

A Motivação para a Certificação de Qualidade: uma análise mediante árvores de decisão

Joaquín Texeira Quirós
jtexeira@unex.es
Universidad de Extremadura
Universidade Autónoma de Lisboa

Maria Rosário Fernandes Justino
mrjustino@iscal.ipl.pt
ISCAL- IPL

Resumo:

Este estudo analisa um questionário respondido por uma amostra de empresas certificadas pela ISO 9000 e uma amostra estratificada de empresas não certificadas, a fim de avaliar se existem diferenças significativas entre os dois grupos de empresas, utilizando modelos preditivos multivariados. Com esta abordagem, avaliamos a percepção dos entrevistados acerca dos elementos de qualidade no questionário que lhe estão associados e qual a probabilidade da empresa ser certificada. Utilizando uma abordagem multivariada levámos em conta a correlação entre as respostas e identificámos o melhor conjunto de respostas ao questionário para prever essa probabilidade. Este estudo mostra, que as empresas em que o entrevistado manifesta maior preocupação com: a relação com clientes; a motivação dos colaboradores e o planeamento estratégico têm maior probabilidade de ser certificadas.

Palavras-chave: Árvores de Decisão; ISO9000; Qualidade; Regressão Logística

Abstract:

This empirical study analyzes a questionnaire answered by a sample of ISO9000 certified companies and a control sample of companies which have not been certified, using multivariate predictive models. With this approach, we assess which quality practices are associated to the likelihood of the firm being certified. It is shown that companies in which the respondent manifests greater concern with respect to the increased competition due to globalization, the motivation of employees, and strategic planning have higher likelihood of being certified.

Keywords: Decision tree; Logistic regression, ISO9000; Quality

1. Introdução

A melhoria da qualidade é actualmente uma das ferramentas que as empresas possuem para promover, ou manter, as suas posições no mercado. Por isso, não é estranho, observarmos que existe um grande número de estudos relacionados com a motivação, os custos e os benefícios da certificação da qualidade, e os seus efeitos no desempenho das empresas. A filosofia denominada Qualidade Total tem vindo a evoluir, desde que surgiu no século XX, à medida que as empresas foram sentindo necessidade de aumentar o seu nível de competitividade no mercado e de garantir até a sua existência. Tudo isto foi provocado pelo aparecimento de mercados mais globalizados, com maiores competências e devido à procura por parte dos clientes de maiores níveis de qualidade nos produtos. Actualmente, num mercado totalmente globalizado, onde as empresas se relacionam com clientes e fornecedores de carácter internacional, já não basta colocar no mercado produtos ou serviços de elevada qualidade. Há que garantir que estes produtos e serviços foram produzidos segundo as normas e os critérios que dão segurança aos clientes de que aquilo que adquirem é fiável. Os clientes estão cada vez mais informados e mais exigentes, procuram comprovar se o produto de determinada marca é certificado, se a sua qualidade está assegurada.

Devido à sua generalizada aceitação, a ISO9000 tornou-se um factor importante no comércio internacional, quase um imperativo para as empresas que exportam para a União Europeia, no qual muitos compradores exigem esta certificação (Erel e Ghosh, 1997). Sendo que, na actualidade, a certificação ISO9000 é a principal unidade de medida e de prova da qualidade a nível internacional, e neste sentido, a implementação de um sistema de qualidade apresenta-se como um excelente suporte da mesma, e a certificação pela ISO9000 a sua bandeira (Escanciano, 2002). Deste modo, em 1987 a International Organization for Standardization (ISO), com sede em Genebra, publica um sistema harmonizado de normas que asseguram a qualidade, denominadas ISO9000, sendo que até ao momento um grande número de empresas em todo o mundo introduziu os aspectos básicos de gestão da qualidade. Neste sentido e de acordo com a ISO Survey de 2009, o número total de certificados emitidos, em todo o mundo em concordância com a ISO 9001, foi de 1.064.785 em 178 países, o que representa um aumento de 81 953 certificações em relação ao ano anterior, ou seja, mais 8%. A China é o país que apresenta um maior número de empresas certificadas. Curiosamente, as empresas americanas não parecem tão interessadas em obter a certificação como os seus concorrentes europeus. Assim o número de certificações nos EUA é menor do que nos países como a Itália, o Japão, a Espanha, a Federação Russa, a Alemanha, o Reino Unido e ainda a Índia. Apesar

disso, cabe-nos assinalar que os EUA juntamente com o Japão foram os pioneiros na aplicação da gestão da qualidade total e conseqüentemente, do seu historial em matéria de gestão da qualidade, centram-se mais neste sistema do que na normativa ISO criada na Europa (Sun, 1999).

Na literatura encontram-se diversos estudos que analisam as motivações para as empresas alcançarem a certificação através da norma ISO9000. De acordo com Tsiotras e Gotzamani (1996), os principais motivos são basicamente quatro: melhorar a imagem e reputação da empresa para o exterior, satisfazer a procura externa e as pressões do mercado, facilitar e simplificar os procedimentos e contratos entre a empresa e os seus clientes, e por último, aumentar a produtividade e o controlo interno da empresa, assim como actualizar os sistemas de gestão de qualidade que a empresa possa ter implementado. Terziovski e Samson (1997) publicaram um dos mais rigorosos estudos, que analisa a relação entre a certificação ISO 9000 e o desempenho organizacional, na presença e na ausência de um ambiente de gestão da qualidade total (TQM). O estudo foi efectuado numa amostra de 952 empresas industriais australianas e de 379 da Nova Zelândia, e concluíram que a certificação ISO9000 não tem por si só um efeito significativo no desempenho organizacional. Os autores afirmam que a principal motivação para as empresas pretenderem obter a certificação da qualidade é a possibilidade de esta lhes abrir portas a novos clientes e a novos mercados que seriam difíceis de atingir de outro modo. Jones et al. (1997) sugerem que as empresas se certificam porque se consideram obrigados a fazê-lo. No entanto, as empresas com um conceito mais desenvolvido da qualidade obtêm mais benefícios.

Terziovski and Samson (1997) constataram ainda, que embora, a ISO9000 tenha pouco impacto, ou mesmo nulo no desempenho da empresa, ela pode contribuir para um melhor desempenho organizacional e criar um clima de mudança dentro da empresa. No entanto, também a liderança, a gestão dos recursos humanos e o enfoque no cliente têm um efeito significativo no desempenho organizacional (Samson and Terziovski, 1999). Do mesmo modo, Sun (1999) mostrou que as práticas TQM, como a liderança, o desenvolvimento dos recursos humanos e a qualidade da informação, contribuem para o aumento da satisfação do cliente e o desempenho empresarial.

Para Gotzamani e Tsiotras (2002), o desempenho é mais influenciado por elementos como: a liderança, as competências dos colaboradores e as relações com clientes. Já Abraham et al. (2000) sugerem que a certificação oferece poucas garantias de resultados de desempenho eficazes, se não forem acompanhados por uma mudança na liderança, comunicação e na estrutura da organização. Para Casadesús e Gimenez (2000), o processo de certificação ISO

9000 representa uma evolução na forma como uma empresa e o seu sistema de gestão da qualidade devem ser geridos, e desta forma ser a chave para o sucesso da gestão empresarial. Deste modo, uma implementação consistente de um sistema de gestão da qualidade contribui significativamente para um melhor desempenho financeiro (Corbett et al., 2005).

No entanto, a forma como a empresa implementa as respectivas normas e introduz as alterações distingue-a em relação às outras empresas, i.e., as empresas que utilizam as normas na sua actuação diária como um catalisador para um melhor desempenho operacional na implementação ISO9000 faz toda a diferença (Naveh and Marcus, 2005). A motivação interna e de gestão para adoptar a ISO9000 tem muitas vezes um efeito positivo sobre a probabilidade de uma organização certificada alcançar uma configuração de eficácia com um melhor desempenho (Boiral e Amara, 2009).

Este estudo contribui para a literatura sobre a motivação interna das empresas para alcançar a certificação pela norma ISO 9000, através da realização de uma análise comparativa de questionários respondidos por uma amostra de empresas certificadas e uma amostra estratificada de empresas não certificadas. Em particular, avaliamos a percepção dos entrevistados acerca dos elementos de qualidade no questionário que lhe estão associados e qual a probabilidade da empresa ser certificada. Na literatura, a comparação de questionários respondidos por empresas certificadas e não certificadas são geralmente realizadas utilizando testes *t* de Student convencionais (veja, e.g., Lima et al. de 2000; Gotzamani e Tsiotras, 2001; Martinez-Costa et al, 2008). Em contraste, neste estudo, comparamos as respostas fornecidas por empresas certificadas e não certificadas utilizando modelos preditivos multivariados. Conceptualmente, há uma diferença sutil entre as respostas em comparação com os testes *t* e modelos preditivos multivariados (ou univariados). O primeiro deve ser utilizado quando a conjectura que a certificação conduz a diferentes respostas ao questionário, este último deve ser empregado quando acreditamos que a percepção dos entrevistados para os itens do questionário conduz à certificação. Ou seja, em vez de perguntar se as empresas certificadas fornecem respostas que sejam estatisticamente diferentes daquelas fornecidas pelas empresas não certificados, tentamos entender se estas respostas preveem, ou estão associadas à probabilidade da empresa ser certificada ou não. Utilizando uma abordagem multivariada levámos em conta a correlação entre as respostas e identificámos o melhor conjunto de respostas ao questionário para prever essa probabilidade. Usando um questionário, respondido por 172 empresas certificadas e 172 empresas não-certificados, nós mostramos que as duas abordagens oferecerem perspectivas diferentes. De acordo com os testes *t*, a maioria das respostas das empresas certificadas e empresas não certificadas são significativamente

diferentes. Por outro lado, a regressão logística multivariada, sugere que apenas três itens do questionário são suficientes para prever se uma empresa é certificada ou não. A análise multivariada de árvores de decisão indica que um quarto elemento, também explica a probabilidade de certificação.

Este artigo está organizado da seguinte forma: a próxima secção descreve o questionário utilizado no estudo empírico, onde uma análise preliminar dos dados também é apresentada. A secção 3 mostra os resultados obtidos pelos dois modelos multivariados: a regressão logística paramétrica e árvores de decisão não paramétricas. Finalmente, a secção 4 apresenta algumas considerações finais.

2. Descrição de dados

2.1. Universo e Amostra

Para a nossa análise utilizámos dados obtidos no IPAC- Instituto Português de Acreditação. O nosso universo e objecto de estudo são as empresas certificadas com sistema de gestão de qualidade, i.e., certificadas pela NP EN ISO9001, em Portugal (Região de Lisboa e Médio Tejo). O nosso universo é composto por 1015 empresas, no entanto, e porque, algumas dessas empresas eram, por exemplo, um grupo de empresas, associações sem fins lucrativos, instituto de emprego e formação profissional, no final, enviámos o questionário a 666 empresas portuguesas, e tentámos obter o máximo possível de respostas válidas. Enviámos o questionário por e-mail para todas essas empresas, e posteriormente fizemos um contacto telefónico, obtivemos 172 respostas de empresas certificadas, o que constitui uma taxa de resposta de 26%. Uma taxa próxima de 25% é um êxito *standard* e é superior ao mínimo sugerido por Malhotra e Grover (1998). Como o objectivo deste estudo é compreender a relação entre as respostas ao questionário e a certificação pela ISO9000, procurámos uma amostra estratificada de empresas dispostas a responder ao questionário. Embora nenhum processo de correspondência seja perfeito, ambas as amostras têm características razoavelmente semelhantes de acordo com sector de actividade, idade aproximada e número aproximado de empregados. No final, a nossa amostra total era composta por 172 empresas certificadas e 172 empresas não-certificadas.

2.2. Questionário

O nosso desenho de investigação determinou a inclusão de dezassete variáveis no questionário, relativas a opiniões dos inquiridos sobre o impacto da implementação da certificação da qualidade nos vários aspectos da actividade da empresa e sobre a envolvente económica. As questões abarcam diferentes elementos de um sistema de gestão da qualidade: relação de clientes (questões 1, 2 e 3), processo de gestão (questões 4 e 5); gestão de recursos humanos (questões 6 e 7); custos da qualidade (questões 8 a 13); planeamento estratégico da qualidade (questões 14 e 15); e relação de fornecedores (questões 16 e 17).

Para medir, a percepção que as empresas da nossa amostra têm relativamente ao questionário apresentado, utilizámos a escala de Likert. Apresentamos afirmações e solicitamos que reflectam a sua opinião, elegendo um dos cinco pontos ou categorias de escala. O formato que utilizamos no nosso questionário, podem ser resumidas numa escala ordinal com cinco categorias, de acordo com a seguinte escala: “1- Discordo totalmente”; “2- Em Desacordo”; “3-Não concordo, nem discordo”; “4- De acordo”; “5- Concordo Totalmente”.

Através do alpha de Cronbach, foi avaliada a consistência interna das respostas. Um alpha de Cronbach de 0.7 a 0.8 são geralmente considerados como satisfatórios (ver, e.g., Bland e Altman, 1997). Foi obtido um coeficiente de confiabilidade da escala de 0,73 para o questionário respondido por empresas certificadas, e 0,81 para o questionário respondido por empresas não-certificadas, o que implica que as escalas de medição são consistentes e fiáveis.

2.3. Análise preliminar

A fim de compreender a relação entre a certificação da qualidade e a opinião dos respondentes ao questionário, iniciamos a análise, estudando as propriedades das respostas individualmente. O quadro 1 apresenta a média da amostra das respostas dadas pelas empresas certificadas e empresas não certificadas, t de Student para a hipótese nula de que a diferença entre estes é zero, e os correspondentes *p*-value.

statement	Mean response		<i>t</i> -statistic	<i>p</i> -value	
	certified	non-certified			
1	3.942	3.494	5.482	<0.001	**
2	3.680	3.355	3.393	0.001	**
3	4.349	3.756	7.481	<0.001	**
4	3.674	3.444	2.288	0.023	*
5	4.326	3.855	6.222	<0.001	**
6	4.234	3.727	5.083	<0.001	**
7	4.368	3.842	6.192	<0.001	**
8	3.426	3.257	1.639	0.102	
9	3.690	3.357	3.335	0.001	**
10	3.936	3.599	3.548	<0.001	**
11	4.006	3.626	4.956	<0.001	**
12	3.686	3.395	2.907	0.004	**
13	3.900	3.637	2.821	0.005	**
14	4.576	3.988	7.724	<0.001	**
15	4.333	3.936	5.102	<0.001	**
16	3.731	3.424	3.226	0.001	**
17	2.824	2.776	0.485	0.628	

Quadro 1 - Médias amostrais das escalas de empresas certificadas e não certificadas, comparação de duas médias amostrais Testes *t* e correspondentes *P*-value. Um (dois) asterisco (s) indicam que a diferença entre as médias é estatisticamente significativa, a 5% (1%) de nível de confiança.

O quadro 1 mostra que, em todos os itens, as respostas médias dadas pelas empresas certificadas é maior do que das não-certificadas. Excluindo as questões 8 (Introdução de produtos inovadores de melhor qualidade e mais baratos, deslocando produtos industriais de mão-de-obra intensiva para países onde é mais barata a força de trabalho) e a questão 17 (Os custos da qualidade da empresa devem ser controlados e distinguidos dos outros custos), há uma diferença estatisticamente significativa entre as respostas das empresas certificadas e não certificadas com nível de significância de 5%. Os resultados sugerem que as empresas certificadas tendem a dar maior importância aos elementos de qualidade medidos por este instrumento.

	st1	st2	st3	st4	st5	st6	st7	st8	st9	st10	st11	st12	st13	st14	st15	st16	st17
	1																
	0.47	1															
	0.31	0.28	1														
	0.14	0.07	0.18	1													
	0.33	0.35	0.34	0.27	1												
	0.32	0.32	0.24	0.07	0.39	1											
	0.32	0.24	0.24	0.23	0.39	0.36	1										
	0.24	0.36	0.13	0.01	0.17	0.25	0.16	1									
	0.23	0.08	0.17	0.21	0.29	0.17	0.25	0.11	1								
0	0.10	0.05	0.19	0.07	0.20	0.07	0.14	0.13	0.20	1							
1	0.25	0.23	0.23	0.12	0.38	0.28	0.20	0.17	0.18	0.46	1						
2	0.27	0.19	0.17	0.05	0.15	0.09	0.11	0.29	0.13	0.19	0.20	1					
3	0.19	0.18	0.21	0.27	0.28	0.22	0.14	0.08	0.18	0.24	0.38	0.19	1				
4	0.34	0.31	0.36	0.25	0.47	0.31	0.38	0.17	0.30	0.30	0.40	0.24	0.42	1			
5	0.29	0.35	0.32	0.24	0.45	0.30	0.33	0.17	0.15	0.29	0.31	0.11	0.36	0.51	1		
6	0.24	0.17	0.20	0.24	0.24	0.24	0.14	0.29	0.20	0.11	0.08	0.10	0.07	0.20	0.13	1	
7	0.07	0.12	0.10	0.07	0.06	0.07	0.03	0.08	0.08	0.15	0.09	0.27	0.08	0.03	0.12	0.12	1

Quadro 2- Matriz de correlação das respostas ao questionário

O quadro 2 mostra as correlações bivariadas das respostas ao questionário. Pode-se observar que vários pares de respostas são substancialmente correlacionadas. Por exemplo, as respostas à questão 1 ("Ao implementar o sistema de gestão da qualidade, ele foi percebido pelos vossos clientes ") e a questão 2 (" Conseguem-se mais facilmente fidelizar clientes ao ser uma empresa com certificação da qualidade. ") apresentam um coeficiente de correlação de 0,47. Este valor não é inesperado uma vez que ambas as questões medem a opinião dos entrevistados sobre a relação com o cliente.

As respostas à questão 14 ("O planeamento estratégico é um processo contínuo que gera os objectivos essenciais da empresa. Estes objectivos deverão conduzir a um maior cultura de desempenho dentro da empresa. ") e a questão de 15 ("Avaliação de desempenho e mais fortes alianças estratégicas, desenvolvendo novas áreas de negócio nas principais competências, deve ser melhorado. ") também apresentam uma alta correlação de 0,51. Estas questões evidenciam a opinião dos respondentes sobre o planeamento estratégico da qualidade, por isso é natural que os respondentes mais (menos) susceptíveis a esse elemento de qualidade irá proporcionar maior (menor) valor em ambas as escalas.

Por outro lado, muitos pares de respostas parecem não ser completamente correlacionadas. Por exemplo, a questão 4 ("Existe necessidade de reduzir o período de desenvolvimento de

novos produtos ou serviços.") e a questão 8, ("custos da qualidade da empresa devem ser controlados e distinguidos dos outros custos. ") têm uma correlação não significativa ($p = 0,83$). Isso não é inesperado, pois a partir da questão 4 a percepção recai sobre a importância do processo de gestão, enquanto questão 8 ajuíza sobre o parecer da medição e controlo dos custos da qualidade. Mais interessante, alguns itens que medem a mesma dimensão da qualidade parecem ser completamente não correlacionados. Por exemplo, a questão 8 ("custos da qualidade da empresa devem ser controlados e distinguidos dos outros custos ") e a questão 13 ("Pode haver uma falha no controlo de custos da qualidade, estando eles a ser mal classificados "), ambas medem a percepção para os custos da qualidade, mas apresentam uma correlação não-significativa bivariada $0,08$ ($p = 0,14$). Estes resultados sugerem que existe uma grande distância entre a vontade de controlar custos da qualidade e da capacidade de realizá-la com sucesso.

3. Modelos Multivariados

3.1. Modelo de Regressão Logística

A regressão logística (também conhecida por modelo logístico ou logit model) é utilizada para prever a probabilidade de ocorrência de um evento por ajustamento de dados a uma curva de função logística. É um modelo generalizado linear utilizado para a regressão binomial. Assim como muitas formas de análise de regressão, ele faz uso de diversas variáveis preditivas que podem ser numéricas ou categóricas.

Deste modo, a regressão logística encontra a equação que melhor prevê o valor de uma variável nominal em função de um conjunto de variáveis independentes. O objectivo é ver se a probabilidade de obtenção de um determinado valor da variável nominal está associada a qualquer uma das variáveis independentes. Suponha que x_1, x_2, \dots, x_n é um conjunto de variáveis independentes, b_0, b_1, \dots, b_n é um conjunto de parâmetros, e $E[y]$ é o valor esperado da variável dependente y .

A equação 1 da regressão logística é:

$$E[y] = \frac{1}{1 + e^{-(b_0 + \sum_{i=1}^n x_i b_i)}}$$

Os estimadores dos parâmetros b_0, b_1, \dots, b_n são obtidos com técnicas de probabilidade máxima.

	Coefficient	t-statistic	p-value	
statement 1	0.257	1.280	0.199	
statement 2	-0.110	-0.620	0.533	
statement 3	0.724	3.990	<0.001	**
statement 4	-0.061	-0.400	0.686	
statement 5	0.210	0.920	0.355	
statement 6	0.189	1.200	0.231	
statement 7	0.448	2.380	0.017	*
statement 8	-0.144	-0.910	0.360	
statement 9	-0.029	-0.190	0.848	
statement 10	0.133	0.780	0.433	
statement 11	0.226	1.030	0.304	
statement 12	0.146	0.950	0.343	
statement 13	-0.224	-1.240	0.214	
statement 14	0.636	2.660	0.008	**
statement 15	0.009	0.040	0.968	
statement 16	0.135	0.850	0.395	
statement 17	-0.119	-0.790	0.432	
Constant	-10.145	-7.350	<0.001	**
LR test		104.710	<0.001	**

Quadro 3- Coeficientes, *t-Statistic* e *P-Values* da Regressão Logística, utilizando as respostas ao questionário. Um (dois) asterisco (s) indicam que a diferença entre as médias é estatisticamente significativa, a 5% (1%) de nível de confiança. A última linha mostra a relação entre probabilidade *t-statistic* e correspondente *p-value*.

Foi realizada uma regressão logística numa variável nominal, que assume o valor 1 se a empresa é certificada e o valor 0 se a empresa não é certificada, em função das respostas ao questionário. O quadro 3 mostra os coeficientes estimados da regressão logística, os testes-*t* para a hipótese nula de que a probabilidade de certificação não está associado com uma resposta ao questionário, e os correspondentes *p-values*. A última linha dá a probabilidade do teste da razão para a hipótese nula de que os coeficientes são conjuntamente zero. O correspondente *p-value* indica que o nulo é fortemente rejeitado. e, portanto, o nosso modelo se adapta significativamente melhor do que um modelo sem variáveis predictoras.

O sinal e a significância dos coeficientes são instrutivos e esclarecedores para a compreensão das diferenças entre as empresas certificadas e não certificadas. O quadro 3 indica, que a maioria dos coeficientes não estão significativamente associados com a probabilidade de certificação.

Na verdade, apenas as questões 3, 7 e 14 são significativos para explicar a probabilidade de uma empresa ser certificada com nível de significância de 5%. Os restantes coeficientes não são estatisticamente significativos nos níveis convencionais. A questão 3 pede a opinião dos entrevistados sobre "Junto com a globalização do mercado, houve um aumento da competitividade, resultando daí clientes mais exigentes". Esta questão avalia a percepção do entrevistado para a qualidade de elemento "relações com clientes". O coeficiente correspondente é positivo, o que implica que os maiores valores das respostas na escala Likert estão associados com a certificação. Por outro lado, a questão 7 ("A maior motivação dos colaboradores proporciona um aumento no compromisso e na responsabilidade individual perante a empresa.") representa as medidas de sensibilização do entrevistado em relação à gestão de recursos humanos. A questão 14 avalia a opinião sobre "O planeamento estratégico é um processo contínuo que gera objectivos essenciais da empresa. E esses objectivos deverão conduzir a uma cultura de alto desempenho dentro da empresa". Esta questão evidencia a percepção do entrevistado acerca do planeamento estratégico da qualidade. Os coeficientes para as questões 7 e 14 também são positivos, indicando que para maiores valores das respostas aumenta a probabilidade de certificação. Curiosamente, todas as questões relacionadas com a relação de fornecedores, processo de gestão e custos da qualidade, não são estatisticamente significativos para explicar a probabilidade de certificação.

3.2. Modelo de Árvore de Decisão

Os classificadores baseados na árvore de decisão são um dos ramos na área da inteligência artificial. Mais concretamente, pertencem a um sub-campo de aprendizagem de máquina, isto deve-se, à sua habilidade de aprender através de exemplos com o objectivo de classificar registos numa base de dados. Assim, uma árvore de decisão é induzida a partir de um conjunto de exemplos de treino onde as classes são previamente conhecidas, a sua estrutura é organizada de tal forma que: (i) cada nó interno (não-folha) é rotulado com o nome de um dos atributos de previsão; (ii) os ramos ou arestas se saíem de um nó interno são rotulados com valores do atributo daquele nó; (iii) cada folha é rotulada com uma classe, a qual é a classe prevista para exemplos que pertençam aquele nó folha. O processo de classificação de um exemplo ocorre fazendo aquele exemplo "caminhar" pela árvore, a partir do nó raiz, procurando percorrer arcos que unem os nós, de acordo com as condições que estes mesmos arcos representam. Ao atingir um nó folha, a classe que rotula aquele nó folha é atribuída àquele exemplo.

Um pressuposto fundamental para a utilização de uma regressão logística é a especificação correcta do valor esperado da variável nominal, o que é mostrado na Equação 1. Utilizando uma incorrecta forma funcional para essa expectativa pode levar a resultados distorcidos na avaliação da relevância estatística das respostas ao questionário. Por esse motivo, é interessante perceber a relação entre a certificação e as respostas ao questionário utilizando um modelo que não requer *a priori* a escolha de uma forma funcional para o valor esperado da variável nominal. Esta função pode ser realizada com modelos não-paramétricos, em que a relação entre a variável de interesse e variáveis explicativas (ou independentes) não é predeterminado por nós, mas é procedente a partir da informação fornecidas pelos dados. Um modelo popular não paramétrico para prever as variáveis nominais é a classificação de árvores (Breiman na al.1984; Quinlan, 1986). Nesta abordagem, o conjunto de dados originais é recursivamente dividido em pequenos subconjuntos mutuamente exclusivos, e os modelos são representados por uma sequência lógica de testes *if-then-else* sobre os atributos das observações.

Suponha-se que temos um conjunto de observações (i.e., empresas), descrito por um vector de atributos (i.e., as respostas ao questionário), e que essas observações pertencem a cada uma das duas classes (i.e., empresas certificadas e empresas não-certificadas). O objectivo de uma árvore de classificação é separar tanto quanto possível as observações que pertencem a cada classe daqueles que pertencem a outra através de uma sequência binária de divisão dos dados. O algoritmo começa com um nó raiz que contém todas as observações. Em seguida, o algoritmo é repetido em todas as divisões binárias possíveis a fim de encontrar o atributo e o correspondente valor de *cut-off* que dá a melhor separação num dos lados que tem principalmente as observações de uma classe e as restantes observações a partir do outro. O atributo ideal e o valor *cut-off* definem-se:

O p indica o número de observações de uma classe e n o número de observações de outra classe contida num determinado nó. A *entropia* $E(p,q)$ daquele nó é definida como:

$$E(p, q) = -\frac{p}{p+n} \log_2 \left(\frac{p}{p+n} \right) - \frac{n}{p+n} \log_2 \left(\frac{n}{p+n} \right)$$

Agora, suponha que uma dada divisão binária dos dados deixa p_1 e n_1 observações de cada classe num nó filha (daughter node), e p_2 e n_2 observações de cada classe no outro. O atributo

de divisão ideal e correspondente cut-off são aqueles que maximizam o ganho de obter informações (*information gain*), deste modo:

$$gain = E(p, q) - \frac{p_1 + n_1}{p + n} E(p_1, q_1) - \frac{p_2 + n_2}{p + n} E(p_2, q_2)$$

Da informação de ganhos positiva resulta em redução de entropia. Desde a entropia caracteriza a diversidade da população num nó, maximizando a informação de ganhos resultando daí que os nós filha são mais homogéneos do que os nós pais.

Este procedimento é repetido para os novos nós filha até que não haja mais melhorias na separação de classes é alcançado ou um critério de cessação seja satisfeito. A não divisão de nós terminais é referida pelo termo figurativo de folhas, e são representadas por retângulos no esquema de representação de árvores de decisão.

A partir do nó raiz, todas as observações são encaminhadas para baixo da árvore de acordo com os valores dos atributos testados nos sucessivos nós e, inevitavelmente, terminar o seu caminho numa folha. No final, as observações são classificadas de acordo com a classe mais predominante na folha onde terminou o seu caminho. Uma descrição detalhada dos algoritmos de árvore está para além do âmbito deste trabalho. Poderá ser consultado em Witten e Frank (2005) para os detalhes técnicos dos algoritmos utilizados nesta investigação.

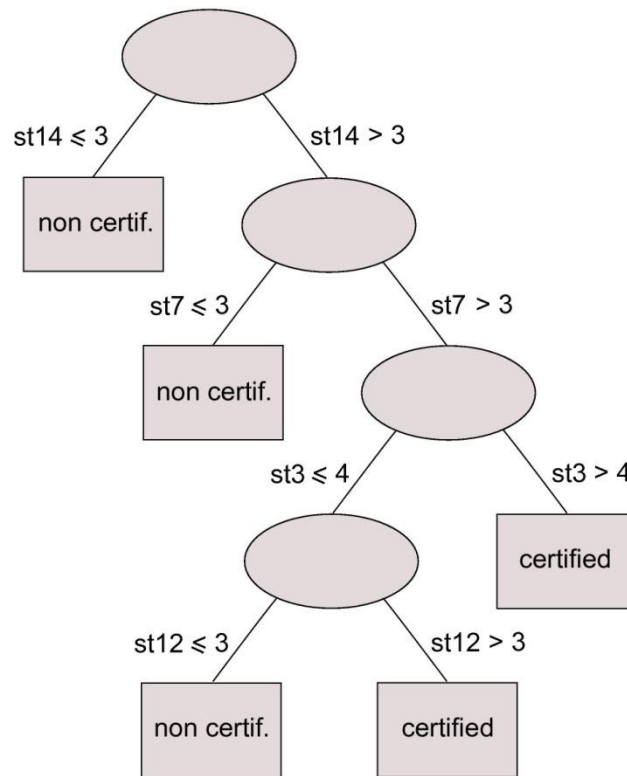


Figura 1- Modelo de Árvores de Decisão para as respostas ao questionário

A Figura 1 mostra o modelo de árvore de classificação para as respostas ao questionário. A interpretação destes modelos é simples e intuitiva. Primeiro, no nó de raiz, o modelo pergunta acerca da resposta à questão 14. Se a resposta à questão 14 na escala de Likert é menor ou igual a 3, o modelo prevê que a empresa não é certificada e o ramo termina aí. Se ocorrer o oposto, o modelo, em seguida, centra-se na resposta à questão 7. Se a resposta a essa afirmação é menor ou igual a 3, o modelo prevê novamente que a empresa não é certificada. Se ocorre o contrário, o modelo, em seguida, centra-se na informação sobre a resposta à afirmação 3. Se a resposta à afirmação 3 é maior que 4, o modelo prevê que a empresa é certificada. Se a resposta a esta questão for menor ou igual a 4, o modelo finalmente tem informações sobre a resposta à afirmação 12. Se a resposta a esta afirmação é menor ou igual a 3, o modelo prevê que a empresa não é certificada, caso contrário, prevê que a empresa é certificada.

Esta estrutura de árvore permite-nos tirar as seguintes conclusões. Primeiro, pequenos valores de resposta à questão de 14 estão associados a empresas não-certificadas, uma vez que o ramo $Q14 \leq 3$ leva a uma folha “ não- certificada”. Pela mesma razão, os pequenos valores da resposta à pergunta 7 também estão associados a empresas não-certificadas. Valores grandes

(pequenos) das respostas às afirmações 3 e 12 estão associados a empresas(não) certificadas, uma vez que as condições $Q3 > 4$ e $Q12 > 3$, respectivamente, levam a folhas “certificadas“, respectivamente.

As restantes questões não são relevantes uma vez que não estão presentes na estrutura da árvore. O modelo de árvore corrobora os resultados da regressão logística, em que apenas as questões 3, 7 e 14 foram significativamente associados à probabilidade de certificação. Além disso, de acordo com a regressão logística, e as árvores de decisão sugerem que para valores mais elevados na escala Likert dessas respostas às questões, estão associadas à certificação. No entanto, o modelo de árvore também sugere que os valores mais elevados nas respostas à questão 12 estão associados à certificação, enquanto esta questão não é estatisticamente significativa no modelo de regressão logística no nível de significância convencional ($p = 0:343$). Esta questão é sobre a opinião dos inquiridos se: "Na realidade, os empresários saberão quanto é que a empresa poderá perder por falta de qualidade", que avalia a percepção para o elemento " custos da qualidade ".

4. Conclusões

Este artigo analisou o impacto da certificação da qualidade nas respostas a um questionário respondido por uma amostra de empresas certificadas pela ISO9001 e uma amostra estratificada de empresas não certificadas, utilizando modelos preditivos multivariados. Com esta abordagem, avaliámos a percepção dos entrevistados acerca dos elementos de qualidade no questionário que lhe estão associados e qual a probabilidade da empresa ser certificada.

Verificámos que, em contraste com a literatura anterior, que normalmente adopta técnicas univariadas, que tratam as respostas ao questionário como uma única variável independente, reconhecemos que existe um alto grau de relacionamento entre as respostas ao questionário e as abordagens multivariadas.

O modelo de regressão logística paramétrico mostrou que a opinião dos entrevistados em três itens do questionário estão significativamente associados à probabilidade de certificação. Em particular, as empresas em que o entrevistado manifesta maior sensibilização: na relação de clientes; no planeamento estratégico da qualidade e na motivação aos colaboradores tiveram maior probabilidade de certificação. Curiosamente, todas as questões que estão relacionadas com os elementos da qualidade, como relação de fornecedores; processo de gestão e custos da qualidade não foram significativos para explicar a probabilidade de certificação. Os resultados fornecidos pelo modelo de regressão logística foram corroborados por um modelo de decisão

não paramétrico. No entanto, o modelo de árvores de decisão sugeriu ainda que na opinião dos entrevistados acerca da percepção dos gestores sobre as perdas devido à falta de qualidade (i.e. custos da qualidade) poderia prever a probabilidade de certificação.

O leitor deve observar que este estudo se baseia em dados de um único país e, claro, estes resultados capturam muitas idiosincrasias do seu ambiente económico e empresarial. Seria interessante perceber se este tipo de análise revela algumas regularidades em diferentes países.

Referências Bibliográficas

- Abraham, M., Crawford, J., Carter, D., Mazotta, F. (2000). Management decisions for effective ISO 9000 accreditation. *Management Decision* 38, 182-193.
- Bland, J.M., Altman, D.G. (1997). Statistics notes: Cronbach's alpha. *British Medical Journal* 314, 572.
- Boiral, O., Amara, N. (2009). Paradoxes of ISO9000 performance: A configurational approach. *Quality Management Journal* 16(3), 37-60.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. (1984). *Classification and regression trees*. Wadworth International Group, Belmont, California.
- Casadésus, M., Giménez, G. (2000). The benefits of the implementation of the ISO9000 standard: Empirical research in 288 Spanish companies. *The TQM Magazine* 12, 432-41.
- Corbett, C., Montes-Sancho, M.J., Kirsch, D.A. (2005). The financial impact of ISO9000 certification in the United States: An empirical analysis. *Management Sciences* 51, 1046-1059.
- Erel, E., Ghosh, G.B. (1997). ISO9000 implementation in Turkish industry. *International Journal of Operations & Production Management* 17, 1233-1246.
- Gotzamani, K.D., Tsiotras, G.D. (2001). An empirical study of the ISO9000 standards' contribution towards total quality management. *International Journal of Operations & Production Management* 21, 1326-1342.
- Gotzamani, K.D., Tsiotras, G.D. (2002). The true motives behind ISO 9000 certification. Their effect on overall certification benefits and their long term contribution towards TQM. *International Journal of Quality & Reliability Management* 19, 151-169.
- Heras, I., Casadésus, M., Garvin, P.M. (2002). ISO 9000 Certification and the bottom line: A comparative study of the profitability of Basque region companies. *Managerial Auditing Journal* 17, 72-78.
- Jones, R., Arndt, G. and Kustin, R. (1997). ISO 9000 among Australian companies: impact of time and reasons for seeking certification on perceptions of benefits received. *International Journal of Quality & Reliability Management* 14, 650-60.
- Lima, M.A.M., Marcelo Resende, M., Hasenclever, L. (2000). Quality certification and performance of Brazilian firms: An empirical study. *International Journal of Production Economics* 66, 143-147.
- Quinlan, J.R. (1986). Induction of decision trees. *Machine Learning* 1, 81-106.
- Martinez-Costa, M., Martinez-Lorente, A.R., Choi, T.Y. (2008). Simultaneous consideration of TQM and ISO 9000 on performance and motivation: An empirical study of Spanish companies. *International Journal of Production Economics* 113, 23-39.
- Sun, H. (1999). Diffusion and contribution of total quality management: An empirical study in Norway. *Total Quality Management* 6, 901-914.
- Terziowski, M., Samson, D. (1997). The business value of quality management systems certification: Evidence from Australia and New Zealand. *Journal of Operations Management* 15, 1-18.
- Terziowski, M., Samson, D. (1999). The link between total quality management practice and organizational performance. *International Journal of Quality & Reliability Management* 16, 226-237.
- Teixeira Quirós, J.; Justino, M.R.(2009): Certification of quality management system under ISO9000 versus business bottom line: empirical evidence..XXIII Congreso, *European Association of Management and Business Economics (AEDEM)*- Sevilla. ISBN: 978-84-7356-609-4.
- Teixeira Quirós, J.; Almacá, J.; Justino, M.R.(2009): Certificação de qualidade versus resultado da empresa: evidencia empirica. *XIX Jornadas Hispano Lusos de Gestão Científica*. Universidad de Jaén. Baeza. ISBN: 978-84-691-8776-0.
- Teixeira Quirós, J.; Justino, M.R.(2010): TQM and ISO9000 versus business income: an empirical study. *TMQ Qualidade*. Número 1. ISSN: 1647-6263. PP. 46-62. Edições Silabo.
- Teixeira Quirós, J.; Almacá, J.A.; Justino, M.R.(2010): How quality affects the bottom line?: A literatura review. *Intangible Capital*. ISSN: 1697-9818. 6(2). 258-271.
- Tsiotras, G., Gotzamani, K. (1996). ISO 9000 as an entry key to TQM: The case of Greek industry. *International Journal of Quality* 13, 64-76.

Witten, I.H., Frank, E. (2005). *Data mining: practical machine learning tools and techniques*. Morgan Kaufmann Publishers.

Curriculum Vitae:

Joaquín Texeira Quirós é doutorado em Ciências Económicas e Empresariais pela Universidade de Extremadura. É professor titular na Faculdade de Ciências Económicas e Empresarias da Universidade da Extremadura e professor na Universidade Autónoma de Lisboa. É director do CEXECI- Centro Extremeño de Estudios y Cooperación con Iberoamérica. As suas áreas de interesse são: Gestão da Qualidade, Custos da Qualidade, Inovação e Qualidade.

Maria Rosário Fernandes Justino é doutorada em Ciências Económicas e Empresariais com menção europeia (Doutor Europeu). É professora no ISCAL – Instituto Superior de Contabilidade e Administração de Lisboa do Instituto Politécnico de Lisboa, e professora convidada na Universidade de Extremadura e na Universidade Autónoma de Lisboa. É membro investigador do Centro de Investigação em Sociologia Económica e das Organizações (SOCIOUS) do ISEG da Universidade Técnica de Lisboa. As suas áreas de interesse são: Gestão da Qualidade, Custos da Qualidade, Inovação e Qualidade.

Authors Profiles:

Joaquín Texeira Quirós is professor at the Faculty of Economics and Business of the University of Extremadura (Spain) and at the University Autónoma of Lisbon. He received his PhD degree from the University of Extremadura in Economics and Business Science. He is director of CEXEXI- Centro Extremeño de Estudios y Cooperación con Iberoamérica. His research interests are in the areas of Quality Management, Quality Costs, Innovation and Quality.

Maria Rosário Justino is professor at ISCAL of the Polytechnical Institute of Lisbon and University of Extremadura. She received her PhD degree from the University of Extremadura (Spain) in Economics and Business Science. Her research interests are in the areas of Quality Management, Quality Costs, Innovation and Quality.

