Test your intuition!

Each of 36 people picks a whole number at random, between 1 and 400. How likely is it that there's a number that gets selected more than once?

A: More than 80% likely

B: Around 60%

C: Around 40%

D: Less than 20% likely

Multiplication and addition principles

Multiplication principle: If an experiment has two stages, with

- a_1 outcomes for the first stage, and
- a_2 outcomes for the second (regardless of what happened at the first stage)

then the total number of outcomes for the experiment is $a_1 \times a_2$.

Multiplication and addition principles

Multiplication principle: If an experiment has two stages, with

- a_1 outcomes for the first stage, and
- a_2 outcomes for the second (regardless of what happened at the first stage)

then the total number of outcomes for the experiment is $a_1 \times a_2$.

Addition principle: If an experiment can proceed in one of two mutually exclusive ways, with

- a_1 outcomes for the first way, and
- a_2 outcomes for the second

then the total number of outcomes for the experiment is $a_1 + a_2$.

Sampling, with and without replacement

A bag has n distinguishable balls.

Sampling: k balls are pulled from the bag, one after the other

Sampling, with and without replacement

A bag has n distinguishable balls.

Sampling: k balls are pulled from the bag, one after the other

Sampling without replacement: When a ball is pulled out, it is *not* returned to the bag. The number of samples of size k that can be drawn is

$$A_{k,n} = (n)_k = n \cdot (n-1) \cdot \cdots \cdot (n-(k-1)) = \frac{n!}{(n-k)!}$$

where

$$n! = n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1.$$

Sampling, with and without replacement

A bag has n distinguishable balls.

Sampling: k balls are pulled from the bag, one after the other

Sampling without replacement: When a ball is pulled out, it is *not* returned to the bag. The number of samples of size k that can be drawn is

$$A_{k,n} = (n)_k = n \cdot (n-1) \cdot \cdots \cdot (n-(k-1)) = \frac{n!}{(n-k)!}$$

where

$$n! = n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1.$$

Sampling with replacement: When a ball is pulled out, it *is* returned to the bag. The number of samples of size k that can be drawn is

$$n^k.$$

Sampling without replacement *or regard for order*

Ordering: A bag has n distinguishable balls. The number of ways to take out the balls one by one and place them in order is

$$n! = n \times (n-1) \times \cdots \times 3 \times 2 \times 1$$

This is also the number of *permutations* of n objects.

Sampling without replacement *or regard for order*

Ordering: A bag has n distinguishable balls. The number of ways to take out the balls one by one and place them in order is

$$n! = n \times (n-1) \times \cdots \times 3 \times 2 \times 1$$

This is also the number of *permutations* of n objects.

Sampling without replacement: The number of ways of drawing k balls from the bag, without replacement, *in order*, is

$$(n)_k = n \cdot (n-1) \cdot \cdots \cdot (n-(k-1)) = \frac{n!}{(n-k)!}$$

Sampling without replacement *or regard for order*

Ordering: A bag has n distinguishable balls. The number of ways to take out the balls one by one and place them in order is

$$n! = n \times (n-1) \times \cdots \times 3 \times 2 \times 1$$

This is also the number of *permutations* of n objects.

Sampling without replacement: The number of ways of drawing k balls from the bag, without replacement, *in order*, is

$$(n)_k = n \cdot (n-1) \cdot \cdots \cdot (n-(k-1)) = \frac{n!}{(n-k)!}$$

Sampling without replacement and without regard for order: The number of ways of drawing k balls from the bag, without replacement, *with order not mattering*, is

$$\frac{(n)_k}{k!} = \frac{n!}{(n-k)!k!} = \binom{n}{k}.$$

This can also be thought of as the number of ways of drawing k balls from the bag in a single draw.

Test your intuition!

On a cold winter's night, 80 000 people trudge into Notre Dame stadium for a Garth Brooks concert. They each leave their scarves at the Knute Rockne gate. As they leave the concert, they each pick up a random scarf from the pile at the gate (they are all too cold to bother looking for their own scarf).

How likely is it that *no one* comes away with the same scarf they arrived with?

A: Very likely (more than 99%)

B: More likely than not — around 2/3

C: Less likely than not — around 1/3

D: Very unlikely (less than 1%)

Dividing a group into smaller groups

Order mattering: The number of ways to split a set of n things into a first subset of size a_1, a second of size a_2, et cetera, up to a kth of size a_k (so $a_1 + \cdots + a_k = n$) is a *multinomial coefficient*:

$$\binom{n}{a_1}\binom{n-a_1}{a_2}\cdots\binom{n-a_1-\cdots-a_{k-1}}{a_k} = \frac{n!}{a_1!a_2!\cdots a_k!} = \binom{n}{a_1, a_2, \ldots, a_k}$$

Dividing a group into smaller groups

Order mattering: The number of ways to split a set of n things into a first subset of size a_1, a second of size a_2, et cetera, up to a kth of size a_k (so $a_1 + \cdots + a_k = n$) is a *multinomial coefficient*:

$$\binom{n}{a_1}\binom{n-a_1}{a_2}\cdots\binom{n-a_1-\cdots-a_{k-1}}{a_k} = \frac{n!}{a_1!a_2!\cdots a_k!} = \binom{n}{a_1, a_2, \ldots, a_k}$$

Order not mattering: The number of ways to split the set into k equal sized subsets (size n/k each), order not mattering, is

$$\frac{1}{k!}\binom{n}{n/k, n/k, \ldots, n/k} = \frac{n!}{((n/k)!)^k k!}$$

Dividing a group into smaller groups

Order mattering: The number of ways to split a set of n things into a first subset of size a_1, a second of size a_2, et cetera, up to a kth of size a_k (so $a_1 + \cdots + a_k = n$) is a *multinomial coefficient*:

$$\binom{n}{a_1}\binom{n-a_1}{a_2}\cdots\binom{n-a_1-\cdots-a_{k-1}}{a_k} = \frac{n!}{a_1!a_2!\cdots a_k!} = \binom{n}{a_1, a_2, \ldots, a_k}$$

Order not mattering: The number of ways to split the set into k equal sized subsets (size n/k each), order not mattering, is

$$\frac{1}{k!}\binom{n}{n/k, n/k, \ldots, n/k} = \frac{n!}{((n/k)!)^k k!}$$

The number of ways to split the set into m_1 subset each of size a_1, m_2 of size a_2, et cetera, up to m_k of size a_k (so $m_1 a_1 + \cdots + m_k a_k = n$) is:

$$\frac{\binom{n}{a_1,\ldots a_1, a_2, \ldots, a_2, \ldots, a_k, \ldots, a_k}}{m_1! m_2! \cdots m_k!} = \frac{n!}{(a_1!)^{m_1}(a_2!)^{m_2}\cdots(a_k!)^{m_k} m_1! m_2! \cdots m_k!}$$

Inclusion-Exclusion

$A_1, A_2, \ldots, A_n$ events in a sample space S (think of A_i as event "(at least) ith thing occurs (and maybe other things, too)")

$A_1 cup A_2 \cup \ldots \cup A_n$ is the event "at least one of the i things occurs"

Inclusion-Exclusion: $P(A_1 \cup A_2 \cup \ldots \cup A_n)$ can be calculated as

$$
\begin{gathered}
P(A_1) + P(A_2) + \cdots + P(A_n) \\
-P(A_1 \cap A_2) - P(A_1 \cap A_3) - \cdots - P(A_{n-1} \cap A_n) \\
+P(A_1 \cap A_2 \cap A_3) + \cdots + P(A_{n-2} \cap A_{n-1} \cap A_n) \\
- \cdots \\
+(-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n)
\end{gathered}
$$

Inclusion-Exclusion

$A_1, A_2, \ldots, A_n$ events in a sample space S (think of A_i as event "(at least) ith thing occurs (and maybe other things, too)")

$A_1 cup A_2 \cup \ldots \cup A_n$ is the event "at least one of the i things occurs"

Inclusion-Exclusion: $P(A_1 \cup A_2 \cup \ldots \cup A_n)$ can be calculated as

$$\begin{gathered} P(A_1) + P(A_2) + \cdots + P(A_n) \\ -P(A_1 \cap A_2) - P(A_1 \cap A_3) - \cdots - P(A_{n-1} \cap A_n) \\ +P(A_1 \cap A_2 \cap A_3) + \cdots + P(A_{n-2} \cap A_{n-1} \cap A_n) \\ - \cdots \\ +(-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n) \end{gathered}$$

$P((A_1 \cup A_2 \cup \ldots \cup A_n)^c)$, event "none of the i things occur", can be calculated as

$$\begin{gathered} 1 - (P(A_1) + \cdots + P(A_n)) \\ +(P(A_1 \cap A_2) - \cdots - P(A_{n-1} \cap A_n)) + \cdots \\ +(-1)^n P(A_1 \cap \cdots \cap A_n) \end{gathered}$$

Putting indistinguishable balls in distinguishable boxes

The number of ways to distribute k indistinguishable balls among n distinguishable boxes, which is the same as number of solutions to

$$a_1 + a_2 + \cdots + a_n = k$$

with all $a_i \geq 0$, is

$$\binom{n+k-1}{n-1} = \binom{n+k-1}{k}.$$

Putting indistinguishable balls in distinguishable boxes

The number of ways to distribute k indistinguishable balls among n distinguishable boxes, which is the same as number of solutions to

$$a_1 + a_2 + \cdots + a_n = k$$

with all $a_i \geq 0$, is

$$\binom{n+k-1}{n-1} = \binom{n+k-1}{k}.$$

The number of ways, if each box should get at least one ball, which is the same as the number of solutions to $a_1 + a_2 + \cdots + a_n = k$ with all $a_i \geq 1$, which is the same as the number of solutions to

$$a_1' + a_2' + \cdots + a_n' = k - n$$

with all $a_i' \geq 0$, is

$$\binom{(k-n)+(n-1)}{n-1} = \binom{k-1}{n-1}.$$

Test your intuition!

The World anti-doping agency (WADA) conducts random drug tests on olympic athletes. One test is for the presence of meldonium, a drug which is estimated to be used by 1 out of every 200 olympic athletes.

The test is 98% accurate: 98% of the time that meldonium is present in a sample, the test will correctly detect it, and 98% of the time that meldonium is absent from a sample, the test will correctly report the absence.

An athlete is selected at random, and tests positive for meldonium. How likely is it that the athlete is actually using meldonium?

A: Around 98%

B: Close to 75%

C: Close to 50%

D: Close to 25%

E: Around 2%

Drug testing example

Meldonium is estimated to be used by 1 out of every 200 athletes.

A certain meldonium test is 98% accurate: 98% of the time that meldonium is present in a sample, the test will correctly detect it, and 98% of the time that meldonium is absent from a sample, the test will correctly report the absence.

An athlete is selected at random, and tests positive for meldonium. How likely is it that the athlete is actually using meldonium?

Drug testing example

Meldonium is estimated to be used by 1 out of every 200 athletes.

A certain meldonium test is 98% accurate: 98% of the time that meldonium is present in a sample, the test will correctly detect it, and 98% of the time that meldonium is absent from a sample, the test will correctly report the absence.

An athlete is selected at random, and tests positive for meldonium. How likely is it that the athlete is actually using meldonium?

$M = \{\text{uses}\}$, $C = \{\text{doesn't}\}$, $P = \{\text{tests positive}\}$, $N = \{\text{tests negative}\}$

$$P(M|P) = \frac{P(P|M)P(M)}{P(P|M)P(M) + P(P|C)P(C)} = \frac{.0049}{.0049 + .0199} = .1975\ldots$$

Drug testing example

Meldonium is estimated to be used by 1 out of every 200 athletes.

A certain meldonium test is 98% accurate: 98% of the time that meldonium is present in a sample, the test will correctly detect it, and 98% of the time that meldonium is absent from a sample, the test will correctly report the absence.

An athlete is selected at random, and tests positive for meldonium. How likely is it that the athlete is actually using meldonium?

$M = \{\text{uses}\}$, $C = \{\text{doesn't}\}$, $P = \{\text{tests positive}\}$, $N = \{\text{tests negative}\}$

$$P(M|P) = \frac{P(P|M)P(M)}{P(P|M)P(M) + P(P|C)P(C)} = \frac{.0049}{.0049 + .0199} = .1975\ldots$$

Suppose instead *two* independent samples test positive? (Event PP)

Drug testing example

Meldonium is estimated to be used by 1 out of every 200 athletes.

A certain meldonium test is 98% accurate: 98% of the time that meldonium is present in a sample, the test will correctly detect it, and 98% of the time that meldonium is absent from a sample, the test will correctly report the absence.

An athlete is selected at random, and tests positive for meldonium. How likely is it that the athlete is actually using meldonium?

$M = \{\text{uses}\}$, $C = \{\text{doesn't}\}$, $P = \{\text{tests positive}\}$, $N = \{\text{tests negative}\}$

$$P(M|P) = \frac{P(P|M)P(M)}{P(P|M)P(M) + P(P|C)P(C)} = \frac{.0049}{.0049 + .0199} = .1975\ldots$$

Suppose instead *two* independent samples test positive? (Event PP)

$$P(M|PP) = \frac{P(PP|M)P(M)}{P(PP|M)P(M) + P(PP|C)P(C)} = \frac{.004802}{.004802 + .000398} = .923\ldots$$

Recall from last time

A *random variable* is a function that assigns a numerical value to each point in a sample space:

$$X : S \to \mathbb{R}.$$

If P is a probability function on the sample space, then for A a set of real numbers

$$P(X \in A) = P(\{s \in S : X(s) \in A\}).$$

Running examples:

- Roll two dice, observe both numbers. X is sum of two numbers.
- Draw cards from deck with replacement until first club is drawn, observe sequence of cards. Y is number of draws.
- Throw dart at random at radius 20cm dartboard, observe point. D is distance from center of board.

Binomial pmf (a.k.a. Binomial distribution)

Bernoulli trial: Single trial, success probability p, failure probability $1 - p$

Random variable X records success or failure

Mass function: $p_X(x) = \begin{cases} p & \text{if } x = 1 \\ q & \text{if } x = 0 \end{cases}$

$X \sim \mathrm{Ber}(p)$

Binomial trial: n independent repetitions of Bernoulli trial, success probability p

Random variable X records number of success

Mass function: $p_X(x) = \binom{x}{k} p^k q^{n-k}$, $x = 0, 1, \ldots, n$

$X \sim \mathrm{Bin}(n, p)$

Test your intuition!

I have two cookie jars, initially with 40 cookies in each one. Each time I want a cookie, I pick a jar at random to take it from.

When I first find a jar empty, which of these is most likely for the number of cookies in the *other* jar?

A: Lots: more than 32

B: Between 24 and 31

C: Between 16 and 23

D: Between 8 and 15

E: Not many: fewer than 8

Poisson process

Models the number of occurrences of a rare event, in unit time:

- Number of earthquakes per year in US
- Number of leap-year babies on campus ("time" interpreted liberally)
- Number of atoms in a sample of francium-223 that decay per minute
- ...

Poisson process

Models the number of occurrences of a rare event, in unit time:

- Number of earthquakes per year in US
- Number of leap-year babies on campus ("time" interpreted liberally)
- Number of atoms in a sample of francium-223 that decay per minute
- ...

Assumptions:

- Average number of occurrences per unit time known — λ
- Disjoint time periods are independent of each other
- Simultaneous occurrences very unlikely

Poisson process

Models the number of occurrences of a rare event, in unit time:

- Number of earthquakes per year in US
- Number of leap-year babies on campus ("time" interpreted liberally)
- Number of atoms in a sample of francium-223 that decay per minute
- ...

Assumptions:

- Average number of occurrences per unit time known — λ
- Disjoint time periods are independent of each other
- Simultaneous occurrences very unlikely

$X =$ number of occurrences in unit time; possible values $0, 1, 2, 3, \ldots$

Poisson process

Models the number of occurrences of a rare event, in unit time:

- Number of earthquakes per year in US
- Number of leap-year babies on campus ("time" interpreted liberally)
- Number of atoms in a sample of francium-223 that decay per minute
- ...

Assumptions:

- Average number of occurrences per unit time known — λ
- Disjoint time periods are independent of each other
- Simultaneous occurrences very unlikely

$X =$ number of occurrences in unit time; possible values $0, 1, 2, 3, \ldots$

X has *Poisson* mass function; $X \sim \mathrm{Poi}(\lambda)$

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

Expectation and variance

X a discrete random variable, range $\{x_1, x_2, \ldots\}$

Expectation:

$$\begin{aligned} E(X) &= \mu \\ &= \mu_1 \\ &=: \sum_{x_i} x_i P(X = x_i) \end{aligned}$$

Expectation and variance

X a discrete random variable, range $\{x_1, x_2, \ldots\}$

Expectation:

$$\begin{aligned} E(X) &= \mu \\ &= \mu_1 \\ &=: \sum_{x_i} x_i P(X = x_i) \end{aligned}$$

Variance:

$$\begin{aligned} \mathrm{Var}(X) &= \sigma^2 \\ &=: E((X - E(X))^2) \\ &= \sum_{x_i} (x_i^2 - \mu)^2 P(X = x_i) \\ &= E(X^2) - E(X)^2 \\ &= \mu_2 - \mu_1^2 \end{aligned}$$

Expectation and variance of common random variables

1. Uniform$(1, \ldots, n)$: $\mu = \frac{n+1}{2}$, $\sigma^2 = \frac{n^2-1}{12}$

Expectation and variance of common random variables

1. Uniform$(1, \ldots, n)$: $\mu = \frac{n+1}{2}$, $\sigma^2 = \frac{n^2-1}{12}$
2. Bernoulli(p): $\mu = p$, $\sigma^2 = pq$

Expectation and variance of common random variables

1. Uniform$(1, \ldots, n)$: $\mu = \frac{n+1}{2}$, $\sigma^2 = \frac{n^2-1}{12}$
2. Bernoulli(p): $\mu = p$, $\sigma^2 = pq$
3. Binomial(n, p): $\mu = np$, $\sigma^2 = npq$

Expectation and variance of common random variables

1. Uniform$(1, \ldots, n)$: $\mu = \frac{n+1}{2}$, $\sigma^2 = \frac{n^2-1}{12}$
2. Bernoulli(p): $\mu = p$, $\sigma^2 = pq$
3. Binomial(n, p): $\mu = np$, $\sigma^2 = npq$
4. Geometric(p): $\mu = \frac{1}{p}$, $\sigma^2 = \frac{q}{p^2}$

Expectation and variance of common random variables

1. Uniform$(1, \ldots, n)$: $\mu = \frac{n+1}{2}$, $\sigma^2 = \frac{n^2-1}{12}$
2. Bernoulli(p): $\mu = p$, $\sigma^2 = pq$
3. Binomial(n, p): $\mu = np$, $\sigma^2 = npq$
4. Geometric(p): $\mu = \frac{1}{p}$, $\sigma^2 = \frac{q}{p^2}$
5. NegativeBinomial(p, k): $\mu = \frac{k}{p}$, $\sigma^2 = \frac{kq}{p^2}$

Expectation and variance of common random variables

1. Uniform$(1, \dots, n)$: $\mu = \frac{n+1}{2}$, $\sigma^2 = \frac{n^2-1}{12}$
2. Bernoulli(p): $\mu = p$, $\sigma^2 = pq$
3. Binomial(n, p): $\mu = np$, $\sigma^2 = npq$
4. Geometric(p): $\mu = \frac{1}{p}$, $\sigma^2 = \frac{q}{p^2}$
5. NegativeBinomial(p, k): $\mu = \frac{k}{p}$, $\sigma^2 = \frac{kq}{p^2}$
6. Hypergeometric(r, g, n): $\mu = \frac{nr}{r+g}$, $\sigma^2 = \dots$

Expectation and variance of common random variables

1. Uniform$(1, \ldots, n)$: $\mu = \frac{n+1}{2}$, $\sigma^2 = \frac{n^2-1}{12}$
2. Bernoulli(p): $\mu = p$, $\sigma^2 = pq$
3. Binomial(n, p): $\mu = np$, $\sigma^2 = npq$
4. Geometric(p): $\mu = \frac{1}{p}$, $\sigma^2 = \frac{q}{p^2}$
5. NegativeBinomial(p, k): $\mu = \frac{k}{p}$, $\sigma^2 = \frac{kq}{p^2}$
6. Hypergeometric(r, g, n): $\mu = \frac{nr}{r+g}$, $\sigma^2 = \ldots$
7. Poison(λ): $\mu = \lambda$, $\sigma^2 = \lambda$

Some non-discrete random variables

- X is a random number between 0 and 1

  $$\textit{range is } \text{all reals } \textit{between } 0 \textit{ and } 1 \text{ — } [0,1]$$

Some non-discrete random variables

- X is a random number between 0 and 1

  range is all reals *between* 0 *and* 1 — $[0, 1]$

- Y is time until a call is answered

  range is all reals *greater than* 0*, plus* ∞ — $(0, \infty]$

Some non-discrete random variables

- X is a random number between 0 and 1

  range is all reals *between* 0 *and* 1 — $[0, 1]$

- Y is time until a call is answered

  range is all reals *greater than* 0*, plus* ∞ — $(0, \infty]$

- Z is temperature 24 hours from now

  range is all reals *at least absolute zero* — $[-459.670, \infty)$

Some non-discrete random variables

- X is a random number between 0 and 1

  range is all reals *between* 0 *and* 1 — $[0, 1]$

- Y is time until a call is answered

  range is all reals *greater than* 0, *plus* ∞ — $(0, \infty]$

- Z is temperature 24 hours from now

  range is all reals *at least absolute zero* — $[-459.670, \infty)$

Definition: X is a *continuous* random variable if there is some function $f : \mathbb{R} \to [0, \infty)$ — the *density function* of X — such that for all real x

$$P(X \leq x) = \int_{-\infty}^{x} f(t)\ dt$$

(so the cdf of X — $F_X(x) = P(X \leq x)$ — is continuous)

Finding the density of $Y = g(X)$ from density of X

Step 1: Use the possible values of X to find the possible values of $g(X)$

E.g., $X \sim \mathrm{Exp}(1)$ *has possible values* $(0, \infty)$*, so* $Y = \log(X)$ *has possible values* $(-\infty, \infty)$

For any values not possible for Y, density is 0

Finding the density of $Y = g(X)$ from density of X

Step 1: Use the possible values of X to find the possible values of $g(X)$

E.g., $X \sim \mathrm{Exp}(1)$ *has possible values* $(0, \infty)$, *so* $Y = \log(X)$ *has possible values* $(-\infty, \infty)$

For any values not possible for Y, density is 0

Step 2: Express $P(Y \leq x)$ as $P(X$ in some range), by asking "where must X be, for $g(X)$ to be at most x?"

E.g., $P(\log X \leq x) = P(X \leq e^x)$

This is the step that usually requires some thinking

Finding the density of $Y = g(X)$ from density of X

Step 1: Use the possible values of X to find the possible values of $g(X)$

E.g., $X \sim \mathrm{Exp}(1)$ *has possible values* $(0, \infty)$, *so* $Y = \log(X)$ *has possible values* $(-\infty, \infty)$

For any values not possible for Y, density is 0

Step 2: Express $P(Y \leq x)$ as $P(X$ in some range), by asking "where must X be, for $g(X)$ to be at most x?"

E.g., $P(\log X \leq x) = P(X \leq e^x)$

This is the step that usually requires some thinking

Step 3: Use the density function of X, together with the result of Step 2, to get the cdf F_Y of Y, for all values that Y can possible take

E.g., $F_Y(x) = P(X \leq e^x) = \int_0^{e^x} e^{-t}\, dt = []_{t=0}^{e^x} = 1 - e^{e^{-x}}$

For all values below the smallest possible value of Y, the cdf is 0; for all values above the largest possible value, it is 1

Finding the density of $Y = g(X)$ from density of X

Step 1: Use the possible values of X to find the possible values of $g(X)$

E.g., $X \sim \mathrm{Exp}(1)$ *has possible values* $(0, \infty)$*, so* $Y = \log(X)$ *has possible values* $(-\infty, \infty)$

For any values not possible for Y, density is 0

Step 2: Express $P(Y \leq x)$ as $P(X$ in some range), by asking "where must X be, for $g(X)$ to be at most x?"

E.g., $P(\log X \leq x) = P(X \leq e^x)$

This is the step that usually requires some thinking

Step 3: Use the density function of X, together with the result of Step 2, to get the cdf F_Y of Y, for all values that Y can possible take

E.g., $F_Y(x) = P(X \leq e^x) = \int_0^{e^x} e^{-t}\, dt = []_{t=0}^{e^x} = 1 - e^{e^{-x}}$

For all values below the smallest possible value of Y, the cdf is 0; for all values above the largest possible value, it is 1

Step 4: Differentiate the cdf of Y to find the pdf of Y for all values that Y can possible take

E.g., $f_Y(x) = \frac{d}{dx}\left(1 - e^{e^{-x}}\right) = e^{-x}e^{e^{-x}}$