¿De dónde salen las expresiones para calcular los límites de control en los gráficos $\bar{X}$ -*R*? ¿Y en los gráficos *I – MR*?

X es una variable aleatoria que sigue una normal centrada en μ y con desviación tipo σ.
$x \sim N(\mu\, ;\sigma)$

Por tanto: $E(x) = \mu$ y $E(s^2) = \sigma^2$.

La desviación tipo de la muestra *s* no es un estimador insesgado de la desviación tipo de la población σ. Se puede demostrar:

$$E(s) = \sqrt{\frac{2}{n-1}}\,\frac{\Gamma\left(n/2\right)}{\Gamma\left((n-1)/2\right)}\,\sigma = c_4\sigma$$

c_4 depende únicamente del tamaño de la muestra *n*, y sus valores están tabulados en la tabla 1.

Como $E(s) = c_4\sigma$, podemos obtener un estimador insesgado de la desviación tipo de la población σ:

$$\hat{\sigma} = \frac{s}{c_4}$$

Consideremos ahora que $x_1, x_2, \ldots, x_n$ es una muestra aleatoria de *n* observaciones de una distribución normal con media μ y varianza σ^2. El rango *R* de la muestra se calcula $R = max(x_i) - min(x_i)$.

Definimos una nueva variable aleatoria denominada rango relativo: $W = \dfrac{R}{\sigma}$

La distribución de *W* es conocida. El apéndice I muestra el aspecto de la distribución *W*.

La esperanza matemática de *W* es:

$$E(W) = E\left(\frac{R}{\sigma}\right) = \frac{1}{\sigma}E(R) = d_2 \Rightarrow E(R) = d_2\sigma$$

Los valores de d_2 dependen del tamaño de muestra, y están en la tabla 1.

Podemos obtener ahora un estimador de la desviación tipo de la población, donde intervendrá el rango:

$$\frac{1}{\sigma}E(R) = d_2 \Rightarrow \hat{\sigma} = \frac{R}{d_2}$$

La varianza de *W* es:

$$V(W) = V\left(\frac{R}{\sigma}\right) = \frac{1}{\sigma^2}V(R) = d_3^{\,2} \Rightarrow V(R) = d_3^{\,2}\sigma^2 \Rightarrow \hat{\sigma}_R = d_3\sigma$$

Los valores de d_3 dependen del tamaño de muestra y se ofrecen en la tabla 1.

Gráfico $\bar{X}$

Los límites de control en un gráfico $\bar{x}$ son:

$$LCS = \mu + 3\frac{\sigma}{\sqrt{n}}$$

$$LC = \mu$$

$$LCI = \mu - 3\frac{\sigma}{\sqrt{n}}$$

Llamando $A = \frac{3}{\sqrt{n}}$ y usando $\hat{\sigma} = \frac{R}{d_2}$, podemos definir una nueva constante $A_2 = \frac{A}{d_2}$, que depende únicamente del tamaño de muestra. Los valores de A y de A_2 están en la tabla 1.

Por tanto, los límites de control quedan, en función del rango:

$$LCS = \mu + A_2R$$

$$LC = \mu$$

$$LCI = \mu - A_2R$$

Gráfico *R*

Sabemos que la esperanza matemática de *R* es $d_2\sigma$ y la desviación tipo de *R* es $d_3\sigma$.

Los límites de control en un gráfico *R* son por tanto:

$$LCS = d_2\sigma + 3d_3\sigma$$

$$LC = d_2\sigma$$

$$LCI = d_2\sigma - 3d_3\sigma$$

Si definimos las constantes D_1 y D_2, que sólo dependen del tamaño de muestra, tenemos:

$$D_1 = d_2 - 3d_3$$

$$D_2 = d_2 + 3d_3$$

Y nos queda:

$LCS = D_2\sigma$
$LC = d_2\sigma$
$LCI = D_1\sigma$

Recordando que $E(R) = d_2\sigma$, y definiendo unas nuevas constantes D_3 y D_4 $\left(D_3 = D_1/d_2 \text{ y } D_4 = D_2/d_2\right)$, tenemos:

$LCS = D_4R$
$LC = R$
$LCI = D_3R$

Gráfico *I*

Los límites de control en un gráfico de valores individuales I son:
$LCS = \mu + 3\sigma$
$LC = \mu$
$LCI = \mu - 3\sigma$

Recordando que $\hat{\sigma} = \dfrac{R}{d_2}$ tenemos:

$LCS = \mu + 3\dfrac{R}{d_2}$

$LC = \mu$

$LCI = \mu - 3\dfrac{R}{d_2}$

Gráfico *MR*

Los límites de control para el gráfico de rangos móviles (*MR*) son los mismos que para el gráfico *R*. La longitud del rango móvil coincide con el tamaño de muestra.

Tabla 1. Constantes para construir gráficos de control para variables.

n	c_4	$1/c_4$	d_2	$1/d_2$	d_3	A	A_2	D_1	D_2	D_3	D_4
2	0,7979	1,2533	1,128	0,8865	0,853	2,121	1,880	0	3,686	0	3,267
3	0,8862	1,1284	1,693	0,5907	0,888	1,732	1,023	0	4,358	0	2,575
4	0,9213	1,0854	2,059	0,4857	0,880	1,500	0,729	0	4,698	0	2,282
5	0,9400	1,0638	2,326	0,4299	0,864	1,342	0,577	0	4,918	0	2,115
6	0,9515	1,0510	2,534	0,3946	0,848	1,225	0,483	0	5,078	0	2,004
7	0,9594	1,0423	2,704	0,3698	0,833	1,134	0,419	0,204	5,204	0,076	1,924
8	0,9650	1,0363	2,847	0,3512	0,820	1,061	0,373	0,388	5,306	0,136	1,864
9	0,9693	1,0317	2,970	0,3367	0,808	1,000	0,337	0,547	5,393	0,184	1,816
10	0,9727	1,0281	3,078	0,3249	0,797	0,949	0,308	0,687	5,469	0,223	1,777
11	0,9754	1,0252	3,173	0,3152	0,787	0,905	0,285	0,811	5,535	0,256	1,744
12	0,9776	1,0229	3,258	0,3069	0,778	0,866	0,266	0,922	5,594	0,283	1,717
13	0,9794	1,0210	3,336	0,2998	0,770	0,932	0,249	1,025	5,647	0,307	1,693
14	0,9810	1,0194	3,407	0,2935	0,763	0,802	0,235	1,118	5,696	0,328	1,672
15	0,9823	1,0180	3,472	0,2880	0,756	0,775	0,223	1,203	5,741	0,347	1,653
16	0,9835	1,0168	3,532	0,2831	0,750	0,750	0,212	1,282	5,781	0,363	1,637
17	0,9845	1,0157	3,588	0,2787	0,744	0,728	0,203	1,356	5,820	0,378	1,622
18	0,9854	1,0148	3,640	0,2747	0,739	0,707	0,194	1,424	5,856	0,391	1,608
19	0,9862	1,0140	3,689	0,2711	0,734	0,688	0,187	1,487	5,891	0,403	1,597
20	0,9869	1,0133	3,735	0,2677	0,729	0,671	0,180	1,549	5,921	0,415	1,585
21	0,9876	1,0126	3,778	0,2647	0,724	0,655	0,173	1,605	5,951	0,425	1,575
22	0,9882	1,0119	3,819	0,2618	0,720	0,640	0,167	1,659	5,979	0,434	1,566
23	0,9887	1,0114	3,858	0,2592	0,716	0,626	0,162	1,710	6,006	0,443	1,557
24	0,9892	1,0109	3,895	0,2567	0,712	0,612	0,157	1,759	6,031	0,451	1,548
25	0,9896	1,0105	3,931	0,2544	0,708	0,600	0,153	1,806	6,056	0,459	1,541

APÉNDICE I: Distribución del rango relativo *W*.

Supongamos que tenemos un proceso de llenado de botellas de agua. El contenido de las botellas de agua en ml es una variable aleatoria Xque se distribuye siguiendo una normal con media 500 ml y desviación tipo 5 ml: $X \sim N(500\,;5)$

Usando MINITAB, vamos a simular la extracción de 10000 muestras de tamaño 4 de este proceso. Para cada muestra, calcularemos su rango, y los iremos guardando.

Éste es el aspecto de la macro para $n = 4$, guardando los valores de los rangos en *C2*.

```
Random 4 c1;
  Normal 500 5.
Range C1 K1.
Let C2(K100)=K1
Let K100=K100+1
```

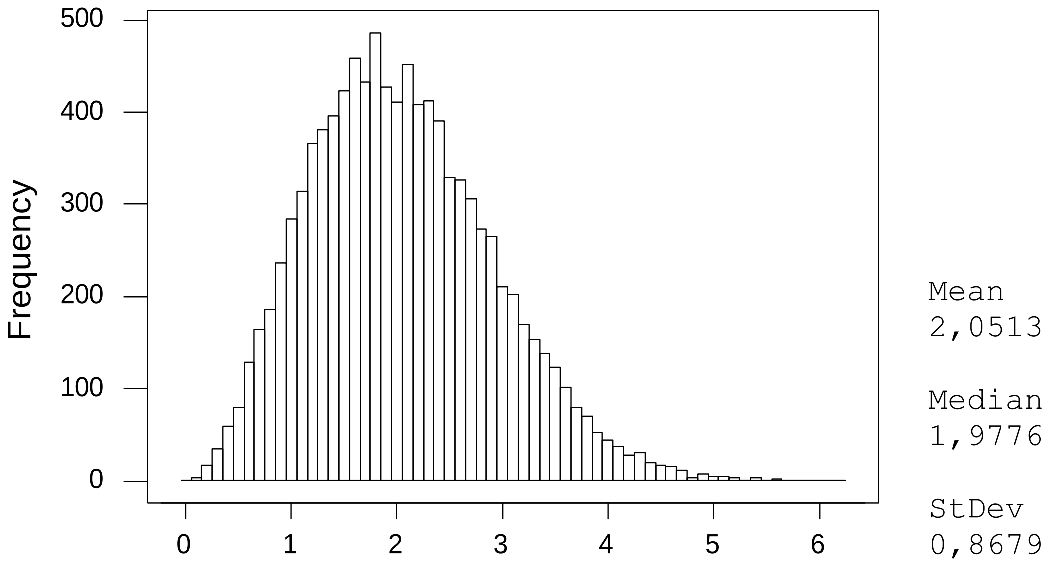
Ponemos el contador *K100* a 1 y ejecutamos 10000 veces la macro.

```
MTB > Let K100=1
MTB > Execute "C:\ rango.mtb" 10000.
```

Los 10000 rangos están ahora guardados en la columna *C2*. Si los dividimos por 5 (que es el valor de σ) tendremos los valores del rango relativo.

```
MTB > Let C2=C2/5
```

Un histograma de los valores en *C2* nos muestra el aspecto del rango relativo *W* para $n = 4$. La media de esta columna es 2,0513 y la desviación tipo es 0,8679. Estos valores son lógicamente muy parecidos a d_2 (esperanza matemática de *W*) y a d_3 (desviación tipo de *W*) para $n = 4$, como puede comprobarse mirando la Tabla 1.



APÉNDICE II: Cálculo directo por simulación de los valores de A_2 en el gráfico $\bar{X}$.

Supongamos que tenemos el proceso de llenado de botellas de agua del apéndice I. Recordemos que el contenido en ml de las botellas de agua se distribuye $N(500;5)$.

Los límites de control para el gráfico $\bar{x}$, si se calcularan a partir de la desviación tipo, serían:

$$LCS = \mu + 3\frac{\sigma}{\sqrt{n}}$$
$$LC = \mu$$
$$LCI = \mu - 3\frac{\sigma}{\sqrt{n}}$$

Pero como se utiliza el rango en lugar de la desviación tipo, los límites de control se calculan a partir de las fórmulas:

$$LCS = \mu + A_2R$$
$$LC = \mu$$
$$LCI = \mu - A_2R$$

Por tanto, debe cumplirse que: $3\frac{\sigma}{\sqrt{n}} = A_2R \Rightarrow A_2 = \frac{3\frac{\sigma}{\sqrt{n}}}{R}$.

Vamos a simular la extracción de 10000 muestras de tamaño 20 de este proceso. Esta es la macro:

```
Random 4 c1;
  Normal 500 5.  #saco una muestra de n=4
Range C1 K1.     #calculo del rango
Let C2(K100)=K1  #guardo el rango en C2
Let K100=K100+1  #incremento el contador
```

Ponemos el contador *K100* a 1 y ejecutamos 10000 veces la macro.

```
MTB > Let K100=1
MTB > Execute "C:\ xbarra.mtb" 10000.
```

En la columna *C2* tenemos ahora el rango de cada una de las 10000 muestras. La media de todos los rangos es:

```
MTB > Mean c2

   Mean of C2 = 10,257
```

Y por tanto, $A_2 = \frac{3\frac{\sigma}{\sqrt{n}}}{R} = \frac{3\frac{5}{\sqrt{4}}}{10{,}257} = 0{,}731$

El valor coincide (redondeado a dos decimales) con el que se ofrece en la tabla 1. Podríamos seguir el mismo procedimiento para otros tamaños de muestra.

Para obtener el valor correcto hasta 3 decimales, es necesario ejecutar la simulación muchas más veces (por ejemplo, 10 millones de veces). Hacer esa ejecución en Minitab es inviable, porque es demasiado lento. Una alternativa es ejecutar esta simple serie de comandos en R:

```
R <- vector(mode="numeric", length=10000000)
for (k in 1:10000000)
{
  x <- rnorm(4, mean=500, sd=5)
  R[k] <- max(x)-min(x)
}
(15/sqrt(4))/(mean(R))
```

Esto nos dará el valor de A2 para n=4