STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S 308 - A 1

De begrippen frequentiequotiënt en kans

Wij stellen ons een experiment E voor, dat tot verschillende uitkomsten S, T, U, enz. kan leiden. Hierbij zullen we het begrip "experiment" zeer ruim opvatten. Het kan bijvoorbeeld zijn een worp met een dobbelsteen of met een munt, een waarneming betrekking hebbend op het al of niet in leven zijn van een bepaalde persoon op een bepaald tijdstip, het controleren van een kasstuk, enz.

Wanneer wij het experiment E enige keren, bijv. n maal, herhalen en het resultaat S treedt hierbij k maal op, dan noemen wij k de <u>frequentie</u> van het resultaat S in de reeks en $\frac{k}{n}$ het <u>frequentiequotiënt</u> van S (afkorting fq, meervoud fqn; notatie fq(S)).

Het frequentiequotiënt bezit o.a. de volgende eigenschappen:

$$0 \leqq fq(S) \leqq 1;$$

$$fq(S) = 0 \text{ indien S nooit optreedt;}$$

$$fq(S) = 1 \text{ indien S altijd optreedt.}$$

Zowel voor verschillende waarden van n als bij een nieuwe reeks van hetzelfde aantal experimenten, zal in het algemeen voor fq(S) niet dezelfde waarde gevonden worden. Voeren wij echter de experimenten E zodanig uit dat de resultaten van de verschillende uitvoeringen elkaar niet beïnvloeden en bepalen wij steeds voor de reeks van reeds plaats gevonden experimenten de grootte van fq(S), dan leert de praktijk, dat bij veel experimenten de fqn voor S minder van elkaar gaan verschillen naarmate n groter wordt. Men noemt dit verschijnsel de "experimentele wet van de grote aantallen".

Een voorbeeld kan een en ander nog verduidelijken. In onderstaande tabel zijn de aantallen zessen vermeld, verkregen in 6 reeksen worpen met een dobbelsteen. De fqn in de vierde kolom tonen nogal wat verschillen. De waarden van n worden nu vergroot door de reeksen achtereenvolgens bij elkaar te nemen. Wij merken op dat de zo verkregen fqn kleinere verschillen vertonen bij toenemende n.

Reeks no	n	k	fq(S)	Reeks no	n	k	fq(S)
1	100	15	0,15	1	100	15	0,150
2	100	16	0,16	1 en 2	200	31	0,155
3	100	13	0,13	1 t/m 3	300	44	0,147
4	100	18	0,18	1 t/m 4	400	62	0,155
5	100	17	0,17	1 t/m 5	500	79	0,158
6	100	20	0,20	1 t/m 6	600	99	0,165
7	100	18	0,18	1 t/m 7	700	117	0,167

Naast het begrip frequentiequotiënt bestaat het begrip kans, welk begrip in de kansrekening of waarschijnlijkheidsrekening van zeer veel belang is. Het begrip kans is een zuiver wiskundig begrip, dat dan ook op de in de wiskunde gebruikelijke wijze wordt ingevoerd, namelijk met behulp van axioma's. In de wiskundige theorie moet het kansbegrip een rol spelen, welke vergelijkbaar is met die van het frequentiequotiënt in de praktijk. Wij voegen daarom aan iedere gebeurtenis S per definitie een getal toe, aangegeven door $P[S]$, dat voldoet aan o.a. de axioma's:

$$0 \leq P[s] \leq 1 \text{ voor iedere } S;$$
$$P[S] = 0 \text{ indien } S \text{ onmogelijk is;}$$
$$P[S] = 1 \text{ indien } S \text{ zeker is.}$$

Deze axioma's komen dus overeen met de bovengenoemde eigenschappen van frequentiequotiënten.

Uit deze en nog twee andere hier niet te noemen axioma's kunnen onder andere de volgende eigenschappen worden afgeleid:

$$P[S \text{ of } T \text{ of } U \text{ of } V] = P[S] + P[T] + P[U] + P[V],$$

indien de gebeurtenissen S, T, U en V niet tegelijk op kunnen treden;

$$P[S \text{ én } T \text{ én } U \text{ én } V] = P[S] \cdot P[T] \cdot P[U] \cdot P[V],$$

indien het optreden van één van de gebeurtenissen S, T, U of V niet afhankelijk is van het optreden van een van de andere drie gebeurtenissen.

Tenslotte noemen we nog een eigenschap, welke uit de axioma's kan worden afgeleid, de zogenaamde <u>theoretische wet van de grote aantallen</u>. Deze beweert dat bij een reeks van n onafhankelijke experimenten met ieder een kans p op succes, de fractie successen $\frac{k}{n}$ voor steeds grotere n steeds dichter bij p komt te liggen en dat de kans dat dit niet gebeurt voor steeds grotere n steeds kleiner wordt. Deze theoretische wet van de grote aantallen voor kansen correspondeert dus met de experimentele wet van de grote aantallen voor frequentiequotiënten.

<u>Voorbeelden</u>

I. Een experiment bestaat uit het éénmaal werpen met een "zuivere" dobbelsteen. Het is onmogelijk tegelijk een vijf en een zes te werpen. De kans dat een vijf of een zes wordt gegooid is dus:

$$P[5 \text{ of } 6] = P[5] + P[6] = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

II. Een experiment bestaat uit het driemaal werpen met een "zuivere" dobbelsteen. De uitkomst van een worp wordt niet beïnvloed door de uitkomst van een vorige worp. De kans dat achtereenvolgens een drie, een vier en een vijf geworpen wordt is dus:

$$P[3 \text{ én } 4 \text{ én } 5] = P[3] \cdot P[4] \cdot P[5] = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216}.$$

W

A

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S308-A2

Stochastische grootheden en kansverdelingen

Een stochastische grootheid (στοχαζεσθαι = raden; een grootheid, naar de waarde waarvan men slechts gissen kan) is een grootheid die -misschien binnen bepaalde grenzen- allerlei waarden aan kan nemen, terwijl men vooruit niet kan voorspellen welke waarde bij een bepaald experiment zal worden aangenomen.

Voorbeelden zijn:
het aantal zessen bij acht worpen met een dobbelsteen; het aantal malen kruis bij zes worpen met een munt; de leeftijd van een willekeurige voorbijganger op straat; het aantal fouten, dat men zal vinden bij een accuratessecontrole.

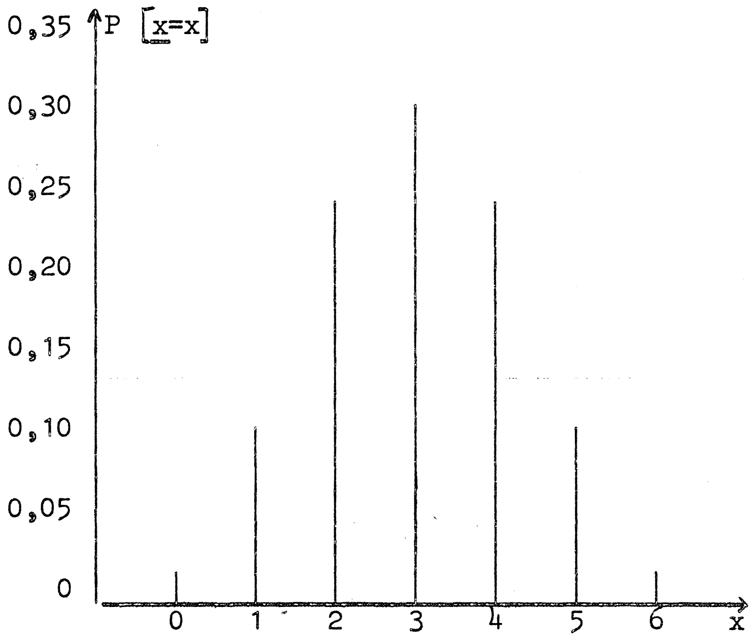
Een stochastische grootheid wordt aangegeven door een onderstreept symbool, bijv. $\underline{x}$.

Stel dat $\underline{x}$ het aantal malen kruis, verkregen bij zes worpen met een zuivere munt vòorstelt, dan kan $\underline{x}$ één van de waarden 0 tot en met 6 aannemen. Indien x de waarde is, welke $\underline{x}$ bij een dergelijk experiment aanneemt, dan behoort bij x een bepaalde kans, die aangegeven wordt met $P[\underline{x}=x]$ [1]). In Figuur 1 is voor iedere x de waarde van $P[\underline{x}=x]$ getekend.

Figuur 1 noemt men de grafische voorstelling van de kansverdeling van $\underline{x}$. Omdat $\underline{x}$ zéker één van de waarden 0 tot en met 6 zal aannemen, is de som van al deze kansen gelijk aan één.

---

1) Voor het begrip kans zie Memorandum S308-A1



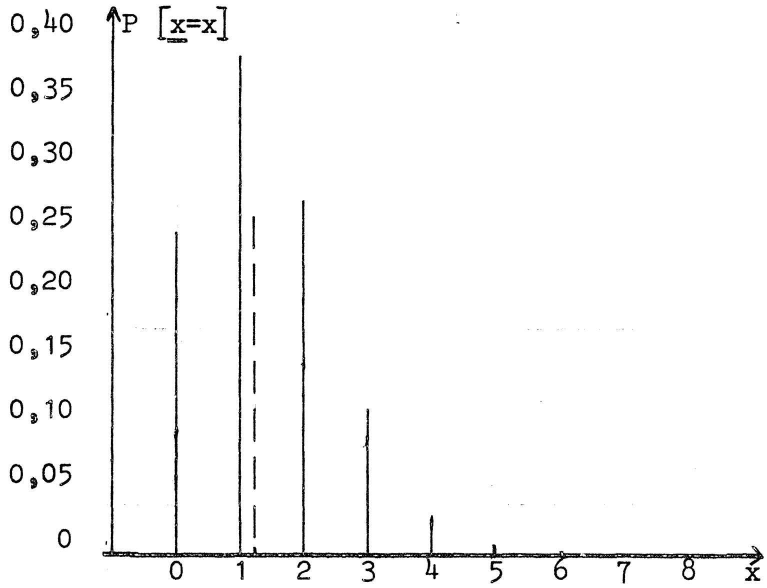
Figuur 1

De kans op x kruizen bij zes worpen met een zuivere munt.

Een andere kansverdeling is die van het aantal zessen, verkregen bij acht worpen met een zuivere dobbelsteen. Ook dit aantal zullen wij met x aanduiden en de waarde, welke x in een experiment heeft, met x. De waarden die x hier kan doorlopen zijn 0 tot en met 8. In Figuur 2 zijn de kansen P [x=x] getekend. De kansen, behorende bij x=6, 7 en 8, zijn zo klein dat ze in de figuur niet van 0 te onderscheiden zijn; zie Voorbeeld II.

De kansverdeling van het aantal te vinden fouten bij een accuratesse-controle is van hetzelfde type als die in de figuren 1 en 2. De grootheid x kan hier alle gehele waarden aannemen tussen 0 en het aantal gecontroleerde handelingen.

In alle bovengenoemde gevallen neemt x slechts gehele waarden aan. Stochastische grootheden, die alleen gehele waarden aan kunnen nemen, of alleen bijv. veelvouden van 1/2 of 1/10, e.d., noemt men discrete stochastische grootheden en de bijbehorende kansverdelingen discrete kansverdelingen.



Figuur 2.

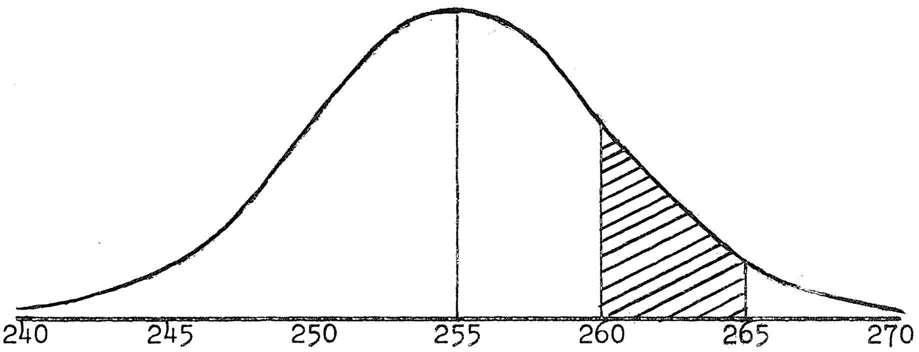
De kans op x zessen bij acht worpen met een zuivere dobbelsteen.

Naast situaties waarin x slechts discrete waarden aan kan nemen, bestaan er veel gevallen waarin x in principe alle waarden aan kan nemen, of alle waarden tussen van tevoren opgegeven grenzen.

Voorbeelden hiervan zijn:

de lengte van een recruut;
het gewicht van een pak koffie;
de temperatuur 's middags om 12 uur;
het gemiddelde bedrag van een aantal inkoopfacturen;
het gemaakte brutowinst-percentage gedurende een bepaald jaar bij een winkelier.

De genoemde grootheden worden in de praktijk met een zekere nauwkeurigheid gemeten, lengten bijvoorbeeld in cm. of mm., temperaturen bijvoorbeeld in $^{\circ}$C of in tienden van $^{\circ}$C, enz. Dit kan de indruk wekken dat ook deze grootheden discreet zijn. Toch blijven deze grootheden continue grootheden, in tegenstelling tot de eerder beschouwde discrete grootheden, daar met meer verfijnde instrumenten nauwkeuriger uitkomsten, bijv. in tienden van mm., in honderste van $^{\circ}$C, enz., bereikt kunnen worden. Men noemt ze continue stochastische grootheden en de bijbehorende verdelingen continue kansverdelingen.



Figuur 3

Kansverdeling van het gewicht van een pak koffie van nominaal 250 gram.

Een mogelijke kansverdeling van het gewicht van een pak koffie is in Figuur 3 geschetst. Het is intuïtief wel in te zien dat bij een continue kansverdeling de kans dat $\underline{x}$ een bepaalde waarde x aanneemt, gelijk is aan 0. Men spreekt bij deze verdelingen dan ook alleen van de kans, dat $\underline{x}$ tussen twee waarden x_1 en x_2 in ligt. Deze kans, $P[x_1 \leq \underline{x} \leq x_2]$, is gelijk aan de oppervlakte onder de kromme tussen x_1 en x_2, gesteld dat de totale oppervlakte onder de kromme gelijk is aan 1. Zo is in ons voorbeeld de kans, dat het gewicht van een willekeurig pak koffie ligt tussen 260 en 265 gram gelijk aan de oppervlakte van het gearceerde gedeelte onder de kromme van Figuur 3 tussen 260 en 265.

De kansverdelingen van de Figuren 1 en 3 zijn symmetrische kansverdelingen. De kansverdeling van Figuur 2 is een scheve discrete kansverdeling. Uiteraard bestaan er ook scheve continue kansverdelingen.

Een belangrijke grootheid van een kansverdeling is zijn (mathematische) verwachting. Bij zes worpen met een dobbelsteen verwacht men gemiddeld één zes, bij acht worpen verwacht men gemiddeld $\frac{8}{6}$ zessen. Hier is de verwachting dus $1\frac{1}{3}$, welke waarde met een stippellijn in Figuur 2 is aangegeven. Het verwachte aantal malen kruis bij zes worpen met een zuivere munt bedraagt $\frac{6}{2} = 3$. In Figuur 3 is het verwachte gewicht gelijk aan 255 gram. In feite is de verwachting een "gewogen gemiddelde" van alle mogelijke waarden van $\underline{x}$, waarbij iedere waarde wordt gewogen met de bijbehorende kans. Bij een discrete verdeling kan deze door sommering berekend worden; bij een continue verdeling met een aan het gewone sommeren verwant procédé, genaamd integratie (zie de Appendix).

De verwachting van een kansverdeling wordt meestal aangegeven met het symbool μ.

Een tweede belangrijke grootheid van een kansverdeling is de <u>variantie</u>, aangeduid met het symbool σ^2. Om deze grootheid te vinden bepaalt men de absolute afwijking van $\underline{x}$ ten opzichte van zijn verwachting μ, dus $|\underline{x} - \mu|$ 2). De grootheid $|\underline{x} - \mu|$ is weer een stochastische grootheid met een zekere kansverdeling en bezit dus ook weer een verwachte waarde. Het rekenen met absolute waarden is echter zeer lastig en daarom geeft men er de voorkeur aan te werken met de verwachting van $(\underline{x} - \mu)^2$. Dit noemt men de variantie van $\underline{x}$. Ook deze grootheid σ^2 is dus niets anders dan het "gewogen gemiddelde" van alle waarden, die $(\underline{x} - \mu)^2$ aan kan nemen.

Een derde belangrijke grootheid van een kansverdeling is de <u>standaardafwijking</u>, ook wel de "spreiding" genoemd, welke met het symbool σ wordt aangeduid. De standaardafwijking σ is inderdaad niets anders dan de wortel uit de variantie σ^2.

---

2) $|a|$ is de "absolute waarde van a". $|a| = a$ indien $a \geq 0$; $|a| = -a$ indien $a \leq 0$. Voorbeelden: $|5| = 5$; $|0| = 0$; $|-3\frac{1}{2}| = 3\frac{1}{2}$.

Voorbeelden

I.

x	$P[\underline{x}=x]$	$x \cdot P[\underline{x}=x]$	$(x-3)^2$	$(x-3)^2 \cdot P[\underline{x}=x]$
0	0,015625	0,000000	9	0,140625
1	0,093750	0,093750	4	0,375000
2	0,234375	0,468750	1	0,234375
3	0,312500	0,937500	0	0,000000
4	0,234375	0,937500	1	0,234375
5	0,093750	0,468750	4	0,375000
6	0,015625	0,093750	9	0,140625
Som	1,000000	3,000000		1,500000

Bovenstaande tabel bevat de exacte waarden van de kansen $P[\underline{x}=x]$ uit Figuur 1 (zie de Appendix). In de praktijk gebruikt men meestal een kleiner aantal decimalen. De verwachting μ is gelijk aan de som van de derde kolom en is dus gelijk aan 3. De variantie σ^2 is gelijk aan de som van de vijfde kolom en is dus gelijk aan 1,5. De standaardafwijking σ is derhalve gelijk aan $\sqrt{1{,}5} = 1{,}2247$.

II

x	$P[\underline{x}=x]$	$x \cdot P[\underline{x}=x]$	$(x-\frac{4}{3})^2$	$(x-\frac{4}{3})^2 \cdot P[\underline{x}=x]$
0	0,2325680	0,0000	1,7778	0,4135
1	0,3721089	0,3721	0,1111	0,0413
2	0,2604762	0,5210	0,4444	0,1158
3	0,1041905	0,3126	2,7778	0,2894
4	0,0260476	0,1042	7,1111	0,1852
5	0,0041676	0,0208	13,4444	0,0560
6	0,0004168	0,0025	21,7778	0,0091
7	0,0000238	0,0002	32,1111	0,0008
8	0,0000006	0,0000	44,4444	0,0000
Som	1,0000000	1,3334		1,1111

De waarden van de kansen $P[\underline{x}=x]$ in Figuur 2 zijn in bovenstaande tabel weergegeven (zie de Appendix). Om ook de kleine kansen voor $x = 7$ en $x = 8$ goed weer te kunnen geven zijn overal voor deze kansen zeven decimalen gebruikt. De berekeningen zijn in vier decimalen uitgevoerd. De verwachting μ is gelijk aan de som van de derde kolom en is dus gelijk aan 1,3334; de exacte waarde is gelijk aan $\frac{4}{3}$. De variantie σ^2 is gelijk aan de som van de vijfde kolom en is dus gelijk aan 1,1111; de exacte waarde is gelijk aan $\frac{10}{9}$. De standaardafwijking σ is derhalve gelijk aan $\sqrt{1,1111} = 1,0541$.

Appendix

De kansverdelingen in de figuren 1 en 2 zijn z.g.n. "binomiale verdelingen", gegeven door

$$P\left[\underline{x} = x\right] = \binom{n}{x} p^{x} (1-p)^{n-x}, \text{ waarin}$$

n = aantal experimenten (bijv. worpen met dobbelsteen of munt),

p = kans dat bij zo'n experiment een "succes" optreedt (bijv. de kans bij een worp op een zes of op kruis),

x = aantal successen bij de n experimenten, en

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!} \quad \text{met } x! = x(x-1)(x-2)\ \ldots\ 2.1 \text{ en } 0! = 1.$$

De kansverdeling van figuur 3 wordt gegeven door

$$p\left[x_1 \leq \underline{x} \leq x_2\right] = \int_{x_1}^{x_2} \frac{1}{5\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-255}{5}\right)^2\right] dx.$$

Dit is de z.g.n. "normale verdeling" of "verdeling van Gauss", in het algemeen gegeven door

$$P\left[x_1 \leq \underline{x} \leq x_2\right] = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx,$$

waarin μ de verwachting en σ de standaardafwijking is (zie ook Memorandum S308-A3).

<u>De verwachting</u> wordt gedefiniëerd door

$$\mu = \sum_{x} x \,.\, P\,[\underline{x} = x]$$

bij discrete kansverdelingen,

en door

$$\mu = \int_{-\infty}^{\infty} x \,.\, f(x)\,dx$$

bij een continue kansverdeling waarvoor

$$P\,[x_1 \leq \underline{x} \leq x_2] = \int_{x_1}^{x_2} f(x)dx.$$

<u>De variantie</u> wordt gedefiniëerd door

$$\sigma^2 = \sum_{x} (x-\mu)^2 \,.\, P\,[\underline{x} = x]$$

bij discrete kansverdelingen,

en door

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \,.\, f(x)dx$$

bij continue kansverdelingen.

W

A

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

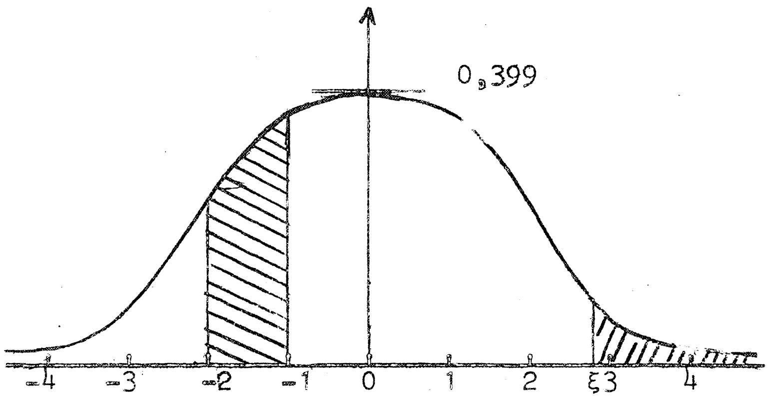
AFDELING MATHEMATISCHE STATISTIEK

S308-A3

De Normale Verdeling

In Memorandum S308-A2 werden de begrippen stochastische grootheid en kansverdeling besproken. Als voorbeeld van een continue kansverdeling is daar reeds de "normale verdeling" gebruikt, zonder dat dit expliciet werd vermeld (zie figuur 3 van dat Memorandum).

De normale verdeling is een continue kansverdeling, die geheel wordt bepaald door de verwachting μ en de standaardafwijking σ. Men duidt deze verdeling, die ook wel de verdeling van Gauss wordt genoemd, daarom kortheidshalve vaak aan met N (μ ; σ). In figuur 1 is een normale verdeling met μ = 0 en σ = 1 getekend. De kromme



Figuur 1

De normale verdeling met verwachting 0 en standaardafwijking 1.

eindigt niet bij +4 en -4, maar loopt naar beide zijden oneindig ver door, waarbij hij steeds dichter tot de horizontale as nadert.

Op deze horizontale as zijn de waarden aangegeven, die de stochastische grootheid $\underline{x}$ aan kan nemen en dat zijn hier dus alle waarden tussen $-\infty$ en $+\infty$. De oppervlakte onder de kromme is, zoals bij alle continue kansverdelingen, gelijk aan 1.

Zoals in Memorandum S308-A2 reeds werd opgemerkt, is de kans, dat $\underline{x}$ een bepaalde waarde ξ aanneemt, gelijk aan 0. Men kan hier slechts spreken over de kans dat $\underline{x}$ tussen twee waarden inligt, stel tussen x_1 en x_2. Deze kans is gelijk aan de oppervlakte tussen x_1 en x_2. In figuur 1 is de kans, dat $\underline{x}$ ligt tussen -2 en -1, dus $P\left[-2 \leq \underline{x} \leq -1\right]$, aangegeven door de gearceerde oppervlakte tussen -2 en -1.

De getekende verdeling is symmetrisch t.o.v. de verticale as door $x = 0$, dus ook t.o.v. de verwachting $\mu = 0$ van de verdeling. Dit geldt niet alleen voor deze normale verdeling, maar ook een willekeurige $N(\mu ; \sigma)$ - verdeling is symmetrisch t.o.v. de verwachting σ, dus t.o.v. de verticale as door $x = \mu$. Bij een $N(0;1)$- verdeling is de kans, dat $\underline{x}$ ligt tussen 1 en 2, dus $P\left[1 \leq \underline{x} \leq 2\right]$, daardoor gelijk aan $P\left[-2 \leq \underline{x} \leq -1\right]$.

Bij een groot aantal praktische problemen is men geïnteresseerd in vragen zoals die naar de kans, dat een stochastische grootheid een waarde groter dan of gelijk aan een bepaald getal ξ aanneemt, of een waarde meer dan zoveel van de verwachting aanneemt, enzovoort. Dergelijke kansen noemt men <u>overschrijdingskansen</u>. Hieronder volgen de nauwkeurige definities van deze kansen.

Onder de <u>rechteroverschrijdingskans</u> van $x = \xi$ verstaan wij de kans, dat $\underline{x}$ groter dan of gelijk is aan ξ, dus $P\left[\underline{x} \geq \xi\right]$. Deze kans is gelijk aan de oppervlakte onder de kromme, rechts van het punt ξ op de horizontale as (zie figuur 1).

Onder de <u>linkeroverschrijdingskans</u> van $x = \xi$ verstaan wij de kans, dat $\underline{x}$ een waarde kleiner dan of gelijk aan ξ aanneemt, dus $P\left[\underline{x} \leq \xi\right]$. Deze kans is gelijk aan de oppervlakte onder de kromme, links van het punt ξ op de horizontale as.

Tenslotte is de <u>tweezijdige overschrijdingskans</u> van $x = \xi$ gelijk aan tweemaal de kleinste éénzijdige overschrijdingskans van $x = \xi$.

De overschrijdingskansen bij een N(0;1) - verdeling zijn uitvoerig getabelleerd. Een tabel van rechteroverschrijdingskansen is aan dit memorandum toegevoegd.

Om de rechteroverschrijdingskans van x = ξ bij een N(μ;σ) - verdeling te vinden, berekenen we eerst de grootheid

$$u = \frac{\xi - \mu}{\sigma} \tag{1}$$

en zoeken daarna de rechteroverschrijdingskans van u op in de tabel van de N(0;1) - verdeling.

De normale verdeling speelt een zeer belangrijke rol in de statistiek:

1. In de eerste plaats geldt voor tal van in de praktijk voorkomende grootheden, dat de normale verdeling een goede benadering vormt voor de daarbij behorende kansverdeling. Dit geldt bijvoorbeeld vaak voor meetfouten, weegfouten, lengten van mensen, diameters van eieren, e.d. Daarom kon ook in het voorbeeld van het gewicht van pakjes koffie in Memorandum S308-A2 een normale verdeling gebruikt worden.

2. Verder is de som van een aantal normaal verdeelde grootheden ook weer normaal verdeeld. Dit geldt zowel voor onderling afhankelijke als voor onderling onafhankelijke stochastische grootheden. Stochastische grootheden zijn onderling onafhankelijk indien de waarde, die één van hen aanneemt, niet beïnvloed wordt door de waarden, die de andere grootheden aannemen.

Zijn n onderling onafhankelijke stochastische grootheden, $\underline{x}_1$, $\underline{x}_2$, ..., $\underline{x}_n$, alle normaal verdeeld, met gemiddelden μ_1, μ_2, ..., μ_n en standaardafwijkingen σ_1, σ_2, ..., σ_n, dan is de grootheid

$$\underline{y} = \underline{x}_1 + \underline{x}_2 + \ldots + \underline{x}_n \tag{2}$$

normaal verdeeld met verwachting

$$\mu = \mu_1 + \mu_2 + \ldots + \mu_n \tag{3}$$

en standaardafwijking

$$\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2} \qquad (4)$$

Dit resultaat geldt voor elke waarde van n.

3. Een dergelijk resultaat geldt bij benadering voor de kansverdeling van de som van een aantal onderling onafhankelijke stochastische grootheden met willekeurige verdelingen, nu echter niet voor alle waarden van n, doch alleen voor grote waarden van n. Men kan namelijk onder zeer algemeen geldende voorwaarden bewijzen, dat voor grote n de verdeling van de som van n onafhankelijk verdeelde stochastische grootheden bij benadering normaal verdeeld is met een verwachting , gegeven door (3) en een standaardafwijking, gegeven door (4).

Een toepassing van punt 3 kan gevonden worden in Voorbeeld I en de daarbij behorende figuur 1 van Memorandum S308-A2. De daar weergegeven kansverdeling van het aantal malen kruis in zes worpen met een zuivere munt is symmetrisch en lijkt reeds veel op een normale verdeling. Bij toeneming van het aantal worpen wordt deze gelijkenis steeds sterker. De mogelijke uitkomsten van één worp kunnen hier gezien worden als één stochastische grootheid, welke de waarde 0 aanneemt wanneer de uitkomst geen kruis is, en de waarde 1, wanneer dit wel het geval is. Het totaal aantal kruisen in n worpen is dan gelijk aan de som van n van deze stochastische grootheden.

Hetzelfde geldt voor de kansverdeling van het aantal zessen bij n worpen met een zuivere dobbelsteen. De kansverdeling van dit aantal zessen bij acht worpen wordt weergegeven in Voorbeeld II en figuur 2 van Memorandum S308-A2. Deze verdeling is niet symmetrisch. Indien men het aantal van acht worpen groter neemt zal de verdeling iets minder asymmetrisch worden. Hoe groter het aantal worpen wordt genomen, hoe sterker de symmetrie wordt en hoe meer de kansverdeling nadert tot een normale verdeling.

Een andere toepassing van de onder 3 genoemde eigenschap is eveneens van groot praktisch belang. Uit deze eigenschap volgt namelijk, dat voor voldoend grote n het steekproefgemiddelde

$$\underline{\bar{x}} = \frac{1}{n}\,(\underline{x}_1 + \underline{x}_2 + \ldots + \underline{x}_n) \qquad (5)$$

van n onafhankelijke trekkingen uit een willekeurige verdeling met verwachting μ en standaardafwijking σ bij benadering een normale verdeling heeft en wel één met verwachting μ en standaardafwijking $\frac{\sigma}{\sqrt{n}}$. Past men dus de in (1) aangegeven herleiding toe, dan blijkt de grootheid

$$\underline{u} = \frac{\underline{\bar{x}} - \mu}{\sigma'}\,\sqrt{n} \qquad (6)$$

een $N(0;1)$ - verdeling te bezitten.

Voorbeelden

I. Gegeven is dat $\underline{x}$ een $N(0;1)$ - verdeling heeft. De rechteroverschrijdingskans van $x = 1,96$, dus $P\,[\underline{x} \geq 1,96]$, is gelijk aan 0,0250 (zie de bijgaande tabel). De linkeroverschrijdingskans van $x = -1,645$, dus $P\,[\underline{x} \leq -1,645]$, is gelijk aan $P\,[\underline{x} \geq 1,645]$, dus 0,0500. De rechteroverschrijdingskans van $x = -1,282$, dus $P\,[\underline{x} \geq -1,282]$, is gelijk aan $1 - P\,[\underline{x} \leq -1,282] = 1 - P\,[\underline{x} \geq +1,282] = 1 - 0,1000 = 0,9000$. De tweezijdige overschrijdingskans van $x = 2,576$ is gelijk aan tweemaal de kleinste waarde van $P\,[\underline{x} \geq 2,576]$ en $P\,[\underline{x} \leq 2,576]$; deze zijn resp. gelijk aan 0,0050 en $1 - 0,0050 = 0,9950$. De gevraagde kans is dus gelijk aan $2 \times 0,0050 = 0,0100$.

De boven gebruikte waarden van x zijn niet willekeurig gekozen maar zijn gelijk aan de waarden van ξ_α in Tabel I van Memorandum S308-A4.

II. Gegeven is dat $\underline{x}$ een N (255 ; 5) verdeling heeft (zie figuur 3 van Memorandum S308-A2). Gevraagd wordt de linkeroverschrijdingskans van $x = 250$.

Wij passen eerst formule (1) toe en vinden:

$$u = \frac{250 - 255}{5} = -1.$$

Dus $P\,[\underline{x} \leq 250] = P\,[\underline{u} \leq -1] = P\,[u \geq 1] = 0,1587$.

De rechter overschrijdingskans van x = 265 bedraagt:

$$P\left[\underline{x} \geq 265\right] = P\left[\underline{u} \geq \frac{265 - 255}{5}\right] = P\left[\underline{u} \geq 2\right] = 0{,}0228.$$

De tweezijdige overschrijdingskans van x = 240 volgt uit

$$P\left[\underline{x} \leq 240\right] = P\left[\underline{u} \leq \frac{240 - 255}{5}\right] = P\left[\underline{u} \leq -3\right] = P\left[\underline{u} \geq 3\right] = 0{,}0013$$

en $P\left[\underline{x} \geq 240\right] = P\left[\underline{u} \geq -3\right] = 1 - P\left[\underline{u} \leq -3\right] =$

$= 1 - P\left[\underline{u} \geq +3\right] = 1 - 0{,}0013 = 0{,}9987.$

De gevraagde kans is dus gelijk aan 2 × 0,0013 = 0,0026.

III. Stel $\underline{x}$ is het aantal zessen in één worp met een zuivere dobbelsteen. De kans op een zes, dus $P\left[\underline{x} = 1\right]$, is gelijk aan $\frac{1}{6}$ en de kans op geen zes, dus $P\left[\underline{x} = 0\right]$, is gelijk aan $\frac{5}{6}$. De verwachting van $\underline{x}$ is dus gelijk aan (zie Memorandum S308-A2):

$$\mu = 0 . P\left[\underline{x} = 0\right] + 1 . P\left[\underline{x} = 1\right] = 0.\frac{5}{6} + 1.\frac{1}{6} = \frac{1}{6} .$$

De variantie is gelijk aan:

$$\sigma^2 = \left(0 - \frac{1}{6}\right)^2 . P\left[\underline{x} = 0\right] + \left(1 - \frac{1}{6}\right)^2 . P\left[\underline{x} = 1\right] =$$

$$= \frac{1}{36} . \frac{5}{6} + \frac{25}{36} . \frac{1}{6} = \frac{30}{36.6} = \frac{5}{36} .$$

De standaardafwijking is dus gelijk aan

$$\sigma = \sqrt{\frac{5}{36}} = \frac{1}{6}\sqrt{5} = 0{,}373.$$

Men werpt 200 maal met deze dobbelsteen. Het aantal zessen dat men bij deze worpen krijgt is dan gelijk aan de som van 200 onafhankelijke stochastische grootheden, welke alle dezelfde kansverdeling bezitten met gemiddelde $\mu = \frac{1}{6}$ en standaardafwijking $\sigma = 0{,}373$. Deze som, $\underline{y}$, heeft volgens formule (3) als verwachting

$$\mu_{\underline{y}} = 200 . \frac{1}{6} = 33\frac{1}{3}$$

en volgens formule (4) als standaardafwijking

$$\sigma_{\underline{y}} = \sqrt{200 . \frac{5}{36}} = \frac{5}{3}\sqrt{10} = 5{,}270 .$$

Het aantal opgetelde stochastische grootheden (200) is groot genoeg om de onderstelling, dat $\underline{y}$ een normale verdeling met verwachting

$\mu = 33\frac{1}{3}$ en standaardafwijking $\sigma = 5{,}270$ bezit, te rechtvaardigen.
Met behulp van formule (1) en de bijgevoegde tabel kan dan bijvoorbeeld berekend worden de kans dat bij 200 worpen met een zuivere dobbelsteen 50 of meer zessen waargenomen zullen worden:

$$P\left[\underline{x} \geq 50\right] = P\left[\underline{u} \geq \frac{50 - 33\frac{1}{3}}{5{,}270} = 3{,}16\right] = 0{,}0008.$$

IV. Uit een groot aantal facturen wordt een steekproef van de omvang 400 genomen. De grootte van zo'n factuur kan gezien worden als een stochastische grootheid, die dus een zekere kansverdeling bezit. Van deze kansverdeling is de verwachting μ en de standaardafwijking σ bekend. Stel μ = ƒ 50,-- en σ = ƒ 30,--. Van de steekproef wordt het gemiddelde $\bar{x}$ bepaald. Volgens het voorgaande heeft $\underline{\bar{x}}$ bij een zo grote steekproefomvang een normale verdeling met verwachting μ = ƒ 50,- en standaardafwijking σ = ƒ $\frac{30,-}{\sqrt{400}}$ = ƒ 1,50, dus een N(50 ; 1,5) - verdeling.

Voor elke waarde van $\underline{\bar{x}}$ kan nu de overschrijdingskans berekend worden. Vindt men bijvoorbeeld $\bar{x}$ = ƒ 46,31, dan is $P\left[\underline{\bar{x}} \leq 46{,}31\right] =$
$= P\left[\underline{u} \leq \frac{46{,}31 - 50}{1{,}5} = -2{,}46\right] = 0{,}0069$; $P\left[\underline{\bar{x}} \geq 46{,}31\right] = P\left[\underline{u} \geq -2{,}46\right] =$
$= 1 - P\left[\underline{u} \geq 2{,}46\right] = 0{,}9931$. De tweezijdige overschrijdingskans is dus gelijk aan $2 \times 0{,}0069 = 0{,}0138$.

Voor $\bar{x}$ = ƒ 52,25 is

$$P\left[\underline{\bar{x}} \leq 52{,}25\right] = P\left[\underline{u} \leq \frac{52{,}25 - 50}{1{,}5} = 1{,}5\right] = 1 - 0{,}0668 = 0{,}9332;$$

$$P\left[\underline{\bar{x}} \geq 52{,}25\right] = P\left[\underline{u} \geq 1{,}5\right] = 0{,}0668.$$

De tweezijdige overschrijdingskans is dus gelijk aan $2 \times 0{,}0668 = 0{,}1336$.

Appendix

De normale N(μ ; σ) - verdeling wordt gegeven door

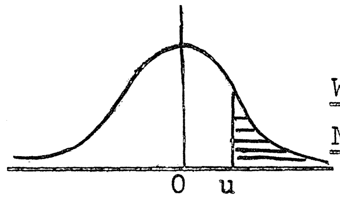
$$P\left[x_1 \leq \underline{x} \leq x_2\right] = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} \exp\left[-\tfrac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx.$$

De rechteroverschrijdingskans van x = ξ is gelijk aan

$$P\left[\xi \leq \underline{x} < \infty\right] = \frac{1}{\sigma\sqrt{2\pi}} \int_{\xi}^{\infty} \exp\left[-\tfrac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx.$$

De linkeroverschrijdingskans van x = ξ is gelijk aan

$$P\left[-\infty \leq \underline{x} \leq \xi\right] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\xi} \exp\left[-\tfrac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx.$$



Waarden van rechteroverschrijdingskansen van de N (0;1) - verdeling, vermenigvuldigd met 10.000.

u	0	1	2	3	4	5	6	7	8	9
0,0	5000	4960	4920	4880	4840	4801	4761	4721	4681	4641
0,1	4602	4562	4522	4483	4443	4404	4364	4325	4286	4247
0,2	4207	4168	4129	4090	4052	4013	3974	3936	3897	3859
0,3	3821	3783	374$\underline{5}$	3707	3669	3632	3594	3557	3520	3483
0,4	3446	3409	3372	3336	3300	3264	3228	3192	3156	3121
0,5	3085	3050	3015	2981	2946	2912	2877	2843	2810	2776
0,6	2743	2709	2676	2643	2611	2578	2546	2514	2483	2451
0,7	2420	2389	2358	2327	2296	2266	2236	2206	2177	2148
0,8	2119	2090	2061	2033	200$\underline{5}$	1977	1949	1922	1894	1867
0,9	1841	1814	1788	1762	1736	1711	1685	1660	1635	1611
1,0	1587	1562	1539	1515	1492	1469	1446	1423	1401	1379
1,1	1357	133$\underline{5}$	1314	1292	1271	1251	1230	1210	1190	1170
1,2	1151	1131	1112	1093	107$\underline{5}$	1056	1038	1020	1003	0985
1,3	0968	0951	0934	0918	0901	0885	0869	0853	0838	0823
1,4	0808	0793	0778	0764	0749	0735	0721	0708	0694	0681
1,5	0668	0655	0643	0630	0618	0606	0594	0582	0571	0559
1,6	0548	0537	0526	0516	0505	049$\underline{5}$	048$\underline{5}$	047$\underline{5}$	046$\underline{5}$	0455
1,7	0446	0436	0427	0418	0409	0401	0392	0384	0375	0367
1,8	0359	0351	0344	0336	0329	0322	0314	0307	0301	0294
1,9	0287	0281	0274	0268	0262	0256	02$\underline{5}$0	0244	0239	0233
2,0	0228	0222	0217	0212	0207	0202	0197	0192	0188	0183
2,1	0179	0174	0170	0166	0162	0158	0154	0150	0146	0143
2,2	0139	0136	0132	0129	0125	0122	0119	0116	0113	0110
2,3	0107	0104	0102	0099	0096	0094	0091	0089	0087	0084
2,4	0082	0080	0078	0075	0073	0071	0069	0068	0066	0064
2,5	0062	0060	0059	0057	0055	0054	0052	0051	0049	0048
2,6	0047	0045	0044	0043	0041	0040	0039	0038	0037	0036
2,7	003$\underline{5}$	0034	0033	0032	0031	0030	0029	0028	0027	0026
2,8	0026	002$\underline{5}$	0024	0023	0023	0022	0021	0021	0020	0019
2,9	0019	0018	0018	0017	0016	0016	0015	001$\underline{5}$	0014	0014
3,0	0013	0013	0013	0012	0012	0011	0011	0011	0010	0010
3,1	0010	0009	0009	0009	0008	0008	0008	0008	0007	0007
3,2	0007	0007	0006	0006	0006	0006	0006	0005	0005	0005
3,3	000$\underline{5}$	000$\underline{5}$	000$\underline{5}$	0004	0004	0004	0004	0004	0004	0003
3,4	0003	0003	0003	0003	0003	0003	0003	0003	0003	0002

Een $\underline{5}$ betekent, dat bij afronding naar beneden afgerond moet worden.

Een 5 wordt naar boven afgerond.

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
AFDELING MATHEMATISCHE STATISTIEK

S 308 - A 4

Betrouwbaarheidsintervallen voor een onbekende fractie

In veel gevallen kan men een onbekende fractie schatten door middel van een aselecte steekproef. Deze fractie kan bijvoorbeeld zijn:

de fractie accuratesse-fouten, gemaakt bij een grote groep van administratieve handelingen;

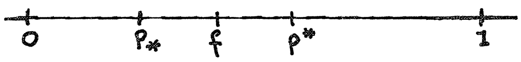
de fractie met omzetbelasting belaste guldens in het totaalbedrag aan verkoopfacturen bij een groothandel;

de fractie onjuiste berekeningen in een loonadministratie.

Laten we het tweede voorbeeld aanhouden. De populatie bestaat dan uit alle guldens in het totaalbedrag van alle verkoopfacturen. Men zoekt nu naar het aantal guldens met het kenmerk: belastbaar uit hoofde van omzetbelasting. Dit kan ook zo uitgedrukt worden: men zoekt naar de fractie belastbare guldens. Deze onbekende fractie noemen wij p.

Uit de populatie wordt nu een steekproef van n stuks op aselecte wijze aangewezen. Stel dat men in deze steekproef k belastbare guldens vindt. De fractie belastbare guldens in de steekproef is dan $f = \frac{k}{n}$ en wij gebruiken nu f als een "schatting" van de onbekende fractie p.

Deze schatting f behoeft uiteraard niet precies gelijk te zijn aan de werkelijke fractie p, maar anderzijds zal het verschil tussen p en f in de meeste gevallen niet zeer groot zijn. Het is dan ook mogelijk een interval op te geven, het zogenaamde "betrouwbaarheidsinterval", waarin de werkelijke fractie belastbare guldens vermoedelijk ligt. Ter illustratie is in figuur 1 het interval van 0 tot 1 getekend. In deze figuur kan de gevonden waarde f precies worden aangegeven. Om deze f is verder het betrouwbaarheidsinterval getekend



Figuur 1

Betrouwbaarheidsinterval voor p

met als ondergrens p_* en als bovengrens p^*. Hoewel p vermoedelijk in dit interval ligt, behoeft de uitspraak: "p ligt tussen p_* en p^*" toch niet altijd juist te zijn. Immers door toevallige omstandigheden zou f wel eens veel af kunnen wijken van p. Men is echter in staat de grenzen p_* en p^* zodanig te bepalen, dat de kans op een onjuiste uitspraak niet groter is dan een van te voren vastgestelde grootte α. Deze kans α wordt de "onbetrouwbaarheid" van de methode genoemd. Men kan dit ook als volgt tot uitdrukking brengen: Indien de beschreven methode vele malen wordt herhaald, dan is de fractie onjuiste uitspraken gemiddeld gelijk aan α.

Verder is wel in te zien, dat, hoe breder men het interval $p_* - p^*$ kiest, hoe kleiner de onbetrouwbaarheid α zal worden. Deze wordt echter slechts dan gelijk aan 0, wanneer men p_* in 0 en p^* in 1 kiest. Indien men van een steekproef gebruik maakt en triviale uitspraken wil vermijden, zal α echter altijd groter dan 0 zijn. Anders gezegd: een niet-triviale uitspraak heeft altijd een zekere onbetrouwbaarheid.

Voor de onbetrouwbaarheid α moet dus vooraf een keuze gedaan worden. De in de praktijk meest gebruikte waarden van α zijn 0,01 en 0,05; verder gebruikt men ook wel 0,001, in gevallen waarin men zeer weinig risico wil nemen.

Voor voldoend grote n kunnen p_* en p^* berekend worden met de volgende, op de zogenaamde "normale benadering" gebaseerde, formules:

$$p_* = f - \xi_\alpha \cdot \sqrt{\frac{f(1-f)}{n}} , \qquad (1)$$

$$p^* = f + \xi_\alpha \cdot \sqrt{\frac{f(1-f)}{n}} . \qquad (2)$$

Hierin is ξ_α een grootheid die van α afhangt; voor verschillende waarden van α is ξ_α opgegeven in tabel I.

Tabel 1

Waarden van ξ_α

α	ξ_α
0,20	1,282
0,10	1,645
0,05	1,960
0,02	2,326
0,01	2,576

Uit de gegeven formules blijkt, dat bij toenemende steekproefomvang n het betrouwbaarheidsinterval voor p zich vernauwt. Ook blijkt uit de formules en de voor ξ_α gegeven tabel dat het betrouwbaarheidsinterval zich verbreedt bij afnemende α. Beide conclusies liggen trouwens ook op intuïtieve gronden reeds voor de hand.

Uit de formules (1) en (2) volgt nog dat de "onnauwkeurigheid" van de schatting gelijk is aan $\xi_\alpha \sqrt{\frac{f(1-f)}{n}}$. Met een onbetrouwbaarheid α zal de werkelijke fractie p dus hoogstens $\xi_\alpha \cdot \sqrt{\frac{f(1-f)}{n}}$ van de schatting f afwijken.

Het hierboven besproken betrouwbaarheidsinterval is tweezijdig begrensd. Er zijn echter ook situaties, waarin men alleen geïnteresseerd is in de ondergrens of de bovengrens van het betrouwbaarheidsinterval. Bij een accuratesse-controle b.v. zal men slechts belangstelling hebben voor de bovengrens van de fractie gemaakte fouten, dus voor p^*. Men kan dezelfde formule (2) en tevens Tabel I gebruiken, mits men de onbetrouwbaarheid α halveert. Voor een eenzijdige onbetrouwbaarheid α zoekt men dan de waarde $\xi_{2\alpha}$ op, die bij 2α behoort. Ook hierbij geldt weer, dat n groot genoeg moet zijn.

In het geval, waarin de boven vermelde benaderingsformules te onnauwkeurig zijn, is wel dezelfde gedachtengang mogelijk, doch dienen andere formules gebruikt te worden. Deze formules worden hier niet vermeld. Voor het geval dat f heel klein en n groot is kunnen p_* en p^* uit Tabel II afgelezen worden, welke gebaseerd is op de zogenaamde Poisson-benadering. Ook hier geldt, dat bij het bepalen van een naar boven of naar beneden eenzijdig begrensd betrouwbaarheidsinterval de onbetrouwbaarheid α gehalveerd dient te worden.

Tabel II

Grenzen voor np_* en np^*, wanneer n groot is en f klein

K \ α	0,01		0,02		0,05		0,10	
	np_*	np^*	np_*	np^*	np_*	np^*	np_*	np^*
0	0	5,30	0	4,61	0	3,69	0	3,00
1	0,005	7,44	0,010	6,64	0,025	5,58	0,051	4,75
2	0,10	9,28	0,14	8,41	0,24	7,23	0,35	6,30
3	0,33	10,98	0,43	10,05	0,61	8,77	0,81	7,76
4	0,67	12,60	0,82	11,61	1,08	10,25	1,36	9,16
5	1,07	14,15	1,27	13,11	1,62	11,67	1,97	10,52
6	1,53	15,66	1,78	14,58	2,20	13,06	2,61	11,85
7	2,03	17,14	2,32	16,01	2,81	14,43	3,28	13,15
8	2,57	18,58	2,90	17,41	3,45	15,77	3,98	14,44
9	3,13	20,00	3,50	18,79	4,11	17,09	4,69	15,71
10	3,71	21,40	4,12	20,15	4,79	18,40	5,42	16,97

Tabel II is ook bruikbaar indien f zeer dicht bij 1 ligt. Men gebruike dan niet de grootheid k maar de grootheid (n-k). Uit de zo gevonden grenzen voor (1-p) zijn die voor f eenvoudig af te leiden.

Voorbeelden

I. Bij een groothandel met een omzet van enkele tientallen millioenen guldens wil men door middel van een steekproef bepalen welk bedrag aan omzetbelasting verschuldigd is. De populatie bestaat hier uit alle guldens van alle verkoopfacturen van een geheel jaar. Men neemt een aselecte steekproef van de omvang $n = 10.000$ en vindt hierin $k = 3.500$ belastbare guldens. Voor de onbetrouwbaarheid wordt 5% ($\alpha = 0{,}05$) gekozen. Om de onder- en bovengrens te vinden vult men in de formules (1) en (2) voor f, n en ξ_α resp. 0,35, 10.000 en 1,960 in. Men vindt dan $p_* = 0{,}3407$ en $p^* = 0{,}3593$. Met een onbetrouwbaarheid van 5% ligt de werkelijke fractie belastbare guldens dus tussen 34,07% en 35,93%.

II. Bij een accuratesse-controle accepteert men een lijst indien de bovengrens van het betrouwbaarheidsinterval van de fractie gemaakte fouten nog beneden 0,05 ligt. In een steekproef van omvang 500 vindt men 16 fouten. Bij een onbetrouwbaarheid van 5% ($\xi_\alpha = 1{,}645$!) vindt men voor de bovengrens $p^* = 0{,}045$. De lijst wordt dus nog juist geaccepteerd.

III. Bij een controle accepteert men een lijst slechts dan indien de bovengrens van het betrouwbaarheidsinterval van de fractie niet door stukken gedekte guldens nog beneden 0,005 ligt.
In een steekproef van omvang 1000 treft men 1 niet gedekte gulden aan. Bij een onbetrouwbaarheid van 1% ($\alpha = 0{,}01$) lezen wij in Tabel II af: $np^* = 6{,}64$, dus $p^* = \frac{6{,}64}{n} = \frac{6{,}64}{1000} = 0{,}00664$. p^* is dus groter dan 0,005 en de lijst wordt niet geaccepteerd.

Appendix

In de literatuur treft men de formules (1) en (2) vaak in een andere vorm aan:

$$p^{*}_{*} = \frac{x}{n} \pm \frac{\xi_{\alpha}}{n} \cdot \sqrt{\frac{x(n-x)}{n}} \qquad \text{(1a en 2a)}$$

Hierin is x het aantal in de steekproef gevonden exemplaren met het gezochte kenmerk. Men verkrijgt deze formules uit (1) en (2) door hierin f door $\frac{x}{n}$ te vervangen.

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S308-A5

Betrouwbaarheidsintervallen voor de verwachting van een normale verdeling.

Een normale verdeling is geheel bepaald door de verwachting μ en de standaardafwijking σ. Voor nadere toelichting verwijzen wij naar Memorandum S308-A3. Indien men van een stochastische grootheid $\underline{x}$ slechts weet, dat hij normaal verdeeld is, dan zal men, om de verdeling geheel te kunnen bepalen, μ en σ moeten "schatten". Stel dat men door middel van een aselecte steekproef de waarnemingen $x_1, x_2, \ldots\ldots, x_n$ van $\underline{x}$, heeft verkregen, dan kunnen μ en σ op basis van deze gegevens "geschat" worden.

Hieronder zullen we eerst het geval bespreken dat de grootte van σ reeds bekend is en daarna het geval dat μ noch σ gegeven zijn.

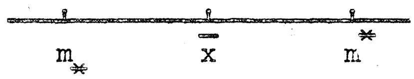
De grootte van σ is reeds bekend.

Een aselecte steekproef van omvang n leverde de waarden x_1, x_2, x_3, ..., x_n op. De grootheid

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + x_3 + \ldots x_n) \qquad (1)$$

wordt nu als een "schatting" van de onbekende waarde μ gebruikt. Deze schatting $\bar{x}$ behoeft uiteraard niet precies gelijk te zijn aan de verwachting μ, maar anderzijds zal het verschil tussen μ en $\bar{x}$ in de meeste gevallen niet erg groot zijn. Het is dan ook mogelijk een interval op te geven, een zogenaamd "betrouwbaarheidsinterval", waarin de verwachting μ vermoedelijk ligt.

Ter illustratie is in onderstaande figuur de gevonden waarde $\bar{x}$ aangegeven. Om deze $\bar{x}$ is verder het betroubaarheidsinterval getekend met als ondergrens m_* en als bovengrens m^*.



Figuur 1

Betrouwbaarheidsinterval voor μ

Hoewel μ vermoedelijk in dit interval ligt, behoeft de uitspraak: "μ ligt tussen m_* en m^* " toch niet juist te zijn. Door toevallige omstandigheden zou $\bar{x}$ wel eens veel van μ af kunnen wijken. Men is echter in staat m_* en m^* zodanig te kiezen, dat de kans op een onjuiste uitspraak niet groter is dan een van te voren vastgestelde grootte α. Deze kans α wordt de "onbetrouwbaarheid" van de methode genoemd. Men kan dit ook als volgt tot uitdrukking brengen: Indien de beschreven methode vele malen wordt herhaald, dan is de fractie onjuiste uitspraken gemiddeld gelijk aan α.

Verder is wel in te zien, dat hoe breder men het interval m_*——m^* kiest, hoe kleiner de onbetrouwbaarheid α zal worden. Deze wordt echter slechts dan gelijk aan 0 wanneer men voor m_* de waarde $-\infty$ en voor m^* de waarde $+\infty$ kiest. Indien men van een steekproef gebruik maakt en triviale uitspraken wil vermijden, zal echter α altijd groter dan nul zijn. Anders gezegd: een niet - triviale uitspraak heeft altijd een zekere onbetrouwbaarheid.

Voor de onbetrouwbaarheid α zal dus vooraf een keuze gedaan moeten worden. De in de praktijk meest gebruikte waarden van α zijn 0,01 en 0,05 (1% en 5%); verder gebruikt men ook wel de waarde 0,001 (1‰) in gevallen waarbij men zeer weinig risico wil nemen.

De grenzen m_* en m^* kunnen berekend worden met de volgende formules:

$$m_* = \bar{x} - \xi_\alpha \cdot \frac{\sigma}{\sqrt{n}}, \qquad (2)$$

$$m^* = \bar{x} + \xi_\alpha \cdot \frac{\sigma}{\sqrt{n}}. \qquad (3)$$

Hierin is ξ_α een grootheid die van α afhangt; voor verschillende waarden van α is ξ_α opgegeven in Tabel I.

Tabel I

Waarden van ξ_α.

α	ξ_α
0,20	1,282
0,10	1,645
0,05	1,960
0,02	2,326
0,01	2,576

Uit de gegeven formules blijkt, dat bij toenemende steekproefomvang n het betrouwbaarheidsinterval voor μ zich vernauwt. Ook blijkt uit de formules en Tabel I dat het betrouwbaarheidsinterval zich verbreedt bij afnemende α, en andersom. Beide conclusies liggen trouwens ook op intuïtieve gronden voor de hand.

Uit (2) en (3) volgt, dat de "onnauwkeurigheid" van de schatting $\bar{x}$ gelijk is aan $\xi_\alpha \frac{\sigma}{\sqrt{n}}$. Met een onbetrouwbaarheid α zal de werkelijke verwachting dus hoogstens $\xi_\alpha \frac{\sigma}{\sqrt{n}}$ van de schatting $\bar{x}$ afwijken.

Het hierboven besproken betrouwbaarheidsinterval is tweezijdig begrensd. Er zijn echter ook situaties waarin men slechts geïnteresseerd is in de ondergrens of de bovengrens van het betrouwbaarheidsinterval. Men kan dan formule (2) resp. formule (3) en tevens Tabel I gebruiken, mits men de onbetrouwbaarheid α in de tabel halveert. Wil men dus een eenzijdig begrensd interval met een onbetrouwbaarheid α, dan dient men de waarde $\xi_{2\alpha}$ te gebruiken, die bij 2α behoort.

De grootte van σ is niet bekend.

Daar σ onbekend is, is het niet mogelijk de grenzen m_* en m^* direct met de formules (2) en (3) te berekenen. Het is daarom noodzakelijk eerst een schatting s van σ te maken. Hiertoe berekenen wij de grootheid

$$s = \sqrt{\frac{1}{n-1}\left\{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots\ldots + (x_n - \bar{x})^2\right\}}. \qquad (4)$$

Voor voldoend grote n wordt het betrouwbaarheidsinterval nu bepaald door de volgende formules:

$$m_* = \bar{x} - \xi_\alpha \cdot \frac{s}{\sqrt{n}} , \qquad (5)$$

$$m^* = \bar{x} + \xi_\alpha \cdot \frac{s}{\sqrt{n}} . \qquad (6)$$

Hierin is ξ_α weer de van α afhankelijke grootheid die in Tabel I gevonden kan worden.

In de meeste gevallen is σ niet bekend en zullen de formules (5) en (6) gebruikt worden. De formules (2) en (3) gelden voor alle waarden van n terwijl voor de toepassing van de formules (5) en (6) n minstens groter dan 30 moet zijn en in sommige gevallen zelfs groter dan 100 teneinde voldoend nauwkeurige grenzen te vinden.

Uit (5) en (6) volgt ook nu weer dat de "onnauwkeurigheid" van de schatting $\bar{x}$ gelijk is aan $\xi_\alpha \cdot \frac{s}{\sqrt{n}}$. Met een onbetrouwbaarheid α zal de werkelijke verwachting dus hoogstens $\xi_\alpha \cdot \frac{s}{\sqrt{n}}$ van de schatting $\bar{x}$ afwijken.

Voorbeelden

I. Uit de waarnemingen van vele jaren is gebleken dat de lengte van recruten normaal verdeeld is. Uit de gemeten lengte van 400 recruten van een bepaalde lichting is voor het gemiddelde 1,76 m en voor de standaardafwijking 0,10 berekend ($\bar{x} = 1{,}76$ en $s = 0{,}10$). Bij een onbetrouwbaarheid van 5% ($\alpha = 0{,}05$) berekent men met (5) en (6) en Tabel I ($\xi_\alpha = 1{,}960$) voor m_* en m^* resp. de waarden 1,750 en 1,770. Met een onbetrouwbaarheid van 5% ligt de gemiddelde lengte van alle recruten van deze lichting dus tussen 1,750 m en 1,770 m.

II. Wenst men in het bovenstaande geval alleen een bovengrens van het gemiddelde te berekenen met een onbetrouwbaarheid van 5%, dan moet in Tabel I ξ_α opgezocht worden voor $\alpha = 0{,}10$. Men vindt $\xi_\alpha = 1{,}645$ en formule (6) leidt tot $m^* = 1{,}768$ m. Met een onbetrouwbaarheid van 5% ligt de gemiddelde lengte van alle recruten dus beneden 1,768 m.

(De gegevens van deze voorbeelden werden niet aan de werkelijkheid ontleend).

Appendix

Een directe toepassing van formule (4) is zeer omslachtig. Het rekenwerk kan bekort worden door (4) enigszins anders te schrijven en wel als volgt:

$$(x_1 - \bar{\bar{x}})^2 + (x_2 - \bar{\bar{x}})^2 + \ldots\ldots + (x_n - \bar{\bar{x}})^2 =$$

$$= (x_1^2 - 2\bar{\bar{x}}x_1 + \bar{\bar{x}}^2) + (x_2^2 - 2\bar{\bar{x}}x_2 + \bar{\bar{x}}^2) + \ldots\ldots\ldots (x_n^2 - 2\bar{\bar{x}}x_n + \bar{\bar{x}}^2) =$$

$$= (x_1^2 + x_2^2 + \ldots + x_n^2) - 2\bar{\bar{x}}(x_1 + x_2 + \ldots + x_n) + n\bar{\bar{x}}^2 =$$

$$= (x_1^2 + x_2^2 + \ldots + x_n^2) - 2\bar{\bar{x}}.n\bar{\bar{x}} + n\bar{\bar{x}}^2 =$$

$$= (x_1^2 + x_2^2 + \ldots + x_n^2) - n\bar{\bar{x}}^2 .$$

Dit ingevuld in (4) leidt tot:

$$s = \sqrt{\frac{1}{n-1} \left\{ (x_1^2 + x_2^2 + \ldots + x_n^2) - n\bar{\bar{x}}^2 \right\}}. \qquad (7)$$

W

A

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S 308-A 6

Betrouwbaarheidsintervallen voor de verwachting van een willekeurige verdeling.

In Memorandum S 308-A 5 werd het bepalen van betrouwbaarheidsintervallen voor de verwachting van een normale verdeling besproken. In veel situaties zal de verdeling, waaruit de steekproef genomen wordt echter niet normaal zijn, of zelfs in het geheel niet bekend. Ook dan is het onder zekere voorwaarden mogelijk betrouwbaarheidsintervallen te construeren.

Stel dat men door middel van een aselecte steekproef het gemiddelde bedrag van de inkoopfacturen in een bepaald boekjaar wil schatten. De kansverdeling van de bedragen is onbekend. Er wordt uit alle inkoop-facturen een aselecte steekproef van de omvang n genomen. De in de steekproef gevonden waarden zijn

$$x_1, x_2, \ldots, x_n.$$

Het gemiddelde bedrag $\bar{x}$ in deze steekproef is

$$\bar{x} = \frac{1}{n} (x_1 + x_2 + \ldots + x_n). \qquad (1)$$

Dit gemiddelde $\bar{x}$ zal in het algemeen niet gelijk zijn aan de verwachting μ van alle inkoopfacturen. Het is echter intuftief duidelijk, dat bij een grote steekproef de gevonden waarde $\bar{x}$ niet veel van μ zal afwijken. Het is dan ook mogelijk een zogenaamd "betrouwbaarheidsinterval" op te geven, waarin de verwachting μ vermoedelijk ligt.

Ter illustratie is in figuur 1 een getallenrechte getekend. Op deze rechte kan de gevonden waarde $\bar{x}$ precies worden aangegeven. Om deze $\bar{x}$ is verder een betrouwbaarheidsinterval getekend met als bovengrens m^* en als ondergrens m_*.

Figuur 1.

Betrouwbaarheidsinterval voor μ.

Hoewel μ vermoedelijk in dit interval ligt, behoeft de uitspraak "μ ligt tussen m_* en m^*" toch niet altijd juist te zijn. Immers door toevallige omstandigheden zou $\bar{x}$ wel eens veel van μ af kunnen wijken. Voor grote waarden van n is men echter in staat de grenzen m_* en m^* zodanig te bepalen dat de kans op een onjuiste uitspraak niet groter is dan een van te voren vastgestelde grootte α. Deze kans α wordt de "onbetrouwbaarheid" van de methode genoemd. Men kan dit ook als volgt tot uitdrukking brengen: Indien de beschreven methode vele malen wordt herhaald, dan is de fractie onjuiste uitspraken gemiddeld gelijk aan α.

Verder is wel in te zien, dat, hoe breder men het interval $m_* \text{---} m^*$ kiest, hoe kleiner de onbetrouwbaarheid α zal worden. Deze wordt echter slechts dan gelijk aan 0, wanneer men voor m_* de kleinst mogelijke waarde van $\bar{x}$ (kleinste factuur) en voor m^* de grootst mogelijke waarde van $\bar{x}$ (grootste factuur) kiest. Indien men echter bij het gebruik maken van een steekproef triviale uitspraken wil vermijden zal α altijd groter dan 0 zijn. Anders gezegd: een niet-triviale uitspraak heeft altijd een zekere onbetrouwbaarheid.

Voor de onbetrouwbaarheid α moet dus vooraf een keuze gedaan worden. De in de praktijk meest gebruikte waarden van α zijn 0,05 en 0,01; verder gebruikt men ook wel 0,001 in gevallen waarin men zeer weinig risico wil nemen.

Voor <u>voldoend grote n</u> kunnen m_* en m^* met de volgende formules berekend worden:

$$m_* = \bar{x} - \xi_\alpha \frac{s}{\sqrt{n}} \tag{2}$$

$$m^* = \bar{x} + \xi_\alpha \frac{s}{\sqrt{n}} \tag{3}$$

Hierin is s:

$$s = \sqrt{\frac{1}{n-1}\left\{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2\right\}} \tag{4}$$

Verder is de grootheid ξ_α in (2) en (3) afhankelijk van de gekozen onbetrouwbaarheid α. Tabel 1 bevat de waarden ξ_α voor verschillende waarden van α.

Tabel I

Waarden van ξ_α

α	ξ_α
0,20	1,282
0,10	1,645
0,05	1,960
0,02	2,326
0,01	2,576

Uit (2) en (3) blijkt wel, dat bij toenemende steekproefomvang n het betrouwbaarheidsinterval voor μ zich vernauwt. Ook blijkt uit deze formules en Tabel I dat het betrouwbaarheidsinterval zich verbreedt bij afnemende α. Beide conclusies liggen op intuïtieve gronden voor de hand.

Uit (2) en (3) volgt nog dat de "onnauwkeurigheid" van de schatting $\bar{x}$ gelijk is aan $\xi_\alpha \cdot \frac{s}{\sqrt{n}}$. Met een onbetrouwbaarheid α zal de verwachting μ dus hoogtens $\xi_\alpha \cdot \frac{s}{\sqrt{n}}$ van de schatting $\bar{x}$ afwijken.

Het hierboven besproken betrouwbaarheidsinterval is tweezijdig begrensd. In sommige situaties zal men slechts geïnteresseerd zijn in de bovengrens, of de ondergrens. Men kan dan dezelfde formules gebruiken, mits men de onbetrouwbaarheid α in de tabel halveert. Wil men dus een eenzijdig begrensd interval met een onbetrouwbaarheid α, dan dient men de waarde $\xi_{2\alpha}$ te gebruiken, die bij 2α behoort.

Voorbeelden

I. Bij een groothandel wil men d.m.v. een aselecte steekproef uit de populatie van alle guldens van alle inkoopfacturen van een bepaald boekjaar een schatting maken van het over dat boekjaar gemaakte brutowinstpercentage. Men neemt een steekproef van de omvang 900 en bepaalt bij iedere aangewezen gulden het winstpercentage. Met (1) en (4) worden $\bar{x}$ en s berekend op resp. 45,50% en 8,4%. De formules (2) en (3) geven bij een onbetrouwbaarheid $\alpha = 0,05$ ($\xi_{\alpha} = 1,960$) nu de grenzen van het betrouwbaarheidsinterval:

$$m_* = 45,50 - 1,960 \times \frac{8.4}{\sqrt{900}} = 45,50 - 0,55 = 44,95\% ;$$

$$m^* = 45,50 + 1,960 \times \frac{8,4}{\sqrt{900}} = 45,50 + 0,55 = 46,05\% .$$

Met een onbetrouwbaarheid van 5% ligt het werkelijke brutowinstpercentage μ dus tussen 44,95 en 46,05.

II. Stel dat men in het vorige voorbeeld slechts in de ondergrens van het betrouwbaarheidsinterval was geïnteresseerd bij dezelfde onbetrouwbaarheid $\alpha = 0,05$. In Tabel I wordt nu de waarde van ξ opgezocht die bij 2α behoort. Met behulp van (2) vinden wij voor de ondergrens m_*:

$$m_* = 45,50 - 1,645 \times \frac{8,4}{\sqrt{900}} = 45,50 - 0,46 = 45,04\% .$$

Met een onbetrouwbaarheid van 5% ligt het werkelijke brutowinstpercentage μ dus boven 45,04.

Appendix

In de Appendix van Memorandum S 308-A 5 wordt aangetoond, dat formule (4) voor s ook als volgt geschreven kan worden:

$$s = \sqrt{\frac{1}{n-1}\left\{(x_1^2 + x_2^2 + \dots + x_n^2) - n\,\bar{x}^{\,2}\right\}} \qquad (5).$$

Meestal kan s met (5) vlugger berekend worden dan met (4).

Het in dit Memorandum vermelde betrouwbaarheidsinterval is afgeleid met behulp van de eigenschap, dat $\bar{\underline{x}}$ voor voldoend grote n bij benadering normaal verdeeld is met verwachting μ en standaardafwijking $\frac{\sigma}{\sqrt{n}}$, wanneer $\underline{x}$ een verdeling heeft met verwachting μ en standaardafwijking σ. (Zie de Memoranda S 308-A 2 en A 3). Voor voldoend grote n kan voor de standaardafwijking σ de schatting s ingevuld worden.

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

W

A

S308 - A7

Bepaling van de steekproefomvang

In de Memoranda S308-A4, A5 en A6 werd besproken hoe een aselecte steekproef gebruikt kan worden om een onbekende fractie (A4) of een onbekende verwachting (A5 en A6) te schatten. Bij schatting in de vorm van betrouwbaarheidsintervallen hangt de lengte van het interval o.a. af van de onbetrouwbaarheid α en de steekproefomvang n.

Uit de in de genoemde Memoranda vermelde formules blijkt, dat de lengte van het betrouwbaarheidsinterval, bij gelijkblijvende onbetrouwbaarheid α, kleiner wordt naarmate de steekproefomvang n groter wordt genomen. In de meeste gevallen heeft men reeds vóór de steekproef genomen is, enig idee omtrent de vereiste nauwkeurigheid van de te maken schattingen. Het is dan van belang na te gaan, hoe groot de omvang van de steekproef moet zijn, om deze nauwkeurigheid te bereiken. Hieronder zullen de verschillende gevallen afzonderlijk besproken worden.

Schatting van een onbekende fractie

Voor grote waarden van n is de lengte van het onbetrouwbaarheidsinterval gelijk aan tweemaal de onnauwkeurigheid van de schatting, te weten

$$\xi_\alpha \sqrt{\frac{f(1-f)}{n}} . \qquad (1)$$

Hierin is f de in de steekproef gevonden fractie met de gezochte eigenschap en wordt ξ_α bepaald door de gekozen onbetrouwbaarheid α. Vergelijk de formules (1) en (2) van A4.

Aan deze onnauwkeurigheid kunnen nu de volgende eisen gesteld worden:

a) men kan eisen, dat deze niet groter is dan een gegeven waarde v, dus onafhankelijk van de gevonden fractie f (absolute onnauwkeurigheid);

b) men kan eisen, dat deze niet groter is dan een bepaalde fractie θ van f (relatieve onnauwkeurigheid).

In geval a) moet dan voldaan zijn aan

$$\xi_\alpha \sqrt{\frac{f(1-f)}{n}} \leqq v,$$

waaruit volgt

$$\xi_\alpha^2 \frac{f(1-f)}{n} \leqq v^2,$$

of

$$n \geqq \frac{\xi_\alpha^2}{v^2} f(1-f). \qquad (2).$$

Hierin is de fractie f nog niet bekend, omdat de steekproef nog niet is genomen. De vorm f(1-f) is echter maximaal voor $f = \frac{1}{2}$ en bedraagt dan $\frac{1}{4}$. Indien men dus n bepaalt door

$$n \geqq \frac{\xi_\alpha^2}{4v^2}, \qquad (3)$$

dan zit men altijd aan de veilige kant. Heeft men, vóór de steekproef genomen wordt, reeds een indruk omtrent de grootte van de onbekende p, dan kan men in (2) voor f de ongunstigste, mogelijk geachte, waarde invullen.

Hierbij is dié waarde het ongunstigst, welke het dichtst bij $\frac{1}{2}$ ligt.

In geval b) moet voldaan zijn aan

$$\xi_\alpha \sqrt{\frac{f(1-f)}{n}} \leqq \theta f,$$

waaruit volgt

$$n \geqq \frac{\xi_\alpha^2}{\theta^2} \left(\frac{1}{f} - 1\right). \tag{4}$$

Ook hierin is f nog niet bekend. De vorm $(\frac{1}{f} - 1)$ heeft nu echter geen maximum, maar wordt groter naarmate men f kleiner kiest. Het is nu dus noodzakelijk bij voorbaat enige kennis omtrent f te bezitten, wil men de vereiste steekproefomvang kunnen bepalen. Neemt men aan, dat f zeker boven een waarde f_o zal liggen, dan vult men deze waarde in (4) in en bepaalt zo de te kiezen n.

In Memorandum S308-A4 werd opgemerkt, dat het betrouwbaarheidsinterval voor de onbekende fractie niet met de formules (1) en (2) bepaald kan worden, indien deze fractie dicht bij 0, of bij 1 ligt. Men dient dan de onder- en/of bovengrens van het interval te bepalen met behulp van Tabel II van genoemd Memorandum. Ook in dit geval kan de vereiste omvang van de te nemen steekproef vastgesteld worden, wanneer men verlangt dat de lengte van het betrouwbaarheidsinterval niet groter mag zijn dan een vooraf gegeven waarde. Verder is het noodzakelijk reeds enige kennis omtrent de werkelijke fractie te bezitten. Zie verder de Voorbeelden IV en V.

Schatting van de verwachting van een kansverdeling.

De onnauwkeurigheid van de schatting

$$\bar{x} = \frac{1}{n} (x_1 + x_2 \ldots + x_n) \tag{5}$$

voor de verwachting van een normale verdeling bedraagt

$$\xi_\alpha \cdot \frac{\sigma}{\sqrt{n}} . \tag{6}$$

Hierin is σ de standaardafwijking van de desbetreffende normale verdeling.

Is σ niet bekend, dan wordt deze geschat door

$$s = \sqrt{\frac{1}{n-1} \quad \{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \ldots + (x_n-\bar{x})^2\}} \tag{7}$$

en wordt de onnauwkeurigheid bepaald door

$$\xi_\alpha \cdot \frac{s}{\sqrt{n}} . \tag{8}$$

De formules (6) en (8) gelden ook voor de onnauwkeurigheid van de schatting (5) van de verwachting van een willekeurige verdeling, wanneer n voldoende groot is om de normale benadering van de verdeling van $\underline{\bar{x}}$ te rechtvaardigen (zie Memorandum A6).

Aan de onnauwkeurigheid kan bijvoorbeeld de eis gesteld worden, dat deze niet groter mag zijn dan een vooraf gegeven waarde d.

Indien σ bekend is, zijn er geen moeilijkheden, want dan wordt n eenvoudig bepaald uit

$$\xi_\alpha \cdot \frac{\sigma}{\sqrt{n}} \leq d,$$

waaruit volgt

$$\xi_\alpha^2 \cdot \frac{\sigma^2}{n} \leq d^2,$$

of

$$n \geq \xi_\alpha^2 \cdot \frac{\sigma^2}{d^2} . \tag{9}$$

Het wordt moeilijker, wanneer σ niet bekend is. Nadat de steekproef genomen is wordt σ geschat met (7). Wil men een eerste indruk van de grootte van s krijgen, dan kan men een kleine steekproef nemen, hieruit een voorlopige schatting van σ maken en deze in (9) invullen. De voorlopige steekproef wordt dan uitgebreid tot de op deze wijze gevonden grootte van n is verkregen. Zijn de resultaten nog niet nauwkeurig genoeg, dan kan men met de nieuwe schatting van s opnieuw de gewenste steekproefomvang berekenen en de reeds gedane waarnemingen tot dit aantal aanvullen.

Behalve langs deze weg is het in veel situaties ook mogelijk met andere methoden een indruk van σ te verkrijgen.

Voorbeelden.

I. Via een enquete wil men de aanhang van een bepaalde politieke partij schatten. De (absolute) onnauwkeurigheid mag niet groter zijn dan 1% bij een onbetrouwbaarheid van 5%.

Het aantal te ondervragen personen wordt berekend met (3). Uit $\alpha = 0,05$ volgt $\xi_\alpha = 1,960$, verder is $v = 0,01$, dus

$$n \geq \frac{(1,960)^2}{4(0,01)^2} = \frac{3,8416}{0,0004} = 9604$$

Er moeten dus minstens 9604 aselect aangewezen personen ondervraagd worden.

In het bovenstaande is ondersteld, dat over de fractie in het geheel geen informatie aanwezig was. Neemt men aan, dat deze zeker niet groter zal zijn dan bijvoorbeeld 0,25, dan vindt men de steekproefomvang door in (2) voor f de waarde 0,25 in te vullen. Dit leidt tot

$$n \geq \frac{(1,960)^2}{(0,01)^2}(0,25)(0,75) = 7203,$$

hetgeen aanzienlijk lager is dan de boven gevonden waarde.

II. Bij een spaarbank wil men het percentage rekeninghouders met saldi beneden *f* 100,-- schatten. De vereiste (relatieve) onnauwkeurigheid is 10% en voor de onbetrouwbaarheid kiest men 1%.

Uit de populatie van alle rekeninghouders wordt een steekproef genomen, waarvan de omvang bepaald werd met (4). Nu is $\alpha = 0,01$, dus $\xi_\alpha = 2,576$, en $\theta = 0,10$. Men vermoedt, dat het gezochte percentage niet lager ligt dan 20% en substitueert daarom $f = 0,20$ en vindt

$$n \geq \frac{(2,576)^2}{(0,10)^2} \left(\frac{1}{0,20} - 1\right) = 2654,31$$

Er moeten dus minstens 2655 rekeninghouders aselect aangewezen worden.

III. Bij een handelsfirma dient via een aselecte guldenssteekproef uit alle inkoopfacturen het brutowinstpercentage bepaald te worden. Bij een onbetrouwbaarheid van 1% mag de onnauwkeurigheid hoogstens $\frac{1}{2}$ % bedragen.

Wanneer de standaardafwijking van de verdeling van de winstpercentages niet bekend is, kan men beginnen met een voorlopige steekproef van de omvang 100. De schatting s van σ wordt hierin berekend op 6,3%. De nieuwe steekproefomvang volgt uit (9) met $\alpha = 0,01$, $\xi_\alpha = 2,576$, $d = 0,5$ en $s = 6,3$:

$$n \geq (2,576)^2 \cdot \frac{(6,3)^2}{(0,5)^2} \approx 1053,5.$$

De oorspronkelijke steekproef moet dus met 954 aselecte trekkingen worden uitgebreid.

IV. Een lijst met geboekte facturen werd reeds gecontroleerd. De fractie fouten die op deze lijst nog is blijven zitten zal derhalve zeer klein zijn en men veronderstelt, dat deze fractie niet groter is dan 0,2%. Men wil deze fractie schatten door middel van een aselecte steekproef. De toegestane onbetrouwbaarheid is 1% en men eist, dat de lengte van het betrouwbaarheidsinterval niet groter is dan 0,5%.

Wanneer de onderstelling juist is, zullen bij een steekproefomvang n = 1000 hoogstens 2 fouten gevonden worden. Voor $\alpha = 0{,}01$ en k = 2 vindt men (vgl. Tabel II van Memorandum S308-A4)

$$np^{\ast} - np_{\ast} = n(p^{\ast} - p_{\ast}) = 9{,}18,$$

dus

$$p^{\ast} - p_{\ast} \approx 0{,}009.$$

Dit interval is langer dan 0,005.

Bij n=2000 zal men niet meer dan 4 fouten vinden. Voor k = 4 vindt men

$$n(p^{\ast} - p_{\ast}) = 11{,}93,$$

dus

$$p^{\ast} - p_{\ast} \approx 0{,}006.$$

Bij n = 3000 vindt men op dezelfde manier $p^{\ast} - p_{\ast} \approx 0{,}0047$. De steekproefomvang kan dus iets kleiner dan 3000 gekozen worden. Bij n = 2850 ($k \leq 6$) vindt men tenslotte $p^{\ast} - p_{\ast} = 0{,}00496$.

Op de hier aangegeven wijze kan men dus door proberen een indruk krijgen van de vereiste steekproefomvang.

V. Bij een bepaalde controle stelt men als eis, dat de bovengrens van het eenzijdige betrouwbaarheidsinterval niet boven 0,1% mag liggen, wanneer men geen enkele ongedekte gulden in de guldenssteekproef aantreft. Men kiest voor de onbetrouwbaarheid $\alpha = 0{,}005$. In Tabel II van Memorandum S308-A4 vindt men bij $\alpha = 0{,}01$ (de tabel is tweezijdig!) en k=0 voor np^* de waarde 5,30. Omdat p^* maximaal 0,001 mag bedragen vindt men voor de minimale waarde van n:

$$\frac{5{,}30}{0{,}001} = 5300.$$

W

A

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S 308 - A 8

Aselecte Getallen

Wanneer men van een bepaalde populatie een eigenschap wil onderzoeken, dan kan men uiteraard deze eigenschap bij alle elementen nagaan. In veel gevallen is dit echter praktisch vrijwel onmogelijk of te kostbaar en men kan dan het volledige onderzoek vervangen door een aselecte steekproef.

Het nemen van een aselecte steekproef wil zeggen, dat men uit de populatie een monster neemt volgens een voorschrift, dat onafhankelijk is van de met de steekproef te onderzoeken eigenschap. Aan deze aan het voorschrift gestelde eis is in ieder geval voldaan indien:
a) elk element dezelfde kans heeft in het monster te worden opgenomen,
en b) de trekkingen onafhankelijk van elkaar plaatsvinden.

Een hulpmiddel om aan deze voorwaarden te voldoen wordt gevonden in een tabel met aselecte cijfers. In deze tabellen zijn de cijfers 0 t/m 9 in grote getale in een willekeurige volgorde opgenomen. Aan dit Memorandum zijn twee van dergelijke tabellen toegevoegd [1].

Voegt men in zo'n tabel twee of meer kolommen van één cijfer bij elkaar, dan ontstaan getallen van twee of meer cijfers welke aselecte getallen worden genoemd. Eenzelfde getal kan meerdere malen op een blad-

---

1) Deze tabellen zijn overgenomen uit Tracts for Computers, No. XXIV, "Tables of Random Sampling Numbers" by M.G. Kendall and B. Babington Smith, blz. 11, series 17 en 18.

zijde voorkomen, zelfs is het mogelijk dat hetzelfde getal tweemaal achter elkaar wordt aangetroffen, zoals het getal 71 in de kolommen 15 en 16 van de eerste tabel.

Bij het gebruik van de tabel is het gewenst, dat de elementen van de populatie genummerd zijn of genummerd kunnen worden. Soms is het mogelijk de elementen door meting aan te wijzen, zoals bij een bak met kaarten waar iedere kaart een bepaalde afstand tot de eerste kaart heeft. Bij een lijst met facturen, waaruit een guldenssteekproef genomen moet worden, kan men vaak van aanwezige subtotalen gebruik maken om het zoeken van een aangewezen gulden te vereenvoudigen.

De steekproef kan er één zijn met teruglegging, zodat ieder element kans loopt meerdere malen getrokken te worden, of één zonder teruglegging, waarbij ieder element dus slechts éénmaal getrokken kan worden. In het eerste geval gebruikt men de tabel zonder meer. In het tweede geval wordt een getal, dat al eens is voorgekomen, overgeslagen. Bij administratieve controles zullen in het algemeen slechts steekproeven zonder teruglegging genomen worden.

Het aantal cijfers, waaruit de aselecte getallen dienen te bestaan, wordt bepaald door de omvang van de populatie. Bij een populatie van 2500 stuks dienen de aselecte getallen uit vier cijfers te bestaan en neemt men derhalve vier kolommen van aselecte cijfers bij elkaar.

Voorbeelden

I De omvang van een steekproef zonder teruglegging is vastgesteld op 65, terwijl de omvang van de populatie 950 bedraagt. Men heeft dus aselecte getallen van drie cijfers nodig. Treft men een getal groter dan 950 aan, dan kan men dit getal gewoon overslaan. Ook een reeds eerder gebruikt getal wordt overgeslagen, wat met zich meebrengt dat men zeker meer dan 65 getallen nodig heeft.

II Bedraagt de omvang van de populatie van het vorige voorbeeld niet 950 maar 1050, dan dient men getallen van vier aselecte cijfers te gebruiken en alle getallen groter dan 1050 over te slaan. Dit heeft het bezwaar dat men slechts ongeveer één van de tien beschikbare getallen kan gebruiken. Dit is te ondervangen door bijvoorbeeld de volgende methode. Van de getallen 2001 t/m 3050 wordt 2000 afgetrokken, van die van 4001 t/m 5050 wordt 4000 afgetrokken, enz., terwijl de getallen van 1051 t/m 2000, die van 3051 t/m 4000, enz. worden overgeslagen. Ook het getal 0000 wordt bij deze methode overgeslagen.

III. De controlebladen van een debiteuren-administratie bestaan uit 316 bladzijden met op elk blad 55 posten. Uit deze postenpopulatie moet een steekproef van de omvang 200 zonder teruglegging genomen worden. De posten zijn niet genummerd. Men kan ze nummeren van 1 t/m 17.380 en daarna een steekproef nemen met aselecte getallen van vijf cijfers.

Een snellere methode gaat als volgt. Van getallen, bestaande uit vijf aselecte cijfers worden de drie meest linkse cijfers gebruikt om een blad aan te wijzen; de andere twee blijven dan beschikbaar om een post op dit blad te bepalen. Vindt men het getal 24037, dan betekent dit, dat post nummer 37 op blad nummer 240 gecontroleerd moet worden. Indien de laatste twee cijfers een getal groter dan 55 vormen, wordt het betreffende getal overgeslagen. Ook de getallen groter dan 94855 en de getallen die eindigen op 00 worden overgeslagen. Van de overige getallen wordt een zodanig veelvoud van 31600 afgetrokken dat deze kleiner of gelijk aan 31655 worden. Dus 41430 wordt 09830, hetgeen betekent dat post 30 van blad 098 gecontroleerd moet worden; 85312 wordt 22112; 74558 wordt overgeslagen.

IV. Voor populaties van kleine omvang bestaat er een eenvoudige methode om de getrokken aselecte getallen op volgorde te zetten. Stel dat de populatie de omvang 200 heeft terwijl uit deze populatie een steekproef van de omvang 50 genomen dient te worden. Op een blad worden de getallen 1 t/m 200 op volgorde uitgetypt. Daarna worden de nummers van de 50 aan te wijzen elementen op dit blad onderstreept, waardoor ze vanzelf op volgorde komen te staan.

Tabel van aselecte cijfers I

	1-4	5-8	9-12	13-16	17-20	21-24	25-28	29-32	33-36	37-40
1	3451	1659	0245	3971	7253	9632	8160	4792	9009	2720
2	2023	3613	2312	6071	9912	6555	0797	0960	5261	8814
3	7668	1033	5810	1670	5496	9481	9937	1266	9993	8834
4	8082	1098	0030	8507	6755	9250	5919	7525	7138	5426
5	4225	4173	4919	5251	5728	3783	2676	8864	7820	1567
6	2646	2874	4763	1367	0378	7182	0210	3966	9235	6225
7	6899	4526	0770	5041	7930	2991	5435	2754	8551	8124
8	8731	6532	4162	0403	2204	5176	6473	3926	2939	8171
9	0614	4415	8378	9951	2001	3027	9705	6264	3574	0600
10	3847	5659	1955	7436	5290	4997	0557	7089	4026	1191
11	8913	9260	3531	0513	6438	4352	2487	0654	8766	6747
12	1551	8878	4584	4606	6748	2182	9435	6300	6703	6710
13	7002	7426	0496	4176	1184	9833	3235	3941	3951	8400
14	5777	0706	5597	1969	5337	3739	7087	7375	2896	2575
15	5107	6868	3699	0642	2374	1556	8628	3087	4808	5348
16	8921	0365	8488	8572	2208	6378	9556	6991	6967	2143
17	2431	9271	5569	0245	1491	4847	3458	5412	0070	2282
18	0033	4541	0323	9765	4967	7563	7402	1324	9266	6921
19	5956	5608	1014	6821	0651	2566	1628	0107	8716	5142
20	5734	5402	7526	6635	5272	1995	2038	9840	3726	2036
21	7297	0322	4761	6415	9932	8971	3192	4163	2436	7675
22	0894	0760	2712	2606	6020	3787	5664	5988	4411	3806
23	0078	9734	1757	0784	1925	2756	1192	2676	9952	9752
24	8938	0716	2025	3778	5285	9075	3507	9712	3978	3747
25	5258	2843	3238	4174	8813	9058	0547	5021	9525	4201

Tabel van aselecte cijfers II

	1-4	5-8	9-12	13-16	17-20	21-24	25-28	29-32	33-36	37-40
1	0859	5165	9829	6966	9049	0724	6763	4147	0475	1790
2	8783	2219	7636	0058	4847	4120	7444	0669	7560	7431
3	7689	4065	9822	7721	2425	0315	6268	6045	6670	8539
4	0897	7559	4194	9131	9200	5229	5032	2520	9889	1938
5	3809	0726	1092	7039	8378	7663	2617	3217	4029	1233
6	3155	1702	2285	4784	9341	0311	7161	2306	2160	4509
7	3093	4460	8021	7628	3721	3180	7896	8556	5605	2728
8	3843	6241	5886	9844	7852	5175	3184	7908	0597	3593
9	9667	8695	7439	2190	8020	4232	7088	7535	1030	9999
10	2662	1926	2816	1324	0243	1665	4134	9380	8127	9829
11	5708	1537	7485	6521	7298	6320	5333	3272	5148	2660
12	9205	7755	7051	4969	5016	2670	9628	4651	4299	9078
13	8833	5035	1080	9789	8738	3120	1974	2526	1224	0349
14	3846	0113	0558	2715	8272	6221	2860	3060	5969	2203
15	2681	2478	4914	8047	7034	1186	6059	3140	0442	8323
16	8390	6231	2249	8242	1653	6029	5233	0650	4127	0941
17	2577	9510	0971	2590	6305	2987	1124	3923	3122	8642
18	4640	9989	3012	1522	8932	2754	7330	6719	3248	8283
19	4274	9544	0116	2594	4434	0667	8733	1698	8360	4163
20	8459	7214	3036	2974	1968	7688	0974	5129	9889	0204
21	2786	3294	7015	1643	1234	6871	5521	9209	3134	7940
22	9794	9204	2878	9960	8702	8851	8858	2078	2075	4582
23	3924	2944	3210	8926	8043	4583	0106	2314	4484	8530
24	5724	5305	1641	3858	9795	5145	8915	1593	8730	7449
25	0400	6069	0896	3588	8258	2422	3972	0489	3220	3717

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S 308-A9

Schatten van het brutowinstpercentage uit inkoopfacturen

Het schatten van een brutowinstpercentage kan in principe geschieden:

a) door middel van een guldenssteekproef;

b) door middel van een postensteekproef.

In het eerste geval bestaat de populatie uit alle guldens van alle inkoopfacturen uit de betreffende periode, in het tweede geval uit alle posten van alle inkoopfacturen uit die periode. In ieder geval dient iedere inkoopgulden resp. iedere inkooppost te kunnen worden opgezocht en moet men bij elk van deze guldens het gemaakte brutowinstpercentage kunnen bepalen. Verder wordt verondersteld dat alle in de betreffende periode ingekochte goederen in diezelfde periode weer worden verkocht.

a) De guldenssteekproef

Bij de guldenssteekproef wordt aselect een aantal guldens aangewezen, hetgeen geschieden kan met behulp van aselecte getallen, zoals beschreven is in Memorandum S308-A8. Men gaat vervolgens na tot welke factuur of tot welke post op een factuur de gulden behoort en bepaalt het winstpercentage v, dat op de gulden is gemaakt.

Deze percentages worden bij een steekproefomvang n voorgesteld door:

$v_1, v_2, \ldots, v_n.$

Het gemiddelde winstpercentage in de steekproef,

$$\bar{v} = \frac{1}{n}(v_1+v_2+\ldots+v_n), \tag{1}$$

is een schatting van het werkelijke winstpercentage V.

Zoals aangegeven is in Memorandum S308-A6 kan voor grote waarden van n op grond van de steekproefuitkomsten een betrouwbaarheidsinterval voor V worden bepaald. Bij een onbetrouwbaarheid α is dit interval

$$\bar{v} - \xi_{\alpha} \cdot \frac{s}{\sqrt{n}} < V < \bar{v} + \xi_{\alpha} \cdot \frac{s}{\sqrt{n}} , \tag{2}$$

waarin de standaardafwijking s gelijk is aan

$$s = \sqrt{\frac{1}{n-1}\left\{(v_1-\bar{v})^2+(v_2-\bar{v})^2+\ldots+(v_n-\bar{v})^2\right\}} , \tag{3}$$

of

$$s = \sqrt{\frac{1}{n-1}\left\{(v_1^2+v_2^2+\ldots+v_n^2)-n\,\bar{v}^2\right\}} ; \tag{4}$$

vergelijk Memorandum S308-A5. Met een onbetrouwbaarheid α ligt het werkelijke brutowinstpercentage dus in het interval, gegeven door (2).

De omvang n van de steekproef wordt berekend aan de hand van de eisen, die men stelt aan de onnauwkeurigheid van de schatting. Verlangt men dat deze onnauwkeurigheid niet groter mag zijn dan d, dan volgt n uit

$$n \geq \xi_{\alpha}^2 \cdot \frac{s^2}{d^2} ; \tag{5}$$

zie Memorandum S308-A7. In dit Memorandum wordt ook beschreven wat men moet doen wanneer s onbekend is.

Het kan voorkomen, dat de verkochte artikelen in verschillende groepen ingedeeld zijn en dat men per groep over schattingen beschikt van de gemaakte winstpercentages. Met behulp van deze gegevens kan, vóórdat de steekproef genomen wordt, een ondergrens s' voor s gevonden worden. Deze gegevens behoeven niet geheel juist te zijn (ze dienen nog door middel van de steekproef gecontroleerd te worden!), maar de onjuistheden zullen in het algemeen op de grootte van s' niet veel invloed hebben.

Stel dat er m artikelgroepen zijn, terwijl de inkoopwaarde van de i^e groep g_i guldens bedraagt en het opgegeven gemiddelde bruto-winstpercentage v'_i. Men berekent s' met

$$s'=\sqrt{\frac{1}{T-1}\left\{g_1(v'_1-\bar{v}')^2+g_2(v'_2-\bar{v}')^2+\ldots+g_m(v'_m-\bar{v}')^2\right\}}\,, \quad (6)$$

of met

$$s'=\sqrt{\frac{1}{T-1}\left\{(g_1v_1'^2+g_2v_2'^2+\ldots+g_mv_m'^2)-T\bar{v}'^2\right\}}\,, \quad (7)$$

waarin T gelijk is aan:

$$T = g_1+g_2+\ldots+g_m, \quad (8)$$

en $\bar{v}'$ gegeven wordt door:

$$\bar{v}'=\frac{1}{T}\,(g_1v'_1+g_2v'_2+\ldots+g_mv'_m). \quad (9)$$

Men weet nu, dat s in ieder geval groter zal zijn dan s' en kan hiermee bij de bepaling van n rekening houden. Bij het opstellen van het betrouwbaarheidsinterval (2) dient uiteraard de met (3) of (4) gevonden waarde van s ingevuld te worden.

b) <u>De postensteekproef.</u>

Bij de postensteekproef wordt aselect een aantal posten aangewezen. Men gaat vervolgens na welke winst w op een aangewezen post is gemaakt. Deze winsten worden bij een steekproefomvang n voorgesteld door:

$$w_1, w_2, \ldots, w_n.$$

Een schatting van de werkelijke gemiddelde winst per post, W, wordt verkregen door

$$w = \frac{1}{n}\,(w_1+w_2+\ldots+w_n). \qquad (10)$$

Voor grote waarden van n is het betrouwbaarheidsinterval voor W, bij een onbetrouwbaarheid α,

$$\overline{w} - \xi_\alpha \cdot \frac{s}{\sqrt{n}} < W < \overline{w} + \xi_\alpha \cdot \frac{s}{\sqrt{n}} \;, \qquad (11)$$

waarin de standaardafwijking s gelijk is aan

$$s=\sqrt{\frac{1}{n-1}\,\{(w_1-\overline{w})^2+(w_2-\overline{w})^2+\ldots+(w_n-\overline{w})^2\}}\;, \qquad (12)$$

of

$$s=\sqrt{\frac{1}{n-1}\,\{(w_1^2+w_2^2+\ldots+w_n^2)-n\overline{w}^2\}}\;; \qquad (13)$$

vergelijk Memorandum S308-A5.

Met een onbetrouwbaarheid α ligt de gemiddelde brutowinst per post dus in het interval, gegeven door (11).

Indien het totale aantal posten in de populatie, N, bekend is, kan een schatting en een betrouwbaarheidsinterval voor de werkelijke totale brutowinst, NW, verkregen worden. Een schatting van de totale winst is dan

$$N\overline{w} = \frac{N}{n}\,(w_1+w_2+\ldots+w_n), \qquad (14)$$

en het betrouwbaarheidsinterval voor de totale brutowinst NW, bij een onbetrouwbaarheid α, is

$$N\overline{w} - \xi_\alpha \cdot \frac{Ns}{\sqrt{n}} < NW < N\overline{w} + \xi_\alpha \cdot \frac{Ns}{\sqrt{n}} \;, \qquad (15)$$

waarin s met (12) of (13) gevonden wordt. Met een onbetrouwbaarheid α ligt de totale brutowinst dus in het interval, gegeven door (15).

Indien ook nog het totale inkoopbedrag G bekend is, wordt een schatting van het werkelijke brutowinstpercentage V verkregen door

$$100\,\frac{N\overline{w}}{G} = 100\,\frac{N}{nG}\,(w_1+w_2+\ldots+w_n). \qquad (16)$$

Het betrouwbaarheidsinterval voor V wordt gegeven door

$$100\,\frac{N\bar{w}}{G} - \xi_\alpha \cdot \frac{100\ Ns}{G\sqrt{n}} < V < 100\,\frac{N\bar{w}}{G} + \xi_\alpha \cdot \frac{100\ Ns}{G\sqrt{n}}\ , \qquad (17)$$

waarin s weer met (12) of (13) gevonden wordt. Met een onbetrouwbaarheid α ligt het brutowinstpercentage dus in het interval, gegeven door (17).

De omvang n van de steekproef kan op dezelfde wijze gevonden worden als hierboven voor de guldenssteekproef werd beschreven (zie Voorbeeld III).

Indien de verkochte artikelen in verschillende groepen ingedeeld zijn en men per groep over schattingen van de gemaakte brutowinst beschikt, kan men ook in dit geval een ondergrens s' voor s gevonden worden. De formules (6) t.m. (9) gaan nu over in:

$$s' = \sqrt{\frac{1}{N-1}\{N_1(w_1'-\bar{w}')^2+N_2(w_2'-\bar{w}')^2+\ldots+N_m(w_m'-\bar{w}')^2\}}\ , \qquad (18)$$

of

$$s' = \sqrt{\frac{1}{N-1}\{(N_1w_1'^2+N_2w_2'^2+\ldots+N_mw_m'^2)-N\bar{w}'^2\}}\ , \qquad (19)$$

waarin

N_i = aantal posten in de i-de groep,

$$N = N_1+N_2+\ldots+N_m, \qquad (20)$$

w_i' = opgegeven gemiddelde winst per post in de i-de groep,

$$\bar{w}' = \frac{1}{N}\,(N_1w_1'+N_2w_2'+\ldots+N_mw_m'). \qquad (21)$$

Ook de in de laatste alinea sub. a gemaakte opmerkingen blijven van kracht.

Vergelijking van de twee methoden

Beide typen steekproeven hebben hun voor- en nadelen. Statistisch gezien heeft een guldenssteekproef in het algemeen de voorkeur omdat de aanwezigheid van een klein aantal grote posten met een winst, die sterk afwijkt van die van de andere posten, bij een postensteekproef wel tot een vertekend beeld kan leiden, maar bij een guldenssteekproef niet. Dit bezwaar vervalt wanneer de winsten per post voor alle posten van dezelfde orde van grootte zijn(zie Memorandum S308-A11).

Het opzoeken van de in de steekproef op te nemen elementen is bij een guldenssteekproef in het algemeen lastiger dan bij een postensteekproef.

Voorbeelden.

I. Bij een handelsfirma wil men door middel van een aselecte guldenssteekproef het brutowinstpercentage bepalen. Men kiest als onbetrouwbaarheid $\alpha=1\%$ en voor de maximaal toegestane onnauwkeurigheid $d=0,5\%$. Om een indruk van de grootte van s te krijgen, begint men aselect uit de inkoopfacturen 50 guldens aan te wijzen. Uit de zo waargenomen 50 brutowinstpercentages wordt s met (3) of (4) berekend op 5,75%. Bij $\alpha=0,01$, dus $\xi_\alpha=2,576$ (tweezijdig), en $d=0,5$, vindt men met (5):

$$n \geq (2,576)^2 \cdot \frac{(5,75)^2}{(0,5)^2} = 877,6.$$

De steekproef van omvang 50 dient dus met minstens 828 trekkingen te worden uitgebreid. In deze uitgebreide steekproef vindt men een gemiddeld percentage $\bar{v}$ van 38,31 en een standaardafwijking s van 5,54%. Het betrouwbaarheidsinterval voor het werkelijke brutowinstpercentage V luidt volgens (2):

$$38,31-2,576 \cdot \frac{5,54}{\sqrt{878}} < V < 38,31+2,576 \cdot \frac{5,54}{\sqrt{878}},$$

of

$$37,83 < V < 38,79.$$

Met een onbetrouwbaarheid van 1% ligt het werkelijke brutowinstpercentage dus tussen 37,83 en 38,79. De onnauwkeurigheid is 0,48%, dus kleiner dan de geëiste 0,50%.

Opmerking: Indien in de uitgebreide steekproef een standaardafwijking s groter dan 5,75% was gevonden, dan had men opnieuw de steekproefomvang n moeten bepalen en de steekproef opnieuw moeten uitbreiden.

II. Bij een groothandel deelt men de artikelen in acht groepen in, waarbij in iedere groep op elk artikel ongeveer hetzelfde winstpercentage wordt gemaakt. Hieronder worden van iedere groep de totale inkoopbedragen en het gemiddelde brutowinstpercentage vermeld:

Groep	Totale inkoop (g_i)	Winstpercentage (v'_i)
1	ƒ 1.326.501,--	37,43
2	" 729.355,--	36,01
3	" 2.071.011,--	25,85
4	" 2.303.886,--	21,66
5	" 1.687.656,--	42,30
6	" 596.003,--	44,73
7	" 942.469,--	38,62
8	" 1.579.376,--	41,25

Men wordt gevraagd een schatting van het brutowinstpercentage te maken, waarbij er niet van mag worden uitgegaan, dat het per artikelgroep opgegeven winstpercentage juist is.

Teneinde de gevraagde schatting te kunnen maken, zal een guldenssteekproef worden genomen. Om een indruk van de standaardafwijking van het brutowinstpercentage per gulden te krijgen wordt met (6) of (7) de grootte van de ondergrens s' berekend op 8,36%. Bij een onbetrouwbaarheid α van 1% en een maximaal toegelaten onnauwkeurigheid d van 0,5% vindt men met (5):

$$n \geq (2{,}576)^2 \cdot \frac{(8{,}36)^2}{(0{,}5)^2} = 1855{,}09.$$

De steekproef moet dus een omvang van minstens 1856 hebben.

Er wordt een steekproef van 2000 aselect gekozen guldens genomen en hierin vindt men voor het gemiddelde percentage $\overline{v}$ per gulden

33,54 en voor de sta d da w jkıng 8,17%. Het betrouwbaarheidsinterva w dt nu gevonden met (2):

$$33{,}54-2{,}576 \quad \frac{8{,}17}{\sqrt{2000}} < V < 33{,}54+2{,}576 \cdot \frac{8{,}17}{\sqrt{2000}} ,$$

of

$$33{,}07 < V < 34{,}01.$$

De onnauwkeurigheid is 0,47% en dus kleiner dan de geëiste 0,50%.

III. Bij een handelsfirma wil men door middel van een aselecte postensteekproef een onderzoek naar de grootte van het brutowinstpercentage verrichten. Men kiest als onbetrouwbaarheid $\alpha=1\%$ en voor de maximaal toegelaten onnauwkeurigheid $d=2\%$. Verder is het totale aantal posten $N=14.616$ met een inkoopwaarde $G=$ ƒ 3.126.440,--.

Men begint met een voorlopige steekproef van 50 posten en bepaalt bij iedere post de brutowinst w_i. Met (12) of (13) berekent men daarna de grootte van s op 18,18.

Uiteindelijk wil men een betrouwbaarheidsinterval construeren voor het brutowinstpercentage met behulp van formule (17). De halve lengte van dit interval bedraagt

$$\xi_\alpha \cdot \frac{100\ Ns}{G\sqrt{n}} ,$$

en deze halve lengte (de onnauwkeurigheid) mag niet meer bedragen dan 2%. Daaruit volgt, dat de vereiste steekproefomvang n wordt berekend uit

$$n \geq \xi_\alpha^2 \cdot \frac{10000\ N^2 s^2}{G^2\ d^2} ,$$

dus

$$n \geq (2{,}576)^2 . \frac{10000.(14616)^2.(18{,}18)^2}{(3126440)^2\ .\ 2^2} = 119{,}8.$$

De voorlopige steekproef dient derhalve met minstens 70 trekkingen te worden uitgebreid tot n=120.

In de uitgebreide steekproef vindt men $\bar{w}=64{,}77$ en $s=18{,}01$. Een schatting voor het brutowinstpercentage is dus

$$100 \cdot \frac{\overline{\overline{Nw}}}{G} = 30{,}28,$$

en het betrouwbaarheidsinterval voor het werkelijke percentage luidt

$$30{,}28-2{,}576 \cdot \frac{100 \cdot 14616 \cdot 18{,}01}{3126440 \cdot \sqrt{120}} < V <$$

$$30{,}28+2{,}576 \cdot \frac{100 \cdot 14616 \cdot 18{,}01}{3126440 \cdot \sqrt{120}},$$

of

$$28{,}30 < V < 32{,}26.$$

De onnauwkeurigheid bedraagt hier 1,98%, dus minder dan de geëiste 2%.

WA

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S308 - A10

De binomiale verdeling

In dit Memorandum zullen we reeksen van "experimenten" beschouwen waarbij elk experiment resulteert in een "succes" of een "mislukking". Voorbeelden van zulke experimenten zijn:

a) worpen met een munt; we kunnen de uitkomst "Kruis" dan als een "succes" en de uitkomst "Munt" als een "mislukking" zien, of andersom;
b) worpen met een dobbelsteen; de uitkomst "5" is een "succes", de uitkomst "geen 5" een "mislukking";
c) trekkingen van knikkers uit een urn met witte en zwarte knikkers; het trekken van een witte knikker is een "succes" en van een zwarte knikker een "mislukking";
d) trekkingen van posten uit een lijst met posten; het trekken van een foute post is een "succes" en van een goede post een "mislukking".

Welke uitkomsten van deze experimenten we als een "succes" willen zien hangt af van het doel van ons onderzoek maar is verder volkomen willekeurig.

Vervolgens eisen we dat de experimenten van de reeks alle "onderling onafhankelijk" zijn, dat wil zeggen, dat de uitkomst van een experiment niet beïnvloed wordt door de uitkomsten van de voorafgaande experimenten en dat de kans op "succes" bij ieder experiment even groot is en gelijk aan, zeg p. Bij achtereenvolgende worpen met een munt of een dobbelsteen is in het algemeen aan deze eisen voldaan.

Bij trekkingen van knikkers uit een urn of van posten uit een boekhouding zullen we het getrokken element na ieder experiment weer moeten "terugleggen" om aan de genoemde eisen te voldoen. We trekken dan "met teruglegging". De kans op een witte knikker is bij iedere trekking $W/(W + Z)$, indien W het aantal witte en Z het aantal zwarte knikkers in de urn is, terwijl de kans op een foute post gelijk is aan de fractie foute posten in de boekhouding.

De binomiale verdeling

Stel dat er in totaal n onderling onafhankelijke experimenten uitgevoerd worden, elk met dezelfde kans p op succes en dus een kans $1 - p$ op een mislukking. Ter vereenvoudiging van de notatie schrijft men dikwijls q in plaats van $1 - p$. Het aantal successen in deze reeks is een stochastische grootheid, zeg $\underline{k}$, welke één van de waarden $0, 1, 2, \ldots, n$ aan kan nemen.

Een bepaald aantal successen kan op verscheidene manieren tot stand komen. Bij een reeks van 5 experimenten kunnen op de volgende manieren 2 successen geboekt worden:

S S M M M	M S S M M	M M S S M	M M M S S
S M S M M	M S M S M	M M S M S	
S M M S M	M S M M S		
S M M M S			

De 5 experimenten kunnen gezien worden als 5 "cellen" waarin 2 "ballen" (= successen) geplaatst moeten worden met in iedere cel hoogstens één bal. Voor de eerste bal zijn er 5 mogelijkheden; is deze geplaatst, dan is één cel bezet en zijn er voor de tweede bal nog 4 mogelijkheden. We kunnen de 2 ballen dus op $5 \cdot 4$ manieren in de 5 cellen plaatsen. Verder zijn er twee candidaten om als "eerste" bal op te treden en als deze eerste is vastgesteld nog slechts 1 candidaat om de "tweede" bal te zijn. Er zijn dus in totaal $2 \cdot 1 = 2!$ mogelijke volgorden waarin de ballen in de cellen geplaatst kunnen worden. Deze volgorden zijn echter niet van elkaar te onderscheiden zodat het boven gevonden aantal manieren om de 2 ballen in de 5 cellen te plaatsen nog door 2! gedeeld moet worden. Twee onderling niet te onderscheiden ballen kunnen derhalve op $\frac{5 \cdot 4}{2!} = \frac{5!}{2!3!} = 10$ verschillende manieren in 5 cellen geplaatst worden.

Iedere reeks experimenten met 2 successen heeft een kans p^2q^3 (zie Voorbeeld II) van Memorandum S308-A1). De reeksen kunnen niet tegelijkertijd optreden en daarom is de kans op één van deze reeksen, dus de kans op 2 successen zonder meer, gelijk aan de som van hun kansen, dus gelijk aan $10p^2q^3$ (zie Voorbeeld I) van genoemd Memorandum).

De methode, waarop hier de kans op 2 successen in 5 experimenten is afgeleid, kan ook toegepast worden om de kans op k successen in n experimenten te berekenen. Het aantal manieren waarop in n experimenten k successen op kunnen treden blijkt dan gelijk te zijn aan

$$\frac{n(n-1)(n-2)\ldots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!} ; \qquad (1)$$

hierin is $n! = n(n-1)(n-2)\cdot \ldots \cdot 3 \cdot 2 \cdot 1$ (zie de Appendix). In verkorte notatie schrijft men voor formule (1) het symbool $\binom{n}{k}$. Elk van de reeksen heeft een kans p^kq^{n-k} en de kans op k successen zonder meer is dus

$$P[\underline{k} = k] = \binom{n}{k}p^kq^{n-k}, \qquad (2)$$

voor $k = 0, 1, 2, \ldots, n$. In plaats van $P[\underline{k} = k]$ wordt ook wel $P[\underline{k} = k \mid n ; p]$ geschreven.

Omdat het zéker is, dat $\underline{k}$ één van de waarden $0, 1, 2, \ldots, n$ aan zal nemen geldt:

$$P[\underline{k} = 0] + P[\underline{k} = 1] + \ldots + P[\underline{k} = n] =$$

$$= \binom{n}{0}p^0q^{n-0} + \binom{n}{1}p^1q^{n-1} + \ldots + \binom{n}{n}p^nq^{n-n} = 1,$$

hetgeen ook op directe wijze geverifieerd kan worden (zie de Appendix).

De kansverdeling van de stochastische grootheid $\underline{k}$ is door (2) volledig bepaald en wordt de binomiale verdeling genoemd. De getallen $\binom{n}{k}$ zijn de zogenaamde "binomiaal-coëfficiënten". In de Appendix worden enkele eenvoudige eigenschappen van deze coëfficiënten afgeleid.

Verwachting

Om de verwachting van $\underline{k}$ te bepalen gaan we eerst terug naar het enkele experiment. De kans op succes is p en het aantal successen is 0 of 1. De verwachting van het aantal successen per experiment is dus (zie Memorandum S308 - A2):

$$0 \cdot q + 1 \cdot p = p.$$

Bij 2 experimenten kunnen 0, 1 of 2 successen optreden met respectievelijk de kansen q^2, $pq + qp = 2pq$ en p^2. De verwachting van het aantal successen is dan

$$0 \cdot q^2 + 1 \cdot 2pq + 2 \cdot p^2 = 2pq + 2p^2 =$$
$$= 2p(q + p) = 2p.$$

Op dezelfde wijze berekent men de verwachting van het aantal successen bij 3 experimenten op 3p en in het algemeen de verwachting van het aantal successen bij n experimenten op np. In formulevorm geldt dus

$$\mu(\underline{k}) = np. \qquad (4)$$

Variantie

Ook hier gaan we eerst terug naar het enkele experiment. De variantie van het aantal successen in één experiment is (zie Memorandum S308 - A2):

$$(0 - p)^2 \cdot q + (1 - p)^2 \cdot p = p^2q + q^2p =$$
$$= pq(p + q) = pq.$$

De variantie van het aantal successen in twee experimenten is

$$(0 - 2p)^2 \cdot q^2 + (1 - 2p)^2 \cdot 2pq + (2 - 2p)^2 \cdot p^2 = 2pq.$$

Op dezelfde wijze berekent men de variantie van het aantal successen bij 3 experimenten op 3pq en die van het aantal successen in n experimenten op npq. Er geldt dus

$$\sigma^2(\underline{k}) = npq. \qquad (5)$$

De standaardafwijking $\sigma(\underline{k})$ van $\underline{k}$ is gelijk aan $\sqrt{npq}$.

Tabellen

De binomiale verdeling is voor lage waarden van n uitvoerig getabelleerd, o.a. in de volgende tabellen:

a) Tables of the Binomial Probability Distribution , National Bureau of Standards, Applied Mathematics, Series-6, Washington 1950; $P[\underline{k} = k]$ en $P[\underline{k} \geq k]$ voor n = 2(1)49 en p = 0,01(0,01)0,50;

b) H.G. Romig: 50 - 100 Binomial Tables, John Wiley and Sons, Inc., New York 1953; $P[\underline{k} = k]$ en $P[\underline{k} \leq k]$ voor n = 50(5)100 en p = 0,01(0,01)0,50;

c) Tables of the Cumulative Binomial Probability Distribution, Harvard University Press, 1955; $P[\underline{k} \geq k]$ voor n = 1(1)50(2)100(10)200(20)500 (50)1000, p = 0,01(0,01)0,50, $p = \frac{1}{12}(\frac{1}{12})\frac{5}{12}$ en $p = \frac{1}{16}(\frac{1}{16})\frac{7}{16}$;

d) S. Weintraub: Tables of the Cumulative Binomial Probability Distribution for small values of p, The Free Press of Glencoe, Collier-Macmillan Limited, London 1963; $P[\underline{k} \geq k]$ voor n = 1(1)100 en p = 0,00001, 0,0001(0,0001)0,001(0,001)0,1.

In de tabellen van de binomiale verdeling is p veelal niet groter dan 0,5. Echter, de kans op k successen is gelijk aan de kans op n - k mislukkingen bij n experimenten, elk met een kans q = 1 - p op mislukking. De kansverdeling van het aantal mislukkingen is dus ook binomiaal en met de volgende, gemakkelijk af te leiden, betrekkingen kunnen we voor p > 0,5 toch met de tabellen werken:

$$P[\underline{k} = k \mid n \,;\, p] = P[\underline{k} = n - k \mid n \,;\, 1 - p], \tag{6}$$

$$P[\underline{k} \leq k \mid n \,;\, p] = P[\underline{k} \geq n - k \mid n \,;\, 1 - p], \tag{7}$$

$$P[\underline{k} \geq k \mid n \,;\, p] = P[\underline{k} \leq n - k \mid n \,;\, 1 - p]. \tag{8}$$

De normale benadering

Voor niet in tabellen van de binomiale verdeling vermelde waarden van n en/of p moet men de gezochte kansen zelf berekenen of trachten, ze met een goede benadering te schatten. De belangrijkste benaderingen zijn in dit verband de zogenaamde "normale benadering" en de "Poisson-benadering". De eerste is vooral van belang wanneer n groot is en p niet te dicht bij 0 of 1 is gelegen, de tweede voor kleine of grote p en grote n. De Poisson-benadering is gebruikt bij het opstellen van Tabel II in Memorandum S308 - A4. In dit Memorandum beperken wij ons verder tot de normale benadering. Voor de Poisson-benadering wordt verwezen naar Memorandum S308 - A19.

Men kan bewijzen, dat voor grote waarden van n het aantal successen bij benadering normaal verdeeld is met verwachting np en standaardafwijking $\sqrt{npq}$, dus dat $\underline{k}$ bij benadering een $N(np\,;\sqrt{npq})$-verdeling bezit (zie voor de notatie Memorandum S308 - A3). Om het berekenen van kansen met behulp van deze benadering toe te kunnen lichten, voeren wij de grootheid $\underline{k}'$ in, die exact een $N(np\,;\sqrt{npq})$ - verdeling bezit. Volgens Memorandum S308 - A3 is

$$\underline{u} = \frac{\underline{k}' - np}{\sqrt{npq}} \tag{9}$$

dan N(0;1) - verdeeld, welke verdeling is getabelleerd in de in genoemd Memorandum opgenomen tabel.

Hoe berekenen we nu $P[\underline{k} = k]$? Men zou als volgt kunnen redeneren ("$\approx$" betekent "is ongeveer gelijk aan"):

$$P\left[\underline{k} = k\right] \approx P\left[\underline{k}' = k\right] = P\left[\underline{u} = \frac{k - np}{\sqrt{npq}}\right] = 0,$$

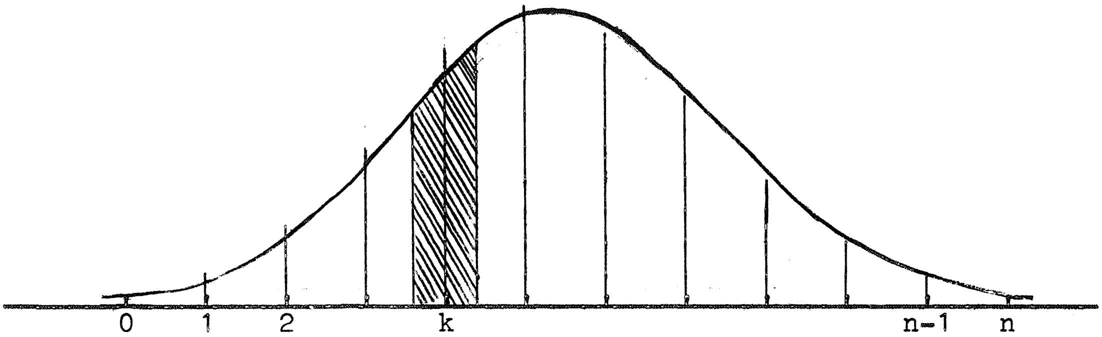
voor $k = 0, 1, 2, \ldots, n$. Het is duidelijk, dat de normale benadering op deze wijze toegepast ons niet veel helpt. De uitkomst 0 wordt veroorzaakt doordat een discrete verdeling (de binomiale) verkeerd benaderd wordt met een continue verdeling (de normale). Figuur 1 geeft aan hoe de moeilijkheid opgelost kan worden. Als benadering van $P[\underline{k} = k]$ nemen we daartoe de oppervlakte onder de normale kromme tussen $k - \frac{1}{2}$ en $k + \frac{1}{2}$. Deze benadering kan toegepast worden voor $k = 1, 2, \ldots, n - 1$. Voor $P[\underline{k} = 0]$ nemen we echter de oppervlakte tussen $-\infty$ en $+\frac{1}{2}$ en voor $P[\underline{k} = n]$ de oppervlakte tussen $n-\frac{1}{2}$ en $+\infty$. In formule:

$$P[\underline{k} = k] \approx P[k - \tfrac{1}{2} < \underline{k}' \leq k + \tfrac{1}{2}] =$$

$$= P\left[\frac{k - np - \frac{1}{2}}{\sqrt{npq}} < \underline{u} \leq \frac{k - np + \frac{1}{2}}{\sqrt{npq}}\right] \quad \text{voor } k = 1, \ldots, n - 1, \tag{10}$$

$$P\left[\underline{k} = 0\right] \approx P\left[-\infty < \underline{k}' \leq + \tfrac{1}{2}\right] = P\left[\underline{u} \leq \frac{-np + \frac{1}{2}}{\sqrt{npq}}\right], \tag{11}$$

$$P\left[\underline{k} = n\right] \approx P\left[n - \tfrac{1}{2} < \underline{k}' < +\infty\right] = P\left[\underline{u} > \frac{n - np - \frac{1}{2}}{\sqrt{npq}}\right]. \tag{12}$$



Figuur 1

Benadering van een binomiale door een normale verdeling

Het is nu ook duidelijk hoe de kansen $P[\underline{k} \leq k]$ en $P[\underline{k} \geq k]$ geschat worden. Met behulp van figuur 1 zien we, dat

$$P[\underline{k} \leq k] = P[\underline{k} = 0] + P[\underline{k} = 1] + \ldots + P[\underline{k} = k] \approx$$

$$\approx P[-\infty < \underline{k}' \leq \tfrac{1}{2}] + P[\tfrac{1}{2} < \underline{k}' \leq 1\tfrac{1}{2}] + \ldots$$

$$+ P[k - \tfrac{1}{2} < \underline{k}' \leq k + \tfrac{1}{2}] =$$

$$= P\left[-\infty < \underline{k}' \leq k + \tfrac{1}{2}\right] = P\left[\underline{u} \leq \frac{k - np + \frac{1}{2}}{\sqrt{npq}}\right], \tag{13}$$

indien $k = 0, 1, 2, \ldots, n - 1$, en op dezelfde wijze, dat

$$P\left[\underline{k} \geq k\right] \approx P\left[\underline{k}' \geq k - \tfrac{1}{2}\right] = P\left[\underline{u} \geq \frac{k - np - \frac{1}{2}}{\sqrt{npq}}\right], \qquad (14)$$

indien $k = 1, 2, \ldots, n$.

De waarden $+\frac{1}{2}$ en $-\frac{1}{2}$ worden continuïteitscorrecties genoemd omdat ze toegepast worden teneinde de benadering van de discrete binomiale verdeling door de continue normale verdeling te verbeteren.

Fracties successen; schatten van p

In plaats van het aantal successen $\underline{k}$ in n experimenten kan men de fractie successen $\underline{f}$, gegeven door

$$\underline{f} = \frac{\underline{k}}{n}, \qquad (15)$$

beschouwen. De stochastische grootheid $\underline{f}$ kan één van de waarden $0, \frac{1}{n}, \frac{2}{n}, \ldots, 1$ aannemen. Verder geldt voor de kans, dat $\underline{f}$ één van deze waarden aanneemt:

$$P\left[\underline{f} = \frac{k}{n}\right] = P\left[\underline{k} = k\right] = \binom{n}{k} p^k q^{n-k}. \qquad (16)$$

Op dezelfde wijze als voor $\underline{k}$ is geschied kan worden aangetoond, dat voor de verwachting $\mu(\underline{f})$ en de variantie $\sigma^2(\underline{f})$ van $\underline{f}$ geldt:

$$\mu(\underline{f}) = p, \; \sigma^2(\underline{f}) = \frac{pq}{n}. \qquad (17)$$

Omdat $\underline{k}$, het aantal successen, bij benadering normaal verdeeld is, is $\underline{f}$, de fractie successen, ook bij benadering normaal verdeeld en wel met verwachting p en standaardafwijking $\sqrt{\frac{pq}{n}}$. Het zal duidelijk zijn, dat we nu als continuïteitscorrecties $+\frac{1}{2n}$ en $-\frac{1}{2n}$ moeten nemen in plaats van $+\frac{1}{2}$ en $-\frac{1}{2}$. Voor een toepassing van deze correcties zie men de formules (1) en (2) van Memorandum S308-A12.

Indien de grootte van p onbekend is kan men deze schatten met de fractie successen $\underline{f}$, welke grootheid daarom een "schatter" van p genoemd wordt. Omdat de verwachting van $\underline{f}$ gelijk is aan p wordt $\underline{f}$ een "zuivere schatter" van p genoemd. (Voor de begrippen "schatter" en "zuiverheid" zie men Memorandum S308 - A11.) Met behulp van de schatter $\underline{f}$ kan een betrouwbaarheidsinterval voor p geconstrueerd worden (zie Memorandum S308 - A4).

Voorbeelden.

I. Voor n = 100 en p = 0,25 wordt gevraagd de kans op 20 of minder successen, dus $P[\underline{k} \leq 20]$, te berekenen. In de tabel, genoemd onder b), vinden we voor de exacte waarde in drie decimalen nauwkeurig: 0,149.

Voor de normale benadering berekenen we eerst de verwachting en de standaardafwijking van $\underline{k}$: $\mu(\underline{k}) = np = 25$, $\sigma(\underline{k}) = \sqrt{npq} = 4{,}3301$. Zonder continuïteitscorrectie vinden we:

$$P[\underline{k} \leq 20] \approx P[\underline{k}' \leq 20] = P[\underline{u} \leq \frac{20 - 25}{4{,}3301}] =$$

$$= P[\underline{u} \leq -1{,}155] = P[\underline{u} \geq +1{,}155] = 0{,}124.$$

Met continuïteitscorrectie vinden we:

$$P[\underline{k} \leq 20] \approx P[\underline{k}' \leq 20{,}5] = P[\underline{u} \leq \frac{20{,}5 - 25}{4{,}3301}] =$$

$$= P[\underline{u} \leq -1{,}032] = P[\underline{u} \geq +1{,}032] = 0{,}151.$$

De benadering met continuïteitscorrectie leidt dus tot een nauwkeuriger uitkomst.

II. Bij 45 worpen met een zuivere dobbelsteen wordt gevraagd naar de kans, dat minstens 3 en hoogstens 6 maal een "6" wordt geworpen.

In dit voorbeeld is n = 45 en $p = \frac{1}{6}$, dus np = 7,5 en $\sqrt{npq} = 2{,}5$. De berekening met de normale benadering gaat als volgt:

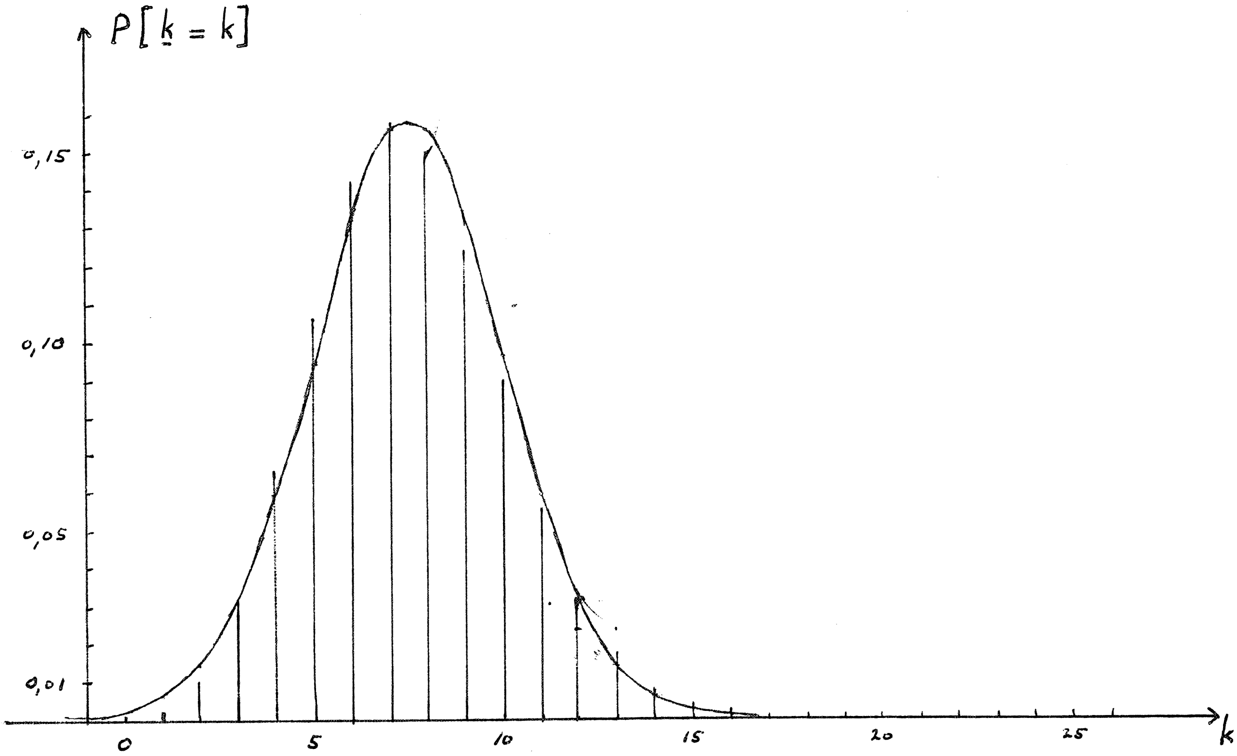
$$P[3 \leq \underline{k} \leq 6] \approx P[2{,}5 < \underline{k}' \leq 6{,}5] =$$

$$= P[\frac{2{,}5 - 7{,}5}{2{,}5} < \underline{u} \leq \frac{6{,}5 - 7{,}5}{2{,}5}] = P[-2{,}0 < \underline{u} \leq -0{,}4] =$$

$$= P[\underline{u} \geq +0{,}4] - P[\underline{u} \geq +2{,}0] =$$

$$= 0{,}4840 - 0{,}0228 = 0{,}4612.$$

De exacte waarde is 0,3457. De benadering is hier niet zo goed omdat de benaderde verdeling nogal "scheef" en n niet groot is. In figuur 2 wordt deze scheve kansverdeling, alsmede de daaraan aangepaste normale verdeling met verwachting 7,5 en standaardafwijking 2,5, in beeld gebracht.



Figuur 2

De binomiale verdeling voor n = 45 en p = $\frac{1}{6}$ en

de aangepaste N(7,5;2,5)-verdeling

III. In een lijst met posten is de fractie foute guldens gelijk aan 0,01. Het totaalbedrag van de lijst is zeer groot, namelijk enkele millioenen guldens. Men vraagt naar de kans, dat in een aselecte steekproef van 100 guldens minstens één foute gulden gevonden wordt.

De guldens worden weliswaar zonder teruglegging getrokken, maar het niet-terugleggen verandert zo weinig aan de populatie, dat we de steekproef als één mét teruglegging mogen beschouwen zonder grote fouten te maken.

Passen wij de normale benadering toe, dan vinden we met $n = 100$, $p = 0{,}01$, $np = 1$, $\sqrt{npq} = 0{,}99499$:

$$P[\underline{k} \geq 1] \approx P[\underline{k}' \geq \tfrac{1}{2}] =$$

$$= P\left[\underline{u} \geq \frac{\frac{1}{2} - 1}{0{,}99499}\right] = P[\underline{u} \geq -0{,}503] =$$

$$= 1 - P[\underline{u} > +0{,}503] = 1 - 0{,}3074 = 0{,}6926.$$

De exacte waarde bedraagt 0,6340. De normale benadering is hier dus niet erg bevredigend, ondanks het toepassen van een continuïteitscorrectie. Met de zogenaamde Poisson-benadering wordt voor $P[\underline{k} \geq 1]$ de waarde 0,6321 gevonden. Zoals te verwachten was geeft deze benadering bij $p = 0{,}01$ betere uitkomsten dan de normale benadering.

Appendix.

I. Het aantal manieren waarop in n experimenten k successen op kunnen treden kan als volgt gevonden worden.
De n experimenten kunnen weer worden gezien als n "cellen" waarin k "ballen" ($0 \leq \underline{k} \leq n$) geplaatst moeten worden met in iedere cel hoogstens één bal. Voor de eerste bal zijn er n mogelijkheden. Voor de tweede bal zijn er nog slechts $n - 1$ mogelijkheden, voor de derde bal $n - 2$ mogelijkheden, terwijl er tenslotte voor de k-de bal $n - k + 1$ mogelijkheden overblijven. We kunnen de k ballen dus op

$$n(n-1)(n-2)\ldots(n-k+1) = \frac{n!}{(n-k)!}$$

manieren in de n cellen plaatsen.

Wij spraken hierboven over de "eerste" bal, de "tweede" bal, enzovoort. We hebben echter k manieren om een "eerste" bal te kiezen, $k - 1$ manieren om een "tweede" bal te kiezen, dus

$$k(k-1)(k-2)\ldots 3 \cdot 2 \cdot 1 = k!$$

mogelijke volgorden waarin de ballen in de cellen geplaatst kunnen worden. Deze volgorden zijn echter niet van elkaar te onderscheiden zodat het boven gevonden aantal manieren om de k ballen in de n cellen te plaatsen nog door k! gedeeld moet worden. Dus kunnen k onderling niet te onderscheiden ballen op

$$\frac{n(n-1)\ldots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!} = \binom{n}{k} \qquad (1)$$

manieren in de cellen geplaatst worden, waarmede formule (1) gevonden is.

II. De kans op 0 successen is gelijk aan de kans op n mislukkingen en dus gelijk aan q^n. Anderzijds volgt uit formule (2)

$$P[\underline{k} = 0] = \binom{n}{0}p^0q^{n-0} = \binom{n}{0}q^n .$$

Dit kan alleen juist zijn wanneer geldt:

$$\binom{n}{0} = 1.$$

Toepassing van formule (1) leert ons verder

$$\binom{n}{0} = \frac{n!}{0!(n-0)!} = \frac{1}{0!},$$

hetgeen slechts klopt voor $0! = 1$. Daarom definiëren wij, of, anders gezegd, spreken wij af, dat de uitdrukking $0!$ gelijk is aan 1.

Verder zijn de volgende relaties gemakkelijk na te gaan:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!k!} = \binom{n}{n-k};$$

$$\binom{n}{1} = \frac{n!}{1!(n-1)!} = \frac{n\,.\,(n-1)!}{(n-1)} = n;$$

$$\binom{n}{n} = \binom{n}{n-n} = \binom{n}{0} = 1;$$

$$\binom{n}{n-1} = \binom{n}{n-n+1} = \binom{n}{1} = n.$$

III. Voor elk tweetal getallen x en y geldt:

$$(x+y)^2 = x^2 + 2xy + y^2 = \binom{2}{0}x^2 + \binom{2}{1}xy + \binom{2}{2}y^2.$$

Zo is ook

$$(x + y)^3 = x^3 + 3x^2y + 3xy^2 + y^3 =$$

$$= \binom{3}{0}x^3 + \binom{3}{1}x^2y + \binom{3}{2}xy^2 + \binom{3}{3}y^3.$$

In het algemeen kan men aantonen, dat voldaan is aan de betrekking

$$(x + y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \ldots$$

$$\ldots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n. \qquad (18)$$

Voor x en y mogen we ook q en p invullen:

$$(q + p)^n = \binom{n}{0}p^0q^n + \binom{n}{1}p^1q^{n-1} + \ldots + \binom{n}{n}p^nq^0.$$

Omdat $p + q = 1$ volgt hieruit direct formule (3).

De relatie (18) noemt men wel het "Binomium van Newton", waarvan de naam "binomiale verdeling" werd afgeleid.

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

Enkele eigenschappen van schatters

S308-A11

In dit Memorandum worden enkele schatters en enige van hun eigenschappen behandeld.

De te bespreken schatters kunnen onder andere gebruikt worden bij het schatten van winstpercentages (zie ook Memorandum S308-A9) en ook in het geval van zogenaamde "positieve controles". Hierbij wordt onder een positieve controle een controle verstaan waarbij men weet of veronderstelt, dat de posten wel te hoog, doch niet te laag geboekt kunnen zijn, zoals bijvoorbeeld bij de controle van onkostendeclaraties of van schadeuitkeringen.

Stel dat het totale aantal posten in de populatie N bedraagt en dat het totaalbedrag T gulden is. Gevraagd wordt een schatting van het te hoog opgevoerde bedrag V door middel van een aselecte steekproef van de omvang n.

Allereerst komt een guldenssteekproef in aanmerking. Men wijst aselect n guldens aan en gaat hiervan na of ze gedekt zijn of niet. Noemen wij het aantal ongedekte guldens in de steekproef $\underline{k}$, dan is een "schatter" van het te hoog opgevoerde bedrag V

$$T \cdot \frac{\underline{k}}{n} . \qquad (1)$$

Verder is een postensteekproef mogelijk. Er worden n posten aangewezen en van elk van deze posten wordt het niet gedekte bedrag bepaald. Bedraagt de i-de post $\underline{x}_i$ guldens en bevat deze post $\underline{v}_i$ niet gedekte guldens, dan is een tweede schatter van het te hoog opgevoerde bedrag V

$$N \cdot \frac{1}{n}\,(\underline{v}_1+\underline{v}_2+\ldots+\underline{v}_n). \qquad (2)$$

Bij deze postensteekproef zijn nog twee andere schatters van het te hoog opgevoerde bedrag V mogelijk. Men kan de fractie ongedekte guldens in de gehele steekproef berekenen en deze fractie vermenigvuldigen met T:

$$T \cdot \frac{\underline{v}_1+\underline{v}_2+\ldots+\underline{v}_n}{\underline{x}_1+\underline{x}_2+\ldots+\underline{x}_n}, \qquad (3)$$

en men kan voor iedere post afzonderlijk de fractie ongedekte guldens bepalen, van al deze fracties het gemiddelde berekenen en dit met T vermenigvuldigen:

$$T \cdot \frac{1}{n}\left(\frac{\underline{v}_1}{\underline{x}_1}+\frac{\underline{v}_2}{\underline{x}_2}+\ldots+\frac{\underline{v}_n}{\underline{x}_n}\right). \qquad (4)$$

Hierboven werden de symbolen k, v en x onderstreept omdat het stochastische grootheden zijn (zie Memorandum S308-A2). De uitdrukkingen (1) t/m. (4) bevatten één of meer van deze grootheden en zijn dus zelf ook stochastische grootheden met een kansverdeling. Aangezien men deze uitdrukkingen kan gebruiken om er schattingen van de werkelijke waarde V mee te maken, noemt men ze <u>schatters</u> van V. Bij iedere steekproef neemt een schatter een zekere waarde aan, welke een schatting van V genoemd wordt.

Hierboven werden vier schatters van V vermeld. De vraag is nu welke schatter de voorkeur verdient.

Een eerste eis die men aan schatter kan stellen luidt: Wordt de steekproef uitgebreid tot ze de gehele populatie omvat (dus indien n=T bij de guldenssteekproef en n=N bij de postensteekproef), dan dient de waarde van de schatter gelijk te zijn aan de te schatten waarde.

Een schatter die aan de bovenvermelde eis voldoet noemt men <u>asymptotisch raak</u> (Engels: consistent) [1].

---

[1] Deze definitie is niet geheel is overeenstemming met die, gegeven in Normaalblad NEN 3117 van het Nederlands Normalisatie-Instituut.

Het is makkelijk in te zien dat de schatters (1), (2) en (3) asymptotisch raak zijn. Met een voorbeeld tonen wij aan dat de schatter (4) niet asymptotisch raak behoeft te zijn.

Laat een populatie bestaan uit vijf posten ten grootte van ƒ 100,-, ƒ 50,-, ƒ 60,-, ƒ 85,- en ƒ 70,-, met ongedekte bedragen van respectievelijk ƒ 37,50, ƒ -,-, ƒ 4,-. ƒ 8,50 en ƒ 10,50. Voor deze populatie is T = 365, N = 5 en V = 60,50. Kiezen wij n = 5, dan vinden wij met (4) als schatting voor V:

$$365 \cdot \frac{1}{5}\left(\frac{37,50}{100} + \frac{0}{50} + \frac{4}{60} + \frac{8,50}{85} + \frac{10,50}{70}\right) = 50,49.$$

Hierboven werd opgemerkt dat een schatter een stochastische grootheid is en dus een kansverdeling heeft. Deze kansverdeling bezit dan weer een verwachting, een standaardafwijking en een variantie (zie Memorandum S308-A2). De schatter zal bij verschillende steekproeven verschillende waarden aannemen, welke waarden met een zekere spreiding in de omgeving van de verwachting zullen liggen. Men kan nu eisen, dat de verwachting van een schatter gelijk is aan de te schatten grootheid, in dit geval V, of anders gezegd, dat de schatter gemiddeld juist is. Een schatter die aan deze eis voldoet noemt men <u>zuiver</u> (Engels: unbiased).

De schatters (1) en (2) zijn zuiver, de schatters (3) en (4) in het algemeen niet. In de Appendix wordt in punt II aangetoond, dat de schatters (3) en (4) onzuiver kunnen zijn.

De variantie van de kansverdeling van de schatter is een maat voor de spreiding, welke de schattingen in verschillende steekproeven zullen aannemen. Hoe kleiner de variantie, hoe groter de kans dat de waarde van een zuivere schatter in een steekproef dicht bij zijn verwachting ligt. Een kleine variantie is daarom een belangrijke eigenschap van een schatter.

De varianties van de schatters (1) en (2) luiden als volgt:

$$\text{Variantie van (1)} = \frac{T-n}{T-1} \cdot \frac{V(T-V)}{n}\ ; \qquad (5)$$

$$\text{Variantie van (2)} = \frac{N-n}{N-1} \cdot \frac{N}{n} \left\{(v_1^2+v_2^2+\ldots+v_N^2) - \frac{V^2}{N}\right\}. \qquad (6)$$

Om deze varianties te kunnen berekenen heeft men de waarde van V en alle waarden v_i van de gehele populatie nodig, welke waarden echter onbekend zijn. Deze varianties dienen derhalve ook weer geschat te worden, hetgeen kan geschieden met de schatters (7) en (8):

$$\text{Schatter van (5)} = \frac{(T-n)T}{n-1} \cdot \frac{\underline{k}(n-\underline{k})}{n^2} ; \qquad (7)$$

$$\text{Schatter van (6)} = \frac{N(N-n)}{n(n-1)} \left\{ (\underline{v}_1^2+\underline{v}_2^2+\ldots+\underline{v}_n^2) + \right.$$

$$\left. - \frac{1}{n} (\underline{v}_1+\underline{v}_2+\ldots+\underline{v}_n)^2 \right\} . \qquad (8)$$

De schatters (7) en (8) zijn zuivere schatters van de varanties van de schatters (1) en (2). Schatters van de standaardafwijkingen van (1) en (2) zijn de wortels uit (7) en (8).

De varianties van de schatters (3) en (4) worden hier niet vermeld.

De vraag, welke van de vier besproken schatters de kleinste variantie heeft, kan niet algemeen beantwoord worden. Eén en ander is namelijk in sterke mate afhankelijk van de eigenschappen van de te onderzoeken populatie.

Conclusie

De schatters (1) en (2) verdienen meestal de voorkeur boven de schatters (3) en (4), omdat de eerstgenoemden altijd asymptotischraak en zuiver zijn. Schatter (4) is nooit aan te raden, maar er zijn situaties, waarin schatter (3) zuiver of nagenoeg zuiver is en waarin hij een kleinere variantie bezit dan de schatters (1) en (2). Een behandeling van deze situaties wordt uitgesteld tot een later Memorandum.

De keuze tussen de schatters (1) en (2) is moeilijker. Om deze keus te kunnen maken zou men na elkaar een guldenssteekproef en een postensteekproef moeten nemen en schattingen van de varianties van (1) en (2) moeten maken (zie de formules (7) en (8)). De schatter met de kleinste variantieschatting verdient dan statistisch gezien de voorkeur om de bovenvermelde redenen. Een verdere overweging bij de keuze zal zijn, dat het aselect kiezen van posten in het algemeen eenvoudiger is dan het uitzoeken van aangewezen guldens.

Appendix

I. De schatters (3) en (4) heten "breukschatters" (Eng.: ratio-estimates) omdat bij deze schatters zowel in de noemer als in de teller stochastische grootheden voorkomen. Ze zijn beide in de meeste situaties onzuiver. Een zuivere en asymptotisch rake breukschatter wordt gegeven door

$$\frac{T}{n}\left(\frac{\underline{v}_1}{\underline{x}_1}+\frac{\underline{v}_2}{\underline{x}_2}+\ldots+\frac{\underline{v}_n}{\underline{x}_n}\right)+$$

$$+\frac{N-1}{n-1}\left\{(\underline{v}_1+\underline{v}_2+\ldots+\underline{v}_n)-\frac{1}{n}(\underline{x}_1+\underline{x}_2+\ldots+\underline{x}_n)\left(\frac{\underline{v}_1}{\underline{x}_1}+\frac{\underline{v}_2}{\underline{x}_2}+\ldots+\frac{\underline{v}_n}{\underline{x}_n}\right)\right\}. \qquad (9)$$

II. Bij een handelsfirma komen in het inkoopboek 10.000 posten à ƒ 1,- en één post à ƒ 10.000,- voor. Op de posten à ƒ 1,- werd 20% en op de post à ƒ 10.000,- werd 10% brutowinst gemaakt. De werkelijk gemaakte brutowinst V bedraagt dus ƒ 3.000,-. Verder geldt T = 20.000 en N = 10.001.

Om de brutowinst V te schatten neemt men een postensteekproef van de omvang 100 (n=100). In deze steekproef bevindt zich al of niet de post à ƒ 10.000,-. Men vindt dus als schattingen voor V:

met (2):	ƒ 101.990,20	of	ƒ 2.000,20;
met (3):	" 2.019,61	of	" 4.000,- ;
met (4):	" 3.980,-	of	" 4.000,- ;
met (9):	-" 96.010,-	of	" 4.000,- .

In totaal zijn er $\binom{10001}{100}$ steekproeven van de omvang 100 uit 10001 posten mogelijk. Hieronder bevinden zich $\binom{10000}{99}$ steekproeven waarin de post à ƒ 10.000,- voorkomt, en $\binom{10000}{100}$ steekproeven waarin deze post niet voorkomt. Ter verduidelijking diene dat

$$\binom{a}{b}=\frac{a!}{b!(a-b)!}=\frac{a(a-1)(a-2)\ldots(a-b+1)}{b!} \qquad (10)$$

en

$$a! = a.(a-1).(a-2).----.\ 3.2.1. \qquad (11)$$

Indien een schatter zuiver is moet het gemiddelde van alle schattingen in alle mogelijke steekproeven van gelijke omvang gelijk zijn aan de te schatten grootheid, in dit geval dus gelijk aan V = ƒ 3.000,-. Dit gemiddelde is bij (2):

$$\left\{\binom{10000}{99}(101990{,}20)+\binom{10000}{100}(2000{,}20)\right\} : \binom{10001}{100} =$$

$$=\left\{(100)(101990{,}20)+(9901)(2000{,}20)\right\} : (10001) = 3.000,-.$$

Op dezelfde wijze vindt men bij (3): 3.980,20,
bij (4): 3.999,80,
en bij (9): 3.000,- .

Het gemiddelde van alle schattingen volgens (2) en (9) is dus gelijk aan de brutowinst V. Bij de schatters (3) en (4) is dit niet het geval. Dit bevestigt dus de beweringen omtrent de zuiverheid van de schatters (2) en (9) en de onzuiverheid van de schatters (3) en (4).

Voor de schatter (2) kan met formule (6) de variantie en dus ook de standaardafwijking berekend worden. In dit voorbeeld kennen wij echter de populatie volledig en kunnen wij dus direct voor alle schatters de definitie van de variantie toepassen (zie Memorandum S308-A2). De variantie van de schatter is gelijk aan het gemiddelde over alle mogelijke steekproeven van het kwadraat van het verschil tussen de steekproefuitkomst en de verwachting. Voor schatter (3) vindt men zo

$$\frac{\binom{10000}{99}(2019{,}61 - 3980{,}20)^2 + \binom{10000}{100}(4000{,}00 - 3980{,}20)^2}{\binom{10001}{100}} =$$

$$= \frac{(100)(1960{,}59)^2 + (9901)(19{,}80)^2}{10001} = 38.823{,}41.$$

Op dezelfde wijze vindt men voor de variantie van

schatter (2): 98.970.403,92,
schatter (4): 3,96,
en schatter (9): 99.010.000,00.

De standaardafwijkingen luiden voor

schatter (2): 9.948,39,
schatter (3): 197,04,
schatter (4): 1,99,
schatter (9): 9.950,38.

Uit deze uitkomsten blijkt dat de schatters (2) en (9) inderdaad zuiver zijn; hun varianties zijn echter groot. De schatters (3) en (4) zijn beide onzuiver, terwijl (4) bovendien niet asymptotisch raak is; de varianties van beide schatters zijn zeer klein.

In situaties, welke lijken op dit, extreem gekozen, voorbeeld verdient een guldenssteekproef zonder meer de voorkeur.

W

A

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S 308 - A 12

Het toetsen van hypothesen omtrent kansen en verwachtingen

In veel steekproefproblemen wordt niet gevraagd een onbekende fractie of een onbekende verwachting te schatten (zie hiervoor de Memoranda S 308 - A 4, A 5 en A 6), doch is het juist vereist een bewering omtrent zo'n fractie of die verwachting aan de hand van de steekproefresultaten op zijn juistheid te onderzoeken.

De bovenbedoelde bewering noemt men de nulhypothese, H_0. Het vergelijken van de nulhypothese H_0 met de steekproefresultaten noemt men toetsen en de daarbij gebruikte procedure de toets.

De eenvoudigste toetsen leiden ófwel tot het resultaat, dat de nulhypothese verworpen wordt, ófwel tot de conclusie, dat deze hypothese niet verworpen kan worden.

Wordt er bij het toetsen voldoende argument tegen H_0 gevonden, dan aanvaardt men een alternatieve hypothese, H_1. Het is niet noodzakelijk dat de alternatieve hypothese slechts één mogelijkheid omvat.

Hierbij dient te worden opgemerkt dat "niet verwerpen" niet betekent "aanvaarden". "Aanvaarden" is sterker dan "niet verwerpen". Veel toetsen worden dan ook uitgevoerd in de hoop dat men de nulhypothese H_0 kan verwerpen en een alternatieve hypothese H_1 kan aanvaarden.

Het eventueel bij het toetsen gevonden argument tegen H_0 is nooit geheel steekhoudend, aangezien bij de boven genoemde toetsen ook onder H_0 alle steekproefresultaten mogelijk zijn; deze zijn echter niet alle even waarschijnlijk. Vinden we dus een steekproefresultaat dat onder H_0 zeer onwaarschijnlijk is, dan verwerpen we H_0 ten gunste van H_1, echter met het voorbehoud, dat we een kleine kans maken een foute beslissing te nemen. Deze kans α op een foute beslissing is de onbetrouwbaarheid α van de toets. De fout, H_0 te verwerpen terwijl H_0 juist is, wordt de fout van de eerste soort genoemd.

Anderzijds kan het voorkomen, dat het steekproefresultaat onder H_0 niet onwaarschijnlijk is, terwijl toch H_1 de juiste hypothese is. Dan wordt H_0 dus ten onrechte niet verworpen. De fout, H_0 niet te verwerpen terwijl H_1 juist is, wordt de fout van de tweede soort genoemd en de kans op die fout wordt aangeduid met β.

Er zijn dus vier verschillende situaties mogelijk, welke in onderstaande tabel schematisch zijn weergegeven.

Tabel I

Mogelijke situaties bij het toetsen van een hypothese

	H_0 juist	H_0 niet juist
H_0 niet verwerpen	terecht	fout 2e soort
H_0 verwerpen	fout 1e soort	terecht

De betekenis van de kansen α en β op fouten van respectievelijk de eerste en de tweede soort kan als volgt worden weergegeven. Indien de toets vele malen wordt toegepast zal in een fractie α van alle beslissingen H_0 ten onrechte verworpen worden en in een fractie β van alle beslissingen ten onrechte niet.

Vóórdat de toets wordt toegepast dient de grootte van α te worden vastgesteld. Men neemt hiervoor veelal 5% en indien men voorzichtig wil zijn 1%. Het bepalen van de waarde van β heeft alleen zin, wanneer men de alternatieve hypothese H_1 nauwkeurig kan specificeren. Voor uitgewerkte voorbeelden verwijzen wij naar Memorandum S 308 - A 13.

Het bovenstaande zullen we nu verduidelijken aan de hand van het toetsen van een hypothese omtrent een kans en een hypothese omtrent de verwachting van een kansverdeling.

Het toetsen van een hypothese omtrent een kans

Stel dat de nulhypothese luidt $p = p_0$, of anders geschreven

$$H_0 : p = p_0 .$$

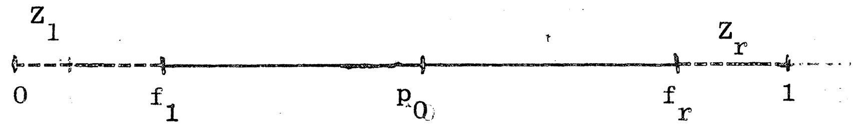
Men veronderstelt dus dat een fractie p_0 van bijvoorbeeld alle geboekte posten fout is. De fractie $\underline{f} = \frac{\underline{k}}{n}$ aan fouten in de steekproef behoeft, ook wanneer de nulhypothese juist is, niet precies gelijk te zijn aan p_0. In principe zijn zelfs alle waarden

$$0 , \frac{1}{n} , \frac{2}{n} , \ldots , 1$$

mogelijk, doch zij zijn onder H_0 niet alle even waarschijnlijk.

Wanneer men de toegestane onbetrouwbaarheid heeft vastgesteld kunnen grenzen f_1 en f_r berekend worden zodanig, dat een procedure, waarbij men H_0 verwerpt indien $f \leq f_1$ of $f \geq f_r$ is, een onbetrouwbaarheid heeft, welke hoogstens gelijk is aan α[1].

In figuur 1 is een en ander in beeld gebracht. Bij een gegeven onbetrouwbaarheid α zijn f_1 en f_r zo bepaald, dat



Figuur 1

Kritieke zones bij $H_0 : p = p_0$

bij gevonden fracties, liggende in de intervallen Z_1 en Z_r, dus tussen 0 en f_1 en tussen f_r en 1 (f_1 behoort tot Z_1 en f_r tot Z_r), de nulhypo-

---

1) Wegens het discrete karakter van de kansverdeling van $\underline{f}$ is het niet altijd mogelijk de grenzen f_1 en f_r zo te kiezen, dat de onbetrouwbaarheid precies gelijk aan α is. De werkelijke onbetrouwbaarheid mag echter nooit groter dan α zijn. Men spreekt dan van de "onbetrouwbaarheidsdrempel" α. In deze Memoranda wordt geen onderscheid gemaakt tussen onbetrouwbaarheid en onbetrouwbaarheidsdrempel.

these H_0 verworpen moet worden. Men noemt de intervallen Z_l en Z_r respectievelijk de <u>linker</u> en <u>rechter kritieke zone</u> van de toets.

Voor grote n en niet te ver van $\frac{1}{2}$ verwijderde p_0 worden de grenzen f_l en f_r van de kritieke gebieden gegeven door

$$f_l = p_0 - \xi_\alpha \sqrt{\frac{p_0(1-p_0)}{n}} - \frac{1}{2n} \qquad (1)$$

en

$$f_r = p_0 + \xi_\alpha \sqrt{\frac{p_0(1-p_0)}{n}} + \frac{1}{2n} \qquad (2)$$

waarbij ξ_α gevonden wordt uit onderstaande tabel II.

Voor zéér grote n mogen de zogenaamde "continuïteitscorrecties" $\pm\frac{1}{2n}$ verwaarloosd worden. Deze correcties zijn aangebracht omdat de discrete verdeling van de fractie $\underline{f}$ benaderd werd door een (continue) normale verdeling. (Zie ook de Memoranda S 308 - A 2 en A 3).

Tabel II

Waarden van ξ_α en $\xi_{2\alpha}$

α	ξ_α	$\xi_{2\alpha}$
0,20	1,282	0,842
0,10	1,645	1,282
0,05	1,960	1,645
0,02	2,326	2,054
0,01	2,576	2,326

Is de nulhypothese juist, dan heeft men een kans α om voor $\underline{f}$ een waarde in één van de kritieke zones te vinden en dus H_0 ten onrechte te verwerpen. Wanneer men H_0 verwerpt geschiedt dit ten gunste van de alternatieve hypothese $H_1 : p \neq p_0$. Vinden we een fractie in Z_l, dan concluderen we tot $H_1 : p < p_0$; ligt de fractie in Z_r, dan besluiten we tot $H_1 : p > p_0$.

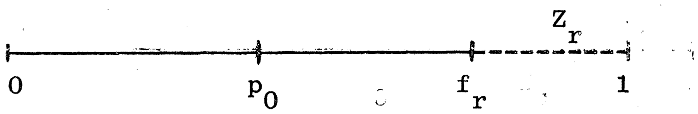
De bovenomschreven toets is een zogenaamde <u>tweezijdige toets</u>. We toetsen $H_0 : p = p_0$ tweezijdig, omdat zowel $H_1 : p < p_0$ als

$H_1 : p > p_0$ als mogelijke alternatieven gezien worden.

Luidt de nulhypothese niet $p = p_0$, maar bijvoorbeeld

$$H_0 : p \leqq p_0$$

dan passen we een zogenaamde <u>rechtseenzijdige toets</u> toe. Onder deze nulhypothese kunnen namelijk fracties kleiner dan p_0 nooit onwaarschijnlijk zijn en zal dus een linker kritieke zone Z_l ontbreken. De nulhypothese wordt dus alleen verworpen voor grote waarden van $\underline{f}$. Ook hier kan een grens f_r berekend worden zodanig, dat H_0 verwerpen voor $f \geqq f_r$ leidt tot een onbetrouwbaarheid α van de methode. Voor grote n en niet te ver van $\frac{1}{2}$ verwijderde p_0 wordt f_r weergegeven door (2), waarbij nu echter niet ξ_α maar de in Tabel II bij α gegeven waarde $\xi_{2\alpha}$ ingevuld moet worden. In figuur 2 is de situatie in beeld gebracht.



<u>Figuur 2</u>

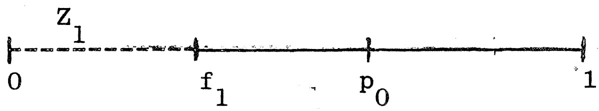
<u>Kritieke zone bij $H_0 : p \leqq p_0$</u>

Vinden we in de steekproef een fractie f in Z_r, dan wordt H_0 verworpen ten gunste van $H_1 : p > p_0$.

De nulhypothese kan ook luiden

$$H_0 : p \geqq p_0.$$

In dit geval wordt een zogenaamde <u>linkseenzijdige toets</u> toegepast. Nu kunnen alleen kleine fracties onder H_0 onwaarschijnlijk zijn en zal een rechter kritieke zone Z_r ontbreken. De bovengrens f_l van de linker kritieke zone Z_l wordt weer berekend met (1), waarbij ook nu weer in plaats van ξ_α de bij α in Tabel II gegeven waarde $\xi_{2\alpha}$ ingevuld moet worden. In figuur 3 wordt deze situatie weergegeven.



<u>Figuur 3</u>

<u>Kritieke zone bij $H_0 : p \geqq p_0$</u>

Vinden we in de steekproef een fractie f in Z_1, dan wordt H_0 verworpen ten gunste van $H_1 : p < p_0$.

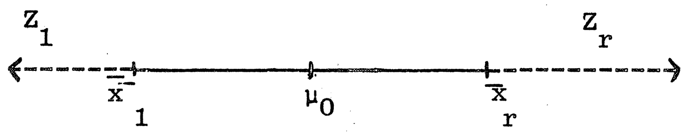
Het toetsen van een hypothese omtrent een verwachting

Laten wij aannemen dat een stochastische grootheid een kansverdeling bezit (zie Memorandum S 308 - A 2) met verwachting μ en standaardafwijking σ en dat de grootte van σ bekend is. Omtrent de verwachting luidt de nulhypothese, dat $\mu = \mu_0$, dus

$$H_0 : \mu = \mu_0 .$$

Het steekproefgemiddelde $\bar{x}$ zal nu, indien de nulhypothese juist is, in het algemeen niet veel van μ_0 afwijken. In principe zijn echter bij de meest voorkomende kansverdelingen onder H_0 alle waarden van $\bar{x}$ mogelijk.

Wanneer men de onbetrouwbaarheid α heeft vastgesteld kunnen weer grenzen $\bar{x}_1$ en $\bar{x}_r$ berekend worden zodanig, dat een procedure, waarbij men H_0 verwerpt indien $\bar{x} \leq \bar{x}_1$ of $\bar{x} \geq \bar{x}_r$, precies een onbetrouwbaarheid α heeft. In figuur 4 is de situatie geschetst. Z_1 en Z_r zijn de linker en



Figuur 4

Kritieke zones bij $H_0 : \mu = \mu_0$

rechter kritieke zones. Vindt men in de steekproef een gemiddelde $\bar{x}$, dat in Z_1 of Z_r ligt, dan verwerpt men H_0.

Bij normale verdelingen (zie Memorandum S 308 - A 3) voor iedere n en bij willekeurige verdelingen voor grote n worden de grenzen $\bar{x}_1$ en $\bar{x}_r$ van de kritieke gebieden Z_1 en Z_r gegeven door

$$\bar{x}_1 = \mu_0 - \xi_\alpha \cdot \frac{\sigma}{\sqrt{n}} \tag{3}$$

en

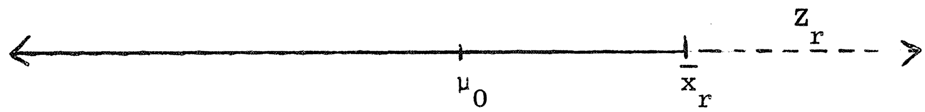
$$x_r - \mu_0 \quad \xi_\alpha \quad \overline{\sqrt{n}} \tag{4}$$

waarbı ' ξ_α gevonden wordt in Tabel II

Is de nulhypothese H_0 juist dan heeft men een kans α om voor $\bar{x}$ een waard in ·én van de kritieke zones te vinden en dus H_0 ten onrechte te verwerpen. Wanneer men H_0 verwerpt geschiedt dit ten gunste van de alternatıeve hypothese $H_1 : \mu \neq \mu_0$ Vindt men een steekpr eıgem ddelde x ın Z_1 dan concludeert men wee tot H_1 $\mu < \mu_0$; ligt $\bar{x}$ ın Z , dan beslu t me t t H_1 $\mu > \mu_0$.

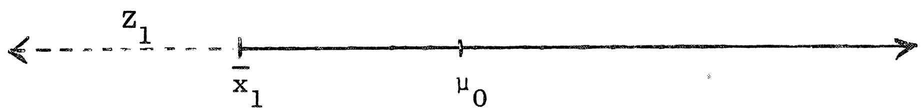
Hierboven werd een tweezijdige toets beschreven, omdat ook ie zowel $H_1 : \mu < \mu_0$ als $H_1 : \mu > \mu_0$ als alternatieven beschouwd worden.

Rechtseenzijdige en linkseenzijdige toetsen worden toegepast bij respectievelijk de nulhypothesen $H_0 : \mu \leqq \mu_0$ en $H_0 : \mu \geqq \mu_0$, dit dus geheel analoog als bij het toetsen van een kan . B'j de rech seenzijd ge toet wordt de rechter kritıeke zone naar beneden begrensd door $\bar{x}_r$; biȷ de linkseenzijdige toets wordt de linker kritieke zone naar boven begrensd door $\bar{x}_l$. Bij de berekening van $\bar{x}_l$ en $\bar{x}_r$ worden weer de formules (3) en (4) gebruikt, waarbij in plaats van ξ_α de bij α behorende waarde $\xi_{2\alpha}$ ingevuld dient te worden. De grootte van $\xi_{2\alpha}$ wordt gevonden in Tabel II De figuren 5 en 6 geven de betreffende situatıes weer.



Figuur 5

Kritieke zone bij $H_0 : \mu \leqq \mu_0$



Figuur 6

Kritieke zone bij $H_0 : \mu \geqq \mu_0$

Meestal zal de grootte van de standaardafwijking σ niet bekend zijn. Voor grote n kan dan in de formules (3) en (4) in plaats van σ de schatting

$$s = \sqrt{\frac{1}{n-1} \; \{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})\}} =$$

$$= \sqrt{\frac{1}{n-1} \; \{(x_1^2 + x_2^2 + \ldots + x_n^2) - n\bar{x}^2\}} \qquad (5)$$

ingevuld worden.

Voorbeelden

I. De directie van een firma is van mening, dat de fractie foutief berekende verkoopfacturen ca. 0,05 bedraagt. Men neemt hier genoegen mee omdat a. het moeilijk is beter personeel aan te trekken en b. er betrekkelijk weinig klachten van cliënten binnenkomen. Een accountant wil deze mening van de directie op haar juistheid toetsen.

De nulhypothese luidt dus $p = 0,05$. Omdat zowel $p < 0,05$ als $p > 0,05$ juist kan zijn dient tweezijdig getoetst te worden. De accountant kiest een onbetrouwbaarheid van 1 % en besluit een steekproef van 400 facturen te nemen. Met $p_0 = 0,05$, $\alpha = 0,01$, $\xi_\alpha = 2,576$ en $n = 400$ worden f_1 en f_r met de formules (1) en (2) berekend op respectievelijk 0,0219 en 0,0781.

De fractie fouten in de steekproef bedraagt 0,0802 en de accountant verwerpt dus de nulhypothese ten gunste van de alternatieve hypothese dat de fractie foutief berekende verkoopfacturen meer dan 0,05 bedraagt.

Indien een fractie tussen 0,0219 en 0,0781 gevonden was, zou de nulhypothese $p = 0,05$ niet verworpen zijn. Bij een gevonden fractie kleiner dan 0,0219 zou de nulhypothese verworpen zijn ten gunste van de alternatieve hypothese $p < 0,05$.

II. De directie van de vorengenoemde firma leeft ook nog in de overtuiging dat de fractie foutief geboekte facturen niet groter is dan 0,02.

De accountant heeft echter de indruk dat deze fractie hoger is en wil een en ander toetsen. Hij dient nu een rechtseenzijdige toets toe te passen.

De nulhypothese luidt $p \leq 0,02$. Als onbetrouwbaarheid wordt weer 1% gekozen en men neemt een steekproef van 900 facturen. Hoewel 0,02 vrij ver verwijderd is van 0,5 kan toch formule (2) toegepast worden omdat n groot is. Met $p_0 = 0,02$, $\alpha = 0,01$, $\xi_{2\alpha} = 2,326$ (!) en $n = 900$ wordt f_r met formule (2) berekend op 0,0314[2].

De fractie fouten in de steekproef blijkt nu 0,0236 te bedragen zodat de nulhypothese niet verworpen kan worden.

III. De directie van een groothandel beweert, dat het brutowinstpercentage in het afgelopen boekjaar ca. 35 is geweest. De accountant wil deze bewering op haar juistheid toetsen door middel van een steekproef.

De nulhypothese luidt nu $\mu = 35$. Omdat zowel $\mu < 35$ als $\mu > 35$ juist kan zijn, dient tweezijdig getoetst te worden. Men kiest een onbetrouwbaarheid van 5% en besluit een steekproef van 625 guldens uit de verkoopfacturen te nemen (vergelijk Memorandum S 308 - A 11). Omdat de standaardafwijking σ van de populatie van winstpercentages per gulden niet bekend is kunnen de grenzen $\bar{x}_1$ en $\bar{x}_r$ van de kritieke zones pas berekend worden wanneer de steekproef is uitgevoerd. Uit de steekproef wordt s met formule (5) berekend op 4,15%.

Met $\mu_0 = 35$, $s = 4,15$, $\alpha = 0,05$, $\xi_\alpha = 1,960$ en $n = 625$ worden $\bar{x}_1$ en $\bar{x}_r$ met de formules (3) en (4) berekend op respectievelijk 34,675 en 35,325. In de steekproef wordt een gemiddeld percentage van 34,316 gevonden en de nulhypothese moet dus verworpen worden ten gunste van de alternatieve hypothese, dat het brutowinstpercentage kleiner is dan 35.

---

2) Hoewel de waarde 0,02 niet dicht bij een $\frac{1}{2}$ ligt, is toch de op de normale benadering berustende formule (2) toegepast. Deze benadering is geoorloofd aangezien n = 900. De exacte waarde van f_r bedraagt 0,0311, zodat de benaderde waarde tot in 3 decimalen nauwkeurig is.

Indien in de steekproef een percentage tussen 34,675 en 35,325 was gevonden, dan had de nulhypothese niet verworpen kunnen worden. Bij een percentage in de steekproef, grot r dan 35,325, was de nulhypothese verworpen ten gunste van de alternatieve hypothese, dat het brutowinstpercentage groter is dan 35.

IV. Bij Voorbeeld III zou het kunnen zijn, dat de accountant verwacht, dat het opgegeven winstpercentage van 35 te laag is. Hij verwacht dan de nulhypothese $\mu \leq 35$ te kunnen verwerpen ten gunste van de alternatieve hypothese $\mu > 35$ en moet daartoe een rechtseenzijdige toets toepassen. Met dezelfde gegevens en dezelfde onbetrouwbaarheid als in Voorbeeld III berekent hij ($\xi_{2\alpha} = 1,645$!) $\bar{x}_r$ met formule (4) op 35,273.

In de steekproef wordt een gemiddelde brutowinstpercentage van 35,412 gevonden. De accountant verwerpt dus de nulhypothese ten gunste van de alternatieve hypothese, dat het winstpercentage groter is dan 35.

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

W

A

S308 - A 13

De fout van de tweede soort bij het toetsen van hypothesen

Bij het toetsen van hypothesen omtrent onbekende kansen of verwachtingen speelt de fout van de tweede soort een belangrijke rol; zie Memorandum S308-A 12.

Onder de fout van de tweede soort verstaat men de fout, de nulhypothese H_0 niet te verwerpen terwijl de alternatieve hypothese H_1 juist is. Het vooruit bepalen van de waarde van de kans β op een fout van de tweede soort heeft, zoals in genoemd Memorandum wordt opgemerkt, alleen zin, wanneer men de alternatieve hypothese H_1 nauwkeurig kan specificeren. We zullen hieronder een en ander eerst bespreken aan de hand van het toetsen van een hypothese omtrent een kans.

Stel dat men de hypothese H_0 wil toetsen, dat een onbekende kans gelijk is aan p_0, dus

$$H_0 : p = p_0 .$$

De alternatieve hypothese luidt bijvoorbeeld

$$H_1 : p = p_1$$

met $p_1 > p_0$.

Vóórdat de steekproef wordt genomen worden zowel de onbetrouwbaarheid α als de kans op een fout van de tweede soort β gekozen.

Omdat $p_1 > p_0$ is ligt het voor de hand alleen een rechter kritieke zone Z_r te kiezen met als ondergrens f_r.

We zullen nu n en f_r zo kiezen, dat aan de eisen betreffende de kansen α en β op fouten van de eerste resp. tweede soort is voldaan. Het verband tussen p_0, α, n en f_r wordt gelegd door formule (2) van Memorandum S308-A 12, waarbij ξ_α vervangen wordt door $\xi_{2\alpha}$:

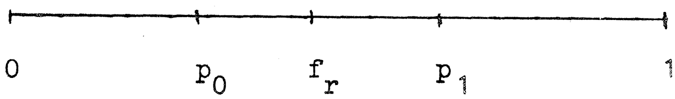
$$f_r = p_0 + \xi_{2\alpha} \sqrt{\frac{p_0(1-p_0)}{n}} \quad . \tag{1}$$

De continuïteitscorrectie $\frac{1}{2n}$ is hierin verwaarloosd.

Een fout van de tweede soort wordt gemaakt indien men, als H_1 juist is, toch in de steekproef geen fractie in de kritieke zone Z_r vindt en de kans hierop moet hoogstens gelijk zijn aan β. Hieruit kan men voor grote n en p_1 niet te ver verwijderd van $\frac{1}{2}$ afleiden, dat moet zijn voldaan aan

$$f_r = p_1 - \xi_{2\beta} \sqrt{\frac{p_1(1-p_1)}{n}} \tag{2}$$

(zie figuur 1).



Figuur 1

Een eenzijdige toets voor H_0 : $p = p_0$ met als alternatief H_1 : $p = p_1$.

De grootte van n berekent men uit de formules (1) en (2) door de rechterleden aan elkaar gelijk te stellen:

$$n = \frac{\left(\xi_{2\alpha}\sqrt{p_0 q_0} + \xi_{2\beta}\sqrt{p_1 q_1}\right)^2}{(p_1 - p_0)^2} , \qquad (3)$$

waarin $q_0 = 1-p_0$ en $q_1 = 1-p_1$. De waarden van $\xi_{2\alpha}$ en $\xi_{2\beta}$ vindt men in Tabel I. Vervolgens wordt f_r gevonden met behulp van formule (1).

Tabel I

Waarden van $\xi_{2\alpha}$ en $\xi_{2\beta}$

α , β	$\xi_{2\alpha}$, $\xi_{2\beta}$
0,20	0,842
0,10	1,282
0,05	1,645
0,02	2,054
0,01	2,326

Indien $p_1 < p_0$ is, dan wordt een linkséénzijdige toets toegepast, met een linker kritieke zone Z_1. Ook dan kan formule (3) toegepast worden om de vereiste steekproefomvang te berekenen.

Uit formule (3) blijkt, dat n groter is naarmate het verschil tussen p_0 en p_1 kleiner is. Het is ook intuïtief duidelijk, dat bij een kleiner verschil tussen p_0 en p_1 een grotere steekproefomvang nodig is om tussen deze twee kansen te kunnen beslissen.

In veel praktijksituaties kan men wel de nulhypothese expliciet formuleren, maar niet de alternatieve. Men stelt dan dat de nulhypothese $p = p_0$ getoetst moet worden en dat men een bepaalde afwijking, bijvoorbeeld $p \geq p_1$ "zeker" wil ontdekken. Ook in deze gevallen vult men de waarde van p_1 in formule (3) in. De kans op een fout van de tweede soort bedraagt dan β voor $p = p_1$ en is kleiner dan β voor $p > p_1$.

Bij het toetsen van een hypothese omtrent een verwachting (zie Memorandum S308-A 12) speelt de fout van de tweede soort eveneens een belangrijke rol.

Stel dat een stochastische grootheid een normale verdeling heeft met bekende standaardafwijking σ (zie Memorandum S308-A3) en dat men de hypothese

$$H_0 : \mu = \mu_0$$

wil toetsen tegen de alternatieve hypothese

$$H_1 : \mu = \mu_1$$

met $\mu_1 > \mu_0$. Allereerst worden weer de onbetrouwbaarheid α en de kans op een fout van de tweede soort β gekozen. Men zal een rechtséénzijdige toets toepassen en er is dus slechts een rechter kritieke zone Z_r met als ondergrens $\overline{x}_r$.

De vrijheid in de keuze van $\overline{x}_r$ en n gebruikt men weer om te voldoen aan de eisen, gesteld ten aanzien van de kansen op fouten van de eerste en de tweede soort. Bij gegeven α, β, μ_0, μ_1 en σ zal n nu moeten voldoen aan

$$n = \sigma^2 \frac{(\xi_{2\alpha} + \xi_{2\beta})^2}{(\mu_1 - \mu_0)^2} . \tag{4}$$

De waarde van $\overline{x}_r$ volgt uit formule (4) van Memorandum S308-A 12, waarin ξ_α vervangen is door $\xi_{2\alpha}$:

$$\overline{x}_r = \mu_0 + \xi_{2\alpha} \frac{\sigma}{\sqrt{n}} . \tag{5}$$

De situatie wordt ingewikkelder, wanneer men σ niet kent, niet weet of de stochastische grootheid normaal verdeeld is, of wanneer ook de standaardafwijking onder de nulhypothese verschilt van die onder de alternatieve hypothese. Ook dan bestaan er echter voor tal van gevallen toetsingsmethoden, welke analoog verlopen aan de boven geschetste procedure.

Indien een steekproef uit een willekeurige populatie wordt genomen, zal voor grote steekproefomvang n het steekproefgemiddelde $\bar{x}$ bij benadering normaal verdeeld zijn met standaardafwijking $\frac{\sigma}{\sqrt{n}}$, waarbij σ de standaardafwijking van de populatie is. Deze σ wordt geschat door de grootheid s (zie Memorandum S308-A6). Voor grote n wordt dan dezelfde toetsingsprocedure toegepast als hierboven werd beschreven (zie voorbeeld II).

Indien $\mu_1 < \mu_0$ is wordt een linkséénzijdige toets gebruikt met een linker kritieke zone Z_1. Verder verloopt de toets op dezelfde wijze als in het geval $\mu_1 > \mu_0$.

Tenslotte merken wij op, dat de grootheid $1-\beta$ het onderscheidingsvermogen van de toets met betrekking tot de alternatieve hypothese H_1 genoemd wordt.

Voorbeelden.

I. De directie van een firma is van mening, dat de fractie foutief berekende verkoopfacturen ca. 0,05 bedraagt. De accountant is echter van mening, dat deze fractie zeker 0,10 bedraagt en wil een en ander toetsen.

De nulhypothese luidt dus $p_0 = 0{,}05$ en de alternatieve hypothese luidt $p_1 \geq 0{,}10$. De accountant kiest een onbetrouwbaarheid van 1% en een onderscheidingsvermogen van 95%. De steekproefomvang wordt met (3), waarin $\alpha = 0{,}01$, $\xi_{2\alpha} = 2{,}326$, $\beta = 0{,}05$, $\xi_{2\beta} = 1{,}645$, $p_0 = 0{,}05$ en $p_1 = 0{,}10$, berekend op 400,3. Voor n = 401 wordt hierna met (1) de grootte van f_r berekend op 0,0753.

Wordt in de steekproef een fractie kleiner dan 0,0753 gevonden, dan wordt H_0 niet verworpen. Indien een fractie gelijk aan of groter dan 0,0753 gevonden wordt, dan verwerpt men H_0 ten gunste van de alternatieve hypothese H_1.

II. De directie van een firma beweert, dat het brutowinstpercentage in het afgelopen boekjaar minder dan 35 is geweest. De accountant wil een en ander toetsen. Hij vermoedt dat dit percentage hoger is en wil, behoudens een kleine kans, de nulhypothese verwerpen als het percentage 37 of meer bedraagt.

De nulhypothese luidt nu $\mu \leq 35$ en de alternatieve hypothese luidt $\mu \geq 37$. De accountant kiest een onbetrouwbaarheid van 5% en een onderscheidingsvermogen van 95%.

Omdat de standaardafwijking σ van de populatie van winstpercentages per gulden niet bekend is, neemt hij een voorlopige steekproef van 200 guldens uit de verkoopfacturen (vergelijk Memorandum S308-A11). Hieruit wordt s berekend op 8,25%.

Uit formule (4) volgt, met $\alpha = 0{,}05$, $\beta = 0{,}05$, $\xi_{2\alpha} = \xi_{2\beta} = 1{,}645$, dat $n \geq 184{,}18$ gekozen moet worden. De voorlopige steekproef behoeft dus niet verder te worden uitgebreid.

Met formule (5) wordt, met $n = 200$, $\bar{x}_r$ bepaald op 35,960. Wordt er in de steekproef een gemiddelde, kleiner dan 35,960 gevonden, dan wordt H_0 niet verworpen. Wordt daarentegen een gemiddeld brutowinstpercentage groter dan of gelijk aan 35,960 gevonden, dan wordt H_1 geaccepteerd.

typ. RMW

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S308-A14

De kans op het bij een steekproef-onderzoek ten onrechte accepteren van populaties

Bij het steekproefsgewijze controleren van een lijst met posten, bijvoorbeeld een accuratessecontrole, zal men de lijst in het algemeen slechts dan accepteren, wanneer bij de gegeven steekproefomvang het aantal gevonden fouten niet boven een zeker toelaatbaar geacht maximum ligt. Wordt in de steekproef hieraan niet voldaan, dan gaat men tot integrale controle van de lijst over.

Stel dat een steekproef van n posten of guldens wordt genomen en dat hierin k fouten worden gevonden. De lijst wordt afgekeurd indien $k > k_0$ en goedgekeurd indien $k \leq k_0$. De grootte van de afkeurgrens k_0 hangt onder andere af van de steekproefomvang n en de maximaal toelaatbaar geachte fractie fouten p_0 in de populatie.

Indien nu de werkelijke fractie fouten p in de populatie groter is dan p_0, dan is er toch nog een kans β om in de steekproef k_0 of minder fouten te vinden. Deze kans β is de kans op het ten onrechte accepteren van de lijst en hangt af van de grootte van n, k_0 en p_0. Terloops merken wij op, dat het ten onrechte accepteren geheel overeenkomt met een fout van de tweede soort bij het toetsen van een hypothese; vgl. Memorandum S308-A13.

Gewoonlijk wordt de kans β niet achteraf berekend, maar eist men bij voorbaat, dat deze kans hoogstens β_0 (bijvoorbeeld 0,01 of 0,05) mag bedragen. Bij gegeven steekproefomvang n stelt men dan vast waar de afkeurgrens k_0 gekozen moet worden om aan deze eis te voldoen.

De grootte van k_0 wordt, voor grote n en p_0 niet te ver van 0,5 verwijderd, berekend met formule (1):

$$k_0 = np_0 - \xi_{2\beta_0} \cdot \sqrt{np_0(1-p_0)} - \tfrac{1}{2}. \qquad (1)$$

Hierin wordt $\xi_{2\beta_0}$ gevonden in Tabel I. Formule (1) mag slechts gebruikt worden als de omvang van de populatie zeer groot is.

Tabel I.

Waarden van $\xi_{2\beta_0}$

β_0	$\xi_{2\beta_0}$
0,20	0,842
0,10	1,282
0,05	1,645
0,02	2,054
0,01	2,326
0,005	2,576
0,001	3,090

Bij de berekening van k_0 met formule (1) zal men meestal geen gehele waarde vinden. Men rondt dan naar beneden af tot het grootste gehele getal k_0', dat kleiner is dan de gevonden waarde k_0. Dit houdt echter in, dat de maximale waarde van β in werkelijkheid kleiner is dan de in formule (1) gebruikte waarde β_0. Wanneer men vast wil houden aan de oorspronkelijk gekozen waarde β_0, dan kan men dit bereiken door de steekproefomvang n te verkleinen. Om dié waarde van n te vinden die, samen met k_0' het maximum van β gelijk aan β_0 maakt, vult men in formule (1) in het linkerlid k_0' en in het rechterlid p_0 en $\xi_{2\beta_0}$ in en lost men vervolgens n weer op. Zie ook Voorbeeld I.

Voor kleine p_0 en grote n wordt k_0 gevonden in Tabel II.[1] Indien in deze tabel een streepje (-) staat, betekent dit, dat voor het bijbehorende produkt np_0 zelfs de kans op geen enkele fout in de steekproef groter is dan de bijbehorende β_0. Ook hier is bij een gekozen steekproefomvang n de maximale waarde van β soms lager dan de gekozen waarde van β_0, zodat het mogelijk is n kleiner te kiezen zonder dat k_0 van waarde verandert. Zie ook Voorbeeld II.

Men dient goed onderscheid te maken tussen een uitspraak in de vorm van een betrouwbaarheidsinterval (zie Memorandum S308-A4) en het type uitspraken, zoals hierboven behandeld is. Geeft men een lange reeks betrouwbaarheidsintervallen op met een onbetrouwbaarheid α, dan betekent dit, dat in een fractie α van alle uitspraken een interval gegeven wordt, dat de ware fractie p niet zal bevatten. Formuleert men de conclusie in de vorm van het goed- of afkeuren van partijen en eist men, dat de kans een "slechte" partij goed te keuren hoogstens gelijk aan β is, dan betekent dit, dat van die gevallen, waarin de partij "slecht" is, hoogstens een fractie β niet ontdekt zal worden. In het eerste geval betreft het aantal onjuiste uitspraken dus een fractie α van alle uitspraken en in het tweede geval een fractie β van alle uitspraken, waarbij een "slechte" partij betrokken is.

Bij een controle op ernstige fouten eist men meestal, dat geen enkele fout gevonden mag worden, dus $k_0 = 0$. Dit geval eist echter een uitvoeriger behandeling, welke plaats vindt in Memorandum S308-A15.

Voorbeelden

I. Bij een groot bedrijf met veel weekloners wordt de lijst met berekende weeklonen iedere week vóór de uitbetaling steekproefsgewijze gecontroleerd. Een fractie fout berekende lonen van 5% wordt nog juist

1) Deze tabel is berekend, gebruik makend van de benadering door de Poisson-verdeling.

acceptabel gevonden en men wil maximaal 1% kans lopen een niet acceptabele lijst goed te keuren. De afkeurgrens k_0 wordt bij een steekproefomvang van 200 met formule (1) bepaald. Met n = 200, p_0 = 0,05, β_0 = 0,01 en $\xi_{2\beta_0}$ = 2,326, vindt men k_0 = 2,33, dus k_0' = 2. Alleen wanneer men minder dan drie fouten vindt zal men de lijst dus accepteren.

Nu is echter het maximum van β lager dan β_0 = 0,01. Om dit maximum te vinden vult men in formule (1) voor k_0, n en p_0 espectievelijk in 2, 200 en 0,05 en lost men $\xi_{2\beta_0}$ op. Men vindt $\xi_{2\beta_0}$ = 2,433, hetgeen overeenkomt met β_0 = 0,075.

Vindt men het niet nodig, dat β_0 daalt to 0,0075 en wil men dus vasthouden aan β_0 = 0,01, dan kan met een kleinere steekproef worden volstaan. De kleinst mogelijke steekproefomvang bij p_0 = 0,05, β_0 = 0,01 en k_0 = 2 wordt gevonden door deze waarden in formule (1) in te vullen, n vervolgens op te lossen en dan af te ronden tot het dichtstbijgelegen grotere gehele getal. In ons geval leidt formule (1) tot n = 189,61 en er kan dus worden volstaan met een steekproef van 190 weeklonen.

II. Bij een accuratesse-controle van een lijst met een zeer groot aantal posten accepteert men deze lijst indien deze maximaal 1% fout berekende posten bevat. Men neemt een steekproef van 700 posten en wil maximaal 1% risico lopen de lijst ten onrechte te accepteren. In Tabel II vindt men, met np_0 = 7 en β_0 = 0,01, k_0 = 1. De lijst zal dus aanvaard worden wanneer men niet meer dan één fout in de steekproef vindt.

Bij n = 700 en p_0 = 0,01 is de kans op minder dan twee fouten in de steekproef gelijk aan 0,0073, dus veel minder dan de gekozen β_0. Deze kans op minder dan twee fouten wordt groter indien de steekproefomvang kleiner wordt. Uit Tabel II blijkt, dat voor n = 665 nog geldt k_0 = 1. In dit geval is de kans op minder dan twee fouten in de steekproef gelijk aan 0,0099, dus bijna gelijk aan de aangenomen β_0.

Tabel II

Waarden van k_0 voor kleine waarden van p_0

np_0			β_0 = 0,001	0,01	0,02	0,05	0,10
	$\leqq$	2,30	-	-	-	-	-
2,35	t/m	2,95	-	-	-	-	0
3,00	"	3,85	-	-	-	0	0
	3,90		-	-	-	0	1
3,95	t/m	4,60	-	-	0	0	1
4,65	"	4,70	-	0	0	0	1
4,75	"	5,30	-	0	0	1	1
5,35	"	5,80	-	0	0	1	2
5,85	"	6,25	-	0	1	1	2
6,30	"	6,60	-	0	1	2	2
	6,65		-	1	1	2	2
6,70	t/m	6,90	-	1	1	2	3
6,95		7,50	0	1	1	[illegible]	3
7,55		7,75	0	1	[illegible]	2	3
7,80	'	7 95	0	1	2	3	3
8,00	"	8,40	0	1	2	3	4
8,45	"	9,05	0	2	2	3	4
9,10	"	9,15	0	2	3	3	4
	9,20		0	2	3	4	4
	9,25		1	2	3	4	4
9,30	t/m	10,00	1	2	3	4	5
10,05	"	10,50	1	3	3	4	5
	10,55		1	3	3	5	6
10,60	t/m	11,20	1	3	4	5	6
11,25	"	11,60	2	[illegible]	4	5	6
11,65		11,75	2	4	[illegible]	[illegible]	6
	11,80		2	4	4	[illegible]	7

Tabel II - Vervolg 1

np_0			β_0 = 0,001	0,01	0,02	0,05	0,10
11,85	t/m	12,00	2	4	4	6	7
12,05	"	12,95	2	4	5	6	7
13,00	"	13,05	2	4	5	6	8
	13,10		3	4	5	6	8
13,15	t/m	13,40	3	5	5	7	8
13,45	"	14,20	3	5	6	7	8
14,25	"	14,40	3	5	6	7	9
14,45	"	14,55	3	5	6	8	9
14,60	"	14,75	3	6	6	8	9
	14,80		4	6	6	8	9
14,85	t/m	15,40	4	6	7	8	9
15,45	"	15,70	4	6	7	8	10
15,75	"	15,95	4	6	7	9	10
16,00	"	16,15	4	7	7	9	10
16,20	"	16,45	4	7	8	9	10
16,50	"	16,55	5	7	8	9	10
16,60	"	16,95	5	7	8	9	11
17,00	"	17,40	5	7	8	10	11
17,45	"	17,50	5	8	8	10	11
17,55	"	17,75	5	8	9	10	11
17,80	"	18,05	5	8	9	10	12
18,10	"	18,20	6	8	9	1	12
18,25	"	18,75	6	8	9	11	12
	18,80		6	9	9	11	12
18,85	t/m	18,95	6	9	10	11	12
19,00	"	19,40	6	9	10	11	13
19,45	"	19,60	6	9	10	12	13
19,65	"	20,00	7	9	10	12	13
20,2	"	20,6	7	10	11	12	14

Tabel II - Vervolg 2

np_0			β_0 = 0,001	0,01	0,02	0,05	0,10
20,8	t/m	21,0	7	10	11	13	14
	21,2		8	10	11	13	14
	21,4		8	10	11	13	15
21,6	t/m	21,8	8	11	12	13	15
22,0	"	22,4	8	11	12	14	15
	22,6		8	11	12	14	16
	22,8		9	11	13	14	16
	23,0		9	12	13	14	16
23,2	t/m	23,6	9	12	13	15	16
	23,8		9	12	13	15	17
	24,0		9	12	14	15	17
	24,2		10	13	14	15	17
24,4	t/m	24,6	1	13	14	16	17
24,8	"	25,2	10	13	14	16	18
	25,4		10	13	1	16	18
25,6	t/m	25,8	11	14	1	17	18
26,0	"	26,4	11	14	1	17	19
	26,6		11	14	16	17	19
26,8	t/m	27,0	11	15	16	18	19
27,2	"	27,6	1	15	16	18	20
	27,8		12	15	17	18	20
	28,0		12	15	17	19	20
28,2	t/m	28,4	12	16	17	19	21
28,6	"	28,8	13	16	17	19	21
	29,0		13	16	18	19	21
	29,2		1	16	18	20	21
29,4	t/m	29,8	13	17	18	20	22
	30,0		14	17	18	20	22

W

A

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S308-A15

Controle op ernstige fouten

Bij een controle op ernstige fouten zal men meestal eisen, dat in de steekproef geen enkele fout mag worden gevonden. De in Memorandum S308-A14 ingevoerde afkeurgrens k_0 is dan dus gelijk aan 0. Vindt men één of meer fouten, dan gaat men tot integrale controle over.

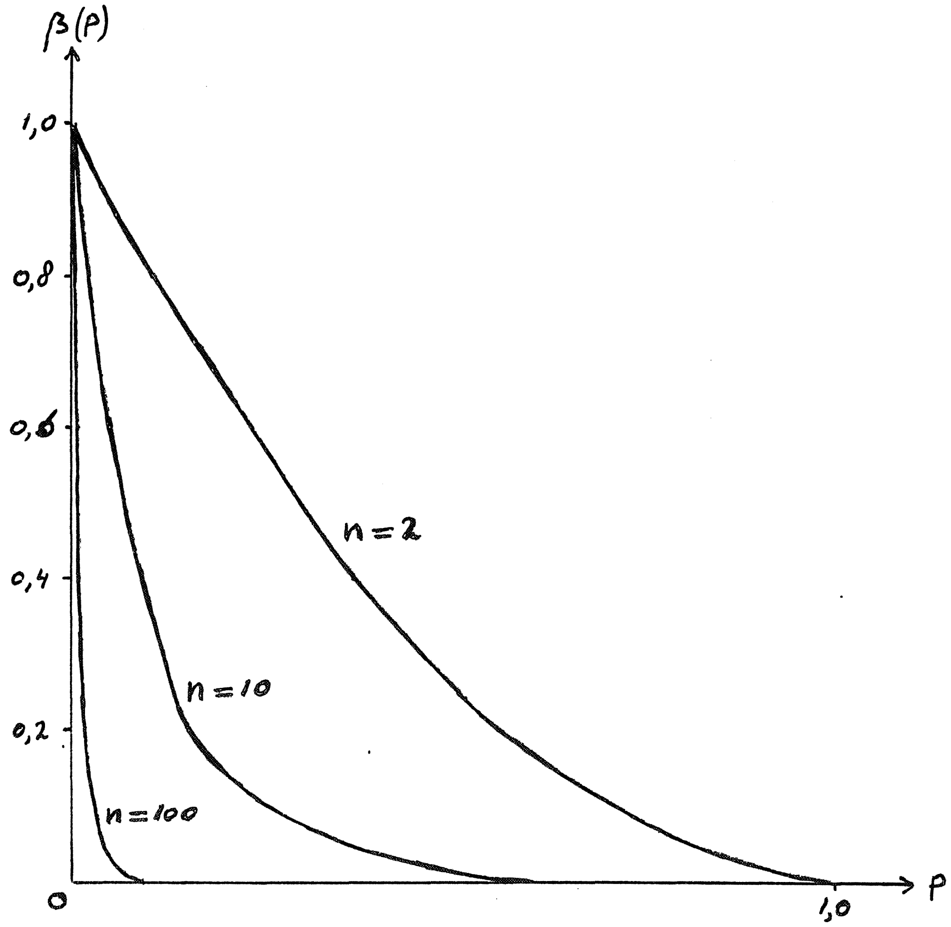
Wanneer men geen enkele fout in een steekproef vindt, dan kan men altijd de bovengrens berekenen van het betrouwbaarheidsinterval, waarin de ware fractie fouten behoudens een onbetrouwbaarheid α ligt (zie Memorandum S308-A4).

In veel gevallen is men echter meer geïnteresseerd in de kans een slechte populatie goed te keuren dan in een betrouwbaarheids-interval voor een onbekende fractie fouten. Een slechte populatie is bijvoorbeeld een populatie, waarin meer dan een fractie p_0 van de posten fout geboekt is, of meer dan een fractie p_0 van het opgegeven totaalbedrag niet gedekt. De kans β op goedkeuren is gelijk aan de kans geen enkele fout in de steekproef aan te treffen en men is dus vooral in deze kans geïnteresseerd wanneer de onderzochte fractie p in de populatie groter is dan p_0.

Op intuïtieve gronden kan men inzien, dat de kans β afneemt naarmate de fractie p groter is en ook naarmate de steekproefomvang n groter wordt. In figuur 1 is β voor enkele waarden van de steekproefomvang n geschetst als functie van de fractie fouten p. Omdat

β van p afhangt schrijven we $\beta(p)$ in plaats van β. Uit de figuur blijkt, dat $\beta(p)$ inderdaad sterk afneemt voor toenemende p en toenemende n.

Aangezien de waarde van n nog niet vastligt, kan men eisen, dat deze zó gekozen wordt, dat de kans $\beta(p)$ voor $p = p_0$, dus $\beta(p_0)$, niet groter mag zijn dan een zekere waarde β_0. Hieruit volgt dan de minimaal vereiste omvang van de steekproef.



Figuur 11

De kans $\beta(p)$ op het niet ontdekken van een fout als functie van de werkelijke fractie fouten p en van de steekproefomvang n.

In Tabel I wordt deze omvang voor enkele waarden van p_0 en β_0 gegeven voor het geval waarin de steekproefomvang klein is ten opzichte van de omvang van de gehele populatie. Uit figuur 1 blijkt dan, dat $\beta(p)$ voor $p > p_0$ zeker kleiner is dan $\beta(p_0)$.

Tabel I

Minimale waarden van n voor verschillende waarden van p_0 en β_0 .

p_0 \ β_0	0,05	0,02	0,01	0,005	0,001
0,05	59	77	90	104	135
0,02	149	194	228	263	342
0,01	299	390	459	528	688
0,005	598	781	919	1058	1379
0,001	2995	3911	4603	5296	6905

Men leest uit Tabel I af, dat β_0 nog aanzienlijk verlaagd kan worden zonder een al te grote toename van de steekproefomvang n, maar dat men bij hogere eisen ten opzichte van p_0 zeer omvangrijke steekproeven moet gaan nemen.

Wanneer een aselecte steekproef van omvang 459 wordt genomen en men een populatie niet accepteert en dus tot volledige controle overgaat als in de steekproef één of meer fouten gevonden worden, dan is de kans, dat een populatie met een fractie fouten groter dan 1% niet geheel wordt onderzocht, hoogstens 0,01. Anders gezegd: wordt de bovenomschreven controlemethode toegepast, dan zal in een lange reeks controles hoogstens 1% van de gevallen, waarin de werkelijke fractie fouten groter is dan 1%, niet worden ontdekt. Overeenkomstige uitspraken kunnen gedaan worden, wanneer men steekproeven neemt van een andere in Tabel I vermelde omvang.

Heeft men eenmaal een bepaalde omvang van de steekproef gekozen, dan is het uiteraard niet alleen van belang te weten hoe groot $\beta(p)$ is voor een bepaalde waarde van p, maar dan is men ook geïnteresseerd in het verdere verloop van $\beta(p)$ als functie van p. Voor een veel gebruikte waarde van n, nl. n = 459, wordt in Tabel II de waarde van $\beta(p)$ voor enkele karakteristieke waarden van p gegeven.

Tabel II

Waarden van $\beta(p)$ voor n = 459
en verschillende waarden van p.

p	$\beta(p)$	p	$\beta(p)$
0,0001	0,9552	0,02	$94 \cdot 10^{-6}$
0,001	0,6318	0,03	$85 \cdot 10^{-8}$
0,005	0,1002	0,04	$73 \cdot 10^{-10}$
0,01	0,0099	0,05	$60 \cdot 10^{-12}$

De bovenbeschreven methoden kunnen worden toegepast, zowel op lijsten, opgevat als een populatie van posten, als op lijsten, opgevat als een populatie van guldens. De lijst dient uiteraard "volledig" te zijn. Bij een postensteekproef dienen alle posten, bij een guldenssteekproef alle guldens aanwezig te zijn, omdat een niet aanwezige post of gulden niet in de steekproef opgenomen en op juistheid gecontroleerd kan worden.

Laten wij nu aannemen, dat de ernstige fouten slechts kunnen bestaan uit het opschrijven van te hoge bedragen, het toevoegen van niet bestaande posten, het opschrijven van hogere totaalsommen dan de werkelijke of het verkeerd transporteren van deeltotalen.

Toepassing van de bovenstaande gedachtengang op deze situatie heeft tot een controlemethode geleid, welke in de literatuur bekend is geworden onder de naam Guldenrangnummermethode [1].

Stel dat de te controleren lijst bestaat uit N posten met een totaalbedrag van T guldens. Uit de lijst worden n guldens aselect aangewezen en men gaat na of hier foute guldens bij zijn. Men controleert echter niet alleen de aangewezen guldens maar tevens de gehele posten of de gedeelter van de posten waartoe deze guldens behoren. Grotere posten hebben een grotere kans om aangewezen te worden en kunnen zelfs verschillende malen aangewezen worden. Men controleert dus hoogstens n posten. Zodra een "ernstige fout" gevonden wordt gaat men alle posten controleren.

Omdat niet alleen de aangewezen guldens, maar tevens de gehele posten of gedeelten van posten, waartoe zij behoren, worden gecontroleerd, wordt er meer gecontroleerd dan alleen de n getrokken guldens. De kans op het niet ontdekken van ernstige fouten is dan ook in het algemeen kleiner dan de bovenstaande tabellen vermelden.

Behalve fouten in posten worden bij deze controle ook fouten in optellingen en fouten, gemaakt bij het transporteren, steekproefsgewijs gecontroleerd Zo vindt men een fout in een optelling van een pagina wanneer bijv. de 6.650-ste gulden zich in de steekproef bevindt en bij het opzoeken blijkt, dat de bladzijde, waarop bovenaan 6.000 en onderaan 6.800 is geschreven, in werkelijkheid deze 6.650-ste gulden niet bevat omdat de optelling slechts tot bijv. 6.500 loopt. Op de bladzijde bevindt zich dan een volledig gefraudeerde "post van 300 gulden'. Een fraude, gepleegd door fout te transporteren, zal men ontdekken wanneer bij het zoeken naar bijv. de 6.530-ste gulden blijkt, dat de ene bladzijde met 6.500 eindigt en de volgende met 70000 begint.

---

1) A. VAN HEERDEN, "Steekproeven als middel van accountantscontrole", Maandblad voor Accountancy en Bedrijfshuishoudkunde 35(1961), pp. 453-475.

Voor het aanwijzen van guldens gebruikt men aselecte getallen (zie Memorandum S308-A8). Dit aanwijzen wordt vereenvoudigd, wanneer men beschikt over lijsten met naar grootte gerangschikte aselecte getallen, welke getallen alle liggen tussen 0 en het totaalbedrag T van de lijst Kiest men voor p_0 en β_0 de waarde 0,01, dan dienen er minstens 459 (zie Tabel I) aselecte getallen op zo'n lijst voor te komen. Het risico een fout van 1% van het totaalbedrag T in één enkele post niet te ontdekken kan geëlimineerd worden door in die gevallen, waarin de sprong tussen twee in grootte opeenvolgende getallen meer is dan 1% van het totaalbedrag T, aselect een getal bij te loten tussen de twee reeds aanwezige aselecte getallen in. Hiervoor zijn in het algemeen slechts enkele nieuwe aselecte getallen nodig, terwijl men de mogelijkheid uitsluit, dat een fout van meer dan 1% van T in een enkele post niet wordt opgemerkt.

Lijsten van aselecte getallen van het bovenbeschreven type werden reeds elders [2] samengesteld Ze bevatten elk ongeveer 500 getallen tussen respectievelijk 0 en 15.000, 0 en 20,000, 0 en 25.000, 0 en 30.000, 0 en 40.000, 0 en 50,000, 0 en 60.000, 0 en 70.000, 0 en 80.000, 0 en 90.000 en 0 en 100.000. Deze lijsten kunnen ook gebruikt worden indien het totaalbedrag T meer dan 100.000 gulden is. Bij bijv. T = 200.000 gebruikt men dan de lijst van 0 tot 20.000 en wijst men tientjes aan in plaats van guldens.

Gaat men bij het gebruik van deze lijsten uit van bekende subtotalen voor het opzoeken van de guldens, dan dient men steeds vanuit de subtotalen verder te tellen; terugtellen is dus niet geoorloofd.

De boven beschreven methode is natuurlijk niet slechts toepasbaar bij een controle op ernstige fouten. Ook als het accuratessefouten en dergelijke betreft, kan de methode worden toegepast. In

[2] Nederlandse Accountants-Maatschap, Afdeling Administratieve Organisatie en Mechanisatie, Rotterdam.

die gevallen zal men echter de afkeurgrens k_0 niet altijd gelijk aan 0 willen stellen, maar kunnen ook andere waarden van k_0 in aanmerking komen.

Voorbeelden van toepassingen

Van de controles waarbij de guldenrangnummermethode toegepast kan worden, worden hieronder enkele genoemd:

uitbetalingen bij een verzekeringmaatschappij (zie Memorandum S334-B3);

declaraties van onkosten, waarbij de verschillende onkosten op één declaratie als posten worden gezien;

loonbestanddelen bij een loonadministratie, waarbij dan de loonbestanddelen als posten worden gezien;

de afloop van debiteurensaldi.

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S308-A16

Steekproeven uit eindige populaties

In de voorgaande Memoranda werd vrijwel overal verondersteld, dat een steekproef genomen werd uit een oneindig grote populatie. In werkelijkheid zullen populaties steeds een eindige omvang bezitten en bijvoorbeeld uit N posten bestaan. Van deze N posten zijn er een aantal, bijvoorbeeld M, "fout" en wij trachten nu de fractie $\frac{M}{N}$ te schatten met behulp van een aselecte steekproef. Deze kan op twee manieren worden genomen.

Steekproef met teruglegging. Er wordt aselect een post aangewezen. De kans dat dit een foute post is, bedraagt $\frac{M}{N}$. Na het constateren of de post al dan niet fout is, wordt deze "teruggelegd", waardoor de oorspronkelijke samenstelling van de populatie is hersteld. Bij de tweede trekking is de kans op een foute post weer $\frac{M}{N}$ en deze kans blijft bij voortdurende teruglegging even groot bij elke volgende trekking.

Steekproef zonder teruglegging. Er wordt aselect een post aangewezen. De kans dat dit een foute post is, bedraagt $\frac{M}{N}$. De post wordt niet "teruggelegd", waardoor de samenstelling van de populatie veranderd is. De kans op een foute post is bij een volgende trekking afhankelijk geworden van het resultaat van de eerste trekking en deze eigenschap handhaaft zich bij alle volgende trekkingen.

De kans op k foute posten in een steekproef van n posten is dus bij een steekproef zonder teruglegging niet gelijk aan die bij een steekproef met teruglegging. Het verschil wordt echter kleiner naarmate de omvang N van de populatie en het aantal foute posten M groter zijn ten opzichte van de omvang n van de steekproef. Het "uitlichten" van een post verandert dan weinig aan de samenstelling van de populatie en de kansen op een foute post veranderen eveneens weinig bij de n opeenvolgende trekkingen.

Bij een oneindig grote populatie verandert een trekking zonder teruglegging niets aan de samenstelling. In dat geval zijn de resultaten bij een steekproef zonder teruglegging precies gelijk aan die bij een steekproef met teruglegging.

Hierboven werden schattingen van fracties gemaakt. Eenzelfde redenering geldt voor het schatten van verwachtingen.

In de voorgaande Memoranda (met uitzondering van Memorandum S308-A11) werd steeds verondersteld, dat de populaties oneindig groot of zeer groot waren. Indien aan deze voorwaarde niet is voldaan ondergaan sommige formules enige wijziging, welke hieruit bestaat, dat standaardafwijkingen met $\sqrt{\frac{N-n}{N-1}}$, dus varianties met $\frac{N-n}{N-1}$, worden vermenigvuldigd (voor de begrippen standaardafwijking en variantie zie Memorandum S308-A2).

De factor $\frac{N-n}{N-1}$ wordt de <u>correctiefactor voor eindige populaties</u> (engels: finite population correction, f.p.c.) genoemd. Hij kan ook geschreven worden als $1-g$, waarin $g = \frac{n-1}{N-1}$.

Wij zullen nu eerst nagaan welke situaties zich kunnen voordoen wanneer de populaties niet oneindig groot zijn:

a). <u>Steekproef met teruglegging</u>: het maakt geen verschil of de populatie eindig of oneindig groot is; de steekproefomvang n kan die van de populatie N zelfs overtreffen; het resultaat is hetzelfde als met een steekproef zonder teruglegging uit een oneindig grote populatie; de in de voorgaande Memoranda gegeven formules blijven onveranderd.

b). Steekproef zonder teruglegging: uit een eindige populatie terwijl de verhouding $\frac{n}{N}$ klein is, zeg < 0,1; de situatie verschilt weinig van die bij een oneindig grote populatie en de formules in de voorgaande Memoranda vormen in het algemeen dan ook goede benaderingen van de exacte formules.

c). Steekproef zonder teruglegging: uit een eindige populatie terwijl de verhouding $\frac{n}{N}$ groot is, zeg $\geq$ 0,1: nu moeten alle standaardafwijkingen met de genoemde correctiefactor voor eindige populaties worden vermenigvuldigd.

Hieronder worden voor alle voorgaande Memoranda de veranderingen in de gegeven formules nagegaan. Ook de gewijzigde formules gelden slechts voor grote steekproefomvang n en dus grote populatieomvang N.

S308-A4

De grenzen voor een betrouwbaarheidsinterval voor een onbekende kans worden

$$p_{*}^{*} = f \pm \xi_\alpha \cdot \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{\frac{f(1-f)}{n}} \qquad \text{(vgl. A4; 1,2)}$$

en de onnauwkeurigheid gaat over in

$$\xi_\alpha \cdot \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{\frac{f(1-f)}{n}}$$

Voor zeer kleine fracties f kan Tabel II slechts dan gebruikt worden indien de verhouding $\frac{n}{N}$ klein is.

S308-A5

Bij een steekproef uit een normale verdeling wordt impliciet altijd verondersteld dat de populatie oneindig groot is en de formules ondergaan dus geen wijziging.

S308-A6

De grenzen voor een betrouwbaarheidsinterval voor een onbekende verwachting luiden nu:

$$m^{*}_{*} = x \pm \xi_\alpha \cdot \sqrt{\frac{N-n}{N-1}} \cdot \frac{s}{\sqrt{n}} \qquad \text{(vgl. A6; 2,3)}$$

en de onnauwkeurigheid wordt

$$\xi_\alpha \cdot \sqrt{\frac{N-n}{N-1}} \cdot \frac{s}{\sqrt{n}} \quad .$$

S308-A7

De steekproefomvang n kan weer bepaald worden door de onnauwkeurigheid van de schatting gelijk te stellen aan de maximaal toegelaten onnauwkeurigheid. Bij de schatting van een onbekende fractie met absolute onnauwkeurigheid v wordt n gevonden met

$$n \geqq \frac{\xi_\alpha^2 . N\ f(1-f)}{\xi_\alpha^2 . f(1-f) + (N-1)v^2} \quad . \qquad \text{(vgl. A7; 2)}$$

Bij de schatting van een onbekende fractie met relatieve onnauwkeurigheid θ wordt n gevonden met

$$n \geqq \frac{\xi_\alpha^2 . N\ (1-f)}{\xi_\alpha^2 . (1-f) + (N-1)\ \theta^2 f} \quad . \qquad \text{(vgl. A7; 4)}$$

Indien de fractie f dicht bij 0 of 1 ligt kan de vereiste steekproefomvang n niet met een eenvoudige formule bepaald worden.

Bij de schatting van de verwachting uit een normaal verdeelde populatie is er nooit sprake van een eindige populatie. Dit kan wel het geval zijn bij de schatting van de verwachting van een willekeurig verdeelde populatie.
Voor een maximaal toegestane absolute onnauwkeurigheid d wordt n bepaald met

$$n \geqq \frac{\xi_\alpha^2 . N\ \sigma^2}{(N-1)d^2 + \xi_\alpha^2 . \sigma^2} \quad . \qquad \text{(vgl. A7; 9)}$$

Indien de standaardafwijking σ van de populatie niet bekend is, wordt σ door de schatting s van σ vervangen.

S308-A9

Bij de toepassing van de in dit Memorandum behandelde methoden dienen zonodig dezelfde correcties aangebracht te worden als in de Memoranda S308-A6 en A7.

S308-A11

In dit Memorandum is met de eindigheid van de populaties rekening gehouden. In de formules (5) en (6) herkent men gemakkelijk de correctiefactor voor eindige populaties.

S308-A12

De formules voor de grenzen van de kritieke gebieden bij het toetsen van een hypothese omtrent een onbekende kans gaan nu luiden:

$$f_l = p_o - \xi_\alpha \cdot \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{\frac{p_o(1-p_o)}{n}} - \frac{1}{2n} ; \qquad \text{(vgl. A12; 1)}$$

$$f_r = p_o + \xi_\alpha \cdot \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{\frac{p_o(1-p_o)}{n}} + \frac{1}{2n} . \qquad \text{(vgl. A12; 2)}$$

Bij het toetsen van een hypothese omtrent een onbekende verwachting van een willekeurige eindige populatie luiden de formules voor de grenzen van de kritieke zones

$$\bar{\bar{x}}_l = \mu_o - \xi_\alpha \cdot \sqrt{\frac{N-n}{N-1}} \cdot \frac{\sigma}{\sqrt{n}} \qquad \text{(vgl. A12; 3)}$$

$$\bar{\bar{x}}_r = \mu_o + \xi_\alpha \cdot \sqrt{\frac{N-n}{N-1}} \cdot \frac{\sigma}{\sqrt{n}} \qquad \text{(vgl. A12; 4)}$$

S308-A13

Bij het toetsen van een hypothese omtrent een onbekende kans wordt de steekproefomvang n nu bepaald met

$$n \geq \frac{N\left(\xi_{2\alpha}\cdot\sqrt{p_o q_o} + \xi_{2\beta}\cdot\sqrt{p_1 q_1}\right)^2}{\left(\xi_{2\alpha}\cdot\sqrt{p_o q_o} + \xi_{2\beta}\cdot\sqrt{p_1 q_1}\right)^2 + (N-1)\left(p_1 - p_o\right)^2} . \quad \text{(vgl. A13; 3)}$$

Bij het toetsen van een hypothese omtrent de verwachting van een normale verdeling wordt de steekproefomvang n steeds bepaald met formule (4) van Memorandum S308-A13.

Voor de berekening van de grenzen van de kritieke zones worden bij het toetsen van een hypothese omtrent eén kans de gewijzigde formules (A12; 1) en (A12; 2), en bij het toetsen van een hypothese omtrent een verwachting van een willekeurige verdeling de gewijzigde formules (A12; 3) en (A12; 4) toegepast.

S308-A14

Voor p_o niet te ver van 0,5 verwijderd en voor grote steekproefomvang n wordt de afkeurgrens k_o gevonden met

$$k_o = np_o - \xi_{2\beta_o} \cdot \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{np_o(1-p_o)} - \tfrac{1}{2} . \quad \text{(vgl. A14; 1)}$$

Voor kleine p_o en grote n mag Tabel II slechts gebruikt worden indien de verhouding $\frac{n}{N}$ zeer klein is.

S308-A15

De in dat Memorandum gegeven Tabel I geldt in feite voor oneindig grote populaties. Bij eindige populatieomvang N kan met kleinere waarden van de steekproefomvang n vol[illegible]aan worden. In Tabel (A15; I) worden deze waarden van n voor verschillende waarden van N vermeld.

Er dient nog op gewezen te worden, dat voor steekproeven zonder teruglegging uit een populatie met kleine omvang N waarbij dus ook de steekproefomvang n klein is, de hierboven gecorrigeerde formules niet meer gelden. In het algemeen zijn er dan geen eenvoudige formules meer beschikbaar. Wel bestaan er tabellen, waarvan men in vele gevallen gebruik kan maken.

Tabel (A15; I)

Minimale waarden van n voor verschillende waarden van p_o, β_o en N.

p_o \ β_o	0,05	0,02	0,01	0,005	0,001
	N = 1.000				
0,05	56	73	85	98	125
0,02	137	175	203	230	289
0,01	257	322	367	409	496
0,005	449	541	600	652	747
0,001	949	979	989	994	998
	N = 2.000				
0,05	57	74	87	100	130
0,02	142	184	215	245	314
0,01	276	353	409	463	581
0,005	516	646	736	820	995
0,001	1552	1716	1799	1858	1936
	N = 5.000				
0,05	58	75	88	102	132
0,02	146	189	222	255	330
0,01	289	374	437	500	642
0,005	563	722	839	952	1204
0,001	2252	2712	3008	3265	3742
	N = 10.000				
0,05	58	75	89	102	133
0,02	147	191	225	258	336
0,01	293	381	447	513	664
0,005	580	750	877	1002	1287
0,001	2587	3236	3688	4111	4985

p_o \ β_o	0,05	0,02	0,01	0,005	0,001
			N = 20.000		
0,05	58	76	89	103	134
0,02	147	192	226	260	338
0,01	295	385	452	520	675
0,005	588	765	897	1029	1331
0,001	2780	3551	4111	4652	5838
			N = 50.000		
0,05	58	76	89	103	134
0,02	148	193	227	261	340
0,01	297	387	456	524	682
0,005	594	774	910	1045	1359
0,001	2906	3761	4397	5024	6448
			N = 75.000		
0,05	58	76	89	103	134
0,02	148	193	227	261	341
0,01	297	388	456	525	684
0,005	595	776	913	1049	1365
0,001	2935	3809	4464	5112	6595
			N = 100.000		
0,05	58	76	89	103	134
0,02	148	193	227	261	341
0,01	297	388	457	525	684
0,005	595	777	914	1051	1368
0,001	2949	3834	4498	5157	6671

A W

β_0 / P_0	0,05	0,02	0,01	0,005	0,001	0,05	0,02	0,01	0,005	0,001
	N = 15.000					N = 20.000				
0,05	59	77	90	103	135	59	77	90	104	135
0,02	148	193	227	260	339	148	193	227	261	339
0,01	296	385	452	518	672	296	386	453	521	676
0,005	586	761	892	1021	1317	589	766	898	1030	1332
0,001	2715	3442	3964	4462	5534	2781	3552	4112	4653	5839
	N = 30.000					N = 40.000				
0,05	59	77	90	104	135	59	77	90	104	135
0,02	148	194	228	262	340	149	194	228	262	341
0,01	297	387	455	523	680	297	388	456	524	682
0,005	592	771	905	1039	1347	594	773	909	1044	1355
0,001	2850	3666	4268	4855	6168	2885	3725	4348	4960	6342
	N = 50.000					N = 75.000				
0,05	59	77	90	104	135	59	77	90	104	135
0,02	149	194	228	262	341	149	194	228	262	342
0,01	298	388	457	525	683	298	389	457	526	685
0,005	595	775	911	1046	1360	596	777	914	1050	1366
0,001	2907	3762	4398	5025	6449	2936	3810	4465	5113	6596
	N = 100.000					N = ∞				
0,05	59	77	90	104	135	59	77	90	104	135
0,02	149	194	228	262	342	149	194	228	263	342
0,01	298	389	458	526	685	299	390	459	528	688
0,005	596	778	915	1052	1369	598	781	919	1058	1379
0,001	2950	3835	4499	5158	6672	2995	3911	4603	5296	6905

W

A

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S308-A17

Het AOQL-keuringssysteem voor populaties

In de Memoranda S308-A14 en A15 werd een keuringssysteem besproken, waarbij op grond van het aantal fouten k in een steekproef van de omvang n werd besloten de betreffende populatie goed of af te keuren. Ligt bij dit systeem k boven k_0, dan wordt de populatie afgekeurd, hetgeen volledige controle tengevolge kan hebben; is $k \leq k_0$, dan wordt de populatie goedgekeurd. De grens k_0 wordt zó bepaald, dat de kans op het goedkeuren van een populatie met een fractie fouten groter dan of gelijk aan p_0 niet groter is dan β_0.

In de bovengenoemde Memoranda was het slechts de bedoeling te onderzoeken of de fractie fouten p in de aangeboden populatie groter of kleiner is dan een van te voren bepaalde fractie p_0. Indien populaties regelmatig op deze wijze worden gekeurd en slechts goedgekeurde populaties worden "afgeleverd", dan weten we zeker, dat van de populaties met een fractie fouten groter dan p_0 maximaal een fractie β_0 ten onrechte wordt goedgekeurd en afgeleverd.

Het AOQL-keuringssysteem

Naast situaties, waarin de bovengenoemde methode goed aansluit bij hetgeen men nastreeft, komen ook situaties voor, waarin men meer geïnteresseerd is in de gemiddelde kwaliteit van alle doorgegeven populaties dan in de kwaliteit van de individuele, goedgekeurde populaties. Onder "kwaliteit" wordt hier verstaan de fractie goede elementen. Het is daarbij niet zo erg, dat er wel eens een populatie van slechte kwaliteit wordt doorgegeven, mits daar populaties van zeer goede kwaliteit tegenover staan, waardoor op de lange duur de gemiddelde kwaliteit aan een van te voren bepaalde minimumeis voldoet.

Deze situatie doet zich onder andere voor wanneer de gekeurde populaties naderhand bij elkaar gevoegd worden en daardoor hun identiteit verliezen. Het kan dan van belang zijn te kunnen garanderen, dat de gemiddelde kwaliteit van de door het bijeenvoegen verkregen populatie aan de genoemde minimumeis voldoet.

Indien het mogelijk is gevonden fouten alsnog te verbeteren kan men een keuringssysteem toepassen van de volgende form:

1e. Trek een aselecte steekproef van n elementen uit de populatie en bepaal het aantal foute elementen k;

2e. Keur de populatie goed indien $k \leq k_0$ en controleer de populatie volledig indien $k > k_0$;

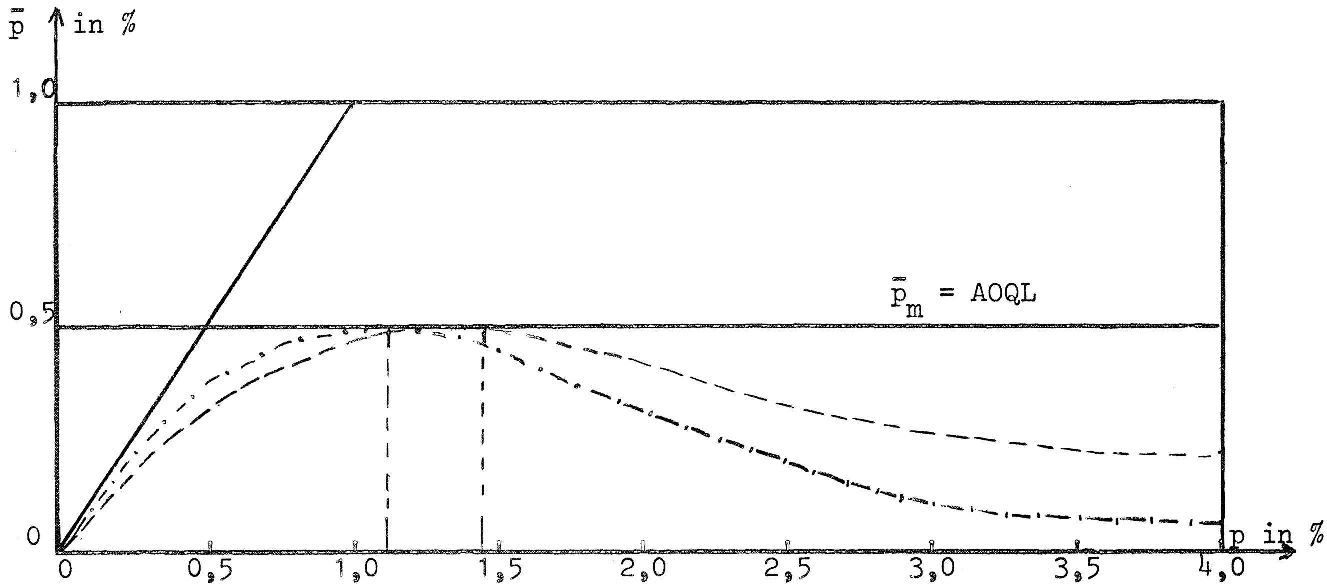
3e. <u>Verbeter alle in de steekproef of tijdens de volledige controle gevonden foute elementen.</u>

Uiteraard kan dit systeem slechts dan toegepast worden indien de keuring van een element niet destructief is.

Bij deze methode worden alle aangeboden populaties, partieel of integraal gecontroleerd, doorgegeven terwijl men er door de keuze van n en k_0 voor kan zorgen dat de <u>gemiddelde</u> kwaliteit van de doorgegeven populaties niet lager ligt dan een van te voren vastgesteld niveau K. De keuze van n en k_0 hangt naast K ook af van de omvang N van de populaties.

Stel dat de ter keuring aangeboden populaties alle dezelfde omvang N bezitten en alle dezelfde fractie p aan foute elementen bevatten. Kiezen wij bepaalde waarden voor n en k_0, dan kunnen wij ons afvragen hoe groot de gemiddelde fractie fouten $\bar{p}$ in de afgeleverde populaties bedraagt. Deze gemiddelde fractie $\bar{p}$, de "Average Outgoing Quality" (AOQ) genoemd, hangt bij gegeven N, n en k_0 nog af van p, zoals in figuur 1 wordt weergegeven.

Bij kleine waarden van p is de kans op meer dan k_0 fouten in de steekproef, en dus de kans op integrale controle, klein. De gemiddelde fractie fouten $\bar{p}$ is dan slechts iets kleiner dan de fractie fouten p vóór de inspectie. Bij grotere waarden van p wordt de kans op meer dan k_0 fouten in de steekproef groter en wordt dus een grotere fractie van de aangeboden populaties integraal gecontroleerd en verbeterd. Het verschil tussen p en $\bar{p}$ wordt dan groter. Bij waarden van p dicht bij 1 wordt de kans op meer dan k_0 fouten in de steekproef zeer groot, de kans op volledige controle dus ook en de gemiddelde fractie fouten $\bar{p}$ van de afgeleverde populaties dientengevolge zeer klein. In het geval p = 1 is $\bar{p}$ gelijk aan 0.



Figuur 1

De gemiddelde uitgaande kwaliteit $\bar{p}$ als functie van de aangeboden kwaliteit p

——————— = AOQ zonder inspectie

- - - - - - - = AOQ met inspectie (N = 1000, n = 70, k_0 = 0)

-.-.-.-.-.- = AOQ met inspectie (N = 1000, n = 145, k_0 = 1)

In het algemeen is de waarde van p niet bekend en kan $\bar{\bar{p}}$ dus niet worden opgegeven, wel echter de maximale waarde $\bar{\bar{p}}_m$ die $\bar{\bar{p}}$ aan kan nemen. Deze maximale waarde $\bar{\bar{p}}_m$ wordt door $\bar{\bar{p}}$ bij een bepaalde waarde, zeg p', aangenomen en wordt de "Average Outgoing Quality Limit" (AOQL) genoemd. De minimale gemiddelde kwaliteit van de afgeleverde populaties is dan gelijk aan $1 - \bar{\bar{p}}_m$, wat ook de kwaliteit 1 - p van de aangeboden populatie is.

Heeft men de minimaal toegelaten uitgaande gemiddelde kwaliteit K_{min} vastgesteld, dan kunnen bij gegeven populatieomvang N combinaties van n en k_0 berekend worden waarvoor K_{min} wordt bereikt. Dit is onder andere gedaan door DODGE en ROMIG [1]. Men vindt in hun tabellen voor iedere populatieomvang N en verschillende waarden van $\bar{\bar{p}}_m$ zes paren waarden van n en k_0 (= c in de tabellen). Bij ieder van deze zes paren waarden (n; k_0) wordt aan de gestelde eis ten aanzien van $\bar{\bar{p}}_m$ voldaan.

Een keuze uit deze paren (n; k_0) heeft alleen zin, wanneer iets bekend is omtrent de (gemiddelde) fractie foute elementen p in de aangeboden populaties. In de tabellen wordt dan dát paar waarden gegeven, waarvoor het gemiddeld te inspecteren aantal elementen minimaal is.

In Tabel I wordt voor $\bar{\bar{p}}_m$ = 0,01 (1% AOQL) de door DODGE en ROMIG gegeven tabel weergegeven. Uit deze tabel is af te lezen, dat niet al te grote fluctuaties in de omvang N van de populatie geen invloed hebben op de keuze van n en k_0. Verandering in de kwaliteit van de aangeboden populatie beïnvloedt hoogstens de optimale keuze van n en k_0, niet de grootte van $\bar{\bar{p}}_m$.

Voorbeeld

Bij een grote onderneming worden iedere week ca. 900 weeklonen berekend. Men wenst deze lonen iedere week steekproefsgewijze te controleren en daarmede te bereiken, dat gemiddeld minstens 99% van de uitbetaalde lonen correct is berekend.

Bij toepassing van het AOQL-keuringssysteem met $\bar{\bar{p}}_m$ = 0,01 vindt men in Tabel I de volgende paren waarden voor n en k_0:

(35;0), (80;1) en (120;2).

1) H.F. Dodge en H.G. Romig, "Sampling Inspection Tables", John Wiley and Sons, New York, Second edition (1959), pp. 197-204.

Indien niets bekend is omtrent de gemiddelde fractie gemaakte berekeningsfouten kan elk paar gekozen worden. Men weet echter uit ervaring, dat ca. 0,75% van de lonen fout berekend wordt. Om nu het gemiddelde aantal te inspecteren weeklonen minimaal te maken wordt $n = 120$ en $k_0 = 2$ gekozen.

Tabel I

Waarden van n en k_0 bij $\bar{\bar{p}}_m = 0{,}01$

p tussen (in %)→	0,00 - 0,02		0,03 - 0,20		0,21 - 0,40		0,41 - 0,60		0,61 - 0,80		0,81 - 1,00	
N	n	k_0	n	k_0	n	k_0	n	k_0	n	k_0	n	k_0
1 - 25	alle	0	alle	0	alle	0	alle	0	alle	0	alle	0
26 - 50	22	0	22	0	22	0	22	0	22	0	22	0
51 - 100	27	0	27	0	27	0	27	0	27	0	27	0
101 - 200	32	0	32	0	32	0	32	0	32	0	32	0
201 - 300	33	0	33	0	33	0	33	0	33	0	65	1
301 - 400	34	0	34	0	34	0	70	1	70	1	70	1
401 - 500	35	0	35	0	35	0	70	1	70	1	70	1
501 - 600	35	0	35	0	75	1	75	1	75	1	75	1
601 - 800	35	0	35	0	75	1	75	1	75	1	120	2
801 - 1.000	35	0	35	0	80	1	80	1	120	2	120	2
1.001 - 2.000	36	0	80	1	80	1	130	2	130	2	180	3
2.001 - 3.000	36	0	80	1	80	1	130	2	185	3	235	4
3.001 - 4.000	36	0	80	1	135	2	135	2	185	3	295	5
4.001 - 5.000	36	0	85	1	135	2	190	3	245	4	300	5
5.001 - 7.000	37	0	85	1	135	2	190	3	305	5	420	7
7.001 - 10.000	37	0	85	1	135	2	245	4	310	5	430	7
10.001 - 20.000	85	1	135	2	195	3	250	4	435	7	635	10
20.001 - 50.000	85	1	135	2	255	4	380	6	575	9	990	15
50.001 - 100.000	85	1	135	2	255	4	445	7	790	12	1520	22

W

A

STICHTING

MATHEMATISCH CENTRUM

2e BOERHAAVESTRAAT 49

AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S308 - A18

Het verband tussen het toetsen van hypothesen en het opstellen van betrouwbaarheidsintervallen.

Tussen het toetsen van een nulhypothese en het opstellen van een betrouwbaarheidsinterval bestaat een nauw verband. In dit Memorandum wordt deze samenhang aangetoond en met numerieke voorbeelden geïllustreerd.

De bedoelde samenhang wordt uitvoerig besproken voor het geval, waarin fracties of kansen in het geding zijn. Aangezien het verband bij verwachtingen van normale of willekeurige verdelingen geheel analoog is, wordt hierop slechts summier ingegaan.

Toetsen en betrouwbaarheidsintervallen betreffende kansen

Toetsen. In de Memoranda S308-A12 en A13 wordt het toetsen van een hypothese uitvoerig behandeld. Vóórdat een steekproef is getrokken wordt reeds een hypothese $H_o : p = p_o$ opgesteld. Verder bepaalt men de toegestane kans α op een fout van de eerste soort en de omvang n van de steekproef. De grenzen f_l en f_r van de kritieke zones volgen dan uit de formules

$$f_1 = p_o - \xi_\alpha \cdot \sqrt{\frac{p_o(1-p_o)}{n}} - \frac{1}{2n}, \qquad (1)$$

$$f_r = p_o + \xi_\alpha \cdot \sqrt{\frac{p_o(1-p_o)}{n}} + \frac{1}{2n}. \qquad (2)$$

Bij een eenzijdige toets wordt slechts één van deze grenzen berekend, waarbij dan ξ_α wordt vervangen door $\xi_{2\alpha}$. De bij α behorende waarden van ξ_α en $\xi_{2\alpha}$ worden gevonden in bijvoorbeeld Tabel II van Memorandum S308-A12.

Vervolgens wordt een aselecte steekproef van de omvang n getrokken en de fractie f in de steekproef bepaald. De nulhypothese H_o wordt verworpen indien f in één van de kritieke zones ligt.

Betrouwbaarheidsintervallen. Het opstellen van betrouwbaarheidsintervallen voor een onbekende kans wordt besproken in Memorandum S308-A4. Men begint met het vaststellen van de toegestane onbetrouwbaarheid α en de omvang n van de steekproef. Daarna wordt de steekproef getrokken en berekent men de fractie f met het gezochte kenmerk in de steekproef. De grenzen van het betrouwbaarheidsinterval voor de onbekende kans p volgen voor grote n uit de formules

$$p^* = f + \xi_\alpha \cdot \sqrt{\frac{f(1-f)}{n}}, \qquad (3)$$

$$p_* = f - \xi_\alpha \cdot \sqrt{\frac{f(1-f)}{n}}, \qquad (4)$$

Bij een eenzijdig betrouwbaarheidsinterval wordt slechts één van deze grenzen berekend, waarbij dan ξ_α door $\xi_{2\alpha}$ wordt vervangen. De bij α behorende waarden van ξ_α en $\xi_{2\alpha}$ worden in bovengenoemde tabel vermeld. Met een onbetrouwbaarheid α ligt de werkelijke waarde p in het zo bepaalde interval.

Onderling verband tussen toets en betrouwbaarheidsinterval bij kansen.

Stel dat de stochastische grootheid $\underline{f}$ in een steekproef van de omvang n de waarde f heeft aangenomen en dat men een betrouwbaarheidsinterval op wil stellen voor de onbekende kans p met een onbetrouwbaarheid α. De vraag is dan welke waarden van p in dit interval moeten worden opgenomen. Om deze vraag voor een bepaalde waarde p_o van p te beantwoorden kijken we nu naar hetgeen wij zouden doen als wij de nulhypothese H_o:$p = p_o$ zouden hebben getoetst. Deze nulhypothese zou zijn verworpen wanneer f in de kritieke zone (berekend met dezelfde waarden van α en n) ligt en niet zijn verworpen wanneer dit niet het geval is. Wij besluiten nu p_o in het eerste geval niet en in het tweede geval wel in het betrouwbaarheidsinterval op te nemen. Dezelfde vraag wordt vervolgens beantwoord voor alle andere mogelijke waarden van p, waarbij de in de steekproef gevonden waarde f wordt vastgehouden.

Het bij de gevonden fractie f behorende betrouwbaarheidsinterval voor de onbekende kans p wordt op deze wijze dus gedefiniëerd als de verzameling van alle waarden van p_o, waarvoor de nulhypothese H_o:$p = p_o$ niet wordt verworpen.

Bij het bepalen van deze verzameling behoeven wij de beschreven methode niet steeds voor alle waarden van p afzonderlijk toe te passen. Immers, als wij de waarde van p_o groter maken, dus in figuur 1 b) naar rechts verschuiven, dan zullen ook de kritieke waarden f_l en f_r naar

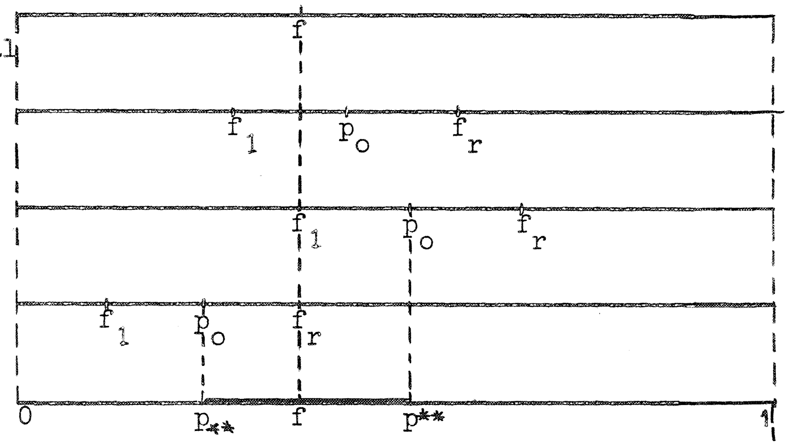
a) gevonden fractie

b) p_o in betrouwbaarheidsinterval voor p

c) bovengrens van betrouwbaarheidsinterval

d) ondergrens van betrouwbaarheidsinterval

e) betrouwbaarheidsinterval voor p



Figuur 1

De constructie van een tweezijdig betrouwbaarheidsinterval voor een onbekende kans p rechts verschuiven. Valt f bij een bepaalde p_o reeds in de linker kritieke zone, dan zal dat voor alle waarden $p > p_o$ ook het geval zijn. De bovengrens p^{**} van het betrouwbaarheidsinterval is dan ook die waarde van p, waarvoor f_l juist samenvalt met de gevonden waarde f; vergelijk de figuren 1 c) en e). Om deze p^{**} te vinden stellen wij in formule (1) p_o gelijk aan p^{**} en f_1 gelijk aan f. Dit geeft een vierkantsvergelijking in p^{**}, waaruit voor p^{**} volgt [1]:

$$p^{**} = \frac{(nf + \frac{1}{2}) + \frac{1}{2}\,\xi_\alpha^{\,2} + \xi_\alpha \cdot \sqrt{\frac{1}{n}(nf + \frac{1}{2})\,(n - nf - \frac{1}{2}) + \frac{1}{4}\,\xi_\alpha^{\,2}}}{n + \xi_\alpha^{\,2}}. \qquad (5)$$

De ondergrens van de bovenbedoelde verzameling waarden van p_o is de waarde p_{**}, waarvoor de fractie f precies samenvalt met de ondergrens f_r van de rechter kritieke zone; vergelijk de figuren 1 d) en e). Voor de berekening van p_{**} dient in formule (2) f_r door f en p_o door p_{**} vervangen te worden. Dit geeft weer een vierkantsvergelijking, nu in p_{**}, waaruit volgt [2]:

$$p_{**} = \frac{(nf - \frac{1}{2}) + \frac{1}{2}\,\xi_\alpha^{\,2} - \xi_\alpha \cdot \sqrt{\frac{1}{n}(nf - \frac{1}{2})\,(n - nf + \frac{1}{2}) + \frac{1}{4}\,\xi_\alpha^{\,2}}}{n + \xi_\alpha^{\,2}}. \qquad (6)$$

Bij de gevonden fractie f wordt geen der hypothesen $H_o: p = p_o$, met p_o tussen p_{**} en p^{**}, verworpen. Een nulhypothese $H_o: p = p_o$, met $p_o \leqq p_{**}$ of $p_o \geqq p^{**}$ wordt wel verworpen. Het interval $p_{**} - p^{**}$ is een betrouwbaarheidsinterval voor de onbekende kans p, behorende bij de gevonden fractie f.

De hierboven bij de constructie van het betrouwbaarheidsinterval getoetste nulhypothesen bezaten alle dezelfde kans α op een fout van de eerste soort, dat is de kans, dat onder H_o de fractie f in één van de kritieke zones terecht komt. Deze kans α is dan echter ook de kans, dat het betrouwbaarheidsinterval $p_{**} - p^{**}$ niet de ware waarde p bevat en is dus de onbetrouwbaarheid van de methode.

---

[1] Zonder bewijs vermelden we, dat voor p^{**} de grootste van de twee wortels genomen moet worden.

[2] Voor p_{**} moet de kleinste wortel genomen worden.

De formules (5) en (6) wijken op het oog sterk af van de formules (3) en (4). Men kan echter aantonen, dat ze voor zeer grote n practisch dezelfde uitkomsten geven. Beide paren formules geven slechts benaderingen voor de exacte grenzen van het betrouwbaarheidsinterval, doch de formules (5) en (6) geven betere benaderingen dan de formules (3) en (4). Beide paren formules gelden verder slechts voor steekproeven met teruglegging, of voor steekproeven zonder teruglegging uit zeer grote of oneindig grote populaties.

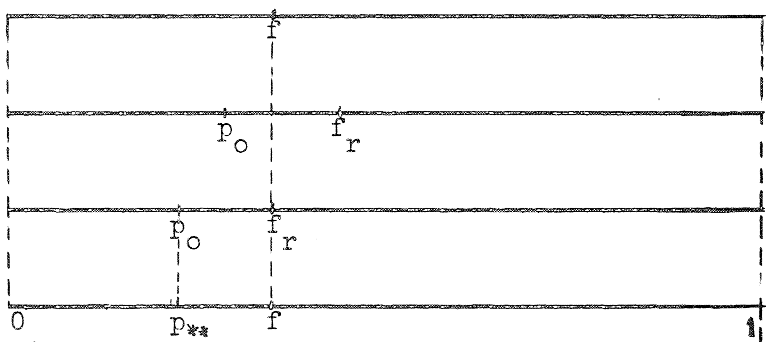
Eenzijdige toetsen en eenzijdige betrouwbaarheidsintervallen bij kansen.

Evenals in het boven behandelde tweezijdige geval, bestaat er ook verband tussen de eenzijdige toets en het eenzijdige betrouwbaarheidsinterval. De afleiding van dit eenzijdige begrensde interval uit de eenzijdige toets is geheel analoog aan die van het tweezijdig begrensde interval uit de tweezijdige toets. Men beperkt zich nu echter tot één grens.

Een nulhypothese van de gedaante $H_o : p \leqq p_o$ wordt een rechtseenzijdige toets genoemd en bezit een rechter kritieke zone Z_r. De ondergrens f_r van deze zone wordt gevonden met formule (2), waarbij ξ_α door $\xi_{2\alpha}$ vervangen dient te worden. Indien de in de steekproef gevonden fractie f in deze zone ligt, dus indien $f \geqq f_r$, dan wordt H_o verworpen met een kans α op een fout van de eerste soort.

Willen wij op grond van dezelfde n, α en f een eenzijdig betrouwbaarheidsinterval voor de onbekende kans p opstellen, dan kiezen we hier-

a) gevonden fractie
b) p_o in betrouwbaarheidsinterval voor p
c) ondergrens van betrouwbaarheidsinterval
d) betrouwbaarheidsinterval voor p



Figuur 2

De constructie van een linkseenzijdig begrensd betrouwbaarheidsinterval voor een onbekende kans p voor weer de verzameling van waarden p_o, waarvoor de hypothese $H_0: p \leq p_o$ niet wordt verworpen. Ook hier behoeft dit niet voor iedere p_o afzonderlijk te worden nagegaan. Aan de hand van hetgeen bij tweezijdig begrensde intervallen is gezegd, verifiëert men eenvoudig, dat het betrouwbaarheidsinterval nu alleen een linkergrens p_{**} heeft, welke dezelfde is als de in formule (6) gegeven grens, mits men ξ_α door $\xi_{2\alpha}$ vervangt; vergelijk figuur 2.

Het bij de gevonden f behorende betrouwbaarheidsinterval luidt dus $p_{**} < p \leq 1$. Men noemt het een linkseenzijdig begrensd betrouwbaarheidsinterval of kortheidshalve een linkseenzijdig betrouwbaarheidsinterval. <u>Bij een rechtseenzijdige toets behoort dus een linkseenzijdig betrouwbaarheidsinterval.</u>

Het verband tussen een linkseenzijdige toets met een linker kritieke zone en een rechtseenzijdig betrouwbaarheidsinterval wordt op dezelfde wijze gelegd. Voor de bovengrens p^{**} van dit betrouwbaarheidsinterval $0 \leq p < p^{**}$ vinden we dan de p^{**} van formule (5), met ξ_α vervangen door $\xi_{2\alpha}$. <u>Bij een linkseenzijdige toets behoort dus een rechtseenzijdig betrouwbaarheidsinterval.</u>

Ook in het eenzijdige geval is de kans α op een fout van de eerste soort gelijk aan de onbetrouwbaarheid α van het bijbehorende betrouwbaarheidsinterval. Voor het verband tussen p_* en p_{**} en tussen p^* en p^{**} geldt hetzelfde als in het tweezijdige geval is opgemerkt.

<u>Voorbeelden.</u>

I. In een steekproef van de omvang n = 300 vindt men een fractie f = 0,200 met het gezochte kenmerk. Leidt men met de formules (5) en (6) een tweezijdig betrouwbaarheidsinterval af met een onbetrouwbaarheid $\alpha = 0{,}05$ ($\xi_\alpha = 1{,}960$), dan luidt dit

$$0{,}157 < p < 0{,}251$$

Volgens de voorgaande theorie betekent dit dus, dat alle hypothesen $H_o: p = p_o$ met p_o tussen 0,157 en 0,251 niet verworpen worden, wanneer men toetst met een onbetrouwbaarheid $\alpha = 0{,}05$ en in een steekproef van de omvang n = 300 voor f de waarde 0,200 vindt.

Berekent men het interval met behulp van de formules (3) en (4), dan is het resultaat

$$0{,}155 < p < 0{,}245 \qquad (7)$$

Het interval is dus iets naar links verschoven en iets korter geworden. Bovendien gaat het genoemde verband met het toetsen niet steeds meer op. De waarde $p_o = 0{,}156$ zou bij toetsen worden verworpen, terwijl deze wel in het betrouwbaarheidsinterval (7) ligt. Dit wordt dus veroorzaakt doordat bij de afleiding van de formules (3) en (4) een benadering meer wordt gemaakt dan bij de afleiding van de formules (5) en (6). Het is eenvoudig in te zien, dat het ook kan voorkomen, dat bij toetsen een zekere waarde p_o niet wordt verworpen (bijvoorbeeld $p_o = 0{,}250$) terwijl dezelfde waarde desondanks niet in het interval (7) ligt.

II. In een steekproef van de omvang n = 300 vindt men een fractie f = 0,200 met het gezochte kenmerk. Leidt men met formule (6) een linkseenzijdig betrouwbaarheidsinterval af met een onbetrouwbaarheid $\alpha = 0{,}05$ ($\xi_{2\alpha} = 1{,}645$), dan luidt dit

$$0{,}163 < p \leqq 1.$$

Dit betekent dus, dat alle hypothesen $H_o : p \leqq p_o$ met $p_o > 0{,}163$ niet verworpen worden, indien men toetst met een onbetrouwbaarheid $\alpha = 0{,}05$ en in een steekproef van de omvang n = 300 voor f de waarde 0,200 vindt.

Berekent men het interval met behulp van formule (4), dan is het resultaat

$$0{,}162 < p \leqq 1. \qquad (8)$$

De ondergrens is dus iets naar links verschoven en het interval is wat langer geworden. Verder gaat het genoemde verband met het toetsen niet steeds meer op. De ondergrens $p_{**} = 0{,}163$ van het hier gevonden eenzijdige betrouwbaarheidsinterval ligt hoger dan de ondergrens $p_{**} = 0{,}157$ van het in Voorbeeld I gevonden tweezijdige betrouwbaarheidsinterval. Dit betekent, dat bijvoorbeeld een nulhypothese $H_o : p = 0{,}160$ bij tweezijdige toetsing <u>niet</u> en een nulhypothese $H_o : p \leqq 0{,}160$ bij rechtseenzijdige toetsing <u>wel</u> verworpen zou worden, bij dezelfde onbetrouwbaarheid $\alpha = 0{,}05$ en dezelfde in een steekproef van 300 stuks gevonden waarde f = 0,200.

De rechtseenzijdige toets leidt voor waarden van f, groter dan 0,160 dus eerder tot verwerping dan de tweezijdige toets. Voor de linkseenzijdige toets geldt een dergelijke conclusie met betrekking tot waarden van f, kleiner dan 0,160.

Toetsen en betrouwbaarheidsintervallen betreffende verwachtingen.

Het verband tussen het toetsen van hypothesen omtrent verwachtingen en het opstellen van betrouwbaarheidsintervallen voor onbekende verwachtingen is geheel analoog aan dat bij kansen. We nemen voorlopig aan, dat de standaardafwijking σ van de populatie, waaruit een steekproef wordt getrokken, bekend is. Vóórdat de steekproef wordt getrokken wordt een nulhypothese $H_o: \mu = \mu_o$ opgesteld. Daarna wordt de toegestane kans α op een fout van de eerste soort en de omvang n van de steekproef bepaald. De grenzen $\bar{x}_l$ en $\bar{x}_r$ van de kritieke zones worden gevonden met de formules

$$\bar{x}_l = \mu_o - \xi_\alpha \cdot \frac{\sigma}{\sqrt{n}}, \tag{9}$$

$$\bar{x}_r = \mu_o + \xi_\alpha \cdot \frac{\sigma}{\sqrt{n}}. \tag{10}$$

Uit de steekproefuitkomsten wordt tenslotte het gemiddelde $\bar{x}$ berekend. Ligt $\bar{x}$ in één van de kritieke zones, dan wordt de hypothese H_o verworpen.

Heeft men omgekeerd uit de steekproefuitkomsten het gemiddelde $\bar{x}$ berekend en wil men, met onbetrouwbaarheid α, een betrouwbaarheidsinterval voor de onbekende verwachting μ opstellen, dan bepaalt men de verzameling waarden μ_o, waarvoor de hypothese $H_o: \mu = \mu_o$ niet wordt verworpen. De ondergrens m_* van deze verzameling vindt men, door in formule (10) $\bar{x}_r$ door $\underline{x}$ en μ_o door m_* te vervangen. Oplossing van m_* geeft direct

$$m_* = \bar{x} - \xi_\alpha \cdot \frac{\sigma}{\sqrt{n}}, \tag{11}$$

Op eenzelfde wijze vindt men de bovengrens m^* met formule (9):

$$m^* = \bar{x} + \xi_\alpha \cdot \frac{\sigma}{\sqrt{n}}, \tag{12}$$

De formules (11) en (12) zijn precies dezelfde formules als die, gegeven in de Memoranda S308-A5 en A6.

Ook hier geldt weer, dat de onbetrouwbaarheid α van de schattingsmethode gelijk is aan de kans α op een fout van de eerste soort bij de toets. Verder bestaat ook hier hetzelfde verband tussen eenzijdig begrensde betrouwbaarheidsintervallen. In het eenzijdige geval wordt ξ_α vervangen door $\xi_{2\alpha}$.

Indien de standaardafwijking σ niet bekend is, dient deze eerst op de bekende wijze geschat te worden met de grootheid s. In de formules wordt dan σ door s vervangen; vergelijk Memorandum S308-A5.

Voorbeelden.

III. In een steekproef van de omvang n = 100 uit een normale verdeling met bekende standaardafwijking, groot 5,11, werd een gemiddelde $\bar{\bar{x}} = 25{,}40$ gevonden. Bij een onbetrouwbaarheid α van 1% ($\xi_\alpha = 2{,}576$) worden met de formules (11) en (12) de grenzen van het betrouwbaarheidsinterval voor de verwachting μ berekend:

$$24{,}08 < \mu < 26{,}72. \tag{13}$$

Een hypothese $H_o: \mu = \mu_o$, met μ_o tussen 24,08 en 26,72, zal dus, bij het vinden in de steekproef van een gemiddelde $\bar{\bar{x}} = 25{,}40$, niet verworpen worden. Om een dergelijke hypothese te toetsen berekenen we met de formules (9) en (10) de grenzen van de kritieke zones:

$$\bar{\bar{x}}_l = \mu_o - 1{,}32; \quad \bar{\bar{x}}_r = \mu_o + 1{,}32.$$

Het is gemakkelijk in te zien, dat $\bar{\bar{x}} = 25{,}40$ voor alle μ's in het interval (13) tussen $\bar{\bar{x}}_l$ en $\bar{\bar{x}}_r$ ligt en dat deze μ's bij toetsing dus niet verworpen zullen worden.

IV. Met de gegevens van Voorbeeld III berekenen we vervolgens, met behulp van formule (12), de bovengrens m^* van het bij $\bar{\bar{x}} = 25{,}40$ behorende rechtseenzijdige betrouwbaarheidsinterval voor de onbekende verwachting μ ($\xi_{2\alpha} = 2{,}326$):

$$\mu < m^* = 26{,}59.$$

Een hypothese $H_o: \mu \geq \mu_o$, met μ_o kleiner dan 26,59, zal dus, bij het vinden in de steekproef van een gemiddelde $\bar{\bar{x}} = 25{,}40$, niet verworpen worden. Om zo'n hypothese te toetsen berekenen we met formule (9) de bovengrens $\bar{\bar{x}}_l$ van de linker kritieke zone:

$$\bar{\bar{x}}_l = \mu_o - 1{,}19.$$

Het is weer gemakkelijk in te zien, dat voor alle μ's kleiner dan 26,59 het steekproefgemiddelde $\bar{x} = 25{,}40$ altijd groter is dan $\bar{x}_1$ en dat deze μ's bij toetsing dus niet verworpen zullen worden.

Opmerkingen.

I. Hoewel een betrouwbaarheidsinterval geconstrueerd wordt met behulp van de toetsingstheorie, is het opstellen van zo'n interval niet equivalent met het toetsen van een hypothese. Bij het toetsen van een hypothese wordt ook nog rekening gehouden met de kans op een fout van de tweede soort, terwijl een hiermede corresponderende kans bij een betrouwbaarheidsinterval ontbreekt.

II. De omvang n van de steekproef wordt bij een betrouwbaarheidsinterval bepaald aan de hand van de toegelaten onnauwkeurigheid, dus door de toegelaten lengte van het interval (zie Memorandum S308-A7). Bij het toetsen van een hypothese wordt deze omvang bepaald door de eisen inzake de kansen α en β op fouten van de eerste, respectievelijk de tweede soort (zie Memorandum S308-A13).