Homework 10 Solution

R-19.3

$\mu = E(X) = \sum_{i=1}^{n} p_i = 0.02 * 10^6 = 2 * 10^4$

$\Pr(X > (1+\delta)\mu \le \left[\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right]^{\mu}$, where δ > 0. Here, $\delta = 1$,

Pr (X > 2*2 * $10^4 \le \left[\frac{e^{\sigma}}{(1+\sigma)^{(1+\sigma)}}\right]^{\mu} = \left[\frac{e^{1}}{(2\,)^{2}}\right]^{2*10^4} = \left[\frac{e}{4}\right]^{20000}$

Using a Chernoff bound, the probability that more than 4% of the 1 million children born in a given large city have this birth defect is bound by 0.6796^{20000}

C-19.4

Suppose n = 3. We start with 123. After 1 step, we get 123, 132, and 321, each with probability 1/3. In step 2, from 123, we get 123, 213, and 132, each with probability 1/9, and so on. In the end we get each permutation of 123 occurring with some probability of the form i/27, where i is the number of times we can get that permutation following the depth-3 tree of possibilities, which has degree-3 and 27 external nodes. But we need each permutation to occur with probability 1/3! = 1/6, and there is no way to make a fraction of the form i/27 equal to 1/6 with i being an integer.

A-19.3

(a) It is given that the probability that a router perform probabilistic packet marking is p $\le$ 1/2. Now, for the packet to survive with a mark, which is generated from a farthest router is only possible when other routers do not perform probabilistic packet marketing on that packet and probability of that is given by (1 – p). Hence, the total probability that the router farthest from the recipient will mark a packet and this mark will survive all the way to the recipient is **$p\,(1-p)^{d-1}$**, where d is the total number of routers.

(b) This problem is same as coupon collector problem. Let $X = X_1 + X_2 + \ldots + X_d$ be a random variable that we need to collect to identify all d routers, where X_i represents the number of packets we need to collect in order to go from having i-1 distinct router addresses to having i distinct addresses. After getting i-1 distinct addresses, the chances of getting new one is $p_i = \frac{d-(i-1)}{d}$ and the expected value of $E[X_i] = 1/\,p_i$. By linearity of expectation, $E[X] = n\,H_n$, where H_n is the nth harmonic number, which, can be approximated as $\ln d \le H_n \le \ln d + 1$. Now, according to tail estimate, recipient will receive all the addresses after collecting **(cd ln d) packets for c ≥ 2**. Thus, the upper bound on the expected number of packets that the recipient needs to collect to get all d routers addresses is **O (d log d).**