com a matéria	17
2.3.1	Refração e reflexão	18
2.3.2	Lente e a sua geometria	19
2.3.2.1	Diâmetro	20
2.3.2.2	Erro de centragem	21
2.3.2.3	Espessura	22
2.4	Processos de produção	22
2.4.1	Tolerâncias	23
2.4.2	Simulações de Monte Carlo	23
3	Metodologias estatísticas	25
3.1	Análise exploratória de dados	25
3.2	Distribuições de probabilidade	26
3.3	Métodos de estimação pontual	30
3.3.1	Método da máxima verosimilhança	31
3.3.2	Método dos momentos	31

3.3.3	Método da correspondência dos quantis	32
3.3.4	Método <i>maximum goodness-of-fit</i>	32
3.4	Critérios de informação	34
3.5	Estimação intervalar	34
3.6	Testes de ajustamento	35
3.7	Testes de hipóteses para normalidade	37
3.8	Regressão linear	37
3.9	Métodos de reamostragem	39
3.9.1	<i>Cross-validation</i>	40
3.9.2	<i>Bootstrap</i>	40
4	Análise Exploratória dos Dados	42
4.1	Diâmetro	44
4.2	Erro de centragem	47
4.2.1	Incertezas das medições	49
4.3	Espessura	51
4.3.1	Incerteza das medições	55
5	Ajustamento das distribuições de probabilidade	58
5.1	Estimação pontual	58
5.1.1	Diâmetro	58
5.1.2	Erro de centragem	59
5.1.3	Espessura	60
5.1.4	Desfasamento entre quantis teóricos e empíricos	62
5.2	Estimação intervalar	64
5.2.1	Incerteza associada à estimação dos parâmetros	68
5.2.2	Desfasamento entre quantis teóricos e empíricos	69
5.3	Testes de ajustamento	72
5.4	Qualidade do ajuste	75
6	Previsão dos parâmetros das distribuições de probabilidade	78
6.1	Tendências da distribuição da espessura	78
6.2	Correlações	82
6.3	Modelos para a previsão dos parâmetros	84

6.4	Previsões	87
6.5	Simulações de Monte Carlo	88
7	Conclusões	92
7.1	Trabalho futuro	94
A	Apêndice	98
A.1	Histogramas e <i>boxplots</i> para as medições da espessura nos Conjuntos E a F	98
A.2	Distribuições de probabilidade com melhor desempenho nos métodos de estimação pontual	100
A.3	Desfasamentos entre quantis teóricos e empíricos	103
A.4	Desfasamentos entre os quantis teóricos e empíricos para as distribuições obtidas por <i>bootstrap</i>	107
A.5	Qualidade do ajuste	109
A.6	Distribuições de probabilidade ajustadas ao Diâmetro e Erro de centragem	111
A.7	Distribuições de probabilidade ajustadas à Espessura	112
A.8	Distribuições de probabilidade previstas	113
A.9	Comparação dos modelos de previsão com as simulações Monte Carlo	115
A.10	Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais . .	116

Índice de Figuras

Primeira câmara 35mm compacta portátil (<i>Leica Portugal, 2023</i>).	2
Fábrica Leica de Portugal em 1973 (<i>Leica Camera Wetzlar, 2023</i>).	3
Fábrica Leica na atualidade, inaugurada em 2013 (<i>Leica Portugal, 2023</i>).	4
Produtos óticos comercializados pela Leica Camera AG (<i>Leica Camera Wetzlar, 2023</i>).	11
Capacidade de um sistema ótico em diferenciar zonas mais brilhantes e mais escuras (Eckhardt Optics, 2023).	12
Capacidade de um sistema ótico em resolver um padrão de linhas (Yu Kikuchi, 2023).	13
Difração consoante a abertura da lente, para o mesmo objeto e à mesma distância (Greg Hollows, Nicholas James, 2023).	14
Tipos de astigmatismo (Bruce MacEvoy, 2013).	15
Coma (Nasim Mansurov, 2019).	15
Curvatura de Campo (Bjorn Petersen, 2016).	15
Distorção.	16
Aberração esférica (Matt Osborne, 2019).	16
Tipos de aberrações cromáticas (Bruce MacEvoy, 2013).	17
Espetro eletromagnético.	17
Comportamento da luz quando interage com a superfície.	18
Comportamentos da luz consoante o tipo de lente com que interage.	19
Medições do diâmetro de uma lente (Bernd et al., 2012)	20
Medição do diâmetro de uma lente (NDC Technologies, 2015).	20
Ângulo de inclinação de uma superfície esférica (Bernd et al., 2012).	21
<i>Opticentric</i> da <i>Trioptics</i> (<i>Trioptics, 2023</i>).	21
Exemplo do processo de medição da espessura (Bernd et al., 2012).	22
Representação gráfica das medições do diâmetro.	45
Gráfico de Cullen e Frey para as medições do diâmetro do Conjunto B (Cullen et al., 1999).	46
Representação gráfica das medições do erro de centragem.	48
Gráfico de Cullen e Frey para as medições do erro de centragem no Conjunto A (Cullen et al., 1999).	49
Intervalos de confiança para as medições do erro de centragem no Conjunto B.	50
Representação gráfica das medições da espessura nos Conjuntos A a D.	53
Gráfico de Cullen e Frey para as medições da espessura no Conjunto G (Cullen et al., 1999)	55
Intervalos de confiança para as medições da espessura no Conjunto B.	56

Valores do parâmetro α estimados para as 1001 amostras do Conjunto B na distribuição Beta . . .	64
Distribuição empírica vs distribuição Weibull ajustada para o conjunto F.	67
Desfasamentos entre os quantis teóricos e empíricos para as medições do diâmetro nos Conjuntos A e B	70
Desfasamentos entre os quantis teóricos e empíricos para as medições do erro de centragem nos Conjuntos A e B	70
Desfasamentos entre os quantis teóricos e empíricos para as medições da espessura nos Conjuntos A ao D.	71
Qualidade de ajuste das distribuições ajustadas ao diâmetro nos Conjuntos A e B.	75
Qualidade de ajuste das distribuições ajustadas ao erro de centragem nos Conjuntos A e B.	76
Qualidade de ajuste das distribuições ajustadas à espessura nos Conjuntos A a D.	77
Distribuições de probabilidade consoante a dureza do material (Gross, 2007).	78
Distribuições de probabilidade para a espessura nos Conjuntos A a G.	80
Distribuições simétricas.	81
Distribuições assimétricas.	82
Correlações entre os parâmetros da lente e as características da distribuição dos dados.	83
Densidade das distribuições de probabilidade previstas para as lentes H a K.	87
Simulações de Monte Carlo geradas pelo <i>software</i> e função de densidade empírica.	88
Distribuições Normais ajustadas aos dados das simulações de Monte Carlo vs previsões pelos modelos.	89
Representação gráfica das medições da espessura do Conjunto E ao H.	98
Representação gráfica das medições da espessura do Conjunto I ao J.	99
Desfasamentos entre quantis teóricos e amostrais para a espessura dos Conjuntos E a H.	107
Desfasamentos entre quantis teóricos e amostrais para a espessura dos Conjuntos I a K.	108
Qualidade de ajuste das distribuições na espessura dos Conjuntos E a H.	109
Qualidade de ajuste das distribuições na espessura dos Conjunto I a J.	110
Distribuições de probabilidade para o diâmetro.	111
Distribuições de probabilidade para o erro de centragem.	111
Distribuições de probabilidade para a espessura dos Conjuntos de H a K.	112
Densidades previstas para as lentes A a D.	113
Densidades previstas para as lentes E a G.	114

Distribuições Normais ajustadas aos dados das simulações de Monte Carlo vs previsões pelos
modelos. 115

Índice de Tabelas

Estatísticas <i>goodness-of-fit</i> clássicas (D'Agostino, 2017)	33
Estatísticas <i>Anderson-Darling</i> modificadas (Luceño, 2006)	34
Valores nominais e tolerâncias para os parâmetros das lentes.	43
Medidas de localização para as medições do diâmetro.	44
Medidas de forma e dispersão para as medições do diâmetro.	44
Testes de normalidade para as medições do diâmetro.	45
Medidas de localização para as medições do erro de centragem.	47
Medidas de forma e dispersão para as medições do erro de centragem.	47
Testes de normalidade para as medições do erro de centragem.	48
Incerteza associada às medições do erro de centragem.	51
Medidas de localização para as medições da espessura.	51
Medidas de forma e dispersão para as medições da espessura.	52
Testes de normalidade para as medições espessura.	54
Incerteza associada às medições da espessura.	57
Distribuições que apresentaram melhor ajustamento para o diâmetro no Conjunto B.	59
Distribuições que apresentaram melhor ajustamento para o erro de centragem no Conjunto A.	59
Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto B.	60
Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto C.	60
Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto E.	61
Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto G.	61
Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto J.	61
Desfasamento entre os quantis teóricos e os empíricos para o diâmetro no conjunto A.	63
Desfasamento entre os quantis teóricos e os empíricos para o erro de centragem no conjunto A.	63
Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto A.	63
Médias dos parâmetros para as amostras do diâmetro	65
Medianas dos parâmetros para as amostras do erro de centragem.	65
Médias e medianas dos parâmetros para as amostras da espessura	66
Intervalos de confiança para os parâmetros das distribuições de probabilidade ajustadas ao diâmetro	68
Intervalos de confiança para os parâmetros das distribuições de probabilidade ajustadas ao erro de centragem.	68
Intervalos de confiança para os parâmetros das distribuições de probabilidade ajustadas à espessura.	69

Testes de ajustamento para as distribuições de probabilidade ajustadas ao diâmetro.	72
Testes de ajustamento para as distribuições de probabilidade ajustadas ao erro de centragem. . .	73
Testes de ajustamento para as distribuições de probabilidade ajustadas à espessura.	74
Distância do parâmetro de localização ao valor nominal.	79
Base de dados para a previsão dos parâmetros das distribuições de probabilidade ajustadas à espessura.	84
Erros associados à previsão dos parâmetros das distribuições de probabilidade da espessura. . .	86
Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto A.	90
Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto D.	90
Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto G.	90
Distribuições que apresentaram melhor ajustamento para o diâmetro no Conjunto A.	100
Distribuições que apresentaram melhor ajustamento para o erro de centragem no Conjunto B. . .	100
Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto A.	100
Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto D.	101
Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto F.	101
Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto H.	101
Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto I.	101
Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto K.	102
Desfasamento entre os quantis teóricos e os empíricos para o diâmetro no conjunto B.	103
Desfasamento entre os quantis teóricos e os empíricos para o erro de centragem no conjunto B. .	103
Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto B.	103
Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto C.	104
Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto D.	104
Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto E.	104
Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto F.	104
Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto G.	105
Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto H.	105
Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto I.	105
Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto J.	105

Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto K.	106
Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto B.	116
Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto F.	116
Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto H.	116
Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto I.	117
Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto J.	117
Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto K.	117

Siglas e Acrónimos

AD – *Anderson-Darling*

AD2 – *Anderson Darling de segunda ordem*

AD2L – *Left-tail Anderson Darling de segunda ordem*

AD2R – *Right-tail Anderson Darling de segunda ordem*

AIC – *Akaike's Information Criterion (em português, Critério de Informação de Akaike)*

ADL – *Left-tail Anderson Darling*

ADR – *Right-tail Anderson Darling*

BIC – *Bayesian Information Criterion (em português, Critério de Informação Bayesiano)*

CV – *Cross-Validation*

CvM – *Cramér-von Mises*

DMA – *Desfasamento Máximo Absoluto*

EAM – *Erro absoluto médio*

EPAM – *Erro percentual absoluto médio*

EQM – *Erro quadrático médio*

HK – *Knoop hardness (em português, Dureza)*

k-Fold CV – *k-Fold Cross-validation*

KS – *Kolmogorov-Smirnov*

LOOCV – *Leave-One-Out-Cross-Validation*

MCQ – *Método da correspondência dos quantis*

MGOF – *Maximum goodness-of-fit*

MM – *Método dos momentos*

MMV – *Método de máxima verosimilhança*

MTF – *Modulation Transfer Function*

REQM – *Raiz do erro quadrático médio*

1 Introdução

Este projeto foi realizado em contexto prático empresarial na Leica - Aparelhos Ópticos de Precisão S.A., no âmbito do Mestrado em Estatística para Ciência de Dados. Os principais objetivos deste estudo foram a recolha de medições das características das lentes, ajustamento de distribuições de probabilidade às medições e criação de um modelo para a previsão destes dados na produção em série.

Ao longo desta secção, será feita uma apresentação da entidade em que foi realizado o estágio, no âmbito deste projeto, onde serão apresentadas as secções e áreas de produção da empresa, bem como a sua missão e os seus valores. Apresenta-se, também, o enquadramento histórico da entidade para ilustrar a evolução da empresa.

Por fim, apresenta-se um resumo do estado da arte, os objetivos deste estudo e a estrutura da dissertação.

1.1 Motivação

Durante o processo de desenvolvimento ótico de novos produtos são feitas simulações Monte Carlo com o objectivo de estimar a variação do desempenho do produto, com todas as tolerâncias e compensações associadas à produção em série.

Esta simulação é tanto mais eficaz quanto melhor aproximadas são as distribuições aleatórias associadas às tolerâncias de produção. Neste momento são utilizadas distribuições normais para todos os parâmetros. No entanto, desconfia-se, com base na literatura, que estas podem não ser as distribuições mais adequadas para a medida das espessuras das lentes e, possivelmente para outras variáveis.

O objetivo deste projecto é, então, estudar as distribuições estatísticas de variáveis, como a medida da espessura das lentes, e definir modelos preditivos para novas produções de produtos diferentes.

1.2 Leica - Aparelhos Ópticos de Precisão S.A

A Leica tem como propósito a produção e o desenvolvimento de aparelhos óticos de precisão, que expandem os limites naturais da percepção visual daquilo que rodeia o utilizador. Para isso, surgem câmaras e lentes, óticas de desporto, onde se incluem binóculos *rangefinder* a *laser* e miras de elevada *performance*.

As fotografias registadas com as lentes Leica são motivo de destaque, pelo seu modo de capturar a

paisagem de forma natural, devido à qualidade do contraste e da resolução.

O foco da Leica tem sido as necessidades reais dos utilizadores, onde as suas opiniões são sempre bem-vindas e ponderadas antes de cada inovação. Os produtos são sempre idealizados para serem um equipamento capacitado para criar experiências visuais únicas.

1.2.1 História

No ano de 1849, em Wetzlar, na Alemanha, o jovem mecânico Carl Kellner funda o “*Optisches Institut*” (Instituto Ótico), onde começa por produzir óculos e telescópios. Mais tarde, é publicado um tratado sobre ótica, com o título “*Das orthoskopische Ocular, eine neu erfundene achromatische Linsencombination*” (“A ortoscopia ocular, uma combinação de lentes acromáticas recentemente inventada”) (Kellner, 1849). O sucesso deste documento impulsionou a produção de microscópios, instrumentos que viriam a destacar-se pela qualidade e precisão.

Em 1911, Oskar Barnack começa a trabalhar na Ernst Leitz - Wetzlar (designação da empresa nessa época). Nas suas horas livres, a sua atenção era direccionada para a construção de uma câmara portátil. No ano seguinte, o físico e matemático, Max Berek desenha a primeira lente para câmaras e, em 1913, Oskar Barnack cria a Ur-Leica (Figura 1.1(a)), o protótipo da primeira câmara de 35 mm considerada realmente portátil.

Seguidamente, em 1923, dá-se a pré-produção de máquinas capazes de suportar filmes de 35 milímetros (Figura 1.1(b)), com o intuito de serem testadas pela fábrica e também por fotógrafos independentes. A 3 de novembro foi registada a patente para um “*Rollfilmkamera*” (rolo de filme para câmara).



(a) Protótipo da Ur-Leica.



(b) Leica A.

Figura 1: Primeira câmara 35mm compacta portátil (Leica Portugal, 2023).

Apesar da avaliação pouco conclusiva, relativamente aos protótipos da Leica-A, em 1924, Ernst Leitz decidiu avançar a produção da mesma câmara, nomeando-a como Leica-I. Este modelo fez sucesso e impulsionou a produção de modelos seguintes, como as câmaras Leica Luxur e a Leica Compur.

No ano de 1938, Ernest Leitz II ajudou judeus a fugir da Alemanha, ao justificar que seriam enviados para departamentos comerciais da empresa noutros países, mesmo quando não se tratava de um funcionário da empresa.

Em 1973, dá-se a criação da fábrica de Portugal, enquanto “*extended Workbench*” da então Ernest Leitz Wetzlar, GmbH, com 92% de capital alemão e 8% a pertencer ao Banco Espírito Santo. A fábrica dá os primeiros passos numas pequenas instalações, em Famalicão, montando microscópios. No ano seguinte, muda de instalações, dentro do mesmo concelho, para o local onde se manteria durante quarenta anos.



Figura 2: Fábrica Leica de Portugal em 1973 (*Leica Camera Wetzlar*, 2023).

No ano de 1976, a fábrica de Portugal, que contava com cento e oitenta funcionários, inicia a produção e a montagem de binóculos e, em 1981, a fábrica ampliava as suas instalações, construindo um novo pavilhão na fábrica, destinado à secção da Montagem. No entanto, a sua expansão não ficava por aqui e, em 1987 deu-se a construção de um terceiro pavilhão, para funcionar a secção da Ótica. Nos anos seguintes, a produção continuava a desenvolver-se e, em 1992, começam a ser produzidas objetivas para os sistemas das máquinas R e M.

No ano de 2002, teve início, em Portugal, a produção das máquinas fotográficas M7 e R9. A M7 é uma máquina fotográfica analógica que teve bastante sucesso, tornando-se inclusive, líder de mercado. Além disso, foi instalada a unidade de cromagem em Portugal. No ano seguinte, iniciou-se a maquinação de magnésio na fábrica de Portugal.

Em 2004, teve início o *Programa A la Carte* com intuito de facilitar combinações personalizadas de acabamento metálico, tipo de cabedal, ampliação da ocular e gravação.

A 31 de março de 2013, a Leica inaugura a nova, e a atual, unidade fabril, também em Famalicão (Figura 3). As instalações ocupam uma área de 52 mil metros quadrados, dos quais 13,6 mil são reservados à produção.



Figura 3: Fábrica Leica na atualidade, inaugurada em 2013 (*Leica Portugal, 2023*).

Em maio do ano seguinte, a Leica Camera AG inaugura a sua nova fábrica em Am Leitz Park 1, na nova zona industrial de Wetzlar. Em Portugal, a produção geral anual atinge 40 mil binóculos, 15 mil objetivas, 4 mil miras telescópicas e 20 mil máquinas fotográficas.

A 1 de dezembro de 2016, foi inaugurada aquela que seria a primeira Leica Store em Portugal, no Porto.

Em 2023, assinalam-se os 50 anos da Leica Portugal e, em 2024, os 100 anos da primeira máquina fotográfica a ser concebida, a *Leica A*.

1.2.2 Áreas de produção

Os produtos Leica são o resultado de diversos processos produtivos complexos, divididos por quatro secções: Mecânica, Ótica, Montagem e Eletrónica.

A secção de mecânica divide-se em duas áreas produtivas: a Maquinagem e Tramentos de Superfície. Aqui, transformam-se matérias-primas, como alumínio, magnésio, latão ou aço, em componentes que servem todos os produtos Leica. A parte da Maquinagem está equipada com máquinas de CNC de última geração, mas também com equipamentos convencionais, necessários para desenvolver operações complexas e específicas que os produtos exigem. Na área do Tratamento de Superfície é realizado o polimento manual, proporcionando um acabamento singular de qualidade, graças a um trabalho artesanal. Após uma texturização da superfície mecânica através de areamento são aplicados diferentes revestimentos como anodização ou cromagem. A fase final é a pintura, feita

manualmente, de forma a garantir o maior rigor e qualidade neste processo.

Por sua vez, a secção da Ótica está dividida, também, em duas áreas produtivas. Na Ótica Plana são produzidos prismas para binóculos e para máquinas fotográficas. Na outra área, a Ótica Esférica, produzem-se as lentes para binóculos, miras e objetivas Leica. As lentes e os prismas são componentes óticos cujo objetivo é guiar e corrigir os raios de luz desde a entrada no aparelho, para que seja formada uma imagem perceptível ao olho do utilizador. Caso seja uma câmara, o objetivo deste processo destina-se a formar uma imagem perceptível num sensor ótico. A Ótica Plana existe apenas em Famalicão.

Por fim, na Montagem é obtido o produto Leica através de um processo de montagem manual e sequencial. Todos os equipamentos produzidos nesta linha dependem da entrada dos mesmos componentes óticos, mecânicos e eletrónicos.

1.2.3 Tecnologia & Inovação

O departamento de Tecnologia & Inovação foi formado em abril de 2015. Esta secção prima pela experiência em desenvolvimento do produto na Leica, assim como na proximidade às secções da Produção e da Montagem. Estes dois fatores potenciam o desenvolvimento dos produtos e beneficiem a produção e o cliente.

As diferentes áreas que integram este departamento são a Engenharia Mecânica, o *Design* Ótico, a Engenharia Eletrónica, a Industrialização e a Gestão de Projetos.

No que diz respeito à Mecânica, desenvolvem-se soluções mecânicas necessárias para a conceção de produtos Opto-mecânicos e eletrónicos. Além disso, é feita a construção em modelagem 3D e desenho técnico.

Já na Eletrónica, as atividades englobam o desenvolvimento de circuitos eletrónicos a incorporar nos produtos desenvolvidos e o desenvolvimento de *firmware*, *software* no dispositivo de *hardware* que executa funções como a entrada/saída ou a conexão a outros dispositivos.

No *Design* Ótico, no âmbito do *Design* de sistemas óticos, analisam-se as aberrações óticas, a luz difusa e os reflexos em sistemas opto-mecânicos. Além disso fazem-se simulações de tolerâncias e simulações de Monte Carlo.

As três áreas acima são responsáveis por desenhar o conceito de um produto capaz de cumprir as especificações do cliente tal como de uma produção em série. No entanto, de forma a tornar o conceito em produto, é necessário o departamento de engenharia industrial, responsável por planear a produção e montagem dos componentes, preparar todas as ferramentas de trabalho necessárias, bem como instruir os trabalhadores que produzirão o novo produto em série.

Por fim, é necessária uma área de gestão de projetos de forma a planejar e garantir a correta execução do projeto para um produto que nasce numa ideia, torna-se num conceito, é industrializado e tem um tempo de vida durante o seu ciclo de produção em série.

1.3 Estado da Arte

Um sistema ótico corresponde a um conjunto de elementos óticos individuais, como lentes ou prismas, e é concebido com o intuito de formar imagens com qualidade para o utilizador. Para isso, é necessário garantir a propagação da luz dentro do sistema (Zhang et al., 2017).

Segundo a teoria da formação da imagem de Fourier, a qualidade de um sistema ótico é descrita pela sua capacidade em transmitir frequências espaciais da estrutura do objeto (Gross, 2007). O desempenho é então quantificado pela *optical transfer function*, que indica o contraste da imagem e os detalhes do objeto em função da frequência, expressa em *line pairs* por milímetros. O valor absoluto desta função é designado por *Modulation Transfer Function* (MTF). Em suma, esta métrica de *performance* ótica avalia a capacidade da lente, ou de um sistema ótico, em transferir o contraste numa dada resolução, do objeto para a imagem (Gross, 2007).

No entanto, formar imagens perfeitas é praticamente impossível visto que qualquer sistema ótico apresenta determinadas aberrações, que vão reduzir a qualidade da imagem. Ao projetar um sistema ótico, o objetivo principal passa por reduzir estas aberrações de modo a que o desempenho do sistema seja satisfatório (Gross, 2005).

As aberrações óticas são fenómenos que degradam a qualidade da imagem, reduzindo o valor de MTF medido. Historicamente, foram definidas a partir de combinações de cinco aberrações primárias e duas relacionadas com dispersões cromáticas. As aberrações primárias ocorrem devido a imperfeições relativas ao *design* ótico da lente, enquanto que as aberrações cromáticas surgem devido ao fenómeno da dispersão para diferentes comprimentos de onda, uma vez que este está dependente do índice de refração do material (Gross, 2005).

Além disso, a qualidade de imagem é também influenciada pela luz, nomeadamente pelo seu comportamento ao atingir a superfície e ao propagar-se pela mesma. A formação de imagens está dependente de fenómenos relacionados com o comportamento da luz, como a reflexão e refração, e com a geometria da lente.

O sistema ótico produzido contém sempre alguns desvios face à versão projetada. Estes erros de produção não podem ser evitados, mas podem ser controlados até certo ponto. Porém, o controlo destes

erros requer custos adicionais e a atribuição de tolerâncias ao *design* de um sistema ótico determinará a que custo o sistema poderá ser produzido (Gross, 2007).

Koch (1978) considera que o objetivo de qualquer processo de atribuir tolerâncias passa por determinar o intervalo ou conjunto de valores para as características da lente que minimizam os custos de produção e mantêm o desempenho pretendido do sistema ótico. Segundo Hasenauer (2015) as tolerâncias mais apertadas devem aplicar-se às características cuja pequena mudança causa maior degradação do desempenho do sistema ótico.

A modelação com distribuições de probabilidade pode ser útil para a escolha de distribuições de tolerâncias adequadas para casos em que não existem dados estatísticos de produção disponíveis. Perceber a relação entre as tolerâncias estabelecidas e o erro de produção é bastante útil para simular a rentabilidade da produção dos sistemas óticos através das Simulações de Monte Carlo. Além disso, tais simulações, são usadas para simular o desempenho esperado do sistema ótico real (Gross, 2007).

As funções de densidade probabilidade mais importantes, no âmbito do *software* de *design* ótico, são a Normal, truncada ou não truncada, a Uniforme e a Bernoulli (referida na literatura como distribuição "*Ping-pong*") (Gross, 2007), a definir na secção (3.2). No entanto, as densidades uniformes podem originar a atribuição de tolerâncias mais conservadoras, desde a espessura, ou os raios, até mesmo ao índice de refração. O conhecimento das distribuições dos erros de fabrico, para os diferentes tipos de produção, torna-se útil caso, por razões de custo, seja necessário diminuir os valores de tolerância (Gross, 2007).

As distribuições podem ser usadas em simulações estatísticas durante o processo de estabelecer as tolerâncias. A Normal ou a Uniforme são distribuições que, geralmente, são assumidas nestes processos de tolerância (Youngworth & Stone, 2000) ou usadas nas simulações Monte Carlo (Smith & Rimmer, 1999).

Segundo Tienvieri & Rich (1998) e Gross (2007) a tendência da distribuição da espessura é que seja deslocada para a direita, conferindo-lhe uma assimetria negativa. Existem dois motivos para isso. Quando a lente é produzida dentro da tolerância para todas as características, é possível que surjam impurezas ou arranhões. Assim sendo, estas superfícies têm de ser polidas, para garantir a qualidade da imagem da lente. No entanto, ao polir, a espessura da lente é reduzida. No pior dos casos, a superfície é perfeita, depois do polimento, mas a espessura da lente é menor que o permitido. Assim sendo, é uma prática comum tentar manter a espessura da lente mais próxima do limite superior de tolerância para permitir o polimento, caso necessário. A outra razão é que parar de polir e lixar assim que a lente esteja dentro dos limites de tolerância evita que estes processos se repitam quando já não são necessários, diminuindo o

custo de produção. Além disso, quanto menos dureza apresentar o vidro, mais ampla será a distribuição em torno do valor central.

1.4 Objetivos

As lentes desempenham um papel muito importante no nosso dia a dia, pela sua propriedade de formar imagens de objetos, graças ao seu efeito de foco. Isto acontece porque a velocidade de propagação da luz na lente é menor do que no ar, de modo a permitir a refração de um feixe de luz, quando entra e sai da lente. Podem ser usadas de forma singular, como em óculos, lentes de contacto, lupas de bolso, condensadores de projeção ou luzes de sinalização. Adicionalmente, podem ainda ser usados conjuntos de lentes, feitas de diferentes materiais, combinadas para corrigir possíveis aberrações que possam surgir. São exemplos disso instrumentos como câmaras, microscópios ou telescópios.

O processo de desenvolvimento do sistema ótico engloba três fases distintas: o *Design* nominal do sistema, o *Design* para produção e, por fim, a Produção do sistema e validação do mesmo.

No *Design* nominal do sistema selecionam-se as características (variáveis) para obter a *performance* final pretendida. Na fase seguinte, no *Design* para produção, variam-se as características definidos de acordo com as tolerâncias garantidas para cada uma no seu processo de produção e verifica-se o desempenho do produto "*as-built*". Depois, na fase final, produz-se e valida-se o sistema ótico.

O foco deste projeto será a fase intermédia e o objetivo principal é estudar as distribuições destas variáveis e desenvolver modelos que permitam prever o comportamento dos dados na produção em série.

Para isso, a fase inicial deste estudo passa pela medição das características da lente, que influenciam o resultado final de *performance* ótica da mesma. As características a serem medidas são o diâmetro, a espessura e o erro de centragem e, posteriormente, será calculada a incerteza associada às medições (erro de centragem e espessura).

Seguidamente, serão ajustadas distribuições de probabilidade às medições dos características das lentes e procurar-se-ão identificar tendências entre as distribuições e as características das lentes. Assim sendo, é pretendido entender a correlação existente entre as diferentes características da lente, que constam na ficha técnica, e os parâmetros das distribuições.

Depois, o objetivo será encontrar e ajustar modelos para a previsão das dimensões das características das lentes para novas produções.

Por fim, pretende-se comparar as simulações de Monte Carlo com as previsões efetuadas, para perceber quais estão mais próximas das medições realizadas.

1.5 Estrutura da Dissertação

Inicialmente, foi feita uma apresentação da Leica - Aparelhos Ópticos de Precisão S.A. Foram abordados conceitos relativos ao passado histórico da entidade, bem como as áreas de produção atuais, passando também pela secção em que o estágio foi desenvolvido. Ainda no primeiro capítulo, efetuou-se um enquadramento das áreas e tópicos a serem estudados e apresentaram-se os objetivos da dissertação.

No Capítulo 2 serão apresentados conceitos relativos aos sistemas óticos. Neste capítulo, o foco passa por entender o funcionamento e a utilidade de um sistema ótico, bem com a finalidade das lentes como parte do mesmo. No final do capítulo, abordar-se-ão conceitos relativos à produção dos sistemas óticos, nomeadamente as tolerâncias das características da lente. Posteriormente, serão abordadas as simulações de Monte Carlo, uma metodologia fundamental para a perceção do desempenho do produto.

O Capítulo 3 é destinado às metodologias estatísticas usadas ao longo deste projeto. Serão apresentadas as distribuições de probabilidade, usadas para o ajuste, bem como os métodos utilizados para a estimação dos seus parâmetros. Além disso, serão abordados os testes de ajustamento e conceitos relativos à regressão linear e aos métodos de reamostragem, que foram utilizados no decorrer deste projeto.

No Capítulo 4 é apresentada a análise exploratória dos dados. Uma breve caracterização dos dados, será realizada, nomeadamente das medições do diâmetro, do erro de centragem e da espessura. Neste capítulo, apresentar-se-ão conclusões relativas à forma dos dados, úteis para o capítulo seguinte, bem como ilações relativamente à incerteza associada às medições do erro de centragem e da espessura.

O Capítulo 5 abrange o processo de ajustamento das distribuições de probabilidade às medições recolhidas. Neste capítulo serão apresentados resultados das distribuições com melhor ajuste, em cada um dos conjuntos de dados, bem como a estimação dos parâmetros das respetivas distribuições. A qualidade do ajuste será averiguada através dos testes de ajustamento e de representações gráficas.

No Capítulo 6 são apresentadas as tendências que se verificaram nas distribuições de probabilidade associadas às medições da espessura, tal como as correlações entre as características das lentes e os parâmetros das distribuições ajustadas. Seguidamente, são apresentados os modelos de previsão para os parâmetros das distribuições da espessura e as respetivas previsões. No final, foram comparadas as previsões dos modelos com as simulações Monte Carlo, tendo como referência as medições que foram efetuadas da espessura das lentes.

Por fim, no sétimo capítulo, serão abordadas as principais conclusões fruto desta dissertação e o

trabalho a ser desenvolvido futuramente.

2 Ótica

Neste capítulo serão apresentados conceitos teóricos de ótica. Ao longo deste capítulo será explicitado o conceito de sistema ótico de imagem, que tem como objectivo principal criar imagens para o utilizador. A fim de entendermos, como funciona um sistema óptico, serão abordados fenómenos de interação da luz com a matéria e, posteriormente, apresentadas as lentes que são os meios utilizados para influenciar esta interação através do seu material e geometria.

Compreendido o funcionamento de um sistema ótico a nível teórico, é importante perceber como funcionam os processos de produção. Quando queremos contruir um sistema ótico produzível em série teremos sempre de associar tolerâncias às características, de acordo com as capacidades dos processos de produção associados. Posteriormente, as tolerâncias são utilizadas em simulações Monte Carlo, para perceber o desempenho do sistema ótico final.

2.1 Sistemas óticos

A Leica Camera AG é uma empresa dedicada à comercialização de produtos óticos, especificamente aqueles que produzem imagens para o utilizador. São exemplos muito conhecidos as câmaras fotográficas e as suas objectivas associadas, os binóculos, os telescópios, as miras telescópicas e, ainda, os binóculos e monóculos com capacidade de medir distâncias. Na figura 4, podemos ver alguns exemplos destes produtos, comercializados pela Leica.



Figura 4: Produtos óticos comercializados pela Leica Camera AG (*Leica Camera Wetzlar, 2023*).

2.2 Qualidade de imagem

Com toda a mística associada, nunca podemos dissociar a marca Leica à sua imagem de marca central: a qualidade de imagem.

A qualidade de imagem é de facto um dos principais factores que leva os produtos Leica a atingirem o nível de reputação que têm.

No entanto, uma imagem real formada através um sistema ótico não é perfeita e essas imperfeições podem ser caracterizadas pela resolução e contraste. A capacidade que um sistema tem a resolver os mais pequenos detalhes é denominada de resolução. E quanto mais elevada for a resolução de um sistema, maior será a sua capacidade de resolver dois objectos próximos. Por outro lado, o contraste está relacionado com a capacidade de um sistema diferenciar zonas mais brilhantes e mais escuras.

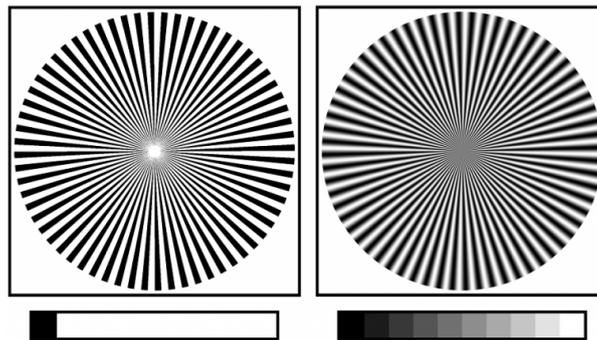


Figura 5: Capacidade de um sistema ótico em diferenciar zonas mais brilhantes e mais escuras (Eckhardt Optics, 2023).

2.2.1 Modulation Transfer Function

Quando descrevemos contraste e resolução anteriormente, utilizamos definições subjectivas para os termos. No entanto, e para podermos controlar e especificar um produto, devemos utilizar métricas objectivas que sejam possíveis de aferir na prática. É aqui que entra a métrica *Modulation Transfer Function* (MTF).

A MTF é uma métrica adimensional, de 0 a 1, ou em percentagem, que descreve a capacidade de um sistema resolver determinado padrão de pares de linhas espaçados de uma certa distância, o qual denominamos frequência.

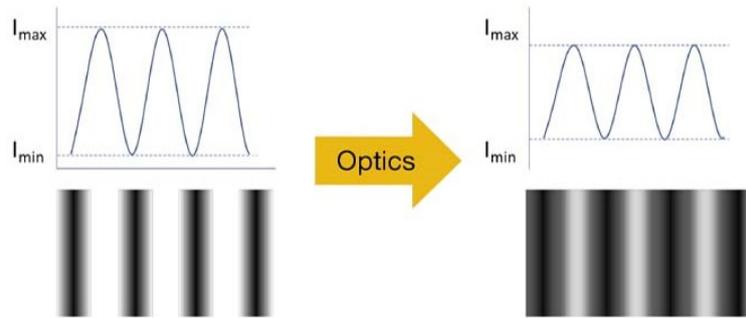


Figura 6: Capacidade de um sistema ótico em resolver um padrão de linhas (Yu Kikuchi, 2023).

Definimos o MTF para uma determinada frequência como na Fórmula (2.1):

$$MTF = \frac{(I_{max} - I_{min})}{(I_{max} + I_{min})}, \quad (2.1)$$

sendo I_{max} e I_{min} as intensidades máximas e mínimas do objeto, respetivamente.

A métrica MTF é bastante útil para descrever a qualidade de imagem de um sistema ótico e pode ser obtida para diferentes frequências, dando uma medida de contraste e resolução. Pode, também, ser medida ao longo do campo de imagem, desde o centro até à periferia pode assumir valores diferentes, mostrando como varia a qualidade de imagem ao longo do campo de imagem.

2.2.2 Aberrações óticas

A dada altura podemos perguntar-nos o porquê da imagem que formamos não replicar perfeitamente o objecto, ou seja, o MTF para frequências superiores a 0 não tomará o valor de 1. Isto acontece devido aos princípios físicos da interação entre luz e matéria.

Os sistemas óticos com a melhor performance são denominados "*diffraction limited*", o que quer dizer que a única limitação que têm é dada pelo fenómeno da difração.

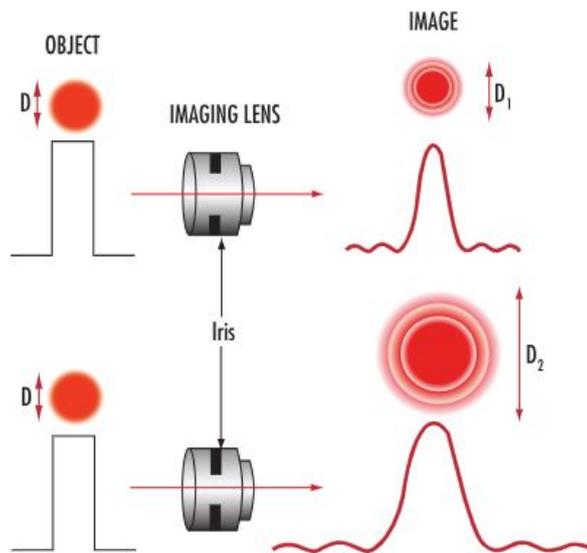


Figura 7: Difração consoante a abertura da lente, para o mesmo objeto e à mesma distância (Greg Hollows, Nicholas James, 2023).

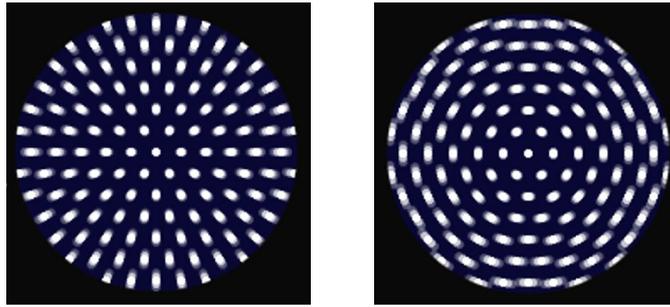
A difração ocorre quando os raios de luz se espalham, quando passam por uma abertura. Este fenómeno, para além de depender do comprimento de onda da luz, depende também da dimensão da abertura lente. A difração será menor para aberturas menores. Relativamente ao comprimento de onda, quanto menor for, menor será a difração. Este conceito é ilustrado na Figura 7.

No entanto, a maior parte dos sistemas não são apenas degradados pelo fenómeno acima descrito, mas também por aquilo a que chamamos de aberrações óticas.

As aberrações óticas são fenómenos óticos que degradam a qualidade de imagem, reduzindo o valor da MTF medido. Historicamente, foram definidas através de combinações de cinco aberrações primárias: Esférica, Coma, Astigmatismo, Curvatura de Campo e Distorção e duas relacionadas com dispersões cromáticas: longitudinal e lateral.

Em seguida, exemplificam-se os vários tipo de aberrações primárias, ilustradas desde a Figura 8 à 12.

A Figura 8 representa o Astigmatismo. Esta aberração provoca que as linhas horizontais e verticais do objecto sejam focados em planos diferentes na imagem. Traduzindo-se, desta forma, num efeito desfocado de uma das direções.



(a) Astigmatismo tangencial.

(b) Astigmatismo sagital.

Figura 8: Tipos de astigmatismo (Bruce MacEvoy, 2013).

A Figura 9 representa o coma. Esta aberração, tal como o nome indica, provoca na imagem que os objectos circulares sejam reproduzidos na imagem com um arrasto que faz lembrar a forma de um cometa.



Figura 9: Coma (Nasim Mansurov, 2019).

Na Figura 10 apresenta-se a curvatura de campo. Esta aberração é responsável pelo desfoque de objectos que se encontram no mesmo plano mas em posições diferentes do campo de imagem.



Figura 10: Curvatura de Campo (Bjorn Petersen, 2016).

A distorção, ilustrada na Figura 11 é uma aberração que não degrada a qualidade de imagem, no entanto, altera a forma da imagem, visto que os pontos do objecto se formam em posições diferentes da imagem. Esta é uma aberração que pode ser corrigida nas câmeras modernas porque não degrada a informação da imagem.

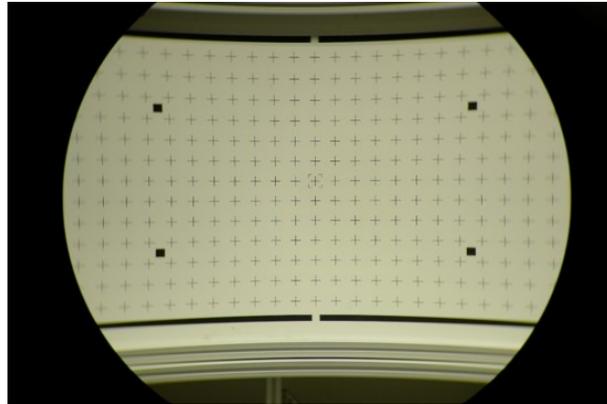


Figura 11: Distorção.

A aberração esférica, apresentada na Figura 12, produz uma degradação da imagem, reduzindo a resolução e o contraste da imagem sem alterar os planos de foco, nem produzir imperfeições assimétricas como são o coma e o astigmatismo.



Figura 12: Aberração esférica (Matt Osborne, 2019).

Por sua vez, as aberrações cromáticas, representadas na Figura 13, são caracterizadas pela incapacidade da lente em focar os vários comprimentos de onda da cor no mesmo ponto. Este fenómeno ocorre com mais frequência em casos de alto contraste, onde a lente é incapaz de corrigir esta aberração, provocando imagens com uma névoa colorida nas extremidades dos objetos (Figura 13).

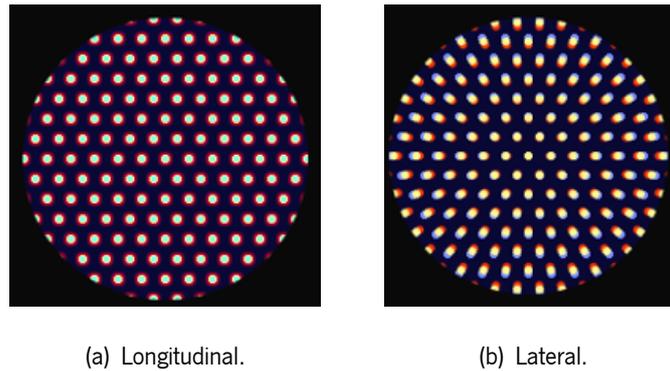


Figura 13: Tipos de aberrações cromáticas (Bruce MacEvoy, 2013).

Na aberração cromática longitudinal, ilustrada na Figura 2.13(a), dá-se um desfoque da cor nas extremidades dos pontos.

Por outro lado, na aberração cromática lateral, representada na Figura 2.13(b), verifica-se uma distorção na cor do ponto, que é notória à medida que os pontos se afastam do centro.

2.3 Interação da luz com a matéria

A luz corresponde a radiação que oscila pelo campo eletromagnético, propagando-se pelo espaço como uma onda. O espectro visível ao olho humano corresponde aos comprimentos de onda entre os 400 e os 750 nanómetros (nm). No que diz respeito à velocidade de propagação, esta varia consoante o meio de propagação. No vácuo, a luz propaga-se à velocidade de 3×10^8 m/s.

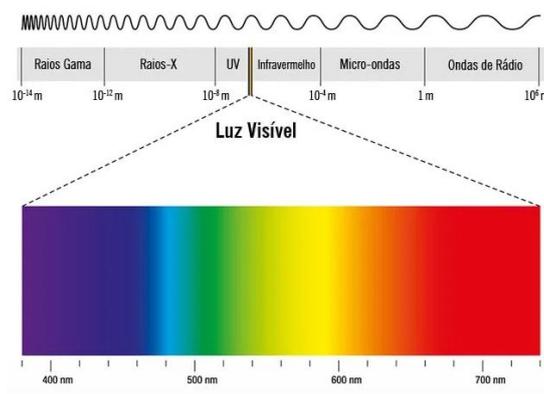


Figura 14: Espectro eletromagnético.

Quando a luz incide numa superfície, entre dois materiais diferentes, existem dois comportamentos possíveis, a refração e a reflexão. Dentro do material é possível que ocorra um outro fenómeno, a

absorção, não tão relevante para o âmbito deste projeto.

2.3.1 Refração e reflexão

A refração e a reflexão são fenômenos óticos relacionados com a forma como a luz se comporta ao atravessar de um material para outro, com diferentes índices de refração.

De acordo com os princípios de ótica geométrica, a lei da refração, ou Lei de Snell, descreve a mudança de direção de um raio de luz ao passar numa superfície entre dois meios. A forma mais comum de a representar é através da equação (2.2), que relaciona os índices de refração dos meios e os respetivos ângulos do raio com a superfície, o de incidência e o de refração.

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (2.2)$$

O índice de refração n_i é definido pela razão entre a velocidade da luz no vácuo (c) e a velocidade da luz nesse meio (c_i). O índice de refração de um determinado material i é dado pela expressão (2.3).

$$n_i = \frac{c_i}{c} \quad (2.3)$$

Na Figura 15 está representado um esquema que ilustra o comportamento de um raio quando atinge uma superfície com um índice de refração diferente, nos fenômenos da refração e da reflexão.

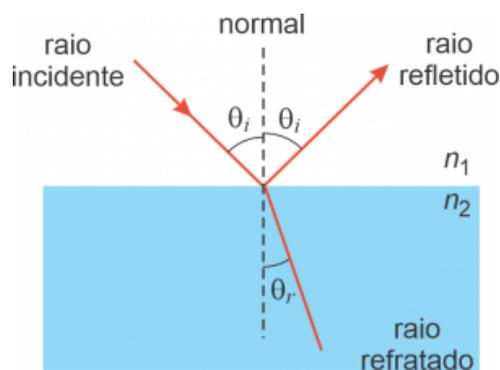


Figura 15: Comportamento da luz quando interage com a superfície.

Assim sendo, como ilustra a Figura 15, a reflexão ocorre quando o raio de luz é projetado no sentido oposto ao raio de incidência. Existem dois tipos de reflexão, a regular ou a difusa. Uma reflexão diz-se regular se, como de acordo com a fig 15, o raio reflectido obedece à condição $\theta = \theta$. Este tipo de reflexão é típico de superfícies polidas com baixos níveis de rugosidade. Superfícies mais rugosas refletem os raios

No âmbito deste estudo, foram medidas as seguintes características da lente: diâmetro, erro de centragem e espessura.

2.3.2.1 Diâmetro

O diâmetro de uma lente corresponde ao comprimento da linha reta que passa pelo centro. No entanto, como a lente não forma um círculo perfeito, podem ser considerados, durante a medição, diâmetros de diferentes dimensões, como ilustra a Figura 17.

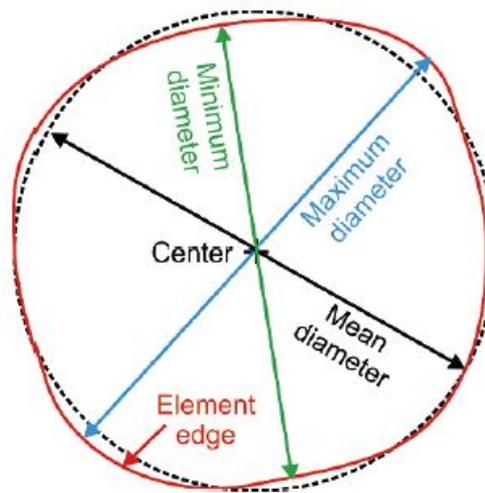
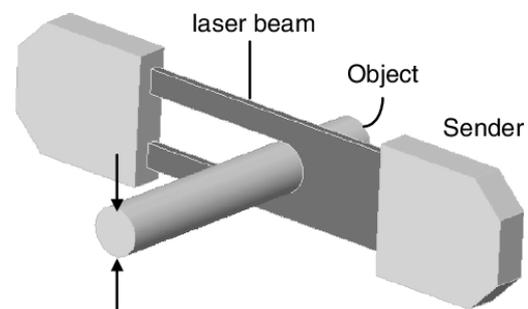


Figura 17: Medições do diâmetro de uma lente (Bernd et al., 2012)

As medições do diâmetro foram efetuadas na Centragem, na secção da Ótica. O instrumento utilizado foi o micrómetro *PrecisionPro 6000 Z-Mike by Beta LaserMike* (Figura 2.18(a)). O valor é obtido através da interrupção do feixe do laser (Figura 2.18(b)).



(a) PrecisionPro 6000 Z-Mike by Beta LaserMike.



(b) Medição do diâmetro de um objeto através da interrupção do laser.

Figura 18: Medição do diâmetro de uma lente (NDC Technologies, 2015).

A superfície da lente com maior curvatura deve estar virada para cima. Depois, verifica-se se o laser está a meio da espessura da lente, na parte lateral, e só depois se faz a medição. Após a primeira medição, roda-se a lente e efetua-se outra medição, com o mesmo procedimento.

2.3.2.2 Erro de centragem

O erro de centragem de uma superfície esférica é dado pelo ângulo de inclinação da superfície, entre o eixo normal à superfície e o eixo de referência (Bernd et al., 2012), que está dependente das características do sistema ótico. A unidade de medida é o minuto ($'$).

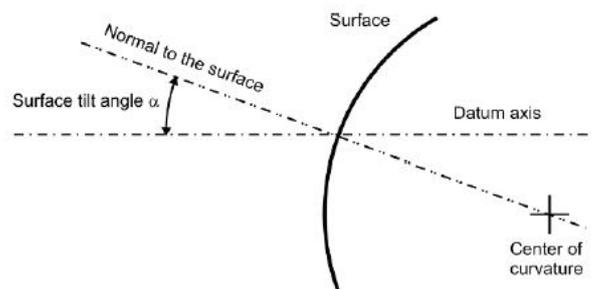


Figura 19: Ângulo de inclinação de uma superfície esférica (Bernd et al., 2012).

As medições foram efetuadas através do *Opticentric* da *Trioptics* (Figura 20), um instrumento de grande precisão para efetuar testes de centragem ótica, alinhamento e colagem, seja ela manual ou automática, para uma lente ou um conjunto.



Figura 20: *Opticentric* da *Trioptics* (Trioptics, 2023).

2.3.2.3 Espessura

A espessura de uma lente diz respeito à distância entre os pontos centrais das duas superfícies. Estes pontos coincidem com o eixo ótico da lente. Para lentes esféricas, o eixo ótico passa pelos centros de curvatura.

Para a medição deste parâmetro foi usado um comparador. Para garantir que a medição é efetuada da forma correta, é necessário que a lente seja colocada no centro e a superfície do instrumento não pode ser plana. A superfície do instrumento que contacta com a superfície da lente deve ser esférica, de forma a garantir o contacto de um ponto apenas e o material usado é normalmente o rubi pela sua dureza para garantir que não é facilmente danificado. Se esta medição for feita com pontas danificadas, facilmente se risca a lente.

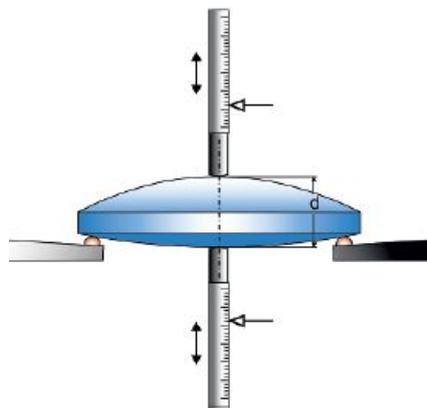


Figura 21: Exemplo do processo de medição da espessura (Bernd et al., 2012).

2.4 Processos de produção

Os processos produtivos não conseguem garantir que os valores de uma determinada característica da lente sejam iguais, daí serem atribuídas tolerâncias segundo as quais esse parâmetro pode estar compreendido.

Conhecendo o processo de fabrico, podemos adequar um sistema ao mesmo. Isto é, sabendo a capacidade do processo podemos dessensibilizar o sistema de forma a que a sua *performance* se mantenha, dada a alteração das características específicas. Desta forma, é possível controlar os custos associados à produção de determinado produto Gross (2007).

É importante definir tolerâncias apertadas apenas para aqueles parâmetros cuja pequena mudança causaria uma grande degradação na *performance* (Hasenauer, 2015).

2.4.1 Tolerâncias

O objetivo é encontrar um determinado conjunto de tolerâncias que definam as precisões com que os constituintes de um instrumento ótico devem ser fabricados e montados, para que o instrumento final produzido tenha a *performance* ótica idealizada.

Numa fase inicial, é realizado um estudo de sensibilidades do sistema. Desta forma, é possível saber qual a *performance* expectável do sistema final com a variação de determinado parâmetro. Considerando as variações dos parâmetros individuais que mantenham o desempenho final dentro dos limites considerados, é avaliada a executabilidade dessas tolerâncias nos processos produtivos que serão utilizados. Quando essas tolerâncias não são possíveis de alcançar com o processo idealizado para tal, é necessário encontrar alternativas, que podem passar por compensadores ou um novo conceito para o produto.

Por fim, um sistema ótico real não é apenas afetado por uma única tolerância, mas sim por todo o conjunto. Para a análise da qualidade de imagem de um sistema produzido, torna-se necessário o cálculo ou a simulação dos efeitos na qualidade de imagem de todas as tolerâncias atribuídas. Este processo pode ser realizado através de modelos estatísticos e simulações Monte Carlo.

2.4.2 Simulações de Monte Carlo

As tolerâncias atribuídas e os processos de compensação escolhidos são o ponto de partida para uma simulação mais detalhada das tolerâncias, as simulações Monte Carlo (Gross, 2007).

As simulações de Monte Carlo são usadas para simular o desempenho esperado de um sistema ótico real. Em cada um dos seus ciclos, são gerados valores para os parâmetros a fabricar, de forma aleatória, dentro das tolerâncias estabelecidas, de acordo com as distribuições de probabilidade, possibilitando a previsão do desempenho do sistema "*as-built*".

Para uma simulação de Monte Carlo é necessário que seja definida a função *Merit*:

$$Merit = \sum_k g_k^2 \cdot (b_k - a_k)^2 \quad (2.4)$$

sendo b_k a medida de desempenho k , g_k o peso da medida de desempenho k e a_k o objetivo para a medida de desempenho k .

É escolhido um valor máximo para esta função que terá efeito no rendimento da produção. Isto é, caso o valor da função seja superior ao máximo estabelecido, o sistema é rejeitado, caso contrário é aceite.

Deste modo, é possível estimar a taxa de rejeição e o rendimento de produção através de um conjunto de distribuições tolerâncias.

Geralmente, a Normal e a Uniforme são as distribuições de probabilidade mais usadas para as simulações de Monte Carlo e para modelar as tolerâncias (Gross, 2007), no entanto, nem sempre refletem de forma correta os dados.

3 Metodologias estatísticas

Neste capítulo são abordados conceitos teóricos relacionados com as metodologias estatísticas usadas ao longo deste projeto.

Na primeira secção, (3.1), explicitam-se conceitos relativos à análise exploratória de dados.

De seguida, na secção (3.2) são apresentadas as distribuições de probabilidade teóricas que foram ajustadas aos dados. Depois, na secção (3.3) são abordados métodos de estimação pontual dos parâmetros das distribuições de probabilidade. Na secção (3.4), são apresentados os critérios de informação usados para quantificar a qualidade da distribuição ajustada.

Na secção (3.5), serão abordados conceitos relacionados à estimação intervalar, úteis para a estimação das incertezas de medição do erro de centragem e da espessura.

Sucessivamente, abordam-se os testes de ajustamento e os testes de hipóteses para normalidade.

Na secção (3.8) abordam-se conceitos relativos à regressão linear e, por fim, na secção (3.9) são referidos os métodos de reamostragem.

3.1 Análise exploratória de dados

O processo de ajustar uma função de distribuição a um conjunto de dados pode ser dividido em três fases. A primeira corresponde à escolha de uma distribuição de probabilidade, que modela a variável aleatória. Na fase seguinte o objetivo é estimar os parâmetros dessa distribuição e na última avalia-se a qualidade do ajuste.

A análise exploratória dos dados pode ser um passo útil numa fase inicial. Medidas estatísticas como a média, o desvio padrão, coeficientes de assimetria e achatamento podem ser importantes para identificar possíveis funções de densidade probabilidade a ajustar. A nível gráfico, histogramas dos dados com a respetiva função de densidade de probabilidade sobreposta podem também ser úteis para identificar possíveis distribuições teóricas candidatas a um bom ajustamento.

Cullen et al. (1999) propuseram o gráfico *skewness-kurtosis*, onde são representados os valores relativos à assimetria e ao achatamento para a distribuição empírica e para algumas distribuições comuns. No entanto, a *skewness* e a *kurtosis* são conhecidas por não serem robustas, isto é, podem assumir valores muito diferentes com modificações na amostra. Assim sendo, para ser tida em conta a incerteza relativa à falta de robustez destas estatísticas, recorre-se ao método de reamostragem não paramétrico *bootstrap* (Efron & Tibshirani, 1994).

Porém, o problema da elevada variabilidade da *skewness* e da *kurtosis* não pode ser resolvido com

o uso do *bootstrap*. Deste modo, o gráfico *skewness-kurtosis* não permite concluir nada relativamente a um possível ajustamento, mas surge como uma boa ferramenta, numa fase inicial, para uma melhor compreensão sobre quais serão as distribuições mais prováveis de efetuar um bom ajuste.

3.2 Distribuições de probabilidade

Para este estudo foi feita uma triagem das distribuições com o intuito de eliminar aquelas que não se adequassem aos dados e eleger um conjunto que fosse mais apropriado. Para começar, escolheram-se apenas distribuições contínuas, uma vez que o conjunto de dados são medições dos parâmetros das lentes.

As distribuições escolhidas foram a Normal, Lognormal, Logística, Loglogística, Weibull, Gama e a Beta. Algumas destas distribuições podem assumir três parâmetros porém, para este estudo, apenas se consideraram as versões com dois parâmetros.

Além disso, serão também definidas as distribuições Normal truncada, Uniforme e Bernoulli, escolhidas para as simulações de tolerâncias, segundo (Gross, 2007).

- **Normal**

A distribuição Normal é caracterizada por um parâmetro de localização μ e pelo de dispersão σ . O gráfico da função de densidade é centrado em μ , correspondente à média, e a dispersão dos valores em torno deste valor é dada pelo σ , o desvio padrão. A função de densidade é dada pela Equação (3.5):

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \mu \in \mathbb{R}, \sigma > 0, x \in \mathbb{R}. \quad (3.5)$$

A função distribuição é dada pela expressão (3.6):

$$F(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt. \quad (3.6)$$

A distribuição Normal pode ser truncada no intervalo $[a, b]$. A função densidade da Normal truncada é dada pela Equação (3.7):

$$f(x; \mu, \sigma, a, b) = \begin{cases} \frac{f(x; \mu, \sigma)}{F(b) - F(a)}, & \text{para } a \leq x \leq b \\ 0 & \text{caso contrário.} \end{cases} \quad (3.7)$$

sendo o μ e o σ os parâmetros da distribuição Normal descrita na equação 3.5.

A função distribuição da Normal truncada é definida pela Equação (3.8)

$$F(x; \mu, \sigma) = \frac{\int_a^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt}{F(b) - F(a)} = \frac{F(x) - F(a)}{F(b) - F(a)} \quad (3.8)$$

- **Log-normal**

Uma variável aleatória X de uma distribuição Normal, quando transformada em e^X segue uma distribuição log-normal, como indicado a seguir:

$$X \sim N(\mu, \sigma) \Rightarrow e^X \sim \text{log-normal}(\mu, \sigma), \quad (3.9)$$

sendo o valor médio e a variância definidos do seguinte modo:

$$E[e^X] = e^{\mu + \frac{\sigma^2}{2}}, \text{ var}[e^X] = (e^{\sigma^2} - 1) \times e^{2\mu + \sigma^2}. \quad (3.10)$$

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2}, \mu \in \mathbb{R}, \sigma > 0, x > 0. \quad (3.11)$$

- **Logística**

A distribuição Logística apresenta dois parâmetros, μ de localização e σ de escala. A distribuição está centrada no parâmetro de localização, correspondente à média amostral. Por sua vez, o parâmetro de escala revela a dispersão dos dados, sendo proporcional ao desvio padrão.

$$f(x; \mu, \sigma) = \frac{e^z}{\sigma(1 + e^z)^2}, z = \frac{x - \mu}{\sigma}, 0 \leq \mu \leq \infty, \sigma > 0, x \in \mathbb{R}. \quad (3.12)$$

- **Log-logística**

Uma variável aleatória X de uma distribuição Logística, quando transformada em e^X segue uma distribuição log-logística, como indicado a seguir:

$$X \sim \text{Logística}(\mu, \sigma) \Rightarrow e^X \sim \text{log-logística}(\alpha, \beta). \quad (3.13)$$

Por sua vez, a distribuição Log-logística é caracterizada por um parâmetro de escala, representado por α , e um parâmetro de forma β .

$$f(x; \alpha, \beta) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{(1 + (x/\alpha)^\beta)^2}, x > 0, \alpha > 0, \beta > 0 \quad (3.14)$$

- **Weibull**

A distribuição Weibull possui um parâmetro de forma (β) e um parâmetro de escala (α). Como o nome indica, o parâmetro de forma é responsável pela forma da distribuição e corresponde ao declive de uma reta num gráfico de probabilidades. Já o parâmetro de escala corresponde ao percentil 63,2% dos dados. Ao mudar o valor deste parâmetro os valores da distribuição mudam. Quanto maior for o valor do parâmetro de forma, mais achatada será a distribuição e maior será o leque de valores que a variável poderá assumir.

$$f(x; \alpha, \beta) = \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} e^{-(x/\alpha)^\beta}, x \geq 0; \alpha, \beta > 0 \quad (3.15)$$

- **Gama**

Os parâmetros da distribuição Gama são o de forma, representado por α , e o de escala, representado por β .

O parâmetro de forma determina a forma da curva da distribuição. À medida que se aumenta este parâmetro a distribuição fica mais concentrada no centro, enquanto que valores menores indicam uma distribuição mais dispersa.

O parâmetro de escala controla a escala da distribuição. Quando o parâmetro assume valores maiores a distribuição fica mais estreita. Pelo contrário, valores menores deste parâmetro originam distribuições mais amplas.

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1} e^{-\beta x} \beta^\alpha}{\Gamma(\alpha)}, \quad x > 0, \quad \alpha, \beta > 0, \quad (3.16)$$

onde $\Gamma(\alpha) = (\alpha - 1)!$ corresponde à função Gama e $\beta = \frac{1}{\theta}$, sendo θ o parâmetro de escala.

- **Beta**

A distribuição Beta possui dois parâmetros de forma positivos (α e β). Ambos são responsáveis por dar forma à distribuição, tornando-a bastante flexível.

Caso $\alpha > \beta$ a cauda esquerda é mais longa que a cauda direita. Neste caso, é mais frequente registarem-se mais observações nos valores superiores e menos nos valores menores, havendo maior probabilidade destes últimos serem *outliers*. Para $\beta > \alpha$ a cauda que se alonga é a esquerda e o raciocínio é análogo ao do caso anterior. Ainda, caso os dois parâmetros sejam iguais, a distribuição é simétrica.

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1} (1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad x \in [0, 1], \quad \alpha, \beta > 0, \quad (3.17)$$

onde $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

- **Uniforme**

A variável aleatória X segue uma distribuição Uniforme, definida no intervalo $[a, b]$, sendo $-\infty < a < b < \infty$, com função de densidade probabilidade dada pela seguinte expressão:

$$f(x; a, b) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \quad \text{para } a \leq x \leq b \\ 0 & \text{caso contrário.} \end{cases} \quad (3.18)$$

- **Bernoulli**

Uma variável aleatória X que segue uma distribuição de Bernoulli pode assumir dois resultados diferentes: $X = 0$ e $X = 1$. Este último valor é caracterizado como "sucesso", sendo $P(X = 1) = p$. Pelo contrário, o acontecimento $X = 0$ tem probabilidade $q = 1 - p$, onde $0 < p < 1$. A função de densidade probabilidade é dada pela seguinte expressão:

$$P(X = i) = p^i (1-p)^{1-i}, \quad i \in \{0, 1\}. \quad (3.19)$$

3.3 Métodos de estimação pontual

Durante o processo de ajustamento, é necessário estimar os parâmetros da distribuição a ajustar aos dados. Para isso, existem diferentes métodos de estimação.

No âmbito deste estudo, foi usado o *package fitdistrplus* (Delignette-Muller & Dutang, 2015), que proporciona funções para o ajustamento de distribuições univariadas. Com esta biblioteca, é possível usar diferentes métodos de estimação, tais como o método de máxima verosimilhança (MMV), método dos momentos (MM), método da correspondência dos quantis (MCQ) e o *maximum goodness-of-fit* (MGOF).

Cullen et al. (1999) sugerem que o MMV e o MM são os mais usados para o ajustamento de distribuições de probabilidade. No entanto, os resultados obtidos por estes dois métodos podem ser diferentes, uma vez que o método dos momentos apresenta uma grande sensibilidade a *outliers*.

Existem algumas distribuições contínuas para as quais a função de densidade probabilidade pode assumir valores elevados em algum ponto, não existindo assim estimadores de máxima verosimilhança para os parâmetros da distribuição. Quando o MMV não pode ser usado, o MM surge como alternativa. No entanto, existem algumas distribuições em que alguns dos primeiros momentos não são finitos, impossibilitando assim a estimação pelo método dos momentos.

Luceño (2006) propõe um método alternativo para estimar os parâmetros de uma função distribuição contínua, através da minimização das estatísticas da função distribuição empírica. Uma vez que estas estatísticas são úteis para testar a qualidade do ajustamento da distribuição, o método é designado por *maximum goodness-of-fit*. Este método pode ser usado quando não existem estimadores pelo método de método de máxima verosimilhança ou pelo método dos momentos. Além disso, disponibiliza diversos estimadores, dependendo da escolha da estatística associada à função distribuição empírica, ao contrário dos métodos MM e MMV, que disponibilizam estimadores únicos.

As diferentes estatísticas têm como interpretação a distância entre a distribuição paramétrica ajustada e a distribuição empírica. No referido *package* este método possui oito distâncias disponíveis. D'Agostino (2017) definiu as estatísticas *goodness-of-fit* clássicas, nomeadamente a *Cramér-von Mises* (CvM), *Kolmogorov-Smirnov* (KS) e *Anderson-Darling* (AD). Luceño (2006), propôs as restantes, variantes da estatística AD. Estas últimas estatísticas modificadas podem oferecer mais peso à cauda direita, à esquerda, ou às duas em simultâneo. Este peso pode ser aumentado, através do uso das respetivas estatísticas modificadas de AD de segunda ordem.

Outra alternativa, é o método da correspondência dos quantis (MCQ), que consiste em igualar os quantis teóricos aos quantis empíricos (Tse, 2009).

3.3.1 Método da máxima verosimilhança

A estimação pelo método de máxima verosimilhança (MMV) permite obter a expressão do estimador que fornecerá, perante uma amostra, o valor mais plausível para o parâmetro a estimar. Para isso, define-se a função de verosimilhança, que avalia a plausibilidade para cada valor de θ , numa dada amostra.

Esta função é definida pela função de densidade probabilidade conjunta da amostra. Como este método pressupõe que as observações sejam independentes e identicamente distribuídas, a função pode ser definida pelo produto das funções de densidade probabilidade de cada observação.

$$L(\theta|x_1, \dots, x_n) = \prod_{i=1}^n f(x_i | \theta), \quad (3.20)$$

sendo x_i a i -ésima observação de uma amostra aleatória com n observações.

O valor de θ que maximiza a função de verosimilhança é denominado por estimador de máxima verosimilhança.

Por vezes, é útil utilizar o logaritmo da função de verosimilhança, uma vez que se trata de uma função monótona não decrescente, maximizar o logaritmo é o mesmo que maximizar a função de verosimilhança.

$$\log L(\theta|x_1, \dots, x_n) = \sum_{i=1}^n \log f(x_i|\theta) \quad (3.21)$$

Por fim, é possível encontrar o máximo global da função de verosimilhança ao resolver $\frac{\partial L}{\partial \theta} = 0$ ou $\frac{\partial \log L}{\partial \theta} = 0$. As soluções obtidas nas expressões anteriores correspondem a mínimos ou máximos locais. Posteriormente, é necessário identificar o ponto que corresponde ao máximo local e conseqüentemente, ao estimador de máxima verosimilhança.

3.3.2 Método dos momentos

A estimação pelo método dos momentos (MM) consiste em igualar os momentos empíricos aos momentos teóricos correspondentes. Poderão ser igualados os momentos centrais, não centrais ou ainda, momentos de qualquer ordem. Tais igualdades traduzem-se em equações e as suas soluções são denominadas de estimadores pelo método dos momentos. Geralmente, são usados os momentos não centrais.

Assim sendo, o valor do parâmetro θ pode ser encontrado ao igualar o primeiro momento empírico ao respetivo momento teórico.

$$E[X^k|\theta] = \frac{1}{n} \sum_{i=1}^n x_i^k, \quad (3.22)$$

para $k = 1, \dots, d$, que representa a ordem dos momentos, sendo d o número de parâmetros a estimar e x_i^k as n observações da variável X .

Adicionalmente, pode ser relevante igualar os momentos centrais, quando estamos perante momentos de ordem igual ou superior a dois. Ou seja, poderão ser igualadas as médias e variâncias empíricas aos correspondentes momentos populacionais.

$$E(X|\theta) = \bar{x}, E((X - E(X))^k|\theta) = m_k, \quad (3.23)$$

sendo que m_k representa os momentos centrais empíricos.

3.3.3 Método da correspondência dos quantis

Os momentos de qualquer ordem podem não existir para as distribuições com caudas pesadas. No entanto, os quantis, ou percentis, existem sempre, e possibilitam a estimação dos parâmetros nestes casos. Igualam-se os quantis teóricos aos quantis da amostra.

Considerando k quantis $0 < \delta_1, \dots, \delta_k < 1$, e sendo $\delta_i = F(x_{\delta_i}; \theta)$. Diz-se que $x_{\delta_i} = F^{-1}(\delta_i, \theta)$, onde θ é o vetor de parâmetros de dimensão k da função densidade. Assim sendo, x_{δ_i} é o quantil δ_i da variável X , representando-se por $x_{\delta_i}(\theta)$, dada a dependência em θ .

Seja \hat{x}_{δ_i} o quantil δ_i calculado a partir da amostra. Obtém-se o valor estimado do parâmetro resolvendo a expressão

$$x_{\delta_i}(\hat{\theta}) = \hat{x}_{\delta_i} \quad (3.24)$$

para $i = 1, \dots, k$.

A solução $\hat{\theta}$ das equações anteriores é chamada de estimador pelo método da correspondência dos quantis (MCQ).

3.3.4 Método maximum goodness-of-fit

O objetivo do método de *maximum goodness-of-fit* (MGOF), ou *minimum distance estimation* (D'Agostino, 2017), é comparar as distribuições ajustadas com a distribuição empírica e identificar aquela que mais se aproxima dos dados. Para isso, as estatísticas *maximum goodness-of-fit* medem a

distância entre a distribuição ajustada e a distribuição empírica. No *package fitdistr* existem oito distâncias diferentes, propostas por Luceño (2006), das quais três são as estatísticas *goodness-of-fit* clássicas e as restantes são variantes de uma das distâncias clássicas.

As três distâncias clássicas são a *Kolmogorov-Smirnov* (KS), a *Cramér-von Mises* (CvM) e a *Anderson-Darling* (AD). Considerando x_i como o conjunto das n observações da variável X , ordenadas por ordem crescente, as estimativas empíricas das três distâncias *goodness-of-fit* são apresentadas na Tabela 1.

Tabela 1: Estatísticas *goodness-of-fit* clássicas (D'Agostino, 2017)

Estatística	Fórmula geral
KS	$\sup_x F_n(x) - F(x) $
CvM	$n \int_{-\infty}^{+\infty} (F_n(x) - F(x))^2 dF(x)$
AD	$n \int_{-\infty}^{+\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1-F(x))} dF(x)$

onde $F_i \triangleq F(x_i)$

A estatística *Anderson-Darling* é de particular interesse porque realça igualmente as caudas e o corpo principal da distribuição, pois exerce mais peso nas caudas da distribuição, em comparação com as estatísticas KS e CvM. Assim sendo, esta distância é frequentemente usada para escolher a melhor distribuição entre as ajustadas.

As estatísticas *Cramer-von Mises* e *Kolmogorov-Smirnov* não têm em consideração a complexidade do modelo, isto é, o número de parâmetros. Isto não constitui um problema quando as distribuições comparadas têm o mesmo número de parâmetros, mas pode surgir uma tendência para a seleção de distribuições mais complexas.

As restantes estatísticas *goodness-of-fit* são variantes da estatística *Anderson-Darling*. São dadas pelo peso que dão às caudas da distribuição, como apresentado na Tabela 2.

Tabela 2: Estatísticas *Anderson-Darling* modificadas (Luceño, 2006)

Estatística	Fórmula geral
ADR	$\int_{-\infty}^{+\infty} \frac{(F_n(x)-F(x))^2}{1-F(x)} dF(x)$
ADL	$\int_{-\infty}^{+\infty} \frac{(F_n(x)-F(x))^2}{F(x)} dF(x)$
AD2L	$ad2r = \int_{-\infty}^{+\infty} \frac{(F_n(x)-F(x))^2}{(1-F(x))^2} dF(x)$
AD2R	$ad2l = \int_{-\infty}^{+\infty} \frac{(F_n(x)-F(x))^2}{(F(x))^2} dF(x)$
AD2	$ad2r + ad2l$

onde $F_i \triangleq F(x_i)$; $\bar{F}_i \triangleq 1 - F(x_i)$

A *Right-tail Anderson Darling* (ADR) destaca a cauda direita, enquanto que a *Left-tail Anderson Darling* (ADL) realça a cauda esquerda. Cada uma das caudas, ou as duas em simultâneo, podem receber um maior destaque, através do uso das estatísticas AD de segunda ordem, a AD2R, AD2L e AD2, respetivamente.

3.4 Critérios de informação

Os critérios de informação medem a qualidade relativa de uma distribuição ajustada a um conjunto de dados. Neste trabalho foram usados os seguintes critérios: AIC (*Akaike's Information Criterion*) e o BIC (*Bayesian Information Criterion*). Quanto menores forem os valores destes critérios mais adequado é o modelo de distribuição ajustado aos dados.

3.5 Estimação intervalar

Na estimação intervalar são determinados intervalos dos valores possíveis para o parâmetro desconhecido. Estes intervalos são constituídos com um nível de confiança, sendo 90%, 95% e 99%, os mais usuais. Um intervalo de confiança de 99% significa que aproximadamente 99% dos intervalos calculados através deste método, vão conter o verdadeiro valor do parâmetro desconhecido (mas fixo).

Considerando uma amostra aleatória proveniente de uma distribuição $N(\mu, \sigma)$, em que μ é desconhecido e σ conhecido. A variável aleatória $\frac{\bar{X}-\mu}{\sigma} \sqrt{n}$ segue a $N(0, 1)$.

Assim sendo, o intervalo de $100(1 - \alpha)\%$ de confiança para o valor médio μ , é dado pela seguinte expressão:

$$\left] \bar{X} - c \frac{\sigma}{\sqrt{n}}, \bar{X} + c \frac{\sigma}{\sqrt{n}} \right[, \quad (3.25)$$

sendo $c = z_{1-\alpha/2}$ o quantil de probabilidade $1 - \alpha/2$ da $N(0, 1)$.

Caso σ seja desconhecido, a variável $\frac{\bar{X}-\mu}{S} \sqrt{n}$ apresenta uma distribuição *t-student* com $n - 1$ graus de liberdade. Desta vez, intervalo de $100(1 - \alpha)\%$ confiança para o valor médio μ , é dado pela seguinte expressão:

$$\left] \bar{X} - c \frac{S}{\sqrt{n}}, \bar{X} + c \frac{S}{\sqrt{n}} \right[, \quad (3.26)$$

sendo $c = t_{n-1;1-\alpha/2}$, que representa o quantil de probabilidade $1 - \alpha/2$ da t_{n-1} , e S o desvio padrão amostral.

3.6 Testes de ajustamento

Os testes de ajustamento são uma categoria dos testes de hipóteses não paramétricas cujo o intuito é testar se os dados são provenientes de uma dada função de distribuição F_0 , que poderá conter ou não parâmetros especificados.

As hipóteses a testar são as seguintes:

$$H_0 : F(X) = F_0(X) \quad vs \quad H_1 : F(X) \neq F_0(X)$$

- **Teste de Kolmogorov-Smirnov**

O teste de Kolmogorov-Smirnov utiliza-se quando a distribuição populacional é contínua e são conhecidos os seus parâmetros.

Considerando uma amostra ordenada, proveniente de uma distribuição desconhecida, mas contínua, o objetivo é testar se a função de distribuição da amostra é idêntica à função de distribuição que se assume como conhecida (F_0).

A estatística de teste dada pela seguinte expressão:

$$D_n = \sup_{x \in \mathbb{R}} |F_n^*(x) - F(x)| \quad (3.27)$$

A estatística de teste D_n mede a distância máxima entre a função distribuição da amostra F_n^* e a função distribuição teórica F . Assim sendo, a distância de Komogorov-Smirnov é uma medida

de discrepância entre as funções de distribuição e é medida verticalmente, considerando todos os pontos da amostra.

- **Teste de ajustamento de Qui-quadrado**

O objetivo do teste de ajustamento de Qui-quadrado é avaliar se os dados se ajustam a um certo modelo teórico correspondente a uma dada função de distribuição.

Agrupam-se os dados em classes e comparam-se as frequências dos valores observados na amostra com os valores esperados das diferentes classes, supondo verdadeira a hipótese nula.

Existem condições que têm de ser verificadas para que este método possa ser aplicado. É necessário que não existam mais de 20% de classes com valores esperados inferiores a 5, e que todas apresentem valor esperado superior ou igual a 1.

No entanto, caso alguma das condições não seja verificada, é possível agregar classes vizinhas.

Para aplicar o teste efetuam-se os seguintes passos:

- Formular as hipóteses

H_0 : Os dados seguem distribuição especificada

vs

H_1 : Os dados não seguem a distribuição especificada

- Dividir as n observações por $j \geq 2$ classes não sobreponíveis. Calculam-se as frequências observadas nas diferentes classes (n_j).
- Calcular as frequências esperadas em cada classe, supondo verdadeira a hipótese nula. São representadas por np_j , sendo p_j a probabilidade de a variável aleatória pertencer à j -ésima classe, sob validade da H_0 .
- Calcular a estatística de teste, definida na equação (3.28). Diz respeito a uma medida de ajuste entre as frequências observadas e esperadas.

$$U = \sum_{j=1}^k \frac{(N_j - np_j)^2}{np_j} \quad (3.28)$$

Espera-se que a estatística de teste U tenha um valor pequeno, sinónimo da pouca discrepância entre as frequências observadas e as teóricas (esperadas). Assim sendo, a hipótese nula deverá ser rejeitada para valores grandes da estatística de teste.

U segue uma distribuição $\chi^2_{(k-1)-r}$, sendo r o número de parâmetros a serem estimados a partir da amostra. No caso da análise realizada $r = 0$.

3.7 Testes de hipóteses para normalidade

Os testes de normalidade são usados para averiguar se a distribuição dos dados é proveniente de uma população com uma distribuição Normal. Assim, é pretendido testar:

$$H_0 : X \sim N(\mu, \sigma) \text{ vs } H_1 : X \not\sim N(\mu, \sigma)$$

- **Teste de Lilliefors**

O teste de Kolmogorov-Smirnov pode ser usado para testar a normalidade, no caso dos parâmetros, valor médio e variância, serem conhecidos. O teste de Lilliefors corresponde a uma adaptação do teste de Kolmogorov-Smirnov, no caso dos referidos parâmetros não estarem especificados em H_0 .

- **Teste de Shapiro-Wilk**

O teste de Shapiro-Wilk surge como uma alternativa ao teste de K-S, para averiguar se a amostra aleatória é proveniente de uma distribuição Normal. A estatística de teste é dada pela seguinte expressão Shapiro & Wilk (1965):

$$W = \frac{(\sum_{i=1}^n \alpha_i X_i)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (3.29)$$

sendo X_i os valores da variável em estudo, ordenados por ordem crescente, \bar{X} a sua média e α_i as constantes geradas a partir da média, variância e covariância de n ordens com a distribuição Normal $N(0, 1)$. Valores elevados para W indicam que a distribuição da variável é próxima da Normal.

3.8 Regressão linear

O modelo de regressão linear é amplamente utilizado em aplicações, de fácil interpretabilidade e que, de um modo geral, permite a obtenção de boas previsões. A variável resposta (ou variável dependente) é quantitativa e as variáveis explicativas (ou independentes) são quantitativas e/ou qualitativas.

A relação entre a variável resposta Y e as variáveis explicativas $X = (X_1, \dots, X_p)$ pode ser descrita através da expressão:

$$Y = f(X) + \varepsilon, \quad (3.30)$$

sendo ε o erro aleatório, com $\varepsilon \sim N(0, \sigma^2)$.

A estimação de f apresenta dois propósitos: a previsão e a inferência.

- **Previsão**

$$\hat{Y} = \hat{f}(X), \quad (3.31)$$

sendo \hat{Y} a previsão de Y e \hat{f} estimativa para f .

A precisão de \hat{Y} depende de dois fatores: erro redutível e erro irredutível. A \hat{f} não será uma estimativa perfeita para f e esta imprecisão irá introduzir algum erro. Este erro será redutível, uma vez que a precisão de \hat{f} pode ser melhorada através de técnicas de aprendizagem estatística para estimar f .

No entanto, mesmo que a estimativa para f seja perfeita, a previsão estará, na mesma, associada a um erro, uma vez que a variável de resposta é definida através de ε (como visto na fórmula (3.30)). A variabilidade associada a ε , conhecida como erro irredutível, afeta as previsões e o seu erro não pode ser reduzido.

O objetivo será estimar f através de técnicas de aprendizagem, minimizando o erro redutível.

- **Inferência**

Quando estamos perante um caso de inferência o interesse é perceber a associação entre as variáveis preditoras e a variável de resposta. O objetivo pode não ser necessariamente a previsão.

Assim sendo, pretende-se indentificar que variáveis estão associadas com a resposta e perceber qual a relação entre a variável resposta e cada preditor. Além disso, é importante perceber se a relação entre Y e X_1, \dots, X_p pode ser descrita de forma adequada através de um uma equação linear.

Ainda, o objetivo poderá ser uma combinação de previsão e inferência.

Depois da seleção do modelo é necessário proceder à divisão da amostra em treino e teste (80% e 20%), respetivamente.

Para estimar f , estabelecemos um conjunto de dados de treino, usado para aplicar o método de aprendizagem estatística para a estimação de f . Os métodos de aprendizagem estatística para este intuito podem ser paramétricos ou não paramétricos

- **Métodos paramétricos**

Os métodos paramétricos efetuam pressupostos quanto à forma de f . Por exemplo, a forma linear:

$$f(X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p. \quad (3.32)$$

Depois da seleção do modelo, é necessário um método que use os dados de treino para ajustar o modelo, ou seja, precisamos de estimar os parâmetros $\beta_1, \beta_2, \dots, \beta_p$.

Frequentemente, o modelo estimado através desta metodologia providencia uma boa aproximação à forma desconhecida de f , caso assumamos uma forma linear.

- **Métodos não paramétricos**

Por outro lado, nos métodos não paramétricos, não existem pressupostos quanto à forma de f . Procuram estimá-la de maneira a que esta se aproxime dos pontos dos dados, sem correr o risco de *overfitting*, ou sobreajustamento.

No entanto, estes métodos apresentam uma desvantagem: é necessário um número elevado de observações para obter uma estimativa de f precisa, uma vez que estes métodos não reduzem o problema da estimação para um número reduzido de parâmetros.

3.9 Métodos de reamostragem

Os métodos de reamostragem consistem em obter diferentes amostras a partir de um conjunto de dados original.

Os métodos *Cross-Validation* (CV) (k-Fold CV e LOOCV) e o *bootstrap* foram os selecionados para este estudo.

Para estimar o erro de teste, a amostra é dividida em duas partes: o conjunto de treino e o de teste (validação). O modelo é ajustado ao conjunto de treino e aplicado, para previsão da variável resposta, aos dados de teste.

Na reamostragem *bootstrap* foram obtidas diferentes amostras a partir da amostra original e com a mesma dimensão, mas com repetição de dados.

3.9.1 Cross-validation

O *k-Fold Cross-validation* (*k-Fold CV*) consiste em dividir aleatoriamente o conjunto de observações em k grupos de dimensão aproximada. O primeiro grupo é considerado como sendo o de validação e os restantes são usados para ajustar o método (treino). O erro quadrático médio (MSE - *Mean Squared Error*) é calculado para o conjunto de validação.

O método é repetido k vezes e, em cada uma, é considerando um grupo diferente para validação. No final, obtêm-se k erros quadráticos médios e a estimativa do erro final é dada pela média destes k erros.

O *Leave-One-Out-Cross-Validation* (LOOCV) é um caso especial do *k-Fold CV*. A amostra é dividida em duas partes, sendo o conjunto de validação apenas constituído por uma única observação e as restantes são o conjunto de treino. O modelo estatístico de aprendizagem é ajustado nas $n - 1$ observações e a previsão é feita para a observação de teste do conjunto de validação. Este processo é repetido n vezes, de modo a que todas as observações sejam o conjunto de validação.

Assim sendo, teremos n erros quadráticos médios, para cada conjunto de validação, e o erro de teste é obtido pela média dos erros.

Este método apresenta algumas vantagens face ao anterior: tende a não sobrestimar o erro de teste e produz sempre os mesmos resultados, pois não existe aleatoriedade na divisão dos conjuntos, como acontece no método anterior.

De um modo geral, o *k-Fold CV* produz estimativas mais precisas da taxa de erro de teste, face ao LOOCV, cuja variância da estimativa do erro tende a ser mais elevada.

3.9.2 Bootstrap

O *Bootstrap* é uma técnica de reamostragem a usar em diversas situações tais como: para cálculo da precisão de um modelo, estimação dos parâmetros ou quantificação da incerteza associada a um determinado estimador, entre outras. Por exemplo, através deste método podem ser obtidas estimativas para a mediana de uma população, bem como o respetivo intervalo de confiança (James et al., 2013).

Em alternativa à recolha de novos dados, independentes da população, obtêm-se, repetidamente, diferentes amostras de igual dimensão provenientes dos dados originais, com reposição.

A incerteza dos parâmetros de uma distribuição ajustada pode ser estimada através da função *bootdist* da biblioteca *fitdist*, como referido em Efron & Tibshirani (1994). A função *bootdist* da biblioteca referida devolve as medianas dos parâmetros das distribuições das n amostras geradas por *bootstrap*. Além disso, devolve também os intervalos de confiança a 95% para as medianas dos parâmetros.

O *bootstrap* usado na forma paramétrica assume que os dados provêm de uma distribuição identificada, enquanto que a forma não paramétrica não faz qualquer suposição acerca da forma dos dados.

4 Análise Exploratória dos Dados

Uma vez recolhidos os dados, a análise exploratória é o primeiro passo a tomar para qualquer estudo estatístico. Esta etapa é importante para caracterizar os dados e resumir as suas principais características, bem como detetar alguns erros ou anomalias na base de dados, que poderão ser prejudiciais numa análise futura. Além disso, permite uma melhor compreensão das variáveis e do modo como se relacionam, útil para perceber que metodologias estatísticas poderão ou não ser úteis.

Para o efeito de caracterizar os dados da amostra foram usadas medidas de tendência central, de dispersão, e de forma, nomeadamente de assimetria e achatamento.

- **Medidas de tendência central:** Média, Mediana e Quantis;
- **Medidas de dispersão:** Desvio padrão, Variância, Coeficiente de variação (CV) e Amplitude interquartis;
- **Medidas de forma:** Assimetria e Achatamento.

Além disso, foram calculadas as incertezas de medição para o erro de centragem e para a espessura.

Foram realizadas medições de 11 tipos de lentes, as quais serão denominadas por conjuntos, de A a K. Para os Conjuntos A e B foram efetuadas medições do diâmetro, erro de centragem e espessura. Para as restantes lentes, de C a K, apenas foi medida a espessura.

As lentes A e B fazem parte de uma mira, enquanto que as nove restantes são constituintes de uma objetiva.

Na Tabela 3 está apresentado um resumo das informações relativas às características das lentes que foram medidas, nomeadamente a dimensão da amostra das medições e as tolerâncias associadas a cada característica.

Tabela 3: Valores nominais e tolerâncias para os parâmetros das lentes.

Conjunto	Dimensão	Diâmetro		Erro de centragem	Espessura	
		Nominal	Tolerância	Tolerância	Nominal	Tolerância
A	210	23,200	0,000	2,000	4,000	+ 0,100
			-0,033			- 0,100
B	150	14,5	0,000	1,000	5,000	+ 0,100
			-0,027			- 0,100
C	96	27,000	0,000	1,200	1,610	+ 0,030
			-0,013			- 0,010
D	300	28,500	-0,040	0,800	3,150	+ 0,030
			-0,100			-0,030
E	75	27,000	-0,065	3,000	6,800	0,000
			-0,117			- 0,030
F	80	27,500	0,000	2,000	6,470	+ 0,030
			-0,100			- 0,030
G	80	29,000	-0,100	2,100	6,600	+ 0,030
			0,000			- 0,030
H	200	20,000	0,000	1,000	3,300	+ 0,030
			- 0,013			- 0,030
I	100	13,900	0,000	1,000	1,200	+ 0,100
			- 0,027			- 0,100
J	90	14,800	0,000	1,000	6,200	+ 0,100
			-0,027			- 0,100
K	70	27,000	-0,040	0,700	4,840	+ 0,030
			-0,100			- 0,030

4.1 Diâmetro

A Tabela 4 contém as medidas de localização para as medições efetuadas às lentes dos Conjuntos A e B.

Tabela 4: Medidas de localização para as medições do diâmetro.

Conjunto	Mínimo	1º Quartil	Média	Mediana	3º Quartil	Máximo
A	23,172	23,180	23,181	23,181	23,182	23,189
B	14,478	14,480	14,480	14,481	14,482	14,483

Consultando a Tabela 4, é possível constatar que os valores médios dos Conjuntos A e B, além de serem próximos, estão aproximadamente à mesma distância do valor nominal (0,019 e 0,020, respetivamente). Além disso, os dois conjuntos possuem a mesma amplitude interquartil.

A Tabela 5 apresenta as medidas de forma e dispersão para as medições efetuadas às lentes dos Conjuntos A e B.

Tabela 5: Medidas de forma e dispersão para as medições do diâmetro.

Conjunto	Desvio padrão	CV(%)	Assimetria	Achatamento
A	0,002	0,008	-0,448	5,540
B	0,001	0,008	-0,217	2,462

No que diz respeito à dispersão, os Conjunto A e B são semelhantes, pois apresentam um coeficiente de variação igual e desvios padrão próximos, tal como sugere a Tabela 5. No entanto, o Conjunto A, que apresenta maior tolerância, tem um desvio padrão superior. A dispersão em termos relativos é igual nos dois conjuntos, mas inferior, em termos absolutos, no Conjunto B.

Relativamente à forma, os dados são simétricos, sendo o Conjunto A aquele que está mais próximo de ser assimétrico. Por outro lado, pelo achatamento, os dois conjuntos apresentam formas diferentes. O Conjunto B apresenta um achatamento inferior comparativamente ao Conjunto A. Neste último, é provável que se verifiquem *outliers*, uma vez que um valor alto do coeficiente de *kurtosis* reflete caudas pesadas. No Conjunto B, a presença de *outliers* já não será tão provável pois o valor de achatamento é inferior, sugerindo caudas mais leves, que implicam que os valores estejam não tão distantes da média.

Na Figura 22 estão representados os histogramas e os *boxplots* para os Conjuntos A e B.

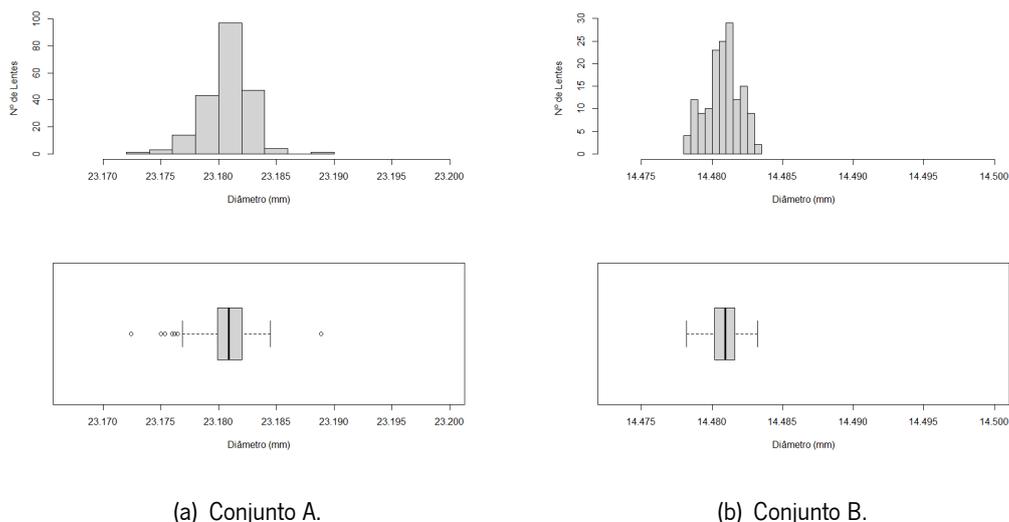


Figura 22: Representação gráfica das medições do diâmetro.

Como sugerido pela Tabela 5 e ilustrado no *boxplot* da Figura 4.22(a), verifica-se a presença de *outliers* no Conjunto A. Além disso, é possível constatar que existem três classes predominantes, que agregam a grande maioria dos valores, sendo que as restantes, que têm poucas observações, ilustram as caudas pesadas, quantificadas pelo elevado coeficiente de *kurtosis* para o Conjunto A.

Por outro lado, no Conjunto B, não se detetou a presença de *outliers*, tal como sugere o *boxplot* da Figura 4.22(b). Seguidamente, as classes mais distantes do centro agregam mais observações face às classes centrais, em comparação com o que se verificou no Conjunto A. Isto é sinónimo de caudas mais leves, tal como foi visto no achatamento para o Conjunto B (Tabela 5).

Os resultados dos testes de normalidade aplicados às medições do diâmetro, em ambos os conjuntos, estão disponíveis na Tabela 6.

Tabela 6: Testes de normalidade para as medições do diâmetro.

Conjunto	Lilliefors		Shapiro-Wilk	
	D	valor- p	W	valor- p
A	0,080	0,002	0,963	$2,807 \times 10^{-5}$
B	0,064	0,133	0,980	0,024

A Tabela 6 revela que testes os produziram resultados diferentes. No teste de Lilliefors, para um nível de significância de 5%, não existem evidências estatísticas para se rejeitar a hipótese nula de que os dados do Conjunto B provêm de uma distribuição Normal. Pelo contrário, pelo teste de Shapiro-Wilk,

conclui-se que existem evidências estatísticas para se rejeitar a hipótese nula, a um nível de significância de 5%.

Para o Conjunto A, ambos os testes sugerem que existem evidências estatísticas para se rejeitar a hipótese nula, para 5% de significância.

Por fim, o último passo na análise inicial das medições do diâmetro das lentes foi a comparação dos valores relativos à assimetria e ao achatamento nos dados empíricos e nas distribuições teóricas, apresentada na Figura 23.

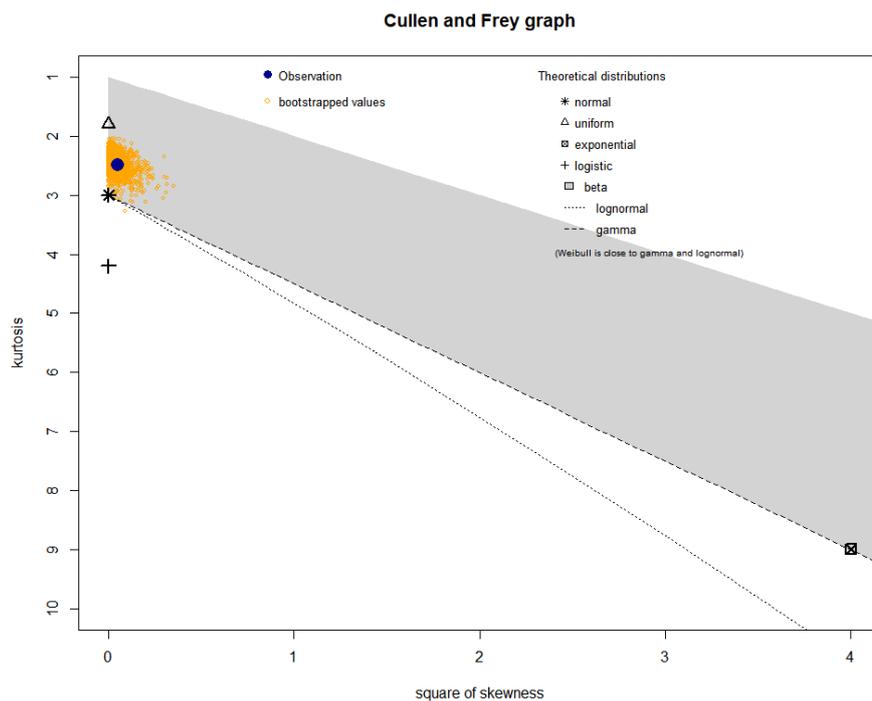


Figura 23: Gráfico de Cullen e Frey para as medições do diâmetro do Conjunto B (Cullen et al., 1999).

Na Figura 23 estão representados os valores de assimetria (*skewness*) e achatamento (*kurtose*) para dados empíricos e para amostras geradas por bootstrap não paramétrico. A Figura 23 sugere que distribuições como a Normal, a Gamma ou a Weibull poderão estar próximas da distribuição empírica dos dados.

4.2 Erro de centragem

As medidas de localização para o erro de centragem, nos Conjuntos A e B, estão disponíveis na Tabela 7.

Tabela 7: Medidas de localização para as medições do erro de centragem.

Conjunto	Mínimo	1º Quartil	Média	Mediana	3º Quartil	Máximo
A	0,192	0,342	0,442	0,447	0,526	0,855
B	0,436	0,526	0,571	0,568	0,620	0,699

Apesar de apresentar uma tolerância superior para ao erro de centragem, em média, as medições efetuadas para as lentes Conjunto A foram inferiores às do Conjunto B, que apresentava metade da tolerância. No entanto, o valor máximo registado no Conjunto B foi superior ao máximo observado no Conjunto A, como é possível conferir na Tabela 7.

A Tabela 8 apresenta as medidas de forma e dispersão para erro de centragem, nos A e B.

Tabela 8: Medidas de forma e dispersão para as medições do erro de centragem.

Conjunto	Desvio padrão	CV(%)	Assimetria	Achatamento
A	0,139	30,734	0,343	2,892
B	0,062	10,885	-0,009	2,372

Pela Tabela 8, é possível concluir que as medições efetuadas às lentes do Conjunto A apresentam maior variabilidade relativamente à média, face ao Conjunto B. No que diz respeito ao achatamento, o Conjunto A apresenta maior coeficiente de *kurtosis* que o Conjunto B. Apesar de não serem valores altos que indiquem a presença de *outliers*, ao acontecer, é mais provável que ocorra no Conjunto A do que no B. Já na assimetria, ambos os conjuntos são simétricos, pois assumem valores reduzidos do coeficiente de *skewness*.

Os histogramas e os *boxplots* para os Conjuntos A e B são apresentados na Figura 22.

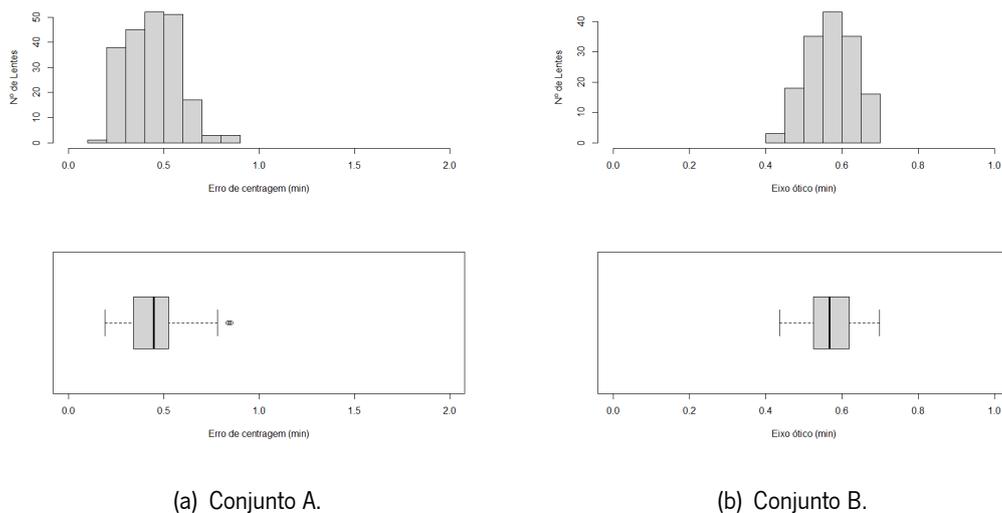


Figura 24: Representação gráfica das medições do erro de centragem.

Relativamente aos *outliers*, apenas surgiram no Conjunto A, como ilustra o *boxplot* da Figura 4.24(a). A nível gráfico, é visível a simetria no Conjunto B, ilustrada no histograma da Figura 4.24(b). Pelo contrário, a simetria no Conjunto A já não é tão evidente, tal como sugere o histograma da Figura 4.24(a), onde é visível cauda mais prolongada do lado direito, que corresponde aos *outliers*.

Tabela 9 contém os resultados dos testes de normalidade aplicados às medições do erro de centragem, em ambos os conjuntos.

Tabela 9: Testes de normalidade para as medições do erro de centragem.

Conjunto	Lilliefors		Shapiro-Wilk	
	D	valor- p	W	valor- p
A	0,054	0,143	0,979	0,004
B	0,048	0,540	0,988	0,225

Como revela a Tabela 9, apenas existiram evidências estatísticas para se rejeitar a hipótese nula, de que os dados provêm de uma distribuição Normal, no teste de Shapiro-Wilk aplicado ao Conjunto A, para 5 % de significância. Os restantes resultados sugerem que não existem evidências estatísticas para se rejeitar a hipótese nula.

Uma vez mais, para finalizar a análise inicial das medições do erro de centragem das lentes, comparou-se as medidas de assimetria e de achatamento, nos dados empíricos e nas distribuições teóricas (Figura 25).

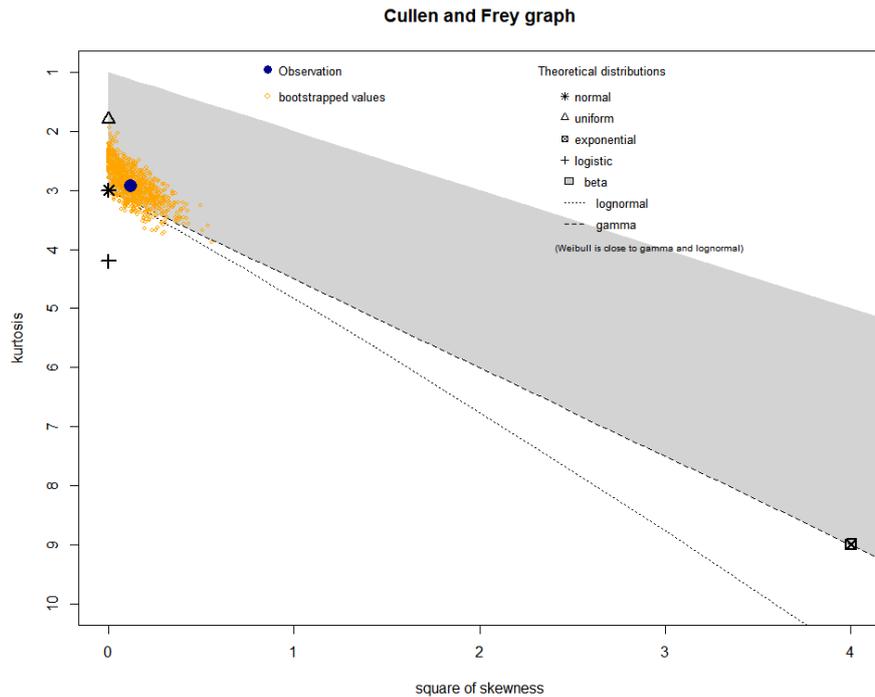


Figura 25: Gráfico de Cullen e Frey para as medições do erro de centragem no Conjunto A (Cullen et al., 1999).

A Figura 23 sugere que a distribuição empírica dos dados poderá apresentar semelhanças com uma distribuição Normal, Beta ou Gamma.

4.2.1 Incertezas das medições

Para o cálculo da incerteza de medição no erro de centragem, foram selecionadas 12 lentes, de forma aleatória, e cada uma foi medida dez vezes. Seguidamente, foram calculados os valores médios para cada lente e estimados os respetivos intervalos de confiança.

Na Figura 26 estão representados os intervalos de confiança para os valores médios das medições do erro de centragem nas 12 lentes. Os intervalos foram calculados para um nível de confiança de 95%, 97,5%, 99% e 99,7%.

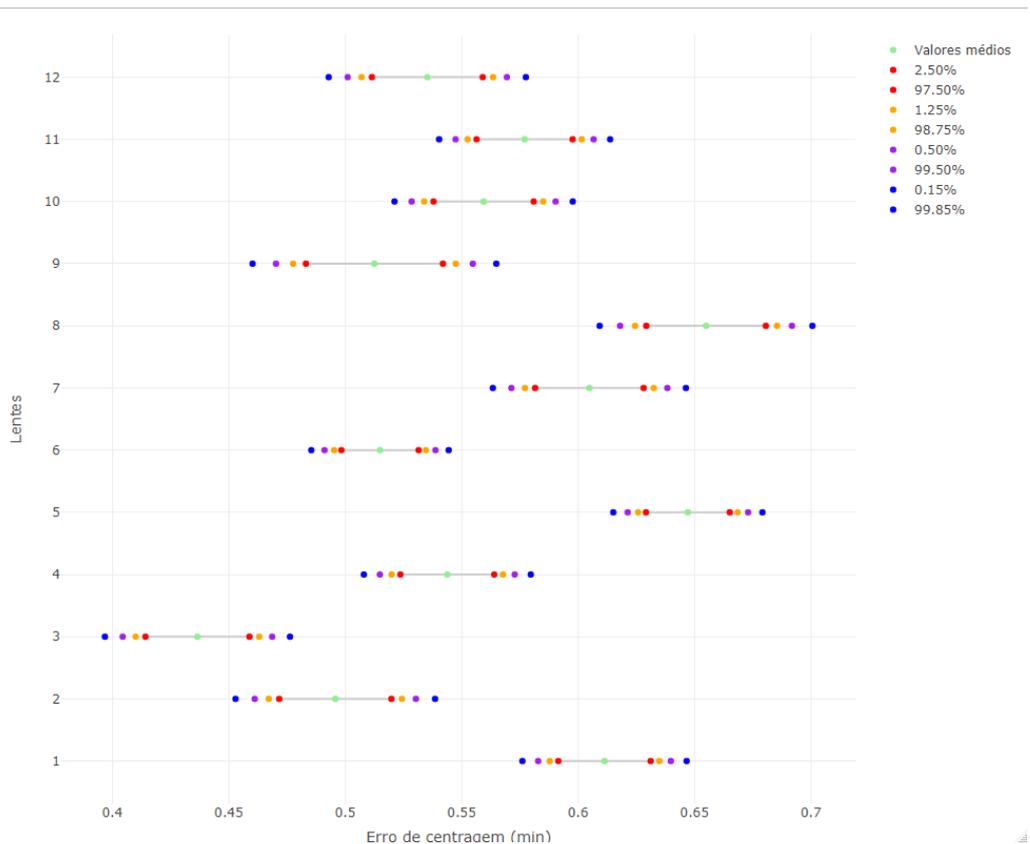


Figura 26: Intervalos de confiança para as medições do erro de centragem no Conjunto B.

A Figura 26 revela alguma variabilidade de lente para lente, que poderá ter sido originada pela folga no suporte em que a lente foi colocada ou por características da própria lente. Uma vez que o aparelho executa a rotação da lente de forma autónoma, a variabilidade pode ser explicada pelo modo como a lente é colocada no aparelho.

O passo seguinte, foi o cálculo das diferenças absolutas entre os extremos dos intervalos e os valores médios. O resultado desta diferença corresponde à incerteza da medição naquela lente, para um determinado nível de confiança. A diferença máxima observada nas 12 lentes foi considerada como sendo a incerteza da medição para o erro de centragem.

A Tabela 10 apresenta as incertezas de medição para o erro de centragem no Conjunto B, para diferentes níveis de confiança.

Tabela 10: Incerteza associada às medições do erro de centragem.

Confiança	Incerteza (min)
95%	0,026
97,5%	0,030
99%	0,034
99,7%	0,039

Concluiu-se que, para um nível de confiança de 99,7%, a incerteza associada à medição do erro de centragem foi de 0,039 minutos.

4.3 Espessura

As medidas de localização relativas às medições da espessura para os Conjuntos A a K constam na Tabela 11.

Tabela 11: Medidas de localização para as medições da espessura.

Conjunto	Mínimo	1º Quartil	Média	Mediana	3º Quartil	Máximo
A	4,001	4,025	4,033	4,035	4,041	4,059
B	4,999	5,016	5,027	5,028	5,037	5,060
C	1,625	1,634	1,635	1,636	1,638	1,640
D	3,138	3,167	3,170	3,171	3,175	3,180
E	6,785	6,793	6,795	6,795	6,798	6,800
F	6,445	6,480	6,483	6,485	6,490	6,495
G	6,576	6,595	6,599	6,602	6,605	6,614
H	3,271	3,313	3,317	3,319	3,324	3,330
I	1,251	1,260	1,264	1,264	1,269	1,278
J	6,186	6,224	6,230	6,232	6,239	6,251
K	4,852	4,861	4,864	4,866	4,868	4,870

Em média, os valores médios das medições das lentes foram superiores ao valor nominal (consultar a Tabela 3) em todos os conjuntos, exceto no Conjunto G e no Conjunto E. Este último era previsível pois

a tolerância superior era zero. Uma vez mais, com exceção do Conjunto E, a mediana foi sempre igual ou superior à média.

A Tabela 12 apresenta as medidas de forma e dispersão para a espessura dos Conjuntos A a K.

Tabela 12: Medidas de forma e dispersão para as medições da espessura.

Conjunto	Desvio padrão	CV(%)	Assimetria	Achatamento
A	0,011	0,274	-0,269	2,632
B	0,013	0,265	0,010	2,288
C	0,004	0,230	-1,195	3,735
D	0,007	0,230	-1,450	6,497
E	0,004	0,050	-0,086	3,836
F	0,009	0,140	-1,613	6,300
G	0,008	0,110	-0,690	3,202
H	0,011	0,332	-1,774	6,595
I	0,007	0,520	0,113	2,379
J	0,013	0,205	-0,939	3,989
K	0,005	0,101	-1,069	3,200

Em termos relativos, a dispersão dos valores em torno da média foi superior na maioria dos conjuntos com tolerâncias igual a 0,1, como é possível verificar pela Tabela 12. Adicionalmente, as lentes com tolerância 0,03 apresentaram dados bastante assimétricos em relação às lentes com tolerâncias maiores. Apenas são exceções o Conjunto E, cuja amplitude de tolerância é a menor e apresenta dados simétricos, e o Conjunto J, cuja tolerância é igual a 0,1, mas os dados são assimétricos. No que diz respeito ao achatamento, as lentes com tolerâncias menores ou iguais a 0,03 apresentaram o coeficiente de *kurtose* superior a 3, sendo nestes conjuntos mais provável encontrarmos *outliers*.

Os histogramas e os *boxplots* para a espessura nos Conjuntos A a D estão representados na Figura 27.

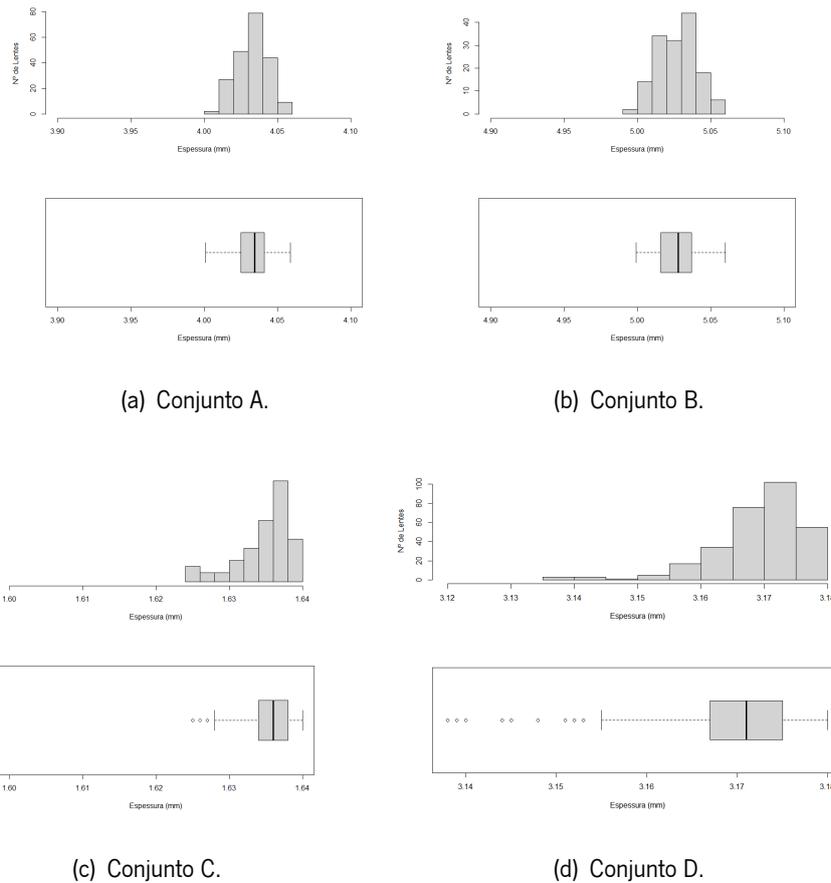


Figura 27: Representação gráfica das medições da espessura nos Conjuntos A a D.

A Figura 27 revela que a maioria das medições da espessura foram superiores ao valor nominal. Para os Conjuntos A e B, os histogramas das Figuras 4.27(a) e 4.27(b) ilustram a simetria dos conjuntos com tolerâncias igual a 0,1, e o baixo achatamento, já que não se verificam, graficamente, grandes diferenças entre as classes centrais e as classes mais distantes da mediana.

A forte assimetria negativa das lentes com tolerâncias inferiores é ilustrada nos histogramas das Figuras 4.27(c) e 4.27(d), na medida em que se observa baixa frequência nas classes menores e mais observações nas classes maiores, superiores à mediana. Nestes conjuntos a distribuição apresentou a cauda esquerda, ou inferior, mais prolongada. Além disso, também é notório o elevado achatamento pois existe uma diferença significativa entre as classes com mais observações e as classes com menos. Consequentemente, verificou-se também que os valores mais baixos correspondiam a *outliers*.

Os histogramas e boxplots para os conjuntos restantes estão representados nas Figuras 1 e 2, remetidas para o Apêndice A.1.

Na Tabela 13 estão disponíveis os resultados dos testes de normalidade, aplicados aos Conjuntos A a K.

Tabela 13: Testes de normalidade para as medições espessura.

Conjunto	Lilliefors		Shapiro-Wilk	
	D	valor- p	W	valor- p
A	0,065	0,033	0,987	0,051
B	0,076	0,033	0,986	0,148
C	0,190	$3,778 \times 10^{-9}$	0,876	$1,995 \times 10^{-7}$
D	0,123	$8,578 \times 10^{-12}$	0,899	$2,771 \times 10^{-13}$
E	0,095	0,088	0,973	0,109
F	0,131	0,002	0,872	$1,001 \times 10^{-6}$
G	0,139	0,001	0,963	0,022
H	0,145	$5,842 \times 10^{-11}$	0,836	$9,793 \times 10^{-14}$
I	0,063	0,459	0,984	0,267
J	0,101	0,023	0,947	0,001
K	0,181	$5,609 \times 10^{-6}$	0,879	$6,305 \times 10^{-6}$

No que diz respeito ao teste de Shapiro-Wilk, para um nível de significância de 5%, não existem evidências estatísticas para rejeitar a hipótese nula de que os dados provêm de uma distribuição Normal nos Conjuntos: A, B, E e I. No entanto, segundo o teste de Lilliefors, apenas não se verificaram evidências estatísticas para a rejeição da H_0 , no mesmo nível de significância, nos Conjuntos E e I.

Novamente, para finalizar a análise inicial das medições da espessura das lentes compararam-se as medidas de forma e assimetria nos dados empíricos e nas distribuições teóricas, apresentadas na Figura 23.

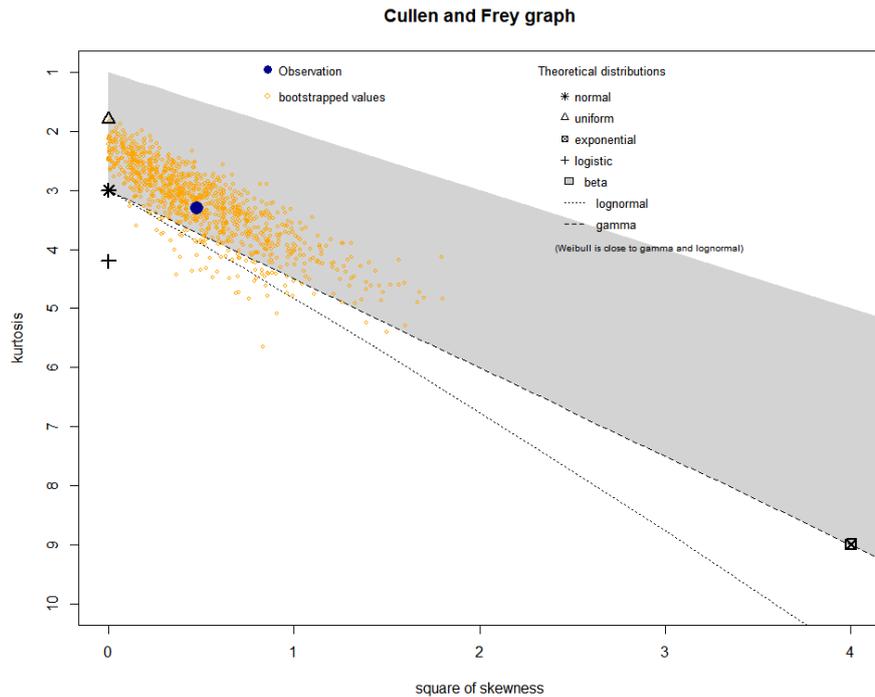


Figura 28: Gráfico de Cullen e Frey para as medições da espessura no Conjunto G (Cullen et al., 1999)

Na Figura 23 estão representados os valores de assimetria (*skewness*) e achatamento (*kurtose*) para dados empíricos e para amostras geradas por bootstrap não paramétrico. A Figura 23 sugere que distribuições como a Gamma, Weibull ou Lognormal poderão estar próximas da distribuição empírica dos dados.

4.3.1 Incerteza das medições

Para o cálculo da incerteza de medição na espessura, foram selecionadas 12 lentes, de forma aleatória, e cada uma foi medida 10 vezes. Seguidamente, foram calculados os valores médios para cada lente e estimados os respetivos intervalos de confiança.

A Figura 29 apresenta graficamente os intervalos de confiança para as medições da espessura para diferentes níveis de confiança.

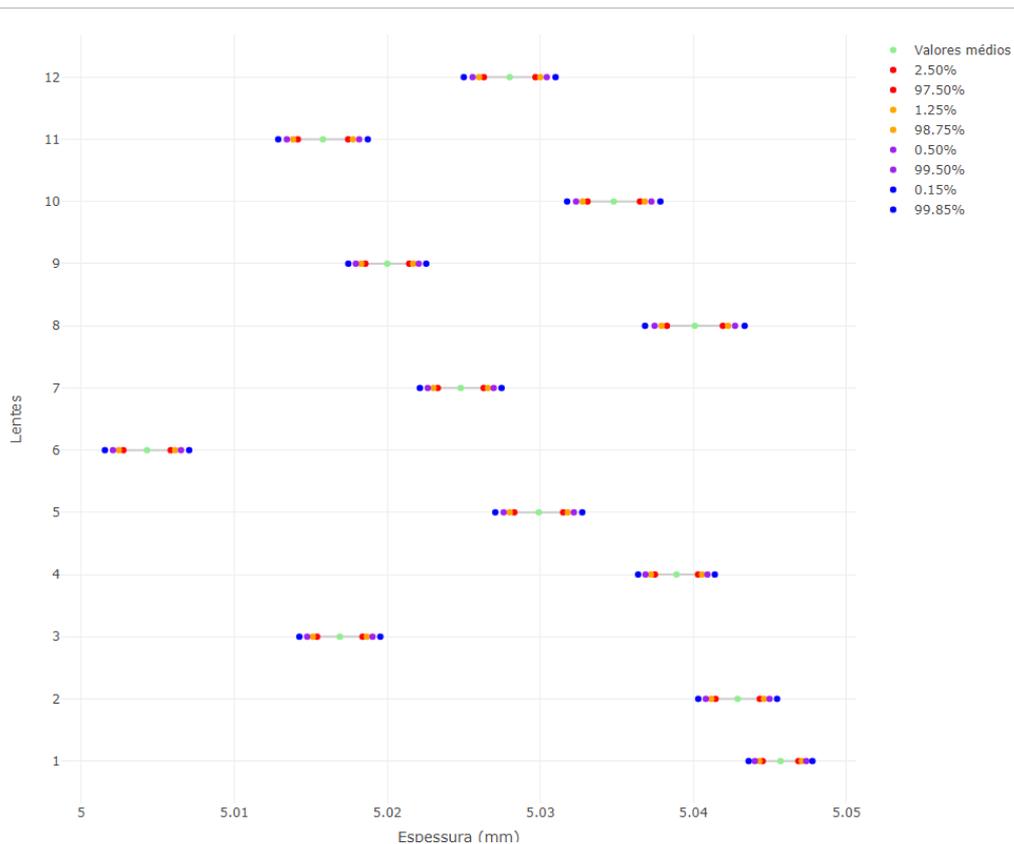


Figura 29: Intervalos de confiança para as medições da espessura no Conjunto B.

Na Figura 29 estão representados os intervalos de confiança para os valores médios das medições do espessura nas 12 lentes. Os intervalos foram calculados para um nível de confiança de 95%, 97,5%, 99% e 99,7%.

Embora não tão acentuada como no erro de centragem, a Figura 29 revela alguma variabilidade nos intervalos de confiança, de lente para lente. Desta vez, o erro humano associado a esta medição é maior, pois o método de medição da espessura é manual.

O passo seguinte, foi o cálculo das diferenças absolutas entre os extremos dos intervalos e os valores médios. O resultado desta diferença corresponde à incerteza da medição naquela lente, para um determinado nível de confiança. A diferença máxima observada de entre as 12 lentes foi considerada como sendo a incerteza da medição para a espessura.

A Tabela 14 apresenta as incertezas de medição para a espessura no Conjunto B, para diferentes níveis de confiança.

Tabela 14: Incerteza associada às medições da espessura.

Confiança	Incerteza (mm)
95%	0,002
97,5%	0,002
99%	0,003
99,7%	0,003

Para um nível de confiança de 99,7% concluiu-se que a incerteza associada à medição da espessura era de 0,003 milímetros.

5 Ajustamento das distribuições de probabilidade

O processo de ajustamento das distribuições de probabilidade foi realizado através dos quatro métodos para a estimação dos parâmetros.

Numa fase inicial, foi efetuada a estimação pontual para todos os conjuntos de dados e para os quatro métodos de estimação. Nesta secção, foram escolhidas as distribuições com melhor desempenho em cada um dos métodos para cada um das características da lente. Posteriormente, verificaram-se as diferenças entre quantis teóricos e empíricos para verificar qual o método com a distribuição mais próxima dos dados. Além disso, a qualidade do ajuste foi verificada através de representações gráficas para a função de distribuição.

Seguidamente, na estimação intervalar, calculam-se as incertezas associadas à estimação dos parâmetros para a distribuição que mais se aproximou dos dados. Para isso, foram geradas, a partir dos dados das medições, amostras por *bootstrap* paramétrico e, para cada, foi estimado o par de parâmetros da distribuição. Para a estimação final dos parâmetros foram usadas as médias e as medianas dos parâmetros nas amostras e foi estimada a incerteza através de um intervalo de confiança para a mediana dos parâmetros.

Para finalizar, efetuaram-se testes de ajustamento às distribuições eleitas e verificou-se a qualidade do ajuste através de representações gráficas.

5.1 Estimação pontual

Nesta secção serão apresentadas, para cada característica das lentes, as distribuições com melhor ajuste para o método de máxima verosimilhança (MMV), o *maximum goodness-of-fit* (MGOF), o método dos momentos (MM) e para o método da correspondência dos quantis (MCQ).

A escola das distribuições com melhor ajuste foi feita com base nos critérios de informação, AIC e BIC, e na distância *Anderson-Darling* (AD).

5.1.1 Diâmetro

No que diz respeito ao diâmetro da lente, as distribuições com melhor desempenho no ajuste foi diferente consoante o método aplicado. No entanto, se compararmos os conjuntos, apenas se verificou diferença na distribuição no método MM. Para o Conjunto A, a distribuição com melhor desempenho neste método foi a Logística, enquanto que no Conjunto B foi a Gama. Nestes dois conjuntos, as distribuições

ajustadas foram as mesmas nos restantes métodos, diferindo apenas na distância *Anderson-Darling*, para o método MGOF e no par de quantis, para o método MCQ.

Tais resultados podem ser consultados nas Tabela 15.

Tabela 15: Distribuições que apresentaram melhor ajustamento para o diâmetro no Conjunto B.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Log-logística	796,658	-1589,316	-1583,295	0,064	0,089	0,734
MGOF	Gama AD2	799,470	-1594,939	-1588,918	0,078	0,127	0,712
MM	Gama	799,614	-1595,227	-1589,206	0,065	0,112	0,786
MCQ	Gama (0,20; 0,90)	799,590	-1595,179	-1589,158	0,069	0,125	0,807

Os resultados para o Conjunto A estão disponíveis para consulta na Tabela 1 do Apêndice A.2.

5.1.2 Erro de centragem

Já no erro de centragem, as distribuições de probabilidade foram as mesmas com melhor ajuste para os quatro métodos.

Como os valores do erro de centragem, para os dois conjuntos, estavam compreendidos entre 0 e 1, a distribuição Beta foi a melhor opção. Além disso, é uma distribuição bastante flexível, capaz de assumir diversas formas, sendo mais provável apresentar um bom ajuste, mesmo quando a forma dos dados muda.

Tabela 16: Distribuições que apresentaram melhor ajustamento para o erro de centragem no Conjunto A.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Beta	123,362	-242,724	-236,030	0,051	0,068	0,641
MGOF	Beta ADL	123,305	-242,610	-235,916	0,044	0,062	0,596
MM	Beta	123,352	-242,703	-236,009	0,048	0,064	0,608
MCQ	Beta (0,25; 0,80)	123,261	-242,522	-235,828	0,043	0,062	0,618

Tal como sugerem as tabelas 16 e 2 (consultar Apêndice A.2), apenas se verificaram diferenças entre os conjuntos nos métodos MGOF, na distância *Anderson-Darling* e num dos quantis escolhidos, no método MCQ.

5.1.3 Espessura

Para as medições da espessura, as distribuições não diferiram consoante o método utilizado, dentro do mesmo conjunto. No entanto, entre conjuntos, as distribuições com melhor ajuste não foram as mesmas. Supeitou-se que a tolerância associada à espessura tinha influência na forma dos dados e, conseqüentemente, na distribuição ajustada.

Tabela 17: Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto B.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Normal	435,234	-866,468	-860,447	0,077	0,114	0,645
MGOF	Normal ADL	434,597	-865,193	-859,172	0,062	0,070	0,502
MM	Normal	435,234	-866,468	-860,447	0,077	0,114	0,645
MCQ	Normal (0,05; 0,85)	435,188	866,375	-860,354	0,070	0,113	0,630

A distribuição que apresentou um melhor ajuste aos dados do Conjunto B foi a distribuição Normal. Tal como sugere a Tabela 17, as distribuições ajustadas pelos quatro métodos estão bastante próximas no que diz respeito à log-verosimilhança, ao AIC e à distância AD.

O Conjunto C apresenta uma tolerância assimétrica, inferior à dos dois primeiros conjuntos. Foi esperado que as lentes deste conjunto apresentassem uma distribuição diferente, distante da Normal. Na Tabela 18 são apresentadas as distribuições com melhor ajuste nos quatro métodos de estimação para este conjunto.

Tabela 18: Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto C.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Weibull	416,036	-828,072	-822,943	0,136	0,222	1,418
MGOF	Weibull ADL	413,695	-823,391	-818,262	0,123	0,228	1,536
MM	Weibull	413,286	-822,572	-817,444	0,120	0,257	1,605
MCQ	Weibull (0,20; 0,65)	416,030	-828,059	-822,930	0,140	0,235	1,468

A Tabela 18 sugere que a distribuição Weibull é a melhor alternativa para o ajuste às medições da espessura. A diferença entre os quatro métodos é mais acentuada, verificando-se maior Log-verosimilhança e menor distância *Anderson-Darling* para os métodos MMV e MCQ.

O Conjunto E apresenta uma tolerância ainda mais reduzida, pois a sua tolerância superior para a espessura é 0. Na Tabela 19 apresentam-se as distribuições com melhor ajuste.

Tabela 19: Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto E.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Log-logística	319,219	-634,438	-629,803	0,094	0,097	0,644
MGOF	Logística AD2L	319,061	-634,123	-629,488	0,101	0,105	0,661
MM	Logística	313,186	-634,371	-629,736	0,108	0,112	0,683
MCQ	Logística (0,25; 0,80)	313,118	-634,236	-629,601	0,092	0,094	0,634

Tal como sugere a Tabela 19, não se obtiveram as mesmas distribuições nos quatro métodos. Ainda que se verifiquem algumas diferenças entre a verosimilhança, o AIC e a estatística *Anderson-Darling* revelam que as quatro alternativas para o ajuste aos dados estão próximas.

Tabela 20: Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto G.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Weibull	282,238	-560,476	-555,712	0,084	0,044	0,268
MGOF	Weibull ADL	282,196	-560,393	-555,629	0,086	0,042	0,261
MM	Weibull	282,186	-560,372	-555,608	0,077	0,044	0,278
MCQ	Weibull (0,20; 0,90)	281,396	-558,791	-554,027	0,117	0,110	0,645

Para o Conjunto G, a Tabela 20 revela que a distribuição Weibull é a melhor candidata ao ajuste e que os métodos estão próximos entre eles, com exceção do método MCQ, cuja distância AD é superior às restantes.

Tabela 21: Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto J.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Weibull	272,096	-540,191	-535,192	0,053	0,028	0,180
MGOF	Weibull ADL	272,059	-540,117	-535,118	0,047	0,026	0,193
MM	Weibull	271,758	-539,516	-534,516	0,063	0,046	0,262
MCQ	Weibull (0,35; 0,75)	272,089	-540,179	-535,179	0,048	0,027	0,179

No Conjunto J a distribuição Weibull apresentou, uma vez mais, o melhor ajuste, para cada um dos

métodos. A Tabela 21 sugere que as distribuições encontram-se próximas entre si e no que diz respeito à distribuição empírica dos dados, uma vez que as estatísticas *Anderson-Darling* assumem valores muito reduzidos.

As distribuições com melhor ajuste para os restantes Conjuntos estão disponíveis no Apêndice A.2, da Tabela 3 à 8.

5.1.4 Desfasamento entre quantis teóricos e empíricos

Com o intuito de avaliar a qualidade do ajustamento foram construídos indicadores baseados nos quantis teóricos e empíricos das distribuições de dados e das distribuições de probabilidade. Cada desfasamento corresponde a uma diferença entre um par de quantis (5.33). Foram calculados os quantis desde o $Q_{0,01}$ até ao $Q_{0,99}$, inclusive, com o espaçamento de 0,01 unidades.

$$Desfasamentos = Q_{teórico} - Q_{empírico}, \quad (5.33)$$

sendo Desfasamento Máximo Absoluto (DMA) dado pela seguinte equação:

$$DMA = \max |Q_{teórico} - Q_{empírico}| \quad (5.34)$$

Quanto menor forem os desfasamentos entre os quantis teóricos e os empíricos melhor a qualidade do ajustamento.

Adicionalmente, foram calculados os seguintes erros, métricas para a qualidade do ajuste:

$$EQM = \frac{1}{n} \sum_{i=1}^n (Q_{teórico_i} - Q_{empírico_i})^2 \quad (5.35)$$

$$REQM = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_{teórico_i} - Q_{empírico_i})^2} \quad (5.36)$$

$$EAM = \frac{1}{n} \sum_{i=1}^n |Q_{teórico_i} - Q_{empírico_i}| \quad (5.37)$$

$$EPAM = \frac{1}{n} \sum_{i=1}^n \frac{|Q_{teórico_i} - Q_{empírico_i}|}{Q_{empírico_i}} \times 100 \quad (5.38)$$

Das quatro métricas utilizadas destacou-se o desempenho da métrica Erro absoluto médio (EAM).

As Tabelas que se seguem, da Tabela 22 à 24, apresentam os desvios máximos absolutas entre os quantis teóricos e empíricos e os respetivos erros.

Tabela 22: Desfasamento entre os quantis teóricos e os empíricos para o diâmetro no conjunto A.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Log-logística	0,0016	$9,9300 \times 10^{-8}$	0,0002	0,0010	0,0003
MGOF	Gama ADL	0,0010	$8,5300 \times 10^{-8}$	0,0002	0,0010	0,0003
MM	Logística	0,0012	$6,2500 \times 10^{-8}$	0,0002	0,0010	0,0003
MCQ	Gama (0,45; 0,65)	0,0009	$8,9400 \times 10^{-8}$	0,0002	0,0010	0,0003

Para as medições do diâmetro no Conjunto A, a distribuição Logística pelo método dos momentos surge como a melhor opção, pois apresenta menor Erro quadrático médio (EQM) e Erro absoluto médio (EAM).

Tabela 23: Desfasamento entre os quantis teóricos e os empíricos para o erro de centragem no conjunto A.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Beta	0,0749	$1,8300 \times 10^{-4}$	0,0094	2,4600	0,0135
MGOF	Beta ADL	0,0750	$1,8000 \times 10^{-4}$	0,0087	2,3000	0,0134
MM	Beta	0,0756	$1,7800 \times 10^{-4}$	0,0089	2,3600	0,0134
MCQ	Gama (0,45; 0,80)	0,0722	$1,8700 \times 10^{-4}$	0,0093	2,4200	0,0137

No que diz respeito ao erro de centragem, para o Conjunto A, a distribuição Beta ADL é a que apresenta menor Erro absoluto médio e *Erro percentual absoluto médio*.

Tabela 24: Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto A.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Normal	0,0049	$1,4800 \times 10^{-6}$	0,0009	0,0200	0,0012
MGOF	Normal ADL	0,0072	$2,5600 \times 10^{-6}$	0,0010	0,0300	0,0016
MM	Normal	0,0049	$1,4800 \times 10^{-6}$	0,0009	0,0200	0,0012
MCQ	Normal (0,20; 0,70)	0,0048	$1,4700 \times 10^{-6}$	0,0009	0,0200	0,0012

Já na espessura, para o Conjunto A, foi escolhida a distribuição Normal pelo método de máxima verosimilhança, pois apresentou menor Desfasamento Máximo Absoluto (DMA) e menor Erro absoluto

médio (EAM).

Os resultados para as restantes lentes foram remetidos para o Apêndice A.3 e estão disponíveis desde a Tabela 9 à 20.

5.2 Estimação intervalar

A estimação intervalar foi efetuada com o propósito de estimar os parâmetros e a incerteza associada à sua estimação. Para isso, foi usado *bootstrap* paramétrico para simular amostras, provenientes dos dados das medições. Para as 1001 amostras foram estimados os parâmetros e calculados os intervalos de confiança a 95% para o valor da mediana.

O primeiro passo foi perceber que valores assumia cada parâmetro nas várias amostras simuladas.

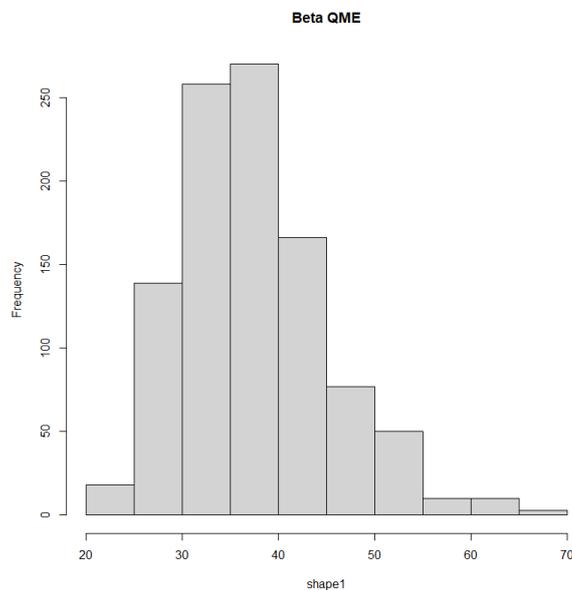


Figura 30: Valores do parâmetro α estimados para as 1001 amostras do Conjunto B na distribuição Beta

A estimação de cada parâmetro será dada pela mediana das 1001 amostras, caso a distribuição empírica do parâmetro seja simétrica. Caso contrário, como no exemplo da Figura 30, a estimação do parâmetro α será dada pelo valor médio dos 1001 valores. Este processo será igualmente aplicado ao segundo parâmetro de cada distribuição.

A Tabela 25 apresenta as medianas dos parâmetros das distribuições para as 1001 amostras do diâmetro.

Tabela 25: Médias dos parâmetros para as amostras do diâmetro

Conjunto	Distribuição	μ	σ
A	Logística MM	23,181	0,001
B	Log-logística MMV	21186,139	14,481

A Tabela 26 apresenta as médias dos parâmetros das distribuições para as 1001 amostras do erro de centragem. Serão usadas as médias para cada parâmetro pois, nos Conjuntos A e B, α e β assumiram distribuições assimétricas.

Tabela 26: Medianas dos parâmetros para as amostras do erro de centragem.

Conjunto	Distribuição	α	β
A	Beta MGOF	5,435	6,906
B	Beta MCQ	37,412	28,042

A Tabela 27 apresenta as médias e as medianas dos parâmetros das distribuições para as 1001 amostras da espessura. Para os parâmetros β da distribuição Weibull dos Conjuntos F e I, e para o σ da distribuição Normal do Conjunto I foram calculados os valores médios, uma vez que as distribuições destes parâmetros eram assimétricas. Para os restantes parâmetros, que apresentam distribuições simétricas, estão representadas as medianas.

Tabela 27: Médias e medianas dos parâmetros para as amostras da espessura

Conjunto	Distribuição	Parâmetros	
A	Normal MMV	$\mu = 4,033$	$\sigma = 0,011$
B	Normal MMV	$\mu = 5,027$	$\sigma = 0,013$
C	Weibull MM	$\alpha = 1,637$	$\beta = 527,406$
D	Weibull MGOF	$\alpha = 3,173$	$\beta = 594,734$
E	Logística MM	$\mu = 6,795$	$\sigma = 0,002$
F	Weibull MMV	$\alpha = 6,487$	$\beta = 1019,351$
G	Weibull MM	$\alpha = 6,603$	$\beta = 1057,507$
H	Weibull MGOF	$\alpha = 3,322$	$\beta = 417,519$
I	Normal MMV	$\mu = 1,264$	$\sigma = 0,007$
J	Weibull MMV	$\alpha = 6,236$	$\beta = 622,843$
K	Weibull MM	$\alpha = 4,866$	$\beta = 1187,936$

É possível comparar as distribuições de probabilidade empíricas com as distribuições de probabilidade teóricas com parâmetros definidos, como os parâmetros estimados na Tabela 27.

A Figura 31 representa os quatro gráficos *goodness-of-fit* para o Conjunto F, onde foi definida a distribuição Weibull($\alpha = 3,173$; $\beta = 594,734$).

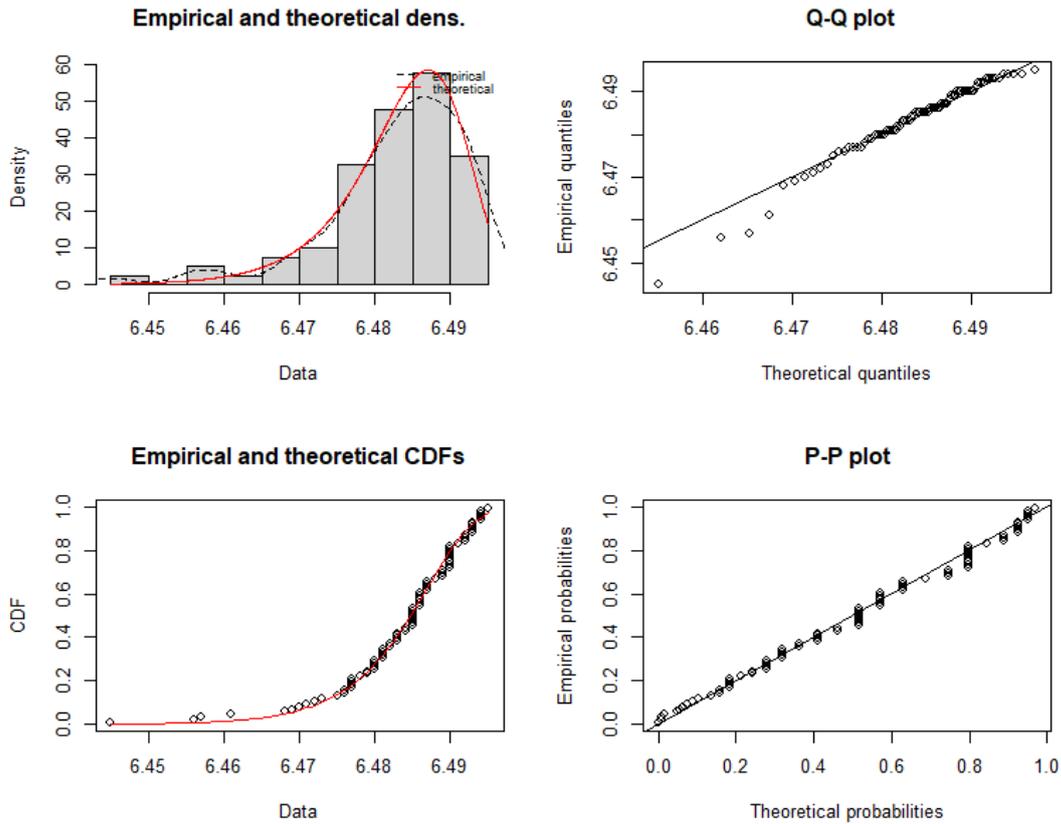


Figura 31: Distribuição empírica vs distribuição Weibull ajustada para o conjunto F.

No que diz respeito à densidade, a distribuição teórica, representada a vermelho, na Figura 31, está próxima da função densidade empírica, conseqüentemente, sobrepõe-se ao histograma das medições da espessura.

Relativamente aos quantis, os da distribuição Weibull, estão próximos dos quantis empíricos. Neste segundo gráfico, da Figura 31, a proximidade entre quantis empíricos e teóricos é traduzida pela sobreposição dos pontos na linha reta. No entanto, existem alguns pontos que não se aproximam da reta, que dizem respeito a valores menores da espessura, distantes dos valores mais frequentes da amostra e portanto, mais difíceis de ajustar.

Seguidamente, no gráfico inferior esquerdo da Figura 31 são apresentadas as funções de distribuição teórica e empírica. Uma vez mais, caso os pontos se aproximem da curva, estamos perante um bom ajustamento da distribuição teórica, algo que se verifica.

Por último, no gráfico *P-P* da Figura 31 comparam-se as probabilidade empíricas e as probabilidades teóricas. Uma vez mais, como os pontos se sobrepõem à linha reta, estamos perante um bom ajuste, pois as probabilidades teóricas estão próximas às probabilidades empíricas.

5.2.1 Incerteza associada à estimação dos parâmetros

A incerteza associada à estimação dos parâmetros das distribuições ajustadas aos parâmetros das lentes foi efetuada através de um intervalo de confiança a 95%. Da Tabela 28 à 30 são apresentados os percentis 2,5 e 97,5 para a mediana dos parâmetros, das 1001 amostras, obtidos por *bootstrap* paramétrico.

Tabela 28: Intervalos de confiança para os parâmetros das distribuições de probabilidade ajustadas ao diâmetro

Conjunto	Distribuição	α		β	
		2,5	97,5	2,5	97,5
A	Logística MM	23,1805	23,1810	0,0009	0,0012
B	Log-logística MMV	14,4807	14,4810	18743,8640	24428,0151

A Tabela 28 revela que as incertezas associadas aos parâmetros de localização, para o Conjunto A, e para o de escala, para o Conjunto B, é bastante reduzida.

No que diz respeito ao segundo parâmetro, a incerteza associada à estimação do parâmetro de dispersão, no Conjunto A, é também reduzido. No entanto, dada à diferente escala do parâmetro de forma da distribuição Log-logística, a incerteza associada a este parâmetro é maior.

Tabela 29: Intervalos de confiança para os parâmetros das distribuições de probabilidade ajustadas ao erro de centragem.

Conjunto	Distribuição	α		β	
		2,5	97,5	2,5	97,5
A	Beta MGOF	4,332	6,776	5,394	8,712
B	Beta MCQ	25,583	54,395	19,356	41,190

A Tabela 29 sugere que a incerteza associada aos parâmetros das distribuição Beta é maior no Conjunto B.

Tabela 30: Intervalos de confiança para os parâmetros das distribuições de probabilidade ajustadas à espessura.

Conjunto	Distribuição	α		β	
		2,5%	97,5%	2,5%	97,5%
A	Normal MMV	4,032	4,035	0,010	0,012
B	Normal MMV	5,025	5,029	0,012	0,015
C	Weibull MM	1,636	1,638	527,177	527,613
D	Weibull MGOF	3,172	3,174	535,353	662,084
E	Logística MM	6,794	6,796	0,0015	0,0023
F	Weibull MMV	6,485	6,489	870,609	1226,245
G	Weibull MM	6,601	6,605	1057,490	1057,490
H	Weibull MGOF	3,321	3,323	366,467	483,532
I	Normal MMV	1,263	1,266	0,006	0,007
J	Weibull MMV	6,233	6,238	535,108	742,715
K	Weibull MM	4,865	4,867	1187,468	1188,351

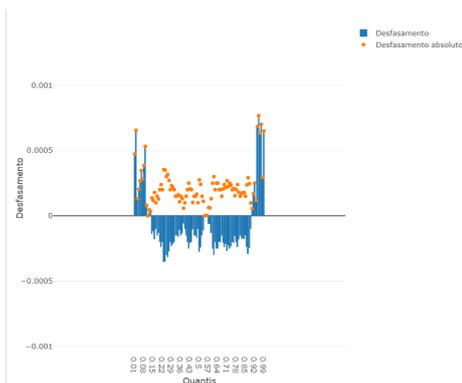
Sendo que α representa o parâmetro de localização e β o de dispersão.

Os intervalos de confiança associados à estimação dos parâmetros são reduzidos, excetuando os Conjuntos D, F, H e J, em que o intervalo de confiança para β tem uma amplitude grande, como revela a Tabela 30.

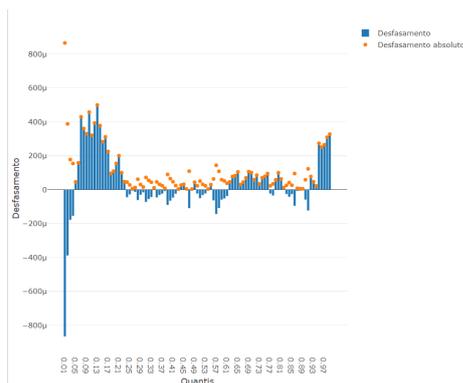
5.2.2 Desfasamento entre quantis teóricos e empíricos

Para verificar a proximidade entre as distribuições ajustadas e os dados simularam-se amostras, com as mesmas dimensões dos conjuntos de lentes medidos, provenientes das distribuições ajustadas e com os parâmetros estimados. Depois, calcularam-se os quantis empíricos e dos dados simulados e a diferença entre os dois, a que chamamos de desfasamentos.

A Figura 32 apresenta os desfasamentos nos Conjuntos A e B para o diâmetro.



(a) Desfasamentos no conjunto A



(b) Desfasamentos no conjunto B

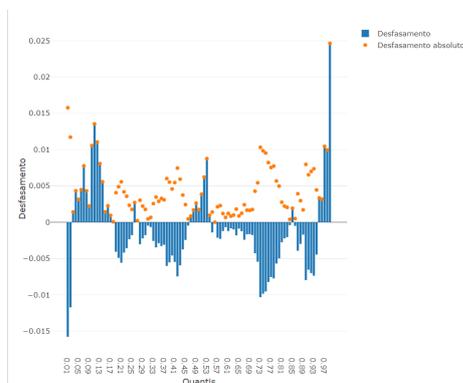
Figura 32: Desfasamentos entre os quantis teóricos e empíricos para as medições do diâmetro nos Conjuntos A e B

O erro absoluto médio dos desfasamentos é de 0,0002 milímetros no Conjunto A, enquanto que para o Conjunto B é de 0,0001 milímetros. Além disso, a Figura 32 revela que os desfasamentos são maiores nas caudas das distribuições, isto é, as diferenças são maiores para os quantis nos extremos, os menores e os maiores. Pelo contrário, na parte central da distribuição, as diferenças são menores.

A Figura 33 apresenta os desfasamentos nos Conjuntos A e B para o erro de centragem.



(a) Desfasamentos no conjunto A



(b) Desfasamentos no conjunto B

Figura 33: Desfasamentos entre os quantis teóricos e empíricos para as medições do erro de centragem nos Conjuntos A e B

No Conjunto A, os desfasamentos foram maiores nos quantis mais baixos, como revela a Figura 5.33(a). Depois, nos restantes quantis, os desfasamentos são relativamente reduzidos, com exceção do último. Neste conjunto o erro absoluto médio foi de 0,009 minutos.

Por outro lado, no Conjunto B, os desfasamentos foram menores, como sugere a Figura 5.33(b).

Registou-se um erro absoluto médio de 0,004 minutos algumas oscilações pelos desfasamentos entre quantis. No entanto, estas diferenças são bastante menores comparadas com o Conjunto A.

A Figura 34 apresenta os desfasamentos nos Conjuntos A a D para a espessura.

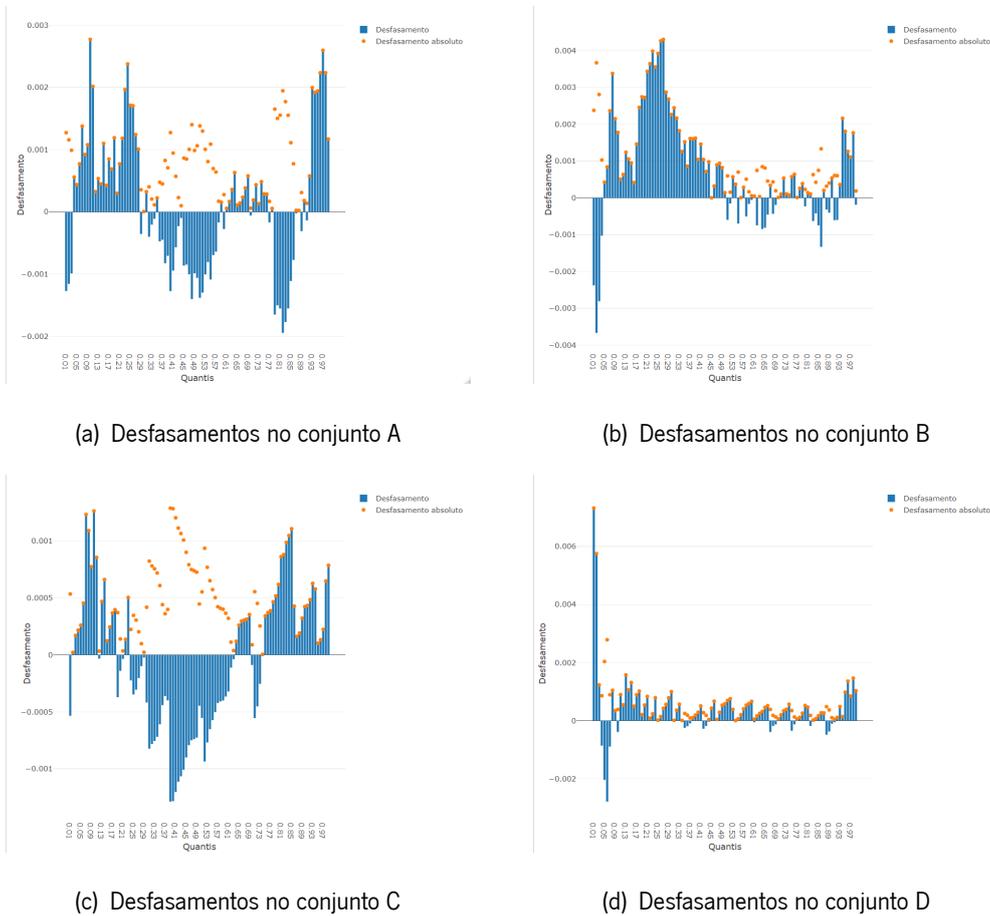


Figura 34: Desfasamentos entre os quantis teóricos e empíricos para as medições da espessura nos Conjuntos A ao D.

Para os dados da espessura, verificaram-se algumas oscilações, que variam de conjunto para conjunto, tal como ilustrado na Figura 34, para os desfasamentos entre os quantis menores.

No Conjunto A, tal como sugere a Figura 5.34(a), verificam-se várias variações nos desfasamentos, sendo maiores nas caudas da distribuição e no centro.

A Figura 5.34(b) revela que os desfasamentos no Conjunto B são superiores nos quantis menores e diminuem ao longo dos quantis da parte central da distribuição. Já na cauda superior, os desfasamentos entre quantis aumentam.

No conjunto C, tal como sugere a Figura 5.34(c) os desfasamentos são maiores na parte central, ainda que não muito elevados.

A Figura 5.34(d) revela que o Conjunto D tem os desfasamentos mais constantes, com exceção dos que são relativos aos quantis iniciais, que assumem um valor absoluto maior.

Os desfasamentos entre os quantis teóricos e empíricos para os restantes conjuntos foram remetidos para o Apêndice A.4 e estão disponíveis para consulta nas Figuras 3 e 4.

5.3 Testes de ajustamento

O ajuste das distribuições foi confirmado através do Teste de ajustamento de Qui-quadrado e do Teste de Kolmogorov-Smirnov.

Para os testes de ajustamento de qui-quadrado, as classes foram construídas baseadas nos decis da amostra empírica e da distribuição ajustada.

Foi testado se os dados das medições efetuadas para o diâmetro, erro de centragem e espessura eram provenientes de de uma função distribuição com parâmetros especificados. Esses parâmetros correspondem aqueles que foram estimados na secção 5.2, apresentados nas Tabelas 25, 26 e 27.

Os resultados dos testes de ajustamento são apresentados nas Tabelas 31, 32 e 33. Para cada teste está especificado a estatística de teste e o respetivo valor- p .

Assim sendo, a Tabela 31 contém os resultados dos testes de ajustamento para as distribuições de probabilidade ajustadas às medições do diâmetro, nos Conjuntos A e B.

Tabela 31: Testes de ajustamento para as distribuições de probabilidade ajustadas ao diâmetro.

Conjunto	Distribuição	Qui-quadrado		Kolmogorov-Smirnov	
		U	valor- p	D	valor- p
A	Logística MME	16,568	0,056	0,063	0,370
B	Log-logística MLE	27,911	0,001	0,063	0,590

No que diz respeito ao teste de ajustamento de Qui-quadrado, a estatística de teste apresenta valores elevados, que indicam uma grande discrepância entre as frequências observadas e as esperadas, embora no Conjunto A o valor- p seja ligeiramente superior a 0,05, tal como sugera a Tabela 31. Deste modo, para o Conjunto A, pode dizer-se que, ao nível de 5% de significância, não existem evidências estatísticas para se rejeitar a hipótese nula de que os dados são provenientes da distribuição ajustada. Pelo contrário, no Conjunto B, para o mesmo nível de significância, existem evidências estatísticas para se rejeitar a hipótese nula.

Por outro lado, no teste de Kolmogorov-Smirnov, a estatística de teste observada, D , apresentou valores bastante reduzidos, que traduzem a proximidade entre as funções de distribuição empírica e as teóricas. Além disso, em ambos os Conjuntos, o valor- p foi superior a 0,05. Assim sendo, pelo teste de Kolmogorov-Smirnov, ao nível de 5% de significância, não existem evidências estatísticas para se rejeitar a hipótese nula de que os dados são provenientes das distribuições ajustadas.

Seguidamente, na Tabela 32 inclui os resultados dos testes de ajustamento para as distribuições de probabilidade ajustadas às medições do diâmetro, nos Conjuntos A e B.

Tabela 32: Testes de ajustamento para as distribuições de probabilidade ajustadas ao erro de centragem.

Conjunto	Distribuição	Qui-quadrado		Kolmogorov-Smirnov	
		U	valor- p	D	valor- p
A	Beta MGE ADL	8,784	0,457	0,046	0,777
B	Beta QME	3,976	0,913	0,049	0,862

Desta vez, como revela a Tabela 32, a estatística de teste de Qui-quadrado apresenta valores menores, indicando uma maior proximidade entre as frequências observadas e as esperadas. Além disso, os valores- p são superiores a 0,05 confirmando que, ao nível de 5% de significância, não existem evidências estatísticas para se rejeitar a hipótese nula de que os dados são provenientes das distribuições ajustadas.

Adicionalmente, no teste de Kolmogorov-Smirnov, a estatística de teste observada, D , apresentou, também, valores bastante baixos, que confirmam a proximidade entre as funções de distribuição empírica e as teóricas. Além disso, em ambos os Conjuntos, o valor- p foi superior a 0,05. Assim sendo, pelo teste de Kolmogorov-Smirnov, ao nível de 5% de significância, não existem evidências estatísticas para se rejeitar a hipótese nula de que os dados são provenientes das distribuições ajustadas.

Depois, a Tabela 33 apresenta os resultados dos testes de ajustamento para as distribuições de probabilidade ajustadas às medições da espessura, do Conjuntos A ao K.

Tabela 33: Testes de ajustamento para as distribuições de probabilidade ajustadas à espessura.

Conjunto	Distribuição	Qui-quadrado		Kolmogorov-Smirnov	
		<i>U</i>	valor- <i>p</i>	<i>D</i>	valor- <i>p</i>
A	Normal MLE	9,043	0,433	0,068	0,299
B	Normal MLE	19,381	0,221	0,078	0,318
C	Weibull MME	25,437	0,003	0,120	0,125
D	Weibull MGE	92,073	$6,242 \times 10^{-16}$	0,058	0,308
E	Logística MME	27,581	0,001	0,112	0,303
F	Weibull MLE	21,886	0,009	0,086	0,601
G	Weibull MME	11,707	0,230	0,073	0,783
H	Weibull MGE	20,682	0,014	0,088	0,089
I	Normal MLE	22,057	0,009	0,061	0,857
J	Weibull MLE	7,889	0,545	0,051	0,973
K	Weibull MME	28,756	0,001	0,114	0,322

Novamente, no que diz respeito ao teste de ajustamento de Qui-quadrado, a estatística de teste apresenta valores elevados para alguns conjuntos (Tabela 33). São exemplos disso os Conjuntos C, D, E, F, H, I ou K, onde se verificou uma grande discrepância entre as frequências observadas e as esperadas. Para estes conjuntos, existem evidências estatísticas, para um nível de significância de 5%, para se rejeitar a hipótese nula de que os dados não provenientes das respectivas distribuição ajustadas, uma vez que os valores-*p* são inferiores a 0,05. Pelo contrário, nos Conjuntos A, B, G e J, não existe evidência estatísticas, para se rejeitar a hipótese nula.

Porém, no teste de Kolmogorov-Smirnov, a estatística de teste observada, *D*, apresentou valores bastante reduzidos, que confirmam a proximidade entre as funções de distribuição empírica e as teóricas, ao contrário do teste anterior. Além disso, em todos, o valor-*p* observado foi superior a 0,05. Deste modo, pelo teste de Kolmogorov-Smirnov, ao nível de 5% de significância, não existem evidências estatísticas para se rejeitar a hipótese nula de que os dados são provenientes das distribuições ajustadas.

5.4 Qualidade do ajuste

Em seguida, os gráficos que se seguem, da Figura 35 à 37, são relativos às funções de distribuição das distribuições ajustadas com intervalos de confiança a 99,7%. A utilização destes intervalos de confiança com amplitude de $3 \times \sigma$ é uma opção a nível empresarial.

Assim sendo, Figura 35 apresenta as funções de distribuição das distribuições ajustadas para as medições do diâmetro, nos Conjuntos A e B.

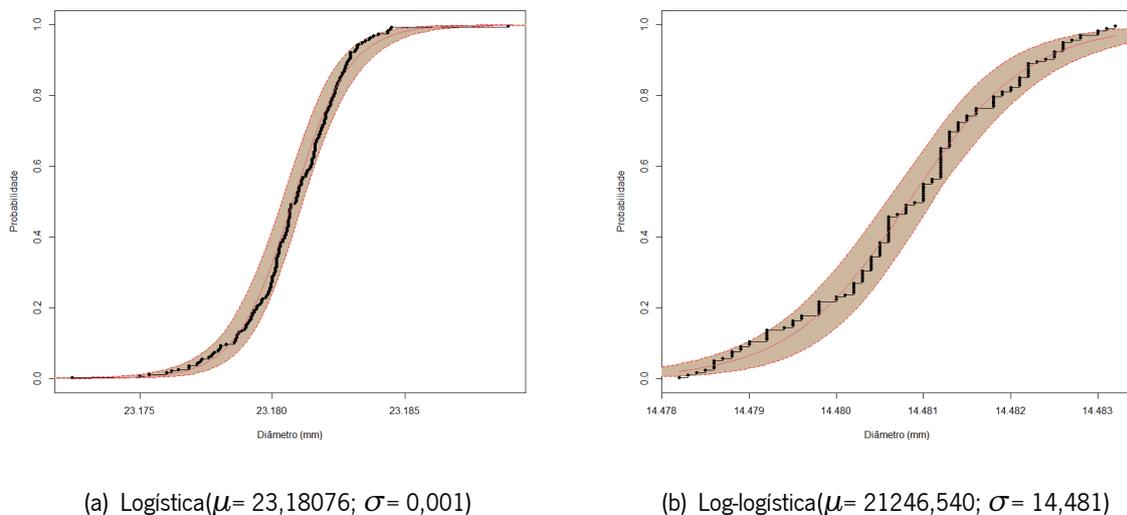
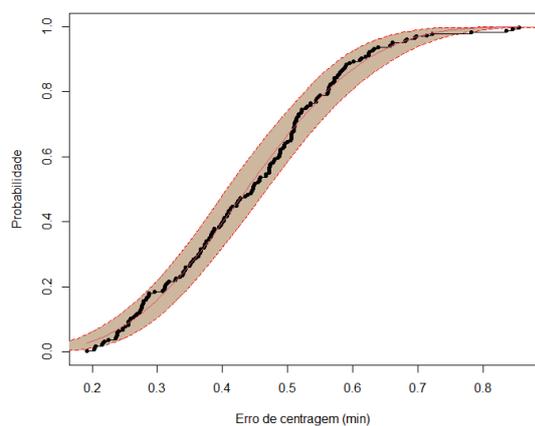


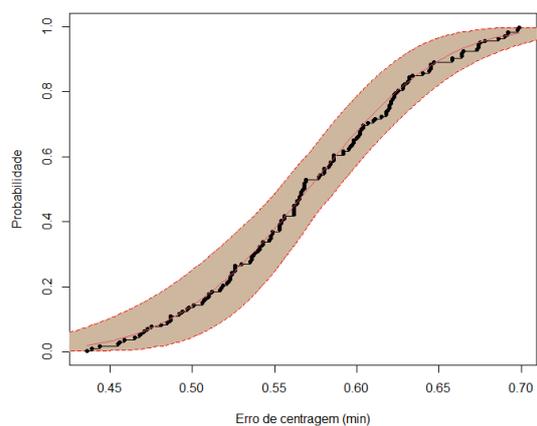
Figura 35: Qualidade de ajuste das distribuições ajustadas ao diâmetro nos Conjuntos A e B.

A Figura 35 sugere que, graficamente, a distribuição ajusta-se bem aos dados, na medida em que apenas os outliers, ou valores mais distantes da parte central da distribuição, estão fora dos intervalos de confiança, nos dois conjuntos.

Seguidamente, a Figura 36 apresenta as funções de distribuição das distribuições ajustadas para as medições do erro de centragem, nos Conjuntos A e B.



(a) Beta($\alpha= 5,485$; $\beta= 6,979$)

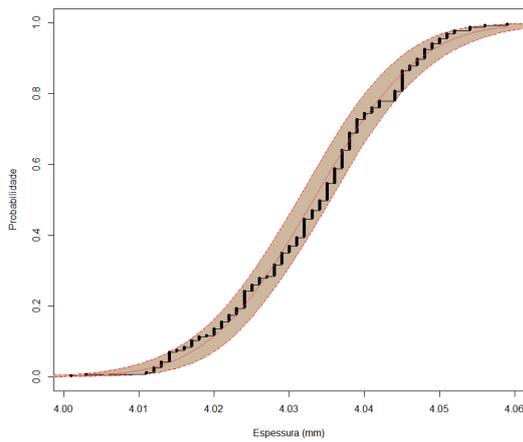


(b) Beta($\alpha= 36,321$; $\beta= 27,351$)

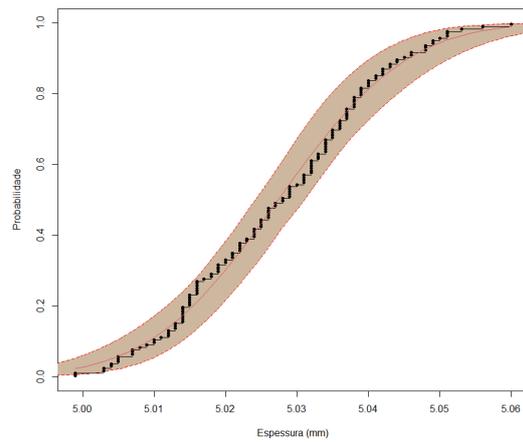
Figura 36: Qualidade de ajuste das distribuições ajustadas ao erro de centragem nos Conjuntos A e B.

Seguidamente, no erro de centragem também se verificaram pontos fora das bandas de confiança no Conjunto A, como revela a Figura 5.36(a). Novamente, foram nas caudas das distribuições que a distribuição falhou no ajuste. No entanto, no Conjunto B, a distribuição efetuou um ajuste melhor, pois fez uma boa aproximação a todos os pontos, situando-se todos dentro das bandas de confiança, tal como sugere a Figura 5.36(b).

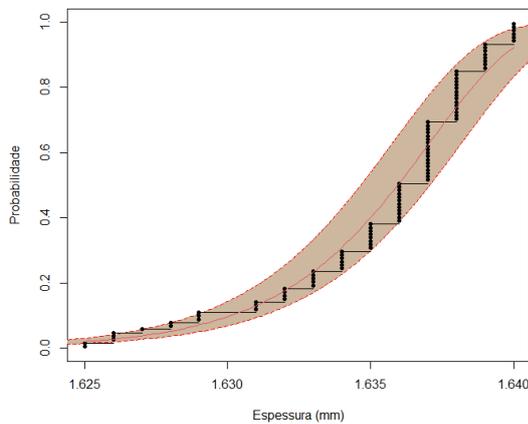
Depois, segue-se a Figura 37, que apresenta as funções de distribuição das distribuições ajustadas para as medições da espessura, nos Conjuntos A, B, C e D.



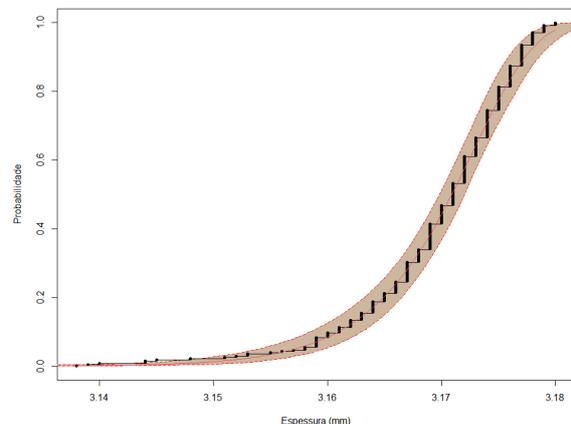
(a) Normal($\mu = 4,033$; $\sigma = 0,011$)



(b) Normal($\mu = 5,027$; $\sigma = 0,013$)



(c) Weibull($\alpha = 1,637$; $\beta = 527,406$)



(d) Weibull($\alpha = 3,173$; $\beta = 594,734$)

Figura 37: Qualidade de ajuste das distribuições ajustadas à espessura nos Conjuntos A a D.

Nas distribuições ajustadas às medições da espessura, a tendência dos pontos pertencentes às caudas da distribuição situarem-se fora das bandas de confiança manteve-se.

No Conjunto A todos os pontos apresentam-se dentro das bandas de confiança, como ilustrado na Figura 5.37(a), enquanto que no Conjunto B, como sugere a Figura 5.37(b), o mínimo surge fora da banda de confiança. Já no Conjunto C surgem pontos nas caudas da distribuição, que assumem as probabilidades menores e maiores, fora do intervalo de confiança a 99,7%, como revela a Figura 5.37(c). Adicionalmente, no Conjunto D, esta tendência mantém-se, com alguns dos quantis menores e maiores a estarem fora das bandas de confiança, como é visível na Figura 5.37(d).

A qualidade de ajuste para as distribuições de probabilidade ajustadas aos restantes conjuntos é ilustrada no Apêndice A.5, nas Figuras 5 e 6.

6 Previsão dos parâmetros das distribuições de probabilidade

O objetivo principal é prever qual o comportamento de uma lente nova na produção em série. Para isso, pretende-se efetuar uma previsão da forma da distribuição de uma lente tendo em conta as características da mesma, como os valores nominais ou as tolerâncias.

Assim sendo, estudou-se a correlação entre os parâmetros das distribuições ajustadas nas medições da espessura nos vários conjuntos e as características da lente, com o intuito de identificar possíveis relações entre a geometria da lente e a forma da distribuição.

Seguidamente, contruíram-se modelos para a previsão dos parâmetros das distribuições de probabilidade, tendo como variáveis preditoras os parâmetros da lente com mais relevância nas correlações.

Por fim, compararam-se as previsões efetuadas através dos modelos com as Simulações Monte Carlo, tendo como referência as medições realizadas.

6.1 Tendências da distribuição da espessura

Como abordado no Capítulo 1, as distribuições de probabilidade associadas às medições da espessura apresentam um deslocamento para a direita face ao valor nominal. Além disso, quanto menos dureza o vidro apresentar, mais ampla será a distribuição da espessura em torno da média. Estas tendências relativas ao comportamento da distribuição da espessura são ilustradas na Figura 38.

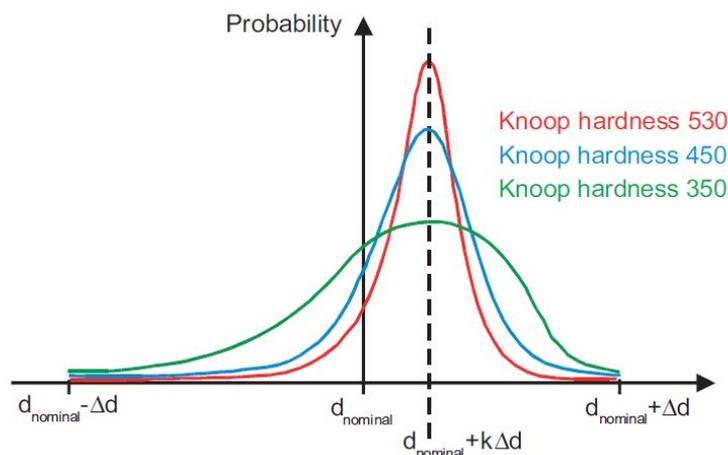


Figura 38: Distribuições de probabilidade consoante a dureza do material (Gross, 2007).

O $d_{nominal} + \Delta d$ corresponde ao máximo enquanto que o $d_{nominal} - \Delta d$ diz respeito ao mínimo. O $d_{nominal} + k\Delta d$ está no centro da distribuição, onde se encontra o parâmetro de localização (média)

nas distribuições Normal e Logística, *scale* na distribuição Weibull). Por consequência, o $k\Delta d$ traduz a distância do parâmetro de localização ao valor nominal.

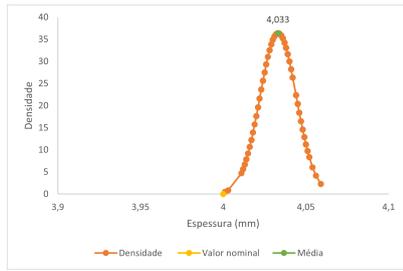
Na Tabela 34 constam os conjuntos das lentes, por ordem crescente relativamente ao valor nominal e as respectivas distâncias ao centro da distribuição.

Tabela 34: Distância do parâmetro de localização ao valor nominal.

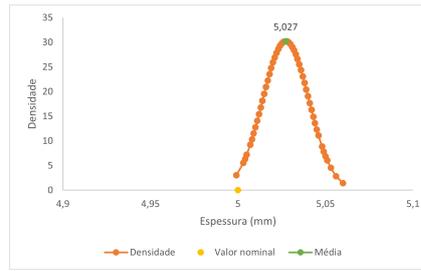
Conjunto	Nominal (mm)	$k\Delta d$(mm)
I	1,200	0,064
C	1,610	0,027
D	3,150	0,023
H	3,300	0,019
A	4,000	0,033
K	4,840	0,024
B	5,000	0,027
J	6,200	0,030
F	6,470	0,017
G	6,600	0,003
E	6,800	-0,005

De um modo geral, à medida que o valor nominal aumenta, a distância entre parâmetro de localização e o valor nominal da espessura foi diminuindo, como sugere a Tabela 34.

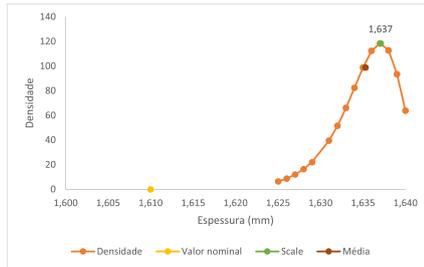
Para a maior parte dos conjuntos medidos, as distribuições de probabilidade ajustadas aos dados da espessura apresentavam um deslocamento para a direita face ao valor nominal. Foram exceções os Conjuntos E e G, sendo que o Conjunto E tinha tolerância superior igual a 0, o que implicou que o valor máximo da tolerância correspondesse ao valor nominal. Já no Conjunto G, a média situa-se ligeiramente à direita do valor nominal, sendo que a distribuição está mais próxima do valor nominal.



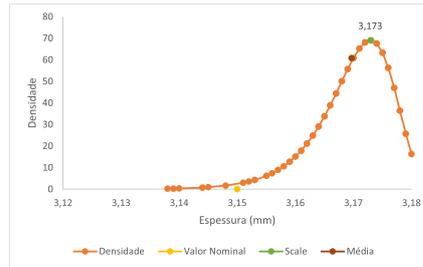
(a) Normal($\mu= 4,033$; $\sigma= 0,011$).



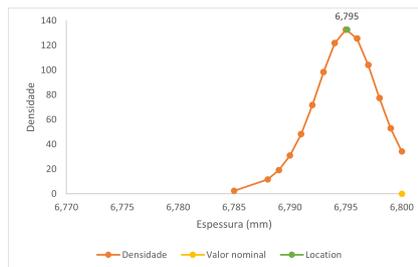
(b) Normal($\mu= 5,027$; $\sigma= 0,013$).



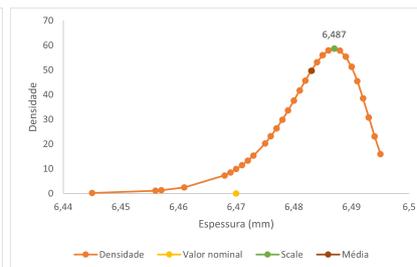
(c) Weibull($\alpha= 527,406$; $\beta= 1,637$).



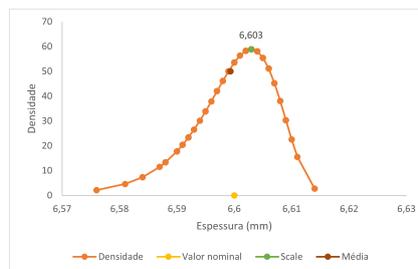
(d) Weibull($\alpha= 594,734$; $\beta= 3,173$).



(e) Logística($\mu= 6,795$; $\sigma= 0,002$).



(f) Weibull($\alpha= 1019,351$; $\beta= 6,487$).



(g) Weibull($\alpha= 1057,507$; $\beta= 6,603$).

Figura 39: Distribuições de probabilidade para a espessura nos Conjuntos A a G.

As lentes com tolerâncias maiores apresentam um conjunto de medições mais próximas da simetria e com um menor achatamento. Nestes casos, a distribuição Normal surgiu como a melhor alternativa.

Por outro lado, a maioria das lentes com tolerâncias menores apresentou assimetria negativa forte, o que impossibilitou o ajuste da distribuição Normal. Para estes conjuntos, a Weibull foi a melhor opção para o ajuste.

No entanto, apesar de apresentar dados muito próximos da simetria, a distribuição Normal não apresentou bons resultados no ajuste aos dados do Conjunto E. Além disso, corresponde à lente com menor amplitude de tolerâncias, uma vez que a sua tolerância superior é nula. Para este conjunto, a distribuição Logística surgiu como a melhor hipótese.

As distribuições de probabilidade ajustadas ao diâmetro e ao erro de centragem estão disponíveis para consulta no Apêndice A.6, nas Figuras 7 e 8, respectivamente. As restantes distribuições de probabilidade para a espessura estão disponíveis no Apêndice A.7, na Figura 9.

Além disso, a Figura 38 sugere que a dureza do vidro (Knoop hardness (HK)) influencia a concentração de valores em torno da média. Gross (2007) revela que quanto menos dureza apresentar o vidro, mais ampla será a distribuição em torno do centro.

Neste caso, a concentração de valores ao redor da média pode ser quantificada pela *kurtosis*, uma métrica estatística que caracteriza o achatamento da distribuição.

Uma vez que Gross (2007) apenas sugere o uso de distribuição Normal e Uniforme para as medições da espessura, a relação entre a dureza e o achatamento poderia não se verificar para distribuições assimétricas. Assim sendo, considerou-se dividir as distribuições em dois grupos, as simétricas (Figura 40) e as assimétricas (Figura 41).

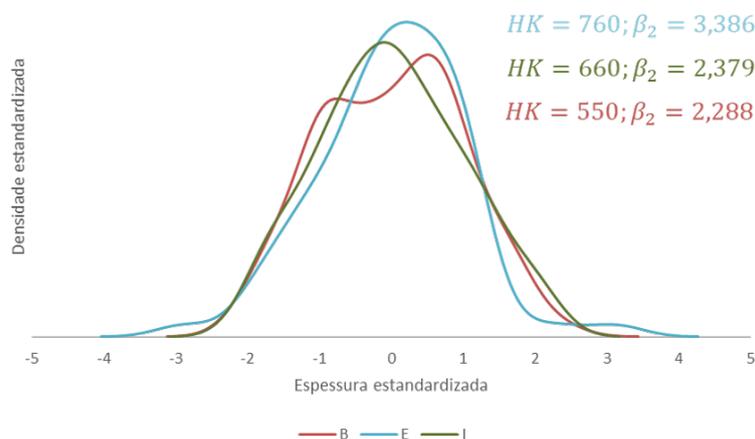


Figura 40: Distribuições simétricas.

No conjunto das distribuições simétricas, a lente com menor coeficiente de dureza (HK) possui uma distribuição com menor coeficiente de *kurtosis*. Nas outras duas lentes o achatamento aumenta quando o coeficiente de dureza é superior. Assim sendo, verifica-se a relação entre dureza e concentração de valores ao redor da média, ilustrada na Figura 40.

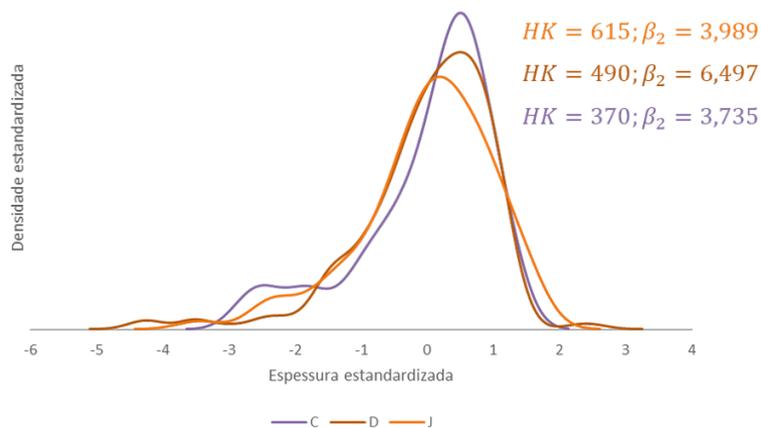


Figura 41: Distribuições assimétricas.

Por outro lado, no conjunto das distribuições assimétricas esta relação não foi evidente. O Conjunto D apresenta menor dureza que o Conjunto J, mas menor concentração de valor em torno da médio, uma vez que possui maior *kurtosis*, tal como sugere a Figura 41.

6.2 Correlações

Nesta secção, o objetivo foi encontrar correlações entre a geometria da lente e as suas tolerâncias com os parâmetros da distribuição ajustada, bem como estatísticas descritivas que caracterizam a forma dos dados. Assim, foram utilizadas as seguintes características da lente, que constam na ficha técnica:

- Amplitude de tolerância - somas das tolerâncias inferior e superior associada ao valor nominal da espessura da lente;
- Dureza (Knoop hardness (HK)) - coeficiente de dureza do vidro da lente;
- Expansão térmica - coeficiente de expansão térmica do vidro da lente;
- n_e - índice de refração do material da lente;
- Nominal - valor nominal da espessura da lente;
- R1 e R2 - raios de curvatura das superfícies da lente;
- ν_e - número de *abbe*, medida que quantifica a dispersão cromática da luz através das lentes.

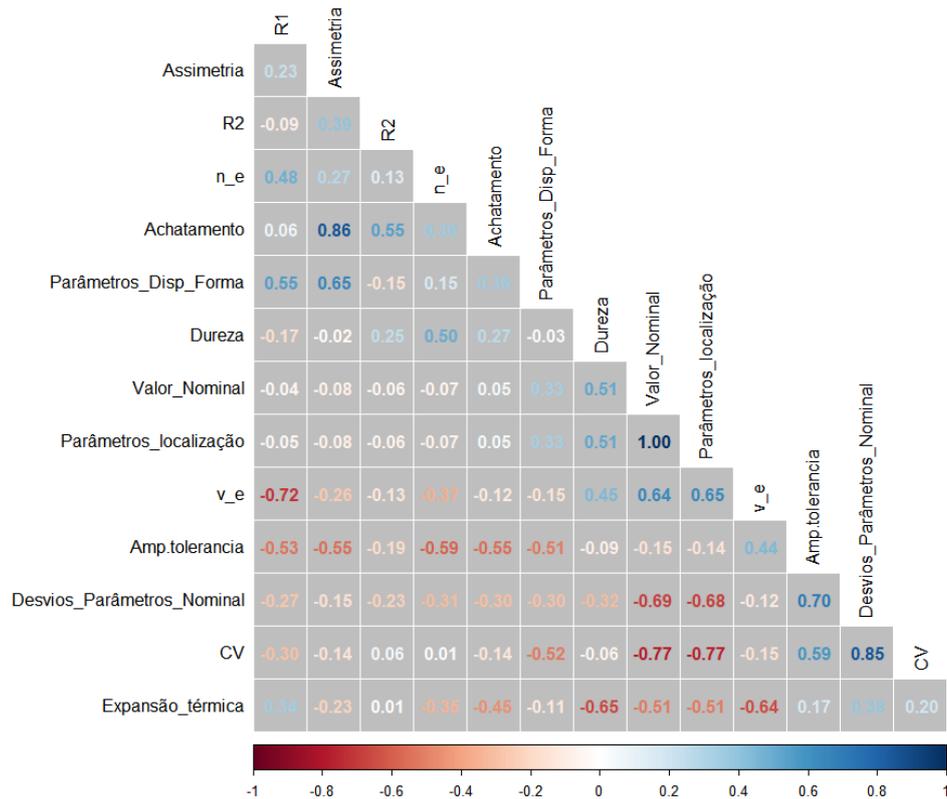


Figura 42: Correlações entre os parâmetros da lente e as características da distribuição dos dados.

Amplitudes de tolerância maiores estão associadas a achatamentos menores, ou seja, a uma maior concentração de valores ao redor da média. Além disso, estão correlacionadas de forma negativa com a assimetria, sugerindo que a distribuição Normal poderá ajustar-se melhor a lentes com tolerância maior. Além disso, tolerâncias maiores estão associadas a maiores desvios relativamente à média e a maiores desvios dos parâmetros de localização relativamente aos valores nominais.

Lentes com valor nominal maior e, conseqüentemente, distribuições com parâmetros de localização superiores, estão associadas a menores variações em torno da média. Adicionalmente, lentes com valores nominais baixos e, conseqüentemente, parâmetros de localização menores, estão associadas a desvios nominais maiores.

Por sua vez, o parâmetro de localização está correlacionado positivamente com o v_e e a dureza. Lentes com maiores parâmetros de localização estão associadas a vidros com maiores coeficientes de dureza e v_e . Por outro lado, este parâmetro está correlacionado de forma negativa com o coeficiente de variação e o desvio do parâmetro ao nominal, traduzindo-se numa associação linear inversa. À semelhança do parâmetro de localização, o valor nominal também se relaciona com o coeficiente de variação e com o desvio dos parâmetros ao nominal.

Vidros com coeficientes de expansão térmica altos estão associados a lentes com coeficiente de variação mais alto e maiores desvios do parâmetro de localização relativamente ao valor nominal.

Em suma, as características da lente que se revelaram mais relevantes para a previsão dos parâmetros das distribuições de probabilidade da espessura foram o valor nominal, a amplitude de tolerância, a dureza e a expansão térmica.

6.3 Modelos para a previsão dos parâmetros

A previsão dos parâmetros das distribuições de probabilidade teve como ponto de partida os dados contidos na Tabela 35. As variáveis de resposta foram os parâmetros das distribuições, enquanto que as variáveis preditoras foram parâmetros da lente que se revelaram mais relevantes nas correlações.

Tabela 35: Base de dados para a previsão dos parâmetros das distribuições de probabilidade ajustadas à espessura.

Lente	Distribuição	Nominal	Amplitude		Expansão	Parâmetros	
			de tolerância	Dureza		térmica	Localização/escala
A	Normal	4,000	0,200	550	9,500	4,033	0,011
B	Normal	5,000	0,200	550	7,000	5,027	0,013
C	Weibull	1,610	0,040	370	8,100	1,637	527,406
D	Weibull	3,150	0,060	500	7,500	3,173	594,492
E	Logística	6,800	0,030	760	6,300	6,795	0,002
F	Weibull	6,470	0,060	750	6,200	6,487	1024,475
G	Weibull	6,600	0,060	710	6,600	6,603	1057,495
H	Weibull	3,300	0,060	750	6,200	3,322	417,729
I	Normal	1,200	0,200	660	7,900	1,264	0,007
J	Weibull	6,200	0,200	615	5,700	6,236	627,843
K	Weibull	4,840	0,030	490	9,000	4,866	1187,948

Os modelos para a previsão dos parâmetros foram separados de acordo com o tipo de distribuição do conjunto. Foram criados dois grupos, um com as lentes A, B, E e I (Normal e Logística) e outro com as restantes lentes, cujas distribuições são Weibull (Tabela 35).

Recorreu-se ainda, a métodos de seleção de variáveis para detetar qual seria o conjunto de variáveis mais importantes para a previsão de cada parâmetro. Tais variáveis seriam incluídas no modelo. Os métodos usados foram o *Forward Stepwise Selection* e o *Backward Stepwise Selection*. Dada a pequena

dimensão da amostra para as distribuições simétricas o método apresentava problemas no cálculo do AIC.

O método de reamostragem que conduziu a um modelo com menor erro de teste foi o método *k-Fold Cross-validation*, com $k = 4$.

A previsão do parâmetros das distribuições Normal e Logística é dada pelas expressões (6.39) e (6.40).

$$\hat{\mu} = 0,061 + 0,990 \times \text{Nominal} + 0,073 \times \text{Amp.tol} \quad (6.39)$$

$$\hat{\sigma} = 0,0381 - 0,00005 \times \text{Dureza} \quad (6.40)$$

A previsão do parâmetros da distribuição Weibull é dada pelas expressões (6.41) e (6.42).

$$\hat{\alpha} = 0,027 + 0,996 \times \text{Nominal} + 0,162 \times \text{Amp.tol} \quad (6.41)$$

$$\hat{\beta} = -919,867 - 1931,204 \times \text{Amp.tol} + 182,828 \times \text{Nominal} + 156,721 \times \text{Exp.term} - 0,162 \times \text{Dureza} \quad (6.42)$$

Através dos modelos estimados, desde (6.39) ao (6.42), efetuaram-se as previsões dos parâmetros para as distribuições e calculou-se o respetivo erro associado (Tabela 36).

Tabela 36: Erros associados à previsão dos parâmetros das distribuições de probabilidade da espessura.

Lentes	Parâmetro de localização		Parâmetro de forma	
	Previsão	Erro	Previsão	Erro
A	4,036	0,002	0,012	0,001
B	5,026	0,002	0,012	0,001
C	1,637	0,000	506,600	31,103
D	3,175	0,002	634,392	36,236
E	6,795	0,000	0,002	0,000
F	6,482	0,005	997,055	31,310
G	6,612	0,009	1090,005	27,862
H	3,324	0,002	417,490	0,239
I	1,264	0,001	0,007	0,000
J	6,236	0,000	620,881	6,962
K	4,858	0,008	1180,076	7,871

Nas distribuições simétricas a média dos erros absolutos para o parâmetro de localização foi de 0,001, enquanto que para as distribuições assimétricas o erro foi de 0,004.

Por outro lado, o erro associado à previsão dos parâmetros de dispersão, foi de 0,004, nas distribuições simétricas, e de 20,018 para o parâmetro de forma, nas distribuições assimétricas.

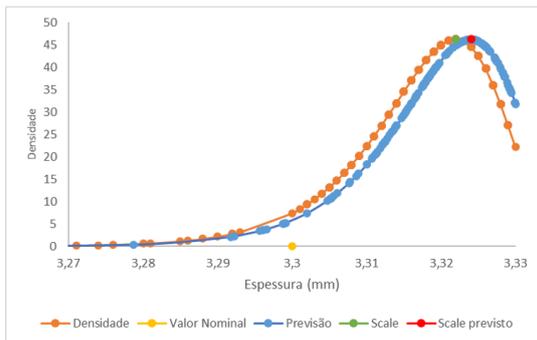
Em relação à previsão dos parâmetros, o modelo construído apresentou um bom desempenho para os parâmetros de localização e de forma. Contudo, a previsão dos parâmetros de escala e dispersão não foi tão eficaz.

Nesta abordagem, a análise de resíduos não verificou os pressupostos exigidos na utilização de um modelo de regressão linear. Este facto pode ser explicado pela pequena dimensão da amostra.

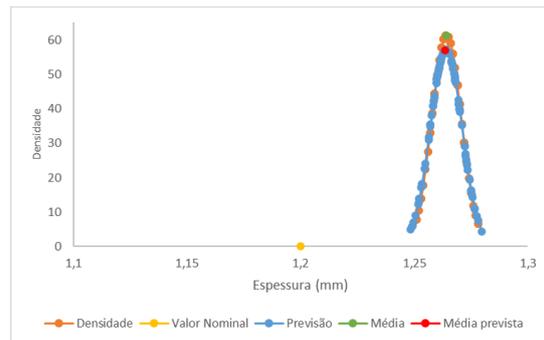
Uma sugestão será a utilização da informação da Tabela 35 como referência e aplicar interpolação linear para a obtenção das estimativas dos parâmetros.

6.4 Previsões

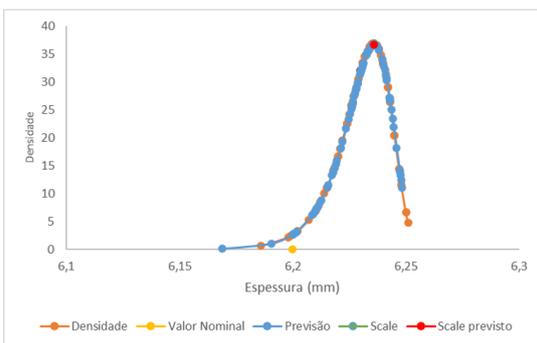
Através dos parâmetros previstos na secção anterior, disponíveis na Tabela 36, efetuaram-se simulações das distribuições, com a mesma dimensão que conjunto original. Na Figura 43 estão representadas, para cada conjunto, as densidades das distribuições de probabilidade ajustadas às medições efetuadas e as densidades das distribuições de probabilidade previstas.



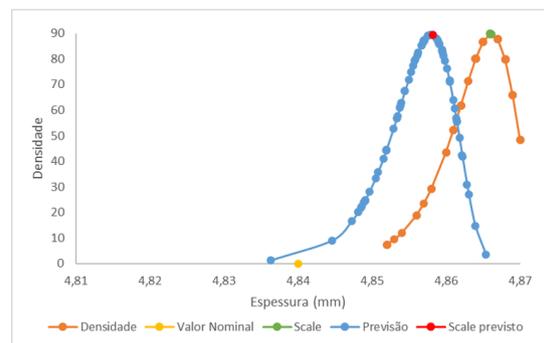
(a) Lente H.



(b) Lente I.



(c) Lente J.



(d) Lente K.

Figura 43: Densidade das distribuições de probabilidade previstas para as lentes H a K.

A Figura 43 sugere que existe um pequeno deslocamento da curva da previsão face à real. Este desvio da densidade prevista é relativo ao erro na estimação do parâmetro de localização. No que diz respeito à forma, esta é mais próxima da real, levando a concluir que os parâmetros de forma e dispersão foram estimados com mais exatidão.

As previsões para a espessura dos restantes conjuntos encontram-se representadas nas Figuras 10 e 11, disponíveis no Apêndice A.8.

6.5 Simulações de Monte Carlo

O método de Monte Carlo implementado no *software* da empresa gera uma grande quantidade de valores a partir de uma distribuição Normal truncada, onde é considerada a tolerância e o valor nominal das lentes com tolerâncias simétricas, valores estes que constam na ficha técnica de uma lente.

Simularam-se 30 000 observações provenientes de uma distribuição Normal truncada com as seguintes características: valor médio igual ao valor nominal da lente, desvio-padrão proporcional à tolerância, e intervalo de valores entre valor nominal \pm tolerância.

A Figura 44 apresenta um histograma para as simulações de Monte Carlo efetuadas para a lente A (Figura 6.44(a)) e a lente D (6.44(b)). Além disso, estão presentes, também, a função de densidade empírica e algumas funções de densidade com o mesmo valor médio, mas diferentes desvios padrão.

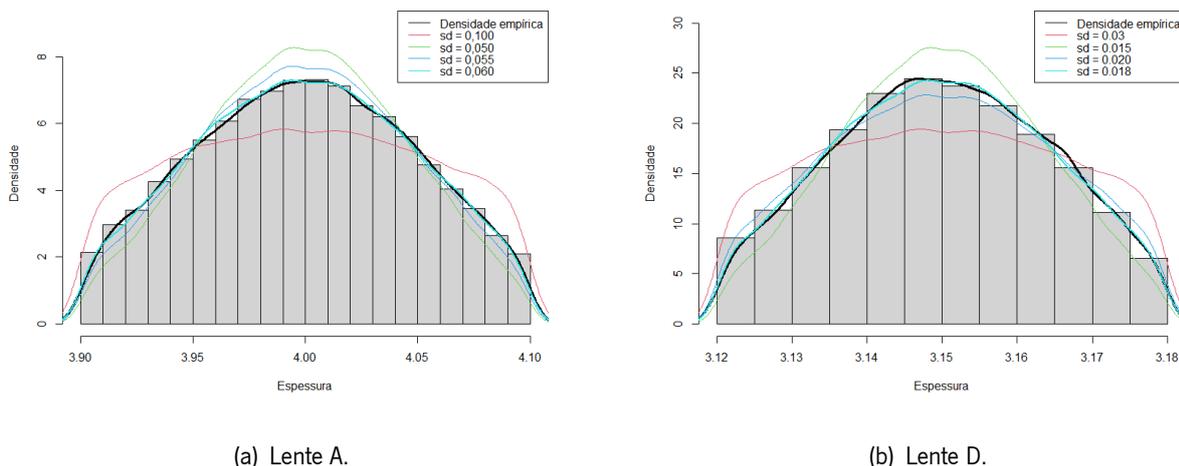


Figura 44: Simulações de Monte Carlo geradas pelo *software* e função de densidade empírica.

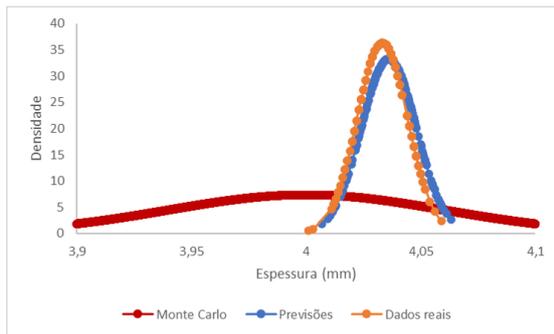
Para a espessura da lente A, com tolerância 0,1, verificou-se que o desvio-padrão da função densidade mais próxima da densidade empírica era de 0,06, como se verifica pela Figura 6.44(a). Por outro lado, para a espessura da lente D, com tolerância 0,03, constatou-se que o desvio-padrão que melhor se ajustou foi o de 0,018, tal como sugere a Figura 6.44(b).

Assim sendo, a razão entre o desvio-padrão e a tolerância de cada tipo de lente foi de 0,6.

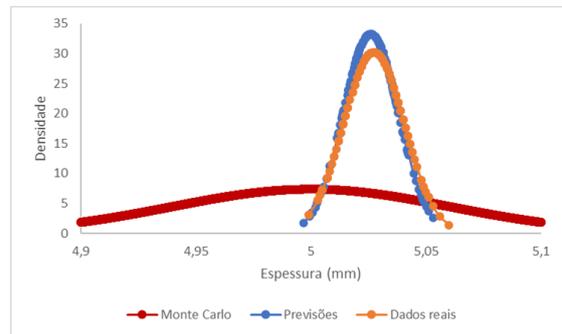
Da análise realizada, concluiu-se que os valores, nominal e tolerância, da ficha técnica são utilizados numa distribuição Normal truncada do seguinte modo: nominal como valor médio e $0,6 \times \text{tolerância}$ como desvio-padrão.

Na Figura 45 estão ilustradas as densidades das simulações Monte Carlo, as densidades previstas

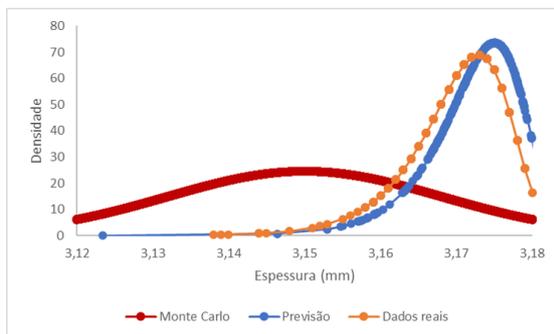
pelos modelos preditivos e as densidades reais.



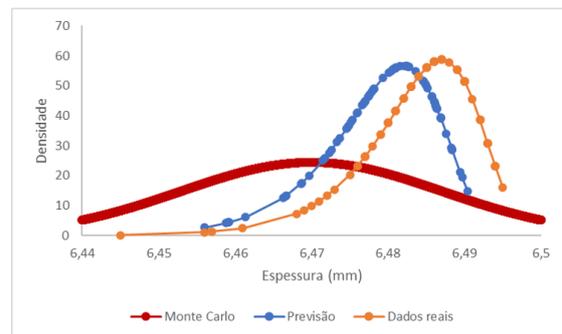
(a) Lente A.



(b) Lente B.



(c) Lente D.



(d) Lente F.

Figura 45: Distribuições Normais ajustadas aos dados das simulações de Monte Carlo vs previsões pelos modelos.

No que diz respeito à forma da curva, a densidade obtida através das simulações Monte Carlo está bastante distante da densidade real, tal como revela a Figura 45. Além disso, também está deslocada, uma vez que as previsões efetuadas por Monte Carlo estão centradas no valor nominal. De um modo geral as distribuições da espessura apresentaram-se à direita do valor nominal.

No entanto, no Conjunto G, a previsão por Monte Carlo surge mais próxima da distribuição real, no que diz respeito ao centro da distribuição (consultar a Figura 12, disponível no Apêndice A.9).

Nas Tabelas 37 à 39 é possível conferir os erros entre as previsões através dos modelos e as previsões através das simulações de Monte Carlo, tendo como referência os valores reais (medições efetuadas).

Tabela 37: Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto A.

Métricas	Previsão	Simulação Monte Carlo
MSE	$8,300 \times 10^{-6}$	0,002
MAE	0,003	0,041
MAPE	0,063%	1,015%
RMSE	0,003	0,050

A Tabela 37 revela que o erro absoluto médio das previsões dos modelo foi de aproximadamente 0,003 enquanto que o das simulações Monte Carlo foi de 0,041. Face às simulações Monte Carlo o erro das previsões dos modelos diminuiu cerca de 94%.

Tabela 38: Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto D.

Métricas	Previsão	Simulação Monte Carlo
MSE	$7,050 \times 10^{-6}$	$4,54 \times 10^{-6}$
MAE	0,002	0,020
MAPE	0,074%	0,622%
RMSE	0,003	0,021

No que diz respeito ao Conjunto D, a 38 sugere que a erro diminuiu, novamente, das simulações Monte Carlo para os modelos. Em termos percentuais a redução foi de 88%.

Tabela 39: Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto G.

Métricas	Previsão	Simulação Monte Carlo
MSE	$8,4109 \times 10^{-6}$	0,0001
MAE	0,0091	0,0062
MAPE	0,1378%	0,0943%
RMSE	0,0092	0,0075

Já no Conjunto G, as Simulações de Monte Carlo apresentam um erro absoluto médio menor, como sugere a Tabela 39. Desta vez, as previsões efetuadas apresentam um aumento de 46% face às simulações Monte Carlo.

No restantes conjuntos as previsões dos modelos preditivos foram mais eficazes e, em termos absolutos, foram sempre menores (consultar Apêndice A.10, Tabelas 21 à 26).

Em média, o erro absoluto médio das previsões dos modelos diminuiu 71% face às simulações Monte Carlo.

7 Conclusões

As simulações de Monte Carlo são realizadas durante o desenvolvimento ótico de novos produtos com a finalidade de estimar a variação do desempenho do produto, tendo em conta as tolerâncias e compensações associadas à produção em série. Deste modo, quanto mais próximas forem as distribuições de probabilidade associadas às tolerâncias de produção, mais eficaz será a simulação do desempenho do produto.

Assim sendo, a motivação para este projeto passou por estudar e ajustar distribuições de probabilidade a características de várias lentes e desenvolver modelos para a sua previsão na produção em série.

Para isso, inicialmente foi medido o diâmetro, o erro de centragem e a espessura de dois conjuntos de lentes diferentes e calculada a incerteza das medições do erro de centragem e da espessura. Depois, foram ajustadas distribuições de probabilidade às respetivas medições das características da lente nos dois conjuntos. Em seguida, foi do interesse da empresa, focar o estudo na característica espessura de modo a estudar a influência da tolerância deste parâmetro na escolha e previsão das distribuições de probabilidade. Deste modo, realizaram-se novas medições da espessura da lente em nove conjuntos diferentes.

Para as medições do diâmetro do Conjunto A a distribuição Logística pelo método dos momentos foi a que melhor se ajustou. Já no Conjunto B, a distribuição com melhor ajuste foi a Log-logística pelo método de máxima verosimilhança.

No que diz respeito ao erro de centragem, para ambos os conjuntos a distribuição com melhor ajuste foi a distribuição Beta. No Conjunto A, o método cujo desempenho foi melhor foi o *maximum goodness-of-fit*, com a distância *Left-tail Anderson Darling (ADL)*. Para o Conjunto B, o método que proporcionou o melhor ajustamento foi o método da correspondência dos quantis. Para esta característica da lente, a incerteza de medição, para um nível de confiança de 99,7%, foi de 0,039 minutos.

Relativamente às medições da espessura, as distribuições ajustadas aos Conjuntos A e B, de tolerância $\pm 0,1$, foram a Normal pelo método de máxima verosimilhança. Para um nível de confiança de 99,7%, a incerteza associada à medição da espessura foi de 0,003 milímetros.

No entanto, suspeitou-se que, para lentes com tolerância reduzidas, a distribuição Normal não seria a mais adequada. Assim sendo, efetuaram-se novas medições, quer para conjuntos com tolerâncias menores, para verificar se existiam diferenças na distribuição, quer para conjuntos com a mesma tolerância, para averiguar se o ajustamento da distribuição Normal se confirmava para lentes com tolerância $\pm 0,1$ para a espessura.

Efetivamente, nas lentes com tolerâncias reduzidas ($\pm 0,03$) as distribuições Weibull foram as que melhor se ajustaram. Por outro lado, para uma das lentes com tolerância $\pm 0,1$ a distribuição Normal foi a melhor, enquanto que na outra lente a Weibull foi a mais próxima da distribuição empírica dos dados. Além desta exceção, existiu uma outra, a Lente E, cuja tolerância inferior para a espessura era 0 e a superior era $-0,03$. Para este conjunto, a distribuição de probabilidade que melhor se ajustou foi a Logística pelo método dos momentos.

Assim sendo, concluiu-se que a distribuição de probabilidade associada às medições da espessura varia consoante a tolerância. No geral, as lentes com maior tolerância para a espessura seguiram uma distribuição Normal, enquanto que lentes com tolerâncias mais reduzidas seguiram uma distribuição Weibull.

Além disso, verificaram-se algumas tendências associadas à distribuição das medições da espessura. A maioria dos conjuntos apresentou a maior parte da distribuição à direita do valor nominal, tornando as medições da espessura mais próximas do limite superior de tolerância. As lentes são produzidas com espessura acima do valor nominal para que depois, caso necessário, sejam retrabalhadas e o valor da espessura seja menor que o nominal.

A relação entre o coeficiente de dureza e a concentração de valores ao redor da média apenas se verifica para distribuições com a forma semelhante às que a literatura sugere. Deste modo, verificou-se que, para as lentes cuja distribuição é Normal (ou simétrica), a concentração de valores de espessura em torno da média é maior para as lentes cujo vidro tem menor coeficiente de dureza. Para as restantes lentes, a relação entre a dureza e a *kurtosis* não foi conclusiva. Além disso, lentes com valor nominal mais baixo estão associadas a maior variação relativa dos valores da espessura em torno da média e a maiores desvios face ao centro da distribuição.

No que diz respeito às previsões, observou-se que a maioria das distribuições de probabilidade estava próxima das medições reais. Os parâmetros σ e β foram previsto com mais exatidão, enquanto que o μ e o α por vezes encontraram-se desviados do valor real, translacionando o centro da distribuição.

As simulações de Monte Carlo implementadas no *software* geraram dados para espessura das lentes a partir de uma distribuição Normal truncada, sendo o valor médio o valor nominal, desvio padrão como $0,60 \times \text{tolerância}$ e intervalo de valores compreendido entre $\text{nominal} \pm \text{tolerância}$. Esta metodologia foi a mesma para todas as lentes com tolerâncias simétricas.

Em suma, concluiu-se que as previsões através das simulações de Monte Carlo conduziam a uma falsa percepção dos valores da espessura face ao intervalo de tolerância. As previsões efetuadas através dos modelos desenvolvidos foram mais eficientes, realistas e apresentaram uma redução no erro absoluto

médio em 71% face às simulações de Monte Carlo.

7.1 Trabalho futuro

Face às dificuldades encontradas durante a elaboração deste projeto e a escrita desta dissertação, surgiram algumas questões que podem levar a desenvolvimentos futuros do trabalho realizado no âmbito deste estágio.

No que diz respeito à incerteza de medição, seria interessante incluir medições dos parâmetros das lentes feitas por diferentes medidores. Isto seria útil para identificar a variação das medições por medidor e entre medidores, sendo mais perceptível o erro humano associado à incerteza das medições.

Além disso, seria importante aumentar a quantidade de informação para as previsões, de forma a que estas fossem feitas com mais exatidão. A recolha de mais medições de outros tipos de lentes seria útil para otimizar as previsões do parâmetro de localização das distribuições de probabilidade.

Por fim, seria interessante ajustar distribuições de probabilidade e desenvolver modelos para a previsão dos valores dos restantes parâmetros da lente, que constam na ficha técnica, e para outros processos como, por exemplo, as peças mecânicas com tolerâncias associadas, pois também influenciam a *performance* do sistema ótico.

Referências Bibliográficas

- Bernd, D., Henriette, M., Gross, H., et al. (2012). *Handbook of optical systems, volume 5: Metrology of optical components and systems*. John Wiley & Sons.
- Bjorn Petersen. (2016). *Optical Anomalies and Lens Corrections Explained*. bhphotovideo.com/explora/photography/tips-and-solutions/optical-anomalies-and-lens-corrections-explained. (Acedido a 23-08-2023)
- Bruce MacEvoy. (2013). *Astronomical Optics*. handprint.com/ASTRO/ae4.html. (Acedido a 6-10-2023)
- Cullen, A. C., Frey, H. C., & Frey, C. H. (1999). *Probabilistic techniques in exposure assessment: a handbook for dealing with variability and uncertainty in models and inputs*. Springer Science & Business.
- D'Agostino, R. (2017). *Goodness-of-fit-techniques*. Routledge.
- Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: An r package for fitting distributions. *Journal of statistical software*, 64, 1–34.
- Eckhardt Optics. (2023). *Resolution Testing Methods*. eckop.com/resources/optical-testing/resolution-testing-methods/. (Acedido a 6-10-2023)
- Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.
- Greg Hollows, Nicholas James. (2023). *The Airy Disk and Diffraction Limit*. edmundoptics.com/knowledge-center/application-notes/imaging/limitations-on-resolution-and-contrast-the-airy-disk/. (Acedido a 15-10-2023)
- Gross, H. (2005). *Handbook of optical systems, volume 1, fundamentals of technical optics* (Vol. 1).
- Gross, H. (2007). *Handbook of optical systems, volume 3: Aberration theory and correction of optical systems* (Vol. 3). Wiley-Vch.
- Hasenauer, D. (2015). *Optical design tolerancing a key to product cost reduction*. Synopsys.
- James, G., Witten, D., Hastie, T., Tibshirani, R., et al. (2013). *An introduction to statistical learning* (Vol. 112). Springer.

- Kellner, C. (1849). \rightarrow *das orthoskopische okular, eine neu erfundene achromatische linsenkombination*. Vieweg.
- Koch, D. G. (1978). A statistical approach to lens tolerancing. In *Computer-aided optical design* (Vol. 147, pp. 71–82).
- Leica Camera Wetzlar. (2023). www.leica-camera.com/en-int. (Acedido a 14-06-2023)
- Leica Portugal. (2023). www.leica.pt. (Acedido a 14-06-2023)
- Luceño, A. (2006). Fitting the generalized pareto distribution to data using maximum goodness-of-fit estimators. *Computational Statistics & Data Analysis*, 51(2), 904–917.
- Matt Osborne. (2019). *Leica Thambar 90mm Review (f2.2 lens) + Portrait Shoot*. mrleica.com/leica-thambar-90mm-review/. (Acedido a 15-10-2023)
- Nasim Mansurov. (2019). *What is Coma?* photographylife.com/what-is-coma. (Acedido a 15-10-2023)
- NDC Technologies. (2015). *Z-MIKE Benchtop systems*. msiviking.com/documents/zmike_precisiondimensionalmeasurementsolutions.pdf. (Acedido a 14-03-2023)
- Quality assurance of precision optics*. (2013). University of Central Florida. Retrieved from <https://books.google.pt/books?id=jiXrnQEACAAJ>
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591–611.
- Smith, D. J., & Rimmer, D. (1999). Manufacture of optical components under statistical process control. In *Optical fabrication and testing* (Vol. 3739, pp. 115–122).
- Tienvieri, C. T., & Rich, T. (1998). Modern lens design using a lens manufacturing database. In *International optical design conference 1998* (Vol. 3482, pp. 508–518).
- Trioptics. (2023). *High-precision centration test devices, alignment, cementing and bonding systems*. trioptics.com/products/opticentric-lens-centration-and-alignment/. (Acedido a 16-03-2023)
- Tse, Y.-K. (2009). *Nonlife actuarial models: theory, methods and evaluation*. Cambridge University Press.

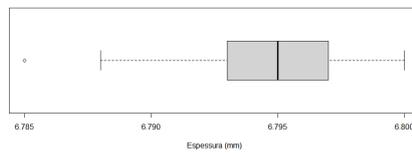
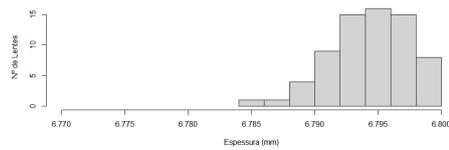
Youngworth, R. N., & Stone, B. D. (2000). Cost-based tolerancing of optical systems. *Applied Optics*, 39(25), 4501–4512.

Yu Kikuchi. (2023). *What Is the Modulation Transfer Function?* olympus-lifescience.com/en/resources/white-papers/what-is-the-modulation-transfer-function/. (Acedido a 23-04-2023)

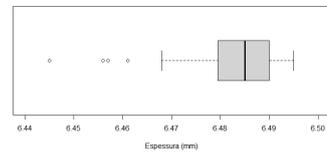
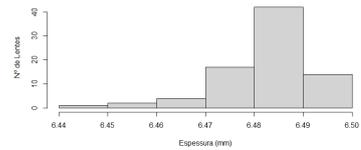
Zhang, S., Li, C., & Li, S. (2017). *Understanding optical systems through theory and case studies*.

Apêndice

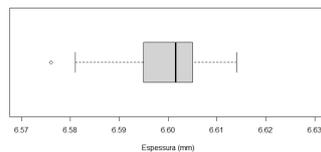
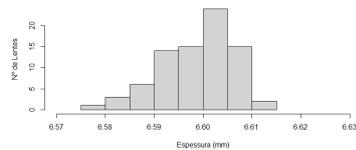
A.1 Histogramas e boxplots para as medições da espessura nos Conjuntos E a F



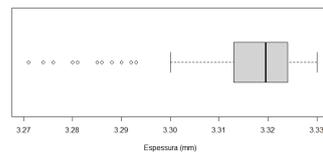
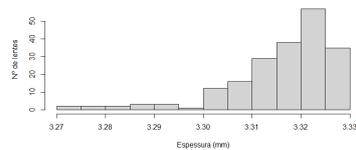
(a) Conjunto E.



(b) Conjunto F.

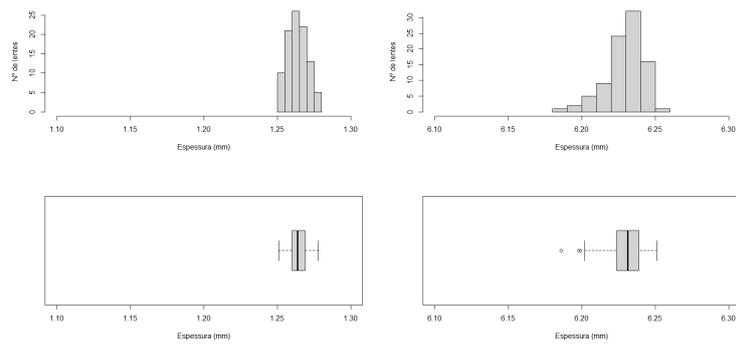


(c) Conjunto G.



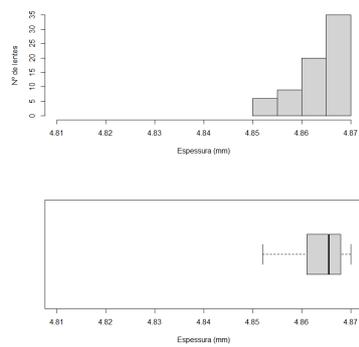
(d) Conjunto H.

Figura 1: Representação gráfica das medições da espessura do Conjunto E ao H.



(a) Conjunto I.

(b) Conjunto J.



(c) Conjunto K.

Figura 2: Representação gráfica das medições da espessura do Conjunto I ao J.

A.2 Distribuições de probabilidade com melhor desempenho nos métodos de estimação pontual

Tabela 1: Distribuições que apresentaram melhor ajustamento para o diâmetro no Conjunto A.

Método	Distribuição	Log-verossimilhança	AIC	BIC	KS	CvM	AD
MMV	Log-logística	1019,710	-2035,410	-2028,720	0,060	0,130	1,040
MGOF	Gama ADL	1013,260	-2022,530	-2015,830	0,060	0,140	1,300
MM	Logística	1020,480	-2036,960	-2030,270	0,060	0,510	0,930
MCQ	Gama (0,45; 0,65)	1013,000	-2021,990	-2015,300	0,110	0,510	2,710

Tabela 2: Distribuições que apresentaram melhor ajustamento para o erro de centragem no Conjunto B.

Método	Distribuição	Log-verossimilhança	AIC	BIC	KS	CvM	AD
MMV	Beta	204,658	-405,315	-399,294	0,053	0,036	0,272
MGOF	Beta ADR	204,311	-404,622	-398,601	0,049	0,023	0,211
MM	Beta	204,657	-405,314	-399,293	0,069	0,125	0,807
MCQ	Beta (0,30; 0,80)	204,601	-405,202	-399,181	0,048	0,029	0,230

Tabela 3: Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto A.

Método	Distribuição	Log-verossimilhança	AIC	BIC	KS	CvM	AD
MMV	Normal	648,670	-1293,340	-1286,650	0,060	0,130	0,830
MGOF	Normal AD2	648,530	-1293,060	-1286,360	0,080	0,110	1,040
MM	Normal	648,670	-1293,340	-1286,650	0,060	0,190	0,830
MCQ	Normal (0,20; 0,70)	648,670	-1293,340	-1286,640	0,060	0,130	0,830

Tabela 4: Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto D.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Weibull	1085,977	-2167,953	-2160,546	0,059	0,169	1,312
MGOF	Weibull ADL	1085,041	-2166,081	-2158,674	0,047	0,097	0,892
MM	Weibull	1083,779	-2163,558	-2156,150	0,081	0,312	2,548
MCQ	Weibull (0,10; 0,70)	1085,127	-2166,254	-2158,846	0,051	0,107	0,991

Tabela 5: Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto F.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Weibull	274,928	-545,856	-541,092	0,084	0,054	0,458
MGOF	Weibull ADR	274,612	-545,225	-540,461	0,066	0,048	0,419
MM	Weibull	272,157	-540,314	-535,550	0,090	0,108	0,897
MCQ	Weibull (0,30 0,95)	274,892	-545,784	-541,020	0,082	0,051	0,466

Tabela 6: Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto H.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Weibull	662,558	-1321,116	-1314,519	0,106	0,289	2,058
MGOF	Weibull ADL	658,557	-1313,115	-1306,518	0,090	0,220	2,215
MM	Weibull	651,283	-1298,566	-1291,970	0,115	0,502	3,849
MCQ	Weibull (0,30; 0,90)	662,533	-1321,065	-1314,469	0,105	0,284	2,077

Tabela 7: Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto I.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Normal	350,230	-696,460	-691,310	0,063	0,044	0,321
MGOF	Normal ADR	349,853	-695,706	-690,556	0,056	0,034	0,277
MM	Normal	350,230	-696,460	-691,310	0,063	0,044	0,321
MCQ	Normal (0,60; 0,95)	350,215	-696,431	-691,281	0,066	0,045	0,310

Tabela 8: Distribuições que apresentaram melhor ajustamento para a espessura no Conjunto K.

Método	Distribuição	Log-verosimilhança	AIC	BIC	KS	CvM	AD
MMV	Weibull	283,253	-562,506	-558,009	0,136	0,176	1,171
MGOF	Weibull ADL	280,907	-557,814	-553,317	0,113	0,161	1,195
MM	Weibull	281,443	-558,885	-554,388	0,116	0,165	1,111
MCQ	Weibull (0,65; 0,85)	283,164	-562,328	-557,831	0,150	0,225	1,402

A.3 Desfasamentos entre quantis teóricos e empíricos

Tabela 9: Desfasamento entre os quantis teóricos e os empíricos para o diâmetro no conjunto B.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Log-logística	0,0009	$3,2800 \times 10^{-8}$	0,0001	0,0008	0,0001
MGOF	Gama AD2	0.0005	$2,0400 \times 10^{-8}$	0,0001	0,0008	0,0002
MM	Gama	0,0005	$1,8200 \times 10^{-8}$	0,0001	0,0007	0,0001
MCQ	Gama (0,20; 0,90)	0,0005	$1,8500 \times 10^{-8}$	0,0001	0,0008	0,0001

Tabela 10: Desfasamento entre os quantis teóricos e os empíricos para o erro de centragem no conjunto B.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Beta	0,016	$1,690 \times 10^{-5}$	0,003	0,510	0,004
MGOF	Beta ADR	0.024	$2,300 \times 10^{-5}$	0,003	0,510	0,005
MM	Beta	0,016	$1,680 \times 10^{-5}$	0,003	0,510	0,004
MCQ	Beta (0,30; 0,80)	0,023	$2,310 \times 10^{-5}$	0,003	0,560	0,005

Tabela 11: Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto B.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Normal	0,005	$1,770 \times 10^{-6}$	0,001	0,020	0,001
MGOF	Normal ADL	0,007	$2,610 \times 10^{-6}$	0,001	0,023	0,002
MM	Normal	0,005	$1,770 \times 10^{-6}$	0,001	0,020	0,001
MCQ	Normal (0,05; 0,85)	0,004	$1,770 \times 10^{-6}$	0,001	0,020	0,001

Tabela 12: Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto C.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Weibull	0,0026	$5,4900 \times 10^{-7}$	0,0005	0,0310	0,0007
MGOF	Weibull ADL	0,0018	$4,4000 \times 10^{-7}$	0,0005	0,0320	0,0007
MM	Weibull	0,0021	$4,6000 \times 10^{-7}$	0,0006	0,0340	0,0007
MCQ	Weibull (0,30; 0,40)	0,0035	$4,0500 \times 10^{-7}$	0,0015	0,0910	0,0020

Tabela 13: Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto D.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Weibull	0,0030	$6,0400 \times 10^{-7}$	0,0006	0,0180	0,0008
MGOF	Weibull ADL	0,0045	$5,8500 \times 10^{-7}$	0,0004	0,0130	0,0008
MM	Weibull	0,0027	$1,0500 \times 10^{-6}$	0,0008	0,0250	0,0010
MCQ	Weibull (0,10; 0,70)	0,0032	$5,8000 \times 10^{-7}$	0,0005	0,0170	0,0008

Tabela 14: Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto E.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Log-logística	0,002	$3,170 \times 10^{-7}$	0,001	0,006	0,001
MGOF	Logística AD2L	0,003	$4,060 \times 10^{-7}$	0,001	0,006	0,001
MM	Logística	0,002	$2,930 \times 10^{-7}$	0,001	0,006	0,001
MCQ	Logística (0,25; 0,80)	0,003	$4,020 \times 10^{-7}$	0,001	0,006	0,001

Tabela 15: Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto F.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Weibull	0,0063	$1,4600 \times 10^{-6}$	0,0007	0,0100	0,0012
MGOF	Weibull ADR	0,0050	$1,1500 \times 10^{-6}$	0,0007	0,0110	0,0011
MM	Weibull	0,0053	$2,3900 \times 10^{-6}$	0,0012	0,0190	0,0015
MCQ	Weibull (0,30; 0,95)	0,0063	$1,5000 \times 10^{-6}$	0,0007	0,0110	0,0012

Tabela 16: Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto G.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Weibull	0,0053	$9,4300 \times 10^{-7}$	0,0006	0,0090	0,0010
MGOF	Weibull ADL	0,0060	$1,1800 \times 10^{-6}$	0,0006	0,0100	0,0011
MM	Weibull	0,0057	$1,1000 \times 10^{-6}$	0,0007	0,0100	0,0010
MCQ	Weibull (0,70; 0,90)	0,0054	$9,8500 \times 10^{-7}$	0,0006	0,0090	0,0010

Tabela 17: Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto H.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Weibull	0,013	$6,820 \times 10^{-6}$	0,001	0,040	0,003
MGOF	Weibull ADL	0,010	$4,820E \times 10^{-6}$	0,001	0,035	0,002
MM	Weibull	0,007	$5,110 \times 10^{-6}$	0,002	0,056	0,002
MCQ	Weibull (0,30; 0,90)	0,013	$7,240 \times 10^{-6}$	0,001	0,040	0,003

Tabela 18: Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto I.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Normal	0,0070	$1,0300 \times 10^{-6}$	0,0006	0,0100	0,0010
MGOF	Normal ADR	0,0070	$9,4600 \times 10^{-7}$	0,0006	0,0100	0,0010
MM	Normal	0,0090	$2,0200 \times 10^{-6}$	0,0010	0,0160	0,0014
MCQ	Normal (0,60; 0,95)	0,0070	$1,1500 \times 10^{-6}$	0,0007	0,0110	0,0011

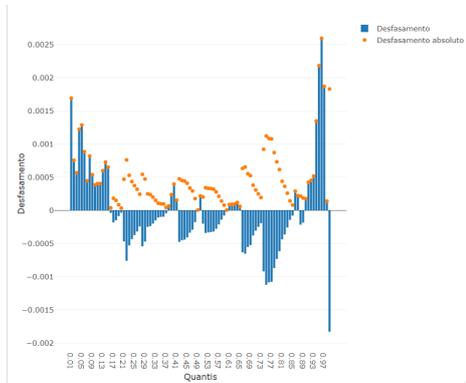
Tabela 19: Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto J.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Weibull	0,0020	$6,0500 \times 10^{-7}$	0,0007	0,0110	0,0008
MGOF	Weibull ADL	0,0020	$6,1200 \times 10^{-7}$	0,0006	0,0100	0,0008
MM	Weibull	0,0020	$9,9000 \times 10^{-7}$	0,0009	0,0140	0,0010
MCQ	Weibull (0,35; 0,75)	0,0010	$6,0400 \times 10^{-7}$	0,0007	0,0110	0,0008

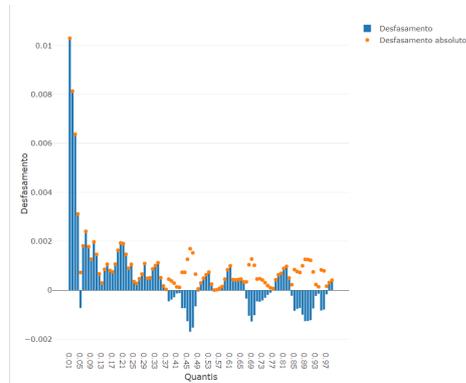
Tabela 20: Desfasamento entre os quantis teóricos e os empíricos para a espessura no conjunto K.

Método	Distribuição	DMA	EQM	EAM	EPAM(%)	REQM
MMV	Weibull	0,0030	$1,0600 \times 10^{-6}$	0,0007	0,0150	0,0010
MGOF	Weibull ADL	0,0050	$9,7000 \times 10^{-7}$	0,0007	0,0150	0,0010
MM	Weibull	0,0030	$1,0600 \times 10^{-6}$	0,0007	0,0150	0,0010
MCQ	Weibull (0,65; 0,85)	0,0030	$1,110 \times 10^{-6}$	0,0008	0,0160	0,0011

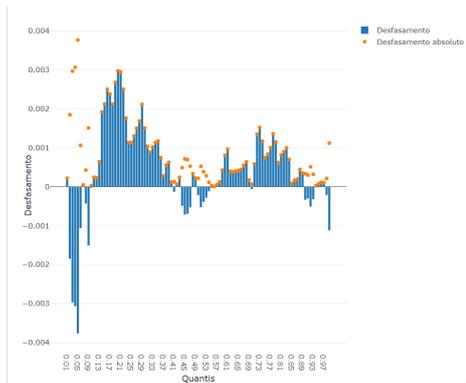
A.4 Desfasamentos entre os quantis teóricos e empíricos para as distribuições obtidas por bootstrap



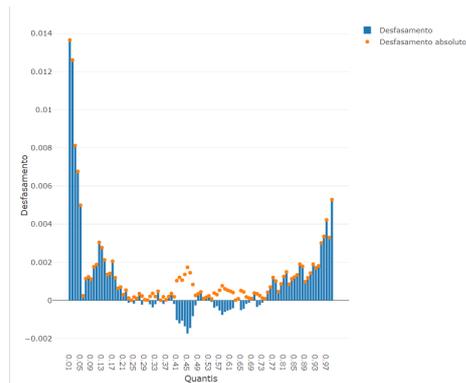
(a) Desfasamentos no conjunto E.



(b) Desfasamentos no conjunto F.

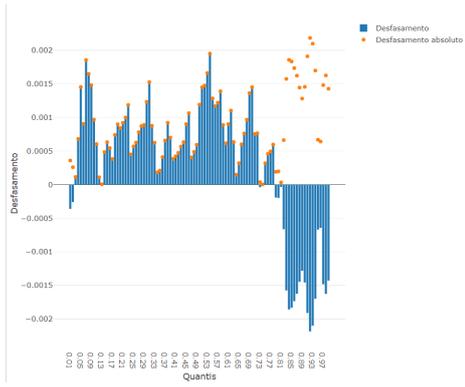


(c) Desfasamentos no conjunto G.

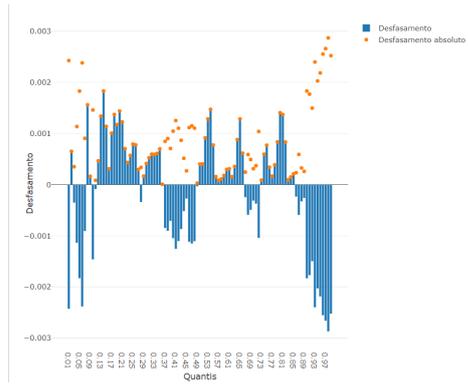


(d) Desfasamentos no conjunto H.

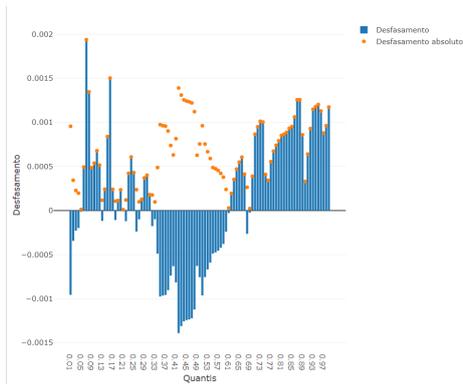
Figura 3: Desfasamentos entre quantis teóricos e amostrais para a espessura dos Conjuntos E a H.



(a) Desfasamentos no conjunto I.



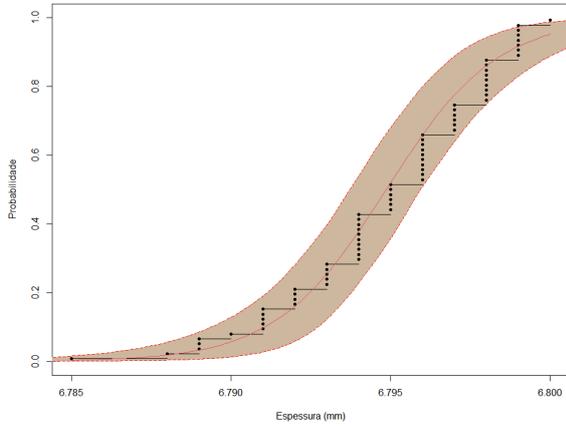
(b) Desfasamentos no conjunto J.



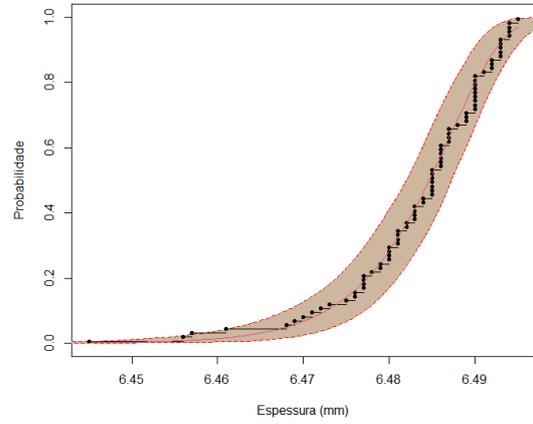
(c) Desfasamentos no conjunto K.

Figura 4: Desfasamentos entre quantis teóricos e amostrais para a espessura dos Conjuntos I a K.

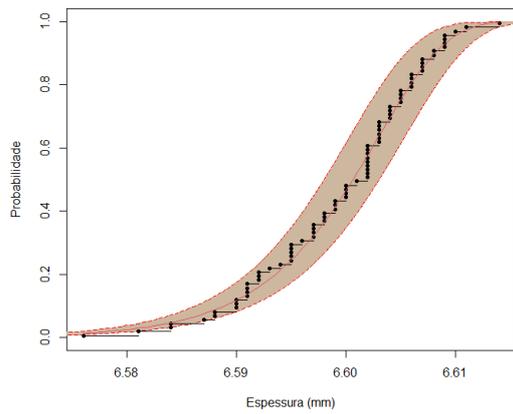
A.5 Qualidade do ajuste



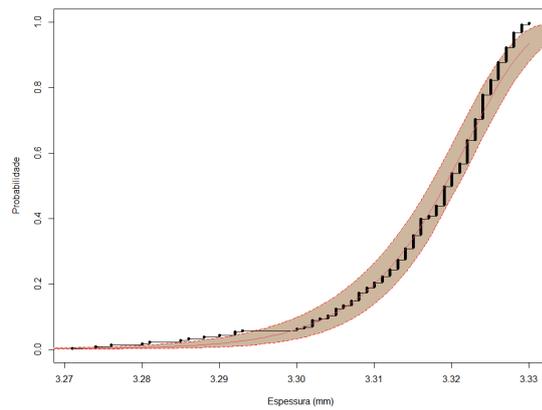
(a) Logística($\mu = 6,795$; $\sigma = 0,002$).



(b) Weibull($\alpha = 6,487$; $\beta = 1019,351$).

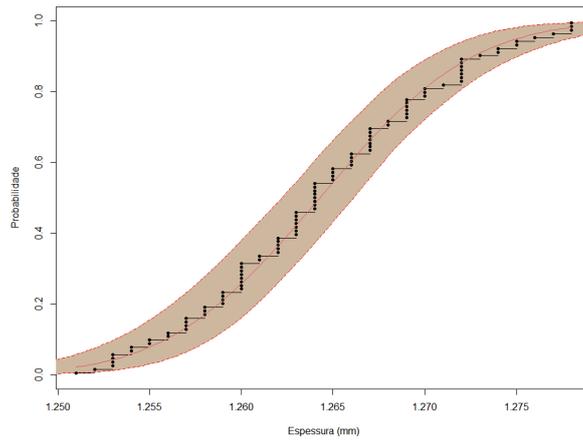


(c) Weibull($\alpha = 6,603$; $\beta = 1057,507$).

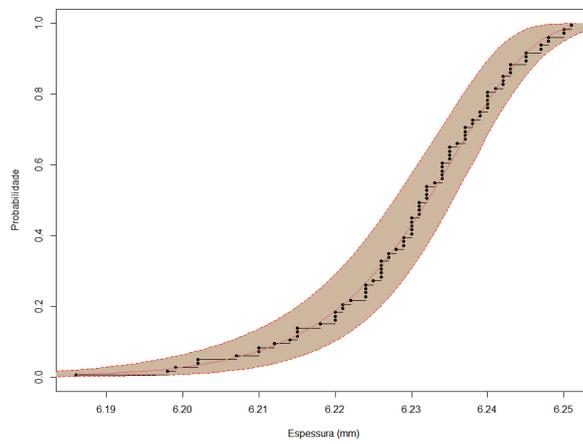


(d) Weibull($\alpha = 3,322$; $\beta = 417,519$).

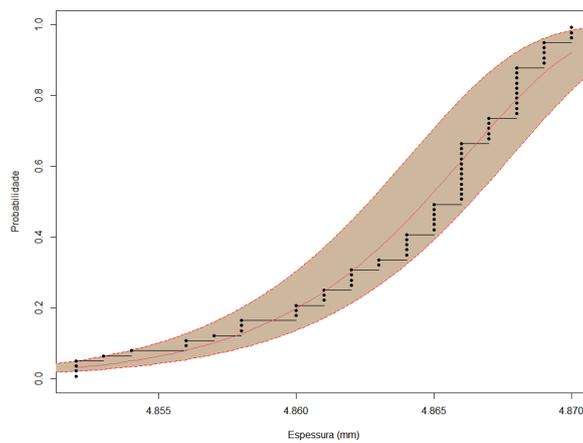
Figura 5: Qualidade de ajuste das distribuições na espessura dos Conjuntos E a H.



(a) Normal($\mu = 1,264$; $\sigma = 0,007$).



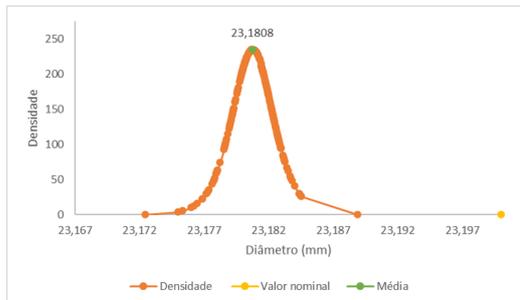
(b) Weibull($\alpha = 6,236$; $\beta = 622,843$)



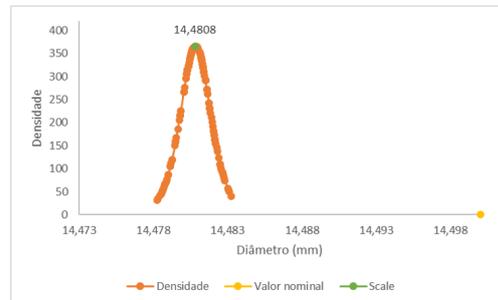
(c) Weibull($\alpha = 4,866$; $\beta = 1187,936$).

Figura 6: Qualidade de ajuste das distribuições na espessura dos Conjunto I a J.

A.6 Distribuições de probabilidade ajustadas ao Diâmetro e Erro de centragem

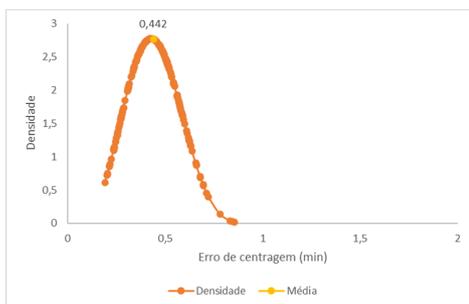


(a) Logística($\mu= 23,181$; $\sigma= 0,001$).

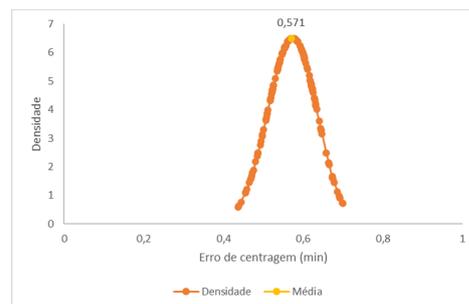


(b) Log-logística($\mu= 21186,139$; $\sigma= 14,481$).

Figura 7: Distribuições de probabilidade para o diâmetro.



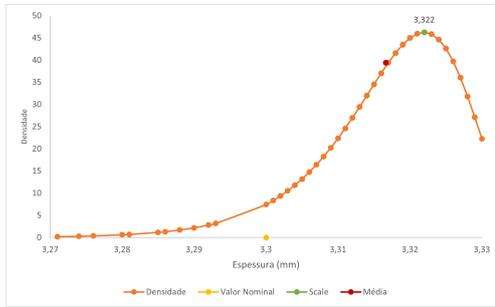
(a) Beta($\alpha= 5,435$; $\beta= 6,906$).



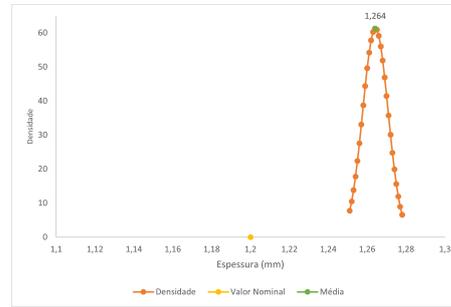
(b) Beta($\alpha= 37,412$; $\beta= 28,042$).

Figura 8: Distribuições de probabilidade para o erro de centragem.

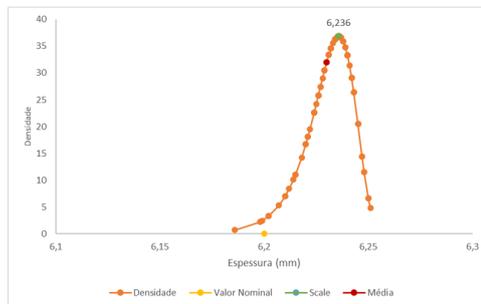
A.7 Distribuições de probabilidade ajustadas à Espessura



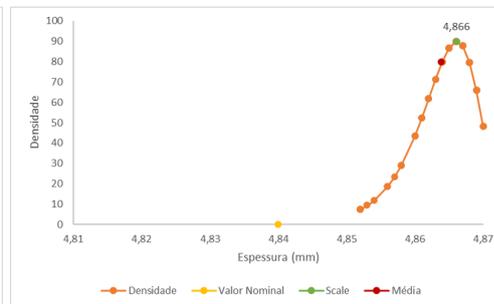
(a) Weibull($\alpha = 417,519$; $\beta = 3,322$).



(b) Normal($\mu = 1,264$; $\sigma = 0,007$).



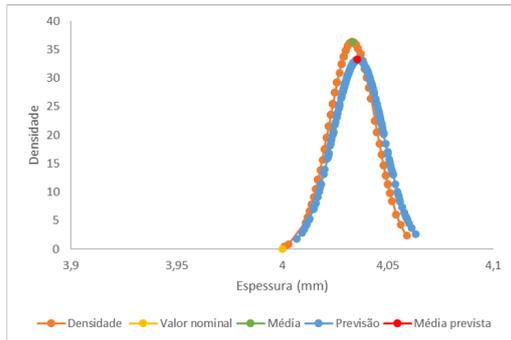
(c) Weibull($\alpha = 627,843$; $\beta = 6,236$).



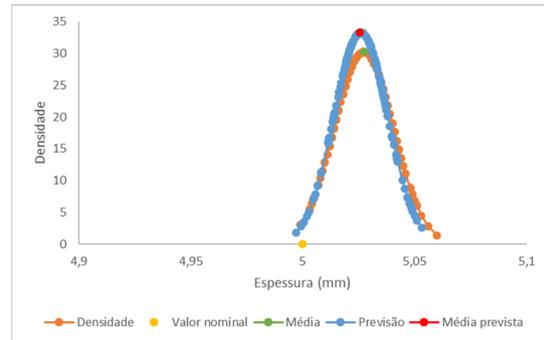
(d) Weibull($\alpha = 1187,936$; $\beta = 4,866$).

Figura 9: Distribuições de probabilidade para a espessura dos Conjuntos de H a K.

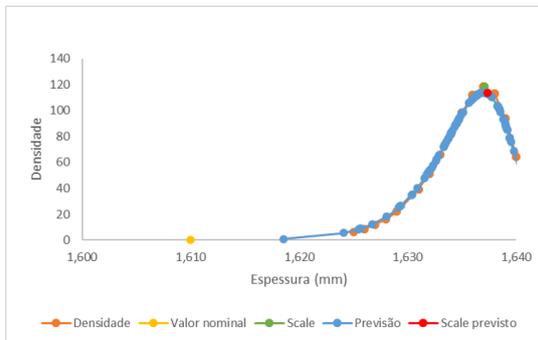
A.8 Distribuições de probabilidade previstas



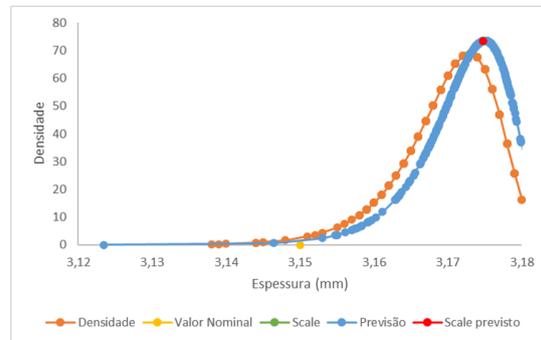
(a) Lente A.



(b) Lente B.

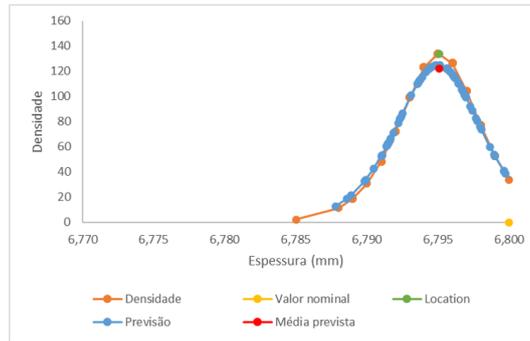


(c) Lente C.

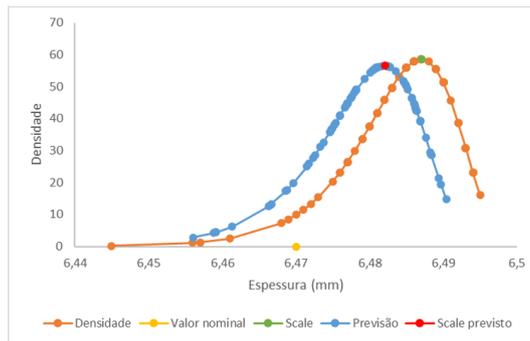


(d) Lente D.

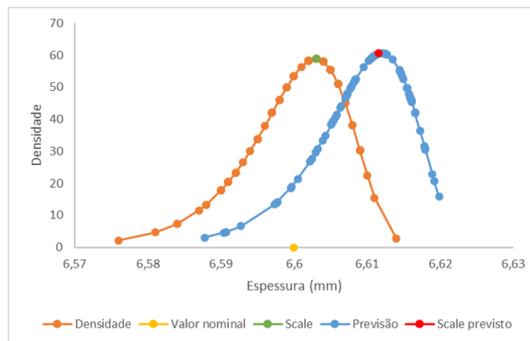
Figura 10: Densidades previstas para as lentes A a D.



(a) Lente E.



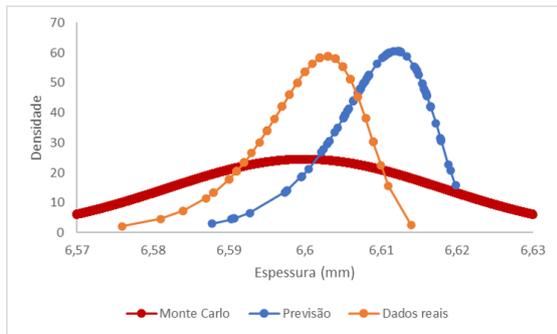
(b) Lente F.



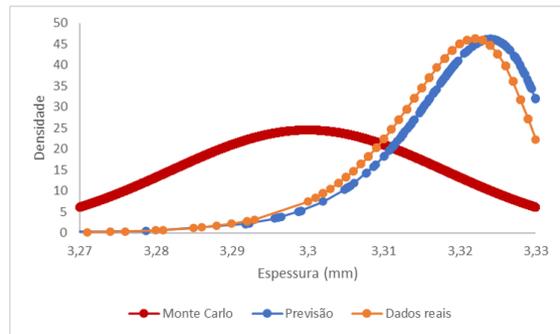
(c) Lente G.

Figura 11: Densidades previstas para as lentes E a G.

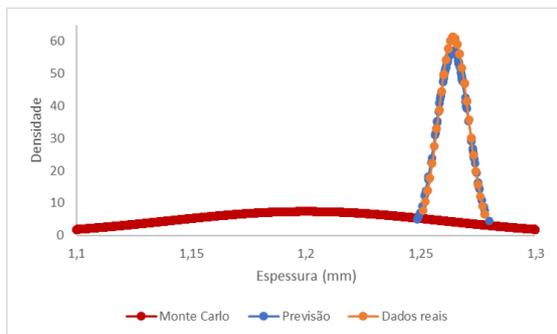
A.9 Comparação dos modelos de previsão com as simulações Monte Carlo



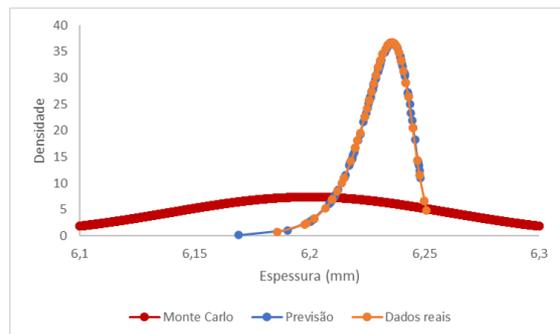
(a) Lente G.



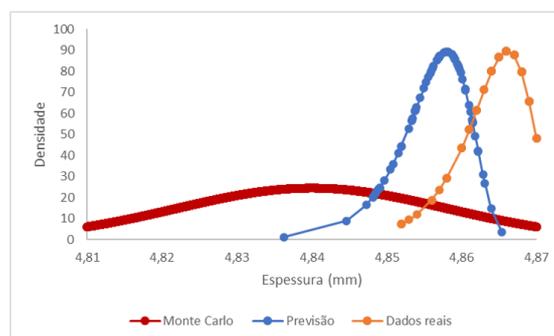
(b) Lente H.



(c) Lente I.



(d) Lente J.



(e) Lente K.

Figura 12: Distribuições Normais ajustadas aos dados das simulações de Monte Carlo vs previsões pelos modelos.

A.10 Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais

Tabela 21: Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto B.

Métricas	Previsão	Simulação Monte Carlo
EQM	$4,970 \times 10^{-6}$	0,002
EAM	0,002	0,036
EPAM	0,039%	0,726%
REQM	0,002	0,044

Tabela 22: Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto F.

Métricas	Previsão	Simulação Monte Carlo
EQM	$2,37 \times 10^{-5}$	$2,257 \times 10^{-4}$
EAM	0,005	0,014
EPAM	0,073%	0,209%
REQM	0,005	0,015

Tabela 23: Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto H.

Métricas	Previsão	Simulação Monte Carlo
EQM	$1,5596 \times 10^{-5}$	0,0003
EAM	0,0030	0,0171
EPAM	0,0960%	0,5153%
REQM	0,0040	0,0183

Tabela 24: Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto I.

Métricas	Previsão	Simulação Monte Carlo
EQM	$8,455 \times 10^{-7}$	0,006
EAM	0,001	0,065
EPAM	0,048%	5,158%
REQM	0,001	0,076

Tabela 25: Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto J.

Métricas	Previsão	Simulação Monte Carlo
EQM	$1,245 \times 10^{-6}$	0,002
EAM	0,001	0,039
EPAM	0,013%	0,626%
REQM	0,001	0,047

Tabela 26: Erros associados às Simulações de Monte Carlo e às previsões face aos dados reais, para o Conjunto K.

Métricas	Previsão	Simulação Monte Carlo
EQM	$6,084 \times 10^{-5}$	0,001
EAM	0,008	0,024
EPAM	0,160%	0,493%
REQM	0,008	0,026