

# ECE313 Final Review - Cramming Carnival Solutions

Author: Members of HKN

## 1 Dutiful Drills

All parts of this problem are unrelated to each other.

### 1.1 Solution:

- (a) **Suppose a random variable  $T$  has the exponential distribution with parameter  $\lambda$ . Suppose it is observed that  $T = t$  for some fixed value of  $t$ . Find the ML estimate,  $\hat{\lambda}_{ML}(t)$ , of  $\lambda$ , based on this observation.**

We know that  $T$  has a distribution  $\lambda e^{-\lambda t}$ . The estimate is simply the value of  $\lambda > 0$  that maximizes this distribution for  $t$  fixed. Therefore, we take the derivative with respect to  $\lambda$  and set it equal to 0.

$$\frac{d}{d\lambda} \lambda e^{-\lambda t} = e^{-\lambda t} - \lambda t e^{-\lambda t} = (1 - \lambda t) e^{-\lambda t} = 0$$

The exponential can never truly equal 0, so we can drop it. We are then left with  $\hat{\lambda}_{ML}(t) = \frac{1}{t}$ .

- (b) **Suppose it is assumed that  $X$  is drawn at random from the uniform distribution on the interval  $[0, b]$ , where  $b$  is some parameter. Find the ML estimator of  $b$  given  $X = u$  is observed.**

Recall that pdf of the uniform distribution is  $\frac{1}{b}$  for all values in the range. In order to maximize this value, we want to minimize  $b$ . However, we can't minimize  $b$  such that  $u$  can't even be observed.

Therefore, we should set  $\hat{b}_{ML}(u) = u$ .

- (c) **Consider a Poisson process on the interval  $[0, T]$  with rate  $\lambda > 0$ , and let  $0 < \tau < T$ . Define  $X_1$  to be the number of counts during  $[0, \tau]$ ,  $X_2$  to be the number of counts during  $[\tau, T]$ , and  $X$  to be the total number of counts during  $[0, T]$ . Let  $i, j, n$  be nonnegative integers such that  $n = i + j$ . Express the following probabilities in terms of  $n, i, j, \tau, T$ , and  $\lambda$ , simplifying your answers as much as possible:**

(i)  $P\{X = n\}$

$$P\{X = n\} = \frac{e^{-\lambda T} (\lambda T)^n}{n!}$$

(ii)  $P\{X_1 = i\}$

$$P\{X_1 = i\} = \frac{e^{-\lambda \tau} (\lambda \tau)^i}{i!}$$

(iii)  $P\{X_2 = j\}$

$$P\{X_2 = j\} = \frac{e^{-\lambda(T-\tau)} (\lambda(T-\tau))^j}{j!}$$

(iv)  $P\{X_1 = i | X = n\}$

Let's use the definition of conditional probability.

$$P\{X_1 = i | X = n\} = \frac{P\{X_1 = i, X = n\}}{P\{X = n\}}$$

Since we know that  $i + j = n$ , the above can be rewritten as

$$= \frac{P\{X_1 = i, X_2 = j\}}{P\{X = n\}}$$

Since  $X_1$  and  $X_2$  are independent, the numerator can be split into two probabilities.

$$= \frac{P\{X_1 = i\} P\{X_2 = j\}}{P\{X = n\}}$$

Now we can just evaluate and multiply. The exponential and the  $\lambda$ s drop off.

$$= \frac{n!}{i!j!} \left(\frac{\tau}{T}\right)^i \left(\frac{T-\tau}{T}\right)^j$$

Note that if we allow  $p = \frac{\tau}{T}$ , this can be written suggestively as

$$\binom{n}{i} p^i (1-p)^{n-i}$$

It's a binomial distribution! Cool.

(v)  $P\{X = n | X_1 = i\}$

Since  $n = i + j$ , the given can be rewritten to become

$$P\{X = n | X_1 = i\} = P\{X_2 = j | X_1 = i\}$$

$X_1$  and  $X_2$  are independent, so the conditional probability does nothing.

$$= P\{X_2 = j\} = \frac{e^{-\lambda(T-\tau)} (\lambda(T-\tau))^j}{j!}$$

## 2 Persistent Particles 1

Hyoun is playing with his quantum particle generator. Specifically, he is interested in the spins of the particles that he generates. A generated particle has a 50% chance of being spin-up and a 50% chance of being spin-down. All generated particles are independent from each other.

Hyoun decides that he will keep generating particles until the following conditions are met. For each of the following conditions, compute the expected number of particles that Hyoun needs to generate in order to meet the condition. As soon as Hyoun satisfies the condition, he will stop generating particles.

### 2.1 Solution:

- (a) **Hyoun gets at least one spin-up particle.**

This is a geometric distribution with  $p = \frac{1}{2}$ . The expected number of particles required is then  $\frac{1}{p} = \boxed{2}$ .

- (b) **Hyoun gets at least one spin-up and at least one spin-down particle.**

The first particle is guaranteed to be either spin-up or spin-down. It is then a geometric distribution with  $p = \frac{1}{2}$  in order to get the other particle. By linearity of expectation, the expected number of particles required is then  $1 + \frac{1}{p} = \boxed{3}$ .

- (c) **Hyoun gets at least two spin-up particles.**

This is the sum two independent geometric distributions with  $p = \frac{1}{2}$ . By linearity of expectation, the expected number of particles required is then  $\frac{1}{p} + \frac{1}{p} = \boxed{4}$ .

- (d) **Hyoun gets at least two spin-up particles in a row.**

Let  $X$  denote the number of flips required to get two spin-up particles in a row. We wish to find  $E[X]$ , which stands for the expected number of particles required to get two spin-up particles in a row. We arrive at the following expression  $E[X]$ :

$$E[X] = 1 + \frac{1}{2}E[X] + \frac{1}{2} \left( 1 + \frac{1}{2}E[X] \right)$$

Where did this expression come from? First, we have to generate a particle (this is where the first 1 comes from). If the particle generated was spin-down, which happens half the time, then we're back to where we started; i.e. we now need to generate two spin-up particles in a row once again. This is why we add  $\frac{1}{2}E[X]$ .

The second term occurs when we get spin-up. We then generate another particle (hence the 1). Half the time it will be spin-up as well, which means we are done. The other half of the time we generate a spin-down particle. This resets our whole problem, and we go back to the situation where we need to generate two spin-up particles in a row once again, which is why we add the  $\frac{1}{2}E[X]$ .

Now we have one equation and one unknown, so we can just solve for  $E[X]$ . This will give us  $\boxed{6}$ .

- (e) **Hyouin gets a spin-up particle followed immediately by a spin-down particle.**

The problem is completely unchanged if the word “immediately” is removed from the problem statement, since if a spin-up particle is not followed by a spin-down particle, then it must’ve been followed by a spin-up particle. Therefore, we first have to generate a spin-up particle, then we have to generate a spin-down particle.

But this is now the sum two independent geometric distributions with  $p = \frac{1}{2}$  once again. By linearity of expectation, the expected number of particles required is then  $\frac{1}{p} + \frac{1}{p} = \boxed{4}$ .

- (f) **Hyouin gets more spin-up particles than spin-down particles.**

Let  $X$  denote the number of flips required to satisfy this condition. The condition can be rewritten as “Hyouin gets one more spin-up particle than spin-down particles.” We are trying to find  $E[X]$ . We then have the following equation:

$$E[X] = 1 + \frac{1}{2}2E[X]$$

The first 1 comes from generating our first particle. If we get a spin-up particle, then we’re done. Half the time, we will not get a spin-up particle. This means we need to satisfy our condition just to get back up to even, then satisfy the condition again in order to finally end the experiment.

This equation has no solution. Therefore,  $E[X] = \boxed{\infty}$ .

If you don’t believe this, try setting up this experiment in Python.

- (g) **Hyouin gets the same number of spin-up particles as spin-down particles.**

This is satisfied at the very beginning when Hyouin has 0 spin-up particles and 0 spin-down particles. Therefore, the answer is  $\boxed{0}$ .

- (h) **Hyouin gets the same positive number of spin-up particles as spin-down particles.**

Suppose the first particle generated is spin-down. We are then back to part f’s condition. So the expected number of particles required is still  $\infty$ .

If the first particle generated is spin-up, then everything flips, but the answer remains the same.

Therefore, in all cases the expected number of particles required is  $\infty$ , so the overall expected number of particles required is  $\boxed{\infty}$ .

- (i) **Hyouin gets more than twice as many spin-up particles as spin-down particles.**

Hyouin has to get more spin-up particles than spin-down particles in order for this condition to be satisfied, so the expected number of particles required must be more than the amount required in part f. So the answer is still  $\boxed{\infty}$ .

### 3 Persistent Particles 2

Suppose that Hyouin generates two quantum particles. Let  $A$  denote the event that the first quantum particle was spin-up. Let  $B$  denote the event that the second quantum particle was spin-up. Let  $C$  denote the event that the quantum particles match each other.

Are  $A$ ,  $B$ , and  $C$  pairwise independent? Are they completely independent?

#### 3.1 Solution:

Event  $A$  happens with probability  $\frac{1}{2}$ . Event  $B$  happens with probability  $\frac{1}{2}$ . Event  $C$  happens with probability  $\frac{1}{2}$ .

Event  $A$  and event  $B$  both happening only occurs with probability  $\frac{1}{4}$  (both particles had to be spin-up). Therefore, events  $A$  and  $B$  are pairwise independent, since  $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ .

Event  $B$  and event  $C$  both happening only occurs with probability  $\frac{1}{4}$  (both particles had to be spin-up). Therefore, events  $B$  and  $C$  are pairwise independent, since  $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ .

Event  $A$  and event  $C$  both happening only occurs with probability  $\frac{1}{4}$  (both particles had to be spin-up). Therefore, events  $A$  and  $C$  are pairwise independent, since  $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ .

Since events  $A$  and  $B$ ,  $A$  and  $C$ , and  $B$  and  $C$  are all pairwise independent,  $A$ ,  $B$ , and  $C$  are pairwise independent.

The probability that all three events happen occurs with probability  $\frac{1}{4}$  (both particles had to be spin-up). This means that since  $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \neq \frac{1}{4}$ ,  $A$ ,  $B$ , and  $C$  are not completely independent.

### 4 Persistent Particles 3

Hyouin continues to generate his quantum particles. Suppose that he generates  $n$  of them. Let  $X_k$  be equal to  $\frac{1}{2}$  if the  $k$ 'th particle is spin-up, and 0 otherwise. Let  $Y_k$  be equal to  $\frac{1}{2}$  if the  $k$ 'th particle is spin-down, and 0 otherwise. Let  $Z_k = X_k^2 + 2Y_k$ .

Now let  $X$  be the sum of all  $X_k$ ,  $Y$  be the sum of all  $Y_k$ , and  $Z$  be the sum of all  $Z_k$ .

#### 4.1 Solution:

(a) Find  $E[X_1]$  and  $E[Z_1]$ .

We know that  $X_1 = \frac{1}{2}$  with probability  $\frac{1}{2}$  and  $X_1 = 0$  with probability  $\frac{1}{2}$ . Therefore,  $E[X_1] = \frac{1}{4}$ .

We know that  $Z_1 = \frac{1}{4}$  with probability  $\frac{1}{2}$  (spin-up case) and  $Z_1 = 1$  with probability  $\frac{1}{2}$  (spin-down case). Therefore,  $E[Z_1] = \frac{5}{8}$ .

(b) **Find  $E[Z]$  and  $\text{Var}(Z)$ .**

The generated particles are independent of each other, so expectation and variance both scale with  $n$ . We still need to find the variance of  $Z_1$  though.

We know that  $Z_1^2 = \frac{1}{16}$  with probability  $\frac{1}{2}$  (spin-up case) and  $Z_1^2 = 1$  with probability  $\frac{1}{2}$  (spin-down case). Therefore,  $E[Z_1^2] = \frac{17}{32}$ .

Now we have  $\text{Var}(Z_1) = E[Z_1^2] - E[Z_1]^2 = \frac{9}{64}$ .

Therefore,  $E[Z] = \frac{5n}{8}, \text{Var}(Z) = \frac{9n}{64}$

(c) **Find  $\text{Cov}(X_i, Y_j)$  and  $\text{Cov}(Y_i, Z_j)$  if  $1 \leq i \leq n$  and  $1 \leq j \leq n$ .**

If  $i \neq j$ , then the covariances are all 0 since different particles are completely independent.

Otherwise,  $\text{Cov}(X_i, Y_i) = E[X_i Y_i] - E[X_i]E[Y_i]$ . We know that  $X_i Y_i = 0$  always, so the expectation is also 0.

We know that  $Y_1 = \frac{1}{2}$  with probability  $\frac{1}{2}$  and  $Y_1 = 0$  with probability  $\frac{1}{2}$ . Therefore,  $E[Y_1] = \frac{1}{4}$ .

Thus,  $\text{Cov}(X_i, Y_j) = -\frac{1}{16}$  when  $i = j$  and 0 otherwise.

$\text{Cov}(Y_i, Z_i) = E[Y_i Z_i] - E[Y_i]E[Z_i]$ . We know that  $Y_i Z_i = \frac{1}{2}$  with probability  $\frac{1}{2}$  (spin-down case) and  $Y_i Z_i = 0$  with probability  $\frac{1}{2}$  (spin-up case). Therefore,  $E[Y_i Z_i] = \frac{1}{4}$ .

Thus,  $\text{Cov}(Y_i, Z_j) = \frac{3}{32}$  when  $i = j$  and 0 otherwise.

(d) **Find  $\text{Cov}(X, Y)$  and  $\text{Cov}(Y, Z)$ .**

We now use the answers from part c. First, we expand  $X$  and  $Y$ .

$$\text{Cov}(X, Y) = \text{Cov} \left( \sum_{i=1}^n X_i, \sum_{j=1}^n Y_j \right)$$

Recall that covariance is a linear operation. Therefore,

$$= \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(X_i, Y_j)$$

However, the covariance is only nonzero when  $i = j$ .

$$= \sum_{i=1}^n \text{Cov}(X_i, Y_i) = \boxed{\text{Cov}(X, Y) = -\frac{n}{16}}$$

A similar derivation shows that  $\text{Cov}(Y, Z) = \frac{3n}{32}$

(e) **Are  $X$  and  $Y$  positively correlated, negatively correlated, or uncorrelated?**

The covariance is negative, so  $X$  and  $Y$  are negatively correlated. This makes sense! If  $X$  is high, then  $Y$  will be low and vice versa.

(f) **Are  $Y$  and  $Z$  positively correlated, negatively correlated, or uncorrelated?**

The covariance is positive, so  $Y$  and  $Z$  are positively correlated. This makes sense! If  $Y$  is high, then  $Z$  will be high as well.

(g) **Are  $X$  and  $Z$  positively correlated, negatively correlated, or uncorrelated?**

By inspection, it seems like  $X$  and  $Z$  should be positively correlated. However, let's check.

We have to find the covariance. We'll start with the covariance between  $X_i$  and  $Z_j$ .

If  $i \neq j$ , then the covariances are all 0 since different particles are completely independent.

Otherwise,  $\text{Cov}(X_i, Z_i) = E[X_i Z_i] - E[X_i]E[Z_i]$ . We know that  $X_i Z_i = 0$  with probability  $\frac{1}{2}$  (spin-down case) and  $X_i Z_i = \frac{1}{8}$  with probability  $\frac{1}{2}$  (spin-up case). Therefore,  $E[X_i Z_i] = \frac{1}{16}$ .

Thus,  $\text{Cov}(X_i, Z_j) = -\frac{3}{32}$  when  $i = j$  and 0 otherwise.

Therefore,  $\text{Cov}(X, Z) = -\frac{3n}{32}$ . The covariance is negative, so  $X$  and  $Z$  are negatively correlated.

Don't trust your intuition...

## 5 Persistent Particles 4

Ryan decides to modify Hyouin's quantum particle generator to have some fun. In particular, Ryan's quantum particle generator now does the following to generate particles:

1. The first particle will have a 50% chance to be spin-up and a 50% chance to be spin-down.
2. The  $i$ 'th particle, where  $i \neq 1$ , will be spin-up with probability  $p_1$  if the  $i - 1$ 'th particle was spin-up (and  $1 - p_1$  probability of spin-down). It will be spin-down with probability  $p_2$  if the  $i - 1$ 'th particle was spin-down (and  $1 - p_2$  probability of spin-up in this scenario).
3. The quantum particle generator completely resets after generating 10 particles.

Hyouin now generates 10000 particles. Note: For parts (b)-(f), use the scenario from part (b).

### 5.1 Solution:

- (a) **Ryan first sets  $p_1 = p_2 = 0.6$ . What is the expected number of spin-up particles?**

Spin-up and spin-down are completely arbitrary labels. In fact, interchanging them does not change the problem at all. Therefore, since they must be symmetric, over the course of 10 particles we expect 5 of them to be spin-up. Since each batch of 10 is independent, we can then simply multiply this by the total number of batches, yielding 5000.

However, this seems mildly dangerous. Let's prove it more rigorously. Let  $P(i)$  denote the probability that the  $i$ 'th particle is spin-up. We'll consider  $i \geq 2$  for now. There are two conditions for the  $i$ 'th particle to be spin-up:

- The  $i - 1$ 'th particle was spin-up. We'll get spin-up with probability  $p_1 = 0.6$ .
- The  $i - 1$ 'th particle was spin-down. We'll get spin-up with probability  $1 - p_2 = 0.4$ .

However, the probability that the  $i - 1$ 'th particle was spin-up is simply given as  $P(i - 1)$ . The probability that the  $i - 1$ 'th particle was spin-down is therefore  $1 - P(i - 1)$ . Thus,

$$P(i) = P(i - 1)p_1 + (1 - P(i - 1))(1 - p_2)$$

We then plug in  $p_1 = p_2 = 0.6$  and use our initial condition:  $P(1) = 0.5$  by definition. This gives us

$$P(i) = P(i - 1)0.6 + 0.4 - 0.4P(i - 1) = 0.2P(i - 1) + 0.4$$

Intuition tells us that  $P(i) = 0.5$  for all  $i$ . This holds true for  $i = 1$ . Let us assume that it holds true for  $i - 1$  and check if it works for  $P(i)$ .

Indeed,  $P(i) = 0.2 \cdot 0.5 + 0.4 = 0.5$ . Therefore,  $P(i) = 0.5$  for all  $i$ . Ryan actually did absolutely nothing to the machine.

- (b) **For the remainder of this problem, Ryan sets  $p_1 = 0.6$  and  $p_2 = 0.55$ . What is the expected number of spin-up particles?**

We get the same recurrence relation as before, but with different values of  $p_1$  and  $p_2$  this time.

$$P(i) = P(i - 1)0.6 + 0.45 - 0.45P(i - 1) = 0.15P(i - 1) + 0.45$$



There is no easy closed form solution for this like last time. However, the particle generator completely resets after generating 10 particles. Thus, we only have to calculate the first 10 values. By linearity of expectation, we only have to calculate the expected value of the first 10 values, then add them up to get the overall expected value. We then multiply by the number of batches, as usual.

In this case, the expected value equals the probability itself since these are all Bernoulli variables (with different values of  $p$ ). The table below lists the 10 values.

$i$	1	2	3	4	5	6	7	8	9	10
$P(i)$	0.5	0.525	0.52875	0.529	0.529	0.529	0.529	0.529	0.529	0.529

The values appear to be converging onto a number (what is the number?) However, we're here to find the expected number of spin-up particles after generating 10 of them. Adding up the probabilities, this gives us about 5.26.

The 1000 batches are still independent, so this gives us an overall expected number of spin-up particles at  $\boxed{5260}$ .

- (c) **Give an upper bound on the probability that the actual number of spin-up particles is more than 8000.**

Let  $X$  denote the number of spin-up particles after generating 10000 of them. We want to find  $\Pr\{X > 8000\}$ .

Since  $X$  is an integer random variable,  $\Pr\{X > 8000\} = \Pr\{X \geq 8001\}$ . We also note that  $X$  is a non-negative random variable, so we can now use Markov's Inequality. Markov's Inequality states the following:

$$\Pr\{X \geq a\} \leq \frac{E[X]}{a}$$

We have the expected value from part (b). Therefore, the upper bound on the probability is simply

$$\boxed{\frac{5260}{8001}}.$$

- (d) **Give an upper bound on the probability that the actual number of spin-up particles vary from the expected value by more than 250.**

Let  $X$  denote the number of spin-up particles after generating 10000 of them. Upper bounds on the probability of deviating from the expected value sounds a lot like Chebyshev's Inequality. However, Chebyshev's Inequality requires knowing the standard deviation. However, we still have our probability table from above. Each particle of the 10 is now an independent Bernoulli random variable, which means we can just sum all the variances. Variance then also scales with the number of independent batches of 10.

The variance of a Bernoulli distribution is  $p(1-p)$ . Therefore, the overall variance of  $X$  is 2493. The standard deviation of  $X$  is simply the square root of the variance, which equals about 50.

Chebyshev's Inequality states that

$$\Pr\{|X - \mu| \geq k\sigma\} \leq \frac{1}{k^2}$$

Since  $\sigma = 50$  and we're looking for 250, this means that we can set  $k = 5$ . Note that our problem asked for strictly more than, which means that we would've had to use a greater than sign, but our rounding errors will account for it anyway so we don't need to worry about it.

Therefore, an upper bound on the probability will be  $\boxed{\frac{1}{25}}$ .

- (e) **What is the estimated probability that the actual number of spin-up particles vary from the expected value by less than 250?**

Let  $X$  denote the number of spin-up particles after generating 10000 of them. This is the sum of a large number of independent, identical batches, which means we can use the central limit theorem! While we should apply the continuity correction, we're not going to for ease of calculations.

We're looking for  $\Pr\{|X - \mu| \leq 250\}$ . However, we need to normalize by the standard deviation.

$$\Pr\{|X - \mu| \leq 250\} = \Pr\left\{\left|\frac{X - \mu}{50}\right| \leq 5\right\} \approx \boxed{1 - 2Q(5)}$$

However,  $Q(5) \approx 2.8 \cdot 10^{-7}$ , so this probability is very close to 1. This means that the probability that the actual number of spin-up particles varies from the expected value by **more** than 250 is  $2Q(5)$ . Recall from part (d) that the upper bound was found to be  $\frac{1}{25}$ . Man was it a bad bound.

- (f) **What is the estimated probability that the actual number of spin-up particles vary from the expected value by less than or equal to 1?**

We use the same process as in part (e) to get  $\approx \boxed{1 - 2Q\left(\frac{1}{50}\right)} \approx 0.016$ . Pretty small.

## 6 Gullible Gaussian

Aaron is drawing from a Gaussian distribution with mean 4 and variance 9.

### 6.1 Solution:

- (a) **Suppose he draws from the Gaussian distribution once and multiplies what he gets by 2. What is the mean of his result? The variance?**

Let the Gaussian distribution be denoted as  $X$ . Aaron is drawing from the Gaussian distribution  $2X$ , which has mean  $4 \cdot 2 = 8$  and variance  $9 \cdot 2^2 = 36$ .

- (b) **Suppose he draws from the Gaussian distribution twice independently and sums what he gets. What is the mean of his result? The variance?**

Let the first Gaussian distribution be denoted as  $X$  and the second Gaussian distribution be denoted as  $Y$ . Aaron is drawing from the Gaussian distribution  $X + Y$ , which has mean  $4 + 4 = 8$  and variance  $9 + 9 = 18$ .

- (c) **Is this a contradiction?**

The key idea here is that summing  $n$  independent random variables is NOT the same as scaling a random variable by  $n$ . Summing independent random variables will have far less variance compared to taking a single random variable and scaling it by a constant.

- (d) **(Unrelated to the previous parts) Give an informal argument for why Poisson processes with large  $\lambda$  can be well-approximated by a Gaussian distribution with mean  $\lambda$  and variance  $\lambda$ .**

The sum of two Poisson processes with mean and variance  $\lambda_1$  and  $\lambda_2$  respectively is another Poisson process with mean and variance  $\lambda_1 + \lambda_2$ . Therefore, Poisson processes with large  $\lambda$  can be treated as the sum of  $\lambda$  Poisson processes with mean and variance 1. Now the law of large numbers/central limit theorem comes into play; summations of a large number of random variables always tends towards the Gaussian distribution with the appropriate mean and variance. Therefore, Poisson processes with large  $\lambda$  can be well-approximated by a Gaussian distribution with mean  $\lambda$  and variance  $\lambda$ .

## 7 Burning Bridges

Esther is building a bridge between ECEB and Beckman. Her bridge will use ten concrete pillars as supports, each of which requires eight concrete blocks. The bridge is designed such that each concrete pillar can tolerate two cracked concrete blocks, and the overall bridge can tolerate a single failed concrete pillar. Thus, in order for Esther's bridge to fall, there has to be at least two concrete pillars that each have at least three cracked concrete blocks. Suppose concrete blocks crack independently with probability  $p$ .

### 7.1 Solution:

- (a) **Find an expression for the exact probability,  $p_0$ , that a particular concrete pillar fails, in terms of  $p$ .**

Let  $X$  denote the number of concrete block cracks.  $X$  is a binomial distribution with  $n = 8$  and  $p$ . Therefore,

$$p_0 = P\{X \geq 3\} = 1 - P\{X = 0\} - P\{X = 1\} - P\{X = 2\}$$

$$= \boxed{1 - (1 - p)^8 - 8(1 - p)^7 p - 28(1 - p)^6 p^2}$$

- (b) **Find an expression for the exact probability,  $p_1$ , that the bridge falls, in terms of  $p_0$ .**

Let  $Y$  denote the number of failed concrete pillars.  $Y$  is a binomial distribution with  $n = 10$  and  $p = p_0$ . Therefore,

$$p_1 = P\{Y \geq 2\} = 1 - P\{Y = 0\} - P\{Y = 1\}$$

$$= \boxed{1 - (1 - p_0)^{10} - 10(1 - p_0)^9 p_0}$$

## 8 Articulate Arrival

Calls arrive to a cell in a certain wireless communication system according to a Poisson process with an arrival rate of  $\lambda = 5$  calls per minutes. We now measure time in minutes and consider an interval of time beginning at  $t = 0$ . Let  $N_t$  denote the number of calls that arrive in  $t$  minutes. For a fixed  $t > 0$ , the random variable  $N_t$  is a Poisson random variable with parameter  $5t$ .

### 8.1 Solution:

- (a) **Find the probability that no calls arrive in the first four minutes.**

We know that  $N_4$  denotes the number of calls that arrive in 4 minutes. In addition,  $N_4$  has a Poisson distribution with mean 20. Therefore, the probability that no calls arrive in the first four minutes is

$$P\{N_4 = 0\} = \frac{e^{-20}(20)^0}{0!} = \boxed{e^{-20}}$$

- (b) **Find the probability that the first call arrives after four minutes.**

This is the same as part (a). Therefore, the probability is still  $\boxed{e^{-20}}$ .

- (c) **Find the probability that two or fewer calls arrive in the first four minutes.**

This is the same as the probability that  $N_4$  is less than or equal to 2.

$$P\{N_4 \leq 2\} = P\{N_4 = 0\} + P\{N_4 = 1\} + P\{N_4 = 2\} = e^{-20} \left( 1 + 20 + \frac{20^2}{2} \right) = \boxed{221e^{-20}}$$

- (d) **Find the probability that the third call arrives before time  $t$ .**

The probability that the third call arrives before time  $t$  is the same as one minus the probability that two or fewer calls occur before time  $t$ . There, we're looking for

$$1 - P\{N_t \leq 2\} = 1 - P\{N_t = 0\} - P\{N_t = 1\} - P\{N_t = 2\} = \boxed{1 - e^{-5t} \left( 1 + 5t + \frac{(5t)^2}{2} \right)}$$

(e) **Derive the pdf of the arrival time of the third call.**

Let  $X_3$  denote the arrival time of the third call. Part (d) had us find the probability that the third call arrives before time  $t$ , which is  $P\{X_3 < t\} = P\{X_3 \leq t\}$ . However, this is the definition of the CDF! Therefore, we simply differentiate what we got in part (d) in order to get the pdf.

$$\frac{d}{dt} \left( 1 - e^{-5t} \left( 1 + 5t + \frac{(5t)^2}{2} \right) \right) = -e^{-5t}(5 + 25t) + 5e^{-5t} \left( 1 + 5t + \frac{(5t)^2}{2} \right) = \boxed{e^{-5t} \frac{5^3 t^2}{2}}$$

(f) **Find the expected arrival time of the sixth call.**

The times between arrivals are exponentially distributed with parameter  $\lambda$ . The expected time between arrivals is thus  $\frac{1}{\lambda}$ . Therefore, the expected time until the sixth arrival is  $\frac{6}{\lambda} = \boxed{\frac{6}{5}}$ .

## 9 Noble Normal

Let  $X$  have the  $N(10, 16)$  distribution. Find the numerical values of the following probabilities (practice using the  $Q$  tables).

### 9.1 Solution:

The idea is to use the fact that  $\frac{X-10}{4}$  is a standard normal random variable.

(a)  $P\{X \geq 15\}$

$$P\{X \geq 15\} = P\left\{\frac{X-10}{4} \geq \frac{15-10}{4}\right\} = Q\left(\frac{15-10}{4}\right) = Q(1.25) \approx \boxed{0.1056}$$

(b)  $P\{X \leq 5\}$

$$P\{X \leq 5\} = P\left\{\frac{X-10}{4} \leq \frac{5-10}{4}\right\} = \Phi\left(\frac{5-10}{4}\right) = Q(1.25) \approx \boxed{0.1056}$$

(c)  $P\{X^2 \geq 400\}$

$$\begin{aligned} P\{X^2 \geq 400\} &= P\{X \geq 20\} + P\{X \leq -20\} \\ &= P\left\{\frac{X-10}{4} \geq \frac{20-10}{4}\right\} + P\left\{\frac{X-10}{4} \leq \frac{-20-10}{4}\right\} \\ &= Q(2.5) + Q(7.5) \approx Q(2.5) = \boxed{0.0062} \end{aligned}$$

Note that  $Q(7.5) < 10^{-12}$ , so the approximation was valid.

(d)  $P\{X = 2\}$

Since  $X$  is a continuous random variable, the probability here is just  $\boxed{0}$ .

## 10 Uniquely Uniform

### 10.1 Solution:

- (a) **Suppose  $X$  is a continuous random variable with CDF  $F_X$ . Let  $Y$  be the result of applying  $F_X$  to  $X$  (aka  $Y = F_X(X)$ ). What is the distribution of  $Y$ ?**

Let's find the CDF of  $Y$ , aka  $F_Y(c)$ .

$$F_Y(c) = \Pr(Y \leq c) = \Pr(F_X(X) \leq c) = \Pr(X \leq F_X^{-1}(c))$$

We know that  $F_X(c) = \Pr(X \leq c)$ . Therefore,  $F_Y(c) = F_X(F_X^{-1}(c)) = c$ .

The only distribution that has this CDF is the uniform distribution from 0 to 1.

- (b) **Find a function  $g$  such that, if  $U$  is uniformly distributed over the interval  $[0, 1]$ ,  $g(U)$  is a Rayleigh distribution with  $\sigma^2 = 9$ .**

The target CDF is

$$1 - \exp\left(-\frac{t^2}{18}\right)$$

We set this equal to  $u$  and solve.

$$1 - \exp\left(-\frac{t^2}{18}\right) = u$$

$$\exp\left(-\frac{t^2}{18}\right) = 1 - u$$

$$-\frac{t^2}{18} = \ln(1 - u)$$

$$\boxed{g(u) = \sqrt{-18 \ln(1 - u)}}$$

Note that  $u$  is a number between 0 and 1, so the  $\ln$  will always be negative, so it will cancel out the other negative sign in the square root.



## 11 Displeased Distribution

Let  $T$  have the pdf  $f_{\sigma^2}(t) = I_{\{t \geq 0\}} \frac{t}{\sigma^2} e^{-\frac{t^2}{2\sigma^2}}$ , where  $\sigma^2$  is unknown. Find the ML estimate of  $\sigma^2$  if it is observed that  $T = 12$ . Note that  $I_{\{t \geq 0\}} = 1$  if  $t \geq 0$  and 0 otherwise.

### 11.1 Solution:

We first plug in  $t = 12$ .

$$f_{\sigma^2}(12) = \frac{12}{\sigma^2} e^{-\frac{72}{\sigma^2}}$$

Now we need to take the derivative with respect to  $\sigma^2$  and set it to 0. This is not the most fun derivative to take, so we introduce a trick. Let  $x = \frac{12}{\sigma^2}$ . We now have

$$f_{\sigma^2}(12) = x e^{-6x}$$

If we know the value of  $x$  that maximizes this, we can then derive the value of  $\sigma^2$  that maximizes this as well. Therefore, we can now take the derivative with respect to  $x$  and set it equal to 0.

$$\frac{d}{dx} x e^{-6x} = 0 = e^{-6x} - 6x e^{-6x}$$

We know that  $e^{-6x}$  will not be exactly equal to 0 for any  $x$ , so we drop it. We're left with  $1 - 6x = 0$ , implying that  $x = \frac{1}{6}$ .

Now we can solve for  $\sigma^2$ , since  $x = \frac{12}{\sigma^2}$ . We therefore get our ML estimate to be 72.

## 12 Gruesome Gaussian

Suppose  $X$ ,  $Y$ , and  $Z$  are independent Gaussian distributions.  $X$  has mean 1 and variance 4.  $Y$  has mean 0 and variance 9.  $Z$  has mean 10 and variance 1. Find the pdf of  $X + Y + Z$ .

Hint: ECE210's frequency domain may be useful.

### 12.1 Solution:

Let's first write out the three pdfs.

$$f_X(c) = \frac{1}{\sqrt{2\pi}4} \exp\left(-\frac{(c-1)^2}{2 \cdot 4}\right)$$

$$f_Y(c) = \frac{1}{\sqrt{2\pi}9} \exp\left(-\frac{(c-0)^2}{2 \cdot 9}\right)$$

$$f_Z(c) = \frac{1}{\sqrt{2\pi}1} \exp\left(-\frac{(c-10)^2}{2 \cdot 1}\right)$$

We know that  $f_{X+Y+Z}(c) = f_X(c) * f_Y(c) * f_Z(c)$ .

However, I am not in the mood to convolve three Gaussian distributions together. Recall in ECE210 that convolution in time is equivalent to multiplication in frequency. So let's convert these into the frequency domain, multiply, then convert back out.

From the ECE210 tables, we know that

$$\exp\left(-\frac{t^2}{2\sigma^2}\right) \iff \sigma\sqrt{2\pi} \exp\left(-\frac{\sigma^2\omega^2}{2}\right)$$

Therefore, let us apply the Fourier Transform to each of the three pdfs separately. We'll denote this function as  $T_X(u)$ . Note that a time shift results in an exponential in the frequency domain.

$$f_X(c) = \frac{1}{\sqrt{8\pi}} \exp\left(-\frac{(c-1)^2}{8}\right) \iff T_X(u) = \frac{\sqrt{2\pi \cdot 4}}{\sqrt{8\pi}} \exp\left(-\frac{4u^2}{2}\right) e^{-ju1}$$

$$f_Y(c) = \frac{1}{\sqrt{18\pi}} \exp\left(-\frac{c^2}{18}\right) \iff T_Y(u) = \frac{\sqrt{2\pi \cdot 9}}{\sqrt{18\pi}} \exp\left(-\frac{9u^2}{2}\right)$$

$$f_Z(c) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(c-10)^2}{2}\right) \iff T_Z(u) = \frac{\sqrt{2\pi}}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) e^{-ju10}$$

Now we multiply everything together.

$$T_{X+Y+Z}(u) = T_X(u)T_Y(u)T_Z(u) = \exp\left(-\frac{4u^2}{2} - \frac{9u^2}{2} - \frac{u^2}{2}\right) e^{-ju11}$$

$$= \exp\left(-\frac{14u^2}{2}\right) e^{-ju11} \iff \boxed{\frac{1}{\sqrt{28\pi}} \exp\left(-\frac{(c-11)^2}{28}\right)}$$

The resulting pdf is also a Gaussian distribution with mean 11 and variance 14. Cool.

## 13 Bombastic Brain Cell

Suppose the failure rate function for someone's brain cell with lifetime  $E$  is  $h(t) = a^3$  for  $0 \leq t \leq 2$  and  $h(t) = bt + c$  for  $t \geq 2$  (you get to decide what units to use). Find the CDF and mean of the lifetime  $E$ . For the mean, it is not necessary to evaluate the integral.

### 13.1 Solution:

We know that the CDF  $F_E(t)$  is given by the following formula:

$$F_E(t) = 1 - e^{-\int_0^t h(s)ds}$$

We can therefore plug in the given failure function. When  $0 \leq t \leq 2$ , the failure function is simply  $a^3$ .

$$F_E(t) = 1 - e^{-\int_0^t a^3 ds} = 1 - e^{-a^3 t}$$

When  $t \geq 2$ , our integral now becomes

$$F_E(t) = 1 - e^{-\int_0^2 a^3 ds - \int_2^t bt + cds}$$

Evaluating, we get

$$F_E(t) = 1 - e^{-2a^3 - \frac{b}{2}t^2 + 2b - ct + 2}$$

Combining, our final answer is

$$F_E(t) = \begin{cases} 1 - e^{-a^3 t} & 0 \leq t \leq 2 \\ 1 - e^{-2a^3 - \frac{b}{2}t^2 + 2b - ct + 2} & t \geq 2 \end{cases}$$

The expected value can be calculated from the CDF by using the following integral:

$$E[E] = \int_0^\infty (1 - F_E(t))dt$$

We can now plug our CDF into this equation. We will not evaluate this integral.

$$= \int_0^2 e^{-a^3 t} dt + \int_2^\infty e^{-2a^3 - \frac{b}{2}t^2 + 2b - ct + 2} dt$$

## 14 Bubbly Balloons

Aaron is blowing up balloons for his birthday party. However, he's not very consistent with how much helium he puts in his balloons. The radius of his balloons follow a uniform distribution with mean 4 meters and variance 3 meters. You may assume that the balloons are all perfect spheres. You may also assume that negative radii, however unlikely, are ok.

Let  $V$  be the random variable denoting the volume of the balloons, while  $S$  is the random variable denoting the surface area of the balloons.

### 14.1 Solution:

Let  $R$  denote the random variable representing the radius. We know that  $S = 4\pi R^2$ , while  $V = \frac{4}{3}\pi R^3$ .

We know that  $R$  is a uniform distribution with mean 4 and variance 3. Let  $a$  denote the smallest possible  $R$  and  $b$  denote the largest possible size of  $R$ . The mean is then  $\frac{a+b}{2}$ , while the variance is  $\frac{(b-a)^2}{12}$ . Solving, we get  $a = 1$  and  $b = 7$ . This means that the pdf of  $R$ ,  $f_R(u)$ , equals  $\frac{1}{6}$  when  $u$  is between 1 and 7, and 0 elsewhere.

(a) **What is the support of  $S$  and  $V$ ?**

Since  $a$  and  $b$  are both positive, we can simply plug in the smallest and biggest values of  $R$  to obtain

the smallest and biggest values of  $S$  and  $V$ . Therefore,  $S \in [4\pi, 196\pi], V \in [\frac{4\pi}{3}, \frac{1372\pi}{3}]$ .

(b) **What is  $E[S]$  and  $E[V]$ ?**

By expectation rules, we know that

$$E[g(X)] = \int_{-\infty}^{\infty} g(u)f_X(u)du$$

We can then plug everything in for  $S$  and  $V$ .

$$E[S] = \int_1^7 4\pi u^2 \frac{1}{6} du = \frac{2\pi}{9} u^3 \Big|_1^7 = \boxed{76\pi}$$

$$E[V] = \int_1^7 \frac{4}{3}\pi u^3 \frac{1}{6} du = \frac{1\pi}{18} u^4 \Big|_1^7 = \boxed{\frac{400\pi}{3}}$$

Note that the bounds of the integral are defined by the support of  $f_R(u)$ .

(c) **Find the CDF of  $V$ .**

We denote the CDF as  $F_V(u)$ . By the definition of CDF, we have

$$F_V(u) = \Pr\{\frac{4}{3}\pi R^3 \leq c\} = \Pr\{R^3 \leq \frac{3}{4\pi}c\} = \Pr\{R \leq \left(\frac{3}{4\pi}c\right)^{\frac{1}{3}}\}$$

Since  $R$  is a uniform distribution, we know that  $F_R(u) = \Pr(R \leq u) = \frac{1}{6}(u - 1)$  for  $u$  within 1 to 7, 0 for  $u \leq 1$ , and 1 for  $u \geq 7$ .

Now let's look carefully at what each expression is saying. The first equation says the probability that  $R$  is less than or equal to something. The second equation also says the probability that  $R$  is less than or equal to  $u$ , and gives a formula for it. Therefore, we can plug the expression we got in the first equation as  $u$  in the second equation to get our final CDF. We then make sure to use our support so that the CDF remains valid.

$$F_V(u) = \begin{cases} 0 & \text{if } u \leq \frac{4\pi}{3} \\ \frac{1}{6} \left( \frac{3}{4\pi} u \right)^{\frac{1}{3}} - \frac{1}{6} & \text{if } \frac{4\pi}{3} < u < \frac{1372\pi}{3} \\ 1 & \text{if } u \geq \frac{1372\pi}{3} \end{cases}$$

(d) **Find the pdf of  $V$ .**

To find the pdf, we simply take the derivative of what we found in part (c).

$$f_V(u) = \begin{cases} 0 & \text{if } u \leq \frac{4\pi}{3} \\ \frac{1}{24\pi} \left( \frac{3}{4\pi} u \right)^{-\frac{2}{3}} & \text{if } \frac{4\pi}{3} < u < \frac{1372\pi}{3} \\ 0 & \text{if } u \geq \frac{1372\pi}{3} \end{cases}$$

## 15 Deft Distribution

Suppose  $X$  and  $Y$  have a bivariate Gaussian joint distribution with  $E[X] = E[Y] = 1$  and  $\text{Var}(X) = 2$ . The variance of  $Y$  and the correlation coefficient are not known. Finally, suppose that  $X$  is independent of  $X - Y$ .

### 15.1 Solution:

(a) **Find  $\text{Cov}(X, Y)$ .**

Since  $X$  is independent of  $X - Y$ , that means that  $X$  and  $X - Y$  are uncorrelated. Therefore,

$$0 = \text{Cov}(X, X - Y) = \text{Cov}(X, X) - \text{Cov}(X, Y) = 2 - \text{Cov}(X, Y)$$

Therefore,  $\boxed{\text{Cov}(X, Y) = 2}$ .

(b) **Find  $E[X|X - Y = 2]$ .**

Since  $X$  and  $X - Y$  are independent,  $E[X|X - Y = 2] = E[X] = \boxed{1}$ .

(c) **Find  $E[Y|X = 2]$ .**

Since  $X$  and  $Y$  are jointly Gaussian, we know that

$$E[Y|X = 2] = \hat{E}[Y|X = 2] = L^*(X = 2) = \mu_Y + \frac{\text{Cov}(X, Y)}{\text{Var}(X)}(2 - \mu_X) = \boxed{2}$$

## 16 Dangerous Diagonal Darts (with Determinants!)

Lauren has challenged Alex to a game of darts. Unfortunately, Alex is quirky and he only owns a triangular dart board, with corners at  $(0,0)$  and  $(2,0)$  and  $(0,w)$  in the  $xy$ -plane. Lauren is annoyed with this—any normal person has a triangular dart board with corners at  $(0,0)$  and  $(2,0)$  and  $(0,2)$  in the  $xy$ -plane. Let  $X$  and  $Y$  be the random variables for the location of Lauren's throw, and suppose their joint density is given by

$$f(x, y) = \begin{cases} 0 & x < 0 \text{ or } x > 2 \text{ or } y < 0 \text{ or } y > 2 \\ \frac{3}{16}x^2(2-y) & \text{otherwise} \end{cases}$$

To be fair to Lauren, Alex will let her define two linear transformations,  $U = g_1(X) = c_1X$  and  $V = g_2(Y) = c_2Y$  before her throw, and will use the transformed position  $(U, V)$  to determine where her throw hit on his wide dartboard.

### 16.1 Solution

- (a) **Show that  $U$  and  $V$  are invertible and differentiable.**

It's easy to see that  $U$  and  $V$  are differentiable.  $dU/dX = c_1$  and  $dV/dY = c_2$ . They're also invertible, since one can define  $U^{-1} = U/c_1$  and  $V^{-1} = V/c_2$ .

- (b) **Find the values of  $c_1$  and  $c_2$  to maximize the probability that  $(U, V)$  lies within the region defined by Alex's dartboard. Hint: it's simpler than you think! If you do want to work out all the math, you'll probably want a computer algebra system.**

Since  $U$  and  $V$  are differentiable invertible transformations we can use the determinant of their Jacobian matrix and find their joint density, and then find values of  $c_1$  and  $c_2$  to maximize the value of the integral for this joint density over Alex's triangle.

We have

$$J_g = \frac{\partial g_1}{\partial X} \frac{\partial g_2}{\partial Y} - \frac{\partial g_1}{\partial Y} \frac{\partial g_2}{\partial X} = c_1 c_2$$

so then the joint density of  $U$  and  $V$  is

$$f(u, v) = f(U^{-1}, V^{-1}) \frac{1}{J_g} = \frac{3}{16} \frac{u^2}{c_1^3 c_2} \left( 2 - \frac{v}{c_2} \right)$$

For this to be a probability distribution we need

$$\int_0^w \int_0^{2-w/v} f(u, v) dv du = \frac{w^3}{c_1^3 c_2^2} (.25c_2 - 1/8) = 1$$

and some algebra quickly tells us

$$c_1^3 = \frac{w^3}{c_2^2} \left( \frac{c_2}{4} - \frac{1}{8} \right)$$

to satisfy the above. Now we consider the probability that the dart lies in Alex's triangle:

$$\int_0^w \int_0^{2-(2/w)v} f(u, v) dv du = \frac{w^3(0.0625c_2 - 0.0125)}{c_1^3 c_2^2}$$

and we can substitute the value of  $c_1$  into this to obtain

$$\text{Pr(in triangle)} = \frac{.0625c_2 - .0125}{.25c_2 - 1/8}$$

We don't need to do any calc to maximize this: we can simply set it equal to 1 and obtain  $c_2 = 0.6$ . Then we can plug back into  $c_1$  to obtain

$$c_1^3 = \frac{5w^3}{72}$$

so  $c_1 = .411w$  and Lauren is guaranteed to win. Sorry, Alex.

(c) **What is the geometric interpretation of your solution to the above?**

Lauren is using  $c_1$  and  $c_2$  to create a box completely contained within Alex's triangle where her dart will be guaranteed to fall. This is why we said the part above was easy: one could have recognized this and solved for a box directly, foregoing all the transformations!

(d) **Bonus: Alex thinks this is too easy: now Lauren must define two *cubic* transformations,  $U = c_1X^3$  and  $V = c_2Y^3$ . Again, find  $c_1$  and  $c_2$  to maximize her chances of hitting Alex's dartboard.**

Apply the same procedure as part (b), but with  $J_g = 9c_1c_2X^2Y^2$ .



## 17 Packed Plane

Ryan is the first person to board a plane that contains  $n$  seats. However, he has forgotten what seat he is supposed to take, and he has also lost his boarding pass (what is the probability of that!) Therefore, he simply chooses a random seat and sits down.

Aaron is the second person to board the same plane. However, he has also forgotten what seat he was supposed to take. Aaron's boarding pass is buried beneath 200 bags of Swedish Fish, so he's not in the mood to pull it out. Therefore, he also picks an empty seat randomly.

The next  $n - 2$  passengers behind Aaron then board the plane. When a passenger boards and finds Ryan or Aaron in their seat, they will ask them to move. Ryan or Aaron will then choose another empty seat randomly. Importantly, Aaron and Ryan don't know that the other person also does not know what their seat is. Therefore, Aaron will never ask Ryan to move, and vice versa.

### 17.1 Solution:

- (a) **What is the probability that Ryan never needs to move after choosing a seat randomly for the first time? What about Aaron?**

Of the  $n$  seats available to Ryan when he boards the plane, 2 of them, namely Ryan's seat and Aaron's seat, will result in Ryan not having to move. Therefore, the probability that Ryan never needs to move after choosing a seat randomly for the first time is  $\boxed{\frac{2}{n}}$ .

Of the  $n - 1$  seats available to Aaron when he boards the plane, 1 of them, namely either Ryan's seat and Aaron's seat that Ryan currently isn't in, will result in Aaron not having to move. Therefore, the probability that Aaron never needs to move after choosing a seat randomly for the first time is  $\boxed{\frac{1}{n - 1}}$ .

- (b) **What is the exact probability that the  $k$ 'th passenger has to ask Ryan to move?**

When the  $k$ 'th passenger boards, passengers 3 through  $k - 1$  are all in their assigned seats (and Aaron is sitting somewhere random). Ryan could be in any of the remaining  $n - k + 3$  seats randomly. Therefore, the probability that Ryan needs to move is then  $\boxed{\frac{1}{n - k + 3}}$ .

Not convinced? Try small values of  $n$  and  $k$ .

- (c) **What is the exact probability that the  $k$ 'th passenger has to ask Aaron to move?**

Same as part (b).  $\boxed{\frac{1}{n - k + 3}}$ .

- (d) **What is the expected number of times that Ryan has to move? What about Aaron?**

After each passenger, Ryan has to move with probability  $\frac{1}{n-k+3}$ . Therefore, we simply add this up for all  $k$ . The final answer is

$$\sum_{k=3}^n \frac{1}{n-k+3} = \boxed{\sum_{j=3}^n \frac{1}{j}}$$

Aaron also needs to move that many times.

- (e) **What is the exact probability that Ryan and Aaron end up in their assigned seats?**

At the very end when all  $n - 2$  passengers are in their seats, there are 2 seats remaining. Ryan is in one of these seats, while Aaron is in the other. Therefore, the exact probability that Ryan and Aaron are in their assigned seats is 50%.

- (f) **Bonus: Is there some other method of choosing seats such that Ryan and Aaron are more likely to end up in their assigned seats?**

They could just look at their boarding passes or ask a flight attendant or something.

However, if this option is unavailable to them, then the answer appears to be no, since Ryan and Aaron don't have any information to go off of. Feel free to message the author (alexnmz2) if you think this solution is wrong.

- (g) **Bonus: Is there some other method of choosing seats such that Ryan and Aaron have to move around less?**

The answer also appears to be no, since Ryan and Aaron don't have any information to go off of. Feel free to message the author (alexnmz2) if you think this solution is wrong.

## 18 Demoniac Dice

Ryan is at board-game night and is learning a fun new game called Dungeons and Dragons. The game requires 20-sided dice. Unfortunately, the party only has 6-sided dice.

Devise a method to roll a 20-sided die such that each number could be rolled with equal probability using only 6-sided die. The method should use the least expected number of rolls possible (for example, solutions that are expected to use 4 rolls are better than solutions that are expected to use 4.3 rolls). Also give the expected number of rolls that your method will use.

### 18.1 Solution 1:

We can generate an integer between 1 and 36 by rolling two dice. If the integer is between 1 and 20, then we're done. Otherwise, we should just try again.

The probability that we succeed is  $\frac{20}{36}$ , so this is a geometric distribution with  $p = \frac{5}{9}$ . However, each trial takes two dice rolls. Therefore, the expected number of rolls is  $\frac{2}{p} = \frac{18}{5} = \boxed{3.6}$ .

### 18.2 Solution 2:

We can instead generate an integer between 1 and 4 and another integer between 1 and 5.

The former is a geometric distribution with  $p_1 = \frac{2}{3}$ . The latter is a geometric distribution with  $p_2 = \frac{5}{6}$ . The two distributions are independent, so, by linearity of expectation, we get that the expected number of rolls is  $\frac{1}{p_1} + \frac{1}{p_2} = \boxed{2.7}$ .

### 18.3 Solution 3:

The following solution appears to be optimal. We first create the following ranges of numbers:  $\{1 - 5\}$ ,  $\{6 - 10\}$ ,  $\{11 - 15\}$ ,  $\{16 - 20\}$ . We'll label them as  $R_1$ ,  $R_2$ ,  $R_3$ , and  $R_4$  respectively.

We then follow the following steps:

1. We first roll one die and get the number  $d$ . If it is between 1 and 4, then choose  $R_d$  as our range and skip the next step.
2. Otherwise, roll a second die. If it is between 1 and 3 and  $d = 5$ , then choose  $R_1$ . If it is between 4 and 6 and  $d = 5$ , then choose  $R_2$ . If it is between 1 and 3 and  $d = 6$ , then choose  $R_3$ . If it is between 4 and 6 and  $d = 6$ , then choose  $R_4$ .
3. Finally, we now have our range. We can therefore roll a die repeatedly until we get a number  $m$  that isn't 6. We then return the  $m$ 'th number in the range chosen.

The first step always takes 1 roll. The second step only happens  $\frac{1}{3}$  of the time, and also only takes 1 roll if it does happen. Therefore, the expected number of rolls here is  $\frac{1}{3}$ . The final step is a geometric distribution with  $p = \frac{5}{6}$ , so the expected number of rolls is  $\frac{1}{p} = \frac{6}{5}$ .

By linearity of expectation, we can add these expectations up. The overall expected number of rolls is then  $\boxed{\frac{38}{15} \approx 2.533}$ .

Feel free to message alexmz2 if you think you can do better.

## 19 Meticulous Matches

### 19.1 Solution:

- (a) **Tabish is going to do what no child should do: play with matches. Tabish takes a match stick and breaks it into three parts. He then attempts to arrange these three parts into a triangle. If Tabish breaks the match at two randomly chosen positions, what is the probability that the three parts can be arranged into a triangle?**

Let us assume that the match has a length of 1 meter (chosen for ease of arithmetic, although it admittedly is a very large match). In order to form a triangle, no piece formed can be bigger than 0.5 meters. Finally, we align the match such that the left side is  $x = 0$  and the right side is  $x = 1$ . We now consider four possible scenarios.

- The first piece is longer than 0.5 meters, i.e. the first piece stretches from 0 to some position bigger than 0.5. For this to happen, both cuts need to be in the second half. Cuts are chosen independently and uniformly across the match, so this happens with probability  $\frac{1}{4}$ . This scenario will not allow us to make a triangle.
- The second piece is longer than 0.5 meters. Let  $X$  denote the position of the first cut and  $Y$  denote the position of the second cut. We want to find  $\Pr\{|Y - X| \geq 0.5\} = 2\Pr\{Y - X \geq 0.5\}$  since  $X$  and  $Y$  are the same random variable in this scenario.

The joint pdf distribution  $X, Y$  is uniform over the unit square. We then simply integrate over the part that is interesting to us.

$$2\Pr\{Y - X \geq 0.5\} = 2 \int_{x=0}^{0.5} \int_{y=x+0.5}^1 1 dy dx$$

The bounds for  $y$  go from  $x + 0.5$  to 1, since that is where  $Y - X \geq 0.5$ . The bounds for  $x$  go from 0 to 0.5, since those are the only values of  $x$  that permit this scenario being true. Now we can just evaluate.

$$\begin{aligned} 2 \int_{x=0}^{0.5} \int_{y=x+0.5}^1 1 dy dx &= 2 \int_{x=0}^{0.5} y|_{x+0.5}^1 dx = 2 \int_{x=0}^{0.5} 0.5 - x dx \\ &= 2 \left( 0.5x - \frac{1}{2}x^2 \right) \Big|_0^{0.5} = \frac{1}{4} \end{aligned}$$

Therefore, this scenario will happen with probability  $\frac{1}{4}$ . This scenario will not allow us to make a triangle.

- The third piece is longer than 0.5 meters, i.e. the third piece stretches from some position smaller than 0.5 up to  $x = 1$ . For this to happen, both cuts need to be in the first half. Cuts are chosen independently and uniformly across the match, so this happens with probability  $\frac{1}{4}$ . This scenario will not allow us to make a triangle.
- None of the pieces are longer than 0.5 meters. This is simply all the other parts subtracted from 1. This scenario is also the only scenario that allows us to make a triangle.

Therefore, the final probability is  $\boxed{\frac{1}{4}}$ .

- (b) **Tabish instead chooses to first break the match at a random position, then break the longest of the two resulting parts at a random position. What is the probability that the three parts can be arranged into a triangle?**

Let us assume that the match has a length of 1 meter (chosen for ease of arithmetic, although it admittedly is a very large match). In order to form a triangle, no piece formed can be bigger than 0.5 meters. Finally, we align the match such that the left side is  $x = 0$  and the right side is  $x = 1$ .

Let us also assume that our first break happens on the 0 to 0.5 meter side, leaving the longer half on the right. This problem is symmetric if our first break was on the 0.5 to 1 meter side, so we'll just multiply our final answer by 2 in the end. Let's denote the position of this first break as  $y$ .

There are now two scenarios that we must consider.

- The second cut occurs such that the rightmost piece is longer than 0.5. This happens with probability  $\frac{1-y-0.5}{1-y}$ , since the length of the longer piece is  $1-y$ , and the cut can happen anywhere except in a region of length 0.5 meters.
- The second cut occurs such that the center piece is longer than 0.5. This happens with probability  $\frac{1-y-0.5}{1-y}$  as well, since the cut can happen anywhere except the region of length 0.5 meters stretching from the first cut rightward.

All other scenarios allow us to form a triangle, so we use  $1 - \frac{1-y-0.5}{1-y} - \frac{1-y-0.5}{1-y} = \frac{y}{1-y}$ .

We then need to integrate this over all possible values of  $y$ , which is from 0 to 0.5. Don't forget the extra factor of 2 as well!

$$2 \int_0^{0.5} \frac{y}{1-y} dy \approx \boxed{0.386}$$

The last step was evaluated using Wolfram Alpha. This is higher than part (a), which is good!

## 20 Scenic Sleepwalk

Sania tends to sleepwalk. However, she sleepwalks in a peculiar way: every minute, starting at minute 1, she will randomly choose a direction  $(+x, -x)$  and walk one foot towards that direction. Suppose that the room she's sleeping in is a  $n$  feet long. Sania will wake up when she runs into a wall (i.e. at  $x = 0$  or  $x = n + 1$ ). Since we don't want this to happen for various reasons, answer the following questions below:

### 20.1 Solution:

- (a) **Given that she starts at  $x = 1$ , what is the probability that she wakes up by running into the  $x = n + 1$  wall?**

By part (b), the probability is  $\boxed{\frac{1}{n+1}}$ .

- (b) **Given that she starts at  $x = k$ , what is the probability that she wakes up by running into the  $x = n + 1$ ? Express your answer in terms of  $k$ .**

Let  $T(k)$  denote the probability that she wakes up by running into the  $x = n + 1$  wall. We allow  $k$  to go from 0 to  $n + 1$ . If  $k = 0$ , then Sania has ran into the  $x = 0$  wall and ended her sleepwalk at the wall we didn't want, so  $T(0) = 0$ . If  $k = n + 1$ , then Sania has ran into the  $x = n + 1$  wall and ended our sleepwalk at the wall we wanted, so  $T(n + 1) = 1$ .

We now have the following expression for  $T(k)$ :

$$T(k) = \frac{1}{2}T(k-1) + \frac{1}{2}T(k+1)$$

This is because if Sania decides to go left, then we'll be at the  $k - 1$ 'th position, so the probability will be  $T(k - 1)$ . Similar situation for when she moves right.

This recurrence relation defines an arithmetic sequence. We conjecture that  $T(k) = \frac{k}{n+1}$ . This satisfies our boundary cases, so it seems like a good candidate.

Now let us check by plugging this into the actual expression. To do this, we'll assume that the expression is correct for all other  $T(k')$  except when  $k' = k$ , which we will then prove.

$$T(k) = \frac{1}{2}T(k-1) + \frac{1}{2}T(k+1)$$

We know that  $T(k-1) = \frac{k-1}{n+1}$  and  $T(k+1) = \frac{k+1}{n+1}$ . We now have

$$T(k) = \frac{1}{2} \frac{k-1}{n+1} + \frac{1}{2} \frac{k+1}{n+1}$$

$$T(k) = \frac{k-1+k+1}{2(n+1)}$$

$$T(k) = \frac{2k}{2(n+1)}$$

$$T(k) = \boxed{\frac{k}{n+1}}$$

- (c) **Given that she starts at  $x = 1$ , what is the expected amount of time that elapses before she runs into a wall?**

By part (d), we have  $\boxed{1(n-1+1) = n}$

- (d) **Given that she starts at  $x = k$ , what is the expected amount of time that elapses before she runs into a wall? Express your answer in terms of  $k$ .**

Let  $T(k)$  denote the expected number of steps before she runs into a wall given that she starts at position  $k$ . We again allow  $k$  to go from 0 to  $n + 1$ . If  $k = 0$ , then Sania has run into a wall, so

$T(0) = 0$ . If  $k = n + 1$ , then Sania has also run into a wall, so  $T(n + 1) = 0$  as well (in both cases, the sleepwalk is already over, so the number of steps required to end the sleepwalk is 0).

We now have the following expression for  $T(k)$ :

$$T(k) = 1 + \frac{1}{2}T(k-1) + \frac{1}{2}T(k+1)$$

At  $T(k)$ , we will take one step. We will then either end up in the  $k-1$  situation or the  $k+1$  situation with equal probability.

We conjecture that the solution must be  $T(k) = k(n-k+1)$  (this comes after just playing with the problem for a while. It can also be solved methodically using generating functions, which are out of the scope of this class, so we'll stick with the playing for now).

We will now prove that it is the correct solution using the same method as part (b). First, we note that this expression satisfies our boundary cases well.

Now we assume the expression is correct for all  $T(k')$  for  $k' \neq k$ . We then prove that the  $k' = k$  case is correct. We start with

$$T(k) = 1 + \frac{1}{2}T(k-1) + \frac{1}{2}T(k+1)$$

We now plug the  $T(k-1)$  case and the  $T(k+1)$  case in.

$$T(k) = 1 + \frac{1}{2}((k-1)(n-(k-1)+1)) + \frac{1}{2}((k+1)(n-(k+1)+1))$$

$$T(k) = 1 + \frac{1}{2}((k-1)(n-k+2)) + \frac{1}{2}((k+1)(n-k))$$

$$T(k) = 1 + \frac{1}{2}(kn - k^2 + 2k - n + k - 2) + \frac{1}{2}(kn - k^2 + n - k)$$

$$T(k) = 1 - 1 + \frac{1}{2}kn + \frac{1}{2}kn - \frac{1}{2}k^2 - \frac{1}{2}k^2 + k + \frac{1}{2}k - \frac{1}{2}k - \frac{1}{2}n + \frac{1}{2}n$$

$$T(k) = kn - k^2 + k$$

$$\boxed{k(n-k+1)}$$

The proof holds, so this is correct.

**Bonus: How do you solve the 2D version of this problem? In the 2D version, Sania will randomly choose a direction  $(+x, -x, +y, -y)$ . The room she's sleeping in is  $n \times n$ . Sania will wake up when she hits any of the four walls. She starts at  $(a, b)$ .**

The author does not know how to solve the bonus (more likely, he just hasn't spent enough time playing with the problem to guess the correct recurrence). If you know the solution to the 2D problem, please message alexmz2.

## 21 Spry Stairs

Sahan needs to climb the four flights of stairs to get to his Academic Advisor meeting (otherwise he can't register for classes again). However, he has a bit of extra time, so he decides to have a little bit of fun. At each step he will take either one stair or two stairs at a time. How many ways can he climb  $n$  stairs?

### 21.1 Solution:

Let  $F(x)$  denote the number ways to reach stair number  $x$ . By definition/intuition,  $F(0) = 1$  and  $F(1) = 1$ .  $F(2) = 2$  since Sahan could either go one stair twice or two stairs once.

In general for  $F(x)$ , we use the following formula:

$$F(x) = F(x - 1) + F(x - 2)$$

To reach stair  $x$ , Sahan can either reach stair  $x - 1$  and then climb one stair up, or he can reach stair  $x - 2$  and climb two stairs up.

This is the formula for the Fibonacci numbers! However, note that  $F(0) = 1$ , whereas the first Fibonacci number is usually regarded as 0. Therefore, the number of ways Sahan can climb  $n$  stairs is the  $n + 1$ 'th Fibonacci number.

## 22 Enigmatic Envelopes

Before you are two envelopes, of which you can only pick one. Both of the envelopes contain money, but one contains twice the amount compared to the other. You first choose one randomly, open it, and see that it contains  $\$Y$ . Should you then change your choice of envelope?

Grant uses the following reasoning:

The expected amount of money gained if you simply walk away is trivially  $\$Y$ . If you switch, half of the time you will walk away with  $\$Y/2$  (because half the time you picked the better envelope) and half the time you will walk away with  $2\$Y$  (because you picked the worse envelope). Adding everything up, switching is expected to give you  $\frac{5Y}{4}$  reward, so you should always switch.

Is Grant correct? Explain why or why not.

### 22.1 Solution:

Grant is not correct! This result defies all intuition. The answer should be “it doesn't matter”, since you technically didn't gain any information.

Let's say that after switching, some Men in Black come by and wipe your memory. You then look at the envelope you just switched to and you are given the option to switch again. The exact same argument will convince you to switch the envelopes. Given enough options and memory wipes, you'll be switching envelopes forever! A paradox!

Another way to view this problem is to instead define the values in the envelopes before you first pick. Let one envelope contain  $\$X$  and the other envelope contain  $2\$X$ . This means that if you switch, half the time you will gain  $\$X$ , while the other half of the time you will lose  $\$X$ . Therefore, the net gain is 0, so it won't matter whether you switch or not switch.

In fact, Grant's analysis is wrong. This is because the two envelopes are already locked in by the time you open an envelope. Therefore, “half the time the other envelope contains  $2\$Y$  and the other half of the time the other envelope contains  $\$Y/2$ ” is invalid - this just simply isn't true.