Massachusetts Institute of Technology

Department of Electrical Engineering & Computer Science

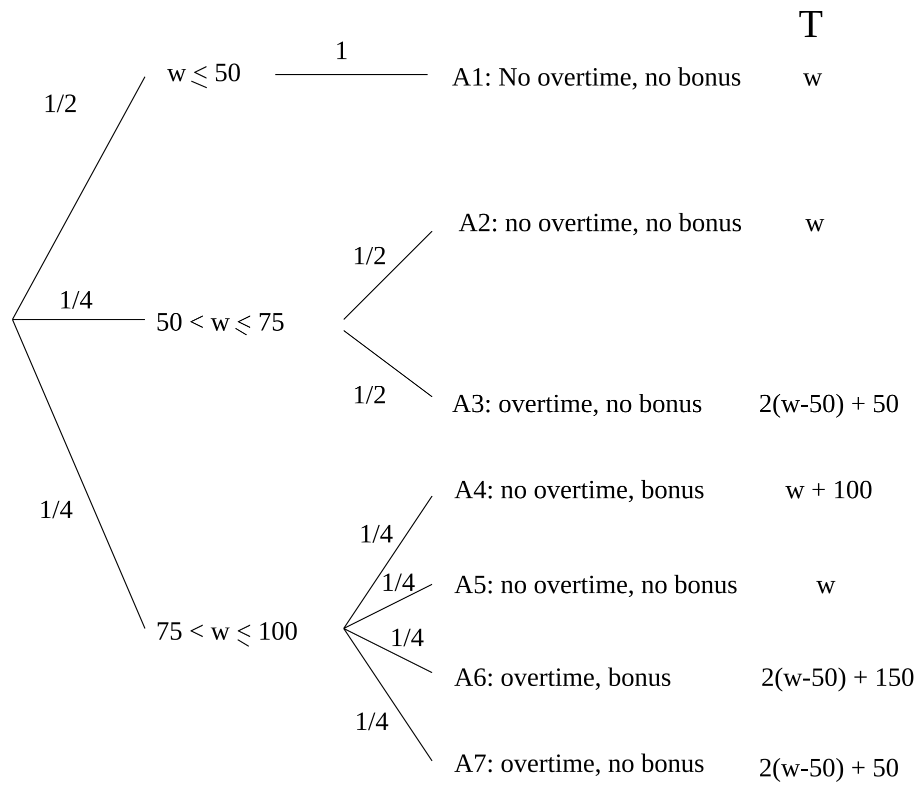
6.041/6.431: Probabilistic Systems Analysis

(Spring 2005)

Problem Set 7 Solutions

Due: April 6, 2005

1. Define the following events and RVs:
   W = number of hours Oscar works in a week,
   T = total amount of Oscar's earnings in a week (including overtime and bonus).
   Now, We want to find $\mathbf{E}[T]$ and $\text{Var}(T)$.



From the tree diagram, we use total expectation theorem,

$$\mathbf{E}[T] = \sum_{i=1}^{7} \mathbf{P}(A_i)\mathbf{E}[T|A_i]$$

Note that $\{A_1, A_2, ..., A_7\}$ is mutually exclusive and collectively exhaustive.

For each A_i, the conditional PDF $f_{T|A_i}(t)$ is constant because any linear function $aX + b$ of a uniformly distributed RV X is also uniformly distributed. Therefore,

$$\begin{aligned}
f_{T|A_1}(t) &= \tfrac{1}{50} && \text{for } 0 \le t \le 50 \\
f_{T|A_2}(t) &= \tfrac{1}{25} && \text{for } 50 < t \le 75 \\
f_{T|A_3}(t) &= \tfrac{1}{50} && \text{for } 50 < t \le 100 \\
f_{T|A_4}(t) &= \tfrac{1}{25} && \text{for } 175 < t \le 200 \\
f_{T|A_5}(t) &= \tfrac{1}{25} && \text{for } 75 < t \le 100 \\
f_{T|A_6}(t) &= \tfrac{1}{50} && \text{for } 200 < t \le 250 \\
f_{T|A_7}(t) &= \tfrac{1}{50} && \text{for } 100 < t \le 150
\end{aligned}$$

and

$$\begin{array}{rclrcl} \mathbf{E}[T|A_1] &=& 25 & \mathbf{E}[T|A_2] &=& \frac{125}{2} \\ \mathbf{E}[T|A_3] &=& 75 & \mathbf{E}[T|A_4] &=& \frac{375}{2} \\ \mathbf{E}[T|A_5] &=& \frac{175}{2} & \mathbf{E}[T|A_6] &=& 225 \\ \mathbf{E}[T|A_7] &=& 125. & & & \end{array}$$

Using the total expectation theorem, the expected salary per week is then equal to

$$\mathbf{E}[T] = \frac{1}{2}\cdot 25 + \frac{1}{8}\cdot\frac{125}{2} + \frac{1}{8}\cdot 75 + \frac{1}{16}\cdot\frac{375}{2} + \frac{1}{16}\cdot\frac{175}{2} + \frac{1}{16}\cdot 225 + \frac{1}{16}\cdot 125 = \boxed{68.75.}$$

For the variance of T, we need to first find $\mathbf{E}[T^2]$.

$$\mathbf{E}[T^2] = \sum_{i=1}^{7}\mathbf{P}(A_i)\mathbf{E}[T^2|A_i].$$

Using the fact that $\mathbf{E}[X^2] = (a^2+ab+b^2)/3 = ((a+b)^2 - ab)/3$ for any uniformly distributed RV X ranging from a to b, we obtain

$$\begin{array}{rclrcl} \mathbf{E}[T^2|A_1] &=& 50^2/3 & \mathbf{E}[T^2|A_2] &=& (125^2 - 50\cdot 75)/3 \\ \mathbf{E}[T^2|A_3] &=& (150^2 - 50\cdot 100)/3 & \mathbf{E}[T^2|A_4] &=& (375^2 - 175\cdot 200)/3 \\ \mathbf{E}[T^2|A_5] &=& (175^2 - 75\cdot 100)/3 & \mathbf{E}[T^2|A_6] &=& (450^2 - 200\cdot 250)/3 \\ \mathbf{E}[T^2|A_7] &=& (250^2 - 100\cdot 150)/3. & & & \end{array}$$

Therefore,

$$\mathbf{E}[T^2] = \frac{1}{2}\frac{2500}{3} + \frac{1}{8}\frac{11875}{3} + \frac{1}{8}\frac{17500}{3} + \frac{1}{16}\frac{105625}{3} + \frac{1}{16}\frac{23125}{3} + \frac{1}{16}\frac{152500}{3} + \frac{1}{16}\frac{47500}{3} = \frac{101875}{12}.$$

$$\mathrm{Var}(T) = \mathbf{E}[T^2] - (\mathbf{E}[T])^2 = \frac{180625}{48} \approx \boxed{3763.}$$

2. (a) The transform $M_J(s)$ given is a transform of a binomial random variable with parameters $n = 10$ and $p = \frac{2}{3}$. Thus the PMF for J is:

$$p_J(j) = \binom{n}{j}(\frac{1}{3})^{n-j}(\frac{2}{3})^j \quad \text{for } j = 0, 1, 2, ...10$$

(b) Again by inspection, K is a geometric random variable shifted to the right by 3 with parameter $p = \frac{1}{5}$. This is because we can rewrite $M_K(s) = e^{3s}\frac{\frac{1}{5}e^s}{1-\frac{4}{5}e^s}$. Thus,

$$p_K(k) = (\frac{4}{5})^{k-4}\frac{1}{5} \quad \text{for } k = 4, 5, 6, \ ...$$

$$\mathbf{E}[K] = 3 + \frac{1}{p} = 3 + 5 = 8$$

$$\mathrm{Var}(K) = \frac{1-p}{p^2} = \frac{\frac{4}{5}}{\frac{1}{25}} = 20$$

(c) Note that $L = K_1 + K_2 + ...K_J$, thus L is a random sum of random variables. So, determining the transform of L is easier than determining the PMF for L.

$$M_L(s) = M_J(s) \mid_{e^s = M_K(s)} = (\frac{1}{3} + \frac{2}{3}(\frac{\frac{1}{5}e^{4s}}{1 - \frac{4}{5}e^s}))^{10}$$

The expectation of L is $\mathbf{E}[L] = \mathbf{E}[K]\mathbf{E}[J] = 8 * \frac{20}{3} = \frac{160}{3}$
The variance of L is

$$\text{Var}(L) = \text{Var}(K)\mathbf{E}[J] + \text{Var}(J)(\mathbf{E}[K])^2 = (20)(10 * \frac{2}{3}) + (10 * \frac{2}{3} * \frac{1}{3})(64) = \frac{2480}{9}$$

(d) $\mathbf{P}(\text{person donates}) = \frac{1}{4}$. Let M = total # of donors from all living groups, and define

$$X_i = \begin{cases} 1 & \text{if ith person donates} \\ 0 & \text{otherwise.} \end{cases}$$

The PMF for X is just

$$p_X(x) = \begin{cases} \frac{1}{4} & \text{if } x = 1 \\ \frac{3}{4} & \text{if } x = 0. \end{cases}$$

Then,

$$M = X_1 + X_2 + ...X_L.$$

Therefore the transform of M is:

$$M_M(s) = M_L(s) \mid_{e^s = M_X(s)}$$

The transform of X is (by inspection)

$$M_X(s) = (\frac{3}{4} + \frac{1}{4}e^s)$$

Therefore,

$$M_M(s) = (\frac{1}{3} + \frac{2}{3}(\frac{\frac{1}{5}(\frac{3}{4} + \frac{1}{4}e^s)^4}{1 - \frac{4}{5}(\frac{3}{4} + \frac{1}{4}e^s)}))^{10}$$

To obtain $\mathbf{P}(M = 0)$, we simply evaluate the transform of M at $e^s = 0$.

$$p_M(0) = M_M(s) \mid_{e^s = 0} = (\frac{1}{3} + \frac{2}{3}(\frac{\frac{1}{5}(\frac{3}{4})^4}{1 - \frac{4}{5}(\frac{3}{4})}))^{10}.$$

The expectation of M is $\mathbf{E}[M] = \mathbf{E}[X]\mathbf{E}[L] = \frac{40}{3}$
The variance of M is

$$\text{Var}(M) = \text{Var}(X)\mathbf{E}[L] + \text{Var}(L)(\mathbf{E}[X])^2 = 27.22$$

3. We know that:

$$\rho(X_1, X_2) = \frac{\mathrm{Cov}(X_1, X_2)}{\sigma_{X_1}\sigma_{X_2}}$$

Therefore we first find the covariance:

$$\begin{aligned} \mathrm{Cov}(A, B) &= \mathbf{E}[AB] - \mathbf{E}[A]\mathbf{E}[B] \\ &= \mathbf{E}[WX + WY + X^2 + XY] \\ &= \mathbf{E}[X^2] = 1 \end{aligned}$$

and

$$\begin{aligned} \sigma_A &= \sqrt{\mathrm{Var}(A)} = \sqrt{2} \\ \sigma_B &= \sqrt{\mathrm{Var}(B)} = \sqrt{2} \end{aligned}$$

and therefore:

$$\rho(A, B) = \frac{1}{2}.$$

We proceed as above to find the correlation of A, C.

$$\begin{aligned} \mathrm{Cov}(A, C) &= \mathbf{E}[AC] - \mathbf{E}[A]\mathbf{E}[C] \\ &= \mathbf{E}[WY + WZ + XY + XZ] \\ &= 0 \end{aligned}$$

and therefore

$$\rho(A, C) = 0.$$

4. (a) First note that $\hat{X}$ should be a r.v., not a number. In particular, we are to minimize over all r.v.'s $\hat{X}$ that can be expressed as functions of Y. From lecture, $\hat{X} = \mathbf{E}[X|Y]$.

   Now, take conditional expectations, to get $Y = \mathbf{E}[Y|Y] = \mathbf{E}[X|Y] + \mathbf{E}[W|Y]$. Since there is complete symmetry between X and W, we also have $\mathbf{E}[X|Y] = \mathbf{E}[W|Y]$, which finally yields $\mathbf{E}[X|Y] = Y/2$.

   (b) In the dependent case, we cannot simply conclude that the distribution $f_{X,W}(x, w)$ is symmetric in its two argument (i.e., $f_{X,W}(x, w) = f_{X,W}(w, x)$), even though the marginals $f_X(x), f_W(w)$ are the same.
   Since $f_{X,W}(x, w)$ is not symmetric, $\mathbf{E}[X|Y] \neq \mathbf{E}[W|Y]$ in general.
   So in this case, one cannot really solve the problem with the available information, we really need the joint distribution in order to compute the conditional expectations.
   The solution given in the independent case still works, though, for any symmetric distribution.

5. (a) Let A be the event that the sender transmitted, and K be the number of photons counted by the photodetector. Using Bayes rule,

$$\mathbf{P}(A \mid K = k) = \frac{p_{K|A}(k)\mathbf{P}(A)}{p_K(k)} = \frac{p_{X+N}(k) \cdot p}{p_N(k) \cdot (1-p) + p_{X+N}(k) \cdot p}$$

The discrete random variables X and N are given by the following PMF's:

$$p_X(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x \geq 0$$

$$p_N(n) = \frac{\mu^n e^{-\mu}}{n!}, \quad n \geq 0$$

The sum of two independent Poisson random variables is also Poisson, with mean equal to the sum of the means of each of the random variables. This fact can be derived by looking at the product of the transforms of X and N. Therefore:

$$p_{X+N}(k) = \frac{(\lambda+\mu)^k e^{-(\lambda+\mu)}}{k!}, \quad k \geq 0$$

Thus,

$$\mathbf{P}(A \mid K = k) = \frac{p \cdot \frac{(\lambda+\mu)^k e^{-(\lambda+\mu)}}{k!}}{p \cdot \frac{(\lambda+\mu)^k e^{-(\lambda+\mu)}}{k!} + (1-p) \cdot \frac{\mu^k e^{-\mu}}{k!}} = \frac{1}{1 + \frac{1-p}{p}\left(\frac{\mu}{\lambda+\mu}\right)^k e^{\lambda}}$$

(b) Let S be the number of photons transmitted by the sender. Then with probability p, $S = X$, and with probability $1-p$, $S = 0$. The least squares estimate of the number of photons transmitted by the sender is simply the mean, in the absence of any additional information:

$$\hat{S}_1 = \mathbf{E}[S] = p \cdot \lambda + (1-p) \cdot 0 = p\lambda.$$

(c) The least squares predictor has a form

$$\hat{S}_2(k) = \mathbf{E}[S \mid K = k].$$

Using Bayes rule,

$$\hat{S}_2(k) = \sum_{s=0}^{k} s p_{S|K}(s|k) = \sum_{s=0}^{k} s \frac{p_{K|S}(k|s) p_S(s)}{p_K(k)} = \frac{1}{p_K(k)} \sum_{s=0}^{k} s p_{K|S}(k|s) p_S(s).$$

From the definitions of S and K, the following are true:

$$\begin{aligned}
p_S(s) &= \begin{cases} (1-p) + pe^{-\lambda}, & s = 0 \\ p\frac{\lambda^s e^{-\lambda}}{s!}, & s = 1, 2, \ldots \end{cases} \\
p_{K|S}(k|s) &= p_N(k-s) = \frac{\mu^{(k-s)} e^{-\mu}}{(k-s)!} \\
p_K(k) &= p\frac{(\lambda+\mu)^k e^{-(\lambda+\mu)}}{k!} + (1-p)\frac{\mu^k e^{-\mu}}{k!}
\end{aligned}$$

In order to obtain the last expression, we observe that $K = S + N$ with probability p, and $K = N$ with probability $1-p$. Substituting into the formula above,

$$\hat{S}_2(k) = \frac{1}{p_K(k)} \left[0 \cdot (1-p) \frac{\mu^{(k-0)} e^{-\mu}}{(k-0)!} + \sum_{s=0}^{k} sp \frac{\lambda^s e^{-\lambda}}{s!} \frac{\mu^{(k-s)} e^{-\mu}}{(k-s)!} \right]$$

$$
\begin{aligned}
&= \frac{1}{p_K(k)} p e^{-\lambda} e^{-\mu} \frac{(\lambda+\mu)^k}{k!} \sum_{s=0}^{k} s \frac{k!}{s!(k-s)!} \left(\frac{\lambda}{\lambda+\mu}\right)^s \left(\frac{\mu}{\lambda+\mu}\right)^{(k-s)} \\
&= \frac{1}{p_K(k)} p e^{-\lambda} e^{-\mu} \frac{(\lambda+\mu)^k}{k!} k \left(\frac{\lambda}{\lambda+\mu}\right) = k \left(\frac{\lambda}{\lambda+\mu}\right) \frac{p e^{-(\lambda+\mu)} (\lambda+\mu)^k}{k! p_K(k)} \\
&= k \left(\frac{\lambda}{\lambda+\mu}\right) \frac{1}{1 + \frac{1-p}{p} \left(\frac{\mu}{\lambda+\mu}\right)^k e^{\lambda}}.
\end{aligned}
$$

Thus

$$\hat{S}_2(k) = \frac{k\lambda}{\lambda+\mu} \frac{1}{1 + \frac{1-p}{p} \left(\frac{\mu}{\lambda+\mu}\right)^k e^{\lambda}}.$$

Note that as k increases, the estimator can be approximated by $\frac{k\lambda}{\lambda+\mu}$.

(d) The linear least squares predictor has the form

$$\hat{S}_3(k) = \mathbf{E}[S] + \frac{\text{Cov}(S,K)}{\sigma_K^2}(k - \mathbf{E}[K]) \tag{1}$$

Note that since X and N are independent, S and N are also independent.

$$
\begin{aligned}
\mathbf{E}[S] &= p\lambda \\
\mathbf{E}[S^2] &= p\mathbf{E}[X^2] + (1-p)(0) = p(\lambda^2 + \lambda) \\
\sigma_S^2 &= \mathbf{E}[S^2] - (\mathbf{E}[S])^2 = p(\lambda^2+\lambda) - (p\lambda)^2 = p(1-p)\lambda^2 + p\lambda. \\
\Rightarrow \quad \mathbf{E}[K] &= \mathbf{E}[S] + \mathbf{E}[N] = p\lambda + \mu \\
\sigma_K^2 &= \sigma_S^2 + \sigma_N^2 \\
&= p(1-p)\lambda^2 + p\lambda + \mu.
\end{aligned}
$$

Finally, we need to find $\text{Cov}(S,K)$.

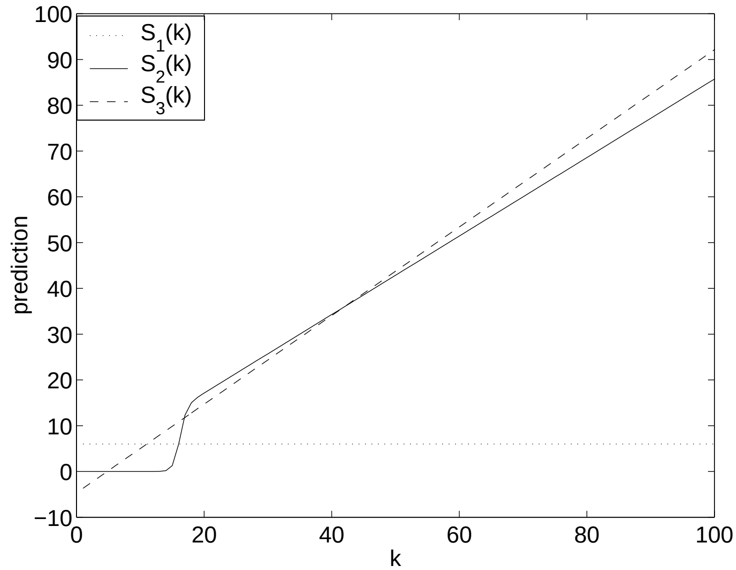
$$
\begin{aligned}
\text{Cov}(S,K) &= \mathbf{E}[(S - \mathbf{E}[S])(K - \mathbf{E}[K])] \\
&= \mathbf{E}[(S - \mathbf{E}[S])(S - \mathbf{E}[S] + N - \mathbf{E}[N])] \\
&= \mathbf{E}[(S - \mathbf{E}[S])(S - \mathbf{E}[S])] + \mathbf{E}[(S - \mathbf{E}[S])(N - \mathbf{E}[N])] \\
&= \sigma_S^2 + \mathbf{E}[(S - \mathbf{E}[S])(N - \mathbf{E}[N])] \\
&= \sigma_S^2 \\
&= p(1-p)\lambda^2 + p\lambda.
\end{aligned}
$$

Note that we have used the fact that $(S - \mathbf{E}[S])$ and $(N - \mathbf{E}[N])$ are independent, and $\mathbf{E}[(S - \mathbf{E}[S])] = 0 = \mathbf{E}[(N - \mathbf{E}[N])]$.

Therefore, substituting all the numbers into the equation above, we get the linear predictor:

$$\hat{S}_3(k) = p\lambda + \frac{p(1-p)\lambda^2 + p\lambda}{p(1-p)\lambda^2 + p\lambda + \mu} (k - p\lambda - \mu).$$

The graph bellow illustrates the three estimators for a particular setting of the parameters $\lambda = 30, \mu = 5, p = 0.2$:


S1(k)
S2(k)
S3(k)
prediction
k
100
90
80
70
60
50
40
30
20
10
0
−10
0
20
40
60
80
100