

Control estadístico de procesos

Parte I

Erick Orozco-Acosta¹

Universidad Simón Bolívar

Facultad de Ingeniería

Programa de Ingeniería Industrial

2024-01-26

¹erick.orozco@unisimon.edu.co

Principios de control estadístico de procesos (CEP)

Introducción del CEP

Según Kumar y Suresh (2009),...

Date	Contribution	Contributor
1776	Specialization of labour in manufacturing	Adam Smith
1799	Interchangeable parts, cost accounting	Eli Whitney & others
1832	Division of labour by skill; assignment of jobs by Skill; basics of time study	Charles Babbage
1900	Scientific management time study and work study Developed; dividing planning and doing of work	Frederick W.Taylor
1900	Motion of study of jobs	Frank B. Gilbreth
1901	Scheduling techniques for employees, machines Jobs in manufacturing	Henry L. Gantt
1915	Economic lot sizes for inventory control	F.W. Harris
1927	Human relations; the Hawthorne studies	Elton Mayo
1931	Statistical inference applied to product quality: quality control charts	W.A. Shewart
1935	Statistical Sampling applied to quality control: inspection sampling plans	H.F.Dodge & H.G.Roming
1940	Operations research applications in world war II	P.M.Blacker & others
1946	Digital Computer	John Mauchly and J.P.Eckert
1947	Linear Programming	G.B.Dantzig, Williams & others
1950	Mathematical programming, on-linear and stochastic processes	A.Charnes, W.W.Cooper & others
1951	Commercial digital computer: large-scale computations available	Sperry Univac
1960	Organisational behaviour: continued study of people at work	L.Cummings, L.Porter
1970	Integrating operations into overall strategy and policy Computer applications to manufacturing, scheduling, and control, Material Requirement Planning (MRP)	W.Skinner J.Orlicky & G. Wright
1980	Quality and productivity applications from Japan: robotics, CAD-CAM	W.E. Deming & J.Juran

Figura 1: Resumen histórico de la administración de operaciones

Introducción del CEP I

Tomando como referencia Vargas (2006), se tiene ...

- ▶ Las compañías son exitosas por la innovación, productividad y calidad de los bienes que ponen en el mercado.
- ▶ Los avances en materia de desarrollo tecnológico genera que se pueda producir más de lo que se vende. Así²,...

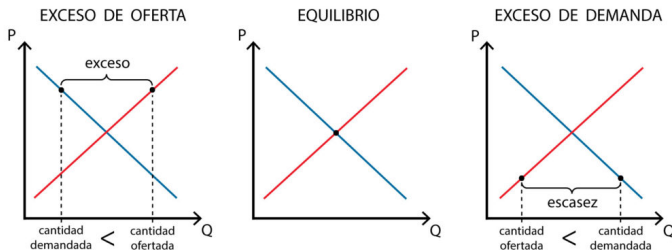


Figura 2: Curva de oferta y demanda

²<https://policonomics.com/es/oferta-demanda/>

Introducción del CEP II

- ▶ Lo anterior, hace muy intensa la lucha por la participación de mercado por la competencia entre calidad y precio.
- ▶ Es claro que mejorar la calidad -» aumenta la productividad -» menor costo por unidad.

Pero, ¿qué se entiende por calidad?, revisemos. . .

- ▶ Trabajos como Duncan (1987), Ryan (2000) y Montgomery (2001) discuten acerca de ello.
- ▶ Hoyer y Hoyer (2001) realiza una comparación de las posturas de diferentes gurús.

Introducción del CEP III

- ▶ **W. Edwards Deming** dice que la calidad se debe definir en términos de satisfacción del cliente y plantea la calidad de forma multidimensional.
- ▶ Para **Joseph M. Juran** la calidad son aquellas características del producto que satisfacen las necesidades del consumidor. Además, calidad=producto libre de defectos.
- ▶ Lo anterior, concuerda con **Philip B. Crosby** que habla de la teoría de “cero defectos”.
- ▶ **Kaoru Ishikawa** concuerda con el aspecto de la satisfacción al cliente, pero además involucra el concepto de calidad en toda de organización.

Introducción del CEP IV

- Para **Genichi Taguchi** la calidad es la pérdida que un producto le causa a la sociedad una vez es embarcado. Menor pérdida monetaria=mayor satisfacción del cliente. Define una función de pérdida así³:

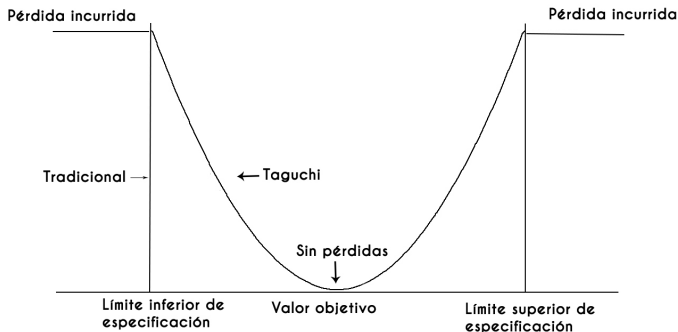


Figura 3: Función de pérdida de Taguchi

³<https://www.lifeder.com/genichi-taguchi/>

Introducción del CEP V

Reto: Mostrar matemáticamente que la función de pérdida $L(x) = k(x - N)^2$ se minimiza cuando el valor de la característica de calidad observado x es igual al valor nominal N . Recuerde que k es una constante de proporcionalidad.

- ▶ En 1924, **Walter A. Shewhart** de *Bell Telephone Laboratories* hizo un primer bosquejo de una carta de control. En 1931 publicó *Economic control quality of manufactured product* donde se sientan las bases del control estadístico de calidad.
- ▶ En 1944, **Harold F. Dodge** y **H. Roming** de *Bell System* publicaron las *Tablas de inspección por muestreo (simple y doble)*.

Introducción del CEP VI

- ▶ Durante la segunda guerra mundial se dieron avances en las técnicas de control de calidad en EEUU e Inglaterra pero al final la aplicación disminuyó considerablemente.
- ▶ A comienzos de los 50's **W. Edwards Deming** viaja a Japón invitado por la *Unión Japonesa de Científicos e Ingenieros* para impartir varios seminarios y cursos. **Deming** desarrolló el concepto de calidad como objetivo económico y estratégico.
- ▶ A partir de lo anterior, en los 70's se observa al resurgir de la industria japonesa y hace que occidente retome el interés en los métodos de control de calidad.

Introducción del CEP VII

- ▶ A finales de los 70's e inicios de los 80's se estructuran sistemas globales conocidos como "Control Total de la Calidad" (TQC) en Japón y el "Manejo de la Calidad Total" (TQM) en EEUU.
- ▶ Algunos consideran el TQC/TQM como una verdadera filosofía empresarial. Mientras que otros lo consideran una cultura. Una parte del TQC/TQM es el "Control estadístico de procesos" (SPC).
- ▶ Según Stoumbos et. al. (2000), el SPC se refiere a "métodos estadísticos usados extensivamente para monitorear y mejorar la calidad y productividad de procesos de manufactura y operaciones de servicios".

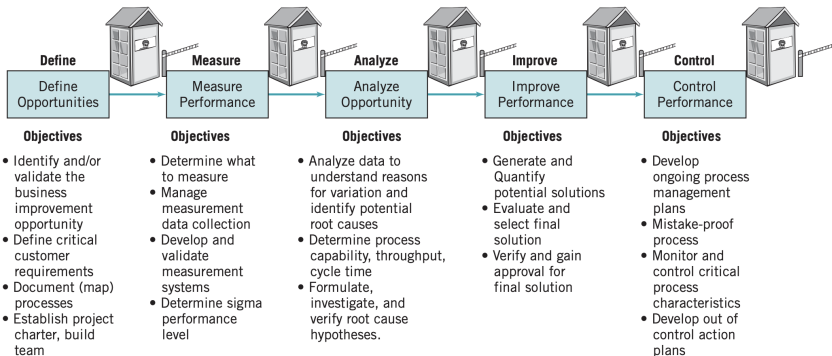
Introducción del CEP VIII

Dos áreas se destacan en el control estadístico de procesos:

- ▶ Planeamiento de la calidad: se distinguen dos etapas,
 - ▶ *Diseño de producto*: Son las especificaciones de las características de calidad del producto y sus regiones de tolerancia.
 - ▶ *Diseño de proceso*: se relaciona con el diseño del proceso de fabricación y la identificación de las causas de variabilidad (uso de diseños factoriales).
- ▶ Inspección de la calidad: determina si un producto satisface con los requerimientos de calidad deseados. Por ejemplo: índices de capacidad de proceso y cartas de control.

Introducción del CEP IX

DMAIC: es una herramienta enfocada en la mejora de los problemas de forma focalizada en la mejora continua. Montgomery (2001) lo ilustra así:



“The magnificent seven” (Los siete magníficos)

Aspectos de los 7's magníficos

- ▶ Se refiere a las siguientes herramientas de CEP:
 1. Histograma
 2. Hoja de creckeo (Check list)
 3. Diagrama de Pareto
 4. Diagrama Causa-Efecto
 5. Diagrama de concentración de defectos
 6. Diagrama de dispersión
 7. Cartas de control

- ▶ La más sofisticada es la carta de control que fue introducida por **Walter A. Shewhart**, que verá más adelante.

Histograma I

- ▶ Es una herramienta gráfica que permite estudiar el centramiento, dispersión y la distribución de los datos en un rango de variación.
- ▶ Surge como resultado de construir una tabla de datos agrupada. Los detalles numéricos y desarrollo en Excel se puede observar [aquí](#).

Ejemplo 1: Toma el conjunto de datos Cars93 del paquete MASS de R, y realice los siguientes:

- (a) Realice una tabla de datos agrupados para la variable Weight.
- (b) Extraiga los principales estadísticos.
- (c) Realice un histograma usando frecuencias y densidades, y en conjunto con lo anterior saque conclusiones.

Histograma II

Solución:

Inciso (a): Tabla de datos agrupados.

```
library(agricolae)
library(MASS)
datos <- Cars93$Weight # vector variable peso
aux <- hist(datos, plot=FALSE) # resultados
Tabla <- table.freq(aux); Tabla # La tabla agrupada
```

##	Lower	Upper	Main	Frequency	Percentage	CF	CPF
## 1	1500	2000	1750	3	3.2	3	3.2
## 2	2000	2500	2250	17	18.3	20	21.5
## 3	2500	3000	2750	25	26.9	45	48.4
## 4	3000	3500	3250	21	22.6	66	71.0
## 5	3500	4000	3750	23	24.7	89	95.7
## 6	4000	4500	4250	4	4.3	93	100.0

Histograma III

Inciso (b): Principales estadísticos.

```
summary(datos)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	1695	2620	3040	3073	3525	4105

```
sd(datos) #desviación estándar
```

```
## [1] 589.8965
```

- ▶ ¿Más o menos entre qué valores se podría tener una aproximación de la moda? ¿Cómo se calcularía?
- ▶ ¿Qué se podría decir de la relación entre Media, Mediana y Moda?

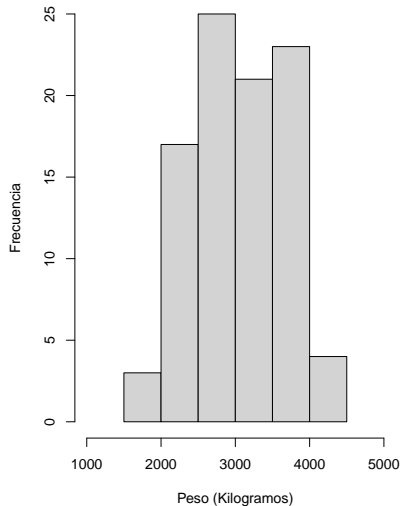
Histograma IV

Inciso (c): histograma usando frecuencias y densidades

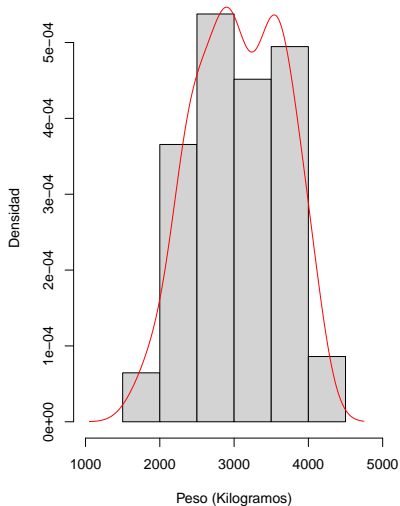
```
par(mfrow=c(1,2)) # ubicación de los gráficos
hist(datos,
      xlab = "Peso (Kilogramos)",
      ylab = "Frecuencia",
      main = "Peso de los carros (Frecuencia)",
      xlim = c(1000, 5000)
)
hist(datos,
      freq = FALSE, # parámetro de densidad
      xlab = "Peso (Kilogramos)",
      ylab = "Densidad",
      main = "Peso de los carros (Densidad)",
      xlim = c(1000, 5000)
)
lines(density(datos), col = "red") # traza densidad
```

Histograma V

Peso de los carros (Frecuencia)



Peso de los carros (Densidad)



Conclusiones:

Histograma VI

- ▶ Debido a que $\text{Moda} < \text{Mediana} < \text{Media}$ el sesgo es positivo.
- ▶ La forma de los histogramas es companoide con dos modas (bimodal) ¿Qué creen que podría estar ocurriendo? ¿Debemos llegar hasta aquí?

Ejemplo 2: Hay dos máquinas extrusoras de polietileno de baja densidad (LDPE) que están fabricando el mismo producto que es un tubular para empaques flexible. Las dos variables fundamentales en este proceso son el ancho y calibre (“espesor”) de la película. Las especificaciones son 60 ± 2 cm de ancho en calibre 2 ± 0.2 milésimas de pulgada. Como ingeniero del proceso ¿tendría argumentos para pensar que las máquinas están bien configuradas cumpliendo las especificaciones? Nota: Usar los datos del objeto `extruders.Rdata`.

solución:

Histograma VII

Parte I: análisis exploratorio de datos.

```
library(dplyr)
load("data/extruders.Rdata")
# Media y SD de cada variable por extrusora
xx <- extruders %>% group_by(extruder) %>%
  summarise(mean_an=mean(ancho), sd_an=sd(ancho),
            mean_ca=mean(calibre), sd_ca=sd(calibre))
xx
```

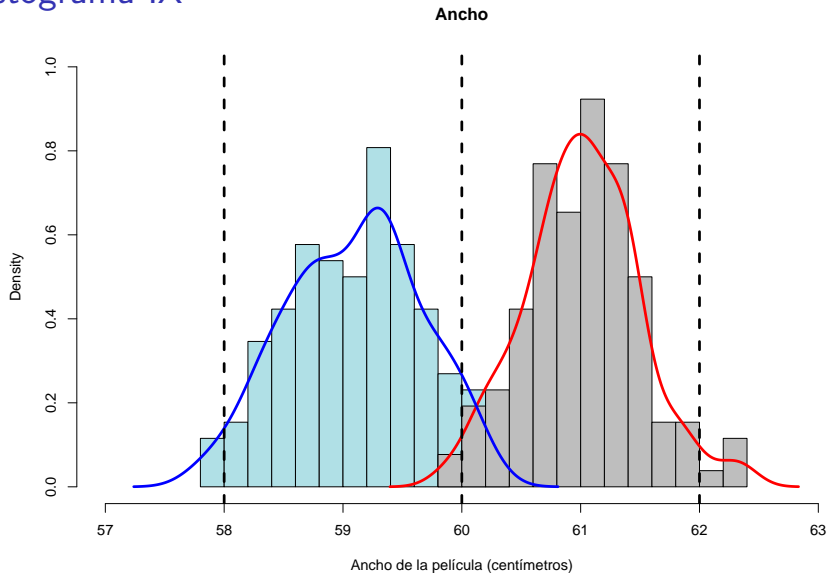
```
## # A tibble: 2 x 5
##   extruder   mean_an sd_an mean_ca sd_ca
##   <chr>      <dbl> <dbl>   <dbl> <dbl>
## 1 extruder_1    59.1 0.560     2.01 0.0977
## 2 extruder_2    61.0 0.481     1.94 0.0760
```

Histograma VIII

Analizando el ancho de la película se tiene:

```
# divide el conjunto de datos
ext <- split(extruders,extruders$extruder)
# histograma y densidad para ancho extrusora 1 y 2
hist(ext[[1]]$ancho, main = "Ancho",
      xlab = "Ancho de la película (centímetros)",
      freq = FALSE, ylab = "Density", col = "powderblue",
      xlim = c(57, 63), ylim = c(0,1))
hist(ext[[2]]$ancho, main = "", xlab = "", ylab = "",
      freq = FALSE, col = "grey", add = TRUE,
      xlim = c(57, 63), ylim = c(0,1))
lines(density(ext[[2]]$ancho), col = "red",lwd=3)
lines(density(ext[[1]]$ancho), col = "blue",lwd=3)
abline(v=58,col="black",lwd=3,lty='dashed') # EI
abline(v=60,col="black",lwd=3,lty='dashed') # Nominal
abline(v=62,col="black",lwd=3,lty='dashed') # ES
```

Histograma IX



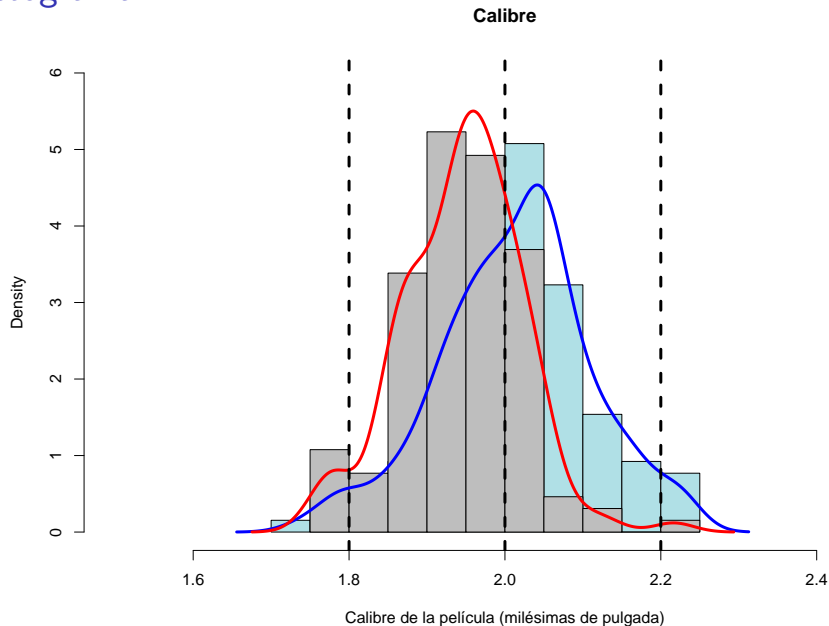
Azul claro y azul oscuro (Extruder 1) y Gris y rojo (Extruder 2)

Histograma X

Luego para el calibre...

```
# histograma y densidad para calibre extrusora 1 y 2
hist(ext[[1]]$calibre, main = "Calibre",
      xlab = "Calibre de la película (milésimas de pulgada)",
      freq = FALSE, ylab = "Density", col = "powderblue",
      xlim = c(1.5, 2.5), ylim = c(0,6))
hist(ext[[2]]$calibre, main = "", xlab = "", ylab = "",
      freq = FALSE, col = "grey", add = TRUE,
      xlim = c(1.5, 2.5), ylim = c(0,6))
lines(density(ext[[1]]$calibre), col = "blue",lwd=3)
lines(density(ext[[2]]$calibre), col = "red",lwd=3)
abline(v=1.8,col="black",lwd=3,lty='dashed') # EI
abline(v=2,col="black",lwd=3,lty='dashed') # Nominal
abline(v=2.2,col="black",lwd=3,lty='dashed') # ES
```


Histograma XI



Histograma XII

A partir de los gráficos anteriores, se puede afirmar:

- ▶ Ambas extrusoras están arrojando producto fuera de las especificaciones del cliente.
- ▶ Descriptivamente, las máquinas muestran configuraciones parecidas en términos de calibre pero diferentes en el ancho de la película.
- ▶ ¿Hay significancia estadística en la igualdad o diferencia en las medias de los anchos de las películas/calibres para las extrusoras?
- ▶ Es decir, para ancho y calibre se debe establecer: $H_0 : \mu_1 = \mu_2$ v.s $H_I : \mu_1 \neq \mu_2$.

Histograma XIII

Para los anterior, se tienen los siguientes supuestos:

- ▶ A partir de la lectura se puede asumir que las extrusoras son independientes.
- ▶ ¿Se puede asumir que las muestras provienen de poblaciones normales? (Recomendación: pruebas de normalidad)
- ▶ Las varianzas son desconocidas.
- ▶ ¿Se puede asumir que las varianzas son iguales? (Recomendación: contraste de varianzas)

Para el ancho...

```
# prueba de Shapiro-Wilk para ancho extruder 1  
shapiro.test(ext[[1]]$ancho)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: ext[[1]]$ancho  
## W = 0.98658, p-value = 0.2332
```

Histograma XIV

```
# prueba de Shapiro-Wilk para ancho extruder 2  
shapiro.test(ext[[2]]$ancho)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  ext[[2]]$ancho  
## W = 0.99032, p-value = 0.5028
```

```
# prueba de igualdad de varianzas para ancho  
var.test(x = ext[[1]]$ancho, y= ext[[2]]$ancho)
```

```
##  
## F test to compare two variances  
##  
## data:  ext[[1]]$ancho and ext[[2]]$ancho  
## F = 1.3523, num df = 129, denom df = 129, p-value = 0.08776  
## alternative hypothesis: true ratio of variances is not equal to 1  
## 95 percent confidence interval:  
##  0.9561314 1.9125945  
## sample estimates:  
## ratio of variances  
##          1.352291
```

Histograma XV

Ahora si se puede evaluar si estadísticamente hay diferencia entre la media de los calibre en las extrusoras.

```
#####  
# Caso muestras normales e independientes, varianzas #  
# desconocidas pero iguales #  
#####  
t.test(ancho ~ extruder, data = extruders, var.equal=T)  
  
##  
## Two Sample t-test  
##  
## data: ancho by extruder  
## t = -29.664, df = 258, p-value < 2.2e-16  
## alternative hypothesis: true difference in means between group extruder_1 and  
## 95 percent confidence interval:  
## -2.047779 -1.792828  
## sample estimates:  
## mean in group extruder_1 mean in group extruder_2  
## 59.09488 61.01518
```

Ahora en el caso del calibre...

```
# prueba de Shapiro-Wilk para ancho extruder 1  
shapiro.test(ext[[1]]$calibre)
```

Histograma XVI

```
##  
## Shapiro-Wilk normality test  
##  
## data:  ext[[1]]$calibre  
## W = 0.98794, p-value = 0.3131  
# prueba de Shapiro-Wilk para ancho extruder 2  
shapiro.test(ext[[2]]$calibre)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data:  ext[[2]]$calibre  
## W = 0.98281, p-value = 0.09929  
# prueba de igualdad de varianzas para ancho  
var.test(x = ext[[1]]$calibre, y= ext[[2]]$calibre)
```

Histograma XVII

```
##
## F test to compare two variances
##
## data:  ext[[1]]$calibre and ext[[2]]$calibre
## F = 1.6529, num df = 129, denom df = 129, p-value = 0.004602
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  1.168697 2.337799
## sample estimates:
## ratio of variances
##                1.65293
```

```
#####
# Caso muestras normales e independientes, varianzas #
# desconocidas pero distintas                        #
#####
t.test(calibre ~ extruder, data = extruders, var.equal=F)
```

Histograma XVIII

```
##
```

```
## Welch Two Sample t-test
```

```
##
```

```
## data: calibre by extruder
```

```
## t = 6.23, df = 243.26, p-value = 2.037e-09
```

```
## alternative hypothesis: true difference in means between group extruder_1 and
```

```
## 95 percent confidence interval:
```

```
## 0.04625237 0.08902283
```

```
## sample estimates:
```

```
## mean in group extruder_1 mean in group extruder_2
```

```
## 2.012478 1.944841
```

Todo lo anterior implica que como ingeniero no se puede estar tranquilo!

Hoja de control o chequeo I

- ▶ Las hojas de control o chequeo son una herramienta que se usa para capturar datos de un proceso desde sus inputs a las outputs.
- ▶ Estas hojas muchas veces se hacen de forma híbrida dependiendo del proceso. Es decir, no hay un estándar general.
- ▶ Suelen también usarse como listas de chequeo en industrias como la aeronáutica. Por ejemplo: Verificación de instrumentos y sistemas de las aeronaves.
- ▶ En la actualidad, gran parte de estas hojas se diligencian en sistemas de información de las compañías que vienen dentro de sus ERP's.

Hoja de control o checkeo II

Un ejemplo de la industria aeroespacial en Montgomery (2001)...

CHECK SHEET DEFECT DATA FOR 2002-2003 YTD																										
Part No.:	TAX-41																									
Location:	Bellevue																									
Study Date:	6/5/03																									
Analyst:	TCB																									
	2002												2003													
Defect	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	Total								
Parts damaged	1		3		1	2	1		10		3	2		2	7	2	34									
Machining problems			3	3					1	8	3		8		3	29										
Supplied parts rusted			1	1	2		9												13							
Masking insufficient	3		6	4	3	1												17								
Misaligned weld	2													2												
Processing out of order	2													2					4							
Wrong part issued	1		2											3												
Unfinished fairing			3													3										
Adhesive failure			1		1											2		1	1	6						
Powdery alodine					1													1								
Paint out of limits					1													1		2						
Paint damaged by etching			1													1										
Film on parts					3		1		1												5					
Primer cans damaged							1													1						
Voids in casting									1	1												2				
Delaminated composite									2													2				
Incorrect dimensions											13	7	13		1	1		1	36							
Improper test procedure									1													1				
Salt-spray failure																						4		2	4	
TOTAL	4	5	14	12	5	9	9	6	10	14	20	7	29	7	7	6	2	166								

Figura 4: Hoja de control para una proceso de fabricación aeroespacial

Hoja de control o checkeo III

Otro ejemplo de fabricación en Gutiérrez Pulido y Vara Salazar (2009)...

HOJA DE VERIFICACIÓN PARA DEFECTOS EN VÁLVULAS			
Periodo: _____ Departamento: _____			
MODELO DE PRODUCTO	ZONA DEL MOLDE		
	ZONA 1	ZONA 2	ZONA 3
A	ooo xxx ++	ooooo xx ++ //	ooooooooo xxxxxx /
B	oooo xx +++ /	ooooo xxxxx /	ooooooooo xxxxxxx ++
C	ooooo x +	oooo xxx	ooooooooo xxxxx /
D	oooo xx ++ //	ooooo xxx /	ooooooooooo xxxxx ++++

Códigos para defectos: o porosidad, + maquinado, x llenado, / ensamble

Figura 5: Hoja de control para una proceso de fabricación de válvulas

Diagrama de Pareto I

A partir de Gutiérrez Pulido y Vara Salazar (2009) se tiene:

- ▶ Se sabe que más del 80% de la problemática en una organización se debe a causas comunes. Pero son pocos los problemas o situaciones vitales que contribuyen en gran medida al problema. (**Premisa de Pareto**)
- ▶ La “ley 80-20” o “Pocos vitales, muchos triviales”. Es decir, poco elementos (20%) generan la mayor parte de los defectos.

```
# librería de control estadístico de calidad
library(qcc)
# crear conjunto de datos de defectos
defectos=c(50, 97, 50, 120, 27, 26)
names(defectos) <- c('Geles', 'Mallas', 'Ajuste',
                    'Pegas', 'Descuelgue', 'Corte')
# diagrama de Pareto
pareto.chart(defectos)
```

Diagrama de Pareto II

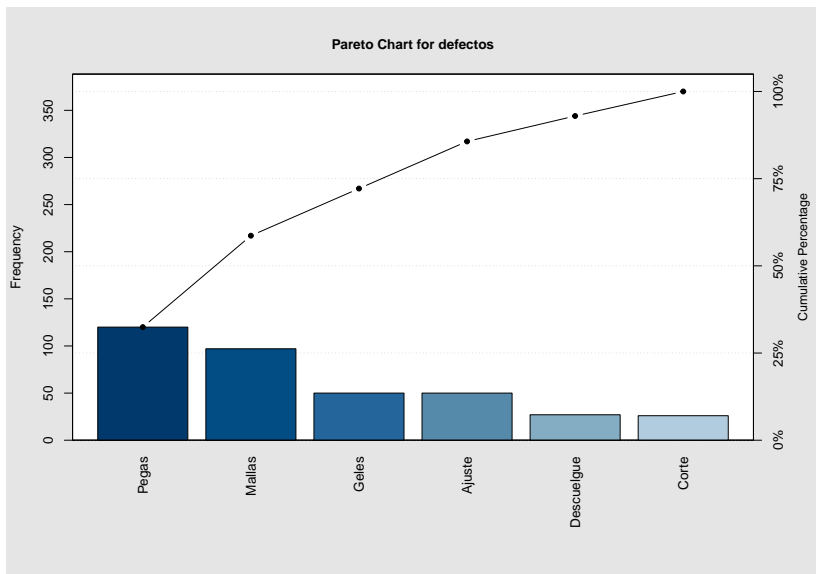


Diagrama de Pareto III

```
##
## Pareto chart analysis for defectos
##      Frequency  Cum.Freq. Percentage Cum.Percent.
##  Pegas      120.000000  120.000000   32.432432    32.432432
##  Mallas      97.000000  217.000000   26.216216    58.648649
##  Geles       50.000000  267.000000   13.513514    72.162162
##  Ajuste      50.000000  317.000000   13.513514    85.675676
##  Descuelgue  27.000000  344.000000    7.297297    92.972973
##  Corte       26.000000  370.000000    7.027027   100.000000
```

Ejemplo (Gutiérrez Pulido y Vara Salazar 2009): De acuerdo con la información de una hoja de verificación en una línea del proceso de envasado de tequila, en el último mes se presentaron los siguientes resultados en cuanto a defectos y frecuencia:

DEFECTO DE ENVASADO	FRECUENCIA
Botella	804
Tapa	715
Etiqueta	1823
Contraetiqueta	742
Sello	916
Otros	102

Diagrama de Pareto IV

```
# crear conjunto de datos de defectos
defectos <- c(804, 715, 1823, 742, 916, 102)
names(defectos) <- c('Botella', 'Tapa',
                    'Etiqueta', 'ContraEtiqu',
                    'Sello', 'Otros')

# diagrama de Pareto
pareto.chart(defectos)
```

Diagrama de Pareto V

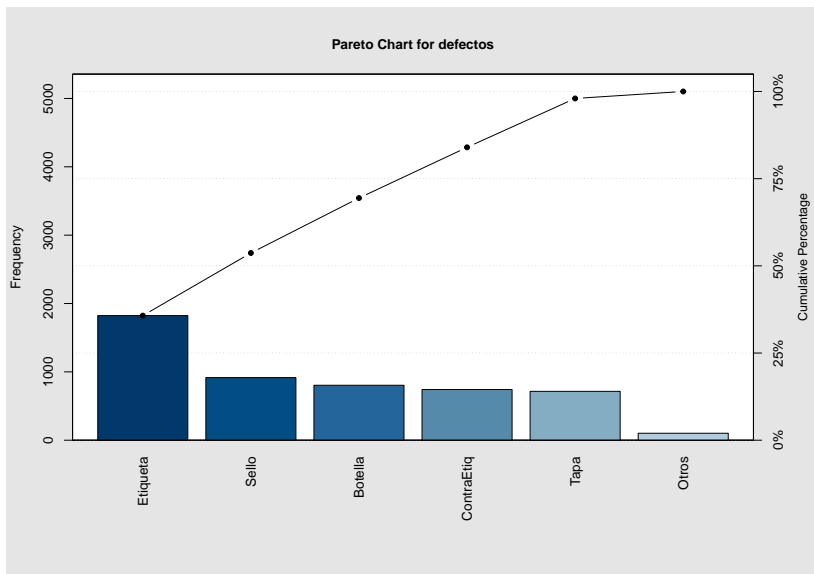


Diagrama de Pareto VI

##

Pareto chart analysis for defectos

##		Frequency	Cum.Freq.	Percentage	Cum.Percent.
##	Etiqueta	1823.000000	1823.000000	35.731086	35.731086
##	Sello	916.000000	2739.000000	17.953744	53.684829
##	Botella	804.000000	3543.000000	15.758526	69.443356
##	ContraEtiqu	742.000000	4285.000000	14.543316	83.986672
##	Tapa	715.000000	5000.000000	14.014112	98.000784
##	Otros	102.000000	5102.000000	1.999216	100.000000

Diagrama causa-efecto I

- ▶ Es una representación gráfica de las causas primarias y secundarias (o derivadas) de un problema en el proceso segmentadas por las 6M's.
- ▶ En la mayoría de las veces, es realizado por las reuniones de círculos de calidad entre los distintos actores del proceso.
- ▶ Se le suele llamar “Diagrama de espina de pescado”, por su forma, o también “Diagrama de Ishikawa”

Diagrama causa-efecto II

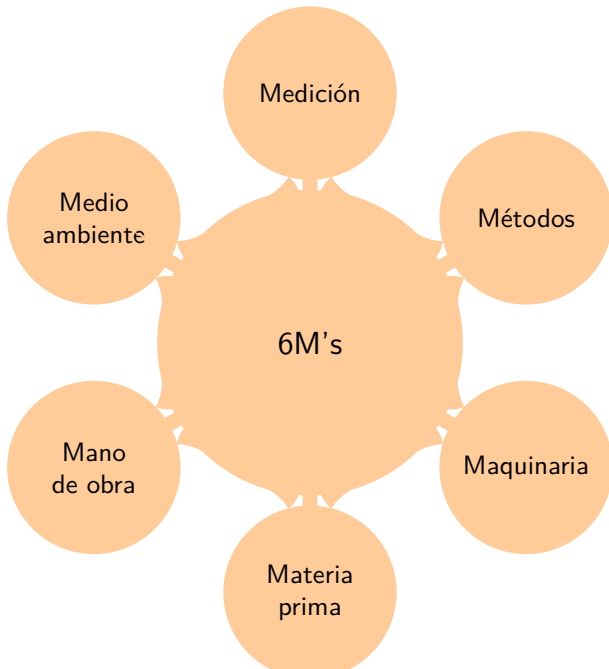


Diagrama causa-efecto III

```
library(SixSigma)
effect <- "Calibre variable"
causes.gr <- c("M.O", "Ambiente", "Medición",
               "Material", "Método", "Maquinaria")
causes <- vector(mode = "list",
                  length = length(causes.gr))
causes[1] <- list(c("operario #1", "operario #2"))
causes[2] <- list(c("Temperatura", "Viento"))
causes[3] <- list(c("Instrumentos"))
causes[4] <- list(c("HDPE", "LHDPE"))
causes[5] <- list(c("Configuración"))
causes[6] <- list(c("Rodamientos"))
ss.ceDiag(effect,
           causes.gr,
           causes,
           sub = "Calibre variable")
```

Diagrama causa-efecto IV

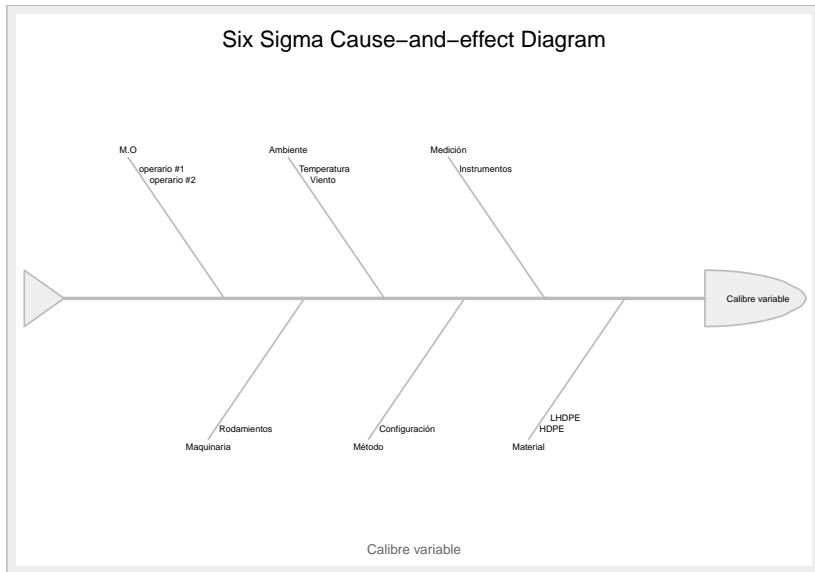


Diagrama de concentración de defectos

De Montgomery (2001)...

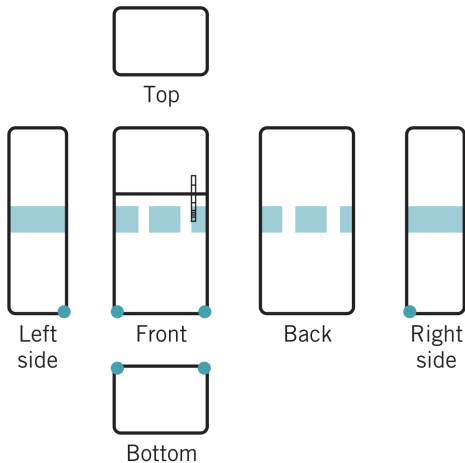


Figura 6: Ejemplo de defectos de un enfriador

Diagrama de dispersión I

- ▶ Es una representación gráfica que muestra la asociación que existe entre dos variables.
- ▶ Los datos son recolectados en pares (x_i, y_i) y son graficados y_i contra la correspondiente x_i .
- ▶ Este diagrama es muy útil en la regresión y guarda relación con la medida de correlación entre dos variables, que se puede medir así:

$$\rho(XY) = \frac{Cov(XY)}{\sqrt{Var(X) \cdot Var(Y)}}$$

- ▶ La correlación entre dos variables no implica causalidad

Diagrama de dispersión II

Ejemplo 4: En un proceso de extrusión se tienen los siguientes datos de 8 órdenes de producción:

Ancho (mm)	120	45	150	200	260	310	185	275
Velocidad (kg/h)	48	125	130	100	90	160	200	250

- (a) Grafique la velocidad (y) contra el ancho (x).
- (b) Calcule la covarianza y el coeficiente de correlación.
- (c) ¿Qué conclusiones extrae de este análisis?

Diagrama de dispersión III

```
ancho <- c(120, 45, 150, 200, 260, 310, 185, 275)  
velocidad <- c(48, 125, 130, 100, 90, 160, 200, 250)  
plot(x = ancho, y = velocidad, pch=19)
```

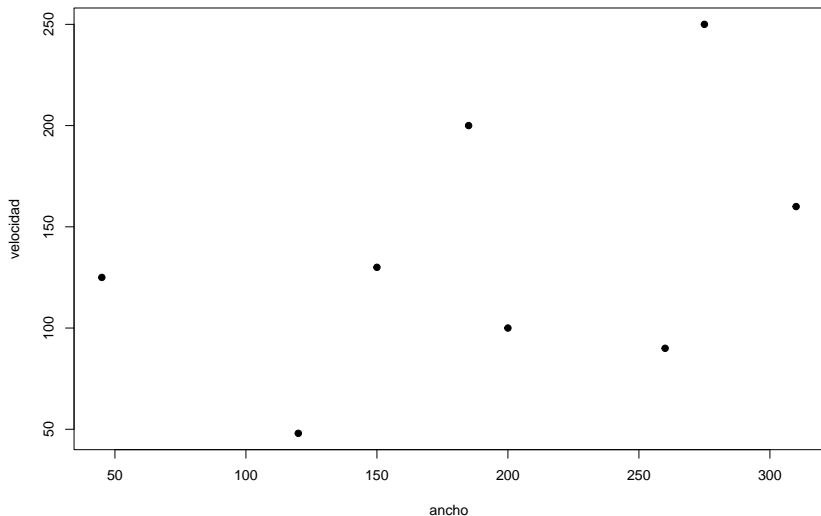


Diagrama de dispersión IV

$$\bar{x} = 193.12$$

$$\bar{y} = 137.88$$

$$\begin{aligned} \text{Cov}(X, Y) &= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n - 1} \\ &= \frac{(120 - \bar{x})(48 - \bar{y}) + \dots + (275 - \bar{x})(250 - \bar{y})}{8 - 1} \\ &= 2374.02 \end{aligned}$$

```
# covarianza
covarianza <- cov(ancho, velocidad)
covarianza

## [1] 2374.018
```

Diagrama de dispersión V

$$\begin{aligned}\rho(X, Y) &= \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}} = \frac{2374.02}{\sqrt{7728.12 \times 4136.12}} \\ &= 0.4199\end{aligned}$$

```
co.correl <- cor(ancho, velocidad)
co.correl
```

```
## [1] 0.4199041
```

****Observaciones**:**

- ▶ La covarianza y correlación es positiva. Sin embargo, esta última es leve.
- ▶ Es normal que la correlación sea leve, debido a que la velocidad de las extrusoras en kg/h también depende del calibre de la película.

Otro magnífico (Despliegue de la función calidad, QFD) I

- ▶ **QFD:** Es una herramienta de planeación que introduce la voz del cliente en el desarrollo de un producto o proyecto.
- ▶ La primera aplicación fue realizada por primera vez en Japón por Mitsubishi en 1972 por Kobe Shipyard.
- ▶ **Voz del cliente:** lista de requerimientos de primer nivel del cliente.
- ▶ **Casa de la calidad:** herramienta que relaciona la voz del cliente con los aspectos técnicos del proceso.

Otro magnífico (Despliegue de la función calidad, QFD) II



Algunas distribuciones de probabilidad

Distribuciones especiales de probabilidad (Discretas) I

Para recordar...

NOMBRE	FUNCION	PARAMETROS	$E(X)$	$V(X)$
Uniforme	$f(x_k) = \frac{1}{n},$ $k = 1, 2, \dots, n$	$x_1 < x_2 < \dots < x_n$ $n \in \mathbb{N}$	$\frac{1}{n} \sum_{k=1}^n x_k$	$\frac{1}{n} \sum_{k=1}^n x_k^2 -$ $-\frac{1}{n} \left(\sum_{k=1}^n x_k \right)^2$
De dos puntos	$f(x_1) = p,$ $f(x_2) = 1 - p$	$x_1 < x_2$ $0 < p < 1$	$x_1 p + x_2 (1 - p)$	$(x_1 - x_2)^2 p (1 - p)$
Bernoulli	$f(0) = p,$ $f(1) = 1 - p$	p	p	$p(1 - p)$
Binomial	$f(k) = \binom{n}{k} p^k (1 - p)^{n-k}$ $k = 0, 1, 2, \dots, n$	$0 < p < 1$ $n \in \mathbb{N}$	np	$np(1 - p)$
Poisson	$f(k) = \frac{1}{k!} e^{-\lambda} \lambda^k$ $k = 0, 1, 2, 3, \dots$	$\lambda > 0$	λ	λ
Hiper-geométrica	$\frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$ $k \in \mathbb{N}_0, k \leq n, k \leq M$	$M \in \mathbb{N}_0, N \in \mathbb{N}, n \in N$ $n \leq M \leq N$	$n \cdot \frac{M}{N}$	$\left(\frac{N-n}{N-1} \right) \cdot n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N} \right)$
Binomial negativa	$f(k) = \binom{k+r-1}{r-1} p^r (1-p)^k$ $k = 0, 1, 2, \dots$	$r > 0,$ $0 < p < 1$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$
Geométrica	$f(k) = p(1-p)^k$ $k = 0, 1, 2, \dots$	$0 < p < 1$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$

Distribuciones especiales de probabilidad (Discretas) II

Ejercicio (Blanco Castañeda 2012, Pág 133, Ej.17): Supóngase que el 5% de los artículos producidos en una fábrica son defectuosos. Se escogen 15 artículos al azar y se inspeccionan, ¿cuál es la probabilidad que se encuentren a lo sumo tres artículos defectuosos?

Solución: $X :=$ “Número de artículos defectuosos”.

$$X \stackrel{d}{=} \mathcal{B}(n = 15, p = 0.05)$$

$$\begin{aligned} P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= (0.95)^{15} + \dots + \binom{15}{3} (0.05)^3 (0.95)^{15-3} \\ &= 0.4633 + 0.3658 + 0.1348 + 0.0307 = 0.9945 \end{aligned}$$

```
`P(X<=3)` <- pbinom(q = 3, size = 15, prob = 0.05)
`P(X<=3)`
```

```
## [1] 0.9945327
```


Distribuciones especiales de probabilidad (Discretas) III

Ejercicio (Blanco Castañeda 2012, Pág 131, Ej.8): En un departamento de control de calidad, se inspeccionan las unidades terminadas que provienen de una línea de ensamble. Se piensa que la proporción de unidades defectuosas es de 0.01. ¿Cuál es la probabilidad de que la quinta unidad inspeccionada sea la segunda que se encuentre defectuosa?

Solución: X := "Número de unidades que es necesario inspeccionar hasta obtener dos unidades defectuosas". $X \stackrel{d}{=} \mathcal{B}_N(k=2, p=0.01)$

Como $f(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$, $x = k, k+1, \dots$, se tiene:

$$P(X=5) = \binom{5-1}{2-1} (0.01)^2 (0.99)^{5-2} = 3.881 \times 10^{-4}$$

```
`P(X=5)` <- dbinom(5 - 2, 2, 0.01); `P(X=5)`
```

```
## [1] 0.0003881196
```

Distribuciones especiales de probabilidad (Discretas) IV

Ejercicio (Blanco Castañeda 2012, Pág 134, Ej.25): En una fábrica se producen cartuchos de tinta para estilografo. Uno de cada 30 cartuchos producidos resulta ser defectuoso. Los cartuchos son empacados en cajitas de 6 cartuchos cada una. Calcular el número esperado de cajitas que contienen, respectivamente, ningún cartucho defectuoso, un cartucho defectuoso, dos o más cartuchos defectuosos, en un envío de 1000 paquetes.

Solución: $X :=$ “Número de cartuchos defectuosos”. $X \stackrel{d}{=} \mathcal{B}(n, p)$

$$P(X = 0) = (29/30)^6 = 0.8159 \rightarrow \approx 816$$

$$P(X = 1) = 0.1688 \rightarrow \approx 169 ; P(X \geq 2) = 0.0153 \rightarrow \approx 15$$

```
`P(X=0)` <- dbinom(0, 6, prob = 1/30)
`P(X=1)` <- dbinom(1, 6, prob = 1/30)
`P(X>=2)` <- pbinom(1, 6, prob = 1/30, lower.tail = FALSE)
```

```
## [1] 0.81594420 0.16881604 0.01523976
```

Distribuciones especiales de probabilidad (Discretas) V

Ejercicio (Blanco Castañeda 2012, Pág 138, Ej.49): Un cargamento de 20 unidades de una cierta mercancía se considera de buena calidad cuando él contiene a los más dos unidades defectuosas y se considera de mala calidad si contiene por lo menos 4 unidades defectuosas. El vendedor y el comprador del cargamento acuerdan tomar una muestra aleatoria de cuatro unidades para analizarla. Si todos los elementos de la muestra son de buena calidad entonces se realiza el negocio. El vendedor corre el riesgo, bajo este acuerdo, de no vender un cargamento de buena calidad, en tanto que el comprador corre el riesgo de comprar un cargamento de mala calidad. ¿Quién corre el mayor riesgo en esta transacción? Si usted es un superintendente de comercio ¿Qué recomendaría para equilibrar los riesgos entre las partes?

Distribuciones especiales de probabilidad (Discretas) VI

Solución: $X :=$ "Número de defectuosos". $X \stackrel{d}{=} H(n, M, N)$. Si $N = 20$ entonces $M = 0, 1, \dots, 20$. Hay acuerdo si $X = 0$ en $n = 4$. Se tiene:

► Para $M = 0$, $P(X = 0) = 1 \longrightarrow P(X \geq 1) = 0$

► Para $M = 1$, $P(X = 0) = \frac{\binom{19}{4}}{\binom{20}{4}} = 0.8 \longrightarrow P(X \geq 1) = 0.2$

► Para $M = 2$, $P(X = 0) = \frac{\binom{18}{4}}{\binom{20}{4}} = 0.632 \longrightarrow P(X \geq 1) = 0.368$

\vdots

► Para $M = 16$, $P(X = 0) = \frac{\binom{4}{4}}{\binom{20}{4}} = \frac{1}{4845} \longrightarrow P(X \geq 1) = \frac{4844}{4845}$

$$P(X \geq 1, M \leq 2) = \frac{1}{21}(0.8 + 0.632) = 0.027 \longrightarrow (\text{R.Vend.})$$

$$P(X = 0, M \geq 4) = \frac{1}{21}\left(\frac{364}{969} + \dots + \frac{1}{4845}\right) = 0.061 \longrightarrow (\text{R.Compr})$$

Distribuciones especiales de probabilidad (Discretas) VII

```
N <- 20
n <- 4
# posibles alternativas de defectuosos
M <- seq(0, 20)
# prob de aceptar distr hipergeométrica
p_0 <- dhyper(x = 0, m = M, n = N-M, k = 4)
p_1m <- 1 - p_0
tt <- data.frame(N=rep(20,length(M)),
                 n=rep(4,length(M)), M=M,
                 Prob.aceptar=p_0,
                 Prob.rechazar= p_1m)

tt
R.vendedor <- sum(tt$Prob.rechazar[1:3])/21
R.vendedor
R.comprador <- sum(tt$Prob.aceptar[5:21])/21
R.comprador
```

Distribuciones especiales de probabilidad (Discretas) VIII

##	N	n	M	Prob.aceptar	Prob.rechazar
## 1	20	4	0	1.0000000000	0.00000000
## 2	20	4	1	0.8000000000	0.20000000
## 3	20	4	2	0.6315789474	0.3684211
## 4	20	4	3	0.4912280702	0.5087719
## 5	20	4	4	0.3756449948	0.6243550
## 6	20	4	5	0.2817337461	0.7182663
## 7	20	4	6	0.2066047472	0.7933953
## 8	20	4	7	0.1475748194	0.8524252
## 9	20	4	8	0.1021671827	0.8978328
## 10	20	4	9	0.0681114551	0.9318885
## 11	20	4	10	0.0433436533	0.9566563
## 12	20	4	11	0.0260061920	0.9739938
## 13	20	4	12	0.0144478844	0.9855521
## 14	20	4	13	0.0072239422	0.9927761
## 15	20	4	14	0.0030959752	0.9969040
## 16	20	4	15	0.0010319917	0.9989680
## 17	20	4	16	0.0002063983	0.9997936
## 18	20	4	17	0.0000000000	1.0000000
## 19	20	4	18	0.0000000000	1.0000000
## 20	20	4	19	0.0000000000	1.0000000
## 21	20	4	20	0.0000000000	1.0000000

```
R.vendedor <- sum(tt$Prob.rechazar[1:3])/21; R.vendedor
```

```
## [1] 0.02706767
```

```
R.comprador <- sum(tt$Prob.aceptar[5:21])/21; R.comprador
```

```
## [1] 0.06081871
```

Distribuciones especiales de probabilidad (Continuas) I

NOMBRE	FUNCION	PARA-METROS	$E(X)$	$V(X)$
Uniforme	$f(x) = \frac{1}{b-a},$ $a < x < b$	$a < b$	$\frac{a+b}{2}$	$\frac{(a-b)^2}{12}$
Normal	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $x \in \mathbb{R}$	$\mu \in \mathbb{R},$ $\sigma^2 > 0$	μ	σ^2
Normal estándar	$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$ $x \in \mathbb{R}$		0	1
Gamma	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x},$ $x > 0$	$\alpha > 0,$ $\beta > 0$	α/β	α/β^2
Exponencial	$f(x) = \lambda e^{-\lambda x},$ $x > 0$	$\lambda > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
t de Student	$f(x) = a_n \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2},$ $a_n := \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}}, \quad x \in \mathbb{R}$	$n \in \mathbb{N}$	0, $n \geq 2$	$\frac{n}{n-2},$ $n \geq 3$
Chi-cuadrada	$\frac{1}{a_n} x^{\frac{n}{2}-1} e^{-x/2},$ $a_n := 2^{n/2} \Gamma\left(\frac{n}{2}\right), \quad x > 0$	$n > 0$	n	$2n$
F de Fisher	$f(x) = \frac{a_n x^{\frac{m}{2}-1}}{(n+mx)^{(m+n)/2}}$ $a_n := \frac{\Gamma(\frac{m+n}{2}) m^{m/2} n^{n/2}}{\Gamma(\frac{m}{2}) \Gamma(\frac{n}{2})}, \quad x > 0$	$m, n \in \mathbb{N}$	$\frac{n}{n-2},$ $n \geq 3$	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)},$ $n \geq 5$

Distribuciones especiales de probabilidad (Continuas) II

Ejercicio (Blanco Castañeda 2012, Pág 174, Ej.12): El volumen de una máquina de llenado automático de latas una bebida gaseosa tiene una distribución normal con media 12.4 onzas de líquido y una desviación estándar de 0.1 onzas de líquido.

- (a) ¿Cuál es la probabilidad de que el volumen depositado sea menor que 12 onzas de líquido?
- (b) Si se desechan todas las latas que tienen menos de 12.1 o más de 12.6 onzas de líquido, ¿cuál es la proporción de latas desechadas?

Distribuciones especiales de probabilidad (Continuas) III

Solución: X := "Cantidad de onzas de líquido". $X \stackrel{d}{=} N(12.4, 0.1)$

Parte (a): $P(X < 12) = ?$

$$\begin{aligned} P(X < 12) &= P\left(Z < \frac{12 - 12.4}{0.1}\right) = P(Z < -4) \\ &= \Phi(-4) = 3.17 \times 10^{-5} \approx 0 \end{aligned}$$

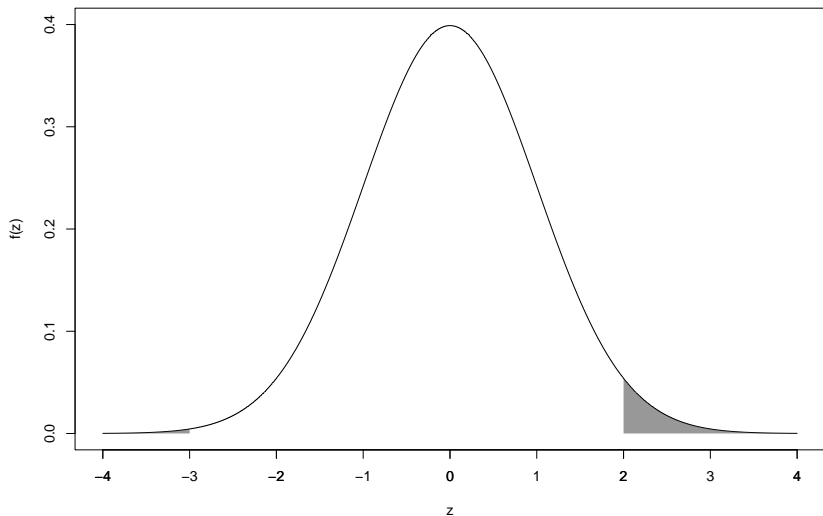
Parte (b): $P(12.1 > X > 12.6) = ?$

$$\begin{aligned} P(12.1 > X > 12.6) &= P\left(Z < \frac{12.1 - 12.4}{0.1}\right) + P\left(Z > \frac{12.6 - 12.4}{0.1}\right) \\ &= P(Z < -3) + [1 - P(Z < 2)] \\ &= \Phi(-3) + (1 - \Phi(2)) \\ &= 0.0013 + [1 - 0.9772] = 0.0241 \approx 2.41\% \end{aligned}$$

```
`P(12.1>X>12.6)` <- pnorm(12.1, mean = 12.4, sd = 0.1) +  
  pnorm(12.6, mean = 12.4, sd = 0.1, lower.tail = F)
```

Distribuciones especiales de probabilidad (Continuas) IV

$$P(-3 > Z > 2) = 0.0241$$



Distribuciones especiales de probabilidad (Continuas) V

Ejercicio (Canavos 1988, Pág 176, Ej.5.8): Una fábrica produce pistones cuyos diámetros se encuentran adecuadamente clasificados por una distribución normal con un diámetro promedio de 5 cm y una desviación estándar igual a 0.001 cm. Para que un pistón sirva, su diámetro debe encontrarse entre 4.998 y 5.002 cm. Si el diámetro del pistón es menor que 4.998 se desecha; si es mayor que 5.002 puede reprocesarse. ¿Qué porcentaje de pistones servirá? ¿Qué porcentaje será desechado? ¿Qué porcentaje será reprocesado?

Distribuciones especiales de probabilidad (Continuas) VI

Solución: $X :=$ "Diametro en cm de los pistones". $X \stackrel{d}{=} N(5, 0.001)$.

- Pistones que sirven: $P(4.998 < X < 5.002) = ?$

$$\begin{aligned} P(4.998 < X < 5.002) &= P(-2 < Z < 2) \\ &= P(Z < 2) - P(Z < -2) \\ &= \Phi(2) - \Phi(-2) \\ &= 0.9772 - 0.0228 = 0.9544 \approx (95.44\%) \end{aligned}$$

- Pistones que se desechan: $P(X < 4.998) = ?$

$$\begin{aligned} P(X < 4.998) &= P(Z < -2) \\ &= \Phi(-2) = 0.0228 \approx (2.28\%) \end{aligned}$$

- Pistones que se reprocesan: $P(X > 5.002) = 0.0228 \approx (2.28\%)$

Distribuciones especiales de probabilidad (Continuas) VII

```
mu <- 5
sig <- 0.001
`P(4.998<X<5.002)` <- pnorm(5.002, mean = mu, sd = sig) -
  pnorm(4.998, mean = mu, sd = sig)
`P(4.998<X<5.002)`*100 # % pistones que sirven
```

```
## [1] 95.44997
```

```
`P(X<4.998)` <- pnorm(4.998, mean = mu, sd = sig)
`P(X<4.998)`*100 # % pistones que se desechan
```

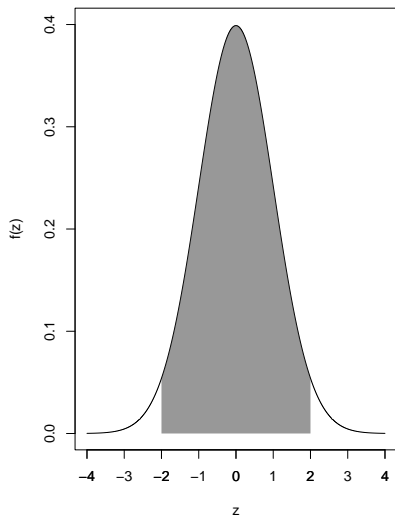
```
## [1] 2.275013
```

```
`P(X>5.002)` <- pnorm(5.002, mean = mu, sd = sig,
  lower.tail = FALSE)
`P(X>5.002)`*100 # % pistones que se reprocesan
```

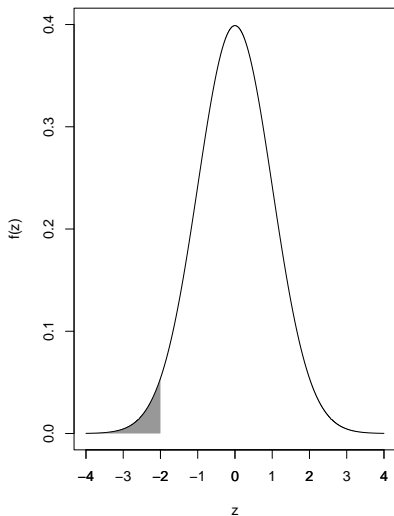
```
## [1] 2.275013
```

Distribuciones especiales de probabilidad (Continuas) VIII

$$P(-2 < Z < 2) = 0.9545$$



$$P(Z < -2) = 0.0228 = P(Z > 2)$$



Capacidad de proceso

Índices de capacidad de proceso I

Para este tópico tomaremos como referencia a Gutiérrez Pulido y Vara Salazar (2009) y Vargas (2006)...

- ▶ El análisis de capacidad de procesos evalúa la habilidad que tiene un proceso de cumplir con las especificaciones que fija el cliente.
- ▶ Un análisis de capacidad de proceso, responde al productor la pregunta: ¿Puedo ofrecer lo que el consumidor pide?
- ▶ Es decir, la idea es comparar las especificaciones del producto con la variabilidad (sin descuidar el centramiento) del proceso.

Índices de capacidad de proceso II

Índice C_p .

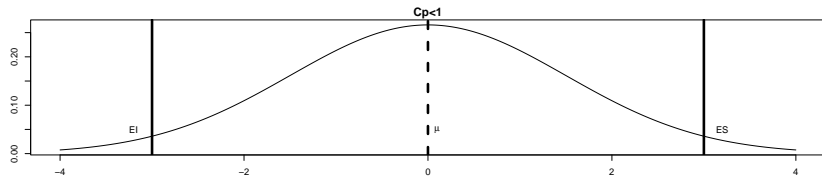
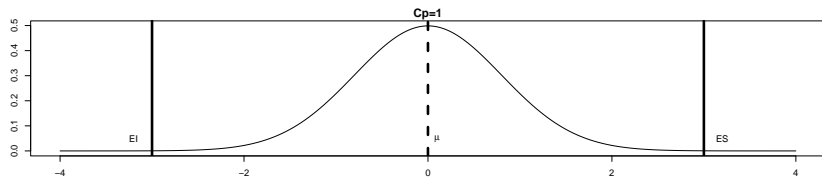
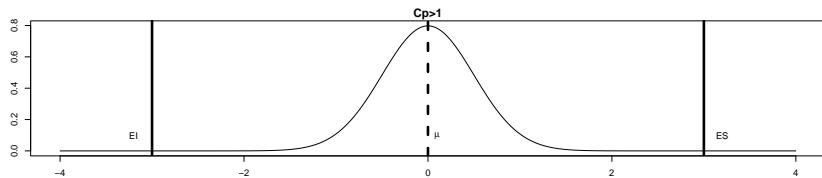
Debido a que la curva normal tiene 6 desviaciones estándar de amplitud, se tiene:

$$C_p = \frac{LES - LEI}{6\sigma},$$

donde, LES y LEI son los límites de especificación superior e inferior, respectivamente y σ la desviación estándar del proceso. Algunas observaciones de este índice son:

- ▶ Cuanto mayor sea el valor de este índice se dice que el proceso cumplirá con las especificaciones. Se dice que un $C_p = 1$ tendría buena pinta, pero en contextos empresariales se tienen estándares de 1.5 y 2.0.
- ▶ Si σ es desconocido, se debe usar un estimador en el denominador.

Índices de capacidad de proceso III



Índices de capacidad de proceso IV

Índice C_{pk} .

El C_p es útil cuando μ es igual al valor nominal objetivo N . Como una medida de protección ante los desplazamientos de la media, se define los siguientes:

$$C_{pi} = \frac{\mu - LEI}{3\sigma}, \quad \text{y} \quad C_{ps} = \frac{LES - \mu}{3\sigma},$$

donde μ es la media del proceso. Entonces se puede definir el siguientes índice:

$$C_{pk} = \min\{C_{pi}, C_{ps}\}.$$

En muchas ocasiones, es posible tener límites de especificación de un solo lado. Por ejemplo: $C_{pk} = C_{ps}$ si se analiza el porcentaje de impurezas en un proceso químico.

Índices de capacidad de proceso V

Si $d = (LES - LEI)/2$ y $M = (LES + LEI)/2$, se puede probar que:

$$C_{pk} = C_p \left\{ 1 - \frac{|\mu - M|}{d} \right\}$$

Índice C_{pm} (Índice de Taguchi).

Este índice involucra en la variabilidad del proceso el desvío de la media con respecto al valor objetivo N y se define:

$$C_{pm} = \frac{LES - LEI}{6\sqrt{\sigma^2 + (\mu - N)^2}}.$$

Dividiendo ambos términos por σ , se obtiene:

$$C_{pm} = \frac{C_p}{\sqrt{1 + \left[\frac{\mu - N}{\sigma} \right]^2}}.$$

Índices de capacidad de proceso VI

Cabe resaltar que $C_{pm} > C_p$. No proporciona buenos resultados si $N \neq M$.

Índice C_{pmk} .

Pearn, Kotz, y Johnson (1992) introdujeron un índice que rescata las mejores propiedades de C_p , C_{pk} y C_{pm} . Se define de la siguiente manera:

$$C_{pmk} = \frac{\min(\mu - LEI, LES - \mu)}{3\sqrt{\sigma^2 + (\mu - N)^2}}.$$

Si $\mu = M$, entonces $C_{pm} = C_{pmk}$ sin importar $|\mu - N|$. De lo contrario $C_{pmk} < C_{pm}$.

A partir de lo anterior se pueden establecer las siguientes relaciones:

$$C_{pmk} \leq C_{pk} \leq C_p$$

Índices de capacidad de proceso VII

y

$$C_{pmk} \leq C_{pm} \leq C_p.$$

No es clara la relación entre C_{pk} y C_{pm} .

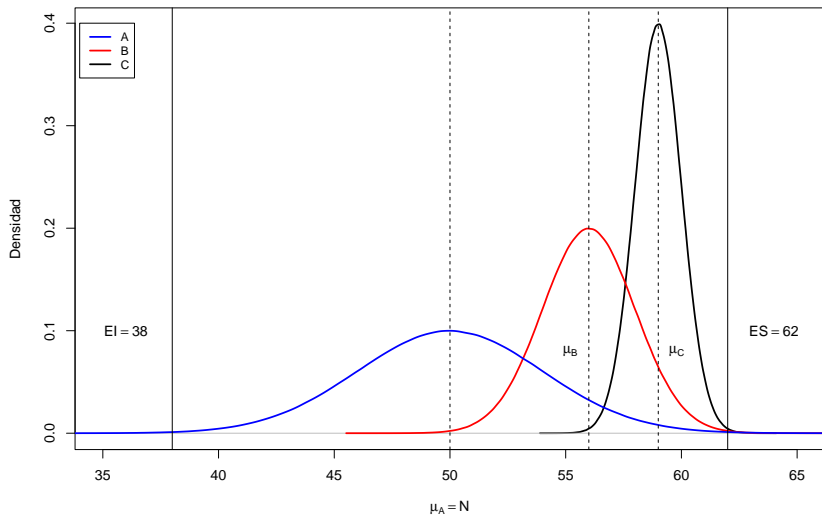
Ejemplo: Considere de a partir de datos simulados de tres procesos A, B y C se obtuvieron los parámetros que están en la tabla:

Proceso	μ	σ
A	50	4
B	56	2
C	59	1

Asuma que $LES = 62$ mm, $LEI = 38$ mm y $N = 50$ mm. Calcule C_p , C_{pk} , C_{pm} y C_{pmk} . A continuación se pueden ver los procesos:

Índices de capacidad de proceso VIII

Gráfico comparativo de los procesos



Índices de capacidad de proceso IX

Proceso	μ	σ	C_p	C_{pk}	C_{pm}	C_{pmk}
A	50	4	1	1	1	1
B	56	2	2	1	0.6324	0.3162
C	59	1	4	1	0.4417	0.1104

Índices de capacidad de proceso X

Estimación de los índices de capacidad

En la mayoría de las veces los parámetros que involucran los índices son desconocidos, por ende, hay que estimarlos. Por ejemplo, en el cálculo del C_p , se podría usar S en lugar de σ . O, también $\frac{\bar{R}}{d_2}$. Un estimador para C_p es:

$$\hat{C}_p = \frac{LES - LEI}{6\hat{\sigma}}.$$

Usando $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$ un intervalo de confianza $(1 - \alpha)100\%$ para C_p es:

$$\hat{C}_p \sqrt{\frac{\chi^2_{(n-1, \frac{\alpha}{2})}}{n-1}} \leq C_p \leq \hat{C}_p \sqrt{\frac{\chi^2_{(n-1, 1-\frac{\alpha}{2})}}{n-1}}.$$

Índices de capacidad de proceso XI

Un estimador para C_{pk} es:

$$\widehat{C}_{pk} = \widehat{C}_p \left\{ 1 - \frac{|\bar{X} - M|}{d} \right\},$$

donde \bar{X} es la media muestral.

Bissell (1990) propuso el siguiente intervalo de confianza para C_{pk} :

$$\begin{aligned} \widehat{C}_{pk} \left[1 - z_{\frac{\alpha}{2}} \sqrt{\frac{1}{9n\hat{C}_{pk}^2} + \frac{1}{2(n-1)}} \right] &\leq C_{pk} \\ &\leq \widehat{C}_{pk} \left[1 + z_{\frac{\alpha}{2}} \sqrt{\frac{1}{9n\hat{C}_{pk}^2} + \frac{1}{2(n-1)}} \right]. \end{aligned}$$

Índices de capacidad de proceso XV

Tiene mucha lógica usar como \widehat{C}_{pm} la especificación de C_{pm} pero con los estimadores S^2 y \bar{X} . Sin embargo, Chan, Cheng, y Spiring (1988) lo critica y propone:

$$\widehat{C}_{pm} = \frac{LES - LEI}{6\sqrt{\sum_{i=1}^n \frac{(X_i - N)^2}{n-1}}}.$$

Un intervalo de confianza $(1 - \alpha)100\%$ para C_{pm} es:

$$\widehat{C}_{pm} \sqrt{\frac{\chi^2_{(k, \frac{\alpha}{2})}}{k}} \leq C_{pm} \leq \widehat{C}_{pm} \sqrt{\frac{\chi^2_{(k, 1 - \frac{\alpha}{2})}}{k}},$$

donde:

$$k = \frac{n \left[1 + \left(\frac{\bar{X} - N}{S} \right)^2 \right]}{\left[1 + 2 \left(\frac{\bar{X} - N}{S} \right)^2 \right]}$$

Índices de capacidad de proceso XVI

Luego, un estimador para C_{pmk} se tiene:

$$\widehat{C_{pmk}} = \frac{d - |\bar{X} - N|}{3\sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 + (\bar{X} - N)^2}}.$$

Índices de capacidad de proceso XVII

Proporción de unidades no conformes

Si $X \sim N(\mu, \sigma^2)$, la proporción p de unidades no conformes se define:

$$\begin{aligned} p &= P(X < LEI, \text{ o } , X > LES) \\ &= P\left(Z < \frac{LEI - \mu}{\sigma}\right) + P\left(Z > \frac{LES - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{LEI - \mu}{\sigma}\right) + \left[1 - \Phi\left(\frac{LES - \mu}{\sigma}\right)\right] \\ &= 1 - \left[\Phi\left(\frac{LES - \mu}{\sigma}\right) - \Phi\left(\frac{LEI - \mu}{\sigma}\right)\right]. \end{aligned}$$

Adicionalmente, si se asume que $\mu = M = \frac{LES+LEI}{2}$. Si $C_p = 1$, implica que $LES - \mu = -(LEI - \mu) = 3\sigma$. Reemplazando en la expresión anterior se obtiene que $p = 0.27\%$ o 27 p.p.m. Con $C_p = 2$ a calidad “6 sigma” ≈ 3.4 p.p.m.

Índices de capacidad de proceso X

Ejemplo (Gutiérrez Pulido y Vara Salazar 2009, Pág 133, Ej.5): Los siguientes datos representan las mediciones de viscosidad de los últimos tres meses de un producto lácteo. El objetivo es tener una viscosidad de 80 ± 10 cps.

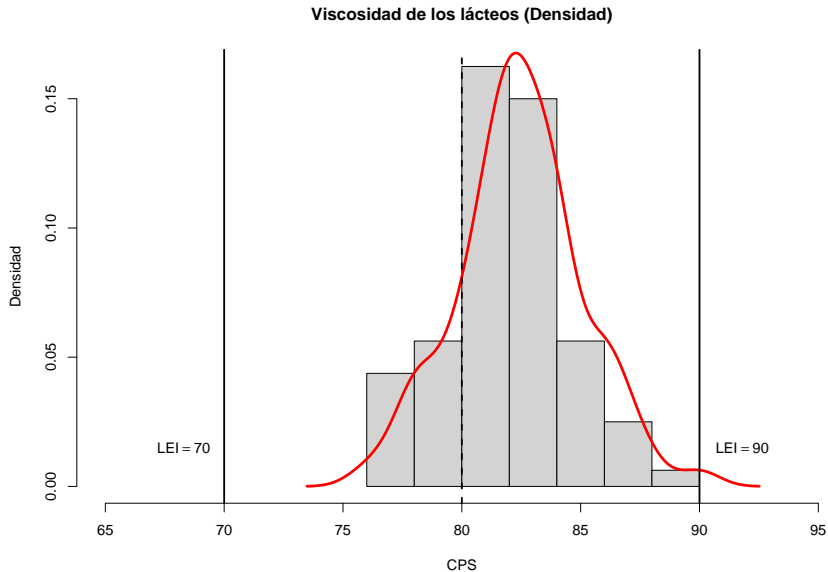
84	81	77	80	80	82	78	83
81	78	83	84	85	84	82	84
82	80	83	84	82	78	83	81
86	85	79	86	83	82	84	82
83	82	84	86	81	82	81	82
87	84	83	82	81	84	84	81
78	83	83	80	86	83	82	86
87	81	78	81	82	84	83	79
80	82	86	82	80	83	82	76
79	81	82	84	85	87	88	90

Índices de capacidad de proceso XI

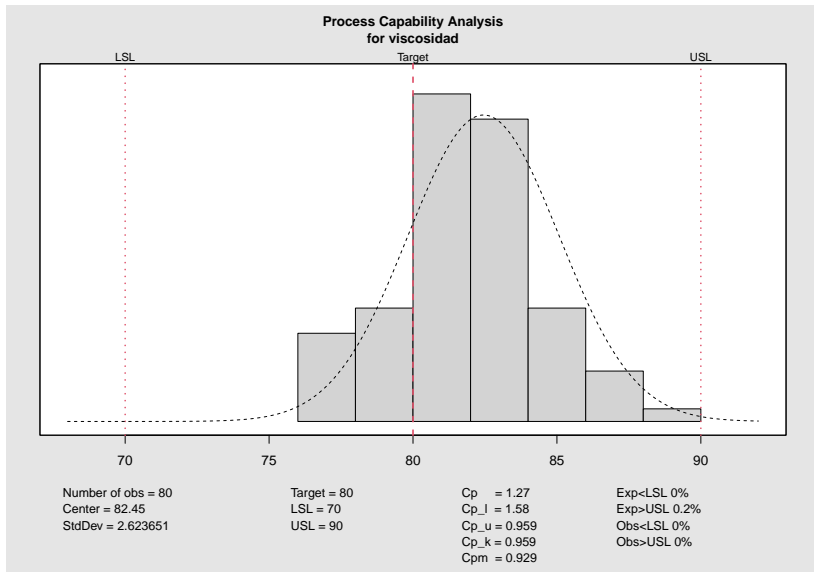
- (a) Construya una gráfica de capacidad de proceso (Histograma con tolerancias).
- (b) Calcule la media y desviación estándar, y tomando a éstos como parámetros poblacionales estime los índices C_p , C_{pk} , C_{pm} y C_{pmk} , y sus intervalos de confianza.
- (c) Estime el porcentaje fuera de las especificaciones.
- (d) ¿Por qué las estimaciones realizadas se deben ver con ciertas reservas?

Índices de capacidad de proceso XII

Solución:



Índices de capacidad de proceso XIII I



Índices de capacidad de proceso XIII II

##

Process Capability Analysis

##

Call:

process.capability(object = qq, spec.limits = c(70, 90), target = 80, st

##

Number of obs = 80 Target = 80

Center = 82.45 LSL = 70

StdDev = 2.624 USL = 90

##

Capability indices:

##

	Value	2.5%	97.5%
--	-------	------	-------

## Cp	1.2705	1.0726	1.468
-------	--------	--------	-------

## Cp_l	1.5818	1.3659	1.798
---------	--------	--------	-------

## Cp_u	0.9592	0.8195	1.099
---------	--------	--------	-------

## Cp_k	0.9592	0.7928	1.126
---------	--------	--------	-------

## Cpm	0.9286	0.7547	1.102
--------	--------	--------	-------

##

Exp<LSL 0% Obs<LSL 0%

Exp>USL 0.2% Obs>USL 0%

Índices de capacidad de proceso XIV

```
# código que genera los índices es R
qq <- qcc(viscosidad,
          type = "xbar.one",
          plot = F)
process.capability(
  qq,
  spec.limits = c(70, 90),
  target = 80,
  std.dev = sd(viscosidad)
)
```

Índices de capacidad de proceso XV

Ejemplo (Gutiérrez Pulido y Vara Salazar (2009), Pág 120, Ejemplo 5.7): Un componente fundamental en las ruedas de automóvil son las llantas. Éstas tienen una capa que en determinada referencia debe tener una longitud de 550 mm con una tolerancia de ± 8 mm. Esto indica que el valor nominal es 550 mm con $ES = 558$ y $EL = 542$. El objeto `tiros.Rdata` contiene 36 muestras cada una de 5 ejemplares, para un total de 180 lecturas. Haga un estudio integral de capacidad de proceso.

Índices de capacidad de proceso XVI

Partes de un estudio integral de capacidad de proceso

1. Medidas de tendencia central (Media, mediana y moda)
2. Medidas de dispersión (Desviación estándar) y rango medio.
3. Comparación entre límites reales con las especificaciones ($\mu \pm 3\sigma$)
4. Gráfica de capacidad (Histograma con N , EL y ES)
5. Porcentaje fuera de las especificaciones ($P(X < LEI \text{ o } X > LES) = ?$)
6. Índices de capacidad (C_p , C_{pk} , C_{pm} y C_{pmk} , con los intervalos de confianza.)

Índices de capacidad de proceso XVII I

```
# cargar datos
load("data/tires.Rdata")

# especificaciones
EI <- 542; ES <- 558; N <- 550

# medidas de tendencia central y dispersión
longitudes <- tires$longitud # vector de datos
tt <- hist(x = longitudes, plot=FALSE)
TT <- table.freq(tt)
TT # tabla de datos agrupada
```

##	Lower	Upper	Main	Frequency	Percentage	CF	CPF
## 1	548	549	548.5	11	6.1	11	6.1
## 2	549	550	549.5	16	8.9	27	15.0
## 3	550	551	550.5	26	14.4	53	29.4
## 4	551	552	551.5	39	21.7	92	51.1
## 5	552	553	552.5	34	18.9	126	70.0
## 6	553	554	553.5	26	14.4	152	84.4
## 7	554	555	554.5	12	6.7	164	91.1
## 8	555	556	555.5	12	6.7	176	97.8
## 9	556	557	556.5	3	1.7	179	99.4
## 10	557	558	557.5	1	0.6	180	100.0

Índices de capacidad de proceso XVII II

```
summary(longitudes)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      548.0   551.0   552.0   552.5   554.0   558.0
```

```
# estimación de la moda para datos agrupados
```

```
moda <- 551 + (((39-26)/((39-26)+(39-34)))*(552-551)
```

```
moda
```

```
## [1] 551.7222
```

```
# estimación del rango medio
```

```
aux <- qcc.groups(data = tires$longitud, sample = tires$subgrupo)
```

```
rango_medio <- mean(apply(aux, 1, function(x) max(x) - min(x)))
```

```
rango_medio
```

```
## [1] 4.694444
```

```
# desviación a largo plazo
```

```
stand.deviaclp <- sd(longitudes); stand.deviaclp
```

```
## [1] 1.998323
```

```
stand.deviaccp <- rango_medio/2.326; stand.deviaccp
```

```
## [1] 2.018248
```

Índices de capacidad de proceso XVII III

```
# límites reales
LRI <- mean(longitudes)-3*stand.deviaccp
LRS <- mean(longitudes)+3*stand.deviaccp
c(LRI=LRI,LRS=LRS)
```

```
##          LRI          LRS
## 546.4786 558.5881
```

```
# proporción de no conformes en %
nc <- pnorm(EI, mean = mean(longitudes), sd = stand.deviaccp) +
      pnorm(ES, mean = mean(longitudes), sd = stand.deviaccp,
            lower.tail = FALSE)
nc*100
```

```
## [1] 0.3378272
```

```
# calculo del Cpmk
C_pmk <- min(ES-mean(longitudes),mean(longitudes)-EI)/
  (3*sqrt(stand.deviaccp^2+(mean(longitudes)-N)^2))
C_pmk
```

```
## [1] 0.5625884
```

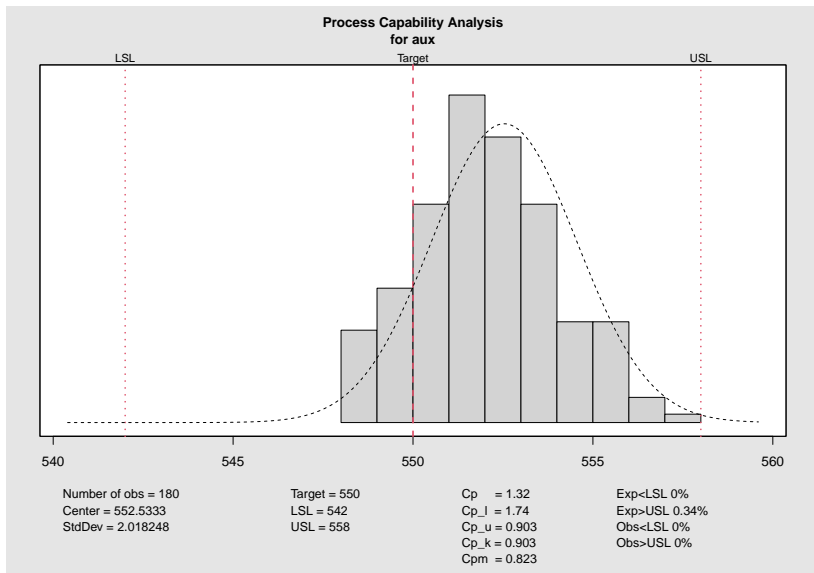
```
# construcción del objeto qcc
xx <- qcc(data = aux, type = "xbar", plot = FALSE)
str(xx,1)
```


Índices de capacidad de proceso XVII IV

```
## List of 11
## $ call      : language qcc(data = aux, type = "xbar", plot = FALSE)
## $ type      : chr "xbar"
## $ data.name : chr "aux"
## $ data      : num [1:36, 1:5] 556 556 556 552 551 552 551 550 556 554 ...
## ..- attr(*, "dimnames")=List of 2
## $ statistics: Named num [1:36] 553 554 553 553 552 ...
## ..- attr(*, "names")= chr [1:36] "1" "2" "3" "4" ...
## $ sizes     : Named int [1:36] 5 5 5 5 5 5 5 5 5 5 ...
## ..- attr(*, "names")= chr [1:36] "1" "2" "3" "4" ...
## $ center    : num 553
## $ std.dev   : num 2.02
## $ nsigmas   : num 3
## $ limits    : num [1, 1:2] 550 555
## ..- attr(*, "dimnames")=List of 2
## $ violations:List of 2
## - attr(*, "class")= chr "qcc"

# función de capacidad de procesos.
yy <- process.capability(object = xx, spec.limits = c(EI, ES))
```

Índices de capacidad de proceso XVII V



Índices de capacidad de proceso XVII VI

```
##
## Process Capability Analysis
##
## Call:
## process.capability(object = xx, spec.limits = c(EI, ES))
##
## Number of obs = 180          Target = 550
##      Center = 552.5          LSL = 542
##      StdDev = 2.018          USL = 558
##
## Capability indices:
##
##      Value    2.5%    97.5%
## Cp      1.3213  1.1844  1.4579
## Cp_l    1.7397  1.5830  1.8963
## Cp_u    0.9029  0.8144  0.9914
## Cp_k    0.9029  0.7974  1.0083
## Cpm     0.8233  0.7154  0.9310
##
## Exp<LSL 0%      Obs<LSL 0%
## Exp>USL 0.34%   Obs>USL 0%
```

Aspectos básicos de R

Instalando R

- ▶ Ingresar al siguiente link.
- ▶ Descargar el paquete ejecutable según el sistema operativo sobre el que se trabaja.
- ▶ Descomprimir (en caso de que sea necesario) e instalar.

¿Cómo funciona R? (Parte I)

- ▶ R es un lenguaje orientado a objetos, por ello es simple y flexible.
- ▶ R es un lenguaje interpretado (como Java) y no compilado (como C, C++, Fortran, Pascal, ...), lo cual significa que los comandos escritos en el teclado son ejecutados directamente sin necesidad de construir ejecutables.
- ▶ La sintaxis de R es muy simple e intuitiva. Por ejemplo: una regresión se ejecuta `lm(y~x)`.
- ▶ Para que una función sea ejecutada en R debe estar siempre acompañada de paréntesis, inclusive en el caso que no haya nada dentro de los mismos (por ej., `ls()`)

¿Cómo funciona R? (Parte II)

- ▶ Si se escribe el nombre de la función sin los paréntesis, R mostrará el contenido (código) mismo de la función.

```
lapply
## function (X, FUN, ...)
## {
##     FUN <- match.fun(FUN)
##     if (!is.vector(X) || is.object(X))
##         X <- as.list(X)
##     .Internal(lapply(X, FUN))
## }
## <bytecode: 0x12b934fd0>
## <environment: namespace:base>
```

¿Cómo funciona R? (Parte III)

- ▶ **Orientado a Objetos** significa que las variables, datos, funciones, resultados, etc., se guardan en la memoria activa del computador en forma de objetos con un nombre específico.
- ▶ Una función en R funciona de la siguiente manera:

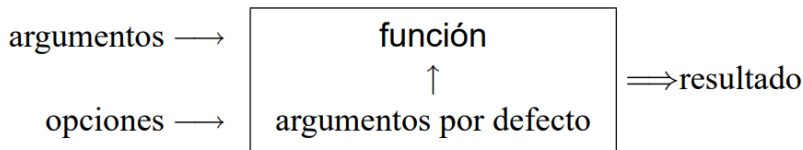


Figura 7: Esquema de función en R

Creación, listado y remoción de objetos en memoria I

- Un objeto puede ser creado con el operador “asignar” el cual se denota como una flecha con el signo menos y el símbolo “>” o “<” dependiendo de la dirección en que asigna el objeto:

```
n <- 10
```

```
n
```

```
## [1] 10
```

```
10 -> n
```

```
n
```

```
## [1] 10
```

Creación, listado y remoción de objetos en memoria II

- ▶ Si el objeto ya existe, y se vuelve a realizar una asignación toma como valor esta última:

```
n <- rpois(10,1)
str(n)
```

```
##  int [1:10] 3 2 0 2 3 0 2 0 0 1
```

```
class(n)
```

```
## [1] "integer"
```

```
n <- diag(3)
str(n)
```

```
##  num [1:3, 1:3] 1 0 0 0 1 0 0 0 1
```

```
class(n)
```

```
## [1] "matrix" "array"
```

Creación, listado y remoción de objetos en memoria III

- ▶ Se puede escribir una expresión sin asignar su valor a un objeto; en este caso el resultado sera visible en la pantalla pero no será guardado en memoria:

```
(10 + 2) * 5
```

```
## [1] 60
```

- ▶ La asignación será omitida de los ejemplos si no es necesaria para la comprensión del mismo. La función `ls` simplemente lista los objetos en memoria: solo se muestran los nombres de los mismos.

```
name <- "Carmen"; n1 <- 10; n2 <- 100; m <- 0.5  
# ls()
```

Creación, listado y remoción de objetos en memoria IV

La ayuda en línea de R proporciona información muy útil de cómo utilizar las funciones. La ayuda se encuentra disponible directamente para una función dada. Por ejemplo:

```
?lm
```

```
#help(lm)
```

```
#help("lm") para caracteres no convencionales
```

Ver página de ayuda o documentación [aquí](#)

References I

- Bissell, AF. 1990. «How reliable is your capability index?» *Journal of the Royal Statistical Society Series C: Applied Statistics* 39 (3): 331-40. <https://doi.org/https://doi.org/10.2307/2347383>.
- Blanco Castañeda, Liliana. 2012. *Probabilidad*. Editorial UN. Universidad Nacional de Colombia-Facultad de Ciencias.
- Canavos, George. 1988. *Probabilidad y estadística, aplicaciones y métodos*. McGraw Hill.
- Chan, Lai K, Smiley W Cheng, y Frederick A Spiring. 1988. «A new measure of process capability: Cpm». *Journal of Quality Technology* 20 (3): 162-75. <https://doi.org/https://doi.org/10.1080/00224065.1988.11979102>.
- Duncan, AJ. 1987. *Quality Control and Industrial Statistics*. Irwin.
- Gutiérrez Pulido, Humberto, y Román de la Vara Salazar. 2009. *Control estadístico de la calidad y Seis Sigma*. McGraw Hill Education.
- Hoyer, RW, y B. Hoyer. 2001. «What is quality». *Quality progress* 34 (7): 53-62.

References II

- Kumar, S. A., y N. Suresh. 2009. *Operations Management*. New Age International.
<https://books.google.com.co/books?id=LYtmlR8UwjUC>.
- Montgomery, Douglas C. 2001. *Introduction to statistical quality control*. John Wiley & Sons.
- Pearn, Wen Lea, Samuel Kotz, y Norman L Johnson. 1992.
«Distributional and inferential properties of process capability indices». *Journal of Quality Technology* 24 (4): 216-31.
<https://doi.org/https://doi.org/10.1080/00224065.1992.11979403>.
- Ryan, Thomas P. 2000. *Statistical methods for quality improvement*. John Wiley & Sons.
- Vargas, JA. 2006. *Control estadístico de calidad*. Universidad Nacional de Colombia, Bogotá, Colombia.

References III

Zachary G. Stoumbos, Thomas P. Ryan, Marion R. Reynolds Jr., y William H. Woodall. 2000. «The State of Statistical Process Control as We Proceed into the 21st Century». *Journal of the American Statistical Association* 95 (451): 992-98.
<https://doi.org/10.1080/01621459.2000.10474292>.