

PROBABILIDADE E ESTATÍSTICA

**MODERNIZAÇÃO DA AVALIAÇÃO NA DISCIPLINA
PROBABILIDADE E ESTATÍSTICA**

RELATÓRIO FINAL

Março de 2020



Universidade de Brasília



A necessidade de modernização do sistema de avaliação da disciplina Probabilidade e Estatística surgiu com o processo de unificação da disciplina no departamento de Estatística da Universidade de Brasília. Buscando manter a consistência na execução do projeto pedagógico nas diversas turmas e reduzir os custos operacionais preservando a qualidade das avaliações, optou-se pela elaboração de um sistema de avaliação automatizado para geração e correção das provas.

O sistema dispõe-se das funcionalidades computacionais da ferramenta R/exams, incluindo a aleatorização do enunciado das questões para geração de provas individualizadas para os alunos. Trabalhamos com a extensão desse pacote para maximizar a automação do processo avaliativo em todas as suas fases: elaboração, impressão, aplicação, correção, divulgação e revisão das provas. Além disso, o sistema conta com um banco com 300 modelos de questões de Estatística básica elaboradas em linguagem R.

O sistema também foi desenvolvido considerando sua possível aplicação em outras disciplinas e Universidades. Desejamos compartilhar e disponibilizar à comunidade do sistema educacional público o resultado desse trabalho. Sendo assim, esse relatório tem por objetivo detalhar e divulgar o produto consolidado. No Capítulo 1, apresentamos uma visão geral do sistema para o leitor. O Capítulo 2 ilustra alguns resultados interessantes da aplicação do novo sistema de avaliação na disciplina Probabilidade e Estatística. O Capítulo 3 traz um manual detalhado de utilização do sistema, incluindo o processo de geração e correção das provas. Por fim, no Apêndice, encontram-se exemplos de uma prova, do resultado apresentado ao aluno e da lista de fórmulas adotada nas avaliações.

Este trabalho contou com a enorme dedicação de professores, alunos e equipe administrativa do departamento de Estatística da UnB, conforme detalhado na página a seguir. Agradecemos também à Coordenação do departamento pelo apoio irrestrito dado à equipe de trabalho, ao Programa Aprendizagem para o 3o Milênio (A3M) da UnB pelo financiamento de dois alunos e ao Programa de Iniciação Científica (PIBIC) que ofertou uma bolsa de estudos.

Brasília, março de 2020
Prof. Guilherme Souza Rodrigues
Profa. Thais Carvalho Valadares Rodrigues

Ilustração da capa
(criada a partir de dados de criminalidade na cidade de Chicago)
Gregory Matthews
The R Graph Gallery

Comissão de Planejamento Unificado da disciplina Probabilidade e Estatística

Prof. Jhames Matos Sampaio

Prof. André Luiz Fernandes Cançado

Profa. Thais Carvalho Valadares Rodrigues

Prof. Guilherme Souza Rodrigues (Coordenador A3M)

Professores envolvidos na criação e revisão do banco de questões

André Luiz Fernandes Cançado

Antônio Eduardo Gomes

Cira Etheowalda Guevara Otiniano

Eduardo Yoshio Nakano

Guilherme Souza Rodrigues

Helton Saulo Bezerra dos Santos

Jhames Matos Sampaio

José Augusto Fiorucci

Leandro Tavares Correia

Lucas Moreira

Peter Zörnig

Roberto Vila Gabriel

Thais Carvalho Valadares Rodrigues

Equipe administrativa

Tathyanna Martins Cordeiro

Edenilson Lopes Carvalho

Lucas Fernandes de Albuquerque Lira

Karen Luíza Silva de Sousa

André Oliveira

Alunos

Álefe Lacerda Gomes Santos (bolsista A3M)

Lucas de Moraes Bastos (bolsista A3M e PIBIC)

Tonny Barbosa Pereira (TCC)

Rodrigo Fernando Murca Barroso (Monitoria)

Vítor Macêdo Rocha (PIBIC)

Sumário

1	Introdução	2
1.1	Objetivos	4
1.2	Desenvolvimento	4
1.3	Aspectos gerais do sistema	5
1.4	Descrição dos produtos entregues	6
2	Análise de resultados	8
2.1	Análise descritiva dos dados	10
2.1.1	Desempenho dos alunos	10
2.1.2	Prova substitutiva	11
2.2	Análise via Teoria da Resposta ao Item	14
2.2.1	Resultados preliminares	15
2.2.2	Avaliação das questões	16
2.3	Conclusões, recomendações e trabalhos futuros	25
3	Manual de procedimentos	26
3.1	Pré-requisitos	27
3.1.1	Programas	27
3.1.2	Pastas e arquivos	27
3.2	Geração das provas	31
3.2.1	Localizando as pastas	31
3.2.2	Carregando as funções suplementares	31
3.2.3	Organizando o cadastro dos alunos	32
3.2.4	Definindo os parâmetros da prova	33
3.2.5	Definindo os temas da prova	33
3.2.6	Sorteando as questões para as turmas	34
3.2.7	Criando múltiplos exames	35
3.2.8	Salvando a área de trabalho	36
3.2.9	Considerações finais	36
3.3	Correção das provas	38

3.3.1	Instruções gerais	38
3.3.2	Carregando os arquivos	39
3.3.3	Localizando as pastas	39
3.3.4	Digitalizando as respostas	40
3.3.5	Checando as leituras	41
3.3.6	Gerando os resultados	42
3.3.7	Salvando a área de trabalho	43
3.3.8	Considerações finais	43
A	Apêndice	45
A.1	Exemplo de Prova	46
A.2	Exemplo de Resultado	51
A.3	Lista de Fórmulas	59
A.4	Tabelas dos parâmetros das questões	61

Introdução

O Departamento de Estatística (EST) da UnB oferece semestralmente, em média, 10 turmas de Probabilidade e Estatística, totalizando aproximadamente 500 alunos de diversos cursos da área das Ciências Exatas, incluindo Ciências Econômicas, Química, Matemática, Computação e diversas Engenharias. A partir do segundo semestre de 2017, com a unificação da disciplina aprovada na 490^a Reunião Ordinária do Colegiado do Departamento de Estatística, todos os professores passaram a adotar a mesma ementa e o mesmo sistema de avaliação, visando a padronização do curso oferecido e as vantagens decorrentes.

Para tanto, surgiu a necessidade da aplicação de provas unificadas a todos os alunos matriculados na disciplina. Obviamente, turmas de dias/horários diferentes não podiam fazer a mesma avaliação. Diante desse desafio, uma opção operacionalmente onerosa era reunir todos os alunos simultaneamente para aplicação dos exames. Algumas tentativas ocorreram aos sábados, mas foram majoritariamente frustradas por diversas razões logísticas. Tal estratégia demandava a elaboração de provas que seguiam certa padronização, mas que deveriam ser suficientemente distintas. Esse processo de elaboração e de correção manual era demasiadamente dispendioso devido ao número de turmas e de alunos envolvidos.

No momento da unificação, decidiu-se adotar provas exclusivamente objetivas, de múltipla escolha. Embora haja mérito na discussão das vantagens e desvantagens de cada tipo de prova: *aberta, fechada* ou *híbrida*, intencionalmente evitamos incluir esse debate no escopo do projeto.

Simultaneamente a esse movimento do Departamento de Estatística da UnB, e diante dos mesmos desafios aqui encontrados, um grupo de pesquisadores da Universidade de Economia e Negócios de Viena, liderados pelo Professor Achim Zeileis, criaram uma biblioteca de funções em linguagem R, batizada como *R/Exams*, que permite a automação das fases de elaboração e correção de provas em cursos de graduação. Entre outras inúmeras funcionalidades, o pacote permite que os valores contidos no enunciado de cada questão sejam gerados aleatoriamente, de modo que cada aluno tenha uma prova específica, diferente das demais. Além do desejável efeito na redução do risco de fraudes, tal abordagem garante longevidade ao banco de questões, uma vez que não basta ao aluno decorar as respostas, mas sim dominar todos os passos necessários para se chegar à solução desejada.

A cada nova aplicação das provas no formato proposto, um valioso conjunto de dados se torna disponível, permitindo, entre outros usos, a calibração prévia do nível de dificuldade das provas, a identificação de temas de ensino com aproveitamento deficitário, a avaliação longitudinal do desempenho dos alunos, e, de maneira mais ampla, a consistência na execução do projeto pedagógico nas diversas turmas.

A necessidade dessa automação se tornou clara para a disciplina de Probabilidade e Estatística. Provas individualizadas e padronizadas se mostraram uma solução robusta para o sistema de avaliação. Todo material produzido está disponível sob consulta para outras Universidades, podendo ser aplicado diretamente em cursos de Estatística básica ou adaptado para qualquer outra disciplina com provas de múltipla escolha.

1.1 Objetivos

Este projeto modernizou e otimizou o sistema de avaliação dos alunos do curso de Probabilidade e Estatística da UnB, incluindo todas as suas fases: elaboração, impressão, aplicação, correção, divulgação e revisão das provas.

Os objetivos específicos estão elencados a seguir.

- Redução dos custos operacionais, em especial na construção e na correção das provas, por meio do reaproveitamento das questões elaboradas e da automação da correção.
- Construção de um extenso e variado banco de questões (com possível aproveitamento em outras disciplinas de Estatística básica e/ou em cursos semipresenciais ou a distância).
- Estabelecimento de metodologias de análise das respostas dos alunos, via Teoria da Resposta ao Item, para o aperfeiçoamento contínuo do sistema e da qualidade do banco de questões;
- Criação de rotinas computacionais para viabilizar e automatizar as soluções desenvolvidas ao longo do projeto (extensão do pacote *R-exams*);
- Elaboração de relatórios gerenciais para professores e coordenadores da disciplina (vide Capítulo 2);
- Promoção da consistência na execução do projeto pedagógico nas diversas turmas, incluindo a equalização do nível de dificuldade das provas nas diversas turmas;
- Aumento da transparência e da agilidade na divulgação dos resultados (os alunos têm acesso ao espelho e à solução detalhada de sua prova);
- Simplificação da logística de impressão, aplicação e revisão dos exames;
- Redução do risco de fraudes;

1.2 Desenvolvimento

O projeto contou com a participação de dois alunos bolsistas (financiados pelo Programa Aprendizagem para o 3º Milênio (A3M) da UnB) por um período de 12 meses. Ambos trabalharam na construção de um banco com 300 questões de Estatística básica no formato requerido pela ferramenta R/Exams, contendo: enunciado, funções matemáticas que geram as alternativas, descrição detalhada da solução e especificação dos procedimentos de geração dos valores aleatórios (que variam de aluno para aluno). A análise apresentada no Capítulo 2 foi elaborada como parte de um Projeto de iniciação científica (PIBIC), financiado pela FAPDF, e de um Trabalho de Conclusão de Curso (TCC).

O projeto teve também participação de uma equipe de professores (ver detalhes na capa), os quais contribuíram enormemente na proposição de modelos de perguntas e num rigoroso processo de revisão das questões. Além de supervisionar os alunos citados, o Professor Guilherme, coordenador do projeto,

trabalhou na extensão das funcionalidades computacionais da ferramenta R/Exams, a fim de facilitar a operacionalização das provas e automatizar as atividades intermediárias inerentes ao processo.

1.3 Aspectos gerais do sistema

Buscando um sistema de avaliação automatizado, com provas individualizadas, a comissão de planejamento unificado da disciplina de Probabilidade e Estatística optou por adotar como ferramenta primária de avaliação o pacote computacional livre Exams (<http://www.r-exams.org>).

Entre as principais funcionalidades do pacote R-exams está a aleatorização dos parâmetros do enunciado das questões conforme especificação do usuário. Ou seja, números aleatórios são gerados para compor o enunciado do problema. Dessa forma, ainda que a mesma questão apareça em provas diferentes, o aluno deverá compreender todos os passos do raciocínio para chegar ao resultado correto.

Todas as questões do banco foram elaboradasmeticulosamente para permitir essa aleatorização. A especificação do procedimento de amostragem dos parâmetros é uma atividade delicada e determinante para garantir que todas as amostras gerem enunciados que sejam válidos e razoáveis. A definição das funções matemáticas que geram as alternativas também requer cuidado - caso contrário há o risco de haver alternativas iguais ou absurdas.

O processo de revisão do banco também foi igualmente importante para suprimir eventuais falhas de lógica ou de programação. Visando demonstrar a aplicabilidade do conteúdo teórico nas diversas áreas de atuação dos futuros profissionais, a contextualização do enunciado das questões foi fator primordial na construção desse material. O banco atualmente conta com 300 modelos de questão, todos eles escritos em linguagem R, versando sobre 30 tópicos de Estatística básica. Este banco pode ser utilizado em outras disciplinas de Estatística básica e está disponível sob consulta.

Além do banco de questões, o sistema conta com um arcabouço de funções em linguagem R, como extensão ao pacote R-exams, para ampliar as suas funcionalidades e facilitar a interface com o usuário. Os parâmetros que podem ser modificados para atender necessidades específicas são:

- Número de questões na prova;
- Pontuação de cada questão;
- Escolha dos tópicos (dentre 30 opções atualmente disponíveis);
- Número de turmas e quantidade de alunos em cada turma.

Quanto à estrutura de avaliação da disciplina Probabilidade e Estatística na UnB, esta é composta por 3 provas regulares (com pesos 30%, 30% e 40%) e uma prova substitutiva (abrangendo todo o conteúdo do curso) para alunos que desejem repor uma eventual prova não realizada ou queiram substituir a menor nota (entre as provas regulares). Cada avaliação contém 10 questões de múltipla escolha sobre 10 tópicos

pré-estabelecidos. Sendo assim, para cada turma, o sistema desenvolvido sorteia 10 questões do banco (uma de cada tópico). Assim, apesar de turmas diferentes realizarem provas diferentes, a fixação dos tópicos garante uniformidade no conteúdo cobrado nas avaliações. Em princípio, também é possível sortear 10 questões diferentes (ainda uma de cada tópico) para cada aluno do curso, mas considerando a aleatorização do enunciado, essa opção foi considerada inapropriada por inviabilizar a revisão em sala das questões feitas pelos alunos da turma. Alternativamente, poder-se-ia garantir que não há questões que se repetem em mais de uma turma. Tal estratégia também foi considerada inoportuna por comprometer o uso da TRI, uma vez que sem questões repetidas, não seria possível comparar os resultados obtidos nas diferentes turmas.

Após a aplicação do exame, as folhas de resposta dos discentes são escaneadas e então corrigidas automaticamente pelo sistema, o qual, ao identificar o número do documento (prova), executa a correção considerando o gabarito que lhe é próprio. Ao suspeitar de erro de marcação ou de leitura do cartão de respostas, a plataforma solicita ao operador a conferência manual dos dados.

Como subproduto do processo de correção das provas, o sistema produz resultados de forma individualizada. Mais especificamente, cada aluno recebe um documento (em formato *pdf*) contendo uma cópia da sua folha de respostas, o gabarito resumido e a solução detalhada da sua prova específica. Para acessar o arquivo, disponibilizado no Moodle, o aluno usa como senha o número da prova, impresso em seu caderno de questões (que o aluno leva consigo ao final da prova). Considerando a dimensão atual do banco de dados e a aleatorização dos enunciados, não há prejuízo que os alunos tenham acesso a provas antigas.

A avaliação do desempenho dos alunos nesse nível de riqueza e profundidade é inédita no âmbito do Departamento de Estatística da UnB, e possivelmente em toda a Universidade de Brasília.

Maiores detalhes a respeito da utilização do sistema (passo-a-passo) encontram-se no Capítulo 3.

1.4 Descrição dos produtos entregues

Os produtos produzidos nesse projeto estão sumarizados a seguir:

1. Banco de questões

O banco possui 300 modelos de questões de Estatística básica, organizadas em 30 temas. Cada questão contém a descrição do procedimento amostral para geração dos parâmetros aleatórios envolvidos, o enunciado, as funções que geram as alternativas erradas e a solução detalhada da questão.

2. Relatório de análise dos dados (Capítulo 2)

Relatório com análise dos dados produzidos com a aplicação do sistema na disciplina de Probabilidade e Estatística (a partir da leitura dos cartões de respostas).

3. Manual de utilização do sistema (Capítulo 3)

O manual descreve detalhadamente o procedimento de geração e correção das provas utilizando o código disponível em linguagem R.

4. Códigos do sistema

Os códigos desenvolvidos estão disponíveis sob consulta.

Análise de resultados

2º/2019

A mensuração do aprendizado dos estudantes a partir da realização de testes é um aspecto fundamental do sistema educacional. O resultado dos testes fornece a professores, educadores e alunos informações de grande importância sobre o aprendizado dos discentes. Tal medida é utilizada, por exemplo, para determinar se o aluno deve ou não ser aprovado no curso.

Tradicionalmente, pela simplicidade de sua formulação, a Teoria Clássica dos Testes (TCT) é adotada para este fim. Em resumo, nessa abordagem, entende-se que a pontuação do aluno (escore observado) corresponde à soma da pontuação esperada (resultado da *habilidade latente* do aluno) e do erro amostral (que faz com que o desempenho varie caso o aluno seja hipoteticamente submetido a diversas provas sobre o mesmo conteúdo e com o mesmo nível de dificuldade). Ou seja, a nota obtida é um estimador natural da *proficiência* do aluno.

A habilidade latente (um atributo mental não observável do respondente), por sua vez, resulta de diversos fatores, dentre os quais destacamos o potencial cognitivo do aluno, sua dedicação ao curso, seus métodos de estudo e a experiência em sala de aula (composição da turma, estrutura física, horário, didática adotada pelo professor, materiais didáticos utilizados, entre outros fatores).

Recentemente, a Teoria de Resposta ao Item (TRI) vem ganhando espaço, em detrimento da Teoria Clássica dos Testes, sobretudo para testes de grande escala. De acordo com o Ministério da Educação, “No Brasil, a TRI é usada desde 1995 nas provas do Sistema Nacional de Avaliação da Educação Básica (Saeb), que mede o desempenho de estudantes do ensino fundamental e médio. Em 2009, foi usada pelo Enem com o objetivo de garantir a comparação das notas do exame daquele ano com os seguintes.” No exterior, exames como o *Test of English as a foreign language* (TOEFL) adotam a TRI há mais de 50 anos. A TRI é aplicável a diversos campos das Ciências e, para além do campo educacional, se mostrou particularmente útil na Medicina e na Psicometria.

Os modelos de TRI buscam estimar, a partir do ajuste de modelos estatísticos, a probabilidade de um aluno responder corretamente a um item em função de suas características (habilidades) e dos atributos do próprio item. A abordagem probabilística permite que se leve em consideração, entre outros aspectos, o nível de dificuldade e o poder de discriminação do item (uma medida do quanto a probabilidade de acerto varia em função do aumento da habilidade do aluno).

Destacamos que não há intenção de, em um futuro próximo, atribuir menção em função da habilidade estimada via TRI. Os modelos são empregados exclusivamente para a investigação de como o processo avaliativo, e em especial o banco de questões, pode ser aprimorado.

Na Seção 2.1 apresentamos alguns gráficos que ilustram o desempenho dos alunos ao longo do semestre e o efeito da prova substitutiva nas menções. Depois, na Seção 2.2, introduzimos o modelo de TRI adotado e apresentamos os principais resultados obtidos. Por fim, na Seção 2.3, listamos as principais conclusões e

recomendações.

2.1 Análise descritiva dos dados

2.1.1 Desempenho dos alunos

A Tabela 2.1 apresenta um conjunto de estatísticas do curso. Neste estudo, apenas são contabilizados os alunos que concluíram o curso, com ou sem sucesso. As menções apresentadas não correspondem necessariamente à menção efetivamente atribuída pelo professor, o qual tem a prerrogativa de considerar outros critérios, como a entrega de listas de exercícios, a participação do aluno em sala, etc.

A média final (5.8) foi ligeiramente superior à nota média prevista (5.5), especificada na construção dos testes pela função que amostra as questões de cada turma, conforme descrito na Subseção 3.2.6. Isso se deve ao descarte das respostas dos alunos que não concluíram o curso. A nota média das diferentes turmas apresentou pequena variação, diferentemente do que era observado antes da unificação da disciplina, quando as notas médias das turmas variavam amplamente. Maior homogeneidade do resultado entre turmas também foi observada no percentual de aprovados. Entretanto, apenas 4% dos alunos atingiram a menção máxima (SS). Cientes da dificuldade de se obter nota final igual ou superior a 9 com esse sistema de avaliação, mesmo após a prova substitutiva, os professores podem considerar outros aspectos do desempenho dos alunos com nota final próxima de 9.

Tabela 2.1: Estatísticas do curso

Turma	Alunos	Nota final média	% de aprovados	% de SS
AA	47	5.8	76.1	8.1
AB	49	5.5	65.2	8.2
BA	38	5.7	72.9	2.4
BB	46	6.0	77.3	6.2
CA	52	6.2	83.9	10.7
CB	44	5.8	74.6	0.0
CC	49	5.0	63.1	0.0
DA	16	5.4	74.1	0.0
DB	39	5.4	64.1	0.0
EA	49	5.2	64.0	0.0
Total	429	5.8	71.5	4.0

A Figura 2.1 apresenta a nota final de cada aluno da disciplina, discriminada pelo curso do discente.

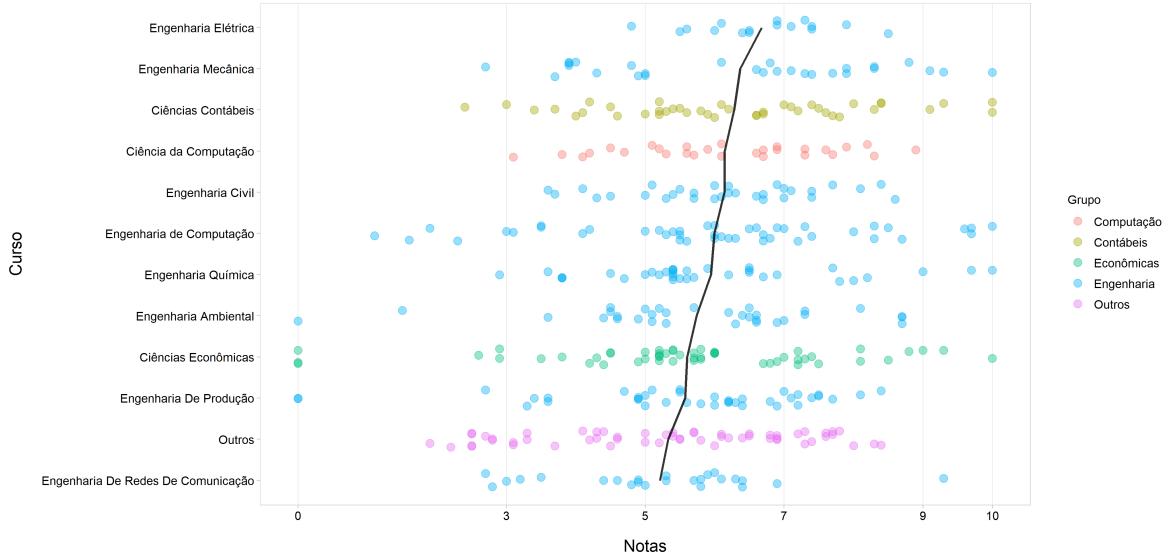


Figura 2.1: Notas finais por curso. Para evitar pontos sobrepostos (quando alunos de um mesmo curso apresentam a mesma nota final), os pontos foram ligeiramente descolados verticalmente (procedimento conhecido como “jitter”). A linha preta indica a média geral de cada curso.

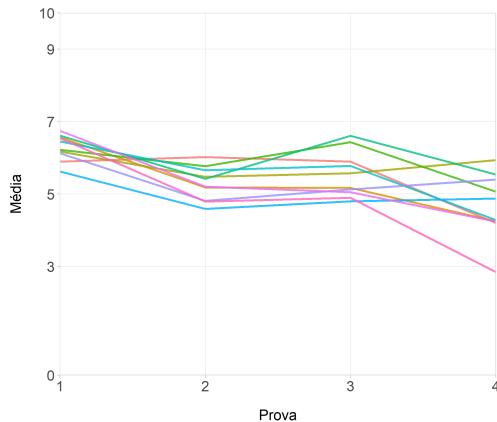
Nota-se que, embora haja certa similaridade entre os cursos, o desempenho dos alunos de Engenharia Elétrica, por exemplo, é substancialmente superior ao de cursos classificados como “Outros”.

A Figura 2.2(a) indica que a Prova 2 é o exame mais difícil ao longo do semestre. Embora as notas da prova substitutiva também sejam mais baixas, cabe ressaltar que isso se deve sobretudo ao fato de que os alunos que se submetem à prova substitutiva, em geral, tiveram desempenho insatisfatório nas provas anteriores. Os painéis (b), (c) e (d) mostram o percentual de acerto de cada tópico das Provas 1, 2 e 3, respectivamente. Ao avaliá-lo com cuidado, os professores podem observar temas cujo desempenho foi superior ou inferior à média. Em que pese o resultado da prova depender em grande medida da composição da turma, temas que se apresentem atípicamente abaixo da média sugerem ao professor avaliar a necessidade de interferir na estratégia de ensino, dedicando mais tempo ao tópico, por exemplo.

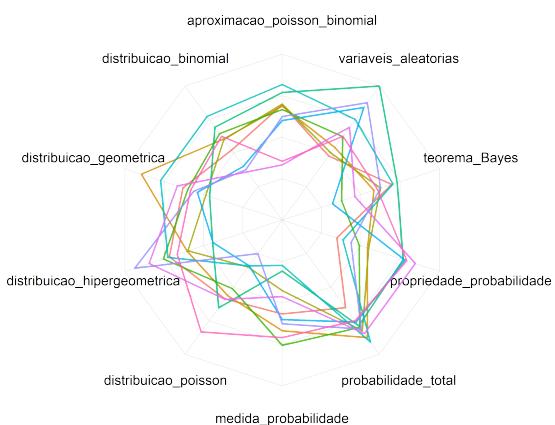
O tema sobre teste de hipóteses para a proporção teve percentual de acerto inferior a 50% em todas as turmas, o que não parece razoável. Sugerimos, portanto, que se investigue as causas deste resultado para que se tome as medidas cabíveis. Caso as questões estejam inapropriadas (voltaremos a esse tema adiante), é preciso melhorar os modelos de questão. Caso contrário, será preciso buscar maneiras de aperfeiçoar o ensino nessa frente. Neste contexto, cabe ressaltar que este é normalmente o último tópico ministrado no curso.

2.1.2 Prova substitutiva

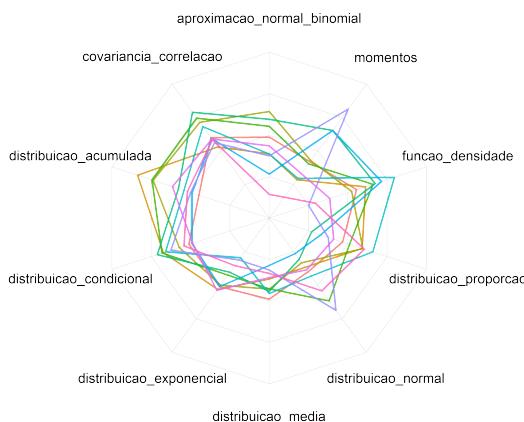
Conforme discutido na Seção 1.3, os alunos são autorizados a realizar uma prova substitutiva no final do semestre, versando sobre todo o conteúdo do curso. A nota da prova substitutiva substitui a menor nota,



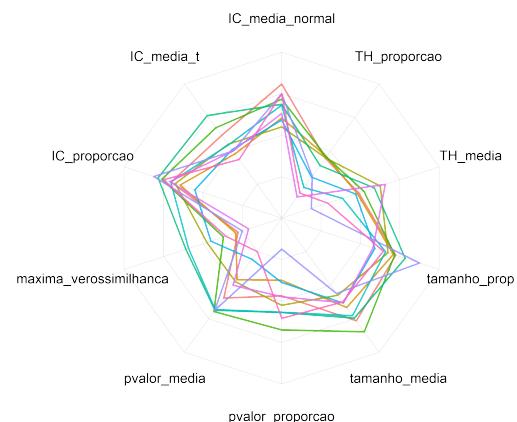
(a) Nota média por turmas e por prova.



(b) Proporção de acertos - Prova 1.



(c) Proporção de acertos - Prova 2.

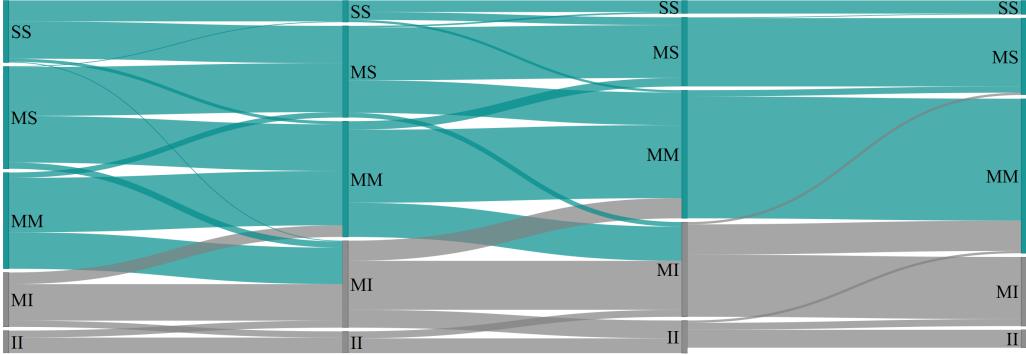


(d) Proporção de acertos - Prova 3.

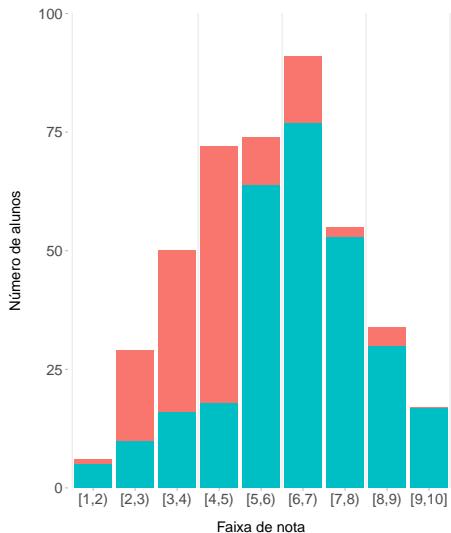
Figura 2.2: (a) Nota final média dos alunos por turma e número da prova. (b), (c) e (d) apresentam, respectivamente, o percentual de acerto por tema nas provas 1, 2 e 3. O centro da figura representa 0% de acerto, enquanto a borda mais externa indica 100%. Questões anuladas estão representadas pela média de acerto no respectivo tema.

desde que a nota final não diminua (por isso ela é conhecida como “sub do bem”).

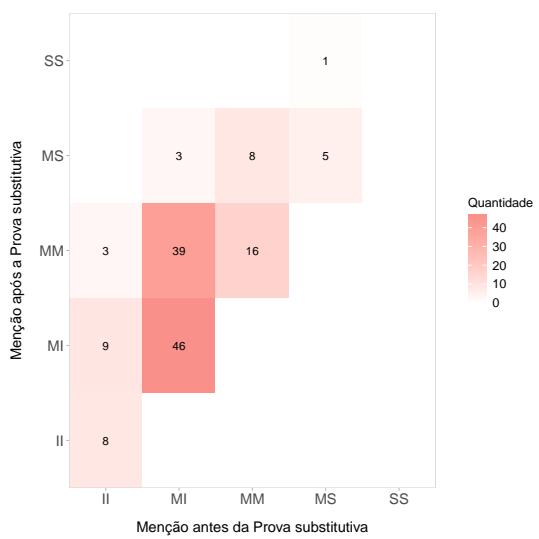
A Figura 2.3 traz um resumo visual dos efeitos da prova substitutiva. Primeiramente, no painel (a), apresentamos o fluxo acumulado parcial das menções ao longo do curso. Observa-se que a mobilidade entre as menções é limitada. A menção máxima, por exemplo, é atingida quase exclusivamente por aqueles que tiveram alto desempenho desde a primeira prova. A figura (b) retrata o número de alunos com nota final nas respectivas faixas, com a cor atribuída de acordo com a participação do aluno na prova substitutiva. Ao todo, 138 fizeram a prova substitutiva, dos quais 78.3% estavam reprovados. Destes, 41.7% conseguiram a aprovação.



(a) Fluxo das menções.



(b) Número de alunos por faixa (após a Prova 3).



(c) Menção antes e depois da Prova substitutiva.

Figura 2.3: (a) Fluxo das menções parciais acumuladas. As quatro barras verticais indicam, na ordem, as 4 provas realizadas no curso. Na Prova 1, temos a menção correspondente à nota da própria prova; na Prova 2, calculou-se a menção parcial acumulada a partir da média aritmética das duas primeiras notas; na Prova 3, a nota final foi definida de acordo com os pesos atribuídos na ementa; na Prova substitutiva, tem-se a menção final obtida pelo aluno. Considerou-se apenas alunos que realizaram as 3 primeiras provas. (b) Número de alunos por faixa de nota final após a Prova 3. (c) Matriz com o número de alunos por menção, dada a situação antes e depois da Prova substitutiva.

2.2 Análise via Teoria da Resposta ao Item

A simples visualização dos dados brutos já é de grande valia. Entretanto, para responder questões mais desafiadoras, lançamos mão de técnicas de modelagem estatística de dados. Entre outras vantagens, o ajuste de modelos de TRI nos permite traçar um diagnóstico mais completo do banco de questões. Enquanto o percentual de acerto nos dá indícios quanto ao nível de dificuldade da questão, ele pouco nos informa sobre a qualidade das alternativas erradas (se são facilmente descartadas por alunos de baixa habilidade) e sobre o poder discriminante do item. Além de evitar questões demasiadamente fáceis ou difíceis, queremos que a habilidade latente do aluno esteja fortemente associada à probabilidade de acerto da questão. Caso contrário podemos concluir que o item é pouco informativo para a habilidade do aluno, o que é o torna sem valor prático. O percentual bruto de acerto também não leva em consideração quais alunos responderam ao item. Se ele foi sorteado para uma turma excelente, por exemplo, um alto valor levaria à falsa impressão de que se trata de uma questão fácil.

Nesse primeiro momento, ajustamos quatro modelos (um para cada prova), todos com a mesma estrutura, definida a seguir.

$$\begin{aligned}
 Y_{ij} &\sim \text{Ber}(p_{ij}) \\
 p_{ij} &= c_j + \frac{1 - c_j}{1 + \exp(-(a_j\theta_i - b_j))} \\
 a_j &\sim \text{Half-normal}(0, 1) \\
 b_j &\sim N(0, 1) \\
 c_j &\sim \text{Beta}(5, 17) \\
 \theta_i &\sim N(0, 1),
 \end{aligned}$$

onde Y_{ij} representa se o aluno i acertou (1) ou errou (0) a questão j , p_{ij} a probabilidade do aluno i acertar a questão j , c_j a probabilidade de acerto de um aluno sem proficiência, a_j o poder discriminante da questão j , b_j a dificuldade do item j e θ_i a habilidade latente do aluno i .

Note que, com este conjunto de modelos, precisamos estimar $3 * 228 + 3 * 435 = 1989$ parâmetros valendo-se das 13779 respostas (pares alunos/questões), o que não é tarefa fácil. Para tanto, usamos o pacote *RStan*, do R, para gerar amostras da distribuição a posteriori.

Assumimos que a aleatorização dos números nos enunciados não altera os atributos (parâmetros) das questões. Além disso, para qualquer par de questões, dada a habilidade do aluno e os parâmetros das questões, os resultados (acertou/errou) são independentes um do outro.

2.2.1 Resultados preliminares

Um dos desafios na construção dos testes é garantir certo equilíbrio entre as turmas. Caso as provas sejam mal calibradas, a nota esperada de um dado aluno pode ser muito diferente dependendo da turma em que ele esteja matriculado. Não é razoável, por exemplo, que a nota dependa mais das questões que foram sorteadas para sua turma do que de sua proficiência.

A Figura 2.4(a) apresenta a nota média estimada de um aluno mediano (com habilidade maior que exatamente 50% dos alunos) nas diversas turmas. Notadamente, quando se observa uma prova específica, a variação entre as turmas é apreciável. Entretanto, quando se avalia a dificuldade média nas 4 provas, conclui-se que a variação entre as turmas é bastante discreta.

Na Figura 2.4(c), apresentamos a *informação* dos testes, uma medida que indica o quanto as respostas “desvendam” a habilidade não-observável do aluno. Além disso, tal medida esclarece em qual intervalo de habilidade as respostas são mais informativas. Se, por um lado, as turmas parecem estar bem balanceadas no que diz respeito ao nível de dificuldade das questões, por outro, percebe-se que algumas provas são menos informativas que outras. Nesses casos, mesmo que o aluno tenha respondido 40 itens ao longo do curso, a incerteza a respeito de sua proficiência será alta, prejudicando o processo avaliativo. Nesse sentido, sugerimos que, no momento do sorteio das questões, se verifique também se as provas estão igualmente informativas.

A Figura 2.4(b) compara a avaliação adotada pela ementa com o que obtemos, mantidas as quantidades de alunos em cada menção, ao ordenar os alunos pela média ponderada das habilidades estimadas. Enquanto a maior parte dos discentes não teria sua menção alterada, há diversos alunos que passariam da condição de reprovado para aprovado.

A diferença se deve, sobretudo, a dois fatores. Primeiro, independentemente do método, há um intervalo de incerteza em torno da estimativa pontual. Por exemplo, quando o modelo estima uma habilidade igual a um, ele também entende que a habilidade poderia ser de 0.9 ou 1.1, digamos. Daí, alunos que estejam no limite das menções acabam eventualmente mudando de faixa dependendo do método adotado. A incerteza pode ser reduzida aumentando-se o número de questões, mas isso tende a aumentar os custos operacionais. De todo modo, isso vale para lembrar que o acaso é um componente não-negligenciável da nota do aluno, e que outros critérios além da nota podem ser adotados.

Segundo, na TRI leva-se em conta exatamente quais questões foram acertadas, e não apenas o número total de acertos. Por isso, o escore estimado tende a ser maior quando o aluno acerta mais questões difíceis ou com maior poder de discriminação. Considere o seguinte exemplo. Suponha que uma das questões da prova de Estatística seja sobre técnicas de fotografia. Nesse caso, a habilidade do aluno em Estatística não alteraria a probabilidade do acerto deste item. Por isso, o modelo de TRI concluiria que o item é

pouco informativo para a habilidade ($a_j \approx 0$). Daí, o acerto ou erro do item pouco alteraria a habilidade estimada pelo modelo. Entretanto, na forma corrente de avaliação, o acerto do item aumentaria em 1 a nota do aluno, da mesma forma que ocorreria com qualquer uma das outras questões da prova.

A Figura 2.5(a) apresenta o correlograma dos *resíduos* do modelo (medida de discrepância entre o valor previsto pelo modelo, p_{ij} e o valor observado Y_{ij}) por tema. Visto que o modelo assume implicitamente que $Y_{ij}|p_{ij}$ e $Y_{lm}|p_{lm}$, são independentes, para qualquer $i \neq l$ e $j \neq m$, é esperado (pelo modelo) que as correlações sejam próximas a zero. Entretanto, como a referida suposição não é plenamente válida, observa-se algumas regiões de coloração mais intensa. Note que a correlação se dá entre os resíduos, e não sobre os temas diretamente. Por isso, não se pode concluir, por exemplo, que os temas *Distribuição normal* e *Intervalo de confiança para a média* são independentes entre si, mas sim que a informação de quanto o modelo “errou” no primeiro tema não está probabilisticamente associado ao quanto o modelo irá errar no segundo tema.

Baseado no correlograma, construímos uma rede de associação entre os resíduos. O resultado está disponível na Figura 2.5(b). Os temas da Prova 1 estão separados por uma linha imaginária dos temas da Prova 3, indicando que, para essas provas, os temas estão mais relacionados entre si. Cabe esclarecer, entretanto, que a distância entre os pontos não corresponde exatamente ao valor da correlação do par.

As questões do tema *Estimação de máxima verossimilhança* sabidamente envolvem procedimentos e entendimentos de certo modo descolados dos demais temas da Prova. No entanto, o mesmo não pode ser dito sobre o tema de *Teste de hipóteses para a proporção*. Por isso, novamente sugerimos que se investigue as os motivos que tornaram o comportamento desse tema contraintuitivo.

2.2.2 Avaliação das questões

Nesta seção, apresentamos um estudo da qualidade das questões, de acordo com o ajuste via TRI.

A figura 2.6 nos indica a probabilidade de um aluno mediano acertar cada questão aplicada durante o semestre. Temas com maior variabilidade entre as probabilidades são justamente aqueles em que os modelos de questão são mais heterogêneos.

As questões de alguns temas são notadamente mais fáceis. Por exemplo, na Prova 1, percebe-se que os pontos referentes ao tema *lei da probabilidade total* etão mais à direita, o que era de se esperar. Na Prova 3, os temas mais fáceis são os que tratam de tamanho da amostra e intervalo de confiança calculados a partir da tabela da distribuição Normal. Nesses casos, mesmo que o aluno não entenda plenamente os conceitos estatísticos, ele poderia acertar as questões desde de que identificasse corretamente a formula que deveria aplicar. Conforme visto anteriormente, o tema *Teste de hipóteses para a proporção* foi o mais difícil. Entre as possíveis explicações está o nível de dificuldade das questões, o fato de este ser o último

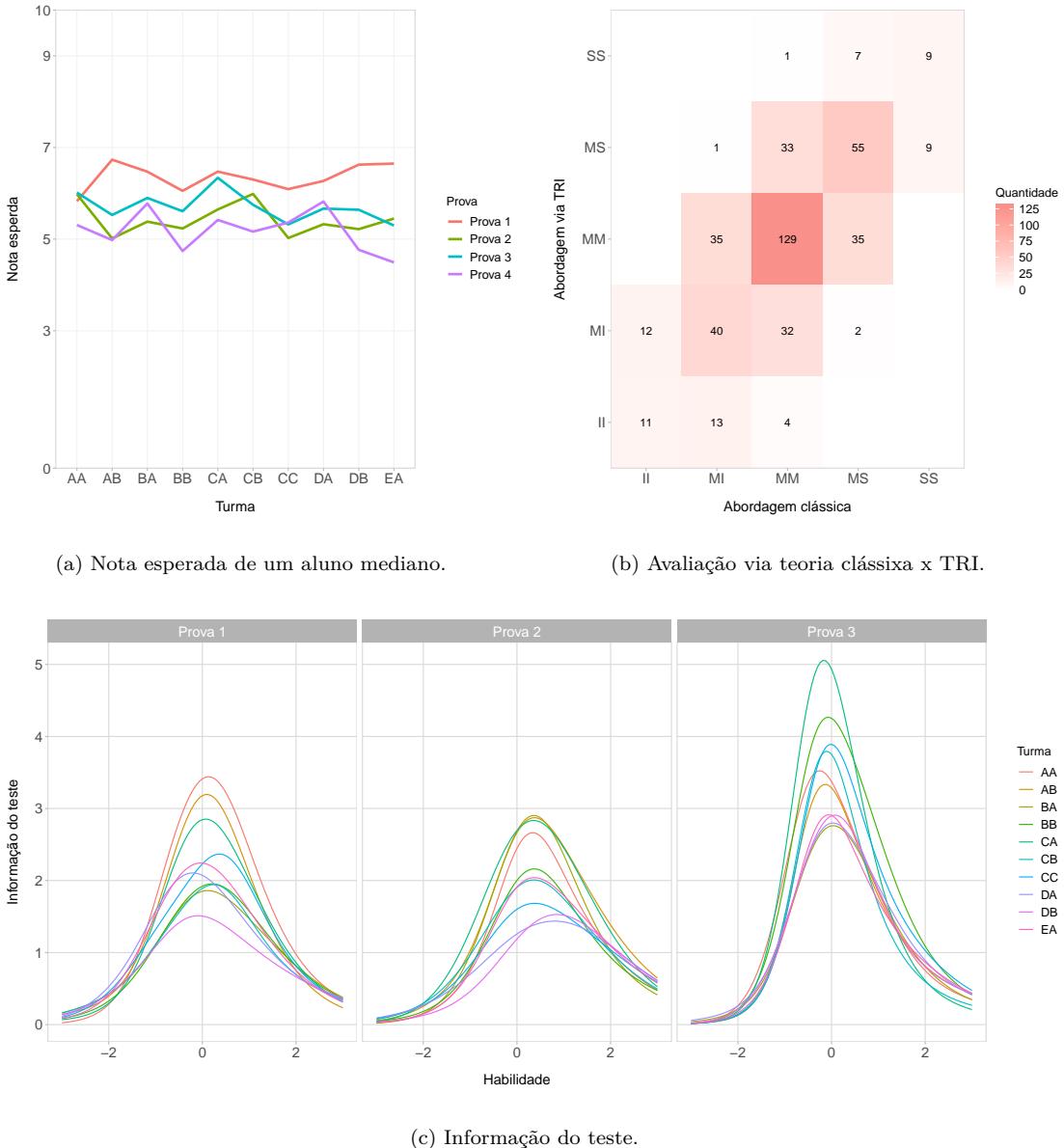
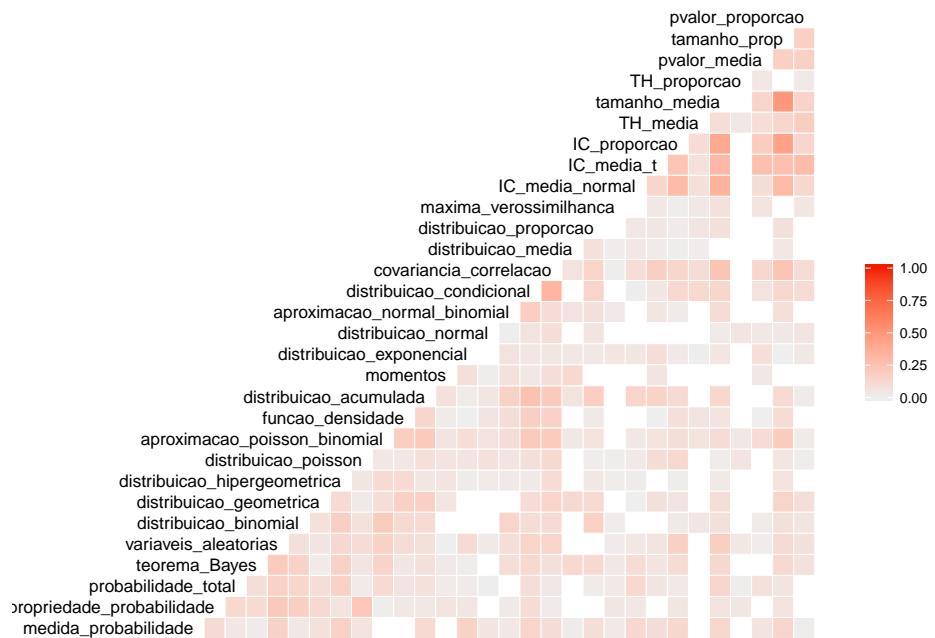
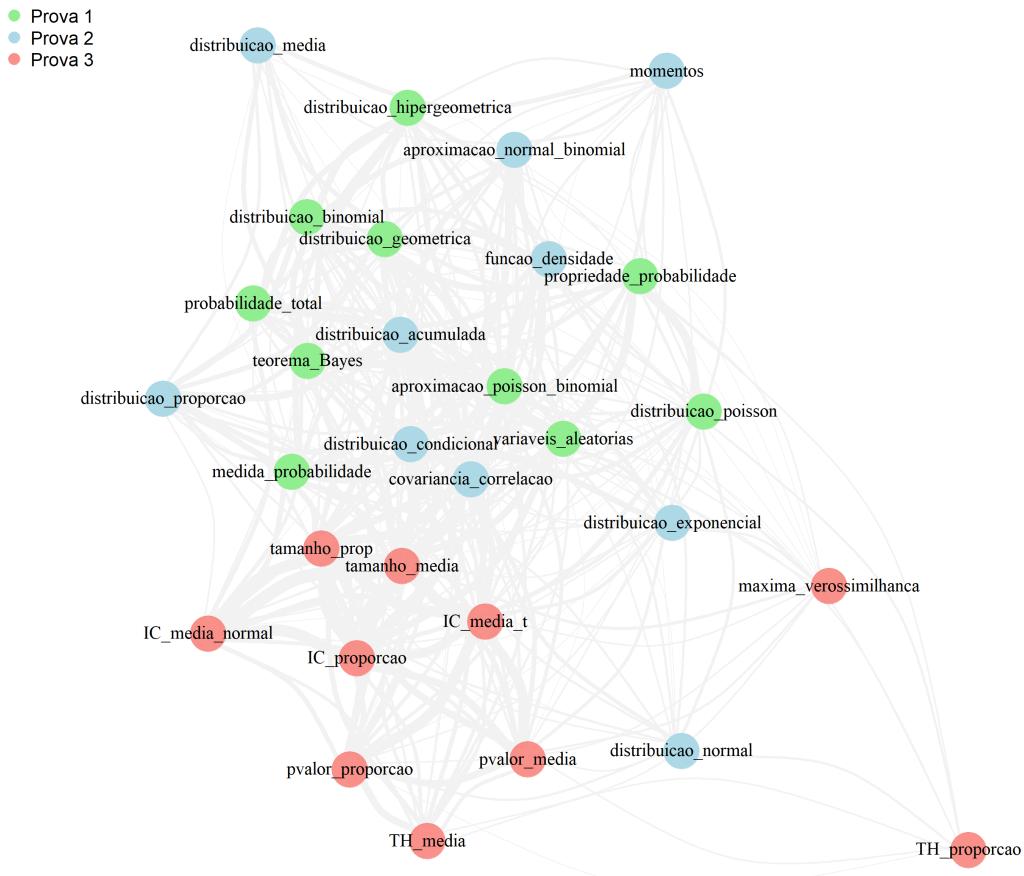


Figura 2.4: (a) Nota esperada de um aluno com habilidade mediana (isso é, com $\theta = 0$). (b) Número de alunos em cada menção, calculando-a com base na nota final observada (menção real) e na habilidade final estimada via TRI (definida como a média ponderada das habilidades estimadas nas provas, usando-se os mesmos pesos e protocolos da prova substitutiva). Para as menções via TRI, ordenamos e dividimos os alunos de modo a manter os percentuais observado nas menções reais. (c) Informação do teste por turma e por número da prova.



(a) Correlograma dos resíduos por tema.



(b) Grafo de associação dos resíduos.

Figura 2.5: (a) Correlograma (de Pearson) dos *resíduos componente do desvio* (RCD) por tema. Valores negativos foram substituídos por zero. (b) Rede de associação dos resíduos por tema.

tema ministrado no curso ou alguma deficiência no processo de ensino. Os dados apresentados neste gráfico também estão disponíveis nas tabelas do Apêndice A.4.

As Figuras 2.7 a 2.9 representam as curvas características das questões aplicadas nas provas 1 a 3. Quando a probabilidade de acerto do aluno mais habilidoso (habilidade=3) é inferior a 0.8 ou a probabilidade de acerto do aluno menos habilidoso (habilidade=-3) é superior a 0.4, sugerimos que a questão seja revisada. Tais questões estão representadas por linhas tracejadas. A leitura cuidadosa das curvas características indicam a origem do problema. Se a curva apresentar pouca variação, o item não está fortemente associado à habilidade do aluno (o parâmetro a é baixo). Se ela começar de uma posição elevada, um aluno pouco proficiente é capaz de deduzir a resposta correta com probabilidade alta. Por fim, quando a curva parece estar horizontalmente deslocada para direita (esquerda), a questão é muito fácil (difícil).

Na tabela 2.2, identificamos as questões que devem ser revisadas. A tabela completa (com todas as questões aplicadas) está disponível no Apêndice A.4.

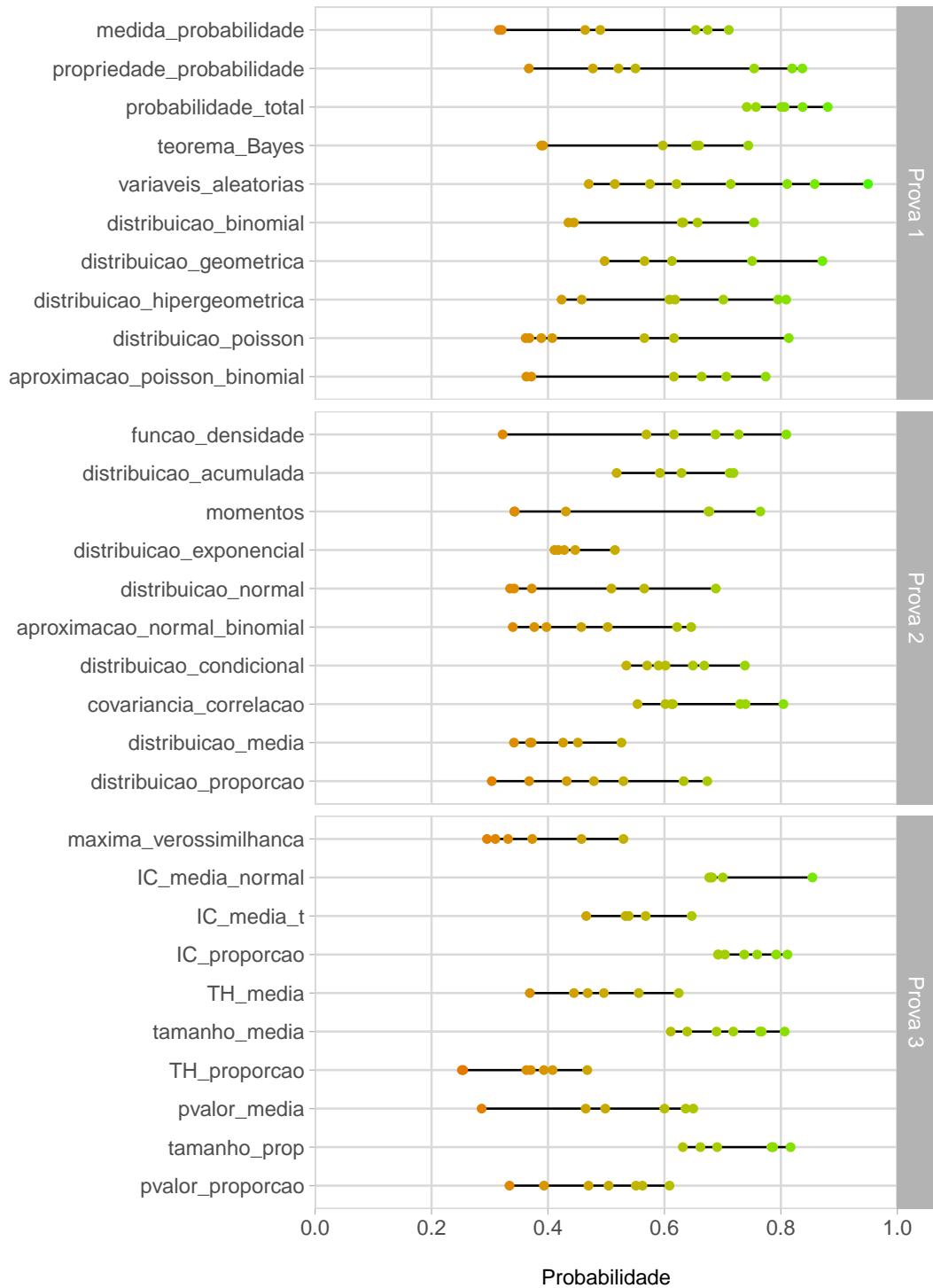


Figura 2.6: Probabilidade de um aluno mediano (habilidade igual a zero) responder corretamente cada questão do banco que tenha sido sorteada a pelo menos uma das turmas nas provas 1 a 3.

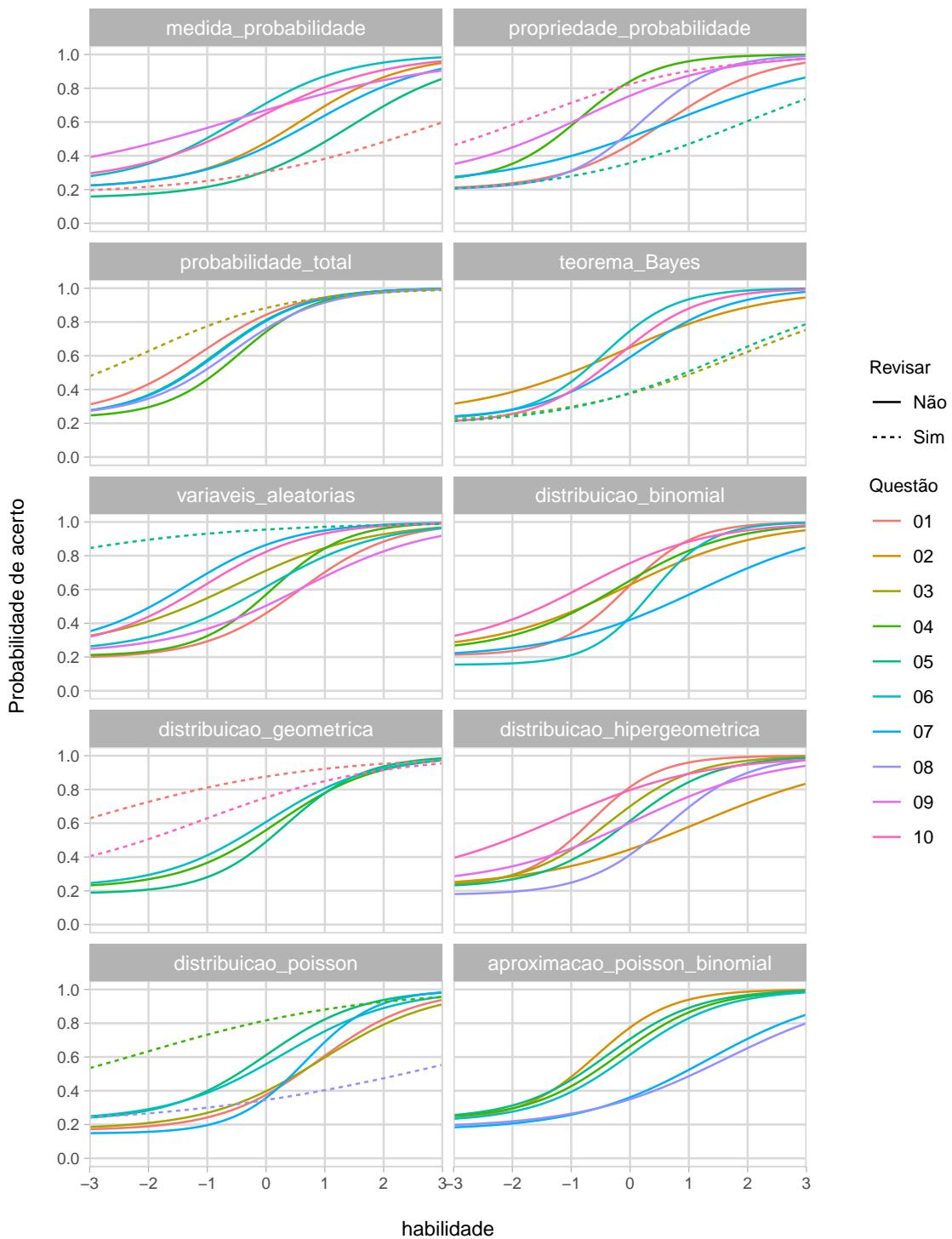


Figura 2.7: Curva característica dos itens (indica a probabilidade de um aluno com dada habilidade acertar ao respectivo item) da Prova 1.

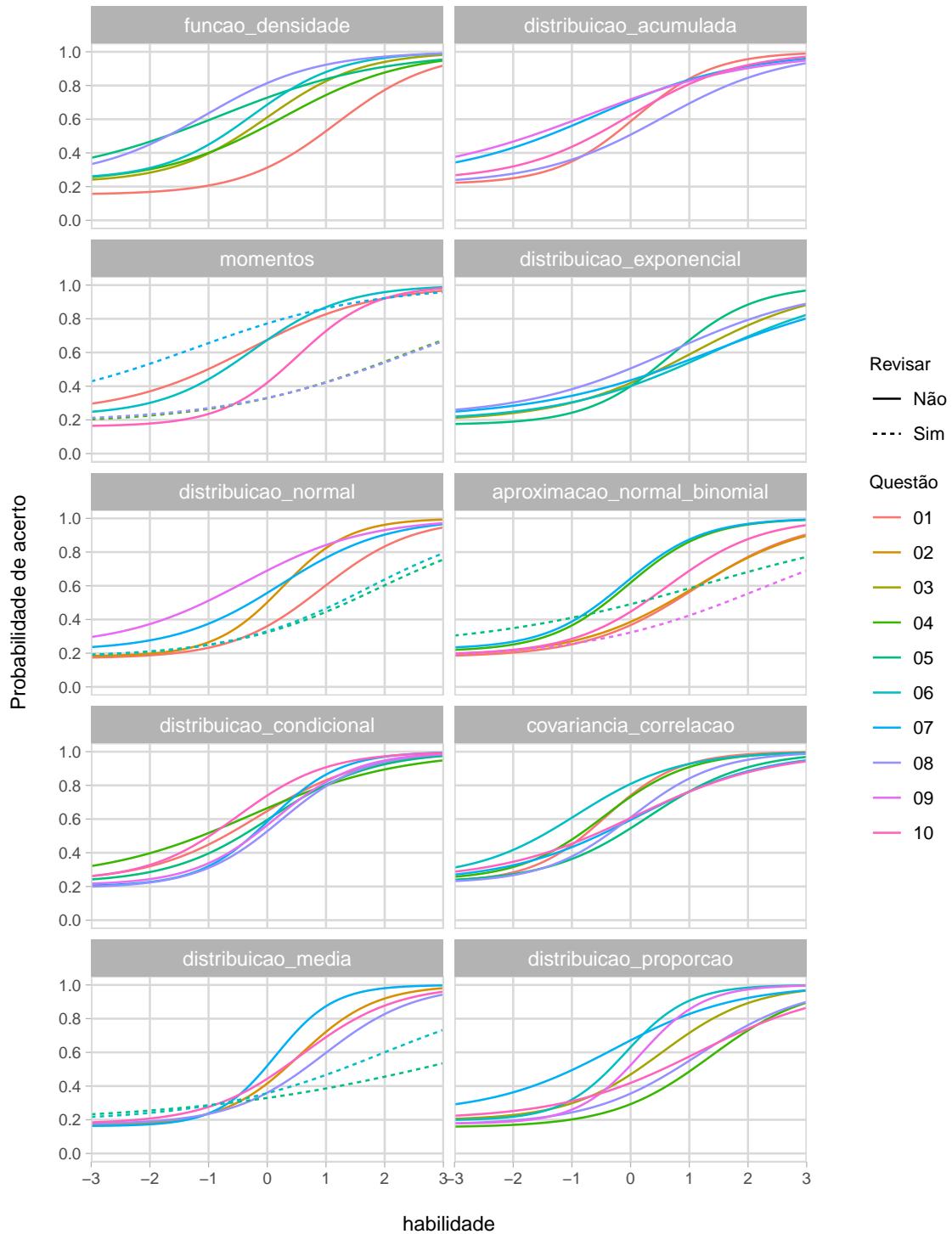


Figura 2.8: Curva característica dos itens (indica a probabilidade de um aluno com dada habilidade acertar ao respectivo item) da Prova 2.

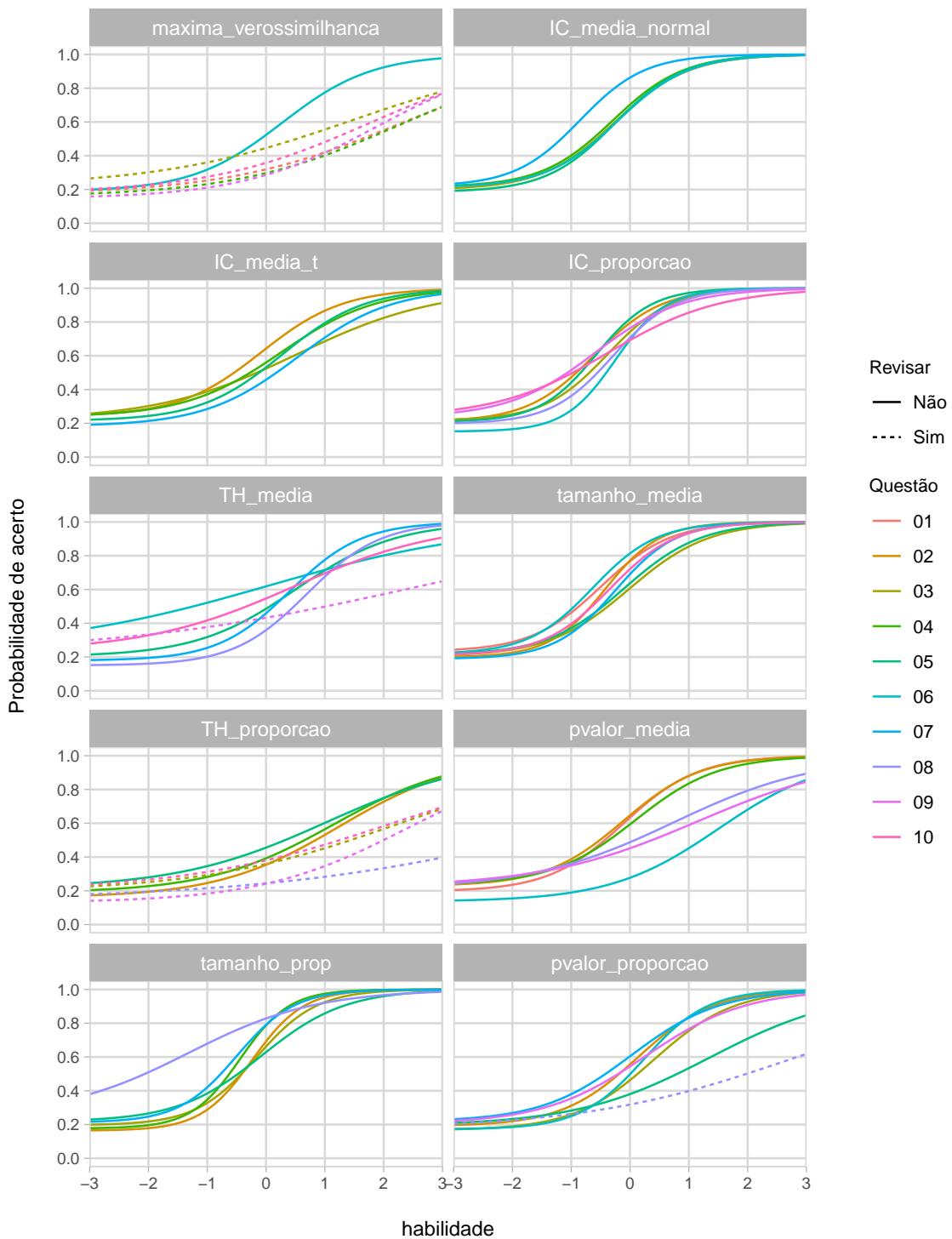


Figura 2.9: Curva característica dos itens (indica a probabilidade de um aluno com dada habilidade acertar ao respectivo item) da Prova 3.

Tabela 2.2: Parâmetros das questão a revisar. n:número de respondentes; p1, p2 e p3: probabilidade de acerto de um aluno com habilidade -3 (mínima), 0 (mediana) e 3 (máxima).

Tema	Q	a	b	c	n	% acerto	p1	p2	p3
medida_probabilidade	01	0.6	1.6	0.17	44	27	20	32	60
propriedade_probabilidade	05	0.7	1.3	0.18	90	37	21	37	74
propriedade_probabilidade	10	0.7	-1.2	0.25	39	85	46	82	97
probabilidade_total	03	0.8	-1.7	0.25	132	87	48	88	99
teorema_Bayes	03	0.7	1.3	0.20	89	38	23	39	75
teorema_Bayes	05	0.7	1.2	0.19	45	38	21	39	79
variaveis_aleatorias	05	0.5	-2.7	0.27	52	100	84	95	99
distribuicao_geometrica	01	0.5	-1.6	0.26	48	90	63	87	97
distribuicao_geometrica	10	0.7	-0.7	0.24	44	77	40	75	95
distribuicao_poisson	04	0.5	-1.1	0.25	48	83	53	81	96
distribuicao_poisson	08	0.4	1.4	0.19	16	25	24	36	55
momentos	04	0.6	1.5	0.18	44	30	20	34	67
momentos	07	0.7	-0.8	0.25	16	81	43	76	96
momentos	08	0.6	1.5	0.19	96	34	21	34	67
distribuicao_normal	05	0.8	1.5	0.18	133	34	19	33	76
distribuicao_normal	06	0.9	1.5	0.18	94	33	19	34	79
aproximacao_normal_binomial	05	0.5	0.7	0.23	63	51	30	50	77
aproximacao_normal_binomial	09	0.7	1.5	0.17	14	14	20	34	69
distribuicao_media	05	0.4	1.5	0.19	88	32	23	34	54
distribuicao_media	06	0.7	1.3	0.19	111	37	22	37	73
maxima_verossimilhanca	01	0.7	1.5	0.17	48	29	19	33	69
maxima_verossimilhanca	03	0.6	0.9	0.22	141	46	27	46	78
maxima_verossimilhanca	04	0.7	1.6	0.16	47	28	18	31	69
maxima_verossimilhanca	09	0.9	1.6	0.15	84	30	16	30	76
maxima_verossimilhanca	10	0.7	1.3	0.18	16	25	20	37	77
TH_media	09	0.4	1.0	0.22	178	44	30	44	65
TH_proporcao	03	0.6	1.4	0.19	131	37	23	37	69
TH_proporcao	08	0.4	2.1	0.15	48	19	18	25	40
TH_proporcao	09	0.8	1.9	0.13	38	16	14	25	67
TH_proporcao	10	0.6	1.2	0.20	51	39	24	39	70
pvalor_proporcao	08	0.6	1.6	0.18	16	19	20	33	62

2.3 Conclusões, recomendações e trabalhos futuros

Neste capítulo investigamos em detalhes o desempenho dos alunos e a qualidade das questões criadas nesse projeto. Para tanto, usamos a Teoria de Resposta ao item, o que, no limite do nosso conhecimento, ainda não havia sido empregada para auxiliar a gestão do processo avaliativo em cursos oferecidos pela UnB.

O percentual de aprovação está compatível com o observado historicamente em disciplinas em cursos de Ciências Exatas. A variação entre turmas, por sua vez, teve uma redução substancial depois da unificação do curso (a análise dos dados históricos poderá ser oportunamente incluída nesse relatório).

O percentual de alunos com menção máxima, SS, sem considerar possíveis ajustes ou arredondamentos, se mostrou demasiadamente baixo, o que sugere a adoção de medidas que o aumentem.

Ao avaliar os dados da prova substitutiva, observamos que 78.3% dos alunos que a fizeram estavam reprovados. Destes, 41.7% lograram aprovação. Apenas 1 aluno conseguiu chegar ao SS após a prova substitutiva.

Independentemente do método de avaliação (clássico ou via TRI), é preciso considerar o componente aleatório dos testes. A título de ilustração, um aluno com 8.8, submetido a outra prova, poderia, por mero acaso, alcançar 9.2.

Apresentamos gráficos que permitem ao professor identificar temas em que os alunos tiveram desempenho deficitário. Tal informação é útil para o planejamento do curso no próximo semestre.

Os dados indicam que, pelo procedimento de sorteio das questões para as turmas, as provas estão bem equilibradas em relação ao nível de dificuldade, mas não quanto ao nível de *informação*. Nesse sentido, sugerimos que se sorteie as questões de modo que o nível de informação dos testes seja parecido em todas as turmas.

Conforme detalhado na subseção 2.2.2, diversas questões deverão ser revisadas, seja por estarem descalibradas no nível de dificuldade, por serem pouco informativas ou por conterem alternativas falsas facilmente detectáveis.

A análise via TRI pode ser aprimorada de diversas maneiras. Por exemplo, ao invés de considerar um modelo independente por prova, seria preferível ajustar um único modelo capaz de acomodar a natureza temporal dos dados. Canalizar a distribuição a posteriori dos parâmetros das questões do semestre anterior para especificar a priori dos parâmetros no semestre corrente traria ganhos ao modelo. Em outras palavras, se em um dado semestre constatarmos que determinado item é difícil, devemos nos valer dessa informação para melhorar a precisão de nossas inferências.

Manual de procedimientos

Este capítulo descreve o passo-a-passo necessário para gerar e corrigir provas de uma disciplina utilizando o sistema elaborado. Este foi desenvolvido para uso no Windows, e NÃO funciona no macOS. Para usá-las em outro sistema operacional, é preciso editar os códigos R. Para correta utilização do produto, é necessário que o usuário tenha noções básicas de R. A disciplina de Probabilidade e Estatística (PE) será usada como referência para facilitar a exposição dos passos, mas o sistema também pode ser aplicado a outras disciplinas, conforme detalhado nas próximas seções.

3.1 Pré-requisitos

Os pré-requisitos necessários para utilização do sistema estão descritos a seguir.

3.1.1 Programas

Para executar os códigos será necessária a instalação dos seguintes programas (todos de livre acesso):

- RStudio (versão utilizada: 3.5.1);
- Pacote R exams (versão utilizada: 2.3-2);
- wkhtmltopdf (versão utilizada: 0.12.5);
- Rtools (versão utilizada: Rtools version 3.5.0.4. Na instalação, marcar a caixa “add app directory to the system path”);
- PDFTk (versão utilizada: 2.0.2);
- GhostScript (versão utilizada: 9.50);
- ImageMagick’s convert (com a opção legacy convert.exe). Veja a Seção “Details” do help da função *nops_scan()* do R.

3.1.2 Pastas e arquivos

Para facilitar a apresentação do sistema e auxiliar a organização dos arquivos, é necessário criar algumas pastas no computador. Este manual utilizará a disciplina Probabilidade e Estatística (PE) como referência, portanto uma pasta chamada “PE” será criada, contendo as seguintes subpastas:

- i. Banco_questoes
- ii. Suplementos
- iii. Prova_1
 - Cadastro dos alunos

A seguir tem-se a descrição do conteúdo necessário em cada uma das subpastas citadas acima.

- i. A pasta “Banco_questoes” deve conter as questões do banco em formato .Rnw e a planilha das respectivas dificuldades “matriz.dificuldades.csv” (mais detalhes desta planilha na seção 1.2.4).

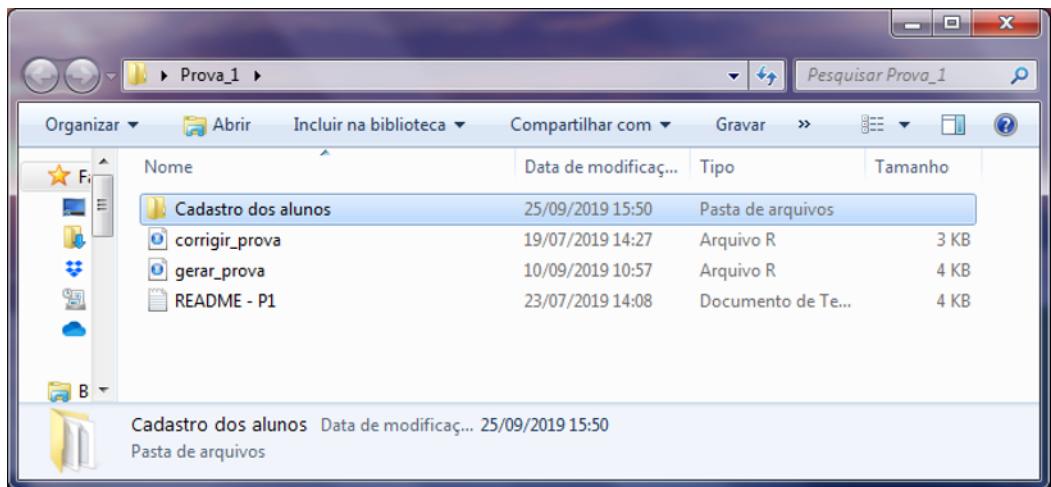


Figura 3.1: Ilustração da pasta “Prova_1”.

ii. A pasta “Suplementos” deve conter os seguintes arquivos:

- “info.tex” (arquivo *.tex* com as instruções impressas na contracapa da prova).
- “Funcoes_Extras.R” (arquivo contendo as funções que complementam o pacote *exams*).
- “Folha_branco.pdf” (arquivo pdf com apenas uma folha em branco).

iii. Além disso, a pasta “Prova_1” deve conter os seguintes arquivos:

- gerar_prova.R: arquivo utilizado na geração das provas.
- corrigir_prova.R : arquivo utilizado na correção das provas.
- README - P1.txt : arquivo de registro dos acontecimentos relevantes durante o processo de correção das provas.
- Para gerar as provas da disciplina de Probabilidade e Estatística (PE) na UnB, a pasta “Cadastro dos alunos” deve conter 2 arquivos com os dados atuais dos alunos matriculados em PE:
 - “Composição de Turmas PE.txt”: solicitar à secretaria do departamento.
 - “PE.txt”: solicitar à secretaria do departamento.

Além disso, a pasta “Cadastro dos alunos” deve conter o dicionário de códigos de cursos, que não precisa ser atualizado semestralmente.

- “Código de opções de curso.txt”.

Observação: Para gerar provas em outras instituições, a pasta “Cadastro dos alunos” pode ficar em branco nesse momento. **Atenção!** Os 3 arquivos devem estar salvos com o encoding UTF-8.

Segue ilustração da organização das pastas:

Universidade de Brasília		Página:	1 de	18
Secretaria de Administração Acadêmica		Emissão:	29/08/2019	14:34
Composição de Turmas				
Período: 2019/2				
Nível: Graduação				
Orgão: Departamento de Estatística				
Disciplina: EST 115045 Probabilidade e Estatística				
Turma: AA				
Matrícula / Opção		Nome	Menção	
17/016035 - 3221		Amanda Brazil Suffert		
19/0010070 - 3221		Ana Carolina Mendes de Oliveira		
19/0024186 - 3221		Ana Clara Rocha Santos		
19/0010665 - 3221		Anderson da Silveira Farias		
19/0042184 - 3221		Augusto Soares Corrêa Oliveira		
19/0042184 - 3221		Beatriz Silva de Araújo		
18/0140084 - 3221		Bruno Costa Santos Gonçalves		
17/016040 - 3221		Cecília Gómez Lacerda Lavenere Bastos		
19/0011289 - 1341		Caio Tomás de Paula		
19/0026090 - 3221		Cesar Freitas Albuquerque		
19/0010665 - 3221		Cláudia Ferreira Souza		
19/0026308 - 6353		Daniel Alexandre Coentreira Junior		
19/0027291 - 3221		Enzo Vinícius Costa Pessoa		
18/0052608 - 3221		Ernani Santos Prospero e Silva		
19/0010665 - 3221		Fernando César da Costa		
18/0017233 - 6220		Gabriela Pereira do Nascimento		
16/0007640 - 3221		Guilherme da Castro e Souza		
19/0010665 - 3221		Gustavo Oliveira Oliveira Castro		
19/0042796 - 3221		Gustavo Coelho		
16/0029864 - 3221		Gustavo Neris Nascimento		
18/0048462 - 3221		Idiana Silveira Tomazelli		
19/0042796 - 3221		Isadora Oliveira Souza		
19/0042974 - 1325		Isadora Silva Teles		
19/0057661 - 1341		Iuri Mateus da Silva Santos		
19/0010665 - 3221		Jeanne Ferreira da Costa		
14/0022988 - 1368		Jelson Tavares da Costa		
18/0113674 - 3221		João Emanuel Nascimento de Paula Rodrigues		
18/0113682 - 60017		João Felipe Castro dos Santos		
19/0010665 - 3221		Jonas Henrique da Cunha Góes		
19/0057840 - 60844		Jose Felipe de Aguiar da Silva		
19/0031310 - 3221		Julia Braga Silveira Mendes		
17/0130886 - 6335		Julia de Lima Sousa		

Figura 3.2: Ilustração do arquivo “Composição de Turmas PE.txt”.

PE - Bloco de notas						
Atenção	Código da Disciplina	Nome da Disciplina	Créditos	Turma	Data da Emissão	Período
	115045	Probabilidade e Estatística	002-002-000	006 AA	02/09/2019	2019/2
Alunos que não constem nesta Relação deverão ter sua situação regularizada na SAA, sob pena						
Nome do Departamento Departamento de Estatística						
Aulas Ministradas – Rubrica dos						
Número da Aula						
Falta	Alunos	Data da				
Número	%	Matrícula	Nome			
17/016035		AMANDA BRAZIL SUFFERT				
		19/0010070	ANA CAROLINA MENDES DE OLIVEIRA			
		19/0024186	ANA CLARA ROCHA SANTOS			

Figura 3.3: Ilustração do arquivo “PE.txt”.

Código de opções de curso - Bloco de notas

Arquivo Editar Formatar Exibir Ajuda

universidade de Brasília
secretaria de Administração Acadêmica
Cursos / Opções : Créditos - Graduação
Período de Referência: 2018/1

Campus	Cursos	Opções	Per. Início	Per. Fim:	Rec- onhecido:	Dura- ção:	Área Acesso:	Límite de créd. por Per. Min	Total de Credito:	Crédi DC
FAC	Darcy Ribeiro	01473								
COM		GR	Comunicação Social (Noturno)							
FACE	ADM	08362	Comunicação organizacional (Bacharel)	2010/1	Sim	Plena		14	26	186 0
		00019	GR	Administração (Diurno)						
	ADM	08001		1971/2 1988 / 2	Sim	Plena	Exatas	12	30	60 0
	ADM	08109		1971/2 1988 / 2	Sim	Plena	Humanas	12	30	60 0
	ADM	08117	ADMINISTRAÇÃO DE EMPRESAS (Bacharel)	1971/2 1988 / 1	Sim	Plena	Exa/Hum	12	30	188 0
	ADM	08117	ADMINISTRAÇÃO DE EMPRESAS (Bacharel)	1988/2 1997 / 1	Sim	Plena		12	30	200 0
	ADM	08117	Administração de Empresas (Bacharel)	1997/2 2006 / 1	Sim	Plena		12	30	198 0
	ADM	08117	ADMINISTRAÇÃO DE EMPRESAS (Bacharel)	2006/2 2010 / 0	Sim	Plena		12	30	198 0
	ADM	08117	ADMINISTRAÇÃO DE EMPRESAS (Bacharel)	2010/1 2010 / 2	Sim	Plena		12	30	200 0
	ADM	08117	Administração (Bacharel)	2011/1 2017 / 1	Sim	Plena		12	28	200 0
		08117								

Figura 3.4: Ilustração do arquivo “Código de opções de curso.txt”.

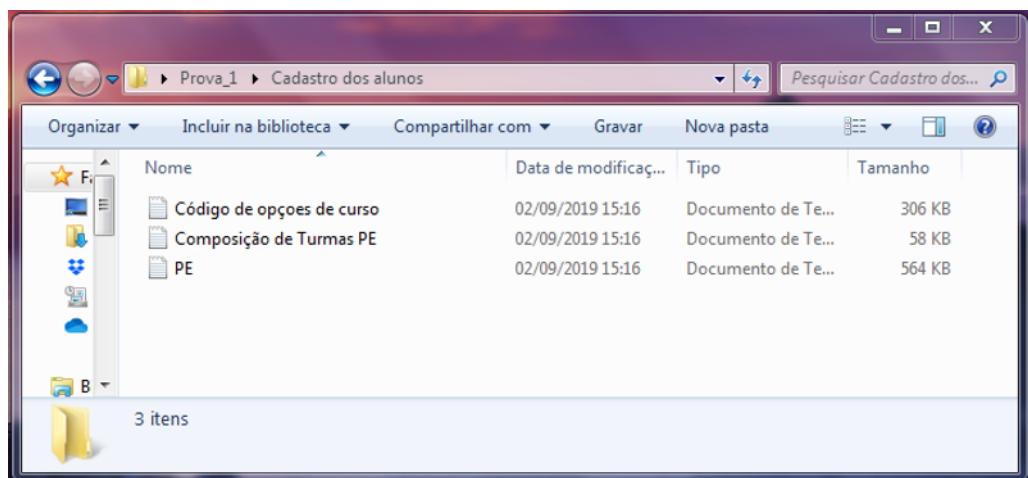


Figura 3.5: Ilustração da pasta “Cadastro dos alunos” antes do processo de geração das provas.

3.2 Geração das provas

Após concluir a etapa de pré-requisitos, o código *gerar_prova.R* pode ser executado conforme descrito a seguir.

3.2.1 Localizando as pastas

Execute as seguintes linhas do programa:

```
### Identificando as pastas necessárias

if(length(ls()) > 0) rm(list = ls()) # Limpando a área de trabalho

pasta.cadastros <- choose.dir(default=getwd(),
                                 caption="Escolha a pasta com os dados dos alunos")

pasta.provas <- choose.dir(default=getwd(),
                               caption="Escolha a pasta onde salvar as provas")

pasta.questoes <- choose.dir(default=getwd(),
                                caption="Escolha a pasta com o banco de questões")

pasta.suplementos <- choose.dir(default=getwd(),
                                    caption="Escolha a pasta com os arquivos suplementares")
```

Nesse passo, o programa solicitará que o usuário indique a localização de alguns arquivos. Mais especificamente, quatro pastas precisam ser especificadas, conforme exemplificado abaixo:

- Escolha a pasta com os dados dos alunos: \PE\Prova_1\Cadastro dos alunos.
- Escolha a pasta onde salvar as provas: \PE\Prova_1.
- Escolha a pasta com o banco de questões: \PE\Banco_questoes.
- Escolha a pasta com os arquivos suplementares: \PE\Suplementos.

3.2.2 Carregando as funções suplementares

Execute a seguinte linha do programa:

```
### Carregando as funções suplementares usadas no código

source(paste0(pasta.suplementos, "\\Funcoes_Extras.R")) # Complementando pacote Exams
```

Além da estrutura presente no pacote *exams*, outras funções foram desenvolvidas para complementar o sistema de avaliação. Nessa etapa, essas funções extras serão carregadas no R.

	A	B	C	D	E	F	G	H	I
1	Matrícula	Nome	Ano	Turma	Horario Professor	Código_curso	Curso		
2	1,4E+08	JEISON TAVA	2014	AA	8/10 CIRA ETHEOWALDA	1368	Outros		
3	1,5E+08	NATALIA VALE	2015	AA	8/10 CIRA ETHEOWALDA	1589	Outros		
4	1,6E+08	GUILHERME I	2016	AA	8/10 CIRA ETHEOWALDA	3221	Outros		
5	1,6E+08	MATHEUS MY	2016	AA	8/10 CIRA ETHEOWALDA	1325	Outros		
6	1,6E+08	GUSTAVO NE	2016	AA	8/10 CIRA ETHEOWALDA	3221	Outros		
7	1,6E+08	ANDRÃ‰ DA	2016	AA	8/10 CIRA ETHEOWALDA	3221	Outros		
8	1,7E+08	LUCAS MESQ	2017	AA	8/10 CIRA ETHEOWALDA	3221	Outros		
9	1,7E+08	LUISA BOUDE	2017	AA	8/10 CIRA ETHEOWALDA	8117	Outros		
10	1,7E+08	PIERRE CARI	2017	AA	8/10 CIRA ETHEOWALDA	1589	Outros		
11	1,7E+08	LUCAS GOMII	2017	AA	8/10 CIRA ETHEOWALDA	3221	Outros		
12	1,7E+08	LUCAS DE AS	2017	AA	8/10 CIRA ETHEOWALDA	8885	Outros		
13	1,7E+08	MARCEL ROC	2017	AA	8/10 CIRA ETHEOWALDA	2321	Outros		
14	1,7E+08	WENDEL DE	2017	AA	8/10 CIRA ETHEOWALDA	3221	Outros		
15	1,7E+08	JULIA DE LIM	2017	AA	8/10 CIRA ETHEOWALDA	6335	Outros		
16	1,7E+08	AMANDA BRA	2017	AA	8/10 CIRA ETHEOWALDA	3221	Outros		

Figura 3.6: Ilustração do arquivo “Tabela de alunos matriculados.csv”.

3.2.3 Organizando o cadastro dos alunos

Aqui, o programa cria o cadastro dos alunos com base nos arquivos disponibilizados pela secretaria do departamento de Estatística da Universidade de Brasília. Em outras instituições, o usuário não precisa rodar essa seção (leia a observação de como proceder no final desse tópico).

```
### Organizando o cadastro a partir dos arquivos disponibilizados pela secretaria.
cadastro <- gerar.cadastro() # Gerando um arquivo csv com o cadastro.
```

Nesse passo, o programa solicitará que o usuário indique a localização dos seguintes arquivos:

- Arquivo .txt com os dados dos alunos: “PE.txt”.
- Arquivo .txt com os dados dos cursos dos alunos: “Composição de Turmas PE.txt”.
- Arquivo .txt com o dicionário de códigos de cursos: “Código de opções de curso.txt”.

O comando acima irá incluir automaticamente na pasta “Cadastro dos alunos” uma planilha de alunos matriculados para cada turma e uma planilha contendo informação de todas as turmas, conforme ilustrado a seguir.

Observação: Para geração de provas em outras universidades, não é necessário rodar essa seção. No entanto, será necessário incluir na pasta “Cadastro dos alunos” as planilhas por turma “Tabela de alunos – XX.csv” seguindo o modelo acima e exatamente essa nomenclatura dos arquivos. Além disso, deve-se combinar os dados de todas as turmas na “Tabela de alunos matriculados.csv” seguindo o mesmo modelo. A pasta “Cadastro dos alunos” deve conter os arquivos .csv ilustrados na figura a seguir (e os arquivos .txt para o caso de PE). Novamente, lembre-se de salvar os arquivos usando o encoding UTF-8.

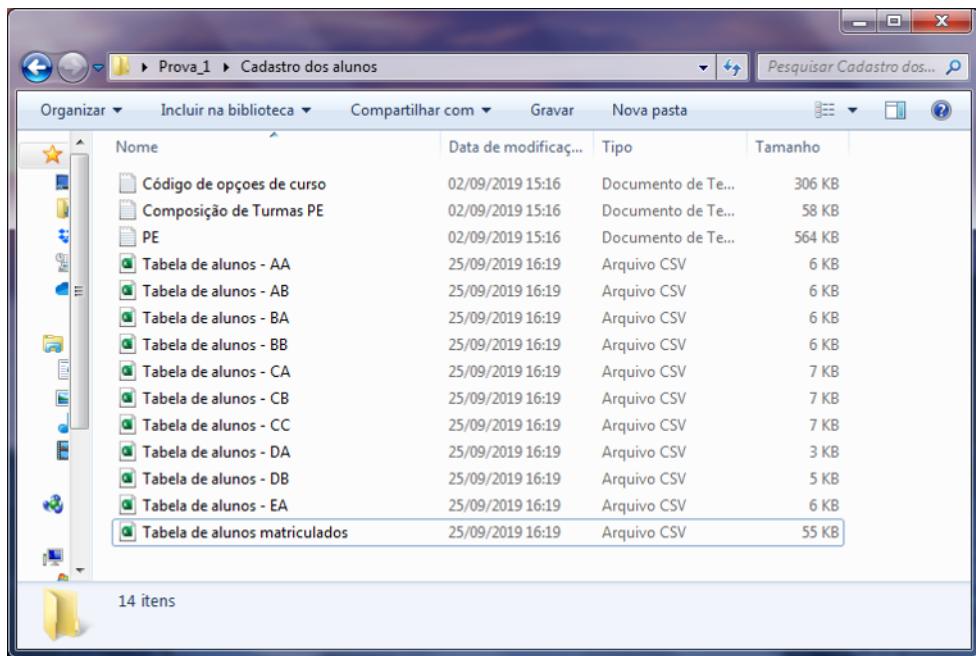


Figura 3.7: Ilustração da pasta “Cadastro dos alunos”.

3.2.4 Definindo os parâmetros da prova

Execute o código a seguir alterando os parâmetros conforme as respectivas informações da prova.

```
### Definindo os parâmetros da prova
semente <- 8768435 # Por segurança, trocar esse número a cada aplicação
set.seed(semente) # Definindo a semente do algoritmo
numero.prova <- 1 # Número da prova
nome.prova <- paste0("Prova_", numero.prova) # Nome da prova
titulo.prova <- paste("Prova", numero.prova) # Título incluído na folha de respostas
data.prova <- "2019-09-18" # aaaa-mm-dd
n.questoes <- 10 # Definindo o número de questões na prova
total.pontos <- 10 # Definindo a pontuação máxima da prova
```

Aqui é essencial modificar a semente do algoritmo a cada nova aplicação de prova. A semente é importante para garantir a reproduzibilidade dos resultados ao rodar novamente o código, se for necessário. Além disso, o uso de uma semente repetida ou previsível poderia, em casos extremos, colocar em risco o sigilo da prova.

3.2.5 Definindo os temas da prova

O banco de questões é organizado por tema, e o nome dos arquivos seguem a seguinte nomenclatura “xx_nome.tema_yy.Rnw”, onde “xx” representa o número do tema e “yy” o número da questão do

respectivo tema. A estrutura de questões da prova é definida nesse estágio, sendo adotada 1 questão por tema, todas com igual pontuação (a primeira questão pode ter sua pontuação ligeiramente alterada para garantir que a soma dos pontos seja exatamente igual ao valor previamente definido). Alterações nesta estrutura padrão devem ser feitas com cautela.

```
### Definindo os temas da prova

aux <- list.files(pasta.questoes, ".Rnw")

banco.questoes <- vector(n.questoes, mode="list")

for(questao in 1:n.questoes) { # Uma questão por tópico
    topico <- formatC(questao + 10*(numero.prova - 1), digits=1, flag="0", format="d")
    banco.questoes[[questao]] <- aux[str_detect(aux, paste0("^", topico))]
}

### Definições padrão

aux <- list.files(pasta.cadastros) # Lendo os cadastros das turmas

cadastro.turmas <- aux[grep("Tabela de alunos - ", aux)]

turmas <- substr(cadastro.turmas, 20, 21) # Identificando a sigla das turmas

n.turmas <- length(cadastro.turmas) # Identificando o número de turmas

n.tempos <- length(banco.questoes) # Definindo o número de temas na prova

questoes.por.tema <- rep(1, n.questoes) # Definindo o número de questões por tema

Pontos <- round(total.pontos*rep(1/n.questoes, n.questoes), 2) # Valor da questão

Pontos[1] <- round(total.pontos-sum(Pontos[-1]), 2)
```

3.2.6 Sorteando as questões para as turmas

Rode o código a seguir.

```
### Sortear as questões para as turmas

questoes.sorteadas <- sortear.questoes()
```

No sorteio das questões, será necessário selecionar o arquivo com a tabela de questões com as respectivas dificuldades “PE\Banco_questoes\2_2019\Completo\matriz.dificuldades.csv”. O arquivo contém o nome da questão, o percentual de acertos, o tamanho da amostra para o qual o percentual foi calculado (apenas para referência) e o número do tema da questão.

Observação: A função “sortear.questoes()” tem o parâmetro “prob.media” que especifica a nota média desejada, com default “prob.media=0.55”. Ao alterar essa nota, certifique-se de que as questões do banco têm um percentual de acerto compatível. Embora esse parâmetro possa ser usado para aumentar ou diminuir a probabilidade de seleção de cada questão com base em seu nível de dificuldade, recomendamos

A	B	C	D	E	F	G	H	I	J
1	Nome.questao	percentual	n	num.tema					
2	1 01_medida_probabilidade_01.Rnw	0,388888889	36	1					
3	2 01_medida_probabilidade_02.Rnw	0,583333333	48	1					
4	3 01_medida_probabilidade_03.Rnw	0,407407407	27	1					
5	4 01_medida_probabilidade_04.Rnw	0,330188679	106	1					
6	5 01_medida_probabilidade_05.Rnw	0,277777778	36	1					
7	6 01_medida_probabilidade_06.Rnw	0,434920455	NA		1				
8	7 01_medida_probabilidade_07.Rnw	0,434343434	99	1					
9	8 01_medida_probabilidade_08.Rnw	0,434920455	NA		1				
10	9 01_medida_probabilidade_09.Rnw	0,622641509	53	1					
11	10 01_medida_probabilidade_10.Rnw	0,434782609	46	1					
12	11 02_propriedade_probabilidade_01.Rnw	0,493142741	NA		2				
13	12 02_propriedade_probabilidade_02.Rnw	0,493142741	NA		2				
14	13 02_propriedade_probabilidade_03.Rnw	0,611111111	144	2					
15	14 02_propriedade_probabilidade_04.Rnw	0,730769231	52	2					
16	15 02_propriedade_probabilidade_05.Rnw	0,291666667	48	2					

Figura 3.8: Ilustração do arquivo “Matriz de dificuldades.csv”.

cautela na definição desse valor. Grandes alterações desse parâmetro, para mais ou para menos, poderiam resultar em provas desconfortavelmente parecidas entre as turmas. As questões sorteadas serão salvas no arquivo “Questoes_sorteadas_Prova_1.csv” na pasta \PE\Prova_1.

3.2.7 Criando múltiplos exames

Aqui as provas serão geradas para cada aluno de cada turma com base nas questões sorteadas no passo 6. É possível adicionar uma (ou mais) página em branco em cada prova alterando a variável “adicionar.paginas”. Esta página adicional pode ser usada para rascunho.

```
### criando múltiplos exames
gerar.provas(
  numero.turmas=1:n.turmas,
  adicionar.paginas=rep(1, n.turmas),
  #n.provas.turmas=rep(1, n.turmas),
  duplex=F
)
```

Caso tenha interesse em testar uma versão preliminar do código, rode essa seção desmarcando a linha “#n.provas.turmas=rep(1, n.turmas)”, fazendo com que apenas uma prova seja gerada em cada turma. Para rodar a versão final, volte a partir do passo 4 para manter a mesma semente.

Nessa etapa será criada uma pasta chamada “Para_impressao” com 1 arquivo pdf por turma para impressão. Para facilitar a impressão frente e verso, a função verifica se a capa de alguma das provas aparece em uma página par. Em caso afirmativo, a função adiciona automaticamente uma página em

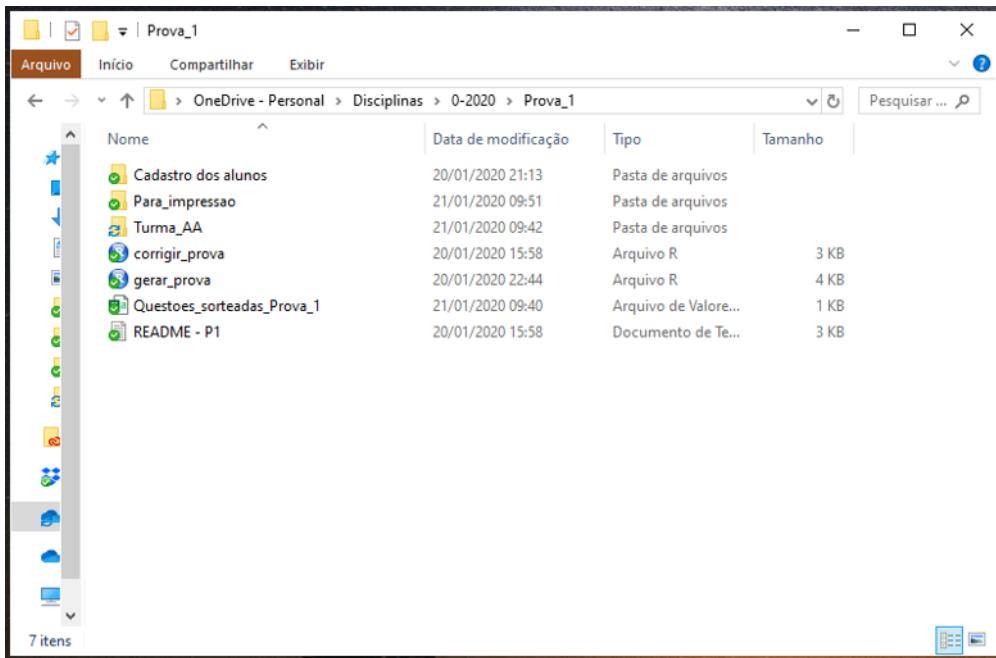


Figura 3.9: Ilustração da pasta “Prova_1”.

branco de modo a garantir que nenhuma uma folha conterá páginas de provas distintas.

Além disso, para cada turma será criada uma pasta chamada “Turma_XX”, conforme ilustrado na figura a seguir.

A pasta de cada turma contém 3 subpastas. Em “Provas” estão as provas individuais de cada aluno separadamente e em “Solucoes” estão suas respectivas resoluções (criadas com base nos arquivos .Rnw da pasta “Banco_questoes”). A pasta “Respostas” será usada posteriormente na etapa de correção das provas.

3.2.8 Salvando a área de trabalho

Execute o código a seguir para salvar a área de trabalho das provas geradas na pasta “\PE\Prova_1”.

```
### Salvando a área de trabalho
save.image(file=paste0(pasta.provas, "\\Area_trabalho_provas_geradas.RData"))
```

O processo de geração das provas foi concluído. Na pasta “Para_impressao” estão os arquivos que devem ser enviados para impressora. Devido ao tamanho dos arquivos, o compartilhamento pode ser realizado na nuvem, por meio de um link do *dropbox*, por exemplo.

3.2.9 Considerações finais

Após a impressão dos arquivos, o caderno de questões de cada aluno é grampeado e a ele anexamos a folha de resposta utilizando um clips para não ser amassada. Adicionalmente, entregamos ao aluno

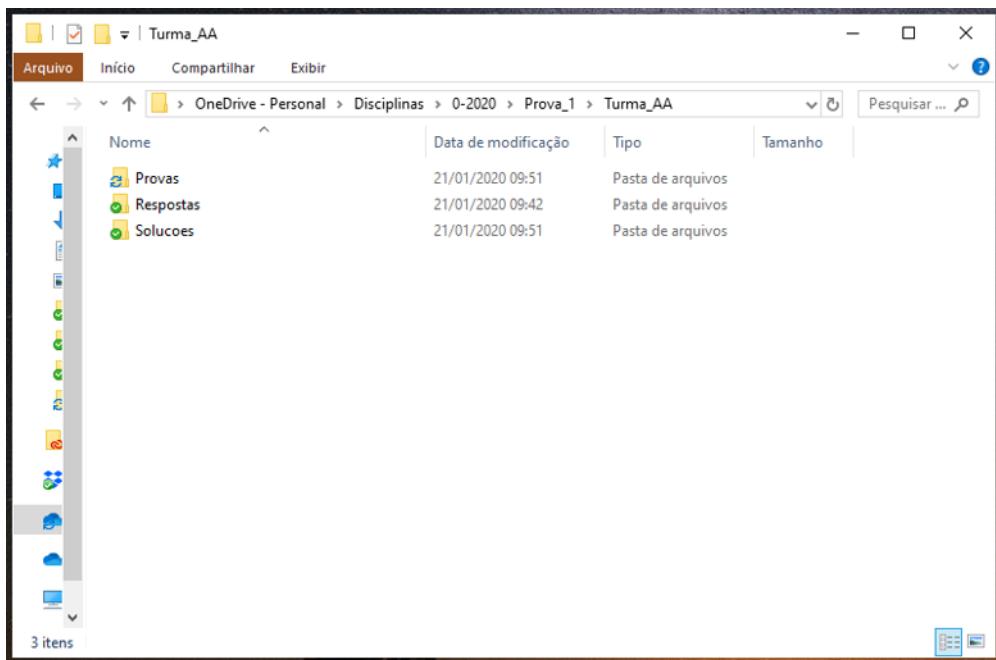


Figura 3.10: Ilustração da pasta “Turma_AA”.

uma folha com o formulário, cujo modelo encontra-se em anexo, e uma folha contendo as tabelas das distribuições Normal e t-student. O formulário e as tabelas compõem o chamado kit prova, documentos que são impressos apenas uma vez no semestre e recolhidos e guardados após cada avaliação. Aos alunos orientamos não escrever nada no kit prova, pois o kit é reaproveitado inúmeras vezes para reduzir os custos de impressão.

Os alunos são orientados a levarem seu caderno de questões ao terminar a prova, e guardá-lo para acessar o resultado. No canto superior direito do caderno de questões de cada aluno há o número de identificação da prova, que é a senha de acesso do resultado na divulgação das notas. Considerando que o banco de questões é grande, não há prejuízo os alunos terem acesso a provas antigas. Além disso, a entrega do caderno de questões é importante para que os alunos confirmam posteriormente a solução detalhada da sua prova, podendo inclusive reportar erro em alguma questão do banco.

O código para gerar as provas substitutivas é ligeiramente diferente do apresentado aqui. No item 3.2.7, adicionamos a variável “n.provas.turmas” com um vetor contendo o número de alunos por turma que irão realizar a prova substitutiva. Os alunos que desejam fazer essa avaliação respondem a um questionário pelo Moodle demonstrando seu interesse, e assim também reduzimos os custos de impressão. Além disso, no item 3.2.5, é possível definir os temas que serão cobrados alterando a variável “banco.questoes”, conforme ilustração a seguir:

```

banco.questoes <- list(
  # Módulo 1
  aux[str_detect(aux, "04_")], # Tema 4
)
  
```

```

aux[str_detect(aux, "^06_")], # Tema 6
aux[str_detect(aux, "^08_")], # Tema 8

# Módulo 2

aux[str_detect(aux, "^1[1-3]_")], # Tema 11, 12 ou 13
aux[str_detect(aux, "^14_")], # Tema 14
aux[str_detect(aux, "^18_")], # Tema 18
aux[str_detect(aux, "^19|20_")], # Tema 19 ou 20

# Módulo 3

aux[str_detect(aux, "^2[2-4]_")], # Tema 22 ou 24
aux[str_detect(aux, "^2[5-7]_")], # Tema 25 ou 27
aux[str_detect(aux, "^28|29|30_")] # Tema 28, 29 ou 30
)

```

3.3 Correção das provas

3.3.1 Instruções gerais

Além dos pré-requisitos detalhados na Seção 3.1, seguem algumas instruções adicionais para o processo de correção das provas.

As folhas de respostas escaneadas da turma XX devem ser salvas, em formato pdf, na pasta “PE\Prova_1\Turma_XX\Respostas”. As configurações utilizadas no escâner interferem fortemente na qualidade da leitura dos documentos pelo pacote *exams*. Após diversos testes, a configuração a seguir mostrou-se a mais adequada, reduzindo o número de erros de leitura e o tamanho dos arquivos gerados.

- Nome do arquivo: Nome da turma (AA, por exemplo).
- Cor: Preto e branco (e não em escala de cinza).
- Resolução: 600 dpi.
- Qualidade: Melhor.
- Opção: Todas as provas em um único arquivo (múltiplas folhas).
- Formato: pdf.
- Orientação: Escanear de cabeça para baixo.
- Claridade: 4 (na escala de 1 a 11).
- Observações: Evitar que as páginas estejam inclinadas.

Além disso, o cuidado no preenchimento do cartão de respostas é primordial para a correta leitura dele.

Os aspectos gerais mais importantes que devem ser reforçados com os alunos são:

- O preenchimento (com caneta azul ou preta) dos campos disponíveis para a matrícula e respostas das questões deve ser realizado, exclusivamente, por meio da marcação de um “X”, e não pintando os campos citados.
- O cartão de respostas deve ser preservado, de modo que não seja amassado, dobrado, manchado ou receba outros tipos de avarias. O aluno não deve escrever no verso da folha de respostas.
- Antes de iniciar a prova, o aluno deve conferir se o número de identificação da prova (Identidade do documento) é o mesmo na folha de respostas e no caderno de questões.

A lista completa de instruções impressas nas provas de PE está no modelo de prova apresentada no apêndice.

Por fim, é recomendável fazer um backup da pasta “PE\Prova_1” antes de iniciar a correção das provas e documentar a ocorrência de problemas durante a correção (e o procedimento executado para correção deles) em um arquivo txt para registro (“README - P1.txt”).

O código `corrigir_prova.R` pode ser executado conforme descrito a seguir.

3.3.2 Carregando os arquivos

A área de trabalho salva ao final do processo de geração das provas será carregada no ambiente de trabalho do R, além das funções complementares ao pacote *exams*.

```
if(length(ls()) > 0) rm(list = ls())
setwd(choose.dir(default=getwd(), caption="Escolha a pasta onde as provas foram salvadas"))
load("Area_trabalho_provas_geradas.RData")
source(paste0(pasta.suplementos, "\\Funcoes_Extras.R")) # Complementando pacote Exams
```

3.3.3 Localizando as pastas

Esta seção deve ser rodada apenas se as provas tiverem sido geradas em outro computador. Caso contrário, pular para o próximo passo.

```
pasta.provas <- choose.dir(default=getwd(),
                               caption="Escolha a pasta onde as provas foram salvadas")
pasta.cadastros <- choose.dir(default=getwd(),
                                 caption="Escolha a pasta com os dados dos alunos")
pasta.suplementos <- choose.dir(default=getwd(),
                                   caption="Escolha a pasta com os arquivos suplementares")
```

Nesse passo, o programa solicitará que o usuário atualize a localização de algumas pastas.

- Escolha a pasta onde as provas foram salvadas: \PE\Prova_1.
- Escolha a pasta com os dados dos alunos: \PE\Prova_1\Cadastro dos alunos.
- Escolha a pasta com os arquivos suplementares: \PE\Suplementos.

3.3.4 Digitalizando as respostas

Aqui o cartão de respostas dos alunos será digitalizado. A seguir note que a função “digitalizar.respostas()” contém a opção “rotate=T” para indicar que os arquivos foram escaneados de cabeça para baixo (caso não seja o caso, modificar para “rotate=F”). Os parâmetros “threshold” e “minrot” são critérios utilizados internamente na codificação das informações contidas nas imagens. Para imagens com boa qualidade (veja Subseção 3.3.1), os valores pré-definidos são satisfatórios, mas caso seja necessário alterá-los, consulte o *help* da função *nops_scan*.

```
### Juntar os metainfos das turmas em um único objeto
juntar.metainfos(pasta.provas)

metainfo <- readRDS(paste0(pasta.provas, "\\Metainfo\\metainfo.RDS"))

### Definir as turmas a serem corrigidas
corrigir.turmas <- 1:n.turmas

### Digitalizar as respostas
renomear.arquivos(pasta.provas, corrigir.turmas)
gerar.cadastros(pasta.provas, corrigir.turmas)
digitalizar.respostas(pasta.provas, corrigir.turmas, rotate=T,
                      threshold=c(0.04, 0.6), minrot=0.001)
```

Nesta etapa será criado um arquivo chamado “metainfo.Rds” na pasta “Metainfo”. Este arquivo contém os metadados produzidos, sendo usado internamente durante o processo de correção das provas.

Além disso, na pasta de cada turma “\Turma_XX\Respostas”, os arquivos “Cadastro.csv” e “nops_scan_x.zip” serão criados. O primeiro contém os dados do aluno (nome, matrícula) e o número de identificação da sua prova. O segundo contém todas as imagens (cartão de resposta) lidas pelo programa e um arquivo “Daten.txt” com as informações extraídas dessas imagens.

Neste momento é necessário abrir o arquivo “Daten.txt” e verificar se há alguma linha com a palavra “ERROR”, indicando erro de leitura em alguma prova. Caso haja alguns erros, recomenda-se escanear as provas novamente. Alternativamente, os dados das provas não lidas poderão ser inseridos manualmente. Para tanto, abra a imagem correspondente no arquivo zipado “nops_scan_x.zip” e insira os dados na linha que apresentou erro. Para facilitar, no lugar do “ERROR”, copie e cole a linha anterior (como

modelo) e altere todos os dados necessários. Note que a codificação da marcação do aluno segue a seguinte lógica: a (10000), b (01000), c (00100), d (00010), e (00001). Após as alterações, salve o arquivo “Daten.txt” e, ao fechar a janela do arquivo zipado “nops_scan_x.zip”, selecione a opção de aceitar mudanças no arquivo zip na caixa de diálogo que aparecerá.

Caso algum aluno tenha errado a marcação da folha de respostas (preenchendo todo o quadrado, por exemplo), o arquivo “Daten.txt” também pode ser alterado nesse momento para inserir manualmente as informações.

3.3.5 Checando as leituras

Execute o código a seguir para checar as leituras dos cartões.

```
### Checar leituras em respostas onde mais de um item foi marcado
checar.leituras(pasta.provas, corrigir.turmas)

leitura.duvidosa <- leitura.correta <- procurar.erros(pasta.provas, corrigir.turmas)
fix(leitura.correta)

write.table(leitura.correta, paste0(pasta.provas, "\\leitura.correta.txt"))
corrigir.leitura(leitura.correta, pasta.provas, corrigir.turmas)
corrigir.provas(corrigir.turmas)
```

Primeiramente, as leituras serão checadas para verificar se as matrículas que os alunos preencheram estão de acordo com as matrículas registradas no cadastro dos alunos. Aqui é normal aparecer algumas discordâncias, seja porque o aluno marcou errado (mais provável), ou deixou em branco, ou o sistema não leu corretamente (pode acontecer). Em todos esses casos, para corrigir a matrícula, deve-se digitar manualmente o número correto. Para isso, digite o número de matrícula que o aluno escreveu na prova com base na imagem apresentada ao lado. Em geral, a marcação com “x” está errada, mas o número informado está correto. Caso este número também tenha sido escrito errado (raro, mas acontece), o sistema vai continuar acusando erro nessa prova, e o número de matrícula correto do aluno pode ser encontrado no cadastro com base no seu nome completo.

Essa etapa de correção das matrículas é a mais dispendiosa, por isso vale reforçar a atenção dos alunos com esse ponto para reduzir os erros.

Em seguida, as leituras serão checadas para verificar se alguma questão tem dupla marcação. Se houver, o programa vai abrir uma janela com os dados da respectiva prova para alteração manual. Lembrando que a codificação segue a seguinte lógica: a (10000), b (01000), c (00100), d (00010), e (00001). Recomenda-se abrir a imagem da prova correspondente (busque o arquivo em “nops_scan_x.zip” na pasta “Prova_1\Turma_XX\Respostas”) para averiguar o motivo da dupla marcação (falha do aluno ou

do sistema).

Após essas correções manuais, as provas serão corrigidas e alguns gráficos gerais serão apresentados no R.

3.3.6 Gerando os resultados

Além das notas, cada aluno recebe um arquivo contendo a sua folha de respostas, o gabarito correto e a solução detalhada da sua prova. Este arquivo é codificado com uma senha, esta é o número de identificação da prova, que também está impressa no canto superior direito do caderno de questões de cada aluno. Lembrando que os alunos levam seu caderno de questões ao terminar a prova, e devem guardá-lo para acessar o resultado com essa senha.

```
### Gerar os resultados

diretorio.resultados <- paste0(pasta.provas, "\\Resultados")
criar.resumos(pasta.provas, diretorio.resultados, corrigir.turmas)
gerar.espehos(pasta.provas, diretorio.resultados, corrigir.turmas)
converter.html2pdf(corrigir.turmas) # Converter os espehos de prova para pdf
banco.respostas <- gerar.banco.dados(corrigir.turmas) # Banco das respostas dos alunos
juntar.arquivos.divulgacao(corrigir.turmas) # Juntar os arquivos para divulgação
inserir.senha.pdf(corrigir.turmas)
```

Nessa seção, o sistema faz a organização desses arquivos para apresentação do resultado da prova aos alunos. Uma pasta chamada “Resultados” é criada em “\Prova_1\” contendo uma planilha por turma com as notas. Além disso, na pasta “Prova_1\Resultados\Turma_XX”, estão os arquivos de cada aluno, cujo nome é o número da matrícula e a senha é o número de identificação da prova. Esses arquivos pdf são disponibilizados no Moodle.

Considerando a estrutura da disciplina de Probabilidade e Estatística, com 3 provas e uma prova substitutiva, as menções provisórias podem ser geradas após a prova 3 rodando a linha a seguir.

```
### Gerar menções provisórias

gerar.mencoes.provisorias(corrigir.turmas)
```

A função “gerar.mencoes.provisorias()” reuni o resultado das 3 provas em uma única planilha e calcula a menção provisória (anterior à prova substitutiva) da disciplina de Probabilidade e Estatística considerando os pesos estabelecidos. Portanto, essa função deve ser usada apenas após a correção da Prova 3. Observação: caso professores façam ajustes da nota em seus arquivos pessoais, as alterações serão perdidas no momento da junção das planilhas.

De forma análoga, durante a correção das provas substitutivas, a menção final pode ser gerada rodando a linha a seguir.

```
gerar.mencoes.finais(1:n.turmas, diretorio.resultados)
```

A função “gerar.mencoes.finais()” reuni o resultado das 4 provas em uma única planilha e calcula a nota final considerando a substituição da menor nota pela nota da prova substitutiva (se beneficiar o aluno). Observação: caso professores façam ajustes da nota em seus arquivos pessoais, as alterações serão perdidas no momento da junção das planilhas.

3.3.7 Salvando a área de trabalho

Execute o código a seguir para salvar a área de trabalho das provas corrigidas na pasta “\PE\Prova_1”.

```
### Salvando a área de trabalho  
save.image(file=paste0(pasta.provas, "\\Area_trabalho_provas_corrigidas.RData"))
```

Pronto! Agora é só disponibilizar os arquivos pdf codificados com senha aos alunos.

3.3.8 Considerações finais

Caso tenha ocorrido algum problema em uma questão após a aplicação das provas, é possível realizar algumas intervenções para ajuste na nota do aluno.

- Caso 1: A resposta correta está em outra alternativa

Nesse caso, é possível alterar o código .Rnw da questão com problema para indicar que a resposta correta está em outra alternativa. Na pasta “\PE\Banco_questoes”, abra o arquivo da questão correspondente e identifique em qual posição do vetor “alt” encontra-se a alternativa correta (ver seção do código “## gerando alternativas”). Troque a linha “solutions <- c(TRUE, rep(FALSE, 4))” por “solutions <- c(F, T, F, F, F)”, se a segunda alternativa estiver correta, e assim por diante. Esta é a única linha que precisa ser alterada. Em seguida, execute novamente o código “gerar_prova.R” sem realizar nenhuma alteração nele (mantenha inclusive a mesma semente). Dessa forma, o arquivo metainfo será atualizado e, ao corrigir novamente as provas, a alternativa correta será devidamente identificada. Portanto, execute o código “corrigir_prova.R” novamente também sem nenhuma modificação. Dessa forma, tanto as notas dos alunos quanto os resultados que são enviados a eles (ver anexo 2: Resultados do exame) serão atualizados com o novo gabarito da questão.

Observação: Em todos os códigos .Rnw, a solução correta é armazenada na primeira posição do vetor “alt” (ou seja, “alt[1] <- solução correta”), e posteriormente essa posição é aleatorizada. Portanto, recomenda-se que faça a alteração detalhada no parágrafo anterior para corrigir o problema dessa avaliação que já foi realizada, mas depois modifique o código para o padrão das demais questões, ou seja, “alt[1] <- solução correta” e “solutions <- c(TRUE, rep(FALSE, 4))”.

- Caso 2: A questão não apresenta nenhuma alternativa correta

Nesse caso a questão precisa ser anulada, e o ponto pode ser dado diretamente ao aluno ou redistribuído para as demais questões da prova. Em ambos os casos, as notas podem ser diretamente alteradas nas planilhas excel da pasta “PE\Prova_x\Resultados”. Entretanto, os arquivos enviados aos alunos (ver anexo 2: Resultados do exame) não serão atualizados. Os alunos podem ser apenas notificados que receberão um ponto a mais caso tenham errado a questão (se o ponto for dado a todos, por exemplo).

Para redistribuir os pontos entre as demais questões, é possível rodar as provas novamente para atualizar tanto as notas quanto os arquivos enviados aos alunos (ver anexo 2: Resultados do exame). Nesse caso, modifique as linhas a seguir do código “gerar_prova.R” (ver seção 3.2.5) para anular a primeira questão da prova, por exemplo, e redistribuir os pontos.

Substituindo

```
# Definindo o valor de cada questão
Pontos <- round(total.pontos*rep(1/n.questoes, n.questoes), 2)
Pontos[1] <- round(total.pontos-sum(Pontos[-1]), 2)
```

por

```
# Definindo o valor de cada questão
Pontos <- c(0,round(total.pontos*rep(1/(n.questoes-1), n.questoes-1), 2))
Pontos[1] <- round(total.pontos-sum(Pontos[-1]), 2)
```

Recomenda-se registrar as intervenções necessárias durante o processo de correção das provas no arquivo “README - P1.txt” para manter controle do processo.

Apêndice

Universidade de Brasília - (Turma AA)

Prova 1 2019-09-18

Dados pessoais

Sobrenome:	
Nome:	
Assinatura:	

Número de matrícula

Neste campo **não** podem ser realizadas modificações dos dados.

Categoria	Identidade do documento
210	19091800008

Marcar cuidadosamente: Não marcado: ou

Este documento é lido à máquina. Por favor não dobrar ou sujar. Utilize uma caneta preta ou azul. Somente cruzes claramente reconhecíveis e em posição exata serão avaliadas!

Respostas 1 - 10

	a	b	c	d	e
1	<input type="checkbox"/>				
2	<input type="checkbox"/>				
3	<input type="checkbox"/>				
4	<input type="checkbox"/>				
5	<input type="checkbox"/>				
6	<input type="checkbox"/>				
7	<input type="checkbox"/>				
8	<input type="checkbox"/>				
9	<input type="checkbox"/>				
10	<input type="checkbox"/>				
	a	b	c	d	e

+

+

Leia atentamente as instruções abaixo:

- Escreva seu nome completo e matrícula na folha de respostas.
- O preenchimento (com caneta azul ou preta) dos campos disponíveis para a matrícula e respostas das questões deve ser realizado, exclusivamente, por meio da marcação de um "X", e não pintando os campos citados.
- O cartão de respostas deve ser preservado, de modo que não seja amassado, dobrado, manchado ou receba outros tipos de avarias. Será de inteira responsabilidade do aluno o correto preenchimento do cartão de respostas.
- Não escreva no verso da folha de respostas e nas folhas contendo as fórmulas e as tabelas das distribuições. Esse material será recolhido ao final da prova.
- Antes de iniciar a prova, confira se o número de identificação da prova (Identidade do documento) é o mesmo na folha de respostas e no caderno de questões.
- O aluno poderá levar consigo o caderno de questões. O resultado será disponibilizado por meio da plataforma Aprender (Moodle). Para acessá-lo, utilize como senha o número de identificação da prova.
- Caso a resposta correta não esteja presente na lista de alternativas por erro de arredondamento, será considerada correta a alternativa que melhor aproxime o resultado exato.
- Tenha em mãos somente: caneta, lápis, borracha e calculadora. Não será permitido empréstimo de material.
- A avaliação terá duração de **1h50** (uma hora e cinquenta minutos), improrrogáveis. O(a) aluno(a) que não entregar a avaliação dentro desse intervalo de tempo, terá nota 0 (zero).
- Não será permitido, em hipótese alguma, que o(a) aluno(a) faça a avaliação em uma turma/horário diferente daquele que em ele(a) está matriculado(a).
- O(a) aluno(a) só poderá sair da sala após 1h (uma hora), mesmo que já tenha finalizado a avaliação. Após a saída do(a) primeiro(a) aluno(a) não será permitido a entrada de nenhum(a) outro(a) aluno(a).
- O(a) aluno(a) deverá portar sua carteira de estudante e apresentá-la, quando for solicitada sua assinatura.
- O ponto decimal (" . " ponto) será adotado como símbolo de separador de decimais.
- Os enunciado das questões apresentam situações hipotéticas e valores fictícios.

Faça uma excelente Prova!

1. (1 ponto) Uma certa fábrica de canetas esferográficas tem encontrado defeito em 1% de sua produção. Assumindo independência entre as falhas, a probabilidade de, entre 113 canetas, pelo menos uma ser defeituosa é:
 - (a) 0.010
 - (b) 0.257
 - (c) 0.679
 - (d) 0.990
 - (e) 0.265
2. (1 ponto) Suponha que de 10 objetos escolhemos 4 ao acaso com reposição. Qual a probabilidade de que nenhum objeto seja escolhido mais de uma vez? Aproxime a resposta com duas casa decimais.
 - (a) 0.94
 - (b) 0.50
 - (c) 0.40
 - (d) 0.70
 - (e) 0.07
3. (1 ponto) O SAC de uma empresa recebe, em média, 95 ligações por dia em horário comercial. A fim de estimar o número adequado de atendentes para trabalhar nesses horários, a empresa deseja estimar a probabilidade de receber mais do que 98 ligações em um único dia no horário comercial. Seja X o número de ligações em um determinado dia, qual das distribuições de probabilidade seria adequada para modelar a variável aleatória X ?
 - (a) Poisson(95)
 - (b) Binomial(0.5, 98)
 - (c) Binomial(0.5, 95)
 - (d) Geométrica(95/98)
 - (e) Poisson(98)
4. (1 ponto) Segundo a empresa de consultoria Kantar no Brasil, a confiança no noticiário político eleitoral visto em redes sociais tem diminuído nos últimos anos por causa da ocorrência de "Fake news". Estima-se que dessas notícias veiculadas nas redes sociais 45% são "Fake news". Se uma pessoa já leu 9 notícias em uma rede social e conseguiu checar a veracidade delas por outra fonte confiável, qual é a probabilidade condicional de que a 11ª notícia que ela ler seja a primeira "Fake news" lida?
 - (a) 0.1114
 - (b) 0.3025
 - (c) 0.2025
 - (d) 0.2475
 - (e) 0.1361
5. (1 ponto) Considere que $P(A) = 1/2$, $P(C) = 1/4$ e $P(A \cap B) = 1/6$, sendo A e C eventos independentes, e B e C eventos disjuntos. Calcule $P((B \cup C)|A)$.
 - (a) 0.021
 - (b) 0.583
 - (c) 0.292
 - (d) 0.042

(e) 0.557

6. (1 ponto) Seja X uma variável aleatória discreta com a seguinte distribuição de probabilidades:

$$P(X = x) = \frac{k}{x}, \quad \text{onde } X \text{ assume os valores } 3, 4, 5 \text{ e } 9.$$

Assinale a alternativa correspondente à variância de X .

- (a) 720/161
- (b) 90180/25921
- (c) 540/23
- (d) 497339/11439
- (e) 180/161

7. (1 ponto) Uma caixa contém 7 bolas azuis e 4 bolas brancas. Uma bola é extraída, sua cor observada e, a seguir, a bola é reposta na caixa com mais 4 bolas da mesma cor. Esse processo é repetido consecutivamente. Qual a probabilidade de se extrair uma bola azul na segunda retirada?

- (a) 0.342
- (b) 0.405
- (c) 0.636
- (d) 0.868
- (e) 0.467

8. (1 ponto) Considere que o DF possui 2.562.963 habitantes (dados do Censo 2010), e que a probabilidade de um habitante da cidade acionar o SAMU em uma hora qualquer do dia é de 0.000002. Supondo que os acionamentos ao SAMU ocorram de forma independente, qual é a probabilidade de observarmos exatamente 7 chamados ao SAMU em determinada hora no DF (utilize a aproximação de poisson da distribuição binomial)?

- (a) 0.875
- (b) 0.110
- (c) 0.117
- (d) 0.125
- (e) 0.890

9. (1 ponto) Para inspecionar um lote de 12 peças, o funcionário de uma empresa sorteia uma amostra de 8 peças ao acaso. Caso nenhuma peça defeituosa seja encontrada na amostra o lote é aceito; caso contrário é devolvido ao fornecedor. Suponha que 2 das 12 peças sejam defeituosas. Se a escolha for realizada sem reposição qual a probabilidade de aceitação do lote?

- (a) 0.233
- (b) 0.167
- (c) 0.139
- (d) 0.028
- (e) 0.091

10. (1 ponto) Em uma cidade em que os carros são testados para emissão de poluentes, 28% deles emitem quantidade considerada excessiva. O teste repara 91% dos carros que emitem excesso de poluentes, mas resulta em falso positivo para 6.000000000000001% dos carros que emitem quantidade considerada normal. Qual é a probabilidade de um carro reprovado no teste realmente emitir quantidade excessiva de poluentes?

- (a) 0.964
- (b) 0.855
- (c) 0.940
- (d) 0.145
- (e) 0.910

Resultados do exame

Nome: [REDACTED]

Número de matrícula: [REDACTED]

Nota: 7

Avaliação

Questão	Nota	Resposta dada	Resposta correta
1	0	a____	__c__
2	0	__c__	b____
3	1	a____	a____
4	1	__d_	__d_
5	1	_b____	__b__
6	0	____e	__b__
7	1	__c__	__c__
8	1	_b____	__b__
9	1	____e	____e
10	1	_b____	__b__

Folha do exame

+ Universidade de Brasília - (Turma AA) +
Prova 1 2019-09-18

Dados pessoais		Número de matrícula	
Sobrenome:	[REDACTED]	[REDACTED]	0
Nome:	[REDACTED]	[REDACTED]	1
Assinatura:	[REDACTED]	[REDACTED]	2
		verificado	
<p>Neste campo não podem ser realizadas modificações dos dados.</p>			
Categoria	Identidade do documento		
210	19091800008		
Marcar cuidadosamente: <input checked="" type="checkbox"/> Não marcado: <input type="checkbox"/> ou <input checked="" type="checkbox"/>			
Este documento é lido à máquina. Por favor não dobrar ou sujar. Utilize uma caneta preta ou azul . Somente cruzes claramente reconhecíveis e em posição exata serão avaliadas!			

Respostas 1 - 10					
	a	b	c	d	e
1	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
5	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
7	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
10	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	a	b	c	d	e

+ +

Probabilidade e Estatística

1. Questão

Uma certa fábrica de canetas esferográficas tem encontrado defeito em 1% de sua produção. Assumindo independência entre as falhas, a probabilidade de, entre 113 canetas, pelo menos uma ser defeituosa é:

- (a) 0.010
- (b) 0.257
- (c) 0.679
- (d) 0.990
- (e) 0.265

Solução

Seja X o número de canetas defeituosas na amostra, sabe-se que $X \sim \text{Binomial}(113, 0.01)$. Portanto,

$$P(X > 0) = 1 - P(X = 0) = 1 - 0.321 = 0.679.$$

- (a) Falso
- (b) Falso
- (c) Verdadeiro
- (d) Falso
- (e) Falso

2. Questão

Suponha que de 10 objetos escolhemos 4 ao acaso com reposição. Qual a probabilidade de que nenhum objeto seja escolhido mais de uma vez? Aproxime a resposta com duas casa decimais.

- (a) 0.94
- (b) 0.50
- (c) 0.40
- (d) 0.70
- (e) 0.07

Solução

A probabilidade desejada é dada por

$$\frac{10 \times 9 \times \dots \times 7}{10^4}.$$

- (a) Falso
- (b) Verdadeiro
- (c) Falso
- (d) Falso
- (e) Falso

3. Questão

O SAC de uma empresa recebe, em média, 95 ligações por dia em horário comercial. A fim de estimar o número adequado de atendentes para trabalhar nesses horários, a empresa deseja estimar a probabilidade de receber mais do que 98 ligações em um único dia no horário comercial. Seja X o número de ligações em um determinado dia, qual das distribuições de probabilidade seria adequada para modelar a variável aleatória X ?

Probabilidade e Estatística

- (a) Poisson(95)
- (b) Binomial(0.5, 98)
- (c) Binomial(0.5, 95)
- (d) Geométrica(95/98)
- (e) Poisson(98)

Solução

Como a probabilidade de sucesso (receber uma ligação de um dado cliente) é pequena e há muitas chances de sucesso (número de clientes que potencialmente ligarão), a distribuição de probabilidade adequada para descrever o comportamento da variável aleatória X é Poisson(95).

- (a) Verdadeiro
- (b) Falso
- (c) Falso
- (d) Falso
- (e) Falso

4. Questão

Segundo a empresa de consultoria Kantar no Brasil, a confiança no noticiário político eleitoral visto em redes sociais tem diminuído nos últimos anos por causa da ocorrência de "Fake news". Estima-se que dessas notícias veiculadas nas redes sociais 45% são "Fake news". Se uma pessoa já leu 9 notícias em uma rede social e conseguiu checar a veracidade delas por outra fonte confiável, qual é a probabilidade condicional de que a 11^a notícia que ela ler seja a primeira "Fake news" lida?

- (a) 0.1114
- (b) 0.3025
- (c) 0.2025
- (d) 0.2475
- (e) 0.1361

Solução

Seja $X \sim Geo(0.45)$ o número de notícias lidas por uma pessoa, em uma rede social, até ler uma "Fake news". Então, usando a propriedade da perda da memória, tem-se:

$$P(X = 11 | X > 9) = P(X = 2) = 0.45 \times 0.55^1 = 0.2475.$$

- (a) Falso
- (b) Falso
- (c) Falso
- (d) Verdadeiro
- (e) Falso

5. Questão

Considere que $P(A) = 1/2$, $P(C) = 1/4$ e $P(A \cap B) = 1/6$, sendo A e C eventos independentes, e B e C eventos disjuntos. Calcule $P((B \cup C)|A)$.

- (a) 0.021
- (b) 0.583

Probabilidade e Estatística

- (c) 0.292
- (d) 0.042
- (e) 0.557

Solução

Pela definição de probabilidade condicional,

$$P((B \cup C)|A) = \frac{P((B \cup C) \cap A)}{P(A)} = \frac{P((B \cap A) \cup (C \cap A))}{P(A)}.$$

Como B e C são disjuntos, então são também disjuntos os eventos $(B \cap A)$ e $(C \cap A)$. Logo, $P((B \cap A) \cup (C \cap A)) = P(B \cap A) + P(C \cap A)$. Além disso, como A e C são independentes, então $P(C \cap A) = P(C)P(A)$. Daí, temos que

$$\frac{P((B \cap A) \cup (C \cap A))}{P(A)} = \frac{P(B \cap A) + P(C)P(A)}{P(A)} = \frac{0.167 + 0.250 \times 0.500}{0.500} = 0.583.$$

- (a) Falso
- (b) Verdadeiro
- (c) Falso
- (d) Falso
- (e) Falso

6. Questão

Seja X uma variável aleatória discreta com a seguinte distribuição de probabilidades:

$$P(X = x) = \frac{k}{x}, \quad \text{onde } X \text{ assume os valores } 3, 4, 5 \text{ e } 9.$$

Assinale a alternativa correspondente à variância de X .

- (a) 720/161
- (b) 90180/25921
- (c) 540/23
- (d) 497339/11439
- (e) 180/161

Solução

Primeiramente devemos determinar o valor de k . Uma vez que a soma das probabilidades deve ser um, basta resolver a equação

$$\frac{k}{3} + \frac{k}{4} + \frac{k}{5} + \frac{k}{9} = 1,$$

resultando em $k = 180/161$. Para o cálculo da variância precisamos, antes, calcular os valores de $E(X)$ e $E(X^2)$:

$$\begin{aligned} E(X) &= 3 \times P(X = 3) + 4 \times P(X = 4) + 5 \times P(X = 5) + 9 \times P(X = 9) = 720/161, \\ E(X^2) &= 3^2 \times P(X = 3) + 4^2 \times P(X = 4) + 5^2 \times P(X = 5) + 9^2 \times P(X = 9) = 540/23. \end{aligned}$$

De modo que $V(X) = E(X^2) - \{E(X)\}^2 = 90180/25921$.

- (a) Falso
- (b) Verdadeiro

Probabilidade e Estatística

- (c) Falso
- (d) Falso
- (e) Falso

7. Questão

Uma caixa contém 7 bolas azuis e 4 bolas brancas. Uma bola é extraída, sua cor observada e, a seguir, a bola é reposta na caixa com mais 4 bolas da mesma cor. Esse processo é repetido consecutivamente. Qual a probabilidade de se extrair uma bola azul na segunda retirada?

- (a) 0.342
- (b) 0.405
- (c) 0.636
- (d) 0.868
- (e) 0.467

Solução

Considere os eventos

$$\begin{aligned}A &= \text{"extraír uma bola azul na segunda retirada"}, \\B &= \text{"extraír uma bola branca na primeira retirada"}.\end{aligned}$$

Note que $\{B, B^c\}$ forma uma partição do espaço amostral, onde $P(B) = 4/11$ e $P(B^c) = 7/11$. Logo, pelo Teorema da probabilidade total tem-se

$$\begin{aligned}P(A) &= P(A|B)P(B) + P(A|B^c)P(B^c) \\&= \frac{7}{15} \times \frac{4}{11} + \frac{11}{15} \times \frac{7}{11} \approx 0.636.\end{aligned}$$

- (a) Falso
- (b) Falso
- (c) Verdadeiro
- (d) Falso
- (e) Falso

8. Questão

Considere que o DF possui 2.562.963 habitantes (dados do CENSO 2010), e que a probabilidade de um habitante da cidade acionar o SAMU em uma hora qualquer do dia é de 0.000002. Supondo que os acionamentos ao SAMU ocorram de forma independente, qual é a probabilidade de observarmos exatamente 7 chamados ao SAMU em determinada hora no DF (utilize a aproximação de poisson da distribuição binomial)?

- (a) 0.875
- (b) 0.110
- (c) 0.117
- (d) 0.125
- (e) 0.890

Probabilidade e Estatística

Solução

O número de chamados ao SAMU tem distribuição binomial com parâmetros

$$X \sim \text{Bin}(2562963, 0.000002).$$

Utilizando a aproximação de Poisson para a Binomial, tem-se, aproximadamente, que

$$X \sim \text{Pois}(np = \lambda = 5.126).$$

Portanto, a probabilidade de observarmos exatamente $X = 7$, é

$$P(X = 7) = \exp(-\lambda) \times \lambda^7 / 7! = 0.110.$$

- (a) Falso
- (b) Verdadeiro
- (c) Falso
- (d) Falso
- (e) Falso

9. Questão

Para inspecionar um lote de 12 peças, o funcionário de uma empresa sorteia uma amostra de 8 peças ao acaso. Caso nenhuma peça defeituosa seja encontrada na amostra o lote é aceito; caso contrário é devolvido ao fornecedor. Suponha que 2 das 12 peças sejam defeituosas. Se a escolha for realizada sem reposição qual a probabilidade de aceitação do lote?

- (a) 0.233
- (b) 0.167
- (c) 0.139
- (d) 0.028
- (e) 0.091

Solução

Seja X a variável relativa ao número de peças defeituosas. A probabilidade de aceitação do lote é dada por

$$P(X = 0) = \frac{\binom{10}{8} \binom{2}{0}}{\binom{12}{8}} = 0.091.$$

- (a) Falso
- (b) Falso
- (c) Falso
- (d) Falso
- (e) Verdadeiro

10. Questão

Em uma cidade em que os carros são testados para emissão de poluentes, 28% deles emitem quantidade considerada excessiva. O teste repara 91% dos carros que emitem excesso de poluentes, mas resulta em falso positivo para 6.00000000000001% dos carros que emitem quantidade considerada normal. Qual é a probabilidade de um carro reprovado no teste realmente emitir quantidade excessiva de poluentes?

- (a) 0.964

Probabilidade e Estatística

- (b) 0.855
- (c) 0.940
- (d) 0.145
- (e) 0.910

Solução

Defina os eventos

$$\begin{aligned} R &= \text{"O carro foi reprovado no teste".} \\ E &= \text{"Emissão excessiva de gases poluentes".} \\ N &= \text{"Emissão normal de gases poluentes".} \end{aligned}$$

Pelo teorema de bayes, a probabilidade desejada é dada por

$$P(E|R) = \frac{P(R|E) \times P(E)}{P(R)}$$

$$= \frac{0.91 \times 0.28}{0.91 \times 0.28 + 0.06000000000000001 \times 0.72} = \frac{0.255}{0.298} = 0.855.$$

- (a) Falso
- (b) Verdadeiro
- (c) Falso
- (d) Falso
- (e) Falso

Fórmulas

Probabilidade:

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
 - $P(A_i | B) = \frac{P(A_i)P(B | A_i)}{\sum_j^k P(A_j)P(B | A_j)}$
 - $P(A \cap B) = P(A|B)P(B)$
 - $P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$
-

Momentos (média, variância e covariância):

$$\bullet E[g(X)] = \begin{cases} \sum_i g(x_i)P(X = x_i) \\ \int_{-\infty}^{\infty} g(x)f(x)dx \end{cases}$$

- $\text{Var}(X) = E(X^2) - [E(X)]^2$
- $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$
- $\text{Corr}(X, Y) = \text{Cov}(X, Y)/(\sigma_X\sigma_Y)$
- $E[H(X, Y)] = \sum_i \sum_j H(x_i, y_j)P(x_i, y_j)$

Distribuições amostrais quando $n \rightarrow \infty$:

- $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
 - $\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$
-

Intervalos de confiança:

- $\left(\bar{x} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$
 - $\left(\hat{p} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$
 - $\left(\bar{x} - t_{1-\frac{\alpha}{2}; n-1} \frac{s}{\sqrt{n}}, \bar{x} + t_{1-\frac{\alpha}{2}; n-1} \frac{s}{\sqrt{n}}\right)$
-

Tamanho de amostra:

- $n = \left(\frac{z_{1-\frac{\alpha}{2}} \sigma}{E}\right)^2$
 - $n = \left(\frac{z_{1-\frac{\alpha}{2}} \sqrt{\tilde{p}(1-\tilde{p})}}{E}\right)^2$
-

Observações:

Na fórmula do Teorema de Bayes, é necessário que $A_i \cap A_j = \emptyset, \forall i \neq j$ e $\Omega = A_1 \cup A_2 \cup \dots \cup A_k$. O nível de confiança dos intervalos apresentados é de $1 - \alpha$. O quantil α da distribuição Normal padrão é denotado por $z_\alpha = P(Z \leq \alpha)$, onde $Z \sim N(0, 1)$. De maneira análoga, o quantil α da distribuição T com n graus de liberdade é representado por $t_{\alpha; n}$. Denotamos por E a margem de erro máxima admitida no cálculo do tamanho de amostra e \tilde{p} é uma estimativa preliminar da proporção com base em amostra-piloto ou estudos anteriores. Distribuição nula é a distribuição da estatística de teste sob H_0 , e μ_0 e p_0 são os valores assumidos dos parâmetros na hipótese nula.

Distribuições de probabilidade:

Bin (n, p)	$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$	$E(X) = np$	$Var(X) = np(1-p)$
Geo (p)	$P(X = k) = p(1-p)^{k-1}$ $P(X = n+k \mid X > n) = P(X = k)$ $P(X > n+k \mid X > n) = P(X > k)$	$E(X) = \frac{1}{p}$	$Var(X) = \frac{1-p}{p^2}$
Hiper (N, b, n)	$P(X = k) = \frac{\binom{b}{k} \binom{N-b}{n-k}}{\binom{N}{n}}$	$E(X) = \frac{nb}{N}$	$Var(X) = \frac{nb}{N} \left(1 - \frac{b}{n}\right) \left(\frac{N-n}{N-1}\right)$
Pois (λ)	$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$	$E(X) = \lambda$	$Var(X) = \lambda$
Exp (λ)	$f(x) = \lambda e^{-\lambda x}, \quad x > 0$ $F(x) = 1 - e^{-\lambda x}, \quad x > 0$ $P(X > s+t \mid X > t) = P(X > s)$	$E(X) = \frac{1}{\lambda}$	$\text{Var}(X) = \frac{1}{\lambda^2}$
N (μ, σ^2)	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$E(X) = \mu$	$\text{Var}(X) = \sigma^2$

Testes de hipóteses:

Parâmetro de interesse	Estatística de teste	Distribuição Nula
Média (Variância conhecida)	$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$Z \sim N(0, 1)$
Média (Variância desconhecida)	$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$	$T \sim T_{n-1}$
Proporção	$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	$Z \sim N(0, 1)$

Tabela A.1: Parâmetros para cada questão da Prova 1 . n=número de respondentes; p1, p2 e p3: probabilidade de acerto de um aluno com habilidade -3 (mínima), 0 (mediana) e 3 (maxima).

Tema	Q	a	b	c	n	% acerto	p1	p2	p3
medida_probabilidade	01	0.6	1.6	0.17	44	27	20	32	60
medida_probabilidade	02	1.1	0.7	0.21	86	50	22	49	95
medida_probabilidade	05	1.0	1.5	0.15	52	31	16	32	86
medida_probabilidade	06	1.1	-0.5	0.24	48	71	28	71	98
medida_probabilidade	07	1.0	0.8	0.20	39	46	22	46	92
medida_probabilidade	09	0.6	-0.3	0.25	95	67	39	67	90
medida_probabilidade	10	0.9	-0.1	0.24	64	66	30	65	96
propriedade_probabilidade	01	1.1	0.7	0.20	45	49	21	48	95
propriedade_probabilidade	04	1.6	-1.3	0.24	98	79	27	84	100
propriedade_probabilidade	05	0.7	1.3	0.18	90	37	21	37	74
propriedade_probabilidade	07	0.7	0.5	0.22	56	52	28	52	86
propriedade_probabilidade	08	1.6	0.3	0.20	48	54	21	55	99
propriedade_probabilidade	09	0.9	-0.7	0.26	52	77	35	75	98
propriedade_probabilidade	10	0.7	-1.2	0.25	39	85	46	82	97
probabilidade_total	01	1.2	-1.3	0.24	55	84	31	84	100
probabilidade_total	03	0.8	-1.7	0.25	132	87	48	88	99
probabilidade_total	04	1.5	-0.7	0.23	98	72	25	74	100
probabilidade_total	05	1.3	-1.1	0.24	45	80	28	81	100
probabilidade_total	07	1.4	-1.1	0.24	50	76	28	80	100
probabilidade_total	08	1.3	-0.8	0.24	48	75	27	76	99
teorema_Bayes	02	0.8	-0.1	0.25	100	66	32	65	95
teorema_Bayes	03	0.7	1.3	0.20	89	38	23	39	75
teorema_Bayes	05	0.7	1.2	0.19	45	38	21	39	79
teorema_Bayes	06	1.7	-0.7	0.23	52	73	24	74	100
teorema_Bayes	07	1.2	0.1	0.22	48	60	24	60	98
teorema_Bayes	10	1.4	-0.3	0.20	94	64	21	66	99
variaveis_aleatorias	01	1.2	0.7	0.19	46	48	20	47	96
variaveis_aleatorias	03	0.9	-0.5	0.25	83	72	33	71	97
variaveis_aleatorias	04	1.6	0.2	0.21	93	58	21	58	99
variaveis_aleatorias	05	0.5	-2.7	0.27	52	100	84	95	99
variaveis_aleatorias	06	1.0	0.0	0.23	48	62	26	62	96

Tabela A.1: Parâmetros para cada questão da Prova 1 . n=número de respondentes; p1, p2 e p3: probabilidade de acerto de um aluno com habilidade -3 (mínima), 0 (mediana) e 3 (maxima). (continuação)

Tema	Q	a	b	c	n	% acerto	p1	p2	p3
variaveis_aleatorias	07	1.1	-1.5	0.25	50	84	35	86	99
variaveis_aleatorias	09	0.9	0.6	0.22	40	50	25	51	92
variaveis_aleatorias	10	1.1	-1.2	0.24	16	88	32	81	99
distribuicao_binomial	01	1.7	-0.1	0.21	88	60	21	63	100
distribuicao_binomial	02	0.9	-0.1	0.23	45	64	29	63	95
distribuicao_binomial	04	1.1	-0.2	0.23	100	66	27	66	98
distribuicao_binomial	06	1.9	0.7	0.15	96	46	16	44	100
distribuicao_binomial	07	0.8	1.0	0.20	16	38	22	44	85
distribuicao_binomial	10	1.0	-0.7	0.25	44	77	32	75	98
distribuicao_geometrica	01	0.5	-1.6	0.26	48	90	63	87	97
distribuicao_geometrica	04	1.2	0.2	0.22	64	58	23	57	97
distribuicao_geometrica	05	1.5	0.5	0.18	97	53	19	50	98
distribuicao_geometrica	06	1.1	0.0	0.22	135	61	25	61	97
distribuicao_geometrica	10	0.7	-0.7	0.24	44	77	40	75	95
distribuicao_hipergeometrica	01	1.7	-1.2	0.23	62	77	24	81	100
distribuicao_hipergeometrica	02	0.7	0.9	0.22	50	44	25	46	84
distribuicao_hipergeometrica	03	1.4	-0.5	0.22	92	70	24	70	99
distribuicao_hipergeometrica	05	1.4	0.0	0.22	48	60	23	62	99
distribuicao_hipergeometrica	08	1.4	0.9	0.18	52	44	18	42	97
distribuicao_hipergeometrica	09	0.9	0.1	0.24	40	60	29	61	94
distribuicao_hipergeometrica	10	0.8	-1.0	0.25	84	80	40	80	98
distribuicao_poisson	01	1.2	1.1	0.16	40	35	17	39	94
distribuicao_poisson	03	1.0	1.0	0.17	89	43	19	41	91
distribuicao_poisson	04	0.5	-1.1	0.25	48	83	53	81	96
distribuicao_poisson	05	1.2	0.0	0.22	91	63	24	62	98
distribuicao_poisson	06	1.0	0.3	0.22	94	57	25	57	96
distribuicao_poisson	07	1.7	1.1	0.15	50	34	15	37	98
distribuicao_poisson	08	0.4	1.4	0.19	16	25	24	36	55
aproximacao_poisson_binomial	02	1.6	-0.9	0.22	181	74	24	77	100
aproximacao_poisson_binomial	04	1.3	-0.2	0.23	96	65	25	66	99
aproximacao_poisson_binomial	05	1.3	-0.5	0.23	48	69	26	71	99

Tabela A.1: Parâmetros para cada questão da Prova 1 . n=número de respondentes; p1, p2 e p3: probabilidade de acerto de um aluno com habilidade -3 (mínima), 0 (mediana) e 3 (maxima). (continuação)

Tema	Q	a	b	c	n	% acerto	p1	p2	p3
aproximacao_poisson_binomial	06	1.3	0.0	0.22	16	62	23	62	98
aproximacao_poisson_binomial	07	0.9	1.2	0.17	48	35	18	37	85
aproximacao_poisson_binomial	08	0.8	1.3	0.18	39	33	20	36	80

Tabela A.2: Parâmetros para cada questão da Prova 2 . n=número de respondentes; p1, p2 e p3: probabilidade de acerto de um aluno com habilidade -3 (mínima), 0 (mediana) e 3 (maxima).

Tema	Q	a	b	c	n	% acerto	p1	p2	p3
funcao_densidade	01	1.2	1.5	0.15	103	32	16	32	92
funcao_densidade	03	1.2	0.0	0.22	49	61	24	62	98
funcao_densidade	04	1.0	0.3	0.22	134	58	26	57	95
funcao_densidade	05	0.7	-0.6	0.24	49	71	37	73	95
funcao_densidade	06	1.3	-0.4	0.24	52	67	26	69	99
funcao_densidade	08	1.1	-1.1	0.25	44	80	33	81	99
distribuicao_acumulada	01	1.5	0.1	0.21	52	58	22	59	99
distribuicao_acumulada	07	0.8	-0.5	0.25	104	70	34	71	96
distribuicao_acumulada	08	1.0	0.5	0.21	49	49	24	52	93
distribuicao_acumulada	09	0.7	-0.5	0.25	47	74	38	72	95
distribuicao_acumulada	10	1.1	0.0	0.24	90	62	27	63	97
momentos	01	0.9	-0.3	0.24	49	65	30	68	97
momentos	04	0.6	1.5	0.18	44	30	20	34	67
momentos	06	1.3	-0.3	0.22	52	65	25	68	99
momentos	07	0.7	-0.8	0.25	16	81	43	76	96
momentos	08	0.6	1.5	0.19	96	34	21	34	67
momentos	10	1.5	0.8	0.16	42	43	16	43	98
distribuicao_exponencial	03	0.9	0.9	0.19	185	44	21	43	88
distribuicao_exponencial	05	1.4	1.0	0.17	52	40	18	41	97
distribuicao_exponencial	06	0.8	1.1	0.19	16	31	22	42	82
distribuicao_exponencial	07	0.7	0.9	0.21	97	43	25	45	80
distribuicao_exponencial	08	0.8	0.5	0.22	81	52	26	51	89
distribuicao_normal	01	1.3	1.2	0.17	49	37	17	37	95
distribuicao_normal	02	1.7	0.4	0.18	91	53	18	51	99

Tabela A.2: Parâmetros para cada questão da Prova 2 . n=número de respondentes; p1, p2 e p3: probabilidade de acerto de um aluno com habilidade -3 (mínima), 0 (mediana) e 3 (maxima). (continuação)

Tema	Q	a	b	c	n	% acerto	p1	p2	p3
distribuicao_normal	05	0.8	1.5	0.18	133	34	19	33	76
distribuicao_normal	06	0.9	1.5	0.18	94	33	19	34	79
distribuicao_normal	07	1.1	0.3	0.22	48	54	24	57	96
distribuicao_normal	09	1.0	-0.4	0.24	16	69	30	69	97
aproximacao_normal_binomial	01	1.1	1.2	0.18	138	38	19	38	90
aproximacao_normal_binomial	02	1.0	1.1	0.18	49	39	20	40	90
aproximacao_normal_binomial	04	1.5	-0.1	0.21	52	60	22	62	99
aproximacao_normal_binomial	05	0.5	0.7	0.23	63	51	30	50	77
aproximacao_normal_binomial	07	1.5	-0.2	0.22	42	64	24	65	99
aproximacao_normal_binomial	09	0.7	1.5	0.17	14	14	20	34	69
aproximacao_normal_binomial	10	1.2	0.8	0.19	39	44	20	46	96
distribuicao_condicional	01	1.1	-0.2	0.23	16	62	26	65	98
distribuicao_condicional	04	0.8	-0.2	0.24	44	66	32	67	95
distribuicao_condicional	05	1.2	0.1	0.22	94	60	24	60	98
distribuicao_condicional	07	1.7	0.1	0.20	96	58	21	59	99
distribuicao_condicional	08	1.4	0.3	0.19	87	52	20	53	98
distribuicao_condicional	09	1.4	0.2	0.21	42	57	22	57	99
distribuicao_condicional	10	1.3	-0.7	0.23	52	71	26	74	99
covariancia_correlacao	01	1.6	-0.7	0.22	86	70	24	74	100
covariancia_correlacao	04	1.4	-0.6	0.24	47	74	26	73	99
covariancia_correlacao	05	1.2	0.3	0.22	88	56	24	55	97
covariancia_correlacao	06	1.2	-1.1	0.25	52	79	31	80	99
covariancia_correlacao	07	0.9	0.1	0.23	48	58	27	60	95
covariancia_correlacao	08	1.4	0.0	0.22	61	59	23	61	99
covariancia_correlacao	10	0.9	0.1	0.24	49	59	29	61	94
distribuicao_media	02	1.5	0.9	0.17	42	43	17	43	98
distribuicao_media	05	0.4	1.5	0.19	88	32	23	34	54
distribuicao_media	06	0.7	1.3	0.19	111	37	22	37	73
distribuicao_media	07	2.0	0.3	0.16	45	49	16	53	100
distribuicao_media	08	1.3	1.2	0.17	49	37	18	37	94
distribuicao_media	10	1.2	0.7	0.18	96	45	18	45	96

Tabela A.2: Parâmetros para cada questão da Prova 2 . n=número de respondentes; p1, p2 e p3: probabilidade de acerto de um aluno com habilidade -3 (mínima), 0 (mediana) e 3 (maxima). (continuação)

Tema	Q	a	b	c	n	% acerto	p1	p2	p3
distribuicao_proporcao	03	1.3	0.7	0.20	47	51	21	48	97
distribuicao_proporcao	04	1.2	1.6	0.16	68	29	16	30	90
distribuicao_proporcao	06	1.9	-0.2	0.20	97	60	20	63	100
distribuicao_proporcao	07	1.0	-0.3	0.24	44	66	29	67	97
distribuicao_proporcao	08	1.1	1.2	0.17	49	33	18	37	90
distribuicao_proporcao	09	1.9	0.3	0.18	87	53	18	53	100
distribuicao_proporcao	10	0.9	1.0	0.20	39	41	22	43	86

Tabela A.3: Parâmetros para cada questão da Prova 3 . n=número de respondentes; p1, p2 e p3: probabilidade de acerto de um aluno com habilidade -3 (mínima), 0 (mediana) e 3 (maxima).

Tema	Q	a	b	c	n	% acerto	p1	p2	p3
maxima_verossimilhanca	01	0.7	1.5	0.17	48	29	19	33	69
maxima_verossimilhanca	03	0.6	0.9	0.22	141	46	27	46	78
maxima_verossimilhanca	04	0.7	1.6	0.16	47	28	18	31	69
maxima_verossimilhanca	06	1.3	0.4	0.19	91	55	20	53	98
maxima_verossimilhanca	09	0.9	1.6	0.15	84	30	16	30	76
maxima_verossimilhanca	10	0.7	1.3	0.18	16	25	20	37	77
IC_media_normal	03	1.6	-0.4	0.20	137	64	21	68	100
IC_media_normal	04	1.7	-0.5	0.22	38	63	22	70	100
IC_media_normal	05	1.7	-0.4	0.18	95	68	19	68	100
IC_media_normal	06	1.7	-0.4	0.21	110	67	22	68	100
IC_media_normal	07	1.8	-1.5	0.22	47	81	24	85	100
IC_media_t	02	1.4	-0.1	0.24	89	65	25	65	99
IC_media_t	03	0.8	0.4	0.22	49	51	26	53	91
IC_media_t	04	1.2	0.3	0.24	63	59	25	57	98
IC_media_t	05	1.4	0.4	0.21	134	57	22	54	98
IC_media_t	07	1.3	0.7	0.18	92	49	19	47	96
IC_proporcao	02	1.7	-1.1	0.21	111	74	22	79	100
IC_proporcao	03	1.9	-0.7	0.22	46	76	22	74	100
IC_proporcao	05	2.1	-1.2	0.21	51	78	21	81	100
IC_proporcao	06	2.3	-0.6	0.15	93	62	15	69	100

Tabela A.3: Parâmetros para cada questão da Prova 3 . n=número de respondentes; p1, p2 e p3: probabilidade de acerto de um aluno com habilidade -3 (mínima), 0 (mediana) e 3 (maxima). (continuação)

Tema	Q	a	b	c	n	% acerto	p1	p2	p3
IC_proporcao	08	1.9	-0.6	0.20	48	65	20	70	100
IC_proporcao	09	1.4	-0.8	0.24	38	71	26	76	99
IC_proporcao	10	1.1	-0.4	0.24	40	68	28	69	98
TH_media	05	1.2	0.6	0.20	49	47	22	50	96
TH_media	06	0.5	0.0	0.24	40	62	37	62	87
TH_media	07	1.6	0.7	0.18	47	49	18	47	99
TH_media	08	1.6	1.1	0.15	62	44	15	37	98
TH_media	09	0.4	1.0	0.22	178	44	30	44	65
TH_media	10	0.8	0.4	0.23	51	59	28	56	91
tamanho_media	01	1.7	-0.8	0.23	128	74	24	77	100
tamanho_media	02	2.1	-0.9	0.20	51	75	20	76	100
tamanho_media	03	1.5	0.0	0.20	40	57	21	61	99
tamanho_media	05	1.5	-0.2	0.21	48	62	22	64	99
tamanho_media	06	1.8	-1.2	0.22	47	77	23	81	100
tamanho_media	07	1.9	-0.5	0.19	65	62	19	69	100
tamanho_media	10	1.8	-0.6	0.21	48	67	22	72	100
TH_proporcao	02	1.0	1.2	0.16	95	38	17	36	87
TH_proporcao	03	0.6	1.4	0.19	131	37	23	37	69
TH_proporcao	04	0.9	1.1	0.19	16	31	20	41	88
TH_proporcao	05	0.8	0.8	0.21	48	46	24	47	86
TH_proporcao	08	0.4	2.1	0.15	48	19	18	25	40
TH_proporcao	09	0.8	1.9	0.13	38	16	14	25	67
TH_proporcao	10	0.6	1.2	0.20	51	39	24	39	70
pvalor_media	01	1.6	-0.2	0.20	98	61	20	64	99
pvalor_media	02	1.5	-0.2	0.23	51	69	24	65	99
pvalor_media	04	1.4	0.1	0.24	93	65	24	60	99
pvalor_media	06	1.1	1.6	0.14	97	28	14	29	86
pvalor_media	08	0.8	0.6	0.21	40	48	24	50	89
pvalor_media	09	0.8	0.9	0.22	48	46	25	46	85
tamanho_prop	02	2.3	-0.5	0.16	92	65	16	69	100
tamanho_prop	03	1.9	-0.3	0.20	49	59	20	66	100

Tabela A.3: Parâmetros para cada questão da Prova 3 . n=número de respondentes; p1, p2 e p3: probabilidade de acerto de um aluno com habilidade -3 (mínima), 0 (mediana) e 3 (maxima). (continuação)

Tema	Q	a	b	c	n	% acerto	p1	p2	p3
tamanho_prop	04	2.4	-1.1	0.18	141	72	18	78	100
tamanho_prop	05	1.4	-0.1	0.22	38	58	23	63	99
tamanho_prop	07	2.0	-1.0	0.21	91	74	22	79	100
tamanho_prop	08	0.9	-1.2	0.25	16	88	38	82	99
pvalor_proporcao	02	1.6	0.2	0.19	40	52	20	56	99
pvalor_proporcao	03	1.5	0.6	0.17	142	51	17	47	98
pvalor_proporcao	05	0.9	1.2	0.19	48	38	21	39	85
pvalor_proporcao	06	1.8	0.4	0.17	89	53	17	50	100
pvalor_proporcao	07	1.3	0.0	0.22	48	60	23	61	98
pvalor_proporcao	08	0.6	1.6	0.18	16	19	20	33	62
pvalor_proporcao	09	1.2	0.3	0.20	44	57	22	55	97