Probability, Odds Ratio and Risk Ratio

Dr. Abbas Adigun (PhD)

Biostatistician

19th May 2017

Probability

Probability is a measure of the chance of getting some outcome of interest from some event.

The event might be rolling a dice and the outcome of interest might be getting a six;

or the event might be performing a biopsy with the outcome of interest being evidence of malignancy and so on.

Some basics of probabilty

1.The probability of a particular outcome from an event will lie between zero and one.

2. The probability of an event that is certain to happen is equal to one. For example, the probability that everybody dies eventually.

3. The probability of an event that is impossible is zero. For example, throwing a seven with a normal dice.

4. If an event has as much chance of happening as of not happening, then it has a probability of $^1/_2$ or 0.5.

5. If the probability of an event happening is *p*, then the probability of the event *not* happening is 1 – *p*.

Calculating probability

The probability of a particular outcome from an event is equal to the number of outcomes that favour that event, divided by the *total* number of possible outcomes.

Example

Consider a simple example: What is the probability of getting an even number when you roll a dice?

Total number of possible outcomes = 6 (1 or 2 or 3 or 4 or 5 or 6)

Total number of outcomes favouring the 'event an even number' = 3 (i.e. 2 or 4 or 6)

So, the probability of getting an even number = 3/6 = $^1/_2$= 0.5

However, In the real world you will often have to use what is called the *proportional frequency* approach, which uses existing frequency data as the basis for probability calculations.

Example: Consider the table below which shows the causes of blunt injury to limbs. The table shows that the sum of proportional frequency is equal to one which implies that it can be interpreted as equivalent to probabilities.

Causes of injury	N=75 No of Patients	Proportional frequency
Falls	46	0.613
Crush	20	0.267
Motor vehicle crash	6	0.08
Other	3	0.040

What is the probability that if you chose one of these 75 patients at random their injury will have been caused by a Motor vehicle crash?

The probability is 0.08

Risk

The risk of any particular outcome from an event is equal to the number of favourable outcomes divided by the total number of outcomes. This implies risk is the same as probability.

As an example, let us look at the contingency table 6.1 from the cohort study of coronary heart disease (CHD) in adult life and the risk factor ‘weighing 8.16 kg or less at the age of one year.

			Weighed less than 8.16kg at age 1		
			yes	no	Total
Has CHD		**yes**	4	38	42
		no	11	237	248
		Total	15	275	290

To find the risk that an individual who weighed <= 8.16kg at one year will have CHD.

$$\text{Risk} = \frac{\text{no of people who weighed} <= 8.16kg \ \ and\ had\ CHD}{\text{no of people who weighed} <= 8.16kg}$$

$$= \frac{4}{15}$$

$$=0.2667$$

Also, find the risk for an individual who weighed > 8.16kg at one year will have CHD

$$\text{Risk} = \frac{\text{no of people who weighed} > 8.16kg \ \ and\ had\ CHD}{\text{no of people who weighed} > \ 8.16kg}$$

$$= \frac{38}{275}$$

$$= 0.1382$$

The risk for a single group, as it is described it above, is also known as the *absolute risk*, mainly to distinguish it from *relative risk*, which is the risk for one group *compared* to the risk for some other group.

The risk ratio

In practice, risks and odds for a single group are not nearly as interesting as a *comparison* of risks and odds between *two* groups. For risk you can make these comparisons by dividing the risk for one group (usually the group exposed to the risk factor) by the risk for the second, non-exposed, group. This gives us the *risk ratio*.

We can generalize the risk ratio calculation with the help of contingency table like the table 6.2 where the cell values a represented as *a b c and d*

		Group exposed to risk factor		
		yes	no	Total
Has disease	yes	a	b	a+b
	no	c	d	c+d
Total		a+c	b+d	

Among those exposed to the risk factor,

$$\text{risk of disease} = \frac{a}{(a + c)}$$

Among those not exposed,

$$\text{risk of disease} = \frac{b}{b + d}$$

The risk ratio is therefore is

$$= \frac{a}{(a + c)} \Big/ \frac{b}{b + d}$$

$$= \frac{a(b + d)}{b(a + c)}$$

Worked example

Find the risk ratio of having CHD between those two groups.

Solution

$$a = 4, b = 38, c = 11, d = 237$$

$$= \frac{4(38 + 237)}{38(4 + 11)}$$

$$= \frac{1100}{570}$$

$$= 1.9298$$

Stata command

```
use "C:\Users\ADIGUM\Desktop\abuja\risk.dta",  clear
 cs diseasevariable exposurevariable, exact
cs chd weight_less_8_16
```

if you already have the contingency table

```
csi 4 38 11 237
```

or csti 4/15 38/275

Odds

The *odds* for an event is equal to the number of outcomes favourable to the event divided by the number of outcomes not favourable to the event.

Notice that:

1. The value of the odds for an outcome can vary from zero to infinity.

2. When the odds for an outcome are less than one, the odds are *unfavourable* to the outcome; the outcome is *less* likely to happen than it is *to* happen.

3. When the odds are equal to one, the outcome is as likely to happen as not.
4. When the odds are greater than one, the odds are *favourable* to the outcome; the outcome is *more* likely to happen than not.

Example

The table A underneath shows the outcome from the exercise and stroke a case-control study for those subjects who had and who had not exercised during the age between 15 and 25.

Table A

Exercised undertaken during the age between 15-25		Cases (Stroke)	Control
	yes	55	130
	no	70	68

From the table we can find the odds of exercise among those who had stroke

Odds of exercise among those who had stroke

$$= \frac{numbers\ of\ cases\ who\ exercise}{number\ of\ cases\ without\ exercise}$$

$$= \frac{55}{70}$$

$$= 0.7857$$

We can as well calculate the odds of exercise among those who hadn't had stroke

$$= \frac{numbers\ of\ control\ who\ exercise}{number\ of\ contol\ without\ exercise}$$

$$= \frac{130}{68}$$

$$= 1.9118$$

In other words, among those who'd had a stroke, the odds that they had exercised was less than half the odds (0.7857/1.9118) of those who hadn't had a stroke.

The odds ratio

An odds ratio (OR) is a measure of association between an exposure and an outcome.

In a case-control study you can compare the odds that those with a disease will have been exposed to the risk factor, with the odds that those who don't have the disease or condition will have been exposed.

If you divide the former by the latter you get the *odds ratio*.

We can generalize the odds ratio calculation with the help of the 2 × 2 table below

			Cases	Control
Exposed to risk factors		Yes	a	b
		no	c	d

The odds of exposure to the risk factor among those with the disease = $\frac{a}{c}$

The odds of exposure to the risk factor among the healthy controls = $\frac{b}{d}$

The formula for calculating odds ratio is therefore

$$\frac{a}{c} \Big/ \frac{b}{d} = \frac{ad}{bc}$$

Example

Find the odds ratio of exercise for those with stroke compared to those without stroke from table A.

Solution

$$a = 55\ b = 130\ c = 70\ d = 68$$

$$\text{Odds ratio} = \frac{ad}{bc} = \frac{55 \times 68}{130 \times 70}$$

$$= 0.4110$$

Stata commands

use "C:\Users\ADIGUM\Desktop\abuja\odd_ratio.dta", clear

cci 55 70 130 68 for cell values

cc diseasevariable exposurevariable

cc stroke exe


CORBIS/Brian Leng (05065)