1 Packet Routing

(a) The distribution of the packets arriving at the switch is $S \sim \text{Poisson}(\lambda)$. We would like to determine the distribution of A, where $S = A + B$. Given that $S = s = a + b$ packets arrive at the switch, where a packets are routed to A and b packets are routed to B, the number of packets sent to A is distributed as $A \sim \text{Binomial}(s, p)$. As such, we have that

$$\begin{aligned}
p_A(a) &= \sum_{s=0}^{\infty} p_{A,S}(a, s) \\
&= \sum_{s=0}^{\infty} p_{A|S}(a \mid s) \cdot p_S(s) \\
&= \sum_{s=a}^{\infty} \binom{s}{a} p^a (1-p)^{s-a} \cdot e^{-\lambda} \frac{\lambda^s}{s!} \\
&= \frac{p^a \cdot e^{-\lambda}}{a!} \sum_{s=a}^{\infty} \frac{s!}{(s-a)!} (1-p)^{s-a} \cdot \frac{\lambda^s}{s!} \\
&= \frac{(\lambda p)^a \cdot e^{-\lambda}}{a!} \sum_{s=a}^{\infty} \frac{((1-p)\lambda)^{s-a}}{(s-a)!} \\
&= \frac{(\lambda p)^a \cdot e^{-\lambda}}{a!} \cdot e^{(1-p)\lambda} \\
&= e^{-\lambda p} \frac{(\lambda p)^a}{a!}.
\end{aligned}$$

This shows that $p_A \sim \text{Poisson}(\lambda p)$.

(b) Yes. In determining the joint distribution, we find that

$$\begin{aligned}
p_{A,B}(a, b) &= p_{A,B|S}(a, b \mid a + b) \cdot p_S(a + b) \\
&= p_{A|S}(a \mid a + b) \cdot p_S(a + b) \\
&= \binom{a+b}{a} p^a (1-p)^b \cdot e^{-\lambda} \frac{\lambda^{a+b}}{(a + b!)} \\
&= \frac{(a+b)!}{a!b!} p^a (1-p)^b \cdot e^{-\lambda} \frac{\lambda^{a+b}}{(a+b)!} \\
&= e^{-\lambda p} \frac{(\lambda p)^a}{a!} \cdot e^{-\lambda(1-p)} \frac{(\lambda(1-p))^b}{b!} \\
&= p_A(a) \cdot p_B(b).
\end{aligned}$$

Where we have substituted $p_B(b) = e^{-\lambda(1-p)} \frac{(\lambda(1-p))^b}{b!}$, which can be derived similarly as in part (a). In particular, since the joint distribution $p_{A,B}(a, b)$ is the product of the marginals $p_A(a)$ and $p_B(b)$, we find that the number of packets routed to A and B are independent.

2 Compact Arrays

Let $A = A_1 + A_2 + \cdots + A_{i-1}$ be a random variable indicating the number of 0s in the first $i - 1$ entries of the array. Each $A_j \mid j \in [1, i-1]$ is then an indicator variable with probability $\frac{1}{10}$ of being

1 (jth entry is 0) and $\frac{9}{10}$ of being 0 (jth entry is not 0). We are then looking for the expectation of $X = i - A$, since the element at index i will move up by however many 0s there are ahead of it. We get that

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}[i - A] \\ &= i - \mathbb{E}[A] \\ &= i - \frac{i-1}{10} \\ &= \frac{9i+1}{10}\end{aligned}$$

by linearity of expectation. Furthermore, since all the A_j are independent, we have that

$$\begin{aligned}\mathrm{var}(X) &= \mathrm{var}(i - A) \\ &= \mathrm{var}(A) \\ &= \sum_{j=1}^{i-1} \mathrm{var}(A_j) \\ &= (i-1)\left(\mathbb{E}[A_j^2] - (\mathbb{E}[A_j])^2\right) \\ &= (i-1)\left(\frac{1}{10} - \left(\frac{1}{10}\right)^2\right) \\ &= \frac{9(i-1)}{100}.\end{aligned}$$

3 Message Segmentation

(a) $p_{Q,R}(q,r) = p_N(mq+r) = (1-p)^{mq+r-1}p$

(b) For $q > 0$,

$$\begin{aligned}p_Q(q) &= \sum_{r=0}^{m-1} p_{Q,R}(q,r) \\ &= \sum_{r=0}^{m-1} (1-p)^{mq+r-1}p \\ &= p \cdot \left(\frac{(1-p)^{mq-1}}{1-(1-p)} - \frac{(1-p)^{mq+m-1}}{-(1-p)}\right) \\ &= \left((1-p)^{mq-1}\right)(1-(1-p)^m)\end{aligned}$$

For $q = 0$,

$$\begin{aligned}p_Q(q) &= \sum_{r=1}^{m-1} p_{Q,R}(q,r) \\ &= \sum_{r=1}^{m-1} (1-p)^{mq+r-1}p\end{aligned}$$

$$
\begin{aligned}
&= p \cdot \left(\frac{(1-p)^{mq}}{1-(1-p)} - \frac{(1-p)^{mq+m-1}}{1-(1-p)} \right) \\
&= ((1-p)^{mq}) \left(1-(1-p)^{m-1}\right) \\
&= \left(1-(1-p)^{m-1}\right)
\end{aligned}
$$

$$p_R(r) = \sum_{q=0}^{\infty} p_{Q,R}(q,r) = \sum_{q=0}^{\infty} (1-p)^{mq+r-1} p = p \cdot \frac{(1-p)^{r-1}}{1-(1-p)^m}$$

(c) For $q > 1$,

$\mathbb{P}(Q = q \mid N > m) = \sum_{r=0}^{m-1} \mathbb{P}(N = mq + r \mid N > m) = \sum_{r=0}^{m-1} \mathbb{P}(N = mq + r) =$ $\left((1-p)^{mq-1}\right)(1-(1-p)^m)$, by the memoryless property.

For $q = 1$,

$\mathbb{P}(Q = q \mid N > m) = \sum_{r=1}^{m-1} \mathbb{P}(N = m + r \mid N > m) = \sum_{r=1}^{m-1} \mathbb{P}(N = r) = \sum_{r=1}^{m-1} (1-p)^{r-1} p =$ $p \cdot \left(\frac{1}{(1-(1-p))} - \frac{(1-p)^{m-1}}{1-(1-p)} \right) = \left(1-(1-p)^{m-1}\right)$, by the memoryless property.

$\mathbb{P}(R = r \mid N > m) = \sum_{q=1}^{\infty} \mathbb{P}(N = mq + r \mid N > m) = \sum_{q=1}^{\infty} \mathbb{P}(N = mq + r) = p \cdot \frac{(1-p)^{m+r-1}}{1-(1-p)^m}$, by the memoryless property.

4 Introduction to Information Theory

(a) Since $p(\cdot)$ is the PMF of X, it can only take on values in the range $[0, 1]$. Under the assumption that the entropy is considered over the support of $p(\cdot)$, we can thus expect $p(\cdot)$ to take on only positive values in this range. In this case, each of the terms in the sum of the expectation will involve $\log p(x)$ scaled by $p(x)$. In the case that $p(x) < 1 \ \forall x$, the negative outside the summation can be brought into the logs to invert each $p(x)$, making them now all greater than 1 and thus positive when their logarithm is taken. When $p(x) = 1$ for some x, all other values of x are necessarily 0 and we only consider this x in our entropy calculation, resulting in $H(X) = 0$. Thus, we get that $H(X) \geq 0$.

(b) We should expect $H(X)$ to be greater when $p = \frac{1}{2}$, as this results in a higher variance in X, which should increase our expected "surprise". For $p = \frac{1}{3}$, we have that $H(X) = -\left(\frac{1}{3} \cdot \log_2 \frac{1}{3} + \frac{2}{3} \cdot \log_2 \frac{2}{3}\right) \approx .9183$. For $p = \frac{1}{2}$, we have that $H(X) = -\left(\log_2 \frac{1}{2}\right) = 1$, which agrees with our reasoning.

(c) $H(Y) = -\left((1-p_e) \cdot \log_2 \frac{1-p_e}{2} + p_e \cdot \log_2 p_e\right)$

(d) $H(X, Y) = -\left((1-p_e) \cdot \log_2 \frac{1-p_e}{2} + p_e \cdot \log_2 \frac{p_e}{2}\right)$

5 Soliton Distribution

(a) Let X be a random variable indicating the number of packets of degree d that are reduced 1 degree after the $(k+1)$st chunk is peeled off. Then $X = X_1 + X_2 + \cdots + X_{f_k(d)}$, where each $X_i \mid i \in [1, f_k(d)]$ is an indicator variable indicating whether the ith packet of degree d is reduced 1 degree. We are looking for $\mathbb{E}[X] = \sum_{i=1}^{f_k(d)} \mathbb{P}(X_i = 1)$, by linearity of expectation. Since each of the X_i packets consist of d chunks and $n-k$ chunks are left after k have been peeled off, we have that $\mathbb{P}(X_i = 1) = \frac{d}{n-k}$, and so we get that $\mathbb{E}[X] = f_k(d) \cdot \frac{d}{n-k}$.

(b) The recurrence relation we get is $f_{k+1}(d) = f_k(d) - f_k(d) \cdot \frac{d}{n-k} + f_k(d+1) \cdot \frac{d+1}{n-k} = f_k(d)\left(1 - \frac{d}{n-k}\right) + f_k(d+1)\left(\frac{d+1}{n-k}\right)$.

Base Case: Given a base case $d = 2$, we have that $f_{k+1}(1) = f_k(2)\left(\frac{2}{n-k}\right) = 1$, so that $f_k(2) = \frac{n-k}{2} = \frac{n-k}{2(2-1)}$.

Inductive Hypothesis: Assume $f_k(d) = \frac{n-k}{d(d-1)}$ for some $d > 2$.

Inductive Step: Prove $f_k(d+1) = \frac{n-k}{(d+1)d}$. From our inductive hypothesis, we have that $f_k(d) = \frac{n-k}{d(d-1)}$ and $f_{k+1}(d) = \frac{n-(k+1)}{d(d-1)}$. Plugging these values into our recurrence relation above and rearranging the terms, we get

$$\begin{aligned}
f_k(d+1) &= \frac{n-k}{d+1}\left(f_{k+1}(d) - f_k(d)\left(1 - \frac{d}{n-k}\right)\right) \\
&= \frac{n-k}{d+1}\left(\frac{n-(k+1)}{d(d-1)} - \frac{n-k}{d(d-1)}\left(1 - \frac{d}{n-k}\right)\right) \\
&= \frac{n-k}{(d+1)(d)(d-1)}\left(n-k-1-(n-k-d)\right) \\
&= \frac{n-k}{(d+1)d} \quad \square
\end{aligned}$$

From this, we have that $f_0(d) = \frac{n}{d(d-1)}$, which is the expected number of degree d packets received. This suggests $p(d) = \frac{1}{d(d-1)}$ for $d \neq 1$, and $p(d) = \frac{1}{n}$ for $d = 1$.

(c) $\mathbb{E}[p(d)] = \sum_{d=1}^{n} d \cdot p(d) = \frac{1}{n} + \sum_{d=2}^{n} \frac{1}{d-1} = \frac{1}{n} + \sum_{k=1}^{n-1} \frac{1}{k}$. As $n \to \infty$, we get $\approx \ln n$.

6 [Bonus] Connected Random Graph