# Chapter 2
# Linear Models

In this chapter, a brief introduction of linear models is presented. Linearity can be interpreted in terms of both linearity in parameters or linearity in variables. In this book, we have considered linearity in parameters of a model. Linear models may generally include regression models, analysis of variance models, and analysis of covariance models. As the focus of this book is to address various generalized linear models for repeated measures data using GLM and Markov chain/process, we have reviewed regression models in this chapter very briefly.

## 2.1 Simple Linear Regression Model

Let us consider a random sample of n pairs of observations $(Y_1,\ X_1), \ldots,\ (Y_n,\ X_n)$. Here, let $Y$ be the dependent variable or outcome and $X$ be the independent variable or predictor. Then the simple regression model or the regression model with a single predictor is denoted by

$$E(Y|X) = \beta_0 + \beta_1 X. \tag{2.1}$$

It is clear from (2.1) that the simple regression model is a population averaged model. Here $E(Y\big|X) = \mu_{Y|X}$ represents conditional expectation of $Y$ for given $X$ . In other words,

$$\mu_{Y|X} = \beta_0 + \beta_1 X \tag{2.2}$$

which can be visualized from the figure displayed below (Fig. 2.1).



M.A. Islam and R.I. Chowdhury, *Analysis of Repeated Measures Data*,
DOI 10.1007/978-981-10-3794-8_2

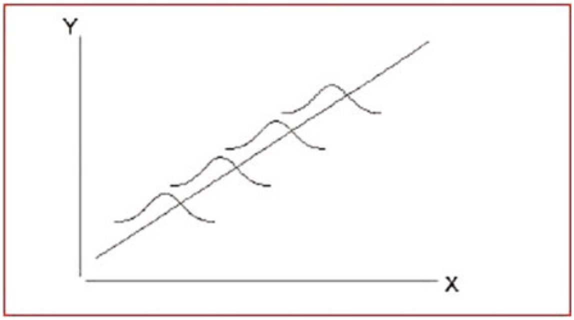


**Fig. 2.1** Population Regression Model

An alternative way to represent model (2.1) or (2.2) is

$$Y = \beta_0 + \beta_1 X + \varepsilon \tag{2.3}$$

where $\varepsilon$ denotes the distance of Y from the conditional expectation or conditional mean, $\mu_{Y|X}$, as evident from expression shown below:

$$Y = \mu_{Y|X} + \varepsilon \tag{2.4}$$

where $\varepsilon$ denotes the error in the dependent or outcome variable, $Y$, attributable to the deviation from the population averaged model and $\varepsilon$ is a random variable as well with $E(\varepsilon) = 0$ and $Var(\varepsilon) = \sigma^2$.

## 2.2 Multiple Regression Model

We can extend the simple regression model shown in Sect. 2.1 for multiple regression model with $p$ predictors $X_1, \ldots, X_p$. The population averaged model can be shown as

$$E(Y|X) = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p. \tag{2.5}$$

Here $E(Y|X) = \mu_{Y|X}$ as shown in Sect. 2.1.

Alternatively,

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \varepsilon \tag{2.6}$$

which can be expressed as

$$Y = \mu_{Y|X} + \varepsilon. \tag{2.7}$$

In vector and matrix notation, the model in Eq. (2.6) for a sample of size n is

$$Y = X\beta + \varepsilon \tag{2.8}$$

where

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, \; X = \begin{pmatrix} 1 & X_{11} & \ldots & X_{1p} \\ 1 & X_{21} & \ldots & X_{1p} \\ \vdots & & & \\ 1 & X_{n1} & \ldots & X_{np} \end{pmatrix}, \; \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \; \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

It is clear from the formulation of regression model that it provides a theoretical framework for explaining the underlying linear relationships between explanatory and outcome variables of interest. A perfect model can be obtained only if all the values of the outcome variable are equal to conditional expectation for given values of predictors which is not feasible in explaining real life problems. However, still it can provide very important insight under the circumstance of specifying a model that keeps the error minimum. Hence, it is important to specify a model that can produce estimate of outcome variable as much close to observed values as possible. In other words, the postulated models in Sects. 2.2 and 2.3 are hypothetical idealized version of the underlying linear relationships which may be attributed to merely association or in some instances causation as well.

The population regression model is proposed under a set of assumptions: (i) $E(\varepsilon_i) = 0$, (ii) $Var(\varepsilon_i) = \sigma^2$, (iii) $E(\varepsilon_i\varepsilon_j) = 0$ for $i \neq j$, and (iv) independence of $X$ and $\varepsilon$. In addition, assumption of normality is necessary for likelihood estimation as well as for testing of hypotheses. Based on these assumptions, we can show the mean and variance of $Y_i$ as follows:

$$E(Y_i|X_i) = X_i\beta, \text{ and } Var(Y_i|X_i) = \sigma^2,$$

where $X_i$ is the ith row vector of the matrix $X$. Using (2.8), we can rewrite the assumptions as follows: (i) $E(\varepsilon) = 0$, and (ii) $Cov(\varepsilon) = \sigma^2 I$. Similarly, $E(Y|X) = X\beta$, and $Cov(Y|X) = \sigma^2 I$.

## 2.3 Estimation of Parameters

For estimating the regression parameters, we can use both method of least squares and method of maximum likelihood. It may be noted here that for extending the concept of linear models to generalized linear models or covariate dependent Markov models, the maximum likelihood method will be used more extensively, hence, both are discussed here although method of least squares is a more convenient method of estimation for linear regression model with desirable properties.

### 2.3.1 Method of Least Squares

The method of least squares is used to estimate the regression parameters by minimizing the error sum of squares or residual sum of squares. The regression model is

$$Y_i = \beta_0 + \beta_1 X_{i1} + \ldots + \beta_p X_{ip} + \varepsilon_i, \quad i = 1,\ 2, \ldots,\ n \tag{2.9}$$

and we can define the deviation between outcome variable and its corresponding conditional mean for given values of $X$ as follows:

$$\varepsilon_i = Y_i - (\beta_0 + \beta_1 X_{i1} + \ldots + \beta_p X_{ip}). \tag{2.10}$$

Then the error sum of squares is defined as a quadratic form

$$Q = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left[Y_i - (\beta_0 + \beta_1 X_{i1} + \ldots + \beta_p X_{ip})\right]^2. \tag{2.11}$$

The sum of squares of error is minimized if the estimates are obtained from the following equations:

$$\left.\frac{\partial Q}{\partial \beta_0}\right|_{\beta=\hat{\beta}} = -2\sum_{i=1}^{n} \left[Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \ldots + \hat{\beta}_p X_{ip})\right] = 0 \tag{2.12}$$

$$\left.\frac{\partial Q}{\partial \beta_j}\right|_{\beta=\hat{\beta}} = -2\sum_{i=1}^{n} \left[Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \ldots + \hat{\beta}_p X_{ip})\right] X_{ij} = 0, \tag{2.13}$$

$j = 1,\ldots,p$. We can consider (2.12) as a special case of Eq. (2.13) for $j = 0$ and $X_0 = 1$.

Using model (2.8), $Q$ can be expressed as

$$Q = \varepsilon'\varepsilon = (Y - X\beta)'(Y - X\beta). \tag{2.14}$$

The right-hand side of (2.14) is

$$Q = Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta$$

where $Y'X\beta = \beta'X'Y$. Hence the estimating equations are

$$\left.\frac{\partial Q}{\partial \beta}\right|_{\beta=\hat{\beta}} = -2X'Y + 2X'X\hat{\beta} = 0. \tag{2.15}$$

Solving Eq. (2.15), we obtain the least squares estimators of regression parameters as shown below:

$$\hat{\beta} = (X'X)^{-1}(X'Y). \tag{2.16}$$

The estimated regression model can be shown as

$$\hat{Y} = X\hat{\beta} \tag{2.17}$$

and alternatively

$$Y = X\hat{\beta} + e \tag{2.18}$$

where

$$\hat{Y} = \begin{pmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_n \end{pmatrix}, \; X = \begin{pmatrix} 1 & X_{11} & \dots & X_{1p} \\ 1 & X_{21} & \dots & X_{2p} \\ \vdots & & & \\ 1 & X_{n1} & \dots & X_{np} \end{pmatrix}, \; \hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{pmatrix}, \; e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}.$$

It may be noted here that **e** is the vector of estimated errors from the fitted model. Hence, we can show that

$$e = Y - \hat{Y} \tag{2.19}$$

and the error sum of squares is

$$e'e = (Y - \hat{Y})'(Y - \hat{Y}). \tag{2.20}$$

#### 2.3.1.1 Some Important Properties of the Least Squares Estimators

The least squares estimators have some desirable properties of good estimators which are shown below.

(i) Unbiasedness: $E(\hat{\beta}) = \beta$.

Proof: We know that $\hat{\beta} = (X'X)^{-1}(X'Y)$ and $Y = X\beta + \varepsilon$. Hence,

$$\begin{aligned} E(\hat{\beta}) &= E[(X'X)^{-1}(X'Y)] \\ &= (X'X)^{-1}X'E(Y) \\ &= (X'X)^{-1}X'E(X\beta + \varepsilon) \\ &= (X'X)^{-1}X'X\beta \\ &= \beta. \end{aligned}$$

(ii) $Cov(\hat{\beta}) = (X'X)^{-1}\sigma^2$.

Proof:

$$\begin{aligned} Cov(\hat{\beta}) &= Cov[(X'X)^{-1}X'Y] \\ &= (X'X)^{-1}X'Cov(Y)X(X'X)^{-1} \end{aligned}$$

where $Cov(Y) = \sigma^2 I$. Hence,

$$\begin{aligned} Cov(\hat{\beta}) &= (X'X)^{-1}X'IX(X'X)^{-1}\sigma^2 \\ &= (X'X)^{-1}\sigma^2. \end{aligned} \tag{2.21}$$

(iii) The least squares estimator $\hat{\beta}$ is the best linear unbiased estimator of $\beta$.
(iv) The mean squared error is an unbiased estimator of $\sigma^2$. In other words,

$$E\left(\frac{e'e}{n-p-1}\right) = \sigma^2 \tag{2.22}$$

Proof: Let us denote $SSE = e'e = (Y - X\hat{\beta})'(Y - X\hat{\beta})$ and $s^2 = \frac{SSE}{n-p-1}$ where $p$ is the number of predictors. Total sum of squares of $Y$ is $Y'Y$. The sum of squares of errors can be rewritten as

$$\begin{aligned} SSE &= Y'Y - Y'X\hat{\beta} - \hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta} \\ &= Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta} \end{aligned}$$

where $Y'X\hat{\beta} = \hat{\beta}'X'Y$. Then replacing $\hat{\beta}$ by $(X'X)^{-1}(X'Y)$, it can be shown that

$$\begin{aligned} SSE &= Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta} = Y'Y - \hat{\beta}'X'Y \\ &= Y'Y - [(X'X)^{-1}X'Y]'X'Y \\ &= Y'Y - Y'X(X'X)^{-1}X'Y \\ &= Y'[I - Y'X(X'X)^{-1}X']Y \end{aligned}$$

It can be shown that the middle term of the above expression is a symmetric idempotent matrix and $\frac{SSE}{\sigma^2}$ is chi-square with degrees of freedom equal to the rank of the matrix $[I - Y'X(X'X)^{-1}X']$. The rank of this idempotent matrix is equal to the $trace[I - Y'X(X'X)^{-1}X']$ which is $n - p - 1$. Hence,

$$E[(n-p-1)(s^2)/\sigma^2] = E(SSE/\sigma^2) = trace[I - Y'X(X'X)^{-1}X'] = n-p-1.$$

This implies $E(SSE) = (n-p-1)\sigma^2$ and $E\left(\frac{SSE}{n-p-1}\right) = \sigma^2$. In other words, the mean square error is an unbiased estimator of $\sigma^2$, i.e. $E(s^2) = \sigma^2$.

### 2.3.2 Maximum Likelihood Estimation

It is noteworthy that the estimation by least squares method does not require normality assumption. However, the estimates of regression parameters can be obtained assuming that $Y \sim N_n(X\beta,\ \sigma^2 I)$ where $E(Y|X) = X\beta$ and $Var(Y|X) = \sigma^2 I$. The likelihood function is

$$\begin{aligned} L(\beta,\ \sigma^2) &= \frac{1}{(2\pi)^{n/2}[\sigma^2 I]^{1/2}} e^{-(Y-X\beta)'(\sigma^2 I)^{-1}(Y-X\beta)/2} \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-(Y-X\beta)'(Y-X\beta)/2\sigma^2}. \end{aligned}$$

The log-likelihood function can be shown as follows:

$$\ln L(\beta,\ \sigma^2) = -\frac{1}{2}n\ \ln(2\pi) - \frac{1}{2}n\ \ln\sigma^2 - \frac{1}{2\sigma^2}(Y-X\beta)'(Y-X\beta). \tag{2.23}$$

Differentiating (2.23) with respect to parameters and equating to zero, we obtain the following equations:

$$\left.\frac{\partial\ \ln L}{\partial\beta}\right|_{\beta=\hat{\beta},\ \sigma^2=\hat{\sigma}^2} = -\frac{1}{2\hat{\sigma}^2}(-2X'Y - 2X'X\hat{\beta}) = 0 \tag{2.24}$$

$$\left.\frac{\partial\ \ln L}{\partial\sigma^2}\right|_{\beta=\hat{\beta},\ \sigma^2=\hat{\sigma}^2} = -\frac{n}{2\hat{\sigma}^2} + \frac{1}{2(\hat{\sigma}^2)^2}(Y-X\hat{\beta})'(Y-X\hat{\beta}) = 0 \tag{2.25}$$

Solving (2.24) and (2.25), we obtain the following maximum likelihood estimators:

$$\hat{\beta} = (X'X)^{-1}(X'Y),$$

and

$$\hat{\sigma}^2 = \frac{1}{n}(Y-X\hat{\beta})'(Y-X\hat{\beta}).$$

#### 2.3.2.1 Some Important Properties of Maximum Likelihood Estimators

Some important properties of maximum likelihood estimators are listed below:

(i) $\hat{\beta} \sim N_{p+1}\left[\beta,\ \sigma^2 (X'X)^{-1}\right]$,
(ii) $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-p-1)$,
(iii) $\hat{\beta}$ and $\hat{\sigma}^2$ are independent,
(iv) If Y is $N_n(X\beta,\ \sigma^2 I)$ then $\hat{\beta}$ and $\hat{\sigma}^2$ are jointly sufficient for $\beta$ and $\sigma^2$, and
(v) If Y is $N_n(X\beta,\ \sigma^2 I)$ then $\hat{\beta}$ have minimum variance among all unbiased estimators.

## 2.4 Tests

In a regression model, we need to perform several tests, such as: (i) significance of the overall fitting of model involving $p$ predictors, (ii) significance of each parameter to test for significant association between each predictor and outcome variable, and (iii) significance of a subset of parameters.

(i) *Test for significance of the model*

In the regression model, $Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \varepsilon$, it is important to examine whether none of the predictors $X_1, \ldots, X_p$ is linearly associated with outcome variable, $Y$, against the hypothesis that at least one of the predictors is linearly associated with outcome variable. As the postulated model represents a hypothetical relationship between population mean and predictors, $E(Y|X) = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p$., the contribution of the model can be tested from the regression sum of squares which indicates the fit of the model for the conditional mean, compared to the error sum of squares that measures deviation of observed values of outcome variable from the postulated linear relationship of predictors with conditional mean. It may be noted here that total sum of squares due to outcome variable can be partitioned into two components for regression and error as shown below:

$$Y'Y = \hat{\beta}'X'Y + (Y - X\hat{\beta})'(Y - X\hat{\beta})$$

where $\hat{\beta}'X'Y$ is the sum of squares of regression (SSR) and $(Y - X\hat{\beta})'(Y - X\hat{\beta})$ is the sum of squares error (SSE).

The coefficient of multiple determination, $R^2$, measures the extent or proportion of linear relationship explained by the multiple linear regression model. This is the squared multiple correlation. The coefficient of multiple determination can be defined as:

$$R^2 = \frac{\text{Regression Sum of Squares}}{\text{Total Sum of Squares}} = \frac{\hat{\beta}'X'Y - n\bar{Y}^2}{Y'Y - n\bar{Y}^2}. \tag{2.26}$$

and the range of $R^2$ is $0 \leq R^2 \leq 1$, 0 indicating that the model does not explain the variation at all and 1 for a perfect fit or 100% is explained by the model.

The null and alternative hypotheses for overall test of the model are:

$H_0 : \beta_1 = \ldots = \beta_p = 0$ and $H_1 : \beta_j \neq 0$, for at least one $j$, $j = 1, \ldots, p$.

Under null hypothesis, sum of squares of regression is $\chi^2_p\sigma^2$ and similarly sum of squares of error is $\chi^2_{n-p-1}\sigma^2$. The test statistic is

$$F = \frac{SSR/p}{SSE/(n-p-1)} \sim F_{p,(n-p-1)}. \tag{2.27}$$

Rejection of null hypothesis indicates that at least one of the variables in the postulated model contributes significantly in the overall or global test.

(ii) *Test for the significance of parameters*

Once we have determined that at least one of the predictors is significant, next step is to identify the variables that exert significant linear relationship with outcome variable. Statistically it is obvious that inclusion of one or more variables in a regression model may result in increase in regression sum of squares and thus decrease in error sum of squares. However, it needs to be tested whether such inclusion is statistically significant or not. These tests will be elaborated in the next section in more details. The first task is to examine each individual parameter separately to identify predictors with statistically significant linear relationship with outcome variable of interest.

The null and alternative hypotheses for testing significance of individual parameters are:

$$H_0 : \beta_j = 0 \text{ and } H_1 : \beta_j \neq 0.$$

The test statistic is

$$t = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \tag{2.28}$$

which follows a $t$ distribution with $(n - p - 1)$ degrees of freedom. We know that $Cov(\hat{\beta}) = (X'X)^{-1}\sigma^2$ and estimate for the covariance matrix is $\hat{C}ov(\hat{\beta}) = (X'X)^{-1}s^2$ where $s^2$ is the unbiased estimator of $\sigma^2$. The standard error of $\hat{\beta}_j$ can be obtained from corresponding diagonal elements of the inverse matrix $(X'X)^{-1}$. In this rejection of null hypothesis implies a statistically significant linear relationship with outcome variable.

(iii) *Extra Sum of Squares Method*

As we mentioned in the previous section that inclusion of a variable may result in increase in SSR and subsequently decrease in SSE, it needs to be tested whether the increase in SSR is statistically significant or not. In addition, it is also possible to test whether inclusion or deletion of a subset of potential predictors result in any statistically significant change in the fit of the model or not. For this purpose, extra sum of squares principle may be a very useful procedure.

Let us consider a regression model

$$Y = X\beta + \varepsilon$$

where $Y$ is $n \times 1$, $X$ is $n \times k$, $\beta$ is $k \times 1$, and $k = p + 1$. If we partition $\beta$ as follows

$$\beta = \begin{pmatrix} \boldsymbol{\beta_1} \\ \hline \boldsymbol{\beta_2} \end{pmatrix}$$

where

$$\beta = \left( \begin{array}{c} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{r-1} \\ \hline \beta_r \\ \vdots \\ \beta_p \end{array} \right), \boldsymbol{\beta}_1 = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{r-1} \end{pmatrix} \text{ and } \boldsymbol{\beta}_2 = \begin{pmatrix} \beta_r \\ \vdots \\ \beta_p \end{pmatrix}.$$

We can express the partitioned regression model as

$$Y = X_1\boldsymbol{\beta}_1 + X_2\boldsymbol{\beta}_2 + \varepsilon \tag{2.29}$$

where

$$X_1 = \begin{pmatrix} 1 & X_{11} & \dots & X_{1,r-1} \\ 1 & X_{21} & \dots & X_{2,r-1} \\ \vdots & & & \\ 1 & X_{n1} & \dots & X_{n,r-1} \end{pmatrix}, \ X_2 = \begin{pmatrix} X_{1,r} & \dots & X_{1,p} \\ X_{2,r} & \dots & X_{2,p} \\ \vdots & & \\ X_{n,r} & \dots & X_{n,p} \end{pmatrix}.$$

Let us consider this model as the full model. In other words, the full model is comprised of all the variables under consideration. We want to test, whether some of the variables or a subset of the variables included in the full model contributes significantly or not. This subset may include one or more variables and the corresponding coefficients or regression parameters are represented by the vector $\boldsymbol{\beta}_2$. Hence, a test on whether $\boldsymbol{\beta}_2 = 0$ is an appropriate null hypothesis here. This can be employed for a single parameter as a special case.

Regression and error sum of squares from full and reduced models are shown below.

**Full Model:**
Under full model, the SSR and SSE are:

$$\text{SSR (full model)} = \hat{\beta}' X' Y$$
$$\text{SSE (full model)} = Y'Y - \hat{\beta}' X' Y$$

**Reduced Model:**
Under null hypothesis, the SSR and SSE are:

$$\text{SSR (reduced model)} = \hat{\boldsymbol{\beta}}_1' X_1' Y$$
$$\text{SSE (reduced model)} = Y'Y - \hat{\boldsymbol{\beta}}_1' X_1' Y$$

Difference between SSR (full model) and SSR (reduced model) shows the contribution of the variables $X_r, \ldots, X_p$ which can be expressed as:

$$SSR(\beta_{\mathbf{2}}|\beta_{\mathbf{1}}) = \hat{\beta}' X' Y - \hat{\boldsymbol{\beta}}'_1 X_1' Y.$$

This is the extra sum of squares attributable to the variables under null hypothesis.

The test statistic for $H_0 : \boldsymbol{\beta}_2 = 0$ is

$$F = \frac{SSR(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)/(k - r + 1)}{s^2} \sim F_{(k-r+1),(n-k)}. \tag{2.30}$$

Acceptance of null hypothesis implies there may not be any statistically significant contribution of the variables $X_r, \ldots, X_p$ and the reduced model under null hypothesis is equally good as compared to the full model.

## 2.5 Example

A data set on standardized fertility measure and socioeconomic indicators from Switzerland is used for application in this chapter. This data set is freely available from ‘datasets’ package in R. Full dataset and description are available for download from the Office of Population Research website (site https://opr.princeton.edu/archive/pefp/switz.aspx). Following variables are available in the ‘swiss’ dataset from datasets package. This data set includes indicators for each of 47 French-speaking provinces of Switzerland in 1888. The variables are:

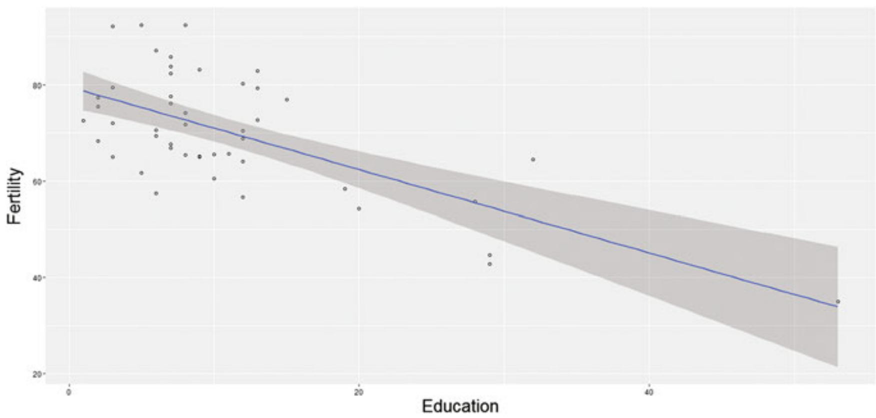


**Fig. 2.2** Simple Linear Regression

| | |
|---|---|
| Fertility | common standardized fertility measure |
| Agriculture | % of males involved in agriculture as occupation |
| Examination | % draftees receiving highest mark on army examination |
| Education | % education beyond primary school for draftees |
| Catholic | % 'catholic' (as opposed to 'Protestant') |
| Infant Mortality | live births who live less than one year. |

Here the first example shows a fitting of a simple regression model where the outcome variable, $Y =$ common standardized fertility measure and $X =$ percent education beyond primary school for draftees. The estimated model is $\hat{Y} = 79.6101 - 0.8624X$. Education appears to be negatively associated with fertility measure in French-speaking provinces (p-value < 0.001). Figure 2.2 displays the negative relationship. Table 2.1 summarizes the results.

**Table 2.1** Estimates and tests of parameters of a simple regression model

| Variable | Estimate | Std. error | t-value | Pr(>\|t\|) |
|---|---|---|---|---|
| Constant | 79.6101 | 2.1041 | 37.836 | 0.000 |
| Education | −0.8624 | 0.1448 | −5.954 | 0.000 |

**Table 2.2** Estimates and tests of parameters of a multiple linear regression model

| Variable | Estimate | Std. error | t-value | Pr(>\|t\|) |
|---|---|---|---|---|
| Constant | 62.10131 | 9.60489 | 6.466 | 0.000 |
| Agriculture | −0.15462 | 0.06819 | −2.267 | 0.029 |
| Education | −0.98026 | 0.14814 | −6.617 | 0.000 |
| Catholic | 0.12467 | 0.02889 | 4.315 | 0.000 |
| Infant Mortality | 1.07844 | 0.38187 | 2.824 | 0.007 |

Using the same data source, an example for the fit of a multiple regression model is shown and the results are summarized in Table 2.2. For the same outcome variable, four explanatory variables are considered, percent males involved in agriculture as profession $(X_1)$, education $(X_2)$, percent catholic $(X_3)$, and infant mortality $(X_4)$. The estimated model for the outcome variable, fertility, is $\hat{Y} = 62.10131 - 0.15462X_1 - 0.98026X_2 + 0.12467X_3 + 1.08844X_4$.

All the explanatory variables show statistically significant linear relationship with fertility, agriculture, and education are negatively but percent catholic and infant mortality are positively related to the outcome variable. The fit of the overall model is statistically significant ($F = 24.42$, D.F. = 4 and 42, p-value $< 0.001$). About 70% ($R^2 = 0.699$) of the total variation is explained by the fitted model.

Springer

http://www.springer.com/978-981-10-3793-1

Analysis of Repeated Measures Data
Islam, M.A.; CHOWDHURY, R.
2017, XIX, 250 p. 6 illus., Hardcover
ISBN: 978-981-10-3793-1