Control Estadístico de Calidad

Entrega 1. 2024-II

Michel Mendivenson Barragán Zabala
Departamento de Estadística
Universidad Nacional de Colombia

mbarraganz@unal.edu.co

Juan Sebastián Huertas Pirajan
Departamento de Estadística
Universidad Nacional de Colombia

juhuertasp@unal.edu.co

Diego Andres Paez Molina
Departamento de Estadística
Universidad Nacional de Colombia

dpaezm@unal.edu.co

19 de marzo de 2024

Ejercicio 1: Sea $X \sim N(\mu, \sigma)$ una característica de calidad. Mediante simulaciones, establezca el comportamiento del ARL (en control y fuera de él) de las Cartas R y S para observaciones normales con límites 3σ y muestras de tamaño (a) $n = 3$ y (b) $n = 10$ ¿Qué regularidades observa?

Para la implementación de la solución, se creará en R una función que nos permita simular cuantás veces querramos el momento en que un proceso da una alerta (bien sea verdadera o falsa) con argumentos que nos permitan modificar tanto el tamaño de muestra n como los límites de la carta de control y su línea central para cada una de las cartas. Las funciones se definen como sigue:

```r
RunLengthS = function(mu = 0, sigma = 1, CorrimientoSigma = 1,n = 3, m = 1000){
  S = c() ; c4 = sqrt(2/(n-1)) * gamma(n/2) / gamma((n-1)/2); CLs = c4 * sigma;
  LCL = sigma * (c4 - 3 * sqrt(1 - c4**2)); UCL = sigma * (c4 + 3 * sqrt(1 - c4**2))

  pb = txtProgressBar(min = 0, max = m, style = 3) # Barra de progreso

  i = 0
  while (length(S) < m){
    s = sd(rnorm(n, mu, CorrimientoSigma))
    i = i + 1
    if (s < LCL | s > UCL){
      S = c(S, i)
      i  = 0; setTxtProgressBar(pb, length(S))
    }}
  return(S)}

RunLengthR = function(mu = 0, sigma = 1, CorrimientoSigma = 1,n = 3, m = 1000){
  # Constantes carta R
  d3 = c(0.853, 0.888, 0.880, 0.864, 0.848, 0.833, 0.820, 0.808, 0.797, 0.787, 0.778, 0.770,
         0.763, 0.756, 0.750, 0.744, 0.739, 0.734, 0.729, 0.724, 0.72, 0.716, 0.712,0.708)
  d2 = c(1.128, 1.693, 2.059, 2.326, 2.534, 2.704, 2.847, 2.970, 3.078, 3.173, 3.258, 3.336,
         3.407, 3.472, 3.532, 3.588, 3.640, 3.689, 3.735, 3.778, 3.819, 3.858, 3.895, 3.931)
  d3 = d3[n-1]; d2 = d2[n-1]; R = c(); UCL = (d2 + 3 * d3) * sigma; LCL = (d2 - 3 * d3) * sigma

  pb = txtProgressBar(min = 0, max = m, style = 3) # Barra de progreso
```

```
i = 0
while(length(R) < m){
  r = diff(range(rnorm(n, mu, CorrimientoSigma)))
  i = i + 1
  if (r < LCL | r > UCL){
    R = c(R, i)
    i = 0; setTxtProgressBar(pb, length(R))
  }}
return(R)}
```

Tenga en cuenta que la salida de la función es un vector con los valores de los tiempos en que se detecto una señal dados los límites y la línea central correspondientes al proceso en control (El proceso en control se definió con $\mu = 0$, $\sigma = 1$ y además para cada uno se tomaron m = 1000 muestras de tiempos en que se generó una alerta). Los resultados para diferentes corrimientos se grafican a continuación:

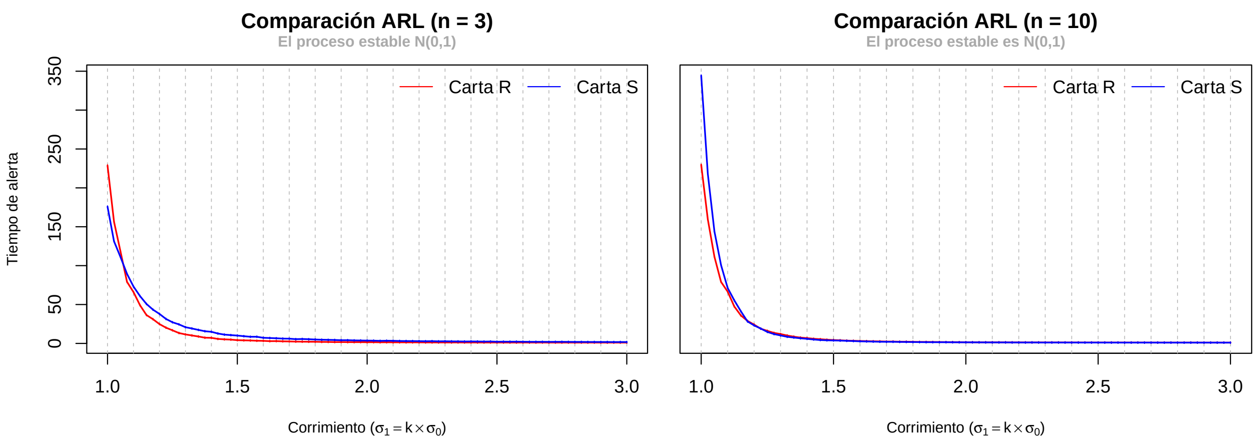

Figura 1: Comparación de ARL entre cartas R y S para n = 3, 10.

Se tomaron 81 corrimientos a intervalos regulares desde 1 hasta 3 por lo que una tabla no sería útil para analizar la información. Para $n = 3$ a la carta R le toma, en promedio, más tiempo para producir falsas alarmas para un proceso estable mientras que para los corrimientos de $1{,}1 \times \sigma$ en adelante, la carta R también tiende a tardar menos en dar una alarma verdadera. Por otro lado, la carta S para un tamaño de subgrupos racionales de 10 tarda en promedio más tiempo que la carta R para dar falsas alarmas y en cuanto a alarmas verdaderas no parece diferenciarse demasiado respecto a la carta R.

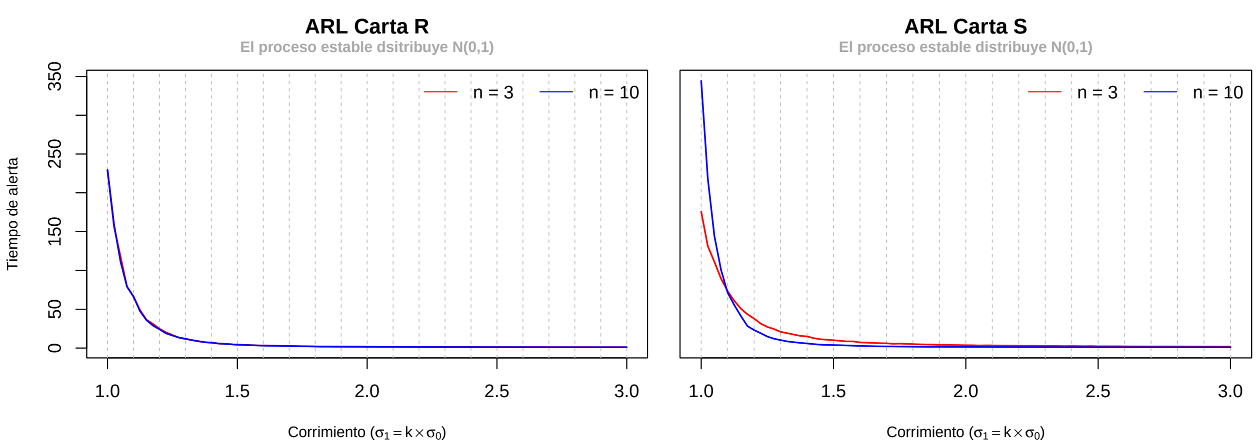

Figura 2: Comparación de ARL para cada carta con n = 3, 10

Además vemos que la carta R tiene un comportamiento más estable respecto al cambio del tamaño de los subgrupos racionales mientras que S tiende a comportarse mejor a medida que este aumenta debido a, en promedio, presentar tanto una menor frecuencia de falsas alarmas como mayor frecuencia de alarmas verdaderas.

Ejercicio 2: Sea $X \sim N(\mu, \sigma)$ una característica de calidad. Se sabe que los valores objetivo de los parámetros del proceso son $\mu = \mu_0$ y $\sigma = \sigma_0$. Construir las curvas OC de la carta S^2 con límites de probabilidad. Interpretar los resultados

Recordemos que los límites de probabilidad de una carta de control S^2 se calculan usando la distribución chi-cuadrada ($UCL = \frac{\sigma_0^2}{n-1}\chi^2_{(1-\alpha/2;\ n-1)}$ y $LCL = \frac{\sigma_0^2}{n-1}\chi^2_{(\alpha/2;\ n-1)}$). Lo que nos interesa calcular es la probabilidad $P(LCL < S^2 < UCL|\sigma^2 = \sigma_1^2)$ y como la variable aleatoria definida como $\frac{(n-1)S^2}{\sigma^2}$ tiene distribución χ^2 con $n-1$ grados de libertad podemos calcular la probai¿bilidad así:

$$
\begin{aligned}
P\left(LCL < \frac{(n-1)S^2}{\sigma^2} < UCL|\sigma^2 = \sigma_1^2\right) &= P\left(LCL < \frac{(n-1)S^2}{\sigma_1^2} < UCL\right) \\
&= P\left(\frac{(n-1)LCL}{\sigma_1^2} < S^2 < \frac{(n-1)UCL}{\sigma_1^2}\right) \\
&= P\left(\chi^2_{(n-1)} < \frac{(n-1)UCL}{\sigma_1^2}\right) - P\left(\chi^2_{(n-1)} < \frac{(n-1)LCL}{\sigma_1^2}\right)
\end{aligned}
$$

Y calculando estas probabilidades podremos graficar las curvas OC, con el fin de realizar esta operación de forma más sencilla se crea la siguiente función en R:

```
OCs2 = function(n = 5, sigma2_0 = 1, corrimiento = 1){
  UCL = sigma2_0/(n-1) * qchisq(0.975, df = n - 1)
  LCL = sigma2_0/(n-1) * qchisq(0.025, df = n - 1)
  beta = pchisq((n-1) * UCL / corrimiento , df = n - 1)
  - pchisq((n-1) * LCL / corrimiento , df = n - 1)
  return(beta)
}
```

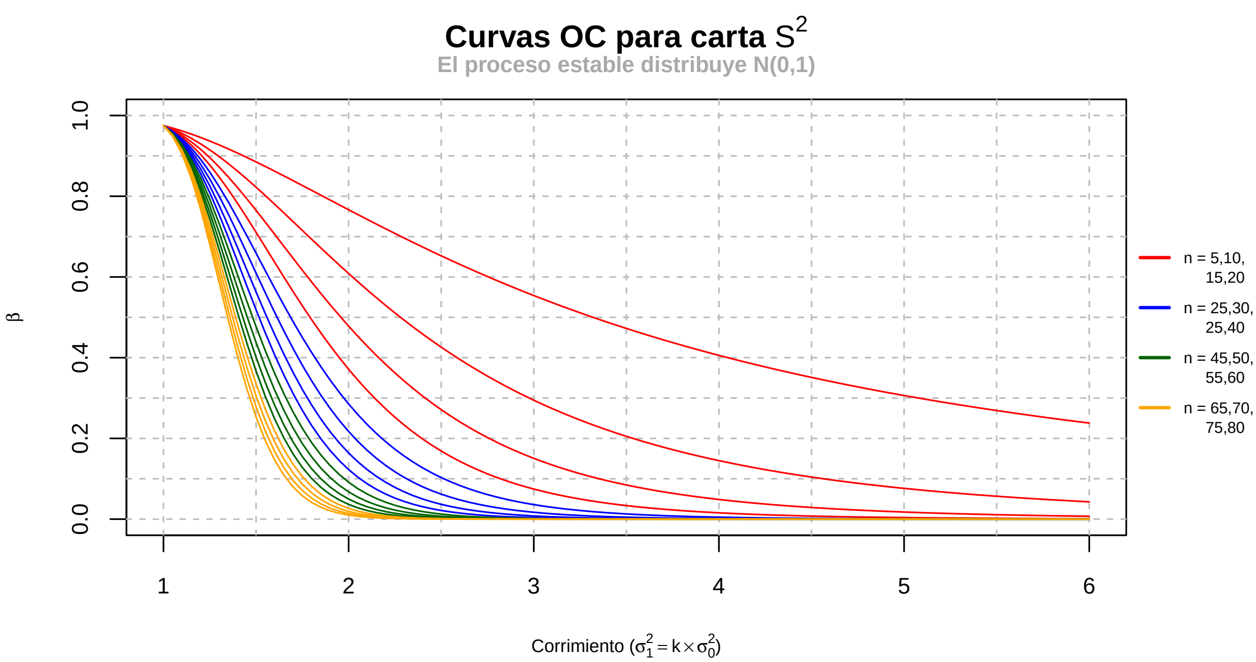

Figura 3: Curvas de operación característica para carta S²

Note que el incremento en el tamaño de los subgrupos racionales se realizo en múltiplos de 5 a cambio de obtener un gráfico más entendible y útil. Respecto al gráfico como tal, vemos que la carta

S^2 es más efectiva para muestras lo suficientemente grandes y además de esto, vemos también que a partir de tamaños de alrededor de 50 no es realmente productivo intentar aumentar n pues la relación costo-beneficio podría no ser óptima. Finalmente, revisando el gráfico nos damos cuenta que para corrimientos $\sigma_1^2 = 1{,}5\ \sigma_0^2$ se obtiene una probabilidad de detectar el cambio en el siguiente subgrupo racional de más del 50 % para $n > 15$ por lo que sería recomensable usar esta carta para procesos en que se permitan tomar más de 15 muestras.

Ejercicio 3: Sea $X \sim N(\mu_0, \sigma_0)$ una característica de calidad. Construya la carta $\bar{X}$ para el monitoreo de la media del proceso. Genere 10 muestras de tamaño n provenientes de X, de tal modo que la media muestral de ninguna de ellas caiga fuera de los límites de control. A partir del undécimo momento de monitoreo se pide generar muestras del mismo tamaño n provenientes de una distribución normal con media $\mu_1 = \mu_0 + k\sigma_0$ y $\sigma_1 = \sigma_0$ (con $k = 1,0$) hasta que la carta emita una señal por primera vez. Si se asume que el proceso caracterizado por X es estable y que se desconoce el momento en el cual se produjo el incremento en el nivel medio, ¿en qué muestra ocurrió el cambio en la media del proceso más probablemente?

Tal como se pide en el ejercicio, se plantea el siguiente algoritmo en R para k = 1.

```r
# Parámetros
mu_0 <- 0; sigma_0 <- 1; k <- 1; n <- 5
LCL <- mu_0 - 3 * (sigma_0 / sqrt(n)); UCL <- mu_0 + 3 * (sigma_0 / sqrt(n))

# Muestras
generar_muestra <- function(mu, sigma, size) {
  return(rnorm(size, mean = mu, sd = sigma))
}

# Primeras 10 muestras
muestras <- list()
for (i in 1:10) {
  muestra <- generar_muestra(mu_0, sigma_0, n)
  media_muestra <- mean(muestra)
  muestras[[i]] <- muestra
}


#  muestra num11 en adelante con cambio en la media del proceso
muestra_num <- 11
set.seed(13)
while (TRUE) {
  muestra <- generar_muestra(mu_0 + k * sigma_0, sigma_0, n)
  media_muestra <- mean(muestra)
  muestras[[muestra_num]] <- muestra
  if (media_muestra < LCL || media_muestra > UCL) {
    cat("Señal fuera de los límites de control en la muestra", muestra_num, "\n")
    break
  }
  muestra_num <- muestra_num + 1
}
```

```
## Señal fuera de los límites de control en la muestra 11
```

```
## La carta dio señal en el tiempo 11
```

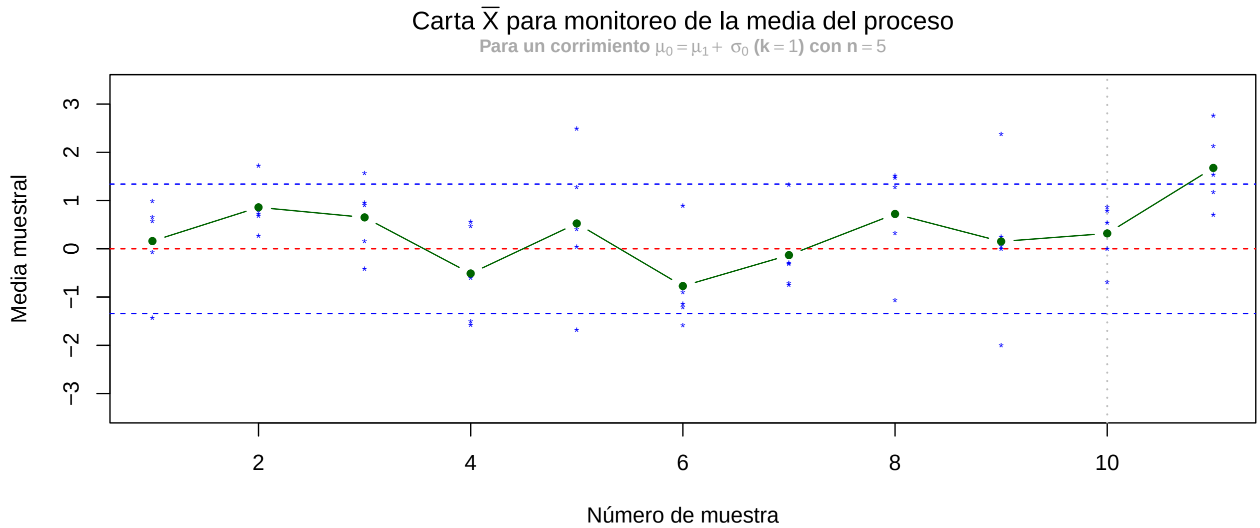

Figura 4: Carta X barra (Con k = 1)

Y para $k = 2$ simplemente se cambia este valor en los parámetros establecidos al inicio del algoritmo. Obteniendo la siguiente gráfica

```
## Señal fuera de los límites de control en la muestra 45

## La carta dio señal en el tiempo 45
```

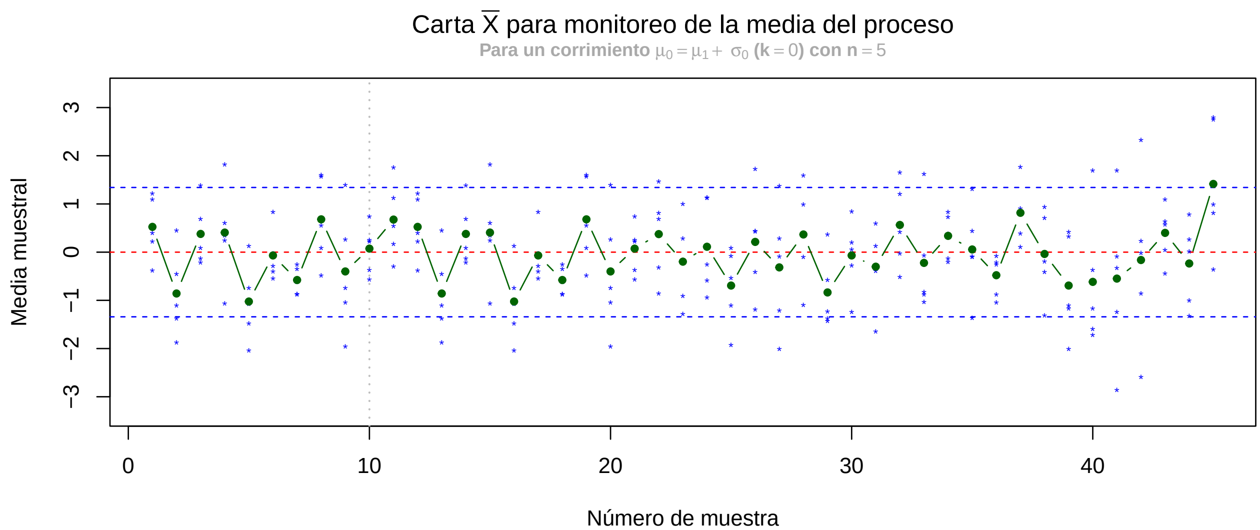

Figura 5: Carta X barra (Con k = 0))

Finalmente el momento en que más probablemente se dio el cambio en la media resultará del estimador

$\hat{\tau}_{MV} = \underset{0 \leq t < T}{\text{argmax}} \left[(T-t)(\bar{\bar{X}}_{T,t} - \mu_0)^2 \right]$ implementado en R así:

```
# Para k = 1
Total = length(medias)
FV = c()
for (i in 0:(Total - 1)){
  media_no_control = sum(medias[(i+1):Total])/(Total - i)
```

```
  FV = c(FV,log((Total-i) * (media_no_control - mu_0)^2))
}
```

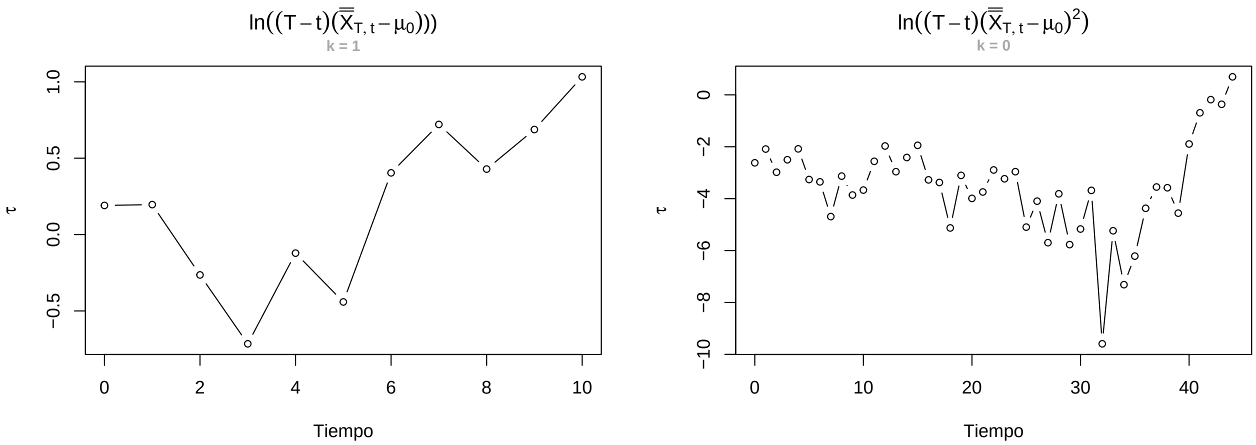

Figura 6: Gráfica función de verosimilitud para punto de cambio (k = 1,0)

Vemos que la función para valores de k pequeños tiende a ser maximizada más atrás de la última media registrada en la carta, pero para valores de k más grandes tiende a maximizarse en los últimos valores de la carta.

Ejercicio 4: Sea $X \sim N(\mu_0, \sigma_0)$ una característica de calidad. Se pide:

a. Mediante simulaciones, establezca el comportamiento del ARL de la Carta $\bar{X}$ con límites tres sigma para observaciones normales.
b. Genere 20 subgrupos racionales de tamaño $n = 3$ provenientes de X. Asúmase que el proceso es estable en cuanto a dispersión y con los subgrupos iniciales, construya la carta $\bar{X}$ como es habitual hasta verificar la estabilidad del proceso. Establezca el comportamiento del ARL para la carta que se obtiene del análisis de Fase I realizado.
c. Repetir lo indicado en el literal (b) con 50 subgrupos racionales de tamaño $n = 3$. Comente los resultados.

Ejercicio 5: Calcular el ARL de la Carta $\bar{X}$ mediante cadenas de Markov. Diseñar la carta con límites de control ubicados a tres desviaciones estándar de la media y dividiendo la región de control estadístico en franjas de ancho igual a una desviación estándar.