Integrating Statistical and Non-Statistical Audit Evidence Using Belief Functions: A Case of Variable Sampling*

Rajendra P. Srivastava
Professor, School of Business
University of Kansas
Lawrence, Kansas, 66045

Glenn R. Shafer
Professor, Graduate School of Management
Rutgers University-Newark
Newark, N.J.

June 23, 1993

---

* This research has been supported in part by grants from the Peat Marwick Foundation, the University of Kansas General Research Fund, and the University of Kansas School of Business Research Fund. The authors would like to thank Peter Gillett and the participants of the AI Seminar at the University of Kansas for their valuable comments on an earlier version of the article.

Integrating Statistical and Non-Statistical Audit Evidence Using Belief Functions: A Case of Variable Sampling

Abstract

The main purpose of this article is to show how one can integrate statistical and non-statistical items of evidence under the belief function framework. First, we use the properties of consonant belief functions to define the belief that the true mean of a variable lies in a given interval when a statistical test is performed for the variable. Second, we use the above definition to determine the sample size for a statistical test when a desired level of belief is needed from the sample. Third, we determine the level of belief that the true mean lies in a given interval when a statistical test is performed for the variable with a given sample size. We use an auditing example to illustrate the process.

Key Words: Audit Judgment, Belief Functions, Non-statistical Evidence, Statistical Evidence, Variable Sampling

Introduction

This article has two closely related objectives. The first objective is to determine from belief-function theory the sample size for a statistical test in auditing when a given level of assurance is needed from the sample. The second objective is to show how belief-function theory allows us to combine statistical and non-statistical evidence.

In general, an auditor accumulates both statistical and non-statistical items of evidence. An example of non-statistical evidence is the level of assurance provided by analytical procedures that the inventory account balance is not materially misstated. The auditor evaluates this evidence in light of what he or she knows about the strength of the analytical procedures and relevance of the data used and makes a judgment about the level of assurance it provides for the assertion that the inventory account balance is not materially misstated. An example of statistical evidence is the physical count and valuation of a sample of inventory items. From such statistical evidence, the auditor determines with a certain level of confidence that the inventory account balance is not materially misstated.

We believe that an objective approach to integrating statistical and non-statistical evidence in auditing should make the audit process more efficient. In this article, we provide the belief-function approach for such an integration. (See Appendix B for a discussion of belief functions. Also, see, e.g., Akresh, Loebbecke, and Scott 1988; Aldersley 1989; Shafer 1976; Shafer, Shenoy, and Srivastava 1988; Shafer and Srivastava 1990; Srivastava and Shafer 1992; Srivastava, Shenoy, and Shafer 1990.)

Recently. Srivastava and Shafer (1992) have derived belief-function formulas for audit risk that integrate audit evidence at various levels of the account. They have considered both statistical and non-statistical items of evidence in their treatment. However, they do not provide any linkage between (1) the statistical evidence and the level of assurance obtained from such an item of evidence, and (2) the extent of testing, i.e., the sample size in a statistical sampling and the desired level of assurance. We will deal with these issues in this article.

The remainder of this article is divided into four sections and two appendices. In Section I, we outline the standard statistical approach to sampling using the mean per unit method. In Section II, we show how to assess potential and actual degrees of assurance from statistical evidence in terms of belief functions, and we derive a formula for the sample size required in order to obtain the needed level of assurance. In Section III, we illustrate how statistical evidence can be integrated with non-statistical evidence using a numerical example. In Section IV, we provide a summary of our results and discuss directions for further research. In Appendix A, we derive the *100x% likelihood interval*. Finally, in Appendix B, we discuss the basics of the belief-function formalism and the relationship between the consonant belief functions and the statistical evidence.

I. The Standard Statistical Approach

There are several statistical approaches used in auditing for determining whether the reported account balance is not materially misstated (e.g., see Bailey 1981; and Arens and Loebbecke 1981). We will use the mean per unit approach to illustrate the decision process and compare it with the belief-function approach. Let us consider the audit of an inventory account. On the basis of the *central limit* theorem, we know that even if the population of the values of inventory items is not normally distributed, the sample mean will be normally distributed with the unknown mean μ_o, provided a sufficiently large sample is taken. Let us assume, for simplicity, that we know the standard deviation σ of the population.

The unknown mean μ_o is the true audited mean, the mean we would find if we were able to audit the entire population accurately. The auditor intends to perform an audit procedure (e.g., physical count of inventory items and their valuation) on a sample of inventory items to help in deciding whether the recorded mean, say μ_r, is within an acceptable range of the true mean μ_o, or whether it is materially in error.

It is sometimes said, in this situation, that the null hypothesis is

$$\mu_o = \mu_r, \tag{1}$$

and the alternative hypothesis is

$$\mu_o \neq \mu_r.$$

In a sense, however, the real null hypothesis is that μ_o is in the interval

$$\mu_r - TE \leq \mu_o \leq \mu_r + TE, \quad (2)$$

where TE is the maximum tolerable error per item—the minimum error regarded as material. The alternative hypothesis is that μ_o falls outside the interval (2)—the difference between μ_o and μ_r is more than TE.

The standard approach to this testing problem (Bailey 1981, Roberts 1978) uses a sample size n such that

$$Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = \frac{Z_{\alpha/2}}{Z_{\alpha/2} + Z_\beta} TE,$$

where $Z_{\alpha/2}$ and Z_β are standard normal deviates. In other words,

$$n = \left[\frac{(Z_{\alpha/2} + Z_\beta)\sigma}{TE} \right]^2. \quad (3)$$

The auditor uses the acceptance region

$$\mu_r - A \leq \bar{y} \leq \mu_r + A, \quad (4)$$

where $\bar{y}$ is the sample mean and the precision A is given by

$$A = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = TE - Z_\beta \frac{\sigma}{\sqrt{n}}. \quad (5)$$

If the sample mean falls in the acceptance region (4), then the hypothesis (2) is accepted; otherwise it is rejected.

The rationale for the acceptance region (4) is that the probability of accepting (2) will be 1-α if μ_r is exactly correct (exactly equal to μ_o) and approximately β if the error in μ_r is just barely material (μ_r is exactly TE from μ_o). In hypothesis testing terminology, α is the probability of Type I error, the probability of rejecting the null hypothesis if it is exactly true (actually, if (1) is true), while β is the maximum probability of Type II error, the probability of accepting the null hypothesis if it is just barely false.

In auditing, the auditor very seldom knows σ. Thus, he or she must use an estimated standard deviation, S, for computing the precision. Since the auditor is more concerned with the incorrect acceptance, he or she determines the precision from (5) based on a desired level of β-risk (AICPA 1981, 1983a, 1983b):

$$A = TE - Z_\beta \frac{S}{\sqrt{n}} . \tag{6}$$

The effective level of risk of incorrect rejection (α-risk or Type I error) is obtained from $Z_{\alpha/2} = \frac{TE}{S/\sqrt{n}} - Z_\beta$ (see Equation 5).

II. Belief-Function Approach

Shafer (Chapter 9, 1976) has proposed a "likelihood method" for determining the level of assurance or belief from statistical evidence. In this approach, the level of confidence in the *likelihood interval* (see Appendix A for the definition) is defined to be the level of belief in the interval. For example, a 100x% likelihood interval provides x level of belief that the true audited mean μ_o lies in the interval. As derived in Appendix A (A-2), the 100x% likelihood interval for μ_o when the sample audited mean is $\bar{y}$ is given by

$$\left[\bar{y} - \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1 - x)},\ \bar{y} + \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1 - x)}\right], \tag{7}$$

and thus the belief, say $\mathbf{Bel}_2$, from the statistical evidence that μ_o lies in the above interval is x (also, see the discussion in Appendix B), i.e.,

$$\mathbf{Bel}_2\left(\left[\bar{y} - \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1 - x)} \le \mu_o \le \bar{y} + \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1 - x)}\right]\right) = x. \tag{8}$$

In general, we consider the recorded account balance to be not materially misstated[1] (fs*:* fairly stated) when the recorded mean is within the tolerable error of the true audited mean, i.e., $|\mu_r - \mu_o| \le TE$, or

$$\mu_r - TE \le \mu_o \le \mu_r + TE. \tag{9}$$

However, since we do not know the true audited mean, we need to express the above condition in terms of the sample audited mean, $\bar{y}$.

We want our degree of belief in the condition (9) (i.e., the account balance is not materially misstated) to be at least x. In order for this to happen, the interval in (7) must be contained in the interval $[\mu_r - TE, \mu_r + TE]$. This means that (see Footnote 1)

$$\mu_r - TE + \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1-x)} \leq \bar{y} \leq \mu_r + TE - \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1-x)}. \qquad (10)$$

In other words, if (10) were true then we will have the belief that the account balance is not materially misstated to be at least equal to x. But, there is some level of uncertainty in achieving (10). Thus, for planning purposes, we want probability $(1 - \alpha)$ of achieving a belief of at least x that the account is not materially misstated. In other words, we want the condition in (10) to hold with probability $(1 - \alpha)$ when the true audited mean μ_o is exactly equal to the recorded value μ_r. This requirement can be written as:

$$\mathbf{P}\left(\mu_o - TE + \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1-x)} \leq \bar{y} \leq \mu_o + TE - \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1-x)}\right) = (1-\alpha),$$

or

$$\mathbf{P}\left(-\frac{TE}{\sigma/\sqrt{n}} + \sqrt{-2\log_e(1-x)} \leq \frac{\bar{y} - \mu_o}{\sigma/\sqrt{n}} \leq \frac{TE}{\sigma/\sqrt{n}} - \sqrt{-2\log_e(1-x)}\right) = (1-\alpha). \qquad (11)$$

In order for (11) to be satisfied, we must have $\frac{TE}{\sigma/\sqrt{n}} - \sqrt{-2\log_e(1-x)} = Z_{\alpha/2}$, or

$$n = \frac{\sigma^2}{TE^2}\left[Z_{\alpha/2} + \sqrt{-2\log_e(1-x)}\right]^2. \qquad (12)$$

Failing to obtain the desired degree of belief of at least x from the statistical evidence can be thought of as rejecting the null hypothesis. The probability of rejecting if the null hypothesis $\mu_o = \mu_r$ is exactly true—the significance level of the test—is still, by design, α. The probability of rejecting if the null hypothesis is just barely materially false—the minimum power of the test—is now the probability of (10) failing when μ_o is equal to μ_r + TE, or

$$1 - \mathbf{P}\left(-2Z_{\alpha/2} - \sqrt{-2\log_e(1-x)} \leq \frac{\bar{y} - (\mu_r + TE)}{\sigma/\sqrt{n}} \leq -\sqrt{-2\log_e(1-x)}\right)$$

$$\geq 1 - \mathbf{P}\left(\frac{\bar{y} - (\mu_r + TE)}{\sigma/\sqrt{n}} \leq -\sqrt{-2\log_e(1-x)}\right) = 1 - \beta, \qquad (13)$$

where β is Type II error discussed earlier and the corresponding normal deviate Z_β is given by

$$Z_\beta = \sqrt{-2\log_e(1-x)}. \qquad (14)$$

From (14), the desired belief x that the condition (10) is true with probability (1- α) and power (1- β) can be given by

$$x = 1 - \exp(-\tfrac{1}{2}Z_{\beta}^{2}). \tag{15}$$

Sample Size Determination

It is noteworthy that the sample size formula (12) for the belief-function approach is the same as (3) for the standard statistical approach. The only difference is that (12) provides us the desired level of belief in the condition (10) that the account balance is not materially misstated which then can be combined with the beliefs from the non-statistical evidence to obtain the overall belief. As discussed in Section III, the belief function approach is relatively more efficient because it aggregates objectively both the statistical and non-statistical items of evidence.

As we have seen, if our goal is to obtain degree of belief x that there is no material error, and if we want to have a probability (1− α) of obtaining this goal when there is no error at all, then the sample size we must use is

$$n = \frac{\sigma^2}{TE^2}\left[Z_{\alpha/2} + \sqrt{-2\log_e(1-x)}\right]^2, \tag{16}$$

and the minimum power of this test is (1- β) where β is given by

$$\beta = \mathbf{P}\left(\frac{\bar{y} - (\mu_r + TE)}{\sigma/\sqrt{n}} \le -\sqrt{-2\log_e(1-x)}\right). \tag{17}$$

Let us consider a specific example. Suppose the auditor wants to determine the sample size to conduct an audit of the inventory account. There is a total of 1,000 items in the stock (N = 1,000). The total recorded value of the inventory is $500,000. The recorded mean μ_r = $500. The estimated standard deviation σ = $75, and the tolerable error TE = $25 per item. Suppose the auditor plans to conduct the audit at 20% risk of incorrect rejection (i.e., α = 20%) and plans to achieve 70% of belief in the decision interval. From (16), the sample size for this interval is:

$$n = \frac{75^2}{25^2}\left[Z_{0.10} + \sqrt{-2\log_e(1-0.7)}\right]^2$$
$$= 9\left[1.28 + \sqrt{-2\log_e(0.3)}\right]^2 = 72.$$

Equation (16) is used to compute values of n as given in Table 1 for σ = \$75, TE = \$25, and for various levels of risk of incorrect rejection, α, and various levels of the desired belief x that the account balance is not materially misstated. This table also shows the corresponding power of the test. As discussed earlier, the standard statistical approach yields the same sample size as the belief-function approach for a given level of risk of incorrect rejection and power.

Table 1

Sample size, n, for a desired level of belief, x, that the true audited mean μ_o lies in the interval B = [\$475, \$525], i.e., the account is not materially misstated, given that the recorded mean μ_r = \$500, σ = \$75, and TE = \$25.

Desired Level of α	$Z_{\alpha/2}$	Desired level of Belief in 'fs' x	Sample Size From (16) n	Z_β From (14)	Corresponding Level of Power (1- β)
0.20	1.28	0.40	47	1.01	0.84
0.20	1.28	0.50	54	1.18	0.88
0.20	1.28	0.60	62	1.35	0.91
0.20	1.28	0.70	72	1.55	0.94
0.15	1.44	0.40	54	1.01	0.84
0.15	1.44	0.50	62	1.18	0.88
0.15	1.44	0.60	70	1.35	0.91
0.15	1.44	0.70	81	1.55	0.94
0.10	1.65	0.40	64	1.01	0.84
0.10	1.65	0.50	72	1.18	0.88
0.10	1.65	0.60	81	1.35	0.91
0.10	1.65	0.70	92	1.55	0.94
0.05	1.96	0.40	79	1.01	0.84
0.05	1.96	0.50	89	1.18	0.88
0.05	1.96	0.60	99	1.35	0.91
0.05	1.96	0.70	111	1.55	0.94

It should also be noted that the sample sizes computed here are all based on the assumption that errors are independent from one sample item to another sample item. In many cases, errors in the sample items will be dependent, and the dependence may justify a smaller sample. Evidence for the control procedure may indicate, for example, that if the procedure is effective, then no material errors will result. In this case, the absence of errors in only a few randomly sampled items will strengthen the auditor's belief that the procedure is effective and that most of the remaining items in the population are also materially correct. We intend to provide a formal analysis of this kind of argument, in both the belief-function and the Bayesian frameworks, in a future article.

Evaluation of Sample Results

Suppose the auditor has performed the statistical test and obtained the sample mean, $\bar{y}$, and the standard error of the mean, $S/\sqrt{n}$. The condition that the recorded account balance is not materially misstated is (see Footnote 1):

$$\mu_r + TE - \bar{y} = \frac{S}{\sqrt{n}} \sqrt{-2\log_e(1-x)} \quad \text{for } \mu_r \le \bar{y} \le \mu_r + TE, \tag{18}$$

and

$$\bar{y} - (\mu_r - TE) = \frac{S}{\sqrt{n}} \sqrt{-2\log_e(1-x)} \quad \text{for } \mu_r - TE \le \bar{y} \le \mu_r. \tag{19}$$

Solving (18) and (19) yields the highest achieved level of belief x that the recorded account balance is not materially misstated. The result can be written in a single expression as:

$$\mathbf{Bel}_2(fs) = x = 1 - \exp\left(-\frac{n}{2S^2}(TE - |\bar{y} - \mu_r|)^2\right), \text{ for } \mu_r - TE \le \bar{y} \le \mu_r + TE. \tag{20}$$

The belief in 'not fs' when the observed sample mean lies in the interval B = [μ_r - TE, μ_r+ TE] is zero as discussed in Appendix B (B-9), i.e.,

$$\mathbf{Bel}_2(\text{not fs}) = 0, \text{ for } \mu_r - TE \le \bar{y} \le \mu_r + TE. \tag{21}$$

Similarly from (B-9) and (B-10), we obtain the following beliefs when the observed mean $\bar{y}$ falls outside the interval B, i.e., for $\bar{y} \ge \mu_r$+ TE or $\bar{y} \le \mu_r$ - TE

$$\mathbf{Bel}_2(fs) = 0, \tag{22}$$

$$\boxed{\mathbf{Bel}_2(\text{not fs}) = 1 - \exp\left(-\frac{n(\bar{y} - \mu_r - TE)^2}{2S^2}\right)}, \text{ for } \bar{y} \geq \mu_r + TE, \tag{23}$$

and

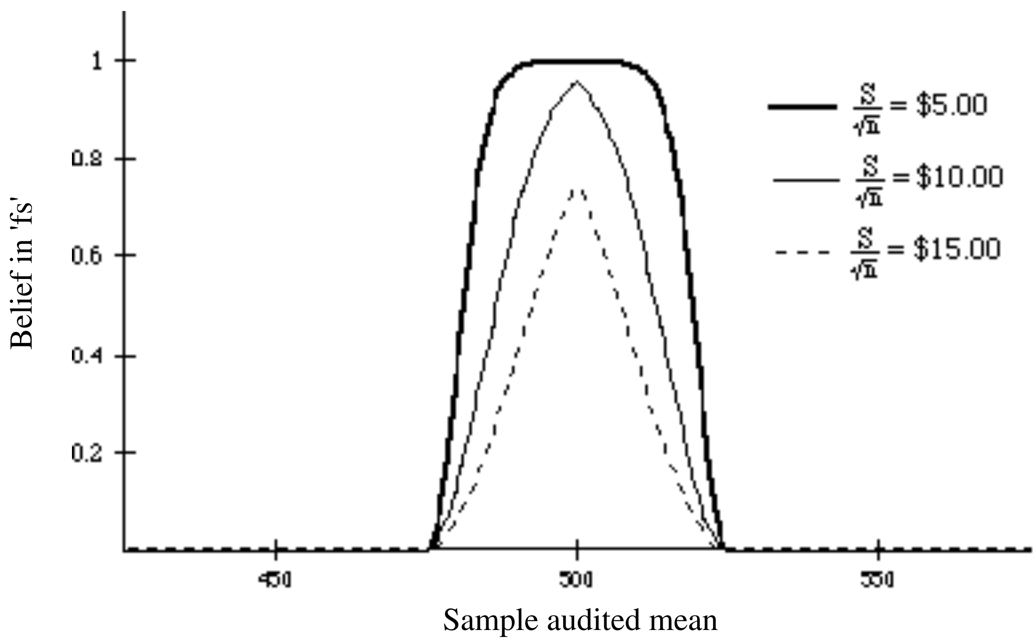
$$\boxed{\mathbf{Bel}_2(\text{not fs}) = 1 - \exp\left(-\frac{n(\mu_r - TE - \bar{y})^2}{2S^2}\right)}, \text{ for } \bar{y} \leq \mu_r - TE. \tag{24}$$

Equations (20-21), and (22-24) give beliefs that the recorded account balance is not materially misstated (fs), and materially misstated (not fs), respectively, when the observed mean $\bar{y}$ falls in the interval B, and outside of the interval B. The results are interesting and intuitive as discussed below.

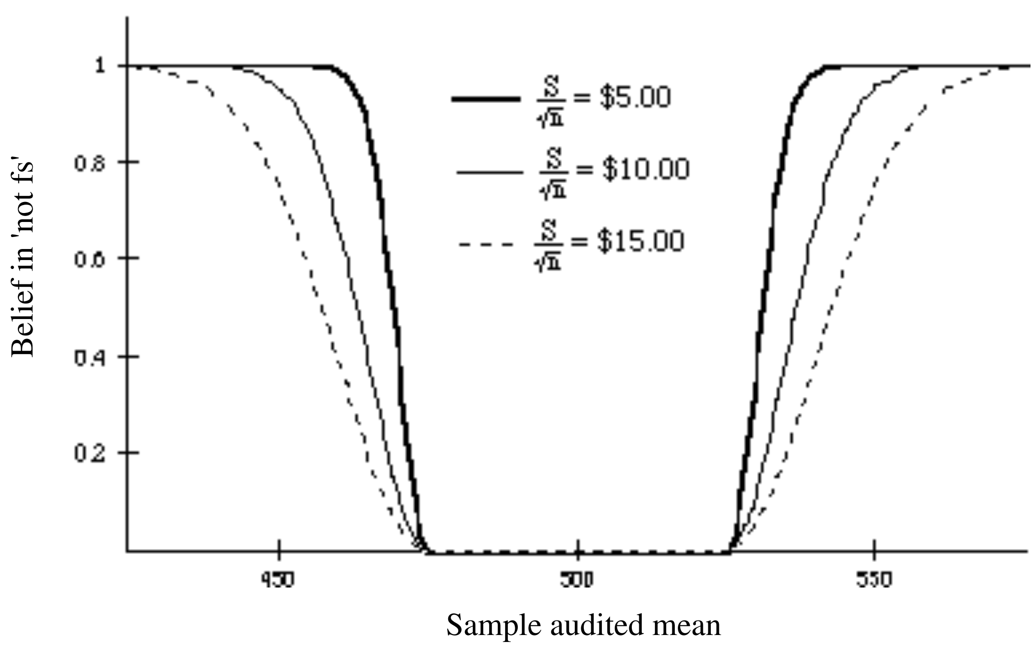
Figure 1

Belief in 'fs' and 'not fs' (Equations 20-24) as a function of the sample audited mean $\bar{y}$ for different values of the observed standard error of the mean ($S/\sqrt{n}$). B = [\$475, \$525], μ_r = \$500, n = 100, and TE = \$25.

Panel A: Belief in 'fs' as a function of the sample audited mean.



Panel B: Belief in 'not fs' as a function of the sample audited mean.



Bel$_2$(fs) and **Bel**$_2$(not fs) are plotted as a function of the sample audited mean in Figure 1 for three values of the standard error of the mean ($5, $10, and $15), and for μ_r =$500, TE =$25 per item, and n = 100. In this case, the interval B = [$475, $525]. As evident from (20-24), **Bel**$_2$(fs) and **Bel**$_2$(not fs) depend on the difference between the observed mean error (μ_r - $\bar{y}$) and the tolerable error TE. When the observed mean error is less than TE, i.e., $|\mu_r - \bar{y}| < TE$, **Bel**$_2$(fs) > 0 and peaks at $\bar{y} = \mu_r$, but **Bel**$_2$(not fs) = 0 (see Panel A of Figure 1). However, when the observed mean error is greater than the tolerable error, i.e., $|\mu_r - \bar{y}| > TE$, **Bel**$_2$(fs) = 0 and **Bel**$_2$(not fs) > 0 with **Bel**$_2$(not fs) approaching 1 at a large mean error (see Panel B of Figure 1). Both **Bel**$_2$(fs) and **Bel**$_2$(not fs) are affected by the size of the standard error of the mean (see Figure 1). The practical implication of this result is that the auditor will have a higher level of belief in 'fs' or 'not fs', as the case may be, for the same observed sample mean but for a lower standard error of the mean, an intuitive result.

It is noteworthy that **Bel**$_2$(fs) = 0 and **Bel**$_2$(not fs) = 0 at the end points of the interval B. This simply means that if the observed mean falls right at one of the end points then we are ignorant about whether the account balance is materially misstated or not materially misstated. If we compute the corresponding Type II error (β-risk) from (14) and (17) we find β-risk = 0.5. This result is similar to what we have established in the belief function framework; it simply means that the probability that the account is materially misstated is 50% and that it is not materially misstated is 50%. That is, we are completely ignorant about its fairness if the observed mean falls right at one of the end points of the interval B.

III. Integrating Statistical and Non-statistical Audit Evidence

In this section, we plan to discuss two situations of integrating statistical evidence with non-statistical evidence. One deals with planning of an audit. The other deals with the evaluation of an audit.

Planning an Audit

Consider the earlier example of the inventory account audit. The auditor has already performed the procedures related to the non-statistical evidence such as certain analytical procedures, and study and evaluation of the accounting system and the control environment related to the account. The auditor's assessment of belief[2] from all such evidence that the account is not materially misstated is, say, 0.7, i.e., $\mathbf{Bel}_1$(fs) = 0.7, and that the account is materially misstated is zero, i.e., $\mathbf{Bel}_1$(not fs) = 0. Suppose the auditor feels that an overall degree of belief of 0.91[3] is needed in order to feel justified in asserting that the account balance is not materially misstated. We want to find the level of belief the auditor should plan to achieve from the statistical test in order to obtain the overall 0.91 degree of belief.

Consider the following level of beliefs from the statistical evidence: $\mathbf{Bel}_2$(fs) = x and $\mathbf{Bel}_2$(not fs) = 0, i.e., $\mathbf{m}_2$(fs) = x, $\mathbf{m}_2$(not fs) = 0, $\mathbf{m}_2$({fs, not fs}) = 1 - x (see Appendix B for the definition of **m**-values). Combining the two items of evidence with no conflict using Dempster's rule[4] we obtain:

$$\mathbf{m}(\text{fs}) = \mathbf{m}_1(\text{fs})\mathbf{m}_2(\text{fs}) + \mathbf{m}_1(\text{fs})\mathbf{m}_2(\{\text{fs, not fs}\}) + \mathbf{m}_1(\{\text{fs, not fs}\})\mathbf{m}_2(\text{fs})$$
$$= 0.7\ x + 0.7\ (1 - x) + 0.3\ x = 0.7 + 0.3\ x.$$

But, we want $\mathbf{m}$(fs) = 0.91 that yields x = 0.7. Thus, $\mathbf{Bel}_2$(fs) = 0.7 and $\mathbf{Bel}_2$(not fs) = 0. The auditor can now determine the extent of statistical testing using (16). For TE = \$25, α-risk = 20% (i.e., $Z_{\alpha/2}$ = 1.28), σ = \$75, and x = 0.7, we obtain n = 72 from (16). The corresponding β-risk as calculated in Table 1 is 6.0% (see row 4 in Table 1).

Evaluation of Audit

Continuing with the above example, assume that the auditor randomly selected 72 items (n=72) from the inventory population and performed the relevant audit procedures. The audited sample mean is \$485 based on the mean per unit method (MPU) of variable sampling, i.e., [illegible] = \$485. The measured standard deviation is \$80. As considered earlier, the recorded mean, μ_r = \$500. Thus with TE =\$25 per item, the interval B = [\$475, \$525]. Our interest here is to find whether the result is good enough to accept the account balance to be not materially misstated.

For comparison purposes, we will treat the above findings using both: (1) the belief-function approach, and (2) the MPU approach (mean per unit method in statistics).

Belief-Function Approach

Using (20) and (21), the belief that the recorded account balance is not materially misstated when the sample audited mean $\bar{y}$ =\$485 is:

$$\mathbf{Bel}_2(\text{fs}) = 1 - \exp\left(-\frac{72}{2(80)^2}(25 - |485 - 500|)^2\right) = 0.4302,$$

and

$$\mathbf{Bel}_2(\text{not fs}) = 0.$$

This implies that

$$\mathbf{m}_2(\text{fs}) = 0.4310,\ \mathbf{m}_2(\text{not fs}) = 0,\ \mathbf{m}_2(\{\text{fs, not fs}\}) = 0.5698.$$

Apparently, the target belief of 0.7 in 'fs' is not available from the statistical evidence. When the above evidence is combined with the non-statistical evidence (with $\mathbf{m}_1$(fs) = 0.7, $\mathbf{m}_1$(not fs) = 0, $\mathbf{m}_1$({fs, not fs}) = 0.3), the overall belief[5] in 'fs' is **Bel**(fs) = 0.8291. This is a smaller belief than the planned value of 0.91. There are several options that the auditor can choose, including: (1) The auditor can reject the account balance as fairly stated and refuse to give an unqualified opinion on the account. (2) The auditor can propose an adjusting entry to the account balance and thus increase the belief that the account is not materially misstated. (3) The auditor can increase the sample size.

In the first alternative, the auditor would issue either a "qualified opinion" or an "adverse opinion" depending on the severity of the misstatement (see, e.g., Arens and Loebbecke 1991). However, under the second alternative, when the auditor decides to propose an adjusting entry, the question is what should be the adjustment under the belief-function framework? The following discussion answers this question.

In order to achieve the desired level of belief that the account balance is not materially misstated, we must obtain 0.7 degree of belief from the statistical evidence. In our example, we will obtain 0.7 degree of belief from such an item of evidence if the recorded mean is reduced to \$495.39 from \$500.00 (use (20) to compute the adjusted recorded mean). This means that the auditor should propose a reduction of the recorded mean by \$4.61 that will then yield the desired level 0.91 for the overall belief in 'fs.'

Under the third alternative, the auditor would increase the sample size to achieve the desired level of belief from the statistical evidence. In our example, we need 0.7 level of belief from the statistical evidence. Since the auditor has already estimated the standard deviation to be \$80, one should use this value to determine the new sample size using (16). The result is[6] n = 82 compared to n = 72, the initial sample size. This new sample size requires that the auditor select ten more items from the inventory stock and perform the audit procedures. Suppose the auditor has done that and obtained a new mean of \$489, i.e., $\bar{y}$ = \$489 and a new standard deviation of \$81, i.e., S = \$81. Using (20-21), the auditor obtains[7] a new belief from the statistical evidence of 0.7062 that the account is fairly stated and no belief that the account is not fairly stated. This value is above the desired value of 0.7, and thus when the auditor combines the two items of evidence, the statistical evidence and the non-statistical evidence, the overall belief that the account is fairly stated becomes[8] 0.9119 which is above the overall desired value of 0.91. Now, the auditor would feel comfortable accepting the account to be fairly stated since the combined belief in 'fs' is above his desired level.

MPU Approach

The computed lower and upper confidence limits (Arens and Loebbecke 1991, p. 520) at the planned level of β-risk of 6.0% (Z_β = 1.552) are:

Computed lower confidence limit = \$470.39

Computed upper confidence limit = \$499.61

The above confidence interval is not contained in the interval [\$475, \$525] and therefore to accept the account to be not materially misstated the auditor must propose an adjusting entry of -\$4.61 to the recorded mean. This conclusion is exactly the same as obtained in the belief-function approach. In fact, the statistical approach will yield the same result as that of the belief-function approach if the level of β-risk used in the statistical approach is based upon the level of belief desired in the belief-function approach (17).

It is important to point out here that the efficiency in the belief-function approach should come from aggregating all the evidence at various levels of the account as discussed by Srivastava and Shafer (1992). For example, comparing their Tables 7 and 8, we find that if the items of evidence at the financial statement level and at the account level are ignored then the overall belief that the account is not materially misstated is only 0.71 instead of 0.95. Thus in order to achieve a higher overall assurance in the traditional approach, the auditor will have to plan the audit at a lower detection risk and hence a larger sample size.

In the above example we have considered only affirmative type non-statistical evidence, i.e., the evidence is providing beliefs to 'fs' and no assurance to 'not fs.' But, there may be situations where the non-statistical evidence may provide assurance for 'not fs' too. In those situations, Dempster's rule will lead to a larger sample size.

IV. Summary and Conclusion

We have demonstrated how beliefs can be assessed from the statistical evidence and how these beliefs can be integrated with the non-statistical evidence. We have derived a formula for the sample size needed for a desired level of belief. Numerical examples are used to illustrate both the sample size determination and the sample result evaluation. It is interesting that the sample size formula is similar to the formula used in the standard statistical approach. As expected, with all the factors held fixed, the sample size increases with the increase in the desired level of belief. Also, we found that when the sample audited mean is not within the acceptable range of the recorded mean, the level of adjustment to be made in the recorded mean for the account to be acceptable is the same in the two approaches for the same level of β-risk.

However, as discussed by Srivastava and Shafer (1992), the belief-function approach, in general, should provide a more efficient audit because it allows the auditor to integrate all the non-statistical evidence in an objective way. If the belief from the non-statistical evidence is large then we need only a small amount of belief from the statistical evidence to achieve the desired level of overall belief. In this situation, the corresponding β-risk will be relatively large and thus the belief-function approach would require a smaller sample size compared to the traditional approach.

In the present article we only deal with the variable sampling and independent errors. There are several areas for extension of this work including: (1) Assessment of beliefs and sample size determination in attribute sampling. (2) Assessment of beliefs and sample size determination in dollar unit sampling. (3) Assessment of beliefs and sample size determination when errors are not assumed to be independent of each other.

Footnotes

1. The condition that the account balance is not materially misstated, i.e., the account balance is fairly stated (fs) is

$$\text{fs: } \mu_r - TE \le \mu_o \le \mu_r + TE.$$

However, μ_o is not known. Therefore, we will use the likelihood interval for determining the condition under which the recorded account balance will not be materially misstated. The 100x% likelihood interval for μ_o is (A-2)

$$\left[\bar{y} - \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1 - x)},\ \bar{y} + \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1 - x)}\right].$$

The condition that the account balance is not materially misstated is obtained by requiring that the likelihood interval be contained in the interval $[\mu_r - TE, \mu_r + TE]$, that is, the upper boundary of the likelihood interval be less than $(\mu_r + TE)$ and the lower boundary be greater than $(\mu_r - TE)$. This requirement gives the following condition for the recorded account balance to be not materially misstated (i.e., fairly stated):

$$\text{fs: } \mu_r - TE + \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1- x)} \le \bar{y} \le \mu_r + TE - \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1- x)}.$$

The condition that the account balance is materially misstated is (i.e., not fairly stated):

$$\text{not fs: } \bar{y} \le \mu_r - TE + \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1- x)} \text{ or } \bar{y} \ge \mu_r + TE - \frac{\sigma}{\sqrt{n}}\sqrt{-2\log_e(1- x)}$$

2. As discussed in Srivastava and Shafer (1992), one can use individual assessment of the strength of the non-statistical items of evidence then combine them using Dempster's rule to obtain an overall strength. We have the following level of beliefs from the non-statistical evidence: $\mathbf{Bel}_1(\text{fs}) = 0.7$ and $\mathbf{Bel}_1(\text{not fs}) = 0$. Since we need the *basic probability assignment functions* or **m**-values in Dempster's rule, we need to express the above beliefs in terms of **m**-values. For the above case we have $\mathbf{m}_1(\text{fs}) = 0.7$, $\mathbf{m}_1(\text{not fs}) = 0$, and $\mathbf{m}_1(\{\text{fs, not fs}\}) = 0.3$.

3. We choose 0.91 for simplicity so that we can use Table 1 for the sample size calculation.

4. In the case of two independent items of evidence with $\mathbf{m}_1$ and $\mathbf{m}_2$ representing the **m**-values on the frame of interest Θ, Dempster's rule yields the combined **m**-value for a subset A of frame Θ to be:

$$\mathbf{m}(A) = K^{-1}\Sigma\{\mathbf{m}_1(B_1)\mathbf{m}_2(B_2) | B_1 \cap B_2 = A, A \ne \emptyset\},$$

where K is the re-normalization constant;

$$K = 1 - \Sigma\{\mathbf{m}_1(B_1)\mathbf{m}_2(B_2) | B_1 \cap B_2 = \emptyset\}.$$

The second term in K above represents the conflict between the two items of evidence. When
the two items of evidence exactly contradict each other, i.e., when K = 0, Dempster's rule cannot be used to combine such evidence. See Shafer (1976) for combining more than two items of evidence.

5. The two items of evidence provide the following **m**-values: $\mathbf{m}_1(\text{fs}) = 0.7$, $\mathbf{m}_1(\text{not fs}) = 0$, and $\mathbf{m}_1(\{\text{fs, not fs}\}) = 0.3$, and $\mathbf{m}_2(\text{fs}) = 0.4302$, $\mathbf{m}_2(\text{not fs}) = 0$, and $\mathbf{m}_2(\{\text{fs, not fs}\}) = 0.5698$. Using Dempster's rule (Footnote 4), we obtain K = 1 and

$$\mathbf{m}(\text{fs}) = \mathbf{m}_1(\text{fs})\mathbf{m}_2(\text{fs}) + \mathbf{m}_1(\text{fs})\mathbf{m}_2(\{\text{fs, not fs}\}) + \mathbf{m}_1(\{\text{fs, not fs}\})\mathbf{m}_2(\text{fs})$$
$$= 0.7\text{x}0.4302 + 0.7\text{x}0.5698 + 0.3\text{x}0.4302 = 0.8291,$$

$\mathbf{m}$(not fs) = 0, and $\mathbf{m}$({fs, not fs}) = $\mathbf{m}_1$({fs, not fs})$\mathbf{m}_2$({fs, not fs}) = 0.3x0.5698 = 0.1709.

6. $n = \frac{80^2}{25^2}\left[1.28 + \sqrt{-2\log_e(1-0.7)}\right]^2 = 82.$

7. $\mathbf{Bel}_2(\text{fs}) = 1 - \exp\left(-\frac{82}{2(81)^2}(25 - |489 - 500|)^2\right) = 0.7062$, $\mathbf{Bel}_2$[not fs] = 0.

8. $\mathbf{m}$(fs) = $\mathbf{m}_1$(fs)$\mathbf{m}_2$(fs) + $\mathbf{m}_1$(fs)$\mathbf{m}_2$({fs, not fs}) + $\mathbf{m}_1$({fs, not fs})$\mathbf{m}_2$(fs)

   = 0.7x0.7062 + 0.7x0.2938 + 0.3x0.7062 = 0.9119,

References

Akresh, A. D., J. K. Loebbecke, and W. R. Scott. 1988. Audit Approaches and Techniques. *Research Opportunities in Auditing: The Second Decade*. edited by A. R. Abdel-khalik and Ira Solomon. Sarasota, FL:AAA.13-55.

Aldersley, Steve. 1989. Discussion of Achieved Audit Risk and the Audit Outcome Space. *Auditing: A Journal of Practice and Theory* (Supplement): 85-97.

American Institute of Certified Public Accountants. 1981. *Statements on Auditing Standards*, *No. 39: Audit Sampling.* New York: AICPA.

American Institute of Certified Public Accountants. 1983a. *Audit and Accounting Guide: Audit Sampling*. New York: AICPA.

American Institute of Certified Public Accountants. 1983b. *Statements on Auditing Standards*, *No. 47: Audit Risk and Materiality in Conducting an Audit.* New York: AICPA.

Arens, A. A. and Loebbecke, J. K. 1981. *Applications of Statistical Sampling to Auditing*, Englewood Cliffs, NJ: Prentice-Hall.

Arens, A. A. and Loebbecke, J. K. 1991. *AUDITING: An Integrated Approach*. Englewood Cliffs, NJ: Prentice-Hall.

Bailey, Jr., A. D. 1981. *Statistical Auditing: Review, Concepts, and Problems.* Hracourt Brace Jovanovich.

Edwards, A.W. F. 1972. *Likelihood*. Cambridge University Press.

Roberts, D. M. 1978. *Statistical Auditing*. American Institute of Certified Public Accountants. New York.

Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton University Press.

Shafer, G. R. P. P. Shenoy, and R. P. Srivastava. 1988. AUDITOR'S ASSISTANT: An Knowledge Engineering Tool for Audit Decisions. *Proceedings of the 1988 Touche Ross/University of Kansas Symposium on Auditing Problems*. Lawrence, KS: School of Business, University of Kansas, 61-84.

Shafer, G. R. and R. P. Srivastava. 1990. The Bayesian And Belief-Function Formalisms: A General Perspective for Auditing. *Auditing: A Journal of Practice and Theory*, (Supplement): 110-48.

Srivastava, R. P. and G. R. Shafer. 1992. Belief-Functions Formulas for Audit Risk. *The Accounting Review* (April): 249-283.

Srivastava, R. P., P. P. Shenoy, and G. R. Shafer. 1990. Belief Propagation in Networks for Auditing. *Audit Judgment Symposium*. University of Southern California (February 19-20).

APPENDIX A

Likelihood Interval

The *likelihood function*, in any parametric statistical problem, is the function of the parameter that assigns to each value of the parameter the probability (or probability density) of the actual observations given that value of the parameter. In our problem, where the probability density is normal with known variance σ^2/n, the unknown parameter is μ_o, and the observation is the sample mean $\bar{y}$, the likelihood function is

$$L(\mu_o) = \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left[- \frac{n\,(\bar{y}-\mu_o)^2}{2\sigma^2}\right] \tag{A-1}$$

The *maximum likelihood estimate* of μ_o is the value of μ_o that maximizes the likelihood function—in this case, $\bar{y}$. Edwards (1972) defines the 100x % confidence interval for μ_o based on the likelihood approach to be the interval consisting of all values of μ_o for which $L(\mu_o)$ is at least (1- x) of its maximum value. In our case, this is the interval consisting of all μ_o such that

$$\frac{L(\mu_o)}{L(\bar{y})} \geq (1 - x),$$

or

$$\exp\left[- \frac{n(\bar{y}-\mu_o)^2}{2\sigma^2}\right] \geq (1 - x),$$

or

$$(\bar{y} - \mu_o)^2 \leq - \frac{2\sigma^2}{n} \log_e(1 - x),$$

where $\log_e$ denotes the natural logarithm. This is a symmetric interval around $\bar{y}$; the endpoints are

$$\bar{y} \pm \frac{\sigma}{\sqrt{n}} \sqrt{- 2 \log_e(1 - x)}.$$

The interval

$$\left[\bar{y} - \frac{\sigma}{\sqrt{n}} \sqrt{- 2 \log_e(1 - x)},\ \bar{y} + \frac{\sigma}{\sqrt{n}} \sqrt{- 2 \log_e(1 - x)}\right] \tag{A-2}$$

is the 100x% likelihood interval.

APPENDIX B

Belief-Function Formalism

The belief-function formalism is based on mathematical probability, just as the Bayesian formalism is, but it allows us to bring probability statements to bear on questions of interest in a less direct way. Belief functions have antecedents in the seventeenth century work of George Hooper and James Bernoulli (Shafer 1986).

This section is an elementary introduction to those aspects of the formalism that we use in the article. For a more comprehensive and detailed introduction, see Shafer (1976).

Frames and Compatibility Relations

We call an exhaustive and mutually exclusive set of possible answers to a question a *frame*. (Some readers may prefer the name *sample space*; but it may be better to reserve this name for the set of outcomes in an experiment with well-defined probabilities.) We will often use the symbol Θ to represent the frame in which we are interested. In the case of a yes-no question, the frame has only two elements;

$$\Theta = \{\text{yes,no}\},$$

or

$$\Theta = \{\text{the account is fairly stated, the account is not fairly stated}\},$$

etc. But in general, a frame may be a very large set, for its question may have many possible answers.

In auditing problems, we typically work with many different frames. The basic question and hence the basic frame of interest may be very simple (Is the financial statement fairly stated? Yes or no.). But it may be necessary to bring in many subsidiary and related questions. If we need to consider these questions together, we may end up working with very large and complex frames.

Belief-function degrees of belief arise when we have probabilities not for the frame Θ that interests us but instead for some related frame, say S. We might, for example, have probabilities for whether a control procedure is being properly followed, but we might not have probabilities for the more basic question of real interest, the question of whether the account to which the control procedure is related is fairly stated. If the relation between the two frames can be expressed by saying that certain elements of the two frames are incompatible with each other, then our probabilities for S can give rise to belief-function degrees of belief for Θ. For example, if we judge that the account will be fairly stated if the control procedure is being properly followed, then our probability for the control procedure being properly followed will lead to a degree of belief that the account is fairly stated.

The knowledge only certain elements of Θ are not compatible with certain elements of S can be represented formally by listing, for each element s of S, those elements of Θ that are compatible with s. We may call such a list the *compatibility relation* between Θ and S.

Consider the following auditing example. Suppose an auditor is interested in determining whether the accounts receivable of a client are fairly stated. The auditor begins with a frame Θ consisting of two possible states; $\Theta = \{\theta_1,\theta_2\}$, where

$$\theta_1 = \text{accounts receivable are fairly stated,}$$

and

$$\theta_2 = \text{accounts receivable are materially misstated.}$$

The auditor performs a certain analytic review procedure that is relevant to the accounts receivable. Thinking about the relevance of this procedure leads the auditor to consider the frame $S = \{s_1,s_2\}$, where

$$s_1 = \text{the procedure will uncover any material error,}$$

and

$$s_2 = \text{the procedure will not uncover material error.}$$

If the auditor performs the procedure and finds no evidence of material error, then he has established a compatibility relationship between these two frames. The element s_1 is compatible with θ_1; when the procedure is reliable and no error is detected, the account is fairly stated. On the other hand, s_2 is compatible with both θ_1 and θ_2; when the procedure is unreliable and no error is detected, the account might be in material error and might not be.

In general, we write $\Gamma(s)$ for the subset of Θ consisting of elements with which the element s of S is compatible. In our example,

$$\Gamma(s_1) = \{\theta_1\} \text{ and } \Gamma(s_2) = \{\theta_1,\theta_2\} = \Theta.$$

In general, $\Gamma(s)$ can be any subset of Θ, except that it cannot be empty. (Were there no elements of Θ compatible with s, s would be impossible, whence its probability would be zero, and we could omit it from our formulation of the problem.) Different s may have the same $\Gamma(s)$.

Basic Probability Assignments (m-values)

It is the probabilities for elements of S together with the compatibility relation between S and Θ that determine belief-function degrees of belief for Θ. The basic idea is that each probability $\mathbf{P}(s)$ should contribute to a degree of belief in the subset $\Gamma(s)$ of Θ consisting of elements with which s is compatible. If several s have the same $\Gamma(s)$, say $\Gamma(s)$ is equal to B for several s, then the probabilities of all these s will contribute to our degree of belief that the answer to the question considered by Θ is somewhere in B.

For each subset B of Θ, let $\mathbf{m}(B)$ be the total probability for all the s whose $\Gamma(s)$ is equal to B:

$$\mathbf{m}(B) = \sum_{\Gamma(s)=B} \mathbf{P}(s). \qquad \text{(B-1)}$$

We call the mapping m our *basic probability assignment (bpa)*. It follows from this definition that the **m**-values for all the subsets of Θ add to one. In symbols,

$$\sum_{B\subseteq\Theta} \mathbf{m}(B) = 1. \qquad \text{(B-2)}$$

Also,

$$\mathbf{m}(\emptyset) = 0, \qquad \text{(B-3)}$$

where ∅ is the empty set. Conditions (B-2) and (B-3) are sufficient as well as necessary for a mapping to be a basic probability assignment (Shafer 1976).

Let us return to our auditing example and assign probabilities to the elements of S. Suppose, based on his or her experience with the procedure (before applying it to this particular case), the auditor feels that the procedure is 70% reliable and 30% not reliable. Formally, this means that the auditor's probabilities are $\mathbf{P}(s_1) = 0.7$ and $\mathbf{P}(s_2) = 0.3$.

Applying (B-1) with these probabilities, we obtain

$$\mathbf{m}(\theta_1) = \mathbf{P}(s_1) = 0.7$$

and

$$\mathbf{m}(\Theta) = \mathbf{P}(s_2) = 0.3.$$

Since there is no element in S that is compatible only with θ_2,

$$\mathbf{m}(\theta_2) = 0.$$

Of course, $\mathbf{m}(\varnothing)$ is also zero.

If the mapping Γ maps every point in S to a point in Θ rather than to a larger subset—i.e., each element of S is compatible with only one element of Θ—then the **m**-values are simply probabilities for the elements of Θ. The **m**-value for each point in Θ is its probability, and the **m**-values for larger subsets of Θ are all zero. In our case, however, one of the points in S is mapped to a larger subset. Thus the **m**-values are not exactly probabilities; the 70% probability is assigned to the point θ_1, but the other 30% is assigned to the whole frame Θ rather than to θ_2.

In this example, the frame Θ is very small. In other examples, where the frame is much larger, the belief-function structure may still be quite simple, because the **m**-values may be zero for most subsets of the frame. In general, we call the subsets for which the **m**-values are not zero *focal elements*. The **m**-values for the focal elements must add to one. Aside from the requirement that the empty set cannot be a focal element, there is no restriction on what subsets can be a focal elements. Two focal elements can overlap or be disjoint, or one can contain the other.

Belief Functions

We have explained the basic probability assignment, which is one way of representing the mathematical information in a belief function, but we have not yet explained the belief function itself. We reserve the term *belief function* for the function that expresses, for each subset of the frame, our *total* belief in that subset.

In general, our total degree of belief in a subset A of Θ will be more than **m**(A). This **m**-value is the total probability for s that are compatible with everything in A and nothing outside of A. But in order for its probability to contribute to belief in A, it is enough for s to be compatible with some of the elements of A and nothing outside of A. So to get a total degree of belief in A, we should add to **m**(A) the **m**(B) for all subsets B of A. In symbols:

$$\mathbf{Bel}[A] = \sum_{B \subseteq A} \mathbf{m}(B) \qquad \text{(B-4)}$$

We call a function **Bel** defined in this way a *belief function*. It follows from this definition that **Bel**[Θ] =1 and that **Bel**[∅] = 0, where ∅ represents the empty set.

Applying definition (4) to our example, we find the degrees of belief

$$\mathbf{Bel}[\theta_1] = \mathbf{m}(\theta_1) = 0.7,$$

$$\mathbf{Bel}[\theta_2] = \mathbf{m}(\theta_2) = 0,$$

and

$$\mathbf{Bel}[\Theta] = \mathbf{m}(\theta_1) + \mathbf{m}(\theta_2) + \mathbf{m}(\{\theta_1, \theta_2\}) = 1.0.$$

In our auditing example, there are only two focal elements. One of them, $\{\theta_1\}$, was a proper subset of the frame Θ; the other was Θ itself. This type of belief function is very common, and it is convenient to have a name for it. We call a belief function that has at most one proper subset of the frame as a focal element a *simple support function*, and we call the proper subset that is a focal element the *focus* of the simple support function. Thus the belief function **Bel** in our example is a simple support function with $\{\theta_1\}$ as its focus.

Plausibility Functions

Given a belief function **Bel**, we can define another interesting function that we call the *plausibility function* for **Bel**. The plausibility function for Bel is denoted by **PL**, and it is defined by

$$\mathbf{PL}[A] = 1 - \mathbf{Bel}[\text{not } A]. \qquad \text{(B-5)}$$

Intuitively, the plausibility of A is the degree to which A is plausible in the light of the evidence—the degree to which we do not disbelieve A or to assign belief to its negation 'not A'. Complete ignorance or lack of opinion about A is represented by **Bel**[A] = 0 and **PL**[A] = 1.

Belief functions differ from Bayesian probability in representation of ignorance. In Bayesian theory, ignorance is represented by assigning equal probability to all the outcomes. In the belief-function framework, ignorance is represented by a *vacuous* belief function. This belief function assigns an **m**-value of one to the entire frame Θ and an **m**-value of zero to all its proper subsets. This results in **Bel**[A] = 0 and **PL**[A] = 1 for every proper non-empty subset A of Θ.

In general, **Bel**[A] ≤ **PL**[A] for every subset A of our frame Θ. If we believe A, then we think A is plausible, but the converse is not necessarily true.

A zero plausibility for a proposition means that we are sure that it is false (like a zero probability in the Bayesian theory), but a zero degree of belief for a proposition means only that we see no reason to believe the proposition.

Consonant Belief Functions

Shafer (1976) has used *consonant belief functions* for determining the level of belief from the statistical evidence. In this appendix, we discuss how consonant beliefs are related to such evidence.

Suppose f is a real-valued function on the frame Θ, with the property that

$$0 \le f(\theta) \le 1 \text{ for all } \theta \text{ in } \Theta,$$

and

$$f(\theta) = 1 \text{ for at least one } \theta \text{ in } \Theta.$$

Then, as it turns out, the function **Bel** defined by

$$\mathbf{Bel}(A) = 1 - \max_{\theta \in \text{notA}} f(\theta)$$

for each non-empty A is a belief function. Its plausibility function is given by

$$\mathbf{PL}(A) = \max_{\theta \in A} f(\theta), \tag{B-6}$$

and therefore the plausibility for a single point θ is simply

$$\mathbf{PL}(\theta) = f(\theta). \tag{B-7}$$

A belief function that is defined in this way is called a *consonant* belief function (Shafer 1976). We can restate (B-6) and (B-7) by saying that the plausibility of a singleton θ is f(θ), and the plausibility of a non-singleton set is the largest plausibility of any of its elements.

A consonant belief function can be described in the continuous case, just as in the discrete case, by a function f that gives the plausibilities for singletons. The plausibilities of larger subsets are then given by

$$\mathbf{PL}(A) = \sup_{\theta \in A} f(\theta). \tag{B-8}$$

This formula is the same as (B-1), except that we write “sup” instead of “max.” This is because there may be an infinite sequence of θ in A such that the numbers f(θ) approach an upper limit without actually reaching it. In this case, the *maximum*, technically, does not exist, but the *supremum* will be the upper limit. The consonant belief functions that we use in this article arise from nested confidence intervals or nested likelihood intervals.

As we have explained above that a consonant belief function **Bel** can be completely specified by specifying a function f that gives the plausibility of singletons:

$$\mathbf{Bel}(A) = 1 - \mathbf{PL}(\text{notA}) = 1 - \sup_{\theta \in \text{notA}} f(\theta).$$

For the sake of clarity, we will now give formulas for f for the consonant belief function based on the likelihood method.

In the case of the likelihood method, f is simply the likelihood, renormalized by dividing by its maximum value, so that its new maximum is one:

$$f(\mu) = \frac{L(\mu)}{L(\bar{y})} = \exp[-\frac{n(\bar{y} - \mu)^2}{2\sigma^2}].$$

Thus the degree of belief for any interval $B = (\mu_1, \mu_2)$ is given by

$$\mathbf{Bel}(B) = 1 - \mathbf{PL}(\text{not}B) = 1 - \sup_{\mu \in \text{not}B} \exp[- \frac{n(\bar{y}-\mu)^2}{2\sigma^2}]. \qquad \text{(B-9)}$$

If $\bar{\mathbf{y}}$ does not fall in B, then (B-9) gives

$$\mathbf{Bel}(B) = 0. \qquad \text{(B-10)}$$

If $\bar{y}$ does fall in B, then (B-9) yields

$$\mathbf{Bel}\,(B) = 1 - \exp\left(- \frac{n(\bar{y} - \mu_i)^2}{2\sigma^2}\right), \qquad \text{(B-11)}$$

where μ_i is the one of the pair μ_1 and μ_2 that comes closest to $\bar{y}$.