

285<sup>o</sup>

# REUNIÃO DO DEPARTAMENTO DE ESTATÍSTICA

Ordinária

Extraordinária

Início: 14:40

DATA: 14/09/18  
Término: 15:54

## Lista de Presença

1. Adrian Heringer Pizzinga \_\_\_\_\_
2. Ana Beatriz Monteiro Fonseca \_\_\_\_\_
3. Ana Maria Lima de Farias \_\_\_\_\_
4. Douglas Rodrigues Pinto \_\_\_\_\_
5. Eduardo Ferioli Gomes \_\_\_\_\_
6. Hugo Henrique Kegler dos Santos \_\_\_\_\_
7. Jessica Quintanilha Kubrusly \_\_\_\_\_
8. Jony Arrais Pinto Junior \_\_\_\_\_
9. José Murilo Ferraz Saraiva \_\_\_\_\_
10. José Rodrigo de Moraes \_\_\_\_\_
11. Keila Mara Cassiano \_\_\_\_\_
12. Karina YurikoYaginuma \_\_\_\_\_
13. Luciane Ferreira Alcoforado \_\_\_\_\_
14. Ludmilla S. Viana Jacobson \_\_\_\_\_
15. Luis Guillermo Coca Velarde \_\_\_\_\_
16. Luz Amanda Melgar Santander \_\_\_\_\_
17. Márcia Marques de Carvalho \_\_\_\_\_
18. Marco Aurélio dos Santos Sanfins \_\_\_\_\_
19. Mariana Albi de Oliveira Souza \_\_\_\_\_
20. Maria Cristina Bessa Moreira \_\_\_\_\_
21. Moisés Lima de Menezes \_\_\_\_\_
22. Núbia Karla de Oliveira Almeida \_\_\_\_\_
23. Patrícia LusiéVeloza da Costa \_\_\_\_\_
24. Valentin Sisko \_\_\_\_\_
25. Wilson Calmon Almeida dos Santos \_\_\_\_\_

①  
Eduardo Ferioli Gomes  
Hugo Henrique Kegler dos Santos  
Jessica Quintanilha Kubrusly

José Rodrigo de Moraes

Karina YurikoYaginuma  
Luciane Ferreira Alcoforado

Ludmilla S. Viana Jacobson

Luz Amanda Melgar Santander

Márcia Marques de Carvalho

Mariana Albi de Oliveira Souza

Núbia Karla de Oliveira Almeida

Patrícia LusiéVeloza da Costa

Valentin Sisko

### Ata da 285ª Reunião Ordinária do Departamento de Estatística

1 Aos catorze dias do mês de setembro de dois mil e dezoito (14/09/2018) foi realizada, na sala  
2 de reuniões do Instituto de Matemática e Estatística, a 285ª (ducentésima octogésima quinta)  
3 reunião ordinária do Departamento de Estatística (GET), que se iniciou às 14h40m horas sob a  
4 presidência da professora Patrícia Lusié Velozo da Costa, chefe do GET, para deliberação  
5 sobre os seguintes itens de pauta: **1) Aprovação da ata da reunião anterior; 2) Aprovação**  
6 **da ata da 135ª reunião extraordinária; 3) Apreciação da decisão ad referendum para**  
7 **inclusão de outros docentes na banca do concurso permanente para preencher a vaga**  
8 **do docente Gilberto Pereira Sassi; 4) Aprovação do projeto de pesquisa dos professores**  
9 **Mariana Albi e Wilson Calmon ; 5) Substituição do prof. Fábio nos colegiados: 05(B)**  
10 **Ciências Sociais Bacharelado (suplente), 38 Engenharia Elétrica (titular) e 41 Engenharia**  
11 **de Telecomunicações (suplente); 6) Substituição da profa. Patrícia como suplente do**  
12 **prof. Moisés no Colegiado da Estatística (54); 7) Calendário de reuniões; 8) Informes.**  
13 Estavam presentes os seguintes professores: Douglas Rodrigues Pinto, Eduardo Ferioli  
14 Gomes, Hugo Henrique Kegler dos Santos, Jessica Quintanilha Kubrusly, José Rodrigo de  
15 Moraes, Karina Yuri Yaginuma, Luciane Ferreira Alcoforado, Ludmilla S. Viana Jacobson,  
16 Luis Guillermo Coca Velarde, Luz Amanda Melgar Santander, Márcia Marques de Carvalho,  
17 Mariana Albi de Oliveira Souza, Núbia Karla de Oliveira Almeida, Patrícia Lusié Velozo da  
18 Costa e Valentin Sisko. **Item 1)** A profa. Patrícia submeteu à votação a aprovação da ata da  
19 284ª reunião ordinária, que foi aprovada por unanimidade. **Item 2)** A profa. Patrícia submeteu à  
20 votação a aprovação da ata da 135ª reunião extraordinária, que foi aprovada por unanimidade.  
21 **Item 3)** A professora Patrícia explicou que para compor a banca do concurso para professor  
22 permanente do GET que ocorrerá entre 05 a 09 de novembro de 2018 (editais 164 e 165/2018)  
23 era necessário atender alguns critérios: ter 5 membros externos a UFF; esses membros  
24 deveriam ser todos de instituições diferentes; os nomes precisavam constar em Ata e atender  
25 aos princípios da impessoalidade e da moralidade administrativa, dentre as quais se incluem:  
26 não ter parentesco até quarto grau; não ter afinidade colateral até terceiro grau ou em linha  
27 reta; não ter relacionamento marital ou afetivo no presente ou no passado; não ter orientado no  
28 Mestrado ou Doutorado nos últimos 10 (dez) anos; não ter coautorias, nos últimos 10 (dez)  
29 anos, que totalizem pelo menos 25% dos trabalhos científicos do membro da Comissão ou do  
30 candidato no referido período; não ter coautorias, nos últimos 10 (dez) anos, entre membros da  
31 Comissão que totalizem pelo menos 25% dos trabalhos científicos de um deles no referido  
32 período. Recorrendo a listagem de nomes constando na Ata da 281ª Reunião Ordinária (RO)  
33 do GET e juntando esses requisitos com as disponibilidades dos professores convidados, não  
34 foi possível ter algum professor da Fiocruz selecionado inicialmente. Porém, na 281ª RO, foram  
35 selecionadas apenas 5 instituições externas. Sendo assim, a Comissão de Pessoal Docente  
36 (CPD) solicitou a Chefia no dia 22 de agosto de 2018 que incluísse um quinto membro externo  
37 e que isso teria que ser realizado no mesmo dia. Para isso, indicaram fazer uma decisão Ad  
38 Referendum. Devido a falta de luz no Gragoatá, a CPD permitiu entregar a documentação  
39 necessária para isso no dia seguinte (23 de agosto de 2018). E, por não ter tempo hábil de  
40 realizar uma reunião, a professora Patrícia tomou a decisão Ad Referendum, conforme  
41 permitido pelo Art. 39 §III do Estatuto e Regimento Geral da UFF, incluindo outros nomes de  
42 professores para serem possíveis membros de banca. E também precisou obter o aceite e a  
43 documentação assinada pelos docentes Iuri da Costa Leite (Fiocruz) e Josiane da Silva  
44 Cordeiro Coelho (UFRRJ) entre os dias 22 e 23 de agosto para serem o segundo e o terceiro

45 membro suplente, respectivamente. Após esses esclarecimentos, a profa. Patrícia submeteu à  
46 votação a aprovação da ata da 1ª decisão ad referendum, que foi aprovada por unanimidade.  
47 **Item 4)** A profa. Patrícia apresentou o parecer da comissão de pesquisa favorável à aprovação  
48 do projeto de pesquisa dos professores Mariana Albi de Oliveira Souza e Wilson Calmon  
49 Almeida dos Santos intitulado “Estimação do Número de Grupos em Contexto de Dados de  
50 Ranking através do Modelo Placket-Luce: uma Abordagem Hierárquica Bayesiana”. A profa.  
51 Patrícia colocou em votação a aprovação do projeto, que foi aprovado por unanimidade. **Item**  
52 **5)** A profa. Patrícia perguntou se alguém se candidatava para ocupar as vagas nos Colegiados  
53 do professor Fábio e foi solicitada a relação de Colegiados de cada docente. Com base nessa  
54 relação, selecionou-se os seguintes membros: José Murilo Ferraz Saraiva para ser suplente do  
55 Colegiado de Ciência Sociais Bacharelado (05 B), Jony Arrais Pinto Junior para ser titular do  
56 Colegiado de Engenharia Elétrica (38) e Jessica Quintanilha Kubrusly para ser suplente do  
57 Colegiado de Engenharia de Telecomunicações (41). **Item 6)** A professora Patrícia era  
58 suplente no Colegiado de Estatística (54), mas, como Chefe do GET, ela tornou-se membro  
59 nato desse Colegiado. Sendo assim, a professora Karina Yuriko Yaginuma se ofereceu para  
60 ser membro suplente do professor Moisés Lima de Menezes no referido Colegiado. **Item 7)** As  
61 próximas reuniões em 2018 ocorrerão nas seguintes datas: 05 de outubro, 09 de novembro e  
62 14 de dezembro. **Item 8)** Foram dados os seguintes informes: As professoras Jéssica, Mariana  
63 e Patrícia, membros da Comissão da X Semana da Estatística (SEMEST) pediram a  
64 colaboração dos demais docentes na divulgação e na participação da X SEMEST que ocorrerá  
65 entre os dias 16 e 19 de outubro de 2018. As inscrições para participação, para submissão de  
66 trabalhos e para trabalhar como monitores já encontram-se abertas no site [www.semest.uff.br](http://www.semest.uff.br).  
67 A professora Patrícia informou que as atas das reuniões departamentais passarão a estar  
68 disponíveis no site <http://www.atas.uff.br/>. A professora Patrícia informou que na próxima  
69 sexta-feira, dia 21 de setembro, ocorrerá o próximo Papos Aleatórios com a docente e ex-aluna  
70 Rebecca de Oliveira Souza sobre o trabalho Processos pontuais espaciais para dados das  
71 unidades prisionais no Brasil em conjunto com os docentes Marina Silva Paez e Vinicius  
72 Pinheiro Israel. Mais detalhes sobre os Papos Aleatórios podem ser visualizados em  
73 <https://sites.google.com/view/paposaleatorios/p%C3%A1gina-inicial?authuser=0>. A professora  
74 Mariana informou que como correspondente da Associação Brasileira de Estatística (ABE) ela  
75 solicita, de tempos em tempos, aos colegas de departamento, notícias para divulgar no boletim  
76 da ABE. Sendo esta uma forma de divulgar nosso departamento dentro da comunidade  
77 estatística, ela reforçou o pedido de colaboração de todos os docentes do departamento  
78 enviando informações relevantes tais como eventos (realizados e futuros), concursos,  
79 publicações recentes, entre outras. O professor Eduardo informou sobre a criação da Revista  
80 de Iniciação Científica Papa-Goiaba do IME-UFF. Esta revista será uma publicação online  
81 (<http://rpg.sites.uff.br/>) de fluxo contínuo, cujo objetivo será divulgar a produção acadêmica dos  
82 trabalhos de iniciação científica, estudos orientados, monografias de conclusão de curso e  
83 demais atividades de iniciação à pesquisa em Matemática, Estatística ou áreas afins. Nada  
84 mais havendo a tratar e ninguém mais desejando fazer uso da palavra, foi encerrada a reunião  
85 às 15h54m, cuja ata vai datada e assinada por mim, Patrícia Lusié Velozo da Costa, chefe do  
86 Departamento de Estatística. Niterói, 14 de setembro de 2018.

*Patrícia Lusié Velozo da Costa*

Patrícia Lusié Velozo da Costa  
Chefe Depto de Estatística - UFF  
SIAPE 1805333

1ª ATA AD REFERENDUM DO DEPARTAMENTO DE ESTATÍSTICA DA UNIVERSIDADE FEDERAL  
FLUMINENSE

1           Aos vinte e três dias do mês de agosto de dois mil e dezoito (23/08/2018) foi realizada, na Sala da Chefia do  
2 Departamento de Estatística (GET), no 3º andar do bloco G do Campus Gragoatá, a 1ª Reunião Ad Referendum do  
3 Departamento de Estatística da Universidade Federal Fluminense, que se iniciou às 08h00min, presidida pela Chefe de  
4 Departamento Patrícia Lusié Velozo da Costa, para deliberação do seguinte item: 1) **Inclusão de novos membros na banca  
5 do concurso para a vaga do professor Gilberto Pereira Sassi que ocorrerá entre os dias 05 a 09 de novembro de 2018.**  
6 Para compor a banca do referido concurso, convidou-se os docentes eleitos pelo Departamento de Estatística na 281ª Reunião  
7 Ordinária ocorrida no Departamento de Estatística no dia vinte e sete de abril de dois mil e dezoito (27/04/2018). O membro da  
8 banca não pode possuir ligações com algum dos candidatos inscritos neste Concurso ou com outro componente da Comissão  
9 Examinadora, dentre as quais se incluem: Parentesco até quarto grau; Afinidade colateral até terceiro grau ou em linha reta;  
10 Relacionamento marital ou afetivo no presente ou no passado; Orientação de Mestrado ou Doutorado nos últimos 10 (dez)  
11 anos; Coautorias, nos últimos 10 (dez) anos, que totalizem pelo menos 25% dos trabalhos científicos do membro da Comissão  
12 ou do candidato no referido período; Coautorias, nos últimos 10 (dez) anos, entre membros da Comissão que totalizem pelo  
13 menos 25% dos trabalhos científicos de um deles no referido período. Sendo assim, é necessário ter no mínimo 5 docentes de  
14 instituições diferentes da Universidade Federal Fluminense para compor a banca (incluindo membro titular e suplente) e na  
15 281ª Reunião, sugeriu-se apenas 5 instituições diferentes. Na Fiocruz haviam 3 nomes possíveis: o Cosme Marcelo Furtado  
16 Passos da Silva, porém ele tem impedimentos com o candidato Carlos Eduardo Hirth Pimentel (orientou no mestrado e no  
17 doutorado), a docente Aline Araújo Nobre que estará indisponível no período do concurso e o docente Leonardo Soares Bastos  
18 que tem impedimento com o candidato Marcio Watanabe Alves de Souza (artigo em vias de publicação). E como não havia  
19 outro docente dessa instituição e nem outra instituição no Rio de Janeiro diferente dos demais membros fez-se necessária a  
20 inclusão de novos membros suplentes para compor a banca. A Coordenação de Pessoal Docente solicitou no dia 22 de agosto  
21 de 2018 a Chefe de Departamento da Estatística a inclusão de novo(s) membro(s) e pediu para a documentação ser entregue no  
22 mesmo dia. Como o campus do Gragoatá ficou sem energia nesse dia, a Cecília, responsável pela Coordenação de Pessoal  
23 Docente, permitiu a entrega dessa documentação no dia 23 de agosto de 2018. Portanto, com a urgência dessa documentação,  
24 foi necessária que essa inclusão fosse tomada ad referendum, conforme permitido pelo Artigo 39 §III do Estatuto e Regimento  
25 Geral da Universidade Federal Fluminense. Sendo assim, eu, Patrícia Lusié Velozo da Costa aprovo ad referendum a inclusão  
26 de um ou mais dos seguintes nomes para compor a banca: na Fiocruz: Iuri da Costa Leite, Geraldo Marcelo da Cunha ou  
27 Cleber Nascimento do Carmo, na UFRRJ: Josiane da Silva Cordeiro Coelho ou Felipe Leite Coelho da Silva e/ou na PUC:  
28 Helio Cortes Vieira Lopes. Nada mais havendo a tratar, foi encerrada a reunião às 08h30m, cuja ata vai datada e assinada por  
29 mim, Patrícia Lusié Velozo da Costa, Chefe do Departamento de Estatística. Niterói, 23 de agosto de 2018.

Patrícia Lusié Velozo da Costa  
Chefe Depto de Estatística - UFF  
SIAPE 1805333



Departamento de Estatística UFF &lt;chefia.est.uff@gmail.com&gt;

**Re: EDITAL 165/2018 – ÁREA: PROBABILIDADE E ESTATÍSTICA – CADASTRO DE MEMBRO DE BANCA**

CPD UFF &lt;cpd.certame@gmail.com&gt;

22 de agosto de 2018 11:43

Para: Chefia GET-UFF &lt;CHEFIA.EST.UFF@gmail.com&gt;, patricialusie@ig.com.br

Cc: Dgl Cpd &lt;concursos.cpd@gmail.com&gt;, CPD UFF &lt;suporte.cpd.uff@gmail.com&gt;, Chefia DGLD &lt;chefia.dgld.cpd@gmail.com&gt;, Carla Aparecida Florentino Rodrigues &lt;carlarodrigues@id.uff.br&gt;, Valéria Lima &lt;cpd.pare@gmail.com&gt;

Prezados,

Retificando o e-mail anterior, informamos que não podem compor a mesma banca membros de mesma instituição externa.

Portanto, **NÃO** podem participar da mesma banca os professores **WALDIR JESUS DE ARAÚJO LOBÃO** e **LARISSA DE CARVALHO ALVES**.

Caso o Departamento tenha interesse em prosseguir com o concurso, Aguardamos a chegada à CPD hoje, até as 16h, da nova documentação de banca aprovando o novo membro.

Para facilitar, os documentos que devem chegar são:

- Ata aprovando o novo membro (pode ser uma ata *ad referendum*)
- Nova Declaração de Composição de Banca
- Declaração de não impedimento do novo membro

Att,

Cecília Martins

Divisão de Gestão e Lotação Docente  
Coordenação de Pessoal Docente  
Pro- Reitoria de Gestão de Pessoas  
Universidade Federal Fluminense  
(21) 2629-5272

Em qua, 22 de ago de 2018 às 11:28, CPD UFF <cpd.certame@gmail.com> escreveu:

Prezados,

Informamos que não foi possível cadastrar o Prof. **ALEXANDRE SOUSA DA SILVA** no concurso para a área PROBABILIDADE E ESTATÍSTICA porque o mesmo não possui cadastro no sistema.

Para realizar o cadastro, é preciso que ele siga os passos abaixo:

- 1) Acessar o site <https://app.uff.br/cpd>
- 2) Clicar na aba "Login" e em "Cadastre-se aqui"
- 3) Preencher a primeira tela de cadastro (Informações pessoais) e clicar em "Avançar"
- 4) Preencher a segunda tela de cadastro (Informações de contato), assinalando SIM à pergunta "Vai ser examinador ou secretário de alguma banca?" e clicar em "Avançar"
- 5) Preencher a terceira tela de cadastro (Informações profissionais) e clicar em "Confirmar".

Informamos que não foi possível cadastrar a Prof. **LARISSA DE CARVALHO ALVES** porque a mesma não possui cadastro como membro de banca no sistema.

Para atualizar o cadastro, é preciso que ela siga os passos abaixo:

- 1) Acessar o site <https://app.uff.br/cpd>
- 2) Fazer o login com CPF e senha
- 3) Clicar na aba "Minha conta" e em "Dados cadastrais"

- 4) Clicar em “Avançar” ao final da primeira tela de cadastro
- 5) Assinalar a opção SIM na pergunta “Vai ser examinador ou secretário de alguma banca?”
- 6) Preencher a terceira tela com os dados atualizados e clicar em CONFIRMAR

Além do cadastro, é necessário que o Professor faça o upload no sistema dos seguintes documentos:

- 1) Comprovante da Titulação declarada
- 2) Comprovante do vínculo com a instituição de ensino e/ou pesquisa declarada
- 3) Resumo do currículo

Pedimos que repasse as orientações aos Professores para que eles possam realizar ou atualizar o cadastro e para que possamos finalizar o cadastro da banca no sistema.

Ressaltamos a necessidade de os membros se cadastrarem tanto para a realização do concurso como para a requisição de pagamento de pró-labore, diárias e passagens, conforme cada caso, uma vez que o cadastro é necessário para fazer o orçamento do que será gasto em cada concurso.

Solicitamos também que nos informem quem será o **PRESIDENTE DA BANCA**.

Contamos com a compreensão dos senhores.  
Em caso de dúvidas, estamos à disposição

Att.,

Cecília Martins

Divisão de Gestão e Lotação Docente  
Coordenação de Pessoal Docente  
Pro- Reitoria de Gestão de Pessoas  
Universidade Federal Fluminense  
(21) 2629-5272



Universidade Federal Fluminense  
Instituto de Matemática e Estatística  
Departamento de Estatística  
Parecer de Proposta sobre Atividade de Pesquisa

05/09/2018

**Proponentes:** Mariana Albi de Oliveira Souza (SIAPE: 1809003)  
Wilson Calmon Almeida dos Santos (SIAPE: 2197625)

**Tipo de atividade:** Projeto de Pesquisa

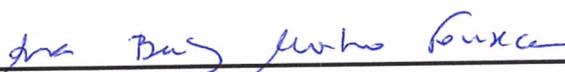
**Título:** Estimaco do Nmero de Grupos em contexto de Dados de Ranking atravs do Modelo Plackett-Luce: uma Abordagem Hierrquica Bayesiana

**PARECER:**

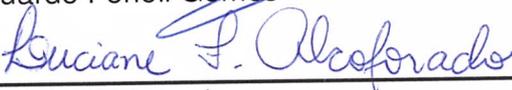
*A Comisso de Pesquisa, no uso de suas atribuices, avaliou o projeto acima identificado. Ao considerar a Instruico de Servio 04-2016 do Departamento de Estatística (GET), por entender que o texto e os objetivos do projeto se enquadram no que é, por essa, determinando, e que a proposta atende aos interesses do Departamento de Estatística, aceita incondicionalmente a atividade proposta pelos docentes.*

*Tendo em vista que o tempo de execuo do projeto proposto é de dois anos, conforme descrito no Art. 9º da Instruico de Servio 04-2016 do GET, os docentes responsveis devero entregar relatrios parciais ao final de cada ano para efeitos de pontuao naquele ano.*

*Niteri, 05 de setembro de 2018*

  
\_\_\_\_\_  
Ana Beatriz Monteiro Fonseca

  
\_\_\_\_\_  
Eduardo Ferioli Gomes

  
\_\_\_\_\_  
Luciane Alcoforado



### Projeto de Pesquisa:

#### *Estimação do Número de Grupos em contexto de Dados de Ranking através do Modelo Plackett-Luce: uma Abordagem Hierárquica Bayesiana*

**Identificação do Projeto:** Projeto de pesquisa intitulado “Estimação do Número de Grupos em contexto de Dados de Ranking através do Modelo Plackett-Luce: uma Abordagem Hierárquica Bayesiana”

**Nome e matrícula SIAPE dos docentes responsáveis:** Mariana Albi de Oliveira Souza (SIAPE: 1809003) e Wilson Calmon Almeida dos Santos (SIAPE: 2197625).

**Introdução:** São abundantes os dados de ranqueamento [doravante, **RD** de *Ranking Data*]. Listas de “10 mais”, “100 mais” são comuns e, em diversas áreas: melhores periódicos, carros, filmes, times de futebol, etc. Além disso, como destacado em Marden [1], uma estratégia não paramétrica usual consiste em transformar dados oriundos de populações supostamente contínuas em RD. RD são essencialmente multivariados e, dessa forma, técnicas tradicionais de análise multivariada podem ser empregadas para analisá-los. Em particular, técnicas de agrupamento podem ser utilizadas para identificar grupos e empates não revelados. Tipicamente, o número de grupos é um insumo de muitos procedimentos de agrupamento e a importância deste parâmetro é frequentemente negligenciada. Todavia, nos últimos 20 anos, métodos interessantes foram propostos para estimá-lo. Ilustrativamente, podemos citar: Tibshirani et al [2], Fang and Wang [3], Fujita et al [4] ou Liang et al [5]. Contudo, nenhum deles foi desenvolvido especificamente para RD. Nosso objetivo é desenvolver métodos específicos para estimação do número de grupos no contexto de RD a partir do modelo Plackett-Luce - [6] e [7] - através de uma abordagem bayesiana.

**Revisão Bibliográfica:** Milligan e Cooper [8] apresentam uma vasta revisão e comparação, via simulação, de métodos estatísticos propostos para lidar com o problema da estimação/identificação do número de grupos em problemas de agrupamento. Como destacado em Gordon [9], os diversos métodos encontrados na literatura até meados da década de 1990 careciam de maior justificativa teórica – em certo sentido, podemos dizer que o problema de estimação do número de grupos era negligenciado. A maioria dos métodos era baseada em inspeção visual ou regras de ouro. Muitos métodos propostos consistiam em adotar medidas baseadas nas dissimilaridades agregadas dentro [*Within*] e entre [*Between*] grupos. Podemos citar, por exemplo, as contribuições de Calinski e Harabasz [10], Hartigan [11] e Krzanowski e Lai [12]. O trabalho de Tibshirani et al [2] deu uma nova dimensão ao problema. Os autores propuseram estimar o número de grupos minimizando uma medida de discrepância definida como estatística *GAP* que mede a distância entre a dissimilaridade *Within* e seu valor esperado sob uma dada distribuição de referência. Alternativas sofisticadas têm sido propostas na literatura desde então. Fang e Wang [3] propõe usar o método de bootstrap para estimar a instabilidade de agrupamento, medida esta que, como discutido em Wang [13], pode ser utilizada para estimar o número de grupo – o número de grupos ideal seria aquele que minimiza a instabilidade do agrupamento. Fujita et al [4] propõe uma abordagem não paramétrica em que o objetivo é minimizar a estatística *Slope*, uma sofisticação da estatística *Silhouette* proposta por Rousseeuw [14]. A estatística *Silhouette* avalia a adequação do agrupamento por comparar a distância de cada objeto ou indivíduo em relação a seu próprio grupo com o grupo distinto mais próximo. A estatística *Slope*, por sua vez, pode ser vista como uma média geométrica entre o nível da estatística *Slope* e sua variação quando o número de grupos aumenta em

uma unidade. Em Liang et al [5] é apresentado um método de Agrupamento que permite também a estimação do número de grupos através da minimização da entropia da informação. Os autores trabalham com um contexto de dados mistos e utilizam uma abordagem não paramétrica para medir a entropia no caso contínuo e no caso discreto. A entropia aparece como uma alternativa às medidas de dissimilaridade usuais e, segundo os autores, tem por objetivo estabelecer uma métrica mais geral do que as usuais.

**Motivação/Justificativa:** Como destacado em Gordon [9], diversos métodos de estimação [ou identificação] foram propostos na literatura para contextos bem gerais. Entretanto, como destacado pelo próprio autor, originalmente os métodos desenvolvidos careciam de uma maior justificativa teórica e apresentavam-se como “regras de ouro”, refletindo a pouca importância dada para tal parâmetro. O trabalho de Tibshirani et al [2] provocou uma reflexão sobre o tema e motivou o desenvolvimento de novas alternativas como, por exemplo, Fang e Wang [3], Fujita et al [4] ou Liang et al [5]. Nenhum destes métodos, porém, foi desenvolvido para lidar com RD. Nesse sentido, buscamos desenvolver um método que preencha a lacuna de um tipo de dado que é abundante, importante e não necessariamente satisfatoriamente tratado com os procedimentos mais gerais [desenvolvidos para dados oriundos de populações contínuas ou categóricas usuais].

**Metodologia/materiais e métodos:** Um modelo para RD [sem empates] é uma coleção de distribuições de probabilidade sobre o conjunto de permutações dos inteiros  $1, \dots, k$  se  $k$  é o número de objetos, conforme Marden [1]. O modelo de Plackett-Luce - [6] e [7] - apresenta uma proposta simples de atribuir probabilidade às permutações no qual é associado a cada objeto um parâmetro escalar não negativo  $\theta_1, \dots, \theta_k$ . A probabilidade de o  $i$ -ésimo objeto ser classificado como melhor que os demais é diretamente proporcional à  $\theta_i$ . Nós pretendemos estender tal modelo de modo a associar à função de probabilidade dos *Rankings* o número de grupos. Em investigação [em andamento], conseguimos descrever tal dependência adotando uma abordagem bayesiana hierárquica na qual o número de grupos  $g$  é um hiperparâmetro que afeta diretamente a quantidade de objetos em cada grupo  $d_1, \dots, d_g$ , onde  $d_l$  representa a quantidade de objetos no grupo  $l$ . A quantidade de objetos em cada grupo é outro parâmetro que afeta a classificação de cada objeto  $c_1, \dots, c_k$ , onde  $c_l$  representa o grupo no qual o objeto  $l$  foi classificado. As classificações definem um hiperparâmetro que afeta, em conjunto a um parâmetro de escala  $s$ , diretamente a distribuição de  $\theta_1, \dots, \theta_k$ . O algoritmo Metropolis-Hasting - Gamerman [15] - está sendo implementado para a obtenção da posteriori de  $g$ , cuja moda produz nosso estimador do número de grupos. Uma ilustração do método proposto deve ser feita com dados das velocidades em voltas mais rápidas feitas por pilotos da fórmula Indy em treinos na temporada 2017. Para avaliar a performance do método proposto em comparação com as principais alternativas encontradas na literatura, será conduzido um estudo de simulação baseado em modelos para RD e outros compatíveis com as metodologias concorrentes.

**Resultados esperados:** Neste trabalho esperamos apresentar estimadores não viesados do número de grupos o contexto de RD que sejam também mais eficientes do que os procedimentos usuais de estimação do número de grupos como aqueles mencionados na introdução e revisão bibliográfica. Esperamos ainda que os métodos propostos produzam bons resultados em contextos mais gerais [não necessariamente, RD], sobretudo, quando os dados forem contaminados pela presença de *outliers*.

**Cronograma de Execução:** na sequência é apresentado o cronograma de execução. O trabalho teve início em Janeiro de 2018 e tem previsão de duração até Junho de 2019.

Tarefas (cronograma)	ANO 1 (2018)												ANO 2 (2019)					
	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	JAN	FEB	MAR	ABR	MAI	JUN
	1 Levantamento Bibliográfico Estimção Número Grupos	X	X	X	X													
2 Estudo do Modelo Plackett-Luce [PL]	X	X	X															
3 Implementação Análise Bayesiana Modelo PL	X	X	X	X	X													
7 Levantamento Bibliográfico de Métodos Alternativos				X	X	X	X	X	X	X	X	X	X					
8 Implementação Simulações PL [Programa R. SIM]					X	X	X	X	X	X	X	X						
9 Implementação Análise Simulações contexto geral							X	X	X	X	X	X						
10 Escrita: Descrição Nossa Metodologia					X	X	X	X	X	X	X	X						
11 Implementação Métodos Alternativos [Programa R. ASD]							X	X	X	X	X	X	X	X				
12 Escrita: Descrição Métodos Alternativos										X	X	X	X	X	X	X	X	
13 Implementação Simulações Alternativas [Programa R. SIM]											X	X	X	X	X	X	X	X
14 Produção de Resultados																		
15 Análise Resultados Simulações e Escrita														X	X	X	X	X
16 Submissão como Working Paper																X	X	X

#### Equipe responsável:

Membro	Matrícula	Categoria
Mariana Albi de Oliveira Souza	1809003	Professor GET
Wilson Calmon Almeida dos Santos	2197625	Professor GET

#### REFERÊNCIAS

- [1] Marden, John I. (1995): "Analyzing and Modeling Rank Data"; Chapman & Hall/CRC Monographs on Statistics & Applied Probability, London, 1st Edition.
- [2] Tibshirani, Robert ; Walther, Guenther; Hastie, Trevor (2001): "Estimating the Number of Clusters in a Data Set Via the Gap Statistic"; Journal of the Royal Statistical Society Series B (Statistical Methodology) 63(2):411-423.
- [3] Fang, Yixin and Wang, Junhui (2012): "Selection of the number of clusters via the bootstrap method"; Computational Statistics and Data Analysis, 56, 468–477.
- [4] Fujita, André; Takahashi, Daniel Y.; Patriota, Alexandre G. (2014): "A non-parametric method to estimate the number of clusters"; Computational Statistics and Data Analysis, 73, 27–39.
- [5] Liang, Jiye; Zhao, Xingwang; Li, Deyu; Cao, Fuyuan; Dang, Chuangyin (2012): "Determining the number of clusters using information entropy for mixed data"; Pattern Recognition, 45, 2251–2265.
- [6] Plackett, Robert L. (1975): "The Analysis of Permutations." Appl. Statist 24 (2):193–202.
- [7] Luce, R. D. (1959). Individual choice behavior. New York: Wiley.
- [8] Milligan, G.W., Cooper, M.C., 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50, 159–179.
- [9] Gordon, A. (1996): "Classification"; 2nd edn. London: Chapman and Hall-CRC.
- [10] Calinski, R.B., Harabasz, J., 1974. A dendrite method for cluster analysis. Commun. Stat. 3, 1–27.
- [11] Hartigan, J. (1975): "Clustering Algorithms"; New York, Wiley.
- [12] Krzanowski, W.J., Lai, Y.T., (1988): "A criterion for determining the number of clusters in a data set". Biometrics 44, 23–34.
- [13] Wang, J. (2010): "Selection of the number of clusters via the bootstrap method"; Biometrika, 97, 4, 893–904.
- [14] Rousseeuw, P. (1987): "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis"; J. Comput. Appl. Math.; 20, 53–65.
- [15] Gamerman, D. (1997): "Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference"; Chapman and Hall; London, 1st Edition.

Niterói, 15 de agosto de 2018.

---

Mariana Albi de Oliveira Souza [SIAPE: 1809003]

---

Wilson Calmon Almeida dos Santos [SIAPE: 2197625]