OXFORD
UNIVERSITY PRESS

Dougherty

Introduction to Econometrics, 5th edition

Chapter 2: Properties of the Regression Coefficients and Hypothesis Testing

A MONTE CARLO EXPERIMENT

True model

$$Y = \beta_1 + \beta_2 X + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X$$

$$\hat{\beta}_2 = \frac{\sum \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum \left(X_i - \overline{X}\right)^2} = \beta_2 + \sum a_t u_i$$

$$a_i = \frac{X_i - \overline{X}}{\sum \left(X_j - \overline{X}\right)^2}$$

In the previous slideshow, we saw that the error term is responsible for the variations of b_2 around its fixed component β_2. We demonstrated mathematically that the expectation of the error term is zero and hence that $\hat{\beta}_2$ is an unbiased estimator of β_2.

A MONTE CARLO EXPERIMENT

True model

$$Y = \beta_1 + \beta_2 X + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X$$

$$\hat{\beta}_2 = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sum (X_i - \overline{X})^2} = \beta_2 + \sum a_t u_i$$

$$a_i = \frac{X_i - \overline{X}}{\sum (X_j - \overline{X})^2}$$

In this slideshow we will investigate the effect of the error term on $\hat{\beta}_2$ directly, using a Monte Carlo experiment (simulation).

A MONTE CARLO EXPERIMENT

Choose model in which Y is determined by X, parameter values, and u

A Monte Carlo experiment is a laboratory-style exercise usually undertaken with the objective of evaluating the properties of regression estimators under controlled conditions.

A MONTE CARLO EXPERIMENT

Choose model in which Y is determined by X, parameter values, and u

We will use one to investigate the behavior of OLS regression coefficients when applied to a simple regression model.

A MONTE CARLO EXPERIMENT

Choose model in which Y is determined by X, parameter values, and u

Choose data for X

Choose parameter values

We will assume that *Y* is determined by a variable *X* and a disturbance term *u*, we will choose the data for *X*, and we will choose values for the parameters.

A MONTE CARLO EXPERIMENT

Choose model in which Y is determined by X, parameter values, and u

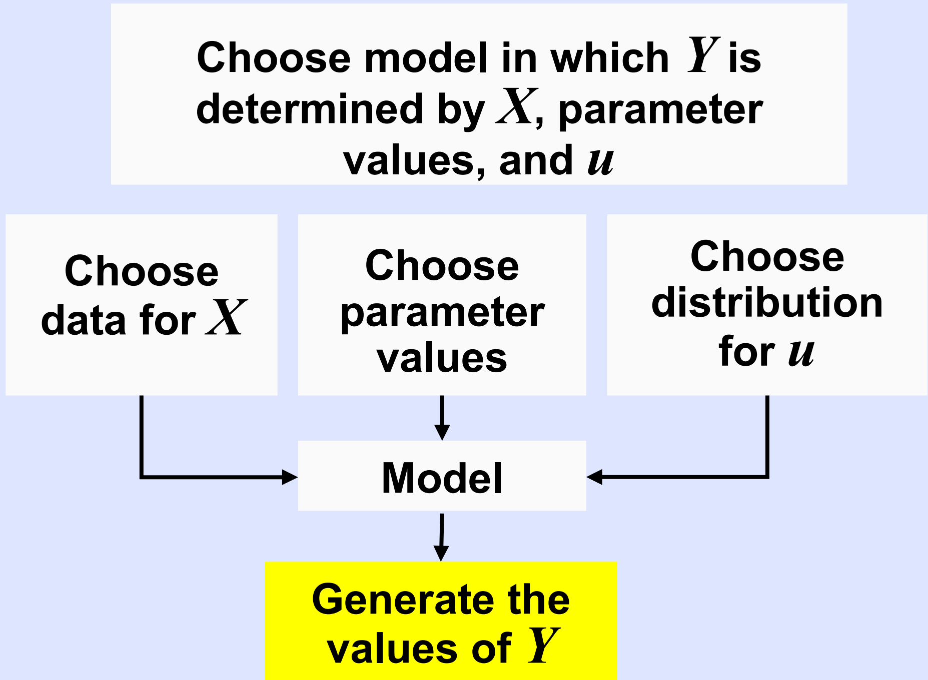
Choose data for X

Choose parameter values

Choose distribution for u

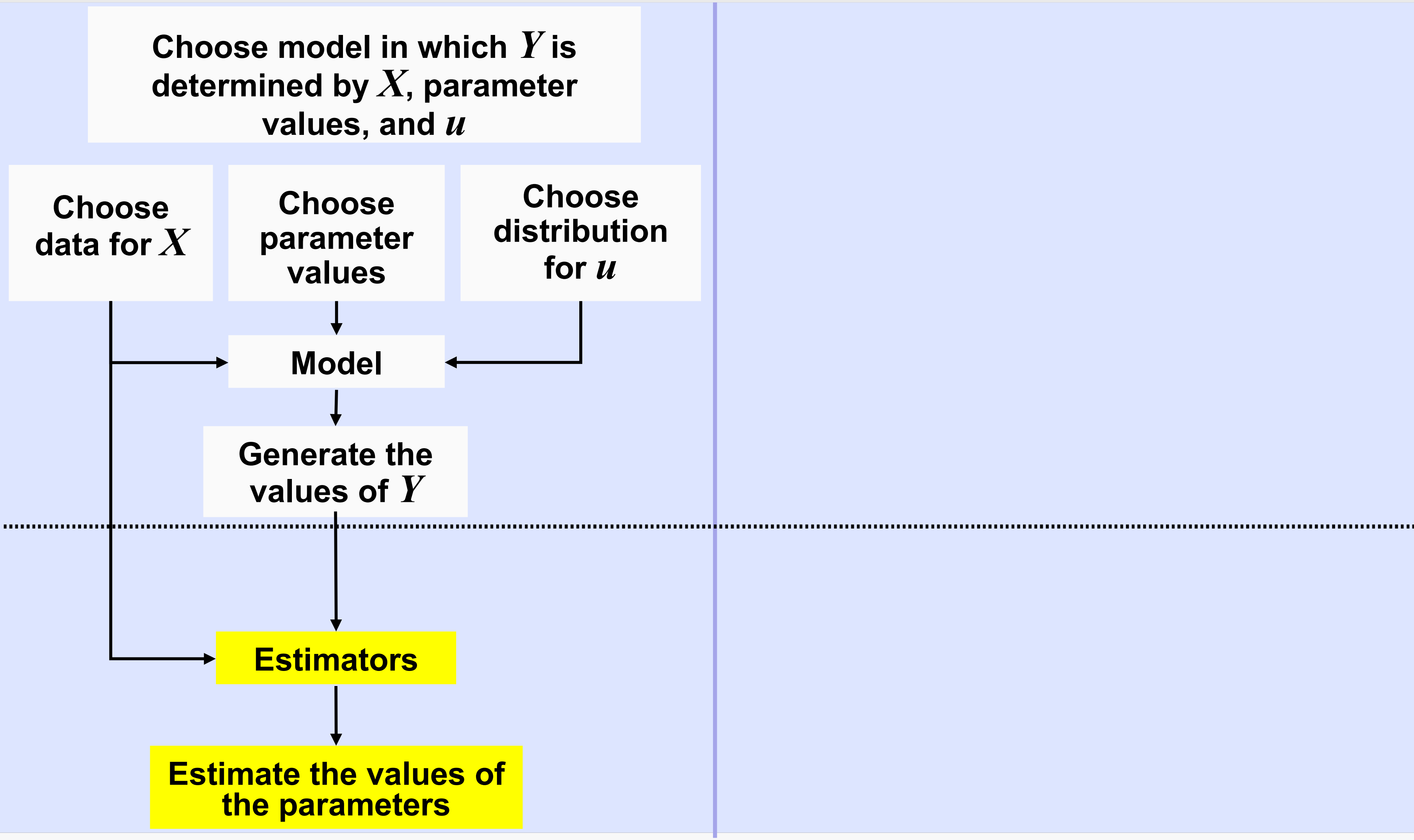
We will also generate values for the disturbance term randomly from a known distribution.

A MONTE CARLO EXPERIMENT



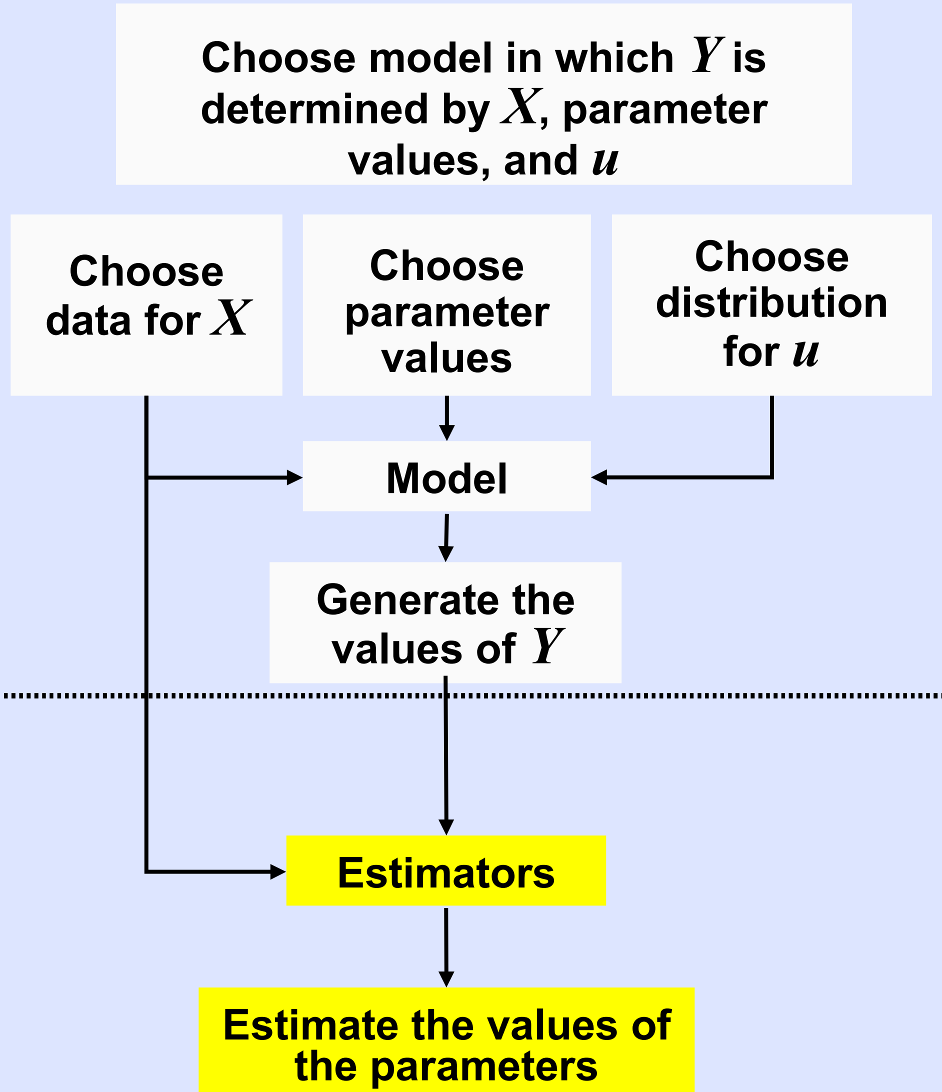
The values of Y in the sample will be determined by the values of X, the parameters and the values of the disturbance term.

A MONTE CARLO EXPERIMENT



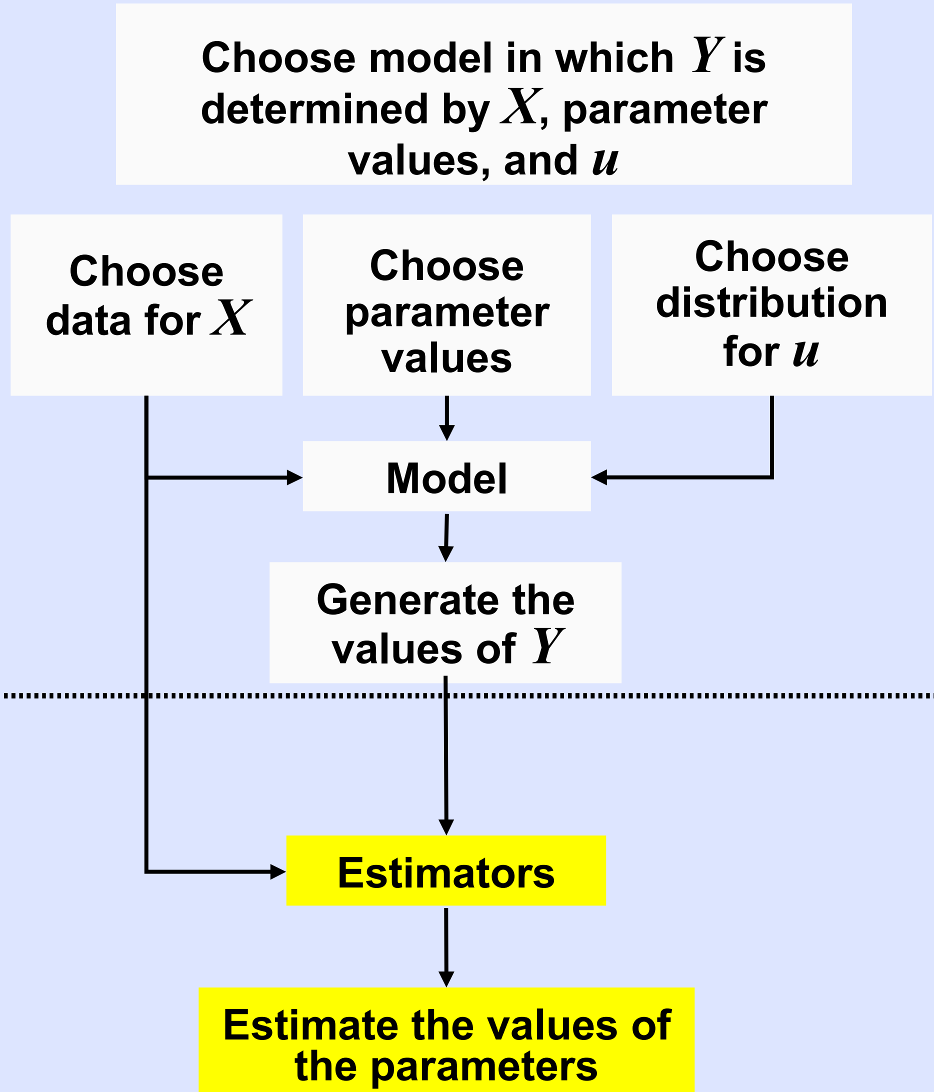
We will then use the regression technique to obtain estimates of the parameters using only the data on Y and X.

A MONTE CARLO EXPERIMENT



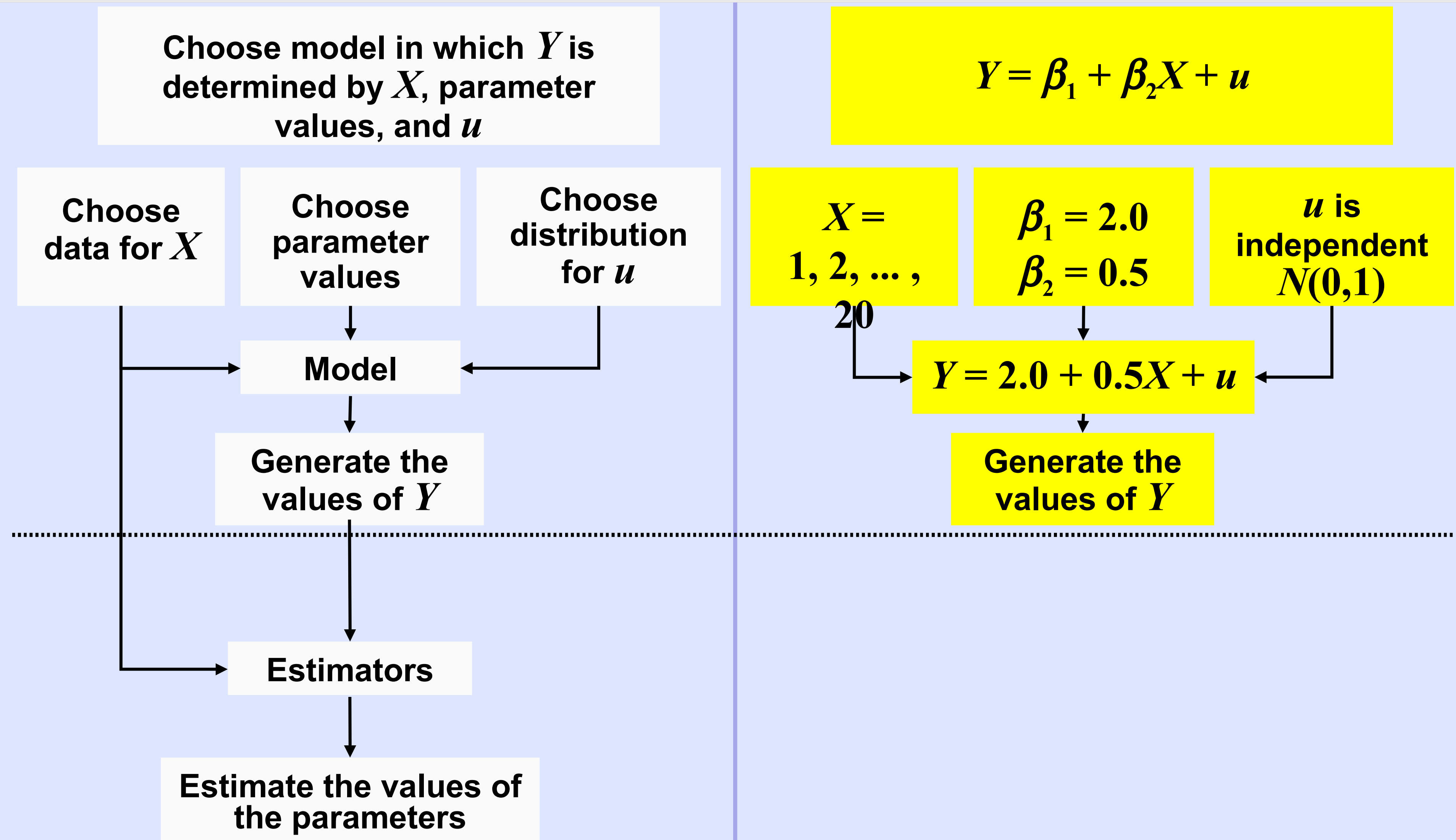
We can repeat the process indefinitely keeping the same data for *X* and the same values of the parameters but using new randomly-generated values for the disturbance term.

A MONTE CARLO EXPERIMENT



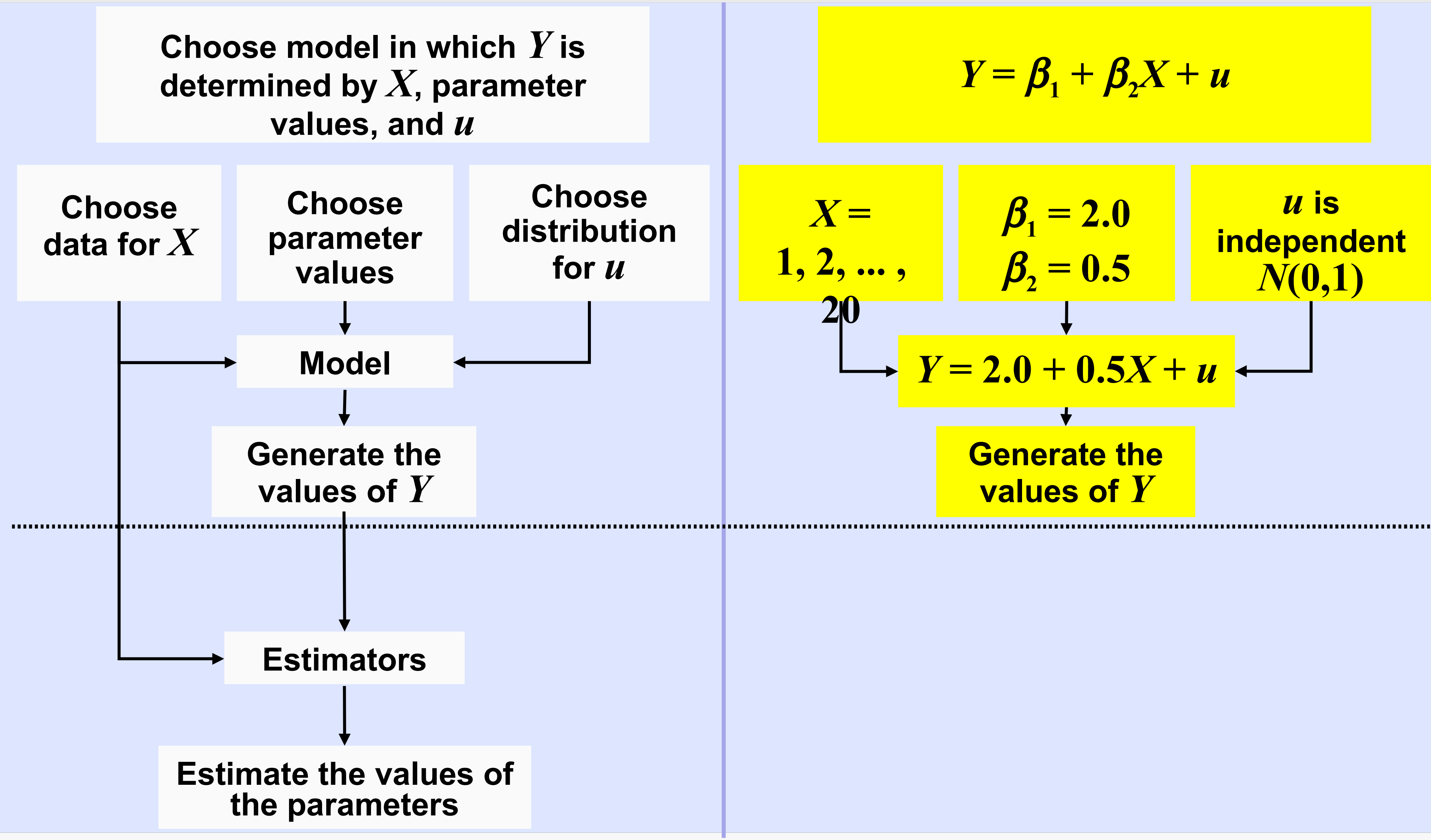
In this way we can derive probability distributions for the regression estimators which allow us, for example, to check up on whether they are biased or unbiased.

A MONTE CARLO EXPERIMENT



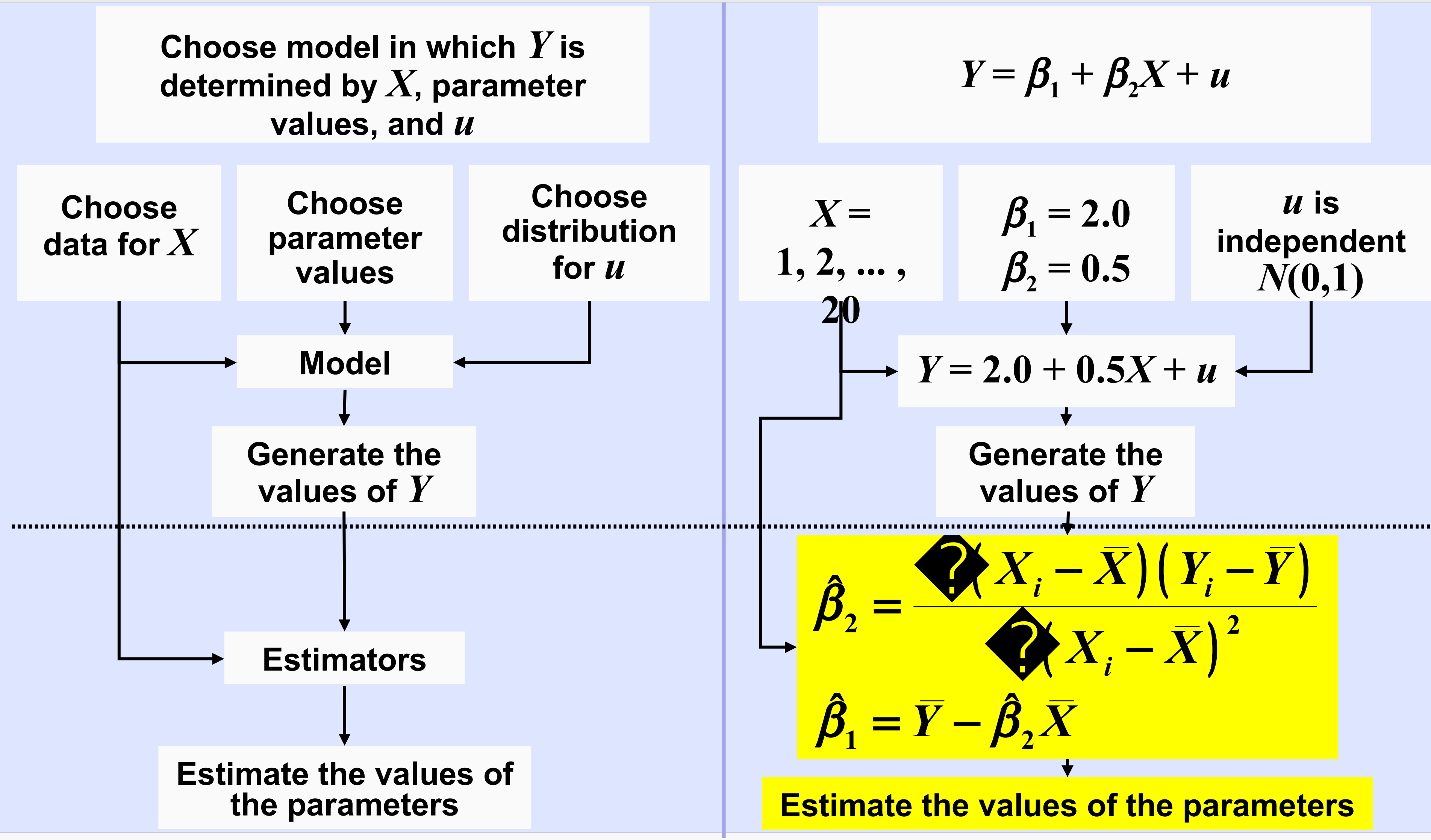
In this experiment we have 20 observations in the sample. X takes the values 1, 2, ..., 20. β_1 is equal to 2.0 and β_2 is equal to 0.5.

A MONTE CARLO EXPERIMENT



The disturbance term is generated randomly using a normal distribution with zero mean and unit variance. Hence we generate the values of *Y*.

A MONTE CARLO EXPERIMENT



We will then regress Y on X using the OLS estimation technique and see how well our estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ correspond to the true values β_1 and β_2.

A MONTE CARLO EXPERIMENT

$$Y = 2.0 + 0.5X + u$$

X	$2.0+0.5X$	u	Y	X	$2.0+0.5X$	u	Y
1				11			
2				12			
3				13			
4				14			
5				15			
6				16			
7				17			
8				18			
9				19			
10				20			

Here are the values of *X*, chosen quite arbitrarily.

A MONTE CARLO EXPERIMENT

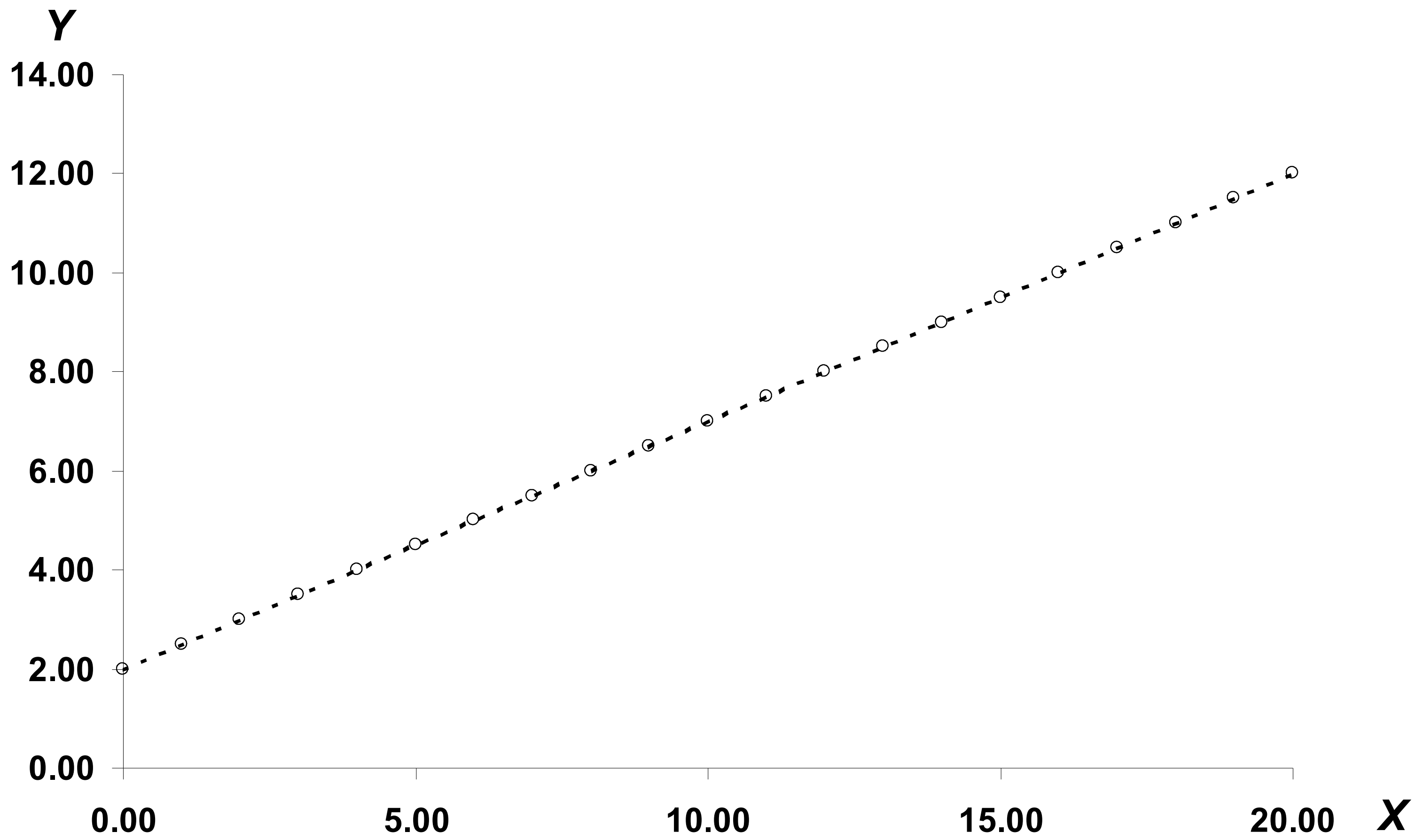
$$Y = 2.0 + 0.5X + u$$

X	$2.0+0.5X$	u	Y	X	$2.0+0.5X$	u	Y
1	2.5			11	7.5		
2	3.0			12	8.0		
3	3.5			13	8.5		
4	4.0			14	9.0		
5	4.5			15	9.5		
6	5.0			16	10.0		
7	5.5			17	10.5		
8	6.0			18	11.0		
9	6.5			19	11.5		
10	7.0			20	12.0		

Given our choice of numbers for β_1 and β_2, we can derive the nonstochastic component of Y.

A MONTE CARLO EXPERIMENT

$$Y = 2.0 + 0.5X + u$$



The nonstochastic component is displayed graphically.

A MONTE CARLO EXPERIMENT

$$Y = 2.0 + 0.5X + u$$

X	$2.0+0.5X$	u	Y	X	$2.0+0.5X$	u	Y
1	2.5	–0.59		11	7.5	1.59	
2	3.0	–0.24		12	8.0	–0.92	
3	3.5	–0.83		13	8.5	–0.71	
4	4.0	0.03		14	9.0	–0.25	
5	4.5	–0.38		15	9.5	1.69	
6	5.0	–2.19		16	10.0	0.15	
7	5.5	1.03		17	10.5	0.02	
8	6.0	0.24		18	11.0	–0.11	
9	6.5	2.53		19	11.5	–0.91	
10	7.0	–0.13		20	12.0	1.42	

Next, we generate randomly a value of the disturbance term for each observation using a $N(0,1)$ distribution (normal with zero mean and unit variance).

A MONTE CARLO EXPERIMENT

$$Y = 2.0 + 0.5X + u$$

X	$2.0+0.5X$	u	Y	X	$2.0+0.5X$	u	Y
1	2.5	–0.59	1.91	11	7.5	1.59	
2	3.0	–0.24		12	8.0	–0.92	
3	3.5	–0.83		13	8.5	–0.71	
4	4.0	0.03		14	9.0	–0.25	
5	4.5	–0.38		15	9.5	1.69	
6	5.0	–2.19		16	10.0	0.15	
7	5.5	1.03		17	10.5	0.02	
8	6.0	0.24		18	11.0	–0.11	
9	6.5	2.53		19	11.5	–0.91	
10	7.0	–0.13		20	12.0	1.42	

Thus, for example, the value of Y in the first observation is 1.91, not 2.50.

A MONTE CARLO EXPERIMENT

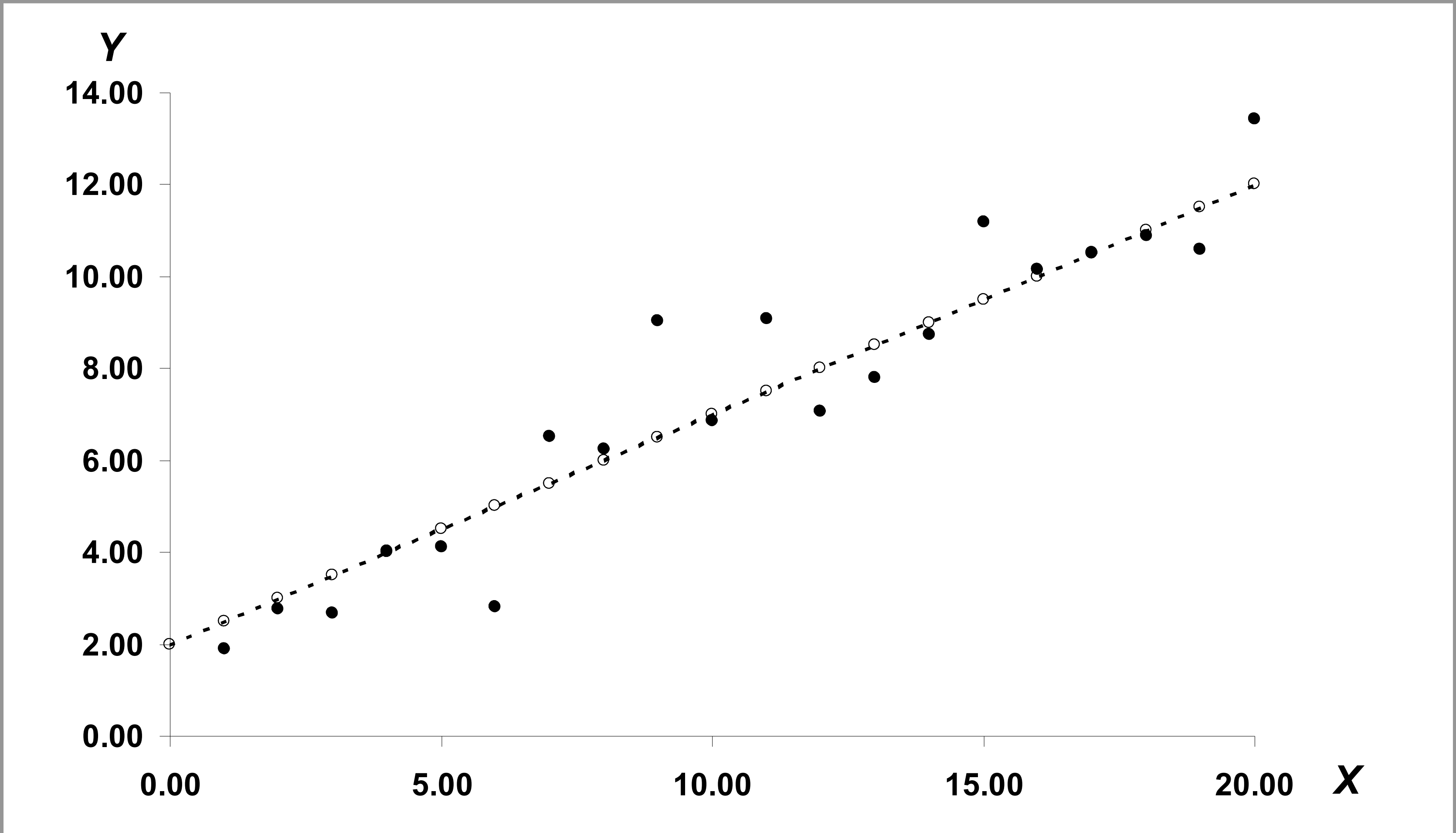
$$Y = 2.0 + 0.5X + u$$

X	$2.0+0.5X$	u	Y	X	$2.0+0.5X$	u	Y
1	2.5	–0.59	1.91	11	7.5	1.59	9.09
2	3.0	–0.24	2.76	12	8.0	–0.92	7.08
3	3.5	–0.83	2.67	13	8.5	–0.71	7.79
4	4.0	0.03	4.03	14	9.0	–0.25	8.75
5	4.5	–0.38	4.12	15	9.5	1.69	11.19
6	5.0	–2.19	2.81	16	10.0	0.15	10.15
7	5.5	1.03	6.53	17	10.5	0.02	10.52
8	6.0	0.24	6.24	18	11.0	–0.11	10.89
9	6.5	2.53	9.03	19	11.5	–0.91	10.59
10	7.0	–0.13	6.87	20	12.0	1.42	13.42

Similarly, we generate values of Y for the other 19 observations.

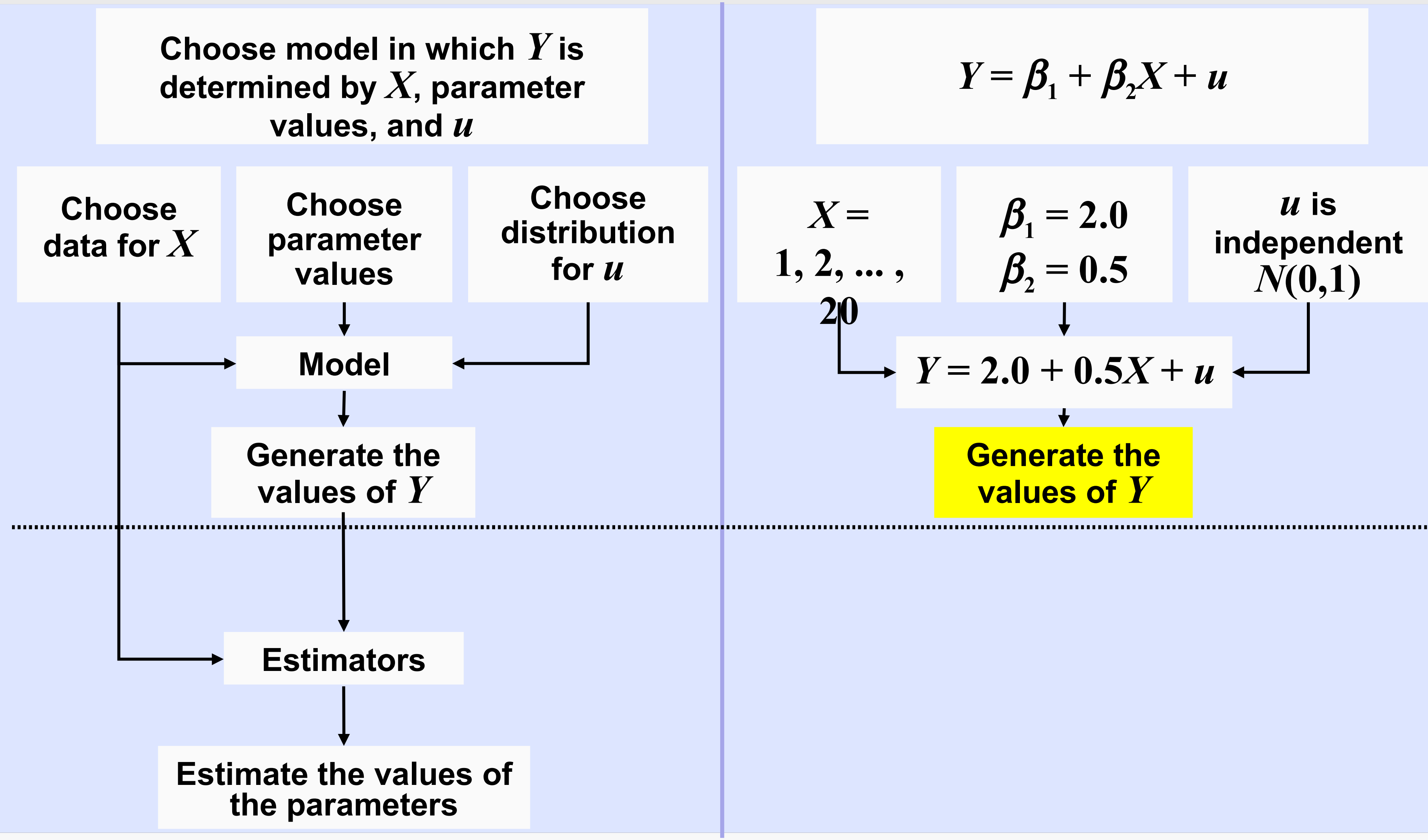
A MONTE CARLO EXPERIMENT

$$Y = 2.0 + 0.5X + u$$



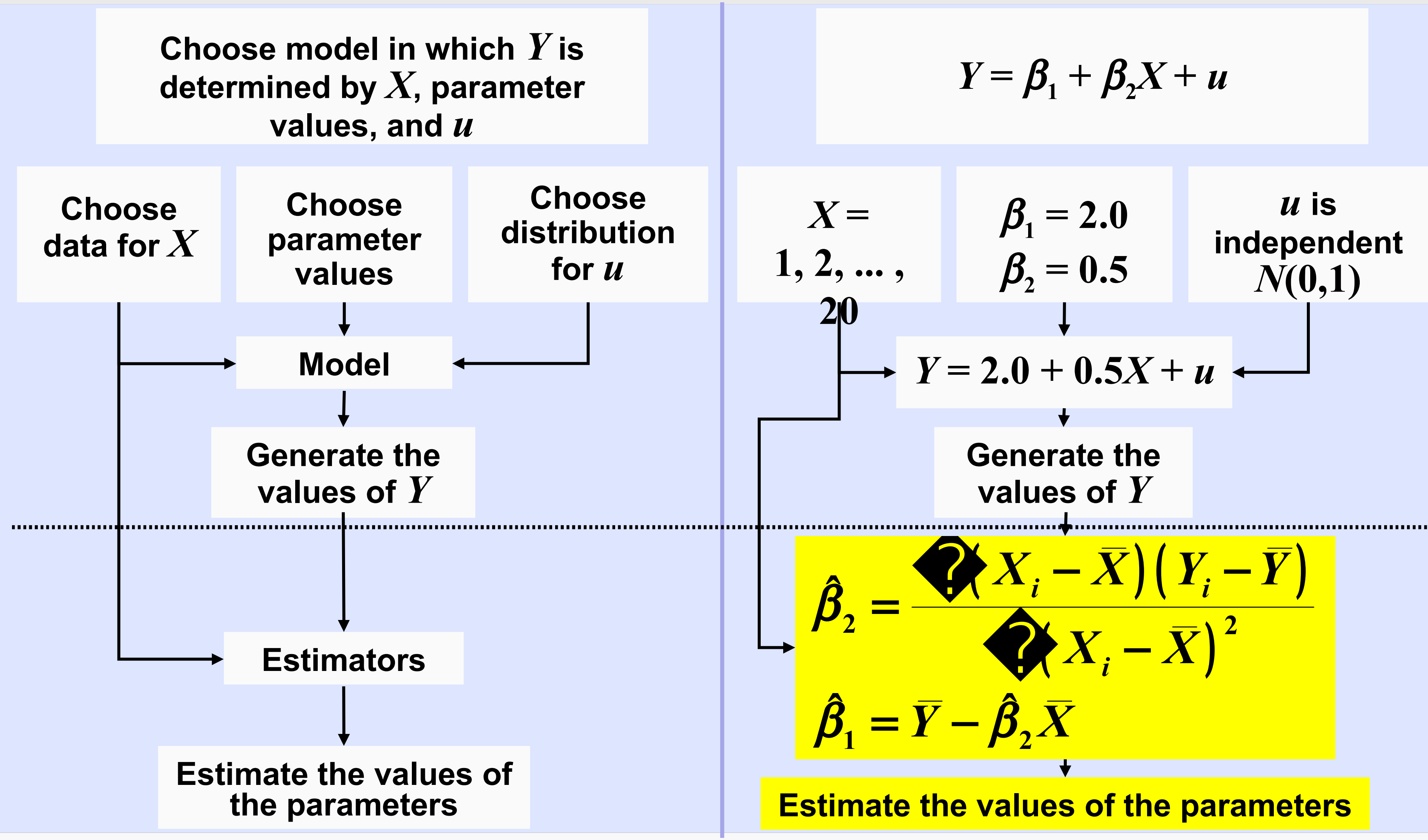
The 20 observations are displayed graphically.

A MONTE CARLO EXPERIMENT



We have reached this point in the Monte Carlo experiment.

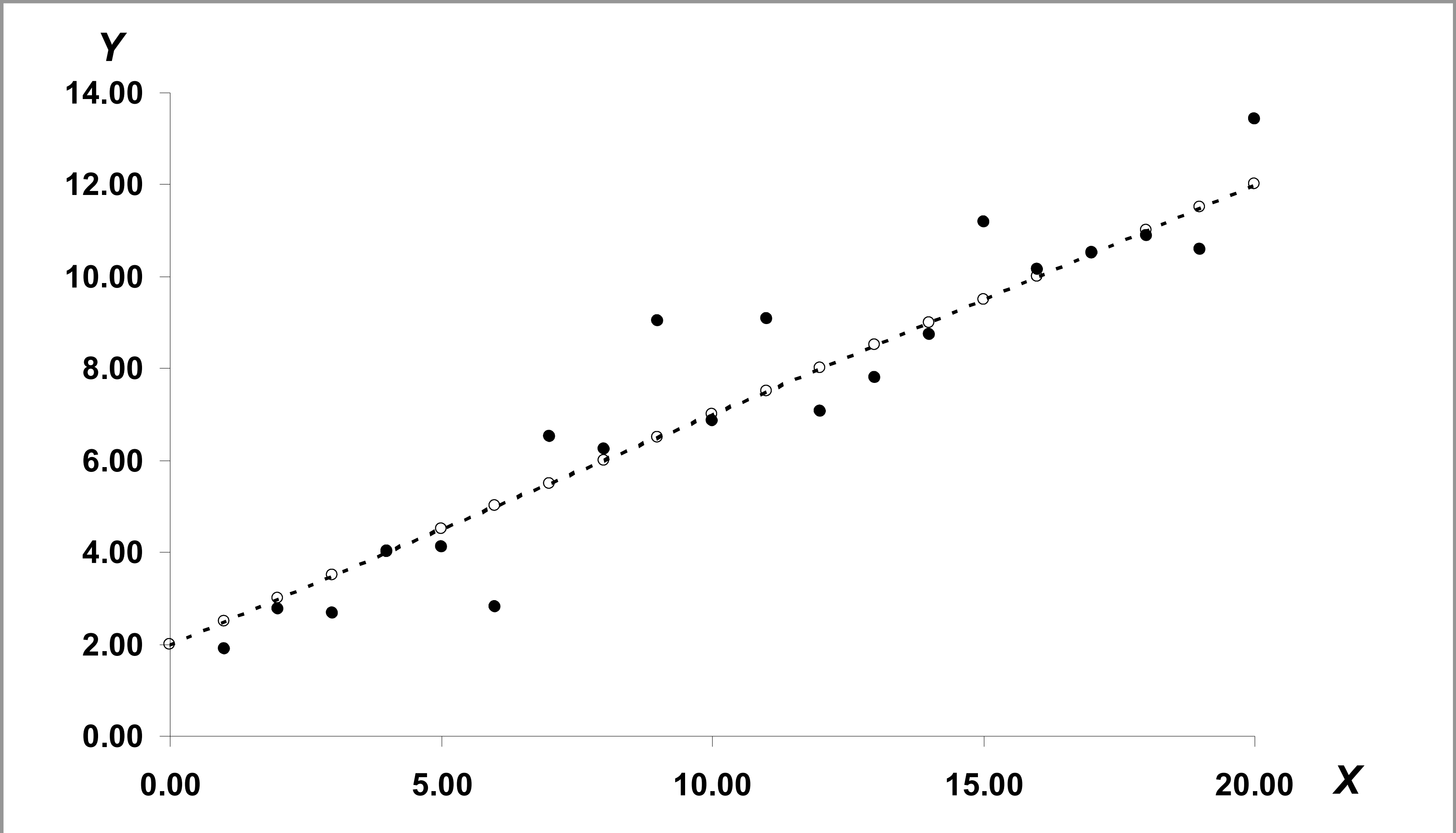
A MONTE CARLO EXPERIMENT



We will now apply the OLS estimators for $\hat{\beta}_1$ and $\hat{\beta}_2$ to the data for *X* and *Y*, and see how well the estimates correspond to the true values.

A MONTE CARLO EXPERIMENT

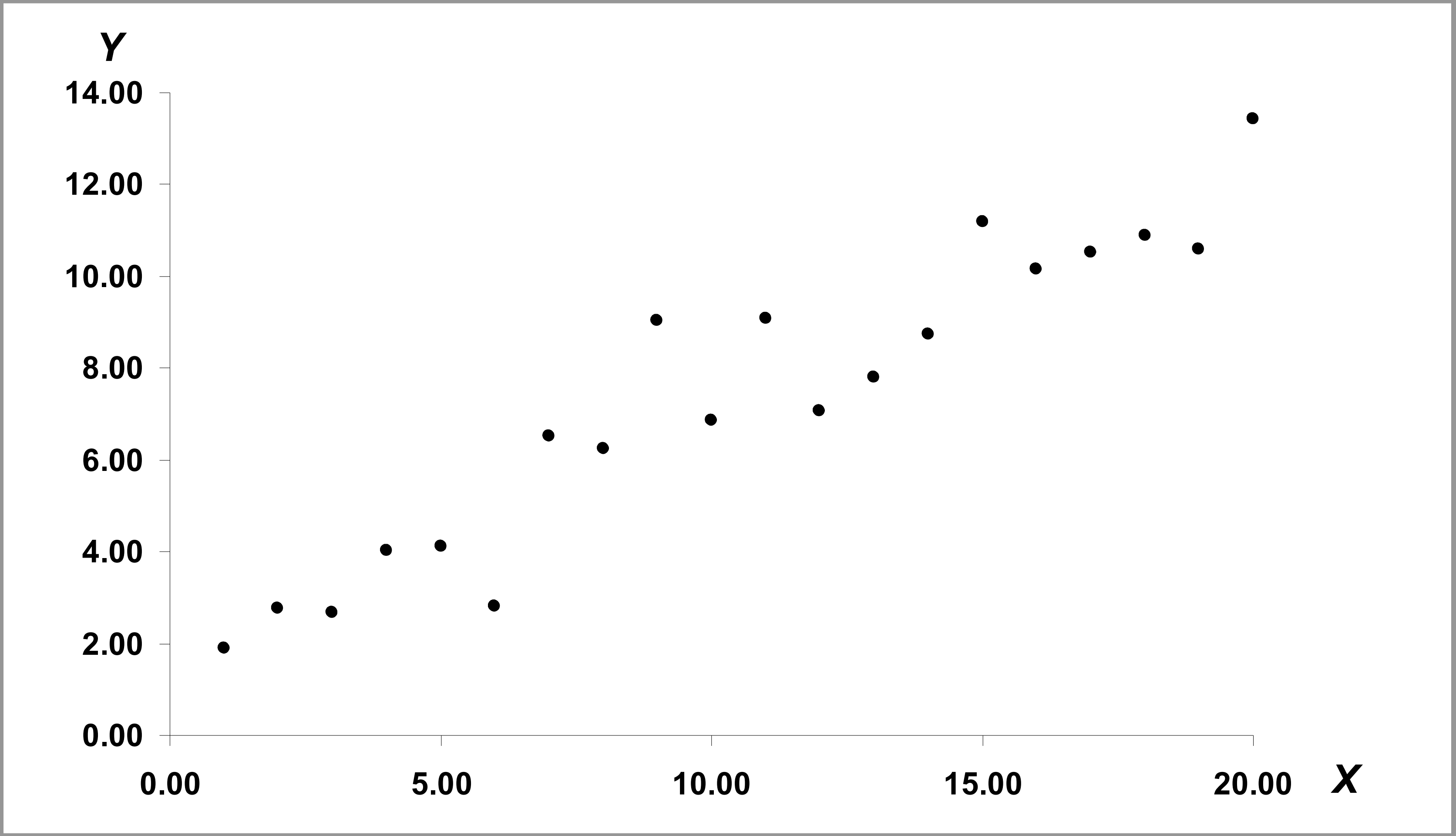
$$Y = 2.0 + 0.5X + u$$



Here is the scatter diagram again.

A MONTE CARLO EXPERIMENT

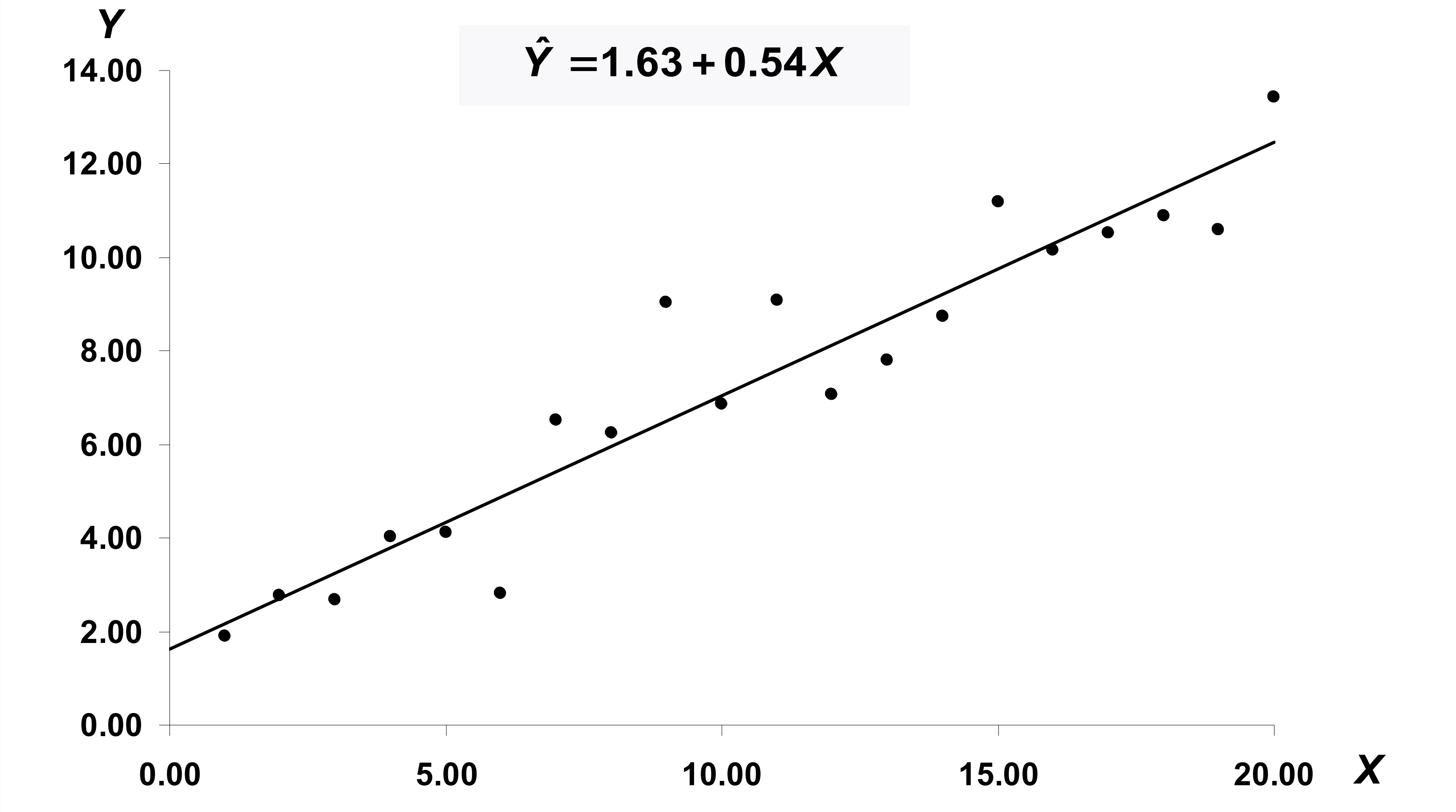
$$Y = 2.0 + 0.5X + u$$



The regression estimators use only the observed data for *X* and *Y*.

A MONTE CARLO EXPERIMENT

$$Y = 2.0 + 0.5X + u$$

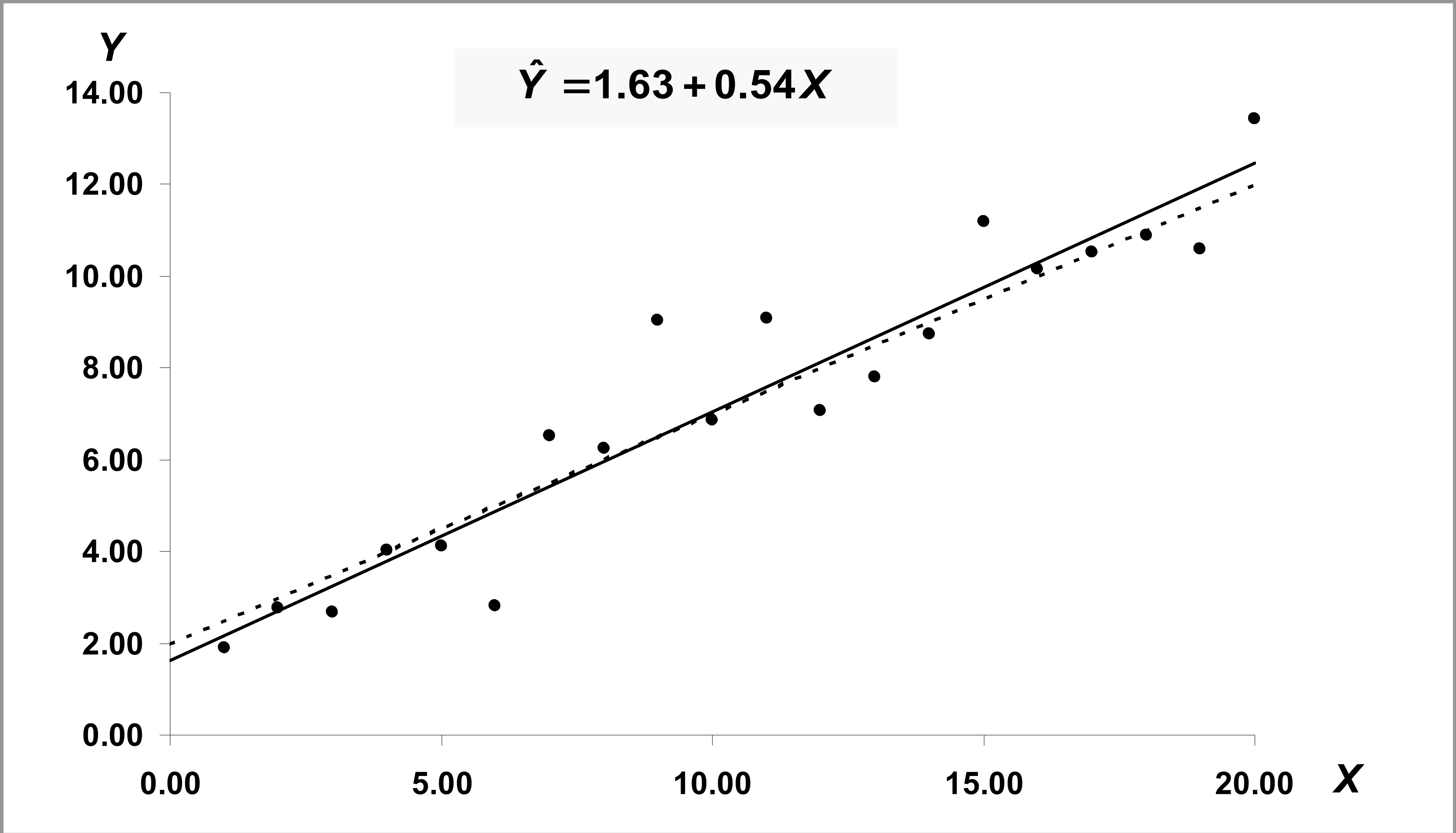


$$\hat{Y} = 1.63 + 0.54X$$

Here is the regression equation fitted to the data.

A MONTE CARLO EXPERIMENT

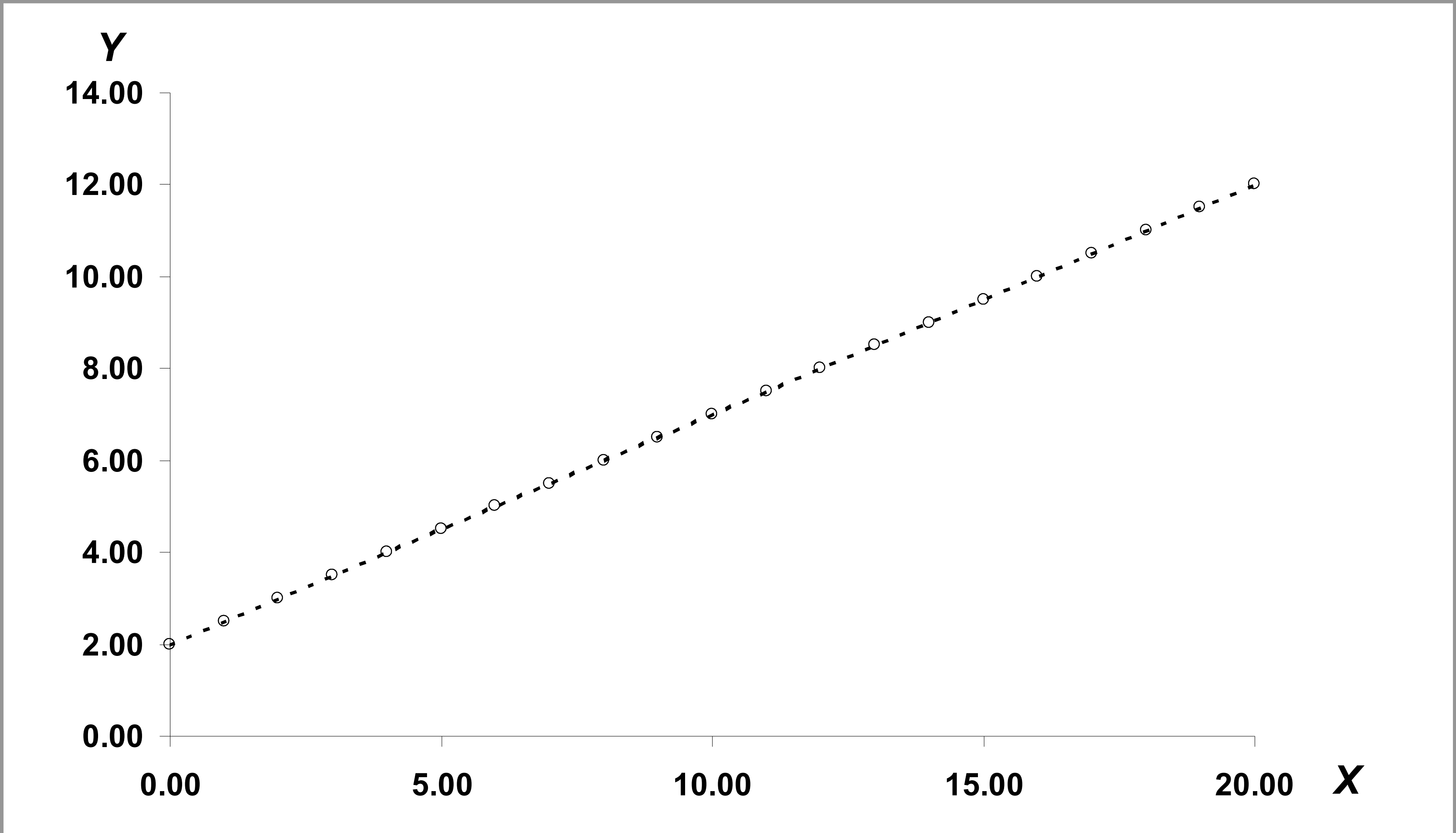
$$Y = 2.0 + 0.5X + u$$



For comparison, the nonstochastic component of the true relationship is also displayed. β_2 (true value 0.50) has been overestimated and β_1 (true value 2.00) has been underestimated.

A MONTE CARLO EXPERIMENT

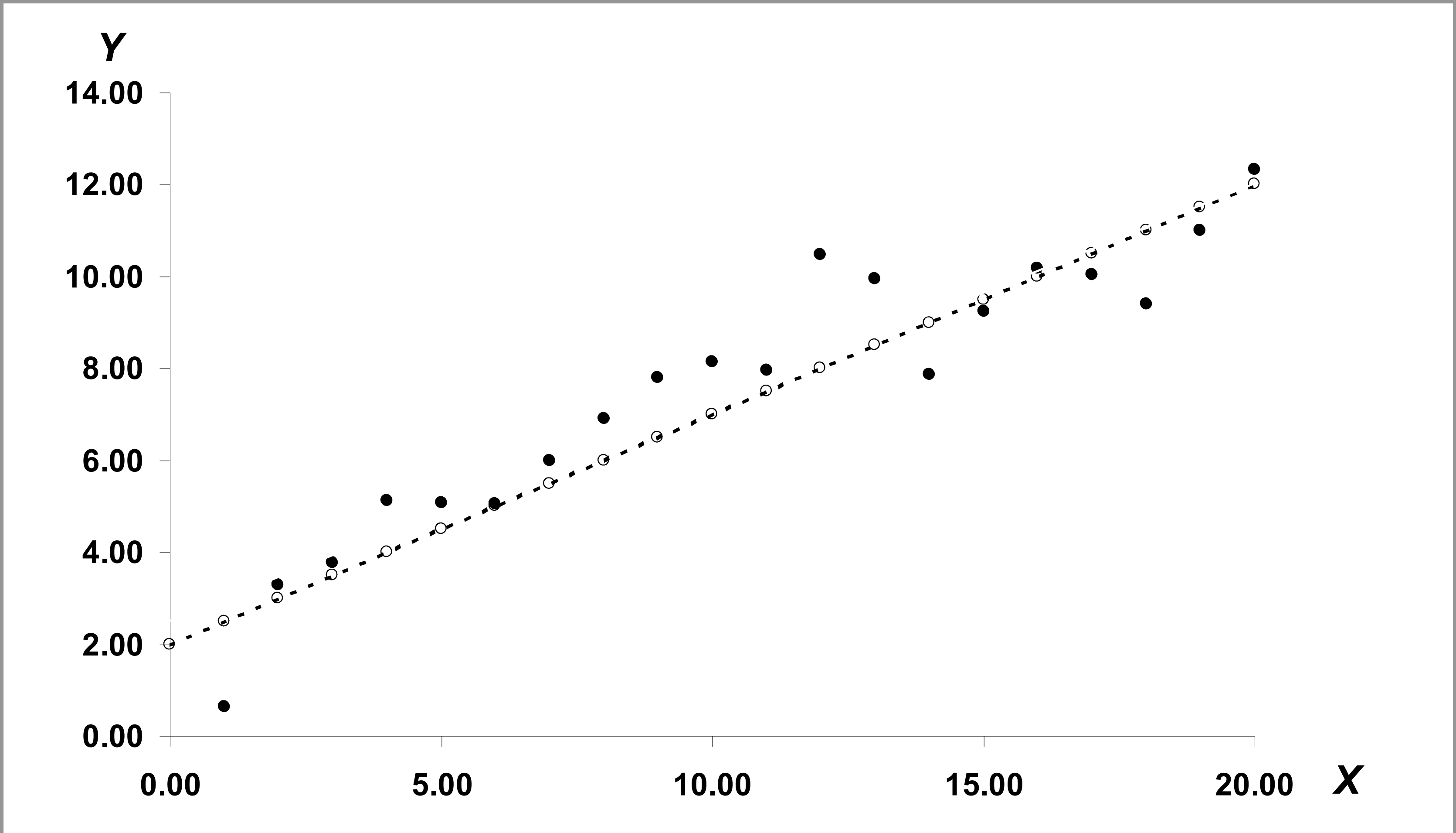
$$Y = 2.0 + 0.5X + u$$



We will repeat the process, starting with the same nonstochastic components of *Y*.

A MONTE CARLO EXPERIMENT

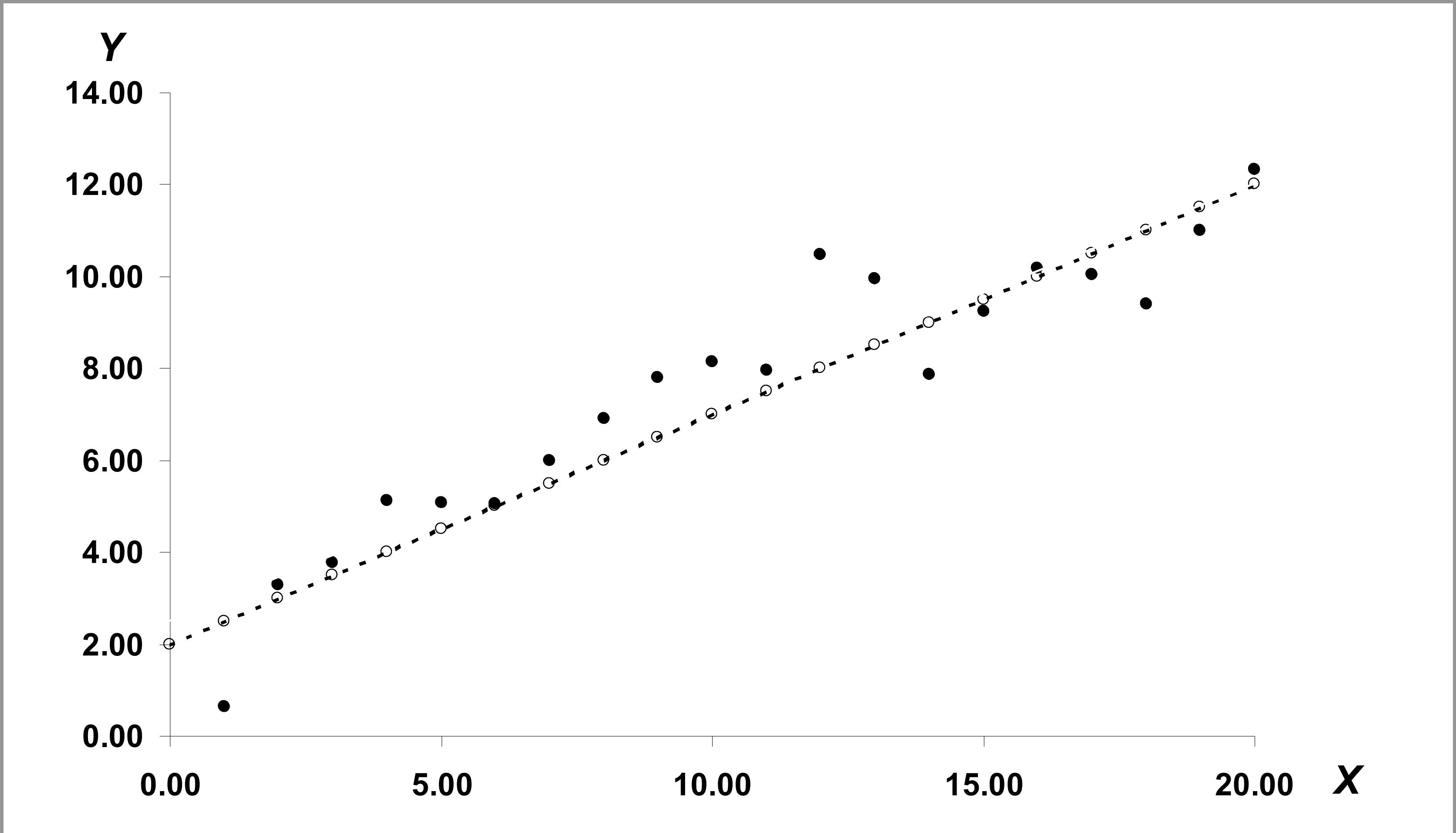
$$Y = 2.0 + 0.5X + u$$



As before, the values of Y are modified by adding randomly-generated values of the disturbance term.

A MONTE CARLO EXPERIMENT

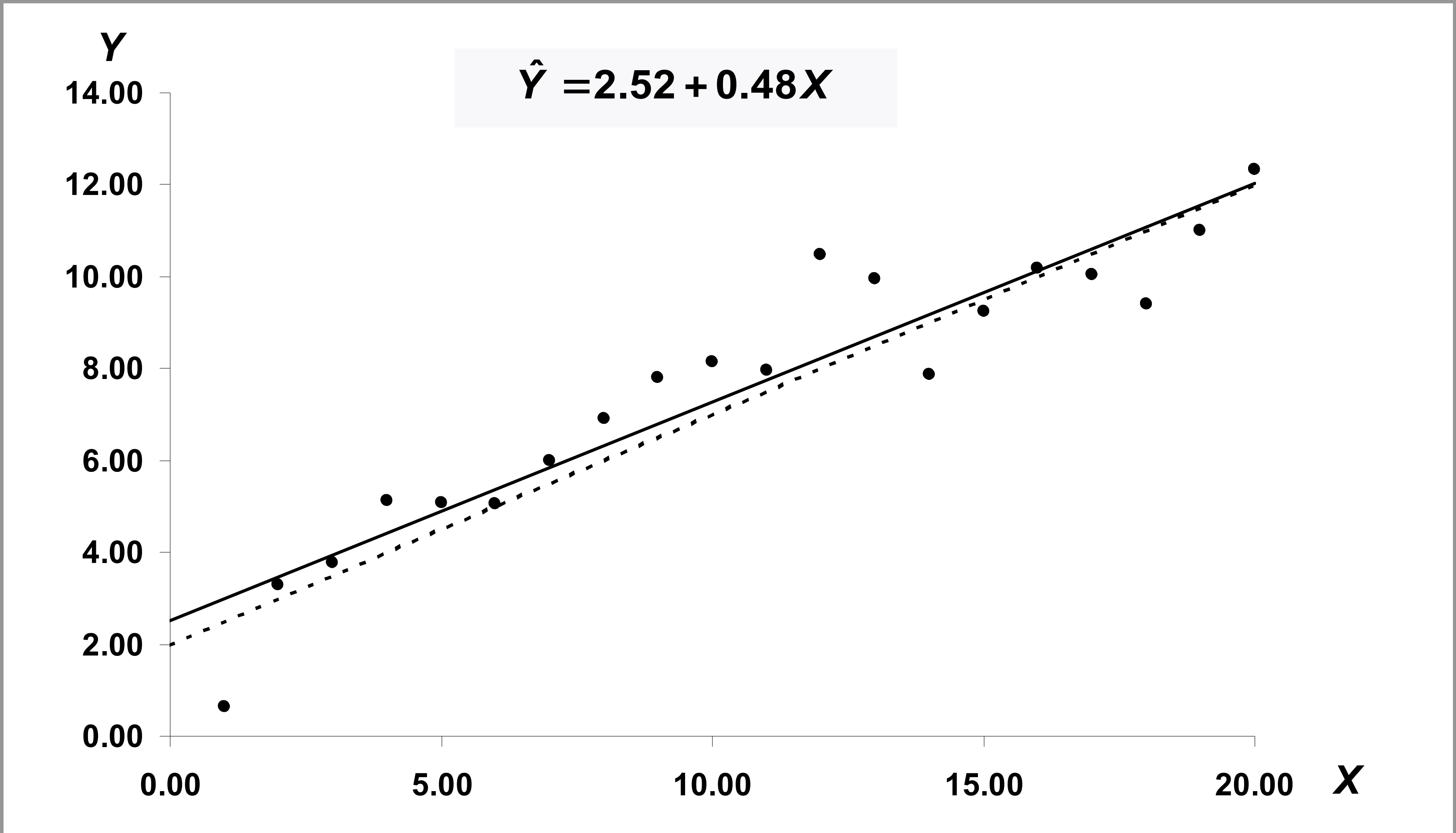
$Y = 2.0 + 0.5X + u$



The new values of the disturbance term are drawn from the same $N(0,1)$ distribution as the previous ones but, except by coincidence, will be different from them.

A MONTE CARLO EXPERIMENT

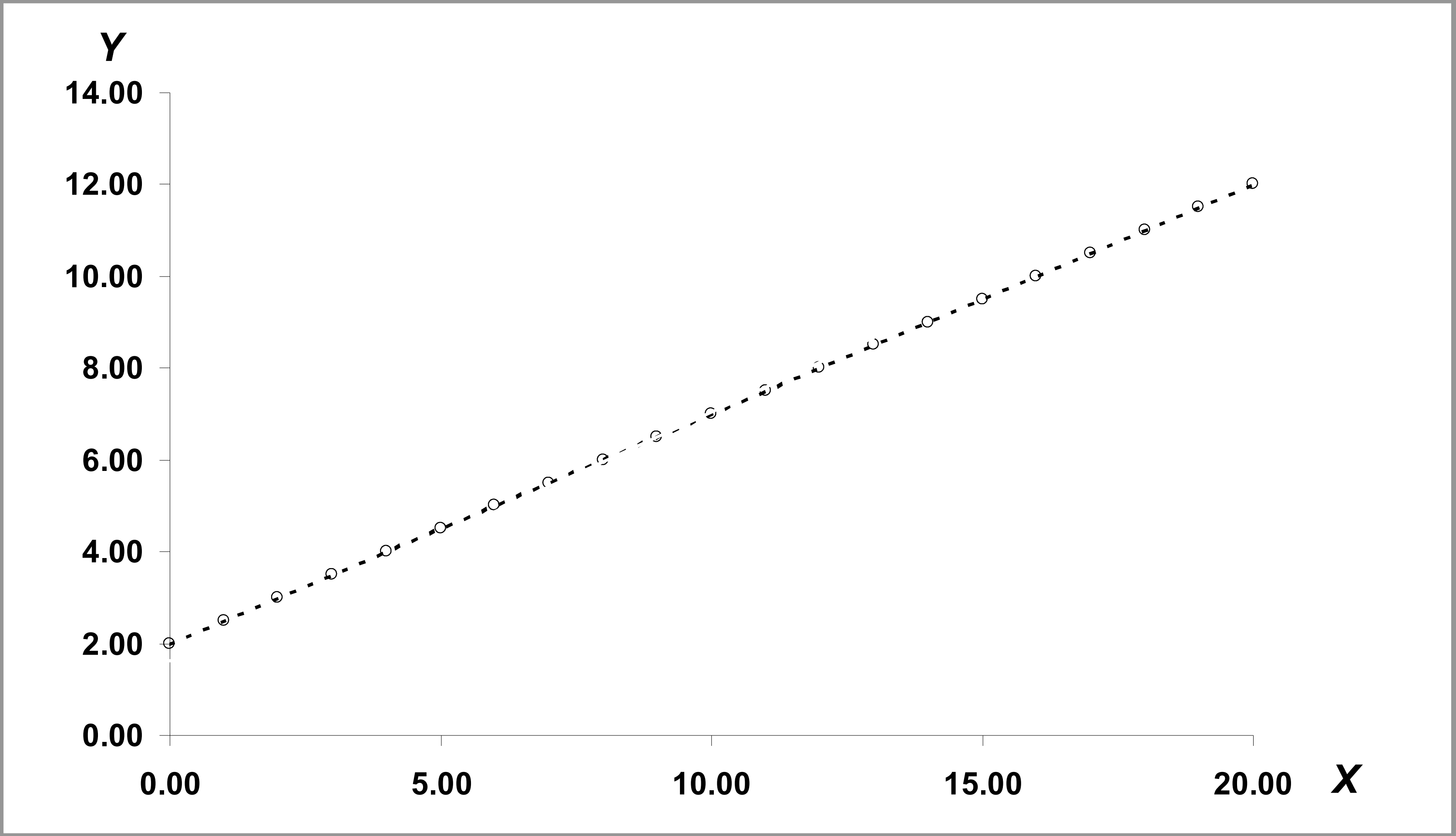
$$Y = 2.0 + 0.5X + u$$



This time the slope coefficient has been underestimated and the intercept overestimated.

A MONTE CARLO EXPERIMENT

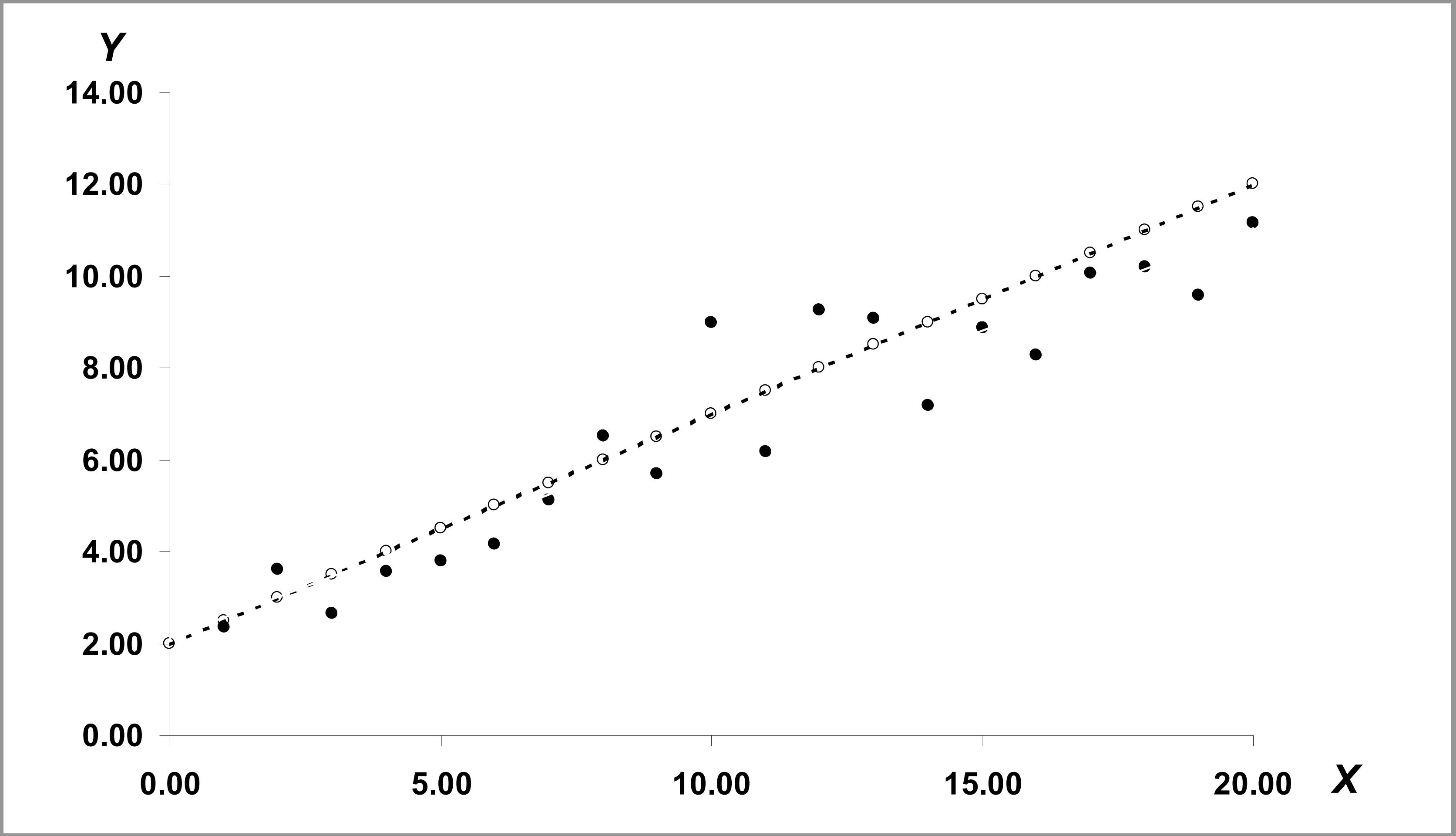
$$Y = 2.0 + 0.5X + u$$



We will repeat the process once more.

A MONTE CARLO EXPERIMENT

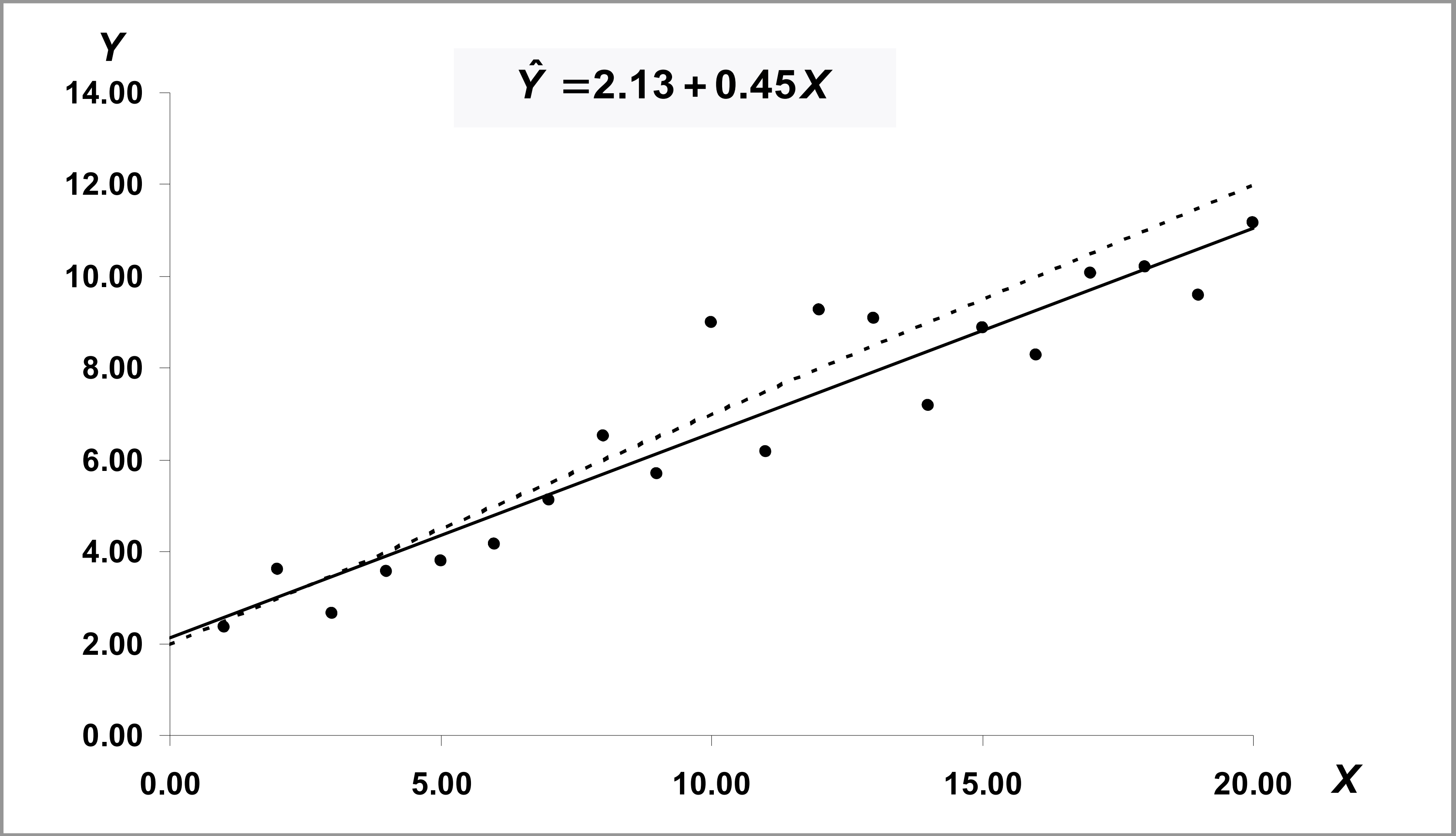
$$Y = 2.0 + 0.5X + u$$



A new set of random numbers has been used in generating the values of *Y*.

A MONTE CARLO EXPERIMENT

$$Y = 2.0 + 0.5X + u$$



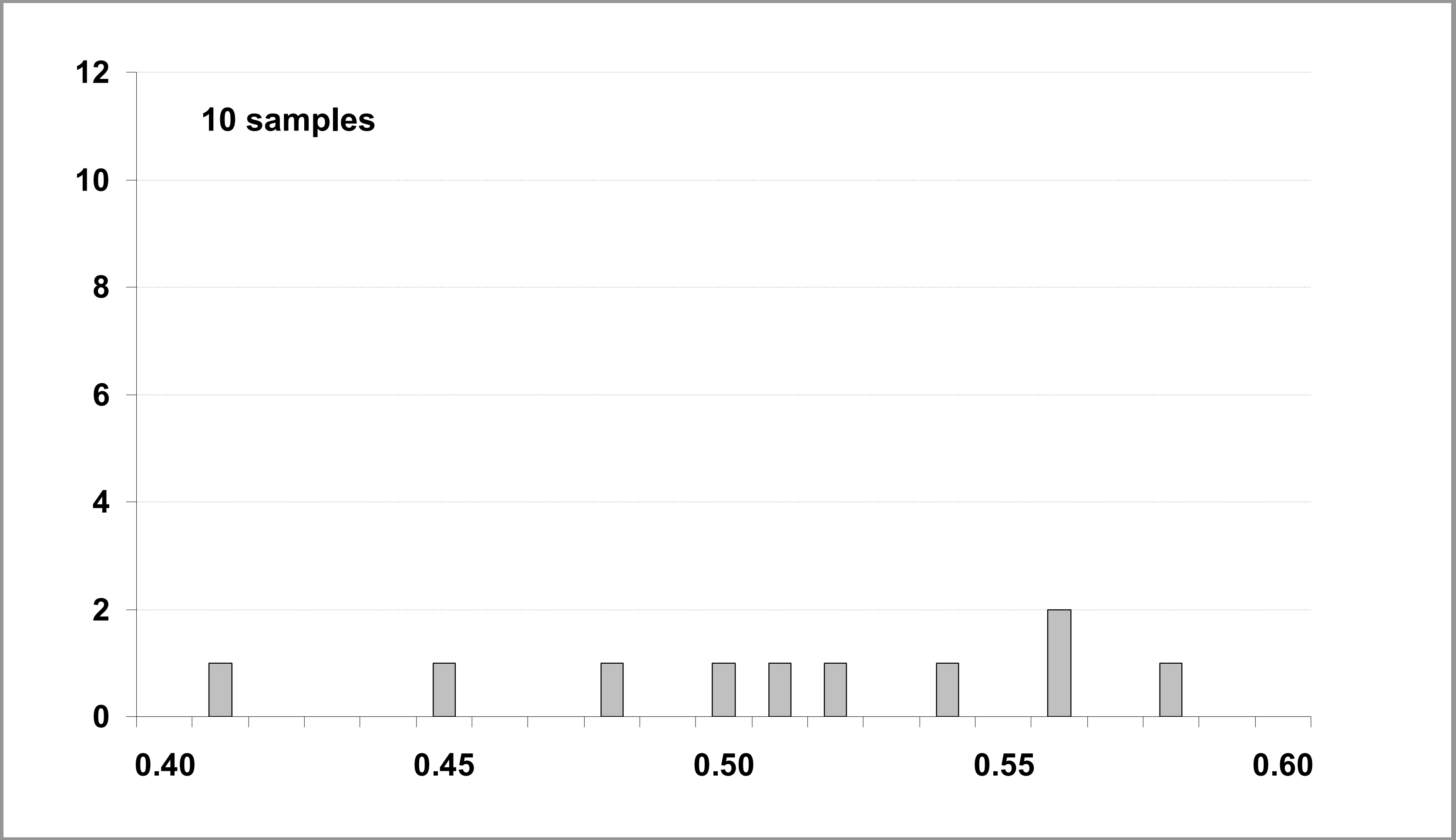
As last time, the slope coefficient has been underestimated and the intercept overestimated.

A MONTE CARLO EXPERIMENT

sample	$\hat{\beta}_1$	$\hat{\beta}_2$
1	1.63	0.54
2	2.52	0.48
3	2.13	0.45
4	2.14	0.50
5	1.71	0.56
6	1.81	0.51
7	1.72	0.56
8	3.18	0.41
9	1.26	0.58
10	1.94	0.52

The table summarizes the results of the three regressions and adds those obtained repeating the process a further seven times.

A MONTE CARLO EXPERIMENT



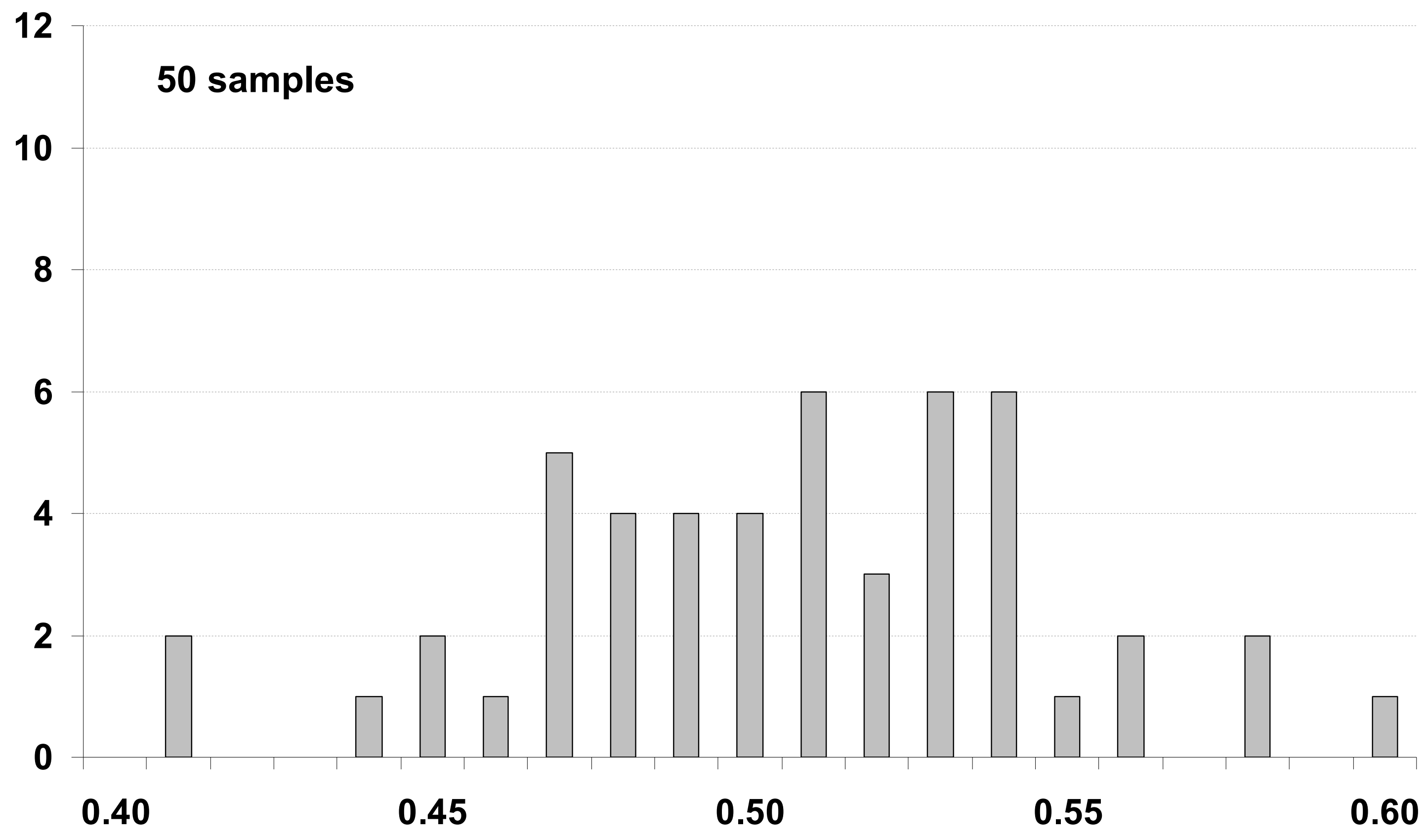
Here is a histogram for the estimates of β_2. Nothing much can be seen yet.

A MONTE CARLO EXPERIMENT

1–10	11–20	21–30	31–40	41–50
0.54	0.49	0.54	0.52	0.49
0.48	0.54	0.46	0.47	0.50
0.45	0.49	0.45	0.54	0.48
0.50	0.54	0.50	0.53	0.44
0.56	0.54	0.41	0.51	0.53
0.51	0.52	0.53	0.51	0.48
0.56	0.49	0.53	0.47	0.47
0.41	0.53	0.47	0.55	0.50
0.58	0.60	0.51	0.51	0.53
0.52	0.48	0.47	0.58	0.51

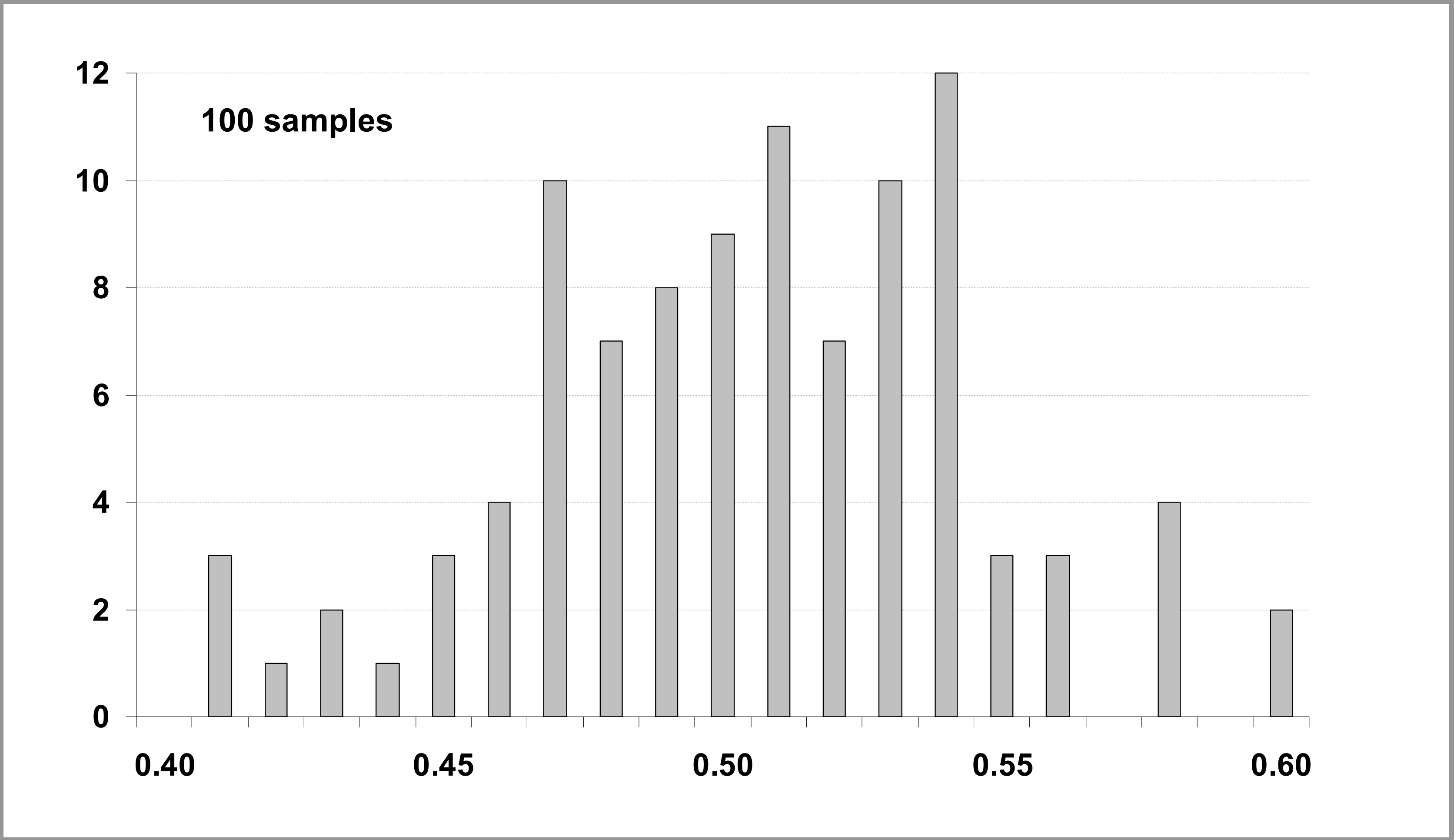
Here are the estimates of β_2 obtained with 40 further replications of the process.

A MONTE CARLO EXPERIMENT



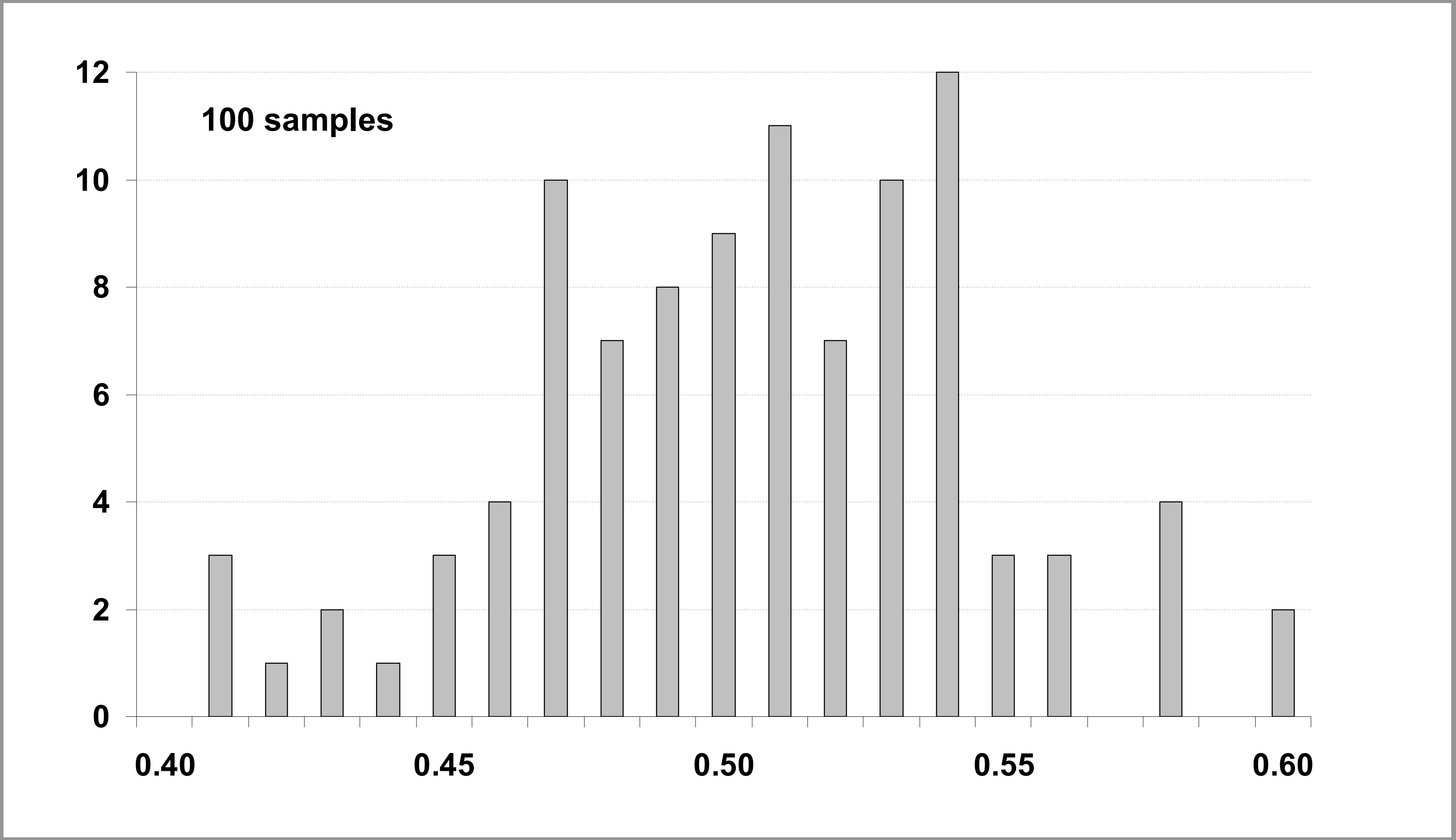
The histogram is beginning to display a central tendency.

A MONTE CARLO EXPERIMENT



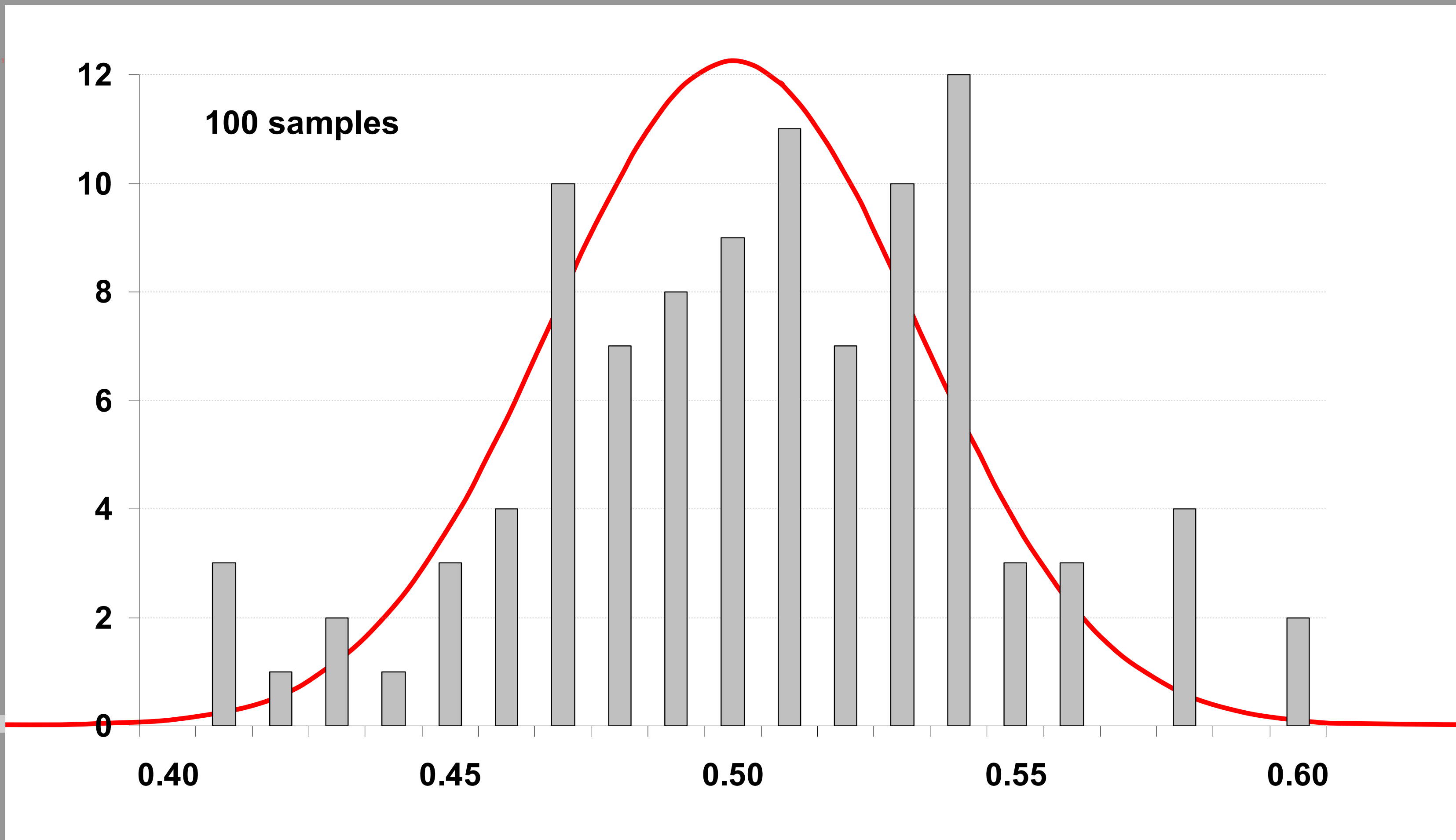
This is the histogram with 100 replications. We can see that the distribution appears to be symmetrical around the true value, implying that the estimator is unbiased.

A MONTE CARLO EXPERIMENT



However, the distribution is still rather jagged. It would be better to repeat the process 1,000,000 times, perhaps more.

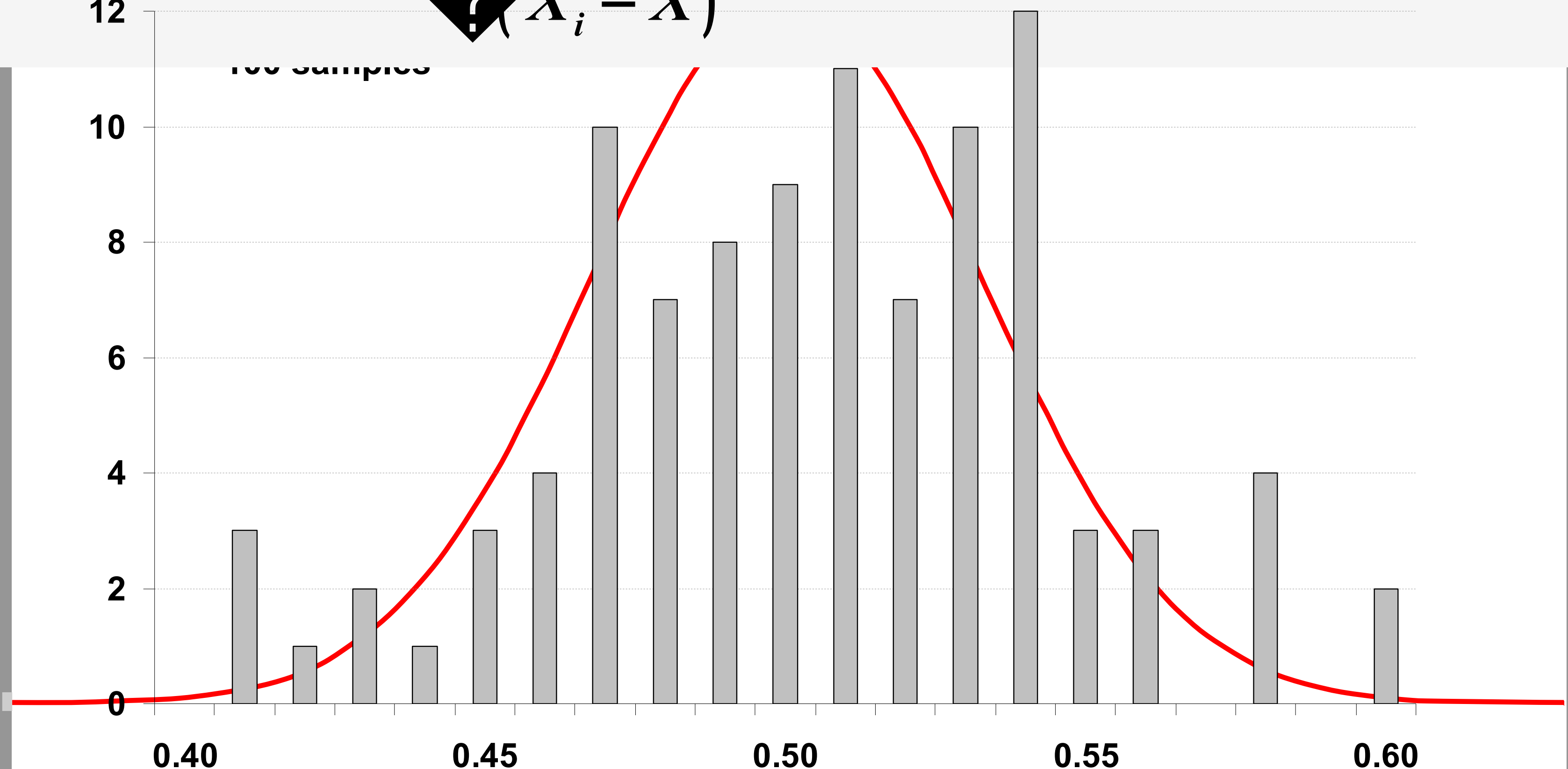
A MONTE CARLO EXPERIMENT



The red curve shows the limiting shape of the distribution. It is symmetrical around the true value, indicating that the estimator is unbiased.

A MONTE CARLO EXPERIMENT

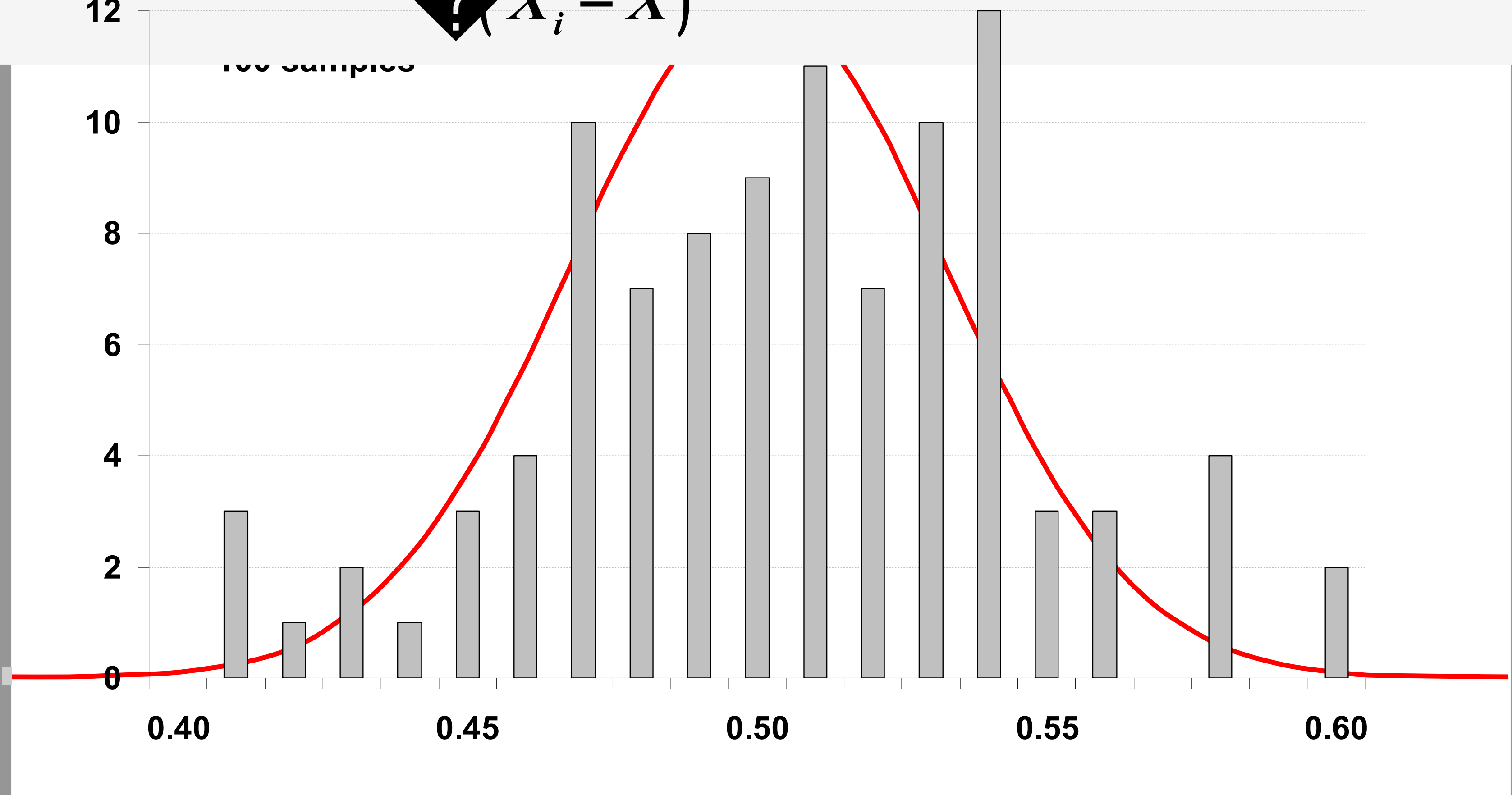
$$\hat{\beta}_2 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \beta_2 + \sum a_t u_i$$



The distribution is normal because the random component of the coefficient is a weighted linear combination of the values of the disturbance term in the observations in the sample. We demonstrated this in the previous slideshow.

A MONTE CARLO EXPERIMENT

$$\hat{\beta}_2 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \beta_2 + \sum a_t u_i$$



We are assuming (Assumption A.6) that the disturbance term in each observation has a normal distribution. A linear combination of normally distributed random variables itself has a normal distribution.

The content of this slideshow comes from Section 2.4 of C. Dougherty, *Introduction to Econometrics*, fifth edition 2016, Oxford University Press. Additional (free) resources for both students and instructors may be downloaded from the OUP Online Resource Centre http://www.oxfordtextbooks.co.uk/orc/dougherty5e//.

Individuals studying econometrics on their own who feel that they might benefit from participation in a formal course should consider the London School of Economics summer school course
EC212 Introduction to Econometrics
http://www2.lse.ac.uk/study/summerSchools/summerSchool/Home.aspx
or the University of London International Programmes distance learning course
EC2020 Elements of Econometrics
www.londoninternational.ac.uk/lse.