

I N S T I T U U T
I N S T I T U U T
C O L L E C T I E
C O L L E C T I E
N E D E R L A N D
N E D E R L A N D

ICN-Informatie Instituut Collectie Nederland

nr. 4

De steekproef als hulpmiddel bij collectiebeheer

Inleiding

Een van de kerntaken van het ICN is het verstrekken van adviezen op het gebied van collectiemanagement. Voor musea wordt het steeds belangrijker om bij het ‘managen’ van hun collectie te beschikken over betrouwbare cijfers. Het ICN biedt de musea met dit informatieblad een instrument om in het kader van collectiemanagement betrouwbare cijfers over hun collectie te verzamelen.

Steekproef als hulpmiddel

Sommige collecties zijn zo groot dat niet elk object (regelmatig) bekeken of gecontroleerd kan worden. De informatie over de objecten is vaak onvolledig. Dit maakt het voor musea lastig betrouwbare uitspraken te doen over omvangrijke collecties en te begroten hoeveel tijd en geld nodig zijn om bepaalde handelingen aan de collectie te verrichten. Immers, het is onmogelijk te begroten wat conservering kost als er geen gegevens voorhanden zijn over de conditie van een collectie.

In zo'n geval biedt een steekproef uitkomst. Door een steekproef te trekken is het mogelijk met de gegevens van een *deel* van de collectie, een uitspraak doen over de *totale* collectie. Het trekken van een steekproef is gebonden aan regels.

In dit informatieblad wordt uiteengezet hoe een *eenvoudige* steekproef kan worden getrokken. Daarbij zijn de volgende begrippen van belang:

- populatieomvang** het aantal objecten in de te beoordelen collectie;
- steekproefomvang** het aantal objecten dat uit de collectie wordt geselecteerd;
- aselecte steekproef** een steekproef waarbij elk object uit de populatie een even grote kans heeft om te worden gekozen.

Voorwaarden

Voor het trekken van een eenvoudige steekproef moet worden voldaan aan twee belangrijke voorwaarden:

1 De steekproef moet getrokken worden uit een homogene groep objecten die voor de steekproef relevante, vergelijkbare kenmerken bezitten. Zo kan het belangrijk zijn om groepen objecten met een verschillende herkomst, die later zijn samengevoegd tot één collectie, afzonderlijk te bekijken. Bijvoorbeeld: *De grote en zorgvuldig genummerde penningencollectie van een particulier verzamelaar is in het verleden toegevoegd aan de matig genummerde penningcollectie van het museum. Beide delen zijn door de wijze van nummeren nog te herkennen. Zomaar een aselecte steekproef uit de totale collectie trekken heeft geen zin, omdat vooraf al bekend is dat het ene deel van de collectie uitvoeriger is geregistreerd dan het andere. De collectie is niet homogeen. Het is dan beter twee aparte steekproeven te trekken.*

2 De populatieomvang moet bekend zijn.

Nauwkeurigheid

Omdat het bij grote populaties onmogelijk is alle objecten te bekijken of te controleren, kan nooit een uitspraak worden gedaan die voor honderd procent nauwkeurig is. Wel kan de gewenste nauwkeurigheid worden aangegeven. In de statistiek is het gebruikelijk te werken met een nauwkeurigheid van 95% of 90%. In het eerste geval betekent dit dat de nauwkeurigheid met een marge van maximaal 5% naar boven *of* beneden kan afwijken. In het tweede geval bedraagt de marge maximaal 10% naar boven *of* beneden. De mate van nauwkeurigheid heeft consequenties voor de steekproefomvang. Anders gezegd: hoe kleiner de gewenste marge, hoe groter de steekproef moet zijn.

In de tabel is voor diverse waarden van de populatieomvang aangegeven hoe groot de steekproef moet zijn bij een gewenste nauwkeurigheid van 95% respectievelijk 90%. Als de populatieomvang 3000 is, moet de steekproef bijvoorbeeld 359 of 77 zijn. Als een populatieomvang 3382 is, kunnen de getallen naar boven worden afgerond en bedraagt de steekproef 370 of 78.

Als de steekproefomvang is bepaald, kan worden onderzocht welk collectiedeel voldoet aan een bepaalde eigenschap. Bijvoorbeeld: welk deel is geregistreerd of

ICN-Informatie
Nummer 4, juli 2000
Verschijnt onregelmatig

Redactiesecretariaat:
Afdeling Cultuurbeheer &
Passieve conservering
Postbus 76709
1070 KA Amsterdam
T 020 305 46 18
F 020 305 46 33
E frank.bergevoet@icn.nl

© 2000 Instituut Collectie Nederland (ICN).
Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande toestemming van het ICN. Het ICN kan niet verantwoordelijk worden gesteld voor schade veroorzaakt door het toepassen van de beschreven methoden en/of materialen.

Druk drukkerij Mart.Spruijt bv, Amsterdam

Issn 1566-760x

verkeert in een slechte conditie? Hoe dat gaat, wordt geïllustreerd aan de hand van het volgende voorbeeld.

Een natuurhistorisch museum overweegt een particuliere schenking te aanvaarden van 1915 schelpen. Of het museum de schenking aanvaardt, is afhankelijk van de kosten die de registratie van de verzameling met zich meebrengt. Het museum wil weten van hoeveel schelpen de vindplaats en -datum genoteerd zijn.

Gekozen wordt een nauwkeurigheid van 95%. De tabel geeft aan dat bij een populatieomvang van 1915 tussen de 334 en 340 objecten moeten worden bekeken om tot een betrouwbare uitspraak te komen. Naar boven afgerond geeft dit aantal een steekproefomvang van 340. Die 340 objecten worden aselect bepaald en bekeken. Na controle blijkt dat van 234 schelpen gegevens over de herkomst en vondstdatum aanwezig zijn. Dat is 0,69 (234/340) oftewel 69% van de steekproef.

Omdat de steekproefomvang gebaseerd is op de tabel en voldoende groot is, kan nu met 95% betrouwbaarheid worden beweerd dat tussen de (69% - 5%) 64% en (69% + 5%) 74% van de totale populatie schelpen vergezeld gaat van de vondstgegevens.

Op basis van deze percentages kan het museum een minimale en maximale begroting opstellen voor het registratieproject.

Bij eenzelfde berekening op basis van een gewenste nauwkeurigheid van 90% hoeven slechts 76 schelpen te worden bekeken. Daarvan gaan er bijvoorbeeld 52 vergezeld van de vondstgegevens. Dat is 0,68 (52/76) oftewel 68% van de steekproef. Echter, omdat gekozen is voor een kleinere nauwkeurigheid, kan slechts worden beweerd dat tussen de (68% - 10%) 58% en (68% + 10%) 78% van de schelpen is geregistreerd. Vertaald in een begroting wordt duidelijk dat de minimale en maximale kosten verder uit elkaar liggen dan bij een gewenste nauwkeurigheid van 95%.

Willekeurig getal

Het is van belang dat de steekproefomvang aselect, dus door het toeval wordt bepaald. Spreadsheetprogramma’s als Excel kunnen helpen met het genereren van aselecte getallen. Om Excel te kunnen gebruiken moeten de objecten uit een populatie gekoppeld zijn aan een nummer. Het is aan te bevelen om de objecten voor de duur van de steekproef een nummer te geven. Een voorbeeld kan dit verduidelijken.

De schelpenverzameling wordt bewaard in vier ladekasten. Elke kast heeft vijf laden die weer zijn onderverdeeld in kleine vakjes. In elk vakje ligt een schelp en al dan niet een kaartje met de vondstgegevens. De vakjes worden tijdelijk genummerd van 1 tot 1915. Met behulp van de computer wordt een lijst van 340 willekeurige cijfers bepaald (de steekproefomvang bij een gekozen nauwkeurigheid van 5%). Vervolgens wordt de inhoud van de 340 vakjes bekeken.

Op vergelijkbare wijze kunnen stellingen, kasten, legborden, dozen, rekken en objecten worden genummerd.

Als de objecten uit een populatie die bekeken moet worden een opeenvolgend registratienummer hebben, kan zo’n nummer ook worden gebruikt als uitgangspunt voor het bepalen van de willekeurige getallen.

Met behulp van Excel kan op de volgende manier een reeks willekeurige getallen worden bepaald:

- Klik in een leeg werkblad een willekeurige cel aan
- Typ in de formulebalk van Excel in =aselect()*(1915-1)+1 en geef enter
- Het programma berekent een toevalscijfer tussen de 1 en 1915. De cijfers (in dit geval 1915 en 1) bepalen waartussen het toevalscijfer wordt berekend. =aselect()*(500-27)+27 betekent bijvoorbeeld dat de computer een getal berekent tussen de 27 en 500
- Kopieer de formule naar het aantal cellen dat overeenkomt met de steekproefomvang.
- Rond de cijfers eventueel af met behulp van ‘rechtermuisknop/celeigenschappen/getal/decimalen’

Geavanceerd De cijfers kunnen eventueel nog oplopend gesorteerd worden. Daartoe moeten de cijfers echter eerst met ‘bewerken/plakken speciaal/waarden’ gekopieerd worden naar vrije cellen. Zie vervolgens ‘help/oplopende volgorde (sorteren)’.

Tabel steekproefomvang

<i>Populatie-omvang</i>	<i>Steekproef-omvang bij een nauwkeurigheid van 95%</i>	<i>Steekproef-omvang bij een nauwkeurigheid van 90%</i>
20	20	20
30	30	25
40	40	30
50	49	34
60	56	39
70	65	41
80	72	44
90	78	44
100	84	47
150	114	54
200	139	59
250	160	62
300	178	64
350	193	66
400	207	67
450	219	69
500	229	69
550	239	70
600	247	71
650	255	71
700	262	72
750	268	72
800	274	73
850	279	73
900	284	73
950	289	74
1000	293	74
1200	307	75
1400	318	75
1600	327	76
1800	334	76
2000	340	76
2250	346	77
2500	351	77
3000	359	77
4000	370	78
5000	377	78
6000	381	78
7000	384	78
8000	387	78
9000	389	78
10000	390	78
20000	398	79
50000	402	79
100000	405	79