The Difference Method: a method for producing probabilistic data samples where one value is known to be greater than another value.  
  
Abstract:  
Background: If two variables are believed to be related in such a manner that one is greater than another then independent sampling within a probabilistic sensitivity analysis (PSA) may be inappropriate.  
  
Objective: To describe and test a method, the ‘Difference Method’ (DM), for generating PSA samples where the constraint that one value is greater than another is maintained and which also satisfy both clinical and statistical validity***.***Method: The DM approach samples the target variables via the difference parameter. If the target variables are bounded, it involves transforming the variable to the real line and then sampling via the difference parameter. A case study was conducted to compare the DM approach with two commonly applied methods in literature (independent sampling and sampling using the same random number generator).   
   
Results: The DM generated PSA samples have summary statistics that were similar to the given summary statistics whilst maintaining the constraint that one value was greater than another. It also implies plausible correlation between the two target variables. An excel workbook was provided to implement the method.   
   
Conclusions: Clinical and statistical validity of PSA samples is important in economic modelling. A standard sampling method may violate these validities. The DM approach provides a solution to overcome the problem with naïve sampling methods and should be considered in PSA.   
  
  
Introduction  
  
Sometimes there is an absolute belief that the value of one variable is greater than the value of another. There may be uncertainty around the true mean values of the variables, but the ordering of the values is known. For example, if an individual rates his/her general health as ‘good’, then later as ‘fair’, we might be uncertain about how to map the ‘good’ and ‘fair’ health evaluations onto a numeric scale, but assume the ‘good’ general health score will be higher than the ‘fair’ general health score. Health related quality of life (HRQoL) for two different severity levels of a disease may be related in this manner.  
  
In the context of health technology assessment, probabilistic sensitivity analysis (PSA) represents the generally accepted approach for characterising the uncertainty in parameters included in an economic model and for producing accurate results in non-linear models. [[[1]](#endnote-1),[[2]](#endnote-2)] This involves simulating a large number of realisations of the economic model, each time sampling values from the distributions applied to each uncertain parameter included in the model.  
  
In a model where the distributions of parameters that we believe to be constrained in relation to their relative values overlap, independent sampling of mean values could result in a PSA that lacks clinical validity, as in some realisations the logical constraint may be violated with the sampled value of parameters may be equivalent to assuming that having a disease makes people healthier. [[[3]](#endnote-3),[[4]](#endnote-4)] Quantile matching between distributions is likely to underestimate the true uncertainty and could still violate the known ordering of the mean values. Another flawed alternative observed in papers sent to the authors for peer-review is a method whereby samples are excluded when the ordering assumption is violated – this results in the mean of the sampled realisations not equalling that of the source data and therefore should be avoided.  
  
The aim of this paper is to describe an approach to generating paired samples for PSA whilst maintaining both clinical and statistical validity. We call this the difference method (DM). To have clinical validity when sampling parameters where it is known that one value is greater than another, all PSA realisations should exhibit this characteristic. To have statistical validity, 1) if the target variables are bounded then all PSA realisations should be in the range as the target variables; 2) the mean and variance of the sampled values should match closely to the mean and variance of the target variables; 3) the induced correlation between the sampled values should be plausible.   
  
Method  
Suppose that there are two variables and with the value of is greater than the value of , where the distribution of has mean and variance , and the distribution of has mean and variance . If , then define

*.* (1)

If , then define

*.* (2)

Let and be the mean and variance of the distribution of the difference . Then assuming and in equation (1) are independent and and in equation (2) are independent, we get and using equation (1) and (2).   
  
If both variable and are not bounded, we assume a Gamma(,) distribution for so that the difference is positive, where is the shape parameter and is the rate parameter. Hence,  
 and. These can be solved simultaneously to give and .

* If , then the sampling procedure involves sampling from Normal(,) and from Gamma(,); sampled values of and are derived from sampled values of using equation (1).
* If , then the sampling procedure involves sampling ; sampled values of and are derived from sampled values of using equation (2).

If both and are bounded between 0 and 1, for example probabilities and most HRQoL; or bounded to be positive, for example costs, we suggest a three-step sampling procedures.

* Step 1: Given mean and variance of and , sample and from Beta distributions if the and sample from Gamma distributions if .
* Step 2: Transform sampled and to the real line, using and if the and using and if the ). Calculate the mean and variance for sampled and .
* Step 3: Use the DM approach described above for unbounded variables to re-define be the difference between and . Sample , then either using sample valued for from Step 2 and sampled values for to derive sampled values for or using sampled values for from Step 2 and sampled values for to derive sampled values for depending on the variance between and . Back transform to obtain sampled values for and .

A case study comparing the DM approach with other standard sampling methods

In health technology appraisals, it is common that the modeller does not have access to the individual patient level data (IPD), but only summary statistics derived from the IPD. In this case study, suppose that there is an active (worse) and remission (better) state in an economic model for a particular condition. The input parameters we are sampling in PSA are HRQoL and cost. Suppose that the modeller is given the information about the input parameters as following.

* Active (worse) state: the input parameter for HRQoL has mean 0.54 and variance 0.019; the input parameter for cost has mean 110 and variance 15.
* Remission (better) state: the input parameter for HRQoL has mean 0.70 and variance 0.016; the input parameter for cost has mean 100 and variance 10.

Results  
Table 1 shows that the mean and variance of 5000 sampled realisations using three sampling approaches (the DM, independent sampling and sampling using the same random number generator) are all closely match the given summary statistics in Table 1. Figure 1 shows the scatter plots of sampled pair of values using the three methods.   
The independent sampling approach lacks of both clinical and statistical validity since it may produce sampled value of HRQoL in the worse state is higher than sampled value of HRQoL in the better state and sampled value of cost in the better state is higher than sampled value of cost in the worse state; there is no between the two variable. The sampling using the same random number generator method lacks of statistical validity since the induced correlation between the two sampled values are implausibly high. The DM approach guarantees to maintain the constraint that one value is greater than another for each sampled pair. The induced correlation also seem to be plausible.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable State | Variable Type | Mean (Variance) known to modeller | | |
| Difference method | Independent sampling | Same random number |
| active (worse) state | HRQoL | 0.54 (0.019) | 0.54 (0.019) | 0.53 (0.019) |
| Cost | 110.02 (15.16) | 110.05 (14.82) | 110.15 (15.31) |
| remission (better) state | HRQoL | 0.70 (0.017) | 0.70 (0.017) | 0.69 (0.016) |
| Cost | 100.04 (9.84) | 100.04 (9.84) | 100.12 (10.20) |

Table 2: mean and variance of 5000 sampled realisations using three sampling approaches.

Figure 1:

Discussion  
  
Failure to account for constraints between parameter values may result in PSA values that do not accurately characterise the uncertainty present in a decision problem. This could result in decisions made on the allocation of scarce health care resources to be sub-optimal, although the direction of bias would depend on the specific model. Other outputs and analyses that are reliant on the PSA, such as cost effectiveness acceptability curves and frontiers, and value of information analyses, are likely to also be flawed if the constraint that the value of one variable is bigger than another is not accounted for appropriately.   
Naïve sampling methods should not be used if it is believed that two target variables have the constraint that value of one variable is bigger than another. In the case study we have shown that the two naïve sampling approaches we used either lack of both clinical and statistical validity (the independent sampling) or lack of statistical validity (the sampling using the same random number generator). When the distribution of the two target variables do not overlap, the independent sampling approach may not violate clinical validity. Hence, the modeller needs to decide whether no correlation between the two target variables is more plausible then positive correlation.  
The DM has been shown in the case study to be effective in generating bivariate estimates which satisfy both clinical and statistical validity. It provides a solution to an issue that may have important implications for the interpretation of economic evaluations of health technologies. An earlier version of the method has been used in recent work for the National Institute for Health and Care Excellence.[[[5]](#endnote-6),[[6]](#endnote-7)] One drawback of the DM approach is that it does not work if utilities are believed below 0. A modification of the method is required.

We have developed an Excel workbook which implements the DM approach, which is included as an online appendix to this paper. We hope it helps those who wish to apply this approach when sampling parameter values under a constraint that the value of one variable is bigger than another..   
   
In conclusion, when producing PSA samples both clinical and statistical validity should be checked. Where there is a strong belief that variables are constrained in that one value is greater than another the DM should be considered as a method to ensure the clinical and statistical validity of PSA samples and analyses derived from these samples.

1. Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health Econ [Internet]. 2005 Apr [cited 2012 Jul 15];14(4):339–47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15736142 [↑](#endnote-ref-1)
2. NICE. Guide to the methods of technology appraisal [Internet]. 2008. Available from: http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf [↑](#endnote-ref-2)
3. Kearns B, Lloyd Jones M, Stevenson M, Littlewood C. Cabazitaxel for the second-line treatment of metastatic hormone-refractory prostate cancer: a NICE single technology appraisal. Pharmacoeconomics [Internet]. 2013 Jun [cited 2014 Nov 24];31(6):479–88. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23580356 [↑](#endnote-ref-3)
4. Carroll C, Stevenson M, Scope A, Evans P, Buckley S. Hemiarthroplasty and total hip arthroplasty for treating primary intracapsular fracture of the hip: a systematic review and cost-effectiveness analysis. Health Technol Assess [Internet]. 2011 Oct [cited 2014 Dec 7];15(36):1–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21978400 [↑](#endnote-ref-4)
5. Harnan SE, Tappenden P, Essat M, Gomersall T, Minton J, Wong R, et al.Measurement of exhaled nitric oxide concentration in asthma: a systematic review and economic evaluation of NIOX MINO, NIOX VERO and NObreath. Health Technol Assess 2015;19(82) [↑](#endnote-ref-6)
6. Archer R, Tappenden P, Ren S, Martyn-St James M, Harvey R, Basarir H, et al.Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262): clinical effectiveness systematic review and economic model. Health Technol Assess 2016;20(39) [↑](#endnote-ref-7)