COMBINING EXPERT OPINIONS FOR DECISION MAKING

E Rosenblueth (I)

SUMMARY

Some uncertainty is inevitable about the validity of all opinons. The treatment that each opinon deserves depends on our prior probability about the phenomena of interest and about the bias and expertise of the corresponding participant. These, together with the opinions stated, furnish posterior probabilities through application of Bayes' theorem. The results allow updating the credibilities of the participants. The same approach holds for calculating the parameters of probability distributions as well as the probabilities of utilities. The approach is applied to several problems in earthquake engineering.

INTRODUCTION

Rational decision making involves assigning probabilities and utilities and choosing the option with maximum utility. A set of independent opinions is never less valuable than any member of the set. (This is not necessarily true of statements by committee members.) To get their full value the statements must be properly combined. Intuitive treatments, such as weighted averages — whether outliers are discarded or not — can seriously mislead. This paper develops bayesian methods for combining independent statements. (A precedent is found in ref 1.)

We work out two kinds of formulation. One applies to qualitative statements of opinions about variables that can only take up a small number of values; the second one applies to probability distributions and can be adapted to a treatment of probabilities or utilities. These methods are well suited for use to conjunction with some informal techniques, such as Delphi, in which the interaction of experts is conveniently dealt with.

The methods are applicable to opinions in the cootext of earthquake engineering, especially in the realm of variables that go into public policy making. They are illustrated accordingly.

PROBLEM FORMULATION FOR INDEPENDENT STATEMENTS

Let H_k, $k = 1, 2, \ldots, N$ be exhaustive, mutually exclusive hypotheses. Out of n witnesses, n_j state that H_j is true and $n-n_j$ that it is not, so $\sum_{j=1}^{N} n_j = n$. We seek to compute the posterior probability P''_k that each of the H_k be true given the prior probabilities P'_k and statements which may be true, erroneous, mistaken, or insincere. We assume that the statements are mutually independent except where defined otherwise.

(I) Instituto de Ingeniería, Universidad Nacional Autónoma de México; Visiting Professor, University of Waterloo.

SOLUTION

Let Q_{ijk} be the probability that witness i says that hypothesis H_j is true when H_k is true and let prime and double prime denote quantities assigned before and after knowledge of the statements respectively. From Bayes' theorem (concepts of probability and bayesian statistics are found in any standard text on the subject, such as ref 2) we find

$$P''_k = \frac{P''_k \prod_{i=1}^{n} Q'_{ijk}}{\sum_{\ell'=1}^{N} P'_\ell \prod_{i=1}^{n} Q'_{ij\ell}} \qquad (1)$$

If we wish to update Q_{ijk} we must recognize at least two situations. In one, for a given subject, Q_{ijk} and Q_{ijh} are independent of each other if $k \neq h$. In the second situation Q_{ijh} — the probability that the subject be truthful — is independent of j. In the first situation

$$Q''_{ijk} = Q'_{ijk} + (1 - Q'_{ijk})P''_k/(n'_{ik} + 1) \qquad (2)$$

In the second situation we write Q_i for Q_{ijj} and n'_i for n'_{ik} and we obtain

$$Q''_i = (n'_i Q'_i + P''_j)/(n'_i + 1) \qquad (3)$$

where j corresponds to the hypothesis which the subject said was true.

Parameter n'_{ik} measures our prior intensity of conviction. It admits a heuristic interpretation in terms of the following fictitious experiment, prior to the real experiment. We know that H_k is true. We ask subject i n'_{ik} times which is the correct hypothesis. On $n'_{ij}Q'_{ijk}$ occasions he answers H_j. Now we perform the real experiment, knowing that the correct hypothesis is again H_k, and again he answers H_j. Since we deal with a Bernoulli process the new expected fraction of the number of times in which i will say H_j will be $Q''_{ijk} = (n'_{ik}Q'_{ijk} + 1)/(n'_{jk} + 1)$. If however, H_k is not true in the real experiment, we have no basis in the first situation for updating and distribution and so $Q''_{ijk} = Q'_{ijk}$. The probability that H_k be true in the real experiment is P''_k while it is $1 - P''_k$ that it be false. Multiplying the first expression for Q''_{ijk} by P''_k and the second one by $1 - P''_k$ and adding we arrive at eq 2. Notice that subscript j corresponds to the hypothesis H_j that each subject claimed was true when H_k was true. In the second situation, if H_k were false the updated Q_i would be $n_i/Q'_i/n_i + 1)$; the probability that H_k be false is $1 - P''_k$; hence eq 3.

PROBABILITY DISTRIBUTIONS

Assume that n witnesses report the values x_i, $i = 1, 2, \ldots, n$ for random variable X, which can be the measured or estimated value of a quantity x. Each reported value can be true (valid) or the result of a deceit, error, mistake, or bias. Every x_i is the value of a specimen of a variable having some probability of possessing each of two or more probability distributions. Suppose there are only two populations: that of the true values, Y and that of the false ones, Z and that the x_i are statistically independent. Let $Q_{i_1 \ldots i_k}$ denote the probability that $x_{i_1}, x_{i_2}, \ldots, x_{i_k} \varepsilon Y$, $p_Y(\cdot)$ denote the probability density functions of Y and Z, and $P_Y(\cdot)$ and $P_Z(\cdot)$ denote the corresponding distribution functions. From Bayes' theorem

$$P''_Y(x) = \sum_{k=0}^{n} \Sigma Q''_{i_1 \ldots i_k} P''_Y(x | x_{i_1} \ldots x_{i_k}) \quad (4)$$

$$Q''_{i_1 \ldots i_k} = Q'_{i_1} \ldots Q'_{i_k} (1 - Q'_{i_{k+1}}) \ldots (1 - Q'_{i_n}) p'_Y(x_{i_k} \ldots x_{i_k}) p'_Z(x_{i_{k+1}} \ldots x_{i_n})/D$$

$$p'_Y(x_{i_1} \ldots x_{i_k}) = p''_Y(x_{i_1} | x_{i_2} x_{i_3} \ldots x_{i_k}) p''_Y(x_{i_2} | x_{i_3} \ldots x_{i_k}) \ldots p''_Y(x_{i_{k-1}} | x_{i_k}) p'_Y(x_{i_k})$$

and there is a similar expression for $p'_Z(x_{i_{k+1}} \ldots x_{i_n})$. The second sum in eq 4 covers all combinations of n data taken k at a time. D is the sum of all terms having the form of the numerator in the expression for $Q''_{i_1 \ldots i_k}$. The posterior distribution of true values is thus a linear combination of 2^n posterior distributions, each obtained from Bayes' theorem.

We may face one of two important situations. Either Y actually varies from one specimen to another and there is no uniquely true value (as with soil properties say, which vary from point to point within a soil formation) or there is such a unique value and the difference between one true x_i and that value is due to experimental or observational error (as with reported earthquake magnitudes: there is only one true magnitude but each station that reports will give a different magnitude). Errors contained in the true x_i do not include "human" or "gross errors" or mistakes. (There are other situations of practical interest in which we will not delve. For instance the actual value of a material property can vary from one specimen to another and the spread in valid data can reflect this variation plus the one due to experimental errors.) In the first case we are often interested in the distribution of Y, but when there is a uniquely true value, only <u>its</u> distribution is of interest and it can be obtained from relations between this value and the parameters of $P''_Y(\cdot)$ For design purposes it may be unnecessary to deal with $P''_Y(\cdot)$ directly as it is often some property of this distribution (such as a utility obtained from the convolution of the probability distribution of Y and that of another variable) that is of interest and can be obtained as a linear combination of the values corresponding to the conditional distributions $P''_Y(x | x_{i_1} \ldots x_{i_k})$.

When n = 1 eq 4 takes a simple form:

$$P''_Y(x) = Q''_1 P''_Y(x | x_1) + (1 - Q''_1) P'_Y(x) \quad (5)$$

$$Q''_1 = \left[1 + (1/Q'_1 - 1) p'_Z(x_1)/p'_Y(x_1)\right]^{-1}$$

The case in which Y and Z are gaussian deserves special attention because situations abound in which this assumption is reasonable and because every continuous variable having a continuous distribution characterized by no more than two parameters can be transformed into a gaussian variable. Suppose then that Y and Z are gaussian and, further, that they have unknown process means and known process variances. Then using the natural conjugate distribution of the mean[2] we can write, for the process $Y \overset{d}{=} N(\mu_Y, \sigma_Y^2)$ and for the prior, $\mu_Y = N(m_Y', \sigma_Y^2/n_Y')$ where n_Y' is a parameter that measures our prior certainty about m_Y' and heuristically can be regarded as the number of specimens in a fictitious prior experiment. The prior bayesian distribution of Y is described by $Y' \overset{d}{=} N[m_Y', \sigma_Y^2(1 + 1/n_Y')]$ and its posterior bayesian distribution by $Y'' \overset{d}{=} N[m_Y'', \sigma_Y^2(1 + 1/n_Y'')]$ where $m_Y'' = (n_Y' m_Y' + n_Y m_Y)/n_Y''$, $n_Y'' = n_Y' + n_Y$, m_Y = sample average, and n_Y = number of data in the sample (in our case $n_Y = k$). Similar expressions hold for Z ($n_Z = n - k$).

Often when there is a unique true value of Y this is the process mean. Its posterior distribution is described by $\mu_Y'' \overset{d}{=} N(m_Y'', \sigma_Y^2/n_Y'')$.

When $n > 0$ and $Q_i' < 1$ for at least one i, $P_Y''(x)$ cannot, in general, be normal, nor can in general the posterior distribution of μ_Y. Each, however, is the linear combination of normal distributions and the pertinent properties of Y and of μ_Y can be obtained with as much simplicity for each term as though the problem of credibility were absent.

The problem can be modified so that each witness has some probability of belonging to one of several groups, and each group its own tendency to under or overestimate the variable of interest. The expressions required for computing posterior expectations and variances are readily derived.

APPLICATIONS

Example 1.1 We will try to assess the intensity of a given earthquake at a certain site on the basis of an expert witness' account. From previous experience with earthquakes of the magnitude in question at sites of similar ground conditions we are sure that the intensity in which we are interested is either V or VI. We thus have $N = 2$, $n = 1$. We will assume that Q_{ijk} depends on k and that Q_{ijh} is independent of Q_{ijk} if $h \neq k$.

Before sending the expert to the site we have fictitiously calibrated him (we have looked at his record of assessments and at those of specialists having a comparable experience) as follows: we show him eight sites where the intensity was V; we record that he correctly classifies the intensity as V on six occasions and he states that it is VI in the other two. Next we take him to two sites where the intensity was VI; we note that he estimates the intensity as V at one site and VI at the second one. Hence $n_{11}' = 8$, $n_{12}' = 2$, $Q_{111}' = 0.75$, $Q_{122}' = 0.5$. Using eq 1 and $P_1'' + P_2'' = 1$ we get

$$Q_{i11}'' = 0.750 \div 0.028 P_1'' \; , \quad Q_{i21}'' = 1 - Q_{i11}''$$

$$Q_{i12}'' = 0.500 - 0.167 P_1'' \; , \quad Q_{i22}'' = 1 - Q_{i12}''$$

The expert says that the intensity was V. If experience with similar seismic and site conditions points to equal probabilities that the intensity be V or VI we have $P_1' = 0.5$. Replacing P_1', Q_{i11}', and Q_{i112}' in eq 1 we get $P_1'' = 0.600$. On the other hand had we begun with $P_1' = 0.667$ for intensity V we would have arrived at $P_1'' = 0.750$ as the probability that the intensity at the site of interest was V. According to the foregoing expression for the updated Q_{ijk}s, if $P_1' = 0.5$ the expert now has a probability of 0.767 of stating that an earthquake of intensity V has indeed this intensity, of 0.233 of assigning it intensity VI, and a probability of 0.600 of assigning intensity VI to an earthquake of intensity VI. if $P_1' =$ 0.667 these probabilities are respectively 0.771, 0.229, and 0.650.

Example 1.2 Using the data in the foregoing imaginary experiment let us solve the same example but assuming that there is no bias on the part of the witness, so $Q_1' = Q_{111}' = Q_{122}'$ is the prior probability that the assessment is correct whatever the intensity. In this case $n_1' = 8 + 2 = 10$, $Q_1' = (6 + 1)/10 = 0.7$. Now $Q_1'' = (7 + P_1''/11)$. From eq 1 the assumption $P_1' = 0.5$ gives $P_1'' = 0.700$ while $P_1' = 0.667$ yields $P_1'' = 0.824$. We notice a significant increase with respect to example 1.1 in the probability that the site intensity was V. This we can attribute to a decrease in the probability that the assessment was incorrect. The updated credibility we would assign the subject is $Q_1'' = 0.700$ for $P_1' = 0.5$, and 0.711 when $P_1' =$ 0.667.

Example 2.1 One witness states that the intensity was V and all the rest that it was VI; N = 2. We take $n_1 = 1$ and $n_2 = n - 1$. Suppose that Q_{ijk} and Q_{ijh} are independent of each other if $k \neq h$, that for all i $Q_{i11}' = Q_{i22} = 0.9$, $Q_{i12}' = 0.1$ (that is, all witnesses are equally trustworthy and the likelihood that any one of them makes a correct statement is nine times that of an incorrect statement), that $n_{i1}' = n_{i2}'$ is independent of i (we are equally sure of our assessments about all the witnesses' expertise), and that $P_1' = P_2'$ (we are completely ignorant as to whether the intensity was V or VI). The solution is straightforward using eqs 1 and 2. Results are depicted in the first of each pair of lines of table 1. We can appreciate the sensitivity of P_1'' to n and that of Q_{i11}'' to n_{i1}' and to n. With n = 2 the witness accounts are not informative in this case because we have taken $P_1' = P_2'$ and $Q_{i11}' = Q_{i22}'$. Had a single subject provided all the statements, each one corresponding, say, to his examination of one square block in the city, we would have updated Q_{ijk} using eq 5.

Example 2.2 Take the data and prior convictions in example 2.1 but suppose that Q_{ijj} is independent of j. Applying eq 3 we find the updated values in the second of each pair of lines in table 1. Differences with results in example 2.1 are spectacular for large n and small n_{ik}' or n_i'. If one subject had made all the assessments we would have used eq 6 to update Q_1.

Example 3.1 Analysis and measurement of the fundamental period of vibration of several nominally identical structures has led to a prior gaussian probability distribution of the variable such that $m_Y' = 1$ s, $\sigma_Y = .1$ s,

$n_Y = 4$. One more structure of this type is tested and a period T_1 is reported. There is doubt as to the validity of this report but essentially none about the previous measurements. The following parameters are assigned, $m'_Z = 1$ s, $\sigma_Z^2(1 + 1/n'_Z)/\sigma_Y^2(1 + 1/n'_Y) = 2$. Bayesian distributions of the fundamental period appear in fig 1 on normal probability paper.

Example 3.2 Data in example 3.1 refer now in appropriate units to the strength of a certain type of structural member. Find the design strengths associated with different probabilities that they not be met. Results in fig 2 were obtained from fig 1. Beyond some values of T_1 a decrease in this value brings about an increase in the design strength. The paradox is due to a decrease in the probability that T_1 be valid. Suppose now we are asked to find the optimum central safety factor for members of this type under the assumption that if through design we multiply the expected strength ET by some factor α then αT will have the distribution that T had; we are given the probability distribution of the load and its recurrence periods as well as the disutility associated with the initial cost as a function of α and that of the loss in case of failure. Under the assumption that utilities are additive we compute the present value of the disutility as a function of α_1 first as though T_1 were a false value and multiply it by $1 - Q''_1$, and then as though it were true and multiply this by $-Q''_1$. Finally we add these two to the disutility associated with the initial cost and we find that α which minimizes the total.

Example 3.3 Four stations report the magnitude of an earthquake as x_i = 6.2, 6.2, 6.8, and 7.2. From the area over which the phenomenon was felt, reported intensities, and other indirect evidence we expected a magnitude of 6.4. We assign this prior magnitude a gaussian distribution with standard deviation of 4.4 and assume that stations which correctly interpret their records have a standard deviation of .2 relative to the true value. Thus $m'_Y = 6.4$, $\sigma_Y = .2$, $n'_Y = .25$. We further take $Q'_i = .9$ for all i, $m'_Z = m'_Y$, $\sigma_Z = 2\sigma_Y$, $n'_Z = n'_Y$. From eq 4 we get 16 gaussian distributions whose linear combination furnishes the posterior of the magnitude, taken to be the expected μ_Y of the true reported values Y. The fact that $x_1 = x_2$ reduces the number of distributions to 12. Table 2 displays values of $Q''_{i_1 \ldots i_k}$, of the expected μ_Y, which is m''_Y and of the standard deviation of μ_Y, which is $\sigma_Y/\sqrt{n''_Y}$. The probabilities $Q''_{i_1 \ldots i_k}$ for the three most probable distributions add up to 99.1. The case in which only x_4 were true also merits consideration because of its large m''_Y and $\sigma_Y/\sqrt{n''_Y}$, save when very low probabilities of exceedance were of interest.

ACKNOWLEDGEMENTS

I am indebted to Luis Esteva for his critical constructive comments and to Carlos Ferregut for computational support.

REFERENCES

1. Pearl, J, "A framework for processing value judgements", IEEE Transactions on Systems, Man and Cybernetics, SMC-7, 5, 1977.5, 349-55.

2. Larson, H L, Introduction to probability theory and statistical inference, 2nd ed, John Wiley & Sons, New York, 1974.

3. Rosenblueth, E, "Procesamiento de información dudosa: teoría general, Instituto de Ingeniería, Universidad Nacional Autónoma de México, in press; "On the processing of doubtful information: Part 1: general theory", Solid Mechanics Division, University of Waterloo, in press.

Table 1 Posterior probabilities in examples 2.1 and 2.2

n	P''_1	Q''_{111} or Q''_1					Example
		$n'_i = 0$	1	2	5	∞	
1	.9000	.9900 .9000	.9450 .9000	.9300 .9000	.9150 .9000	.9000 .9000	2.1 2.2
2	.5000	.9500 .5000	.9250 .7000	.9167 .7667	.9083 .8333	.9000 .9000	2.1 2.2
3	.1000	.9100 .1000	.9050 .5000	.9033 .6333	.9017 .7667	.9000 .9000	2.1 2.2
4	.0122	.9012 .0122	.9006 .4561	.9004 .6041	.9002 .7520	.9000 .9000	2.1 2.2
5	.0014	.9001 .0014	.9001 .4507	.9000 .6005	.9000 .7502	.9000 .9000	2.1 2.2

Table 2 Component distributions in example 3.3

i: $x_i \epsilon Y$	$Q''_{i_1 \ldots i_n}$	m''_Y	$\sigma_Y/\sqrt{n''}$
-	.000025	6.4000	.4000
1 or 2 3 4	.002988 .000775 .003539	6.2400 6.7200 7.0400	.1789 .1789 .1789
1,2 1,3 or 2,3 1,4 or 2,4 3,4	.856791 .001407 .000003 .108941	6.2222 6.4889 6.6667 6.9333	.1333 .1333 .1333 .1333
1,2,3 1,2,4 1,3,4 or 2,3,4	.025529 2×10^{-7} .000001	6.4000 6.5231 6.7077	.1109 .1109 .1109
1,2,3,4	10^{-8}	6.5882	.0970

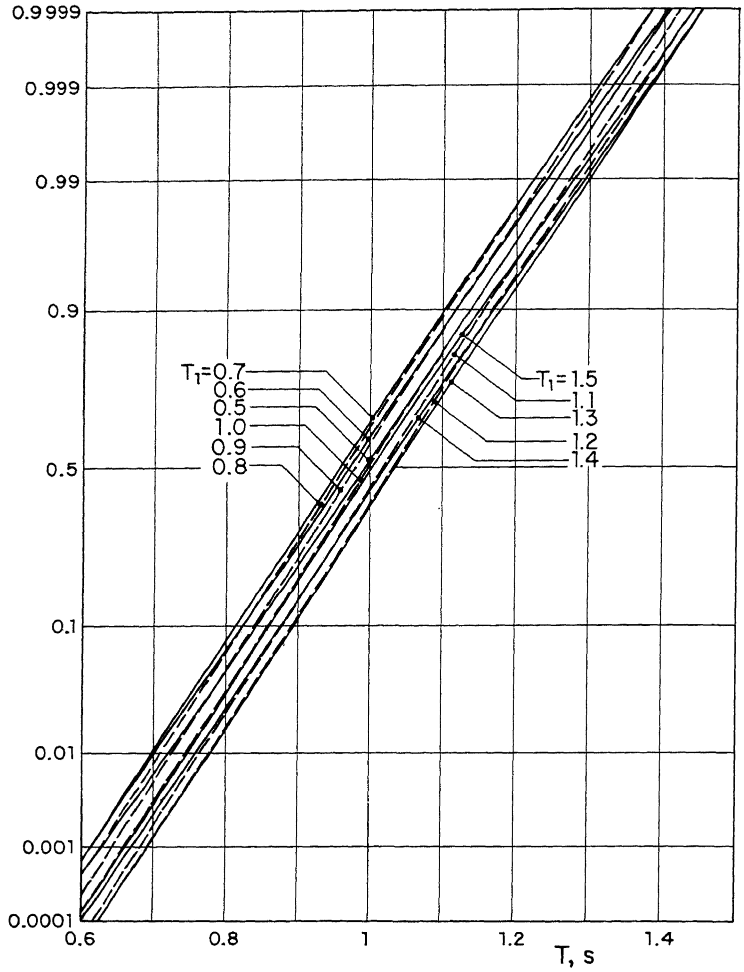

Fig 1. Probability distribution functions, example 3.1

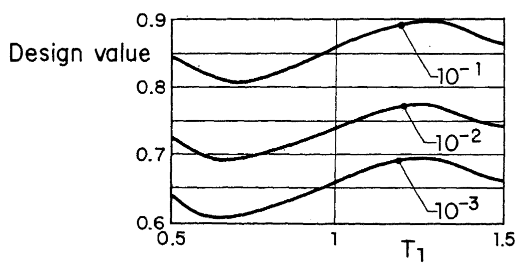

Fig 2. Design values, example 3.2