

CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ

Rio de Janeiro - RJ - Brasil

ISSN impresso 1413-9022 / ISSN on-line 2317-4536 - v. 32, p. 01 - 14, 2012

APLICAÇÕES DO CONTROLE ESTATÍSTICO MULTIVARIADO DA QUALIDADE: CONTROLANDO GARRAFEIRAS PLÁSTICAS NUMA EMPRESA INDUSTRIAL

Márcio Botelho da Fonseca Lima
Universidade Federal da Paraíba
tismalu@uol.com.br

Roberta de Lourdes Silva dos Santos
Universidade Federal da Paraíba
robertalss@globocom

Erivaldo Lopes de Souza
Universidade Federal da Paraíba
elopesouza@hotmail.com

Luiz Bueno da Silva
Universidade Federal da Paraíba
bueno@ct.ufpb.br

Resumo

Este artigo tem como objetivo utilizar técnicas estatísticas multivariadas para aperfeiçoar o processo de monitorar e controlar a qualidade numa empresa que produz garrafeiras plásticas. Assim, a utilização das técnicas de análise de componentes principais (ACP) e da carta T^2 de Hotelling pode sumarizar parte das informações relevantes desse processo. Esta pesquisa foi realizada com interação de empresas, pois se baseou numa auditoria de uma grande companhia multinacional de bebidas. O modelo de ACP pode ser obtido via software R. Através de funções usadas neste programa, tem-se o valor da proporção da variância explicada e dos coeficientes usados no modelo. Assim pode-se obter o novo conjunto de dados com variáveis pouco correlacionadas. A carta de controle T^2 de Hotelling também pode ser construída via software R. Produziram-se, enfim, dois resultados de considerável importância: os escores dos componentes principais e um gráfico T^2 de Hotelling adaptado, evidenciando a relação entre as dez variáveis analisadas.

Palavras-chave: Controle estatístico da qualidade, Carta T^2 de Hotelling, Análise de componentes principais.

1. Introdução

No meio empresarial, o advento das parcerias entre empresas, firmadas com o objetivo de melhor gerenciar a cadeia de valor dos produtos, foi um fator relevante que contribuiu para o aumento da exigência de melhores níveis de qualidade.

Quando o foco da qualidade está na conformidade das características obtidas com aquelas planejadas antes da sua fabricação, surge a necessidade de utilização de um controle eficiente dessas características. É fundamental que sejam monitoradas com ferramentas adequadas para aquela situação.

Se o monitoramento da conformidade envolve mais de uma característica da qualidade, e se essas características estão correlacionadas, não será adequada a utilização de cartas de controle univariadas. A eficiência desse monitoramento precisa levar em conta a correlação entre as variáveis. Nesses casos, é aconselhável o uso de cartas de controle multivariadas. Além disso, geralmente um produto industrial possui mais do que uma característica de qualidade (PAN; LEE, 2010).

Nas linhas que se seguem, primeiramente, na seção 2, é apresentada a abordagem teórica, englobando noções de Controle Estatístico do Processo (CEP), o CEP multivariado, a Análise de Componentes Principais (ACP) e o gráfico T^2 de Hotelling. Em seguida, na seção 3, descreve-se a metodologia adotada, um estudo de caso. Logo após, na seção 4, são apresentados os resultados da pesquisa e finalmente na seção 5 apresentam-se as conclusões.

2. Fundamentação Teórica

Paladini (2004) afirma que a ideia principal do Controle Estatístico do Processo (CEP) é que melhores processos de produção com menos variabilidade propiciam níveis melhores de qualidade nos resultados de produção. Esse autor acrescenta que melhores níveis de qualidade resultam em custos menores.

2.1. Controle estatístico do processo (CEP)

Lopes (2007) considera que o CEP trata do uso de técnicas estatísticas em todos os estágios considerados críticos no desenvolvimento de um projeto ou da produção. Montgomery (2009), por sua vez, afirma que o principal objetivo do CEP é monitorar o desempenho de processos ao longo do tempo e detectar eventos incomuns que exerçam influência nas propriedades determinantes da qualidade do produto final.

2.2. Análise de eficiência do Gráfico de \bar{X} via *Average Run Length* (ARL)

O ARL, ou seja, o comprimento médio da sequência, também pode ser traduzido por número médio de amostras até o alarme (NMA).

2.2.1. Erros dos tipos 1 e 2 e poder de detecção do gráfico de \bar{X} barra

O erro do tipo 1, α , que ocasiona um alarme falso, consiste na probabilidade de que, estando o processo livre de causas especiais, o gráfico de \bar{X} indique uma observação na zona de ação, ou seja, acima do Limite Superior de Controle ($LSC_{\bar{X}}$) ou abaixo do Limite Inferior de Controle ($LSC_{\bar{X}}$). O erro do tipo 2, β , ocorre quando o processo, mesmo estando sujeito à influência de causas especiais, for representado por um Gráfico de \bar{X} exibindo uma observação dentro dos limites de controle.

O poder de detecção, Pd, do gráfico de \bar{X} é definido como a probabilidade deste gráfico indicar um alarme verdadeiro, ou seja, quando ocorrer uma observação acima ou abaixo dos limites de controle, caso a média do processo tenha se alterado.

Para o valor usual de $k = 3$, pode-se verificar que o valor de α é igual a 0,0027. O NMA_0 até um alarme falso é $1/\alpha$, ou seja, aproximadamente 370 amostras. Da mesma maneira, o NMA_1 até um alarme verdadeiro é $1/Pd$, pois ambas as distribuições são geométricas de parâmetros α e Pd, respectivamente, sendo suas médias iguais aos inversos desses parâmetros.

2.2.2. Escolhas do intervalo de tempo entre amostras (h), do tamanho da amostra (n), de número de amostras (m) e separação das fases I e II

O intervalo de tempo entre amostras(h) depende de vários fatores, entre os quais se destacam o tempo de produção de cada item, o custo de inspeção, a estabilidade do processo e o prejuízo de se operar um processo que sofre influência de causas especiais.

Na prática, geralmente o valor de h é escolhido de maneira arbitrária, mas deve-se considerar que valores pequenos de h implicam custos elevados com amostragem e maior incidência de alarmes falsos (COSTA *et al.*, 2004).

O NMA_1 correspondente ao alarme verdadeiro de que o processo mudou depende da intensidade do deslocamento da média do processo: por um lado, no caso em que o processo sofra uma causa especial grave e a média do processo se desloque de uma distancia de vários desvios-padrão, o NMA_1 será pequeno, pois o gráfico detectará rapidamente a presença dessa causa; por outro lado, para pequenos deslocamentos da

média, o NMA_1 será grande, indicando que esses deslocamentos dificilmente serão detectados rapidamente (SAMOHYL, 2009).

Segundo Montgomery (2009), recentemente o uso do NMA para descrever o desempenho do gráfico de \bar{X} tem sido criticado, pelo fato de a distribuição do MNA ser geométrica. Assim há duas preocupações em relação ao NMA: por um lado, o desvio-padrão do NMA é muito grande; por outro lado, a distribuição geométrica é muito assimétrica, de modo que a média da distribuição ($1/Pd$ ou $1/\alpha$) não é necessariamente um valor “típico” do comprimento da sequência.

O tamanho da amostra geralmente é pequeno, sendo $n=1, 2, 3... 9$. O número mínimo de amostras (m) necessárias à construção de novos gráficos na linha de produção (a primeira fase de implantações de gráficos de controle) é $m=25$.

De fato, a questão da escolha do tamanho da amostra é, sobretudo, de ordem econômico-financeira e não estatística. Amostras maiores são sempre mais precisas pois têm maior probabilidade de detectar alterações no processo, mas acarretam gastos mais elevados em termos de tempo e recursos além de erros de mensuração mais frequentes. Nesse sentido, na indústria em geral, as mensurações devem ser efetuadas com extremo cuidado e o valor de m deve ser mínimo. Além disso, é raro observar um valor de $n>9$, e o valor de $n=4$ parece ser de uso universal.

Na prática, a implantação e utilização do gráfico de \bar{X} passam por 2 fases de operações. Na fase I, a montagem do gráfico, inicialmente os dados são levantados e posteriormente estatísticas tais como média e desvio-padrão são calculadas. Com esses valores o gráfico de controle é montado. Na fase II, geralmente se assume que o processo é razoavelmente estável. De uma maneira geral, as causas especiais que ocorrem na fase II resultam de pequenos deslocamentos do processo, pois as verdadeiras fontes ruins de variabilidade foram sistematicamente removidas durante a fase I (MONTGOMERY, 2005).

Somente a fase I foi desenvolvida nesta pesquisa, e inúmeras observações (quatro grupos de 25 amostras de $n=4$, com $h=0,5$ horas) foram coletadas para se determinar também o valor nominal e os limites de especificação para uma determinada percentagem de itens fora das especificações, estabelecida pela empresa em estudo, totalizando 98 amostras, pois o aplicativo R detectou duas observações atípicas.

2.3. Controle estatístico do processo e dados multivariados

Para Suchs e Kennet (1998), a maioria dos dados referentes à qualidade em empresas industriais é naturalmente multivariada. Para tais autores, é comum utilizar uma abordagem onde os dados em diferentes dimensões são agregados e transformados. Entretanto, eles afirmam que o uso dessa abordagem resulta em perda de informações, dificultando a identificação de causas especiais e prejudicando o entendimento da capacidade do processo. Assim, os dados multivariados são muito mais informativos do que dados univariados referentes a diversas variáveis analisadas separadamente.

De acordo com Treasure *et al.* (2011), o controle estatístico multivariado é usado com o objetivo de eliminar a redundância em dados observados. Os autores ainda afirmam que no *Multivariate Statistical Process Control* (MSPC) é comum utilizar-se de abordagens para obter um conjunto resumido de dados que posteriormente são usados para calcular estatísticas como a T^2 de Hotelling.

Lyu e Chen (2008) afirmam que as indústrias modernas envolvem inúmeras situações que requerem o monitoramento ou controle simultâneo de duas ou mais características da qualidade relacionadas. Os autores compartilham a mesma ideia de Montgomery (2005) ao afirmarem que monitorar separadamente as características da qualidade pode levar a conclusões equivocadas.

2.4. Análise de componentes principais no controle estatístico de processo

Suchs e Kennet (1998) salientam que uma abordagem comum no controle de qualidade consiste em reduzir o número de dimensões de um conjunto de dados através de uma transformação que identifique as combinações lineares significativas entre as p variáveis disponíveis. Para esses autores, a construção de combinações lineares específicas é chamada de componentes principais. Eles afirmam que, para observações multivariadas, as componentes principais são as combinações lineares dos desvios das p variáveis em relação aos seus respectivos alvos.

Montgomery (2005) corrobora tal abordagem ao citar que os componentes principais de um conjunto de variáveis de um processo são justamente as combinações lineares obtidas a partir dessas variáveis. Esse autor também afirma que a intenção básica de obter componentes principais é propiciar um novo conjunto de direções ortogonais que defina o máximo de variabilidade dos dados originais, o que conduz a uma descrição do processo através de um número consideravelmente menor de variáveis em comparação com o conjunto original.

A principal utilidade da análise de componentes principais para o controle estatístico do processo consiste em obter um conjunto de dados onde as variáveis não apresentem correlações. A matriz de coeficientes $A_{p \times p}$, usados no sistema de equações do modelo, é a matriz composta pelos autovetores da matriz de covariância $\Sigma_{p \times p}$ ou, alternativamente, pela matriz de autovetores da matriz de correlação $P_{p \times p}$ das variáveis.

Um modelo de análise de componentes principais (ACP) apresenta a forma:

$$\begin{cases} z_1 = a_{11}x_1 + a_{21}x_2 + \dots + a_{p1}x_p \\ z_2 = a_{12}x_1 + a_{22}x_2 + \dots + a_{p2}x_p \\ \vdots \\ z_p = a_{1p}x_1 + a_{2p}x_2 + \dots + a_{pp}x_p \end{cases} \quad (1)$$

Por meio desse modelo, é possível transformar um conjunto de dados referentes às variáveis originais x_1, x_2, \dots, x_n , em um novo conjunto de dados referentes aos componentes principais z_1, z_2, \dots, z_n . Essas novas variáveis obtidas apresentam baixa correlação entre si.

Sendo $a_1 = (a_{11}, a_{12}, \dots, a_{1p})$ o autovetor correspondente ao autovalor λ_1 da matriz de covariância do conjunto de dados, tem-se $z_1 = a_1 X$, $z_2 = a_2 X$, ..., $z_p = a_p X$, onde $X = (x_1, x_2, \dots, x_p)$. Sendo I uma matriz identidade, os autovalores da matriz de covariância S podem ser obtidos utilizando a equação característica:

$$|S - \lambda I| = 0 \quad (2)$$

A seleção dos componentes principais que serão consideradas no monitoramento pode ser feita verificando a proporção de variância explicada (PVE) de cada componente. O valor do PVE_i correspondente ao i -ésimo componente principal é dado por:

$$PVE_i = \frac{\text{var}(z_i)}{\sum_{i=1}^p \text{var}(z_i)} \quad (3)$$

Se $\text{var}(z_1) = \lambda_1$, $\text{var}(z_2) = \lambda_2$, ..., $\text{var}(z_p) = \lambda_p$, então se tem:

$$PVE_i = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \quad (4)$$

2.5. Cartas de controle T^2 de Hotelling

Para Montgomery (2009), um gráfico de controle multivariado popular é o gráfico T^2 de Hotelling. Segundo esse autor, nesse gráfico monitora-se o valor da estatística T^2 para cada uma das amostras. Caso algum dos valores de T^2 esteja fora dos limites de controle, e identificada uma causa especial, então se conclui que o processo está fora de controle. Sendo n o tamanho de cada uma das amostras, \bar{x}_{jk} a média da j -ésima variável na k -ésima amostra, $\bar{\bar{x}}_j$ a média das médias das amostras da j -ésima variável, tem-se a estatística T^2 igual a:

$$T^2 = n(\bar{x}_{jk} - \bar{\bar{x}}_j)' s^{-1} (\bar{x}_{jk} - \bar{\bar{x}}_j) \quad (5)$$

Para a construção dos limites de controle do gráfico de controle T^2 de Hotelling, duas fases distintas devem ser consideradas. A primeira fase consiste na obtenção de amostras retiradas do processo com o objetivo de construir um conjunto de dados, denominados dados históricos, que serão utilizados na segunda fase. Conforme Montgomery (2005), nessa última fase são calculados, a partir dos dados obtidos na primeira fase, os parâmetros (estatísticas) utilizados para estabelecer os limites de controle para a “produção futura” com base numa análise retrospectiva dos dados.

Para Mason e Young (2002), a primeira etapa do desenvolvimento de um conjunto de dados históricos consiste na obtenção de uma série de dados preliminares, que são representados por amostras retiradas do processo quando este se encontra em controle. Em caso contrário, devem-se procurar medidas para estabilizar o processo, para só então fazer a retirada das amostras.

Para a construção dos limites de controle considerando m subgrupos preliminares de tamanho n , a expressão dos limites de controle para a primeira fase é fornecida pelas seguintes equações :

$$LSC = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha, p, mn-m-p+1} \quad (6)$$

$$LIC = 0 \quad (7)$$

Onde n representa o tamanho do subgrupo, m o número de amostras, p o número de variáveis e $F_{\alpha, p, mn-m-p+1}$ representa o percentil da distribuição F com p e $(mn-m-p+1)$ graus de liberdade, obtido a partir do α adotado.

Na segunda fase, ainda considerando subgrupos preliminares de tamanho n , $\bar{\bar{X}}$ e S são calculados a partir dos dados obtidos na primeira fase e, ao calculá-los, é necessário verificar se existem pontos fora dos limites de controle e se as causas especiais foram determinadas. Essas observações devem ser omitidas do cálculo de $\bar{\bar{X}}$ e S . Assim, para monitorar a produção futura, os limites de controle para a segunda fase são dados por:

$$LSC = \frac{p(m+1)(n-1)}{mn-m-p+1} F_{\alpha, p, mn-m-p+1} \quad (8)$$

$$LIC = 0 \quad (9)$$

O limite superior de controle do gráfico depende do valor da estatística F para p variáveis, com nível de significância α e cujo número de graus de liberdade é igual a $gl=mn-m-p+1$. O monitoramento do valor da estatística T^2 pode se tornar errôneo se as variáveis consideradas na análise forem correlacionadas. Nesses casos, é importante usar técnicas como a ACP, para obter novas variáveis que, conservando as propriedades importantes das variáveis originais, apresentem correlações tão baixas quanto possíveis.

3. Metodologia – um estudo de caso

O caso analisado consiste na determinação dos limites de controle para construção de um diagrama de controle multivariado baseado na estatística T^2 de Hotelling, para as dimensões de uma garrafeira como a indicada na figura 1. Para o esboço do gráfico de controle, foram retiradas 98 amostras de $n = 4$.

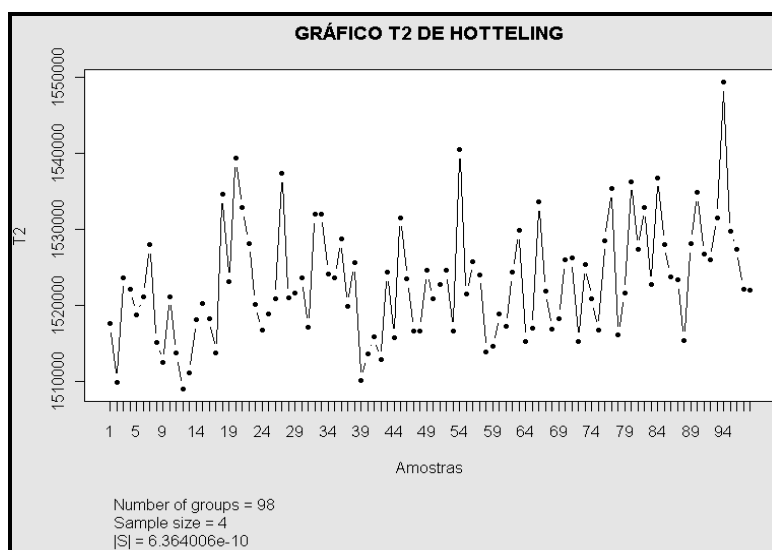
Figura 1 - Garrafeira



A determinação do intervalo de tempo entre a retirada das amostras depende de vários fatores, dentre eles o tempo de produção de cada item, custos de inspeção, a estabilidade do processo e o desperdício de recursos ao se trabalhar sob a influência de causas especiais. Considerando esses fatores, determinou-se um intervalo de 25 minutos entre as retiradas das amostras. As características-alvo que foram medidas para obtenção das amostras são listadas a seguir:

1. CE – comprimento externo;
2. CI – comprimento interno;
3. LE – largura externa;
4. LI – largura interna;
5. AT – altura total;
6. AE – altura de empilhamento;
7. CF – comprimento de fundo (Módulo de Encaixe do Fundo – Comprimento);
8. LF – largura de fundo (Módulo de Encaixe do Fundo – Largura);
9. AN – altura do ninho (Altura das Divisórias Internas);
10. AF – altura de encaixe do fundo.

Figura 2 – Diagrama de controle T^2 para as variáveis monitoradas



Fonte: Autores

Os limites de controle para as 98 amostras do diagrama T^2 de Hotelling são $LIC=0$ e $LSC=21.16$, mostrados na figura 2. Pode-se perceber que os valores da estatística T^2 para todas as amostras estão muito acima dos limites de controle calculados. É possível que os valores altos da estatística T^2 tenham sido distorcidos pela presença de correlações significativas entre as variáveis e pelo autocorrelacionamento de cada variável considerada isoladamente.

Verifica-se, via software R, que a maior parte dos pares de variáveis apresenta uma correlação acima de 0,1. Isso pode levar a erros na construção do diagrama de controle T^2 de Hotelling. É necessário substituir esse conjunto de variáveis por outro conjunto que represente bem a maior parte das variabilidades do conjunto original. Utilizando a análise de componentes principais, pode-se chegar a esse objetivo.

O modelo obtido para substituir os dados originais fornece 10 componentes principais. Ainda por meio do software R, pode-se perceber que as sete primeiras

componentes representam mais de 86 % das variabilidades do conjunto original de dados. Pode-se utilizar, então, esses sete componentes agrupados em 98 amostras de 4 elementos para conservar a estrutura original das amostras coletadas.

A partir dos coeficientes do modelo, pode-se fazer a transformação dos dados originais em dados referentes aos componentes principais que serão utilizados.

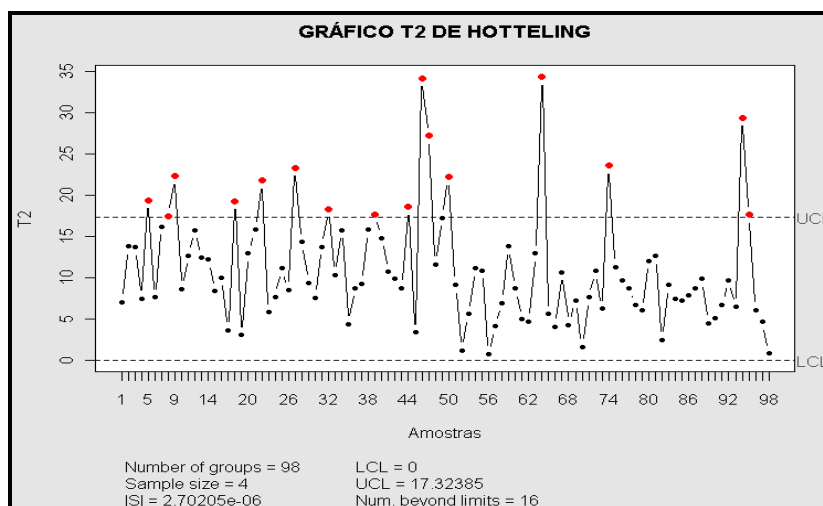
Assim, o modelo de componentes principais que será usado na transformação dos dados consiste em:

$$\begin{aligned} z1 &= -0.141*AE - 0.195*AF - 0.149*AT - 0.856*AN - 0.121*CF + 0.35*CI + 0.127*LI - 0.178*LE \\ z2 &= 0.159*AF - 0.204*AN - 0.129*CE - 0.222*CF - 0.498*CI - 0.651*LI - 0.438*LE \\ z3 &= -0.396*AN - 0.611*CI + 0.135*LI + 0.660*LE \\ z4 &= -0.287*AE - 0.431*AT + 0.174*AN - 0.655*CE - 0.155*CI + 0.316*LI + 0.363*LF \\ z5 &= -0.139*AT - 0.189*CE - 0.456*CF + 0.131*LI - 0.835*LF \\ z6 &= -0.779*AF + 0.258*AT - 0.250*CE + 0.400*CF - 0.196*CI - 0.105*LF - 0.220*LE \\ z7 &= -0.488*AF + 0.131*AN - 0.549*CF + 0.178*CI - 0.376*LI + 0.299*LF + 0.421*LE \end{aligned}$$

(10)

Utilizando o novo conjunto de dados disponível com a estrutura inicial dos dados (98 amostras de 4 elementos), tem-se na Figura 3 um diagrama com valores mais plausíveis da estatística T^2 para as amostras coletadas.

Figura 3 – Diagrama de controle T^2 para os componentes principais



Fonte: Autores

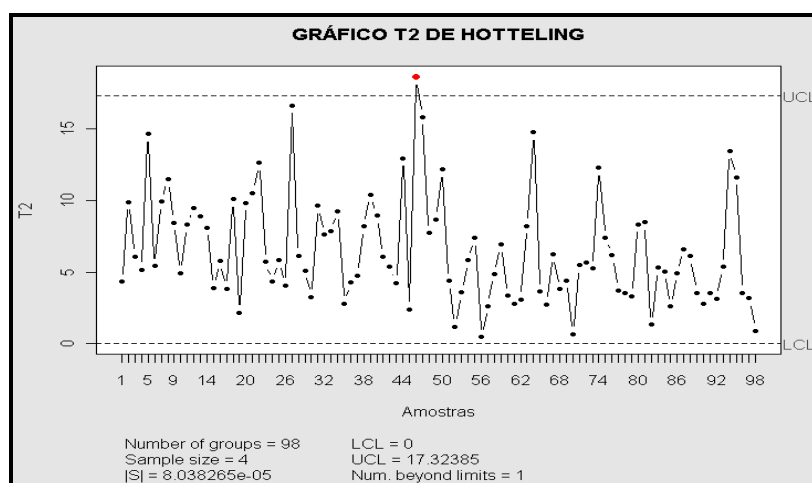
Os limites estabelecidos utilizando-se dos dados referentes aos sete componentes principais são LIC (LCL) = 0 e LSC (UCL) = 17.32, como mostra a Figura 3. Ao se adotar esses limites, têm-se 16 amostras fora dos limites de controle e 82 amostras dentro desses limites.

4. Resultados

Como a produção de garrafeiras consiste em um processo discreto parcialmente automatizado, foram construídos gráficos de controle da média com limites alargados para cada uma das variáveis em estudo (para evitar alarmes falsos), pois, segundo Costa *et al.*(2004), nesse caso ocorre o fenômeno de autocorrelacionamento de cada variável, constatado durante a realização da pesquisa. Também foram construídos gráficos da amplitude habituais e verificou-se que todas as variáveis se encontravam em controle estatístico.

O cálculo da estatística T^2 é feito a partir da matriz de covariância(s) (e variâncias). Os elementos presentes na diagonal principal da matriz representam as variâncias das observações para cada variável. Um meio relevante de reduzir o número de alarmes falsos é substituir o cálculo a partir da variabilidade dentro das amostras por aquele das variâncias das médias amostrais na diagonal principal da matriz de covariância. Ao utilizar-se esse procedimento, o novo gráfico obtido para as componentes principais é mostrado na Figura 4.

Figura 4 - Novo gráfico T^2 de Hotelling para o modelo de ACP



Fonte: Autores

Na Figura 4, somente a amostra 47 apresentou um alarme de descontrole do processo. Verificou-se que era apenas um erro de digitação devido ao grande número de amostras preliminares utilizadas.

5. Conclusões

Este trabalho buscou mostrar uma aplicação de um controle simultâneo num processo de produção de garrafeiras de plástico. A análise mostrou que os valores elevados das correlações entre as variáveis distorceram os valores calculados da

estatística T^2 para as amostras, gerando uma inconsistência entre os valores dessa estatística e os valores obtidos para os limites de controle. Essa distorção pôde ser contornada através do uso da análise de componentes principais, que possibilitou obter um novo conjunto de variáveis pouco correlacionadas, mantendo a maior parte das variabilidades do conjunto original.

Após a utilização da análise de componentes principais, mostrou-se que 16 das 98 amostras analisadas apresentaram sinais de que o processo não estava em controle estatístico, indicando ainda a presença de alarmes falsos. A redução do número de alarmes falsos pôde ser efetuada usando, na matriz de covariância, variâncias entre amostras em vez das variâncias entre observações. Realizou-se dessa forma uma nova carta de controle e constatou-se que processo estava em controle estatístico.

Em relação a um possível desdobramento desta pesquisa, torna-se importante destacar um artigo de Maranhão e Epprecht (2011), que demonstraram que um esquema de controle baseado em variáveis transformadas é preferível àquele que utiliza componentes principais, pois, geralmente, constata-se que o primeiro esquema é mais eficaz em se tratando de sinalizar descontroles no processo.

Contudo, no que tange ainda a essa sinalização, vale ressaltar que ambos os esquemas supracitados são mais eficazes do que o controle individual das variáveis observadas, além de propiciarem as vantagens de requererem uma menor quantidade de cartas de controle e de levarem em consideração a estrutura de correlação entre variáveis. Além disso, a utilização de um algoritmo de busca para frente com o objetivo de detectar observações atípicas (*outliers*) pode ser recomendada para aperfeiçoar o CEP multivariado (GARRET, 1989; FILMOZER *et al.*, 2005; GUH e SHIUE, 2008).

Referências

- COSTA, A. F. B.; EPPRECHT, E. K.; CARPINETTI, L. C. R. **Controle Estatístico de Qualidade**. São Paulo: Atlas. 2004.
- FILZMOSE, P.; GARRETT, R. G.; REIMANN, C. Multivariate Outlier Detection in Exploration Geochemistry. **Computers & Geosciences**.v.31, p.579-587, 2005.
- FUCHS, C.; KENETT, R. S. **Multivariate Quality Control: Theory and Application**. MarcelDekker:New York. 1998.
- GARRET, R. G. The Chi-Square Plot: a Tool for Multivariate Outlier Recognition. **Journal of Geochemical Exploration**.v.32,p.319-341,1989.

GUH, R-S.; SHIUE, Y-R. An Effective Application of Decision Tree Learning for on-line Detections of Mean Shifts in Multivariate Control Charts. **Computer & Industrial Engineering**. v.55,p.475-493, 2008.

LOPES, L. F. D. **Controle Estatístico de Processo** (Apostila). 2007. Disponível em: <http://www.ebah.com.br/cep-control-e-estatistico-de-processo-pdf-a6788.html>. Acesso em 10/04/2009.

LYU, Jr J.; CHEN, M. N. Automated Visual Inspection Expert System for Multivariate Statistical Process Control Chart. **Expert Systems With Applications**, Volume 36, issue 3, Part 1, pages 5113-5118. 2009.

MARANHÃO, P. H. C.; EPPRECHT, E. K. O Uso de Variáveis Transformadas no Controle Estatístico de Processos Multivariados. **Revista eletrônica – Pesquisa Operacional para o Desenvolvimento**, Rio de Janeiro, v.3,n.3, p.201-217. 2011.

MONTGOMERY, D. C. **Statistical Quality Control**.ed.5.Wiley: Danvers. 2005.

MONTGOMERY, D. C. **Introdução ao Controle Estatístico da Qualidade**. Rio de Janeiro: LTC. 2009.

PALADINI, E. P. **Gestão da Qualidade: teoria e prática**. Atlas, São Paulo. 2004.

PAN J.; LEE, C. New Capability Indices for Evaluating the performance of Multivariate Manufacturing Processes. **Quality and Reliability Engineering International**, 26 3-15. 2010.

SAMOHYL, R. W. **Controle Estatístico de Qualidade**, Rio de Janeiro; Elsevier. 2009.

TREASURE, R. J.; KRUGER, U.; COOPER J. E. Dynamic Multivariate Statistical Process Control using Subspace Identification. **Journal of Process Control**, Volume 14, issue 3, pages 279-292. 2004.

APPLICATIONS OF MULTIVARIATE STATISTICAL QUALITY- CONTROL: CONTROLLING PLASTIC CRATES OF BOTTLES IN AN INDUSTRIAL ENTERPRISE

Abstract

This paper aims to use the multivariate statistical techniques to improve the process of monitoring and control the quality in an enterprise able to work in the plastic market. Thus, the use of principal components analysis (PCA) and the Hotelling T square chart can summarize relevant information of this process. This research was realized with interaction of enterprises because it counted on the audit of a worldwide company of drinks. The PCA model can be obtained via R software. Through functions used in this program, one obtains the value of the ratio of the variance explained and the coefficients used in the model. Thus, one can get the new data set with little variables correlated. The Hotelling T^2 control chart can also be built via software R. In conclusion, they were produced two results of considerable importance: the scores of the principal components and an adapted Hotelling T square chart, highlighting the relationship between the ten variables analyzed.

Keywords: *Statistical quality control, Hotelling T square chart, Principal components analysis.*