Re: quality control

Source: http://sci.tech−archive.net/Archive/sci.stat.math/2006−11/msg00387.html

---

- *From*: David Winsemius <doe_snot@xxxxxxxxxxx>
- *Date*: Tue, 14 Nov 2006 10:40:53 −0600

"Frank" <deps_bear@xxxxxxxxx> top−posted:

> Richard Ulrich wrote:
>
>> "Frank" <deps_bear@xxxxxxxxx> wrote:
>>
>>> If I know a product fails .01% of the time and I have 1500 items I'm
>>> running through a process. How many items do I need to check with,
>>> say, 99% confidence that all the items are built correctly.
>>
>> How many failures do you expect? Almost always, zero.
>> This is dealing with exact probabilities. For a higher failure
>> rate, you might want to look at the p of success, and raise
>> to a power, e.g., (.9999)^n . For the tiny p of 0.01%,
>> the figuring can be pretty much additive
>>
>> You want to have only so many items *unchecked* that there
>> will be, on the average, only 1 bad item in 100 samplings −−
>> so that 99 times out 100, there will be none.
>>
>> You expect 1 failure in 10,000. One hundred samplings
>> that each fail to test 100 items will meet that condition.
>> So you need to check 1400 of each 1500.
>
> Richard, thanks for your reply, however checking 1400 out of the 1500
> items does not seem like a cost−effective approach. I may not have
> phrased the question correctly and also did some additional research.
>
> I'm thinking with the binary outcome of fail/not fail, then this is
> more of just a sampling question and knowing there are 1500 items is
> irrelevant. So, if I suspect that 1% of the items fail, and put a

> confidence interval around this estimate with a width of plus/minus 10% (that is, the interval would go from .09% to 1.1%), then I would need to sample 38 with 95% confidence. That is, I can say that my sample of 38 will contain 1% of the proportion that fails 95% of the time.

You cannot force the variance (or standard deviation) of a binomial problem by just saying "it will be +/− 10%". You cannot specify the width of a confidence interval just by wishful thinking. We will leave aside, of course, the case of the midwestern state legislator who introduced a bill to make pi some rational number.

I also cannot quite figure out where the 38 came from. What you wrote sounds like something cobbled together to please a boss who told you to design a system to do the impossible. "Cost−effective" analysis requires that you specify the costs of each error which are then compared to the cost of doing the inspections. You appear to be scrambling to satisfy a manager that knows nothing about statistics. Just tell her that the problem (as you posed it, anyway) has no solution. Capitalize the message if necessary.

If you sampled 38 and got zero defects, your results would be:
(output of Frank Harrell's implementation in Hmisc):

> binconf(0,38) (the default value of alpha is 0.05)

PointEst Lower Upper
Exact 0 0 0.09251276

So you would have 95% confidence that the defect rate was no more than 9.3%.

Q: What is the exact 99% confidence interval for the binomial proportion around an observed value of zero in a sample size of 1500?

A:> binconf(0,1500,method="all",alpha=0.01)
PointEst Lower Upper
Exact 0 0 0.003525981
Wilson 0 0 0.004403785
Asymptotic 0 0 0.000000000

So in your orignal problem you sampled the entire batch and you ended up 99% "confident" that the fault rate is no more than 0.0035 (or so). Even examining the entire lot you are still a long way from your specified confidence that the true failure rate be less than 0.0001.

−−
David Winsemius

.