Risk, relative risk, odds, relative odds

Principal Question:

Is there a **relationship** between the two variables, so that the category into which individuals fall for one variable seems to depend on the category they are in for the other variable?

6.1 Displaying Relationships Between Categorical Variables

- Data displayed in a **contingency** or two-way table.
- If one variable is explanatory, use it to define the rows of the table.
- Two types of **conditional percents**: row percents and column percents.
- Use row percents if the explanatory variable is the row variable.

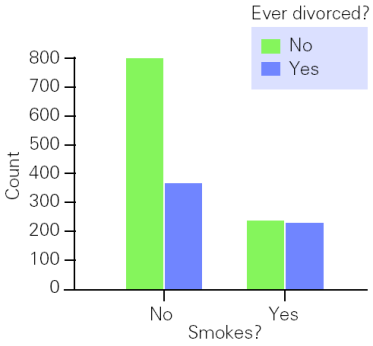
Example 1 *Smoking and Divorce Risk*

Data on smoking habits and divorce history for the 1669 respondents who had ever been married.

TABLE 6.1 Smoking and Divorce, GSS Surveys 1991–1993

	Ever Divorced?		
Smoke?	*Yes*	*No*	*Total*
Yes	238	247	485
No	374	810	1184
Total	612	1057	1669

Data Source: SDA archive at UC Berkeley web site (www.csa.berkeley.edu:7502/).



Among smokers, 49% have been divorced, 51% have not.
Among nonsmokers, only 32% have been divorced, 68% have not.
The difference between row percents indicates a relationship.

Example 2 *Tattoos and Ear Pierces*

Responses from $n = 565$ men to two questions:
1. Do you have a tattoo?
2. How many total ear pierces do you have?

TABLE 6.2 Ear Pierces and Tattoos for Men ($n = 565$)

Ear Pierces	No Tattoo	Tattoo	Total
0	381	43	424
1	54	16	70
2 or more	45	26	71
Total	480	85	565

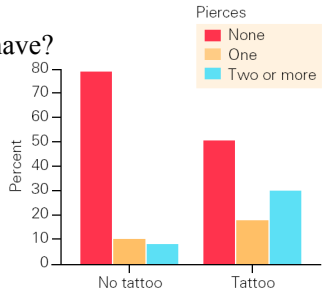


FIGURE 6.2 Column percents for the ear pierce and tattoo data

Among men with no ear pierces, 43/424 = 10% have a tattoo.
Among men with one ear pierce, 16/70 = 23% have a tattoo.
Among men with two or more ear pierces, 26/71 = 37% have a tattoo.
% with a tattoo ↑ as number of ear pierces ↑ => relationship
Could examine column percents (see graph above) or overall percents too.

Example 3 *Gender and Reasons for Taking Care of Your Body*

1997 poll (random-digit dialing) of 1218 southern CA residents.
Question: What is the most important reason why you try to take care of your body? Is it mostly to be attractive to others, mostly to keep healthy, or mostly to help your self-confidence, or what?

TABLE 6.3 Reasons for Taking Care of Body and Gender

	Healthy	Self-Confidence	Attractive	Don't Know
Men	76%	16%	7%	1%
Women	74%	20%	4%	2%

Source: www.latimes.com, poll archives, study #401.

Percent distribution of responses shown for men and women.
Pattern of responses is **very similar**.
Response does not seem to be related to gender.

2. Risk, Relative Risk, Odds Ratio, and Increased Risk

$$\textbf{Risk} = \frac{\text{Number in category}}{\text{Total number in group}}$$

Example:

Within a group of 200 individuals, asthma affects 24 people. In this group the *risk* of asthma is 24/200 = 0.12 or 12%.

$$\textbf{Relative Risk} = \frac{\text{Risk in category 1}}{\text{Risk in category 2}}$$

Example:

For those who drive under the influence of alcohol, the relative risk of an accident is 15. => The risk of an accident for those who drive under the influence is 15 times the risk for those who don't drive under the influence.

- Relative risk = 1 => two risks are the same.
- Risk in denominator often the baseline risk.

Example 1 *Smoking and Divorce Risk (cont)*

TABLE 6.1 Smoking and Divorce, GSS Surveys 1991–1993

	Ever Divorced?		
Smoke?	*Yes*	*No*	*Total*
Yes	238	247	485
No	374	810	1184
Total	612	1057	1669

Data Source: SDA archive at UC Berkeley web site (www.csa.berkeley.edu:7502/).

- For smokers: risk of divorce = 238/485 = 0.491 or 49.1%.
- For nonsmokers: risk of divorce = 374/1184 = 0.316 or 31.6%.

$$\text{Relative Risk of divorce} = \frac{49\%}{32\%} = 1.53$$

In this sample, the risk of divorce for smokers is 1.53 times the risk of divorce for nonsmokers.

Percent increase in risk

$$= \frac{\text{Difference in risks}}{\text{Baseline risk}} \times 100\%$$

$$= (\text{relative risk} - 1) \times 100\%$$

Note:

When risk is smaller than baseline risk, relative risk < 1 and the percent "increase" will actually be negative, so we say *percent decrease* in risk.

Example 1 *Smoking and Divorce Risk (cont)*

Relative Risk of divorce for smokers = 1.53

Percent increase in risk of divorce for smokers
= (1.53 – 1) x 100% = 53%

$$= \frac{\text{Difference in risks}}{\text{Baseline risk}} \times 100\% = \frac{(49 - 32)}{32} \times 100\%$$

= 53%

The risk of divorce is 53% higher for smokers than it is for nonsmokers.

Odds

= Number in category 1 to Number in category 2
= (Number in category 1/Number in category 2) to 1

Odds Ratio

= (Odds for group 1) / (Odds for group 2)

Example:

Odds of getting a divorce to *not* getting a divorce for smokers are 238 to 247 or 0.96 to 1.

Odds of getting a divorce to *not* getting a divorce for nonsmokers are 374 to 810 or 0.46 to 1.

Odds Ratio = 0.96 / 0.46 = 2.1 => the odds of divorce for smokers are about double the odds for nonsmokers.

3 Misleading Statistics About Risk

Questions to Ask:

- What are the actual risk? What is the baseline risk?
- What is the population for which the reported risk or relative risk applies?
- What is the time period for this risk?

Example 4 *Disaster in the Skies?*

"**Errors by air traffic controllers** climbed from 746 in fiscal 1997 to 878 in fiscal 1998, an **18% increase**." *USA Today*

Look at risk of controller error *per flight*:

In 1998: 5.5 errors per million flights

In 1997: 4.8 errors per million flights

Risk of error increased but the **actual risk is *very* small**.