

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

S 111 (M 43)

Statistische kwaliteitsbeheersing.

drs. J. de Boer



1953

Statistische kwaliteitsbeheersing

door Drs J. DE BOER

Overdruk uit Conserva 6 (1953), 165

Statistische kwaliteitsbeheersing

door Drs J. DE BOER

*Mathematisch Centrum, Amsterdam, Statistische Afdeling. Leiding: Prof. Dr D. van Dantzig
Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk*

STATISTICAL QUALITY CONTROL

The usefulness of statistical sampling as a tool for quality control is discussed and illustrated by means of some examples. Three methods to obtain a given operating characteristic (i. e. a curve where the percentage of defective items in a lot is plotted against the percentage of such lots accepted by the sampling procedure) are considered: single, double and multiple sampling.

1. Het nut van statistische monsterneming

Bij statistisch onderzoek van een partij goederen wordt een steekproef of monster genomen en op het resultaat van deze steekproef worden conclusies omtrent de gehele partij gebaseerd. Op zichzelf is dit niets bijzonders; iedereen past dit toe als hij bv. appels wil kopen en eerst één of twee appels bekijkt om de desbetreffende soort te beoordelen.

Het verschil tussen deze alledaagse en de wetenschappelijke (dat wil hier zeggen: wiskundige) statistische monsterneming is, behalve in een verschil in uitvoering, gelegen in het feit dat men bij laatstgenoemde methode kan aangeven hoe vaak de conclusies, waartoe men komt, fout zijn.

Hier stuiten we op de vraag naar het nut van monsterneming. Waarom, vraagt men misschien, zou men monsterneming toepassen, waarbij een reële kans op een foute conclusie over de kwaliteit van de partij bestaat (bv. als de partij zeer behoorlijk is, maar de steekproef toch grotendeels uit slechte exemplaren bestaat, wat altijd voor kan komen), terwijl een volledig onderzoek van alle exemplaren uit de partij een resultaat zou geven dat ten volle betrouwbaar zou zijn?

In het geval van destructieve keuring, d.w.z. dat gekeurde exemplaren verder onbruikbaar zijn, is monsterneming natuurlijk noodzakelijk. Verder ligt in het algemeen één antwoord zeer voor de hand: de uitvoering van een monsterneming is belangrijk goedkoper in tijd en geld dan die van een volledig onderzoek van de gehele partij. (We zullen steeds onderstellen dat de steekproef slechts een zeer klein gedeelte van de partij omvat.) Een ander praktisch voordeel van monsterneming is: het verkregen materiaal kan sneller bewerkt worden, zodat uit de resultaten van een steekproef vlugger conclusies getrokken kunnen worden dan uit die van een volledig onderzoek.

Tegenover deze voordelen staat, zoals gezegd, het nadeel van de kans op foute conclusies, waarvan de frequentie met een geschikte steekproefmethode (bij monsterneming van een groot aantal verschillende partijen, vgl. onder afb. 1) echter aangegeven kan worden. Dit is natuurlijk van zeer groot belang, omdat de praktische betekenis

van monsterneming staat of valt met de mogelijkheid of onmogelijkheid om grenzen voor deze fouten aan te geven.

Verder is er het bezwaar, dat de personen, belast met de uitvoering en (of) bewerking van de resultaten der monsterneming, speciaal tot dit doel geïnstrueerd moeten worden. De te geven voorschriften zijn streng, maar bijna altijd eenvoudig en onveranderlijk.

Statistische monsterneming wordt toegepast voor twee verschillende doeleinden:

1. Keuring van de kwaliteit van partijen goederen.

2. Contrôle op het fabricageproces.

We beperken ons in dit artikel tot één aspect van het eerste onderwerp, nl. het geval dat een partij door een producent geleverd wordt aan een consument. Daarvoor is echter nodig, dat we eerst wat dieper ingaan op de uitvoering van de monsterneming.

2. Het nemen van een steekproef

Het spreekt vrijwel vanzelf dat men zoveel mogelijk er naar zal streven in een te keuren partij slechts exemplaren samen te brengen die uit dezelfde partij grondstof(fen) en door hetzelfde productieproces verkregen zijn. Dit houdt ook in dat de verschillende exemplaren door hetzelfde personeel vervaardigd zijn. Kort gezegd: men zal zoveel mogelijk rekening houden met alle factoren die kunnen leiden tot kwaliteitsverschil. Maar gewoonlijk is dit om twee redenen niet geheel door te voeren:

- a) Bij een bepaalde partij kan niet altijd meer worden nagegaan welke exemplaren hetzelfde productieproces enz. doorlopen hebben;

- b) Soms is splitsing van een partij in homogene groepen van exemplaren (waarvoor dus productieproces enz. gelijk is) wel mogelijk, maar deze groepen kunnen zo klein worden dat de totale steekproef naar verhouding veel te groot zou moeten zijn, wil men een redelijke betrouwbaarheid voor de eindconclusie verkrijgen.

Algemeen kan gezegd worden: *Neem zo groot mogelijke partijen, maar houd tegelijkertijd rekening met de belangrijkste factoren die kwaliteitsverschillen kunnen veroorzaken.*

Een te keuren partij kan dan gesplitst worden in zoveel mogelijk homogene gedeelten naar bovengenoemde factoren en uit deze gedeelten worden dan steekproeven genomen, waarvan de grootten zich verhouden als die van de bijbehorende gedeelten van de partij. Als bv. de partij gesplitst is in drie gedeelten, die resp. de helft, één derde en één zesde van het geheel zijn, dan moet van de totale steekproef ook resp. de helft, één derde en één zesde uit deze drie gedeelten genomen worden.

Het belangrijkste is echter, dat alle exemplaren uit een partij of een gedeelte van een partij dezelfde kans moeten hebben om in de steekproef te komen. Hieraan zou eigenlijk voldaan moeten worden door een loting, waarbij de exemplaren van de partij genummerd worden en de uitgelote nummers de steekproef vormen. Soms gebruikt men hiervoor tabellen van getallen die volgens een lotingsprocédé verkregen zijn en dus niet meer „geproduceerd” behoeven te worden. In de praktijk volstaat men vaak met minder precieze methoden, die, naar men dan aanneemt, toch wel ongeveer op hetzelfde neerkomen.

3. „Producenten risico”, „consumenten risico” en karakteristiek

We nemen aan dat een te verhandelen partij goedgekeurd moet worden op het percentage defecte exemplaren dat er in voorkomt. Hierbij dient bedacht te worden dat we slechts één bepaald defect beschouwen; is er sprake van meerdere, dan moet voor elk afzonderlijk beslist worden welk type van monsterneming in aanmerking komt.

De steekproef wordt genomen en al naar gelang het aantal defecte exemplaren daarin kleiner of groter is dan een van te voren vastgesteld getal, wordt de partij goedgekeurd of afgekeurd. Dit getal wordt bepaald door „producenten risico” en „consumenten risico”. De betekenis van deze termen is de volgende:

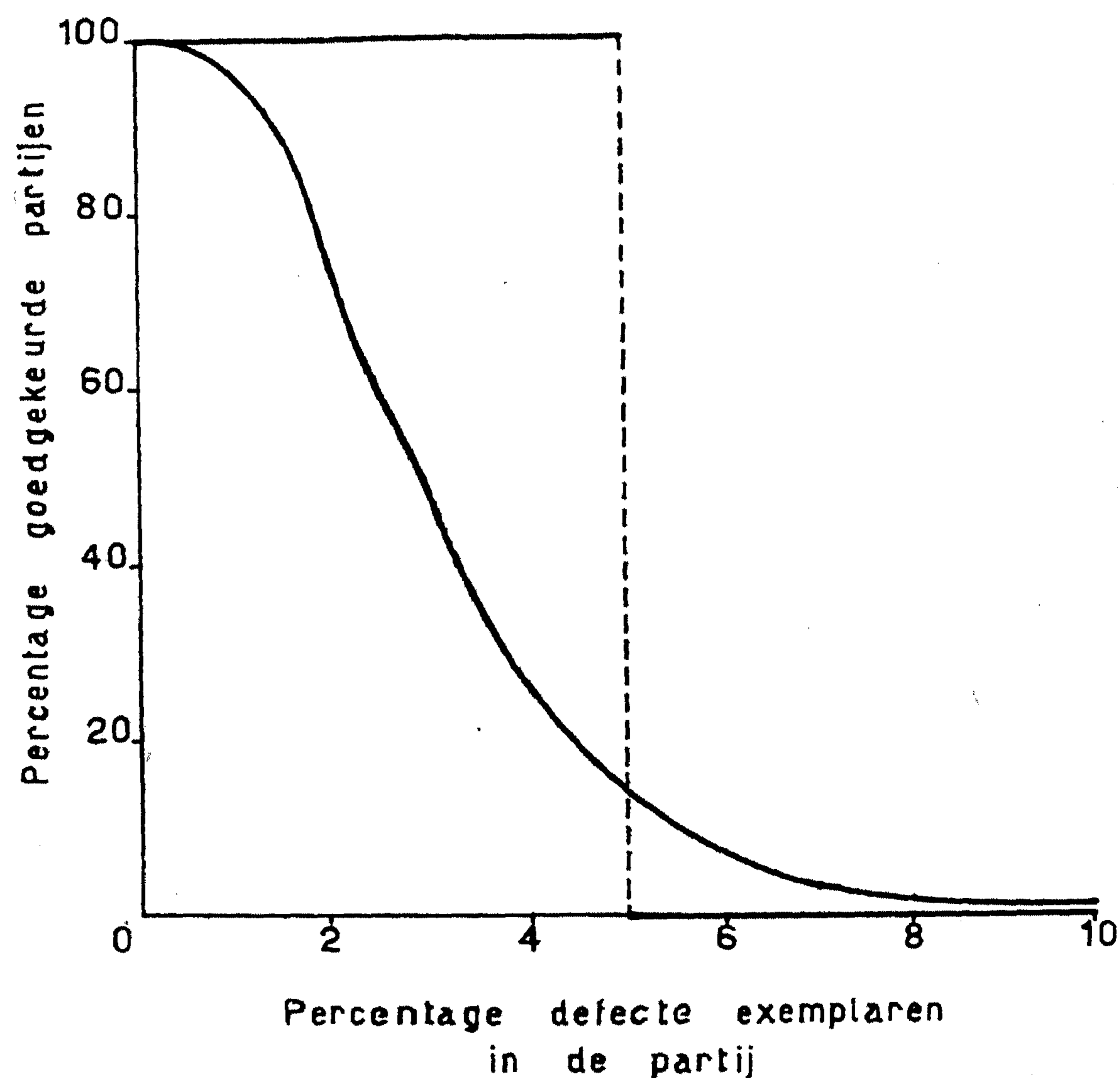
De producent kan meestal twee percentages opgeven (bv. 1% en 4%), zò, dat hij van de monsterneming eist, dat een partij die niet meer dan 1% defecten bevat, slechts in hoogstens 4% van de gevallen waarin zo'n partij zich voordoet, afgekeurd zal worden. Deze 4% wordt dan „producenten risico” genoemd (ook als hij een andere grens dan 1% stelt). Men begrijpt waarom: het is het risico voor de producent dat een partij, die z. i. zeer behoorlijk is (dat wil in bovengenoemd geval dus zeggen: een partij die niet meer dan 1% defecten bevat), toch afgekeurd wordt.

Evenzo kan de afnemer de eis stellen dat een z. i. slechte partij, bv. met 6% of meer defecten, slechts in bv. hoogstens 7% van de gevallen waarbij zo'n partij voorkomt, goedgekeurd

wordt. In dat geval heet deze 7% dan „consumenten risico”.

Deze twee eisen, tezamen met de grootte van de steekproef, bepalen voor elk type monsterneming het verloop van de zg. karakteristiek. Deze geeft bij genoemde monsterneming het verband aan tussen het percentage defecte exemplaren in een partij en de kans dat een partij met dat bepaalde percentage defecte exemplaren goedgekeurd zal worden.

Een karakteristiek die aan bovengenoemde twee eisen van producenten en consumenten risico voldoet, zou er bij een bepaalde steekproefgrootte bv. kunnen uitzien als de vloeiende lijn in afb. 1. Opgemerkt dient te worden dat in § 1 de verhouding tussen de grootten van partij en steekproef zeer groot ondersteld is, zodat de omvang van de partij geen invloed heeft op het verloop van de karakteristiek.



Afb. 1. Keuringskarakteristieken

Hieruit ziet men dus dat dan bv. een partij met 4% defecte exemplaren 25% kans heeft om goedgekeurd te worden. Dit betekent niet dat er precies 25 van 100 dergelijke partijen goedgekeurd zullen worden, maar wel dat het percentage goedgekeurde onder zulke partijen steeds dichterbij 25 zal komen te liggen als het aantal partijen stijgt, dus dat „op den langen duur” ongeveer 25% van dit soort partijen goedgekeurd zal worden.

Men merke op dat een ideale karakteristiek uit twee horizontale stukken zou bestaan, zoals bv. de dikke lijnen in de figuur, verbonden door de stippellijn. (Deze ideale karakteristiek dient hier slechts als voorbeeld en staat niet in verband met de eerste.) Immers deze geeft aan dat partijen

met (in dit geval) 5% of minder defecte exemplaren altijd, en partijen met meer dan 5% defecte exemplaren nooit goedgekeurd zullen worden. Maar men begrijpt dat hiervoor een volledig onderzoek zonder fouten nodig is i. p. v. monsterneming, die altijd tot foute conclusies kan leiden. Omdat men een steekproef neemt, is een ideale karakteristiek niet te bereiken en om dezelfde reden zijn de termen „producenten risico” en „consumenten risico” ontstaan die, in de taal van de karakteristiek overgebracht, betekenen dat men zich met een minder steil verloop van de karakteristiek tevreden stelt. Hoe kleiner de twee risico's bij twee vaste percentages defecten, hoe steiler de karakteristiek loopt. Een steiler verloop valt slechts te bereiken door de steekproef groter te nemen.

Vaak gebruikt men andere namen voor dezelfde grootheden of andere grootheden dan de twee risico's. We zullen ons hier echter tot deze suggestieve termen beperken.

4. De types van monsterneming

Om aan de onder 3 behandelde eisen van producenten en consumenten risico te voldoen, kunnen we verschillende steekproefmethodes toepassen.

We onderscheiden de volgende drie types van monsterneming:

- Enkelvoudige steekproefmethode
- Dubbele steekproefmethode
- Meervoudige steekproefmethode.

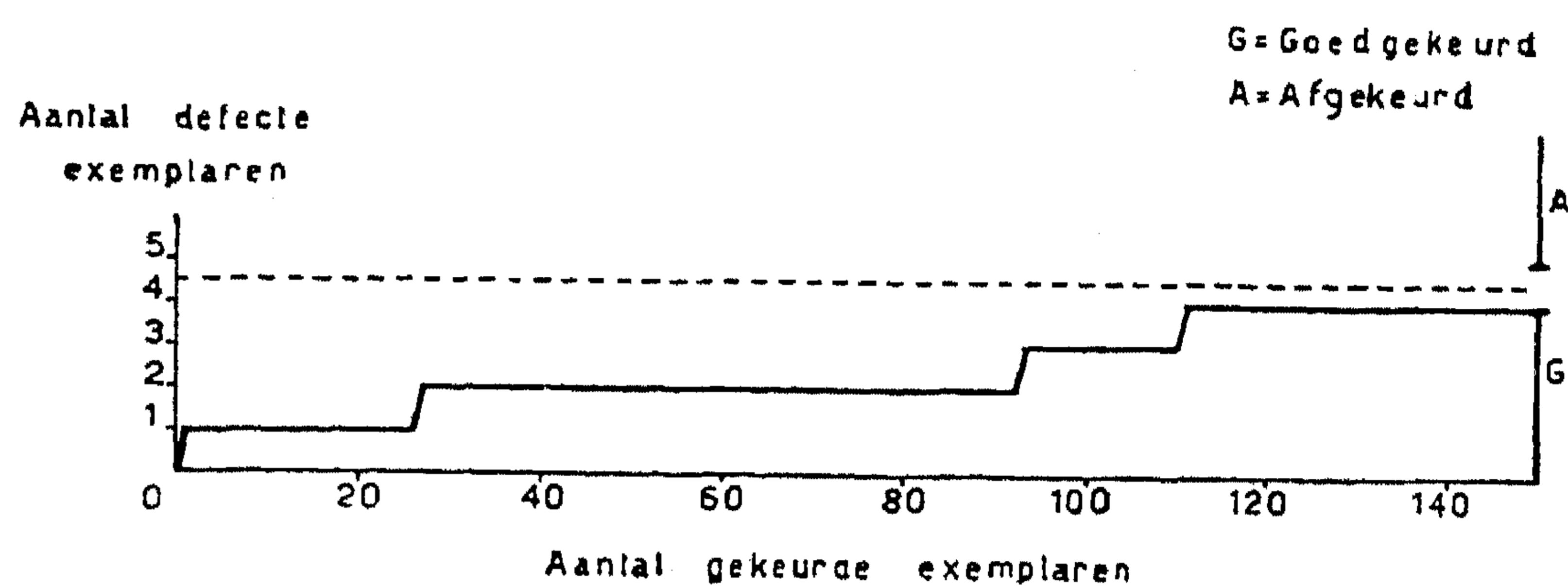
Soms gebruikt men een zgn. sequente methode, waarbij de grootte van de steekproef in het geheel niet van tevoren vaststaat en men stapsgewijze te werk gaat, maar deze mogelijkheid, die nog iets verder gaat dan de meervoudige steekproefmethode, zullen we hier buiten beschouwing laten. Ook gaan we niet in op de wiskundige achtergrond van de drie types, maar beperken ons tot de verschillen in de uitvoering van de monsterneming bij de types.

Bij elk type behoren zgn. *goedkeuringsgetallen* en *afkeuringsgetallen*. De betekenis hiervan is deze:

Ligt het aantal defecte exemplaren in een steekproef beneden het goedkeuringsgetal G (resp. boven het afkeuringsgetal A) of is het er gelijk aan, dan wordt de partij goedgekeurd (resp. afgekeurd).

Bij een *enkelvoudige steekproefmethode* wordt één steekproef genomen die relatief groot is ten opzichte van de steekproeven bij de andere types. Het afkeuringsgetal is in dit geval één meer dan het goedkeuringsgetal ($A = G + 1$), zodat men dus altijd direct na telling van het aantal defecte exemplaren in de steekproef een beslissing over

goed- of afkeuring van de partij kan nemen. In een schema waar het aantal achtereenvolgens gekeurde exemplaren van de steekproef uitgezet is tegen het aantal defecten daaronder, kan men een zg. „keuringsweg” aangeven, d. i. een gebroken lijn die ontstaat door bij elk defect exemplaar een stap schuin naar rechts boven en bij elk niet-defect exemplaar een stap horizontaal naar rechts te maken. Het hieronder gegeven voorbeeld zal dit duidelijk maken (zie afb. 2).

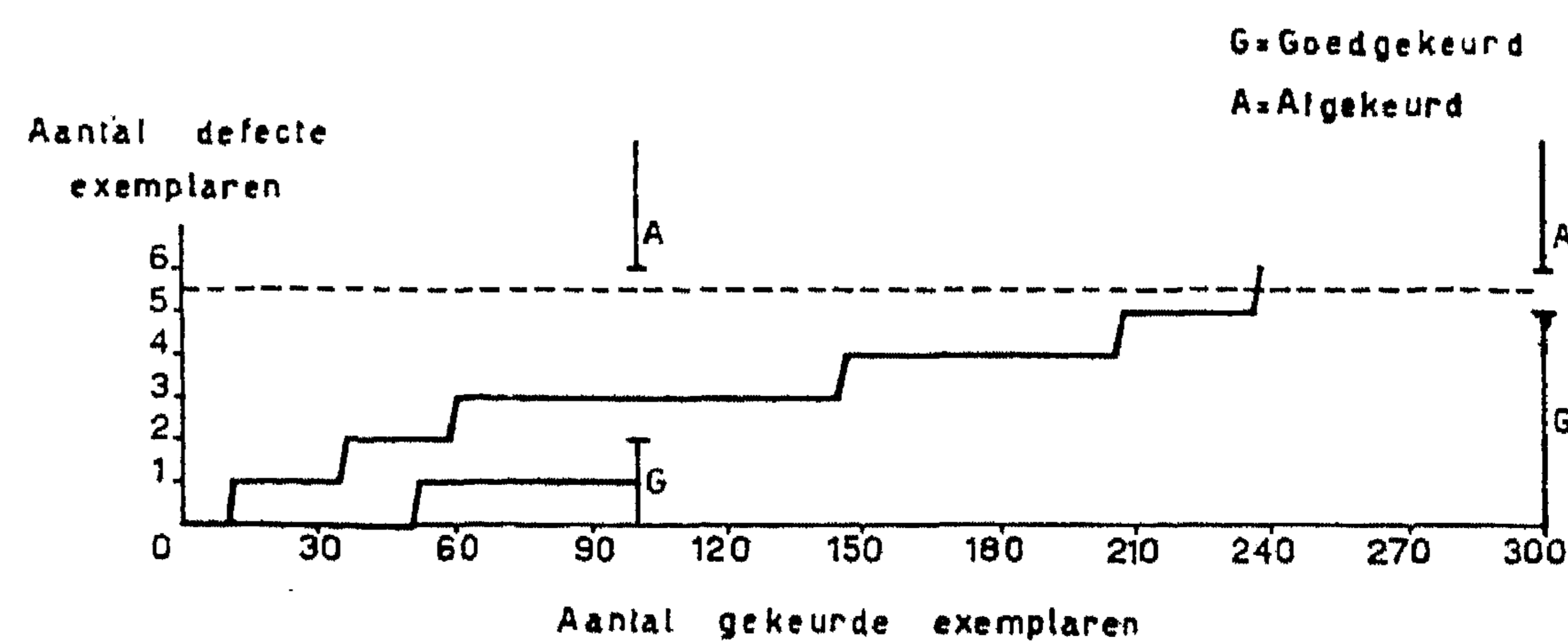


Afb. 2. Keuring met enkelvoudige steekproefmethode

In dit schema hebben we een steekproef van 150 exemplaren met goedkeuringsgetal 4 en afkeuringsgetal 5. We vinden defecte exemplaren bij de 1e, 26e, 92e en 110e waarneming, zodat de keuringsweg beneden de stippellijn blijft, d.w.z. er zijn niet meer dan 4 defecten en de partij wordt goedgekeurd.

Het spreekt vanzelf dat de keuring stopgezet kan worden zodra de keuringsweg boven de stippellijn uitkomt. Dan is het aantal defecte exemplaren immers minstens 5, zodat de partij afgekeurd kan worden.

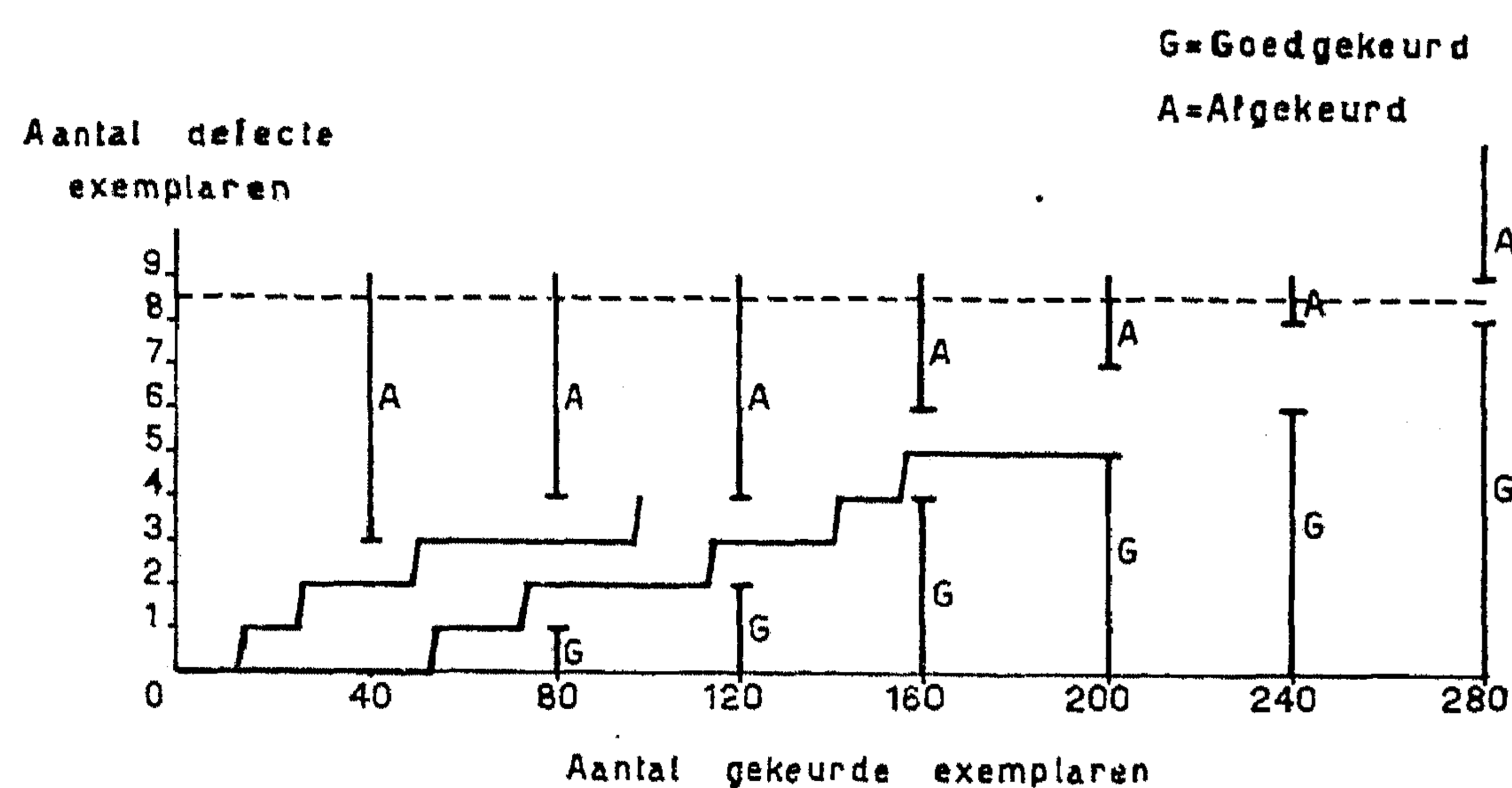
Bij een dubbele steekproef wordt een kleinere steekproef genomen, maar nu is $A > G + 1$. Ligt het aantal defecte exemplaren tussen A en G in (en alléén dan), dan wordt een tweede steekproef genomen. De getallen A en G voor de twee steekproeven *samen* (die in de regel zullen afwijken van de A en G voor de eerste steekproef) verschillen dan één, zodat een beslissing kan worden genomen. De naam „dubbele steekproefmethode” is dus in zoverre misleidend, dat alleen in een bepaald geval inderdaad twee steekproeven genomen worden. In afb. 3 is een voorbeeld van een schema voor een dubbele steekproef getekend met 2 keuringswegen erin.



Afb. 3. Keuring met dubbele steekproefmethode

Hier is een dubbele steekproef opgesteld: eerst 100 waarnemingen en dan eventueel (nl. als er meer dan 2 en minder dan 6 defecte exemplaren bij die 100 zijn) nog eens 200. Voor de eerste steekproef is dus $A = 6$ en $G = 2$; voor de twee steekproeven samen is $A = 6$ en $G = 5$. Is het aantal defecten bij de eerste 100 niet meer dan 2 (zoals bij de ene keuringsweg), dan wordt de partij al na deze eerste steekproef goedgekeurd. Is het aantal defecten 6 of meer, dan is er ook geen tweede steekproef nodig: de keuringsweg komt boven de stippellijn uit en de partij wordt afgekeurd. De andere keuringsweg is een voorbeeld van een geval waarin de tweede steekproef nodig is. Omdat deze bij de 24e waarneming de stippellijn overschrijdt, wordt op dat moment de keuring beëindigd en de partij afgekeurd.

In het geval van een meervoudige steekproef ten slotte wordt een nog kleinere eerste steekproef genomen. Ook hier is $A > G + 1$, zodat soms een tweede nodig is. Dan geldt weer $A > G + 1$ voor de getallen A en G die behoren bij de twee steekproeven samen. Kan nu geen beslissing genomen worden, dan volgt een derde steekproef, enz. Vaak geldt pas na de zevende of achtste steekproef: $A = G + 1$, maar meestal heeft men al veel eerder kunnen beslissen (zie afb. 4).



Afb. 4. Keuring met meervoudige steekproefmethode

De hier weergegeven meervoudige steekproef bestaat dus uit 7 steekproeven, elk van 40 waarnemingen. De goedkeuringsgetallen zijn na de 2e, 3e, ..., 7e steekproef resp. 1, 2, 4, 5, 6, 8 (na de 1e steekproef kan de partij nog niet goedgekeurd worden); de afkeuringsgetallen na de 1e, 2e, ..., 7e steekproef zijn resp.: 3, 4, 4, 6, 7, 8, 9. De keuringswegen corresponderen met steekproeven die resp. „in de loop van de 3e steekproef” tot afkeuring en na de 5e steekproef tot goedkeuring van de desbetreffende partij leiden.

We kunnen de drie bovengenoemde voorbeelden in een tabel (zie tabel I) samenvatten, zoals gedaan is op blz. 313 van het boek „Sampling Inspection”¹⁾. Dit boek bevat een zeer uitgebreide verzameling tabellen, geschikt voor de praktische statistische monsterneming.

1) Zie literatuuropgave.

TABEL I
Drie steekproefschema's

type monsterneming	steekproef	steekproefgrootte	gecombineerde steekproeven		
			grootte	goedk. getal	afk. getal
enkelvoudig		150	150	4	5
dubbel	eerste	100	100	2	6
	tweede	200	300	5	6
meervoudig	eerste	40	40	x)	3
	tweede	40	80	1	4
	derde	40	120	2	4
	vierde	40	160	4	6
	vijfde	40	200	5	7
	zesde	40	240	6	8
	zevende	40	280	8	9

x) Hier is goedkeuring bij de eerste steekproef nog niet toegestaan.

Deze drie monsternemingen hebben nagenoeg dezelfde karakteristiek, nl. die van afb. 1. Ze behoren dus bij een producenten risico van 4% (voor 1% defecten) en een consumenten risico van 7% (voor 6% defecten). Omdat de partij zeer groot ondersteld is t.o.v. de steekproef (vgl. § 1), speelt de omvang van de partij geen rol bij deze tabel.

We zien hier de bovengenoemde verhouding van de groottes der eerste steekproeven bij de verschillende types.

Bij bepaalde aan de karakteristiek gestelde eisen (producenten en consumenten risico) zijn monsternemingen van elk van deze types te vinden die aan deze eisen voldoen. In het genoemde boek is bij elk tweetal eisen een schema gevoegd, zoals het hierboven gegevene. De drie types steekproeven van één tabel hebben dan een vrijwel gelijk verlopende karakteristiek, die ook bij de tabel gegeven wordt en de keuze tussen de types wordt dan bepaald door factoren als:

1. kosten van de monsterneming,
2. informatie over de kwaliteit, die de verschillende types geven,
3. moeilijkheid van instructie van het personeel, enz.

We kunnen hier niet dieper op ingaan; in het ene geval zal bv. de eerste factor belangrijker zijn dan de andere; in andere gevallen kan het weer heel anders zijn. Daarom zijn geen algemene regels voor de keuze van het type monsterneming te geven. Meestal zullen de kosten van de monsterneming de belangrijkste rol spelen. In dit verband geldt de overweging, dat het totale aantal geïnspecteerde exemplaren gemiddeld afneemt, naarmate men van een meer samengestelde steekproefmethode gebruik maakt.

Samenvatting

Statistische monsterneming als kwaliteitsbeheersing verdient veelal de voorkeur boven volledig onderzoek van een partij omdat het minder kostbaar is. Bij destructieve keuring is het de enige methode die in aanmerking komt.

Bij het nemen van een steekproef moet elk exemplaar dezelfde kans hebben om in de steekproef te komen.

De grootte van producenten risico en consumenten risico bepalen bij een bepaalde steekproefmethode de omvang van de steekproef en het verloop van de karakteristiek, d. i. de kromme die het verband aangeeft tussen de kwaliteit van een partij (d.w.z. het percentage defecte exemplaren erin) en de kans op goedkeuring van de partij.

Te kleine risico's zouden te grote steekproeven of zelfs keuring van de gehele partij vergen.

Op grond van de kosten van monsterneming en andere van belang zijnde factoren wordt uitgemakt welk van de drie types monsterneming toegepast zal worden: de enkelvoudige, de dubbele of de meervoudige steekproefmethode.

Naar aanleiding van het resultaat van de steekproef wordt de partij goedgekeurd resp. afgekeurd al naar gelang het aantal defecte exemplaren in de steekproef kleiner resp. groter is dan of gelijk aan het bijbehorende goedkeuringsgetal resp. afkeuringsgetal.

Literatuur:

Statistical Research Group, Columbia University, Sampling Inspection, Mc Graw Hill Book Company, Inc., New York, 1948.

H. F. Dodge en H. G. Romig: Sampling Inspection Tables, John Wiley & Sons, Inc., New York, 3e druk, 1946.

H. F. Dodge en H. G. Romig: Sequential analysis of statistical data: Applications, Columbia University Press, 1945.