

ECONOMETRIA I
TAREA 7
MAESTRIA EN ECONOMIA, 2023-2025
EL COLEGIO DE MEXICO

Profr. Eneas A. Caldiño García
Entrega: 6 de marzo de 2024

1. La función de densidad del vector aleatorio $\begin{pmatrix} Y \\ X \end{pmatrix}$ es:

$$f_{y,x}: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$f_{y,x}(y, x) = \begin{cases} \lambda^2 e^{-\lambda y} & \text{si } 0 \leq x \leq y \\ 0 & \text{de lo contrario} \end{cases}$$

$\lambda > 0$. El valor de λ es desconocido.

Se sabe que la *función de regresión* de Y sobre X es:

$$g: [0, \infty) \rightarrow \mathbb{R}$$

$$g(x) = \frac{1}{\lambda} + x$$

i) En un mismo plano graficar las *funciones de regresión*

$$g(\cdot | \lambda): (0, \infty) \rightarrow \mathbb{R}$$

correspondientes a $\lambda = 0.5$, $\lambda = 1$ y $\lambda = 2$.

ii) Escribir los mejores modelos para explicar Y en función de X cuando se sabe que $\lambda = 0.5$, $\lambda = 1$ y $\lambda = 2$.

2. Distribución logística univariada.

Definición: La v.a. X tiene distribución logística ($X \sim Lg(\alpha, \beta)$) si y solo si

$$F_x: \mathbb{R} \rightarrow \mathbb{R}$$

$$F_x(x) = \frac{\exp\left(\frac{x-\alpha}{\beta}\right)}{1 + \exp\left(\frac{x-\alpha}{\beta}\right)} = \frac{1}{1 + \exp\left(-\frac{x-\alpha}{\beta}\right)}$$

donde $\alpha \in \mathbb{R}$ y $\beta \in (0, \infty)$.

$$E(X) = \alpha, \quad \text{var}(X) = \frac{\beta^2 \pi^2}{3}.$$

i) Encontrar la función de densidad $f_x: \mathbb{R} \rightarrow \mathbb{R}$

ii) En un mismo plano graficar la función de densidad f_x con $\alpha = 0$ y $\beta = 1$ y la función de densidad ϕ de la Normal univariada estándar, $N(0,1)$.

iii) En un mismo plano graficar la función de distribución acumulativa F_x con $\alpha = 0$ y $\beta = 1$ y la función de distribución acumulativa Φ de la $N(0,1)$.

3. Modelo logit. Ilustración de lo visto en clase.

i) Generar (usando algún paquete estadístico) una muestra $\{U_i\}_{i=1}^{200}$ i.i.d. donde $U_i \sim Lg(0,1) \quad \forall i \in \{1, 2, \dots, 200\}$.

ii) Generar (usando algún paquete estadístico) una muestra $\{X_i\}_{i=1}^{200}$ i.i.d. (de cualquier distribución), tal que $E[U_i | X_i] = 0 \quad \forall i \in \{1, 2, \dots, 200\}$.

iii) Generar $\{Y_i^*\}_{i=1}^{200}$, donde $Y_i^* = a + bX_i + U_i \quad i = 1, 2, \dots, 200$; con $a = 0.5$ y $b = 0.4$

iv) Generar la muestra $\left\{ \begin{pmatrix} Y_i \\ X_i \end{pmatrix} \right\}_{i=1}^{200}$

$$\text{donde } Y_i = \begin{cases} 1 & \text{si } Y_i^* > 0 \\ 0 & \text{si } Y_i^* \leq 0 \end{cases} \quad i = 1, 2, \dots, 200$$

v) En un mismo plano, graficar (usando los valores dados de a y b) la función de regresión (obtenida en clase)

$$E[Y|X=x] = 1 - \frac{1}{1 + \exp(a+bx)}$$

junto con la muestra $\left\{ \begin{pmatrix} Y_i \\ X_i \end{pmatrix} \right\}_{i=1}^{200}$ obtenida en iv).

4. Modelo probit (normit). Ilustración de lo visto en clase.

i) Generar (usando algún paquete estadístico) una muestra $\{U_i\}_{i=1}^{200}$ i.i.d. donde

$$U_i \sim N(0, 1) \quad \forall i \in \{1, 2, \dots, 200\}.$$

ii) Generar (usando algún paquete estadístico) una muestra $\{X_i\}_{i=1}^{200}$ i.i.d. (de cualquier distribución continua), tal que $E[U_i|X_i] = 0 \quad \forall i \in \{1, 2, \dots, 200\}$.

iii) Generar $\{Y_i^*\}_{i=1}^{200}$, donde $Y_i^* = a + bX_i + U_i \quad i = 1, 2, \dots, 200$; con $a = 0.5$ y $b = 0.4$

iv) Generar la muestra $\left\{ \begin{pmatrix} Y_i \\ X_i \end{pmatrix} \right\}_{i=1}^{200}$

$$\text{donde } Y_i = \begin{cases} 1 & \text{si } Y_i^* > 0 \\ 0 & \text{si } Y_i^* \leq 0 \end{cases} \quad i = 1, 2, \dots, 200$$

v) En un mismo plano, graficar (usando los valores dados de a y b) la función de regresión (obtenida en clase)

$$g(x) = E[Y|X=x] = 1 - \Phi(-(a+bx)) = \Phi(a+bx)$$

junto con la muestra $\left\{ \begin{pmatrix} Y_i \\ X_i \end{pmatrix} \right\}_{i=1}^{200}$ obtenida en iv).

5. $\begin{pmatrix} Y \\ X \end{pmatrix}$ vector aleatorio bidimensional.

$$f_{x,y}: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$f_{x,y}(x,y) = \begin{cases} x+y & \text{si } x \in (0,1) \text{ e } y \in (0,1) \\ 0 & \text{de lo contrario} \end{cases}$$

Sabemos que la función de regresión es:

$$g: (0,1) \rightarrow \mathbb{R}$$

$$g(x) = \frac{x + 2/3}{2x + 1}$$

i) Graficar la función de regresión $g: (0,1) \rightarrow \mathbb{R}$

ii) Verificar que la función de cedasticidad es:

$$h: (0,1) \rightarrow \mathbb{R}$$

$$h(x) = \frac{2x^2 + 2x + 1/3}{6(2x+1)^2}$$

iii) Graficar la función de cedasticidad $h: (0,1) \rightarrow \mathbb{R}$

6. Estudiar las sección 2-1 Definition of the Simple Regression Model del capítulo 2 de Wooldridge J.M., Introductory Econometrics: A Modern Approach, 5th, 6th or 7th Edition, South-Western CENGAGE Learning.

1. La función de densidad del vector aleatorio $\begin{pmatrix} Y \\ X \end{pmatrix}$ es:

$$f_{Y,X}: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$f_{Y,X}(y, x) = \begin{cases} \lambda^2 e^{-\lambda y} & \text{si } 0 \leq x \leq y \\ 0 & \text{de lo contrario} \end{cases}$$

$\lambda > 0$. El valor de λ es desconocido.

Se sabe que la función de regresión de Y sobre X es:

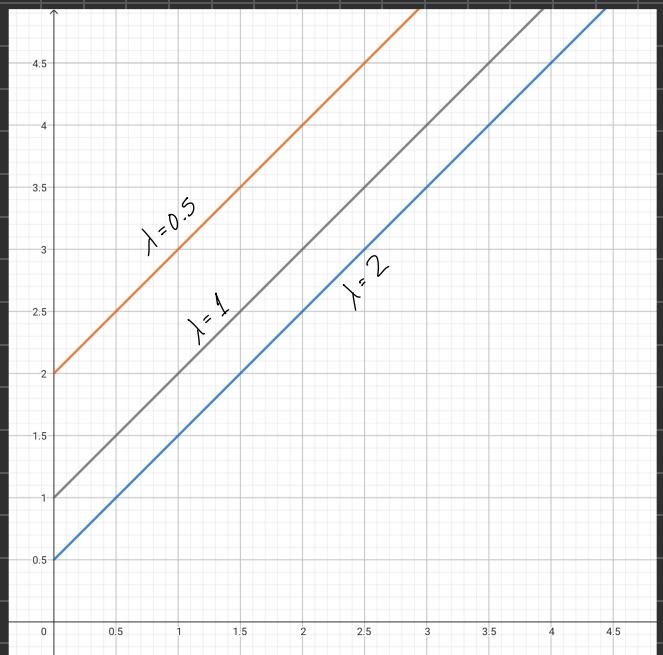
$$g: [0, \infty) \rightarrow \mathbb{R}$$

$$g(x) = \frac{1}{\lambda} + x$$

i) En un mismo plano graficar las funciones de regresión

$$g(\cdot | \lambda): (0, \infty) \rightarrow \mathbb{R}$$

correspondientes a $\lambda = 0.5$, $\lambda = 1$ y $\lambda = 2$.



ii) Escribir los mejores modelos para explicar Y en función de X cuando se sabe que $\lambda = 0.5$, $\lambda = 1$ y $\lambda = 2$.

Mejor modelo: $Y = \gamma_\lambda + X + U$. Entonces:

$$\text{Si } \lambda = 0.5 \Rightarrow Y = 2 + X + U$$

$$\lambda = 1 \Rightarrow Y = 1 + X + U$$

$$\lambda = 2 \Rightarrow Y = 0.5 + X + U$$

2. Distribución logística univariada.

Definición: La v.a. X tiene distribución logística ($X \sim Lg(\alpha, \beta)$) si y solo si

$$F_X: \mathbb{R} \rightarrow \mathbb{R}$$

$$F_X(x) = \frac{\exp\left(\frac{x-\alpha}{\beta}\right)}{1 + \exp\left(\frac{x-\alpha}{\beta}\right)} = \frac{1}{1 + \exp\left(-\frac{x-\alpha}{\beta}\right)}$$

donde $\alpha \in \mathbb{R}$ y $\beta \in (0, \infty)$.

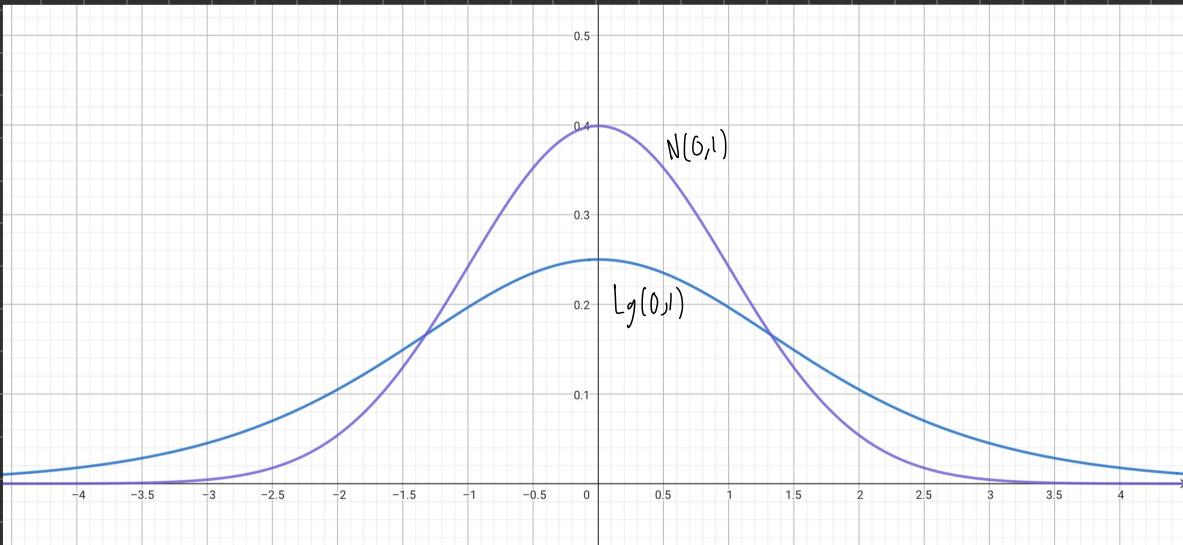
$$E(X) = \alpha, \quad \text{var}(X) = \frac{\beta^2 \pi^2}{3}.$$

i) Encontrar la función de densidad $f_X: \mathbb{R} \rightarrow \mathbb{R}$

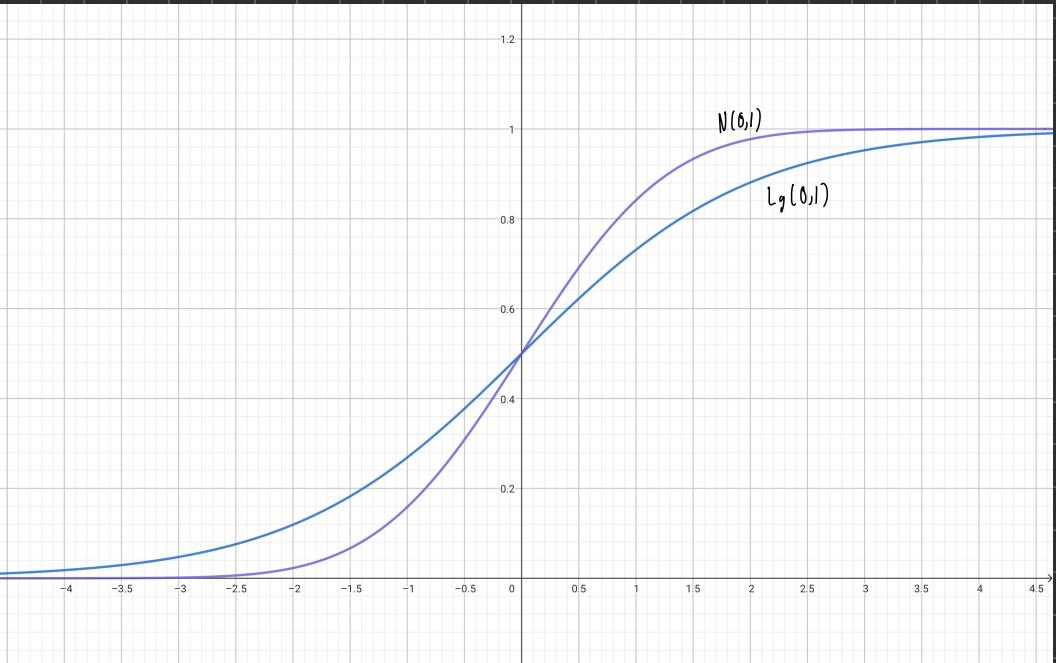
$$f_X(x) = \frac{dF_X(x)}{dx} = \frac{d}{dx} \frac{1}{1 + \exp\left(-\frac{x-\alpha}{\beta}\right)} = \left(\frac{1}{1 + \exp\left(-\frac{x-\alpha}{\beta}\right)} \right)^{-2} \left(\exp\left(-\frac{x-\alpha}{\beta}\right) \left(-\frac{1}{\beta} \right) \right)$$

$$\therefore f_x(x) = \frac{\exp\left(-\frac{x-\alpha}{\beta}\right)}{\beta \left(1 + \exp\left(-\frac{x-\alpha}{\beta}\right)\right)^2}$$

iii) En un mismo plano graficar la función de densidad f_x con $\alpha = 0$ y $\beta = 1$ y la función de densidad ϕ de la Normal univariada estándar, $N(0,1)$.



iii) En un mismo plano graficar la función de distribución acumulativa F_x con $\alpha = 0$ y $\beta = 1$ y la función de distribución acumulativa Φ de la $N(0,1)$.



3. Modelo logit. Ilustración de lo visto en clase.

- i) Generar (usando algún paquete estadístico) una muestra $\{U_i\}_{i=1}^{200}$ i.i.d. donde $U_i \sim Lg(0, 1)$
 $\forall i \in \{1, 2, \dots, 200\}$

```
set.seed(1)
U <- rlogis(n = 200, location = 0, scale = 1)
```

- ii) Generar (usando algún paquete estadístico) una muestra $\{X_i\}_{i=1}^{200}$ i.i.d. (de cualquier distribución), tal que $\mathbb{E}[U_i|X_i] = 0 \forall i \in \{1, 2, \dots, 200\}$

```
X <- rnorm(n = 200, mean = 0, sd = 4)
```

- iii) Generar $\{Y_i^*\}_{i=1}^{200}$, donde $Y_i^* = a + bX_i + U_i$, $i = 1, 2, \dots, 200$; con $a = 0.5$ y $b = 0.4$

```
a <- 0.5
b <- 0.4
Y_star <- a + b*X + U
```

- iv) Generar la muestra $\left\{ \begin{pmatrix} Y_i \\ X_i \end{pmatrix} \right\}_{i=1}^{200}$ donde $Y_i = \begin{cases} 1 & \text{si } Y_i^* > 0 \\ 0 & \text{si } Y_i^* \leq 0 \end{cases}, i = 1, 2, \dots, 200$

```
Y <- ifelse(Y_star>0, yes = 1, no = 0)
Rvec <- data.frame(X, Y)
```

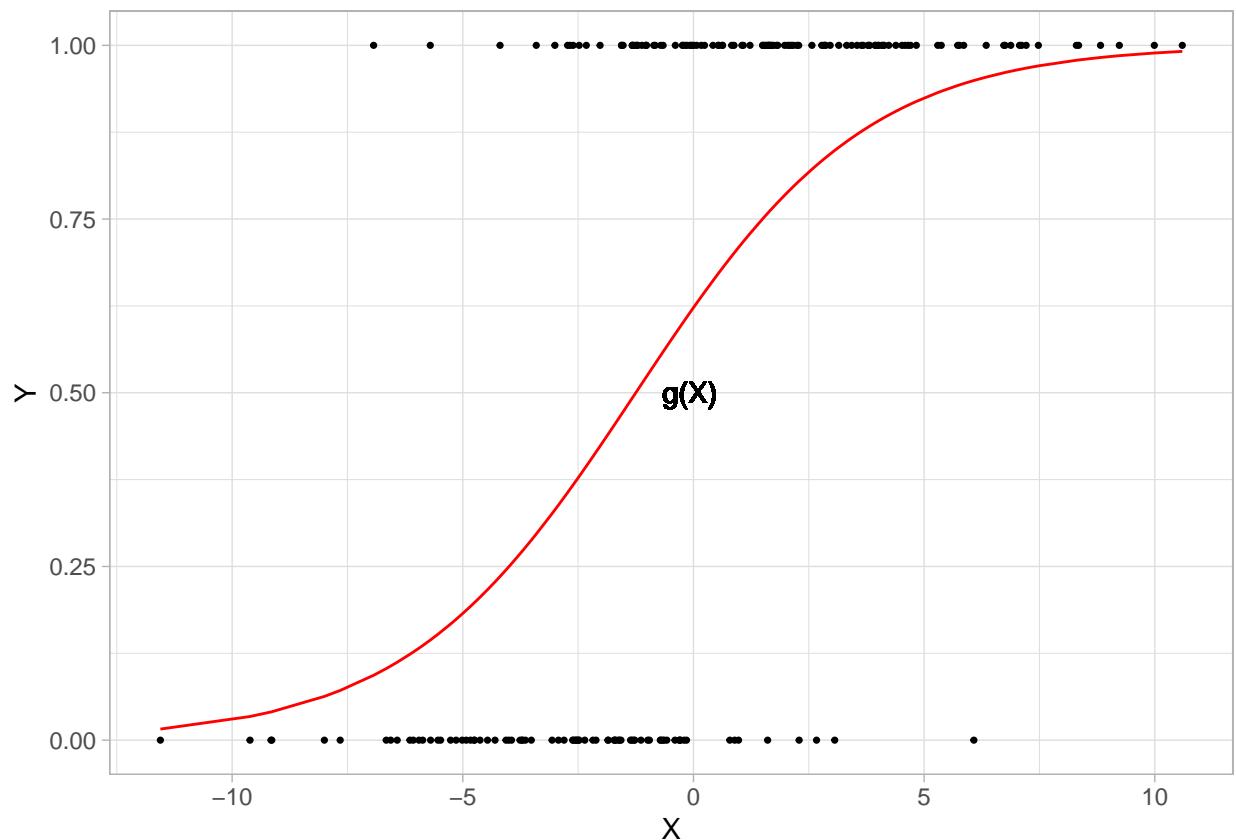
- v) En un mismo plano, graficar (usando los valores dados de a y b) la función de regresión (obtenida en clase)

$$\mathbb{E}[Y|X = x] = 1 - \frac{1}{1+exp(a+bx)}$$

junto con la muestra $\left\{ \begin{pmatrix} Y_i \\ X_i \end{pmatrix} \right\}_{i=1}^{200}$ obtenida en iv)

```
library(ggplot2)
g <- 1- (1/(1+exp(a+b*X)))

ggplot(data.frame(X,Y,g)) +
  geom_point(aes(x=X, y=Y),
             size=0.6) +
  geom_line(aes(x=X,y=g),
            col = "red") +
  theme_light() +
  geom_label(
    label="g(X)",
    x=X[match(100, order(X))]+0.5,
    y=.5,
    size=4,
    label.size = NA,
    fill = NA
  )
```



```
rm(list = ls())
```

4. Modelo probit (normit). Ilustración de lo visto en clase.

- i) Generar (usando algún paquete estadístico) una muestra $\{U_i\}_{i=1}^{200}$ i.i.d. donde $U_i \sim N(0, 1)$
 $\forall i \in \{1, 2, \dots, 200\}$

```
U <- rnorm(n = 200, mean = 0, sd = 1)
```

- ii) Generar (usando algún paquete estadístico) una muestra $\{X_i\}_{i=1}^{200}$ i.i.d. (de cualquier distribución), tal que $\mathbb{E}[U_i|X_i] = 0 \forall i \in \{1, 2, \dots, 200\}$

```
X <- rnorm(n = 200, mean = 0, sd = 4)
```

- iii) Generar $\{Y_i^*\}_{i=1}^{200}$, donde $Y_i^* = a + bX_i + U_i$, $i = 1, 2, \dots, 200$; con $a = 0.5$ y $b = 0.4$

```
a <- 0.5
b <- 0.4
Y_star <- a + b*X + U
```

- iv) Generar la muestra $\left\{ \begin{pmatrix} Y_i \\ X_i \end{pmatrix} \right\}_{i=1}^{200}$ donde $Y_i = \begin{cases} 1 & \text{si } Y_i^* > 0 \\ 0 & \text{si } Y_i^* \leq 0 \end{cases}, i = 1, 2, \dots, 200$

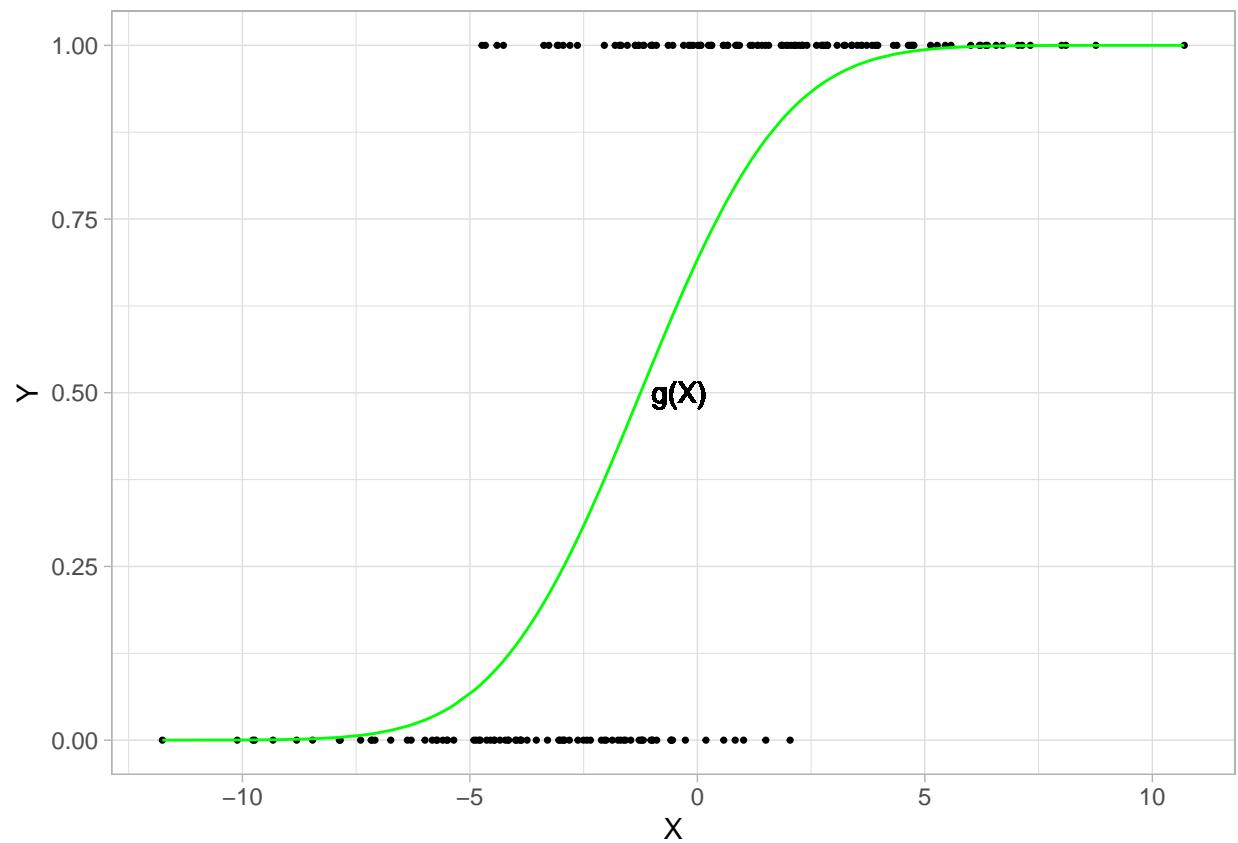
```
Y <- ifelse(Y_star>0, yes = 1, no = 0)
Rvec <- data.frame(X, Y)
```

- v) En un mismo plano, graficar (usando los valores dados de a y b) la función de regresión (obtenida en clase)

$g(x) = \mathbb{E}[Y|X = x] = 1 - \Phi(-(a + bx)) = \Phi(a + bx)$
junto con la muestra $\left\{ \begin{pmatrix} Y_i \\ X_i \end{pmatrix} \right\}_{i=1}^{200}$ obtenida en iv)

```
library(ggplot2)
g <- pnorm(a+b*X)

ggplot(data.frame(X,Y,g)) +
  geom_point(aes(x=X, y=Y),
             size=0.6) +
  geom_line(aes(x=X,y=g),
            col = "green") +
  theme_light() +
  geom_label(
    label="g(X)",
    x=X[match(100, order(X))] + 0.5,
    y=.5,
    size=4,
    label.size = NA,
    fill = NA
  )
```



5. $\begin{pmatrix} Y \\ X \end{pmatrix}$ vector aleatorio bidimensional.

$$f_{x,y}: \mathbb{R}^2 \rightarrow \mathbb{R}$$

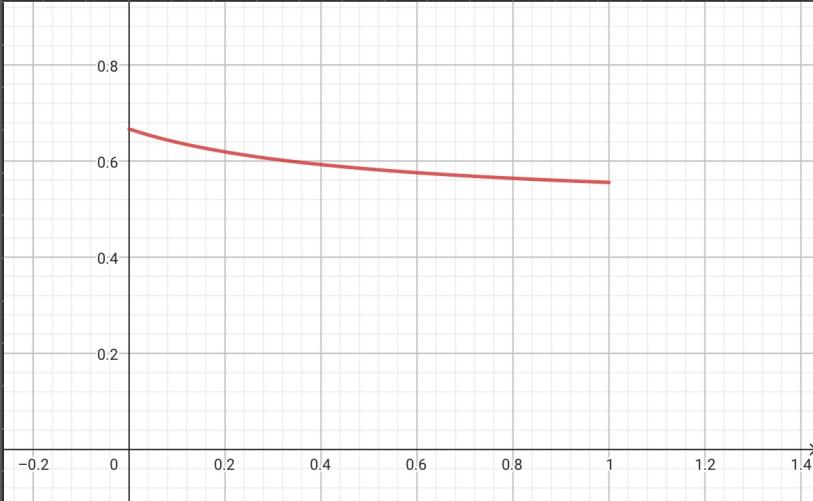
$$f_{x,y}(x, y) = \begin{cases} x+y & \text{si } x \in (0, 1) \text{ e } y \in (0, 1) \\ 0 & \text{de lo contrario} \end{cases}$$

Sabemos que la función de regresión es:

$$g: (0, 1) \rightarrow \mathbb{R}$$

$$g(x) = \frac{x + 2/3}{2x + 1}$$

i) Graficar la función de regresión $g: (0, 1) \rightarrow \mathbb{R}$



ii) Verificar que la función de cedasticidad es:

$$h: (0, 1) \rightarrow \mathbb{R}$$

$$h(x) = \frac{2x^2 + 2x + 1/3}{6(2x+1)^2}$$

1) (Calculamos) $\int_x(y)$:

$$\int_x(y) = \int_0^1 (x+y) dy = \left[xy + \int_0^1 y dy \right]_0^1 = \left[xy + \frac{y^2}{2} \right]_0^1 = x + \frac{1}{2}$$

2) (Calculamos) la f. de cedasticidad ($h(x)$):

$$\int_y(x) = \text{var}(Y|X) = E(Y^2|X) - E^2(Y|X) = \int_0^1 y^2 \frac{x+y}{x+\frac{1}{2}} dy - \frac{(x+\frac{1}{2})^2}{(2x+1)^2} = \frac{1}{2x+1} \int_0^1 (y^2 x + y^3) dy - \frac{x^2 + \frac{4}{3}x + \frac{4}{9}}{(2x+1)^2}$$

$$\begin{aligned}
&= \frac{1}{x+1} \left[\int_0^1 y^2 x \, dy + \int_0^1 y^3 \, dy \right] - \frac{x^2 + \frac{4}{3}x + \frac{4}{9}}{(2x+1)^2} = \frac{1}{x+1} \left(x \left. \frac{y^3}{3} \right|_0^1 + \left. \frac{y^4}{4} \right|_0^1 \right) - \frac{x^2 + \frac{4}{3}x + \frac{4}{9}}{(2x+1)^2} \\
&= \frac{1}{x+1} \left(\frac{x}{3} + \frac{1}{4} \right) - \frac{x^2 + \frac{4}{3}x + \frac{4}{9}}{(2x+1)^2} = \frac{\frac{x}{3} + \frac{1}{4}}{x+1} - \frac{x^2 + \frac{4}{3}x + \frac{4}{9}}{(2x+1)^2} = \frac{\frac{2}{3}x + \frac{1}{2}}{2x+1} - \frac{x^2 + \frac{4}{3}x + \frac{4}{9}}{(2x+1)^2} \\
&= \frac{(2x+1) \left(\frac{2}{3}x + \frac{1}{2} \right) - x^2 - \frac{4}{3}x - \frac{4}{9}}{(2x+1)^2} = \frac{\frac{4}{3}x^2 + x + \frac{2}{3}x + \frac{1}{2} - x^2 - \frac{4}{3}x - \frac{4}{9}}{(2x+1)^2} \\
&= \frac{\frac{1}{3}x^2 + \frac{1}{3}x + \frac{1}{18}}{(2x+1)^2} \cdot \frac{6}{6} = \frac{\frac{6}{3}x^2 + \frac{6}{3}x + \frac{6}{18}}{6(2x+1)^2} = \boxed{\frac{2x^2 + 2x + 1/3}{6(2x+1)^2}}
\end{aligned}$$

iii) Graficar la función de cedasticidad $h:(0, 1) \rightarrow \mathbb{R}$

