Stochastic studies on the atomic bomb casualties

Masuyama, Motosaburo
Institute of Physical Theapy & Internal Medicine, Tokyo University | Hygieno-meteorological Laboratory, Central Meteorological Observatory

STOCHASTIC STUDIES ON THE ATOMIC BOMB CASUALTIES.

By Motosaburo MASUYAMA[1]

Contents

Part I. Death rate distance curve at Hiroshima.

1. Introduction

It would be convenient to distinguish between "statistics" (descriptive statistics) and "stochastics" (inductive statistics) in what follows. Statistics consists of a mere collection and array of data, but stochastics consists of estimation, testing hypotheses and designs. In classical statistics the population means the total of individuals which has certain characteristics and investigation of individual in the population should be done correctly and exhaustively. After the investigation of each individual is over, the results are represented by tables, diagrams etc. That is all and there is no objective method of estimation or test criterion of hypotheses. Therefore the classical statistics may be suitable to describe historical facts but not scientific laws.

We investigate to know scientific laws. To know scientific laws, it is impossible for infinite population and trivial for finite population to investigate exhaustively the population. As for infinite population the statement is clear, I will illustrate what I say for finite population. If we could investigate exhaustively the individual of the finite population P and get a law L, then this law would do nothing from scientific point of view. For there remains no other individual which follows the law L except that in P. That is to say, the law L should not be applied in future or in any other region etc. Such a limited law is not worthy of a scientific law. In another words, it is essential for the establishment of scientific laws that only a part of the population, viz. a sample is permitted to be investigated. As there is a finite margin between sample and population, it is important to

estimate the population characteristics, viz. parameters and to test the hypotheses concerning parameters as well as we can. These two main problems are studied in stochastics based on the modern mathematical theory of probability.

2. Theoretical

It is well known fact that the dosage-mortality curve for many toxic substances are represented by the normal distribution function

$$(1)\qquad P(x) = \int_{-\infty}^{x} exp. \{-(x-m)^2/(2\sigma^2)\} dx/(2\pi\sigma^2)^{1/2}, \quad x = log\ D$$

where D means the concentration of substance and $P(x)$ the survival fractions of animals[1]. Two parameters m and σ are respectively the population mean and the population standard deviation of resistance of animal against the toxic substance.

I have supposed that if we could know the killing power K of the atomic bomb — it may be caused by neutron, by gamma ray or by blast etc. — as a function of distance R, we might expect that the above formula (1) would hold putting $x = K$. Such an abscissa x will be called hereafter the intensity scale.

As a matter of fact we could not know K itself and we have assumed that K follows the inverse square law. This assumption might be approximately valid at Hiroshima, for the neutron effect was limited in very small region near the bomb centre where almost all people died and within a circle of radius 2 km from the bomb centre there is no hill.

As the functional form of $P(x)$ does not alter by linear transformation of the variable x, we put $x = 1/R^2$. In what follows we have assumed the height of explosion $h = 0.5\ km$ and calculated R from the horizontal distance r and h $(R^2 = h^2 + r^2)$.

Let the size of random sample at distance r_1 be N_1 and the number of dead at this point be k_1. Then the maximum likelihood estimate of $P(x_1)$ is equal to k_1/N_1 [2]. As we do not know a priori the numerical values of two parameters, we must at first estimate those values from the sample as follows. If we put

$$(2)\qquad\qquad t = (x - m)/\sigma,$$

then we have instead of (1)

$$(3)\qquad P(x) = \int_{-\infty}^{t} exp. \{- t^2/2 \} d\,t/(2\pi)^{1/2}.$$

The right hand side of this equation contains neither m nor σ and if we give the numerical value of $P(x)$, which should be in the range 0 to 1, we have exactly one value of t and vice versa. Therefore if we put

$$(4) \qquad k_1/N_1 = \int_{-\infty}^{t_1} exp. \; \{-t^2/2\} \, dt/(2\pi)^{1/2},$$

then there must exist linear relation

$$(5) \qquad t_1 = (x_1 - m)/\sigma,$$

provided that there is no sampling errors.

Actually there is always sampling errors and we estimate two parameters by some means, for example by the method of least square on the (x_1, t_1)-diagram. The method used in the following is that illustrated in Fisher-Yates' table [2] . The only difference lies in the point that the weighting coefficient here is $w = N Z^2/(PQ)$ but not $Z^2/(PQ)$.

After getting the estimates of two parameters we can calculate the theoretical — strictly speaking semi-empirical — value of $P(x)$.

3. Basic data

Basic data were collected by the Japanese Medical Party of the Committee for the Atomic Bomb Casualties under the direction of the present author on 22 November '45. The results are represented in the following table where r is arranged in a functional scale. We selected at random 28 small districts utilizing a random sampling number on the map of Hiroshima and selected at random some families in each district and asked about the fate of each member of the family and that of adjacent families.

r	k	N
0.49	121	124
0.52	33	41
0.55		
0.58		
0.60		
0.64		
0.68	13	15
0.71		
0.75		
0.79		
0.86	32	34
0.91	124	131
0.98		
1.06	5	7
1.16	16	27
1.28	28	47
1.44	20	51
1.67	77	217
2.03	38	180

4. Results

Neglecting three points near the bomb centre (v.i.) and starting from the following provisional probit Y, we have for goodness of fit

$$\chi^2 = 1.78, \qquad n = 8 - 2,$$
$$P = 0.94.$$

Y	q
6.56	2/34
6.26	7/131
5.66	2/71
5.36	1/27
5.06	47/19
4.76	31/51
4.46	140/217
4.16	142/180

There is no indication of departure from the presumed functional form of the population. As the fitting is good, we can test whether the rejection of three points near the bomb centre is permissible or not.

Let the value to be rejected be η and the numerical value of F in the F-distribution for $n_1 = 1$ and $n_2 = n$ and level of significance α be $F_n^1(\alpha)$, and we have

$$(6) \qquad F_n^1(\alpha) \leq n \, (\eta - Y)^2/[x^2\{1 + 1/S(w) + (x - \bar{x})^2/A\}].$$

Being based on this formula we can conclude that the observed death rates

near the bomb centre are too small to be expected from the theory. The reason why the fitting is not good near the centre may be explained in two ways. 1) To answer to the questionnaire there must remain at least one person alive in a selected family. 2) There are many buildings near the bomb centre and the concrete walls act as shelters. Actually the geographical distribution of death rates shows that our formula is valid in the areas near the bomb centre where there is no concrete wall.

5. Conclusion

Neglecting various attributes — sex, age, kind of shelter, kind of cause of death etc.— we can represent the death rate distance curve by the formula (1) with $x = 1/R^2$. It is remarkable that this formula contains only two parameters. The present author supposes that the formula would be valid for other symptomes, if we could choose a suitable intensity scale experimentally or theoretically.

Part II. Death rate date curve at Hiroshima.

1. Introduction

In the first paper [1] I have studied the atomic bomb casualties as a function of distance. In this second paper I have intended to study it as a function of time.

2. Theoretical

I have pointed out rather intuitively during the war that the appearance of some symptoms in acute diseases especially in contagious diseases follows approximately the exponential law. Analogically I have supposed that the death rate date curve is represented by the probability density function

$$(1)\qquad f(t) = e^{-t/\sigma}/\sigma \qquad \text{after the bombardment}$$
$$= 0 \qquad \text{before the bombardment.}$$

The probability distribution function, which corresponds to empirical cumulative relative frequency, is

$$(2)\qquad P(t) = 1 - e^{-t/\sigma} \qquad \text{for} \quad t \geq 0$$
$$= 0 \qquad \text{for} \quad t < 0.$$

If the assumption on the specification of the population is true, the maximum likelihood estimate of the unique parameter σ, the decay constant, or the mean survival period is equal to the sample mean of t, viz. $\bar{t}$.

3. Basic data and method of calculation

The basic data was obtained from the reports of post mortem examinations, submitted to the Kabe police station. At that time to bury the bombed people there needed generally the post mortem examination and the permission of policeman.

date	u	N	Nu
6/VIII	0	119	0
	1	85	85
	2	43	86
	3	31	93
	4	25	100
25	5	7	35
	6	8	48
	7	9	63
	8	2	16
	9	2	18
18/IX	10	0	0

$$331 \overline{)544}$$
$$1.644$$

The method of calculation will be obvious from the table. Putting $t = 4u$, we have as an estimate

$$\bar{u} = 1.644 + 1/2 = 2.144.$$

Solving numerically the equation

$$e^{-d/\sigma} = 1/2,$$

We have

$$d = 5.944 \text{ (days)}.$$

That is to say, daily death becomes one half after every 5.944 days. To test the goodness of fit, we must have at first the theoretical frequencies, which are obtained as follows.

u	$u/\bar{u}$	$z = e^{-u/\bar{u}}$	$1 - z$	$\Delta(1 - z)$	$N_c = 331 \, \Delta(1 - z)$
0	0.0000	1.0000	0.0000	0.3728	123.40
1	0.4665	0.6272	0.3728	0.2338	77.39
2	0.9328	0.3934	0.6066	0.1467	48.56
3	1.3993	0.2467	0.7533	0.0920	30.45
4	1.8657	0.1547	0.8453	0.05765	19.08
5	2.3821	0.09705	0.90295	0.03618	11.98
6	2.7985	0.06087	0.93913	0.02269	7.51
7	3.2649	0.03818	0.96182	0.01424	4.71
8	3.7314	0.02394	0.97606	0.00892	2.95
9	4.1978	0.01502	0.98498	0.00560	1.35
10	4.6642	0.00942	0.99058		3.12
					331.00

Thus we have

$$\chi^2 = S(N_0 - N_c)^2/N_c = 4.43, \ n = 7 - 2, \ P = 0.50.$$

The fitting is very good so that we can not reject the hypothesis that the population is exponential. For any convex function there exists a functional relation

$$\{f(x_1) + f(x_2)\}/2 \geqq f\{(x_1 + x_2)/2\}.$$

Therefore it leaves some room for consideration that non-grouped data follows the exponential law. At any rate we might suppose that the grouped data follow the exponential law and it is sufficient for practical purpose.

4. Application

Hiroshima had been visited by the violent typhoon in the middle of September '45 and the patients with radiation sickness suffered afterwards from various complications, such as malnutrition, pneumonia or septicaemia etc., for the means of transportation, humble barracks and crops were destroyed and they could hardly go to hospital or get something to eat in the heavy rain and in the strong wind. Therefore it would be interesting, if we could estimate the total number of dead who would not have died, if only there had not been the violent typhoon.

Our method of estimation is as follows. Having got the formula (1) utilizing the earlier readings before typhoon, it is now possible to forecast the total number of dead and if the amount of dead people does not match with that forecasted, the difference means the dead from the complications due to the typhoon.

date	u	N_x
6/VIII	0	
25	5	46,185
30/XI	28	78,150

According to the reports to the Pacification Section of the Hiroshima Prefectural Office there were 46185 deaths and 17429 missing on 25 August '45 within Hiroshima Prefecture.

As $1/\bar{u} = 0.4664$, we have

$$P(u) = 1 - e^{-0.4664\,u}$$

for the semi-empirical cumulative relative frequency. Putting $u = 5$, which corresponds to the date 25 August '45, we obtain

$$P(5) = 0.9029 .$$

That is to say, according to the formula we can expect about 10 % increment of death after 25 August '45. For 30 November '45 we have approximately $u = 28$ and

$$P(28) = 0.999978 .$$

That is to say $P(28)$ might be considered as one. We can suppose that 46,185 and $(46,185 + 17,429) = 63,614$ are respectively the minimum and the maximum of dead for $u = 5$, and

$$N_\infty P(5) = 46,185 \text{ or } 63,614 ,$$

we have as the expected total number of dead at the end of November

$$N_\infty = 51,150 \text{ (min.) or } 70,460 \text{ (max.)} .$$

According to the reports on 30 November '45 there were 78,150 deaths and 13,983 missing (total $= 92,133$). Therefore the differences 7,690 (min.) and 40,983 (max.) might be supposed to survive, if we could have given sufficient drugs, nutrition or suitable protection against typhoon at that time. We must, of course, take into consideration the situation that the bereaved

families could receive from the city hall, from the municipal or village office condolence money, if they announced that the dead had been at Hiroshima when the bomb fell. This situation might exagerate the total number of dead.

5. Conclusion

The death rate date curve is represented well by the exponential function. The reason is not clear at present. We may consider the analogy of mean free path of the kinetic theory of gas, that of the energy distribution of brain wave [5] etc. Without any further experiment we can not decide which is true.

Part III. Comparison of the death rate of male
with that of female at Hiroshima

1. Introduction

According to the experimental investigation of malnutrition in USSR during the war the female is more resistant than the male. It is also a well known fact that the female is more resistant to bleeding and to cancer. I have intended in this paper to compare stochastically the death rate of the male with that of the female at the same distance from the bomb centre.

2. Basic data

The basic data were collected by the method of questionnaire which was designed mainly by LT. M. E. HABEL and was slightly modified by the present author. The received questionnaires were sorted in our Institute of Statistical Mathematics, Ministry of Education. The sorted cards were arranged according to sex and distance. The distance r used here is represented in a functional scale, which I have used in my first paper. The date of sampling was the middle of December '45.

3. Results

The results are shown in the following table. Almost all military forces are excluded in those figures. The fitting is very good for $r < 1.28\,km$, but not for $r < 1.28\,km$. For $r < 1.28\,km$ the death rate for male and that for female are almost equal but for $r < 1.28\,km$ the death rate for male is significantly higher than that for female. For example let us take as a cross section the death rate in the Hiroshima Army Hospital near the bomb centre. According to the report of the Army Medical School[2] we have the follow-

r	Male		Female	
km	k	N	k	N
0.49	266	283	299	322
0.52				
0.55	55	60	52	65
0.58				
0.60				
0.64	178	223	163	220
0.71				
0.75				
0.79	121	143	110	145
0.85				

0.91	252	296	283	340
0.98				
1.03				
1.16	370	487	267	415
1.28				
1.44	210	493	284	642
1.67	148	509	206	589
2.03	134	930	168	1,033
2.73	112	1,349	105	1,731
5.75	71	2,757	73	3,501

ing figures of the date 19 November '45, where male and female mean respectively the reserve army surgeon and the nurse. To underestimate the death rate of male we add the number of male missing to that of healthy and to overestimate the death rate of female we add the number of female missing to that of dead, then we obtain

$$\chi_c^2 = 9.335, \qquad n = 1, \qquad P \ll 0.01,$$

where the suffix c means the Yates' correction.

	dead	missing	healthy
male	95	23	4
female	64	4	48

To our regret, we can not know in detail from the report how they were protected at that time. This result could merely be the fact that there were more male outside without any protection when the bomb fell. But it would be noteworthy that for $r < 1.28$ km there is no highly significant difference between the male death rate and the female one.

It is not clear at present what is meant by $r = 1.28$ km. It might mean the effective radius of neutron action on the human body.

There is another noteworthy fact. That is to say, there is a low death rate region within the circle of radius 1 km[3]. For example, the death rate for $r = 0.62$ km is approximately equal to that for $r = 1.10$ km. We could assume that this means the protecting action of concrete buildings, for there were many concrete buildings concentrated within this low death rate region. If this assumption be right, we could deduce from this fact a crude estimate of, so to speak, mean air-equivalent thickness of concrete wall and accordingly an estimate of hardness of Gamma ray. Such an estimation will be done soon by Mr. Murati, member of the Institute of Physical and Chemical Research.

4. Conclusion

My formula introduced in the first paper is valid for large sample gathered by Lt. Habel and me a month later than the previous sample gathered by the Japanese Medical Party.

It is pointed out that the female seems to be more resistant than the male but the difference is significant only near the bomb centre.

(The figure 12.15 in " The Effects of Atomic Weapons "McGraw Hill Book Co. Inc., 1950, seems to be the " smoothed " curve of our findings (note added on 10th April, 1951)).

2) Gensi Bakudan ni yoru Hirosima Sensai Igakuteki Tyôsa Hôkoku, 30 Nov. '45.

3) It might not be attributable to the sampling error, for the same tendency is obvious for three groups, viz. male group and female group of the present survey and mixed group of the previous survey.

Part IV. Death rate date curve at Nagasaki.

1. Introduction

I have found at Hiroshima that the cumulative death rate date curve is represented by the formula

$$(1) \qquad P(t) = 1 - e^{-t/\sigma},$$

where the only parameter σ means the mean survival period [6] . It is the purpose of this paper to verify whether or not this law holds at Nagasaki and to compare two parameters σ_H and σ_N if the law holds. The suffixes H and N mean respectively Hiroshima and Nagasaki.

2. Basic data

The basic data is obtained from the records of post mortem examinations reported to the Section of Criminal Matters[4] of the Prefectural Office.

date	u	N
10/VIII	0	122
	1	67
	2	57
	3	26
26	4	22
	5	13
	6	14
	7	2

The bomb fell on 9 August '45. According to the reports the total amount of dead on this day is 2,189. But this is too large compared with that on the next day. The reason will become obvious if we consider the orographical situation of the afflicted area. (c.f. the map of the city). It lies in the valley. After the bombardment a conflagration broke out and many people escaped into pathless hills. Accordingly most of the severely wounded or irradiated people could scarcely be found just after the disaster and died unnoticed. As far as I know there are many dead who were found after one month or later. As the exact date of death was uncertain, such people were regarded by the police to have been killed instantaneously.

I have assumed therefore that the people in the hills died following to the same law as the people received into hospitals. If this assumption is true, we can cut off the data of the date 9 August. The functional form and the numerical value of paramater σ do not alter by this process. Putting $u = 4t$, we have as an estimate of σ

$$\bar{u} = 2.08.$$

For the goodness of fit we have

$$\chi^2 = 4.97, \qquad n = 7 - 2, \qquad P = 0.50.$$

after grouping the last two observed frequencies. The fitting is good and our assumption seems to be rational. It is a rather unexpected result that the numerical value of σ_N differs from σ_H by less than 3 %.

To test the null-hypothesis

$$H_0 : \sigma_H = \sigma_N,$$

---

4) The death by a bomb is unnatural one for civilian.

we can use F-distribution putting

$$F = \bar{u}_H/\bar{u}_N, \qquad n_1 = 2N_H, \qquad n_2 = 2N_N \ [7].$$

In our case $N_H = 331$, $N_N = 323$ and

$$F = 1.02 < F_{646}^{662} \ (0.05)^{5)}$$

viz. we can not reject the hypothesis H_0.

3. Conclusion

The exponential law found at Hiroshima holds at Nagasaki and the unique parameters are the same.

References

[1] S. KATO: Nihon Yakubutugaku Zassi (Japanese Journal of Pharmacology), 36 (1942), 588.

[2] S. HURUYA: Tôkei Sûri Kenkyû (Bulletin of Mathematical Statistics), 1 (1942), 124.

[3] R. A. FISHER-F. YATES: Statistical Tables, 1938, c.f. Garwood: Biometrika, 32 (1941), 46.

[4] M. MASUYAMA: Stochastic Studies on the Atomic Bomb Casualties at Hiroshima, I. Death rate distance curve.

[5] M. MASUYAMA: Igaku to Seibutugaku (Medicine and Biology), 5 (1944), 483, 488, 529, 583; M. MASUYAMA: Kwagaku (Science), 14 (1944), 266; K. MOTOKAWA: ibid. 15 (1945), 35.

[6] M. MASUYAMA: Stochastic Studies on the Atomic Bomb Casualties at Hiroshima. II. Death rate date curve.

[7] M. MASUYAMA: Stochastic Studies on Brain Waves, Kwagaku (Science), 14 (1944), 266.

---

5) New notation introduced by the author and used in Japan. c.f. M. MASUYAMA: Syôsûrei no Matomekata to Zikken Keikwaku no Tatekata (How to manage small sample and how to design experiments), 2nd ed. (1944), 127. c.f. my first paper on the atomic bomb casualties.