

De pilot steekproef: een gids die je ook de verkeerde weg kan wijzen!

Bij sommige (schattings-)methoden voor het opzetten van steekproeven is vooraf een indicatie nodig van de spreiding tussen de (nog te vinden!) foutbedragen. Als je die spreiding weet, kan je uitrekenen hoeveel waarnemingen er aanvullend nodig zijn om de maximale fout aan de (uitvoerings-)materialiteit te laten voldoen. Om een indicatie van deze spreiding te krijgen wordt dan een pilot steekproef aangeraden. En, zoals bekend, als accountants niet weten wat ze moeten doen, doen ze het 25 keer.

Wat kan hier mis gaan?

Hoe kleiner de steekproef, des te groter de kans dat de uitkomst niet representatief is. Als de steekproef minder fouten oplevert dan het verwachte aantal op grond van de werkelijke foutkans, wordt de spreiding tussen de foutbedragen onderschat waardoor de pilot een te kleine aanvullende steekproef adviseert.

Een voorbeeld: de populatie bevat 4% posten met fouten. Dat betekent nog niet dat de populatie een materiele fout bevat want niet alle posten zullen 100% fout zijn. Een pilot van 25 moet zal verwachting 1 fout opleveren. Echter, de kans op 0 fouten in een steekproef van 25 uit een populatie met 4% fouten is 0,36. Met 36% kans zal deze pilot dus 0 fouten opleveren. De conclusie is dan dat er geen aanvullende waarnemingen meer worden omdat de spreiding tussen de foutbedragen 0 is.

De kans dus dat deze pilot van 25 leidt tot een te kleine steekproef is 36%.

Hierbij een tabel met meer uitkomsten:

omvang pilot	25	
werkelijk fout% posten	geen aanvullend onderzoek bij uitkomst	kans
1	0	0,78
2	0	0,60
3	0	0,47
4	0	0,36
5	0	0,28
6	0	0,21
7	0	0,16

Hoe moet het dan wel?

Al in de jaren '80 van de vorige eeuw heeft prof. J.Kriens met collega's van de Universiteit van Tilburg simulatiestudies gedaan op integraal gecontroleerde voorraden. Voorraad is een typisch voorbeeld van een populatie met fouten naar twee kanten waar je de spreiding tussen de foutbedragen nodig hebt voor de evaluatie en dus ook voor de opzet. Zij kwamen tot drie conclusies:

1. dat bij zo'n 150 tot 180 waarnemingen je meestal voldoende fouten vindt voor een bruikbare schatting van de spreiding tussen de foutbedragen. Voor alle zekerheid adviseer ik daarom 200 waarnemingen;

2. dat er dan zelden aanvullende waarnemingen nodig zijn. Het is dus beter om de pilot over te slaan;
3. dat hoe minder fouten de populatie bevat, des te meer waarnemingen nodig zijn om zeker te weten dat een steekproef (met ook weinig fouten) representatief is. En andersom: een populatie met heel veel fouten kan met een kleinere steekproef (zeg, 100 waarnemingen) worden gecontroleerd. Dat betekent dat het risicoanalyse model bij dit soort schattingsmethoden niet werkt.

Kriens, J. & Timmermans, H. & Wildenberg, H. & Kleijnen, Jack. (1989). Regression Sampling in Statistical Auditing: A Practical Survey and Evaluation. Tilburg University, Open Access publications from Tilburg University. 43. 10.1111/j.1467-9574.1989.tb01262.x.