

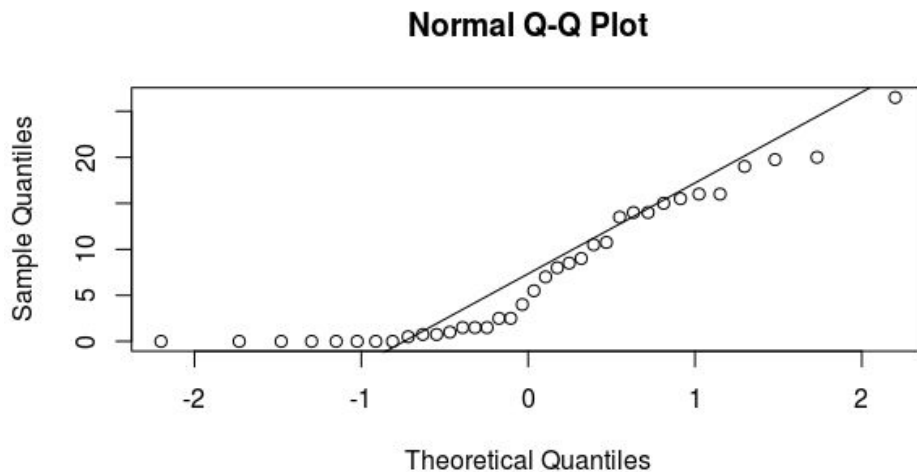
Laboratorio 3
Estadística Para Ingenieros
Francisco Sucre 10-10717
José Cipagauta 05-38040

Se tienen las notas del primer examen de tres secciones de un mismo curso. Los exámenes son evaluados sobre 100 puntos.

```
> # Cargamos la data  
> Notas11_12 <- read.table("Notas11-12.txt", header = T, fill=NA)  
> Notas12_13 <- read.table("Notas12-13.txt", header = T, fill=NA)  
> attach(Notas11_12)  
> attach(Notas12_13)
```

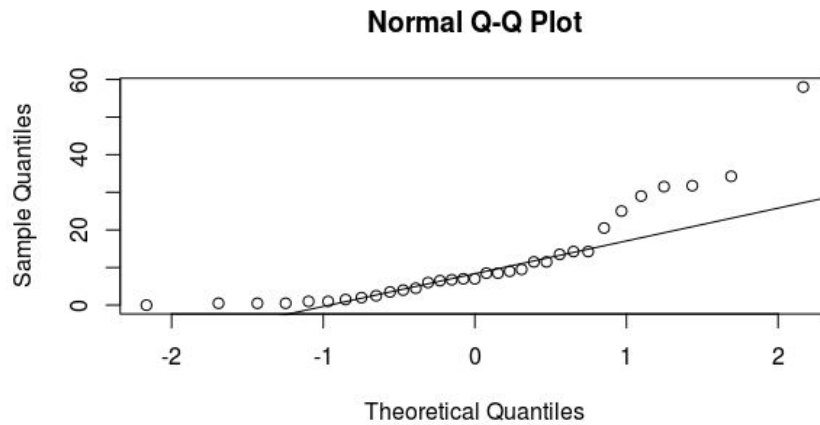
1. Utilizando las graficas vistas en clases, ¿cuáles de las secciones tienen notas que se distribuyen normal?

```
> # Grafico Notas11_12  
> # S5  
> qqnorm(S5)  
> qqline(S5)
```



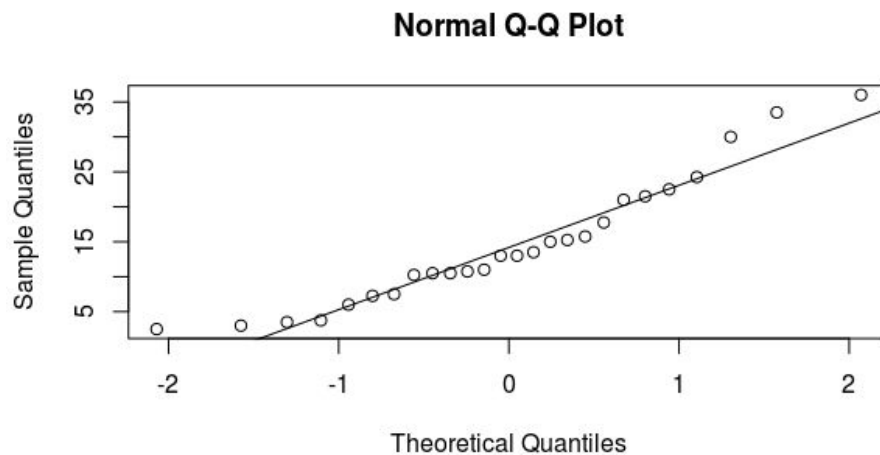
No es normal ni se aproxima a serlo, excepto en secciones muy pequeñas

```
> # S6  
> qqnorm(S6)  
> qqline(S6)
```



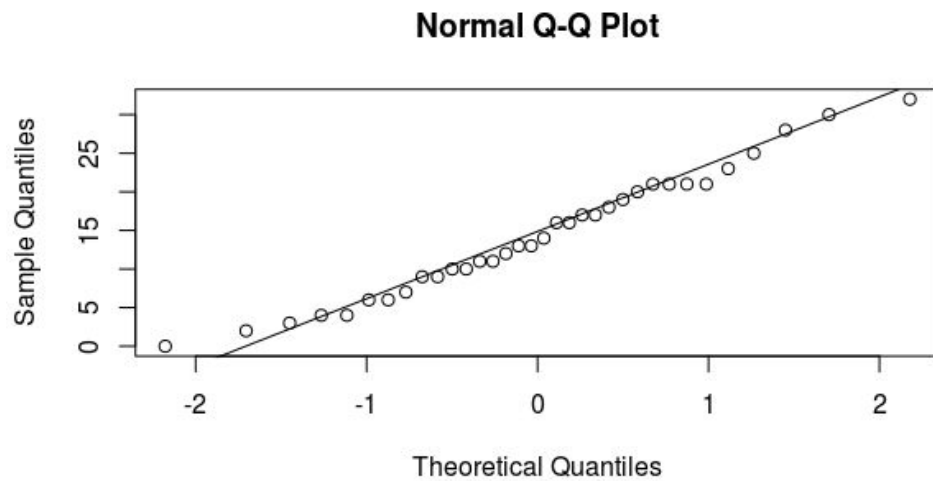
Tiene unas secciones que no se acercan a estar distribuidas de forma normal, pero en la sección del medio se comporta de manera muy parecida a una distribución normal

```
> # S13  
> qqnorm(S13)  
> qqline(S13)
```



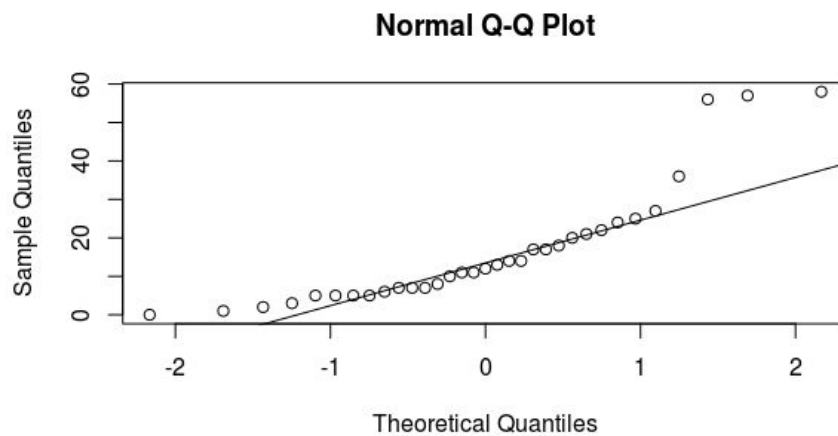
Se aproxima a la distribución normal, excepto en las colas, aun así se podría tomar como una distribución normal, pero las conclusiones obtenidas al asumir esto podrían ser erradas

```
> # S3  
> qqnorm(S3)  
> qqline(S3)
```



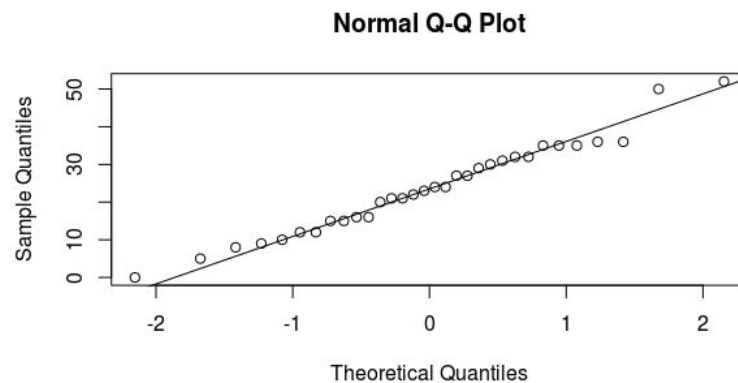
Se aproxima bastante a la distribución normal así que si se puede tomar como una distribución normal

```
> # S4  
> qqnorm(S4)  
> qqline(S4)
```



Se aproxima a la distribución normal, excepto en las colas con datos atípicos, no se recomienda tomar como una distribución normal

```
> # S9
> qqnorm(S9)
> qqline(S9)
```



Se aproxima a la distribución normal así que si se puede tomar como una distribución normal

**2. Calcule intervalos de confianza del 97% para la media de la nota de cada sección.
¿Qué consideraciones se deben tomar en cuenta para construir estos intervalos?**

Consideraciones: Que los datos pertenezcan a una distribución normal, puede que el intervalo no sea el más correcto o que los resultados no nos presenten información útil si la distribución no es normal

```
> S5_tTest = t.test(S5, conf.level = 0.97)
> S6_tTest = t.test(S6, conf.level = 0.97)
> S13_tTest = t.test(S13, conf.level = 0.97)
> S3_tTest = t.test(S3, conf.level = 0.97)
> S4_tTest = t.test(S4, conf.level = 0.97)
> S9_tTest = t.test(S9, conf.level = 0.97)
```

```
> S5_tTest
```

One Sample t-test

```
data: S5
t = 5.8223, df = 35, p-value = 1.324e-06
alternative hypothesis: true mean is not equal to 0
97 percent confidence interval:
 4.501025 10.221197
sample estimates:
mean of x
 7.361111
```

> S6_tTest

One Sample t-test

```
data: S6
t = 5.172, df = 32, p-value = 1.205e-05
alternative hypothesis: true mean is not equal to 0
97 percent confidence interval:
 6.547662 16.800822
sample estimates:
mean of x
11.67424
```

> S13_tTest

One Sample t-test

```
data: S13
t = 8.1085, df = 25, p-value = 1.836e-08
alternative hypothesis: true mean is not equal to 0
97 percent confidence interval:
10.42633 18.68906
sample estimates:
mean of x
14.55769
```

> S3_tTest

One Sample t-test

```
data: S3
t = 10.234, df = 33, p-value = 9.046e-12
alternative hypothesis: true mean is not equal to 0
97 percent confidence interval:
11.19496 17.56975
sample estimates:
mean of x
14.38235
```

> S4_tTest

One Sample t-test

```
data: S4
t = 6.1278, df = 32, p-value = 7.501e-07
alternative hypothesis: true mean is not equal to 0
97 percent confidence interval:
10.37494 22.59475
sample estimates:
mean of x
16.48485
```

> S9_tTest

One Sample t-test

```
data: S9
t = 11.034, df = 31, p-value = 2.906e-12
alternative hypothesis: true mean is not equal to 0
97 percent confidence interval:
18.85399 28.64601
sample estimates:
mean of x
23.75
```

Conclusiones interesantes es que considerando que el examen es evaluado sobre 100 puntos y en el mejor de los casos, con un 97% de confiabilidad el mayor promedio es 28.64 entre todas las secciones, lo cual indica un promedio terriblemente bajo

3. ¿Qué conclusiones puede sacar a partir de los intervalos de confianza de sus variables? En particular, ¿cómo compararía los resultados entre secciones?

Las comparaciones se harían con intervalos de confianza de la diferencia entre medias

```
> # Comparaciones  
> t.test(S5,S6,conf.level = 0.97)
```

Welch Two Sample t-test

```
data: S5 and S6  
t = -1.6671, df = 50.669, p-value = 0.1017  
alternative hypothesis: true difference in means is not equal to 0  
97 percent confidence interval:  
-10.090081  1.463819  
sample estimates:  
mean of x mean of y  
7.361111 11.674242
```

Podemos ver que la media de S6 podría ser mayor que la del S5 Pero al incluir el 0, podemos decir que las medias son relativamente iguales.

```
> t.test(S5,S13,conf.level = 0.97)
```

Welch Two Sample t-test

```
data: S5 and S13  
t = -3.2773, df = 47.585, p-value = 0.001961  
alternative hypothesis: true difference in means is not equal to 0  
97 percent confidence interval:  
-12.108965 -2.284198  
sample estimates:  
mean of x mean of y  
7.361111 14.557692
```

Podemos ver que la media de S13 siempre es mayor que la de S5, por al menos 2 puntos

```
> t.test(S6,S13,conf.level = 0.97)
```

Welch Two Sample t-test

```
data: S6 and S13  
t = -0.99976, df = 56.402, p-value = 0.3217  
alternative hypothesis: true difference in means is not equal to 0  
97 percent confidence interval:  
-9.304656  3.537756  
sample estimates:  
mean of x mean of y  
11.67424  14.55769
```

Podemos ver que la media de S13 podría ser mayor que la del S6 Pero al incluir el 0, podemos decir que las medias son relativamente iguales

```
> t.test(S3,S4,conf.level = 0.97)
```

Welch Two Sample t-test

```
data: S3 and S4
t = -0.69272, df = 48.357, p-value = 0.4918
alternative hypothesis: true difference in means is not equal to 0
97 percent confidence interval:
-8.889096  4.684105
sample estimates:
mean of x mean of y
14.38235  16.48485
```

Podemos ver que la media de S4 podría ser mayor que la de S3 Pero al incluir el 0, podemos decir que las medias son relativamente iguales

```
> t.test(S3,S9,conf.level = 0.97)
```

Welch Two Sample t-test

```
data: S3 and S9
t = -3.6441, df = 53.868, p-value = 0.0006044
alternative hypothesis: true difference in means is not equal to 0
97 percent confidence interval:
-15.097822 -3.637472
sample estimates:
mean of x mean of y
14.38235  23.75000
```

Podemos ver que la media de S9 siempre es mayor que la de S3, por al menos 3 puntos

```
> t.test(S4,S9,conf.level = 0.97)
```

Welch Two Sample t-test

```
data: S4 and S9
t = -2.1087, df = 60.494, p-value = 0.03912
alternative hypothesis: true difference in means is not equal to 0
97 percent confidence interval:
-14.9222502  0.3919472
sample estimates:
mean of x mean of y
16.48485  23.75000
```

Podemos ver que la media de S9 es generalmente mayor que la de S4 pero aun asi incluye el 0, por lo cual podrían ser iguales

4. Si realizan la comparación de los intervalos de confianza entre años, ¿cómo los compararía? ¿Qué puede concluir al hacer la comparación?

Se compararian los intervalos de confianza de la diferencia de promedio de todas las notas de las 3 secciones de cada año

```
> t.test(Notas11_12, Notas12_13, conf.level = 0.97)
```

Welch Two Sample t-test

```
data: Notas11_12 and Notas12_13
t = -4.3467, df = 187.4, p-value = 2.263e-05
alternative hypothesis: true difference in means is not equal to 0
97 percent confidence interval:
-10.945643 -3.618684
sample estimates:
mean of x mean of y
10.82895 18.11111
```

Podemos concluir que los promedios del año 12_13 es mayor a los promedios del año 11_12 por lo menos en 3 puntos, con un intervalo de confianza del 97%