# MATH181 Solution Sheet 1

**1.** *A set of students obtains the marks: 70, 75,50, 25, and 81. Compute the mean and standard deviation of the student's marks.*

There are 5 students. From the definitions, the mean is

$$\bar{x} = \frac{1}{5}(70 + 75 + 50 + 25 + 81) = \underline{\underline{60.2}}$$

Variance is

$$\sigma^2 = \frac{1}{5}\left((70 - 60.2)^2 + (75 - 60.2)^2 + (50 - 60.2)^2 + (25 - 60.2)^2 + (81 - 60.2)^2\right) = 418.16$$

The standard deviation is

$$\sigma = \sqrt{\sigma^2} = \underline{\underline{20.45}}$$

As a rule-of-thumb, we expect about $\frac{2}{3}$ of the students to score between $\bar{x} - \sigma$ and $\bar{x} + \sigma$.

**2.** *A tennis club has 10 members of which 7 are men and 3 are women. Determine the number of ways in which a committee of 3 can be chosen.*

The committee of three people from a pool of 10 can be chosen in ${}^{10}C_3$ ways.

$$\underline{\underline{{}^{10}C_3 = \frac{10!}{3!7!} = 120}}$$

*Also find the number of ways that the committee can be chosen if it includes at least one man.*

To find the number of committees with at least one man, we have to remove the all-female committees from the list. Since there are 3 women in the club, there is only ${}^{3}C_3 = 1$ all-female committee. Thus, the number of committees with at least one man is

$$\underline{\underline{{}^{10}C_3 - {}^{3}C_3 = 120 - 1 = 119}}$$

*How many possible committees are there with at least one woman?*

Likewise, the easiest way to find how many committees have at least one woman, is to count the possible all male committees, and take them away from the total. Since there are 7 men, the number to subtract is

$${}^{7}C_3 = \frac{7!}{3!4!} = 35$$

and the answer is

$$\underline{\underline{{}^{10}C_3 - {}^{7}C_3 = 120 - 35 = 85}}$$

**3.** *A pencil manufacturer finds that 4% of the pencils produced are defective. These defects occur randomly in the manufacturing process.*

These probabilities can all be found from the binomial distribution. We use the expansion

$$(a+b)^6 = b^6 + 6ab^5 + 15a^2b^4 + 20a^3b^3 + 15a^4b^2 + 6a^5b + a^6 = \sum_{k=0}^{6} {}^6C_k\, a^k\, b^{6-k}$$

with $a = 0.04$ and $b = 1 - 0.04 = 0.96$.

*Determine the probability that in a sample of 6 pencils*

(i) *all the pencils are defective;*

This probability is

$$\underline{\underline{P(6 \text{ defective}) = a^6 = (0.04)^6 = 4.096 \times 10^{-9}}}$$

(ii) *all the pencils are good;*

$$\underline{\underline{P(0 \text{ defective}) = b^6 = (0.96)^6 = 0.783}}$$

(iii) *exactly 2 pencils are defective;*

This is the probability corresponding to $15a^2b^4$ in the binomial:

$$P(2 \text{ defective}) = {}^6C_2\,(0.04)^2\,(0.96)^4 = \underline{\underline{15\,(0.04)^2\,(0.96)^4 = 0.0204}}$$

(iv) *fewer than 3 pencils are defective.*

We have to add probabilities,

$$\begin{aligned} P(\text{ fewer than 3 defective}) &= P(0 \text{ defective}) + P(1 \text{ defective}) + P(2 \text{ defective}) \\ &= b^6 + 6ab^5 + 15a^2b^4 = 0.7828 + 0.1957 + 0.0204 = \underline{\underline{0.9988}} \end{aligned}$$

**4.** *The letters FIIINNTY are written on 8 otherwise identical cards. The cards are chosen randomly one at a time and placed in order. What is the probability that they spell INFINITY?*

In the word **INFINITY** there are three **I**s and two **N**s. All the other letters only occur once. There are eight letters in the word **INFINITY**.

Method 1

Hence the probability of choosing **I** first out of the bag is $\frac{3}{8}$. The probability of next getting **N** is then $\frac{2}{7}$. This allows me to build up the probability of getting the word **INFINITY** as:

$$p = \frac{3}{8}\cdot\frac{2}{7}\cdot\frac{1}{6}\cdot\frac{2}{5}\cdot\frac{1}{4}\cdot\frac{1}{3}\cdot\frac{1}{2}\cdot\frac{1}{1} = \underline{\underline{\frac{3!\,2!}{8!} = \frac{1}{3360} = 2.98 \times 10^{-4}}}$$

Method 2

There are 8! orders in which we can pull the letters out of the bag, all equally likely. The number which correctly spell INFINITY are $3! \cdot 2!$ because there are 3! orders in which we could take out the three $I$ cards, and 2! orders for the two cards with $N$ on.

$$p = \frac{\text{successful outcomes}}{\text{all outcomes}} = \underline{\underline{\frac{3!\,2!}{8!} = \frac{1}{3360} = 2.98 \times 10^{-4}}}$$

**5.** *Five seeds are weighed, the results (in grams) are*

$$20.97, \ 19.17, \ 19.47, \ 21.60, \ 19.92$$

*What is the mean and standard deviation of these weights?*

The mean weight is $(20.97 + 19.17 + 19.47 + 21.60 + 19.92)/5 = \underline{\underline{20.23 \text{ g}}}$.

The standard deviation is $\underline{\underline{\sigma = 0.92 \text{ g}}}$.

*What error should you give on the mean weight of a seed from this sort of plant?*

By chance, the mean of our sample of 5 seeds might be a little higher or lower than the average over a really large sample with millions of seeds. Our best guess for the "sampling error" with a sample of size $N$ is

$$\varepsilon = \frac{\sigma}{\sqrt{N-1}} = \frac{0.92}{\sqrt{5-1}} = \underline{\underline{0.46 \text{ g}}}$$

so we should say that

<u>the average weight of a seed is $20.23 \pm 0.46$ g</u>.

This is a relative error of 2.3%.

(It's OK if you divided by $\sqrt{N}$ instead of $\sqrt{N-1}$.)

*How many seeds must I weigh to determine the mean to 1%? How many to know the mean to 1 part in a thousand?*

A 1% error in a quantity means that the error should be 100 times smaller than the quantity itself. Since our seeds average 20.2 g, a 1% error is 0.202 g. However, almost everybody thought that a 1% error meant an error of 0.01 g.

As we increase the sample size, the individual weights will still scatter by about the same amount, so the standard deviation will not change much. The estimated error will drop, because of the square-root factor, and our average will become more representative of the true average over all seeds of this species.

The error $\varepsilon$ drops like $1/\sqrt{N}$, to get the error down from 2.3% to 1% we need to increase our sample size by a factor $2.3^2 = 5.2$, so we need about <u>26 seeds for a 1% error.</u>

(You might get a slightly different answer, depending on where you used $\sqrt{N}$ or $\sqrt{N-1}$, but it should be about this size). The sample size needed for a given error depends on the standard deviation - if we are looking at a quantity where the individual measurements scatter more, we would need a bigger sample.

*How many to know the mean to 1 part in a thousand?*

Decreasing the error by a factor of 10 means increasing the sample by a factor of 100. We need $\approx 2\,600$ seeds for an error of 1 part in $10^3$.