The Institute of Internal Auditors Australia

Dear Q&A

What is best practice for determining sample sizes and selecting a sample to identify fraud?

Answer

Perhaps the best way to approach this is to ask yourself "what is the organisation's risk appetite?" What would be a tolerable error in the population? In the case of fraud or corruption, your organisation clearly has a target of zero but realistically this is not possible, so at what level does the organisation want to detect it. The organisation might for example want to assure itself that fewer than 1% of transactions contain an error. Noting of course, that not all errors are fraud – but in this context all frauds are errors. It might be that your organisation wants to establish that fewer than 1 in 1,000 transactions are fraudulent – in which case the sample sizes will be very large.

You will need to establish whether you are using sampling to find error / fraud or are using sampling to estimate the level of error / fraud. The approaches are quite different.

If the purpose is to find error / fraud, simple sampling is not necessarily the best approach. If there is some way other than sampling to identify erroneous transactions, it might be best to use this first to concentrate the errors into reduced population for examination.

Consider whether data analysis techniques can be used to identify those transactions that might be erroneous. You could for example use Benford's Law analysis https://www.journalofaccountancy.com/issues/2017/apr/excel-and-benfords-law-to-detect-fraud.html to identify transactions that do not fit the normal pattern of numbers.

Or you could use frequency analysis to look for transactions that are too common for their value – most commonly this will be transactions that are unusually large or moderately sized transactions that occur too frequently. It is not uncommon for transactions valued slightly below a delegation threshold to be more common that might otherwise be expected. This may reduce the number of items to be examined to a reasonable number, but sampling might still be necessary.

Having concentrated the potential errors, the likelihood of detecting fraud is increased, but unfortunately you will not be able readily to project your results across the entire population of transactions. Your result will apply to the concentrated set of transactions you are using as the base for sampling. This concentration may mean you can afford to establish that fewer than 2% of these transactions are fraudulent or otherwise improper. This approach helps you to find fraud / error, rather than to estimate what the level of fraud / error is.

The discussion that follows is more targeted at estimating the level of error / fraud rather than at finding it. Clearly there is some chance that error / fraud will be detected as a by-product, but this is not the most efficient way of going about it.

Having what you are looking for and the proportion of the population that you can afford to fall into that category, you then need to decide what your appetite for error as an auditor is. If you are sampling, you can never be certain because you have not looked at everything. There is error inherent in sampling. In fact, even if you look at everything there is a possibility of making an error in testing, so certainty is hard to achieve. We express the auditor's risk appetite in sampling using two measures – precision and confidence.

› Precision defines an interval within which the auditor is attempting to contain their estimate of the result. This is usually expressed as being ±A in relation to the actual estimate. The more precise you want to be, the larger will be the sample. Often the precision is expressed as we want to establish that the rate of error / fraud is less than x%. In this case the calculations are simpler, but the sample size is still driven by the target error rate.
› Confidence is the probability the sample is truly representative of the underlying population. The more confident you want to be, the larger will be the population.

In the IIA-Australia White Paper 'Internal Audit Sampling', we discuss the sampling approach for estimating the actual value of error / fraud rate – the occurrence of error / fraud is an attribute of a transaction. In the formula for attributes sampling:

› Use the established organisational risk appetite as p – this should be expressed as a decimal, so 1% occurs as 0.01
› Use the internal auditors desired precision (also a decimal number) as A
› Use the desired confidence to obtain Z – we commonly use:
  › 95% confidence – Z = 1.96
  › 96% confidence – Z = 2.00
  › 98% confidence – Z = 2.33

If we are simply trying to establish that the error does not exceed a particular value (we call this p), then the necessary sample size is:

NO ERRORS are found, then the error rate is less than p

$$n = \frac{\log(1-confidence)}{\log(1-p)}$$

If errors are found, the results from a random sample can be projected across the whole population. The IIA-Australia White Paper discusses how to do this.

The Institute of Internal Auditors Australia