Problem Set 3: Solutions
Due: September 28, 2011

1. Since we have n choices for the first integer and n choices for the second, the total number of ways to pick the two integers is n^2, each of these ways being equally likely. Of these, we can count the pairs in which the integers differ by exactly two: (1,3), (2,4), (3,5),$\cdots$, $(n-2, n)$. There are $n-2$ pairs of this form, and another $n-2$ pairs of the reverse order, that is, (3,1), (4,2), etc. We have a total of $2(n-2)$ pairs that satisfy our requirement. Hence, by the discrete uniform law, the probability we are looking for is $\frac{2(n-2)}{n^2}$. We can verify that our formula works for any $n \geq 2$.

2. To find the probability, we will find the number of favorable outcomes and divide by the total number of possible outcomes. There are $\binom{10}{8}$ favorable outcomes, i.e., successful combinations that will open the lock. There are $\binom{20}{8}$ total number of ways to choose 8 numbers out of 20, and therefore the probability that the burglar will open the vault on his first try is:
$$\frac{\binom{10}{8}}{\binom{20}{8}} = \frac{45}{125970} \approx 0.00036.$$

3. *Count the number of different letter arrangements you can make by changing the order of the letters in the word* **aardvark** *(count the original word, too).*

   Answer: $\binom{8}{3}\binom{5}{2}3! = \frac{8!}{3!2!1!1!1!} = 3360.$

   Solution: choose 3 places out of the total 8 for the 3 "a"-s (the $\binom{8}{3}$ multiplier), then choose 2 places out of the remaining 5 for the 2 "r"-s (the $\binom{5}{2}$ multiplier), then place "d","v", and "k" in the remaining 3 spaces (the 3! multiplier).

   Equivalently, we can take the number of ways to order 8 *different* letters (8!), and then divide it by the numbers of ways to re-order groups of 3 (letter "a"), 2 (letter "r"), 1 (letter "d"), 1 (letter "v"), 1 (letter "k").

4. Without prior knowledge on whether the exit of campus lies East or West, the exact answers of the passerby are not as important as whether a string of answers is similar or not. Let R_r denote the event that we receive r similar answers and T denote the event that these repeated answers are truthful. Let S denote the event that the questioned passerby is a student. Note that, because a professor always gives a false answer, $T \cap S^c = \emptyset$ and thus $\mathbf{P}(T \cap S^c) = 0$. Therefore,
$$\mathbf{P}(T|R_r) = \frac{\mathbf{P}(T \cap R_r)}{\mathbf{P}(R_r)} = \frac{\mathbf{P}(T \cap R_r \cap S)}{\mathbf{P}(R_r)} = \frac{\mathbf{P}(T \cap R_r|S)\mathbf{P}(S)}{\mathbf{P}(R_r)}$$
where the stated independence of a passerby's successive answers implies $\mathbf{P}(T \cap R_r|S) = \left(\frac{3}{4}\right)^r$. Applying the Total Probability Theorem and again making use of independence, we also deduce
$$\mathbf{P}(R_r) = \mathbf{P}(R_r|S)\mathbf{P}(S) + \underbrace{\mathbf{P}(R_r|S^c)}_{1}\mathbf{P}(S^c) = \left(\left(\frac{3}{4}\right)^r + \left(\frac{1}{4}\right)^r\right)\frac{2}{3} + \frac{1}{3} \quad .$$

   (a) Applying the above formulas for $r = 1$, we have $\mathbf{P}(R_1) = 1$ and thus
$$\mathbf{P}(T|R_1) = \frac{\frac{3}{4} \cdot \frac{2}{3}}{1} = \boxed{\frac{1}{2}} \quad .$$

(b) For $r = 2$, the formulas yield $\mathbf{P}(R_2) = \frac{3}{4}$ and thus

$$\mathbf{P}(T|R_2) = \frac{\left(\frac{3}{4}\right)^2 \frac{2}{3}}{\frac{3}{4}} = \boxed{\frac{1}{2}}.$$

(c) For $r = 3$, the formulas yield $\mathbf{P}(R_3) = \frac{15}{24}$ and thus

$$\mathbf{P}(T|R_3) = \frac{\left(\frac{3}{4}\right)^3 \frac{2}{3}}{\frac{15}{24}} = \boxed{\frac{9}{20}} \quad .$$

(d) For $r = 4$, the formulas yield $\mathbf{P}(R_4) = \frac{35}{64}$ and thus

$$\mathbf{P}(T|R_4) = \frac{\left(\frac{3}{4}\right)^4 \frac{2}{3}}{\frac{35}{64}} = \boxed{\frac{27}{70}} \quad .$$

(e) As soon as we receive a dissimilar answer from the same passerby, we know that this passerby is a student; a professor will always give the same (false) answer. Let D denote the event of receiving the first dissimilar answer. Given D on the fourth answer, either the student has provided three truthful answers followed by one untruthful answer, occurring with probability $\left(\frac{3}{4}\right)^3 \frac{1}{4}$, or the student has provided three untruthful answers followed by one truthful answer, occurring with probability $\left(\frac{1}{4}\right)^3 \frac{3}{4}$. Note that event T corresponds to the former; thus,

$$\mathbf{P}(T|R_3 \cap D) = \frac{\left(\frac{3}{4}\right)^3 \frac{1}{4}}{\left(\frac{3}{4}\right)^3 \frac{1}{4} + \left(\frac{1}{4}\right)^3 \frac{3}{4}} = \boxed{\frac{9}{10}} \quad .$$

(f) In parts (a) - (d), notice the decreasing trend in the probability of the passerby being truthful as the number of similar answers grows. Intuitively, our confidence that the passerby is a professor grows as the sequence of similar answers gets longer, because we know a professor will always give the same (false) answer while a student has a chance to answer either way. However, as part (e) demonstrates, the first indication that the passerby is a student will boost our confidence that the previous string of similar answers are truthful, because any single answer by the student has a 3-to-1 chance of being a truthful one.

For the remainder of this problem, let E and W represent the events that a passerby provides East and West, respectively, as an answer and let T_E represent the event that East is the correct answer. We are told Ima's *a priori* knowledge is $\mathbf{P}(T_E) = \epsilon$.

(g) Using Bayes's Rule and all the arguments used in parts (a) - (e), we have

$$\begin{aligned}
\mathbf{P}(T_E|E) &= \frac{\mathbf{P}(E|T_E)\mathbf{P}(T_E)}{\mathbf{P}(E)} = \frac{\left(\frac{2}{3}\right)\left(\frac{3}{4}\right)\epsilon}{\left(\frac{2}{3}\right)\left(\frac{3}{4}\right)\epsilon + \left(\left(\frac{2}{3}\right)\left(\frac{1}{4}\right) + \frac{1}{3}\right)(1-\epsilon)} = \boxed{\epsilon} \quad \text{and} \\
\mathbf{P}(T_E|W) &= \frac{\mathbf{P}(W|T_E)\mathbf{P}(T_E)}{\mathbf{P}(W)} = \frac{\left(\left(\frac{2}{3}\right)\left(\frac{1}{4}\right) + \frac{1}{3}\right)\epsilon}{\left(\left(\frac{2}{3}\right)\left(\frac{1}{4}\right) + \frac{1}{3}\right)\epsilon + \left(\frac{2}{3}\right)\left(\frac{3}{4}\right)(1-\epsilon)} = \boxed{\epsilon} \quad .
\end{aligned}$$

In particular, we have used that $\mathbf{P}(E) = \mathbf{P}(E|T_E)\mathbf{P}(T_E) + \mathbf{P}(E|T_E^c)\mathbf{P}(T_E^c)$ (and similarly for $\mathbf{P}(W)$)

(h) Likewise, given two consecutive and similar answers from the same passerby, we have

$$\mathbf{P}(T_E|EE) = \frac{\left(\frac{2}{3}\right)\left(\frac{3}{4}\right)^2\epsilon}{\left(\frac{2}{3}\right)\left(\frac{3}{4}\right)^2\epsilon + \left(\left(\frac{2}{3}\right)\left(\frac{1}{4}\right)^2 + \frac{1}{3}\right)(1-\epsilon)} = \boxed{\epsilon} \quad \text{and}$$

$$\mathbf{P}(T_E|WW) = \frac{\left(\left(\frac{2}{3}\right)\left(\frac{1}{4}\right)^2 + \frac{1}{3}\right)\epsilon}{\left(\left(\frac{2}{3}\right)\left(\frac{1}{4}\right)^2 + \frac{1}{3}\right)\epsilon + \left(\frac{2}{3}\right)\left(\frac{3}{4}\right)^2(1-\epsilon)} = \boxed{\epsilon} \quad .$$

(i) Finally, given three consecutive and similar answers from the same passerby,

$$\mathbf{P}(T_E|EEE) = \frac{\left(\frac{2}{3}\right)\left(\frac{3}{4}\right)^3\epsilon}{\left(\frac{2}{3}\right)\left(\frac{3}{4}\right)^3\epsilon + \left(\left(\frac{2}{3}\right)\left(\frac{1}{4}\right)^3 + \frac{1}{3}\right)(1-\epsilon)} = \boxed{\frac{9\epsilon}{11-2\epsilon}} \quad \text{and}$$

$$\mathbf{P}(T_E|WWW) = \frac{\left(\left(\frac{2}{3}\right)\left(\frac{1}{4}\right)^3 + \frac{1}{3}\right)\epsilon}{\left(\left(\frac{2}{3}\right)\left(\frac{1}{4}\right)^3 + \frac{1}{3}\right)\epsilon + \left(\frac{2}{3}\right)\left(\frac{3}{4}\right)^3(1-\epsilon)} = \boxed{\frac{11\epsilon}{9+2\epsilon}} \quad .$$

For $\epsilon = \frac{9}{20}$, we calculate $\mathbf{P}(T_E|EEE) = \frac{81}{202}$ and $\mathbf{P}(T_E|WWW) = \frac{1}{2}$.

Notice that the E, EE and EEE answers to parts (f) - (h) match the answers to parts (a)-(c) when $\epsilon = \frac{1}{2}$, or when Ima's prior knowledge does not favor either possibility.

5. We assume that all choices for length-r lottery numbers out of n integers are equally-likely and use the discrete uniform probability law and counting arguments. In other words, if A denotes the event of interest,

$$\mathbf{P}(A) = \frac{\text{number of elements in } A}{\text{number of elements in } \Omega} \equiv \frac{\text{number of favorable outcomes}}{\text{number of total outcomes}}$$

in each part of this problem. We also assume that winning this particular lottery does not depend on the order that the integers are chosen and that $n > 2r$.

(a) Event A corresponds to the subset L being drawn in increasing order. Note that any particular length-r subset L may result from drawing the same r integers in $r!$ different orderings, and only one of these orderings can be in increasing order. Combining this observation with the total probability theorem, we obtain

$$\mathbf{P}(A) = \sum_{\text{subsets}} \mathbf{P}(\text{subset})\underbrace{\mathbf{P}(A|\text{subset})}_{\frac{1}{r!}} = \frac{1}{r!}\underbrace{\sum_{\text{subsets}} \mathbf{P}(\text{subset})}_{1} = \boxed{\frac{1}{r!}}$$

(b) Event A corresponds to selecting the r integers such that a match occurs with the subset L selected by the lottery. In other words, given our choice of r integers, we wish to determine the probability that L is the same. Clearly, among all the $\binom{n}{r}$ choices only one is favorable and hence the probability is

$$\mathbf{P}(A) = \boxed{\frac{1}{\binom{n}{r}}}$$

(c) Event A corresponds to having exactly k numbers out of the r we choose to match members of L. The number of favorable outcomes is then the *product* of (i) the number of unique length-k subsets of L, or $\binom{r}{k}$, and (ii) the number of ways we could choose the remaining $r-k$ integers from the set that the lottery did *not* select, or $\binom{n-r}{r-k}$. Therefore the desired probability is equal to

$$\mathbf{P}(A) = \boxed{\frac{\binom{r}{k}\binom{n-r}{r-k}}{\binom{n}{r}}}$$

(d) Event A corresponds to choosing r numbers out of n such that no two numbers are consecutive. Consider the equivalent problem of trying to order r black balls and $n-r$ white balls in a sequence such that no two black balls are consecutive. (Having a black ball in position i is equivalent to choosing integer i in the original problem).
We can count the number of sequences as follows:
Envision $n-r$ white balls in a sequence. There are $n-r-1$ gaps between the white balls, and 2 gaps on both ends; a total of $n-r+1$ gaps. Assume that we can only place one black ball in each gap to make sure no two are consecutive. The number of ways we can place r black balls into the $n-r+1$ gaps is $\binom{n-r+1}{r}$.
Therefore, $\mathbf{P}(A) = \frac{\binom{n-r+1}{r}}{\binom{n}{r}}$.

(e) Event B corresponds to choosing r numbers out of n such that exactly two numbers are consecutive. Again, consider the equivalent problem of trying to order r black balls and $n-r$ white balls in a sequence such that exactly two black balls are consecutive. First we place $r-1$ black balls in the $n-r+1$ gaps between white balls. This can be achieved in $\binom{n-r+1}{r-1}$ ways.
The r^{th} black ball should be paired with one of the $r-1$ already placed black balls. This can be achieved in $\binom{r-1}{1}$ ways.
By the counting principle, the total number of such sequences with exactly two consecutive black balls is $(r-1)\binom{n-r+1}{r-1}$ ways.
Therefore, $\mathbf{P}(B) = (r-1)\,\frac{\binom{n-r+1}{r-1}}{\binom{n}{r}}$.

6. (a) Because each X_k has a continuous distribution, the probability that $X_i = X_j$ is 0, for any $i \neq j$. Therefore, the probability of repeated values in the sequence is also 0. In particular, we can effectively consider only sequences with distinct values and ignore sequences with "ties". To be somewhat more rigorous, note that for $j \neq k$, $\mathbf{P}(X_j = X_k \mid X_k = a) = \mathbf{P}(X_j = a \mid X_k = a) = \mathbf{P}(X_j = a)$ since X_j and X_k are independent by assumption. Because we are dealing with continuous distributions, this final probability is 0.

(b) The answers below are all based on the observation that, given any value of $m \geq 1$ and any set of numbers $\{X_n \mid 1 \leq n \leq m\}$, all permutations of the m numbers are equally likely, since the choices are independent by assumption.

X_n is a record if and only if it is the largest among the first n X_k's. Since the X_k's are all i.i.d., any of them is equally likely to be the largest among them, by symmetry. Hence, the probability that it is X_n is just $1/n$.

Alternatively, we can using a counting method. With n distinct values, there are exactly $n!$ permutations into different orderings. In order for X_n to be a record, its value must be the maximum and therefore there is only 1 choice for its value. However, the remaining $n-1$ values can be arranged in any order, and thus there are $(n-1)!$ favorable outcomes. Since each outcome (i.e., permutation) is equally likely, the desired probability is $(n-1)!/n! = 1/n$.

Thus, for all $n > 0$,

$$\mathbf{P}(X_n \text{ is a record}) = \frac{1}{n}.$$

(c) Yes, these two events are independent. Let E_i be the event that X_i is a record, for each $i \geq 1$. Without loss of generality, let us assume that $k < n$. To show that E_k and E_n are independent, we need to show that

$$\mathbf{P}(E_k \cap E_n) = \mathbf{P}(E_k)\mathbf{P}(E_n).$$

We can equivalently write the left-hand-side as $\mathbf{P}(E_k \cap E_n) = \mathbf{P}(E_k \mid E_n)\mathbf{P}(E_n)$. From part (a), we know that $\mathbf{P}(E_n) = 1/n$. Now consider the conditional probability term. Given that X_n is a record (event E_n) and $k < n$, the first k values of the sequence must be less than X_n. But otherwise, there are still $k!$ total permutations and still $(k-1)!$ favorable permutations where the largest of the k values is assigned to X_k. Therefore, $\mathbf{P}(E_k \mid E_n) = 1/k$ and we have the desired equality

$$\mathbf{P}(E_k \cap E_n) = \mathbf{P}(E_k \mid E_n)\mathbf{P}(E_n) = \frac{1}{k} \cdot \frac{1}{n} = \mathbf{P}(E_k)\mathbf{P}(E_n).$$

(d) Yes, these events are independent. We need to show that

$$\mathbf{P}(E_{j_1} \cap \cdots \cap E_{j_m}) = \mathbf{P}(E_{j_1}) \cdot \cdots \cdot \mathbf{P}(E_{j_m}).$$

We use the multiplication rule to write the left hand side as

$$\mathbf{P}(E_{j_1} \cap \cdots \cap E_{j_m}) = \mathbf{P}(E_{j_1} \mid E_{j_2} \cap \cdots \cap E_{j_m})\mathbf{P}(E_{j_2} \mid E_{j_3} \cap \cdots \cap E_{j_m}) \cdot \cdots \cdot \mathbf{P}(E_{j_{m-1}} \mid E_{j_m})\mathbf{P}(E_{j_m}).$$

Using the same reasoning as in part (c), we know that future records do not affect the probability of past records, i.e., $\mathbf{P}(E_k \mid E_n) = \mathbf{P}(E_k) = 1/k$ if $k < n$. Therefore, all of these conditional probabilities simplify so that we obtain

$$\mathbf{P}(E_{j_1} \cap \cdots \cap E_{j_m}) = \mathbf{P}(E_{j_1})\mathbf{P}(E_{j_2}) \cdot \cdots \cdot \mathbf{P}(E_{j_{m-1}})\mathbf{P}(E_{j_m}) = \frac{1}{j_1}\frac{1}{j_2} \cdot \cdots \cdot \frac{1}{j_m}.$$

Hence, the events are independent.

(e) Following the hint, we calculate the integral

$$\int_1^y \frac{1}{x^a} dx = \begin{cases} \frac{y^{1-a}-1}{1-a}, & \text{if } a \neq 0, \\ y-1, & \text{if } a = 0. \end{cases}$$

Suppose $a > 1$. Then, taking limits, we obtain that

$$\int_1^\infty \frac{1}{x^a}dx = \lim_{y\to\infty}\int_1^y \frac{1}{x^a}dx = \lim_{y\to\infty}\frac{y^{1-a}-1}{1-a} = \frac{0-1}{1-a} = \frac{1}{a-1} < \infty.$$

For the converse, suppose $a \le 1$. Consider the subcase where $a \neq 0$. Then

$$\int_1^\infty \frac{1}{x^a}dx = \lim_{y\to\infty}\int_1^y \frac{1}{x^a}dx = \lim_{y\to\infty}\frac{y^{1-a}-1}{1-a} = \frac{\infty-1}{1-a} = \infty.$$

For the remaining subcase of $a = 0$, we have

$$\int_1^\infty \frac{1}{x^a}dx = \lim_{y\to\infty}\int_1^y \frac{1}{x^a}dx = \lim_{y\to\infty} y - 1 = \infty.$$

Therefore, we have proven the facts given in the hint.
Now, consider the series in the stated fact. Suppose $a > 1$ and let $f(x) = (1/x)^a$. Note that the function $f(x)$ is continuous, monotonically decreasing, and non-negative. We can therefore apply the integral test for convergence of series, which states that the series $\sum_{k=1}^\infty f(k)$ converges if and only if the integral $\int_1^\infty f(x)dx$ is finite. Since we have already shown that this integral is finite, we conclude that the series converges.
For an alternate, visual argument, see Figure 1. We define the functions $f_u(x)$ and $f_\ell(x)$ as

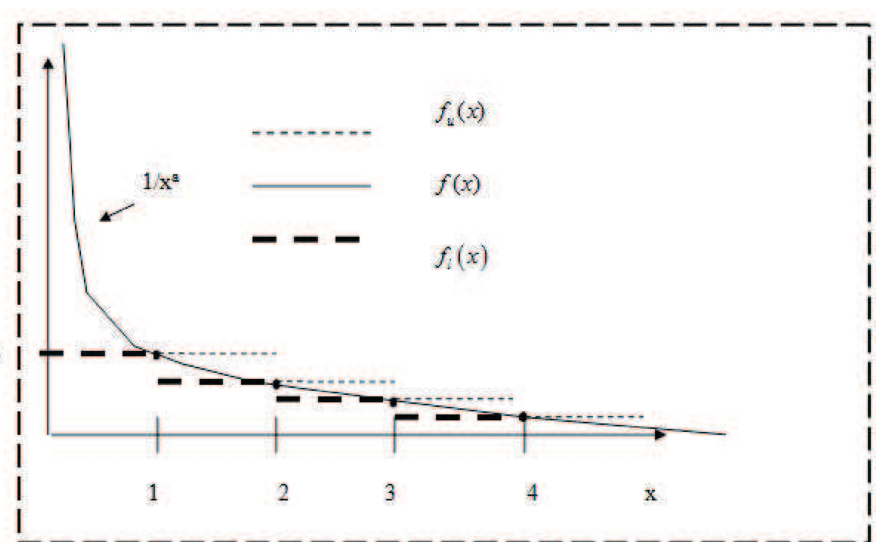

Figure 1: $f(x) = 1/x^a$ is the continuous approximation of the terms in the series. $f_u(x)$ gives the upper (maximum) value $f(x)$ takes on in each interval $[k, k+1]$, where k is a positive integer. Similarly, $f_\ell(x)$ gives the lower (minimum) values of $f(x)$ over the same intervals.

the maximum and minimum values, respectively, of $f(x)$ over each interval $[k, k+1]$, where k are the positive integers. Thus, these two functions are step functions. For $a \ge 0$, it is true that $f_u(x) \ge f(x)$ for all x and $\int_1^{n+1} f_u(x)dx = \sum_{k=1}^n \frac{1}{k^a}$. Similarly, $f_\ell(x) \le f(x)$ for all x and $\int_1^n f_\ell(x)dx + 1 = \sum_{k=1}^n \frac{1}{k^a}$. It therefore follows that

$$\int_1^{n+1} f(x)dx \le \int_1^{n+1} f_u(x)dx = \sum_{k=1}^n \frac{1}{x^a} = 1 + \int_1^n f_\ell(x)dx \le 1 + \int_1^n f(x)dx.$$

We have then shown that the series is bounded by the integral. Therefore, the series $\sum_{k=1}^{n} \frac{1}{k^a}$ converges as $n \to \infty$ if and only if the integral $\int_1^n f(x)dx$ converges as $n \to \infty$, which we have already shown is true if and only if $a > 1$. A similar argument proves the other case where we assum $a < 0$.

For the converse, suppose $a \leq 1$ and compare the series to the harmonic series $\sum_{k=1}^{\infty} 1/k$. Note that each term of our series is at least as large as the corresponding term in the harmonic series. Since the harmonic series diverges, we conclude by the comparison test that our series also diverges.

(f) From part (b), we have $\mathbf{P}(E_i) = 1/i$, and from part (d) we know that the events $E_1, E_2, \ldots$ are independent. Moreover, by the fact given above part (e), we know that $\sum_{i=1}^{\infty} \mathbf{P}(E_i) = \sum_{i=1}^{\infty} \frac{1}{i} = \infty$. Therefore, by the Borel-Cantelli Lemma for independent events, we can infer that

$$\mathbf{P}(\text{the sequence has an infinite number of records})$$
$$= \mathbf{P}(\text{an infinite number of } E_i\text{'s occur}) = 1.$$

(g) By independence of the two events, we have

$$\mathbf{P}((X_n, X_{n-1}) \text{ is a double record}) = \mathbf{P}(E_n \cap E_{n-1}) = \mathbf{P}(E_n)\mathbf{P}(E_{n-1}) = \frac{1}{n(n-1)}.$$

(h) Note that unlike records, the occurrences of double record pairs are not independent. This is seen clearly because the probability of having a double record $\{X_n, X_{n-1}\}$ is $\frac{1}{n(n-1)}$ and $\{X_{n+1}, X_n\}$ is $\frac{1}{n(n+1)}$, whereas the probability of having both double record pairs is $\frac{1}{(n-1)n(n+1)}$ (since we would need to have records at $n-1, n$, and $n+1$).

(i) Let D_n be the event that (X_n, X_{n-1}) is a double record pair. We know from part (g) that $\mathbf{P}(D_n) = \frac{1}{n(n-1)}$ and from part (h) that the events $D_1, D_2, \ldots$ are not independent. But we can still use the part of the Borel-Cantelli lemma that doesn't require independence. Since the sum $\sum_{n=1}^{\infty} \mathbf{P}(D_n) = \sum_{n=1}^{\infty} \frac{1}{n(n-1)} < \infty$, we conclude that

$$\mathbf{P}(\{X_n\} \text{ has an infinite number of double record pairs}) = 0.$$

(j) In fact, we have not used the actual distribution of the X_n's in any part of our derivations. As long as we draw the X_n's from a continuous i.i.d. distribution, the above results apply.