33. Find $P(x = 2)$ $= 0$

34. Find the probability that a student takes $\geq$ *at least* five years to earn a Bachelor of Science (B.S.) degree.

35. Find the probability that a student takes $\leq$ *not more than* four years to earn a Bachelor of Science (B.S.) degree.

36. Find the probability that a student takes $\leq$ *at most* five years to earn a Bachelor of Science (B.S.) degree.

37. Find the probability that a student takes $>$ *more than* four years to earn a Bachelor of Science (B.S.) degree.

38. Find the probability that a student takes $\geq$ *no less than* five years to earn a Bachelor of Science (B.S.) degree.

39. Find the probability that a student $\leq$ *does not exceed* six years to earn a Bachelor of Science (B.S.) degree.

40. Find the probability that a student earns a Bachelor of Science (B.S.) degree in *under* four years. $<$

41. Find the probability that a student takes $>$ *over* 4 years to earn a Bachelor of Science (B.S.) degree.

42. Find the probability that a student takes $<$ *fewer than* 5 years to earn a Bachelor of Science (B.S.) degree.

Outliers

To determine if a point is an outlier, do the following:

1. Input the following equations into the TI 83, 83+,84, 84+:

   $y_1 = a + bx$

   $y_2 = (a + bx) + 2s$ where s is the standard deviation of the residuals

   $y_3 = (a + bx) - 2s$

   If any point is above y_2 or below y_3 then the point is considered to be an outlier.

A point is an outlier if it is more than 2 st. dev lengths above or below the regression line

Note: The ~~calculator function~~ LinRegTTest (STATS TESTS LinRegTTest) calculates s.

S = the st dev of the regression line = a measure of how far a typical point will be above or below the regression line

FORMULA REVIEW

Linear Equations

$y = a + bx$ where a is the y-intercept and b is the slope. The variable x is the independent variable and y is the dependent variable.

Least Squares Line or Line of Best Fit:

$$\hat{y} = a + bx$$

a = y-intercept

b = slope

Standard deviation of the residuals:

$$s = \sqrt{\frac{SSE}{n-2}}$$

where

SSE = sum of squared errors $(\hat{y}_i - y_i)^2$

n = the number of data points

① Mult. Rule

$$P(\underbrace{A \text{ and } B}_{\text{A and B are occurring in a sequence}}) = P(A) \cdot P(B)$$

② Addn Rule

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

(for mutully exclusive events)

③ Addn Rule

$$P(A \text{ or } B) = P(A) + P(B) - P(\overbrace{A \text{ and } B}^{\text{A and B are occ. simult.}})$$

A and B not M.E. (they can occur simultaneously)

④ conditional probability

occurs when the question uses the phrase "given that"

27 (i) The least squares line <u>cannot</u> be used to predict the area of a 51^{st} state.

- Since $x = 51$ is outside the range of the x-data ($min = 3$, $max = 50$).

(1) $\frac{220}{450} = \boxed{0.4\overline{8}}$

(2) $\frac{125}{450} = \boxed{0.2\overline{7}}$

(3) $\frac{26}{61} \doteq \boxed{0.426}$

(4) A: subject left ICU after 2 weeks

B: subject left ICU after 3 weeks

A and B are mutually exclusive, so we use

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

$$= \frac{105}{450} + \frac{220}{450}$$

$$= 325/450 = \boxed{0.7\overline{2}}$$

(5) A: subject left ICU after 2 weeks

B: subject was treated with treatment 2

A and B are <u>not</u> mutually exclusive events, so we use

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

$$= \frac{105}{450} + \frac{234}{450} - \frac{44}{450}$$

$$= \frac{105+234-44}{450} = \frac{295}{450}$$

$$= \boxed{0.6\overline{5}}$$

(6) $\frac{44}{234} \doteq \boxed{0.188}$

(7) $\frac{75+45}{155} = \frac{120}{155}$

$$\doteq \boxed{0.774}$$

(8) "didn't leave after 2 weeks" means that the subject ~~left~~ in either 3 or 4 weeks.

A = the selected subject left after 3 weeks

B: the selected subject left after 4 weeks

Since A and B are mutually exclusive events, we use

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

$$= \frac{75}{155} + \frac{45}{155}$$

$$= \frac{120}{155} \doteq \boxed{0.774}$$

the denominator is 155 since the phrase "given that" in the statement of the question indicates that the selected subject was given treatment 1

(9) $\boxed{\begin{array}{l} x = \text{variable 2} \\ y = \text{variable 1} \end{array}}$

(10) predictor variable $= x = \boxed{\text{Variable 2}}$

response variable $= y = \boxed{\text{Var. 1}}$

(11) As x increases, y tends to decrease, therefore I expect a $\boxed{\text{negative correlation}}$

(12) As x (the interest rate) increases, y tends to decrease, so I expect a $\boxed{\text{negative correlation}}$

(13) Assuming that x = height and that y = weight,

It seems like as x (ht.) increases, y (wt.) increases, So I expect a $\boxed{\text{positive correlation}}$ between height and wait. The taller a person is, the more they would tend to weigh.

14 As x (min. temp) increases, y (heating) cost tends to decrease, So I expect a $\boxed{\text{negative}}$ Correlation

15 x = amount of time that the customer wants to rent the bike for. ← independent variable

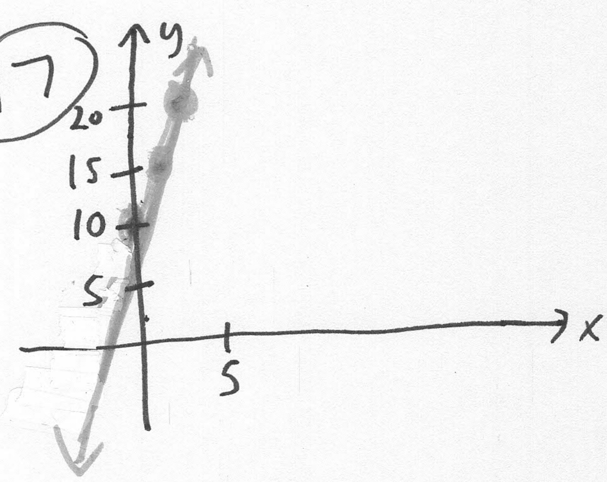
y = the cost to rent ← dependent variable

16

x	y
0	$10+5(0)=10$
1	$10+5\cdot(1)=15$
2	$10+5(2)=20$
3	$10+5(3)=25$
4	$10+5(4)=30$
5	$10+5(5)=35$
⋮	⋮
x	$10+5x$

$$\boxed{y = 10+5x}$$

17



(18) The y-int is 10, since when $x=0$, $y=\$10$.
The y-int represents the up front fee.

(19) the slope is 5, or more importantly, is $\$5/1\text{hr}$ or \$5 per hour.

The slope is the average rate of change of y (cost) for every 1 unit increase (1hr) in time.

(20) ind. var. = x = year
dep. var = y = amount of soil lost in pounds

(21) When $x=1$, $y=14{,}000$, so $\boxed{\text{14,000 pounds of shoreline}}$

(22) The y-int is 0. or zero pounds.
When $x=0$, the time is when the soil starts eroding

(23) Since $y=15-1.5x$ is equivalent to $y=15+(-1.5)x$ the slope is $\boxed{-1.5}$, or more importantly, $\frac{-\$1.5}{1\text{hr}}$, the rate at which the stock price is dropping; i.e., \$1.50 per hour.

(24) y-int = $15

this represents the stock price at the start/opening of the stock exchange that day.

(25) $y = 101.32 + 2.48x$

when $x = 60$,

$y = 101.32 + 2.48(60)$

$= \boxed{250.12 \text{ sales}}$

(26) $y = 101.32 + 2.48(90)$

$= 324.52$ sales

(27) (a) Let rank be the independent variable and area be the dependent variable.

(b) see attached graph

(c) There appears to be a linear relationship, with one outlier.

(d) $\boxed{\hat{y} = 24177.06 + 1010.478x}$

(e) $r = 0.50047$

There is a moderate, positive correlation between rank and area.

27

f Alabama: 46,407.576
Colorado: 62,575.224

g If the outlier is removed, there is a linear relationship

h Hawaii is an outlier since it is more than 2 standard deviations below the least squares (regression) line. (See attached graph)

i rank: 51 (x)
area: 75,711 mi^2 (y)

j $\hat{y} = -87{,}065.3 + 7{,}828.532x$

k ↗

l Alabama: 85,162.404; The prior estimate was closer. Alaska is an outlier!

m yes, with the exception of Hawii

n Question Interpret the meaning of the standard deviation of the least squares (regression) line.

Answer: The standard deviation is a measure of how far a typical point will be above or below the least squares line.

(28) x represents the number of years it takes a person to earn a B.S.

(29) $\eta = \boxed{4.85 \text{ yrs}}$

(30) $\boxed{0.30}$

(31) $P(x \le 5) =$

$= P(3) + P(4) + P(5)$

$= 0.05 + 0.40 + 0.30$

$= \boxed{0.75}$

(32) $P(x > 3) = P(x \ge 4)$

$= P(4) + P(5) + P(6) + P(7)$

$= 0.40 + 0.30 + 0.15 + 0.1$

$= \boxed{0.95}$

(33) 0; no one gets the degree in 2 yrs.

(34) $P(x \ge 5) = P(5) + P(6) + P(7$

$= 0.30 + 0.15 + 0.10$

$= \boxed{0.55}$

(35) $P(x \le 4)$

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

$= 0.05 + 0.40$

$= \boxed{0.45}$

(36) $P(x \le 5)$

$= P(3) + P(4) + P(5)$

$= 0.05 + 0.40 + 0.30$

$= \boxed{0.75}$

37 $P(x > 4)$

$= P(x \geq 5)$

$= P(5) + P(6) + P(7)$

$= 0.30 + 0.15 + 0.10$

$= \boxed{0.55}$

38

$= P(x \geq 5)$

$= \boxed{0.55}$

39 $P(x \leq 6)$

$= P(3) + P(4) + P(5) + P(6)$

$= \ldots$

$= \boxed{0.90}$

41 $P(x > 4)$

$= P(x \geq 5)$

$= 0.55$

40 $P(x < 4)$

$= P(x \leq 3)$

$= P(3) = \boxed{0.05}$

42 $P(x < 5)$

$= P(x \leq 4)$

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

$= 0.05 + 0.40$

$= \boxed{0.45}$

43

$\mu = 4.85$

$\sigma = 1.06$