Probability and Computing 2 Sep, 2014

Problem Set 1

- Due Date: **16 Sep (Tue), 2014**
- Turn in your problem sets electronically (LaTeX, pdf or text file) by email. If you submit handwritten solutions, start each problem on a fresh page.
- Collaboration is encouraged, but all writeups must be done individually and must include names of all collaborators.
- Refering sources other than the text book and class notes is strongly discouraged. But if you do use an external source (eg., other text books, lecture notes, or any material available online), ACKNOWLEDGE all your sources (including collaborators) in your writeup. This will not affect your grades. However, not acknowledging will be treated as a serious case of academic dishonesty.
- The points for each problem are indicated on the side.
- Be clear in your writing.
- 15 problems - 10 marks each

1. Exercises Ex 1.7, 1.11, 1.12, 1.10, 1.14, 2.8, 2.9, 2.18, 2.19, 2.22, 2.25, 2.26, in Mitzenmacher-Upfal
2. Problems 27 and 35 in Chap 1, Grimmet-Stirzaker
3. **Sampling a uniform number and its factorization**

   Consider the following algorithm presented in class to sample a number uniformly in the range $\{1, 2, \ldots, N\}$ along with its prime factorization.

```
• Input: Positive integer $N$
  (a) Set $S_0 \leftarrow N, i \leftarrow 0$.
  (b) While $S_i \neq 1$ do
      i. $S_{i+1} \leftarrow RandInt[S_i]$
      ii. Set $i \leftarrow i + 1$
  (c) Set $R \leftarrow \prod_{i:i\geq 1, S_i is \text{ prime}} S_i$
  (d) If $R \leq N$
      i. $X \leftarrow Bernouli(R/N)$
      ii. If $X = 1$, output $R$ with its factorization else goto Step 3a
  (e) else goto Step 3a.
```

Let $(S_1, S_2, \ldots, , S_k = 1)$ be the sequence of integers obtained as mentioned in the above algorithm. For integers $i \in \{2, 3, \ldots, N\}$ let X_i be the number of times i appears in the sequence. Let $2 = p_1 < p_2 \cdots < p_k$ be the sequence of prime numbers less than or equal to N.

(a) Show that $\Pr[X_N > 0] = 1/N$.

(b) Show that $\Pr[X_N = k] = \frac{1}{N^k}\left(1 - \frac{1}{N}\right)$.

(c) Generalize to show that for all i, $\Pr[X_i = k] = \frac{1}{i^k}\left(1 - \frac{1}{i}\right)$.

(d) Show that the above is unaltered when conditioned on events depending on X_j for $j > i$. In particular, show that

$$\Pr[X_i = k | X_{i+1} = k_{i+1}, X_{i+2} = k_{i+3}, \ldots, X_N = k_N] = \frac{1}{i^k}\left(1 - \frac{1}{i}\right).$$

(e) Argue from above that

$$\Pr[R = p_1^{e_1} \cdot p_2^{e_2} \ldots p_k^{e_k}] = \frac{1}{\prod_{i=1}^{k} p_i^{e_i}} \cdot \prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right).$$

(f) Conclude that the above algorithm samples a number uniformly in $[N]$ along with its factorization.