Rate: A quantity, amount, or degree of something measured per unit of something else.

- Webster's Dictionary

Proportion: The comparative relation in size, amount, etc. a part, share, or fraction thereof.

-Webster's Dictionary

A rate has units, while a proportion does not. A proportion must lie between 0 and 1, while a rate may be any real number.

Illustrative example of comparing rates

Fact: In 1962 there were 73,555 deaths in Sweden, but only 7,781 in Panama (Rotham, Modern Epidemiology).

Why would this be so?

- Sweden has a larger population?
- To compare populations, use a death measure that accounts for the size of the population, i.e., the death rate.

$$\text{Death Rate} = \frac{\text{Number of deaths}}{\text{Total population}} = DR$$

$$\text{Death Rate (per 1,000)} = DR \times 1{,}000$$

	Sweden	Panama
Deaths	73,555	7,781
Population	7,496,000	1,075,000
Death Rate	0.009813	0.007238
Rate (/1000)	9.813	7.238

Death rate in Sweden is larger, by 36%
(9.813/7.238 = 1.356)

Why? Epidemic in Sweden?

Environmental differences?

Different age distributions?

	Sweden		Panama	
Age group	Population	relative Frequency	Population	relative Frequency
0 - 29	3,145,000	0.42	741,000	0.69
30 - 59	3,057,000	0.41	275,000	0.26
60 +	1,294,000	0.17	59,000	0.05
	7,496,000		1,075,000	

Sweden's population is older. Is this was accounts for the difference between the two death rates?

	Sweden		
Age group	Population	deaths	rate (per 1,000)
0 - 29	3,145,000	3,523	1.12
30 - 59	3,057,000	10,928	3.57
60+	1,294,000	59,104	45.68
	7,496,000	73,555	

	Panama		
Age group	Population	deaths	rate (per 1,000)
0 - 29	741,000	3,904	5.27
30 - 59	275,000	1,421	5.17
60+	59,000	2,456	41.63
	1,075,000	7,781	

The overall rate, 9.813 and 7.238 are the crude death rates, and the rates for the three age groups are called the age-specific death rates.

Q: How do we compare these two populations? (the problem is that **two** variables - age distribution and age-specific rates - are different)

A: Hold one of the variables constant.

Two possibilities:

(1) Hold the age distribution constant
(2) Hold the age-specific rate constant

Number (1) attempts to identify the effect of different age-specific rates.

Number (2) attempts to identify the effect of different age distributions.

Standardization:

Q: What would the two crude rates be if Sweden and Panama had the same age distribution?

For example, suppose that both countries had 1 million people distributed as follows:

Age group	Population	Relative frequency
0 – 29	500,000	0.50
30 – 59	400,000	0.40
60+	100,000	0.10
	1,000,000	

If we apply Sweden's age-specific death rates to this imaginary population, we can calculate a hypothetical (imaginary) death rate. It would represent what Sweden's crude death rate would be, if Sweden had the same age distribution as this imaginary population, but the age-specific rates remained unchanged.

			Sweden	
	standard	relative	rate	expected
Age group	population	frequency	per 1000	deaths
0 – 29	500,000	0.50	1.12	560
30 – 59	400,000	0.40	3.57	1428
60+	100,000	0.10	45.68	4568
	1,000,000			6556

500,000 x (1.12 / 1000) = 500 x 1.12 = 560
400,000 x (3.57 / 1000) = 400 x 3.57 = 1428
100,000 x (45.68 / 1000) = 100 x 45.68 = 4568

If Sweden had the age structure of the imaginary population, its death rate would be 6556/1,000,000, or 6.556 per thousand:

$$\frac{6{,}556}{1{,}000{,}000} = \frac{6.566}{1000}$$

this is the "standardized" or "age-adjusted" death rate for Sweden. It is the result of applying the Swedish age-specific rates to the population that we chose as the standard.

- Why is the age-adjusted rate (6.6) less than the crude rate (9.8)?

If Panama also had the age structure of the standard (imaginary) population, we would have:

			Panama	
Age group	standard population	relative frequency	rate per 1000	expected deaths
0 – 29	500,000	0.50	5.27	2635
30 – 59	400,000	0.40	5.17	2068
60+	100,000	0.10	41.63	4163
	1,000,000			8866

Standardized death rate for Panama:
8.866 per 1,000

After adjusting for the difference between the age distributions in this way, we now see that Panama has the larger death rate, by 35%
(8.866/6.556 =1.352).

Remember: Using only the observed crude death it appeared that Sweden's rate was larger, by 36%
(9.813/7.238 = 1.356).

Aside: Confounding

This example of comparing two crude death rates provides a nice illustration of confounding.

The differences in age distribution between Sweden and Panama confound our comparison about the death rates.

In general, confounding occurs when we fail to control for an important variable, in this case, age structure. Adjusting forces the age structure to be equal, so that we may compare the resulting death rates.

Take home message: If two crude death rates are different it could be because the age-specific death rates are different or because the age distributions are different or both. We cannot tell which without examining the age-specific data.

- The need for adjustment comes from the variability of the age-specific rates!

Take a look at the calculation of the age-adjusted death rate for Panama:

			Panama	
Age group	standard population	relative frequency	rate per 1000	expected deaths
0 – 29	500,000	0.50	5.27	2635
30 – 59	400,000	0.40	5.17	2068
60+	100,000	0.10	41.63	4163
	1,000,000			8866

Age-adjusted rate = 8.866 (per thousand)

$$= 0.50(5.27) + 0.40(5.17) + 0.10(41.63)$$

$$= w_1 r_1 + w_2 r_2 + w_3 r_3 = \sum_{i=1}^{3} w_i r_i \qquad (**)$$

Where w_i is the relative frequency of strata i and r_i is strata specific death rate

The w_i's are non-negative numbers between 0 and 1, such that $w_1 + w_2 + w_3 = 1$.

Expression (**) shows that the Age-adjusted death rate is a ***weighted average*** *of the three age-specific rates.*

Each country's crude death rate is also a weighted average of it age-specific rates. The weights in this case are the relative frequencies of the age groups in the actual population.

Panama's crude death rate = 7.238 (per thousand)

$$= 0.69(5.27) + 0.26(5.17) + 0.05(41.63)$$

$$= w_{p1}r_{p1} + w_{p2}r_{p2} + w_{p3}r_{p3} = \sum_{i=1}^{3} w_{pi}r_{pi}$$

Sweden's crude death rate = 9.813 (per thousand)

$$= 0.42(1.12) + 0.41(3.57) + 0.17(45.68)$$

$$= w_{s1}r_{s1} + w_{s2}r_{s2} + w_{s3}r_{s3} = \sum_{i=1}^{3} w_{si}r_{si}$$

Each countries Crude rate is a weighted average of its age-specific rates.

If the weights represent

- that country's actual age distribution,
  then the adjusted rate is just the Crude rate.

- some other age distribution,
  then the adjusted rate is a hypothetical (imaginary) rate representing what that country's crude rate would be if it had the other age distribution.

It <u>does</u> matter which standard population is chosen.

By manipulating the weights, we can change the ratio of Sweden's age-adjusted rates to Panama's age-adjusted rate.

That is, how much greater Panama's adjusted rate is over Sweden's adjusted rate depends on the standard population!

Important note:

Whenever one study population has an age-specific rate that is larger than the corresponding rate in the other, we can, by choosing a standard that puts all of the weight on that age group, make the first population have a larger standardized rate.

If in some other age group the <u>second</u> population has the larger rate, then by putting all of the weight on <u>that</u> age group, we can make the <u>second</u> population's standardized rate be the larger one.

But if one population's age-specific rates are larger in <u>every</u> age group, then that population will have the larger standardized rate <u>no matter what standard population is selected</u>.

Important note (con't): Example

	Sweden		Panama	
Age group	relative Frequency	Age rate	relative Frequency	Age rate
1: 0 - 29	w_{s1}	$20 = r_{s1}$	w_{p1}	$1 = r_{p1}$
2: 30 – 59	w_{s2}	$5 = r_{s2}$	w_{p2}	$5 = r_{p2}$
3: 60 +	w_{s3}	$1 = r_{s3}$	w_{p3}	$20 = r_{p3}$

Panama's adjusted rate = $\sum_{i=1}^{3} w_{pi} r_{pi}$ (per same as r_{si})

$$= w_{p1}\,(20) + w_{p2}\,(5) + w_{p3}\,(1)$$

$$= w_{p1}r_{p1} + w_{p2}r_{p2} + w_{p3}r_{p3}$$

Sweden's adjusted rate = $\sum_{i=1}^{3} w_{si} r_{si}$ (per same as r_{pi})

$$= w_{s1}\,(1) + w_{s2}\,(5) + w_{s3}\,(20)$$

$$= w_{s1}r_{s1} + w_{s2}r_{s2} + w_{s3}r_{s3}$$

Note our adjustment implies that $w_{si} = w_{pi}$ for all i.

- Try $w_1 = 0$, $w_2 = 0$, $w_3 = 1$

  Panama's adjusted rate =

  Sweden's adjusted rate =

- Try $w_1 = 0$, $w_2 = 1$, $w_3 = 0$

  Panama's adjusted rate =

  Sweden's adjusted rate =

- Try $w_1 = 1$, $w_2 = 0$, $w_3 = 0$

  Panama's adjusted rate =

  Sweden's adjusted rate =

Important special adjustment case: $r_{pi} = c\, r_{si}$ for some constant c. (adjustment implies $w_{si} = w_{pi}$ for all i)

Then $$\frac{\sum_{i=1}^{3} w_i r_{pi}}{\sum_{i=1}^{3} w_i r_{si}} = \frac{\sum_{i=1}^{3} w_i r_{pi}}{\sum_{i=1}^{3} w_i r_{si}} = \frac{c\sum_{i=1}^{3} w_i r_{si}}{\sum_{i=1}^{3} w_i r_{si}} = c$$

- What does this mean, in terms of implications for choosing the standard population (w_i's)?

Overview:

The crude rate is a weighted average of stratum specific rates, where the weights are the relative frequency of that stratum.

If there are k strata:

$$\text{Crude rate} = \sum_{i=1}^{k} w_i r_i = w_1 r_1 + \ldots + w_k r_k$$

Where w_i is the relative frequency in strata i and
r_i is strata specific rate

The w_i's are non-negative numbers between 0 and 1, such that $w_1 + \ldots + w_k = 1$.

The method of adjustment where we change the weights to make the age-distribution comparable is called the **Direct adjustment** method.

Another method of adjusting would be to force the stratum specific rates to be the same. This is called the **Indirect adjustment** method.

Indirect adjustment

Indirect adjustment uses rates from the standard population and the relative frequency from the two countries.

	Sweden		Panama	
Age group	relative Frequency	Age rate	relative Frequency	Age rate
1: 0 - 29	0.42 = w_{s1}	1.12 = r_{s1}	0.69 = w_{p1}	5.27 = r_{p1}
2: 30 – 59	0.41 = w_{s2}	3.57 = r_{s2}	0.26 = w_{p2}	5.17 = r_{p2}
3: 60 +	0.17 = w_{s3}	45.68 = r_{s3}	0.05 = w_{p3}	41.63 = r_{p3}

Use the standard population rates of (r_1=2, r_2=5, r_3=40 per 1000) to indirectly adjust the death rates above:

Sweden's IA death rate = $\sum_{i=1}^{3} w_{si} r_i$ (per 1,000)

$= w_{s1}r_{s1} + w_{s2}r_{s2} + w_{s3}r_{s3}$

$= 0.42\ (2) + 0.41\ (5) + 0.17\ (40) = 9.69$

Panama's IA death rate = $\sum_{i=1}^{3} w_{pi} r_i$ (per 1,000)

$= w_{p1}r_{p1} + w_{p2}r_{p2} + w_{p3}r_{p3}$

$= 0.69\ (2) + 0.26\ (5) + 0.05\ (40) = 4.68$

The resulting IA rate for Sweden is the hypothetical crude death rate of the standard population if the age distribution was the same as Sweden's.

Thus we are effectively holding the strata-specific rates constant and examining the effect of the age-distribution.

The Indirectly Adjusted rate (IA) is used for a different purpose than the Directly adjusted rate, even though the ultimate goal is to compare two different populations.

Direct adjustment compares the strata specific rates between two populations, while Indirectly Adjusted rates compare the age structure between two populations.

SMR: Standardized mortality ratio

The standardized mortality ratio for a population is defined as:

SMR = Observed # of deaths/Expected # of deaths

Note: The expected # of deaths is calculated using the rates from the standard population and the age-distribution of the population in question.

This is similar to Indirect adjustment. In fact, there is a simple relation:

IA rate = Crude rate / SMR

Or

SMR = Crude rate / IA

Where

Crude rate = observed # deaths / total population

Indirect Adj. = expected # deaths / total population

Example: Expected number of deaths

	Sweden		Panama	
Age group	Frequency	Observed deaths	Frequency	Observed Deaths
1: 0 - 29	3,145,000	3,523	741,000	3,904
2: 30 – 59	3,057,000	10,928	275,000	1,421
3: 60 +	1,294,000	59,104	59,000	2,456
	7,496,000	73,555	1,075,000	7,781

Use the standard population rates of ($r_1=2$, $r_2=5$, $r_3=40$ per 1000) to find the expect number of deaths:

Sweden's expected # of deaths

= 3,145 (2) + 3,057 (5) + 1,294 (40) = 73,335

SMR = 73,555 / 73,335 = 1.003 =approx= 9.813 / 9.69

Panama's expected # of deaths

= 741 (2) + 275 (5) + 59 (40) = 5,217

SMR = 7,781 / 5,217 = 1.4915 =approx= 7.238 / 4.68

Approximations arise from initially rounding off relative frequencies.

To Think about:

In Epidemiology, the denominator of a rate is often expressed in person-years. For the following table, calculate the **crude death rate** and **age-specific death rates** for each population. The death rates should be expressed as deaths per 100,000 person-years at risk. (Modified source: Selvin)

	Population 1		Population 2	
Age	Deaths	P-Years	Deaths	P-Years
40- 49	1	1,000	1	1,000
50 – 59	3	1,500	10	5,000
60 - 69	8	2,000	40	10,000
70 - 79	20	2,500	160	20,000
Total	32	7,000	211	36,000

(a) What would you conclude regarding risk of death in the two populations by examining the crude death rates only?

(b) What would you conclude regarding risk of death in the two populations by examining age-specific death rates only?

To Think about (continued):

(c) It is necessary to control for the effects of age through an adjustment procedure when comparing crude death rates between these two populations? Why or why not?

(d) If the table was changed to the following, would it be necessary to calculate adjusted rates?

	Population 1		Population 2	
Age	Deaths	P-Years	Deaths	P-Years
40- 49	1	1,000	3	2,000
50 – 59	3	1,500	10	3,000
60 - 69	8	2,000	40	4,000
70 - 79	20	2,500	160	5,000
Total	32	7,000	211	36,000