LA-UR-77-2420

MASTER

TITLE: STATISTICAL ASPECTS OF NUCLEAR SAFEGUARDS

CONF-771042--1

AUTHOR(S): Gary L. Tietjen

SUBMITTED TO: Third ERDA Statistical Symposium, October 26-28, 1977, Battelle Pacific Northwest Laboratories, Richland, WA

los alamos
scientific laboratory
of the University of California
LOS ALAMOS, NEW MEXICO 87545

An Affirmative Action/Equal Opportunity Employer

Form No. 836
St. No. 2629
1/75

UNITED STATES
ENERGY RESEARCH AND
DEVELOPMENT ADMINISTRATION
CONTRACT W-7405-ENG. 36

STATISTICAL ASPECTS OF NUCLEAR SAFEGUARDS

by

Gary L. Tietjen
Los Alamos Scientific Laboratory
Los Alamos, New Mexico

ABSTRACT

A nuclear fuel reprocessing cycle is used to illustrate problems encountered by a statistician when trying to reconcile total amounts of an element at different stages in the recovery cycle. Calculation of errors are discussed along with problems of biases, holdup, and simulation.

Introduction

Each of the ERDA laboratories and contractors is already, or soon will be, deeply immersed in nuclear safeguards and accountability. As I hear the problem discussed from a political viewpoint, there are frequent official references to a "malevolent act," but the term seems to refer more to blackmail threats to a civilian population than to the use of weapons in war, though the latter possibility is always present. The questions are: how can we keep unauthorized persons from getting nuclear material and how do we tell whether some of it is missing? A uniform system of keeping track of our inventory will be necessary since some international control seems imminent and perhaps desirable.

The task is of enormous proportions. Some of the reactors going on line will process or reprocess 50 kg/day of plutonium. Every item, every drop of solution, every piece of scrap metal, and every whiff of powder will have to be accounted for. Moreover, the transactions from one place to another or from one form to another will take place rapidly, so that the accountability will have to be automated on the computer. There will not be time to mull over decisions on a case-by-case basis as we have hitherto done.

When one mentions the word safeguards, he may be completely misunderstood. There are many who think of safeguards wholly in terms of physical security. At the new plutonium facility at Los Alamos, there probably will be a computer check of your badge, your signature, your fingerprints and perhaps of your voice before entering the facility. At the same time, you would be monitored for radio-activity, of course. The chemists think the problem of safeguards

solved if they have devised accurate and precise methods of analysis for minute quantities of material. The physicists think of safeguards problems in terms of very rapid methods of analysis using nondestructive methods of analysis not requiring lengthy sample preparation. (In all of their methods they either count the sample directly or irradiate it in some way and then count it.) The computer people believe they have solved the problem of safeguards if they are able to get the numbers quickly onto a data base with rapid retrieval capability. It is left for the statisticians to try to make some sense of the thousands of numbers that will be generated.

I have chosen one small segment of an actual reprocessing cycle at Los Alamos to illustrate the problems faced by the statistician. I shall neither exaggerate nor minimize the difficulties. Of course, material is lost during processing. The public and the press don't seem to understand this and have not been sympathetic. The losses are not so large as at first they seem. Losses of uranium at Los Alamos over the last 25 years, if put in metallic form, could be placed in one of the larger women's purses. Of course, she would have difficulty walking out with it.

The situation is somewhat like making cookies. Suppose you were given a certain amount of flour, sugar, etc., for this purpose, and that the ingredients were weighed out to you. After the cookies are baked, they are weighed. You are allowed a certain loss for evaporation, but still there is material missing. Where is it? On the beaters, the spoon, in the bowl, and on the dishrag that wiped up the spillage. Taking all this into account, one still has to decide whether the kids running through the kitchen have licked the spoon or made off with a cookie.

Let me get into the example (Slide 1). We start out with a uranium metal alloy. The concentration of uranium in the metal is determined chemically, and the metal is weighed. A part is then machined from the alloy, and the finished part is weighed. The difference in the two weighings is the weight of the scrap that is gathered up and put into cans. The scrap itself cannot be weighed because it is oily; it can neither be dissolved safely nor stored safely because it is pyrophoric. Consequently the scrap is then burned to an impure oxide and stored in cans in a vault until such time as there is enough of it accumulated for a batch to be reprocessed and until the facilities are ready. Then the ash is dissolved in an acid. The volume of the solution is

measured and the concentration determined by an NDA device called the USAS. At this point we make our first check: concentration × volume = total uranium. The total uranium in solution should be equal to the weight of the scrap metal × concentration of the metal.

Next, something is done to the solution to precipitate the uranium oxide About 90% of the uranium is precipitated and 9% remains in the filtrate solution and 1% is on the rags used for cleanup. The precipitate is weighed and its concentration determined chemically. (For obscure reasons, the concentration of the batch is not used directly. Not every batch is assayed. Instead, the annual average concentration is used. For reasons of the chemistry involved, this should be quite close to the analysis for any one batch. It is the 90% figure that will vary considerably.) The filtrate solution has its volume measured and its concentration determined by the USAS device. Finally, the collection of rags for an entire month (rather than a batch) is burned and the amount of uranium in them determined by the random driver, which has about ten times the error of the USAS device. Fortunately. only a small amount of material is involved.

We then add up the total uranium in the precipitate, the filtrate, and the rags and it should check with the amount found in the solution before precipitation. The differences in the consecutive totals are called MUFS (now BPID's) and each check makes up an account. There may be 75-100 such accounts at an R & D facility such as Los Alamos.

At the end of each month, we close the books on the scrap metal. but it may be some time before we have all the figures with which to reconcile the books. What error shall we associate with the three totals we now have? Certainly the totals have different variances We can enter the figures as shown here (Slide 2) in a system of multiple entry bookkeeping, but we must allow an extra column for the error or variance of each figure. With the aid of the error column it is our job to decide whether the books balance. If not, there has been an arithmetical error or a diversion of material and an investigation ensues. This system we might refer to as STATISTICAL bookkeeping with the statistician acting as the AUDITOR.

How do we calculate these errors? Each entry is a product: concentration x volume. The variance of the product cv is $\mu_c^2\sigma_v^2 + \mu_v^2\sigma_c^2 + \sigma_c^2\sigma_v^2$. If we replace variances with sample variances, we have one estimator of the sample variance

of cv, but it is biased. An unbiased estimate is $\bar{c}^2 s_v^2 + \bar{v}^2 s_c^2 - s_c^2 s_v^2 (1-1/m-1/n)$.

The propagation of error estimate is $\bar{c}^2 s_v^2 + \bar{v}^2 s_c^2$, and it, too, is biased. Which do we use? There is still some argument among statisticians. No minimum mean square estimator seems available, the problem being seemingly intractable.

These estimates, however, do not take into account the error in fitting the calibration lines, which could be considerable. To be more explicit let us get to the details. There is a linear calibration line set up for the USAS device (Slide 3) and the equation of the regression line is $y = \hat{\alpha} + \hat{\beta}x$ where the x's are regarded as fixed (they are known standards). We use this regression line in reverse, i.e., we observe y and solve for the corresponding x-value: $x = (y-\hat{\alpha})/\hat{\beta}$. This gives us the concentration. We multiply this by the volume v of the solution and sum over the several solutions processed during the month to obtain the total uranium $\Sigma v_i(y_i-\hat{\alpha})/\hat{\beta}$. For the precipitates, we have another calibration line for the chemical results (Slide 4) with equation $z = \hat{\gamma} + \hat{\delta}x$. Again this is used in reverse: $x = (z-\hat{\gamma})/\hat{\delta}$. In this case, though, we observe a large number of z's and get an annual average x, of the corresponding results as the concentration factor. Multiplying this average concentration by the weight w of a particular precipitate, we obtain $\Sigma w_i(z-\hat{\gamma})/\hat{\delta}$. For the filtrate, a different calibration curve is used with the USAS device (a different set of standards) and we obtain $\Sigma q_i(y_i-\hat{\rho})/\hat{\theta}$ where q_i are volumes and the regression line is $y = \hat{\rho} + \hat{\theta}x$. Finally, for the rags, we use still another line, $y = \hat{\tau} + \hat{\xi}x$ and use the single figure $x = (y-\hat{\tau})/\hat{\xi}$. The difference between the two sums that should balance is then (Slide 5):

$$Q = \underbrace{\Sigma v_i(y_i-\hat{\alpha})}_{\hat{\beta}} - \underbrace{\Sigma w_i(z-\hat{\gamma})}_{\hat{\delta}} - \underbrace{\Sigma q_i(y_i-\hat{\rho})}_{\hat{\theta}} - \underbrace{(y_i-\hat{\tau})}_{\hat{\xi}} .$$

MUF = solution - precipitate - filtrate - rags

We could, by propagation of error, find an approximate variance s_Q^2. If Q is unbiased, we would like to test whether it is zero, and if we had enough faith, we might rely upon asymptotic normality and look at the ratio Q/s_Q. If Q has estimated bias B, we might form the ratio $(Q-B)/(s_Q^2+s_B^2)^{1/2}$ and compare it to a t-distribution. Is propagation of error the proper tool here?

Some of sample variances needed for s_Q^2 may be difficult to obta'n. The statistician will have to obtain the calibration results and obtain variances for each piece of equipment used. He will need to familiarize himself thoroughly with each step in the process, which will be time-consuming. The variances for volumes can be a real headache. I had always thought that a tank volume, once calibrated, would stay calibrated, but that is not the case here. The calibration is constantly drifting. The tank used has to be filled with hollow boron glass cylinders that act as moderators to keep a solution from going critical. The acid solutions eat the glass away, causing the volume to continually increase until recalibrated. Thus we get a curve somewhat like this (Slide 6). It is not trivial to recalibrate some of these tanks. Even if you fill a tank with a measured container of water, how much air is in that water? What is the density of the water? At Idaho, one large tank there holding about 2000 liters has to be shielded and is sometimes calibrated as follows: Pour a known volume with a known concentration of strontium into the filled tank. Observe the concentration of the dilute solution. The ratio of the two concentrations is proportional to that of the two volumes. We have had the same type of problem with uranium foil. Every year the weight of the foil, which is sitting in storage, increases! That gives an apparent increase in uranium. The auditors didn't know what to do with that one until it was discovered the foil was simply oxidizing.

Another problem is "hold-up" in these tanks. Depending upon the acidity, some of the uranium may adhere to the glass cylinders. When a more acid solution is used, you flush this off and get more uranium than you started with. A common case of hold-up occurs in glove boxes. A little uranium oxide may be spilled during weighing and left in the glove box. Eventually, perhaps months later, the glove box is thoroughly cleaned and this buildup added to the account. The result can be observed by watching the account as a function of time. Nearly every loss or low value is followed by a high value in the succeeding month. How do we model this hold-up? How do we take it into account?

Another approach we have tried is simulation. We need a confidence interval for the quantity Q. We don't wish to rely either upon normality of Q nor upon the propagation of error approximation for the variance of Q. To do simulation, however, we shall have to assume certain distributions and

parameter values for the random variables involved in Q. Mark and Myrle Johnson at Los Alamos have done some simulation work on this problem. To keep the results from being overly dependent upon a given distribution, one needs a family of reasonable distributions for the random variables. They have come up with a family each member of which has mean zero, unit variance and zero skewness (i.e., they are all symmetric). There is a parameter α that governs the kurtosis (Slide 7). The family includes the uniform distribution at one extreme (β_2=1.8), the normal distribution (β_2=3) and a very peaked distribution with β_2=5.4. A single algorithm permits easy generation of the variables for any member of the family. They first decided to call this family the NEW DIST until someone pronounced the acronym too rapidly. (I have copies of their paper should anyone be interested.) By varying the kurtosis α, we can study the length of the resulting confidence interval on Q (Slide 8). We can then choose the longest interval for which we think the kurtosis is reasonable. Of course, one could study a family of asymmetric distributions by exponentiating the random variable we generate.

The simulation approach requires the same amount of work in gathering parameters and variances but has seemed a bit more reasonable and flexible to us than the straight propagation of error. We are looking for further suggestions along these lines.

The picture may be still further complicated by frequent (say, weekly) calibration, which will be insisted upon at the new plutonium facility at LASL. Then we will have to add a few more but similar terms to our expression for Q.

A more disturbing problem is bias. What is bias? Some of you were raised on the concept that the bias of an estimator $\hat{\theta}$ is $E(\hat{\theta}) - \theta$. That, by definition, seems to make the bias a constant. In a series of influential papers, Churchill Eisenhart at NBS gave a very similar definition. but he has replaced $E(\hat{\theta})$ with the limiting mean μ of a set of measurements (under identical circumstances) on a quantity. He then says "the systematic error or bias...of a measurement process will ordinarily have both constant and variable components." That makes bias a random variable. He illustrates by considering a distance measured with a steel tape. The temperature on the day on which the measurements are made adds a random variable into the limiting mean, hence into the bias. The term "limiting mean" is not so well defined and from this the concept has expanded into "long-term" and "short-term" systematic errors. which

may be either constants or random variables. Not understanding each other. there have been vociferous arguments among statisticians within our ERDA community about bias and systematic error and how to correct for them and when not to correct for them, etc.--with everyone using their own definition of bias. May I give an example of what confuses the statisticians, and even more the experimenter? It is to get a form like some we see from the EPA and NBS asking for a series of measurements to be used in standardizing a new method. say. Here are the questions the experimenter is required to supply under the heading of CALIBRATION RESULTS: (1) What is the overall uncertainty on the value of the activity? (2) What is the standard error? (3) Give a 99% confidence limit; (4) The total estimated systematic error is __, comprised of __% due to __ and __% due to ___ etc. (5) How are the systematic errors combined? (6) How are the random and systematic errors combined? To fill out such a form requires agreement on what the terms mean, and I don't think we have yet reached that argument among ourselves. We need to do some housecleaning. Indeed, we may be a little disturbed about filling out the form because we think they might misinterpret or misuse what we say.
I am trying to say that this chemistry business is swarming with biases and systematic errors and I would like to get a colloquium started on that issue.