University of Technology Sydney
School of Mathematical and Physical Sciences

Probability and Random Variables (37161) –
Class 2 Preparation Work

1. Consider a random experiment which consists of repeatedly drawing cards at random from a standard deck of playing cards.

   If selected cards are not placed back in the deck, calculate the probabilities of the following events:

   i) the first three cards drawn are all red, given that the first card drawn is red
   ii) the first three cards drawn are all red, given that the first card drawn is a diamond
   iii) the fifth and six card drawn are both aces, given that only one of the first four cards drawn was an ace
   iv) the first picture card (jack, queen or king) is a jack, given that the first four cards selected are all numbered 2, 3 or 4.

2. A factory has three employees who all make a particular product. 10% of the objects manufactured by Employee A are rejected as faulty, 15% of the objects manufactured by Employee B are rejected as faulty and 25% of the objects made by Employee C are rejected as fault.

   Employee A makes twice as many parts as both Employee B and Employee C (who each make an equal number.)

   Let A, B and C be the events that a randomly selected part was made by employee A, B or C respectively. That gives $P(A) = 0.5$, $P(B) = 0.25$ and $P(C) = 0.25$.

   Let F be the probability that a randomly chosen item is faulty.

   i) Write down the conditional probabilities $P(F|A), P(F|B)$ and $P(F|C)$.

   ii) A part is selected at random and is observed to be non-faulty. Using Bayes' Theorem, show that the probability that part was made by Employee C is $\approx 22\%$.

3. People eating the chemical compound phenylthiocarbamide (PTC) either report that it has a very bitter taste or no taste at all.

The inability to taste PTC is controlled by a single recessive gene. Let the "taster" allele be *T* and the "non-taster" allele be *t*. Each individual has two alleles - one inherited from his/her mother and one from his/her father. Assume that both of each parent's alleles are equally likely to pass to each child.

Each individual is classified by the two letters (in order) describing the allele inherited from his/her mother and the one from his/her father. For example, someone classified as *tT* received a "non-taster" allele from his/her mother and a "taster" allele from his/her father.

Only individuals who are *tt* (two "non-taster" alleles) record PTC as tasteless, all others record it as bitter.

One study in the United States estimated that 70% of the population were "tasters" and the remaining 30% "non-tasters".

i) Show that the expected proportion of *t* alleles ("non-taster" alleles) in the population is approximately 54.8%.

ii) What is the probability that the child of two *tt* "non-taster" parents is a "non-taster"?

iii) A mother with alleles *tt* and a father with alleles *Tt* have three children. What is the probability that all three of their children are "non-tasters"?

4. Simpson's paradox in sports

In baseball, batting averages are calculated as the number of times a player gets a base hit (hits the ball and safely gets to first base) as a proportion of the total times he goes to bat. As such, a batting average calculated over a time period can be taken to be the probability that, during one randomly selected occasion at bat during that period, a player manages a base hit.

The following is a real example concerning the Major League Baseball performances of Derek Jeter of the New York Yankees and David Justice of the Atlanta Braves and Cleveland Indians during the 1995, 1996 and 1997 seasons.

In 1995, 1996 and 1997, Jeter's batting averages were 0.250, 0.314 and 0.291 respectively.
In 1995, 1996 and 1997, Justice's batting averages were 0.253, 0.321 and 0.329 respectively.

Jeter went to bat 1284 times during that three year period (48 times in 1995, 582 times in 1996 and 654 times in 1997).
Justice went to bat 1046 times during that three year period (411 times in 1995, 140 times in 1996 and 495 times in 1997).

a)

i) Show that, correct to three decimal places, the batting average of Derek Jeter over the three seasons is 0.300.

That is, show that for a randomly selected occasion at bat from any season, the probability that he managed a base hit on that occasion was approximately 0.300.

ii) Calculate the batting average of David Justice over the three seasons. That is, for a randomly selected occasion at bat from any season, what is the probability that he managed a base hit.

b)

i) Which player had the better batting average during each individual season?

ii) Which player had the better batting average calculated over all three seasons?

c) Discuss the above result and reasons for how it might have arisen.