



**HONG KONG STATISTICAL SOCIETY**  
**2016 EXAMINATIONS – SOLUTIONS**  
**HIGHER CERTIFICATE – MODULE 1**

The Society has published these solutions to assist candidates preparing for the examinations in 2017.

The solutions are intended as learning aids and should not be seen as "model answers".

Users of the solutions should always be aware that in many cases there are valid alternative methods. Also, in the many cases where discussion is called for, there may be other valid points that could be made.

While every care has been taken with the preparation of these solutions, the Society will not be responsible for any errors or omissions.

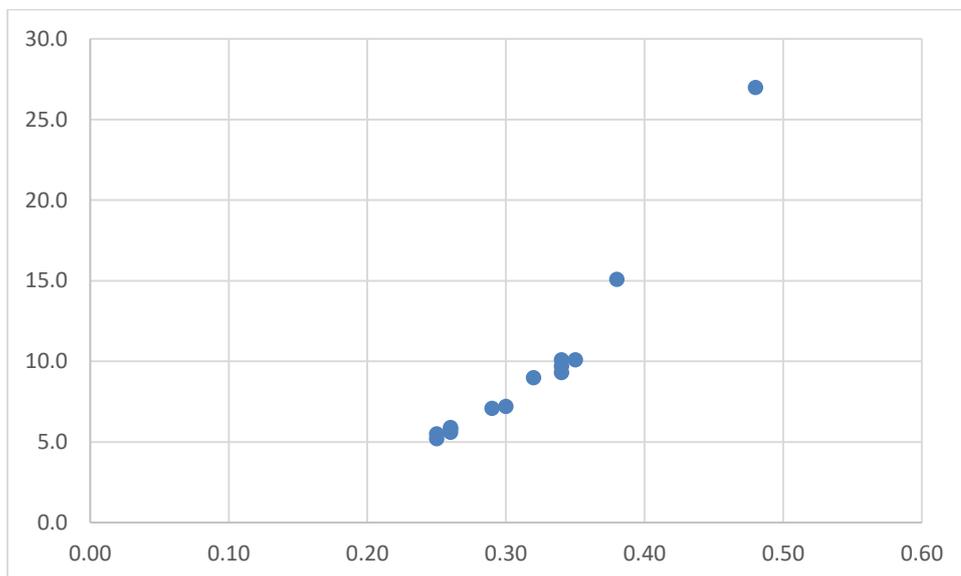
The Society will not enter into any correspondence in respect of these solutions.

*The HCl paper is, by its nature, open-ended: the questions expect candidates to determine their own approach to analysis and interpretation. For this reason, it is not possible to give definitive solutions. The mark scheme below therefore gives, for each question, one possible approach and the corresponding mark allocation. Other approaches will require the marker to exercise judgement.*

---

1(i) The report should address salient points, such as the following, looking at comparisons within and between groups and over time.

- **Within** the Anglo-Saxon group the US generally has a higher Gini index, and a higher 10/10 ratio.
- **Within** the Northern Europe group both measures suggest the countries are broadly the same.
- **Within** the Western Europe group Portugal and Italy have higher inequality as indicated by both measures.
- **Within** comparisons are not really meaningful in the ‘other countries’ group as it is heterogeneous, but Mexico has much greater inequality than Japan.
- **Between** groups, it is broadly the case that the Northern European countries have lower inequality than the rest. Mexico has far greater inequality than other countries.
- **Trends over time:** the broad pattern is for inequality to increase. (Simple line graphs – sparklines – for the 10/10 ratio would be useful here. Or identify the peak inequality year for each country and note that it tends to be recent.)  
Some slight evidence that Mexico goes against this trend. E.g. its peak 10/10 was in 1995.
- **Consistency.** There is a strong positive relationship between Gini and 10/10. E.g. for 2010 figures:



Not a linear relationship if Mexico and US are included, (but pmcc is 0.95).

So the two measures do send a consistent message.

*(Award up to 16 marks for distinct points and/or graphs addressing the specified issues. Award up to 4 marks for the quality of the written report. E.g. 2, 1 or 0 for the structure of the report, 2, 1 or 0 for clarity of wording.)*

2. (i) The purpose of a census is to enumerate, and collect data on, every member of a population (1)
- The census is a snapshot whereas various administrative records span continuous time and would have to be analysed, with some difficulty, to get a spot figure for a particular date. (1)
- It differs from a sample survey in that, by definition, a sample survey does not attempt to reach the whole population (1)
- A census will differ from records held by government departments in that it aims to be complete whereas government department records will not be (1)
- Also it addresses particular questions (e.g. language spoken, religion) which government records are unlikely to contain (1)
- And participation in the census is a legal requirement (1)
- (ii) A participation rate of 94% is high (1)
- and to that extent might be regarded as giving very good information when compared with other data (1)
- However, the nature of the missing 6% is an issue. These people are likely to be untypical of the 94% who participate (1)
- E.g. estimating the homeless rate from the 94% reached would be very inaccurate (1)
- (iii) Since 'Jedi Knight' is not in any real sense a religion (1)
- this indicates that people do not always take the census seriously (1)
- This may therefore cast doubt on the accuracy of other responses they give (1)
- It may also indicate a contempt for, or a distrust of, government and the collection of data by government agencies (1)
- While this example indicates that not all responses can be taken seriously, there may still be value in asking the question (1)
- For example, the 2011 census quantified the decline in Christianity and the rise in Islam: these are likely to be real phenomena (1)
- (iv) Conducting the census online in 2021 will present problems for excluded groups, those without internet access or with limited internet capability such as those in poverty and the old (1)
- So enumerators will still need to be used to reach these groups (1)
- It may also be more difficult to persuade people to complete the census by going online than it is to persuade them to fill in a printed form (1)
- (There may be additional concerns about security of information when it is supplied online.)
- Additional information held by government agencies is unlikely to be complete (1)
- Record matching in combining databases is a notorious problem (1)

3. (i) Equal number in the four groups, so the overall mean is  
 $(204 + 259 + 266 + 317) / 4 = 261.5$  for men (1)  
 $(178 + 235 + 266 + 304) / 4 = 245.75$  for women (1)  
The SDs are  $0.9 \times \sqrt{1308} = 32.5$  for men, and  $0.8 \times \sqrt{1540} = 31.4$  for women (1,1)  
The overall mean is  $(1308 \times 261.5 + 1540 \times 245.75) / 2848 = 253.0$  (2)

*(Award 1 for an attempt at weighted mean that results in the wrong answer)*

- (ii) Milk consumption is very low for both men and women in Q1 and Q2 (1)  
but it rises sharply in Q3 and again in Q4. (1)  
So those who eat most fresh vegetables consume much more milk than those who eat less fresh vegetables. (1)
- (iii) Rice consumption falls, for men and for women, as fresh vegetable consumption rises (1)  
It is the only food group to show this pattern. (1)
- (iv) Parallel bar chart, back-to-back bar charts or simple line graphs to compare men and women. (1)  
*(Charts must be appropriate (1), accurate (1), labelled (1) and well executed (1))* (4)
- (v) Divide men's figures by 1.2 (or multiply women's figures by 1.2) (1)

to obtain comparison figures:

Fresh veg	170	178
Fruit	26	28
Rice	306	315
Wheat flour	66	56
Meat	58	48
Fish	19	19

- (1)  
No very large differences in patterns of consumption (1)  
but men eat more meat and wheat flour (1)  
while women eat more rice and fresh vegetables (1)

4(a) **Strengths**

The information is clearly categorised by race. (1)

The categorisation by educational level is harder to extract but it is there. (1)

**Weaknesses**

There is no indication of the relative numbers in each race group. (1)

There is no indication of the numbers at each educational level. (1)

It is tempting, but wrong, to read the vertical dimension of each segment within a bar as conveying information about the size of the group. (1)

There is no indication of variability from the various group means. (1)

**Additional data**

Sizes of various groups, both between and within race groups. (1)

Some measure of variation within groups: e.g. SD or quartiles (1)

Parallel bars, grouped by race, with widths representing size of group. (1)

Each bar to have whiskers for SD or for quartiles. (1)

4(b) *Award a maximum of 5 marks for strengths and weaknesses such as those stated here or other sensible suggestions.*

**Strengths**

Chart is clear and (mostly) easy to read (1)

Comparisons can be made readily (1)

**Weaknesses**

Slight problem when symbols overlap (e.g.USA). (1)

This could be remedied by offsetting symbols. (1)

The lines joining symbols could be seen as suggesting a continuous distribution. (1)

The symbols alone would be better. (1)

No indication of numbers in groups, but not easy to see how to deal with this (1)

**Summary of data**

Across all OECD countries, with the possible exception of the USA, obesity rates are inversely related to educational attainment levels (1)

In many countries, there are about twice the proportion of obese people in the lowest educational group as in the highest (1)

Differences in obesity rates between educational levels seem to be least in Nordic countries (1)

Overall obesity rates are highest in (some) Anglo-Saxon countries, but no clear pattern as to where they are lowest (1)

Without information on numbers within the different groups it isn't possible to give an accurate average or overall figure for obesity, but about 20% seems a reasonable figure. (1)