D R A F T

Conf-730110--1

RECEIVED BY TIC JAN 29 1973

A STATISTICAL STUDY OF RADIATION
DOSE FOR FISHERMEN*

1117

by

Thomas M. Beetle

BATTELLE
PACIFIC NORTHWEST LABORATORY
RICHLAND, WASHINGTON

December 1972

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Atomic Energy Commission, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

- * This paper is based on work performed under United States Atomic Energy Commission Contract AT(45-1)-1830.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Jey

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

A STATISTICAL STUDY OF RADIATION
DOSE FOR FISHERMEN*

Thomas M. Beetle

BATTELLE MEMORIAL INSTITUTE
Pacific Northwest Laboratories
Richland, Washington 99352

ABSTRACT

The evaluation of the low level radiation dose to a population residing in the vicinity of a nuclear reactor installation is complex because there are several pathways and radioactive materials that constitute potential sources of exposure. The existence of these pathways and materials suggests a natural division of research efforts for evaluating the cumulative dose to the population.

During the operation of the production reactors at the Hanford installation in Washington, the Columbia River water that was used for cooling carried small quantities of radionuclides into the river. This created pathways for potential external and internal exposure for the local population, and, as a result, the river became the subject for several research projects. One project involved a study of local sport-fishing activities for the purpose of estimating a component of radiation dose associated with such activities.

A time/space frame of units was defined for the river area and time period of interest. Fishing activity quantities were defined for the units in a manner that would allow estimation of annual area totals from data for a sample of units. A multistage sampling plan and related estimation procedures were designed for the purpose of estimating the

* This paper is based on work performed under United States Atomic Energy Commission Contract AT(45-1)-1830.

-ii-

number of fishermen, the number of fish caught, the time spent fishing, and other quantities of interest. The study provides an illustration of an application of statistical concepts for investigating a relationship between an environmental disturbance and a human population.

A STATISTICAL STUDY OF RADIATION DOSE FOR FISHERMEN

INTRODUCTION

The normal operation of nuclear reactors may result in the introduction of minute amounts of radioactive matter into the environment. For example, during the operation of the production reactors at the Hanford installation in Washington, the Columbia River water that was used for cooling carried small quantities of radionuclides into the river. These materials were deposited in the soil, vegetation, and fish, and constituted a potential source of radiation exposure for the human population residing in the vicinity of the plant. Because of this, the United States Atomic Energy Commission sponsors research for monitoring the environment around AEC installations and evaluating radiation dose to the local populations.

Safety standards for radiation dose, as established by various agencies, are expressed in terms of rems per year per individual. Since it is not practical to measure dose for everyone residing in the vicinity of a nuclear installation, the standards are applied in comparisons to estimates of quantities associated with the dose distribution for the population. A quantity of special interest is the average dose per individual in the population. Since there is more than one pathway for potential exposure and different radioactive materials give rise to different doses, the data collection and estimation procedures for investigating cumulative dose can be quite complex. The usual research approach is to study each pathway and material separately and synthesize dose estimates. This presentation is a description of a part of one of these studies.

The study was conducted during 1966-67. The specific problem was to collect data on the habits and experiences of Columbia River fishermen for use in estimating radiation doses due to sportfishing by the Tri-City population. A person could receive external exposure while fishing or internal exposure through consumption of fish. Some of the quantities of interest were number of fishermen, number of fishing trips, length of fishing trips, and number of fish caught. The statistical problems were associated with the sampling designs and estimation procedures.

DEFINITIONS

The population of interest was the residents of the Tri-Cities (Richland, Kennewick, and Pasco), Washington and nearby towns (Figure 1). The three cities are located along a ten mile stretch of the Columbia River below the AEC Hanford installation in southeastern Washington. A thirty mile segment of the river is used for sportfishing by the local residents. Since the fishing activity varies seasonally and it was desirable to collect information by directly observing fishing activities, it was decided that the whole thirty miles of the river would be studied for a full year. A sampling frame was constructed by defining time/space units for the period and area of interest (Figure 2). Since fishing activity during non-daylight hours is almost nonexistent, the nighttime periods were omitted from the frame. Daily observation time was divided into four-hour periods - three periods for the late fall through early spring seasons, and four periods for the rest of the year. The shoreline was divided into 21 areas on the basis of judgments concerning the ability

of an observer to completely cover an area during a four hour assignment period (Figure 3). Each time period-area combination defined a unit for sampling purposes. There were 26,817 units in the frame.

Relationships between the time/space units and the quantities of interest were constructed by specifying activities that would be uniquely associated with a unit. For example, a fisherman's experiences were associated with the unit that encompassed the time and place for the completion of his fishing trip. So, the length of his trip, the number of fish he caught, his place of residence, and other information were associated with a specific unit. When summed over all units, the resulting quantities would equal the total time spent fishing in the area, the total number of fish caught, and so forth. For the total number of fishermen using the area, we observed first trips for the year. Information on these activities and characteristics were obtained by direct observation and interview, and the results for a sample of the units constituted the data for the study.

SAMPLING

Sample design considerations had to include recognition of the resources available for the field work. The constraints imposed were:

1. The work was to be performed by one interviewer/observer (a Washington State Department of Game Biologist);
2. He was to work 239 eight-hour days; and
3. He was not to work split shifts.

In addition to allowing for these constraints, the sample design had the usual aims of increasing precision for the estimates through stratification. The result was a stratified, multistage sampling design.

Prior information on seasonal differences in fishing activities and favorite fishing areas suggested stratification on both the time and space variables. Unfortunately, since one observer could not be in two places at one time, it was not feasible to stratify on the space variable. Therefore, eight strata were constructed on the time variable - weekdays versus weekends and holidays for each of four "quarters" (Figure 4). The purpose of the stratification was to allow for larger samples in the strata that were expected to have more variable units. Thus, the aim was to do proportionately more sampling on weekends and holidays and during the first and third quarters than during other parts of the year.

Since split shifts were not to be allowed, it was not possible to select simple random samples of units within strata (Figure 5). Instead, a three-stage sampling procedure was used: a simple random sample of days was selected; for each day, one of the initial two or three time periods was selected at random; and, for that time period and the following one, an area was selected at random. The overall plan yielded a schedule and locations for 478 units to be visited by the observer during his 239 work days.

ANALYSIS

Unbiased estimates for this sampling plan can be constructed by building on unbiased estimates for each stage of sampling (Figure 6).

Let

x_{hijk} = the observed quantity for area k during time period j on day i in stratum h .

D_h = number of days in stratum h .

d_h = number of days in the sample for stratum h .

(Let h denote a stratum with days divided into three time periods, i.e., $j = 1, 2, 3$).

$21 x_{hijk}$ is an unbiased estimate of the total over the whole river segment for time period j on day i in stratum h .

$2(21 x_{hi1k} \text{ or } 21 x_{hi3k}) + 21 x_{hi2k}$ is an unbiased estimate for the total over the whole river segment on day i in stratum h .

$\sum_i 2[21(x_{hi1k} \text{ or } x_{hi3k})] + 21 x_{hi2k}$ is an unbiased estimate for the total over the whole river segment for stratum h .

A similar construction gives unbiased estimates for strata which have days divided into four time periods, and, since strata are sampled independently, the sum of the estimates is an unbiased estimate of the quantity of interest.

An expression for the variance of an estimate of a total for a stratum with days divided into three periods is (Figure 7):

$$\begin{aligned} V_h = & k_1 \sum_{i=1}^{D_h} (\bar{x}_{hi..} - \bar{x}_{h...})^2 + \\ & + k_2 \sum_{i=1}^{D_h} (\bar{x}_{hi1.} - \bar{x}_{hi3.})^2 + \\ & + k_3 \sum_{i=1}^{D_h} \sum_{j=1}^3 \sum_{k=1}^{21} w_j (x_{hijk} - \bar{x}_{hij.})^2, \end{aligned}$$

where the k 's and w 's are functions of sample sizes and the "bars" and "dots" denote means per unit over the dotted subscripts. This expression reveals the parts of the variance that are associated with the three

-6-

sampling stages. The first term is a variation among daily means; the second term is a variation among the means for first and third time periods (the only ones being sampled); and the third term is a variation among areas. A similar expression can be derived for the strata which have days divided into four periods. The variance for an estimate of a total over all strata is the sum of the variances for the strata.

Without further assumptions, it is not possible to obtain unbiased estimates of the three parts of V_h from the sample data. The sampling at the second stage (periods within days) is a kind of systematic sampling where the first and third periods do not both get sampled on any given day. Also, the third stage sample (areas within periods) consists of only one area for each period. Because of these features, there was no way of computing unbiased estimates of the second and third terms in V_h . Similar remarks can be made about the V_h for other strata.

There are various ways of obtaining "approximate" estimates of the variances. An estimate which is an average of an overestimate and an underestimate is (Figure 8):

$$\begin{aligned} \hat{V}_h = & (2D_h - d_h) (2D_h)^{-1} \sum_{i=1}^{d_h} \left\{ 2(21 x_{hi1k} \text{ or } 21 x_{hi3k}) + \right. \\ & \left. + 21 x_{hi2k} - \left[\sum_{i=1}^{d_h} (2(21 x_{hi1k} \text{ or } 21 x_{hi3k}) + 21 x_{hi2k}) \right. \right. \\ & \left. \left. (d_h)^{-1} \right] \right\}^2 (d_h - 1)^{-1} d_h^{-1} . \end{aligned}$$

-7-

This amounts to a "corrected" value of an estimate of variation among estimates of daily totals. A similar estimate is appropriate for other strata, and a sum of estimates approximates the variance for an estimated total.

Estimates of dose were computed from information from more than one study (Figure 9). In the following expressions, capital letters denote information from the fishermen study and lower case letters denote information from other studies.

Average External Dose per Fisherman = $H\bar{r}/F$

H = estimated total fishing hours

$\bar{r}$ = estimated dose rate (mrem/hr)

F = estimated total number of fishermen

$\bar{r}$ is a background corrected estimate of average dose rate from samples of water, soil, and vegetation from the river.

Average Internal Dose per "Columbia River" Fish Eater =

$$= \sum_{\text{Species}} N_S w_S c_S f_S / e$$

N_S = estimated number of fish of species S caught during the year.

w_S = estimated average edible weight of flesh per fish for species S (gm/fish).

c_S = estimated average radionuclide concentration per gram of edible weight for species S (pCi/gm).

f_S = estimated dose factor for each pCi of intake for species S (mrem/pCi).

w_S and c_S come from studies of fish samples from the river; f_S comes from physics, chemistry and biology studies; and e comes from a household survey in the Tri-Cities.

-8-

ESTIMATES

Some estimates from the studies are as follows (Figure 10):

	<u>Estimate</u>	<u>"Standard Deviation"</u>
Number of Tri-City Fishermen	13,500	1,300
Number of Fishing Trips by Tri-City Fishermen	63,000	4,000
Numbers of Fish Caught by Tri-City Fishermen		
Bass	12,000	3,000
Trout	1,200	500
Steelhead	1,000	300
Spiney Ray	35,000	7,000

Average External Dose per Tri-City Fisherman = .6 mrem/yr.

Average Internal Dose per "Columbia River"
Fish Eater (Tri-City Residents) = 8.25 mrem/yr.

The one "standard deviation" error for these fishermen survey estimates varies from seven percent to 42 percent of the values of the estimates. This gives a rough idea of the adequacy of the sampling and provides some guidance for design of similar research projects.

CONCLUSIONS AND COMMENTS

It might have been possible to improve the fishermen study by using unequal probability sampling of areas at the third stage. That is, allowing higher probabilities for selection of heavily fished areas and adjusting the estimation procedure might have resulted in a reduced variance comparable to what might have been achieved had we been able to stratify areas. Also, a "creel census" approach, involving selection of time/space points at

random and travel of the interviewer through the area of interest, might have been a more efficient use of resources; but there were some administrative difficulties associated with such an approach.

This time/space frame sampling and estimation procedure illustrates a technique that can be used in studies of the "effects" of pollution on human populations. The frame definitions and sampling procedures are affected by the research resources that are available, and, as a result, the analysis procedures may not be standard or straightforward.

Finally, though the total research problem is probably best approached by studying separate parts and involving several disciplines, it should be a true interdisciplinary effort with close communications among researchers. Combining data collection and analysis results from several independent studies tends to produce biased estimates of the quantities interest. That is, a product of means is not in general equal to a mean product. Statistical theory can provide a unifying thread for all of the data collection and analysis procedures of the several studies, and can help to eliminate biases.

STUDY REGION

FIGURE 1

TIME/SPACE FRAME

AREA	FEB. 10				FEB. 11				FEB. 12				FEB. 13				FEB. 14				FEB. 15			
	6-10	10-2	2-6		6-10	10-2	2-6		6-10	10-2	2-6		6-10	10-2	2-6		6-10	10-2	2-6		6-10	10-2	2-6	
1																								
2																								
3																								
4																								
5																								
6																								
7																								
19																								
20																								
21																								

AREA	JUNE 10				JUNE 11				JUNE 12				JUNE 13				JUNE 14				JUNE 15			
	5-9	9-1	1-5	5-9	5-9	9-1	1-5	5-9	5-9	9-1	1-5	5-9	5-9	9-1	1-5	5-9	5-9	9-1	1-5	5-9	5-9	9-1	1-5	5-9
1																								
2																								
3																								
4																								
5																								
6																								
20																								
21																								

FIGURE 2

SAMPLE AREAS

FIGURE 3

STRATIFICATION

FIGURE 4

THREE STAGE SAMPLING

DRAFT

FIGURE 5

UNBIASED ESTIMATES

X_{hijk} = Quantity observed for the K^{th} area during the j^{th} time period for the i^{th} day in stratum h

(1) $2 \sum_i X_{hijk}$ = unbiased estimate for total for j^{th} time period for day i in stratum h

(2) $2 \left[2 \sum_i (X_{hi1K} \text{ or } X_{hi3K}) \right] + 2 \sum_i X_{hi2K}$ = unbiased estimate for total for i^{th} day in stratum h

(3) $D_h \sum_{i=1}^{d_h} \left\{ 2 \left[2 \sum_i (X_{hi1K} \text{ or } X_{hi3K}) \right] + 2 \sum_i X_{hi2K} \right\} / d_h$ = unbiased estimate of total for h^{th} stratum, where

D_h = number of days in h^{th} stratum

d_h = number of days in sample from h^{th} stratum

Add estimates over strata.

VARIANCE OF ESTIMATE

~~Variance of Estimate~~

$$\begin{aligned}
 V_h = & K_1 \sum_{i=1}^{D_h} (\bar{X}_{hi...} - \bar{X}_{h...})^2 + \\
 & + K_2 \sum_{i=1}^{D_h} (\bar{X}_{hi1.} - \bar{X}_{hi3.})^2 + \\
 & + K_3 \sum_{i=1}^{D_h} \sum_{j=1}^3 \sum_{k=1}^{21} w_j (X_{hijk} - \bar{X}_{hij.})^2
 \end{aligned}$$

Center

ESTIMATE OF VARIANCE OF ESTIMATE

~~"Estimate" of Variance of Estimate~~

$$\hat{V}_h = (2D_h - d_h)(2D_h)^{-1}.$$

$$\sum_{i=1}^{d_h} \left\{ 2 \left[21(X_{hi1K} \text{ or } X_{hi3K}) + 21 X_{hi2K} \right] - \right.$$

$$\left. - \left[\sum_{i=1}^{d_h} 2(21(X_{hi1K} \text{ or } X_{hi3K})) + 21 X_{hi2K} \right] (d_h)^{-1} \right\}$$

$$(d_h - 1)^{-1} d_h^{-1}$$

Cont'd

DOSE CALCULATIONS

Tri-City

Average External Dose Per _^ Fisherman =

$$= H \bar{r} / F$$

$$(\text{Hours})(\text{mrem/Hr.}) / (\text{Number})$$

Average Internal Dose Per "Columbia
River" Fish Eater =

$$= \sum_{\text{Species}} N_s w_s c_s f_s / e$$

$$= \sum_{\text{Species}} (\text{No. of species})(\text{gm./fish})(\text{pCi/gm.})(\text{mrem/pCi})$$

(Number)

SOME RESULTS

DRAFT

~~Some~~
Estimates

	<u>Estimate</u>	<u>"Standard Deviation"</u>
Number of Tri-City Fishermen	13,500	1,300
Number of Fishing Trips by Tri-City Fishermen	63,000	4,000
Number of Fish Caught by Tri-City Fishermen		
Bass	12,000	3,000
Trout	1,200	500
Steelhead	1,000	300
Spiney Ray	35,000	7,000

Average External Dose per Tri-City Fisherman = .6 mrem
per year

Average Internal Dose per "Columbia River"
Fish Eater (Tri-City Residents) = 8.3
mrem per
year

FIGURE 10