THE PILL PROBLEM

Suppose you have invented a pill to cure a certain disease. Sometimes the pill works and the patient is cured; sometimes it doesn't work.

We want to make that last statement more quantitative. Let p be the fraction of the cases in which in the pill works. We want to determine p, as accurately as we can.

So we plan to pick a random sample of 100 patients and give the pill to each. Then we will count the number S of successful cures, and use $S/100$ as our estimate of p. For example, if we find that the pill works for 41 of the patients, then we will figure that p must be about 0.41. But we realize that $S/100$ may not be exactly right; a larger study might give better accuracy.

Will we at least be able to say that p is within 0.1 of $S/100$? That is, can we say that $|(S/100) - p| \leq 0.1$? Not with complete certainty. Maybe the real value of p is very low and we just get lucky with our 100 patients. Or maybe the real value of p is very high but our luck is rotten on the day of the experiment.

Ok then, can we at least say that we are 95% sure that p is within 0.1 of $S/100$? That is, we hope that our study supports the statement:

$$P(|(S/100) - p| \leq 0.1) = 0.95 \text{ or more.}$$

The *pill problem* is to determine when such a statement is justified.

To generalize the situation, suppose that we give the pill to n patients (instead of 100), and that S/n is the fraction of the cases in which the pills work. We want to say, with some high degree of certainty, that the observed value S/n is within some allowed error ε of the true value p. (In the above, ε was 0.1.) That is, we want to say that $|(S/n) - p| \leq \varepsilon$ with some high degree of certainty. The pill problem is to determine when a statement such as

$$P(|(S/n) - p| \leq \varepsilon) = 0.95 \text{ or more}$$

or

$$P(|(S/n) - p| \leq \varepsilon) = 0.90 \text{ or more}$$

or

$$P(|(S/n) - p| \leq \varepsilon) = 0.99 \text{ or more}$$

will be justified.

There are three interrelated quantities here:

- the sample size n
- the allowable error ε
- the degree of certainty

Typically we know two of the three, and we want to determine the third. Example 8 on page 895 is the pill problem where $\varepsilon = 0.01$, the degree of certainty is 95%, and n is to be calculated. Problems 24 and 26 on page 897 are similar. In other problems, the unknown might be ε or the degree of certainty.

We proceed to apply the central limit theorem to the pill problem. Define

$$X_i = \begin{cases} 1 & \text{if the pill works for the } i\text{th patient} \\ 0 & \text{otherwise.} \end{cases}$$

Then the X_i's are i.i.d. with mean p (which is unknown). (We can assume that in a random sample the results for each patient are independent of the other patients.) Their variance is $p(1-p)$, also unknown to us, but we can use (as on page 890) the worst case: The variance for a binary variable is a maximum at $p = 1/2$, in which case the variance is $1/4$ and the standard deviation is $1/2$. (If the true variance is less than this, so much the better; S/n is then likely to be even closer to p. In §12.7 we will find another possible approach that can be used after we have the data in hand.) Then the number S of successes is equal to $X_1 + \cdots + X_n$, and the fraction of the cases in which the pill works is

$$\frac{S}{n} = \overline{X} = \frac{X_1 + \cdots + X_n}{n}$$

which has mean p (unknown) and standard deviation (in the worst case) $1/(2\sqrt{n})$. By the central limit theorem, $\overline{X}$ is approximately normally distributed:

$$\begin{aligned} P(|\overline{X} - p| \le \varepsilon) &= P\left(\left|\frac{\overline{X} - p}{1/(2\sqrt{n})}\right| \le \frac{\varepsilon}{1/(2\sqrt{n})}\right) \\ &= P(|Z| \le 2\varepsilon\sqrt{n}) \end{aligned}$$

where Z has the standard normal distribution (and the equality holds approximately, and in the worst case).

We need to examine the "z-value":

$$z = 2\varepsilon\sqrt{n}.$$

If $z = 1.96$ or more, than because $P(|Z| \le 1.96) = 0.95$, we can be at least 95% sure that $\overline{X}$ is within ε of the unknown mean p.

degree of certainty	z value
90%	1.64
95%	1.96
98%	2.33
99%	2.58

To return to the original example, we had $n = 100$ and $\varepsilon = 0.1$. So $z = 2\varepsilon\sqrt{n} = 2$ and the degree of certainty is $P(|Z| \le 2)$, which is 0.9544, according to the table on page 919. (Details: The table gives $\Phi(2) = 0.9772$. How did I get 0.9544 from this? Be sure you can do this, too.) So yes, we can indeed assert with 95% certainty that p is within 0.1 of $S/100$.

In Example 8 on page 895, we need 95% certainty, so we need $z = 1.96$ or better. Since $\varepsilon = 0.01$, we need $\sqrt{n}$ to be $1.96/0.02 = 98$ or better. The conclusion is that we need a sample size of 9,604.

In Example 2 on page 890, exactly the same problem is discussed, but with use of Chebyshev's inequality instead of the central limit theorem. And in that example, a sample size of 50,000 is called for. The moral here should not be ignored.

—H. B. Enderton