1. In class, we discussed the Buffon's needle experiment. Let's explore a small extension here. A needle of length l is dropped randomly on a plane surface that is partitioned in rectangles by horizontal lines that are a apart and by vertical lines that are b apart. Suppose that $l < a$ and $l < b$. What is the expected number of rectangle sides crossed by the needle? What is the probability that the needle crosses at least one side of a rectangle?

We can model the problem similar to the original Buffon's Needle. Let the distance to the nearest horizontal line be denoted by y_{dist} and let the distance to the nearest vertical line be x_{dist}. Let θ represent the orientation of the needle. Like in class, we can bound these variables:

$$y_{\text{dist}} \leq \frac{a}{2}$$

$$x_{\text{dist}} \leq \frac{b}{2}$$

$$\theta \leq \pi$$

Without loss of generality (because the intersection of lines is rotation-invariant), we can say that we've crossed a rectangular line if either $y_{\text{dist}} \leq \frac{l}{2} sin\theta$ or $x_{\text{dist}} \leq \frac{l}{2} cos\theta$.

Since we went over it in class, we know that probability of crossing a horizontal line is given by:

$$\int_{\theta=0}^{\frac{\pi}{2}} \int_{y_{\text{dist}}=0}^{\frac{l}{2} sin\theta} \frac{4}{a\pi} \, \mathrm{d}y \, \mathrm{d}\theta = \frac{2l}{a\pi}$$

Similarly, for the crossing of a vertical line:

$$\int_{\theta=0}^{\frac{\pi}{2}} \int_{x_{\text{dist}}=0}^{\frac{l}{2} cos\theta} \frac{4}{b\pi} \, \mathrm{d}y \, \mathrm{d}\theta = \frac{2l}{b\pi}$$

Now, probability of crossing no lines is the complement of crossing at least one line. And that is the probability of crossing a horizontal line plus the probability of crossing a vertical line minus the probability of crossing both a horizontal and a vertical line. And we know that it's impossible to cross more than 2 lines since l is shorter than a and b.

The needle crosses both lines when both $y_{\text{dist}} \leq \frac{l}{2} sin\theta$ and $x_{\text{dist}} \leq \frac{l}{2} cos\theta$. So the probability of crossing both is given by:

$$\int_{\theta=0}^{\frac{\pi}{2}} \int_{y_{\text{dist}}=0}^{\frac{l}{2} sin\theta} \int_{x_{\text{dist}}=0}^{\frac{l}{2} cos\theta} \frac{8}{ab\pi} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}\theta$$

$$= \frac{8}{ab\pi} \int_{\theta=0}^{\frac{\pi}{2}} \frac{l^2}{4} \cos\theta \sin\theta \, \mathrm{d}\theta$$

$$= \frac{l^2}{ab\pi}$$

Finally, the probability of crossing at least one side is $\frac{2l}{a\pi} + \frac{2l}{a\pi} - \frac{l^2}{ab\pi} = \frac{l(2a+2b-l)}{ab\pi}$. So probability of crossing one line is:

$$P(1 \text{ lines}) = \frac{l(2a + 2b - l)}{ab\pi}$$

And probability of crossing two lines is:

$$P(2 \text{ lines}) = \frac{l^2}{ab\pi}$$

And the probability of no lines doesn't matter for expectation, since it has no weight. So the expected number of lines crossed is:

$$\frac{l(2a + 2b - l)}{ab\pi} + \frac{2l^2}{ab\pi}$$
$$= \frac{l(2a + 2b + l)}{ab\pi}$$

2. A scientist takes measurements of cell activity at discrete time intervals. Let X_i denote the measurement taken at time i, where the X_i are IID continuous random variables with pdf $f_X(x)$. We call X_k a record if $X_k > X_i$ for all $i < k$.

   (a) Find the probability that X_n is a record where $n \geq 1$.

   Since all orderings of $X_1, \cdots X_n$ are equally likely, and since we have a continuousdistribution, ties are impossible, so we know that the probability that X_n is a record is simply $\frac{1}{n}$.

3. To efficiently sell advertising space, Google uses a (generalized) second-price auction. In a second-price auction, the bidder with the highest bid wins the item, but only pays as much as the second highest bidder offered. Suppose we have a seller, Google, and n buyers. The buyer's bids are IID exponentials, where i bids $X_i \approx \text{Exp}(\lambda)$. Let P be the price of the item. Find the distribution of P.

   The probability of P taking the value x is if x is the second highest price out of all $X_1, \cdots X_n$. Concretely, this is the probability that $n - 2$ of the buyers have less than x, and that 1 of the n buyers has more than x. This can be calculated by the CDF, or the complement of the CDF. So let's say X_k is the second-highest buyer. Then probability that his price is x, and that x is the second-highest price is given by:

   $$P(P = x; k) = f(x) * ((n - 1)F(x)^{n-2})(1 - F(x))$$

   And since this event can't **just** happen to the kth buyer, we have to multiply this by n, because all buyers are equally likely to be the second-highest bid. So we have:

   $$P(P = x) = n(n - 1)f(x)F(x)^{n-2}(1 - F(x))$$

4. n graduate students at Berkeley Research Lab come to work by bicycles every day, parking their bikes in an unlocked bike rack in their lab. On a given day, after a hard day's work, the students begin leaving one by one, taking their bikes out of the rack on their way out. The first student has forgotten what his bike looks like, and picks up one of the n bikes from the rack uniformly at random. Subsequently the other $(n - 1)$ students who leave one by one, follow the "honest if possible" policy of picking up their bike if it is there, else picking a bike uniformly at random from the remaining collection in the rack.

   (a) What is the probability that the first student finds his/her own bike?

   Since the first student picks a bike uniformly out of n bikes, we can just count to find that the probability is $\frac{1}{n}$.

(b) List the bikes that the last student may possibly take. What is the probability that athe last student finds his/her bike?

We know that the last student can only take the first or last bike. This is because for any bike belonging to student i, (except for the first student,) student i has already had a chance to try to get their bike back. If it hadn't already been stolen, they would've taken it.

For the students before the last student, the probability of them picking the first or the last bike is the same. If their bike is on the rack, the probabilities are 0, and if their bike has been stolen, they are equally likely to pick the first or last bike. Since there are only 2 possibilities for the last student, and they are equally likely, by symmetry, we know that the probability he gets his bike back is 0.5.

5. Consider random variables X and Y have a joint PDF uniform on the triangle with vertices at $(0,0), (1,0), (0,1)$.

   (a) Find the joint PDF of X and Y.

   Since every point is equally likely, and the area of the triangle is $\frac{1}{2}$, we have:

   $$f_{X,Y}(x,y) = 2, (x,y) \in \text{triangle}$$

   $$f_{X,Y}(x,y) = 0, \text{ else}$$

   (b) Find the marginal PDF of Y.

   $$f_Y(y) = \int_{x=\infty}^{\infty} 2dx$$

   $$= \int_{x=0}^{1-y} 2dx = 2 - 2y$$

   (c) Find the conditional PDF of X given Y.

   $$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y))}{f_Y(y)}$$

   $$= \frac{2}{2-2y} = \frac{1}{1-y}, \text{ for } x \leq y$$

   (d) Find $\mathbb{E}[X]$ in terms of $\mathbb{E}[Y]$

   By geometry, we can see that $\mathbb{E}[X|Y=y] = \frac{1-y}{2}$. By the total expectation theorem, we see:

   $$\mathbb{E}[X] = \int_{y=-\infty}^{\infty} f_Y(y) \frac{1-y}{2} dy$$

   $$= \int_{y=0}^{1} f_Y(y) \frac{1-y}{2} dy$$

   Distributing and splitting up the integrals we get:

   $$= \int_{y=0}^{1} \frac{f_Y(y)}{2} - \frac{f_Y(y)y}{2} dy$$

$$= \int_{y=0}^{1} \frac{f_Y(y)}{2} - \int_{y=0}^{1} \frac{f_Y(y)y}{2} \, dy$$

Notice that the second term is just the expectation of Y, so we have

$$= \frac{1}{2} \int_{y=0}^{1} f_Y(y) - \frac{1}{2} \mathbb{E}[Y]$$

Since the integral integrates over every possible value of y, we know that the value of the integral must be 1. So we end up with

$$\mathbb{E}[X] = \frac{1}{2} - \frac{1}{2} \mathbb{E}[Y]$$

(e) Find $\mathbb{E}[X]$.

Since the triangle is symmetric along the line $y = x$, and it's a uniform distribution, we know that $\mathbb{E}[X] = \mathbb{E}[Y]$. Then, we have:

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

$$\frac{3}{2} \mathbb{E}[X] = \frac{1}{2}$$

$$\mathbb{E}[X] = \frac{1}{3}$$

6. Figure 1 (not shown) shows the joint density $f_{X,Y}$ of random variables X and Y.

(a) Find A and sketch f_X, f_Y, and $f_{X|X+Y\leq 3}$

Since the total sum of the probabilities is 4A, which has to be 1, we know that $A = \frac{1}{4}$.

(b) Find $\mathbb{E}[X|Y = y]$ for $1 \leq y \leq 3$ and $\mathbb{E}[Y|X = x]$ for $1 \leq x \leq 4$

By the total expectation theorem, this is just $\mathbb{E}[X]$. Since the object is symmetric along $x = 2.5$, we know that the expected value of x is just 2.5.

Similar reasoning with $\mathbb{E}[Y]$. Except we don't have symmetry anymore. But above $y = 2$, there's a weight of $3A$ with the center being at $y = 2.5$, and below $y = 2$ there's a weight of A with center point at $y = 1.5$. So in total:

$$\mathbb{E}[Y] = \frac{2.5 * 3A + 1.5 * A}{4A} = 2.25$$

(c) Find $\operatorname{cov}(X, Y)$.

$$\operatorname{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$