13

A Biometric Analysis of Some Insemination-Labour and Menstrual-Labour Curves in Certain Mammalia.

BY

D. BERRY HART, M.D., F.R.C.P.E.,

Lecturer on Midwifery and Diseases of Women, Surgeons' Hall, Edinburgh.

Reprinted from the EDINBURGH MEDICAL JOURNAL, *May* 1914.

A BIOMETRIC ANALYSIS OF SOME INSEMINATION-LABOUR AND MENSTRUAL-LABOUR CURVES IN CERTAIN MAMMALIA.

By D. BERRY HART, M.D., F.R.C.P.E.,
Lecturer on Midwifery and Diseases of Women, Surgeons' Hall, Edinburgh.

In a previous communication * I showed that from Tessier's statistics as to the insemination-labour duration in ewes, from Earl Spencer's in cattle, and from von Winckel's and Reid's

* "On the Duration of the Interval between Insemination and Parturition in Certain Mammals as Studied by Biometric Curves, with Special Reference to the Calculation of the Onset of Labour in Human Pregnancy," *Edin. Obstet. Trans.*, xxxviii. 107 ; *Edin. Med. Journ.*, 1913, xi. 291.

menstrual-labour durations in women, a frequency polygon could in each case be constructed.

In Tessier's ewes the curve from 912 labours was a lofty symmetrical one, slightly skew on the right side. The dates were given in 24-hour intervals, and were continuous over 11 days.

In the others (Spencer's, von Winckel's, and Reid's) the curve was interrupted by peaks when grouped in 24 - hour intervals, but was still of a frequency nature. When Spencer's were taken in 48-hour groupings, the curve smoothed out (Fig. 1). In von Winckel's and Reid's a 96-hour grouping gave the same result, *i.e.*, a fairly smooth and symmetrical frequency curve was obtained (Fig. 3 gives Reid's).

The irregularities in the Spencer, von Winckel, and Reid statistics seemed to me to be due to the close 24-hour grouping and to the consequent separation of births happening near midnight from those quite close to them shortly after midnight. By grouping them in 48- or 96-hour periods this mal-allotment may have been avoided to a great extent.

The question still arose, however, as to the propriety of smoothing out these curves by a means of which the method was not quite clear, and therefore one had to face this point—Given 24-hour groupings with the peaked irregularities, what could be made of them by biometric treatment—would Gauss's method of least squares be of use?

I therefore consulted my friends Mr. J. D. Hamilton Dickson, Fellow and Tutor of Peterhouse, Cambridge, and Dr. A. Daniell, whose *Text-Book of Physics* is well known, as to this, and was fortunate enough to enlist their sympathy and active help in the question. Mr. Dickson worked out Spencer's and Reid's data, and Dr. Daniell treated the question from the curve point of view, so that from their most valuable help I am now enabled to bring the matter up again with benefit to the elucidation of the practical obstetrical questions—what is the significance and accuracy of a labour date calculated from a single insemination in cows or from the last day of the last menstrual period in the human female? What is the most probable insemination-labour duration in cattle and the most probable menstrual-labour duration in the human female? These data may be regarded as the varying measurements between insemination and labour or between the last menstruation and labour, and can therefore be treated by Gauss's method of least squares, a form of mathematical inquiry used in astronomical observation data, and, in general, by statistical observers, to settle, in any series of observations of the same or similar objects, the

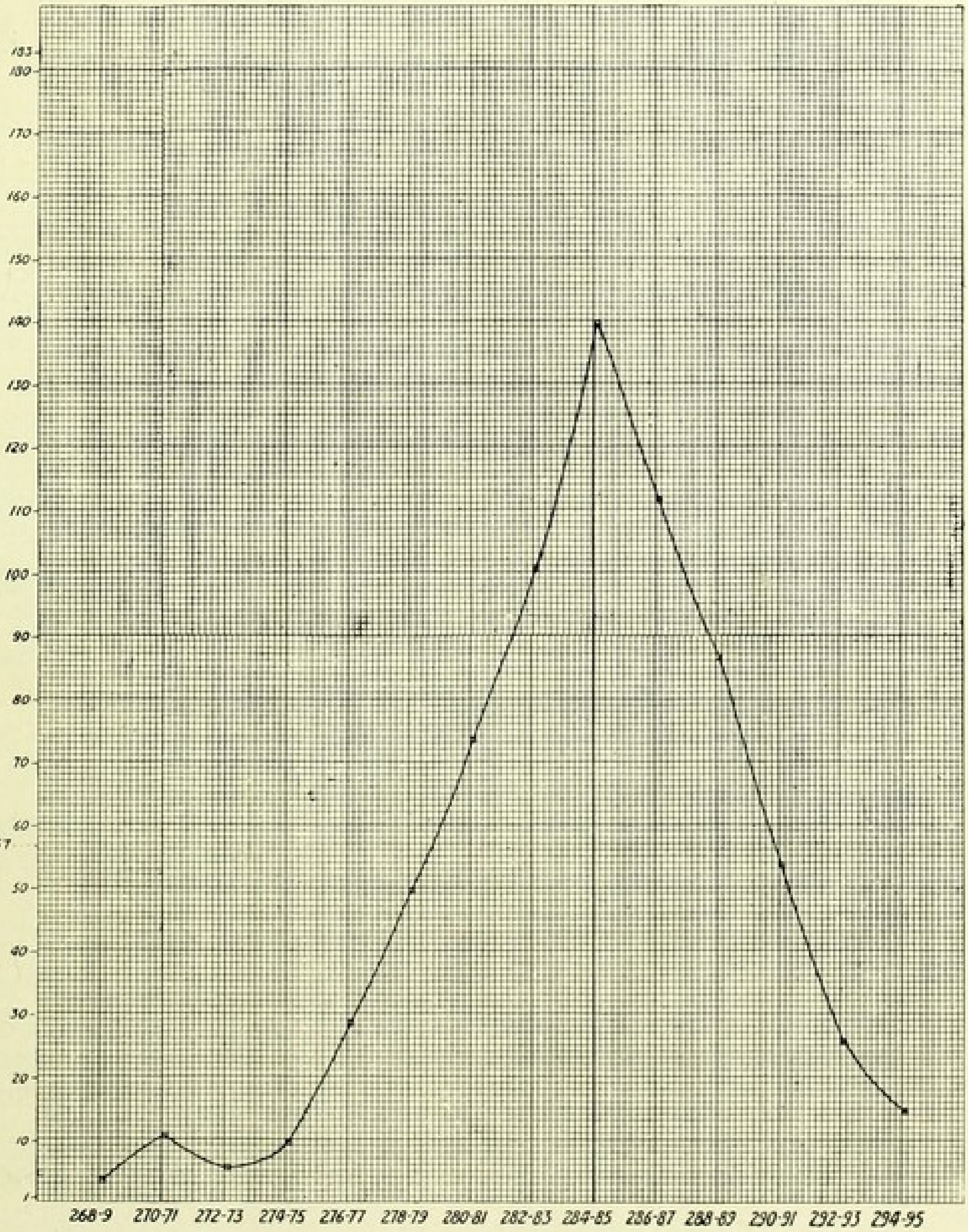


FIG. 1.—Frequency polygon based on Earl Spencer's birth statistics of 764 cattle: in 48-hour groupings.

most probable result, *i.e.*, the arithmetical mean of the observations, its probable error as well as the probable error of a single observation. I do not propose to explain this method in all its bearings, as that would involve points unnecessary here, and is best studied in such works as Merriman's *Method of Least Squares*, Jevon's *Principles of Science*, and Herschell's article in the *Edinburgh Review* for 1850.

Jevons gives the following summary of Gauss's method, and I have added a few explanatory remarks :—

1. Draw the mean of all the observed results (by dividing the sum of the measurements by their number).

2. Find the excess or defect, that is, the error of each result from the mean (call this v).

3. Square each of these results (and call each v^2 and the sum of them Σ).

4. Add together all these squares of the errors, which, of course, are all positive.

5. Divide by one less than the number of observations. This gives the *square of the mean error.*

6. Take the square root of the last result; it is the *mean error of a single observation.*

7. Divide now by the square root of the number of observations, and we get the *mean error of the mean result.*

8. Lastly, multiply by the natural constant 0·6745 (or approximately by 0·674, or even by $\frac{2}{3}$), and we arrive at the *probable error of the mean result.*

I now go on to give Mr. Hamilton Dickson's calculations in Spencer's cases of 764 cattle, and in Reid's 500 human menstrual-labour durations in 24-hour intervals. Mr. Dickson also plotted out these results on graph paper, and drew the frequency polygon threading Reid's (Fig. 4).

If we deal first with the numerical results of Spencer's cattle statistics, it is to be noted in Mr. Dickson's analysis below that in the first column (left hand) the separate number of days is given in each case; in the second the number of cattle which had the same duration; in the third the product of the number of cattle and the number of days these cattle gestated, and at the foot the sum,

216,451, or the Σ in such an expression as $\sqrt{\frac{\Sigma(v^2)}{n-1}}$;

in the fourth the difference between the mean and each result (v); in the fifth the squares of the differences (v^2); and in the sixth these multiplied by the number of cases. The sum of the squares of the differences is 60,629. I now go on to quote Mr. Hamilton Dickson's text, and his interesting comment.

Spencer's Observations Calculated from Insemination.

			v	v^2	
220	1	220	63	3,969	3,969
...					
226	1	226	57	3,249	3,269 *
...					
233	1	233	50	2,500	2,500
234	1	234	49	2,401	2,401
235	1	235	48	2,304	2,304
...					
239	1	239	44	1,936	1,936
...					
242	1	242	41	1,681	1,681
...					
245	2	490	38	1,444	2,888
246	2	492	37	1,369	2,738
...					
248	1	248	35	1,225	1,225
...					
250	1	250	33	1,089	1,089
...					
252	2	504	31	961	1,922
253	1	253	30	900	900
254	1	254	29	841	841
255	2	510	28	784	1,568
...					
257	2	514	26	676	1,352
258	3	774	25	625	1,875
259	1	259	24	576	576
...					
262	1	262	21	441	441
263	2	526	20	400	800
...					
266	1	266	17	289	289
...					
268	2	536	15	225	450
269	2	538	14	196	382
270	5	1,350	13	169	845
271	6	1,626	12	144	864
272	3	816	11	121	363
273	3	819	10	100	300
274	5	1,370	9	81	405
275	5	1,375	8	64	320
276	15	4,140	7	49	735
277	14	3,878	6	36	504
278	18	5,004	5	25	450
279	32	8,928	4	16	512
280	35	9,800	3	9	315
281	39	10,959	2	4	156
282	47	13,254	1	1	47
283	54	15,282	0	0	0
284	66	18,744	1	1	66
285	74	21,090	2	4	296
286	60	17,160	3	9	540
287	52	14,924	4	16	832
288	42	12,096	5	25	1,050
289	45	13,005	6	36	1,620
290	23	6,670	7	49	1,127
291	31	9,021	8	64	1,984
292	16	4,672	9	81	1,296
293	10	2,930	10	100	1,000
294	8	2,352	11	121	968
295	7	2,065	12	144	1,008
296	6	1,776	13	169	1,014
297	2	594	14	196	392
...					
299	1	299	16	256	256
...					
304	1	304	21	441	441
305	1	305	22	484	484
306	3	918	23	529	1,587
307	1	307	24	576	576
...					
313	1	313	30	900	00
	764	216,451			60,649
		Mean = 283·3			

* This is a *lapsus pennæ* for 3249, and it has been left, as it involves only $\frac{1}{3000}$ error.

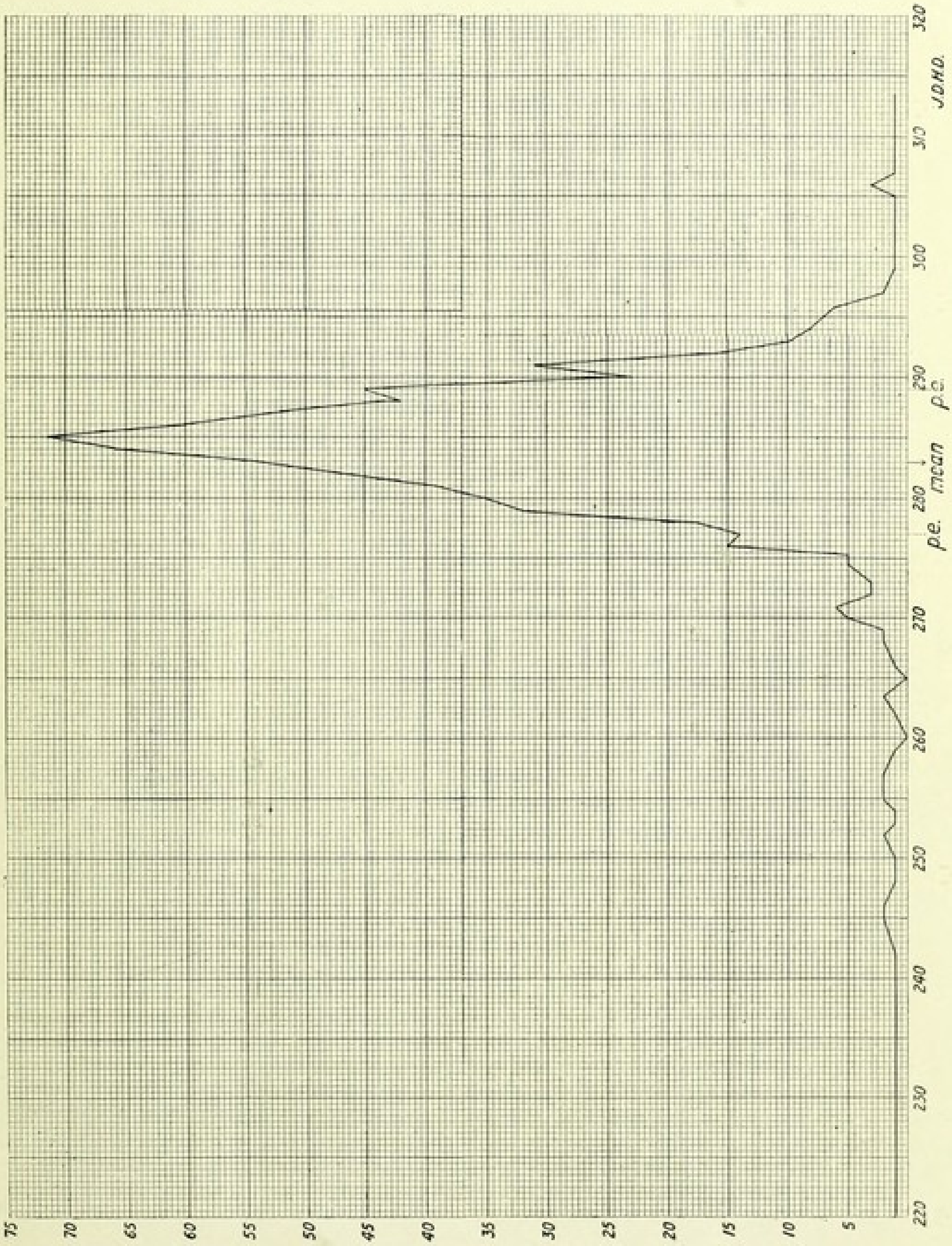


FIG. 2.—Spencer's cattle statistics in 24-hourly observations.

The probable error of *one* observation is (r),

$$0{\cdot}6745\sqrt{\frac{\Sigma(v^2)}{n-1}}$$

where n is the number of observations, and v is as defined above. This arithmetical mean is 283·3; but to simplify the calculation (we shall see that it does not affect the result later) it was taken as 283.

Hence $$\Sigma(v^2) = 60{,}649.$$

$$\frac{\Sigma(v^2)}{763} = 79{\cdot}48.$$

$$\sqrt{\frac{\Sigma(v^2)}{763}} = 8{\cdot}915.$$

and then $$r = 0{\cdot}6745 \times 8{\cdot}915 = 6{\cdot}012.$$

This quantity r is mainly useful for drawing the probability curve; but as your purpose is the period of gestation, and the accuracy to be obtained from these observations, we are more concerned with the probable error of the *mean* of the observation, viz. 283·3; this *p. e.* is (r_0), given by

$$r_0 = 0{\cdot}6745\sqrt{\frac{\Sigma(v^2)}{n(n-1)}}$$
$$= 0{\cdot}6745 \times \sqrt{0{\cdot}1041}$$
$$= 0{\cdot}6745 \times 0{\cdot}323 = 0{\cdot}218.$$

The result is that, taking all the observations, their mean is 283·3 (or 283) days, with a probable error amounting to 0·2 of a day, and therefore we may say 283 days net.

On the curve it *looks* as if 285 is nearer it; but there is no good reason for rejecting the 25 observations from the 220th to the 259th day inclusive. *If*, however, we reject them, then the mean day is 285·8—say the 286th—which errs again by being (from the curve) apparently a little too high.

For my own part, I do not value 286 as so good a result as 283; but the physical reasons may perhaps justify the preference for 286. I have not calculated the *p. e.* for 286, but, in any case, it can only be a small part of a day.

The smalln ss of r (only ± 6 days, on 283) indicates that the observations give a very close approximation to the truth; it means that half of the births fell between the 276th and 289th days. The curve would be a very *thin*, *stand-up* one, but I have not calculated it.

REID'S DURATIONS IN 500 LABOURS CALCULATED FROM THE LAST DAY OF MENSTRUATION.

			v.	*v*².	
252	4	1,008	26	676	2,704
253	1	253	25	625	625
254	3	762	24	576	1,728
255	1	255	23	529	529
256	2	512	22	484	968
257	4	1,028	21	441	1,764
258	4	1,032	20	400	1,600
259	4	1,036	19	361	1,444
260	6	1,560	18	324	1,944
261	5	1,305	17	289	1,445
262	3	786	16	256	768
263	9	2,367	15	225	2,025
264	10	2,640	14	196	1,960
265	5	1,325	13	169	845
266	10	2,660	12	144	1,440
267	9	2,403	11	121	1,089
268	13	3,484	10	100	1,300
269	5	1,345	9	81	405
270	13	3,510	8	64	832
271	12	3,252	7	49	588
272	13	3,536	6	36	468
273	16	4,368	5	25	400
274	21	5,754	4	16	336
275	20	5,500	3	9	180
276	16	4,416	2	4	64
277	16	4,432	1	1	16
278	22	6,116	...	...	...
279	21	5,859	1	1	21
280	15	4,200	2	4	60
281	18	5,058	3	9	162
282	25	7,050	4	16	400
283	14	3,682	5	25	350
284	15	4,260	6	36	540
285	14	3,990	7	49	686
286	15	4,290	8	64	960
287	11	3,157	9	81	891
288	17	4,896	10	100	1,700
289	8	2,312	11	121	965
290	9	2,610	12	144	1,296
291	14	4,074	13	169	2,366
292	6	1,752	14	196	1,176
293	3	879	15	225	675
294	6	1,764	16	256	1,536
295	2	590	17	289	578
296	5	1,480	18	324	1,620
297	8	2,376	19	361	2,888
298	6	1,788	20	400	2,400
299	1	299	21	441	441
300	2	600	22	484	968
301	4	1,204	23	529	2,116
302	1	302	24	576	576
303	1	303	25	625	625
304	1	304	26	676	676
305	2	610	27	729	1,458
306	0	0	28	784	0
307	1	307	29	841	841
308	2	616	30	900	1,800
309	0	0	31	961	0
310	1	310	32	1,024	1,024
311	1	311	33	1,089	1,089
..					
314	1	314	36	1,296	1,296
315	2	630	37	1,369	2,738
316	1	316	38	1,444	1,444
	500	139,138			65,829
	Mean =	278·3			

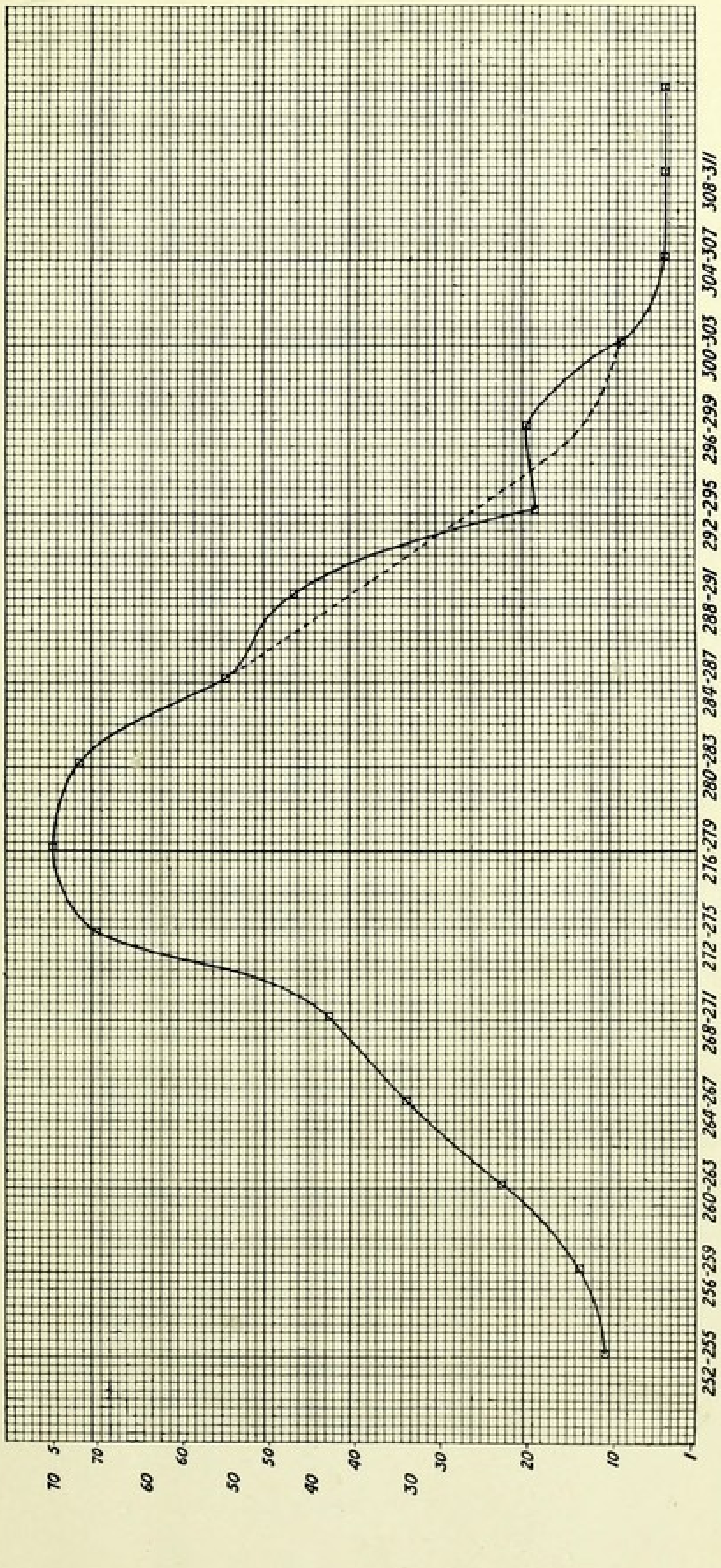


FIG. 3.—Frequency curve based on Reid's statistics of 500 human female births in 96-hour groupings, and calculated from last day of last menstruation.

The probable error (r) of one observation is

$$r = 0{\cdot}6745 \sqrt{\frac{\Sigma(v^2)}{n-1}}$$

Here $\Sigma(v^2) = 65{,}829$, with the *mean* 278 (more accurately 278·276), $n = 500$.

$$\therefore \quad \sqrt{\frac{\Sigma(v^2)}{n-1}} = \sqrt{\frac{65{,}829}{499}} = \sqrt{131{\cdot}922} = 11{\cdot}49$$

hence $$r = 0.6745 \times 11{\cdot}49 = 7{\cdot}748.$$

Also probable error (r_0) of the mean is

$$r_0 = 0{\cdot}6745 \sqrt{\frac{\Sigma(v^2)}{n(n-1)}}$$

$$= 0{\cdot}6745 \times 0{\cdot}514 = 0{\cdot}3466.$$

The result is—the most probable menstrual-labour duration of gestation is 278 days (or, if you like, 278·276 days), with a probable range of 0·35 (0·3466) of a day, more or less. In other words, the most probable duration lies between

278·28 – 0·35 and 278·28 + 0·35.
277·93 days and 278·63 days.

The following is Mr. Dickson's comment on the results:—

Peterhouse, Cambridge,
22nd November 1913.

Dear Doctor Berry Hart,—I have calculated *Spencer* and *Reid* in each case for the *whole* number of observations given.

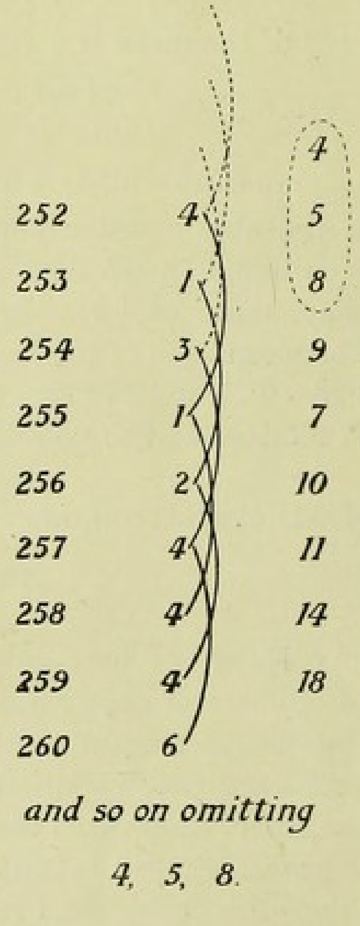
Spencer.—The most probable length of the insemination - labour period is 283·3 days, with a probable variation, up or down, of 0·22 of a day. If we use only the observations from the 262nd to the 313th day inclusive, then the most probable result is 285·8 days, with a somewhat similar variation as in the other case, but I have not calculated it, as the calculation would have to be made almost *de novo*, and it is long. Besides, cows are not so interesting as humans, and for a like reason I have not calculated the probability curve (Fig. 2).

Reid.—Here the most probable period of gestation is 278·276—say 278—days, with a probable variation, up or down, of 0·347 of a day. I have plotted the probability curve * in the midst of the observations, and you see it fits in very comfortably.

Of course the humans' observations cannot be so exactly got as those of cows, so that the cows' curve is much sharper than the humans'; that is to say, we can determine more accurately for cows than for humans—a result of which we have the measure in the two, probable errors, 0·22 and 0·347, almost in the ratio of 2 : 3, *i.e.*, cows are $1\frac{1}{2}$ times more accurate than humans.

* Calculated from these observations.

With regard to "smoothing" observations, I am always against it, my view being that the observations treated (as here) by least squares will look after themselves better than any plan we can adopt. But if we want to save calculation and (as in the present case) reduce * the amount of numbers to be employed (say) by grouping in 48 hours or in 4 days, then the plan is (for 4 days) to take *every* adjacent 4 days possible. Thus Reid, p. 406, I should treat as below.



The general theory of probability allows for no exceptional cases, *i.e.*, it expects to involve every case, however bad *we* may think it. But if we know of any reason attached to a particular case which makes it obviously unreliable, then, of course, cut it out. At the same time Chauvenet (an American mathematician) has devised a criterion by which a test can be applied for cutting out doubtful observations. However, it practically amounts almost to a recalculation for each observation cut out, and is too laborious to be used in actual observations.—Yours sincerely,

J. D. HAMILTON DICKSON.

* 1° plot all these 4-day points, then 2° use judgment in cutting out very astray ones, and calculate with the rest. But, after all, I prefer taking the actual observations themselves.

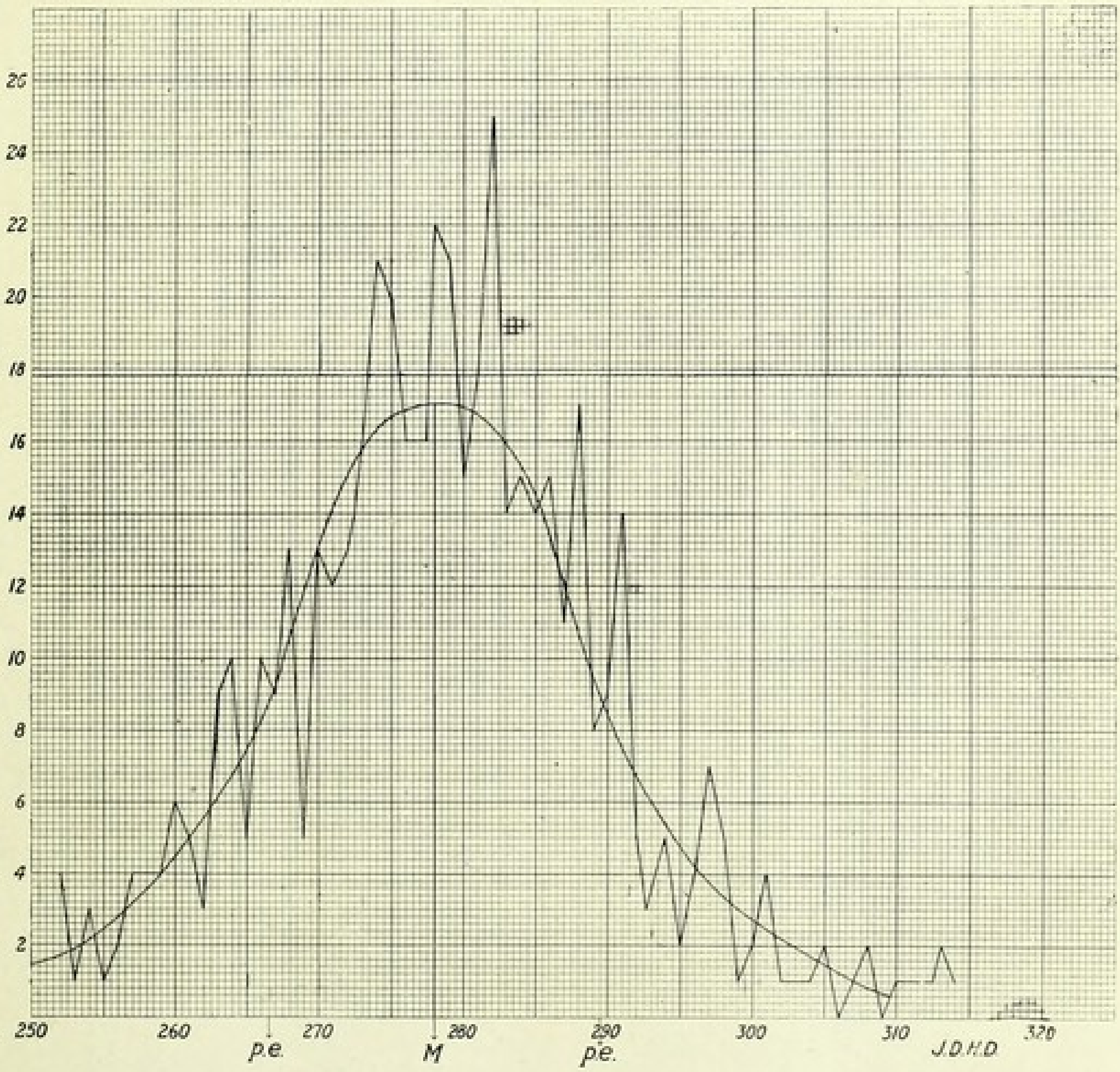


FIG. 4.—Reid's statistics of 500 cases in 24-hour dates with symmetrical frequency-curve threaded through them.

In the graph of Reid's statistics Mr. Dickson has threaded a symmetrical frequency curve, and this shows clearly the theoretical curve and the actual data plotted out in the peaked lines. Any point on the curve would satisfy the equation to the curve. Gauss's method of least squares gives us the power of settling the probable error in about half the cases, those grouped round M and between *p.-e.* and *p. e.*

Spencer's 24-hour graph is seen at Fig. 2 and the smoothed-out one at Fig. 1.

I may now discuss some points relative to the value of the results given by the smoothed-out curve and the 24-hour unsmoothed results when treated by the method of least squares. On this Dr. Daniell has given me valuable notes. In some respects they are too technical for actual reproduction, but I have greatly benefited by his discussion. Reid's smoothed-out data (Fig. 3) show that they form a frequency polygon, and that the most births occur between the 271st and 291st day. Mr. Dickson's results are for the same data, and give a mean of 278 days, with a probable error of 7 days on each side. The most probable duration of the menstrual-labour period is 277·93 and 278·63 days—practically 278 days.

It will be noticed that in the smoothed-out frequency polygon of Reid's data there is a humped irregularity in the 288-291-day group, and also a balancing one in the 292-295 day group. This probably means some disturbance of the average position of the ovum at a late period after insemination. Thus Venn remarks in his *Logic of Chance* on height data in English, French, and Belgian adults, according to Quetelet's statistics, that the mean in Englishmen is 5 ft. 9 ins., in Belgians about 5 ft. 7 ins., and in Frenchmen about 5 ft. 4 ins. The mixture of all these data would give a disturbing hump on the right side owing to their heterogeneity. The same would happen if the wafer on the target were displaced a certain distance after a few thousand shots had been fired, *i.e.*, the curve derived from the whole data would be skew, as Venn figures. Thus the skew nature of Reid's polygon might be explained by a late altered situation of the shed ovum, *e.g.*, its being in a deep fold of the mucosa.

It is evident, then, that one will find Gauss's method of greater value, but the plotting and smoothing out of curves has some advantages, and is a graphic appeal to the eye, of the generalisation, that the dates of labour after insemination or after a menstrual date conform to the law of probability within certain

Obstetrical Table.

January	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
October	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**1**	**2**	**3**	**4**	**5**	**November**
February	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28				
November	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**1**	**2**	**3**				**December**
March	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
December	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**1**	**2**	**3**	**January**
April	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30		
January	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**1**	**2**		**February**
May	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
February	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**1**	**2**	**3**	**4**	**5**	**March**
June	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30		
March	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**1**	**2**	**3**	**4**		**April**
July	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
April	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**1**	**2**	**3**	**4**	**5**	**May**
August	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
May	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**1**	**2**	**3**	**4**	**5**	**June**
September	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30		
June	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**1**	**2**	**3**	**4**	**5**		**July**
October	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
July	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**1**	**2**	**3**	**4**	**5**	**August**
November	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30		
August	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**1**	**2**	**3**	**4**		**September**
December	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
September	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**1**	**2**	**3**	**4**	**5**	**October**

This table is based on a 278-day menstrual-labour duration for the human female. The initial date taken is the last day of a menstrual period.

limits; that beyond these limits we get results of a less calculable nature, but all the same, results that can be understood on the principles just laid down.

I here append an obstetric calendar based on the results given above. It will be readily understood. Thus, if the last day of the last period was 6th January, 11th October, the date below the 6th, would be in the centre of the most probable fortnight for the occurrence of labour, that is, the most probable date would be seven days before or after the 278th day—in this case 11th October. It might, however, occur earlier or later than this fortnight, but this could not be foreseen, although understandable on the principles laid down.

LITERATURE.—Galton, Sir Francis, *Memories of my Life*, p. 302 (London: Methuen & Co.); also "Family Likeness in Eye Colour," *Proc. Roy. Soc. Lond.*, 1886, for Mr. Dickson's appendix to Galton's paper. Jevons, W. S., *The Principles of Science*, p. 208 (Macmillan & Co., 1877). Venn, *Logic of Chance*, third edition (Macmillan & Co., 1888). For additional literature see author's paper, *Edin. Med. Journ.*, October 1913.