



# Steekproef voor controletellingen bij Tweede Kamerverkiezing 2021

datum 18 maart 2021

samenvatting Deze notitie beschrijft het ontwerp en de trekking van de steekproef ten behoeve van controles van de software die wordt gebruikt bij de Tweede Kamerverkiezing van 2021.

trefwoorden Enkelvoudig aselechte steekproef, uniforme verdeling

## 1. Inleiding

Bij de Tweede Kamerverkiezing op 17 maart 2021 worden optellingen in de software OSV2020 steekproefsgewijs gecontroleerd. Deze software sommeert onder andere in elke gemeente per politieke partij<sup>1</sup> de aantallen stemmen per stembureau tot stemtotalen per politieke partij in de betreffende gemeente. In elke gemeente worden voor een steekproef van drie politieke partijen de stemtotalen per politieke partij nageteld en vergeleken met de uitkomsten van de software.

De Kiesraad heeft aan het Centraal Bureau voor de Statistiek gevraagd om de steekproef ten behoeve van deze controles te trekken. Voor elke gemeente is een steekproef met omvang drie gewenst uit de politieke partijen die in de kieskring van de betreffende gemeente deelnemen aan de Tweede Kamerverkiezing van 2021. De steekproeven per gemeente moeten onafhankelijk van elkaar worden getrokken. Deze notitie bevat informatie over welke politieke partijen per kieskring mee doen, het toegepaste steekproefontwerp, en het resultaat van de trekking volgens dat ontwerp.

## 2. Politieke partijen 2021

In een openbare zitting heeft de Kiesraad bekend gemaakt dat 37 politieke partijen deelnemen aan de Tweede Kamerverkiezing van 17 maart 2021. De namen met bijbehorende nummering zoals vastgesteld door de Kiesraad zijn opgenomen in Tabel 1.

Voor de Tweede Kamerverkiezing 2021 is Nederland verdeeld in 20 kieskringen, die elk uit een of meer gemeenten bestaan. De 352 gemeenten in Europees Nederland zijn opgenomen in de eerste negentien kieskringen. De drie gemeenten van Caribisch Nederland: Bonaire, Sint-Eustatius en Saba, vormen de twintigste kieskring. Politieke partijen kunnen in verschillende kieskringen met verschillende kandidatenlijsten uitkomen, en kunnen ervoor kiezen om niet in alle kieskringen deel te nemen aan de verkiezing. Tabel 2 bevat per kieskring de politieke partijen die niet deelnemen aan de Tweede Kamerverkiezing 2021.

---

<sup>1</sup> Formeel gaat het, zoals in de Kieswet beschreven, om 'lijsten' die aan de verkiezingen deelnemen. Ter wille van de leesbaarheid wordt in deze nota het begrip 'politieke partij' of 'partijlijst' gehanteerd.

Tabel 1 Politieke partijen Tweede Kamerverkiezing maart 2021

Partijnummer	Partijnaam
1	VVD
2	PVV (Partij voor de Vrijheid)
3	CDA
4	D66
5	GROENLINKS
6	SP (Socialistische Partij)
7	Partij van de Arbeid (P.v.d.A.)
8	ChristenUnie
9	Partij voor de Dieren
10	5OPLUS
11	Staatkundig Gereformeerde Partij (SGP)
12	DENK
13	Forum voor Democratie
14	BIJ1
15	JA21
16	CODE ORANJE
17	Volt
18	NIDA
19	Piratenpartij
20	LP (Libertaire Partij)
21	JONG
22	Splinter
23	BBB
24	NLBeter
25	Lijst Henk Krol
26	OPRECHT
27	JEZUS LEEFT
28	Trots op Nederland (TROTS)
29	UBuntu Connected Front
30	Blanco lijst met als eerste kandidaat Zeven, A.J.L.B.
31	Partij van de Eenheid
32	DE FEESTPARTIJ (DFP)
33	Vrij en Sociaal Nederland
34	Wij zijn Nederland
35	Modern Nederland
36	De Groenen
37	Partij voor de Republiek

Tabel 2 Niet deelnemende politieke partijen per kieskring

Kieskring	Partijnummer
1	30 31 34 35 36 37
2	29 30 31 32 34 35
3	29 31 32 33 34 36 37
4	32 33 35 36 37
5	28 30 32 33 34 35 36 37
6	29 31 33 34 35 36 37
7	31 32 33 34 35 36 37
8	30 31 34 35 36 37
9	30 32 34 35
10	27 30 31 35 36 37
11	31 33 35 36 37
12	27 30 32 34 35 36 37
13	28 31 32 33 36 37
14	28 29 31 32 33 36 37
15	28 31 32 33 34 36 37
16	28 29 32 33 34 35 36 37
17	28 33 34 35 36 37
18	29 33 34 35 36 37
19	27 29 33 35 36 37
20	21 22 23 24 25 26 27 28 30 31 32 33 34 35 36 37

### 3. Ontwerp en trekking van de steekproef

Het meest eenvoudige en meest bekende steekproefontwerp is het enkelvoudig aselechte steekproefontwerp. Dit ontwerp staat het dichtst bij wat intuïtief onder het trekken van een steekproef wordt verstaan. Ieder element heeft dezelfde kans om in de steekproef te komen en er wordt geen gebruik gemaakt van hulpinformatie. Het ontwerp kan zowel met als zonder terugleggen worden toegepast. In de praktijk hebben steekproeven zonder terugleggen de voorkeur omdat deze de meeste informatie geven.

Een voorbeeld van een enkelvoudig aselechte steekproef zonder terugleggen is de lottomachine. Een voor een worden balletjes getrokken, waarbij alle balletjes die in de machine zitten dezelfde kans hebben om te worden geselecteerd en getrokken balletjes worden niet in de machine teruggelegd.

De formele definitie van het enkelvoudig aselechte steekproefontwerp zonder terugleggen voor steekproeven met vaste omvang  $n$  uit een populatie is: elke deelverzameling van  $n$  verschillende elementen uit de populatie heeft dezelfde kans om als steekproef te worden gerealiseerd. Als  $N$  de populatieomvang is, dan heeft elk element uit de populatie kans  $n/N$  om te worden geselecteerd. De kans dat een vooraf aangewezen tweetal elementen in de steekproef terecht komt is  $n(n-1)/(N(N-1))$ .

Voor de steekproef van te controleren partijlijsten per gemeente is gekozen om in elke gemeente, onafhankelijk van elkaar, het enkelvoudig aselechte steekproefontwerp zonder terugleggen toe te passen. Per gemeente is volgens dit ontwerp een steekproef met omvang drie getrokken uit de deelnemende politieke partijen in de kieskring van de betreffende gemeente.

Bij de trekking van de steekproef is softwarepakket R, versie 4.0.3, gebruikt. Door de standaardfunctie 'sample' aan te roepen met argumenten  $x =$  de vector met te kiezen partijnummers in een gemeente, en  $n = 3$ , wordt een enkelvoudig aselechte steekproef zonder terugleggen getrokken met omvang 3 uit de te kiezen partijnummers in de betreffende gemeente.

De sample functie maakt gebruik van een pseudorandomgenerator die bij benadering een uniform verdeelde output produceert. Ten behoeve van de reproduceerbaarheid van de output is gebruik gemaakt van een random seed. De getallenreeks van de pseudorandomgenerator wordt volledig bepaald door de keuze van de random seed. Als de pseudorandomgenerator opnieuw wordt geïnitieerd met dezelfde random seed, zal hij dezelfde getallenreeks produceren. De waarde van de random seed hoeft niet willekeurig te zijn. Vanwege de aard van het algoritme zullen de getallen die het algoritme genereert de uniforme verdeling volgen, dit met uitzondering van het eerste getal. Door de random seed willekeurig te kiezen verdwijnt deze uitzondering. Dat is gedaan door de seconden, nauwkeurig tot op de milliseconde, van de op een willekeurig moment gekozen systeemtijd van de computer te vermenigvuldigen met duizend. Voor de reproduceerbaarheid is het ook van belang te melden dat de gemeentesteekproeven in volgorde van het invoerbestand van de Kiesraad zijn getrokken en geleverd.

Het resultaat van de steekproeftrekking is een steekproef van te controleren partijnummers per gemeente. Hierbij is de partijnummering van Tabel 1 genomen. De steekproef is opgenomen in een bestand dat door het CBS is geleverd aan de Kiesraad. Dit bestand bevat vijf kolommen die gescheiden zijn met een puntkomma. De kolommen bevatten achtereenvolgens gemeentecode, gemeentenaam en de drie te controleren partijnummers.

Tabel 3 is een samenvatting van de steekproef. Per partij is opgenomen in hoeveel gemeenten de stemmen op de betreffende partij moeten worden gecontroleerd. In de tabel valt op dat met name partijnummers 33, 35, 36 en 37 relatief weinig in de steekproef voorkomen. Dat is plausibel omdat deze partijen in veel kieskringen niet meedoen aan de Tweede Kamerverkiezing, zie Tabel 2.

Tabel 3 Frequentie van politieke partijen in de steekproef

<u>Partijnummer</u>	<u>Aantal gemeenten</u>
1	39
2	38
3	28
4	32
5	35
6	29
7	32
8	46
9	38
10	35
11	40
12	37
13	41
14	24
15	42
16	34
17	33
18	30
19	43
20	34
21	34
22	22
23	35
24	34
25	40
26	30
27	29
28	25
29	24
30	23
31	16
32	24
33	2
34	12
35	2
36	2
37	1