Navorsings- en Oorsigartikels

Die fundamentele belang van differensiaalvergelykings met drie singulariteite vir die Wiskundige Statistiek

H.S. Steyn
Departement Statistiek, Universiteit van Suid-Afrika

UITTREKSEL

Dit is bekend dat die oplossings van tweede-orde lineêre differensiaalvergelykings met hoogstens vyf singulariteite 'n fundamentele rol speel in die Wiskundige Fisika. In hierdie artikel word aangetoon dat dieselfde bewering ook geld vir die Wiskundige Statistiek maar met die verskil dat 'n vergelyking met drie reguliere singulariteite voldoende is. Twee wye klasse van waarskynlikheidsverdelings, een vir kontinue verdelings en een vir diskrete verdelings, word gedefinieer as oplossings van so 'n differensiaalvergelyking. Daarna word aangetoon dat al die verdelings wat normaalweg in die Wiskundige Statistiek as van belang geag word in hierdie klasse bevat word. In die geval van kontinue verdelings is die waarskynlikheidsfunksies oplossings van die genoemde differensiaalvergelyking terwyl vir diskrete verdelings die waarskynlikheidsvoortbringende funksies wel oplossings van so 'n vergelyking is. Deur geskikte meerdimensionale uitbreidings te definieer volg ooreenkomstige differensiaalvergelykings vir kontinue en diskrete meerveranderlike verdeling.

ABSTRACT

The fundamental importance of differential equations with three singularities in Mathematical Statistics

It is well-known that the solution of a second order linear differential equation with at most five singularities plays a fundamental role in Mathematical Physics. In this paper it is shown that this statement also applies to Mathematical Statistics but with the difference that an equation with three singularities will suffice. Two wide classes of probability distributions are defined as solutions of such a differential equation, one for continuous distributions and one for discrete distributions. These two classes contain as members all the distributions which are normally considered as of importance in Mathematical Statistics. In the continuous case the probability functions are solutions of the relevant second order equation, while in the discrete case the probability generating functions are solutions there-of. By defining appropriate multidimentional extensions corresponding differential equations are obtained for continuous and discrete multivariate distributions.

1. INLEIDING

Reeds aan die einde van die vorige eeu het wiskundiges soos Klein en Bôcher (kyk Whittaker en Watson[20] bl. 203) aangetoon dat al die lineêre differensiaalvergelykings, wat toe van belang was en tans nog in die Wiskundige Fisika van belang is, spesiale gevalle of konfluente gevalle is van 'n algemene tweede-orde differensiaalvergelyking met vyf singulariteite a_1, a_2, a_3, a_4 en ∞ waar al vyf hierdie punte reguliere punte is.

Dat lineêre differensiaalvergelykings ook in die Wiskundige Statistiek 'n belangrike rol speel, is welbekend. Die bekendste hiervan is die differensiaalvergelyking van die eerste-orde wat ten grondslag lê aan Pearson se stelsel van frekwensiekrommes, naamlik:

$$(b_0 + b_1x + b_2x^2)\frac{d}{dx} f(x) + (a_0 + a_1x)f(x) = 0 . \qquad (1)$$

Die bekendste kontinue eenveranderlike waarskynlikheidsverdelings van die Wiskundige Statistiek soos die normale, die beta- en die gammaverdelings, is almal oplossings van die vergelyking (1). Verder, indien F(t) die waarskynlikheidsvoortbringende funksie (wvf) is van enkele van die bekendste diskrete eenveranderlike waarskynlikheidsverdelings soos die Poisson-, die binomiale, die negatiewe binomiale of trinomiale verdelings dan kan direk aangetoon word dat F(t) 'n vergelyking soortgelyk aan (1) bevredig. Dit is ook bekend (kyk Steyn[16]) dat 'n omvattende klas van diskrete verdelings, die sogenaamde hipergeometriese klas, voortgebring word deur $K\,_2F_1(a;b;c;t)$, waar K 'n konstante is en $_2F_1(\cdot)$ 'n oplossing is van die hipergeometriese vergelyking, naamlik:

$$t(1-t)\frac{d^2}{dt^2}F(t) + [c-(a+b+1)t]\frac{d}{dt}F(t) - ab\,F(t) = 0 \qquad (2)$$

In die afgelope twee dekades het lineêre parsiële differensiaalvergelykings van die tweede-orde sporadies verskyn in die veld van kontinue meerveranderlike waarskynlikheidsverdelings (kyk bv. Muirhead[12] bl. 259-281). 'n Omvattende klas van diskrete meerveranderlike verdelings wat voortgebring word deur veralgemeende hipergeometriese funksies, en wat 'n stelsel van tweede-orde parsiële differensiaalvergelykings bevredig, is gedefinieer deur Steyn.[16]

In die volgende afdelings word aangetoon dat, soos in die geval van die Wiskundige Fisika, daar ook in die Wiskundige Statistiek ten opsigte van waarskynlikheidsverdelings vir alle praktiese redes volstaan kan word met oplossings van tweede-orde lineêre differensiaalverdelings, maar anders as in die geval van Wiskundige Fisika besit hierdie differensiaalvergelykings *drie* (in plaas van vyf) reguliere singulariteite. Die parsiële differensiaalvergelykings wat ten opsigte van meerveranderlike verdelings hanteer sal word, is logiese uitbreidings van die eenveranderlike differensiaalvergelyking met drie singulariteite.

Die doel van hierdie artikel is om redelik oorsigtelik te wees en om aan te toon hoe omvattend die klas van waarskynlikheidsverdelings is wat met die genoemde differensiaalvergelyking geassosieer is. Daarom sal al die wiskundige verwerkings nie aangetoon word nie. Daar kan aanvaar word dat òf die definisiegebiede van die toevalsveranderlikes en die genererende simbole òf die parameters wat voorkom, sodanig is dat die singuliere punte nie die konvergensie-eienskappe beïnvloed nie.

2. RIEMANN SE VERGELYKING EN UITBREIDING VAN DIE PEARSONSTELSEL VIR EENVERANDERLIKE VERDELINGS

Soos reeds genoem, omvat Pearson se stelsel, wat deur vergelyking (1) gedefinieer word, wel 'n aantal van die bekendste eenveranderlike kontinue verdelings, maar 'n wyer klas verdelings is om baie redes nodig. 'n Logiese uitbreiding van (1) sou wees om die tweede-orde lineêre differensiaalvergelyking:

$$(c_0+c_1x+c_2x^2+c_3x^3)\frac{d^2}{dx^2}f(x) + (b_0+b_1x+b_2x^2)\frac{d}{dx}f(x) + (a_0+a_1x)f(x) = 0$$

in oënskou te neem. Indien hierdie vergelyking in die standaardvorm en met $u = f(x)$ geskryf word as

$$\frac{d^2u}{dx^2} + p(x)\frac{du}{dx} + q(x)u = 0\,, \tag{3}$$

dan is dit op grond van die faktor $c_0+c_1x+c_2x^2+c_3x^3$ wat nou in die noemer verskyn, duidelik dat dit nodig word om tweede-orde lineêre vergelykings met drie singuliere punte te beskou. Laat a, b en c hierdie punte wees, terwyl daar aanvaar word dat hierdie punte regulier is met $\alpha,\alpha';\beta,\beta';\gamma,\gamma'$ as die oplossings van die indeksvergelykings, sodat (3) dan geskryf kan word (kyk Whittaker & Watson[20]) in die vorm:

$$\frac{d^2u}{dx^2} + \left\{\frac{1-\alpha-\alpha'}{x-a} + \frac{1-\beta-\beta'}{x-b} + \frac{1-\gamma-\gamma'}{x-c}\right\}\frac{du}{dx} + \left\{\frac{\alpha\alpha'(a-b)(a-c)}{x-a} + \frac{\beta\beta'(b-c)(b-a)}{x-b} + \frac{\gamma\gamma'(c-a)(c-b)}{x-c}\right\}\frac{u}{(x-a)(x-b)(x-c)} = 0 \tag{4}$$

waar $\alpha+\alpha'+\beta+\beta'+\gamma+\gamma' = 1$.

Die vergelyking (4) staan in die literatuur bekend as Riemann se vergelyking en die oplossing hiervan, aangedui deur,

$$u = P\begin{Bmatrix} a & b & c & \\ \alpha & \beta & \gamma & x \\ \alpha' & \beta' & \gamma' & \end{Bmatrix}, \tag{5}$$

word Riemann se P-vergelyking genoem.

Enkele spesiale gevalle van (5) sal nou hanteer word.

(i) Die bekende hipergeometriese vergelyking (2) is 'n spesiale geval van (4) en word gedefinieer deur die skema

$$P\begin{Bmatrix} 0 & \infty & 1 & \\ 0 & a & 0 & x \\ 1-c & b & c-a-b & \end{Bmatrix} \text{ en reduseer tot}$$

$$x(1-x)\frac{d^2u}{dx^2} + \{c-(a+b+1)x\}\frac{du}{dx} - abu = 0\,, \tag{6}$$

met singulariteite in die punte 0, ∞ en 1. Die bekendste partikuliere oplossing van (6) heet die hipergeometriese funksie en word aangedui deur ${}_2F_1(a;b;c;x)$ wat as 'n reeks geskryf kan word:

$${}_2F_1(a;b;c;x) = 1 + \frac{ab}{c}x + \frac{a(a+1)b(b+1)}{c(c+1)}\frac{x^2}{2!} + \ldots = \sum_{r=0}^{\infty}\frac{a^{[r]}b^{[r]}}{c^{[r]}}\frac{x^r}{r!}\,,$$

$$\text{waar } y^{[r]} = y(y+1)\ldots(y+r-1). \tag{7}$$

Hierdie reeks asook die sogenaamde veralgemeende hipergeometriese reeks

$${}_pF_q\begin{bmatrix} a_1,a_2,\ldots,a_p; x \\ c_1,c_2,\ldots,c_q \end{bmatrix} = \sum_{r=0}^{\infty}\frac{a_1^{[r]}a_2^{[r]}\ldots a_p^{[r]}}{c_1^{[r]}c_2^{[r]}\ldots c_q^{[r]}}\frac{x^r}{r!}\,,$$

is intensief gedurende die afgelope twee eeue deur wiskundiges bestudeer. Dit is welbekend in die teorie van spesiale funksies dat (6) wel 24 partikuliere oplossings besit, wat almal in terme van hipergeometriese reekse ${}_2F_1(\cdot)$ (kyk byv. Erdelyi[2] bl. 104) geskryf kan word.

(ii) Die konfluente hipergeometriese vergelyking word verkry indien $c \to \infty$ in die skema

$$P\begin{Bmatrix} 0 & \infty & c & \\ \tfrac{1}{2}+m & -c & c-k & x \\ \tfrac{1}{2}-m & 0 & k & \end{Bmatrix}$$

en reduseer dan tot

$$\frac{d^2u}{dx^2} + \frac{du}{dx} + \left(\frac{k}{x} + \frac{\tfrac{1}{4}-m^2}{x^2}\right)u = 0\,. \tag{8}$$

Die vergelyking (8) word deur $u = e^{-\frac{1}{2}x}W_{k,m}(x)$ omvorm tot

$$\frac{d^2w}{dx^2} + \left\{-\tfrac{1}{4} + \frac{k}{x} + \frac{\tfrac{1}{4}-m^2}{x^2}\right\}w = 0\,. \tag{9}$$

In die meer algemene vorm kan vergelyking (8) geskryf word as:

$$\frac{d^2y}{dx^2} + \left(a + \frac{b}{x}\right)\frac{dy}{dx} + \left(\alpha + \frac{\beta}{x} + \frac{\gamma}{x^2}\right)y = 0\,, \tag{10}$$

wat nog steeds 'n homogene tweede-orde differensiaalvergelyking is met 'n reguliere singuliere punt in

$x = 0$, wat die enigste eindige singuliere punt is.

Dit is welbekend (kyk Erdelyi[2] bl. 257) dat (10) opgelos kan word in terme van òf konfluente hipergeometriese funksies, aangedui deur ${}_1F_1(\cdot)$, òf in terme van Besselfunksies, aangedui deur ${}_0F_1(\cdot)$, wat oplossings is van 'n spesiale geval van (10) naamlik,

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} - (n^2 - x^2)y = 0. \quad (11)$$

(iii) Vir latere gebruik is dit nodig om daarop te let dat die skema

$$P\left\{\begin{matrix} 0 & \infty & 1 & \\ d & a & 0 & x \\ 1-c-d & b & c-a-b & \end{matrix}\right\},$$

die vergelyking (4) reduseer tot

$$x(1-x)\frac{d^2u}{dx^2} + \{c-(a+b+1)x\}\frac{du}{dx}$$

$$+\left\{\frac{d(1-c-d)}{x}\lim_{p\to\infty}\frac{(-p)(-1)}{(x-p)} + ab\lim_{p\to\infty}\frac{(p-1)p}{(x-p)^2}\right\}u = 0$$

$$\text{of}\quad x(1-x)\frac{d^2u}{dx^2} + \{c-(a+b+1)x\}\frac{du}{dx}$$

$$+\left\{\frac{d(1-c-d)}{x} - ab\right\}u = 0. \quad (12)$$

Die waarde van vergelykings (6) tot (12) vir die daarstelling van verskillende bekende waarskynlikheidsverdelings, sal in die volgende afdelings hanteer word. Die diskrete verdelings wat hanteer sal word, het almal waarskynlikheidsvoortbringende funksies wat oplossings is van hierdie vergelykings of van die uitbreidings daarvan na meerveranderlikes, terwyl die kontinue verdelings se waarskynlikheidsfunksies oplossings van die betrokke vergelykings of hulle uitbreidings sal wees.

3. DISKRETE VERDELINGS GEASSOSIEER MET RIEMANN SE VERGELYKING

3.1 Diskrete eenveranderlike verdelings

3.1.1 *Die eenveranderlike hipergeometriese familie van verdelings*

Die funksie $F(t) = K{}_2F_1(a,b;c;t)$ (13)

sal 'n wvf wees mits $K = [{}_2F_1(a,b;c;1)]^{-1}$, waar a, b en c sodanig is dat die reeks nie-negatief is. Aangesien dit welbekend is dat

$${}_2F_1(a,b;c;1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)},$$

volg dat die waarskynlikheidsfunksies wat deur (13) voortgebring word, gegee word deur

$$f(x) = \frac{\Gamma(c-a)\Gamma(c-b)}{\Gamma(c)\Gamma(c-a-b)}\frac{a^{[x]}b^{[x]}}{x!\,c^{[x]}}.$$

Die volgende bekende waarskynlikheidsverdelings is enkele voorbeelde uit die bogenoemde familie van verdelings wat deur (13) gedefinieer word:

(i) Die hipergeometriese waarskynlikheidsfunksie

$$\frac{n!}{x!(n-x)!}(Np)^{(x)}(Nq)^{(n-x)}/N^{(n)},$$

waar $r^{(s)} = r(r-1)\ldots(r-s+1)$, wat verkry word vir

$a = -n;\ b = -Np;\ c = Nq - n + 1$ in (13).

(ii) Die negatiewe hipergeometriese waarskynlikheidsfunksie

$$\frac{(m+x-1)!}{(m-1)!x!}(Np)^{(x)}(Nq)^{(m)}/N^{(m+x)},$$

wat volg uit (13) vir $a = m;\ b = -Np;\ c = -N + m$.

(iii) Die Eggenberger-Pólya- en inverse Eggenberger-Pólya-verdelings (kyk Steyn[16]).

(iv) Die Beta-binomiaal en die Beta-Pascalverdelings (kyk Ord[14]).

(v) Die waarskynlikheidsverdelings van Waring en van Yule (kyk Johnson & Kotz[6] bl. 251).

Ander bekende diskrete verdelings soos die binomiale, die negatiewe binomiale, die Poisson- en die Poission-gammaverdelings is almal grensvorme van die bogenoemde spesiale gevalle vir (13).

3.1.2 *Die verdelings van Katz en van Kemp en Kemp*

Die klas van waarskynlikheidsfunksies van Katz[7] is oplossings van die differensievergelyking

$$(x+1)f(x+1) = (\alpha+\delta x)f(x);\ x = 0,1,\ \ldots$$

terwyl die uitgebreide klas bestaan uit oplossings van die vergelyking

$$(x+\mu)f(x+1) = (\alpha+\delta x)f(x);\ x = 0,1,\ldots,$$

met $\alpha > 0$, $\delta < 1$, $\mu > 0$. (Kyk Gurland en Tripathi[3]). Die wvf volg direk as

$$G(t) = {}_2F_1(1,\alpha/\delta;\mu;\delta t)/{}_2F_1(1,\alpha/\delta;\mu;\delta) \quad (14)$$

en is gevolglik 'n oplossing van die hipergeometriese vergelyking (6) met $a = 1$, $b = \alpha/\delta$, $c = \mu$ en $x = \delta t$.

Wanneer $\delta = 1$ herlei (14) tot 'n spesiale geval van (13). Wanneer $\delta \to 0$, kan aangetoon word dat die konfluente hipergeometriese vorm

$$G(t) = {}_1F_1(1;\beta;\gamma t)/{}_1F_1(1;\beta;\gamma);\quad \gamma>,\ \beta>0 \quad (15)$$

verkry word en wat 'n oplossing is van die vergelyking (10). Dit is welbekend dat (15) die wvf is van die hiper-Poissonstelsel van verdelings wat bestaan uit die sub-Poisson-, Poisson- of super-Poissonverdeling al na $\beta < 1$, $\beta = 1$ of $\beta > 1$.

Kemp en Kemp[8] het 'n interessante klas van diskrete verdelings met wvf

${}_2F_1(a,b;c;\lambda(t-1))$, $\lambda = \pm 1$, as uitgangspunt gedefinieer.

Hierdie sogenaamde hipergeometriese faktoriaalmomentverdelings is oplossings van die hipergeometriese vergelyking, met veranderlike en parameters soos aangedui.

3.1.3 *Ord se stelsel van diskrete verdelings*

Ord[14] se stelsel is oplossings van die differensievergelyking

$$\Delta f(x-1) = \frac{a-x}{b_0+b_1x+b_2x(x-1)} f(x-1),\ x=0,1,\ldots$$

of

$$\{b_0+b_1x+b_2x(x-1)\}\{f(x)-f(x-1)\}=(a-x)f(x-1),$$

en lewer na vermenigvuldiging met t^x en sommering oor alle x-waardes die volgende differensiaalvergelyking in die wvf F(t):

$$t[b_2t(1-t)\frac{d^2F}{dt^2}+\{b_1-t(b_1+2b_2-1)\}\frac{dF}{dt}$$

$$-(b_0+b_1+a-1)F] = -b_0F$$

waar

$$F = F(t) = \sum_x f(x)t^x ,$$

of

$$t(1-t)\frac{d^2F}{dt^2}+\{c_1-t(c_1+2-c_2)\}\frac{dF}{dt}$$

$$+\{\frac{c_0}{t}-(c_0+c_1-c_2+d)\}F=0 \qquad (16)$$

waar

$$c_0=\frac{b_0}{b_2},\ c_1=\frac{b_1}{b_2},\ c_2=\frac{1}{b_2},\ d=\frac{a}{b_2} .$$

Wanneer $c_0=0$ het (16) die vorm van die hipergeometriese vergelyking (6) maar vir $c_0 \neq 0$ het (16) die vorm van vergelyking (12). Die relevante konstantes vir die Riemannskema (5) kan bepaal word deur die gelykstelling van konstantes in (12) en (16) volgens hulle posisies in die vergelykings. Ord se stelsel omvat gevolglik nog meer as die hipergeometriese stelsel maar bly steeds binne die oplossings van Riemann se vergelyking.

3.1.4 *Magreeksverdelings*

'n Aantal van die waarskynlikheidsfunksies wat in 3.1.1-3.1.3 genoem is, is ook bekende lede van die sogenaamde familie van magreekswaarskynlikheidsfunksies wat gedefinieer word deur $A(x)\theta^x/B(\theta)$ met

$$B(\theta) = \sum_{x=0}^{\infty} A(x)\theta^x.$$

'n Bekende lid van hierdie familie wat nog nie genoem is nie, is die logreeksverdeling met wvf

$$F(t) = \log(1-\theta t)/\log(1-\theta) .$$

Daar kan direk deur differensiasie aangetoon word dat F(t) 'n oplossing is van 'n spesiale geval van (10). In die algemeen sal die vorm van die differensiaalvergelyking natuurlik afhang van die vorm van A(x).

3.2 **Diskrete meerveranderlike verdelings**

Dit is vooraf nodig om 'n meerdimensionale reeks van hipergeometriese tipe te definieer wat as wvf kan optree vir die meerveranderlike uitbreidings van die voorafgaande eenveranderlike verdelings van 3.1.1 en 3.1.2. Hierdie reeks is:

$$KF_1(a;b_1,b_2,\ldots,b_k;c;t_1,t_2,\ldots,t_k)$$

$$= K\sum_{x_1}\ldots\sum_{x_k}\frac{a^{\left[\sum_1^k x_i\right]}}{c^{\left[\sum_1^k x_i\right]}}\left(\prod_{i=1}^{k} b_i^{[x_i]}\frac{t_i^{x_i}}{x_i!}\right), \qquad (17)$$

$$\text{waar } K^{-1} = F_1(a;b_1,b_2,\ldots,b_k;c;1,1,\ldots,1)$$

$$= {}_2F_1(a;\sum_{i=1}^{k} b_i;c;1)$$

$$= \frac{\Gamma(c)\Gamma(c-a-\sum_{i=1}^{k} b_i)}{\Gamma(c-a)\Gamma(c-\sum_{i=1}^{k} b_i)},$$

en waar die parameters in (17) sodanig is dat die terme van die reeks eindig en nie-negatief bly en sodat $F_1(\cdot)$ konvergeer vir $|t_i| \leqslant 1$; $i=1,2,\ldots,k$. Die meerdimensionale reeks (17) is 'n verdere uitbreiding van Appell se F_1 hipergeometriese reeks (kyk Erdelyi[2] bl. 224) en dit kan geredelik aangetoon word dat (17) 'n oplossing is van die parsiële differensiaalvergelykings (kyk Steyn[16]):

$$(1-t_i)\sum_{j=1}^{k} t_j\frac{\partial^2F}{\partial t_i\partial t_j}+\{c-(a+b_i+1)t_i\}\frac{\partial F}{\partial t_i}-b_i\sum_{j\neq i} t_j\frac{\partial F}{\partial t_j}$$

$$= a\,b_iF \quad , \quad i=1,2,\ldots,k . \qquad (18)$$

Dit is duidelik dat (18) 'n meerdimensionale uitbreiding is van die hipergeometriese vergelyking (6).

Ooreenkomstiglik kan daar meerdimensionale uitbreidings gedefinieer word vir die diskrete stelsels van Katz, Kemp en Kemp en van Ord waaruit parsiële lineêre differensiaalvergelykings vir die wvf volg. Vir die huidige doel sal egter volstaan word met (17) en (18) aangesien hierdie stelsel reeds die bekendste diskrete meerveranderlike verdelings omvat.

Enkele voorbeelde van diskrete meerveranderlike verdelings wat deur (17) gegenereer word is die volgende:

(i) Die meerveranderlike hipergeometriese verdeling met waarskynlikheidsfunksie:

$$(n!/\prod_0^k x_i!)\prod_0^k (Np_i)^{(x_i)}/N^{(n)}$$

$$\text{waar} \quad \sum_0^k x_i=n\ ,\ \sum_0^k Np_i=N .$$

(ii) Die negatiewe meerveranderlike hipergeometriese verdeling met waarskynlikheidsfunksie:

$$\left[(m+\sum_1^k x_i-1)!/((m-1)!\prod_1^k x_i!)\right]\prod_1^k (Np_i)^{(x_i)}Np_0^{(m)}/N^{\left(m+\sum_1^k x_i\right)}$$

$$\text{met} \quad \sum_0^k Np_i = N .$$

(iii) Die meerveranderlike Eggenberger-Pòlyaverdelings wat met (i) en (ii) ooreenkom (Kyk Steyn[16]).

(iv) Die multinomiaal-Dirichlet en negatiewe-multinomiaal-Dirichletverdeling (Mosimann[9,10]).

(v) Die verdeling van die selfrekwensies van 'n $2 \times k$ gebeurlikheidstabel met vaste randtotale. (Die verdeling geassosieer met die selfrekwensies van 'n $h \times k$ gebeurlikheidstabel berus op 'n verdere uitbreiding van die $F_1(\cdot)$ reeks wat in (17) gedefinieer is. Die betrokke parsiële differensiaalvergelykings is egter nog steeds lineêre differensiaalvergelykings van die tweede-orde wat direkte meerdimensionale uitbreidings van (6) is).

Die bekende multinomiale, negatiewe multinomiale en meerveranderlike Poissonverdeling is almal limietvorme van (i) en (ii) hierbo wanneer n tot oneindige groei. Verdere voorbeelde kan nog gevind word in Janardan[4] en in Janardan en Patil.[5]

4. KONTINUE VERDELINGS GEASSOSIEER MET RIEMANN SE VERGELYKING

4.1 Eenveranderlike kontinue verdelings

Soos reeds genoem is in die inleidende paragraaf is die meeste van die bekende eenveranderlike kontinue verdelings oplossings van Pearson se vergelykings (1). Vir die huidige doel word in hierdie afdeling gevolglik gelet op verdelings wat nie onder die Pearsonstelsel val nie, maar wat wel onder die oplossings van Riemann se vergelykings tel. Slegs enkele van die bekendste verdelings word as voorbeelde genoem.

(i) Die nie-sentrale χ^2-verdeling:

Indien $z = \sum_{i=1}^{n} x_i^2$, waar x_i 'n waarneming is uit 'n normaal, $N(m_i,1)$, populasie en die x_i's onafhanklik verdeel is dan word die mvf van z gegee deur:

$$M(\alpha) = C\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} \exp\{-\tfrac{1}{2}\sum_{i=1}^{n}(x_i - m_i)^2\}\exp(\alpha\sum_{i=1}^{n}x_i^2)\, dx_1 dx_2 \ldots dx_n$$

met C 'n konstante. Differensiasie van $M(\alpha)$ lewer na enkele bewerkings die vergelyking

$$(1-2\alpha)^2 \frac{dM(\alpha)}{d\alpha} = \{n(1-2\alpha) + \lambda^2\}M(\alpha) ,$$

met $\lambda^2 = \sum_{i=1}^{n} m_i^2$. Uit hierdie vergelyking word die volgende differensiaalvergelyking in die waarskynlikheidsfunksie (kyk Steyn[18]) verkry:

$$(z - n - \lambda^2 + 4)\phi(z) + (4z + 8 - 2n)\frac{d}{dz}\phi(z) + 4z\frac{d^2}{dz^2}\phi(z) = 0 , \qquad (19)$$

wat 'n homogene lineêre differensiaalvergelyking van die tweede-orde is en wel van die konfluente hipergeometriese tipe (10) is en dus opgelos kan word in terme van òf 'n ${}_1F_1(\cdot)$ òf 'n ${}_0F_1(\cdot)$ funksie. Die oplossing van (19) is egter welbekend in statistiese literatuur as die waarskynlikheidsfunksie vir die niesentrale χ^2-verdeling (Muirhead[12] bl. 22) en word gegee deur

$$\phi(z) = [e^{-\frac{1}{2}\lambda^2}\, {}_0F_1(\tfrac{1}{2}n;\lambda^2 z/4)][e^{-\frac{1}{2}z} z^{\frac{1}{2}n-1}/2^{\frac{1}{2}n}\,\Gamma(\tfrac{1}{2}n)] .$$

Die faktor in die tweede stel vierkanthakies is die sentrale χ^2-verdeling. Die oplossing van (19) kan ook gevind word deur substitusie van

$$\phi(z) = e^{-\frac{1}{2}z} z^{\frac{1}{2}n-1}\psi(z)$$

waardeur (19) herlei word tot die Besselvergelyking (11).

Dit is dikwels nuttig om die differensiaalvergelyking vir 'n nie-sentrale verdeling $\phi(z)$ te reduseer deur die transformasie van die afhanklike veranderlike $\phi(z)$, naamlik,

$\phi(z)$ = (bekende sentrale waarskynlikheidsfunksie) $\psi(z)$.

(ii) Die nie-nulverdeling van die korrelasiekoëffisiënt:

In hierdie geval kan soortgelyk aan (i) 'n differensiaalvergelyking in die waarskynlikheidsfunksie afgelei word. Dit is egter vir die huidige doel voldoende om uit te gaan vanuit die bekende waarskynlikheidsverdeling (Muirhead[12] bl. 152) naamlik:

$$\left[\frac{\Gamma(n)\Gamma(n-1)}{\Gamma(n+\frac{1}{2})(2\pi)^{\frac{1}{2}}}(1-r^2)^{\frac{1}{2}\overline{n-3}}\right]\left[(1-\varrho^2)^{\frac{1}{2}n}(1-\varrho r)^{-n+\frac{1}{2}}\, {}_2F_1(\tfrac{1}{2},\tfrac{1}{2};n+\tfrac{1}{2};\tfrac{1}{2}(1+\varrho r))\right]$$

Die faktor binne die eerste stel vierkanthakies toon die waarskynlikheidsfunksie vir die nulverdeling. Die faktor binne die tweede stel vierkanthakies is 'n oplossing van die hipergeometriese vergelyking (6) met $c-a = \frac{1}{2}$, $c-b = \frac{1}{2}$, $c = n+\frac{1}{2}$, $z = \frac{1}{2}(1+\varrho r)$. Dit is gevolglik duidelik dat die waarskynlikheidsfunksie van nie-nulverdeling 'n homogene lineêre differensiaalvergelyking van die tweede-orde bevredig en dat hierdie vergelyking na transformasie van die afhanklike veranderlike, soos aangetoon in (i) hierbo, herlei na die hipergeometriese vergelyking (6).

(iii) Die nie-sentrale F-verdeling en die nie-sentrale Beta-verdeling van die eerste en tweede soort.

In elk van hierdie bostaande gevalle kan, òf uitgaande van eerste beginsels soos in (i), òf uitgaande van die bekende nie-sentrale verdeling, aangetoon word dat die betrokke waarskynlikheidsfunksies oplossings is van die konfluente hipergeometriese vergelyking.

4.2 Meerveranderlike kontinue verdelinge

4.2.1 *Parsiële differensiaalvergelykings van die eerste orde:*

'n Algemene tipe van kontinue verdelings met lineêre regressievergelykings, wat 'n logiese uitbreiding is van die eenveranderlike stelsel van Pearson, soos gedefinieer in (1), word verkry as oplossings van die eerste order vergelykings (Steyn[19]).

$$(b_{ioo} + \sum_{j=1}^{k} b_{ijo}\, x_j + \sum_{j,\ell=1}^{k} b_{ij\ell}x_j x_\ell)\frac{\partial f}{\partial x_i} = (a_{io} + \sum_{j=1}^{k} a_{ij}x_j)f ,$$

met $i = 1,2,\ldots, k$ en $b_{ij\ell} = b_{i\ell j}$.

Hierdie stelsel bevat, onder andere, die meerveranderlike normaalverdeling, 'n meerveranderlike t-verdeling (kyk Pearson,[15] Narumi[13]) en die Dirichletverdeling.

Die welbekende verdeling van Wishart (dit is, die verdeling van die elemente van die variansie-kovariansiematriks) word nie bevat in die bovermelde stelsel nie, maar is 'n spesiale geval van 'n veralgemening van die Pearsonstelsel in terme van matriksveranderlikes. In hierdie laasgenoemde stelsel is die waarskynlikheidsfunksie van die verdeling van Wishart 'n oplossing van die lineêre vergelyking (Steyn[17]):

$$\Sigma^{-1}A\Psi + A\frac{\partial^*}{\partial a}\Psi = (n-k-2)I\Psi \tag{20}$$

met $\frac{\partial^*}{\partial a} = \begin{bmatrix} 2\frac{\partial}{\partial a_{11}} & \frac{\partial}{\partial a_{12}} & \cdots \\ \frac{\partial}{\partial a_{12}} & 2\frac{\partial}{\partial a_{22}} & \cdots \\ \vdots & \vdots & \end{bmatrix}$,

waar A die matriks van som van kwadrate en produkte is soos bereken uit 'n steekproef van grootte n uit 'n meerdimensionele normaalpopulasie met var.-kov. matriks Σ en gemiddelde vektor nul. (Vir notasie kyk Muirhead[12]).

Vergelyking (20) het betrekking op die sogenaamde sentrale verdeling van A. Vir die bestudering van, onder andere, die nie-sentrale verdeling van A is dit vooraf nodig om enkele parsiële differensiaalvergelykings van die tweede orde in te voer.

4.2.2 *Parsiële differensiaalvergelykings van hipergeometriese funksies met matriksargumente*

Die tweede-orde differensiaalvergelykings wat tot dusver in die kontinue meerveranderlike analise voorkom, is almal uitbreidings van die hipergeometriese vergelyking (2) of (6) asook van (10), en dit is gevolglik tans nie nodig om na die meer algemene Riemann-vorm uit te brei nie. Dit is nodig om vooraf 'n hipergeometriese funksie met matriksargument Y te definieer (kyk Muirhead[12] bl. 258).

$${}_2F_1(a,b;c;Y) = \sum_{k=0}^{\infty}\sum_{\kappa}\frac{(a)_\kappa(b)_\kappa}{(c)_\kappa}\frac{C_\kappa(Y)}{k!} \tag{21}$$

waar $C_\kappa(Y)$ simmetriese homogene polinome (sonale polinome) van graad k in die karakteristieke wortels $y_1,y_2,\ldots,y_m$ van Y is en κ is partisie $\kappa=(k_1,k_2\ldots)$ van k is, terwyl

$$(a)_\kappa = \prod_{i=1}^{m}(a-\tfrac{1}{2}\overline{i-1})^{[k_i]}, \quad (a)_0 = 1 .$$

'n Stel parsiële differensiaalvergelykings wat sodanig is dat ${}_2F_1(\cdot)$ soos gedefinieer in (21) 'n oplossing van elk van die vergelykings is, is bekend in die literatuur (Muirhead[12] bl. 271), maar om die verband met die hipergeometriese vergelyking (6) aan te dui is dit insiggewend om eers vergelyking (6) te skryf in terme van die volgende operatore:

$$\Delta = x^2\frac{d^2}{dx^2}, \quad \delta = x\frac{d^2}{dx^2}, \quad E = x\frac{d}{dx} \text{ en } \varepsilon = \frac{d}{dx},$$

sodat (6) dan gegee word as:

$$\delta F - \Delta F + c\varepsilon F - (a+b+1)EF = ab\,F. \tag{22}$$

Vir die geval van die hipergeometriese funksie (21) word nou die volgende operatore gedefinieer:

$$\Delta_Y = \sum_{i=1}^{m} y_i^2\frac{\partial^2}{\partial y_i^2} + \sum_{i=1}^{m}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{y_i^2}{y_i-y_j}\frac{\partial}{\partial y_i}$$

$$\partial_Y = \sum_{i=1}^{m} y_i\frac{\partial^2}{\partial y_i^2} + \sum_{i=1}^{m}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{y_i}{y_i-y_j}\frac{\partial}{\partial y_i}$$

$$E_Y = \sum_{i=1}^{m} y_i\frac{\partial}{\partial y_i}, \quad \varepsilon_Y = \sum_{i=1}^{m}\frac{\partial}{\partial y_i}.$$

Dit is nou bekend dat $F \equiv {}_2F_1(a,b;c;Y)$ die volgende parsiële differensiaalvergelyking bevredig (Muirhead[11,12])

$$\delta_Y F - \Delta_Y F + [c-\tfrac{1}{2}(m-1)]\varepsilon_Y F - [a+b+1-\tfrac{1}{2}(m-1)]E_Y F = mabF. \tag{23}$$

Dit is duidelik dat vir m = 1 vergelyking (23) dieselfde is as (22). Die vergelyking (23) kan ook geskryf word as:

$$y_i(1-y_i)\frac{\partial^2 F}{\partial y_i^2} + \Big\{c-\tfrac{1}{2}(m-1)-[a+b+1-\tfrac{1}{2}(m-1)]y_i + \tfrac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{y_i(1-y_i)}{y_i-y_j}\Big\}\frac{\partial F}{\partial y_i} - \tfrac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{y_j(1-y_j)}{y_i-y_j}\frac{\partial F}{\partial y_j} = ab\,F \tag{24}$$

$(i=1,2.\ \ldots;m)$, waar $y_1,y_2,\ldots,y_m$ die karakteristieke wortels is van Y.

Ooreenkomstig is dit ook bekend (Muirhead[12] bl. 277) dat ${}_1F_1(a;c;Y)$ en ${}_0F_1(c;Y)$ wat, soos in afdeling 2, soortgelyk aan ${}_2F_1(\cdot)$ gedefinieer word, oplossings is van die volgende parsiële differensiaalvergelykings (25) en (26) respektiewelik:

$$y_i\frac{\partial^2 F}{\partial y_i^2} + \Big\{c-\tfrac{1}{2}(m-1)-y_i+\tfrac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{y_i}{y_i-y_j}\Big\}\frac{\partial F}{\partial y_i} - \tfrac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{y_j}{y_i-y_j}\frac{\partial F}{\partial y_j} = aF \tag{25}$$

en

$$y_i\frac{\partial^2 F}{\partial y_i^2} + \Big\{c-\tfrac{1}{2}(m-1)+\tfrac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{y_i}{y_i-y_j}\Big\}\frac{\partial F}{\partial y_i} - \tfrac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{y_j}{y_i-y_j}\frac{\partial F}{\partial y_j} = F \tag{26}$$

4.2.3 *Enkele nie-sentrale meerveranderlike verdelings*

(i) Indien die r-de steekproef vektor 'n waarde is wat verkry is uit 'n m-dimensionale normale populasie $N(h_r,\Sigma)$ waar $h_r = [h_{1r},h_{2r},\ldots,h_{mr}]$, dan word die waarskynlikheidsfunksie van A, die matriks van die som van die steekproefkwadrate en -produkte, gegee deur die bekende nie-sentrale Wishart-waarskynlikheidsfunksie (Muirhead,[12] bl. 442)

$$\left[\frac{1}{2^{\frac{1}{2}mn}\Gamma_m(\tfrac{1}{2}n)|\Sigma|^{\frac{1}{2}n}}(\mathrm{etr}(-\tfrac{1}{2}\Sigma^{-1}A))|A|^{\frac{1}{2}\overline{n-m-1}}\right]\left[\mathrm{etr}(-\tfrac{1}{2}\Omega)\,{}_0F_1(\tfrac{1}{2}n;\tfrac{1}{4}\Omega\Sigma^{-1}A)\right] \tag{27}$$

waar $\Omega = \Sigma^{-1}H'H$, $h = [h_{ir}]$ en $\Gamma_m(\cdot)$ die meerveranderlike gammafunksie voorstel. Uit die bostaande skryfwyse in (27) is dit duidelik dat (27) uit die volgende twee faktore bestaan: Die faktor binne die eerste stel vierkanthakies is die bekende waarskynlikheidsfunksie van die sentrale verdeling van A (Wishart se verdeling), terwyl die faktor binne die tweede stel vierkanthakies die nie-sentrale faktor is. Dit is verder duidelik dat (27) die oplossing is van 'n parsiële lineêre differensiaalvergelyking wat na transformasie van die afhanklike matriksveranderlike deur gebruikmaking van die bekende sentrale verdeling (kyk afdeling 4.1) reduseer tot vergelyking (26).

(ii) Enkele verdere waarskynlikheidsfunksies wat geassosieer is met die bogenoemde parsiële hipergeometriese vergelykings is die volgende:

(a) Die waarskynlikheidsfunksie van die matriks (Muirhead[12] bl. 449, de Waal[1])

$$F = A^{1/2} B^{-1} A^{1/2}$$

waar A en B onafhanklik is en A 'n $W_m(r,I,\Omega)$ en B 'n $W_m(n-p,I)$, $r \geq m$, $n-p \geq m$ as verdelings besit, is 'n oplossing van vergelyking (25).

(b) Die gesamentlike waarskynlikheidsfunksie van die karakteristieke wortels van AB^{-1} waar A 'n niesentrale Wishart- en B 'n sentrale Wishartverdeling besit (Muirhead,[12] bl. 450).

5. ENKELE ALGEMENE OPMERKINGE

In die voorafgaande afdelings is 'n poging aangewend om aan te toon hoedat die teoretiese waarskynlikheidsverdelings van die Wiskundige Statistiek in die eenveranderlike geval geassosieer is met die oplossings van 'n tweede-orde lineêre differensiaalvergelyking met drie reguliere singulariteite (Reimann se vergelyking) en in die meerveranderlike geval met logiese meerdimensionale uitbreidings van so 'n vergelyking. Vir kontinue waarskynlikheidsverdelings kan die tweede-orde vergelyking beskou word as 'n uitbreiding van die bekende eerste-orde vergelyking van Pearson. Vir die diskrete verdelings is die wvf die oplossing van die betrokke vergelyking. Hieruit volg, onder andere, dat:

(i) die transendente funksies van die suiwer wiskunde (dit wil sê, die gamma-, die hipergeometriese, die Bessel- en ander funksies) noodwendig 'n fundamentele rol moet speel by die bestudering van waarskynlikheidsverdelings;

(ii) daar vir 'n wye klas van kontinue verdelings normaalweg 'n redelik eenvoudige verband bestaan tussen die waarskynlikheidsfunksie en sy eerste en tweede orde afgeleides;

(iii) vir 'n wye klas van diskrete verdelings die wvf sodanig is dat daar 'n redelik eenvoudige verband bestaan tussen die wvf en sy eerste en tweede afgeleides, wat heenwys na 'n redelik eenvoudige rekursieformule tussen die momente weens die feit dat die vergelyking vir die wvf deur die transformasie $t = e^{\alpha}$ omvorm kan word tot 'n tweede-orde differensiaalvergelyking in die momentvoortbringende funksie $M(\alpha)$;

(iv) Riemann se vergelyking 'n veel wyer klas van verdelings bevat as wat in hierdie uiteensetting hanteer is.

LITERATUURVERWYSINGS

1. De Waal, D.J. (1972). An asymptotic distribution of noncentral multivariate Dirichlet variates, *S.A. Stat. Journal*, 6(1), 31-40.
2. Erdėlyi, A. (ed.) (1953). *Higher Transcendental Functions*, Vol. I (McGraw-Hill).
3. Gurland, J. & Trimpathi, R. (1974). Estimation of parameters on some extensions of the Katz family of discrete distributions involving hypergeometric functions. In *A Modern Course on Statistical Distributions*. Vol. I, Patil, G.P.; Kotz, S & Ord, J.K. (eds.)
4. Janardan, K.G. (1973). Chance mechanisms for multivariate hypergeometric models, *Sankhya, Series A*, Vol. 35, 465-478.
5. Janardan, K.G. & Patil, G.P. (1972). A unified approach for a class of multivariate hypergeometric models, *Sankhya, Series A*, Vol. 34, 363-376.
6. Johnson, N.L. and Kotz, S. (1969). *Discrete Distributions* (Houghton Mifflin, Boston).
7. Katz, L. (1946). On the class of functions defined by the difference equation $(x+1)f(x+1) = (a+bx)f(x)$ (Abstract), *Annals of Math, Stat.*, Vol. 17, 501.
8. Kemp, A.W. & Kemp, C.D. (1974). A family of discrete distributions defined via their factorial moments, *Comm Statist*, Vol. 3, 1187-1196.
9. Mosimann, J.E. (1962). On the compound multinomial distribution, the multivariate β-distribution and correlations among proportions, *Biometrika*, Vol. 49, 65-82.
10. Mosimann, J.E. (1963). On the compound negative multinomial distribution and correlations among inversely sampled pollen counts, *Biometrika*, Vol. 50, 47-54.
11. Muirhead, R.J. (1970). Partial differential equations for hypergeometric functions of matrix argument, *Annals of Math. Stat.*, Vol. 41, 991-1001.
12. Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory* (John Wiley, New York).
13. Narumi, S. (1923). On the general forms of bivariate frequency distributions which are mathematical possible when regression and variation are subjected to limiting conditions, *Biometrika*, Vol. XV, 209-221.
14. Ord, J.K. (1967). On a system of discrete distributions, *Biometrika*, Vol. 54, 649-656.
15. Pearson, Karl (1923). On non-skew frequency surfaces, *Biometrika*, Vol. XV, 231-244.
16. Steyn, H.S. (1951). On discrete multivariate probability functions, *Koningkl. Nederl. Akademie van Wetenschappen, Series A*, Vol. 54, No. 1, 23-30.
17. Steyn, H.S. (1951). The Wishart distribution derived by solving simultaneous linear differential equations, *Biometrika*, Vol. 38, 470-472.
18. Steyn, H.S. (1968). Eienskappe van waarskynlikheidsverdelings deur gebruik van differensiaalvergelykinge, *S.A. Stat. Journal*, Vol. 2, No. 1, 1-8.
19. Steyn, H.S. (1960). On regression properties of Pearson's types, *Koninkl. Nederl. Akademie van Wetenschappen, Series A*, Vol. 63, No. 3, 302-311.
20. Whittaker, E.T. & Watson, G.N. (1920). *A Course of Modern Analysis Third ed.* (Cambridge Univ. Press).