QUESTION BOOKLET
EE 126 Spring 2006 Midterm #2
Thursday, April 13, 11:10-12:30pm

DO NOT OPEN THIS QUESTION BOOKLET UNTIL YOU ARE TOLD TO DO SO

- You have 80 minutes to complete the midterm.
- The midterm consists of three problems, provided in the question booklet (THIS BOOKLET), that are in no particular order of difficulty.
- Write your solution to each problem in the space provided in the solution booklet (THE OTHER BOOKLET). Try to be neat! If we can't read it, we can't grade it.
- You may give an answer in the form of an arithmetic expression (sums, products, ratios, factorials) that could be evaluated using a calculator. Expressions like $\binom{8}{3}$ or $\sum_{k=0}^{5}(1/2)^k$ are also fine.
- A correct answer does not guarantee full credit and a wrong answer does not guarantee loss of credit. You should concisely explain your reasoning and show all relevant work. The grade on each problem is based on our judgment of your understanding as reflected by what you have written.
- This is a closed-book exam except for two sheets of handwritten notes (one 8.5×11 page, both sides OR two pages, single side only), plus a calculator.

Problem 1: (12 points)

Diana the Daredevil is trying to break the world land-speed record using her rocket-powered motorcycle. In order to do so, she needs to cover more than 500 yards in 10 seconds.

If her motorcycle has Z pounds of rocket fuel to start, then it travels a distance of $X = 50\sqrt{Z}$ yards in 10 seconds.

(a)(5 points) Suppose that the amount of rocket fuel Z is a random variable uniformly distributed over $[50, 150]$. Compute the PDF and CDF of the distance X.

(b)(2 points) Compute the probability that Diana breaks the world record on any given trial.

(c)(5 points) Now suppose that Diana is allowed to take a total of n trials in order to try and break the world record. (Here n is some fixed positive integer: i.e., $n \in \{1, 2, 3, \ldots\}$). The amount of rocket fuel at the start is independent from trial to trial. What is the smallest possible integer n such that Diana has a better than 90% chance of breaking the world record over the n trials?

Solution:

(a) (5pt) We first compute the CDF

$$\begin{aligned} F_X(x) &= P(X < x) \\ &= P(50\sqrt{(}Z) < x) \\ &= P\left(Z < \left(\frac{x}{50}\right)^2\right) \\ &= = \begin{cases} 0 & \text{if } x < 50\sqrt{50} \\ \frac{\left(\frac{x}{50}\right)^2 - 50}{100} & \text{if } 50\sqrt{50} \le x < 50\sqrt{150} \\ 1 & \text{if } x > 50\sqrt{150} \end{cases} \end{aligned}$$

From which, we can derive the PDF

$$\begin{aligned} f_X(x) &= \frac{d}{dx} F_X(x) \\ &= = \begin{cases} 0 & \text{if } x < 50\sqrt{50} \\ \frac{x}{50^3} & \text{if } 50\sqrt{50} \le x < 50\sqrt{150} \\ 0 & \text{if } x > 50\sqrt{150} \end{cases} \end{aligned}$$

(b) (2pt) The probability that Diana breaks the world record is equal to the probability that her motorcycle travels a distance greater than 500 yards in 10 seconds. So,

$$P(X > 500) = 1 - F_X(500) = 1 - \frac{1}{2} = \frac{1}{2}$$

(c) (5pt) The probability that Diana breaks the world record in n trials is equal to the probability that the maximum distance traveled by her motorcycle in the n trials is greater than 500 yards in 10 seconds. Let Y denote the maximum distance traveled by her motorcycle. We have that $Y = \max(X_1, X_2, \ldots, X_n)$, where the X_i are independent and equally distributed random variables.

$$\begin{aligned} F_Y(y) &= P(Y < y) \\ &= P(\max(X_1, X_2, \ldots, X_n) < y) \\ &= P(X_1 < y, X_2 < y, \ldots, X_n < y) \\ &= P(X_1 < y)^n \\ &= F_X(y)^n \end{aligned}$$

$$P(Y > 500) = 1 - F_Y(500) = 1 - F_X(500)^n = 1 - \frac{1}{2}^n$$

and the minimum integer n for which $1 - \frac{1}{2}^n \geq 0.9$ is $n = 4$.

Problem 2: (14 points)

A car dealer sells two models of cars, Ray and Sprint. A Ray car breaks after time R days, where R is exponentially distributed with parameter λ_R. Similarly, a Sprint car breaks after a random time S that is exponentially distributed with parameter λ_S. The random variables R and S are independent.

When it is broken, a car is brought to the dealer for repair. The cost (in dollars) of fixing a broken Ray is a random variable uniformly distributed over the interval $[100, 300]$, whereas the cost (in dollars) of fixing a broken Sprint is a random variable that is uniformly distributed over $[200, 400]$ dollars. The costs of fixing different cars are independent.

(a)(3 points) Given that that a Ray did not break in the first h days, what is the expected time before the Ray breaks?

(b)(2 points) What is the probability that a Ray breaks before a Sprint?

(c)(4 points) Compute the mean lifetime of a Ray car given that Ray fails before Sprint.

(d)(5 points) On some day, the dealer has to fix N_R cars of the Ray model and N_S cars of the Sprint model. Here N_R and N_S are independent Poisson random variables with parameters μ_R and μ_S respectively. The dealer is interested in estimating the total revenue Z that he will earn from repairing all the cars that day. Compute the expected value, variance, and the moment generating function of Z.

Solution:

(a) (3pt) The expected time before the Ray breaks is given by

$$E(R|R > r) = r + \int_0^\infty P(R > r + h|R > r)dh$$

From the memoryless property of the exponential rand of variable $P(R > r + h|R > r) = P(R > h)$, we have

$$E(R|R > r) = r + \int_0^\infty P(R > h)dh = r + \int_0^\infty e^{\lambda_R h}dh = r + \frac{1}{\lambda_R}$$

(b) (2pt) The probability that a Ray breaks before a Sprint is given by

$$\begin{aligned} P(R \leq S) &= 1 - P(R > S) = 1 - \int_0^\infty \int_s^\infty \lambda_R e^{-\lambda_R r}\lambda_S e^{-\lambda_S s}drds \\ &= 1 - \int_0^\infty e^{-\lambda_R s}\lambda_S e^{-\lambda_S s}ds \\ &= 1 - \frac{\lambda_S}{\lambda_R + \lambda_S} \\ &= \frac{\lambda_R}{\lambda_R + \lambda_S} \end{aligned}$$

(c) (4pt)

$$\begin{aligned}
E(R|R<S) &= \int_0^\infty P(R>r|R<S)dr \\
&= \int_0^\infty \frac{P(r<R<S)}{P(R<S)}dr \\
&= \frac{1}{P(R<S)}\int_0^\infty \left[\int_r^\infty \int_r^S \lambda_R e^{-\lambda_R v}\lambda_S e^{-\lambda_S S}dvdS\right]dr \\
&= \frac{1}{P(R<S)}\int_0^\infty \left[\int_r^\infty [e^{-\lambda_R r}-e^{-\lambda_R S}]\lambda_S e^{-\lambda_S S}dS\right]dr \\
&= \frac{1}{P(R<S)}\int_0^\infty \left[e^{-(\lambda_R+\lambda_S)r}-\frac{\lambda_S}{\lambda_R+\lambda_S}e^{-(\lambda_R+\lambda_S)r}\right]dr \\
&= \int_0^\infty e^{-(\lambda_R+\lambda_S)r}dr \\
&= \frac{1}{\lambda_R+\lambda_S}
\end{aligned}$$

(d) (5pt) The moment generating function of random variable X which is uniformly distributed over $[a;b]$ is

$$M_X(s) = \int_a^b e^{sx}\frac{1}{b-a}dx = \frac{e^{sb}-e^{sa}}{s(b-a)}$$

The moment generating function of a Poisson random variable N with parameter μ is $M_N(s) = e^{\mu(e^s-1)}$.

The total revenue Z can be expressed as the sum of independent random variables

$$Z = \sum_{i=1}^{N_S} Y_{Si} + \sum_{j=1}^{N_T} Y_{Tj}$$

where the Y_{Si} are independent random variables uniformly distributed over $[100;300]$, and the Y_{Tj} are independent random variable uniformly distributed over $[200;400]$. Y_{Si} and Y_{Tj} are independent for any $i = 1,\ldots,N_S$ and $j = 1,\ldots,N_T$

By linearity of expectation, and law of iterated expextations, we obtain

$$
\begin{aligned}
E[Z] &= E\left[\sum_{i=1}^{N_S} Y_{Si} + \sum_{j=1}^{N_T} Y_{Tj}\right] \\
&= E\left[\sum_{i=1}^{N_S} Y_{Si}\right] + E\left[\sum_{j=1}^{N_T} Y_{Tj}\right] \\
&= E\left[E\left[\sum_{i=1}^{N_S} Y_{Si}|N_S\right]\right] + E\left[E\left[\sum_{j=1}^{N_T} Y_{Tj}|N_T\right]\right] \\
&= 200E[N_S] + 300E[N_T] \\
&= 200\mu_S + 300\mu_T
\end{aligned}
$$

Similarly, using the independence property of the random variables

$$
\begin{aligned}
\mathrm{Var}[Z] &= \mathrm{Var}\left[\sum_{i=1}^{N_S} Y_{Si}\right] + \mathrm{Var}\left[\sum_{j=1}^{N_T} Y_{Tj}\right] \\
&= E\left[\mathrm{Var}\left[\sum_{i=1}^{N_S} Y_{Si}|N_S\right]\right] + \mathrm{Var}\left[E\left[\sum_{i=1}^{N_S} Y_{Si}|N_S\right]\right] + E\left[\mathrm{Var}\left[\sum_{j=1}^{N_T} Y_{Tj}|N_T\right]\right] + \\
&\quad + \mathrm{Var}\left[E\left[\sum_{j=1}^{N_T} Y_{Tj}|N_T\right]\right] \\
&= \mathrm{Var}[Y_S]E[N_S] + (E[Y_S])^2\mathrm{Var}[N_S] + \mathrm{Var}[Y_T]E[N_T] + (E[Y_T])^2\mathrm{Var}[N_T] \\
&= \frac{200^2}{12}(\mu_S + \mu_T) + 200^2\mu_S + 300^2\mu_T
\end{aligned}
$$

The moment generating function of Z is

$$
\begin{aligned}
M_Z(s) &= E[e^{sZ}] \\
&= E\left[e^{s\left(\sum_{i=1}^{N_S} Y_{Si} + \sum_{j=1}^{N_T} Y_{Tj}\right)}\right] \\
&= E\left[e^{s\sum_{i=1}^{N_S} Y_{Si}}\right] E\left[e^{s\sum_{j=1}^{N_T} Y_{Tj}}\right] \\
&= E\left[E\left[e^{s\sum_{i=1}^{N_S} Y_{Si}}|N_S\right]\right] E\left[E\left[e^{s\sum_{j=1}^{N_T} Y_{Tj}}|N_T\right]\right] \\
&= E\left[M_{Y_S}(s)^{N_S}\right] E\left[M_{Y_T}(s)^{N_T}\right] \\
&= \left[M_{N_S}(s)|_{e^s = M_{Y_S}(s)}\right] \left[M_{N_T}(s)|_{e^s = M_{Y_T}(s)}\right] \\
&= e^{\mu_S(M_{Y_S}(s)-1)} e^{\mu_T(M_{Y_T}(s)-1)} \\
&= e^{\mu_S\left(\frac{e^{300b}-e^{100s}}{s200}-1\right)+\mu_T\left(\frac{e^{400s}-e^{200s}}{s200}-1\right)}
\end{aligned}
$$

Problem 3: (14 points)

George the Gambler is a good poker player: every round that he plays, he wins a random amount Z of money, distributed as $Z \sim N(\mu, \sigma^2)$ with $\mu > 0$. Every night, George goes to the local poker club, and plays $T = 1 + V$ rounds of poker, where V is a Poisson random variable with parameter $\lambda = 5$. Let X be the total amount of money that George wins in a given night.

(a)(2 points) In the absence of any further information, what is the best estimate of X? (Here "best" is measured by the minimum mean-squared error.)

(b)(3 points) Suppose that you observe that George plays $T = t$ rounds of poker (where t is some positive integer). Now what is the best estimate of X (again measured in terms of minimim mean-squared error)?

(c)(5 points) Now suppose that by peeking into George's wallet at the end of the night, you make a noisy observation Y of the amount of money X that he won—say of the form

$$Y \;=\; X + W$$

where $W \sim N(0,1)$ is Gaussian noise independent of X. Compute the Bayes' least squares estimate of X based on observing $T = t$ and $Y = y$. Also compute the *linear* least squares estimate (LLSE) of X based on observing $T = t$ and $Y = y$, and the error variance of the LLSE.

(d)(4 points) Now suppose that you observe that $\{T \leq 2\}$ and $Y = y$. Compute the Bayes' least squares estimate of X based on this information.

Solution:

(a) (2pt) We know that the best estimate of X is $E[X]$. E[X] can be computed using the law od iterated expectation: $E[X] = E[E[X|T]] = E[T\mu] = \mu(1 + E[V]) = \mu(1 + 5) = 6\mu$

(b) (3pt) Observing $T = t$ the best estimate of X becomes $E[X|T = t]$. When $T = t$, $X = Z_1 + Z_2 + \ldots + Z_t$ is a Gaussian random variable with mean $t\mu$ and variance $t\sigma^2$. So,

$$E[X|T = t] = E[Z_1 + Z_2 + \ldots + Z_t] = t\mu$$

(c) (5pt) Given that $T = t$, we know that X is a gaussian random variable. Y is a noisy observation of X through gaussian noise. So, Y itself gaussian and X, Y are jointly gaussian random variables. In this case, we know that the Bayes' least square estimate is the LLSE. We have that

$$E[X|T = t, Y = y] = t\mu + \frac{t\sigma^2}{t\sigma^2 + 1}(y - t\mu).$$

The error variance is $(1 - \rho^2)\sigma_X^2 = \left(1 - \frac{\sigma_X^2}{\sigma_Y^2}\right)\sigma_X^2 = \left(1 - \frac{t\sigma^2}{t\sigma^2+1}\right)t\sigma^2$

(d) (5pt) Using the conditional version of the law of total expectation, we have

$$\begin{aligned} E[X|T \leq 2, Y = y] &= E[E[X|T \leq 2, Y = y, T]|T \leq 2, Y = y] \\ &= E\left[T\mu + \frac{T\sigma^2}{T\sigma^2 + 1}(y - T\mu) \mid T \leq 2, Y = y\right] \\ &= \sum_{i=1}^{2} P(T = i|T \leq 2)\left[i\mu + \frac{i\sigma^2}{i\sigma^2 + 1}(y - i\mu)\right] \end{aligned}$$

where

$$\begin{aligned} P(T = i|T \leq 2, Y = y) &= \frac{P(T = i, T \leq 2, Y = y)}{P(T \leq 2, Y = y)} \\ &= \frac{f_{Y|T}(y|T = i)P(T = i)}{f_{Y|T}(y|T = 1)P(T = 1) + f_{Y|T}(y|T = 2)P(T = 2)} \end{aligned}$$

and

$$\begin{aligned} P(T = 1) &= P(V = 0) = e^{-5} \\ P(T = 2) &= P(V = 1) = 5e^{-5} \\ f_{Y|T}(y|T = i) &= \mathcal{N}(i\mu, i\sigma^2) \end{aligned}$$

Using the law of total variance we obtain

$$\begin{aligned} \mathrm{Var}[X|T \leq 2, Y = y] &= E[\mathrm{Var}[X|T \leq 2, Y = y, T]|T \leq 2, Y = y] + \\ &\quad + \mathrm{Var}[E[X|T \leq 2, Y = y, T]|T \leq 2, Y = y] \end{aligned}$$