The University of Sydney
School of Mathematics and Statistics

Solutions to Tutorial Week 4

MATH1905: Statistics (Advanced) Semester 2, 2017

Web Page: `http://sydney.edu.au/science/maths/MATH1905`
Lecturer: Michael Stewart

Be sure to complete questions 7 and 8 from week 3 and discuss any difficulties with your tutor.

1. (**Multiple Choice**) A six-sided die is loaded in such a way that each of the (equally likely) even numbers is twice as likely to occur as each of the (equally likely) odd numbers. The die is thrown (independently) two times. The probability that a total of 5 is observed is:

(a) 1/9 (b) 4/81 (c) 2/9 (d) 2/81 (e) 8/81

Solution: Interpret the question to mean that for some value v, at a *single roll* we have

No.	1	2	3	4	5	6	Total
Prob.	v	$2v$	v	$2v$	v	$2v$	$9v = 1$

So $v = \frac{1}{9}$, in fact for a single roll we have

No.	1	2	3	4	5	6
Prob.	1/9	2/9	1/9	2/9	1/9	2/9

Now the probability of obtaining a total of 5 in 2 rolls is

$$\begin{aligned} P(\text{Total} = 5) &= P(1,4) + P(2,3) + P(3,2) + P(4,1) \\ &= 2\left\{P(1)P(4) + P(2)P(3)\right\} \\ &= 2\left[\left(\frac{1}{9} \times \frac{2}{9}\right) + \left(\frac{2}{9} \times \frac{1}{9}\right)\right] \\ &= \frac{8}{81}\,. \end{aligned}$$

So the correct answer is (e).

2. An unknown number N of animals of a certain species are present in a certain habitat. To try to "estimate" N, 10 animals are captured and tagged. Some time later (to allow the tagged animals to "randomly mix" in the habitat) a further 10 animals are captured and it turns out only 2 of these 10 had tags.

(a) Write the probability of getting 2 tagged animals in the second sample as a function of N.

Solution: Since there are 10 tagged and $N - 10$ untagged in the habitat, *assuming each sample of size 10 is equally likely* (suggested by the phrase "randomly mix"), the desired probability (according to the "classical definition") is the ratio

$$\frac{\text{No. different samples of size 10 with 2 tagged and 8 untagged}}{\text{Total no. different samples of size 10}}$$

The denominator is simply $\binom{N}{10}$; to see this note that this is the number of subsets of (i.e. samples taken without replacement from) $\{1, 2, \ldots, N\}$ of size 10, and we can establish a one-to-one correspondence between the integers 1,2,...,N and the animals in the habitat.

As for the numerator, we can evaluate this by outlining a two-step procedure for constructing such a sample and then appeal to the multiplication principle. Suppose the tagged animals are numbered 1,2,...,10 and that the untagged are numbered 11,12,...,N. We construct an appropriate sample by

1. choosing 2 elements from $\{1, 2, \ldots, 10\}$, which can be done in $\binom{10}{2}$ ways;
2. choosing 8 elements from $\{11, 12, \ldots, N\}$; the number of ways of doing this is the same as choosing 8 from $\{1, 2, \ldots, N-10\}$, which can be done in $\binom{N-10}{8}$ ways.

Thus the total number of ways of constructing an appropriate sample is the product $\binom{10}{2}\binom{N-10}{8}$ and so the desired probability is the ratio

$$\frac{\binom{10}{2}\binom{N-10}{8}}{\binom{N}{10}}.$$

(b) What is the smallest possible value for N?

Solution: Note that 18 distinct animals have "been seen": the 10 from the original sample plus the 8 untagged animals in the follow-up sample. Thus there must be at least 18 in the habitat.

The same conclusion can be reached in a slightly different way by noting that the number of either tagged or untagged in the sample can be no more than the corresponding number in the population. Focussing on the untagged animals we have

- 8 in the sample and
- $N - 10$ in the population,

thus the inequality

$$8 \leq N - 10$$

holds, that is $N \geq 18$.

3. In a large factory there are four machines M_1, M_2, M_3 and M_4 all producing identical items for computers. The production from all four machines for a day is collected in a large storage bin. The quality control manager has noted the following information on percentages of daily production and defective items produced by each machine:

Machine:	M_1	M_2	M_3	M_4
Production:	10%	20%	30%	40%
Defectives:	0.1%	0.05%	0.50%	0.20%

An inspector selects an item at random from the storage bin in a particular day.

(a) What is the probability that the selected item is defective?

Solution: We can interpret the information given to mean that, with

$$\begin{aligned} D &= \text{“selected item is defective” and, for } j = 1, 2, 3, 4, \\ M_j &= \text{“selected item is from machine } j\text{”} \end{aligned}$$

and also that

$$\begin{aligned} P(M_1) &= 0.1\,, \quad P(D|M_1) = 0.001\,, \\ P(M_2) &= 0.2\,, \quad P(D|M_2) = 0.0005\,, \\ P(M_3) &= 0.3\,, \quad P(D|M_3) = 0.005\,, \\ P(M_4) &= 0.4\,, \quad P(D|M_4) = 0.002\,. \end{aligned}$$

Since

$$D = (D \cap M_1) \cup (D \cap M_2) \cup (D \cap M_3) \cup (D \cap M_4)\,,$$

and these are mutually exclusive we have

$$P(D) = P(D \cap M_1) + P(D \cap M_2) + P(D \cap M_3) + P(D \cap M_4)\,.$$

Also for each $j = 1, 2, 3, 4$,

$$P(D \cap M_j) = P(M_j)P(D|M_j)\,.$$

So

$$P(D) = 0.0001 + 0.0001 + 0.0015 + 0.0008 = 0.0025\,.$$

(b) What is the conditional probability that the selected item is from machine M_4 given that it is defective?

Solution:

$$P(M_4|D) = \frac{P(D \cap M_4)}{P(D)} = \frac{0.0008}{0.0025} = 0.32\,.$$

4. Suppose A and B are two independent events. Noting that $A = (A \cap B) \cup (A \cap B^c)$ prove (using the third axiom) that A and B^c are also independent.

Solution: Since $A \cap B$ and $A \cap B^c$ are mutually exclusive, by the third axiom

$$P(A) = P\{(A \cap B) \cup (A \cap B^c)\} = P(A \cap B) + P(A \cap B^c)\,.$$

Therefore

$$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) \\ &= P(A) - P(A)P(B) \quad \text{(by independence)} \\ &= P(A)\,[1 - P(B)] \\ &= P(A)P\,(B^c) \end{aligned}$$

and so we see A and B^c are independent.

5. A fair coin is flipped 3 times and a fair 6-sided die is rolled twice, all independently. Let X equal the number of heads obtained in the coin flips *plus* the number obtained in the die rolls. Find $P(X = 10)$.

Solution: Let Y = no. heads and Z = total showing on the 2 dice. Then by independence

$$\begin{aligned} P(X = 10) &= P(Y = 0, Z = 10) + P(Y = 1, Z = 9) + P(Y = 2, Z = 8) + P(Y = 3, Z = 7) \\ &= P(Y = 0)P(Z = 10) + P(Y = 1)P(Z = 9) + P(Y = 2)P(Z = 8) + P(Y = 3)P(Z = 7)\,. \end{aligned}$$

As shown in lectures we have

y	0	1	2	3
$P(Y = y)$	$1/8$	$3/8$	$3/8$	$1/8$

Also for rolling 2 fair (6-sided) dice independently, we have the same setup as for questions 6 and 7 in the week 3 Tutorial , that is 36 ordered pairs

$$\begin{matrix} (1,1) & \cdots & (1,6) \\ \vdots & \ddots & \vdots \\ (6,1) & \cdots & (6,6) \end{matrix}$$

that are equally likely. So

$$\begin{aligned} P(Z = 10) &= P(4,6) + P(5,5) + P(6,4) = \frac{3}{36}\,; \\ P(Z = 9) &= P(3,6) + P(4,5) + P(5,4) + P(6,3) = \frac{4}{36}\,; \\ P(Z = 8) &= P(2,6) + P(3,5) + P(4,4) + P(5,3) + P(6,3) = \frac{5}{36}\,; \\ P(Z = 7) &= P(1,6) + P(2,5) + P(3,4) + P(4,3) + P(5,3) + P(6,3) = \frac{6}{36}\,. \end{aligned}$$

So

$$P(X = 10) = \left(\frac{1}{8} \times \frac{3}{36}\right) + \left(\frac{3}{8} \times \frac{4}{36}\right) + \left(\frac{3}{8} \times \frac{5}{36}\right) + \left(\frac{1}{8} \times \frac{6}{36}\right) = \frac{3 + 12 + 15 + 6}{8 \times 36} = \frac{1}{8}\,.$$

6. *Does a flush beat a straight?* In the card game of Poker, all possible subsets of size 5 (called "hands") from the 52 standard playing cards[1] are classified into different types. The *more difficult* (i.e. unlikely) the type of hand is to get, the higher it is ranked in the hierarchy of hands.

Two such types of hands are a *flush* (where all cards are of the same suit) and a *straight* (where the numerical values on the cards are in sequence, e.g. 2,3,4,5,6 or 8,9,10,J,Q *irrespective of the suits.*

(a) Count how many possible hands give a flush by enumerating how many ways each of the following steps may be performed:

(i) choose a suit;

Solution: 4

(ii) choose 5 of the 13 cards in that suit.

Solution: $\binom{13}{5}$.
So there are

$$4\binom{13}{5} = 4 \times \frac{13 \times 12 \times 11 \times 10 \times 9}{5 \times 4 \times 3 \times 2 \times 1} = 4 \times 13 \times 11 \times 9 = 5,148$$

different ways to get a flush.

(b) Count how many possible hands give a straight by enumerating how many ways each of the following steps may be performed:

(i) choose a lowest numerical value;

Solution: The possible "lowest numerical values" are A,2,. . . ,10, so there are 10 ways to choose the "lowest numerical value".

(ii) for each value, choose 1 of the 4 possible cards of that value.

Solution: There are 4 ways to choose the suit for each of the 5 numerical values, so the total number of ways to choose these is 4^5. Thus there are

$$10 \times 4^5 = 10,240$$

different ways to get a straight.

(c) Does a flush beat a straight?

Solution: Since a flush is *harder* to get, it beats a straight.

Note: these totals both include the 40 different possible "straight flushes", which in turn include the 4 "royal flushes".

[1] There are 13 numerical values A,2,3,. . . ,9,10,J,Q,K in each of 4 suits. Note that A can also be "high", that is a straight can be 10,J,Q,K,A.