Practice Problems

Problem:

Dive	Mark	Myrna
1	28	27
2	22	27
3	21	28
4	26	6
5	18	27

- Find the mean, median, mode, range?
- What can be said about this data?
- Which diver was more consistent?

Problem:

Dive	Mark	Myrna
1	28	27
2	22	27
3	21	28
4	26	6
5	18	27

Find the mean, median, mode, range?

mean	**23**	**23**
median	**22**	**27**
range	**10**	**22**

What can be said about this data?

Due to the outlier, the median is more typical of overall performancee.

Which diver was more consistent?

Dive	*Marks* x	$\bar{x}$	$x - \bar{x}$	$(x - \bar{x})^2$
1	28	23	5	25
2	22	23	-1	1
3	21	23	-2	4
4	26	23	3	9
5	18	23	-5	25
	Sum=115		0	Sum=64

Dive	*Myrna* x	$\bar{x}$	$x - \bar{x}$	$(x - \bar{x})^2$
1	27	23	4	16
2	27	23	4	16
3	28	23	5	25
4	6	23	-17	289
5	27	23	4	16
	Sum=115		0	Sum=362

Mark's Variance = 64 / 5 = 12.8
Myrna's Variance = 362 / 5 = 72.4

Conclusion: Mark has a lower variance therefore he is more consistent.

A study of the effects of smoking on sleep patterns is conducted. The measure observed is the time, in minutes, that it takes to fall asleep. These data are obtained:

Smokers:
69.3 56.0 22.1 47.6 53.2 48.1 52.7 34.4 60.2 43.8 23.2 13.8
Nonsmokers:
28.6 25.1 26.4 34.9 29.8 28.4 38.5 30.2 30.6 31.8 41.6 21.1 36.0 37.9 13.9

(a) Find the sample mean for each group.
(b) Find the sample standard deviation for each group.
(c) Make a dot plot of the data sets A and B on the same line.
(d) Comment on what kind of impact smoking appears to have on the time required to fall asleep.

Smokers: 69.3 56.0 22.1 47.6 53.2 48.1 52.7 34.4 60.2 43.8 23.2 13.8

Nonsmokers: 28.6 25.1 26.4 34.9 29.8 28.4 38.5 30.2 30.6 31.8 41.6 21.1 36.0 37.9 13.9

(a) Find the sample mean for each group.

(b) Find the sample standard deviation for each group.

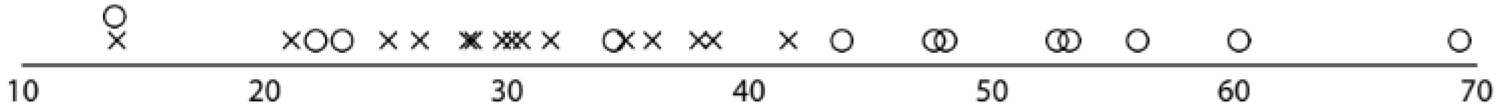
(c) Make a dot plot of the data sets A and B on the same line.

(d) Comment on what kind of impact smoking appears to have on the time required to fall asleep.

1.17 (a) $\bar{X}_{\text{smokers}} = 43.70$ and $\bar{X}_{\text{nonsmokers}} = 30.32$;

(b) $s_{\text{smokers}} = 16.93$ and $s_{\text{nonsmokers}} = 7.13$;

(c) A dot plot is shown below.



In the figure, "×" represents the nonsmoker group and "○" represents the smoker group.

(d) Smokers appear to take longer time to fall asleep and the time to fall asleep for smoker group is more variable.

1.30 Below are the lifetimes, in hours, of fifty 40-watt, 110-volt internally frosted incandescent lamps, taken from forced life tests:

919 1196 785 1126 936 918 1156 920 948 1067 1092 1162
1170 929 950 905 972 1035 1045 855 1195 1195 1340 1122
938 970 1237 956 1102 1157 978 832 1009 1157 1151 1009
765 958 902 1022 1333 811 1217 1085 896 958 1311 1037
702 923

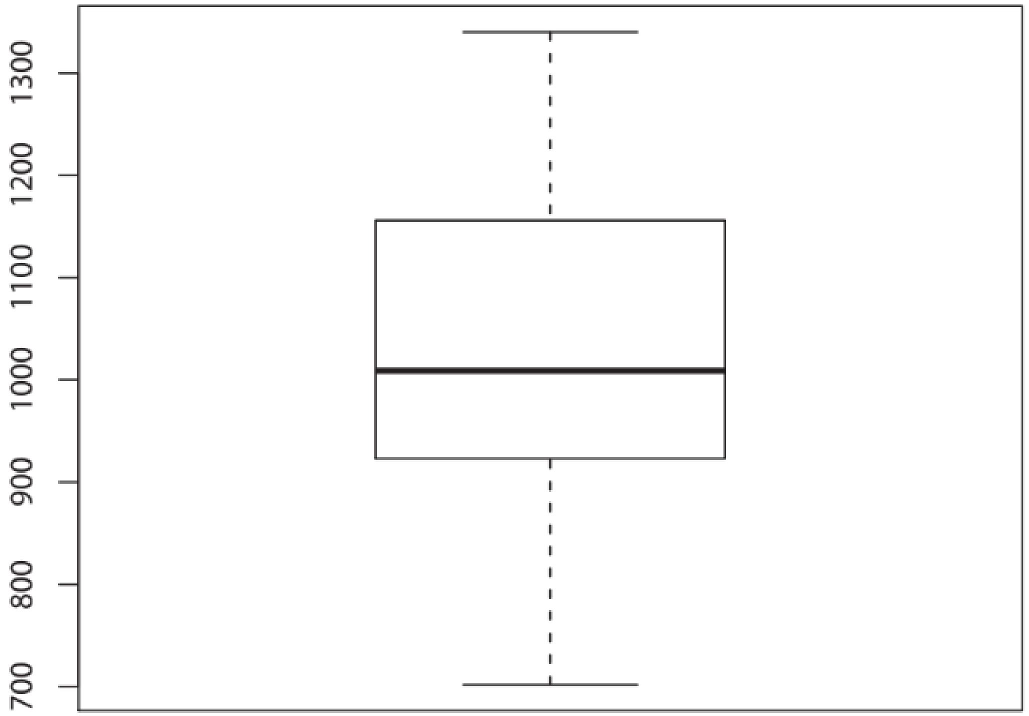
Construct a box plot for these data.

1.30 Below are the lifetimes, in hours, of fifty 40-watt, 110-volt internally frosted incandescent lamps, taken from forced life tests:

919 1196 785 1126 936 918 1156 920 948 1067 1092 1162 1170 929 950 905 972 1035 1045 855 1195 1195 1340 1122 938 970 1237 956 1102 1157 978 832 1009 1157 1151 1009 765 958 902 1022 1333 811 1217 1085 896 958 1311 1037 702 923

Construct a box plot for these data.

1.30 A box plot plot is shown next.



Problem:

1.28 Many manufacturing companies in the United States and abroad use molded parts as components of a process. Shrinkage is often a major problem. Thus, a molded die for a part is built larger than nominal size to allow for part shrinkage. In an injection molding study it is known that the shrinkage is influenced by many factors, among which are the injection velocity in ft/sec and mold temperature in ◦C. The following two data sets show the results of a designed experiment in which injection velocity was held at two levels (low and high) and mold temperature was held constant at a low level. The shrinkage is measured in cm × 10^4.

Shrinkage values at low injection velocity:
72.68 72.62 72.58 72.48 73.07
72.55 72.42 72.84 72.58 72.92
Shrinkage values at high injection velocity:
71.62 71.68 71.74 71.48 71.55
71.52 71.71 71.56 71.70 71.50

(a) Construct a dot plot of both data sets on the same graph. Indicate on the plot both shrinkage means, that for low injection velocity and high injection velocity.

(b) Based on the graphical results in (a), using the location of the two means and your sense of variability, what do you conclude regarding the effect of injection velocity on shrinkage at low mold temperature?

Problem:

Shrinkage values at low injection velocity:
72.68 72.62 72.58 72.48 73.07
72.55 72.42 72.84 72.58 72.92
Shrinkage values at high injection velocity:
71.62 71.68 71.74 71.48 71.55
71.52 71.71 71.56 71.70 71.50

(a) Construct a dot plot of both data sets on the same graph. Indicate on the plot both shrinkage means, that for low injection velocity and high injection velocity.

(b) Based on the graphical results in (a), using the location of the two means and your sense of variability, what do you conclude regarding the effect of injection velocity on shrinkage at low mold temperature?

1.28 (a) A dot plot is shown next.

In the figure, "×" represents the low-injection-velocity group and "○" represents the high-injection-velocity group.

(b) It appears that shrinkage values for the low-injection-velocity group is higher than those for the high-injection-velocity group. Also, the variation of the shrinkage is a little larger for the low injection velocity than that for the high injection velocity.

Problem:

Consider the situation of Exercise 1.28. But now use the following data set, in which shrinkage is measured once again at low injection velocity and high injection velocity. However, this time the mold temperature is raised to a high level and held constant.

Shrinkage values at low injection velocity:
76.20 76.09 75.98 76.15 76.17
75.94 76.12 76.18 76.25 75.82
Shrinkage values at high injection velocity:
93.25 93.19 92.87 93.29 93.37
92.98 93.47 93.75 93.89 91.62

(a) As in Exercise 1.28, construct a dot plot with both data sets on the same graph and identify both means (i.e., mean shrinkage for low injection velocity and for high injection velocity).

(b) As in Exercise 1.28, comment on the influence of injection velocity on shrinkage for high mold temperature. Take into account the position of the two means and the variability around each mean.

(c) Compare your conclusion in (b) with that in (b) of Exercise 1.28 in which mold temperature was held at a low level. Would you say that there is an interaction between injection velocity and mold temperature? Explain.

Problem:

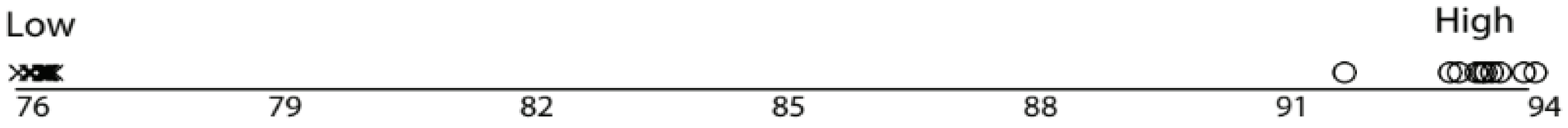
Shrinkage values at low injection velocity:
76.20 76.09 75.98 76.15 76.17
75.94 76.12 76.18 76.25 75.82
Shrinkage values at high injection velocity:
93.25 93.19 92.87 93.29 93.37
92.98 93.47 93.75 93.89 91.62

(a) As in Exercise 1.28, construct a dot plot with both data sets on the same graph and identify both means (i.e., mean shrinkage for low injection velocity and for high injection velocity).

(b) As in Exercise 1.28, comment on the influence of injection velocity on shrinkage for high mold temperature. Take into account the position of the two means and the variability around each mean.

(c) Compare your conclusion in (b) with that in (b) of Exercise 1.28 in which mold temperature was held at a low level. Would you say that there is an interaction between injection velocity and mold temperature? Explain.

1.31 (a) A dot plot is shown next.



In the figure, "×" represents the low-injection-velocity group and "∘" represents the high-injection-velocity group.

(b) In this time, the shrinkage values are much higher for the high-injection-velocity group than those for the low-injection-velocity group. Also, the variation for the former group is much higher as well.

(c) Since the shrinkage effects change in different direction between low mode temperature and high mold temperature, the apparent interactions between the mold temperature and injection velocity are significant.

The following are historical data on staff salaries (dollars per pupil) for 30 schools sampled in the eastern part of the United States in the early 1970s.

3.79 2.99 2.77 2.91 3.10 1.84 2.52 3.22

2.45 2.14 2.67 2.52 2.71 2.75 3.57 3.85

3.36 2.05 2.89 2.83 3.13 2.44 2.10 3.71

3.14 3.54 2.37 2.68 3.51 3.37

(a) Compute the sample mean and sample standard deviation.

(b) Construct a relative frequency histogram of the data.

(c) construct a box plot.

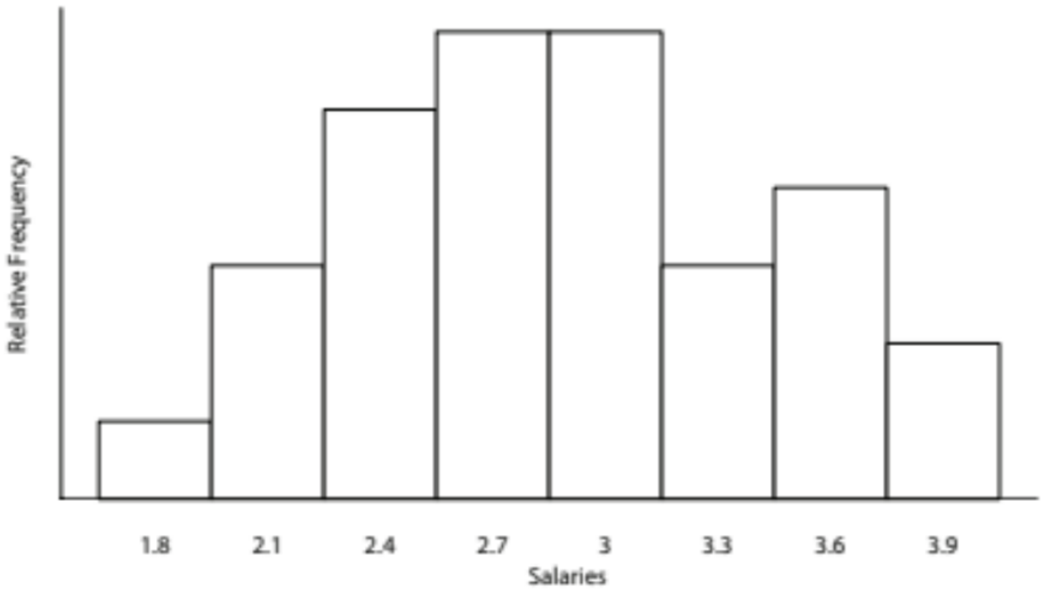
The following are historical data on staff salaries (dollars per pupil) for 30 schools sampled in the eastern part of the United States in the early 1970s.

3.79 2.99 2.77 2.91 3.10 1.84 2.52 3.22
2.45 2.14 2.67 2.52 2.71 2.75 3.57 3.85
3.36 2.05 2.89 2.83 3.13 2.44 2.10 3.71
3.14 3.54 2.37 2.68 3.51 3.37

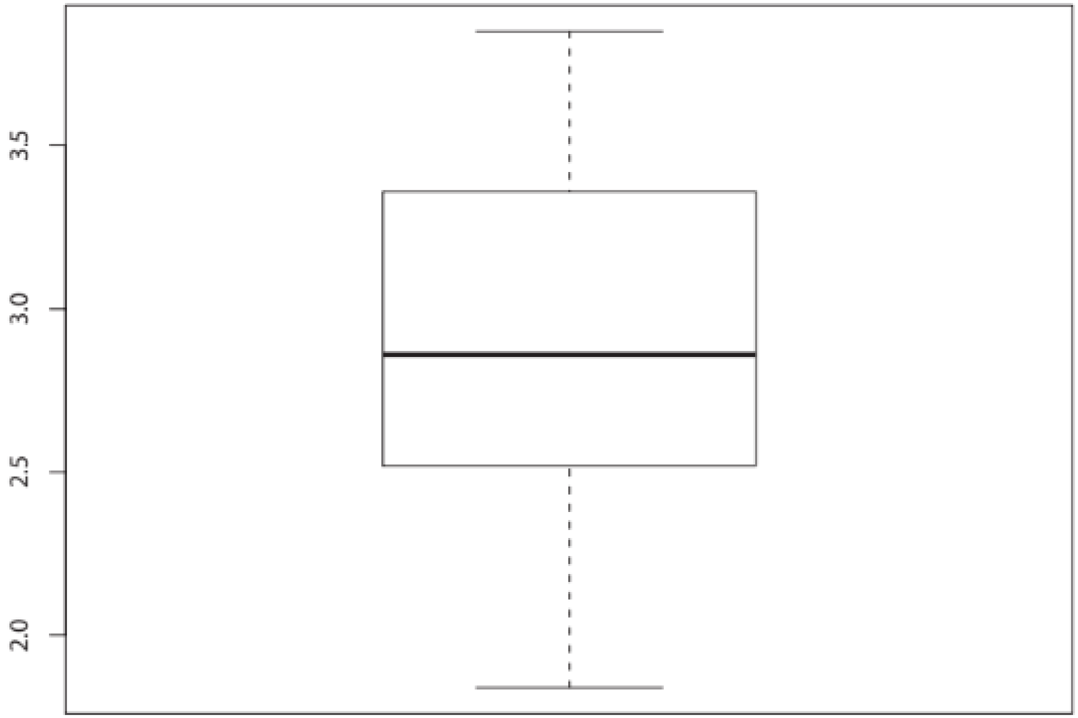
(a) Compute the sample mean and sample standard deviation.
(b) Construct a relative frequency histogram of the data.
(c) construct a box plot.

Solution: (a) $\bar{X} = 2.8973$ and $s = 0.5415$.

(b) A histogram plot is shown next.



(c) construct a box plot.



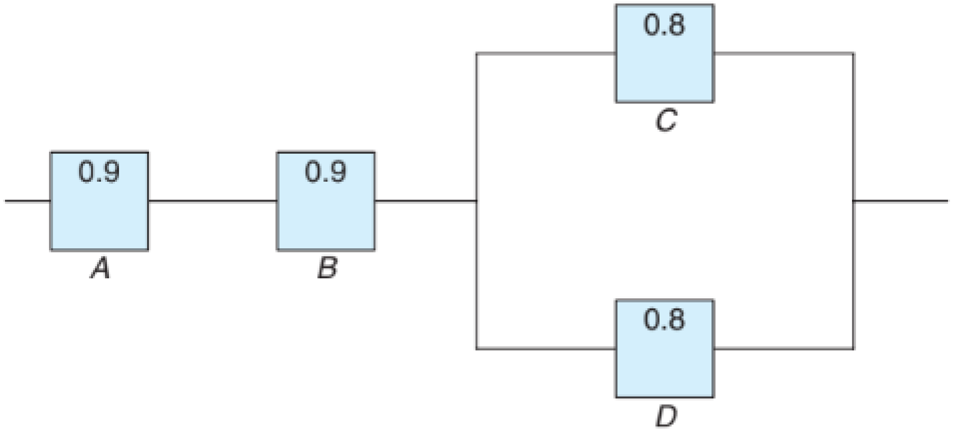
Problem: A small town has one fire engine and one ambulance available for emergencies. The probability that the fire engine is available when needed is 0.98, and the probability that the ambulance is available when called is 0.92. In the event of an injury resulting from a burning building, find the probability that both the ambulance and the fire engine will be available, assuming they operate independently.

Problem: A small town has one fire engine and one ambulance available for emergencies. The probability that the fire engine is available when needed is 0.98, and the probability that the ambulance is available when called is 0.92. In the event of an injury resulting from a burning building, find the probability that both the ambulance and the fire engine will be available, assuming they operate independently.

***Solution*:** Let A and B represent the respective events that the fire engine and the ambulance are available. Then

$$P(A \cap B) = P(A)P(B) = (0.98)(0.92) = 0.9016.$$

Example 2.39: An electrical system consists of four components as illustrated in Figure 2.9. The system works if components A and B work and either of the components C or D works. The reliability (probability of working) of each component is also shown in Figure 2.9. Find the probability that (a) the entire system works and (b) the component C does not work, given that the entire system works. Assume that the four components work independently.



An electrical system consists of four components as illustrated in Figure 2.9. The system works if components A and B work and either of the components C or D works. The reliability (probability of working) of each component is also shown in Figure 2.9. Find the probability that (a) the entire system works and (b) the component C does not work, given that the entire system works. Assume that the four components work independently.

Solution: In this configuration of the system, A, B, and the subsystem C and D constitute a serial circuit system, whereas the subsystem C and D itself is a parallel circuit system.

(a) Clearly the probability that the entire system works can be calculated as follows:

$$\begin{aligned} P[A \cap B \cap (C \cup D)] &= P(A)P(B)P(C \cup D) = P(A)P(B)[1 - P(C' \cap D')] \\ &= P(A)P(B)[1 - P(C')P(D')] \\ &= (0.9)(0.9)[1 - (1 - 0.8)(1 - 0.8)] = 0.7776. \end{aligned}$$

The equalities above hold because of the independence among the four components.

(b) To calculate the conditional probability in this case, notice that

$$\begin{aligned} P &= \frac{P(\text{the system works but } C \text{ does not work})}{P(\text{the system works})} \\ &= \frac{P(A \cap B \cap C' \cap D)}{P(\text{the system works})} = \frac{(0.9)(0.9)(1 - 0.8)(0.8)}{0.7776} = 0.1667. \end{aligned}$$

2.88 Before the distribution of certain statistical software, every fourth compact disk (CD) is tested for accuracy. The testing process consists of running four independent programs and checking the results. The failure rates for the four testing programs are, respectively, 0.01, 0.03, 0.02, and 0.01.

(a) What is the probability that a CD was tested and failed any test?

(b) Given that a CD was tested, what is the probability that it failed program 2 or 3?

(c) In a sample of 100, how many CDs would you expect to be rejected?

(d) Given that a CD was defective, what is the probability that it was tested?

2.88 Before the distribution of certain statistical software, every fourth compact disk (CD) is tested for accuracy. The testing process consists of running four independent programs and checking the results. The failure rates for the four testing programs are, respectively, 0.01, 0.03, 0.02, and 0.01.

(a) What is the probability that a CD was tested and failed any test?

(b) Given that a CD was tested, what is the probability that it failed program 2 or 3?

(c) In a sample of 100, how many CDs would you expect to be rejected?

(d) Given that a CD was defective, what is the probability that it was tested?

2.88 Consider the events:

F: failed the test,

P: passed the test.

(a) $P(\text{failed at least one tests}) = 1 - P(P_1P_2P_3P_4) = 1 - (0.99)(0.97)(0.98)(0.99) = 1 - 0.93 = 0.07$,

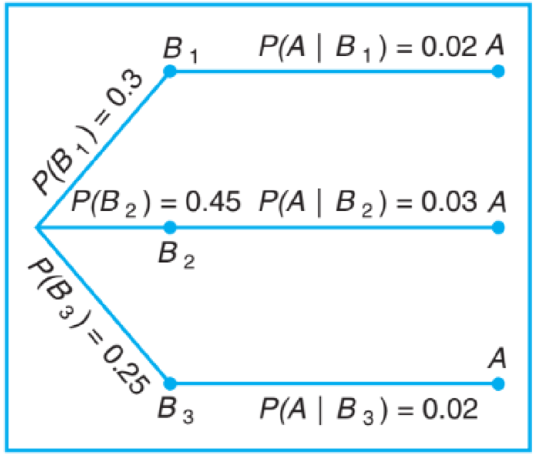
(b) $P(\text{failed 2 or 3}) = 1 - P(P_2P_3) = 1 - (0.97)(0.98) = 0.0494$.

(c) $100 \times 0.07 = 7$.

(d) 0.25.

In a certain assembly plant, three machines, B1, B2, and B3, make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective. Now, suppose that a finished product is randomly selected. What is the probability that it is defective?

In a certain assembly plant, three machines, B1, B2, and B3, make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective. Now, suppose that a finished product is randomly selected. What is the probability that it is defective?



***Solution*:** Consider the following events:

A: the product is defective,

B_1: the product is made by machine B_1,

B_2: the product is made by machine B_2,

B_3: the product is made by machine B_3.

Applying the rule of elimination, we can write

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3).$$

Referring to the tree diagram of Figure 2.15, we find that the three branches give the probabilities

$$P(B_1)P(A|B_1) = (0.3)(0.02) = 0.006,$$
$$P(B_2)P(A|B_2) = (0.45)(0.03) = 0.0135,$$
$$P(B_3)P(A|B_3) = (0.25)(0.02) = 0.005,$$

and hence

$$P(A) = 0.006 + 0.0135 + 0.005 = 0.0245.$$

A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varying times. In fact, plans 1, 2, and 3 are used for 30%, 20%, and 50% of the products, respectively. The defect rate is different for the three procedures as follows:
$P(D|P1)=0.01$, $P(D|P2)=0.03$, $P(D|P3)=0.02$,

where $P(D|Pj)$ is the probability of a defective product, given plan j. If a random product was observed and found to be defective, which plan was most likely used and thus responsible?

A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varying times. In fact, plans 1, 2, and 3 are used for 30%, 20%, and 50% of the products, respectively. The defect rate is different for the three procedures as follows:
P(D|P1)=0.01, P(D|P2)=0.03, P(D|P3)=0.02,

where P(D|Pj) is the probability of a defective product, given plan j. If a random product was observed and found to be defective, which plan was most likely used and thus responsible?

***Solution*:** From the statement of the problem

$$P(P_1) = 0.30, \quad P(P_2) = 0.20, \quad \text{and} \quad P(P_3) = 0.50,$$

we must find $P(P_j|D)$ for $j = 1, 2, 3$. Bayes' rule (Theorem 2.14) shows

$$\begin{aligned} P(P_1|D) &= \frac{P(P_1)P(D|P_1)}{P(P_1)P(D|P_1) + P(P_2)P(D|P_2) + P(P_3)P(D|P_3)} \\ &= \frac{(0.30)(0.01)}{(0.3)(0.01) + (0.20)(0.03) + (0.50)(0.02)} = \frac{0.003}{0.019} = 0.158. \end{aligned}$$

Similarly,

$$P(P_2|D) = \frac{(0.03)(0.20)}{0.019} = 0.316 \text{ and } P(P_3|D) = \frac{(0.02)(0.50)}{0.019} = 0.526.$$

The conditional probability of a defect given plan 3 is the largest of the three; thus a defective for a random product is most likely the result of the use of plan 3.

In the field of quality control, the science of statistics is often used to determine if a process is "out of control."
Suppose the process is, indeed, out of control and 20% of items produced are defective.
(a) If three items arrive off the process line in succession, what is the probability that all three are defective?
(b) If four items arrive in succession, what is the probability that three are defective?

(a) To calculate the probability that all three items are defective, we can use the multiplication rule of probability.
Let D denote a defective item, then the probability of getting a defective item is $P(D) = 0.20$, and the probability of getting a non-defective item is $P(N) = 0.80$.
The probability of getting three defective items in a row is:
$P(DDD) = P(D) \times P(D) \times P(D) = 0.20 \times 0.20 \times 0.20 = 0.008$
Therefore, the probability that all three items are defective is 0.008 or 0.8%.

(b) To calculate the probability that three out of four items are defective, we can use the combination formula:
Let's call the probability of getting a defective item P(D) and the probability of getting a non-defective item P(N). Then the probability of getting three defective items and one non-defective item in any order is:
$P(DDD*N) = C(4,3) \times P(D)^3 \times P(N) = 4 \times 0.20^3 \times 0.80 = 0.0512$
where $C(4,3) = 4$ is the number of ways to choose 3 out of 4 items.
Therefore, the probability that three out of four items are defective is 0.0512 or 5.12%.

Before the distribution of certain statistical software, every fourth compact disk (CD) is tested for accuracy. The testing process consists of running four independent programs and checking the results. The failure rates for the four testing programs are, respectively, 0.01, 0.03, 0.02, and 0.01.

(a) What is the probability that a CD was tested and failed any test?

(b) Given that a CD was tested, what is the probability that it failed program 2 or 3?

(c) In a sample of 100, how many CDs would you expect to be rejected?

(d) Given that a CD was defective, what is the probability that it was tested?

(a) The probability that a CD is tested and fails any test can be found by using the complement rule of probability. The complement of the event "a CD is tested and fails any test" is "a CD is not tested or passes all four tests."

The probability that a CD is not tested is 3/4. The probability that a CD passes all four tests is:

$(1 - 0.01) \times (1 - 0.03) \times (1 - 0.02) \times (1 - 0.01) = 0.938875$

Therefore, the probability that a CD was tested and failed any test is:

$1 - 0.938875 \times 3/4 = 0.07353125$

Before the distribution of certain statistical software, every fourth compact disk (CD) is tested for accuracy. The testing process consists of running four independent programs and checking the results. The failure rates for the four testing programs are, respectively, 0.01, 0.03, 0.02, and 0.01.

(a) What is the probability that a CD was tested and failed any test?

(b) Given that a CD was tested, what is the probability that it failed program 2 or 3?

(c) In a sample of 100, how many CDs would you expect to be rejected?

(d) Given that a CD was defective, what is the probability that it was tested?

(b) Given that a CD was tested, the probability that it failed program 2 or 3 can be found using Bayes' theorem:

P(failed program 2 or 3 | tested) = P(tested | failed program 2 or 3) × P(failed program 2 or 3) / P(tested)

We know that P(tested) = 1/4, P(failed program 2 or 3) = P(failed program 2) + P(failed program 3) - P(failed both) = 0.03 + 0.02 - (0.03 × 0.02) = 0.0494.

To find P(tested | failed program 2 or 3), we can use the conditional probability formula:

P(tested | failed program 2 or 3) = P(tested and failed program 2 or 3) / P(failed program 2 or 3)

We know that P(tested and failed program 2 or 3) = P(failed program 2 or 3) × P(failed program 1) × P(failed program 4) = 0.0494 × 0.01 × 0.01 = 0.0000494.

Therefore, P(tested | failed program 2 or 3) = 0.0000494 / 0.0494 = 0.001.

Plugging all the values into Bayes' theorem, we get:

P(failed program 2 or 3 | tested) = 0.001 × 0.0494 / 0.25 = 0.0001976

Therefore, given that a CD was tested, the probability that it failed program 2 or 3 is 0.01976%.

Before the distribution of certain statistical software, every fourth compact disk (CD) is tested for accuracy. The testing process consists of running four independent programs and checking the results. The failure rates for the four testing programs are, respectively, 0.01, 0.03, 0.02, and 0.01.

(a) What is the probability that a CD was tested and failed any test?

(b) Given that a CD was tested, what is the probability that it failed program 2 or 3?

(c) In a sample of 100, how many CDs would you expect to be rejected?

(d) Given that a CD was defective, what is the probability that it was tested?

(c) The expected number of CDs that would be rejected in a sample of 100 can be found by multiplying the probability of a CD being rejected by the total number of CDs in the sample:

Expected number of rejected CDs = 100 × P(rejected CD)

We know from part (a) that P(rejected CD) = 0.07353125, so:

Expected number of rejected CDs = 100 × 0.07353125 = 7.353125

Therefore, we would expect approximately 7 CDs to be rejected in a sample of 100.

Before the distribution of certain statistical software, every fourth compact disk (CD) is tested for accuracy. The testing process consists of running four independent programs and checking the results. The failure rates for the four testing programs are, respectively, 0.01, 0.03, 0.02, and 0.01.

(a) What is the probability that a CD was tested and failed any test?

(b) Given that a CD was tested, what is the probability that it failed program 2 or 3?

(c) In a sample of 100, how many CDs would you expect to be rejected?

(d) Given that a CD was defective, what is the probability that it was tested?

To find the probability that a CD was tested given that it was defective, we can use Bayes' theorem:

P(tested | defective) = P(defective | tested) × P(tested) / P(defective)

We know that P(defective) = 0.2, because 20% of the CDs produced are defective. To find P(tested), we can use the law of total probability:

P(tested) = P(tested | defective) × P(defective) + P(tested | non-defective) × P(non-defective)

We know that every fourth CD is tested, so P(tested | non-defective) = 0.25. Also, P(non-defective) = 0.8, because 80% of the CDs produced are not defective. Plugging in the values, we get:

P(tested) = 1/4 × 0.2 + 1/4 × 0.8 = 0.25

Now, we need to find P(defective | tested), which we calculated in part (d) of the previous question:

P(defective | tested) = 0.60003

Plugging in the values, we get:

P(tested | defective) = P(defective | tested) × P(tested) / P(defective) = 0.60003 × 0.25 / 0.2 = 0.7500375

Therefore, given that a CD is defective, the probability that it was tested is 75.00375%.