Exercise 5: Analysing and Reporting Trial-Based Cost-Effectiveness Analysis Results

Acknowledgments

This exercise has been developed more extensively by Prof. Andrea Manca over the years, but it is based on earlier material produced by Prof. Andrew Briggs. We are grateful to Prof. Briggs for letting us use his exercise as a starting point. Errors and omissions are our own.

Overview

Please open the Excel file *Exercise 5 – Template 2020*. On loading this spreadsheet you will find that the file contains different worksheets, each relating to a particular element of the exercise.

Upon opening the spreadsheet you may get a security warning that Macros have been disabled. In order to complete the exercise you must enable this content.

Now, let's inspect the spreadsheet's content.

The first worksheet, <*Raw data* >, contains a small dataset from a hypothetical RCT-based CEA. Patient-level observations relating to two treatment groups are provided. Each observation contains a cost and effect for each trial participant.

The <*Descriptive analysis*> worksheet contains some basic descriptive statistics relating to the data included in the <*Raw data* > worksheet, which you are asked to calculate (more on this in subsequent sections). **Please note that the cells in this worksheet are all named, which means that formulae cannot be dragged across to estimate the statistic for the other group. Remember this when you are asked to calculate the descriptive statistics.**

The worksheet <*Confidence box*> is already set up to plot the five $(\Delta\bar{C}, \Delta\bar{E})$ pairs needed to obtain the box. Please note that this worksheet is for illustrative purposes only, and you will not be asked to work with it.

The worksheet <*Bootstrap*> contains two macro buttons ('Run bootstrap macro' and 'Run CEAC Macro'). The first generates 1,000 bootstrap replications of the cost and effect differences from the dataset. The second generates the data points needed to build the cost-effectiveness acceptability curve (CEAC).

The worksheet <*Histogram*> contains a graph template for plotting the empirical distribution of the ICER. Similarly, the sheet <*CE Plane*> allows you to graph the scatter plot of the joint (empirical) distribution of the pair $(\Delta\overline{C}, \Delta\overline{E})$ obtained from bootstrap.

The <*Net Benefit Histogram*> worksheet graphs the empirical distribution of the net (monetary) benefit for a given value of λ, and the sheet <*CEAC*> plots the cost-effectiveness acceptability curve.

Finally, the <*Parametric*> worksheet asks you to calculate the mean net (monetary) benefit, its standard error, and the probability that the net (monetary) benefit is positive for a range of different λ values. This is an optional task.

Note: The aim of this exercise is for you to analyse the cost-effectiveness data and present the results, therefore to help you with your task, many of the worksheets have pre-drawn graphs.

The following list aims to guide you through the stages you should cover, but remember, the overall aim of the exercise is for you to get comfortable with analysing and presenting cost-effectiveness information.

To help you with the above task, the Excel cells you are asked to complete (as you progress through the exercise) have been highlighted in pale yellow. The cells in pale green are optional and refer to either more advanced issues or to issues that have not been discussed in the lecture. These have been included for future reference.

A more detailed 'step-by-step' guide to the exercise follows this summary.

a) Start by calculating and examining the summary statistics for the cost-effectiveness data (i.e. see section below titled 'getting to know the data').

b) Take a look at the confidence box, and make sure you understand its derivation.

c) Undertake the bootstrapping using the macro.

d) Estimate the bootstrap confidence intervals for the ICER. Calculate the descriptive statistics for the ICER, and consider the potential limitations of using the ratio in this exercise (i.e. see sections 'Bootstrapping the ICER', 'Plotting the data' and 'Summary statistics for the bootstrapped ICER').

e) Employ the bootstrap replicates to produce a cost-effectiveness acceptability curve (i.e. see section 'Generating a cost-effectiveness acceptability curve').

f) Choose a possible value of the threshold value and use this to estimate the net (monetary) benefit statistic from the data. Estimate the confidence interval for this quantity.

g) Use the same bootstrap information on the difference in mean costs and mean effect differences to calculate replicates of the net (monetary) benefit statistic (i.e. see section 'Non-parametric net benefit estimation').

h) Estimate the cost-effectiveness acceptability curve using a parametric approximation from the net (monetary) benefit statistic (optional task) (i.e. see section 'Parametric net benefit and acceptability curve').

You should aim to spend approximately 90 minutes on the exercise.

Steps to complete the exercise

An outline Excel spreadsheet for the exercise can be found on the workshop's webpage with the filename *Exercise 5 – Template 2020*. This spreadsheet provides the raw data for the exercise and gives labels and headings for the summary information that you will be asked to complete as part of the exercise. All cells you are required to complete are coloured pale yellow.

The numbered steps below refer to each stage of the exercise.

1. Getting to know the data

Spreadsheet commands employed:

- AVERAGE = calculates mean
- STDEV = calculates standard deviation
- SQRT = calculates squared root

The <*Raw data*> worksheet contains data from a hypothetical trial-based CEA comparing a new intervention (i.e. 'treatment' group) with a control (i.e. 'control' group). The RCT recruited a total of 100 patients (50 in each arm) whose cost and effect pairs are listed in columns C, D, G, and H.

Please do not change the location of these data in the worksheet, as it is the primary source of data for the macros and formulas in the other worksheets.

Your first task is to calculate the descriptive statistics in the <*Descriptive analysis*> sheet. If necessary, remind yourself of how the standard analyses (i.e. estimates of mean, standard error, confidence intervals, etc.) are undertaken. The Excel help option (F1) is a useful source.

We have said in the lecture that stochastic CEA is concerned not only with the quantification of the point estimates of the differential mean costs and effects (used to calculate the mean ICER), but also with estimating their confidence intervals. We learned that confidence intervals for ratio statistics pose particular problems for parametric techniques (i.e. we cannot estimate the variance of a ratio very easily). Hence, this exercise shows how we can use the non-parametric bootstrapping technique to estimate the confidence interval for this ICER.

[**Hint:** The standard error of the sample mean can be derived using the following formula:

$$SE(\bar{X}) = \frac{SD(X)}{\sqrt{n}}$$

where X is the random variable of interest (e.g. cost), $\bar{X}$ its mean and n is the number of observations in the trial arm of interest.

Also note that, because the two groups are independent, the standard error of the difference in mean costs can be obtained as:

$$SE(\Delta\bar{C}) = \sqrt{\left(\frac{SD(C_T)^2}{n_T}\right) + \left(\frac{SD(C_C)^2}{n_C}\right)}$$

i.e.

$$= \sqrt{SE(\bar{C}_T)^2 + SE(\bar{C}_C)^2}$$

Where $\bar{C}_T$ and $\bar{C}_C$ are respectively the mean cost in the treatment (T) and control (C) arms and n_T and n_C are their sample size.]

For simplicity, we suggest you report in this worksheet all the results of your analysis to facilitate their comparison.

2. Bootstrapping the ICER

Spreadsheet commands employed:

- VLOOKUP = vertical look up

If necessary, use the Excel help (F1) to make sure you are comfortable with what this command does.

Now go to the <*Bootstrap*> worksheet. The bootstrap macro has been written for you and can be launched by pressing the '**Run bootstrap macro**' button at the top of the <*Bootstrap*> worksheet. However, before running the macro, make sure you understand this worksheet.

On the left of the worksheet a bootstrap resample of the treatment and control groups are generated using a random number generator and the VLOOKUP command. The random numbers are generated in columns B and F, referring to the control group and treatment groups, respectively. These random numbers are generated to take an integer value between 1 and 50, and they refer to the patient id number on the <*Raw data*> sheet (cells B and F). The VLOOKUP command then searches for (and assigns) the cost and effect pairs corresponding to the patient ID from the <Raw data> sheet.

At the top of each group of 50 resamples (where 50 is the original sample size in each arm), the average cost and effect of the resamples are calculated. This is the bootstrap replication of the original data. The macro simply takes these average values and pastes them into the bootstrap resample part of the worksheet on the right. The macro should be run for a large number of replicates, here we will run it for 1,000 resamples.

(a) When you are happy with this process, press the '**Run Bootstrap Macro**' button. The macro will undertake 1,000 resamples and it will take a few minutes to complete depending on the speed of your PC's processor. When the macro has ended its run, you should be left with 1,000 replicates of the mean costs and effects from the bootstrap resamples corresponding to each arm of the trial (columns L, M, P and Q).

(b) Use these bootstrap data to calculate the cost and effect differences in each of the 1,000 replications in columns S and T.

(c) Now use these differences to estimate the vector of 1,000 bootstrapped ICERs.

(d) Now inspect the worksheet <*Histogram*>, which shows the empirical distribution of the mean ICER. This graph is already set up for you.

Note: The table below the descriptive statistics for the ICER is used to plot the histogram of the empirical distribution of the ICER. Please do not modify this table.

3. Plotting the data

You are now asked to plot the results of the bootstrap on the cost-effectiveness plane. This graph is extremely useful as it gives you direct visual information about the location of the joint density of the differential mean costs and effects.

(a) Select the <*CE plane*> worksheet. This worksheet has already been set up (using (X-Y) chart) for you to generate the graph once you fill in the vectors of the differential mean costs and effect. So, to do this, click the right hand mouse button and choose 'Select data' from the pop-up menu.

(b) The ' Select data' dialogue should now pop up and the cursor should be in the 'Data range' part of the 'Data range' tab. Press the button with the red arrow, then select the effect and cost vectors of bootstrapped differences (columns S and T) from the 'Bootstrap' worksheet that you calculated in 2(b) above. Press 'Enter' to return to the dialogue box, then press OK and the chart should appear.

What is the relationship between each of the cost and effect pairs and the ICERs that you calculated in terms of the cost-effectiveness plane?

4. Summary statistics for the bootstrapped ICER

Spreadsheet commands required: AVERAGE
MEDIAN
MIN
MAX
PERCENTILE

We can use the vector of bootstrapped ICERs to obtain descriptive statistics concerning the empirical distribution of the ICER and, most importantly, estimate the confidence intervals around its mean value.

In column Z of the worksheet <*Bootstrap*>, calculate the mean, median, minimum and maximum values of the bootstrap vector. Also use the PERCENTILE command to calculate the 2.5th and 97.5th percentile values of the bootstrap vector. These are the bootstrap estimates of the confidence interval. If necessary use the Excel help (F1) to review the PERCENTILE function.

Although the summary information you calculated for the bootstrapped ICERs is useful for descriptive purposes, the mean of the bootstrap replications will not equal the ICER calculated from the original data. It is the ICER obtained from the sample mean ΔC and ΔE and not the mean ICER obtained from the mean bootstrapped ΔC and ΔE, which is the appropriate estimate.

How useful are the estimated confidence intervals for the ICER in this case?

5. Generating a cost-effectiveness acceptability curve

Spreadsheet commands required:
IF = logical function if
COUNTIF = count the number of data points that satisfy the condition if

We know from Presentation 5.4 that confidence limits on ICERs do not give the information required when there is a non-negligible probability that the ratio could take a negative value. The alternative is to use cost-effectiveness acceptability curves. These curves show the proportion of the joint density of costs and effects falling in the acceptable region of the cost-effectiveness plane, conditional on a particular value for the ceiling ratio.
Your task is to create the data needed to build the CEAC. To do this you first need to create an indicator variable (0=no,1=yes) to help you identify whether a given point on the CE plane falls in the acceptable region or not.
Note: This is probably the most difficult part of the exercise. Our advice is to try to think through what the steps would need to look like before trying to implement this in Excel.

Now, you may want to use a nested ‘IF’ statement, and relate this to the cost-effectiveness threshold value given in cell AC3. Once you have done this, you need to calculate the proportion of bootstrap replications that identify cost-effective scenarios. (Hint: use the COUNTIF command, and remember we want the proportion, not the absolute number). Enter this information in the cell AC4. Change the threshold value and see the proportion change.
Note: In trials evaluating only 2 strategies, you could use the function AVERAGE( ) instead of COUNTIF(). The latter provides a more general solution.

The ‘Run CEAC Macro’ button can be used to produce results for the proportion of bootstrap replicates that are cost-effective for different threshold values. Once the macro has finished running, plot the results on the <*CEAC*> sheet.

(a) The cells to be completed in column AB should indicate whether the bootstrap cost-effect pair in that row represents a cost-effective point on the plane (1=yes, 0=no) in relation to the given threshold value in cell AC3.

Note: In our example a simple nested IF statement will still be appropriate as all the cost differences are positive. More generally - when the joint distribution of differential costs and effects lies on more than one quadrant - to get the correct answer you must use three 'IF' statements in combination to sort out the issue of negative ratios. (Hint: we cannot use a simple statement like IF(V12<AC3,1,0) because negative ratios in the NW quadrant of the plane will satisfy such a statement but are clearly not cost-effective). However, this is not very important as the exercise demonstrates the superiority of using net benefits instead!

(b) Once you have successfully generated the indicator in column AB, calculate the proportion of replicates that are cost-effective and enter this result in cell AC4.

(c) A macro is provided that takes the value of the threshold from column AE, puts the value into cell AC3 and records the result from AC4 in column AF. Run this macro in order to generate the data for the acceptability curve.

Note: The macro here is not strictly necessary. You could, in alternative, obtain the values for the CEAC using data table facilities in Excel.

(d) Now plot the acceptability curve. Select the *<CEAC>* worksheet and use the right mouse button to get 'Source data'. Choose the data in columns AE and AF for the 'Data range' field to plot the CE acceptability curve.

6. Non-Parametric Net-Benefit estimation

We can use the bootstrap replicates of the differential mean cost and effect to calculate the incremental net benefit statistic instead of the ICER. For each of the 1,000 bootstrap replications, calculate the incremental net benefit statistic using a decision rule of λ=1,500. Now complete the summary statistics for incremental net benefits and, using the values given, plot a histogram on the *<Net Benefit Histogram>* worksheet showing the distribution of incremental net-benefits.

Compare this distribution to that for the ICER based on the same bootstrap replications. Is the incremental net-benefit statistic roughly normally distributed?
Now calculate the proportion of replicates for which the intervention is cost-effective based on the bootstrap of the mean incremental net benefits.

(a) The incremental net-benefit replicates are obtained by multiplying the bootstrap estimate of the effect difference by the decision rule (λ=1,500) and subtracting the bootstrap estimate of the cost difference. Make sure you change cell AC3 to be £1,500.
(b) Complete the summary statistics for incremental net benefits. Note how much better behaved the statistic is.
(c) Complete the 'Is it cost-effective?' column by noting that the replication is cost-effective if the incremental net-benefit is positive. Here, you can use a simple IF statement without any nesting – much simpler than the previous approach based on ICERs and the CE plane.

7. Parametric net-benefits and acceptability curves*

*More advanced task if you have time – and for future reference!
Spreadsheet commands required: NORMDIST

Assuming that the incremental net benefit statistic follows roughly a normal distribution, we can now be more confident in employing the parametric assumptions in order to generate cost-effectiveness acceptability curves.

Switch to the <*Parametric*> worksheet.

Using the formulae for incremental net benefit and the standard error of incremental net benefit given in Presentation 5.4, complete columns B and C of the worksheet, using the values of the threshold given between £0 and £5,000. Standardise the incremental net benefit statistic in column D and use this value and the NORMDIST function to calculate the probability that the intervention is cost-effective in column E.

The variance of the incremental net benefit, for a specific value of λ, can be estimated as:

$$Var(N\hat{M}B) = \lambda^2 \cdot Var(\Delta\bar{E}) + Var(\Delta\bar{C}) - 2\lambda Cov(\Delta\bar{E}, \Delta\bar{C})$$

and the confidence interval around the mean incremental net benefit, for a give λ value is:

$$\left(N\bar{M}B - z_{\alpha/2} \cdot SE(N\bar{M}B), N\bar{M}B + z_{\alpha/2} \cdot SE(N\bar{M}B)\right)$$

Finally, plot this data on the CEAC chart to give the parametric estimate of the acceptability curve. How do the parametric and non-parametric curves compare?

(a) Calculate the incremental net-benefit statistic for the values of "lambda" given between £0 and £5,000 using the cost and effect differences from the <*Descriptive analysis*> worksheet.

(b) Also calculate the corresponding standard error of incremental net-benefits, using the estimates of standard error for the cost and effect differences from the <*Descriptive Statistics*> worksheet, together with the estimate of correlation between them.

(c) In column D, calculate the standardised value of net benefit by dividing the statistic through by its standard error. Standardisation produces a normal variable with mean 0 and variance 1. This gives the test statistic for testing net benefits as significantly different from zero.

(d) Use the test-statistic from the previous step as an input to the NORMDIST function in order to estimate the probability that net benefits is positive (i.e. that the intervention is cost-effective). The NORMDIST function has four arguments. Again, remind yourself about this function using the Excel help (F1). Use the test statistic from the previous step as the first, set the mean and the standard deviation to zero and one respectively and choose the cumulative distribution by entering the logical value 1.

(e) Finally, add the data that you have just calculated to the CEA curve figure and compare the parametric and non-parametric estimated curves.

8. Other methods for calculating confidence intervals around ICER: Fiellers' method and Taylor Series approximation*

*More advanced task if you have time – and for future reference!

This is a more advanced task and you are invited to calculate these confidence intervals only if you have completed the points 1-7 above.

To obtain these confidence intervals you will need to apply the formulas given in the appendix slide of the lecture.

Are the results different from those obtained using non-parametric bootstrap?

What can you tell about the use of parametric-based approaches as opposed to the non-parametric bootstrap in this exercise?

Are you reaching different conclusions?

References

Briggs A, Claxton K and Sculpher MJ. *Decision Modelling for Health Economic Evaluation*. 2006. Oxford University Press. Oxford. ISBN 9780198526629

Bucher HC, Guyatt GH, Griffith LE and Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; 50(6): 683-691.

Curtis L. *Unit Costs of Health and Social Care 2018*. Personal Social Services Research Unit (PSSRU); 2018. URL: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2018/ (Cited 07/05/19)

de Vries F, Setakis E, Zhang B, van Staa TP. Long-acting beta2-agonists in adult asthma and the pattern of risk of death and severe asthma outcomes: a study using the GPRD. *Eur Respir J*. 2010, 36:494-502. DOI: 10.1183/09031936.00124209

Goeree R, O'Brien BJ, Blackhouse G, Marshall J, Briggs AH, Lad R. "Cost-effectiveness and cost-utility of long-term management strategies for heartburn". *Value in Health* 2002; 5:4 312-28. doi:10.1046/j.1524-4733.2002.54145.x

Humbert M, Beasley R, Ayres J, Slavin R, Hébert J, Bousquet J, Beeh KM, Ramos S, Canonica GW, Hedgecock S, Fox H, Blogg M, Surrey K. Benefits of omalizumab as add-on therapy in patients with severe persistent asthma who are inadequately controlled despite best available therapy (GINA 2002 step 4 treatment): INNOVATE. *Allergy*. 2005 Mar;60(3):309-16. DOI: 10.1111/j.1398-9995.2004.00772.x

Lloyd A, Price D, Brown R. The impact of asthma exacerbations on health-related quality of life in moderate to severe asthma patients in the UK. *Prim Care Respir J*. 2007 Feb;16(1):22-7. DOI: 10.3132/pcrj.2007.00002

National Institute for Health and Care Excellence. NICE technology appraisal guidance 278: Omalizumab for the treatment of severe persistent allergic asthma (review of TA133 and TA201). Final scope. URL: https://www.nice.org.uk/guidance/ta278/documents/asthma-severe-persistent-patients-aged-6-adults-omalizumab-rev-ta133-ta201-final-scope2 (cited 09/05/19).

NHS reference costs 2017–2018. URL: https://improvement.nhs.uk/resources/reference-costs/#rc1718 (Cited 07/05/19)

Norman G, Faria R, Paton F, Llewellyn A, Fox D, Palmer S, Clifton I, Paton J, Woolacott N, McKenna C. Omalizumab for the treatment of severe persistent allergic asthma: a systematic review and economic evaluation. *Health Technol Assess*. 2013 Nov;17(52):1-342. doi: 10.3310/hta17520.

Novartis Pharma. NICE technology appraisal guidance 133: Omalizumab for severe persistent allergic asthma. Manufacturer's submission, 2007. URL: https://www.nice.org.uk/guidance/ta133/documents/manufacturers-submission2 (Cited 07/05/19)

Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparisons for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003; 326: 472. DOI: 10.1136/bmj.326.7387.472

Tice J, Walsh JME, Synnott P. et al. Biologic Therapies for Treatment of Asthma Associated with Type 2 Inflammation: Effectiveness, Value, and Value-Based Price Benchmarks. Draft Evidence Report. Institute For Clinical and Economic Review. 2018. URL: https://icer-review.org/wp-content/uploads/2018/04/Asthma-Draft-Report-FOR-PUBLICATION-9.24.pdf (Cited 07/05/19).

Tsuchiya A, Brazier J, McColl E. *et al*. Deriving preference-based single indices from non-preference based condition-specific instruments: Converting AQLQ into EQ5D indices. *HEDS Discussion Paper 02/01*. 2002.The University of Sheffield. Sheffield.

APPENDIX: Excel functions

Excel functions you will need to use in the practical exercises

All of the practical exercises use Excel and no other software is required. For the majority of the exercises you should be fine if you are using a Mac. However, there are some exercises which involve visual basic coding (VBA) in Excel to create macros, e.g. for simulation/bootstrapping. If you are using Excel 2004 or Excel 2011 for Mac then you should be fine since both versions incorporate VBA. However, if you are using Excel 2008 for Mac this version does not include VBA functionality and so you will not be able to run the macros, and we would therefore encourage you to either use another laptop, or (if that is not a feasible option for you) consider working jointly with another workshop delegate for the parts of the practical exercises that use macros.

Although many of you will be familiar with Excel functions, we thought it would be useful to let you know which functions you will be using during the practical exercises. It is worthwhile taking a look at the help sections for these prior to the workshop.

NB: The functions listed on the following pages are relevant for Excel 2016. If you are using another version of Excel, some of the functions may be different.

AVERAGE: This function allows you to calculate the mean of a range or set of data, e.g. =AVERAGE(a1:d1)

COUNTIF: This counts the number of cells in a data range or set that satisfy a given criteria, e.g. =COUNTIF(a1:d1,"<1")

CORREL: This function allows you to calculate the correlation between two random variables.

INT: This function returns number down to the nearest integer

SUM: This function calculates the total of a range or set of data,
e.g. =SUM(a1:d1)

IF: This is a logical function, such that Excel returns a value or "FALSE" or "TRUE" if a criteria is met, e.g. =IF(a1<1,"TRUE","FALSE")

STDEV.S: This calculates the standard deviation based on a sample of data,
e.g. =STDEV.S(a1:d1)

SQRT: This is used to calculate the square root of a value, e.g. =SQRT(a1)

MEDIAN: This function is used to calculate the median value in a range or set of values, e.g. =MEDIAN(a1:d1)

MIN: This function is used to calculate the minimum value in a range or set of values, e.g. = MIN(a1:d1)

MAX: This function is used to calculate the maximum value in a range or set of values, e.g. = MAX(a1:d1)

RAND(): Returns an evenly distributed random number greater than or equal to 0 and less than 1. A new random number is returned every time the worksheet is calculated (i.e. when you press F9). This function will be used in the Bootstrap exercise.

NORM.DIST: This allows you to draw for a normal distribution for each realisation of a particular variable. The normal distribution is defined using 4 data points: X (the data point we want the distribution for), the mean, standard deviation, and cumulative (logical function that returns the cumulative distribution function if "TRUE" and the probability mass function if "FALSE"),
e.g. =NORM.DIST (a1,a2,a3,TRUE)

NORM.INV: This allows calculating the inverse of the normal cumulative distribution for a specified probability, mean and standard deviation. When a random draw is required the probability should be specified using the RAND() function. It is useful in the calculation of parametric confidence intervals.

PERCENTILE: This function returns the k-th percentile of values in a range. You can use this function to establish a threshold of acceptance. For example, you can decide to examine candidates who score above the 90th percentile. In our exercises this function will be used to calculate non-parametric confidence intervals, e.g. = PERCENTILE.EXC(a1, a2, a3, k)

VLOOKUP: This stands for vertical look up. This function searches for a value in the leftmost column of a table, and then returns a value in the same row from a column you specify in the table. VLOOKUP is defined using 4 data points, lookup value (value in the first column of the array), table array (range or name of table that you are looking up information in), column index number (column number in the table) and range lookup (logical function that allows to return either an exact or approximate match), e.g. =VLOOKUP(1,A1:C10,2)

Macros: Visual basic can be used to write programs to undertake repetitive tasks in Excel. We often use these to run simulations. You will not be required to write macros on the workshop but to run these you will need to go to view>macros or use the automatic links created on the sheet.
You will be required to enable macros when opening Excel. To do this go to 'macro security' on the developer tab and choose 'enable all macros'.
Plotting graphs: You will need to plot graphs in some of the exercises. The graph function can be found at insert>chart. Excel will lead you through the process of building the graph. If you want to make changes after you have inserted the graph just right click on the graph and you will be given a drop down menu.

Recommended reading

http://www-users.york.ac.uk/~mb55/pubs/pbstnote.htm
Statistics Notes in the *British Medical Journal*, with particular reference to:

- Altman DG and Bland JM. The normal distribution. *BMJ* 1995; 310: 298.
- Bland JM and Altman DG. The odds ratio. *BMJ* 2000; 320: 1468.
- Vickers AJ and Altman DG. Analysing controlled trials with baseline and follow up measurements. *BMJ* 2001; 323: 1123-1124.
- Altman DG and Bland JM. Variables and parameters. *BMJ* 1999; 318: 1667.
- Altman DG and Bland JM. How to randomise. *BMJ* 1999; 319: 703-4.
- Altman DG and Bland JM. Treatment allocation in controlled trials: why randomise? *BMJ* 1999; 318: 1209.
- Bland JM and Altman DG. Bayesians and frequentists. *BMJ* 1998; 317: 1151.

Spiegelhalter DJ *et al.* An introduction to Bayesian methods in health technology assessment. *BMJ* 1999; 319: 508. DOI: 10.1136/bmj.319.7208.508
https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one
Statistics at Square One (BMJ electronic book), with particular reference to the following topics:

- Mean and standard deviation (https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/2-mean-and-standard-deviation)
- Populations and samples (https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/3-populations-and-samples)
- Statements of probability and confidence intervals (https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/4-statements-probability-and-confiden)

Differences between means: type I and type II errors and power (https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/5-differences-between-means-type-i-an)
Bayesian methods in health services research
Spiegelhalter David J, Myles Jonathan P, Jones David R, and Abrams Keith R. Methods in health service research: An introduction to bayesian methods in health technology assessment. *BMJ*, Aug 1999; 319: 508. DOI: 10.1136/bmj.319.7208.508
Gill Christopher J, Sabin Lora, and Schmid Christopher H. Why clinicians are natural Bayesians. *BMJ*, May 2005; 330: 1080. DOI: 10.1136/bmj.330.7499.1080
Lilford RJ and Braunholtz D. For Debate: The statistical basis of public policy: a paradigm shift is overdue. *BMJ*, Sep 1996; 313: 603-607. JSTOR, www.jstr.org/stable/29732771
Goodman SN. Toward Evidence-Based Medical Statistics. 1: The P Value Fallacy. *Ann Intern Med*, June 15, 1999; 130(12): 995-1004. DOI: 10.7326/0003-4819-130-12-199906150-00008
http://annals.org/aim/fullarticle/712762/toward-evidence-based-medical-statistics-1-p-value-fallacy
Goodman SN. Toward Evidence-Based Medical Statistics. 2: The Bayes Factor. *Ann Intern Med*, June 15, 1999; 130(12): 1005-1013. DOI: 10.7326/0003-4819-130-12-199906150-00019
http://annals.org/aim/fullarticle/712763/toward-evidence-based-medical-statistics-2-bayes-factor
Useful websites for refreshing statistical concepts:
http://www.statsoft.com/Textbook
This Electronic Statistics Textbook offers training in the understanding and application of statistics
https://faculty.chass.ncsu.edu/garson/PA765/statnote.htm
On line book in statistics
Further reading:
Briggs AH (2001) Handling uncertainty in economic evaluation and presenting the results. In *Economic evaluation in health care. Merging theory with practice*. MF Drummond and A McGuire. Oxford, Oxford University Press. 172-214
Briggs AH, O'Brien BJ, Blackhouse G (2002) Thinking outside the box: Recent Advances in the

Analysis and Presentation of Uncertainty in Cost-Effectiveness Studies. *Annual Review of Public Health*; Vol. 23: 377-401. DOI: 10.1146/annurev.publhealth.23.100901.140534