SA

stichting
mathematisch
centrum

MC

SA

AFDELING MATHEMATISCHE STATISTIEK SC 22/71 MAART

J. OOSTERHOFF
SYLLABUS VAN HET COLLEGE BIOMETRIKA

2e boerhaavestraat 49 amsterdam

I n h o u d

Hoofdstuk 1. Inleiding in de Wiskunde

§1. De getallenrechte

Uit de schoolwiskunde zijn de reële getallen bekend. We kunnen deze eenvoudig voorstellen met behulp van de getallenrechte. Hiertoe tekenen we een rechte lijn en plaatsen ergens op deze lijn een punt: het nulpunt, dat we aanduiden met 0. Vervolgens kiezen we een lengte-eenheid, b.v. 1 cm, en zetten rechts van 0 de gehele positieve getallen *) 1, 2, 3, ... uit op eenheidsafstanden van elkaar, en links van 0 de gehele negatieve getallen -1, -2, -3, ... op eenheidsafstanden van elkaar. Zie fig. 1.1. De breuken en de irrationale getallen (zoals $\frac{1}{2}$, $\sqrt{2}$, π) kunnen we ook op de getallenrechte aangeven.

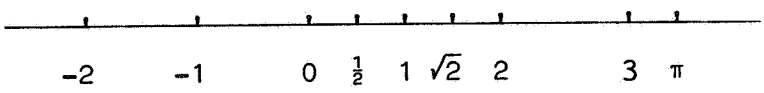


fig. 1.1. De getallenrechte

Met elk punt op de rechte correspondeerd één reëel getal; omgekeerd behoort bij elk reëel getal één punt op de getallenrechte.

De afstand van een punt op de getallenrechte, dat correspondeert met een reëel getal a, tot het nulpunt 0 noemen we de absolute waarde van het getal a en duiden we aan met $|a|$. Uit deze definitie volgt, dat

$$|a| = \begin{cases} a & \text{als } a \geq 0 \\ -a & \text{als } a < 0. \end{cases}$$

We vermelden de volgende rekenregels voor absolute waarden:

$$|a+b| \leq |a| + |b|, \quad |a-b| \leq |a| + |b|,$$
$$|a.b| = |a| \,.\, |b|,$$
$$\left|\frac{a}{b}\right| = \frac{|a|}{|b|} \quad (b \neq 0).$$

*) De gehele positieve getallen noemt men vaak de natuurlijke getallen.

Het is eenvoudig deze eigenschappen te verifiëren.

Een open interval is de verzameling van alle reële getallen, die liggen tussen twee vaste getallen a en b. Is $a < b$, dan geven we dit open interval aan met (a,b). Een getal x ligt dus in (a,b), als $a < x < b$. Met een open interval (a,b) correspondeert een lijnstuk op de getallenrechte; de randpunten a en b behoren er niet toe.

Een gesloten interval is de verzameling van alle reële getallen, die liggen tussen twee vaste getallen a en b, doch met inbegrip van a en b zelf. Is $a < b$, dan duiden we dit gesloten interval aan met $[a,b]$. Een getal x ligt dus in $[a,b]$, als $a \leq x \leq b$. Ook met $[a,b]$ correspondeert een lijnstuk op de getallenrechte; de randpunten a en b behoren er nu wel toe.

Worden verzamelingen reële getallen gekarakteriseerd door éénzijdige ongelijkheden, $x < a$, $x \leq a$ of $x > a$ of $x \geq a$, dan spreekt men van een onbegrensd interval. Men geeft deze intervallen vaak aan met $(-\infty,a)$ resp. $(-\infty,a]$ resp. (a,∞) resp. $[a,\infty)$. De verzameling van alle reële getallen geeft men vaak aan met $(-\infty,\infty)$.

Met een omgeving van een reëel getal a bedoelt men een open intervalletje, dat a bevat, b.v. $(a-h, a+h)$ met positieve h.

§2. Bewijzen met volledige inductie

Zij E een eigenschap, waarvan we willen nagaan of deze voor alle natuurlijke getallen n waar is.

Voorbeeld 1.1. De som van de eindige meetkundige reeks is volgens de schoolwiskunde

$$(1.1) \qquad 1 + r + r^2 + \ldots + r^{n-1} = \frac{1-r^n}{1-r} \quad (r \neq 1).$$

Dit is een eigenschap, die we voor elk natuurlijk getal n willen aantonen.

We kunnen in dergelijke gevallen vaak gebruik maken van bewijzen met volledige inductie, die berusten op de volgende stelling.

Stelling 1.1. Als

(i) de eigenschap E juist is voor $n = 1$, en

(ii) de juistheid van E voor $n + 1$ is af te leiden uit de juistheid van

E voor n (voor alle natuurlijke n),

dan is de eigenschap E juist voor alle natuurlijke getallen n.

Bewijs:

Het bewijs berust op de eigenschap, dat elke verzameling van natuurlijke getallen een kleinste bezit. Stel dat E niet juist is voor zekere n; dan is er dus een verzameling (van één of meer) natuurlijke getallen waarvoor E niet juist is. Deze verzameling heeft een kleinste element, zeg k (k>1). Dan is E dus wel juist voor k - 1. Maar gegeven is, dat als E juist is voor k - 1, we kunnen aantonen dat E ook juist is voor k. Dit levert een tegenspraak op, zodat E juist is voor elk natuurlijk getal.

Voorbeeld 1.1. (vervolg) We bewijzen formule (1.1) nu met volledige inductie. Voor n = 1 hebben we: $1 = \frac{1-r}{1-r}$, zodat voor n = 1 de formule juist is. Zij de formule juist voor n. Dan is

$$1 + r + r^2 + \ldots + r^{n-1} + r^n = \frac{1-r^n}{1-r} + r^n = \frac{1-r^n}{1-r} + \frac{r^n-r^{n+1}}{1-r} = \frac{1-r^{n+1}}{1-r}$$

zodat de formule ook juist is voor n + 1. Aan de voorwaarden van Stelling 1.1 is nu voldaan, zodat (1.1) juist is voor elk natuurlijk getal n.

Voorbeeld 1.2. Gevraagd te bewijzen

$$(1.2) \qquad \frac{1}{1.2} + \frac{1}{2.3} + \frac{1}{3.4} + \ldots + \frac{1}{n(n+1)} = 1 - \frac{1}{n+1} \quad \text{voor } n = 1, 2, 3, \ldots .$$

Voor n = 1 leveren beide leden van (1.2) juist $\frac{1}{2}$ op, zodat de formule dan juist is. Stel dat (1.2) juist is voor n. Dan is (1.2) ook juist voor n + 1, immers

$$\frac{1}{1.2} + \frac{1}{2.3} + \ldots + \frac{1}{n(n+1)} + \frac{1}{(n+1)(n+2)} = 1 - \frac{1}{n+1} + \frac{1}{(n+1)(n+2)} = 1 - \frac{1}{n+2}.$$

Volgens Stelling 1.1 is (1.2) nu juist voor elk natuurlijk getal n.
In dit geval kunnen we (1.2) overigens eenvoudig rechtstreeks bewijzen, gebruik maken van $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$ voor elk natuurlijk getal k. Het rechterlid van (1.2) kunnen we nu ook schrijven als

$$(\frac{1}{1} - \frac{1}{2}) + (\frac{1}{2} - \frac{1}{3}) + (\frac{1}{3} - \frac{1}{4}) + \ldots + (\frac{1}{n} - \frac{1}{n+1}) = 1 - \frac{1}{n+1}.$$

Voorbeeld 1.3. Toon aan dat

$$(1.3) \qquad 1 + 2^2 + 3^2 + \ldots + n^2 = \frac{1}{6}\, n(n+1)(2n+1) \text{ voor alle } n = 1, 2, 3, \ldots .$$

Voor n = 1 zijn beide leden gelijk aan 1. Als de formule juist is voor n, dan is

$$1 + 2^2 + 3^2 + \ldots + n^2 + (n+1)^2 = \frac{1}{6}\, n(n+1)(2n+1) + (n+1)^2 =$$

$$= \frac{1}{6}(n+1)\,\{n(2n+1) + 6(n+1)\} = \frac{1}{6}(n+1)(n+2)(2n+3)$$

$$= \frac{1}{6}(n+1)\,\{(n+1) + 1\}\,\{2(n+1) + 1\},$$

zodat (1.3) dan ook juist is voor n + 1. Hieruit volgt weer het gevraagde m.b.v. Stelling 1.1.

Opmerking. Ter verkorting van de notatie maakt men vaak gebruik van het sommatiesymbool $\sum$. We schrijven

$$a_1 + a_2 + \ldots + a_n = \sum_{i=1}^{n} a_i,$$

waarmee we de som van alle a_i's bedoelen met indices i van 1 t/m n.
Zo zouden we de linkerleden van (1.1), (1.2) en (1.3) kunnen schrijven als

$$\sum_{i=0}^{n-1} r^i, \quad \sum_{i=1}^{n} \frac{1}{i(i+1)}, \quad \sum_{i=1}^{n} i^2.$$

Met r^0 bedoelt men steeds 1 (alle reële r).

Teneinde het gebruik van worteltekens en breuken zoveel mogelijk te vermijden schrijft men ook vaak

$$\sqrt[n]{r^m} = r^{m/n}, \text{ bijv. } \sqrt{r} = r^{\frac{1}{2}}, \; \sqrt[4]{r^3} = r^{\frac{3}{4}} \text{ etc.}$$

en

$$\frac{1}{\sqrt[n]{r^m}} = r^{-m/n}, \text{ bijv. } \frac{1}{r^2} = r^{-2}, \; \frac{1}{r\sqrt{r}} = r^{-3/2} \text{ etc.}$$

Deze definities van machten van r zijn in overeenstemming met de regel,

dat men bij het vermenigvuldigen van verschillende machten van eenzelfde getal de exponenten optelt.

§3. Rijen. Limieten

Zij a_1, a_2, a_3, ... een oneindig voortlopende rij reële getallen. Ter verkorting van de notatie geven we zo'n rij vaak weer met $\{a_n\}$, waarin a_n het n-de element van de rij voorstelt.

Beschouw de rij $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$ of kortweg $\{2^{-n}\}$. Naarmate n groter wordt, nadert 2^{-n} tot nul zonder ooit precies nul te worden.
Kiezen we een willekeurig klein interval $(-\varepsilon,\varepsilon)$ om nul ($\varepsilon>0$), dan is er altijd een index N aan te geven, zodat alle getallen 2^{-n} met $n > N$ binnen dit interval liggen. We noemen nul de limiet van de rij $\{2^{-n}\}$ en schrijven $\lim_{n\to\infty} 2^{-n} = 0$. Algemener definiëren we:

Definitie 1.1. Het getal a heet de limiet van de rij $\{a_n\}$, als bij elk positief getal ε een index N bestaat, zodat

$$|a_n - a| < \varepsilon \quad \text{voor alle } n > N.$$

Notatie: $\lim_{n\to\infty} a_n = a$.

Deze definitie houdt dus in, dat a limiet is van de rij $\{a_n\}$, als alle getallen a_n, die ver genoeg in de rij staan (voorbij een zekere index N), minder dan ε van a verschillen. De index N in definitie 1.1 zal in 't algemeen van ε afhangen (hoe verder in de rij, des te dichter liggen de getallen a_n bij a).

Een rij $\{a_n\}$ met een limiet heet convergent. Heeft een rij géén limiet, dan heet de rij divergent.

Voorbeeld 1.4. De rij $1, 2, 3, \ldots$ is divergent.

Voorbeeld 1.5. De rij $1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \frac{1}{5}, \ldots$ is convergent met limiet nul. Een dergelijke rij, waarvan de elementen afwisselend positief en negatief zijn, heet een alternerende rij.

Voorbeeld 1.6. De rij $1, -1, 1, -1, 1, \ldots$ is divergent. Ook deze rij is alternerend.

Een fundamentele stelling over convergentie van rijen is

Stelling 1.2. (Convergentiecriterium van CAUCHY). De rij $\{a_n\}$ is d.e.s.d.*) convergent, als voor elke positieve ε een index N bestaat zodanig, dat

$$|a_n - a_m| < \varepsilon \quad \text{voor alle } n, m > N.$$

Het bewijs van deze stelling laten we achterwege. Merk op, dat de limiet a van de rij in de formulering van de stelling niet optreedt. Men kan dus tot convergentie van een rij besluiten zonder de limiet te kennen.

Men noemt een rij $\{a_n\}$ monotoon stijgend als $a_1 < a_2 < a_3 < \ldots$ en monotoon dalend als $a_1 > a_2 > a_3 > \ldots$. De rij $\{a_n\}$ heet naar boven begrensd als er een getal C bestaat zodat $a_n < C$ voor alle n, en naar beneden begrensd als er een getal C bestaat zodat $a_n > C$ voor alle n.

Als een rij $\{a_n\}$ monotoon stijgt, kunnen de elementen onbeperkt toenemen dan wel naderen tot een limiet, in welk geval de rij naar boven begrensd is. Als $\{a_n\}$ monotoon daalt, nemen de elementen of wel onbeperkt af of wel ze naderen tot een limiet, in welk geval de rij naar beneden begrensd is. Dit wordt uitgedrukt door de volgende stelling.

Stelling 1.3. Een monotoon stijgende rij is d.e.s.d. convergent, als ze naar boven begrensd is. Een monotoon dalende rij is d.e.s.d. convergent, als ze naar beneden begrensd is.

Een formeel bewijs van deze stelling geven we niet. Merk op, dat ook in deze stelling de limiet van de rij niet genoemd wordt.

Tenslotte noemen we nog enige bekende eigenschappen van limieten. Als de rij $\{a_n\}$ limiet a en de rij $\{b_n\}$ limiet b heeft, dan heeft de rij

$\{a_n + b_n\}$ limiet a + b,

$\{a_n - b_n\}$ limiet a - b,

$\{a_n . b_n\}$ limiet a . b,

$\{c . a_n\}$ limiet c . a (c is constante),

$\{\frac{a_n}{b_n}\}$ limiet $\frac{a}{b}$ (mits alle $b_n \neq 0$ en $b \neq 0$).

*) De uitdrukking "dan en slechts dan" zullen we steeds afkorten tot d.e.s.d.

Voorbeeld 1.7. Beschouw de rij 1, r, r^2, r^3, ... of kortweg $\{r^n\}$. Als $|r| < 1$, dan is de rij convergent met limiet nul; als $|r| > 1$, dan is de rij divergent. Om dit aan te tonen beschouwen we de rij van absolute waarden 1, $|r|$, $|r|^2$, $|r|^3$, ... Als $|r| < 1$ is, dan is deze rij monotoon dalend en naar beneden begrensd door nul, zodat volgens Stelling 1.3 de rij convergent is; noem de limiet a.
De rij $|r|$, $|r|^2$, $|r|^3$, ... heeft dezelfde limiet a, maar volgens de bovenstaande vierde eigenschap ook limiet $|r|a$. Dit kan alleen als $a = 0$. Maar als $\lim_{n\to\infty} |r|^n = 0$, dan is ook $\lim_{n\to\infty} r^n = 0$. Indien $|r| > 1$, dan neemt $|r|^n$ met toenemende n onbegrensd toe, zodat $\{|r|^n\}$ en ook $\{r^n\}$ divergeert.

§4. Oneindige reeksen

Uit de schoolwiskunde is bekend, dat de (oneindige) meetkundige reeks $1 + r + r^2 + r^3 + \ldots$ som $\frac{1}{1-r}$ heeft, mits $|r| < 1$. Wat bedoelen we echter met de som van een oneindige reeks? Het is onmogelijk oneindig veel termen op te tellen; we zouden hiermee nooit klaar komen. Beschouw eens een bijzonder geval: $r = \frac{1}{2}$, en schrijf de rij eindige sommen op:

$$1$$
$$1 + \frac{1}{2}$$
$$1 + \frac{1}{2} + \frac{1}{4} = 1 + \frac{3}{4}$$
$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = 1 + \frac{7}{8}$$
$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = 1 + \frac{15}{16}, \text{ etc.}$$

We zien dat deze sommen steeds dichter naderen tot het getal 2 naarmate we meer termen van de reeks meenemen. Het is daarom intuïtief aantrekkelijk 2 de som van de reeks te noemen. Ook op een andere wijze kan men dit aanvaardbaar maken. Immers vermenigvuldig alle termen van de reeks eens met 2, dan vinden we

$$2(1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots) = 2 + (1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots),$$

in overeenstemming met $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = 2$.
Algemeen definiëren we

Definitie 1.2. Zij $a_1, a_2, a_3, \ldots$ een rij getallen en zij $S_1 = a_1$, $S_2 = a_1 + a_2$, $S_3 = a_1 + a_2 + a_3, \ldots$. Men noemt de rij $\{S_n\}$ de rij van partiële sommen van de reeks $a_1 + a_2 + a_3 + \ldots$. Als de rij $\{S_n\}$ convergeert met limiet S, dan noemt men de reeks $a_1 + a_2 + a_3 + \ldots$ convergent met som S. Is de rij $\{S_n\}$ divergent, dan heet ook de reeks divergent.

Notatie: $\sum_{n=1}^{\infty} a_n = S$ (als de reeks convergeert).

Beschouw wederom de meetkundige reeks $1 + r + r^2 + \ldots$. Als $r \neq 1$, dan is $S_n = 1 + r + r^2 + \ldots + r^{n-1} = \frac{1-r^n}{1-r}$, zie voorbeeld 1.1. De rij van partiële sommen heeft dus de gedaante $\{\frac{1-r^n}{1-r}\}$. Indien $|r| < 1$, dan is $\lim_{n\to\infty} r^n = 0$ (zie voorbeeld 1.7), zodat in dat geval $\lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{1-r^n}{1-r} = \frac{1}{1-r}$.
Indien $|r| > 1$, dan is de rij $\{r^n\}$ en dus ook de rij $\{S_n\}$ divergent, zodat de reeks dan divergeert.
Is $r = 1$, dan heeft de reeks de gedaante $1 + 1 + 1 + 1 + \ldots$ en is ten duidelijkste divergent. Als $r = -1$, dan heeft de reeks de gedaante $1 - 1 + 1 - 1 + - \ldots$; de rij van partiële sommen is dan $1, 0, 1, 0, 1, \ldots$, een divergente rij, zodat ook nu de reeks divergeert.

Conclusie: de meetkundige reeks $1 + r + r^2 + \ldots$ is convergent met som $\frac{1}{1-r}$ als $|r| < 1$, en is divergent als $|r| \geq 1$.

Voorbeeld 1.8. De reeks $\frac{1}{1.2} + \frac{1}{2.3} + \frac{1}{3.4} + \ldots$ is convergent met som 1. Immers de rij van partiële sommen heeft volgens voorbeeld 1.2 de gedaante $\{1 - \frac{1}{n+1}\}$ en convergeert naar de limiet 1.

Voorbeeld 1.9. De harmonische reeks $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots$ is divergent. Immers de n-de partiële som heeft de vorm $S_n = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$. Teneinde Stelling 1.2 toe te passen beschouwen we het verschil $S_n - S_m$. Als $n > 2m$, dan is $|S_n - S_m| = S_n - S_m = \frac{1}{m+1} + \frac{1}{m+2} + \ldots + \frac{1}{n} > \frac{1}{n} + \frac{1}{n} + \ldots \frac{1}{n} = \frac{n-m}{n} > \frac{1}{2}$.
Kies $\varepsilon = \frac{1}{2}$, dan is bij elke m dus een n te vinden zodat $|S_n - S_m| > \varepsilon$. De index N uit Stelling 1.2 bestaat dus niet, zodat de rij van partiële sommen divergeert.

Uit voorbeeld 1.9 volgt, dat een reeks $a_1 + a_2 + a_3 + \ldots$ niet noodzakelijk convergeert als $\lim_{n\to\infty} a_n = 0$. Als de reeks echter convergeert, dan kan men aantonen dat de rij $\{a_n\}$ convergeert met limiet nul. Een reeks is dus zeker divergent als de rij $\{a_n\}$ niet naar nul convergeert.

Voorbeeld 1.10. De hyperharmonische reeks $1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \ldots$ is convergent. Immers de rij van partiële sommen is monotoon stijgend. Daar voor elke $k > 1$ geldt dat $\frac{1}{k^2} < \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}$, is $S_n = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots + \frac{1}{n^2} < 1 + (\frac{1}{1} - \frac{1}{2}) + (\frac{1}{2} - \frac{1}{3}) + \ldots + (\frac{1}{n-1} - \frac{1}{n}) = 2 - \frac{1}{n} < 2$ voor alle n, zodat volgens Stelling 1.3 de rij $\{S_n\}$ convergeert. Men kan laten zien, dat de som van de reeks $\frac{1}{6}\pi^2$ is ($\pi^2 \approx 9{,}870$; $\frac{\pi^2}{6} \approx 1{,}645$).

Men noemt $a_1 + a_2 + a_3 + \ldots$ absoluut convergent, als de reeks van absolute waarden $|a_1| + |a_2| + |a_3| + \ldots$ convergeert. Zonder bewijs vermelden we

Stelling 1.4. Elke absoluut convergente reeks is ook convergent.

Een alternerende reeks is een reeks waarvan de termen beurtelings positief en negatief zijn. Een alternerende reeks is op grond van bovenstaande stelling zeker convergent als ze absoluut convergeert.
De volgende stelling kan men vaak toepassen als een alternerende reeks niet absoluut convergeert.

Stelling 1.5. Een alternerende reeks is convergent als de termen in absolute waarde monotoon naar nul convergeren.

Voorbeeld 1.11. De reeks $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - + \ldots$ is convergent volgens Stelling 1.5, omdat $1 > \frac{1}{2} > \frac{1}{3} > \ldots$ en $\lim_{n\to\infty} \frac{1}{n} = 0$. Deze reeks is niet absoluut convergent (voorbeeld 1.8).

Voorbeeld 1.12. De reeks $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + 1 \ldots$ is convergent volgens Stelling 1.5. Ook deze reeks is niet absoluut convergent. Men kan bewijzen dat de som van deze reeks gelijk is aan $\frac{\pi}{4}$.

Voorbeeld 1.13. De meetkundige reeks $1 + (-\frac{1}{2}) + (-\frac{1}{2})^2 + (-\frac{1}{2})^3 + \ldots$ is absoluut convergent en dus convergent (Stelling 1.4). De convergentie volgt ook uit Stelling 1.5.

We merken nog op, dat men de termen van een convergente reeks, die niet absoluut convergeert, niet straffeloos van volgorde mag verwisselen. De reeks kan dan nl. een andere som krijgen of zelfs gaan divergeren! Beschouw b.v. de reeks uit voorbeeld 1.11:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} +- \ldots$$

en schrijf hieronder dezelfde reeks na vermenigvuldiging met $\frac{1}{2}$

$$\frac{1}{2} \quad - \frac{1}{4} \quad + \frac{1}{6} \quad - \frac{1}{8} +- \ldots$$

tellen we deze beide reeksen op (combineer termen die onder elkaar staan), dan krijgen we

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} ++- \ldots,$$

welke reeks ook ontstaat als we de termen uit de eerste reeks in andere volgorde schrijven.

§5. Binomiaalcoëfficiënten

In de wiskunde hebben we vaak te maken met producten van de vorm $n(n-1)(n-2) \ldots 3.2.1$, waarin n een natuurlijk getal is.
Men heeft hier een verkorte notatie voor ingevoerd:

$$n! = n(n-1)(n-2) \ldots 3.2.1 \quad \text{(voor alle natuurlijke n)}.$$

(Men spreke n! uit als n faculteit.) Deze getallen nemen met opklimmende n zeer snel toe, zoals uit het volgende lijstje blijkt:

$$1! = 1.\ 2! = 2,\ 3! = 6,\ 4! = 24,\ 5! = 120,\ 6! = 720,\ 7! = 5040,$$
$$8! = 40320,\ 9! = 362880,\ 10! = 3628800,\ \ldots\ .$$

Tenslotte definieert men 0! als 1. Merk op, dat $n! = n.(n-1)!$

We kunnen n! ook interpreteren als het aantal mogelijke volgorden,

waarin men n objecten kan rangschikken. Immers voor de eerste plaats heeft men de keuze uit n objecten, voor de tweede plaats de keuze uit de n-1 resterende objecten, voor de derde plaats de keuze uit de n-2 resterende objecten, etc. Bij het samenstellen van een bepaalde volgorde kiest men dus uit n(n-1)(n-2) ... 3.2.1 mogelijkheden.

In tal van gebieden van de wiskunde, o.a. ook in de mathematische statistiek, spelen binomiaalcoëfficiënten een grote rol.
We definiëren:

Definitie 1.3. Laat n en k natuurlijke getallen zijn, $n \geq k$. Dan is per definitie

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

de binomiaalcoëfficiënt van n boven k.
Voorts definiëren we $\binom{n}{0} = 1$ en $\binom{0}{0} = 1$.

Men kan de definitie van binomiaalcoëfficiënten ook uitbreiden tot willekeurige reële getallen n in bovenstaande formule, maar daar zullen we in het kader van dit college geen gebruik van maken.

Ook binomiaalcoëfficiënten kan men een practische interpretatie geven. Stel dat we uit n objecten er k kiezen. Dan zijn er $\binom{n}{k}$ verschillende keuzen van k objecten mogelijk (als we niet op de volgorde letten waarin de objecten worden gekozen). Immers bij het eerste te kiezen object hebben we n mogelijkheden, bij het tweede te kiezen object nog n-1 mogelijkheden, ..., bij het k-de te kiezen object nog n-(k-1) mogelijkheden, zodat we uit n(n-1)(n-2) ... (n-k+1) mogelijkheden kunnen kiezen. Hierbij hebben we echter wél op de volgorde gelet: we hebben gespecificeerd welk object bij de eerste resp. tweede resp. ... resp. k-de keuze wordt getrokken. We kunnen, zoals we zagen, dezelfde k objecten in k! verschillende volgorden rangschikken en dus ook in k! verschillende volgorden trekken. Letten we niet op de volgorde, dan mogen we tussen deze k! verschillende trekkingsvolgorden géén onderscheid maken en moeten we n(n-1)(n-2) ... (n-k+1) dus nog door k! delen om het aantal verschillende keuzen te verkrijgen. Dit leidt volgens definitie 1.3 juist tot $\binom{n}{k}$ verschillende mogelijkheden.

We noemen enige eenvoudige eigenschappen van binomiaalcoëfficiënten

die het rekenwerk vereenvoudigen.

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

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

(ii) $\binom{n}{n} = 1$, $\binom{n}{1} = \binom{n}{n-1} = n$ (volgt direct uit definitie).

(iii) $\binom{n}{k} = \binom{n}{n-k}$, immers beide leden zijn gelijk aan $\frac{n!}{k!(n-k)!}$

(iv) $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$,

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

$= \frac{n(n-1)(n-2)\ \ldots\ (n-k+1)}{k!} \cdot \left(\frac{k}{n} + \frac{n-k}{n}\right) = \binom{n}{k}$.

Deze laatste eigenschap stelt ons in staat binomiaalcoëfficiënten voor niet te grote n en k eenvoudig te berekenen m.b.v. de <u>driehoek van PASCAL</u>:

$\binom{0}{0}$	1
$\binom{1}{0}\ \binom{1}{1}$	1 1
$\binom{2}{0}\ \binom{2}{1}\ \binom{2}{2}$	1 2 1
$\binom{3}{0}\ \binom{3}{1}\ \binom{3}{2}\ \binom{3}{3}$	1 3 3 1
$\binom{4}{0}\ \binom{4}{1}\ \binom{4}{2}\ \binom{4}{3}\ \binom{4}{4}$	1 4 6 4 1
$\binom{5}{0}\ \binom{5}{1}\ \binom{5}{2}\ \binom{5}{3}\ \binom{5}{4}\ \binom{5}{5}$	1 5 10 10 5 1
. . . . .	. . . . .

De getalwaarden van de links vermelde binomiaalcoëfficiënten staan op de overeenkomstige plaatsen in de rechterfiguur. Elk getal op een nieuwe regel onstaat door de twee getallen, die er schuin boven staan, op te tellen, behalve de getallen aan de rand die steeds 1 zijn (hier staat ook

maar één getal schuin boven, nl. 1).

Uit de schoolwiskunde is bekend, dat $(1+x)^2 = 1 + 2x + x^2$ en $(1+x)^3 = 1 + 3x + 3x^2 + x^3$. Het is echter veel minder eenvoudig om bijv. $(1+x)^{20}$ op dergelijke wijze snel in machten van x uit te drukken. M.b.v. binomiaalcoëfficiënten kan men hier echter een algemene uitdrukking voor geven.

Stelling 1.6. Zij n een natuurlijk getal. Dan is

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

Deze relatie staat bekend als het binomium van NEWTON.

Bewijs:
We zullen het binomium bewijzen met volledige inductie naar n. Voor n = 1 is de formule correct (ga na). Zij de formule correct voor n - 1. Dan is

$$(1+x)^n = (1+x)(1+x)^{n-1} = (1+x)\sum_{k=0}^{n-1}\binom{n-1}{k}x^k = \sum_{k=0}^{n-1}\binom{n-1}{k}x^k + \sum_{k=0}^{n-1}\binom{n-1}{k}x^{k+1} =$$

$$= 1 + \sum_{k=1}^{n-1}\{\binom{n-1}{k} + \binom{n-1}{k-1}\}x^k + x^n = 1 + \sum_{k=1}^{n-1}\binom{n}{k}x^k + x^n = \sum_{k=0}^{n}\binom{n}{k}x^k,$$

(de één na laatste overgang volgt uit eigenschap (iv)), zodat de formule ook juist is voor n. Hiermee is de stelling bewezen.

Het is nu eenvoudig $(x+y)^n$, met twee variabelen x en y, te ontwikkelen. Immers $(x+y)^n = y^n(1 + \frac{x}{y})^n$, terwijl volgens het binomium van NEWTON

$$(1 + \frac{x}{y})^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{-k}.$$

Dit geeft

$$(x+y)^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}.$$

Is x+y = 1, dus y = 1-x, dan volgt hieruit onmiddellijk dat

$$\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k} = 1.$$

Voorbeeld 1.14. Kiezen we in het binomium van NEWTON x = 1, dan vinden we, dat de som van de getallen in een rij van de driehoek van PASCAL gelijk is aan

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Voorbeeld 1.15. Zij c een willekeurig positief getal. Dan is $\lim_{n\to\infty} \frac{c^n}{n!} = 0$. Immers zij n_0 het kleinste natuurlijke getal $\geq 2c$. Dan kunnen we, voor $n > n_0$, schrijven

$$\frac{c^n}{n!} = \frac{c^{n_0}}{n_0!} \cdot \frac{c^{n-n_0}}{(n_0+1)(n_0+2)\ldots n} = \frac{c^{n_0}}{n_0!} \cdot \frac{c}{n_0+1} \cdot \frac{c}{n_0+2} \cdots \frac{c}{n} < \frac{c^{n_0}}{n_0!} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2} = \frac{c^{n_0}}{n_0!} \left(\frac{1}{2}\right)^{n-n_0}.$$

Daar $\frac{c^{n_0}}{n_0!}$ een vast getal is en $\lim_{n\to\infty} \left(\frac{1}{2}\right)^{n-n_0} = 0$, nadert $\frac{c^n}{n!}$ inderdaad tot nul als $n \to \infty$. Uit dit voorbeeld blijkt wel hoe snel n! stijgt bij toenemende n.

§6. Machtreeksen. De exponentiële functie

In §4 van dit hoofdstuk hebben we gesproken over reeksen. De termen van de besproken reeksen waren gegeven getallen. Men kan echter ook functiereeksen beschouwen, waarvan de termen functies van een variabele zijn, bijv.

$$1 + x + x^2 + x^3 + \ldots,$$

$$\sin x + \sin 2x + \sin 3x + \ldots .$$

Zeer belangrijke functiereeksen zijn de machtreeksen, die de algemene gedaante

$$c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \ldots \quad (c_0, c_1, c_2, \ldots \text{ constante coëfficiënten})$$

hebben. Het bekendste voorbeeld van een dergelijke machtreeks is de meetkundige reeks $1 + x + x^2 + x^3 + \ldots$. Ook bij een machtreeks is het van grote betekenis, voor welke waarden van x de reeks convergeert. De volgen-

de stelling (die we niet bewijzen) geeft hierin reeds enig inzicht.

Stelling 1.7. Bij de convergentie van een machtreeks zijn er drie mogelijkheden:

(i) de reeks convergeert voor geen enkele x (behalve voor x = 0),

(ii) " " " " alle waarden van x,

(iii) er bestaat een getal R zodanig, dat de reeks convergeert voor alle x met $|x| < R$ en divergeert voor alle x met $|x| > R$. Men noemt R de convergentie-straal van de machtreeks. Men zou kunnen zeggen, dat in geval (i) R = 0 en in geval (ii) $R = \infty$.

Voorbeeld 1.16. De meetkundige reeks $1 + x + x^2 + x^3 + \ldots$ heeft convergentiestraal R = 1.

Voorbeeld 1.17. De machtreeks $1 + x + 2^2x^2 + 3^3x^3 + 4^4x^4 + \ldots$ convergeert voor geen enkele x (behalve x = 0); immers opdat voor een vaste x de reeks convergeert, moet in elk geval $\lim_{n\to\infty} n^n x^n$ nul zijn (zie opmerking na voorbeeld 1.9). Welke waarde x ook heeft, nx neemt met toenemende n onbeperkt toe, zodat de limiet van $(nx)^n$ niet eens bestaat.

Binnen het convergentiegebied bestaat de som van een machtreeks en deze is uiteraard ook een functie van x. We zullen zo'n som daarom aanduiden met S(x). Zo is bij de meetkundige reeks $S(x) = \frac{1}{1-x}$, mits $|x| < 1$.

Beschouw een machtreeks $c_0 + c_1x + c_2x^2 + c_3x^3 + \ldots$. Differentiëren we deze machtreeks term voor term naar x, dan ontstaat een nieuwe machtreeks $c_1 + 2c_2x + 3c_3x^2 + \ldots$. Men kan laten zien, dat deze nieuwe machtreeks dezelfde convergentiestraal heeft als de oorspronkelijke reeks. Is de som van de oorspronkelijke reeks S(x), dan is de som van de afgeleide reeks juist gelijk aan S'(x), de afgeleide van S(x) (binnen het convergentiegebied). Binnen het convergentiegebeid is de som van een machtreeks dus differentieerbaar en de afgeleide wordt gevonden door de machtreeks term voor term te differentiëren. Dit is een zeer belangrijke eigenschap van machtreeksen.

Voorbeeld 1.18. Beschouw de meetkundige reeks $1 + x + x^2 + x^3 + \ldots$ voor $|x| < 1$. De som van deze reeks is $\frac{1}{1-x}$, met de afgeleide $\frac{1}{(1-x)^2}$. Diffenrentiëren we de reeks term voor term, dan is volgens bovenstaande

eigenschap de reeks $1 + 2x + 3x^2 + \ldots$ dus ook convergent voor $|x| < 1$ met als som $\frac{1}{(1-x)^2}$.

Een centrale rol speelt in de wiskunde de reeks

$$1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots = \sum_{n=0}^{\infty} \frac{x^n}{n!} .$$

Deze machtreeks is absoluut convergent voor alle x, zoals we thans zullen bewijzen. Zij x_0 een willekeurig reëel getal en ε een willekeurig positief getal. We zullen aantonen, dat voor voldoend grote natuurlijke getallen m en n $(n > m)$

$$\frac{(x_0)^m}{m!} + \frac{(x_0)^{m+1}}{(m+1)!} + \ldots + \frac{(x_0)^n}{n!} < \varepsilon$$

waarna de absolute convergentie van de reeks (voor x_0) volgt uit Stelling 1.2. Zij $m > 2x_0$. Dan is

$$\frac{|x_0|^m}{m!} + \frac{|x_0|^{m+1}}{(m+1)!} + \ldots + \frac{|x_0|^n}{n!} = \frac{|x_0|^m}{m!} \cdot \{1 + \frac{|x_0|}{m+1} + \ldots + \frac{|x_0|^{n-m}}{(m+1)(m+2)\ldots n}\} <$$

$$< \frac{|x_0|^m}{m!} \cdot \{1 + \frac{1}{2} + \ldots + (\frac{1}{2})^{n-m}\} = \frac{|x_0|^m}{m!} \cdot \frac{1-(\frac{1}{2})^{n-m+1}}{1-\frac{1}{2}} < 2\,\frac{|x_0|^m}{m!} .$$

Volgens voorbeeld 1.15 is $\lim_{m\to\infty} \frac{|x_0|^m}{m!} = 0$, zodat voor alle voldoend grote m $\frac{|x_0|^m}{m!} < \frac{\varepsilon}{2}$, waarmee de aangekondigde ongelijkheid bewezen is. Daar x_0 willekeurig was gekozen, is de beschouwde reeks dus inderdaad (absoluut) convergent voor alle x.

De som van bovenstaande reeks zullen we voorlopig aanduiden met E(x). De functie E(x) heeft de merkwaardige eigenschap, dat de afgeleide E'(x) gelijk is aan E(x) zelf (voor alle x), immers differentiëren we de reeks term voor term, dan ontstaat weer dezelfde reeks (ga na!).

Om het karakter van E(x) te leren kennen,leiden we een aantal eigenschappen van de functie E(x) af.

(i) Beschouw de functie $g(x) = E(x).E(-x)$. Differentiatie naar x geeft:

$$g'(x) = E'(x).E(-x) - E(x).E'(-x) = E(x)E(-x) - E(x)E(-x) = 0,$$

m.a.w. $g(x)$ heeft in elk punt afgeleide nul en is dus constant. Daar $g(0) = \{E(0)\}^2 = 1$, is dus

$$E(x).E(-x) = 1 \quad \text{voor alle } x.$$

(ii) Uit eigenschap (i) blijkt, dat $E(x) \neq 0$ voor alle x, terwijl uit de definitie van $E(x)$ volgt, dat $E(x) > 0$ is voor alle $x \geq 0$. Uit (i) volgt dan weer, dat $E(x) > 0$ is voor $x < 0$, zodat

$$E(x) > 0 \quad \text{voor alle } x.$$

(iii) Beschouw de functie $h(x) = E(x)/E(x+a)$, waarin a een willekeurige constante. Differentiatie naar x levert:

$$h'(x) = \frac{E'(x)\,E(x+a) - E'(x+1)\,E(x)}{\{E(x+a)\}^2} = \frac{E(x)\,E(x+a) - E(x+a)\,E(x)}{\{E(x+a)\}^2} = 0,$$

m.a.w. $h(x)$ heeft in elk punt afgeleide nul en is dus constant. Daar $h(0) = E(0)/E(a) = 1/E(a)$, is dus $E(x)/E(x+a) = 1/E(a)$ of wel

$$E(x+a) = E(x)\,E(a) \quad \text{voor alle } x \text{ en alle } a.$$

(iv) Noemen we $E(1) = e$, dan volgt door herhaald toepassen van eigenschap (iii), dat voor elk natuurlijk getal n geldt

$$e = E(1) = E(n \cdot \tfrac{1}{n}) = E(\tfrac{1}{n} + \tfrac{1}{n} + \ldots + \tfrac{1}{n}) = \{E(\tfrac{1}{n})\}^n,$$

of wel

$$E(\tfrac{1}{n}) = e^{1/n}.$$

Is ook m een natuurlijk getal, dan volgt op dezelfde wijze

$$E(\tfrac{m}{n}) = E(m \cdot \tfrac{1}{n}) = \{E(\tfrac{1}{n})\}^m = e^{m/n},$$

zodat voor alle rationale getallen $r \geq 0$

$$E(r) = e^r.$$

Is r een negatief rationaal getal, schrijf r dan als $-\frac{m}{n}$ (m en n natuurlijke getallen), dan is (zie (i))

$$E(r) = 1/E(\tfrac{m}{n}) = e^{-m/n} = e^{r},$$

zodat voor elk rationaal getal de eigenschap $E(r) = e^{r}$ juist is.
Men kan deze eigenschap ook generaliseren tot willkeurige reële getallen r, we gaan hier verder niet op in.

We vatten deze eigenschappen samen in de volgende stelling:

Stelling 1.8. Zij het getal e gedefinieerd door

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots,$$

dan is voor elk reëel getal x

$$1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots = e^{x};$$

deze functie heeft de eigenschap, dat

$$\frac{d}{dx} e^{x} = e^{x}.$$

Het getal e speelt een even belangrijke rol in de wiskunde als het getal π; evenals π is ook e irrationaal. In vijf decimalen: e = 2,71828... .

Het getal e kan men ook op geheel andere wijze te voorschijn krijgen. Er geldt nl. de volgende limiet-relatie:

$$\lim_{n\to\infty} (1 + \frac{1}{n})^{n} = e.$$

We tonen eerst aan, dat de limiet in het linkerlid bestaat. Volgens het binomium van NEWTON is

$$(1 + \frac{1}{n})^n = 1 + n \cdot \frac{1}{n} + \frac{n(n-1)}{2!} \cdot \frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!} \frac{1}{n^3} + \ldots + \frac{n!}{n!} \cdot \frac{1}{n^n} =$$

$$= 1 + 1 + \frac{1}{2!}(1 - \frac{1}{n}) + \frac{1}{3!}(1 - \frac{1}{n})(1 - \frac{2}{n}) + \ldots$$

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

Noemen we $T_n = (1 + \frac{1}{n})^n$, dan is de rij $\{T_n\}$ monotoon stijgend, immers T_{n+1} ontstaat uit T_n door in de laatste som de factoren 1-1/n, 1-2/n, ... te vervangen door de grotere factoren 1-1/(n+1), 1-2/(n+1), ... en er tenslotte nog een positieve term bij op te tellen.
Noem $S_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}$, dan volgt uit de ontwikkeling van T_n bovendien, dat $T_n \leq S_n$ voor elke n. Daar de rij $\{S_n\}$ ook monotoon stijgt en $\lim_{n\to\infty} S_n = e$, is $T_n < e$ voor elke n. Uit Stelling 1.3 volgt nu, dat de rij $\{T_n\}$ convergeert. Noemen we de (nog onbekende) limiet T, dam is in elk geval $T \leq e$, daar $T_n < e$ voor elke n. Om te bewijzen dat $T = e$, merken we op dat voor elke positieve gehele $m < n$ geldt

$$T_n > 1 + 1 + \frac{1}{2!}(1 - \frac{1}{n}) + \frac{1}{3!}(1 - \frac{1}{n})(1 - \frac{2}{n}) + \ldots + \frac{1}{m!}(1 - \frac{1}{n})(1 - \frac{2}{n})$$

$$\ldots (1 - \frac{m-1}{n})$$

(we hebben nu immers in het rechterlid de laatste termen uit de ontwikkeling van T_n weggelaten). Nemen we in beide leden de limiet voor $n \to \infty$ (m vast) dan nadert het linkerlid tot T en het rechterlid tot S_m, zodat $T \geq S_m$. Door n voldoende groot te kiezen, kunnen we deze laatste ongelijkheid voor elke $m = 1, 2, 3, \ldots$ bewijzen. Maar dan volgt, dat ook
$T \geq \lim_{m\to\infty} S_m = e$.
We hebben nu bewezen, dat $T \leq e$ en tegelijk $T \geq e$, zodat $T = e$, waarmee de limiet-relatie bewezen is.

Men noemt e^x een <u>exponentiële functie</u>, omdat de variabele x in de exponent optreedt. De functie e^x is geschetst in fig. 1.1 en uitvoerig getabelleerd voor een groot aantal waarden van x. Exponentiële functies komen niet alleen voor in de zuivere wiskunde, maar spelen ook een belang-

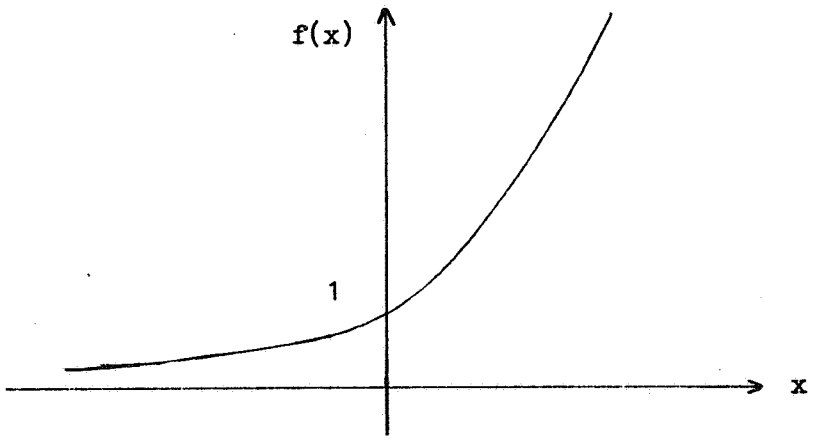


fig. 1.1. De grafiek van $f(x) = e^x$.

rijke rol bij het beschrijven van fysische, chemische of biologische processen door middel van een wiskundig model. Ze zijn tevens van groot belang in de mathematische statistiek.

Daar e^x ook e^x tot een primitieve functie heeft, kan men eenvoudig bepaalde integralen van de functie e^x berekenen, immers

$$\int_a^b e^x dx = e^x \Big|_a^b = e^b - e^a.$$

In de mathematische statistiek speelt de functie $e^{-\frac{1}{2}x^2}$ een grote rol. Men kan bewijzen, dat de (oneigenlijke) integraal van deze functie over het integratie-interval $(-\infty,\infty)$ bestaat en gelijk is aan

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi}.$$

Hoofdstuk 2. Elementaire Statistiek

§1. Populatie en steekproef

Het prototype, of, zoals men meestal zegt, het model van een grote groep statistische problemen is het volgende:
Men heeft een vaas, waarin zich een groot aantal rode en witte knikkers bevindt. Men wil een schatting hebben van het percentage rode knikkers in de vaas.

Een iets andere vorm van dit zelfde probleem hebben we, als men om een of andere reden vermoedt dat de fraktie rode knikkers een bepaalde waarde p heeft, en wenst na te gaan of dit vermoeden juist is.

Voorbeeld 2.1. Een partij goederen bevat een onbekende fraktie ondeugdelijke exemplaren. Hoe groot is deze fraktie?

Voorbeeld 2.2. Een deel van de chimpansees bereikt de volwassen leeftijd, de andere overlijden voordien. Hoe groot is deze laatste fraktie?

Voorbeeld 2.3. Als men ratten een bepaalde dosis van een zekere stof inspuit, ondervinden sommige ratten schadelijke gevolgen, andere niet. Hoe groot is het percentage dat schadelijke gevolgen ondervindt van de injektie?

Voorbeeld 2.4. Op grond van de erfelijkheidswetten van MENDEL mag men verwachten, dat een bepaalde fraktie p van alle planten, die door kruising van twee planten verkregen kunnen worden, een bepaalde eigenschap bezit. Hoe kan men de juistheid van dit vermoeden controleren?

In al deze voorbeelden gaat het erom een uitspraak te doen over een groep van objekten. Deze groep van objekten noemt men de populatie die beschouwd wordt. Zo is de populatie in ons model de verzameling knikkers in de vaas, in vb. 2.1 de partij goederen, in vb. 2.2 de verzameling van alle chimpansees, in vb. 2.3 alle ratten van een bepaalde soort en in vb. 2.4 de verzameling van alle planten die door kruising uit de twee planten verkregen kunnen worden. Het aantal objekten, waaruit de populatie bestaat, noemt men de omvang van de populatie.

Het is bij het toepassen van statistische methoden van groot belang,

dat men de populatie nauwkeurig omschrijft. Bij vb. 2.2 zou men zich zeer wel kunnen voorstellen, dat de fraktie chimpansees die overlijdt alvorens de volwassen leeftijd te bereiken geheel anders ligt voor in gevangenschap en in de vrije natuur levende dieren, terwijl er ook nog verschillen zouden kunnen bestaan tussen deze frakties bij dieren in verschillende streken. Het is dan essentiëel vast te stellen welke populatie men zal beschouwen.

In het geval van de vaas met knikkers, en ook in vb. 2.1, zou men alle rode en witte knikkers uit de vaas, resp. alle deugdelijke en ondeugdelijke exemplaren uit de partij, kunnen tellen. We spreken dan van een volledige waarneming. Maar bij zeer grote populatie-omvang kan dit bezwaarlijk zijn, en in vb. 2.1 kan het gebeuren, dat het middel erger is dan de kwaal, nl. als we een destruktieve keuringsmethode moeten gebruiken om uit te maken of een exemplaar al dan niet deugdelijk is. In andere gevallen (vb. 2.3) is volledige waarneming zelfs in het algemeen onmogelijk!

Het essentiële van de wiskundige statistiek is nu, dat deze ons methoden geeft, die ons in staat stellen uit onvolledige waarnemingen, d.w.z. waarnemingen die slechts een (vaak onbekend) deel van de populatie omvatten, conclusies te trekken over de gehele populatie. Om een onvolledige waarneming te doen moeten we eerst een aantal objekten uit de populatie trekken waaraan we de waarnemingen gaan verrichten. In het geval van de vaas met knikkers moeten we een aantal knikkers eruit pakken, in vb. 2.3 een aantal ratten afzonderen, etc., etc. We spreken in dit verband van het nemen van een steekproef (Engels: sample) uit de populatie. Een steekproef kan bestaan uit een willekeurig aantal n objekten; we noemen dit aantal de omvang van de steekproef.

Een steekproef van de omvang n kan men op twee verschillende manieren verkrijgen:

(i) Men trekt telkens één objekt uit de populatie, doet aan dit ene objekt de waarneming, en voegt het weer bij de populatie. Eénzelfde objekt kan op deze wijze méér dan eens in de steekproef voorkomen. We spreken van een steekproef met teruglegging.

(ii) Men trekt n objekten (al of niet tegelijkertijd) uit de populatie, zonder de getrokken elementen weer aan de populatie toe te voegen. Nu spreekt men van een steekproef zonder teruglegging.

In de praktijk past men gewoonlijk steekproeftrekkingen zonder teruglegging toe. Voor de theorie maakt het verschil of men steekproeven met, dan wel zonder teruglegging neemt, tenminste als de omvang der populatie eindig is. Als de steekproefomvang klein is t.o.v. de populatieomvang wordt het verschil verwaarloosbaar. Daar de theorie het gemakkelijkst is voor steekproeven met teruglegging, zullen we ons in het vervolg vooral bezig houden met steekproeven met teruglegging of - als geen teruglegging plaats vindt - tot populaties wier omvang zeer groot is t.o.v. de steekproefomvang. Een dergelijke steekproef ontstaat per definitie door achtereenvolgens n trekkingen uit dezelfde populatie te nemen. Immers na elke trekking en de bijbehorende waarneming aan het getrokken objekt wordt dit objekt weer bij de populatie gevoegd, zodat de populatie bij de volgende trekking onveranderd is.

We zullen nu twee eisen geven waaraan trekkingen met teruglegging moeten voldoen.

E1. Elke trekking moet aselect zijn, d.w.z. alle objekten uit de populatie moeten gelijkwaardig zijn bij de trekking; we mogen niet een bepaald objekt selekteren, maar moeten het te trekken objekt lukraak kiezen.

E2. De trekkingen moeten onderling onafhankelijk zijn, d.w.z. de keuze van een objekt mag niet afhangen van het resultaat van de vorige trekkingen.

Heeft men bij de trekkingsprocedure deze beide eisen in acht genomen dan spreekt men van een aselekte steekproef met teruglegging. Trekt men een steekproef zonder teruglegging, dan dient men er zorg voor te dragen dat elke trekking aselekt is m.b.t. de populatie die na de voorgaande trekkingen nog resteert, zonder acht te slaan op de uitkomsten van de voorgaande trekkingen. In dit geval spreekt men van een aselekte steekproef zonder teruglegging.

Bij het nemen van een steekproef dient men er steeds voor te waken dat aan deze eisen voldaan is. Bij het trekken van knikkers uit een vaas (of van kaarten uit een spel) kan men proberen dit te verwezenlijken door voor elke trekking zo goed mogelijk te schudden.

Het volgende hulpmiddel zal vaak uitkomst brengen. We nummeren de objekten van de populatie in één of andere volgorde en bepalen door loting welke nummers we in de steekproef zullen opnemen. Daartoe kunnen we een vaas met tien gelijke knikkers, genummerd 0, 1, 2, ..., 9 nemen en hieruit met teruglegging aselekte trekkingen doen.

Wenst men een steekproef met omvang 50 uit een populatie met omvang 1000, dan nummeren we de objekten van de populatie van 000 t/m 999. Uit de vaas met 10 knikkers doen we nu 50 maal telkens drie trekkingen met teruglegging. Zo krijgen we 50 groepjes van drie cijfers, dus 50 nummers onder de 1000. In de steekproef nemen we nu die 50 objekten uit de populatie op, wier nummers overeenkomen met de 50 groepjes van drie cijfers die we getrokken hebben. Zo wordt dan een aselekte steekproef met teruglegging verkregen, want onder de 50 nummers kunnen gelijke voorkomen. Wenst men een steekproef zonder teruglegging, dan schrappe men nummers, die reeds eerder voorgekomen zijn en zette het lotingsprocédé voort tot 50 verschillende nummers onder 1000 zijn verkregen.

Vooral bij grote steekproeven kost het tijd om alle nodige lotingen uit te voeren. Er zijn echter tabellen geconstrueerd die de resultaten van grote aantallen aselekte trekkingen uit een vaas met tien genummerde knikkers geven. (Deze getallen zijn niet verkregen door werkelijk knikkers uit een vaas te trekken, maar m.b.v. snellere methoden die op hetzelfde neerkomen.) Zo bevatten bijv. de "Tables of random sampling numbers" van M.G. KENDALL en B. BABINGTON SMITH (Cambridge Univ. Press 1946) 100 000 trekkingen uit de populatie 0, 1, 2, ..., 9. Men noemt dergelijke cijfers gewoonlijk toevalscijfers of aselekte cijfers. Wenst men nu weer een steekproef van 50 elementen uit de populatie van omvang 1000, dan hoeft men slechts, te beginnen op een willekeurig punt, de 50 volgende groepjes van drie cijfers te nemen.

Soms is het niet mogelijk deze methode toe te passen, eenvoudig omdat het onmogelijk is de te onderzoeken populatie te nummeren, daar niet alle objekten uit de populatie bereikbaar zijn. Men zoekt dan naar een steekproefmethode waarvan men verwacht dat toch zo goed mogelijk aan de eisen van aselektheid voldaan is. Is aan de eisen slechts ten dele voldaan, dan zijn de nog te bespreken eenvoudige statistische technieken niet zonder voorbehoud van toepassing.

§2. Enige experimenten

Keren wij thans terug tot de vaas met knikkers, waarvan een onbekende fraktie rood gekleurd is. Uit deze vaas nemen we een aselekte steekproef

van n knikkers met teruglegging en noteren het aantal rode knikkers dat wij daarbij aantreffen. Wij geven dit aantal aan met x. Herhalen we dit experiment, dan zullen we in het algemeen een andere waarde voor x vinden. Dit is nu een typisch statistisch verschijnsel: herhaling van eenzelfde experiment naar beste weten op dezelfde wijze uitgevoerd geeft afwijkende uitkomsten. Het aantal rode knikkers in een steekproef van n stuks varieert van steekproef tot steekproef. Een dergelijke variërende grootheid noemen we een stochastische grootheid (Engels: random variable). We zullen in het vervolg stochastische grootheden steeds onderstrepen: $\underline{x}$ is het aantal rode knikkers in een (nog te trekken) steekproef van n knikkers uit onze vaas.

Na uitvoering van een experiment geven we de waargenomen waarde van $\underline{x}$ aan zonder onderstreping. Trekken we bijv. 50 keer een steekproef van n = 25 (met teruglegging) uit de vaas, dan kan $\underline{x}$ elke keer één van de waarden 0, 1, 2, ..., 25 aannemen. In een bepaald geval vonden we voor $\underline{x}$ de volgende waarden (in volgorde van waarneming):

I 4 3 1 2 2 2 7 3 3 0 5 4 3 2 4 3 5 4 3 4 0 1 2 2 6 5 0 4 1 6 2 1 2 6

3 5 5 2 3 2 4 3 2 3 1 1 1 4 3 3.

Het is duidelijk, dat de waargenomen waarden ieder op zichzelf reeds enige informatie verschaffen over de onbekende fraktie p van rode knikkers. Als p dicht bij 1 lag, d.w.z. als bijna alle knikkers in de vaas rood waren, zouden we niet zulke lage uitkomsten hebben gekregen. Het is dus aannemelijk dat p klein is, d.w.z. dat zich weinig rode knikkers in de vaas bevinden.

Om een duidelijk overzicht te krijgen van de resultaten, zien we van de volgorde daarvan af en vatten ze samen in een histogram (zie fig. 2.1). Dit is een staafdiagram: in horizontale richting zijn de waarden uitgezet die $\underline{x}$ kan aannemen, in vertikale richting de relatieve frekwentie waarmee $\underline{x}$ deze waarde heeft aangenomen. Het totale oppervlak van de staafjes is gelijk aan 1, immers de breedte van elk staafje is 1 evenals de som van de hoogten. Dergelijke figuren geven in één oogopslag een duidelijk overzicht van de waarnemingen; alleen de volgorde der waarnemingen gaat hierbij verloren.

Herhaling van het experiment gaf de reeks waarnemingen:

II 3 0 2 0 3 4 3 1 5 3 2 2 2 2 5 4 1 1 1 2 2 3 3 4 3 2 4 4 0 1 1 2 1 5

2 4 6 0 3 3 2 2 3 4 4 3 3 3 0 8.

Ook deze waarnemingen zijn in de vorm van een histogram samengevat (zie fig. 2.2).

We zien dat er ondanks de verschillen toch een duidelijke overeenkomst is tussen de beide waarnemingsreeksen I en II. In beide gevallen varieert x slechts van 0 tot hoogstens 8, en de grote meerderheid der gevonden waarden van x is niet kleiner dan 1 en niet groter dan 5. Bij deze experimenten ligt de fraktie rode knikkers in een steekproef van 25 stuks dus voornamelijk tussen 0,04 en 0,20.

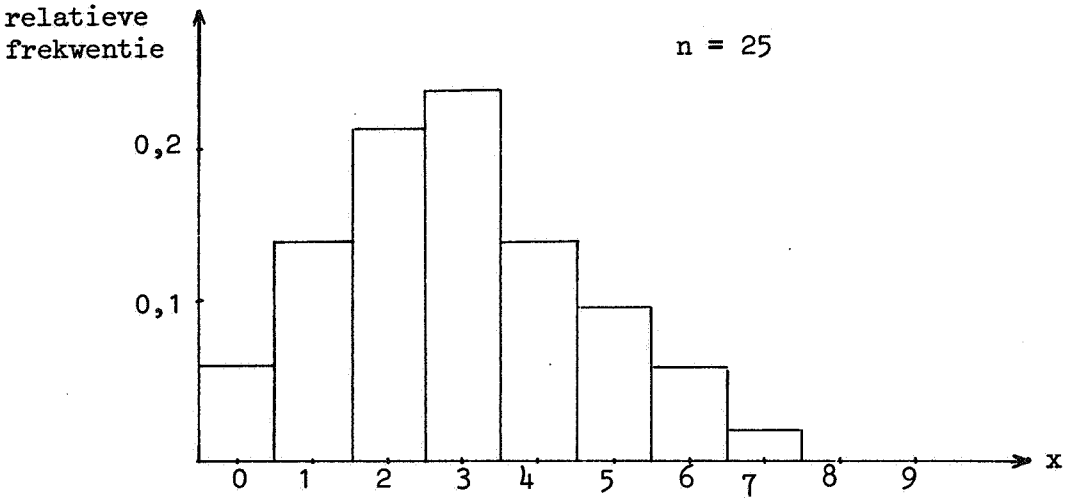


fig. 2.1. Histogram van reeks I

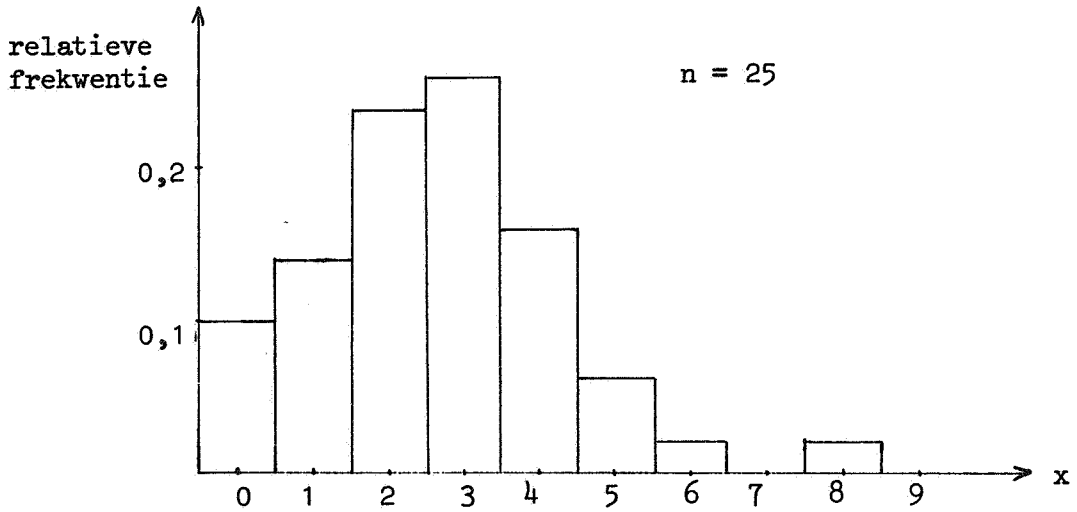


fig. 2.2. Histogram van reeks II

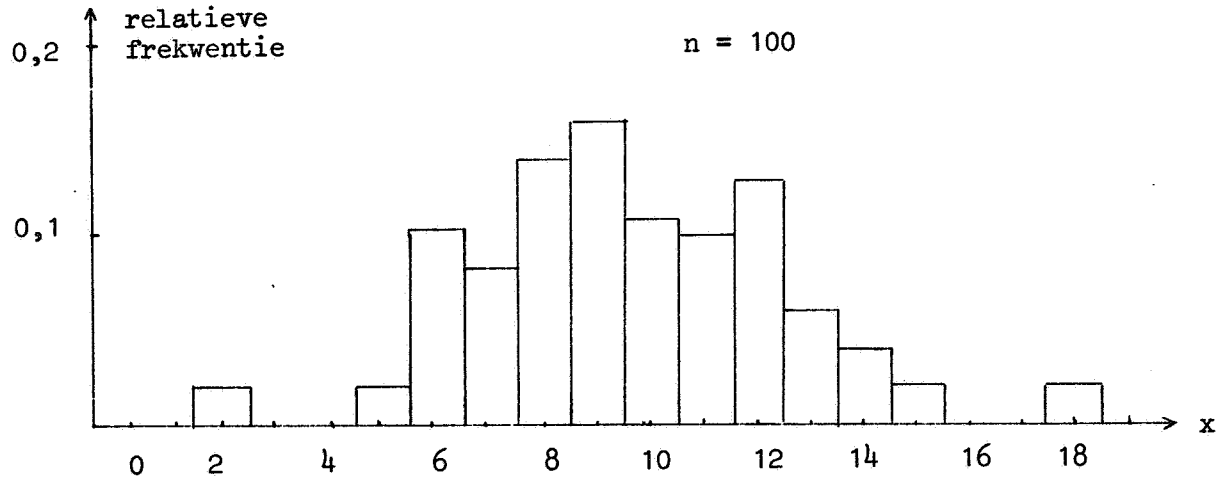


fig. 2.3. Histogram van derde reeks

Deze statistische regelmaat krijgt een iets ander karakter als we elke steekproef vergroten van 25 tot bijv. 100 trekkingen met teruglegging. De waarden van x zullen dan over een groter gebied variëren en in doorsnede ook groter uitvallen dan bij kleinere n. De fraktie x/n rode knikkers in de steekproef zal echter minder variëren. Om dit te demonstreren is een reeks van 50 steekproeven, elk van omvang 100 (met teruglegging) uit dezelfde vaas getrokken. De resultaten zijn weer samengevat in een histogram, zie fig. 2.3. In deze reeks ligt x/n in de grote meerderheid der gevallen tussen 0,06 en 0,13.

Daar het voor de hand ligt om de fraktie rode knikkers in de steekproef te gebruiken als schatting voor de onbekende fraktie rode knikkers in de vaas, kunnen we concluderen dat de nauwkeurigheid van deze schatting blijkbaar toeneemt naarmate we meer trekkingen doen. Dit is een zeer algemeen ervaringsfeit, dat we overal waar aselekte, onafhankelijke trekkingen uit een populatie genomen worden, kunnen vaststellen. In de statistiek wordt bijna voortdurend van deze experimenteel vastgestelde eigenschap gebruik gemaakt.

Voorbeeld 2.5. Teneinde de kwaliteit van een bepaald soort zaden te beoordelen, neemt een kweker 100 maal 100 zaden aselekt uit een grote partij. Vervolgens vacht hij af, hoeveel zaden in elk honderdtal binnen een zeker

tijdsbestek ontkiemen. Dit aantal is weer een stochastische variabele, x. De resultaten van het experiment zijn in het histogram van fig. 2.4 weergegeven.

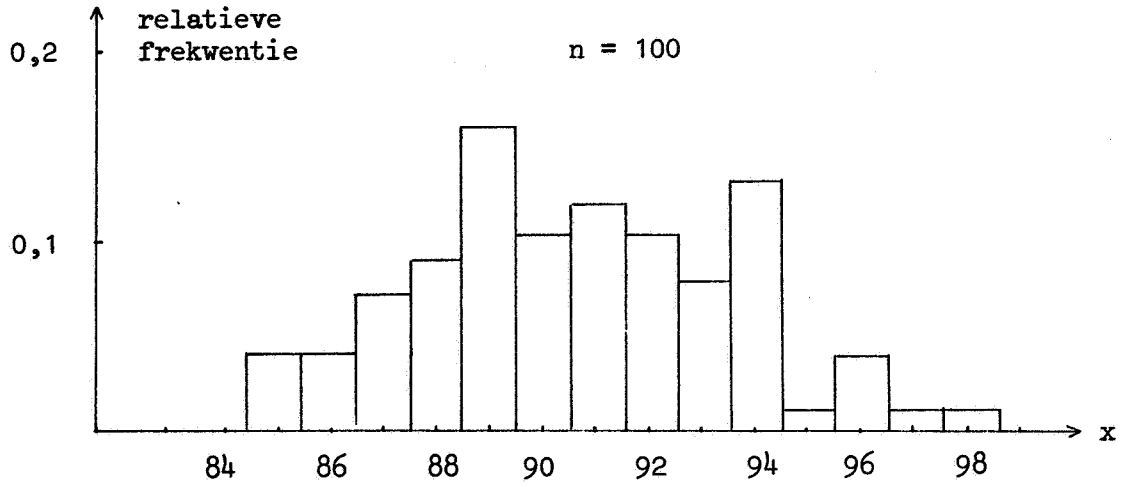


fig. 2.4. Histogram van aantallen ontkiemde zaden, zie vb. 2.5.

We zien dat bij veruit de meeste honderdtallen de fraktie ontkiemde zaden tussen 0,87 en 0,94 ligt.

Tot nu toe hebben we steeds populaties beschouwd, waarvan de elementen een bepaalde eigenschap al of niet bezitten (al dan niet rode ballen, al dan niet ontkiemde zaden). Het komt echter ook vaak voor, dat bij elk objekt uit een populatie Π één of meer getallen behoren. Bij een aselekte trekking krijgen we één objekt uit Π, en het (de) bijbehorende getal(len) kunnen we dan waarnemen. Aangezien de bijbehorende getallen bij andere trekkingen andere waarden hebben, zijn dit stochastische grootheden op de populatie Π. Trekken we een steekproef uit een dergelijke populatie, dan kunnen we de resultaten weer in een histogram uitzetten.

Voorbeeld 2.6. Zij Π de denkbeeldige populatie van alle mogelijke worpen met één dobbelsteen. Bij elke worp behoort een getal, nl. het aantal ogen x, een stochastische grootheid. Het is duidelijk, dat x alleen de waarden 1, 2, 3, 4, 5 of 6 kan aannemen. De resultaten van 25 worpen met een dobbelsteen zijn weergegeven in het histogram van fig. 2.5.

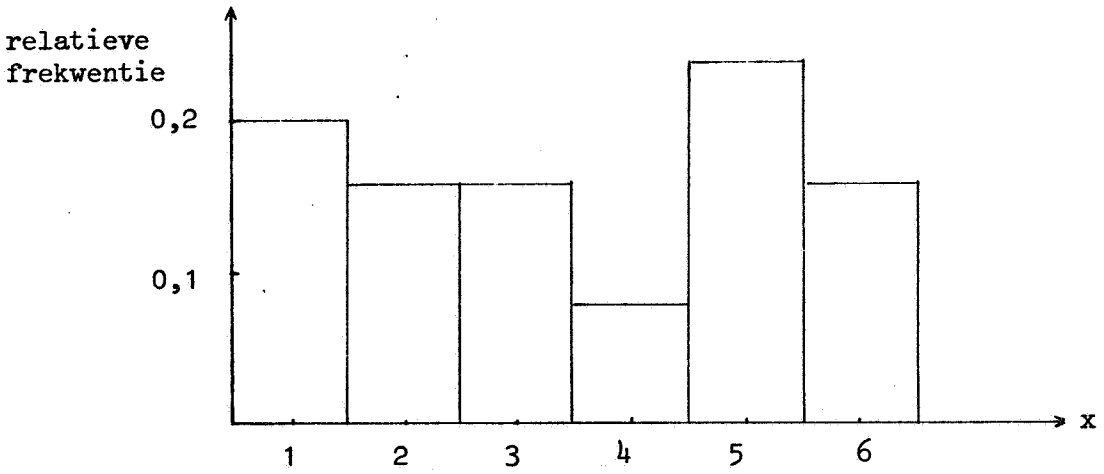


fig. 2.5. Histogram van resultaten bij worpen met dobbelsteen, zie vb. 2.6.

In bovenstaand voorbeeld kan de stochastische grootheid slechts een beperkt aantal verschillende waarden aanemen. Men spreekt dan van een diskrete stochastische grootheid. In vele gevallen kan een stochastische grootheid echter, althans in theorie, alle mogelijke waarden in een bepaald interval aannemen. Is Π bijv. de populatie van alle volwassen Nederlanders, dan is het mogelijk de lengte van elk individu uit de populatie te bepalen. Deze stochastische grootheid kan in principe elke reële waarde in een zeker interval aannemen; men noemt zo'n grootheid een continue stochastische grootheid. Trekken we een steekproef uit deze populatie, en willen we de meetresultaten op overzichtelijke wijze weergeven, dan delen we het interval van mogelijke uitkomsten in een aantal "klassen" (deelintervalletjes) in en turven, hoeveel van de gemeten lengten in elk van deze klassen terecht komen. Aldus ontstaat weer een histogram.

Voorbeeld 2.7. Bij een aselekte steekproef van 30 volwassen Nederlanders werden de volgende lengten x gemeten (in mm):

1809 1766 1667 1762 1703 1748 1852 1844 1766 1781 1694 1712 1481 1709 1765
1664 1506 1646 1540 1822 1539 1712 1690 1720 1705 1577 1718 1668 1734 1780.

Deze lengten zijn weergegeven in het histogram in fig. 2.6 (klassebreedte 50 mm).

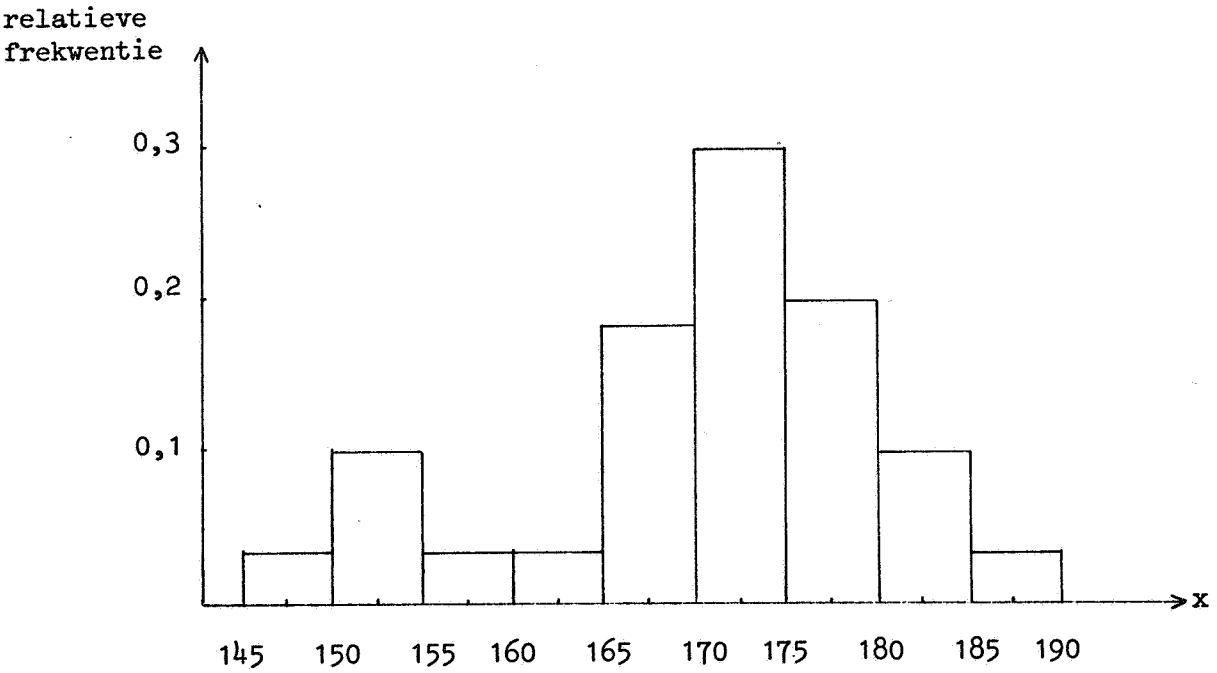


fig. 2.6. Histogram van de lengten van 30 Nederlanders, zie vb. 2.7.

De klasseindeling van het interval van mogelijke waarden van een continue stochastische grootheid is altijd min of meer willekeurig. Men streeft er steeds naar de indeling zo te kiezen, dat de typische verdeling van de uitkomsten zo goed mogelijk tot uitdrukking komt. Naarmate de steekproefomvang n toeneemt, zal men gewoonlijk een fijnere klasseindeling kiezen.

§3. Kansrekening

Zij Π een eindige populatie met omvang N, terwijl N(A) objekten van Π een bepaald kenmerk A bezitten. Doen we nu een aselekte trekking uit Π, dan kan het zijn dat we een objekt met kenmerk A trekken, doch het is ook mogelijk dat dit niet het geval is. "Het vinden van kenmerk A bij aselekte trekking" is blijkbaar een mogelijke, doch niet zekere, gebeurtenis. Een dergelijke gebeurtenis wordt daarom vaak een eventualiteit genoemd.

De kans of waarschijnlijkheid van deze gebeurtenis (eventualiteit) wordt gedefinieerd als

$$P(A) \stackrel{def}{=} \frac{N(A)}{N} .$$

Het is duidelijk dat voor elk kenmerk A de kans P(A) tenminste 0 en ten hoogste 1 is. Een onmogelijke eventualiteit, d.w.z. het vinden van een kenmerk B dat geen enkel objekt uit Π bezit, krijgt blijkbaar kans 0 en evenzo heeft de zekere eventualiteit, d.w.z. het vinden van een kenmerk C dat elk objekt uit Π bezit, kans 1.

Als de omvang van Π oneindig is, dan is het quotient $\frac{N(A)}{A}$ onbepaald, zodat we het kansbegrip niet op dezelfde wijze als bij eindige populaties kunnen invoeren. Men kan echter nog wel spreken over een fraktie p van de populatie van objekten die het kenmerk A bezitten. (Op een streng-mathematische definitie van het begrip fraktie gaan we hier niet in, daar dit ons te ver zou voeren.) We definieëren nu: $P(A) \stackrel{def}{=} p$.

Voorbeeld 2.8. Zij Π de vaas met witte en rode knikkers. Als Π in totaal N knikkers bevat, waarvan er R rood zijn, dan is de kans dat een aselekte trekking een rode knikker oplevert dus $P(\text{rood}) = \frac{R}{N}$.

Voorbeeld 2.9. Zij Π de verzameling van alle Nederlanders. Daar ongeveer de helft van de Nederlandse bevolking van het manlijk geslacht (M) is, is de kans dat een aselekte trekking uit deze populatie een persoon van het manlijk geslacht oplevert $P(M) \approx \frac{1}{2}$.

Voorbeeld 2.10. Zij Π de denkbeeldige populatie van alle mogelijk worpen met één dobbelsteen, en x het aantal ogen bij een worp (zie vb. 2.6). Bij een zuivere dobbelsteen is dan de kans op elk van de uitkomsten 1, 2, ..., 6 gelijk aan $\frac{1}{6}$.

De kansrekening houdt zich bezig met het berekenen van onbekende kansen uit bekende kansen. Fundamenteel zijn daarbij een aantal rekenregels, die wij hier zonder bewijs geven. In het geval van een eindige populatie zijn ze echter eenvoudig te verifiëren.

Als A en B twee gebeurtenissen zijn, dan is

(i) $P(A \text{ of } B) = P(A) + P(B) - P(A \text{ en } B)$. *)

Een bijzonder geval krijgen we als A en B elkaar uitsluiten, d.w.z. als A en B niet tegelijkertijd kunnen optreden. We noemen de gebeurtenissen A en B dan disjunkt. Daar in dit geval $P(A \text{ en } B) = 0$, is

(ii) $P(A \text{ of } B) = P(A) + P(B)$ als A en B disjunkt.

Algemeen: als $A_1, A_2, \ldots, A_k$ alle disjunkte gebeurtenissen zijn, d.w.z. als géén tweetal gebeurtenissen gelijktijdig kan optreden, dan is

(iii) $$P(A_1 \text{ of } A_2 \text{ of } \ldots \text{ of } A_k) = \sum_{i=1}^{k} P(A_i).$$

Als van de disjunkte gebeurtenissen $A_1, A_2, \ldots, A_k$ er altijd één optreedt, dan is $(A_1 \text{ of } A_2 \text{ of } \ldots \text{ of } A_k)$ dus een zekere eventualiteit, zodat in dat geval

(iv) $$\sum_{i=1}^{k} P(A_i) = 1.$$

Als A en B twee gebeurtenissen zijn, die elkaars al dan niet optreden niet beïnvloeden, dan is

$$P(A \text{ en } B) = P(A) \cdot P(B).$$

Is aan deze relatie voldaan, dan noemt men A en B onafhankelijk.

Voorbeeld 2.11. Zij bij één worp met een zuivere dobbelsteen A het gooien van een even aantal ogen en B het gooien van hoogstens vier ogen. Dan is, zie vb. 2.10 en rekenregel (iii), $P(A) = \frac{1}{2}$ en $P(B) = \frac{2}{3}$. Uit rekenregel (i) volgt nu: $P(A \text{ of } B) = P(A) + P(B) - P(A \text{ en } B) = \frac{1}{2} + \frac{2}{3} - \frac{1}{3} = \frac{5}{6}$. De gebeurtenissen A en B zijn onafhankelijk, want $P(A \text{ en } B) = \frac{1}{3}$ en $P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$. Zij C het gooien van een zes. Dan zijn de gebeurtenissen A en C niet onafhankelijk, immers als C optreedt, dan treedt A

*) Met "of" is steeds het niet-exclusieve "of/en" bedoeld. De eventualiteit "A of B" treedt dus op, zodra minstens één van de beide eventualiteiten A en B optreedt.

zeker op. Dit blijkt uit de kansen: P(A en C) = P(C) = $\frac{1}{6}$, terwijl P(A) . P(C) = $\frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$. De gebeurtenissen B en C zijn evenmin onafhankelijk, immers P(B en C) = 0 terwijl P(B) . P(C) = $\frac{1}{9}$. Algemeen geldt, dat disjunkte gebeurtenissen nooit onafhankelijk zijn, aangezien het optreden van de ene gebeurtenis impliceert dat de andere gebeurtenis niet optreedt.

Voorbeeld 2.12. Men doet zes onafhankelijke worpen met één zuivere dobbelsteen. Hoe groot is de kans op minstens één 6? Zij A de gebeurtenis, dat bij de zes worpen minstens één 6 optreedt, B_i de gebeurtenis, dat bij de i-de worp een 6 boven komt en $\bar{B}_i$ de gebeurtenis, dat bij de i-de worp juist géén zes bovenkomt (i = 1, 2, ..., 6). De gevraagde kans P(A) = P(B_1 of B_2 of ... of B_6) is nu niet gelijk aan P(B_1) + P(B_2) + ... + P(B_6), omdat de gebeurtenissen B_1, B_2, ..., B_6 niet disjunkt zijn. We passen daarom een foefje toe. Zij $\bar{A}$ de gebeurtenis, dat bij géén der worpen een 6 optreedt. Dan zijn A en $\bar{A}$ disjunkt, terwijl één van beiden met zekerheid optreedt, zodat P(A) + P($\bar{A}$) = 1 of wel P(A) = 1 - P($\bar{A}$). Nu is P($\bar{A}$) = P($\bar{B}_1$ en $\bar{B}_2$ en ... en $\bar{B}_6$) = P($\bar{B}_1$) . P($\bar{B}_2$) ... P($\bar{B}_6$ = $(\frac{5}{6})^6$, immers de worpen zijn onafhankelijk uitgevoerd, zodat de gebeurtenissen $\bar{B}_1$, $\bar{B}_2$, ..., $\bar{B}_6$ onafhankelijk zijn. Maar dan volgt direkt, dat P(A) = 1 - $(\frac{5}{6})^6$.

§4. Een voorbeeld uit de genetica

We zullen de in §3 besproken elementaire kansrekening thans toepassen op de door MENDEL experimenteel gevonden kruisingswetten. In elk individu zijn de genen, de dragers van de erfelijke eigenschappen, in paren aanwezig. We zullen ons in het vervolg beperken tot de beschouwing van één paar genen. Bestaan van een gen in zo'n paar twee typen, A en a, dan zijn er drie mogelijke genotypen: AA, aa en Aa. In beide eerste gevallen spreekt men van homozygote, in het laatste geval van heterozygote individuen. Met het volgende model kan men de experimentele kruisingswetten nu goed verklaren:

a. een nakomeling van een heterozygoot individu erft met kans $\frac{1}{2}$ een A gen en met kans $\frac{1}{2}$ een a gen;
b. de overerving van genen van beide ouders geschiedt onafhankelijk;
c. de overdraging van genen aan verschillende nakomelingen geschiedt

onafhankelijk (afgezien van ééneiïge tweelingen, die dezelfde genen erven).

We zullen in het vervolg onderstellen, dat het gen A dominant en het gen a recessief is, zodat de genotypen Aa en AA hetzelfde phenotype hebben, maar het phenotype behorend bij aa afwijkt van het phenotype behorend bij AA en Aa. Als het genotype aa ernstige afwijkingen met zich meebrengt, is het van belang te weten wat men bij kruisingen mag verwachten.

We merken allereerst op, dat een nakomeling van twee AA-ouders weer AA is en van twee aa-ouders weer aa. Is de ene ouder AA en de andere ouder aa, dan is een nakomeling steeds heterozygoot. Als echter de ene ouder AA is en de ander Aa, dan staat het genotype van nakomelingen niet éénduidig vast. In dat geval zal een nakomeling met kans $\frac{1}{2}$ heterozygoot zijn en met kans $\frac{1}{2}$ homozygoot AA (zie punt a). Mutatis mutandis geldt hetzelfde voor nakomelingen van ouders van genotypen aa en Aa. Zijn beide ouders heterozygoot, dan is op grond van de modelonderstelling a en b een nakomeling met kans $\frac{1}{4}$ van genotype AA, met kans $\frac{1}{2}$ van genotype Aa en met kans $\frac{1}{4}$ van genotype aa (ga na!).

Stel nu, dat in een zeer grote populatie van volwassen individuen van het eerste phenotype (behorend bij AA en Aa) het genotype Aa voorkomt bij een fraktie λ en het genotype AA bij een fraktie $1-\lambda$ van de individuen. Dan is dus de kans, dat een aselekt getrokken individu van genotype Aa (resp. AA) is, gelijk aan λ (resp. $1-\lambda$). We zullen voorts onderstellen, dat de partnerkeuze in de populatie aselekt plaats vindt (!).

We berekenen nu in de eerste plaats de kans, dat een nakomeling van een ouderpaar uit deze populatie genotype aa heeft. Een dergelijk genotype kan uiteraard alleen optreden, als beide ouders heterozygoot zijn (aa-ouders komen in de populatie immers niet voor). De kans dat beide ouders genotype Aa bezitten (bij aselekte trekking van ouders uit de populatie) is gelijk aan λ^2, terwijl de kans dat een nakomeling van heterozygote ouders genotype aa bezit gelijk is aan $\frac{1}{4}$. De gevraagde kans is dus $\frac{1}{4}\lambda^2$.

In de tweede plaats vragen we ons af, hoe groot de kans is dat een nakomeling van ouders van het eerste phenotype, die broer en zuster zijn, het genotype aa bezit. In dit geval gaan we uit van aselekte partnerkeuze van de grootouders in dezelfde populatie. Opdat een nakomeling genotype aa

heeft, moeten beide ouders weer heterozygoot zijn; daar deze laatsten geen aselekt getrokken paar vormen uit de populatie (broer en zuster!), moeten we voor het berekenen van kansen terug gaan naar de grootouders. Zijn beide grootouders van genotype AA, dan kunnen de ouders niet heterozygoot zijn, zodat we dit geval buiten beschouwing kunnen laten. Andere mogelijke combinaties van de grootouders zijn: AA en Aa met kans $\lambda(1-\lambda)$, Aa en AA kans $\lambda(1-\lambda)$, Aa en Aa met kans λ^2 (genotype aa komt niet voor in de populatie). In alle drie gevallen zijn beide ouders heterozygoot met kans $(\frac{1}{2})^2$, zodat de kans, dat beide ouders heterozygoot zijn thans gelijk is aan $\frac{1}{2}\lambda - \frac{1}{4}\lambda^2$. Een nakomeling van deze ouders is dus van genotype aa met kans $\frac{1}{8}\lambda - \frac{1}{16}\lambda^2$. Deze redenering is nog niet geheel juist, immers grootouders van genotype Aa en Aa kunnen ook twee nakomelingen hebben van genotype Aa (of aa) en aa, die op hun beurt een nakomeling van genotype aa kunnen voortbrengen. Ter wille van een zo goed mogelijke vergelijkbaarheid met de eerst beschouwde situatie, hebben we echter als bekend ondersteld, dat de beide ouders van het eerste phenotype zijn (dus niet aa). Hiermee vervalt niet alleen de laatst genoemde mogelijkheid, doch dit gegeven heeft ook invloed op de reeds berekende kans. Men kan laten zien, dat in de hier beschouwde situatie de kans op een nakomeling van type aa gelijk is aan $(\frac{1}{8}\lambda - \frac{1}{16}\lambda^2)/(1 - \frac{7}{16}\lambda^2)$. Voor kleine waarden van λ valt $\frac{1}{16}\lambda^2$ in het niet vergeleken bij $\frac{1}{8}\lambda$, evenals $\frac{7}{16}\lambda^2$ vergeleken bij 1, zodat de gevraagde kans bij benadering gelijk is aan $\frac{1}{8}\lambda$ (bij kleine λ).

Als bijv. $\lambda = 0{,}02$, dan is $\frac{1}{4}\lambda^2 = 0{,}0001$ en $\frac{1}{8}\lambda = 0{,}0025$, zodat duidelijk blijkt dat de kans op nakomelingen van het genotype aa veel kleiner is in de beschouwde populatie bij aselekte partnerkeuze van de ouders dan bij ouders die broer en zuster zijn. Zijn individuen van het genotype aa ernstig gehandicapt, dan is het dus van biologisch standpunt veel gunstiger om kruising van nauwe verwanten te voorkomen.

§5. Kansverdeling van stochastische grootheden

We beschouwen een diskrete stochastische grootheid $\underline{x}$. Neem eens aan dat deze alleen de waarden $x_1, x_2, \ldots, x_m$ kan aannemen. Elk van deze waarden zal met een bepaalde kans worden aangenomen. De kans p_i, dat $\underline{x}$ de waarde x_i aanneemt, noteren we dan als $p_i = P(\underline{x} = x_i), i = 1, 2, \ldots, m$.

Daar $\underline{x}$ maar één van deze waarden tegelijk kan aannemen en altijd één van deze waarden aanneemt, geldt steeds (zie (iv), §3)

$$\sum_{i=1}^{m} p_i = 1 \quad \text{of wel} \quad \sum_{i=1}^{m} P(\underline{x} = x_i) = 1.$$

De kansen $p_1, p_2, \ldots, p_m$ beschrijven tezamen de kansen, waarmee $\underline{x}$ bepaalde waarden aanneemt: ze karakteriseren de kansverdeling van $\underline{x}$. Een dergelijke kansverdeling kan men grafisch weergeven in een diagram: op de horizontale as zet men de waarden uit die $\underline{x}$ kan aannemen, terwijl in elk van deze punten een verticaal streepje wordt opgericht met als lengte de grootte van de kans waarmee deze waarde wordt aangenomen. Indien $\underline{x}$ alleen een aantal opeenvolgende gehele getallen als waarden kan aannemen, dan kan men de kansverdeling ook weer m.b.v. een histogram weergeven; op de vertikale as zet men dan kansen uit i.p.v. relatieve frekwenties.

Voorbeeld 2.13. Zij $\underline{x}$ het aantal ogen verkregen bij één worp met een zuivere dobbelsteen. Dan kan $\underline{x}$ de waarden 1, 2, ..., 6 aannemen, elk met kans $\frac{1}{6}$. De kansverdeling van $\underline{x}$ is grafisch weergegeven in fig. 2.7a en fig. 2.7b.

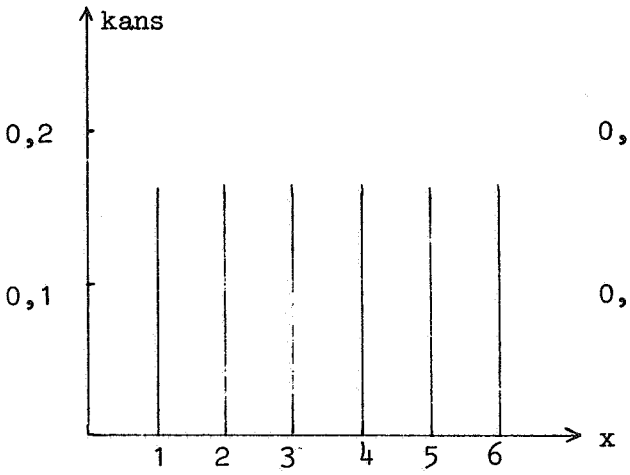


fig. 2.7a. Kansverdeling van $\underline{x}$ uit vb. 2.13.

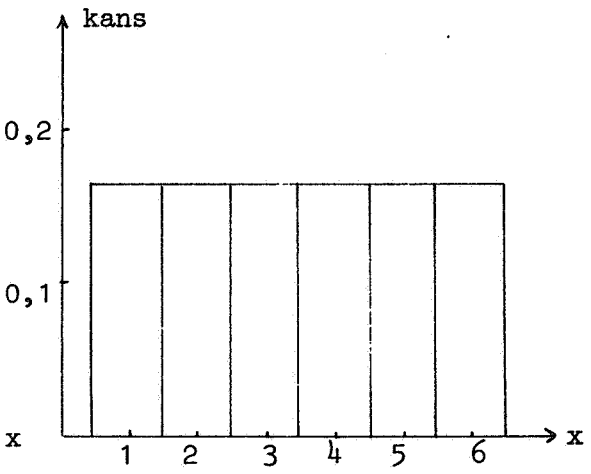


fig. 2.7b. Kansverdeling van $\underline{x}$ uit vb. 2.13.

Voorbeeld 2.14. Beschouw een vaas met evenveel rode als witte knikkers. Men trekt tienmaal aselekt met teruglegging een knikker uit de vaas. Zij $\underline{x}$ het aantal rode knikkers in deze steekproef. De kansverdeling van $\underline{x}$, die we later nog zullen bespreken (§11), is weergegeven in fig, 2.8a en fig. 2.8b.

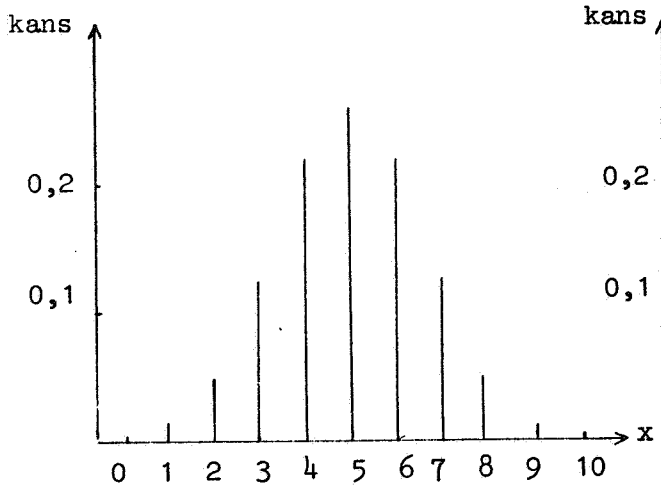


fig. 2.8a. Kansverdeling van x uit vb. 2.14.

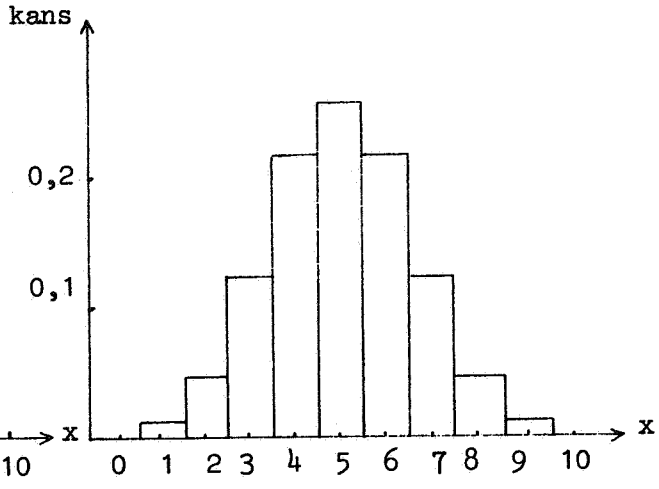


fig. 2.8b. Kansverdeling van x uit vb. 2.14.

Voorbeeld 2.15. We beschouwen weer de vaas met witte en rode knikkers, doch onderstellen nu, dat er vier maal zoveel rode als witte knikkers in de vaas zijn. De kansverdeling van x, het aantal rode knikkers in een steekproef met teruglegging van omvang 10, is weergegeven in de figuren 2.9a en 2.9b.

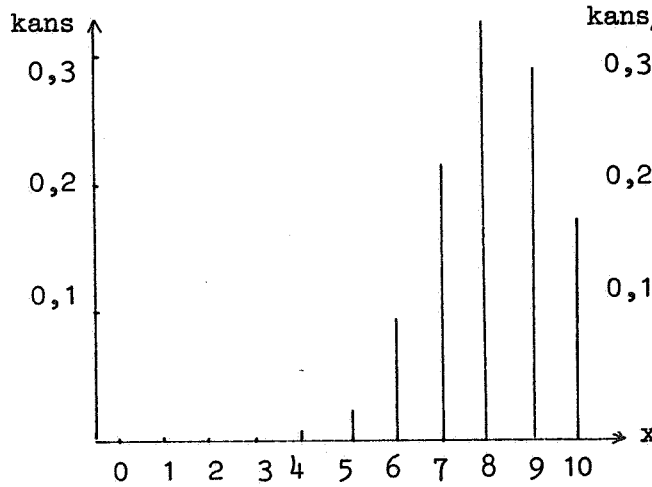


fig. 2.9a. Kansverdeling van x uit vb. 2.15.

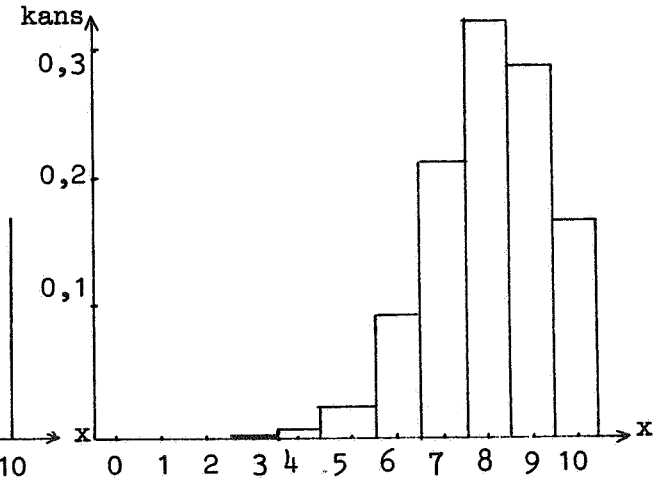


fig. 2.9b. Kansverdeling van x uit vb. 2.15.

Beschouwen we de figuren 2.8 en 2.9 wat nader, dan zien we, dat de kansverdeling van $\underline{x}$ uit vb. 2.14 symmetrisch is om het punt $x = 5$, terwijl de kansverdeling van $\underline{x}$ uit vb. 2.15 scheef is. Ook de kansverdeling van $\underline{x}$ uit vb. 2.13 zou men symmetrisch kunnen noemen om het punt $x = 3\frac{1}{2}$. Een dergelijke kansverdeling, waarbij elke mogelijke waarde met dezelfde kans wordt aangenomen, wordt een uniforme verdeling genoemd.

Bij een gegeven kansverdeling van $\underline{x}$ kan men nu alle mogelijke kansen berekenen. Is $\underline{x}$ diskreet, dan is bijv. voor elke waarde van x de kans $P(\underline{x} \leq x)$ gelijk aan de som van alle p_i's behorend bij x_i's die voldoen aan $x_i \leq x$. Men noemt $P(\underline{x} \leq x)$ als funktie van x wel de verdelingsfunktie van $\underline{x}$

Zij $\underline{x}$ thans een continue stochastische grootheid. In welke situaties treden continue stochastische grootheden op? Indien men een lengte, gewicht, temperatuur, stroomsterkte of tijdsverloop tussen twee gebeurtenissen meet, dan is het aantal mogelijke verschillende uitkomsten in principe oneindig groot, nl. alle reële getallen in een zeker (evt. onbegrensd) interval. Ook dergelijke waarnemingen, waarvan de uitkomsten een schijn van grote precisie suggereren, kunnen een stochastisch karakter hebben, d.w.z. bij herhaling van hetzelfde experiment tot verschillende resultaten leiden. Deze variabiliteit kan bijv. veroorzaakt worden door een bij de meting optredende toevallige meetfout of door een oncontroleerbare verandering van de omstandigheden waaronder het experiment wordt uitgevoerd. De populatie bestaat nu uit de verzameling van alle mogelijke waarnemingsuitkomsten, die men met dezelfde meetapparatuur onder zo goed mogelijk gelijk gehouden omstandigheden zou kunnen verkrijgen. Voert men een dergelijk experiment uit, dan beschouwt men de waarneming als een aselekte trekking uit deze fiktieve populatie. Herhaalt men een dergelijk experiment enige malen, er daarbij zorg voor dragend dat de waarnemingen elkaar niet beïnvloeden, dan kan men de waarnemingen veelal beschouwen als een aselekte steekproef uit de denkbeeldige populatie. In deze situatie, die vooral bij fysische metingen vaak optreedt, zijn de waarnemingen nu op te vatten als continue stochastische grootheden.

Uiteraard kan men in de praktijk nooit metingen verrichten met een oneindig grote precisie, zodat niet elk reëel getal in een zeker interval als uitkomst kan optreden. In vele gevallen vormt dit "model" echter toch

een zo goede beschrijving (benadering) van de werkelijke situatie, dat men hier van uit kan gaan bij de statistische analyse.

Is op alle elementen van een zeer grote populatie Π een getal gedefinieerd (bijv. een lengte of gewicht), dat alle reële waarden in een bepaald interval kan aannemen en dat waargenomen kan worden na aselekte trekking van een individu uit Π, dan is deze waarneming eveneens op te vatten als een continue stochastische grootheid, ook al meten we zonder meetfout. Het is duidelijk, dat thans een aselekte steekproef van waarnemingen wordt verkregen, indien men de te onderzoeken elementen aselekt uit Π trekt. Deze situatie doet zich vaak voor bij het onderzoek van grote biologische populaties. Een voorbeeld hiervan hebben we reeds ontmoet in §2 (zie vb. 2.7).

De kansverdeling van een continue stochastische grootheid $\underline{x}$ kunnen we karakteriseren m.b.v. een grafiek. Bij elke continue kansverdeling behoort nl. een funktie $p(x)$ van x met de volgende eigenschappen:

a. $p(x) \geq 0$ voor alle x, d.w.z. de grafiek van de funktie $p(x)$ ligt geheel boven (of op) de x-as.

b. het oppervlak tussen de x-as en de grafiek van de funktie $p(x)$ is precies 1.

De kans $P(\underline{x} \leq x_0)$ is nu gelijk aan het oppervlak links van het punt x_0 onder de grafiek van de funktie $p(x)$. In de figuren 2.10 en 2.11 zijn voor twee kansverdelingen de funkties $p(x)$ geschetst en de kansen $P(\underline{x} \leq x_0)$ voor zekere x_0 door arcering van het betreffende oppervlak aangegeven.

Men noemt de funktie $p(x)$, die de kansverdeling van $\underline{x}$ karakteriseert, de <u>verdelingsdichtheid</u> van $\underline{x}$. De verdelingsfunktie $P(\underline{x} \leq x)$ is blijkens bovenstaande kennelijk de bepaalde integraal van de verdelingsdichtheid tussen de grenzen $-\infty$ en x. Aldus is de verdelingsfunktie direkt bepaald door de verdelingsdichtheid $p(x)$.

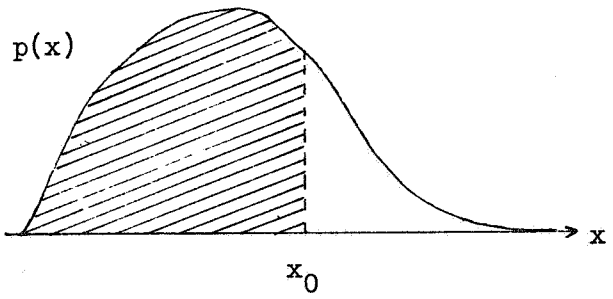


fig. 2.10. Verdelingsdichtheid van een scheve kansverdeling. Het gearceerde oppervlak is gelijk aan $P(\underline{x} \leq x_0)$.

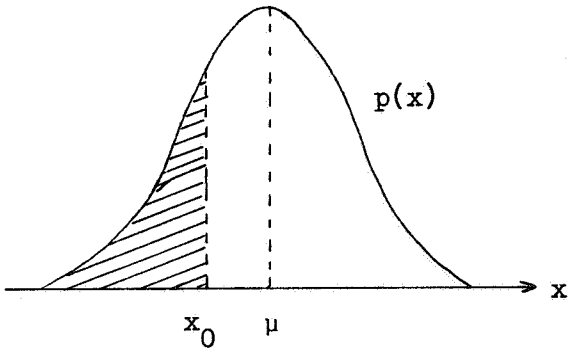


fig. 2.11. Verdelingsdichtheid van een symmetrische kansverdeling. Het gearceerde oppervlak is gelijk aan $P(\underline{x} \leq x_0)$.

Niet alleen een kans $P(\underline{x} \leq x_0)$, maar ook een kans $P(\underline{x} > x_0)$ is uit de grafiek af te lezen. Deze laatste kans is nl. gelijk aan het oppervlak onder de grafiek van p(x) rechts van het punt x_0. Daar het totale oppervlak onder grafiek van p(x) gelijk is aan 1, blijkt $P(\underline{x} \leq x_0) + P(\underline{x} > x_0) = 1$, in overeenstemming met het feit dat de eventualiteiten $\underline{x} \leq x_0$ en $\underline{x} > x_0$ disjunkt zijn. altijd één van beiden optreden en de som van hun kansen dus 1 moet zijn.

Tenslotte is een kans $P(x_1 \leq \underline{x} \leq x_2)$ eveneens uit de grafiek af te lezen. Deze is nl. gelijk aan het oppervlak onder de grafiek van p(x) tussen de punten x_1 en x_2 (mits $x_1 < x_2$ is, anders is de kans nul!).

Bij een continue kansverdeling is de kans $P(\underline{x} = x_0)$, dat $\underline{x}$ de waarde x_0 aanneemt, voor elke x_0 gelijk aan nul. Dit volgt uit de interpretatie van een kans als een oppervlak. Hieruit blijkt tevens, dat bij een continue stochastische grootheid $\underline{x}$ de kansen $P(\underline{x} \leq x_0)$ en $P(\underline{x} < x_0)$ gelijk zijn; evenzo is $P(\underline{x} \geq x_0) = P(\underline{x} > x_0)$, etc.

Beschouwen we fig. 2.10 wat nader, dan zien we dat de funktie p(x) niet symmetrisch is; de kansverdeling wordt daarom scheef genoemd. Bovendien suggereert de grafiek, dat p(x) voor toenemende x steeds dichter tot nul nadert zonder ooit exakt nul te worden. Dit betekent, dat $\underline{x}$ willekeurig grote waarden nog met een (zeer) kleine kans aanneemt. De kansverdeling uit

fig. 2.11 is daarentegen symmetrisch om het symmetriepunt μ, terwijl $\underline{x}$ alleen waarden in een begrensd interval kan aannemen.

Het werkt soms verhelderend bij een kansverdeling te denken aan de verdeling van één gram massa. Bij een diskrete grootheid $\underline{x}$, die alleen de waarden $x_1, x_2, \ldots, x_m$ kan aannemen met bepaalde kansen, denkt men zich de gram massa over deze diskrete punten verdeeld naar rato van de bijbehorende kansen. Bij een continue grootheid $\underline{x}$ denkt men zich de gram massa uitgesmeerd over de x-as met een dikte evenredig aan de grootte van de kansdichtheid ter plaatse.

De waarde van een stochastische grootheid bij één experiment valt niet precies te voorspellen. Wel kunnen we een soort doorsnee-waarde aangeven. Hebben we nl. te maken met een populatie van elementen op elk waarvan een getal x is gedefinieerd, dan kunnen we deze getallen x middelen over alle populatie-elementen. Het aldus ontstane populatiegemiddelde van de grootheid $\underline{x}$ noemen we de verwachting (Engels: "expectation" of "mean") van $\underline{x}$, aangeduid met $E\underline{x}$. Is $\underline{x}$ een diskrete grootheid, die slechts de waarden $x_1, x_2, \ldots, x_m$ kan aannemen, dan is $p_i = P(\underline{x} = x_i)$ de fraktie van de populatie-elementen waarop $\underline{x}$ de waarde x_i heeft. Hieruit volgt, dat in dit geval

$$E\underline{x} = \sum_{i=1}^{m} x_i \, P(\underline{x} = x_i).$$

Is $\underline{x}$ een continue grootheid, dan is de mathematische definitie van $E\underline{x}$ wat ingewikkelder *). Maar ook in dat geval kan men $E\underline{x}$ steeds interpreteren als een populatiegemiddelde. Wij vestigen er de aandacht op, dat $\underline{x}$ een stochastische grootheid is, dus aan toevalsfluctuaties onderhevig, terwijl $E\underline{x}$ een vast getal is.

We kunnen de verwachting $E\underline{x}$ ook anders interpreteren. Hierbij denken we aan de analogie tussen de verdeling van een gram massa en een kansverdeling. Met de verwachting van $\underline{x}$ correspondeert nu het zwaartepunt van de massaverdeling.

---

*) Als $\underline{x}$ een continue stochastische grootheid is met kansdichtheid $p(x)$, dan definieert men de verwachting van $\underline{x}$ door $E\underline{x} = \int_{-\infty}^{\infty} x \, p(x)dx$.

De gemiddelde waarde van $\underline{x}$ in een lange reeks experimenten blijkt doorgaans dicht bij het getal $E\underline{x}$ te liggen, en wel des te dichter naarmate de reeks langer is. Weging van 10 aselekt gekozen volwassen manlijke Nederlanders geeft meestal een gemiddelde tussen 70 en 80 kg, maar zo nu en dan zal het gemiddelde de 85 kg overschrijden. Het gemiddelde gewicht van 1000 Nederlandse mannen zal echter bij nagenoeg elk aselekt getrokken duizendtal minder dan 1 kg van het populatiegemiddelde $E\underline{x}$ afwijken. Deze statistische regelmaat, die van hetzelfde aard is als besproken in §2, is van groot belang in de statistiek.

Is de kansverdeling van een (diskrete of continue) stochastische grootheid $\underline{x}$ symmetrisch, dan valt $E\underline{x}$ samen met het symmetriepunt. In dat geval is de verwachting dus een goede maat voor het centrum van de kansverdeling. Is de kansverdeling van $\underline{x}$ echter scheef, dan is dit niet steeds het geval. Denken we bijv. aan de inkomstenverdeling van Nederlandse kostwinners, dan zal het populatiegemiddelde door een kleine groep zeer grote inkomens omhoog getrokken worden en de grote meerderheid van de kostwinners zal minder verdienen dan het populatiegemiddelde (d.i. het verwachte inkomen). In dergelijke gevallen kan men het centrum van een kansverdeling beter beschrijven met de mediaan M van de kansverdeling, d.i. een getal met de eigenschap dat bij de helft van de populatie-elementen de grootheid een grotere waarde en bij de andere helft een kleinere waarde aanneemt, m.a.w.

$$P(\underline{x} < M) = P(\underline{x} > M) = \frac{1}{2}$$

(bij een diskrete grootheid $\underline{x}$ is dit niet altijd mogelijk en is een kleine wijziging van de definitie van de mediaan nodig). In de statistiek speelt de mediaan een minder grote rol dan de verwachting, omdat de mediaan wiskundig lastiger te hanteren is.

Vaak wil men ook aangeven in welke mate een willekeurige waarde van $\underline{x}$ van de verwachting $E\underline{x}$ zal afwijken. Zo maakt het voor een kippefokker een groot verschil of alle kippeëieren ongeveer hetzelfde gewicht hebben (dus in gewicht weinig van het populatiegemiddelde afwijken) dan wel zeer sterk in gewicht variëren. Als maat voor de afwijking tussen $\underline{x}$ en $E\underline{x}$ gebruikt men de verwachting van de kwadratische afwijking, variantie ge-

naamd en aangeduid met $\sigma^2(\underline{x})$. Men definieert als het ware een nieuwe stochastische grootheid $(\underline{x} - E\underline{x})^2$ op dezelfde populatie als $\underline{x}$ en berekent daarvan het populatiegemiddelde:

$$\sigma^2(\underline{x}) = E(\underline{x} - E\underline{x})^2.$$

Uiteraard is $\sigma^2(\underline{x}) \geq 0$ voor elke stochastische grootheid $\underline{x}$. Uit de definitie blijkt, dat de eenheid waarin de variantie wordt uitgedrukt het kwadraat is van de eenheid waarin $\underline{x}$ is uitgedrukt. De positieve vierkantswortel uit de variantie is de <u>standaardafwijking</u> $\sigma(x)$. Deze heeft dezelfde "dimensie" als $\underline{x}$.

De volgende belangrijke eigenschappen van verwachtingen en varianties geven we zonder bewijs:

$E(c + \underline{x}) = c + E\underline{x}$ voor elke constante c;

$E(c\underline{x}) = c \cdot E\underline{x}$;

$E(\underline{x} + \underline{y}) = E\underline{x} + E\underline{y}$

(algemener: $E(\sum_{i=1}^{n} \underline{x}_i) = \sum_{i=1}^{n} E\underline{x}_i$);

$\sigma^2(c + \underline{x}) = \sigma^2(\underline{x})$;

$\sigma^2(c\underline{x}) = c^2 \cdot \sigma^2(\underline{x})$.

Men noemt twee stochastische grootheden <u>onafhankelijk</u>, als de getalwaarde van de ene grootheid bij een experiment geen invloed heeft op de kansverdeling van de andere grootheid. Zo zijn de aantallen ogen bij twee opeenvolgende worpen met een zuivere dobbelsteen wel onafhankelijk; ook bij de tweede worp heeft elk der waarden 1 t/m 6 een kans $\frac{1}{6}$, ongeacht de uitkomst van de eerste worp. Beschouwen we echter de populatie van alle Nederlanders, dan zijn lengte $\underline{l}$ en gewicht $\underline{g}$ géén onafhankelijke stochastische grootheden, want lange personen zullen in het algemeen zwaarder zijn dan kleine.

Als $\underline{x}$ en $\underline{y}$ onafhankelijke stochastische grootheden zijn, dan geldt

$$\sigma^2(\underline{x} + \underline{y}) = \sigma^2(\underline{x}) + \sigma^2(\underline{y}).$$

Algemener, als $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ alle onderling onafhankelijke stochastische grootheden zijn, is

$$\sigma^2(\sum_{i=1}^{n} \underline{x}_i) = \sum_{i=1}^{n} \sigma^2(\underline{x}_i).$$

Als de stochastische grootheden afhankelijk zijn, behoeven deze relaties niet te gelden. In het extreme geval $\underline{y} = 1 - \underline{x}$ bijv. heeft de som van $\underline{x}$ en $\underline{y}$ steeds de waarde 1, dus variantie 0, ook al zijn de varianties van $\underline{x}$ en $\underline{y}$ wel degelijk groter dan nul.

Als toepassing van deze formules bepalen we nu verwachting en variantie van een steekproefgemiddelde. Laat $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ een n-tal waarnemingen voorstellen van een stochastische grootheid $\underline{x}$, bijv. de uitkomsten van een n maal uitgevoerd experiment of waarneming van een grootheid aan n aselekt gekozen elementen uit een grote populatie. We onderstellen dan dat $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ alle onafhankelijk zijn en dezelfde kansverdeling hebben (als $\underline{x}$). Dan geldt voor elke index i uit de rij 1 t/m n, dat $E\underline{x}_i = E\underline{x}$ en $\sigma^2(\underline{x}_i) = \sigma^2(\underline{x})$. Men definieert het steekproefgemiddelde als

$$\bar{\underline{x}} = \frac{1}{n} \sum_{i=1}^{n} \underline{x}_i.$$

Dan geldt

$$E\bar{\underline{x}} = E(\frac{1}{n} \sum_{i=1}^{n} \underline{x}_i) = \frac{1}{n} E(\sum_{i=1}^{n} \underline{x}_i) = \frac{1}{n} \sum_{i=1}^{n} E\underline{x} = E\underline{x}.$$

Anderzijds is

$$\sigma^2(\bar{\underline{x}}) = \sigma^2(\frac{1}{n} \sum_{i=1}^{n} \underline{x}_i) = \frac{1}{n^2} \sigma^2(\sum_{i=1}^{n} \underline{x}_i) = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2(\underline{x}_i) = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2(\underline{x}) =$$

$$= \frac{1}{n^2} \cdot n\sigma^2(\underline{x}) = \frac{\sigma^2(x)}{n}.$$

Voorbeeld 2.16. Men voert één worp uit met een zuivere dobbelsteen. Zij $\underline{x}$ het aantal ogen bij deze worp. Gevraagd wordt de verwachting en standaard-

afwijking van $\underline{x}$ te bepalen. Daar elk van de mogelijke uitkomsten 1, 2, ..., 6 met kans $\frac{1}{6}$ optreedt, vinden we direkt

$$E\underline{x} = \frac{1}{6}(1 + 2 + \ldots + 6) = 3\tfrac{1}{2},$$

$$\sigma^2(\underline{x}) = E(\underline{x} - E\underline{x})^2 = E(\underline{x}^2 - 2\underline{x}E\underline{x} + (E\underline{x})^2) = E\underline{x}^2 - (E\underline{x})^2 =$$

$$= \frac{1}{6}(1^2 + 2^2 + \ldots + 6^2) - (3\tfrac{1}{2})^2 = 91/6 - 49/4 = 35/12,$$

zodat

$$\sigma(\underline{x}) = \sqrt{35/12} = \frac{1}{6}\sqrt{105} \approx 1{,}71.$$

(De in bovenstaande toegepaste schrijfwijze van $\sigma^2(\underline{x})$ is bij berekeningen vaak handig.)

Voorbeeld 2.17. Men voert twee onafhankelijke worpen met één zuivere dobbelsteen uit. Zij $\underline{z}$ de som van het aantal ogen bij beide worpen. Bereken verwachting en standaardafwijking van $\underline{z}$. Zij $\underline{x}$ het aantal ogen bij de eerste en $\underline{y}$ het aantal ogen bij de tweede worp. Dan is $\underline{z} = \underline{x} + \underline{y}$ en $\underline{x}$ en $\underline{y}$ hebben dezelfde verwachting en variantie (zie vb. 2.16). Uit de rekenregel volgt nu

$$E\underline{z} = E(\underline{x} + \underline{y}) = E\underline{x} + E\underline{y} = 7,$$

$$\sigma^2(\underline{z}) = \sigma^2(\underline{x} + \underline{y}) = \sigma^2(\underline{x}) + \sigma^2(\underline{y}) = 35/6,$$

zodat

$$\sigma(\underline{z}) = \frac{1}{6}\sqrt{210}.$$

§6. De normale verdeling

Een zeer belangrijke continue kansverdeling is de standaard-normale verdeling. Een stochastische grootheid met deze kansverdeling zullen we vaak aanduiden met $\underline{u}$. De kansdichtheid van deze verdeling is geschetst in fig. 2.12; mathematisch heeft deze kansdichtheid de gedaante

$$p(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} .$$

Men kan bewijzen dat het oppervlak onder de grafiek van p(u) 1 is, zoals een kansdichtheid betaamt.

Beschouwen we fig. 2.12 wat nader, dan zien we dat de grafiek van de kansdichtheid symmetrisch is om het punt 0 en zich uitstrekt van $-\infty$ tot $+\infty$. Deze symmetrie impliceert, dat de verwachting $E\underline{u}$ en de mediaan van de kansverdeling samenvallen en gelijk zijn aan nul. Voorts is de standaardafwijking van $\underline{u}$ juist gelijk aan 1, d.i. de afstand van het centrum van de kansverdeling tot de plaatsen waar p(u) zijn buigpunten heeft.

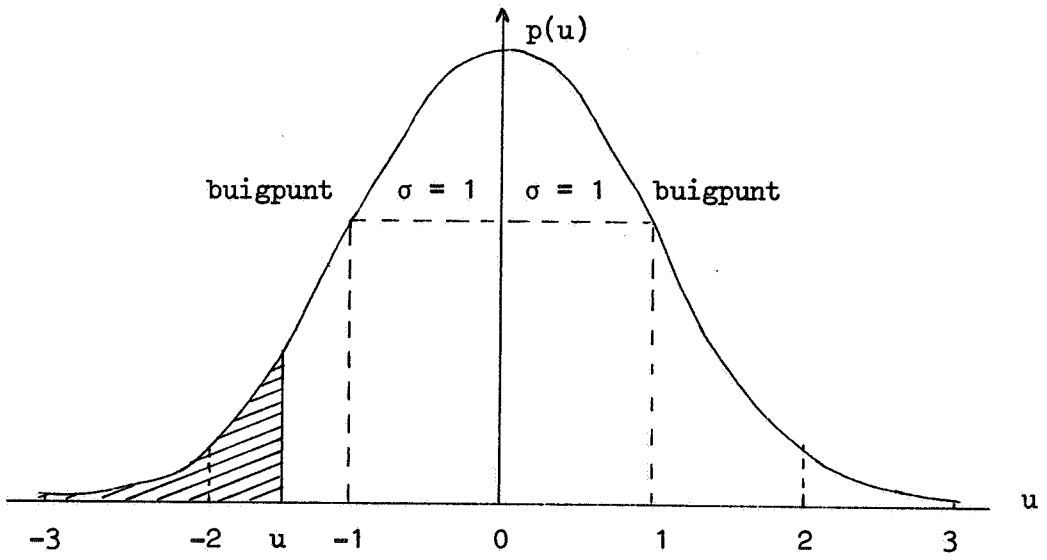


fig. 2.12. De kansdichtheid van de standaardnormale verdeling. Het gearceerde oppervlak is gelijk aan de kans $P(\underline{u} < u)$.

In principe kan men kansen van de gedaante $P(\underline{u} < u)$ nu bepalen uit de grafiek in fig. 2.12 door het oppervlak onder de grafiek van de kansdichtheid links van het punt u te berekenen; hetzelfde geldt voor kansen van de vorm $P(\underline{u} > u)$ of $P(u < \underline{u} < u_2)$. Praktisch is dit moeilijk uitvoerbaar. Voor een groot aantal waarden van u kan men kansen van de gedaante $P(\underline{u} \leq u)$ echter aflezen in tabel 1, achter in de syllabus.

Daar de kansverdeling van $\underline{u}$ symmetrisch is om 0, is $P(\underline{u} < -u) = P(\underline{u} > u)$ voor elke u. Anderzijds is het totale oppervlak onder de grafiek

van p(u) gelijk aan 1, zodat $P(\underline{u} > u) = 1 - P(\underline{u} < u)$. Deze eigenschappen stellen ons in staat kansen te berekenen, die niet in tabel 1 vermeld staan; deze tabel geeft immers alleen linkse kansen $P(\underline{u} < u)$ voor positieve waarden van u. Zo is

$$P(\underline{u} > 1) = 1 - P(\underline{u} \leq 1) = 0{,}1587;$$

$$P(\underline{u} > 3) = 1 - P(\underline{u} \leq 3) = 0{,}0013;$$

$$P(\underline{u} < -2) = P(\underline{u} > 2) = 1 - P(\underline{u} \leq 2) = 0{,}0228;$$

$$P(-2 < \underline{u} < 2) = P(\underline{u} < 2) - P(\underline{u} < -2) = 0{,}9772 - 0{,}0228 = 0{,}9544.$$

De grootheid $\underline{u}$ neemt dus met een kans van ruim 0,95 waarden aan tussen -2 en +2, terwijl waarden groter dan 3 (en kleiner dan -3) zeer onwaarschijnlijk zijn.

In tabel 1 zijn bij verschillende waarden van u de kansen $P(\underline{u} \leq u)$ vermeld. Omgekeerd is het vaak belangrijk bij een gegeven kans ter grootte α de waarde van u te bepalen, zodat $P(\underline{u} \geq u) = \alpha$. We noemen dergelijke waarden van u <u>rechtse α-punten</u> van de standaard-normale verdeling en geven ze aan met u_α, dus

$$P(\underline{u} \geq u_\alpha) = \alpha.$$

Daar $P(\underline{u} < u_\alpha) = 1 - \alpha$, kunnen we in tabel 1, door te zoeken bij kansen ter grootte $1 - \alpha$, deze rechtse α-punten ook wel vinden; zo is bijv. $u_{0,5} = 0$; $u_{0,05} = 1{,}645$; $u_{0,025} = 1{,}96$.

De grootheid $\underline{u}$ heeft verwachting 0 en variantie 1; de standaard-normale verdeling wordt daarom vaak aangeduid met N(0,1). Er bestaat echter een hele familie van normale verdelingen, waarvan de standaard-normale verdeling er één is. Telt men bijv. bij $\underline{u}$ een vast getal μ op, dan verschuift de kansdichtheid van de grootheid $\underline{u} + \mu$ over een afstand μ (naar rechts als $\mu > 0$, naar links als $\mu < o$). Men kan voorts $\underline{u}$ (of $\underline{u} + \mu$) ook in een andere schaal meten. Is bijv. $\underline{u}$ een lengte, die in meters gemeten standaard-normaal verdeeld is, en gaat men over op centimeters, dan dient men de horizontale schaal met een faktor 100 uit te rekken. Dit heeft echter tot gevolg, dat het oppervlak onder de grafiek van de aldus uitge-

rekte kansdichtheid niet meer 1, maar 100 is. Daarom dient men in dit geval de hoogte van de grafiek van de kansdichtheid met $\frac{1}{100}$ te vermenigvuldigen. De standaardafwijking gaat hierbij weer over van 100 in 1.

Alle kansverdelingen, die door verschuiving over een afstand μ of schaalverandering met een faktor σ uit de standaard-normale verdeling ontstaan, noemt men ook normale verdelingen. Bij elke μ en σ^2 behoort dus een normale verdeling, die door deze twee getallen wordt vastgelegd. Men noemt μ en σ^2 de parameters van de normale verdeling. Deze normale verdelingen worden vaak aangeduid met $N(\mu,\sigma^2)$ en zijn symmetrisch om het punt μ. De mediaan van een $N(\mu,\sigma^2)$ verdeling is dan ook gelijk aan de verwachting μ, terwijl σ^2 de variantie is van de bijbehorende stochastische grootheid.

Het is zeer eenvoudig kansen met betrekking tot een normaal verdeelde grootheid $\underline{x}$ uit te rekenen door reduktie op de $N(0,1)$ verdeling. Immers als $\underline{x}$ $N(\mu,\sigma^2)$ verdeeld is, dan is de grootheid

$$\frac{\underline{x} - \mu}{\sigma}$$

standaard-normaal verdeeld. Hieruit volgt dat voor elke x

$$P(\underline{x} < x) = P(\frac{\underline{x}-\mu}{\sigma} < \frac{x-\mu}{\sigma}) = P(\underline{u} < \frac{x-\mu}{\sigma})$$

en deze laatste kans kan men, bij gegeven x, μ en σ bepalen met behulp van de tabel van de $N(0,1)$ verdeling.

Voorbeeld 2.18. Bij het vullen van grote partijen pakjes met bloemzaadjes kan men de vulmachine op een bepaald gemiddeld gewicht μ instellen, bijv. 10 gram. Er blijft echter een zekere toevallige variatie in de vulgewichten $\underline{x}$ om μ bestaan; uit langdurige ervaring is bekend dat $\underline{x}$ bij goede benadering normaal verdeeld is met standaardafwijking $\sigma = 0{,}06$, althans voor instellingen van de machine in de buurt van 10 gram. Op de pakjes staat als netto-inhoud 10 gram vermeld; volgens een (fiktief) voorschrift van de Keuringsdienst van waren mogen dan niet meer dan 1% van de pakjes minder dan 10 gram zaadjes bevatten. Op welk gewicht μ dient de vulmachine nu te worden ingesteld?
Het is duidelijk dat $\mu = 10$ te laag is; immers dan is $P(\underline{x} < 10) = \frac{1}{2}$.

Kiest men $\mu = 10{,}1$, dan is $P(\underline{x} < 10) = P(\frac{\underline{x}-\mu}{\sigma} < \frac{10-\mu}{\sigma}) = P(\underline{u} < -1{,}67) = 1 - P(\underline{u} < 1{,}67) = 0{,}0475$ $(> 0{,}01)$, zodat ook deze instelwaarde nog te laag is. We berekenen μ rechtstreeks als volgt.
$P(\underline{x} < 10) = P(\underline{u} < \frac{10-\mu}{0{,}06})$; daar $P(\underline{u} < -2{,}33) = P(\underline{u} > 2{,}33) = 0{,}01$, moet dus $\frac{10-\mu}{0{,}06} \leq -2{,}33$ zijn ofwel $\mu \geq 10{,}14$. Men zal de vulmachine dus, zo zuinig mogelijk, op $\mu = 10{,}14$ gram afstellen.

We noemen nog de volgende belangrijke eigenschappen van normaal verdeelde grootheden:

(i) als $\underline{x}$ normaal verdeeld is en c een constante, dan zijn $c + \underline{x}$ en $c\underline{x}$ ook normaal verdeeld;

(ii) als $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ alle normaal verdeeld zijn, dan is hun som ook normaal verdeeld.

In de statistiek spelen de normale verdelingen een centrale rol. Hiervoor zijn twee redenen aan te voeren. In de eerste plaats is de meetfout bij fysische metingen vaak normaal verdeeld. De ware (onbekende) waarde van de te meten grootheid is μ, doch door de meetfout vindt men uitkomsten, die om μ varieren. Men gaat er hierbij van uit, dat de meetfout een $N(0,\sigma^2)$ verdeling heeft, dus niet tot systematisch te hoge of te lage waarden leidt. De onderzoeker zal nu uit een aantal uitkomsten deze onbekende waarde μ willen schatten en misschien ook hypothesen omtrent de waarde van μ willen toetsen. Ook vertonen kenmerken van objekten van grote biologische populaties vaak een normale verdeling, bijv. lengten van bruine bonen of van personen van een bepaald ras. In dit geval zal de onderzoeker door steekproef trekking informatie trachten te verkrijgen over het populatie-gemiddelde μ. Soms is ook de variantie σ^2 van belang.

De tweede reden is van meer wiskundige aard. Vele stochastische grootheden, die bij de statistische analyse van belang zijn en uit steekproeven van waarnemingen worden bepaald, blijken nl. verdelingen te bezitten, die bij grote steekproefomvang veel gelijken op normale verdelingen. Bij het berekenen van kansen maakt men hier dan dankbaar gebruik van.

§7. Schattingstheorie bij normale verdelingen

Zij $\underline{x}$ een normaal verdeelde stochastische grootheid, bijv. een fysische meting of een grootheid, die gedefinieerd is op de elementen van

een grote biologische populatie, met onbekende μ en σ^2. Hoe zal men het populatiegemiddelde μ en de variantie σ^2 nu kunnen schatten aan de hand van een aantal onafhankelijke waarnemingen $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ van deze grootheid (die dus alle dezelfde onbekende $N(\mu,\sigma^2)$ verdeling bezitten)?

Het ligt voor de hand de verwachting (populatiegemiddelde!) van $\underline{x}$ te schatten met het steekproefgemiddelde

$$\overline{\underline{x}} = \frac{1}{n} \sum_{i=1}^{n} \underline{x}_i .$$

De grootheid $\overline{\underline{x}}$, waarmee we μ willen schatten, is zelf ook weer een stochastische grootheid met een kansverdeling. We noemen $\overline{\underline{x}}$ een schatter (Engels: estimator) van μ. Na uitvoering van de experimenten en vaststelling van de uitkomsten, krijgt $\overline{\underline{x}}$ een bepaalde numerieke waarde $\overline{x}$, die we de schatting van μ noemen.

Aan het einde van §5 hebben we gezien, dat

$$E\overline{\underline{x}} = E\underline{x} = \mu$$

Dit betekent, dat μ het centrum van de kansverdeling van $\overline{\underline{x}}$ is, m.a.w. we schatten μ door $\overline{\underline{x}}$ niet systematisch te hoog of te laag. Als de verwachting (populatiegemiddelde) van een schatter gelijk is aan de (onbekende) waarde van de te schatten parameter, dan noemt men de schatter zuiver (Engels: unbiased). Dit is een gunstige eigenschap van een schatter. I.h.b. is $\overline{\underline{x}}$ een zuivere schatter van μ. Daar bovenstaande formule is afgeleid in §5 zonder gebruik te maken van de normaliteit van de waarnemingen, is $\overline{\underline{x}}$ ook een zuivere schatter van het populatiegemiddelde van $\underline{x}$ bij niet-normaal verdeelde waarnemingen!

Men zou μ echter ook wel op andere wijze kunnen schatten, bijv. met de middelste van de naar opklimmende grootte geordende $\underline{x}_i$'s (bij even n is er geen middelste; men zou dan het gemiddelde van de twee middelste waarnemingen kunnen nemen). Men kan laten zien, dat deze middelste waarneming ook een zuivere schatter van μ is. Welke schatter zal men nu kiezen?

In het algemeen kan men zeggen, dat men de voorkeur zal geven aan een schatter, waarvan de uitkomsten zo dicht mogelijk bij μ liggen. M.a.w. het is gunstig als de kansverdeling van de schatter van μ geconcentreerd is.

Daar de variantie een maat is voor de variabiliteit van een stochastische grootheid, is het dus om de genoemde reden gunstig onder de zuivere schatters te zoeken naar een schatter met een zo klein mogelijke variantie.
De variantie van $\bar{\underline{x}}$ bedraagt (zie §5)

$$\sigma^2(\bar{\underline{x}}) = \frac{\sigma^2}{n} .$$

Men kan nu laten zien, dat elke andere zuivere schatter van μ een grotere variantie heeft, zodat $\bar{\underline{x}}$ de best mogelijke schatter is van μ.
Dit geldt echter niet algemeen voor niet-normaal verdeelde waarnemingen!

Uit de formule voor de variantie van $\bar{\underline{x}}$ blijkt, dat $\sigma^2(\bar{\underline{x}})$ omgekeerd evenredig is met het aantal waarnemingen n, m.a.w. de nauwkeurigheid waarmee μ geschat wordt door $\bar{\underline{x}}$ neemt toe met toenemend aantal waarnemingen. De verdeling van $\bar{\underline{x}}$ is op grond van de eigenschappen (i) en (ii) uit §6 ook weer normaal en volgens bovenstaande formules dus $N(\mu, \frac{\sigma^2}{n})$ verdeeld. Voor een tweetal waarden van n zijn de kansdichtheden van $\bar{\underline{x}}$ geschetst in fig. 2.13, tezamen met de kansdichtheid van de $\underline{x}_i$'s zelf.

Als bijv. $\mu = 5$ en $\sigma^2 = 9$, dan is bij $n = 4$ de kans op een waarde van $\bar{\underline{x}}$ kleiner dan 4 gelijk aan (zie tabel 1):

$$P(\bar{\underline{x}} < 4) = P\left(\frac{\bar{\underline{x}} - \mu}{\sqrt{\sigma^2/n}} < \frac{4 - 5}{\sqrt{9/4}}\right) = P(\underline{u} < -\frac{2}{3}) = 0{,}252;$$

bij $n = 25$ wordt deze kans

$$P(\bar{\underline{x}} < 4) = P\left(\frac{\bar{\underline{x}} - \mu}{\sqrt{\sigma^2/n}} < \frac{4 - 5}{\sqrt{9/25}}\right) = P(\underline{u} < -\frac{5}{3}) = 0{,}048.$$

Hoe zullen we σ^2 schatten? De variantie σ^2 van $\underline{x}$ is de verwachte waarde van $(\underline{x} - \mu)^2$. Bij bekende μ is het daarom intuïtief aantrekkelijk om σ^2 te schatten met het corresponderende steekproefgemiddelde

$$\frac{1}{n} \sum_{i=1}^{n} (\underline{x}_i - \mu)^2.$$

Daar μ onbekend is ondersteld, zou men in deze uitdrukking μ kunnen schatten met $\bar{\underline{x}}$, hetgeen leidt tot de schatter

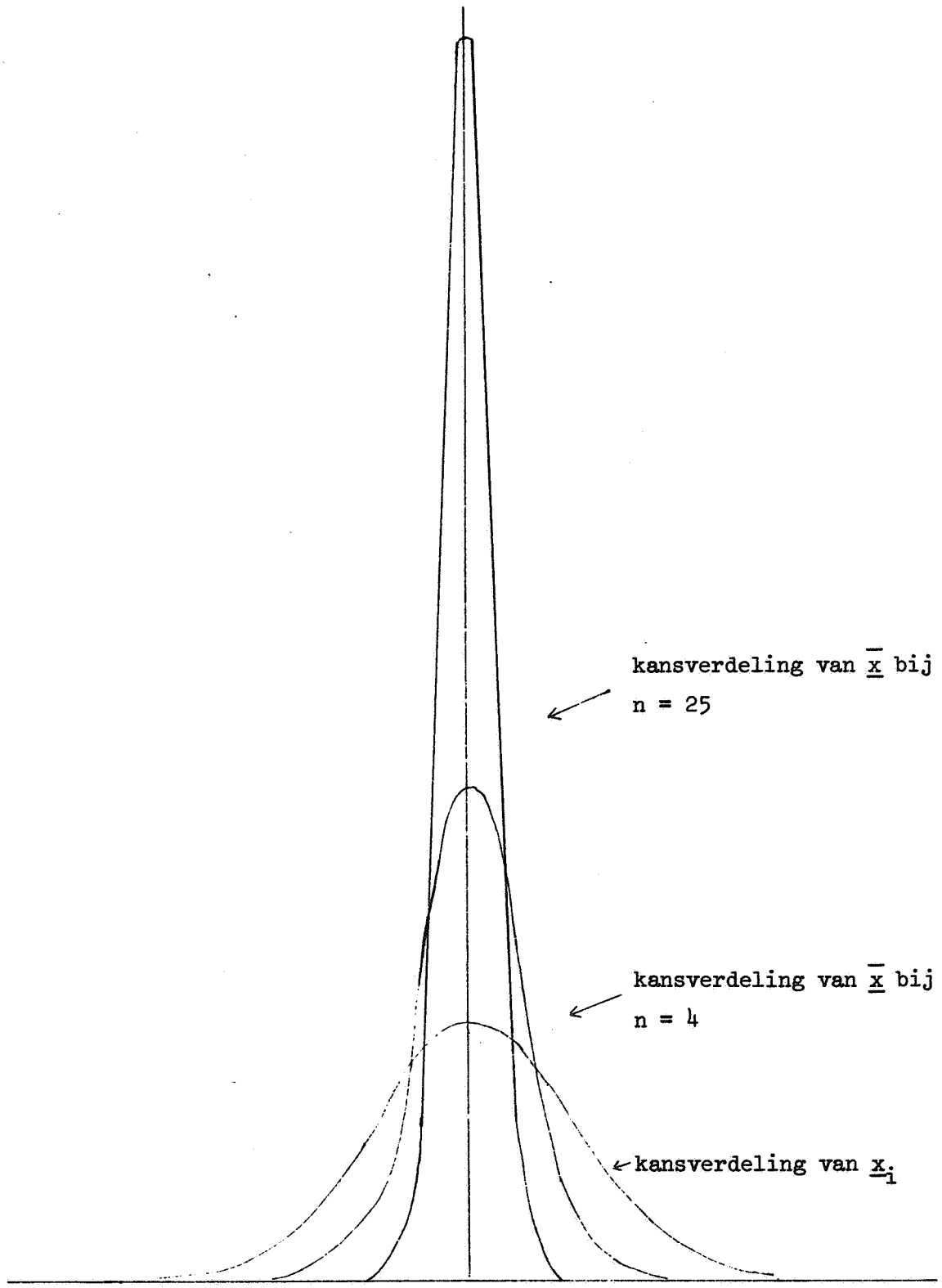


fig. 2.13. Kansverdeling van x_i, van $\bar{x}$ bij n = 4 en van $\bar{x}$ bij n = 25.

$$\underline{s}'^2 = \frac{1}{n} \sum_{i=1}^{n} (\underline{x}_i - \overline{\underline{x}})^2.$$

Men kan echter laten zien, dat de kwadratische vorm

$$\underline{S}^2 = \sum_{i=1}^{n} (\underline{x}_i - \overline{\underline{x}})^2$$

verwachting

$$E\underline{S}^2 = (n - 1)\sigma^2$$

heeft, zodat

$$E\underline{s}'^2 = \frac{n-1}{n} \sigma^2$$

en $\underline{s}'^2$ dus géén zuivere schatter van σ^2 is. Het is nu echter duidelijk, dat

$$\underline{s}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\underline{x}_i - \overline{\underline{x}})^2$$

wel een zuivere schatter is van σ^2; s^2 heet de steekproefvariantie. Men zegt wel, dat $\underline{S}^2$ een kwadratische vorm is met n-1 vrijheidsgraden. Een zuivere schatter van de variantie van de $\underline{x}_i$ ontstaat, als men de kwadratische vorm deelt door zijn vrijheidsgraden. De naam vrijheidsgraden wordt duidelijker, wanneer men bedenkt dat van de grootheden $x_1 - \overline{x}$, $x_2 - \overline{x}$, ..., $x_n - \overline{x}$ er slechts n-1 vrij gekozen kunnen worden, daar hun som nul is. Het aantal vrijheidsgraden van $\underline{S}^2$ is één kleiner dan het aantal waarnemingen, omdat in $\underline{S}^2$ de verwachtingen van de $\underline{x}_i$ (die hier alle gelijk zijn) geschat zijn door één parameter μ te schatten.

De verdeling van $\underline{s}^2$ is niet normaal, doch hangt samen met de zogenaamde chikwadraat-verdeling. We gaan hier verder niet op in en merken slechts op, dat de verdeling van $\underline{s}^2$ alleen van σ^2 en van het aantal vrijheidsgraden n-1 afhangt, doch niet van μ.

Uiteraard is $\underline{s}$, de positieve wortel van $\underline{s}^2$, een schatter van de standaardafwijking σ van $\underline{x}$.

We merken nog op, dat steekproefgemiddelde $\overline{\underline{x}}$ en steekproefvariantie $\underline{s}^2$ onafhankelijk zijn, d.w.z. de waarde van $\overline{\underline{x}}$ beïnvloedt de waarde, die $\underline{s}^2$

aanneemt, niet. Deze eigenschap, die alleen geldt voor normaal verdeelde waarnemingen $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ met dezelfde verwachting en variantie, speelt in de statistiek een grote rol.

Ook als de waarnemingen $\underline{x}_1, \ldots, \underline{x}_n$ niet normaal verdeeld zijn, is $\underline{s}^2$ een zuivere schatter van σ^2.

Schat men μ met $\bar{\underline{x}}$, dan kan men zich afvragen hoe nauwkeurig deze schatting is. Teneinde hiervan een indruk te geven zou men bijv. de standaardafwijking van $\bar{\underline{x}}$ kunnen vermelden, doch deze is gelijk aan $\sigma/\sqrt{n}$ en hangt dus af van de onbekende σ. Schat men σ met $\underline{s}$, dan geeft de uitkomst $s/\sqrt{n}$ wel een idee van de grootte van $\sigma(\bar{\underline{x}})$ en dus van de nauwkeurigheid waarmee men μ schat door $\bar{\underline{x}}$. Men noemt $s/\sqrt{n}$ de <u>standaardfout</u> van de schatting $\bar{x}$. In publikaties ziet men naast de schatting ook vaak de standaardfout op de volgende (verwarrende) wijze vermeld:

$$\bar{x} \pm s/\sqrt{n},$$

bijv. $2{,}53 \pm 0{,}08$.

Weliswaar geeft de standaardfout een zekere indruk van de nauwkeurigheid waarmee men μ schat, maar we kunnen er geen precieze uitspraken aan verbinden. Dit is wel mogelijk bij gebruikmaking van de zogenaamde <u>betrouwbaarheidsintervallen</u>.

Men berekent uit de waarnemingen volgens een bepaald voorschrift twee getallen μ_l en μ_r ($\mu_l < \mu_r$) en doet vervolgens de uitspraak dat de onbekende waarde van de parameter μ tussen de beide grenzen μ_l en μ_r ligt. Hoe men μ_l en μ_r ook bepaalt, deze uitspraak zal niet steeds juist zijn. Als in een fraktie α van een lange reeks gevallen, dat men de waarnemingen verricht en de grenzen μ_l en μ_r volgens hetzelfde voorschrift bepaalt, de waarde van μ niét in het interval (μ_l, μ_r) ligt, dan zegt men dat het interval (μ_l, μ_r) een <u>betrouwbaarheidsinterval</u> is voor de parameter μ met een <u>onbetrouwbaarheid</u> α. De onbetrouwbaarheid α, die steeds tussen 0 en 1 ligt, kan men zelf kiezen door geschikte keuze van het berekeningsvoorschrift voor μ_l en μ_r.

Bij een kleine waarde van α zal de uitspraak, dat μ in het berekende betrouwbaarheidsinterval ligt, bijna steeds juist zijn, zodat een kleine onbetrouwbaarheid gunstig lijkt. Hier tegenover staat echter, dat het be-

trouwbaarheidsinterval wijder wordt naarmate men α kleiner kiest, zodat de uitspraak, dat μ in dit interval ligt, steeds minder informatie omtrent μ geeft naarmate α kleiner is. Als compromis kiest men voor α gewoonlijk 1%, 5% of 10%.

We merken nog op, dat het betrouwbaarheidsinterval nauwer wordt met toenemende steekproefomvang n, zodat bij dezelfde onbetrouwbaarheid α het interval leidt tot steeds preciezere uitspraken omtrent de waarde van μ. Dit is geheel in overeenstemming met het feit, dat ook de schatting $\bar{x}$ van μ steeds nauwkeuriger wordt bij toenemende n.

Alvorens het berekeningsvoorschrift voor μ_l en μ_r te beschrijven, introduceren we een nieuwe familie van continue kansverdelingen, de t-verdelingen van STUDENT. Deze verdelingen hangen af van één parameter ν, het aantal vrijheidgraden van de t-verdeling ($\nu = 1,2,3,\ldots$). De grafiek van de kansdichtheid van een t-verdeling is symmetrisch om 0 en gelijkt bij toenemend aantal vrijheidsgraden ν steeds meer op de grafiek van de standaard-normale verdeling, doch de "staarten" van de t-verdelingen zijn iets dikker dan van de N(0,1) verdeling, zie fig. 2.14.

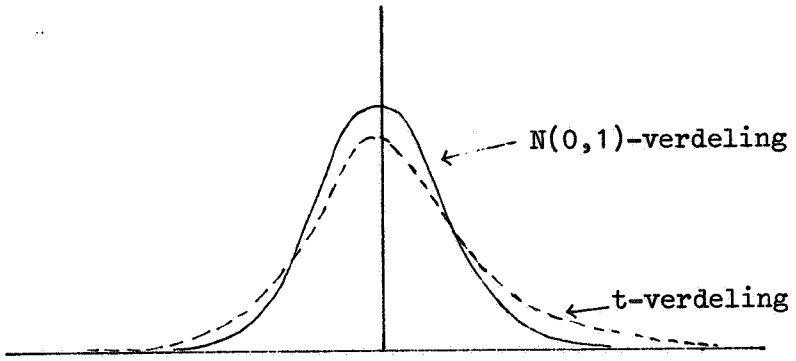


fig. 2.14. Grafieken van de kansdichtheden van de N(0,1) verdeling en een t-verdeling.

Een grootheid met een t-verdeling met ν vrijheidsgraden zullen we in het vervolg aangeven met $\underline{t}_\nu$. Het rechtse α-punt van een t-verdeling met ν vrijheidsgraden noteren we met $t_{\nu;\alpha}$. In tabel 2 achter in de syllabus staan voor een aantal vrijheidsgraden en enige waarden van α dergelijke rechtse α-punten vermeld. Als het aantal vrijheidsgraden onbeperkt toeneemt, naderen deze α-punten tot de overeenkomstige α-punten van de N(0,1)

verdeling. Voor grote ν zal men $t_{\nu;\alpha}$ dan ook benaderen met het rechtse α-punt u_α van de standaard-normale verdeling.

Een betrouwbaarheidsinterval voor μ bij gekozen onbetrouwbaarheid heeft nu als ondergrens

$$\mu_l = \bar{x} - t_{n-1;\frac{\alpha}{2}} \cdot s/\sqrt{n}$$

en als bovengrens

$$\mu_r = \bar{x} + t_{n-1;\frac{\alpha}{2}} \cdot s/\sqrt{n}.$$

Merk op, dat de beste schatting $\bar{x}$ van μ juist in het midden van dit betrouwbaarheidsinterval ligt.

Voorbeeld 2.19. Een kippefokker wenst een idee te krijgen van het gemiddelde gewicht van de door zijn kippen gelegde eieren. Hiertoe kiest hij aselekt 20 eieren uit de in een bepaalde periode gelegde eieren en weegt deze. De gevonden gewichten waren (in grammen)

40,3; 48,6; 45,1; 43,7; 42,5; 39,6; 44,0; 50,1; 47,7; 42,5;

42,3; 46,1; 44,2; 51,9; 46,0; 41,8; 43,2; 38,4; 49,5; 40,4.

De gewichten van alle door zijn kippen in deze periode gelegde eieren vormen dus de populatie. Een schatting van het populatiegemiddelde wordt verkregen door het steekproefgemiddelde te bepalen:

$$\bar{x} = 44{,}395 \text{ gram}.$$

De standaardafwijking $\sigma(\underline{x})$ van het gewicht van een aselekt gekozen ei schatten we met

$$s = \sqrt{s^2} = \sqrt{13{,}768} = 3{,}711 \text{ gram},$$

zodat de geschatte standaardafwijking van $\bar{x}$ gelijk is aan $s/\sqrt{20} = 0{,}830$ gram. Nemen we aan, dat de gewichten van de eieren bij benadering normaal verdeeld zijn, dan kunnen we een betrouwbaarheidsinterval voor μ bepalen.

Als onbetrouwbaarheid kiezen we $\alpha = 0{,}05$. Daar de steekproefomvang $n = 20$ bedraagt, moeten we nu $t_{19;0,025}$ opzoeken in de tabel; we vinden 2,093. Dit geeft

$$\mu_l = 44{,}395 - 2{,}093 \times 0{,}830 = 42{,}658 \approx 42{,}7 \text{ gram},$$
$$\mu_r = 44{,}395 + 2{,}093 \times 0{,}830 = 46{,}132 \approx 46{,}1 \text{ gram}.$$

Met een onbetrouwbaarheid van 5% kunnen we nu dus de uitspraak doen, dat het gemiddelde gewicht van alle eieren uit de populatie ligt tussen 42,7 en 46,1 gram.

§8. Toetsingstheorie bij normale verdelingen

Vaak is men niet zo zeer geïnteresseerd in een schatting van een onbekende parameter, maar in het toetsen van hypothesen omtrent de waarde van de parameter. Dit doet zich bijv. voor, indien men een nieuwe theorie aan de hand van waarnemingen experimenteel wil verifiëren. Hebben op de theorie betrekking hebbende waarnemingen een deterministisch karakter (zodat men bij herhaling van het experiment steeds dezelfde uitkomst vindt), dan kan één experiment voldoende zijn om het gestelde doel te bereiken. Hebben de waarnemingen echter een stochastisch karakter, bijv. door het optreden van toevallige meetfouten of door variatie binnen een grote populatie, dan is het trekken van conclusies veel lastiger. Experimentele verificatie van een theorie, dat meisjes eerder het melkgebit verliezen dan jongens, of dat roken het ontstaan van longkanker bevordert, of dat sperwers zich aanpassen aan hun omgeving door bepaalde veranderingen in hun lichaamsbouw, is niet eenvoudig en statistische methoden kunnen hierbij goede diensten bewijzen. Maar ook al verzamelt men een aselekte steekproef van waarnemingen, volmaakt ondubbelzinnig is de uitslag in zo'n geval nooit. Wanneer de waarnemingen echter veel beter in overeenstemming zijn met de geponeerde theorie dan met haar tegendeel betekent dit een statistische bevestiging. Deze is minder sterk dan een logisch bewijs, dat ieder overtuigt tenzij er fouten tegen de logika in optreden. Bij een statistische bevestiging van de vroegere gebitswisseling van meisjes bestaat altijd een kleine kans, dat de meisjes in onze steekproef toevallig eerder

wisselen dan de jongens, hoewel dat in de populatie niet het geval was. Op dit risiko van fouten komen wij nog terug.

Beschouwen we eens het bloedsuikergehalte van personen, die aan een bepaalde ziekte lijden. Men vermoedt dat dit bloedsuikergehalte hoger ligt bij personen die aan de ziekte lijden dan bij gezonde personen.
We veronderstellen, dat uit ervaring bekend is dat bij gezonde personen het gemiddelde bloedsuikergehalte μ_0 mg per 100 ml bedraagt, waarin μ_0 een bekend getal is. We kunnen nu, om het vermoeden experimenteel te onderzoeken, een steekproef trekken uit een grote groep van lijders aan de ziekte en bij elke persoon een bepaling van het bloedsuikergehalte verrichten. We hebben hier ten duidelijkste te maken met stochastische waarnemingen, daar het bloedsuikergehalte van persoon tot persoon enigszins varieert, terwijl men bij de meting bovendien nog te maken krijgt met meetfouten. Liggen de uitkomsten door elkaar genomen flink boven de μ_0, dan zal men niettemin geneigd zijn te zeggen, dat het vermoeden experimenteel bevestigd is. Fluctueren de uitkomsten echter om de μ_0 of liggen ze zelfs lager, dan wordt het vermoeden door de waarnemingen niet bevestigd. Het precizeren van een dergelijke min of meer intuïtieve (en dus ook subjektieve) conclusie is het doel van de statistische toetsingstheorie.

Een statistische <u>toets</u> is een procedure om na te gaan in hoeverre waarnemingen in overeenstemming zijn met een tevoren opgestelde hypothese omtrent een populatiegrootheid (meestal een parameter van een kansverdeling, bijv. de verwachting, d.i. het populatiegemiddelde). Deze te toetsen hypothese wordt <u>nulhypothese</u> genoemd en met H_0 aangeduid. Bij slechte overeenstemming tussen H_0 en de waarnemingen verwerpt men H_0. Bij een goede overeenstemming wordt H_0 <u>niet verworpen</u>. Dit wil niet zeggen dat H_0 nu bewezen is en aanvaard mag worden, maar alleen dat onze waarnemingen H_0 niet tegenspreken.

Het verwerpen van H_0 is een uitspraak, waaraan veel meer betekenis gehecht kan worden dan aan het niet-verwerpen van H_0. Men tracht daarom H_0 als het tegendeel te formuleren van wat op theoretische gronden wordt verwacht. De gang van zaken doet dus enigszins denken aan een "bewijs uit het ongerijmde" in de wiskunde. In bovenstaand voorbeeld zal men bijv. als H_0 kiezen de hypothese, dat het bloedsuikergehalte bij patienten niet verhoogd is. Als nu in de steekproef deze getallen veel hoger liggen dan

de genoemde μ_0, dan wordt H_0 verworpen en mag men het vermoeden statistisch bevestigd achten.

Hoe meet men nu de verenigbaarheid van H_0 en de waarnemingen? Na de formulering van H_0 kiest men een onbetrouwbaarheid , d.i. het toelaatbaar geachte risiko dat H_0 ten onrechte wordt verworpen (terwijl H_0 waar is). Een gangbare keuze is $\alpha = 0{,}05$, terwijl ook $\alpha = 0{,}01$ geregeld voorkomt. Vervolgens zoekt men een toetsingsgrootheid $\underline{t}$; dit is een funktie van de waarnemingen die zich onder H_0 anders zal gedragen dan bij onjuistheid van H_0. Zolang de steekproef niet getrokken en waargenomen is, zijn de waarnemingen stochastisch en zal dus ook $\underline{t}$ een stochastische grootheid zijn. In het waardebereik van $\underline{t}$ zonderen we nu een kritieke zone Z af van waarden, die onder H_0 zeer onwaarschijnlijk zijn, maar bij onjuistheid van H_0 zeer wel kunnen optreden. Blijkt nu bij invulling van de getalwaarde van de waarnemingen de waarde van t in de kritieke zone te liggen, dan verwerpen wij H_0. De keuze van Z wordt zo gemaakt, dat

(i) $\quad P(\underline{t} \text{ valt in Z als } H_0 \text{ juist is}) \leq \alpha$

is, zodat er een kans van hoogstens α is dat we H_0 verwerpen hoewel H_0 juist is.

Keren we thans terug tot ons voorbeeld. Laten we eens aannemen, dat de bloedsuikergehalten van patienten normaal verdeeld zijn met onbekende verwachting μ en bekende variantie σ^2. De nulhypothese formuleren we als

$$H_0: \mu \leq \mu_0.$$

Noemen we de bloedsuikergehalten bij een steekproef van n patienten $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$, en onderstellen we dat deze waarnemingen onafhankelijk zijn (hetgeen bij aselekte keuze van patienten en onafhankelijke meting wel het geval zal zijn), dan is het steekproefgemiddelde $\overline{\underline{x}}$ de beste schatter van μ en $N(\mu, \frac{\sigma^2}{n})$ verdeeld. Het ligt daarom voor de hand de toets te baseren op de grootheid $\overline{\underline{x}}$. Als toetsingsgrootheid kiezen we

$$\underline{t}^* = \frac{\overline{\underline{x}} - \mu_0}{\sigma} \sqrt{n}.$$

Als $\mu = \mu_0$, dan is deze toetsingsgrootheid N(0,1) verdeeld. Indien $\mu = \mu_1 < \mu_0$, dan verschuift de kansverdeling van $\underline{t}^*$ over een afstand $\frac{\mu_0-\mu_1}{\sigma}\sqrt{n}$ naar links; is daarentegen $\mu = \mu_2 > \mu_0$, dan verschuift de kansverdeling over een afstand $\frac{\mu_2-\mu_0}{\sigma}\sqrt{n}$ naar rechts. In fig. 2.15 is de kansdichtheid $\underline{t}^*$ voor een drietal waarden van μ geschetst.

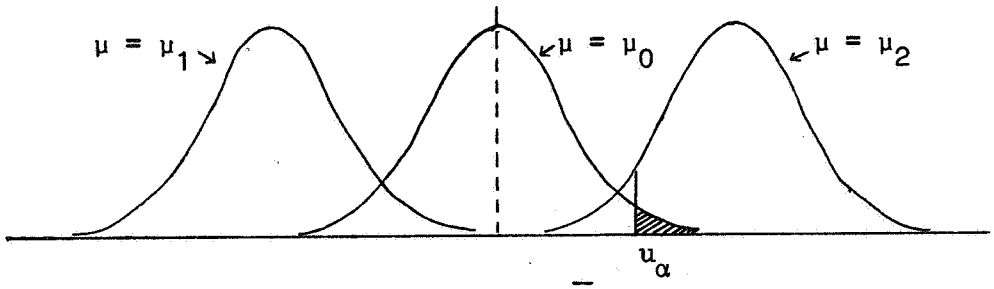


fig. 2.15. Kansdichtheid van $\frac{\bar{x} - \mu_0}{\sigma}\sqrt{n}$ voor $\mu = \mu_0$, $\mu = \mu_1 < \mu_0$ en $\mu = \mu_2 > \mu_0$. Het gearceerde oppervlak is gelijk aan α.

We kiezen thans een onbetrouwbaarheid α, bijv. $\alpha = 0{,}05$. Daar relatief grote waarden van $\bar{x}$ (en dus van $\underline{t}^*$) slecht in overeenstemming zijn met de hypothese H_0, zullen we grote waarden van de toetsingsgrootheid in de kritieke zone willen opnemen. We beweren nu, dat de toetsingsprocedure

$$\text{verwerp } H_0 \text{ als } \frac{\bar{x} - \mu_0}{\sigma}\sqrt{n} > u_\alpha$$

juist aan de eis (i) voldoet. De kritieke zone heeft thans de vorm $Z = \{\text{alle waarden van } \underline{t}^* \text{ groter dan } u_\alpha\}$. Tot de nulhypothese behoren alle waarden van $\mu \leq \mu_0$. Indien $\mu = \mu_0$, dan is $P(\underline{t}^* \text{ in } Z) = P(\underline{t}^* > u_\alpha) = \alpha$, en als $\mu = \mu_1 < \mu_0$, dan is de kansverdeling $\underline{t}^*$ naar links verschoven, zodat in dit geval $P(\underline{t}^* \text{ in } Z) = P(\underline{t}^* > u_\alpha) < \alpha$. Aan de eis (i) is dus inderdaad voldaan.

Bij de beschreven toets loopt men het risiko H_0 ten onrechte te verwerpen. Deze foute beslissing wordt een fout van de <u>eerste soort</u> genoemd. De kans op deze foute beslissing is hoogstens gelijk aan α en door een kleine keuze van de onbetrouwbaarheid kan men dit risiko dus beperken.

Naarmate men α kleiner kiest, wordt u_α groter en de kritieke zone dus kleiner.

Men loopt echter ook het risiko H_0 niét te verwerpen, hoewel de alternatieve hypothese H_1: $\mu > \mu_0$ waar is. Een dergelijke foute beslissing noemt men een fout van de tweede soort. Als de ware (onbekende) waarde van μ gelijk is aan $\mu_2 > \mu_0$, dan is de kans op deze fout juist gelijk aan het oppervlak links van u_α onder de rechtse kromme in fig. 2.15. De kans op deze foute beslissing kan men verkleinen door u_α te verkleinen (dus de kritieke zone te vergroten); hierbij tast men echter de onbetrouwbaarheid α van de toets aan. Voert men de steekproefomvang n op, dan blijft u_α ongewijzigd, maar de rechtse kromme verschuift verder naar rechts bij dezelfde waarde $\mu = \mu_2$, immers de rechtse kromme was t.o.v. de middelste kromme verschoven over een afstand $\frac{\mu_2 - \mu_0}{\sigma}\sqrt{n}$. Dit leidt tot een kleiner oppervlak onder de rechtse kromme links van u_α, en dus tot een kleinere kans op een fout van de tweede soort.

Het is in de toetsingstheorie meer gebruikelijk over het onderscheidingsvermogen (Engels: power) te spreken dan over een kans op een fout van de tweede soort. Het onderscheidingsvermogen is per definitie gelijk aan één min de kans op een fout van de tweede soort. Dit onderscheidingsvermogen is géén vast getal, maar hangt af van het speciale alternatief dat men beschouwt. Bij $\mu = \mu_0$ is het gelijk aan het oppervlak onder de rechtse kromme rechts van het punt u_α in fig. 2.15. Hoe groter het onderscheidingsvermogen is, des te beter is de toets (bij vaste α). We hebben reeds gezien dat het onderscheidingsvermogen stijgt bij toenemende steekproefomvang. Als funktie van μ neemt het onderscheidingsvermogen ook toe naarmate μ toeneemt; ook in dit geval verschuift de rechtse kromme immers verder naar rechts.

Bovenstaande uiteenzetting van een toets is gebaseerd op het feit, dat zeer grote waarden van de toetsingsgrootheid $\underline{t}^*$ zeer onwaarschijnlijk zijn als H_0 waar is. Het is duidelijk, dat naarmate $\underline{t}^*$ grotere waarden aanneemt, de hypothese H_0 onaannemelijker wordt. Stel eens dat $\underline{t}^*$ in een experiment de waarde t^* aanneemt. Dan kunnen we de kans bepalen dat $\underline{t}^*$ een waarde aanneemt nog groter dan t^* terwijl $\mu = \mu_0$. Deze is immers gelijk aan

$$P(\underline{t}^* > t^* \text{ als } \mu = \mu_0) = P(\underline{u} > t^*);$$

deze laatste kans kunnen we opzoeken in tabel 1 van de N(0,1) verdeling. Men noemt deze kans de overschrijdingskans (bij t^*); ze wordt vaak aangeduid met het symbool p. Deze overschrijdingskans is op te vatten als een maat voor de verenigbaarheid van de hypothese H_0 met de waarnemingen. Wordt de toets uitgevoerd met een onbetrouwbaarheid α, dan zal H_0 worden verworpen als $p < \alpha$ is, zoals men eenvoudig inziet. Bij rapportering geeft men vaak de overschrijdingskans p op; de lezer kan dan nagaan bij welke onbetrouwbaarheid α de nulhypothese H_0 nog wordt verworpen.

Voorbeeld 2.20. Uit een biologische populatie worden 6 individuen aselekt getrokken. Aan elk van deze individuen wordt een grootheid gemeten, waarvan bekend is dat deze normaal verdeeld is met een standaardafwijking $\sigma = 2$. Gevraagd wordt de hypothese te toetsen dat de verwachte waarde μ van de grootheid kleiner of gelijk is aan nul. De uitkomsten van de metingen zijn:

$$1{,}03;\ 1{,}78;\ 0{,}31;\ -0{,}86;\ 1{,}69;\ 0{,}99.$$

Hieruit volgt dat $t^* = 0{,}82 \cdot \sqrt{6}/2 = 1{,}00$. Wordt de toets uitgevoerd met een onbetrouwbaarheid van 5%, dan wordt de hypothese niet verworpen, immers $u_{0,05} = 1{,}645$ en $1{,}00 < 1{,}645$. De overschrijdingskans bedraagt

$$P(\underline{t}^* > 1{,}00) = P(\underline{u} > 1{,}00) = 0{,}1587,$$

zodat de hypothese ook bij $\alpha = 0{,}1$ nog niet verworpen zou worden.

De onderstelling, dat σ^2 bekend is, is weinig realistisch, daar dit slechts zelden het geval zal zijn. Laten we deze onderstelling vallen, dan kunnen we de beschreven toets echter niet meer toepassen, omdat σ in de toetsingsgrootheid optreedt. Het ligt nu voor de hand σ te schatten met $\underline{s}$, de wortel uit de steekpreofvariantie $\underline{s}^2$, en

$$\underline{t} = \frac{\overline{\underline{x}} - \mu_0}{\underline{s}} \sqrt{n}$$

als toetsingsgrootheid te gebruiken. Deze toetsingsgrootheid is nu echter niet meer normaal verdeeld, maar heeft, als $\mu = \mu_0$, een t-verdeling van STUDENT met n-1 vrijheidsgraden. We verwerpen de nulhypothese

$$H_0: \mu \leq \mu_0$$

thans ten gunste van de alternatieve hypothese

$$H_1: \mu > \mu_0$$

als

$$\frac{\underline{\bar{x}} - \mu_0}{\underline{s}} \sqrt{n} > t_{n-1;\alpha} \text{ ,}$$

waarin $t_{n-1;\alpha}$ het rechtse α-punt van de $\underline{t}_{n-1}$-verdeling voorstelt. Evenals bij de toetsingsgrootheid $\underline{t}^*$ kan men weer laten zien, dat voor elke $\mu \leq \mu_0$ de hypothese H_0 wordt verworpen met een kans van ten hoogste α. Het onderscheidingsvermogen tegen een alternatief μ_2 neemt weer toe als μ_2 stijgt en als, bij vaste μ_2, de steekproefomvang n stijgt.

De hypothese H_0 was een éénzijdige hypothese, die leidde tot een rechtséénzijdige kritieke zone Z. Men noemt een dergelijke toets een rechtséénzijdige toets. Wensen we de nulhypothese

$$H_0: \mu \geq \mu_0$$

te toetsen met als alternatieve hypothese

$$H_1: \mu < \mu_0,$$

dan verwerpen we hypothese H_0 ten gunste van H_1 als

$$\frac{\underline{\bar{x}} - \mu_0}{\underline{s}} \sqrt{n} < -t_{n-1;\alpha} \text{ ;}$$

dit is een linkséénzijdige toets met onbetrouwbaarheid α, die overigens een soortgelijke gedaante heeft als de rechtséénzijdige toets.

Het komt ook voor, dat men een nulhypothese van de gedaante

$$H_0: \mu = \mu_0$$

wil toetsen met als alternatieve hypothese

$$H_1: \mu < \mu_0 \quad \text{of} \quad \mu > \mu_0.$$

In dit geval zijn zowel relatief grote als relatief kleine waarden van de toetsingsgrootheid $\underline{t}$ slecht in overeenstemming met de nulhypothese. Men stelt dan ook een kritieke zone Z op, die uit twee stukken bestaat: Z_1 (kleine waarden van $\underline{t}$) en Z_2 (grote waarden van $\underline{t}$). We moeten er thans voor zorgen, dat bij de gekozen (tweezijdige) onbetrouwbaarheid onder H_0

$$P(\underline{t} \text{ valt in } Z) = P(\underline{t} \text{ valt in } Z_1) + P(\underline{t} \text{ valt in } Z_2) \leq \alpha$$

is. Dit kan men bereiken door de kans dat $\underline{t}$ in Z_1 resp. Z_2 valt beide gelijk aan $\frac{\alpha}{2}$ te kiezen. Dit leidt tot de volgende <u>tweezijdige toets</u>: verwerp H_0

ten gunste van $\mu < \mu_0$ als $\dfrac{\underline{\bar{x}} - \mu_0}{\underline{s}} \sqrt{n} < t_{n-1;\frac{\alpha}{2}}$ is

en

ten gunste van $\mu > \mu_0$ als $\dfrac{\underline{\bar{x}} - \mu_0}{\underline{s}} \sqrt{n} > t_{n-1;\frac{\alpha}{2}}$ is.

De keuze van een éénzijdige of tweezijdige toets hangt steeds af van het gestelde probleem en ligt meestal wel voor de hand. De keuze van de onbetrouwbaarheid α is lastiger; meestal kiest men $\alpha = 0{,}05$; maar als een fout van de eerste soort zeer ernstig is, komt ook $\alpha = 0{,}01$ of $\alpha = 0{,}001$ in aanmerking. Er zij nog opgemerkt, dat men zich bij een grote steekproefomvang een kleine waarde van α meestal wel kan permiteren, omdat het omderscheidingsvermogen dan toch relatief groot is. In géén geval mag men zich bij deze beslissing laten leiden door de uitkomsten van de waarnemingen zelf!

De behandelde toetsen, die berusten op de toetsingsgrootheid $\underline{t}$, heten <u>één-steekproeftoetsen van STUDENT</u>.

<u>Voorbeeld 2.21</u>. Een kweker brengt bloemzaadjes op de markt in een bepaalde verpakking. De inhoud van de zakjes varieert enigszins, maar de kweker

verzekert, dat de zakjes gemiddeld minstens 10 gram zaadjes bevatten. Een grote afnemer besluit te onderzoeken of de bewering van de kweker juist is. Mocht dit niet het geval zijn, dan is hij van plan te reclameren. Hij neemt een aselekte steekproef van 10 zakjes uit een grote partij en weegt de inhoud. We onderstellen, dat de gewichten onafhankelijk zijn en normaal verdeeld met onbekende verwachting μ en variantie σ^2. Daar de afnemer bij reclame zo zeker mogelijk van zijn zaak wil zijn, wordt als nulhypothese H_0: $\mu \geq 10$ en als alternatieve hypothese H_1: $\mu < 10$ gekozen, terwijl de toetsing uitgevoerd zal worden met een onbetrouwbaarheid $\alpha = 0{,}01$. De gevonden gewichten zijn (in grammen):

$$9{,}6;\ 9{,}9;\ 10{,}2;\ 10{,}0;\ 9{,}5;\ 9{,}9;\ 10{,}1;\ 10{,}5;\ 9{,}8;\ 9{,}4.$$

Uit de waarnemingen berekenen we $\bar{x} = 9{,}89$ en $s = \sqrt{s^2} = \sqrt{0{,}1121} = 0{,}335$, zodat $t = -0{,}11.\sqrt{10}/0{,}335 = -1{,}04$. Daar volgens tabel 2 $t_{9;0,01} = 2{,}821$ en dus $t > -t_{9;0,01}$ is, wordt H_0 niet verworpen en is er op grond van de steekproef geen aanleiding tot reclame.

De één-steekproeftoetsen van STUDENT worden meestal in een iets andere situatie toegepast. Men is nl. vaak geïnteresseerd in de invloed van een bepaalde behandeling op individuen uit een grote populatie en verricht daarom metingen aan een steekproef van individuen voor en na de behandeling met het doel deze uitkomsten te vergelijken. Wenst men bijv. de invloed van een bepaalde injektie na te gaan op de bloeddruk dan kan men voor en na de injektie deze bloeddruk bepalen aan een steekproef van individuen uit de beschouwde populatie. Men verkrijgt aldus een n-tal gepaarde waarnemingen

$$(\underline{y}_1,\underline{z}_1),\ (\underline{y}_2,\underline{z}_2),\ \ldots,\ (\underline{y}_n,\underline{z}_n).$$

De $\underline{y}_i$ en $\underline{z}_i$ stellen dan de waarnemingen vóór en na injektie voor bij het i-de individu uit de steekproef. Nemen we aan, dat deze waarnemingen alle onafhankelijk en normaal verdeeld zijn met verwachtingen

$$E\underline{y}_i = \mu_i,\ E\underline{s}_i = \nu_i \quad (i=1,2,\ldots,n),$$

dan zijn de verschillen

$$\underline{x}_1 = \underline{y}_1 - \underline{z}_1,\ \underline{x}_2 = \underline{y}_2 - \underline{z}_2,\ \ldots,\ \underline{x}_n = \underline{y}_n - \underline{z}_n$$

ook onafhankelijk en normaal verdeeld met verwachtingen

$$E\underline{x}_i = \mu_i - \nu_i \quad (i=1,2,\ldots,n).$$

Indien de $\underline{x}_i$ alle dezelfde variantie hebben (hetgeen bijv. het geval is als alle $\underline{y}_i$ dezelfde variantie hebben en alle $\underline{z}_i$ ook), dan kan men de hypothese

$$H_0: \mu_i - \nu_i = 0 \quad (\text{resp. } \leq 0 \text{ resp. } \geq 0) \text{ voor alle } i$$

toetsen met de boven beschreven toetsen van STUDENT. In dit geval is $\mu_0 = 0$ en de toetsingsgrootheid is dus

$$\frac{\bar{\underline{x}}}{\underline{s}}\sqrt{n}.$$

Het is van grote betekenis, dat de μ_i, evenals de ν_i, onderling best mogen verschillen; niveauverschillen tussen de paren verstoren de toets niet (we letten immers allen op de verschillen $\mu_i - \nu_i$).

In vele praktische situaties zullen de waarnemingen niet normaal verdeeld zijn, zodat de toetsen van STUDENT niet mogen worden toegepast. Men kan dan echter gebruik maken van verdelingsvrije toetsen, waarbij slechts zeer weinig onderstellingen omtrent de verdeling van de waarnemingen worden gemaakt. In plaats van de besproken toets van STUDENT kan men bijv. de symmetrietoets van WILCOXON toepassen; hierbij onderstelt men alleen dat de waarnemingen een symmetrische verdeling bezitten. In §10 zullen we voorts nog de tekentoets ontmoeten. Door hun algemene toepasbaarheid spelen verdelingsvrije toetsen een grote rol bij het statistisch onderzoek van experimenten.

§9. De twee-steekproeven toets van WILCOXON

De twee-steekproeven toets van WILCOXON is de meest gebruikelijke verdelingsvrije toets voor het probleem van twee steekproeven. We zullen

deze toets voortaan kortweg aanduiden als de toets van WILCOXON.

Het probleem van twee steekproeven kan als volgt worden omschreven. Men beschouwt een bepaald kenmerk (bijv. een lengte) dat aan elk element van twee populaties kan worden gemeten. Uit elk van beide populaties trekt men een aselekte steekproef; zij m de omvang van de steekproef uit de eerste populatie en n de omvang van de steekproef uit de tweede populatie. Bij bepaling van het kenmerk aan deze elementen zal men twee rijtjes getallen vinden:

$$x_1, x_2, \ldots, x_m \qquad \text{(eerste steekproef)}.$$
$$y_1, y_2, \ldots, y_n \qquad \text{(tweede steekproef)}.$$

Zijn de waarnemingen nog niet verricht, dan zijn dit stochastische grootheden. De waarnemingen $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_m$ hebben alle dezelfde kansverdeling, daar ze uit dezelfde populatie afkomstig zijn, en hetzelfde geldt voor $\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_n$. We zullen aannemen dat deze waarnemingen alle onafhankelijk zijn (dit is bij trekking zonder teruglegging niet het geval als de steekproefomvang groot is t.o.v. populatie-omvang). Men wenst de hypothese H_0 te toetsen dat het kenmerk in beide populaties dezelfde kansverdeling heeft (m.a.w. er zijn tussen beide populaties géén systematische verschillen m.b.t. dit kenmerk).

We zullen niets onderstellen omtrent de kansverdeling van de waarnemingen (dus geen normaliteit). We toetsen H_0 dan met behulp van de toetsingsgrootheid van WILCOXON, $\underline{W}$ genaamd, die gelijk is aan tweemaal het aantal paren $(\underline{x}_i, \underline{y}_j)$ waarvoor $\underline{x}_i > \underline{y}_j$ is, vermeerderd met het aantal paren $(\underline{x}_r, \underline{y}_s)$ waarvoor $\underline{x}_r = \underline{y}_s$.

Het is duidelijk dat $\underline{W}$ een relatief kleine waarde zal aannemen indien in de steekproeven de x-waarden overwegend kleiner zijn dan de y-waarden, een relatief grote als het omgekeerde het geval is, een een intermediare waarde als geen van deze beide situaties zich voordoet.

We zullen in deze paragraaf gemakshalve alleen situaties beschouwen waar alle waarnemingen verschillend zijn.

Voorbeeld 2.22. Men meet de lengte van 8 aselekt gekozen zaaddozen (in mm) in elk van twee populaties planten. De uitkomsten zijn:

1e steekproef (x): 2,9; 3,4; 4,1; 4,5; 4,8; 5,3; 5,8; 6,5

2e steekproef (y): 3,1; 3,3; 4,0; 4,7; 5,1; 5,2; 6,0; 6,8.

Men gaat nu eenvoudig na dat in dit geval $\underline{W}$ de waarde $W = 62$ aanneemt.

Bij het uitvoeren van de toetsingsprocedure dient men de verdeling van $\underline{W}$ te kennen onder de nulhypothese, d.w.z. als H_0 waar is.

Als m en n beide klein zijn, is de kansverdeling van $\underline{W}$ eenvoudig uit te rekenen. Men maakt daarbij gebruik van het feit dat onder H_0 alle waarnemingen dezelfde kansverdeling hebben en dus elke ordening van x-waarden en y-waarden even waarschijnlijk is. Zo hebben we in het geval $m = 2$, $n = 3$ de volgende 10 even waarschijnlijke mogelijkheden (in volgorde van opklimmende grootte):

xxyyy → W = 0		
xyxyy → W = 2		$P(\underline{W}=0) = 0,1$
xyyxy → W = 4		$P(\underline{W}=2) = 0,1$
xyyyx → W = 6		$P(\underline{W}=4) = 0,2$
yxxyy → W = 4	$\Longrightarrow$	$P(\underline{W}=6) = 0,2$
yxyxy → W = 6		$P(\underline{W}=8) = 0,2$
yxyyx → W = 8		$P(\underline{W}=10) = 0,1$
yyxxy → W = 8		$P(\underline{W}=12) = 0,1$
yyxyx → W = 10		
yyyxx → W = 12		

Voor kleine m en n is de kansverdeling van $\underline{W}$ (onder H_0) uitvoerig getabelleerd.

Zijn m en n beide vrij groot, dan is onder H_0 $\underline{W}$ bij benadering normaal verdeeld met verwachting en variantie

$$E\underline{W} = \mu = mn$$

$$\sigma^2(\underline{W}) = \sigma^2 = \frac{1}{3}\, mn\, (m+n+1),$$

mits de waarnemingen geen gelijke waarden aannemen. Voor niet te kleine m en n is dus onder H_0 de grootheid

$$\underline{W}^* = \frac{\underline{W}-\mu}{\sigma}$$

bij benadering standaard-normaal verdeeld.

We beschouwen thans eerst tweezijdige toetsing. Daar zowel relatief grote als relatief kleine waarden van $\underline{W}$ slecht in overeenstemming zijn met H_0, kiezen we als kritieke zone Z juist deze grote en kleine waarden en verwerpen H_0 als $\underline{W}$ een waarde in deze kritieke zone aanneemt. Zij α de onbetrouwbaarheid van deze toets. Dan mag de kans dat $\underline{W}$ een waarde in Z aanneemt als H_0 waar is dus niet groter zijn dan α:

$$P(\underline{W} \text{ valt in Z terwijl } H_0 \text{ juist}) \leq \alpha.$$

De kritieke zone Z bestaat uit 2 stukken:

$$\text{waarden } W \leq W_l \quad \text{en} \quad \text{waarden } W \geq W_r.$$

De getallen W_l en W_r, de grenzen van de kritieke zone, hangen af van m, n en α. Aan de gestelde eis is zeker voldaan als

$$P(\underline{W} \leq W_l \text{ terwijl } H_0 \text{ juist}) \leq \frac{1}{2}\alpha$$

én

$$P(\underline{W} \geq W_r \text{ terwijl } H_0 \text{ juist}) \leq \frac{1}{2}\alpha.$$

Voor kleine waarden van m en n en $\alpha = 5\%$ (dus $\frac{1}{2}\alpha = 0{,}025$) kan men W_l opzoeken in de volgende tabel; W_r vindt men dan uit de relatie $W_r = 2mn - W_l$.

Tabel van W_1 bij $\frac{1}{2}\alpha = 0{,}025$; $m,n \leq 13$.

n \ m	1	2	3	4	5	6	7	8	9	10	11	12	13
3	-	-	-										
4	-	-	-	0									
5	-	-	0	2	4								
6	-	-	0	4	6	10							
7	-	-	0	6	10	12	16						
8	-	0	4	8	12	16	20	26					
9	-	0	4	8	14	20	24	30	34				
10	-	0	6	10	16	22	28	34	40	46			
11	-	0	6	12	18	26	32	38	46	52	60		
12	-	2	8	14	22	28	36	44	52	58	66	74	
13	-	2	8	16	24	32	40	48	56	66	74	82	90

(een streepje geeft aan dat men H_0 nooit kan verwerpen, omdat voor deze m en n ook $P(\underline{W}=0) > 0{,}025$ is).
De tabel geeft alleen waarden van W_1 voor $m \leq n$; voor $m > n$ verwissele men eenvoudig m en n, de tabel blijft dan geldig.

Zij m en n beide niet te klein, dan kan men W_1 en W_r bij $\alpha = 0{,}05$ bepalen m.b.v. de normale benadering

$$\frac{W_1-\mu}{\sigma} = -1{,}96 \quad \text{en} \quad \frac{W_r-\mu}{\sigma} = 1{,}96,$$

waarin μ en σ uit de voorgaande formules berekend dienen te worden.

Vindt men na uitvoering van het experiment een waarde W die niet in de kritieke zine ligt, dus $W_1 < W < W_r$, dan wordt H_0 niet verworpen (dit betekent echter niet dat men tot "H_0 is waar" mag besluiten). Vindt men een waarde $W \leq W_1$, dan wordt H_0 verworpen ten gunste van de hypothese dat het beschouwde kenmerk in de eerste populatie systematisch kleinere waarden aanneemt dan in de tweede populatie. Is daarentegen $W \geq W_r$, dan verwerpt men H_0 ten gunste van de hypothese dat het kenmerk in de eerste populatie systematisch grotere waarden aanneemt dan in de tweede populatie.

Voorbeeld 2.22 (vervolg). Eerder vonden we $W = 62$ bij $m = n = 8$. Bij $\alpha = 0{,}05$ is volgens de tabel $W_1 = 26$ en dus $W_r = 2mn - W_1 = 102$. De hypothese H_0 wordt dus niet verworpen. Daar m en n vrij klein zijn, is toepassing van de normale benadering hier niet aan te bevelen. Doen we dit toch, dan vinden we $\mu = 64$ en $\mu = \sqrt{\frac{1}{3}} \times 8 \times 8 \times 17 = 19{,}04$, waaruit volgt dat $W_1 = 64 - 1{,}96 \times 19{,}04 \approx 26$ en dus $W_r \approx 102$. De normale benadering blijkt hier dus nog het juiste resultaat te geven.

Soms wenst men niet tweezijdig maar éénzijdig te toetsen. Dit kan het geval zijn als men er uitsluitend in geïnteresseerd is H_0 te verwerpen ten gunste van de alternatieve hypothese dat het kenmerk in de eerste populatie systematisch kleiner (bij linkséénzijdige toetsing) dan wel systematisch groter (bij rechtséénzijdige toetsing) is dan in de tweede populatie.

We beschouwen hier kort het geval van linkséénzijdige toetsing. In dit geval bestaat de kritieke zone alleen uit kleine waarden van W:

$$\text{waarden } W \leq W_1,$$

en moet bij de gekozen onbetrouwbaarheid α deze kritieke zone voldoen aan

$$P(\underline{W} \leq W_1 \text{ terwijl } H_0 \text{ juist}) \leq \alpha$$

(let op het verschil met tweezijdige toetsing, waar in het rechterlid $\frac{1}{2}\alpha$ stond i.p.a. α). Om bij kleine steekproevenomvangen en $\alpha = 5\%$ de kritieke waarde W_1 te bepalen kan men de gegeven tabel niet gebruiken. Er bestaan echter ook soortgelijke tabellen voor andere waarden van α, die men wel kan toepassen. Bij grote m en n kan men weer de normale benadering toepassen; voor $\alpha = 5\%$ volgt W_1 dit maal uit

$$\frac{W_1 - \mu}{\sigma} = -1{,}645.$$

Voor rechtséénzijdige toetsing verloopt de redenering analoog; bij toepassing van de normale benadering vindt men in dat geval

$$\frac{W_r - \mu}{\sigma} = 1{,}645.$$

§10. Lineaire regressie

In vele praktische situaties is men geïnteresseerd in de invloed van een grootheid op een andere grootheid. We beschouwen hier het geval dat $\underline{y}$ een stochastische grootheid is, waarvan de kansverdeling afhangt van een andere, niet-stochastische grootheid x. Zo kan $\underline{y}$ de lengte van de remweg voorstellen bij een bepaalde beginsnelheid x, of $\underline{y}$ is de levensduur van een gloeilamp bij een aangebrachte spanning x. Een ander bekend voorbeeld is de responsie $\underline{y}$ bij een dosis x van een stof, indien de verschillende doses aan aselekt gekozen groepen proefdieren worden toegediend. In deze gevallen kan men de waarde van x zelf kiezen, terwijl $\underline{y}$ bij deze waarde van x vervolgens wordt waargenomen. Bij vele ijkingsproblemen is dit model eveneens van betekenis; x stelt dan de meting van een grootheid voor volgens een zeer nauwkeurige (maar tijdrovende of kostbare) methode, terwijl $\underline{y}$ een meting van dezelfde grootheid is met een grove methode, waarbij een meetfout optreedt. In al deze gevallen noemt men x de onafhankelijke variabele en $\underline{y}$ de (van x) afhankelijke variabele.

De vraag rijst nu, hoe de kansverdeling van $\underline{y}$ afhangt van x. We zullen alleen het eenvoudigste geval bespreken, nl. dat de verwachting van $\underline{y}$ lineair afhangt van x:

$$E(\underline{y}|x) = \alpha + \beta x, \quad ^{*)}$$

waarin α en β onbekende grootheden zijn, en de variantie σ^2 van $\underline{y}$ niet van x afhangt (dus steeds even groot is). We spreken van lineaire regressie van $\underline{y}$ op x. We hebben hier dus te maken met drie onbekenden, de parameters α, β (die men de regressiecoëfficiënten noemt) en σ^2. Op grond van verkregen waarnemingen kan men deze parameters schatten of hypothesen omtrent deze parameters toetsen.

De waarnemingen bestaan thans uit paren:

$$(x_1,\underline{y}_1),\ (x_2,\underline{y}_2),\ \ldots,\ (x_n,\underline{y}_n);$$

---

*) Met $E(\underline{y}|x)$ bedoelen we de verwachting van $\underline{y}$ bij gegeven waarde x.

bij iedere beschouwde waarde x_i van x behoort een waarneming y_i. We zullen steeds onderstellen, dat de y_i onafhankelijk zijn. De waarnemingsuitkomsten kan men grafisch uitzetten in een spreidingsdiagram, zoals bijv. in fig. 2.16 is geschied. In het geval van lineaire regressie zal de zo verkregen puntenwolk een duidelijk lineair verloop vertonen.

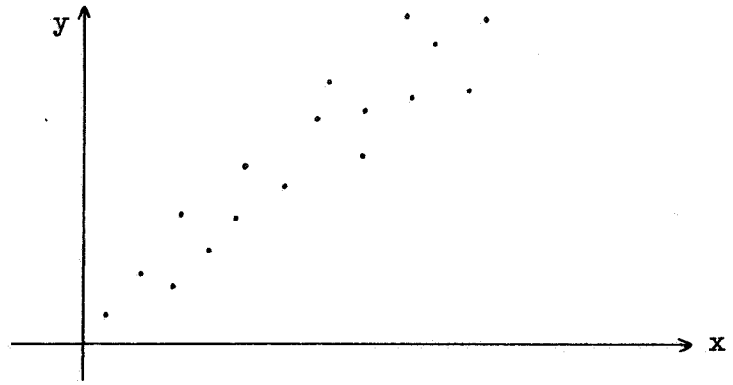


fig. 2.16. Spreidingsdiagram.

Teneinde de onbekende parameters te schatten, zoeken we de rechte lijn (de regressielijn van y op x), die het beste bij de puntenparen in de grafiek past. De aangewezen methode om dit te bereiken is de methode der kleinste kwadraten, die we thans zullen bespreken. Zij

$$y = a + bx$$

de vergelijking van een rechte lijn in het strooiïngsdiagram. Men noemt de verticale afstand van het i-de waargenomen punt tot deze rechte lijn,

$$r_i = y_i - a - bx_i,$$

een residu. We zullen na a en b zo trachten te bepalen, dat de kwadraatsom van de residuen

$$Q = \sum_{i=1}^{n} r_i^2$$

minimaal is (zie fig. 2.17).

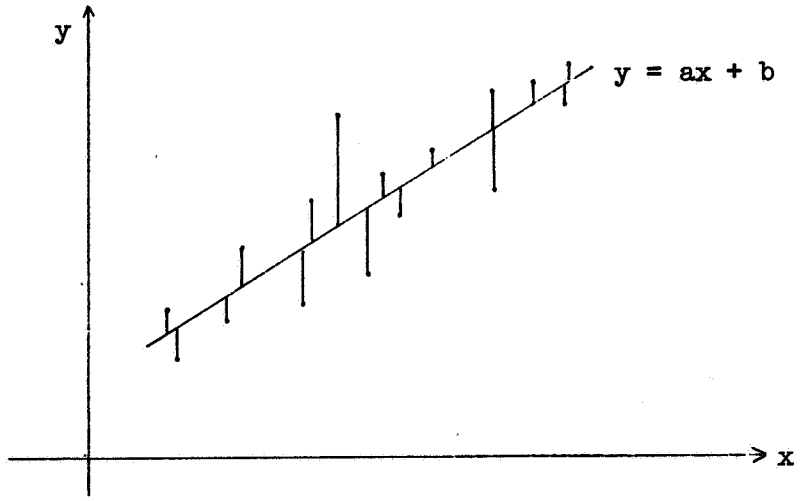


fig. 2.17. Spreidingsdiagram met aan puntenwolk aangepaste rechte lijn en residuen.

Zij $\bar{x}$ het gemiddelde der x_i en $a' = a + b\bar{x}$ (zodat $a = a' - b\bar{x}$). Dan kunnen we Q ook als volgt schrijven:

$$Q = \sum_{i=1}^{n} \{y_i - a' - b(x_i - \bar{x})\}^2 =$$

$$= \sum_{i=1}^{n} (y_1 - a')^2 + b^2 \sum_{i=1}^{n} (x_i - \bar{x})^2 - 2b \sum_{i=1}^{n} y_i(x_i - \bar{x}),$$

waarbij we gebruik hebben gemaakt van de relatie $\sum_{i=1}^{n} (x_i - \bar{x}) = 0$, De vorm Q bestaat dus uit twee stukken, waarvan het ene alleen a' en het andere alleen b bevat, zodat deze afzonderlijk geminimaliseerd kunnen worden. De eerste term is minimaal, als men voor a' het gemiddelde der y_i kiest, d.w.z.

$$a = \bar{y} - b\bar{x}.$$

De laatste twee termen van Q zijn minimaal, als men voor b kiest

$$b = \sum_{i=1}^{n} y_i(x_i - \bar{x}) \Big/ \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Men kan deze resultaten eenvoudig verifiëren door naar a' resp. b te differentiëren en de afgeleide nul te stellen.

Kunnen de aldus gevonden waarden a en b nu ook dienst doen als schattingen van de parameters α en β? We merken allereerst op, dat a en b stochastische grootheden zijn, immers beide hangen af van de stochastische waarnemingen $\underline{y}_i$ (bij bovenstaande afleiding zijn we van gegeven waarnemingsuitkomsten uitgegaan en hebben daarom niet onderstreept). Vroeger zijn zuiverheid en een kleine variantie reeds als gunstige eigenschappen van schatters ter sprake gekomen. We zullen thans laten zien, dat $\underline{a}$ en $\underline{b}$ zuivere schatters zijn van α resp. β. Immers

$$E\{\sum_{i=1}^{n} \underline{y}_i(x_i-\bar{x})\} = \sum_{i=1}^{n} (x_i-\bar{x})\, E\underline{y}_i = \sum_{i=1}^{n} (x_i-\bar{x})(\alpha+\beta x_i) =$$

$$= \beta \sum_{i=1}^{n} (x_i-\bar{x})x_i = \beta \sum_{i=1}^{n} (x_i-\bar{x})^2,$$

zodat

$$E\underline{b} = \beta,$$

en anderzijds

$$E\underline{a} = E(\bar{\underline{y}}-\underline{b}\bar{x}) = E(\frac{1}{n}\sum_{i=1}^{n} \underline{y}_i) - \beta\bar{x} = \frac{1}{n}\sum_{i=1}^{n} E\underline{y}_i - \beta\bar{x} =$$

$$= \frac{1}{n}\sum_{i=1}^{n} (\alpha+\beta x_i) - \beta\bar{x} = \alpha + \frac{\beta}{n}\sum_{i=1}^{n} x_i - \beta\bar{x} = \alpha.$$

Zonder bewijs vermelden we, dat $\underline{a}$ en $\underline{b}$ bovendien onder alle zuivere schatters die lineair zijn in de $\underline{y}_i$, de kleinste variantie hebben.

De kleinste kwadraten methode leidt dus tot schatters van α en β met zeer gunstige eigenschappen. We zoeken nu nog een schatter van de derde onbekende parameter, σ^2. Hiertoe substitueren we in de definitie van de kwadraatsom Q van residuen de kleinste kwadraten schatters $\underline{a}$ en $\underline{b}$. De aldus ontstane grootheid noemt men gewoonlijk de residuele kwadraatsom of "error sum of squares" en wordt vaak aangeduid met $\underline{SS}_e$:

$$\underline{SS}_e = \sum_{i=1}^{n} (\underline{y}_i-\underline{a}-\underline{b}x_i)^2.$$

Men kan laten zien, dat

$$E(\underline{SS}_e) = (n-2)\sigma^2,$$

zodat

$$\underline{SS}_e/(n-2)$$

een zuivere schatter is van σ^2. Hier is n-2 het aantal vrijheidsgraden van de kwadraatsom $\underline{SS}_e$. Intuïtief is dit resultaat wel te begrijpen. Bij bekende verwachtingen μ_i van de $\underline{y}_i$ zou men de variantie σ^2 immers kunnen schatten met het gemiddelde

$$\frac{1}{n} \sum_{i=1}^{n} (\underline{y}_i - \mu_i)^2.$$

Deze μ_i zijn in het regressiemodel gelijk aan $\alpha + \beta x_i$ en kunnen we schatten met $\underline{a} + \underline{b}x_i$. Doen we dit, dan ontstaat juist $\underline{SS}_e/n$. Daar we hier voor het schatten van de μ_i twee parameters α en β hebben geschat, lijkt het plausibel, dat het aantal vrijheidsgraden twee kleiner is dan het aantal waarnemingen. Dit is inderdaad het geval en we moeten $\underline{SS}_e$ dus door n-2 delen om een zuivere schatter te verkrijgen van σ^2 (vergelijk de opmerking op pag. 53).
De geschatte regressielijn

$$Y = a + bx$$

is een schatting van de ware regressielijn $E(\underline{y}|x) = \alpha + \beta x$. *)
Deze geschatte regressielijn stelt ons in staat bij gegeven waarde van x de bijbehorende waarde van $\underline{y}$ te "voorspellen". Eigenlijk voorspellen we aldus de verwachte waarde van $\underline{y}$. De waarden, die $\underline{y}$ van geval tot geval aanneemt, zullen om de verwachte waarde (dus om de ware regressielijn) heen schommelen met variantie σ^2. Naarmate σ^2 kleiner is (hetgeen we kunnen beoordelen aan de hand van $\underline{SS}_e/(n-2)$ of zonder formules ook aan de mate van concentratie van de puntenwolk om de geschatte lijn), zullen we

*) We merken nog op, dat het punt $(\overline{x},\overline{y})$ steeds op de geschatte regressielijn ligt.

uitkomsten $\underline{y}$ dus beter kunnen voorspellen. Dit voorspellen van $\underline{y}$ bij gegeven x is een van de belangrijkste toepassingen van regressie. Bij ijkingsproblemen doet zich echter de omgekeerde situatie voor; uit een waargenomen $\underline{y}$, de uitkomst van een grove meting met een meetfout, zou men graag de juiste waarde x willen afleiden. De geschatte regressielijn stelt ons in staat bij elke waarde van $\underline{y}$ de bijbehorende x te schatten.

Tot nu toe hebben we over de precieze verdeling van de $\underline{y}_i$ geen onderstellingen gemaakt. Bij het schatten van de parameters is dit ook overbodig. Zijn de $\underline{y}_i$ normaal verdeeld, dan kan men ook betrouwbaarheidsintervallen voor α en β, voorspellingsintervallen voor een toekomstige waarde van $\underline{y}$ bij gegeven x (waarbinnen de waarde van $\underline{y}$ met een bepaalde kans zal liggen) en betrouwbaarheidsintervallen voor x bij een uitkomst $\underline{y}$ construeren. Tevens kunnen dan hypothesen omtrent de waarden van α en β worden getoetst. We gaan hier verder niet op in.

In het begin van dit hoofdstuk hebben we ondersteld, dat de regressie van $\underline{y}$ op x lineair is. De schatters van α, β en σ^2 zijn ook hier op gebaseerd. In vele situaties is het a priori echter geenszins zeker, dat de regressie lineair is. Aan de hand van een spreidingsdiagram kan men met enige ervaring vaak wel beoordelen of dit een redelijke onderstelling is, maar men dient dan wel over een goed "timmermansoog" te beschikken. Als men bij elke x_i niet één, maar meerdere waarnemingen $\underline{y}_i$ verricht en deze waarnemingen normaal verdeeld zijn, dan kan men de hypothese H_0, dat de regressie lineair is, bovendien ook toetsen.

We besluiten deze korte beschouwing over lineaire regressie met een waarschuwing. Ook al blijkt een rechte lijn heel aardig te passen bij de waargenomen puntenparen, dan nog heeft men geen enkele garantie, dat deze regressielijn ook dienst kan doen voor waarden van x, die ver buiten het gebied liggen waarbinnen men bij de uitgevoerde experimenten x heeft gekozen. Men denke in dit verband aan regressielijnen van geboortecijfers of verkoopcijfers van auto's op de tijd, die soms gehanteerd worden om prognoses te maken voor tijdstippen in een ver verwijderde toekomst. De enorme fouten, die met deze extrapolatie gemaakt zijn, spreken wel voor zichzelf.

Voorbeeld 2.23. Bij de bepaling van stikstofgehalten in het serum (in gr per 100 cc serum) bij vijf aselekt gekozen groepen van elk vijf ratten op vijf verschillende leeftijden (in dagen) vond men de volgende resultaten:

Leeftijd x	Stikstofgehalte y				
25	0,70	0,76	0,78	0,80	0,84
50	0.92	0.96	0.98	1,00	1,03
80	0,98	1,03	1,05	1,06	1,11
130	1,09	1,13	1,14	1,14	1,22
180	1,13	1,19	1,20	1,21	1,25

In fig. 2.18 zijn deze cijfers in een spreidingsdiagram uitgezet, tezamen met de best passende lineaire regressielijn. Toepassing van de formules voor a en b geeft als kleinste-kwadraten schattingen van de regressiecoëfficiënten a = 0,802 en b = 0,00239.
Onderstellen we, dat de y's normaal verdeeld zijn, dan kunnen we toetsen of de regressie lineair is. Op het oog ziet het er naar uit dat dit niet het geval zal zijn en bij toetsing wordt de lineariteit inderdaad verworpen. De bovenstaande schattingen zijn in deze situatie dus van weinig waarde.

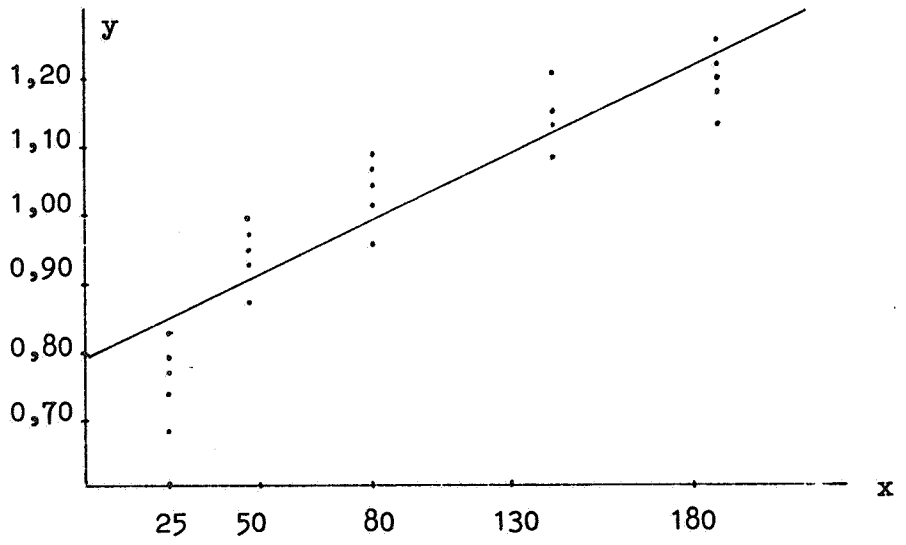


fig. 2.18. Spreidingsdiagram van de gegevens uit vb. 2.23.

§11. De binomiale verdeling; schatten en toetsen van een kans

Onderstel eens, dat in een zeer grote populatie van individuen een bepaalde fraktie p een zekere eigenschap A bezit. Over deze fraktie p, die onbekend is, willen we graag nadere informatie verkrijgen. Het is in de statistiek gebruikelijk het optreden van A een succes (S) te noemen en het niet-optreden van A een mislukking (M). Met behulp van deze terminologie kunnen we dus een groot aantal verschillende problemen beschrijven.

Stel we trekken een aselekte steekproef van n individuen uit onze populatie en tellen het aantal "successen". Zij $\underline{x}$ het aantal successen in de steekproef, dan kan $\underline{x}$ dus één van de waarden 0, 1, 2, ..., n aannemen. Bij elk individu hebben we een kans p op succes.

Hoe ziet de kansverdeling van de diskrete grootheid $\underline{x}$ er nu uit? We zullen voor een willekeurige x (één van de waarden 0, 1, ..., n) de kans $P(\underline{x}=x)$ bepalen. Er treden nu x successen en n-x mislukkingen in één of andere volgorde op. Beschouw eens één speciaal rijtje van successen en mislukkingen:

SSMSMMMS ... MSM.

Op elke plaats treedt een S op met kans p, een M met kans q = 1-p. Trekken we met teruglegging, of is bij trekking zonder teruglegging de populatie zeer groot t.o.v. de steekproefomvang, dan kunnen we de trekkingen als onafhankelijk beschouwen. De kans op een bepaald rijtje uitkomsten is dan gelijk aan het produkt van de kansen op elk der afzonderlijke uitkomsten (zie §3). Voor een vast rijtje met x maal S en n-x maal M komt in het produkt dus ook x maal het symbool p en n-x maal het symbool q voor, m.a.w.

$$P\text{ (vast rijtje uitkomsten met x successen)} = p^x q^{n-x}.$$

Nu bestaan er, zoals we vroeger gezien hebben, $\binom{n}{x}$ verschillende rijtjes ter lengte n met x successen, immers we kunnen x successen op $\binom{n}{x}$ verschillende manieren over de n beschikbare plaatsen verdelen. Daar verschillende rijtjes niet tegelijk als uitkomst van het experiment kunnen optreden (ze zijn dus disjunkt), mogen we hun kansen optellen, zodat

$$P(\underline{x}=x) = \binom{n}{x} p^x q^{n-x} \text{ voor } x = 0, 1, \ldots, n.$$

Bij de behandeling van het binomium van NEWTON hebben we gezien, dat de som van deze kansen juist 1 is, zoals een kansverdeling betaamt.

De bovenstaande kansverdeling noemt men een binomiale verdeling. Deze hangt af van twee parameters; de kans p op succes en de steekproefomvang n. Als $p = \frac{1}{2}$, dan is deze kansverdeling symmetrisch om het punt $E\underline{x}$.

Voorbeeld 2.24. Men doet drie onafhankelijke worpen met een zuivere dobbelsteen. Hoe groot zijn nu de kansen op nul, één, twee of drie-maal een even uitkomst? Noemen we het optreden van een even uitkomst bij een worp een succes, dan is de kans op succes $p = \frac{1}{2}$. Is $\underline{x}$ het aantal seccessen, dan is dus

$$P(\underline{x}=0) = (\tfrac{1}{2})^3 = \tfrac{1}{8},\ P(\underline{x}=1) = \binom{3}{1}(\tfrac{1}{2})^3 = \tfrac{3}{8},\ P(\underline{x}=2) = \binom{3}{2}(\tfrac{1}{2})^3 = \tfrac{3}{8},$$
$$P(\underline{x}=3) = (\tfrac{1}{2})^3 = \tfrac{1}{8}.$$

Voorbeeld 2.25. Beschouw een vaas met vier maal zoveel rode als witte knikkers. Zij $\underline{x}$ het aantal rode knikkers in een steekproef met teruglegging van omvang 10, dan kunnen we m.b.v. bovenstaande formule de kansverdeling van $\underline{x}$ weer uitrekenen (het trekken van een rode knikker is nu een succes, de kans op succes is $\frac{4}{5}$). Deze kansverdeling is geschetst in fig. 2.9, zie vb. 2.15.

We bepalen thans verwachting en variantie van $\underline{x}$. Dit geschiedt het eenvoudigst door invoering van de stochastische grootheden $\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_n$; $\underline{y}_i$ is het aantal successen bij de i-de trekking. Elke $\underline{y}_i$ kan dus alleen de waarde 1 (bij S, met kans p) of 0 (bij M, met kans q = 1-p) aannemen. Nu is

$$E\underline{y}_i = 0.\ P(\underline{y}_i=0) + 1.\ P(\underline{y}_i=1) = p$$

en

$$\sigma^2(\underline{y}_i) = E(\underline{y}_i - p)^2 = E\underline{y}_i^2 - 2pE\underline{y}_i + p^2 = p - p^2 = pq.$$

Daar $\underline{x} = \underline{y}_1 + \underline{y}_2 + \ldots + \underline{y}_n$, volgt nu direkt

$$E\underline{x} = E(\underline{y}_1+\underline{y}_2+\ldots+\underline{y}_n) = E\underline{y}_1 + E\underline{y}_2 + \ldots + E\underline{y}_n = np.$$

De grootheden $\underline{y}_1$, $\underline{y}_2$, ..., $\underline{y}_n$ zijn onafhankelijk ondersteld, dus

$$\sigma^2(\underline{x}) = \sigma^2(\underline{y}_1+\underline{y}_2+\ldots+\underline{y}_n) = \sigma^2(\underline{y}_1) + \sigma^2(\underline{y}_2) + \ldots + \sigma^2(\underline{y}_n) = npq.$$

De binomiale verdeling is voor grote steekproefomvang n lastig te hanteren. Daarom is het van groot belang, dat bij grote n en p niet te dicht bij 0 of 1 de binomiale verdeling veel gelijkt op een normale verdeling met dezelfde verwachting en variantie. Zo is voor n = 100 en p = 0,4 de kansverdeling van $\underline{x}$ in de vorm van een histogram geschetst in fig. 2.19, tezamen met de kansdichtheid van de normale verdeling met verwachting 40 en variantie 24.

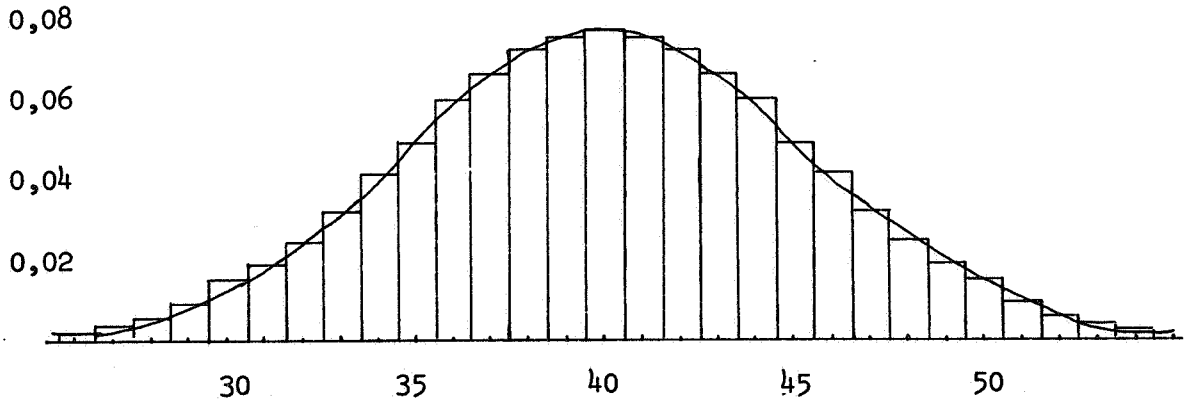


fig. 2.19. Histogram van de kansverdeling van een binomiaal verdeelde $\underline{x}$ (n=100, p=0,4) en de kansdichtheid van de N(40, 24) verdeling.

We kunnen de kansen $P(\underline{x} \leq x|\underline{x}$ binomiaal) dus zeer goed benaderen met $P(\underline{x} \leq x|\underline{x}$ normaal). Uit de figuur zien we, dat deze benadering nog wat beter wordt, als we de binomiale kans benaderen met $P(\underline{x} \leq x + \frac{1}{2}|\underline{x}$ normaal). Deze $\frac{1}{2}$ wordt de "continuïteitscorrectie" genoemd. Zo is in het onderhavige geval $P(\underline{x} \leq 50|\underline{x}$ binomiaal) = 0,9832, terwijl

$P(\underline{x} \leq 50\frac{1}{2}|\underline{x}$ normaal$) = P(\underline{u} \leq \frac{10\frac{1}{2}}{\sqrt{24}}) = P(\underline{u} \leq 2{,}14) = 0{,}9838$, zodat de benadering inderdaad zeer redelijk is.

Wordt een kans $P(\underline{x} \geq x|\underline{x}$ binomiaal$)$ gevraagd, dan werkt de continuiteitscorrectie juist de andere kant op en benadert men met $P(\underline{x} \geq x - \frac{1}{2}|\underline{x}$ normaal$)$.

Willen we de kans p schatten, dan ligt het voor de hand deze te schatten met de fraktie successen in onze steekproef, dus met $\underline{x}/n$. Deze schatter is zuiver, want

$$E(\frac{\underline{x}}{n}) = \frac{1}{n} E\underline{x} = \frac{1}{n} \cdot np = p.$$

De variantie van deze schatter is

$$\sigma^2(\frac{\underline{x}}{n}) = \frac{1}{n^2} \sigma^2(\underline{x}) = \frac{pq}{n},$$

zodat de variantie afneemt met toenemende steekproefomvang. Onder alle zuivere schatters is $\underline{x}/n$ bovendien de schatter met de kleinste variantie.

Het is ook mogelijk voor p betrouwbaarheidsintervallen te bepalen. Het is niet zo eenvoudig de onder- en bovengrens in een formule (gebaseerd op de waarnemingsuitkomsten) weer te geven; we volstaan met het geven van twee figuren 2.20 en 2.21. In beide figuren kan men bij gevonden steekproeffraktie successen de ondergrens en bovengrens van het betrouwbaarheidsinterval voor p aflezen bij een aantal verschillende steekproefomvangen, in de eerste figuur bij een onbetrouwbaarheid $\alpha = 10\%$, in de tweede figuur bij een onbetrouwbaarheid $\alpha = 5\%$.

We zien in beide figuren, dat het interval steeds de schatting x/n van p bevat en nauwer wordt naarmate n toeneemt. De intervallen bij $\alpha = 0{,}10$ zijn nauwer dan bij $\alpha = 0{,}05$ bij dezelfde n en x/n: wie meer risiko op een foute uitspraak neemt kan ook nauwkeuriger uitspraken doen!

Voorbeeld 2.26. In een groot park vangt men 50 konijnen, waarvan er vijf besmet blijken te zijn met bepaalde bacillen. Vatten we de gevangen konijnen op als een aselekte steekproef uit de populatie van alle konijnen uit het park (of dit geoorloofd is hangt af van de wijze waarop de dieren zijn gevangen), dan kunnen we voor de populatiefraktie p van besmette

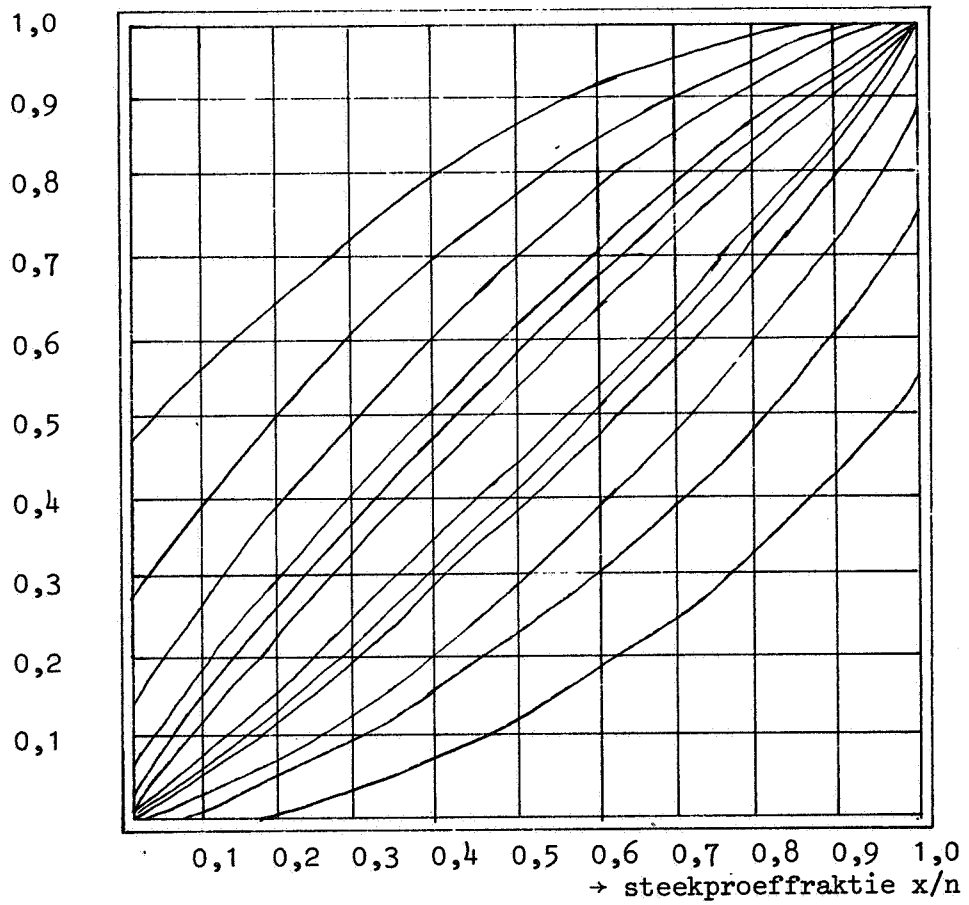


fig. 2.20. Betrouwbaarheidsgrenzen voor een kans p met een onbetrouwbaarheid 0,10 van het interval, als funktie van de steekproeffraktie x/n.

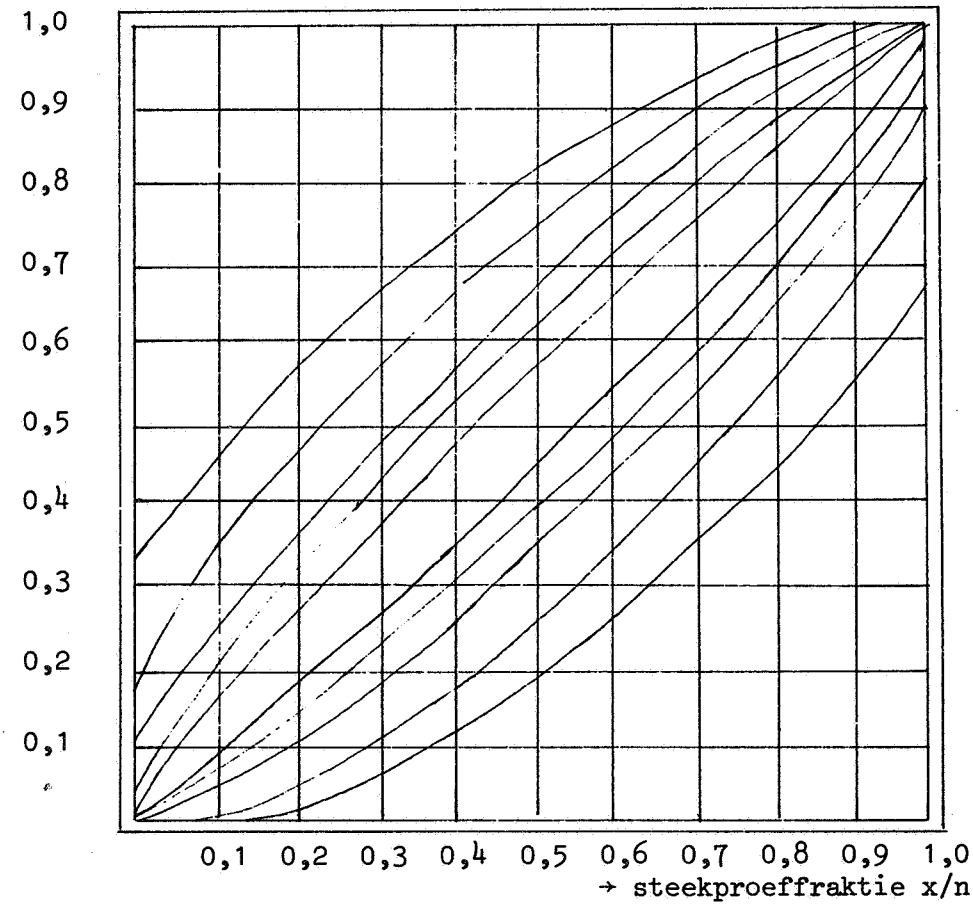


fig. 2.21. Betrouwbaarheidsgrenzen voor een kans p met een onbetrouwbaarheid 0,05 van het interval, als funktie van de steekproeffraktie x/n.

konijnen een betrouwbaarheidsinterval opstellen. Bij een onbetrouwbaarheid $\alpha = 0{,}10$ heeft dit interval de gedaante $0{,}04 < p < 0{,}20$ (zie fig. 2.20), bij $\alpha = 0{,}05$ is het interval $0{,}03 < p < 0{,}22$ (zie fig. 2.21).

Ook hypothesen over de waarde van p kan men toetsen. De toetsingsgrootheid is steeds het aantal successen $\underline{x}$ in de getrokken steekproef. Wenst men bij een gegeven onbetrouwbaarheid α de nulhypothese

$$H_0: p \leq p_0 \ (p_0 \text{ is een gegeven getal})$$

te toetsen, dan zal men in de kritieke zone grote waarden van $\underline{x}$ willen opnemen, daar deze slecht in overeenstemming zijn met H_0 en veeleer zullen optreden als $p > p_0$ is. De toetsingsprocedure heeft dus de gedaante:

$$\text{verwerp } H_0 \text{ als } \underline{x} \geq x \ ,$$

waarin x_α een geheel getal is met de eigenschap

$$P(\underline{x} \geq x_\alpha \text{ als } H_0 \text{ waar}) \leq \alpha.$$

Daar de kans op grote waarden van $\underline{x}$ afneemt bij dalende p, is aan deze eis zeker voldaan wanneer x_α voldoet aan

$$P(\underline{x} \geq x_\alpha \text{ als } p = p_0) \leq \alpha.$$

Het onderscheidingsvermogen van deze toets is zo groot mogelijk als de kritieke zone zo groot mogelijk is; we zullen dus voor x het kleinste gehele getal zoeken dat nog aan bovenstaande voorwaarde voldoet.

Wenst men omgekeerd de hypothese

$$H_0: p \geq p_0$$

te toetsen, dan zal men op grond van dezelfde redenering H_0 verwerpen als $\underline{x} \leq x_\alpha$ is, waarin x_α het grootste gehele getal is dat voldoet aan

$$P(\underline{x} \leq x_\alpha \text{ als } p = p_0) \leq \alpha.$$

Als men de tweezijdige hypothese

$$H_0: p = p_0$$

wilt toetsen met als alternatieve hypothese

$$H_1: p < p_0 \quad \text{of} \quad p > p_0,$$

dan zal men ook een kritieke zone kiezen die uit twee stukken bestaat: kleine waarden en grote waarden van $\underline{x}$. De toetsing verloopt dan als volgt:

verwerp de nulhypothese H_0

ten gunste van $p < p_0$ als $\underline{x} \leq x_{1,\alpha}$

of

ten gunste van $p > p_0$ als $\underline{x} \geq x_{2,\alpha}$;

hierbij moet de tweezijdige kritieke zone Z voldoen aan de voorwaarde $P(\underline{x}$ neemt waarde in Z aan als $p = p_0) \leq \alpha$, hetgeen betekent

$$P(\underline{x} \leq x_{1,\alpha} \text{ als } p = p_0) + P(\underline{x} \geq x_{2,\alpha} \text{ als } p = p_0) \leq \alpha.$$

Dit kan men bereiken door er voor te zorgen, dat $x_{1,\alpha}$ het grootste gehele getal is dat nog voldoet aan

$$P(\underline{x} \leq x_{1,\alpha} \text{ als } p = p_0) \leq \frac{\alpha}{2}$$

en $x_{2,\alpha}$ het kleinste gehele getal is dat nog voldoet aan

$$P(\underline{x} \geq x_{2,\alpha} \text{ als } p = p_0) \leq \frac{\alpha}{2}.$$

Aan de hand van een tweetal voorbeelden zullen we demonstreren hoe deze berekening verloopt.

Voorbeeld 2.27. Volgens de wetten van MENDEL zouden onder bepaalde genetische omstandigheden uit kruising van een plant met rode en een plant met

witte bloemen planten moeten ontstaan, die elk met kans $p = \frac{1}{2}$ rose bloemen hebben. Een onderzoeker meent echter, dat deze theorie in het onderhavige geval niet juist is, zonder zich overigens te durven uitspreken over de vraag of $p > \frac{1}{2}$ dan wel $p < \frac{1}{2}$ is. Om de theorie te toetsen kiest hij aselekt 11 zaadjes uit een groot aantal zaadjes ontstaan bij de kruising van beide planten, kweekt hieruit 11 planten op en stelt vast hoeveel van deze planten rose bloemen hebben. Zij dit aantal $\underline{x}$; dan kan $\underline{x}$ dus één van de waarden 0, 1, ..., 11 aannemen. Als nulhypothese kiezen we H_0: $p = \frac{1}{2}$, met als alternatieve hypothese H_1: $p < \frac{1}{2}$ of $p > \frac{1}{2}$. Voor de onbetrouwbaarheid van de toets wordt $\alpha = 0{,}05$ gekozen. Om thans het grootste getal $x_{1,\alpha}$ te vinden zodat voor $p = \frac{1}{2}$ voldaan is aan $P(\underline{x} \leq x_{1,\alpha}) \leq \frac{\alpha}{2} = 0{,}025$, bepalen we de volgende kansen (steeds voor het geval $p = \frac{1}{2}$):

$$P(\underline{x}=0) = (\tfrac{1}{2})^{11} = \frac{1}{2048} \approx 0{,}0005;$$

$$P(\underline{x}=1) = \binom{11}{1}(\tfrac{1}{2})^{11} \approx 0{,}0054;$$

$$P(\underline{x}=2) = \binom{11}{2}(\tfrac{1}{2})^{11} \approx 0{,}0269;$$

$$P(\underline{x}=3) = \binom{11}{3}(\tfrac{1}{2})^{11} \approx 0{,}0806.$$

Hieruit lezen we af dat $x_{1,\alpha} = 1$, immers $P(\underline{x} \leq 1) \approx 0{,}006 < 0{,}025$, maar $P(\underline{x} \leq 2) \approx 0{,}032 > 0{,}025$. Op dezelfde wijze gaat men na, dat $x_{2,\alpha} = 10$. We verwerpen dus H_0 ten gunste van $p < \frac{1}{2}$ als $\underline{x}$ de waarde 0 of 1 aanneemt en ten gunste van $p > \frac{1}{2}$ als $\underline{x}$ de waarde 10 of 11 aanneemt; in de overige gevallen wordt H_0 niet verworpen. Vermoedde de onderzoeker reeds vóór het experiment, dat $p < \frac{1}{2}$ is, dan is het beter als nulhypothese H_0: $p > \frac{1}{2}$ te kiezen; de toets is dan een linkséénzijdige toets met als kritieke zone de waarden $\underline{x} \leq x_\alpha$ met $x_\alpha = 2$, zoals eveneens uit de bovenstaande kansen volgt.

Voorbeeld 2.28. Er wordt wel eens beweerd, dat in een gezin, waarin alle kinderen jongens zijn, de kans p op een volgende jongensgeboorte groter is dan $\frac{1}{2}$. Om deze bewering te onderzoeken trekt men uit de Nederlandse bevolking aselekt 200 gezinnen met twee jongens (en géén meisjes) en wacht af of in deze gezinnen nog een kind geboren wordt. In 82 gezinnen blijkt dit het geval; hierbij worden 49 jongensgeboorten en 33 meisjesgeboorten

geteld (géén tweelingen). Als nulhypothese kiezen we H_0: $p \leq \frac{1}{2}$, als alternatieve hypothese H_1: $p > \frac{1}{2}$, terwijl voorts $\alpha = 0{,}01$ wordt genomen. Noemen we het aantal jongensgeboorten in de 82 gezinnen $\underline{x}$, dan bestaat de kritieke zone uit alle waarden $\underline{x} \geq x_\alpha$, waarin x het kleinste gehele getal is dat voldoet aan $P(\underline{x} \geq x_\alpha \text{ als } p = \frac{1}{2}) \leq 0{,}01$. Berekening van x_α als in vb. 2.22 vereist nu veel rekenwerk *); daar de steekproefomvang n vrij groot is, kunnen we de normale benadering toepassen. Als $p = \frac{1}{2}$, is $E\underline{x} = 41$ en $\sigma^2(\underline{x}) = 20\frac{1}{2}$, zodat $P(\underline{x} \geq x_\alpha \text{ als } p = \frac{1}{2}) \approx P(\underline{u} \geq \frac{x_\alpha - \frac{1}{2} - 41}{\sqrt{20{,}5}}) \leq 0{,}01$. Daar volgens tabel 1 $P(\underline{u} \geq 2{,}33) \approx 0{,}01$, moet dus $x_\alpha - \frac{1}{2} - 41 \geq 2{,}33\sqrt{20{,}5}$ of $x_\alpha \geq 52{,}07$, dus $x_\alpha = 53$. (Uit tabellen blijkt dat we $x_\alpha = 52$ mogen kiezen, daar $P(\underline{x} \geq 52) = 0{,}0099$ als $p = \frac{1}{2}$.) Daar in het experiment $x = 49$, wordt H_0 niet verworpen en kan men de bewering experimenteel (nog) niet bevestigd achten.

Het is ook mogelijk het onderscheidingsvermogen van de toets bij een specifiek alternatief uit te rekenen. Toetst men bijv. de nulhypothese H_0: $p \leq \frac{1}{2}$ tegen de alternatieve hypothese H_1: $p > \frac{1}{2}$, met als kritieke zone $Z = \{\text{uitkomsten } x \geq x_\alpha\}$, dan is het onderscheidingsvermogen bij een speciaal alternatief p_1 ($\frac{1}{2} < p_1 < 1$) gelijk aan

$$P(\underline{x} \text{ valt in } Z \text{ als } p = p_1) = P(\underline{x} \geq x_\alpha \text{ als } p = p_1)$$

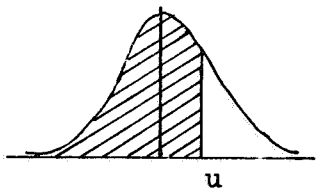
welke kans we, hetzij exakt, hetzij met de normale benadering kunnen berekenen.

De hier besproken <u>binomiale toetsen</u> voor een kans worden ook vaak toegepast in een andere situatie, nl. als men n onafhankelijke waarnemingen $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ verricht en de hypothese H_0 wil toetsen, dat de mediaan van de verdeling van de waarnemingen nul is. Zijn de waarnemingen normaal verdeeld (met onbekende variantie σ^2), dan kan men de toets van STUDENT toepassen (zie §8). Is dit echter niet het geval (bijv. als de waarnemingen diskreet zijn), dan kan men als volgt te werk gaan. Laten we

*) Er bestaan voor $n \leq 100$ echter uitvoerige tabellen van de binomiale verdeling voor een groot aantal waarden van p.

de uitkomsten nul buiten beschouwing, dan geldt onder H_0: $P(\underline{x} > 0) = P(\underline{x} < 0) = \frac{1}{2}$. Het aantal positieve uitkomsten heeft dus onder H_0 een binomiale verdeling met $p = \frac{1}{2}$. Dit betekent dat we H_0 met een binomiale toets kunnen toetsen. Men spreekt dan meestal van een tekentoets.

Tabel 1. Verdelingsfunktie $P(\underline{u} \leq u)$ voor de standaard-normale verdeling



Waarden van 10^4. $P(\underline{u} \leq u)$
voor u = 0,00 (0,01) 3,49

u	0	1	2	3	4	5	6	7	8	9
0,0	5000	5040	5080	5120	5160	5199	5239	5279	5319	5359
0,1	5398	5438	5478	5517	5557	5596	5636	5675	5714	5753
0,2	5793	5832	5871	5910	5948	5987	6026	6064	6103	6141
0,3	6179	6217	6255	6293	6331	6368	6406	6443	6480	6517
0,4	6554	6591	6628	6664	6700	6736	6772	6808	6844	6879
0,5	6915	6950	6985	7019	7054	7088	7123	7157	7190	7224
0,6	7257	7291	7324	7357	7389	7422	7454	7486	7517	7549
0,7	7580	7611	7642	7673	7704	7734	7764	7794	7823	7852
0,8	7881	7910	7939	7967	7995	8023	8051	8078	8106	8133
0,9	8159	8186	8212	8238	8264	8289	8315	8340	8365	8389
1,0	8413	8438	8461	8485	8508	8531	8554	8577	8599	8621
1,1	8643	8665	8686	8708	8729	8749	8770	8790	8810	8830
1,2	8849	8869	8888	8907	8925	8944	8962	8980	8997	9015
1,3	9032	9049	9066	9082	9099	9115	9131	9147	9162	9177
1,4	9192	9207	9222	9236	9251	9265	9279	9292	9306	9319
1,5	9332	9345	9357	9370	9382	9394	9406	9418	9429	9441
1,6	9452	9463	9474	9484	9495	9505	9515	9525	9535	9545
1,7	9554	9564	9573	9582	9591	9599	9608	9616	9625	9633
1,8	9641	9649	9656	9664	9671	9678	9686	9693	9699	9706
1,9	9713	9719	9726	9732	9738	9744	9750	9756	9761	9767
2,0	9772	9778	9783	9788	9793	9798	9803	9808	9812	9817
2,1	9821	9826	9830	9834	9838	9842	9846	9850	9854	9857
2,2	9861	9864	9868	9871	9875	9878	9881	9884	9887	9890
2,3	9893	9896	9898	9901	9904	9906	9909	9911	9913	9916
2,4	9918	9920	9922	9925	9927	9929	9931	9932	9934	9936
2,5	9938	9940	9941	9943	9945	9946	9948	9949	9951	9952
2,6	9953	9955	9956	9957	9959	9960	9961	9962	9963	9964
2,7	9965	9966	9967	9968	9969	9970	9971	9971	9973	9974
2,8	9974	9975	9976	9977	9977	9978	9979	9979	9980	9981
2,9	9981	9982	9982	9983	9984	9984	9985	9985	9986	9986
3,0	9987	9987	9987	9988	9988	9989	9989	9989	9990	9990
3,1	9990	9991	9991	9991	9992	9992	9992	9992	9993	9993
3,2	9993	9993	9994	9994	9994	9994	9994	9995	9995	9995
3,3	9995	9995	9995	9996	9996	9996	9996	9996	9996	9997
3,4	9997	9997	9997	9997	9997	9997	9997	9997	9997	9998

Een 5 betekent, dat bij afronding naar omlaag moet worden afgerond
Een 5 wordt naar boven afgerond.

Tabel 2. Rechtse α-punten van de t-verdelingen van STUDENT

Bij gegeven α en aantal vrijheidsgraden ν is getabelleerd $t_{\nu;\alpha}$ gedefinieerd door $P(\underline{t}_\nu > t_{\nu;\alpha}) = \alpha$.

ν \ α	0,10	0,05	0,025	0,01	0,005
1	3,078	6,314	12,706	31,821	63,657
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,576	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
25	1,316	1,708	2,060	2,485	2,787
30	1,310	1,697	2,042	2,457	2,750
40	1,303	1,684	2,021	2,423	2,704
60	1,296	1,671	2,000	2,390	2,660
∞	1,282	1,645	1,960	2,326	2,576