STATISTICAL AUDITING (89) | *gisteren*

De Janboerenfluitjesmethode

Accountant, bent u ook zo enthousiast over de JBF-methode om steekproeven te evalueren? Ik niet.

De AICPA Audit Sampling Guide (tabel 5.1, pagina 94 van de 2020 editie) geeft een voorbeeld van een evaluatie van een niet-statistische steekproef. Dit blijkt de methode te zijn die accountants kennen als 'de janboerenfluitjesmethode'. In de voetnoot onderaan die bladzijde wordt verwezen naar een betere methode die uitgaat van een statistische steekproef. Ik zal laten zien waarom deze JBF-methode niet deugt en hoe de betere methode, die uit de voetnoot in de Guide, werkt.

Een voorbeeld:

Uit een populatie van 10.000 transacties die samen een bedrag van 1.000.000 euro vertegenwoordigen, wordt een (aselecte) steekproef van 3 waarnemingen gedaan. Voor de JBF-methode maakt het niet uit of er transacties (postensteekproeven) of euro’s (geldsteekproeven) zijn gestoken. De gegevens zijn:

Waarneming	Geboekt in €	Correct in €	Fout in €
1	50	43	7
2	100	95	5
3	150	150	0
	300	288	12

De JBF-methode stelt nu dat van de 300 gecontroleerde euro’s er 12 fout zijn, en dat daarom dus ook 4% van de populatie van 1.000.000 euro, dus 40.000 euro fout is.

Wat is er tegen? Straks zal ik vanuit de wiskunde aangeven hoe het wel moet, en dan vergelijk ik de wiskundig juiste formule met de formule die hier wordt gehanteerd. Maar eerst, waarom deze aanpak niet kan kloppen:

- De JBF-methode gaat er in dit voorbeeld vanuit dat er 300 waarnemingen zijn gecontroleerd (waarvan er 12 fout zijn). Dat is niet zo: bij een postensteekproef zijn er 3 van de 10.000 transacties geselecteerd en bij een geldsteekproef zijn er 3 van de 1.000.000 euro’s gestoken en er kwamen 297 euro’s onvrijwillig mee. Door de gecontroleerde euro’s te tellen baseert de JBF-methode in dit voorbeeld de conclusie voor 50% op waarneming 3 (zie tabel hierna)!
- Dat in een geldsteekproef van 3 meer dan 3 euro’s zijn gecontroleerd geeft inderdaad meer zekerheid. Het effect daarvan zit niet in de geprojecteerde fout (de meest waarschijnlijke uitkomst vergeleken met een integrale controle op basis van de bevindingen) maar in de maximale fout (een zo pessimistische raming van de uitkomst van een integrale controle dat het steekproefrisico tot een aanvaardbaar laag niveau is teruggebracht). Dat effect is overigens te verwaarlozen als de gecontroleerde transacties klein zijn ten opzichte van het bedrag dat door één gestoken euro wordt vertegenwoordigd (1.000.000/3 = 333.333).
- Zowel de uitkomst 12 (de totale fout) als de uitkomst 300 (het totaal gecontroleerde bedrag) zijn door toeval tot stand gekomen steekproefresultaten. Als waarneming 3 geen 150 maar 200 euro (geboekt en correct) was geweest dan was de breuk veranderd.

Dat betekent dat deze JBF-methode - die in de Audit Sampling Guide van de AICPA "niet statistisch" wordt genoemd – neerkomt op een van de lastigste problemen in de statistiek: hoe om te gaan met de breuk tussen twee aan kans onderhevige resultaten?

Hoe dan wel?

Hoe moet het dan wel? Het idee van een foutprojectie is het gevonden foutbedrag te corrigeren voor de kans op ontdekken. Voor elk element van de steekproef is de bijdrage tot de foutprojectie gelijk aan het foutbedrag gedeeld door de kans op selectie. Omdat postensteekproeven en geldsteekproeven anders werken qua selectiekansen, moeten we ze apart behandelen.

Postensteekproef
Item 1 heeft een foutbedrag van 7. De kans op selectie van item 1 is bij een postensteekproef van 3 uit een populatieomvang van 10.000 posten: 3/10.000. Let op: de omvang van de transacties is bij een postensteekproef niet relevant. De bijdrage tot de foutprojectie is dus 7 gedeeld door 3/10.000.

Interessant: de juiste methode geeft dezelfde uitkomst als de onjuiste methode! Hoe kan dat? Nou, dat is omdat de gemiddelde omvang van de gestoken posten (300/3) gelijk is aan de gemiddelde omvang van alle posten (1.000.000/10.000). Bij een postensteekproef is dat naar verwachting het geval, maar zeker niet gegarandeerd. Iets wat gemiddeld klopt, klopt per toepassing zelden. Kijk maar wat er gebeurt als waarneming 3 een correct geboekt bedrag van 200 was geweest.

Waarneming	Geboekt in €	Correct in €	Fout in €	Selectiekans	Bijdrage tot projectie in €
1	50	43	>7	0,0003	23333,33333
2	100	95	5	0,0003	16666,66667
3	150	150	0	0,0003	0
	300	288	12		40.000

Geldsteekproef
Waarneming 1 heeft een foutbedrag van 7. De kans op selectie van waarneming 1 van 50 euro is bij een geldsteekproef van 3 uit een populatieomvang van 1.000.000 euro: 3 maal 50/1.000.000. De omvang van de transactie is nu dus wel relevant: hoe meer euro in een transactie, des te groter de kans op selectie. De bijdrage tot de foutprojectie is dus 7 gedeeld door 3 maal 50/1.000.000.

Waarneming	Geboekt in €	Correct in €	Fout in €	Selectiekans	Bijdrage tot projectie in €
1	50	43	7	0,00015	46666,66667
2	100	95	5	0,0003	16666,66667
3	150	150	0	0,00045	0
	300	288	12		63.333

Wat ik in getallen heb gedaan, kan ook in symbolen; de uitkomst verandert niet maar het geeft wel een altijd bruikbare formule. Postensteekproef in symbolen:
T posten hebben samen boekwaarde M en de steekproef bestaat uit n waarnemingen. Voor elke waarneming weten we de geboekte waarde x en de juiste waarde y, dus het foutbedrag (dat ook 0 kan zijn!) is x - y.

De kans op selectie van een transactie is bij een postensteekproef (n/T), dus gelijk voor elke waarneming. De bijdrage aan de geprojecteerde fout is gelijk aan foutbedrag gedeeld door selectiekans, dus (x - y) / (n/T) euro. Voor alle waarnemingen samen kun je dit schrijven als T * SOM((x - y) / n) oftewel:
Aantal posten maal gemiddeld foutbedrag per steekproefpost bij een postensteekproef.

In het voorbeeld: 10.000 maal (12/3) = 40.000.
Geldsteekproef in symbolen:
De totale boekwaarde van de T transacties is M. De kans op selectie van een transactie van omvang x is bij een geldsteekproef: x * n / M (de kans hangt nu immers af van de boekwaarde van de post ten opzichte van de totale boekwaarde).
Voor elke waarneming is de bijdrage aan de geprojecteerde fout nu: (x - y) / (x * n/M) euro. Voor alle waarnemingen samen kun je dit schrijven als (M/n) * SOM(x - y) / x.
We herkennen (M/n) als het selectie-interval voor de geldsteekproef, en SOM (x - y) / x als de totale foutfractie (ook wel *taint* genoemd).
De formule wordt dus:
Interval maal totale foutfractie bij een geldsteekproef.

In het voorbeeld: de foutfracties zijn 7/50 =0,14 en 5/100 =0,05 dus samen 0,19. Het interval is 1.000.000/3 = 333.333 dus de uitkomst wordt 333.333 maal 0,19 = 63.333

Conclusies

Bij een postensteekproef is het steekproefgemiddelde van de transacties (de noemer die niet gebruikt mocht worden) gemiddeld gelijk aan het populatiegemiddelde (de noemer die wel gehanteerd had moeten worden). De JBF-methode geeft dus een gemiddeld juiste, maar onzuivere schatting van de foutprojectie waarbij geen maximale fout is te bepalen en dus de onzuiverheid niet kan worden gekwantificeerd.

Bij een geldsteekproef is het ondenkbaar dat de gemiddelde steekproefwaarneming gelijk is aan de gemiddelde omvang van de transacties in de populatie. De steekproef is immers getrokken met kansen evenredig met de omvang van de transacties en de gemiddelde steekproefwaarneming is dus (veel) groter dan de gemiddelde transactie in de populatie.

Deze JBF-methode (evaluatie postensteekproef als was het een geldsteekproef) deelt dus door een te groot gemiddelde een geeft daarmee een onderschatting van de geprojecteerde fout.
Vandaar dat de voetnoot op bladzij 94 van de AICPA Audit Sampling Guide meldt dat als de niet-statistische steekproef is gemaakt om op een geldsteekproef te lijken (door transacties met hogere bedragen meer kans te geven dan kleinere bedragen), er beter met een andere formule kan worden gewerkt. Dat blijkt de formule die ik heb uitgelegd: foutbedrag gedeeld door selectiekans.

Voor de liefhebbers (en die blijken er inderdaad te zijn!): de wiskundige fout in de JBF-methode is dat men SOM((x-y)/x) heeft verward met SOM(x-y)/SOM(x). In het voorbeeld: 7/50+5/100+0/150 is niet hetzelfde als (7+5+0)/300. Bij een postensteekproef is SOM((x-y)/n) wel gelijk aan (1/n) SOM(x-y). In het voorbeeld: 7/3+5/3+0/3 = (7+5+0)/3. Dat is het hele verschil.

Deel dit artikel

Drs. Paul van Batenburg is zelfstandig adviseur die als statisticus met verstand van controleren de eenmanszaak en website steekproeven.eu voert.

GERELATEERD

STATISTICAL AUDITING (88) | *15 december 2020*

Inventarisatie ten behoeve van de controle op de NOW regeling: plan B in actie!

Tot nu toe heb ik altijd accountants aangeraden om bij de controle op de hoeveelheden in de voorraadadministratie geen geldsteekproef maar een postensteekproef te... →

STATISTICAL AUDITING (87) | *08 oktober 2020*

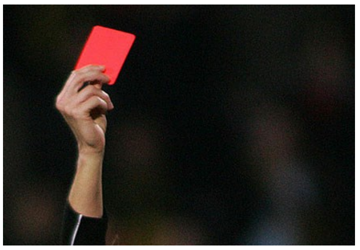
Het gebruik van cijferanalyse ten tijde van Corona: zorg voor een plan B!

Gegevensgerichte cijferanalyse is te zien als een vorm van machine learning. Er wordt een relatie verondersteld tussen te controleren gegevens en andere gegevens.... →

STATISTICAL AUDITING (86) | *21 augustus 2020*

Controle van een bestand: de steekproefomvang is niet het probleem, maar de populatie!

Het gebeurt nog wel eens dat een accountant wordt gevraagd om, los van de jaarrekeningcontrole, een uitspraak te doen over de kwaliteit van een bestand. →



STATISTICAL AUDITING (85) | *17 juli 2020*

Boete? Waarom en waarover?

Onlangs kreeg ik een vraag van een fiscalist over een boete die werd opgelegd naar aanleiding van een correctie die was gebaseerd op een statistische steekproef.... →

STATISTICAL AUDITING (84) | *22 juni 2020*

IPE testing: 'wat' is belangrijker dan 'hoeveel'

Onlangs had ik een discussie met accountants over de steekproefomvang bij IPE-testing (IPE=Information provided by entity). Conclusie: gebruik je eigen professionele... →