Section 7: Solutions

Review of Main Concepts

- **Central Limit Theorem (CLT)**: Let $X_1, \ldots, X_n$ be iid random variables with $\mathbb{E}[X_i] = \mu$ and $Var(X_i) = \sigma^2$. Let $X = \sum_{i=1}^{n} X_i$, which has $\mathbb{E}[X] = n\mu$ and $Var(X) = n\sigma^2$. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, which has $\mathbb{E}\left[\overline{X}\right] = \mu$ and $Var(\overline{X}) = \frac{\sigma^2}{n}$. $\overline{X}$ is called the *sample mean*. Then, as $n \to \infty$, $\overline{X}$ approaches the normal distribution $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$. Standardizing, this is equivalent to $Y = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ approaching $\mathcal{N}(0,1)$. Similarly, as $n \to \infty$, X approaches $\mathcal{N}(n\mu, n\sigma^2)$ and $Y' = \frac{X-n\mu}{\sigma\sqrt{n}}$ approaches $\mathcal{N}(0,1)$.

  It is no surprise that $\overline{X}$ has mean μ and variance σ^2/n – this can be done with simple calculations. The importance of the CLT is that, for large n, regardless of what distribution X_i comes from, $\overline{X}$ is *approximately normally distributed with mean μ and variance σ^2/n*. Don't forget the continuity correction, only when $X_1, \ldots, X_n$ are discrete random variables.

- **Multivariate: Discrete to Continuous:**

	Discrete	**Continuous**
Joint PMF/PDF	$p_{X,Y}(x,y) = \mathbb{P}(X=x, Y=y)$	$f_{X,Y}(x,y) \neq \mathbb{P}(X=x, Y=y)$
Joint range/support $\Omega_{X,Y}$	$\{(x,y) \in \Omega_X \times \Omega_Y : p_{X,Y}(x,y) > 0\}$	$\{(x,y) \in \Omega_X \times \Omega_Y : f_{X,Y}(x,y) > 0\}$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \le x, s \le y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(t,s)\,ds dt$
Normalization	$\sum_{x,y} p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_y p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$
Expectation	$\mathbb{E}[g(X,Y)] = \sum_{x,y} g(x,y) p_{X,Y}(x,y)$	$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy$
Independence must have	$\forall x,y, p_{X,Y}(x,y) = p_X(x)p_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$	$\forall x,y, f_{X,Y}(x,y) = f_X(x)f_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$

- **Law of Total Probability (r.v. version):** If X is a discrete random variable, then

$$\mathbb{P}(A) = \sum_{x \in \Omega_X} \mathbb{P}(A|X=x) p_X(x) \qquad \text{discrete } X$$

- **Law of Total Expectation (Event Version)**: Let X be a discrete random variable, and let events $A_1, \ldots, A_n$ partition the sample space. Then,

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X|A_i]\,\mathbb{P}(A_i)$$

- **Conditional Expectation**: See table. Note that linearity of expectation still applies to conditional expectation: $\mathbb{E}[X+Y|A] = \mathbb{E}[X|A] + \mathbb{E}[Y|A]$

- **Law of Total Expectation (RV Version)**: Suppose X and Y are random variables. Then,

$$\mathbb{E}[X] = \sum_{y} \mathbb{E}[X|Y=y]\,p_Y(y) \qquad \text{discrete version.}$$

- **Conditional distributions**

	Discrete	**Continuous**
Conditional PMF/PDF	$p_{X\|Y}(x\|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$	$f_{X\|Y}(x\|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
Conditional Expectation	$\mathbb{E}[X\|Y=y] = \sum_x x p_{X\|Y}(x\|y)$	$\mathbb{E}[X\|Y=y] = \int_{-\infty}^{\infty} x f_{X\|Y}(x\|y) dx$

- **Continuous Law of Total Probability:**

$$\mathbb{P}(A) = \int_{x \in \Omega_X} \mathbb{P}(A|X=x) f_X(x) dx$$

- **Continuous Law of Total Expectation:**

$$\mathbb{E}[X] = \int_{y \in \Omega_Y} \mathbb{E}[X|Y = y] f_Y(y) dy$$

1. Content Review

(a) Select one: Suppose we have n independent and identically distributed random variables $X_1, X_2, \ldots, X_n$, each with mean μ and variance σ^2. Let $X = \sum_{i=1}^{n} X$. Then as n grows large, the Central Limit Theorem tells us that X behaves similarly to which normal distribution?

◯ $X \sim \mathcal{N}(n\mu, n\sigma^2)$

◯ $X \sim \mathcal{N}(\mu, n\sigma^2)$

◯ $X \sim \mathcal{N}(n\mu, \sigma^2)$

◯ $X \sim \mathcal{N}(n\mu, n^2\sigma^2)$

Solution:

The first one. By linearity of expectation, $\mathbb{E}[X] = n\mu$. Now since each of the rvs are independent, we may say that Var $(()\, X) = n\sigma^2$. Then as n grows large, X behaves similarly to a normal random variable with the same expectation and variance as itself.

(b) Suppose C and D are discrete random variables. Then $\mathbb{E}[C|D = d] =$

◯ $\sum_d dp_{D|C}(d|c)$

◯ $\sum_c cp_{C|D}(c|d)$

◯ $\int_{-\infty}^{\infty} cf_{c|d}dx$

◯ $\frac{\mathbb{E}[C]}{\mathbb{E}[D]}$

Solution:

Choice b is the correct answer from the definition of conditional expectation for discrete random variables.

(c) Suppose X and Y are random variables and A is an event. Given that $\mathbb{E}[X|A] = 4$ and $\mathbb{E}[Y|A] = 10$, what is $\mathbb{E}[2X + Y/2|A]$?

◯ 14

◯ 18

◯ 9

◯ 13

Solution:

Choice d is the correct answer since linearity of expectation still applies to conditional expectation:

$$\mathbb{E}[2X + Y/2|A] = \mathbb{E}[2X|A] + \mathbb{E}[Y/2|A] = 2\mathbb{E}[X|A] + \mathbb{E}[Y|A]/2 = 2 \cdot 4 + 10/2 = 8 + 5 = 13.$$

Central Limit Theorem Problems

The next few problems are CLT focused problems. Here's a general template for that! Sometimes we'll be trying to solve for the probability of something (e.g., $P(X \le 10)$, and sometimes, we'll be trying to find a value of some parameter that will allow for the probability to be in a certain range (e.g., $P(X \le 10) \le 0.2$). Regardless, we still will want to apply CLT on X, and follow the same process (the only difference is that we may be solving for different things).

(a) Setup the problem - write event you are interested in, in terms of sum of random variables. (what do we want to solve for/what is the probability we want to be true?)
  - Write the random variable we're interested in as a sum of i.i.d., random variables
  - Apply CLT to $X = X_1 + X_2 + ... + X_n$ (we can approximate X as a normal random variable $Y \sim N(\mu, \sigma^2)$)
  - Write the probability we're interested in

(b) *If the RVs are discrete*, apply continuity correction.

(c) Normalize RV to have mean 0 and standard deviation 1: $Z = \frac{Y-\mu}{\sigma}$

(d) Replace RV in probability expression with $Z \sim N(0,1)$

(e) Write in terms of $\Phi(z) = P(Z \le z)$

(f) Look up in the Phi table (or do a reverse Phi table lookup if we're looking for a value of z that gives us a certain probability)

2. Round off error

Let X be the sum of 100 real numbers, and let Y be the same sum, but with each number rounded to the nearest integer before summing. If the roundoff errors are independent and uniformly distributed between -0.5 and 0.5, what is the approximate probability that $|X - Y| > 3$? **Solution:**

Let $X = \sum_{i=1}^{100} X_i$, and $Y = \sum_{i=1}^{100} r(X_i)$, where $r(X_i)$ is X_i rounded to the nearest integer. Then, we have

$$X - Y = \sum_{i=1}^{100} X_i - r(X_i)$$

Note that each $X_i - r(X_i)$ is simply the round off error, which is distributed as $Unif(-0.5, 0.5)$. Since $X - Y$ is the sum of 100 i.i.d. random variables with mean $\mu = 0$ and variance $\sigma^2 = \frac{1}{12}$, $X - Y \approx W \sim \mathcal{N}(0, \frac{100}{12})$ by the Central Limit Theorem. For notational convenience let $Z \sim \mathcal{N}(0,1)$

$$\begin{aligned}
\mathbb{P}(|X-Y|>3) &\approx \mathbb{P}(|W|>3) && \text{[CLT]} \\
&= \mathbb{P}(W>3) + \mathbb{P}(W<-3) && \text{[No overlap between } W>3 \text{ and } W<-3\text{]} \\
&= 2\,\mathbb{P}(W>3) && \text{[Symmetry of normal]} \\
&= 2\,\mathbb{P}\left(\frac{W}{\sqrt{100/12}} > \frac{3}{\sqrt{100/12}}\right) \\
&\approx 2\,\mathbb{P}(Z>1.04) && \text{[Standardize } W\text{]} \\
&= 2\,(1-\Phi(1.04)) \approx 0.29834
\end{aligned}$$

3. Bad Computer

Each day, the probability your computer crashes is 10%, independent of every other day. Suppose we want to evaluate the computer's performance over the next 100 days.

(a) Let X be the number of crash-free days in the next 100 days. What distribution does X have? Identify $\mathbb{E}[X]$ and $Var(X)$ as well. Write an exact (possibly unsimplified) expression for $\mathbb{P}(X \ge 87)$. **Solution:**

Since X counts the number of crash-free days (successes) in 100 days (trials), where each trial is a success with probability 0.9, we can see that X is binomial with $n = 100$ and $p = 0.9$, or $X \sim \text{Binomial}(100, 0.9)$. Hence, $\mathbb{E}[X] = np = 90$ and $Var(X) = np(1-p) = 9$. Finally,

$$\mathbb{P}(X \geq 87) = \sum_{k=87}^{100} \binom{100}{k} (0.9)^k (1-0.9)^{100-k}$$

(b) Approximate the probability of at least 87 crash-free days out of the next 100 days using the Central Limit Theorem. Use continuity correction.

Important: continuity correction says that if we are using the normal distribution to approximate

$$\mathbb{P}(a \leq \sum_{i=1}^{n} X_i \leq b)$$

where $a \leq b$ are integers and the X_i's are i.i.d. **discrete** random variables, then, as our approximation, we should use

$$\mathbb{P}(a - 0.5 \leq Y \leq b + 0.5)$$

where Y is the appropriate normal distribution that $\sum_{i=1}^{n} X_i$ converges to by the Central Limit Theorem.[1]

For more details see pages 209-210 in the book. **Solution:**

From the previous part, we know that $\mathbb{E}[X] = 90$ and $\text{Var}(() X) = 9$.

$$\begin{aligned} \mathbb{P}(X \geq 87) &= \mathbb{P}(86.5 < X < 100.5) = \mathbb{P}(\frac{86.5 - 90}{3} < \frac{X - 90}{3} < \frac{100.5 - 90}{3}) \\ &\approx \mathbb{P}(-1.17 < \frac{X - 90}{3} < 3.5) \approx \Phi(3.5) + \Phi(1.17) - 1 \approx 0.9998 + 0.8790 - 1 = 0.8788 \end{aligned}$$

Notice that, if you had used $86.5 < X$ in place of $86.5 < X < 100.5$, your answer would have been nearly the same, because $\Phi(3.5)$ is so close to 1.

4. Tweets

A prolific twitter user tweets approximately 350 tweets per week. Let's assume for simplicity that the tweets are independent, and each consists of a uniformly random number of characters between 10 and 140. (Note that this is a discrete uniform distribution.) Thus, the central limit theorem (CLT) implies that the number of characters tweeted by this user is approximately normal with an appropriate mean and variance. Assuming this normal approximation is correct, estimate the probability that this user tweets between 26,000 and 27,000 characters in a particular week. (This is a case where continuity correction will make virtually no difference in the answer, but you should still use it to get into the practice!).

Solution:

Let X be the total number of characters tweeted by a twitter user in a week. Let $X_i \sim Unif(10, 140)$ be the number of characters in the ith tweet (since the start of the week). Since X is the sum of 350 i.i.d. rvs with

[1] The intuition here is that, to avoid a mismatch between discrete distributions (whose range is a set of integers) and continuous distributions, we get a better approximation by imagining that a discrete random variable, say W, is a continuous distribution with density function

$$f_W(x) := p_W(i) \quad \text{when } i - 0.5 \leq x < i + 0.5 \text{ and } i \text{ integer}$$

mean $\mu = 75$ and variance $\sigma^2 = 1430$, $X \approx N \sim \mathcal{N}(350 \cdot 75, 350 \cdot 1430)$. Thus,

$$\begin{aligned}\mathbb{P}(26,000 \leq X \leq 27,000) &= \mathbb{P}(25,999.5 \leq X \leq 27,000.5)\\ &\approx \mathbb{P}(25,999.5 \leq N \leq 27,000.5)\end{aligned}$$

Standardizing this gives the following formula

$$\begin{aligned}\mathbb{P}(25,999.5 \leq N \leq 27,000.5) &= \mathbb{P}\left(\frac{25,999.5 - 350 \cdot 75}{\sqrt{350 \cdot 1430}} \leq \frac{N - 350 \cdot 75}{\sqrt{350 \cdot 1430}} \leq \frac{27000.5 - 350 \cdot 75}{\sqrt{350 \cdot 1430}}\right)\\ &\approx \mathbb{P}\left(-0.35 \leq \frac{N - 350 \cdot 75}{\sqrt{350 \cdot 1430}} \leq 1.06\right)\\ &\approx \mathbb{P}\left(-0.35 \leq Z \leq 1.06\right)\\ &= \Phi(1.06) - \Phi(-0.35)\\ &\approx 0.85543 - (1 - 0.63683)\\ &= 0.49226\end{aligned}$$

So the probability that this user tweets between 26,000 and 27,000 characters in a particular week is approximately 0.4923.

5. Trinomial Distribution

A generalization of the Binomial model is when there is a sequence of n independent trials, but with three outcomes, where $\mathbb{P}(\text{outcome } i) = p_i$ for $i = 1, 2, 3$ and of course $p_1 + p_2 + p_3 = 1$. Let X_i be the number of times outcome i occurred for $i = 1, 2, 3$, where $X_1 + X_2 + X_3 = n$. Find the joint PMF $p_{X_1,X_2,X_3}(x_1, x_2, x_3)$ and specify its value for all $x_1, x_2, x_3 \in \mathbb{R}$. **Solution:**

We use a similar argument as for the binomial PMF. $\binom{n}{x_1,x_2,x_3}$ is the number of ways to select which of the n outcomes result in each of the 3 outcomes. Then, we multiply the probabilities of each trial being the corresponding outcome (e.g., $p_1^{x_1}$ is the probability that all x_1 trials end up being outcome 1). This gives use the following PMF:

$$p_{X_1,X_2,X_3}(x_1, x_2, x_3) = \binom{n}{x_1, x_2, x_3} \prod_{i=1}^{3} p_i^{x_i} = \frac{n!}{x_1! x_2! x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

where $x_1 + x_2 + x_3 = n$ and are nonnegative integers.

6. Do You “Urn” to Learn More About Probability?

Suppose that 3 balls are chosen without replacement from an urn consisting of 5 white and 8 red balls. Let $X_i = 1$ if the i-th ball selected is white and let it be equal to 0 otherwise. Give the joint probability mass function of

(a) X_1, X_2 **Solution:**

Here is one way of defining the joint pmf of X_1, X_2

$$\mathbb{P}(X_1 = 1, X_2 = 1) = \mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 1 \mid X_1 = 1) = \frac{5}{13} \cdot \frac{4}{12} = \frac{20}{156}$$

$$\mathbb{P}(X_1 = 1, X_2 = 0) = \mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 0 \mid X_1 = 1) = \frac{5}{13} \cdot \frac{8}{12} = \frac{40}{156}$$

$$\mathbb{P}(X_1 = 0, X_2 = 1) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 1 \mid X_1 = 0) = \frac{8}{13} \cdot \frac{5}{12} = \frac{40}{156}$$

$$\mathbb{P}(X_1 = 0, X_2 = 0) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 0 \mid X_1 = 0) = \frac{8}{13} \cdot \frac{7}{12} = \frac{56}{156}$$

(b) X_1, X_2, X_3 **Solution:**

Instead of listing out all the individual probabilities, we could write a more compact formula for the pmf. In this problem, the denominator is always $P(13, k)$, where k is the number of random variables in the joint pmf. And the numerator is $P(5, i)$ times $P(8, j)$ where i and j are the number of 1s and 0s, respectively.

If we wish to compute $p_{X_1,X_2,X_3}(x_1, x_2, x_3)$, then the number of 1s (i.e., white balls) is $x_1 + x_2 + x_3$, and the number of 0s (i.e., red balls) is $(1 - x_1) + (1 - x_2) + (1 - x_3)$. Then, we can write the pmf as follows:

$$p_{X_1,X_2,X_3}(x_1, x_2, x_3) = \frac{10!}{13!} \cdot \frac{5!}{(5 - x_1 - x_2 - x_3)!} \cdot \frac{8!}{(5 + x_1 + x_2 + x_3)!}$$

7. Successes

Consider a sequence of independent Bernoulli trials, each of which is a success with probability p. Let X_1 be the number of failures preceding the first success, and let X_2 be the number of failures between the first 2 successes. Find the joint pmf of X_1 and X_2. Write an expression for $E[\sqrt{X_1X_2}]$. You can leave your answer in the form of a sum. **Solution:**

X_1 and X_2 take on two particular values x_1 and x_2, when there are x_1 failures followed by one success, and then x_2 failures followed by one success. Since the Bernoulli trials are independent the joint pmf is

$$p_{X_1,X_2}(x_1, x_2) = (1 - p)^{x_1} p \cdot (1 - p)^{x_2} p = (1 - p)^{x_1+x_2} p^2$$

for $(x_1, x_2) \in \Omega_{X_1,X_2} = \{0, 1, 2, \dots\} \times \{0, 1, 2, \dots\}$. By the definition of expectation

$$E[\sqrt{X_1X_2}] = \sum_{(x_1,x_2)\in\Omega_{X_1,X_2}} \sqrt{x_1x_2} \cdot (1 - p)^{x_1+x_2} p^2.$$

8. Continuous joint density

The joint density of X and Y is given by

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

and the joint density of W and V is given by

$$f_{W,V}(w, v) = \begin{cases} 2 & 0 < w < v, 0 < v < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent? Are W and V independent?

Solution:

For two random variables X, Y to be independent, we must have $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ for all $x \in \Omega_X,\ y \in \Omega_Y$. Let's start with X and Y by finding their marginal PDFs. By definition, and using the fact that the joint PDF is 0 outside of $y > 0$, we get:

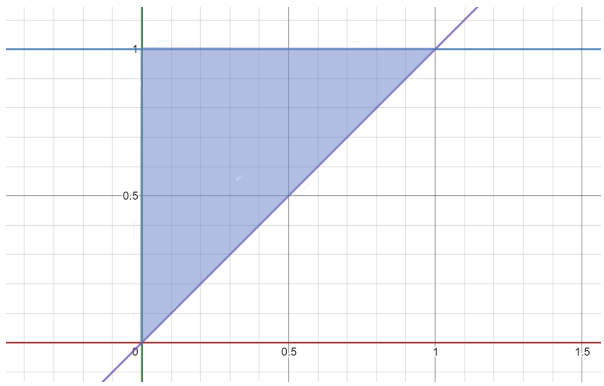
$$f_X(x) = \int_0^\infty xe^{-(x+y)} dy = e^{-x}x$$

We do the same to get the PDF of Y, again over the range $x > 0$:

$$f_Y(y) = \int_0^\infty xe^{-(x+y)}dx = e^{-y}$$

Since $e^{-x}x \cdot e^{-y} = xe^{-x-y} = xe^{-(x+y)}$ for all $x, y > 0$, X and Y are independent.

We can see that W and V are not independent simply by observing that $\Omega_W = (0, 1)$ and $\Omega_V = (0, 1)$, but $\Omega_{W,V}$ is not equal to their Cartesian product. Specifically, looking at their range of $f_{W,V}(w, v)$. Graphing it with w as the "x-axis" and v as the "y-axis", we see that :



The shaded area is where the joint pdf is strictly positive. Looking at it, we can see that it is not rectangular, and therefore it is not the case that $\Omega_{W,V} = \Omega_W \times \Omega_V$. Remember, the joint range being the Cartesian product of the marginal ranges is not sufficient for independence, but it is *necessary*. Therefore, this is enough to show that they are not independent.

9. Trapped Miner

A miner is trapped in a mine containing 3 doors.

- D_1: The 1^{st} door leads to a tunnel that will take him to safety after 3 hours.
- D_2: The 2^{nd} door leads to a tunnel that returns him to the mine after 5 hours.
- D_3: The 3^{rd} door leads to a tunnel that returns him to the mine after a number of hours that is Binomial with parameters $(12, \frac{1}{3})$.

At all times, he is equally likely to choose any one of the doors. What is the expected number of hours for this miner to reach safety?

Solution:

Let T = number of hours for the miner to reach safety. (T is a random variable)
Let D_i be the event the i^{th} door is chosen. $i \in \{1, 2, 3\}$. Finally, let T_3 be the time it takes to return to the mine in the third case only (a random variable). Note that the expectation of T_3 is $12 * \frac{1}{3}$ because it is binomially distributed with parameters $n = 12, p = \frac{1}{3}$. By Law of Total Expectation, linearity of expectation, and by applying

the conditional expectations given by the problem statement:

$$\begin{aligned}\mathbb{E}[T] &= \mathbb{E}[T|D_1]\,\mathbb{P}(D_1) + \mathbb{E}[T|D_2]\,\mathbb{P}(D_2) + \mathbb{E}[T|D_3]\,\mathbb{P}(D_3) \\ &= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (\mathbb{E}[T_3 + T]) \cdot \frac{1}{3} \\ &= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (\mathbb{E}[T_3] + \mathbb{E}[T]) \cdot \frac{1}{3} \\ &= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (4 + \mathbb{E}[T]) \cdot \frac{1}{3}\end{aligned}$$

Solving this equation for $\mathbb{E}[T]$, we get

$$\mathbb{E}[T] = 12$$

Therefore, the expected number of hours for this miner to reach safety is 12.

10. Lemonade Stand

Suppose I run a lemonade stand, which costs me \$100 a day to operate. I sell a drink of lemonade for \$20. Every person who walks by my stand either buys a drink or doesn't (no one buys more than one). If it is raining, n_1 people walk by my stand, and each buys a drink independently with probability p_1. If it isn't raining, n_2 people walk by my stand, and each buys a drink independently with probability p_2. It rains each day with probability p_3, independently of every other day. Let X be my profit over the next week. In terms of n_1, n_2, p_1, p_2 and p_3, what is $\mathbb{E}[X]$?

Solution:

Let R be the event it rains. Let X_i be how many drinks I sell on day i for $i = 1, ..., 7$. We are interested in $X = \sum_{i=1}^{7} (20X_i - 100)$. We have $X_i|R \sim \text{Binomial}(n_1, p_1)$, so $\mathbb{E}[X_i|R] = n_1p_1$. Similarly, $X_i|R^C \sim \text{Binomial}(n_2, p_2)$, so $\mathbb{E}\left[X_i|R^C\right] = n_2p_2$. By the law of total expectation,

$$\mu = \mathbb{E}[X_i] = \mathbb{E}[X_i|R]\,\mathbb{P}(R) + \mathbb{E}\left[X_i|R^C\right]\mathbb{P}(R^C) = n_1p_1p_3 + n_2p_2(1-p_3)$$

Hence, by linearity of expectation,

$$\begin{aligned}\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{7}(20X_i - 100)\right] &= 20\sum_{i=1}^{7}\mathbb{E}[X_i] - 700 = 140\mu - 700 \\ &= 140 \cdot (n_1p_1p_3 + n_2p_2(1-p_3)) - 700.\end{aligned}$$

11. 3 points on a line

Three points X_1, X_2, X_3 are selected at random on a line L (continuous independent uniform distributions). What is the probability that X_2 lies between X_1 and X_3? **Solution:**

Let $X_1, X_2, X_3 \sim Unif(0,1)$.

$$\begin{aligned}
\mathbb{P}(X_1 < X_2 < X_3) &= \int_{-\infty}^{\infty} \mathbb{P}(X_1 < X_2 < X_3 \mid X_2 = x)\ f_{X_2}(x)\ dx && \text{Continuous LoTP} \\
&= \int_{-\infty}^{\infty} \mathbb{P}(X_1 < x, X_3 > x)\ f_{X_2}(x)\ dx && \text{Independence of } X_1, X_2, X_3 \\
&= \int_{-\infty}^{\infty} \mathbb{P}(X_1 < x)\ \mathbb{P}(x < X_3)\ f_{X_2}(x)\ dx && \text{Independence of } X_1, X_3 \\
&= \int_{-\infty}^{\infty} F_{X_1}(x)\ (1 - F_{X_3}(x))\ f_{X_2}(x)\ dx \\
&= \int_0^1 x\ (1-x)\ 1\ dx \\
&= \frac{x^2}{2} - \frac{x^3}{3}\Big|_0^1 = \frac{1}{6}
\end{aligned}$$