Pertanika J. Sci. & Technol. 2(2): 165-173 (1994)

ISSN: 0128-7680

Penggunaan Pendekatan Sistem-S dan ESSYNS dalam Analisis Taburan Normal

M. H. Lee and B. R. Ahmad Mahir

Jabatan Statistik
Fakulti Sains Matematik & Komputer,
Universiti Kebangsaan Malaysia,
43600 UKM Bangi,
Selangor Darul Ehsan, Malaysia.

Received 9 July 1993

ABSTRAK

Makalah ini bertujuan menunjukkan penggunaan kaedah sistem-S dan perisian ESSYNS dalam bidang statistik penghitungan. Pertama, kami akan tunjukkan bagaimana ketumpatan dirumus sebagai sistem-S dan seterusnya bagaimana perwakilan bagi longgokan dan kuantil dapat diterbitkan. Longgokan diperolehi hanya dengan sedikit usaha tambahan dan kuantil dapat dihitung daripada songsangan sistem-S. Perwakilan ini selanjutnya diselesaikan secara berangka dengan menggunakan ESSYNS. Lengkung ketumpatan dan longgokan bagi taburan normal dijana untuk nilai dalam julat (-3,3). Kaedah sistem-S menghasilkan banyak kuantil dalam satu larian komputer. Penggunaan pengitlakan dan penghampiran sistem-S menunjukkan kaedah alternatif ini berguna untuk menganalisis taburan.

ABSTRACT

The purpose of this paper is to illustrate the use of S-systems methodology and ESSYNS software in the area of computational statistics. First, we demonstrate how the densities are formulated as S-systems, then we derive representations for cumulatives and quantiles. Cumulatives are obtained with little extra effort and quantiles are readily calculated from the inverted S-system. These representations are then solved numerically using ESSYNS. The density and the cumulative curves for the normal distribution are generated for the values in the range (-3,3). The S-system method produces many quantiles in a single computer run. The use of S-system generalizations and approximation has shown to be a useful alternative method for analysing distributions.

Katakunci: taburan, sistem-S, persamaan pembeza, perhitungan berstatistik

PENGENALAN

Apabila analisis atau penilaian sesuatu taburan hendak dikaji lazimnya seseorang penyelidik akan terfikir untuk menggunakan pendekatan penakrifan statistik. Sebagai contoh, apabila taburan F hendak dikaji, nisbah dua taburan khi kuasa dua akan digunakan. Pilihan lain yang mungkin adalah perumusan dalam sebutan fungsi-fungsi siri atau fungsi-fungsi khas. Walau bagaimanapun, perwakilan matematik yang padat tidak selalunya sesuai untuk tujuan penghitungan dan adalah lebih berkesan apabila berasaskan perwakilan yang sesuai untuk penilaian berangka walaupun perwakilan tersebut tidak begitu menarik secara matematik.

Sebagai contoh, perhatikan perwakilan yang berlainan bagi fungsi polinomial yang tertentu. Lazimnya, takrifan dalam sebutan kuasa, iaitu $f(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3$ akan digunakan. Kos penghitungan bagi kuasa adalah tinggi dan kebanyakan buku pengaturcaraan mencadangkan kepada kita untuk menggunakan konsep hasil darab, iaitu $f(t) = \alpha_0 + \alpha_1 \times t + \alpha_2 \times t \times t + \alpha_3 \times t \times t \times t$. Perwakilan ini lebih berkesan berbanding yang pertama tetapi boleh dipertingkatkan secara bererti dengan perwakilan $f(t) = \alpha_0 + (\alpha_1 + (\alpha_2 + \alpha_3 \times t) \times t) \times t$ walau pun ini kurang bermakna secara intuisi. Ini menjelaskan bahawa perwakilan yang berlainan bagi masalah yang sama mempunyai kelebihan masing-masing. Pertama, dapat dilihat dengan baik secara intuisi, keduanya sesuai untuk analisis dengan kertas dan pensil dan yang ketiga mempunyai alkhwarizmi yang optimum.

Melalui makalah ini kita perkenalkan perwakilan matematik yang sesuai untuk analisis secara berangka bagi ketumpatan, longgokan dan kuantil taburan normal. Pertama, akan ditunjukkan bagaimana fungsi ketumpatan dapat dirumus semula kepada satu set persamaan terbitan yang khusus dikenali sebagai sistem-S. Seterusnya ditunjukkan terbitan bagi perwakilan fungsi longgokan dan kuantil. Akhir sekali dipaparkan beberapa hasil penyelesaian berangka yang jitu diperolehi melalui program ESSYNS yang diadakan khusus untuk analisis sistem-S.

PENDEKATAN SISTEM-S

Sistem-S pada peringkat awal digunakan untuk analisis bidang yang lain daripada bidang statistik, iaitu untuk analisis sistem-sistem biologi dan sistem lain yang terurus secara kompleks (lihat Savageau 1976; Savageau and Voit 1982; Voit 1990 untuk rujukan lanjut). Lanjutan daripada itu, sistem-S berjaya digunakan dalam pelbagai bidang termasuk biokimia, imunologi, genetik dan perhutanan (lihat Voit 1991). Sistem-S terdiri daripada set persamaan pembeza peringkat pertama tak linear dengan setiap persamaan mempunyai struktur yang homogen, iaitu terbitan setiap pembolehubah bersamaan dengan beza antara dua hasil darab fungsi hukum kuasa. Jika terdapat n pembolehubah bersandar dan t merupakan pembolehubah bebas, maka sistem-S dapat ditulis seperti berikut:

$$\dot{X}_i = \alpha_i \prod_{j=1}^{n} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n} X_j^{h_{ij}}$$

$$i \in \{1,2, \ldots, n\} \qquad (1)$$

dengan $\dot{X}_i = dX_i/dt$; X_i, α_i dan β_i boleh mengambil sebarang nilai nyata tak negatif, dan g_{ij} dan h_{ij} pula boleh mengambil sebarang nilai nyata. Syarat $X_i > 0$ bukan merupakan kekangan yang serius kerana pembolehubah yang negatif boleh dijelmakan kepada positif dengan mudah (Savageau and Voit 1987).

PERUMUSAN SEMULA

Pembolehubah rawak T adalah tertabur secara normal jika dan hanya jika ketumpatan kebarangkalian,

$$n(t; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \tag{2}$$

$$-\infty < t < \infty, \qquad -\infty < \mu < \infty, \qquad \sigma > 0$$

dan T dirujuk sebagai pembolehubah rawak normal. μ dan σ mewakili min dan sisihan piawai taburan normal. Lanjutan daripada ini, fungsi longgokan normal dapat dinyatakan seperti berikut:

$$F_n(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma^2}\right)^2} dx \tag{3}$$

$$-\infty < x < \infty,$$

Terbitan persamaan (2) terhadap t menghasilkan

$$\begin{aligned} \frac{dn(t; \mu, \sigma)}{dt} &= n'(t; \mu, \sigma) \\ &= -\frac{t-\mu}{\sigma^3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \\ &= -\frac{t-\mu}{\sigma^2} n(t; \mu, \sigma) \end{aligned} \tag{4}$$

Rumus fungsi ketumpatan kebarangkalian bagi taburan normal ini dapat dibentuk semula sebagai sistem-S dengan menakrifkan dua pembolehubah baru, iaitu $X_1 = t/\sigma^2 + c$ dan $X_2 = n(t; \mu, \sigma)$. Terbitan setiap pembolehubah ini terhadap t akan menghasilkan perwakilan dalam sistem-S dengan $c' = c + \mu/\sigma^2$

$$\begin{aligned} \dot{X}_1 &= \frac{1}{\sigma^2} \\ \dot{X}_2 &= (c' - X_1)X_2 = c'X_2 - X_1X_2 \end{aligned} \tag{5}$$

Penambahan pemalar $c > 0$ pada X_1 bertujuan memastikan semua pembolehubah bersandar sentiasa positif. Ubah suai ini tidak mempengaruhi persamaan $\dot{X}_1$ tetapi mempengaruhi nilai awal X_1 dan persamaan $\dot{X}_2$. Jika $t_0 = 0$, maka syarat awal kedua-dua pembolehubah bersandar adalah $X_{1,0} = c$ dan $X_{2,0} = (2\pi\sigma^2)^{-1/2}\,\text{eksp}(-\mu^2/2\sigma^2)$. Apabila sistem ini ingin diselesaikan untuk t yang meningkat ke arah positif, pemalar c boleh mengambil sebarang nilai, misalnya $c = 1$ adalah memadai. Jika t ingin dilaksanakan pada arah negatif, maka nilai c yang agak besar perlu dipilih.

Tatacara perumusan semula ini adalah tidak bitara dan perumusan yang lain boleh dicuba untuk memenuhi kesesuaian yang dikehendaki. Misalnya, seseorang penyelidik boleh membuat dua takrifan baru, iaitu $X_1 = n(t;\mu,\sigma)$ dan $X_2 = (t-\mu)/\sigma$ untuk menghasilkan sistem-S yang baru (Voit 1991).

Lanjutan daripada ini, taburan longgokan normal berhubung dengan (2) mudah diperolehi kerana terbitan fungsi longgokan bersamaan fungsi ketumpatan taburan yang berkaitan. Oleh itu, ketumpatan dan longgokan normal dapat dirumus secara serentak seperti berikut:

$$
\begin{aligned}
X_1 &= \frac{t}{\sigma^2} + c & \dot{X}_1 &= \frac{1}{\sigma^2} \\
X_2 &= n(t;\mu,\sigma) & \dot{X}_2 &= c'X_2 - X_1X_2 \\
X_3 &= F_n(t) & \dot{X}_3 &= X_2
\end{aligned}
\tag{6}
$$

dengan X_3 merupakan fungsi longgokan bagi taburan normal. Apabila t_0 bersamaan μ, maka syarat awal bagi X_3 bersamaan 0.5 diperolehi. Jika X_3 ditakrif sebagai $1 - F_n(t)$, maka nilai kebarangkalian hujung atas yang berhubung dengan rantau penolakan akan diperolehi.

Sebagai contoh, perwakilan sistem-S bagi fungsi ketumpatan dan fungsi longgokan taburan normal piawai ($\mu = 0$, $\sigma = 1$) adalah

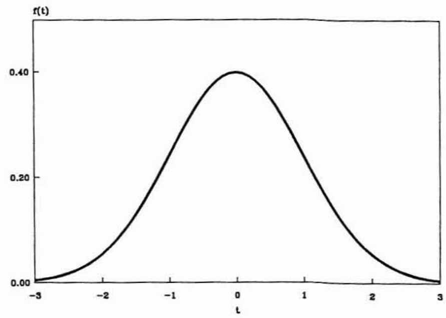
$$
\begin{aligned}
X_1 &= t & \dot{X}_1 &= 1 \\
X_2 &= n(t;0,1) & \dot{X}_2 &= -X_1X_2 \\
X_3 &= F_z(t) & \dot{X}_3 &= X_2
\end{aligned}
\tag{7}
$$

untuk $t > 0$. Jika fraktil $t < 0$ yang diminati, maka $X_1 = -t$ ditakrifkan sebagai gantian bagi X_1 pada sistem persamaan (7) dan ini bertujuan memastikan kepositifan pembolehubah. Oleh itu, perwakilan sistem-S untuk takrifan X_2 dan X_3 yang sama pada julat nilai t yang negatif adalah seperti berikut:

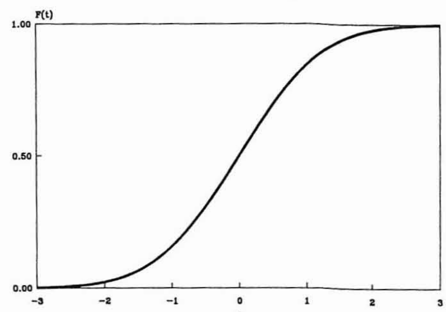
$$
\begin{aligned}
\dot{X}_1 &= -1 & X_1(t_0) &= t_0 \\
\dot{X}_2 &= X_1X_2 & X_2(t_0) &= (2\pi)^{-1/2}\,\text{eksp}(-t_0/2) \\
\dot{X}_3 &= X_2 & X_3(t_0) &= F_z(t_0)
\end{aligned}
\qquad (8)
$$

Apabila perwakilan sistem-S bagi fungsi taburan berjaya diadakan, penyelesaian berangka boleh diperolehi daripada sebarang pengamir berangka. Secara khususnya, ESSYNS (Voit *et al.* 1990) dipilih kerana perisian ini diadakan khas untuk analisis sistem-S. Ia berasaskan kepada penjelmaan logaritma pembolehubah X_i dan menggunakan kaedah Taylor yang diubah-suai. Alkhwarizmi ini berupaya memberikan kejituan penyelesaian yang tinggi dengan dua atau tiga kali ganda lebih cepat daripada kaedah biasa (Irvine and Savageau 1990).

Sehubungan ini, *Rajah 1* menunjukkan contoh penyelesaian sistem persamaan (7) dan (8) dengan ESSYNS. Nilai bagi parameter-parameter $\alpha_1, \alpha_3, \beta_2$ dan g_{32}, h_{21}, h_{22} adalah bersamaan satu dan selainnya bersamaan sifar untuk $0 < t < \infty$. Sebaliknya untuk kasus $-\infty < t < 0, \alpha_2, \alpha_3, \beta_1$, dan g_{21}, g_{22}, g_{32} bernilai satu dan yang selainnya sifar. Daripada *Rajah 1* dapat diperhatikan bahawa sebahagian besar lengkung fungsi ketumpatan dan fungsi longgokan dapat dijana dengan baik dan mudah oleh ESSYNS.



Fungsi ketumpatan kebarangkalian normal piawai $n(0,1)$



Fungsi taburan normal piawai $n(0,1)$

Rajah 1. Ketumpatan dan longgokan bagi taburan normal piawai

SONGSANGAN SISTEM-S

Lanjutan daripada hasil keputusan yang langsung, pendekatan sistem-S membenarkan penghitungan ciri-ciri taburan kebarangkalian yang sukar diperolehi dengan cara biasa. Pada bahagian ini akan ditunjukkan bagaimana kuantil dapat dihitung melalui songsangan sistem-S. Apabila taburan normal dibahagikan kepada empat bahagian yang sama maka setiap bahagian ini dikenali sebagai kuantil. Nilai tengahnya dikenali sebagai median dan terdapat dua kuantil di kiri dan kanan median ini. Songsangan sistem-S bermakna menggantikan pembolehubah bebas T dengan pembolehubah yang ditakrif sebagai fungsi longgokan dan setiap pembolehubah sistem-S diungkapkan sebagai fungsi longgokan. Apabila sistem persamaan (7) dibahagikan dengan persamaan terbitan yang terakhir, kuantil bagi taburan normal piawai diperolehi melalui pembolehubah X_1 songsangan. Songsangan sistem-S berhubung dengan kuantil dengan X_3 sebagai pembolehubah bebas dapat dinyatakan seperti berikut:

$$
\begin{aligned}
\frac{dX_1}{dX_3} &= \overset{\circ}{X}_1 = X_2^{-1} \\
\frac{dX_2}{dX_3} &= \overset{\circ}{X}_2 = -X_1
\end{aligned}
\tag{9}
$$

Jika seseorang penyelidik berminat untuk mendapatkan kuantil hujung atas, maka beliau hanya perlu menukar X_3 yang lama kepada $X_3 = 1 - F_z(t)$. Pembolehubah X_1 ditambah dengan pemalar c (katakan c = 1) untuk memastikan pemboleh ubah sistem persamaan sentiasa positif seperti untuk perwakilan taburan normal. Perwakilan sistem-S bagi kuantil hujung atas adalah seperti berikut:

$$
\begin{aligned}
&\dot{X}_1 = 1 & & \\
&\dot{X}_2 = X_2 - X_1X_2 & \qquad &\overset{\circ}{X}_1 = -X_2^{-1} \\
&\dot{X}_3 = -X_2 & \qquad &\overset{\circ}{X}_2 = X_1 - 1
\end{aligned}
\tag{10}
$$

Apabila pembolehubah bebas X_3 ditingkatkan daripada $F_z(t_0)$ ke kuantil K yang diminati, pembolehubah $X_1(K)$ menghasilkan titik-titik penyelesaian termasuk yang ke K dengan bergantung kepada saiz peningkatan, $\pm\Delta t$ yang digunakan. Nilai awal pada $t_0 = 0$ bagi $X_{2,0}$ ialah $(2\pi)^{-1/2}$ tetapi nilai awal bagi pembolehubah bebas X_1 bersamaan 0.5 dan tidak lagi mengikuti nilai awal t. Misalkan, nilai akhir yang diminati adalah 10^{-6} pada saiz peningkatan yang berlainan (lihat Jadual 1) maka ini dapat diubah dengan menggunakan arahan *Extend.* Jadual 1 menyenaraikan hasil keputusan bagi kuantil taburan normal piawai antara nilai 0.5 hingga 10^{-6}. Kejituan penyelesaian dikawal oleh toleransi ralat setempat maksimum yang boleh dipilih antara nilai 10^{-3} hingga 10^{-16}. Toleransi yang digunakan adalah 10^{-12}. Titik penyelesaian yang terjana sehubungan dengan pembolehubah X_1 adalah untuk t + 1. Oleh itu, untuk mendapatkan statistik z berhubung dengan kuantil adalah perlu minus 1 daripada semua titik penyelesaian bagi X_1. Keputusan yang dihasilkan dapat dilihat dalam Jadual 1. ESSYNS tidak memperlihatkan masalah malah memberikan kejituan yang tinggi bagi variat yang mempunyai beberapa sifar di hadapan.

Hasil keputusan yang diperolehi bersamaan dengan sifir statistik (Pearson and Hardley 1976; Murdoch and Barnes 1985) pada kejituan empat titik perpuluhan. Penyelesaian melalui sistem-S bukan sahaja dapat memberi kesemua nilai seperti dalam jadual statistik tetapi juga nilai yang wujud antara mereka pada kejituan yang tinggi. Ini merupakan salah satu kelebihan apabila menggunakan kaedah penyelesaian sistem-S. Sebaliknya jika seseorang penyelidik ingin mendapatkan nilai yang tidak terdapat dalam sifir statistik yang biasa, lazimnya pendekatan interpolasi linear digunakan dan kejituan yang akan diperolehi adalah kurang memuaskan jika kejituan tinggi dikehendaki.

JADUAL 1
Persentil bagi taburan normal (0.5 hingga 0.000001)

α	**kuantil**	α	**kuantil**	α	**kuantil**	α	**kuantil**	α	**kuantil**
0.50	0.00000000	0.050	1.64485363	0.020	2.05374891	0.0010	3.09023231	0.00010	3.71901648
0.45	0.12566135	0.048	1.66456286	0.019	2.07485473	0.0009	3.12138915	0.00009	3.74554859
0.40	0.25334710	0.046	1.68494077	0.018	2.09692743	0.0008	3.15590676	0.00008	3.77501194
0.35	0.38532047	0.044	1.70604340	0.017	2.12007169	0.0007	3.19465105	0.00007	3.80816826
0.30	0.52440051	0.042	1.72793432	0.016	2.14441062	0.0006	3.23888012	0.00006	3.84612614
0.25	0.67448975	0.040	1.75068607	0.015	2.17009038	0.0005	3.29052673	0.00005	3.89059188
0.20	0.84162123	0.038	1.77438191	0.014	2.19728638	0.0004	3.35279478	0.00004	3.94440008
0.15	1.03643339	0.036	1.79911811	0.013	2.22621177	0.0003	3.43161440	0.00003	4.01281081
0.10	1.28155157	0.034	1.82500682	0.012	2.25712924	0.0002	3.54008380	0.00002	4.10747965
0.05	1.64485363	0.032	1.85217986	0.011	2.29036788	0.0001	3.71901648	0.00001	4.26489079
		0.031	1.86629574	0.010	2.32634787			0.000009	4.28835653
		0.030	1.88079361	0.009	2.36561813			0.000008	4.31445101
		0.029	1.89569792	0.008	2.40891555			0.000007	4.34386117
		0.028	1.91103565	0.007	2.45726339			0.000006	4.37758784
		0.027	1.92683657	0.006	2.51214433			0.000005	4.41717340
		0.026	1.94313375	0.005	2.57582930			0.000004	4.46518390
		0.025	1.95996398	0.004	2.65206981			0.000003	4.52638930
		0.024	1.97736843	0.003	2.74778139			0.000002	4.61138233
		0.023	1.99539331	0.002	2.87816174			0.000001	4.75342425
		0.022	2.01409081	0.001	3.09023231				
		0.021	2.03352015						

KESIMPULAN

Perumusan semula taburan normal sebagai sistem-S tidak memperlihat kelebihan pada peringkat awal, iaitu penggantian rumus asal yang padat oleh satu set persamaan terbitan yang agak besar, tak linear dan yang tidak mempunyai penyelesaian analisis yang mudah. Walau bagaimanapun, penilaian bagi taburan memerlukan kaedah berangka tidak kira pendekatan mana yang dipilih. Perwakilan dalam sistem-S menawarkan penghitungan yang cepat dan yang boleh diyakini, khususnya apabila taburan yang ingin dinilai adalah pada keseluruhan julat pembolehubah bebas dan tidak hanya pada beberapa titik penyelesaian sahaja. Perumusan semula taburan statistik yang lain dapat dilihat pada makalah-makalah seperti Savageau (1982); Voit (1992); Lee dan Ahmad Mahir (1993). Savageau and Voit (1987) pula berjaya menunjukkan hampir kesemua persamaan terbitan dan fungsi-fungsi khas dapat diwakilkan dengan baik sebagai submodel sistem-S.

Apabila perwakilan sistem-S bagi taburan normal berjaya dibentuk, perisian ESSYNS menawarkan keputusan penyelesaian yang jitu. Setiap pemboleh ubah hanya memerlukan satu nilai awal sahaja untuk tujuan penilaian yang dipertimbangkan dan sebahagian besar lengkung boleh diperolehi cuma dalam beberapa saat sahaja. Lanjutan daripada itu, songsangan sistem-S menawarkan penilaian kuantil dengan hanya sedikit pengubahsuaian pada sistem persamaan taburan statistik yang terbentuk untuk berbagai nilai parameter yang diminati. Penyelesaian yang diperolehi adalah jitu sehingga sepuluh titik perpuluhan atau lebih.

RUJUKAN

IRVINE, D.H. and M.A. SAVAGEAU. 1990. Efficient solution of nonlinear ordinary differential equations expressed in S-system canonical form. *SIAM J. Num. Anal.* **27(3)**: 704-735.

LEE, M.H. dan B.R. AHMAD MAHIR. 1993. Pembentukan semula fungsi taburan ke dalam bentuk sistem-S kanonik. *Pascasidang Seminar Siswazah*: pp. 41-54.

MURDOCH, J. and J.A. BARNES. 1985. *Statistical Tables for Science, Engineering, Management and Business Studies.* Macmillan.

PEARSON, E.S. and H.O. HARDLEY. 1976. *Biometrika Tables for Statisticians Vol. 2.* Cambridge: University Printing House.

SAVAGEAU, M.A. 1976. *Biochemical System Analysis: A Study of Function and Design in Molecular Biology.* Reading, Mass.: Addison-Wesley.

SAVAGEAU, M.A. 1982. A suprasystem of probability distribution. *Biom. J.* **24**: 323-330.

SAVAGEAU, M.A. and E.O. VOIT. 1982. Power-law approach to modeling biological systems: I. Theory. *J. Ferment. Technol.* **60(3)**: 221-228.

SAVAGEAU, M.A. and E.O. VOIT. 1987. Recasting nonlinear differential equations as S-systems: A canonical nonlinear form. *Mathem. Biosci.* **87**: 83-115.

VOIT, E.O. 1990. S-system analysis of endemic infection. *Compt. Math. Appl.* **20**(**4-6**): 161-173.

VOIT, E.O. 1991. *Canonical Nonlinear Modeling: S-System Approach to Understanding Complexity.* New York: Van Nostrand Reinhold.

VOIT, E.O. 1992. The S-distribution: A tool for approximation and classification of univariate, unimodal probability distributions. *Biom. J.* **34**(**7**): 855-878.

VOIT, E.O., D.H. IRVINE and M.A. SAVAGEAU. 1990. *The User's Guide to ESSYNS.* Charleston, S.C.: Medical University of South Carolina Press.