Audit Sampling

for Microsoft Excel

Reproducible statistical selection for tests of detail
and tests of controls

Version 1.0 | 02 November 2026

Monetary unit sampling (MUS) and random sampling
AICPA reliability-factor and exact binomial methodology

Contents

1	What this tool does	3
2	Setting it up on a laptop	3
3	Drawing a sample	5
4	How the sample size is calculated	6
4.1	Monetary unit sampling	6
4.2	The reliability factor	7
4.3	The expansion factor	7
4.4	A worked example	8
4.5	The certainty stratum	8
4.6	Random sampling: attribute sample size	9
4.7	What is AICPA, and what is a choice made here	10
5	Reproducibility	11
6	The workpaper	12
7	Evaluating the sample	13
8	Verification, limitations and review points	14
9	If something goes wrong	15

1 What this tool does

The add-in draws a statistical sample from a range you select in Excel and writes a complete workpaper to a new worksheet: the parameters used, how the sample size was derived, the items selected, live evaluation formulas, every row it excluded and why, and the date and time at the foot of the sheet.

It offers two methods.

- **Monetary unit sampling (MUS)** - probability-proportional-to-size selection for tests of detail. Every monetary unit in the population has an equal chance of selection, so larger balances are more likely to be picked, and any item at or above the sampling interval is examined in full.
- **Random sampling** - equal-probability selection for tests of controls, and for any test where value should not drive the chance of selection. Simple random, systematic or stratified, with the sample size entered directly or derived from a tolerable rate of deviation.

Everything happens on the laptop. The workbook is read in place, the sample is computed locally, and no client data is transmitted anywhere.

> **Haphazard selection is deliberately not offered**
>
> Picking items by eye cannot be reproduced and cannot be shown to be unbiased. Offering it beside these methods would make it look like a statistical sample.

2 Setting it up on a laptop

The add-in runs entirely from the laptop. Everything it needs is in the folder you were given: **there is nothing to install, no administrator rights are needed, and no internet connection is required.**

Copy the folder somewhere local - not a network drive - and double-click:

```
Start-AuditSampling.bat
```

The first run creates an HTTPS certificate for this laptop and registers the add-in with Excel. Windows may ask you to confirm the certificate; choose Yes. It takes a few seconds.

Every time after that, run the same file. It starts the local server in a second window and leaves it running.

- Leave the server window open while you work. Closing it stops the add-in loading.
- If Excel was already open, close and reopen it - Excel only looks for add-ins at startup.
- The add-in appears on the **Audit Sampling** tab of the ribbon.

> **Why there is a server window at all**
>
> An Office add-in is a web page: Excel fetches it from a web address rather than reading it off the disk. Because everything here runs locally, that address is served by a small program on the laptop itself - the second window. It is reachable only from this machine, and it serves only the add-in's own pages. Your workbook is read by Excel, never by the server, and no data leaves the laptop.

What is in the folder:

Item	What it is
Start-AuditSampling.bat	Sets up and starts. Run it every time.
Remove-AuditSampling.bat	Unregisters the add-in and deletes the certificate.
node\	The runtime the local server needs. Nothing is installed on the laptop.
dist\	The add-in itself, already built.
manifest.xml	What tells Excel about the add-in.
tools\	The local server and the certificate script.
docs\	This document.

The certificate is created on the laptop, not shipped with the folder

Office refuses to load an add-in over an unencrypted connection, even from the same machine, so a certificate for localhost is needed. Each laptop generates its own on first run, and it is trusted only for the user who ran the setup. No private key is shared between laptops, and none travels in the distribution folder.

To remove it: run `Remove-AuditSampling.bat`. That unregisters the add-in from Excel and deletes the certificate. Workpapers already produced are ordinary worksheets and are untouched.

3 Drawing a sample

1. Select the population. Highlight the whole range, including the header row if there is one. The task pane shows the address, the row and column counts, whether it found a header, and which column it thinks holds the amount.

2. Set the population. Press *Use this selection as the population*. The range is pinned, and from then on the cursor can go anywhere - including onto the workpaper the add-in writes - without losing it. The pinned range is the one that gets sampled; the pane always shows which range that is.

3. Choose a method on the Monetary unit or Random tab, and press *Set up sample*.

4. Fill in the dialog. It shows a live plan as you type - reliability factor, sampling interval, resulting sample size, the certainty threshold - so you approve the numbers before the sample is drawn. The identification block (entity, period, account or assertion, preparer, objective) is required: the sample cannot be drawn until it is complete.

5. Draw. The workpaper appears on a new sheet.

> **The seed**
>
> Every run needs a random seed, and it is recorded on the workpaper. The same seed over the same population always produces the same sample. A seed is suggested for you, and free text such as FY25-AR-TOD-1 works as well as a number - and reads better on a workpaper.

Each run is listed in the task pane with its seed and population fingerprint. That history is deliberate: repeated runs over the same population with different seeds are visible to anyone reviewing the file.

4 How the sample size is calculated

This section sets out the methodology in full, so that it can be reviewed, challenged and tied back to the guidance your firm applies.

4.1 Monetary unit sampling

Sample size follows the reliability-factor method of the AICPA Audit Guide *Audit Sampling*:

$$n = \frac{\text{book value} \times \text{reliability factor}}{\text{tolerable} - (\text{expected} \times \text{expansion factor})}$$

$$\text{interval} = \frac{\text{tolerable} - (\text{expected} \times \text{expansion factor})}{\text{reliability factor}}$$

where *book value* is the total recorded amount of the sampling frame, *tolerable* is tolerable misstatement, and *expected* is the misstatement you expect to find.

The tool derives the **interval first**. That matters: the interval depends only on the assurance parameters and not on the population value, so the certainty stratum can be separated without the circular reasoning of removing items, recomputing the interval, and finding that a different set of items now qualifies.

All selection arithmetic then runs in whole minor units - cents, fils, pence. Cumulative totals over a long ledger accumulate floating-point error, and an ambiguous boundary in the cumulative walk would make the sample impossible to reproduce. Monetary unit *m* selects the first item whose cumulative total is strictly greater than *m*: half-open ranges, no ties.

Three selection methods are available once the size is fixed:

Method	How items are selected
Fixed interval	One random start within the first interval, then every interval-th monetary unit after it.
Cell	The population is cut into interval-sized cells and one unit is drawn at random inside each. The short final ce
Random	The required number of distinct monetary units is drawn at random across the whole cumulative value.

One item, more than one hit

Under cell and random selection an item can be selected by more than one monetary unit. Each hit is a separate sampling unit and is projected separately - the workpaper lists them in their own block and shows a hit count against each item. Projecting such an item only once understates the upper misstatement limit by roughly 8% under cell selection and 26% under unrestricted random selection.

4.2 The reliability factor

Reliability factors are Poisson factors. RF(k, c) is the value of lambda for which the probability of observing at most k misstatements equals 1 - c; equivalently it is half the chi-square quantile with 2k + 2 degrees of freedom.

They are **computed, not looked up**. The add-in is therefore not restricted to the confidence levels that happen to appear in a printed table, and the workpaper shows the derivation as a live Excel formula - `=CHISQ.INV(confidence, 2*k+2)/2` - so a reviewer can click the cell and see where the number came from.

The computed values reproduce the published table; the published figures are rounded to two decimals, the tool carries full precision:

Misstatements k	Computed (95%)	Published (95%)
0	2.996	3.00
1	4.744	4.75
2	6.296	6.30
3	7.754	7.76
4	9.154	9.16

The same agreement holds at 99%, 90%, 85%, 80%, 75%, 70% and 50%, and is locked down by the test suite.

4.3 The expansion factor

The expansion factor inflates the allowance for expected misstatement. These are AICPA table constants, used verbatim:

Confidence	99%	95%	90%	85%	80%	75%	70%	50%
Expansion factor	1.90	1.60	1.50	1.40	1.30	1.25	1.20	1.00

If you choose a confidence level that is not in this table, the factor is interpolated linearly between the two nearest levels and **the workpaper says so**. Interpolation is a choice made here, not AICPA guidance - see section 4.7.

4.4 A worked example

A receivables ledger of 1,000,000, tolerable misstatement 50,000, no expected misstatement, at 95% confidence:

```
reliability factor  RF(0, 95%)  =  2.9957
expansion factor                =  1.60   (not used: expected = 0)

interval  =  (50,000 - 0) / 2.9957      =  16,690.41
n         =  1,000,000 x 2.9957 / 50,000 =  59.9  ->  60 items
```

The sample size is always rounded up. Had expected misstatement been 10,000, the denominator would become 50,000 - (10,000 x 1.6) = 34,000, giving an interval of 11,349.53 and a sample of 89 items - the cost of expecting to find something.

4.5 The certainty stratum

Any item at or above the sampling interval is certain to be hit by the cumulative walk, so it is separated out and examined in full. These items are reported separately on the workpaper, and their misstatement is **known, never projected** - projecting a 100%-examined item would invent sampling risk that does not exist. On the selected-items table their Projected misstatement column shows that actual misstatement as it is - never taint x interval - and the evaluation counts it once, as known misstatement. Their taints never enter the incremental allowance ladder.

You can set a different threshold in the dialog. If you set it above the interval, items between the two remain in the projectable stratum and can be selected more than once; the workpaper warns you. If you set it below the interval, more items are examined in full than the interval requires - more work, not less assurance - and it says that too.

When the parameters ask for less than one item

If the interval exceeds the whole population value, a fixed-interval walk can legitimately select nothing at all. An empty workpaper is never a useful answer, so the tool refuses and explains that the tolerable misstatement is too large relative to the population.

4.6 Random sampling: attribute sample size

For tests of controls the size can be entered directly, or derived from the tolerable rate of deviation. The derivation uses the **exact binomial** distribution, not the Poisson approximation: the sample size is the smallest n for which observing the expected number of deviations still supports the tolerable rate at the chosen confidence.

```
n  =  smallest n such that
      P(X <= expected deviations | n, tolerable rate)  <=  1 - confidence
```

The results reproduce the published AICPA attribute sample size table exactly, at a 5% risk of overreliance:

Tolerable rate	0 dev.	1 dev.	2 dev.	3 dev.	4 dev.
5%	59	93	124	153	181
10%	29	46	61	76	89
20%	14	22	30	-	-

Computed values match the published table in every cell tested.

Note the input. The dialog asks for the expected *number of deviations*, not an expected rate. The published table is indexed by rate, and its mapping from a rate to a number of deviations is not a simple formula; asking for the number directly removes the ambiguity. Enter 1 and you get 93 at a 5% tolerable rate, exactly as the table's 0.5% row does.

Selection is then simple random (unbiased partial Fisher-Yates, without replacement), systematic with a random start, or stratified. Stratified samples are allocated proportionally, equally, or by Neyman - stratum size times standard deviation - with largest-remainder rounding so the allocations sum to exactly the sample size and no stratum is allocated more items than it holds.

A population with no monetary column at all - a despatch log, say - can be sampled by ticking *No amount column*, which weights every populated row equally.

4.7 What is AICPA, and what is a choice made here

Worth being explicit about, because a reviewer will ask.

Element	Basis
MUS sample size and interval	AICPA Audit Guide, Audit Sampling - reliability-factor method.
Reliability factors	Poisson, computed rather than looked up. Reproduces the published table.
Expansion factors	AICPA table constants, used verbatim.
Interpolating expansion factors between tabled confidence levels	A choice made here. AICPA publishes only the eight levels listed in 4.3. The workpaper fla
Attribute sample size	Exact binomial. Reproduces the published AICPA table.
Expected deviations as an integer input	A choice made here, to avoid guessing at the table's rate-to-deviations mapping.
Rounding the interval down to whole minor units	A choice made here. It makes the cumulative walk exact and reproducible, and errs towar
Upper misstatement limit	AICPA - basic precision, projected misstatement and the incremental allowance ladder.
Upper deviation rate	Exact one-sided binomial (Clopper-Pearson) bound.

Before it is used on a live file

This is the AICPA framework. If your engagements are performed under ISA 530, the mechanics are the same but the standard does not prescribe these tables. Whoever owns audit methodology at the firm should review this section and confirm the approach before the tool is relied on.

5 Reproducibility

A sample is only defensible if someone else can reproduce it. Three things make that true here.

The seed. Selection is driven by a seeded generator, never by the system random number generator. The same seed, over the same population, with the same parameters, always produces the same sample - on any machine, in any version of Excel. The seed is recorded on the workpaper.

The fingerprint. Each run records a hash of the sampling frame: the row numbers, the whole minor units the selection actually walks over, and the options that determine them. Same fingerprint means the same data; a different fingerprint means the population changed and the two samples are not comparable. It turns "I re-ran it and got the same sample" into a checkable claim.

An independent implementation. A separate Python implementation of the whole engine ships with the tool, in `reference/`. A reviewer can reproduce a selection *without running Excel and without trusting the add-in's code*:

```
python reference/sampling_reference.py mus population.csv \
    --amount-column 3 --seed FY25-AR-TOD-1 \
    --tolerable 50000 --expected 5000
```

The two implementations are checked against each other on every build: random number streams, all three MUS selection methods, the population fingerprint and the random samples must agree exactly.

Technical notes for anyone re-performing the work independently:

- The generator is sfc32, seeded through the xmur3 string hash. Both use only 32-bit integer operations, so they can be re-implemented exactly in Python, R or VBA.
- Every integer draw uses rejection sampling. A plain remainder operation biases the low values, which over many samples is a real statistical defect.
- Distinct items are drawn with a partial Fisher-Yates shuffle; the loop direction is fixed, because changing it would silently invalidate every sample issued previously.

6 The workpaper

Written to a new worksheet, top to bottom:

Section	Contents
Identification	Entity, period, account or assertion, preparer, objective, run ID.
Population	Source range, items in the frame, book value, exclusions, fingerprint.
Sampling parameters	Confidence, reliability factor (as a live formula), expansion factor, tolerable and expected misstatement
Sample size	Book value, certainty stratum, residual population, sampling units planned.
Matters for the reviewer	Warnings raised during the run.
Selected items	Sequence, source row, your identifier columns and the book amount; then the shaded testing columns
Sampling units	One row per monetary unit drawn, which is what the allowance ladder ranks. Each unit looks up its iter
Evaluation	Basic precision, known and projected misstatement, most likely misstatement, the incremental allowan
Exclusions	Every dropped row, with the reason.
Footer	Date and time prepared, seed, fingerprint, run ID, version.

The evaluation is built from **live Excel formulas**, not values pasted in once. Enter the audited amounts and the projected misstatement, the allowance and the upper limit all update. A reviewer can click any cell and see the derivation.

Text copied from the ledger is written as text: a customer name beginning with an equals sign cannot execute as a formula in the workpaper, and an invoice number such as 00123 keeps its leading zeros.

7 Evaluating the sample

Fill in the audited amount against each selected item. Everything below follows automatically.

Monetary unit sampling.

```
taint            = (book - audited) / book, capped at 100%
projected        = hits x taint x interval        (representative items)
known            = book - audited                 (certainty items;
                   their Projected cell shows the same figure)
basic precision  = RF(0) x interval
incremental      = sum over ranked taints of
                   (RF(k) - RF(k-1) - 1) x interval x taint(k)
-------------------------------------------------------------------
upper misstatement limit = basic precision + known
                         + projected + incremental
```

An item hit more than once is tested once, on one row, and projected once per hit: its Projected misstatement is hits x taint x interval, with the hit count shown beside it, and the ladder ranks its taint once per hit. Each sampling unit looks its item up by source row, so you can sort and filter the table without changing the evaluation.

The conclusion compares the upper misstatement limit with tolerable misstatement and states the result in words.

Understatements are not netted into that limit. They are counted and totalled separately, because netting them into an overstatement limit understates the risk. Evaluate them in their own right.

Tests of controls. Record Y against each deviation. The upper deviation rate is the exact one-sided binomial bound, written as a live formula (`=BETA.INV(confidence, k+1, n-k)`), and is compared with the tolerable rate to state whether the control can be relied on.

> **Zero sampling risk when nothing was sampled**
>
> If every item turned out to be a certainty item, the whole population was examined and there is no sampling risk to allow for. Basic precision is then zero, not the reliability factor times the interval, and the workpaper says why.

8 Verification, limitations and review points

What has been verified.

- 112 automated tests covering the statistics, the selection methods, the random number generator, the workpaper structure and every formula it emits.
- Reliability factors, expansion factors, attribute sample sizes and deviation-rate bounds checked against the published AICPA tables and, independently, against SciPy.
- The independent Python implementation reproduces the TypeScript engine exactly - random number streams, all three MUS methods, the fingerprint and the random samples.

Limitations you should know about.

- The methodology is the AICPA framework. Under ISA 530 the mechanics are the same but the tables are not prescribed - have methodology confirm the approach.
- Expansion factors are interpolated between tabled confidence levels; runs that rely on this are flagged on the workpaper.
- The population is whatever range you pin. The tool cannot know whether that range is complete - reconciling the population to the general ledger remains your work, and the workpaper leaves room to record it.
- Sample sizes above 100,000 items are refused as a guard against mistyped parameters.

What a reviewer will look for.

- Is the population complete, and does it reconcile?
- Were negatives, zeros and blanks excluded, and is that appropriate? Every exclusion is listed on the workpaper with its reason.
- Are the seed and method recorded? Both are, in the parameters block and again in the footer.
- Was the sample re-run after the data changed? Compare the fingerprint.
- Was the population sampled repeatedly with different seeds until a clean sample appeared? Every run in a session is listed in the task pane with its seed and fingerprint.

9 If something goes wrong

What you see	What it means
"Sorry, we can't load the add-in. Please make sure your network de…	The local server is not running. Despite what the message says, this is not about your interne
The Audit Sampling tab is missing	Excel only looks for add-ins at startup. Close Excel completely and reopen it. If it is still missin
The task pane opens but stays blank	The server is running but the pages are not reachable, or the certificate was refused. Close th
"No Node runtime found"	The folder is incomplete - it should contain a node sub-folder. Ask for a fresh copy of the distri
"Port 3000 is already in use"	A server is already running - possibly left over from an earlier session. Close any window titlec
The Set up sample button is greyed out	No population is pinned and the current selection is a single cell, or you are on a workpaper th
Draw sample stays disabled in the dialog	Something in the identification block is still empty, or the parameters are impossible. The plan
Rows appear in Exclusions that you expected to be sampled	The item could not be selected - blank, zero, negative, an error, or text. The reason is given against each

The distribution folder also contains README.md, which covers the same ground for a technical reader, and installer/README.md for deployment by IT.