SCIE1000, Solutions to Tutorial Week 6.

(1) 1. Here are some important points.

   - 'Flukes' or 'coincidences' can occur, but we are more likely to hear about these flukes than results that are not interesting.
   - The published literature is a biased sample of all studies ever completed, as negative results and repetitions of previous findings are unlikely to be published.
   - About half of the negative results go missing and positive results are twice as likely to be published
   - No field of medicine (or science) is immune from of publication bias.
   - This problem is compounded by fake fixes:
     * Trial registers - people didn't use them
     * Journals did not enforce their rules about refusing to publish results of trials that were not registered before the trial was commenced.
     * Only 1 in 5 trials have their results posted online.
   - We cannot know the effect of medicines if we do not have access to all of the information.
   - The consequences of not publishing negative results can be many thousands of unnecessary deaths.
   - The best solution would be for all trials conducted in humans, for all drugs, to be published, including older trials.

2. How this article made you feel will be different for each person, but here are some selected quotes which could stimulate discussion.

   - "the success of many drugs has been driven by sophisticated marketing rather than verifiable evidence"
   - "Journal publications of clinical trials may generate media attention, propel researchers careers, and generate some journals a revenue stream."
   - "at present, industry seems extremely reluctant to make its clinical study reports freely available."
   - "It is the public who take and pay for approved drugs, and therefore the public should have access to complete information about those drugs."
   - "Non-disclosure of complete trial results undermines the philanthropy of human participants and sets back the pursuit of knowledge."
   - "Potentially valid reasons for restricting the full public release of clinical study reports include: ensuring patient confidentiality (although this could be remedied with redaction); commercial secrets (although these should be clear after drugs have been registered, and we found no commercially sensitive information in the clinical study reports received from EMA with minimal redactions); and finally, industry concerns over adversaries malicious cherry picking over large datasets (which could be reduced by requiring the prospective registration of research protocols)."

(2) Functions are somewhat like separate programs which can be run as part of a larger program. They will take some input values (known as arguments), perform some predetermined calculations and produce some output. Often in a program, the same calculations will be repeated, so to make the program easier to understand and to debug, these repeated calculations will be written into a separate function.

(3) The outputs are:

Program 1:

```
55
10
```

Program 2:

```
GPA is 5.25
Grade is: 2
```

(4) This is a discussion question.

(5) This is a discussion question.

(6) pH$= 2.5$ means there are $10^{-2.5}$ mols/L of H^+ ions in Coca Cola. Now, assume the population of Australia is about 2×10^7 people, that they all drink Coca Cola (we will allow for this later in reduced personal consumption), and that the average consumption is about 0.5 L/person/week, or about 25 L/person/year. So we have:

$10^{-2.5}$ mols/L $\times 2 \times 10^7$ people $\times$ 25 L/person/year $\approx 1.5811 \times 10^6$ mol $\approx 1.6 \times 10^6$ mol of H^+ ions.

(7) The trick is to first fill in the totals for rows and columns. Of 200 women, 155 are pregnant and 45 are not. These are the totals for the two columns. Also, 143 tested positive and 57 tested negative; these are the row totals. The value of A is 139, so now it is easy to find the remaining values: $B = 4$, $C = 16$ and $D = 41$.

Now it is easy to calculate the values. Sensitivity equals $139/155$, specificity equals $41/45$ and accuracy equals $180/200$.

(8)
1. Students scored high marks on this question but tended to write too much, and hence wasted time.
   A. Reasons include: medical tests are not perfect; individual patients vary; operator errors; interpretation errors; different conditions may be present more or less strongly.
   B. False negative: treatment is delayed; early and/or unexpected death; untreated suffering; increased risk of permanent harm; ultimate treatment may be more expensive.
   False positive leads to unnecessary: worry and psychological suffering; treatment; side effects; costs.
2. A. Consider 100 men. We have $A + C = 10$, so a sensitivity of 34.9% means $A = 3.49$ and hence $C = 6.51$. Similarly, $B + D = 90$, so a specificity of 63.1% means $D/90 = 0.631$, so $D = 56.79$ and hence $B = 33.21$.
   B. Of the 100 men, $A + B = 3.49 + 33.21 = 36.7$ test positive, of whom $3.49/36.7 = 9.5\%$ have cancer.
   C. Of the 100 men, 63.3 test negative, of whom $6.51/63.3 = 10.3\%$ have cancer.
   D. Whichever way you test, you have about a 10% chance of having cancer. In fact, you are more likely to have prostate cancer if you test **negative** than if you test positive. Essentially, there seems to be little or no point in taking the test.
3. A. The population is probably roughly half male and half female, so about half of the time the test accurately determines that a randomly selected Somalian is not pregnant (because the person is male and so cannot be pregnant). However, if the Somalian is female, which happens about half of the time, then the test is wrong 90% of the time.
   B. To calculate the sensitivity, we need to know the values of A and C. The people in C are the people where the test was negative (so they are male), but the condition (pregnancy) was present. It is clear that there can be no pregnant men, so $C = 0$. Regardless of what A is, we can calculate the sensitivity to be
   $$\frac{A}{A+C} = \frac{A}{A} = 1.$$
   This is why the sensitivity must be 100%.
   C. As noted in the answer to Part 2C, regardless of whether the test is positive or negative, an individual man has about a 10% chance of having cancer, which is what ws known about the population before he took the test. For the pregnancy test in Part 3, however, when the test is negative, it is known that the result is 100% correct. Thus, a person who tests negative certainly cannot have the condition (in this case, a man cannot be pregnant). Despite the absurd simplicity of the gender-based pregnancy test, it is arguably more useful than the genuine test for prostate cancer with low specificity and sensitivity.
4. A test with very high sensitivity must have the values of A and $A + C$ of similar size. This means that C must be very small. The value C is the number of times that the test is negative and the condition is present, that is, the number of false negatives. If this number is low, then if a test comes back negative, it is very likely that the condition is actually not present, rather than a false negative reading. This means, for example, that a doctor can be quite confident in ruling out a condition if the sensitivity of the test for that condition is high.

(9)
1. The output is

```
Enter numbers of people with:
Condition Yes, Test +ve:86
Condition Yes, Test -ve:10
Sensitivity = 0.895833333333
```

2. This question involved running program to verify the previous answers.
3. One possible solution is

```
def getSpec(B,D):
    # Calculate specificity
    s = D/(B+D)
    return s

def getAcc(A,D,N):
    # Calculate accuracy
    a = (A+D)/N
    return a
```

4. This program is given in the answer to the next question, with some minor additions specific to that question.

1. Here is one possible solution

```
# A program to investigate effectiveness of a medical test
from pylab import *

def getSens(A,C):
    # Calculate sensitivity
    s = A/(A+C)
    return s

def getSpec(B,D):
    # Calculate specificity
    s = D/(B+D)
    return s

def getAcc(A,D,N):
    # Calculate accuracy
    a = (A+D)/N
    return a

#Asks user for input
print("This program calculates the sensitivity, specificity")
print("and accuracy of a medical test.")
print("Enter numbers of people with:")
A = eval(input("Condition Yes, Test +ve:"))
B = eval(input("Condition No, Test +ve:"))
C = eval(input("Condition Yes, Test -ve:"))
D = eval(input("Condition No, Test -ve:"))

#Calculates total population
N = A + B + C + D

#Uses functions to calculate sensitivity, specificity and accuracy
sens = getSens(A, C)
spec = getSpec(B, D)
acc = getAcc(A, D, N)

#Prints results
print("Sensitivity =",sens)
print("Specificity =",spec)
print("Accuracy =",acc)
```

```
if sens>spec:
    print("The sensitivity of this test is larger than the specificity.")
elif spec>sens:
    print("The specificity of this test is larger than the sensitivity.")
else:
    print("The sensitivity and specificity of this test are equal.")
```

Here is the output from running the program.

```
This program calculates the sensitivity, specificity
and accuracy of a medical test.
Enter numbers of people with:
Condition Yes, Test +ve:3.49
Condition No, Test +ve:33.21
Condition Yes, Test -ve:6.51
Condition No, Test -ve:56.79
Sensitivity = 0.349
Specificity = 0.631
Accuracy = 0.6028
The specificity of this test is larger than the sensitivity.
```

This agrees with the values stated in Question 7.

(11) Some comments:

- terminology such as “1 in 10” and “18%” may be correct, but why mix it up? It may just confuse the reader.
- 13000 cases out of 304000 cases is inconsistent with the other numbers. This is about 4%, yet 10% of male and 4% of female cancers are caused by alcohol, it says.
- 40% to 98% is a meaningless range; it is so broad that it is useless.
- Another comment on that range: what is 100% of people drank more than the recommended maximum? Then the fact that 40% to 98% of people with cancer drank too much might mean that the alcohol was a protective factor.

Other comments are also possible.

(12)

1. The life expectancy of a Nigerien woman is 53.85 years, and she will have on average 7.52 children. We are assuming each of those children were born after 9 months of pregnancy, so a typical Nigerien woman spends $9 \times 7.52 = 67.68$ months, or 5.64 years pregnant. This is $\dfrac{5.64}{53.85} = 0.1047$, or 10.47%, of her life pregnant.
2. The proportion of the population of Niger that are women is 50%, so, on average, only 500 of the sample of Nigeriens should be female. Only 10.47% of these women will be pregnant, which gives $500 \times 0.1047 = 52.37$ women. This is about 52, the figure given in the question.
3. Every woman in the sample will test positive, but only 52 women will actually be pregnant, so $A = 52$ and $B = 500 - 52 = 448$. All 500 of the men will test negative, and all of them are definitely not pregnant, so $C = 0$ and $D = 500$. From here the table can be filled in and the sensitivity, specificity and accuracy can easily be calculated. The sensitivity equals $\dfrac{52}{52+0} = 1$, specificity equals $\dfrac{500}{500+448} = 0.530$ and accuracy equals $\dfrac{500+52}{1000} = 0.552$.
4. Intuitively, this test should be better, because it excludes a few more people who we already know won't be pregnant. From the table, $48 + 9 = 57\%$ of the female population is younger than 15 or older than 49, so only 43% of the female population will test positive, which is 215 women. Now, only 52 of these women are actually pregnant, so we still have that $A = 52$ but now $B = 215 - 52 = 163$. All of the remaining 785 Nigeriens test negative and we know are not pregnant, so we still have that $C = 0$ but now $D = 785$. So we have that the sensitivity equals $\dfrac{52}{52+0} = 1$, specificity equals $\dfrac{785}{785+163} = 0.832$ and accuracy equals $\dfrac{52+785}{1000} = 0.837$.

(13) Here are a few points and comments:

- The first sentence states "The problem with mathematics is that it treats everything as being exact." Not all mathematics is concerned about exact phenomena - the entire field of probability theory is concerned with random events, and while the probability of an event may be exact, there is usually no way to predict when and if it will occur. In addition, the whole idea of a mathematical model is to *approximate* something, to a useful level of accuracy. Even though the model will not be exact, it still allows useful conclusions to be drawn.
- The second sentence states "But the 'real world' doesn't follow exact equations, and nothing is certain or precise." A lot of the real world DOES follow exact equations, and to a very high degree of accuracy. For example, GPS satellites work with such high precision according to times that are corrected due to the effects of special and general relativity - these values match up exactly. Our internet and banking transactions are secure due to cryptographic methods that just as secure in the 'real world' as it is on paper.

  Even if the real world is not exact, mathematical models still provide a useful mechanism for understanding and interpreting phenomena.
- The last statement, "Thus, mathematics may be nice in theory, but not very useful in practice" is completely false. For example, the 2009 Australian Prime Minister's prize for science was awarded to Dr. John O'Sullivan for creating the underlying science to make WiFi communication possible - the science he 'created' was to apply mathematics to the real world problem of fast and reliable data transmission.

(14) Many aspects of this report are quite good, such as: identifying the journal containing the research; giving both absolute and percentage figures; being easy to understand; noting that watching television is not the cause; and giving information about the disease itself and how it might be mitigated. However, there are a number of serious problems which become apparent on closer examination. These include:

- At the end of the second paragraph it says "the odds were still small, about 2 in 100". What does this mean, given that 6% of the children developed asthma? The quoted statement seems to make no sense at all.
- It says that children who watch more than 2 hours of television per day are almost twice as likely to develop asthma as those who watch less television, but it does not say how much less television. Presumably children who watch 1 hour and 59 minutes of television per day do not have a substantially lower risk.
- The article does say that that television is not the cause of the increased risk, but the title certainly implies that it is.
- Most importantly, consider the statement "Of the children with asthma, 2% did not watch TV, 20% watched TV daily for less than an hour, 34% watched 1-2 hours a day and 44% watched more than two hours daily." This presumably is taken to prove that watching more television is associated with asthma (although not necessarily causing it). However, we cannot deduce anything without knowing statistics about the whole population. For example, what if 80% of **all** children watched more than 2 hours of television per day? Then this would mean that watching lots of television is associated with a **lower** incidence of asthma.

(15) While the article is good in that it interviews the investigating scientist, there a few points where both the author and the expert are inconsistent:

- The article claims that swine flu is "rampant" throughout the Northern territory, however does not mention in any overall increase in the number of influenza cases - only its proportion relative to regular influenza.
- The time scale varies in the story - at the top it is "between June and July" but at the end of the article the authors only mention a time period of 6 weeks.
- They states that "Seasonal influenza began to climb in late May this year when surveillance increased in a bid to reduce swine flu cases." Therefore, swine flu **might** have actually been more prevalent in the general community before the surveillance started, and as more and more people became aware (and fearful) of the disease, they might be reporting more.
- On the same token, they say "Up to 8.2 per cent of all reported cases were swine flu...", which implies that this number might not be accurate as the number of swine flu cases might have been grossly under-reported. The words "Up to 8.2 per cent..." also implies that they are not exactly certain about the number they are quoting, and in reality, might be very different.

(16) One comment is: 56% of students who had had an accident admitted they had driven while sleepy. But consider the 259 students who had not had an accident; maybe 100% of them had driven while sleepy. If so, driving while sleepy would not be a risk factor, but would instead be a protective factor.

(17) Answers will vary. Noteworthy points include:

- The study only applies to men, yet there is a claim of a 60% reduction without any gender restriction.
- the phrase "steadily decreased" is used: does this imply linear? If so, does drinking a large amount give a negative risk?
- It claims a "2% reduction" for each additional glass drunk per month. Does this mean drinking 50 glasses a month (which is not particularly excessive) gives a 0% risk?
- The final statement about a "4% reduction" appears to be at odds with the rest of the article; it suggests that only 1/2 of a glass per day would give a 60% reduction. Then, presumably, 1 glass per day would give 120% reduction, which means a negative probability?