1st extra-mural use of Student's *z* distribution

Material courtesy of Warren Winkelstein, Jr.

Vignettes of the history of epidemiology:
Three firsts by Janet Elizabeth Lane-Claypon.,
American Journal of Epidemiology, 2004 Jul 15;160(2):97-101.

VIGNETTE: Janet Elizabeth Lane-Claypon
A Forgotten Epidemiologic Pioneer
Epidemiology 2006; 17(6):705.

Presentation by James Hanley

Department of Epidemiology, Biostatistics & Occupational Health
McGill University
October 27, 2008.

After he met Fisher for the first time, in September 1922, Gosset sent Fisher a copy of Student's tables ...

> *"as you are the only man that's ever likely to use them!"*

The tables had been in constant use at the brewery for the past 14 years, but they were almost unknown elsewhere. Gosset accepted their neglect by the statistical establishment with a good grace. In contrast, Fisher's sense of justice was outraged. In a "Historical Note" at the beginning of Statistical Methods for Research Workers (14th ed., 1970, p. 23), we read:

> *"Student's" work was not quickly appreciated (it had, in fact, been totally ignored in the journal in which it had appeared), and from the first edition it has been one of the chief purposes of this book to make better known the effect of his researches, and of mathematical work consequent upon them.*

Joan Fisher Box, "Gosset, Fisher, and the t Distribution" *The American Statistician* 1981; 35(2):61-66.

Janet Elizabeth Lane-Claypon

A Forgotten Epidemiologic Pioneer

Warren Winkelstein, Jr.

Epidemiology

17(6)

2006

p705

In 1959, Mantel and Haenszel published their classic paper on "Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease."[1] The first sentence of their paper states "The present-day controlled retrospective studies of cancer date from the Lane-Claypon paper on breast cancer published in 1926."

So who was Lane-Claypon?

Janet Elizabeth Lane-Claypon was born into a wealthy English family in rural Lincolnshire in 1877. She entered University College, London, in 1899, receiving her Bachelor's degree with first-class honors in 1902, a DSc in physiology in 1905, and an MD in 1910.

She was the first woman ever to receive a research scholarship from the British Medical Society. Her doctoral research (on the developmental histology of the ovary and the hormonal control of lactation) was carried out in the laboratories of the renowned physiologist Ernest Starling. Her work was extensively cited in Marshall's 1910 textbook on *The Physiology of Reproduction*, the first textbook on this topic. In 1907, Lane-Claypon joined the staff of the Lister Institute of Preventive Medicine, where she researched the bacteriology and biochemistry of milk. In 1909, she received a Jenner Fellowship from the Institute to study maternal and child health programs in Europe. Her subsequent career involved epidemiologic research, educational administration, and advocacy for maternal and child welfare.

Her 1926 paper cited by Mantel and Haenszel was titled *A Further Report on Cancer of the Breast, With Special Reference to Its Associated Antecedent Conditions.*[2] This 135-page document described a study of 500 hospitalized cases and 500 controls. The methodology (including a detailed questionnaire) was meticulously described, and the data were exhaustively analyzed and interpreted using contingency tables and standard statistical procedures. Remarkably, the findings included most of the currently recognized risk factors for breast cancer.

Lane-Claypon published 3 books and 30 papers, 2 of which (besides the breast cancer report) might be considered classics. A 1912 paper assessed weight gain in infants fed boiled cows' milk compared with human breast milk. The study used for the first time the historical cohort design, used

Janet Elizabeth Lane-Claypon

Student *t* to analyze the data, and controlled for confounding. Her 1926 study of survival from breast cancer surgery took into account competing risks and used a life-table survival analysis.

Lane-Claypon married at the age of 52. Restrictions on the employment of married women forced her from the civil service, whereupon she terminated her professional activities and moved to the country. She lived quietly with her husband until her death in 1967 at the age of 90.

REFERENCES

1. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*. 1959;22:719–748.
2. Lane-Claypon JE. *A Further Report on Cancer of the Breast, With Special Reference to Its Associated Antecedent Conditions*. Reports on Public Health and Medical Subjects No. 32. Ministry of Health. London: Published by His Majesty's Stationary Office; 1926.

Report to the Local Government Board upon the available data in regard to the value of boiled milk as a food for infants and young animals. By Janet E. Lane-Claypon, M.D., D.Sc. (Lond.).

February 24th, 1912.

(This report embodies the result of an inquiry undertaken in connection with the Board's Grant for Auxiliary Scientific Investigations.)

PART IV.—METHOD AND RESULTS OF WORKING UP THE BERLIN MATERIAL.

The material for this research was obtained, as already mentioned (p. 4), from the Infant Consultation of the Naunyn Strasse in Berlin. This consultation is conducted by Dr. Ballin, to whom I am deeply indebted for permission to use his material.

Source of the Material.—Six years ago, infant consultations were started by the municipality of Berlin, under the auspices of a special fund, the Schmidt-Gallisch Stiftung.

Four were first started, and then another, and finally two more, thus making seven in all, in different parts of Berlin.

Each of these consultations is in charge of a medical officer who has made a special study of the diseases and ailments of children. The attendance is so large that assistants have been appointed to assist the medical officer in the discharge of his duties. The consultations are held daily, and at the Naunyn Strasse (where the material here dealt with was obtained) the average daily attendance is about 100 babies.

Each Consultation has its own staff of health visitors attached. These are women who have been thoroughly trained for work among children, and are appointed by the municipality to visit the homes of the babies who are brought up to the consultation. Their duty is to instruct the mothers in the general care and hygiene of the infant in accordance with the directions given by the medical officers at the consultation.

The clientèle of the consultation consists exclusively of the working classes. The fathers of the children who are brought up to the consultation are for the

Selection of Material

[...] Two main series of infants :-

(1) Healthy babies of the average artisan class, fed upon milk in various forms, in order to have a control consisting of the average baby.

(2) Healthy babies of the same class but fed only upon boiled cows' milk, in order to study the difference, if any, produced upon the average baby of the class by feeding it exclusively upon boiled milk, as compared with the infant of class (1).

It was decided to exclude from the control series all babies who had received less than four months breast feeding, taking into consideration further points described below.

Table I.

Showing the age of first attendance and of leaving the Consultation of the Babies of the Control or Breast-fed Series.

I. Age in weeks.	II. No. brought at each day.	III. Age on leaving (in months).							IV. Total in each week.
		4–5.	5–6.	6–7.	7–8.	8–9.	9–10.	10–12.	
1	3. 2. 0. 3. 0. 4. 4.	2	1	1	0	4	0	8	16
2	5. 7. 8. 6. 6. 7. 10.	9	4	5	3	1	3	24	49
3	8, 11, 9, 15, 12, 5, 7.	8	11	5	6	8	9	20	67
4	9. 4. 3. 8. 0. 3. 4.	2	3	7	5	1	2	11	31
5	1. 2. 4. 5 2. 3. 2.	—	1	3	1	2	1	11	19
6	5. 2. 3. 5. 2. 2. 3.	—	6	3	1	3	0	9	22
7	3. 1. 2. 3. 4. 2. 2.	—	4	1	5	1	2	4	17
8	1. 2. 1. 0. 3. 3. 0.	—	3	2	2	1	0	2	10
9	2. 2. 3. 4. 3. 2. 2.	—	3	6	2	2	1	4	18
10	1. 1. 1. 1. 2. 0. 1.	—	—	3	1	1	0	2	7
11	2. 4. 3. 3. 0. 0. 1.	—	—	4	1	1	1	6	13
12	3. 0. 0. 0. 3. 5. 1.	—	—	—	4	1	2	5	12
13	2. 1. 0. 1. 2. 0. 2.	—	—	—	1	0	2	5	8
14	1. 1. 0. 0. 0. 0. 2.	—	—	—	1	0	1	2	4
15–18	——	—	—	—	—	3	1	3	7
		21	36	40	33	29	25	116	300

TABLE II.

Showing the age of first attendance and of leaving the Consultation of the Babies of the boiled cows' Milk Series.

I. Age in weeks.	II. No. brought at each day.	III. Age on leaving (in months). 4–5.	5–6.	6–7.	7–8.	8–9.	9–10.	10–12.	IV. Total in each week.
1	2. 0. 0. 1. 2. 0. 2.	1	0	1	1	0	0	4	7
2	2. 1. 1. 1. 5. 7. 2.	0	1	4	2	0	3	9	19
3	6. 4. 3. 8. 2. 4. 2.	0	1	3	0	1	2	22	29
4	6. 2. 4. 4. 4. 4. 3.	3	3	2	1	3	1	14	27
5	1. 1. 3. 1. 4. 2. 4.	0	0	2	0	1	3	10	16
6	1. 2. 1. 2. 4. 1. 1.	3	0	2	3	0	1	3	12
7	1. 2. 2. 2. 2. 0. 5.	—	—	1	2	0	1	10	14
8	1. 3. 2. 3. 3. 2. 4.	—	—	—	2	0	2	14	18
9	2. 0. 3. 4. 2. 0. 2.	—	—	2	1	2	2	6	13
10	0 3. 2. 1. 1. 0. 2.	—	—	—	1	0	1	7	9
11	0. 3. 1. 1. 1. 0. 2.	—	—	1	2	2	1	2	8
12	0. 3. 1. 0. 0. 1. 1.	—	—	—	2	0	0	4	6
13	1. 2. 4. 0. 1. 1. 1.	—	—	1	1	2	1	5	10
14	1. 1. 1. 1. 1. 1. 0.	—	—	—	—	—	1	5	6
15–18	——	—	—	—	—	3	1	6	10
		7	6	19	18	13	20	121	204

I. (control) Breastfed series

TABLE III.

Showing the average weights of the babies of the control or breast-fed series, grouped in periods of eight days, and the number of observations made.

I. Age in days.	II. No. of observations on each day.	III. Total No.	IV. Average Weight.
			Grammes.
1–8	3. 2. 0. 4. 0. 4. 6. 5.	24	3,185
9–16	10 10 9. 11. 11. 12. 12. 22.	97	3,317
17–24	16. 21. 23. 19. 19. 20. 13. 13.	144	3,507
25–32	20. 24. 16. 19. 20. 19. 31. 18.	167	3,746
33–40	16. 19. 19. 27. 23. 25. 18. 15.	162	3,939
41–48	29. 18. 27. 19. 19. 24. 29. 27.	192	4,119
49–56	18. 19. 33. 25. 19. 18. 26. 17.	175	4,291
57–64	29. 29. 31. 24. 30. 29. 25. 27.	224	4,443
65–72	23. 20. 35. 27. 27. 25. 27. 35.	219	4,638
73–80	26. 24. 28. 19. 32. 34. 31. 21.	215	4,737
81–88	34. 40. 27. 29. 31. 25. 26. 26.	238	4,937
89–96	35. 22. 41. 32. 27. 30. 23. 22.	232	5,079
97–104	26. 27. 25. 33. 61. 52.	224	5,191
105–112	54. 58. 45. 60.	217	5,380
113–120	54. 54. 58. 50.	216	5,666
121–128	60. 57. 61. 53.	231	5,659
129–136	52. 61. 59. 55.	227	5,757
137–144	60. 48. 56. 56.	220	5,929
145–152	47. 50. 51. 43.	191	6,033
153–160	47. 56. 45. 43.	191	6,237
161–168	50. 50. 45. 39.	184	6,274
169–176	47. 47. 48. 40.	182	6,312
177–184	41. 41. 46. 43.	171	6,434
185–192	41. 32. 27. 47.	147	6,458
193–200	32. 28. 41. 27.	128	6,664
201–208	32. 37. 29. 35.	133	6,709
209–216	31. 31. 30. 28.	120	6,734
217–224	32. 37. 35. 32.	136	6,798
225–232	54. 56.	110	6,778
233–240	55. 54.	109	6,886
241–248	58. 42.	100	6,891
249–256	42. 47.	89	7,118
257–264	38. 40.	78	7,276
265–272	46. 40.	86	7,217
273–280	43. 38.	81	7,388
281–288	36. 41.	77	7,281
289–296	28. 35.	63	7,608
297–304	34. 38.	72	7,567
305–312	36. 32.	68	7,801
313–320	22. 36.	58	7,555
321–328	32. 26.	58	7,753
329–336	37. 20.	57	7,704
337–344	26. 22.	48	7,752
345–352	21. 22.	43	8,034
353–360	22. 15	37	8,077
361–368	19. 15.	34	8,274

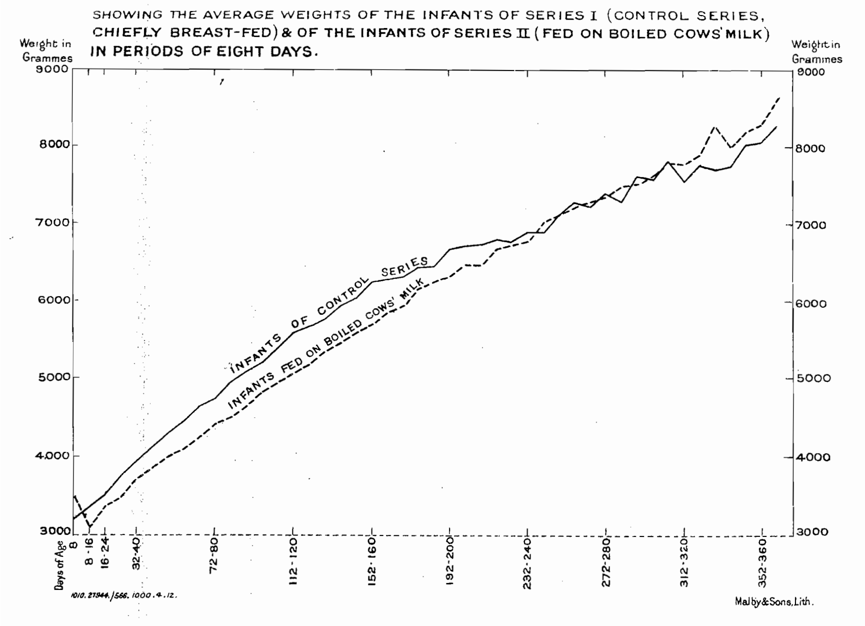
II. Boiled cows' milk series

TABLE IV.

howing the average weights of the babies of the boiled cows' milk series, grouped in periods of eight days, and the number of observations made.

I. Age in days.	II. No. of observations on each day.	III. Total No.	IV. Average weights.
			Grammes.
1–8	2. 0. 0. 1. 2. 0. 3. 2.	10	3515
9–16	2. 3. 2. 6. 7. 5. 9. 6.	40	3090
17–24	7. 11. 6. 15. 11. 13. 4. 9.	76	3358
25–32	15. 17. 15. 14. 14. 13. 11. 14.	113	3472
33–40	15. 14. 15. 14. 17. 17. 15 18.	125	3708
41–48	12. 13. 21. 13. 17. 15. 23. 17.	131	3848
49–56	15. 23. 16. 16. 17. 21. 17. 18.	143	3991
57–64	23. 14. 26. 21. 18. 12. 19. 19.	152	4082
65–72	21. 20. 21. 20. 15. 23. 15. 20.	155	4240
73–80	22. 20. 24 20. 20. 22. 22. 9.	159	4407
81–88	25. 16. 26. 31. 20 24. 26. 28.	196	4486
89–96	18. 13. 34. 16. 26. 21. 32. 25.	185	4628
97–104	15. 22. 30. 10. 55. 30.	162	4814
105–112	47. 49. 39. 38.	173	4935
113–120	46. 46. 44. 42.	178	5052
121–128	41. 53. 40. 45.	179	5171
129–136	46. 38. 46. 41.	171	5326
137–144	41. 45. 43. 35.	164	5436
145–152	40. 40. 41. 41.	162	5569
153–160	47. 40. 37. 36.	160	5669
161–168	36. 45. 40. 27.	148	5831
169–176	45. 41. 39. 39.	164	5915
177–184	29. 38. 41. 36.	144	6146
185–192	33. 33. 37. 37.	140	6242
193–200	36. 37. 34. 31.	138	6319
201–208	67. 62.	129	6475
209–216	77. 53.	130	6467
217–224	68. 65.	133	6677
225–232	66. 67.	133	6721
233–240	52. 59.	111	6770
241–248	62. 52.	114	7010
249–256	53. 70,	123	7112
257–264	37. 57.	94	7204
265–272	53. 41.	94	7274
273–280	39. 41.	95	7347
281–288	48. 43.	91	7481
289–296	37. 48.	85	7512
297–304	40. 39.	79	7610
305–312	42. 42.	84	7788
313–320	28. 34.	62	7765
321–328	40. 27.	67	7887
329–336	36 28.	64	8281
337–344	28. 31.	59	7985
345–352	20. 23.	43	8194
353–360	27. 21.	48	8281
361–368	25. 13.	38	8613

Diagram I.


SHOWING THE AVERAGE WEIGHTS OF THE INFANTS OF SERIES I (CONTROL SERIES, CHIEFLY BREAST-FED) & OF THE INFANTS OF SERIES II (FED ON BOILED COWS' MILK) IN PERIODS OF EIGHT DAYS.
Weight in Grammes
9000
8000
7000
6000
5000
4000
3000
INFANTS OF CONTROL SERIES
INFANTS FED ON BOILED COWS' MILK
Days of Age
8
8·16
16·24
32·40
72·80
112·120
152·160
192·200
232·240
272·280
312·320
352·360
1010. 27944./566. 1000.4.12.
Malby & Sons, Lith.

Analysis of the curves of Diagram I. (1)

Diagram I shows at once that a considerable divergence between the two curves starts in the early days of life, and continues well-marked up to about the 208th day, after which it disappears fairly rapidly.

The question suggested by these curves is, — Is the difference between the average weight of breast-fed and of babies of the same age fed upon boiled cows' milk due to the method of feeding?

Analysis of the curves of Diagram I. (2)

Diagram I. would seem to have answered this question affirmatively. Before, however, stating this definitely to be the case, it is advisable to consider whether some other factor may not be concerned, to which this difference can be attributed.

Such a factor might be the error due to the so-called "Error of Sampling." If this error is significant, then the curves may have a different interpretation to the apparently obvious one, and it therefore becomes essential to examine the importance of this factor, before proceeding to draw deductions from the curves as they stand in Diagram I.

*Analysis of the Data by Statistical Methods.**—In dealing with the error of sampling the important point will evidently be to ascertain how much the mean value obtained from the observations as shown on the curves is likely to differ from the mean of all babies in the same class, that is to say what is the probable error of the mean.

Suppose M_1 and M_2 are the means of the two sets of observations, then the accuracy of each must evidently depend upon

(*a*) The number of observations upon which it is based, and

(*b*) The divergence of these observations from their mean value.

In statistical work the expression $\cdot 67449 \frac{s}{\sqrt{N}}$ is taken to represent the probable error, where s = the square-root of the average of the squares of the distances of the observations from the mean, and is known as the "Standard Deviation," and where N = the number of observations. (Cp. Yule. Introduction to the Theory of Statistics. Chaps. VII. and XVII.)

The measures of reliability or the "probable errors" for the two means will be $\cdot 67449 \frac{s_1}{\sqrt{N_1}}$ and $\cdot 67449 \frac{s_2}{\sqrt{N_2}}$ respectively. These expressions may be called E_1 and E_2.

Experience has shown that unless the difference between M_1 and M_2 is at least two or three times as great as $\sqrt{E_1^2 + E_2^2}$ then it is not safe to assert that the difference found is really significant; it might be due to an error of sampling.

This method is only strictly speaking applicable when the variables, *i.e.*, the observations are "normally" distributed (vide Yule, op. cit. Chap X.) but it may fairly be used as a sufficiently accurate test for material such as the present.

This test of the error of sampling has been applied over three periods of eight days, in each of the series. The three periods selected were the three consecutive periods included from the 137th to the 160th day after birth. These periods were selected as being those where there were a large number of observations in both series, and where the numbers of each series were most nearly equal.

---

[a] For instruction in the statistical methods employed and for supervision of the results obtained I am deeply indebted to Dr. Major Greenwood, Junr., of the Lister Institute, and have much pleasure in thanking him for his most valuable help.

The unit of grouping taken was 200 grammes, and the results obtained are given in the accompanying table.

Days of age.	Mean weight (in grammes).		Difference $M_1 \sim M_2$ (in grammes).	Probable error.		Value of $\sqrt{E_1^2+E_2^2}$.	Ratio of $\frac{M_1 \sim M_2}{\sqrt{E_1^2+E_2^2}}$.
	Series I.	Series II.		Series I.	Series II.		
137–144	5,929	5,436	495	44·98	44·59	63·6	7·8
145–152	6,033	5,569	455	43·84	41·20	60·1	7·6
153–160	6,237	5,669	548	48·25	44·00	65·1	8 4

The mean of these observations bears therefore such a ratio to the value of $\sqrt{E_1^2+E_2^2}$ as to show clearly that the difference between the mean values of the two series can hardly be due to an error of sampling.

It appears that there is a difference between the values obtained for the series of babies fed upon the breast and for those fed upon boiled cows' milk, and that this difference can hardly be attributed to errors of sampling.

It does not, however, necessarily follow that the difference of food has been the causative factor, and it becomes necessary to ask whether there can be any other factor at work which is producing the difference found.

- Health of the children...

- Social class of the children...

The following **inferences** may be drawn as to the divergence of the two curves in Diagram I up to the 208th day :-

(1) There is a significant difference between the average weight of infants fed upon the breast and upon boiled cows’ milk, in favour of the former; and

(2) An important factor in this result is the method of feeding.

The Curves of Diagram I. maybe divided into **three parts**, namely:-

(i) The first part where the curves cross and then diverge; the curve of the boiled milk series, which starts above the curve of the breast-fed series; falling rapidly below this latter curve.

(2) The second part of the curves where the two curves run approximately parallel from about the 24th to the 200th day of life, and

(3) The last part of the Curves where the divergence is obliterated, the subsequent tendency being for the curve of the boiled cows' milk series to show a value a little above that of the breast-fed series.

From the preceding statistical analysis it appears that the divergence of the middle part of the curve is to be attributed essentially to the difference in the method of feeding of the two series.

Further Analysis of the **First Part of the Curves** of Diagram I.

At no part of the curves is the tendency to diverge so markedly shown as in the first part of the curves, during a period extending over the **first three of the eight-day periods of life.**

The average weight of the breast-fed babies shows a rise from the first, while that of the babies fed upon boiled cows' milk falls throughout the two first eight-day periods, and shows no rise above the first eight day period until the 33-40th days of life.

It is a matter of common knowledge that every baby loses weight during the first few days of life, and a drop in the average weight of the breast-fed babies in the second eight-day period was almost to be expected. This possible fall in the curve is concealed to some extent by the grouping of the weights into periods of eight days, the first period including the period of fall in weight. In many of these cases the observations would commence at a time when the loss of weight after birth had already taken place, and the child was again beginning to increase in weight.

The absence of fall in the curve of the breast-fed babies can therefore be explained.

When a comparison is made between the two curves, it appears that while one curve rises the other falls, and evidently there is either some fundamental factor or factors at work producing this difference, or some source of error has crept into one or both of the curves.

It was considered desirable first to eliminate any possible source of error. The same source of error as was sought for in the middle part of the curves may evidently be at work in this part of the curves, namely, the error of sampling, and this was therefore investigated.

Statistical Examination of the Average Weights obtained in the first four periods of Eight Days.

The same method and notation as were used in dealing with the middle part of the curves was applied, viz. :—

N = Number of observations.

s = Standard deviation.

E = Probable error, and is represented by the expression

$$\cdot 67449 \frac{s}{\sqrt{N}}$$

M_1 and M_2 are the means of the two series, their difference being "D."

By this method the following values were obtained and are tabulated below :—

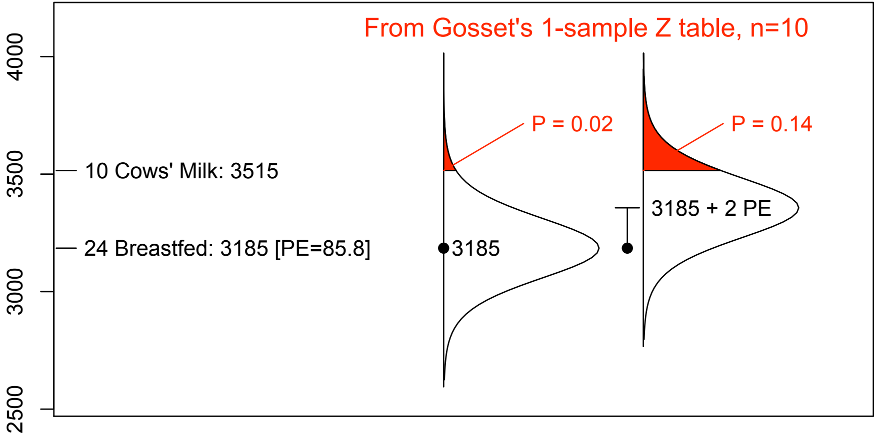
—	Days of Age.	Mean (in grammes).	Standard Deviation.	Probable Error.	$\sqrt{E_1^2+E_2^2}$	$\frac{D}{\sqrt{E_1^2+E_2^2}}$
Series I ...	1–8	3,185	622	85·8		
		D=330			122·0	2·7
Series II ...	1–8	3,515	410	87·4		
Series I ...	9–16	3,312	544	37·3		
		D=222			60·5	3·7
Series II ...	9–16	3,090	452	48·3		
Series I ...	17–24	3,512	632	35·4		
		D=145			49·5	2·92
Series II ...	17–24	3,367	460	35·7		
Series I ...	25–32	3,745	652	34·1		
		D=272			47·5	5·7
Series II ...	25–32	3,473	522	33·1		

The average weight of the babies fed upon boiled cows' milk is higher for the first eight-day period than that of the breast-fed babies. The former value is based upon **10** observations, and the latter upon **24**; it becomes a question whether any importance can be attributed to this difference in average weight or **whether it may not be due to an error introduced by the extremely small number of observations available for the boiled cows' milk series.**

Ten observations are not sufficient for the formula given in the above table of results to be justifiably employed, since the reliability of the method is exaggerated when the number of observations is very small.

The method introduced by "Student" (ref) is applicable for small number of observations. It is based upon the probability of the occurrence of the mean value obtained by the ordinary method among the average population.

Taking 3185 ± 85·8 (the "probable error" of 3185 is 85·8) grammes as the mean weight of babies in the average population it appears that the chance of 10 observations from such a population having a mean of 3515 grammes with a standard deviation of 410 is 1 in 50. Suppose, however, that the mean weight of the average baby in the population were 3357 grammes, it is then found that the probability that a population with a mean weight of the babies of this age of 3357 grammes (3185 + twice the probable error, *i.e.*, 172) should give in 10 observations a mean of 3515 is 1 in 7. It may be remarked that so far as the evidence goes, there is about 1 chance in 10 that the mean weight of the controls is not less than 3359.


From Gosset's 1-sample Z table, n=10
4000
3500
3000
2500
10 Cows' Milk: 3515
24 Breastfed: 3185 [PE=85.8]
P = 0.02
P = 0.14
3185
3185 + 2 PE

It seems therefore that the difference between the weights of the two series for the first eight-day period, might he considered as due to an error of sampling brought about by the extremely small number of observations available for the series of babies fed upon boiled cows' milk. It may be taken that the babies of both series whose weights were observed during this period of life can be considered as average samples of the population, the influence of other factors, if present, which would tend to cause a divergence of the two series, being inappreciable compared wlth that caused by the error of sampling.

The figures of the later periods, are based upon sufficiently large number of observations for the ordinary method to be reliable.

The tabulated results show that the ratio of the difference of the means to the measure of the sum or difference of the probable errors ($\sqrt{E_1^2 + E_2^2}$) is in all cases greater than 2, and hence the difference in weight of the two series, may fairly be attributed to some factor other than the error of sampling.

A source of error might arise in respect of the distribution of the variables.

In applying the usual method, it is assnmed that these are "normally" distributed; inspection of the distribution of the individual weights suggests that this condition is not accurately fulfilled, and the process is not then strictly reliable (84).

The figures, however, approximate sufficiently to the normal type for it to be unlikely that an appreciable error is introduced in basing the results obtained upon the application of this formula.

Some factor other than the error of sampling must therefore be sought for.

The possible influence of the social conditions has already been dealt with fully in a previous section of this report (see pp. 41, 42) in connection with the middle part of the curves, and it has been shown that in this part of the population, which is to a great extent a selected population, this is a negligible factor. It need not therefore be raised again.

The main factor for consideration will evidently be that of the feeding and it seems not unreasonable to suppose that the loss in weight which occurs in all children is on the average more prolonged in babies fed upon boiled cows' milk, than in babies fed upon the breast.