# Metodología Six Sigma

Herramientas para la fase de medición: Sistema de análisis de medición (continuación)

*Profesor: Gustavo Ahumada*

## Evaluando los sistemas de medición

Para que un sistema de medición sea preciso, la contribución de la variabilidad R&R del instrumento a la variabilidad total debe ser inferior al 10%. Un valor entre 10 y 30% puede ser aceptable. Un valor superior al 30% representa un mal sistema de medición.

Para evaluar este hecho, empleamos la raíz cuadrada de las variabilidades, eso es, la desviación estándar. La contribución de cada fuente de variación es estonces calculada sobre la varianza total del estudio:

$\% \quad Variacion \quad medición = \frac{\sigma medicion}{\sigma Total}$.

Otra medición útil es el número de categorías distintas. Es calculada empleando la siguiente formula:

$Numero \quad categorias =$ ($\sigma$para a parte / $\sigma$medición)$\times$1.41

y redondeando al entero inferior más cercano (1 si es inferior a 1). El número de categorías distintas debe ser mayor o igual que cuatro. Este valor mide la relación entre la variabilidad debida al sistema de medición y la variabilidad inherente. Si es inferior a 4, entonces la variabilidad R&R del instrumento es grande en comparación con la variabilidad inherente. De lo contrario, la relación entre ambas variabilidades puede considerarse adecuada.

**Ejemplo**. Voltaje de baterías (continuación). La contribución de la variabilidad de la medición R&R a la variación total del proceso $\% \quad Variacion \quad medición$ es como sigue:

```
voltimetro <- factor(rep(1:2, each = 9))
bateria <- factor(rep(rep(1:3, each = 3), 2))
run <- factor(rep(1:3, 6))
voltaje <- c(1.4727, 1.4206, 1.4754, 1.5083, 1.5739,
             1.4341, 1.5517, 1.5483, 1.4614, 1.3337,
             1.6078, 1.4767, 1.4066, 1.5951, 1.8419,
             1.7087, 1.8259, 1.5444)
baterias <- data.frame(voltimetro, bateria, run, voltaje)
baterias
```

```
##    voltimetro bateria run voltaje
## 1           1       1   1  1.4727
## 2           1       1   2  1.4206
## 3           1       1   3  1.4754
## 4           1       2   1  1.5083
## 5           1       2   2  1.5739
## 6           1       2   3  1.4341
## 7           1       3   1  1.5517
## 8           1       3   2  1.5483
## 9           1       3   3  1.4614
## 10          2       1   1  1.3337
## 11          2       1   2  1.6078
## 12          2       1   3  1.4767
## 13          2       2   1  1.4066
```

```
## 14          2       2   2  1.5951
## 15          2       2   3  1.8419
## 16          2       3   1  1.7087
## 17          2       3   2  1.8259
## 18          2       3   3  1.5444
```

```
anova(lm(voltaje ~ bateria + voltimetro +
           bateria * voltimetro,
         data = baterias))
```

```
## Analysis of Variance Table
##
## Response: voltaje
##                    Df   Sum Sq  Mean Sq F value Pr(>F)
## bateria             2 0.063082 0.031541  1.9939 0.1788
## voltimetro          1 0.044442 0.044442  2.8095 0.1195
## bateria:voltimetro  2 0.018472 0.009236  0.5839 0.5728
## Residuals          12 0.189821 0.015818
```

```
(sqrt(0.0197) / sqrt(0.0234)) * 100
```

```
## [1] 91.75404
```

Esto significa que mucha de la variación del proceso es debido a la variabilidad de la medición R&R. El número de categorías distintas $Ncategorías$ es

```
(0.0037 / sqrt(0.0197)) * 1.41
```

```
## [1] 0.03716959
```

Como fue indicado antes, se toma este valor como 1. Este resultado confirma lo inadecuado del actual sistema de medición.

# Empleando el paquete SixSigma

El gráfico y el calculo numérico son recogidos en la función **ss.rr** del paquete **SixSigma**. La sintaxis de la función es la siguiente

```
# {ss.rr}(var,part, appr, data, main, sub)
```

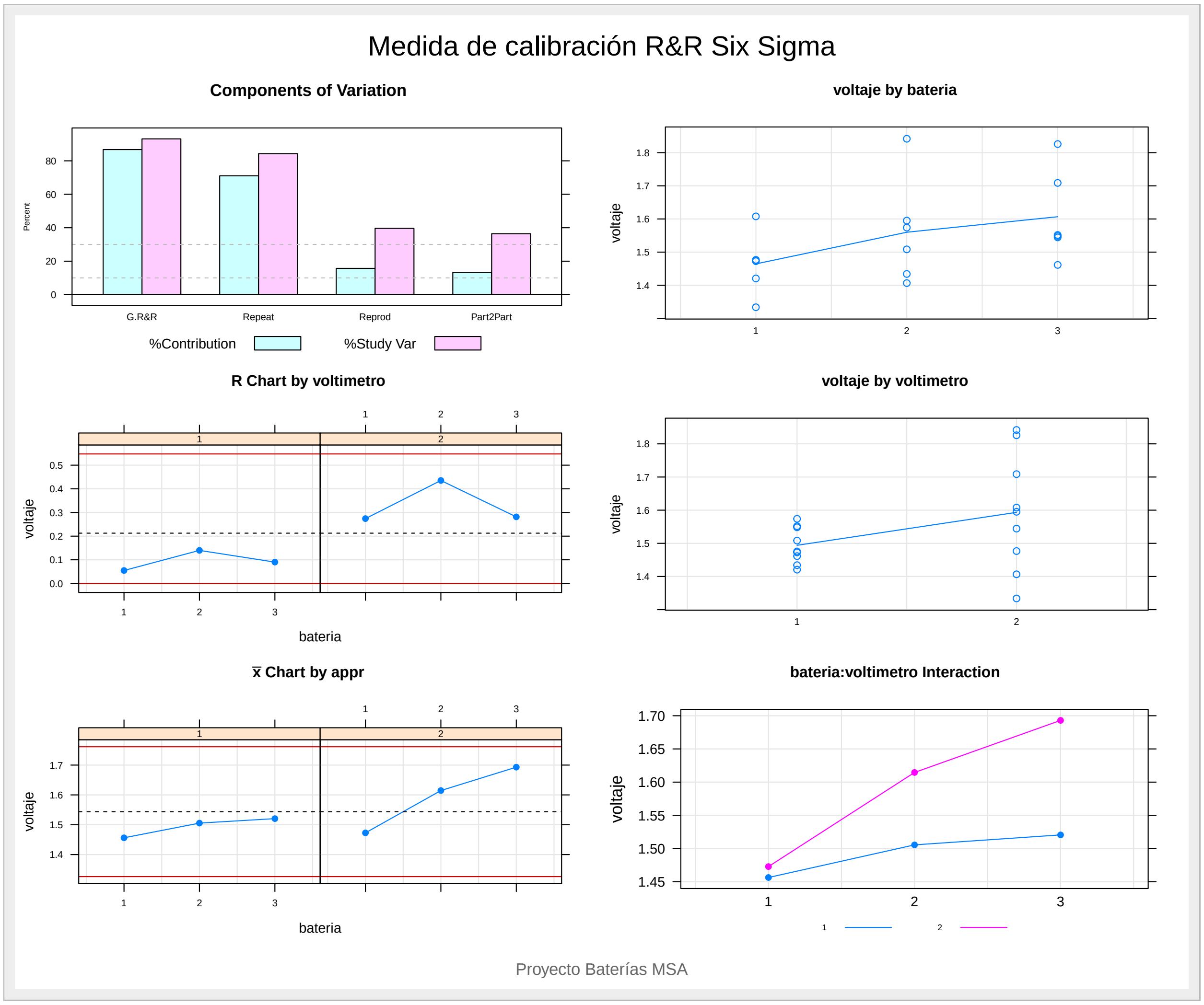
Después de utilizar esta función, se obtiene un texto con una tabla ANOVA, donde se presenta en detalle las fuentes de variación, y algunos gráficos en particular:

- Gráfico de barras para el porcentaje de contribución de cada fuente de variación
- Valores medidos por gráfico de tasador
- Valores medidos por parcela parcial,
- Valores medidos medios por parte y evaluador,
- Gráfico de control medio para el estudio R&R,
- Tabla de control de rango para estudio R&R.

**Ejemplo**. Voltaje de baterías (continuación). El texto y el gráfico resultante para los datos de baterías pueden ser obtenidos y guardados en un objeto en $R$ empleando la siguientes linea de comando:

```
library(SixSigma)
my.rr <- ss.rr(var = voltaje, part = bateria,
appr = voltimetro,
data = baterias,
main = "Medida de calibración R&R Six Sigma",
sub = "Proyecto Baterías MSA")
```

```
## Complete model (with interaction):
##
##                      Df  Sum Sq Mean Sq F value Pr(>F)
## bateria               2 0.06308 0.03154   3.415  0.227
## voltimetro            1 0.04444 0.04444   4.812  0.160
## bateria:voltimetro    2 0.01847 0.00924   0.584  0.573
## Repeatability        12 0.18982 0.01582
## Total                17 0.31582
##
## alpha for removing interaction: 0.05
##
##
## Reduced model (without interaction):
##
##                 Df  Sum Sq Mean Sq F value Pr(>F)
## bateria          2 0.06308 0.03154   2.120  0.157
## voltimetro       1 0.04444 0.04444   2.987  0.106
## Repeatability 14 0.20829 0.01488
## Total           17 0.31582
##
## Gage R&R
##
##                        VarComp %Contrib
## Total Gage R&R    0.018162959    86.74
##   Repeatability   0.014878111    71.05
##   Reproducibility 0.003284848    15.69
##     voltimetro    0.003284848    15.69
## Part-To-Part      0.002777127    13.26
## Total Variation   0.020940086   100.00
##
##                        StdDev  StudyVar %StudyVar
## Total Gage R&R    0.13477002 0.8086201     93.13
##   Repeatability   0.12197586 0.7318552     84.29
##   Reproducibility 0.05731359 0.3438816     39.61
##     voltimetro    0.05731359 0.3438816     39.61
## Part-To-Part      0.05269846 0.3161907     36.42
## Total Variation   0.14470690 0.8682414    100.00
##
## Number of Distinct Categories = 1
```




Los resultados muestran que % *Variacion medición* debido a R&R (cuarta tabla, primera fila, tercera columna) es 93.13%, variabilidad mayor que 30%. En adición, el número el número de categorías distintas es igual a 1. Este pequeño número de categorías, junto con tan amplio porcentaje de variabilidad, es el peor resultado posible para un sistema de medición.

Para descubrir donde se encuentran los problemas del sistema de medición, podemos hacer uso de algunas herramientas gráficas.

## Interpretación de los gráficos

Se puede generar un diagrama de barras para ver la contribución de cada componente a la varianza total. Así podremos detectar de un vistazo si los problemas de medición provienen de la repetibilidad o reproducibilidad.

**Ejemplo**. Voltaje de baterías (continuación). La gráfica en la parte superior izquierda de la Medición de calibración R&R es un gráfico de barras que representa la contribución de cada componente a la varianza total. Su objetivo es detectar de un vistazo si los problemas de medición provienen de la repetibilidad o la reproducibilidad. En este caso, está claro que el proceso tiene problemas con la repetibilidad y los valores de reproducibilidad, ya que su contribución es en ambos casos superior al 10%, siendo la repetibilidad aproximadamente el doble de reproducibilidad

Usando gráficos de franjas y gráficos de líneas que representan los efectos, podemos ver si la diferencia entre los evaluadores es el problema o si la interacción entre los evaluadores y las partes es importante.

**Ejemplo**. Voltaje de baterías (continuación). Los primeros dos gráficos en la columna derecha de la Medición de calibración R&R muestran cada medida como un punto en el gráfico. En la gráfica superior, el eje $x$ representa las baterías, y se ha trazado una línea que une los medios de cada prototipo. En el diagrama central, el eje $x$ representa los voltímetros. La gráfica inferior muestra la interacción entre los dos factores operadores y prototipos. Los medios de los pares *batería* $\times$ *voltímetro* están representados por puntos y unidos por líneas.

Con estos gráficos podemos detectar si existe alguna interacción entre operadores y prototipos (lo que sería un problema) o diferencias entre operadores. En el caso en cuestión, detectamos que la primera batería parece tener un voltaje más bajo que el resto, pero su variabilidad es similar a la de las otras, por lo que esto no es un problema para el sistema de medición. Sin embargo, la trama del tasador indica diferencias aparentes entre los voltímetros. Las medias y las variabilidades son diferentes, y este es un problema para nuestro sistema de medición. Las líneas en el diagrama de interacción no se cruzan, por lo que la interacción entre factores es irrelevante.

Finalmente, los gráficos de control se pueden trazar por grupo y con los límites de control adaptados para los estudios de R&R. Para el gráfico de control medio, la línea central y los límites son:

$$Lineacentral = \bar{\bar{x}},$$

$$Límitesuperior = \bar{\bar{x}} + \frac{3}{d_2\sqrt{n}}\bar{R},$$

$$Límiteinferior = \bar{\bar{x}} - \frac{3}{d_2\sqrt{n}}\bar{R};$$

y para el gráfico de control por rangos

$$Lineacentral = \bar{R},$$

$$Límitesuperior = \bar{R} \times \left(1 + \frac{d3}{d2}\right),$$

$$Límiteinferior = \bar{R} \times \left(1 - \frac{d3}{d2}\right)$$

donde d2 y d3 son las constantes famosas de Shewart para construir gráficos de control, x es el promedio general y R es el rango promedio.

Dentro de un estudio R&R del sistema de medición, la mayoría de los puntos en el gráfico $x$ deben estar fuera de los límites de control. Esto se debe al hecho de que el gráfico representa una variación de parte a parte (mismo operador, mismo prototipo), mientras que los límites se aplican a los datos generales. De lo contrario, el sistema de medición se consideraría inadecuado. Sin embargo, en la tabla de control de rango, todos los puntos deben estar dentro de los límites de control.

## Análisis de la capacidad de los procesos

El análisis de capacidad es una herramienta muy importante en la fase de medición. Dado que se espera que la salida de un proceso cumpla con los requisitos del cliente, las especificaciones o las tolerancias de ingeniería. Puede ser realizado un estudio de la capacidad del proceso para determinar en qué medida el proceso puede cumplir con estas expectativas.

La capacidad del proceso es una propiedad medible de un proceso según la especificación, expresada como Índice de capacidad de proceso $C_{pk}$.

Creamos datos para el análisis de la capacidad del proceso en R. Suponer que se tiene una muestras de 20 alimentos con valores de calorías.

```
# Creamos los datos para el análisis de capacidad del procesos
muestra_alimentos <- c(755.81, 750.54, 751.05, 749.52, 749.21, 748.38,
     748.11, 753.07, 749.56, 750.08, 747.16, 747.53,
     749.22, 746.76, 747.64, 750.46, 749.27, 750.33,
     750.26, 751.29)
```

## Índices $C_p$ y $C_{pk}$

$C_p$ es una medida de la capacidad del proceso que compara la especificación anchura (límite superior - límite inferior) a la variación total del proceso ($6\sigma$) para una proceso dado o parámetros del producto. Cuando mayor es el $C_p$, más potencial tiene un proceso para ajustarse dentro de los límites de especificación.

$$C_p = \frac{USL - LSL}{6\sigma}$$

$C_{pk}$ es muy útil porque permite variación ($\sigma$) y posición ($\bar{X}$) entre los límites superior e inferior. Si la media esta más cercana al límite superior (USL), entonces $C_{pk}$ es $\frac{USL-\bar{X}}{3\sigma}$. Si la media está más cerca al límite inferior entonces $C_{pk}$ es $\frac{\bar{X}-USL}{3\sigma}$

$$C_{pk} = min\left(\frac{USL - \bar{X}}{3\sigma}, \frac{\bar{X} - LSL}{3\sigma}.\right)$$

Un proceso centrado perfectamente, $C_p = C_{pk}$.

En el ejemplo de muestra_alimentos (después de establecer $LSL = 740$ y $USL = 760$), se calcula $C_p$ y $C_{pk}$ empleando el paquete **SixSigma**:

```
library(SixSigma)
ss.ca.cp(muestra_alimentos,740,760)
```

```
## [1] 1.584136
```

```
ss.ca.cpk(muestra_alimentos,740,760)
```

```
## [1] 1.546513
```

La capacidad del proceso actual es aceptable pero puede ser mejorada para alcanzar el valor deseado para $C_{pk}$ (1.67). Los intervalos de confianza obtenidos cuando se adicionan los parámetros **ci** a la funciones.

```
ss.ca.cp(muestra_alimentos,740,760, ci = TRUE)
```

```
## [1] 1.084600 2.083046
```

```r
ss.ca.cpk(muestra_alimentos,740,760, ci = TRUE)
```

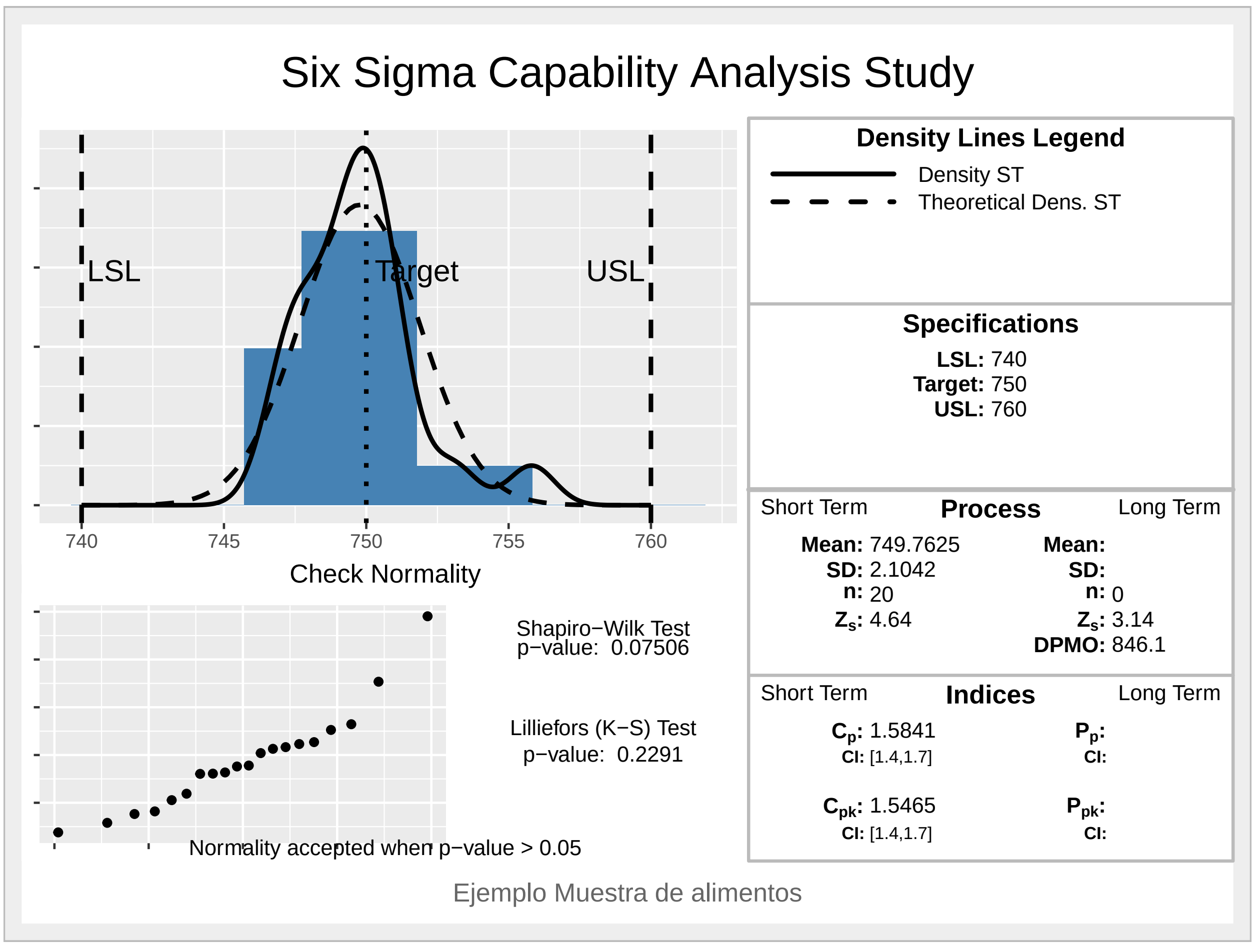
```
## [1] 1.033560 2.059466
```

### Estudio de capacidad con el paquete SixSigma

La difusión de los resultados Six Sigma es muy importante para los profesionales, especialmente para el Master Black Belts, para difundir el pensamiento Six Sigma en toda la organización. Se puede obtener un resumen de un análisis de capacidad en R con la función **ss.study** en el paquete SixSigma.

```r
# realizar Estudio de capacidad
ss.study.ca(muestra_alimentos, LSL = 740, USL = 760,
            Target = 750, alpha = 0.5,
            f.su = "Ejemplo Muestra de alimentos")
```

```
## Warning in ss.study.ca(muestra_alimentos, LSL = 740, USL = 760, Target =
## 750, : Normality test/s failed
```




### Interpreación del gráfico

- El gráfico superior es un histograma de los datos de muestra, que incluye el objetivo y los límites de especificación. Las líneas de densidad se trazan para las funciones de densidad empírica y teórica.

- El gráfico inferior es un gráfico cuantil-cuantil (gráfico Q – Q) para verificar si los datos se distribuyen normalmente. Cuando lo son, los puntos están aproximadamente en línea recta. Además, también se muestran las pruebas numéricas más comunes. La normalidad se acepta cuando el valor p de la prueba de hipótesis es mayor que 0.05. En el ejemplo de muestra de alimentos anterior, los datos son normales ya que el valor $p$ es superior a 0.05.

- El rendimiento del proceso y los índices se calculan con los datos proporcionados a corto y largo plazo. Para datos a largo plazo, también muestra los valores de $P_p$ y $P_{pk}$. En este ejemplo, no tenemos datos a largo plazo. El índice de capacidad (1.547) es bastante aceptable, aunque se puede mejorar para alcanzar el valor deseado de 1.67 (para el proceso 6 sigma).