Question

The Normal Distribution

A-level Maths Topic Summaries - Statistics

Subject & topics: Maths | Statistics | Hypothesis Tests **Stage & difficulty:** A Level P3

Fill in the blanks to complete these notes about the normal distribution.

Part A

Normal distributions and their properties

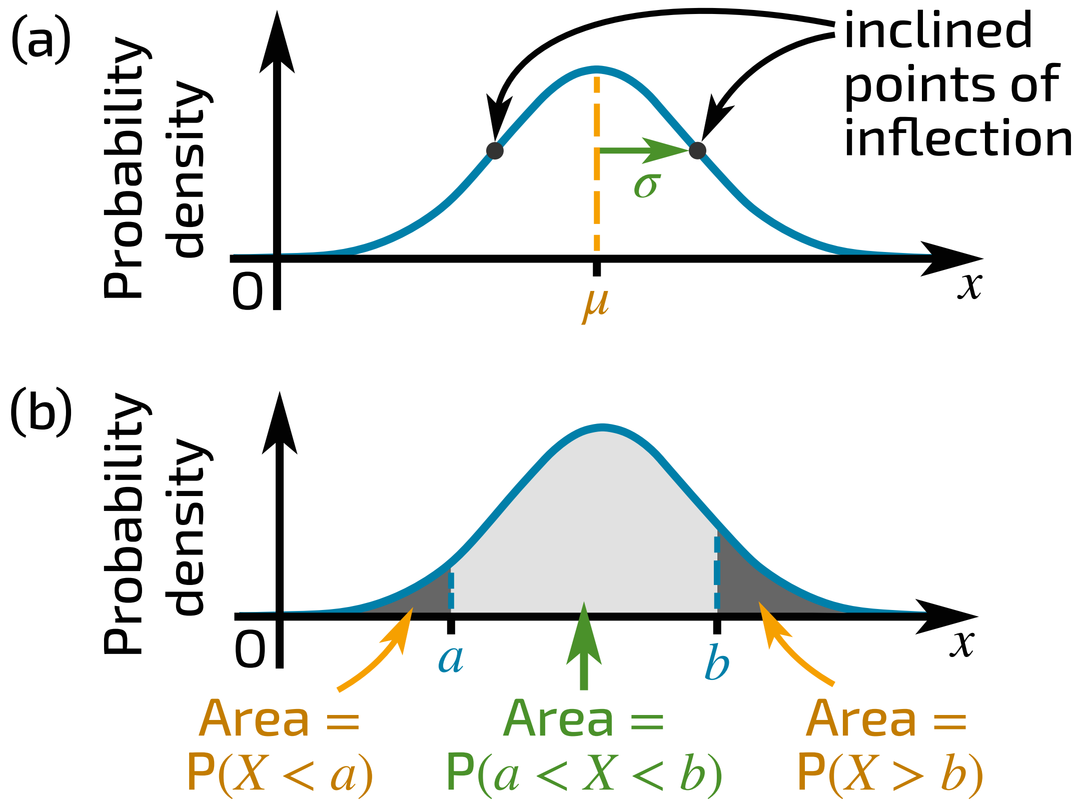

Figure 1: The normal distribution: (a) key features (b) how to calculate probabilities.

Normal distributions are a type of probability distribution for [ ] random variables. Normal distributions have a [ ] curve shape. They are symmetric about the mean [ ]. The standard deviation [ ] is a measure of the spread of the distribution. It is the distance from the mean to the (non-stationary) points of [ ].

We write the distribution for a random variable X as $X \sim \mathrm{N}(\mu, \sigma^2)$. Note that it is the [ ] σ^2 that is used in this description, not the standard deviation.

We calculate probabilities by finding the [ ] of regions under the curve using a calculator.

- [ ] is the area under the curve to the left of $x = a$.
- [ ] is the area under the curve between $x = a$ and $x = b$.
- [ ] is the area under the curve to the right of $x = b$.

The total area under the curve is [ ].

Items:

1 | areas | bell | continuous | inflection | μ | P($X < a$) | P($a < X < b$) | P($X > b$) | σ | variance

Part B

The standard normal distribution

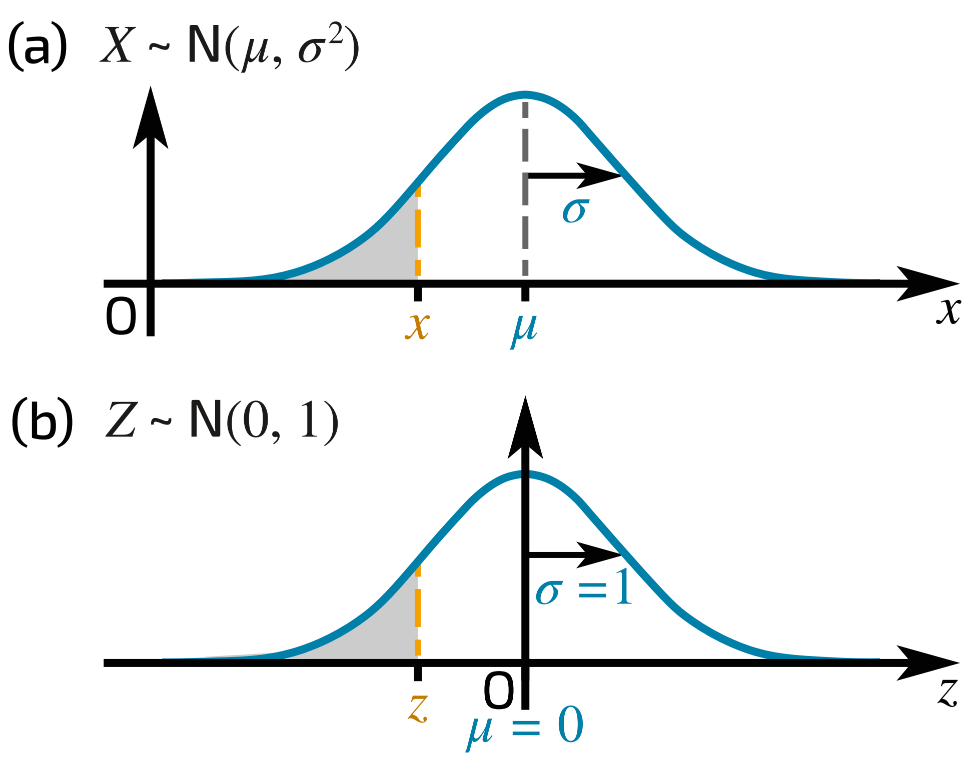

Figure 2: (a) a normal distribution (b) the standard normal distribution.

The **standard normal distribution** is a normal distribution with mean [ ] and standard deviation [ ]. We write this distribution as $Z \sim \mathrm{N}(0, 1)$.

Any normal distribution can be transformed into the standard normal distribution by performing two transformations: a [ ] parallel to the x-axis by [ ] followed by a stretch parallel to the x-axis by a scale factor [ ].

Every value x for a normal distribution has an equivalent value z on the standard normal distribution to which it is related by the formula

$$\boxed{} = \frac{\boxed{} - \mu}{\sigma}$$

These values of x and z are equivalent in the sense that the area under the curve to the left of x is the same as the area under the curve to the left of z, and hence $\mathrm{P}(X < x) = \mathrm{P}(Z < z)$.

Items:

translation $\frac{1}{\sigma}$ 0 1 $-\mu$ x z

Part C

The inverse normal distribution

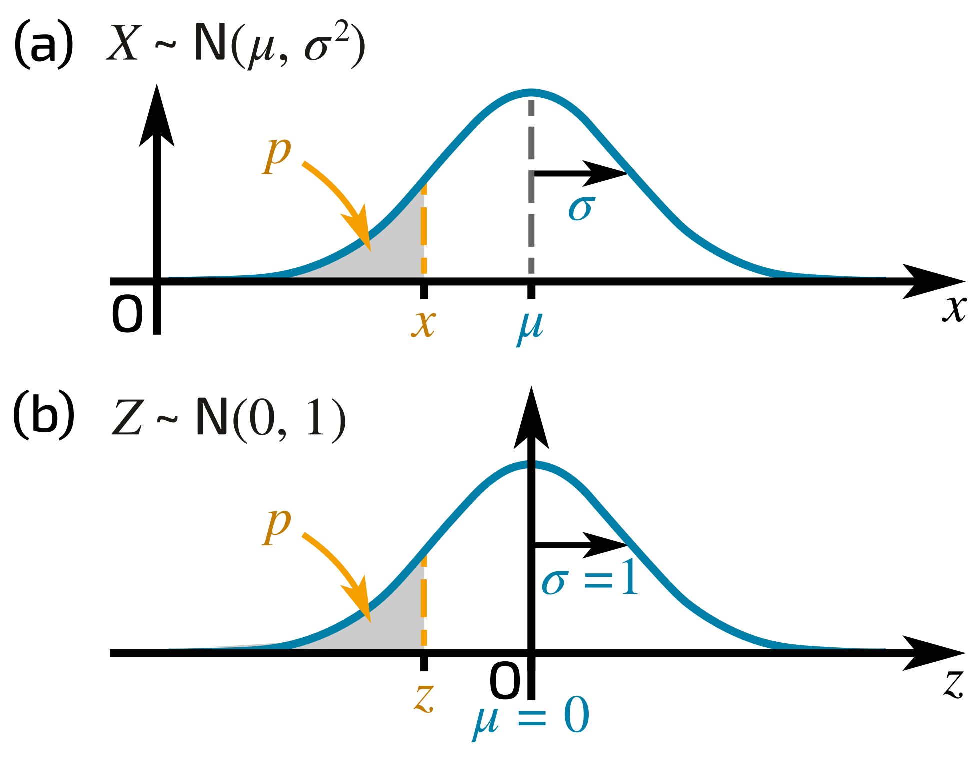

Figure 3: The inverse normal distribution can produce a value of (a) x or (b) z.

If a random variable X is described by a [ ] distribution, and we know the mean and standard deviation, we can calculate the probability that X is less than x. We put in values of μ, σ and x, and get out the probability $p = \mathrm{P}(X < x)$.

The [ ] distribution is used to go the other way:

- If we know both μ and σ, we put in the [ ], μ and σ, and get out the value of x for which [ ] $= p$.
- If we do not know both μ and σ, we put in the probability p and set $\mu =$ [ ] and $\sigma =$ [ ]. We get out the value of the standard normal distribution for which [ ] $= p$. If we have extra information, we can then use [ ] to find the unknown value(s) of μ and σ.

You may see the inverse normal distribution written as [ ].

Items:

0 | 1 | normal | inverse normal | $\mathrm{P}(X < x)$ | $\mathrm{P}(Z < z)$ | probability p | $\Phi^{-1}(p)$ | $z = \dfrac{x - \mu}{\sigma}$

Created for isaacscience.org by Jonathan Waugh

Question

Normal Distribution: Distribution of Sample Means

A-level Maths Topic Summaries - Statistics

Subject & topics: Maths | Statistics | Hypothesis Tests **Stage & difficulty:** A Level P3

Fill in the blanks to complete these notes about the distribution of the mean values of samples taken from a normally distributed population.

Let us suppose that we are trying to find the mean value of a particular quantity within a population. If we take samples, and calculate a mean value for each sample, the mean values we end up with will not all be identical. They will form a sampling distribution.

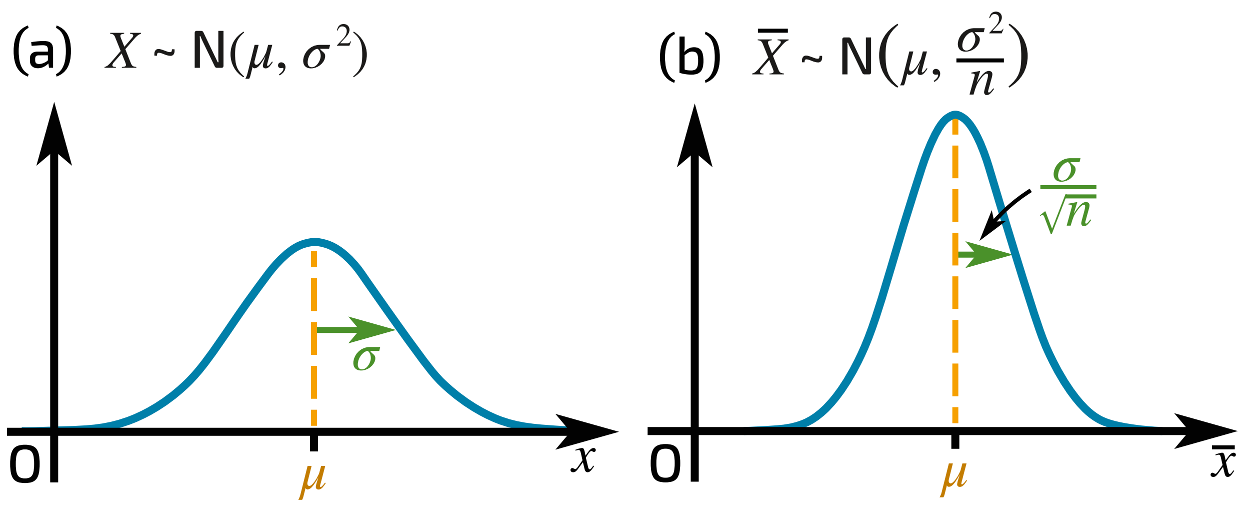

Figure 1: Sampling a normal distribution. (a) the normal distribution for the population, (b) the distribution of sample mean values.

If a quantity X is normally distributed in the population, and we take random samples, the mean values of the samples will also obey a ____ distribution.

The **distribution of sample means** for samples of size ____ is

$$\overline{X}_n \sim \mathrm{N}\left(\mu, \square\right)$$

The mean value of the distribution of sample means is ____. It is ____ the mean value of the normal distribution for the population.

The standard deviation of the distribution of sample means is ____. For $n > 1$, this value is ____ the standard deviation of the normal distribution for the population. The ____ the value of n, the smaller the standard deviation for the distribution of sample means. This reflects the fact that taking a larger sample will give a more reliable estimate for the mean value of the population.

Items:

μ n $\frac{\sigma}{\sqrt{n}}$ $\frac{\sigma^2}{n}$ normal smaller than the same as larger

Created for isaacscience.org by Jonathan Waugh

Question deck:

STEM SMART Single Maths 41 - The Normal Distribution

Question

Probability 5.5

Subject & topics: Maths | Statistics | Probability **Stage & difficulty:** A Level P3, Further A P1

Assuming they all follow a normal distribution find the following probabilities.

Part A

$X \sim \mathrm{N}(10, 4)$

If $X \sim \mathrm{N}(10, 4)$ find the following probabilities, giving your answers to 3 sf:

$\mathrm{P}(X \leq 9) =$

$\mathrm{P}(X \geq 15) =$

Part B

$X \sim \mathrm{N}(15, 0.3)$

If $X \sim \mathrm{N}(15, 0.3)$ find the following probabilities, giving your answers to 3 sf:

$\mathrm{P}(X < 14) =$

$\mathrm{P}(14 < X < 16) =$

Part C

S normally distributed

Assuming that the variable S follows a normal distribution with mean 40 and variance 16 find the following probabilities, giving your answers to 3 sf:

$\mathrm{P}(S > 44) =$

$\mathrm{P}(|S - 40| < 5) =$

Part D

Points of inflection

The points of inflection of a normal distribution curve occur at -2 and 6. Deduce the mean and variance of the distribution.

The mean is and the variance is .

Created by for isaacphysics.org by Julia Riley

Question deck:

STEM SMART Single Maths 41 - The Normal Distribution

isaac You work it out.

Question

Probability 5.6

Subject & topics: Maths | Statistics | Probability **Stage & difficulty:** A Level P3, Further A P1

Assuming they all follow a normal distribution find the following.

Part A

$X \sim \mathrm{N}(18, 9)$

Consider the normal distribution $X \sim \mathrm{N}(18, 9)$.

Find the Z-score of the value $x = 12$.

Find the Z-score of the value $x = 21$.

Part B

$X \sim \mathrm{N}(90, 100)$

Consider the normal distribution $X \sim \mathrm{N}(90, 100)$.

Find the value of X that is 3 standard deviations above the mean.

If $x = 0$, find how many standard deviations it is from the mean. (If $x = 0$ is above the mean, enter a positive number; if $x = 0$ is below the mean, enter a negative number.)

Part C

$X \sim \mathrm{N}(\mu, 16)$

Consider the normal distribution $X \sim \mathrm{N}(\mu, 16)$. When $x = 17$, it is 2.5 standard deviations below the mean μ; find μ.

Part D

$X \sim \mathrm{N}(100, \sigma^2)$

Consider the normal distribution $X \sim \mathrm{N}(100, \sigma^2)$. When $x = 82$, it is 3 standard deviations below the mean; find σ^2.

Created for isaacphysics.org by Julia Riley

Question deck:

STEM SMART Single Maths 41 - The Normal Distribution

Question

Probability 5.8

Subject & topics: Maths | Statistics | Probability **Stage & difficulty:** A Level P3, Further A P1

The number of gamma-rays emitted by a radioactive sample in a given time period follows a normal distribution and has a mean of 500.8 and standard deviation of 7.1. Find the following.

Part A

30% of the measurements $< p$

30% of the measurements have a value less than p; find the value of p. Give your answer to 4 sf.

Part B

15% of the measurements $> q$

15% of the measurements have values greater than q; find the value of q. Give your answer to 4 sf.

Part C

Six successive measurements

Six measurements are made successively.

Find the probability that none of the 6 measurements are greater than 515. Give your answer to 3 sf.

Find the probability that exactly 4 of the 6 are less than 510. Give your answer to 3 sf.

Created for isaacphysics.org by Julia Riley

Question deck:

STEM SMART Single Maths 41 - The Normal Distribution

Question

Probability 5.9

Subject & topics: Maths | Statistics | Probability **Stage & difficulty:** A Level P3, Further A P1

Answer the following questions. You may assume in all cases that the values of the quantities follow a normal distribution.

Part A

A batch of batteries

A batch of batteries has a mean voltage of $1.50\,\mathrm{V}$ and standard deviation σ. It is found that 20% have a voltage less than $1.47\,\mathrm{V}$. Deduce the value of σ, giving your answer to 3 sf.

Part B

Falling times

A number of students measure the time it takes for an object to fall a certain distance. The distribution of times has a standard deviation of $0.095\,\mathrm{s}$. It is found that 10% of the measurements exceed $5.80\,\mathrm{s}$. Find the mean time taken, giving your answer to 3 sf.

Part C

A batch of lenses

An experimenter has a batch of 150 lenses. They find that 10 of them have focal lengths less than $14.7\,\mathrm{cm}$ and 6 have focal lengths greater than $15.6\,\mathrm{cm}$.

Find the mean and the variance of the focal lengths, giving your answers to 3 sf.

The mean is ____ and the variance is ____.

Created for isaacphysics.org by Julia Riley

Question deck:

STEM SMART Single Maths 41 - The Normal Distribution

isaac You work it out.

Question

Data Analysis 5.9

Subject & topics: Maths | Statistics | Data Analysis **Stage & difficulty:** A Level P3

To indicate that an observation X comes from a normal distribution with mean μ and standard deviation σ the following notation is used

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

If we look at a sample of n independent observations of X and calculate the mean $\overline{X}_n$, then $\overline{X}_n$ also comes from a normal distribution where

$$\overline{X}_n \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

i.e. the standard deviation for the distribution of the means is $\frac{\sigma}{\sqrt{n}}$.

Using this notation answer the following questions.

Part A
$X \sim \mathrm{N}(0,1)$

An observation X comes from a normal distribution $X \sim \mathrm{N}(0,1)$.

(i) The distribution of the sample means is given by $\overline{X}_5 \sim \mathrm{N}\left(0, a\right)$. What is the value of a?

(ii) Find the probability that $X > 0.8$. Give your answer to 3 sf.

(iii) Find the probability that $\overline{X}_5 > 0.8$. Give your answer to 2 sf.

Part B

$X \sim \mathrm{N}(10, 4)$

An observation X comes from a normal distribution $X \sim \mathrm{N}(10, 4)$.

(i) The distribution of the sample means is given by $\overline{X}_{10} \sim \mathrm{N}\left(10, b\right)$. What is the value of b?

(ii) Find the probability that $X < 11$. Give your answer to 3 sf.

(iii) Find the probability that $\overline{X}_{10} < 11$. Give your answer to 3 sf.

Part C

$X \sim \mathrm{N}(8, 100)$

An observation X comes from a normal distribution $X \sim \mathrm{N}(8, 100)$.

(i) The distribution of the sample means is given by $\overline{X}_{4} \sim \mathrm{N}\left(8, c\right)$. What is the value of c?

(ii) Find the probability that $-5 < \overline{X}_{4} < 13$. Give your answer to 3 sf.

Part D

$X \sim \mathrm{N}(-10, 3)$

An observation X comes from a normal distribution $X \sim \mathrm{N}(-10, 3)$.

(i) The distribution of the sample means is given by $\overline{X}_6 \sim \mathrm{N}\left(-10, d\right)$. What is the value of d?

(ii) Find the probability that $-10 < \overline{X}_6 < -8.5$. Give your answer to 3 sf.

Created for isaacphysics.org by Julia Riley

Question deck:

STEM SMART Single Maths 41 - The Normal Distribution

isaac You work it out.

Question

Hypothesis Testing: Normal Distribution 2

Subject & topics: Maths | Statistics | Hypothesis Tests **Stage & difficulty:** A Level P3

Using a single black ink cartridge, a specific type of large printer is known to print a mean of $10\,750$ pages of black and white text, with a variance of $90\,000\,\text{pages}^2$.

The company that makes the printer changes some of its software in an attempt to make the use of ink by their printers more efficient. They perform 80 tests on printers using the updated software and calculate a mean of $10\,834$ pages per ink cartridge.

Test at the 2% significance level whether the new software has improved the ink efficiency.

Part A

Assumptions

Which of the following options do you need to assume in order to perform a hypothesis test? Select all that apply.

- [ ] That the variance in the number of pages printed in the sample of 80 tests is much smaller than the variance when using the old software, because 80 is much smaller than the number of printers using the old software.
- [ ] That the variance of the sample obeys a Poisson distribution.
- [ ] That the number of pages printed using one ink tank has a normal distribution.
- [ ] That the variance in the number of pages printed using the new software is the same as the variance when using the old software.

Part B

Null and alternative hypotheses

Fill in the blanks to state the distribution of the number of pages, and the null and alternative hypotheses.

Let X be the number of pages printed using one cartridge with the new software. Then $X \sim \mathrm{N}(\mu, \_\_\_\_)$, where μ is the mean.

The null hypothesis is that the new software does not improve ink efficiency, and the mean number of pages per cartridge is the same as before. The alternative hypothesis is that the new software improves ink efficiency.

$$\mathrm{H}_0 : \mu = 10\,750 \qquad \mathrm{H}_1 : \_\_\_\_$$

Items:

33.541 300 1125 $90\,000$ $\mu < 10\,750$ $\mu \neq 10\,750$ $\mu > 10\,750$

Part C

Carrying out the test

Fill in the blanks to complete the hypothesis test.

Let $\overline{X}$ be the mean number of pages printed using one cartridge with the new software. Assuming that the null hypothesis is true, then $\overline{X} \sim \mathrm{N}(10\,750, \frac{\_\_\_\_}{80})$. Therefore, the p-value for a sample mean of $10\,834$ pages is found to be

$$p = \mathrm{P}(\overline{X} \geq 10\,834) = \_\_\_\_$$

For a one-tailed test at the 2% significance level, p-values of less than 0.02 are in the critical (rejection) region. The calculated p value for the sample is ____ 0.02.

Therefore, ____ the null hypothesis. There ____ significant evidence that the new software improves ink efficiency.

Items:

is not 0.3897 is 300 reject less than 0.006133 greater than do not reject $90\,000$

Created for isaacphysics.org by Jonathan Waugh

Question deck:

STEM SMART Single Maths 41 - The Normal Distribution

Question

Hypothesis Testing: Normal Distribution 1

Subject & topics: Maths | Statistics | Hypothesis Tests **Stage & difficulty:** A Level P3

A species of chicken is known to produce eggs which have a mass that follows a normal distribution with a mean mass of $55.0\,\mathrm{g}$ and a variance of $5.00\,\mathrm{g}^2$.

A chicken-keeper has 1000 birds of this species. Each bird lays one egg per day, and on one particular day the chicken-keeper collects the eggs from 125 birds and finds a mean mass of $54.0\,\mathrm{g}$. The chicken-keeper wants to know whether their birds produce eggs with a below average mass.

Test, at the 5% significance level, whether this is the case.

Part A
Assumptions

In order to carry out a hypothesis test, which of the following do you need to assume?

- The variance of the masses of the eggs produced by the chicken-keeper's birds is the same as that of the whole chicken population.
- The variance of the masses of the eggs produced by the chicken-keeper's birds is 1000 times smaller than that of the population as the chicken keeper has 1000 birds.
- The variance of the masses of the eggs produced by the chicken-keeper's birds is 1000 times that of the population as the chicken keeper has 1000 birds.

Part B

The null and alternative hypotheses

Drag and drop into the spaces below to define the variables and state the null and alternative hypotheses for this test.

Let M be the mass of an egg produced by the chicken-keeper's birds, and let μ be the mean mass of the eggs produced by the chicken-keeper's birds. Then $M \sim$ [ ] $(\mu,$ [ ] $)$.

H_0: [ ] [ ] [ ]

H_1: [ ] [ ] [ ]

Items:

B $>$ $55.0\,\mathrm{g}$ $<$ 5.00 Po N $54.0\,\mathrm{g}$ $=$ μ

Part C

The distribution of the sample means

Fill in the blanks below to complete the description of the distribution of the sample means.

The masses of the eggs produced by the chicken-keeper's birds have a [ ] distribution. Therefore, if samples of these eggs are taken, the mean values of these samples, $\overline{M}$, have a [ ] distribution.

The masses of the eggs produced by the chicken-keeper's birds have a variance of [ ] g^2. For samples containing 125 eggs, the variance of the sample is therefore [ ] g^2.

Hence, under the assumption that the null hypothesis is true, the sample distribution is $\overline{M} \sim$ [ ] $($ [ ] $,$ [ ] $)$.

Items:

0.005 B binomial 5.00 N 55.0 normal 0.04

Part D

Carrying out the test

Choose three options and put them into order to complete the hypothesis test.

Available items

Therefore, as the value of the sample mean ($54.0\,\mathrm{g}$) lies in the critical region, we reject the null hypothesis. There is evidence that the chicken-keeper's birds produce eggs with a below average mass.

We need to carry out a one-tailed test at the 5% significance level. Using the inverse normal distribution, the z-value at the upper boundary of the critical (rejection) region is $z = \Phi^{-1}(0.95) = 1.645$.

$z = \dfrac{\overline{M} - \mu}{\sqrt{\dfrac{\sigma^2}{n}}}$. Hence, the boundary of the critical region is at $\overline{M} = 55.0 + 0.2 \times 1.645 = 55.33$. The critical region is therefore $\overline{M} < 55.33$.

Therefore, as the value of the sample mean ($54.0\,\mathrm{g}$) lies in the acceptance region, we do not reject the null hypothesis. There is no significant evidence that the chicken-keeper's birds produce eggs with a below average mass.

$z = \dfrac{\overline{M} - \mu}{\sqrt{\dfrac{\sigma^2}{n}}}$. Hence, the boundary of the critical region is at $\overline{M} = 55.0 + 0.2 \times -1.645 = 54.67$. The critical region is therefore $\overline{M} < 54.67$.

We need to carry out a one-tailed test at the 5% significance level. Using the inverse normal distribution, the z-value at the upper boundary of the critical (rejection) region is $z = \Phi^{-1}(0.05) = -1.645$.

Created for isaacphysics.org by Jonathan Waugh