



Est. Cristina Barbiero.⁽¹⁾
crisbarbiero@hotmail.com
Est. María I. Flury⁽¹⁾
carla@cimero.org.ar
Est. Marta Ruggieri⁽¹⁾
scavone@arnet.com.ar

*(1) Instituto de Investigaciones Teóricas y Aplicadas. Escuela de Estadística.
Proyecto: "Métodos multivariados aplicados a procesos industriales".*

CONTROL ESTADÍSTICO DE PROCESOS MULTIVARIADOS . ESQUEMAS DE MUESTREO PARA GRÁFICOS T^2 DE HOTELLING.

1.-INTRODUCCION

En los últimos años, el control estadístico de procesos multivariados se ha convertido en una herramienta extremadamente útil para el control simultáneo de varias características de calidad relacionadas.

Decidida la implementación de un gráfico de control multivariado para monitorear un proceso, el paso siguiente es diseñar el esquema de muestreo a adoptar y para ello es necesario definir ciertos parámetros entre los cuales están: el tamaño de muestra o subgrupo de observaciones con que se habrá de generar cada punto en el gráfico y la frecuencia o intervalo de muestreo, es decir el tiempo que se dejará transcurrir entre puntos sucesivos.

En el presente trabajo se analizan y comparan distintos criterios para seleccionar un determinado plan de muestreo, evaluando su rapidez para detectar cambios en el proceso, de una magnitud dada. Algunas de las propuestas analizadas corresponden a esquemas de muestreo estándar que utilizan tamaños de muestra e intervalos de muestreo fijos y por lo tanto con costos asociados a cada plan de inspección, constantes por unidad de tiempo. En estos casos se evalúa la incidencia de los parámetros utilizados en el esquema de muestreo, en la capacidad del gráfico de control para detectar desviaciones de los valores objetivo, de distinta magnitud. Otras propuestas consideran un tamaño de muestra fijo e intervalo de muestreo variable. Por último, también se presenta el caso más general de tamaño de muestra variable e intervalo de muestreo variable. En cada situación alternativa se compara el procedimiento de control multivariado a través de la T^2 de Hotelling, con el procedimiento estándar.

Las propuestas presentadas se encuentran aún en etapa de implementación, pero son muy buenas las perspectivas de lograr beneficios con su utilización.



2.- GRAFICO T^2 DE HOTELLING CON TAMAÑO DE MUESTRA E INTERVALO DE MUESTREO FIJOS

Sean p características de calidad correlacionadas que se miden simultáneamente y que cuando el proceso está bajo control, pueden ser modeladas por una distribución de probabilidad normal p -variante, con vector de medias $\bar{\mu}_0 = (\mu_{0.1}, \mu_{0.2}, \dots, \mu_{0.p})'$ y matriz de varianzas y covarianzas Σ_0 de orden $p \times p$.

Se plantea un plan de control estadístico en base a muestras o subgrupos de tamaño n sobre los que se miden las p características, calculando la estadística T^2

$$T_i^2 = n(\bar{X}_i - \bar{\mu}_0)' \Sigma_0^{-1} (\bar{X}_i - \bar{\mu}_0)$$

siendo:

$\bar{X}_i = (\bar{X}_{i1}, \bar{X}_{i2}, \dots, \bar{X}_{ip})$: el vector de medias para las p variables, en la i ésima muestra.

Si $\bar{\mu}_0$ y Σ_0 son conocidos y el proceso está bajo control, la estadística T_i^2 se distribuye como una chi-cuadrado central con p grados de libertad ($T_i^2 \approx \chi_p^2$).

Cuando el proceso no está bajo control, la estadística T_i^2 se distribuye según una chi-cuadrado no central con p grados de libertad y con parámetro de no centralidad: $\lambda = n(\mu_i - \mu_0)' \Sigma_0^{-1} (\mu_i - \mu_0)$, con $\mu_i \neq \mu_0$. Una medida del cambio en el vector de medias puede definirse a partir de la distancia de Mahalanobis: $d = \sqrt{(\mu_i - \mu_0)' \Sigma_0^{-1} (\mu_i - \mu_0)}$, resultando $\lambda = n d^2$.

Esta expresión del parámetro λ de no centralidad indica que es más probable detectar un cambio "d" en el vector de medias, cuando se toma un n más grande, ya que ello conduce a un λ mayor y similarmente, que para un n dado, es más fácil detectar un cambio "d", cuando éste es más grande.

Una medida de la efectividad del gráfico para detectar un cambio, viene dada por la magnitud del **ARL** (average run length). El **ARL** es el *número de puntos* que en promedio deberán graficarse desde que se produce una salida de control, hasta obtener una señal de fuera de control.

Otra medida de efectividad, viene dada por la magnitud del **AvD** (average duration), que es el *tiempo promedio* transcurrido hasta que el gráfico produzca una señal, desde el momento en que ha habido un cambio en el vector de medias.

La relación que liga al **ARL** y el **AvD** es la siguiente

$$\text{AvD} = (\text{ARL} - 1) f + t,$$

donde f : frecuencia de muestreo o tiempo entre dos muestras consecutivas



t : tiempo promedio entre la ocurrencia del cambio y la obtención de la primera muestra, luego de ocurrido el cambio.

Si la probabilidad de un cambio en el proceso, en un intervalo de muestreo particular, es pequeña, puede suponerse que el momento en que se produce la salida de control será en la mitad del intervalo de muestreo, con lo cual $t \cong f / 2$, resultando

$$AvD = (ARL - 0,5) f$$

Aparisi (1998) compara cinco esquemas de muestreo, con un mismo costo de inspección, suponiendo que el costo es proporcional al número de items a inspeccionar. El criterio que utiliza para seleccionar el plan más efectivo es el de buscar aquél que tiene asociado el **menor valor de AvD, para un cambio "d"**.

Los planes propuestos son:

Plan 1: $n = 1$ $f = 1$ unidad de tiempo

Plan 2: $n = 2$ $f = 2$ unidades de tiempo

Plan 3: $n = 3$ $f = 3$ unidades de tiempo

Plan 4: $n = 5$ $f = 5$ unidades de tiempo

Plan 5: $n = 7$ $f = 7$ unidades de tiempo

Para comparar estos esquemas se impone como condición que todos ellos tengan el mismo **AvD**, cuando *el proceso está bajo control*, lo cual determina distintos **ARL** y distintas probabilidades de falsa alarma (α) a definir en cada caso.

Una vez identificadas estas probabilidades se establece el límite superior de control a utilizar en el gráfico T^2 , dado que:

$$\alpha = P(T^2 > LC / d = 0) = P(\chi_p^2 > LC) \text{ y por lo tanto: } LC = \chi_{p,1-\alpha}^2, \text{ donde } \chi_{p,1-\alpha}^2 \text{ es el } (1-\alpha) \text{ percentil.}$$

En el mencionado artículo, se comparan los valores de **AvD** obtenidos para $p = 2,3,4,5$, y 10 variables y para cambios "d" iguales a: 0; 0.5; 1.0; 1.5; 2.0; 2.5 y 3.

Un análisis de los resultados obtenidos, permite establecer criterios para la selección del tamaño de muestra adecuado y la frecuencia de muestreo, los cuales dependen del orden de magnitud en la desviación de los promedios de las variables de calidad que se pretenden detectar (**d**). Los mismos pueden enunciarse como:

- Para detectar cambios pequeños ($d \in [0.5, 1]$), es mejor usar tamaño de muestra grande, en forma más espaciada ($n = 7, f = 7$), independientemente del número de variables considerado.
- Para detectar cambios intermedios ($d \in [1, 2]$), los esquemas con $n = 5, f = 5$ y $n = 7, f = 7$, son los mejores, independientemente del número de variables considerado.
- Para detectar cambios grandes ($d \in [2,3]$), es mejor utilizar tamaño de muestra pequeño, muy frecuentemente. Si se consideran pocas variables, es preferible el plan con $n = 2, f = 2$. Si se consideran muchas variables, es mejor incrementar el tamaño de la muestra.



Como consideraciones adicionales, se establece que para elegir adecuadamente un plan de muestreo también se debe considerar la capacidad del proceso. Se entiende que un proceso es "capaz", cuando además de tener su variabilidad acotada, produce dentro de las especificaciones de calidad para él impuestas. En procesos con buena capacidad, pequeños cambios ejercen muy poca influencia sobre la cantidad de items disconformes y por lo tanto se deberían elegir planes que sean capaces de detectar solamente cambios grandes. Por el contrario, en procesos no capaces, aún cambios muy pequeños, pueden tener influencia en la cantidad de artículos disconformes y por lo tanto, se deberían elegir planes que sean capaces de detectar rápidamente dichos cambios.

En todo caso el tamaño de la muestra y la frecuencia de muestreo, dependerá del número de variables considerado.

Resumiendo, para que un esquema de muestreo sea adecuado, para su definición se deben considerar tres aspectos: la magnitud del cambio que desea detectarse, la capacidad del proceso productivo y el número de variables de calidad consideradas para el control.

3.- GRÁFICO T^2 DE HOTELLING CON TAMAÑO DE MUESTRA FIJO E INTERVALO DE MUESTREO VARIABLE

Como fuera señalado, el procedimiento usual de gráfico de control T^2 requiere tomar muestras de tamaño fijo a intervalos de tiempo fijo. Cuando un valor supera el límite de control, se debe tomar una acción correctiva. Sin embargo diferentes valores de T^2 sugieren distintas perspectivas respecto de la permanencia del proceso en estado de control. Aún cuando la muestra observada produzca un valor de la estadística T^2 inferior al LC, si el mismo está cercano a tal límite, podría tomarse como indicativo de la existencia de un problema. Por el contrario, si T^2 es muy inferior al LC, no hay indicios de alejamientos de los valores estándares de calidad.

Por lo tanto se propone (Haro,C.,2001) establecer una zona de advertencia de modo que si el valor observado de T^2 se encuentra próximo al LC, se vuelva a tomar una muestra luego de un período más corto de tiempo. En síntesis, al diseñar el gráfico, se deberán elegir dos tiempos entre muestras sucesivas distintos y la decisión de usar uno u otro después de cada muestra dependerá del valor observado de la estadística T^2 obtenido con esa muestra.

Gráficamente:

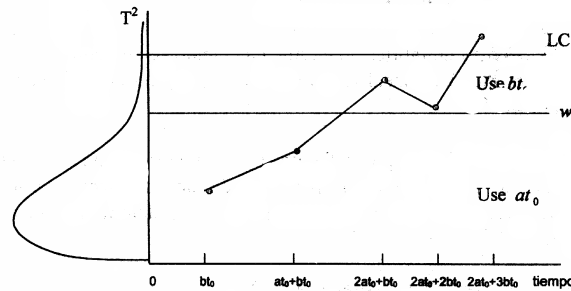


Figura 1.- Gráfico de control T^2 de Hotelling con intervalo de muestreo variable.

El procedimiento que se propone, entonces, es fijar un límite adicional de advertencia (w), y dos variantes de intervalos de muestreo (at_0 y bt_0), de modo que:

Si $T_{i-1}^2 \leq w \Rightarrow$ la siguiente muestra se toma después de at_0 unidades de tiempo

Si $w < T_{i-1}^2 \leq LC \Rightarrow$ la siguiente muestra se toma después de bt_0 unidades de tiempo

Si $T_{i-1}^2 > LC \Rightarrow$ el gráfico señala una situación de fuera de control;

Siendo t_0 el tiempo entre muestras sucesivas en un gráfico T^2 estándar y a y b tales que $0 < b < 1 < a$. Por las restricciones en las constantes a y b ; $at_0 > t_0$ y $bt_0 < t_0$.

Este procedimiento coincidiría con el de un gráfico T^2 estándar si $w = 0$ y $a = b = 1$.

Para que los procedimientos con intervalo de muestreo fijo o variable sean equivalentes cuando el proceso está bajo control, se debe cumplir que, en promedio, los intervalos de muestreo sean iguales en ambos.

El tiempo variable entre muestras se definió como:

$$t(T_i^2) = \begin{cases} bt_0 & \text{si } w < T_{i-1}^2 \leq LC \\ at_0 & \text{si } 0 < T_{i-1}^2 \leq w \end{cases}$$

y por lo tanto, el requerimiento anterior implica que :

$$E[(t(T_i^2) / T_{i-1}^2 < LC; d = 0)] = \\ = a t_0 P(T_{i-1}^2 < w / T_{i-1}^2 < LC) + b t_0 P(w < T_{i-1}^2 < LC / T_{i-1}^2 < LC) = t_0$$

Los tres parámetros del diseño resultan: w , a , y b que están relacionados a partir de la ecuación anterior. Operando se obtiene:



$$w = F^{-1}\left[\left(\frac{1-b}{a-b}\right)F(LC,p); p\right] \quad 0 < b < 1 < a,$$

donde: $F(x,p) = P(\chi_p^2 \leq x)$, es la función de distribución Chi- cuadrado central con p grados de libertad.

Por lo tanto, dados dos intervalos entre muestras: at_o y bt_o , se puede obtener w y producir un gráfico de control T^2 que tenga un tiempo promedio entre muestras igual al gráfico T^2 estándar con el mismo **ARL**, cuando el proceso está centrado en el valor objetivo.

La efectividad de un plan de control queda determinada por la rapidez con que el mismo detecta una condición de fuera de control, por lo que para comparar este nuevo procedimiento con el estándar, se considera como medida apropiada, el **tiempo medio hasta la señal de falta de control** (Average Time to Signal: **ATS(d)**).

El **ATS** es función de la magnitud de salida de control, la cual puede evaluarse mediante la distancia de Mahalanobis " d " definida en la sección anterior.

Se supone que el cambio producido en los promedios de las variables de calidad consideradas se mantiene constante desde el tiempo cero en adelante.

Para calcular el **ATS** se puede considerar al gráfico de control como una cadena de Markov con tres estados:

Estado 1: $T_{i-1}^2 \in [0, w]$

Estado 2: $T_{i-1}^2 \in (w, LC]$

Estado 3: $T_{i-1}^2 \in (LC, \infty)$

El estado 3 es un estado absorbente puesto que indica que el proceso está fuera de control. La matriz de probabilidades de transición, habiendo ocurrido en el proceso un cambio de magnitud d , es:

$$P_d = \begin{bmatrix} P_{11}^d & P_{12}^d & P_{13}^d \\ P_{21}^d & P_{22}^d & P_{23}^d \\ P_{31}^d & P_{32}^d & P_{33}^d \end{bmatrix},$$

donde P_{ik}^d , con $i, k = 1, 2, 3$, denota la probabilidad de transición del estado anterior i al actual k , cuando el vector de medias sufrió un cambio " d ".



Por ejemplo, la probabilidad P_{12}^d es la probabilidad de que estando en un período cualquiera, en el Estado 1: ($0 < T_{i-1}^2 < w$), pase en el período siguiente al Estado 2:

($w < T_i^2 < LC$), habiendo ocurrido una salida de control que ocasionó un corrimiento " d " en el vector de medias.

Las muestras entre tiempos consecutivos son independientes y por lo tanto el valor de T^2 en un período no influye sobre el valor de T^2 en el próximo, sino solamente en el intervalo de tiempo transcurrido entre las muestras, que en este caso habrá sido at_0 , el más largo.

Así la probabilidad condicional del estado E_2 dado el estado E_1 es igual a la probabilidad del estado E_2 .

Por lo tanto:

$$P_{12}^d = P^d(w < T_i^2 < LC / 0 < T_{i-1}^2 \leq w) = P^d(w < T_i^2 \leq LC) = P(\chi_{p(\lambda)}^2 < LC) - P(\chi_{p(\lambda)}^2 < w)$$

donde $\lambda = nd^2$ (parámetro de no centralidad).

De modo similar se obtendrían las otras combinaciones de probabilidades de transición entre estados ,que resultan iguales a distintas probabilidades sobre la distribución de T^2 cuando el proceso se salió de control.

Page (1954) indica que el valor ATS se calcula mediante la expresión:

$$ATS(d) = b'(I - Q_d)^{-1}t \quad (1)$$

donde:

b' es el vector de probabilidades iniciales (p_1, p_2), tal que $p_1 + p_2 = 1$; siendo p_1 y p_2 las proporciones de veces que el proceso permanece en los estados 1 y 2, usando los tiempos de espera at_0 y bt_0 , r, antes que el gráfico dé una señal de fuera de de control,

I es la matriz identidad de orden 2,

Q_d es la matriz de probabilidades de transición donde se han eliminado los elementos asociados con el estado absorbente y

t' es el vector de intervalos de tiempo de muestreo (at_0, bt_0).

La ecuación (1) puede escribirse en función de p, d, w, b y a como:

$$ATS(p, d, w, b, a) =$$

$$\left[\begin{array}{c} \frac{P(\chi_p^2 < w)}{P(\chi_p^2 < LC)} \\ 1 - \frac{P(\chi_p^2 < w)}{P(\chi_p^2 < LC)} \end{array} \right] \left(\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] - \left[\begin{array}{cc} P(\chi_{p(\lambda)}^2 < w) & P(w < \chi_{p(\lambda)}^2 < LC) \\ P(\chi_{p(\lambda)}^2 < w) & P(w < \chi_{p(\lambda)}^2 < LC) \end{array} \right] \right)^{-1} \left[\begin{array}{c} at_0 \\ bt_0 \end{array} \right] \quad (2)$$

Dado que la efectividad de un proceso de control será mayor cuanto menor sea el ATS para un particular d , se establece como criterio para fijar los parámetros del diseño n y b , la condición de hacer mínima esta expresión para valores fijos de p y d .



Esta búsqueda de extremos no puede hacerse por medios analíticos y se requieren procedimientos computacionales intensivos.

Haro C.(2001) obtiene los valores de a , b y w necesarios para una variedad de valores de n , p , d y α , comparando luego el valor ATS obtenido para el procedimiento con intervalo de muestreo variable, con el obtenido para el gráfico T^2 estándar. Se proporcionan tablas con valores de ATS, indicando los *planes de control óptimos* que brindan el *mínimo* valor posible de ATS para un descentrado (d) determinado.

Entre las recomendaciones señaladas por el autor, se menciona que el gráfico de control T^2 de Hotelling con intervalos de muestreo variable es más eficiente para detectar cambios en el vector de medias dado que origina un ATS menor.

Además, la eficiencia es mayor para descentrados pequeños, especialmente cuando el tamaño de muestra es grande y si el descentrado es moderado ($d = 1.0$ y $d = 1.5$) se logran valores de ATS hasta dos veces menores. Con descentrados grandes, la mejora obtenida respecto al procedimiento estándar es pequeña.

Estas mejoras en eficiencia se detectan al observar el tiempo medio hasta la señal. Sin embargo es interesante señalar que ambos procedimientos producen el mismo número medio de muestras a tomar $E[N]$ hasta que el gráfico dé una señal de fuera de control, para cualquier descentrado d .

4.- GRÁFICO T^2 DE HOTELLING CON TAMAÑO DE MUESTRA VARIABLE E INTERVALO DE MUESTREO VARIABLE

En este punto se considera la posibilidad de fijar un plan de muestreo aún más general, permitiendo variar tanto el intervalo de muestreo como el tamaño de la muestra, dependiendo de la posición que ocupa el valor observado T^2_{i-1} en el gráfico de control. El objetivo es diseñar un gráfico todavía más potente, que permita reducir el tiempo para detectar un apartamiento en el vector de medias.

Haro, C. (2001) analiza esta propuesta considerando en el diseño dos combinaciones de tamaños de muestra e intervalos de tiempo entre muestras.

El gráfico propuesto es:

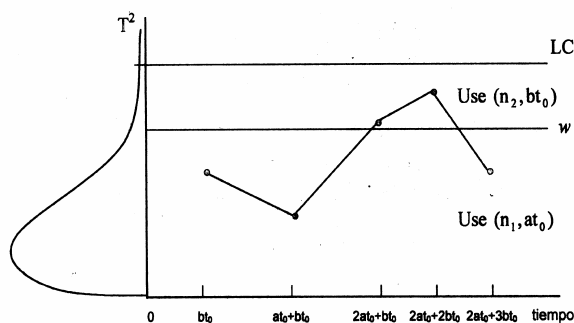




Gráfico de control T^2 de Hotelling con tamaño de muestra variable e intervalo de muestreo variable

Siendo n_0 : tamaño de muestra fijo en el gráfico T^2 estándar
 t_0 : intervalo de muestreo fijo en el gráfico T^2 estándar
 n_1 y n_2 : dos tamaños de muestra posibles, tales que $n_1 < n_0 < n_2$
 at_0 y bt_0 : dos intervalos entre muestras posibles, tales que $a > 1$ y $0 < b < 1$
 w : límite de aviso, que determina cuándo cambiar el tamaño de la muestra y la frecuencia de muestreo
LC : límite de control

Cuando:

$w=0$	$n_1=n_2=n_0$	$a=b=1$	\Rightarrow gráfico T^2 estándar
$w \neq 0$	$n_1 < n_0 < n_2$	$a=b=1$	\Rightarrow gráfico T^2 con tamaño de muestra variable
$w \neq 0$	$n_1 < n_0 < n_2$	$a > 1, 0 < b < 1$	\Rightarrow gráfico T^2 con tamaño de muestra variable e intervalo de muestreo variable

El procedimiento de control que se propone es evaluar cada muestra y decidir según el siguiente criterio:

Si $T_{i-1}^2 \leq w$ \Rightarrow la siguiente muestra será más pequeña, de tamaño n_1 y se toma después de at_0 unidades de tiempo (intervalo entre muestras más largo)

Si $w < T_{i-1}^2 \leq LC \Rightarrow$ la siguiente muestra es grande, de tamaño n_2 y se toma después de bt_0 unidades de tiempo (intervalo entre muestras más pequeño)

Si $T_{i-1}^2 > LC$ \Rightarrow el gráfico señala una situación de fuera de control.

Los parámetros de diseño de este gráfico de control son cinco: n_1 , n_2 , a , b y w y para fijarlos es necesario tener en cuenta el ATS que se pretende. Además, a los efectos de comparación con el gráfico T^2 estándar, se requiere que el tamaño de muestra promedio y el intervalo de muestreo promedio sean iguales, cuando el proceso está bajo control, es decir que debe verificarse:

$$E(n(T_i^2) / T_{i-1}^2 < LC; d=0) = n_0$$
$$E(t(T_i^2) / T_{i-1}^2 < LC; d=0) = t_0$$

Nuevamente, no es un procedimiento sencillo seleccionar estos parámetros, de modo que se cumplan las restricciones impuestas y que consigan el mínimo valor de ATS posible para detectar una desviación d en las características de calidad. Haro, C. (2001) considera



algunas variantes de p y α y compara la potencia de esta alternativa con la estándar y con la propuesta intermedia presentada en la sección anterior, donde sólo se puede variar el intervalo de muestreo. En cada caso especifica los cinco parámetros que permiten obtener la misma performance que el gráfico T^2 estándar en una situación bajo control y que a la vez minimizan el ATS en una situación fuera de control.

A partir de las comparaciones realizadas concluye que, cuando el descentrado es pequeño, se logran importantes reducciones en los valores de ATS cuando se utiliza el gráfico T^2 combinado. Los ATS obtenidos son entre tres y cuatro veces más pequeños, que los que se obtienen cuando se emplea el método estándar. También se logra una reducción muy importante en el número medio de muestras a tomar, hasta detectar el cambio en el vector de medias.

Sin embargo, cuando el descentrado es grande, hay muy poca diferencia en los valores de ATS de ambos procedimientos. Esto no es relevante, porque en ambos casos el valor de ATS es pequeño. Tampoco son importantes las diferencias logradas en el número medio de muestras a tomar.

5.- CONSIDERACIONES FINALES

En todo proceso industrial moderno, la evaluación de calidad on line es un importante aspecto a considerar, no sólo por la obtención de un mejor producto, sino por los réditos económicos a que esto conlleva. Frente a procesos complejos posiblemente sea necesario implementar procedimientos de control más complejos que tengan en cuenta varias características de calidad simultáneamente, como lo es el basado en la estadística T^2 . En cualquier caso, es importante considerar aspectos de diseño que permitan trabajar con una alta probabilidad de detectar una salida de control en forma relativamente rápida.

En este trabajo se han presentado planes de muestreo alternativos propuestos por distintos autores en estos últimos años con el fin de mejorar la eficiencia del gráfico T^2 , especialmente frente a pequeñas desviaciones de los estándares de calidad.

A través de los planes comentados se logra una reducción muy importante en el tiempo promedio transcurrido desde la salida de control hasta su detección.

Si bien estas mejoras en ocasiones requieren de un tamaño de muestra superior para detectar la salida de control cuando ésta se ha producido, garantiza el mantener el promedio de unidades a inspeccionar, mientras el proceso está bajo control.

Aunque las propuestas presentadas son muy nuevas y se encuentran aún en etapas de desarrollo, aparecen con muy buena perspectiva de lograr beneficios con su implementación.

6.- BIBLIOGRAFIA

- Aparisi, Francisco: "Sampling Plans for the multivariate T^2 control chart". Quality Engineering, vol.10 N° 1 (1997-1998).



- Flury, María I.; Barbiero, Cristina: "Aplicación de técnicas de control multivariado en procesos industriales". Sextas Jornadas de Investigación de la FCEE.(2001).
- Haro, César: "Selección de planes de muestreo para los gráficos T^2 de Hotelling y de varianza generalizada". Tesis doctoral. (2001).
- Fuchs Camil ; Kenett Ron: "Multivariate Quality Control". Rocher Institute of Technology. (1998).
- Page E. S. : "Continuous inspection schemes".Biometrika 41, p. 100-115. (1954).