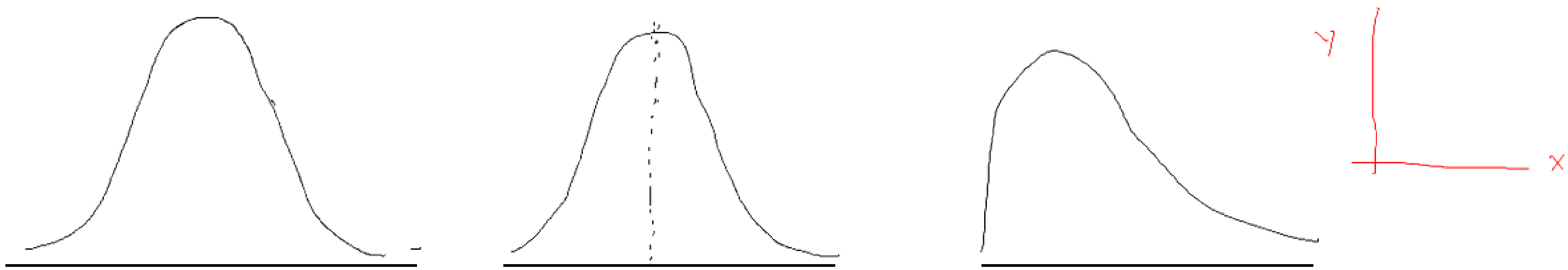
$$e_i = y_i - \hat{y}_i, \quad i = 1, \ldots, n$$

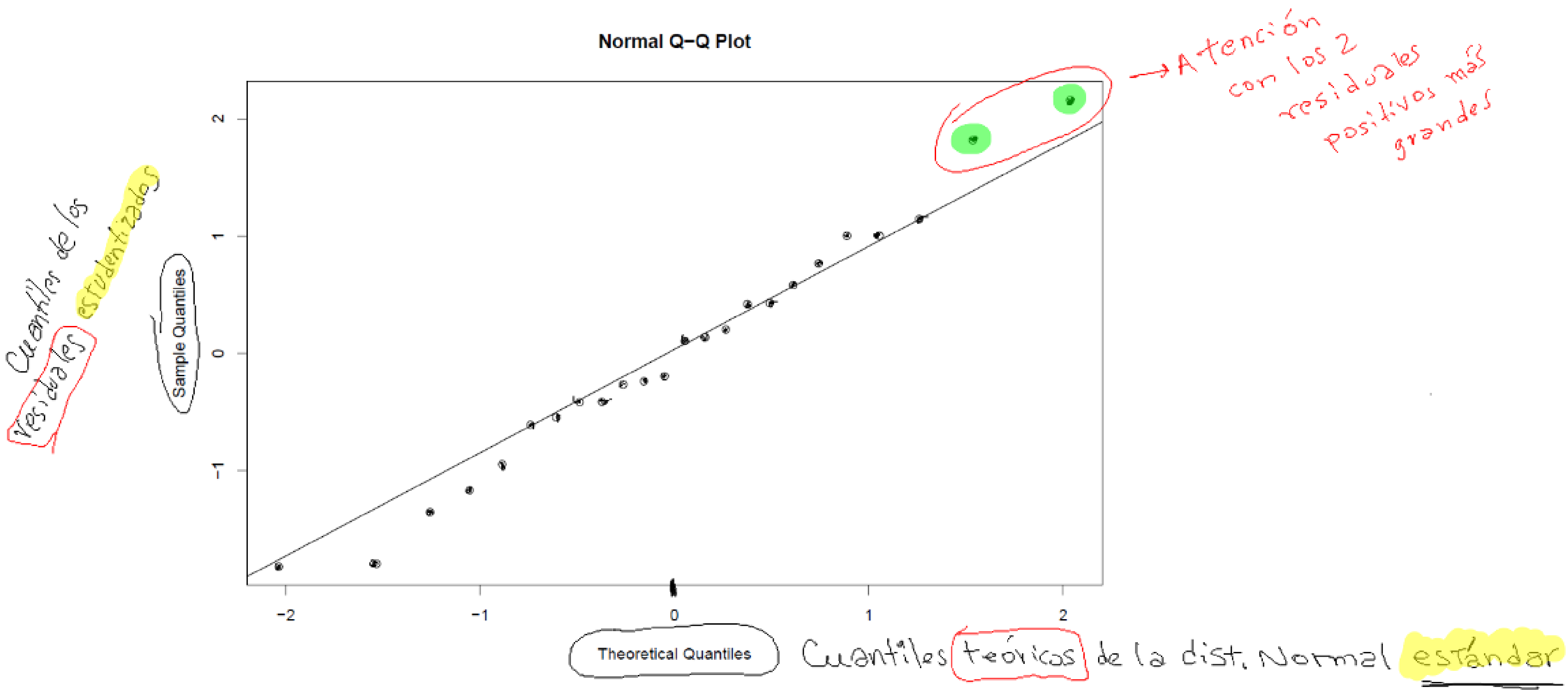
$$x'_i\hat{\beta}$$

y x $\hat{y}$ e residual estandarizado

```
> modelo |> augment()
# A tibble: 24 × 11
   Sueldo Educacion Sexo   Edad    X4 .fitted .resid  .hat .sigma .cooksd .std.resid
    <dbl>     <dbl> <chr> <dbl> <dbl>   <dbl>  <dbl> <dbl>  <dbl>   <dbl>      <dbl>
 1   4.22      12   M        31    10    4.76 -0.545 0.205  1.16  0.0150      -0.538
 2   5.89      18   M        23    -3    5.46  0.431 0.417  1.16  0.0354       0.497
 3   9.88      20   M        50    -6    8.73  1.15  0.217  1.13  0.0720       1.14
 4   2.35       7.5 F        19    -4    2.62 -0.266 0.235  1.16  0.00443     -0.268
 5   7         20   F        39    -2    6.58  0.420 0.123  1.16  0.00437      0.395
 6   1.25       9   M        18     7    3.07 -1.82  0.255  1.05  0.236       -1.86
 7   6.78       3   M        39    -7    4.96  1.82  0.536  0.981 1.28         2.35
 8   5.19      15   F        32     8    4.61  0.581 0.144  1.16  0.0103       0.554
 9   8.16      21   F        35     5    6.00  2.16  0.162  1.02  0.167        2.08
10   6.1       18   F        34     7    5.33  0.766 0.131  1.15  0.0158       0.724
# ℹ 14 more rows
```

1. Normalidad de errores
2. Homocedasticidad de errores
3. Independencia de errores


y
x


Normal Q–Q Plot
Sample Quantiles
Theoretical Quantiles
−2
−1
0
1
2
Cuantiles de los residuales estudentizados
Cuantiles teóricos de la dist. Normal estándar
→ Atención con los 2 residuales positivos más grandes

```
> library(moments)
> residuales |> skewness()
[1] 0.1257975
> residuales |> agostino.test()

        D'Agostino skewness test

data:  residuales
skew = 0.1258, z = 0.3019, p-value = 0.7627
alternative hypothesis: data have a skewness

> residuales |> kurtosis()
[1] 2.592246
> residuales |> anscombe.test()

        Anscombe-Glynn kurtosis test

data:  residuales
kurt = 2.592246, z = -0.028832, p-value =
0.977
alternative hypothesis: kurtosis is not equal to 3
```

$as \to$ estimador
$As \to$ parámetro

* $as = 0.1258$, ¿es lo suficientemente grande para decir que es distinto de cero?

* $H_0 : As = 0$
$H_1 : As \neq 0$
$\left.\begin{array}{l} \alpha = 0.05 \\ pV = 0.7627 \end{array}\right\}$ No se rechaza Ho

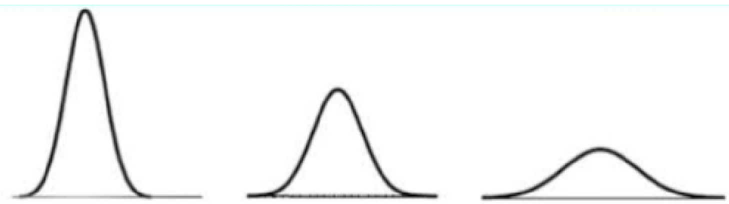
Existe evidencia estadística de que los residuales son simétricos

$k \to$ estimador
$K \to$ parámetro

* $k = 2.59$

* $H_0 : K = 3$
$H_1 : K \neq 3$
$\left.\begin{array}{l} \alpha = 0.05 \\ pV = 0.977 \end{array}\right\}$ No se rechaza Ho

Existe evidencia estadística de que los residuales presentan una distribución mesocúrtica.



Leptocúrtica — $g_2 > 3$

Mesocúrtica — $g_2 = 3$

Platicúrtica — $g_2 < 3$

COEFICIENTE DE CURTOSIS

Mide el grado de apuntamiento o achatamiento de la distribución de frecuencia, respecto a la curva de distribución normal que tiene coeficiente igual a 0.

$$g_2 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{s_x^4} - 3$$

Pruebas de normalidad

H0: Los errores siguen una distribución Normal

H1: Los errores no siguen una distribución Normal

Kolmogorov (1903 - 1987)
Smirnov (1900 - 1966)
Lilliefors (1928 - 2008)
Anderson (1918 - 2016)
Darling (1915 - 2014)
Shapiro (1930 - 2023)
Wilk (1922 - 2013)

- Shapiro Wilk es la prueba más potente (con mayor $\underbrace{\text{potencia de prueba}}_{1-\beta}$)

	H_0 cierta	H_0 falsa
Acepta H_0	$1-\alpha$	$II \to \beta$
Rechaza H_0	$I \to \alpha$	$1-\beta$

Si H_0 es falsa (errores no son normales)

SW detecta mejor esas desviaciones

$n \downarrow \quad 1-\beta \downarrow$

Pruebas de normalidad
H0: Los errores siguen una distribución Normal
H1: Los errores no siguen una distribución Normal
$\alpha = 0.05$

```
> residuales |> shapiro.test()

	Shapiro-Wilk normality test

data:  residuales
W = 0.98304, p-value = 0.9446
```

```
> residuales |> ad.test()

	Anderson-Darling normality test

data:  residuales
A = 0.13547, p-value = 0.9741
```

```
> residuales |> lillie.test()

	Lilliefors (Kolmogorov-Smirnov) normality
	test

data:  residuales
D = 0.075434, p-value = 0.977
```

No se rechaza H0, por lo tanto los errores siguen una distribución Normal

F distribution

①the definition is given by:

$F=\frac{\chi^2_{\nu_1}}{\nu_1}/\frac{\chi^2_{\nu_2}}{\nu_2}$, where $\chi^2_{\nu_i}$ is the chi-square PDF of DOF(degree of freedom) ν_i, for $i=1,2$.

SCReg

SCError

GLE

GLReg

②the F distribution PDF is expressed in below equality:

$$h(f)=\frac{\Gamma(\frac{\nu_1+\nu_2}{2})\cdot(\frac{\nu_1}{\nu_2})^{\frac{\nu_1}{2}}}{\Gamma(\frac{\nu_1}{2})\cdot\Gamma(\frac{\nu_2}{2})}\cdot\frac{f^{\frac{\nu_1}{2}-1}}{(1+\frac{\nu_1}{\nu_2}\cdot f)^{\frac{\nu_1+\nu_2}{2}}}$$

t^2

$F^{1/2}$

t distribution

$$T=\frac{\overline{X}_n-\mu}{S/\sqrt{n}}$$
$$=\frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}}/\frac{S/\sqrt{n}}{\sigma/\sqrt{n}}$$
$$=\frac{Z}{S/\sigma},\text{ where } Z\sim\phi(0,1)$$
$$=\frac{Z}{\sqrt{S^2/\sigma^2}}$$
$$=\frac{Z}{\sqrt{\frac{\chi^2_{n-1}}{n-1}}}$$

The PDF of t distribution is given by:

$$f_T(t)=\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\cdot\nu}\cdot\Gamma(\frac{\nu}{2})}\cdot(1+\frac{t^2}{\nu})^{-\frac{\nu+1}{2}},$$

where ν is the degree of freedom, $-\infty<t<\infty$.

$H_0: \beta_1 = 4$

$H_1: \beta_1 \neq 4$

$t = \frac{\hat{\beta}_1 - \beta_1}{s_{\hat{\beta}_1}}$

Gamma distribution

$$f(x)=\frac{1}{\beta^\alpha\cdot\Gamma(\alpha)}\cdot x^{\alpha-1}\cdot e^{-\frac{x}{\beta}}$$
$$=\frac{1}{\beta\cdot\Gamma(\alpha)}\cdot(\frac{x}{\beta})^{\alpha-1}\cdot e^{-\frac{x}{\beta}}$$
$$=\frac{\frac{1}{\beta}\cdot(\frac{x}{\beta})^{\alpha-1}\cdot e^{-\frac{x}{\beta}}}{\Gamma(\alpha)}$$

, where $\alpha>0,\beta>0$

Chi-Square distribution

$f(x)=\frac{1}{2^{\frac{\nu}{2}}\cdot\Gamma(\frac{\nu}{2})}\cdot x^{\frac{\nu}{2}-1}\cdot e^{-\frac{x}{2}}$, for $x>0$

, where ν is a positive integer, the chi-square PDF.

$Z^2\sim\chi^2_1$

Gamma function

$\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}\cdot e^{-x}\,\mathrm{d}\,x$,

where $\alpha>0$

Standard normal distribution

$\boxed{Z\sim\phi(0,1)}$

Beta function

$\beta(x,y)=\frac{\Gamma(x)\cdot\Gamma(y)}{\Gamma(x+y)}$

Exponential distribution

$F_T(t)=1-e^{-\lambda\cdot t}=P(T\leq t)$

$f_T(t)=\frac{\mathrm{d}\,F_T(t)}{\mathrm{d}\,t}=\lambda\cdot e^{-\lambda\cdot t}$

Beta distribution

$f_X(x)=\frac{1}{\beta(a,b)}\cdot x^{a-1}\cdot(1-x)^{b-1}$

$F_X(k)=\frac{\beta(k,a,b)}{\beta(a,b)}=\frac{\int_0^k x^{a-1}\cdot(1-x)^{b-1}}{\beta(a,b)}$

Poisson distribution

$P(X=k)=\frac{(\mu)^k}{k!}\cdot e^{-\mu}$, for k=0,1,2,...,