TOPIC 4
PROBABILITY AND STATISTICS

SUBTOPIC 4.9
THE NORMAL DISTRIBUTION AND CURVE

Index

The big picture

Butterfly cocoons are mostly similar in size.

GLady Pixabay

In the picture above, you can see several butterfly cocoons, and most of them are approximately the same length. However, if you look closely, you will notice that a couple of them are very short and a couple are very long. Most things grow in this way, whether it's the heights of oak trees, the weights of apples at the market or the

Overview (/study/ap 122- cid- 754029/

life-spans of people in your country. Most of the population is clustered close to the mean, and the further from the mean the values are, the fewer members of the population have those values.

This distribution is what we call the normal distribution. The normal distribution is probably the most widely known probability distribution function because it is so useful for modelling real-world data.

Concept

The normal distribution uses one of the most common **models** in statistics: the normal curve, or bell curve. This model allows you to visualise probability using area under the curve, while functions and integration enable you to perform calculations for probabilities analytically. You can also construct tables to list probabilities of multiple intervals that the random variable may reside in.

4. Probability and statistics / 4.9 The normal distribution and curve

The normal distribution

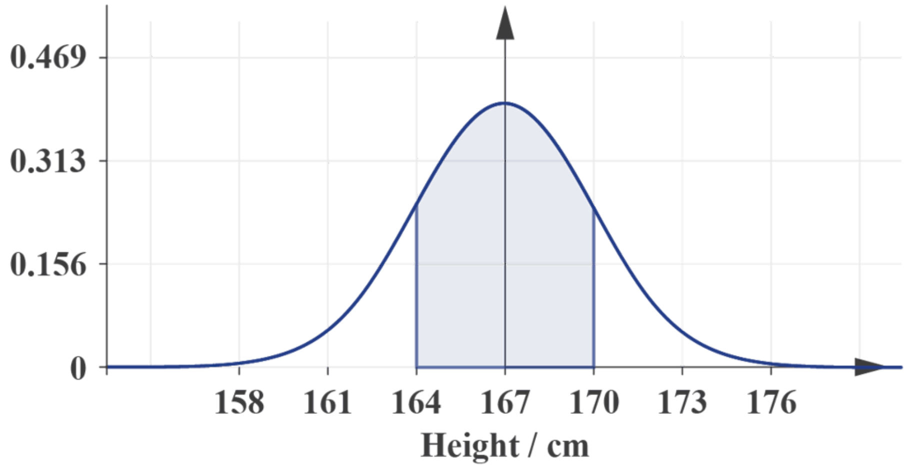
The normal (bell) curve

Expressing the normal distribution graphically

The normal distribution is often modelled graphically with a normal (or bell) curve. Take, for example, the heights of adult women living in Europe, where the mean height is 167 cm (see figure below).

Student view

Overview
(/study/ap
122-
cid-
754029/



More information

The image shows a bell curve representing the normal distribution of the heights of European adult women. The horizontal axis is labeled "Height / cm" with intervals from 158 to 176, increasing by 3 cm. The vertical axis ranges from 0 to 0.469, with intervals at approximately 0.156, 0.313, and 0.469. The curve peaks at 167 cm, which is the mean, and also the mode and median due to the symmetry of the normal distribution. The standard deviation is labeled as σ = 3, and it's depicted that one standard deviation from the mean is approximately 164 cm at the lower side and 170 cm at the higher side.

Section Student... (0/0) Feedback Print (/study/app/m/sid-122-cid-754029/book/the-big-picture-id-26265/print/) Assign

[Generated by AI]

The bell curve above shows the shape of the normal distribution of the heights of European adult women. The normal curve always has its highest point in the centre, and that point is the mean of the data, μ. Because the distribution is symmetric, μ is also its mode and median. On graphs of the curve, the horizontal axis is often labelled in intervals using the standard deviation σ (the square root of the variance). For example, the above figure is drawn this way, meaning the mean of the data is $\mu = 167$ and the standard deviation is $\sigma = 3$.

Student view

While subtopic 4.8 (/study/app/m/sid-122-cid-754029/book/the-big-picture-id-26260/) dealt with discrete random variables in the binomial distribution, the normal distribution deals with continuous random variables. In this subtopic, you

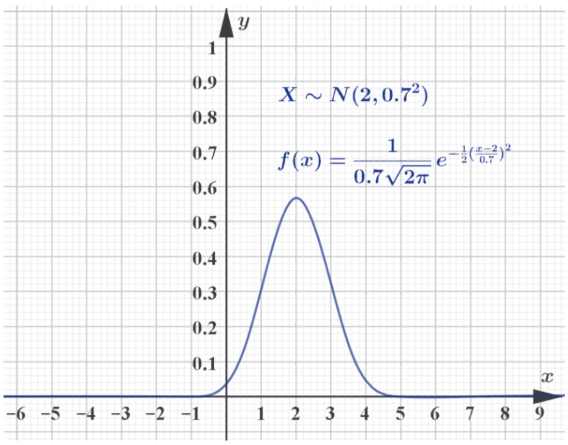
Overview (/study/ap 122- cid- 754029/

will explore how to use the normal distribution to find probabilities. As you will discover, the probability of the random variable falling within a specific range is illustrated by the area under the curve. The shaded area in the figure above, therefore, represents $\mathrm{P}(164 \leqslant \mathrm{X} \leqslant 170)$.

International Mindedness

While the figure above illustrates the heights of European women, the normal curve could be used to model the heights of women in any country or region in the world, despite the fact that the mean height could be different. The model will maintain its shape but will vary as the mean and standard deviation change. How will the model change? Why do you think the distributions of heights of people from different parts of the world are different?

The figure below shows a normal distribution for various values of μ. Manipulating μ is equivalent to performing a horizontal translation of the graph.



More information

The image is a graph depicting a normal distribution curve. The X-axis ranges from -5 to 5 and is labeled with numbers accordingly. The Y-axis shows probabilities, with values marked from 0 to 0.4 at intervals of 0.05. A blue curve, representing the normal distribution, peaks at around the X value of 2 and decreases symmetrically on both sides. The mathematical representation and equations such as (X \sim

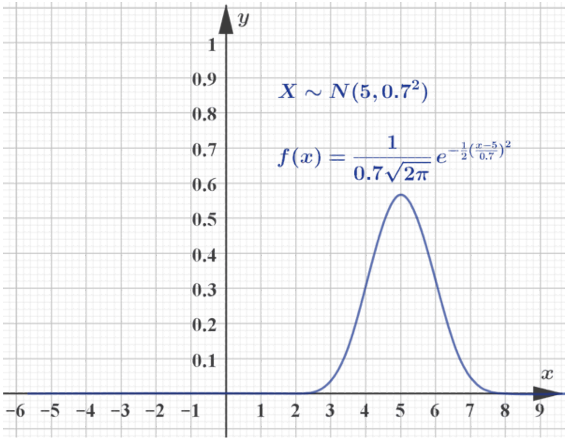
Student view

N(2, 0.5^2)) and (f(x) = \frac{1}{0.5\sqrt{2\pi}} e^{-\frac{(x - 2)^2}{0.5^2}}) are written within the graph.

These equations describe the properties of the normal distribution, specifically its mean and variance.

The overall shape of the curve is symmetric, with a bell-like appearance centered at X = 2.

[Generated by AI]

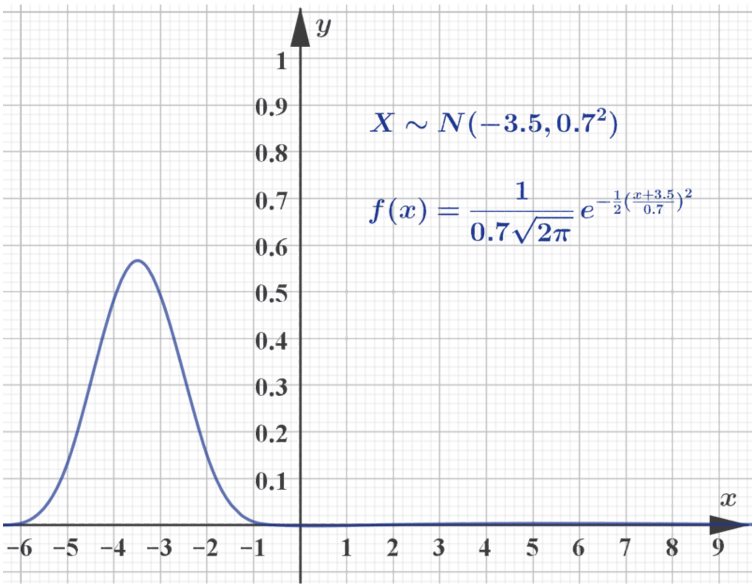


More information

This is a graph displaying a normal distribution curve. The x-axis is labeled with values ranging from -6 to 6 in increments of 1, representing the variable X. The y-axis, labeled with f(x), ranges from 0 to approximately 0.5 in increments of 0.05. The curve peaks around the center at the value of 0 on the x-axis, indicating that the distribution is symmetric. The notation X ~ N(0,1^2) is used to describe a normal distribution with a mean of 0 and a standard deviation of 1. There is also an equation displayed: f(x) = (1 / √2π) e^(-x^2/2), representing the probability density function of the normal distribution.

[Generated by AI]

Overview
(/study/ap
122-
cid-
754029/



For a given standard deviation, σ, the mean, μ, sets the centre of the normal distribution. Thus, manipulating μ is equivalent to a horizontal translation; for a larger view, click on the individual graphs.

More information

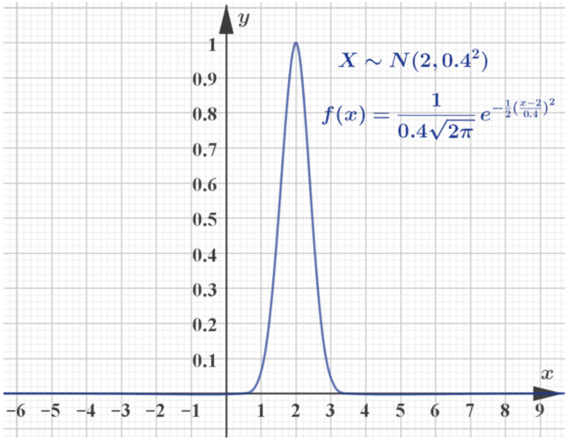
This image displays a graph of a normal distribution. The X-axis represents random variable values with a range marked from -6 to 7. The Y-axis represents probability density, labeled from 0 to 0.4. The graph shows a bell curve centered at -3.5, indicating the mean of the distribution. The curve rises sharply to a peak near 0.04 and then declines symmetrically. The equation displayed is the probability density function, which is (f(x) = \frac{1}{0.7\sqrt{2\pi}} e^{-\frac{(x + 3.5)^2}{2 \times 0.7^2}}), specifying the mean ((\mu = -3.5)) and standard deviation ((\sigma = 0.7)) of the distribution.

[Generated by AI]

The standard deviation determines the spread of the data around the mean. A small σ indicates a close clustering around μ, whereas a larger σ leads to a broad normal curve with data spread further away from the mean. The figure below shows the normal distribution for various values of σ . Manipulating σ is equivalent to performing both horizontal and vertical stretches of the graph.

Student
view

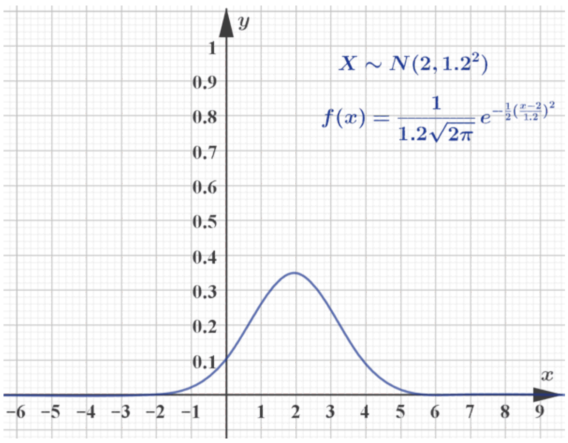
Overview
(/study/ap
122-
cid-
754029/



More information

The image is a graph showing a normal distribution curve. The X-axis ranges from -6 to 7 and represents a variable, while the Y-axis represents probability density, ranging from 0 to 0.35. The text in the graph includes the notation for a normal distribution, "X ~ N(2, 0.4²)", and the probability density function, "f(x) = 1/0.1√2π * e^(-1/2(x-2)²/0.4²)", indicating a normal distribution with mean 2 and standard deviation 0.4. The curve sharply peaks at x = 2, which is the mean of the distribution, and tails off symmetrically on both sides, illustrating typical characteristics of a normal distribution.

[Generated by AI]



More information

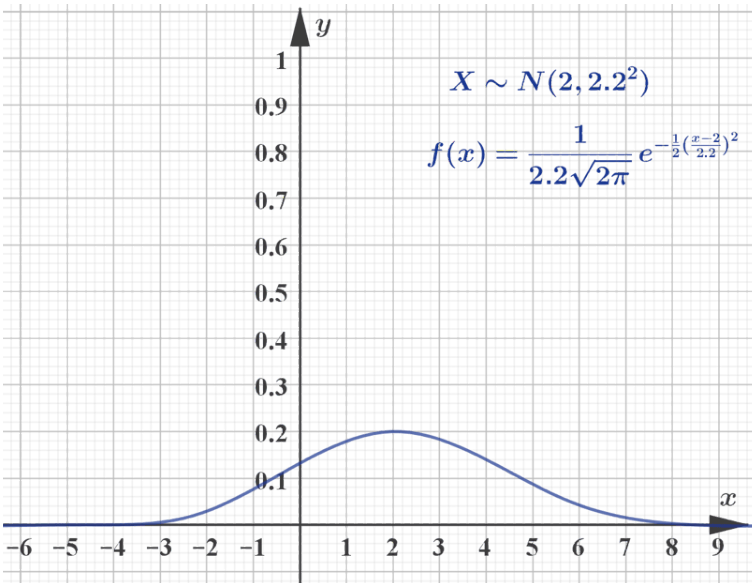
The image depicts a graph representing a probability distribution. The X-axis is labeled with numbers ranging from -5 to 5. The Y-axis has values ranging from 0 to 0.4 in increments of 0.1. The graph features a blue curve with a peak around zero, indicating the highest probability density at this point. The text

Student view

Overview
(/study/ap
122-
cid-
754029/

within the image includes formulas: "X ~ N(2,1.5²)" and "f(x) = 1/(1.5√2π) e^{-(x-2)²/4.5}". This suggests a normal distribution with a mean of 2 and a variance of 1.5 squared.

[Generated by AI]



The standard deviation, σ, determines the width of the normal curve. Manipulating σ is equivalent to performing a horizontal stretch and a vertical stretch. For a larger view, click on an individual graph.

More information

This graph illustrates a normal distribution curve. The X-axis represents the data points in terms of standard deviations from the mean, ranging from -5 to +5. The Y-axis indicates the probability density, with intervals marked at 0.02, 0.04, 0.06, 0.08, and 0.1. The curve peaks at the mean, with probability density decreasing symmetrically as it moves away from the center. A mathematical equation and the notation 'X ~ N(2, 2^2)' are present, describing the normal distribution.

[Generated by AI]

Student view

You can try changing the mean and standard deviation to transform the normal curve in the applet below. Can you anticipate how the model will change before you move the slider?

Overview
(/study/ap
122-
cid-
754029/

Interactive 1. Explore How Changing the Mean and Standard Deviation Affects the Normal Curve.

More information for interactive 1

The interactive provides a dynamic way to explore the effects of changing the mean (m) and standard deviation (s) on the normal distribution curve. Users can adjust the mean within a range from $-10$to 10, allowing them to observe how the curve shifts horizontally. Increasing the mean moves the curve to the right, while decreasing it shifts the curve to the left. This horizontal translation illustrates how the central tendency of the data changes with different mean values.

The standard deviation can be adjusted from 0.1 to 3, enabling users to see how the spread of the distribution varies. A smaller standard deviation, such as 0.1, results in a taller and narrower curve, indicating that the data points are closely clustered around the mean. Conversely, a larger standard deviation, like 3, produces a shorter and wider curve, showing that the data is more spread out. This manipulation demonstrates both horizontal and vertical stretching of the graph, illustrating how variability in the data impacts the distribution's shape.

By experimenting with these ranges, users can anticipate how the model will change. They can observe that the mean sets the center of the distribution, while the standard deviation controls its width and height.

Student
view

Overview
(/study/ap
122-
cid-
754029/

Making connections

The probability is not just illustrated as the area under a normal curve, it actually **equals** the area. In subtopic 5.5 (/study/app/m/sid-122-cid-754029/book/the-big-picture-id-26293/) you will learn how to calculate a definite integral, which in its most basic interpretation is the area between a function and the x-axis.

Theory of Knowledge

Treating data as 'normally distributed' requires an assumption that the data in question is indeed normally distributed. Is mathematics unique as an area of knowledge in regard to its integration of assumptions into knowledge construction? What is the **justification** for the inclusion of assumptions and do other areas of knowledge **justify** assumptions in the same way?

Expressing the normal distribution algebraically

As you saw in subtopic 4.8.1 (/study/app/m/sid-122-cid-754029/book/the-binomial-distribution-id-26261/), the notation $X \sim$ denotes that a random variable X **is distributed** a certain way. In the figures above, you saw that the shape of the normal curve depends on the mean, μ, and the standard deviation, σ (or the variance σ^2). The normal distribution can therefore be defined by:

$$X \sim N(\mu, \sigma^2).$$

This notation tells us that we have a normal distribution (N) with a certain mean and variance. As you will see, the standard deviation (σ) will be used in most work with the normal distribution, so rather than expressing a distribution with $\mu = 10$ and $\sigma = 3$ as $X \sim N(10, 9)$, you will commonly see it written $X \sim N(10, 3^2)$ to signal you to use the standard deviation.

Student
view

Another way to express the normal distribution is with the **normal probability density function**. A probability density function is the equation of a curve that has the probabilities as the area underneath. The normal probability density function is a two-parameter function: it depends on the mean, μ, and the standard deviation, σ. These numbers define the normal distribution completely. If $X \sim N(\mu, \sigma^2)$, we have

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},$$

where $-\infty < x < \infty$.

Be aware

Notice that the sample space, theoretically, for a normal distribution is the set of all real numbers. When modelling real life there may be limitations for that model, but the distribution theoretically remains the same.

Exam tip

You do not need to remember this formula. Your calculator has built-in applications to do calculations involving normal distribution.

It is not easy to verify, but this probability density function is a valid probability distribution, meaning that the sum of the probabilities for the sample space equals 1. Since probability is represented by the area under a curve and the integral of a function gives the area bound by the function and the x-axis, we know that

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1\,.$$

Overview
(/study/ap
122-
cid-
754029/

Probabilities in the normal distribution

The probability of an event within a normal distribution is found by taking the integral of the probability density function (that is, the area beneath the curve). In the next section we will see how to accomplish this with the graphing calculator. But first, we will examine the relationship between probabilities and area in general.

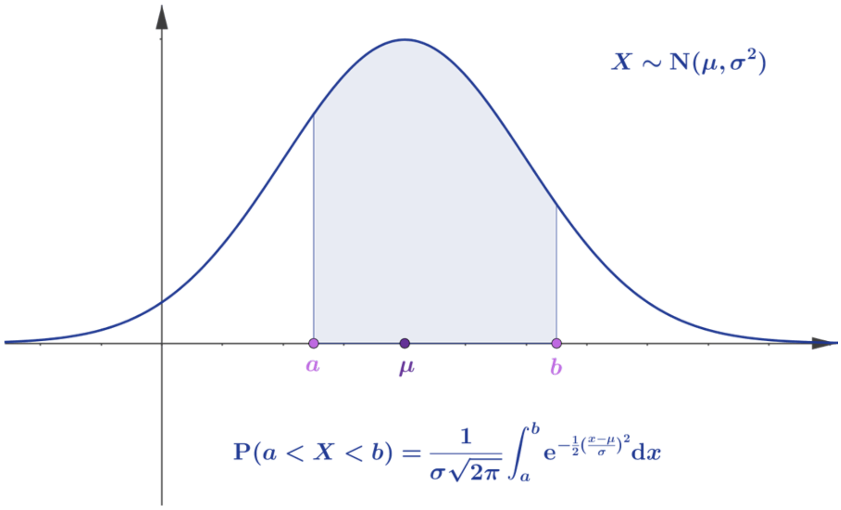
In the table below, we explore three general scenarios associated with finding the probability that a normally distributed variable lies within a given range if $X \sim N(2, 1.5^2)$: (1) the probability that X has a value within an interval $a \leqslant x \leqslant b$, (2) the probability that X has a value greater than (or less than) a certain measure and (3) the probability that X has a value beyond a certain distance from the mean.

Using the information in the table below, we explore three general scenarios associated with finding the probability that a normally distributed variable lies within a given range.

Three scenarios for the probability that a normally distributed variable lies in a range, u
symmetry properties of the normal distribution and the property of the probability densit
For a larger view of a graph, click on it.

Student
view

Overview
(/study/ap
122-
cid-
754029/

Scenario	Example for $X \sim \mathrm{N}($
	The probability that X lie between 1.2 and 3.5: $\mathrm{P}(1.2 < X < 3.5) = \frac{1}{1.5\sqrt{2\pi}} \int_{1.2}^{3.5} \mathrm{e}^{-\frac{1}{2}(\frac{x}{1}}$ ≈ 0.544

More information

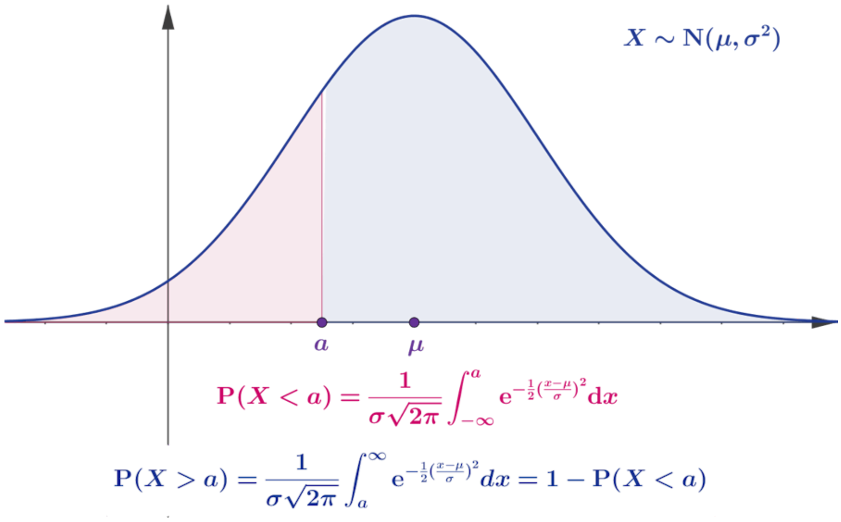
The image shows a graph representing a normal distribution curve. The X-axis represents the possible values of a variable X, marked by points labeled 'a,' 'μ,' and 'b,' with 'μ' being the mean. The Y-axis represents the probability density. The bell-shaped curve is centered at the mean, μ. The area between points 'a' and 'b' under the curve is shaded, representing the probability P(a < X < b).

Text in the image includes: - "X ~ N(μ,σ²)" indicating the variable X follows a normal distribution with mean μ and variance σ². - The probability formula at the bottom of the image: "P(a < X < b) = 1 / (σ√2π) ∫ from a to b of e^(-1/2 * (x-μ/σ)²) dx" indicates the mathematical expression to calculate the shaded area under the curve.

[Generated by AI]

Student view

Overview (/study/ap 122- cid- 754029/

Scenario	Example for $X \sim \mathrm{N}(2$
 More information	The probability that X is than 1.2: $\mathrm{P}(X < 1.2) = \frac{1}{1.5\sqrt{2\pi}} \int_{-\infty}^{1.2} \mathrm{e}^{-\frac{1}{2}(\frac{x}{1}}$ ≈ 0.297
	The probability that X is than 1.2: $\mathrm{P}(X > 1.2) = \frac{1}{1.5\sqrt{2\pi}} \int_{1.2}^{\infty} \mathrm{e}^{-\frac{1}{2}(\frac{x-}{1.}}$ $= 1 - \mathrm{P}(X < 1.2)$ $\approx 1 - 0.297$ $= 0.703$

The image is a graph depicting a normal distribution, represented by a bell-shaped curve. The graph includes two distinct shaded areas under the curve: one on the left shaded in pink and another on the right in blue. These shaded areas represent probabilities in a normal distribution, often for a random variable X. The x-axis is labeled with "a" and "μ" (mean) as key points, marked by dots. The y-axis has no numerical labels.

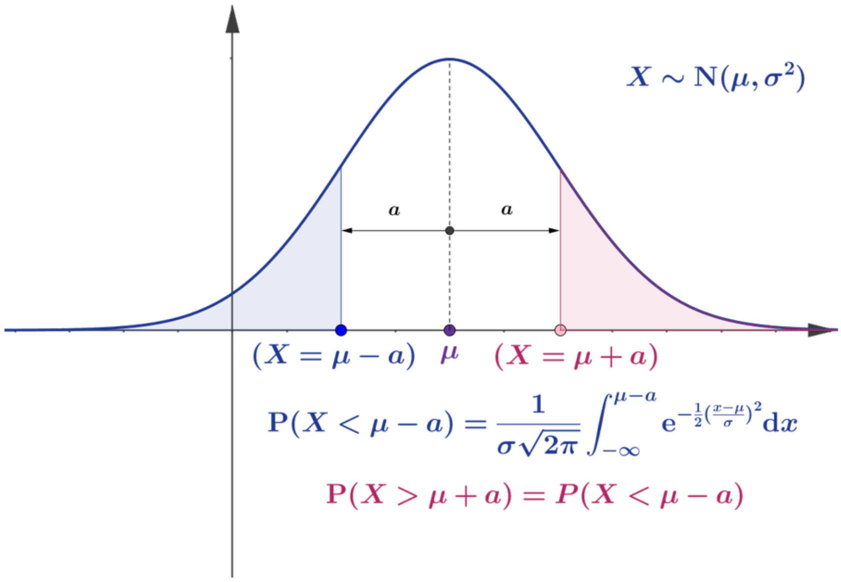
Text over the graph indicates the distribution as "X ~ N(μ,σ²)" or normally distributed with mean μ and variance σ². Equations at the bottom express cumulative probabilities: - P(X < a) = (\frac{1}{\sigma \sqrt{2\pi}}\int_{-\infty}^{a} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx) - P(X > a) = (1 - P(X < a)) = (\frac{1}{\sigma \sqrt{2\pi}}\int_{a}^{\infty} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx)

The graph highlights the concept of area under the curve corresponding to probabilities, emphasizing how these probabilities are computed using the standard normal distribution equation.

[Generated by AI]

Student view

Overview
(/study/ap
122-
cid-
754029/

Scenario	Example for $X \sim \mathrm{N}(2$
 More information The image is a graph depicting a normal distribution, represented by a bell-shaped curve, centered at the mean (μ). The X-axis is labeled with points (X = μ - a), μ, and (X = μ + a), indicating the mean and its deviations. The Y-axis has no specific label but represents probability density. The graph includes shaded areas on the left and right tails, which correspond to probabilities P(X < μ - a) and P(X > μ + a) respectively. Below the graph, two equations are shown: 1. P(X < μ - a) = (1/σ√2π)∫ from -∞ to (μ - a) of e^(-1/2((x-μ)/σ)^2) dx. 2. P(X > μ + a) = P(X < μ - a). These equations calculate the probabilities of occurrences within the shaded regions under the curve, demonstrating the characteristics of a normal distribution. [Generated by AI]	Values symmetric aroun mean μ correspond to th probabilities. For exampl $\mathrm{P}(X < 2 - 1.5 = 0.5)$ $= \frac{1}{1.5\sqrt{2\pi}} \int_{-\infty}^{0.5} \mathrm{e}^{-\frac{1}{2}(\frac{x}{1}}$ ≈ 0.159 And $\mathrm{P}(X > 2 + 1.5 = 3.5)$ $= \frac{1}{1.5\sqrt{2\pi}} \int_{3.5}^{\infty} \mathrm{e}^{-\frac{1}{2}(\frac{x-}{1.}}$ $= \mathrm{P}(X < 2 - 1.5 = 0.$ ≈ 0.159

As you can see, each probability depends on the mean and standard deviation, and each can be calculated with an integral evaluated from a lower bound to an upper bound. We will use technology to calculate these integrals in section 4.9.2

Student view

Overview (/study/app/m/sid-122-cid-754029/

(/study/app/m/sid-122-cid-754029/book/the-normal-distribution-and-calculator-functions-id-26267/), but there are some instances when we can approximate probabilities without directly using the integral.

When we know some probabilities, we can use the graph to find others. Here are two examples taken from what we found in the table .

- We discovered that $\mathrm{P}(\mathrm{X}<1.2)=0.297$. Using the complement to an event, we can also find $\mathrm{P}(\mathrm{X}>1.2)=1-0.297=0.703$.
- You can also use the complement to find the $\mathrm{P}(\mu-\mathrm{a}<\mathrm{X}<\mu+\mathrm{a})$. Since $\mathrm{P}(\mathrm{X}<0.5)=\mathrm{P}(\mathrm{X}>3.5)\approx 0.159$, we find that $\mathrm{P}(0.5<\mathrm{X}<3.5)\approx 1-2(0.159)\text{ approx}0.682$.

Making connections

Previously you learned about the percentile of a value within a data set. In this example, since $\mathrm{P}(X<1.2)=0.297$, the value 1.2 is close to the 30 th percentile. Likewise, 0.5 is close to the 16th percentile and 3.5 is close to the 84th percentile.

Example 1

★★☆

If $X\sim N(12,\sigma^2)$, $\mathrm{P}(\mathrm{X}<11)=0.369$ and $\mathrm{P}(\mathrm{X}>14)=0.252$, find the following:

a) $\mathrm{P}(\mathrm{X}>11)$

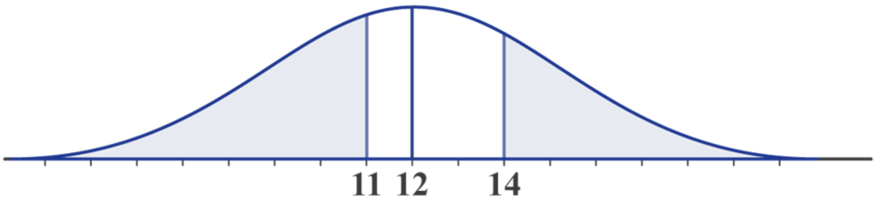
b) $\mathrm{P}(11<\mathrm{X}<14)$

c) The approximate percentile of $X=14$.

Student view

Overview
(/study/ap
122-
cid-
754029/

First, make a sketch of the normal curve and shade in the given probabilities.



Sketch the normal curve and identify the mean and given probabilities.

a) Use the complement principle to find
$\mathrm{P}(\mathrm{X} > 11) = 1 - \mathrm{P}(\mathrm{X} < 11) = 1 - 0.369 = 0.631$.

b) You can see that $11 < X < 14$ lies between the two given probabilities. Since you already know the area above $X = 11$ is 0.631, subtract the given $\mathrm{P}(\mathrm{X} > 14) = 0.252$. Hence,
$\mathrm{P}(11 < \mathrm{X} < 14) = \mathrm{P}(\mathrm{X} > 11) - \mathrm{P}(\mathrm{X} > 14) = 0.631 - 0.252 = 0.379$
.

c) The percentile corresponds to the percentage of data that lies below a certain value. Since you are given $\mathrm{P}(\mathrm{X} > 14) = 0.252$, you know that $\mathrm{P}(\mathrm{X} < 14) = 1 - \mathrm{P}(\mathrm{X} > 14) = 1 - 0.252 = 0.748$. Therefore, 74.8% of the data lies below 14, meaning 14 is approximately the 75th percentile.

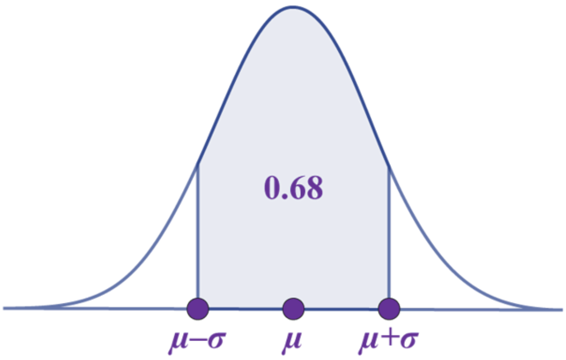
Finally, we can find approximate probabilities using intervals defined by exact standard deviations from the mean.

For any normally distributed data set, which has mean μ with standard deviation σ, the proportion of the diagram shaded between 1 standard deviation less than the mean $(\mu - \sigma)$ and 1 standard deviation more than the mean $(\mu + \sigma)$ is 0.68

Student view

Overview
(/study/ap
122-
cid-
754029/

(correct to 2 significant figures) and denoted by $\mathrm{P}(\mu - \sigma < X < \mu + \sigma) = 0.68$. This is illustrated below.



Approximately 68% of normally distributed data is within 1 standard deviation of the mean.

More information

The image is a bell-shaped curve representing a normal distribution. The curve is centered on the mean, denoted as μ, with lines extending horizontally in both directions to μ - σ and μ + σ, where σ is the standard deviation. The area under the curve between μ - σ and μ + σ is shaded, representing approximately 68% of the total area under the curve. This visually demonstrates the probability that a normally distributed variable falls within one standard deviation from the mean.

[Generated by AI]

Example 2

★☆☆

X represents a catch of fish that have a mean weight of 1.8kg and a standard deviation of 0.5kg such that $X \sim N(1.8, 0.5^2)$. Find an interval, centred around the mean, that contains the weights of approximately 68% of the fish.

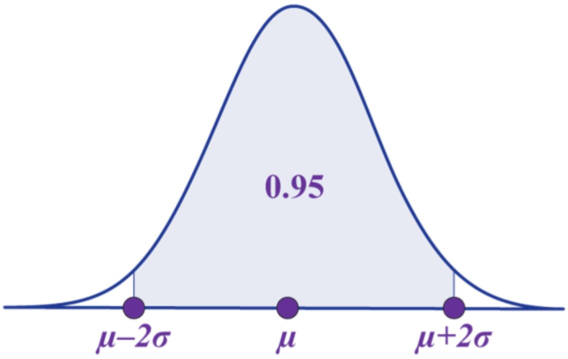
Student view

Since approximately 68% of the fish in a normal distribution will have a weight within 1 standard deviation of the mean, the interval is 1.8 ± 0.5.

Overview
(/study/ap
122-
cid-
754029/

This means $\mathrm{P}(1.3 < \mathrm{x} < 2.3) \approx 0.68$.

Similarly, we estimate the probability of being within 2 standard deviations of the mean as 95% and the probability of being within 3 standard deviations of the mean as 99.7%. These are illustrated in the figures below.



Approximately 95% of normally distributed data lies within 2 standard deviations of the mean.

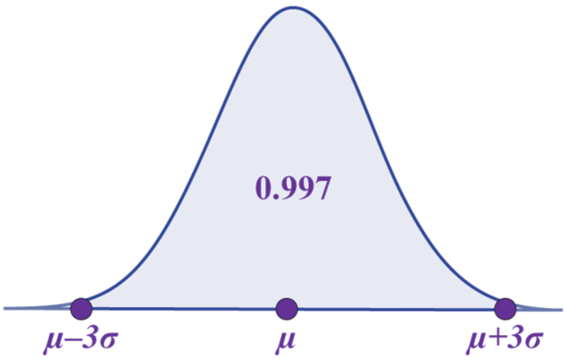
More information

The image shows a bell curve representing a normal distribution. This graph depicts a symmetrical, bell-shaped curve centered at the mean (μ). It emphasizes that approximately 95% of the data lies between two points: μ-2σ and μ+2σ, which are two standard deviations away from the mean on either side. The central section of the curve is highlighted, showing the area where 95% of data points reside. The X-axis is labeled with μ (mean) at the center, with μ-2σ on the left and μ+2σ on the right. A label inside the curve reads "0.95," indicating the probability of data within this range.

[Generated by AI]

Student
view

Overview
(/study/ap
122-
cid-
754029/



Approximately 99% of normally distributed data lies within 3 standard deviations of the mean.

More information

The image displays a bell-shaped curve, representing a normal distribution. The horizontal axis marks the mean at the center with the symbol 'μ'. On either side of the mean are symbols 'μ-3σ' and 'μ+3σ', indicating three standard deviations below and above the mean respectively. The graph highlights that approximately 99.7% of the data under a normal distribution curve falls within this range. The figure 0.997 is prominently displayed in the body of the curve, reinforcing the percentage of data capture within these bounds.

[Generated by AI]

Important

In summary, we have these estimated probabilities:

- $\mathrm{P}(\mu - \sigma < \mathrm{X} < \mu + \sigma) = 0.68$
- $\mathrm{P}(\mu - 2\sigma < \mathrm{X} < \mu + 2\sigma) = 0.95$
- $\mathrm{P}(\mu - 3\sigma < \mathrm{X} < \mu + 3\sigma) = 0.997.$

Regardless of the values of the mean and standard deviation, these probabilities remain the same due to the nature of the normal distribution. You need to memorise them.

Student
view

Overview
(/study/app/m/sid-122-cid-754029/

Example 3

★★☆

The time T it takes Yasmin to get to school is normally distributed, such that $T \sim N(35, 2.5^2)$. She is 99.7% confident that it will take her at least 27.5 minutes and less than t minutes to make it to school. Find t.

First, since $\sigma = 2.5$, you can determine that $27.5 = \mu - 3\sigma$.

A 99.7% probability corresponds to the probability of being within 3 standard deviations of the mean, so that means t must be the value exactly 3σ above the mean.

Therefore, $t = \mu + 3\sigma = 35 + 7.5 = 42.5$.

Example 4

★★☆

The number of chocolates in a bag is normally distributed with mean 43 and standard deviation 3. Find the probability that there are more than 37 chocolates in a bag.

First, recognise that for this situation $37 = \mu - 2\sigma$.

The probability that the number of chocolates is within 2 standard deviations of the mean is 0.95, which means
$\mathrm{P}(\text{chocolates } > 37) = 0.95 + \mathrm{P}(\text{more than } 2\sigma)$.

Since there is 0.05 probability the number is beyond 2 standard deviations, and the normal distribution is symmetrical, $\mathrm{P}(\text{more than } 2\sigma) = 0.025$.

Student view

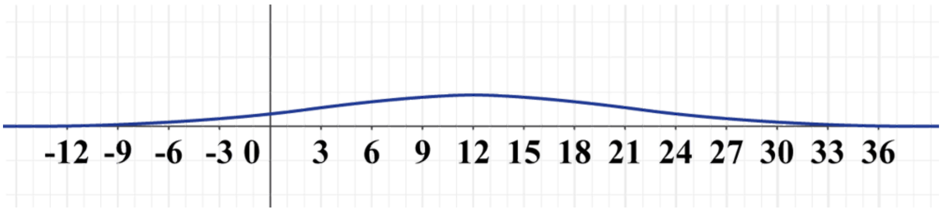
Overview
(/study/ap
122-
cid-
754029/

Therefore, $\mathrm{P}(\text{chocolates } > 37) = 0.95 + 0.025 = 0.975.$

Example 5

★★☆

Determine which of these normal curves represents $X \sim N(12, 3^2)$.

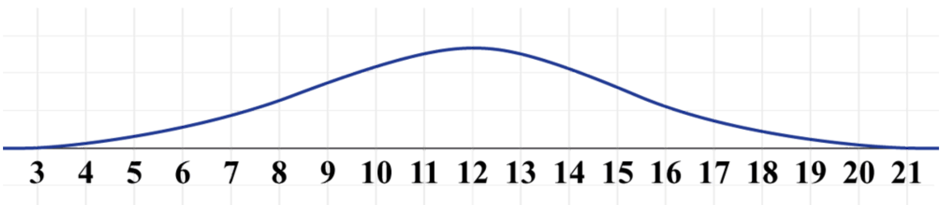


a

More information

The image depicts a normal distribution curve on a graph. The X-axis represents the variable X, which follows a normal distribution with a mean of 12 and a standard deviation of 3 (denoted X ~ N(12, 3²)). The Y-axis typically represents the probability density or frequency. The curve is bell-shaped, symmetrical around the mean value of 12, indicating the normal distribution characteristics. Values near the mean are more frequent than those far away, illustrating the properties of the normal curve.

[Generated by AI]



b

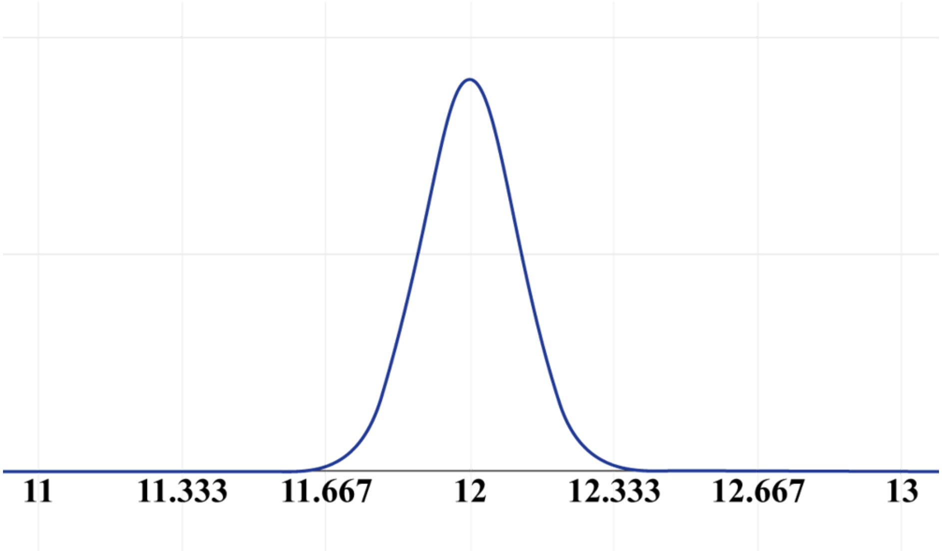
More information

Student
view

Overview
(/study/ap
122-
cid-
754029/

The image is a graph displaying a bell curve plotted on a grid. The X-axis is horizontal and does not have labeled values visible in the image provided. The Y-axis is vertical, and like the X-axis, it lacks visible labels or specific values. The overall trend of the graph shows a rise to a peak at the center and then a symmetrical decline on either side, typical of a normal distribution curve.

[Generated by AI]



c

More information

The image is a graph showing a bell curve, which is a representation of a normal distribution.

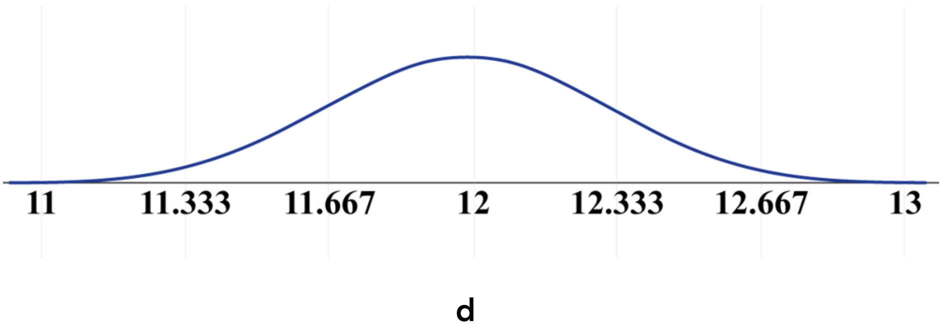
- **X-axis:** The horizontal axis represents a range of values, centered around 0, and evenly spaced at intervals of 1 and 2 units, extending from approximately -2 to 2.
- **Y-axis:** The vertical axis shows the probability density, which peaks at the center and gradually decreases to near zero as it moves towards the edges, indicating a normal distribution curve.
- **Curve Details**: The blue line peaks sharply at the center, meaning the highest probability density occurs at the mean 0, and symmetrically decreases as it moves outward.

This graph visually represents a probability density function of a standard normal distribution, showing how data is distributed around a mean with most of the values clustering around the central peak.

[Generated by AI]

Student
view

Overview
(/study/ap
122-
cid-
754029/



More information

The image displays a bell curve graph representing a symmetrical distribution. On the X-axis, the values range from -3 to 3 with increments of 1, suggesting standard deviations from the mean. The Y-axis is labeled from 0.0 to 0.4, likely representing probability density. The peak of the curve is at the center of the graph, aligning with the mean value of 0, at the highest probability density around 0.4, indicating most data points cluster around the mean with symmetric tails tapering towards the lower and upper extremes. The curve demonstrates the classic characteristics of a normal distribution.

[Generated by AI]

First, determine that since $X \sim N(\mu, \sigma^2)$, we can see that $\mu = 12$ and $\sigma = 3$.

All of the choices are centred at $x = 12$, so we must look at the spread of the data. Since the standard deviation is 3, approximately 68% of the area under the curve should be within 1 standard deviation of the mean, i.e. between $12 - 3 = 9$ and $12 + 3 = 15$. Also, 95% of the area should be within $\mu \pm 2\sigma$, or between 6 and 18.

Considering this, choices c and d are easy to eliminate, because nearly all of the area under the curve lies between 11 and 13, making σ for those graphs far too small.

Student
view

Overview
(/study/ap
122-
cid-
754029/

For choice a, look at the area to the left of 6 and to the right of 18. It is clearly more than 5% of the total area, so a can be eliminated too.

This means that choice b is the best choice. Clearly more than half of the area is within the range of 9 to15, so that is consistent with the 68% we are looking for within 1 standard deviation of the mean. Looking at the areas beyond 6 and 18, we see only a small area, giving us a reasonable picture of the remaining 5%.

3 section questions

4. Probability and statistics / 4.9 The normal distribution and curve

The normal distribution and calculator functions

Finding probabilities on the calculator

In section 4.9.1 (/study/app/m/sid-122-cid-754029/book/the-normal-distribution-id-26266/), you learned how to find some probabilities within the normal distribution, but you were limited to situations involving values that were exactly 1, 2 or 3 standard deviations from the mean. In this section you will learn how to use technology to find the probability of a random variable falling within any range of values in a normal distribution. The normal distribution is so commonly used that most graphic display calculators (GDCs) have a built-in **normal cumulative distribution function**, or Normal CDF.

Suppose $X \sim N(2, 1.5^2)$. The instructions below show how to use your GDC to find three probabilities.

Student
view

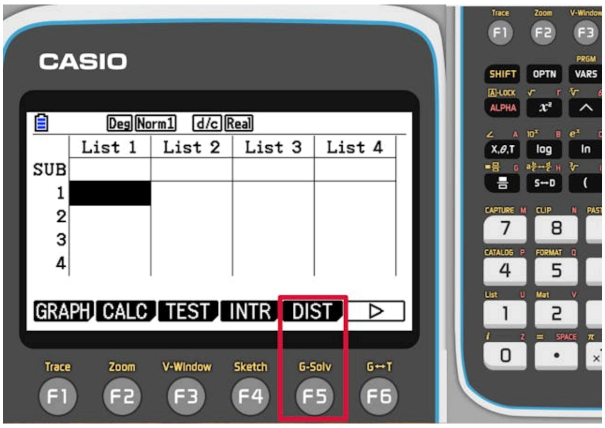
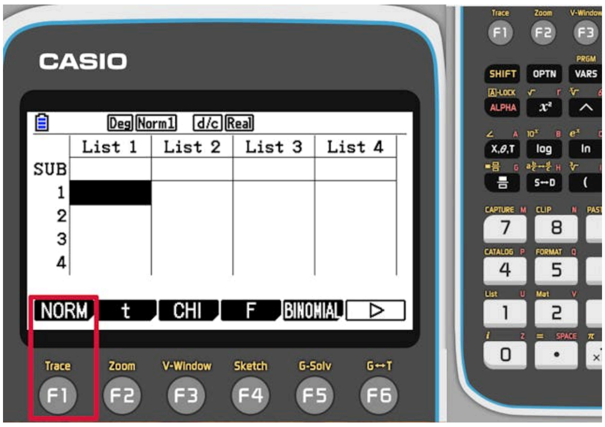
Overview
(/study/app/m/sid-122-cid-754029/

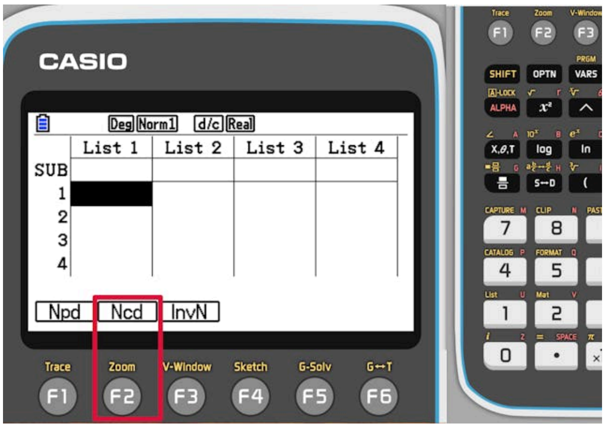
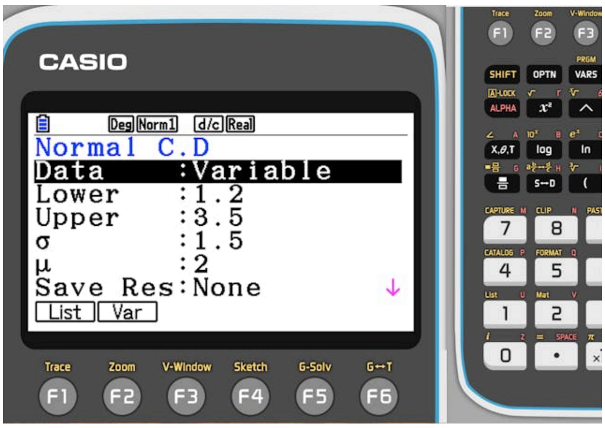
- $\mathrm{P}(1.2 < \mathrm{X} < 3.5)$
- $\mathrm{P}(\mathrm{X} < 1.2)$
- $\mathrm{P}(\mathrm{X} < 0.5)$ or $\mathrm{P}(\mathrm{X} > 3.5)$

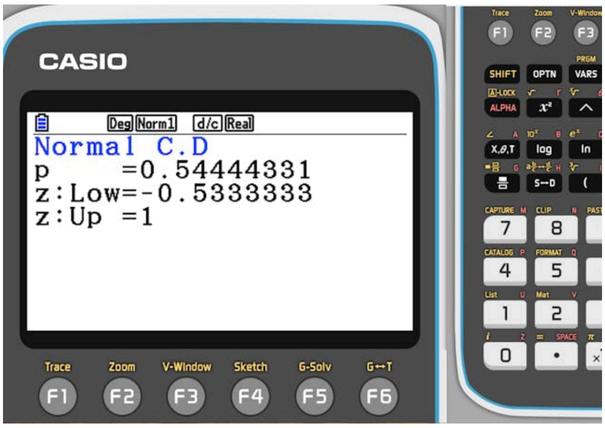
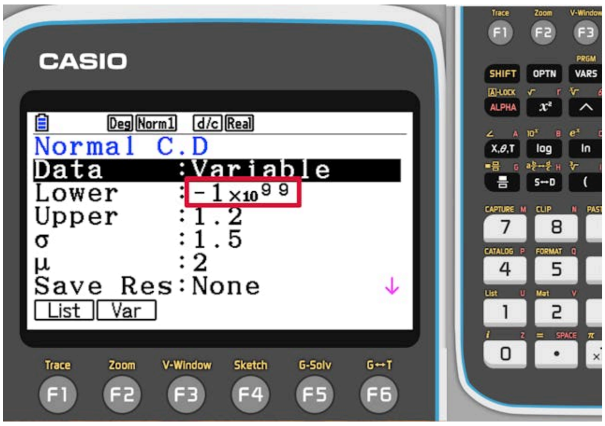
You saw these three probabilities illustrated previously in section 4.9.1 (/study/app/m/sid-122-cid-754029/book/the-normal-distribution-id-26266/).

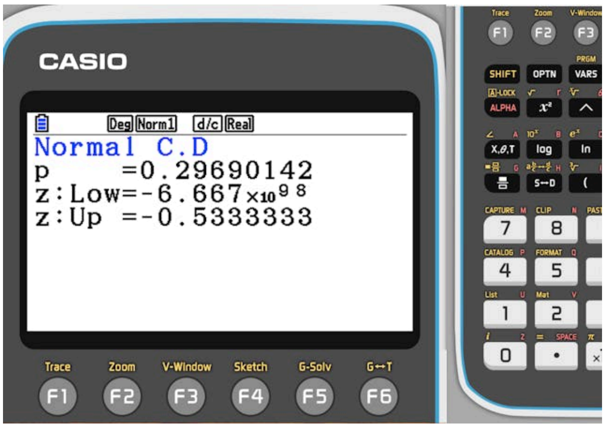
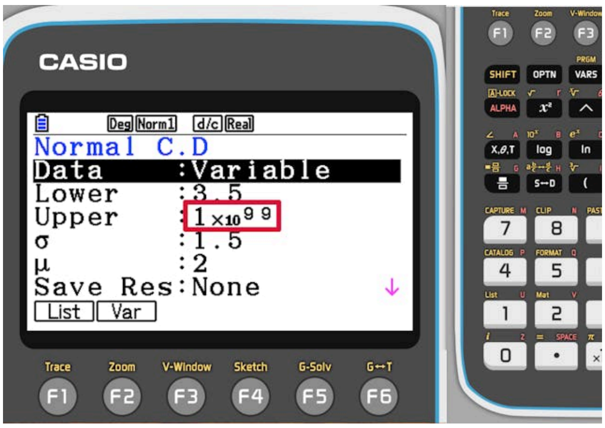
Step	Explanation
In these instructions you will see how to find $\mathrm{P}(1.2 < X < 3.5)$, $\mathrm{P}(X < 1.2)$ and $\mathrm{P}(X > 3.5)$ for a normally distributed X with mean $\mu = 2$ and standard deviation $\sigma = 1.5$. Choose the statistics mode.	

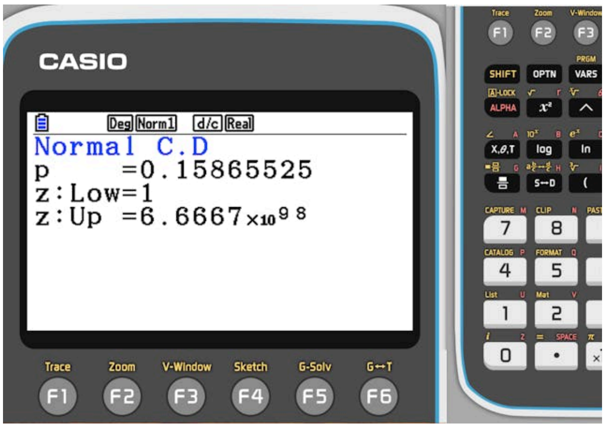
Student view

Step	Explanation
Press F5 to select distributions.	
Press F1 to select normal distribution.	

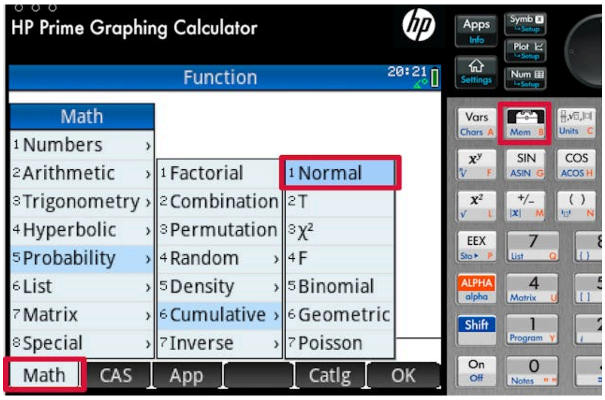
Step	Explanation
Press F2 to select the cumulative normal distribution (Ncd) function.	
Enter the mean $\mu = 2$ and standard deviation $\sigma = 1.5$. Enter also the appropriate lower and upper bounds to find the probability $\mathrm{P}(1.2 < X < 3.5)$.	

Step	Explanation
The calculator tells you the probability, p.	
You can use similar steps to find $\mathrm{P}(X < 1.2)$. In entering the parameters, you need to be careful with the lower bound. In the probability you are trying to find there is no lower bound. The calculator needs a lower bound, so you need to enter a large negative number.	

Step	Explanation
The calculator tells you the probability, p.	
In finding $\mathrm{P}(x > 3.5)$, there is no upper bound, so you need to enter a large positive number here.	

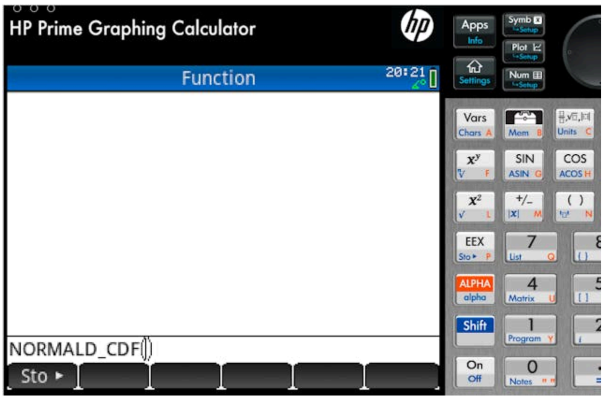
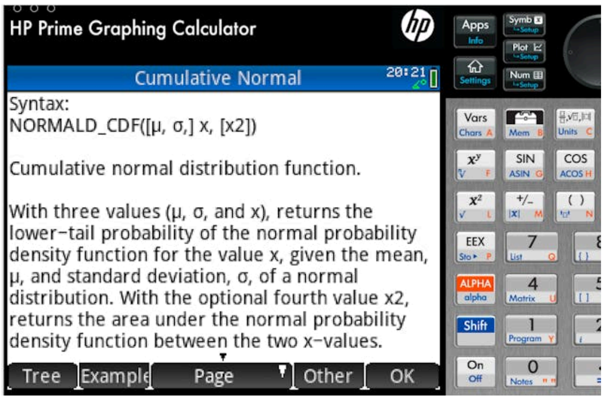
Step	Explanation
The calculator tells you the probability, p.	

Overview
(/study/ap
122-
cid-
754029/

Step	Explanation
In these instructions you will see how to find $\mathrm{P}(1.2 < X < 3.5)$, $\mathrm{P}(X < 1.2)$ and $\mathrm{P}(X > 3.5)$for a normally distributed X with mean $\mu = 2$ and standard deviation $\sigma = 1.5$. Enter the home screen of any application.	
Open the toolbox and select the cumulative normal distribution function through the menu system.	

Student
view

Overview
(/study/ap
122-
cid-
754029/(

Step	Explanation
You will need to tell the calculator some parameters. If you are not sure what the calculator expects, open the help screen.	
This is the first page of the help screen for the cumulative normal distribution function. On the next page you can also find some examples.	

Student
view

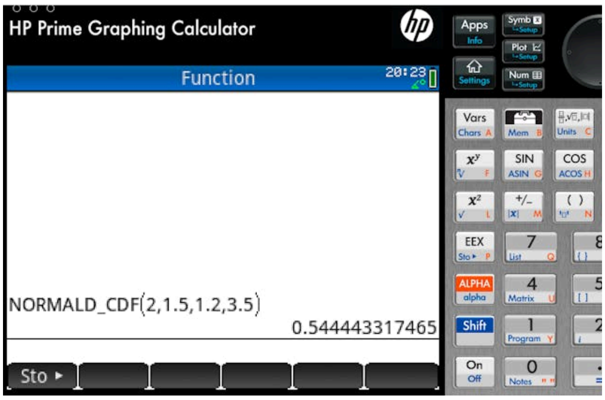
Overview
(/study/ap
122-
cid-
754029/

<table>
<tr><th>Step</th><th>Explanation</th></tr>
<tr><td>

Enter the mean $\mu = 2$ and standard deviation $\sigma = 1.5$. Enter also the appropriate lower and upper bounds to find the probability $\mathrm{P}(1.2 < X < 3.5)$.

The calculator expects these numbers in this order.

</td><td>

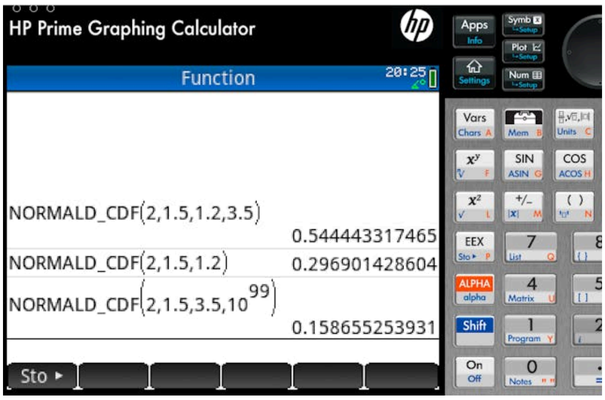


</td></tr>
<tr><td>

You can use similar steps to find $\mathrm{P}(X < 1.2)$. In entering the parameters, you need to be careful with the lower bound. In the probability you are trying to find there is no lower bound. In the second line of the screen you can see that if only three parameters entered, the calculator will treat this as a probability with no lower bound.

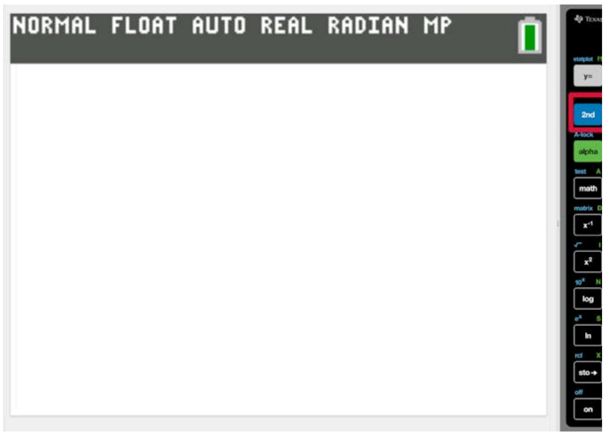
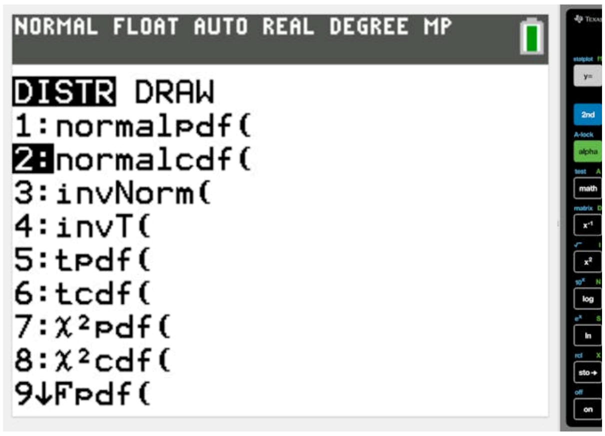
In finding $\mathrm{P}(x > 3.5)$, there is no upper bound. You can tell this to the calculator by entering a large positive number (for example 10^{99}) as the upper bound.

</td><td>

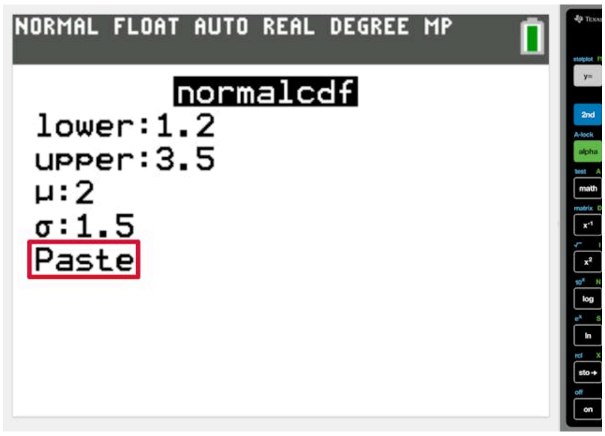
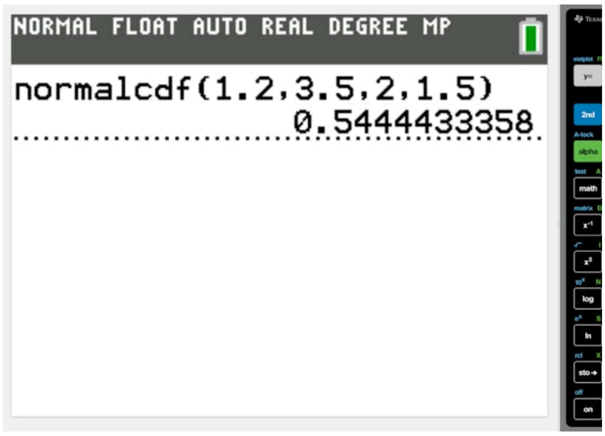


</td></tr>
</table>

Student
view

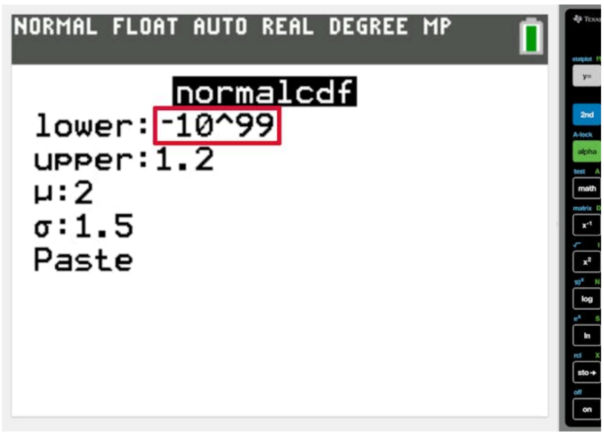
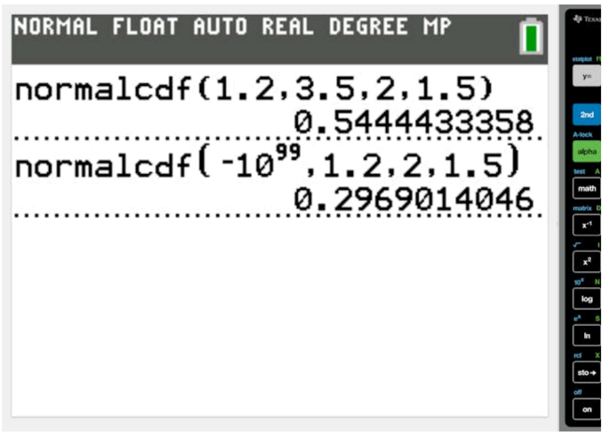
Step	Explanation
In these instructions you will see how to find $\mathrm{P}(1.2 < X < 3.5)$, $\mathrm{P}(X < 1.2)$ and $\mathrm{P}(X > 3.5)$for a normally distributed X with mean $\mu = 2$ and standard deviation $\sigma = 1.5$. Open up the distributions menu.	
Choose the cumulative normal distribution (normalcdf) option.	

Overview
(/study/ap
122-
cid-
754029/

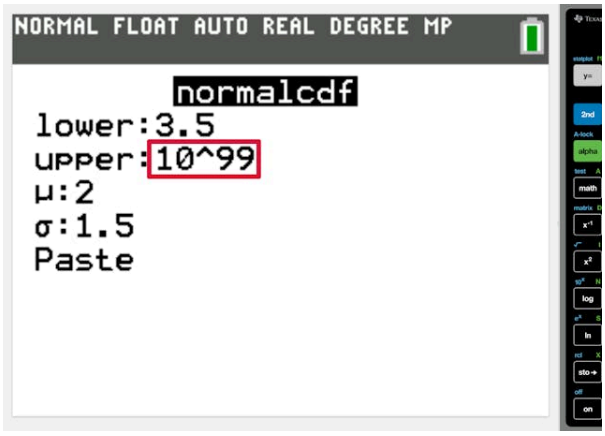
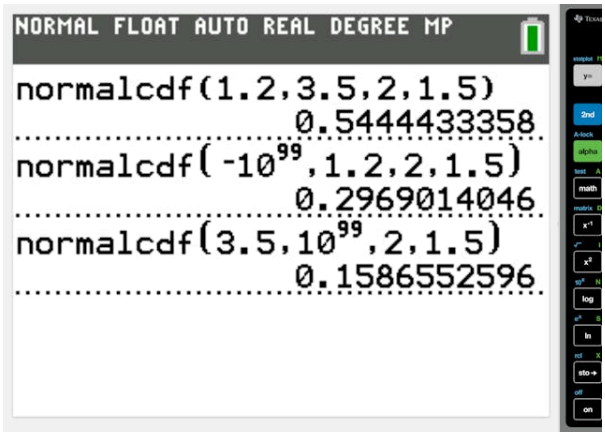
Step	Explanation
Enter the mean $\mu = 2$ and standard deviation $\sigma = 1.5$. Enter also the appropriate lower and upper bounds to find the probability $\mathrm{P}(1.2 < X < 3.5)$. When you are done, scroll down to "paste" and press enter.	
After pressing enter one more time, the calculator tells you the probability.	

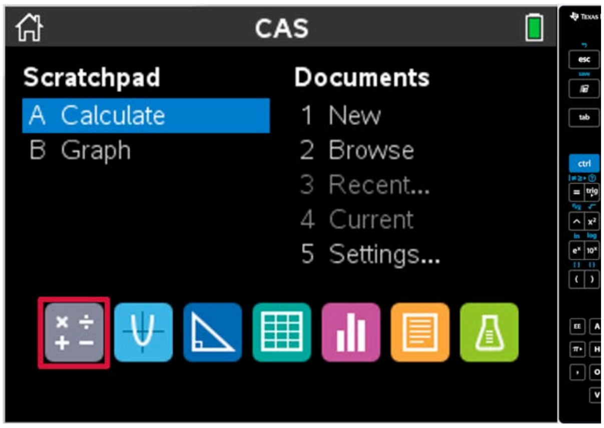
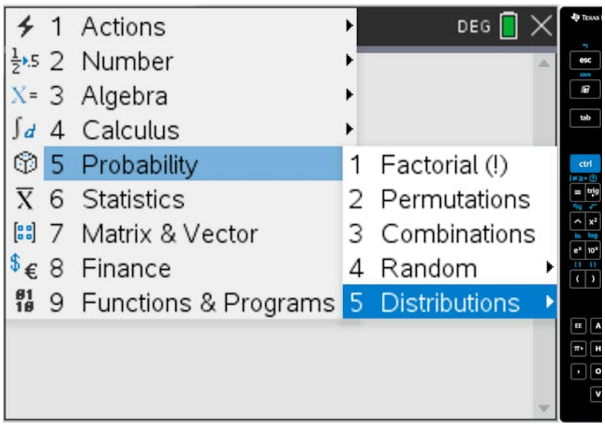
Student
view

Overview
(/study/ap
122-
cid-
754029/

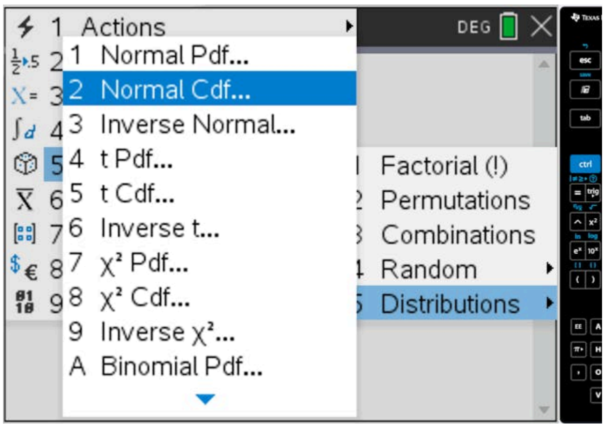
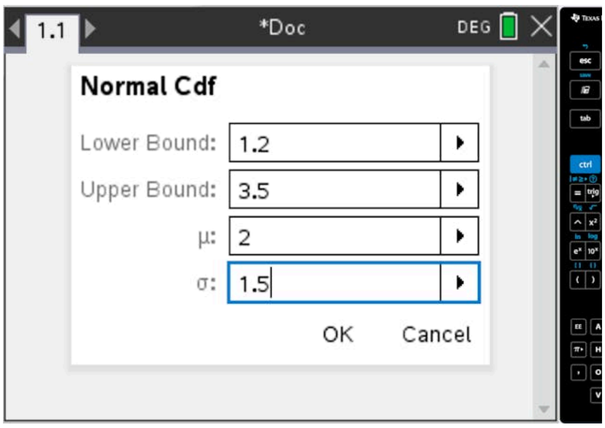
Step	Explanation
You can use similar steps to find $\mathrm{P}(X < 1.2)$. In entering the parameters, you need to be careful with the lower bound. In the probability you are trying to find there is no lower bound. The calculator needs a lower bound, so you need to enter a large negative number.	
After pressing enter one more time, the calculator tells you the probability.	

Student
view

Step	Explanation
In finding $\mathrm{P}(x > 3.5)$, there is no upper bound, so you need to enter a large positive number here.	
After pressing enter one more time, the calculator tells you the probability.	

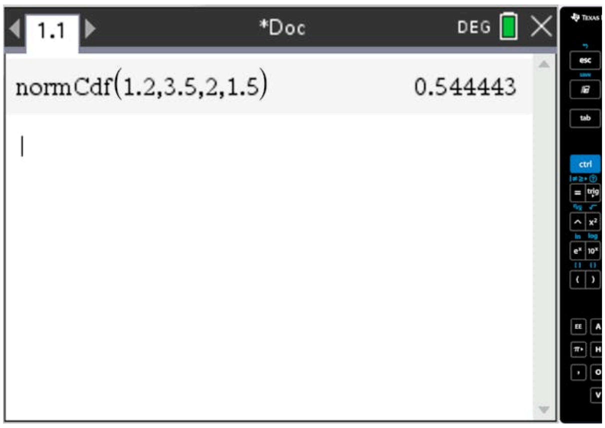
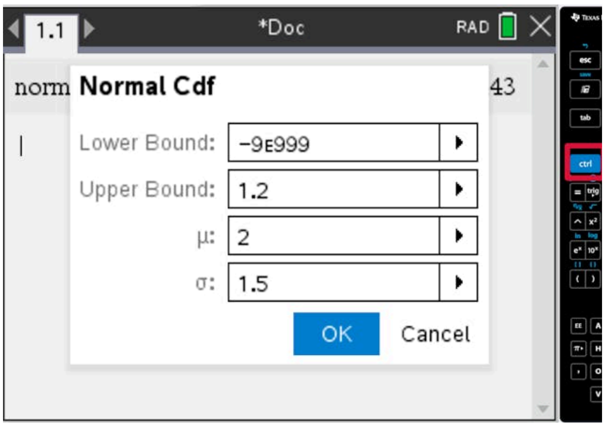
Step	Explanation
In these instructions you will see how to find $\mathrm{P}(1.2 < X < 3.5)$, $\mathrm{P}(X < 1.2)$ and $\mathrm{P}(X > 3.5)$for a normally distributed X with mean $\mu = 2$ and standard deviation $\sigma = 1.5$. Open a calculator page.	
Choose probability distributions through the menu system ...	

Overview
(/study/ap
122-
cid-
754029/

Step	Explanation
... and choose the cumulative normal distribution (Normal Cdf) option.	
Enter the mean $\mu = 2$ and standard deviation $\sigma = 1.5$. Enter also the appropriate lower and upper bounds to find the probability $\mathrm{P}(1.2 < X < 3.5)$.	

Student
view

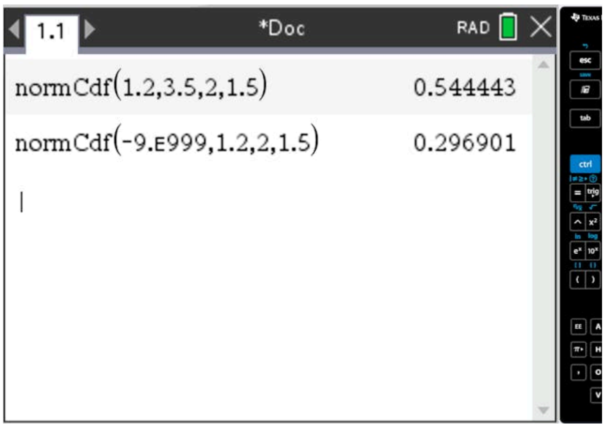
Overview
(/study/ap
122-
cid-
754029/

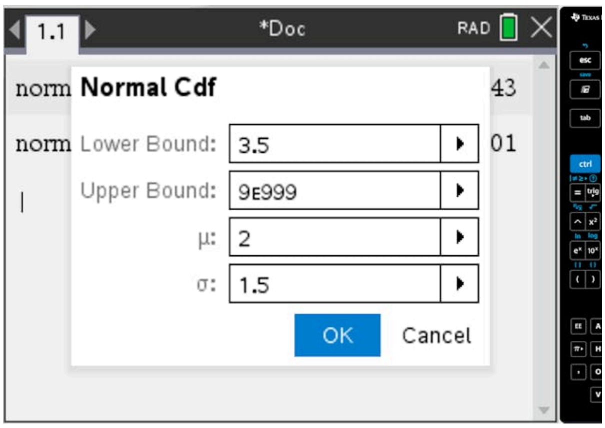
Step	Explanation
The calculator tells you the probability.	
You can use similar steps to find $\mathrm{P}(X < 1.2)$. In entering the parameters, you need to be careful with the lower bound. In the probability you are trying to find there is no lower bound. The calculator needs a lower bound, so you need to enter a large negative number. On some models you can enter the negative infinity symbol ($-\infty$) to indicate that there is no lower bound. If it is available on your calculator, you can find this symbol (among others) in the symbol list.	

Student
view

Overview
(/study/app/m/sid-122-cid-754029/
Section
Student... (0/0)
Feedback
Print (/study/app/m/sid-122-cid-754029/book/the-normal-distribution-id-26266/print/)
Assign

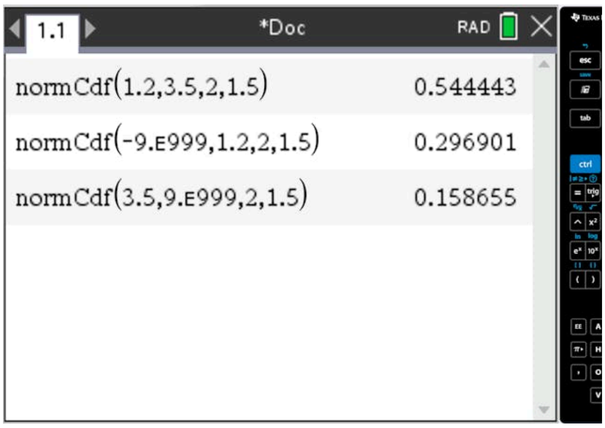
Step	Explanation
The calculator tells you the probability.	(see image 1 below)
In finding $\mathrm{P}(x > 3.5)$, there is no upper bound, so you need to enter a large positive number here. If it is available on your model, you can also enter the infinity symbol to indicate that there is no upper bound.	





Student view

Overview
(/study/ap
122-
cid-
754029/

Step	Explanation
The calculator tells you the probability.	

Exam tip

You must be able to use your GDC to obtain probabilities for the normal distribution. Tables are not provided in the formula booklet, and the normal distribution is so common that it is unlikely an examination would not include the normal distribution.

Activity

Your calculator also probably has a Normal PDF application. This is used only for graphing a normal curve, and you do not need to know how to use it for this course. However, why do you think it might be helpful to see the graph of a normal distribution when working on problems?

Student
view

Overview
(/study/ap
122-
cid-
754029/

Search how to use the Normal PDF on your calculator and practise graphing some. Can you anticipate the settings you should use for the graph window to see the curve clearly?

Using the inverse normal option

The instructions above show how to use the calculator to find values of the cumulative distribution function of the normal distribution. Your calculator also has an option to find values of the inverse of this cumulative distribution function. The following examples illustrate typical situations where this inverse normal option can be helpful to find the answer.

Example 1

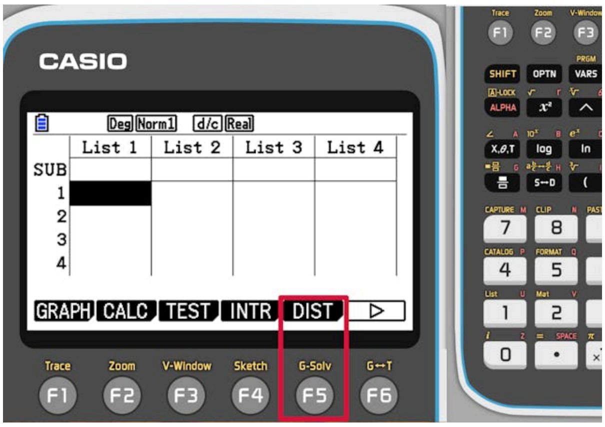
★☆☆

Let $X \sim N(2.2, 0.5^2)$. Find the value of b such that $\mathrm{P}(\mathrm{X} < \mathrm{b}) = 0.72$.

The instructions below show the method for using the calculator to find $b = 2.49$

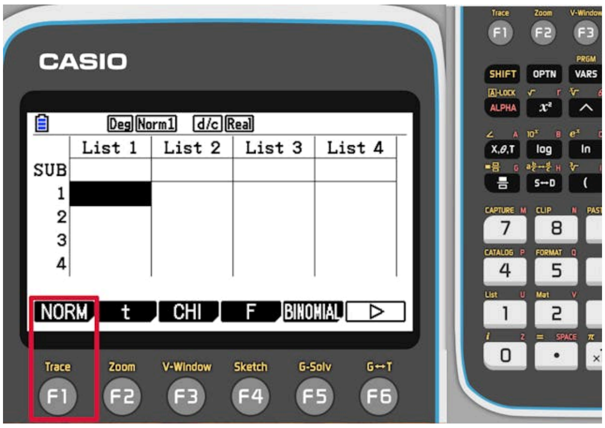
Student
view

Overview
(/study/ap
122-
cid-
754029/

Step	Explanation
These instructions will show you how to find the value of b such that $\mathrm{P}(X < b) = 0.72$, where X is a normal random variable with mean $\mu = 2.2$ and standard deviation $\sigma = 0.5$. Choose the statistics option.	
Press F5 to choose the distribution option.	

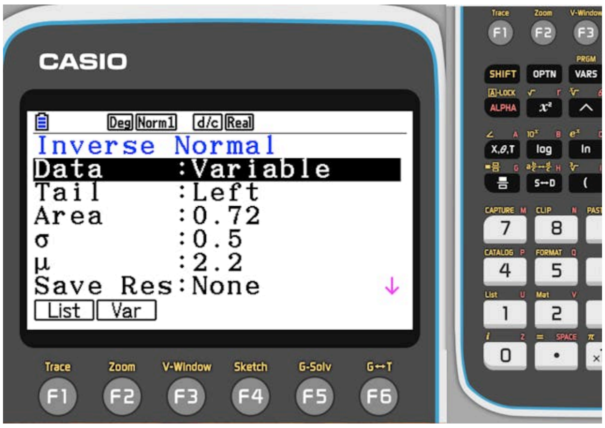
Student
view

Overview
(/study/ap
122-
cid-
754029/

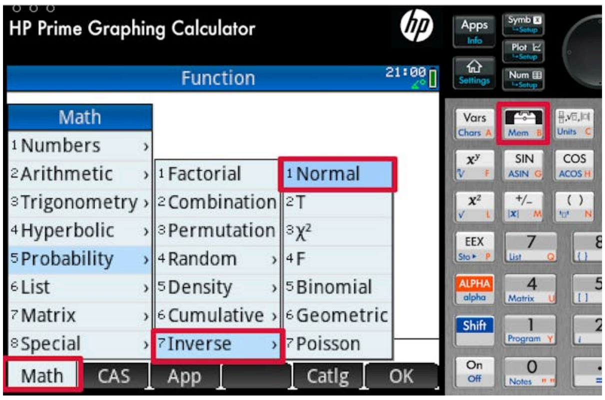
Step	Explanation
Press F1 to choose to work with normal distributions.	
Press F3 to choose the inverse normal (InvN) option.	

Student
view

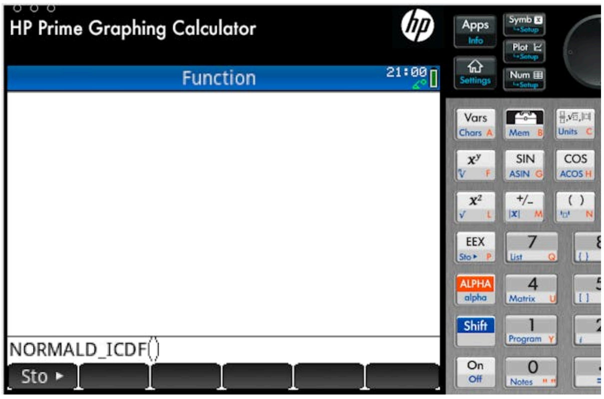
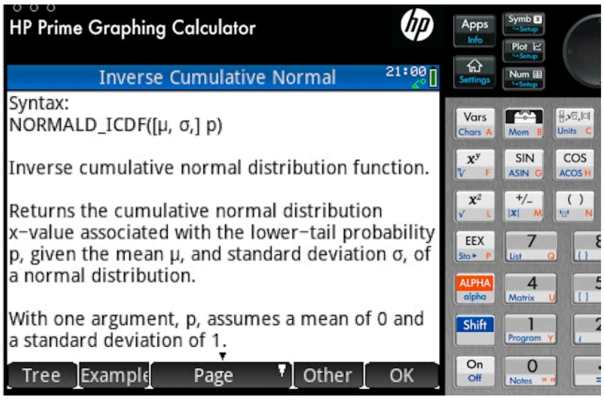
Overview
(/study/ap
122-
cid-
754029/

Step	Explanation
Enter the mean ($\mu = 2.2$) and the standard deviation ($\sigma = 0.5$) of the distribution. You also need to tell the value of the probability in the row that asks for area. The tail option is needed as left when a probability of the form $\mathrm{P}(X < b)$ is given. This calculator can also work out the bound if the probability is given in the form $\mathrm{P}(X > b)$. In this case the tail needs to be set to right.	
The calculator gives you the value of the bound, b.	

Student view

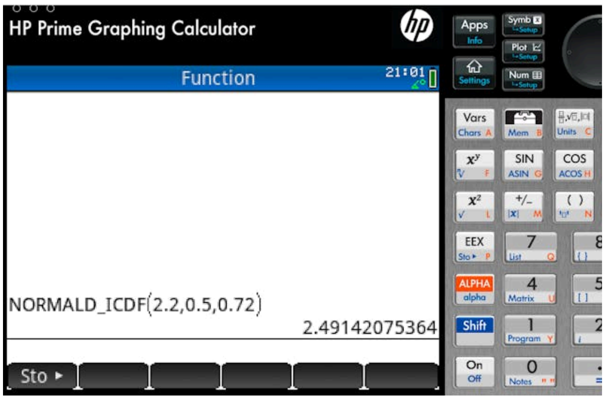
Step	Explanation
These instructions will show you how to find the value of b such that $\mathrm{P}(X < b) = 0.72$,where X is a normal random variable with mean $\mu = 2.2$ and standard deviation $\sigma = 0.5$. Enter the home screen of any application.	
Open the toolbox and find the inverse normal option through the menu system.	

Overview
(/study/ap
122-
cid-
754029/

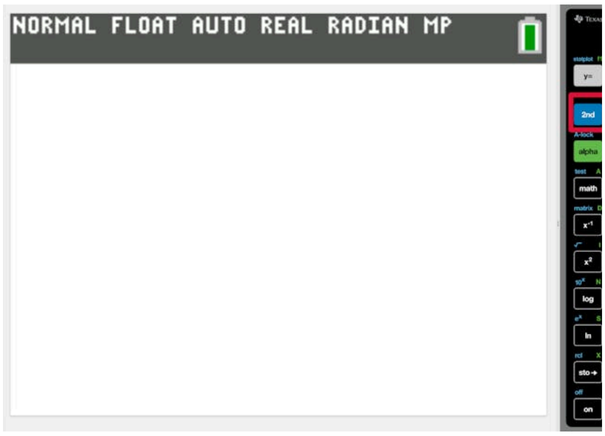
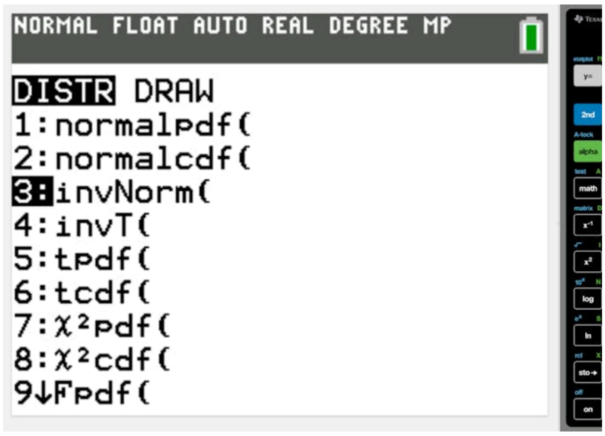
Step	Explanation
You need to tell the calculator the parameters (mean, standard deviation ...). If you do not remember what the calculator expects, open the help screen.	
This is the first page of the help screen of the inverse cumulative normal distribution function.	

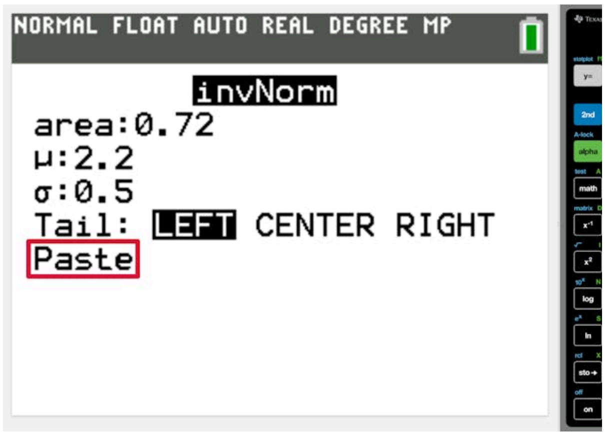
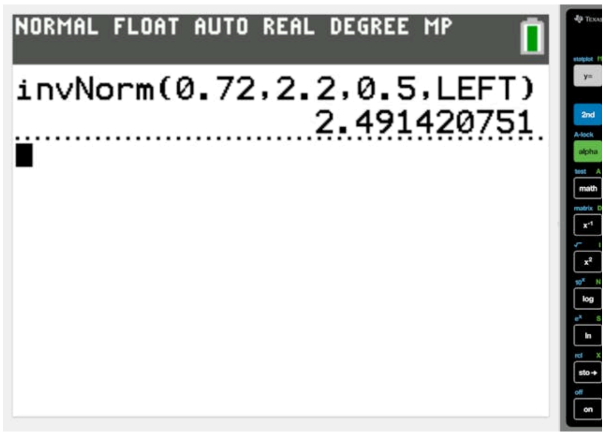
Student
view

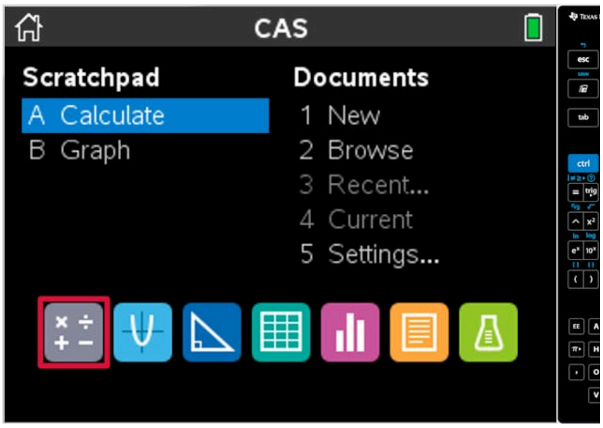
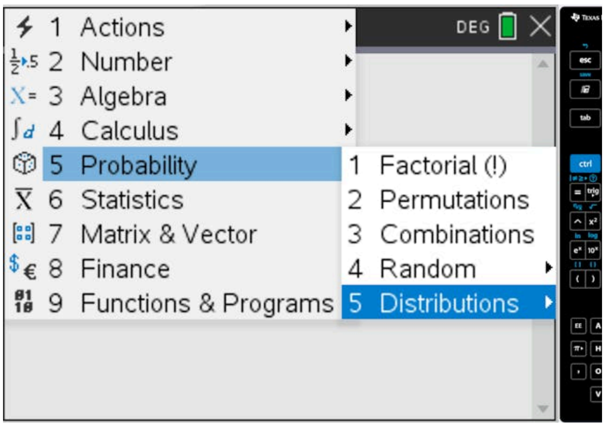
Overview
(/study/ap
122-
cid-
754029/

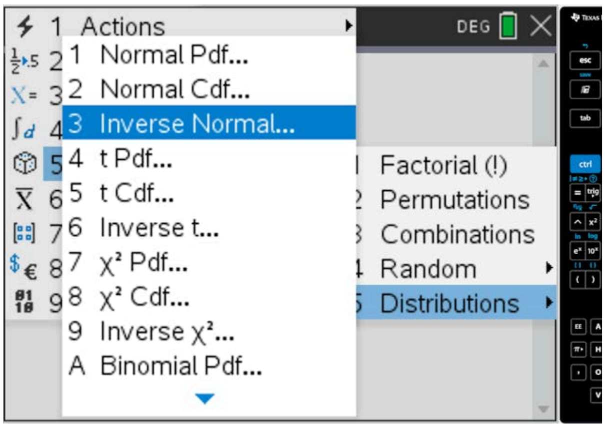
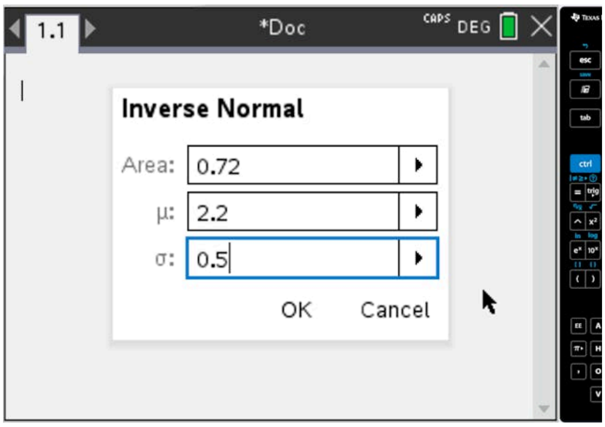
Step	Explanation
Enter the mean ($\mu = 2.2$) and the standard deviation ($\sigma = 0.5$) of the distribution. You also need to tell the value of the probability ($p = 0.72$). The calculator gives you the value of the bound, b.	

Student
view

Step	Explanation
These instructions will show you how to find the value of b such that $\mathrm{P}(X < b) = 0.72$,where X is a normal random variable with mean $\mu = 2.2$ and standard deviation $\sigma = 0.5$. Choose the distribution menu.	
Choose the inverse normal (InvNorm) option.	

Step	Explanation
Enter the mean ($\mu = 2.2$) and the standard deviation ($\sigma = 0.5$) of the distribution. You also need to tell the value of the probability in the row that asks for area. The tail option is needed as left when a probability of the form $\mathrm{P}(X < b)$ is given. This calculator can also work out the bound if the probability is given in the form $\mathrm{P}(X > b)$. In this case the tail needs to be set to right. Once done entering the parameters, move down to "paste" and press enter.	
The calculator gives you the value of the bound, b.	

Step	Explanation
These instructions will show you how to find the value of b such that $\mathrm{P}(X < b) = 0.72$,where X is a normal random variable with mean $\mu = 2.2$ and standard deviation $\sigma = 0.5$. Open a calculator page.	
Open the menu system and choose distributions ...	

Step	Explanation
... and select the option to work with the inverse normal option.	
Enter the mean ($\mu = 2.2$) and the standard deviation ($\sigma = 0.5$) of the distribution. You also need to tell the value of the probability in the row that asks for area.	

Step	Explanation
The calculator gives you the value of the bound, b.	

Example 2

★★☆

Let $X \sim N(2.2, 0.5^2)$. Find the value of b such that $\mathrm{P(X > b) = 0.43}$.

Since we are looking for a value of X and know the mean and standard deviation, we can use Inverse Normal to find b, where 43% of the distribution is greater than the value b.

Recall that the inverse normal function requires the area less than the value we are trying to find, so we first need to find the complement of the given probability. If $\mathrm{P(X > b) = 0.43}$, then $\mathrm{P(X < b) = 1 - 0.43 = 0.57}$.

Overview (/study/ap 122- cid- 754029/

Now we can use Inverse Normal on the calculator to find $b = 2.29$ accurate to 3 significant figures.

Example 3

★★★

Let $X \sim N(2.2, 0.5^2)$. If you know that $\mathrm{P}(1.8 < \mathrm{X} < \mathrm{b}) = 0.6$, find the value of b.

In order to use the inverse normal function to find b in this case, we first need to find $\mathrm{P}(\mathrm{X} < \mathrm{b})$. To do this, consider the fact that $\mathrm{P}(\mathrm{X} < \mathrm{b}) = \mathrm{P}(\mathrm{X} < 1.8) + \mathrm{P}(1.8 < \mathrm{X} < \mathrm{b})$.

You can use the Normal CDF to find $\mathrm{P}(\mathrm{X} < 1.8) = 0.211855$.

Now that you know $\mathrm{P}(\mathrm{X} < \mathrm{b}) = 0.211855 + 0.6 = 0.811855$, you can use Inverse Normal to find $b = 2.64$.

Important

Do not round values within the process of solving a problem. Only round your final answer.

Exam tip

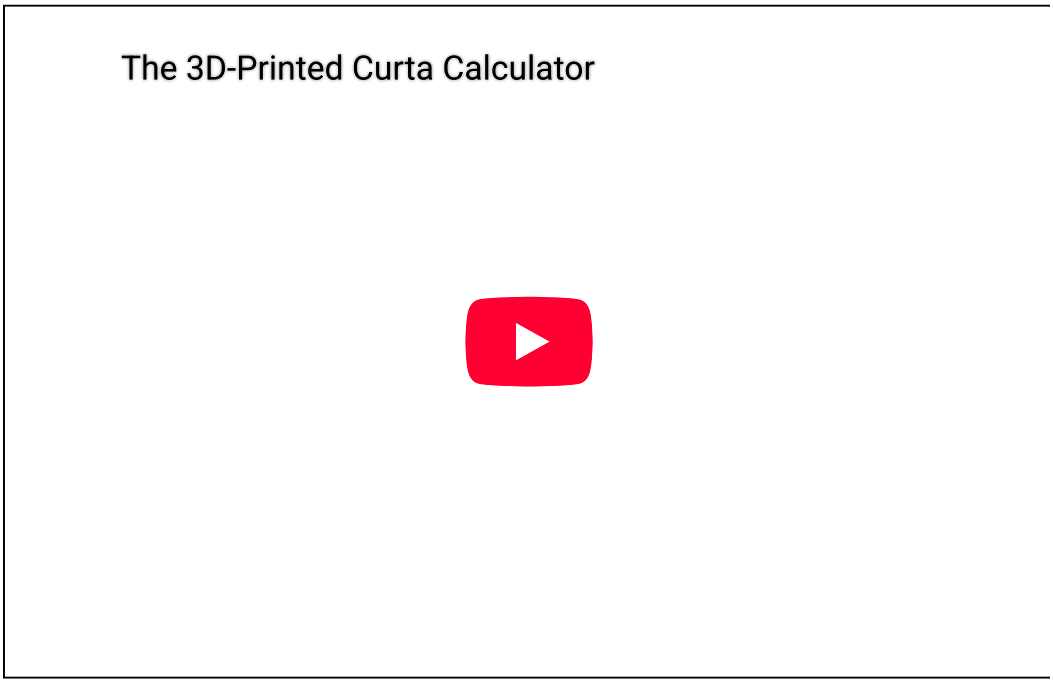
To answer the questions about normal distributions on Paper 2 of the exam, you will only need to know how to access and use Normal CDF and Inverse Normal. If you are curious about how to graph the normal curve, you can explore using the Normal PDF application on your calculator.

Student view

Theory of Knowledge

This section discusses how calculators can be used to calculate probabilities. From the TOK perspective , what makes a calculator, especially a GDC, a calculator and not a computer?

Consider, for example, the following video about the Curta 'Computer' — or is it a calculator?



How does the mechanical Curta 'Computer' work?

After watching this video, push beyond the mathematics and science connections to consider fully the following related TOK questions.

Is an abacus a computer? Is your calculator a computer? Your phone? Your watch? In other words, what is (not) a computer? Is this device, the Curta 'Computer', a computer? What computers might you call tools? Why is an encoding machine not a computer?

Press this inquiry a little further. When is your computer not a computer? How do the names of other devices and objects determine how we view and analyse their use? In other words, when is a chair not a chair (if you see some college dorm rooms, for example, you will see how a chair has become a working closet)?

4 section questions

Overview
(/study/ap
122-
cid-
754029/

4. Probability and statistics / 4.9 The normal distribution and curve

Checklist

Section Student... (0/0) Feedback Print (/study/app/m/sid-122-cid-754029/book/checklist-id-26268/print/) Assign

What you should know

By the end of this subtopic you should be able to:

- understand the notation $X \sim N(\mu, \sigma^2)$
- sketch a normal curve given the mean and standard deviation
- use symmetry and complements with given probabilities to find probabilities of other ranges
- use the 0.68, 0.95 and 0.997 estimates to find probabilities involving values 1, 2 and 3 standard deviations from the mean
- use the Normal CDF application on the calculator to find the probability of a range of values given a mean and standard deviation
- use the Inverse Normal application on the calculator to find a value (the quantile) that is above a certain proportion of the data in a distribution.

4. Probability and statistics / 4.9 The normal distribution and curve

Investigation

Section Student... (0/0) Feedback Print (/study/app/m/sid-122-cid-754029/book/investigation-id-26269/print/) Assign

Student view

Overview (/study/ap 122- cid- 754029/

Nearly everything that grows or develops naturally has features whose measurements form normal distributions. Investigate this trend by selecting something in your everyday life that you can measure. Maybe you are interested in food and could measure the size or weight of all the grapes on a bunch. Or maybe you participate in track sports and want to measure how long it takes the members of your team to run a particular distance. You can measure anything, but make sure you can collect a lot of data (at least 50 measurements if possible).

Divide your data into classes and create a histogram. Does the data look like a normal distribution? You may have to adjust the class ranges a few times to be able to see it clearly.

Once you have determined whether your data forms a normal distribution, calculate the mean and standard deviation. Use them with your calculator to find several probabilities using different ranges of measurements and compare those with the percentage of your data that fall into each range.

How close were the probabilities you found with the calculator to the percentages you found from your data? Do you think you were correct to assume there was a normal distribution? Why or why not? Do you think collecting more data would change your results?

Student view