



AGRUPAMENTO DE ESCOLAS DE  
CISTER - ALCOBÇA

Observatório da  
Comunidade Escolar

# Análise de Variância

*Linhas  
Interpretativas*

2015

Departamento de Avaliação e Planeamento

---

*“Ausência de evidência não é evidência de ausência.”*

Carl Sagan

*“Statistical models are sometimes misunderstood [...]. Statistical models for data are never true. The question whether a model is true is irrelevant. A more appropriate question is whether we obtain the correct scientific conclusion if we pretend that the process under study behaves according to a particular statistical model.”*

Scott Zeger

## Conteúdos

Conteúdos.....	2
I. Introdução.....	3
II. Intenções e objetivos do modelo .....	4
III. Análise de Variância Revisitada.....	5
IV. Linhas interpretativas .....	8
V. Conclusão .....	12

## I. Introdução

Este documento pretende apresentar as intenções do modelo de Análise de Variância, assim como sugerir linhas interpretativas **possíveis** dos dados por ele gerados. Os resultados são apresentados periodicamente nos relatórios do Observatório da Comunidade Escolar.

Para uma leitura mais aprofundada sobre o modelo de Análise de Variância, recomenda-se o documento “Estatísticas e Modelos”, publicado em 2014.

No presente, optou-se por criar um instrumento de trabalho menos focado na definição técnico-científica do modelo, mais centrado nos seus fundamentos conceituais e **mais preocupado com o que fazer com os resultados do que com os resultados em si**. Para tal, tentaremos mostrar algumas linhas possíveis de interpretação e ação subsequente. A apresentação dessas linhas, bem como uma possível estratégia de ação em função de casos específicos, será feita sobre um conjunto de exemplos, para que a exposição se torne mais concreta.

Os destinatários deste documento são, em primeiro lugar, todos os docentes, para que possam melhorar a percepção que têm sobre o modelo de Análise de Variância dos professores. Aliás, os relatórios do Observatório tentam satisfazer as necessidades de dados estatísticos de diferentes pontos de vista: do professor, do diretor de turma, de delegados e coordenadores a vários níveis, até às estruturas mais elevadas, como o Conselho Pedagógico.

No entanto, a intenção da qual nasce, de facto, este documento, é a de proporcionar às estruturas de coordenação de ano, envolvendo os delegados de grupo e coordenadores de departamento, maior compreensão do modelo. A informação que este produz é estrategicamente valiosa e tem sido, de certa forma, desaproveitada. Assim, propõe-se também uma estratégia concreta de ação que pode e deve ser dinamizada no terreno por estes agentes.

Evidentemente, supõe-se que tanto o Observatório da Comunidade Escolar como as estruturas diretivas e pedagógicas estarão disponíveis para colaborar numa implementação em larga escala de um modelo simples e eficaz, que nos permita, de modo continuado, melhorar a qualidade e a justiça do processo chave da organização, no seu sentido mais lato: o processo de educar.

## II. Intenções e objetivos do modelo

De uma forma muito concisa, o principal objetivo do modelo de Análise de Variância é **estimar a influência do professor nas classificações**.

Em primeiro lugar, deve desmistificar-se este conceito, por vezes mal recebido por parte dos professores, por dele serem feitas interpretações erradas sobre as intenções. Assim, vale a pena parar para explicar que:

- A intenção do modelo não é, de forma alguma, servir como base a qualquer tipo de avaliação do desempenho docente.
- Nem os professores que aparecem com estimativas de influência desviantes são maus profissionais nem os que aparecem com estimativas de influência alinhadas são bons profissionais. De facto, isso nunca foi um pressuposto de qualquer modelo produzido pelo Observatório. **Não há nenhuma relação entre a estimativa da influência do professor e a qualidade da sua prestação docente.**
- Os professores com estimativas desviantes não têm, necessariamente, que convergir com o resto do grupo. Cada situação deve ser cuidadosamente analisada em relação às suas causas, e a ação a desenvolver deverá ser ponderada e ajustada em função da realidade específica que constitui cada caso.
- Os professores com estimativas alinhadas não estão, necessariamente, em convergência de estratégias e métodos com o resto do grupo. Podem ser desviantes nas práticas apesar de convergentes nos resultados.

Quais são, então, as intenções por detrás deste modelo? O trabalho enquadra-se no seguinte objetivo geral do Observatório:

*“A atuação do Observatório visa, de uma forma construtiva e respeitadora da autonomia dos agentes educativos, criar uma cultura de autoavaliação e de permanente otimização da qualidade da educação.”*

Dos objetivos específicos do tratamento estatístico periódico dos resultados escolares, destacam-se:

1. Como parte do processo contínuo de autoavaliação da escola, **monitorizar de forma regular os resultados do processo-chave** da organização, que é o processo de ensino-aprendizagem.
2. Fornecer aos grupos disciplinares uma ferramenta de trabalho, para que possam:
  - a. de forma contínua, **melhorar o desempenho dos seus elementos** através de confronto dos resultados individuais e aprendizagem recíproca;
  - b. trabalhar no sentido de **observar o princípio fundamental de que os alunos têm direito a condições de aprendizagem semelhantes independentemente da turma ou escola que integram.**

Em suma, a ideia nuclear é dar suporte e monitorização a ações (nomeadamente, por parte dos departamentos e grupos) cujo objetivo seja melhorar, de forma continuada, a qualidade e a equidade do processo ensino-aprendizagem.

### III. Análise de Variância Revisitada

Nesta secção, relembrem-se, de forma mais concetual, os pressupostos e medidas do modelo de Análise de Variância.

#### Pressupostos do modelo

Cada classificação atribuída é função de uma série de variáveis (ou fatores). Entre elas, encontram-se, por exemplo, o aluno, a natureza da disciplina, o professor, o contexto socioeconómico da família, a distância de casa à escola *et cetera*.

Entre esses fatores, os três cujas influências são predominantes na “produção” de uma classificação são (e este facto foi confirmado durante o desenvolvimento do modelo): o **aluno**, a **disciplina** e o **professor**.

Os restantes fatores são determinados por estes e é, por isso, impossível, através de métodos estatísticos, quantificar a sua influência. Dito de outro modo, as variáveis “*distância de casa à escola*” e “*aluno*” são dependentes, e esse facto faz com que seja impossível separar as suas influências.

Em relação às três variáveis consideradas, provou-se, durante o desenvolvimento do modelo, que:

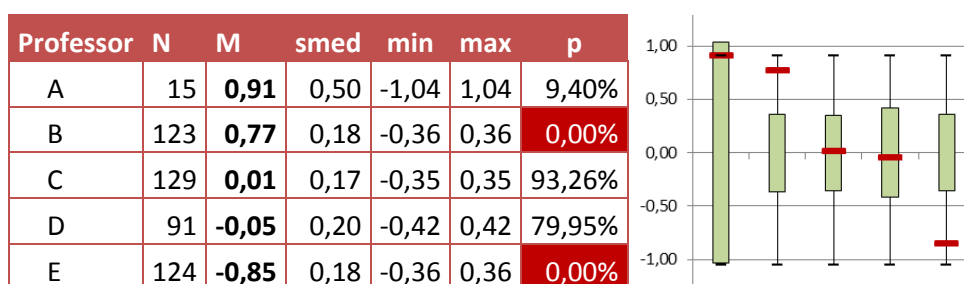
- As variáveis **aluno**, **disciplina** e **professor** são independentes entre elas;
- As variáveis **aluno**, **disciplina** e **professor** têm todas uma influência significativa nas classificações.

Vale a pena reler e refletir sobre as duas últimas afirmações. Elas contêm uma parte importante da essência do modelo. Em primeiro lugar, refletem um trabalho de validação que foi feito durante o desenvolvimento do modelo que não é observável nos resultados publicados.

Em segundo, estabelecem uma fundação importante: de que estamos a olhar para as variáveis certas. Adicionalmente, sabemos que, para além disso, podemos quantificar os seus efeitos e, finalmente, aferir uma medida que o modelo do valor esperado, criado com a mesma intenção, nunca conseguiu produzir: um valor estimado da influência do nosso trabalho no processo-chave da organização, despido das contribuições do aluno e da disciplina, que sempre dificultaram a sua observação.

#### Medidas

Os dados normalmente apresentados são, para cada grupo disciplinar, uma tabela e um gráfico:



A medida mais importante aferida por este modelo é, claramente, a estimativa da influência do professor. Esse valor é dado na coluna **M** da tabela e pela barra horizontal vermelha, no gráfico.

De onde provém este valor? A cada classificação atribuída por um professor, é subtraída a influência do aluno e da disciplina. Ao valor que resta chamamos resíduo. A média de todos os resíduos de um professor é o valor **M**.

Concretizando, um valor de 0.91 para o professor A significa que este professor, para a mesma disciplina e para alunos com perfil semelhante, está a atribuir classificações, *grosso modo*, 0.91 valores mais altas que os seus homólogos. É muito importante que se lembre a premissa de que os efeitos do aluno e da disciplina já foram retirados. Só assim se pode atribuir o resíduo ao professor.

No entanto, como qualquer estimativa, esta é sensível ao erro amostral (desvio padrão da distribuição dos resíduos). Isso significa que a influência do professor será estimada com mais qualidade quanto maior for o tamanho da amostra (número de classificações atribuídas pelo professor).

Sendo assim, este valor por si só é insuficiente para uma análise correta. Igualmente importante é saber a probabilidade de, tendo em conta o tamanho da amostra, a média dos resíduos se encontrar tão ou mais afastada de zero devido a fatores meramente aleatórios. Esse valor de prova é dado na coluna *p* da tabela. Em relação à representação gráfica, é preferível introduzir o conceito de intervalo de confiança.

Assim, uma análise completa é como se segue:

- **Hipótese nula:** o professor X não influencia as classificações.
- Se o professor não influencia as classificações, a sua influência (*M*) é estimada perto de zero.
- É calculado o intervalo de confiança (centrado em zero), em função da variância dos dados e do tamanho da amostra do professor.
- Um intervalo de confiança de 95% (p.e.) significa que:
  - Se a influência do professor não é significativa, 95% das vezes a estimativa será obtida dentro do intervalo de confiança.
  - Se a estimativa é obtida fora desse intervalo, a hipótese nula é, em princípio, falsa (a probabilidade de a estimativa ser obtida fora dele, sendo a hipótese verdadeira, é inferior a 5%). *Assumindo* que é falsa, o professor X influencia as classificações.
- Estatisticamente, considera-se a influência de uma variável:
  - Significativa, quando a estimativa é obtida fora do intervalo de confiança a 95%;
  - Altamente significativa, quando a estimativa é obtida fora do intervalo de confiança a 99%;
- Adicionalmente, num teste desta natureza, pode sempre haver:
  - Falsos positivos: estimativa fora do intervalo de confiança quando a hipótese nula é verdadeira;
  - Falsos negativos: estimativa dentro do intervalo de confiança quando a hipótese nula é falsa.

Uma leitura mais direta pode ser feita através do gráfico:

- Para cada professor, o retângulo verde representa o intervalo de confiança e a barra vermelha a estimativa da sua influência.
- Quando a barra vermelha está fora do intervalo, deve tomar-se como falsa a hipótese nula, ou seja, deve partir-se do princípio que o professor influencia as classificações de forma discrepante em relação ao resto do grupo (o desvio é estatisticamente significativo).

O modelo tem um grau complexidade mais ou menos elevado e a sua compreensão profunda pode tornar-se pouco acessível dependendo da bagagem científica de cada um. Os dados que se apresentam são, no entanto, de fácil interpretação utilizando estas heurísticas para a análise.

### *Ciclo de melhoria*

Porque foi a aferição das influências centrada na variável professor? Apenas por ser a variável que podemos controlar. De que nos vale concluir que a disciplina de Matemática é diferente da disciplina de Educação Física? De que nos vale concluir que existem alunos com rendimentos distintos? Em que medida podemos agir em relação a essas diferenças (se é que faz sentido agir nestes casos)?

Já em relação à nossa própria influência, podemos e devemos agir. E devemos agir no sentido das principais intenções do trabalho do Observatório: por um lado, melhorar, de forma contínua, o nosso desempenho, aprendendo entre nós e aferindo o sucesso das nossas estratégias; por outro, trabalhar no sentido de observar o princípio fundamental de que os alunos têm direito a condições de aprendizagem semelhantes independentemente da turma ou escola que integram.

Idealmente, do processo de análise dos resultados pelos grupos e departamentos, deveriam emergir propostas de ações de melhoria/convergência **concretas**, que podem ser geridas, com ligeireza, através de ciclos PDCA (Plan, Do, Check, Act):

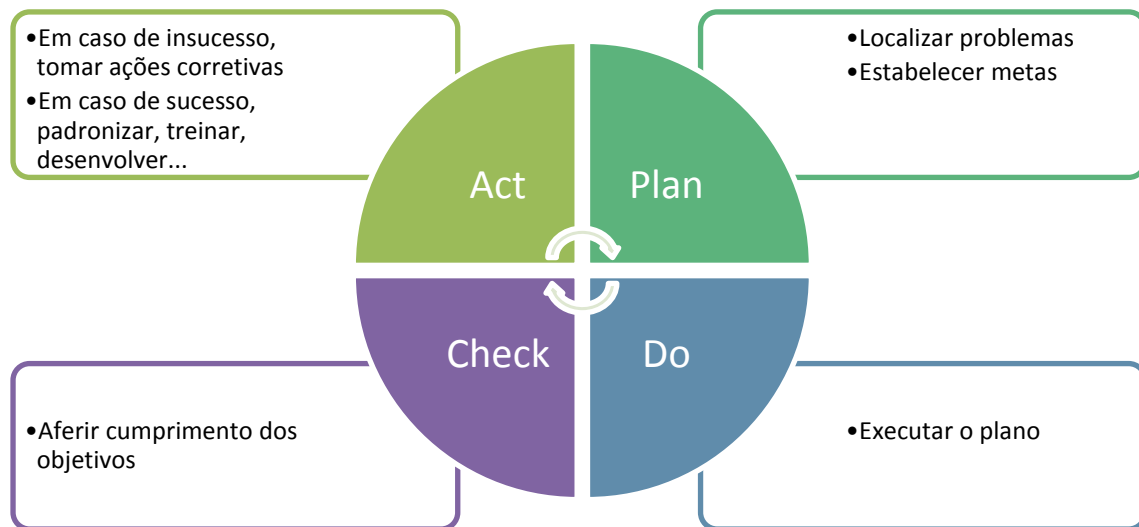
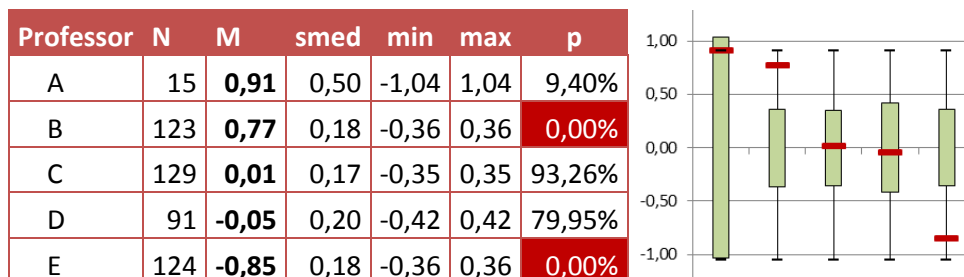


Figura 1 - Concetualização de um plano de acompanhamento de uma ação de melhoria



## IV. Linhas interpretativas

Relembremos os dados mostrados na secção anterior. Que possíveis interpretações é possível extrair destes resultados?



Como método de trabalho, propõe-se uma análise dos resultados de cada professor, seguida de uma análise geral aos resultados em grupo.

### Professor A

N	M	smed	min	max	p
15	<b>0,91</b>	0,50	-1,04	1,04	9,40%

O professor A apresenta o maior de todos os desvios observados. A estimativa da sua influência é de 0.91 valores. No entanto, este professor atribuiu apenas 15 classificações. A qualidade da estimativa é fraca e, portanto, o seu intervalo de confiança é grande. Vale a pena seguir este caso nas próximas publicações de resultados para clarificar a influência do professor.

Com base apenas nestes dados, não se pode concluir que a influência do professor seja significativa.

### Professor B

N	M	smed	min	max	p
123	<b>0,77</b>	0,18	-0,36	0,36	0,00%

O professor B tem uma influência marcada sobre as classificações dos seus alunos. Com uma amostra desta dimensão (123 alunos), a probabilidade de o professor não ter influência e se obter a estimativa da influência em 0,77 valores é aproximadamente igual a 0.00% (neste caso concreto é de 0,000025%). Clarifiquemos esta interpretação, que muitas vezes é verbalizada de forma incorreta, mas antes disso, relembremos a hipótese nula do modelo:

**Hipótese nula: o professor B não influencia as classificações.**

Apresentam-se agora várias leituras possíveis (corretas e incorretas) destes dados:

#### Leitura 1

*A probabilidade de a hipótese ser verdadeira é quase nula e, portanto, esta deve ser rejeitada. O professor B influencia as classificações e a sua influência é de 0.77.*

Esta leitura é tecnicamente incorreta. O modelo não consegue produzir a probabilidade de a hipótese nula ser verdadeira ou falsa. Essa asserção é interpretativa, em função dos dados e da realidade conhecida. Também é incorreto dizer que a influência do professor é exatamente 0.77. Essa estimativa apenas localiza um intervalo de confiança onde se encontra o valor real.

## Leitura 2

*Se a hipótese for verdadeira, a probabilidade de se obter  $M \geq 0.77$  é quase zero. A hipótese nula deve ser rejeitada. Trabalhar-se-á sobre o pressuposto que a hipótese é falsa, ou seja, este professor influencia as classificações de forma discrepante em relação ao grupo disciplinar.*

Esta leitura é correta. **O grupo deve, a partir deste ponto, trabalhar com o pressuposto que o professor influencia as classificações de forma discrepante em relação aos restantes elementos.** O valor da estimativa em 0.77 significa que o valor real da influência do professor se situa, com 95% de certeza, no intervalo [0.41; 1.13], que é nada mais que o mesmo intervalo de confiança, mas agora centrado em 0.77.

O verdadeiro trabalho de interpretação do grupo disciplinar começa aqui. Até este ponto, a interpretação é (estatisticamente) objetiva. Neste caso, é objetivo que há fatores relacionados com os professores do grupo que estão a escoriar a equidade das classificações. Isto não é a análise dos resultados, são os resultados em si.

A questão a fazer a seguir é: **quais as causas deste desvio?**

As causas podem ser as mais variadas. A sua identificação apenas pode ser feita pelo grupo disciplinar, conhecedor da realidade que originou os dados em análise. Deixam-se aqui *algumas* causas possíveis, em forma de questão:

- O professor utiliza instrumentos de avaliação divergentes?
- O professor aplica os critérios de classificação de provas de forma menos exigente que os seus pares?
- O professor aplica os critérios de classificação final da mesma forma que os seus pares?
- O professor tem estratégias para o processo ensino-aprendizagem que são, de facto, mais eficazes que as dos seus pares?
- ...

Há, no entanto, questões que se podem considerar incorretas. Por exemplo:

- O professor tem alunos “melhores” ou “piores” que os seus pares?
- Os resultados são assim porque o professor tem uma disciplina “difícil”?

**Ao identificar as causas, é sempre incorreto relacioná-las com o fator aluno ou com o fator disciplina.** Essas influências já foram retiradas. Melhor ainda, retirar essas influências é o princípio fundador de todo o modelo.

Muitas outras questões podem ser feitas para tentar encontrar as causas e em nenhuma circunstância pode ser encontrado um conselho melhor que o do grupo disciplinar, desdobrado nas suas coordenações de ano, para as responder. O grupo tem necessariamente que ser uma peça fundamental nesta engrenagem de melhoria. Sem uma análise séria e consequente, todos estes dados serão, simplesmente, números.

Encontradas as causas, pode rascunhar-se um plano de melhoria/convergência cujos resultados da implementação possam ser, preferencialmente, quantificados.

Por exemplo, vamos pressupor, para efeitos desta exposição, que o grupo identificou como causa da discrepância o facto de este professor elaborar provas com um grau de dificuldade menor que o do resto do grupo. Neste caso, o plano de melhoria/convergência pode passar por qualquer coisa semelhante à figura que se mostra a seguir:

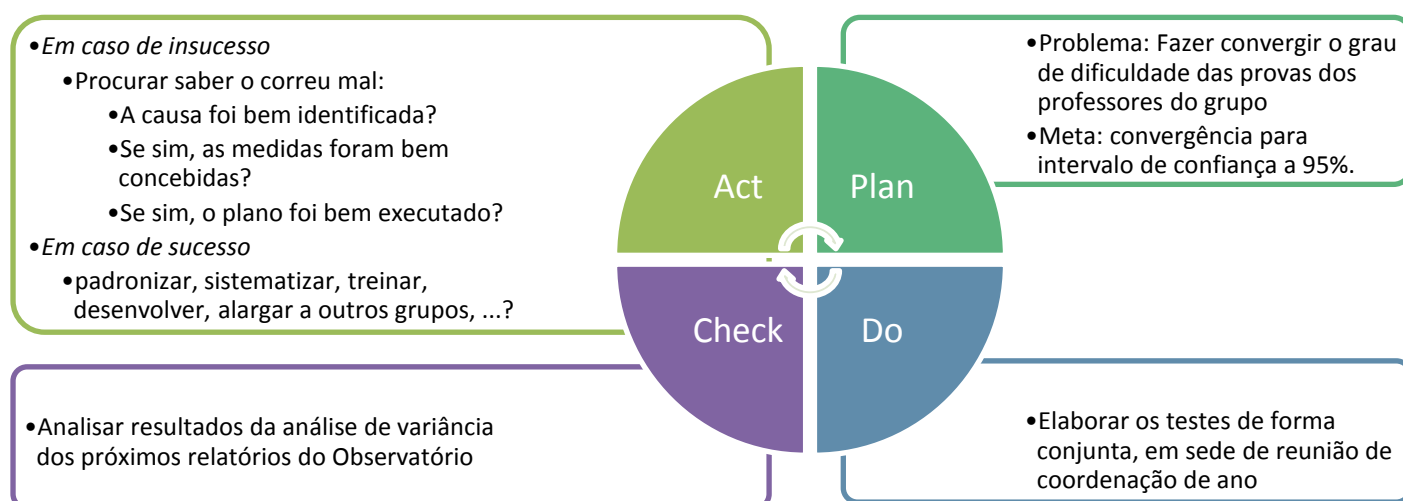


Figura 2 - Ciclo de acompanhamento de uma ação de melhoria

### Professores C e D

N	M	smed	min	max	p
129	<b>0,01</b>	0,17	-0,35	0,35	93,26%
91	<b>-0,05</b>	0,20	-0,42	0,42	79,95%

Estes professores têm influências médias dentro deste grupo. Não é verdade que não tenham influência (todos influenciámos), mas que a sua influência dentro deste grupo se situa em valores médios.

Isto não significa, que este seja o ponto a atingir por todos os professores do grupo. O que se pretende é que não haja desvios significativos. Deve haver convergência, mas o ponto para o qual o grupo deve convergir não é necessariamente o ponto médio. Da análise feita pelo grupo disciplinar, pode concluir-se que todos devem convergir com o professor B ou E, por exemplo.

Podemos também estar em presença de falsos negativos, isto é, professores que apesar de serem divergentes nas suas práticas, não sejam divergentes nas classificações atribuídas. A análise destas idiossincrasias só pode ser feita no contexto do próprio grupo.

### Professor E

N	M	smed	min	max	p
124	<b>-0,85</b>	0,18	-0,36	0,36	<b>0,00%</b>

O caso deste professor é semelhante ao do professor B, mas a sua influência sobre as classificações tem sinal negativo. Neste caso, o grupo deve também rejeitar a hipótese nula e trabalhar com o pressuposto que o professor influencia as classificações de forma discrepante em relação aos restantes elementos. A sua influência situa-se, com 95% de certeza, no intervalo  $[-0,49; -1,21]$ .

À semelhança do professor B, o grupo deve procurar encontrar as causas da divergência e elaborar um plano de melhoria.

### Análise geral do grupo

Numa análise geral ao grupo ressalta o facto de existirem dois professores com resultados divergentes do resto do grupo de forma estatisticamente muito significativa. De facto, estes dois casos estariam fora do intervalo de confiança, mesmo que este fosse calculado a 99.99%.

O primeiro ponto que vale a pena sublinhar pode ser exposto da seguinte forma: os professores divergentes não são, necessariamente, os professores que devem mudar o seu curso de ação. Como já foi dito, é importante que haja convergência, mas o ponto de convergência não é necessariamente o ponto médio.

Por exemplo, se o grupo chegar à conclusão que o professor B tem resultados discrepantes porque utiliza estratégias mais eficazes em aula (dito de um modo mais grosseiro, consegue ensinar *mais*), a estratégia a adotar pode ser a de dissimular essas estratégias pelo grupo (melhoria de desempenho), na esperança de convergir sobre os resultados do professor B. Evidentemente, os resultados de um plano dessa natureza devem ser aferidos a médio-longo prazo. Pode ainda não haver convergência assinalável no próximo relatório.

Se por outro lado, a causa e a estratégia a seguir forem as descritas anteriormente (na análise dos resultados do professor B), é expectável que haja convergência num prazo muito curto.

## V. Conclusão

De um modo geral, o grupo deve tentar definir como objetivo o cumprimento, por parte de todos, de parâmetros de convergência essenciais, ao nível instrumental, de avaliação, de classificação *et cetera*, deixando espaço para a liberdade individual de cada professor como agente educativo e como decisor *primo* do seu espaço pedagógico.

Essa liberdade não pode, de modo algum, ser absoluta, pois é contraditória com a observância de um princípio fundamental da filosofia do nosso Agrupamento de Escolas: os alunos têm direito a condições de aprendizagem semelhantes independentemente da turma ou escola que integram.

A expressão “condições de aprendizagem semelhantes” é suficientemente vaga para permitir liberdade individual e satisfatoriamente definida para que ela não seja absoluta. Voltámos à premissa inicial, em que todo este trabalho deve ser centrado: **melhorar, de forma continuada, a qualidade e a equidade do processo ensino-aprendizagem.**

---

Qualquer esclarecimento acerca dos modelos utilizados ou dos dados apresentados pode ser dado pelo Observatório através de [observatorio@aecister.pt](mailto:observatorio@aecister.pt).