NAME:

MAE 108 – Probability and Statistical Methods for Engineers - Spring 2013
Final Exam - June 12, 2013
3 hours, open book, open notes, calculator allowed, no cell phones or computers.
Write your answers directly on the exam. Each problem is worth 10 points; 100 points total.
The exam is **11** pages long. **Justify all your answers**.

---

(1) You ask your neighbor to water a sickly plant while you are on vacation. Without water it will die with probability 80%, and with water it will die with probability 15%. You are 90 percent certain that your neighbor will remember to water the plant.

(a) What is the probability that the plant will be alive when you return?

Solution: Let us call W the event that the neighbor did water the plant and A the event that the plan is alive. We want $P(A)$ which is calculated using the theorem of total probability: $P(A) = P(W)P(A|W) + P(\bar{W})P(A|\bar{W}) = .9 \times .85 + .1 \times .2 = 78.5\%$

(b) If the plant is dead when you return, what is the probability that your neighbor forgot to water it?

Solution: We now want $P(\bar{W}|\bar{A}) = P(\bar{W}\bar{A})/P(\bar{A})$. Clearly $P(\bar{A}) = 1 - 0.785 = 0.215$ and $P(\bar{W}\bar{A}) = P(\bar{A}|\bar{W})P(\bar{W}) = .8 \times .1$ and thus $P(\bar{W}|\bar{A}) = 37.2\%$.

(2) Fifty-two percent of the students at a certain college are females. Five percent of the students in this college are majoring in computer science. Two percent of the students are women majoring in computer science.

(a) If a student is female, what is the probability that the student is majoring in computer science?

Solution: Let us call F the event that the student is female and C the event that the student is majoring in computer science. Here we want to calculate the conditional probability $P(C|F) = P(CF)/P(F) = 0.02/0.52 = 3.9\%$.

(b) If a student is majoring in computer science, what is the probability that the student is female?

Solution: $P(F|C) = P(FC)/P(C) = 0.02/0.05 = 40\%$.

(c) What is the probability that a student selected at random is a male not majoring in computer science?

Solution: $P(\bar{F}\bar{C}) = 1 - P(F \cup C)$ and we have, using the addition rule, $P(F \cup C) = P(F) + P(C) - P(FC) = 0.52 + 0.05 - 0.02 = 0.55$ therefore $P(\bar{F}\bar{C}) = 45\%$.

(3) A student in MAE 108, William, enjoys playing the TETRIS video game. He keeps track of his best scores over a period of a week. His highest scores, in each day of the week, were: 110, 120, 100, 140, 110, 130, and 130.

(a) Estimate the mean and standard deviation for his scores using the method of moments.

Solution: We just have to apply for formula for $\bar{x}$ and s^2 and find $\bar{x} = 840/7 = 120$ and $s^2 = (1200)/6 = 200$ and thus $s = 14.14$

(b) If the variance for William's high score is known and equal to $\sigma^2 = 15^2$, determine the 98% confidence interval for the mean of the high scores.

Solution: The confidence interval is $[\bar{x} - k_{.99}\frac{\sigma}{\sqrt{n}}, \bar{x} + k_{.99}\frac{\sigma}{\sqrt{n}}]$. With $k_{.99} = 2.33$ we have the interval $[120 - 2.33\frac{15}{\sqrt{7}}, 120 + 2.33\frac{15}{\sqrt{7}}] = [106.8, 133.2]$.

(c) William wants to play the game every day until he knows the true value of his mean high score within ± 1 point with 98% confidence. How many additional days does he need to play the game?

Solution: $w = \frac{\sigma}{\sqrt{n}} k_{.99} \leq 1$ and thus the number of games he has to play is $n \geq (\sigma \times k_{.99})^2$. With $k_{.99} = 2.33$ and $\sigma = 15$ we get $n = 1222$ days, so he needs to play 1215 additional days which is over 3 years!

(4) The lifetime of interactive computer chips produced by a certain semiconductor manufacturer are normally distributed with mean 4.4×10^6 hours and standard deviation of 3×10^5 hours. All chips produced are statistically independent from each other.

(a) If a mainframe manufacturer requires that at least 90 percent of the chips from a large batch have lifetimes of at least 4×10^6 hours, can he buy the chips from the semiconductor firm?

Solution: We call T the lifetime of the chips. Given the value of μ and σ we have

$$P(T \geq 4.10^6) = 1 - P(T \leq 4.10^6) = 1 - \Phi\left(\frac{4 \times 10^6 - 4.4 \times 10^6}{3.10^5}\right) = 1 - \Phi(-4/3) = \Phi(4/3) \approx 90.1\%$$

and thus yes, he should contract with the firm.

(b) What is the probability that a batch of 5 chips contains at least 3 chips whose lifetimes are less than 3.8×10^6 hours?

Solution: First we have to calculate the probability of having a lifetime less than 3.8×10^6 hours. It is $P(T \leq 3.8 \times 10^6) = \Phi\left(\frac{3.8 \times 10^6 - 4.4 \times 10^6}{3.10^5}\right) = \Phi(-2) = 1 - \Phi(2) = 1 - 0.9772 = 0.0228$. We then have a Bernoulli sequence with $p = 2.28\%$. We want to compute $P(X \geq 3)$ among $n = 5$ trials trials, this is equal to $1 - P(X = 0) - P(X = 1) - P(X = 2)$. Applying the binomial we have:
$P(X = 0) = (1-p)^5 = 0.891081$, $P(X = 1) = 5p(1-p)^4 = 0.103953$, $P(X = 2) = \frac{5!}{3!2!}p^2(1-p)^3 = 10p^2(1-p)^3 = 0.004850$. Thus the probability we are looking for is given by $1 - 0.891081 - 0.103953 - 0.004850 \approx .0116\%$ (very small).

(5) We are monitoring radioactivity in a plant. Every day, someone measures the radioactivity level in the plant for a total of 6 seconds. We assume that the number of radioactive particles emitted by the plant follows a Poisson process at a rate of 0.5 particles per second. We look at the data for 4 consecutive days (so 4 non-overlapping time intervals of 6 seconds each).

(a) What is the probability that 2 or more particles are counted each day in all four days?

Solution: Let X be the number of particles counted. We must first find the probability that any given day emits 2 or more particles. Since the particles follow a Poisson process with $\nu = 0.5$ particles/second and t=6 seconds, we have $P(X \geq 2) = 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1) = 1 - \frac{(0.5\times6)^0 e^{-(0.5\times6)}}{0!} - \frac{(0.5\times6)^1 e^{-(0.5\times6)}}{1!} = 0.8009$.The probability that all four days will emit 2 or more particles is therefore $0.8009^4 = 0.4113$.

(b) What is the probability that 2 or more particles are counted in at least one of the days?

Solution: This is a binomial distribution using the probability that each day emits 2 or more particles, which was found in part (a) to be 0.8009. Therefore $P($2 or more particles emitted in at least one of 4 days$)=1 - P($2 or more particles emitted in 0 periods$) = 1 - \binom{4}{0} p^0 (1-p)^4 = 1 - (1 - 0.8009)^4 = 0.9984$.

(c) What is the probability that the first time in which 2 or more particles are counted is the third day?

Solution: This is a geometric distribution with $P(N = 3) = pq^{N-1} = 0.8009 \times (1 - 0.8009)^2 = 0.0317$.

(6) An inspector in charge of a criminal investigation is convinced with probability 60% that a certain suspect, named Eric L., is guilty. A new piece of information has just arrived showing that the crime was committed by a Frenchman. A person not guilty would have a 20% chance of being French. If Eric L. turns out to be French, how certain of the guilt of the suspect should the inspector now be?

Solution: Let us call G the event that the suspect is guilty and F the event that the suspect is French. We want to compute $P(G|F)$, since we now know that the suspect is in F. Using the multiplication rule we have $P(G|F) = P(GF)/P(F)$. To get $P(GF)$ we write again the multiplication rule as $P(GF) = P(F|G)P(G) = 1 \times 0.6 = 0.6$. In order to get $P(F)$ we have to use the theorem of total probability using the sets of events G and $\bar{G}$: $P(F) = P(F|G)P(G) + P(F|\bar{G})P(\bar{G}) = 1 \times 0.6 + 0.2 \times 0.4 = 0.68$. (Note: the probability that the suspect is French is NOT 20% because you know some additional information about him, namely that he is 60% probable to be guilty). Finally we have $P(G|F) = P(GF)/P(F) = .6/.68 \approx 88.2\%$.

(7) A radioactive mass emits α-particles from time to time. The time between two emissions is random. Let T represent the time in seconds between two emissions. Assume that the probability density function for T is given by

$$f_T(t) = \begin{cases} k^2 t\, e^{-kt} & t > 0 \\ 0 & t \leq 0, \end{cases}$$

where k is a positive constant.

(a) Find the cumulative distribution function (CDF) of the time between emissions.

Solution: Given the PDF, we can find the CDF by integration: $F_T(t) = \int_0^t k^2 u e^{-ku} du = -kue^{-ku}|_0^t + \int_0^t ke^{-ku} du = 1 - (1 + kt)e^{-kt}$.

(b) Use the result from (a) to show that f_T is a proper probability density function (PDF) for all values of $k > 0$. Calculate the mean time between emissions.

Solution: f is a positive function. Its integral on the whole domain is

$$\int_0^\infty k^2 u e^{-ku} du = F_T(\infty) = 1$$

and thus it is a pdf for all k. To compute the mean we have to calculate the integral

$$E(T) = \int_0^\infty t f_T(t) dt = \int_0^\infty k^2 t^2\, e^{-kt} dt = \frac{1}{k} \int_0^\infty u^2 e^{-u} du = \frac{2}{k}.$$

(c) You want to estimate the value of the parameter k. In order to do this, you carry out 100 different experiments, among which you find that 50 experiments show re-emission times of less than 10 seconds. Use this information to derive an equation to solve for k. Show graphically that this equation has a unique solution for $k > 0$.

Solution: We have $P(0 \leq T \leq 10) = F_T(T = 10) = 1 - (1 + 10k)e^{-10k} \approx 50/100$. Therefore k satisfies the equation $2(1 + 10k) = e^{10k}$. The function $2(1 + 10k)$ starts at the value 2 when $k = 0$ and goes up linearly. The function e^{10k} starts at the value 1 (below 2) when $k = 0$ and goes up exponentially (much faster than linearly). They meet therefore once at a positive value for k.

(8) This weekend you have decided to watch all three IRON MAN movies. You are, however, worried about their quality. Given your taste in movies, you have been told that you only have a probability 40% of liking each movie, which you assume to be statistically independent from each other. Denote I the random variable counting the number of Iron Man movies you end up enjoying.

(a) Derive the probability mass function for I (PMF).

Solution: I can take 4 different values: 0, 1, 2, and 3. Since this is a Bernoulli sequence, the PMF is given by the binomial distribution. Let us call $p = .4$ the probability of enjoying each movie:
$P(I = 0) = (1 - p)^3 = 21.6\%$
$P(I = 1) = 3p(1 - p)^2 = 43.2\%$
$P(I = 2) = 3p^2(1 - p) = 28.8\%$
$P(I = 3) = p^3 = 6.4\%$

(b) What is the mean value and standard deviation for I?

Solution: For the mean we have $E(I) = 0 + .432 + 2 \times .288 + 3 \times 0.064 = 1.2$. To get the variance we first have to compute $E(I^2) = 0 + .432 + 2^2 \times .288 + 3^2 \times 0.064 = 2.16$ and thus $Var(I) = 2.16 - 1.2^2 = 0.72$ and $\sigma_I = \sqrt{Var(I)} \approx 0.85$.

(c) If you like at least one of the Iron Man movies, what is the probability that you like all three?

Solution: We are being asked to compute $P(I = 3|I \geq 1) = P(I = 3)/P(I \geq 1)$. Clearly $P(I \geq 1) = 1 - P(I = 0) = 0.784$ and thus the probability is $0.064/0.784 \approx 8.2\%$.

(9) We consider a Poisson process with mean rate of occurrence ν. We call X_t the random variable measuring the number of occurrences of the event during the time interval $[0, t]$. Recall that for a Poisson process we have

$$P(X_t = x) = \frac{(\nu t)^x}{x!} e^{-\nu t}$$

and that the mean value of occurrences is $E(X_t) = \nu t$.

(a) If $\nu t < 1$, show that the PMF always decreases with x.

This is trivial because $\nu t < 1$ then $(\nu t)^{x+1} < (\nu t)^x$ therefore $P(X_t = x + 1) < P(X_t = x)$.

(b) If $\nu t > 1$, by comparing the values of $P(X_t = x)$ and $P(X_t = x - 1)$ show that the PMF first increases with x then decreases.

Solution: It is trivial to show that $P(X_t = x)/P(X_t = x - 1) = \nu t/x$ which is > 1, and thus increases, as long as $x < \nu t$ and goes down after that.

(c) If $\nu t > 1$ is not an integer number, what is the value of X at which the PMF reaches its maximum value?

Solution: The maximum is reached for the largest value of x such that $\nu t/x \geq 1$, which is the largest integer less than νt.

(d) If νt is an integer number larger than one, show that the PMF reaches its maximum value for two consecutive values of X.

Solution: Let us call n the integer equal to νt. In that case have $P(X_t = n)/P(X_t = n - 1) = \nu t/n = 1$ and thus the maximum for the PMF is reached at two points: $X = n - 1$ and $X = n$.

(10) [HARD PROBLEM - you are strongly advised you finish #1 to # 9 before starting this one]

The efficient frontier is a concept in finance quantifying the relationship between expected return and standard deviation; in this problem we derive some of the ideas of this financial concept. We consider three potential investment options, which are all statistically independent:

– Risk free investment $\mathbf{A_0}$ with expected return rate R_a and zero standard deviation;
– Risky investment **A** with expected return rate R_a and standard deviation σ_a;
– Risky investment **B** with expected return rate $R_b > R_a$ and standard deviation σ_b.

By expected return we mean that if you invest an amount x in one asset with expected return rate R, your expected return will be $x \times R$. A portfolio is generated by putting portions of your money into different investments and the expected return rate and standard deviation of the portfolio is derived in the following problems. Assume that the total amount you wish to invest is 1.

(a) Combination of risky and risk-free investments: You invest everything in $\mathbf{A_0}$ and **B** to form portfolio 1 (denoted $P1$). Derive the linear relation between expected return, E_1, and standard deviation, σ_1, of $P1$:

$$E_1 = \frac{R_b - R_a}{\sigma_b} \cdot \sigma_1 + R_a \tag{1}$$

Solution: Assuming invest (1-y) in A_0 and y in B,
$E_1 = R_a \cdot (1-y) + R_b \cdot y$,
$Var(P1) = (1-y)^2 \cdot 0^2 + y^2 \cdot \sigma_b^2 = (y \cdot \sigma_b)^2$, $\Rightarrow \sigma_1 = y \cdot \sigma_b$, $y = \frac{\sigma_1}{\sigma_b}$
$E_1 = (R_b - R_a)y + R_a = \frac{R_b - R_a}{\sigma_b} \cdot \sigma_1 + R_a$

(b) Combination of two risky investments: You invest everything in **A** and **B** to form portfolio 2 (denoted $P2$). Derive the quadratic relationship between variance, $Var(P2)$, and expected return, E_2, for $P2$:

$$Var(P2) = \left(\frac{E_2 - R_b}{R_b - R_a}\right)^2 \sigma_a^2 + \left(\frac{E_2 - R_a}{R_b - R_a}\right)^2 \sigma_b^2 \tag{2}$$

Solution: Assuming invest (1-y) in A and y in B,
$E_2 = R_a \cdot (1-y) + R_b \cdot y = (R_b - R_a)y + R_a$, $y = \frac{E_2 - R_a}{R_b - R_a}$

$$\begin{aligned} Var(P2) &= (1-y)^2\sigma_a^2 + y^2\sigma_b^2 \\ &= \left(1 - \frac{E_2 - R_a}{R_b - R_a}\right)^2 \sigma_a^2 + \left(\frac{E_2 - R_a}{R_b - R_a}\right)^2 \sigma_b^2 \\ &= \left(\frac{R_b - E_2}{R_b - R_a}\right)^2 \sigma_a^2 + \left(\frac{E_2 - R_a}{R_b - R_a}\right)^2 \sigma_b^2 = \left(\frac{E_2 - R_b}{R_b - R_a}\right)^2 \sigma_a^2 + \left(\frac{E_2 - R_a}{R_b - R_a}\right)^2 \sigma_b^2 \end{aligned}$$

[TURN TO NEXT PAGE]

(c) Assuming that $P1$ and $P2$ have the same expected return, show that $Var(P2)$ is always larger than, or equal to, $Var(P1)$.

Solution: Assume $E_1 = E_2 = E$, $Var(P1) = \sigma_1^2 = \left(\frac{E-R_a}{R_b-R_a}\right)^2 \sigma_b^2$,
$Var(P2) = \left(\frac{E-R_b}{R_b-R_a}\right)^2 \sigma_a^2 + \left(\frac{E-R_a}{R_b-R_a}\right)^2 \sigma_b^2$
$\because \left(\frac{E-R_b}{R_b-R_a}\right)^2 \sigma_a^2 \geq 0 \rightarrow Var(P2) \geq Var(P1)$

(d) Assuming again that $P1$ and $P2$ have the same expected return, E, find the value of E which minimizes the variance of $P2$.

Solution: we need $\frac{\partial Var(P2)}{\partial \hat{E}} = 0 \rightarrow 2(E - R_b)\left(\frac{\sigma_a}{R_b-R_a}\right)^2 + 2(E - R_a)\left(\frac{\sigma_b}{R_b-R_a}\right)^2 = 0$
multiply by $\frac{(R_b-R_a)^2}{2} \rightarrow E(\sigma_a^2 + \sigma_b^2) = R_b\sigma_a^2 + R_a\sigma_b^2 \rightarrow E = \frac{R_b\sigma_a^2+R_a\sigma_b^2}{\sigma_a^2+\sigma_b^2}$.