expert: Modeling Without Data Using Expert Opinion

by Vincent Goulet, Michel Jacques and Mathieu Pigeon

Introduction

The **expert** package provides tools to create and manipulate empirical statistical models using expert opinion (or judgment). Here, the latter expression refers to a specific body of techniques to elicit the distribution of a random variable when data is scarce or unavailable. Opinions on the quantiles of the distribution are sought from experts in the field and aggregated into a final estimate. The package supports aggregation by means of the Cooke, Mendel–Sheridan and predefined weights models.

We do not mean to give a complete introduction to the theory and practice of expert opinion elicitation in this paper. However, for the sake of completeness and to assist the casual reader, the next section summarizes the main ideas and concepts.

It should be noted that we are only interested, here, in the mathematical techniques of expert opinion aggregation. Obtaining the opinion from the experts is an entirely different task; see Kadane and Wolfson (1998); Kadane and Winkler (1988); Tversky and Kahneman (1974) for more information. Moreover, we do not discuss behavioral models (see Ouchi, 2004, for an exhaustive review) nor the problems of expert selection, design and conducting of interviews. We refer the interested reader to O'Hagan et al. (2006) and Cooke (1991) for details. Although it is extremely important to carefully examine these considerations if expert opinion is to be useful, we assume that these questions have been solved previously. The package takes the opinion of experts as an input that we take here as available.

The other main section presents the features of version 1.0-0 of package **expert**.

Expert opinion

Consider the task of modeling the distribution of a quantity for which data is very scarce (e.g. the cost of a nuclear accident) or even unavailable (e.g. settlements in liability insurance, usually reached out of court and kept secret). A natural option for the analyst[1] is then to ask experts their opinion on the distribution of the so-called *decision variable*. The experts are people who have a good knowledge of the field under study and who are able to express their opinion in a simple probabilistic fashion. They can be engineers, physicists, lawyers, actuaries, etc. Typically,

a group of 12 to 15 experts is consulted.

The experts express their opinion on the distribution of the decision random variable by means of a few quantiles, usually the median and a symmetric interval around it. Let us recall that the $100rth$ quantile of a random variable is the value q_r such that

$$\lim_{h \to 0} F(q_r - h) \le r \le F(q_r),$$

where F is the cumulative distribution function of the random variable. Hence, by giving the quantiles q_u and q_v, $u < v$, an expert states that, in his opinion, there is a probability of $v - u$ that the true value of the variable lies between q_u and q_v. We generally avoid asking for quantiles far in the tails (say, below 5% and above 95%) since humans have difficulty evaluating them.

In addition to the distribution of the decision variable, the experts are also queried for the distribution of a series of *seed* (or calibration) variables for which only the analyst knows the true values. By comparing the distributions given by the experts to the latter, the analyst will be able to assess the quality of the information provided by each expert. This step is called *calibration*.

In summary, expert opinion is given by means of quantiles for k seed variables and one decision variable. The calibration phase determines the influence of each expert on the final *aggregated distribution*. The three methods discussed in this paper and supported by the package are simply different ways to aggregate the information provided by the experts.

The aggregated distribution in the classical model of Cooke (1991) is a convex combination of the quantiles given by the experts, with weights obtained from the calibration phase. Consequently, if we asked for three quantiles from the experts (say the 10*th*, 50*th* and 90*th*), the aggregated distribution will consist of three "average" quantiles corresponding to these same probabilities. This model has been used in many real life studies by the analysts of the Delft Institute of Applied Mathematics of the Delft University of Technology (Cooke and Goossens, 2008).

The predefined weights method is similar, except that the weights are provided by the analyst.

On the other hand, the model of Mendel and Sheridan (1989) adjusts the probabilities associated with each quantile by a (somewhat convoluted) Bayesian procedure to reflect the results of the calibration phase. Returning to the context above, this means that the probability of 10% associated with the first quantile given by an expert may be changed to, say, 80% if the calibration phase proved that this ex-

---

[1]Also called *decision maker* in the literature.

pert systematically overestimates the distributions. Combining intersecting adjusted probabilities from all the experts usually results in an aggregated distribution with far more quantiles than were asked for in the beginning.

It is well beyond the scope of this paper to discuss the actual assumptions and computations in the calibration and aggregation phases. The interested reader may consult the aforementioned references, Goulet et al. (2008) or Pigeon (2008, in French).

The expert package

To the best of our knowledge, the only readily available software for expert opinion calculations is Excalibur, a closed source, Windows-only application available from the Risk and Environmental Modelling website of the Delft University of Technology (http://dutiosc.twi.tudelft.nl/~risk/). Excalibur does not support the Mendel–Sheridan model. Data has to be entered manually in the application, making connection to other tools inconvenient. Furthermore, exporting the results for further analysis requires an ad hoc procedure.

expert is a small R package providing a unified interface to the three expert opinion models exposed above along with a few utility functions and methods to manipulate the results. Naturally, the analyst also benefits from R's rich array of statistical, graphical, import and export tools.

Example dataset

Before looking at the functions of the package, we introduce a dataset that will be used in forthcoming examples. We consider an analysis where three experts must express their opinion on the $10th$, $50th$ and $90th$ quantiles of a decision variable and two seed variables. The true values of the seed variables are 0.27 and 210,000, respectively. See Table 1 for the quantiles given by the experts.

Main function

The main function of the package is `expert`. It serves as a unified interface to all three models supported by the package. The arguments of the function are the following:

1. `x`, the quantiles of the experts for all seed variables and the decision variable — in other words the content of Table 1. This is given as a list of lists, one for each expert. The interior lists each contain vectors of quantiles: one per seed variable and one for the decision variable (in this order). For the example above, this gives

```
> x <- list(
+   E1 <- list(
+     S1 <- c(0.14, 0.22, 0.28),
+     S2 <- c(130000, 150000, 200000),
+     X  <- c(350000, 400000, 525000)),
+   E2 <- list(
+     S1 <- c(0.2, 0.3, 0.4),
+     S2 <- c(165000, 205000, 250000),
+     X  <- c(550000, 600000, 650000)),
+   E3 <- list(
+     S1 <- c(0.2, 0.4, 0.52),
+     S2 <- c(200000, 400000, 500000),
+     X  <- c(625000, 700000, 800000)))
```

2. `method`, the model to be used for aggregation. Must be one of 'cooke' for Cooke's classical model, 'ms' for the Mendel–Sheridan model, or 'weights' for the predefined weights model;

3. `probs`, the vector of probabilities corresponding to the quantiles given by the experts. In the example, we have

```
> probs <- c(0.1, 0.5, 0.9)
```

4. `true.seed`, the vector of true values of the seed variables. Obviously, these must be listed in the same order as the seed variables are listed in argument `x`. In the example, this is

```
> true.seed <- c(0.27, 210000)
```

5. `alpha`, parameter α in Cooke's model. This argument is ignored in the other models. Parameter α sets a threshold for the calibration component of an expert below which the expert automatically receives a weight of 0 in the aggregated distribution. If the argument is missing or `NULL`, the value of α is determined by an optimization procedure (see below);

6. `w`, the vector of weights in the predefined weights model. This argument is ignored in the other models. If the argument is missing or `NULL`, the weights are all equal to $1/n$, where n is the number of experts.

If the calibration threshold α in Cooke's model is not fixed by the analyst, one is computed following a procedure proposed by Cooke (1991). We first fit the model with $\alpha = 0$. This gives a weight to each expert. Using these weights, we then create a "virtual" expert by aggregating the quantiles for the seed variables. The optimal α is the value yielding the largest weight for the virtual expert. This is determined by trying all values of α just below the calibration components obtained in the first step. There is no need to try other values and, hence, to conduct a systematic numerical optimization.

When the experts do not provide the $0th$ and $100th$ quantiles, `expert` sets them automatically following the ad hoc procedure exposed in Cooke (1991). In a nutshell, the quantiles are obtained by

| | | Variable | | |
Expert	Quantile	Seed 1	Seed 2	Decision
1	10*th*	0.14	130,000	350,000
	50*th*	0.22	150,000	400,000
	90*th*	0.28	200,000	525,000
2	10*th*	0.20	165,000	550,000
	50*th*	0.30	205,000	600,000
	90*th*	0.40	250,000	650,000
3	10*th*	0.20	200,000	625,000
	50*th*	0.40	400,000	700,000
	90*th*	0.52	500,000	800,000

Table 1: Quantiles given by the experts for the seed and decision variables

removing and adding 10% of the smallest interval containing all quantiles given by the experts to the bounds of this interval. Therefore, the minimum and maximum of the distribution are set the same for all experts. In our example, the 0*th* and 100*th* quantiles of the first seed variable are set, respectively, to

$$q_0 = 0.14 - 0.1(0.52 - 0.14) = 0.102$$
$$q_1 = 0.52 + 0.1(0.52 - 0.14) = 0.558,$$

those of the second seed variable to

$$q_0 = 130,000 - 0.1(500,000 - 130,000) = 93,000$$
$$q_1 = 500,000 + 0.1(500,000 - 130,000) = 537,000$$

and, finally, those of the decision variable to

$$q_0 = 350,000 - 0.1(800,000 - 350,000) = 305,000$$
$$q_1 = 800,000 + 0.1(800,000 - 350,000) = 845,000.$$

Note that only the Mendel–Sheridan model allows non-finite lower and upper bounds.

The function `expert` returns a list of class 'expert' containing the vector of the knots (or bounds of the intervals) of the aggregated distribution, the vector of corresponding probability masses and some other characteristics of the model used. A `print` method displays the results neatly.

Consider using Cooke's model with the data of the example and a threshold of $\alpha = 0.3$. We obtain:

```
> expert(x, "cooke", probs, true.seed,
+       alpha = 0.03)

Aggregated Distribution Using Cooke Model

        Interval Probability
(305000, 512931]         0.1
(512931, 563423]         0.4
(563423, 628864]         0.4
(628864, 845000]         0.1

 Alpha: 0.03
```

As previously explained, if the value of α is not specified in the call to `expert`, its value is obtained by optimization and the resulting aggregated distribution is different:

```
> expert(x, "cooke", probs, true.seed)

Aggregated Distribution Using Cooke Model

        Interval Probability
(305000, 550000]         0.1
(550000, 600000]         0.4
(600000, 650000]         0.4
(650000, 845000]         0.1

 Alpha: 0.3448
```

The Mendel–Sheridan model yields an entirely different result that retains many of the quantiles given by the experts. The result is a more detailed distribution. Here, we store the result in an object `fit` for later use:

```
> fit <- expert(x, "ms", probs, true.seed)
> fit

Aggregated Distribution Using
Mendel-Sheridan Model

        Interval Probability
(305000, 350000]     0.01726
(350000, 400000]     0.06864
(400000, 525000]     0.06864
(525000, 550000]     0.01726
(550000, 600000]     0.06864
(600000, 625000]     0.06864
(625000, 650000]     0.53636
(650000, 700000]     0.06864
(700000, 800000]     0.06864
(800000, 845000]     0.01726
```

Utility functions and methods

In addition to the main function `expert`, the package features a few functions and methods to ease manipulation and analysis of the aggregated distribution. Most are inspired by similar functions in package **actuar** (Dutang et al., 2008).

First, a `summary` method for objects of class 'expert' provides more details on the fitted model

than the `print` method, namely the number of experts, the number of seed variables and the requested and set quantiles:

```
> summary(fit)

Call:
expert(x = x, method = "ms", probs = probs,
       true.seed = true.seed)

Aggregated Distribution Using
Mendel-Sheridan Model

          Interval Probability
 (305000, 350000]     0.01726
 (350000, 400000]     0.06864
 (400000, 525000]     0.06864
 (525000, 550000]     0.01726
 (550000, 600000]     0.06864
 (600000, 625000]     0.06864
 (625000, 650000]     0.53636
 (650000, 700000]     0.06864
 (700000, 800000]     0.06864
 (800000, 845000]     0.01726

 Number of experts: 3,
 Number of seed variables: 2
 Quantiles: 0*, 0.1, 0.5, 0.9, 1*
   (* were set)
```

Methods of `mean` and `quantile` give easy access to the mean and to the quantiles of the aggregated distribution:

```
> mean(fit)

[1] 607875

> quantile(fit)

    0%    25%    50%    75%   100%
305000 600000 625000 625000 845000
```

Of course, one may also specify one or many quantiles in a second argument:

```
> quantile(fit, c(0.1, 0.9))

   10%    90%
400000 650000
```

Moreover, the `quantile` method can return linearly interpolated quantiles by setting the `smooth` argument to `TRUE`:

```
> quantile(fit, smooth = TRUE)

    0%    25%    50%    75%   100%
305000 603478 633898 645551 845000
```

The inverse of the quantile and smoothed quantile functions are, respectively, the cumulative distribution function (cdf) and the ogive (Klugman et al., 1998; Goulet and Pigeon, 2008). In a fashion similar to `ecdf` of the base R package **stats**, the `cdf` function returns a function object to compute the cdf of the aggregated distribution at any point:

```
> F <- cdf(fit)
> knots(F)

 [1] 305000 350000 400000 525000 550000
 [6] 600000 625000 650000 700000 800000
[11] 845000

> F(knots(F))

 [1] 0.00000 0.01726 0.08590 0.15455
 [5] 0.17181 0.24045 0.30909 0.84545
 [9] 0.91410 0.98274 1.00000

> F(c(450000, 750000))

[1] 0.0859 0.9141
```

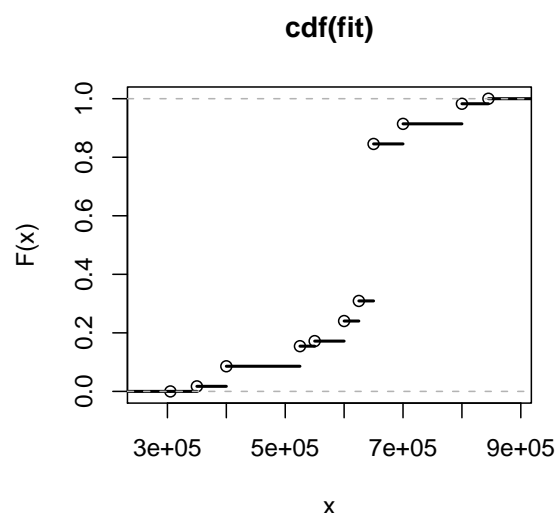A `plot` method yields the graphic in Figure 1:

```
> plot(F)
```


Figure 1: Cumulative distribution function of the aggregated distribution

Function `ogive` is the analogue of `cdf` for the linearly interpolated cdf:

```
> G <- ogive(fit)
> G(knots(G))

 [1] 0.00000 0.01726 0.08590 0.15455
 [5] 0.17181 0.24045 0.30909 0.84545
 [9] 0.91410 0.98274 1.00000

> G(c(450000, 750000))

[1] 0.1134 0.9484

> plot(G)
```

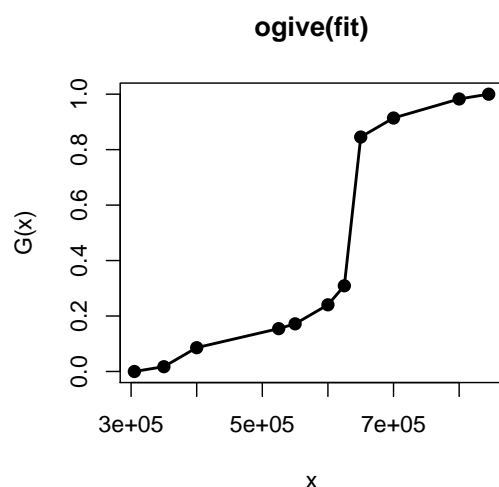See Figure 2 for the graphic created with the last expression.

Figure 2: Ogive of the aggregated distribution

Finally, a method of `hist` draws the derivative of the ogive, that is the histogram. The arguments of the method are the same as those of the default method in package **stats**. See Figure 3 for the graphic created by the following expression:
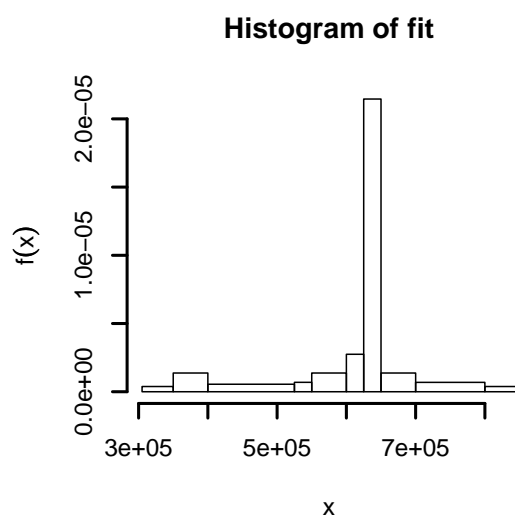
```
> hist(fit)
```


Figure 3: Histogram of the aggregated distribution

Further model fitting

Depending on the application, the aggregated distribution one obtains from `expert` might be too crude for risk analysis, especially if one used Cooke's model with very few quantiles. Now, the expert aggregated distribution actually plays the role of the empirical distribution in usual statistical modeling.

That means one can further fit a continuous or discrete distribution to the aggregated "data".

For example, consider the task of fitting a gamma distribution to the Mendel–Sheridan aggregated distribution obtained above. Among many other options (see, e.g., Klugman et al., 1998; Goulet and Pigeon, 2008), we choose our parameters to minimize the Cramér-von Mises distance

$$d(\alpha, \lambda) = \sum_{i=1}^{n} (G(x_i) - F(x_i; \alpha, \lambda))^2,$$

where $G(\cdot)$ is the ogive of the aggregated distribution, $x_1, \ldots, x_n$ are its knots and $F(\cdot; \alpha, \lambda)$ is the cdf of a gamma distribution with shape parameter α and rate parameter λ. The optimal parameters are simple to obtain with the function `optim`:

```
> f <- function(p, x) drop(crossprod(G(x) -
+     pgamma(x, p[1], p[2])))
> optim(c(100, 0.00016), f,
        x = knots(G))$par

[1] 4.233e+02 6.716e-04
```

Conclusion

The paper introduced a package to elicit statistical distributions from expert opinion in situations where real data is scarce or absent. The main function, `expert`, is a unified interface to the two most important models in the field — the Cooke classical model and the Mendel–Sheridan Bayesian model — as well as the informal predefined weights model. The package also provides functions and methods to compute from the object returned by `expert` and plot important results. Most notably, `cdf`, `ogive` and `quantile` give access to the cumulative distribution function of the aggregated distribution, its ogive and the inverse of the cdf and ogive.

For modeling purposes, the package can be used as a direct replacement of Excalibur. However, coupled with the outstanding statistical and graphical capabilities of R, we feel **expert** goes far beyond Excalibur in terms of usability and interoperability.

Acknowledgments

Bibliography

R. Cooke. *Experts in Uncertainty*. Oxford University Press, 1991.

R. Cooke and L. Goossens. TU Delft expert judgment data base. *Reliability Engineering and System Safety*, 93(5):657–674, 2008.

C. Dutang, V. Goulet, and M. Pigeon. **actuar**: An R package for actuarial science. *Journal of Statistical Software*, 25(7), 2008. URL http://www.jstatsoft.org/v25/i07.

V. Goulet and M. Pigeon. Statistical modeling of loss distributions using **actuar**. *R News*, 8(1):34–40, May 2008. URL http://cran.r-project.org/doc/Rnews/Rnews_2008-1.pdf.

V. Goulet, M. Jacques, and M. Pigeon. Modeling of data using expert opinion: An actuarial perspective. 2008. In preparation.

J. B. Kadane and R. L. Winkler. Separating probability elicitation from utilities. *Journal of the American Statistical Association*, 83(402):357–363, 1988.

J. B. Kadane and L. J. Wolfson. Experiences in elicitation. *The Statistician*, 47(1):3–19, 1998.

S. A. Klugman, H. H. Panjer, and G. Willmot. *Loss Models: From Data to Decisions*. Wiley, New York, 1998. ISBN 0-4712388-4-8.

M. Mendel and T. Sheridan. Filtering information from human experts. *IEEE Transactions on Systems, Man and Cybernetics*, 36(1):6–16, 1989.

A. O'Hagan, C. Buck, C. Daneshkhah, J. Eiser, P. Garthwaite, D. Jenkinson, J. Oakley, and T. Rakow. *Uncertain Judgements*. Wiley, 2006.

F. Ouchi. A literature review on the use of expert opinion in probabilistic risk analysis. Policy Research Working Paper 3201, World Bank, 2004. URL http://econ.worldbank.org/.

M. Pigeon. Utilisation d'avis d'experts en actuariat. Master's thesis, Université Laval, 2008.

A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. *Science, New Series*, 185(4157):1124–1131, 1974.

Vincent Goulet
École d'actuariat
Université Laval
Québec, Canada
vincent.goulet@act.ulaval.ca

Michel Jacques
École d'actuariat
Université Laval
Québec, Canada
michel.jacques@act.ulaval.ca

Mathieu Pigeon
Institut de Statistique
Université Catholique de Louvain
Louvain-la-Neuve, Belgium
mathieu.pigeon@uclouvain.be