Risks and Odds

Andy Grogan-Kaylor

23 Jun 2020 13:28:58

Setup

```
. clear all
```

We are going to set up a table with 10 rows of information.

```
. set obs 10 // 10 rows of information
number of observations (_N) was 0, now 10

. generate events = 100 // 100 hypothetical events
```

In each row of the table our *event of interest* happened a different number of times.

In the code below we make use of Stata's special variable for the row number of a data set: `_n`.

```
. generate it_happened = _n * 10
```

As a result, the *event of interest didn't* happen 100 − happened times.

```
. generate it_didnt_happen = 100 - it_happened
```

> As you think through the rest of this example, it might be worth giving yourself a concrete example of the *event of interest*. What is a concrete example of a good thing that might happen, or a bad thing that might happen?

Our Table of Information So Far

Let's list out our table of information so far:

```
. list, abbreviate(20)
```

	events	it_happened	it_didnt_happen
1.	100	10	90
2.	100	20	80
3.	100	30	70
4.	100	40	60
5.	100	50	50
6.	100	60	40
7.	100	70	30
8.	100	80	20
9.	100	90	10
10.	100	100	0

Risk

Now let's think about risk: risk $= P(\text{event}) = \frac{\text{number of events}}{\text{number of events+number of non-events}}$

```
. generate risk_it_happened = it_happened / (it_happened + it_didnt_happen)
```

There is also a risk that the event didn't happen.

```
. generate risk_it_didnt_happen = it_didnt_happen / (it_happened + it_didnt_happen)
```

Odds

The odds are the probability that an event happened divided by the probability that it did not happen

odds $= \frac{P(\text{happened})}{P(\text{didn't happen})}$

$= \frac{\frac{\text{number of events}}{\text{number of events+number of non-events}}}{\frac{\text{number of non-events}}{\text{number of events+number of non-events}}}$

```
. generate odds = risk_it_happened / risk_it_didnt_happen
(1 missing value generated)
```

which incidentally reduces to

$= \frac{\text{number of events}}{\text{number of non-events}}$

Look At Our Table Of Information

Let's look at our table of information.

Notice how the odds start to overstate the risk, as the risk becomes more common.

```
. list it_happened it_didnt_happen risk_it_happened risk_it_didnt_happen odds, abbreviate(15)
```

	it_happened	it_didnt_happen	risk_it_happe~d	risk_it_didnt~n	odds
1.	10	90	.1	.9	.1111111
2.	20	80	.2	.8	.25
3.	30	70	.3	.7	.4285715
4.	40	60	.4	.6	.6666666
5.	50	50	.5	.5	1
6.	60	40	.6	.4	1.5
7.	70	30	.7	.3	2.333333
8.	80	20	.8	.2	4
9.	90	10	.9	.1	9
10.	100	0	1	0	.

We can even graph this.

```
. twoway scatter odds risk_it_happened it_happened, ///
> title("Risk and Odds") ///
> xtitle("How many times out of 100 did this happen?") ///
> scheme(michigan)

. quietly: graph export myscatter.png, width(500) replace
```

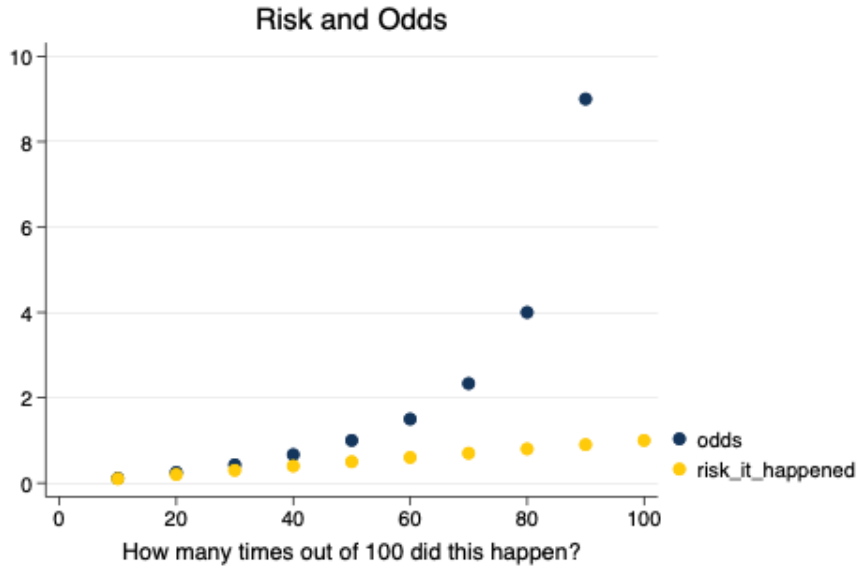

Figure 1: Risk and Odds