Probability and Statistics, Spring 2009
Assignment 2 Solution

Q1

(a) $P_x(x) = p(1-p)^{x-1}, x = 1, 2, 3, ...$

$$E[X] = \sum_{x=1}^{\infty} xp(1-p)^{x-1} = \frac{1}{p}$$

pf: Let

$$S = p + 2p(1-p)^1 + 3p(1-p)^2 + ...$$
$$(1-p)S = p(1-p) \ + 2p(1-p)^2 + ...$$

Subtracting this two equations, we have $pS = \frac{p}{1-(1-p)} = 1, \ S = \frac{1}{p}$.

$$E[X^2] = \sum_{x=1}^{\infty} x^2 p(1-p)^{x-1} = \frac{2-p}{p^2}$$

pf: Let

$$H = p + 4p(1-p) + 9p(1-p)^2 + ...$$
$$(1-p)H = \ p(1-p) + 4p(1-p)^2 + ...$$

so

$$pH \ = p + 3p(1-p) + 5p(1-p)^2 + ...$$
$$(1-p)pH \ = p(1-p) + 3p(1-p)^2 + ...$$

then $ppH = p + 2p(1-p) + 2p(1-p)^2 + ... = p + \frac{2p(1-p)}{1-(1-p)}$ finally we got $H = \frac{2-p}{p^2}$. So, $Var[X] = E[X^2] - E[X]^2 = \frac{1-p}{p^2}$.

(b) Given $P_x(x) = C_x^n p^x (1-p)^{n-x}$ and $\sum_{x=0}^{n} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} = 1$.

$$\begin{aligned}
E[X] &= \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} ......(x = 0 \text{ can be neglected}) \\
&= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{n-x} = np \cdot 1 = np.
\end{aligned}$$

$$
\begin{aligned}
E[X^2] &= \sum_{x=0}^{n} x^2 \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} = \sum_{x=1}^{n} x^2 \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} (1+x-1) \frac{(n-1)!n}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} \frac{(n-1)!n}{(x-1)!(n-x)!} p^x (1-p)^{n-x} + \sum_{x=1}^{n} (x-1) \frac{(n-1)!n}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{n-x} + \sum_{x=2}^{n} (x-1) \frac{(n-1)!n}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= np + \sum_{x=2}^{n} \frac{(n-2)!n(n-1)}{(x-2)!(n-x)!} p^x (1-p)^{n-x} \\
&= np + p^2 n(n-1) \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x} \\
&= np + p^2 n(n-1) \cdot 1.
\end{aligned}
$$

So, $Var[X] = E[X^2] - E[X]^2 = np(1-p)$.

(c) Given$P_x(x) = C_{k-1}^{x-1} p^k (1-p)^{x-k}$ and $\sum_{x=k}^{\infty} \frac{(x-1)!}{(k-1)!(x-k)!} p^k (1-p)^{x-k} = 1$

$$
E[X] = \sum_{x=k}^{\infty} x \frac{(x-1)!}{(k-1)!(x-k)!} p^k (1-p)^{x-k} = \frac{k}{p} \sum_{x=k}^{\infty} \frac{x!}{k!(x-k)!} p^{k+1} (1-p)^{x-k} = \frac{k}{p}.
$$

$$
\begin{aligned}
E[X^2] &= \sum_{x=k}^{\infty} x^2 \frac{(x-1)!}{(k-1)!(x-k)!} p^k (1-p)^{x-k} \\
&= \sum_{x=k}^{\infty} (x+1) \frac{x!}{(k-1)!(x-k)!} p^k (1-p)^{x-k} - \sum_{x=k}^{\infty} \frac{x!}{(k-1)!(x-k)!} p^k (1-p)^{x-k} \\
&= \sum_{x=k}^{\infty} \frac{(x+1)!}{(k-1)!(x-k)!} p^k (1-p)^{x-k} - \sum_{x=k}^{\infty} \frac{x!}{(k-1)!(x-k)!} p^k (1-p) (\text{the latter is } E[X]) \\
&= \frac{k(k+1)}{p^2} \sum_{x=k}^{\infty} \frac{(x+1)!}{(k+1)!(x-k)!} p^{k+2} (1-p)^{x-k} - \frac{k}{p} = \frac{k(k+1)}{p^2} \cdot 1 - \frac{k}{p}.
\end{aligned}
$$

So, $Var(X) = E[X^2] - E[X]^2 = \frac{k(1-p)}{p^2}$.

(d) Given $P_x(x) = \frac{\alpha^x e^{-\alpha}}{x!}, x = 0, 1, 2, ...$ and $\sum_{x=0}^{\infty} \frac{\alpha^x e^{-\alpha}}{x!} = 1$.

$$
E[X] = \sum_{x=0}^{\infty} x \frac{\alpha^x e^{-\alpha}}{x!} = \sum_{x=1}^{\infty} x \frac{\alpha^x e^{-\alpha}}{x!} = \alpha \sum_{x=1}^{\infty} \frac{\alpha^{x-1} e^{-\alpha}}{(x-1)!} = \alpha.
$$

$$
\begin{aligned}
E[X^2] &= \sum_{x=0}^{\infty} x^2 \frac{\alpha^x e^{-\alpha}}{x!} = \sum_{x=1}^{\infty} x^2 \frac{\alpha^x e^{-\alpha}}{x!} \\
&= \sum_{x=1}^{\infty} (1+x-1) \frac{\alpha^x e^{-\alpha}}{(x-1)!} = \sum_{x=1}^{\infty} \frac{\alpha^x e^{-\alpha}}{(x-1)!} + \sum_{x=1}^{\infty} (x-1) \frac{\alpha^x e^{-\alpha}}{(x-1)!} \\
&= E[X] + \sum_{x=2}^{\infty} (x-1) \frac{\alpha^x e^{-\alpha}}{(x-1)!} \\
&= E[X] + \alpha^2 \sum_{x=2}^{\infty} \frac{\alpha^{x-2} e^{-\alpha}}{(x-2)!} = \alpha + \alpha^2.
\end{aligned}
$$

So, $Var(X) = E[X^2] - E[X]^2 = \alpha$.

Q2

(a) The pmf of N is given by $p_N(N = n) = (1-p)^n p$.

(b)

$$
\begin{aligned}
E[N] &= \sum_{n=0}^{\infty} np(1-p)^n \\
&= p(\sum_{n=0}^{\infty} (n+1)(1-p)^n - \sum_{n=0}^{\infty} (1-p)^n) \\
&= p(-\frac{d}{dp}(\sum_{n=0}^{\infty} (1-p)^{n+1}) - \sum_{n=0}^{\infty} (1-p)^n) \\
&= p(-\frac{d}{dp}(\frac{1}{p} - 1) - \frac{1}{p}) \\
&= p(\frac{1}{p^2} - \frac{1}{p}) = \frac{1}{p} - 1 = 100 - 1 = 99
\end{aligned}
$$

(c) The appropriate p should satisfy the following inequality

$$
\sum_{n=0}^{999} p(1-p)^n = 1 - (1-p)^{1000} \leq 0.01.
$$

Thus, $1 - \sqrt[1000]{0.99} \geq p$.

Q3

(a) The pmf of N is $p(x) = p(1-p)^{(x-1)}$ for $x = 1, 2, \ldots$, so the conditional probability is

$$
\begin{aligned}
P[N &= k | N \leq m] = \frac{P[N \leq m, N = k]}{P[N \leq m]} \\
&= \frac{P[N = k]}{P[N \leq m]} = \frac{p(1-p)^{k-1}}{P[N \leq m]} \\
&= \frac{p(1-p)^{k-1}}{\sum_{n=1}^{m} p(1-p)^{(n-1)}} = \frac{p(1-p)^{k-1}}{1-(1-p)^m}
\end{aligned}
$$

when $k \le m$ and zero when $k > m$. Thus, the answer is $\begin{cases} \frac{p(1-p)^{k-1}}{1-(1-p)^m}, k \le m \\ 0, \quad \text{otherwise} \end{cases}$.

(b) The probability is $\sum_{i=0}^{\infty} p(1-p)^{2i+1} = \frac{1-p}{2-p}$.

Q4

$p(k) = C_k^{12} \left(\frac{1}{12}\right)^k \left(\frac{11}{12}\right)^{12-k}$. By matlab, we got $p(0) = 0.35, p(1) = 0.38, p(2) = 0.19, p(3) = 0.058$, $p(4) = 0.012, ...$, and found that accumulative pmf from 0 to 4 is 0.99, ie, $\sum_{k=5}^{12} p(k)$ is less than 1%. So, once the manufacturer charges no less than \$130 , which is equal to \$50 + 4*\$20(4 times reparations), he will lose money with probability 1% even less. Average cost per player is 50+[expectation of reparation fee] expectation of reparation fee = $\sum_{k=0}^{12} 20kC_k^{12} \left(\frac{1}{12}\right)^k \left(\frac{11}{12}\right)^{12-k} = 20 \cdot 12 \cdot \frac{1}{12} = 20$, so average cost is \$70 per player.

Q5

(a) $(1-0.001)^{10000}$.

(b) If the broken ones is not replaced, the probability that a disk does not fail in two day is $(1-0.001)^2 = 0.999^2$. Thus, the probability that there are fewer than 10 failures in two days is given by

$$\sum_{k=0}^{9} \binom{10000}{k} (0.999^2)^{10000-k}(1-0.999^2)^k.$$

If the broken disk drives in a day are replaced, the equation becomes

$$\sum_{k=0}^{9} \binom{20000}{k} (0.999)^{20000-k}(1-0.999)^k.$$

(c) Assume that we require K spare disk drives. Then, K should satisfy

$$\sum_{k=0}^{K} \binom{10000}{k} (0.999)^{10000-k}(0.001)^k \ge 0.99.$$

Q6

(a) mean $= \frac{1}{2}$, variance $= \frac{21}{4}$

(b) mean $= -8$, variance $= 105$

(c) mean $= 0.6285$, variance $= 0.1051$

(d) mean $= 0.5$, variance $= 0.125$

Q7

(a) According to the problem statement,

p(detected defective | actually defective) $= a$

p(actually defective) $= p$

p(detected nondefective | actually nondefective) $= 1$

p(detected defective) $=$ p(detected defective | actually defective)p(actually defective)+ p(detected defective | actually nondefective)p(actually nondefective) $= ap + 0 \cdot (1-p) = ap$.

so the result is $(1-ap)^k ap$.

(b) p(actually defective | detected nondefective) $=$ p(actually defective and detected nondefective)/ p(detected nondefective) $= \frac{p(1-a)}{1-pa}$.

(c) Given p(detected defective | actually nondefective) $= b$.

$p(\text{detected nondefective}) = p(1-a) + (1-p)(1-b)$

the answer is $p(\text{actually defective} \mid \text{detected nondefective}) = \frac{p(1-a)}{p(1-a)+(1-p)(1-b)}$.

Q8

(a)Assume the observation interval is τ

$$\begin{aligned} P[\text{signal present} \mid X &= k] = \frac{P[\text{signal present, } X = k]}{P[X = k]} \\ &= \frac{P[X = k \mid \text{signal present}]P[\text{signal present}]}{P[X = k]} \\ &= \frac{\frac{(\tau\lambda_1)^k e^{-\tau\lambda_1}}{k!} \cdot p}{\frac{(\tau\lambda_1)^k e^{-\tau\lambda_1}}{k!} \cdot p + \frac{(\tau\lambda_0)^k e^{-\tau\lambda_0}}{k!} \cdot (1-p)} \end{aligned}$$

$$\begin{aligned} P[\text{signal absent} \mid X &= k] = 1 - P[\text{signal present} \mid X = k] \\ &= \frac{\frac{(\tau\lambda_0)^k e^{-\tau\lambda_0}}{k!} \cdot (1-p)}{\frac{(\tau\lambda_1)^k e^{-\tau\lambda_1}}{k!} \cdot p + \frac{(\tau\lambda_0)^k e^{-\tau\lambda_0}}{k!} \cdot (1-p)} \end{aligned}$$

(b) The decision rule is equivalent to

$$\frac{(\tau\lambda_1)^k e^{-\tau\lambda_1}}{k!} \cdot p \underset{absent}{\overset{present}{\gtrless}} \frac{(\tau\lambda_0)^k e^{-\tau\lambda_0}}{k!} \cdot (1-p)$$

and thus can be simplify to

$$k \underset{absent}{\overset{present}{\gtrless}} \frac{\ln\frac{1-p}{p} + (\lambda_1 - \lambda_0)\tau}{\ln(\frac{\lambda_1}{\lambda_0})} \triangleq T.$$

(c) The probability of error is given by

$$\begin{aligned} P_e &= P[\text{signal present}]P[k \le T \mid \text{signal present}] + P[k > T \mid \text{signal absent}]P[\text{signal absent}] \\ &= p\sum_{k=0}^{T} \frac{(\tau\lambda_1)^k e^{-\tau\lambda_1}}{k!} + (1-p)\sum_{k=T+1}^{\infty} \frac{(\tau\lambda_0)^k e^{-\tau\lambda_0}}{k!}. \end{aligned}$$

Q9

(a) The numbers that will have k zeros is $n = kM \sim kM + (M-1)$. The probability is the summation

$$\sum_{i=0}^{M-1} p^i (\frac{1}{2})^k (1-p) = (\frac{1}{2})^k (1-p^M) = (\frac{1}{2})^{k+1}.$$

(b) The avarage codeword length is

$$\sum_{k=0}^{\infty} (k+1+m)(\frac{1}{2})^{k+1} = \sum_{k=0}^{\infty} (k+1)(\frac{1}{2})^{k+1} + m\sum_{k=0}^{\infty} (\frac{1}{2})^{k+1} = m+2.$$

(c) The compression ratio is $\frac{\frac{p}{1-p}}{m+2}$.