



**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**

**Facultad de Ciencias Naturales y Matemáticas**

**“Desarrollo de un paquete en R para Gráficos de Control  
Multivariante”**

**INFORME DE PROYECTO INTEGRADOR**

**Previo a la obtención del Título de:**

**INGENIERO EN ESTADÍSTICA INFORMÁTICA**

**Presentado por:**

**JHONNY FABRICIO BARAHONA ALVARADO**

**GUAYAQUIL – ECUADOR**

**2016**



## AGRADECIMIENTOS

*Un grande agradecimiento a Dios, por permitirme culminar está etapa de mi vida educacional, a mi padre Raúl Barahona Orozco por brindarme su apoyo incondicional para seguir adelante, a mis hermanos y esposa por brindarme su ayuda en aquellos momentos de cansancio.*

*A mis estimados profesores PhD. Omar Ruiz Barzola y M.Sc Sonnia Reyes Ramos por su gran aporte para la elaboración de este trabajo.*

## DEDICATORIA

*El presente proyecto que con sacrificio y dedicación culmino con éxito va dedicado  
al ser que guía mis pasos, mi Dios.*

*A mi padre Raúl, mi esposa Ruth y mi hijo Fabricio Jonael.*



ESCUELA SUPERIOR POLITECNICA DEL LITORAL

Facultad de Ciencias Naturales y Matemáticas

Ingeniería en Estadística Informática

INFORME DEL PROFESOR DE LA MATERIA INTEGRADORA

Habiendo sido nombrada PROFESOR DE LA MATERIA INTEGRADORA del señor

JHONNY FABRICIO BARAHONA ALVARADO

Con el tema del proyecto integrador “DESARROLLO DE UN PAQUETE EN R PARA GRÁFICOS DE CONTROL MULTIVARIANTE”, previa a la obtención del título de INGENIERO EN ESTADÍSTICA INFORMÁTICA, me permito informar que he leído el contenido del proyecto integrador; luego de lo cual indico que estoy de acuerdo en que el mismo se lo ha desarrollado conforme a los lineamientos de la Unidad de Titulación Especial de la ESPOL.

M.Sc. Sonia Paola Reyes Ramos

PROFESOR DE LA MATERIA INTEGRADORA

Guayaquil, 13 de septiembre del 2016

## DECLARACIÓN EXPRESA

"La responsabilidad del contenido de este Trabajo de Titulación, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la Escuela Superior Politécnica del Litoral"

---

Jhonny Fabricio Barahona Alvarado

## RESUMEN

Los gráficos de control multivariante son una valiosa aportación en el campo del control de procesos, el gráfico de control  $T^2$  de Hotelling es el más utilizado en este campo por su fácil aplicación, aunque no es el más óptimo al momento de dar señales de fuera de control. Para mejorar esa debilidad del gráfico  $T^2$  de Hotelling, se han realizado variaciones de él, dos de ellas son los gráficos de doble dimensión  $DDT^2$  y dimensión variable  $VDT^2$ . Sin embargo, estas variaciones son más complejas al momento de aplicarlas en producción.

Por el motivo antes expuesto se desarrolló un paquete en R que realice los gráficos de control multivariantes y calcule sus respectivos ARL, utilizando los gráficos  $T^2$  de Hotelling, de doble dimensión  $DDT^2$  y dimensión variable  $VDT^2$  estos dos últimos son variaciones realizadas al tradicional gráfico  $T^2$  que permite obtener un mejor rendimiento y a la vez reducir el costo del muestreo evitando el uso innecesario de aquellas variables complejas de medir y/o caras.

Para la validación de los resultados se realiza la simulación de  $p=p_1+p_2$  variables normales multivariante, de  $p_1$  variables fáciles de medir y/o baratas y  $p_2$  variables difíciles de medir y/o costosa.

Se obtuvo como resultado el paquete en R, el cual consta de varias funciones que realizan los tres gráficos de control, con una interfaz sencilla para usuarios noveles de R. Además se confirmó que el gráfico de dimensión variable en todos los escenarios planteados en este proyecto supera al rendimiento del grafico de control  $T^2$  tradicional con las  $p$  variables, el grafico de dimensión supera en todos los escenarios planteados el rendimiento del gráfico  $T^2$  con las  $p_1$  variables y en algunas ocasiones supera el tradicional  $T^2$  con las  $p$  variables. Tanto el gráfico de control  $VDT^2$  y  $DDT^2$  reducen el costo del muestreo ya que utilizan las variables costosas no siempre sino solo cuando es necesario.

## ÍNDICE GENERAL

AGRADECIMIENTOS.....	iii
DEDICATORIA.....	iv
DECLARACIÓN EXPRESA.....	vi
RESUMEN.....	vii
ÍNDICE GENERAL.....	viii
ÍNDICE DE TABLAS.....	x
ÍNDICE DE GRÁFICOS.....	xi
ÍNDICE DE ILUSTRACIONES.....	xii
ABREVIATURAS.....	xiii
SIMBOLOGIAS.....	xiv
CAPÍTULO 1.....	1
1. INTRODUCCIÓN.....	1
1.1 Planteamiento del Problema.....	1
1.2 Justificación.....	1
1.3 Objetivos.....	2
1.3.1 Objetivo General.....	2
1.3.2 Objetivos Específicos.....	2
CAPÍTULO 2.....	3
2. METODOLOGÍA.....	3
2.1 Contraste de Hipótesis.....	3
2.2 Distancias.....	4
2.2.1 Distancia Euclidiana.....	4
2.2.2 Distancia de Mahalanobis.....	5
2.3 Distribución Probabilística Chi-cuadrado no Centralizada.....	6
2.4 Gráficos de Control Multivariante basados en el gráfico de control T2 de Hotelling.....	7
2.4.1 Gráfico de Control T2 de Hotelling.....	7
2.4.2 Gráfico de Control T2 con Dimensión Variable.....	8
2.4.3 Gráfico de Control T2 con Doble Dimensión.....	13

2.5	Simulación de variables aleatorias con Distribución Normal Multivariante	17
2.5.1	Distribución Normal Multivariante $N(\mu, \Sigma)$	17
2.5.2	Método basado en la triangulación de Cholesky	17
2.6	Desarrollo de paquetes en R	18
2.6.1	Estructura de paquetes	18
2.6.2	DESCRIPTION	19
2.6.3	NAMESPACE	19
2.6.4	Documentación (Carpeta man)	20
2.6.5	Código (Carpeta R)	20
2.6.6	Comprobar, construir y Publicar	21
CAPÍTULO 3		22
3.	RESULTADOS	22
3.1.	Desarrollo y publicación del Paquete Gráficos de control multivariante	23
3.2.	Gráficos de control T2 de Hotelling	26
3.3.	Gráficos de control T2 con dimensión variable	31
3.4.	Gráficos de control T2 con doble dimensión	34
3.5.	Análisis del rendimiento de los gráficos de control	37
CONCLUSIONES Y RECOMENDACIONES		39
BIBLIOGRAFÍA		40
ANEXOS		42
a.	Tabla de datos simulados	42
b.	Código fuente del grafico de control T2 de Hotelling	44
c.	Código fuente del grafico de control de doble dimensión DDT2	46
d.	Código fuente del grafico de control de dimensión variable VDT2	51
e.	DESCRIPTION del paquete Gcmulti	56

## ÍNDICE DE TABLAS

Tabla 2.1 Tipos de Errores.....	4
Tabla 3.1 Análisis de rendimiento de diferentes tamaños de muestra.....	37
Tabla 3.2 Análisis de rendimiento de varios casos con tamaño de muestra uno.....	38

## ÍNDICE DE GRÁFICOS

Gráfico 2.1 Función de Densidad Chi Cuadrado no Centralizada.....	6
Gráfico 2.2 Gráfico de Control T2 de Hotelling.....	8
Gráfico 2.3 Gráfico de Control con Dimensión Variable.....	12
Gráfico 2.4 Gráfico de control con doble dimensión.....	17
Gráfico 3.1 Gráfico de control T2 con p1 variables y n=1.....	27
Gráfico 3.2 Gráfico de control T2 con p1 variables y n=5.....	28
Gráfico 3.3 Gráfico de control T2 con p variables y n=1.....	29
Gráfico 3.4 Gráfico de control T2 con p variables y n=5.....	30
Gráfico 3.5 Gráfico de control T2 con dimensión variable con n=1.....	32
Gráfico 3.6 Gráfico de control T2 con dimensión variable con n=5.....	33
Gráfico 3.7 Gráfico de control T2 con doble dimensión con n=1.....	35
Gráfico 3.8 Gráfico de control T2 con doble dimensión con n=5.....	36

## ÍNDICE DE ILUSTRACIONES

Ilustración 2.1 Estructura de un paquete de R.....	18
Ilustración 2.2 DESCRIPTION de un paquete de R.....	19
Ilustración 2.3 Estructura de la carpeta man.....	20
Ilustración 3.1 Paquete Gcmulti Gráficos de Control Multivariante.....	23
Ilustración 3.2 Documentación gráfico T2 de Hotelling.....	23
Ilustración 3.3 Documentación gráfico de doble dimensión DDT2.....	24
Ilustración 3.4 Documentación gráfico de dimensión variable VDT2.....	24
Ilustración 3.5 Paquete desarrollado en GitHub.....	25
Ilustración 3.6 Ejemplo de instalación del paquete Gcmulti.....	25

## ABREVIATURAS

ARL	Promedio de longitud de rachas
$ARL_0$	Promedio de longitud de rachas cuando el proceso está bajo control
$ARL_d$	Promedio de longitud de rachas cuando el proceso está fuera de control
$DDT^2$	Gráfico $T^2$ con doble dimensión
$VDT^2$	Gráfico $T^2$ con dimensión variable

## SIMBOLOGIAS

$\mu$	Vector de medias de las variables
$\mu_0$	Vector de medias de las variables cuando el proceso está bajo control
$w$	Límite de advertencia
$CLP_1$	Límite de control de las $p_1$ variables
$CLP$	Límite de control de las $p$ variables
$T^2$	Estadístico $T^2$ de Hotelling
$T_{p_1}^2$	Estadístico $T^2$ para las $p_1$ variables
$T_p^2$	Estadístico $T^2$ para las $p$ variables
$LCS$	El límite de control superior del gráfico $T^2$ de Hotelling
$\lambda$	Parámetro de no centralidad de la distribución Chi Cuadrada no centralizada
$\alpha$	Error tipo I
$\beta$	Error tipo II
$d$	Distancia de Mahalanobis
$dis_{p_1}$	Distancia de Mahalanobis para las $p_1$ variables
$dis_p$	Distancia de Mahalanobis para las $p$ variables

## CAPÍTULO 1

### 1. INTRODUCCIÓN

#### 1.1 Planteamiento del Problema

La industria en general maneja varios procesos mediante los cuales realizan sus productos o servicios, en estos procesos se encuentran involucradas varias características de calidad que pueden ser monitoreadas por medio de los gráficos de control univariantes y multivariantes, y así cuando los procesos se encuentren en “estado fuera de control estadístico” tomar las acciones pertinentes de forma instantánea. Lo óptimo es hacer un gráfico de control multivariante y no p gráficos univariantes.

Aunque los gráficos multivariantes son una mejor opción para el control de procesos, en muchas ocasiones los costos asociados al muestreo son muy elevado, por ello se busca realizar gráficos de control multivariante que utilicen las variables que reflejen altos costos sólo cuando sea necesario, esto ayudaría a reducir los costos de muestreo.

Existe una variedad de gráficos de control y varios programas estadísticos donde se encuentran implementadas funciones para realizar gráficos de control pero de manera univariante. Lamentablemente no existen programas estadísticos que realicen gráficos de control multivariantes, lo cual facilitaría su aplicación y el control en los procesos.

#### 1.2 Justificación

Considerando que no existen programas estadísticos libres que ofrezcan la posibilidad de realizar gráficos de control multivariantes, es necesario el desarrollo de una aplicación que realice estos gráficos en un software estadístico de costo cero para el usuario final y con la interfaz más amigable posible.

Este programa estadístico sería de mucha utilidad para toda empresa que desee mejorar el control en sus procesos y reducir su variación.

## 1.3 Objetivos

### 1.3.1 Objetivo General

Desarrollar un paquete en R que realice gráficos de control multivariantes para mejorar el control en los procesos de las empresas industriales, comerciales y de servicio financiero.

### 1.3.2 Objetivos Específicos

- Conocer el estado del arte en el tema de los Gráficos de Control a través de fuentes bibliográficas fiables para poder realizar las funciones en el software estadístico.
- Seleccionar los Gráficos de Control Multivariante más idóneos tomándose como referencia la tesis doctoral “GRÁFICOS DE CONTROL DE CALIDAD MULTIVARIANTE CON DIMENSIÓN VARIABLE” para desarrollar el paquete.
- Realizar la programación de las funciones en R a través de programación modular para realizar los gráficos de control multivariante seleccionados.
- Probar las funciones con datos simulados para validar las funciones a través del Método basado en la triangulación de Cholesky.
- Desarrollar el paquete en el software estadístico para poner a disposición de los usuarios de R la aplicación desarrollada siguiendo los pasos descritos en la web CRAN.

## CAPÍTULO 2

### 2. METODOLOGÍA

#### 2.1 Contraste de Hipótesis

Un contraste de hipótesis es un método o procedimiento para decidir si una hipótesis se debe rechazar o no.

Una hipótesis estadística es una suposición que se plantea respecto a los parámetros o la distribución de probabilidad de una población.

En los contrastes de hipótesis se distinguen dos hipótesis estadísticas:

- La hipótesis nula designada por  $H_0$ , que es la que se establece en principio con el único propósito de rechazarla o no rechazarla.
- La hipótesis alterna  $H_a$ , que es complementaria de la primera. Cuando se habla de contrastar una hipótesis nula contra una alterna, esto siempre se hace suponiendo que la nula es verdadera.

En general la forma de las hipótesis nula y alterna es:

- $H_0$ : Parámetro poblacional = ( $<$ ,  $>$ ) valor supuesto
- $H_a$ : Parámetro poblacional = ( $<$ ,  $>$ ) valor supuesto

#### Posibles errores en el contraste de hipótesis

El contraste de hipótesis no establece la verdad de la hipótesis sino un criterio que nos permite decidir si una hipótesis se rechaza o no, o el determinar si las muestras observadas difieren significativamente de los resultados esperados. En este proceso podemos incurrir en dos tipos de errores según sea la situación real y la decisión que tomemos.

Si rechazamos una hipótesis cuando no debiera hacerlo, cometemos un error de tipo I, mientras que si no la rechazamos debiendo ser rechazada diremos que hemos cometido un error de tipo II.

En la siguiente tabla se resumen los 2 tipos de errores:

Decisiones		
Eventos	No rechazar $H_0$	Rechazar $H_0$
$H_0$ es verdadera	Decisión Correcta	ERROR TIPO I $P(\text{Error I}) = \alpha$
$H_0$ es falsa	ERROR TIPO II $P(\text{Error II}) = \beta$	Decisión Correcta

**Tabla 2.1 Tipos de Errores**  
Fuente: El Autor

## 2.2 Distancias

El concepto de distancia es uno de los más importantes sobre el cual se han elaborado muchos conceptos matemáticos como la convergencia y los espacios métricos.

Para un investigador puede ser importante determinar si dos objetos, con determinadas características, se deben considerar cercanos o no. El interés puede consistir en la ubicación de los individuos en alguna de las varias poblaciones con base en su proximidad a ellas.

### 2.2.1 Distancia Euclidiana

Está definida de la siguiente manera: dados 2 puntos de  $\mathbb{R}^p$ ,  $X_h = (X_{h1}, \dots, X_{hp})$  y  $X_i = (X_{i1}, \dots, X_{ip})$ , se define su distancia euclidiana como:

$$d_{hi} = \sqrt{\sum_{j=1}^p (X_{hj} - X_{ij})^2} = (X_h - X_i)'(X_h - X_i) \quad (2.1)$$

La métrica euclidiana pese a ser la más simple de las distancias a determinar y calcular por su relación con otros resultados bien conocidos de las matemáticas; presenta evidentes inconvenientes de aplicación fuera de espacios donde la línea recta sea la menor distancia que conecta a dos puntos.

### 2.2.2 Distancia de Mahalanobis

Cuando en un estudio nos encontramos con variables que suelen estar en escalas de medición diferente y además son correlacionadas, resulta conveniente utilizar la distancia de Mahalanobis ya que esta considera tanto el efecto de las unidades de medición como la asociación lineal entre las variables.

La distancia de Mahalanobis entre los objetos  $X_h = (X_{h1}, \dots, X_{hp})$  y  $X_i = (X_{i1}, \dots, X_{ip})$ , se define como:

$$D_{hi}^2 = (X_h - X_i)' S^{-1} (X_h - X_i), \text{ con } h, i = 1, \dots, n \quad (2.2)$$

La distancia de Mahalanobis es usada frecuentemente para medir la distancia entre una observación multivariada (individuo) y el centro de la población de donde procede la observación.

Si  $x_i = (x_{i1}, \dots, x_{ip})'$  representa un individuo particular seleccionado aleatoriamente de una población con centro  $\mu = (\mu_1, \dots, \mu_p)'$ , y matriz de covarianzas  $\Sigma$ , entonces:

$$D_i^2 = (x_i - \mu)' \Sigma^{-1} (x_i - \mu) \quad (2.3)$$

Se considera como una medida de la distancia entre el individuo  $x_i$  y el centroide  $\mu$  de la población.

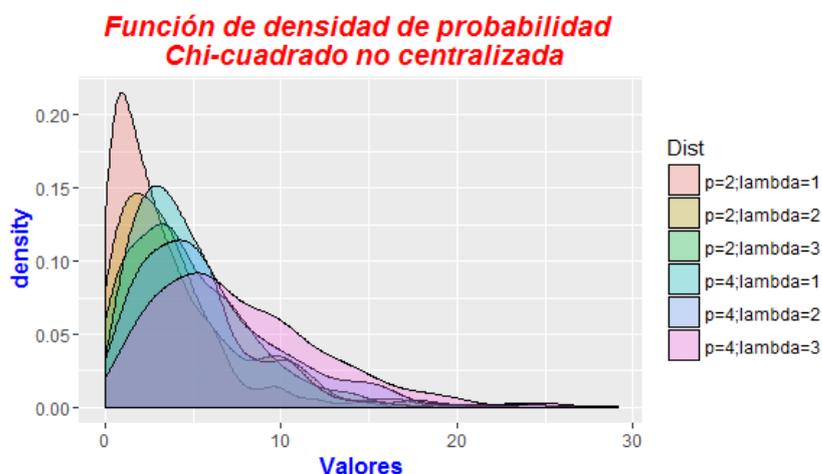
### 2.3 Distribución Probabilística Chi-cuadrado no Centralizada

Sean  $X_1, X_2, \dots, X_p$ ,  $p$  variables aleatorias independientes, normalmente distribuidas con media  $\mu_i$  y la unidad de varianzas. Entonces la variable aleatoria  $\sum_{i=1}^p X_i^2$  se distribuye de acuerdo a la distribución Chi-cuadrado no centralizada. Tiene dos parámetros:

$p$  que indica el número de grados de libertad (es decir, el número de  $X_i$ ).

$\lambda$  que esta relacionada con la media de las variables aleatorias  $X_i$  por:  $\lambda = \sum_{i=1}^p \mu_i^2$ ;  $\lambda$  es llamado parámetro de no centralidad. Su densidad es:

$$f(x; p; \lambda) = \frac{e^{-\left(\frac{x-\lambda}{2}\right)} x^{\left(\frac{p}{2}-1\right)}}{2^{p/2}} \sum_{j=0}^{\infty} \frac{(x\lambda)^j}{4^j j! \Gamma\left(j + \frac{p}{2}\right)} \quad (2.4)$$



**Gráfico 2.1 Función de Densidad Chi Cuadrado no Centralizada**  
Fuente: El Autor

Con  $p$  grados de libertad y  $\lambda$  como parámetro de no centralidad.

## 2.4 Gráficos de Control Multivariante basados en el gráfico de control $T^2$ de Hotelling

### 2.4.1 Gráfico de Control $T^2$ de Hotelling

(Fuchs y Kenett, 1998). La principal aportación en el campo de los gráficos de control multivariantes fue la realizada por el Profesor Harold Hotelling, quien en 1947 basándose en la distancia de Mahalanobis, utilizó el vector de promedios y la matriz de covarianzas de una variable aleatoria con distribución normal multivariada para proponer un gráfico de control denominado gráfico de control  $T^2$  de Hotelling, el cual monitoriza la distancia entre el vector de promedios esperado y el vector de promedios observados teniendo en cuenta su matriz de covarianzas y un límite de control (CL) este último indica si la distancia entre los dos vectores es lo suficientemente grande para declarar al proceso como fuera de control.

#### Obtención del ARL del gráfico de control $T^2$

Sea  $X_i \sim N_p(\mu, \Sigma)$ , procedente de una población con distribución normal multivariada, con parámetros  $\mu$  y  $\Sigma$ , de la cual se extrae la muestra  $i$ , obteniéndose de ellas  $p$  características de calidad correlacionadas.

El estadístico utilizado es el  $T^2$  de Hotelling tiene la forma

$$T_i^2 = n(\bar{x}_i - \mu_0)' \Sigma^{-1} (\bar{x}_i - \mu_0) \quad (2.5)$$

Donde,  $n$  es el tamaño de la muestra,  $\bar{x}_i$  es el vector de promedios de las variables y  $\mu_0$  es el vector de medias de las variables cuando el proceso está bajo control.

Si  $\mu_1 \neq \mu_0$ ,  $T_i^2$  tiene función de distribución de probabilidad  $\chi^2$  no centralizada y parámetro de no centralidad  $\lambda = nd^2$ ; siendo  $d$  (la distancia de Mahalanobis) el cambio o desplazamiento ocurrido durante el proceso.

Para un estado bajo control  $\mu_1 = \mu_0$  se tiene;

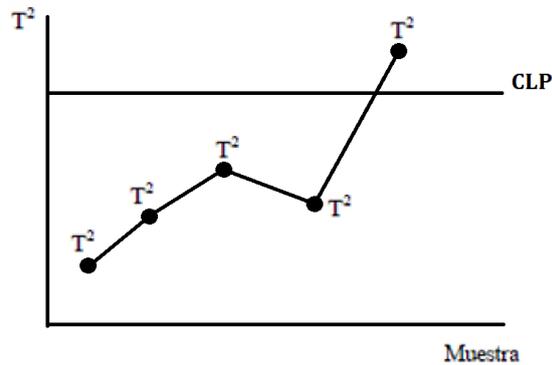
$$ARL_0 = \frac{1}{\int_{UCL}^{\infty} f(x)dx} = \frac{1}{\alpha} \quad (2.6)$$

$f(x)$  se distribuye como una  $\chi_{gl=p}^2$ .

Cuando  $\mu_1 \neq \mu_0$  ha ocurrido un desplazamiento  $d$ , entonces:

$$\begin{aligned} ARL_d &= \frac{1}{1 - \beta} = \frac{1}{1 - P\left(T_i^2 < \frac{UCL}{d}\right)} \\ &= \frac{1}{1 - F_x\left(\frac{UCL}{d}\right)} = \frac{1}{\int_{UCL}^{\infty} f(x; p, \lambda = pd^2)dx} \end{aligned} \quad (2.7)$$

### Gráfico de control $T^2$ de Hotelling



**Gráfico 2.2 Gráfico de Control  $T^2$  de Hotelling**  
Fuente: El Autor

#### 2.4.2 Gráfico de Control $T^2$ con Dimensión Variable

El gráfico de control  $T^2$  con dimensión variable es considerado cuando el número de variables aleatorias a ser monitorizadas es variable. Por este motivo, se lo denominó gráfico de control con dimensión variable ( $VDT^2$ ). La dimensionalidad del gráfico depende básicamente de la complejidad o economía

que representa medir un grupo adicional de  $p_2$  variables de un proceso en el cual se deben controlar  $p = p_1 + p_2$  variables.

- El grupo de  $p_1$  variables corresponde a las menos costosas o menos complejas de monitorizar.
- El grupo restante de  $p_2$  variables son más complejas o más costosas de tratar, y se utilizarán en el momento en que las  $p_1$  variables iniciales indiquen un riesgo de que el proceso pueda salirse de control.

### **Generalidades del gráfico de control $VDT^2$ :**

El procedimiento estándar para la construcción del gráfico de control  $T^2$  implica tomar de forma periódica, una muestra de tamaño  $n$ . Posteriormente se realiza el cálculo del vector de promedios muestrales  $\bar{X}$  de las  $p$  variables a ser monitorizadas, a continuación se calcula el estadístico  $T^2 = n(\bar{X} - \mu_0)' \Sigma^{-1} (\bar{X} - \mu_0)$ .

Donde  $\mu_0 = (\mu_{01}, \mu_{02}, \dots, \mu_{0p})$  es el vector de promedios y  $\Sigma$  es la matriz de covarianzas del vector de variables  $X$ ; los dos parámetros multivariados representan un estado bajo control del proceso.

El gráfico de control  $T^2$  con dimensión variable y función de distribución  $\chi_{p,\alpha}^2$ , actúa y se interpreta de la misma manera que el gráfico de control de Shewhart, cuando este tiene un límite de control (CL) y probabilidad de error tipo I ( $\alpha$ ).

Cuando el proceso está bajo control, es decir  $\mu_1 - \mu_0 = 0$  se tiene que  $\lambda = 0$ , esto reduce la distribución a una chi-cuadrado centralizada  $\chi_p^2$ . Así pues, existe una probabilidad  $\alpha$  de que el estadístico  $T^2$  supere un punto crítico dado por  $\chi_{p,\alpha}^2$  por lo que la tasa de error global puede mantenerse exactamente en el nivel  $\alpha$  mediante el uso de un gráfico de control de Hotelling. Cuando el proceso está fuera de control, es decir  $\mu_1 \neq \mu_0$  se obtiene  $\lambda \neq 0$ .

Para determinar la probabilidad de que el proceso se salga de control, se utiliza la función de distribución chi-cuadrado no centralizada, dada por:

$$\lambda_{p,\lambda}^2(x) = e^{(-\lambda/2)} \sum_{j=0}^{\infty} \left[ \frac{\left(\frac{\lambda}{2}\right)^j}{j!} * \frac{1}{\left[2^{\left(\frac{p}{2}+j\right)} * \Gamma\left(\frac{p}{2}+j\right)\right]} * \int_0^x y^{\left(\frac{p}{2}+j+1\right)} * e^{(-y/2)} dy \right] \quad (2.8)$$

Para obtener el promedio de la longitud de las rachas (ARL) como indicador del rendimiento del gráfico de control, se hace uso de las cadenas de Markov.

### **Obtención del ARL a través del gráfico $VDT^2$ utilizando cadenas de Markov**

Para medir el rendimiento de un gráfico de control, se utiliza el promedio de la longitud de rachas (ARL) o el número esperado de muestras hasta que el gráfico presente una señal de fuera de control. El ARL del gráfico de control estándar  $T^2$  se obtiene fácilmente como  $ARL = 1/Pr$ , donde  $Pr$  es la probabilidad para cualquier muestra dada que el estadístico  $T^2$  exceda el límite de control CL.

Cuando el proceso está bajo control, se tiene  $ARL = \frac{1}{\alpha}$

Y cuando el proceso está fuera de control,  $ARL = \frac{1}{1-\beta} = \frac{1}{[1-p(\chi_p^2(\lambda) < CL)]}$

Donde  $\chi_p^2(\lambda)$  es la distribución chi-cuadrada con  $p$  grados de libertad y parámetro de no centralidad  $\lambda = n \cdot d^2$ ,  $n$  es el tamaño de la muestra y  $d$  es la distancia de Mahalanobis.

El ARL se puede obtener a través de la cadena de Markov (Ross, 1996). Este proceso requiere definir elementos básicos tales como:

- Los estados que alcanza la variable aleatoria
- Las probabilidades de pasar de un estado a otro
- El estado inicial del proceso.

Se puede obtener el ARL para el gráfico  $VDT^2$ , a través de matrices (Brooks y Evans, 1972), tal como se indica en la ecuación:

$$ARL = \mathbf{b}'(\mathbf{I} - \mathbf{Q})' \mathbf{1} \quad (2.9)$$

Siendo  $\mathbf{b}$  el vector correspondiente al estado inicial del proceso markoviano el cual se detalla más adelante,  $\mathbf{I}$  la matriz identidad,  $\mathbf{Q}$  la matriz de transición y  $\mathbf{1}$  el vector unitario.

El vector  $\mathbf{b}$  de probabilidades iniciales es:

- $\mathbf{b} = [1 \ 0]$  Si la monitorización empieza siempre con las  $p_1$  variables.
- $\mathbf{b} = [0 \ 1]$  Si la monitorización empieza siempre con todas las  $p$  variables.

El ARL del gráfico de control  $VDT^2$  es calculado a través de la cadena de Markov con tres estados:

- Estado 1: El estadístico  $T^2$  cae bajo la línea de alerta  $w$  entonces la siguiente muestra contendrá solo  $p_1$  variables.
- Estado 2: El estadístico  $T^2$  cae entre la línea de alerta y el límite de control entonces la siguiente muestra contendrá todas las  $p$  variables.
- Estado 3: El estadístico  $T^2$  excede el respectivo límite de control (estado absorbente de la cadena de Markov).

La matriz de probabilidades de transición para la cadena de Markov es:

$$P^s = \begin{pmatrix} p_{11}^s & p_{12}^s & p_{13}^s \\ p_{21}^s & p_{22}^s & p_{23}^s \\ 0 & 0 & 1 \end{pmatrix}$$

Cuando el superíndice  $s$  indica el estado del proceso:

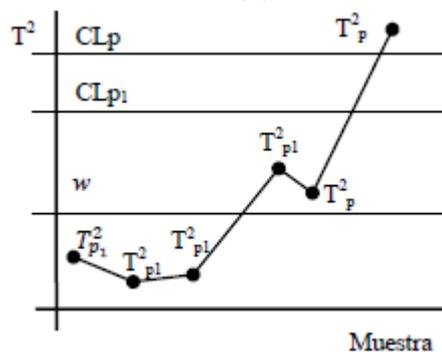
- $s = 0$ , si el proceso está bajo control.
- $s = 1$ , si el proceso está fuera de control.

con un cambio en la media correspondiente a la distancia  $d_1$  en el subespacio de las  $p_1$  variables y para una distancia  $d$  en el caso de utilizar el espacio  $p$ -dimensional de todas las variables.

Los elementos de la matriz  $\mathbf{P}$  están dados por:

- $p_{11}^s = P(T_{p_1}^2 < w_{p_1}) = P(\chi_{p_1}^2(\lambda_{p_1}) < w_{p_1})$
- $p_{21}^s = P(T_p^2 < w_p) = P(\chi_p^2(\lambda_p) < w_p)$
- $p_{12}^s = P(w_{p_1} \leq T_{p_1}^2 < CL_{p_1}) = P(w_{p_1} \leq \chi_{p_1}^2(\lambda_{p_1}) < CL_{p_1})$
- $p_{22}^s = P(w_p \leq T_p^2 < CL_p) = P(w_p \leq \chi_p^2(\lambda_p) < CL_p)$
- $p_{13}^s = P(CL_{p_1} \leq T_{p_1}^2) = P(CL_{p_1} \leq \chi_{p_1}^2(\lambda_{p_1}))$
- $p_{23}^s = P(CL_p \leq T_p^2) = P(CL_p \leq \chi_p^2(\lambda_p))$

### Gráfico de control VDT<sup>2</sup>



**Gráfico 2.3 Gráfico de Control con Dimensión Variable.**

**Fuente:** Artículo "T2 Control Charts with Variable Dimension" de Aparisi et.al., 2012.

La metodología del Gráfico de control multivariante con dimensión variable se obtuvo del artículo: “ $T^2$  Control Charts with Variable Dimension” de Aparisi et.al., 2012.

### 2.4.3 Gráfico de Control $T^2$ con Doble Dimensión

El gráfico de control multivariante  $T^2$  de doble dimensión ( $DDT^2$ ) en cada tiempo muestrea las  $p_1$  variables que son baratas o rápidas de medir, se miden y se calcula el estadístico  $T_{p_1}^2$ . Si su valor es inferior a un umbral dado (denominado límite de advertencia), el proceso se considera bajo control. Si su valor es demasiado alto (por encima del límite de control), el proceso es declarado fuera de control. Si su valor está entre el límite de advertencia y el límite de control se miden las  $p_2$  variables restantes (siendo  $p = p_1 + p_2$ ) y el estadístico global se compara con el límite de control apropiado (hay diferentes límites de control para el conjunto de variables  $p_1$  de bajo costo y para el conjunto total de variables  $p$ ).

Por lo tanto, si se emplea el gráfico de control  $DDT^2$  las variables  $p_2$  que son difíciles o costosas de medir, se miden sólo cuando existe la necesidad de adquirir más información. De esta manera, el coste de muestreo se reduce en comparación con el procedimiento de muestrear siempre todas las  $p$  variables para calcular la estadística. Sin embargo, la potencia del gráfico propuesto para detectar cambios en el proceso es mayor que la potencia de un gráfico  $T_2$  que utiliza sólo el conjunto de  $p_1$  variables que son fáciles y/o baratas de medir puesto que la información proporcionada por las  $p_2$  variables que son caras o difíciles de medir cuando se miden, la capacidad de detectar cambios en el proceso aumenta significativamente.

#### Generalidades del gráfico de control de Doble Dimensión ( $DDT^2$ )

El gráfico de control estándar  $T^2$  consiste en tomar periódicamente una muestra de tamaño  $n$ . Posteriormente se realiza el cálculo del vector de promedios

muestrales  $\bar{X}$  de las  $p$  variables a ser monitorizadas, a continuación se calcula el estadístico  $T^2 = n(\bar{X} - \mu_0)' \Sigma^{-1} (\bar{X} - \mu_0)$ .

Donde  $\mu_0 = (\mu_{01}, \mu_{02}, \dots, \mu_{0p})$  es el vector de promedios y  $\Sigma$  es la matriz de covarianzas del estado bajo control del vector de variables  $X$ .

La monitorización consiste en muestrear  $n$  ítems del proceso y medir en primera instancia solo el subgrupo formado por las  $p_1$  variables que son fáciles y/o baratas de medir. Con esta información, se calcula y grafica el estadístico  $T_{p_1}^2$  de Hotelling sobre el gráfico de control.

- Si  $T_{p_1}^2 < w$  no se toma ninguna acción futura.
- Si  $T_{p_1}^2 \geq CL_{p_1}$  el proceso es considerado fuera de control.

La segunda etapa del procedimiento se hace necesaria cuando  $w \leq T_{p_1}^2 < CL_{p_1}$ , como el valor del estadístico  $T_{p_1}^2$  es alto pero no lo suficiente como para considerar que el proceso está fuera de control entonces debe ser analizada más información estadística de la muestra. Así, las  $p_2$  variables restantes ( $p_2 = p - p_1$ ) se miden a partir de las mismas unidades de la muestra de tamaño  $n$  antes considerada y todas las mediciones de las  $p$  variables de las  $n$  unidades se emplean para calcular el estadístico  $T_p^2$ . Si  $T_p^2 > CL_p$  el proceso se considera fuera de control de lo contrario, se considera bajo control.

### **Obtención del número promedio de muestras necesarias para detectar un cambio en el proceso, ARL**

La medida más común del rendimiento estadístico es el promedio de la longitud de rachas (ARL) que es el número esperado de muestras hasta que una señal de fuera de control es señalada por el gráfico de control.

Cuando los puntos marcados en el gráfico son independientes, el ARL es simplemente el inverso de la probabilidad de que un punto caiga fuera de los límites de control.

El cálculo del ARL para el gráfico de control estándar  $T^2$  cuando el proceso está bajo control es:

$$ARL = \frac{1}{\alpha}$$

Donde  $\alpha$  es la probabilidad de falsa alarma.

$$ARL = \frac{1}{1 - \beta} = \frac{1}{1 - P[\chi_p^2(\lambda) < CL]}$$

Donde  $\beta$  es la probabilidad de error tipo II y  $\chi_p^2(\lambda)$  es la función de distribución chi-cuadrado con  $p$  grados de libertad y parámetro de no centralidad  $\lambda = n \cdot d^2$ , y  $d$  es la distancia de Mahalanobis. Por lo tanto, el ARL del gráfico de control  $DDT^2$  se calcula como sigue:

$$ARL = \frac{1}{1 - P_a}$$

Siendo  $P_a$  la probabilidad de ausencia de señal, dada por

$$P_a = P_1 + P_2 \quad (2.10)$$

Y a su vez,  $P_1$  es la probabilidad de que el primer estadístico calculado con sólo el conjunto de  $p_1$  variables no de señales de fuera de control ni indique la necesidad de medir las variables restantes es decir,

$$P_1 = P(T_{p_1}^2 < w) \quad (2.11)$$

Dado que el primer estadístico cayó en la región de advertencia entonces  $P_2$  es la probabilidad de que el primer estadístico caiga entre el límite de alerta y el

límite de control y que el estadístico  $T^2$  obtenido con las  $p$  variables sea menor que el límite de control.

Por lo tanto, existe la necesidad de calcular un estadístico  $T^2$  con todas las  $p$  variables, y este estadístico no supera el límite de control de  $CL_p$ :

$$P_2 = P(w \leq T_{p1}^2 < CL_{p1}) * P(T_p^2 < CL_p | w \leq T_{p1}^2 < CL_{p1}) \quad (2.12)$$

Para poder calcular  $P_2$ , es necesaria la distribución condicional de  $T_p^2$ . Murphy (1987) muestra que cuando el proceso está bajo control, con  $\mu = 0$ , el estadístico  $D = T_p^2 - T_{p1}^2$  se distribuye como  $\chi_{p2}^2$ , donde  $p_2 = p - p_1$ .

Por lo tanto, cuando el proceso está bajo control, lo calcularemos por medio de:

$$P(T_p^2 < CL_p | w \leq T_{p1}^2 < CL_{p1}) = P(D < CL_p - T_{p1}^2) = P(\chi_{p2}^2 < CL_p - T_{p1}^2) \quad (2.13)$$

Cuando el proceso está fuera de control  $D \sim \chi_{p2}^2[\lambda = n(d_p^2 - d_{p1}^2)]$ .

$$\begin{aligned} P(T_p^2 < CL_p | w \leq T_{p1}^2 < CL_{p1}) &= P(D < CL_p - T_{p1}^2) \\ &= P(T_p^2 < CL_p | w \leq T_{p1}^2 < CL_{p1}) \\ &= P(\chi_{p2}^2[\lambda = n(d_p^2 - d_{p1}^2)] < CL_p - T_{p1}^2) \end{aligned} \quad (2.14)$$

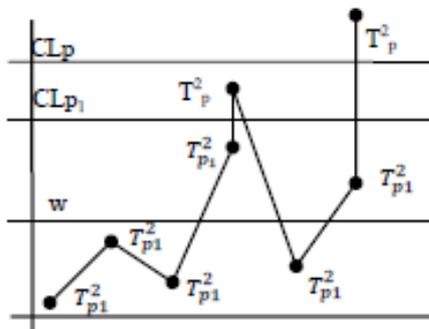
Combinando (2.12) y (2.13), cuando el proceso está bajo control,

$$P_2 = [F_{p1}(CL_{p1}) - F_{p1}(w)] * F_{p2}(CL_p - T_{p1}^2) \quad (2.15)$$

Donde  $F_v$  es la función de distribución chi-cuadrado con  $v$  grados de libertad; y cuando está fuera de control se obtiene:

$$P_2 = [F_{p1,nd_1^2}(CL_{p1,nd_1^2}) - F_{p1}(w)] * F_{p2,nd_1^2}(CL_p - T_{p1}^2) \quad (2.16)$$

### Gráfico de control DDT<sup>2</sup>



**Gráfico 2.4 Gráfico de control con doble dimensión**

**Fuente:** Artículo "Reducing Sampling Costs of Multivariate SPC with a Double-Dimension T2 Control Chart" de Eppretch et.al., 2013.

La metodología del Gráfico de control multivariante con doble dimensión se obtuvo del artículo: "Reducing Sampling Costs of Multivariate SPC with a Double-Dimension T2 Control Chart" de Eppretch et.al., 2013.

## 2.5 Simulación de variables aleatorias con Distribución Normal Multivariante

### 2.5.1 Distribución Normal Multivariante $N(\mu, \Sigma)$

Un vector aleatorio  $X = (X_1, \dots, X_p)'$  que toma valores en  $\mathbb{R}^p$  tiene distribución normal p-variente si su densidad es de la forma:

$$f(x) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}, \quad x = (x_1, \dots, x_p) \in \mathbb{R}^p$$

Donde  $\Sigma_{p \times p}$  es una matriz simétrica definida positiva y  $\mu \in \mathbb{R}^p$ .

### 2.5.2 Método basado en la triangulación de Cholesky

Sea el vector aleatorio  $Z = (Z_1, \dots, Z_n)' \sim N(0, I)$  siendo  $I$  la matriz identidad, entonces  $X = \mu + AZ \sim N(\mu, AA^T = \Sigma)$  donde  $A$  es la matriz triangular inferior de Cholesky.

Algoritmo:

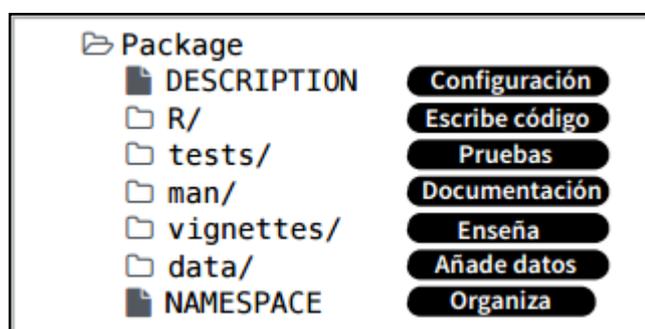
- Calcular  $A$  matriz triangular inferior de Cholesky.
- Generar  $Z_1, \dots, Z_n$  con distribución  $N(0,1)$ . Hacer  $Z = (Z_1, \dots, Z_n)^T$
- Salida:  $X = \mu + AZ$

## 2.6 Desarrollo de paquetes en R

Los paquetes de R son la forma ideal de distribuir código R para que otras personas los puedan reutilizar. RStudio incluye un conjunto de herramientas que facilitan la creación y desarrollo de paquetes en R de forma más productiva.

### 2.6.1 Estructura de paquetes

Las fuentes de un paquete de R están contenida en un directorio como se muestra en la ilustración 2.1.



**Ilustración 2.1 Estructura de un paquete de R**  
Fuente: Rstudio (2015). Desarrollo de Paquetes

El contenido del paquete se puede guardar en discos como:

- *Fuente*: es una carpeta con subcarpetas como se muestra en el grafico ###.
- *Paquete*: es un solo archivo comprimido.

- *Binario*: es un solo archivo comprimido optimizado para un sistema operativo específico.

### 2.6.2 DESCRIPTION

Contiene la información básica del paquete con un formato preestablecido. Los campos Package, Version, License, Title, Author y Maintainer son obligatorios.

```

Package: mipaquete
Title: Título del paquete
Version: 0.1.0
Authors@R: person("Hadley", "Wickham", email =
  "hadley@me.com", role = c("aut", "cre"))
Description: Lo que el paquete hace (en un párrafo)
Depends: R (>= 3.1.0)
License: GPL-2
LazyData: true
Imports:
  dplyr (>= 0.4.0),
  ggvis (>= 0.2)
Suggests:
  knitr (>= 0.1.0)
Imports: paquetes que tu paquete
necesita tener para funcionar. R los
instalará cuando instalas tu paquete.
Suggests: paquetes que no son
esenciales para el tuyo. Usuarios los
pueden instalar manualmente, si
gustan, o no instalarlos.

```

**Ilustración 2.2 DESCRIPTION de un paquete de R**  
Fuente: Rstudio (2015). Desarrollo de Paquetes

### 2.6.3 NAMESPACE

Aquí el autor puede especificar las variables del paquete que se exportan y van a ser utilizadas por los usuarios. Las variables importadas de otros paquetes, las clases y métodos.

Para exportar las variables var1 y var2:

- `export(var1, var2)`

Para importar las variables var1 y var2 del paquete Pkgext:

- `importFrom(Pkgext, var1, var2)`

Para importar métodos y clases de otro paquete:

- `importClassesFrom(package,...)`
- `importMethodsFrom(package,...)`

#### 2.6.4 Documentación (Carpeta man)

Las páginas de ayuda de los objetos R se escriben usando el formato “R documentation” (.Rd), un lenguaje similar a LATEX.

Todos los comandos disponibles en el manual Writing R Extensions de la página de R.

```

\name{load}
\alias{load}
\title{Reload Saved Datasets}
\description{
  Reload the datasets written to a file with the function
  \code{save}.
}
\usage{
  load(file, envir = parent.frame())
}
\arguments{
\item{file}{a connection or a character string giving the
  name of the file to load.}
\item{envir}{the environment where the data should be
  loaded.}
}
\seealso{
  \code{\link{save}}.
}
\examples{
## save all data
save(list = ls(), file = "all.RData")

## restore the saved values to the current environment
load("all.RData")

## restore the saved values to the workspace
load("all.RData", .GlobalEnv)
}
\keyword{file}

```

#### Ilustración 2.3 Estructura de la carpeta man

Fuente: Oscar Perpiñan Lamigueiro (2015). Crear paquetes en R.

#### 2.6.5 Código (Carpeta R)

En la carpeta R se debe agregar todas las líneas de código que deseamos compartir en nuestro paquete.

### 2.6.6 Comprobar, construir y Publicar

Para comprobar el paquete desarrollado: R CMD check MiPaquete/

Para construir el paquete: R CMD build MiPaquete/

El principal repositorio de paquetes estables es CRAN publicar en este repositorio conlleva la aceptación de varias condiciones que se deben cumplir estas se encuentran especificadas en el *CRAN Repository Policy*.

Otros repositorios en los que podemos compartir un paquete son: GitHub, R-Forge, RForge, Bioconductor.

## CAPÍTULO 3

### 3. RESULTADOS

Se realizó la simulación de variables aleatorias normales multivariantes, tomando como parámetros  $\mu$  y  $\Sigma$

Donde:

- $\mu$ : El vector de las medias de las variables (Parámetros del proceso cuando está bajo control).
- $\Sigma$ : La matriz de varianzas y covarianzas esta debe ser una matriz simétrica y definida positiva (Parámetros del proceso cuando está bajo control).

Se realizó la simulación de 250 observaciones y se tomó como parámetros los siguientes valores:

$$\mu = \begin{bmatrix} 4.5 \\ 7 \\ 8.45 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 2.0 & 1.5 & 2.4 \\ 1.5 & 3.0 & 3.1 \\ 2.4 & 3.1 & 4.0 \end{bmatrix}$$

La tabla de los datos simulados y las líneas de código de la simulación de las variables las encontraremos en anexos.

#### Parámetros del proceso cuando está bajo control

$$\mu_0 = [ 5.4, 6.8, 8.5 ]$$

$$\Sigma_0 = \begin{bmatrix} 2.0 & 1.5 & 2.4 \\ 1.5 & 3.0 & 3.1 \\ 2.4 & 3.1 & 4.0 \end{bmatrix}$$

Posterior a la simulación de variables, se realizó las funciones que permiten realizar los gráficos de control en el software estadístico R. Y se procedió a desarrollar el paquete.

### 3.1. Desarrollo y publicación del Paquete Gráficos de control multivariante

#### Documentación del Paquete Gráficos de Control Multivariante

R Documentation

Gcmulti-package {Gcmulti}

## Gráficos de Control Multivariante

**Description**

Permite realizar los gráficos de control multivariante T2 de Hotelling, de doble dimension y de dimension variable. Además calcula los ARL bajo y fuera de control.

**Details**

Package: Gcmulti  
 Type: Package  
 Version: 1.0  
 Date: 2016-08-29  
 License: GPL (>= 2)

**Author(s)**

Jhonny Barahona  
 Omar Ruiz  
 Maintainer: Jhonny Barahona <jfbaraho@espol.edu.ec>

**References**

[1] Aparisi, F.; Epprecht, E.; Ruiz, O.; Veiga A., "Reducing Sampling Costs of Multivariate SPC with a Double-Dimension T<sup>2</sup> Control Chart", International Journal of Production Research, vol. 1, no. 1, pag. 1-15, 2013.  
 [2] Aparisi, F.; Epprecht, E.; Ruiz, O., "T2 Control Charts with Variable Dimension", Journal of Quality Technology, vol. 44, no. 4, pp. 375-393, 2012.  
 [3] Ruiz, O., "Gráficos de control de calidad multivariantes con dimensión variable" Tesis doctoral, Dept. de Estadística e Investigación Operativa, Univ. Pol. de Valencia, España, 2013.

**Ilustración 3.1 Paquete Gcmulti Gráficos de Control Multivariante**  
**Fuente: El autor**

#### Documentación de la función que realiza el gráfico de control T<sup>2</sup> de Hotelling

R Documentation

grafica\_T2 {Gcmulti}

## Grafica T2 de Hotelling

**Description**

Esta función permite realizar la grafica de control T2 de Hotelling y obtener sus ARL's bajo y fuera de control.

**Usage**

```
grafica_T2(mu, sigma, X, LCS, d, k)
```

**Arguments**

**mu:** El vector de las medias de las variables (Parametros del proceso cuando esta bajo control).

**sigma:** La matriz de varianzas y covarianzas (Parametros del proceso cuando esta bajo control).

**X:** es la matriz que contiene las variables

**LCS:** El límite de control superior.

**d:** La distancia de mahalanobis.

**k:** El tamaño de la muestra.

**Value**

Grafico de control multivariante T2 de Hotelling;  
 ARL0: Promedio de longitud de rachas cuando el proceso esta bajo control;  
 ARL1: Promedio de longitud de rachas cuando el proceso esta fuera de control.

**Ilustración 3.2 Documentación gráfico T2 de Hotelling**  
**Fuente: El autor**

Documentación de la función que realiza el gráfico de Doble Dimensión DDT<sup>2</sup>.

R Documentation

**grafica\_T2DD (Gcmulti)**

**Grafico de doble dimension DDT2**

**Description**

Esta funcion permite realizar la grafica de control de doble dimension y obtener sus ARL's bajo y fuera de control.

**Usage**

```
grafica_T2DD(p1, p, w, CLP1, CLP, displ, disp, mu, sigma, X, k, inter = 0.05)
```

**Arguments**

p1: Numero de variables p1 (Variables faciles de medir o baratas).  
 p2: Numero de variables p2 (Variables complejas de medir o costosas).  
 p: Numero total de variables.  
 w: Limite de alerta o de advertencia.  
 CLP1: Limite de control p1.  
 CLP: Limite de control p.  
 displ: Distancia de mahalanobis de p1.  
 disp: Distancia de mahalanobis de p.  
 mu: El vector de las medias de las variables (Parametros del proceso cuando esta bajo control).  
 sigma: La matriz de varianzas y covarianzas (Parametros del proceso cuando esta bajo control).  
 X: es la matriz que contiene las variables.  
 k: El tamaño de la muestra.

**Value**

Grafico de control multivariante de doble dimension.  
 ARL0: Promedio de longitud de rachas cuando el proceso esta bajo control.  
 ARL1: Promedio de longitud de rachas cuando el proceso esta fuera de control.

**Ilustración 3.3 Documentación gráfico de doble dimensión DDT<sup>2</sup>**  
**Fuente: El autor**

Documentación de la función que realiza el gráfico de control de dimensión variable VDT<sup>2</sup>.

R Documentation

**grafica\_T2DV (Gcmulti)**

**Grafico de dimension variable VDT2**

**Description**

Esta funcion permite realizar graficos de control de dimension variable y obtener sus ARL's bajo y fuera de control.

**Usage**

```
grafica_T2DV(p1, p, w, CLp1, CLp, displ, disp, mu, sigma, X, k, ProbInic_p1 = 1)
```

**Arguments**

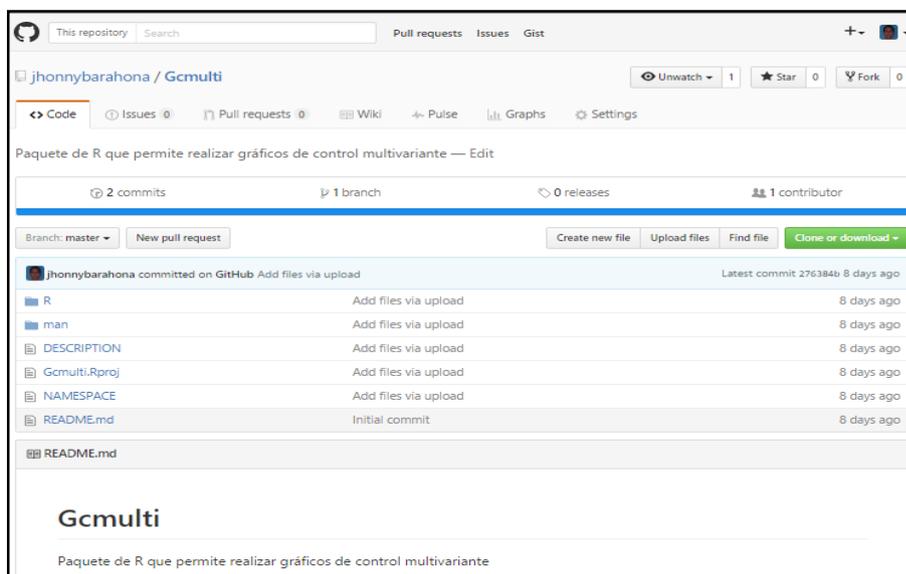
p1: Numero de variables p1 (Variables fáciles de medir o baratas).  
 p2: Numero de variables p2 (Variables complejas de medir o costosas).  
 p: Numero total de variables.  
 w: Limite de alerta o de advertencia.  
 CLP1: Limite de control p1.  
 CLP: Limite de control p.  
 displ: Distancia de mahalanobis de p1.  
 disp: Distancia de mahalanobis de p.  
 mu: El vector de las medias de las variables (Parametros del proceso cuando esta bajo control).  
 sigma: La matriz de varianzas y covarianzas (Parametros del proceso cuando esta bajo control).  
 X: es la matriz que contiene las variables.  
 k: El tamaño de la muestra.

**Value**

Grafico de control multivariante de dimension variable.  
 ARL0: Promedio de longitud de rachas cuando el proceso esta bajo control.  
 ARL1: Promedio de longitud de rachas cuando el proceso esta fuera de control.

**Ilustración 3.4 Documentación gráfico de dimensión variable VDT<sup>2</sup>**  
**Fuente: El autor**

La publicación del paquete que permite realizar gráficos de control multivariante fue realizada en el repositorio de GitHub y está a disposición de los usuarios del R.



**Ilustración 3.5 Paquete desarrollado en GitHub**  
Fuente: El autor

El paquete lo podemos descargar con las siguientes instrucciones:

```
library(devtools)
install_github("jhonnybarahona/Gcmulti")
```

```
> library(devtools)
> install_github("jhonnybarahona/Gcmulti")
Downloading GitHub repo jhonnybarahona/Gcmulti@master
From URL https://api.github.com/repos/jhonnybarahona/Gcmulti/zipball/master
Installing Gcmulti
"C:/PROGRAMAS/R/R-33-1.1/bin/x64/R" --no-site-file --no-enviro --no-save --no-restore --quiet \
CMD INSTALL "C:/Users/Fabricao \
Barahona/AppData/Local/Temp/RtmpMJUzyP/devtoolsff07cd923c3/jhonnybarahona-Gcmulti-6e7a2a1" \
--library="C:/Users/Fabricao \
Barahona/Documents/R/win-library/3.3" --install-tests

* installing *source* package 'Gcmulti' ...
** R
** preparing package for lazy loading
** help
*** installing help indices
** building package indices
** testing if installed package can be loaded
* DONE (Gcmulti)
```

**Ilustración 3.6 Ejemplo de instalación del paquete Gcmulti**  
Fuente: El autor

En caso de no tener descargada la librería *devtools* se deberá descargar. Una vez instalado el paquete ya se puede hacer uso de las funciones desarrolladas.

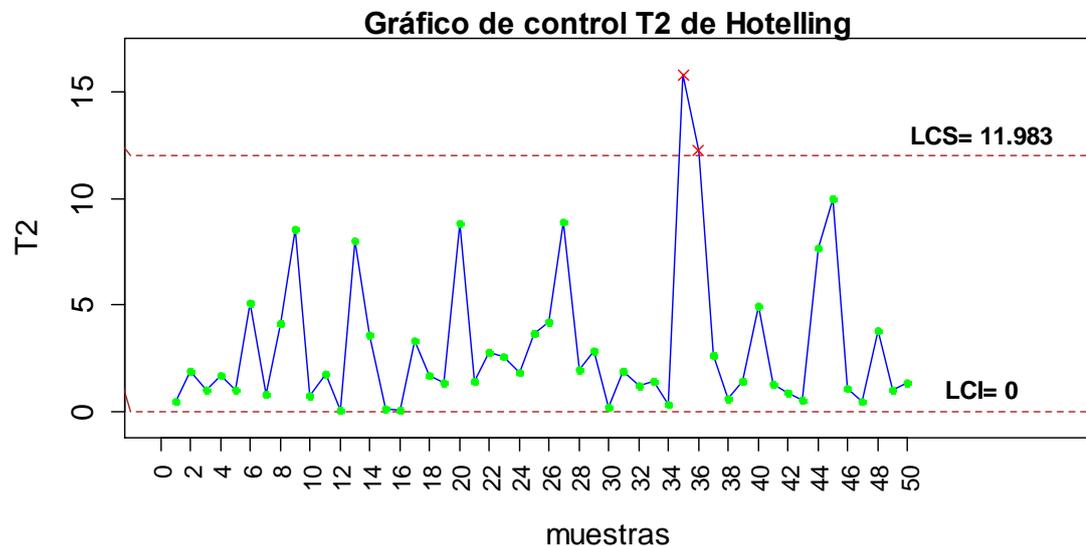
### 3.2. Gráficos de control $T^2$ de Hotelling

Los parámetros utilizados para el gráfico de control  $T^2$  de Hotelling son:

- $n$ : El tamaño de la muestra.
- $\mu$ : El vector de las medias de las variables (Parámetros del proceso cuando está bajo control).
- $\Sigma$ : La matriz de varianzas y covarianzas (Parámetros del proceso cuando está bajo control).
- $d$ : La distancia de mahalanobis.
- $LCS$ : El límite de control superior.
- $p1$ : Número de variables  $p1$  (Variables fáciles de medir o baratas).
- $p2$ : Número de variables  $p2$  (Variables complejas de medir o costosas).
- $p$ : Número total de variables.  $p = p1 + p2$

Con los datos simulados procedemos a realizar gráficos de control  $T^2$  de Hotelling con diferentes escenarios. Las líneas de código en R de la gráfica de control  $T^2$  de Hotelling las encontraremos en los anexos.

Gráfico de control  $T^2$  de Hotelling con  $p_1$  variables utilizando los siguientes parámetros:  $\mu, \Sigma$  cuando el proceso está bajo control,  $n = 1$ ,  $LCS = 11.983$ ,  $d = 1$



**Gráfico 3.1 Gráfico de control T2 con  $p_1$  variables y  $n=1$**

$$ARL_0 = 400, ARL_d = 71.451$$

En el gráfico de control  $T^2$  de Hotelling podemos apreciar que en las muestras 35 y 36 superan el límite de control superior por lo que se concluye que el proceso está fuera de control.

En promedio se requerirá muestrear 400 veces hasta tener una falsa alarma o hasta obtener una muestra que de señal de fuera de control cuando realmente el proceso está bajo control.

En promedio se requerirá muestrear 72 veces hasta obtener una muestra que de señal de bajo control cuando realmente el proceso está fuera de control.

Gráfico de control  $T^2$  de Hotelling con  $p=1$  variables, utilizando los siguientes parámetros:  $\mu, \Sigma$  cuando el proceso está bajo control,  $n = 5$ ,  $LCS = 11.983$ ,  $d = 1.5$



**Gráfico 3.2 Gráfico de control T2 con  $p=1$  variables y  $n=5$**

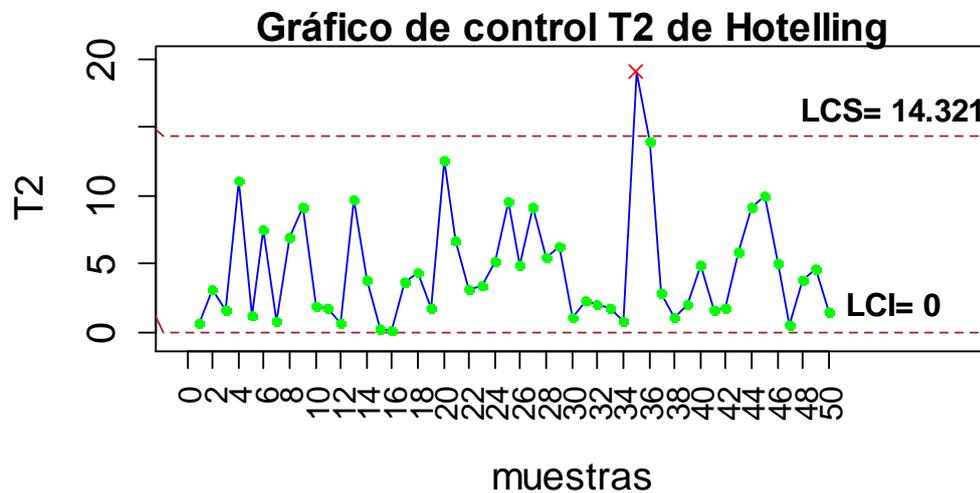
$$ARL_0 = 400, ARL_d = 1.936$$

En el gráfico de control  $T^2$  de Hotelling podemos apreciar que ninguna de las muestras supera el límite de control superior por lo que se concluye que el proceso está bajo control.

En promedio se requerirá muestrear 400 veces hasta tener una falsa alarma o hasta obtener una muestra que de señal de fuera de control cuando realmente el proceso está bajo control.

En promedio se requerirá muestrear 2 veces hasta obtener una muestra que de señal de bajo control cuando realmente el proceso está fuera de control.

Gráfico de control  $T^2$  de Hotelling con  $p$  variables utilizando los siguientes parámetros:  $\mu, \Sigma$  cuando el proceso está bajo control,  $n = 1$   $LCS = 14.321, d = 1.2$



**Gráfico 3.3 Gráfico de control T2 con  $p$  variables y  $n=1$**

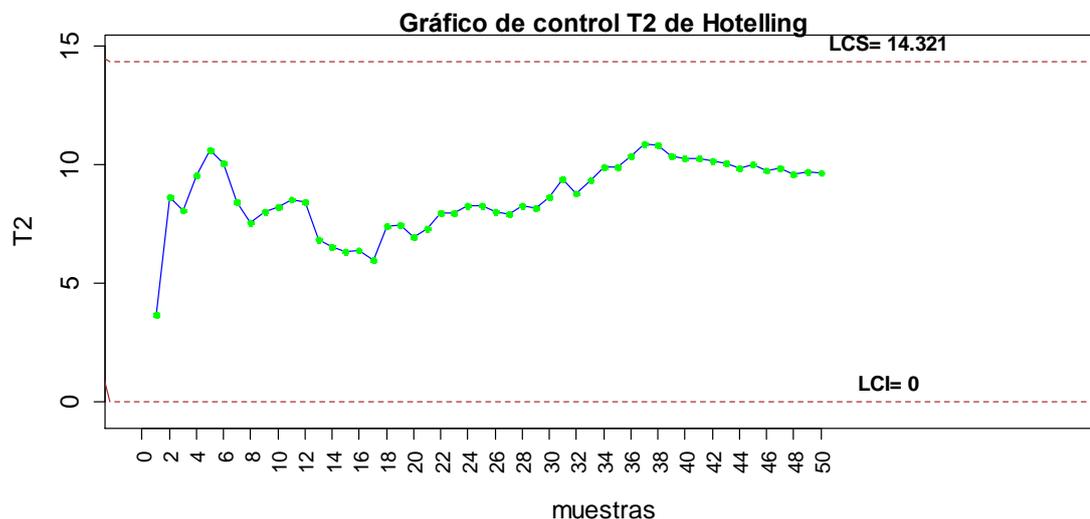
$$ARL_0 = 400, ARL_d = 59.99$$

En el gráfico de control  $T^2$  de Hotelling podemos apreciar que en la muestra 35 es la única que supera el límite de control superior por lo que se concluye que el proceso está fuera de control.

En promedio se requerirá muestrear 400 veces hasta tener una falsa alarma o hasta obtener una muestra que de señal de fuera de control cuando realmente el proceso está bajo control.

En promedio se requerirá muestrear 60 veces hasta obtener una muestra que de señal de bajo control cuando realmente el proceso está fuera de control.

Gráfico de control  $T^2$  de Hotelling con  $p$  variables utilizando los siguientes parámetros:  $\mu, \Sigma$  cuando el proceso está bajo control,  $n = 5$   $LCS = 14.321, d = 2$



**Gráfico 3.4 Gráfico de control T2 con  $p$  variables y  $n=5$**

$$ARL_0 = 400, ARL_d = 1.213$$

En el gráfico de control  $T^2$  de Hotelling podemos apreciar que ninguna de las muestras supera el límite de control superior por lo que se concluye que el proceso está bajo control.

En promedio se requerirá muestrear 400 veces hasta tener una falsa alarma o hasta obtener una muestra que de señal de fuera de control cuando realmente el proceso está bajo control.

En promedio se requerirá muestrear 2 veces hasta obtener una muestra que de señal de bajo control cuando realmente el proceso está fuera de control.

### 3.3. Gráficos de control $T^2$ con dimensión variable

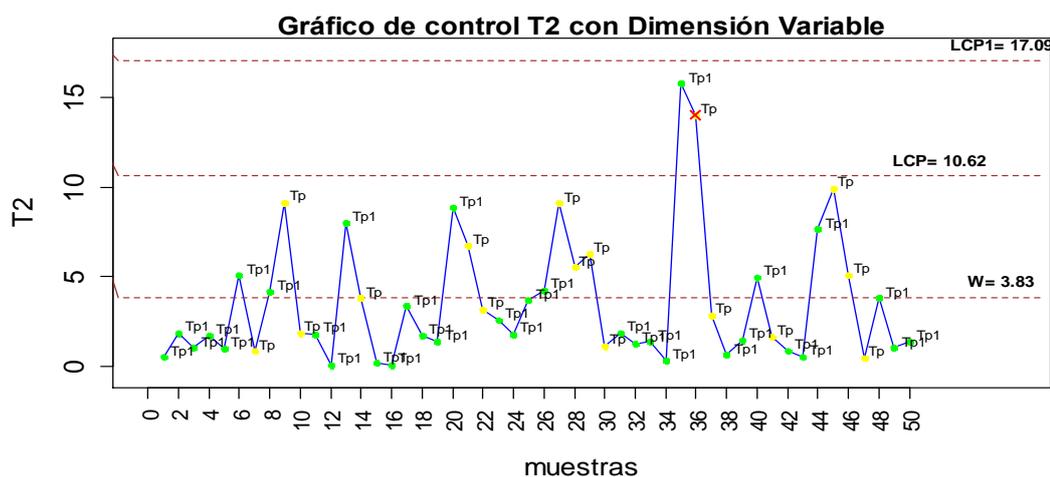
Los parámetros utilizados para el gráfico de control  $T^2$  con dimensión variable son:

- $\mu$ : El vector de las medias de las variables (Cuando el proceso está bajo control).
- $\Sigma$ : La matriz de varianzas y covarianzas (Cuando el proceso está bajo control).
- $p_1$ : Número de variables  $p_1$  (Variables fáciles de medir o baratas).
- $p_2$ : Número de variables  $p_2$  (Variables complejas de medir o costosas).
- $p$ : Número de variables.
- $w$ : Límite de alerta o de advertencia.
- $CLP_1$ : Límite de control  $p_1$ .
- $CLP$ : Límite de control  $p$ .
- $dis_{p_1}$ : Distancia de mahalanobis de  $p_1$ .
- $dis_p$ : Distancia de mahalanobis de  $p$ .
- $n$ : El tamaño de la muestra.

Con los datos simulados procedemos a realizar gráficos de control  $T^2$  con dimensión variable con diferentes escenarios. Las líneas de código en R de la gráfica de control de dimensión variable  $VDT^2$  las encontraremos en los anexos.

Gráfico de control  $T^2$  con dimensión variable utilizando los siguientes parámetros:

$\mu, \Sigma, p_1 = 2, p = 3, w = 3.83, CLP_1 = 17.09, CLP = 10.62, dis_{p_1} = 1, dis_p = 1.2, n = 1.$



**Gráfico 3.5 Gráfico de control T2 con dimensión variable con  $n=1$**

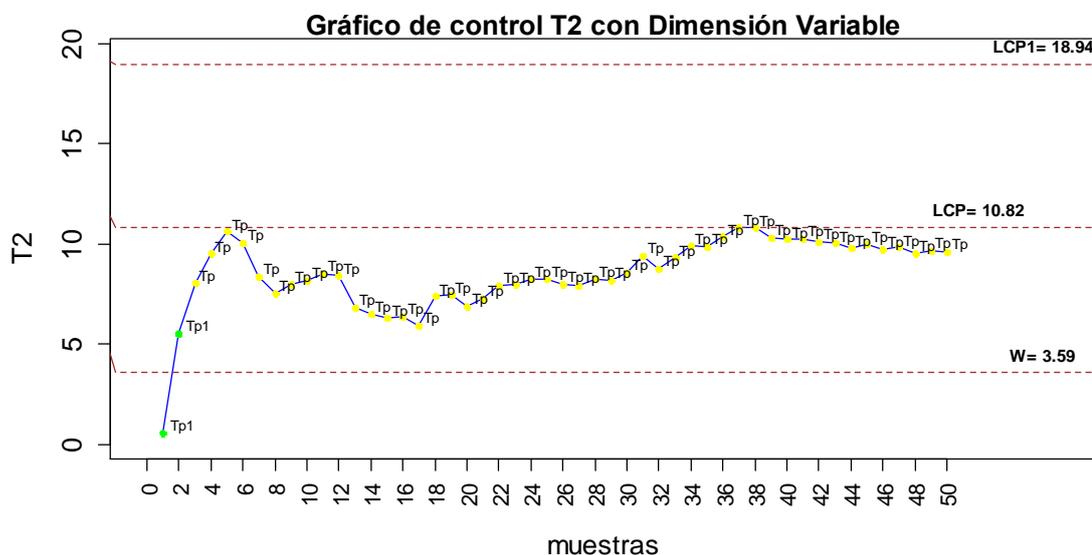
$$ARL_0 = 400, ARL_d = 47.622$$

En el gráfico de control de control  $T^2$  con dimensión variable podemos observar la muestra 36 supera el límite de control para las  $p$  variables por lo que el proceso salió de control estadístico, se debió restablecer el proceso y tomar acciones correctivas.

En promedio se requerirá muestrear 48 veces hasta obtener una muestra que de señal de bajo control cuando realmente el proceso está fuera de control.

Gráfico de control  $T^2$  con dimensión variable utilizando los siguientes parámetros:

$\mu, \Sigma, p_1 = 2, p = 3, w = 3.59, CLP_1 = 18.94, CLP = 10.82, dis_{p_1} = 1.5, dis_p = 2, n = 5.$



**Gráfico 3.6 Gráfico de control T2 con dimensión variable con  $n=5$**

$$ARL_0 = 400, ARL_d = 1.09$$

En el gráfico de control de control  $T^2$  con dimensión variable podemos observar que ninguna muestra supera el supera el límite de control de las  $p_1$  variables.

En promedio se requerirá muestrear 2 veces hasta obtener una muestra que de señal de bajo control cuando realmente el proceso está fuera de control.

### 3.4. Gráficos de control $T^2$ con doble dimensión

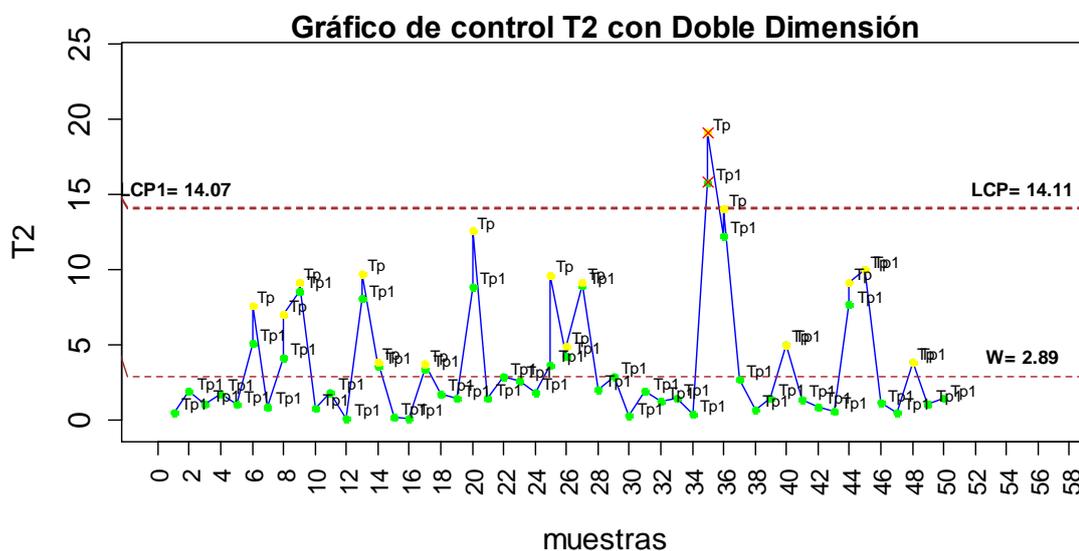
Los parámetros utilizados para el gráfico de control  $T^2$  con doble dimensión son:

- $\mu$ : El vector de las medias de las variables (Cuando el proceso está bajo control).
- $\Sigma$ : La matriz de varianzas y covarianzas (Cuando el proceso está bajo control).
- $p_1$ : Número de variables  $p_1$  (Variables fáciles de medir o baratas).
- $p_2$ : Número de variables  $p_2$  (Variables complejas de medir o costosas).
- $p$ : Número de variables.
- $w$ : Límite de alerta o de advertencia.
- $CLP_1$ : Límite de control  $p_1$ .
- $CLP$ : Límite de control  $p$ .
- $dis_{p_1}$ : Distancia de mahalanobis de  $p_1$ .
- $dis_p$ : Distancia de mahalanobis de  $p$ .
- $n$ : El tamaño de la muestra.

Con los datos simulados procedemos a realizar gráficos de control  $T^2$  con doble dimensión con diferentes escenarios. Las líneas de código en R de la gráfica de control de doble dimensión DDT<sup>2</sup> las encontraremos en los anexos.

Gráfico de control  $T^2$  con doble dimensión utilizando los siguientes parámetros:

$\mu, \Sigma, p_1 = 2, p = 3, w = 2.89, CLP_1 = 14.07, CLP = 14.11, dis_{p_1} = 1, dis_p = 1.2, n = 1.$



**Gráfico 3.7 Gráfico de control T2 con doble dimensión con  $n=1$**

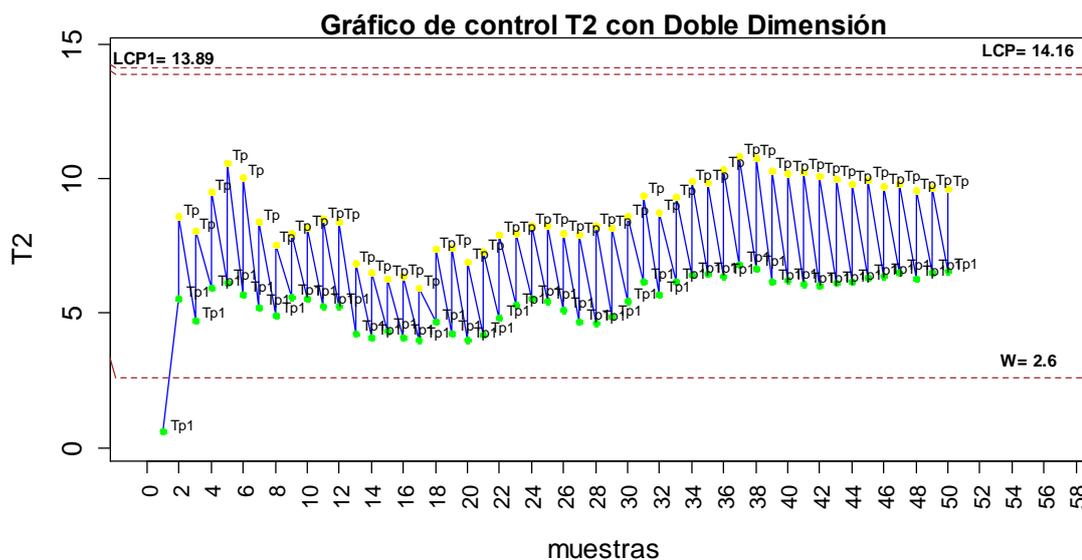
$$ARL_0 = 400, ARL_d = 59.531$$

En el gráfico de control de control  $T^2$  con doble dimensión podemos observar que 2 de las muestras superan los límites de control por lo que en la muestra 35 se debió restablecer la muestra y tomar acciones pertinentes.

En promedio se requerirá muestrear 59 veces hasta obtener una muestra que de señal de bajo control cuando realmente el proceso está fuera de control.

Gráfico de control  $T^2$  con doble dimensión utilizando los siguientes parámetros:

$\mu, \Sigma, p_1 = 2 \quad p = 3, w = 2.60, CLP_1 = 13.89, CLP = 14.16, dis_{p_1} = 1.5, dis_p = 2, n = 5.$



**Gráfico 3.8 Gráfico de control T2 con doble dimensión con  $n=5$**

$$ARL_0 = 400, ARL_d = 1.214$$

En el gráfico de control de control  $T^2$  con doble dimensión podemos observar ninguna de las observaciones supera los límites de control para las  $p_1$  variables; aunque excepto la primera muestra, todas las demás caen encima del límite de alerta  $w$ , por lo que es necesario utilizar todas las  $p$  variables en cada caso, sin que se salga de control el proceso.

En promedio se requerirá muestrear 2 veces hasta obtener una muestra que de señal de bajo control cuando realmente el proceso está fuera de control.

### 3.5. Análisis del rendimiento de los gráficos de control

Parámetros					T2		VDT2				DDT2			
n	p1	p	dp1	dp	ARL1_p1	ARL1_p	W	LCP1	LCP	ARL1	W	LCP1	LCP	ARL1
1	2	3	1	1,2	<b>71,451</b>	<b>59,99</b>	3,83	17,09	10,62	<b>47,622</b>	2,89	14,07	14,11	<b>59,531</b>
5	2	3	1,5	2	<b>1,936</b>	<b>1.213</b>	3,59	18,94	10,82	<b>1,09</b>	2,6	13,89	14,16	<b>1.214</b>

**Tabla 3.1 Análisis de rendimiento de diferentes tamaños de muestra**  
Fuente: El autor

Se realizó el análisis de rendimiento cada una de las gráficas de control realizadas anteriormente para un ARL bajo control de 400, los resultados se muestran en la Tabla 3.1. Se utilizó un tamaño de muestra de 1 y 5, con distancias  $(dp1, dp) = (1, 1.2), (1.5, 2)$  respectivamente.

Notamos que el gráfico VDT2 se comporta mejor que el gráfico estándar T2 con todas las p variables ya que el gráfico de control VDT2 reduce en el ARL fuera de control.

El gráfico DDT2 supera en rendimiento al gráfico estándar T2 para las p1 variables, lo que es notorio, ya que en esos casos el gráfico DDT2 es más rápido y más económico a la vez.

PARAMETROS				T2		VDT2				DDT2			
p1	p	dp1	dp	ARLp1	ARLp	W	LCP1	LCP	ARL1	W	LCP1	LCP	ARL1
2	3	1	1,2	71,45152	59,9999	3,83	17,09	10,62	47,6222	2,89	14,07	14,11	59,531
			1,5	71,45152	32,5174	2,86	38,77	11,65	26,8615	1,52	14,05	14,23	33,74221
			2	71,45152	12,8555	2,19	42,01	12,44	10,5593	2,02	14,03	14,19	14,92437
			2,5	71,45152	5,9008	1,88	26	12,8	4,89408	2,56	14,06	14,14	8,169503
		1,5	1,8	24,52308	18,3291	4,14	16,44	10,31	12,97705	3,23	14,06	14,08	18,00462
			2	24,52308	12,8555	3,59	18,94	10,82	9,53497	2,6	13,89	14,16	13,08638
			2,5	24,52308	5,9008	2,43	20,11	12,19	4,61765	2,57	14,05	14,14	6,535004
			3	24,52308	3,1826	2,14	53,22	12,5	2,59832	3,22	13,96	14,09	4,156582

**Tabla 3.2 Análisis de rendimiento de varios casos con tamaño de muestra uno**  
**Fuente: El autor**

También, se realizó un análisis del rendimiento varios casos, combinando diferentes valores de  $(dp1, dp) = (1, 1.2), (1, 1.5), (1, 2), (1, 2.5), (1.5, 1.8), (1.5, 2), (1.5, 2.5), (1.5, 3)$  para un tamaño de muestra de 1 y un ARL bajo control de 400.

Para todos los casos analizados el gráfico de control VDT2 se comporta mejor que el gráfico estándar T2 con todas las p variables ya que el gráfico de control VDT2 reduce en el ARL fuera de control.

El gráfico DDT2 siempre supera en rendimiento al gráfico estándar T2 para las p1 variables, e incluso en alguno de los casos analizados su rendimiento es mayor o igual al del gráfico estándar T2 para las p variables, lo que es notorio, ya que en esos casos el gráfico DDT2 es más rápido y más económico a la vez.

## CONCLUSIONES Y RECOMENDACIONES

### Conclusiones

1. Se logró realizar las funciones en el software estadístico R que permitan realizar las gráficas y el cálculo de los ARL bajo y fuera de control para cada uno de los gráficos de control analizados en este proyecto.
2. Se desarrolló el paquete Gcmulti Gráficos de control Multivariante en R
3. Se logró confirmar que el gráfico de control de doble dimensión  $DDT^2$  siempre supera el rendimiento del gráfico de control multivariante  $T^2$  de Hotelling con  $p_1$  variables y en algunos casos iguala y supera el gráfico de control  $T^2$  de Hotelling con las  $p$  variables, por lo que es notorio que este grafico es más rápido y más económico a la vez.
4. Se logró confirmar que el gráfico de control de dimensión variable  $VDT^2$  siempre supera el rendimiento de los gráficos de control multivariante  $T^2$  de Hotelling  $p_1$  y  $p$  variables.
5. Se logró confirmar que a medida que se aumenta el tamaño de la muestra aumenta el rendimiento de los gráficos de control  $VDT^2$  y  $DDT^2$ , ya que obtenemos un ARL fuera de control más bajo en los gráficos
6. Se logró publicar el paquete Gcmulti en el repositorio de GitHub y se encuentra disponible para el público en general.

### Recomendaciones

1. Realizar funciones que permita realizar la optimización de los parámetros para cada uno de los gráficos de control analizados en el proyecto.
2. Difundir el paquete a las empresas que deseen mejorar el control de sus procesos.
3. Mejorar los gráficos de control empleando la librería ggplot2 de R.
4. Incluir más gráficos de control multivariantes en el paquete Gcmulti realizado en R.

## BIBLIOGRAFÍA

- [1] Aparisi, F.; Epprecht, E.; Ruiz, O.; Veiga A., "Reducing Sampling Costs of Multivariate SPC with a Double-Dimension  $T^2$  Control Chart", International Journal of Production Research, vol. 1, no. 1, pag. 1-15, 2013.
- [2] Aparisi, F.; Epprecht, E.; Ruiz, O., " $T^2$  Control Charts with Variable Dimension", Journal of Quality Technology, vol. 44, no. 4, pp. 375-393, 2012.
- [3] Brook D. & Evans D.A., An approach to the probability distribution of cusum run length. Biometrika 59, 3, p. 539, 1972
- [4] El blog de los erreros (2013). Creación de paquetes con Rstudio [Online]. Disponible en: <http://erre-que-erre-paco.blogspot.com/2013/03/creacion-de-paquetes-con-rstudio.html>
- [5] Fuchs, C.; Kenett, R., "Multivariate Quality Control- Theory and Applications". Marcel Dekker Inc.,1998.
- [6] Hotelling, H., "Multivariate Quality Control". In C. Eisenhart, M. W. Hastay, and W. A. Wallis, eds. Techniques of Statistical Analysis. New York: McGraw-Hill, 1947.
- [7] Mahalanobis, P.C., "On the generalised distance in statistics". Proceedings of the National Institute of Science of India 12, 49-55,1936.
- [8] Marín Diazaraque J. M., (2012). Cadenas de Markov [Online]. Departamento de Estadística, Univ, Carlos III de Madrid. Disponible en: <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/PEst/tema4pe.pdf>.
- [9] Ross, S. D., "Stochastics Processes". 2th Edition. John Wiley & Son,1996
- [10] O. Perpiñan Lamigueiro (2015). Crear Paquetes en R [Online]. Disponible en: <https://oscarperpinan.github.io/intro/Paquetes.pdf>
- [11] R (2016). Writing R Extensions [Online]. Disponible en: <https://cran.r-project.org/doc/manuals/r-release/R-exts.html>

- [12] Rstudio (2015). Desarrollo de Paquetes [Online]. Disponible en:  
<https://www.rstudio.com/wp-content/uploads/2015/03/devtools-spanish.pdf>
- [13] Ruiz, O., “Gráficos de control de calidad multivariantes con dimensión variable”  
Tesis doctoral, Dept. de Estadística e Investigación Operativa, Univ. Pol. de  
Valencia, España, 2013.
- [14] Tracy, N. D.; Young, J. C.; Mason, R. L., “Multivariate Control Charts for  
Individual Observations”. *Journal of Quality Technology*, 24, 88-95, 1992.
- [15] Wikipedia (2016), Noncentral chi-squared distribution [Online]. Disponible en:  
[https://en.wikipedia.org/wiki/Noncentral\\_chi-squared\\_distribution](https://en.wikipedia.org/wiki/Noncentral_chi-squared_distribution)

## ANEXOS

## a. Tabla de datos simulados

N°	va1	va2	va3
1	3,15	6,49	7,35
2	3,65	8,59	9,83
3	5,57	7,41	9,23
4	4,48	7,01	9,17
5	5,57	8,81	9,72
6	5,86	7,83	10,40
7	4,63	7,86	9,08
8	3,83	4,43	6,46
9	4,83	9,55	9,80
10	6,36	9,20	11,22
11	5,21	9,48	10,63
12	2,04	6,07	6,05
13	4,05	6,78	7,31
14	4,55	7,73	8,46
15	6,46	8,41	10,41
16	3,39	6,80	7,86
17	2,71	6,23	6,09
18	5,97	9,60	10,29
19	7,62	8,43	11,99
20	3,04	5,93	6,26
21	3,39	9,35	8,85
22	2,80	5,57	6,17
23	5,09	9,26	10,06
24	7,95	7,64	10,71
25	5,76	9,39	11,15
26	3,83	4,70	6,00
27	3,48	4,96	6,42
28	3,63	8,70	9,07
29	4,10	8,68	8,65
30	5,33	6,70	8,37
31	4,11	8,47	9,06
32	5,52	8,05	10,31
33	3,19	6,65	7,25
34	6,78	8,99	10,78
35	5,04	7,58	9,19
36	3,46	7,81	7,87
37	4,17	6,03	8,19
38	5,71	8,06	10,17
39	5,10	6,78	8,02
40	4,25	8,03	9,00
41	5,36	6,14	7,62
42	5,33	6,29	8,11
43	2,46	2,99	4,12
44	4,09	6,39	8,00
45	4,41	7,18	8,39
46	3,26	7,05	8,34
47	3,70	7,00	8,16
48	4,38	6,09	8,49
49	5,92	6,16	8,32
50	5,88	9,09	10,91
51	3,92	9,10	8,57
52	5,32	5,59	7,70
53	3,01	6,04	6,46
54	5,44	7,58	9,31
55	2,23	4,34	5,06
56	5,48	6,46	8,21
57	5,52	8,53	10,64
58	2,82	7,01	7,88
59	4,86	8,56	9,83
60	3,92	8,34	8,65
61	5,47	8,29	10,44
62	4,88	6,92	8,98
63	3,05	5,17	5,82
64	4,88	9,45	11,13
65	5,15	7,39	8,45
66	3,43	5,06	6,35
67	5,75	9,50	11,25
68	5,58	3,91	7,16
69	5,37	6,73	9,17
70	3,49	4,76	6,04
71	5,48	6,65	8,67
72	4,48	6,55	9,57
73	3,47	8,21	8,98
74	1,58	5,92	5,11
75	2,87	5,60	6,44
76	3,25	5,77	7,04
77	3,03	8,77	8,58
78	5,90	8,25	9,42
79	3,57	5,44	6,64
80	1,79	5,79	5,18
81	4,02	9,35	9,75
82	6,97	9,45	12,34
83	3,58	6,49	6,90
84	5,40	8,55	10,23
85	3,95	6,35	6,99
86	4,43	7,56	8,28
87	4,20	7,48	9,53
88	4,07	7,66	9,51
89	5,43	4,69	8,40
90	6,09	6,64	9,25
91	5,96	9,87	11,31
92	1,44	6,85	6,27
93	5,08	5,64	7,72
94	6,35	6,91	9,38
95	5,50	7,50	9,41
96	2,76	6,44	7,04
97	3,63	6,60	6,44
98	2,02	3,61	4,60
99	4,85	5,44	7,70
100	4,67	8,72	9,22
101	3,46	6,99	7,06
102	4,14	5,78	6,92
103	3,41	5,72	7,02
104	3,67	7,46	8,67
105	4,96	8,96	9,60
106	3,20	4,46	5,26
107	5,41	9,05	10,88
108	5,26	7,73	10,10
109	7,07	8,86	11,95
110	4,75	5,51	7,77
111	6,56	7,84	10,41
112	4,38	5,16	7,00
113	5,52	7,83	9,43
114	5,73	7,22	9,27
115	3,48	5,28	6,21
116	3,84	8,41	8,74
117	3,47	5,44	6,70
118	3,70	4,25	5,53
119	1,45	3,55	2,93
120	5,48	6,28	8,73
121	3,65	7,05	8,79
122	3,45	7,36	7,62
123	5,11	5,51	7,82
124	5,17	6,68	8,97
125	5,33	6,13	8,79
126	4,24	8,10	9,10
127	6,74	8,81	11,26
128	5,77	9,47	10,23
129	5,22	6,75	9,05
130	5,11	5,83	7,94
131	4,88	7,20	9,16
132	3,54	7,32	7,95
133	4,25	5,01	6,89
134	4,99	9,02	9,57
135	6,45	9,49	11,75
136	4,52	5,58	7,36
137	5,39	6,62	8,43
138	5,79	9,57	10,95
139	4,31	9,17	9,81

140	2,99	5,97	5,98
141	4,05	6,28	7,96
142	4,79	7,57	9,21
143	4,27	7,96	9,78
144	4,17	6,12	7,66
145	4,62	6,99	8,39
146	3,62	6,37	7,19
147	4,53	7,83	8,23
148	5,35	7,62	9,40
149	5,32	7,99	10,25
150	5,03	8,49	9,74
151	4,91	5,21	7,04
152	4,41	6,78	7,63
153	4,71	8,36	10,02
154	4,65	8,88	10,08
155	6,60	9,09	10,63
156	3,18	6,18	7,94
157	4,25	7,98	8,66
158	2,91	5,52	6,48
159	4,38	5,79	7,50
160	7,61	11,34	13,75
161	5,22	6,07	8,15
162	2,51	3,22	4,15
163	6,15	6,37	8,40
164	7,24	8,98	11,45
165	4,61	6,95	8,20
166	3,45	6,81	7,45
167	2,69	6,06	6,69

168	5,09	7,24	9,52
169	6,55	6,39	9,32
170	3,76	5,50	6,33
171	4,40	5,50	7,02
172	6,09	6,36	9,28
173	2,41	3,59	4,84
174	4,81	8,45	9,08
175	3,15	8,72	8,66
176	6,79	7,74	9,80
177	2,08	4,54	4,60
178	4,68	8,85	9,46
179	4,70	5,47	7,00
180	6,31	9,75	12,29
181	4,55	6,37	7,11
182	7,34	9,02	11,71
183	2,60	6,15	6,28
184	5,35	6,49	8,23
185	1,46	7,60	6,71
186	0,81	5,07	5,21
187	3,59	8,58	8,28
188	2,92	6,31	6,77
189	3,63	4,57	6,02
190	5,87	8,10	9,78
191	3,50	4,52	6,25
192	4,60	5,20	7,50
193	4,50	6,98	8,60
194	3,72	6,02	7,35
195	5,14	8,00	9,02

196	5,74	9,23	10,49
197	3,17	6,62	6,95
198	4,56	7,66	9,65
199	6,10	6,74	8,62
200	5,14	8,67	9,41
201	5,15	6,33	7,98
202	5,71	5,43	8,27
203	3,72	6,40	7,44
204	4,18	4,77	6,67
205	4,07	5,82	6,88
206	1,99	4,24	4,39
207	7,12	9,59	13,18
208	3,56	6,49	7,80
209	4,67	5,83	7,82
210	4,68	4,86	6,63
211	5,37	7,37	9,08
212	5,65	8,84	9,91
213	2,12	7,47	7,47
214	2,15	7,43	6,77
215	3,07	6,97	7,77
216	2,68	4,49	5,54
217	5,36	7,23	9,35
218	3,92	8,62	10,33
219	3,81	6,47	7,59
220	4,83	5,46	7,69
221	6,07	7,09	9,41
222	5,87	7,04	9,11
223	4,72	7,51	8,86

224	4,95	6,88	9,21
225	5,87	10,15	11,83
226	3,90	6,87	7,50
227	2,03	4,17	5,21
228	4,54	7,41	8,66
229	3,64	6,94	7,72
230	3,07	7,42	7,70
231	5,06	6,93	9,20
232	6,45	7,41	9,43
233	1,75	4,41	5,31
234	1,87	5,63	5,42
235	4,16	6,62	7,64
236	3,53	7,28	7,74
237	1,57	4,32	3,75
238	5,52	9,10	10,53
239	4,02	6,79	8,66
240	3,00	7,05	6,78
241	6,79	9,25	12,47
242	4,26	5,09	7,46
243	2,93	2,57	4,24
244	4,06	5,34	7,10
245	6,27	8,09	10,14
246	4,38	6,31	8,36
247	3,41	7,54	8,58
248	6,36	11,17	12,43
249	6,98	8,48	11,31
250	4,98	6,56	8,37

## b. Código fuente del gráfico de control $T^2$ de Hotelling

```

# X: Conjunto de Variables muestra
#k: número de muestras
#mu:
#sigma:
#n tamaño de la muestra
#numero de muestras
grafica_T2<-function(mu,sigma,X,alfa=0.001,k){
  if (k>1){
    covarianzas<-sigma
    medias=matrix(c(rep(0,ncol(X)*l(nrow(X)/k))),ncol=ncol(X))
    for(i in 1:l(nrow(X)/k)){
      medias[i,]=colMeans(X[l(abs(2*i-k+1)):l(i*k),])
    }
    T2<-c(rep(0,nrow(medias)))
    muestras<-1:nrow(medias)
    for(i in 1:(nrow(medias))){
      T2[i]<-k*(t(medias[i,]-mu)%*%solve(covarianzas)%*(medias[i,]-mu))
    }
    T2
    LCS<-((ncol(X)*((nrow(X)/k)-1)*(k-1))/((nrow(X)/k)*k-(nrow(X)/k)-
      ncol(X)+1))*qf(1-alfa,ncol(X),(nrow(X)/k)*k-(nrow(X)/k)-ncol(X)+1)
    LCI<-0

    plot(muestras,T2,type='l',col="blue",xaxt='n',xlim=c(0,(length(T2)+18)),ylim=c(min(LCI,min(T
    2))-0.5,max(max(T2),LCS)+0.5))
    axis(1,seq(from=0,to=length(T2),by=2),cex.axis=0.8,las=2)

    abline(h=LCS,lty=2,col="brown")
    text((max(muestras)+5),LCS,paste('LCS=',round(LCS,digits = 6)),pos=3,font=2,cex=0.8)
    abline(h=LCI,lty=2,col="brown")
    text((max(muestras)+5),LCI,paste('LCI=',LCI),pos=3,font=2,cex=0.8)
    for(i in 1:length(muestras)){
      if((T2[i]<LCI)|(T2[i]>LCS)){points(muestras[i],T2[i],pch=4,col="red")}
      else{points(muestras[i],T2[i],pch=20,col="green")}
    }
    mtext('Gráfico de control T2 de Hotelling',side=3,font=2)}
  if(k==1){
    if(nrow(X)<=100){
      covarianzas<-sigma
      T2<-vector(length=nrow(X))
      muestras<-1:nrow(X)
      for(i in 1:(nrow(X))){T2[i]<-t(X[i,]-mu)%*%solve(covarianzas)%*(X[i,]-mu)}
      T2

```

```

LCS<-((ncol(X)*(nrow(X)+1)*(nrow(X)-1))/(nrow(X)*nrow(X)-
      nrow(X)*ncol(X)))*qf(1-alfa,ncol(X),nrow(X)-ncol(X))
LCI<-0
plot(muestras,T2,type='l',col="blue",xaxt='n',xlim=c(0,(length(T2)+10)),
     ylim=c(min(LCI,min(T2))-0.5,max(max(T2),LCS)+1))
axis(1,seq(from=0,to=length(T2),by=2),cex.axis=0.8,las=2)

abline(h=LCS,lty=2,col="brown")
text((max(muestras)+5),LCS,paste('LCS=',round(LCS,digits = 6)),pos=3,font=2,cex=0.8)
abline(h=LCI,lty=2,col="brown")
text((max(muestras)+5),LCI,paste('LCI=',LCI),pos=3,font=2,cex=0.8)
for(i in 1:length(muestras)){
  if((T2[i]<LCI)|(T2[i]>LCS)){ points(muestras[i],T2[i],pch=4,col="red")}
  else {points(muestras[i],T2[i],pch=20,col="green")} }
}
mtext('Gráfico de control T2 de Hotelling',side=3,font=2)
}
if(nrow(X)>100){
  covarianzas<-sigma
  T2<-vector(length=nrow(X))
  muestras<-1:nrow(X)
  for(i in 1:(nrow(X))){
    T2[i]<-t(X[i,]-mu)%*%solve(covarianzas)%*%(X[i,]-mu)
  }
  T2
  LCS<-(((nrow(X)-1)^2)/nrow(X))*qf(1-alfa,(ncol(X)/2),(nrow(X)-ncol(X)-1)/2)
  LCI<-0
  plot(muestras,T2,type='l',col="blue",xaxt='n',xlim=c(0,(length(T2)+18)),
       ylim=c(min(LCI,min(T2))-0.5,max(max(T2),LCS)+1))
  axis(1,seq(from=0,to=length(T2),by=2),cex.axis=0.8,las=2)

  abline(h=LCS,lty=2,col="brown")
  text((max(muestras)+5),LCS,paste('LCS=',round(LCS,digits = 6)),pos=3,font=2,cex=0.8)
  abline(h=LCI,lty=2,col="brown")
  text((max(muestras)+5),LCI,paste('LCI=',LCI),pos=3,font=2,cex=0.8)
  for(i in 1:length(muestras)){
    if((T2[i]<LCI)|(T2[i]>LCS)){points(muestras[i],T2[i],pch=4,col="red")}
    else {points(muestras[i],T2[i],pch=20,col="green")} }
  }
  mtext('Gráfico de control T2 de Hotelling',side=3,font=2)
}}
#CALCULADO LOS ARL
ARL0=1/alfa
media=colMeans(X)
d2<-(t(mu-media)%*%solve(sigma)%*%(mu-media))

```

```

beta=pchisq(LCS,df=ncol(X),ncp=ncol(X)*d2,lower.tail=FALSE)
ARL1=1/beta
return(list(ARLO=ARLO,ARL1=ARL1))
}

```

### c. Código fuente del grafico de control de doble dimensión DDT<sup>2</sup>

```

grafica_T2DD<-function(p1,p,w,CLP1,CLP,disp1,disp,mu,sigma,X,k,inter=0.05)
{
#####Para las p1 variables#####
if(k>1){
muestra_p1variables=X[,1:p1]
mu_p1variables=mu[1:p1]
sigma_p1variables=sigma[1:p1,1:p1]
medias_p1variables=matrix(c(rep(0,p1*(nrow(X)/k))),ncol=p1)
for(i in 1:(nrow(muestra_p1variables)/k)){
medias_p1variables[i,]=colMeans(muestra_p1variables[(i*k):(i*k+1),])
}
T2_p1variables<-c(rep(0,nrow(medias_p1variables)))
muestras_p1variables<-1:nrow(medias_p1variables)
for(i in 1:(nrow(medias_p1variables))){
T2_p1variables[i]<-k*(t(medias_p1variables[i,]-
mu_p1variables)%*%solve(sigma_p1variables)%*%(medias_p1variables[i,]-
mu_p1variables))
}
#####para las p variables
covarianzas<-sigma
medias=matrix(c(rep(0,ncol(X)*I(nrow(X)/k))),ncol=ncol(X))
for(i in 1:(nrow(X)/k)){
medias[i,]=colMeans(X[(i*k):(i*k+1),])
}
T2<-c(rep(0,nrow(medias)))
muestras<-1:nrow(medias)
for(i in 1:(nrow(medias))){
T2[i]<-k*(t(medias[i,]-mu)%*%solve(covarianzas)%*%(medias[i,]-mu))
}
#####T2 para graficar#####33
t2p1=T2_p1variables
t2p=T2
t2=NULL;lb=NULL
cuenta=0;po=NULL
for (i in 1:length(t2p1)){
if(t2p1[i]>w){
cuenta=cuenta+1
}
}
}

```

```

    po=c(po,i)}}

tem=NULL
for ( i in po){
  tem=c(tem,t2p[i])
}
v1t2=c(1:length(t2p1),po)
v2t2=c(t2p1,tem)
mt2=matrix(c(v1t2,v2t2),ncol=2)
mt2=mt2[order(mt2[,1]),]
muestras=mt2[,1]
T2=mt2[,2]

lb[1]="Tp1"
for(i in 1:(length(muestras)-1)){
if(muestras[i]==muestras[i+1]){
  lb[i+1]="Tp"}
else{lb[i+1]="Tp1"}
}
##### Graficando
LCS<-CLP
LCI<-CLP1
W<-w

plot(muestras,T2,type='l',col="blue",xaxt='n',xlim=c(0,(nrow(X)/k)+7),ylim=c(min(W,min(T2)
)-0.5,max(max(T2),LCS)+0.5))
axis(1,seq(from=0,to=length(T2),by=2),cex.axis=0.8,las=2)
abline(h=LCS,lty=2,col="brown")
text((max(muestras)+5),LCS,paste('LCP=',LCS),pos=3,font=2,cex=0.6)
abline(h=LCI,lty=2,col="brown")
text(1,LCI,paste('LCP1=',LCI),pos=3,font=2,cex=0.6)
abline(h=W,lty=2,col="brown")
text((max(muestras)+5),W,paste('W=',W),pos=3,font=2,cex=0.6)
for(i in 1:length(lb)){
  if(lb[i]=="Tp1"){ points(muestras[i],T2[i],pch=20,col="green")}
  else{points(muestras[i],T2[i],pch=20,col="yellow")}
}
for(i in 1:length(muestras)){
  if(T2[i]>=LCS){points(muestras[i],T2[i],pch=4,col="red")}
}

#library calibrate
library(calibrate)
textxy(muestras,T2,lb)
mtext('Gráfico de control T2 con Doble Dimensión',side=3,font=2)}

```

```

if(k==1){
  covarianzas<-sigma
  T2<-vector(length=nrow(X))
  muestras<-1:nrow(X)

  #####Para las p1 variables#####
  muestra_p1variables=X[,1:p1]
  mu_p1variables=mu[1:p1]
  sigma_p1variables=sigma[1:p1,1:p1]
  T2_p1variables<-c(rep(0,nrow(muestra_p1variables)))
  muestras_p1variables<-1:nrow(muestra_p1variables)
  for(i in 1:(nrow(muestra_p1variables))){
    T2_p1variables[i]<-k*(t(muestra_p1variables[i,-
mu_p1variables]%%solve(sigma_p1variables)%%(muestra_p1variables[i,-
mu_p1variables))
  }
  #####para las p variables
  T2<-c(rep(0,nrow(X)))
  for(i in 1:(nrow(X))){
    T2[i]<-k*(t(X[i,]-mu)%%solve(covarianzas)%%(X[i,]-mu))}
  #####T2 para graficar#####33
  t2p1=T2_p1variables
  t2p=T2
  t2=NULL;lb=NULL
  cuenta=0;po=NULL
  for (i in 1:length(t2p1)){
    if(t2p1[i]>w){
      cuenta=cuenta+1
      po=c(po,i)}}

  tem=NULL
  for ( i in po){
    tem=c(tem,t2p[i])
  }
  v1t2=c(1:length(t2p1),po)
  v2t2=c(t2p1,tem)
  mt2=matrix(c(v1t2,v2t2),ncol=2)
  mt2=mt2[order(mt2[,1]),]
  muestras=mt2[,1]
  T2=mt2[,2]

  lb[1]="Tp1"
  for(i in 1:(length(muestras)-1)){
    if(muestras[i]==muestras[i+1]){
      lb[i+1]="Tp"}
  }

```

```

    else{lb[i+1]="Tp1"}
  }
#####Graficando k=1
LCS<-CLP
LCI<-CLP1
W<-w

plot(muestras,T2,type='l',col="blue",xaxt='n',xlim=c(0,(nrow(X)/k)+7),ylim=c(min(LCI,min(T2
))-0.5,max(max(T2),LCS)+5))
axis(1,seq(from=0,to=length(T2),by=2),cex.axis=0.8,las=2)
abline(h=LCS,lty=2,col="brown")
text((max(muestras)+5),LCS,paste('LCP=',LCS),pos=3,font=2,cex=0.6)
abline(h=LCI,lty=2,col="brown")
text(1,LCI,paste('LCP1=',LCI),pos=3,font=2,cex=0.6)
abline(h=W,lty=2,col="brown")
text((max(muestras)+5),W,paste('W=',W),pos=3,font=2,cex=0.6)
for(i in 1:length(lb)){
  if(lb[i]=="Tp1"){ points(muestras[i],T2[i],pch=20,col="green")}
  else{points(muestras[i],T2[i],pch=20,col="yellow")}
}
for(i in 1:length(muestras)){
  if(T2[i]>=LCS){points(muestras[i],T2[i],pch=4,col="red")}
}
#library calibrate
library(calibrate)
textxy(muestras,T2,lb)
mtext('Gráfico de control T2 con Doble Dimensión',side=3,font=2)}

##### Calculando ARL
#(p1,p,w,CLP1,CLP,disp1,disp,mu,sigma,X,k,inter=0.05)
ARLDDT2=function(p1, p, n, iw, iCLP1, iCLP, ides1, ides2){
#-----ARL DDT-----
INT1=function(v, iCLP,lamb1,lamb2,p,p1){
  arriba = (iCLP-v)
  if(arriba<0){
    resul=0}
  else{
    if(lamb2==0){
      resul=pchisq(arriba,p-p1)}
    else{
      if(l(p-p1)==1 | p1==1){
        resul=pnorm(sqrt(arriba)-sqrt(lamb2-lamb1))- pnorm(-sqrt(arriba)-sqrt(lamb2-
lamb1))}
      else{dchisq(arriba, p-p1, lamb2-lamb1) - 0}
    }
  }
}

```

```

}
return(resul)
}

#-----
interv= function(iCLP1, iW){
  if(lamb2==0){
    inter=0.0001}
  else{
    if (iCLP1<5.5) {inter=0.2}
    if ((iCLP1>=5.5) & (iCLP1<11)){inter= 0.0129*iCLP1*iCLP1 - 0.2378*iCLP1 + 1.1241}
    if ((iCLP1>=11) & (iCLP1<18)) {inter= -0.0009*iCLP1*iCLP1*iCLP1 + 0.0393*iCLP1*iCLP1 -
0.5987*iCLP1 + 3.0314}
    if (iCLP1>=18){ inter=0.005}
    if (p1==1){
      if ((ides1<1) & (ides2<1)){ inter=0.0001}}
    else{ inter= 0.0005}
  }
  if (inter<=0){ inter=0.001}
  res2=inter
  return(res2)
}

#-----
lamb1 = n*ides1*ides1; lamb2 = n*ides2*ides2;
pnsignal=0; part1=0
if (p1==1){
  part1=pnorm(sqrt(iw)-sqrt(lamb1))-pnorm(-sqrt(iw)-sqrt(lamb1))}
else{
  if (lamb2>0){ part1 = pchisq(iw, p1, lamb1)}
  else{part1= pchisq(iw,p1)}
  inter=interv(iCLP,iw)}
xi=iw
repeat{
  if (lamb2>0) {
    alto = INT1(xi, iCLP, lamb1,lamb2,p,p1)* dchisq(xi, p1, lamb1)}
  else{
    alto = INT1(xi, iCLP, lamb1,lamb2,p,p1)* dchisq(xi,p1)}
  pnsignal = pnsignal + alto*inter
  xi = xi + inter
  if(xi>iCLP1)break} #(iCLP1 - xi) < inter

pnsignal = pnsignal + part1
ARL = 1 / (1- pnsignal)
return(ARL)

```

```

}
ARL_DDT2=ARLDDT2(p1=p1,p=p,n=k,iw=w,iCLP1=CLP1,iCLP=CLP,ides1=disp1,ides2=disp)
return(ARL_DDT2)
}

```

#### d. Código fuente del grafico de control de dimensión variable VDT<sup>2</sup>

```

#p1: numero de variables p1
#p: numero de variables p ==> p=p1+p2
#w: limite de alerta
#CLp1: limite de control p1
#CLp: limite de control p
#disp1: dist mahalanobis p1
#disp: dist mahalanobis p
#Problnic_p1: Probabilidad Inicial p1=1 ó p1=0
#mu: vector promedios cuando el proceso esta bajo control
#sigma: matriz covarianzas cuando el proceso esta bajo control
#X: la muestra
#n/k: tamaño de la muestra

#### La muestra X debe estar ordenada [p1,p] primero las variables p1: faciles de medir ; y
luego las p2: complejas de medir

grafica_T2DV<-funcion(p1,p,w,CLp1,CLp,disp1,disp,mu,sigma,X,k,Problnic_p1=1){
  if(k>1){
    #####Para las p1 variables#####
    muestra_p1variables=X[,1:p1]
    mu_p1variables=mu[1:p1]
    sigma_p1variables=sigma[1:p1,1:p1]
    medias_p1variables=matrix(c(rep(0,p1*(nrow(X)/k))),ncol=p1)
    for(i in 1:(nrow(muestra_p1variables)/k)){
      medias_p1variables[i,]=colMeans(muestra_p1variables[(abs(2*i-k+1)):i*k,])
    }
    T2_p1variables<-c(rep(0,nrow(medias_p1variables)))
    muestras_p1variables<-1:nrow(medias_p1variables)
    for(i in 1:(nrow(medias_p1variables))){
      T2_p1variables[i]<-k*(t(medias_p1variables[i,]-
mu_p1variables)%%solve(sigma_p1variables)%%(medias_p1variables[i,]-
mu_p1variables))
    }
    #####para las p variables
    covarianzas<-sigma
    medias=matrix(c(rep(0,ncol(X)*i*(nrow(X)/k))),ncol=ncol(X))
    for(i in 1:(nrow(X)/k)){

```

```

    medias[i,]=colMeans(X[l(abs(2*i-k+1)):l(i*k),])
  }
  T2<-c(rep(0,nrow(medias)))
  muestras<-1:nrow(medias)
  for(i in 1:(nrow(medias))){
    T2[i]<-k*(t(medias[i,]-mu)%*%solve(covarianzas)%*(medias[i,]-mu))
  }
  #####T2 para graficar#####33
  t2p1=T2_p1variables
  t2p=T2
  t2=rep(0,length(t2p));lb=NULL
  i=1
  t2[1]=t2p1[1];lb[1]="Tp1"
  repeat{
    if(t2[i]>=w){
      t2[i+1]=t2p[i+1]
      lb[i+1]="Tp"
    }else{t2[i+1]=t2p1[i+1]
      lb[i+1]="Tp1"}
    i=i+1
    if(i >=length(t2))break}
  #####Graficando para k>1#####
  T2=t2
  LCS<-CLp
  LCI<-CLp1
  W<-w

plot(muestras,T2,type='l',col="blue",xaxt='n',xlim=c(0,(length(T2)+7)),ylim=c(min(W,min(T2)
)-0.5,max(max(T2),LCS,LCI)+0.5))
axis(1,seq(from=0,to=length(T2),by=2),cex.axis=0.8,las=2)
abline(h=LCS,lty=2,col="brown")
text((max(muestras)+2),LCS,paste('LCP=',LCS),pos=3,font=2,cex=0.6)
abline(h=LCI,lty=2,col="brown")
text((max(muestras)+6),LCI,paste('LCP1=',LCI),pos=3,font=2,cex=0.6)
abline(h=W,lty=2,col="brown")
text((max(muestras)+6),W,paste('W=',W),pos=3,font=2,cex=0.6)

for(i in 1:length(lb)){
  if(lb[i]=="Tp1"){ points(muestras[i],T2[i],pch=20,col="green")}
  else{points(muestras[i],T2[i],pch=20,col="yellow")}
}
for(i in 1:length(muestras)){
  if(T2[i]>=LCS & T2[i]>=LCI){points(muestras[i],T2[i],pch=4,col="red")}
}

```

```

#library calibrate
library(calibrate)
textxy(muestras,T2,lb)
mtext('Gráfico de control T2 con Dimensión Variable',side=3,font=2)}

if(k==1){
  covarianzas<-sigma
  T2<-vector(length=nrow(X))
  muestras<-1:nrow(X)

  #####Para las p1 variables#####
  muestra_p1variables=X[,1:p1]
  mu_p1variables=mu[1:p1]
  sigma_p1variables=sigma[1:p1,1:p1]
  T2_p1variables<-c(rep(0,nrow(muestra_p1variables)))
  muestras_p1variables<-1:nrow(muestra_p1variables)
  for(i in 1:(nrow(muestra_p1variables))){
    T2_p1variables[i]<-k*(t(muestra_p1variables[i,-
mu_p1variables)%*%solve(sigma_p1variables)%*%(muestra_p1variables[i,-
mu_p1variables))
  }
  #####para las p variables
  T2<-c(rep(0,nrow(X)))
  for(i in 1:(nrow(X))){
    T2[i]<-k*(t(X[i,]-mu)%*%solve(covarianzas)%*%(X[i,]-mu))}
  #####T2 para graficar#####
  t2p1=T2_p1variables
  t2p=T2
  t2=rep(0,length(t2p));lb=NULL
  i=1
  t2[1]=t2p1[1];lb[1]="Tp1"
  repeat{
    if(t2[i]>=w){
      t2[i+1]=t2p[i+1]
      lb[i+1]="Tp"}
    else{t2[i+1]=t2p1[i+1]
      lb[i+1]="Tp1"}
    i=i+1
    if(i >=length(t2))break}
  #####Graficando k=1
  T2=t2
  LCS<-CLp
  LCI<-CLp1
  W<-w

```

```

plot(muestras,T2,type='l',col="blue",xaxt='n',xlim=c(0,(length(T2)+7)),ylim=c(min(LCI,min(T2
))-0.5,max(max(T2),LCS,LCI)+0.5))
axis(1,seq(from=0,to=length(T2),by=2),cex.axis=0.8,las=2)
abline(h=LCS,lty=2,col="brown")
text((max(muestras)+2),LCS,paste('LCP=',LCS),pos=3,font=2,cex=0.6)
abline(h=LCI,lty=2,col="brown")
text((max(muestras)+6),LCI,paste('LCP1=',LCI),pos=3,font=2,cex=0.6)
abline(h=W,lty=2,col="brown")
text((max(muestras)+6),W,paste('W=',W),pos=3,font=2,cex=0.6)

for(i in 1:length(lb)){
  if(lb[i]=="Tp1"){ points(muestras[i],T2[i],pch=20,col="green")}
  else{points(muestras[i],T2[i],pch=20,col="yellow")}
}
for(i in 1:length(muestras)){
  if(T2[i]>=LCS&T2[i]>=LCI){points(muestras[i],T2[i],pch=4,col="red")}
}

#library calibrate
library(calibrate)
textxy(muestras,T2,lb)
mtext('Gráfico de control T2 con Dimensión Variable',side=3,font=2)
}

##Calculo del ARL
n=k
lamb1 = n*disp1*disp1; lamb2 = n*disp*disp
a= pchisq( w, p1, lamb1)
b= pchisq( CLp1, p1, lamb1) - a
c= pchisq( w, p, lamb2)
d= pchisq( CLp, p, lamb2) - c
if(ProbInic_p1==0){
  b1=1;b2=0}
else{b1=0;b2=1-b1}

##Calculo de Probabilidades para Qo
a0= pchisq(w, p1, 0)
b0= pchisq(CLp1,p1,0)-a0
c0= pchisq(w, p, 0)
d0= pchisq(CLp,p,0)-c0
##multiplicacion del vector B por la Inversa de I-Qo
k=1/(1-d0-a0+(a0*d0)-(b0*c0)) #determinante de I-Qo
s11=(b1*(1-d0)+b2*c)*k
s12=(b1*b0+b2*(1-a0))*k

```

```

k=s11+s12 #calculo ARLo utilizando la informacion anterior
s11=s11/k #calculo el vector S
s12=s12/k
#calculo ARLss(d)
deno = a+d-(a*d)+(b*c)-1
num = s12*(a-1) + (s11*(d-1)) - (b*s11) - (s12*c)
ARL = num / deno
return(ARL)
}

```

Código fuente de la simulación de variables aleatorias

```

mu=c(4.5,7,8.45)
sigma=matrix(c(2,1.5,2.4,1.5,3,3.1,2.4,3.1,4),ncol=3)
#la matriz de covarianzas debe ser definida positiva
eigen(sigma)
library(MASS)
dmulti=mvrnorm(250,mu,sigma)
colnames(dmulti)=c("va1","va2","va3")
View(dmulti)
#Datos a utilizar
#n1---->50 dmulti[1:50,]
#n2---->100 dmulti[1:100,]
#n5---->250 dmulti
dmulti_n1=dmulti[1:50,]
dmulti_n2=dmulti[1:100,]
#vector de medias bajo control
mu0=c(5.4,6.8,8.5)
Código fuente de las gráficas realizadas en el proyecto
#uso de la función T2 de Hotelling
#para p1
grafica_T2(mu0[1:2],sigma[1:2,1:2],dmulti_n1[,1:2],k=1,LCS=11.983,d=1) #n=1
grafica_T2(mu0[1:2],sigma[1:2,1:2],dmulti[,1:2],k=5,LCS=11.983,d=1.5) #n=5
#para p
grafica_T2(mu0,sigma,dmulti_n1,k=1,LCS=14.321,d=1.2)#n=1
grafica_T2(mu0,sigma,dmulti,k=5,LCS=14.321,d=2)#n=5
#uso de la función T2 con dimensión variable
grafica_T2DV(p1=2,p=3,w=3.83,CLp1=17.09,CLp=10.62,disp1=1,disp=1.2,mu=mu0,sigma=sigma,X=dmulti_n1,k=1)
grafica_T2DV(p1=2,p=3,w=3.59,CLp1=18.94,CLp=10.82,disp1=1.5,disp=2,mu=mu0,sigma=sigma,X=dmulti,k=5)
#Uso de la función T2 con doble dimensión
grafica_T2DD(p1=2,p=3,w=2.89,CLP1=14.07,CLP=14.11,disp1=1,disp=1.2,mu=mu0,sigma=sigma,X=dmulti_n1,k=1)

```

grafica\_T2DD(p1=2,p=3,w=2.6,CLP1=13.89,CLP=14.16,disp1=1.5,disp=2,mu=mu0,sigma=sigma,X=dmulti,k=5)

### e. DESCRIPTION del paquete Gcmulti



```
DESCRIPTION *
1 Package: Gcmulti
2 Type: Package
3 Title: Graficos de control Multivariantes
4 Version: 1.0
5 Date: 2016-08-29
6 Author: Jhonny Barahona
7 Maintainer: Jhonny Barahona <jfbaraho@espol.edu.ec>
8 Description: Permite realizar los graficos de control multivariante T2 de
9 Hotelling, de doble dimension y de dimension variable.
10 License: GPL (>=2)
11 Imports: calibrate (>= 1.7.2), graphics (>= 3.1.3), stats (>= 3.1.3)
12 Suggests:
13 MASS
14 URL: https://github.com/jhonnybarahona
15 Encoding: UTF-8
16 RoxygenNote: 5.0.1
17

1:1 DCF ↕
```