Score: ___________ **/ 90**

PSTAT 120A / **FINAL EXAM** / Summer 2022 Instructor: **Ethan Marzban**

NAME: ______________________________________ **Perm Number:** ______________________

SECTION (circle one): **3:30 - 4:20pm (Lucas)** **5 - 5:50pm (Moya)** **8 - 8:50am (Moya)**

Instructions:

- You will have **170 minutes** to complete this exam.
- You are allowed the use of two **8.5 $\times$ 11-inch** sheets, front and back, of notes. You are also permitted the use of **calculators**; the use of any and all other electronic devices (laptops, cell phones, etc.) is prohibited.
- Unless otherwise specified, simplification is not needed; however, all integrals and infinite sums (unless otherwise specified) must be evaluated.
  - One exception is that, whenever applicable, answers may be left in terms of Φ, the standard normal c.d.f..
- Good Luck!!!

Honor Code: In signing my name below, I certify that all work appearing on this exam is entirely my own and not copied from any external source. I further certify that I have not received any unauthorized aid while taking this exam.

$\times$ ____

Multiple Choice Questions:

Question:	1	2	3	Total
Points:	1	1	1	3
Score:				

Short-Answer Questions:

Question:	4	5	6	7	8	9	10	11	12	13	Total
Points:	7	9	11	11	10	9	5	11	9	5	87
Score:											

1 Multiple Choice Questions

Please fill in the bubble(s) ***on the exam below*** *corresponding to your answer. You do not need to submit any additional work for these questions.*

1. For $(X, Y) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ and a fixed constant $z \in [0,1]$, which of the following correctly provides $\mathbb{P}(X + Y \leq z)$? [1pts.]

   ○ $\displaystyle\int_0^z \int_0^{z-x} (1)\ \mathrm{d}x\ \mathrm{d}y$

   ○ $\displaystyle\int_0^1 \int_0^{z-x} (1)\ \mathrm{d}x\ \mathrm{d}y$

   ○ $\displaystyle\int_0^z \int_0^{z-y} (1)\ \mathrm{d}x\ \mathrm{d}y$

   ○ $\displaystyle\int_0^1 \int_0^{z-y} (1)\ \mathrm{d}x\ \mathrm{d}y$

   $\surd$ **None of the above**

   Solution: I admit this was a bit tricky; note that X and Y are i.i.d. $\mathcal{N}(0,1)$; i.e. they are **normally** distributed, not uniformly distributed. If instead $X, Y \overset{\text{i.i.d.}}{\sim} \text{Unif}[0,1]$ then we would have exactly the same setup as we had in Quiz 4, meaning answer choice 3 would be correct. But since $X, Y \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ the correct answer is "none of the above." (Because it was fairly tricky, I awarded half-credit for answer choice 3.)

2. Given a bivariate random vector (X, Y) with joint probability density function (p.d.f.) given by $f_{X,Y}(x,y)$, which of the following correctly computes $\mathbb{E}[\cos(X+Y)]$? (Only one answer choice is correct.) [1pts.]

   ○ $\displaystyle\int_{-\infty}^{\infty} \cos(x+y) f_{X,Y}(x,y)\ \mathrm{d}x$

   ○ $\displaystyle\iint_{\mathbb{R}^2} \cos(x+y) f_X(x) f_Y(y)\ \mathrm{d}A$

   ○ $\displaystyle\iint_{\mathbb{R}^2} \cos(x+y) f_{X,Y}(x, z-x)\ \mathrm{d}A$

   $\surd$ $\displaystyle\iint_{\mathbb{R}^2} \cos(x+y) f_{X,Y}(x,y)\ \mathrm{d}A$

   ○ None of the above

   Solution: This is simply an application of the multivariate LOTUS.

3. A hand of 5 cards is dealt without replacement from a standard deck of 52 cards. What is the probability that the hand contains exactly 2 pairs (by "pair" we mean "matching in denomination, not suit.")? [1pts.]

√ $$\frac{\binom{13}{2}\binom{4}{2}^2\binom{11}{1}\binom{4}{1}}{\binom{52}{5}}$$

○ $$\frac{\binom{13}{2}\binom{4}{2}^2\binom{11}{1}\binom{4}{1}}{(52)_5}$$

○ $$\frac{\binom{13}{4}\binom{4}{1}^4\binom{12}{1}\cdot(5!)}{\binom{52}{5}}$$

○ $$\frac{\binom{13}{4}\binom{4}{1}^4\binom{12}{1}\cdot(5!)}{(52)_5}$$

○ None of the above

Solution:

- Pick the 2 denominations to be in the two pairs: $\binom{13}{2}$
- Pick 2 suits (from a total 4) to be in the first pair; $\binom{4}{2}$. Pick another 2 (from a total 4) to be in the second pair; $\binom{4}{2}$.
- From the remaining $13 - 2 = 11$ denominations, pick 1 to be the rank of the non-paired card: $\binom{11}{1}$
- Finally, pick 1 suit (from a total 4) to be the non-paired card: $\binom{4}{1}$
- Since all hands are assumed to be equally likely, we divide by the total number of 5-card hands: $\binom{52}{5}$.

2 Short Answer Questions

Please mark your final answers in the spaces provided below each question. ***Be sure to show all of your work!***

4. The archipelago of *Gauchonia* is a collection of 12 smaller islands and one "Main Island," on which the capital lies. A ferry exists to help locals commute from island to island. On one particular trip, 10 people board the ferry at the capital on the Main Island, and then request to disembark on a random island, independently of all other passengers. Assume the following:
   - Nobody disembarks at the Main Island
   - Nobody can stay on the ferry forever (i.e. everyone disembarks at *some* island)
   - The ferry only stops at an island if someone is disembarking at that island.

   Let X denote the number of islands at which the ferry makes a stop.

   (a) What is $\mathbb{P}(X = 12)$? [2pts.]

   Solution: Since there are only 10 passengers, we know that $S_X = \{1, \cdots, 10\}$ Therefore, $\mathbb{P}(X = 12) = 0$.

   (b) Compute $\mathbb{E}[X]$. As a hint: Elevator Problem. [5pts.]

   Solution: The idea was to solve this exactly like the Elevator problem. Define indicators

$$\mathbb{1}_j := \begin{cases} 1 & \text{if ferry stops at } j^{\text{th}} \text{ island} \\ 0 & \text{otherwise} \end{cases}$$

   so that

$$X = \sum_{j=1}^{12} \mathbb{1}_j$$

   and, by linearity of expectation,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{j=1}^{12} \mathbb{1}_j\right] = \sum_{j=1}^{12} \mathbb{E}[\mathbb{1}_j] = \sum_{j=1}^{12} \mathbb{P}(\text{ferry stops at } j^{\text{th}} \text{ island})$$

   Furthermore,

$$\begin{aligned} \mathbb{P}(\text{ferry stops at } j^{\text{th}} \text{ island}) &= 1 - \mathbb{P}(\text{nobody gets off on } j^{\text{th}} \text{ island}) \\ &= 1 - \mathbb{P}(\text{all 10 people get off one of the other 11 islands}) \\ &= 1 - \left(\frac{11}{12}\right)^{10} \end{aligned}$$

   meaning

$$\mathbb{E}[X] = 12\left[1 - \left(\frac{11}{12}\right)^{10}\right]$$

A Note On Grading: I admit that there was actually an error in this problem. My intent was to have this problem be a direct copy of the Elevator Problem; however, it turns out that in the Elevator Problem we did **not** make the assumption that "nobody can stay on the elevator forever." What I'm getting at is this: it is actually not quite correct to write X as a sum of indicators in this problem, because doing so includes 0 in the state space of X whereas the assumption "nobody can stay on the ferry forever" definitionally removes 0 from the state space of X.

For this reason, I was extremely generous in grading this problem. If you did things like in the solutions above, I gave the full 5 points. If you set up the correct indicators, you automatically got 4 points. (The more detailed rubric will become available once grades become available in Gradescope)

I was, however, still somewhat strict about not awarding points when people tried to recognize the distribution of X, as X does not follow any of the distributions we covered in this class.

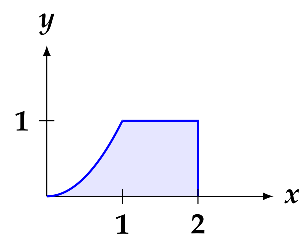
- It is not binomial because trials are not independent (once people disembark the ferry they cannot re-board)
- It is not hypergeometric, because we do not have an *a priori* division of our population into two distinct categories
- It is not discrete uniform, because though we assume a <u>fixed</u>, specific person has an equal chance of disembarking at any one of the islands, it is not the case that the events $\{$ferry stops at 2 islands$\}$ and $\{$ferry stops at 3 islands$\}$ have the same probabilities [which they would, if X were in fact distributed according to a discrete uniform distribution]

5. Let (X, Y) be a bivariate random vector with joint probability density function (p.d.f.) given by

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{6}{5} \cdot xy & \text{if } \sqrt{y} \le x \le 2,\ 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that $f_{X,Y}(x,y)$ is a valid joint p.d.f. **Be sure to show ALL of your steps fully;** large lapses in logic will result in point penalties. [4pts.]

Solution: It will be useful to first sketch the support:



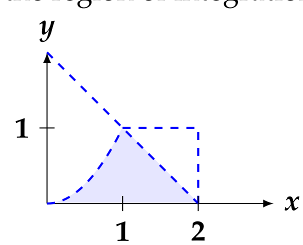
We can now see that integrating in the order dx dy will be easiest:

$$\begin{aligned}\iint_{\mathbb{R}^2} f_{X,Y}(x,y)\, \mathrm{d}A &= \int_0^1 \int_{\sqrt{y}}^2 \frac{6}{5} \cdot xy \, \mathrm{d}x\, \mathrm{d}y \\ &= \frac{6}{5} \int_0^1 y \cdot \frac{1}{2}(4-y)\, \mathrm{d}y \\ &= \frac{3}{5} \int_0^1 (4y - y^2)\, \mathrm{d}y = \frac{3}{5}\left[2 - \frac{1}{3}\right] = \frac{3}{5} \cdot \frac{5}{3} = 1 \checkmark\end{aligned}$$

Hence, $f_{X,Y}(x,y)$ integrates to unity. Additionally, we can see that both x and y are constraing to be (at least) nonnegative, meaning $xy > 0$ for all $(x,y) \in \mathbb{R}^2$ which completes the proof that $f_{X,Y}(x,y)$ is a valid joint p.d.f.

(b) **Set up, but DO NOT EVALUATE** a double integral (or set of double integrals) corresponding to $\mathbb{P}(X + Y \le 2)$. [3pts.]

Solution: Let's sketch the region of integration:



Once again, dx dy will be easiest:

$$\mathbb{P}(X + Y \le 2) = \int_0^1 \int_{\sqrt{y}}^{2-y} \frac{6}{5} \cdot xy \, \mathrm{d}x\, \mathrm{d}y$$

(c) Are X and Y independent? Justify your answer. [2pts.]

> **Solution:** There are two possible ways to answer this question. The first is to simply note that, from part (a), the support of (X, Y) is non-rectangular meaning $\boxed{X \text{ and } Y \text{ are NOT independent}}$. The other would be to note that $f_X(x) \cdot f_Y(y) \neq f_{X,Y}(x,y)$.

6. A point is selected in the first quadrant such that the $x-$ and $y-$ coordinates of the point are independent and both follow the distribution with probability density function (p.d.f.) given by

$$f(t) = \begin{cases} 2\lambda t e^{-\lambda t^2} & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\lambda > 0$ is a constant.

(a) Show that $f(t)$ is a valid p.d.f. [3pts.]

> **Solution:** First, we check nonnegativity: $e^{-y^2} > 0$ for all y, and when $s \geq 0$ we clearly have $s > 0$. Therefore, $f(s)$ is in fact nonnegative over the entire real line.
>
> Now, we also require that the function integrates to unity: as such, we compute
>
> $$\begin{aligned} \int_{-\infty}^{\infty} f(s)\ \mathrm{d}s &= \int_0^{\infty} 2\lambda s e^{-\lambda s^2}\ \mathrm{d}s \\ &= \lambda \int_0^{\infty} e^{-\lambda u}\ \mathrm{d}u = \frac{\lambda}{\lambda} = 1 \checkmark \end{aligned}$$
>
> where we have made the $u-$substitution $u = s^2$ so that $\mathrm{d}u = 2s\ \mathrm{d}s$. Therefore, $f(s)$ is in fact a valid p.d.f.

(b) Identify the distribution (by name!!) of X^2 (remember that X, the $x-$coordinate of the point, follows the distribution with p.d.f. given above). **Include any/all relevant parameter(s).** [5pts.]

Solution: We can use either the Change of Variable formula, or the CDF method.

....

CDF Method: set $W := X^2$ so that

$$\begin{aligned} F_W(w) &= \mathbb{P}(W \le w) = \mathbb{P}(X^2 \le w) = \mathbb{P}(-\sqrt{w} \le X \le \sqrt{w}) \\ &= F_X(\sqrt{w}) - F_X(-\sqrt{w}) \end{aligned}$$

$$\begin{aligned} f_W(w) &= \frac{\mathrm{d}}{\mathrm{d}w} F_W(w) = \frac{1}{2\sqrt{y}} f_X(\sqrt{w}) + \cancelto{0}{\frac{1}{2\sqrt{w}} f_X(-\sqrt{w})} = \frac{1}{2\sqrt{w}} f(\sqrt{w}) \\ &= \frac{1}{2\sqrt{w}} \cdot 2\lambda(\sqrt{w}) \cdot e^{-\lambda(\sqrt{w})^2} = \lambda e^{-\lambda w} \end{aligned}$$

which, coupled with the fact that $S_W = [0, \infty)$, allows us to conclude $(X^2) \sim \text{Exp}(\lambda)$.

....

Change of Variable: Note that $S_X = [0, \infty)$ and so the function $g(t) = t^2$ is in fact invertible over S_X, meaning we do not need to split into cases. Firstly we identify the state space of $W := X^2$ as $S_W = g(S_X) = [0, \infty)$; additionally, $g^{-1}(w) = \sqrt{w}$ so, for a fixed $w \in S_W$

$$\left| \frac{\mathrm{d}}{\mathrm{d}w} g^{-1}(w) \right| = \left| \frac{\mathrm{d}}{\mathrm{d}w} (\sqrt{w}) \right| = \left| \frac{1}{2\sqrt{w}} \right| = \frac{1}{2\sqrt{w}}$$

where we have eliminated the absolute values since $w > 0$. Therefore, by the Change of Variable formula,

$$\begin{aligned} f_W(w) &= f_X[g^{-1}(w)] \cdot \left| \frac{\mathrm{d}}{\mathrm{d}w} g^{-1}(w) \right| \\ &= \cancel{2}\lambda \cancel{(\sqrt{w})} e^{-\lambda(\sqrt{w})^2} \cdot \mathbb{1}_{\{(\sqrt{w}) \ge 0\}} \cdot \frac{1}{\cancel{2}\cancel{\sqrt{w}}} \\ &= \lambda e^{-\lambda w} \cdot \mathbb{1}_{\{w \ge 0\}} \end{aligned}$$

which, again, allows us to conclude $W := (X^2) \sim \text{Exp}(\lambda)$.

(c) Identify the distribution (by name!!) of $X^2 + Y^2$. **Include any/all relevant parameter(s).** As a reminder, you can use previously-derived results without proof so long as you cite them. [3pts.]

Solution: In part (a), we saw that $(X^2) \sim \text{Exp}(\lambda)$; since Y follows the same distribution as X, we also have $(Y^2) \sim \text{Exp}(\lambda)$. Additionally, functions of independent random variables are independent, meaning $(X^2) \perp (Y^2)$ and hence $(X^2) + (Y^2)$ is really the sum of two i.i.d. $\text{Exp}(\lambda)$ distribution, which we know results in the $\boxed{\text{Gamma}(2, \lambda)}$ distribution.

. . . .

In case you didn't remember that the sum of two independent $\text{Exp}(\lambda)$ random variables is distributed according to a Gamma distribution, you could have verified this using MGF's:

$$\begin{aligned} M_{X^2+Y^2}(t) &= M_{X^2}(t) \cdot M_{Y^2}(t) \\ &= \left(\begin{cases} \frac{\lambda}{\lambda - t} & \text{if } t < \lambda \\ \infty & \text{otherwise} \end{cases} \right) \cdot \left(\begin{cases} \frac{\lambda}{\lambda - t} & \text{if } t < \lambda \\ \infty & \text{otherwise} \end{cases} \right) \\ &= \begin{cases} \left(\frac{\lambda}{\lambda - t} \right)^2 & \text{if } t < \lambda \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

which shows that $(X^2 + Y^2) \sim \text{Gamma}(2, \lambda)$. [Note that we did still have to use the fact that functions of independent random variables are independent; i.e. that $X \perp Y \implies X^2 \perp Y^2$.]

. . . .

As an Aside: The distribution of X [i.e. the distribution whose p.d.f. is given by $f(t)$ as defined in the problem statement] is called the **Rayleigh Distribution**, which, perhaps surprisingly, actually arises in the context of MRI's (Magnetic Resonance Imaging): see **this** Wikipedia page for more information!

7. Let (X, Y) be a bivariate random vector with joint probability density function (p.d.f.) given by

$$f_{X,Y}(x, y) = \begin{cases} \lambda y e^{-y(x+\lambda)} & \text{if } x \geq 0,\ y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\lambda > 0$ is a fixed constant.

(a) Find $f_Y(y)$, the marginal p.d.f. of Y and use this to identify Y as belonging to a known distribution. **Be sure to include any/all relevant parameter(s)!** [4pts.]

> **Solution:** Integrating out X, we find
>
> $$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, \mathrm{d}x \\ &= \int_0^{\infty} \lambda y e^{-y(x+\lambda)}\, \mathrm{d}x \\ &= \lambda e^{-y\lambda} \cdot \int_0^{\infty} y e^{-yx}\, \mathrm{d}x = \lambda e^{-\lambda y} \end{aligned}$$
>
> which, noting that $S_Y = [0, \infty)$ allows us to write
>
> $$f_Y(y) = \lambda e^{-\lambda y} \cdot \mathbb{1}_{\{y \geq 0\}}$$
>
> which in turn implies $Y \sim \text{Exp}(\lambda)$.

(b) Find $f_{X|Y}(x \mid y)$, the density of $(X \mid Y = y)$, and use this to identify $(X \mid Y = y)$ as belonging to a known distribution. **Be sure to include any/all relevant parameter(s)!** [3pts.]

> **Solution:** We have
>
> $$\begin{aligned} f_{X|Y}(x \mid y) &= \frac{f_{X,Y}(x, y)}{f_Y(y)} \\ &= \frac{\lambda y e^{-y(x+\lambda)} \cdot \mathbb{1}_{\{x \geq 0,\ y \geq 0\}}}{\lambda e^{-\lambda y} \cdot \mathbb{1}_{\{y \geq 0\}}} \\ &= \frac{y e^{-yx} \cdot \cancel{\lambda e^{-\lambda y}} \cdot \mathbb{1}_{\{x \geq 0\}} \cdot \cancel{\mathbb{1}_{\{y \geq 0\}}}}{\cancel{\lambda e^{-\lambda y}} \cdot \cancel{\mathbb{1}_{\{y \geq 0\}}}} = y e^{-yx} \cdot \mathbb{1}_{\{x \geq 0\}} \end{aligned}$$
>
> which allows us to conclude
>
> $$(X \mid Y = y) \sim \text{Exp}(y)$$

(c) Set up **but do not evaluate** an integral corresponding to $\mathbb{E}[X]$, that only involves the marginal p.d.f. of Y. [4pts.]

Solution: By the Law of Iterated Expectations (i.e. Tower Property), we have

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]]$$

Since $(X \mid Y) \sim \text{Exp}(Y)$, we have

$$\mathbb{E}[X \mid Y] = \frac{1}{Y}$$

and hence

$$\mathbb{E}[X] = \mathbb{E}\left[\frac{1}{Y}\right] = \int_{-\infty}^{\infty} \frac{1}{y} f_Y(y)\ \mathrm{d}y = \int_0^{\infty} \lambda \cdot \frac{1}{y} e^{-\lambda y}\ \mathrm{d}y$$

where we have applied the LOTUS in the third equality from the left. (If you're curious, this integral actually diverges to infinity!)

8. In a class of highly motivated PSTAT 120A students, the average time taken for a randomly selected student to complete the final exam is 30 minutes.

(a) Use Markov's Inequality to bound the probability that a randomly selected student will complete the exam in under 31 minutes. **Be sure to clearly state whether this is an upper or lower bound!** [3pts.]

Solution: Let X denote the time a randomly selected PSTAT 120A student takes to complete the final; then $\mathbb{E}[X] = 30$. Additionally, we seek to bound $\mathbb{P}(X \leq 31)$; by Markov's Inequality, we have that

$$\mathbb{P}(X \geq 31) \leq \frac{\mathbb{E}[X]}{31} = \frac{30}{31}$$

meaning

$$\mathbb{P}(X < 31) = 1 - \mathbb{P}(X \geq 31) \geq 1 - \frac{30}{31} = \frac{1}{31}$$

which is clearly a lower bound.

(b) Suppose that it is also known that the standard deviation of the time a randomly selected students takes to complete the exam is 5 minutes. Use Chebyshev's Inequality to bound the probability that a randomly selected student will complete the exam in under 31 minutes. **Be sure to clearly state whether this is an upper or lower bound!** [3pts.]

Solution: Now that we have access to the variance of X, we can try applying Chebyshev's Inequality. First, note that

$$\begin{aligned}\mathbb{P}(X < 31) &= \mathbb{P}(X - 30 < 31 - 30) = \mathbb{P}(X - 30 < 1) \\ &\geq \mathbb{P}(|X - 30| < 1)\end{aligned}$$

Now, a direct application of Chebyshev's Inequality tells us

$$\mathbb{P}(|X - 30| \geq 1) \leq \frac{\mathrm{Var}(X)}{1^2} = \frac{25}{1}$$

This means that

$$\mathbb{P}(|X - 30| < 10) \geq 1 - \frac{25}{1} = -24$$

and so, combining this with our previous work, we see that

$$\mathbb{P}(X < 31) \geq \mathbb{P}(|X - 30| < 10) \geq -24$$

Now, this is a completely useless lower bound; thus, if this were a "real-world" setting we would report a lower bound of 0. However, for the purposes of this exam, -24 was also accepted as a valid correct answer.

(c) Now, suppose 64 students are sampled at random (and with replacement), and the average time to complete the final exam among these 64 students is recorded. Approximately what is the probability that the average completion time was lower than 31 minutes? [4pts.]

Solution: Let X_i denote the time of the i^{th} student, and let

$$\bar{X}_{64} := \frac{1}{64}\sum_{i=1}^{64} X_i$$

denote the average completion time of the sample of $n = 64$ students. With this notation, we seek $\mathbb{P}(\bar{X}_{64} < 31)$. Additionally, $\mathbb{E}[\bar{X}_{64}] = \mathbb{E}[X_1] = 30$ and $\mathrm{Var}(\bar{X}_{64}) = \mathrm{Var}(X_1)/64 = 25/64$. By the Central Limit Theorem [which we can invoke, since $n = 64 \geq 25$],

$$\left(\frac{\bar{X}_{64} - 30}{\sqrt{25/64}}\right) \overset{d}{\approx} \mathcal{N}(0,\ 1)$$

meaning

$$\begin{aligned}\mathbb{P}(\bar{X}_{64} < 31) &= \mathbb{P}\left(\frac{\bar{X}_{64} - 30}{\sqrt{25/64}} < \frac{31 - 30}{\sqrt{25/64}}\right) \\ &\approx \Phi\left(\frac{1}{5/8}\right) = \Phi(1.6) \approx 94.5\%\end{aligned}$$

9. Doctor Strange is travelling through the Multiverse, visiting parallel universes one by one. There is, however, a 25% chance that any one of his trips will cause an Incursion, independently of all of his other trips.

   (a) What is the probability that exactly 3 of the first 5 trips Doctor Strange makes result in Incursions? [3pts.]

   Solution: Let X denote the number of trips, out of 5, that result in an incursion; then $X \sim \mathrm{Bin}(5, 0.25)$ and

   $$\mathbb{P}(X = 3) = \binom{5}{3}(0.25)^3(0.75)^2$$

   (b) What is the probability that the 7th trip Doctor Strange makes is the first trip that results in an Incursion? [3pts.]

   Solution: Let Y denote the number of trips, including the final trip, needed before Doctor Strange causes his first incursion; then $Y \sim \mathrm{Geom}(0.25)$ and

   $$\mathbb{P}(Y = 7) = (0.75)^6(0.25)$$

(c) What is the expected number of trips Doctor Strange must make before causing his third Incursion? [3pts.]

> **Solution:** Let Z denote the number of trips, including the final trip, needed before Doctor Strange causes his third Incursion; then $Z \sim \text{NegBin}(3, 0.25)$ and
>
> $$\mathbb{E}[Z] = \frac{3}{0.25} = \boxed{12 \text{ trips}}$$

10. At a boutique store, 60% of items are *Gaucci* brand. Additionally, 60% of all items sold at this boutique are found to be counterfeits; 50% of these counterfeit items are *Gaucci* brand. Suppose you purchase an item that is *Gaucci* brand- what is the probability that your item is counterfeit? [5pts.]

> **Solution:** We define the following events:
>
> $$G = \text{"a randomly selected item was } \textit{Gaucci} \text{ brand"}$$
> $$C = \text{"a randomly selected item was counterfeit"}$$
>
> Then, from the problem statement, we have
>
> $$\mathbb{P}(G) = 0.6; \quad \mathbb{P}(C) = 0.6; \quad \mathbb{P}(G \mid C) = 0.5$$
>
> We seek $\mathbb{P}(C \mid G)$, which can be computed using Bayes' Theorem:
>
> $$\mathbb{P}(C \mid G) = \frac{\mathbb{P}(G \mid C)\mathbb{P}(C)}{\mathbb{P}(G)} = \frac{0.5 \cdot 0.6}{0.6} = \boxed{0.5 = 50\%}$$

11. Suppose (X, Y) is a discrete bivariate random vector with joint probability mass function (p.m.f.) given by

		Y		
		1	2	3
X	0	0.1	0.2	0.1
	1	0	0.2	0
	2	0.4	0	0

Furthermore, define events A, B, C in the following manner (recall that 0 is an even number):

$$A = \{X \text{ is even}\}$$
$$B = \{Y \text{ is odd}\}$$
$$C = \{Y \text{ is a multiple of } 3\}$$

(a) Compute $\mathbb{P}(A)$ [3pts.]

Solution:

$$\mathbb{P}(A) = p_X(0) + p_X(2) = (0.1 + 0.2 + 0.1) + (0.4) = 0.8$$

(b) Compute $\mathbb{P}(A \cup B)$ [4pts.]

Solution: By the Addition Rule,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

We computed $\mathbb{P}(A)$ in part (a) above; now all that remains is to compute the remaining quantities:

$$\begin{aligned} \mathbb{P}(B) &= p_Y(1) + p_Y(3) = (0.1 + 0.4) + 0.1 = 0.6 \\ \mathbb{P}(A \cap B) &= p_{X,Y}(0,1) + p_{X,Y}(0,3) + p_{X,Y}(2,1) + p_{X,Y}(2,3) \\ &= 0.1 + 0.4 + 0.1 = 0.6 \\ \mathbb{P}(A \cup B) &= 0.8 + 0.6 - 0.6 = 0.8 \end{aligned}$$

(c) Are A, B, C mutually independent? Justify your answer mathematically, using the definition of mutual independence. [4pts.]

Solution: We can actually already conclude that A, B, C are *not* mutually independent by noting, using our work from part (b),

$$\mathbb{P}(A \cap B) = 0.6 \neq 0.8 \cdot 0.6 = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

Therefore, the three events are not even pairwise independent meaning they are not mutually independent.

12. The speed of a randomly selected car travelling along I-5 is a random variable X with MGF (Moment-Generating Function) given by

$$M_X(t) = \exp\left\{70t + \frac{40}{2}t^2\right\}; \quad \forall t \in \mathbb{R}$$

Additionally, the posted speed limit is 65mph.

(a) What is the average speed of a randomly selected car travelling along I-5? [2pts.]

Solution: We recognize the form of the MGF provided as that associated with a normally distributed random variable; specifically, $X \sim \mathcal{N}(70, 40)$ meaning $\mathbb{E}[X] = 70$.

Alternatively, we could have differentiated the MGF and set $t = 0$:

$$M_X'(t) = (70 + 40t)\exp\left\{70t + \frac{40}{2}t^2\right\}$$
$$M_X'(0) = (70 + 0)e^0 = 70 = \mathbb{E}[X]$$

(b) What is the probability that a randomly selected car will be speeding? (Here, "speeding" means "travelling faster than the speed limit") [3pts.]

Solution: Here we must recognize the distribution of X. We seek $\mathbb{P}(X > 65)$, which we can compute using standardization:

$$\begin{aligned}
\mathbb{P}(X > 65) = 1 - \mathbb{P}(X \leq 65) &= 1 - \mathbb{P}\left(\frac{X-70}{\sqrt{40}} \leq \frac{65-70}{\sqrt{40}}\right) \\
&= 1 - \Phi\left(-\frac{5}{\sqrt{40}}\right) = \Phi\left(\frac{5}{\sqrt{40}}\right) = \Phi\left(\frac{5}{2\sqrt{10}}\right)
\end{aligned}$$

(c) Given that a car was speeding, what is the chance that it was travelling slower than 72 mph? [4pts.]

Solution: We seek $\mathbb{P}(X < 72 \mid X > 65)$. We compute this by first using the definition of conditional probability:

$$\begin{aligned}
\mathbb{P}(X < 72 \mid X > 65) &= \frac{\mathbb{P}(\{X < 72\} \cap \{X > 65\})}{\mathbb{P}(X > 65)} \\
&= \frac{\mathbb{P}(65 < X < 72)}{\mathbb{P}(X > 65)} \\
&= \frac{\mathbb{P}\left(\frac{65-70}{\sqrt{40}} < \frac{X-70}{\sqrt{40}} < \frac{72-70}{\sqrt{40}}\right)}{\Phi\left(\frac{5}{2\sqrt{10}}\right)} \\
&= \frac{\mathbb{P}\left(\frac{X-70}{\sqrt{40}} < \frac{72-70}{\sqrt{40}}\right) - \mathbb{P}\left(\frac{X-70}{\sqrt{40}} < \frac{65-70}{\sqrt{40}}\right)}{\Phi\left(\frac{5}{2\sqrt{10}}\right)} \\
&= \frac{\Phi\left(\frac{1}{\sqrt{10}}\right) - \Phi\left(\frac{5}{2\sqrt{10}}\right)}{\Phi\left(\frac{5}{2\sqrt{10}}\right)}
\end{aligned}$$

13. Consider a random variable X with mean μ and variance σ^2. If f is an appropriately differentiable function, use a second-order Taylor Series Expansion (i.e. a Taylor Series Expansion taken out to the third term in the sum) to show that [5pts.]

$$\mathbb{E}[f(X)] \approx f(\mu) + \frac{\sigma^2}{2} f''(\mu)$$

Solution: Taylor Expanding $f(X)$ about μ yields

$$f(X) = f(\mu) + (X - \mu)f'(\mu) + \frac{1}{2}(X - \mu)^2 f''(\mu) + \cdots$$

Taking the expectation of both sides, and disregarding the higher-order terms (as the problem suggests), we find

$$\begin{aligned}
\mathbb{E}[f(X)] &\approx \mathbb{E}\left[f(\mu) + (X - \mu)f'(\mu) + \frac{1}{2}(X - \mu)^2 f''(\mu)\right] \\
&= f(\mu) + \cancelto{0}{(\mathbb{E}[X] - \mu)f'(\mu)} + \frac{1}{2}\mathbb{E}[(X - \mu)^2] f''(\mu) \\
&= f(\mu) + \frac{1}{2}\sigma^2 \cdot f''(\mu)
\end{aligned}$$

where we have noted that $\mathbb{E}[(X - \mu)^2]$ is definitionally $\mathrm{Var}(X)$.

COMMON MGF's:

Distribution	MGF at t
$\text{Bin}(n,p)$	$(1-p+pe^t)^n, \quad \forall t \in \mathbb{R}$
$\text{Geom}(p)$	$\begin{cases} \dfrac{pe^t}{1-(1-p)e^t} & \text{if } t < -\ln(1-p) \\ \infty & \text{otherwise} \end{cases}$
$\text{NegBin}(r,p)$	$\begin{cases} \left(\dfrac{pe^t}{1-(1-p)e^t}\right)^r & \text{if } t < -\ln(1-p) \\ \infty & \text{otherwise} \end{cases}$
$\text{Pois}(\lambda)$	$e^{\lambda(e^t-1)}, \quad \forall t \in \mathbb{R}$

Distribution	MGF at t
$\text{Exp}(\lambda)$	$\begin{cases} \dfrac{\lambda}{\lambda - t} & \text{if } t < \lambda \\ 0 & \text{otherwise} \end{cases}$
$\text{Gamma}(r,\lambda)$	$\begin{cases} \left(\dfrac{\lambda}{\lambda - t}\right)^r & \text{if } t < \lambda \\ 0 & \text{otherwise} \end{cases}$
$\mathcal{N}(\mu,\sigma^2)$	$\exp\left\{\mu t + \dfrac{\sigma^2}{2}t^2\right\}; \quad \forall t \in \mathbb{R}$
$\text{Unif}[a,b]$	$\begin{cases} \dfrac{e^{tb}-e^{ta}}{t(b-a)} & \text{if } t \neq 0 \\ 1 & \text{if } t = 0 \end{cases}$

USEFUL RESULTS FROM MATHEMATICS:

(***Please note that these are by no means comprehensive; I expect you to know all of the sums from the Calculus Review Series, as well as common mathematical results.***)

- $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$
- $\cosh(x) := \dfrac{e^x + e^{-x}}{2}; \quad \sinh(x) := \dfrac{e^x - e^{-x}}{2}$
- $f(x) = \sum_{n=0}^{\infty} \dfrac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + (x-a)f'(a) + \dfrac{1}{2}(x-a)^2 f''(a) + \dfrac{1}{3!}(x-a)^3 f'''(a) + \cdots$

You may use this page for scratch work, if necessary.

You may use this page for scratch work, if necessary.

You may use this page for scratch work, if necessary.