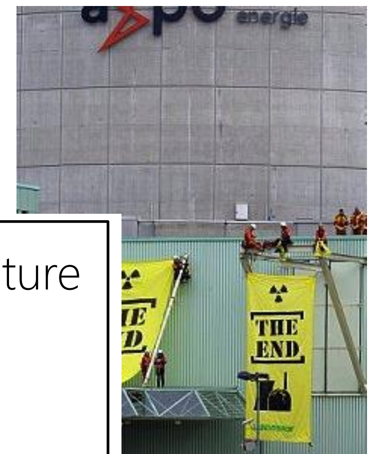
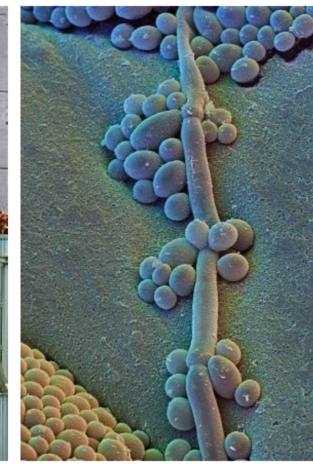
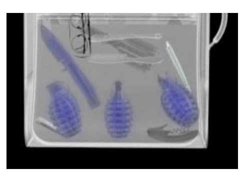


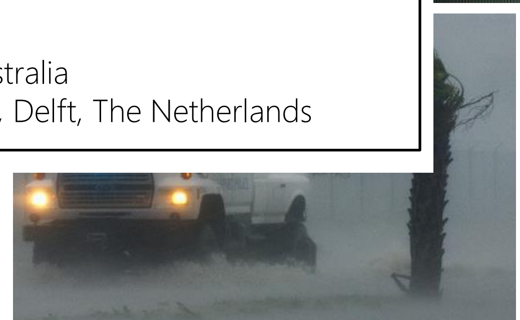


Expert Judgement for Dependence Elicitation: A Literature Review and Future Research Directions

COST Action IS1304 “Expert Judgement in Health”, Malta, October 2015

Christoph Werner | University of Strathclyde, Glasgow, UK
Anca M. Hanea | University of Melbourne, Melbourne, Australia
Oswaldo Morales Napoles | Delft University of Technology, Delft, The Netherlands

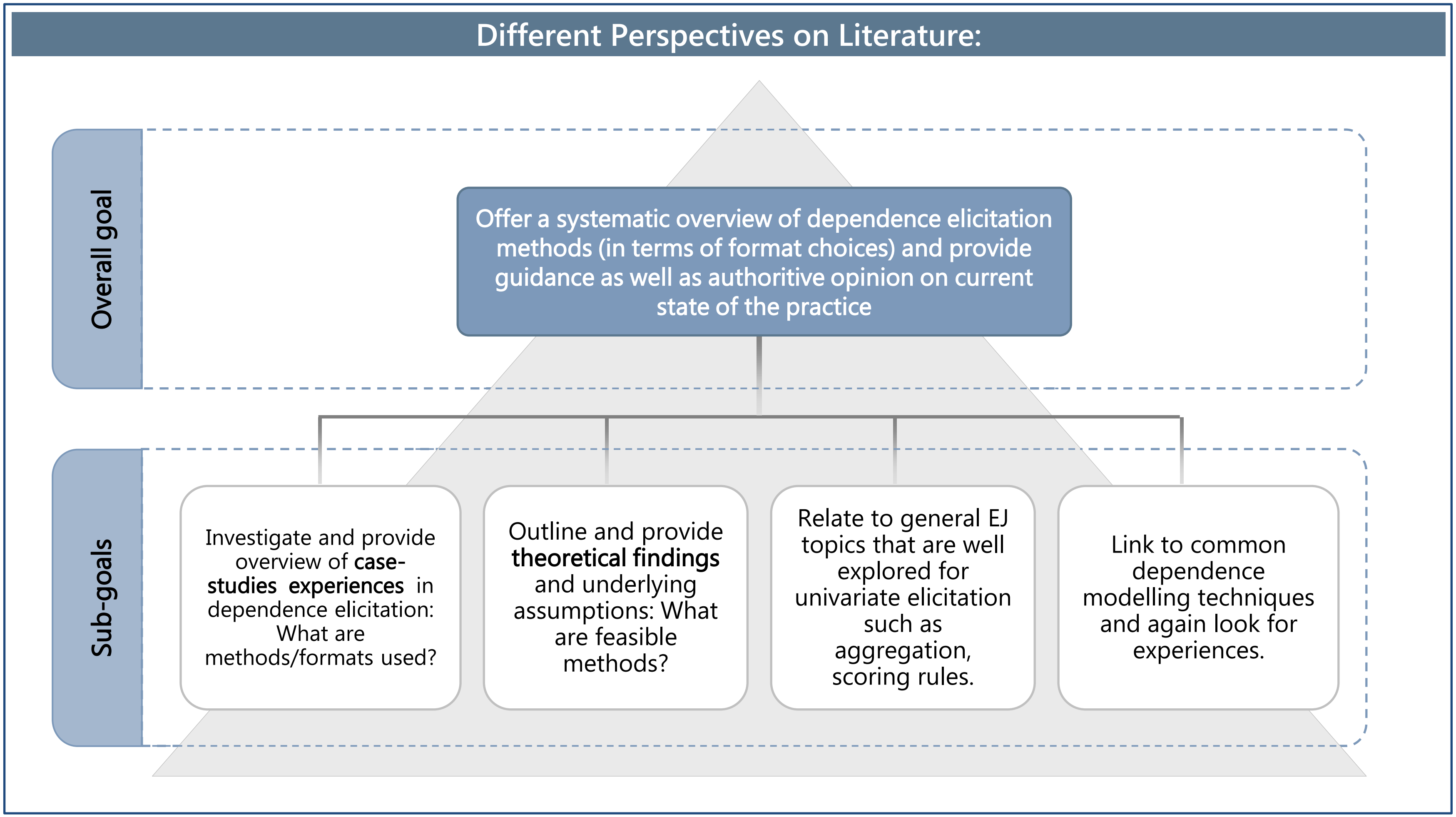




Today's content...

Today's content...

Aims & Objectives:



Today's content…

1. Motivation & Objectives
2. Preliminaries
3. Overview of Case-Study Literature
4. Findings (1): Format Choices
5. Findings (2): Link to Dependence Modelling
6. Further Discussion & Conclusion

Preliminaries:

Definition of Dependence:

Dependence in Subjective Uncertainty:

As in Daneshkhah and Oakley (2010): Two random variables, X and Y, are independent if experts do not change their beliefs about X given information about Y. Hence, dependence means that new information on Y changes the belief on X. It is important to note that here it is not necessary for X and Y to be causally or physically related, but dependence is rather a property of an expert's belief about X and Y. In fact, a main desirable property of an elicitation method is that experts consider the information they provided as reflected in the final representation of the joint/multivariate distribution

Cooke and Kraan (1996) make a distinction between *lumpy* and *smooth* dependence. The former refers to the case when switching values for Y has some effect on various processes that influence the value for X but the exact connection between the two variables is not (completely) understood. Thus, the connection between X and Y is uncertain itself. For smooth dependence on the other hand, this connection is well understood.

Preliminaries:

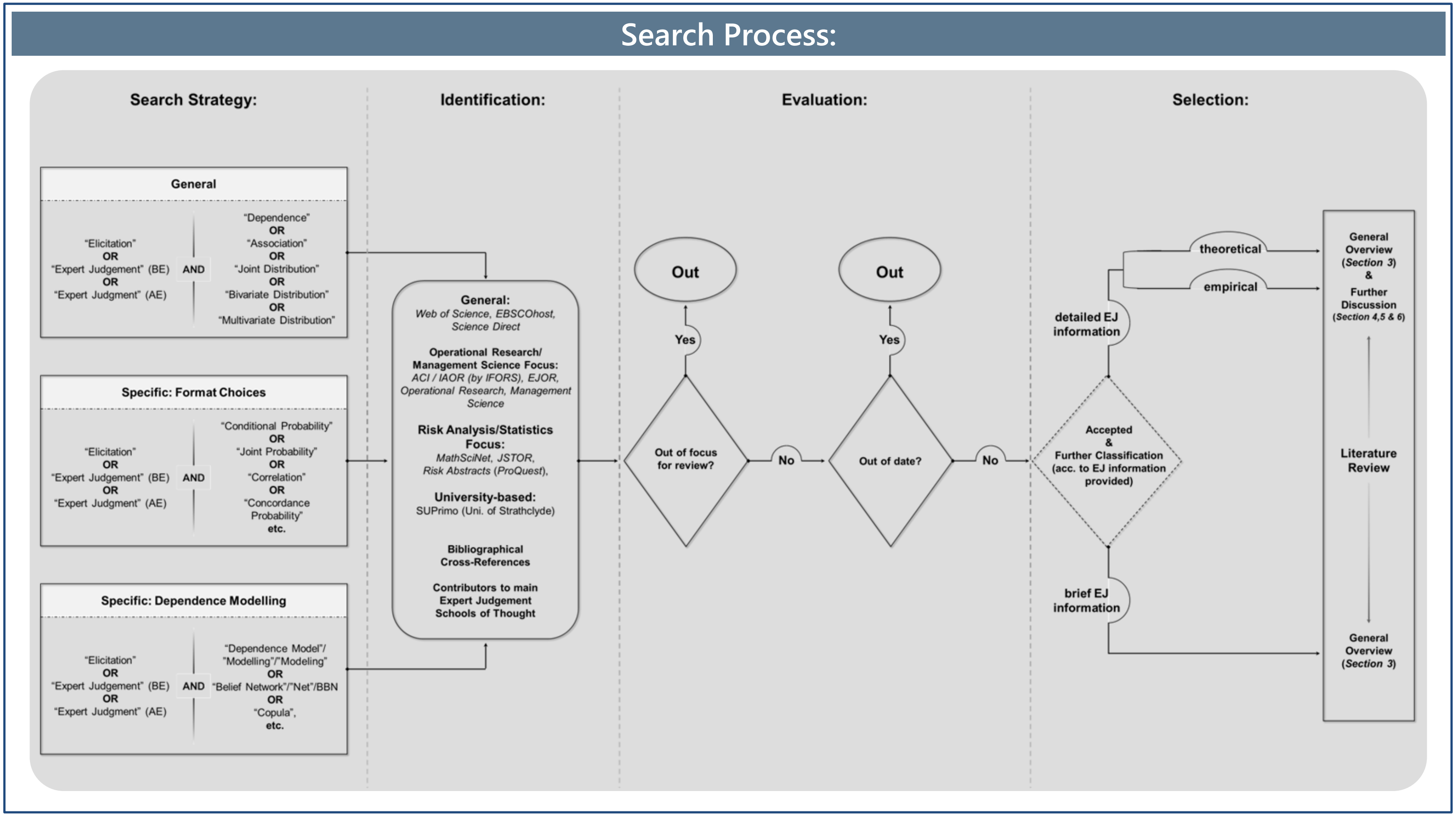
Omitted Approaches:

Transformation and Restructuring omitted:

Some popular methods (and main references) concern: joint probabilities expressed as univariate distributions through isoprobability contours (Abbas et al., 2010), probabilistic inversion for assessing model parameters by estimating its output value (Kraan and Bedford, 2005), predictive elicitation and eliciting hyper-parameters of statistical dependence models such as (Bayes) linear or regression models (Farrow, 2002).

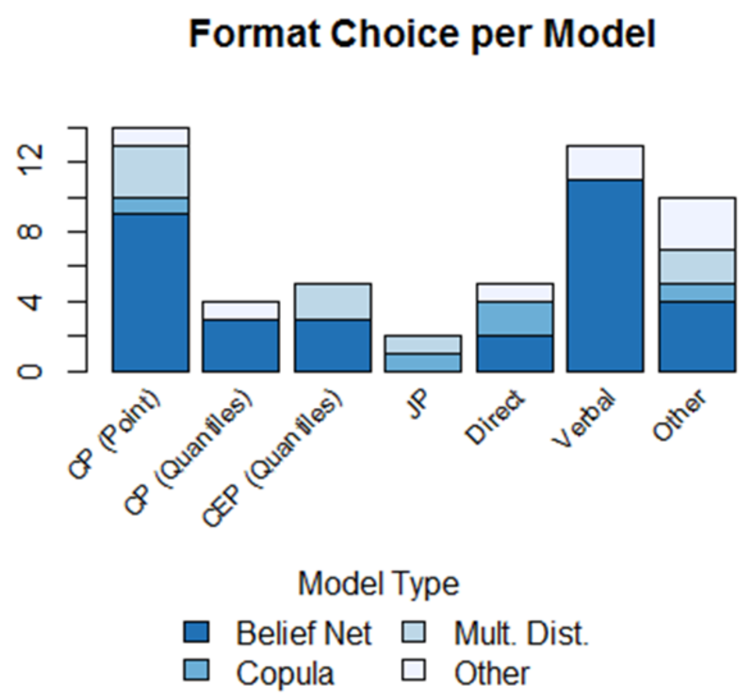
Today's content...

Design of Search Process:



Overview of Literature:

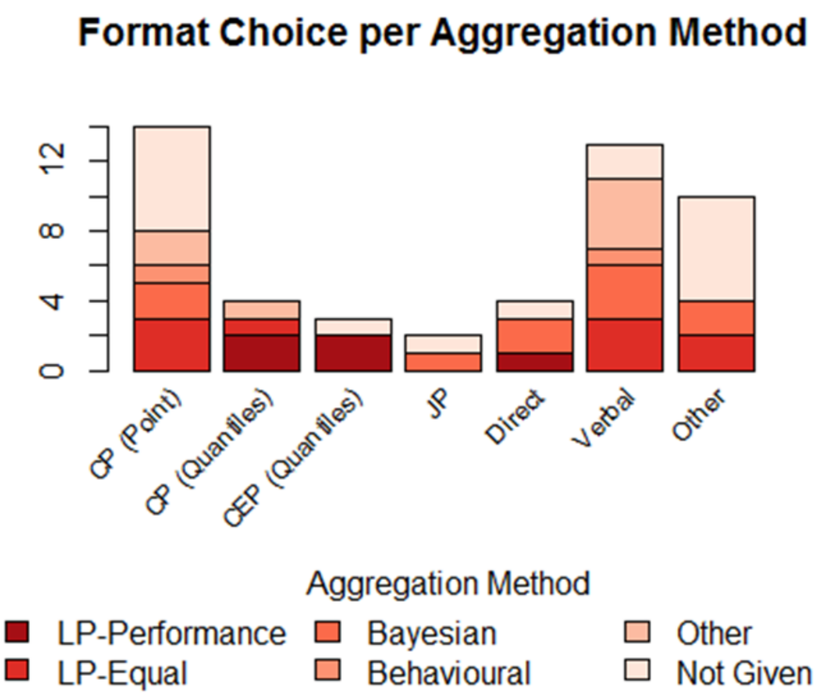
Categorisation of Indentified Literature:



** n=53; Joint Probability (with/without CIs; one each)*

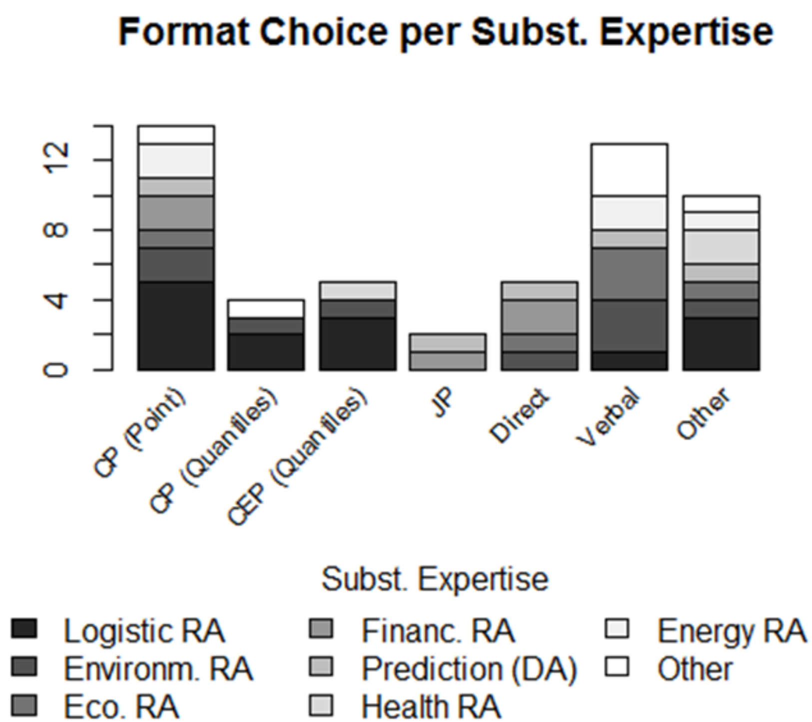
Overview of Literature:

Categorisation of Indentified Literature:



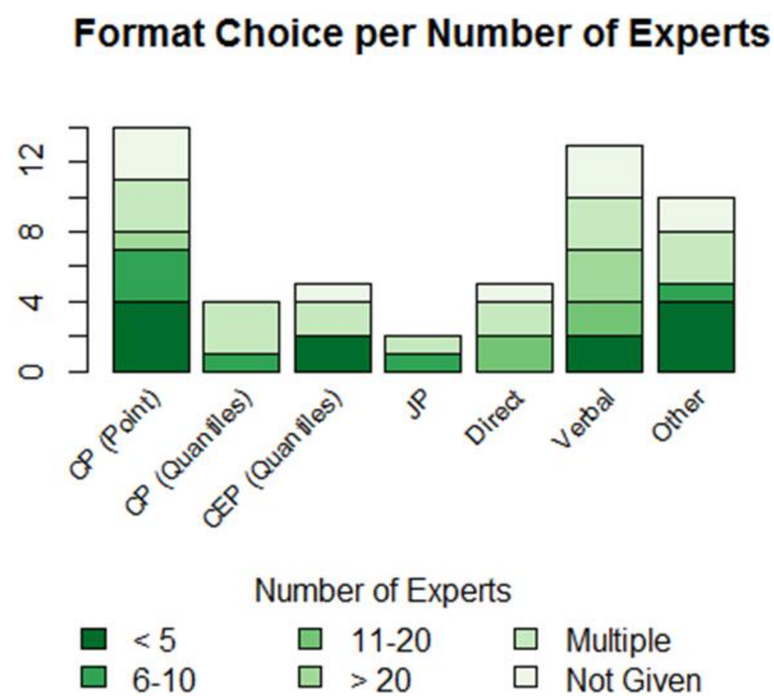
Overview of Literature:

Categorisation of Indentified Literature:



Overview of Literature:

Categorisation of Indentified Literature:



Today's content...

Desiderata of Format Choices:

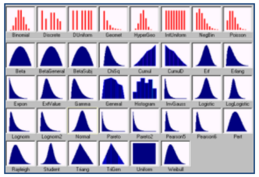
Some Considerations

Modelling

Rigorous Foundations

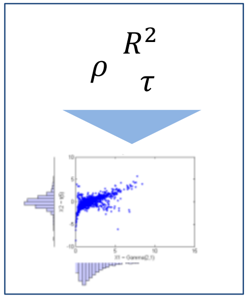
Defensible in terms of Probability Theory

Flexibility

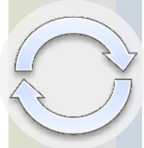


Possible to use assessments in various situations

Little Manipulation

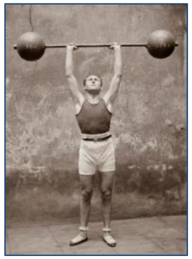


Direct linkage to modelling procedure; little probabilistic inversion



Assessment

Intuitive Interpretation



Task should be seen easy and credible

Accuracy/Coherence

Individuals with similar knowledge should come to similar results as empirical data

Format Choices: Probabilistic

Condtional Probability

Framing

Consider the pair of variables, X and Y. Suppose now that Y has been observed to be above your median value for it. What is the probability that X lies also above your median value for it?

Notation

$$P_{CP}(x, y) := P(X \geq x_i | Y \geq y_i)$$

$$P_{CP}(x, y) := P(X \geq x_i | Y = y_i)$$

Independence: $P(X \geq x)$

Pos. Dependence: $P_{CP} \in (P(X \geq x), 1]$

Neg. Dependence: $P_{CP} \in (P(X \geq x), 1]$

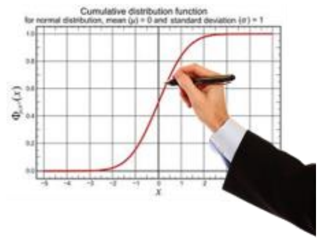
Transformation

Relation to Rank Correlation

Assessment Burden

In an experimental setting (Clemen et al., 2000), conditional probability was among the worst performance for coherence, so mathematical feasibility, and on fourth out of six places in terms of accuracy against empirical data; recommended to alterate (Cooke and Kraan, 1996)

Format Choices: Probabilistic



Condtional Exceedance Probability

Framing

Suppose that not only Y_1 but also Y_2 has been observed above your median value for it. What is now your probability that also X will be observed above your median value?

Notation

$$P_{CEP}(x, y_1, y_2) := P(X \geq x_{0.5} | Y_1 \geq y_{1,0.5}, Y_2 \geq y_{2,0.5})$$

Independence: $P(X \geq x_{0.5} | Y_1 \geq y_{1,0.5})$
Pos. Dependence: $as\ before$
Neg. Dependence: $as\ before$

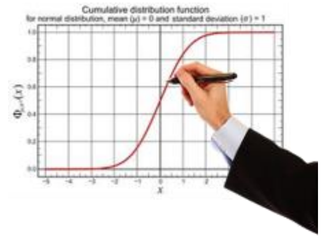
ExampleBBN

[Example BN]

Assessment Burden

Morales et al. (2013) examined probabilistic against direct approaches; fine particular matter exposer near five different power generating stations in Alabama, USA; when averaging out the absolute difference of empirical data and all individual answers, direct elicitation outperformed CEP; for d-Calibration this was vice versa; associative strength competition (Shanks, 2004)

Format Choices: Probabilistic



Joint Probability

Framing

Consider the pair of variables X and Y. What is the probability that both are within the lower (upper) k_{th} percentage of their respective distributions?

Notation

$P_{JP}(x, y) := P(X \geq x, Y \geq y$

Independence: $P_{JP}(x, y) = F_X(x) F_Y(y)$

Pos. Dependence: $P_{JP}(x, y) = F_X(x)$ or $P_{JP}(x, y) = F_Y(y)$

Neg. Dependence: $P_{JP}(x, y)$ approximates 0

Alternative

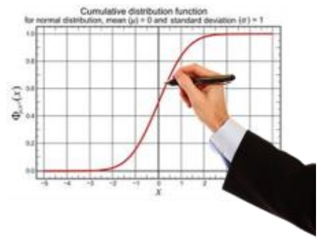
Moala and O'Hagan (2010): $P_{JP}(x, y) := P(x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2)$

Fackler (1991): $P_{MDC}(x, y) := P\left((X - x_{0.5}) - (Y - y_{0.5}) > 0\right)$

Framing Alternative

Consider the pair of variables X and Y. You have indicated that there is a 50/50 chance of X being above or below $x_{0.5}$ and Y above or below $y_{0.5}$. What is the probability that X and Y will either both be above or both be below their medians?

Format Choices: Probabilistic



Concordance Probability

Framing

Suppose we randomly choose the two variables X_i and X_j from their common underlying population. Given that $X_i > X_j$ for category a, what is your probability that the relation $X_i > X_j$ also holds for category b?

$$\mathrm{P_C}(x,y) = \frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 1_{C^*}\left((x_i,y_i),(x_j,y_j)\right)}{\binom{n}{2}}$$

Independence: $\mathrm{P_C}(x,y)$ approximates 0.5
Pos. Dependence: $\mathrm{P_C}(x,y)$ approximates 1
Neg. Dependence: $\mathrm{P_C}(x,y)$ approximates 0

Notation

$$C^* = \{(x_i - x_j)(y_i - y_j) > 0\}$$

Assessment Burden

Clemen et al., (2000), this technique performed reasonably accurate in comparison to other methods and only rarely incoherent assessments outside the mathematically feasible bounds were made. Similarly Gokhale and Press (1982) as well as Garthwaite et al. (2005); Kunda and Nisbett, (1986) came to the conclusion that this method is reasonably accurate and therefore might be preferred ***if feasible***; degree of relatedness in psychological studies

Format Choices: Conditional Expectation

Condtional Quantile (Fractile/Percentile)

Framing

Consider variables X and Y. Given the value for X has been observed at its i^{th} quantile, q_i. What is your expectation of Y's value in terms of its quantile?

Notation

$$\mu_{\min} \leq \mathrm{E}(\mathrm{F_X}(\mathrm{x})|\mathrm{Y} = \mathrm{y_k}) \leq \mu_{\max}$$
$$\mu_{\min} = \min\{\mathrm{F_Y}(\mathrm{y}), 1 - \mathrm{F_Y}(\mathrm{y})\}$$
$$\mu_{\max} = \max\{\mathrm{F_Y}(\mathrm{y}), 1 - \mathrm{F_Y}(\mathrm{y})\}$$

Independence: $\mathrm{E}(\mathrm{F_X}(\mathrm{x})|\mathrm{Y} = \mathrm{y_k}) = 0.5$

Pos. Dependence: $close\ to\ max$

Neg. Dependence: $close\ to\ min$

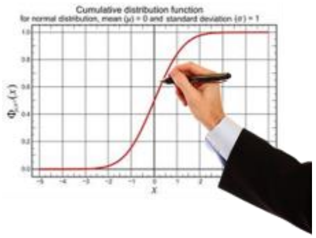
Transformation

Relation to Rank Correlation

Assessment Burden

Clemen et al., (2000) found that the performance of conditional quantile (fractile/percentile) methods is similar to that of joint and conditional probability methods; overall performance, similarly difficult as they do for probabilistic formats; further expert must understand fractiles and the notion of regression towards the mean which might induce additional cognitive difficulties (Clemen and Reilly, 1999)

Format Choices: Statistical



Direct Correlation Coefficient

Framing

Consider variables X and Y. What is the (rank) correlation between the two?

Notation

$\rho_{X,Y}$ defined on interval [0,1]

Independence: close to $\rho = 0$
Pos. Dependence: *close to* $\rho = +1$
Neg. Dependence: *close to* $\rho = -1$

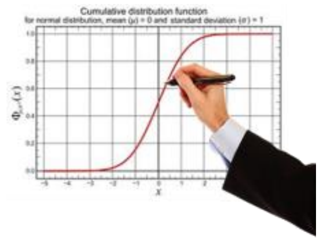
Alterations

rank correlations (in contrast to product-moment one) independent of its marginal distributions implying that its values are always in the aforementioned interval; for choosing the appropriate correlation coefficient, a facilitator/analyst has to take into account what kind of relationship is needed; Rank correlations, such as Spearman's version, assume monotonicity while Pearson's product moment coefficient needs linearity (Reilly, 2000)

Assessment Burden

several (conflicting) conclusions made from research on this format choice; some studies such as Kadane and Wolfson (1998); Morgan and Henrion (1990); Gokhale and Press (1982) view a direct method as unreliable; even trained statisticians will have difficulties with this method and even with the graphical output in form of a scatterplots; Revie et al. (2010); Clemen et al. (2000); Clemen and Reilly (1999) found out that these actually outperformed probabilistic ones and Bayes Linear ones

Format Choices: Statistical



Ratios of Rank Correlation

Framing

Given your previous estimate, what is the ratio of r_{X,Y_2} to r_{X,Y_1}?

Notation

$\rho_{X,Y_1|Y_2}$, corresponds to ratio $R = \frac{r_{X,Y_2}}{r_{X,Y_1}}$

Independence: $E(F_X(x)|Y = y_k) = 0.5$
Pos. Dependence: $close\ to\ max$
Neg. Dependence: $close\ to\ min$

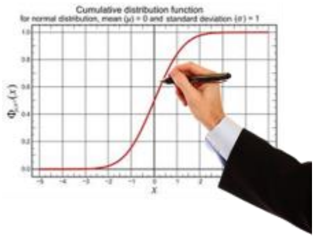
Transformation

Relation to Rank Correlation

Assessmnet Burden

Empirical comparisons to other techniques have not shown any superior nor inferior performance of this method; the authors claim that experts often actually think in terms of unconditional correlations anyway which then facilitates the assessment of the ratio; Delgado et al. (2012) and Roelen et al. (2008)

Format Choices: Statistical



Verbal

Framing

Consider variables X and Y. Given the value for X has been observed at its i^{th} quantile, q_i. What is your expectation of Y's value in terms of its quantile?

$$\rho = \frac{S_{X,Y} - 4}{3}$$

Independence: $close\ to\ 4$
Pos. Dependence: $close\ to\ 7$
Neg. Dependence: $close\ to\ 1$

Notation

Transformation

Relation to Rank Correlation

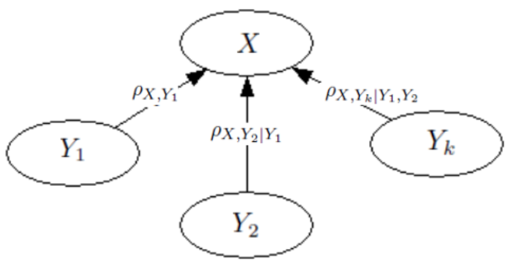
Technique for Human Error Rate Prediction THERP

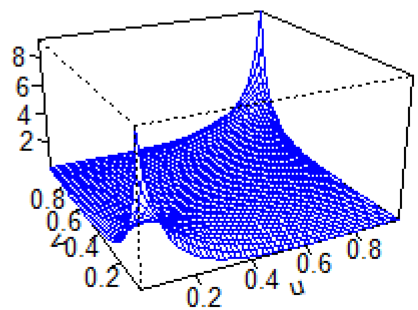
Despite its obvious subjectivity in determining the scale; popular and a strong performance with this technique in terms of coherence and accuracy can be observed and indeed is this method intuitively understandable

Today's content…

Link to Dependence Modelling:

SEJ for Dependence Models:





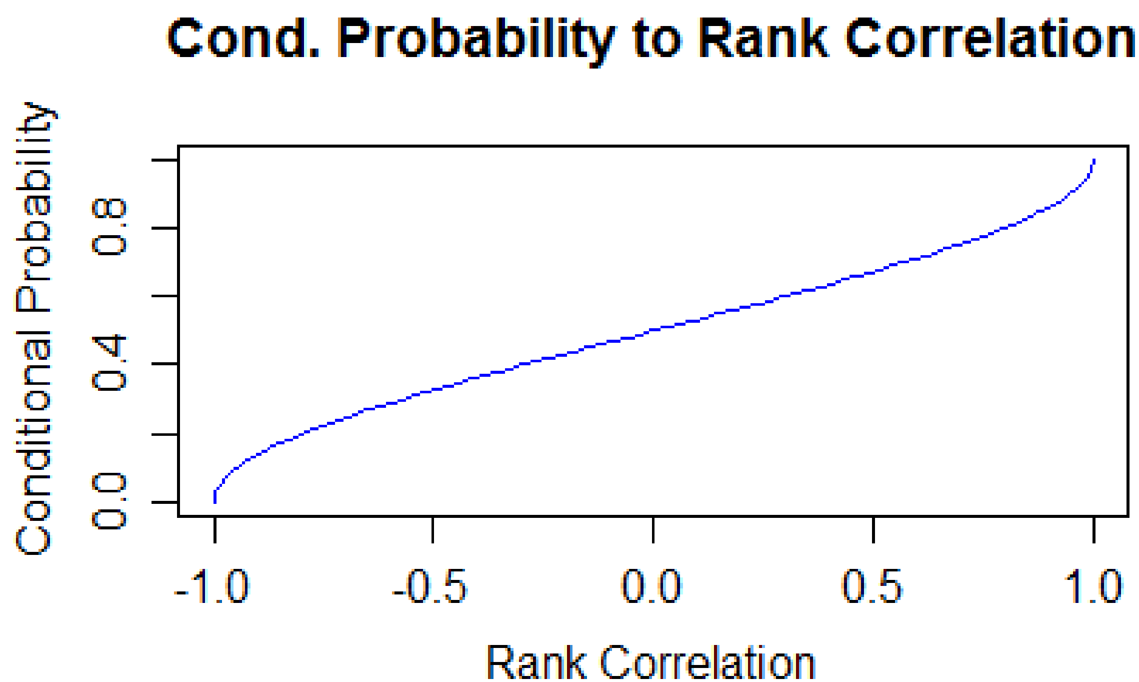
Other

Today's content...

Thank you for your attention!

Appendix (1):

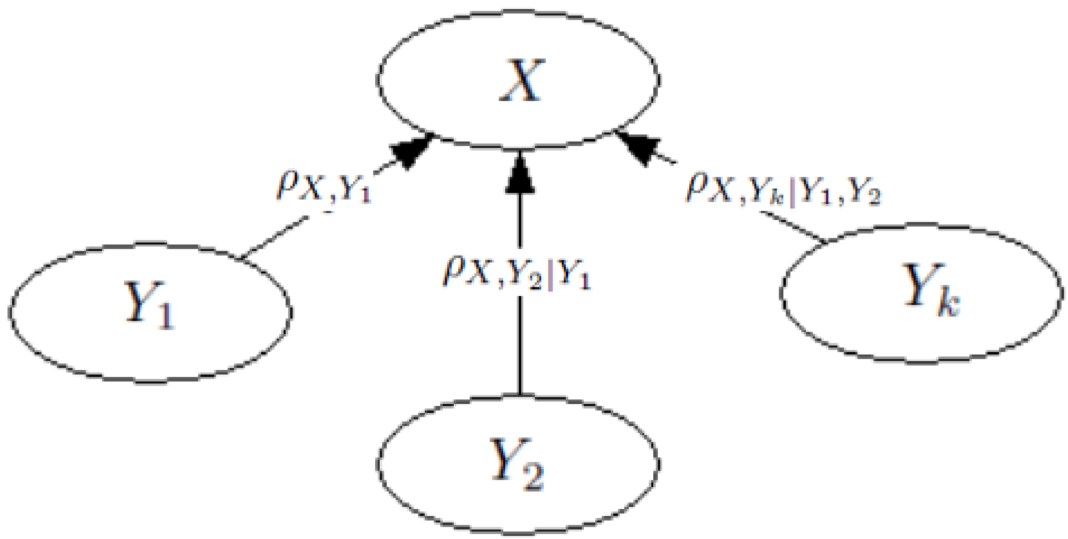
Relation of Cond. Probability to Rank Correlation



Back

Appendix (2):

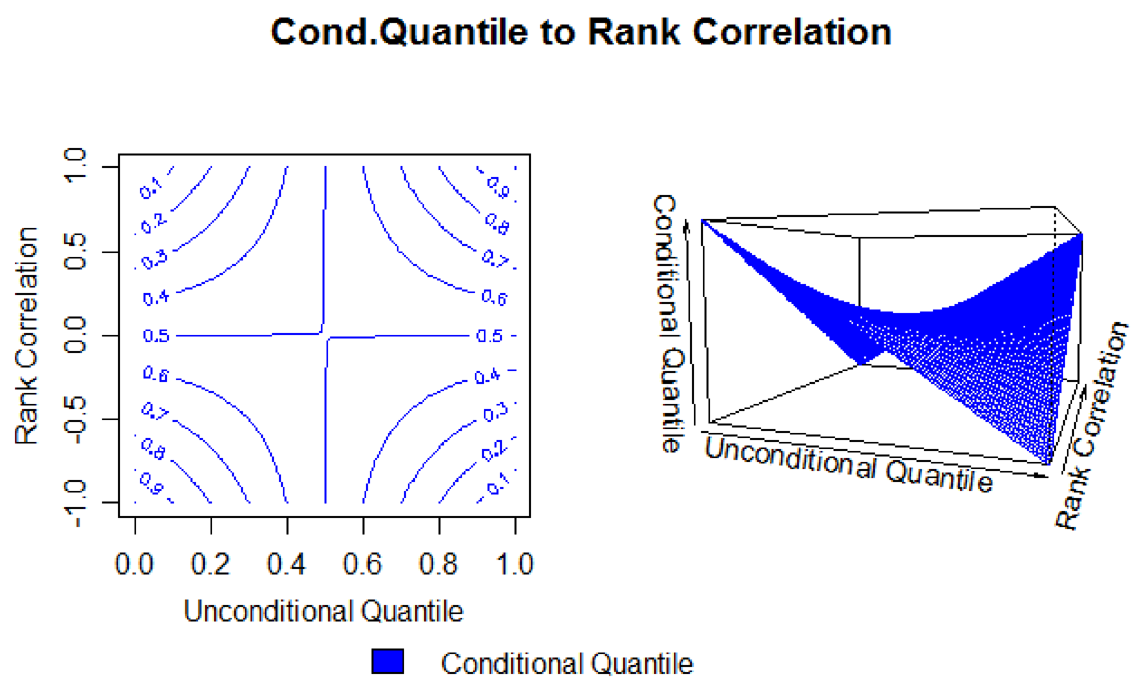
Example BN for Cond. Exceedance Probability:



Back

Appendix (3):

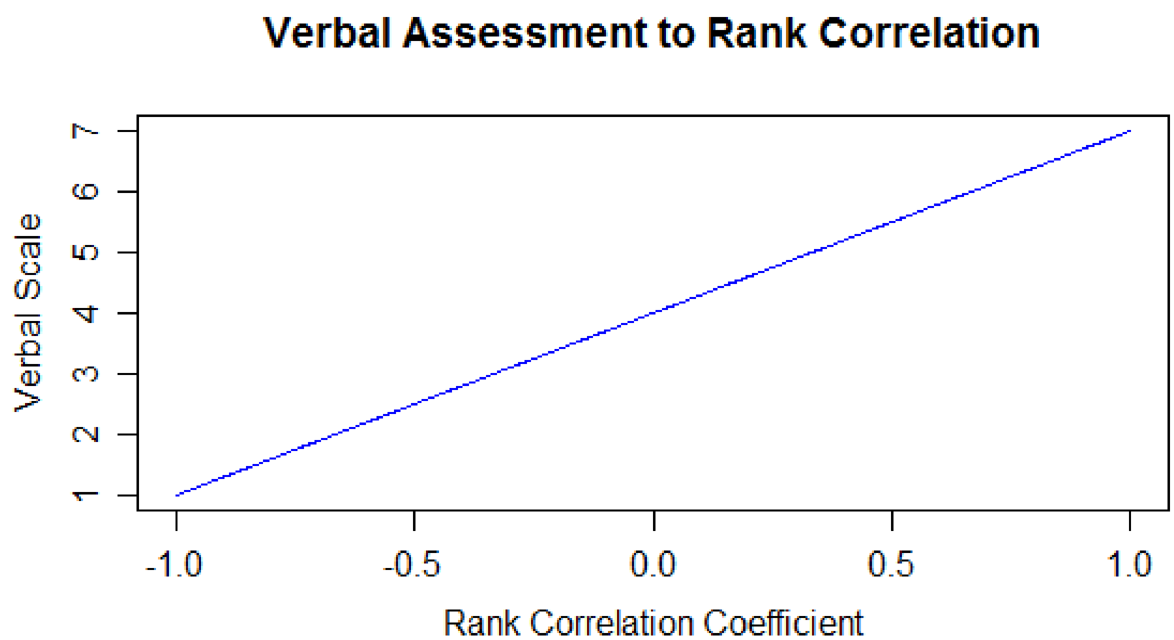
Relation of Conditional Expectation to Rank Correlation:



Back

Appendix (4):

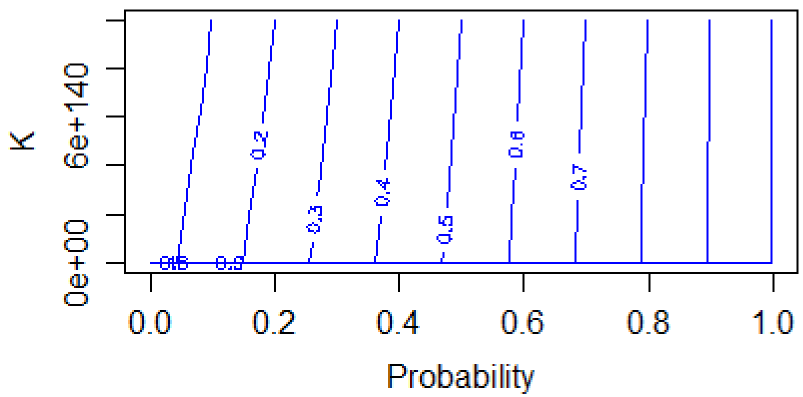
Relation of Verbal Assessment to Rank Correlation:

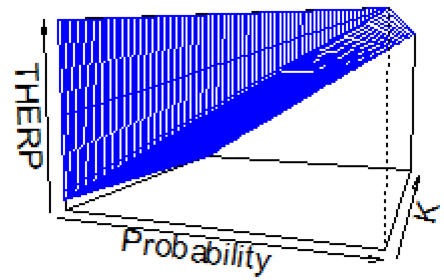


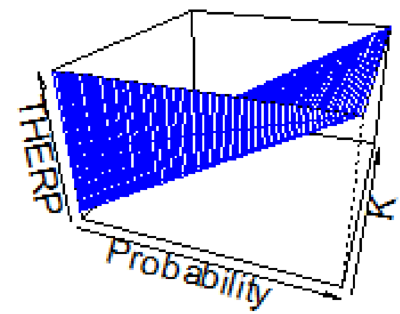
[Back](#)

Appendix (5):

Relation of Verbal THERP Assessment to Rank Correlation:







Back