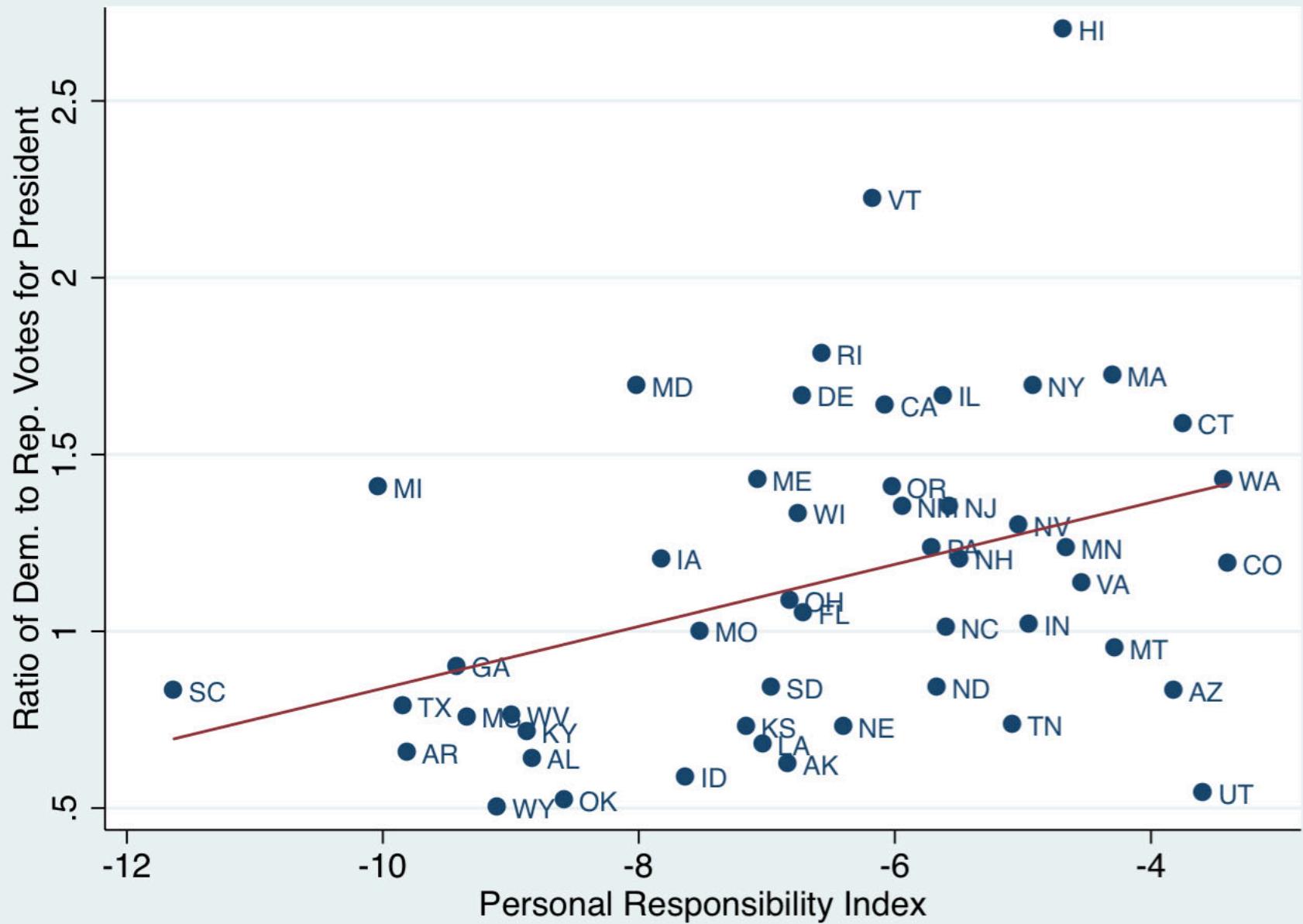


# Who said?

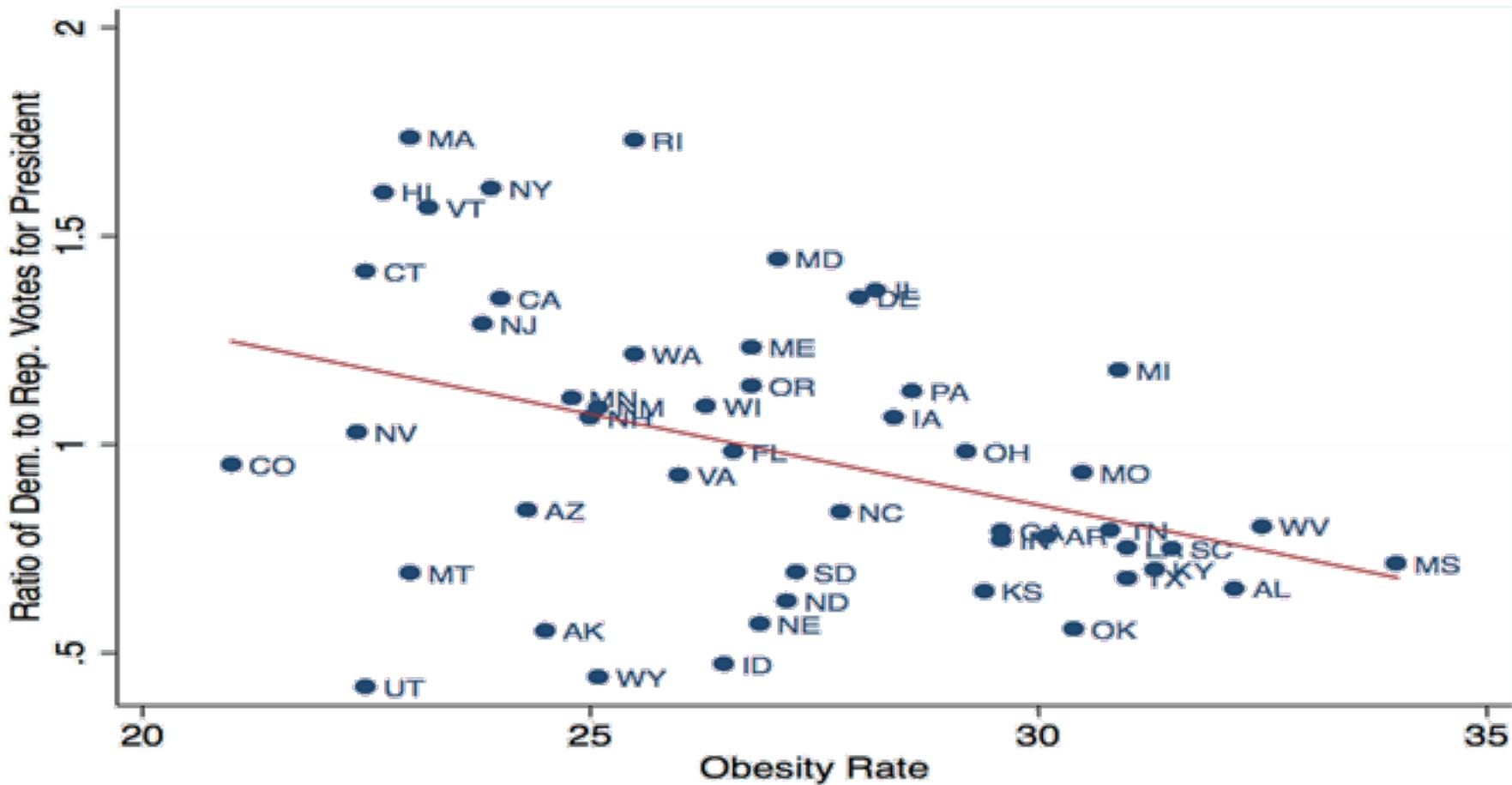
- « There are 47 percent of the people who will vote for the president no matter what. All right, there are 47 percent who are with him, who are dependent upon government, who believe that they are victims, who believe the government has a responsibility to care for them, who believe that they are entitled to health care, to food, to housing, to you-name-it. That that's an entitlement. And the government should give it to them. And they will vote for this president no matter what... »

# Pour Mitt Romney, 47% des Américains sont des assistés

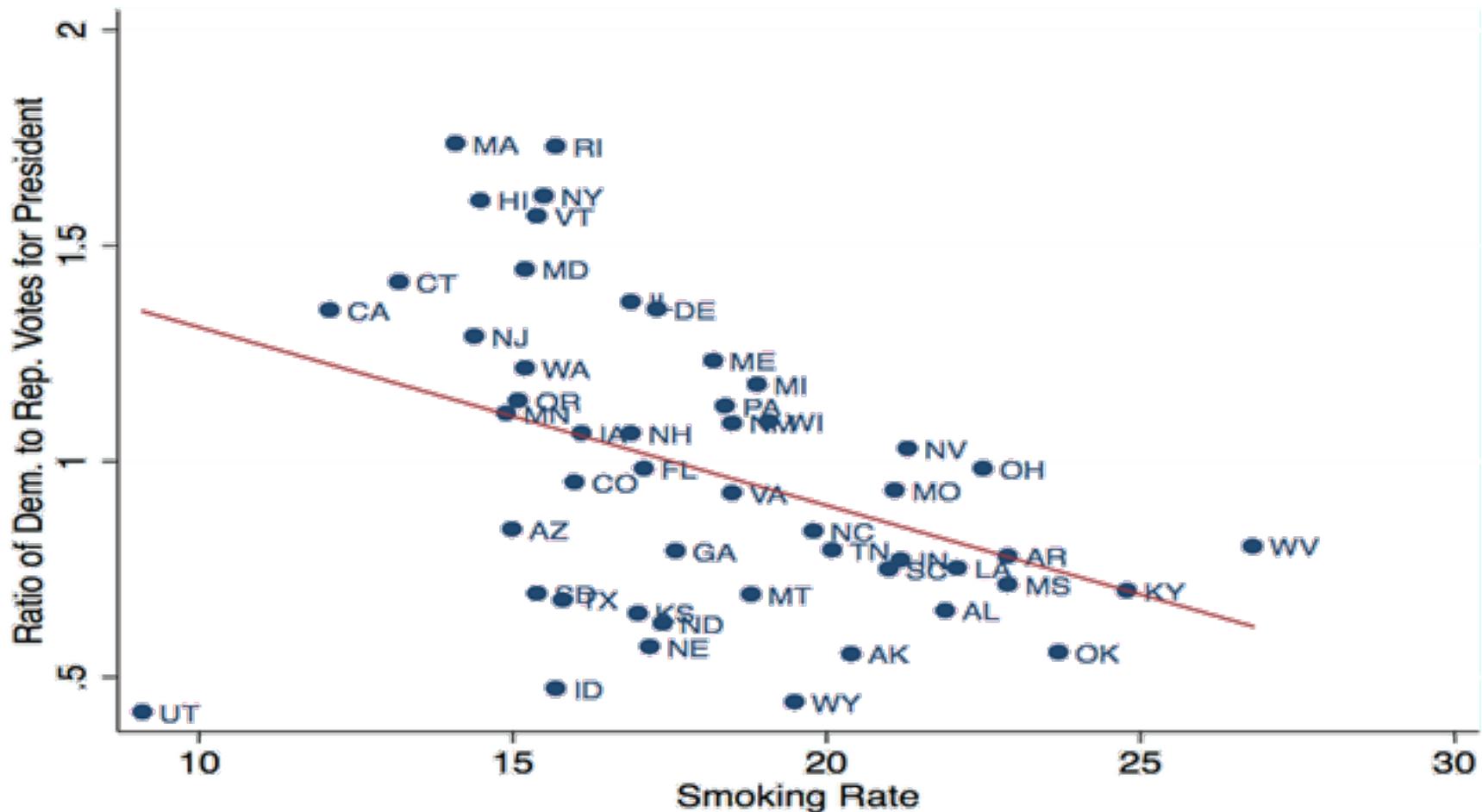
- La première vidéo filmée à son insu le montre répondant à une question d'un de ses contributeurs. Alors que celui-ci lui demande comment il compte l'emporter en novembre, Mitt Romney explique que 47% des Américains sont déjà perdus pour lui. « Ils voteront pour le Président quoi qu'il arrive. 47% des votants sont derrière lui, ils dépendent du gouvernement, ils se prennent pour des victimes, ils pensent que le gouvernement a l'obligation de s'occuper d'eux, ils pensent avoir droit à une couverture maladie, à être nourris, logés, et à tout le reste. Ils pensent que c'est un droit. »



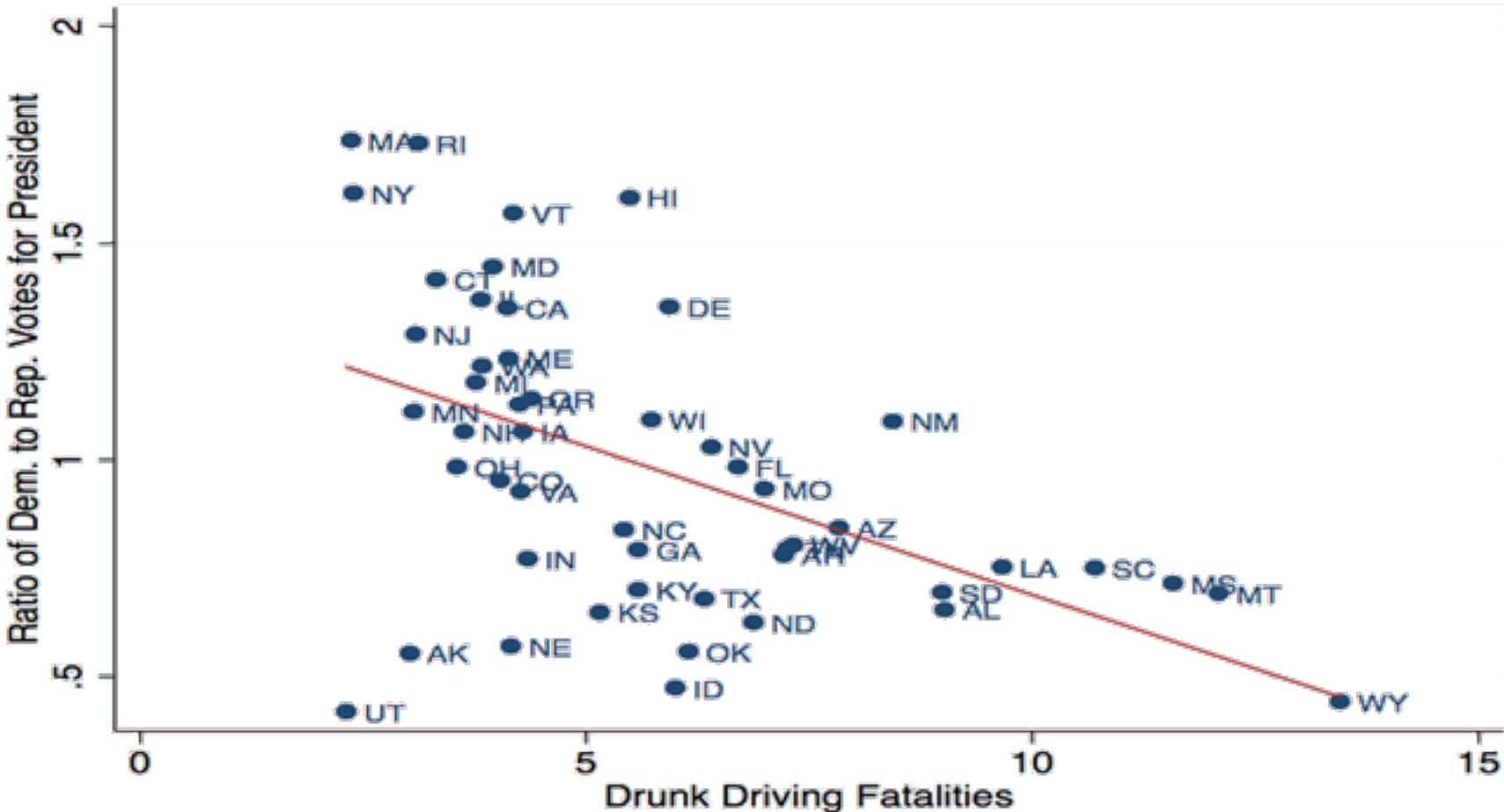
# Liaison entre deux variables quantitatives



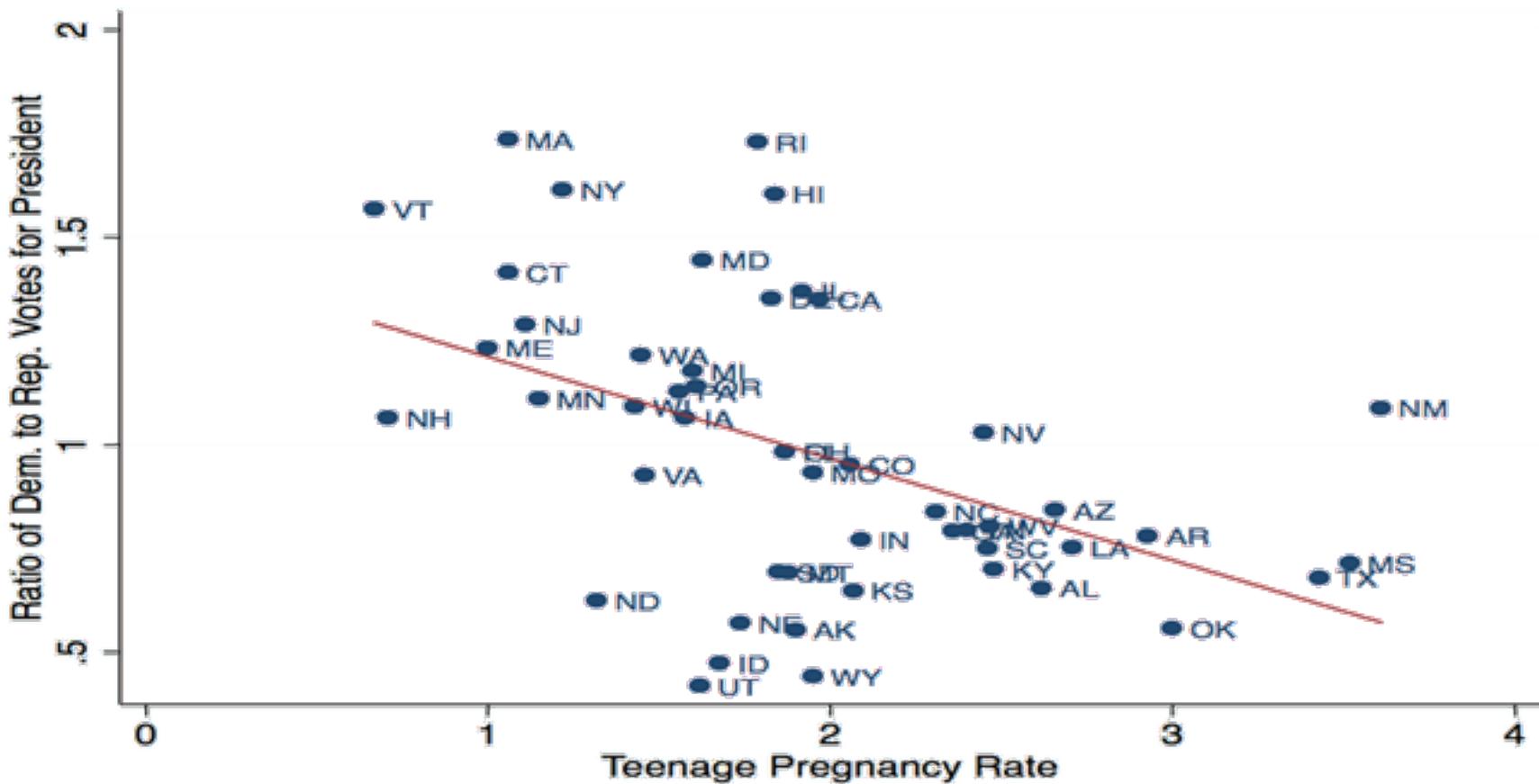
# Liaison entre deux variables quantitatives



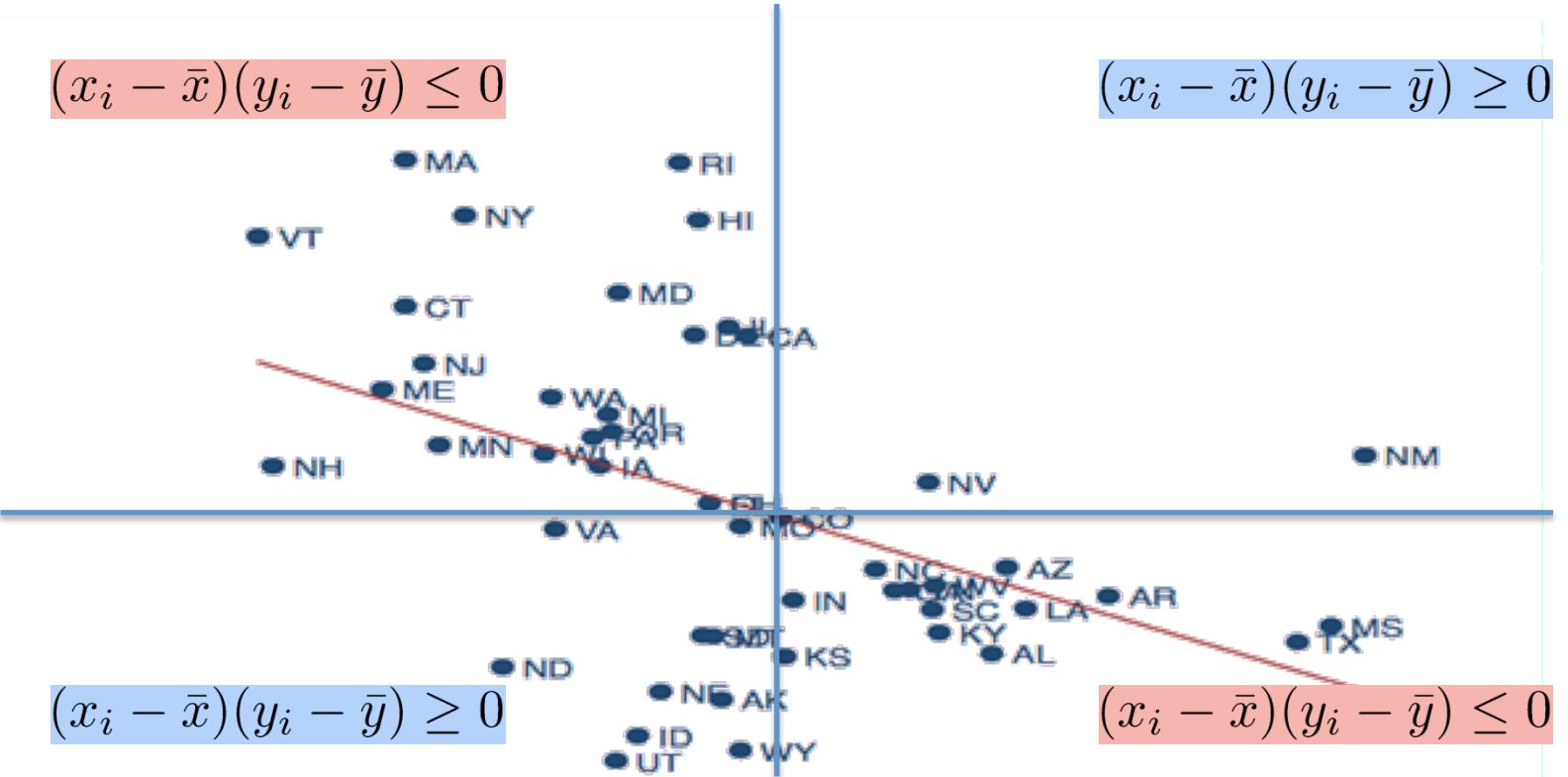
# Liaison entre deux variables quantitatives



# Liaison entre deux variables quantitatives



# Liaison entre deux variables quantitatives



# Liaison entre deux variables quantitatives

$$r = \frac{1}{n} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{s_x} \right) \left( \frac{y_i - \bar{y}}{s_y} \right)$$

# Liaison entre deux variables quantitatives

$$r = \frac{1}{n} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{s_x} \right) \left( \frac{y_i - \bar{y}}{s_y} \right)$$

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{Cov(x, y)}{s_x s_y}$$

# Modèle de régression simple

ANTHROPOLOGICAL MISCELLANEA.

---

REGRESSION *towards MEDIOCRITY in HEREDITARY STATURE.*

By FRANCIS GALTON, F.R.S., &c.

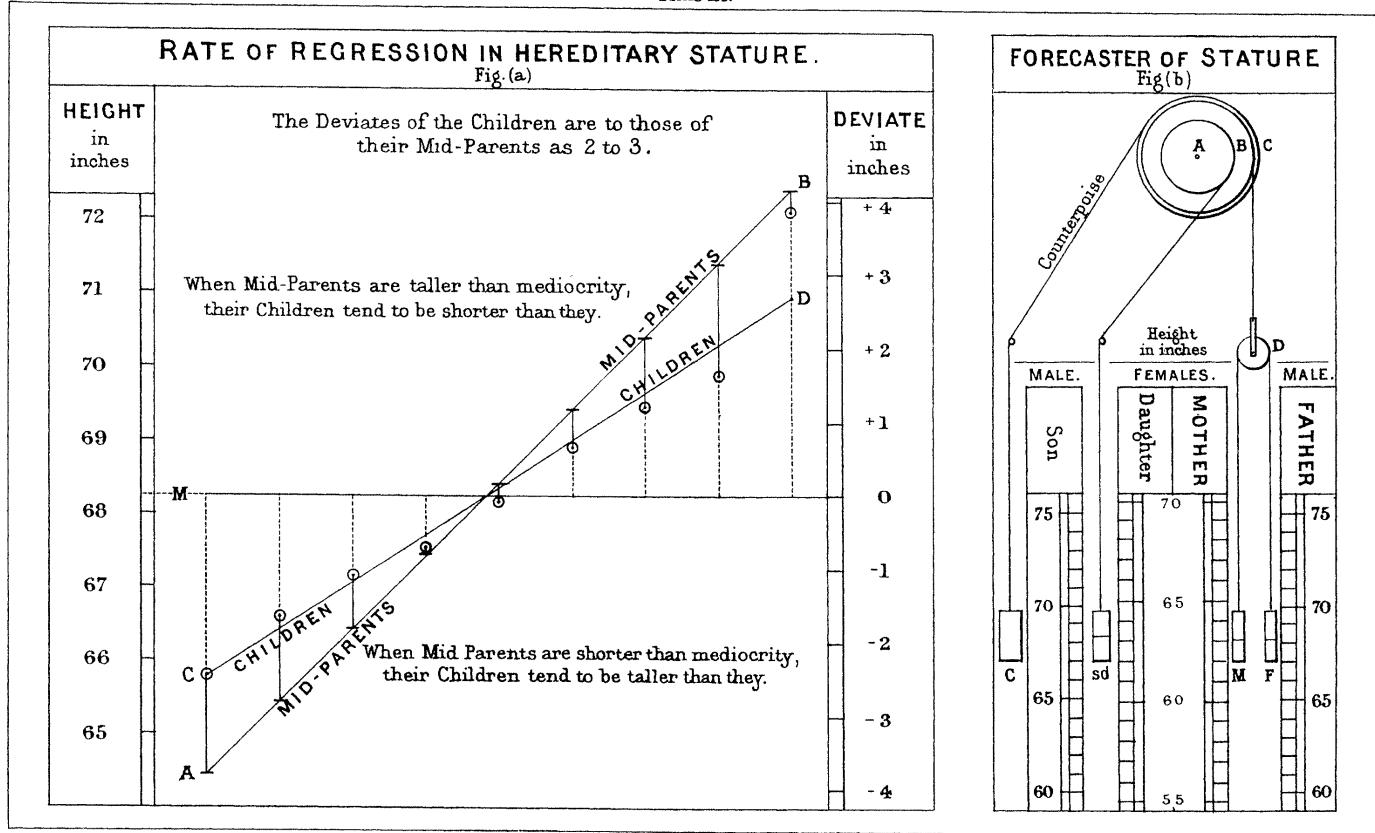
[WITH PLATES IX AND X.]

# Modèle de régression simple

My data consisted of the heights of 930 adult children and of their respective parentages, 205 in number. In every case I transmuted the female statures to their corresponding male equivalents and used them in their transmuted form, so that no objection grounded on the sexual difference of stature need be raised when I speak of averages. The factor I used was 1·08, which is equivalent to adding a little less than one-twelfth to each female height. It differs a very little from the factors employed by other anthropologists, who, moreover, differ a trifle between themselves; anyhow, it suits my data better than 1·07 or 1·09. The final result is not of a kind to be affected by these minute details, for it happened that, owing to a mistaken direction, the computer to whom I first entrusted the figures used a somewhat different factor, yet the result came out closely the same.

# Modèle de régression simple

Plate IX.



# Modèle de régression simple

$$Y_i = \alpha + \beta x_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2)$$

# Estimation des paramètres

$$y_i = a + bx_i + e_i$$

# Estimation des paramètres

- Critère des moindres carrés

$$y_i = a + bx_i + e_i$$

$$\sum e_i^2 \text{ minimum}$$

# Pente (mathématiques)

Pour les articles homonymes, voir [Pente](#).



Cet article est une ébauche concernant les mathématiques.

Vous pouvez partager vos connaissances en l'améliorant ([comment ?](#)) selon les recommandations des projets correspondants.

La **pente** est le coefficient directeur dans un repère cartésien orthonormé.

En géométrie cartésienne, le coefficient directeur désigne le coefficient  $a$  de l'équation d'une droite,  $y = ax + b$ . Cette quantité représente la variation de l'ordonnée  $y$  lorsque l'abscisse  $x$  augmente d'une unité. (Cette définition exclut les droites parallèles au deuxième axe de coordonnée).

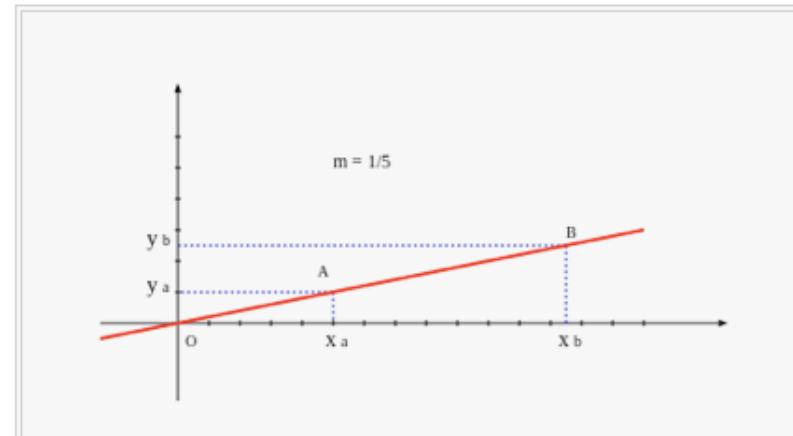
La pente d'une droite (non parallèle à l'axe  $Oy$ ) correspond au rapport entre la variation de  $y$  et la variation correspondante de  $x$ . Cela correspond donc également à la tangente de l'angle que fait la droite avec l'axe  $Ox$ . L'axe  $Ox$  étant interprété comme un axe horizontal, la pente représente le rapport entre la distance verticale et la distance horizontale lorsqu'on suit le mouvement d'un point sur la droite. Cette pente peut être exprimée par un pourcentage : une pente de 20% correspond par exemple à un coefficient directeur de 1/5.

Si une fonction réelle est dérivable en un point, sa courbe représentative admet une tangente en ce point dont la pente est égale au nombre dérivé de l'abscisse en ce point.

## Calcul du coefficient directeur d'une droite déterminée par deux de ses points [modifier]

- Soit une droite parallèle à l'axe  $Oy$ , le coefficient directeur est alors  $+\infty$ .
- Si la droite n'est pas parallèle à l'axe  $Oy$ , et si l'on connaît deux points distincts  $A(x_A, y_A)$  et  $B(x_B, y_B)$ , le coefficient directeur  $m$  de cette droite vaut :

$$m = \frac{y_B - y_A}{x_B - x_A}$$



# Estimation des paramètres

$$\begin{aligned} b &= \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} \\ &= \frac{Cov(x, y)}{Var(x)} \end{aligned}$$

# Estimation des paramètres

$$\begin{aligned} b &= \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} \\ &= \frac{Cov(x, y)}{Var(x)} \end{aligned}$$

$$a = \bar{y} - b\bar{x}$$

# Propriétés des estimateurs

$$B = \frac{\sum_i (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_i (x_i - \bar{x})^2}$$

$$A = \bar{Y} - B\bar{x}$$

# Propriétés des estimateurs

$$B = \frac{\sum_i (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_i (x_i - \bar{x})^2}$$

$$E(B) = \beta$$

$$Var(B) = \frac{\sigma^2}{\sum_i (x_i - \bar{x})^2} = \frac{\sigma^2}{n s_x^2}$$

# Propriétés des estimateurs

$$B = \frac{\sum_i (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_i (x_i - \bar{x})^2}$$

$$A = \bar{Y} - B\bar{x}$$

$$E(B) = \beta \quad E(A) = \alpha$$

# Ecart au modèle

$$\hat{y}_i = a + b x_i$$

$$e_i = y_i - \hat{y}_i$$

Par construction  $\sum_i e_i = 0$

# Ecart au modèle

$$\hat{y}_i = a + b x_i$$

$$e_i = y_i - \hat{y}_i$$

Par construction  $\sum_i e_i = 0$

$$\hat{\sigma}^2 = \frac{\sum_i (y_i - \hat{y}_i)^2}{n - 2} = \frac{\sum_i e_i^2}{n - 2}$$

# Test de signification du coefficient de régression

$$H_0 : \beta = 0$$

$$H_1 : \beta \neq 0$$

# Test de signification du coefficient de régression

$$Var(B) = \frac{\sigma^2}{\sum_i (x_i - \bar{x})^2} = \frac{\sigma^2}{ns_x^2}$$

$$\hat{\sigma}_B^2 = \frac{\hat{\sigma}^2}{\sum_i (x_i - \bar{x})^2} = \frac{\hat{\sigma}^2}{ns_x^2}$$

# Test de signification du coefficient de régression

$$Var(B) = \frac{\sigma^2}{\sum_i (x_i - \bar{x})^2} = \frac{\sigma^2}{ns_x^2}$$

$$\hat{\sigma}_B^2 = \frac{\hat{\sigma}^2}{\sum_i (x_i - \bar{x})^2} = \frac{\hat{\sigma}^2}{ns_x^2}$$

$$t_{obs} = \frac{B}{\hat{\sigma}_B}$$

# Décomposition de la variabilité

$$y_i - \bar{y} = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)$$

$$\sum_i (y_i - \bar{y})^2 = \sum_i (\hat{y}_i - \bar{y})^2 + \sum_i (y_i - \hat{y}_i)^2$$

# « Bibliographie »

- <http://www.jolpress.com/international-mitt-romney-video-cachee-embarrassante-levee-de-fonds-marketing-barack-obama-candidat-republicain-elections-americaines-article-813435.html>
- <http://www.project-syndicate.org/commentary/mitt-romney-rejects-his-natural-voters-by-jeffrey-frankel>
- <http://www.ldjackson.net/mitt-romney-47-of-americans-are-dependent-on-government/>
- <http://www.voxeu.org/article/which-50-states-practise-personal-responsibility>