1 States, Events and Probabilities

Question 1:

You are given 13 cards chosen at random from a standard deck of 52 cards. "At random" means that all elementary events in this process are equally likely. What matters is the set of cards you get, not the order in which you get them. How many elementary events are there? What is the probability that you have all four aces?

Question 2:

In Bridge, a standard deck of 52 cards is dealt at random to four identifiable players (West, North, East and South). Once again, random means that all elementary events are equally likely, and for each player what matters is the set of cards he/she is dealt, not the order in which they are received. How many elementary events (deals) are there? What is the probability that West has all four aces?

Question 3:

For any two events A and B, prove that

$$Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B).$$

You are allowed to use only one property of probability, namely in the basic definition, Equation 2 on p. 3 of the class handout No. 1.

2 Conditional Probabilities and Bayes' Formula

Question 4:

Prove that $Pr(A_2 \cap A_3 | A_1) = Pr(A_3 | A_1 \cap A_2) * Pr(A_2 | A_1)$.

Question 5:

There are four distinct elementary events, A, B, C, D, with equal probabilities 1/4 each. Three other events are defined as follows:

$$X = A \cup D, \quad Y = B \cup D, \quad Z = C \cup D.$$

Show that any pair of X, Y, and Z constitutes two independent events, but all three jointly are not independent.

Question 6:

Suppose 2 percent of a population is affected by the HIV virus. A test is devised that produces 100 x percent false positives and 100 y percent false negatives. That is, if a person does not have the virus, the probability that the test comes out positive anyway is 100 x percent, and if a person does have the virus, the probability that the test comes out negative is 100 y percent. A randomly chosen person from the population takes the test. You want to ensure that the test is very effective at detecting the virus: if the test is negative, you want the chances that the person has the virus to be less than 0.1 percent. But you don't want to avoid frightening people without the virus: if the test is positive, the chances that the person does not have the virus should be less than 1 percent.What are the maximum permissible values of x and y consistent with your requirements?

3   Random Variables

Question 7:

A two fair dice are tossed. Define a random variable as the sum of the numbers that show on the two. What is the CDF of this random variable? What are the expected value and the variance of this random variable?

1   States, Events, and Probabilities

Question 1:

The first card can be chosen in 52 ways, the second in 51, the third in 50, . . . the 13th in 40. Thus there are

$$52 \ * \ 51 \ * \ \ldots * 40 = 52! \, / \, 39!$$

possibilities. But the order does not matter, so appropriate collections 13! of these possibilities comprise just one elementary event for our purpose. Thus there are

$$\frac{52!}{13! \; 39!} \,, \quad \text{denoted by} \quad \binom{52}{13}$$

elementary events.

The set of elementary events in which you have all four aces can be found by giving you the four aces, and then 9 randomly selected cards out of the remaining 48. There are

$$\binom{48}{9} = \frac{48!}{9! \; 39!}$$

elementary events in this event. Therefore the probability is

$$\binom{48}{9} \Big/ \binom{52}{13} = \frac{48!}{9! \; 39!} \; \frac{13! \; 39!}{52!} = \frac{48! \; 13!}{9! \; 52!} = \frac{13 \ * \ 12 \ * \ 11 \ * \ 10}{52 \ * \ 51 \ * \ 50 \ * \ 49} \approx 0.00264$$

Question 2:

Since the order does not matter, we can think of the process as starting with West being given 13 cards out of the 52, then North being given 13 cards out of the remaining 39, then East being given 13 cards out of the remaining 26. This fixes the 13 cards South gets. Thus there are

$$\binom{52}{13} \binom{39}{13} \binom{26}{13} = \frac{52!}{13! \; 39!} \; \frac{39!}{13! \; 26!} \; \frac{26!}{13! \; 13!} = \frac{52!}{(13!)^4}$$

elementary events.

To characterize the event where West has all four aces, start by giving West all four aces. From the remaining cards, give West 9 at random, and divide the 39 that are left randomly giving 13 each to North, East, and South. There are

$$\binom{48}{9} \binom{39}{13} \binom{26}{13} = \frac{48!}{9! \; 39!} \; \frac{39!}{13! \; 26!} \; \frac{26!}{13! \; 13!} = \frac{48!}{9! \; (13!)^3}$$

elementary events of this kind. Therefore the probability that West gets all four aces is

$$\frac{48!}{9!\,(13!)^3}\,\frac{(13!)^4}{52!} = \frac{48!\;13!}{9!\;52!}\,.$$

The numerical answer is the same as in Question 1, but the situations or "experiments" are different, so the underlying sample spaces, elementary events etc. are also different.

It is intuitive that the probabilities in problems 1 and 2 should be the same, since the probability that the 13 cards dealt to one designated player ("you" in Question 1, "West" in Question 2) have some stated property shouldn't depend on how, or even whether, the remaining 13 cards are dealt to the other three players.

Question 3:

Define $C = A - (A \cap B)$, that is, the set of all elementary events that are in A but not in $A \cap B$. Then

$$A = (A \cap B) \cup C, \qquad (A \cap B) \cap C = \emptyset\,.$$

Therefore by the definition,

$$Pr(A) = Pr(A \cap B) + Pr(C)\,.$$

Similary, define $D = B - (A \cap B)$ to get

$$Pr(B) = Pr(A \cap B) + Pr(D)\,.$$

Finally, a Venn diagram shows that C, D, and $(A \cap B)$ are pairwise disjoint events, and

$$A \cup B = C \cup D \cup (A \cap B)\,.$$

Therefore

$$Pr(A \cup B) = Pr(C) + Pr(D) + Pr(A \cap B)$$

Combining these three equations, the result follows.

2 Conditional Probabilities and Bayes' Formula

Question 4:

By the definition of conditional probabilities (equation (3) of class handout No. 1:)

$$Pr(A_2 \cap A_3 | A_1) = \frac{Pr\big((A_2 \cap A_3) \cap A_1\big)}{Pr(A_1)} = \frac{Pr(A_1 \cap A_2 \cap A_3)}{Pr(A_1)}\,,$$

$$Pr(A_3 | A_1 \cap A_2) = \frac{Pr\big(A_3 \cap (A_1 \cap A_2)\big)}{Pr(A_1 \cap A_2)} = \frac{Pr(A_1 \cap A_2 \cap A_3)}{Pr(A_1 \cap A_2)}\,,$$

$$Pr(A_2 | A_1) = \frac{Pr(A_2 \cap A_1)}{Pr(A_1)} = \frac{Pr(A_1 \cap A_2)}{Pr(A_1)}\,.$$

The result follows immediately.

Question 5:

A and D are distinct elementary events, so $A \cap D = \emptyset$. Therefore $Pr(X) = Pr(A) + Pr(D) = 1/2$. Similarly for Y and Z.

Next, $X \cap Y = D$, so $Pr(X \cap Y) = 1/4$. Similarly for $Y \cap Z$ and $X \cap Z$.

Thus $Pr(X \cap Y) = Pr(X) * Pr(Y)$, that is, X and Y are independent. Similarly for the pairs Y and Z, X and Z.

But $X \cap Y \cap Z = D$, so $Pr(X \cap Y \cap Z) = 1/4 \neq 1/8 = Pr(X) * Pr(Y) * Pr(Z)$.

To remember this example, suppose a tetrahedron has three of its faces each given a single distinct color, red, green and blue. The fourth face is split into three regions, one colored red, the second green, and the third blue. A face is chosen at random. Then X stands for the event "the chosen face has some red on it," and similarly for Y and Z.

Question 6:

This is an application of Bayes' Formula. Use the table from the class handout:

		Test outcome	
		Positive	Negative
True HIV	Present	0.02 * $(1 - y)$	0.02 * y
Status	Absent	0.98 * x	0.98 * $(1 - x)$
	Sum over rows	$0.02\,(1 - y) + 0.98\,x$	$0.02\,y + 0.98\,(1 - x)$

So we want

$$\frac{0.02 * y}{0.02\,y + 0.98\,(1 - x)} \leq 0.001\,, \qquad \frac{0.98 * x}{0.02\,(1 - y) + 0.98\,x} \leq 0.01\,.$$

This translates to

$$49\,x + 999\,y \leq 49, \qquad 4851\,x + y \leq 1\,.$$

This has many possible solutions; the values that satisfy both with equality are

$$x = 1.9603 * 10^{-4}, \qquad y = 4.9039 * 10^{-2}\,.$$

So the test needs to be made much more accurate as regards avoiding false positives. This is so even though the error where a person with a negative test is actually infected is 10 times as important to avoid as the opposite kind of error. The reason is that the fraction of truly affected people in the population is very small.

3   Random Variables

Question 7:

This random variable can take on any values between 2 and 12. The sum can be 2 in only one way (one elementary event), namely when each die shows 1. The sum can be 3 in two

ways $(1 + 2$ and $2 + 1)$. Similarly, the number of elementary events for the sum to be 4, 5, 6, ... 12 is respectively 3, 4, 5, 6, 5, 4, 3, 2, 1. So, for the random variable X, we have

$$Pr(X < 2) = 0; \qquad Pr(X = 2) = 1/36, \qquad Pr(X = 3) = 2/36, \ldots .$$

The CDF over the real line is therefore a step function:

$$
\begin{aligned}
F(t) \;&=\; 0 \quad \text{for } t < 2 \\
&=\; 1/36 \quad \text{for } 2 \le t < 3 \\
&=\; 3/36 \quad \text{for } 3 \le t < 4 \\
&=\; 6/36 \quad \text{for } 4 \le t < 5 \\
&=\; 10/36 \quad \text{for } 5 \le t < 6 \\
&=\; 15/36 \quad \text{for } 6 \le t < 7 \\
&=\; 21/36 \quad \text{for } 7 \le t < 8 \\
&=\; 26/36 \quad \text{for } 8 \le t < 9 \\
&=\; 30/36 \quad \text{for } 9 \le t < 10 \\
&=\; 33/36 \quad \text{for } 10 \le t < 11 \\
&=\; 35/36 \quad \text{for } 11 \le t < 12 \\
&=\; 36/36 = 1 \quad \text{for } 12 \le t
\end{aligned}
$$

The expected value is

$$
\begin{aligned}
\mathrm{E}[X] \;&=\; 2\,\frac{1}{36} + 3\,\frac{2}{36} + 4\,\frac{3}{36} + 5\,\frac{4}{36} + 6\,\frac{5}{36} + 7\,\frac{6}{36} \\
&\quad +8\,\frac{5}{36} + 9\,\frac{4}{36} + 10\,\frac{3}{36} + 11\,\frac{2}{36} + 12\,\frac{1}{36} \\
&=\; \frac{2 + 6 + 12 + 20 + 30 + 42 + 40 + 36 + 30 + 22 + 12}{36} = \frac{252}{36} = 7
\end{aligned}
$$

If you had guessed the answer by intuition, you could prove it more easily:

$$
\begin{aligned}
\mathrm{E}[X] - 7 \;&=\; (-5)\,\frac{1}{36} + (-4)\,\frac{2}{36} + (-3)\,\frac{3}{36} + (-2)\,\frac{4}{36} + (-1)\,\frac{5}{36} \\
&\quad + 1\,\frac{5}{36} + 2\,\frac{4}{36} + 3\,\frac{3}{36} + 4\,\frac{2}{36} + 5\,\frac{1}{36} = 0\,.
\end{aligned}
$$

The variance is

$$
\begin{aligned}
\mathrm{V}[X] \;&=\; (2-7)^2\,\frac{1}{36} + (3-7)^2\,\frac{2}{36} + (4-7)^2\,\frac{3}{36} + (5-7)^2\,\frac{4}{36} + (6-7)^2\,\frac{5}{36} + (7-7)^2\,\frac{6}{36} \\
&\quad +(8-7)^2\,\frac{5}{36} + (9-7)^2\,\frac{4}{36} + (10-7)^2\,\frac{3}{36} + (11-7)^2\,\frac{2}{36} + (12-7)^2\,\frac{1}{36} \\
&=\; 2\left[25\,\frac{1}{36} + 16\,\frac{2}{36} + 9\,\frac{3}{36} + 4\,\frac{4}{36} + 1\,\frac{5}{36}\right] \\
&=\; 2\,\frac{25 + 32 + 27 + 16 + 5}{36} = \frac{210}{36} = 5.8333\,.
\end{aligned}
$$

So the standard deviation (square root of the variance) is 2.4152 .