EE 126 Fall 2006 Solutions to midterm #2
Thursday, November 16: 3:30–5pm

DO NOT TURN THIS PAGE OVER UNTIL YOU ARE TOLD TO DO SO

- You have 90 minutes to complete the quiz.
- Write your solutions in the exam booklet. We will not consider any work not in the exam booklet.
- This quiz has three problems that are in no particular order of difficulty.
- The following page in the question booklet has some formulae that may be useful in your solutions.
- A correct answer does not guarantee full credit and a wrong answer does not guarantee loss of credit. You should concisely indicate your reasoning and show all relevant work. The grade on each problem is based on our judgment of your level of understanding as reflected by what you have written.
- This is a closed-book exam except for one double-sided, handwritten, 8.5×11 formula sheet plus a calculator.
- Be neat! If we can't read it, we can't grade it.

Problem	Score
1 [11 points]	
2 [13 points]	
3 [16 points]	
Total	

Some useful formulae

(a) For $|a| < 1$, $\sum_{k=0}^{\infty} a^k = 1/(1-a)$.

(b) For any $n = 1, 2, 3, \ldots$, $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.

(c) For any $n = 1, 2, 3, \ldots$, $\sum_{k=1}^{n} k^2 = \frac{2n^3+3n^2+n}{6}$.

(d) For any γ and $0 \leq a < b < +\infty$, we have

$$\int_a^b x \exp(\gamma x) dx = \frac{be^{\gamma b} - ae^{\gamma a}}{\gamma} - \frac{1}{\gamma^2}\left[e^{\gamma b} - e^{\gamma a}\right].$$

(e) For any $\mu \in \mathbb{R}$, we have

$$\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(x-\mu)^2\right) dx = \sqrt{2\pi}.$$

Problem 1: Each Acme Brand Chocolate egg contains an action figure, chosen uniformly from a set of four different action figures. The price of any given egg (in dollars) is an exponentially distributed random variable with parameter λ, and the prices of different eggs are independent. For $i = 1, \ldots, 4$, let T_i be a random variable corresponding to the number of eggs that you purchase in order to have i different action figures. E.g., after after purchasing T_3 eggs (and not before), you have at least one copy of exactly three different action figures.

(a) 2 pts What is the PMF and expected value of T_2?

(b) 3 pts Compute $\mathbb{E}[T_4]$ and $\text{var}(T_4)$ *without* using moment generating functions. (*Hint:* Find a representation of T_4 as a sum of independent RVs.)

(c) 3 pts Compute the moment generating function of T_4, and use it to verify your answer for $\mathbb{E}[T_4]$ in part (b).

(d) 3 pts You keep buying eggs until you have collected all four eggs. Compute the moment-generating function of the total amount of money that you spend.

Solutions:

(a) On the first trial, you certainly receive one egg. Thereafter, you are waiting for your second egg, and this waiting time is geometric with parameter 3/4. Thus, we have $T_2 = 1 + Z_2$ where $Z_2 \sim \text{Geo}(3/4)$, and

$$p_{T_2}(t) = \begin{cases} (1/4)^{t-2}\,(3/4) & \text{for } t = 2, 3, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, we have $\mathbb{E}[T_2] = 1 + \mathbb{E}[Z_2] = 1 + 1/(3/4) = 7/3$.

(b) Following the same idea as part (a), we can write

$$T_4 = Z_1 + Z_2 + Z_3 + Z_4$$

where $Z_1 \sim \text{Geo}(1)$, $Z_3 \sim \text{Geo}(1/2)$ and $Z_4 \sim \text{Geo}(1/4)$, and all of Z_i's are independent. Hence we have

$$\begin{aligned} \mathbb{E}[T_4] &= 1 + 1/(3/4) + 1/(1/2) + 1/(1/4) \ = 25/3 \\ \text{var}[T_4] &= 0 + (1-3/4)/(3/4)^2 + (1-1/2)/(1/2)^2 + (1-1/4)/(1/4)^2 \ = 130/9 \end{aligned}$$

(c) From independence of Z_i, we have that $M_{T_4}(s)$ is the product of 4 geometric transforms:

$$M_{T_4}(s) = e^s \frac{3e^s/4}{1-e^s/4} \frac{e^s/2}{1-e^s/2} \frac{e^s/4}{1-3e^s/4}$$

Now to find the expected value of T_4 we first take the derivative with respect to s, and then set $s = 0$. By product rule, we have

$$\frac{d}{ds} M_{T_4}(s) = \sum_{i=1}^{4} \frac{d}{ds} M_{Z_i}(s) \left[\prod_{j=1, j\neq i}^{4} M_{Z_i}(s) \right]$$

where each Z_i is geometric as in part (b). Since $M_{Z_i}(0) = 1$ for MGFs, and $\frac{d}{ds}M_{Z_i}(0) = \mathbb{E}[Z_i]$, we have

$$\begin{aligned}
\mathbb{E}[T_4] &= \sum_{i=1}^{4} \mathbb{E}[Z_i](1) \\
&= 1 + 1/(3/4) + 1/(1/2) + 1/(1/4) \ = 25/3,
\end{aligned}$$

as in (b).

(d) In this part, we have a sum of a random number T_4 of exponential RVs. If X is the total amount of money spent and Y_i the price of any particular egg i, the transform of X is obtain from the transform of T_4 by replacing e^s with the transform of Y_i. So we have:

$$\begin{aligned}
M_X(s) &= \left[\frac{\lambda}{\lambda - s}\right] \left[\frac{3\lambda/(4(\lambda - s))}{1 - \lambda/(4(\lambda - s))}\right] \left[\frac{\lambda/(2(\lambda - s))}{1 - \lambda/(2(\lambda - s))}\right] \left[\frac{\lambda/(4(\lambda - s))}{1 - 3\lambda/(4(\lambda - s))}\right] \\
&= \left[\frac{\lambda}{\lambda - s}\right] \left[\frac{3\lambda}{3\lambda - 4s}\right] \left[\frac{\lambda}{\lambda - 2s}\right] \left[\frac{\lambda}{\lambda - 4s}\right].
\end{aligned}$$

Problem 2: Xavier and Yolanda are competing in the long jump. Xavier jumps a distance X whereas Yolanda jumps a distance Y, where X and Y are independent exponential random variables with parameter λ (in meters).

(a) 3 pts The *team score* of Xavier and Yolanda on any given trial is measured by the sum $Z = X + Y$ of their distances. Compute the PDF $f_Z(z)$ and the moment generating function $M_Z(s)$, and evaluate $f_Z(1)$ and $M_Z(2)$ (which will be functions of λ).

(b) 3 pts Xavier and Yolanda win the competition if their team score on at least one of three independent trials exceeds 3 meters. What is the probability as a function of λ that Xavier and Yolanda win the competition? Evaluate your expression for $\lambda = 1$.

(c) 4 pts Suppose that you observe $Z = X + Y$, and are interested in predicting X. Compute the linear least squares estimate of X based on Z; simplify it to the form $f(z) = az + b$ for quantities a and b that can depend on λ.

(d) 3 pts Now suppose that $\lambda = 1$. Compute the PDF of $U = X/(X + Y)$, corresponding to the fraction that Xavier contributes to the team score.

Solutions:

(a) Since X and Y are independent, we obtain the solution by convolution

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)dx \\ &= \int_0^x \lambda^2 \exp(-\lambda(x + (z-x)))dx \\ &= \lambda^2 z \exp(-\lambda z) \end{aligned}$$

for $z \geq 0$, and 0 otherwise. By independence, the MGF is given by the product of the MGFs for X and Y—that is, $M_Z(s) = M_X(s)M_Y(s) = \lambda^2/(\lambda - s)^2$.

(b) The specified event is the complement of having a score less than or equal to 3 on three independent trials, so that the specified probability is equal to $1 - [\mathbb{P}(Z \leq 3)]^3$. We then compute

$$\begin{aligned} \mathbb{P}[Z \leq 3] &= \int_0^3 \lambda^2 z \exp(-\lambda z)dz \\ &= 1 - e^{-3\lambda} - 3\lambda e^{-3\lambda} \end{aligned}$$

(c) We compute $\mathbb{E}[X] = 1/\lambda$ and $\text{var}(X) = 1/\lambda^2$ by known properties of the exponential distribution. Moreover, we have $\mathbb{E}[Z] = 2/\lambda$ and $\text{var}(Z) = \text{var}(X) + \text{var}(Y) = 2/\lambda^2$ by independence. Lastly, we compute

$$\begin{aligned} \text{cov}(X, Z) = &= \text{cov}(X, X + Y) \\ &= \text{cov}(X, X) + \text{cov}(X, Y) \\ &= \text{var}(X) = 1/\lambda^2, \end{aligned}$$

using independence of X and Y. Hence, we have

$$\begin{aligned}
\widehat{X}(z) &= \mathbb{E}[X] + \frac{\operatorname{cov}(X,Z)}{\operatorname{var}(Z)}\,[z - \mathbb{E}[Z]] \\
&= \frac{1}{\lambda} + \frac{1/\lambda^2}{2/\lambda^2}\,[z - 2/\lambda] \\
&= \frac{z}{2}
\end{aligned}$$

(d) Let us compute the CDF F_U of U. By inspection, we have $F_U(u) = 0$ for all $u \leq 0$ and $F_U(u) = 1$ for all $u \geq 1$. Otherwise, for any $u \in (0,1)$, we compute

$$\begin{aligned}
\mathbb{P}[U \leq u] &= \mathbb{P}[\frac{X}{X+Y} \leq u] \\
&= \mathbb{P}[X \leq \frac{u}{1-u} Y].
\end{aligned}$$

Using the independence of X and Y and $\lambda = 1$, we have

$$\begin{aligned}
\mathbb{P}[X \leq \frac{u}{1-u} Y] &= \int_0^\infty \int_0^{\frac{u}{1-u}y} \exp(-x-y)\,dx\,dy \\
&= \int_0^\infty \exp(-y)\left[1 - \exp(-\frac{u}{1-u}y)\right] dy \\
&= 1 - \frac{1}{1+\frac{u}{1-u}} \\
&= 1 - \frac{(1-u)}{1-u+u} \\
&= u.
\end{aligned}$$

Therefore, the RV U is actually uniform on $[0,1]$.

Problem 3: Bob has gone hiking, and is lost in the forest. In order to try and find a road, he decides on the following *distance/coin-flip* strategy. At time instants $t = 1, 2, 3 \ldots$, he chooses a distance uniformly at random between t and $t+1$. Independently of the chosen random distance, he then flips a fair coin; if it comes up heads, he moves the chosen random distance to the right (positive on the real line), and otherwise for a tails toss, he moves the chosen random distance to the left (negative on the real line). Both the random distance and the coin flip are independent random variables for different time instants. Assume that he starts at the origin at time instant $t = 0$.

(a) 3 pts Let $Y_s \in \mathbb{R}$ be Bob's position after repeating his distance/coin-flip strategy for a fixed number of s time instants. Compute its expected value and variance as a function of s.
Hint: Write Y_s as a sum of independent random variables.

Now suppose that Bob repeats his distance/coin-flip strategy for a *random number* S of time rounds, after which he stops. Assume that $S \sim \text{Geo}(p)$ has a geometric distribution with parameter p, and let $X \in \mathbb{R}$ be his final position. For any question below, you may feel free to express your answer (if appropriate) in terms of the moments $\mu_i = \mathbb{E}[S^i], i = 1, 2, 3 \ldots$.

(b) 2 pts Suppose that you observe that $S = s$. What is the minimum mean squared error (MMSE) estimator of X given this information?

(c) 4 pts What is the expected value and variance of his position X?
Hint: Iterated expectations, and your answers to part (a) could be helpful here.

(d) 3 pts Now suppose that you observe that Bob finishes at position $X = x$. Given this information, what is the linear least squares estimator (LLSE) of the number of time rounds S that he repeated his distance/coin-flip strategy?

(e) 4 pts What is the LLSE of S based on X^2?

Solutions:

(a) Let X_t be the change in his position at round t; it is a random variable with PDF

$$f_{X_t}(x_t) = \begin{cases} \frac{1}{2} & \text{for } x_t \in [t, t+1] \cup [-t-1, -i] \\ 0 & \text{otherwise.} \end{cases}$$

Note that we have $\mathbb{E}[X_t] = 0$ and

$$\begin{aligned} \text{var}(X_t) &= \mathbb{E}[X_t^2] \\ &= \int_t^{t+1} x^2 dx \\ &= \frac{1}{3}\left[(t+1)^3 - t^3\right] \\ &= \frac{1}{3}\left[3t^2 + 3t + 1\right]. \end{aligned}$$

After s steps, his position is given by $Y_s = \sum_{t=1}^{s} X_t$, so that $\mathbb{E}[Y_s] = 0$ and (by independence)

$$
\begin{aligned}
\text{var}(Y_s) &= \sum_{t=1}^{s} \text{var}(X_t) \\
&= \sum_{t=1}^{s} t^2 + \sum_{t=1}^{s} t + \frac{1}{3}\sum_{t=1}^{s}(1) \\
&= \frac{2s^3 + 3s^2 + s}{6} + \frac{s(s+1)}{2} + \frac{s}{3} \\
&= \frac{2s^3 + 6s^2 + 6s}{6} \\
&= \frac{1}{3}s^3 + s^2 + s.
\end{aligned}
$$

where we have used some of the useful formulae given.

(b) Give that we observe $S = s$, then the best MMSE estimate of X is the conditional expection $\mathbb{E}[X \mid S = s] = \mathbb{E}[\sum_{i=1}^{s} X_i] = 0$.

(c) Now we have the random sum $X = \sum_{i=1}^{S} X_i$. By iterated expectation, we have

$$
\mathbb{E}[X] \;=\; \mathbb{E}\left[\mathbb{E}[\sum_{i=1}^{s} X_i \;\mid\; S = s]\right] \;=\; \mathbb{E}[0] \;=\; 0.
$$

Moreover, since $\mathbb{E}[X^2|S = s] = \mathbb{E}[Y_s^2] = \text{var}(Y_s)$, we can use iterated expectations and part (a) to write

$$
\begin{aligned}
\text{var}(X) &= \mathbb{E}[X^2] \;=\; \mathbb{E}\left\{\mathbb{E}\left[X^2 \;|S\right]\right\} \\
&= \mathbb{E}\left[\frac{1}{3}S^3 + S^2 + S\right] \\
&= \frac{\mu_3}{3} + \mu_2 + \mu_1.
\end{aligned}
$$

(d) Again using iterated expectations, we have

$$
\text{cov}(X, S) \;=\; \mathbb{E}[XS] \;=\; \mathbb{E}\left[\mathbb{E}[XS \;\mid\; S]\right] = \mathbb{E}\left[S\mathbb{E}[X \;\mid\; S]\right] \;=\; 0,
$$

since $\mathbb{E}[X \mid S] = 0$ from part (b). Hence the LLSE estimate of S given X is $\mathbb{E}[S] = \mu_1$ since S and X are uncorrelated.

(e) In this case, the best linear estimator of S based on X^2 is given by

$$
\mathbb{E}[S] + \frac{\text{cov}(X^2, S)}{\text{var}(X^2)}(X^2 - \mathbb{E}[X^2]).
$$

Note that we have $\mathbb{E}[S] = \mu_1$, $\mathrm{var}(S) = \mu_2 - \mu_1^2$, and $\mathbb{E}[X^2] = \mathrm{var}(X) = \frac{\mu_3}{3} + \mu_2 + \mu_1$ from part (c). Lastly, we compute

$$\begin{aligned}
\mathrm{cov}(X^2, S) &= \mathbb{E}[X^2 S] - \mathbb{E}[S]\mathbb{E}[X^2] \\
&= \mathbb{E}[\mathbb{E}[SX^2 \mid S]] - (1/p)(2/p^2 + 1/p) \\
&= \mathbb{E}\left[S\,\frac{1}{3}S^3 + S^2 + S\right] - \mu_1\,\mathbb{E}[X^2] \\
&= \frac{\mu_4}{2} + \mu_3 + \mu_2 - \mu_1\left[\frac{\mu_3}{3} + \mu_2 + \mu_1\right].
\end{aligned}$$