HOOFDSTUK 7

Verdelingen en betrouwbaarheid

Wiskunde A · havo 5



IN DIT HOOFDSTUK LEER JE:

verdelingen beschrijven en normale vuistregels gebruiken;
steekproefuitkomsten en betrouwbaarheid beoordelen;
intervallen en een passende steekproefomvang berekenen.

Wat weet je nog?

Maak deze vragen voordat je begint. Noteer je berekeningen en redeneringen in je schrift. Alle onderzoeksgegevens in dit hoofdstuk zijn voor de oefeningen samengesteld. Bij normale modellen gebruik je de benaderende vuistregels, tenzij anders vermeld.

1 Tijden samenvatten

Vijf sorteertijden zijn 8, 9, 10, 11 en 17 minuten.

a Bereken het gemiddelde en de mediaan.

b De tijd van 17 minuten wordt vervangen door 12 minuten. Welke centrummaat verandert en welke blijft gelijk?

2 Aandelen vergelijken

In groep A gebruiken 36 van de 80 deelnemers een leenapparaat. In groep B zijn dat 42 van de 120 deelnemers.

a In welke groep is het aandeel gebruikers groter? Onderbouw met percentages.

3 Wie wordt onderzocht?

Een organisatie heeft 6000 leden. Zij loot 300 leden uit het volledige ledenbestand voor een vragenlijst over de openingstijden.

a Noem de populatie, de beoogde steekproef en de variabele.

b Van de 300 leden antwoorden er 180. Welk aandachtspunt ontstaat?

4 Een gemiddelde vergelijken

Bij locatie P zijn de wachttijden 4, 6 en 8 minuten. Bij locatie Q zijn ze 2, 6 en 10 minuten.

a Bereken de twee gemiddelden en vergelijk de spreiding.

5 Een deel van een geheel

Van 240 kaarten hebben er 90 een blauwe rand.

a Schrijf het aandeel als breuk, decimaal getal en percentage.

6 Een midden en twee grenzen

Een interval loopt van 0,31 tot 0,47.

a Bereken het midden, de afstand van het midden tot een grens en de totale breedte.

b Geef de grenzen als percentages.

7 Wortels en breuken

Reken zonder tussentijds af te ronden.

a Bereken $\frac{24}{\sqrt{144}}$ en $\frac{24}{\sqrt{576}}$.

b Met welke factor verandert de uitkomst als de waarde onder de wortel viermaal zo groot wordt?

8 Altijd genoeg plaatsen

Een berekening voor de benodigde capaciteit geeft 72,25 plaatsen.

a Hoeveel hele plaatsen zijn minstens nodig? Leg de afronding uit.

7.1 Verdelingsvormen en normale vuistregels

Leerdoel

Ik kan verdelingen vergelijken op vorm, centrum, spreiding, staarten en uitschieters.

Theorie

Verdelingen beschrijven

Een verdeling laat zien hoe waarden voorkomen. Een klokvorm heeft één top met aflopende flanken; een uniforme verdeling is ongeveer vlak en een meertoppige verdeling heeft meerdere duidelijke toppen. Beschrijf een scheve verdeling door te zeggen aan welke kant de lange staart ligt.

Vergelijk bij dezelfde eenheid ook het centrum en de spreiding. Een veraf gelegen losse waarde kan een uitschieter zijn. Een histogram kan een uitschieter suggereren, maar de oorspronkelijke waarnemingen geven meer zekerheid. Zie voorbeeld 7.1A, Voorbeelden blz. 1.

Onthoud: Een gemiddelde en een standaardafwijking bepalen op zichzelf de verdelingsvorm niet.

1 Rijen vergelijken

De drie rijen bevatten gemeten wachttijden in minuten.

Rij	Waarden
A	4, 7, 10, 10, 13, 16
B	8, 9, 10, 10, 11, 12
C	8, 9, 10, 10, 11, 12, 31

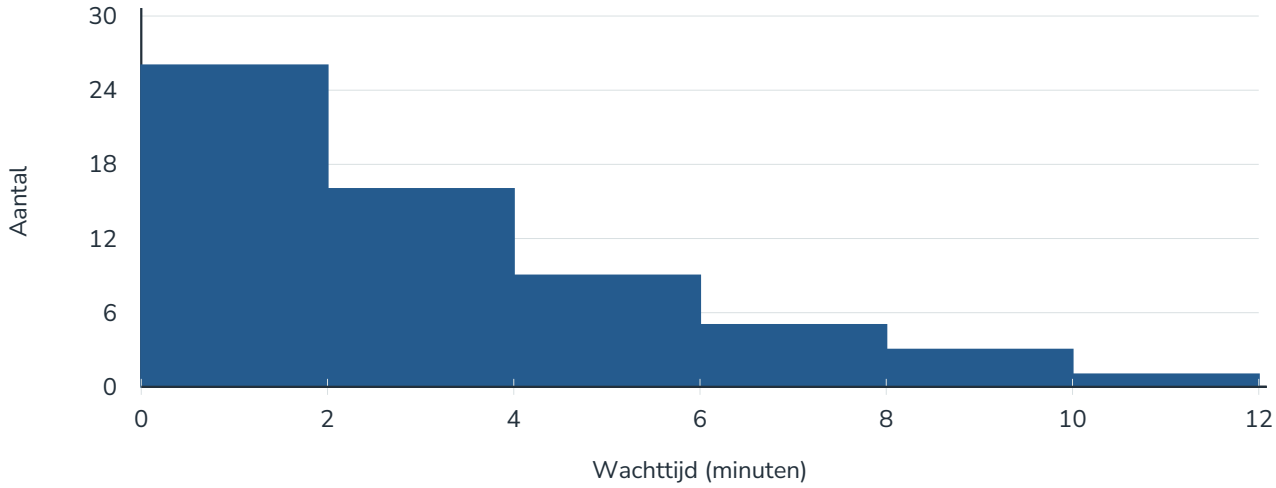
a Vergelijk A en B op centrum en spreiding. Kijk eerst naar het midden en vervolgens naar de uiterste waarden.

b Beschrijf het bijzondere aan C. Leg uit waarom alleen het gemiddelde een onvolledig beeld geeft.

2 De uitgiftebalie

Het histogram toont 60 samengestelde wachttijden.

Wachttijd bij de balie



a Beschrijf de vorm en de kant van de langste staart.

b Een verslag gebruikt een symmetrisch klokmodel voor deze tijden. Vind je de figuur daarvoor voldoende steun?

3 De gemengde meting

Van twee typen sensoren worden evenveel metingen samengevoegd. De metingen van type A liggen dicht bij 20; die van type B dicht bij 40. De groepen overlappen nauwelijks.

a Welke vorm verwacht je bij alle metingen samen? Leg uit.

b Het gezamenlijke gemiddelde is 30. Waarom kan dat als beschrijving van een gebruikelijke meting misleiden?

4 Twee productielijnen

De tabel geeft exacte metingen in millimeters; elke rij bevat acht metingen.

Lijn	Metingen
P	18, 19, 20, 20, 20, 20, 21, 22
Q	20, 20, 21, 22, 22, 23, 24, 40

a Vergelijk de centra en spreidingen en beoordeel of in één rij een opvallend hoge waarde voorkomt.

Ligging en spreiding

Leerdoel

Ik kan bij een normale verdeling het gemiddelde en de standaardafwijking verbinden aan de ligging en spreiding.

Theorie

Ligging en spreiding

Een normale verdeling is een symmetrisch klokmodel. Het gemiddelde μ ligt in het midden. De standaardafwijking σ beschrijft de spreiding in dezelfde eenheid als de gemeten grootheid. Een grotere σ geeft een bredere, lagere klok bij dezelfde totale oppervlakte.

Een verandering van μ verschuift de klok. Een verandering van σ verandert de breedte. De oppervlakte onder de volledige klok vertegenwoordigt alle waarnemingen; de hoogte op één plek is geen percentage. Zie voorbeeld 7.1B, blz. 559.

Onthoud: Bij een normale verdeling liggen gemiddelde en mediaan op dezelfde plaats.

5 Instellingen van een vulmachine

Model A heeft μ = 250 gram en σ = 4 gram. Model B heeft μ = 250 gram en σ = 7 gram. Beide modellen zijn normaal.

a Beschrijf welk kenmerk gelijk blijft en welk kenmerk verandert.

b Model C ontstaat als elke vulling van A precies 6 gram zwaarder wordt. Geef μ en σ van C.

6 Een andere eenheid

De lengte van een onderdeel is volgens een normaal model gemiddeld 8,2 cm met standaardafwijking 0,15 cm.

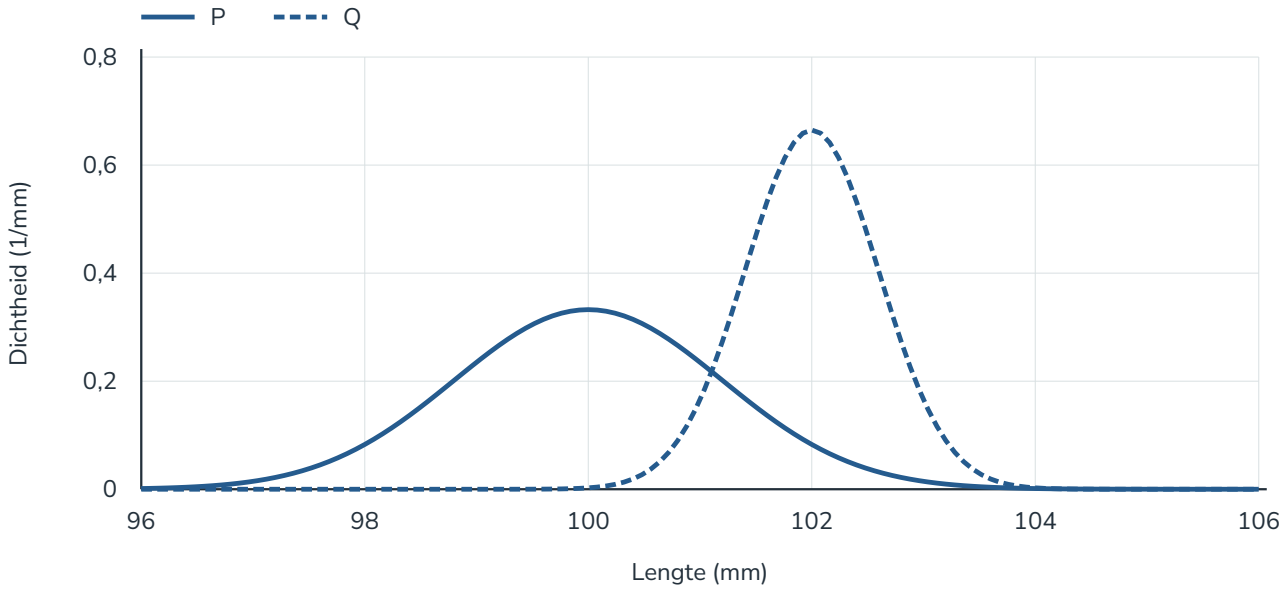
a Geef beide kengetallen in millimeters.

b Een leerling rekent alleen het gemiddelde om. Welke onjuiste conclusie over de spreiding kan daardoor ontstaan?

7 Nauwkeurige plaatjes

Plaatjes worden gemaakt volgens normale modellen P en Q. P: $\mu = 100$ mm, $\sigma = 1{,}2$ mm. Q: $\mu = 102$ mm, $\sigma = 0{,}6$ mm. De gewenste maat is 100 mm.

Twee normale modellen



a Welke lijn ligt gemiddeld het dichtst bij de gewenste maat, en welke lijn werkt met de kleinste spreiding?

b Van ieder Q-plaatje wordt precies 2 mm afgehaald. Vergelijk het nieuwe model met P.

De normale vuistregels

Leerdoel

Ik kan de normale vuistregels toepassen op intervallen rond het gemiddelde.

Theorie

De normale vuistregels

Bij een normaal model ligt ongeveer 68% tussen $\mu - \sigma$ en $\mu + \sigma$, ongeveer 95% tussen $\mu - 2\sigma$ en $\mu + 2\sigma$, en nagenoeg 100% tussen $\mu - 3\sigma$ en $\mu + 3\sigma$. Die laatste benadering is circa 99,7%, dus niet letterlijk alles.

De intervallen zijn symmetrisch rond μ. Voor een verwacht aantal vermenigvuldig je het aandeel met het totale aantal. Werkelijke aantallen kunnen daarvan afwijken. Zie voorbeeld 7.1B, blz. 559.

Onthoud: Controleer eerst of een normale verdeling gegeven of verdedigbaar is.

8 Een interval invullen

Droogtijden volgen een normaal model met $\mu = 50$ minuten en $\sigma = 4$ minuten.

a Bereken eerst $\mu - \sigma$ en $\mu + \sigma$. Welk aandeel hoort bij dit interval?

b Bereken het symmetrische interval voor ongeveer 95%.

9 Controle van een partij

Een normaal model voor de inhoud van potjes heeft $\mu = 360$ ml en $\sigma = 5$ ml. Een partij bevat 800 potjes.

a Geef de intervallen voor ongeveer 68% en 95% van de inhoudsmaten.

b Hoeveel potjes verwacht je buiten het tweede interval? Geef ook aan of dit aantal vaststaat.

10 Grenzen terugvinden

De spanning van batterijen volgt een normaal model. Het symmetrische interval voor ongeveer 95% is 11,4 tot 12,6 volt.

a Bepaal μ en σ.

b Welke grenzen horen bij ongeveer 68%?

c Geef een interval met nagenoeg alle spanningen.

11 Een te sterke belofte

Een folder schrijft: “Bij een normaal model ligt 95% binnen twee standaardafwijkingen. Buiten drie standaardafwijkingen kan dus geen enkele waarde liggen.”

a Leg uit welke gevolgtrekking onjuist is en verbeter de laatste zin.

Aandelen en grenzen

Leerdoel

Ik kan met symmetrie en normale vuistregels een aandeel of grens benaderen en interpreteren.

Theorie

Aandelen en grenzen

Door symmetrie ligt aan elke kant van μ ongeveer 50%. Buiten het middelste 68%-interval ligt samen ongeveer 32%, dus circa 16% aan elke kant. Buiten het middelste 95%-interval ligt circa 2,5% per staart.

Tussen $\mu + \sigma$ en $\mu + 2\sigma$ ligt ongeveer $\frac{(95\% - 68\%)}{2} = 13{,}5\%$. Combineer zulke gebieden. Met alleen deze vuistregels bepaal je geen betrouwbaar percentage bij elke willekeurige grens. Zie voorbeeld 7.1B, blz. 559.

Onthoud: Een grens in één staart hoort bij een eenzijdig aandeel.

12 Te laat klaar

Een normaal model voor doorlooptijd heeft $\mu = 30$ minuten en $\sigma = 3$ minuten.

a Welk aandeel duurt meer dan 33 minuten?

b Welk aandeel ligt tussen 30 en 36 minuten?

13 De bovenste groep

Een normaal model voor de massa van een product heeft $\mu = 75$ gram en $\sigma = 2$ gram.

a Boven welke grens ligt ongeveer 2,5% van de producten?

b Kun je met uitsluitend de drie vuistregels het aandeel boven 78 gram bepalen? Leg uit.

14 Een garantiegrens kiezen

Een normaal model voor de montagetijd heeft $\mu = 42$ minuten en $\sigma = 5$ minuten. Een bedrijf kiest één van drie tijdgrenzen: 47, 52 of 57 minuten. Het wil volgens het model minstens 97% op tijd afronden.

a Bepaal met de vuistregels het aandeel dat uiterlijk bij elk van de grenzen klaar is.

b Kies de kleinste van de drie passende grenzen. Is de doelstelling daarmee voor elke werkdag gegarandeerd?

15 Koelen van onderdelen

Koeltijden volgen een normaal model. Ongeveer 95% ligt tussen 18 en 30 minuten; deze grenzen liggen even ver van het gemiddelde.

a Bepaal het gemiddelde en de standaardafwijking.

b Van 600 onderdelen wordt geteld hoeveel tussen 27 en 30 minuten nodig hebben. Welk aantal verwacht je?

c Een medewerker zegt dat 30 minuten de langst mogelijke koeltijd is. Beoordeel dit.

16 Een andere vraagrichting

Volgens een normaal model ligt ongeveer 16% van de diameters onder 28 mm en ongeveer 16% boven 36 mm.

a Bepaal het gemiddelde en de standaardafwijking.

b Welk aandeel ligt volgens de vuistregels tussen 24 en 36 mm?

7.2 Steekproefvariatie en verdelingen van steekproefuitkomsten

Leerdoel

Ik kan verklaren waarom aselecte steekproeven uit dezelfde populatie verschillende uitkomsten geven.

Theorie

Toeval bij het trekken

In een aselecte steekproef bepaalt toeval welke eenheden worden gekozen. Daardoor verschillen aantallen, proporties en gemiddelden tussen herhaalde steekproeven, ook als de populatie niet verandert. Dit heet steekproefvariatie.

Vergelijk uitkomsten bij dezelfde populatie, vraagstelling en meetwijze. Een verschil tussen twee uitkomsten is op zichzelf geen bewijs dat de populatie is veranderd of dat één meting fout is. Zie voorbeeld 7.2A, blz. 560.

Onthoud: Aselect betekent niet dat elke afzonderlijke steekproef de populatie exact weerspiegelt.

1 Herhaald loten

In een oefenpopulatie heeft 40% een ja-antwoord. Er worden telkens opnieuw 100 personen aselect gekozen. De populatie blijft gelijk. Vijf steekproeven bevatten 37, 44, 39, 41 en 35 ja-antwoorden.

a Bereken de vijf steekproefpercentages.

b Waarom hoeven de vijf percentages niet allemaal 40% te zijn?

2 Twee avondmetingen

Twee teams loten onafhankelijk elk 200 personen uit dezelfde grote, onveranderde populatie. Team A vindt 118 voorstanders, team B 106.

a Bereken het verschil in procentpunten.

b Team A noemt het verschil een bewezen verandering van mening. Wat ontbreekt in deze redenering?

3 Meer herhalingen

Een computer trekt eerst 20 en daarna nog 980 steekproeven van steeds 80 personen uit een vaste grote populatie. Na elke trekking wordt het bestand opnieuw volledig beschikbaar.

a Wordt een afzonderlijke steekproef hierdoor groter? Wat wordt er wel nauwkeuriger in beeld gebracht?

b Is het gemiddelde van 1000 steekproefproporties noodzakelijk exact gelijk aan de populatieproportie?

Wat stelt één punt voor?

Leerdoel

Ik kan individuele waarnemingen onderscheiden van steekproefgemiddelden en -proporties en de normale benadering uitleggen.

Theorie

Wat stelt één punt voor?

Een verdeling van individuele tijden bevat één waarde per persoon. In een steekproevenverdeling van gemiddelden bevat elke waarde het gemiddelde van een hele steekproef. Bij proporties geeft elke waarde het aandeel met een bepaald kenmerk in één steekproef.

Bij voldoende grote aselecte, voldoende onafhankelijke steekproeven is de verdeling van gemiddelden of proporties doorgaans ongeveer normaal. Individuele waarnemingen hoeven daarvoor niet normaal te zijn. Wat voldoende groot is, hangt onder andere af van scheefheid, uitschieters en hoe zeldzaam een kenmerk is. Zie voorbeeld 7.2A, blz. 560.

Onthoud: Controleer bij elk diagram eerst wat één waarneming op de horizontale as betekent.

4 Twee bestanden

Bestand A bevat de leestijd van elk van 1200 bezoekers. Bestand B bevat van 400 herhaalde aselecte steekproeven van elk 100 bezoekers de gemiddelde leestijd.

a Wat stelt één waarde in A voor? En in B?

b Hoeveel waarden bevat B en wat is de steekproefomvang per waarde?

5 Scheve tijden

Individuele reparatietijden hebben een lange staart naar hoge waarden en een eindige standaardafwijking. Een computer berekent gemiddelden uit 1000 grote, onafhankelijke aselecte steekproeven.

a Moet de verdeling van die gemiddelden dezelfde scheve vorm hebben? Licht toe.

b Mag je daaruit concluderen dat de individuele reparatietijden normaal zijn?

6 Drie soorten grafieken

Een simulatie heeft 1500 runs. In iedere run worden 250 stemmen aselect uit een groot vast bestand getrokken. Het percentage voorstanders wordt opgeslagen.

a Wat is de grootheid op de horizontale as van het histogram van de opgeslagen uitkomsten?

b Leg uit waarom 1500 geen juiste steekproefomvang voor één opgeslagen percentage is.

c Het kenmerk komt voor bij ongeveer de helft van de populatie. Welke verdelingsvorm is hier aannemelijk bij veel runs?

Grotere steekproeven

Leerdoel

Ik kan uit simulatiegegevens beschrijven hoe de steekproefomvang de spreiding van uitkomsten beïnvloedt.

Theorie

Grotere steekproeven

Grotere aselecte steekproeven leveren meestal minder uiteenlopende proporties en gemiddelden op. De verdeling van de steekproefuitkomsten wordt smaller; het centrum blijft bij dezelfde populatie ongeveer gelijk.

Meer herhalingen geven een stabieler beeld van die verdeling. Dat is iets anders dan een grotere steekproef per herhaling. Een smalle verdeling beschrijft toevalsvariatie en bewijst niet dat de manier van kiezen eerlijk is. Zie voorbeeld 7.2B, blz. 561.

Onthoud: Vergelijk de steekproefomvang per uitkomst, niet alleen het aantal simulaties.

7 **Simulatietabel**

Elke rij vat 5000 onafhankelijke trekkingen uit dezelfde grote populatie samen. De intervallen bevatten de middelste 95% van de gesimuleerde percentages.

Omvang per run	Gemiddeld %	Middelste 95%
25	50	30% tot 70%
100	50	40% tot 60%
400	50	45% tot 55%

a Vergelijk de centra en bereken de breedte van elk interval.

b Welke verandering zie je als de steekproefomvang telkens viermaal zo groot wordt?

8 Een onderzoeksbudget

Voor een vaste populatie zijn de simulatie-uitkomsten hieronder beschikbaar. Eén echte meting kost € 3 per deelnemer. Het team wil een intervalbreedte van hoogstens 12 procentpunt en heeft maximaal € 1500.

n	Middelste 95%
100	40% tot 60%
400	45% tot 55%
1600	47,5% tot 52,5%

a Welke van de getoonde steekproefomvangen voldoet aan beide eisen?

b Iemand stelt voor bij n = 100 viermaal zoveel simulaties uit te voeren en dan dezelfde nauwkeurigheid te verwachten als bij n = 400. Beoordeel dit.

Variatie of vertekening?

Leerdoel

Ik kan toevallige steekproefvariatie onderscheiden van systematische vertekening.

Theorie

Variatie of vertekening?

Bij vertekening worden sommige groepen te veel of te weinig meegenomen, of beïnvloedt de meetwijze de antwoorden. Denk aan vrijwillige deelname, een onvolledig bestand, selectieve non-respons of een sturende vraag.

Een grotere aselecte steekproef verkleint meestal de toevalsvariatie. Zij herstelt geen systematische selectie. Beoordeel daarom eerst wie kon meedoen en wie werkelijk antwoordde. Zie voorbeeld 7.2B, blz. 561.

Onthoud: Een heel precies getal kan toch een verkeerde groep beschrijven.

9 De avondgroep

Een bibliotheek vraagt alleen bezoekers na 19.00 uur of de bibliotheek later open moet blijven. Er antwoorden 2400 bezoekers.

a Welk selectieprobleem ontstaat voor een uitspraak over alle leden?

b Beschrijf een betere manier om leden te benaderen en noem een aandachtspunt bij de antwoorden.

10 Heel veel reacties

Een online peiling met vrijwillige deelname krijgt 50.000 reacties. Een bericht zegt: “Met dit aantal kan er geen vertekening meer zijn.”

a Verbeter de redenering.

b Waarom lost twintigmaal dezelfde peiling herhalen dit niet automatisch op?

11 Het bezoekersrapport

Uit hetzelfde grote bezoekersbestand zijn 800 keer aselect 225 bezoekers gekozen. Van elke steekproef is de gemiddelde bezoekduur bepaald. Er zijn daarnaast 800 gemiddelden berekend met 900 bezoekers per steekproef.

a Verklaar waarom de gemiddelden binnen de eerste reeks verschillen.

b Wat stelt één uitkomst in de tweede reeks voor, en welke vorm is bij veel runs doorgaans aannemelijk?

c Welke reeks zal meestal minder spreiding vertonen?

d Het bezoekersbestand bevat alleen leden met een jaarpas. Welke beperking blijft gelden bij een uitspraak over alle bezoekers?

12 Van 200 naar 800

Een onderzoek loot deelnemers uit een compleet bestand. De onderzoeker verviervoudigt de steekproef van 200 naar 800. Alleen mensen met een uitgesproken mening blijken vaak te reageren.

a Welke verbetering is bij volledige respons aannemelijk en welk probleem ontstaat bij deze respons?

7.3 Betrouwbaarheidsintervallen voor proporties

Leerdoel

Ik kan de steekproefproportie, steekproefomvang en bijbehorende populatie vaststellen.

Theorie

De juiste proportie

Een proportie is het aandeel met een kenmerk, uitgedrukt als getal tussen 0 en 1. Bij k eenheden met het kenmerk uit n onderzochte eenheden is $p = \frac{k}{n}$. De noemer hoort bij de groep waarover de vraag gaat.

De steekproefproportie is bekend uit de onderzochte gegevens. De populatieproportie gaat over de hele bedoelde groep en moet meestal worden geschat. Niet iedereen die een uitnodiging ontvangt, heeft ook een bruikbaar antwoord gegeven. Zie voorbeeld 7.3A, blz. 562.

Onthoud: Een percentage van 37% voer je in de intervalformule in als 0,37.

1 Een nieuwe ophaalplek

Er worden 500 inwoners uitgenodigd. Van de 400 geldige antwoorden zijn er 148 vóór een nieuwe ophaalplek.

a Welke n en p horen bij de geldige antwoorden?

b Welke aanvullende vraag stel je vóór je dit aandeel op alle inwoners betrekt?

2 De relevante groep

Een aselecte steekproef van 800 reizigers bevat 320 treinreizigers. Van hen gebruiken er 128 een abonnement. Van de overige reizigers hebben er 96 een abonnement.

a Bepaal het aandeel abonnementhouders onder de treinreizigers.

b Bepaal het aandeel abonnementhouders onder alle onderzochte reizigers.

3 Een definitie veranderen

Een onderzoek naar tevredenheid bevat 600 geldige antwoorden: 270 tevreden, 180 neutraal en 150 ontevreden.

a Bepaal de proportie tevreden deelnemers.

b Een verslag rekent met $\frac{270}{420}$ en noemt dat eveneens “het aandeel tevreden deelnemers”. Leg uit welk aandeel dit wél beschrijft.

Een interval berekenen

Leerdoel

Ik kan een benaderend 95%-betrouwbaarheidsinterval voor een populatieproportie berekenen en de voorwaarden beoordelen.

Theorie

Een interval berekenen

Gebruik de formule $p \pm 2\sqrt{\frac{p(1-p)}{n}}$. Bereken eerst de foutmarge: het getal achter ±. Trek die van p af en tel die bij p op. Dit geeft twee grenzen voor de geschatte populatieproportie.

De benadering vraagt een geschikte aselecte steekproef en voldoende onafhankelijke waarnemingen. Er moeten voldoende waarnemingen mét en zonder het kenmerk zijn. Als praktische controle gebruiken we hier $np \geq 10$ en $n(1-p) \geq 10$. Selectieve respons wordt door deze formule niet gecorrigeerd. Zie voorbeeld 7.3A, blz. 562.

Onthoud: De formule beschrijft toevalsvariatie, niet alle mogelijke onderzoeksfouten.

4 Een eerste interval

Van 400 aselect gekozen gebruikers zeggen er 160 ja. Er is volledige respons. De populatie is groot en de waarnemingen zijn voldoende onafhankelijk.

a Bepaal p en controleer de aantallen met en zonder het kenmerk.

b Bereken de foutmarge en de intervalgrenzen. Geef de grenzen op vier decimalen.

5 De reservevoorziening

Uit een grote populatie worden 625 huishoudens aselect onderzocht, met volledige respons en voldoende onafhankelijkheid. 275 hebben een reservevoorziening.

a Bereken een benaderend 95%-interval voor de populatieproportie; geef grenzen op vier decimalen.

b Welke twee aantallen controleer je bij de normale benadering?

6 Een zeldzaam probleem

Van 80 aselect onderzochte apparaten heeft er één een storing. De rest werkt goed.

a Voldoet de steekproef aan de praktische aantallencontrole in de theorie?

b De formule geeft een negatieve ondergrens. Is dat een mogelijke populatieproportie? Wat betekent dit voor het gebruik van dit interval?

Proporties en percentages

Leerdoel

Ik kan intervalgrenzen als proporties en percentages met passende afronding weergeven.

Theorie

Proporties en percentages

Reken de foutmarge en de grenzen met ongeronde tussenuitkomsten uit. Vermenigvuldig grenzen met 100 voor percentages. Een foutmarge van 0,04 is 4 procentpunt.

Rond pas af zoals gevraagd. Gebruik voor het beoordelen van een scherpe grens de ongeronde uitkomsten. Een afgeronde ondergrens kan anders ten onrechte precies gelijk lijken aan een getoetste waarde. Zie voorbeeld 7.3A, blz. 562.

Onthoud: De foutmarge is de halve intervalbreedte.

7 Een interval vertalen

Een berekening levert als interval [0,4637; 0,5763].

a Geef de grenzen in procenten, op één decimaal.

b Bereken de foutmarge in procentpunten vóór het afronden.

8 Het ledenpanel

Een aselect onderzoek uit een grote populatie bevat 900 geldige, voldoende onafhankelijke antwoorden. Alle geselecteerde leden hebben geantwoord; 558 antwoorden ja.

a Bereken het 95%-interval en geef het in procenten op één decimaal.

b Bereken de foutmarge in procentpunten en de totale breedte. Rond beide af op één decimaal.

9 Eerst afronden?

Een onderzoek heeft n = 360 en k = 191. De opzet voldoet aan de voorwaarden. Onderzoeker A rekent met $p = \frac{191}{360}$; onderzoeker B rondt p eerst af op 0,5.

a Bereken het interval met de oorspronkelijke gegevens, in procenten op één decimaal.

b Leg uit waarom de werkwijze van B onnodig informatie verliest.

10 Een grensgeval

In een geschikte aselecte steekproef met volledige respons zijn 560 van de 1000 deelnemers voorstander. Gebruik de intervalformule.

a Bereken de ongeronde ondergrens en geef die als percentage op één decimaal.

b Een verslag gebruikt de afgeronde ondergrens om te stellen dat 52,9% onder het hele interval ligt. Controleer dit met de ongeronde grens.

Een uitspraak beoordelen

Leerdoel

Ik kan een populatieclaim met een interval beoordelen en de betekenis van 95% betrouwbaarheid uitleggen.

Theorie

Een uitspraak beoordelen

Een waarde binnen het interval is verenigbaar met de steekproef volgens deze methode; zij is daarmee niet bewezen. Een waarde buiten het interval past minder goed. Als het hele interval boven 0,50 ligt, ondersteunt het onderzoek de uitspraak dat een meerderheid het kenmerk heeft, onder de onderzoeksvoorwaarden.

Bij veel herhaalde geschikte steekproeven zal ongeveer 95% van de zo berekende intervallen de vaste populatieproportie bevatten. Dit betekent niet dat 95% van de personen in één interval valt, of dat één gerealiseerd interval een garantie geeft. Zie voorbeeld 7.3B, blz. 563.

Onthoud: Betrouwbaarheid hoort bij de herhaalde werkwijze.

11 Een waarde in het interval

Een geschikte peiling rapporteert als 95%-interval 0,46 tot 0,54.

a Is een populatieproportie van 0,50 hiermee verenigbaar? Is die waarde bewezen?

b Ondersteunt het hele interval de uitspraak dat meer dan de helft voorstander is?

12 Honderd nieuwe peilingen

Een organisatie herhaalt een geschikte onderzoeksprocedure 100 keer voor dezelfde vaste populatie. Iedere keer wordt een nieuw 95%-interval berekend.

a Hoeveel intervallen verwacht je ongeveer die de populatieproportie bevatten? Is dat aantal gegarandeerd?

b Een medewerker zegt: “In elk interval valt 95% van de deelnemers.” Verbeter de uitspraak.

13 De verkeerde wortel

Bij p = 0,60 en n = 400 berekent iemand de foutmarge als $2 \times \frac{\sqrt{(0{,}60 \times 0{,}40)}}{400}$.

a Waar zit de fout? Bereken de juiste foutmarge.

b Leg uit waarom een veel te kleine foutmarge tot een te sterke conclusie kan leiden.

14 Een meerderheidsclaim

In een aselecte steekproef van 600 inwoners steunen 354 inwoners een voorstel. De populatie is groot; er is volledige respons en voldoende onafhankelijkheid.

a Bepaal p en bereken het 95%-interval in procenten op één decimaal.

b Ondersteunt dit interval de uitspraak dat een meerderheid van de populatie het voorstel steunt? Licht toe.

15 Een toelichting voor lezers

Een correct berekend 95%-interval voor een populatieproportie is 0,521 tot 0,599.

a Beoordeel de claim dat de populatieproportie 0,61 is.

b Schrijf een correcte uitleg van 95% betrouwbaarheid voor herhaald onderzoek.

16 Een vrijwillige enquête

Een vrijwillige enquête geeft 64% voorstanders en een met de formule berekend smal interval van 63% tot 65%.

a Waarom mag de smalle uitkomst op zichzelf niet als betrouwbare schatting voor alle inwoners worden gepresenteerd?

7.4 Intervallen voor gemiddelden, omvang en nauwkeurigheid

Leerdoel

Ik kan een benaderend 95%-betrouwbaarheidsinterval voor een populatiegemiddelde berekenen en de voorwaarden beoordelen.

Theorie

Het populatiegemiddelde schatten

Gebruik $\overline{X} \pm 2\frac{S}{\sqrt{n}}$. Hier is $\overline{X}$ het steekproefgemiddelde, S de standaardafwijking van de individuele waarnemingen in de steekproef en n de steekproefomvang. De grenzen hebben dezelfde eenheid als $\overline{X}$ en S.

De steekproef moet passend aselect zijn, met voldoende onafhankelijke waarnemingen en een voldoende normale benadering voor de steekproefgemiddelden. Grote steekproeven helpen daarbij; sterke uitschieters en selectieve respons vragen extra aandacht. Het interval schat het gemiddelde van de populatie, niet de grenzen voor individuele waarden. Zie voorbeeld 7.4A, blz. 564.

Onthoud: Gebruik $\frac{S}{\sqrt{n}}$ in de foutmarge; S alleen beschrijft de spreiding van individuele metingen.

1 De wachtruimte

Een geschikte aselecte steekproef van 100 bezoekers heeft gemiddelde wachttijd 18,4 minuten en standaardafwijking 6,5 minuten. De normale benadering voor gemiddelden is geschikt.

a Bereken eerst $\frac{S}{\sqrt{n}}$ en daarna de foutmarge.

b Bereken het 95%-interval en benoem wat je ermee schat.

2 Een vulsysteem

Uit een groot productieproces zijn 144 verpakkingen aselect en voldoende onafhankelijk bemonsterd. $\overline{X}$ = 502 gram en S = 9 gram; de benadering voor gemiddelden is geschikt.

a Bereken het interval voor het populatiegemiddelde.

b Waarom mag je niet zeggen dat 95% van de verpakkingen tussen deze grenzen weegt?

3 De standaardinstelling

Een geschikt aselect onderzoek bevat 225 onderdelen, met gemiddelde lengte 79,2 mm en standaardafwijking 3,6 mm. De normale benadering voor gemiddelden is geschikt. De fabrikant claimt een populatiegemiddelde van 80 mm.

a Bereken het interval en beoordeel de claim.

b Een medewerker stelt voor alle lengten als één batchgemiddelde in te voeren met n = 1. Waarom verliest hij daarmee noodzakelijke informatie?

Foutmarge en breedte

Leerdoel

Ik kan foutmarge en intervalbreedte onderscheiden en grenzen met de juiste eenheid interpreteren.

Theorie

Foutmarge en breedte

De foutmarge m is de afstand van het midden tot één grens. De intervalbreedte is 2m. Bij grenzen L en U geldt $\overline{X} = \frac{(L + U)}{2}$ en $m = \frac{(U - L)}{2}$.

Bij veel herhaalde geschikte steekproeven bevat ongeveer 95% van de berekende intervallen het vaste populatiegemiddelde. Een interval is geen garantie voor een individuele meting. Zie voorbeeld 7.4A, blz. 564.

Onthoud: Een interval van 12 tot 18 heeft breedte 6 en foutmarge 3.

4 Terug naar het midden

Een interval voor een gemiddelde verwerkingstijd loopt van 7,8 tot 9,0 minuten.

a Bepaal het midden, de foutmarge en de breedte.

b Een nieuw interval met hetzelfde midden is half zo breed. Geef de nieuwe grenzen.

5 Seconden in het verslag

Een geschikte steekproef levert een gemiddelde van 135 seconden en een 95%-interval van 129 tot 141 seconden.

a Geef het gemiddelde, de foutmarge en de grenzen in minuten.

b Verandert de relatieve nauwkeurigheid door deze andere eenheid?

6 Midden en spreiding terugvinden

Een met de standaardformule berekend interval loopt van 24,6 tot 25,8 gram. De steekproefomvang is 400.

a Bepaal het steekproefgemiddelde en de foutmarge.

b Bereken de gebruikte steekproefstandaardafwijking S.

De omvang veranderen

Leerdoel

Ik kan met een intervalformule bepalen hoe de steekproefomvang de foutmarge beïnvloedt.

Theorie

De omvang veranderen

Als S ongeveer gelijk blijft, is $m = \frac{2S}{\sqrt{n}}$. Bij viermaal zoveel waarnemingen wordt $\sqrt{n}$ tweemaal zo groot en m half zo groot. Bij negenmaal zoveel waarnemingen wordt m driemaal zo klein.

Voor twee onderzoeksplannen met dezelfde S geldt: $m_{nieuw} = m_{oud} \times \sqrt{(\frac{n_{oud}}{n_{nieuw}})}$. Dit is een planningsvergelijking. De feitelijk gemeten standaardafwijking kan in nieuw onderzoek anders uitvallen. Zie voorbeeld 7.4B, blz. 565.

Onthoud: Een halvering van de marge vraagt viermaal zoveel waarnemingen als S gelijk blijft.

7 Een plan doorrekenen

Met S = 12 minuten is bij n = 100 de foutmarge 2,4 minuten.

a Bereken de marge bij n = 400 en n = 900, als S gelijk blijft.

b Welke aanname is nodig om alleen de verandering van n te vergelijken?

8 Verdubbelen is niet halveren

Een foutmarge is 3,0 minuten bij n = 80. S wordt bij het nieuwe onderzoeksplan gelijk verondersteld.

a Bereken de verwachte marge bij n = 160 op twee decimalen.

b Welke omvang is nodig om de marge precies te halveren?

9 Twee veranderingen tegelijk

Onderzoek A heeft S = 8 en n = 100. Onderzoek B heeft S = 12 en n = 400. De eenheden en de betrouwbaarheidsfactor zijn gelijk.

a Vergelijk de foutmarges.

b Waarom wordt de marge ondanks de verviervoudigde omvang niet gehalveerd?

10 Een onderzoekskeuze

Een onderzoek schat met n = 144 en S = 18 uur een gemiddelde levensduur. Een extra meting kost € 8. S wordt voor de planning gelijk gehouden. Het team kan maximaal € 2400 extra besteden.

a Kan het team met dit budget de foutmarge halveren?

b Bereken de verwachte marge bij volledig gebruik van het extra budget. Rond af op twee decimalen.

Een benodigde omvang berekenen

Leerdoel

Ik kan bij een gewenste maximale foutmarge de benodigde steekproefomvang berekenen en naar boven afronden.

Theorie

Een benodigde omvang berekenen

Bij $m = \frac{2S}{\sqrt{n}}$ en een maximale foutmarge $d > 0$ zoek je n waarvoor $\frac{2S}{\sqrt{n}} \leq d$. Dat geeft $n \geq (\frac{2S}{d})^2$.
Gebruik een onderbouwde schatting voor S en rond de benodigde n naar boven af.

Bij proporties is $m = 2\sqrt{(\frac{p(1-p)}{n})}$. Met een planningswaarde voor p geeft dit $n \geq \frac{4p(1-p)}{d^2}$.
Zonder bruikbare schatting geeft $p = 0{,}50$ de grootste benodigde omvang: $p(1-p)$ is dan 0,25. Hier is d een proportie, geen percentagegetal. Zie voorbeeld 7.4B, blz. 565.

Onthoud: Controleer het gekozen gehele aantal én, bij een minimumvraag, één waarneming minder.

11 Een meetplan

Een geschikt meetonderzoek wordt gepland met $S = 7{,}5$ gram. De maximale foutmarge voor het gemiddelde mag 0,8 gram zijn.

a Bereken het kleinste benodigde gehele aantal metingen.

b Controleer dat één meting minder onvoldoende is.

12 Een onbekend aandeel

Een aselecte peiling moet worden gepland met maximale foutmarge 0,025. Voor de populatieproportie is vooraf geen bruikbare schatting beschikbaar.

a Bereken de benodigde omvang met de veilige planningswaarde $p = 0{,}50$.

b Er wordt 80% respons verwacht. Hoeveel uitnodigingen worden dan minimaal gepland om gemiddeld dat aantal antwoorden te krijgen? Geeft dat een garantie?

13 Een al bekende proportie

Voor de planning van een geschikte peiling is p = 0,30 gekozen. De maximale foutmarge is 3 procentpunt.

a Bereken de benodigde omvang.

b Hoe verandert de veilige benodigde omvang als p vooraf onbekend is?

14 Een offerte kiezen

Voor een meetonderzoek schat men S op 4,2 minuten. De gewenste maximale foutmarge is 0,5 minuut. Elke meting kost € 6. Een offerte biedt 280 metingen voor € 1680.

a Voldoet de offerte aan de gewenste marge?

b Bepaal de minimale extra kosten bij hetzelfde tarief en noem de aanname over S.

15 Verkeerd afgerond

Een berekening geeft $n \geq 187{,}06$. Een onderzoeker schrijft: “n = 187, want 0,06 rond je naar beneden af.”

a Leg uit waarom deze afronding niet past en geef het juiste minimum.

b Dezelfde onderzoeker gebruikt bij een proportiemarge van 4 procentpunt d = 4. Verbeter dit.

16 Lampen testen

Een geschikte aselecte steekproef van 196 lampen heeft $\overline{X} = 810$ uur en S = 28 uur. De normale benadering voor het gemiddelde is geschikt.

a Bereken het interval, de foutmarge en de breedte.

b Welk aantal is nodig voor een marge van 2 uur als S gelijk blijft?

17 Een strakkere planning

Bij een gepland onderzoek wordt S geschat op 11 eenheden. Men wil een intervalbreedte van hoogstens 3 eenheden.

a Bereken de kleinste benodigde steekproefomvang.

b Waarom is dit aantal een planningsuitkomst en geen onvoorwaardelijke belofte?

18 De kleinste marge

Twee geschikte plannen gebruiken dezelfde meeteenheid. Plan A: S = 5 en n = 100. Plan B: S = 9 en n = 400.

a Kies het plan met de kleinste verwachte foutmarge en onderbouw.

b Hoe groot moet n bij S = 9 minstens zijn om marge 0,8 te halen?

Alles door elkaar

Gebruik zo nodig het formuleblad verderop. Geef steeds aan welke grootheid je berekent.

1 De verpakkingslijn

Een normaal model voor massa heeft $\mu = 300$ gram en $\sigma = 8$ gram.

a Geef het middelste 95%-interval en het aandeel boven de bovengrens.

b Een geschikte steekproef van 100 verpakkingen heeft $\overline{X} = 302$ gram en $S = 8$ gram. Bereken het interval voor het populatiegemiddelde. De normale benadering is geschikt.

c Leg uit waarom de twee intervallen verschillende breedten hebben.

2 Het geluidsonderzoek

In een geschikte aselecte steekproef van 625 bewoners zeggen 350 bewoners hinder te ervaren. Er is volledige respons; de populatie is groot en de waarnemingen zijn voldoende onafhankelijk.

a Bereken het 95%-interval in procenten op één decimaal.

b Ondersteunt het hele interval een meerderheid?

c Waarom zou een vrijwillige klachtenwebsite, ook met veel meer reacties, dit onderzoek niet eenvoudig vervangen?

3 Een meetonderzoek uitbreiden

Een geschikt onderzoek heeft $\overline{X} = 40$, $S = 10$ en $n = 100$. De normale benadering is geschikt.

a Geef het 95%-interval, de foutmarge en de breedte.

b Welke omvang is bij dezelfde S nodig voor breedte hoogstens 2?

4 Een reeks uitkomsten

Een simulatie van aselecte steekproeven uit een vaste grote populatie levert voor $n = 64$ een veel bredere verdeling van gemiddelden op dan voor $n = 256$. Beide verdelingen liggen rond 75.

a Verklaar het gelijke centrum en de verschillende spreiding.

b De individuele waarden hebben twee duidelijke toppen. Waarom sluiten ongeveer normale steekproefgemiddelden dit niet uit?

Gericht verder oefenen

Kies op basis van je gemaakte werk welke opgaven je nodig hebt.

1 Gebieden aan weerszijden

Een normaal model heeft $\mu = 60$ en $\sigma = 5$.

a Bepaal eerst de grenzen $\mu - \sigma$ en $\mu + \sigma$. Welk aandeel ligt daarbuiten?

b Welk aandeel ligt boven 65? Gebruik de symmetrie.

2 De juiste invoer

In een geschikte aselecte steekproef zijn 96 van de 240 deelnemers voorstander. De respons is volledig en de waarnemingen zijn voldoende onafhankelijk.

a Bepaal n en p. Bereken daarna de foutmarge en de twee grenzen op vier decimalen.

b Geef het interval in procenten op één decimaal en controleer of 50% erin ligt.

3 Van breedte naar omvang

Voor een geschikt meetplan is S = 6 minuten geschat. De maximale intervalbreedte is 2 minuten.

a Welke maximale foutmarge hoort daarbij?

b Vul in $n \geq (\frac{2S}{d})^2$ en bepaal de benodigde omvang.

c Wat gebeurt er met de marge als n bij dezelfde S wordt verviervoudigd?

Nieuwe controle

Maak deze vragen zonder terug te kijken naar de herstelopgaven.

1 Het onthaalteam

Een geschikte peiling heeft 784 geldige antwoorden, waarvan 392 positief. De respons is volledig, de populatie is groot en de antwoorden zijn voldoende onafhankelijk.

a Bereken het 95%-interval in procenten op één decimaal.

b Ondersteunt dit interval de uitspraak dat meer dan de helft positief is? Leg uit.

2 Een nieuwe tijdmeting

Een geschikte steekproef van 81 waarnemingen heeft $\overline{X} = 26$ minuten en $S = 4{,}5$ minuten. De normale benadering is geschikt.

a Bereken het interval en de breedte.

b Bereken de kleinste omvang voor een maximale marge van 0,7 minuut bij dezelfde S.

Formuleblad

Dit blad mag je bij Test jezelf gebruiken. $\overline{X}$ is het steekproefgemiddelde, S de steekproefstandaardafwijking, p de steekproefproportie en n de steekproefomvang.

95%-betrouwbaarheidsinterval voor een populatieproportie: $p \pm 2\sqrt{\frac{p(1-p)}{n}}$.

95%-betrouwbaarheidsinterval voor een populatiegemiddelde: $\overline{X} \pm 2\frac{S}{\sqrt{n}}$.

Leerdoelenoverzicht

7.1 · Leerdoel 1: Ik kan verdelingen vergelijken op vorm, centrum, spreiding, staarten en uitschieters.

7.1 · Leerdoel 2: Ik kan bij een normale verdeling het gemiddelde en de standaardafwijking verbinden aan de ligging en spreiding.

7.1 · Leerdoel 3: Ik kan de normale vuistregels toepassen op intervallen rond het gemiddelde.

7.1 · Leerdoel 4: Ik kan met symmetrie en normale vuistregels een aandeel of grens benaderen en interpreteren.

7.2 · Leerdoel 1: Ik kan verklaren waarom aselecte steekproeven uit dezelfde populatie verschillende uitkomsten geven.

7.2 · Leerdoel 2: Ik kan individuele waarnemingen onderscheiden van steekproefgemiddelden en -proporties en de normale benadering uitleggen.

7.2 · Leerdoel 3: Ik kan uit simulatiegegevens beschrijven hoe de steekproefomvang de spreiding van uitkomsten beïnvloedt.

7.2 · Leerdoel 4: Ik kan toevallige steekproefvariatie onderscheiden van systematische vertekening.

7.3 · Leerdoel 1: Ik kan de steekproefproportie, steekproefomvang en bijbehorende populatie vaststellen.

7.3 · Leerdoel 2: Ik kan een benaderend 95%-betrouwbaarheidsinterval voor een populatieproportie berekenen en de voorwaarden beoordelen.

7.3 · Leerdoel 3: Ik kan intervalgrenzen als proporties en percentages met passende afronding weergeven.

7.3 · Leerdoel 4: Ik kan een populatieclaim met een interval beoordelen en de betekenis van 95% betrouwbaarheid uitleggen.

7.4 · Leerdoel 1: Ik kan een benaderend 95%-betrouwbaarheidsinterval voor een populatiegemiddelde berekenen en de voorwaarden beoordelen.

7.4 · Leerdoel 2: Ik kan foutmarge en intervalbreedte onderscheiden en grenzen met de juiste eenheid interpreteren.

7.4 · Leerdoel 3: Ik kan met een intervalformule bepalen hoe de steekproefomvang de foutmarge beïnvloedt.

7.4 · Leerdoel 4: Ik kan bij een gewenste maximale foutmarge de benodigde steekproefomvang berekenen en naar boven afronden.

Eindevaluatie

Kijk je werk na met de uitwerkingen. Noteer per leerdoel: dit lukt / dit oefen ik nog. Wijs bij “dit lukt” een passende deelvraag aan in je eigen werk.

Vorm en normale vuistregels: 1a en 3a-c → paragraaf 7.1 en herstelopgave 1. Steekproeven: 1b, 2a, 3d en 4a → paragraaf 7.2.

Proportie-intervallen: 1d en 2b-d → paragraaf 7.3 en herstelopgave 2. Gemiddelden en omvang: 1c en 4b-d → paragraaf 7.4 en herstelopgave 3.

Maak daarna de passende Nieuwe controle: opgave 1 voor proporties en opgave 2 voor gemiddelden en omvang. Kies bij verdelingen en steekproefvariatie een nog niet gemaakte opgave uit 7.1 of 7.2.

Eindevaluatie

Kijk na met de uitwerkingen. Kies bij een fout de bijbehorende opgaven onder Oefen opnieuw. Maak daarna de opgaven onder Controleer zonder hulp.

Toetsvraag	Leerdoel	Oefen opnieuw	Controleer
Test jezelf opgave 1a; Test jezelf opgave 3a–c	Verdelingsvorm, μ/σ en normale vuistregels	7.1 opgave 1a–b; 7.1 opgave 5a–b; Gericht verder oefenen opgave 1a–b	7.1 opgave 15a–c; 7.1 opgave 16a–b; 7.1 opgave 4a
Test jezelf opgave 1b; Test jezelf opgave 2a; Test jezelf opgave 4a	Steekproefvariatie, normale benadering en omvang per run	7.2 opgave 1a–b; 7.2 opgave 4a–b; 7.2 opgave 7a–b	7.2 opgave 11a–c
Test jezelf opgave 3d	Variatie tegenover selectievertekening	7.2 opgave 9a–b; 7.2 opgave 10a–b	7.2 opgave 11d; 7.2 opgave 12a
Test jezelf opgave 2b–d	Proportie-interval, relevante n, voorwaarden en percentages	Gericht verder oefenen opgave 2a–b; 7.3 opgave 2a–b; 7.3 opgave 6a–b	Nieuwe controle opgave 1a; 7.3 opgave 5b
Test jezelf opgave 1d	Populatieclaim en betekenis van 95%	7.3 opgave 11a–b; 7.3 opgave 12a–b	Nieuwe controle opgave 1b; 7.3 opgave 15a–b
Test jezelf opgave 1c; Test jezelf opgave 4b–d	Gemiddeldeninterval, breedte, n-effect en benodigde omvang	7.4 opgave 1a–b; Gericht verder oefenen opgave 3a–c	Nieuwe controle opgave 2a–b; 7.4 opgave 16a–b
7.4 opgave 12a-b	Benodigde omvang voor proportie met bekende of onbekende p	7.4 opgave 15a-b	7.4 opgave 13a–b

Voorbeeld 7.1A

Een opvallende meting

Zeven metingen zijn 14, 15, 15, 16, 16, 17 en 33 seconden.

a Beschrijf het centrum en de opvallende waarde. Kies een centrummaat voor een gebruikelijke meting.

1. Wat zoek ik?

Een centrummaat die een gebruikelijke tijd beschrijft, met aandacht voor 33.

2. Welke basis en aanpak kies ik, en waarom?

Orden de waarden en vergelijk mediaan en gemiddelde; een veraf gelegen hoge waarde beïnvloedt het gemiddelde sterker.

3. Berekening

De mediaan is de vierde waarde: 16. De som is 126, dus het gemiddelde is 18. De waarde 33 ligt ver boven 14 tot 17.

4. Antwoord in context

De mediaan van 16 seconden beschrijft de gebruikelijke meting hier beter. 33 seconden is een mogelijke uitschieter.

5. Controle

Zes van de zeven waarden liggen tussen 14 en 17; 16 ligt in dat gebied. Onderzoek de meting van 33 voordat je die eventueel uitsluit.

Voorbeeld 7.1B

Vullingen beoordelen

De vulmassa volgt een normaal model met $\mu = 480$ gram en $\sigma = 6$ gram. Er zijn 1200 verpakkingen.

a Geef het middelste 95%-interval en bereken het verwachte aantal verpakkingen boven de bovengrens.

1. Wat zoek ik?

Twee massagrenzen en een verwacht aantal in de bovenste staart.

2. Welke basis en aanpak kies ik, en waarom?

Het model is normaal. Twee standaardafwijkingen aan elke kant geven ongeveer 95%; de resterende 5% wordt door symmetrie gelijk verdeeld.

3. Berekening

$\mu \pm 2\sigma = 480 \pm 12$, dus 468 en 492. Boven 492 ligt ongeveer 2,5%; $0{,}025 \times 1200 = 30$.

4. Antwoord in context

Ongeveer 95% ligt tussen 468 en 492 gram. Ongeveer 30 verpakkingen zijn zwaarder dan 492 gram.

5. Controle

De grenzen liggen elk 12 gram van 480; 30 is 2,5% van 1200. De verwachte aantallen aan beide kanten zijn gelijk, niet noodzakelijk de werkelijk getelde aantallen.

Voorbeeld 7.2A

Eén punt, veel personen

Een computer trekt 600 keer een aselecte steekproef van 144 personen uit een grote vaste populatie. Per trekking wordt de gemiddelde reistijd bewaard.

a Leg uit wat één waarde in het resultatenbestand voorstelt en waarom de waarden kunnen verschillen.

1. Wat zoek ik?

De betekenis van één opgeslagen waarde en de oorzaak van verschillen.

2. Welke basis en aanpak kies ik, en waarom?

Maak onderscheid tussen personen per steekproef en het aantal herhalingen. Elke trekking heeft een andere samenstelling.

3. Berekening

Er zijn 600 waarden. Elke waarde is het gemiddelde van 144 reistijden; verschillende gelote personen leveren verschillende gemiddelden.

4. Antwoord in context

Eén waarde is een steekproefgemiddelde van 144 personen. De variatie tussen de 600 gemiddelden ontstaat door het loten.

5. Controle

600 is het aantal gemiddelden, niet de omvang van één steekproef. De populatie hoeft niet te veranderen om verschillen te krijgen.

Voorbeeld 7.2B

Twee simulatiereeksen

Bij n = 100 liggen de middelste 95% van gesimuleerde percentages tussen 32% en 48%. Bij n = 400 liggen ze tussen 36% en 44%. Beide reeksen gebruiken dezelfde grote populatie.

a Vergelijk de nauwkeurigheid en geef aan of dit bewijst dat een peiling met vrijwillige deelname representatief is.

1. Wat zoek ik?

Het verschil in spreiding en de beperking van de conclusie.

2. Welke basis en aanpak kies ik, en waarom?

Vergelijk intervalbreedten bij dezelfde populatie. Houd toevalsvariatie gescheiden van de selectie van deelnemers.

3. Berekening

De centra zijn beide 40%. De breedte daalt van 16 naar 8 procentpunt wanneer n verviervoudigt.

4. Antwoord in context

De grotere steekproeven leveren minder verspreide uitkomsten. Dit zegt niet dat vrijwillige deelnemers representatief zijn.

5. Controle

Beide intervallen liggen rond 40%; de lagere breedte beschrijft alleen het aselecte simulatiemodel.

Voorbeeld 7.3A

De volledige respons

Een grote vereniging loot 500 leden. Iedereen antwoordt; 231 leden willen een extra afhaalmoment. De antwoorden zijn voldoende onafhankelijk.

a Bereken een benaderend 95%-interval voor het aandeel onder alle leden, in procenten op één decimaal.

1. Wat zoek ik?

Twee grenzen voor de onbekende populatieproportie.

2. Welke basis en aanpak kies ik, en waarom?

Gebruik $n = 500$ en $p = \frac{231}{500} = 0{,}462$. Er zijn 231 ja en 269 nee; de aantallen en onderzoeksopzet zijn geschikt voor de normale benadering.

3. Berekening

$p = \frac{231}{500} = 0{,}4620$. De foutmarge is $2 \times \sqrt{(0{,}4620 \times \frac{0{,}5380}{500})} \approx 0{,}0446$. De ongeronde grenzen zijn ongeveer 0,417408 en 0,506592.

4. Antwoord in context

Het interval is ongeveer [0,4174; 0,5066], oftewel 41,7% tot 50,7%.

5. Controle

Het interval ligt rond 46,2%, de grenzen liggen binnen 0% en 100%, en de foutmarge is ongeveer 4,5 procentpunt. De uitkomst schat het aandeel leden, niet hun aantal.

Voorbeeld 7.3B

Een uitspraak onderbouwen

Een geschikte peiling levert voor het aandeel voorstanders het 95%-interval 0,535 tot 0,605.

a Beoordeel “een meerderheid is voor” en licht de betrouwbaarheid toe.

1. Wat zoek ik?

Of het hele interval een meerderheid ondersteunt, en wat 95% inhoudt.

2. Welke basis en aanpak kies ik, en waarom?

Een meerderheid betekent meer dan 0,50. Vergelijk die grens met de ongeronde intervalgrenzen.

3. Berekening

0,535 > 0,50; dus elk aandeel in het gegeven interval is groter dan de helft.

4. Antwoord in context

De peiling ondersteunt een meerderheid. Bij veel herhalingen van deze geschikte procedure bevat ongeveer 95% van de berekende intervallen de populatieproportie.

5. Controle

Er blijven onderzoeksvoorwaarden en toevalsrisico bestaan. Een interval is geen bewijs dat de populatieproportie exact het midden 0,57 is.

Voorbeeld 7.4A

Leestijd schatten

Een geschikte aselecte steekproef van 256 bezoekers heeft $\overline{X} = 64$ seconden en $S = 12$ seconden. De normale benadering voor steekproefgemiddelden is geschikt.

a Bereken het 95%-interval, de foutmarge en de breedte.

1. Wat zoek ik?

De grenzen voor de gemiddelde leestijd van de populatie en de bijbehorende nauwkeurigheid.

2. Welke basis en aanpak kies ik, en waarom?

Gebruik de formule voor een gemiddelde. S beschrijft individuele tijden; delen door $\sqrt{n}$ vertaalt de spreiding naar gemiddelden.

3. Berekening

$m = 2 \times \frac{12}{\sqrt{256}} = \frac{24}{16} = 1{,}5$. De grenzen zijn $64 - 1{,}5 = 62{,}5$ en $64 + 1{,}5 = 65{,}5$.

4. Antwoord in context

Het interval is 62,5 tot 65,5 seconden. De foutmarge is 1,5 seconde en de breedte 3 seconden.

5. Controle

Het midden blijft 64. Het interval gaat over het populatiegemiddelde en is geen interval waarin 95% van de individuele leestijden moet liggen.

Voorbeeld 7.4B

Genoeg metingen plannen

Voor een nieuw onderzoek wordt S op 8,4 seconden geschat. De maximale foutmarge is 1 seconde. De opzet is geschikt.

a Bereken het minimumaantal metingen en controleer het minimum.

1. Wat zoek ik?

Een geheel aantal n dat de maximale foutmarge niet overschrijdt.

2. Welke basis en aanpak kies ik, en waarom?

Los $\frac{2S}{\sqrt{n}} \leq 1$ op met S = 8,4. Omdat er minstens genoeg metingen nodig zijn, rond ik naar boven af.

3. Berekening

$n \geq (\frac{16{,}8}{1})^2 = 282{,}24$. Bij n = 283 is m ≈ 0,99866; bij n = 282 is m ≈ 1,00043. Bij een halvering van de gewenste marge zou het ongeronde benodigde aantal viermaal zo groot worden.

4. Antwoord in context

Er zijn minstens 283 metingen nodig binnen dit planningsmodel.

5. Controle

283 voldoet en 282 niet. De planning veronderstelt dat S = 8,4 een bruikbare schatting blijft.