

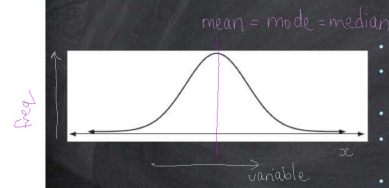
Statistical Applications

1: Normal Distribution

Chapter 5: p204 - 215

Normal Distributions

- Almost all continuous variation that occurs in nature and in man-made objects has a normal distribution. In Biology it is often called a "bell-shaped" curve.



Examples include:

- Weights of 16 year old males
 - Exact volume in soft drink cans
 - Lengths of adult sharks
 - Scores on tests taken by large populations
 - Life times of batteries
- In a normal distribution, the mean = mode = median.

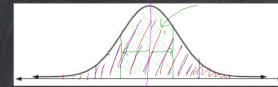
Properties of the Curve

- The distribution is symmetrical about the mean.
- The area under the curve represents the entire population and is equal to 100% (or 1).
- So 50% of the population has a value above the mean and 50% will be less than the mean.

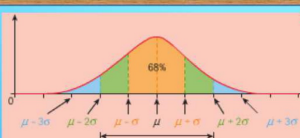


Standard Deviations

- The area bounded by 1 standard deviation on either side of the mean is about 68% of the population

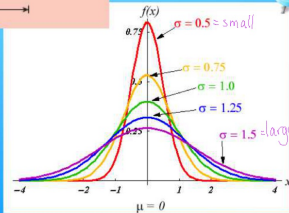


- The area bounded by 2 standard deviations on either side of the mean is about 95% of the population
- The area bounded by 3 standard deviations on either side of the mean is about 99% of the population
- In distribution questions, the mean has the symbol μ and the standard deviation has the symbol σ .



Virtually the entire variation within the population is inside $\pm 3\sigma$

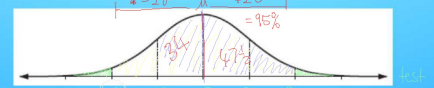
When σ is large, the distribution is flat and spread out, when σ is small, the distribution is tall and narrow.



Example to try:

- 500 year 10 students sat a mathematics examination. Their marks were normally distributed with a mean of 75 and a standard deviation of 8.

a) Assign scores to the markings on the horizontal axis on the curve below.



b) How many students would you expect scored marks:

- more than 83
- less than 59
- between 67 & 91?

less than 83 = 50% + 34% = 84%
 $100 - 84 = 16\%$
 $500 \times 16 = 80$ students

more than 59 = 47.5% + 34% = 81.5%
 $500 \times 81.5 = 407.5 = 408$ students

between 67 & 91 = 47.5% + 47.5% = 95%
 $500 \times 95 = 475$ students

Calculator Questions

Finding Area *Normal cdf*

- The area under the curve represents the percentage of the population between values of the variable. Questions asking for either the percentage of the population, the probability or the number of a sample within certain boundaries are solved by using **normal cdf** on the GDC.

2. The mass of apples are normally distributed with a mean of 100g and a standard deviation of 15g.

a) Find the percentage of apples that have a mass between 90g and 110g.

(This Q continues on the next slide)

In calculate scratchpad

- Press MENU, then 6:Statistics → 5:distributions → 2:Normal Cdf
- Enter details from question, press OK
- Answer is a decimal, x100 to change to %

Calculator Questions

Finding Area and Expected Number

2. The mass of apples are normally distributed with a mean of 100g and a standard deviation of 15g.

b) Find the probability that an apple is heavier than 125g.

The mass of 130 apples are measured.

c) Calculate the expected number of apples that have a mass less than 80g.

In calculate scratchpad

- Press MENU, then 6:Statistics → 5:distributions → 2:Normal Cdf
- Enter details from question, press OK
- Answer is a decimal, x100 to change to %

Example to try:

You should sketch a distribution and shade in the relevant information for every question (whether you are required to or not)

3. The length of time waiting to see a counselor at DAA is normally distributed with a mean of 5 minutes and standard deviation of 1.5 minutes.

a) Find the probability that you will have to wait more than 8 minutes to see your counselor.

b) What proportion of students wait less than 3 minutes?

c) What percentage of students wait between 4 - 7.5 minutes?

d) If 35 students try to see a counselor at lunch time, what is the expected number who wait more than 6 minutes?

Calculator Questions

Finding a Value given the Area *Inverse Normal*

- If you are given the percentage or probability and need to find the value along the x-axis that it corresponds to, use **Inverse Normal** on the GDC.

4. The volume in Oasis water bottles is normally distributed with a mean of 100 ml and a standard deviation of 2 ml.

a) 5% of the bottles are rejected for containing too little water. Find the minimum volume, to the nearest ml, that a bottle must contain if it is to be accepted.

In calculate scratchpad

- Press MENU, then 6:Statistics → 5:distributions → 3:Inverse Normal
- Enter details from question (area MUST be a decimal), press OK
- Answer is the value on the x-axis.

Calculator Questions

Finding a Value given the Area

- Inverse Normal on the GDC can **ONLY** find the value for an area starting on the left hand side. If the question gives an area on the right of the value, you MUST subtract from 1 (or 100%).

5. The mass of watermelons are normally distributed with a mean of 960g and a standard deviation of 125g.

a) It is known that 15% of watermelons have a mass greater than x g. Find the value of x.

b) 300 watermelons are weighed. Find the expected number of watermelons with a mass less than 800g.

always calculates from left

In calculate scratchpad

- Press MENU, then 6:Statistics → 5:distributions → 3:Inverse Normal
- Enter details from question (area MUST be a decimal), press OK
- Answer is the value on the x-axis.

Example to try:

You should sketch a distribution and shade in the relevant information for every question (whether you are required to or not)

6. The marks of 500 candidates in an exam are normally distributed with a mean of 55 marks and a standard deviation of 15 marks.

a) If 5% of the candidates obtain a distinction by scoring d marks or more, find the value of d.

b) If 10% of students fail by scoring f marks or less, find the value of f.

c) What is the expected number of candidates who score more than 75 on the exam?

Calculator Questions

Finding 2 Values given the Area

- If the area is symmetrical about the mean, we use a sketch and symmetry properties to find both values.

7. The volume in large Oasis water bottles is normally distributed with a mean of 1002 ml and a standard deviation of 3.5 ml.

a) 90% of the bottles contain between a ml and b ml, where a and b are symmetrical about the mean. What are the values of a and b to the nearest ml?

In calculator scratchpad

- Press MENU, then 6:Statistics \rightarrow 5:distributions \rightarrow 3:Inverse Normal
- Enter details from question (area MUST be a decimal), press OK
- Answer is the value on the x-axis.

$a = \text{INV normal area} = 0.05$
 $b = \text{INV normal area} = 0.95$

Example to try:

You should sketch a distribution and shade in the relevant information for every question (whether you are required to or not)

7. The marks of 500 candidates in an exam are normally distributed with a mean of 55 marks and a standard deviation of 15 marks.

Find the lower quartile mark and the upper quartile mark and hence the IQR for the exam (round answers to the nearest whole mark).

$IQR = 65 - 45 = 20$

LQ area = 0.25 LQ = 44.88 marks = 45 marks

UQ area = 0.75 UQ = 65.117 marks = 65 marks

Practice

Ex 5A: p207 (do any 2)
 Ex 5B: p210 (do 3 exam-style Q's)
 Ex 5C: p214 (do 3 exam-style Q's)