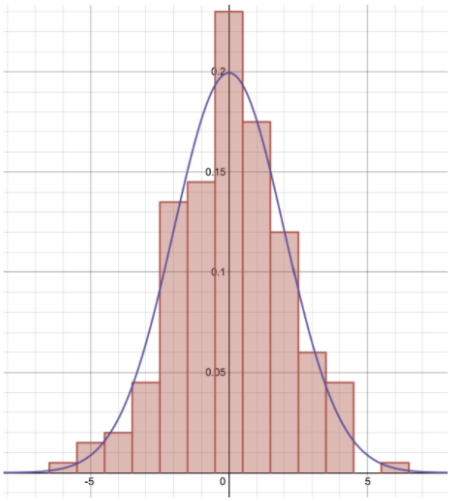
Statistics Guide

A guide to help with the exploration:
Maths IB Standard Level and Higher Level
Applications and Interpretations
Analysis and Approaches

(For first examination in 2021).



Author: Andrew Chambers

Table of Contents

Introduction

I've written this guide to supplement the main Exploration Guide I put together. You should consult the main guide for guidance on choosing topics, an explanation of the marking criteria, common student mistakes and technology advice. In this guide I look at a number of different statistical tests and how to effectively use technology. In many cases these are taught in textbooks simply using technology, whereas it is often desirable to demonstrate a greater understanding through non-calculator methods in your maths exploration. So, where possible I've included non-calculator techniques.

It's important to note that these methods are **not** intended to be exemplars - there are many different ways of explaining the following techniques and ideas, these are just my ideas! You should attempt to put your methods into your own words so that you can demonstrate a good personal understanding. The students who do best in their exploration consult from a variety of sources, collate the ideas and are therefore able to show a deep understanding.

If you do use this guide then it is essential that you correctly cite this source in your exploration - failure to cite sources correctly can lead to malpractice investigations by the IB, so make sure everything is done correctly.

The exploration is a great opportunity to apply your maths knowledge to an area of personal interest - so choose something you are passionate about, and enjoy it!

Pearson's Product Correlation investigation: height and arm span.

Correlation investigations are very common - but also have lots of things that can go wrong with them, so here is an example where I highlight common mistakes and show good practice.

Step 1: Personal engagement

Step 1. You need to work quite hard to justify a personal interest in order to get more than C1 on correlation topics. Two ways of showing personal engagement will be to do some reasonably time consuming data collection and creating a narrative as to why you are investigating this topic.

"Is there a correlation between the height and arm span of Y13 boys?" is a reasonable topic question which will be possible to complete, but it's quite depersonalised. Why do you care about this?

"Can understanding the relationship between height and arm span help me design better fitting suits for Y13 boys?" is immediately more engaging. Now there is a genuine purpose, and plenty of scope for reflection based on this topic question.

Step 2: Collecting data.

The two main problems here are not collecting enough data for the investigation to be meaningful, and not showing any awareness of sampling methods. I would recommend trying to collect 40-50 data points if you are collecting your own data. If you are using secondary data then 50-75 would be better.

You should show a clear explanation of the method used to collect data. For example, "I borrowed the height measuring machine from the school nurse and during a Y13 PE lesson asked my sample to line up with a straight back (no shoes). I measured in cm to 1 decimal place. etc."

Your choice of sampling methods are simple random sampling, convenience, systematic, quota and stratified. You need to show an awareness of which one you are using, a justification for why you are using it and a discussion about potential limitations.

For example, if you decide that the population you are interested in is limited to Y13 boys then you could conduct a simple random sample by assigning a number to every Y13 boy in the school and then using a random number generator to generate your sample.

Step 3: Data presentation. If you have a lot of data then you probably are best including the first part of a table in the main body and then the full table in the appendix. I'll work through some maths with 10 data points as an example.

Height (cm). Rounded to 4 sf.	Arm Span (cm). Rounded to 4 sf.
156.4	162.0
177.7	176.5
161.1	160.8
170.9	170.4
173.3	185.2
173.0	176.5
162.9	170.8
161.2	162.3
188.7	190.9
178.6	180.0

Here we have arm span on the y-axis therefore we are investigating if arm span is dependent on height. We can clearly see a positive linear correlation so it is relevant to do a Pearson's Product correlation calculation.