

1. Scientists measured the annual forest loss (in square kilometers) in Indonesia from 2000 to 2013. They found the regression line **forest loss = 7500+(1021 x year since 2000)** for predicting forest loss in square kilometers from years since 2000.

What is the slope of this line? Say in words what the numerical value of the slope tells you.

2. TRUE or FALSE: Given a regression line with a small slope, it means that the correlation between the two variables is weak.
3. Researchers measured the percent body fat and the preferred amount of salt level for several children. It shows  $\hat{y} = 24.2 + 6.0x$  where  $x$  represents the preferred salt level and  $y$  is the percent body fat.
  - a. Find the predicted amount of body fat for a child when her preferred amount of salt level is 0.6.
  - b. Assume that one of the children surveyed has body fat of 38 with the salt level 0.6. What would be the residual (prediction error) when the regression line was used?
4. A call-in poll conducted by a local radio station concluded that race would not be an issue in the 2008 presidential election. This conclusion was based on data collected from 450 calls made by local listeners. The sampling technique being used is
  - a. simple random sampling.
  - b. stratified random sampling.
  - c. volunteer sampling.
  - d. multistage sampling.
5. The Excite Poll is an online poll at [poll.excite.com](http://poll.excite.com). You click on an answer to become part of the sample. One poll question was "Do you prefer watching first-run movies at a movie theater, or waiting until they are available on home video or pay-per-view?" A total of 8896 people responded with 1118 saying they preferred theaters. From this survey you can conclude that
  - a. Americans prefer watching movies at home.
  - b. a larger sample is necessary.
  - c. the poll uses voluntary response, so the results tell us little about the population of all adults.
  - d. movie theaters should lower their prices.

6. The following table describes the opinions of the 570 people that returned the questionnaire in the survey on their opinion of campus residence quality (high quality, medium quality, low quality).

Class	High	Medium	Low	Total
Freshman	65	25	20	110
Sophomore	55	30	45	130
Junior	60	40	70	170
Senior	30	60	70	160
Total	210	155	205	570

- Find the conditional probability distribution of the opinions among seniors in the college?
  - Is the opinion more likely to be High among Freshman or sophomores?
  - If a student is selected at random, find  $P(\text{Senior AND Medium})$
  - If a student is selected at random, find  $P(\text{Senior} \mid \text{Medium})$
  - If a student is selected at random, find  $P(\text{Senior OR Medium})$
7. A public opinion poll in Ohio wants to determine whether registered voters in the state approve of a measure to ban smoking in all public areas. The researchers select a simple random sample of 50 registered voters from each county in the state and ask whether they approve or disapprove of the measure. This is an example of
- a systematic county sample.
  - a stratified sample.
  - a multistage sample.
  - a simple random sample.
8. I toss a penny and observe whether it lands heads up or tails up. Suppose the penny is fair, i.e., the probability of heads is  $1/2$  and the probability of tails is  $1/2$ . This means
- every occurrence of a head must be balanced by a tail in one of the next two or three tosses.
  - if I flip the coin many, many times, the proportion of heads will be approximately  $1/2$ , and this proportion will tend to get closer and closer to  $1/2$  as the number of tosses increases. (Law of Large Numbers)
  - regardless of the number of flips, half will be heads and half tails.
  - All of the above

9. I roll a four-sided die. The possible outcomes are 1, 2, 3, or 4, depending on the number of spots on the side of the die that is face down. This collection of all possible outcomes is called
- a census.
  - the probability.
  - the sample space.
  - the distribution.

For Questions 10-12, assume that event A occurs with probability 0.4 and event B occurs with probability 0.5. Assume that A and B are disjoint events.

10. The probability that either event occurs (A or B) is
- 0.0
  - 0.7
  - 0.9
  - 1.0

11. The probability that both events occur (A and B) is
- 0.0
  - 0.2
  - 0.7
  - 1

12. If A and B are independent events, the probability that both events occur (A and B) is \_\_\_\_\_

13. The density curve for a continuous random variable  $X$  has which of the following properties?
- The probability of any event is the area under the density curve and above the values of  $X$  that make up the event.
  - The total area under the density curve for  $X$  must be exactly 1.
  - The probability of any event of the form  $X = \text{constant}$  is 0.
  - All of the above

14. The amount of milk sold each day by a grocery store varies according to the Normal distribution with mean 130 gallons and standard deviation 12 gallons. On a randomly selected day, the probability that the store sells at least 154 gallons is
- 0.0228
  - 0.1587
  - 0.8413
  - 0.9772

Nationwide, the amount charged by doctors for performing a particular minor surgical procedure averages \$1220 and varies with a standard deviation of \$300.

15. Find the probability of a randomly selected bill exceeding \$1500.
16. Find the probability of a randomly selected bill being between \$800 and \$1300?
17. Event  $A$  occurs with probability 0.1. Event  $B$  occurs with probability 0.6. If  $A$  and  $B$  are independent, then
- $P(A \text{ and } B) = 0.70$ .
  - $P(A \text{ or } B) = 0.64$ .
  - $P(A \text{ and } B) = 0.64$ .
  - $P(A \text{ or } B) = 0.70$ .
18. An event  $A$  will occur with probability 0.5. An event  $B$  will occur with probability 0.6. The probability that both  $A$  and  $B$  will occur is 0.1. The conditional probability of  $B$ , given  $A$ , is
- $5/6$ .
  - $1/5$ .
  - $1/6$ .
  - It cannot be determined from the information given.

A manufacturing process produces bags of cookies. The distribution of content weights of these bags is Normal with mean 16.0 oz and standard deviation 0.8 oz.

19. Find the weight of a bag of cookies so that it is in the top 10%
20. An event  $A$  will occur with probability 0.5. An event  $B$  will occur with probability 0.6. The probability that both  $A$  and  $B$  will occur is 0.1. We may conclude
- events  $A$  and  $B$  are independent.
  - events  $A$  and  $B$  are disjoint.
  - either  $A$  or  $B$  always occurs.
  - None of the above

21. What is the difference between an experiment and an observational study?

22. You flip a coin and roll a 6 sided die. Show the sample space

23. If a coin is flipped 5 times, then find the probability that at least 1 flip results in a TAIL

24. Create a scatterplot for the table below:

x	1	2	4	6	8
y	4	9	20	33	36

25. What are the features of a density curve that follows a normal distribution?

26. Consider the probability distribution shown in the table below:

x	0	1	2	3	4
P(x)	.4	.25	.15	.15	.05

a) Create a histogram for the distribution of x

b) Find the probability of an x value greater than 1

c) The probability of event B is found with  $P(B) = 1 - .05$  What is event B?

d) Find the expected value for x:

27. In a research experiment involving testosterone levels, a researcher recommends splitting the subjects up based on weight. This is called \_\_\_\_\_ and is done because

\_\_\_\_\_

28. What would this look like if the experiment on the group of male subjects ran as a matched pair study?

1. Scientists measured the annual forest loss (in square kilometers) in Indonesia from 2000 to 2013. They found the regression line  $\text{forest loss} = 7500 + (1021 \times \text{year since 2000})$  for predicting forest loss in square kilometers from years since 2000.

a. What is the slope of this line? Say in words what the numerical value of the slope tells you.

1021 As years inc 1  
forest loss inc 1021 Km<sup>2</sup>

b. ~~If we measure forest loss in square meters per year, what would the slope be? Note that there are 10<sup>6</sup> square meters in a square kilometer.~~

c. ~~If we measure forest loss in thousands of square kilometers per year, what would the slope be?~~

2. TRUE or FALSE: Given a regression line with a small slope, it means that the correlation between the two variables is weak.

Direction

Strength



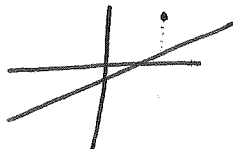
3. Researchers measured the percent body fat and the preferred amount of salt level for several children. It shows  $\hat{y} = 24.2 + 6.0x$  where  $x$  represents the preferred salt level and  $y$  is the percent body fat.

a. Find the predicted amount of body fat for a child when her preferred amount of salt level is 0.6.

$$\hat{\text{fat}} = 24.2 + 6(\text{salt})$$

$$= 27.8\%$$

b. Assume that one of the children surveyed has body fat of 38 with the salt level 0.6. What would be the residual (prediction error) when the regression line was used?



prediction = 27.8

residual = real - pred.

$$y - \hat{y}$$

$$38 - 27.8$$

10.2

4. The following table describes the opinions of the 570 people that returned the questionnaire in the survey on their opinion of campus residence quality (high quality, medium quality, low quality).

Class	High	Medium	Low	Total
Freshman	65	25	20	110
Sophomore	55	30	45	130
Junior	60	40	70	170
Senior	30	60	70	160
Total	210	155	205	570

- a. Find the conditional probability distribution of the opinions among seniors in the college?

- b. Is the opinion more likely to be High among Freshman or sophomores?

Fresh

$$\frac{65}{110} = .59$$

$$\frac{55}{130} = .42$$

$$\frac{30}{160} = .1875$$

$$\frac{60}{160} = .375$$

$$\frac{70}{160} = .4375$$

- c. If a student is selected at random, find P(Senior AND Medium)

$$\frac{60}{570} = .105$$

- d. If a student is selected at random, find P(Senior | Medium)

$$\frac{60}{155} = .387$$

- e. If a student is selected at random, find P(Senior OR Medium)

$$\frac{255}{570} = .447$$

5. A call-in poll conducted by a local radio station concluded that race would not be an issue in the 2008 presidential election. This conclusion was based on data collected from 450 calls made by local listeners. The sampling technique being used is

- simple random sampling.
- stratified random sampling.
- volunteer sampling.
- multistage sampling.

6. The Excite Poll is an online poll at poll.excite.com. You click on an answer to become part of the sample. One poll question was "Do you prefer watching first-run movies at a movie theater, or waiting until they are available on home video or pay-per-view?" A total of 8896 people responded with 1118 saying they preferred theaters. From this survey you can conclude that

- Americans prefer watching movies at home.
- a larger sample is necessary.
- the poll uses voluntary response, so the results tell us little about the population of all adults.
- movie theaters should lower their prices.

7. A public opinion poll in Ohio wants to determine whether registered voters in the state approve of a measure to ban smoking in all public areas. The researchers select a simple random sample of 50 registered voters from each county in the state and ask whether they approve or disapprove of the measure. This is an example of
- a. a systematic county sample.  
 b. a stratified sample.  
 c. a multistage sample.  
 d. a simple random sample.
- Clusters  
SRS
8. I toss a penny and observe whether it lands heads up or tails up. Suppose the penny is fair, i.e., the probability of heads is  $1/2$  and the probability of tails is  $1/2$ . This means
- a. every occurrence of a head must be balanced by a tail in one of the next two or three tosses.  
 b. if I flip the coin many, many times, the proportion of heads will be approximately  $1/2$ , and this proportion will tend to get closer and closer to  $1/2$  as the number of tosses increases. (Law of Large Numbers)  
 c. regardless of the number of flips, half will be heads and half tails.  
 d. All of the above
9. I roll a four-sided die. The possible outcomes are 1, 2, 3, or 4, depending on the number of spots on the side of the die that is face down. This collection of all possible outcomes is called
- a. a census.  
 b. the probability.  
 c. the sample space.  
 d. the distribution.

For Questions 10-12, assume that event A occurs with probability 0.4 and event B occurs with probability 0.5. Assume that A and B are disjoint events.

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

$$.4 + .5 - \underbrace{0}$$

10. The probability that either event occurs (A or B) is
- a. 0.0  
 b. 0.7  
 c. 0.9  
 d. 1.0

11. The probability that both events occur (A and B) is
- a. 0.0  
 b. 0.2  
 c. 0.7  
 d. 1

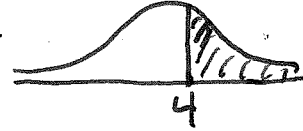
$$(.4)(.5) = .2$$

$$P(A \text{ and } B) = P(A) \cdot P(B)$$

12. If A and B are independent events, the probability that both events occur (A and B) is .2



13. The density curve for a continuous random variable  $X$  has which of the following properties?
- The probability of any event is the area under the density curve and above the values of  $X$  that make up the event.
  - The total area under the density curve for  $X$  must be exactly 1.
  - The probability of any event of the form  $X = \text{constant}$  is 0.
  - All of the above



14. The amount of milk sold each day by a grocery store varies according to the Normal distribution with mean 130 gallons and standard deviation 12 gallons. On a randomly selected day, the probability that the store sells at least 154 gallons is
- 0.0228
  - 0.1587
  - 0.8413
  - 0.9772

Mean 130    St D. 12    Value 154 and up  
 .023

15. The incomes in a certain large population of college teachers have a normal distribution with mean \$75,000 and standard deviation \$10,000. 16 teachers are selected at random from this population to serve on a committee. What is the probability that their average salary is more than \$77,500?
- 0.0228
  - 0.1587
  - 0.8413
  - Essentially 0

In a large population of college-educated adults, the mean IQ is 112 with standard deviation 25. Suppose 300 adults from this population are randomly selected for a market research campaign.

16. The distribution of the sample mean IQ is
- approximately Normal, mean 112, standard deviation 25.
  - approximately Normal, mean 112, standard deviation 1.443.
  - approximately Normal, mean 112, standard deviation 0.083.
  - approximately Normal, mean equal to the observed value of the sample mean, standard deviation
17. The probability that the sample mean IQ is greater than 115 is
- 0.019
  - 0.452
  - 0.528
  - 0.981

Use the following to answer Questions 19 and 20.

Nationwide, the amount charged by doctors for performing a particular minor surgical procedure averages \$1220 and varies with a standard deviation of \$300.

18. Find the probability of a randomly selected bill exceeding \$1500.

Mean 1220    St. D. 300  
 1 val .175

19. Find the probability of a randomly selected bill being between \$800 and \$1300?

2 val. .524

20. Event A occurs with probability 0.1. Event B occurs with probability 0.6. If A and B are independent, then

- a.  $P(A \text{ and } B) = 0.70$ .
- b.  $P(A \text{ or } B) = 0.64$ .
- c.  $P(A \text{ and } B) = 0.64$ .
- d.  $P(A \text{ or } B) = 0.70$ .

$$P(A \cap B) = P(A) \cdot P(B) = (.1)(.6) = .06$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$.1 + .6 - .06 = .64$$

21. An event A will occur with probability 0.5. An event B will occur with probability 0.6. The probability that both A and B will occur is 0.1. The conditional probability of B, given A, is

- a.  $5/6$ .
- b.  $1/5$ .
- c.  $1/6$ .
- d. It cannot be determined from the information given.

$$P(B|A) = \frac{.1}{.5} = .2$$

	A	-A	
B	.1	.5	.6
-B	.4	0	.4
	.5	.5	1

A manufacturing process produces bags of cookies. The distribution of content weights of these bags is Normal with mean 16.0 oz and standard deviation 0.8 oz.

22. Find the weight of a bag of cookies so that it is in the top 10%

$$\text{Area Right} = .1 \quad 17.03 \text{ oz}$$

23. An event A will occur with probability 0.5. An event B will occur with probability 0.6. The probability that both A and B will occur is 0.1. We may conclude

- a. events A and B are independent.
- b. events A and B are disjoint.
- c. either A or B always occurs.
- d. None of the above

If indep.  $(.5)(.6) = .1$  NO  
 If disjoint  $A \cap B = 0$  NO  
 $.5 + .6 - .1 = 1$  YES

	A	-A	
B	.1	.5	.6
-B	.4	0	.4
	.5	.5	1

24. What is the difference between an experiment and an observational study?

↓  
 treatment  
 placebo

↓  
 survey

25. You flip a coin and roll a 6 sided die. Show the sample space

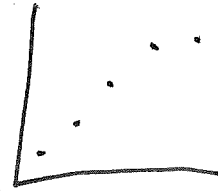
H1 H2 H3 H4 H5 H6  
 T1 T2 T3 T4 T5 T6

26. If a coin is flipped 5 times, then find the probability that at least 1 flip results in a TAIL

Opposite: All are Heads  
 $(\frac{1}{2})^5 = .03125$   
 $1 - .03125 = .96875$

27. Create a scatterplot for the table below:

x	1	2	4	6	8
y	4	9	20	33	36



28. What are the features of a density curve that follows a normal distribution?

Area = 1  
Symmetric  
Bell shaped

Centered at Mean  
Mean = median  
~~95-99-99.8 Rule~~  
68-95-99.6 Rule

29. Consider the probability distribution shown in the table below:

x	0	1	2	3	4
P(x)	.4	.25	.15	.15	.05

a) Create a histogram for the distribution of x



b) Find the probability of an x value greater than 1

$$.15 + .15 + .05 = .35$$

c) The probability of event B is found with  $P(B) = 1 - .05$  What is event B?

Not 4  
value less than 4

$$1 - .05 = .95$$

d) Find the expected value for x:

$$0(.4) + 1(.25) + 2(.15) + 3(.15) + 4(.05)$$

$$1.2$$

30. In a research experiment involving testosterone levels, a researcher recommends splitting the subjects up based on weight. This is called Blocking and is done because

to reduce variation. We expect variation based on weight

31. What would this look like if the experiment on the group of male subjects ran as a matched pair study?

each person is their own control  
1 time get treatment, 1 time get placebo 6