Sundry

Before you start your homework, write down your team. Who else did you work with on this homework? List names and email addresses. (In case of homework party, you can also just describe the group.) How did you work on this homework? Working in groups of 3-5 will earn credit for your "Sundry" grade.

Please copy the following statement and sign next to it:

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up.

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up. (*signature here*)

1 Exponential Practice II

(a) Let $X_1, X_2 \sim \text{Exponential}(\lambda)$ be independent, $\lambda > 0$. Calculate the density of $Y := X_1 + X_2$. [*Hint*: One way to approach this problem would be to compute the CDF of Y and then differentiate the CDF.]

(b) Let $t > 0$. What is the density of X_1, conditioned on $X_1 + X_2 = t$? [*Hint*: Once again, it may be helpful to consider the CDF $\mathbb{P}(X_1 \le x \mid X_1 + X_2 = t)$. To tackle the conditioning part, try conditioning instead on the event $\{X_1 + X_2 \in [t, t+\varepsilon]\}$, where $\varepsilon > 0$ is small.]

Solution:

(a) Let $y > 0$. Observe that if $X_1 + X_2 \le y$, then since $X_1, X_2 \ge 0$, it follows that $X_1 \le y$ and $X_2 \le y - X_1$.

$$\begin{aligned}
\mathbb{P}(Y \le y) = \mathbb{P}(X_1 \le y, X_2 \le y - X_1) &= \int_0^y \int_0^{y-x_1} \lambda \exp(-\lambda x_1) \lambda \exp(-\lambda x_2) \, dx_2 \, dx_1 \\
&= \lambda^2 \int_0^y \exp(-\lambda x_1) \cdot \frac{1 - \exp(-\lambda(y - x_1))}{\lambda} \, dx_1 \\
&= \lambda \int_0^y \big(\exp(-\lambda x_1) - \exp(-\lambda y)\big) \, dx_1 = \lambda \Big(\frac{1 - \exp(-\lambda y)}{\lambda} - y \exp(-\lambda y) \Big).
\end{aligned}$$

Upon differentiating the CDF, we have

$$\begin{aligned} f_Y(y) &= \frac{\mathrm{d}}{\mathrm{d}y}\mathbb{P}(Y \le y) = \lambda \exp(-\lambda y) - \lambda \exp(-\lambda y) + \lambda^2 y \exp(-\lambda y) \\ &= \lambda^2 y \exp(-\lambda y), \qquad y > 0. \end{aligned}$$

(b) Let $0 \le x \le t$. Following the hint, we have

$$\begin{aligned} \mathbb{P}(X_1 \le x \mid X_1 + X_2 \in [t, t+\varepsilon]) &= \frac{\mathbb{P}(X_1 \le x, X_1 + X_2 \in [t, t+\varepsilon])}{\mathbb{P}(X_1 + X_2 \in [t, t+\varepsilon])} \\ &= \frac{\mathbb{P}(X_1 \le x, X_2 \in [t - X_1, t - X_1 + \varepsilon])}{f_Y(t) \cdot \varepsilon} \\ &= \frac{\int_0^x \int_{t-x_1}^{t-x_1+\varepsilon} \lambda \exp(-\lambda x_1) \lambda \exp(-\lambda x_2)\, \mathrm{d}x_2\, \mathrm{d}x_1}{\lambda^2 t \exp(-\lambda t) \cdot \varepsilon} \\ &= \frac{\lambda^2 \int_0^x \exp(-\lambda x_1) \exp(-\lambda (t - x_1)) \varepsilon\, \mathrm{d}x_1}{\lambda^2 t \exp(-\lambda t) \cdot \varepsilon} = \frac{\int_0^x \mathrm{d}x_1}{t} = \frac{x}{t}. \end{aligned}$$

This means that the density is

$$f_{X_1 \mid X_1 + X_2}(x \mid t) = \frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}(X \le x \mid X_1 + X_2 = t) = \frac{1}{t}, \qquad x \in [0, t],$$

which means that conditioned on $X_1 + X_2 = t$, X_1 is actually uniform on the interval $[0, t]$!

2 Normal Distribution

Recall the following facts about the normal distribution: if $X \sim \mathcal{N}(\mu, \sigma^2)$, then the random variable $Z = (X - \mu)/\sigma$ is standard normal, i.e. $Z \sim \mathcal{N}(0, 1)$. There is no closed-form expression for the CDF of the standard normal distribution, so we define $\Phi(z) = \mathbb{P}[Z \le z]$. You may express your answers in terms of $\Phi(z)$.

The average jump of a certain frog is 3 inches. However, because of the wind, the frog does not always go exactly 3 inches. A zoologist tells you that the distance the frog travels is normally distributed with mean 3 and variance $1/4$.

(a) What is the probability that the frog jumps more than 4 inches?

(b) What is the probability that the distance the frog jumps is between 2 and 4 inches?

Solution:

(a) First, we write down the probability we want to find, then transform the probability in order to work with the standard normal.

$$\mathbb{P}[X > 4] = \mathbb{P}[X - 3 > 1] = \mathbb{P}\left[\frac{X - 3}{1/2} > 2\right] = \mathbb{P}[Z > 2] = 1 - \Phi(2) \approx 0.0228$$

(b) Since the mean of the jump is 3, and the normal distribution is symmetric, we can rewrite the desired probability as

$$\mathbb{P}[2 < X < 4] = 1 - (\mathbb{P}[X > 4] + \mathbb{P}[X < 2]) = 1 - 2 \cdot \mathbb{P}[X > 4].$$

We have computed $\mathbb{P}[X > 4] = 0.0228$ in Part (a), so we can plug this in to get 0.9544.

3 Noisy Love

Due to the Central Limit Theorem, the Gaussian distribution is often used as a model for noise. In this problem, we will see how to perform calculations with Gaussian noise models.

Suppose you have confessed to your love interest on Valentine's Day and you are waiting to hear back. Your love interest is trying to send you a binary message: "0" means that your love interest is not interested in you, while "1" means that your love interest reciprocates your feelings. Let X be your love interest's message for you. Your current best guess of X has $\mathbb{P}(X = 0) = 0.7$ and $\mathbb{P}(X = 1) = 0.3$. Unfortunately, your love interest sends you the message through a noisy channel, and instead of receiving the message X, you receive the message $Y = X + \varepsilon$, where ε is independent Gaussian noise with mean 0 and variance 0.49.

(a) First, you decide upon the following rule: if you observe $Y > 0.5$, then you will assume that your love interest loves you back, whereas if you observe $Y \leq 0.5$, then you will assume that your love interest is not interested in you. What is the probability that you are correct using this rule? (Express your answer in terms of the CDF of the standard Gaussian distribution $\Phi(z) = \mathbb{P}(\mathcal{N}(0,1) \leq z)$, and then evaluate your answer numerically.)

(b) Suppose you observe $Y = 0.6$. What is the probability that your love interest loves you back? [*Hint*: This problem requires conditioning on an event of probability 0, namely, the event $\{Y = 0.6\}$. To tackle this problem, think about conditioning on the event $\{Y \in [0.6, 0.6 + \delta]\}$, where $\delta > 0$ is small, so that $f_Y(0.6) \cdot \delta \approx \mathbb{P}(Y \in [0.6, 0.6 + \delta])$, and then apply Bayes Rule.]

(c) Suppose you observe $Y = y$. For what values is it more likely than not that your love interest loves you back? [*Hint*: As before, instead of considering $\{Y = y\}$, you can consider the event $\{Y \in [y, y + \delta]\}$ for small $\delta > 0$. So, when is $\mathbb{P}(X = 1 \mid Y \in [y, y + \delta]) \geq \mathbb{P}(X = 0 \mid Y \in [y, y + \delta])$?]

(d) Your new rule is to assume that your love interest loves you back if (based on the value of Y that you observe) it is more likely than not that your love interest loves you back. Under this new rule, what is the probability that you are correct?

Solution:

(a) The probability that you are correct is

$$\begin{aligned}
&\mathbb{P}(X=0)\mathbb{P}(Y \le 0.5 \mid X=0) + \mathbb{P}(X=1)\mathbb{P}(Y > 0.5 \mid X=1) \\
&\quad = 0.7\mathbb{P}(\mathcal{N}(0,0.49) \le 0.5) + 0.3\mathbb{P}(\mathcal{N}(1,0.49) > 0.5) \\
&\quad = 0.7\mathbb{P}\left(\mathcal{N}(0,1) \le \frac{0.5}{0.7}\right) + 0.3\mathbb{P}\left(\mathcal{N}(0,1) > -\frac{0.5}{0.7}\right) = \mathbb{P}\left(\mathcal{N}(0,1) \le \frac{5}{7}\right) = \Phi\left(\frac{5}{7}\right) \\
&\quad \approx 0.762.
\end{aligned}$$

(b) By conditioning on $\{Y \in [0.6, 0.6+\delta]\}$, we have

$$\begin{aligned}
&\mathbb{P}(X=1 \mid Y \in [0.6,0.6+\delta]) \\
&\quad = \frac{\mathbb{P}(X=1)\mathbb{P}(Y \in [0.6,0.6+\delta] \mid X=1)}{\mathbb{P}(X=0)\mathbb{P}(Y \in [0.6,0.6+\delta] \mid X=0) + \mathbb{P}(Y=1)\mathbb{P}(Y \in [0.6,0.6+\delta] \mid X=1)} \\
&\quad = \frac{\mathbb{P}(X=1) f_{Y|1}(0.6)}{\mathbb{P}(X=0) f_{Y|0}(0.6) + \mathbb{P}(X=1) f_{Y|1}(0.6)},
\end{aligned}$$

where $f_{Y|0}$ is the density of a Gaussian with mean 0 and variance 0.49, and $f_{Y|1}$ is the density of a Gaussian with mean 1 and variance 0.49. Although the expression above may look intimidating, *this is just Bayes rule where $\mathbb{P}(Y=0.6 \mid X=x)$ has been replaced with $f_{Y|x}(0.6)$*. The moral of the story is that conditioning in continuous probability seems strange at first, but it is essentially the same as conditioning in discrete probability, with densities taking the place of probability mass functions.

Now, filling in the probabilities, we have

$$\begin{aligned}
&\mathbb{P}(X=1 \mid Y=0.6) \\
&\quad = \frac{0.3 \cdot (2\pi)^{-1/2} \exp(-0.4^2/(2 \cdot 0.49))}{0.7 \cdot (2\pi)^{-1/2} \exp(-0.6^2/(2 \cdot 0.49)) + 0.3 \cdot (2\pi)^{-1/2} \exp(-0.4^2/(2 \cdot 0.49))} \approx 0.345.
\end{aligned}$$

See what happened here? Before, you thought $\mathbb{P}(X=1) = 0.3$. Observing $Y = 0.6$ gives you slightly more evidence in favor of your love interest loving you back, which increases your belief to $\mathbb{P}(X=1 \mid Y=0.6) = 0.345$.

(c) We are looking for

$$\mathbb{P}(X=1 \mid Y \in [y, y+\delta]) \ge \mathbb{P}(X=0 \mid Y \in [y, y+\delta])$$

which is equivalent to

$$\mathbb{P}(X=1 \mid Y \in [y, y+\delta]) \ge \frac{1}{2}.$$

Now, we can compute the LHS as in the previous part:

$$\begin{aligned}
\mathbb{P}(X=1 \mid Y \in [y,y+\delta]) &= \frac{\mathbb{P}(X=1) f_{Y|1}(y)}{\mathbb{P}(X=0) f_{Y|0}(y) + \mathbb{P}(X=1) f_{Y|1}(y)} \\
&= \frac{0.3\exp(-(1-y)^2/0.98)}{0.7\exp(-y^2/0.98) + 0.3\exp(-(1-y)^2/0.98)} \\
&= \frac{1}{1 + (0.7/0.3)\exp(((1-y)^2 - y^2)/0.98)}.
\end{aligned}$$

In order to make the RHS $\geq 1/2$, we need:

$$\begin{aligned}
\frac{0.7}{0.3}\exp\left(\frac{(1-y)^2-y^2}{0.98}\right) &\leq 1 \\
\exp\left(\frac{(1-y)^2-y^2}{0.98}\right) &\leq \frac{3}{7} \\
\frac{(1-y)^2-y^2}{0.98} &\leq \ln\frac{3}{7} \\
1-2y &\leq 0.98\ln\frac{3}{7}
\end{aligned}$$

which gives the condition

$$y \leq \frac{1}{2}\left(1-0.98\ln\frac{3}{7}\right) \approx 0.915.$$

So, the new rule is to assume that your love interest loves you back if and only if you observe a message which is ≥ 0.915.

(d) As in the first part,

$$\begin{aligned}
&\mathbb{P}(X=0)\mathbb{P}(Y\leq 0.915 \mid X=0)+\mathbb{P}(X=1)\mathbb{P}(Y>0.915 \mid X=1) \\
&= 0.7\mathbb{P}(\mathcal{N}(0,0.49)\leq 0.915)+0.3\mathbb{P}(\mathcal{N}(1,0.49)>0.915) \\
&= 0.7\mathbb{P}\left(\mathcal{N}(0,1)\leq\frac{0.915}{0.7}\right)+0.3\mathbb{P}\left(\mathcal{N}(0,1)>-\frac{0.085}{0.7}\right) \\
&= 0.7\Phi\left(\frac{0.915}{0.7}\right)+0.3\Phi\left(\frac{0.085}{0.7}\right)\approx 0.798.
\end{aligned}$$

As you can see, this strategy performs better than the first part.

4 Deriving Chebyshev's Inequality

Recall Markov's Inequality, which applies for non-negative X and $\alpha > 0$:

$$\mathbb{P}[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}$$

Use an appropriate substitution for X and α to derive Chebyshev's Inequality, where μ denotes the expected value of Y.

$$\mathbb{P}[|Y-\mu| \geq k] \leq \frac{\operatorname{var}(Y)}{k^2}$$

Solution:

Let $X = (Y-\mu)^2$. Note that this satisfies the criterion that X is non-negative. Let $\alpha = k^2$ for $k > 0$. Again, this satisfies the criterion that $\alpha > 0$. Note also that the event $|Y-\mu| \geq k$ is equivalent to the event $(Y-\mu)^2 \geq k^2$. Then

$$\mathbb{P}(|Y-\mu| \geq k) = \mathbb{P}\big((Y-\mu)^2 \geq k^2\big) = \mathbb{P}(X \geq \alpha) \leq \frac{\mathbb{E}(X)}{\alpha} = \frac{\mathbb{E}((Y-\mu)^2)}{k^2} = \frac{\mathrm{var}(Y)}{k^2}.$$

This is equivalent to Chebyshev's Inequality.

5 Easy A's

A friend tells you about a course called "Laziness in Modern Society" that requires almost no work. You hope to take this course next semester to give yourself a well-deserved break after mastering CS 70. At the first lecture, the professor announces that grades will depend only a midterm and a final. The midterm will consist of three questions, each worth 10 points, and the final will consist of four questions, also each worth 10 points. He will give an A to any student who gets at least 60 of the possible 70 points.

However, speaking with the professor in office hours you hear some very disturbing news. He tells you that, in the spirit of the class, the GSIs are very lazy, and to save time the grading will be done as follows. For each student's midterm, the GSIs will choose a real number randomly from a distribution with mean $\mu = 5$ and variance $\sigma^2 = 1$. They'll mark each of the three questions with that score. To grade the final, they'll again choose a random number from the same distribution, independent of the first number, and will mark all four questions with that score.

If you take the class, what will the mean and variance of your total class score be? Use Chebyshev's inequality to conclude that you have less than a 5% chance of getting an A.

Solution:

Let X be the total number of points you receive in the class. Then $X = X_m + X_f$ where X_m are the points you receive on the midterm and X_f are the points you receive on the final. Your midterm score is generated as $X_m = 3Y_m$, where the r.v. Y_m represents the real number that the GSI chose when grading your midterm. Similarly, $X_f = 4Y_f$. The problem statement tells us that Y_m has mean 5 and variance 1 and Y_f has mean 5 and variance 1, so $\mathbb{E}[Y_m] = \mathbb{E}[Y_f] = 5$ and $\mathrm{var}(Y_m) = \mathrm{var}(Y_f) = 1$. Thus,

$$\mathbb{E}[X] = \mathbb{E}[X_m] + \mathbb{E}[X_f] = 3\,\mathbb{E}[Y_m] + 4\,\mathbb{E}[Y_f] = 35,$$
$$\mathrm{var}(X) = \mathrm{var}(X_m) + \mathrm{var}(X_f) = 9\,\mathrm{var}(Y_m) + 16\,\mathrm{var}(Y_f) = 25.$$

Using Chebyshev's Inequality, we get

$$\mathbb{P}[X \geq 60] \leq \mathbb{P}[|X-35| \geq 25] \leq \frac{\mathrm{var}(X)}{25^2} = \frac{1}{25}.$$

Unfortunately, you have at most a 4% chance of getting an A. So, the answer is: your mean score will be 35, the variance will be 25, and yes, you can conclude that you have less than a 5% chance of getting an A.

Note that although we calculated a bound for $\mathbb{P}[|X-35| \geq 25]$, which is the probability that you will get 60 or above or 10 or below, we cannot divide by 2 to refine our bound unless the distribution is symmetric about its mean. In this case, the distribution is not symmetric.

6 Practical Confidence Intervals

(a) It's New Year's Eve, and you're re-evaluating your finances for the next year. Based on previous spending patterns, you know that you spend \$1500 per month on average, with a standard deviation of \$500, and each month's expenditure is independently and identically distributed. As a poor college student, you also don't have any income. How much should you have in your bank account if you don't want to go broke this year, with probability at least 95%?

(b) As a UC Berkeley CS student, you're always thinking about ways to become the next billionaire in Silicon Valley. After hours of brainstorming, you've finally cut your list of ideas down to 10, all of which you want to implement at the same time. A venture capitalist has agreed to back all 10 ideas, as long as your net return from implementing the ideas is positive with at least 95% probability.

Suppose that implementing an idea requires 50 thousand dollars, and your start-up then succeeds with probability p, generating 150 thousand dollars in revenue (for a net gain of 100 thousand dollars), or fails with probability $1-p$ (for a net loss of 50 thousand dollars). The success of each idea is independent of every other. What is the condition on p that you need to satisfy to secure the venture capitalist's funding?

(c) One of your start-ups uses error-correcting codes, which can recover the original message as long as at least 1000 packets are received (not erased). Each packet gets erased independently with probability 0.8. How many packets should you send such that you can recover the message with probability at least 99%?

Solution:

(a) Let T be the random variable representing the amount of money we spend in the year.

We have $T = \sum_{i=1}^{12} X_i$, where X_i represents the spending in the i-th month. So, $\mathbb{E}[T] = 12 \cdot \mathbb{E}[E_1] = 18000$.

And, since the X_is are independent, $\text{var}(T) = 12 \cdot \text{var}(X_1) = 12 \cdot 500^2 = 3{,}000{,}000$.

We want to have enough money in our bank account so that we don't finish the year in debt with 95% confidence. So, we want to keep some money ε more than the mean expenditure such that the probability of deviating above the mean by more than ε is less than 0.05.

Let's use Chebyshev's inequality here to express this.

$$\mathbb{P}(|T-\mathbb{E}[T]| \geq \varepsilon) \leq \frac{\text{var}(T)}{\varepsilon^2} \leq 0.05$$

This gives us $\varepsilon^2 \geq \frac{3,000,000}{0.05}$. So, $\varepsilon \geq 7746$. This means that we want to have a balance of $\geq \mathbb{E}[T] + \varepsilon = 25746$.

Observe that here, while we wanted to estimate $\mathbb{P}(T - \mathbb{E}[T] \geq \varepsilon)$, Chebyshev's inequality only gives us information about $\mathbb{P}(|T - \mathbb{E}[T]| \geq \varepsilon)$. But since

$$\mathbb{P}(|T - \mathbb{E}[T]| \geq \varepsilon) \geq \mathbb{P}(T - \mathbb{E}[T] \geq \varepsilon),$$

this is fine. We just get a more conservative estimate.

(b) For this question, to keep the numbers from exploding, let's work in thousands of dollars. Let X_i be the profit made from idea i, and T be the total profit made. We have $T = \sum_{i=1}^{10} X_i$.

Here, $\mathbb{E}[X_1] = 100p - 50(1-p) = 150p - 50$.

And $\text{var}(X_1) = 150^2 p(1-p)$ as the distribution of X_1 is a shifted and scaled Bernoulli distribution. Using $\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2$ yields the same answer.

We have, $\mathbb{E}[T] = 10 \cdot \mathbb{E}[X_1]$. Similarly, $\text{var}(T) = 10 \cdot \text{var}(X_1)$.

Now, we want to bound the probability of T going below 0 by 0.05. In other words, we want $\mathbb{P}(T < 0) \leq 0.05$.

But, in order to apply Chebyshev's inequality, we need to look at deviation from the mean. We use the assumption that to get our funding we obviously need $\mathbb{E}[T] > 0$. Then:

$$\mathbb{P}(T < 0) \leq \mathbb{P}(T \leq 0 \cup T \geq 2\mathbb{E}[T]) = \mathbb{P}(|T - \mathbb{E}[T]| \geq \mathbb{E}[T]) \leq \frac{\text{var}(T)}{\mathbb{E}[T]^2} \leq 0.05$$

Looking at just the last inequality, we have:

$$\frac{\text{var}(T)}{\mathbb{E}[T]^2} = \frac{10 \cdot \text{var}(X_1)}{100 \cdot \mathbb{E}[X_1]^2} = \frac{\text{var}(X_1)}{10 \cdot \mathbb{E}[X_1]^2} \leq 0.05$$

$$\therefore \frac{\text{var}(X_1)}{\mathbb{E}[X_1]^2} \leq 0.5$$

Now, substituting what we have for variance and expectation, we get the following:

$$-22500p^2 + 22500p \leq 0.5(150p - 50)^2$$

which gives us the quadratic:

$$33750p^2 - 30000p + 1250 \geq 0$$

The solutions for p are $p \geq \frac{1}{9}(4 + \sqrt{13})$ and $p \leq \frac{1}{9}(4 - \sqrt{13})$. So $p \geq 0.845$ or ≤ 0.0438.

The relevant solution here is to pick $p \geq 0.845$, since the other solution yields negative expectation (contradicting the earlier assumption of positive expectation).

(c) We want $k = 1000$ packets to get across without being erased. Say we send n packets. Let X_i be the indicator random variable representing whether the ith packet got across or not.

Let the total number of unerased packets sent across be T. We have $T = \sum_{i=1}^{n} X_i$ and we want $T \geq 1000$.

We want $\mathbb{P}(T < 1000) \leq 0.01$. Now, let's try to get this in a form so that we can use Chebyshev's inequality. We know that $\mathbb{E}[T] > 1000$, so we can say that

$$\begin{aligned}\mathbb{P}(T < 1000) &\leq \mathbb{P}\big(T \leq 1000 \cup T \geq \mathbb{E}[T] + (\mathbb{E}[T] - 1000)\big) \\ &= \mathbb{P}\big(|T - \mathbb{E}[T]| \geq (\mathbb{E}[T] - 1000)\big) \leq \frac{\operatorname{var}(T)}{(\mathbb{E}[T] - 1000)^2} \leq 0.01.\end{aligned}$$

What is $\mathbb{E}[T]$? $\mathbb{E}[T] = n\,\mathbb{E}[X_1] = n(1-p) = 0.2n$.

Next, what is $\operatorname{var}(T)$? $\operatorname{var}(T) = n\operatorname{var}(X_1) = np(1-p) = 0.16n$.

Now, $\dfrac{\operatorname{var}(T)}{(\mathbb{E}[T]-k)^2} \leq 0.01 \implies 16n \leq (0.2n - 1000)^2$. This gives us the quadratic:

$$0.04n^2 - 416n + 1000000 \geq 0$$

Solving the last quadratic, we get $n \geq 6629$ or $n \leq 3774$. Since the second inequality doesn't make sense for our situation, our answer is $n \geq 6629$.