range listed was (1.0 - 2.0). Most people would consider this interval (1.0 - 2.0) to imply that values close to 1.5 to be "normal" and a test result close to the boundaries, either 1 or 2, suggests something to watch for. Indeed, my physician asked me to schedule an appointment with an endocrinologist for further studies.

Since the quality of doctors and I showed him the test results. He laughed "For years I have told them that they shouldn't provide "normal limits" as such when the distribution is highly skewed! You actually have the most typical value in the population!

Suppose that a University is trying to discriminate in favor of women when hiring taff. It advertises positions in the Department of History and in the Department of Geography, and only those departments. Five men apply for the positions in History and one is hired, and eight women apply and two are hired. The success rate for men is twenty percent, and the success rate for women is twenty-five percent. The History Department has favored women over men. In the Geography Department eight men apply and six are hired, and five women apply and four are hired. The success rate for men is seventy-five percent and for women it is eighty percent. The Geography Department has favored women over men. Yet across the University as a whole 13 men and 13 women applied for jobs, and 7 men and 6 women were hired. The success rate for male applicants is greater than the success rate for female applicants.

	Men		Women
History	1/5	<	2/8
Geography	6/8	<	4/5
University	7/13	>	6/13

How can it be that each Department favors women applicants, and yet overall men fare better than women? There is a 'bias in the sampling', but it is not easy to see exactly where this bias arises. There were 13 male and 13 female applicants: equal sample sizes for both groups. Geography and History had 13 applicants each: equal sample sizes again. Nor does the trouble lie in the fact that the samples are small: multiply all the numbers by 1000 and the puzzle remains.

range listed was (1.0 - 2.0). Most people would consider this interval (1.0 - 2.0) to imply that values close to 1.5 to be "normal" and a test result close to the boundaries, either 1 or 2, suggests something to watch for. Indeed, my physician asked me to schedule an appointment with an endocrinologist for further studies.

Since the quality of doctors and I showed him the test results. He laughed "For years I have told them that they shouldn't provide "normal limits" as such when the distribution is highly skewed! You actually have the most typical value in the population!

Suppose that a University is trying to discriminate in favor of women when hiring taff. It advertises positions in the Department of History and in the Department of Geography, and only those departments. Five men apply for the positions in History and one is hired, and eight women apply and two are hired. The success rate for men is twenty percent, and the success rate for women is twenty-five percent. The History Department has favored women over men. In the Geography Department eight men apply and six are hired, and five women apply and four are hired. The success rate for men is seventy-five percent and for women it is eighty percent. The Geography Department has favored women over men. Yet across the University as a whole 13 men and 13 women applied for jobs, and 7 men and 6 women were hired. The success rate for male applicants is greater than the success rate for female applicants.

	Men		Women
History	1/5	<	2/8
Geography	6/8	<	4/5
University	7/13	>	6/13

How can it be that each Department favors women applicants, and yet overall men fare better than women? There is a 'bias in the sampling', but it is not easy to see exactly where this bias arises. There were 13 male and 13 female applicants: equal sample sizes for both groups. Geography and History had 13 applicants each: equal sample sizes again. Nor does the trouble lie in the fact that the samples are small: multiply all the numbers by 1000 and the puzzle remains.

Surgeon A vs. Surgeon B

- 95 patients out of 100 survived with surgeon A, so 95/100 = 95% of them survived.
- 72 patients out of 80 survived with surgeon B, so 72/80 = 90% of them survived.

Which Surgeon would you choose?

Surgeon A:

Of the 100 patients that surgeon A treated, 50 were high risk, of which three died. The other 50 were considered routine, and of these 2 died. This means that for a routine surgery, a patient treated by surgeon A has a 48/50 = 96% survival rate

Surgeon B:

Now we look more carefully at the data for surgeon B and find that of 80 patients, 40 were high risk, of which seven died. The other 40 were routine and only one died. This means that a patient has a 39/40 = 97.5% survival rate for a routine surgery with surgeon B.

6.3 Simpson's paradox

As is the case with quantitative variables, the effects of lurking variables can change or even reverse relationships between two categorical variables. Here is an example that demonstrates the surprises that can await the unsuspecting user of data.

EXAMPLE 6.4 Do Medical Helicopters Save Lives?


© Ashley Cooper/Corbis

Accident victims are sometimes taken by helicopter from the accident scene to a hospital. Helicopters save time. Do they also save lives? Let's compare the percents of accident victims who die with helicopter evacuation and with the usual transport to a hospital by road. Here are hypothetical data that illustrate a practical difficulty:[4]

	Helicopter	Road
Victim died	64	260
Victim survived	136	840
Total	200	1100

We see that 32% (64 out of 200) of helicopter patients died, but only 24% (260 out of 1100) of the others did. That seems discouraging.

The explanation is that the helicopter is sent mostly to serious accidents, so that the victims transported by helicopter are more often seriously injured. They are more likely to die with or without helicopter evacuation. Here are the same data broken down by the seriousness of the accident:

Serious Accidents		
	Helicopter	Road
Died	48	60
Survived	52	40
Total	100	100

Less Serious Accidents		
	Helicopter	Road
Died	16	200
Survived	84	800
Total	100	1000

Inspect these tables to convince yourself that they describe the same 1300 accident victims as the original two-way table. For example, 200 (= 100 + 100) were moved by helicopter, and 64 (= 48 + 16) of these died.

Among victims of serious accidents, the helicopter saves 52% (52 out of 100) compared with 40% for road transport. If we look only at less serious accidents, 84% of those transported by helicopter survive, versus 80% of those transported by road. Both groups of victims have a higher survival rate when evacuated by helicopter. ■

How can it happen that the helicopter does better for both groups of victims but worse when all victims are lumped together? Examining the data makes the explanation clear. Half the helicopter transport patients are from serious accidents, compared with only 100 of the 1100 road transport patients. So the helicopter carries patients who are more likely to die. The seriousness of the accident was a lurking variable that, until we uncovered it, hid the true relationship between survival and mode of transport to a hospital. Example 6.4 illustrates *Simpson's paradox*.

EXAMPLE

In one study, Kanter and his colleagues asked doctors in Germany and the United States to estimate the probability that a woman with a positive mammogram actually has breast cancer with no symptoms or family history of breast cancer.

The probability that one of these women has breast cancer is 1 percent. If a woman has breast cancer, the probability is 90 percent that she will have a positive mammogram. If a woman does not have breast cancer, the probability is 10 percent that she will still have a positive mammogram. Imagine a woman who has a positive mammogram. **What is the probability that she actually has breast cancer?**

When Kanter asked 24 other German doctors the same question, their estimates whipsawed from 1 percent to 90 percent. Eight of them thought the chances were 10 percent or less, 8 more said 90 percent, and the remaining 8 guessed somewhere between 50 and 80 percent. Imagine how upsetting it would be as a patient to hear such divergent opinions.

As for the American doctors, 95% estimated the woman's probability of having breast cancer to be somewhere around 75 percent.

$$P(A_1 \mid B) = \frac{P(A_1)\,P(B \mid A_1)}{P(A_1)\,P(B \mid A_1) + P(A_2)\,P(B \mid A_2)}$$

The right answer is 9 percent.