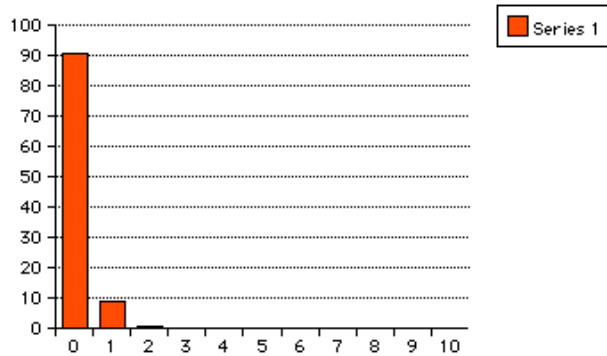


Say you poll 1000 people and ask them "Do you prefer chocolate or vanilla ice cream?" If 300 say vanilla and 700 say chocolate, then  $p$  is  $300/1000$  and  $1-p$  is  $700/1000$ .

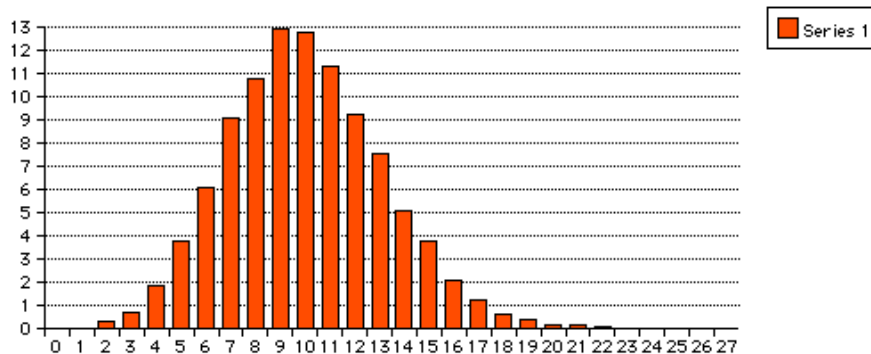
But how much faith should you have that  $p = 300/1000 = 30\%$  is a good measure of the population proportion (i.e., the proportion of the whole population that favors vanilla)?

You need three things in order to have faith that  $p = 30\%$  is a good measure of not just the 1000 people you polled but of all 200 million adult Americans.

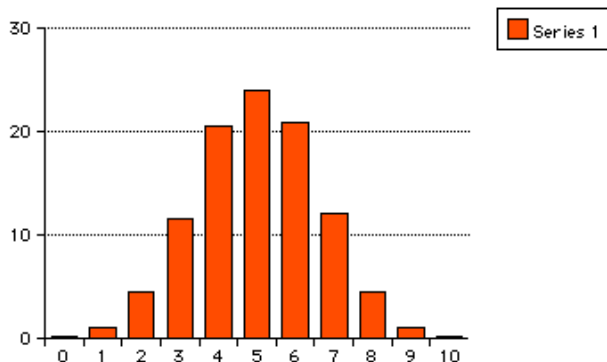
1. The sample needs to have been chosen randomly.
2. The sample size ( $n=1000$  in this case) needs to be big enough that the sample distribution is bell shaped. If  $p$  is small (close to 0) or big (close to 100%, in which case  $1-p$  is small) you'll need a much bigger sample size for the distribution to be bell shaped. For example, if  $p = 1\%$ , here's what the distribution looks like if  $n = 10$ :



Note that this is not bell shaped at all. But when  $n = 1000$ , the distribution is bell shaped:



It turns out that if  $p=50\%$ , then a sample size of  $n=10$  is already big enough to get a bell shaped curve, as you can see:



Thus for  $p = 1\%$  you need  $n$  to be at least 1000 or so to get a bell shaped curve, but if  $p = 50\%$ , you can get by with  $n = 10$ . However, there is one more consideration.

3. The margin of error shouldn't be too big. A big margin of error means that there's a good chance the population proportion is not that close to the 30% measurement you got from the sample, so you shouldn't have much faith that 30% is a good measure of the population proportion. As an example, say the margin of error is  $\pm 25\%$ ; then the population proportion could be anywhere from 5% to 55% (i.e., vanilla might even be preferred over chocolate).

Usually, pollsters want the margin of error to be around  $\pm 2\%$ . The bigger the sample size  $n$ , the smaller the margin of error. So even if  $n$  is big enough that the distribution is bell shaped it might need to be even bigger to get the margin of error as small as you want it. For example, if in a presidential poll one candidate is at 48% and the other at 52%, a  $\pm 5\%$  margin of error is too big to be confident of who's ahead. In this case you might want a margin or error of only  $\pm 1\%$ , but that means you'll need a bigger sample size.