matriz de diseño

$$\mathbf{y}_{n\times 1} = \mathbf{X}_{n\times k}\beta_{k\times 1} + \epsilon_{n\times 1} = \mathbf{X}_{n\times(p+1)}\beta_{(p+1)\times 1} + \epsilon_{n\times 1}$$

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{n\times 1}}_{\checkmark} = \underbrace{\begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{pmatrix}_{n\times K}}_{\checkmark} \times \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}_{K\times 1} ¿? + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} ¿?$$

```
(Intercept)         IMC
  64.557707    3.095683
```

```
(Intercept)          Edad Minutos_ejercicio         IMC
110.47710380   0.45472167       -0.08487366  0.27956285
```

X_1+1

$$\hat{y} = \cancel{\hat{\beta}_0} + \hat{\beta}_1 X_1 + \cancel{\hat{\beta}_2 X_2} + \cancel{\hat{\beta}_3 X_3}$$

$$\hat{y}^* = \cancel{\hat{\beta}_0} + \hat{\beta}_1 (X_1 + 1) + \cancel{\hat{\beta}_2 X_2} + \cancel{\hat{\beta}_3 X_3}$$

$$\hat{y}^* - \hat{y} = \hat{\beta}_1$$

Ejemplo $H_0 : \beta_1 = \beta_2 = \beta_3 = 0 \qquad H_1$: Al menos un $\beta_j \neq 0 \qquad \alpha = 0.05$

```
X = cbind(1,datos$Edad, datos$Minutos_ejercicio, datos$IMC)
modelo2_ = lm(Presion_sistolica ~ X, datos)
modelo2_ |> aov() |> summary()
```

GLReg

```
            Df Sum Sq Mean Sq F value Pr(>F)
X            3  11282    3761   76.65 <2e-16 ***
Residuals   96   4710      49
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

GLError

SC Total = SC Reg + SCE

= 11 282 + 4 710 = 15 992

GLTotal = 99

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	110.477104	8.598105	12.849	< 2e-16	***
Edad	0.454722	0.052810	8.610	1.43e-13	***
Minutos_ejercicio	-0.084874	0.009152	-9.274	5.41e-15	***
IMC	0.279563	0.421363	0.663	0.509	

X_1

X_3

$$\hat{y} = \beta_0 + \underbrace{\overset{\neq 0 \ \neq 0}{\beta_1 X_1}} + \dots + \underbrace{\overset{=0 \ \neq 0}{\beta_3 X_3}}$$

✓

X_1 es signif.

X_1 contribuye signif.

X_3 no es signif.

X_3 no contribuye signif.

$H_0 : \beta_1 = 0$ ✗

$H_1 : \beta_1 \neq 0$ ✓

$\alpha = 0.05$

$pV = 1.43 \times 10^{-13}$

$pV < \alpha$

$\Downarrow$

Rech. H_0

$\Downarrow$

$\beta_1 \neq 0$

$H_0 : \beta_3 = 0$ ✓

$H_1 : \beta_3 \neq 0$

$\alpha = 0.05$

$pV = 0.509$

$pV > \alpha$

$\Downarrow$

No Rech. H_0

$\Downarrow$

$\beta_3 = 0$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

Ejm

		R^2	R^2_{aj}
	$Y \sim X_1 + X_2$	70%	69%
(X_3 sí aporta)	$Y \sim X_1 + X_2 + X_3$	75%	73%
(X_4 no aporta)	$Y \sim X_1 + X_2 + X_3 + X_4$	77%	72%

$\boxed{\text{parsimonia}}$

$$AIC = -2log(L) + 2r$$

$+$ parámetros $\Rightarrow$ $\log(L) \uparrow$

$-2\log(L) \downarrow$

pero $2r \uparrow$

$$VIF_j = \frac{1}{1 - R_j^2}$$

$\boxed{\text{VIF} \uparrow \text{ Multicol} \uparrow \quad \text{VIF} = 1 \text{ Multicol} \downarrow}$

Por ejm: $Y \sim \boxed{X_1 + X_2 + X_3 + X_4}$

$VIF_1 = \frac{1}{1 - R_1^2}$, $R_1^2 \rightarrow R^2$ del modelo $X_1 \sim X_2 + X_3 + X_4$

$R_1^2 = 0.93$

$VIF_1 = \frac{1}{1-0.93} = \frac{1}{0.07} = 14.3$

$VIF_2 = \frac{1}{1-R_2^2}$, $R_2^2 \rightarrow R^2$ del modelo $X_2 \sim X_1 + X_3 + X_4$

$R_2^2 = 0.10$

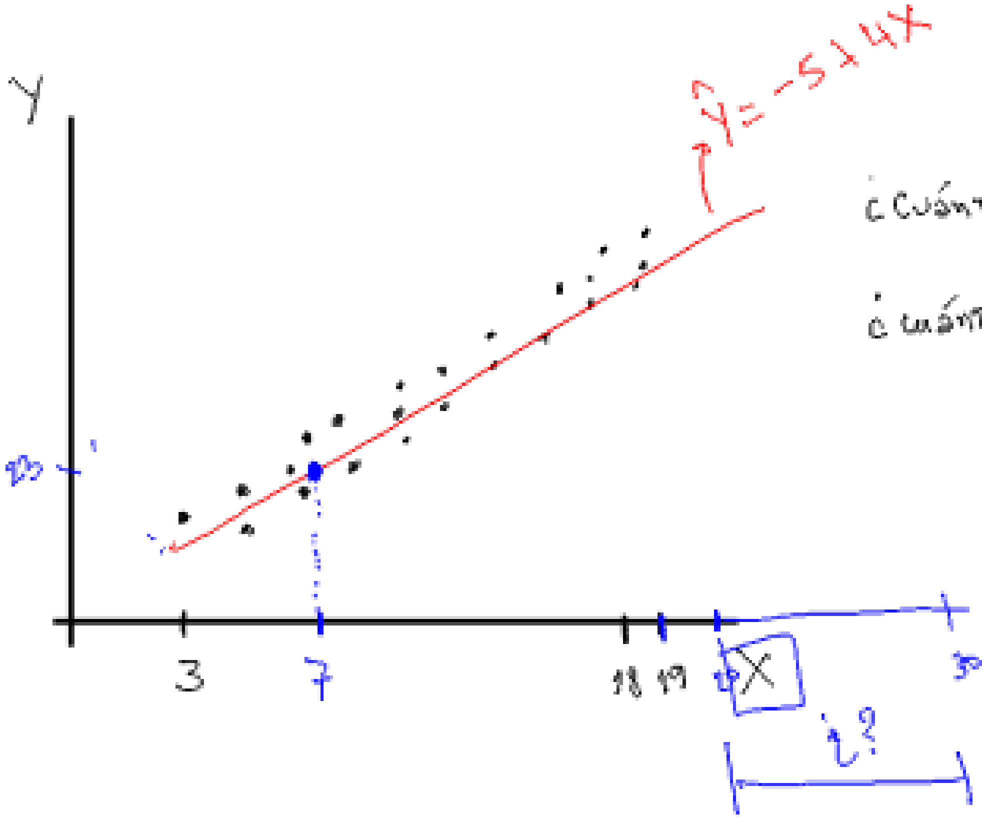
$VIF_2 = \frac{1}{1-0.10} = \frac{1}{0.90} = 1.11$

```
> # ESTIMACIÓN INTERVALAR DE LA MEDIA
> modelo2 |>
+   predict(data.frame(Edad = 30,
+                      Minutos_ejercicio = 60,
+                      IMC = 23),
+           interval = "confidence",
+           level = 0.90)
       fit      lwr      upr
1 125.4563 122.1979 128.7147
```

```
> # PREDICCION INTERVALAR DE Y
> modelo2 |>
+   predict(data.frame(Edad = 30,
+                      Minutos_ejercicio = 60,
+                      IMC = 23),
+           interval = "prediction",
+           level = 0.90)
       fit      lwr      upr
1 125.4563 113.3747 137.5378
```

$IC(\mu \mid Edad=30, Ej=60, IMC=23) = (122.2, 128.7)$

$IP(y \mid Edad=30, Ej=60, IMC=23) = \underbrace{(113.4, 137.5)}_{\text{más ancho}}$



¿Cuánto vale $\hat{y}$ cuando $x = 7$? $\bullet \rightarrow \boxed{23}$

¿Cuánto vale $\hat{y}$ cuando $x = 30$? $-5 + 4(30) = 115$

(extrapolación)