EE 126 Fall 2006 Final Exam
Tuesday, December 19: 12:30–3:30pm

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

- You have 3 hours to complete the exam.
- Write your solutions in the exam booklet. We will not consider any work not in the exam booklet.
- This quiz has five (5) problems that are in no particular order of difficulty.
- 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 two handwritten, 8.5×11 formula sheets plus a calculator.
- Be neat! If we can't read it, we can't grade it.

Some useful formulae

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

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

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

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

(d) **Gaussian normalization:** For any $\mu \in \mathbb{R}$, we have

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

(e) **Binomial formula:** For any $a, b \in \mathbb{R}$ and positive integer n, we have

$$(a+b)^n \quad = \quad \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Problem 1: (20 points) Consider a pair of continuous random variables (X, Y) with joint PDF uniform over the shaded region in Figure 1:

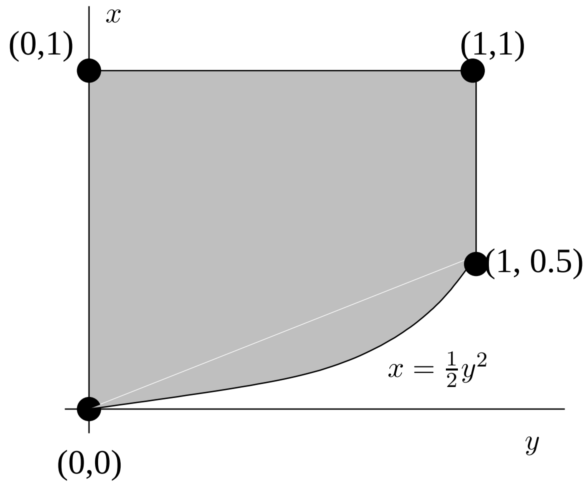

Figure 1: Joint PDF of random variables X and Y.

(a) (4 pt) Compute the conditional PDF $p_{X|Y}(x \mid y)$.

(b) (5 pt) Compute the linear least-squares estimator (LLSE) of X based on Y.

(c) (5 pt) Compute the minimum mean-squared error estimator (MMSE) of X based on Y.

(d) (6 pt) Now suppose that in addition to observing Y, you are also told that $X \geq \frac{1}{2}Y$. Compute the LLSE and MMSE estimators of X based on both Y and $\{X \geq \frac{1}{2}Y\}$.

Solution:

(a) We note $p_{X,Y}$ is uniform over the area where it is non-zero, so once Y is fixed, X will be uniform between $y^2/2$ and 1.

(b) First compute the value of the constant joint function in the region where it is nonzero:

$$\begin{aligned} p_{X,Y}^{-1} &= \int_0^1 (\int_{y^2/2}^1 dx) dy \\ &= 5/6 \end{aligned}$$

Now we compute expectations:

$$\begin{aligned}
\mathbb{E}[X] &= \frac{6}{5}\int_0^1 (\int_{y^2/2}^1 x dx) dy \\
&= 57/100 \\
E[Y] &= \frac{6}{5}\int_0^1 (\int_{y^2/2}^1 dx) y dy \\
&= 9/20 \\
E[XY] &= \frac{6}{5}\int_0^1 (\int_{y^2/2}^1 x dx) y dy \\
&= 11/40 \\
E[Y^2] &= \frac{6}{5}\int_0^1 (\int_{y^2/2}^1 dx) y^2 dy \\
&= 7/25
\end{aligned}$$

So $\text{var}(Y) = 31/400$. We have that the LLSE of X based on Y is

$$E[X] + cov(X,Y)/Var(Y)(Y - E[Y]) = 57/100 + 37/155(Y - 9/20)$$

(c) The MMSE of X based on Y is $E[X|Y = y] = 1/2 + y^2/4$, using the fact that X (conditioned on $Y = y$) is uniform on the interval 1 and $y^2/4$, derived in part (a).

(d) In this case the PDF of X based on $X \geq Y/2$ is uniform between $y/2$ and 1, so both the MMSE and the LLSE are equal to $1/2 + y/4$.

Problem 2: Rolling dice: (20 points)

You play a game by rolling six fair dice simultaneously. Each die has sides $\{1, 2, 3, 4, 5, 6\}$, and each of your rolls are independent. You win a pet frog if at least 2 of the dice shows the same number, and you win a pet elephant if there are at least 4 sixes. What is the probability of

(a) (4 pt) rolling exactly 2 sixes?

(b) (4 pt) winning a pet frog?

(c) (4 pt) winning a pet elephant?

(d) (3 pt) winning a pet elephant given that you have won a pet frog?

(e) (5 pt) winning a pet elephant given that the total score (on all your six rolls) is greater than or equal to 33?

Solution:

(a) $\frac{\binom{6}{2}\times 5^4}{6^6}$

(b) $1 - \frac{6!}{6^6}$

(c) $\frac{\binom{6}{4}\times 5^2+\binom{6}{5}\times 5+1}{6^6}$

(d) $\frac{\binom{6}{4}\times 5^2+\binom{6}{5}\times 5+1}{6^6-6!}$

(e) The configurations that the total score is greater than or equal to 33 are:
36: 6 sixes
35: 5 sixes, 1 five
34: 5 sixes, 1 four
34: 4 sixes, 2 fives
33: 5 sixes, 1 three
33: 4 sixes, 1 five, 1 four
33: 3 sixes, 3 fives
And only the last configuration (3 sixes, 3 fives) can not win a pet elephant. Thus, the probability is

$$\frac{1+\binom{6}{5}\times 3+\binom{6}{4}+\binom{6}{4}\binom{2}{1}}{1+\binom{6}{5}\times 3+\binom{6}{4}+\binom{6}{4}\binom{2}{1}+\binom{6}{3}} = \frac{1+18+15+30}{1+18+15+30+20} = 64/84 = 16/21$$

Problem 3: Loaded elevators: (20 points)

An elevator is designed to tolerate a maximum weight of at most 5000 pounds. We are interested in the probability that the elevator is *overloaded*, meaning that the total weight of all people onboard exceeds the threshold of 5000 pounds. Assume the weight W_i of person i can be modeled as an exponential variable with parameter $\lambda = 1/150$, so that $f_{W_i}(w) = \lambda \exp(-\lambda w)$ for $w \geq 0$, and that the weights of different people are independent.

First assume that exactly 26 people climb onboard the elevator.

(a) (4 pt) Using Markov's inequality, compute an upper bound on the probability that the elevator is overloaded.

(b) (5 pt) Using the the central limit theorem, compute an approximation to the probability that the elevator is overloaded. (You may specify your answer in terms of the CDF $\Phi(z) = \mathbb{P}(Z \leq z)$ of a standard normal variable $Z \sim N(0, 1)$.)

Now suppose that a *random number* T of people board the elevator, where T is a geometric random variable with parameter $q \in (0, 1)$.

(c) (5 pt) Using Chebyshev's inequality, compute an upper bound on the probability that the elevator is overloaded, as a function of q.

(d) (6 pt) Now suppose that you observe that at least 20 people have boarded the elevator (so you know that $T \geq 20$). Compute the minimum mean-squared estimator (MMSE) of the total weight of people on the elevator given this information.

Solution:

(a) We want an upper bound on $\mathbb{P}(26 people \geq 5000 lbs)$ The expected weight of 26 people is $150 \cdot 26 = 3900$, so we can use Markov to bound it to $\mathbb{P}(26 people \geq 5000 lbs) \leq 3900/5000$

(b) The normalized value is $\frac{5000-3900}{\sqrt{585000}} = 1.438$ So $\mathbb{P}(\text{overload}) = 1 - \Phi(1.438)$

(c) We let W represent the total weight, so that $W = \sum_{i=1}^{T} W_i$. This is a standard random sum:

$$\mathbb{E}[W] = 150 \cdot \frac{1}{q}$$
$$\text{var}(W) = 150^2 \cdot \frac{1}{q} + 150^2 \cdot \frac{1-q}{q^2}$$

Chebyshev gives us that

$$\begin{aligned}
\mathbb{P}(W \geq 5000) &= \mathbb{P}(W - 150\frac{1}{q} \geq 5000 - \frac{150}{q} \\
&\leq \mathbb{P}(\left| W - \frac{150}{q} \right| \geq 5000 - \frac{150}{q}) \\
&\leq \frac{150^2 \cdot \frac{1}{q} + 150^2 \cdot \frac{1-q}{q^2}}{(5000 - \frac{150}{q})^2}
\end{aligned}$$

(d) We want $\mathbb{E}[W|T \geq 20]$, but by the restarting property of a geometric, we simply have $20 \cdot 150 + \mathbb{E}[W] = 150(20 + \frac{1}{q})$.

Problem 4: Homer Simpson at the power plant: (20 points)

One day, as Homer Simpson is working at the nuclear power plant, one of the reactors begins to melt down. The emission of radioactive particles can be modeled as a Poisson process with rate $\lambda = 100$ particles/second.

(a) (2 pt) What is the expected *total* number of particles that escape the time window $[10, 20] \cup [50, 60]$ seconds?

(b) (3 pt) Suppose that exactly 300 particles escaped in the first second. What is the PDF of the number of particles that escape in the next second?

In order to stop the spread of radiation, the plant is equipped with a set of n shields that are either "ON" or "OFF". Any OFF shield has no effect on the particle stream, whereas any ON shield blocks each particle with probability p, and let it through with probability $1 - p$, independently of all other particles. Suppose that each shield acts independently of all the other shields.

(c) (4 pt) For some fixed integer k (with $1 \leq k \leq n$), suppose that exactly k of the shields are ON, and consider the stochastic process defined by the particles that end up escaping. Prove that the expected number of particles that escape per second is equal to $100\,(1-p)^k$. Is this a Poisson process?

(d) (7 pt) Now suppose that Homer is distracted by eating donuts, so he only turns ON each shield (independently of all other shields) with probability 0.50.

Let X_n represent the number of particles that escape in the first second. (Recall that fixed integer n is the total number of shields.)

(i) Show that $\mathbb{E}[X_n] = 100\,(1 - \frac{p}{2})^n$.

(ii) Compute $\text{var}[X_n]$.

(*Hint:* Using conditional expectation, the result of (c) could be helpful, and a formula from p. 2 could be useful to you.)

(e) (4 pt) Does the sequence $\{X_1, X_2, X_3 \ldots\}$ converge in probability to any real number c? If not, explain why not. If so, give the value of c, and justify the convergence in probability.

Solution:

(a) The total number of particles that escape in the combined intervals is Poisson with parameter $\lambda\,(10+10) = 2000$ particles. Hence, the expected number of particles emitted is 2000.

(b) Due to the independence of the Poisson process for disjoint time intervals, this PDF is again Poisson with parameter $\lambda\,(1) = 100$ particles.

(c) Any given particle escapes only if it is *not* blocked by any of the k active shields, which happens with probability $\pi = (1-p)^k$, using independence of shields. Hence, the new process is Poisson with parameter $(1-p)^k\lambda = (1-p)^k(100)$. The expected number of particles escaping per second is thus $(1-p)^k\,(100)$.

(d) Letting K be a random variable representing the number of activated shields, then K is binomial (n, q) by definition. Using conditional expectation, we have

$$\begin{aligned}
\mathbb{E}[X_n] &= \mathbb{E}[\mathbb{E}[X \mid K]] \\
&= \mathbb{E}\left[(1-p)^k\,100\right] \\
&= 100\sum_{k=0}^{n}\binom{n}{k}\left(\frac{1}{2}\right)^n (1-p)^k \\
&= \left(\frac{1}{2}\right)^n 100\,(2-p)^n \\
&= 100\,(1-\frac{p}{2})^n.
\end{aligned}$$

using the binomial formula. Similarly, we have

$$\begin{aligned}
\mathbb{E}[X_n^2] &= \mathbb{E}[\mathbb{E}[X \mid K]] \\
&= \mathbb{E}\left[(1-p)^{2k}\,100^2 + (1-p)^k\,100\right] \\
&= 100^2\left(\frac{1}{2}\right)^n\sum_{k=0}^{n}\binom{n}{k}\left[(1-p)^2\right]^k + 100\left(\frac{1}{2}\right)^n\sum_{k=0}^{n}\binom{n}{k}(1-p)^k \\
&= 100^2\left(\frac{1}{2}\right)^n\left(1+(1-p)^2\right)^n + 100\left(\frac{1}{2}\right)^n\left(1+(1-p)\right)^n.
\end{aligned}$$

again using the binomial formula.

Hence, putting the pieces together, we have

$$\text{var}(X_n) = 100^2\left\{\left(\frac{1}{2}\right)^n\left(1+(1-p)^2\right)^n + 100\left(\frac{1}{2}\right)^n(1+\ (1-p)) - (1-\frac{p}{2})^{2n}\right\}.$$

(e) We claim that the sequence $\{X_n\}$ converges in probability to 0. To prove this, we use Markov's inequality to write

$$\mathbb{P}[X_n > \epsilon] \le \frac{\mathbb{E}[X_n]}{\epsilon},$$

and then note that $\mathbb{E}[X_n] \to 0$ as $n \to 0$.

Problem 5: (20 points) Every day that he leaves work at 2pm, Albert the Absent-minded Professor toggles his light switch according to the following protocol: (i) if the light is on, he switches it off with probability 0.80; and (ii) if the light is off, he switches it on with probability 0.30. At no other time (other than the end of each day) is the light switch touched.

(a) (2 pt) Suppose that on Monday night, Albert's office is equally likely to be light or dark. What is the probability that his office will be lit all five nights of the week (Monday through Friday) at 11pm each night?

(b) (3 pt) Suppose that you observe that his office is lit on both Monday and Friday nights. Compute the expected number of nights, from that Monday through Friday inclusive (5 possible nights total), that his office is lit.

(c) (3 pt) Suppose that Albert's office is lit on Monday night. Compute the expected number of days until the first night that his office is dark.

Now suppose that Albert has been working for five years.

(d) (3 pt) Is his light more likely to be on or off at the end of a given workday?

(e) (4 pt) What is the probability of his office light staying off for five consecutive nights?

Now suppose that for each night i the office light is left on, the electricity cost is exponentially distributed with parameter $\lambda = 1$ dollars, whereas if the light is left off, then the electricity cost is 0 dollars. Let Z_i represent the electricity cost on night i, and let $Z = \sum_{i=1}^{365} Z_i$ be the total electricity cost over one year.

(f) (5 pt) What is the expected total electricity cost for Albert's office light over the course of one year? Can you use the central limit theorem presented in class to approximate the probability $\mathbb{P}[Z \geq 365]$? If so, compute such an approximation; if not, explain why not.

Solution:

(a) Let L = light and D = dark. We want the sequence LLLLL; Pr[LLLLL] $= 1/2 \cdot 0.2^4 = 0.0008$

(b) E[days lit] = 2 Pr[DDDL] + 3 (Pr[DDLL] + Pr[DLDL] + Pr[LDDL]) + 4 (Pr[DLLL] + Pr[LLDL] + Pr[LDLL]) + 5 Pr[LLLL], where for each probability we start from an L state. So: E[days lit] = .7328

(c) The PDF is geometric with parameter p=0.8, so expected number of days is 1/p = 1.25

(d) Find steady state probabilities: $\pi_L = 3/11$ and $\pi_D = 8/11$, so the light is more likely to be off.

(e) Pr[DDDDD] $= 8/11 \cdot 0.7^4 = 0.17$

(f) $E(Z_i) = 3/11$ and $E(Z) = 365 \times 3/11$. We can not approximate Z using the central limit theorem. Because Z_i are not independent random variables.