

1. A robot is located at the center of a square world that is 8 kilometers on each side. A package is dropped off in the robot's world at a point (x,y) that is uniformly (continuously) distributed in the square. If the robot's starting location is designated to be $(0,0)$ and the robot can only move up, down, left, and right parallel to the sides of the square, the distance the robot must travel to get to the package at point (x,y) is $|x| + |y|$. Let D be the distance the robot travels to get to the package. Compute $E[D]$.

X is the absolute distance from the robot to the package along the x axis

Y is the absolute distance from the robot to the package along the y axis

$$\begin{aligned} X &\sim \text{Uni}(\alpha, \beta) \text{ where } \alpha = 0, \beta = 4\text{km} \\ Y &\sim \text{Uni}(\alpha, \beta) \text{ where } \alpha = 0, \beta = 4\text{km} \\ E[D] &= E[X] + E[Y] \\ E[X] &= E[Y] = \frac{\alpha + \beta}{2} = 2\text{km} \\ E[D] &= 4\text{km} \end{aligned}$$

2. Say that of all the students who will attend Stanford, each will buy at most one laptop computer when they first arrive at school. 40% of students will purchase a PC, 30% will purchase a Mac, 10% will purchase a Linux machine, and the remaining 20% will not buy a laptop at all. If 15 students are asked which, if any, laptop they purchased, what is the probability that exactly 6 students will have purchased a PC, 4 will have purchased a Mac, 2 will have purchased a Linux machine, and the remaining 3 students will have not purchased a laptop?

$$P_{PC} = .4, P_{Mac} = .3, P_{Linux} = .1, P_{none} = .2$$

$$P_{\text{event occurs}} = \frac{15!}{6!4!2!3!} P_{PC}^6 P_{Mac}^4 P_{Linux}^2 P_{none}^3 = 1.674 \times 10^{-2}$$

3. Let X and Y be two independent random variables such that $X \sim \text{Geo}(p)$ and $Y \sim \text{Geo}(p)$. Mathematically derive an expression for $P(X = k | X + Y = n)$.

$$\begin{aligned} P(X = k | X + Y = n) &= \frac{P(X = k, Y = n - k)}{P(X + Y = n)} \\ &= \frac{P(X = k)P(Y = n - k)}{P(X + Y = n)} \\ P(X + Y = n) &= \sum_y P(X + Y = n, Y = y) && \text{by convolution} \\ &= \sum_y P(X = n - y)P(Y = y) \\ &= \sum_y (1 - p)^{n-y-1} p (1 - p)^{y-1} p \\ &= \sum_{y=1}^{n-1} p^2 (1 - p)^{n-2} \\ &= p^2 (1 - p)^{n-2} \sum_{y=1}^{n-1} 1 \\ &= p^2 (1 - p)^{n-2} (n - 1) \\ P(X = k | X + Y = n) &= \frac{(1 - p)^{k-1} p (1 - p)^{n-k-1} p}{p^2 (1 - p)^{n-2} (n - 1)} \\ &= \frac{1}{n - 1} \end{aligned}$$

4. Choose a number X uniformly at random from the set of numbers $\{1, 2, 3, 4, 5\}$. Now choose a number uniformly at random from the subset no larger than X , that is from $\{1, \dots, X\}$. Let Y denote the second number chosen.

(a) Determine the joint probability mass function of X and Y .

$$P(X = x, Y = y) = \begin{cases} \frac{1}{5} \frac{1}{x} & \text{if } y \leq x \\ 0 & \text{if } y > x \end{cases}$$

(b) Determine the conditional mass function of X given $Y = i$. Do this for $i = 1, 2, 3, 4, 5$.

$$P(X = x|Y = i) = \frac{P(X = x, Y = i)}{P(Y = i)}$$

$$P(Y = i) = \frac{1}{5} \sum_{k=i}^5 \frac{1}{k}$$

$$P(X = x|Y = i) = \frac{1}{x \sum_{k=1}^5 \frac{1}{k}}$$

$$P(X = x|Y = 1) = .438 \frac{1}{x}$$

$$P(X = x|Y = 2) = .779 \frac{1}{x}$$

$$P(X = x|Y = 3) = 1.277 \frac{1}{x}$$

$$P(X = x|Y = 4) = 2.222 \frac{1}{x}$$

$$P(X = x|Y = 5) = 5 \frac{1}{x}$$

(c) Are X and Y independent?

$$X \perp Y \text{ iff } \forall x, y : P(X = x, Y = y) = P(X = x)P(Y = y)$$

$$P(X = 2, Y = 1) = \frac{1}{10} \neq \frac{1}{5} * \frac{137}{300} = .091$$

$$\therefore X \not\perp Y$$

5. Let X be the outcome of a fair die roll. Let $Y = X^2$. What is the covariance $\text{Cov}(X, Y)$?

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

$$E[X] = \sum_{i=1}^6 i * \frac{1}{6} = 3.5$$

$$E[Y] = \sum_{i=1}^6 i^2 * \frac{1}{6} = 15.167$$

$$E[XY] = \sum_{i=1}^6 i^3 * \frac{1}{6} = 73.5$$

$$\text{Cov}(X, Y) = 20.42$$

6. You are tracking the distance to a satellite. An instrument reports that the satellite is 100 a.u. from Earth. Before you observed the instrument reading, your belief for the distance D of the satellite was a Gaussian $D \sim N(\mu = 98, \sigma^2 = 16)$. The instrument gives a reading that is true distance plus Gaussian noise with mean 0 and variance 4.

- (a) What is the PDF of your prior belief of the true distance of the satellite?

D is the true distance

N is the noise

R is the instrument reading

$$R = D + N$$

$$D \sim N(\mu_D = 98, \sigma_D^2 = 16)$$

$$N \sim N(\mu_N = 0, \sigma_N^2 = 4)$$

$$R \sim N(\mu_D + \mu_N = 98, \sigma_D^2 + \sigma_N^2 = 20)$$

$$f_{D|R}(d|r) = \frac{P(R=r|D=d)f_D(d)}{P(R=r)} \quad \text{by Bayes' Theorem}$$

$$f_D(d) = \frac{1}{\sigma_D \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{d-\mu_D}{\sigma_D})^2}$$

- (b) What is the probability density of seeing an observation of 100 a.u. from your instrument, given that the true distance of the satellite is equal to t ?

$$P(R=r|D=d) = f_N(n=r-d)$$

$$f_N(n) = \frac{1}{\sigma_N \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{n-\mu_N}{\sigma_N})^2}$$

$$P(R=100|D=t) = f_N(n=100-t)$$

$$f_N(n=100-t) = \frac{1}{\sigma_N \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{(100-t)-\mu_N}{\sigma_N})^2}$$

- (c) What is the PDF of your posterior belief (after observing the instrument reading) of the true distance of the satellite?

$$f_{D|R}(d|r) = \frac{1}{c} \cdot \frac{1}{2\pi\sigma_N\sigma_D} e^{-\frac{1}{2}((\frac{100-t-\mu_N}{\sigma_N})^2 + (\frac{d-\mu_D}{\sigma_D})^2)}$$

7. Consider the following function which simulates repeatedly rolling a 6-sided die (where each integer value from 1 to 6 is equally likely to be rolled) until a value ≥ 3 is rolled.

```
int roll()
{
    int total = 0;
    while (true) // infinite loop
    {
        int roll = randomInteger(1, 6);
        total += roll;
        if (roll >= 3) // exit condition
        {
            break;
        }
    }
    return total;
}
```

- (a) Let X be the value returned by the function. What is $E[X]$?

$$E[X] = 3(\frac{1}{4}) + 4(\frac{1}{4}) + 5(\frac{1}{4}) + 6(\frac{1}{4}) = \frac{9}{2}$$

- (b) Let Y be the number of times that the die is rolled (i.e. the number of times that a random number is generated) in the function. What is $E[Y]$?

$$Y \sim \text{Geo}(p = \frac{2}{3})$$

$$E[Y] = \frac{1}{p} = \frac{3}{2}$$

8. You go on a camping trip with two friends who each have a mobile phone. Since you are out in the wilderness, mobile phone reception isn't very good. One friend's phone will independently drop calls with 10% probability. Your other friend's phone will independently drop calls with 25% probability. Say you need to make 6 phone calls, so you randomly choose one of the two phones and you will use that same phone to make all your calls (but you don't know which has a 10% versus 25% chance of dropping calls). Of the first 3 (out of 6) calls you make, one of them is dropped. What is the conditional expected number of dropped calls in the 6 total calls you make (conditioned on already having one of the first 3 calls dropped)?

X is the number of dropped calls out of the 6 phone calls

Y is the number of dropped calls out of the first 3 calls

$p_{\text{phone},\alpha}$ is the probability of choosing phone α

$p_{\text{phone},\beta}$ is the probability of choosing phone β

$p_{\text{drop},\alpha}$ is the probability a call is dropped from phone α

$p_{\text{drop},\beta}$ is the probability a call is dropped from phone β

$$p_Y(y) = p_{\text{phone},\alpha} \cdot \binom{3}{y} p_{\text{drop},\alpha}^y (1 - p_{\text{drop},\alpha})^{3-y} + p_{\text{phone},\beta} \cdot \binom{3}{y} p_{\text{drop},\beta}^y (1 - p_{\text{drop},\beta})^{3-y}$$

$$p_{X,Y}(x, y) = p_{\text{phone},\alpha} \cdot \binom{3}{y} p_{\text{drop},\alpha}^y (1 - p_{\text{drop},\alpha})^{3-y} \cdot \binom{3}{x-y} p_{\text{drop},\alpha}^{x-y} (1 - p_{\text{drop},\alpha})^{3-(x-y)}$$

$$+ p_{\text{phone},\beta} \cdot \binom{3}{y} p_{\text{drop},\beta}^y (1 - p_{\text{drop},\beta})^{3-y} \cdot \binom{3}{x-y} p_{\text{drop},\beta}^{x-y} (1 - p_{\text{drop},\beta})^{3-(x-y)}$$

$$p_{X|Y}(x, y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

$$E[X|Y = y] = \sum_x x \cdot p_{X|Y}(x, y)$$

$$p_Y(Y = 1) = .332$$

$$E[X|Y = 1] = 1.586$$

9. Let X_1, X_2, X_3 , and X_4 be a set of pairwise uncorrelated random variables (i.e. $\rho(X_i, X_j) = 0$ when $i \neq j$), which all have the same mean, μ and variance, σ^2 .

An identity that will be useful for this problem is that $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$.

- (a) What is the correlation $\rho(X_1 + X_2, X_3 + X_4)$?

$$\rho(X_1 + X_2, X_3 + X_4) = \frac{\text{Cov}(X_1 + X_2, X_3 + X_4)}{\sqrt{\text{Var}(X_1 + X_2)\text{Var}(X_3 + X_4)}}$$

$$= \frac{\text{Cov}(X_1, X_3) + \text{Cov}(X_1, X_4) + \text{Cov}(X_2, X_3) + \text{Cov}(X_2, X_4)}{\sqrt{(\text{Var}(X_1) + \text{Var}(X_2) + 2\text{Cov}(X_1, X_2))(\text{Var}(X_3) + \text{Var}(X_4) + 2\text{Cov}(X_3, X_4))}}$$

$$= 0 \text{ since } \text{Cov}(X_i, X_j) = 0 \text{ when } i \neq j$$

(b) What is the correlation $\rho(X_1 + X_2, X_2 + X_3)$?

$$\begin{aligned}
 \rho(X_1 + X_2, X_2 + X_3) &= \frac{\text{Cov}(X_1 + X_2, X_2 + X_3)}{\sqrt{\text{Var}(X_1 + X_2)\text{Var}(X_2 + X_3)}} \\
 &= \frac{\text{Cov}(X_1, X_2) + \text{Cov}(X_1, X_3) + \text{Cov}(X_2, X_2) + \text{Cov}(X_2, X_3)}{\sqrt{(\text{Var}(X_1) + \text{Var}(X_2) + 2\text{Cov}(X_1, X_2))(\text{Var}(X_2) + \text{Var}(X_3) + 2\text{Cov}(X_2, X_3))}} \\
 &= \frac{\text{Var}(X_2)}{\sqrt{(\text{Var}(X_1) + \text{Var}(X_2))(\text{Var}(X_2) + \text{Var}(X_3))}} \\
 &= \frac{\sigma^2}{2\sigma^2} \\
 &= \frac{1}{2}
 \end{aligned}$$

(c) What is the correlation $\rho(2X_1, X_1 + X_2)$?

$$\begin{aligned}
 \rho(2X_1, X_1 + X_2) &= \frac{\text{Cov}(2X_1, X_1 + X_2)}{\sqrt{\text{Var}(2X_1)\text{Var}(X_1 + X_2)}} \\
 &= \frac{2\text{Cov}(X_1, X_1 + X_2)}{\sqrt{4\text{Var}(X_1)(\text{Var}(X_1) + \text{Var}(X_2) + 2\text{Cov}(X_1, X_2))}} \\
 &= \frac{2(\text{Cov}(X_1, X_1) + \text{Cov}(X_1, X_2))}{\sqrt{4\text{Var}(X_1)(\text{Var}(X_1) + \text{Var}(X_2) + 2\text{Cov}(X_1, X_2))}} \\
 &= \frac{2\text{Var}(X_1)}{\sqrt{4\text{Var}(X_1)(\text{Var}(X_1) + \text{Var}(X_2))}} \\
 &= \frac{1}{\sqrt{2}}
 \end{aligned}$$

10. In class, we considered the following recursive function.

```

int recurse()
{
    int x = randomInteger(1, 3);
    if (x == 1)
    {
        return 3;
    }
    else if (x == 2)
    {
        return (5 + recurse());
    }
    else
    {
        return (7 + recurse());
    }
}
    
```

Let Y be the value returned by `recurse()`. We previously computed $E[Y] = 15$. What is $\text{Var}(Y)$?

$$\text{Var}(Y) = E[Y^2] - E[Y]^2$$

$$E[Y^2] = E[E[Y^2|X]] = E[Y^2|X=1]P(X=1) + E[Y^2|X=2]P(X=2) + E[Y^2|X=3]P(X=3)$$

$$E[Y^2|X=1] = 9$$

$$E[Y^2|X=2] = E[(5+Y)^2] = E[25 + 10Y + Y^2] = 25 + 10E[Y] + E[Y^2]$$

$$E[Y^2|X=3] = E[(7+Y)^2] = E[49 + 14Y + Y^2] = 49 + 14E[Y] + E[Y^2]$$

$$E[Y^2] = 443$$

$$E[Y]^2 = 225$$

$$\text{Var}(Y) = 218$$

11. **[coding]** In this question, you are going to learn how to calculate p-values for experiments that are called A/B tests. These experiments are ubiquitous. They are a staple of both scientific and user interaction design.

Massive online courses have allowed for distributed experimentation into what practices optimize students' learning. Coursera, a free online education platform that started at Stanford is testing out new ways of teaching a concept in probability. They have 2 different learning activities, "activity1" and "activity2" and they want to figure out which activity leads to better learning outcomes. After interacting with a learning activity, Coursera evaluates a student's learning outcome by asking them to solve a set of questions. After a 2 week period, Coursera randomly assigns each student to either be given "activity1" (Group A) or "activity2" (Group B). The activity that is shown to each student and the student's measured learning outcomes can be found in the file "learningOutcomes.csv".

- (a) What is the difference in sample means of learning outcomes between students who were given "activity1" and students who were given "activity2"?

8.398

- (b) Calculate the p-value for the observed difference in means reported in (a). In other words, assuming the learning objectives for students who had been given "activity1" and "activity2" were identically distributed, what is the probability that you could have sampled 2 groups of students such that you could have observed a difference of means as extreme or more extreme than the one calculated from your data?

.049

- (c) The file "background.csv" stores the background of each user. Student backgrounds fall under 3 categories: more experience, average experience, and less experience. For each of the 3 backgrounds, calculate a difference in means in learning outcome between "activity1" and "activity2" and the p-value of that difference.

More experience: mean difference = 28.42 p-value = 0

Average experience: mean difference = 24.98 p-value = .0063

Less experience: mean difference = 26.02 p-value = .0006