HW Solution 2 — Due: Sep 7, 4 PM

Lecturer: Prapun Suksompong, Ph.D.

Problem 1. [Montgomery and Runger, 2010, Q2-54] Each of the possible five outcomes of a random experiment is equally likely. The sample space is $\{a, b, c, d, e\}$. Let A denote the event $\{a, b\}$, and let B denote the event $\{c, d, e\}$. Determine the following:

(a) $P(A)$

(b) $P(B)$

(c) $P(A^c)$

(d) $P(A \cup B)$

(e) $P(A \cap B)$

Solution: Because the outcomes are equally likely, we can simply use classical probability.

(a) $P(A) = \frac{|A|}{|\Omega|} = \boxed{\frac{2}{5}}$

(b) $P(B) = \frac{|B|}{|\Omega|} = \boxed{\frac{3}{5}}$

(c) $P(A^c) = \frac{|A^c|}{|\Omega|} = \frac{5-2}{5} = \boxed{\frac{3}{5}}$

(d) $P(A \cup B) = \frac{|\{a,b,c,d,e\}|}{|\Omega|} = \frac{5}{5} = \boxed{1}$

(e) $P(A \cap B) = \frac{|\emptyset|}{|\Omega|} = \boxed{0}$

Problem 2. (**Classical Probability and Combinatorics**) Shuffle a deck of cards and cut it into three piles. What is the probability that (at least) a court card will turn up on top of one of the piles.

Hint: There are 12 court cards (four jacks, four queens and four kings) in the deck.

Solution: In [Lovell, 2006, p. 17–19], this problem is named "Three Lucky Piles".

Method 1: When somebody cuts three piles, they are, in effect, randomly picking three cards from the deck. There are $52 \times 51 \times 50$ possible outcomes. The number of outcomes that do not contain any court card is $40 \times 39 \times 38$. So, the probability of having at least one court card is

$$\boxed{\frac{52 \times 51 \times 50 - 40 \times 39 \times 38}{52 \times 51 \times 50} \approx 0.553.}$$

Method 2: Note that our solution above, especially the part where we use the words "in effect", may not be so evident to some of you. If you want to solve this question directly, you can approach it using the total probability theorem which is studied in Chapter 6. In the beginning, we shuffle the cards. So, after the shuffling, we will have a deck of 52 cards with all the possible 52! permutations being equally likely. (In our mind,) we label the cards with #1 to #52 from the top to bottom. Now, the next step is to cut it into three piles. Note that this is the same as choosing two cards (from #2 (top) to #52 (bottom)) to indicate where the two boundaries (which are the same as the two cards at the top of second and third piles) are. Note also that this process is usually biased. Most will try to divide the deck into three piles of approximately equal size. So, it is *unlikely* that you will have the first piles with 50 cards, the second with only one card, and the third with only one card. So, classical probability can not be used here. We only know that there are $\binom{51}{2} = 1,275$ ways to perform the cutting for a particular deck and they are not equally likely. Let event $B_1, \ldots, B_{1275}$ denote each of these cases. For example, B_{134} may be the case in which the cutting positions are at cards #32 and #45. So, the top cards on the three piles are cards #1, #32, and #45. Let A be the event that at least one of these cards is a court card. Of course, the "at least one" counting problem can be simplified by considering the opposite case. A^c is the event that none of the three top cards is a court card. So, there are $52 - 12 = 40$ choices for card #1. There are $40 - 1 = 39$ choices for card #32. There are $39 - 1 = 38$ choices for card #45. For the remaining $52 - 3 = 49$ cards, there is no restriction. So, there are 49! choices. In total, we have $40 \times 39 \times 38 \times (49!)$ shuffled patterns among the 52! equally likely possibilities that satisfy A^c. Therefore,

$$P(A|B_{134}) = \frac{52! - 40 \times 39 \times 38 \times (49!)}{52!} = 1 - \frac{40 \times 39 \times 38}{52 \times 51 \times 50} \approx 0.553.$$

The same reasoning applies to any cutting positions. So, $P(A|B_k) \approx 0.553$ for any k. By the total probability theorem,

$$P(A) = \sum_{k=1}^{1275} P(A|B_k)\,P(B_k) \approx \sum_{k=1}^{1275} 0.553 P(B_k) = 0.553 \sum_{k=1}^{1275} P(B_k) = 0.553 \times 1 = 0.553.$$

Observe that we still don't know the value of each $P(B_k)$ but we know that the sum of them is 1.

Problem 3. (**Classical Probability**) There are three buttons which are painted red on one side and white on the other. If we tosses the buttons into the air, calculate the probability that all three come up the same color.

Remarks: A *wrong* way of thinking about this problem is to say that there are four ways they can fall. All red showing, all white showing, two reds and a white or two whites and a red. Hence, it seems that out of four possibilities, there are two favorable cases and hence the probability is 1/2.

Solution: There are 8 possible outcomes. (The same number of outcomes as tossing three coins.) Among these, only two outcomes will have all three buttons come up the same color. So, the probability is $2/8 = \boxed{1/4}$.

Problem 4. (**Classical Probability and Combinatorics**) A Web ad can be designed from four different colors, three font types, five font sizes, three images, and five text phrases.

(a) How many different designs are possible? [Montgomery and Runger, 2010, Q2-51]

(b) A specific design is randomly generated by the Web server when you visit the site. If you visit the site five times, what is the probability that you will not see the same design? [Montgomery and Runger, 2010, Q2-71]

Solution:

(a) By the multiplication rule, total number of possible designs

$$= 4 \times 3 \times 5 \times 3 \times 5 = \boxed{900}.$$

(b) From part (a), total number of possible designs is 900. The sample space is now the set of all possible designs that may be seen on five visits. It contains $(900)^5$ outcomes. (This is ordered sampling with replacement.)

The number of outcomes in which all five visits are different can be obtained by realizing that this is ordered sampling without replacement and hence there are $(900)_5$ outcomes. (Alternatively, On the first visit any one of 900 designs may be seen. On the second visit there are 899 remaining designs. On the third visit there are 898 remaining designs. On the fourth and fifth visits there are 897 and 896 remaining designs, respectively. From the multiplication rule, the number of outcomes where all designs are different is $900 \times 899 \times 898 \times 897 \times 896$.)

Therefore, the probability that a design is not seen again is

$$\frac{(900)_5}{900^5} \approx \boxed{0.9889.}$$

Problem 5. (**Classical Probability and Combinatorics**) A bin of 50 parts contains five that are defective. A sample of two parts is selected at random, without replacement. Determine the probability that both parts in the sample are defective. [Montgomery and Runger, 2010, Q2-49]

Solution: The number of ways to select two parts from 50 is $\binom{50}{2}$ and the number of ways to select two defective parts from the 5 defective ones is $\binom{5}{2}$ Therefore the probability is

$$\frac{\binom{5}{2}}{\binom{50}{2}} = \frac{2}{245} = \boxed{0.0082}.$$

Alternatively, if the two parts in the sample are selected one by one, then we may also consider their ordering as well. In such case, we use the formula for "ordered sampling without replacement" instead of "unordered sampling without replacement":

$$\frac{(5)_2}{(50)_2} = \frac{5 \times 4}{50 \times 49} = \frac{2}{245} = \boxed{0.0082}.$$

Problem 6. (**Combinatorics**) Consider the design of a communication system in the United States.

(a) How many three-digit phone prefixes that are used to represent a particular geographic area (such as an area code) can be created from the digits 0 through 9?

(b) How many three-digit phone prefixes are possible in which no digit appears more than once in each prefix?

(c) As in part (a), how many three-digit phone prefixes are possible that do not start with 0 or 1, but contain 0 or 1 as the middle digit?

[Montgomery and Runger, 2010, Q2-45]

Solution:

(a) From the multiplication rule (or by realizing that this is ordered sampling with replacement), $10^3 = \boxed{1,000}$ prefixes are possible

(b) This is ordered sampling without replacement. Therefore $(10)_3 = 10 \times 9 \times 8 = \boxed{720}$ prefixes are possible

(c) From the multiplication rule, $8 \times 2 \times 10 = \boxed{160}$ prefixes are possible.

Problem 7. (**Classical Probability and Combinatorics**) We all know that the chance of a head (H) or tail (T) coming down after a fair coin is tossed are fifty-fifty. If a fair coin is tossed ten times, then intuition says that five heads are likely to turn up.

Calculate the probability of getting exactly five heads (and hence exactly five tails).

Solution: There are 2^{10} possible outcomes for ten coin tosses. (For each toss, there is two possibilities, H or T). Only $\binom{10}{5}$ among these outcomes have exactly heads and five tails. (Choose 5 positions from 10 position for H. Then, the rest of the positions are automatically T.) The probability of have exactly 5 H and 5 T is

$$\boxed{\frac{\binom{10}{5}}{2^{10}} \approx 0.246.}$$

Note that five heads and five tails will turn up more frequently than any other single combination (one head, nine tails for example) but the sum of all the other possibilities is much greater than the single 5 H, 5 T combination.

Extra Questions

Here are some optional questions for those who want more practice.

Problem 8. ***An Even Split at Coin Tossing***: Let p_n be the probability of getting exactly n heads (and hence exactly n tails) when a fair coin is tossed $2n$ times.

(a) Find p_n.

(b) Sometimes, to work theoretically with large factorials, we use Stirling's Formula:

$$n! \approx \sqrt{2\pi n} n^n e^{-n} = \left(\sqrt{2\pi e}\right) e^{\left(n+\frac{1}{2}\right)\ln\left(\frac{n}{e}\right)}. \tag{2.1}$$

Approximate p_n using Stirling's Formula.

(c) Find $\lim_{n\to\infty} p_n$.

Solution: Note that we have worked on a particular case ($n = 5$) of this problem earlier.

(a) Use the same solution as Problem 7; change 5 to n and 10 to $2n$, we have

$$p_n = \boxed{\frac{\binom{2n}{n}}{2^{2n}}}.$$

(b) By Stirling's Formula, we have

$$\binom{2n}{n} = \frac{(2n)!}{n!n!} \approx \frac{\sqrt{2\pi 2n}(2n)^{2n} e^{-2n}}{\left(\sqrt{2\pi n} n^n e^{-n}\right)^2} = \frac{4^n}{\sqrt{\pi n}}.$$

Hence,

$$p_n \approx \boxed{\frac{1}{\sqrt{\pi n}}}. \tag{2.2}$$

[Mosteller, *Fifty Challenging Problems in Probability with Solutions*, 1987, Problem 18]

See Figure 2.1 for comparison of p_n and its approximation via Stirling's formula.

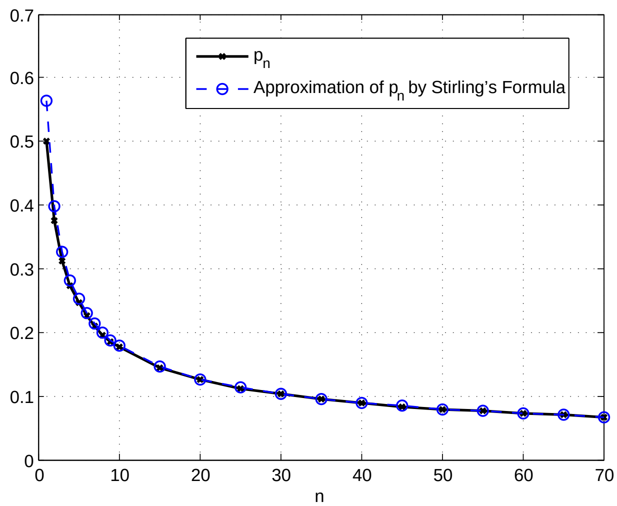

Figure 2.1: Comparison of p_n and its approximation via Stirling's formula

(c) From (2.2), $\lim_{n\to\infty} p_n = \boxed{0}$. A more rigorous proof of this limit would use the bounds

$$\frac{4^n}{\sqrt{4n}} \le \binom{2n}{n} \le \frac{4^n}{\sqrt{3n+1}}.$$

Problem 9. (**Classical Probability and Combinatorics**) Suppose n integers are chosen with replacement (that is, the same integer could be chosen repeatedly) at random from $\{1, 2, 3, \ldots, N\}$. Calculate the probability that the chosen numbers arise according to some non-decreasing sequence.

Solution: There are N^n possible sequences. (This is ordered sampling with replacement.) To find the probability, we need to count the number of non-decreasing sequences among these N^n possible sequences. It takes some thought to realize that this is exactly the counting problem that we called "unordered sampling with replacement". In which case, we can conclude that the probability is $\boxed{\frac{\binom{n+N-1}{n}}{N^n}}$. The "with replacement" part should be clear from the question statement. The "unordered" part needs some more thought.

To see this, let's look back at how we turn the "ordered sampling *without replacement*" into "unordered sampling *without replacement*". Recall that there are $(N)_n$ distinct samples for "ordered sampling without replacement". When we switch to the "unordered" case, we see that many of the original samples from the "ordered sampling without replacement" are regarded as the same in the "unordered" case. In fact, we can form "groups" of samples whose members are regarded as the same in the "unordered" case. We can then count the number of groups. In class, we found that the size of any individual group can be calculated easily from permuting the elements in a sample and hence there are $n!$ members in each group. This leads us to conclude that there are $(N)_n/n! = \binom{N}{n}$ groups.

We are in a similar situation when we want to turn the "ordered sampling *with replacement*" into "unordered sampling *with replacement*". We first start with N^n distinct samples from "ordered sampling with replacement". Now, we again separate these samples into groups. Let's consider an example where $n = 3$. Then sequences "1 1 2", "1 2 1", and "2 1 1" are put together in the same group in the "unordered" case. Note that the size of this group is 3. The sequences "1 2 3", "1 3 2", "2 1 3", "2 3 1", "3 1 2", and "3 2 1" are in another group. Note that the size of this group is 6. Therefore, the group sizes are not the same and hence we can not find the number of groups by N^n/(group size) as in the sampling *without replacement* discussed in the previous paragraph. To count the number of groups, we look at the sequences from another perspective. We see that the "unordered" case, the only information the characterizes each group is "how many of each number there are". This is why we can match the number of groups to the number of nonnegative-integer solutions to the equation $x_1 + x_2 + \cdots + x_N = n$ as discussed in class. Finally, note that for each group, we have only one possible nondecreasing sequence. So, the number of possible nondecreasing sequence is the same as the number of groups.

If you think about the explanation above, you may realize that, by requiring the "order" on the sequence, the counting problem become "unordered sampling".

Here, we present two direct methods that leads to the same answer.

Method 1: Because the sequence is non-decreasing, the number of times that each of the integers $\{1, 2, \ldots, N\}$ shows up in the sequence is the only information that characterizes each sequence. Let x_i be the number of times that number i shows up in the sequence. The number of sequences is then the same as the number of solution to the equation $x_1+x_2+\cdots+x_N = n$ where the x_i are all non-negative integers. We have seen in class that the number of solutions is $\binom{n+N-1}{n}$.

Method 2: [DasGupta, 2010, Example 1.14, p. 12] Consider the following construction of such non-decreasing sequence. Start with n stars and $N - 1$ bars. There are $\binom{n+N-1}{n}$ arrangements of these. For example, when $N = 5$ and $n = 2$, one arrangement is $| * || * |$. Now, add spaces between these bars and stars including before the first one and after the last one. For our earlier example, we have $\_|\_*\_|\_|\_*\_|\_$. Now, put number 1 in the leftmost space. After this position, the next space holds the same value as the previous on if you pass a $*$. On the other hand, if you pass a $|$ then the value increases by 1. Note that because there are $N - 1$ bars, the last space always gets the value N. What you now have is a sequence of $n + N$ numbers with bars between consecutive distinct numbers and stars between consecutive equal numbers. For example, our example would gives $1|2 * 2|3|4 * 4|5$. Note that this gives a non-decreasing sequence of $n + N$ numbers. The corresponding non-decreasing sequence of n numbers for this arrangement of stars and bars is (2,4); that is we only take the numbers to the right of the stars. Because there are n stars, our sequence will have n numbers. It will be non-decreasing because it is a sub-sequence of the non-decreasing

$n + N$ sequence. This shows that any arrangement of n stars and $N - 1$ bars gives one nondecreasing sequence of n numbers.

Conversely, we can take any nondecreasing sequence of n numbers and combine it with the full set of numbers $\{1, 2, 3, \ldots, N\}$ to form a set of $n + N$ numbers. Now rearrange these numbers in a nondecreasing order. Put a bar between consecutive distinct numbers in this set and a star between consecutive equal numbers in this set. Note that the number to the right of each star is an element of the original n-number sequence. This shows that any nondecreasing sequence of n numbers corresponds to an arrangement of n stars and $N - 1$ bars.

Combining the two paragraphs above, we now know that the number of non-decreasing sequences is the same as the number of arrangement of n stars and $N - 1$ bars, which is $\binom{n+N-1}{n}$.

Remark: There is also a method— which will not be discussed here, but can be inferred by finding the pattern of the sums that lead to the number of non-decreasing sequences as we increase the value of n— that would interestingly give the number of non-decreasing sequences as

$$\sum_{k_{n-1}=1}^{N} \cdots \sum_{k_2=1}^{k_3} \sum_{k_1=1}^{k_2} k_1.$$

This sum can be simplified into $\binom{n+N-1}{n}$ by the "parallel summation formula" which is well-known but we didn't discuss in class because this is not a class on combinatorics.