

This document is an excerpt from
Resampling Stats in MATLAB
Daniel T. Kaplan

Copyright (c) 1999 by Daniel T. Kaplan, All Rights Reserved

This document differs from the published book in pagination and in the omission (unintentional, but unavoidable for technical reasons) of figures and cross-references from the book. It is provided as a courtesy to those who wish to examine the book, but not intended as a replacement for the published book, which is available from Resampling Stats, Inc.

www.resample.com
703-522-2713

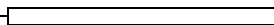
Chapter 2: Making Statements with Precision: Confidence Intervals

Whenever you make a measurement it is important to have some idea of how precise that measurement is. Without information about the precision of a measurement, it is impossible to know whether the measurement is applicable to the purpose you have in mind. There's no point, for instance, in taking the temperature of a possibly sick child if the thermometer is precise only to within ± 5 degrees; if you measured 100 degrees you would have no way to know if the child were perfectly healthy or had a possibly fatal disease. The measurement of the distance between two cities might need to be precise only to a mile or so for the purposes of planning a driving trip, but might need to be precise to within a centimeter for the purpose of detecting continental drift or for predicting an earthquake.

A *confidence interval* is a standard way of describing the precision of a measurement of some parameter. There are two common forms for stating confidence intervals both of which provide exactly the same information, for example:

- The measurement is 10 ± 3 , with 95% confidence.
- The measurement will be between 7 and 13, with 95% confidence.

There are three pieces of information in a confidence interval: the lower bound, the upper bound, and the *confidence level*. In the above example, the lower bound is 7, the upper bound is 13, and the confidence level is 95%.



The confidence level provides information about how sure you are that the true value of the parameter lies between the upper and lower bounds. This is not a subjective indication of confidence: one doesn't merely substitute the number 95% for the phrase "I'm quite sure." Instead, the confidence interval is calculated using mathematical techniques that insure that the confidence level is objective. This section shows some examples of how to compute confidence intervals using re-sampling.

The use of a confidence level of 95% is the convention in science — other levels can be used but whatever level you use it should always be stated explicitly. Confidence intervals computed using a smaller confidence level will be smaller than those computed with a larger confidence level; it is not possible directly to compare two confidence intervals made at differing confidence levels. Although it might seem best to use a confidence level of 100%, in most cases this would result in a confidence interval that was meaninglessly big and is more difficult to compute (see Chapter ??).

EXAMPLE 1: THE CAMPAIGN ADVISOR STATES CONFIDENTLY....

The political advisor to a fringe group's candidate for President has news just 10 days before the election. "I'm afraid the latest poll looks grim, sir. Only 8% of voters support us. We pledged to withdraw from the election if we appealed to fewer than 10% of voters — perhaps now is the time to get out of the race."

Interrupting, the legal advisor opines, "But if we get more than 10% of the vote, we are guaranteed federal campaign money in the next election. I think we should stay in."

The candidate, Ross Anderson, summarizes, "So if we get out now we won't be criticized for splitting the vote and we remain viable for the next election. But if we stay in and get 10% of the vote, we'll get campaign money for next time. [Expletive deleted]. How confident are you in that 8% figure?"

"Quite confident, sir," the advisor responds. "The pollsters talked to 500 randomly selected registered voters. Only 40 of them supported us. That makes ...," he taps on his calculator, "exactly 8%."

The above fictional exchange is unrealistic: modern political advisors know enough about statistics to realize that the 8% figure isn't completely precise. It is based on a random sample. If a different group of 500 had been picked the result would have been somewhat different.

More to the point, there is some probability that a random sample of 500 would show only 8% support even if the true support level were greater than 10%. The candidate's question about confidence in the results is really a question about whether there is any possibility that the true support level is more than 10%.

Here's one way to answer that question: collect more data. But the candidate needs to make a decision about withdrawing now and it takes time (and money!) to collect more data. The candidate needs to infer from the data at hand how precise is the 8% figure. So, rather than actually running several more polls, we use resampling to simulate the polls. Then we look at how much sampling variability there is in the simulated polls.

The following commands are in the file `poll.m`.

```
pollsize = 500;
voters = [92 0; 8 1]; % The data showed 8% support.
                % Code support as 1.
z = starttally; % keep the results in variable z
for trials = 1:1000
    poll = sample(pollsize, voters);
        % take a random sample.
        % This is done with replacement so that
        % the sample is being taken from a virtually
        % infinite population.
    support = count(poll==1)/pollsize;
    tally support z;
end
```

After running this script, the variable `z` contains 1000 simulations of polling 500 voters. We can compute the 95% confidence interval by taking the lower bound to be the 2.5% percentile and the upper bound to be the 97.5% percentile. Only 2.5% of the values in `z` are less than the lower bound, and another 2.5% of the values in `z` are greater than the upper bound, making altogether 5% of the values in `z` that would be outside of the bounds of the confidence interval.¹

```
>> percentile(z, [.025 .975])
ans:    0.058 0.106
```

¹To compute a confidence interval at, say, a 90% confidence level we would set the bounds as the 5% and 95% percentiles, i.e.,

```
>> percentile(z, [.05 .95])
```

and likewise for other confidence levels.

This 95-percent confidence interval suggests some possibility that the candidate has more than the magic 10% level needed to qualify for federal funding.

“Those poll data are worthless!” the candidate shouts. “Get me some solid numbers now!” he demands.

EXAMPLE 2: UNCERTAINTY IN THE MEAN

A light-bulb manufacturer is designing new packaging for a light bulb. They need to print on the package how long the bulb will last. To estimate this value, they take 10 bulbs from the production line and put them in a machine which turns them on and off once each hour. When each bulb burns out, the total number of hours that the bulb lasted is recorded. The data are:

```
>> data = [ 2103 2786 2543 1987 7 2380 3102 2452 3453 2543];
```

In presenting these data to consumers, the company wants to give an accurate portrayal of how long a given bulb will last, but it is understandably reluctant to say that there is some chance that the bulb will burn out almost immediately. Reasoning that customers buy lots of light bulbs and that what matters is not how long an individual bulb lasts but how long they last on average, the company decides to print the “Average Lifetime” of this type of bulb on the package. The average is:

```
>> mean(data) ⇒ ans: 2335.6
```

This number is sent down to the marketing division to be put on the new package.

The package’s graphic designer telephones to the research department. “Is it alright if we round the number to 2500 hours? That looks better on the package.” The researcher knows that the number 2335.6 is not exact, but isn’t quite sure that 2500 wouldn’t be misleading. Saying 2000 hours seems like it would be conservative, but perhaps that is less accurate than 2500. To find out, the researcher could do some more bulb-life testing runs, but these take about 4000 hours to set up and run — about 6 months.

The lifetime of each bulb is a random variable that depends on factors such as fluctuations in the diameter of the filament wire. This random variable has a distribution: we have 10 samples from that distribution that provide our best estimate of what the distribution looks like.

The mean lifetime of 10 bulbs is also a random variable; it varies depending on the specific random sample of 10 bulbs taken from produc-

tion. The researcher wants to find out something about this distribution of the mean. Ideally, given 6 months to do the work, the researcher would sample more bulbs from the production line. But, being unable to do this, the researcher can resample from the data she has already collected:

```
z = starttally;
for trials = 1:1000
    b = sample(10, data);
    mn = mean(b);
    tally mn z;
end
hist(z);
percentile(z, [.025 .975])
```

The result is that the 95% confidence interval for the mean ranges from about 1770 hours to about 2840. (If you run this example, you will find that the exact values will vary a bit from one run to the next since the resampling is done at random. More details on this are given in Sec. ??.) The researcher writes up another memo to marketing: “The lifetime of any one bulb is random, but our best information is that the mean of 10 bulbs selected randomly from our production line will be between 1770 and 2840 hours, 95% of the time.” The marketing department puts 2500 on the package.

The resampling method can be used to compute the confidence interval of many different types of statistics. The general term used for this type of calculation is *bootstrapping*. With the advent of bootstrapping and computers, it has become possible, and even easy, to compute confidence intervals on all sorts of statistics, even ones that you make up for a particular purpose. Historically, statistics and statistics textbooks have emphasized certain statistics, such as the mean or proportion, for which there are algebraic formulas for computing confidence intervals. The two examples just presented could easily be handled by these historically conventional techniques. Since we are now freed from the computational restrictions that have favored the choice of such statistics, one can expect to see increasing diversity in the use of types of descriptive statistics. Below we present three simple examples involving some simple descriptive statistics: the minimum, the median, the total distance. A fourth example shows how confidence intervals can be computed when more than one type of measurement is involved and you’re interested in

the relationship between the measurements.

EXAMPLE 3: CONFIDENCE IN THE EXTREME

Several months after the release of its new package, the light bulb company receives many complaints from consumers that its bulbs last only a few hours rather than the 2500 promised. They are sued by a man who broke his collar bone when replacing a light bulb above his stairway; he was so angry to be replacing a bulb that burned out on its first night of use that he fell off his ladder. The company decides to repackage. Rather than reporting the typical lifetime of a single bulb, they will report the total lifetime of all the bulbs in the 4-bulb package. Since this is a random number, they will help consumers by giving a value that is 99% guaranteed to be exceeded by the bulbs in the package. The following commands are in file `bulblife.m`:

```
z = starttally;
for trials = 1:5000 % See the caution below
    b = sample(4, data);
    totallife = sum(b);
    tally totallife z;
end
percentile(z,.01)
```

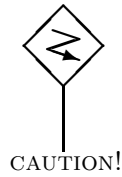
We execute the commands in the script

```
>> bulblife ⇒ ans: 4475
```

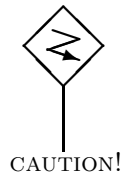
Only in 1% of the cases is the total lifetime less than 4475 hours. The lightbulb company sets the 99% guarantee level at 4500 hours and marks the package:

“The total lifetime of these 4 bulbs will be about 9000 hours but will vary from package to package. There is only a one-percent chance that the bulbs in this package will have a total lifetime of less than 4500 hours; if this happens you get your money back.”

This example differs from many of the others in that we are interested in the 1% percentile rather than 2.5% or 5%. It is harder to make reliable estimates for smaller percentiles. In Section ?? it is suggested that while 1000 trials is adequate for a 2.5%, it isn't adequate for a 1% percentile where 5000 trials should be used. Even so, as we look at more extreme percentiles, the estimates become less reliable. (See [?] for an analysis of this situation, and a way to correct the estimates.)



The bulb packaging conclusion is based on only 10 data points. One of these points, 7, is quite different from the others. The data suggest that about 10% of bulbs have very short lifetimes. The data, however, are not inconsistent with the possibility that an even larger proportion, say 20%, have very short lifetimes. If this were the case — we don't know for sure whether it is or isn't — our conclusion would be flawed: rather than only one percent of packages having a total bulb lifetime of less than 4500 hours, approximately 3% of packages would fail to satisfy the guarantee. This might eat up all the company's profits. The issue here is whether 10 data points is enough to give us a reliable idea of what is the distribution of bulb lifetimes. See Section ?? for a discussion of how to detect when there is too little data.



EXAMPLE 4: HOUSING PRICES: CONFIDENCE IN THE MEDIAN

A magazine is preparing its annual report on “The 10 Best Cities in America.” One of the criteria for inclusion in the list is the cost of housing. The report's author estimates the cost of housing by going through the records in the city courthouse and tabulating the price of each house sold in the last month. All such sales are recorded, whether the house sold for \$50,000 or for \$5,000,000. In the magazine article they could present the average sales price of a house. This would be misleading to most of their readers who would never think of buying a house that costs \$5,000,000. The magazine decides to report the median house price.

The magazine reporter collects the following prices for Cleveland, Ohio in July 1998. (The prices are in \$1000's.)

```
>> prices=[110 63 78 137 490 82 97 119 852 105 110 134 212 98];
```

For the sake of comparison, we compute the 95% confidence intervals for both the mean and median:

```
Ntrials=1000;
Ndata = length(prices);
mn = starttally; % tally of the mean
md = starttally; % tally of the median
for trials = 1:Ntrials
    samp = sample(Ndata, prices);
    a = mean(samp);
    tally a mn;
    a = median(samp);
    tally a md;
end
disp('95% conf. interval on mean:');
percentile(mn, [.025, .975]);
disp('95% conf. interval on median:');
percentile(md, [.025, .975]);
```

The resulting confidence interval on the median is 97 to 137. For the mean, the confidence interval is much larger: 102 to 316.

EXAMPLE 5: CONFIDENCE IN THE DISTANCE

You are involved in planning emergency services for a rural district in northern Minnesota. There is one long road that runs through the district and you are trying to evaluate a plan to build an emergency medicine clinic at a particular site on the road. One criterion in building the clinic is to provide fast access to emergency services. You will measure this by looking at the median distance from the clinic to the patients. You have information about where on the road people live but, since many of the emergencies arise from hunting, snowmobiling and farming accidents, you decide to use data collected over the past year about the place where each emergency patient originated: where the accident occurred, or where the patient fell ill. For simplicity, we'll look at a small part of the data, which gives the location of origin of the patient measured according to the mile-markers on the district road.

```
>> location = [78 145 132 94 67 44 93 78 120 34 78 53 119 95];
```

The new clinic is to be located at mile 117. For example, the accident

victim at mile 93 would need to travel `abs(93-117)` miles to get to the clinic. (`abs` is the absolute value function; it makes a negative number positive. For instance `abs(-7)` gives 7.) The median distance from the accidents to the planned clinic is

```
>> median( abs( location - 117 ) ) => ans: 33.5
```

A real-estate developer, with land to sell at mile 75, points out that the median distance to his land is shorter

```
>> median( abs( location - 75 ) ) => ans: 21
```

This is a situation where it is helpful to know confidence intervals. The 95% confidence intervals on the median distance can be computed using bootstrapping, as in the file `clinic.m`:

```
z = starttally;
for trial = 1:1000
    a = median( abs( location - 117 ) );
    score a z;
end
percentile(a,[.025, .975])
```

We run the commands

```
>> clinic
ans: 22 50
```

The resulting confidence interval is surprisingly wide: 22 to 50 miles. This information about the uncertainty in our estimate can be useful. By repeating the above analysis for hypothetical locations at places other than mile 117, one can see that the exact location of the clinic doesn't affect things much: the clinic could be located just about anywhere between miles 60 to 120.

EXAMPLE 6: CONFIDENCE IN CORRELATIONS: STORKS AND BABIES

We are interested in the possibility that storks deliver babies. We have some data on the population of storks and the yearly number of human births (in millions) from Germany.²

```
years = [1965 1971 1974 1977 1978 1979 1980];
storks = [1920 1700 1090 990 1030 995 930];
babies = [1.04 0.78 0.62 0.58 0.57 0.58 0.62];
```

²These data were published in [?].

When we graph `storks` and `babies` versus `years` we see that both populations are decreasing. Another way to look at the situation is to plot `babies` versus `storks`. We see (Fig. ??) that when `storks` are up, so are `babies`. When `storks` are down, the number of births falls.

One way to quantify the association between `storks` and `babies` is through the r^2 statistic. When $r^2 = 1$, the association is perfect. When $r^2 = 0$, no association is detected. In our case, we will compute r^2 for a linear model: do the data in Fig. ??(right panel) lie on a straight line? The closer r^2 is to 1, the closer the data lie to a straight line. If r^2 equals 1, the data are exactly on a straight line. If r^2 equals zero, a line doesn't describe the data at all. (It is also possible to compute r^2 for other sorts of models. See, e.g., `REGRESS`.)

Figure 1:
The population of storks and the number of
babies born per year in West Germany.

```
In MATLAB,  $r^2$  is easy to compute for the straight line model
>> rsq = corr(storks, babies).^2
ans:    0.8921
```

We now want to compute the confidence intervals of this estimate of r^2 . We can do this with resampling. Keep in mind, though, that r^2 is calculated from two data sets. When we resample the data, we do not want to pick a number at random from `babies` and to pick another number completely at random from `storks`. Instead, we want to pick a year at random and sample *both* the babies and storks that year: to make a paired sample. (In Example ?? we'll see a situation where we do want to pick the babies and stork samples independently.)

In MATLAB, we can select a value from a data set by *indexing*. For example, to reference the 5th value from `babies`, we use the command `babies(5)`. To implement the resampling of the 7 data points we will resample the numbers 1 to 7 and use these as indices to sample the data themselves. Here is an excerpt from the file `storksbabies.m`:

```
Ndata = length(years); % how many data points are there?
z = starttally;
for trials = 1:1000 % Do 1000 trials
    inds = sample( Ndata, 1:Ndata );
    stks = storks(inds);
    babs = babies(inds);
    r = corr(stks, babs);
    rsq = r.*r;
    tally rsq z;
```

```

end
histogram(z);
disp('The 95 percent confidence interval is:');
[percentile(z,.025), percentile(z,.975)]

```

Run the script:

```

>> storksbabies
    The 95 percent confidence interval is:
    0.041 0.998

```

We see that the 95% confidence interval is quite broad, reaching from 0.041 to 0.998, and does come close to zero. But there is more to the story than just this. Although most of the resamples have r^2 close to 1, a few are close to 0. If you modify the program to print out `yrs` for those resamples where r^2 is close to zero, you will discover that they are ones where 1965 and 1971 have been omitted.³ There is definitely an association between storks and babies in this data, but it depends strongly on just a few of the data points. In those resamples where the 1965 and 1971 data are omitted, the perceived association disappears.

The above estimation of the confidence interval of r^2 makes an important assumption: the pairs are independent of one another. This assumption is important because if it is violated the results of the resampling computation do not reflect the real confidence interval. It is hard in this example to know whether this assumption is justified.

To illustrate the consequences of violating the assumption of independence, we will make up some data where we know independence doesn't hold. Imagine that we sent out 2 teams in each of the years to measure births and the stork population. If these teams are accurate, they will get about the same data as each other. The two data points in any one year — one from each team — are obviously not independent. If we wrongly assume that they are independent and compute the confidence interval through resampling, we will get a very narrow confidence interval: 0.813 to 0.9803.⁴

³To see this for yourself, modify the program to add the following lines just after `rsq = r.*r`,

```

if rsq < .1
    years(inds)
end

```

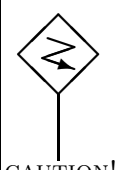
This will print out the years for those cases where the correlation is weak.

⁴You can check this out by making new data sets, as in `babies = concat(babies, babies)`; and likewise for `storks` and `years`.

The assumption of independent samples can often be justified when data are collected from randomly selected cases. For example, in studying the relationship between height and weight in people, you might pick people at random and measure their heights and weights. Although height and weight will presumably be correlated, the weights of the subjects are independent from one subject to the next as are their heights.

The babies and storks data were not selected randomly: measurements were made sequentially. This renders the assumption of independent samples suspect.

See the documentation for `REGRESSION` for an example of how to find the best fitting line relating `storks` and `babies`.



CAUTION!

Those readers with a passing familiarity with the biology of reproduction may protest that there is good theoretical reason to deny the existence of a relationship between the stork population and the arrival of babies. But “relationship” is not the same thing as “correlation.” The finding of a correlation between the size of the stork population and human births does not imply a causal relationship. Correlation and association are the rightful objects of statistical analysis of data. Whether a detected correlation or association can be interpreted as a causal relationship depends on the context in which the data were collected. Generally, it is necessary to carry out an experimental intervention in order to reach conclusions about causality. Effective design of experiments involves a control group and random assignment of cases to be in the control or experimental group. None of this exists in the stork and baby data.

In the preceding examples, we have assessed the extent of sampling variability by mimicking the process of sampling. We resample from our data to reproduce the sorts of sampling variability that come from sampling in the real world. The implicit assumption is that the real world is quite like our data.

Another perspective is this: our data are a random sample of the world and they may not precisely reflect the world. We ask, how much different might the world be from our data while still giving us a reasonable likelihood that our data are really a random sample of the world? This approach is closely related to hypothesis testing, which is covered in Chapter ???. Below, we give a simple example of how resampling can be used for this approach to confidence intervals.

EXAMPLE 7: HOW SAFE IS THE SPACE SHUTTLE?

The space shuttle Challenger flew 24 missions without a serious accident. It was deemed safe enough that an elementary-school teacher, Christa McAuliffe, was sent on the 25th mission. In a tragic accident, televised live to millions of school children, the shuttle exploded, killing all on board.

Put yourself in the place of a mission planner trying to decide how safe is the shuttle. You have data from the first 24 missions in order to plan for the 25th. Writing a safe trip as 0 and a disaster as 1, the data are:

```
>> flights = [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0];
```

Since there were no accidents, the accident rate is 0. We can try to put confidence intervals on this by resampling, but all of the resamples will consist only of 0s and so the resampled accident rate will always be zeros. Common sense tells us that the fact that there has never been an accident doesn't mean that there never will be one. How can we give a meaningful confidence interval to the accident rate?

One way to look at the confidence interval is to ask, "What range of accident rates are consistent with the observed data?" Clearly the lowest possible accident rate is 0 and this is consistent with the data. On the high side, the situation requires some analysis. If the "true" accident rate were 50%, it seems quite unlikely that we would sample a string of 24 consecutive flights without an accident. (This is the same probability as flipping a coin 24 times and never getting tails: $\frac{1}{8192}$. Possible, but not at all likely.) But what if the accident rate were 5%?

We can solve this problem using a simulation. We will hypothesize an accident rate and use a simulation to compute many sequences of 24 flights. Some of these sequences will have no accidents, and are therefore match the observed data. If the fraction of no-accident sequences in the simulation is more than 5% of the total number of sequences, we will declare that the hypothesized accident rate is consistent with the observed data.

We'll try out a number of hypothetical accident rates, and use the largest and smallest rates consistent to set the upper and lower bounds of the accident rate's 95% confidence interval. But, since the lowest possible accident rate is zero, and since this is obviously consistent with the observation of no accidents, half of our work is already done and we need only compute the upper bound of the confidence interval.

Since we will be carrying out the same computation for many different accident rates, this is a natural application for an m-file function that

takes the true accident rate as an argument. (See Step ?? of the tutorial in Appendix ??.) Actually, to make the program even more general, we will give as arguments the number of flights, and the number of accidents. This will allow you to use exactly the same program to set confidence intervals for any observed proportion.

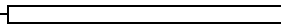
The following listing is the program `shuttle.m`. The first line of the file instructs MATLAB that `SHUTTLE` is a function that takes three arguments (`hypothrate`, `nflights`, and `naccidents`). Looking at the last line of the file, we see that the returned result of the function, held in the variable `res`, is a vector of two numbers: the first gives the probability of seeing the observed number of accidents or lower given the hypothetical true rate, the second is the probability of seeing the observed number of accidents or greater.

```
function res = shuttle(hypothrate, nflights, naccidents)
% shuttle(hypothrate, nflights, naccidents)
% compute the probability that <naccidents>
% or fewer
% will be seen in <nflights> flights if the
% hypothesized accident rate is <hypothrate>
% Returns two numbers:
% -- the probability of seeing naccidents or fewer
% -- the probability of seeing naccidents or greater.
data = [(1-hypothrate) 0; hypothrate 1];
z = starttally;
for trials = 1:5000
    % simulate nflights
    s = sample(nflights, data);
    % count how many crashes
    crash = count(s==1);
    tally crash z;
end
res = [proportion(z<=naccidents); proportion(z>=naccidents)]
```

Let's try a hypothetical accident rate of 5%.

```
>> shuttle(0.05) ⇒ ans: 0.286 1.0
```

The first number in the answer indicates that there is a likelihood of almost 30% that we would see no accidents in 24 flights if the underlying accident rate were 5%. The second number indicates what we already know: the probability of seeing no accidents or more is 100%. So, the 5% hypothetical rate is quite consistent with observing no accidents. Let's try a higher rate:



```
>> shuttle(0.10) ⇒ ans: 0.079 1.0
```

Now the likelihood of our data is 7.9%, just a little greater than 5%. Our data are consistent with this hypothetical rate of 10%, but just barely.

By experimenting with different accident rates, we can find the highest rate consistent with the data. We find, in this case, that an accident rate of 0.12 is consistent with the data at a level of 5% and we therefore can take the 95% confidence interval on the shuttle's pre-Challenger accident rate to be 0 to 0.12.

To illustrate a more general use of this technique to find confidence intervals on a proportion, where we need to find both the upper and lower bounds of the confidence interval, let's find the confidence interval on the accident rate when one accident is observed out of 40 flights. This gives an observed accident rate of $\frac{1}{40} = .025$, so we'll start by looking at hypothetical rates higher than this to set the upper bound of the confidence interval.

```
>> shuttle(.05, 40, 1) ⇒ ans: 0.4310 0.8820
```

There's a 43% likelihood of seeing 1 accident or fewer in 40 flights if the hypothetical accident rate is 0.05. So, we'll try a higher hypothetical rate.

```
>> shuttle(.10, 40, 1) ⇒ ans: 0.085 0.982
```

Not quite there...

```
>> shuttle(.11, 40, 1) ⇒ ans: 0.054 0.990
```

Still higher ...

```
>> shuttle(.12, 40, 1) ⇒ ans: 0.034 0.994
```

Too high. So we'll take the upper bound on the confidence interval to be an accident rate of 0.11. This may be surprising: the upper bound of the confidence interval is smaller for 1 accident in 40 flights than it was for 0 accidents in 25 flights.

Now to look for the lower bound on the confidence interval. The strategy is similar to what we have been doing, but now we try a hypothetical accident rate lower than that observed, and lower it until the calculated probability is 5% of observing 1 accident or *greater*. We'll start with a hypothetical accident rate of 1%.

```
>> shuttle(.01, 40, 1) ⇒ ans: 0.9366 0.3356
```

The likelihood of seeing 1 accident or more in 40 flights is 33%. Try an even lower rate:

```
>> shuttle(.001, 40, 1) ⇒ ans: 0.9990 0.0326
```

This is a bit too low. Back up a bit:

```
>> shuttle(.002, 40, 1) ⇒ ans: 0.9970 0.0748
```

So, the lower bound on the confidence interval is somewhere between 0.001 and 0.002. We'll estimate it to be half way between, at about

0.0015.

In summary, the confidence interval we found for the accident rate when we observe 1 crash out of 40 flights is 0.0015 to 0.11. Since we computed 5% likelihoods at *each end* of the confidence interval, this is a 90% confidence interval.

You may legitimately question whether the seeming precision of the number 0.0015 is justified. The question of how precise are the answers produced by resampling is addressed in Section ??.

In Example ?? we examine another way of looking at the plausible range of accident rates, and how our knowledge of the shuttle's accident rate changes as each flight brings in new evidence.

