

Time Clustering of the Occurrence of Sleep Apnea

Nur Zakiah Mohd Saat

Department of Biomedical Science, Faculty of Allied Health Science,
University Kebangsaan Malaysia, Kuala Lumpur

Abstract: Problem statement: In this study, the inequality of apnea rates during sleep was investigated. **Approach:** This study used a database comprising 13 patients with apnea problems. The data was extracted from the MIT-BIH database. The Lorenz curve and Gini index were used to determine the inequality of apnea rates. In order to test for cluster, random or regular patterns of apnea, an index of runs was used. **Results:** The Lorenz curve showed that the distribution of apnea events for every 10% period of sleep varied between each patient. The Gini index ranged between 0.148 and 0.614. Most of the patients had a Gini coefficient lower than 0.5. The index of runs was between 0.115 and 0.353, which indicates that most patients had a regular pattern of apnea and that it was not random. **Conclusion:** The distribution of the apnea is almost equal and most patients have regular patterns of occurrence of apnea.

Key words: Apnea, gini index, lorenz, inequality, run index

INTRODUCTION

Apnea is a pause in breathing during sleep and it disrupts the quality of sleep. Good sleep quality is important as it ensures a productive day. Sleep disorders, especially apnea, may cause sleepiness during the day. Previous studies indicate that sleep apnea may cause automobile accidents and a low workplace focus (Barbe *et al.*, 1998). Moreover, sleepiness can disrupt any daytime activity. The prevalence of apnea is higher among male patients than among female patients. Patients aged older than 60 also show a higher prevalence of apnea compared to younger adult age groups (Vagiakis *et al.*, 2006). Previous studies have shown that apnea is related to hypertension, diabetes and cardiac disease.

The Apnea Hypopnea Index (AHI) is used to indicate the severity of apnea. The AHI categories of apnea severity are <5, <15 and <30 (Duran *et al.*, 2001). The AHI is defined as the number of apnea and hypopnea events per hour (Young *et al.*, 1993). Even though a higher AHI indicates more serious apnea for a particular patient, further study is required in order to determine the proportional distribution of apnea over a given period of time, e.g. a 10% period of sleep. This information may shed some light on the behavior occurrence of apnea among adults.

There has been an increase in research into sleep apnea since the 1980s due to the rapid increase of community awareness of apnea (Tsai *et al.*, 1999; Drazen, 2002). Some of these studies have undertaken

modelling of the distribution of apnea and the severity of apnea (Saat *et al.*, 2010a; 2010b). However, there has been no study on the similarities of the patterns of apnea during sleep. A study of the distribution of apnea is essential in order to determine if the occurrence of apnea has a cluster (aggregation) distribution, a regular distribution, or a random distribution. From a statistical perspective, this analysis will give more information to patients regarding their apnea problem. Although they may know the severity of their apnea using the AHI, the AHI does not measure the distribution of apnea.

In order to study the disproportional or the inequality occurrence of apnea, the Lorenz curve, the Gini index and the runs test were used. The Lorenz curve and the Gini index are widely used in the economy sector, such as in the study of inequality of poverty distribution in Malaysia (Anand, 1977). Previous studies have used the Lorenz curve to determine the inequality of plant size (Damgaard and Weiner, 2000; He *et al.*, 2005). In the medical field, the Gini index has been used to describe the inequality of healthcare among low income patients (Weich *et al.*, 2001).

MATERIALS AND METHODS

The database comprised of 13 patients with apnea. The data was extracted from the MIT-BIH database which is <http://www.physionet.org/physiobank/database/>. The data consisted of records of apnea

events and non-apnea events every 30 seconds (epoch) (Goldberger *et al.*, 2000). The total apnea was calculated for every 10% period of sleep according to the Apnea Hypopnea Index (AHI). The AHI is calculated by dividing the total apnea experienced during sleep by the number of hours of sleep (Young *et al.*, 1993) and indicates the severity of apnea among patients. An AHI of less than 15 indicates apnea of moderate severity while an AHI of more than 25 indicates a high severity of apnea. The sleeping period was divided into 10. Then the rate of apnea for every 10% period of sleep was calculated. This was done in order to standardize the group of data for each patient. Each patient has sleeps for a different period.

In order to determine the inequality of the distribution of apnea for each patient, the Lorenz curve and the Gini index were used. The Lorenz curve compares the cumulative proportion of the ordered data of apnea with a uniform distribution. Patients that have an equal distribution of apnea for every 10% period of sleep will have a cumulative proportion line that is close to a uniform distribution. Furthermore, graphically the line of the Lorenz curve is close to a 45° line, indicating a uniform distribution. This study also aimed to determine whether patterns of the rate of apnea were cluster, random or regular.

The Gini coefficient: The Gini coefficient was also used to measure the inequality or the degree of variation in the 10% period of sleep for each patient. It is calculated as the sum of the difference between the total area under the uniform distribution (which is the 45° line) and the Lorenz curve. The equation for the Gini coefficient is given in Eq. 1, where x_i is the number of apnea events in the i th interval, i is the rank based on the length of data, μ is the mean of an apnea event for particular patient and n is the length of the apnea event. For the purposes of this study n is 10 because the data was grouped into 10 periods of sleep duration (Dixon *et al.*, 1987):

$$G = \frac{\sum_{i=1}^n 2i - n - 1}{n^2 \mu} \quad (1)$$

The minimum value of the Gini coefficient is 0, indicating equality of apnea distribution for every 10% period of sleep and the maximum value is 1, indicating that distributions are unequal for every 10% period of sleep.

Runs test: Let x_1, x_2, \dots, x_n be an order sequence of apnea and no apnea events. An apnea event is represented by 1 and a non-apnea event is represented

by 0. One run is defined as one symbol, either 0 or 1 and is followed by a different symbol. For example, the sample 0, 1, 0, 0, 1, 1 is said to have 4 runs. The runs are determined as '0', '1', '00' and '11'. A sequence of 0s or 1s which is persistent, meaning that the sequence of 0s followed by 0s is high and the sequence of 1s followed by 1s is also high, will give a smaller number of runs. In this study the index of runs was evaluated and compared with the Gini coefficient.

The runs test is given by the equation below with the mean (μ), variance (σ^2) and index of runs (I):

$$\begin{aligned} \mu &= \frac{N + 2n_1 n_2}{N} \\ \sigma^2 &= \frac{2n_1 n_2 (2n_1 n_2 - N)}{N^2 (N - 1)} \\ I &= \frac{R - 2}{\max R - 2} \\ \max R &= 2n_1 + 1 \end{aligned}$$

Where:

- N_1 = Number of occurrences of no apnea event
- N_2 = Number of occurrences of apnea event
- R = Number of runs
- N = Length of data

RESULTS

The descriptive information of patients is shown in Fig. 1. The bar chart shows that patient 7 had the highest AHI (>40) indicating severe apnea. The Lorenz curve shown in Fig. 2 revealed that the distribution of apnea events between each patient is variable. In Fig. 3, the cumulative graph shows that for every 10% period of sleep the apnea events vary. The graph indicates that a particular patient may have high apnea rates at the beginning of sleep and low apnea rates at the end of sleep.

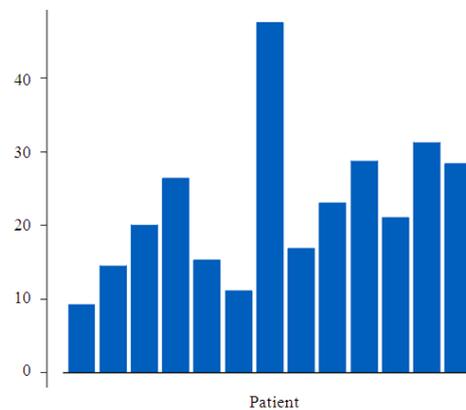


Fig. 1: Graph of AHI for every patient

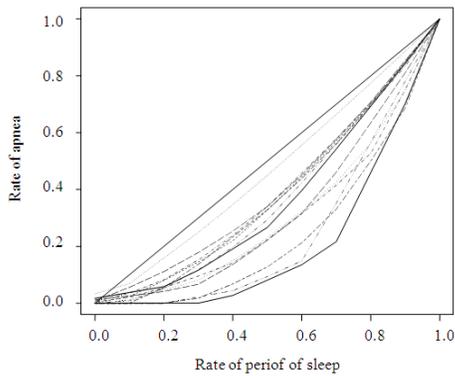


Fig. 2: Lorenz curve of the occurrence of apnea for every 10% period of sleep

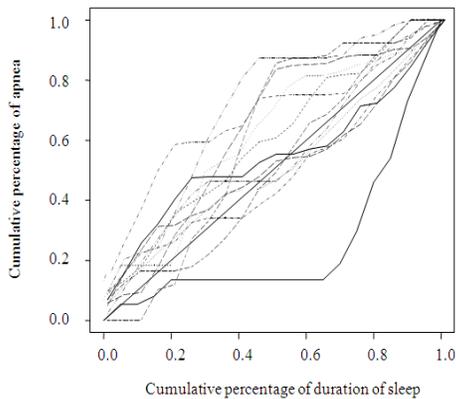


Fig. 3: The cumulative proportion of apnea events

Distribution of apnea events among patients is shown in Table 1. The results show that the highest run was 161 from a patient with moderate apnea. The lowest run was 18, which indicates that the difference between apnea and non-apnea events is low. However, the index of runs shows that the patient with severe apnea had the lowest run of 0.11. This suggests that the particular patient had less frequent alternating between apnea and non-apnea events. The lowest run indicates that there is cluster distribution of apnea and non-apnea events. On the other hand, the higher value of the index of runs suggests that apnea and non-apnea event occurrence is regular.

For example, modelling for the patient with an index of runs of 0.439 shows that when the data is simulated 10000 times, the percentile of 2.5, 5, 95 and 97.5% is between 0.075 and 0.249; this indicates that the occurrence of apnea was regular. All the patients had regular patterns of occurrence of apnea. However, for those patients in the mild apnea category, the occurrence of apnea was random.

Table 1: Permutations for 10000 simulations for test of randomness

AHI	Runs	Index of run	Percentile			
			2.5%	5%	95%	97.5%
15-30	18	0.219	0.123	0.137	0.370	0.397
Mild	78	0.439	0.075	0.087	0.231	0.249
	121	0.247	0.042	0.047	0.125	0.138
≥30	134	0.209	0.044	0.042	0.109	0.113
Severe	130	0.353	0.060	0.063	0.157	0.160
	94	0.353	0.068	0.077	0.209	0.226
	129	0.115	0.017	0.019	0.046	0.047
	149	0.344	0.043	0.049	0.129	0.144
	104	0.174	0.044	0.047	0.106	0.110
	161	0.233	0.032	0.038	0.095	0.110
	81	0.156	0.050	0.054	0.119	0.139
111	0.115	0.017	0.019	0.046	0.047	
43	0.283	0.069	0.069	0.207	0.224	

Table 2: Permutation for 100 simulation of Gini index

AHI	Gini index	Percentile			
		2.5%	5%	95%	97.5%
15-30	0.614	0.143	0.157	0.349	0.369
moderate	0.441	0.092	0.100	0.219	0.230
	0.556	0.082	0.091	0.194	0.206
>30	0.446	0.052	0.057	0.124	0.131
severe	0.414	0.042	0.046	0.100	0.106
	0.544	0.063	0.070	0.151	0.161
	0.148	0.019	0.021	0.046	0.048
	0.299	0.057	0.063	0.137	0.144
	0.334	0.041	0.044	0.100	0.105
	0.284	0.039	0.043	0.093	0.098
	0.267	0.050	0.055	0.120	0.127
	0.331	0.045	0.050	0.107	0.114
	0.309	0.082	0.092	0.202	0.212

In order to determine the distribution of apnea rates according to time, a cumulative graph was plotted (Fig. 3). The graph indicates that apnea rates varied at the beginning, the middle and the end of the sleeping period.

The randomness of apnea rates was analyzed using a runs test. Table 1 indicates that the highest run was 161 from a patient with severe apnea and the lowest run was 18 from a patient with mild apnea. This suggests that the alternating of apnea and non-apnea events was higher in patients that had a higher index of runs. The index of runs varied among mild and severe apnea patients. For example, Table 1 demonstrates that a patient with mild apnea had an index of runs of between 0.219 and 0.439. Meanwhile, the Gini index is in the range of 0.148 to 0.614 (Table 2). Previous studies suggest that a Gini coefficient below 0.2 is indicative of low inequality and a Gini index higher than 0.6 is indicative of a large inequality (Weich *et al.*, 2001). According to Table 2, the lowest Gini coefficient is for the patient that has severe apnea (0.148). This indicates that the distribution of apnea for every 10% period of sleep is almost equal for that patient. Moreover, the

Gini indices for moderate and severe apnea show that the distribution of apnea is almost equal. Most of the patients have Gini coefficients lower than 0.5. This suggests that the gaps between the patients in terms of distribution of apnea are small, as they have almost similar Gini coefficients. On the other hand, the highest Gini coefficient of 0.614 is from a patient who has mild apnea. This indicates that the distribution of apnea for every 10% period of sleep is unequal.

DISCUSSION

This study discusses the use of existing methods to perform inequality tests, namely the Gini index and the runs test for testing randomness of apnea distribution. Earlier studies used the Gini index to study the inequality of plant size (He *et al.*, 2005). Previous approaches used the AHI to classify the apnea. The AHI is defined as the number of apnea events per hour of sleep. However, the AHI reflects only the severity of apnea; it does not describe the disaggregation of apnea during sleep.

The present study has shown that the occurrence of apnea events for every patient varies for every 10% period of sleep. The results of the runs test indicate a higher index of runs among patients that have mild and moderate apnea severity. This suggests that the process of alternating between apnea and non-apnea events is more frequent in patients that have at least mild severity of apnea. Moreover, the occurrence of apnea showed less inequality between patients with mild and severe apnea. In this study, a runs test was used to evaluate the randomness of the distribution of the occurrence of apnea and it was found that the occurrence of apnea is not random (index of runs). This suggests that the occurrence of apnea is regular.

In this study, the Gini index was used to evaluate the inequality of the occurrence of apnea for every 10% period of sleep. The results indicate that the distribution of apnea has low inequality for most of the patients. Compared with all categories of AHI, the Gini index for patients with severe apnea is almost equal for every 10% period of sleep. Meanwhile, the Gini index for mild apnea varies for every 10% period of sleep. This suggests that patients with severe apnea according to the AHI have almost the same pattern for every period of sleep, while the pattern for patients with mild apnea varies for every period of sleep. This is likely due to the older age group and health status of the patients. Previous studies on apnea indicate that the AHI increases as age increases (Weich *et al.*, 2001). Furthermore, the AHI is high among patients with obesity problems, smokers and alcoholics (Vgontzas *et al.*, 1998; Kim *et al.*, 2004).

CONCLUSION

Research into inequality has been of great interest to those in the field of economics. Various studies have used the Gini coefficient in order to find inequality. In this study, the Lorenz curve, the Gini coefficient and a runs test were used to determine the inequality of apnea distribution. In this study, disproportional distribution of apnea events was low for the majority of apnea patients. However, the Gini coefficient was high, especially for patients with mild severity of apnea. In terms of randomness, we found that most patients have clustering of apnea and non-apnea events for both categories of mild and severe apnea. In general, it can be said that the inequality is also due to the age and health status of the patients. Based on these findings we can conclude that apnea distributions can be categorized into two groups: low inequality and mild inequality. The low inequality of distribution of apnea is among the severe apnea patients while mild inequality is among patients with mild apnea.

ACKNOWLEDGEMENT

This research was supported by the research grant ERGS/1/2011/SKK/UKM/03/12 from Ministry of Higher Education Malaysia and Universiti Kebangsaan Malaysia.

REFERENCES

- Anand, S., 1977. Aspects of poverty in Malaysia. *Rev. Income Wealth*, 23: 1-16. DOI: 10.1111/j.1475-4991.1977.tb00001.x
- Barbe, F., J. Pericas, A. Munoz, L. Findley and J.M. Anto *et al.*, 1998. Automobile accidents in patients with sleep apnea syndrome. An epidemiological and mechanistic study. *Am. J. Respir. Crit. Care Med.*, 158: 18-22. PMID: 9655701
- Damgaard, C. and J. Weiner, 2000. Describing inequality in plant size or fecundity. *Ecology*, 81: 1139-1142. DOI: 10.1890/0012-9658(2000)081[1139:DIIPSO]2.0.CO;2
- Dixon, P.M., J. Weiner, T. Mitchell-Olds and R. Woodley, 1987. Bootstrapping the Gini coefficient of inequality. *Ecology*, 68: 1548-1551. DOI: 10.2307/1939238
- Drazen, J.M. 2002. Sleep Apnea Syndrome. *N. Eng. J. Med.*
- Duran, J., S. Esnaola R. Rubio and A. Iztueta, 2001. Obstructive sleep apnea-hypopnea and related clinical features in a population-based sample of subjects aged 30-70 Yr. *Am. J. Respir. Crit. Care Med.*, 163: 685-689.

- Goldberger, A.L., L.A. Amaral, L. Glass, J.M. Hausdorff and P.C. Ivanov *et al.*, 2000. PhysioBank, PhysioToolkit and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, 101: E215-E220. PMID: 10851218
- He, Z., Z. Ma, K.M. Brown and J.P. Lynch, 2005. Assessment of inequality of root hair density in *Arabidopsis thaliana* using the gini coefficient: A close look at the effect of phosphorus and its interaction with ethylene. *Ann. Bot.*, 95: 287-293. DOI: 10.1093/aob/mci024 PMID: 15546931
- Kim, J., K. In, J. Kim, S. You and K. Kang *et al.*, 2004. Prevalence of sleep-disordered breathing in middle-aged Korean men and women. *Am. J. Respir. Crit. Care Med.*, 170: 1108-1113. PMID: 15347562
- Saat, N.Z.M., A.A. Jemain and K. Ibrahim, 2010a. Statistical model of the occurrence of sleep apnea. *J. Applied Sci.*, 10: 752-758.
- Saat, N.Z.M., K. Ibrahim and A.A. Jemain, 2010b. Bayesian methods for ranking the severity of apnea among patients. *Am. J. Applied Sci.*, 7: 167-170. DOI: 10.3844/ajassp.2010.167.170
- Tsai, W.H., W.W. Flemons, W.A. Whitelaw and J.E. Remmers, 1999. A comparison of apnea-hypopnea indices derived from different definitions of hypopnea. *Am. J. Respir. Crit. Care Med.*, 159: 43-48. PMID: 9872816
- Vagiakis, E., F. Kapsimalis, I. Lagogianni, H. Perraki and A. Minaritzoglou *et al.*, 2006. Gender differences on polysomnographic findings in Greek subjects with obstructive sleep apnea syndrome. *Sleep Med.*, 7: 424-430. DOI: 10.1016/j.sleep.2005.12.014
- Vgontzas, A.N., E.O. Bixler, T.L. Tan, D. Kantner and L.F. Martin *et al.*, 1998. Obesity without sleep apnea is associated with daytime sleepiness. *Arch. Int. Med.*, 158: 1333-1337. PMID: 9645828
- Weich, S., G. Lewis and S.P. Jenkins, 2001. Income inequality and the prevalence of common mental disorders in Britain. *Br. J. Psy.*, 178: 222-227. DOI: 10.1192/bjp.178.3.222
- Young, T., M. Palta, J. Dempsey, J. Skatrud and S. Weber *et al.*, 1993. The occurrence of sleep-disordered breathing among middle-aged adults. *N. Eng. J. Med.*, 328: 1230-1235. PMID: 8464434