Problem 1: Let A and B be two subsets of a set S. Let C be the subset of S consisting of elements that are in A or B, but not both. Express C in terms of A and B using only the basic set operations.

We have
$$C = (A \cup B) \smallsetminus (A \cap B).$$

Problem 2: Draw Venn diagrams to illustrate DeMorgan's law, which states that, given two subsets A and B of a set S, that

$$(A \cup B)^c = A^c \cap B^c.$$

We'll draw the Venn diagram in class.

Problem 3: Identify appropriate sample spaces for each of the following scenarios.

(a) Rolling a die.

$S = \{1, 2, 3, 4, 5, 6\}$

(b) Flipping a coin.

$S = \{H, T\}$, where H and T stand for "heads" and "tails," respectively.

(c) Flipping a coin *three* times.

$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

(d) Flipping a coin until it lands heads.

$S = \{H, TH, TTH, TTTH, TTTTH, \ldots\}$

(e) Forecasting whether it will rain tomorrow.

$S = \{\text{rain}, \text{no rain}\}$

(f) Conducting an experiment to measure the speed of light.

$S = [0, \infty)$. Theoretically, any non-negative real number could be obtained.

Problem 4: Describe some events in the following scenarios. Use the sample spaces that you identified in the previous problem.

(a) Rolling a die.

The event that we roll an even number is the subset
$$A = \{2, 4, 6\} \subset S.$$

(b) Flipping a coin.

The only subsets of S are
$$\emptyset, H, T, S,$$
so there aren't really any interesting events besides the simple ones.

(c) Flipping a coin *three* times.

The event that we flip exactly two heads is the subset
$$A = \{HHT, HTH, THH\} \subset S.$$

(d) Flipping a coin until it lands heads.

The event that we flip an even number of tails before we see a head corresponds to the subset
$$A = \{H, TTH, TTTTH, TTTTTTH, \ldots\} \subset S.$$

(e) Forecasting whether it will rain tomorrow.

The only subsets of S are
$$\emptyset, \text{rain}, \text{no rain}, S,$$
so, there aren't really any interesting events besides the simple ones.

(f) Conducting an experiment to measure the speed of light.

The event that you measure a number between 600 and 1000 (inclusive) is the subset
$$A = [600, 1000].$$

Problem 5: Suppose that you choose at random one of the twelve months of the year.
(a) Identify an appropriate sample space.

The sample space is the set of all twelve months:
$$S = \{\text{January}, \text{February}, \ldots, \text{December}\}.$$

(b) Identify the event that you choose a month beginning with a "J".

This is the event
$$A = \{\text{January}, \text{June}, \text{July}\}.$$

(c) Identify the event that you choose a month whose name is four letters long.

This is the event
$$B = \{\text{June}, \text{July}\}.$$

(d) Using one of the basic set operations and your answers to the two previous parts, identify the event that you choose a month beginning with a "J" and whose name is four letters long.

This is the event
$$A \cap B = \{\text{June}\}.$$

Problem 6: Suppose that A and B are two events. Write expressions involving only the basic set operations for each of the following:

(a) Both events occur. $A \cap B$
(b) At least one occurs. $A \cup B$
(c) Neither occurs. $S \smallsetminus (A \cup B)$
(d) Exactly one occurs. $(A \cup B) \smallsetminus (A \cap B)$

Problem 7: Let P be the discrete probability measure on the sample space $S = \mathbb{R}$ with probability mass function

$$p(s) = \begin{cases} 1/4 & : s = 0, \\ 1/4 & : s = 2, \\ 1/2 & : s = 4, \\ 0 & : \text{otherwise.} \end{cases}$$

Compute the probabilities $P(A)$ of the following events.

(a) $A = [-2, -1]$

We have
$$P(A) = 0,$$
since A does not contain $s = 0$, 2, or 4.

(b) $A = (-1, 1)$

We have
$$P(A) = 1/4,$$
since A contains $s = 0$.

(c) $A = (-1, 1) \cup (3, 5]$

We have
$$P(A) = 1/4 + 1/2 = 3/4,$$
since A contains $s = 0$ and 4.

(d) $A = \mathbb{R}$

We have
$$P(A) = 1/4 + 1/4 + 1/2 = 1,$$
since A contains $s = 0$, 2, and 4.

Problem 8: Let P be the discrete probability measure on the sample space $S = \mathbb{R}$ with probability mass function

$$p(s) = \begin{cases} (0.25)(0.75)^{s-1} & : s = 1, 2, 3, \ldots, \\ 0 & : \text{otherwise.} \end{cases}$$

Compute the probabilities $P(A)$ of the following events.

(a) $A = (-\infty, 1)$

We have
$$P(A) = 0,$$
since A does not contain any positive integer.

(b) $A = (-10, 4]$

We have
$$P(A) = 0.25 + (0.25)(0.75) + (0.25)(0.75)^2 + (0.25)(0.75)^3 \approx 0.68,$$
since A contains $n = 1, 2, 3, 4$.

(c) $A = \mathbb{R}$

We have
$$P(A) = \sum_{s=1}^{\infty} (0.25)(0.75)^{s-1} = \frac{0.25}{0.75} \cdot \frac{0.75}{1 - 0.75} = 1,$$
since A contains all positive integers.

(d) $A = \{\text{all positive even integers}\}$

We have
$$P(A) = \sum_{s=1}^{\infty} (0.25)(0.75)^{2s-1} = \frac{0.25}{0.75} \cdot \frac{(0.75)^2}{1 - (0.75)^2} \approx 0.43.$$

Problem 9: Describe in complete detail a probability space that models each of the following scenarios. Be sure to check that the two requirements in the Discrete Probability Construction Lemma are both satisfied.

(a) Rolling a fair six-sided die.

Remember, there are three things that go into a probability space: The sample space, the collection of events, and the probability measure. For this problem, the sample space is
$$S = \{1, 2, 3, 4, 5, 6\}$$
where each integer $s \in S$ represents an outcome of rolling the die. The collection of events will be *all* subsets. The probability measure P is uniform with probability function given by
$$p(s) = 1/6,$$
for each $s \in S$. Since these probabilities sum to 1 over all six sample points, the Construction Lemma guarantees that we have a valid probability measure.

(b) Flipping an **unfair** coin, with probability 0.25 of obtaining heads, and probability 0.75 of obtaining tails.

The probability space has sample space

$$S = \{H, T\}$$

where H represents the outcome of obtaining heads and T the outcome of obtaining tails. The collection of events is *all* subsets. The probability measure is discrete with probability function given by

$$p(H) = 0.25 \quad \text{and} \quad p(T) = 0.75.$$

Since $0.25 + 0.75 = 1$, the Construction Lemma guarantees that we have a valid probability measure.

(c) Flipping an **unfair** coin twice, with probability 0.25 of obtaining heads, and probability 0.75 of obtaining tails. What is the probability that you flip *exactly* one head?

The probability space has sample space

$$S = \{HH, HT, TH, TT\}$$

where H represents the outcome of obtaining heads and T the outcome of obtaining tails. The collection of events is *all* subsets. The probability measure is discrete with probability function given by

$$p(HH) = (0.25)^2 \quad p(HT) = (0.25)(0.75), \quad p(TH) = (0.25)(0.75), \quad p(TT) = (0.75)^2.$$

Since

$$(0.25)^2 + (0.25)(0.75) + (0.25)(0.75) + (0.75)^2 = 1,$$

the Construction Lemma guarantees that we have a valid probability measure. The probability that we flip *exactly* one head is

$$P(\{HT, TH\}) = p(HT) + p(TH) = 2(0.25)(0.75) = 0.375.$$

(d) Flipping an **unfair** fair coin until you obtain a head, with probability 0.25 of obtaining heads, and probability 0.75 of obtaining tails. What is the probability that you flip an odd number of tails before you see a head?

The probability space has sample space

$$S = \{H, TH, TTH, TTTH, \ldots\}.$$

The collection of events is *all* subsets. The probability measure is discrete with probability function given by

$$p(\underbrace{T \cdots TH}_{n \text{ flips}}) = (0.25)(0.75)^{n-1}.$$

But is this *really* a valid probability measure? If you remember your rules for infinite series from calculus, you'll notice that:

$$\sum_{n=1}^{\infty} (0.25)(0.75)^{n-1} = \frac{0.25}{0.75} \cdot \frac{0.75}{1 - 0.75} = 1,$$

so indeed the Construction Lemma guarantees that we obtain a valid probability measure. The probability that we flip an odd number of tails before we see a head is

$$\sum_{n=1}^{\infty}(0.25)(0.75)^{2n-1} = \frac{0.25}{0.75} \cdot \frac{(0.75)^2}{1-(0.75)^2} \approx 0.43.$$

Problem 10: A sample space consists of five simple events, E_1, E_2, E_3, E_4, and E_5.

(a) If $P(E_1) = P(E_2) = 0.15$, $P(E_3) = 0.4$, and $P(E_4) = 2P(E_5)$, find the probabilities of E_4 and E_5.

We know that the total probability must sum to 1, so

$$1 = \sum_{i=1}^{5} P(E_i) = 0.15 + 0.15 + 0.4 + P(E_4) + 2P(E_4).$$

Solving this equation for $P(E_4)$ gives 0.1, from which it follows that $P(E_5) = 0.2$.

(b) If $P(E_1) = 3P(E_2) = 0.3$, find the probabilities of the remaining simple events if you know that the remaining simple events are equally probable.

Same strategy as the first part. We know that

$$1 = \sum_{i=1}^{5} P(E_i) = 0.3 + 0.3/3 + P(E_3) + P(E_3) + P(E_3),$$

and solving this for $P(E_3)$ yields 0.2. But then this must also be the probability $P(E_4)$ and $P(E_5)$, as well.

Problem 11: A survey classified a large number of adults according to whether they were diagnosed as needing eyeglasses to correct their reading vision and whether they use eyeglasses when reading. The proportions falling into the four resulting categories are given in the following table:

	Uses Eyeglasses for Reading	
Needs glasses	Yes	No
Yes	.44	.14
No	.02	.40

If a single adult is selected from the large group, find the probabilities of the events defined below. The adult

(a) needs glasses.

Since the adult is selected at random, any one adult has the same chance as being selected as any other. Therefore, the proportions listed in the table *are* the probabilities that a randomly chosen adult falls into one of the four groups in the table. Therefore, for part (a), the probability is $0.44 + 0.14 = 0.58$, which is the sum of the probabilities of the two groups that both need glasses.

(b) needs glasses but does not use them.

0.14

(c) uses glasses whether the glasses are needed or not.

$$0.44 + 0.02 = 0.46$$

Problem 12:   Let P be a continuous probability measure on the sample space $S = \mathbb{R}$ with probability density function

$$f(s) = \begin{cases} \frac{2}{9}s & : 0 \leq s \leq 3, \\ 0 & : \text{otherwise.} \end{cases}$$

Compute the probabilities $P(A)$ of the following events.

(a) $A = (-1, 1]$

We have:
$$P(A) = \int_{-1}^{1} f(s) \; \mathrm{d}s = \frac{2}{9} \int_{0}^{1} s \; \mathrm{d}s = \frac{1}{9}.$$

(b) $A = (-1, 1) \cup [2, 4)$.

We have:
$$P(A) = \int_{-1}^{1} f(s) \; \mathrm{d}s + \int_{2}^{4} f(s) \; \mathrm{d}s = \frac{2}{9} \int_{0}^{1} s \; \mathrm{d}s + \frac{2}{9} \int_{2}^{3} s \; \mathrm{d}s = \frac{2}{3}.$$

(c) $A = [0, 3]$.

We have:
$$P(A) = \int_{0}^{3} f(s) \; \mathrm{d}s = \frac{2}{9} \int_{0}^{3} s \; \mathrm{d}s = 1.$$

(d) $A = \mathbb{R}$.

We have:
$$P(A) = \int_{-\infty}^{\infty} f(s) \; \mathrm{d}s = \frac{2}{9} \int_{0}^{3} s \; \mathrm{d}s = 1.$$

Problem 13:   Let P be a continuous probability measure on the sample space $S = \mathbb{R}$ with probability density function

$$f(s) = \begin{cases} \dfrac{1}{s^2} & : s \geq 1, \\ 0 & : \text{otherwise.} \end{cases}$$

Compute the probabilities $P(A)$ of the following events.

(a) $A = (-10, 0)$

We have:
$$P(A) = \int_{-10}^{0} f(s) \; \mathrm{d}s = 0.$$

(b) $A = \mathbb{R}$.

We have:

$$P(A) = \int_{-\infty}^{\infty} f(s) \, ds = \int_{1}^{\infty} \frac{1}{s^2} \, ds = \left(\lim_{s \to \infty} \frac{1}{s} \right) + 1 = 1.$$

Problem 14: A particle is located somewhere along the real line $\mathbb{R}$, but the only information that you have about its location is that it is at some point in the interval $[0, 4]$.

(a) Describe in complete detail a probability space that models this scenario.

This scenario may be modeled by a continuous probability distribution P. We may define P by using the Continuous Probability Construction Lemma by defining its probability density function. Since we know that the particle can only be in the interval $[0, 4]$, we set $f(s) = 0$ for all s outside this interval. In the absence of any more information about its location in $[0, 4]$, it makes sense to assume that the density function $f(y)$ is uniform (i.e., constant) across this interval. (This is called a *uniform* or *flat prior*, in "bayesian" lingo.) Since we must have

$$\int_{-\infty}^{\infty} f(s) \, ds = 1,$$

our density function must therefore be given by

$$f(s) = \begin{cases} \frac{1}{4} & : 0 \le s \le 4, \\ 0 & : \text{otherwise.} \end{cases}$$

(b) Suppose some new information comes to light that suggests the particle is closer to 4 than it is 0. Alter your probability model to reflect this new information.

Let's change the density function $f(s)$ so that it has higher values toward $s = 4$, remembering that we must still have

$$\int_{-\infty}^{\infty} f(s) \, ds = 1.$$

The following function does the job:

$$f(s) = \begin{cases} \frac{1}{8}s & : 0 \le s \le 4, \\ 0 & : \text{otherwise.} \end{cases}$$

Problem 15: Let P be the discrete probability measure on $\mathbb{R}$ with probability function

$$p(s) = \begin{cases} \frac{2!}{s!(2-s)!} \left(\frac{1}{4} \right)^s \left(\frac{3}{4} \right)^{2-s} & : s = 0, 1, 2, \\ 0 & : \text{otherwise.} \end{cases}$$

Compute the distribution function $F(s)$.

We have:

$$F(s) = \begin{cases} 0 & : s < 0, \\ 9/16 & : 0 \le s < 1, \\ 15/16 & : 1 \le s < 2, \\ 1 & : 2 \le s. \end{cases}$$

Problem 16: Let P be the continuous probability measure on $\mathbb{R}$ with density function

$$f(s) = \begin{cases} 3s^2 & : 0 \leq s \leq 1, \\ 0 & : \text{otherwise.} \end{cases}$$

Compute the distribution function $F(s)$.

We have:

$$F(s) = \begin{cases} 0 & : s < 0, \\ s^3 & : 0 \leq s \leq 1, \\ 1 & : 1 < s. \end{cases}$$

Problem 17: Let P be a probability measure on $\mathbb{R}$ with distribution function

$$F(s) = \begin{cases} 0 & : s < 0, \\ s & : 0 \leq s \leq 1, \\ 1 & : s > 1. \end{cases}$$

Is P continuous? If so, compute its density function $f(s)$.

Yes, the measure P is continuous since $F(s)$ is continuous. To compute the density $f(s)$, we use the Fundamental Theorem of Calculus:

$$f(s) = F'(s) = \begin{cases} 0 & : s < 0, \\ 1 & : 0 < s < 1, \\ 0 & : s > 1. \end{cases}$$

Note that the density $f(s)$ is undefined at $s = 0, 1$, since $F(s)$ is differentiable everywhere except these points. This is typical of these types of problems. If we want to make $f(s)$ defined on the *entire* real line $\mathbb{R}$, then we simply define $f(0)$ and $f(1)$ to be *any* particular pair of numbers. It does not matter which numbers we choose, since the choice will not affect the definite integrals $\int_a^b f(s) \, ds$, which is how we measure probabilities.

Problem 18: Let P be the discrete probability measure on $\mathbb{R}$ with probability function

$$p(s) = \begin{cases} 3/8 & : s = 1, \\ 7/16 & : s = 2, \\ 3/16 & : s = 3, \\ 0 & : \text{otherwise.} \end{cases}$$

Compute the following quantiles:

(a) $Q(0.25)$

Since $F(s) = 0$ for all $s < 1$ and $F(1) = 3/8 > 0.25$, we have $Q(0.25) = 1$.

(b) $Q(0.5)$ (the median of P)

Since $F(s) = 3/8 < 0.5$ for all $s < 2$ and $F(2) = 3/8 + 7/16 > 0.5$, we have $Q(0.5) = 2$.

(c) $Q(0.75)$

Since $F(s) = 3/8 < 0.75$ for all $s < 2$ and $F(2) = 3/8 + 7/16 > 0.75$, we have $Q(0.75) = 2$.

Problem 19: Let P be the continuous probability measure on $\mathbb{R}$ with density function

$$f(s) = \begin{cases} cs(1-s) & : 0 \leq s \leq 1, \\ 0 & : \text{otherwise.} \end{cases}$$

(a) Find the value of c that makes $f(s)$ a valid density function.

We must have

$$\int_{-\infty}^{\infty} f(s) \, ds = 1.$$

But

$$\int_{-\infty}^{\infty} f(s) \, ds = \int_0^1 cs(1-s) \, ds = \frac{c}{6},$$

and hence we must have $c = 6$.

(b) Compute the quantile $Q(0.95)$.

We compute the value $s = Q(0.95)$ by solving $F(s) = 0.95$ for s. To do this, we first compute a formula for the distribution function $F(s)$:

$$F(s) = \int_{-\infty}^{s} f(t) \, dt = \int_0^s 6t(1-t) \, dt = 3s^2 - 2s^3,$$

provided that $s \geq 0$. Then, using a computer, we solve:

$$F(s) = 0.95 \quad \Rightarrow \quad 3s^2 - 2s^3 = 0.95 \quad \Rightarrow \quad s \approx 0.86.$$

Thus, $Q(0.95) \approx 0.86$.

Problem 20: Let P be a continuous probability measure on the sample space $S = \mathbb{R}^2$ with probability density function

$$f(s,t) = \begin{cases} \frac{1}{8}(s+t) & : 0 \leq s, t \leq 2, \\ 0 & : \text{otherwise.} \end{cases}$$

Compute the probabilities $P(C)$ of the following events.

(a) $C = [-1, 1] \times (-1, 1)$

We have:

$$P(C) = \int_{-1}^{1} \int_{-1}^{1} f(s,t) \, dsdt = \frac{1}{8} \int_0^1 \int_0^1 (s+t) \, dsdt = \frac{1}{8}.$$

(b) $C = \mathbb{R}^2$.

We have:

$$P(C) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s,t) \, dsdt = \frac{1}{8} \int_0^2 \int_0^2 (s+t) \, dsdt = 1.$$

Problem 21:   Let P be a continuous probability measure on the sample space $S = \mathbb{R}^2$ with probability density function

$$f(s,t) = \begin{cases} e^{-(s+t)} & : s, t > 0, \\ 0 & : \text{otherwise.} \end{cases}$$

Show that $P(\mathbb{R}^2) = 1$.

We have

$$\begin{aligned} P(\mathbb{R}^2) &= \int_0^\infty \int_0^\infty e^{-(s+t)} \, \mathrm{d}s\mathrm{d}t \\ &= \int_0^\infty \left[\left(\lim_{s\to\infty} -e^{-(s+t)} \right) + e^{-t} \right] \mathrm{d}t \\ &= \int_0^\infty e^{-t} \, \mathrm{d}t \\ &= \left(\lim_{t\to\infty} -e^{-t} \right) + 1 \\ &= 1. \end{aligned}$$

Problem 22:   A particle is now located somewhere in the plane $\mathbb{R}^2$, but the only information that you have about its location is that it is at some point in the square

$$R = [0,3] \times [0,3].$$

(a) Describe in complete detail a probability space that models this scenario.

As in the first version of this problem, this scenario may be modeled by a continuous probability distribution P. We may define P by using the Continuous Probability Construction Lemma by defining its probability density function. Since we know that the particle can only be in the square R, we set $f(s,t) = 0$ for all (s,t) outside this square. In the absence of any more information about its location in R, it makes sense to assume that the density function $f(s,t)$ is uniform across this square. Since we must have

$$\int_{-\infty}^\infty \int_{-\infty}^\infty f(s,t) \, \mathrm{d}s\mathrm{d}t = 1,$$

our density function must therefore be given by

$$f(s,t) = \begin{cases} \frac{1}{9} & : 0 \leq s, t \leq 3, \\ 0 & : \text{otherwise.} \end{cases}$$

(b) Suppose some new information comes to light that suggests the particle is closer to the corner $(2,2)$ than to the origin $(0,0)$. Alter your probability model to reflect this new information.

This is just like the first version of this problem. You may check that the following function works:

$$f(s,t) = \begin{cases} \frac{1}{27}(s+t) & : 0 \leq s, t \leq 3, \\ 0 & : \text{otherwise.} \end{cases}$$