Pre-Midterm Review Questions

Mark all correct answers in each of the following questions.

1. Consider the lazy secretary problem. Given an assignment of letters to envelopes, a set $\{i_1, i_2, \ldots, i_k\}$ is a *k-cycle* if the letter to person i_1 is sent to person i_2, the letter to i_2 is sent to i_3, ..., the letter to i_k is sent to i_1. Denote by p_k the probability that there exists at least one k-cycle. Thus, for example, we have

$$p_1 = 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + \frac{(-1)^{n-1}}{n!} .$$

   (a) For sufficiently large n we have $p_2 < p_1$.

   (b) We have $p_2 \underset{n\to\infty}{\longrightarrow} 1 - \frac{1}{\sqrt{e}}$.

   (c) We have $p_2 - p_1 \underset{n\to\infty}{\longrightarrow} 0$.

   (d) We have $p_k = \frac{1}{k}$ for $\frac{n}{2} < k \leq n$ and $p_k < \frac{1}{k}$ for $k < \frac{n}{2}$.

   (e) The probability that there exists a k-cycle for some $k > \frac{n}{2}$ is larger than $\frac{1}{2}$ for sufficiently large n.

   (f) The probability that there exists a k-cycle for some $k > \frac{n}{2}$ converges to $\ln 2$ as $n \to \infty$. (Hint: The sum of the first n terms of the harmonic series behaves asymptotically as $\ln n + \gamma$, where $\gamma \approx 0.577$ is Euler's constant.)

   (g) None of the above.

2. Let $A_1, A_2, \ldots$ be events in a probability space $(\Omega, \mathcal{B}, P)$.

   (a) If $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$ for a certain n, then $A_i \bigcap A_j = \emptyset$ for $1 \leq i < j \leq n$.

(b) If $A_i \bigcap A_j = \emptyset$ for $1 \le i < j \le n$, then $P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$.

(c) Denoting $a_r = \sum_{1 \le j_1 < j_2 < \ldots < j_r < \infty} P\left(\bigcap_{i=1}^r A_{j_i}\right)$ for each r, we have

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{r=1}^{\infty} (-1)^{r-1} a_r.$$

(d) The formula in the preceding part is valid if $P(A_n) \underset{n\to\infty}{\longrightarrow} 0$, but is false in general.

(e) If there exists a non-empty event A such that $P(A) = 0$, then P is not a probability function.

(f) Given a set $A \subseteq \mathbf{N}$, put $P(A) = \limsup_{n\to\infty} \frac{|A \bigcap \{1,2,\ldots,n\}|}{n}$ (where $|F|$ denotes the cardinality of a finite set F). Then P defines a probability function on the collection of all subsets of $\mathbf{N}$.

(g) None of the above.

3. Consider the following code fragment:

```
int counter1 = 0, counter2 = 0;
for (int i = 1; i <= n; i++)
   if (Math.random() < 0.5)
      counter1++;
for (int i = 1; i <= n + 1; i++)
   if (Math.random() < 0.5)
      counter2++;
```

Assume that `n` is a variable of type `int`, having some positive value when the above code is executed.

(a) The probability that after the code is executed `counter1` is (strictly) smaller than `counter2` is $1/2$ for any value of `n`.

(b) The probability that after the code is executed `counter1` is smaller than `counter2` is $1/2$ for infinitely many values of `n` and is smaller than $1/2$ for infinitely many others.

(c) The probability that after the code is executed `counter1` is smaller than `counter2` is $1/2$ for infinitely many values of `n` and is larger than $1/2$ for infinitely many others.

(d) If the condition `i <= n` in the first loop is replaced by `i <= m` for some `m` larger than `n`, then the probability in question (strictly) increases.

(e) If the condition `i <= n+1` in the second loop is replaced by `i <= m` for some `m` larger than `n+1`, then the probability in question increases.

(f) If the code fragment in question is embedded in a loop executed repeatedly until `counter1` is smaller than `counter2`, then there is a positive probability the loop will never terminate.

(g) None of the above.

4. Two cards are drawn randomly without replacement from a full deck of cards, and then returned to the deck. The experiment is carried out over and over until for the first time both cards drawn are kings. For $m \geq 0$, denote by K_m the event that during the experiment a total of m kings are drawn (not including the two kings drawn at the last stage). Let Q_m be the event that a total of m queens have been drawn during the experiment, and T_n, $n \geq 1$, the event that the number of pairs of cards drawn is n (including the last drawing).

(a) Since the probability of drawing two kings at any stage is the same, the probability of having an even number of drawings is the same as that of having an odd number of drawings, that is $P(\bigcup_{n=1}^{\infty} T_{2n}) = P(\bigcup_{n=1}^{\infty} T_{2n-1})$.

(b) Since at each drawing (except for the last) we may have two queens, but not two kings, we should expect to see more queens than kings. Consequently, $P(Q_m) \geq P(K_m)$ for each $m \geq 0$, with a strict inequality for at least one value of m.

(c) $P(Q_0) = 1/34$.

(d) $P(Q_m) > 0$ for every $m \geq 0$. In particular, $P(Q_m|T_n) > 0$ for every $m \geq 0$, $n \geq 1$.

(e) $P(K_m) = \left(\frac{31}{32}\right)^{m-1} \cdot \frac{1}{32}, \qquad m \geq 0.$

(f) If the two cards drawn at each stage were drawn with replacement instead of without replacement, $P(T_1)$ was bigger.

(g) None of the above.

5. Alice performs the following experiment. First she rolls 6 dice, then 12 dice, then 18 dice, etc. She stops the experiment at the first stage n at which exactly n out the dice showed "1", exactly n showed "2", etc. Bob performs the same experiment, but stops at the first stage n at which exactly $2n$ out the dice showed either "1" or "2", $2n$ showed either "3" or "4", and $2n$ showed either "5" or "6". Caroline also performs the same experiment, but stops at the first stage n at which exactly $3n$ out the dice showed either one of the numbers "1", "2" and "3".

   (a) Each of the three players will stop with probability 1 within a finite time.

   (b) Alice has a positive probability of never stopping, but Bob and Caroline will stop with probability 1 within a finite time.

   (c) Alice and Bob have a positive probability each of never stopping, but Caroline will stop with probability 1 within a finite time.

   (d) Each of the three players has a positive probability of never stopping.

   (e) The probability that Alice will stop exactly after the second stage is $\frac{5^2\cdot 7\cdot 11^2\cdot 29}{2^{11}\cdot 3^{12}}$.

   (f) There exists a constant $0 < p < 1$ such that, for each $n \geq 1$, the probability that Alice will stop exactly after the nth stage is $(1-p)^{n-1}p$.

   (g) None of the above.

Solutions

1. For any $1 \leq i < j \leq n$, let A_{ij} denote the event that $\{i, j\}$ is a 2-cycle. Thus:

$$p_2 = P\left(\bigcup_{1\leq i<j\leq n} A_{ij}\right).$$

We would like to employ the formula for calculating the probability of a union. Obviously:

$$P(A_{ij}) = \frac{(n-2)!}{n!}, \qquad 1 \le i < j \le n.$$

Moreover, a typical intersection $P\left(\bigcap_{l=1}^{r} A_{i_l j_l}\right)$ is of probability $\frac{(n-2r)!}{n!}$ if all the sets $\{i_l, j_l\}$ are disjoint and is empty otherwise. Since the number of such non-empty intersections is $\frac{1}{r!} \cdot \binom{n}{2,2,\ldots,2,n-2r}$, we have:

$$\begin{aligned} p_2 &= \binom{n}{2} \cdot \frac{(n-2)!}{n!} - \frac{1}{2!}\binom{n}{2,2,n-4} \cdot \frac{(n-4)!}{n!} \\ &\quad + \frac{1}{3!}\binom{n}{2,2,2,n-6} \cdot \frac{(n-6)!}{n!} - \ldots \\ &\quad + (-1)^{r-1}\frac{1}{r!}\binom{n}{2,2,\ldots,2,n-2r} \cdot \frac{(n-2r)!}{n!} + \ldots \\ &\quad + (-1)^{[n/2]-1}\frac{1}{[n/2]!}\binom{n}{2,2,\ldots,2,n-2[n/2]} \cdot \frac{(n-2[n/2])!}{n!} \\ &= \frac{1}{2} - \frac{1}{2!}\cdot\frac{1}{2^2} + \frac{1}{3!}\cdot\frac{1}{2^3} - \ldots + (-1)^{[n/2]-1}\frac{1}{[n/2]!}\cdot\frac{1}{2^{[n/2]}} \,. \end{aligned}$$

Similarly to the lazy secretary problem we obtain:

$$p_2 \underset{n\to\infty}{\longrightarrow} 1 - \frac{1}{\sqrt{e}} \,.$$

As $p_1 \underset{n\to\infty}{\longrightarrow} 1 - \frac{1}{e} > 1 - \frac{1}{\sqrt{e}}$, we have $p_2 < p_1$ for sufficiently large n.

Now notice that any two cycles either coincide or are disjoint. Consequently, for $k > n/2$ there may exist at most one k-cycle. In other words, denoting by $A_{i_1 i_2 \ldots i_k}$ the event that $\{i_1, i_2, \ldots, i_k\}$ is a cycle, all distinct events of this form are disjoint. Now $\{i_1, i_2, \ldots, i_k\}$ is a cycle if and only if any cyclic permutation thereof $\{i_j, i_{j+1}, \ldots, i_k, i_1, \ldots, i_{j-1}\}$ is such. Therefore the number of distinct k-cycles is $n(n-1)\ldots(n-k+1)/k$, which implies:

$$p_k = \frac{n(n-1)\ldots(n-k+1)}{k} \cdot \frac{(n-k)!}{n!} = \frac{1}{k} \,.$$

For $k < n/2$ (and also $k = n/2$ for even n) there may be more than one k-cycle. Thus, the events of the form $A_{i_1 i_2 \ldots i_k}$ are not disjoint. The above calculation still gives that the sum of probabilities of these sets is $1/k$. However, as they intersect the required probability is strictly smaller than the sum of probabilities, namely $p_k < 1/k$.

Obviously, we cannot have more than one cycle of length exceeding $n/2$, so that the probability of the event in part (e) is

$$\begin{aligned} \sum_{n/2<k\le n} \frac{1}{k} &= \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=1}^{[n/2]} \frac{1}{k} = (\ln n + \gamma + o(1)) - (\ln[n/2] + \gamma + o(1)) \\ &= \ln 2 + o(1) \underset{n\to\infty}{\longrightarrow} \ln 2. \end{aligned}$$

In particular, this probability is larger than $\frac{1}{2}$ for sufficiently large n.

Thus, (a), (b), (d), (e) and (f) are true.

2. There may exist non-empty events with 0 probability. For example, take $\Omega = \{1, 2\}$ and define P on all subsets of Ω by:

$$P(A) = \begin{cases} 1, & 1 \in A, \\ 0, & 1 \notin A. \end{cases}$$

P is clearly a probability function, yet $P(\{2\}) = 0$.

The union of (finitely or countably many) sets of 0 probability is also such, even if the sets intersect. Thus the equality $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$ is possible for intersecting sets. (More generally, the equality holds if all intersections $A_i \bigcap A_j$ are of 0 probability.) Of course, as we showed in class, it follows easily from the definition of a probability function that the probability of a finite union of disjoint sets is the sum of probabilities of the sets in the union.

The value of $\sum_{1\le j_1<j_2<\ldots<j_r<\infty} P\left(\bigcap_{i=1}^{r} A_{j_i}\right)$ may well be infinite for an infinite collection $A_1, A_2, \ldots$ of events. For example, taking $A_n = \Omega$ for each n, all intersections in the expression are of probability 1, so that the sum diverges. Even the additional condition $P(A_n) \underset{n\to\infty}{\longrightarrow} 0$ does not guarantee that the sum does not diverge. For example, let (A_n) be a descending sequence of events with $P(A_n) = 1/n$ for each n.

(A possibility for that is to take $\Omega = [0,1]$, the probability of each subinterval of $[0,1]$ being its length, and putting $A_n = [0, 1/n]$.) Then:

$$\sum_{1 \le j_1 < j_2 < \ldots < j_r < \infty} P\left(\bigcap_{i=1}^{r} A_{j_i}\right) \ge \sum_{n=r}^{\infty} \frac{1}{n} = \infty.$$

Taking $P(A) = \limsup_{n\to\infty} \frac{|A \bigcap \{1,2,\ldots,n\}|}{n}$ clearly gives $P(A) = 0$ for every finite set A and $P(\mathbf{N}) = 1$. Thus

$$P\left(\bigcup_{n=1}^{\infty} \{n\}\right) = P(\mathbf{N}) = 1,$$

whereas

$$\sum_{n=1}^{\infty} P(\{n\}) = \sum_{n=1}^{\infty} 0 = 0.$$

It follows that P is not σ-additive. (Actually, with slightly more work one can show that P is not even finitely additive. Take a set constructed by taking large consecutive blocks of integers, followed by much larger holes, followed by much larger blocks, etc. For example, take

$$A = \bigcup_{n=1}^{\infty} \{(2n)!, (2n)! + 1, \ldots, (2n+1)! - 1\}.$$

Then both $P(A) = 1$ and $P(\overline{A}) = 1$. The point is that the sets A and $\overline{A}$ "seem" large when measured at certain times, and small at others. Taking the lim sup gives both of them the measure they tend to assume when they are large, and ignores times when they are small. Since these sets are alternately small and large, with one being small exactly when the other is large, the lim sup "misses" when measuring their sizes.)

Thus, only (b) is true.

3. The first loop puts in `counter1` the number of heads obtained out of n tosses of a coin. The second loop does the same for `counter2`, but for $n + 1$ tosses instead of n. Thus the probability that `counter1` is smaller than `counter2` is the same as that of a player tossing a coin n

times to obtain less heads than a player tossing it $n+1$ times. This probability has been computed in class to be $1/2$ for any value of n.

Replacing n in the first loop by a larger integer is tantamount to letting the first player toss his coin more times, so that his probability of having less heads than his opponent decreases. If $n+1$ in the second loop is increased, then the opposite happens.

Since the probability for `counter1` to be smaller than `counter2` is $1/2$, if the experiment is held repeatedly, the probability for this event to happen m consecutive times is $1/2^m$. It follows that the probability for the loop to never terminate is 0.

Thus, only (a) and (e) are true.

4. The probability of obtaining two kings at a single drawing of two cards is $\frac{4 \cdot 3}{52 \cdot 51} = \frac{1}{221}$. Hence:

$$P(T_n) = \left(\frac{220}{221}\right)^{n-1} \cdot \frac{1}{221}, \qquad n \geq 1.$$

Consequently:

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} T_{2n-1}\right) &= \frac{1}{221} \cdot \left(1 + \left(\frac{220}{221}\right)^2 + \left(\frac{220}{221}\right)^4 + \dots\right) \\ &= \frac{1}{221} \cdot \frac{1}{1-(220/221)^2} = \frac{221}{441}, \end{aligned}$$

and therefore

$$P\left(\bigcup_{n=1}^{\infty} T_{2n}\right) = 1 - \frac{221}{441} = \frac{220}{441}.$$

Expecting to see more queens than kings does not mean that $P(Q_m) \geq P(K_m)$ for each m. It means rather (roughly speaking) that this inequality holds for large values of m, whereas for small values of m the converse inequality holds. Indeed, since $\sum_{m=0}^{\infty} P(Q_m) = \sum_{m=0}^{\infty} P(K_m) = 1$, the inequalities suggested in (b) are impossible.

To calculate $P(Q_0)$ we use conditioning on the stage at which the two drawn cards are kings:

$$
\begin{aligned}
P(Q_0) &= \sum_{n=1}^{\infty} P(T_n) P(Q_0 | T_n) \\
&= \sum_{n=1}^{\infty} \left(\frac{220}{221}\right)^{n-1} \cdot \frac{1}{221} \cdot \left(\frac{48 \cdot 47 - 4 \cdot 3}{52 \cdot 51 - 4 \cdot 3}\right)^{n-1} \\
&= \frac{1}{221} \cdot \sum_{n=0}^{\infty} \left(\frac{187}{221}\right)^{n-1} = \frac{1}{34} .
\end{aligned}
$$

(A simpler way of finding $P(Q_0)$ is as follows. Suppose we stop the experiment at the first stage when either one of the cards is a queen or both are kings. Our question is equivalent to asking for the probability that this new experiment will be stopped due to drawing two kings. Now in the new experiment we may ignore all outcomes except for either drawing at least one queen or drawing two kings. Drawing two queens arises in $52 \cdot 51 - 48 \cdot 47$ possibilities, whereas drawing two kings arises in $4 \cdot 3$ possibilities. Hence $P(Q_0) = \frac{4 \cdot 3}{52 \cdot 51 - 48 \cdot 47} = \frac{1}{34}$.)

It is true that $P(Q_m) > 0$ for every $m \geq 0$. However, under T_n we draw only $2n-2$ cards before the final two kings, so that $P(Q_m | T_n) = 0$ for $m \geq 2n-1$.

We calculate $P(K_m)$ similarly to the calculation of $P(Q_0)$ above:

$$
\begin{aligned}
P(K_m) &= \sum_{n=1}^{\infty} P(T_n) P(K_m | T_n) \\
&= \sum_{n=1}^{\infty} \left(\tfrac{220}{221}\right)^{n-1} \cdot \frac{1}{221} \\
&\qquad \cdot \binom{n-1}{m} \left(\tfrac{52 \cdot 51 - 48 \cdot 47 - 4 \cdot 3}{52 \cdot 51 - 4 \cdot 3}\right)^{m} \left(\tfrac{48 \cdot 47}{52 \cdot 51 - 4 \cdot 3}\right)^{n-1-m} \\
&= \frac{1}{221} \cdot \left(\frac{8}{47}\right)^{m} \sum_{n=0}^{\infty} \binom{n}{m} \left(\frac{188}{221}\right)^{n} \\
&= \frac{1}{221} \cdot \left(\frac{8}{47}\right)^{m} \frac{(188/221)^m}{(33/221)^{m+1}} = \left(\frac{32}{33}\right)^{m} \cdot \frac{1}{33} .
\end{aligned}
$$

If the two cards at each stage are drawn with replacement, then $P(T_1) = (4/52)^2 = 1/169 > 1/221$, so that we would obtain a larger value for $P(T_1)$ than is the case for our experiment.

Thus, only (c) and (f) are true.

5. Let α_n be the probability of Alice to stop the experiment at stage n, assuming she indeed got to this stage, and let β_n and γ_n be the corresponding probabilities for Bob and Caroline, respectively. Clearly:

$$\alpha_n = \frac{\binom{6n}{n,n,n,n,n,n}}{6^{6n}}, \qquad \beta_n = \frac{\binom{6n}{2n,2n,2n}}{3^{6n}}, \qquad \gamma_n = \frac{\binom{6n}{3n,3n}}{2^{6n}} .$$

The probability for Alice to continue the experiment beyond the first n stages is $\prod_{k=1}^{n}(1-\alpha_k)$. Thus her probability to continue forever is 0 if and only if $\prod_{k=1}^{\infty}(1-\alpha_k)=0$. According to Exercise 38(b), since all α_k's are strictly between 0 and 1, we only have to find whether the series $\sum_{k=1}^{\infty}\alpha_k$ diverges or converges. Now, according to Stirling's formula

$$\binom{6n}{n,n,n,n,n,n} = \frac{(6n)!}{(n!)^6} \approx \frac{\sqrt{2\pi\cdot 6n}\left(\frac{6n}{e}\right)^{6n}}{\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right)^6} = \frac{C_1\cdot 6^{6n}}{n^{5/2}}$$

for some positive constant C_1, which yields

$$\alpha_n \approx \frac{C_1}{n^{5/2}} .$$

Since the series $\sum_{n=1}^{\infty}\frac{1}{n^{5/2}}$ converges, Alice will continue the experiment indefinitely with positive probability. Similarly

$$\beta_n \approx \frac{C_2}{n} ,$$

and

$$\gamma_n \approx \frac{C_3}{n^{1/2}} .$$

As the series $\sum_{n=1}^{\infty}\frac{1}{n}$ and $\sum_{n=1}^{\infty}\frac{1}{n^{1/2}}$ both diverge, Bob and Caroline will stop with probability 1 after a finite time.

Alice will stop after the nth stage with probability $\left(\prod_{k=1}^{n-1}(1-\alpha_k)\right)\alpha_n$. For $n=2$ we obtain

$$\left(1-\frac{6!}{6^6}\right)\cdot\frac{\binom{12}{2,2,2,2,2,2}}{6^{12}} = \frac{5^2\cdot 7\cdot 11^2\cdot 29}{2^{10}\cdot 3^{11}}.$$

Thus, only (b) is true.