4-1-2005

Additivity of information value in two-act linear loss decisions with normal priors

Jeffrey Keisler
University of Massachusetts Boston, jeff.keisler@umb.edu

Additivity of information value in two-act linear loss decisions with normal priors.

Jeffrey M. Keisler
Management Sciences and Information Systems Department

College of Management
University of Massachusetts Boston
5th Floor McCormack Hall, Room 230
100 Morrissey Boulevard
Boston, MA 02125
jeff.keisler@umb.edu
phone: 617-287-7738
fax: 617-287-7787

In preparation for Risk Analysis.

Additivity of information value in two-act linear loss decisions with normal priors.

***Abstract**: For the two-act linear loss decision problem with normal priors, conditions are derived for which the expected value of perfect information about two independent risks is super-additive in value. Several applications show how a variety of decision problems can reduce to the canonical problem, and how the general results obtained here can be translated simply to prescriptions for specific situations.*

***Keywords**: Probability, uncertainty, value of information.*

1. Introduction

The likelihood of extreme events can increase when there are multiple uncertainties. Does this mean that the value of information increases in the same proportion?

In a simple two-act linear loss decision problem under uncertainty, the decision maker (DM) has a choice between doing nothing and taking some action, A and receiving an uncertain value x. The DM chooses A if the E(x) > 0. The actual value x typically differs from E(x), but if the x >= 0, then nothing has been lost by choosing A. If the x < 0, then for every additional dollar by which the x differs from E(x), the DM loses an additional dollar. Many real decisions are in this class.

If the DM can obtain information about x prior to deciding, then downside risk can be reduced. Specifically,if the best estimate of x is less than 0, then the decision changes – A is rejected and an expected loss is averted. Otherwise, the decision does not change and no loss is averted. The expected value of perfect information (EVPI) is the expected loss averted if perfect information becomes available. If there is only one possible outcome to the uncertainty that would switch the decision, the value of information is the product of the probability of a switch and the loss averted when the switch occurs (Behn & Vaupel, 1982). When the uncertainty is a continous probability distribution over x, the analogous formula is -INTEGRAL –infinity to 0 of x fx dx, where f is the probability density function. If the distribution, i.e., **fx = exp(-1/2 mu/sigma etc, t**he resulting formula is called normal linear loss integral, and its behavior was well characterized by Raiffa and Schlaifer (1961).

Results about EVPI are known to be difficult to generalize (Hilton, 1981), but specific analyses have yielded substantial qualitative insights. Felli and Hazen (1998) demonstrated that we get significant improvement in estimates of EVPI when the entire decision context is considered, rather than simply the one-way sensitivities of the value of an alternative to each uncertain variable. Hammitt and Shlyakhter (1999), among other things, explored ways in which the interaction of variables affects value of information. In particular, they showed how having different prior information can drastically affect the value of new information about a variable (in two-act linear loss decisions. They went so far as to consider the value of information (**in a ___ decision** – something other than TALL?) about variables having a multiplicative relationship, i.e., x = yz, essentially finding that percentage error matters. Clearly, one factor that often raises or lowers the EVPI about an uncertainty is the presence of other information about other uncertainties.

Thus, if we add a second uncertainty or risk, extreme outcomes (and downsides), become more likely. For example, a lender may have concerns about two risks facing a borrower, where if both risks turn out badly, default is much more likely. In such cases, it may be desirable to obtain all potentially available information so as to avoid the downside. Conversely, it might be that the safety mechanism (not acting) is already adequate. In such cases, an appropriate switch in the decision may be so likely, even under partial information, that the incremental value of having complete information is small. If information is costly, as with experiments or polls, then it is important to know which of these cases obtains. That knowledge would improve the allocation of risk management resources applied to information acquistion.

We can ask when is the value of information about two such uncertainties super-additive and when is it sub-additive, i.e., what makes the value of information about two sources of uncertainty combined greater than the sum of the value of information from each source independently.. This question was addressed with some success by Samson *et al*. (1989). They considered decisions with discrete probabilities on two separate events each with two discrete outcomes. That paper left as an open question the additivity of EVPI with the continuous distributions that may better characterize many uncertainties.

In this paper, we shall consider the additivity question for the normal distribution, both as first case of continuous distributions, and an important distribution in its own right. EVPI here is driven by summary statistics of mean μ and standard deviation σ. We find both a simple approximation and a precise condition for super-additivity of the expected value of information about independent events in two-act linear loss decisions with normal prior distributions.

A more concrete motivation is to guide practice in decision and risk analysis. The value of analysis has been modeled as value of information, as in Watson and Brown (1984), i.e., where there are variables with estimated values and the estimates are improved by using assessment techniques. In planning an analytic intervention, questions often arise regarding how much analysis should be done (e.g., decomposition). In a related vein, super-additive information value has been posited as an explanation for certain organizational structures by Milgrom and Geanakoplos (1991).

2. PROBLEM STATEMENT

This section considers the following decision problem: A decision maker has a choice between receiving \$x or refraining from action and receiving \$0, where $x = x_1 + x_2$, x_1 and x_2 are independent, $x_1 = \mu_1 + \varepsilon_1$, $x_2 = \mu_2 + \varepsilon_2$, ε_1 follows normal distribution with mean 0 and standard deviation σ_1, and ε_2 is independent of ε_1 and is also normal with mean 0 and standard deviation σ_2. Let μ denote $\mu_1 + \mu_2$. The case where the choice is instead between \$$x_1$ and \$$x_2$ is similar.

The decision maker has the information gathering options of either: purchasing no information; purchasing perfect information about x_1 (or x_2) with an option to later purchase information about x_2 before making the decision (paying a premium for the flexibility); or purchasing perfect information about x_1 and x_2 right away without paying a premium.

Let V_1 denote the EVPI on x_1 alone, V_2 denote the EVPI on x_2, and V_{12} denote the EVPI on x_1 and x_2 together. Assume that information about both uncertainties is fairly priced, that is, V_1 is equal to the cost of information about x_1, and likewise for V_2 and x_2. When we consider implications of the EVPI results, we shall also assume that the decision maker wishes to maximize expected monetary value. This is reasonable when the value and costs involved are small relative to the wealth of the decision maker (or, often, of the decision maker's employer).

Almost by definition, if the EVPI about x_1 and x_2 is super-additive ($V_{12} > V_1 + V_2$), the decision maker would tend toward an all-or-nothing approach to information acquisition. Information would be obtained about both uncertainties or neither, but not about one and not the other. If the EVPI is close to additive, then the decision maker can purchase one piece of information and perhaps purchase the other later in the (somewhat unlikely) event that the value of the second piece of information conditional on the outcome of the first is substantially increased. If the EVPI is sub-additive, the decision maker may be willing to preclude the opportunity to get the second piece of information.

For the current problem, Raiffa and Schlaifer's analysis gives directly relevant discussion of the general behavior of EVPI curves in sampling problems for two-act decisions. This paper uses their notation where f is the normal probability density function and G is 1 minus the cumulative probability density function. The subscripts N and N* denote that the functions are for normal or standard normal distributions, respectively, and these subscripts are often suppressed. Raiffa and Schlaifer analyze EVPI as a function of the number of observations of a normal process, yielding the following key fact: EVPI = Max EV achieved with PI – Max EV achieved without PI. Assume $\mu > 0$. Without perfect information, the expected value is Max(E(x),0) = μ. With perfect information about x_1 only, $E(x) = x_1 + E(x_2) = x_1 + \mu_2$. If x_1 is to be revealed, the expected value to the decision maker before learning x_1 increases to $V_1 = E[\mathrm{Max}(x_1 + \mu_2, 0)]$.

Now, let $u_1 = \mu/\sigma_1$. Then V_1 is equal to the linear loss integral $L_{N*}(u_1)$ as in Raiffa and Schlaifer, and the above expectation works out to

$$= \sigma_1[f_{N*}(u_1) - u_1 G_{N*}(u_1)] \qquad (1)$$

The formula for V_2 is the same, of course, with a change in subscripts. If we let

$$\sigma_{12} = \sqrt{(\sigma_1^2 + \sigma_2^2)}. \qquad (2)$$

Then,

$$V_{12} \quad = \sigma_{12}[f_{N*}(u_{12}) - u_{12} G_{N*}(u_{12})] \qquad (3)$$

What drives EVPI is the quantity $E(x_1+x_2)$, and the variance of this quantity is the sum of the variances on its parts. The question of super-additivity comes down to the question of whether the tails of the two small bell curves (for information about x_1 or x_2) on the value of x are together "larger" than the tail of the wider bell curve corresponding to perfect information about both x_1 and x_2.

Raiffa and Schlaifer's analysis provides intuition about a closely related problem. A graph showing the expected value of information in a sampling problem is an S-shaped function of the number of observations. The first few observations have almost no value because they are generally insufficient to switch the decision. Value then increases for a while as observations are added. Eventually there are enough observations and further observations do not add value because if the decision were going to switch, it most likely would have by this point. Considering two sets of observations of size m and n, if the slope of the graph of expected value of information vs. number of observations from 0 to m observations is less than the slope from m to m+n observations, then the information from the two sets of observations is super-additive in value (and it does not matter if m and n are switched here).

In our current problem, the drivers of additivity in value of information are similar. it is known that in "toss-up" situations, where μ is close to 0, or σ_1 and σ_2 are large, a single piece of information can be enough to decide on one alternative or the other, and the second piece of information is less likely to make a difference. In "long-shot" situations, where μ is large or where σ_1 and σ_2 are small, it is hard for the first piece of information to swing the decision, but two pieces of information might. However, intuition does not tell precisely what the words "small," "large," "close" and "far" mean in this paragraph, e.g., is 2 small? Therefore, concrete results are needed when conditions are not so extreme as to make the additivity situation obvious.

3. THE ADDITIVITY RATIO

To make the intuitive statements from the last section more precise, a spreadsheet was used to plot the additivity ratio defined by Samson et al, $V_{12}/(V_1+V_2)$, against σ_1 for different values of $\mu = \mu_1-\mu_2$, while holding σ_2 fixed at 1. Figure 1 shows the additivity ratio for several values of μ and σ terms.

---- INSERT FIGURE 1 ABOUT HERE -----

This curve (viewed as a vertical slice of the surface in Figure 1) can have two shapes. For large μ, the ratio starts at 1 (where $\sigma_1 = 0$), because the smaller piece of information (i.e., the one with smaller σ) is likely to be too small to make a difference once the larger piece of information has changed the expected value and moved it farther from 0. The additivity ratio then increases to a maximum at $\sigma_1 = \sigma_2$, and then decreases asymptotically to 1 again as σ_1 takes on the role of the larger piece of information. This maximum increases without bound as μ, increases and σ_2 is held constant, e.g., for $\mu = 3$, $\sigma_1 = \sigma_2 = 1$, the additivity ratio is 12.5. For small μ, the ratio still starts at 1, increases for a short while (or not at all if $\mu = 0$), then decreases to below 1, reaches a minimum near $\sigma_1 = \sigma_2$, and returns asymptotically toward 1. For a given σ_2, and for a large enough value of μ, the switch from super-additive to sub-additive looks as though it will never occur. As shown later, the level of σ_1 where the switch does occur, increases without bound as a function of μ. A couple of observations: in the first case the additivity ratio is everywhere greater than 1 and increasing in μ, and increases in σ_1 until a level that is also increasing in μ. In the second case, the minimum point of the additivity ratio is decreasing in μ, and reaches its lowest possible value of $1/\sqrt{2}$ for $\mu = 0$, where $(V_{12})^2 = (V_1)^2 + (V_2)^2$. This is consistent with the notion that toss-up decisions lead to the most sub-additivity, and it holds because at $\mu = 0$, V is exactly proportional to standard deviation.

The fact that the additivity ratio is lowest when $\sigma_1 = \sigma_2$ is to some extent an artifact of the way we have defined the ratio. An alternate measure could be $(V_{12} - V_1)/V_2$, the ratio of the expected incremental value of resolving the second uncertainty after the first has been resolved to the expected value of resolving the second uncertainty alone. With either additivity measure, a value above 1 implies super-additivity and a value under 1 implies sub-additivity. The value of the alternate measure varies more both below 1 and above 1, e.g., it is 0.414 (= $\sqrt{2}$-1), when $\mu = 0$ and $\sigma_1 = \sigma_2$, but as σ_1 grows relative to σ_2 the ratio approaches 0, and as σ_2 grows relative to σ_1 the ratio approaches 1.

4. A CONDITION FOR SUPER-ADDITIVITY

The behavior of the additivity ratio curve seems systematic enough to justify a search for the rules that govern it. First, more notation is needed:

$$u_1 = \mu/\sigma_1; \tag{4}$$

$$u_2 = \mu/\sigma_2; \tag{5}$$

$$u_2^+ = \text{the value of } u_2 \text{ at which EVPI is additive;} \tag{6}$$

$$S = \sigma_1/\sigma_2; \tag{7}$$

$$R = \sqrt{(1+S^2)}; \tag{8}$$

and

$$u_R = u_2/R. \tag{9}$$

Constraining the additivity ratio to be equal to 1 and then taking the full differential in terms of u_2^+ and S yields the following equation derived below:

$$du_2^+/dS = [f(u_1)- (u_1/u_R)\ f(u_R)] / [G(u_2)+G(u_1)-G(u_R)] \tag{10}$$

Derivation: The boundary between sub and super-additivity will be found by constraining u_2^+ so that $V_{12} = V_1 + V_2$. We could equivalently constrain u_1, or μ. Letting L_{N*} L denote the unit normal linear loss integral, writing out the V terms yields

$$\sigma_2 R \times L(u_2/R) = \sigma_2 S \times L(u_2/S) + \sigma_2 L(u_2), \tag{11}$$

$$\rightarrow \quad R \times L(u_2/R) - S \times L(u_2/S) - L(u_2) = 0. \tag{12}$$

Letting Q represent the left hand side of the equation 12, full differentiation of Q will identify du_2^+/dS.

Recall from Raiffa and Schlaifer that,

$$L_{N*}(u) = f_{N*}(u) - uG_{N*}(u), \text{ and } dL_{N*}(u)/du = -G_{N*}(u). \tag{13}$$

We shall now suppress the subscript $_{N*}$. We set

$$(\partial Q/\partial u_2)du_2 + (\partial Q/\partial S)dS = 0, \tag{14}$$

$$\rightarrow \quad du_2^+/dS = -(\partial Q/\partial S)/(\partial Q/\partial u_2). \tag{15}$$

Defining $u_R = u_2/R$, and noting that $u_1 = u_2/S$ yields,

$$\partial Q/\partial u_2 = R(\partial L(u_R)/\partial u_2) - S(\partial L(u_1)/\partial u_2) - \partial L(u_2)/\partial u_2, \tag{16}$$

$$= \quad \partial L(u_R)/\partial u_R - \partial L(u_1)/\partial u_1 - \partial L(u_2)/\partial u_2. \tag{17}$$

$$= \quad -G(u_R) + G(u_1) + G(u_2). \tag{18}$$

Next, it is necessary to obtain the other half of the differential:

$$\partial Q/\partial S = (\partial R/\partial S)\ L(u_R) + R\ (\partial L(u_R)/\partial S) - L(u_1) - S\ (\partial L(u_1)/\partial S). \quad (19)$$

$$\rightarrow\ (S/R)\ L(u_R) + R\ (-u_2/R^2)\ (S/R)\ \partial L(u_R)/\partial(u_R) - L(u_1) + S(u_2/S^2)\ \partial L(u_1)/\partial(u_1). \quad (20)$$

Completing the last two partial derivatives yields,

$$(S/R)\ L(u_R) + R(-u_2/R^2)\ (S/R)\ (-G(u_R)) - L(u_1) + S(u_2/S^2)\ (-G(u_1)) \quad (21)$$

$$= \quad (S/R)\ L(u_R) + u_2S/R^2\ G(u_R)) - L(u_1) - u_1(G(u_1)). \quad (22)$$

Expanding the L terms yields,

$$(S/R)\ [f(u_R) - u_RG(u_R)] + u_2(S/R^2)\ G(u_R) - [f(u_1) - u_1G(u_1)] - u_1G(u_1)] \quad (23)$$

$$= \quad (S/R)\ f(u_R) - f(u_1). \quad (24)$$

Replacing S/R with u_R/u_1, the total differential du_2^+/dS from (15) is,

$$[f(u_1) - (u_R/u_1)f(u_R)]\ /\ [G(u_1) + G(u_2) - G(u_R)]. \quad (25)$$

The differential equation for u_2^+ can be graphed numerically without a neat closed form solution. In this figure 2, σ_2 is set at 1, but for different values we would simply scale the x and y axes by a factor of σ_2. Note that that du_2^+/dS contains only "f" terms in the numerator and "G" terms in the denominator, which is reminiscent of the hazard function, usually denoted $H(u) = f_N(u)/G_N(u)$. Figure 2 graphs u_2^+ against S (or, if units are defined so that $\sigma_2 = 1$, the graph is of μ against σ_1). The starting point is $u_2^+(0) = 0$, because it is clear that in a sampling problem where $\mu = 0$ there are no values for the σ terms that lead to super-additivity.

----- INSERT FIGURE 2 ABOUT HERE -----

5. AN APPROXIMATION FOR SUPER-ADDITIVITY

Calculating the integral from the previous section would be inconvenient in practice, and it could help to use an approximation for the solution. One possibility is to focus on the similar structure of the numerator and denominator in (25), and, noting that $\int(h'/h) = \ln(h)$ for some function h, try something like $\ln[[-G(u_1)-G(u_R)]]$. Unfortunately, because of the dissimilarities in the numerator and the denominator, nothing of this sort proves useful. Another possibility works better. We start with our definition of additivity, $V_{12} = V_1 + V_2$. Noting that $V_{12} = \sigma_{12}L(u_{12})$, $V_1 = \sigma_1 L(u_1)$ and $V_2 = \sigma_2 L(u_2)$. Let us temporarily fix $\sigma_1 = \sigma_2 = 1$. Then the additivity border is $\sigma_{12}L(u_{12}) = 2\sigma_1 L(u_1)$. Because $\sigma_{12} = \sqrt{2}\sigma_1$ (and therefore, $u_{12} = u_1/\sqrt{2}$) we can say that additivity occurs when $\sqrt{2}\sigma_1 L(u_{12}) = 2\sigma_1 L(u_1)$, that is, where $L(u_{12}) = \sqrt{2}L(\sqrt{2}u_{12})$, which occurs at $u_{12} \approx 0.5$.

No longer fixing the values of σ_1 and σ_2, we find that for S within a factor of ≈ 3 of 1.0 (i.e., between 0.35 and 2.95), the equivalent approximation $u_2^+ = \sqrt{(S/2)}$, is accurate to within 5% of the actual u_2^+. Above that u_2^+ flattens out more quickly than this approximation. Near 0, du_2^+/dS, for both the actual function and the approximation, approaches infinity. It may be more convenient in some circmstances to rewrite the approximation above as one of the following:

$$\mu^+ = \sqrt{(\sigma_1\sigma_2/2)}, \tag{26}$$

or equivalently,

$$(u_1u_2)^+ = 1/2. \tag{27}$$

The form for u_2^+ is close enough to linear that the approximation holds over a substantial range. For any $\mu > 0$, when S is small enough, there is super-additivity, and for any S, when μ is large enough there is super-additivity. Note, if x_1 and x_2 are not judgmentally independent, (e.g., if they are bivariate normal), the curve for super-additivity would lie above this one by an amount increasing in the correlation between the two variables.

This approximation has an appealing intuitive interpretation. The right side of the approximation is the geometric mean of the two standard deviations, modified by a constant. This sheds light on the earlier assertion that EVPI is sub-additive when the standard deviations are large and super-additive when they are large. It appears that the terms small and large do not mean very small and very large. On the other hand, when there is super-additivity or sub-additivity in value between one quite small experiment and one quite large one, the magnitude of the effect and its value in planning analyses is smaller than when σ_1/σ_2 is near 1.

An exogenous variable, k, may have affect the mean and the standard deviations. We may then ask for what value of k would information value be superadditive. The general first-order case is where $\mu = ak+b$, $\sigma_1 = ck+d$, and $\sigma_2 = ek+f$. Substituting these equations into (27), with a bit of manipulation we get an approximation for superadditivity that is the following quadratic function of k: EVPI is super-additive if

$$(2a^2-ce)k^2 + (4ab-de-fc)k + (2b^2-df) > 0. \qquad (28)$$

This equation has two solutions in k (not necessarily both positive), implying that as k increases, EVPI may switch from sub-additive to super-additive and back, which can happen if σ_1/σ_2 starts out greater than 1 and then decreases to less than 1 as k increases.

6. NUMERICAL EXAMPLE

The following example illustrates the use of this result for several sets of numerical parameter values. Consider a company that has the opportunity to build a plant for a known fixed cost F in order to sell a fixed quantity Q of a new product, and facing risk on both cost and price. The price is uncertain P and the cost C is uncertain. Before deciding to build the plant, the marketing department has asked for extra market research funding to better estimate P, while the manufacturing department wants funding to commission an engineering study to better estimate C.

The decision maker wonders whether either study should be funded, or both, or neither. The decision maker considers P and C to be independent, and assigns the pre-posterior distributions P ~N(E(P), σ_P), and C ~N(E(C),σ_C). If P and C were known, the profit from this opportunity would be easily calculated as $\pi = (P-C) \times Q - F$. Because the relationship here is linear and additive,

$$E(\pi) = (E(P)-E(C)) \times Q - F. \qquad (30)$$

The preposterior standard deviation of the mean profit estimate with respect to the market research study is $\sigma_1 = \sigma_P \times Q$, and for the engineering study $\sigma_2 = \sigma_C \times Q$. In this case, the value of one alternative is known with certainty, while there are two pieces of information relevant to the value of the other alternative. This decision structure has the same conditions for super-additivity as the one in which there are two alternatives each involving one uncertainty.

This example illustrates the use of the approximation, predicting sub- (super-) additivity when $u_2 < (>) \sqrt{(S/2)}$. It also illustrates some of the major characteristics of the behavior of the additivity ratio with normally distributed variables. We consider cases 1-6 in which, respectively: 1) both uncertainties are small, 2) where both uncertainties are large, 3) one uncertainty is large and one is small, 4) Parameters as in the previous case, but standard deviations are the same and set so that V_{12} is approximately the same as in the previous case, 5) $\mu = 0$, standard deviations are not equal, and 6) $\mu = 0$, standard deviations are equal. The numerical assumptions for each case are shown in the first table, and the predictions and results are shown in the second table.

----- INSERT TABLE 1 ABOUT HERE -----

----- INSERT TABLE 2 ABOUT HERE -----

Examples such as these can be used to strengthen intuition about value of information from multiple sources. Note, in case 3, the alternative ratio $(V_{12}-V_1)/V_2$ gives a stronger measure of super-additivity (3.970).

7. OTHER APPLICATIONS

The results in this paper could also be relevant to questions involving use of information from multiple sources, e.g., organizational structure or bundling and pricing of information products. Some possibilities along these lines follow. The conditions for the super-additivity approximation and formal condition are in general form. The fact that the conjugate distribution for the normal is itself the normal distribution allows many common problems to reduce to the form for which these results apply.

1) Variables that are sums of other variables: Consider the case in which there are two alternative, each the sum of independent, normally distributed variables, that is $x_1 = \Sigma y_{i \varepsilon I}$, $x_2 = \Sigma y_{j \varepsilon J}$. To apply the results, we would assign $\sigma = | x_1 - x_2 |$. If I and J do not overlap and K and L are non-overlapping subsets of I ∪ J, then $\sigma_1^2 = \Sigma \sigma_{i\ \varepsilon\ K}^2$, and $\sigma_2 =$

$\Sigma\sigma_{i\ \varepsilon\ L}{}^2$. Thus, the additivity condition applies to value of information about any partition of K, L.

This could occur in a situtation like the following: There is an intervention that can be applied in a number of countries. The intervention will be applied in either the Eastern Hemisphere (I) or the Western Hemisphere (J), and the benefit from either strategy will be the sum of the benefit in each of the countries in the targeted region. Demographic research that would allow precise estimates of benefits within a given country, on the other hand, is most convenient to do in English speaking countries (K) or Spanish speaking countries (L). By translating to the conditions of the additivity approximation, one can quickly identify whether the two sets of demographic research activities are super-additive in value, with respect to the decision about where to intervene. It would also be possible to predict the additivity of the value of market research as the partitions of I and J and of K and L are toggled simultaneously. This structure could even be combined with something like the numerical example, where research is done on the cost and benefit of the intervention in each region, where cost and benefit are normally distributed so that the net value in each region is still normal.

2) Sampling: Another example is the sampling problem discussed at length by Raiffa and Schlaifer, in which it is possible to take a sample of n observations (with standard deviation σ), m observations, or both. A possible action has value that correlates with an uncertain parameter, e.g., $y = ax - b$. Two possible samples about x may be taken to resolve some of the uncertainty. The approximation would only work in this case if the number of prior observations, k, is large so that the pre-posterior variance per observation stays nearly constant at σ^2/k. In this case, $\sigma_1^2 = m(\sigma^2/k)$, and $\sigma_2^2 = n(\sigma^2/k)$, so the two samples are super-additive in value if (approximately)

$$\mu > (\sigma/k)\sqrt{2}(nm)^{1/4}. \tag{31}$$

3) Mixed discrete and continuous probability distributions: Another example is a variation on the problem analyzed by Samson *et al*. In a choice between X and 0, where the value of X has one normal distribution in one state of nature (with a mean value of $E(X_{1)}$ in state ω_1) and another normal distribution in a second state of nature (mean $E(X_2)$ in state ω_2), and the probabilities (p and 1-p) of these two states of nature are known, the problem reduces to the first example above, where

$$\mu = p\ X_1 + (1\text{-}p)\ X_2, \tag{32}$$

while

$$\sigma_1^2 = p\ (\sigma^2|\omega_1), \tag{33}$$

and

$$\sigma_2^2 = (1\text{-}p)\ (\sigma^2|\omega_2). \tag{34}$$

Here, the super-additivity approximation can be stated as a function of p: Information about the value of X in state 1 and information about the value of X in state 2 is super-additive in value if:

$$[p(X_1)+(1\text{-}p)(X_2)]/[(p)(1\text{-}p)]^{1/4} > \sqrt{[(\sigma|\omega_1) \times (\sigma|\omega_2)]}. \tag{35}$$

This result could be combined with Samson et al's discrete conditions to determine whether to analyze resolve the discrete chance nodes or uncertainty about endpoint value distributions or both, for either or both states.

8. CONCLUSIONS

There are three main findings in this paper. The first finding is a full characterization of the behavior of the additivity ratio curves for two-act linear loss decisions with normal priors. The fact that the additivity ratio is no less than $1/\sqrt{2}$ and unbounded from above suggests that it may be more beneficial than one might intuit to obtain information about multiple uncertainties. The second finding is the differential equation defining the formal condition for super-additivity. The last finding is the approximation for this condition, which is more compact and accurate than might be expected. These specific results have benefits over the intuitive directional statements that could already be made regarding the additivity of value. The approximation is easy to remember and to compute by hand. This facilitates identification of super-additivity in cases where the situation is not so extreme as to make it obvious.

By exploiting well-known characteristics of normally distributed variables, a variety of problems can be translated to this template in order to analyze information value super-additivity. Such analysis can have implications for risk management problems involving the collection of information from multiple sources. The results can determine when it makes sense for decision makers to use limited information and when it makes sense for them to seek extensive information if they are going to use information at all.

References:

Anderson, R.M., and Hobbs, B.F., "Using a Bayesian Approach to Quantify Scale Compatibility Bias," Management Science, **48,** 1555-1568 (2002).

Behn, R. D. and Vaupel, J. W. (1982). *Quick analysis for busy decision makers.* New York: Basic Books.

Epstein, B.J., and King, W.R., "An Experimental Study of the Value of Information," *Omega* **10,** 249-258 (1982).

Felli, J. C., and Hazen, G. B., "Sensitivity analysis and the economic value of perfect information," *Medical Decision Making*, **18,** 95-109 (1998).

Hammitt, J.K., and Shlyakhter, A.I., "The Expected Value of Information and the Probability of Surprise," *Risk Analysis*, **19,** 135-152 (1999).

Hilton, R. "The Determinants of Information Value: Synthesizing Some General Results," *Management Science*, **27,** 57-65 (1981).

Matheson, J. E., "The Economic Value of Analysis and Computation," *IEEE Transactions on Systems Science and Cybernetics*, **88C-4,** 797-804 (1968).

Milgrom, P., and Geanakoplos, J., “A theory of managerial hierarchies based on limited attention,” *Journal of the Japanese and International Economies*, **5,** 205-225 (1991).

Raiffa, H., and Schlaifer, R., *Applied statistical decision theory* (Harvard University Graduate School of Business Administration, Boston, 1961).

Samson, D., Wirth, A., and Rickard, J., “The value of information from multiple sources of uncertainty in decision analysis,” *European Journal of Operational Research*, **39,** 254-260 (1989).

Thon, D. and Thorlund-Peterson, L., “The value of perfect information in a simple investment problem.” *Information Economics and Policy*, **5,** 51-71 (1993).

Watson, S.R., and Brown, R.V., "The Valuation of Decision Analysis," Journal of the Royal Statistical Society, **A, 141,** 69-78 (1978).

Figure 1: The additivity ratio $V_{12}/(V_1+V_2)$

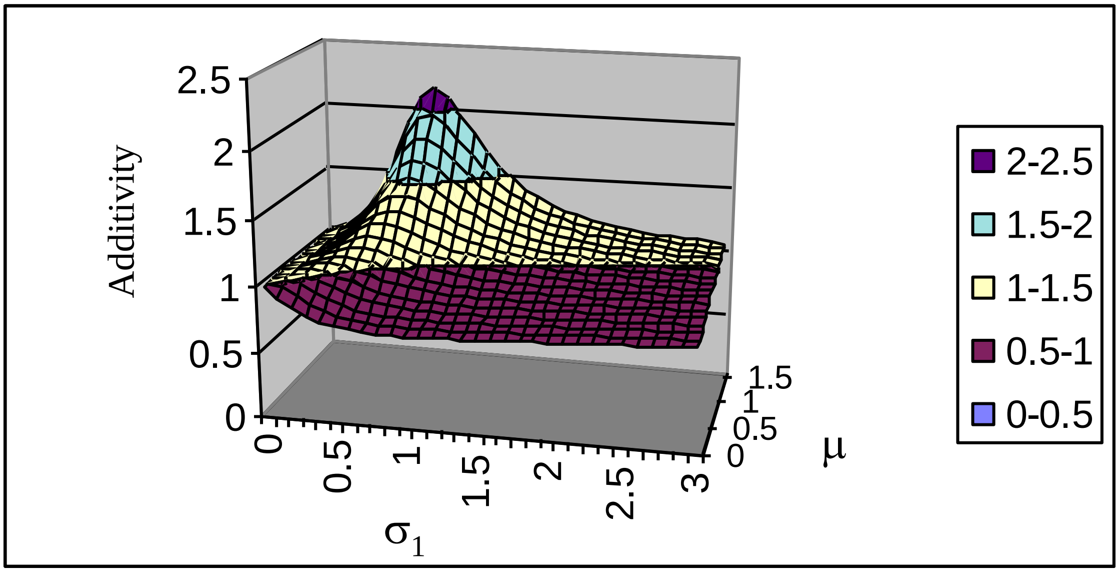

Figure 2: The boundary between super-additivity and sub-additivity ($\sigma_2 = 1$)

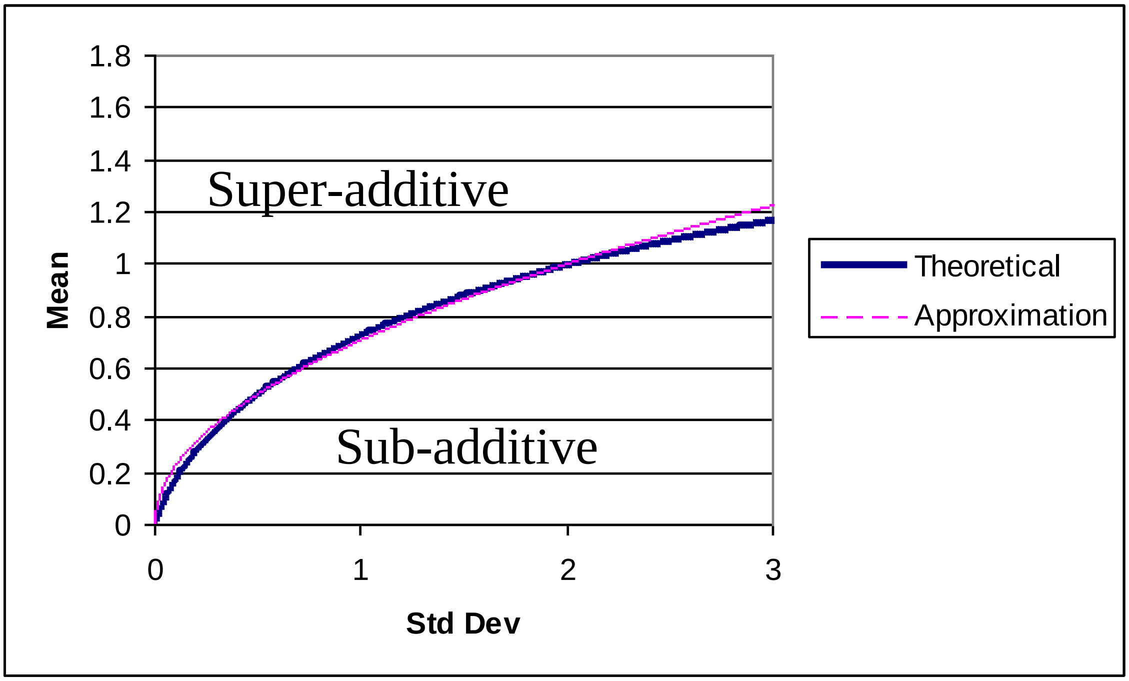

TABLE 1: Assumed values for numerical example.

Case	E(P)	σ_P	E(C)	σ_C	Q	F
1	35	1	30	1	20	45
2	35	10	30	10	20	45
3	35	1	30	5.5	15	45
4	35	4	30	4	15	45
5	35	5	30	1	10	50
6	35	4	30	4	10	50

TABLE 2: Predictions and results for numerical example.

Case	u_2	S	√(S/2)	Prediction	V_1	V_2	V_{12}	Ratio
1	2.75	1	0.707	Super-additive.	0.018	0.018	0.278	7.738
2	0.275	1	0.707	Sub-additive	55.287	55.287	87.465	0.791
3	0.364	0.182	0.302	Super-additive	0.127	20.065	20.573	1.019
4	0.354	1	0.707	Sub-additive	11.868	11.868	29.945	0.882
5	0	5	1.581	Sub-additive	19.947	19.947	20.342	0.850
6	0	1	0.707	Sub-additive	15.958	15.958	22.568	0.707