CSE 312 Section 5

More Random Variable and Discrete Zoo

Administrivia

Announcements & Reminders

- HW3
  - Grades released on gradescope – check your submission to read comments
  - Regrade requests open ~24 hours after grades are released and close after a week

- HW4
  - Written is due tomorrow, Friday 2/7 @ 11:59pm
  - Late deadline Sunday 2/9 @ 11:59pm

- HW5
  - Will be released on the course website today
  - Due Friday 2/14 @ 11:59pm
  - Late deadline Sunday 2/16 @ 11:59pm

Review & Questions

Any lingering questions from this last week?

Each week in section, we'll be reviewing the main concepts from this week and putting them into action by going through some practice problems together. But before we get into that review, we'll try to start off each section with some time for you to ask questions. Was anything particularly confusing this week? Is there anything we can clarify before we dive into the review? This is your chance to clear things up!

Review of Main Concepts

- **Expectation (expected value, mean, or average):** The expectation of a discrete random variable is defined to be
$$\mathbb{E}[X] = \sum_x x\, p_X(x) = \sum_x x\, \mathbb{P}(X = x)$$
The expectation of a function of a discrete random variable $g(X)$ is
$$\mathbb{E}[g(X)] = \sum_x g(x)\, p_X(x)$$

- **Linearity of Expectation:** Let X and Y be random variables, and $a, b, c \in \mathbb{R}$. Then, $\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$. Also, for any random variables $X_1, \ldots, X_n$,
$$\mathbb{E}[X_1 + X_2 + \cdots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n]$$

Review of Main Concepts

- **Variance:** Let X be a random variable and $\mu = E[X]$. The variance of X is defined to be

$$Var(X) = \mathbb{E}[(X-\mu)^2]$$

Notice that since this is an expectation of a non-negative random variable ($(X-\mu)^2$), variance is always nonnegative. With some algebra, we can simplify this to

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

- **Standard Deviation:** Let X be a random variable. We define the standard deviation of X to be the square root of the variance, and denote it $\sigma = \sqrt{Var(X)}$
- **Property of Variance:** Let $a, b \in \mathbb{R}$ and let X be a random variable. Then,

$$Var(aX + b) = a^2 Var(X)$$

Review of Main Concepts

- **Independence:** Random variables X and Y are independent iff

$$\forall x \forall y, \mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x)\,\mathbb{P}(Y = y)$$

  In this case, we have $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ (the converse is not necessarily true).

- **i.i.d. (independent and identically distributed):** Random variables $X_1, \ldots, X_n$ are i.i.d. (or iid) iff they are independent and have the same probability mass function.

- **Variance of Independent Variables:** If X is independent of Y,

$$Var(X + Y) = Var(X) + Var(Y)$$

  This depends on independence, whereas linearity of expectation always holds. Note that this combined with the above shows that $\forall a, b, c \in \mathbb{R}$ and if X is independent of Y,

$$Var(aX + bY + c) = a^2 Var(X) + b^2 Var(Y)$$

Review Questions

Question #1

$$Var(A + B) = Var(A) + Var(B)$$

- True
- False

Question #1

$$Var(A + B) = Var(A) + Var(B)$$

- True
- False

Question #2

What is $Var(3A+4)$?

- $3Var(A)+4$
- $3Var(A)$
- $9Var(A)$
- $Var(A)$

Question #2

What is $Var(3A + 4)$?

- $3Var(A) + 4$
- $3Var(A)$
- $9Var(A)$
- $Var(A)$

Question #3

$$E[A + B] = E[A] + E[B]$$

- True
- False

Question #3

$$E[A + B] = E[A] + E[B]$$

- True
- False

Question #4

What is $E[3A+4]$?

- $3E[A]+4$
- $3E[A]$
- $9E[A]$
- $E[A]$

Question #4

What is $E[3A + 4]$?

- $3E[A] + 4$
- $3E[A]$
- $9E[A]$
- $E[A]$

Problem 2 - Pond fishing

Part a

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + R + G = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

How many of the next 10 fish I catch are blue, if I catch and release

- $$\mathsf{Bin}\left(10, \frac{B}{N}\right)$$
- $$\mathsf{Ber}\left(\frac{B}{N}\right)$$
- $$\mathsf{Bin}\left(1, \frac{B}{N}\right)$$

Part a

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + R + G = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

How many of the next 10 fish I catch are blue, if I catch and release

- $$\text{Bin}\left(10, \frac{B}{N}\right)$$

- $$\text{Ber}\left(\frac{B}{N}\right)$$

- $$\text{Bin}\left(1, \frac{B}{N}\right)$$

Part b

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + R + G = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

How many fish I had to catch until my first green fish, if I catch and release

- $$\text{Ber}\left(\frac{G}{N}\right)$$
- $$\text{Bin}\left(1, \frac{G}{N}\right)$$
- $$\text{Geo}\left(\frac{G}{N}\right)$$

Part b

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + R + G = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

How many fish I had to catch until my first green fish, if I catch and release

- $$\text{Ber}\left(\frac{G}{N}\right)$$
- $$\text{Bin}\left(1, \frac{G}{N}\right)$$
- $$\text{Geo}\left(\frac{G}{N}\right)$$

Part c

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + R + G = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

How many red fish I catch in the next five minutes, if I catch on average r red fish per minute

- $$\text{Poi}(5R)$$
- $$\text{Bin}\left(5, \frac{R}{N}\right)$$
- $$\text{Poi}(5r)$$

Part c

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + R + G = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

How many red fish I catch in the next five minutes, if I catch on average r red fish per minute

- $$\text{Poi}(5R)$$
- $$\text{Bin}\left(5, \frac{R}{N}\right)$$
- $$\text{Poi}(5r)$$

Part d

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + R + G = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

Whether or not my next fish is blue

- $$\text{Poi}(5B)$$
- $$\text{Bin}\left(1, \frac{R}{N}\right)$$
- $$\text{Ber}\left(\frac{B}{N}\right)$$

Part d

Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + R + G = N$. Each fish is equally likely to be caught. For each of the following scenarios, identify the most appropriate distribution (with parameter(s)):

Whether or not my next fish is blue

- $\mathsf{Poi}(5B)$
- $\mathsf{Bin}\left(1, \dfrac{R}{N}\right)$
- $\mathsf{Ber}\left(\dfrac{B}{N}\right)$

Problem 3 - Balls and Bins

Problem 3 - Balls and Bins

Let X be the number of bins that remain empty when m balls are distributed into n bins randomly and independently. For each ball, each bin has an equal probability of being chosen. (Notice that two bins being empty are not independent events: if one bin is empty, that decreases the probability that the second bin will also be empty. This is particularly obvious when $n = 2$ and $m > 0$.) Find $\mathbb{E}[X]$.

Work on this problem with the people around you, and then we'll go over it together!

Problem 3 - Balls and Bins

Let X be the number of bins that remain empty when m balls are distributed into n bins randomly and independently. For each ball, each bin has an equal probability of being chosen. (Notice that two bins being empty are not independent events: if one bin is empty, that decreases the probability that the second bin will also be empty. This is particularly obvious when $n = 2$ and $m > 0$.) Find $\mathbb{E}[X]$.

Let X_i be 1 if bin i is empty, and 0 otherwise.

$$X = \sum_{i=1}^{n} X_i$$

$$\mathbb{E}[X_i] = 1 * \mathbb{P}(X_i = 1) + 0 * \mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = \left(\frac{n-1}{n}\right)^m$$

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n * \left(\frac{n-1}{n}\right)^m$$

Problem 4 - 3 Sided Die

Problem 4 - 3 Sided Die

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(a) What is the PMF of X?

(b) Find $E[X]$.

(c) Find $Var(X)$.

Work on this problem with the people around you, and then we'll go over it together!

Problem 4 - 3 Sided Die

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(a) What is the PMF of X?

First let us define the range of X. A three sided-die can take on values {1,2,3}. Since X is the sum of two rolls, the range of X is $\Omega_X = \{2, 3, 4, 5, 6\}$.

We must define two random variables R_1, R_2 with R_1 being the roll of the first die, and R_2 being the roll of the second die. Then, $X = R_1 + R_2$. Note that $\Omega_{R1} = \Omega_{R2} = \{1,2,3\}$\$. With that in mind we can find the PMF of X:

Work on this problem with the people around you, and then we'll go over it together!

Problem 4 - 3 Sided Die

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(a) What is the PMF of X?

This gives us the following:

$$\begin{aligned}
p_X(k) = \mathbb{P}(X = k) &= \sum_{i \in \Omega_{R1}} \mathbb{P}(R_1 = i, R_2 = k - i) \\
&= \sum_{i \in \Omega_{R1}} \mathbb{P}(R_1 = i) \cdot \mathbb{P}(R_2 = k - i) \qquad \text{(By independence of the rolls)} \\
&= \sum_{i \in \Omega_{R1}} \frac{1}{3} \cdot p_{R2}(k - i) \\
&= \frac{1}{3} \left(p_{R2}(k-1) + p_{R2}(k-2) + p_{R2}(k-3) \right)
\end{aligned}$$

Problem 4 - 3 Sided Die

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(a) What is the PMF of X?

At this point, we can evaluate the pmf of X for each value in the range of X, noting that $p_{R2}(k - i) = 0$ if $k - i \notin \Omega_{R2}$, $1/3$ otherwise. We get:

$$p_X(k) = \begin{cases} 1/9 & k = 2 \\ 2/9 & k = 3 \\ 3/9 & k = 4 \\ 2/9 & k = 5 \\ 1/9 & k = 6 \\ 0 & \text{otherwise} \end{cases}$$

Problem 4 - 3 Sided Die

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(b) Find $E[X]$.

There are two ways to find the expected value of X. We could apply the *definition of expectation* using the PMF found in part (a). This gives us

$$\mathbb{E}[X] = \sum_{k=2}^{6} k p_X(k) = 2 \cdot \frac{1}{9} + 3 \cdot \frac{2}{9} + 4 \cdot \frac{3}{9} + 5 \cdot \frac{2}{9} + 6 \cdot \frac{1}{9} = \boxed{4}$$

Problem 4 - 3 Sided Die

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(b) Find $E[X]$.

Alternatively, we can use *linearity of expectation* here. Let R_1 be the roll of the first die, and R_2 the roll of the second. Then, $X = R_1 + R_2$.
By linearity of expectation, we get:

$$\mathbb{E}[X] = \mathbb{E}[R_1 + R_2] = \mathbb{E}[R_1] + \mathbb{E}[R_2]$$

We compute:

$$\mathbb{E}[R_1] = \sum_{i \in \Omega_{R1}} i \cdot \mathbb{P}(R_1 = i) = \sum_{i \in \Omega_{R1}} i \cdot \frac{1}{3} = \frac{1}{3}(1 + 2 + 3) = 2$$

Similarly, $E[R_2] = 2$, since the rolls are independent.

Plugging into our expression for the expectation of X gives us:

$$E[X] = 2 + 2 = \boxed{4}$$

Problem 4 - 3 Sided Die

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

(c) Find $\mathrm{Var}[X]$.

We know from the definition of variance that

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

We can compute the $\mathbb{E}[X^2]$ term as follows:

$$\mathbb{E}[X^2] = \sum_{x=2}^{6} x^2 p_X(x) = \frac{2^2 \cdot 1 + 3^2 \cdot 2 + 4^2 \cdot 3 + 5^2 \cdot 2 + 6^2 \cdot 1}{9} = \frac{52}{3}$$

Plugging this into our variance equation gives us

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{52}{3} - 4^2 = \boxed{\frac{4}{3}}$$

Problem 10 - Poisson Practice

Problem 10 – Poisson Practice

Seattle averages 3 days with snowfall per year. Suppose the number of days with snowfall follows a Poisson distribution.

a) What is the probability of getting exactly 5 days of snow in a year?

b) According to the Poisson model, what is the probability of getting 367 days of snow?

Work on this problem with the people around you, and then we'll go over it together!

Problem 10 – Poisson Practice

Seattle averages 3 days with snowfall per year. Suppose the number of days with snowfall follows a Poisson distribution.

a) What is the probability of getting exactly 5 days of snow in a year?

Problem 10 – Poisson Practice

- Seattle averages 3 days with snowfall per year. Suppose the number of days with snowfall follows a Poisson distribution.

a) What is the probability of getting exactly 5 days of snow in a year?

Let $X \sim \text{Poi}(3)$ Then $p_X(5) = \frac{3^5 e^{-3}}{5!} \approx 0.1008$

Problem 10 – Poisson Practice

Seattle averages 3 days with snowfall per year. Suppose the number of days with snowfall follows a Poisson distribution.

b) According to the Poisson model, what is the probability of getting 367 days of snow?

Problem 10 – Poisson Practice

- Seattle averages 3 days with snowfall per year. Suppose the number of days with snowfall follows a Poisson distribution.

b) According to the Poisson model, what is the probability of getting 367 days of snow?

Let $X \sim \text{Poi}(3)$ Then $p_X(367) = \frac{3^{367} e^{-3}}{367!} \approx 1.08 \times 10^{-610}$

Problem 10 – Poisson Practice

- Seattle averages 3 days with snowfall per year. Suppose the number of days with snowfall follows a Poisson distribution.

b) According to the Poisson model, what is the probability of getting 367 days of snow?

Let $X \sim \text{Poi}(3)$ Then $p_X(367) = \frac{3^{367}e^{-3}}{367!} \approx 1.08 \times 10^{-610}$

That's a very small estimate, but of course the true probability is 0. Recall that using a Poisson distribution is a modeling assumption, it may produce nonzero probabilities for events that are practically impossible.

Problem 7 – True or False?

Problem 7 – True or False?

- a) For any random variable X, we have $E[X^2] \geq E[X]^2$

Problem 7 – True or False?

a) For any random variable X, we have $E[X^2] \geq E[X]^2$

True

Since $0 \leq \mathrm{Var}(X) = E[(X - E[X])^2]$, since the squaring necessitates the result is non-negative.

Problem 7 – True or False?

- b) Let X, Y be random variables. Then, X and Y are independent if and only if $E[XY] = E[X]E[Y]$

Problem 7 – True or False?

b) Let X, Y be random variables. Then, X and Y are independent if and only if $E[XY] = E[X]E[Y]$

False

The forward implication is true, but the reverse is not. For example, if X is the discrete uniform random variable on the set $\{-1, 0, 1\}$ such that $P(X = -1) = P(X = 0) = P(X = 1) = \frac{1}{3}$, and $Y = X^2$, we have $E[X] = 0$, so $E[X]E[Y] = 0$. However, since $X = X^3$, $E[XY] = E[XX^2] = E[X^3] = E[X] = 0$, we have that $E[X]E[Y] = 0 = E[XY]$. However, X and Y are not independent; indeed, $P(Y = 0|X = 0) = 1 \neq \frac{1}{3} = P(Y = 0)$

Problem 7 – True or False?

e) Let $X \sim \text{Binomial}(n, p)$ and $Y \sim \text{Binomial}(m, p)$ be independent. Then, $X + Y \sim \text{Binomial}(n + m, p)$

Problem 7 – True or False?

c) Let $X \sim \text{Binomial}(n, p)$ and $Y \sim \text{Binomial}(m, p)$ be independent. Then, $X + Y \sim \text{Binomial}(n + m, p)$

True

X is the sum of n independent Bernoulli trials, and Y is the sum of m. So $X + Y$ is the sum of $n + m$ independent Bernoulli trials, so $X + Y \sim \text{Binomial}(n + m, p)$.

Problem 7 – True or False?

d) Let $X_1, \ldots, X_{n+1}$ be independent $\text{Bernoulli}(p)$ random variables. Then, $\mathrm{E}\left[\sum_{i=1}^{n} X_i X_{i+1}\right] = np^2$.

Problem 7 – True or False?

d) Let $X_1, \ldots, X_{n+1}$ be independent $\text{Bernoulli}(p)$ random variables. Then, $\mathrm{E}[\sum_{i=1}^{n} X_i X_{i+1}] = np^2$.

True

Notice that $X_i X_{i+1}$ is also Bernoulli (only takes on 0 and 1), but is 1 if and only if both are 1, so $X_i X_{i+1} \sim \text{Bernoulli}(p^2)$. The statement holds by linearity, since $E[X_i X_{i+1}] = p^2$.

Problem 7 – True or False?

- e) Let $X_1, \ldots, X_{n+1}$ be independent $\text{Bernoulli}(p)$ random variables. Then, $\mathrm{Y} = \sum_{i=1}^{n} X_i X_{i+1} \sim \text{Binomial}(n, p^2)$.

Problem 7 – True or False?

e) Let $X_1, \ldots, X_{n+1}$ be independent $\mathrm{Bernoulli}(p)$ random variables. Then, $\mathrm{Y} = \sum_{i=1}^{n} X_i X_{i+1} \sim \mathrm{Binomial}(n, p^2)$.

False

They are all Bernoulli p^2 as determined in the previous part, but they are not independent. $P(X_1X_2 = 1 | X_2X_3 = 1) = P(X_1 = 1) = p \neq p^2 = P(X_1X_2 = 1)$.

Problem 7 – True or False?

- f) If $X \sim \text{Bernoulli}(p)$, then $nX \sim \text{Binomial}(n, p)$.

Problem 7 – True or False?

f) If $X \sim \text{Bernoulli}(p)$, then $nX \sim \text{Binomial}(n, p)$.

False

The range of X is $\{0, 1\}$, so the range of nX is $\{0, n\}$. nX cannot be $\text{Bin}(n, p)$, otherwise its range would be $\{0, 1, \ldots, n\}$.

Problem 7 – True or False?

- g) If $X \sim \text{Binomial}(n, p)$, then $\frac{X}{n} \sim \text{Bernoulli}(p)$.

Problem 7 – True or False?

- g) If $X \sim \mathrm{Binomial}(n, p)$, then $\frac{X}{n} \sim \mathrm{Bernoulli}(p)$.

False

Again, the range of X is $\{0, 1, \ldots, n\}$, so the range of $\frac{X}{n}$ is $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}$. Hence it cannot be $\mathrm{Ber}(p)$, otherwise its range would be $\{0, 1\}$.

Problem 7 – True or False?

h) For any two independent random variables. X, Y, we have $\mathrm{Var}(X - Y) = \mathrm{Var}(X) - \mathrm{Var}(\mathrm{Y})$.

Problem 7 – True or False?

h) For any two independent random variables. X, Y, we have $\mathrm{Var}(X - Y) = \mathrm{Var}(X) - \mathrm{Var}(\mathrm{Y})$.

False

$$\mathrm{Var}(X - Y) = \mathrm{Var}\big(X + (-Y)\big) = \mathrm{Var}(X) + (-1)^2\mathrm{Var}(Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

Problem 5 (Midterm Review)

5 (Midterm Review)

- How many integers in $\{1,2,\ldots,360\}$ are divisible by one or more of the numbers 2, 3, and 5?

Work on this problem with the people around you, and then we'll go over it together!

5 (Midterm Review)

- How many integers in $\{1,2,...,360\}$ are divisible by one or more of the numbers 2, 3, and 5?

Inclusion-exclusion:

$$\frac{360}{2}+\frac{360}{3}+\frac{360}{5}-\frac{360}{2\times 3}-\frac{360}{2\times 5}-\frac{360}{3\times 5}+\frac{360}{2\times 3\times 5}=80+120+72-60-36-24+12=264$$

Problem 3 (Midterm Review)

3 (Midterm Review)

Consider the following inequality: $a_1 + a_2 + a_3 + a_4 + a_5 + a_6 \leq 70$. A solution to this inequality over the nonnegative integers is a choice of a nonnegative integer for each of the 6 variables $a_1, a_2, a_3, a_4, a_5, a_6$ that satisfies the inequality. To be different, two solutions have to differ on the value assigned to some a_i. How many different solutions are there to the inequality?

Work on this problem with the people around you, and then we'll go over it together!

3 (Midterm Review)

Consider the following inequality: $a_1 + a_2 + a_3 + a_4 + a_5 + a_6 \leq 70$. A solution to this inequality over the nonnegative integers is a choice of a nonnegative integer for each of the 6 variables $a_1, a_2, a_3, a_4, a_5, a_6$ that satisfies the inequality. To be different, two solutions have to differ on the value assigned to some a_i. How many different solutions are there to the inequality?

This is equivalent to asking how many different solutions are there to the equation $a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 \leq 70$.

Using stars and bars, we see that the answer is $\binom{76}{6}$

Problem 12 (Midterm Review)

12 (Midterm Review)

You are working on a difficult passage from a new piece you are learning on the piano. You wish to play it correctly 4 times before stopping for the day. If your probability of playing it correctly on each attempt is 2/3, and the attempts are independent (unfortunately!), what is the probability that you have to play it at least 8 times?

Work on this problem with the people around you, and then we'll go over it together!

12 (Midterm Review)

You are working on a difficult passage from a new piece you are learning on the piano. You wish to play it correctly 4 times before stopping for the day. If your probability of playing it correctly on each attempt is 2/3, and the attempts are independent (unfortunately!), what is the probability that you have to play it at least 8 times?

This is equivalent to the probability that, in the first 7 attempts, you play it correctly 3 or fewer times. Let X be the number of times you play it correctly in the first 7 attempts. Then $X \sim Bin(7, \frac{2}{3})$

$$\mathbb{P}(X \leq 3) = \binom{7}{0}\left(\frac{2}{3}\right)^0\left(\frac{1}{3}\right)^7 + \binom{7}{1}\left(\frac{2}{3}\right)^1\left(\frac{1}{3}\right)^6 + \binom{7}{2}\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^5 + \binom{7}{3}\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^4$$

$$= \frac{379}{3^7} \approx 0.173$$

Problem 10 (Midterm Review)

10 (Midterm Review)

Suppose you record the birthdays of a large group of people, one at a time, until you have found a person whose birthday matches your own birthday. What is the probability that it takes exactly 20 people for this to occur? Assume that there are 365 possible birthdays and each one is equally probable for a randomly chosen person.

Work on this problem with the people around you, and then we'll go over it together!

10 (Midterm Review)

Suppose you record the birthdays of a large group of people, one at a time, until you have found a person whose birthday matches your own birthday. What is the probability that it takes exactly 20 people for this to occur? Assume that there are 365 possible birthdays and each one is equally probable for a randomly chosen person.

This is a geometric distribution:

$$\left(\frac{364}{365}\right)^{19} \frac{1}{354} \approx 0.0026$$

Problem 9 (Midterm review)

9 (Midterm Review)

The space shuttle has 6 O-rings: these were involved in the Challenger disaster. When the space shuttle is launched, each O-ring has a probability of failure of 0.0137, independent of whether other O-rings fail.

(a) What is the probability that, during 23 launches, no O-ring will fail, but that at least one O-ring will fail during the 24th launch?

Work on this problem with the people around you, and then we'll go over it together!

9 (Midterm Review)

The space shuttle has 6 O-rings: these were involved in the Challenger disaster. When the space shuttle is launched, each O-ring has a probability of failure of 0.0137, independent of whether other O-rings fail.

(a) What is the probability that, during 23 launches, no O-ring will fail, but that at least one O-ring will fail during the 24th launch?

The probability that no O-ring fails on a single launch is $(1-0.0137)^6 \approx 0.921$. The probability that this happens for 23 launches and doesn't happen on the 24th launch is $0.921^{23}(1-0.921) \approx 0.0118$.

9 (Midterm Review)

The space shuttle has 6 O-rings: these were involved in the Challenger disaster. When the space shuttle is launched, each O-ring has a probability of failure of 0.0137, independent of whether other O-rings fail.

(b) What is the probability that no O-ring fails during 24 launches?

Work on this problem with the people around you, and then we'll go over it together!

9 (Midterm Review)

The space shuttle has 6 O-rings: these were involved in the Challenger disaster. When the space shuttle is launched, each O-ring has a probability of failure of 0.0137, independent of whether other O-rings fail.

(b) What is the probability that no O-ring fails during 24 launches?

$$0.921^{24} \approx 0.137$$

Problem 7 (Midterm review)

7 (Midterm Review)

You are trying to diagnose the probability that a patient with a positive blood sugar test result has diabetes, even though she is in a low-risk group. The probability of a woman in this group having diabetes is 0.8%. 90% of women with diabetes will test positive in the blood
sugar test. 7% of women without diabetes will test positive in the blood sugar test. Your patient tests positive in the blood sugar test. What is the probability that she has diabetes?

Work on this problem with the people around you, and then we'll go over it together!

7 (Midterm Review)

You are trying to diagnose the probability that a patient with a positive blood sugar test result has diabetes, even though she is in a low-risk group. The probability of a woman in this group having diabetes is 0.8%. 90% of women with diabetes will test positive in the blood
sugar test. 7% of women without diabetes will test positive in the blood sugar test. Your patient tests positive in the blood sugar test. What is the probability that she has diabetes?

Let D be the event that she has diabetes and + be the event of a positive test.

$$\mathbb{P}(D\mid +) = \frac{\mathbb{P}(+\mid D)\mathbb{P}(D)}{\mathbb{P}(+\mid D)\mathbb{P}(D) + \mathbb{P}(+\mid \bar{D})\mathbb{P}(\bar{D})} = \frac{0.9 \times 0.008}{0.9 \times 0.008 + 0.07 \times 0.992} \approx 0.09$$

Problem 1 (Midterm review)

1

Let A and B be events in the same sample space that each have nonzero probability. For each of the following statements, state whether it is always true, always false, or it depends on information not given.

a) If A and B are mutually exclusive, then they are independent.

b) If A and B are independent, then they are mutually exclusive.

c) If $\mathbb{P}(A) = \mathbb{P}(B) = 0.75$, then A and B are mutually exclusive.

d) If $\mathbb{P}(A) = \mathbb{P}(B) = 0.75$, then A and B are independent.

Work on this problem with the people around you, and then we'll go over it together!

1

Let A and B be events in the same sample space that each have nonzero probability. For each of the following statements, state whether it is always true, always false, or it depends on information not given.

a) If A and B are mutually exclusive, then they are independent.

1

Let A and B be events in the same sample space that each have nonzero probability. For each of the following statements, state whether it is always true, always false, or it depends on information not given.

a) If A and B are mutually exclusive, then they are independent.

False

1

Let A and B be events in the same sample space that each have nonzero probability. For each of the following statements, state whether it is always true, always false, or it depends on information not given.

b) If A and B are independent, then they are mutually exclusive.

1

Let A and B be events in the same sample space that each have nonzero probability. For each of the following statements, state whether it is always true, always false, or it depends on information not given.

b) If A and B are independent, then they are mutually exclusive.

False

1

Let A and B be events in the same sample space that each have nonzero probability. For each of the following statements, state whether it is always true, always false, or it depends on information not given.

c) If $\mathbb{P}(A) = \mathbb{P}(B) = 0.75$, then A and B are mutually exclusive.

1

Let A and B be events in the same sample space that each have nonzero probability. For each of the following statements, state whether it is always true, always false, or it depends on information not given.

c) If $\mathbb{P}(A) = \mathbb{P}(B) = 0.75$, then A and B are mutually exclusive.

False

1

Let A and B be events in the same sample space that each have nonzero probability. For each of the following statements, state whether it is always true, always false, or it depends on information not given.

d) If $\mathbb{P}(A) = \mathbb{P}(B) = 0.75$, then A and B are independent.

1

Let A and B be events in the same sample space that each have nonzero probability. For each of the following statements, state whether it is always true, always false, or it depends on information not given.

d) If $\mathbb{P}(A) = \mathbb{P}(B) = 0.75$, then A and B are independent.

Depends whether $\mathbb{P}(A \cap B) = 9/16$

Problem 2 (Midterm review)

2

Given any set of 18 integers, show that one may always choose two of them so that their difference is divisible by 17.

Work on this problem with the people around you, and then we'll go over it together!

2

Given any set of 18 integers, show that one may always choose two of them so that their difference is divisible by 17.

By the pigeonhole principle, two of them, say x and y, must have the same remainder when divided by 17. That means $x \equiv y \pmod{17}$, which in turn means 17 divides $x - y$.

Problem 4 (Midterm Review)

4

You roll three fair dice, each with a different numbers of faces: die 1 has six faces (numbered $1 \ldots 6$), die 2 has eight faces (numbered $1 \ldots 8$), and die 3 has twelve faces (numbered $1 \ldots 12$). Let the random variable X be the sum of the three values rolled. What is $\mathbb{E}[X]$?

Work on this problem with the people around you, and then we'll go over it together!

4

You roll three fair dice, each with a different numbers of faces: die 1 has six faces (numbered 1 … 6), die 2 has eight faces (numbered 1 … 8), and die 3 has twelve faces (numbered 1 … 12). Let the random variable X be the sum of the three values rolled. What is $\mathbb{E}[X]$?

Let D_1, D_2, D_3 be the values of die 1, die 2, and die 3, respectively.

$\mathbb{E}[D_1] = 3.5$
$\mathbb{E}[D_2] = 4.5$
$\mathbb{E}[D_3] = 6.5$

Therefore,

$$\mathbb{E}[X] = \mathbb{E}[D_1 + D_2 + D_3] = \mathbb{E}[D_1] + \mathbb{E}[D_2] + \mathbb{E}[D_3] = 3.5 + 4.5 + 6.5 = 14.5$$

That's All, Folks!

Thanks for coming to section this week!
Any questions?