PART G. PROBABILITY AND STATISTICS

CHAPTER 22. Data Analysis. Probability Theory

Sec. 22.1 Data: Representation, Average, Spread

Problem Set 22.1. Page 1054

1. **Representation of data.** For the present purpose first order the given data

$$12 \quad 11 \quad 9 \quad 5 \quad 12 \quad 6 \quad 7 \quad 9 \quad 11 \quad 11,$$

obtaining

$$5 \quad 6 \quad 7 \quad 9 \quad 9 \quad 11 \quad 11 \quad 11 \quad 12 \quad 12. \qquad \text{(A)}$$

As on p. 1051 you obtain a stem-and-leaf plot of these very simple data in the form

1	0	5
2	0	6
3	0	7
5	0	9 9
8	1	1 1 1
10	1	2 2

The first column gives the cumulative absolute frequency. Division by 10 (the number of data values) gives the cumulative relative frequency. The absolute frequencies of the values are 1, 1, 1, 2, 3, 2, respectively. Division of these values by 10 gives the relative frequencies. You can summarize these numbers in the following table.

Value	5	6	7	9	11	12
Absolute frequency	1	1	1	2	3	2
Relative frequency	0.1	0.1	0.1	0.2	0.3	0.2
Cumulative absolute frequency	1	2	3	5	8	10
Cumulative relative frequency	0.1	0.2	0.3	0.5	0.8	1.0

An example of a histogram is shown on p. 1051. For drawing it in the present problem you need the relative frequencies in your table. The figure shows the result. For drawing the boxplot of the data you need the quartiles q_L (the lower quartile), q_M (the median or middle quartile), and q_U (the upper quartile). If the number of data values is odd, then the median is one of the values, for instance, if 5 values are given, the median is the third value. If the number of data values is even, it is customary to take the average of the two values in the middle. In the present problem you have 10 values and take the sum of the fifth and the sixth, divided by 2, that is, by (A),

$$q_M = \frac{1}{2}(9 + 11) = 10.$$

There are 5 values 5, 6, 7, 9, 9 below the median; hence the third of them is the lower quartile

$$q_L = 7.$$

Similarly, there are 5 values 11, 11, 11, 12, 12 above the median, and the third of these values is the upper quartile

$$q_U = 11.$$

Observing that the smallest data value is 5 and the largest is 12, you can now draw the boxplot as shown in the figure on the next page.

11. Mean and variance. These two quantities (and the standard deviation, which is the square root of the variance) will be much more important in our further work than the quartiles just considered. The mean measures the average size of the data values. It is the arithmetic mean of the values, as defined in (5). In the present problem you have $n = 10$ values and obtain

$$\bar{x} = \frac{1}{10}(5 + 6 + ... + 12 + 12) = 9.3.$$

This is a sum of 10 values. More simply, you can take the numerically different values in the table in Prob. 1 just considered, multiply each by its absolute frequency, and take the sum. This gives

$$\bar{x} = \frac{1}{10}(5 + 6 + 7 + 2 \cdot 9 + 3 \cdot 11 + 2 \cdot 12) = 9.3.$$

This differs from the median 10 (see Prob. 1 above). It is more useful for measuring the average size of the data values because it takes into account the size of each value, not merely its position in the ordered data.

Using $\bar{x} = 9.3$, you calculate from (6) the variance

$$s^2 = \frac{1}{9}[(5-9.3)^2 + (6-9.3)^2 + ... + (12-9.3)^2 + (12-9.3)^2] = 6.455556.$$

More simply, by taking equal data values together as before, you obtain

$$s^2 = \frac{1}{9}[(5-9.3)^2 + (6-9.3)^2 + (7-9.3)^2 + 2 \cdot (8-9.3)^2 + 3 \cdot (11-9.3)^2 + 2 \cdot (12-9.3)^2] = 6.455556.$$

By taking the square root you obtain the standard deviation

$$s = 2.540779.$$

This differs considerably from the interquartile range IQR= $11 - 7 = 4$ (see Prob. 1 above).

The variance and standard deviation are better for measuring the spread (the variation) of the data because, as in the case of $\bar{x}$, these quantities take into account the size of each data value, not merely its position in the data ordered in ascending order (as in (A) in Prob. 1).

17. Reduction of data by deleting outliers. In data such as

$$4, \quad 1, \quad 3, \quad 10, \quad 2$$

the value 10 lies so far away from the other values that you can suspect an error in measuring or recording the data. If you are reasonably sure that some error has occurred, human or by the apparatuses used, you can reduce the data by omitting the obviously erroneous value (or values), called an *outlier*. This will generally change both the mean as well as the variance and hence the standard deviation.

Suppose that those data come from a problem such that you must assume that 10 is the outlier. Its omission will decreases the mean

$$\bar{x} = \frac{1}{5}(4 + 1 + 3 + 10 + 2) = 4$$

to

$$\bar{x}_{red} = \frac{1}{4}(4 + 1 + 3 + 2) = 2.5.$$

Similarly, the variance

$$s^2 = \frac{1}{4}[(4-4)^2 + (1-4)^2 + (3-4)^2 + (10-4)^2 + (2-4)^2] = 12.5$$

is reduced to

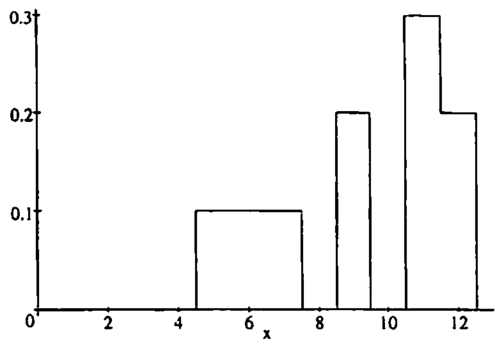
$$s^2_{red} = \frac{1}{3}[(4-2.5)^2 + (1-2.5)^2 + (3-2.5)^2 + (2-2.5)^2] = 1.666667.$$

Now take square roots to see that the standard deviation

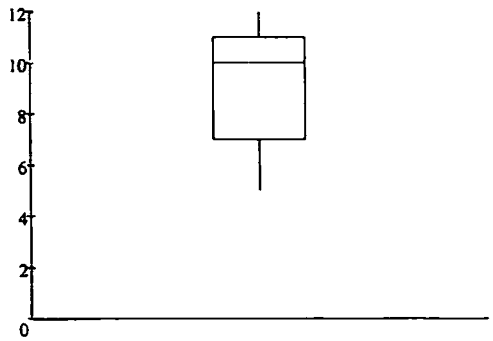
$$s = \sqrt{12.5} = 3.535534$$

is reduced to

$$s_{red} = \sqrt{\frac{5}{3}} = 1.290994.$$



Section 22.1. Problem 1. Histogram



Section 22.1. Problem 1. Boxplot

Sec. 22.2 Experiments, Outcomes, Events

Problem Set 22.2. Page 1057

5. **Sample space.** This sample space S consists of 10 points (outcomes), which you can list as follows (D = Defective, N = Nondefective).

Number of trials	Outcome
1	*D*
2	*ND*
3	*NND*
...	...
10	*NNNNNNNNND*

This is the end of the possibilities because you have only 10 bolts in the lot. If you drew with replacement, your list of outcomes would go on and your sample space would be infinite.

9. **Subsets.** The empty set, three one-point subsets, three two-point subsets, and S itself.

13. **De Morgan's laws.** Without referring to any diagrams and just using the definitions you can obtain the first law by starting on the left:

$A \cup B$ is the set of all points in A or B (or both).

Hence its complement

$(A \cup B)^c$ is the set of points neither in A nor in B.

Now on the right,

A^c is the set of points not in A,

B^c is the set of points not in B.

Hence the intersection

$A^c \cap B^c$ is the set of points simultaneously not in A and not in B,

that is, the set of points neither in A nor in B. This proves De Morgan's first law.

In the second law on the left,

$A \cap B$ is the set of points simultaneously in both A and B.

Call this intersection C. Hence the complement

$(A \cap B)^c$ is the set of points not in C.

On the right of the second law,

A^c is the set of points not in A,

B^c is the set of points not in B.

Hence the union

$D = A^c \cup B^c$ is the set of points either not in A or not in B.

Now if a point is in A but not in B, it is in B^c, hence it is in D. Similarly, if a point is in B but not in A, it is in D. This shows that D consists of the points that are not simultaneously in both A and B, that is, D consists of the points not in the intersection $C = A \cap B$. This proves De Morgan's second law.

Sec. 22.3 Probability

Problem Set 22.3. Page 1063

1. **Rolling two fair dice.** Problems like this amount to counting cases that favor the event in question, namely, the event of obtaining 4, 5, or 6. You can arrange the $6 \cdot 6$ possibilities in a square of 6 rows and 6 columns,

$$\begin{matrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{matrix}$$

where the first number is the number the first die turns up and the second number is the number the second die turns up. The sum 4 corresponds to 3 pairs of values (1, 3), (2, 2), (3, 1) in a sloping line. The sum 5 is obtained from the 4 pairs of values (1, 4), (2, 3). ... in the neighboring sloping line, and the sum 6 from the 5 pairs of values (1, 5), (2, 4), ... in the next sloping line. Together,

$$3 + 4 + 5 = 12$$

of the 36 equally likely cases favor the event in question, which thus has the probability 12/36 = 1/3.

Of course, you need not write down all these cases; you can find the answer simply by noting that

$$4 = 1 + 3 = 2 + 2 = 3 + 1 \qquad 3 \text{ cases}$$

$$5 = 1 + 4 = 2 + 3 = 3 + 2 = 4 + 1 \qquad 4 \text{ cases}$$

$$6 = 1 + 5 = 2 + 4 = 3 + 3 = 4 + 2 = 5 + 1 \qquad 5 \text{ cases}$$

which adds up to 12 cases, as before.

3. **Sampling with replacement.** "At least one" often suggests considering the complementary event "None". Now you have 30 screws, 10 of them left-handed. Hence the probability of obtaining a left-handed screw is $10/30 = 1/3$. You draw with replacement. Hence before drawing the second screw, you put the screw drawn (which may be left- or right-handed) into the box and mix the content thoroughly, so that the probability of drawing another left-handed screw is the same as before, namely, $10/30 = 1/3$. The point of drawing "with replacement" is that the previous drawings do not influence the further ones. This is the feature of independent events, and you now get from (14) the probability

$$(1/3)\cdot(1/3) = 1/9 = 11\%.$$

This is the probability of the complement "No right-handed screw" of the event "At least one right-handed screw", whose probability you are supposed to determine. Hence the Complementation Rule (Theorem 1) gives the answer $8/9 = 89\%$, approximately.

15. **Shooting.** Here assume independence in case (b), which is more or less realistic. "At least once" suggests the use of the complementation rule and gives the answer in the form $1 - (3/4)^2 = 0.4375$ shown on p. A47 of Appendix 2 of the book, which is substantially less than in case (a). You can check the result by calculating and summing the probabilities of the three possible outcomes (H = hit, M = miss)

$$HH \quad HM \quad MH.$$

Because of independence you obtain

$$\frac{1}{4}\cdot\frac{1}{4} + \frac{1}{4}\cdot\frac{3}{4} + \frac{3}{4}\cdot\frac{1}{4} = \frac{7}{16} = 0.4375.$$

In this and similar problems you may do well to check the results by calculating the probability of the outcomes not yet considered, so that you have available all the probabilities and you can see whether they add up to 1. In the present problem this is quite simple. The probability of MM in case (b) is $(3/4)^2 = 9/16$, which, together with 7/16 above adds up to 1.

19. **Extension of Theorem 4 to three events.** Theorem 4 states that under the given conditions,

$$P(A \cap B) = P(A)P(B|A). \tag{13}$$

(You will not need the other formula in (13).) The idea of the proof is introducing suitable notations so that you can reduce the present case to that in the theorem. Write D and E instead of A and B in (13). Then you have

$$P(D \cap E) = P(D)P(E|D). \tag{13*}$$

Now comes the trick that will do it. Put

$$D = A \cap B.$$

Since forming intersections is associative, that is,

$$D \cap E = (A \cap B) \cap E = A \cap B \cap E,$$

formula (13*) takes the form

$$P(A \cap B \cap E) = P(A \cap B)P(E|(A \cap B)). \tag{13**}$$

You are almost done. Set $E = C$ on both sides of (13**). (So here you made a little detour because you could have retained C and get away without introducing E; this is typical of many proofs that one first gets enough elbowroom, perhaps more than one eventually needs in completing a proof, but this often not clear at the beginning.) Furthermore, on the right side of (13**) insert $P(A \cap B)$ from (13). Then you obtain precisely the formula to be proved.

Go over the proof before you leave the problem, so that you fully understand the idea of this proof.

Sec. 22.4 Permutations and Combinations

Problem Set 22.4. Page 1068

3. Number of different samples. The order in which you obtain the four objects of a certain sample does not matter. Hence you are concerned with *combinations*, not permutations. A certain object may appear in a sample only once, not twice or more times. Hence you are dealing with combinations *without repetition*. Theorem 3 now gives the answer

$$\binom{50}{4} = \frac{50\cdot 49\cdot 48\cdot 47}{1\cdot 2\cdot 3\cdot 4} = \frac{5\,527\,200}{24} = 230\,300.$$

From the answer you see that counting cases would be hopeless, even in the case of smaller populations.

You could check the result as follows. You have 50 possibilities for picking a first object, then 49 possibilities of picking a second, then 48 for picking a third, and finally 47 for picking a fourth. Hence you have

$$50\cdot 49\cdot 48\cdot 47 = 5527200 \qquad \text{(A)}$$

possibilities of picking 4 objects from the given 50 objects in a given order. Now the order does not matter in sampling; what matters is just which objects you got in your sample. Hence you have to divide (A) by the number of permutations of 4 objects, which is 4! This is the number of different orders in which you can arrange the 4 objects. You thus obtain the same answer as before.

13. Defectives. (a) The lot consists of 6 nondefectives and 2 defectives. The number of samples of size 3 is

$$\binom{8}{3} = \frac{8\cdot 7\cdot 6}{1\cdot 2\cdot 3} = 56.$$

This follows as in Prob. 3.

(b) If you want no defectives, put them aside and conclude that you now sample from a population of the remaining 6 objects. There are

$$\binom{6}{3} = \frac{6\cdot 5\cdot 4}{1\cdot 2\cdot 3} = 20$$

different samples of size 3.

(c) 1 defective can be selected from 2 in 2 ways. The other 2 objects in a sample of size 3 can be selected from the 6 nondefective objects in

$$\binom{6}{2} = \frac{6\cdot 5}{1\cdot 2} = 15$$

ways. Hence the total number of possibilities is $2\cdot 15 = 30$ (combine each of the two former with each of the 15 latter possibilities).

(d) 2 defectives leaves you with 1 nondefective in a sample of size 3, that is, with 6 choices.

You can now check your result. (b), (c), (d) exhaust all possibilities; hence the sum of the three values should equal the total number of possibilities, which is 56. Indeed,

$$20 + 30 + 6 = 56.$$

Keep this way of checking in mind; it is useful in other perhaps more involved applications.

16e. Binomial coefficients satisfy a very large number of relationships, a small selection of which is included in Chap. 24 of Ref. [1] in Appendix 1. Formula (14) is one of the most useful ones of them. To prove it, start from

$$(1+x)^p(1+x)^q = (1+x)^{p+q}$$

and develop $(1+x)^p$, $(1+x)^q$ as well as the right side in powers of x by means of the binomial theorem. Then equate the coefficients of the power x^r on both sides. On the right you have just one term, namely,

$$\binom{p+q}{r}x^r.$$

One the left you have

$$\left[\binom{p}{0}1+\binom{p}{1}x+\binom{p}{2}x^2+\dots+\binom{p}{p}x^p\right]\left[\binom{q}{0}1+\binom{q}{1}x+\binom{q}{2}x^2+\dots+\binom{q}{q}x^q\right].$$

Now you obtain x^r by multiplying $x^0 \cdot x^r$, then $x \cdot x^{r-1}$, then $x^2 \cdot x^{r-2}$, ..., finally, $x^r \cdot x^0$. These are $r+1$ products. The corresponding coefficients are

$$\binom{p}{0}\binom{q}{r},\ \binom{p}{1}\binom{q}{r-1},\ \binom{p}{2}\binom{q}{r-2},\ \binom{p}{3}\binom{q}{r-3},\dots,\binom{p}{r-1}\binom{q}{1},\ \binom{p}{r}\binom{q}{0}.$$

The sum of these terms is exactly the left side of (14).

Sec. 22.5 Random Variables, Probability Distributions

Problem Set 22.5. Page 1074

1. **Discrete distribution.** This distribution is discrete. It has the possible values 1, 2, 3. Its probability function $f(x)$ has the values

$$f(1) = 1/14, \quad f(2) = 4/14, \quad f(3) = 9/14;$$

see the figure. These are the probabilities with which the values 1, 2, 3 are assumed. $f(x)$ satisfies (6); otherwise it could not serve as a probability function.

From $f(x)$ you can obtain the distribution function $F(x)$ by summing values of $f(x)$, as indicated in (4), by which the distribution function is defined. Since $x = 1$ is the smallest possible value, you see that $F(x) = 0$ for $x < 1$. At $x = 1$ it jumps up to $1/14 = f(1)$; that is, $F(1) = 1/14$. Then it remains constant and equal to 1/14 until it reaches $x = 2$, where the next jump occurs, this time of height $4/14 = f(2)$. In formulas,

$$F(x) = f(1) = 1/14 \quad \text{if} \quad 1 \leqq x < 2$$

and

$$F(2) = f(1) + f(2) = 1/14 + 4/14 = 5/14.$$

This value is retained until the next (and last) jump occurs at $x = 3$, this time of height $f(3) = 9/14$. Thus

$$F(x) = 5/14 \quad \text{if} \quad 2 \leqq x < 3$$

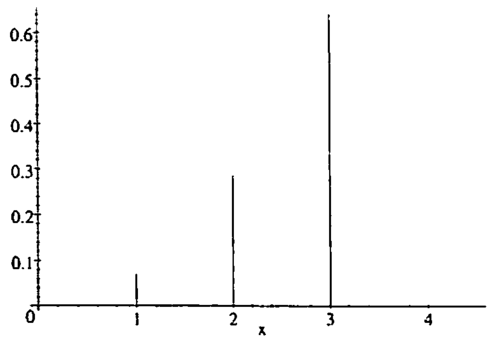
and

$$F(3) = f(1) + f(2) + f(3) = 1/14 + 4/14 + 9/14 = 1.$$

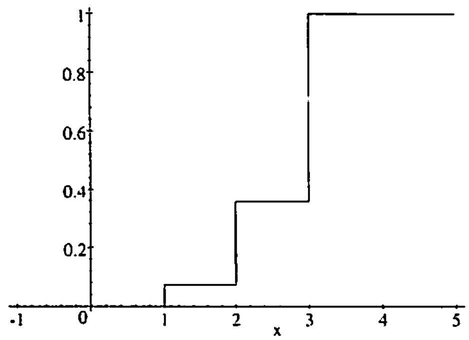
This last value 1 is typical. It is always reached at the last jump and is then retained for all larger x; thus

$$F(x) = 1 \quad \text{if} \quad x \geqq 3.$$

If a distribution has infinitely many possible values, then $F(x) = 1$ may not be reached but appear as the sum of the infinite series of the probabilities of the infitely many possible values. An illustration is given in Example 4. Waiting problems are of interest in connection with ticket counters, telephone services, and so on.



Section 22.5. Problem 1. Probability function



Section 22.5. Problem 1. Distribution function

7. Continuous distribution. Percentage points. Uniform distribution. The figures show the density and the distribution function

$$F(x) = \begin{cases} 0 & \text{if } x < 2 \\ \int_2^x 0.25\,dv = 0.25x - 0.5 & \text{if } 2 \leqq x \leqq 6 \\ 1 & \text{if } x > 6. \end{cases}$$

This distribution is called the *uniform distribution* (on the interval from 2 to 6). Note that the area under the curve of $f(x)$ equals 1, as it should be because of (10) on p. 1073. The solutions of (a), (b), and (c) follow from (1), that is,

$$P(X \leqq x) = F(x).$$

(a) $F(c) = 0.25c - 0.5 = 0.9$, hence $c = 4(0.9 + 0.5) = 5.6$. Indeed, $c = 5.6$ is the point such that 90% of the area under the curve of $f(x)$ lies to the left and 10% to the right of this point.

(b) Here you have to use that the complement of $X \geqq x$ is $X < x$. Hence the probability of $X \geqq x$ is 1 minus the probability of $X < x$. Now the given distribution is *continuous*. Hence probabilities are given by integrals, so that $X < x$ has the same probability as $X \leqq x$. But the latter is given by $F(x)$, as follows from (1). Setting $x = c$, you thus have

$$P(X \geqq c) = 1 - F(c) = 0.5,$$

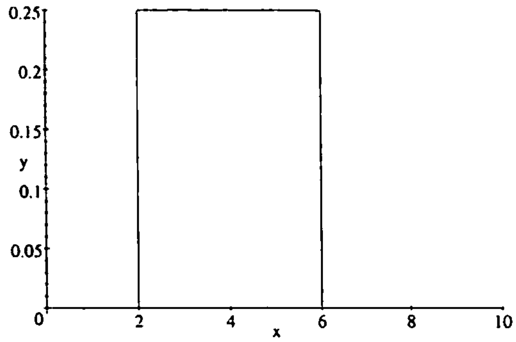
hence

$$F(c) = 0.25c - 0.5 = 0.5, \qquad c = 4.$$

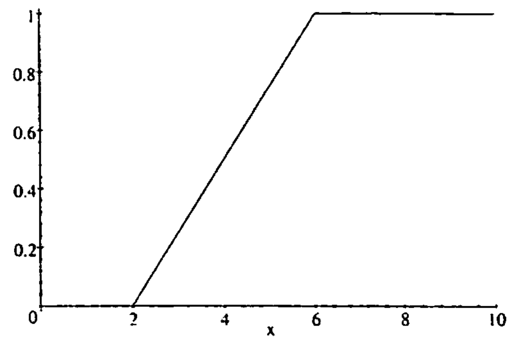
This can also be seen by inspection because c is the point such that 50% of the area under the curve of

$f(x)$ lies to the left and 50% to the right of this point.

(c) $F(c) = 0.25c - 0.5 = 0.05$, $c = (0.50 + 0.05)/0.25 = 2.2$. Indeed, 0.2 are 5% of the length 4 of the interval on which $f(x)$ is not zero. The determination of percentage points as in the present problem will occur quite frequently in Chap. 23 on statistical methods, so that you may do well to make sure that you understand the details of this problem, which is particularly simple, due to the fact that $f(x)$ is piecewise constant.



Section 22.5. Probiem 7. Density of the uniform distribution on the interval from 2 to 6



Section 22.5. Problem 7. Distribution function of the uniform distribution on the interval from 2 to 6

19. **Complements of events,** such as those in the present problem, must be considered quite frequently in applications. Now the complement of a set S is the set of all points not in S; see Sec. 22.2. With this and the real line in mind you will obtain the following answers. The complement of $X \leqq b$ is $X > b$. Note that $X = b$ is included in the given event, hence it must not occur in the complement. The complement of $X < b$ is $X \geqq b$. Here, b belongs to the complement. Similarly in the next two cases. The complement of $X \geqq c$ is $X < c$, and the complement of $X > c$ is $X \leqq c$. If you have an interval, such as $b \leqq X \leqq c$, the complement consists of all numbers that lie outside this interval, so that they are either smaller than b or larger than c. The given interval is *closed*; this means that the points b and c are regarded as points of the interval; hence they do not belong to the complement, which therefore is called *open*. The complement consist of the two open infinite intervals $X < b$, stretching to $-\infty$, and $X > c$, stretching to ∞. These are called *infinite intervals* because they extend to $-\infty$ and ∞, respectively. An important point in this problem is the careful distinction between strict inequalities, such as $X < b$, and those that include equality, such as $X \leqq b$. In the case of *continuous* distributions it does not matter, as just explained in Prob. 7(b), but in the case of *discrete* distributions it becomes essential.

Sec. 22.6 Mean and Variance of a Distribution

Problem Set 22.6. Page 1078

1. **Discrete distribution.** The sum of the probabilities of all four possible values is

$$k\left[\binom{3}{0}+\binom{3}{1}+\binom{3}{2}+\binom{3}{3}\right] = k[1+3+3+1] = 8k.$$

According to (6) in Sec. 22.5, this sum must be equal to 1. Hence $k = 1/8 = 0.125$. This gives the probabilities

$$f(0) = 0.125,\ f(1) = 0.375,\ f(2) = 0.375,\ f(3) = 0.125.$$

To obtain the mean, multiply each probability $f(x)$ by the corresponding x and take the sum over all possible values $x = 0, 1, 2, 3$; see (1a). This gives

$$\mu = 0 \cdot 0.125 + 1 \cdot 0.375 + 2 \cdot 0.375 + 3 \cdot 0.175 = 1.5.$$

You can now obtain the variance from (2a) in the form

$$\sigma^2 = (0-1.5)^2 \cdot 0.125 + (1-1.5)^2 \cdot 0.375 + (2-1.5)^2 \cdot 0.375 + (3-1.5)^2 \cdot 0.125 = 0.75.$$

Note that the values $x = 0$ and 3 contribute 75% to the variance, three times as much as the other two values, because they lie farther away from the mean than the latter. This is typical.

3. **Continuous distribution.** The density is

$$f(x) = 2x \quad \text{if} \quad 0 \leqq x \leqq 1$$

and 0 otherwise. (This cannot be the probability function of a discrete distribution because the possible values of the latter are discrete.) The mean μ is obtained from (1b), where you have to integrate from 0 to 1 only, because $f(x) = 0$ for other values. Check that the area under the curve of $f(x)$ equals 1. This is easy. The integral of $2x$ is x^2, which gives 1 when evaluated at the limits of integration. Hence (10) in Sec. 22.5 is satisfied. By integration you now obtain from (1b)

$$\mu = \int_0^1 x 2x\,dx = \frac{2}{3}x^3\Big|_0^1 = \frac{2}{3}.$$

You can now obtain the variance from (2b) by integration, again from 0 to 1 only, for the same reason as before.

$$\sigma^2 = \int_0^1 \left(x - \frac{2}{3}\right)^2 (2x)\,dx = \int_0^1 \left(2x^3 - \frac{8x^2}{3} + \frac{8x}{9}\right)dx = \frac{1}{18}.$$

15. **Expected value.** The expected value (expected gain in a game, expected profit in a business) is the mean value of the corresponding random variable. Actually, in the present problem the random variable is

$$X = \textit{Number of turkeys sold},$$

not the profit directly. Notice that 5, 6, 7, 8 are possible values. Now the sum of the corresponding probabilities is 1. Hence 5, 6, 7, 8 are *all* the possible values of X. Equation (1a) thus gives the expected number of turkeys sold on any given day, namely,

$$\mu = 5 \cdot 0.1 + 6 \cdot 0.3 + 7 \cdot 0.4 + 8 \cdot 0.2 = 6.7.$$

Hence the expected sale per day is 6 or 7 turkeys. Multiplication by 3.50 gives the expected profit $6.7 \cdot 3.50 = \$23.45$.

Sec. 22.7 Binomial, Poisson, and Hypergeometric Distributions

Problem Set 22.7. Page 1083

1. **Coin tossing, binomial distribution with $p = 1/2$.** This is a typical application of the binomial

distribution with equal probabilities of success and failure (equal probabilities of heads and tails), provided the coins are fair, so that $p = q = 1/2$, and tossing is done in an orderly way, so that independence of trials is guaranteed. The probability function of the binomial distribution is given by (2). If $p = q = 1/2$, you see that in (2) you have

$$p^x q^{n-x} = \left(\frac{1}{2}\right)^{x+n-x} = \left(\frac{1}{2}\right)^n.$$

Hence in this case, (2) takes the simple form (2*).

In the problem, you toss 5 coins. Hence $n = 5$, and (2*) becomes

$$f(x) = \binom{5}{x}\left(\frac{1}{2}\right)^5, \quad x = 0, 1, ..., 5. \tag{A}$$

This is the probability of obtaining precisely x heads in a trial, which consists of simultaneously tossing the 5 coins. This is the answer to the first question in the problem.

Since you are supposed to calculate several probabilities, it may be wise to calculate all the 6 probabilities from (A). For the binomial coefficients you obtain

$$1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1. \tag{B}$$

Multiplication by $(1/2)^5 = 1/32$ gives the corresponding probabilities. The sum of the values in (B) equals 32. Hence the sum of all the probabilities equals 1, as it should be.

From (B) you can readily obtain the answers to the further questions, as follows. You obtain the event "*No heads*" with the probability $f(0) = 1/32$. The complement of this event is "*At least one head.*" By the complementation rule (Sec. 22.3) it has the probability $1 - f(0) = 31/32$.

Similarly, the event "*More than* 4 *heads*" is the same as "5 *heads*" and has the probability $f(5) = 1/32$. The complement is "*Not more than* 4 *heads*". By the complementation rule, the corresponding probability is 31/32.

9. **Poisson distribution.** The information that the resistors have 60 ohms and the characterization of defective ones are irrelevant. Essential are the probability of defectives, 0.1% or $p = 0.001$, and the lot size $n = 200$. Since p is small and n is large, you can expect good approximations from the Poisson distribution. Example 2 will give you some help in doing the present problem, which is even simpler than Example 2.

Obviously, the guarantee will be violated if a lot contains 1 or several defectives. This suggests considering the event

"*At least one defective in a lot.*" (C)

"At least one" often is the signal for switching to the complementary event

"*No defectives in a lot*" (D)

and then using the complementation rule (Sec. 22.3). Since $p = 0.001$ and $n = 200$, you will be dealing with the probability function (5) with the mean $\mu = np = 0.2$, that is,

$$f(x) = \frac{0.2^x}{x!}e^{-0.2} = 0.818731\frac{0.2^x}{x!}. \tag{E}$$

Hence the probability of obtaining no defectives in a given lot is $f(0) = 0.818731$. You thus obtain the answer that the probability of at least 1 defective in a lot is $1 - f(0) = 0.181269$; practically, 18%. This is the probability that a given lot will violate the guarantee.

This would be considered as much too large in practice. What you could do is to give the guarantee that there will be at most 1 defective in a lot. This would reduce the violation rate from 18% to about 1.8%. (Can you calculate this?)

In the present problem the exact distribution is the binomial distribution with $p = 0.001$ and $n = 200$, hence $\mu = 0.2$, that is,

$$f(x) = \binom{200}{x}0.001^x(1 - 0.001)^{200-x}.$$

It follows that the probability of obtaining no defectives in a lot is

$$f(0) = (1 - 0.001)^{200} = 0.818649.$$

Hence your approximate value obtained from the Poisson distribution is practically the same.

15. **Multinomial distribution.** You should first note that for $k = 2$ the multinomial distribution reduces to the binomial distribution. Indeed, you first obtain, writing $x_1 = x, x_2 = n - x$ (which follows from $x_1 + ... + x_k = n$ with $n = 2$), $p_1 = p, p_2 = q = 1 - p$ (which follows from $p_1 + ... + p_k = 1$ with $k = 2$)

$$f(x_1, x_2) = \frac{n!}{x_1!x_2!}p_1^{x_1}p_2^{x_2} = \frac{n!}{x!(n-x)!}p^x q^{n-x} = \binom{n}{x}p^x q^{n-x}.$$

From this you see that the idea of the derivation of the formula in the problem is the same as that in the text given for the binomial distribution. Namely, the product of the powers of $p_1, ..., p_k$ is the probability of obtaining in n trials

x_1 times the event A_1,

x_2 times the event A_2,

and so on, in one particular order, for instance, in each of the first x_1 trials you obtain A_1, then in each of the next x_2 trials you obtain A_2,..., finally, in each of the last x_k trials you obtain A_k. This is one order of obtaining those events. This probability must now be multiplied by the number of permutations of n things divided into k classes of alike things differing from class to class, because this is the number of permutations of the n outcomes obtained, as given in Theorem 1(b) in Sec. 22.4 (which was also used in the derivation of the binomial distribution, for precisely the same purpose). This number is given by the quotient of the factorials in the formula of the problem. This solves the problem.

Sec. 22.8 Normal Distribution

Problem Set 22.8. Page 1090

1. **Use of normal tables.** If x-values are given and probabilities wanted, use Table A7 in Appendix 5. If probabilities are given and x-values wanted, use Table A8. First make sure that you do understand the difference between these two tasks. Furthermore, those tables are given for the normal distribution with mean 0 and variance 1. This is called the *standardized normal distribution*. And your task in most cases is to apply formula (4), which expresses the distribution function F of the given normal distribution in terms of the distribution function Φ of the standardized normal distribution. Before you do the problem, take another look at Example 2, which is of the same type. In the problem, the given mean is $\mu = 10$, the variance is $\sigma^2 = 4$, hence the standard deviation is $\sigma = 2$. Can you visualize this distribution? Its density is as in Fig. 487, but shifted 10 units to the right (why?) and its shape is even flatter than the flattest curve in that figure (why?). With those values for μ and σ, formula (4) takes the form

$$F(x) = \Phi\left(\frac{x - 10}{2}\right).$$

For x you now have to insert the given value 12 and note that $P(X > 12)$, not $P(X \leqq 12)$ is wanted. This calls for the use of the complementation rule. From Table 7 you thus obtain

$$P(X > 12) = 1 - P(X \leqq 12) = 1 - F(12) = 1 - \Phi\left(\frac{12 - 10}{2}\right) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587.$$

This is about 16%.

Is there a connection to the 16% in Fig. 489? There is. You have $\mu + \sigma = 10 + 2 = 12$. Hence you were asking for the probability corresponding to the right "tail" in the left part of Fig. 489.

$P(X < 10)$ asks for the probability that the normal random variable X assume any value between $-\infty$ and the mean. This probability is always 50%, regardless of the size of the mean and the variance, as follows from the fact that the bell-shaped curve of the normal density is symmetric with respect to the mean.

In the next case you calculate $P(X < 11) = P(X \leqq 11) = F(11) = \Phi(0.5)$ and then use Table A7, obtaining 0.6915.

Similarly in the last case, where you have to find the probability corresponding to an interval and must convert two expressions from F to Φ. Intervals occur often. For this reason the text gives the corresponding formula (5). With the given $a = 9$ and $b = 13$, and with μ and σ as before you obtain

$$F(13) - F(9) = \Phi(1.5) - \Phi(-0.5).$$

Table A7 has no values for negative z, for the good reason that because of the symmetry of the normal density,

$$\Phi(-z) = 1 - \Phi(z).$$

Hence you obtain

$$F(13) - F(9) = \Phi(1.5) - (1 - \Phi(0.5)) = 0.9332 - 1 + 0.6915 = 0.6247.$$

Can you find out whether this last result seems reasonable? Well, the interval has length $4 = 2\sigma$. In Fig. 489(a) (not (b)!) the probability corresponding to the interval of length 2σ equals 68%, as indicated. Your present interval is not symmetrically located with respect to the mean, and this causes a slight loss in probability, as seems obvious from the density curve (the bell-shaped curve). Hence 0.6247 seems reasonable.

13. **Unknown value for a given probability.** This is a problem of the second kind, where you need Table A8. Read first the beginning of Prob. 1, where the distinction is again explained. You will be concerned with a normal distribution whose mean is 1000 and whose standard deviation is 100. Hence its distribution function $F(x)$ is expressed by that of the stadardized normal distribution as shown in (4). With those values you have

$$F(x) = \Phi\left(\frac{x - 1000}{100}\right). \tag{A}$$

Now the problem asks for an x such that the probability of observing any value of X (which is the sick-leave time) greater than x is 20% = 0.2. Also, since the distribution function $F(x)$ gives the probability $P(X \leqq x)$, not $P(X > x)$, you have to look for the probability of the event of *not* exceding that "critical" x to be determined. Thus the equation for determining that x is

$$P(X \leqq x) = F(x) = 80\% = 0.8.$$

From this and (A) you have

$$\Phi\left(\frac{x - 1000}{100}\right) = 0.8. \tag{B}$$

From this and Table A8 in Appendix 5 you find, corresponding to 80%,

$$z(\Phi) = 0.842.$$

Because of (B) this means that $(x - 1000)/100 = 0.842$. Solving algebraically for x, you finally obtain the answer $x = 1000 + 0.842 \cdot 100 = 1084$. This means that the company should budget about 1084 hours (practically: 1100 hours) of monthly sick-leave time. Note that this rounded figure corresponds to the point $x = \mu + \sigma$, for which the normal probability to the right of it is 16%; see Fig. 489 on p. 1088. This gives you an opportunity of checking whether your result is reasonable.

Sec. 22.9 Distributions of Several Random Variables

Problem Set 22.9. Page 1099

1. **Probabilities for a two-dimensional random variable.** This is a continuous random variable and distribution. Ordinarily you would have to perform integration because in the general case such probabilities are given by (7) or other formulas involving double integrals. In the present problem, however, the density $f(x, y)$ is constant in the rectangle $4 \leqq x \leqq 10$, $0 \leqq y \leqq 5$, whose sides are 6 and 5, respectively. The volume under the surface of $f(x, y)$ (which is a portion of a plane over the rectangle and

coincides with the xy-plane outside the rectangle) equals the area $5 \cdot 6 = 30$ of the rectangle times the height k. This volume must equal 1; this is the two-dimensional analog of (10) in Sec. 22.5. Hence $k = 1/30$. The two probabilities can also be obtained by considering rectangles. $X \leqq 8$ and $3 \leqq Y \leqq 4$ corresponds to the rectangle $4 \leqq X \leqq 8$ and $3 \leqq Y \leqq 4$, which has width 4 and height 1, hence area 4, so that the first answer is

$$P(X \leqq 8,\ 3 \leqq Y \leqq 4) = \frac{1}{30} \cdot 4 = \frac{2}{15}.$$

Similarly, the portion of the rectangle $9 \leqq X \leqq 13$ and $0 \leqq Y \leqq 1$ inside the large rectangle in which $f(x, y)$ is not 0 is $9 \leqq X \leqq 10$ and $0 \leqq Y \leqq 1$ and has the area 1. This gives the probability

$$P(9 \leqq X \leqq 13,\ Y \leqq 1) = \frac{1}{30} \cdot 1 = \frac{1}{30}.$$

5. **Marginal density.** The given two-dimensional distribution (called the *uniform distribution in a rectangle*) has the density

$$f(x, y) = \frac{1}{(\beta_1 - \alpha_1)(\beta_2 - \alpha_2)}. \tag{A}$$

for (x, y) in the rectangle, and $f(x, y) = 0$ outside the rectangle. The marginal distributions of a continuous distribution are continuous. Hence they have a density. The density $f_2(y)$ of the marginal distribution of Y is obtained from (16) by integration over x. Since $f(x, y)$ is different from 0 only from $x = \alpha_1$ to $x = \beta_1$ (see Fig. 492), you have to integrate the constant function $f(x, y)$ over x between these limits. This gives $\beta_1 - \alpha_1$ times $f(x, y)$ in (A). But the factor $\beta_1 - \alpha_1$ cancels $\beta_1 - \alpha_1$ in the denominator of $f(x, y)$, so that you obtain

$$f_2(y) = \frac{1}{\beta_2 - \alpha_2} \quad \text{if} \quad \alpha_2 < y < \beta_2$$

and $f_2(y) = 0$ outside this y-interval. Hence the two-dimensional uniform distribution has as its marginal distribution with respect to y a one-dimensional uniform distribution (see Example 2 in Sec. 22.6, except for the notations). Similarly for the marginal distribution with respect to the other random variable, X.

7. **Addition of means and variances.** For adding means use Theorem 1, which is valid regardless of whether the random variables concerned are independent or dependent. You thus obtain for the mean thickness of a core

$$50 \cdot 0.5 + 49 \cdot 0.05 = 27.45 \text{ [mm]},$$

which equals a little over 1 in. The first term results from the metal layers and the second from the paper layers. To obtain the standard deviation of the cores you have to use the addition theorem for variances (Theorem 3). Hence from the given standard deviation you must first find the variance of a single metal sheet, which is $0.05^2 = 0.0025$ [mm^2], and of a single paper layer, which is $0.02^2 = 0.0004$ [mm^2]. Theorem 3 requires independence of the random variables concerned. It seems reasonable to assume that this requirement is (practically) satisfied in the present problem. Hence you obtain the variance of a core $50 \cdot 0.0025 + 49 \cdot 0.0004 = 0.1446$ [mm^2]. Note that the paper layers contribute to this value much less than the metal layers do. From this by taking the square root you obtain the standard deviation of the thickness of a core $\sqrt{0.1446} = 0.3803$ [mm].