

**Análise e caracterização
estatística das provas
Vestibular Unicamp 2012**

Introdução

Histograma

Nessa sessão, serão apresentados, de forma resumida e simplificada, os tratamentos estatísticos utilizados na análise e caracterização das provas do Vestibular Unicamp.

Um histograma é uma apresentação gráfica da distribuição de frequências das notas em cada questão (ou prova). Na primeira fase, 48 pontos são atribuídos à parte de Redação e 48 pontos à parte de Conhecimentos Gerais, perfazendo o total de 96 pontos. Cada questão da parte de Conhecimentos Gerais vale um ponto. Cada texto da prova de Redação vale 16 pontos. A segunda fase foi constituída de provas realizadas em três dias consecutivos, que obedeceram à seguinte distribuição:

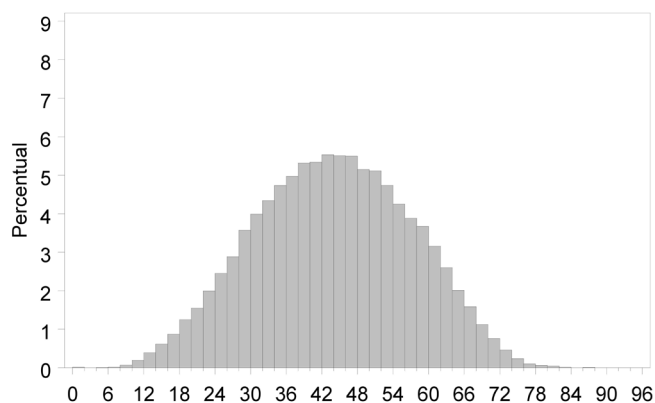
I - Prova de Língua Portuguesa e Literaturas de Língua Portuguesa e prova de Matemática, no primeiro dia (15/01/2012);

II - Prova de Ciências Humanas e Artes e prova de Língua Inglesa, no segundo dia (16/01/2012);

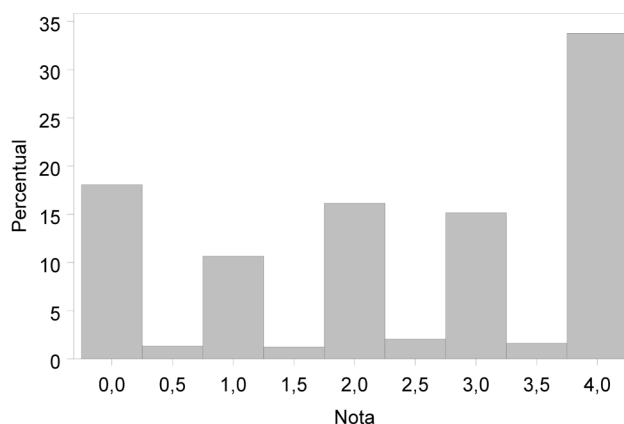
III - Prova de Ciências da Natureza, no terceiro dia (17/01/2012).

A prova de Língua Portuguesa e Literaturas de Língua Portuguesa e a prova de Matemática foram compostas de 12 questões cada, totalizando 48 pontos cada prova. A prova de Ciências Humanas e Artes foi composta por 18 questões, sendo 9 questões de História e 9 de Geografia, que totalizam 72 pontos. A de Língua Inglesa foi composta por 6 questões que totalizam 24 pontos. A prova de Ciências da Natureza foi composta por 24 questões (Biologia, Física e Química possuíram 8 questões cada), totalizando 96 pontos. Cada questão da 2ª fase vale quatro pontos, cada uma contendo dois itens (dois pontos por item) e são corrigidas duas vezes por corretores distintos. Caso a divergência de pontos seja igual a um, o valor final da questão é a média entre as duas notas, podendo gerar notas intermediárias entre os valores inteiros.

Presentes na primeira fase



Presentes na segunda fase



Ao se descrever a distribuição de notas de uma prova, a ideia é fornecer a média aritmética simples (M) e o Desvio Padrão da prova (DP) das notas das prova, de todos os candidatos presentes à prova. Como veremos adiante, essas informações também são necessárias para que se possam calcular as notas padronizadas em cada prova, e conseqüentemente, a nota padronizada de opção.

A média M é uma medida que tenta descrever o centro da distribuição das notas individuais de cada um dos N candidatos presentes à prova. Ela é expressa pela soma de todas as notas divididas pelo número de participantes e é calculada através de:

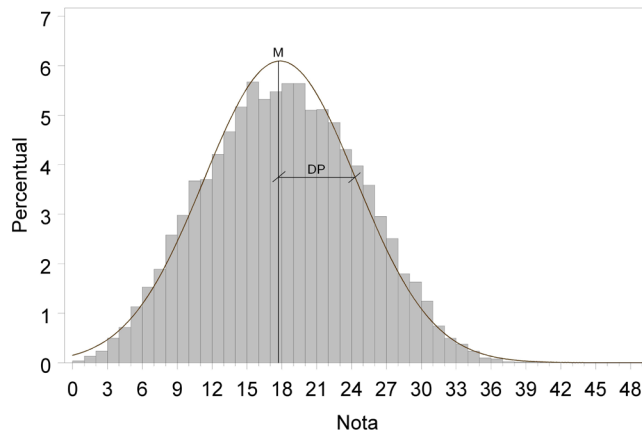
$$M = \frac{\sum_{i=1}^N X_i}{N}$$

O desvio padrão quantifica a dispersão das notas em relação à sua média e, em distribuições simétricas, é uma medida aproximada de largura da distribuição de notas da prova na metade de sua altura. É calculado pela expressão:

$$DP = \sqrt{\frac{\sum_{i=1}^N (X_i - M)^2}{N - 1}}$$

Em provas onde os valores das notas oscilam simetricamente em torno da média, a faixa limitada por um desvio-padrão em torno da média contém de 70% a 80% das notas da prova.

Presentes na segunda fase



Tabelas de frequência por item

A prova de conhecimentos gerais, pertencente à primeira fase, por ser uma prova composta de 48 questões de múltipla escolha com apenas uma alternativa correta cada será apresentada então uma tabela das frequências relativas de respostas de cada item para cada uma das 48 questões que compõem a prova de Conhecimentos Gerais para os três grupos de candidatos de nosso interesse, conforme pode ser notado no exemplo abaixo:

	porcentagem por item				
	A	B	C	D	branco
presentes na 1ª fase	24,09	25,87	26,71	23,25	0,07
presentes na 2ª fase	23,66	24,53	27,15	24,63	0,03
matriculados	23,78	23,58	28,45	24,20	0,00

Padronização das notas

No processo de correção, as provas recebem uma nota bruta. Porém, para calcular a sua classificação, são utilizadas as notas padronizadas. A padronização evita que uma prova muito difícil ou muito fácil num determinado ano faça a diferença no desempenho dos candidatos daquele ano.

O processo de padronização atribui 500 pontos à média geral de cada prova e 100 pontos para cada desvio padrão. A Nota Padronizada (NP) para cada candidato em cada prova é calculada a partir da fórmula:

$$NP_i = \frac{(X_i - M) \times 100}{DP} + 500$$

onde NP_i e X_i são respectivamente a nota padronizada e a nota bruta do i -ésimo candidato na prova em questão. A nota padronizada da primeira fase utiliza a média e o desvio padrão calculados para o grupo de todos os candidatos presentes na primeira fase. As notas padronizadas são utilizadas para compor a nota padronizada de opção (NPO).

Índices de facilidade e discriminação

O objetivo primordial do Vestibular Unicamp é o de classificar candidatos que se enquadram no perfil do aluno desejado, ou seja, que dominem os conteúdos e demonstrem capacidade de exprimir-se com clareza, de organizar suas ideias, de estabelecer relações, interpretar dados e elaborar hipóteses. Logo, uma prova de seleção, como a do Vestibular Unicamp, que abrange várias áreas de interesse - portanto envolve populações com habilidades distintas -, deve ser constituída por questões de dificuldade variada e que tenham o poder de distinguir os grupos de candidatos.

Uma das formas de se analisar o quanto as questões foram adequadas a essa separação é analisar suas características através das medidas dos índices de facilidade e de discriminação. Esses índices foram desenvolvidos para análises de questões objetivas. O Vestibular da Unicamp contempla questões objetivas e dissertativas, e por isso iremos apresentar também uma adaptação dos resultados clássicos sobre esses índices para questões dissertativas.

O Índice de Facilidade (IF) (às vezes também denominado de índice de dificuldade) mede o quanto a questão foi fácil ou difícil, para um conjunto de indivíduos que se submetem àquela prova. Este índice varia entre zero e um, sendo que o IF próximo de zero indica que a questão foi considerada muito difícil pelos candidatos; enquanto IF próximo de um mostra que a questão foi muito fácil, para aquele grupo. Esse índice pode ainda ser interpretado como a nota média da prova ou da questão numa escala de zero a um ponto.

O Índice de Discriminação (ID) mede a capacidade de uma dada questão separar os candidatos de melhor desempenho dos candidatos de pior desempenho na prova.

Os grupos de candidatos com pior e melhor desempenho serão representados pelos candidatos com nota na faixa inferior (27% com pior desempenho) e com nota na faixa superior (27% com melhor desempenho) da prova, respectivamente.

Esses índices podem ser aplicados à caracterização das provas em relação à nota final do Vestibular ou das questões em relação à nota final de uma dada prova.

Cálculo dos índices de facilidade e discriminação

No caso de provas de múltipla escolha, o cálculo do índice de facilidade é apresentado abaixo:

$$IF = PA$$

onde PA é a proporção de acerto de cada questão, entre zero e um.

No caso de provas composta por questões dissertativas, o cálculo do índice de facilidade é apresentado abaixo:

$$IF = \frac{\sum_{i=1}^N X_i}{N \times NM}$$

onde NM representa a nota máxima possível, a qual vale 4 pontos para as questões dissertativas. O valor do IF é igual à nota média da prova normalizada entre zero e um.

Para calcular o índice de discriminação, o conjunto de alunos que prestaram todas as provas é ordenado por sua nota média em cada prova. Os candidatos são então separados nas faixas inferior e superior por sua nota bruta na prova em questão, sendo que cada faixa possui 27% dos candidatos presentes em todas as provas daquela fase. Para cada prova, é calculado o índice de facilidade em cada faixa:

$$IF_{Inf} = \frac{\sum_{i=1}^{27\% \times N} X_{(i)}}{(27\% \times N) \times NM} \quad \text{e} \quad IF_{Sup} = \frac{\sum_{i=73\% \times N + 1}^N X_{(i)}}{(27\% \times N) \times NM}$$

onde $X_{(i)}$ indica o i-ésimo valor ordenado (isto é, $X_{(1)}$ refere-se ao menor valor, $X_{(2)}$ ao segundo menor, e assim por diante).

O cálculo do índice de discriminação é simplesmente a diferença entre os índices de facilidade dos dois grupos:

$$ID = IF_{Sup} - IF_{Inf}$$

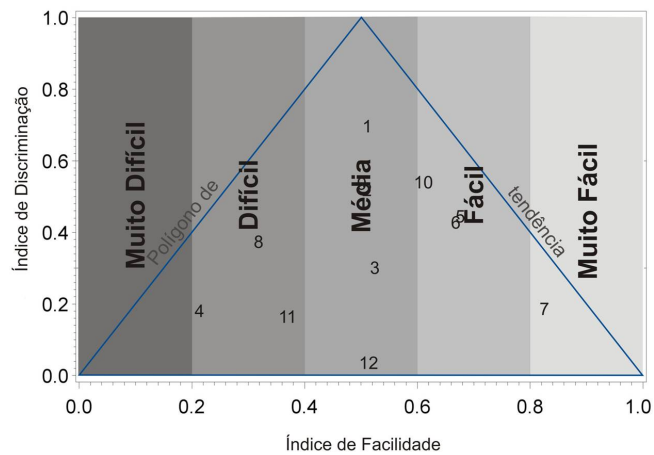
Quanto mais distante os desempenhos dos dois grupos, mais discriminativa é a prova. O ideal é obter o maior índice de discriminação possível. Um índice próximo de zero indica que os dois grupos tiveram desempenhos aproximadamente iguais, portanto a questão não discriminou.

Vale ainda lembrar que o valor obtido para o índice de discriminação indica o quanto a questão discriminou para aquele grupo que se submeteu à prova, ou seja, para outro grupo a questão pode ser mais ou menos discriminativa, ou ainda, ter o mesmo poder de discriminação.

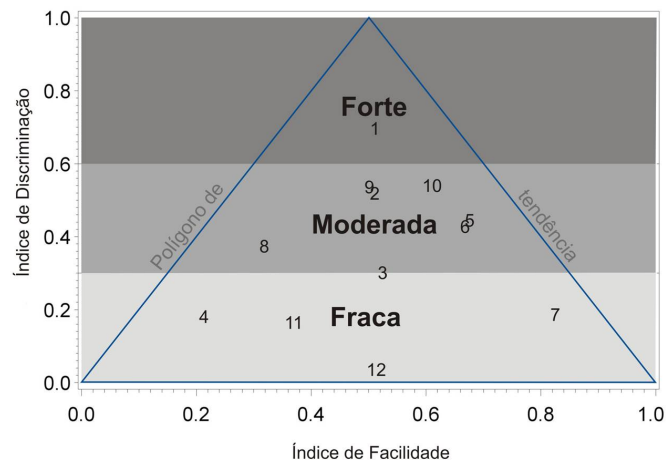
Matriz de ID vs. IF

Para simplificar e resumir a leitura desses índices, esses resultados costumam ser apresentados graficamente como matrizes, com o índice de discriminação sendo apresentado como função do índice de facilidade.

O índice de facilidade apresenta variações entre uma questão considerada muito fácil (candidatos de ambas as faixas, superior e inferior, acertam) e muito difícil (candidatos de ambas as faixas erram).



Já o índice de discriminação apresenta uma dependência maior do tipo de prova, da área do conhecimento aferida, da formulação da questão. O que se observa é que, questões de ciências da natureza, matemática e suas tecnologias apresentam índices de discriminação mais elevados que as questões relativas às ciências humanas e suas tecnologias. Os índices de discriminação devem ser observados considerando a área da questão que está sendo avaliada. Se a maioria acerta ou erra a questão, a discriminação da questão é fraca. Se o grupo com as maiores notas acerta a questão e o com as piores notas erra, a discriminação é forte.



Polígono de tendência

O polígono de tendência apresenta uma situação de discriminação máxima mais provável de ocorrer em exames, construído a partir de algumas conjecturas:

O índice de facilidade da prova é igual à média entre IF_{Sup} e IF_{Inf}

$$IF = \frac{(IF_{Sup} + IF_{Inf})}{2}$$

Para questões com IF menor que 0,5, supomos que todos os candidatos da faixa inferior zeraram na questão. Essa afirmação é matematicamente descrita por $IF_{Inf} = 0$. A partir dessas informações, podemos calcular o valor do índice de discriminação no polígono de tendências em função do valor de IF:

$$ID_{IF < 0,5} = IF_{Sup} - IF_{Inf} = 2 \times IF - 0 = 2 \times IF$$

Para questões com IF maior que 0,5 consideramos que todos os candidatos da faixa superior obtiveram nota máxima na questão ($IF_{Sup} = 1$), sendo que para essa faixa de notas calculamos o polígono de tendência por:

$$ID_{IF > 0,5} = IF_{Sup} - IF_{Inf} = 2 \times (1 - IF)$$

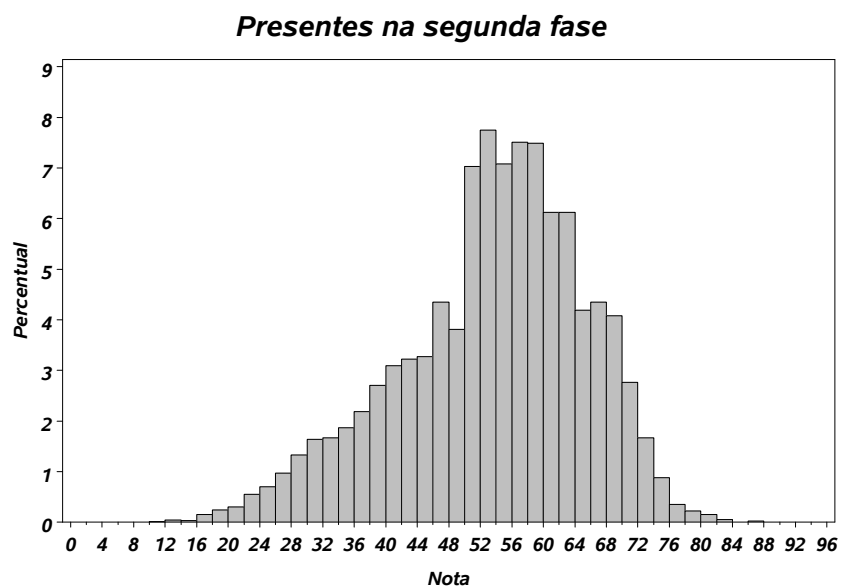
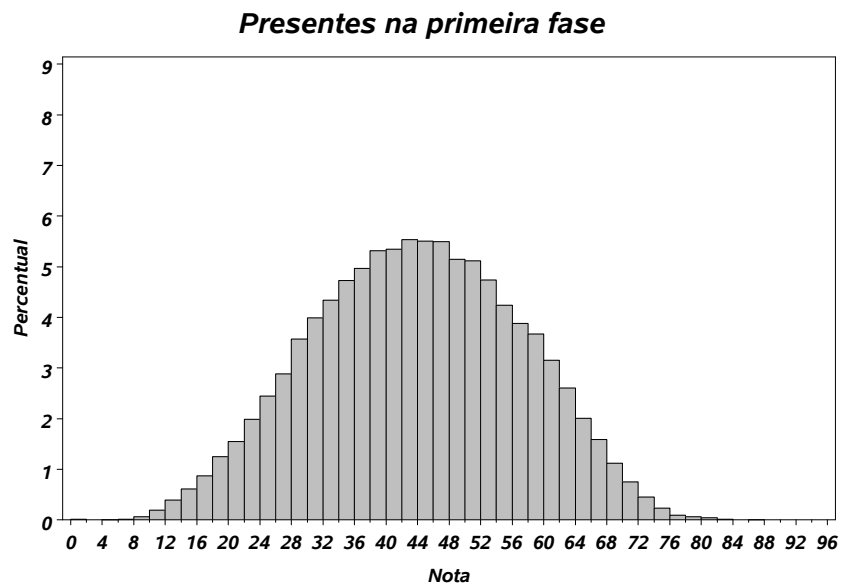
O polígono de tendência é construído a partir dos pares $ID_{IF < 0,5}$ ($ID_{IF > 0,5}$) e IF.

Primeira fase

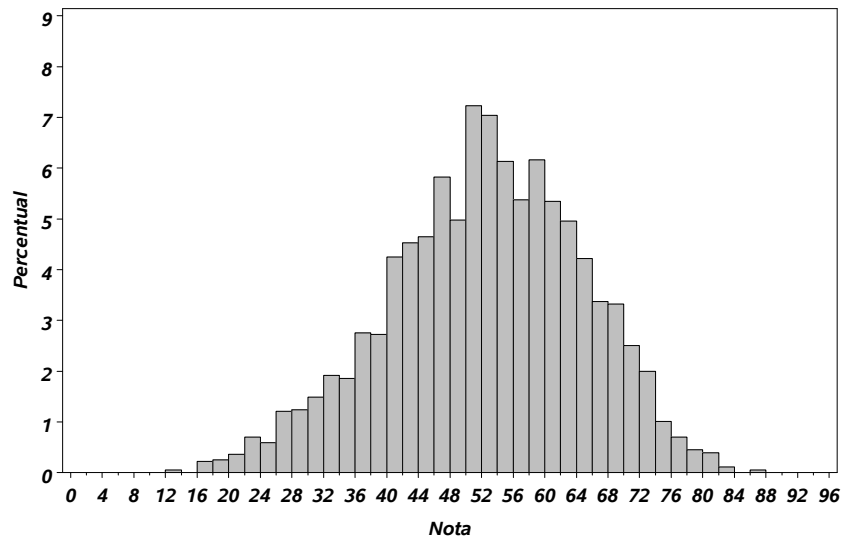
Primeira fase

Caracterização estatística

Distribuição de frequências das notas da primeira fase



Matriculados



Média e desvio padrão da primeira fase

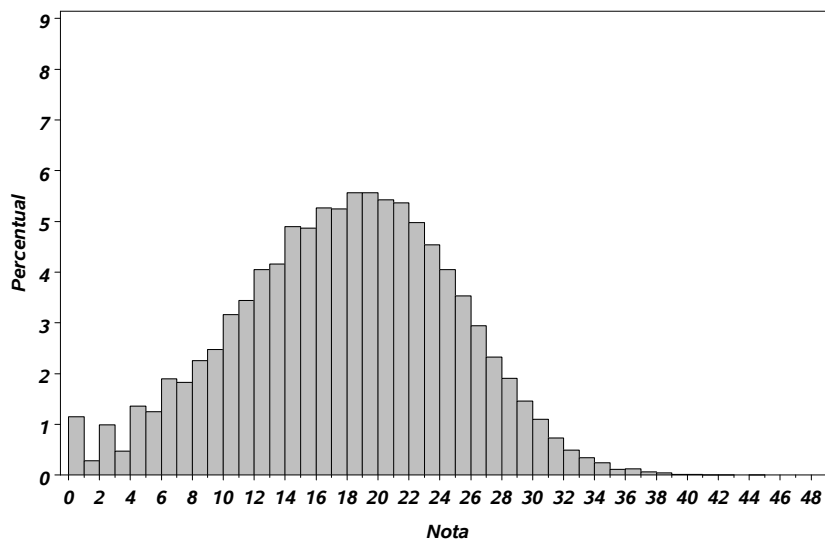
	número de candidatos	média	desvio padrão
Presentes na 1ª fase	56880	43,5	13,2
Presentes na 2ª fase	14592	53,0	11,9
Matriculados	3554	52,0	12,3

Redação

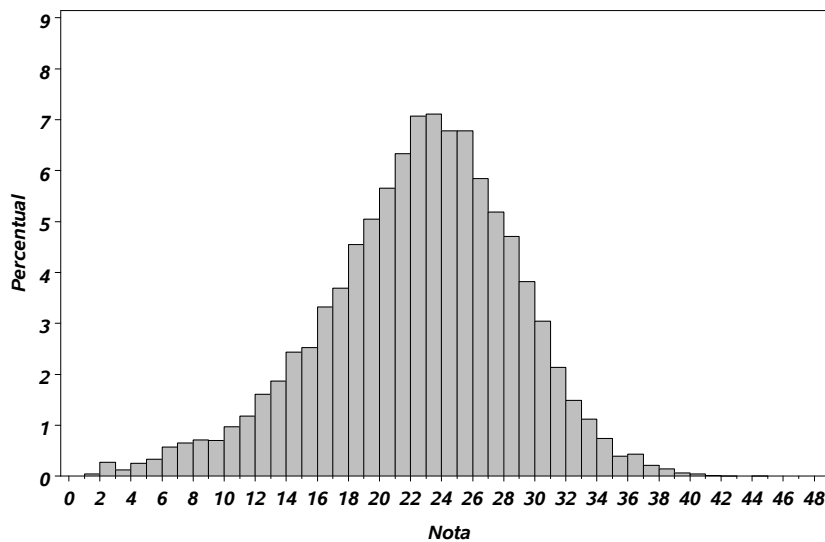
Caracterização estatística

Distribuição de frequências das notas do conjunto de redações

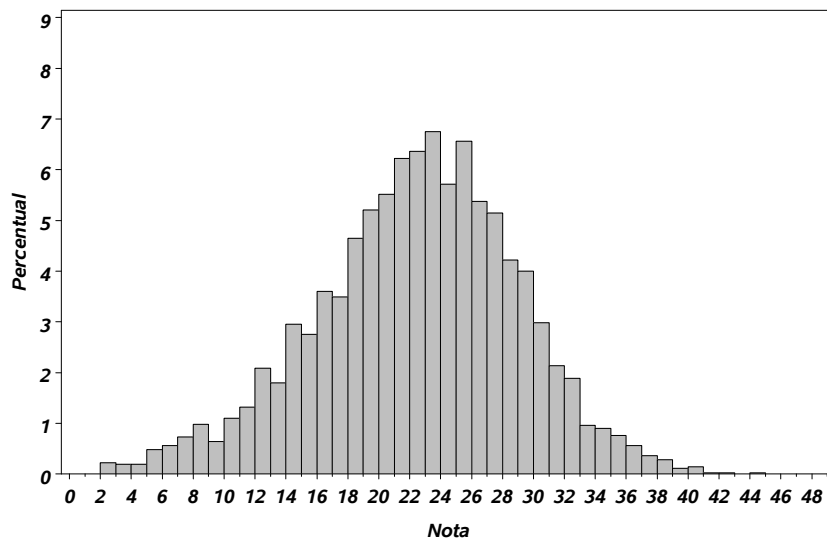
Presentes na primeira fase



Presentes na segunda fase



Matriculados



Média e desvio padrão do conjunto de redações

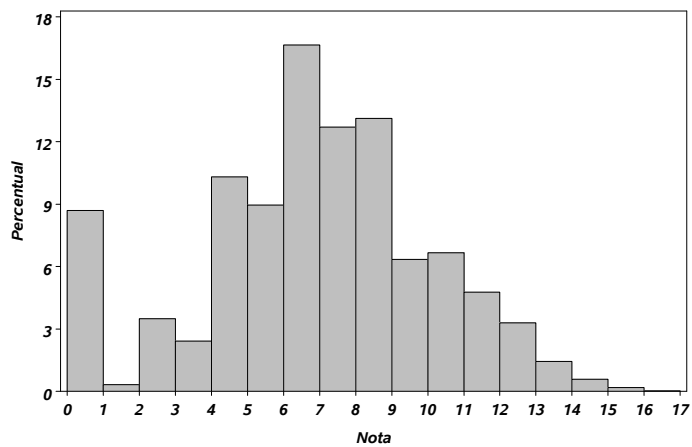
	número de candidatos	média	desvio padrão
Presentes na 1ª fase	56880	17,3	7,0
Presentes na 2ª fase	14592	22,0	6,2
Matriculados	3554	22,0	6,6

Média e desvio padrão por redação

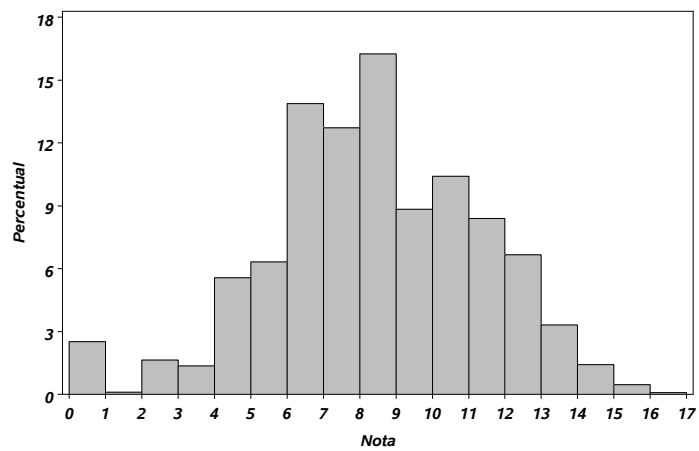
texto	presentes 1ª fase		presentes na 2ª fase		matriculados	
	média	desvio padrão	média	desvio padrão	média	desvio padrão
1	6,4	3,2	7,9	2,9	7,9	3,0
2	5,7	3,1	7,1	2,8	7,2	2,8
3	5,2	3,5	7,0	3,3	6,8	3,5

TEXTO 1

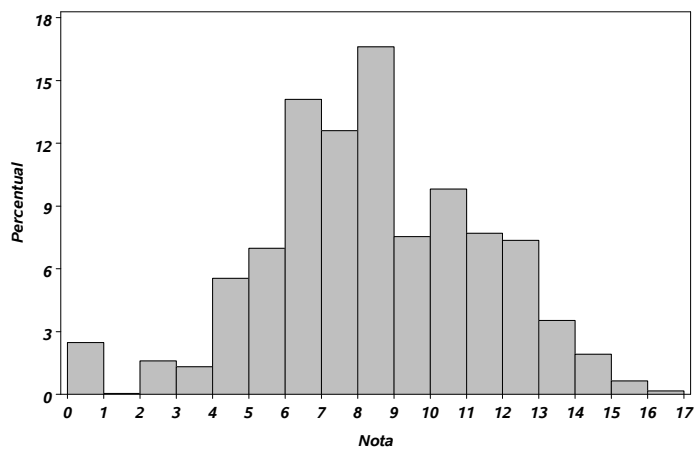
Presentes na primeira fase



Presentes na segunda fase

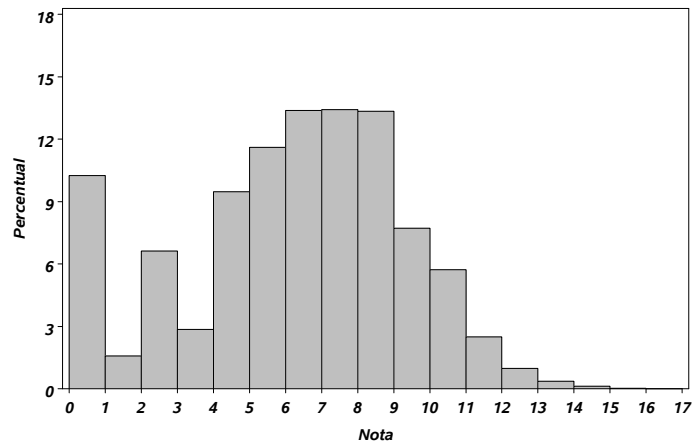


Matriculados

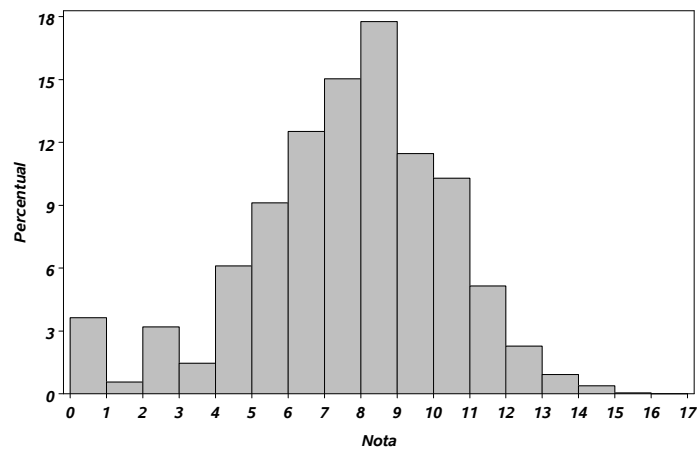


TEXTO 2

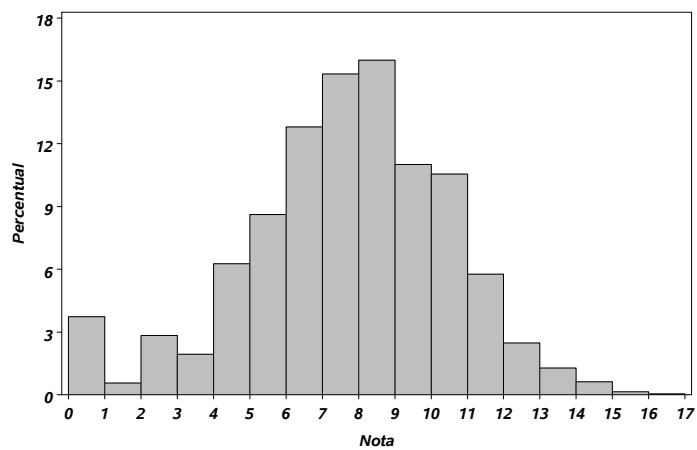
Presentes na primeira fase



Presentes na segunda fase

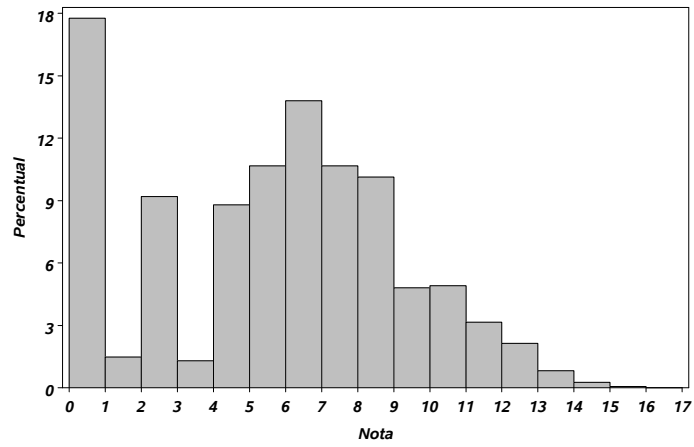


Matriculados

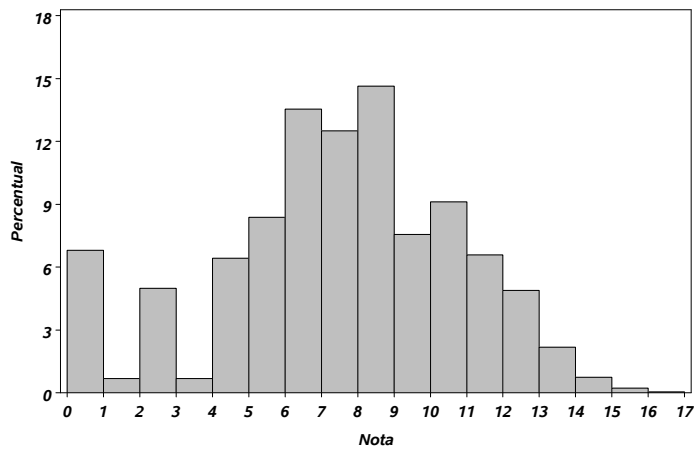


TEXTO 3

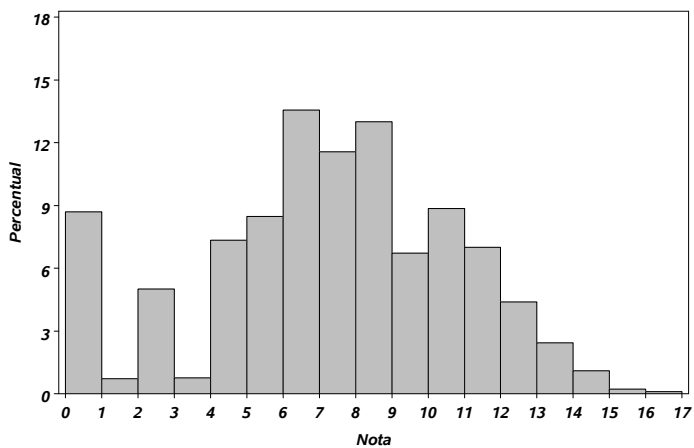
Presentes na primeira fase



Presentes na segunda fase



Matriculados

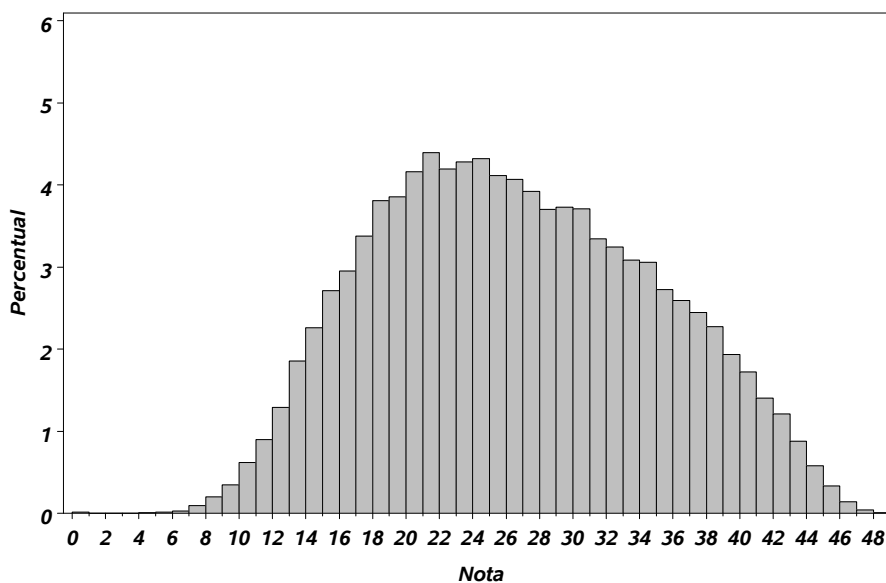


Questões Gerais

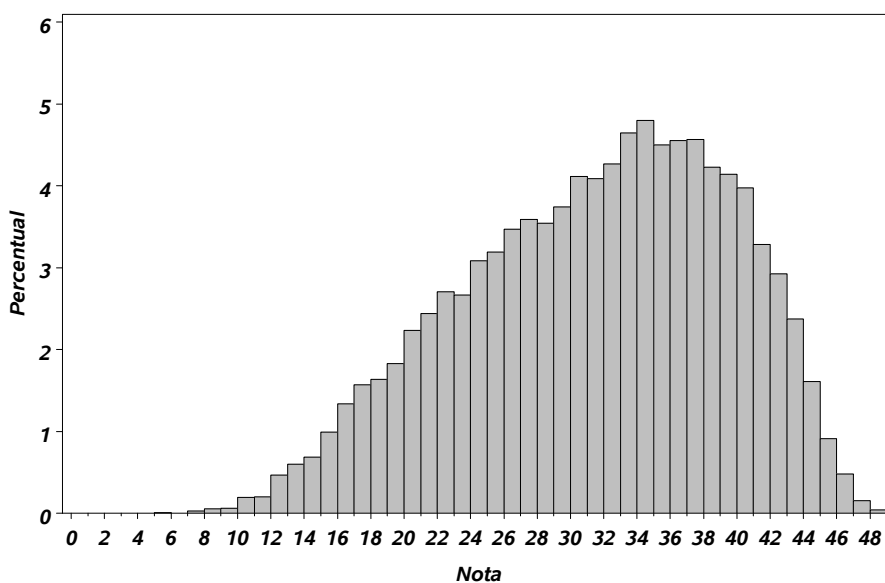
Caracterização estatística

Distribuição de frequências das notas do conjunto de questões gerais

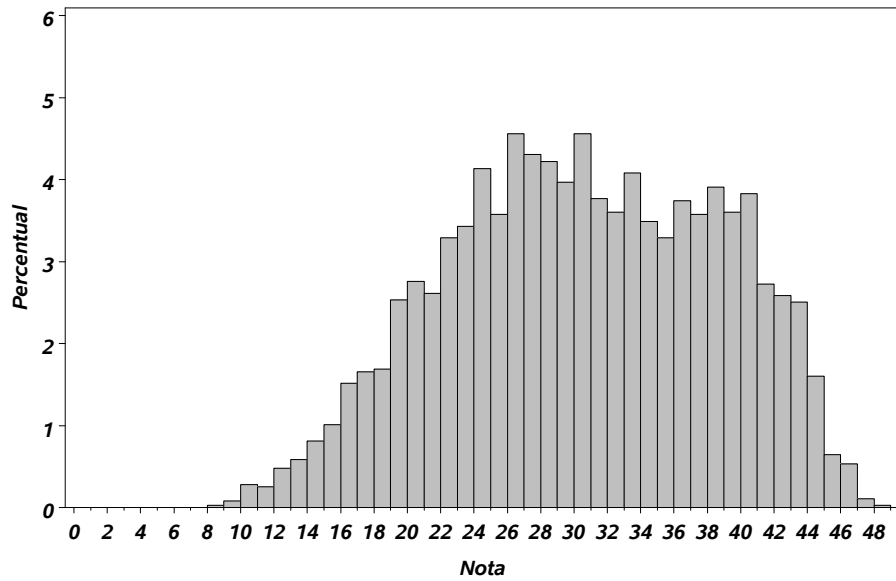
Presentes na primeira fase



Presentes na segunda fase



Matriculados



Média e desvio padrão do conjunto de questões gerais

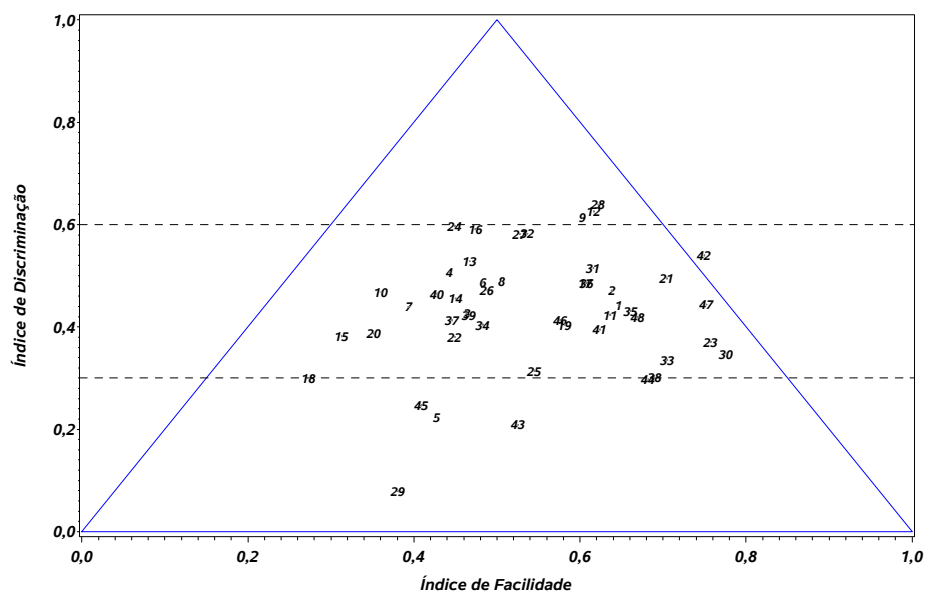
	número de candidatos	média	desvio padrão
Presentes na 1ª fase	56880	26,1	8,3
Presentes na 2ª fase	14592	31,0	8,0
Matriculados	3554	30,0	8,1

Índices de facilidade e discriminação das questões gerais

questão	disciplina	IF	ID
1	Matemática	0,647	0,442
2	Matemática	0,639	0,471
3	Matemática	0,464	0,427
4	Matemática	0,443	0,507
5	Matemática	0,428	0,224
6	Matemática	0,483	0,487
7	Matemática	0,394	0,440
8	Matemática	0,506	0,490
9	Matemática	0,603	0,614
10	Matemática	0,361	0,468
11	Física	0,636	0,424
12	Física	0,617	0,626
13	Física	0,467	0,529
14	Física	0,451	0,456
15	Matemática	0,313	0,382
16	Física	0,474	0,592
17	C. Biológicas	0,607	0,487
18	Matemática	0,273	0,300
19	C. Biológicas	0,582	0,405
20	C. Biológicas	0,352	0,387
21	C. Biológicas	0,705	0,495
22	C. Biológicas	0,449	0,381
23	C. Biológicas	0,758	0,371
24	Química	0,449	0,597

questão	disciplina	IF	ID
25	Química	0,545	0,315
26	Química	0,489	0,472
27	Física	0,528	0,582
28	Química	0,622	0,639
29	Química	0,381	0,079
30	Química	0,777	0,348
31	História	0,616	0,515
32	História	0,537	0,582
33	História	0,705	0,336
34	História	0,483	0,405
35	Filosofia	0,661	0,432
36	História	0,608	0,486
37	História	0,446	0,414
38	História	0,690	0,301
39	História	0,467	0,422
40	Geografia	0,428	0,465
41	Geografia	0,624	0,395
42	Geografia	0,750	0,540
43	Geografia	0,526	0,209
44	Geografia	0,683	0,297
45	Sociologia	0,409	0,247
46	Geografia	0,576	0,414
47	Geografia	0,752	0,444
48	Geografia	0,670	0,420

Múltipla Escolha



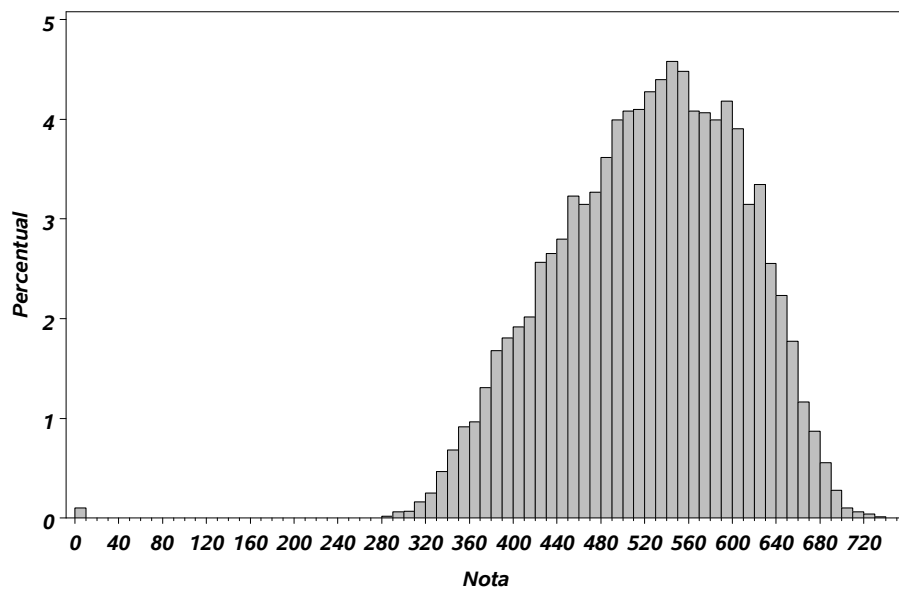
Segunda fase

Segunda fase

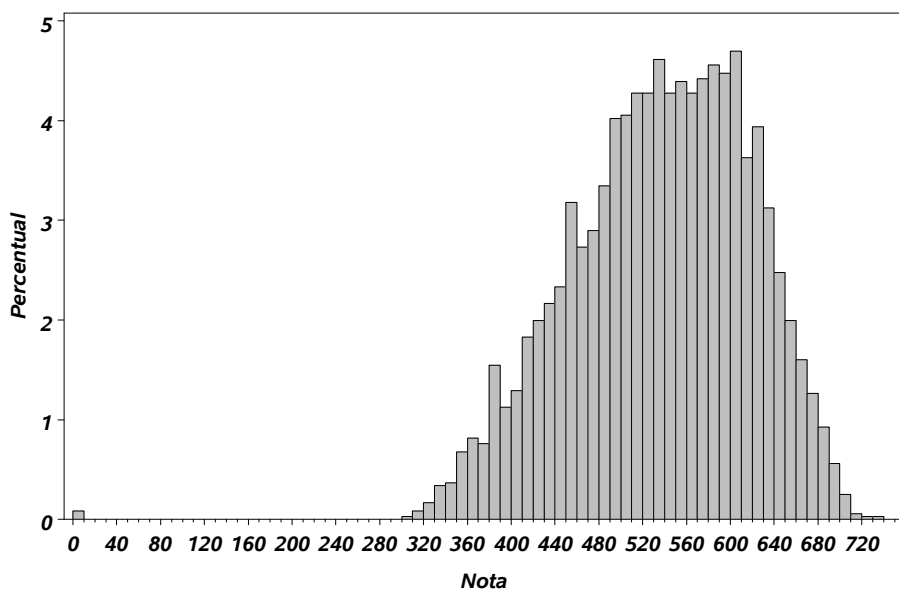
Caracterização estatística

Distribuição de frequências das notas padronizadas de opção (NPO) para a primeira opção

Presentes na segunda fase



Matriculados



Média e desvio padrão

	candidatos	média	desvio padrão
presentes na 2ª fase	14592	523	84
matriculados	3554	537	80

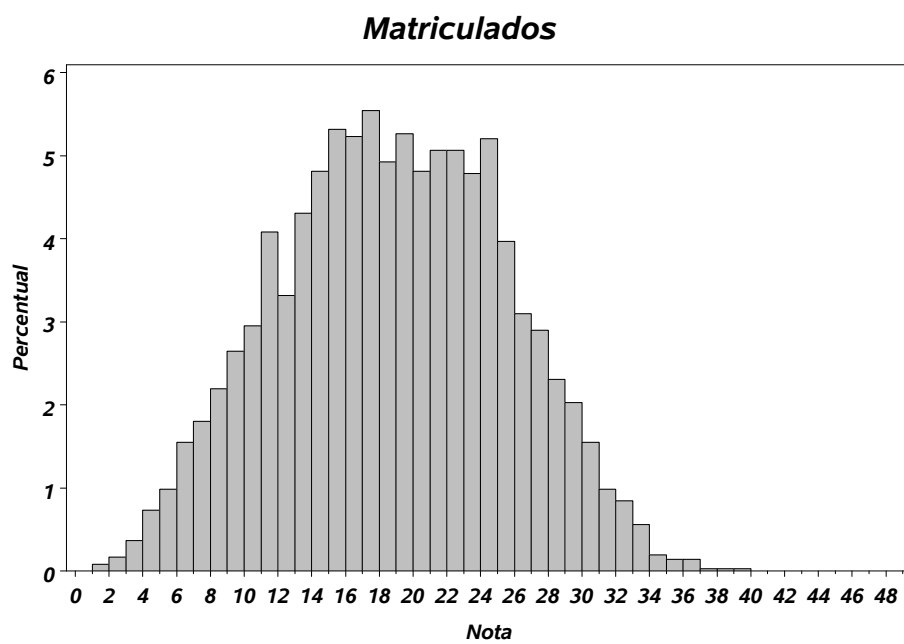
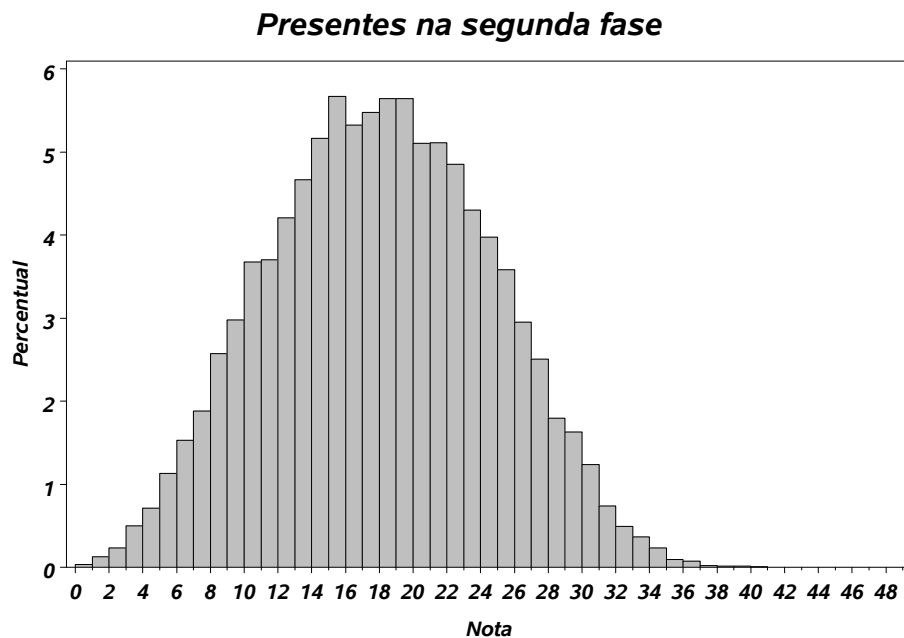
Média e desvio padrão da nota bruta por disciplina

	pontos possíveis	presentes na 2ª fase		matriculados	
		média	desvio padrão	média	desvio padrão
Língua Portuguesa	48	17,8	6,5	18,5	6,7
Matemática	48	20,4	12,1	20,3	12,6
História	36	19,3	5,9	19,6	5,7
Geografia	36	16,6	4,7	17,1	4,6
Língua Estrangeira	24	12,9	5,3	12,9	5,5
Física	32	13,1	9,4	13,0	9,8
Biologia	32	16,2	6,8	16,2	6,7
Química	32	11,6	5,6	12,0	5,6

Língua Portuguesa e Literaturas de Língua Portuguesa

Caracterização estatística

Distribuição de frequências das notas



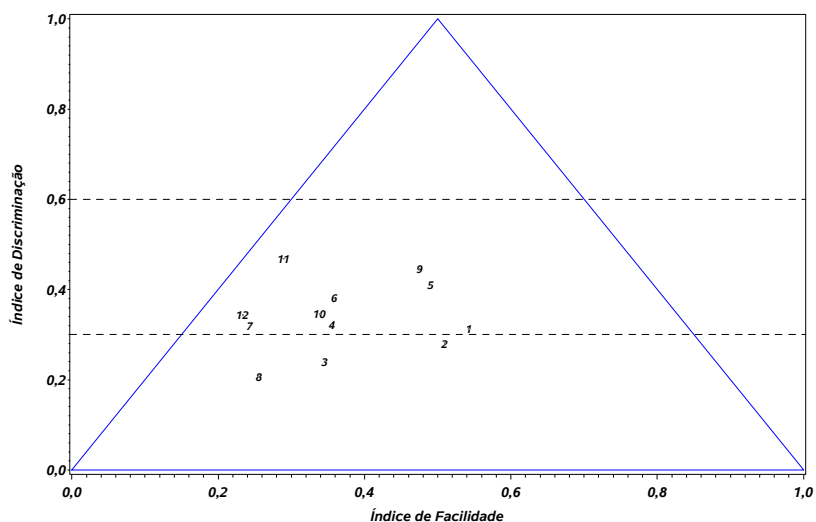
Média e desvio padrão

	candidatos	média	desvio padrão
presentes na 2ª fase	14592	17,8	6,5
matriculados	3554	18,5	6,7

Índices de facilidade e discriminação

questão	IF	ID
1	0,543	0,313
2	0,510	0,281
3	0,346	0,241
4	0,356	0,323
5	0,490	0,410
6	0,359	0,381
7	0,243	0,320
8	0,255	0,207
9	0,475	0,446
10	0,339	0,346
11	0,290	0,469
12	0,234	0,344

Língua Portuguesa e Literaturas de Língua Portuguesa



Média e desvio padrão por questão

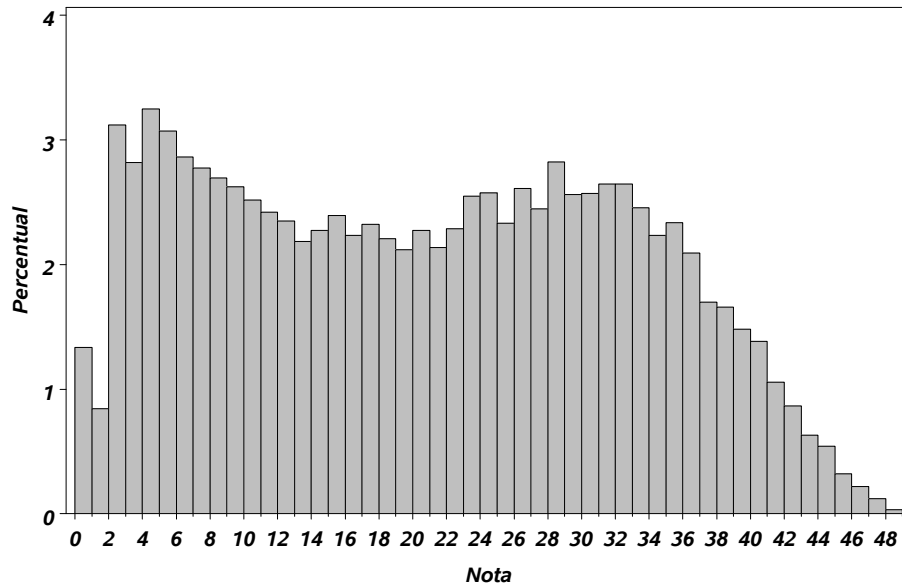
questão	presentes na 2ª fase		matriculados	
	média	desvio padrão	média	desvio padrão
1	2,2	1,0	2,2	1,1
2	2,0	0,9	2,1	0,9
3	1,4	0,9	1,4	0,9
4	1,4	1,0	1,5	1,0
5	2,0	1,1	2,0	1,1
6	1,4	1,0	1,5	1,0
7	1,0	1,0	1,1	1,0
8	1,0	0,8	1,1	0,8
9	1,9	1,2	2,0	1,2
10	1,4	1,0	1,4	1,0
11	1,2	1,2	1,3	1,2
12	0,9	1,0	1,0	1,0

Matemática

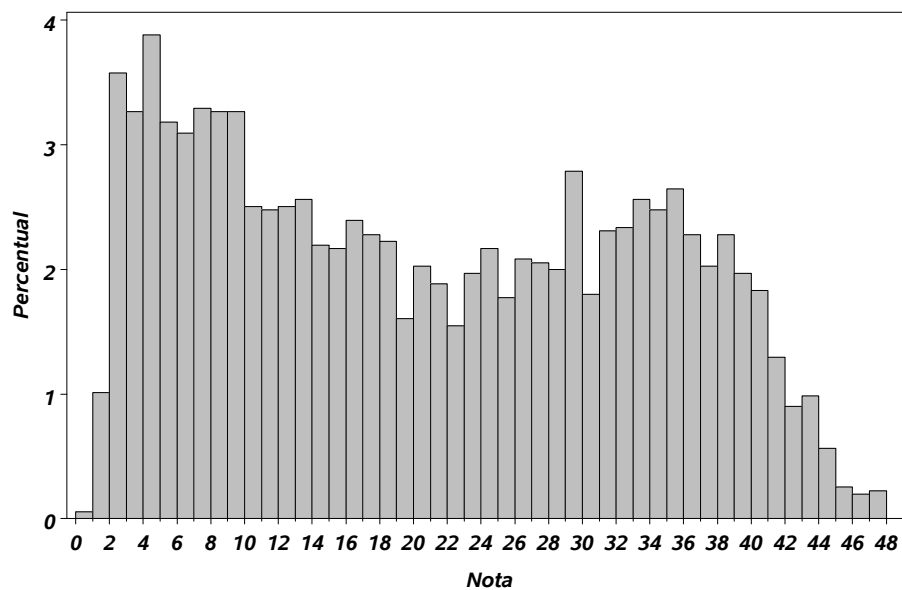
Caracterização estatística

Distribuição de frequências das notas

Presentes na segunda fase



Matriculados

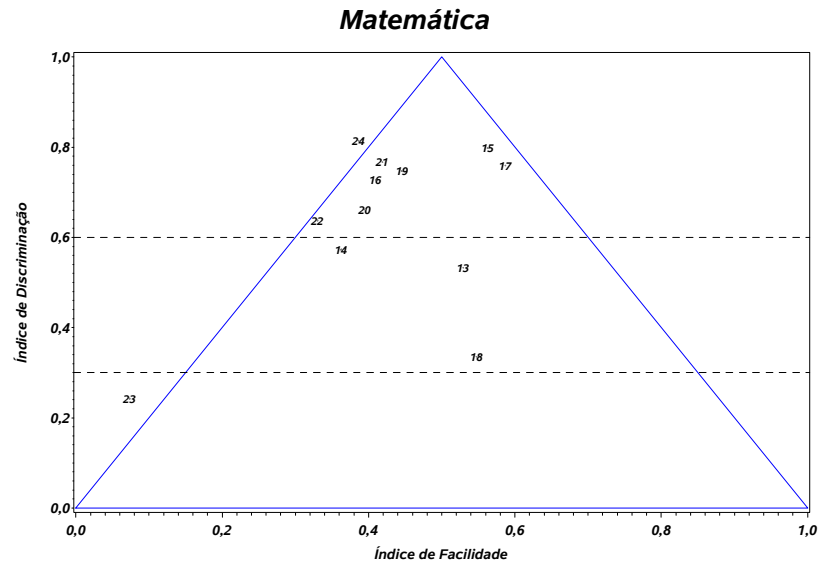


Média e desvio padrão

	candidatos	média	desvio padrão
presentes na 2ª fase	14592	20,4	12,1
matriculados	3554	20,3	12,6

Índices de facilidade e discriminação

questão	IF	ID
13	0,530	0,533
14	0,363	0,573
15	0,563	0,798
16	0,410	0,728
17	0,587	0,760
18	0,548	0,335
19	0,446	0,747
20	0,395	0,662
21	0,419	0,767
22	0,331	0,638
23	0,074	0,243
24	0,387	0,814



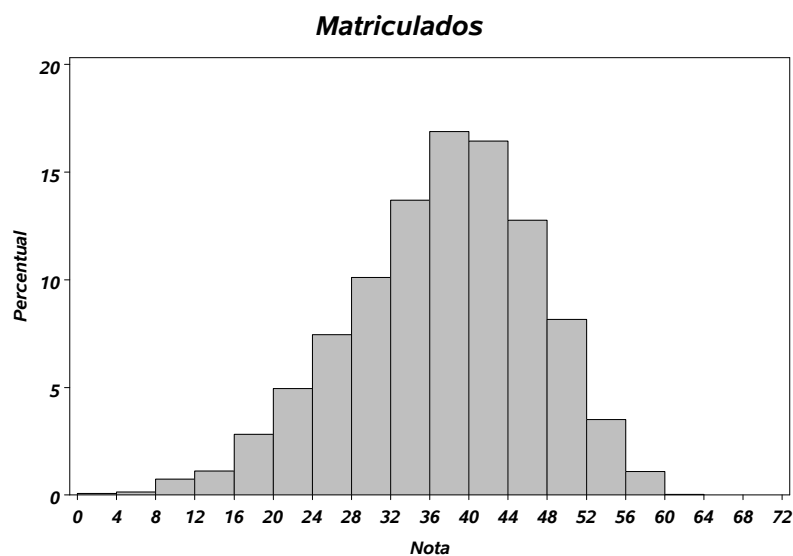
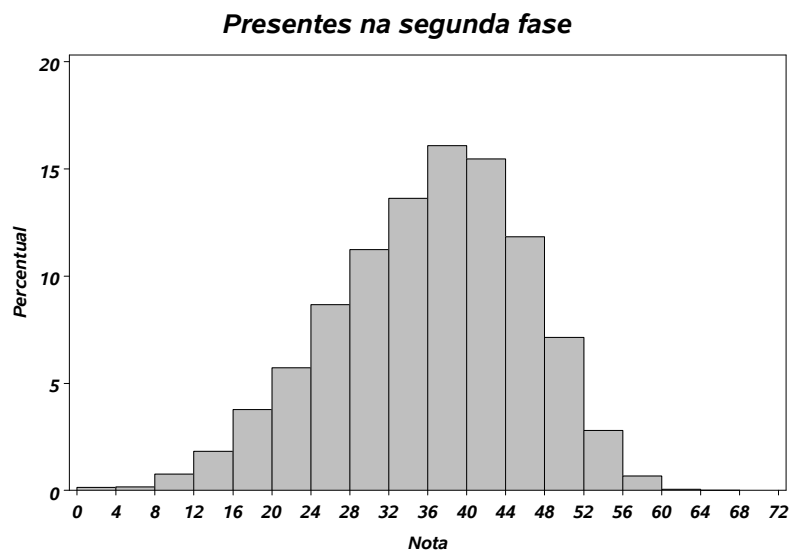
Média e desvio padrão por questão

questão	presentes na 2ª fase		matriculados	
	média	desvio padrão	média	desvio padrão
1	2,1	1,3	2,1	1,3
2	1,5	1,5	1,5	1,5
3	2,3	1,6	2,3	1,6
4	1,7	1,6	1,6	1,6
5	2,4	1,5	2,3	1,5
6	2,2	1,1	2,2	1,1
7	1,8	1,5	1,8	1,5
8	1,6	1,4	1,6	1,4
9	1,7	1,5	1,7	1,5
10	1,3	1,4	1,4	1,4
11	0,3	0,8	0,3	0,9
12	1,6	1,7	1,5	1,7

Ciências Humanas e Artes*

Caracterização estatística

Distribuição de frequências das notas



	candidatos	média	desvio padrão
presentes na 2ª fase	14592	35,9	9,9
matriculados	3554	37,0	9,6

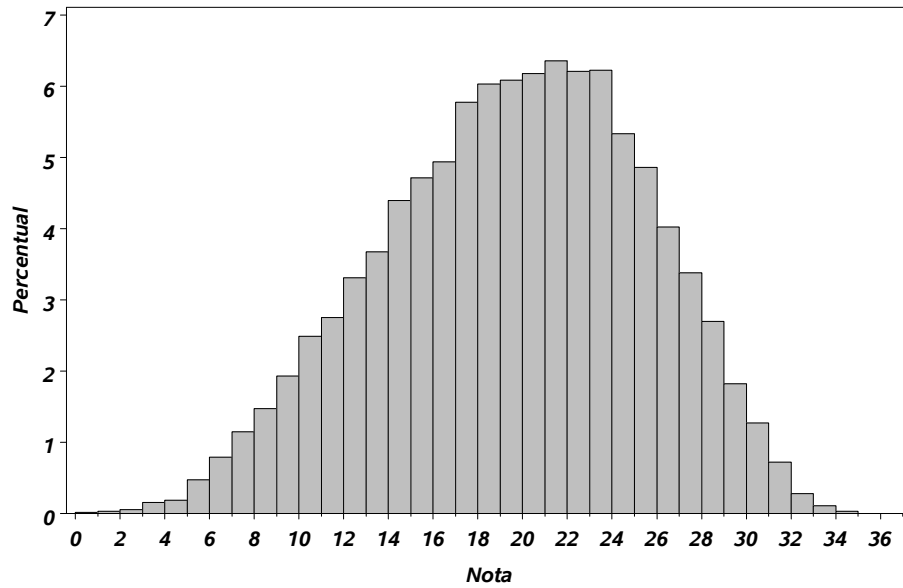
* Essa prova incorpora as disciplinas de Geografia e História

História

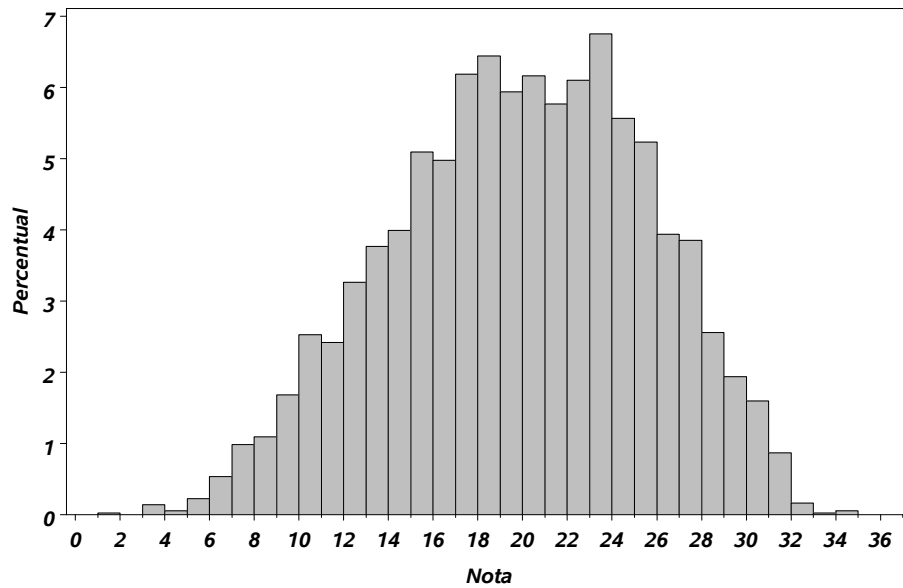
Caracterização estatística

Distribuição de frequências das notas

Presentes na segunda fase



Matriculados

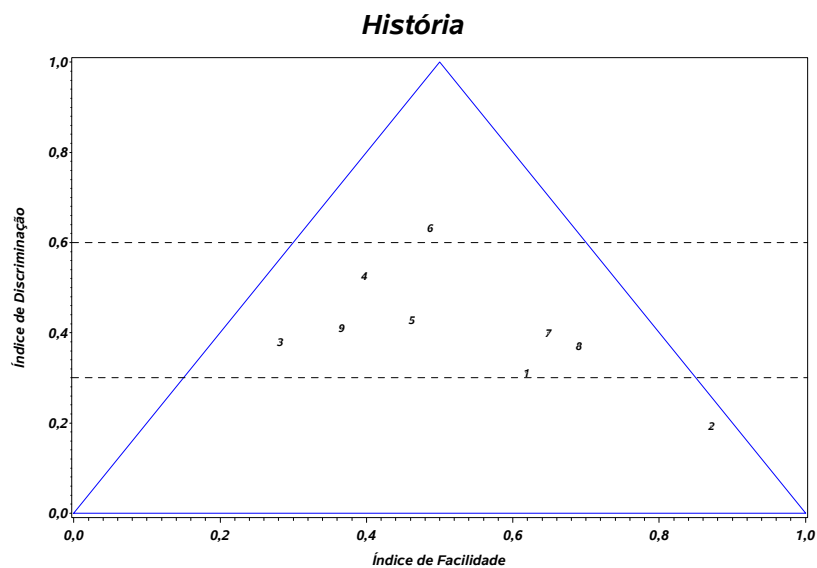


Média e desvio padrão

	candidatos	média	desvio padrão
presentes na 2ª fase	14592	19,3	5,9
matriculados	3554	19,6	5,7

Índices de facilidade e discriminação

questão	IF	ID
1	0,619	0,312
2	0,871	0,193
3	0,282	0,380
4	0,397	0,526
5	0,462	0,429
6	0,487	0,633
7	0,648	0,400
8	0,690	0,373
9	0,366	0,411



Média e desvio padrão por questão

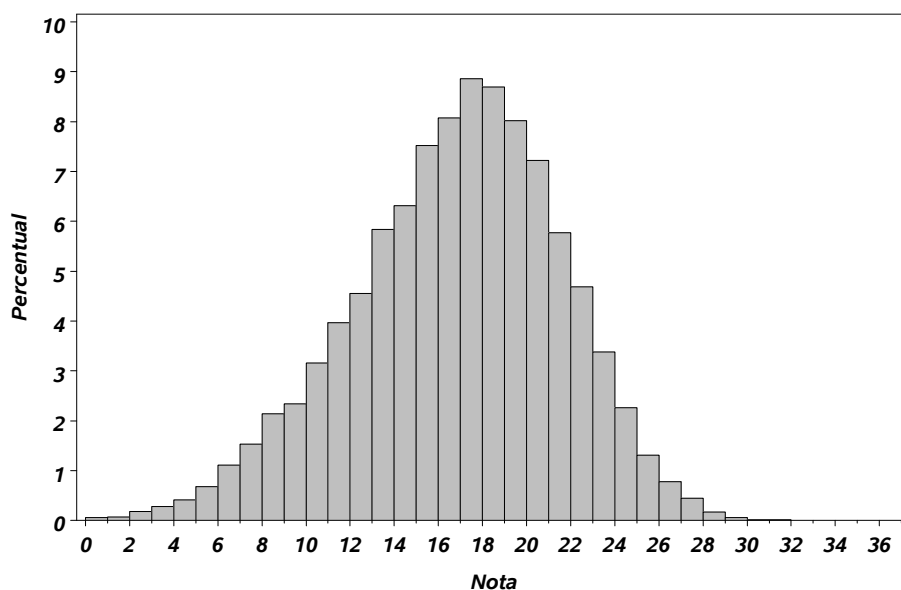
questão	presentes na 2ª fase		matriculados	
	média	desvio padrão	média	desvio padrão
1	2,5	1,0	2,5	0,9
2	3,5	0,6	3,5	0,6
3	1,1	1,0	1,1	1,0
4	1,6	1,2	1,6	1,2
5	1,9	1,1	1,9	1,1
6	2,0	1,4	2,0	1,4
7	2,6	1,2	2,6	1,2
8	2,8	1,0	2,8	1,0
9	1,5	1,0	1,5	1,0

Geografia

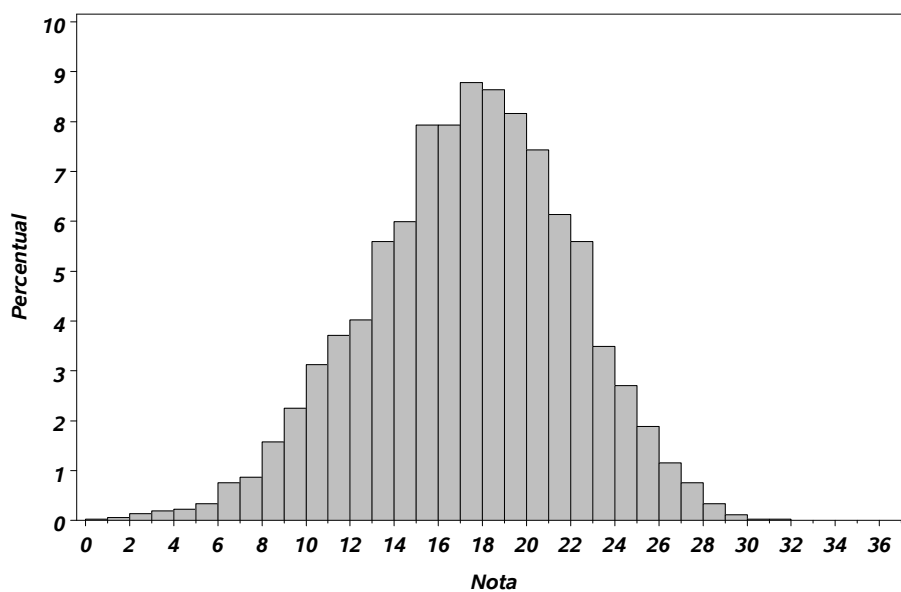
Caracterização estatística

Distribuição de frequências das notas

Presentes na segunda fase



Matriculados

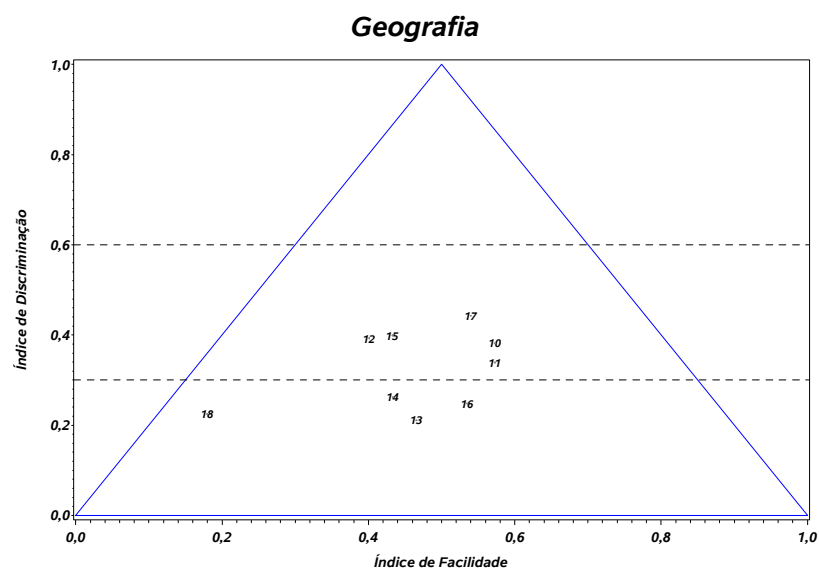


Média e desvio padrão

	candidatos	média	desvio padrão
presentes na 2ª fase	14592	16,6	4,7
matriculados	3554	17,1	4,6

Índices de facilidade e discriminação

questão	IF	ID
10	0,573	0,382
11	0,573	0,339
12	0,401	0,393
13	0,466	0,211
14	0,434	0,262
15	0,433	0,399
16	0,535	0,248
17	0,540	0,443
18	0,180	0,225



Média e desvio padrão por questão

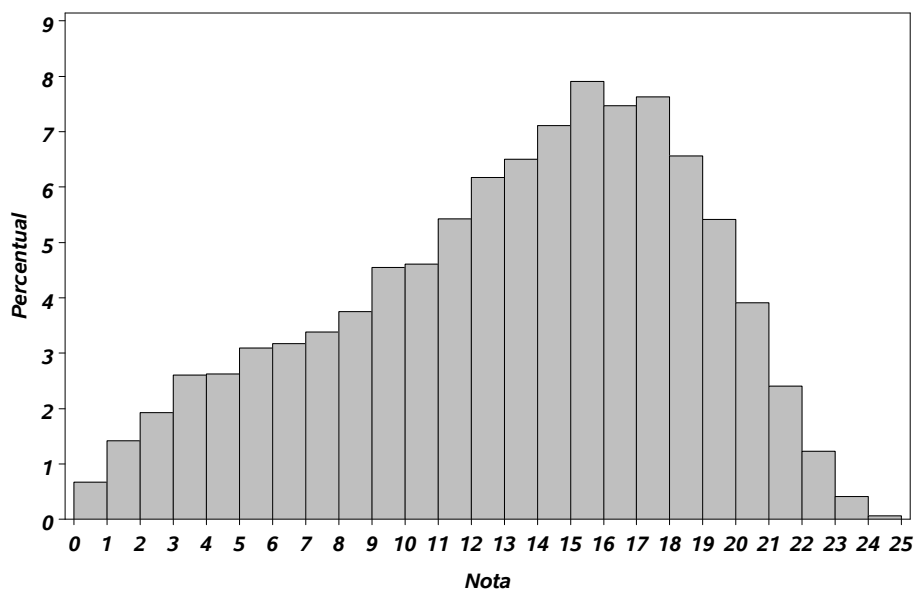
questão	presentes na 2ª fase		matriculados	
	média	desvio padrão	média	desvio padrão
1	2,3	1,1	2,4	1,1
2	2,3	1,0	2,4	1,0
3	1,6	1,0	1,6	1,0
4	1,9	0,7	1,9	0,7
5	1,7	0,9	1,8	0,9
6	1,7	1,1	1,7	1,1
7	2,1	0,8	2,2	0,8
8	2,2	1,1	2,2	1,1
9	0,7	0,7	0,8	0,7

Língua Inglesa

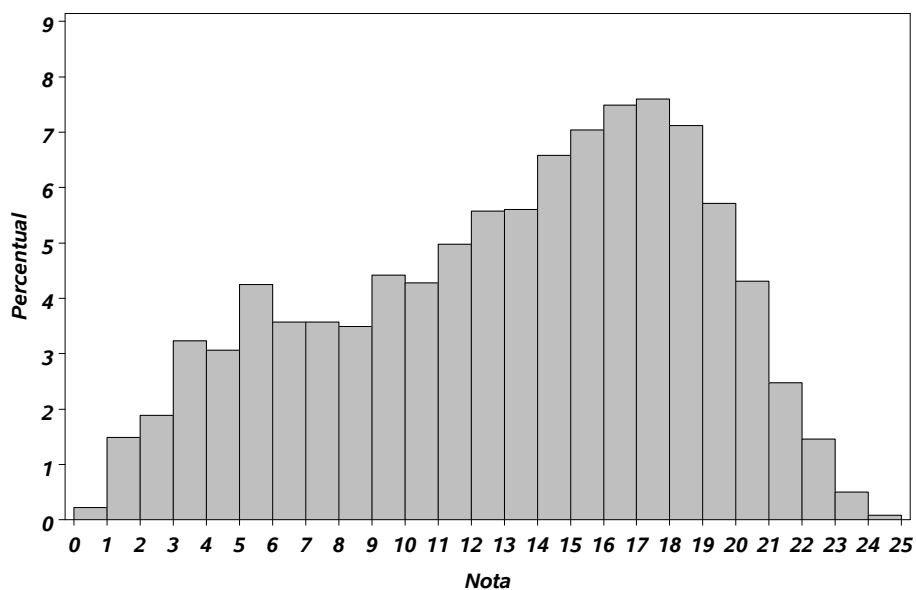
Caracterização estatística

Distribuição de frequências das notas

Presentes na segunda fase



Matriculados

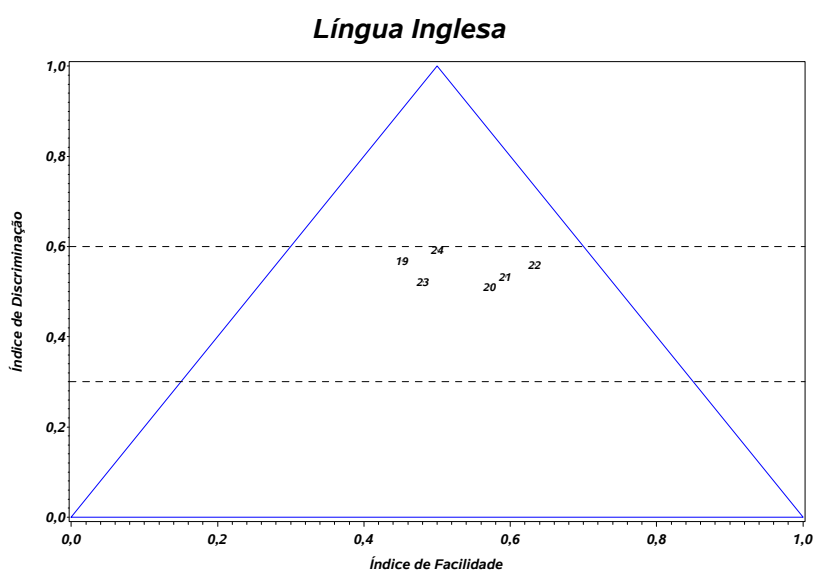


Média e desvio padrão

	candidatos	média	desvio padrão
presentes na 2ª fase	14592	12,9	5,3
matriculados	3554	12,9	5,5

Índices de facilidade e discriminação

questão	IF	ID
19	0,453	0,569
20	0,572	0,512
21	0,594	0,534
22	0,634	0,560
23	0,481	0,522
24	0,501	0,592



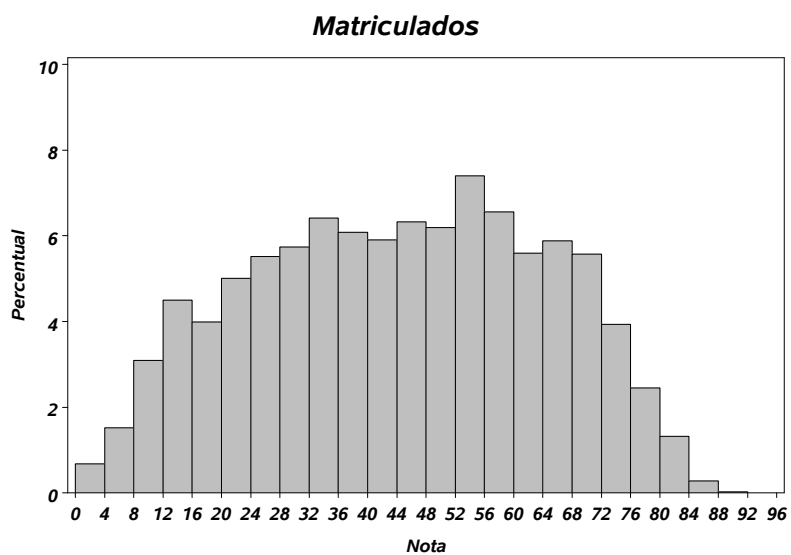
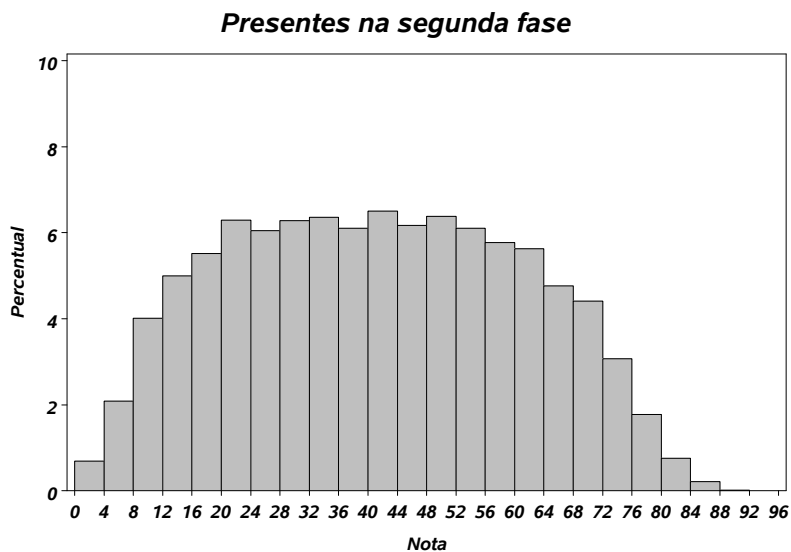
Média e desvio padrão por questão

questão	presentes na 2ª fase		matriculados	
	média	desvio padrão	média	desvio padrão
1	1,8	1,2	1,8	1,2
2	2,3	1,1	2,3	1,1
3	2,4	1,2	2,4	1,2
4	2,5	1,2	2,5	1,2
5	1,9	1,1	1,9	1,2
6	2,0	1,4	2,0	1,4

Ciências da Natureza*

Caracterização estatística

Distribuição de frequências das notas



	candidatos	média	desvio padrão
presentes na 2ª fase	14592	40,9	19,7
matriculados	3554	44,0	19,8

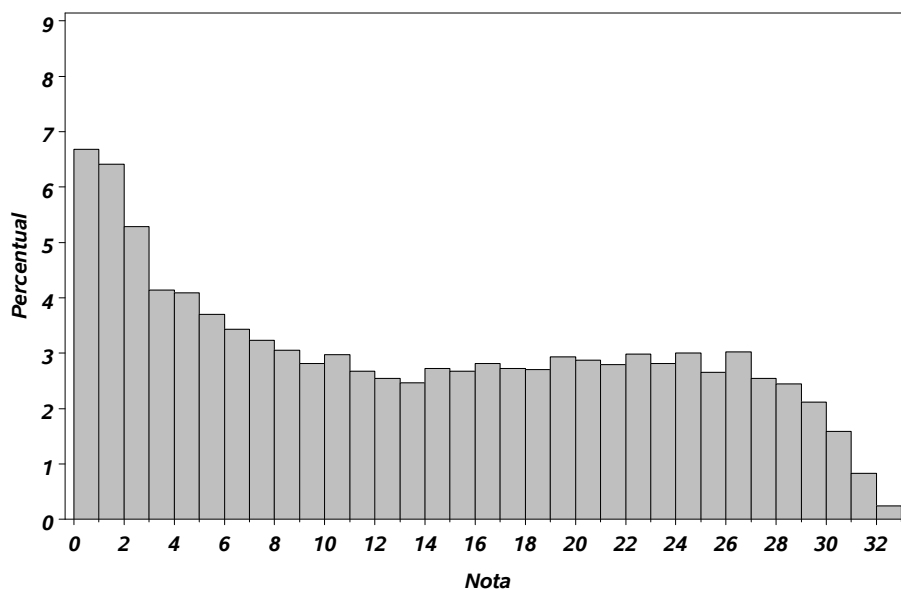
* Essa prova incorpora as disciplinas de Ciências Biológicas, Física e Química.

Física

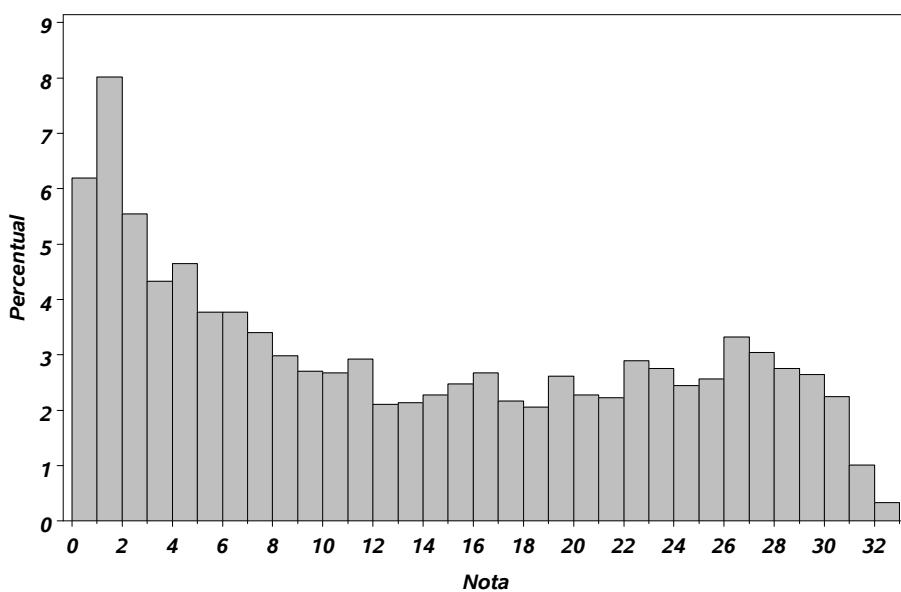
Caracterização estatística

Distribuição de frequências das notas

Presentes na segunda fase



Matriculados

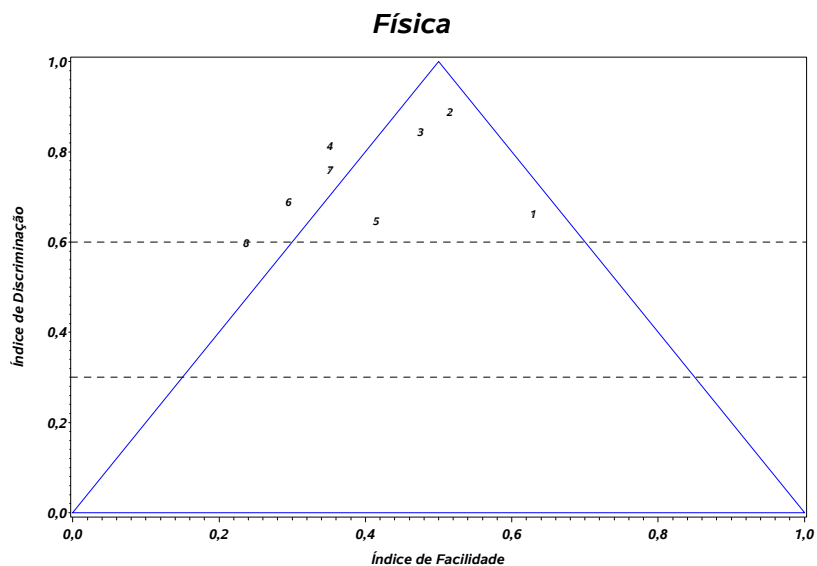


Média e desvio padrão

	candidatos	média	desvio padrão
presentes na 2ª fase	14592	13,1	9,4
matriculados	3554	13,0	9,8

Índices de facilidade e discriminação

questão	IF	ID
1	0,629	0,663
2	0,516	0,890
3	0,475	0,845
4	0,352	0,815
5	0,415	0,648
6	0,295	0,689
7	0,351	0,761
8	0,237	0,598



Média e desvio padrão por questão

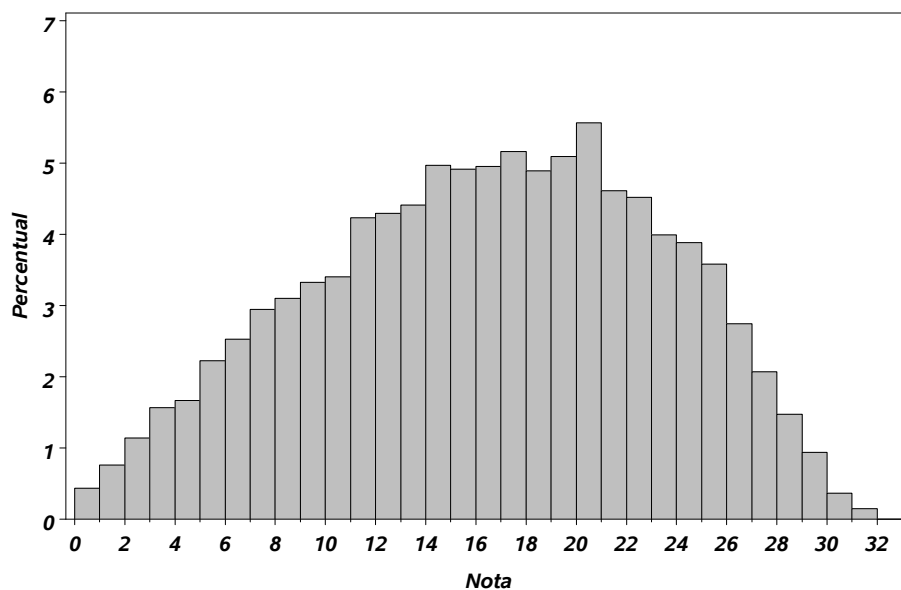
questão	presentes na 2ª fase		matriculados	
	média	desvio padrão	média	desvio padrão
1	2,5	1,3	2,5	1,3
2	2,1	1,7	2,0	1,7
3	1,9	1,6	1,9	1,6
4	1,4	1,6	1,4	1,6
5	1,7	1,3	1,7	1,3
6	1,2	1,4	1,3	1,5
7	1,4	1,5	1,4	1,5
8	1,0	1,3	1,0	1,3

Ciências Biológicas

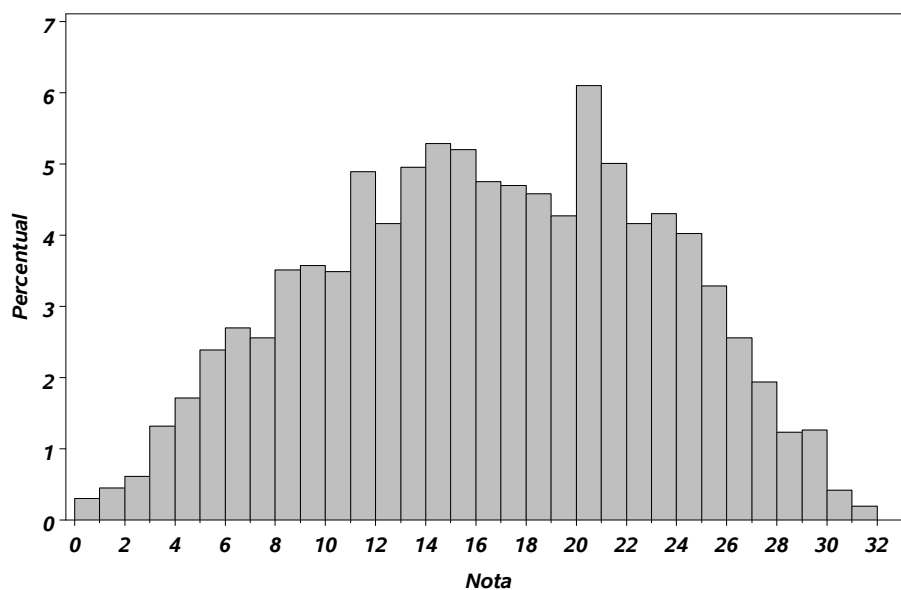
Caracterização estatística

Distribuição de frequências das notas

Presentes na segunda fase



Matriculados

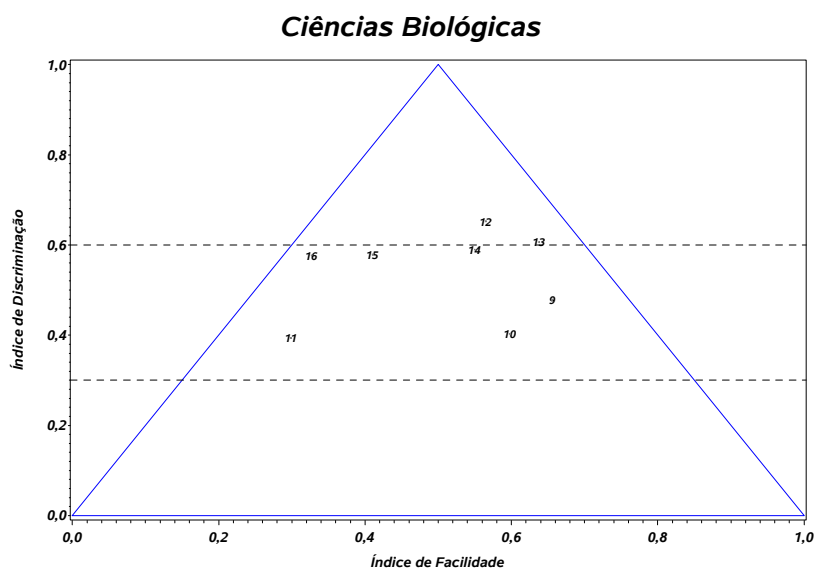


Média e desvio padrão

	candidatos	média	desvio padrão
presentes na 2ª fase	14592	16,2	6,8
matriculados	3554	16,2	6,7

Índices de facilidade e discriminação

questão	IF	ID
9	0,656	0,479
10	0,598	0,403
11	0,299	0,394
12	0,565	0,653
13	0,638	0,607
14	0,550	0,590
15	0,410	0,579
16	0,327	0,577



Média e desvio padrão por questão

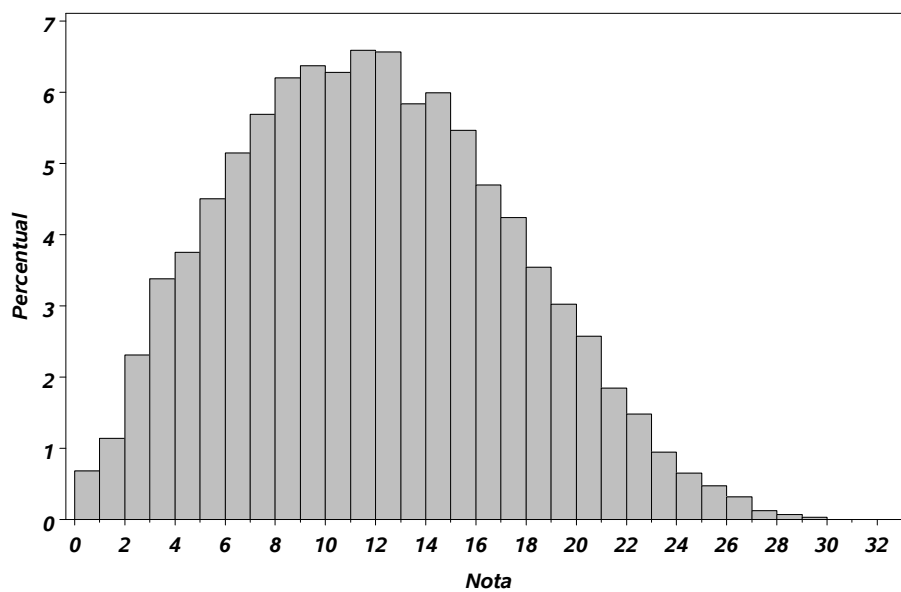
questão	presentes na 2ª fase		matriculados	
	média	desvio padrão	média	desvio padrão
1	2,6	1,1	2,7	1,1
2	2,4	1,1	2,4	1,1
3	1,2	1,1	1,2	1,1
4	2,3	1,3	2,2	1,3
5	2,6	1,4	2,6	1,4
6	2,2	1,3	2,2	1,3
7	1,6	1,2	1,7	1,2
8	1,3	1,3	1,3	1,3

Química

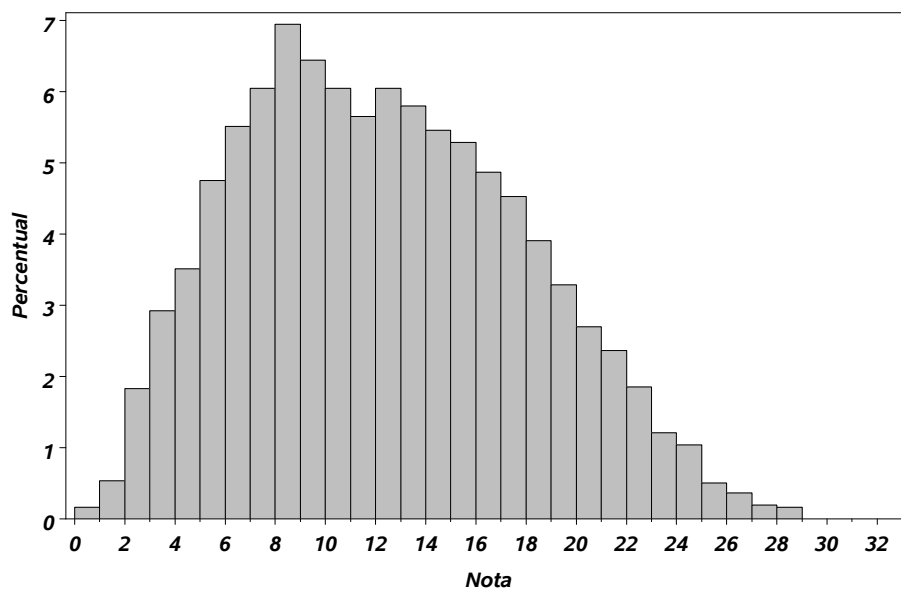
Caracterização estatística

Distribuição de frequências das notas

Presentes na segunda fase



Matriculados

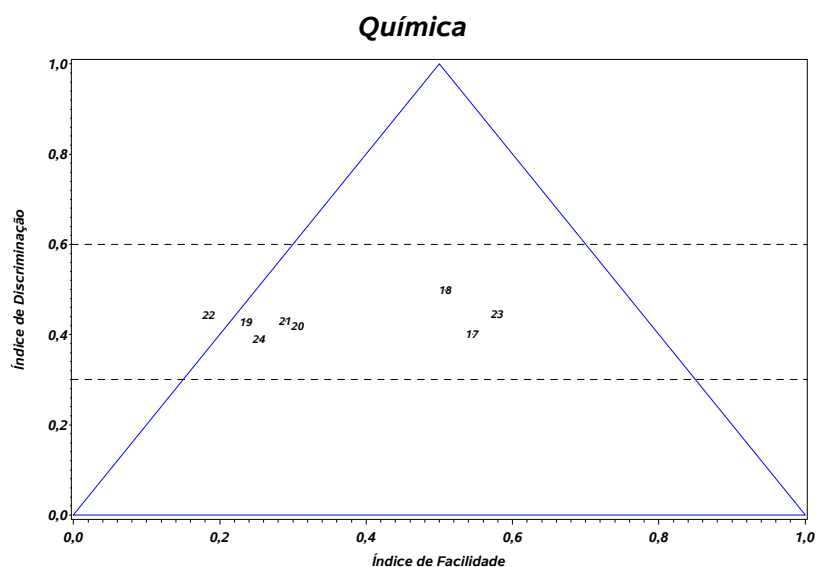


Média e desvio padrão

	candidatos	média	desvio padrão
presentes na 2ª fase	14592	11,6	5,6
matriculados	3554	12,0	5,6

Índices de facilidade e discriminação

questão	IF	ID
17	0,545	0,401
18	0,509	0,499
19	0,237	0,430
20	0,307	0,419
21	0,290	0,430
22	0,186	0,444
23	0,580	0,447
24	0,254	0,392



Média e desvio padrão por questão

questão	presentes na 2ª fase		matriculados	
	média	desvio padrão	média	desvio padrão
1	2,2	1,3	2,3	1,2
2	2,0	1,2	2,1	1,3
3	0,9	1,2	1,0	1,2
4	1,2	1,3	1,4	1,3
5	1,2	1,1	1,2	1,1
6	0,7	1,1	0,7	1,1
7	2,3	1,3	2,4	1,2
8	1,0	1,0	1,0	1,0

Habilidades Específicas

Provas de Habilidades Específicas

Caracterização estatística

curso	situação	candidatos	média	desvio padrão
Arquitetura e Urbanismo	Presentes	264	18,9	10,7
	Matriculados	31	29,8	5,7
Artes cênicas	Presentes	89	13,4	14,0
	Matriculados	25	28,5	6,0
Artes visuais	Presentes	101	18,3	15,3
	Matriculados	30	31,1	9,6
Dança	Presentes	92	15,2	15,0
	Matriculados	25	30,5	10,3
Música	Presentes	217	13,2	16,2
	Matriculados	77	27,3	11,8