Graphs Using Log Scales/ Linear and Exponential Growth

"logscale"

Figure 4.2, on page 71 in Pagano and Gauvreau, shows how age-standardized death rates from various causes have changed in the U.S. since 1950. The creators of this graph have not used a regular arithmetic scale on the vertical axis. If they had, equal <u>distances</u> along that axis would represent equal <u>differences</u> between rates. Instead, equal distances represent greater and greater differences as you move up the axis. They have used a logarithmic (or "log") scale.

What they have done, in effect, is to plot their data on semilog graph paper instead of ordinary (arithmetic) graph paper. Why? Just to annoy and confuse? ...or to show off how they can use fancy mathematics?

The log scale is really just as simple as the arithmetic scale. Its meaning is different, but it is just as simple and straightforward. And it has some advantages.

On an ordinary arithmetic scale the distance between any two values, like 5 and 15, represents their difference, 15 − 5 = 10. It is the same as the distance between 100 and 110, between 923 and 933, or between 0.14 and 10.14. This distance represents a difference of 10 units, no matter where it is placed along the axis, down near zero or up around 1000.

On a logarithmic scale the distance between 5 and 15 has a different meaning. It represents, not the <u>difference</u> between the two values, but their <u>ratio</u>, 15/5 = 3—the larger one is 3 times the smaller. If we start at some other value, say at 50, and go up the same distance, we will be at 3x50 = 150. If we start at 0.1 and go up this same distance, we will be at 0.3.

On the arithmetic scale a fixed distance, like 1/2 inch, represents a fixed <u>amount</u> .

On the log scale, on the other hand, a fixed distance represents a fixed <u>multiple</u>.

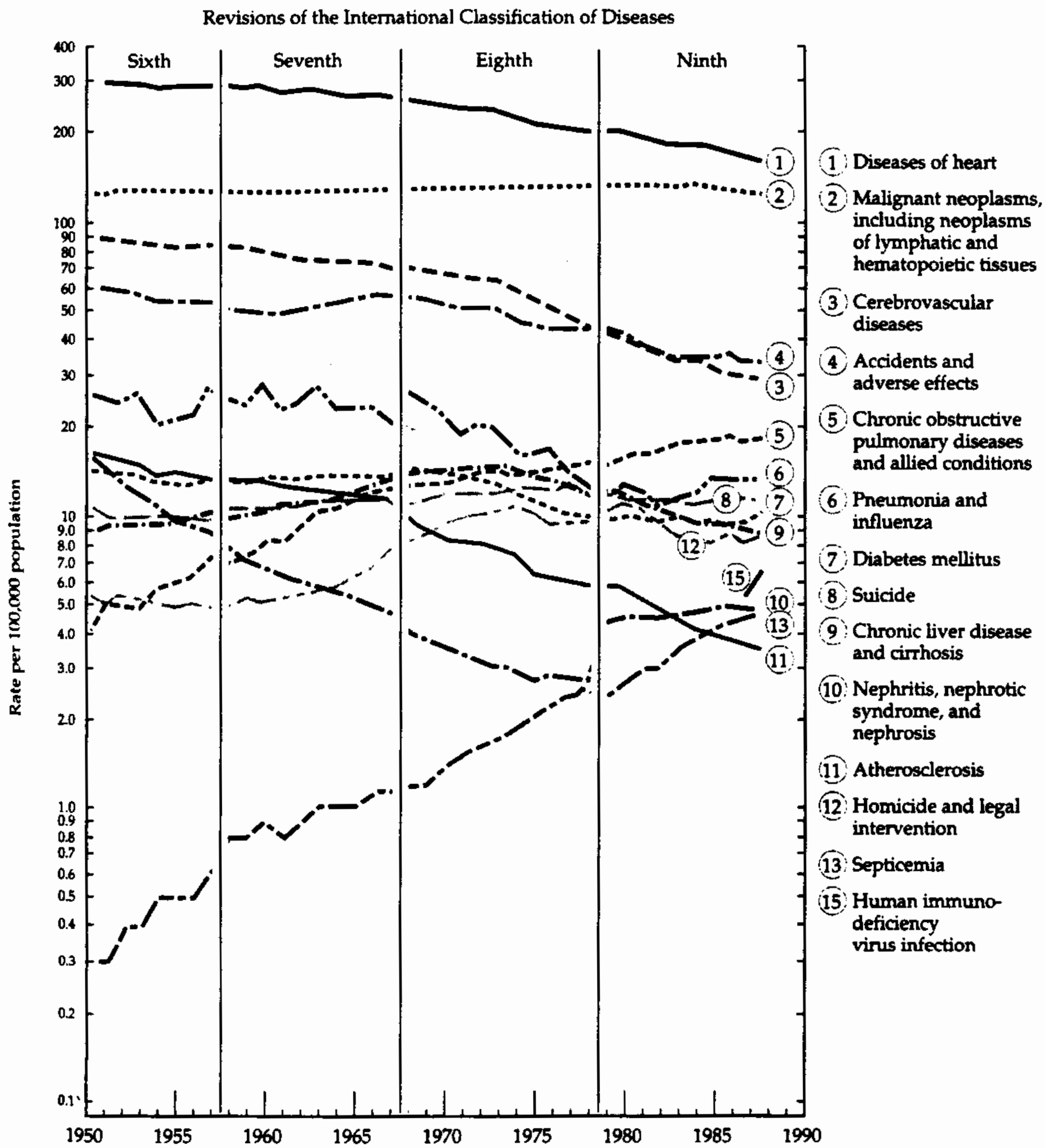


Figure 4.2. Age-adjusted death rates for 14 of the 15 leading causes of death, United States, 1950–1988

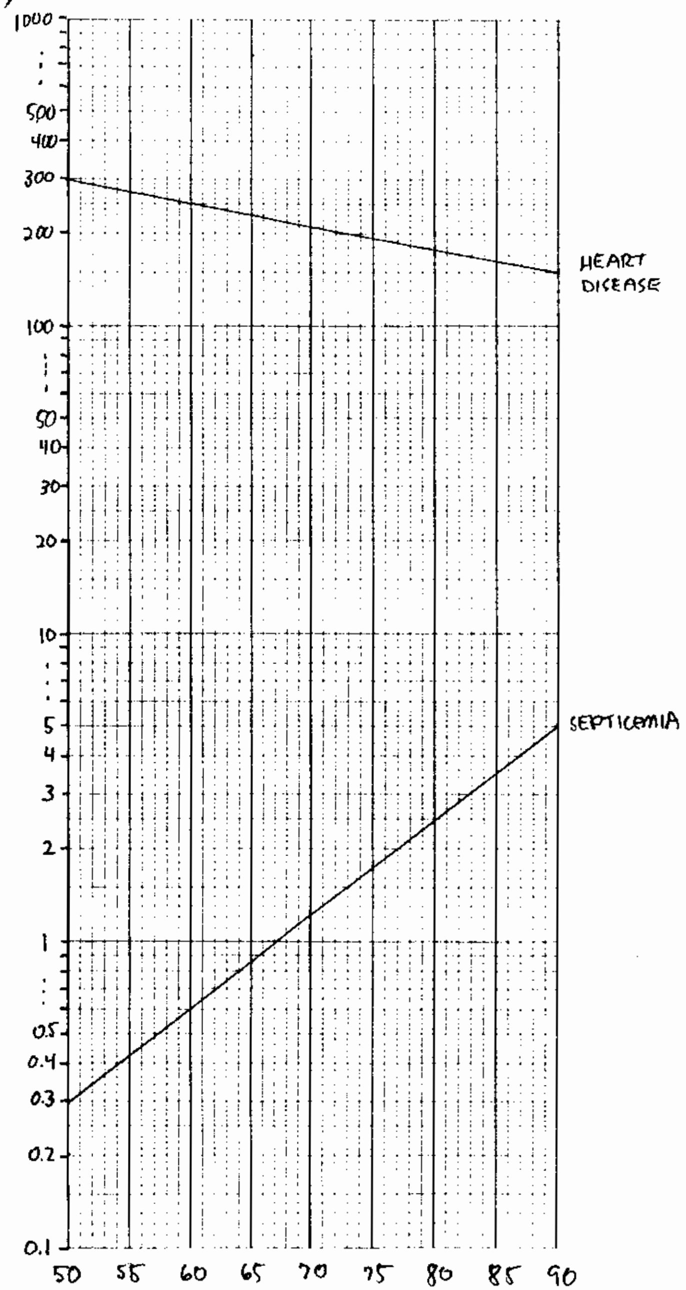
- If one value is 1.25 times another (one is 25% larger than the other), then the distance between those two points on a log scale will be the same, regardless of whether the two values are 1.00 and 1.25, 1000 and 1250, 0.00200 and 0.00250, etc.

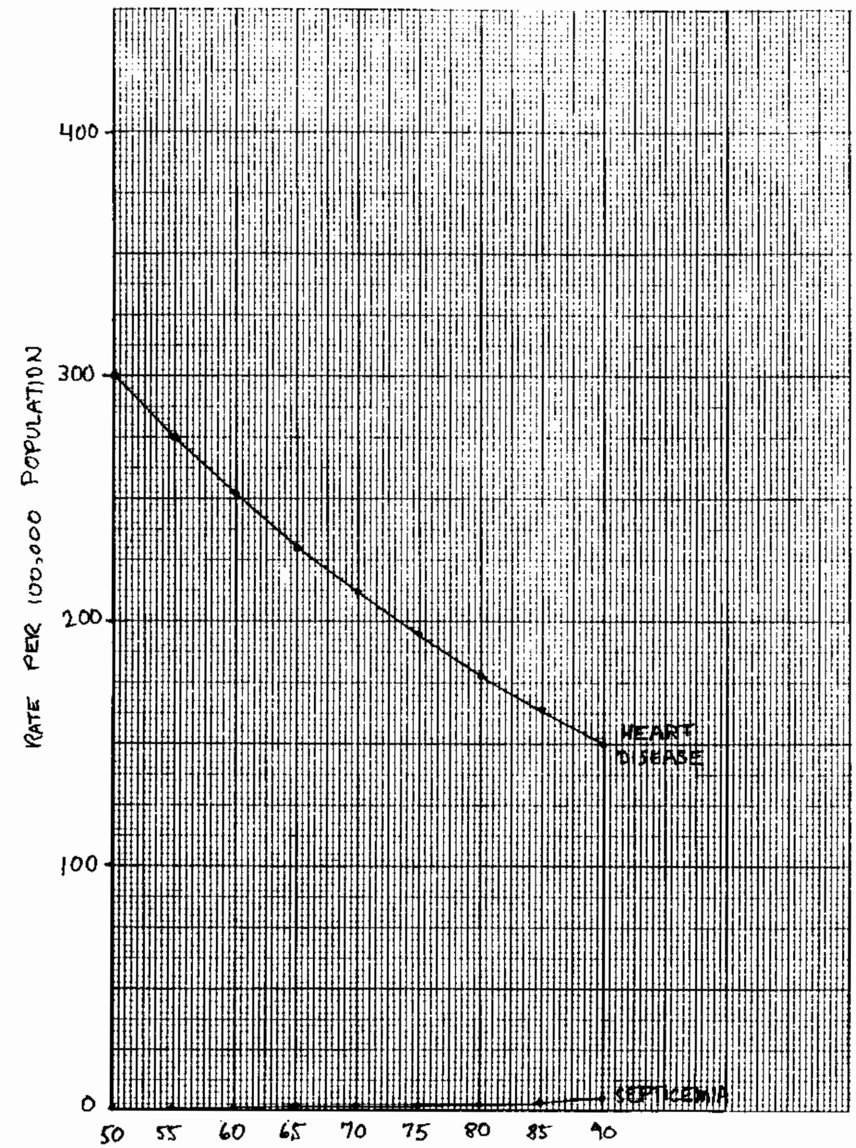
- If one value is 0.50 times another (one is half the other), then the distance between those two points on a log scale will be the same, regardless of whether the two values are 1.00 and 0.50, 1000 and 500, 0.00200 and 0.00100, etc.

How would the data shown in Figure 4.2 look if they were presented on an arithmetic scale? Let's look at a simplified version, using these artificial data:

year	Septicemia death rate	Heart disease death rate
1950	0.300	300
1955	0.426	275
1960	0.606	252
1965	0.862	231
1970	1.225	212
1975	1.741	195
1980	2.475	178
1985	3.518	164
1990	5.000	150

If we graph these rates on semilog paper, we get a picture very much like the one in the book. But if we graph them on ordinary arithmetic paper, we get a very different picture. (See the two attached graphs.)


1000
500
400
300
200
100
50
40
30
20
10
5
4
3
2
1
0.5
0.4
0.3
0.2
0.1
50 55 60 65 70 75 80 85 90
HEART DISEASE
SEPTICEMIA


RATE PER 100,000 POPULATION
400
300
200
100
0
50 55 60 65 70 75 80 85 90
HEART DISEASE
SEPTICEMIA

Focussing on the septicemia death rates, we see that

- Compared to heart disease, as a cause of death, septicemia is not terribly important—reducing the heart disease death rate by just 10% would save more lives than eliminating septicemia altogether. This is obvious from the arithmetic graph. To get this message from the log graph, the reader must study it carefully.

- The <u>trend</u> in the septicemia death rate <u>is</u> very important, however. It is shown clearly on the log scale, but is nearly imperceptible on the arithmetic one.

Why is this trend in a small rate important? Look what will happen if it continues. We now understand that on the log scale the distance between the 1950 and 1960 death rates represents their ratio. Reading these values from the graph, we see that the ratio is approximately $0.6/0.3 = 2$. Since on the log graph the septicemia death rate increases by this same distance every ten years, it means that it has been doubling every ten years. (The doubling time is actually slightly less than ten years, but let's approximate it by ten.)

It starts at 0.3, and

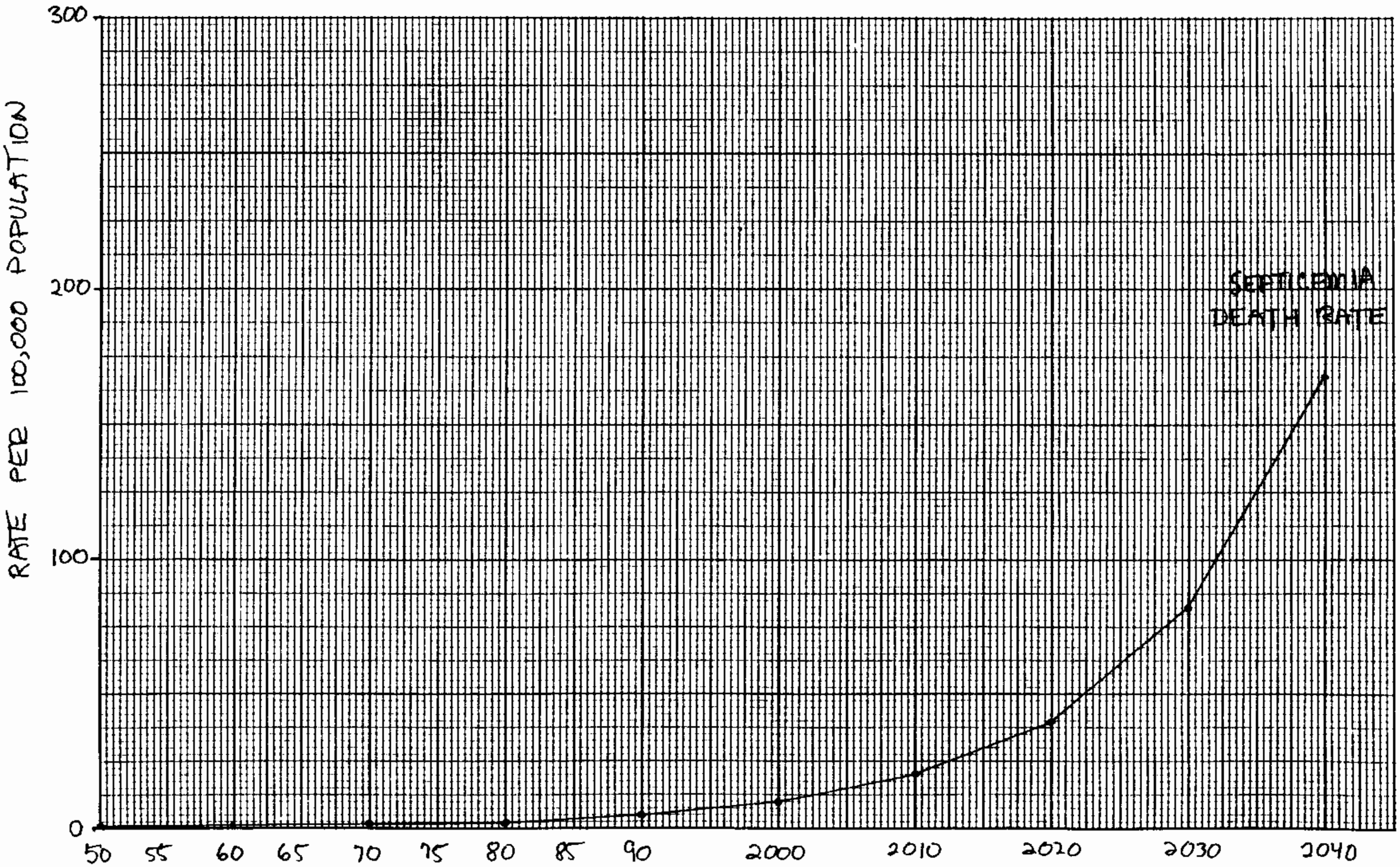
after 10 years it is $0.3x2 = 0.6$ $\quad 0.3x2 \quad = 0.3x2^1$
after 20 years it is $0.6x2 = 1.2$ $\quad 0.3x2x2 \quad = 0.3x2^2$
after 30 years it is $1.2x2 = 2.4$ $\quad 0.3x2x2x2 \quad = 0.3x2^3$
...etc.

After t 10-year time periods the rate is

$$0.3x(2x2x...x2) = 0.3 \text{ x } 2^t$$

The extended arithmetic graph shows that a rate that grows according to this formula grows <u>faster and faster</u> over time. A rate which lies on a sharply increasing curve of this form (on an arithmetic scale) is described as "increasing <u>exponentially</u>."

(The terminology comes from where the time variable, t, appears in the formula. When it is in the <u>exponent</u>, as in $0.3(2^t)$, the rate is growing exponentially over time.)


SEPTICEMIA
DEATH RATE
RATE PER 100,000 POPULATION
300
200
100
0
50
55
60
65
70
75
80
85
90
2000
2010
2020
2030
2040

Now compare this to a different pattern of growth. If the rate, which grew from 0.3 to 0.6 in the first decade, increases by the <u>same amount</u> (0.3) in each succeeding decade (instead of by the same multiple, 2), then

after 10 years it is 0.3+0.3 = 0.6	0.3+1(0.3)
after 20 years it is 0.6+0.3 = 0.9	0.3+2(0.3)
after 30 years it is 0.9+0.3 = 1.2	0.3+3(0.3)
...etc.	

After t 10-year time periods the rate would be

$$0.3 + t(0.3)$$

If we plot these values on arithmetic paper, we see that they lie on a <u>straight line</u>. The rate is increasing <u>linearly</u>. (It does not grow nearly as fast as it would if it were increasing exponentially.)

The reason epidemiologists get excited when when they see rates, like the septicemia death rates, which appear as increasing straight lines on a <u>log</u> plot (vs time) is because that means that even though the rates may be low, they are <u>growing exponentially</u>, and if this continues, the rates will soon become very large. Of course this does not apply only to rates. If a <u>population</u> grows by a fixed percentage every year, then it is growing exponentially. Let's see how this applies to an investment (a population of dollars):

Suppose you put $100 into a bank account that pays 5% interest each year. If every year you take the $5 interest and save it in a coffee can in your closet, then your fortune will increase linearly, by $5 per year. After one year you will have $105, and after two you will have $110.

But if you add the interest to the <u>bank account</u>, then you will have after

1 year: $100+(0.05)100 = 105 = (100)(1.05)$
2 years: $105+(0.05)105 = 110.25 = (100)(1.05)(1.05)$

It doesn't look like much of a difference, but after t years your fortune will be $(100)(1.05)^t$*—<u>reinvesting the interest makes your fortune grow exponentially.</u>*

How long will it take to double your \$100 under these two strategies?

- *coffee can strategy:* *20 years*
- *reinvest interest: We must find the value of t that satisfies the equation:* $(1.05)^t(100) = 2(100)$, *or*

$$\begin{aligned} (1.05)^t &= 2 \\ t \log(1.05) &= \log(2) \\ t &= \log(2)/\log(1.05) \\ &= 14.2 \text{ years.} \end{aligned}$$

How long will it take for your fortune to grow to \$1000 (i.e., to increase ten-fold) under these two strategies?

- *coffee can strategy:* $20 \times 9 = 180$ *years*
- *reinvest interest: Now we must find the value of t that satisfies the equation:* $(1.05)^t(100) = 10(100)$, *or*

$$\begin{aligned} (1.05)^t &= 10 \\ t \log(1.05) &= \log(10) \\ t &= \log(10)/\log(1.05) \\ &= 47.2 \text{ years.} \end{aligned}$$

When something grows at a fixed <u>rate</u>, i.e., by a fixed percentage, in each time period, like 5% per year (as the bank account was growing when the interest was reinvested), or 100% per decade (as the septicemia death rate was growing) it increases <u>exponentially</u>.

When it grows by a fixed <u>amount</u> in each time period, like the \$5 per year for the fortune under the coffee can strategy, it increases <u>linearly</u>, which is much slower.

One reason for displaying the death rates on a log scale is that a large percentage change shows up as a large change on the graph, even if the rate that is changing is small. The 16-fold increase in the septicemia death rate between 1950 and 1990, from 0.3 to 5 ($5/0.3 > 16$), is prominent on the log-scale graph. But because the size of the change ($5-0.3 = 4.7$) is small in comparison to the value, 300, of the heart disease death rate, it is represented by only a tiny rise on the arithmetic graph.

Furthermore, the fact that the septicemia rates appear as a straight line on the log scale tells us immediately that the rates are increasing exponentially.

(There. That wasn't so bad, was it?)

<u>P.S.</u> **(Do <u>not</u> read this if you're already confused, unless you enjoy getting <u>really</u> confused. You don't have to know the following material.)** If you want to see if your data form a straight line on semilog paper, but you happen to be marooned on a deserted island with only some ordinary graph paper and a calculator, you can get the result you want by plotting, not the rates themselves, but the <u>logarithms</u> of the rates.

Mathematically, it is very simple: if something is growing exponentially, it means that in its mathematical formula, the time variable appears as an exponent. We saw this in the case of the septicemia death rate, whose formula was rate = $0.3(2^t)$. If we take the logarithm of a rate like this, we get a formula for a straight line:

$$\log(\text{rate}) = \log(\,0.3(2^t)\,) = \log(0.3) + t \cdot \log(2).$$

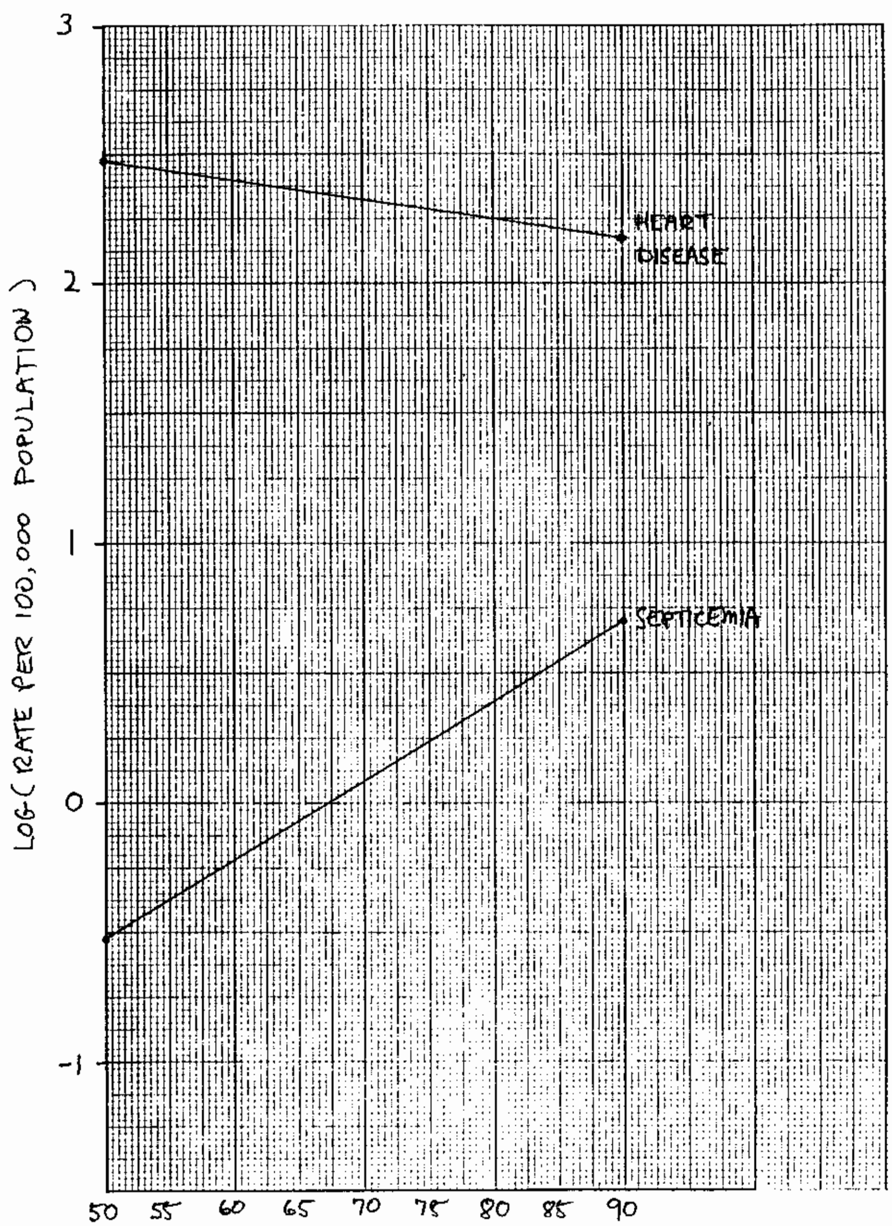
$$\{\text{rate exponential}\} \;\Leftrightarrow\; \{\log(\text{rate}) \text{ linear}\}$$

This is illustrated on the next graph. The data are

year	Sep. death rate	Heart dis. death rate	log(Sep.)	log(Heart)
1950	0.300	300	-0.523	2.477
1955	0.426	275	-0.370	2.439
1960	0.606	252	-0.217	2.402
1965	0.862	231	-0.065	2.364
1970	1.225	212	0.088	2.327
1975	1.741	195	0.241	2.289
1980	2.475	178	0.393	2.251
1985	3.518	164	0.546	2.214
1990	5.000	150	0.699	2.176

Plotting the rates on semilog paper is equivalent to plotting the logarithms of the rates (shown in the last two columns) on ordinary graph paper—you get the same picture either way.

In particular, if the rates form a straight line on semilog paper, then the logarithms of the rates form a straight line on ordinary graph paper, and vice versa. So if you plot the logarithms of the rates on ordinary graph paper, and you see an increasing straight line, then you know that's exactly what you would see if you could plot the rates themselves on semilog paper. You conclude that the rates are increasing exponentially, and you have something exciting to report to your rescuers, if they ever come.



3
2
1
0
-1
LOG (RATE PER 100,000 POPULATION)
HEART DISEASE
SEPTICEMIA
50
55
60
65
70
75
80
85
90