Radiation-Sterilization of Food: Treating Scanty Data from Inoculated Packs

EDWARD W. ROSS, JR.

Office, Technical Director, U.S. Army Natick Research and Development Command, Natick, Massachusetts 01760

(Received for publication January 26, 1979)

ABSTRACT

This paper is about the mathematical methods used in calculating the 12D radiation-dose for a food which has been subjected to an inoculated pack. It concerns the case in which only one partial-spoilage data point is obtained from the experiment. A simple, partly graphical procedure is described. This method is based on binomial confidence limits and furnishes an estimate that lies above the (unknown) true 12D value with probability greater than 90%. The method is applied to an inoculated pack for low-level nitrate/nitrite ham described by Anellis et al. (2).

The purpose of this paper is to describe a mathematical method that may occasionally be useful in treating data on radiation-sterilization obtained from an inoculated pack. The emphasis is less on the microbiological implications of the method than upon explaining it and showing how to use it.

The specific question addressed here is, how can a “safe” dose be estimated when the data contain exactly one dose at which some, but not all, sample units (cans) are inactivated. In most inoculated packs the data contain two or more such “partial-spoilage” points, and in that event the safe (12D) dose is best estimated (in the author’s opinion) by the maximum-likelihood method (8) or the least-squares procedures (6). Both these methods fail (i.e., they do not yield a unique estimate of the 12D dose) when only one partial spoilage point is obtained, as in Anellis et al. (1,2).

In this case the older methods of estimating a safe dose (e.g. Schmidt-Nank or Spearman Kaerber) are still usable, in the sense that they provide a unique estimate of 12D. However, they are based on the assumption that organism-inactivation follows a simple, one-parameter, exponential distribution, an assumption which may sometimes be too restrictive; see the discussion by the author (6). In fact, this assumption may be inconsistent with the data of the experiment, but these methods furnish no way of determining this. The procedure proposed here avoids these difficulties.

This method is based on the general probability theory for inoculated packs given by the author (6) and assumes that organism inactivation follows the shifted (two-parameter) exponential distribution (5,7). The computations are not excessively burdensome, and will lead to roughly the same 12D that would have been obtained by assuming a simple-exponential distribution if that assumption is consistent with the data. That is, the method abandons the simple-exponential assumption only when the data force it to do so.

METHODS

The following nomenclature will be used in describing the experimental results of the inoculated pack.

- $x_i$: $i$-th experimental dose in grays
- $T_i$: number of cans exposed at dose $x_i$
- $n_i$: number of organisms per can at dose $x_i$
- $S_i$: number of spoiled cans at dose $x_i$
- $K_i$: number of sterilized cans at dose $x_i$
- $S_i = T_i - K_i$
- $S_0$: lowest dose with 5 cans at dose $x_i$
- $S_1$: lowest dose for which all cans are sterilized
- Any of $S_i$: lowest dose for which $S_i = 0$
- $S_2$: highest dose with 5 cans
- $S_3$: highest dose for which all cans are sterilized
- $S_4$: lowest dose for which $S_4 = 0$

In graphing results, the quantity

$$Y(x) = \log_{10} \left( 1 - G(x) \right)$$

is also used. It is well known, e.g. Ross (6), that

$$G(x) = 1 + \left( 1/n \right) \ln \Phi(x)$$

and therefore

$$Y(x) = \log_{10} \left[ 1 - \left( 1/n \right) \ln \Phi(x) \right]$$

In this method, based on binomial confidence limits. A significance level, $a = 0.05$, is chosen first. Given that $K_i = T_i - S_i$ out of $T_i$ cans were sterilized at dose $x_i$, one finds from charts or tables of binomial confidence limits for proportions two numbers, $\Phi_{L}(x)$ and $\Phi_{U}(x)$, which are the lower and upper ($1-a$) confidence limits for the true value of $\Phi = \Phi(x)$. Similarly, at dose $x_1$ an upper $\left[ 1 - (a/2) \right]$ confidence limit, $\Phi_{U}(x_1)$, is found for $\Phi_1 = \Phi(x_1)$, and at dose $x_2$ a lower $\left[ 1 - (a/2) \right]$ confidence limit, $\Phi_L(x_2)$, is found for $\Phi_2 = \Phi(x_2)$. That is

- $\Phi_L(x_1) < \Phi_U(x_1)$
- $\Phi_U(x_2) < \Phi_U(x_1)$
- $\Phi_L(x_2) < \Phi_U(x_2)$

By means of formula [2] this set of inequalities is transformed into a set of inequalities on $Y_1 = Y(x_1)$.

- $\Phi_L(x_1) \geq Y(x_1) \geq \Phi_U(x_1) - (a/2)$
- $\Phi_L(x_2) \geq Y(x_2) \geq \Phi_U(x_2) + (a/2)$

This is the First International Congress on Engineering and Food, Boston, Massachusetts, August 9-13, 1976.
The two-parameter exponential form satisfies the inequalities at the doses $x_1, x_2, \ldots$ or give straight lines when plotted on a graph of $Y$ versus $x$, and a death-kinetic function, $G(x)$, is sought such that $G(x) = 1 - 10^{-A(x-A)} = 1 - e^{-2.303A(x-A)}$.

The 12D estimate $X_c$ is determined as follows:

(i) Find the least steep line which satisfies all the inequalities [3], [4] and [5] and has $A > 0$.

(ii) Then $X_c$ is the x-value (dose) at which that line intersects $Y = -12$. Because of the way $X_c$ has been derived, it is true that the (unknown) value of $X_c$ satisfies $X_c < X_c$ with probability exceeding $1-2a = .90$.

**EXAMPLE**

An inoculated pack was carried out with low level nitrite-nitrate ham as substrate, Anellis et al. (2), and gave the following data:

- $x_1 = 17, S_1 = 100, k_1 = 0$
- $x_2 = 20, S_2 = 76, k_2 = 24$
- $x_3 = 23, S_3 = 0, k_3 = 100$

where $T_i = 100$ and $n_i = 2.41 \times 10^4$ for all $i$. The binomial confidence limits are

\[ \hat{\phi}_U = .037; \hat{\phi}_L = .34; \hat{\phi}_U = .963. \]

and the bounds are found, for example, as

\[ B_{1U} = \log_{10} \left[ \frac{1}{n} \ln \left( \hat{\phi}_U \right) \right] = \log_{10} \left[ \frac{1}{n} \ln (.037) \right] = -5.86 \]

Similarly, $2.41 \times 10^6 B_{1L} = -6.12 B_{2U} = -6.35 B_{2L} = -7.80$.

These four bounds are depicted by arrows in the customary semi-logarithmic plot of survival fraction as a function of dose, Fig. 1. On this graph the line (R) was obtained by visually finding the least-steepest straight line which passes above the arrow at $x = 17$, between the arrows at $x = 20$ and below the arrow at $x = 23$. This line is associated with the parameter values

- $A =$ shoulder width $= 6.4$
- $\lambda = -$ (slope) $= .469$

and has

$\hat{X}_c = 12D = 32.0$

In this case we assert with 90% confidence that the (unknown) true value of $X_c$ satisfies $X_c < \hat{X}_c = 32.0$.

Notice that many lines can be found to satisfy all the inequalities. In Fig. 1 (Q) is another such line. However, the graph shows that any line steeper than (R) leads to an $X_c$ that is lower (i.e. less safe) than $\hat{X}_c$. In choosing the least-steepest line we are demanding the highest 12D value that is consistent with the data. Also, that choice leads to the line that has the smallest possible shoulder-width (least value of $A$) and is therefore closest to a simple exponential among all the possible lines.

In choosing the $X_c$ that is lower (i.e. less safe) than $\hat{X}_c$, we are demanding the highest 12D value that is consistent with the data. Also, that choice leads to the line that has the smallest possible shoulder-width (least value of $A$) and is therefore closest to a simple exponential among all the possible lines.

**Figure 1. Graphical determination of $X_c = 12D$ based on binomial confidence limits. $Y = \log_{10}$ (survival probability).**

Notice also that step (a) of the procedure specifically excludes the possibility of a negative shoulder, $A < 0$. For, if a straight line can be found that satisfies the inequalities and has $A < 0$, then other acceptable lines can be found with $A = 0$, and the least-steepest line of the latter family is the one that should be used. This will of course correspond to using a simple-exponential death-kinetic hypothesis.

We might ask whether a simple-exponential form of death-kinetics is possible in this case. That is, can a line be found that satisfies all the inequalities and intersects the origin, $x = 0, Y = 0$? The answer is no. The line (P) in Fig. 1 represents an attempt to do so; (P) and any steeper line violate the lower inequality at $x = 20$ and any less steep line violates the inequality at $x = 23$.

We may examine this question from another point of view. Suppose we assume that $G(x)$ is a simple-exponential distribution, i.e. $A = 0$ and

\[ 1 - G(x) = 10^{-Ax} \]

$Y(x) = -\lambda x$.

The best estimate of $\lambda$ is found, see Ross (6), by using the lone partial-spoilage point and Equation [2].

\[ \hat{\phi} (20) = 24/100 = .24 \]

\[ Y (20) = \log_{10} \left[ \frac{1}{n} \ln (.24) \right] = -6.23. \]

Equation [6] is then solved (at $x = 20$) to obtain $\lambda = .311$. Hence the simple exponential function which fits the one partial-spoilage point is

\[ 1 - G(x) = 10^{-311x} \]

$Y(x) = -.311x$.

It is easy to see that this function is inconsistent with the data at $x = 23$. For,

\[ \hat{\phi} (23) = e^{-n [1 - G(23)]} = e^{-[2.41 \times 10^4 \times 10^{-311} \times 23]} = .844 \]

The probability of getting the observed experimental
result, i.e. all 100 cans sterilized at \( x = 23 \), given that the probability of getting one sample sterilized is \( .844 \), is given by the binomial theorem as \( .844^{100} = 4.3 \times 10^{-8} \). That is, if \( G(x) \) is simple-exponential and consistent with the partial spoilage data, it is extremely unlikely that all sample cans would have been sterilized at \( x = 23 \). Again, we are forced by the data to conclude that \( G(x) \) is unlikely to be simple-exponential.

**DISCUSSION**

To use this method, one has to have charts or tables of binomial confidence limits for proportions. Such charts and tables can be found in Dixon and Massey (4) or Beyer (3) or many other sources. It is not always easy to read the charts accurately in finding \( \hat{\theta}_{1U} \) and \( \hat{\theta}_{3L} \). The following approximate formulas are helpful provided \( T_1 \), \( T_3 \geq 20 \) and \( \alpha = .05:

\[
\hat{\theta}_{1U} = (3.7/T_1) [1 - (1.7/T_3)],
\hat{\theta}_{3L} = 1 - (3.7/T_1) [1 - (1.7/T_3)].
\]

The procedure described here may be criticized for several reasons. Primarily, one is likely to feel uncomfortable about the stability of the estimates. For example, if the pack were repeated, we would not be surprised if \( k_1 \) were, say, 20 or 30 instead of 24; nor would it be startling if \( k_1 \) or \( S_2 \) were 1, 2 or 3 instead of zero. These could lead to substantially different values of \( \hat{\theta}_C \). How can we rely upon these estimates in this case? Secondarily, it is unpleasantly arbitrary to base so much of the unknown true \( 12D \) dose, \( x_C \), on the choice

\[
\hat{x}_C.
\]

rather, with probability \( \geq .9 \), it is an upper bound on \( x_C \). This means that, if we were to repeat the entire inoculated pack many times, the true \( x_C \) would exceed \( \hat{x}_C \) in less than 10% of the repetitions.

We can view this in different ways. Suppose a pack is run and the estimate \( \hat{x}_C \) obtained by this procedure. Then a second pack is run and the estimate \( \hat{x}_C^2 \) is found, where \( \hat{x}_C > \hat{x}_C^2 \). Does this mean that the higher estimate, \( \hat{x}_C^2 \), must be used? Not at all. The statement that

\[
x_C \leq \hat{x}_C \text{ with probability } \geq .9
\]

is logically consistent with the statement

\[
x_C \leq \hat{x}_C \text{ with probability } \geq .9
\]

because

\[
x_C \leq \hat{x}_C \text{ implies } x_C \leq \hat{x}_C.
\]

In other words, probably \( \hat{x}_C \) is merely a worse (higher) upper bound than \( \hat{x}_C \), and we are not forced to use the worse bound when a better one is known.

A looser, but simpler, description is merely to say that by using the \( [1 - (a/2)] \) confidence limits, we are building a substantial safety-factor into the estimate \( \hat{x}_C \), which provides reasonably high (90%) assurance against underestimating \( x_C \) because of random errors. In fact, we may regard \( x_C \) as roughly a 13D estimate because of this extra 90% safety factor. Naturally, if we used a smaller \( \alpha \)-value (say .01) we would be building-in even more of a safety factor. This seems excessive and unnecessary, but the method is not changed in any essential way by a different choice of \( \alpha \). The only effect is that a different graph or formula must be used to find \( \hat{\theta}_{1U} \), \( \hat{\theta}_{3L} \), \( \hat{\theta}_{1U} \), and \( \hat{\theta}_{3L} \).

It is clear that this method makes crucial use of information from outside the partial-spoilage range, i.e. the points \( x_1 \) and \( x_3 \). This is natural because in the present case the partial-spoilage range consists of only one point, \( x_2 \), and furnishes relatively little information.

It is perhaps distasteful to arrive at a unique estimate of \( 12D \) by introducing the extra condition that \( 12D \) itself be maximized. However, procedures of this general type are common in mathematics. For example, one of the basic steps in defining the pseudoinverse of a linear equation system is to make an otherwise non-unique solution unique by demanding that the vector solution itself be minimized. Thus, there are both theoretical and practical reasons (safety) why this extra condition is a plausible one.

In a larger context, we should recognize that this procedure probably ought to be viewed as making the best of a bad situation. An alternative procedure would be to re-run the inoculated pack, using more closely spaced doses, so that several partial-spoilage data points are obtained. The present method is intended to avoid doing this, if possible, since repeating the pack may involve considerable loss of time or money. However, the dilemma as to whether this repetition is needed emphasizes the importance of avoiding the difficulty by designing the original inoculated pack as wisely as possible. In this connection the suggestions of the author (7) may be helpful.

**REFERENCES**