

Diário matemático de uma epidemia em Portugal

JOSÉ PAULO VIANA

Com a chegada da pandemia, muitos de nós pensámos logo que esta seria uma oportunidade de aplicar a Matemática a uma situação real, observando a evolução de covid-19 no nosso país.

A situação mais simples de analisar é a de uma doença epidémica que não provoca mortes, que se torna crónica nos indivíduos infetados que mantêm a sua contagiosidade ao longo do tempo, e que não provoca alterações de comportamentos numa população isolada (por exemplo, uma doença destas numa comunidade de coelhos numa ilha). Neste caso, a evolução do número de elementos infetados ao longo do tempo é, aproximadamente, uma função logística.

A função logística é do tipo $f(x) = \frac{M}{1 + e^{-k(x-x_i)}}$, em que k é um parâmetro que define a rapidez com que a doença se espalha, x_i é a abcissa do ponto de inflexão e M é o total de indivíduos que serão afetados (o efetivo da população menos, eventualmente, o número de indivíduos naturalmente imunes à doença).

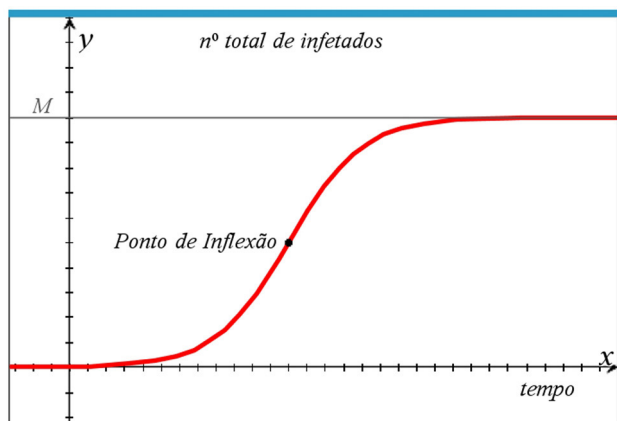


Figura 1. Função logística

Podemos distinguir três fases. Na primeira, a logística confunde-se com uma função exponencial e o crescimento é cada vez mais rápido. A segunda é uma fase intermédia que inclui o ponto de inflexão. Na terceira, o crescimento é cada vez mais lento, a epidemia vai apaziguando e a função tende para M .

Observemos a sua derivada, que nos dá o número diário de novos infetados.

O gráfico é uma curva em forma de sino, mas que não se pode confundir com a curva da distribuição normal. O máximo da derivada, que corresponde ao ponto de inflexão da função

logística, indica-nos o dia em que a epidemia atingiu o seu pico. A partir daí, o número de novos infetados vai diminuindo progressivamente e tender para zero.

A evolução de uma epidemia entre humanos é bastante mais complicada mas as considerações anteriores podem servir de ponto de partida para o nosso caso.

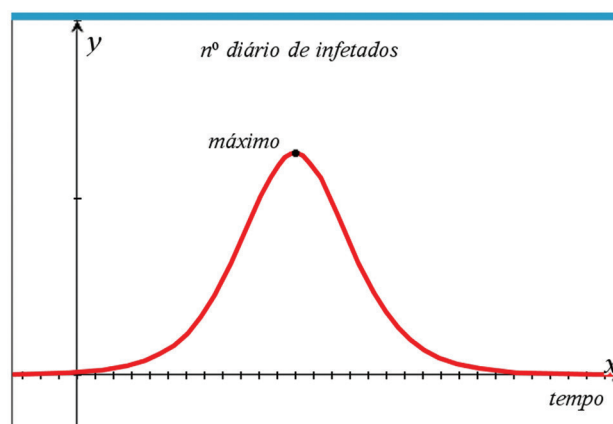


Figura 2. Derivada da função logística

Para fazer o estudo da evolução de covid-19 foram recolhendo-se os dados referentes ao total de infetados desde o início, divulgados diariamente pelas autoridades (ver anexo). Por coincidência, os primeiros casos conhecidos da doença em Portugal foram detetados no dia 1 de março, pelo que a variável "dia" começou a contar a partir dessa data.

OS PRIMEIROS DIAS

Foi preciso esperar algum tempo até se ter um número suficiente de dados. Como se pode ver na figura 3, o nº total de infetados em função do tempo (em dias) é uma exponencial quase perfeita. A nuvem de pontos coincide quase exatamente com o gráfico da função e isso é confirmado pelo coeficiente de determinação r^2 . Este coeficiente pode variar entre 0 (quando não existe qualquer relação entre duas variáveis) e 1 (quando os pontos da nuvem pertencem ao gráfico da função) e, neste caso, ele é altíssimo: 0,9876.

A base da potência da função exponencial (1,404) corresponde ao grau de crescimento da função. Nestes 15 dias, o número de infetados pela doença cresceu 40,4% em cada dia.

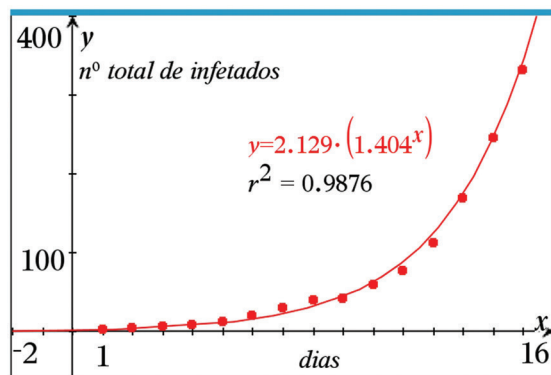


Figura 3. Os primeiros 15 dias

Com os dados atualizados diariamente, fomos fazendo novos modelos, mas os valores obtidos pouco se alteravam. A taxa de crescimento rondava sempre os 40% e r^2 passou a ser superior a 0,99. Os modelos descreviam com precisão a evolução da doença e era bonito ver a matemática a funcionar tão bem. Contudo, isto não permitia prever o mais importante: quando se dariam as alterações para entrar na fase de abrandamento nem, muito menos, quando estaria controlada a epidemia.

A SEGUNDA FASE

A partir do dia 20 os novos pontos começaram a afastar-se do modelo exponencial. Na figura 4 temos a situação dia 22.

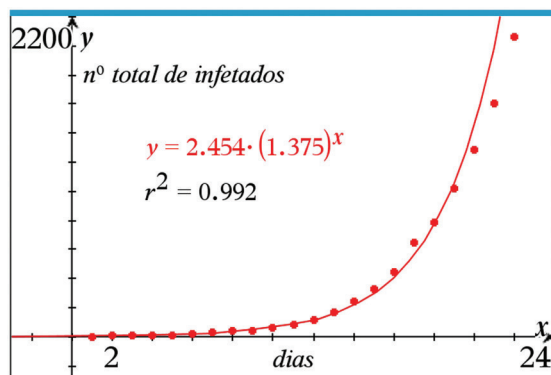


Figura 4. Exponencial no dia 22

Os últimos pontos estão já bastante afastados da exponencial e, apesar de r^2 continuar muito elevado, nota-se claramente que este modelo deixou de servir.

Tínhamos entrado na segunda fase e o modelo logístico tornou-se muito melhor. Como se vê na figura 5, o gráfico da curva passa muito perto dos pontos, ou seja, os desvios (ou resíduos) são muito pequenos.

A vantagem, agora, é que já podíamos começar a fazer previsões. Procuramos o ponto de inflexão do gráfico, que corresponde ao dia 24, o que significaria que o pico da epidemia seria daí a dois dias. Pela expressão da função, o número de infetados no fim da epidemia seria de 5897 e, consultando a tabela de valores da logística, via-se que isso aconteceria a 25 de abril. Parecia demasiado bom.

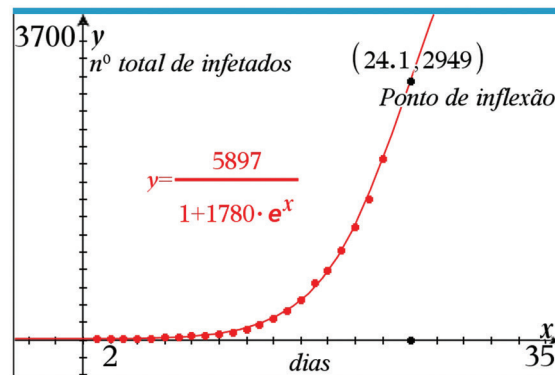


Figura 5. Modelo logístico no dia 22

Nos dias seguintes, continuámos a determinar o modelo logístico com os dados atualizados:

- dia 24 - fim no dia 28 de abril, 6814 infetados
- dia 25 - fim no dia 29 de abril, 7770 infetados
- dia 26 - fim no dia 30 de abril, 9581 infetados
- dia 27 - fim no dia 01 de maio, 12918 infetados
- dia 28 - fim no dia 09 de maio, 13397 infetados

A cada dia que passava, o final da epidemia ficava mais longe e o número total de infetados aumentava. Ou seja, estes modelos descreviam muito bem o passado mas, claramente, não eram nada bons a prever o futuro. Era preciso investigar mais sobre o assunto. A ajuda de João Sardo, que me indicou várias locais da internet, foi essencial.

A evolução de uma epidemia entre a população humana depende de uma grande quantidade de parâmetros. Alguns são fixos (grau de contagiosidade da doença, tempo de incubação, intervalo de tempo que um doente pode contagiar outras pessoas, distribuição populacional por meios urbanos e rurais), outros são variáveis com o tempo (percentagem de pessoas imunes, existência de doentes assintomáticos) e outros vão mudando devido a decisões individuais ou coletivas (isolamento social, distanciamento social, quarentenas preventivas, isolamento dos doentes). Como os comportamentos globais se alteram devido às decisões que se tomam, a doença não progride naturalmente e a função logística deixa de servir.

Outros fatores vêm complicar um estudo que se pretenda rigoroso. Há vários parâmetros associados à doença que não podem ser medidos com um mínimo de rigor e alguns deles só no final da epidemia serão conhecidos. Por outro lado, os dados oficiais divulgados estão sujeitos a erros (quando se fazem poucos testes, detetam-se menos infetados; pessoas doentes com sintomas ligeiros podem não contactar as autoridades de saúde; devido ao excesso de trabalho, os técnicos podem esquecer-se ou adiar o registo de doentes, etc.).

Por estes motivos, todas as tentativas de, nesta fase, encontrar um modelo seguro e fiável estão votadas ao fracasso. As previsões sobre a evolução da doença ficam sujeitas a um enorme grau de imprecisão. Dois vídeos com excelentes explicações sobre este assunto são os de Henry Reich [2] e 3Blue1Brown [3].

Não era só eu que sabia pouco sobre o futuro. Estava explicado o facto de as autoridades irem anunciando datas muito diversas para o aparecimento do pico da infeção. Por outro lado, percebi que os artigos catastrofistas que iam aparecendo em certos jornais e no Facebook não tinham qualquer credibilidade (como se foi confirmando depois).

Voltemos ao nosso estudo. Durante mais uns tempos foram existindo funções logísticas sempre bem adaptadas aos dados do momento e com previsões cada vez mais longínquas para o final da epidemia. Mas, a certa altura, a logística começou a falhar. Os desvios começaram a acentuar-se.

Ora, um artigo no Blog Que Treta! [4] e o vídeo de Henry Reich [3] explicavam que o gráfico da evolução de uma epidemia começa por ser uma logística mas depois passa a ser uma reta. Era o que estava a acontecer aqui. Veja-se, na figura 6, o gráfico referente ao dia 39 (8 de abril), em que já aparecem vários pontos afastados do gráfico da função.

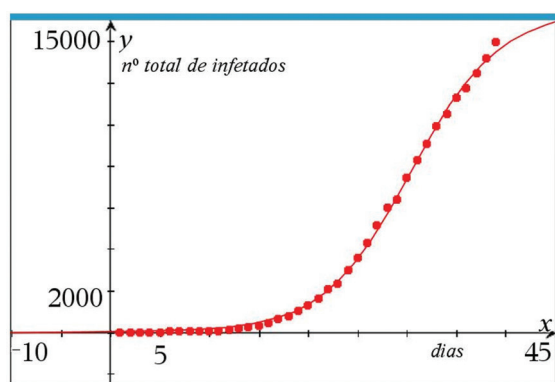


Figura 6. Modelo logístico no dia 39

Podemos confirmar, fazendo o gráfico entre os dias 24 e 39, que estamos já num crescimento retilíneo (figura 7).

A expressão analítica da reta indica-nos que, em média, o crescimento de novos infectados por dia foi de 741.

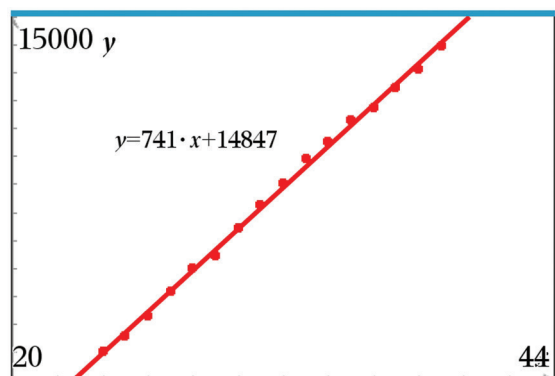


Figura 7. Dias 24 a 39

Henry Reich sugere uma outra forma gráfica de ver a evolução da epidemia, representado, em escalas logarítmicas, a relação diária entre o total de infectados e o número de novos infectados. Nota: como a variabilidade destes números é muito grande, opta-se por

usar o número de infectados nos últimos quatro a sete dias (ou a sua média). O comportamento é muito parecido de país para país. Inicialmente e durante algum tempo, os pontos dispõem-se segundo uma reta crescente, depois começam a aglomerar-se numa zona mais alta e, finalmente, descem rapidamente, marcando o final da epidemia.

Como os meios técnicos de que disponho não permitem fazer gráficos com escalas logarítmicas, foi preciso usar um artifício, colocando, no eixo horizontal, o logaritmo do total de infectados e, no eixo vertical, o logaritmo do número de novos infectados nos últimos quatro dias.

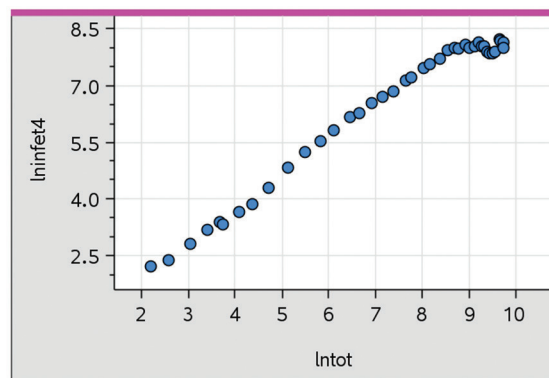


Figura 8. Gráfico logarítmico (até dia 44)
Total de infectados / Infectados em 4 dias

Como se vê na figura 8, depois do período inicial com pontos alinhados segundo uma reta, entrou-se já na segunda fase (aglomeração de pontos na parte superior).

Regressemos às representações habituais. A partir do dia 40 (9 de abril), a situação voltou a alterar-se. Os novos dados indicavam um menor crescimento do número de infectados, com os pontos alinhados numa reta de declive inferior. O gráfico da figura 9 mostra precisamente isso, quando se estava no dia 56.

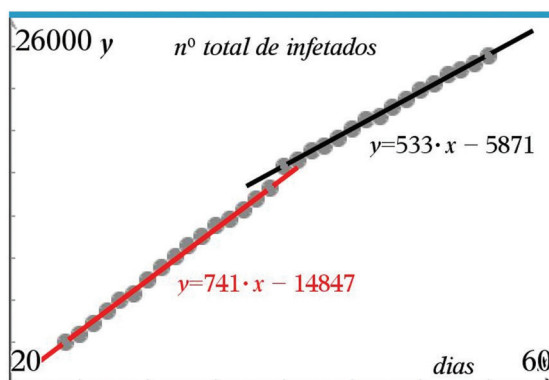


Figura 9. Dias 24 a 56

Nos últimos 16 dias, o crescimento médio de novos infectados diário tinha descido para 533, uma melhoria sensível quando comparada com o crescimento de 741 no período anterior.

Vejam, na figura 10, o gráfico logarítmico nesta altura.

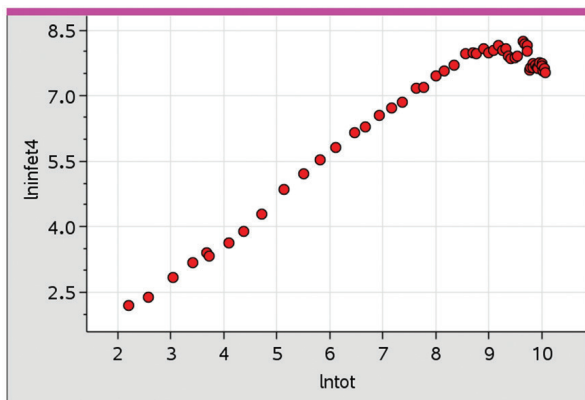


Figura 10. Gráfico logarítmico (até dia 56)
Total de infectados / Infectados em 4 dias

Havia nessa altura um segundo aglomerado de pontos, já abaixo do primeiro. Era mais um bom sinal.

Como, aparentemente, se tinha entrado na fase de apaziguamento do covid-19, podemos analisar o passado e tentar situar o pico da epidemia, isto é, o momento em que o crescimento do número de novos infectados atingiu o seu máximo. Isso terá acontecido entre os dias 8 e 10 de abril, embora não se possa ter a certeza.

A FASE DE APAZIGUAMENTO

Os dados dos dias que se seguiram foram bastante animadores, tanto mais que no dia 62 (2 de maio) foi anunciado que os dados referentes à semana anterior estavam inflacionados devido à duplicação de algumas participações de doentes infectados. Como não foi possível descobrir os verdadeiros valores para esses sete dias, a solução foi distribuir o erro proporcionalmente. Mesmo que isso não corresponda exatamente aos valores reais, as alterações aos modelos e às conclusões serão mínimas.

Nos dias 66 a 68 os números voltaram a subir, quebrando a tendência de abrandamento. Será isto um fenómeno isolado ou, com o cansaço, as pessoas terão começado a ter menos cuidados? Vai ser preciso esperar para se saber.

Como este artigo tem de seguir para a paginação, vejamos os gráficos referentes à situação até agora, dia 68.

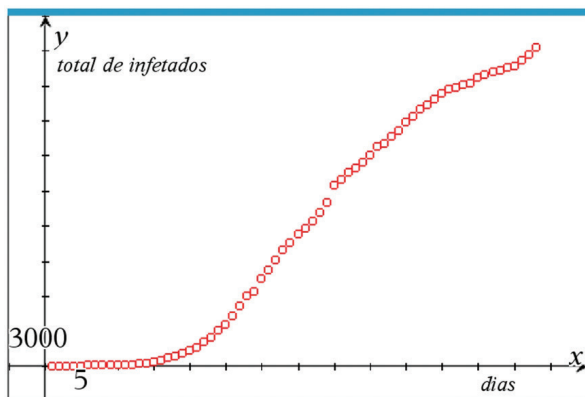


Figura 11. Evolução do número total de infectados até 7 de maio

Como se pode verificar, os três últimos pontos estão acima do alinhamento dos anteriores.

Com mais pormenor, temos o gráfico dos períodos de crescimentos lineares após o período logístico.

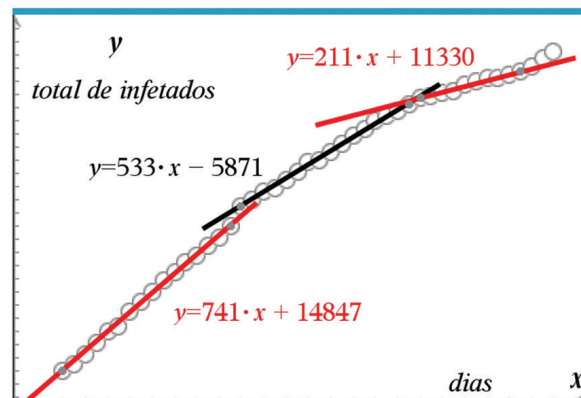


Figura 12. Dias 24 a 68

Entre os dias 56 e 65, o crescimento médio foi de 211 novos infectados diários mas, como é visível, os pontos dos três dias seguintes estão desalinhados e acima da reta. O crescimento médio nos dias 66 a 68 foi de 522, bastante maior que antes.

Finalmente, temos o gráfico em escalas logarítmicas.

Os novos pontos estavam a descer de forma sensível, aproximando-se do eixo horizontal e mostrando o abrandamento da doença. Pena é a existência das três exceções finais. Veja-se o ponto 68 (a cheio) bem acima dos anteriores. Tudo voltou atrás. Resta-nos agora esperar pelos próximos tempos, em que possamos voltar a estar ao vivo com os nossos alunos e a analisar com eles toda a matemática associada a este e a outros fenómenos reais.

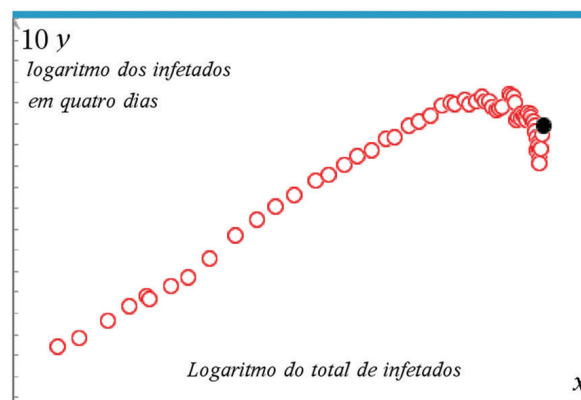


Figura 13. Gráfico logarítmico (até dia 68)
Total de infectados / Infectados em 4 dias

Referências

- [1] IHME COVID-19 Projections, <https://covid19.healthdata.org/portugal>

Mostra graficamente a evolução diária do nº de mortes, do total de mortes e dos meios hospitalares disponíveis, indicando, em

cada caso, as projeções para o futuro (sempre com uma amplitude muito grande, dado o grau de incerteza).

- [2] Henry Reich Minute Physics <https://youtu.be/54XLXg4fYsc>
Explica como analisar a evolução da epidemia usando gráficos com escala logarítmica e chama a atenção para as imprecisões dos dados e para as limitações e precauções a ter, quer com os resultados obtidos, quer com as projeções.
- [3] 3Blue1Brown <https://www.youtube.com/watch?v=gxAaO2rsdIs>
Explica o que é o modelo SIR de uma epidemia, com a evolução das três classes em que a população se pode dividir: Suscetíveis (os que podem apanhar a doença), Infeciosos (os que estão doentes) e Removidos (os recuperados e os falecidos). Mostra muitas simulações da forma como uma epidemia alastra, de acordo com a influência de muitos parâmetros: contagiosidade, grau de contactos entre pessoas, isolamento social (e quando algumas pessoas começam a quebrá-lo), distanciamento social, isolamento de doentes, existência de casos assintomáticos, meio urbano versus meio rural, existência locais de aglomeração de pessoas, viagens entre comunidades separadas, etc.
- [4] Blog Que Treta! <https://ktreta.blogspot.com/2020/04/inflexao.html>
Explica como o gráfico da evolução de uma epidemia começa por ser uma logística e depois passa a ser uma reta.

JOSÉ PAULO VIANA

ANEXO

Número total de infetados, contados a partir de 1 de março de 2020

Dia	1	2	3	4	5	6	7	8	9	10
Nº	2	4	6	9	13	21	30	39	41	59
Dia	11	12	13	14	15	16	17	18	19	20
Nº	78	112	169	245	331	448	642	785	1020	1280
Dia	21	22	23	24	25	26	27	28	29	30
Nº	1600	2060	2362	2995	3544	4268	5170	5962	6408	7443
Dia	31	32	33	34	35	36	37	38	39	40
Nº	8251	9034	9886	10524	11278	11730	12442	13141	13956	15472
Dia	41	42	43	44	45	46	47	48	49	50
Nº	15987	16585	16934	17448	18091	18841	19022	19685	20206	20863
Dia	51	52	53	54	55	56	57	58	59	60
Nº	21379	21982	22353	22797	23307	23712	23852	24105	24262	24725
Dia	61	62	63	64	65	66	67	68	69	70
Nº	24987	25190	25282	25524	25702	26182	26715	27268		

Nota: os valores dos dias 55 a 61 são estimados.

ESTATUTO EDITORIAL DA EDUCAÇÃO E MATEMÁTICA

A Educação e Matemática é uma publicação periódica da Associação de Professores de Matemática (APM). A sua periodicidade atual é de cinco números anuais, sendo um deles temático e duplo. A revista aborda questões relacionadas com o ensino e aprendizagem da matemática. Dirige-se aos professores de matemática, de todos os níveis de ensino, em especial aos sócios da APM, constituindo um meio de comunicação privilegiado da Associação, em Portugal e no estrangeiro. Os principais objetivos da Educação e Matemática são:

- Promover a troca de ideias e experiências entre professores;
- Estimular a reflexão sobre problemas e desafios da educação matemática;
- Discutir temas atuais e importantes da educação matemática e da educação em geral;
- Fornecer elementos de trabalho para as práticas dos professores;
- Divulgar informação relevante para os professores.

A Educação e Matemática publica textos de natureza diversa. Vive muito da contribuição dos sócios, que são autores da maior parte dos artigos. Estas contribuições passam por ideias, pontos de vista, comentários, relatos de experiências, artigos de opinião, resenhas de livros, resolução de problemas, notícias... A Educação e Matemática tem um conjunto de secções de natureza diversificada, algumas das quais com carácter permanente. A revista tem uma equipa redatorial a quem compete desenvolver todo o trabalho de receção e revisão de artigos, bem como organizar a própria revista. À semelhança das outras revistas informativas, a Educação e Matemática assegura o respeito pelos princípios deontológicos e pela ética profissional dos jornalistas, assim como pela boa fé dos leitores.

A DIRETORA DA EDUCAÇÃO E MATEMÁTICA