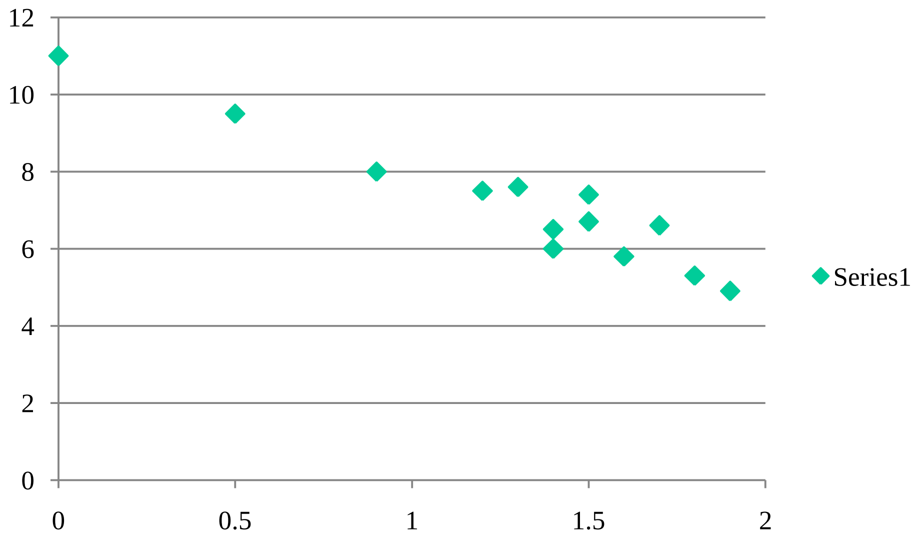
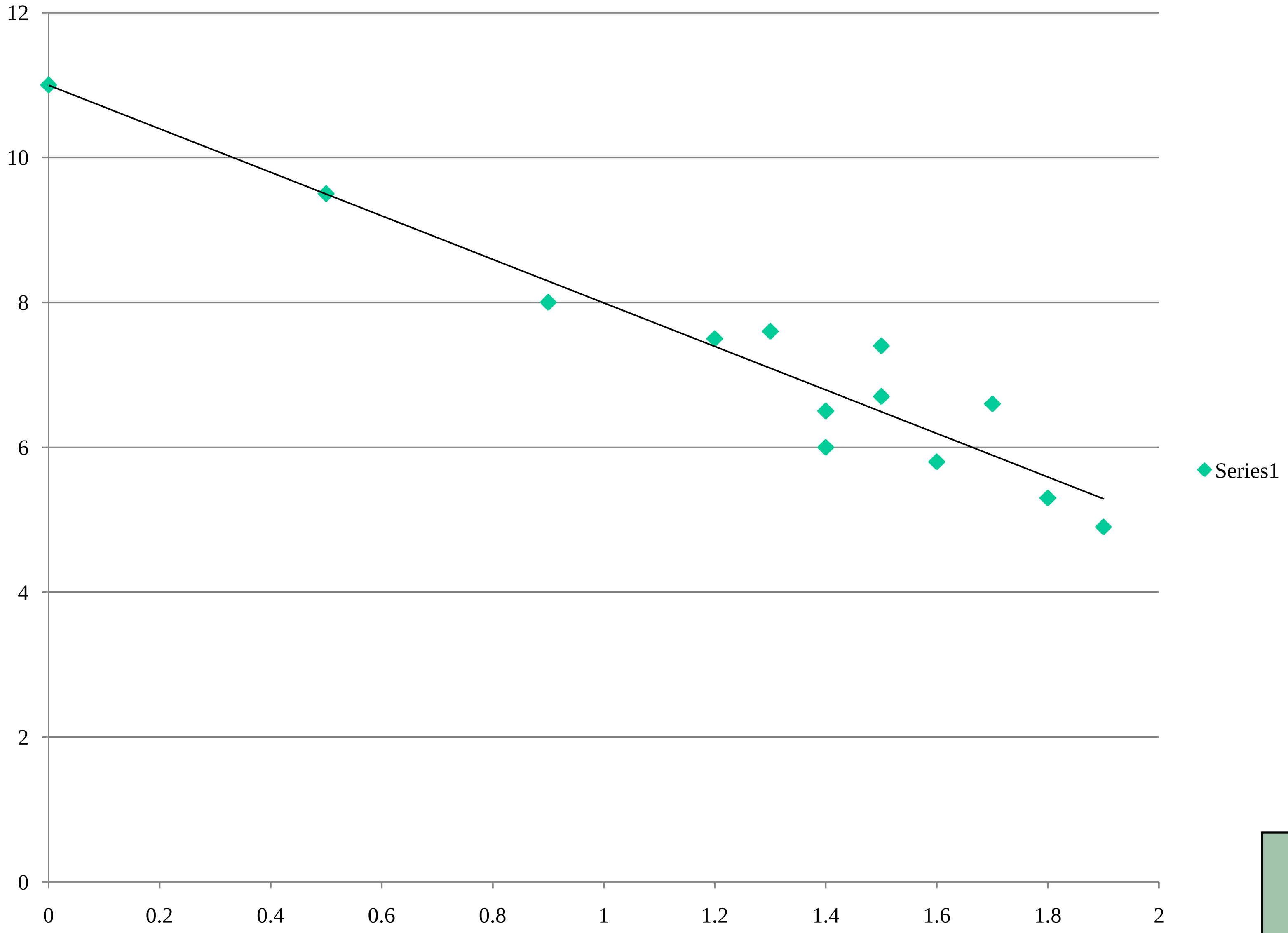
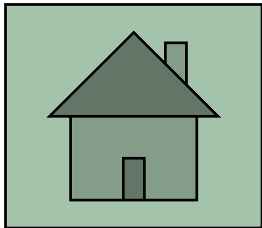
Plot the LS regression line on the scatterplot.

$$\hat{y} = 11 - 3x$$



$$\hat{y} = 11 - 3x$$





If the correlation between two variables is $r=-0.7$, how well does the LS regression line explain the data?

$$r = -0.7$$

$$r^2 = (-0.7)^2 = 0.49$$

so the LS reg line explains 49% of the variability.

Let x represent time since midnight and y represent the number of cars on the road

Given the LS reg line:

$$\hat{y} = 1200 + 300x$$

Find and interpret the slope and intercept.

$$\hat{y} = 1200 + 300x$$

slope is 300. The number of cars on the road increases by 300 every hour.

intercept is 1200. At midnight there are 1200 cars on the road.

Find the equation of the LS regression line given,

$$\bar{x} = 1.304$$

$$\bar{y} = 1.878$$

$$s_x = 3.392$$

$$s_y = 7.554$$

$$r = 0.5251$$

$$\bar{x} = 1.304, \bar{y} = 1.878, s_x = 3.392, s_y = 7.554, r = 0.5251$$

$$\hat{y} = a + bx, \quad b = r\frac{s_y}{s_x}, a = \bar{y} - b\bar{x}$$

$$b = r\frac{s_y}{s_x} = (0.5251)\frac{(7.554)}{(3.392)} = 1.169$$

$$a = \bar{y} - b\bar{x} = 1.878 - (1.169)(1.304) = 0.354$$

$$\hat{y} = a + bx = 0.354 + 1.169x$$

A baseball enthusiast believes pitchers who strike out a lot of batters also walk a lot of batters. He reached this conclusion by going to the library and examining the records of all major-league pitchers between 1990 and 1995. What type of study is his decision based on?

An observational study
(from available data)

In a recent study, children in grades 2 through 4 were randomly assigned to groups and each group was given a different amount of homework. The study showed a significant *negative* relationship between the amount of homework assigned and student attitudes.

What type of study is this?

What type of variable is "amount of homework"

This is an experiment.

The "amount of homework" is the explanatory variable.

Will a fluoride mouthwash used after brushing reduce cavities? Twenty sets of twins were used to investigate this question. One member of each set of twins used the mouthwash after each brushing, the other did not. After 6 months, the difference in the number of cavities of those using the mouthwash was compared with the number of cavities of those who did not use the mouthwash. What type of design does this experiment use?

matched pairs design

The head of the quality control department at a publishing company is studying the effect of type of glue and type of binding on the strength of the bookbinding. The company has three possible glues to choose from and the book can either be bound as a paperback or a hardback.

What are the factors?

How many treatments are there?

What are the factors?

- type of glue (3 types)
- type of binding (2 types)

How many treatments are there?

- 6

Should you have a cup of coffee to make you more alert when studying for a big test? A study on the effect of caffeine involved asking volunteers to take a memory test 20 minutes after drinking cola. Some volunteers were randomly assigned to drink caffeine-free cola; some to drink regular cola (with caffeine), and others a mixture of the two (getting a half dose of caffeine). For each volunteer, a test score (the number of items recalled correctly) was recorded. The volunteers were not told which type of cola they had been given, but the researchers for the study prepared the cups of cola right on the spot (out of sight of the volunteers). Give at least 2 of the basic principles of statistical design were used in this study?

Which of the basic principles of statistical design were used in this study?

Control, randomization, repetition, and blinding.

A news program is doing a report on a controversial court case. Daniel VanHowen is being prosecuted for killing a woman in the park, but all the evidence is circumstantial. A behind the scenes at this news program, decides to do a small opinion poll among the staff members. He walks around and asks everybody he can find: “Don’t you agree that Mr. VanHowen is guilty?”

What type of bias would this question lead to?

response bias

A news program is doing a report on a controversial court case. Daniel VanHowen is being prosecuted for killing a woman in the park, but all the evidence is circumstantial. The news program decides to poll its viewers. They post a question on their web site and ask the TV viewers to respond.

Why would the results of this poll be questionable?

they are using voluntary response
-possible undercoverage
-people who feel strongly will respond

David A. Miller owns a small advertising business. He has nine employees. The names of the employees are given below.

1. Becker 2. Brown 3. Chasten
4. Ito 5. Kiefer 6. Spitzer
7. Taylor 8. Walt 9. Weiss

Use the list of random digits below to select a simple random sample of three names from the list of employees.

11793 20495 05907 11384 44982 20751

1, 7, 9 so
Becker, Taylor, Weiss

A free cholesterol screening program is set up in the downtown area during the lunch hour. Individuals can walk in and have their cholesterol measured for free. One hundred seventy-three people use the service, and their average cholesterol is 217.8.

Is 217.8 a parameter or a statistic? Why?

217.8 is a statistic because it describes the sample

How can bias and variability be reduced?

To reduce bias, use random sampling.

To reduce variability, use a larger sample.

Event A occurs with probability 0.2. Event B occurs with probability 0.8. If A and B are disjoint (mutually exclusive), then which of the following is correct?

a) P(A and B) = 0.16
b) P(A or B) = 1.0
c) P(A and B) = 1.0
d) P(A or B) = 0.16

b) $P(A \text{ or } B) = 1.0$

Ignoring twins and other multiple births, assume babies born at a hospital are independent events with the probability that a baby is a boy and the probability that a baby is a girl both equal to 0.5.

What is the probability that the next three babies are of the same sex?

$$P(G) = 0.5, P(B) = 0.5$$

$$P(\text{next three are same sex}) = P(GGG \text{ or } BBB)$$

$$= P(GGG) + P(BBB) \quad \text{because disjoint}$$

$$= P(G)P(G)P(G) + P(B)P(B)P(B)$$

$$= (0.5)(0.5)(0.5) + (0.5)(0.5)(0.5)$$

$$= 0.125 + 0.125$$

$$= 0.25$$

Ignoring twins and other multiple births, assume babies born at a hospital are independent events with the probability that a baby is a boy and the probability that a baby is a girl both equal to 0.5.

Define event B = {at least one of the next two babies is a boy}. What is the probability of the complement of event B?

$$P(boy) = 0.5, P(girl) = 0.5$$

$$P(\mathrm{B}^{\mathrm{C}}) = P(\text{neither of next two babies are boys})$$

$$= P(\text{girl and girl}) = P(girl)P(girl)$$

$$= (0.5)(0.5)$$

$$= 0.25$$

Ignoring twins and other multiple births, assume babies born at a hospital are independent events with the probability that a baby is a boy and the probability that a baby is a girl both equal to 0.5.

What is the probability that at least one of the next three babies is a boy?

$$P(B) = 0.5, P(G) = 0.5$$

$$P(\text{at least one of next 3 are boys})$$

$$= 1 - P(\text{none of next 3 are boys})$$

$$= 1 - P(GGG)$$

$$= 1 - P(G)P(G)P(G)$$

$$= 1 - (0.5)(0.5)(0.5)$$

$$= 1 - 0.125$$

$$= 0.875$$

Suppose that a college determines the following distribution for X = number of courses taken by a full-time student this semester.

Value of X	3	4	5	6
Probability	0.07		0.25	0.28

What is the probability for $X = 4$?
What is $P(X > 4.7)$?

What is the probability for $X = 4$?

$$P(X=4) = 1 - P(X=3 or X=5 or X=6)$$
$$= 1 - (P(X=3) + P(X=5) + P(X=6))$$
$$= 1 - (0.07 + 0.25 + 0.28)$$
$$= 1 - (0.6)$$
$$= 0.4$$

What is $P(X > 4.7)$?

$$P(X > 4.7) = P(X=5 or X=6)$$
$$= P(X=5) + P(X=6)$$
$$= 0.25 + 0.28$$
$$= 0.53$$

A commuter must pass through 5 traffic lights on her way to work. Let X = the number of red lights she stops at on her way to work. She estimates the distribution for X to be as shown below.

Value of X	1	2	3	4	5
Probability	0.40	0.25	0.15	0.15	0.05

Find $P(2 < X \leq 4)$

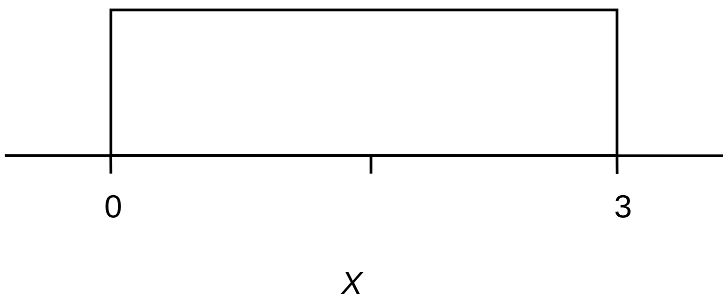
$$P(2 < X \leq 4) = P(X = 3 \text{ or } X = 4)$$

$$= P(X = 3) + P(X = 4)$$

$$= 0.15 + 0.15$$

$$= 0.3$$

Given the uniform density function



find,

a) $P(X = 2.5)$

b) $P(1.2 < X < 2.3)$

a) $P(X = 2.5) = 0$

b) $P(1.2 < X < 2.3) = \left(\frac{1}{3}\right)(2.3 - 1.2)$

$= \left(\frac{1}{3}\right)(1.1)$

$= 0.367$

Let the random variable X represent money spent at a store on a randomly selected day. Assume X is normal with a mean of $420 and a standard deviation of $25.

Find the probability that the amount of money spent is more than 400.

$$P(X > 400) = P\left(Z > \frac{400 - 420}{25} \right)$$

$$= P(Z > -0.8)$$

$$= 1 - P(Z < -0.8)$$

$$= 1 - 0.2119$$

$$= 0.7881$$