"THE IMPACT OF TESTING ERRORS ON VALUE OF INFORMATION: A QUALITY-CONTROL EXAMPLE"

by

A. GABA*
and
R.L. WINKLER**

93/84/TM

* Associate Professor of Decision Sciences, INSEAD, Boulevard de Constance, Fontainebleau 77305 Cedex, France.

** Professor at, Fuqua School of Business, Duke University.

Printed at INSEAD, Fontainebleau, France

THE IMPACT OF TESTING ERRORS ON VALUE OF INFORMATION:

A QUALITY-CONTROL EXAMPLE

Anil Gaba
INSEAD

and

Robert L. Winkler
Fuqua School of Business
Duke University

November 1, 1993

Abstract

In this paper, we extend recent work on the inferential impact of errors in data to a decision-making setting. In the context of a simple quality-control example, we illustrate how errors can cause substantial reductions in the value of information from a sample and uncertainty about error rates can lead to yet further reductions in EVSI. Moreover, we extend the notion of an equivalent error-free sample size (which indicates the reduction in effective sample size due to errors) from an inferential framework to a decision-making framework and find that as uncertainty about error-rate parameters increases, reductions in effective sample size are even greater for a decision maker than the inferential measures suggest.

1. Introduction

When a sample is taken from a dichotomous process, various sources of error may cause some items to be classified incorrectly. For example, the existence of errors in survey data is well documented and widely discussed (e.g., Groves, 1989). Errors can also occur in other contexts, such as observations of good and defective items in a quality-control setting or positive and negative results in medical tests. Not only may the testing of items or patients be imperfect, but possible mistakes in recording, coding, and handling the data add to the problem.

Standard, off-the-shelf procedures for making inferences about a proportion such as the proportion of defective items from a manufacturing process are based on the assumption of a Bernoulli process. Extensions of the Bernoulli model to allow for possible misclassification errors are straightforward, but making inferences is more complicated because of the inclusion of error-rate parameters (Winkler and Gaba, 1990; Gaba and Winkler, 1992). Probability distributions for the error-rate parameters play important roles in identifying the model and influencing the ultimate inferences about the proportion of interest. Even if these parameters are known, their presence can lead to large shifts in estimates and substantial increases in the uncertainty about the proportion. The increases in uncertainty can be viewed in terms of loss of information that reduces the effective sample size. A sample of size n with errors may only be as informative as a much smaller sample without errors. Moreover, the fact that there is often only limited evidence regarding the error parameters can exacerbate this loss of information.

In a decision-making problem, the primary measure of interest when obtaining new information is not some inferential measure of informativeness. Decision analysis provides a formal measure, the expected value of information, that evaluates the new information directly in terms of how much better off the decision maker is likely to be with the information than without the information. In deciding whether to adjust a machine before running a large batch of items, for example, a quality-control manager might contemplate taking a sample of output from the machine to obtain more information about the proportion of defectives currently being produced. The expected value of information from a sample of any given size can be computed and compared with the cost of the sample to determine whether and how much to sample. The impact of possible errors in data from samples in the context of the manager's decision-making problem can be studied by investigating how the possible errors affect the expected value of sample information.

In this paper, we show how errors in dichotomous data can reduce the expected value of information. The situation we consider is the quality-control example from the previous paragraph. The decision-making problem is presented briefly and the inferential model with errors is summarized in Section 2. The value of information under different prior distributions is determined in Section 3. In Section 4, we discuss the notion of "equivalent error-free sample size," comparing inferential and decision-making viewpoints. Section 5 concludes with a brief summary and discussion.

2. To Adjust or Not to Adjust?

Suppose that a manager is about to fabricate 100,000 units of a product on a particular machine. The per-unit net profit will be $0.50 per unit, but for any unit that turns out to be defective, there is an additional expected expense of $2.00 to cover the cost not only of the replacement of the defective item with a good one, but also of possible loss of goodwill and other considerations. The generation of good and defective units is felt to follow a Bernoulli process with parameter p, the proportion of defective units that will be produced by the machine in its current state. The manager has the option of adjusting the machine at a cost of $10,000 before the batch of 100,000 units is run, with the benefit of adjustment being an expected halving of p. Once the large batch is started, it will be run without interruption.

If the manager proceeds without adjusting the machine, the payoff (in dollars) will be 50,000-200,000p. The other option, adjustment, yields a payoff of 40,000-100,000p. This is a standard problem with linear payoff functions, and the breakeven value of p is $p_b=0.10$. Of course, p is generally not known with certainty in such situations; the manager's judgments about p can be represented in terms of a probability distribution f(p). For a risk-neutral manager, adjusting is optimal if the mean of this distribution is greater than p_b.

Obtaining more information in the form of a test run of n units before deciding whether to adjust is possible. The expected value of this sample information is

$$\text{EVSI}(n) = \begin{cases} 100{,}000 \int_{p_b}^{1} (\mu''_n - p_b) f(\mu''_n) d\mu''_n & \text{if } \mu' \le p_b, \\ 100{,}000 \int_{0}^{p_b} (p_b - \mu''_n) f(\mu''_n) d\mu''_n & \text{if } \mu' > p_b, \end{cases} \tag{1}$$

where μ' is the prior mean of p and μ''_n is the posterior mean of p after a test run of n units is observed (Raiffa and Schlaifer, 1961). Before the test run, there is uncertainty about r, the number of units classified as defective in the test run. This translates into uncertainty about μ''_n, represented by a preposterior distribution $f(\mu''_n)$.

If the classification of units in the test run as defective or good is without error, then the calculation of the preposterior distribution of μ''_n and the linear loss integrals in (1) is relatively straightforward. Standard formulas involving the beta-binomial distribution apply if the prior distribution is beta, the natural-conjugate family (Raiffa and Schlaifer, 1961), and the necessary computations can be done numerically for other prior distributions. The numerical examples in this paper use beta prior distributions.

Allowing the possibility of misclassification error, we let e_1 denote the probability that a defective unit is incorrectly classified as good and e_2 denote the probability that a good unit is incorrectly classified as defective. Then the marginal probability that a unit is classified as defective is not p, but

$$q = p(1-e_1) + (1-p)e_2. \tag{2}$$

The posterior distribution after a test run of n items with r classified as defective is a mixture,

$$f(p|r,n) = \sum_{d=o}^{n} w_d f(p|d,n), \tag{3}$$

where w_d is the posterior probability that exactly d of the n units tested are *actually* defective given that r are *reported* as defective and $f(p|d,n)$ is the posterior distribution of p *if* d were known for certain (Gaba and Winkler, 1992). [When the prior distribution is beta,

$$f_\beta(p|\alpha,\beta) = p^{\alpha-1}(1-p)^{\beta-1}/B(\alpha,\beta), \tag{4}$$

where B represents the beta function, the posterior distribution in (3) is a mixture of beta distributions.] The preposterior distribution of μ''_n can therefore also be expressed as a mixture,

$$f(\mu_n'') = \sum_{d=o}^{n} w_d f(\mu_n'' | d), \tag{5}$$

where $f(\mu_n'' | d)$ is the preposterior distribution given a known d. The possibility of misclassfication makes the computation of the preposterior distribution, hence the computation of EVSI from (1), more difficult.

We believe that when errors are possible, it will usually be unrealistic to assume that the error rates are known exactly. Instead, there will be uncertainty about the error rates, and the relevant prior distribution to be assessed is $f(p,e_1,e_2)$. Under this prior distribution, the posterior and preposterior distributions are still of the forms given in (3) and (5), but the uncertainty about e_1 and e_2 further complicates the computation of the weights w_d. One possible specification of the prior distribution is for p, e_1, and e_2 to be independent a priori, each with a beta distribution:

$$f(p,e_1,e_2) = f_\beta(p | \alpha,\beta) f_\beta(e_1 | \alpha_1,\beta_1) f_\beta(e_2 | \alpha_2,\beta_2). \tag{6}$$

Here the prior distribution of p is as in (4), with mean $\mu' = \alpha/(\alpha+\beta)$, and the prior distributions of e_1 and e_2 are beta distributions with means $\mu_1' = \alpha_1/(\alpha_1+\beta_1)$ and $\mu_2' = \alpha_2/(\alpha_2+\beta_2)$, respectively.

3. Value of Information

The actual value of information in the quality-control example will depend on the sample size and the prior information about p, e_1, and e_2. To give the flavor of the impact of errors on the value of information, we present results for some selected values of n and prior distributions. A more detailed look at the specific relationships involving the value of information in this problem would not only require a considerable amount of space, but the details (as opposed to the general flavor) of such value-of-information results are difficult to generalize (Hilton, 1981).

First, we consider the case of known error rates e_1 and e_2. Suppose that the prior distribution of p is a beta distribution with $\alpha=1$ and $\beta=19$. The prior mean, $\mu'=0.05$, is less than p_b, and the optimal strategy without further information is not to adjust. The value of information as a proportion of EVSI when $e_1=e_2=0$ (the usual EVSI without errors) is presented in Table 1 for test runs of size 10 and 100 and for error rates of 0(0.1)0.3.

As expected, the EVSI decreases as the error rates increase, and the reductions in EVSI are often quite large. Furthermore, the relative reduction in value of information is more pronounced for $n=10$ than for $n=100$. For example, the $e_1=e_2=0.1$ case yields an 80% reduction from the error-free case when $n=10$ but only a 27% reduction when $n=100$. As seen in Figure 1, the value of information as a function of n in the error-free case is relatively steep for small n and flattens out for larger n. The EVSI curve with $e_1=e_2=0.1$ is less steep at first but is still climbing when the error-free curve flattens out. The n for which the value of information with errors starts gaining on the value of information without errors will vary with the situation; for the $e_1=e_2=0.1$ case in the quality-control problem this happens in the neighborhood of $n=50$.

Table 1 also illustrates a strong asymmetry. Increases in e_1 cause relatively minor reductions in EVSI, whereas increases in e_2 lead to much more severe reductions. More items are expected to be good than defective, which means that e_2, operating on good units, will have more of a chance to generate erroneous test results than will e_1, which operates on defective units. Nonetheless, the magnitude of some of the differences, especially for $n=10$, is noteworthy. Even with $n=100$, the EVSI when $e_1=0$ and $e_2=0.3$ is less than half the EVSI with these error rates reversed.

Moving to situations with uncertain error rates, we begin by using the prior distribution in (6) with the same distribution as above for p ($\alpha=1,\beta=19$) and varying prior distributions for e_1 and e_2. The prior means of e_1 and e_2 are held at 0.1, so the relevant known-error-rates comparison is with $e_1=e_2=0.1$. As α_i and β_i increase with their ratio remaining constant, the uncertainty about e_i decreases, as can be seen from the beta variance $\mu_i'(1-\mu_i')/(\alpha_i+\beta_i+1)$. The (α_i,β_i) pairs for which results are given in Table 2 are (1,9), (3,27), and (10,90), with standard deviations of 0.090, 0.054, and 0.030, respectively. The same prior distributions for e_1 and e_2 are also used with other prior distributions for p: $(\alpha,\beta)=(0.5,9.5)$ and (0.25,4.75). These distributions share a common mean of 0.05 with the initial $(\alpha,\beta)=(1,19)$, and the standard deviation of p is 0.048 for (1,19), 0.066 for (0.5,9.5), and 0.089 for (0.25,4.75).

Values of EVSI for test runs of 10 and 100 with selected combinations of prior distributions for p, e_1, and e_2 are given in Table 2. First, note that in every case, the EVSI with uncertainty about error rates and mean error rates of $\mu_1' = \mu_2' = 0.10$ is less than the EVSI with known error rates $e_1 = e_2 = 0.1$. (See also the curves in Figure 1.) Moreover, for given distributions of p and e_1, EVSI consistently decreases with greater uncertainty about e_2. For example, with $(\alpha,\beta)=(1,19)$, $(\alpha_1,\beta_1)=(1,9)$, and $n=100$, the EVSI relative to EVSI without errors moves from 51% to 19% to 0 as (α_2,β_2) shifts from (10,90) to (3,27) to (1,9). On the other hand, similar variations in (α_1,β_1) with everything else held constant have no impact on EVSI. This is consistent with the fact that changes in e_2 have a much greater influence on EVSI than changes in e_1 when the error rates are known. As for n, relative shifts in EVSI with greater uncertainty about error rates are roughly comparable for $n=10$ and $n=100$.

Looking next at changes in f(p), we find from Table 2 that the EVSI consistently increases with greater prior uncertainty about p. Keep in mind that the payoff function for the decision-making problem here is a function of p, and information about p is directly relevant. In contrast, information about the error rates is only indirectly relevant in the sense that it reduces the inferential strength of information in terms of learning about p.

4. Equivalent Error-Free Sample Size

It is not surprising that errors generally cause increased uncertainty about the proportion of interest and decreases in the value of information for a given sample size. One way of thinking about the increased uncertainty or reduced value is to ask the following question: If we could obtain error-free information, how large a sample would be equivalent in "information content" or "value" to a sample of size n with errors? This question can be looked at from an inferential viewpoint (Winkler and Gaba, 1990) by finding the error-free sample size that is expected to reduce the uncertainty about p by the same amount as the sample of size n with errors. The reduction in uncertainty might be measured by, for instance, the expected increase in $\alpha+\beta$ from prior to posterior in the case in which the prior distribution for p is beta, or the expected decrease in the variance of p. For a decision maker, however, a more appropriate measure of equivalent error-free sample size is the error-free sample size that yields the same EVSI as a sample of size n with errors. In this section, we present some equivalent error-free sample sizes determined via both inferential and decision-making viewpoints and compare the results from the two viewpoints.

For a prior distribution with $(\alpha,\beta)=(1,19)$ and known error rates $e_1,e_2=0(0.1)0.3$, equivalent error-free sample sizes from the inferential and decision-making viewpoints are given in Tables 3 and 4, respectively. The values in Table 3 were obtained by fitting beta distributions with parameters α^* and β^* to all possible posterior distributions and then finding the a priori expected $\alpha^*+\beta^*$. The equivalent error-free sample size is then $n^*=E(\alpha^*+\beta^*)-(\alpha+\beta)$. From Table 3, we see that increases in e_2 cause greater reductions in n^* than increases in e_1, as would be expected from the discussion in Section 3. There appears to be a tendency for n^*/n to be a bit higher for $n=10$ than for $n=100$, suggesting that the relative information loss in this problem is greater when $n=100$. The reduction in effective sample size is not trivial. When $e_1=e_2=0.1$, the equivalent sample size is only 40% of n when $n=10$ and 30.5% of n when $n=100$. When both error rates are at least 0.2, n^* is less than 20% of n, falling to under 10% of n when $e_1=e_2=0.3$.

The equivalent error-free sample sizes in Table 4 were computed by taking the EVSI for the samples with possible errors and, in each case, finding the error-free n^* that yields the same EVSI. These n^* values exhibit most of the same tendencies as those in Table 3. On the whole, the inferential and decision-making approaches to finding n^* yield very similar values, as can be seen by comparing Tables 3 and 4. The equivalent error-free sample sizes are a bit larger in Table 4 when the error rates are smaller and a bit larger in Table 3 when the error rates are larger, but the differences seem surprisingly small.

When we move to the case of unknown error rates, the results given in Table 5 for selected prior distributions for e_1 and e_2 differ more substantially between the inferential and decision-making viewpoints, with the former yielding larger values of n^*. Moreover, the dropoffs in n^* as the prior variances of e_1 and e_2 increase are greater under the decision-making approach. For example, holding $\alpha_1=1$ and $\beta_1=9$ and letting (α_2,β_2) move from (10,90) to (1,9), n^* with $n=100$ moves from 21.6 to 17.0 via the inferential approach and from 16.6 to effectively zero via the decision-making approach.

5. Summary

For a variety of reasons, data from a test run of a batch of items or from other sampling processes can be subject to error. If we want to make inferences about parameters, potential errors can decrease the accuracy of those inferences by causing greater spread in likelihood functions and hence posterior distributions. Furthermore, ignoring the possibility of errors can result in highly misleading parameter estimates as well as unwarranted claims of accuracy.

If we want to make decisions and are thinking about obtaining new information that might have errors, our concern about the impact of the errors should focus on changes in the expected value of the information, not on some more informal inferential measure. In this paper we have examined the influence of errors on value of information in the context of a simple quality-control problem. The results are admittedly specific to the example chosen, but they are indicative of what can possibly be large reductions in EVSI as a result of errors. Ignoring errors, then, can cause substantial overstatement of the value of information and poor decisions about information gathering. As anticipated, the reductions in EVSI in the example are greater as the error rates increase or the uncertainty about the error rates increases.

In thinking about the loss of information associated with errors, we find it convenient to think in terms of reductions in the effective sample size caused by errors. This has previously been studied from an inferential viewpoint (Winkler and Gaba, 1990), and in this paper we extend the notion to a decision-making framework by finding the error-free sample size that yields the same EVSI as a sample of size n with possible errors. Consistent with the reductions in EVSI noted above, we find that some of the reductions in effective sample size are by no means trivial. For the case of known errors rates, the inferential and decision-making approaches to equivalent error-free sample size give similar results. When uncertainty about error rates is introduced, however, differences between the methods increase as the uncertainty increases, with the decision-making approach providing smaller equivalent sample sizes. When this occurs, the implication is that the loss of information is even more severe from a decision-making viewpoint than it might seem from a strictly inferential viewpoint.

<u>References</u>

Groves, R.M. (1989), *Survey Errors and Survey Costs*. New York: Wiley.

Winkler, R.L., and Gaba, A. (1990), "Inference with imperfect sampling from a Bernoulli process," in S. Geisser, J.S. Hodges, S.J. Press, and A. Zellner (Eds.), *Bayesian and Likelihood Methods in Statistics and Econometrics*. Amsterdam: North-Holland, 303-317.

Gaba, A., and Winkler, R.L. (1992), "Implications of errors in survey data: A Bayesian model," *Management Science*, 38, 913-925.

Raiffa, H., and Schaifer, R. (1961), *Applied Statistical Decision Theory*. Boston: Graduate School of Business Administration, Harvard University.

Hilton, R.W. (1981), "The determinants of information value: Synthesizing some general results," *Management Science*, 27, 57-64.

Table 1. EVSI as a proportion of EVSI when $e_1=e_2=0$ for n=10(100), with $(\alpha,\beta)=(1,19)$ and known error rates $e_1,e_2=0(0.1)0.3$.

		e_2: 0	0.1	0.2	0.3
e_1	0	1(1)	.29(.80)	.06(.62)	.01(.45)
	0.1	.98(.98)	.20(.73)	.03(.52)	.00(.34)
	0.2	.93(.95)	.13(.66)	.01(.41)	.00(.22)
	0.3	.85(.92)	.07(.56)	.00(.29)	.00(.11)

Table 2. EVSI as a proportion of EVSI when $e_1=e_2=0$ for n=10(100), with various combinations of (α,β), (α_1,β_1), and (α_2,β_2).

		(α,β)		
α_1,β_1	α_2,β_2	(1,19)	(0.5,9.5)	(0.25,4.75)
1,9	1,9	.00(.00)	.09(.17)	.36(.45)
3,27	3,27	.04(.18)	.26(.51)	.49(.69)
10,90	10,90	.15(.52)	.35(.74)	.54(.83)
1,9	3,27	.04(.19)	.26(.50)	.48(.68)
1,9	10,90	.09(.51)	.35(.73)	.54(.83)
3,27	1,9	.00(.00)	.09(.17)	.36(.45)
10,90	1,9	.00(.00)	.09(.17)	.36(.45)
$e_1=e_2=0.10$		.20(.73)	.40(.87)	.56(.91)
EVSI when $e_1=e_2=0$		147.8(498.2)	668.1(1091.0)	1417.6(1891.1)
EVPI		607.9	1182.0	1897.9

Table 3. Equivalent error-free n* in the sense of equating increases in $\alpha+\beta$ for n=10(100) with $(\alpha,\beta)=(1,19)$ and known error rates $e_1, e_2=0(0.1)0.3$.

		e_2 = 0	0.1	0.2	0.3
e_1	0	10(100)	4.8(36.2)	3.1(22.8)	2.2(16.1)
	0.1	8.9(89.3)	4.0(30.5)	2.5(18.6)	1.6(12.6)
	0.2	7.8(78.8)	3.2(25.1)	1.9(14.7)	1.2(9.4)
	0.3	6.7(68.4)	2.5(20.1)	1.4(11.0)	0.8(6.5)

Table 4. Equivalent error-free n* in the sense of equating EVSIs for n=10(100) with $(\alpha,\beta)=(1,19)$ and known error rates $e_1,e_2=0(0.1)0.3$.

e_1 \ e_2	0	0.1	0.2	0.3
0	10(100)	3.0(42.8)	1.5(23.4)	1.1(14.0)
0.1	8.6(86.6)	2.5(33.6)	1.3(17.4)	1.0(11.0)
0.2	7.6(75.7)	2.1(26.1)	1.1(12.7)	1.0(5.8)
0.3	6.7(65.1)	1.6(20.9)	1.0(8.0)	0-1*(3.5)

*0-1 when EVSI=0 because EVSI=0 for n=0 <u>and</u> n=1 in the noise-free case.

Table 5. Equivalent error-free n* determined by equating increases in $\alpha+\beta$ and equating EVSIs for n=10(100) with $(\alpha,\beta)=(1,19)$ and various combinations of (α_1,β_1) and (α_2,β_2).

α,β_1	α_2,β_2	Equating $\alpha+\beta$ Increases	Equating EVSIs
1,9	1,9	3.6 (17.0)	0-1* (0-1*)
3,27	3,27	3.7 (17.5)	1.4 (4.9)
10,90	10,90	3.9 (22.5)	2.2 (16.9)
1,9	3,27	3.7 (17.0)	1.4 (4.9)
1,9	10,90	3.9 (21.6)	1.7 (16.6)
3,27	1,9	3.6 (17.6)	0-1* (0-1*)
10,90	1,9	3.7 (17.8)	0-1* (0-1*)

*0-1 when EVSI=0 because EVSI=0 for n=0 <u>and</u> n=1 in the noise-free case.

Figure 1. EVSI without errors, with known error rates, with uncertain error rates, and EVPI for $(\alpha,\beta) = (1,19)$.