4.1

$4(1/4) = 1$

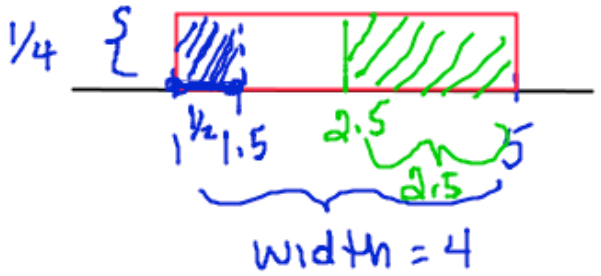
4. Consider a uniform distribution that is defined for $1 \le X \le 5$.

a. Sketch the distribution.

b. What is the probability that X falls below 1.5? $(1/2)(1/4) = 1/8$

c. What is the probability that X lies above 2.5? $(2.5)(1/4) = 5/8$

d. What is the probability that X lies below 1? 0



ehw7

r r close to 1 or close to -1

5. Generally, the larger the correlation (either positive or negative) between the variables for the simple linear regression model, the (better, worse) _____ will be the predictions of y for given values of *x*.
   a. Better
   b. Worse

r^2

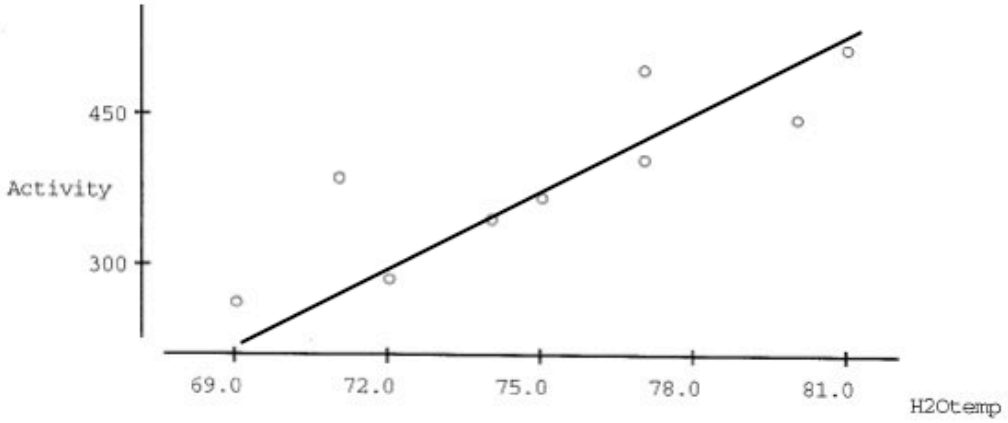
8. True or False: The coefficient of determination can assume negative values.
   a. True
   b. False

14. The correlation between two scores X and Y equals 0.8. If both the X scores and the Y scores are converted to z-scores, then the correlation between z-scores for X and z-scores for Y would be

$r = .8$

then they should still have correlation if both converted w/ same formula

15. Joey read in his biology book that fish activity increases with water temperature, and he decided to investigate this issue by conducting an experiment. On nine successive days, he measures fish activity and water temperature in his aquarium. Larger values of his measure of fish activity denote more activity. The scatter plot of his data is below:



What does the scatter plot reveal?

a. A strong negative linear association
b. A strong positive linear association
c. A weak negative linear association
d. A weak positive linear association
e. No linear association

5.3

12. Suppose that you are given the following results. Find the correlation coefficient of the data

$b = .222 \quad s_x = 6.32 \quad s_y = 1.45$

$b = r\left(\frac{s_y}{s_x}\right)$

$.222 = r\left(\frac{1.45}{6.32}\right)$

Solve for r

5.2

2. The number of motor vehicles registered (in millions) in the U.S. has grown as follows:

	0	5	10	15	20	25	30	35	40	45
Year	1940	1945	1950	1955	1960	1965	1970	1975	1980	1985
Vehicles	32.4	31.0	49.2	62.7	73.9	90.4	108.4	132.9	155.8	171.7

a. Present a scatter plot for the data. (Hint: let $x = 0$ represent the year 1940)

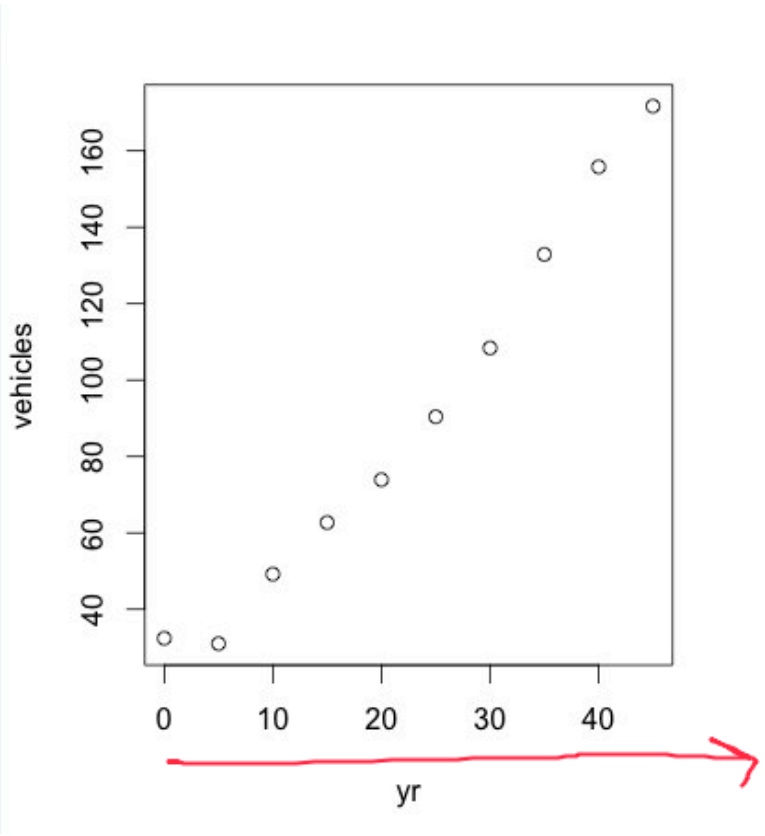
b. Determine the correlation coefficient for the data and interpret the value.

explanatory - x - year

response - y - vehicles

```
> vehicles
 [1]  32.4  31.0  49.2  62.7  73.9  90.4 108.4 132.9 155.8 171.7
> yr=c(0,5,10,15,20,25,30,35,40,45)
> yr
 [1]  0  5 10 15 20 25 30 35 40 45
> plot(yr,vehicles)
> cor(yr,vehicles)
[1] 0.9874813
>
```

r



5.3

18. The table below shows the Men's 800 Meter Run World Records.

	Year	Record
0	1925	111.9
10	1935	109.7
20	1945	106.6
30	1955	105.7
40	1965	104.3
50	1975	104.1
60	1985	101.73
70	1995	101.73

a. Make a scatter-plot of the data. ✓
b. Compute the LSRL. →
c. Provide an interpretation of the slope of this line in the context of these data.
d. Find the correlation coefficient for the relationship. Interpret this number. r
e. Find the coefficient of determination for the relationship. Interpret this number. r^2

For every 1 increase in <year> there is <value of slope> incr/decr of <y-value>.

x - year
y - record

TI calc.
L_1 L_2 ← enter lists
(x) (y)
make sure DiagnosticOn
Stat - Calc - 4 or 8
a + bx

R-Studio
- year = c(0, 10, 20, ...
- record = c(111.9, 109.7, ...
- plot(year, record)
- lm(record ~ year)

```
Call:
lm(formula = vehicles ~ yr)

Coefficients:
(Intercept)          yr
```

x variable name

y-int. Slope

$y = mx + int$

cor(year, record)
x y

5.1

10. The weights of children in the Egyptian village of Nahya were recorded. Here are the mean weights of the 170 children in that village:

Age(months)	1	2	3	4	5	6	7	8	9	10	11	12
Weight (kg)	4.3	5.1	5.7	6.3	6.8	7.1	7.2	7.2	7.2	7.2	7.5	7.8

a. Make a scatter plot of mean weight against age.
b. Is there a positive or negative association between the two variables? If so, which?
c. Does this data describe a linear trend?
d. Would you be willing to predict the weight when age = 25 months? Why?

extrapolation

4. When there is not a significant relationship between the variables, the value of r will be approximately
   a. 0.5
   b. 1
   c. -1
   d. 0
   e. none of theses

r = 0

as x changes, y doesn't change

r = 0

r ≈ 0

Use the following set of observations for the independent variable x and the dependent variable y in questions #9-13:

X	-3	-1	1	3
Y	8	4	5	-1

Q6

Question 14

$n = 64$

Lloyd's Cereal company packages cereal in 1 pound boxes (16 ounces). A sample of 64 boxes is selected at random from the production line every hour, and if the average weight is less than 15 ounces, the machine is adjusted to increase the amount of cereal dispensed. If the mean for 1 hour is 1 pound and the standard deviation is 0.2 pound, what is the probability that the amount dispensed per box will have to be increased?

$\mu = 16oz.$ $\quad \sigma = .2(16) = 3.2oz$ $\quad \sigma_{\bar{x}} = \frac{3.2}{\sqrt{64}}$

$P(\bar{X} < 15)$

Question 6

Which of the following statements is not true?

(4.4)

a) The sampling distribution of sample mean is approximately normal, mound-shaped, and symmetric for $n > 30$ or $n = 30$. T

b) The expected value of the sample mean, $\overline{X}$, is always the same as the expected value of X, the distribution of the population from which the sample was taken. $\mu_{\overline{x}} = \mu$ T

c) The sampling distribution of the sample mean, $\overline{X}$, is always reasonably like the distribution of X, the distribution from which the sample is taken. not always true

(Central limit thm)

X could be any kind of distr.

If $n \geq 30$ then distr of $\overline{x}$ will look normal

d) The larger the sample size, the better will be the normal approximation to the sampling distribution of sample mean. T

e) The standard deviation of the sampling distribution $\overline{X}$ of sample mean $= \sigma/\sqrt{n}$ where σ is the standard deviation of X. T

$\sigma_{\overline{x}} = \sigma/\sqrt{n}$

f) None of the above

Question 11

What effect does decreasing the sample size have on a distribution of sample means?

n

doesn't change $\mu_{\bar{x}}$

does change st. dev. $\Rightarrow \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

a) It will not make any difference

b) It will have more variation

c) It will have less variation

> ?

11. A study compared the body weight of a child to his/her metabolic rate. Use the following statistics to find the equation of the LSRL.

$\bar{x} = 12.5 \quad s_x = 6.568 \quad \bar{y} = 5.888 \quad s_y = 2.687 \quad r = .984$

The least squares regression line formula is $\hat{y} = a + bx$

The slope, b is calculated using $b = r\left(\frac{s_y}{s_x}\right)$ and the y-intercept is $a = \bar{y} - b\bar{x}$.

Slope