

Finite Mathematics

Review 5.

1. Let A and B be two events and $p(A) = .4$, $p(B) = .3$, $p(A \cup B) = .6$.
Find $p(A \cap B)$, $p(B/A)$, $p(A/B)$.

Solution. $p(A \cup B) = p(A) + p(B) - p(A \cap B)$. Therefore $0.6 = 0.3 + 0.4 - p(A \cap B)$
and $p(A \cap B) = 0.1$.

$$p(A|B) = \frac{p(A \cap B)}{p(B)} = \frac{1}{3}$$

$$p(B|A) = \frac{p(A \cap B)}{p(A)} = \frac{1}{4}$$

2. Five cards are taken at random from a standard deck of 52 cards. What is the probability to have at least two cards of the same denomination?

Solution. First we find probability of the complementary event E' : all cards are of different denominations. We can take 5 cards of different kind from a standard deck in $C(13, 5) \times 4^5 = 1317888$ ways. Total number of ways to take 5 cards

$$\text{is } C(52, 5) = 2598960. \text{ Hence } p(E') = \frac{1317888}{2598960} = 0.507$$

$$\text{and } p(E) = 1 - p(E') = 0.493.$$

3. A signal can go through one of three channels with probabilities 0.95, 0.93, and 0.97, respectively. The condition of each channel does not depend on the condition of two others. What is the probability that the signal will go through?

Solution. Let C_1 , C_2 , and C_3 be the events that the signal goes through channels 1, 2, or 3, respectively. Notice that $p(C_1) = 0.95$, $p(C_2) = 0.93$, and $p(C_3) = 0.97$. The signal will go through if it goes through at least one of three channels. Therefore the event E that the signal will go through can be represented as $E = C_1 \cup C_2 \cup C_3$. By De Morgan's law the complementary event (the signal will not go through) is $E' = C_1' \cap C_2' \cap C_3'$. By condition the events C_1' , C_2' , and C_3' are independent, whence

$$\begin{aligned} p(E') &= p(C_1')p(C_2')p(C_3') = (1 - p(C_1))(1 - p(C_2))(1 - p(C_3)) = \\ &= (1 - 0.95)(1 - 0.93)(1 - 0.97) = 0.05 \times 0.07 \times 0.03 = 0.000105. \end{aligned}$$

$$\text{Finally, } p(E) = 1 - p(E') = 0.999895.$$

4. A box contains 4 blue marbles, 5 red marbles, and 7 green marbles. Three times a marble is taken at random and returned to the box. What is the probability to have at least two marbles of different color?

Solution. First compute probability of the complementary event that all three marbles are of the same color. The event that we have three blue marbles in a row is $B_1 \cap B_2 \cap B_3$ where B_1, B_2 , and B_3 are the events that the marble number 1, 2, or 3 will be blue, respectively. Because marbles are returned to the box and the events B_1, B_2, B_3 are independent, the probability to have all three blue marbles

is $p(B_1)p(B_2)p(B_3) = \frac{4}{16} \times \frac{4}{16} \times \frac{4}{16} = \frac{4^3}{16^3}$; in the same way the probability to have all red marbles is $\frac{5^3}{16^3}$, and the probability to have all green ones is $\frac{7^3}{16^3}$.

Therefore $p(E') = \frac{4^3 + 5^3 + 7^3}{16^3} = 0.130$ and $p(E) = 0.870$.

5. The same question as in problem 4 but the marbles are not returned into the box.

Solution. The reasoning is similar to the previous problem but pay attention that after we take, for example, one blue marble from the box there remain 15 marbles and only three of them are blue. This way we obtain

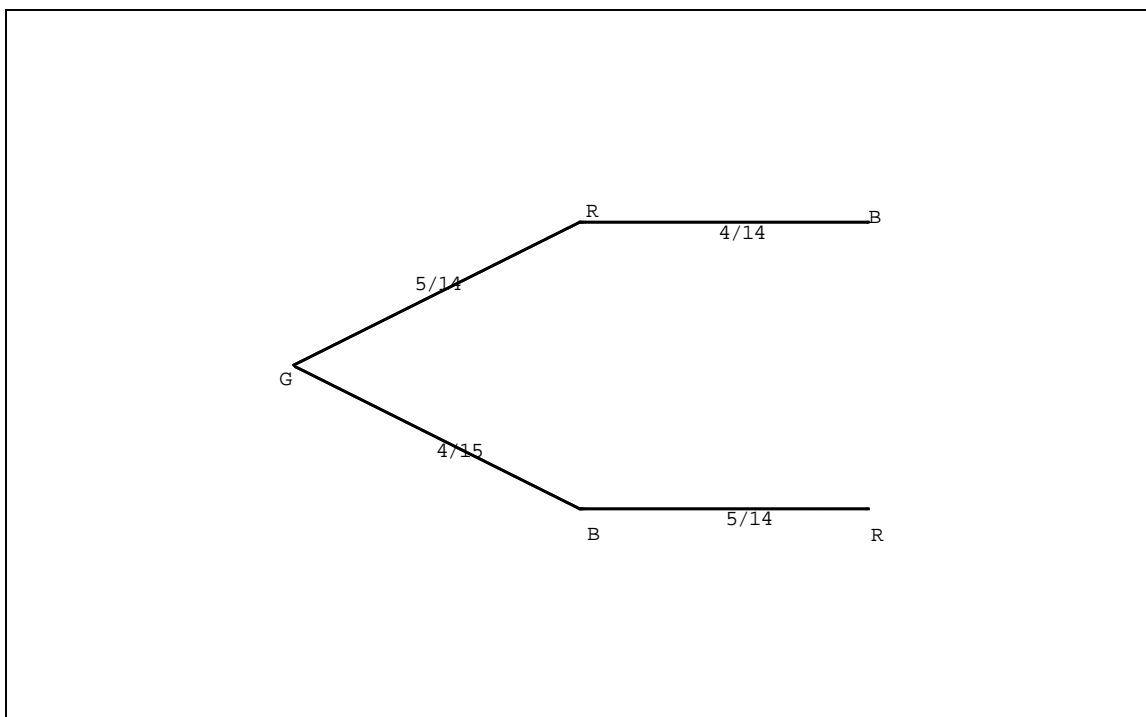
$$p(E') = \frac{4 \times 3 \times 2}{16 \times 15 \times 14} + \frac{5 \times 4 \times 3}{16 \times 15 \times 14} + \frac{7 \times 6 \times 5}{16 \times 15 \times 14} = .0875 \text{ and } p(E) = 0.9125.$$

6. In the conditions of problem 5 what is the probability to have marbles of three different colors if the first marble is green?

Solution. If the first marble is green then the probability that the second marble is red and the third one is blue is $\frac{5}{15} \times \frac{4}{14}$; and the probability that the second marble is blue

and the third one is red is $\frac{4}{15} \times \frac{5}{14}$. Therefore, $p(E) = \frac{20}{210} + \frac{20}{210} = \frac{4}{21}$.

The probability tree below illustrates these calculations.



7. Two fair dice are rolled 8 times. What is the probability to have the sum of the top numbers equal to 6 exactly 5 times?

Solution. We apply binomial probabilities. The number of times the experiment is repeated is 8. The probability of success (to have the sum of the top numbers equal to 6) is $5/36$. Indeed, when we roll two fair dice there are 36 possible outcomes from (1,1) to (6,6) and five of them: (1,5), (2,4), (3,3), (4,2), and (5,1) have sum equal to 6. The number of successes is 5. The probability we are looking for

is $P = C(n, m)p^m(1-p)^{n-m}$, where $p = \frac{5}{36}$, $1-p = \frac{31}{36}$, $n = 8$, $m = 5$. Therefore

$$P = C(8, 5) \left(\frac{5}{36} \right)^5 \left(\frac{31}{36} \right)^3 = C(8, 3) \left(\frac{5}{36} \right)^5 \left(\frac{31}{36} \right)^3 = \frac{8 \times 7 \times 6}{3!} \left(\frac{5}{36} \right)^5 \left(\frac{31}{36} \right)^3 = 56 \frac{5^5 31^3}{36^8} \approx .00185.$$

8. Find the area under the standard normal curve over the interval $-0.83 < z < 0.46$.

Solution. We use the table on page 671. The area we are interested in is the sum of two areas – one over the interval $0 < z < 0.46$ and the second over the interval $-0.83 < z < 0$. The first area we find directly from the table; it is 0.177. The second area by symmetry is the same as the area over the interval $0 < z < 0.83$ and from the table we get for it the value of 0.297. The answer is $0.177 + 0.297 = 0.474$.

9. For a large group of people of the same age and gender their weight is measured. The mean of the measurements is $m = 150$ lb and the standard deviation is $s = 15$ lb; assume the normal distribution of the measurements and find the percentage of people who weigh between 130 lb and 160 lb.

Solution. First we will find the z -scores of the numbers 130 and 160 according to the formula $z_x = \frac{x - m}{s}$. Therefore $z_{130} = \frac{130 - 150}{15} = -1.33$ and $z_{160} = \frac{160 - 150}{15} = 0.67$.

The probability we are looking for equals to the area under the standard normal curve over the interval $-1.33 < z < 0.67$. From the table on page 671 we see that this area is equal to $0.408 + 0.249 = 0.657$; hence the answer is 65.7%.

10. The results of rolling a die 20 times are 1, 3, 3, 6, 1, 4, 5, 3, 2, 3, 5, 1, 1, 5, 6, 4, 2, 3, 4, 1.

(a) Find the frequencies and the relative frequencies.

Solution.

item	frequency	Relative frequency
1	5	5/20
2	2	2/20
3	5	5/20
4	3	3/20
5	3	3/20
6	2	2/20

(b) Find the median and the mode.

Solution. The items 1 and 3 have the greatest frequency; the distribution is bimodal with the modes 1 and 3.

The number of items is even and the median is the average of two middle items (number 10 and number 11) when the items are arranged in the increasing order. We see from the table of frequencies that both middle items are equal to 3; hence the median is 3.

(c) Find the sample mean m and the sample standard deviation s .

Solution.

We can find the mean using the table of frequencies.

$$m = \frac{1 \cdot 5 + 2 \cdot 2 + 3 \cdot 5 + 4 \cdot 3 + 5 \cdot 3 + 6 \cdot 2}{20} = \frac{63}{20} = 3.15.$$

We will find the sample standard deviation s using the formula

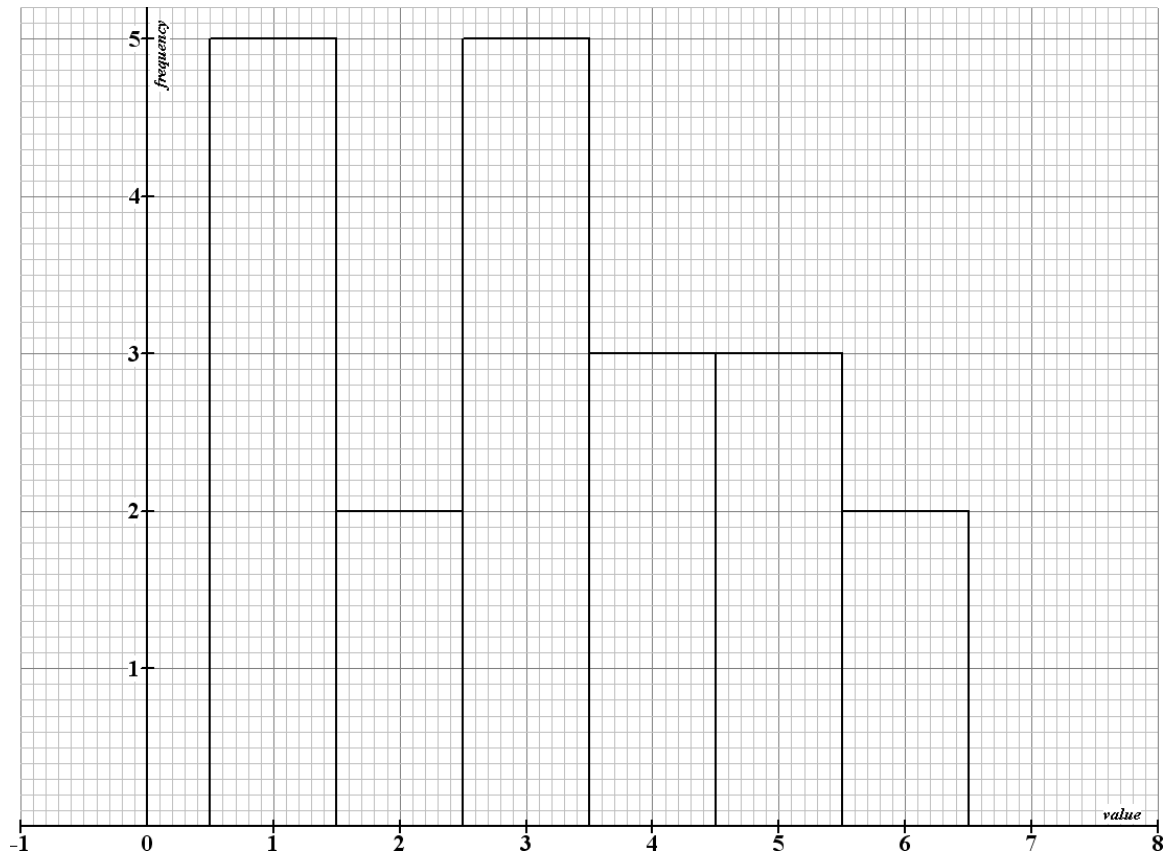
$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - m)^2}{n-1}} =$$

$$= \sqrt{\frac{5(1-3.15)^2 + 2(2-3.15)^2 + 5(3-3.15)^2 + 3(4-3.15)^2 + 3(5-3.15)^2 + 2(6-3.15)^2}{19}} =$$

$$= \sqrt{\frac{54.55}{19}} \approx \sqrt{2.87} \approx 1.69$$

(d) Construct a histogram.

Solution. We construct the histogram putting the values of the data items on the horizontal axis and frequencies (or relative frequencies) on the vertical axis. The histogram is shown below.



The next two problems are similar to the extra credit problems on the test.

11. Suppose that we have a coin that is not fair (e.g. the center of mass is misplaced) so that the probability to have a head is 0.6; respectively the probability to have a tail is 0.4. Suppose the coin is tossed 60 times. What is the probability that the number of heads will be between 30 and 40?

Solution: We will use the normal approximation to binomial distribution. First we have to see if the approximation will be reasonably good. The number of experiments is $n = 60$ and the probability of success (head) is 0.6. Therefore $np = 60 \cdot 0.6 = 36$ and $n(1 - p) = 60 \cdot 0.4 = 24$. These numbers are large enough to guarantee a reasonably good approximation. The parameters of the normal curve we use for approximation are: the mean $\mu = np = 36$ and the standard deviation

$\sigma = \sqrt{np(1 - p)} = \sqrt{60 \cdot 0.6 \cdot 0.4} = \sqrt{14.4} \approx 3.79$. If we take into consideration the continuous correction factor we see that the probability we are looking for is approximately equal to the area under the normal curve with $\mu = 36$, $\sigma = 3.79$ and over the interval from 30.5 to 39.5. Next we compute the z -scores.

$z_{30.5} = \frac{30.5 - 36}{3.79} \approx -1.45$, and $z_{39.5} = \frac{39.5 - 36}{3.79} \approx 0.92$. From the table on page 671 we see that the corresponding area under the standard normal curve is approximately $0.426 + 0.321 = 0.747$.

Remark. You do not have to do it on the test, but now we will compare the approximation we got with the exact binomial probability in this problem. The exact probability is

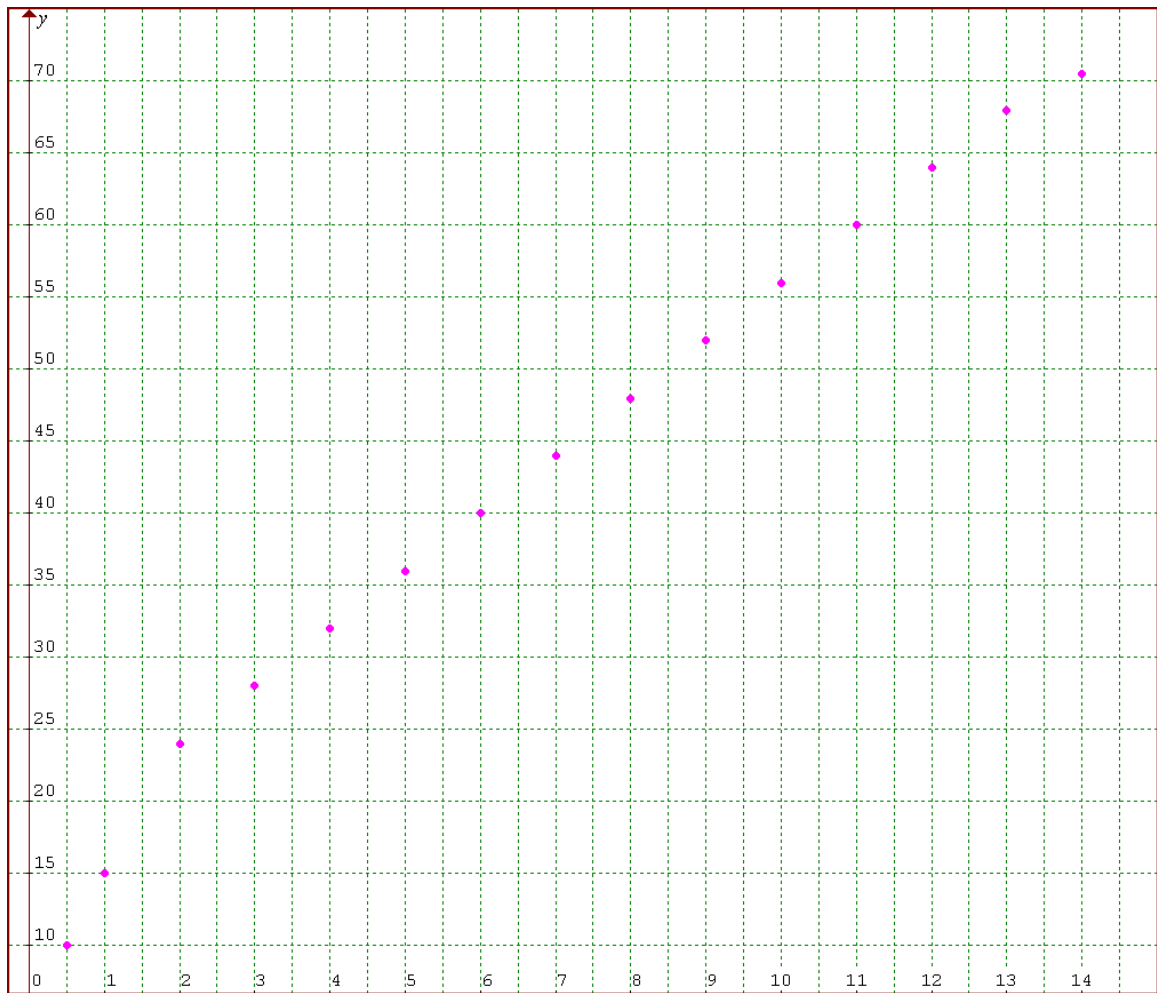
$$P = C(31, 60)0.6^{31}0.4^{29} + C(32, 60)0.6^{32}0.4^{28} + C(33, 60)0.6^{33}0.4^{27} + C(34, 60)0.6^{34}0.4^{26} \\ + C(35, 60)0.6^{35}0.4^{25} + C(36, 60)0.6^{36}0.4^{24} + C(37, 60)0.6^{37}0.4^{23} + C(38, 60)0.6^{38}0.4^{22} \\ + C(39, 60)0.6^{39}0.4^{21} \approx 0.7468060676.$$

12. Solve Problems 17 – 20 on page 682.

Dog Age (x)	1/2	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Equivalent human age (y)	10	15	24	28	32	36	40	44	48	52	56	60	64	68	70.5

(a) Plot a scatter diagram for the given data.

Below you see a computer generated scatter diagram.



(b) Find the equation of the regression line and graph the line on the scatter diagram.

Solution. To find the slope and the y-intercept of the regression line we have to compute the following.

The sum of all x -items, $\sum_{i=1}^{15} x_i = 105.5$.

The previous sum squared, $\left(\sum_{i=1}^{15} x_i\right)^2 = 105.5^2 = 11130.25$.

The sum of squares of all x -items, $\sum_{i=1}^{15} x_i^2 = 1015.25$.

The sum of all y -items, $\sum_{i=1}^{15} y_i = 647.5$.

The sum of the products of x and y -items, $\sum_{i=1}^{15} x_i y_i = 5719$.

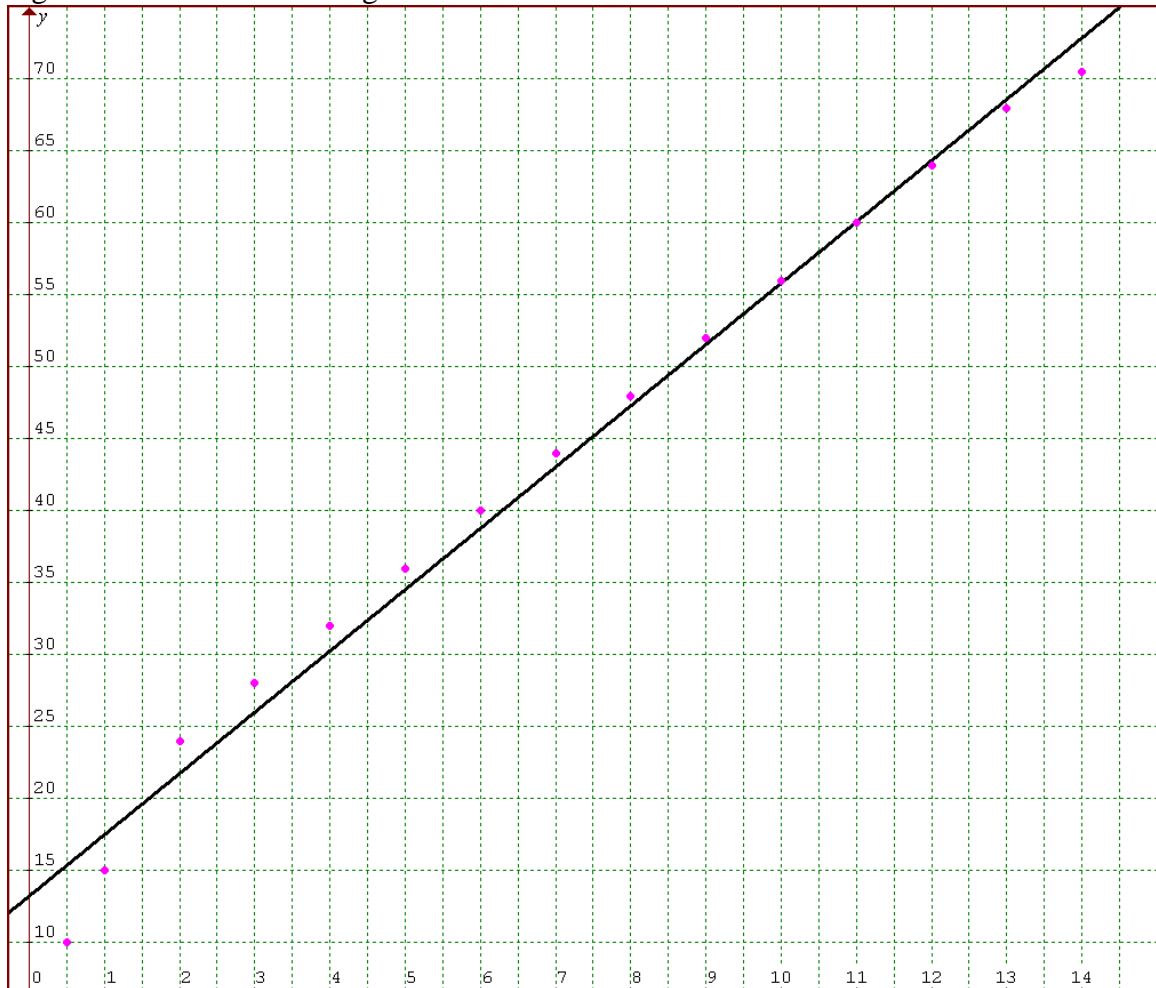
Next we compute the slope of the regression line, a , according to the formula on page 678.

$$a = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} = \frac{15 \times 5719 - 105.5 \times 647.5}{15 \times 1015.25 - 11130.25} \approx 4.26.$$

The y-intercept, b , can be computed by the formula on the same page.

$$b = \frac{\sum y - a \sum x}{n} = \frac{647.5 - 4.26 \times 105.5}{15} \approx 13.20.$$

The slope intercept equation of the regression line is $y = 4.26x + 13.20$; its graph together with the scatter diagram is shown below.



(c) Describe where the data points show the most pronounced departure from the regression line, and explain why this might be so.

Solution. Three points most departing from the regression line are (0.5, 10), (1, 15), and (14, 70.5) which correspond to the very young and the very old dog's age.

(d) Compute the coefficient of correlation.

We have to compute additionally the following.

The sum of all y-items squared, $\left(\sum_{i=1}^{15} y_i\right)^2 = 647.5^2 = 419256.25$.

The sum of squares of y-items, $\sum_{i=1}^{15} y_i^2 = 32975.25$.

We compute the correlation coefficient, r , using the formula on page 679.

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}} =$$

$$= \frac{15 \times 5719 - 105.5 \times 647.5}{\sqrt{15 \times 1015.25 - 11130.25} \sqrt{15 \times 32975.25 - 419256.25}} \approx 0.9942.$$

The correlation coefficient is very close to 1 indicating a very strong correlation between x and y -items.