Statistiek voor tandartsen (IV)

Daar gooien we om; één steekproef, dichotome meting

A.H.B. Schuurs, tandarts[1]
M.A. van 't Hof, statisticus[2]
H.J. Duivenvoorden, methodoloog[3]

Uit de [1]vakgroep Cariologie en Endodontologie van het Academisch Centrum Tandheelkunde Amsterdam (ACTA), de [2]Medisch Statistische Afdeling van de faculteit der Geneeskunde te Nijmegen en de [3]vakgroep Medische Psychologie en Psychotherapie, faculteit der Geneeskunde van de Erasmus Universiteit te Rotterdam.

Samenvatting. Bijna iedereen dobbelt wel eens of gooit eens kruis of munt. Hier is beschreven hoe, met behulp van alle theoretisch mogelijke combinaties van kruis en munt (binomiale verdeling), een 'vals' geldstuk is te onderscheiden van een 'eerlijk' door een aantal keren te werpen. Naarmate vaker wordt geworpen neemt de kans op een juiste representatie toe en daarmee de zekerheid of het geldstuk al dan niet vals is. Pas als kruis (of munt) zeer vaak bovenkomt, mag de nulhypothese dat 'het geldstuk eerlijk is' worden verworpen.

SCHUURS AHB, VAN 'T HOF MA, DUIVENVOORDEN HJ. Statistiek voor tandartsen (IV). Daar gooien we om; één steekproef, dichotome meting. Ned Tijdschr Tandheelkd 1991; 98: 94-7.

Trefwoord: Statistiek

Datum van acceptatie: 20 september 1990.

Adres: Dr. A.H.B. Schuurs, ACTA, 1066 EA Amsterdam.

1 Beschrijving probleem

Een tandarts aarzelt of hij een klinische avond zal bijwonen. Enerzijds lijken de aangekondigde onderwerpen interessant, anderzijds is hij moe en blijft liever thuis. Zijn vrouw stelt voor kruis of munt te gooien: komt kruis boven dan zal hij gaan en bij munt niet. Na werpen blijkt kruis boven te liggen. Maar omdat ook de tandarts niets menselijks vreemd is, stelt hij voor driemaal te gooien en de regel 'de meeste keren gelden' toe te passen. Tot driemaal toe blijkt hij kruis te gooien, wat hem doet opmerken dat het gebruikte geldstuk 'vals' is. Maar is dat zo? Die vraag wordt hierna besproken.

2 Aard gegevens

Het hier gebruikte voorbeeld is ontleend aan Siegel en Castellan (1988).[1] De gegevens, afkomstig uit één steekproef, zijn dichotoom gemeten (zie eveneens Statistiek voor tandartsen III, december 1990). Bij het toetsen wordt gebruik gemaakt van de binomiale verdeling, een begrip dat later wordt toegelicht.

3 Formulering vraagstelling

Omdat de tandarts er van uitgaat dat het geldstuk 'eerlijk' is, verwacht hij dat de kans op het gooien van kruis even groot is als de kans op munt (nulhypothese, H_0: $P_{kruis} = P_{munt} = 0{,}5$). Omdat kruis driemaal bovenkwam, zal de tandarts als alternatieve hypothese (H_1) stellen dat de kans op kruis groter is dan 50% ($H_1 : P_{kruis} > 0{,}5$). Via een experiment zal hij trachten H_0 te verwerpen ten gunste van H_1.

4 Welke toets?

Een aselecte steekproef van 100 personen uit een stad, waarin precies evenveel mannen als vrouwen leven, kan bijvoorbeeld 49 mannen en 51 vrouwen omvatten of 45 mannen en 55 vrouwen, in plaats van 50 van beiden. Een steekproef geeft maar zelden precies de verhouding in de populatie weer. Hetzelfde geldt uiteraard voor een geldstuk dat een aantal keren (= steekproef) wordt opgegooid. Kleine afwijkingen zijn niet verrassend, maar grote afwijkingen zijn onwaarschijnlijker, zij het niet onmogelijk. Er kan zelfs 100 maal opeenvolgend kruis en 0 maal munt verschijnen, maar dan rijst er toch wel twijfel of het geldstuk 'eerlijk' is, dus of de steekproef een afspiegeling is van een 'populatie' met de te verwachten verhouding (50:50) of van een andere populatie, namelijk van die van een vals geldstuk.

We moeten dus de populatie kennen. Hoe die *theoretische* populatieverdeling van kruis of munt van een eerlijk geldstuk is samengesteld, is beschreven in paragraaf 6.

5 Het experiment

Het experiment houdt in dat het betreffende geldstuk een aantal malen wordt opgeworpen om gegevens te verzamelen. De tandarts denkt dat 12 maal gooien wel voldoende zal zijn, hetgeen als een kleine steekproef moet worden beschouwd. Hij blijkt tienmaal kruis en tweemaal munt te werpen.

6 Kenmerkende uitkomsten van een eerlijk geldstuk

Als we de vraag van de tandarts omschrijven zou deze luiden: als ik 12 maal met mijn geldstuk gooi, is het resultaat dan globaal in overeenstemming met 12 maal gooien met een eerlijk geldstuk of zijn er afwijkende resultaten van zo grote orde dat mag worden betwijfeld of het geldstuk 'eerlijk' is.

Bij éénmaal opwerpen van een eerlijk geldstuk bestaan er twee in theorie mogelijke uitkomsten: kruis (50%) of munt (50%). (Gemakshalve vergeten we dat het geldstuk op zijn rand kan blijven staan.) Het zou een misverstand zijn te denken dat bij tweemaal werpen de *drie* mogelijke uitkomsten 1. tweemaal kruis boven; 2. éénmaal kruis boven en 3. nul maal kruis boven, een gelijke kans hebben te verschijnen. Wordt tweemaal geworpen, dan zijn *vier* theoretische uitkomsten mogelijk, namelijk:

1. eerste worp kruis, tweede worp kruis,
2. eerste worp kruis, tweede worp munt,
3. eerste worp munt, tweede worp kruis,
4. eerste worp munt, tweede worp munt.

Elk van de vier genoemde uitkomsten treedt bij een eerlijk geldstuk met een waarschijnlijkheid van 25% op. De vier in theorie mogelijke uitkomsten noemt men de *kansverdeling*. Bij driemaal werpen bestaan er acht theoretisch mogelijke combinaties van uitkomsten, elk met een waarschijnlijkheid van 1/8 (tab. I). Uit deze voorbeelden komt naar voren dat het aantal van alle mogelijke combinaties bij:

- éénmaal werpen gelijk is aan 2^1, namelijk éénmaal kruis of eenmaal munt;
- tweemaal werpen is het aantal mogelijke combinaties $2^2 = 4$;
- driemaal werpen $2^3 = 8$.

Na te gaan is dat bij 12 maal werpen, zoals de tandarts deed, het aantal van alle in theorie mogelijke combinaties gelijk is aan $2^{12} = 4096$. Deze 4096 combinaties vertegenwoordigen de kansverdeling; er zijn niet méér mogelijkheden. Een dergelijke kansverdeling wordt *binomiaal* genoemd.

Als alle combinaties zouden worden uitgeschreven, net als in tabel I, dan zou blijken dat in slechts één van die combinaties bij alle 12 worpen kruis bovenligt (en analoog in één combinatie 12 maal munt).

Tabel I toont dat bij een van de acht combinaties driemaal kruis gegooid werd, bij drie van de acht combinaties tweemaal kruis en bij eveneens drie combinaties éénmaal kruis. Wederom één combinatie bestaat uit driemaal kruis. Tabel I kan dan ook op een andere manier worden weergegeven (tab. II). Met de driehoek van Pascal (tab. IIa) zijn de kansen op de verschillende combinaties na te gaan (maar zij zijn ook met een formule te berekenen); vergelijk ter illustratie de percentages van tabel II met de derde regel van tabel IIa.

Analoog aan tabel II, is tabel III opgesteld. Hierin is te zien in hoeveel van de 4096 combinaties 12 maal kruis, 11 maal kruis, tienmaal ... 0 maal kruis bovenligt. In theorie komt bij één van de 4096 combinaties kruis 12 maal boven, dus met een waarschijnlijkheid van 0,00024 (een kans van 0,024%). En bij 12 van de 4096 combinaties zal kruis 11 maal bovenliggen, dus met een waarschijnlijkheid van 12/4096 = 0,0029.

De gegevens van tabel III kunnen ook grafisch worden weergegeven (afb. 1). Op de Y-as is het aantal combinaties uitgezet voor respectievelijk 12 maal kruis, 11 maal, tienmaal, enzovoorts. Een eigenschap van deze grafiek is dat het oppervlak onder de kromme gelijk is aan 100%, dat wil zeggen gelijk is aan de gesommmeerde percentages van de derde kolom in tabel II. Normaliter worden geen percentages maar proporties (= fracties) gebruikt; dan is het oppervlak gelijk aan de som van alle proporties (vierde kolom) en dus gelijk aan 1.

De balken in afbeelding 1 bevatten dus alle in theorie mogelijke uitkomsten, de *kansverdeling* van de populatie. Voor een dichotome variabele waarbij de twee klassen in de populatie even groot (50:50) en onafhankelijk van elkaar zijn, is de binomiale kansverdeling symmetrisch, met de top in het midden, waar de meeste waarden gesitueerd zijn; naar de zijkanten neemt het aantal 'waarnemingen' progressief af. De gestippelde kromme (polygoon) wekt de indruk dat we met een normale verdeling, de klokvormige Gausse-curve, van doen hebben en dat is ook bij benadering waar. Er bestaan ook andere binomiale verdelingen. Men denke slechts aan die waarbij de verhouding tussen de twee klassen niet 50:50 is.

7 Besluitvorming

De tandarts besloot H_0 te verwerpen als de kans om dat ten onrechte te doen klein is: hij stelde het *significantieniveau* α op 0,01. Verwerping van H_0 houdt acceptatie van

Tabel I. Alle in theorie mogelijke combinaties (8) van kruis en munt bij driemaal werpen met een geldstuk.

Mogelijke combinaties	Eerste worp	Tweede worp	Derde worp
1.	kruis	kruis	kruis
2.	kruis	kruis	munt
3.	kruis	munt	kruis
4.	kruis	munt	munt
5.	munt	kruis	kruis
6.	munt	kruis	munt
7.	munt	munt	kruis
8.	munt	munt	munt

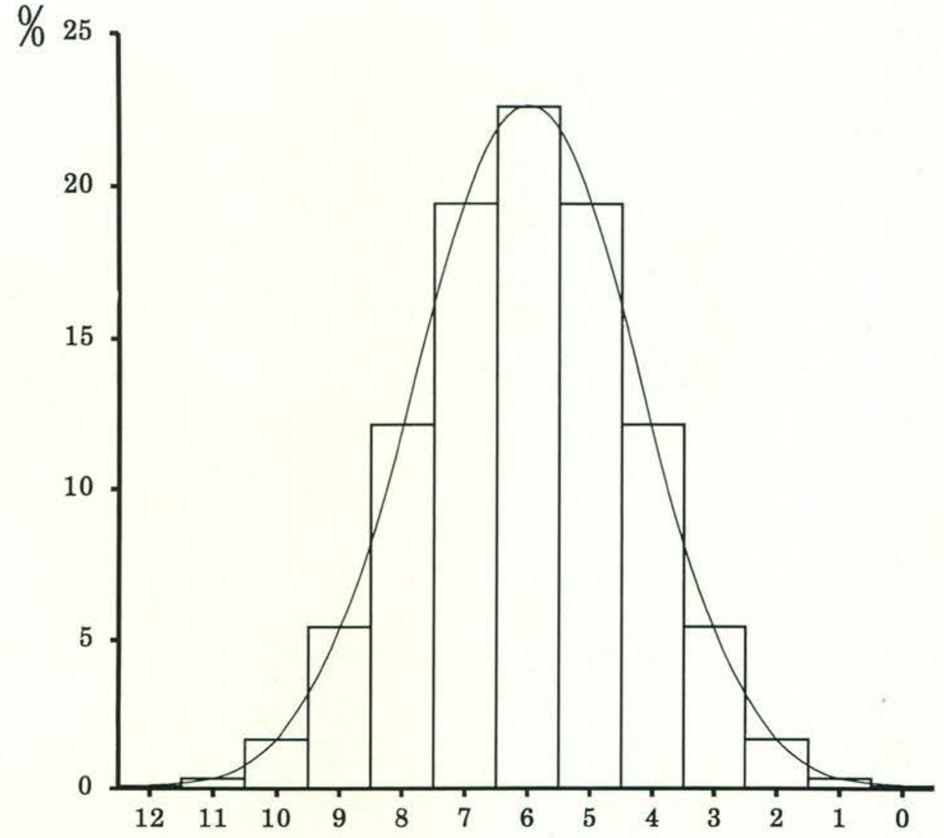
Tabel II. De gegevens van tabel I samengevat. Het aantal combinaties waarin respectievelijk 3, 2, 1 en 0 maal kruis boven ligt.

Aantal combinaties	Aantal malen kruis boven	Kansen van deze combinaties (%)
1	3	12,5
3	2	37,5
3	1	37,5
1	0	12,5
—— +		—— +
8		100%

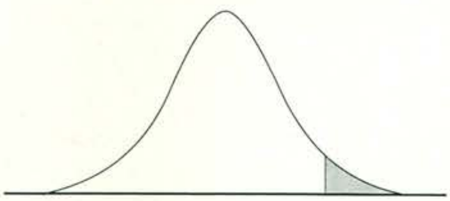
Tabel IIa. Het aantal combinaties voor P = 0,5* bij respectievelijk 1, 2, 3, 4 en 5 worpen. Tussen haakjes staat onder elk aantal combinatie vermeld hoe veel kruisen bovenliggen en cursief zijn de kansen van die uitkomsten vermeld (in %).

Aantal worpen	Aantal combinaties						Aantal mogelijkheden
1.	1	1					2
	(1)	(0)					
	50	*50*					
	—	—					
2.	1	2	1				4
	(2)	(1)	(0)				
	25	*50*	*25*				
3.	1	3	3	1			8
	(3)	(2)	(1)	(0)			
	12,5	*37,5*	*37,5*	*12,5*			
4.	1	4	6	4	1		16
	(4)	(3)	(2)	(1)	(0)		
	6,25	*25*	*37,5*	*25*	*6,25*		
5.	1	5	10	10	5	1	32
	(5)	(4)	(3)	(2)	(1)	(0)	
	3,125	*15,625*	*31,25*	*31,25*	*15,625*	*3,125*	

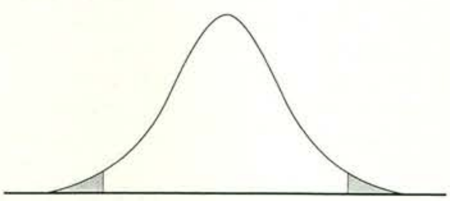
* De driehoek van Pascal is ook geldig voor andere kansen dan P = 0,5.
Noot: de aantallen in de combinaties heten binomiaal coëfficiënten en hebben als zodanig niet met statistiek of meetniveau te maken.



Afb. 1. Kansverdeling van alle mogelijke combinaties 'kruis-munt' bij 12 maal werpen. Op de X-as is te zien hoe groot de kans is op respectievelijk (van links naar rechts) 12 maal, 11 maal, 10 maal . . . éénmaal, nul maal munt gooien.



Afb. 2. Eenzijdige toetsing. Het verwerpingsgebied (gearceerd) van de nulhypothese omvat 5% van het oppervlak onder de kromme.



Afb. 3. Tweezijdige toetsing. Het verwerpingsgebied (5%) is opgedeeld in een linker- (2,5%) en rechterzijde (2,5%).

H_1 in: bij deze waarde van α draagt de verwerping van H_0 een grote mate van zekerheid in zich. De verwerping van H_0 houdt NIET in dat het geldstuk VALS is. Het betekent dat onder de aanname dat H_0 juist is, deze bevinding (de munt is vals) minder dan 1% kans van voorkomen heeft.

8 Is de munt vals?

De kans op 12 maal kruis bij 12 worpen is, aannemende dat het geldstuk eerlijk is, 0,00024 (tab. III). Dat is dus minder dan α, die vooraf op 0,01 was gesteld. Dus moet H_0 worden verworpen. Ook als de tandarts in 12 worpen 11 of 12 maal kruis gooit (waarschijnlijkheid respectievelijk 0,00024 en 0,0029), moet H_0 worden verworpen, omdat beide waarschijnlijkheden bij elkaar opgeteld kleiner zijn dan α (namelijk $0{,}00024 + 0{,}0029 < 0{,}01$).

Komt echter, zoals in ons experiment, tienmaal kruis boven bij de 12 worpen, dan moet de tandarts H_0 aanvaarden, daar de waarschijnlijkheid van tienmaal kruis gooien (of minder vaak) met een eerlijke munt $0{,}00024 + 0{,}0029 + 0{,}0161 = 0{,}0192$, een waarde $> \alpha = 0{,}01$. De waarschijnlijkheid waarmee H_0 ten onrechte wordt verworpen is klein.

De tandarts had natuurlijk een mildere norm kunnen kiezen en vaker (of minder vaak) kunnen gooien. Dikwijls wordt α op 0,05 gesteld. Dat zou consequenties voor de conclusie hebben gehad, want de kans op 12 of 11 of tien kruisen is minder groot dan 0,05 (tab. III) = H_0 wordt verworpen.

9 Verwerpingsgebied

In afbeelding 2 is een deel van het oppervlak onder de kromme gearceerd. Dat deel komt overeen met de waarschijnlijkheid dat 12 maal kruis en 12 of 11 maal kruis was

Tabel III. Aantal combinaties waarin kruis 12 keer, 11 keer, 10 keer en 0 keer boven ligt bij 12 maal gooien met een eerlijk geldstuk.

Aantal combinaties	Aantal malen kruis boven	Kans %	Waarschijnlijkheid van aantal malen kruis
1	12	0,02	0,00024
12	11	0,29	0,0029
66	10	1,61	0,0161
220	9	5,37	0,0537
495	8	12,08	0,1208
792	7	19,36	0,1936
924	6	22,56	0,2256
792	5	19,36	0,1936
495	4	12,08	0,1208
220	3	5,37	0,0537
66	2	1,61	0,0161
12	1	0,29	0,0029
1	0	0,02	0,00024
—— +		—— +	—— +
4096		100,00	1,0000

gegooid, het zogenaamde verwerpingsgebied.

H_0 hield in dat het geldstuk eerlijk was en H_1 dat het geldstuk vals was en de kop te vaak bovenkwam. Als we echter geen enkele indicatie hebben voor de eerlijkheid of valsheid van een geldstuk, zijn in wezen drie hypothesen mogelijk:
1. het geldstuk is vals, waarbij de mogelijkheid wordt opengelaten dat of kruis of munt te vaak bovenkomt, 2. kruis komt vaker boven dan munt, en 3. munt komt vaker boven dan kruis.

In het tweede (en derde) geval wordt eenzijdig getoetst, omdat H_1 een 'richting aangeeft'. In het eerste geval geschiedt toetsing tweezijdig, omdat er geen richting in de formulering is opgenomen, hetgeen in afbeelding 3 tot uiting komt. We zien daar een linker- en een rechtergebied gearceerd (elk de helft van het verwerpingsgebied bevattend).

10 Slot

De binomiale toets, ook elders beschreven,[2 3] komt men niet vaak tegen in de literatuur. Toch wordt bespreking ervan nuttig geacht, omdat het een goede illustratie is van de gedachtengang in de statistiek met betrekking tot toetsen. De toets kan echter goed worden toegepast in situaties als: 1. het toetsen van enkelvoudige overerving volgens de wetten van Mendel, 2. het toetsen van 'failures' in een clinical trial met twee restauratietechnieken, en 3. optreden van bijwerkingen van geneesmiddelen. Generalisatie is ook mogelijk naar meer dan twee categorieën (kruis/munt), bijvoorbeeld 'carieus', 'gevuld' en 'gaaf'.

Problemen zoals hier beschreven, kunnen veelal worden opgelost met de 'tekentoets'. Maar, als N (= aantal worpen, enzovoorts) groot is, wordt de tekentoets op zijn beurt meestal uitgevoerd met behulp van de chi-kwadraat (X^2) benadering, want anders is het rekenwerk overweldigend. De X^2-toets wordt in een vervolgartikel aan de orde gesteld.

Summary

STATISTICS (IV)

Tossing for decision making; one sample, dichotomy

Key words: Statistics

A coin, suspected to be biased, may be distinguished from a fair one by tossing it a number of times and applying the binomial test. Taking into account the probabilities of all possible outcomes (the theoretical sampling distribution), a too large occurrence of heads (or tails) will lead to rejection of the null hypothesis implying that the coin is not fair.

Literatuur

[1]SIEGEL S, CASTELLAN NJ. Non parametric statistics. New York: McGraw-Hill book company, 1988: 15-8 en 38-44.
[2]GUILFORD JP, FRUCHTER B. Fundamental statistics in psychology and education. Tokyo: McGraw-Hill Kogakusha, 1978: 102-5.
[3]VAN DEN BRINK WP, KOELE P. Statistiek. Deel 2. Meppel: Boom, 1985: 65-70.

PAO-toets

Elders in deze aflevering is een bijdrage opgenomen over diabetes mellitus. Wanneer u deze nog niet hebt gelezen, kunt u uw huidige kennis over dit onderwerp toetsen aan de hand van onderstaande vragen. Hebt u de bijdrage wel al bestudeerd, dan zal het beantwoorden van de vragen u niet moeilijk vallen en, indien dat toch blijkt tegen te vallen, u wellicht tot herlezen aansporen. De goede antwoorden staan elders in deze aflevering vermeld.

1. Hoe vaak komt diabetes mellitus onder de Nederlandse bevolking voor?
   a. 2%
   b. 5%
   c. 10%
   d. 15%

2. Er is een type I en een type II diabetes mellitus. Welke bewering uit onderstaand rijtje is onjuist?
   a. Bij type I diabetes mellitus bestaat een absoluut tekort aan insuline
   b. Type I diabetes mellitus wordt niet veroorzaakt door overgewicht
   c. Type I diabetes mellitus komt meestal voor bij oudere patiënten
   d. Bij type II diabetes zijn de organen waar de insuline moet werken vaker minder gevoelig voor insuline

3. Bij een niet-behandelde diabetes mellitus ontstaan op korte termijn door verlies van glucose en water via de urine dorst, uitdroging en gewichtsverlies. Bij langbestaande diabetes mellitus is een bekende complicatie:
   a. levercirrose
   b. retinopathie die tot blindheid leidt
   c. verhoogde kans op botbreuken
   d. ontsteking van de pancreas

4. Bij type I diabetes mellitus bestaat de behandeling uit toediening van insuline. Dit kan niet per os worden gegeven omdat:
   a. de smaak zeer onaangenaam is
   b. insuline zeer slecht wordt geresorbeerd
   c. schadelijk is voor het mondslijmvlies
   d. in de maag wordt afgebroken en daarom niet in de darm kan worden geresorbeerd

5. Wanneer bij de behandeling van type II diabetes de glucosespiegels ondanks een optimaal dieet te hoog blijven, worden bloedsuikerverlagende middelen voorgeschreven. Deze zijn bedoeld om:
   a. de pancreas aan te zetten tot een hogere produktie van insuline
   b. de uitscheiding van glucose door de nieren te bevorderen
   c. de eetlust te remmen
   d. de leverfunctie te stimuleren

6. Een patiënt wordt onwel in de tandartsstoel. Bekend is dat hij insuline gebruikt. Vermoedelijk is dan ook sprake van hypoglykemie. Wanneer de patiënt nog in staat is om te drinken, geef dan:
   a. glas kraanwater
   b. koffie of thee met enkele scheppen suiker
   c. glaasje sterke drank
   d. geen drank in welke vorm dan ook

7. Wanneer moet begonnen worden met een scherpe instelling van het glucosegehalte bij een vrouw met kinderwens?
   a. voor de conceptie
   b. zodra bekend is dat de vrouw zwanger is
   c. in het tweede trimester
   d. voor de partus