

Pengujian Kesesuaian Taburan Normal Berdasarkan Statistik Cramer-Von Mises (Test of Suitability of Normal Distribution Based on Cramer-Von Mises Statistics)

ANI BIN SHABRI & ABDUL AZIZ JEMAIN

ABSTRAK

Sejak taburan normal ditemui dan ianya merupakan salah satu taburan yang penting dalam statistik, terdapat banyak pengujian statistik yang dibangunkan untuk menguji kenormalan data. Namun begitu masih tidak banyak kajian yang dilakukan untuk melihat kembali keupayaan pengujian statistik yang sedia ada. Sebahagian daripada pengujian statistik didapati mudah tetapi hanya sesuai untuk sesuatu keadaan. Dalam kajian ini, pengujian statistik berdasarkan statistik Cramer-von Mises cuba diperbaiki berdasarkan rumus Weibull. Kekuatan statistik yang baru ini dibandingkan kekuatan dengan statistik tradisional Anderson-Darling (AD), Cramer von-Mises (CR), Kolmogorov-Smirnov (KS) dan Shapiro-Wilk (SW). Kajian simulasi berdasarkan beberapa taburan yang berbeza menunjukkan pengujian statistik yang dicadangkan paling sesuai untuk menguji kenormalan.

Kata kunci: Pengujian Kenormalan; Cramer von-Mises; Kolmogorov-Smirnov; Shapiro-Wilk

ABSTRACT

Since normal distributions are the most important ones in statistics, there are large number of tests for normality. However they have less some drawbacks. Some of these tests are simple but suitable for some situations. In this study, the traditional Cramer-von Mises test statistics is modified based on Weibull formula. The new goodness-of-fit test is compared with the traditional Anderson-Darling (AD), Cramer von-Mises (CR), Kolmogorov-Smirnov (KS) and Shapiro-Wilk (SW) test statistics. A simulation study using several different distributions shows that the proposed test is very powerful for testing normality.

Keywords: Test of normality; Cramer von-Mises; Kolmogorov-Smirnov; Shapiro-Wilk

PENGENALAN

Ujian penyuaiian terbaik taburan secara statistik merupakan tajuk yang mendapat perhatian dalam kajian statistik. Pengujian ini digunakan untuk mengukur darjah kesesuaian antara taburan sampel cerapan dengan taburan teori. Antara taburan yang menjadi tunggak statistik adalah taburan normal, justeru ujian kenormalan perlu dilakukan bagi memastikan pentadbiran yang dibuat berasaskan sifat kenormalan sah. Walaupun banyak pengujian statistik yang berupaya menguji kenormalan data, tetapi masih sedikit kajian yang dilakukan untuk melihat kembali kesesuaian pengujian statistik yang sedia ada. Pengujian statistik tradisional Shapiro-Wilk (SW) merupakan pengujian yang paling sesuai berbanding ujian Anderson Daling (AD), Cramer-von Mises (CR) dan Kolmogorov-Smirnov (KS) dalam menguji kenormalan data (D'Agostino & Stephens 1986; Swanepoel & Graan 2002; Zhang 2004). Bagaimanapun statistik SW hanya sesuai untuk sampel bersaiz kecil, $n \leq 50$. Pengujian statistik yang lain seperti Anderson-Darling didapati kurang sesuai berbanding statistik CR tetapi lebih baik berbanding statistik KS dalam menguji kenormalan data. Kajian ini bertujuan untuk memperbaiki uji traditional CR berdasarkan rumus Weibull

(CRW). Kekuatan dan kesesuaian ujian CRW dalam menguji taburan normal, diuji dan dibandingkan dengan ujian AD, KS, CR dan SW.

STATISTIK CRAMER-VON MISES

Pengujian statistik berdasarkan fungsi taburan empirik (FTE) secara umumnya boleh ditakrifkan sebagai:

$$Q = n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \psi(x) dF(x) \quad (1)$$

dengan $\psi(x)$ adalah fungsi pemberat, $F_n(x)$ adalah fungsi taburan empirik bagi sampel rawak, n adalah bilangan sampel dan $F(x)$ adalah fungsi taburan kumulatif bagi sebarang taburan selanjur yang sepadan (Pavur et al. 1992). Apabila $\psi(x) = 1$, pengujian statistik dalam persamaan (1) dikenali sebagai statistik Cramer-von Mises dan boleh ditulis sebagai:

$$Q = n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 dF(x) \quad (2)$$

D'Agostino & Stephens (1986) mendapati bahawa rumus statistik CR untuk pengiraan dapat ditulis sebagai:

$$CR = \left[F(X_{(i)}) - \frac{2i-1}{2n} \right]^2 + \frac{1}{12n} \quad (3)$$

dengan $X_{(i)}$ adalah statistik tertib ke- i .

STATISTIK CRAMER VON-MISES BERDASARKAN RUMUS WEIBULL

Rumus statistik CR dalam persamaan (1) bergantung kepada parameter bagi taburan $F(x)$ yang dianggarkan daripada sampel dan fungsi taburan empirik $F_n(x)$. Dalam bahagian ini kami cuba melihat kesesuaian dan ketepatan statistik CR jika pengubahsuaian ke atas FTE dilakukan dengan menggantikan FTE berdasarkan rumus Weibull. FTE bagi rumus Weibull ditakrifkan sebagai:

$$F_w(X_{(i)}) = \frac{i}{n+1}, \quad i = 1, 2, \dots, n+1 \quad (4)$$

Dengan menggantikan $F_w(x) = F_n(x)$ dan $z = F(x)$ ke dalam persamaan (1) diperolehi

$$\begin{aligned} W^2 &= n \int_{-\infty}^{\infty} \{F_n(x) - F(x)\}^2 dF(x) \\ &= n \int \left\{ \frac{i}{n+1} - z \right\}^2 dz = n \sum_{i=1}^n \int_{z_{i-1}}^{z_i} \left\{ \frac{i}{n+1} - z \right\}^2 dz \\ &= \frac{n}{n+1} \sum_{i=1}^n \left(z_i - \frac{2i+1}{2(n+1)} \right)^2 + \frac{n(n+4)}{12(n+1)^3} \end{aligned} \quad (5)$$

dengan $z_0 = 0$, dan $z_{n+1} = 1$. Menggantikan $z_i = F(X_{(i)})$ untuk $i = 1, 2, \dots, n$ dalam persamaan (5), rumus statistik CR berdasarkan rumus Weibull boleh ditulis sebagai:

$$CRW = \frac{n}{n+1} \sum_{i=1}^n \left[F(X_{(i)}) - \frac{2i+1}{2n} \right]^2 + \frac{n(n+4)}{12(n+1)^3} \quad (6)$$

TABURAN NORMAL

Taburan normal merupakan taburan terpenting dalam statistik. Fungsi ketumpatan kebarangkalian taburan Normal boleh ditulis sebagai:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right), \quad -\infty < x < \infty \quad (7)$$

dengan μ dan σ^2 masing-masing adalah min dan varians (Cohen & Whitten 1988). Fungsi taburan kumulatif bagi taburan normal boleh ditulis sebagai:

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \quad (8)$$

dengan $\Phi(x)$ adalah fungsi taburan kumulatif bagi taburan normal piawai. Pekali kepencongan, γ adalah sifar. Bentuk fungsi bagi taburan normal adalah bersimetri terhadap min dan berbentuk loceng. Anggaran parameter bagi taburan normal adalah:

$$\hat{\mu} = \bar{X} = n^{-1} \sum_{i=1}^n X_i$$

dan

$$\hat{\sigma}^2 = s^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (9)$$

NILAI GENTING

Dalam kajian ini, simulasi Monte-Carlo digunakan untuk mendapatkan nilai genting pengujian statistik CRW. Sampel bersaiz $n = 6, 8, \dots, 20$ bagi mewakili sampel bersaiz kecil dan $25, 30, \dots, 60$ untuk sampel bersaiz besar dijana dari taburan normal piawai dengan parameter taburan ditetapkan iaitu $m = 0$ dan $s = 1$. Parameter setiap sampel dianggar menggunakan persamaan (9). Nilai setiap pengujian statistik CRW dikira dan proses ini diulangi sehingga 50,000 kali. Untuk setiap pengujian statistik yang dijana, nilai 50,000 statistik ini disusun. Nilai kritikal statistik CRW pada aras keertian $\alpha = 0.99, 0.95, 0.90, 0.85, 0.75, 0.50, 0.25, 0.15, 0.10, 0.05$ dan 0.01 diperolehi berdasarkan nilai statistik tertib ke 1, 5, 10, 15, 25, 50, 75, 85, 90, 95 dan 99 persentil bagi nilai 50,000 statistik yang disusun. Nilai kritikal pengujian statistik CRW ini dapat ditunjukkan dalam Jadual 1. Nilai kritikal pengujian statistik SW, AD, CR dan KS diperolehi dari D'Agostino & Stephens (1986).

KEKUATAN PENGUJIAN STATISTIK

Dalam bahagian ini, perbandingan kekuatan pengujian statistik CRW berbanding dengan statistik Anderson-Darling (AD), Kolmogorov-Smirnov, Cramer-von Mises (CR) dan Shapiro-Wilk (SW) dilakukan berdasarkan simulasi Monte-Carlo. Kekuatan setiap pengujian statistik dinilai berdasarkan kadar penolakan hipotesis nol. Hipotesis yang diuji adalah:

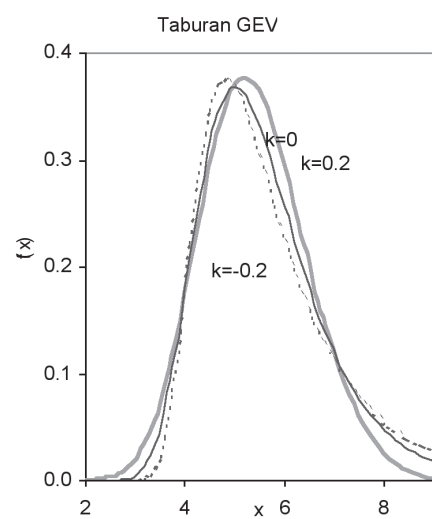
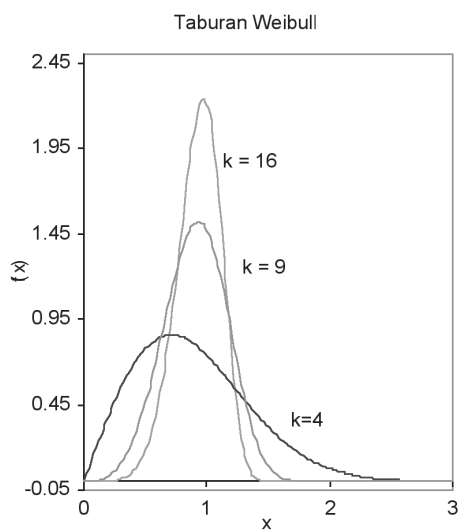
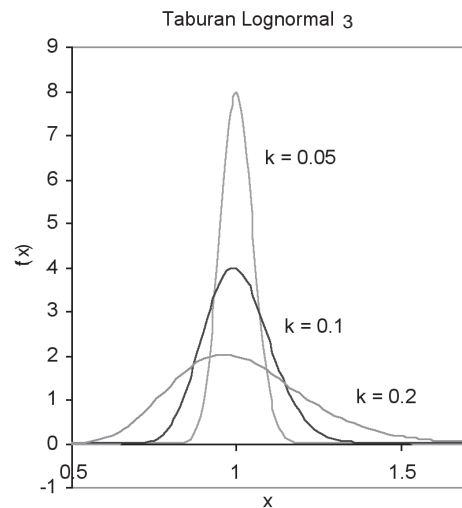
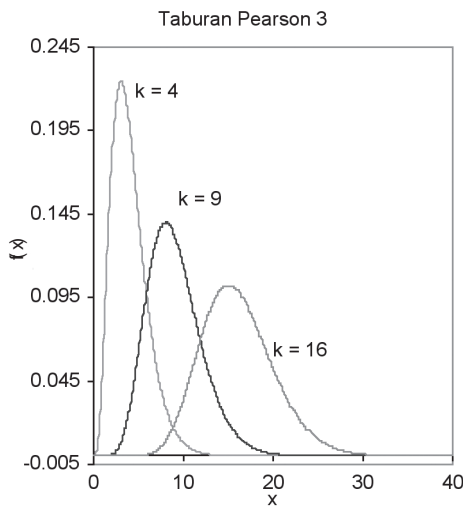
$$H_0 : F(x) = F_0(x) \text{ untuk kesemua } x \in (-\infty, \infty)$$

melawan

$$H_1 : F(x) \neq F_0(x).$$

JADUAL 1. Nilai kritikal bagi pengujian statistik CRW

n	Aras Keertian, α										
	0.01	0.05	0.10	0.15	0.25	0.50	0.75	0.85	0.90	0.95	0.99
6	0.047	0.051	0.055	0.058	0.063	0.076	0.093	0.105	0.114	0.129	0.163
8	0.039	0.044	0.049	0.052	0.058	0.071	0.090	0.103	0.113	0.130	0.168
10	0.034	0.040	0.045	0.048	0.054	0.068	0.088	0.102	0.112	0.130	0.171
12	0.031	0.038	0.042	0.045	0.051	0.066	0.087	0.101	0.112	0.131	0.172
14	0.029	0.036	0.040	0.044	0.049	0.065	0.086	0.100	0.112	0.131	0.175
16	0.027	0.034	0.038	0.042	0.048	0.063	0.084	0.099	0.110	0.131	0.176
18	0.027	0.033	0.037	0.041	0.047	0.062	0.083	0.098	0.110	0.130	0.178
20	0.025	0.032	0.036	0.040	0.046	0.061	0.082	0.097	0.110	0.129	0.176
25	0.024	0.030	0.034	0.038	0.044	0.059	0.081	0.096	0.108	0.128	0.175
30	0.023	0.029	0.033	0.036	0.042	0.057	0.080	0.095	0.107	0.128	0.176
35	0.022	0.028	0.032	0.035	0.041	0.056	0.079	0.095	0.107	0.128	0.176
40	0.021	0.027	0.031	0.035	0.041	0.056	0.078	0.095	0.107	0.128	0.178
45	0.021	0.027	0.031	0.034	0.040	0.055	0.078	0.094	0.107	0.128	0.178
50	0.020	0.026	0.031	0.034	0.040	0.055	0.077	0.093	0.106	0.127	0.177
60	0.020	0.026	0.030	0.033	0.039	0.054	0.077	0.093	0.106	0.128	0.181



RAJAH 1. Keluk fungsi taburan Pearson 3, Lognormal 3, Weibull dan GEV

Pengujian statistik yang menghasilkan kadar penolakan tertinggi adalah pengujian statistik yang terbaik. Bagi menentukan kesesuaian pengujian statistik yang digunakan, sebanyak 10,000 sampel bersaiz $n = 10, 30$ dan 50 dijana daripada taburan alternatif dan nilai setiap statistik ditentukan. Sebanyak 4 taburan yang berbeza digunakan sebagai taburan alternatif iaitu Taburan Generalized Extreme Value (GEV), Taburan Weibull (W), Taburan Lognormal 3 (LN3) dan Taburan Pearson 3.

TABURAN GENERALIZED EXTREME VALUE (GEV)

Fungsi ketumpatan kebarangkalian bagi taburan GEV diberikan oleh

$$f(x) = \frac{1}{\alpha} \left[1 - k \left(\frac{x - \xi}{\alpha} \right) \right]^{\frac{1}{k} - 1} \text{eks} \left(\left[1 - k \left(\frac{x - \xi}{\alpha} \right) \right]^{\frac{1}{k}} \right) \quad (10)$$

dengan parameter ξ, α dan k masing-masing adalah parameter lokasi, skala dan bentuk (Chowdhury et al. 1991). Pekali kepencongan bagi taburan GEV adalah:

$$\gamma = \frac{k}{|k|} \frac{-\Gamma(1 + 3k) + 3\Gamma(1 + k)\Gamma(1 + 2k) - 2\Gamma^3(1 + k)}{[\Gamma(1 + 2k) - \Gamma^2(1 + k)]^{3/2}} \quad (11)$$

yang bergantung kepada parameter k sahaja. Bentuk fungsi ketumpatan bagi taburan GEV untuk $\alpha = 1, \xi = 0$ dan $k = -0.2, 0$ dan 0.3 yang berpadanan dengan pekali kepencongan $g = 3.535, 1, 1396$ dan 0.069 masing-masing dapat ditunjukkan dalam Rajah 1. Fungsi taburan GEV didapati mempunyai bentuk pencong ke kanan.

Jadual 2 memberikan hasil simulasi perbandingan kesesuaian pengujian statistik bagi taburan normal pada aras keertian 5%. Jadual 2 menunjukkan bahawa pengujian statistik terbaik adalah ditunjukkan oleh statistik CWR bila $k = -0.2$ untuk kesemua nilai n dan $k = 0$ dan 0.2 bila $n = 30$ dan 50 . Manakala statistik SW terbaik diperolehi untuk $k = 0$ dan 0.2 bila $n = 30$ dan 50 . Kesemua pengujian statistik kecuali statistik KS memberikan keputusan yang sama untuk $k = -0.2$. Secara keseluruhannya didapati statistik CRW terbaik bila $k = 0.2$ dan statistik SW bila $k = -0.2$ dan 0 .

TABURAN WEIBULL (W)

Fungsi ketumpatan bagi taburan Weibull boleh ditulis sebagai

$$f(x) = \frac{k}{\beta^k} (x - \xi)^{k-1} \text{eks} \left(- \left(\frac{x - \xi}{\beta} \right)^k \right), \quad \xi < x < \infty, k > 0, \beta > 0 \quad (12)$$

Pekali kepencongan bagi taburan Weibull adalah:

$$\gamma = \frac{\Gamma_3 - 3\Gamma_2\Gamma_1 + 2\Gamma_1^3}{(\Gamma_2 - \Gamma_1^2)^{3/2}}$$

dengan $\Gamma_i = \Gamma \left(1 + \frac{i}{k} \right)$ dan $\Gamma ()$ adalah fungsi gamma (Cohen & Whitten 1988) dan ditakrifkan sebagai

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$$

Bentuk fungsi ketumpatan bagi taburan Weibull untuk $\xi = 0$ dan parameter $k = 2, 4$ dan 6 berpadanan dengan pekali kepencongan $\gamma = 0.631, -0.087$ dan -0.373 masing-masing dapat ditunjukkan dalam Rajah 1. Bentuk fungsi taburan Weibull pencong kekanan bila $k = 2$ dan pencong kekiri bila $k = 4$ dan 6 .

Jadual 2 menunjukkan statistik SW memberikan kadar penolakan tertinggi untuk hampir kesemua nilai k diikuti oleh AD, CR, CRW dan KS. Statistik CRW didapati memberikan kadar penolakan yang terendah bila bentuk taburan pencong kekiri. Statistik AD adalah terbaik untuk $n = 10$ manakala statistik SW untuk $n = 50$.

TABURAN LOGNORMAL 3 (LN3)

Fungsi ketumpatan bagi taburan LN3 boleh ditulis sebagai:

$$f(x) = \frac{1}{\sigma(x - \xi)\sqrt{2\pi}} \text{eks} \left[- \frac{1}{2} \left(\frac{\ln(x - \xi) - \mu}{\sigma} \right)^2 \right], \quad \xi < x < \infty, \sigma^2 > 0 \quad (13)$$

Pekali kepencongan bagi taburan LN3 adalah

$$\gamma = 3\phi + \phi^3 \quad (14)$$

dengan $\phi = \sqrt{\exp(\sigma^2) - 1}$ bergantung hanya kepada σ (Maidment 1992). Bentuk fungsi ketumpatan bagi taburan LN3 untuk $\xi = 0, \mu = 0$ dan $\sigma = 0.05, 0.1$ dan 0.2 yang berpadanan dengan pekali kepencongan $\gamma = 0.15, 0.302$ dan 0.614 masing-masing dapat ditunjukkan dalam Rajah 1. Taburan LN3 didapati mempunyai bentuk pencong kekanan dan menghampiri normal untuk kesemua nilai σ .

Hasil kajian menunjukkan bahawa statistik CRW memberikan kadar penolakan tertinggi diikuti oleh statistik WS, AD, CR dan KS untuk hampir kesemua nilai σ . Keputusan yang sama juga ditunjukkan dalam Jadual 2.

JADUAL 2. Perbandingan kekuatan pengujian statistik bagi taburan normal berdasarkan 10000 lelaran dari taburan GEV, Weibull, Lognormal dan Pearson

Saiz Sampel	Taburan	Parameter Bentuk	Pengujian Statistik					
			AD	CR	CRW	KS	SW	
10	GEV	0.2	0.06	0.06	0.08*	0.06	0.05	
		0	0.15	0.13	0.20*	0.12	0.15	
		-0.2	0.32	0.30	0.39*	0.26	0.33	
	Weibull	2	0.08	0.08	0.123*	0.07	0.08	
		4	0.19*	0.18	0.15	0.17	0.18	
		6	0.06*	0.06	0.04	0.06	0.06	
	Lognormal	0.05	0.05	0.05	0.07*	0.05	0.05	
		0.1	0.06	0.06	0.08*	0.06	0.06	
		0.2	0.08	0.08	0.12*	0.07	0.08	
	Pearson	4	0.13	0.12	0.19*	0.11	0.13	
		9	0.08	0.08	0.12*	0.08	0.08	
		16	0.07*	0.07	0.10	0.07	0.07	
		0.2	0.07	0.07	0.10*	0.06	0.07	
	30	GEV	0.2	0.07	0.07	0.10*	0.06	0.07
			0	0.40	0.35	0.43	0.28	0.45*
			-0.2	0.79	0.75	0.81	0.65	0.83*
Weibull		2	0.19	0.16	0.24*	0.13	0.22	
		4	0.44*	0.41	0.36	0.34	0.43	
		6	0.10	0.09	0.06	0.08	0.10*	
Lognormal		0.05	0.06	0.06	0.07*	0.05	0.05	
		0.1	0.08	0.08	0.10*	0.07	0.08	
		0.2	0.16	0.14	0.19*	0.12	0.18	
Pearson		4	0.37	0.33	0.41	0.25	0.42*	
		9	0.18	0.17	0.23*	0.14	0.21	
		16	0.12	0.11	0.15*	0.10	0.13	
		0.2	0.09	0.09	0.12*	0.08	0.08	
50		GEV	0.2	0.09	0.09	0.12*	0.08	0.08
			0	0.61	0.55	0.62	0.44	0.67*
			-0.2	0.95	0.93	0.95	0.86	0.97*
	Weibull	2	0.32	0.27	0.35	0.21	0.42*	
		4	0.60*	0.56	0.51	0.47	0.53	
		6	0.13	0.12	0.08	0.11	0.14*	
	Lognormal	0.05	0.07	0.06	0.07*	0.06	0.06	
		0.1	0.09	0.08	0.11*	0.07	0.09	
		0.2	0.25	0.22	0.27	0.18	0.28*	
	Pearson	4	0.57	0.50	0.59	0.39	0.66*	
		9	0.29	0.26	0.32	0.21	0.33*	
		16	0.18	0.16	0.20*	0.13	0.20	
		0.2	0.09	0.09	0.12*	0.08	0.08	

*Pengujian Statistik Terbaik

TABURAN PEARSON 3

Fungsi ketumpatan bagi taburan P3 boleh ditulis sebagai:

$$f(x) = \frac{\beta^{-k}}{\Gamma(k)} (x - \xi)^{k-1} \exp\left(\frac{-(x - \xi)}{\beta}\right),$$

$$\xi < x < \infty, k > 0 \tag{13}$$

Pekali kepencongan bagi taburan P3 adalah

$$\gamma = \frac{2}{\sqrt{k}}$$

Bentuk fungsi ketumpatan bagi taburan P3 untuk parameter $\xi = 0$ dan $k = 4, 9$ dan 16 berpadanan dengan pekali kepencongan $\gamma = 1, 0.67$ dan 0.5 masing-masing ditunjukkan dalam Rajah 1. Bentuk taburan P3 didapati mempunyai bentuk pencong kekanan bila $k = 4$ dan 9 , menghampiri normal apabila $k = 16$. Jadual 2 menunjukkan bahawa kadar penolakan tertinggi diperolehi berdasarkan statistik CRW diikuti oleh statistik SW, AD, CR dan KS.

PERBINCANGAN DAN KESIMPULAN

Dalam kajian ini, pengujian statistik Cramer-von Mises berdasarkan rumus Weibull (CRW) dalam menguji kesesuaian bagi taburan Normal cuba diperbaiki. Nilai

kritikal bagi pengujian statistik CRW dibangunkan berdasarkan simulasi Monte-Carlo. Kesesuaian dan kekuatan pengujian statistik CRW diuji dan dibandingkan dengan statistik AD, CR, SW dan KS.

Hasil perbandingan pengujian statistik, didapati statistik SW lebih baik berbanding dengan statistik yang lain untuk sampel n kurang daripada 30, manakala statistik SW dalam kebanyakan kes lebih baik berbanding statistik yang lain untuk sampel bersaiz besar.

Keputusan yang diperolehi adalah setara dengan kajian yang telah dilakukan oleh D'Agostino & Stephens (1986) dan Zhang & Wu (2005). Bagaimanapun apabila statistik CRW digunakan, statistik ini didapati lebih baik berbanding statistik SW dalam menguji kenormalan data.

RUJUKAN

- Chowdhury, J.U., Stedinger, J.R. & Lu, L. 1991. Goodness-of-fit tests for regional generalized extreme value flood distributions. *Water Resources Research* 27(7): 1765-1776.
- Cohen, A.C. & Whitten, B.J. 1988. *Parameter estimation in reliability and life span models*. New York: Marcel Dekker, Inc.
- D'Agostino, R.B. & Stephens, M.A. 1986. *Goodness-of-fit Techniques*. New York: Dekker.
- Maidment, D.R. 1992. *Handbook of hydrology*. New York: McGraw-Hill, Inc.
- Pavur, R.J., Edgeman, R.L. & Scott, R.C. 1992. Quadratic statistics for the goodness-of-fit test of the inverse gaussian distribution. *IEEE Transactions on Reliability* 41: 118-123.
- Swanepoel, J.W.H. & Graan, F.C.V. 2002. Goodness-of-fit tests based on estimated expectations of probability integral transformed order statistics. *Annals of the Institute of Statistical Mathematics* 54(3): 531-542.
- Zhang, J. & Wu, Y. 2005. Likelihood-ratio tests for normality. *Computational Statistics & Data Analysis* 49: 709-721.

Ani bin Shabri
Jabatan Matematik
Fakulti Sains
Universiti Teknologi Malaysia
81310 Skudai, Johor
Malaysia

Abdul Aziz Jemain
Pusat Pengajian Sains Matematik
Fakulti Sains & Teknologi
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor DE.
Malaysia

Diserahkan: 19 Januari 2007
Diterima : 10 April 2007