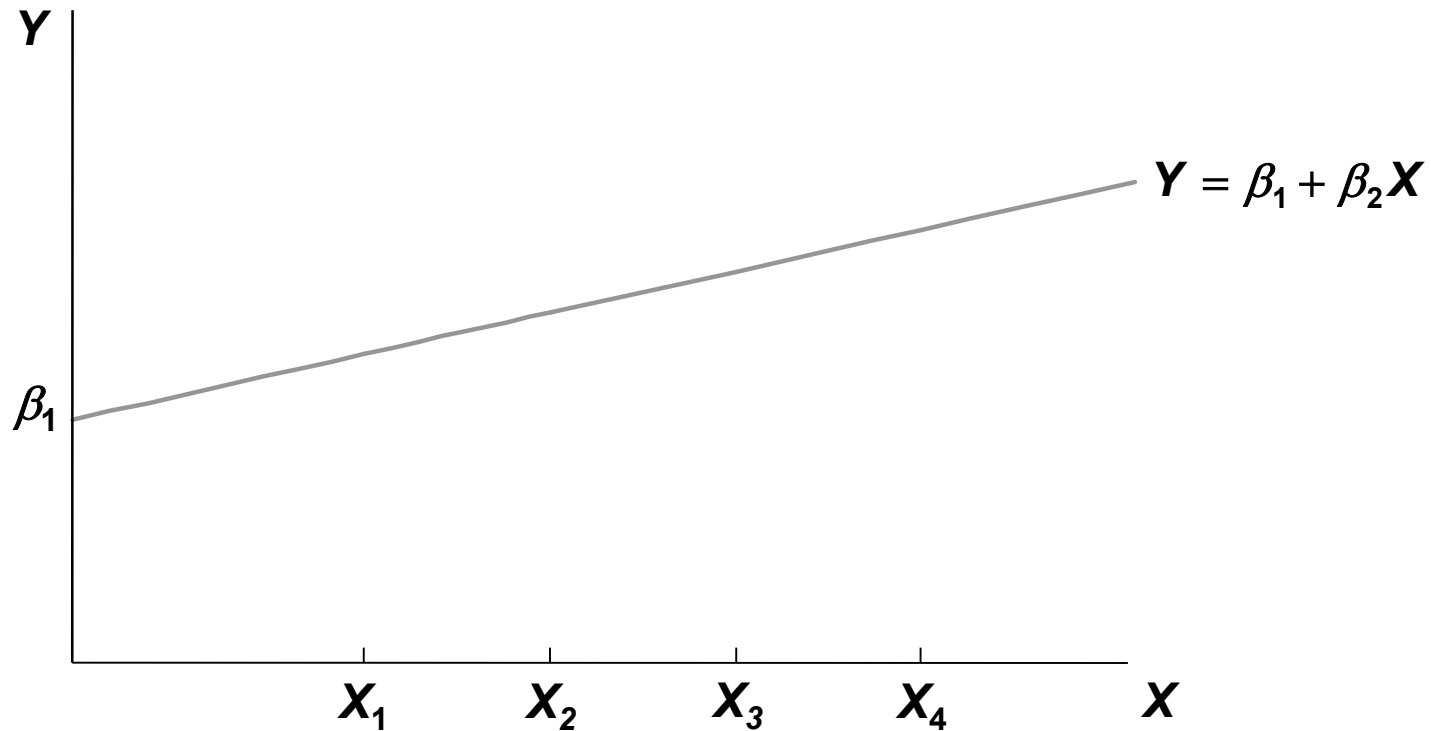


ជំពូកទី២ សេចក្តីផ្តើមជំរឿននាម

Introduction to Simple linear regression

រៀបរៀងនិងដេញដោលយល់ស្តីអំពី ដំណោះស្រាយ
សម្រាប់បញ្ហាដែលសាកលវិទ្យាល័យកូម៉ង់ដ៍នីស្ត្រា
និងវិទ្យាសាស្ត្រសេដ្ឋកិច្ច

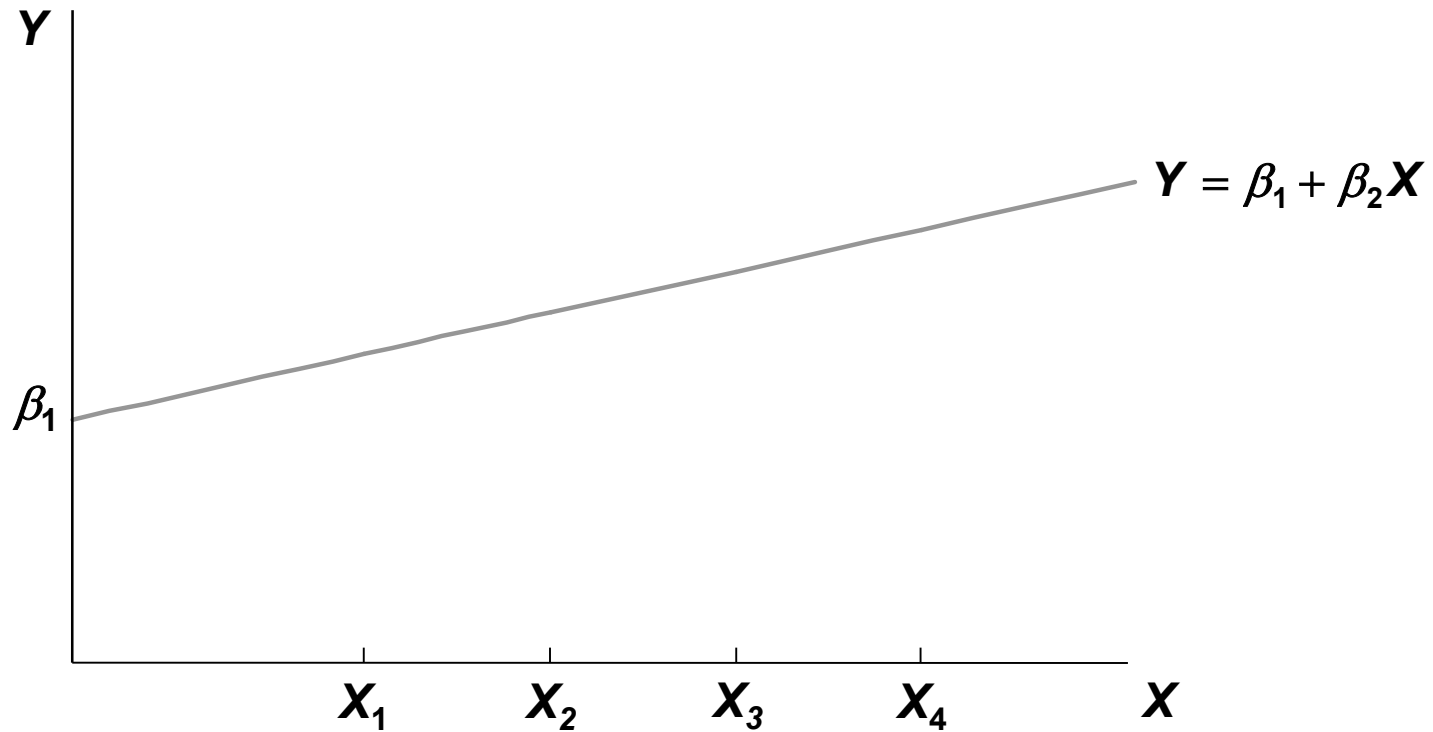
Introduction to SIMPLE REGRESSION MODEL



ឧបមាថាអថេរ Y គឺជាមុខងារលីនេអ៊ែរនៃអថេរ X មួយទៀតដែលមានប៉ារ៉ាម៉ែត្រមិនស្គាល់ b_1 និង b_2 ដែលយើងចង់ប៉ាន់ស្មាន។

Suppose that a variable Y is a linear function of another variable X , with unknown parameters β_1 and β_2 that we wish to estimate.

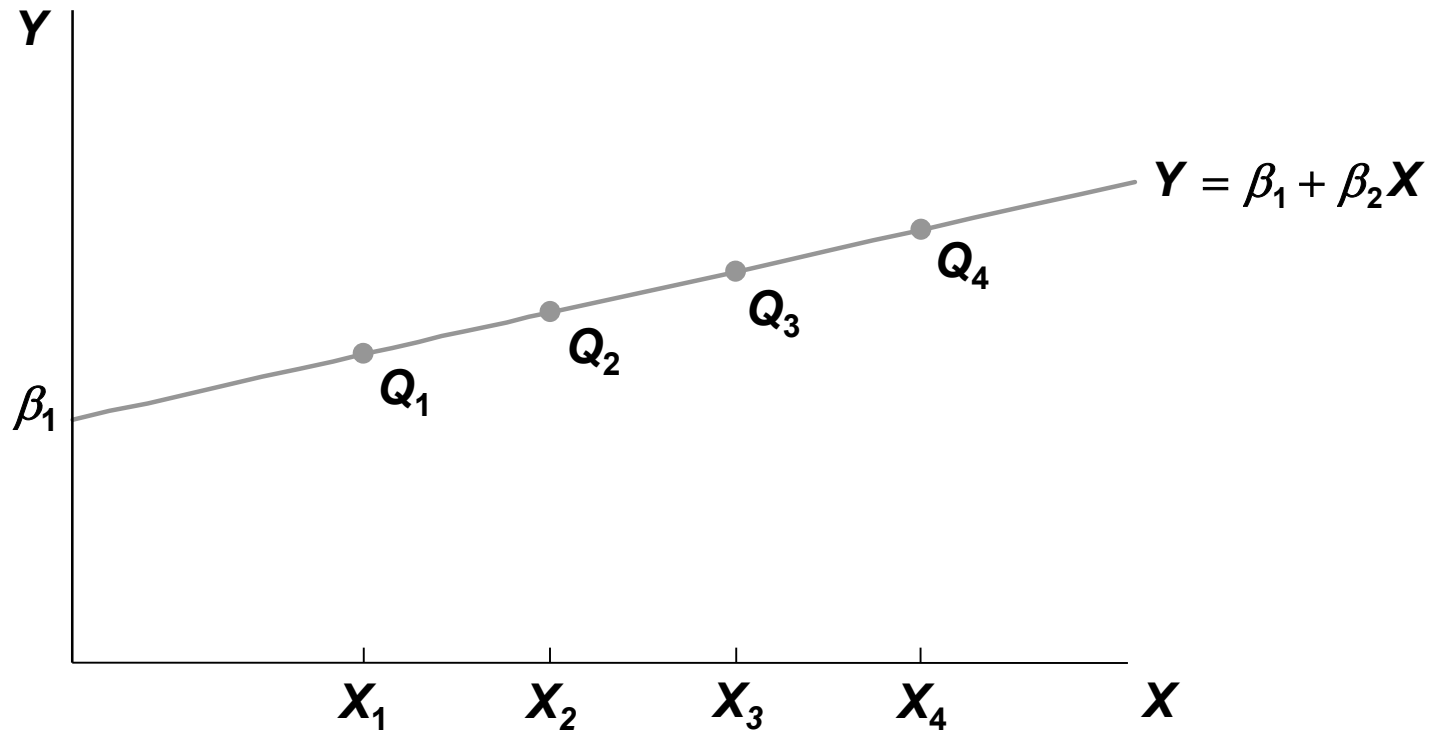
SIMPLE REGRESSION MODEL



ឧបមាថាយើងមានគំរូនៃការសង្កេត ៤ ជាមួយតម្លៃ X ដូចដែលបានបង្ហាញ។

Suppose that we have a sample of 4 observations with X values as shown.

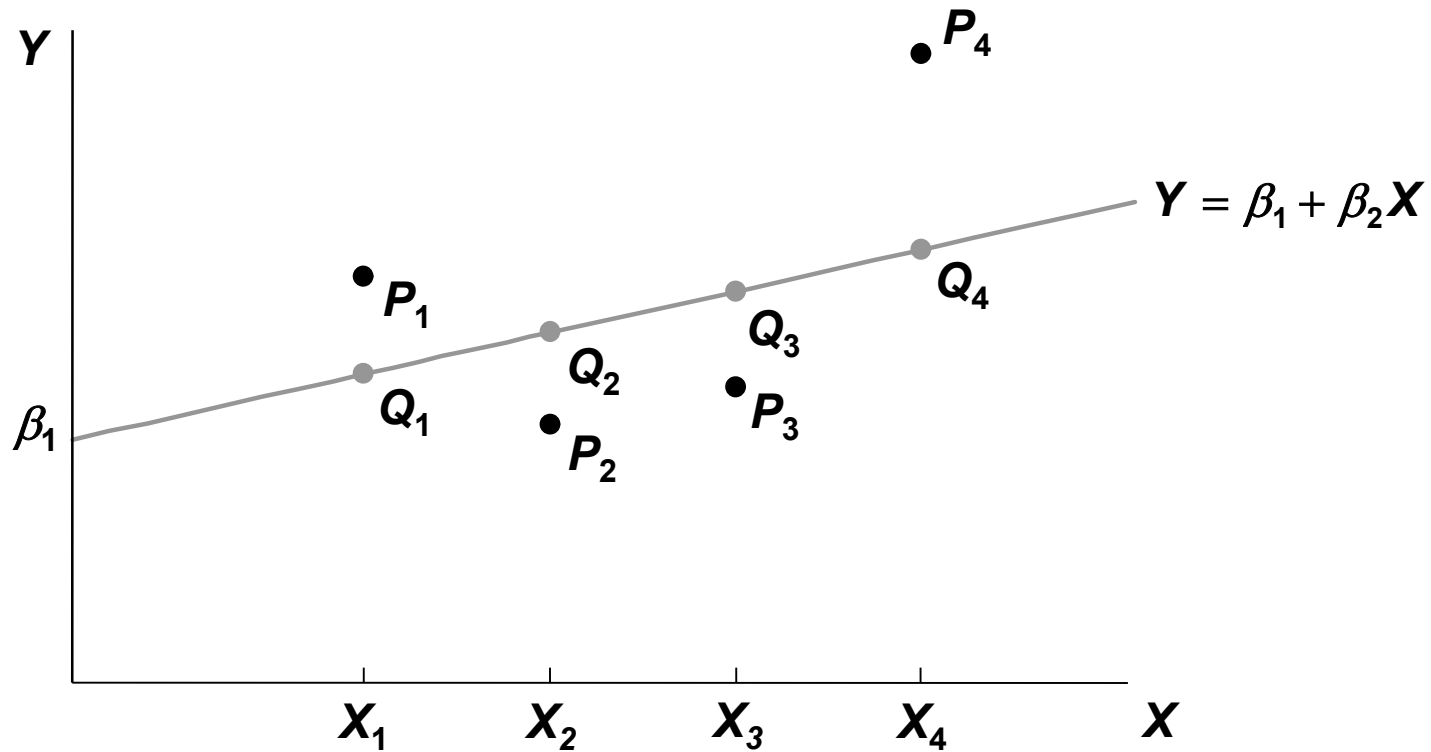
SIMPLE REGRESSION MODEL



ប្រសិនបើទំនាក់ទំនងពិតជាពិតប្រាកដការសង្កេតនឹងស្ថិតនៅលើបន្ទាត់ត្រង់ហើយ យើងនឹងមិនមានបញ្ហាក្នុងការទទួលបានការប៉ាន់ស្មានត្រឹមត្រូវនៃ b_1 និង b_2 ទេ ។

If the relationship were an exact one, the observations would lie on a straight line and we would have no trouble obtaining accurate estimates of β_1 and β_2 .

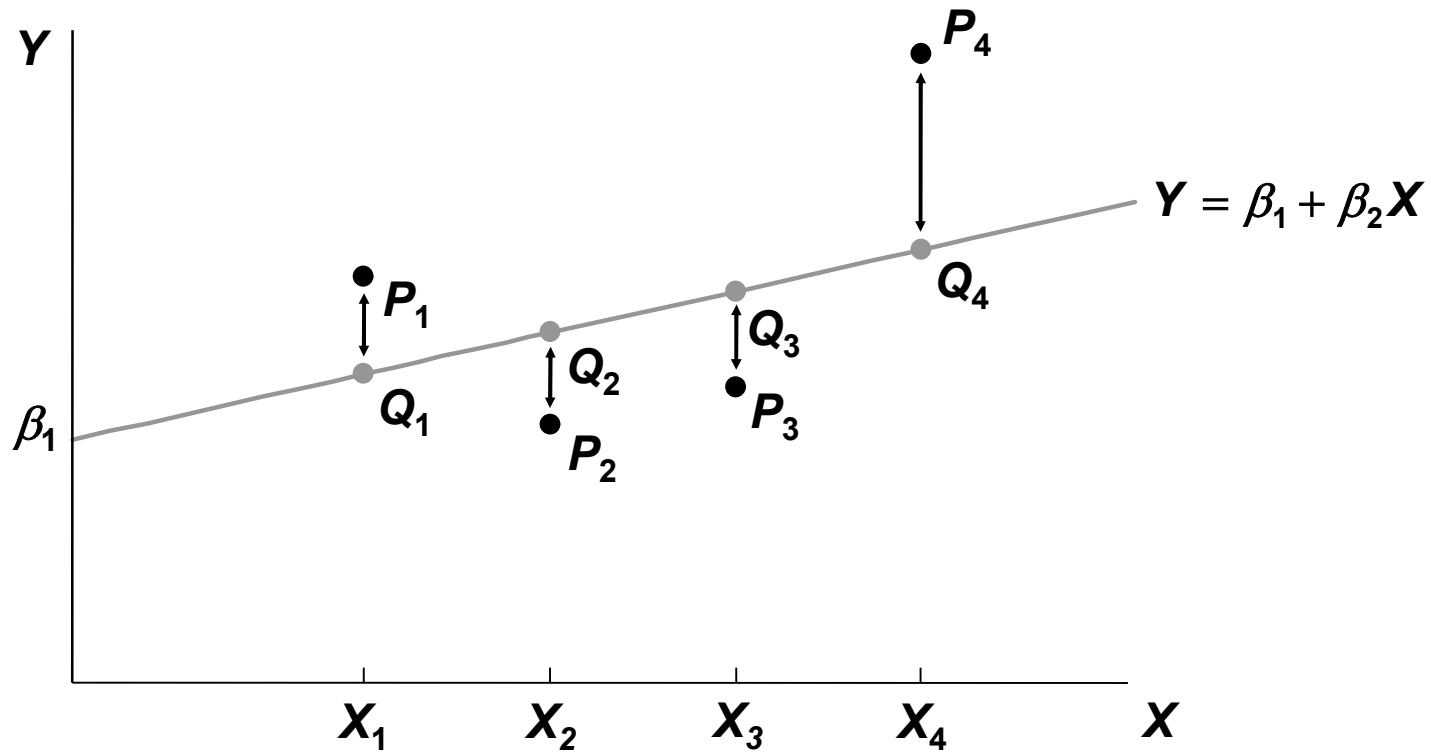
SIMPLE REGRESSION MODEL



នៅក្នុងការអនុវត្តជាក់ស្តែងទំនាក់ទំនងសេដ្ឋកិច្ចភាគច្រើនមិនមានលក្ខណៈពិតប្រាកដទេ ហើយតម្លៃជាក់ស្តែងរបស់ Y គឺខុសគ្នាពីអ្វីដែលត្រូវនឹងបន្ទាត់ត្រង់។

In practice, most economic relationships are not exact and the actual values of Y are different from those corresponding to the straight line.

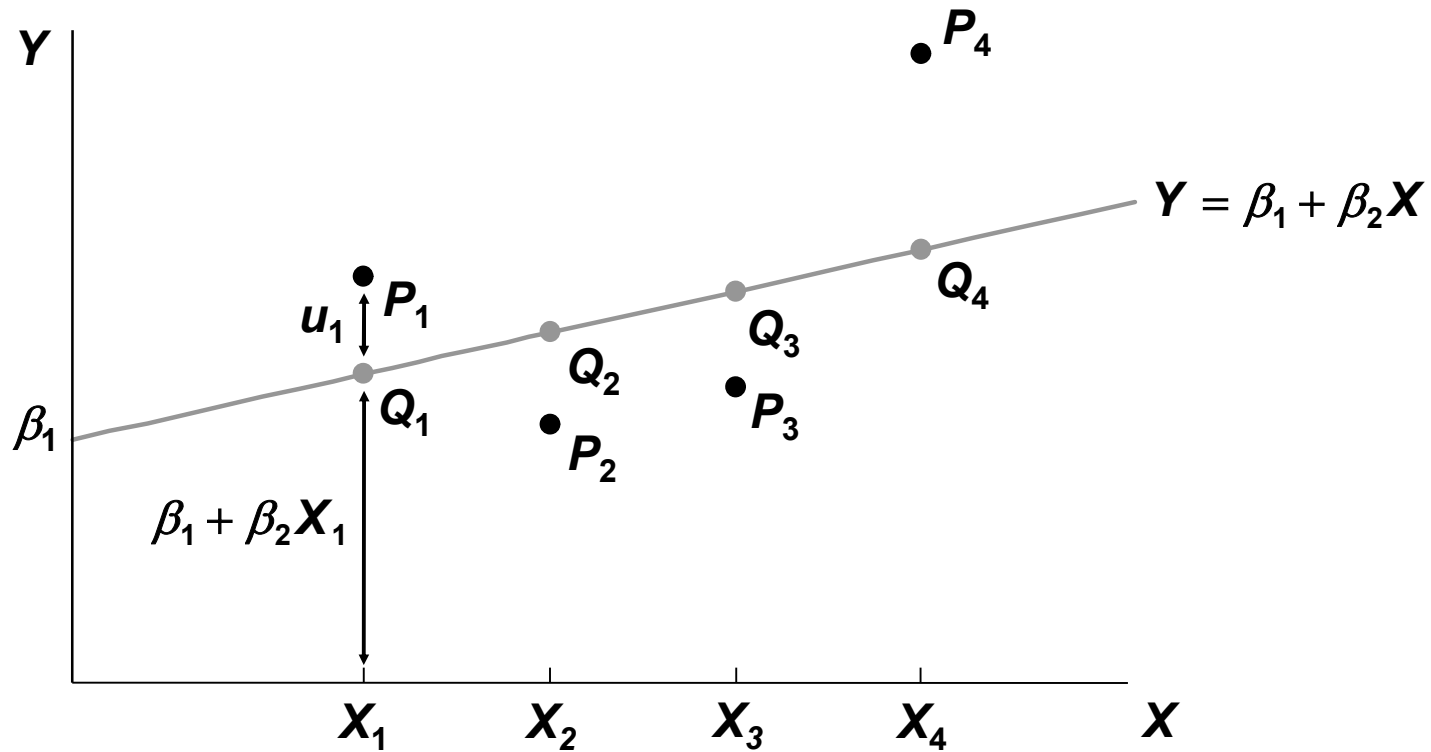
SIMPLE REGRESSION MODEL



ដើម្បីអនុញ្ញាតឱ្យមានភាពខុសគ្នាបែបនេះយើងនឹងសរសេរគំរូដូចជា $Y = b_1 + b_2X + u$ ដែល u គឺជាអថេរមិនអាចវាស់បាន ឬពាក្យរំខាន។

To allow for such divergences, we will write the model as $Y = \beta_1 + \beta_2X + u$, where u is a disturbance term.

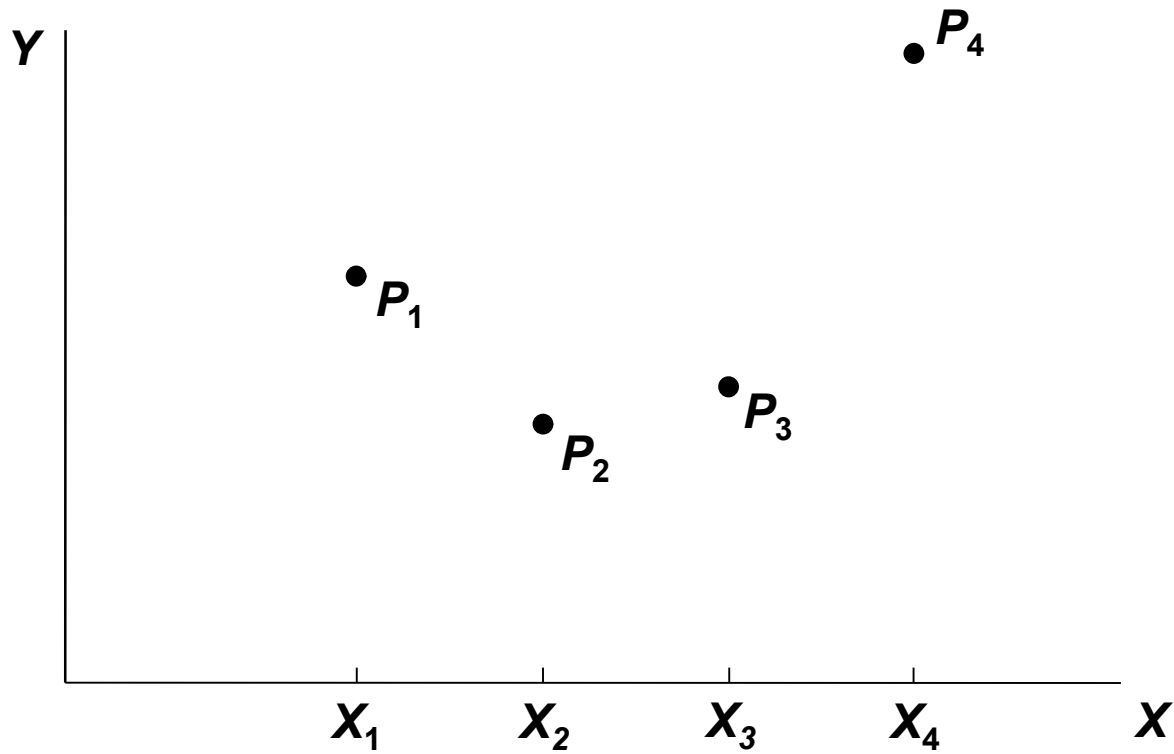
SIMPLE REGRESSION MODEL



ជួរច្នៃតម្លៃនីមួយៗនៃ Y មានសមាសធាតុមិនទាក់ទងគ្នា, $b_1 + b_2 X$ និងសមាសធាតុចៃដន្យមួយ u ។ ការសង្កេតដំបូងត្រូវបានរលាយទៅជាសមាសធាតុទាំងពីរនេះ។

Each value of Y thus has a nonrandom component, $\beta_1 + \beta_2 X$, and a random component, u . The first observation has been decomposed into these two components.

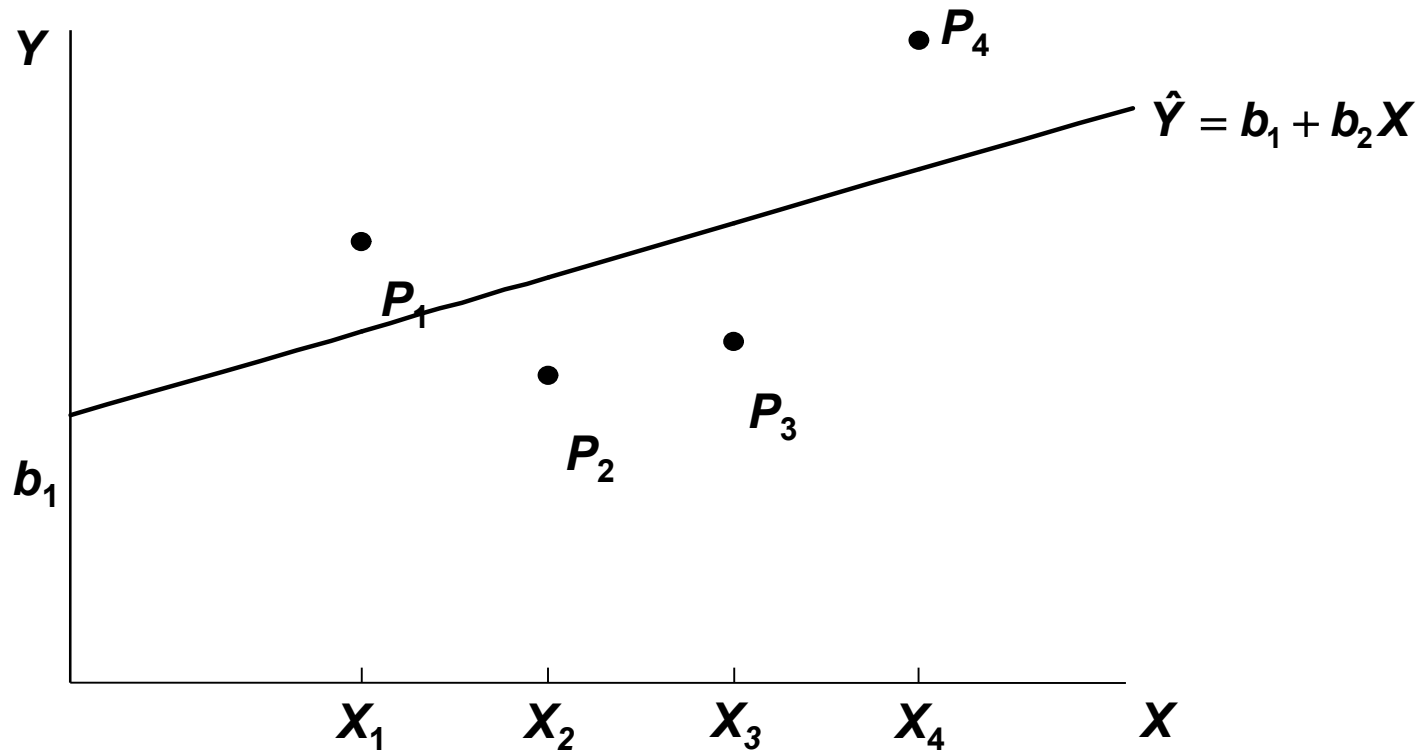
SIMPLE REGRESSION MODEL



នៅក្នុងការអនុវត្តជាក់ស្តែងយើងអាចមើលឃើញតែចំណុច P ប៉ុណ្ណោះ។

In practice we can see only the P points.

SIMPLE REGRESSION MODEL

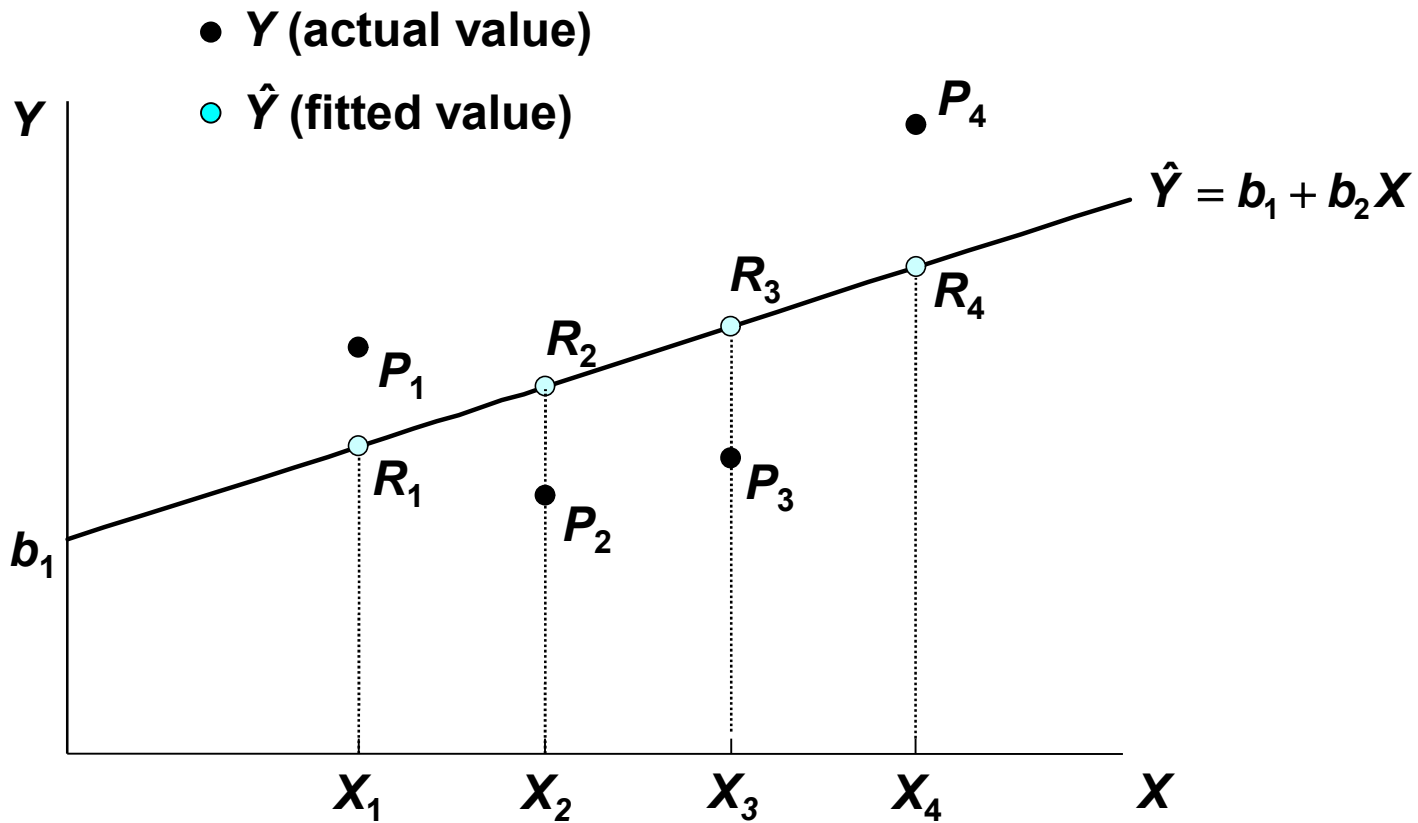


ជាក់ស្តែងយើងអាចប្រើចំណុច P ដើម្បីគូរបន្ទាត់ដែលជាការប៉ាន់ស្មានទៅនឹងបន្ទាត់

$\hat{Y} = b_1 + b_2X$ ។ ប្រសិនបើយើងសរសេរបន្ទាត់ $Y = b_1 + b_2X$, b_1 គឺជាការប៉ាន់ស្មាននៃ β_1 និង b_2 គឺជាការប៉ាន់ស្មាននៃ β_2 ។

Obviously, we can use the P points to draw a line which is an approximation to the line $Y = \beta_1 + \beta_2X$. If we write this line $Y = b_1 + b_2X$, b_1 is an estimate of β_1 and b_2 is an estimate of β_2 .

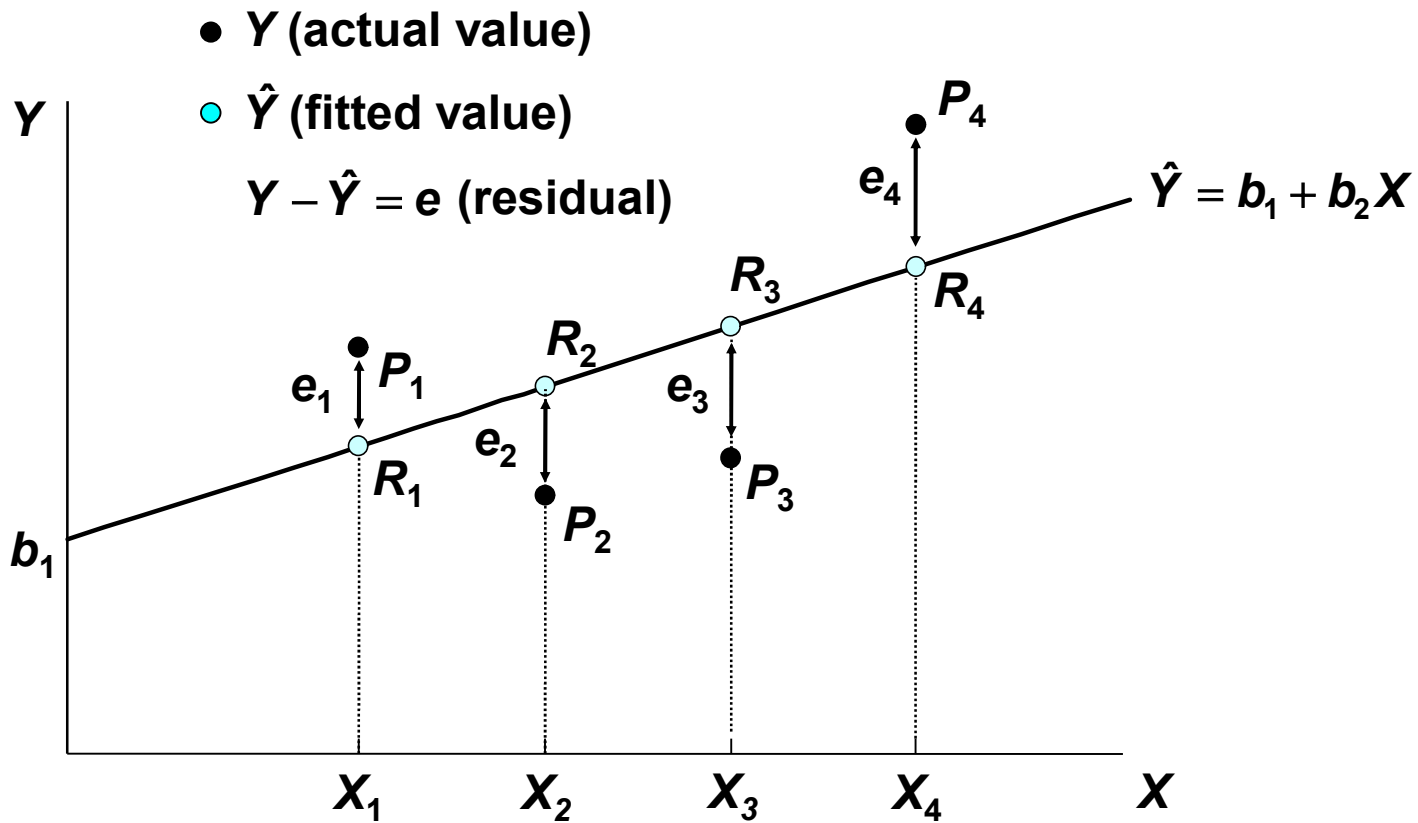
SIMPLE REGRESSION MODEL



បន្ទាត់ត្រូវបានគេហៅថាម៉ូដែលសមនឹងតម្លៃរបស់ Y ដែលបានព្យាករណ៍ដោយវាត្រូវបានគេហៅថាតម្លៃប៉ាន់ស្មាននៃ Y ។ ពួកវាត្រូវបានផ្តល់ដោយកំពស់នៃចំនុច R ។

The line is called the fitted model and the values of Y predicted by it are called the fitted values of Y . They are given by the heights of the R points.

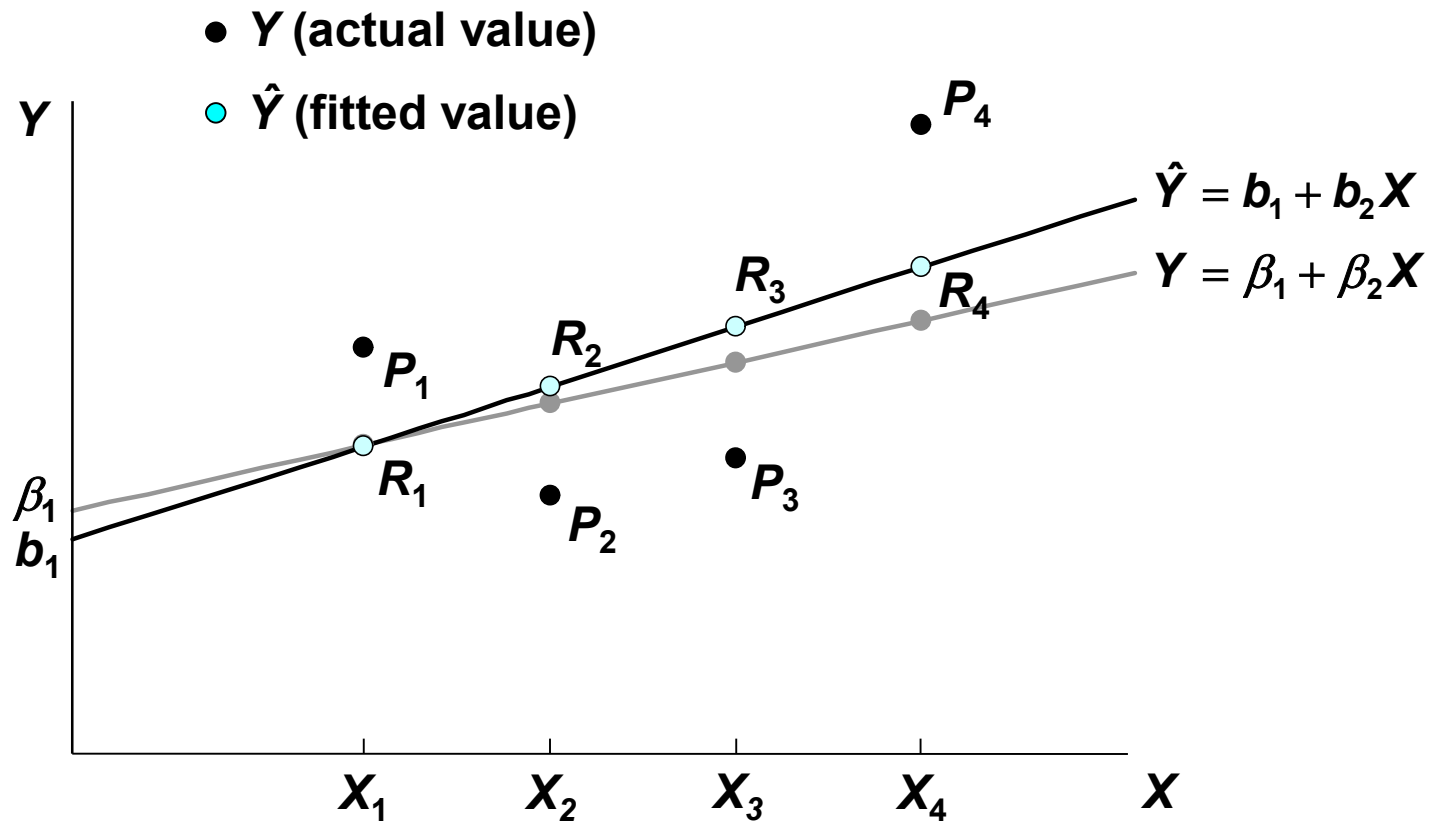
SIMPLE REGRESSION MODEL



ភាពខុសគ្នារវាងតម្លៃជាក់ស្តែងនិងប៉ាន់ស្មាននៃ Y ត្រូវបានគេស្គាល់ថាជាសំណល់។

The discrepancies between the actual and fitted values of Y are known as the residuals.

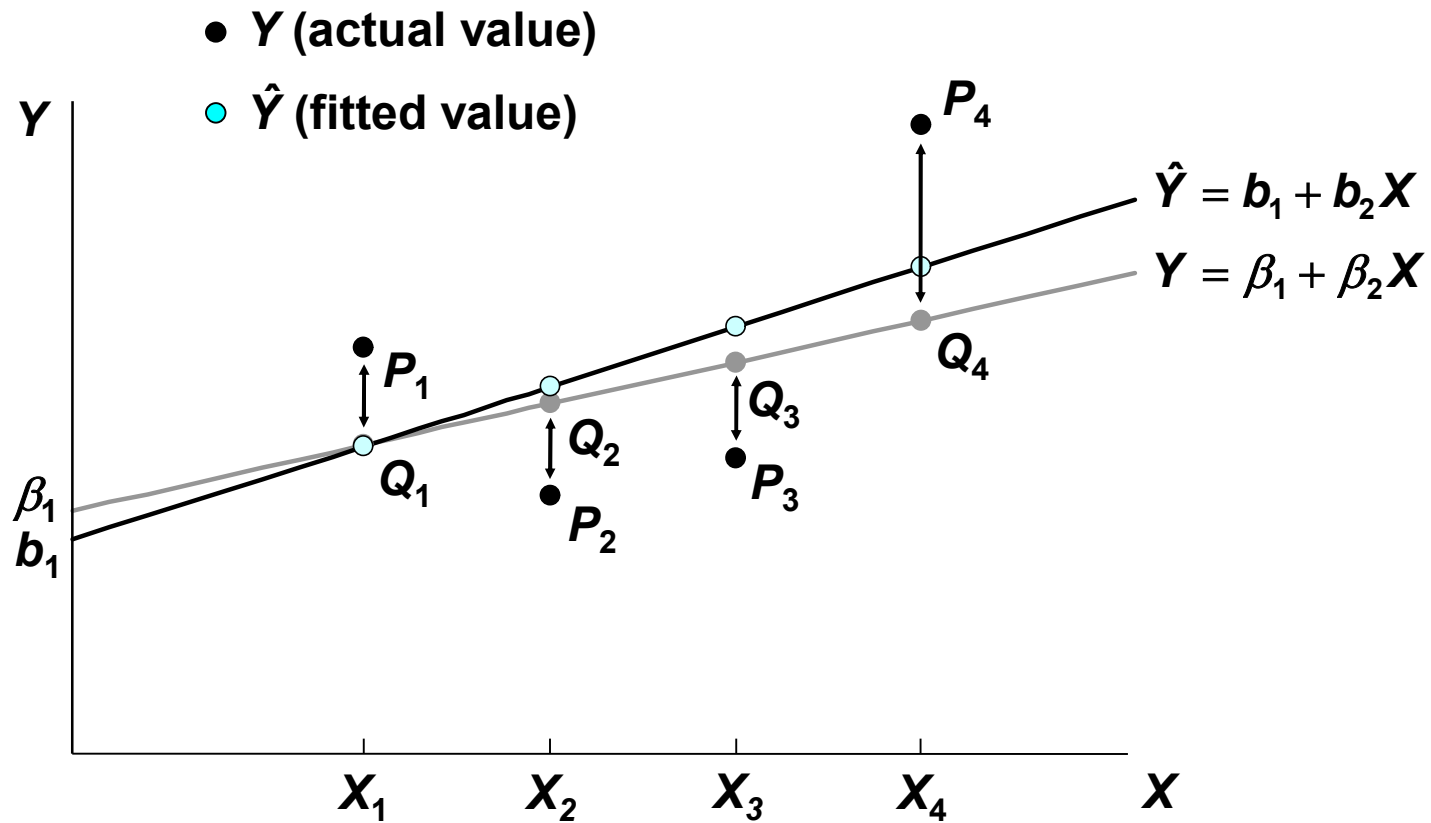
SIMPLE REGRESSION MODEL



ចំណាំថាតម្លៃនៃសំណល់មិនដូចគ្នានឹងតម្លៃនៃពាក្យរំខានទេ។ ដ្យាក្រាមឥឡូវនេះបង្ហាញពីទំនាក់ទំនងដែលមិនស្គាល់ពិតប្រាកដក៏ដូចជាខ្សែដែលសមស្រប។

Note that the values of the residuals are not the same as the values of the disturbance term. The diagram now shows the true unknown relationship as well as the fitted line.

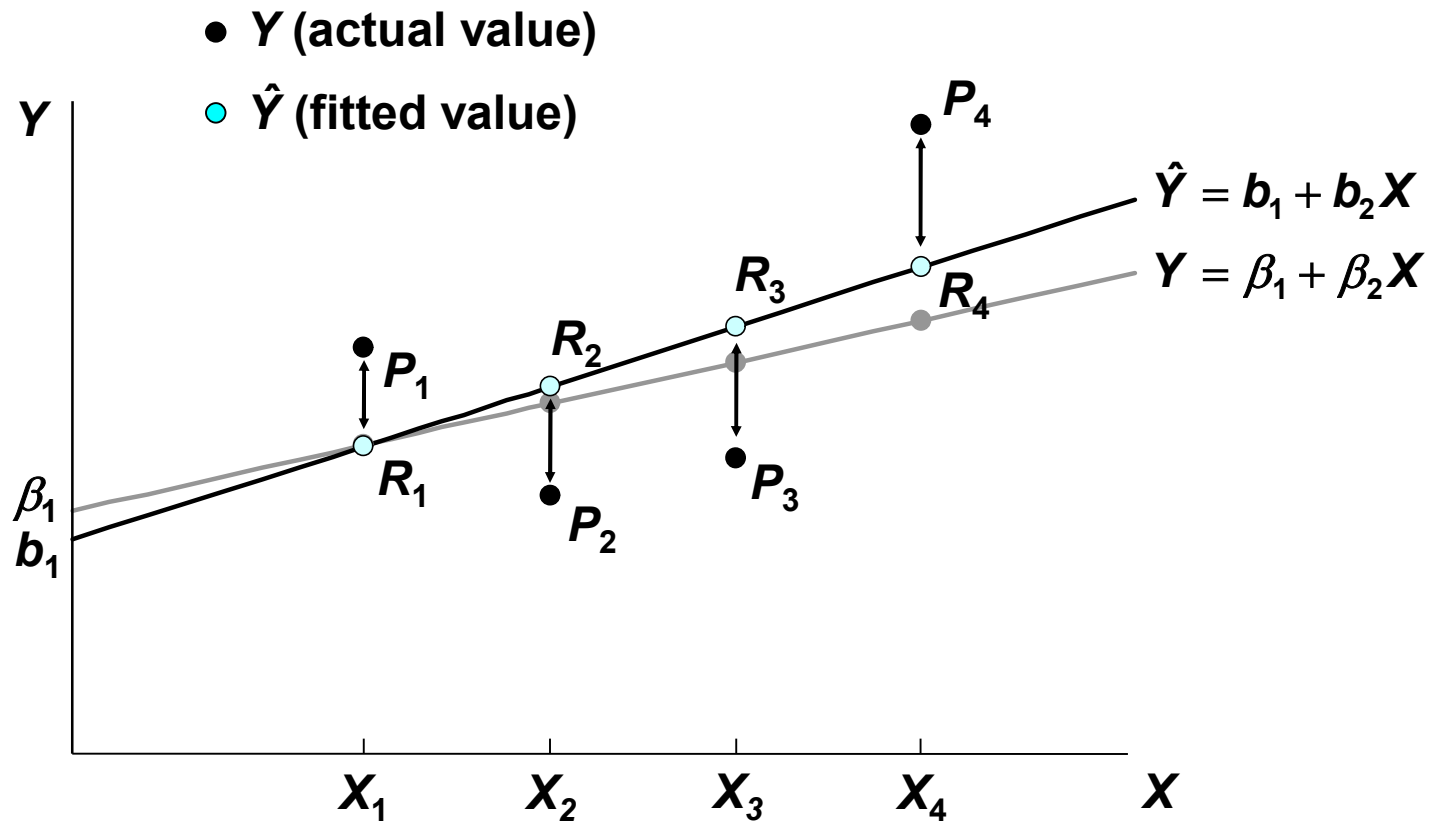
SIMPLE REGRESSION MODEL



ពាក្យរំខាននៅក្នុងការសង្កេតនីមួយៗគឺទទួលខុសត្រូវចំពោះភាពផ្សេងគ្នារវាងសមាសធាតុមិនផ្លូវការនៃទំនាក់ទំនងពិតនិងការសង្កេតជាក់ស្តែង។

The disturbance term in each observation is responsible for the divergence between the nonrandom component of the true relationship and the actual observation.

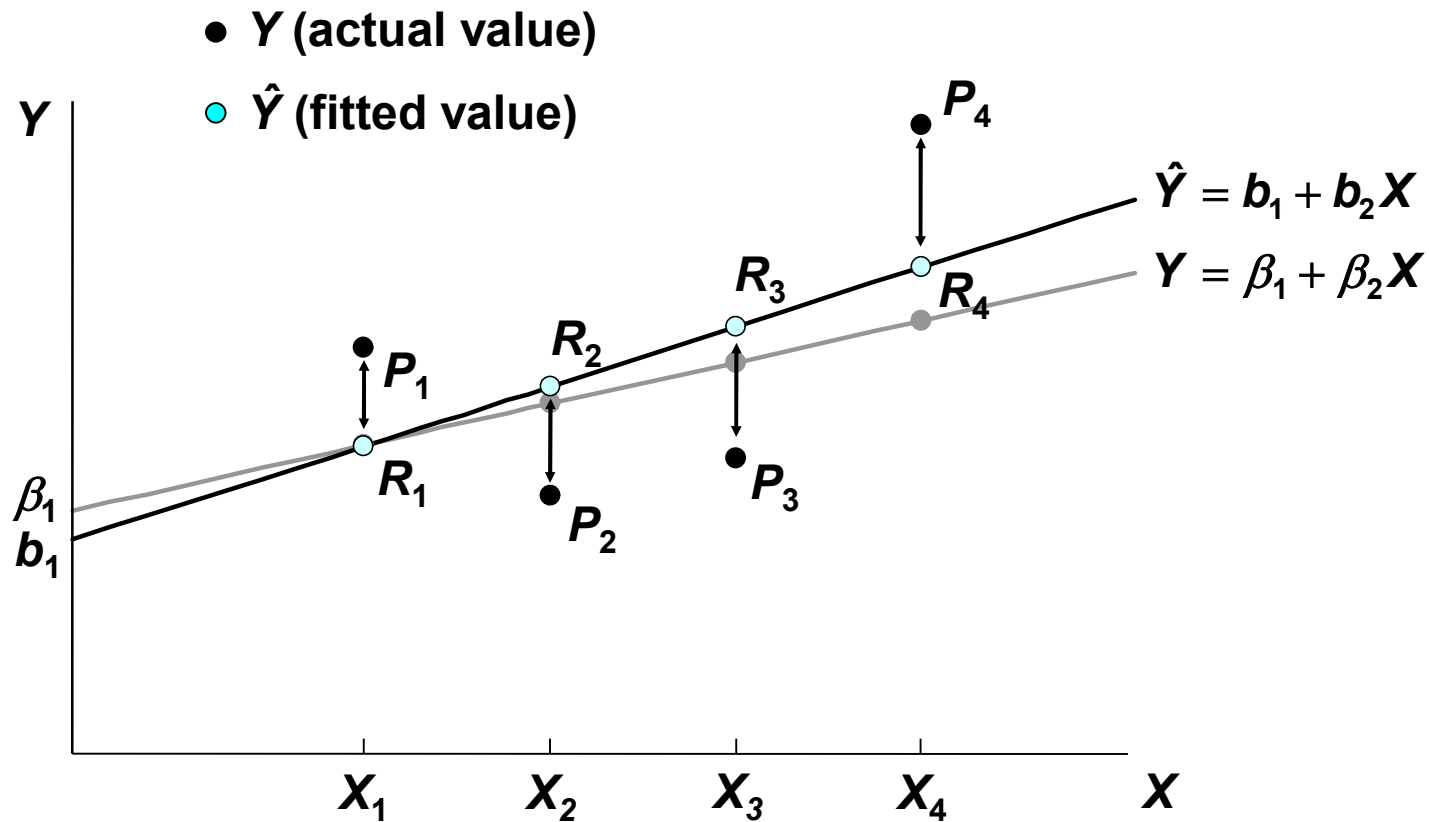
SIMPLE REGRESSION MODEL



សំណល់គឺជាភាពខុសគ្នារវាងតម្លៃជាក់ស្តែងនិងតម្លៃសមស្រប។

The residuals are the discrepancies between the actual and the fitted values.

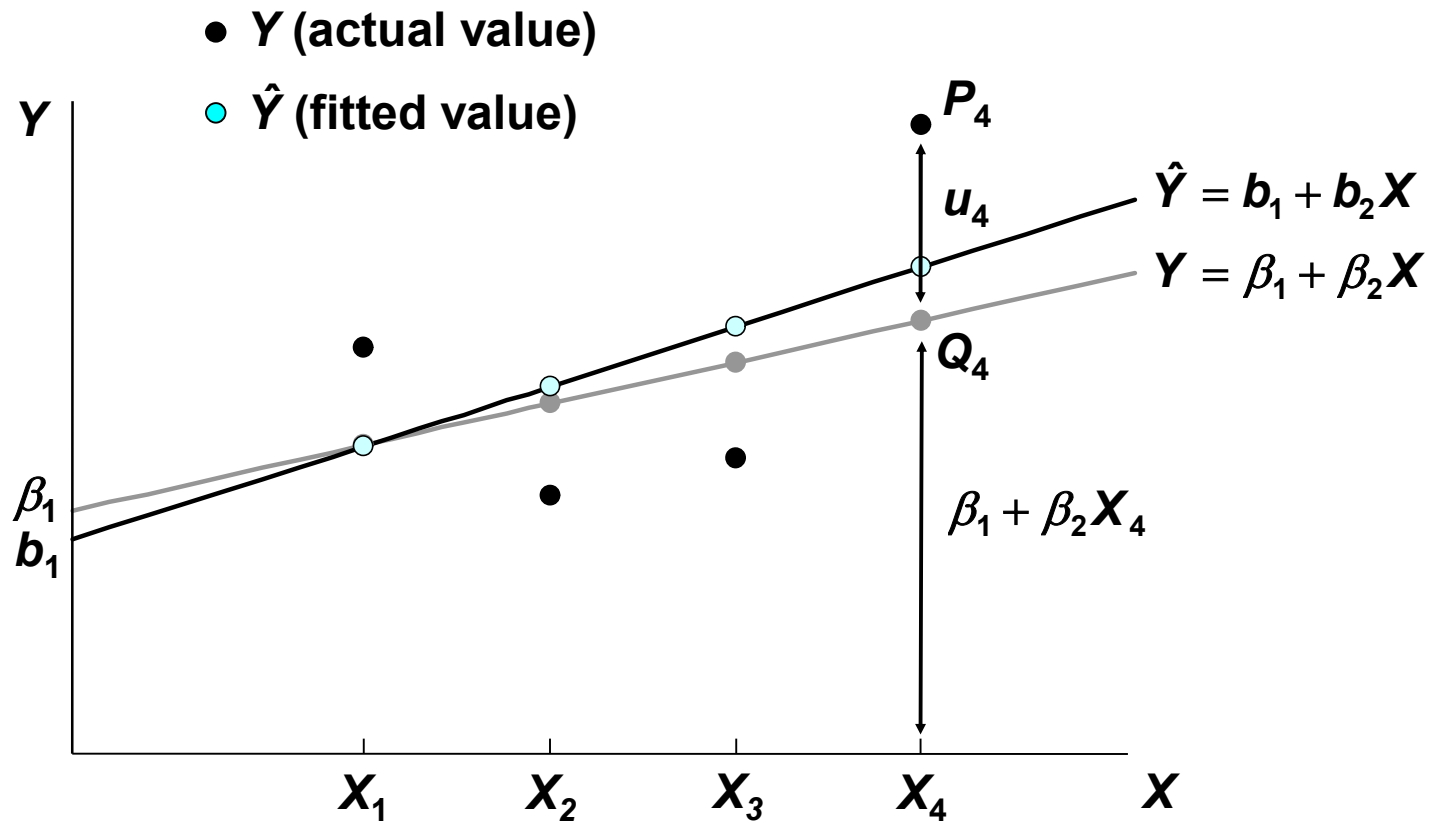
SIMPLE REGRESSION MODEL



ប្រសិនបើតម្លៃប៉ាន់ស្មានមួយគឺល្អ នោះសំណល់និងតម្លៃនៃពាក្យរំខាននឹងមានលក្ខណៈប្រហាក់ប្រហែលគ្នា ប៉ុន្តែពួកគេត្រូវតែទុកឱ្យនៅដាច់ដោយឡែកពីគ្នា។

If the fit is a good one, the residuals and the values of the disturbance term will be similar, but they must be kept apart conceptually.

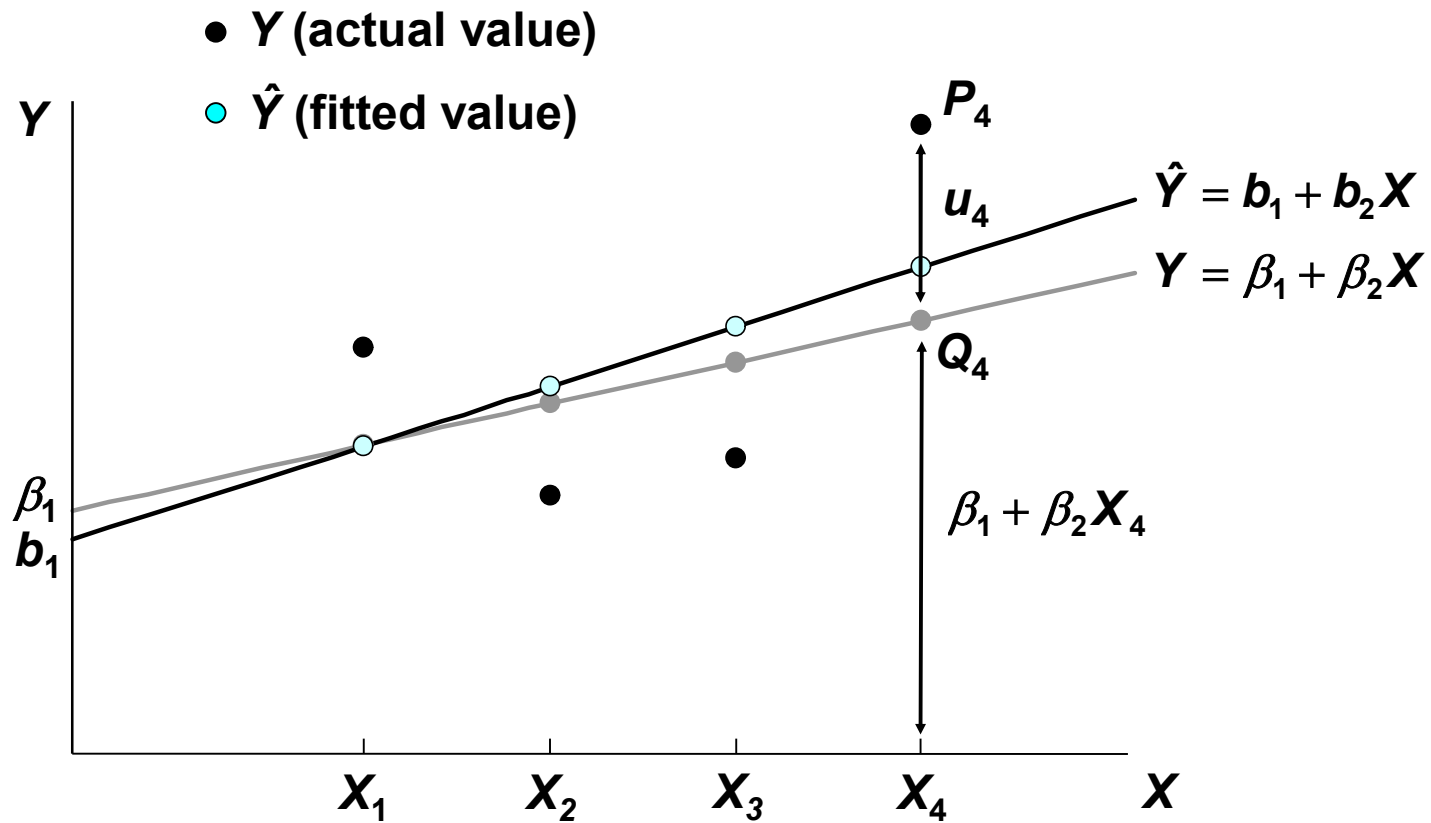
SIMPLE REGRESSION MODEL



បន្ទាត់ទាំងពីរនេះនឹងត្រូវប្រើក្នុងការវិភាគរបស់យើង។ ការបំបែកធាតុនីមួយៗនៃតម្លៃរបស់ Y នឹងត្រូវបានបង្ហាញជាមួយនឹងការសង្កេតទី ៤ ។

Both of these lines will be used in our analysis. Each permits a decomposition of the value of Y . The decompositions will be illustrated with the fourth observation.

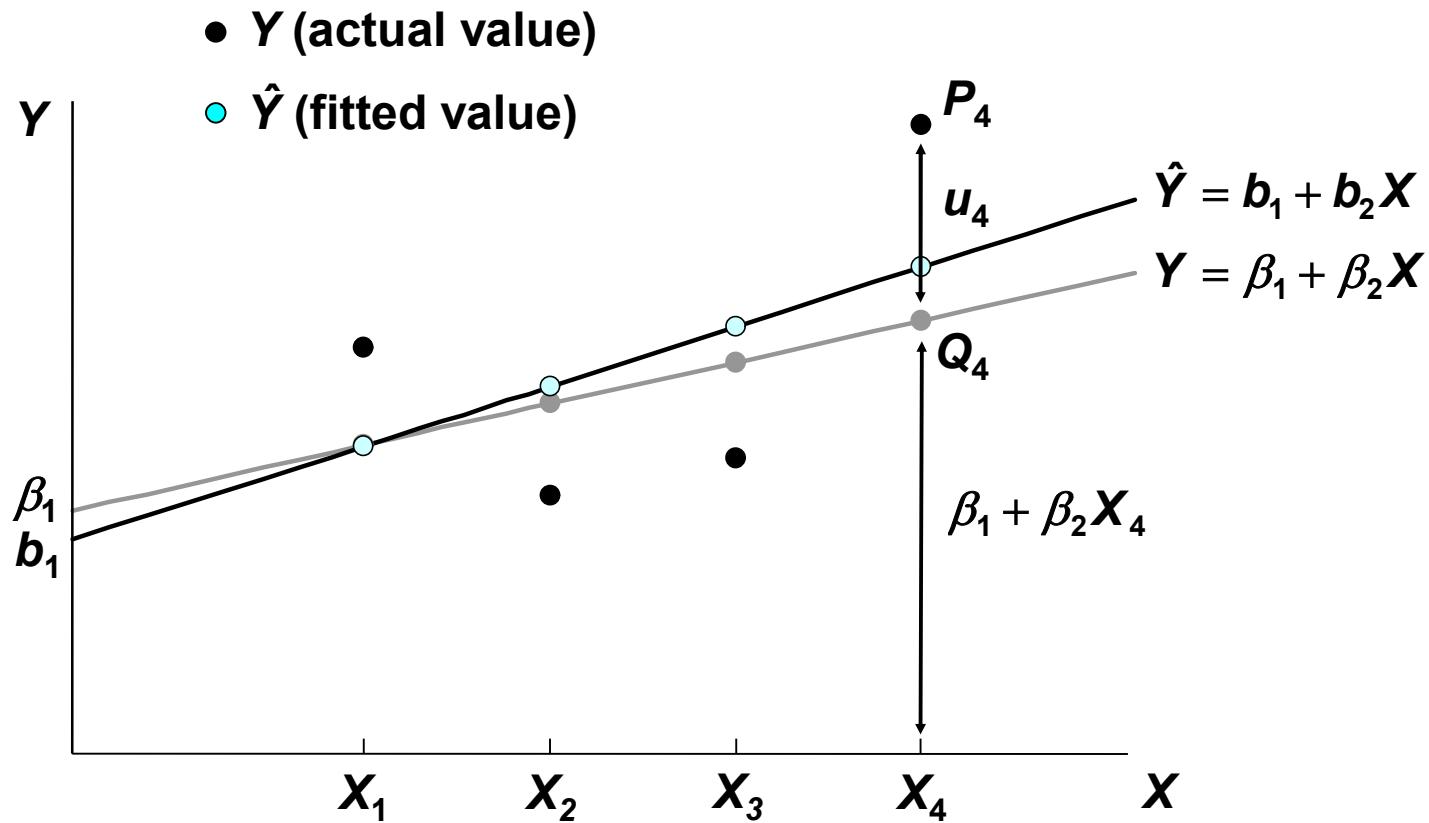
SIMPLE REGRESSION MODEL



ដោយប្រើទ្រឹស្តីទំនាក់ទំនង Y អាចត្រូវបានរលាយចូលទៅក្នុងសមាសធាតុពាក់ព័ន្ធរបស់វាគឺ $b_1 + b_2X$ និងសមាសធាតុចៃដន្យ u ។

Using the theoretical relationship, Y can be decomposed into its nonstochastic component $\beta_1 + \beta_2X$ and its random component u .

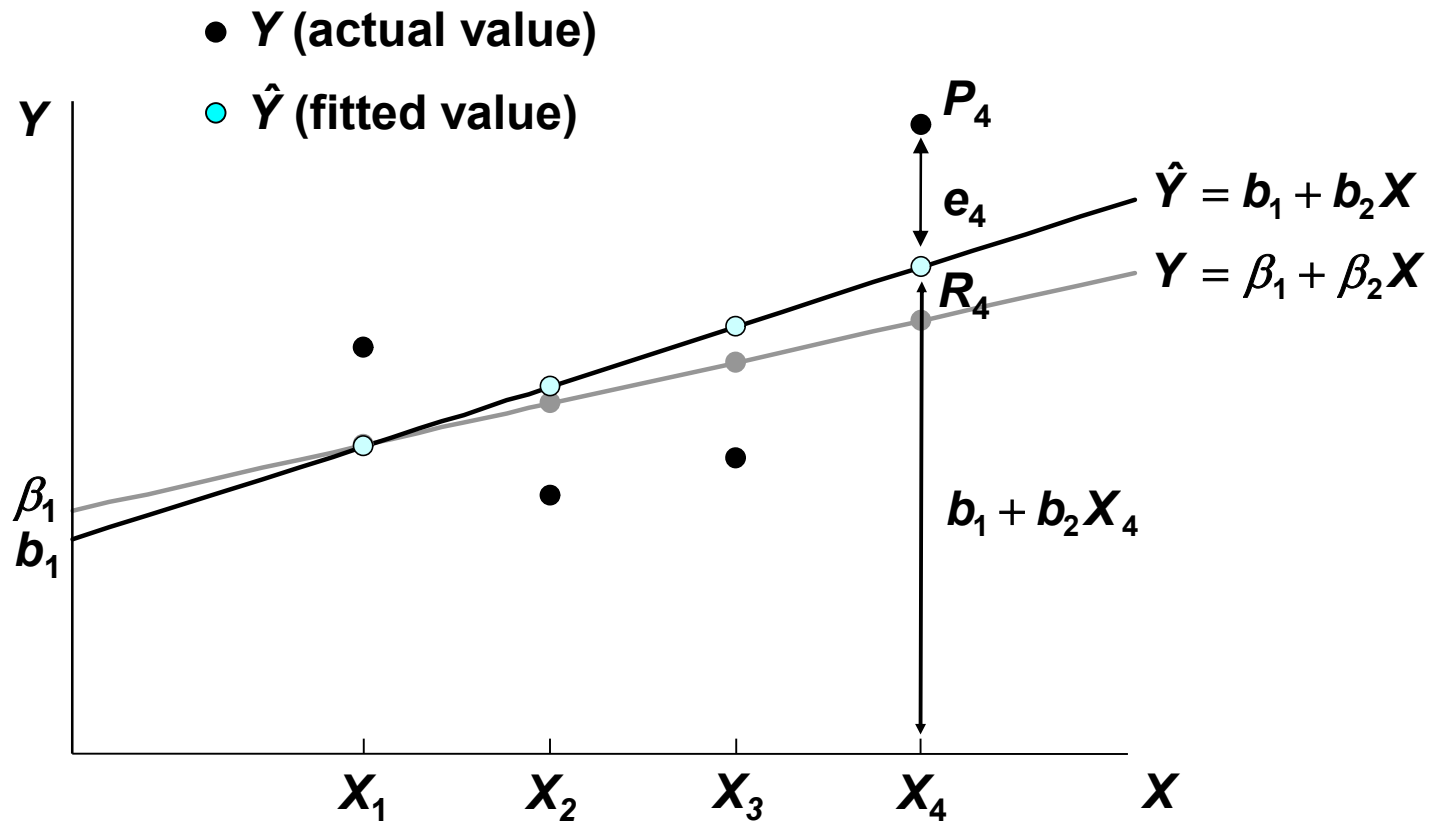
SIMPLE REGRESSION MODEL



នេះគឺជាការបំប្លែងទ្រឹស្តីពីព្រោះយើងមិនដឹងពីតម្លៃនៃ b_1 ឬ b_2 ឬតម្លៃនៃពាក្យរំខាន។ យើងត្រូវប្រើវានៅក្នុងការវិភាគរបស់យើងអំពីលក្ខណៈសម្បត្តិនៃមេគុណតំរែតំរង់។

This is a theoretical decomposition because we do not know the values of β_1 or β_2 , or the values of the disturbance term. We shall use it in our analysis of the properties of the regression coefficients.

SIMPLE REGRESSION MODEL



ការបំបែកផ្សេងទៀតគឺយោងទៅលើបន្ទាត់ដែលត្រូវនឹងតម្លៃប៉ាន់ស្មាន។ តាមការសង្កេតនីមួយៗ តម្លៃជាក់ស្តែងរបស់ Y ស្មើនឹងតម្លៃប៉ាន់ស្មានបូកនឹងសំណល់។ នេះគឺជាការបំផ្លែងប្រតិបត្តិការ ដែលយើងនឹងប្រើសម្រាប់គោលបំណងអនុវត្តជាក់ស្តែង។

The other decomposition is with reference to the fitted line. In each observation, the actual value of Y is equal to the fitted value plus the residual. This is an operational decomposition which we will use for practical purposes.

SIMPLE REGRESSION MODEL

Least squares criterion:

Minimize RSS (residual sum of squares), where

$$RSS = \sum_{i=1}^n e_i^2 = e_1^2 + \dots + e_n^2$$

ដើម្បីចាប់ផ្តើម យើងនឹងគូសខ្សែបន្ទាត់ប៉ាន់ស្មានដើម្បីកាត់បន្ថយផលបូកការេនៃសំណល់ដែលនៅសល់ RSS ។ នេះត្រូវបានពិពណ៌នាថាជាលក្ខណៈវិនិច្ឆ័យការេតូចបំផុត ឬតម្លៃអប្បបរមា។

To begin with, we will draw the fitted line so as to minimize the sum of the squares of the residuals, RSS . This is described as the least squares criterion.

SIMPLE REGRESSION MODEL

Least squares criterion:

Minimize RSS (residual sum of squares), where

$$RSS = \sum_{i=1}^n e_i^2 = e_1^2 + \dots + e_n^2$$

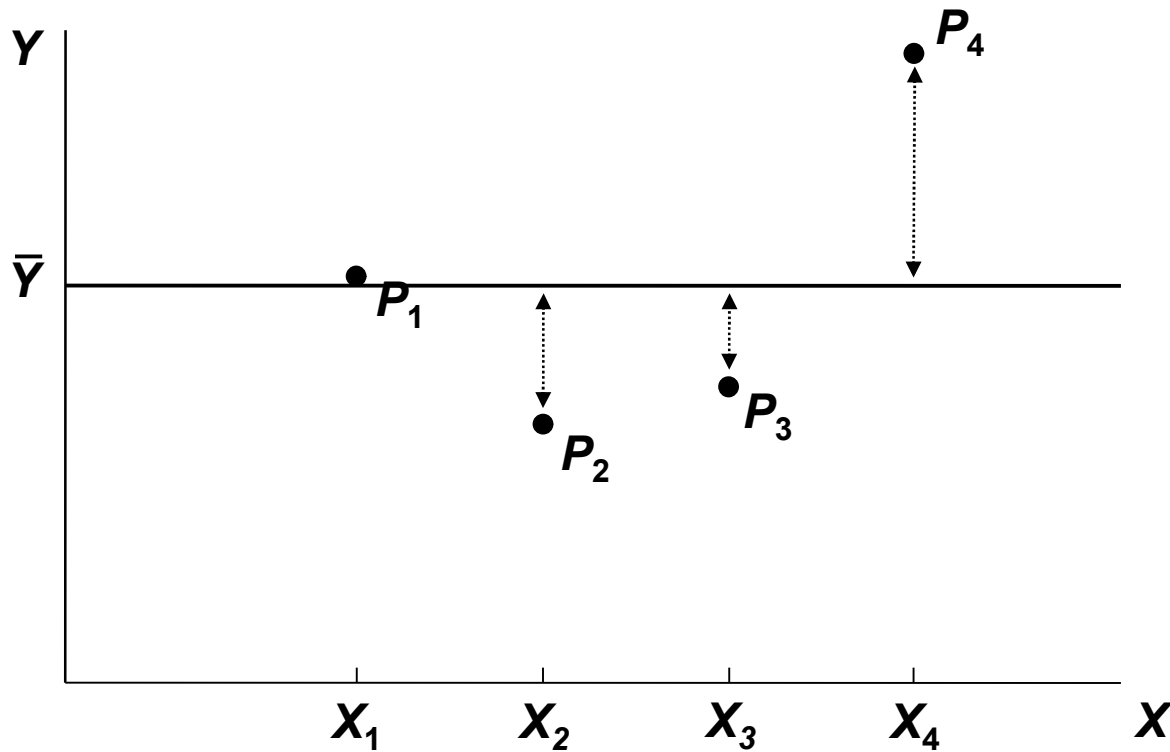
Why not minimize

$$\sum_{i=1}^n e_i = e_1 + \dots + e_n$$

ហេតុអ្វីបានជាគណនាការ៉េនៃសំណល់? ហេតុអ្វីមិនគ្រាន់តែធ្វើអប្បបរមាចំនួនសរុបនៃសំណល់?

Why the squares of the residuals? Why not just minimize the sum of the residuals?

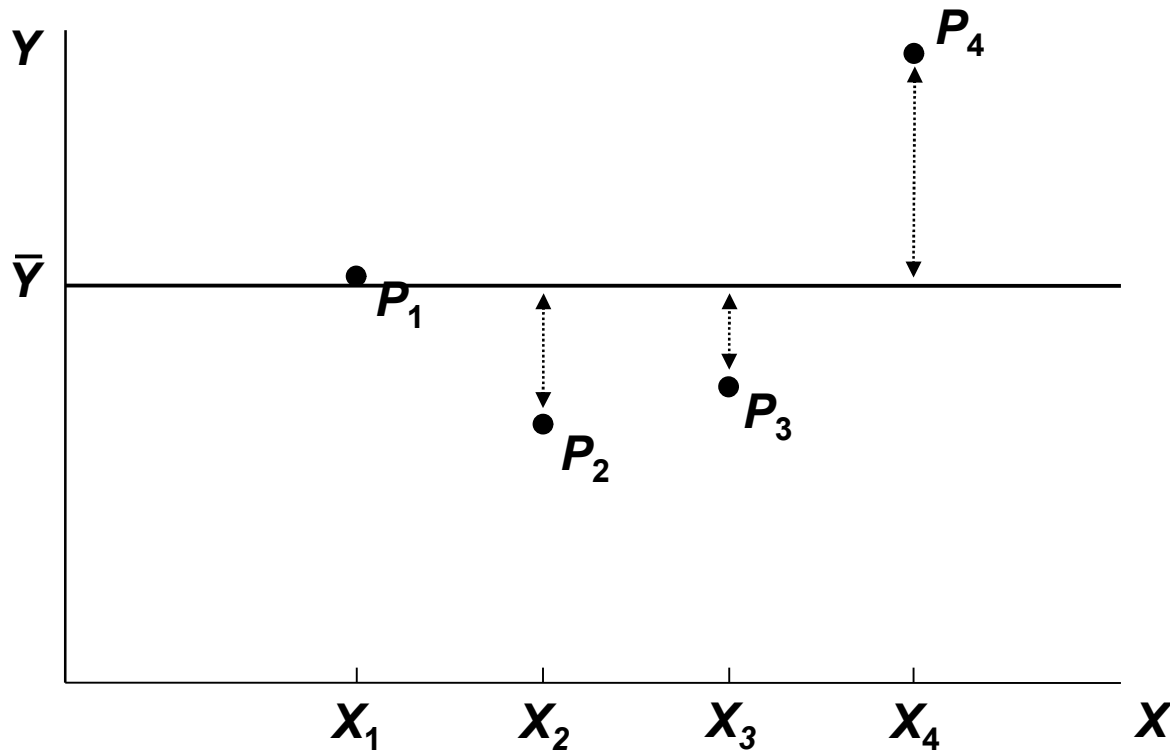
SIMPLE REGRESSION MODEL



ចម្លើយគឺថាអ្នកនឹងទទួលបាននូវប៉ាន់ស្មានដ៏ល្អឥតខ្ចោះមួយដោយគូរបន្ទាត់ផ្អែកតាមរយៈតម្លៃមធ្យមរបស់ Y ។ ផលបូកនៃសំណល់នឹងមានតម្លៃសូន្យ។

The answer is that you would get an apparently perfect fit by drawing a horizontal line through the mean value of Y . The sum of the residuals would be zero.

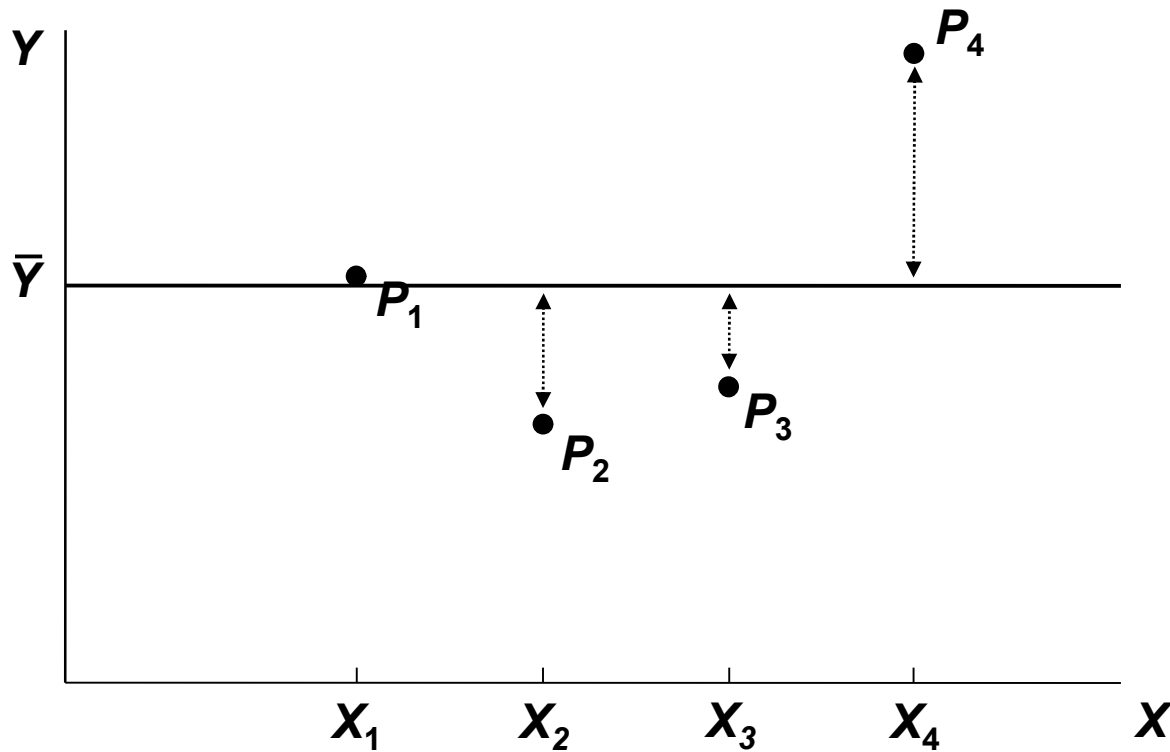
SIMPLE REGRESSION MODEL



ហើយដើម្បីការពារសំណល់អវិជ្ជមានពីការលុបចោលនូវចំនុចវិជ្ជមាន អ្នកត្រូវតែប្រើវិធីមួយគឺធ្វើការ៉េនៃសំណល់។

You must prevent negative residuals from cancelling positive ones, and one way to do this is to use the squares of the residuals.

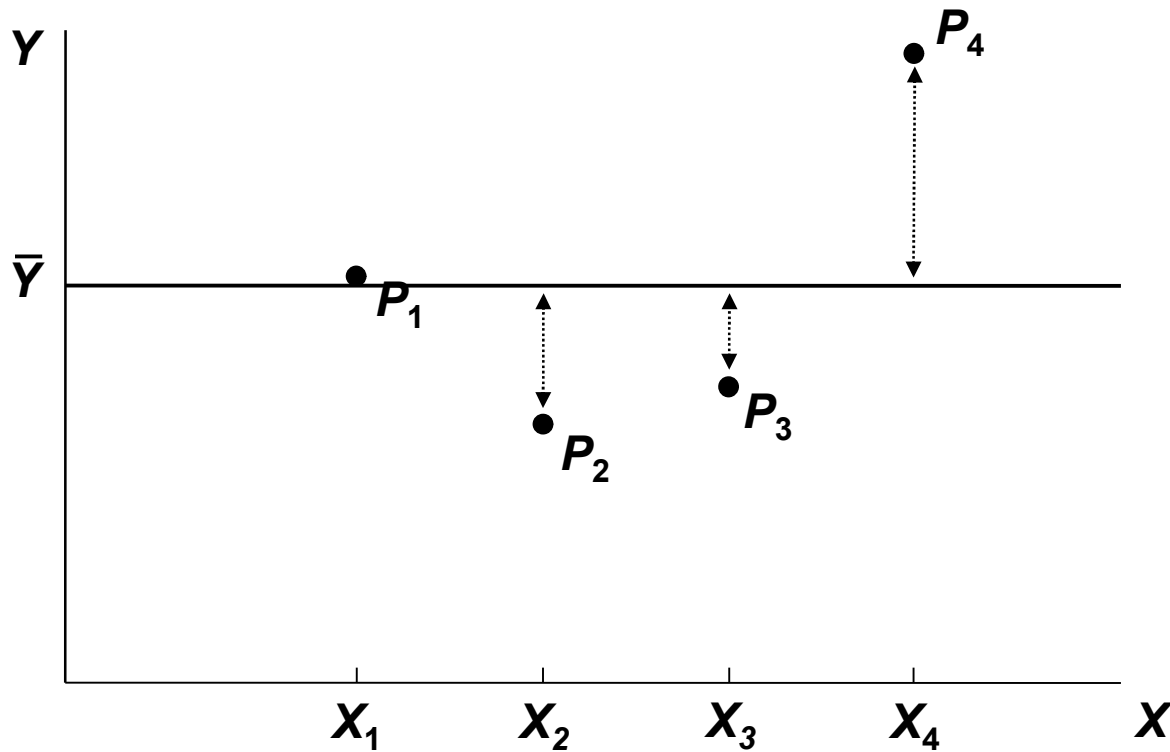
SIMPLE REGRESSION MODEL



ជាការពិតមានវិធីផ្សេងទៀតក្នុងការដោះស្រាយបញ្ហា។ លក្ខណៈវិនិច្ឆ័យការវែតូចបំផុតមានចំណុចទាក់ទាញដែលអ្នកប៉ាន់ស្មានបានមកជាមួយ វាមានលក្ខណៈសម្បត្តិគួរឱ្យចង់បានដែលផ្តល់នូវលក្ខខណ្ឌជាក់លាក់។

Of course there are other ways of dealing with the problem. The least squares criterion has the attraction that the estimators derived with it have desirable properties, provided that certain conditions are satisfied.

SIMPLE REGRESSION MODEL



លំដាប់បន្ទាប់បង្ហាញពីរបៀបដែលលក្ខណៈវិនិច្ឆ័យការរើតូចបំផុតត្រូវបានប្រើដើម្បីគណនាមេគុណ
នៃបន្ទាត់ដែលត្រូវបានស្មាន។

The next sequence shows how the least squares criterion is used to calculate the coefficients of the fitted line.

SIMPLE REGRESSION MODEL

- ភាពចៃដន្យតែងកើតមានជានិច្ច ហេតុនេះយើងត្រូវរៀនសូត្រយល់ពីលទ្ធភាពដែលវាអាចកើតឡើង។

- បើមិនបានស្វែងយល់ពីភាពចៃដន្យ ពេលជួបវា អ្នកតែងតែយល់ថាវាគឺជាព្រហ្មលិខិត។