A NOTE ON SKEW-NORMAL DISTRIBUTION APPROXIMATION TO THE NEGATIVE BINOMAL DISTRIBUTION

JYH-JIUAN LIN[1], CHING-HUI CHANG[2*] AND ROSEMARY JOU[3]

[1]Department of Statistics
Tamkang University
151 Ying-Chuan Road, Tamsui, Taipei County 251,
Taiwan
117604@mail.tku.edu.tw

[2]Department of Applied Statistics and Information Science
Ming Chuan University
5 De-Ming Road, Gui-Shan District, Taoyuan 333,
Taiwan
chchang@mcu.edu.tw

[3] Graduate Institute of Management Sciences,
Tamkang University
151 Ying-Chuan Road, Tamsui, Taipei County 251,
Taiwan
893560135@s93.tku.edu.tw

Abstract: This article revisits the problem of approximating the negative binomial distribution. Its goal is to show that the skew-normal distribution, another alternative methodology, provides a much better approximation to negative binomial probabilities than the usual normal approximation.

Key-Words: Central Limit Theorem, cumulative distribution function (*cdf*), skew-normal distribution, skew parameter.

[*] Corresponding author

1. INTRODUCTION

This article revisits the problem of the negative binomial distribution approximation. Though the most commonly used solution of approximating the negative binomial distribution is by normal distribution, Peizer and Pratt (1968) has investigated this problem thoroughly and proposed a different method which outperforms the normal approximation. However, the methodology used in Peizer and Pratt (1968) is rather complicated and difficult. With different motive, we propose another alternative, the skew-normal distribution, and show that it provides a much better approximation to negative binomial probabilities than the usual normal approximation.

The skew-normal distribution is first introduced by O'Hagan and Leonard (1976), which is a generalization of the normal distribution. Its definition is as follows.

A *r.v.* X is said to follow a skew-normal distribution with location parameter μ, scale parameter σ and skew parameter λ (the $SN(\mu, \sigma^2, \lambda)$ distribution) if the *pdf* of X is given as

$$f(x \mid \mu, \sigma, \lambda) = (2/\sigma)\phi((x-\mu)/\sigma)\Phi(\lambda(x-\mu)/\sigma), \quad (1.1)$$

where ϕ and Φ denote the standard normal *pdf* and *cdf* respectively. The parameter space is: $\mu \in (-\infty, \infty)$, $\sigma > 0$ and $\lambda \in (-\infty, \infty)$. Besides, the distribution $SN(0,1,\lambda)$ with *pdf* $f(x \mid 0,1,\lambda) = 2\phi(x)\Phi(\lambda x)$ is called the standard skew-normal distribution. When $\lambda = 0$, the above *pdf* (1.1) boils down to the $N(\mu, \sigma^2)$ *pdf*. Also parameter $\lambda > 0$ (< 0) implies positively (negatively) skewed shape of $f(x \mid \mu, \sigma, \lambda)$ above. For more skew-normal distribution's properties, one can see Azzalini (1985), Gupta, Nguyen and Sanqui (2004), Arnold and Lin (2004) or Chang et al. (2008).

It seems that the skew-normal distribution is a better choice than the normal distribution to approximate the discrete distribution. The reason is that the skew-normal distribution can take negatively skewed, symmetric or positively skewed shape through its skew parameter and hence is more versatile than the normal distribution. Comparing to the normal distribution, Chang et al. (2008) has shown that the extra (skew) parameter λ would provide a better approximation of the binomial distribution by the $SN(\mu, \sigma^2, \lambda)$ distribution. Besides, Chang et al. (2008) has also mentioned that it should also work for negative binomial distribution approximation. However, a lot of details of the numerical results are not reported in Chang et al. (2008). Therefore, we will show that the extra (skew) parameter λ would provide a better approximation of the negative binomial distribution by the $SN(\mu, \sigma^2, \lambda)$ distribution than the normal distribution in this paper along with the details.

The skew-normal distribution and its usefulness to approximate the negative binomial distribution is discussed in the next section. The usual normal approximation method is also introduced in Section 2 and numerical results and comparisons are provided in Section 3.

2. APPROXIMATION BY THE NORMAL AND SKEW-NORMAL DISTRIBUTIONS

In this section we'll see how the $SN(\mu, \sigma^2, \lambda)$ distribution can be used to approximate a negative binomial distribution.

2.1. Normal Approximation by the CLT

It is a common practice to approximate the sampling distribution of the sample mean by the fact of the Central Limit Theorem (CLT). The CLT is a very powerful theorem and is applying to various areas. For the applications of the CLT, please see Wang and Wei (2008), Giribon, Revetria and Oliva (2007), Takacs, Nagy and Vujacic (2007) and Mceachen et al (2005). By the Central Limit Theorem, the *cdf* of $NB(r, \ p)$ (negative binomial distribution where Y represents the number of failures before the $r^{\underline{\text{th}}}$ success) at k, $k \in \{0, 1, 2, \ldots\}$, $F_{r,\,p}(k)$, i.e., $P(Y \le k)$, is approximated by the normal *cdf* as

$$F_{r,\,p}(k) \approx \Phi((p(k+0.5) - rq)/\sqrt{rq})$$

where $q = 1 - p$.

2.2. Skew-Normal Approximation

For a given $NB(r, \ p)$ distribution, the parameters of the approximating (or, matching) $SN(\mu, \sigma^2, \lambda)$ distribution are found by equating the first three central moments, i.e.,

(a) $E(Y) = rq/p = E(X)$

$$= \mu + \sigma(\sqrt{2/\pi})(\lambda/\sqrt{1+\lambda^2});$$

(b) $E(Y - E(Y))^2 = rq/p^2 = E(X - E(X))^2$

$$= \sigma^2\{1 - (2/\pi)\lambda^2/(1+\lambda^2)\};$$

(c) $E(Y - E(Y))^3 = rq(1+q)/p^3$

$$= E(X - E(X))^3$$

$$= \sigma^3(\sqrt{2/\pi})(\lambda/\sqrt{1+\lambda^2})^3((4/\pi) - 1).$$

Solving the above three equations is straight forward. We obtain the values of λ, σ and μ as follows.

(A) Given (r, p), the above last two equations yield

$$\lambda = \{(\sqrt{rq}\sqrt{(2/\pi)}((4/\pi)-1)/(1+q))^{2/3} + (2/\pi)-1\}^{-1/2}$$

(B) Given (r, p) and after obtaining λ as above, σ is found as

$$\sigma = (\sqrt{rq}\ /\ p)/\{1-(2/\pi)\lambda^2/(1+\lambda^2)\}^{1/2}\ .$$

(C) Finally, after obtaining λ and σ as above, get μ as

$$\mu = (\ rq\ /\ p) - \sigma(\sqrt{2/\pi})\lambda / \sqrt{1+\lambda^2}\,.$$

After obtaining λ, σ and μ as given above, the negative binomial *cdf* is approximated as

$$F_{r,\,p}(k) \approx \int_{-\infty}^{k+0.5} f(x\,|\,\mu,\sigma,\lambda)dx = \Psi_\lambda((k+0.5-\mu)/\sigma),$$

where $\Psi_\lambda(c) = \int_{-\infty}^{c} 2\phi(z)\Phi(\lambda z)dz$ is the standard skew-normal *cdf*.

2.3. Restriction

A mild restriction is put on r and p, which is

$$(2-p)/\sqrt{r(1-p)} < 0.99527,$$

in order to make the skew-normal approximation work, i.e., $F_{r,\,p}(k) \approx \Psi_\lambda((k+0.5-\mu)/\sigma)$. Since the skewness of the skew-normal distribution lies in (-0.99527, 0.99527), the skew-normal approximation is applied only when the skewness of $NB(r,\ p)$, $(2-p)/\sqrt{r(1-p)}$, lies in this interval. However, the negative binomial distribution is always positively skewed. Hence, strictly speaking, the restriction is that the skewness of $NB(r, p)$ lies in (0, 0.99527).

2.4. Evaluation Criterion

Chang et. al. (2008) uses the "maximal absolute error (*MABS*)" of the skew-normal approximation to investigate the binomial approximation $B(n,\ p)$ as in Schader and Schmid (1989) (though their work is related to the normal approximation to the binomial distribution). *MABS* can be substantial when p is near 0 or 1. In this paper we will also adopt the *MABS* criterion for this negative binomial approximation problem and it is defined as

$$MABS(r,\ p) = \max_{k \in R_Y} |\,F_{r,\,p}(k) - F^*{}_{r,\,p}(k)\,| \tag{2.1}$$

where $F^*_{r,p}(k) = \Phi((p(k+0.5)-rq)/\sqrt{rq}\,)$ is for the normal case and $F^*_{r,p}(k) = \Psi_\lambda((k+0.5-\mu)/\sigma)$ is for the skew-normal case.

3. EXAMPLES AND NUMERICAL RESULTS

As a demonstration of the improved approximation of the negative binomial distribution by the skew-normal distribution, we first show the matching skew-normal (as in Eq. (3.1) and normal (as in Eq. (3.2)) *pdf*s for the negative binomial distribution ($Y \sim NB(r=20,\ p=0.75)$) in the Figure 3.1 and the errors are reported in the Table 3.1. Here ENA and ESA stand for "Error in Normal Approximation" and "Error in Skew-Normal Approximation", and they are defined as in (3.3) and (3.4) respectively. It can be seen that the skew-normal curve follows the asymmetric negative binomial distribution closely.

$$Y \sim SN(\mu = 3.4107, \sigma^2 = (4.4148)^2, \lambda = 2.4224) \tag{3.1}$$

$$Y \sim N(r(1-p)/p = 6.6667, r(1-p)/p^2 = 8.8889) \tag{3.2}$$

$$ENA = F_{r,p}(k) - \Phi((p(k+0.5)-rq)/\sqrt{rq}) \tag{3.3}$$

$$ESA = F_{r,p}(k) - \Psi_\lambda((k+0.5-\mu)/\sigma) \tag{3.4}$$

Next, we calculate the *MABS* (see Eq. (2.1)) for the skew-normal approximations as well as the normal approximations (henceforth called *MABS*(*SN*) and *MABS*(*N*), respectively) for various *r* and *p*. The results are reported in Table 3.2. On the other hand, Figure 3.2 shows the plots of *MABS*(*SN*) and Figure 3.3 shows the plots of *MABS*(*N*) as a function of *p* (for fixed *r* = 20) and *r* (for fixed *p* = 0.75). Plots for other values of *r* showed almost identical patterns. The results and plots both clearly show that the skew-normal approximation is much superior to the usual normal approximation.

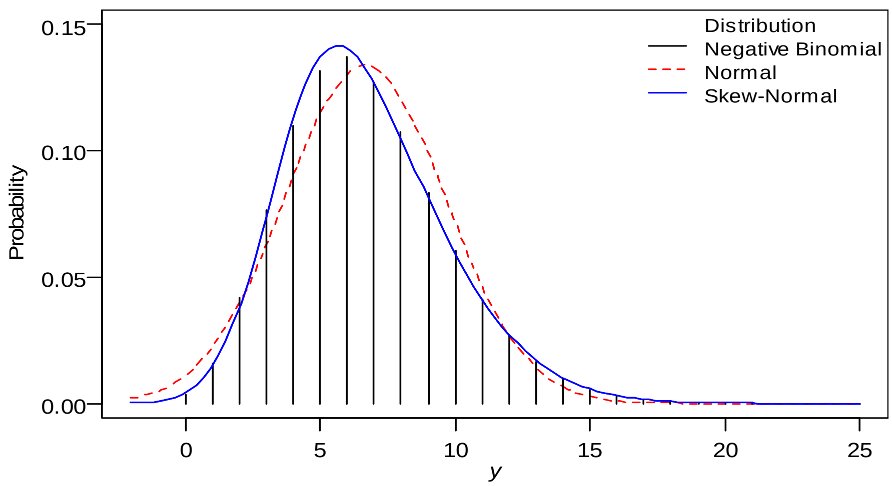

Figure 3.1. *NB*(20, 0.75) *pmf* with matching normal *pdf* and skew-normal *pdf*.

Table 3.1. Errors in approximating the *NB*(*r*, *p*) *cdf* for *r* = 20 and *p* = 0.75.

k	$F_{r,\,p}(k)$	$\Phi((p(k+0.5)-rq)/\sqrt{rq})$	$\Psi_\lambda((k+0.5-\mu)/\sigma)$	ENA	ESA
0	0.003171	0.019303	0.005514	-0.016132	-0.002343
1	0.019027	0.041552	0.020538	-0.022525	-0.001511
2	0.060649	0.081125	0.058806	-0.020476	0.001843
3	0.136957	0.144088	0.132844	-0.007131	0.004113
4	0.246648	0.233698	0.244519	0.012950	0.002129
5	0.378279	0.347783	0.380532	0.030495	-0.002253
6	0.515393	0.477710	0.520158	0.037683	-0.004765
7	0.642714	0.610073	0.646510	0.032641	-0.003796
8	0.750141	0.730696	0.751118	0.019445	-0.000977
9	0.833695	0.829027	0.832210	0.004668	0.001485
10	0.894272	0.900733	0.891686	-0.006461	0.002586
11	0.935574	0.947507	0.933097	-0.011933	0.002477
12	0.962249	0.974800	0.960490	-0.012552	0.001758
13	0.978664	0.989046	0.977708	-0.010383	0.000956
14	0.988337	0.995698	0.987991	-0.007361	0.000346
15	0.993818	0.998476	0.993826	-0.004657	-0.000007
16	0.996816	0.999513	0.996972	-0.002697	-0.000156
17	0.998403	0.999860	0.998584	-0.001457	-0.000181
18	0.999219	0.999964	0.999369	-0.000745	-0.000150
19	0.999626	0.999992	0.999732	-0.000365	-0.000106
20	0.999825	0.999998	0.999892	-0.000173	-0.000066
21	0.999920	1.000000	0.999958	-0.000080	-0.000038
22	0.999964	1.000000	0.999985	-0.000036	-0.000021
23	0.999984	1.000000	0.999995	-0.000016	-0.000011
24	0.999993	1.000000	0.999998	-0.000007	-0.000005
25	0.999997	1.000000	0.999999	-0.000003	-0.000002
26	0.999999	1.000000	1.000000	-0.000001	-0.000001
27	0.999999	1.000000	1.000000	-0.000001	0.000000
≥28	1.000000	1.000000	1.000000	0.000000	0.000000

Table 3.2. *MABS*[†] of Normal (*N*) and Skew-normal (*SN*) approximations.

p	*r* = 10		*r* = 20		*r* = 30		*r* = 40	
	N	*SN*	*N*	*SN*	*N*	*SN*	*N*	*SN*
0.05	0.042129	0.007047	0.029768	0.003545	0.024300	0.002732	0.021041	0.002406
0.10	0.042172	0.007056	0.029799	0.003549	0.024325	0.002735	0.021063	0.002408
0.15	0.042246	0.007070	0.029855	0.003556	0.024374	0.002741	0.021103	0.002413
0.20	0.042398	0.007096	0.029951	0.003572	0.024447	0.002749	0.021168	0.002420
0.25	0.042572	0.007173	0.030074	0.003582	0.024547	0.002760	0.021255	0.002431
0.30	0.042713	0.007183	0.030255	0.003614	0.024681	0.002775	0.021361	0.002440
0.35	0.043171	0.007339	0.030404	0.003652	0.024876	0.002797	0.021510	0.002462
0.40	0.043457	0.007326	0.030720	0.003666	0.025081	0.002818	0.021719	0.002488
0.45	0.043747	0.007607	0.031128	0.003746	0.025423	0.002834	0.021989	0.002515
0.50	0.044510	0.007731	0.031506	0.003823	0.025734	0.002903	0.022290	0.002557
0.55	0.044975	0.008115	0.032180	0.003870	0.026319	0.002922	0.022774	0.002611
0.60	0.047113	0.007898	0.032937	0.004032	0.026806	0.003031	0.023343	0.002678
0.65	0.048293	0.008851	0.034127	0.004064	0.027507	0.003116	0.024082	0.002736
0.70	0.049695	0.009698	0.035662	0.004423	0.028867	0.003281	0.024766	0.002881
0.75	0.052746	0.010689	0.037683	0.004765	0.030088	0.003503	0.026319	0.003033
0.80	0.058346	0.011237	0.039137	0.004894	0.032966	0.003708	0.027936	0.003247
0.85	0.065066	0.011371	0.045388	0.006149	0.035987	0.004373	0.030451	0.003570
0.90	0.060526	-	0.049883	0.007261	0.042437	0.005136	0.037231	0.004355
0.95	0.112839	-	0.058694	-	0.061012	0.007171	0.044428	0.007024

- (r, p) doesn't satisfy the inequality $(1+q)/\sqrt{q} < \sqrt{r}/(1.00475)$.

[†] Take k = 0, 1, …, 10^5 while *MABS* is calculated.

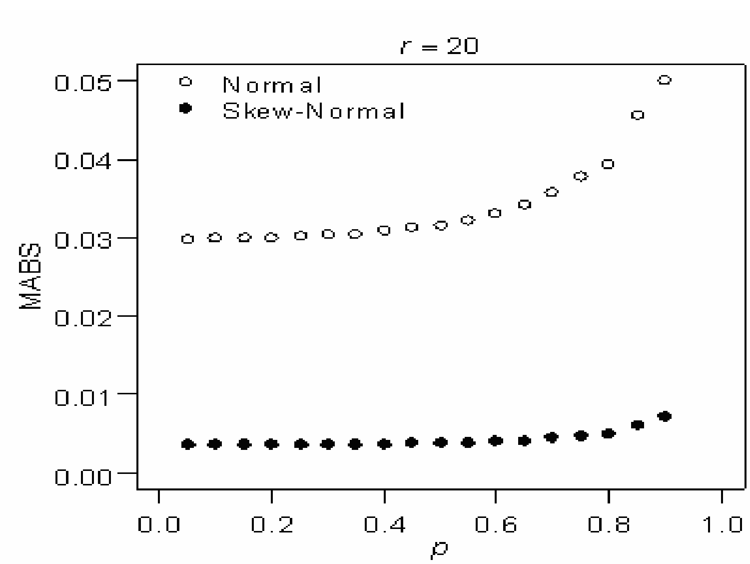

Figure 3.2. Plots of *MABS*(*SN*)

Table 3.2. (cont.) *MABS*[†] of Normal (*N*) and Skew-normal (*SN*) approximations.

p	$r = 50$		$r = 100$		$r = 150$		$r = 200$	
	N	*SN*	*N*	*SN*	*N*	*SN*	*N*	*SN*
0.05	0.018819	0.002172	0.013305	0.001546	0.010863	0.001250	0.009407	0.001069
0.10	0.018838	0.002174	0.013319	0.001548	0.010874	0.001251	0.009417	0.001070
0.15	0.018874	0.002178	0.013344	0.001552	0.010895	0.001254	0.009435	0.001073
0.20	0.018931	0.002186	0.013384	0.001557	0.010927	0.001258	0.009463	0.001076
0.25	0.019009	0.002194	0.013439	0.001564	0.010972	0.001264	0.009502	0.001082
0.30	0.019118	0.002207	0.013510	0.001574	0.011032	0.001273	0.009555	0.001089
0.35	0.019257	0.002223	0.013613	0.001586	0.011113	0.001283	0.009623	0.001098
0.40	0.019426	0.002248	0.013735	0.001604	0.011214	0.001298	0.009712	0.001110
0.45	0.019632	0.002277	0.013891	0.001625	0.011347	0.001315	0.009828	0.001126
0.50	0.019939	0.002301	0.014102	0.001649	0.011515	0.001338	0.009973	0.001145
0.55	0.020332	0.002347	0.014381	0.001689	0.011742	0.001368	0.010168	0.001171
0.60	0.020822	0.002422	0.014736	0.001734	0.012011	0.001406	0.010408	0.001203
0.65	0.021450	0.002471	0.015182	0.001787	0.012398	0.001457	0.010737	0.001247
0.70	0.022381	0.002608	0.015790	0.001869	0.012883	0.001523	0.011169	0.001306
0.75	0.023637	0.002750	0.016632	0.001986	0.013549	0.001611	0.011770	0.001380
0.80	0.025416	0.002878	0.017771	0.002131	0.014603	0.001742	0.012591	0.001505
0.85	0.028043	0.003165	0.019858	0.002382	0.016174	0.001945	0.013959	0.001663
0.90	0.033356	0.003633	0.022730	0.002654	0.019003	0.002233	0.016306	0.001969
0.95	0.046076	0.005705	0.031012	0.003402	0.025441	0.003156	0.022313	0.002745

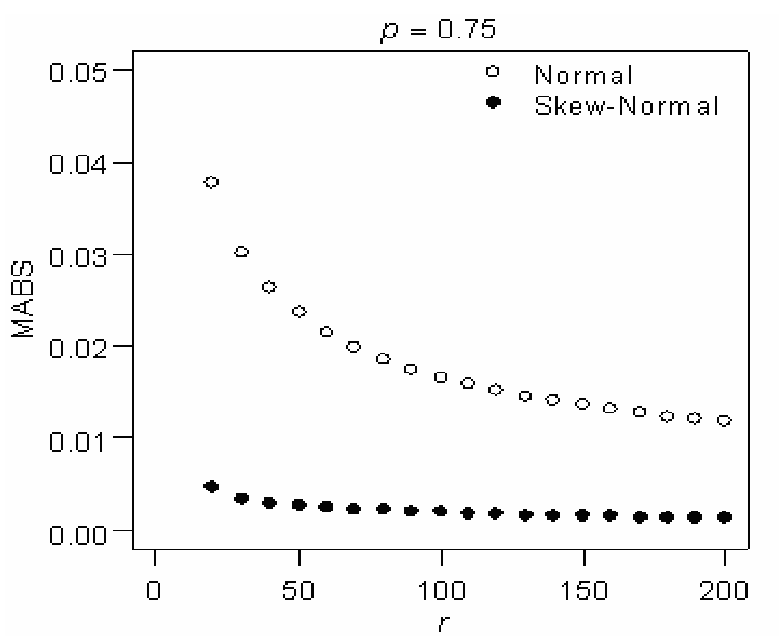

Figure 3.3. Plots of *MABS*(*N*).

Concluding Remark: In this paper we have shown that the skew-normal distribution can provide a far better approximation than the normal distribution, and we believe that it can also be used to approximate discrete distributions other than the negative binomial distribution.

***Acknowledgement*：** The second author's research has been supported partially by a research grant from the National Science Council (98-2118-M-130 -002).

References

[1] Arnold, B. C. and Lin, G. D. (2004). Characterizations of the Skew-Normal and Generalized Chi Distributions. *Sankhyā: The Indian Journal of Statistics*, Vol. 66, Part - 4, 593 - 606.

[2] Azzalini, A. (1985). A Class of Distributions Which Includes the Normal Ones. *Scandinavian Journal of Statistics*, 12, 171 - 178.

[3] Chang, C. H., Lin, J. J., Pal, N. and Chiang, M. C. (2008). A Note on Improved Approximation of the Binomial Distribution by the Skew-Normal Distribution. *The American Statistician*, 62(2), 167 - 170.

[4] Giribone, P., Revetria, R. and Oliva, F. (2007). A Stochastic Simulation Model for Representing a Clinical Pathology Laboratory Workout. 6th *WSEAS International Conference on System Science and Simulation in Engineering*, Venice, Italy, 346-351.

[5] Gupta, A. K., Nguyen, T. T. and Sanqui, J. A. T. (2004). Characterization of the Skew-Normal Distribution. *Annals of the Institute of Statistical Mathematics*, 56(2), 351 - 360.

[6] Marta Takacs, Karoly Nagy and Vladimir Vujacic (2007). Type-2 Fuzzy Set Representation of Stochastic Adding A/D Conversion. Proc. of the 9th WSEAS International Conference on Mathematical and Computational Methods in Science and Engineering, 198-201.

[7] Mceachen, J. C., Zachary J. M., Wang, J. and Cheng, K. W. (2005). Conversation Exchange Dynamics: A New Signal Primitive For Visualizing Network Intrusion Detection. 4th *WSEAS International Conference On Electronics, Control And Signal Processing*, Miami, Florida, USA, 199-204.

[8] O'Hagan, A. and Leonard, T. (1976). Bayes Estimation Subject to Uncertainty about Parameter Constraints. *Biometrika*, 63, 201 - 203.

[9] Peizer, D. B. and Pratt, J. W. (1968). A Normal Approximation for Binomial, *F*, Beta, and Other Common, Related Tail Probabilities, I, *Journal of the American Statistical Association*, 63(324), 1416 – 1456.

[10] Schader, M. and Schmid, F. (1989). Two Rules of Thumb for the Approximation of the Binomial Distribution by the Normal Distribution, *The American Statistician*, 43(1), 23 - 24.

[11] Wang J. and Wei C. (2008). The asymptotical Behavior of Probability Measures for the Fluctuations of Stochastic Models, *WSEAS Transactions On Mathematics*, Issue 5, Volume 7, 273-282.