Randomness and Computation 2018/19
Coursework 2 (summative)

Issue date: Monday, 4th March, 2019

The deadline for this coursework is 4pm on Tuesday, 19th March 2019 (Tuesday of week 9). Please submit your solutions either *electronically via* `submit` *or by hand to the ITO in Appleton Tower.* **Remember that the School's policy on late coursework means that late coursework incurs a penalty unless you have good justification for the lateness (as per the "good reasons" definitions on our webpages). If you believe you do satisfy the "good reasons" conditions, you should request an extension (well in advance) through the ITO.**

This coursework should be your own individual work. *You may discuss understanding of the questions with your classmates, but may not share solutions, or give strong hints. If you use any resources apart from the course slides/notes, or the book, you* must *cite these in detail (on a per-question basis.)*

This coursework is worth 20% of your final course average.

1. We consider the number of *acyclic orientations* in (undirected) Erdős-Rényi graphs.

   Definition 1 *For any directed graph* $\vec{G} = (V, \vec{E})$*, we say that* $\vec{G}$ *is* acyclic *if there is no directed cycle in* $\vec{G}$*. Such a graph is sometimes called a* directed acyclic graph (DAG)*.*

   Definition 2 *For any given (simple) undirected graph* $G = (V, E)$*, an* orientation $\vec{G}$ *of* G *is any directed graph* $\vec{G} = (V, \vec{E})$ *such that* $|\vec{E}| = |E|$ *and such that for every* $(u, v) \in E$*, exactly* one *of the arcs* $(u \to v)$ *and* $(v \to u)$ *belongs to* $\vec{E}$*.*

   Definition 3 *For any simple* undirected *graph* $G = (V, E)$*, the set of* acyclic orientations (AOs) *of* G*, denoted* $AO(G)$*, is the set of all orientations* $\vec{G}$ *of* G *which are acyclic. Such a graph is sometimes called a* directed acyclic graph (DAG)*.*

   The problem of *counting the number of acyclic orientations* of a given undirected graph is a well-known $\sharp$P-complete ("hard to count") problem. We will consider the problem when the input graph is drawn from the *random model* $\mathcal{G}_{n,p}$ and show how to evaluate the *expected* number of AOs in polynomial-time. Unlike some simple structures that we will count in $G_{n,p}$ in our lectures (eg, clique on 4 vertices), we will not prove an exact value for $E_{n,p}[|AO(G)|]$; instead we will derive an algorithm to compute the expectation, given n, p.

   Definition 4 *The Erdős-Rényi model* $\mathcal{G}_{n,p}$ *of random graphs is parametrized by the number of vertices* n*, and an edge addition probability* $p \in [0, 1]$*.*

   We generate a undirected simple graph $G = (V, E)$ *from* $\mathcal{G}_{n,p}$ *by setting* $V = \{1, \ldots, n\}$*. To construct the edge set* E*, we consider each* (i, j) *pair* $1 \le i < n$*,* $i < j \le n$ *independently, and we add the undirected edge* (i, j) *to* E *with probability* p *(omitting it with probability* $1 - p$*).*

The Erdős-Rényi process always creates a simple graph without loops or parallel edges. The particular number of edges added will vary depending on the results of the independent trials for the edges. The *expected* number of edges, written $\mathbb{E}_{n,p}[|E|]$ has the value $p \cdot \binom{n}{2}$. Also, for a particular "number of edges" m which might be generated, all simple graphs with m edges have the same probability (which is $p^m(1-p)^{\binom{n}{2}-m}$) in the Erdős-Rényi model for n, p.

We now present a recurrence which was proved independently by two famous combinatoricists in the early 1970s. We need one more definition first:

Definition 5 *Let $n, m \in \mathbb{N}$. The number of different* directed acyclic graphs (DAGs) *on n vertices with m arcs, is denoted by $A_{n,m}$.*

We never compute the $A_{n,m}$ values; we just need to work with the definition.

Theorem 6 (Robinson, Stanley) *Suppose we define the quantity $A_n(x)$ by*

$$A_n(x) = \sum_{m=0}^{\binom{n}{2}} A_{n,m} x^m.$$

(the co-efficients of $A_n(x)$ are the counts of DAGs with m arcs for various m). Then

$$A_n(x) \quad = \quad \sum_{i=1}^{n} (-1)^{i+1} \binom{n}{i} (1+x)^{i(n-i)} A_{n-i}(x).$$

We will *not* be proving the correctness of this recurrence, as it is difficult. We will be *applying* it to compute the *expected number of AOs* of a random graph G drawn from $G_{n,p}$.

The expected number of AOs for a random graph from $\mathcal{G}_{n,p}$ is defined as

$$\mathbb{E}_{n,p}[|AO(G)|] = \sum_{G=(V,E),|V|=n} \mathbb{P}r_{n,p}[G] \cdot |AO(G)|,$$

where $\mathbb{P}r_{n,p}[G]$ is the probability that G is generated by $\mathcal{G}_{n,p}$.

(a) Prove that the Robinson-Stanley polynomial $A_n(x)$, when evaluated at $x = \frac{p}{1-p}$ for any $p \in (0, 1)$, satisfies the following equality: *[15 marks]*

$$A_n\left(\frac{p}{1-p}\right) = (1-p)^{-\binom{n}{2}} \times \mathbb{E}_{n,p}[|AO(G)|].$$

You just need the definition of $A_n(x)$ to show this. You will not need to use induction*, or* proof by contradiction*, just manipulation and simplification of equations. The proof will only be about 5-6 lines long if done right.*

(b) Design an $O(n^2)$-time (or at worst, an $O(n^3)$-time) algorithm) to evaluate $\mathbb{E}_{n,p}[|AO(G)|]$ exactly for given values n, p. Justify the running time in detail. *[10 marks]*

For this you will need the result of part (a) (not the proof), and the recurrence of Thm 6. Think about how to calculate the $A_i(x)$ values for given x in order of increasing i.

2. Suppose we have access to some Bernoulli random variable Y with unknown parameter $p \in (0,1)$ ($p = \Pr[\mathsf{Y} = 1]$), and we want to solve two problems using repeated sampling from Y, given an natural number N as input:

   (a) To compute a value $\widehat{p} \in (0,1)$ such that $\Pr\left[|\widehat{p} - p| \leq \frac{1}{N}\right]$ is at least $1 - \frac{1}{N}$; *[10 marks]*

   (b) To determine, with probability at least $1 - \frac{1}{N}$, whether $p \leq \frac{1}{4}$ or not. *[10 marks]*

   For each of these two problems, determine whether the result can be achieved by drawing a polynomial-number of samples from Y. If the result is possible, compute the minimum number of samples needed to guarantee the result, subject to the use of the Chernoff bounds of Corollary 4.6. If you think the result cannot be achieved with polynomially-many samples, give your reasons.

3. Again consider the Erdős-Rényi model $G_{n,p}$ model of random graphs (where every edge is included independently and identically with probability p).

   We will be interested in the presence (or absence) of a *4-cycle* in the random graph - a 4-cycle being a set of four vertices which can be arranged so that all 4 outer induced edges belong to the generated graph G, but neither of the two "crossing edges" belong to G. We will consider the set of indicator variables $\{X_f : f \subseteq [n], |f \cap [n]| = 4\}$, where X_f will be 1 for $f = \{u_f, v_f, w_f, z_f\}$ iff these vertices can be ordered so that the 4 "outer edges" lie in the generated graph $G \leftarrow G_{n,p}$, but the two "crossing edges" are absent from G.

   We will be interested in the total number of 4-cycles $X = \sum_{f \subseteq [n], |f \cap [n]| = 4} X_f$, and the expectation and variance of X.

   (a) Derive an exact expression for the expected number of 4-cycles $\mathrm{E}[X]$ in the random graph. This will depend on n and p. *[5 marks]*

   (b) Using the $\mathrm{E}[X]$ value from (b), show that for any $p = p(n)$ that satisfies $pn \to 0$ as $n \to \infty$, that $\Pr[X > 0] \to 0$. *[5 marks]*

   (c) Derive a close estimate of the second moment $\mathrm{E}[X^2]$, taking into account the different cases for pairs of 4-cycles to overlap or not-overlap. Hence derive a bound on the variance $\mathrm{Var}[X]$. *[10 marks]*

   (d) Using your result from (c), use the special case of Chebyshev's Inequality with $a = \mathrm{E}[X]$ to derive a sufficient condition on $p = p(n)$ (with respect to n) to ensure $\Pr[X > 0] \to 1$ (equivalently, that $\Pr[X = 0] \to 0$). *[5 marks]*

   This question requires the content of Lectures 11-12 on "the probabilistic method".

4. In this question we consider a *Markov chain* on the set of *contingency tables*. A collection of *contingency tables* is defined in terms of two lists $r = (r_1, \dots, r_m)$ and $c = (c_1, \dots, c_n)$ of positive integer values. These lists are considered to be the *row sums* (the r_is) and *column sums* (the c_js) of hypothetical $m \times n$ tables of non-negative integers. For the given lists r, c, the *set of contingency tables* $\Sigma_{r,c}$ is defined as the set of all $X \in \mathbb{N}_0{}^{mn}$ that satisfy the following conditions:

$$X_{i,j} \geq 0 \quad \text{for all } i \in [m], j \in [n] \tag{1}$$

$$\sum_{j=1}^{n} X_{i,j} = r_i \quad \text{for all } i \in [m] \tag{2}$$

$$\sum_{i=1}^{m} X_{i,j} = c_i \quad \text{for all } j \in [n] \tag{3}$$

In diagrammatic form, we are interested in all the $m \times n$ tables $X \in \mathbb{N}_0{}^{mn}$ that have the given row rums $r_1, \dots, r_m$ and given column sums $c_1, \dots, c_n$:

$X_{1,1}$	$X_{1,2}$	$X_{1,3}$	$\dots$	$X_{1,n}$	r_1
$X_{2,1}$	$X_{2,2}$	$X_{2,3}$	$\dots$	$X_{2,n}$	r_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$X_{m,1}$	$X_{m,2}$	$X_{m,3}$	$\dots$	$X_{m,n}$	r_m
c_1	c_2	c_3	$\dots$	c_n	

Note that our given row and column sums *must* satisfy $\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j$, otherwise we have $\Sigma_{r,c} = \emptyset$ (no feasible solutions).

Above is the description of the set of contingency tables $\Sigma_{r,c}$. We now define our Markov chain on the elements of $\Sigma_{r,c}$. We name the chain M, and for every two contingency tables $X, Y \in \Sigma_{r,c}$, there will be some probability $M[X, Y]$ (maybe 0) that we move from table X to table Y in a single transition.

The transitions (single-step moves) of the Markov chain are a result of the following process: we choose two rows $i, i', i \neq i'$ from $[m]$ independently at random, and two columns $j, j', j \neq j'$ from $[n]$ independently at random. This random draw identifies a "mini-table" of dimensions 2×2 (note i, i' don't have to be adjacent, and neither do j, j'):

$X_{1,1}$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$X_{1,n}$	r_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\dots$	$\dots$	$\mathbf{X}_{i,j}$	$\dots$	$\mathbf{X}_{i,j'}$	$\dots$	$\dots$	r_i
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\dots$	$\dots$	$\mathbf{X}_{i',j}$	$\dots$	$\mathbf{X}_{i',j'}$	$\dots$	$\dots$	$r_{i'}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$X_{m,1}$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$X_{m,n}$	r_m
c_1	$\dots$	c_j	$\dots$	$c_{j'}$	$\dots$	c_n	

This 2×2 "mini-table" can be visualised independently of the remainder of that table. It will have some individual small row and column sums (depending on the current X, of course):

$X_{i,j}$	$X_{i,j'}$	a_1
$X_{i',j}$	$X_{i',j'}$	a_2
b_1	b_2	

Claim: Focusing on the 2×2 table mapped out by i, i' and j, j', the set of all 2×2 tables with row sums a_1, a_2, and column sums b_1, b_2 is

$$\Sigma_{a,b} = \left\{ \begin{bmatrix} i & (a_1 - i) \\ (b_1 - i) & i + (b_2 - a_1) \end{bmatrix} : \max\{0, a_1 - b_2\} \leq i \leq \min\{a_1, b_1\} \right\}.$$

Transitions: Conditional on having already chosen i, i', j, j' as our "rows and columns", we replace the 2×2 "mini-table" by a uniform random element of $\Sigma_{a,b}$.

The Markov chain M on $\Sigma_{r,c}$: We will have $M[X, Y] > 0$ for any X, Y such that the non-0 entries of $X - Y$ (the differing entries of X versus Y) lie in some 2×2 sub-matrix. Note this includes the possibility of $X = Y$. Note that for any such pair X, Y that satisfies this condition, the probability $M[X, Y]$ will be $\frac{2}{m(m-1)} \frac{2}{n(n-1)}$, multiplied by $\frac{1}{|\Sigma_{a,b}|}$ for the specific a_1, a_2, b_1, b_2 of that 2×2 subtable.

If X, Y are such that the non-zero values of $X - Y$ *don't* fit into a 2×2 mini-table, then $M[X, Y] = 0$.

Your task in this exercise is to prove the necessary steps to show that M has a *unique stationary distribution* over $\Sigma_{r,c}$, and that this is the uniform distribution. We will discussing Markov chains in class starting week 7 (Friday 8th March and onwards); however, you will be guided through this exercise, and don't actually need the material from those lectures to start working on this question.

(a) Prove the **"Claim"** above (that the elements listed for $\Sigma_{a,b}$ are exactly the subtables of non-negative integer 2×2 tables which satisfy row sums a_1, a_2 and column sums b_1, b_2). *[5 marks]*

(b) Prove that, for any pair of contingency tables $X, Y \in \Sigma_{r,c}$ such that $X \neq Y$, that we can find a sequence of tables *[10 marks]*

$$X = Z^0, Z^1, \ldots, Z^{\ell-1}, Z^\ell = Y$$

leading from X to Y such that $Z^j \in \Sigma_{r,c}$ for all $j = 0, \ldots, \ell$, and such that $M[Z^j, Z^{j+1}] > 0$ for every $j = 0, \ldots, \ell - 1$.

note: Once this has been proven, you will have shown the property of *"irreducibility"* for this specific Markov chain.

[**Hint:** The way to prove this to consider the "metric" $d(\cdot, \cdot)$ on $\Sigma_{r,c} \times \Sigma_{r,c}$ defined as $d(Z, W) =_{\text{def}} \sum_{i=1}^{m} \sum_{j=1}^{n} |W_{ij} - Z_{ij}|$, and show that for every pair $X, Y \in \Sigma_{r,c}, X \neq Y$, we can find a transition of the Markov chain that we can apply to $Z^0 = X$ to get Z^1 such that $d(Z^1, Y) < d(X, Y)$. Then apply induction to build the entire sequence.]

(c) For any $X, Y \in \Sigma_{r,c}$, there are probably many sequences $X = Z^0, Z^1, \ldots, Z^\ell = Y$ of M-transitions which could map X to Y. These can be of varying lengths, and also the probabilities of a particular path to Y being taken will vary. Let $\mathcal{P}_{X,Y}$ be the set of all sequences of states which will connect X to Y via M. *[10 marks]*

We say that the Markov chain is *aperiodic* if it is the case that for *every* $X, Y \in \Sigma_{r,c}$, that $\gcd\{\ell(P) : P \in \mathcal{P}_{X,Y}\} = 1$.

Prove that our chain M is aperiodic.

[**Hint:** This is not a long unwieldy proof. You need one observation about this Markov chain, applied carefully.]

(d) Markov chains like this one are designed in such a way as to encourage changes in the current state X. We do not have any concept of "quality" of a particular state, and over a number of steps of the Markov chain we are not aiming to arrive at any particular state, but instead to become more and more "random" in the pool of possible states $\Sigma_{r,c}$. There is a concept of a *stationary distribution* in probability theory, which is any row vector π of length $|\Sigma_{r,c}|$, such that $\pi(X) \geq 0$ for all $X \in \Sigma_{r,c}$, and $\sum_{X \in \Sigma_{r,c}} \pi(X) = 1$, and *[5 marks]*

$$\pi \cdot M = \pi.$$

Note that in the equation above, we use M to denote the $|\Sigma_{r,c}| \times |\Sigma_{r,c}|$ matrix such that $M[X, Y]$ is the probability of transitioning from X to Y in a single step.

It is well-known from standard probability theory, that any Markov chain on a finite state space which is *both* irreducible and aperiodic, has a *unique* stationary distribution. By our work in (b) and (c), we have already shown that M has a unique stationary distribution.

In this part (d), you are asked to prove that this stationary distribution is the *uniform* stationary distribution.

[**Hint:** Consider the uniform distribution on $\Sigma_{r,c}$, with $\pi(X) = \frac{1}{|\Sigma_{r,c}|}$, and verify that it satisfies the condition of a stationary distribution for M.]

Please submit your work in advance of 4pm, Tuesday 18th March *either* in person to the ITO (in Appleton Tower), *or* electronically from DICE using the following `submit` command (for your file named `coursework2.pdf`):

```
submit rc cw2 coursework2.pdf
```

Mary Cryan, 4th March 2018