Practice Final 1

1. The joint probability density function of the lifetimes X and Y of two connected components in a machine is

$$f_{X,Y}(x,y) = \begin{cases} xe^{-x(1+y)}, & x \geq 0, y \geq 0; \\ 0, & \text{otherwise.} \end{cases}$$

   (a) What is the probability that the lifetime X of the first component exceeds 3?

   Solution: $\mathrm{P}(X > 3) = \int_3^\infty \int_0^\infty xe^{-x(1+y)} dy dx = \int_3^\infty e^{-x} dx = e^{-x}|_3^\infty = 0.05$.

   (b) Are X and Y independent? Justify your answer.

   Solution: No. If X and Y were independent then

$$f_{X,Y}(1,1) f_{X,Y}(2,2) = 2e^{-8} \quad \text{and} \quad f_{X,Y}(1,2) f_{X,Y}(2,1) = 2e^{-7}$$

   should both be equal to the same value $f_X(1) f_X(2) f_Y(1) f_Y(2)$.
   Alternatively one can calculate the marginal PDFs

$$f_X(x) = \int_0^\infty xe^{-x(1+y)} dy = e^{-x},$$
$$f_Y(y) = \int_0^\infty xe^{-x(1+y)} dx = \frac{1}{(1+y)^2}$$

   for $x, y \geq 0$ and conclude that $f_X(x) f_Y(y)$ is not the same function as $f_{X,Y}(x,y)$.

2. A radio station gives a gift to the third caller who knows the birthday of the radio talk show host. Each caller has a 0.7 probability of guessing the host's birthday, independently of other callers.

   (a) What is the probability mass function of the number of calls necessary to find the winner?

   Solution: n calls are necessary to find the winner if the n-th guess is correct and there are exactly two correct guesses among the first $n-1$. The probability of this is

$$\mathrm{P}(N = n) = \binom{n-1}{2} \cdot 0.7^3 \cdot 0.3^{n-3}.$$

   (b) What is the probability that the station will need five or more calls to find a winner?

   Solution: No winner has been found in the first four calls if the number of correct guesses in those calls is 0, 1, or 2. The probability of this is

$$\mathrm{P}(N \geq 5) = 0.3^4 + 4 \cdot 0.7 \cdot 0.3^3 + \binom{4}{2} \cdot 0.7^2 \cdot 0.3^2 = 0.3483.$$

3. Alice sends a message a that equals -1 or 1. Bob receives the value B which is a Normal random variable with mean a and standard deviation 0.5. Bob guesses that Alice sent 1 if $B > 0.5$, that Alice sent -1 if $B < -0.5$, and declares failure otherwise (when $|B| \le 0.5$).

   (a) What is the probability that Bob declares failure?

   Solution: For either message, failure occurs when a normal random variable is between 1 and 3 standard deviations from the mean on one side. If N is a Normal$(0,1)$ random variable then
$$\mathrm{P}(1 \le N < 3) = \mathrm{P}(N < 3) - \mathrm{P}(N < 1) \approx 0.9987 - 0.8413 = 0.1574.$$

   (b) Given that Bob didn't declare failure, what is the probability that his guess is correct?

   Solution: By symmetry we may assume Alice sent 1. The event "Bob's guess is correct *and* failure didn't occur" happens when B takes value 0.5 or larger, or when a Normal$(0,1)$ random variable N takes value at most 1, which is approximately 0.8413. Therefore the conditional probability that Bob's guess is correct is about $0.8413/(1-0.1574) \approx 0.9985$.

4. The number of people who enter an elevator on the ground floor is a Poisson random variable with mean 10. There are 20 floors above (not including) the ground floor and each person is equally likely to get off on any one of them, independently of all others.

   (a) What is the probability p that the elevator doesn't stop on the seventh floor?

   Solution: The number of passengers bound for the seventh floor is a Poisson random variable with mean $10/20 = 1/2$, so the probability that no passengers land there is the probability that this random variable takes value zero, which is $p = e^{-1/2} \approx 0.6065$.

   Alternatively, by conditioning on the number N of passengers who enter the elevator and applying the total probability theorem,
$$p = \sum_{n=0}^{\infty} (19/20)^n \cdot \mathrm{P}(N = n) = \sum_{n=0}^{\infty} (19/20)^n \cdot e^{-10} \cdot \frac{10^n}{n!}$$
$$= e^{-10} \cdot \sum_{n=0}^{\infty} \frac{(10 \cdot 19/20)^n}{n!} = e^{-10} \cdot e^{10 \cdot 19/20} = e^{-1/2}.$$

   (b) What is the expected number of stops that the elevator will make?
   (Express the answer in terms of p in case you didn't complete part (a).)

   Solution: By linearity of expectation, the expected number of stops is the sum of the probabilities that the elevator stops on floor 1 up to floor 20. By part (a) each of these probabilities is $1 - e^{-1/2}$ so the answer is $20 \cdot (1 - e^{-1/2}) \approx 7.869$.

5. 500 balls are drawn without replacement from a bin with 600 black balls and 400 white balls.

   (a) What is the expected number of black balls drawn?

   Solution: In any given draw the probability that the ball is black is $600/1000 = 3/5$. By linearity of expectation the expected number of black balls drawn is $(3/5) \cdot 500 = 300$.

(b) What is the variance of the number of black balls drawn?

Solution: The variance of the number of black balls is the sum of the 500 variances v indicating that any given ball is black plus the $500 \cdot 499$ covariances c indicating that any two given balls are black. Each individual variance is $v = 3/5 \cdot 2/5 = 6/25$, while each of the covariances equals

$$c = \frac{600}{1000} \cdot \frac{599}{999} - \left(\frac{600}{1000}\right)^2 \approx -2.402 \cdot 10^{-4},$$

from where the variance is $500v + 500 \cdot 499 \cdot c \approx 60.06$.

(c) Is the probability you drew fewer than 200 black balls more than 2%? Justify your answer.

Solution: The standard deviation of the number of black balls drawn is about $\sqrt{60.06} \approx 7.750$. In order to draw fewer than 200 black balls, the number of black balls drawn must be $100/7.750 \approx 12.90$ standard deviations lower than the mean. By Chebyshev's inequality, the probability of this is at most $1/12.90^2 \approx 0.006$. This is much less than 2%.

Practice Final 2

1. Urn A has 4 blue balls. Urn B has 1 blue ball and 3 red balls.

(a) You draw a ball from a random urn and it is blue. What is the probability that it came from urn A?

Solution: Let B_1 be the event the ball is blue and A be the event the ball came from urn A. By Bayes' rule

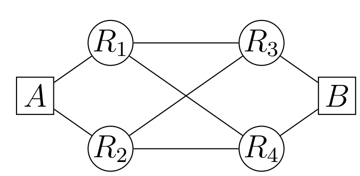
$$\mathrm{P}(A|B_1) = \frac{\mathrm{P}(B_1|A)\,\mathrm{P}(A)}{\mathrm{P}(B_1|A)\,\mathrm{P}(A) + \mathrm{P}(B_1|A^c)\,\mathrm{P}(A^c)} = \frac{1 \cdot (1/2)}{1 \cdot (1/2) + (1/4) \cdot (1/2)} = \frac{4}{5}.$$

(b) You draw another ball from the same urn. What is the probability that the second ball is also blue?

Solution: Let B_2 be the event that the second ball is blue. By the total probability theorem and Bayes' rule

$$\mathrm{P}(B_2|B_1) = \frac{\mathrm{P}(B_2 \cap B_1)}{\mathrm{P}(B_1)} = \frac{\mathrm{P}(B_2 \cap B_1|A)\,\mathrm{P}(A) + \mathrm{P}(B_2 \cap B_1|A^c)\,\mathrm{P}(A^c)}{\mathrm{P}(B_1|A)\,\mathrm{P}(A) + \mathrm{P}(B_1|A^c)\,\mathrm{P}(A^c)}$$
$$= \frac{1 \cdot (1/2) + (1/4)^2 \cdot (1/2)}{1 \cdot (1/2) + (1/4) \cdot (1/2)} = \frac{17}{20}.$$

2. Computers A and B are linked through routers R_1 to R_4 as in the picture. Each router fails independently with probability 10%.



(a) What is the probability there is a connection between A and B?

Solution: Let R_i be the event that router i is operational. The event "there is a connection between A and B" is $(R_1 \cup R_2) \cap (R_3 \cup R_4)$. By independence

$$\begin{aligned}
\mathrm{P}((R_1 \cup R_2) \cap (R_3 \cup R_4)) &= \mathrm{P}(R_1 \cup R_2)\,\mathrm{P}(R_3 \cup R_4) \\
&= (1 - \mathrm{P}(R_1^c \cap R_2^c))(1 - \mathrm{P}(R_1^c \cap R_2^c)) \\
&= (1 - \mathrm{P}(R_1^c)\,\mathrm{P}(R_2^c))(1 - \mathrm{P}(R_3^c)\,\mathrm{P}(R_4^c)) \\
&= (1 - 0.1^2)^2 \\
&= 0.9801.
\end{aligned}$$

(b) Are the events "there is a connection between A and B" and "exactly two routers fail" independent? Justify your answer.

Solution: No. The probability that there is a connection between A And B given that exactly two routers fail is $2/3$: Given that exactly two routers fail, the failed routers are equally likely to be any of the 6 pairs $R_1R_3, R_1R_4, R_2R_3, R_2R_4, R_1R_2, R_3R_4$, and there is a connection between A and B in the first 4 out of these 6 possibilities. This probability is not equal to the unconditional probability from part (a) and so the two events are not independent.

3. A bus takes you from A to B in 10 minutes. On average a bus comes once every 5 minutes. A taxi takes you in 5 minutes, and on average a taxi comes once every 10 minutes. Their arrival times are independent exponential random variables. A bus comes first.

(a) If you want to minimize the (expected) travel time, should you take this bus?

Solution: Yes. If you waited for a taxi your expected travel time would be the expected waiting time for the next taxi which is 10 minutes plus its travel time which is another 5 minutes for a total of 15 minutes.

(b) If you do take the bus, what is the probability that you made the wrong decision?

Solution: The probability of a wrong decision is the probability that a taxi arrives within the next five minutes, which is the probability that an Exponential(1/10) random variable is less than 5, which is $1 - e^{-5/10} = 1 - e^{-1/2} \approx 39.35\%$.

4. 10 people toss their hats and each person randomly picks one. The experiment is repeated one more time.

(a) What is the probability that Bob picked his own hat both times?

Solution: By independence, the probability that Bob picked his hat both times is the product of the probabilities that he picked it in each trial, so it is $(1/10) \cdot (1/10) = 1/100$.

(b) Let A be the event that at least one person picked their own hat both times. True or false: $P(A) > 25\%$? Justify your answer.

Solution: False. Let X_i take value 1 if person i picked their hat both times. A occurs if $X = X_1 + \cdots + X_{10} \geq 1$. By part (a) and linearity of expectation, $\mathrm{E}[X] = 10 \cdot (1/100) = 0.1$. By Markov's inequality, $\mathrm{P}[X \geq 1] \leq \mathrm{E}[X]/1 = 0.1$ which is less than 25%.

5. X is a Normal$(0, \Theta)$ random variable, where the prior PMF of the parameter Θ is $\mathrm{P}(\Theta = 1/2) = 1/2$, $\mathrm{P}(\Theta = 1) = 1/2$. You observe the following three independent samples of X: $1.0, 1.0, -1.0$.

   (a) What is the posterior PMF of Θ?

   Solution: By Bayes' rule

$$f_{\Theta|X_1X_2X_3}(\theta|1.0, 1.0, -1.0) \propto f_{X_1X_2X_3|\Theta}(1.0, 1.0, -1.0|\theta)\,\mathrm{P}(\Theta = \theta) \propto \frac{1}{\theta^3}e^{-3/2\theta^2}\,\mathrm{P}(\Theta = \theta).$$

   As Θ is equally likely to take values $1/2$ and 1, the posterior PMF is

$$f_{\Theta|X_1X_2X_3}(1/2|1.0, 1.0, -1.0) = \frac{8e^{-6}}{8e^{-6} + e^{-3/2}} \quad f_{\Theta|X_1X_2X_3}(1|1.0, 1.0, -1.0) = \frac{e^{-3/2}}{8e^{-6} + e^{-3/2}}.$$

   (b) What is the MAP estimate of Θ?

   Solution: As $e^{-3/2} \approx 0.2231$ is larger than $8e^{-6} \approx 0.0198$ the MAP estimate is $\hat{\Theta} = 1$.

   (c) What is the posterior probability that $|X| \geq 1$?

   Solution: The posterior probabilities of Θ are $1/2$ with probability about $0.0198/(0.2231 + 0.0198) \approx 0.0815$ and 1 with probability about $0.2231/(0.2231 + 0.0198) \approx 0.9185$. By the total probability theorem the posterior probability of $|X| \geq 1$ is about

$$\begin{aligned}
0.0815 \cdot \mathrm{P}(|\mathrm{Normal}(0, 1/2)| \geq 1) &+ 0.9185\,\mathrm{P}(|\mathrm{Normal}(0, 1)| \geq 1) \\
&\approx 0.0815 \cdot 2\,\mathrm{P}(\mathrm{Normal}(0, 1) \geq 2) + 0.9185 \cdot 2\,\mathrm{P}(\mathrm{Normal}(0, 1) \geq 1) \\
&\approx 0.0815 \cdot 2 \cdot 0.023 + 0.9185 \cdot 2 \cdot 0.159 \\
&\approx 0.2958.
\end{aligned}$$

Practice Final 3

1. Let X, Y, Z be independent Binomial$(2, \frac{1}{2})$ random variables.

   (a) What is the conditional PMF of X conditioned on $X \neq Z$?

   Solution: The joint PMF is

$$\begin{aligned}
\mathrm{P}(X = 0, X \neq Z) &= \mathrm{P}(X = 0, Z = 1) + \mathrm{P}(X = 0, Z = 2) = \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{4} = \frac{3}{16} \\
\mathrm{P}(X = 1, X \neq Z) &= \mathrm{P}(X = 1, Z = 0) + \mathrm{P}(X = 1, Z = 2) = \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{4} = \frac{4}{16} \\
\mathrm{P}(X = 2, X \neq Z) &= \mathrm{P}(X = 2, Z = 0) + \mathrm{P}(X = 2, Z = 1) = \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{2} = \frac{3}{16}.
\end{aligned}$$

   The conditional PMF is the joint PMF normalized by $\mathrm{P}(X \neq Z)$, which is

$$\mathrm{P}(X = 0|X \neq Z) = \frac{3}{10}, \quad \mathrm{P}(X = 1|X \neq Z) = \frac{4}{10}, \quad \mathrm{P}(X = 2|X \neq Z) = \frac{3}{10}.$$

(b) Are X and Y independent conditioned on $(X \neq Z)$ AND $(Y \neq Z)$?

Solution: No. We show that

$$P(X = 1 \mid X \neq Z, Y \neq Z)\,P(Y = 1 \mid X \neq Z, Y \neq Z) \neq P(X = 1, Y = 1 \mid X \neq Z, Y \neq Z).$$

By symmetry the two probabilities on the left are the same. We calculate these expressions:

$$\begin{aligned}
P(X \neq Z, Y \neq Z) &= \sum_{z\in\{0,1,2\}} P(X \neq Z, Y \neq Z|Z = z)\,P(Z = z) \\
&= (3/4)^2 \cdot (1/4) + (1/2)^2 \cdot (1/2) + (3/4)^2 \cdot (1/4) \\
&= 13/32, \\
P(X = 1, X \neq Z, Y \neq Z) &= P(X = 1, Z \neq 1, Y \neq Z) \\
&= P(X = 1)\,P(Z \neq 1, Y \neq Z) \\
&= (1/2) \cdot (P(Y \neq 0, Z = 0) + P(Y \neq 2, Z = 2)) \\
&= (1/2) \cdot 2 \cdot (1/4) \cdot (3/4) \\
&= 3/16, \\
P(X = 1, Y = 1, X \neq Z, Y \neq Z) &= P(X = 1, Y = 1, Z \neq 1) \\
&= (1/2)(1/2)(1/2) \\
&= 1/8.
\end{aligned}$$

By the conditional probability formula the expression on the left is $((3/16)/(13/32))^2 \approx 0.2130$ and the one on the right is $(1/8)/(13/32) \approx 0.3077$. These are not equal.

2. Alice and Bob decide to meet somewhere. Alice's arrival time A is uniform between 12:00 and 12:45. Bob's arrival time B is uniform between 12:15 and 1:00. Their arrival times are independent.

(a) Let f_{A-B} be the PDF of $A - B$. What is $f_{A-B}(0)$?

Solution: We model A and B as Uniform$(0, 3/4)$ and Uniform$(1/4, 1)$ random variables respectively (at the hour scale). By the convolution formula, $f_{A-B}(0) = \int_{-\infty}^{\infty} f_A(t)f_B(t)dt$, where f_A, f_B are the PDFs of A and B. $f_A(t)f_B(t)$ takes value $(4/3)^2$ when t is between $1/4$ and $3/4$ and 0 otherwise, so the integral equals $(1/2) \cdot (4/3)^2 = 8/9$. (If time is scaled in minutes the answer is 60 times smaller.)

(b) What is the probability that Bob arrives before Alice?

Solution: The event that Bob arrives before Alice is the value of the integral $\int_{a>b} f_A(a)f_B(b)dadb$. The value of the integrand is $(4/3)^2$ when (a, b) is in the interior of the triangle with vertices $(1/4, 1/4), (1/4, 3/4), (3/4, 3/4)$ and zero elsewhere. The triangle has area $(1/2)^2/2 = 1/8$. Therefore $P(A > B) = (1/8)(4/3)^2 = 2/9$.

3. Let $Y = AX + B$ where A, B, X are independent Normal$(0, 1)$ random variables.

(a) What is $\mathrm{Var}[E[Y|X]]$?

Solution: By linearity of expectation, $E[AX+B|X] = E[A]X+E[B] = 0$ so $\mathrm{Var}[E[Y|X]] = 0$.

(b) What is $E[\mathrm{Var}[Y|X]]$?

Solution: By independence, $\mathrm{Var}[AX+B|X] = \mathrm{Var}[AX|X]+\mathrm{Var}[B] = X^2\,\mathrm{Var}[A]+\mathrm{Var}[B] = X^2 + 1$, so $E[\mathrm{Var}[Y|X]] = E[X^2 + 1] = \mathrm{Var}[X] + 1 = 2$.

4. Boys and girls arrive independently at a meeting point at a rate of one boy per minute and one girl per minute, respectively. Let T be the first time at which both a boy and a girl have arrived.

   (a) Find the cumulative distribution function (CDF) of T.

   Solution: The probability that a boy has arrived by time t is $1 - e^{-t}$, i.e. the CDF of an Exponential(1) random variable. The probability that a boy has arrived by time t is therefore $1 - e^{-t}$, and same for a girl. The events are independent, the probability that both have arrived by time t is $\mathrm{P}(T \le t) = (1 - e^{-t})^2$ if $t \ge 0$ and 0 if not.

   (b) What is the expected value of T? (**Hint:** You don't *have* to use calculus.)

   Solution: We can write $T = T_1 + T_2$ where T_1 is the arrival time of the first person and T_2 is the arrival time of the next person *of the opposite gender*. As people arrive at a rate of two per minute, T_1 is an Exponential(2) random variable. By the memoryless property T_2 is an Exponential(1) random variable. Therefore $\mathrm{E}[T] = \mathrm{E}[T_1] + \mathrm{E}[T_2] = 1/2 + 1 = 3/2$.

5. A deck of cards is divided into 26 pairs. Let X be the number of those pairs in which both cards are of the same suit. (A deck of cards has 4 suits and each suit has 13 cards.)

   (a) What is the expected value of X?

   Solution: We can write $X = X_1 + \cdots + X_{26}$ where X_i is 1 if the cards in the i-th pair are of the same suit and 0 if not. Then $\mathrm{E}[X_i] = \mathrm{P}(X_i = 1)$ is the probability that the i-th pair's cards are of the same suit, which is $12/51$ because conditioned on the first card's suit, there are 12 out of 51 identical choices for the second one. By linearity of expectation $\mathrm{E}[X] = \mathrm{E}[X_1] + \cdots + \mathrm{E}[X_{26}] = 26 \cdot 12/51 \approx 6.118$.

   (b) What is the variance of X?

   Solution: The variance of X is the sum of the 26 variances of X_i and the $26 \cdot 25$ covariances of X_i and X_j. The variance of X_i is $v = (12/51) \cdot (1 - 12/51) \approx 0.1799$. The covariance of X_i and X_j is

   $$\mathrm{E}[X_i X_j] - \mathrm{E}[X_i]\,\mathrm{E}[X_j] = \mathrm{P}(X_i = 1, X_j = 1) - \mathrm{P}(X_i = 1)\,\mathrm{P}(X_j = 1).$$

   The term $\mathrm{P}(X_i = 1, X_j = 1)$ is the probability of the event A that within both the i-th pair and the j-th pair, both cards are of the same suit. We can calculate this using the total probability theorem applied to the event E that the first card of the i-th pair and the first card of the j-th pair are of the same suit:

   $$\mathrm{P}(X_i = 1, X_j = 1) = \mathrm{P}(A) = \mathrm{P}(A|E)\,\mathrm{P}(E) + \mathrm{P}(A|E^c)\,\mathrm{P}(E^c).$$

   The probability of E is $12/51$. Conditioned on E, A happens if the second cards of both pairs are also of the same suit, which is $11/50 \cdot 10/49$. Conditioned on E^c—for example, if the i-th pair's first card is a heart and the j-th pair first card is a spade—A happens if the second cards are a heart and a spade respectively, which happens with probability $(12/50) \cdot (12/49)$, and so

   $$\mathrm{P}(A) = \frac{11}{50} \cdot \frac{10}{49} \cdot \frac{12}{51} + \frac{12}{50} \cdot \frac{12}{49} \cdot \left(1 - \frac{12}{51}\right).$$

   Therefore the covariance of X_i and X_j equals

   $$c = \mathrm{P}(A) - \left(\frac{12}{51}\right)^2 \approx 0.0001469.$$

Finally, $\mathrm{Var}[X] = 26 \cdot v + 26 \cdot 25 \cdot c \approx 4.7737$.