

Normalidad en R

```
# Cargando paquetes
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5      v purrr  0.3.4
```

```
## v tibble  3.1.4      v dplyr  1.0.7
```

```
## v tidyr   1.1.3      v stringr 1.4.0
```

```
## v readr   2.0.1      v forcats 0.5.1
```

```
## Warning: package 'tibble' was built under R version 4.1.1
```

```
## Warning: package 'readr' was built under R version 4.1.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
```

```
## x dplyr::lag()     masks stats::lag()
```

```
library(ggpubr)
```

```
## Warning: package 'ggpubr' was built under R version 4.1.1
```

```
library(rstatix)
```

```
## Warning: package 'rstatix' was built under R version 4.1.1
```

```
##
```

```
## Attaching package: 'rstatix'
```

```
## The following object is masked from 'package:stats':
```

```
##
```

```
##      filter
```

```
# Cargando una base de datos de R
```

```
head(ToothGrowth)
```

```
##      len supp dose
```

```
## 1  4.2   VC  0.5
```

```
## 2 11.5   VC  0.5
```

```
## 3  7.3   VC  0.5
```

```
## 4  5.8   VC  0.5
```

```
## 5  6.4   VC  0.5
```

```
## 6 10.0   VC  0.5
```

```
# Tomando una muestra aleatoria.
set.seed(2021)
ToothGrowth%>% sample_n_by(supp,dose,size = 1)
```

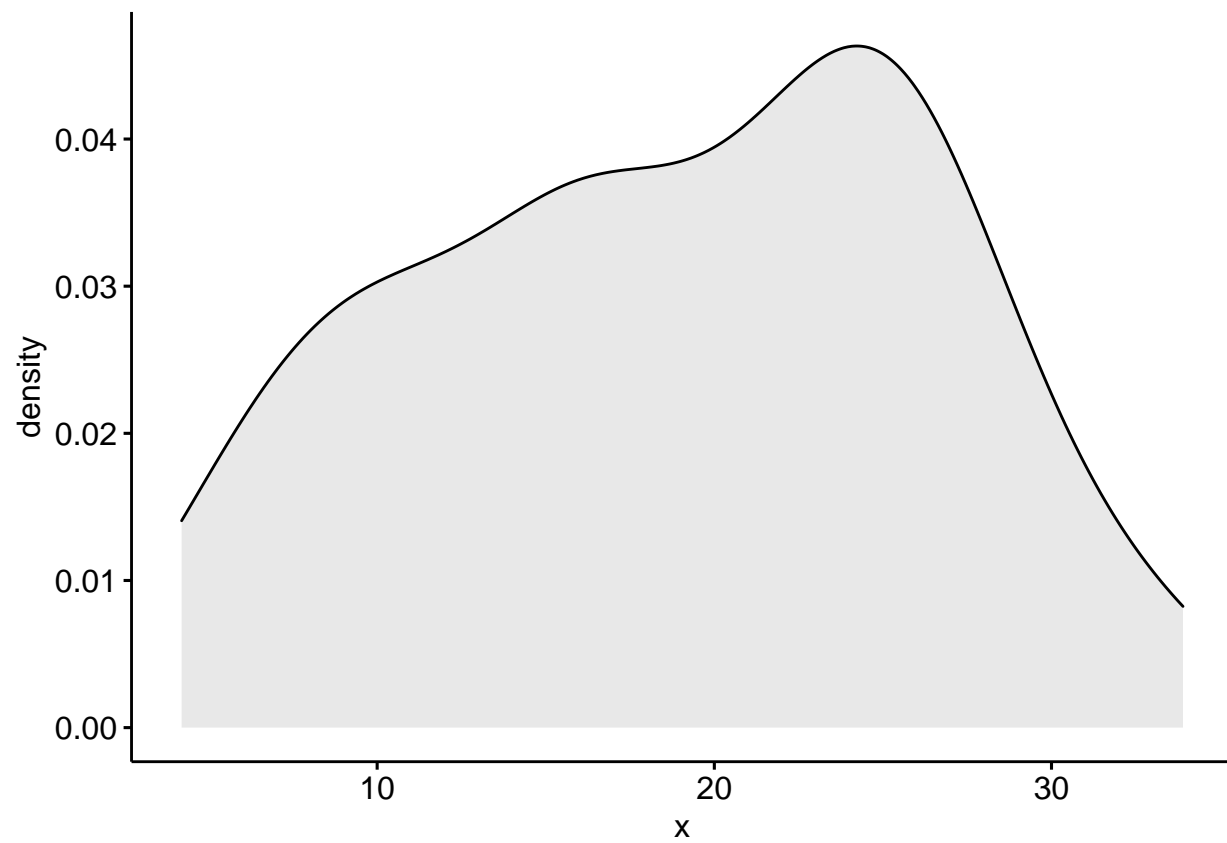
```
## # A tibble: 6 x 3
##   len supp  dose
##   <dbl> <fct> <dbl>
## 1   8.2 OJ    0.5
## 2  25.2 OJ    1
## 3   23  OJ    2
## 4  11.2 VC    0.5
## 5  17.3 VC    1
## 6  32.5 VC    2
```

Verificar normalidad en R

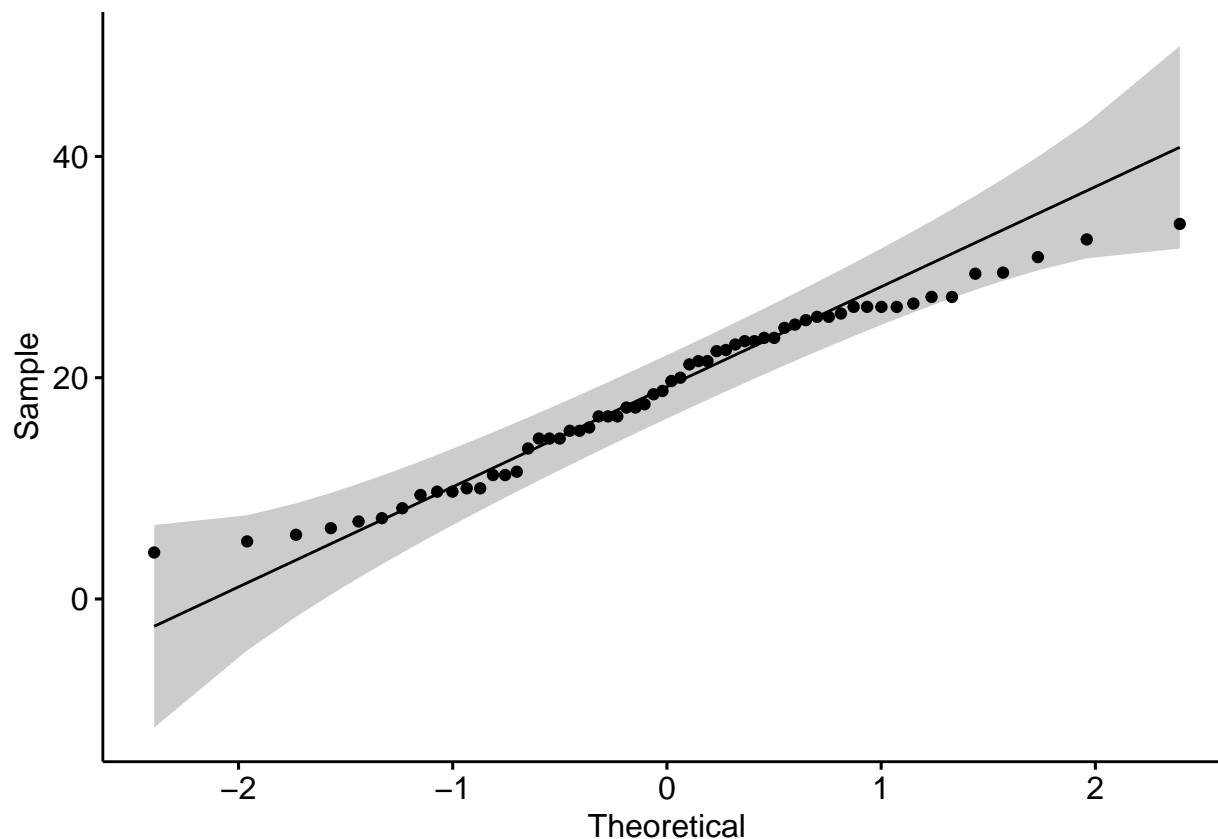
Pregunta: Queremos probar si la variable len (longitud del diente) es normalmente repartido. ## Métodos visuales El gráfico de densidad y el gráfico Q-Q se pueden utilizar para comprobar la normalidad visualmente.

1. **Gráfico de densidad:** el gráfico de densidad proporciona un juicio visual sobre si la distribución tiene forma de campana.
2. **Gráfico de QQ:** gráfico de QQ (o gráfico de cuantiles-cuantiles) dibuja la correlación entre una muestra dada y la distribución normal. A 45 grados También se traza la línea de referencia. En una gráfica QQ, cada observación se representa como un solo punto. Si los datos son normales, los puntos deben formar una línea recta.

```
library("ggpubr")
# Density plot
ggdensity(ToothGrowth$len, fill = "lightgray")
```



```
# QQ plot  
ggqqplot(ToothGrowth$len)
```



```
ToothGrowth %>% shapiro_test(len)
```

```
## # A tibble: 1 x 3
##   variable statistic      p
##   <chr>      <dbl> <dbl>
## 1 len          0.967 0.109
```

p-valor mayor que 0.05, por lo tanto asumimos normalidad en la variable **len**.

Normalidad para agrupacion de dose(por cada categoria o respuesta de dose)

```
ToothGrowth %>%
group_by(dose) %>%
shapiro_test(len)
```

```
## # A tibble: 3 x 4
##   dose variable statistic      p
##   <dbl> <chr>      <dbl> <dbl>
## 1 0.5 len          0.941 0.247
## 2 1 len          0.931 0.164
## 3 2 len          0.978 0.902
```

Todos los p-valores son mayores que 0.05, por lo tanto la variable **len** siguen una distribucion Normal para cada respuesta o categoria de **dose**