รายงานสรุปเนื้อหาและการนำไปใช้ประโยชน์

จากการเข้าอบรม สัมมนา หรือประชุมวิชาการ

ข้าพเจ้า นางศิริพร สมุทรวชิรวงษ์ ตำแหน่ง อาจารย์ สังกัดสาขาวิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้ ขอนำเสนอรายงานสรุปเนื้อหาและการนำไปใช้ประโยชน์จากการเข้าร่วมนำเสนอและประชุมวิชาการระดับนานาชาติ "The 2nd International Conference of Multidisciplinary Approaches on UN Sustainable Development Goals (UNSDGs)" ระหว่างวันที่ 28-29 ธันวาคม 2560 ณ จังหวัดกรุงเทพมหานคร ตามหนังสือขออนุญาตเดินทางไปราชการ เลขที่ ศธ.0523.4.7/244 ลงวันที่ 30 ตุลาคม 2560 ซึ่งเข้าร่วมนำเสนอผลงานและประชุมวิชาการดังกล่าว ข้าพเจ้าได้เลือกใช้งบประมาณการพัฒนาบุคลากรกรณีที่ 3 และข้าพเจ้าขอนำเสนอสรุปเนื้อหาและการนำไปใช้ประโยชน์ของการนำเสนอผลงานและร่วมประชุมวิชาการ ดังต่อไปนี้

สรุปเนื้อหาจากการเข้าร่วมนำเสนอผลงานและประชุมวิชาการ

ข้าพเจ้าร่วมนำเสนอผลงานวิชาการในหัวข้อเรื่อง "A negative binomial-new weighted Lindley distribution for count data and its application to hospitalized patients with diabetes at Ratchaburi hospital, Thailand" ซึ่งเป็นการศึกษาข้อมูลเชิงนับ (Count Data) และสร้างการแจกแจงใหม่ ซึ่งเป็นแนวทางหนึ่งที่มีความสำคัญในการวิเคราะห์ข้อมูล ที่ไม่เป็นไปตามข้อสมมติของการแจกแจงปัวซง ที่นิยมใช้ในข้อมูลเชิงนับคือ ค่าเฉลี่ย (mean) ต้องเท่ากับค่าความแปรปรวน (variance) ข้อมูลเชิงนับที่สนใจในการศึกษาครั้งนี้ คือจำนวนผู้ป่วยโรคเบาหวานที่เข้ารับการรักษาในโรงพยาบาล โรคเบาหวานคือโรคที่เซลร่างกายมีความผิดปรกติในขบวนการเปลี่ยนน้ำตาลในเลือดให้เป็นพลังงาน เมื่อน้ำตาลไม่ได้ถูกใช้จึงทำให้ระดับน้ำตาลในเลือดสูงขึ้นกว่าระดับผิดปรกติ ในปัจจุบันหลายประเทศใช้เกณฑ์ระดับน้ำตาลที่มากกว่า 126 มก./ดล. โดยมีข้อแม้ว่าเป็นค่าของน้ำตาลในน้ำเลือดหลังจากอดอาหารมาอย่างน้อย 8 ชม. ดังนั้นการพัฒนาการแจกแจงใหม่ที่จะช่วยในการแก้ปัญหาข้อมูลเชิงนับที่มีความแปรปรวนมากกว่าค่าเฉลี่ยของข้อมูลดังกล่าว จึงควรมีการศึกษาวิจัย เพื่อสร้างตัวแบบที่เหมาะสมกับข้อมูลที่มีความแปรปรวนมากกว่าค่าเฉลี่ย อันจะส่งผลต่อการวิเคราะห์ข้อมูลเชิงนับที่มีประสิทธิภาพมากขึ้น และนำมาประยุกต์ใช้ในข้อมูลเชิงนับเพื่อการอธิบายจำนวนผู้ป่วยโรคเบาหวานที่เข้ารับการรักษาในโรงพยาบาล เพื่อกำหนดแนวทางในการรักษาให้มีประสิทธิภาพขึ้นได้ต่อไป

ในการร่วมรับฟังกานำเสนอผลงานวิจัยของวิทยากรผู้ทรงคุณวุฒิ (Keynote speakers) และผู้เข้าร่วมนำเสนอผลงานวิชาการท่านอื่น ๆ มีหัวข้อที่น่าสนใจ หลากหลาย และสอดคล้องกับวิสัยทัศน์เชิงนโยบายการพัฒนาเศรษฐกิจของประเทศไทย “ไทยแลนด์4.0 ภายใต้การนำของพลเอกประยุทธ์ จันทร์โอชา นายกรัฐมนตรีและหัวหน้าคณะรักษาความสงบแห่งชาติ (คสช.) ที่เข้ามาบริหารประเทศบนวิสัยทัศน์ที่ว่า “มั่นคง มั่งคั่ง และยั่งยืน” ที่มีภารกิจสำคัญในการขับเคลื่อนปฏิรูปประเทศด้านต่าง ๆ เพื่อปรับแก้ จัดระบบ ปรับทิศทาง และสร้างหนทางพัฒนาประเทศให้เจริญ สามารถรับมือกับโอกาสและภัยคุกคามแบบใหม่ ๆ ที่เปลี่ยนแปลงอย่างเร็ว รุนแรงในศตวรรษที่ 21 ได้

เพื่อให้เกิดผลจริงต้องมีการพัฒนาวิทยาการ ความคิดสร้างสรรค์ นวัตกรรม วิทยาศาสตร์ เทคโนโลยี และการวิจัยและพัฒนา แล้วต่อยอดในกลุ่มเทคโนโลยีและอุตสาหกรรมเป้าหมาย ดังนี้

1. กลุ่มอาหาร เกษตร และเทคโนโลยีชีวภาพ เช่น สร้างเส้นทางธุรกิจใหม่ (New Startups) ด้านเทคโนโลยีการเกษตร เทคโนโลยีอาหาร เป็นต้น

2. กลุ่มสาธารณสุข สุขภาพ และเทคโนโลยีทางการแพทย์ เช่น พัฒนาเทคโนโลยีสุขภาพ เทคโนโลยีการแพทย์ สปา เป็นต้น

3. กลุ่มเครื่องมือ อุปกรณ์อัจฉริยะ หุ่นยนต์ และระบบเครื่องกลที่ใช้ระบบอิเล็กทรอนิกส์ควบคุม เช่น เทคโนโลยีหุ่นยนต์ เป็นต้น

4. กลุ่มดิจิตอล เทคโนโลยีอินเตอร์เน็ตที่เชื่อมต่อและบังคับอุปกรณ์ต่าง ๆ ปัญญาประดิษฐ์และเทคโนโลยี สมองกลฝังตัว เช่น เทคโนโลยีด้านการเงิน อุปกรณ์เชื่อมต่อออนไลน์โดยไม่ต้องใช้คน เทคโนโลยีการศึกษา อี–มาร์เก็ตเพลส อี–คอมเมิร์ซ เป็นต้น

5. กลุ่มอุตสาหกรรมสร้างสรรค์ วัฒนธรรม และบริการที่มีมูลค่าสูง เช่น เทคโนโลยีการออกแบบธุรกิจ ไลฟ์สไตล์ เทคโนโลยีการท่องเที่ยว การเพิ่มประสิทธิภาพการบริการ เป็นต้น

โดยผู้นำเสนองานประชุมวิชาการให้ครั้งนี้ มีงานวิจัยที่หลากหลาย และนำสถิติเข้ามาใช้ในการวิเคราะห์ข้อมูลแสดงให้เห็นถึงความสำคัญของสาขาวิชาสถิติได้เป็นอย่างดี

การนำไปใช้ประโยชน์จากการเข้าร่วมนำเสนอผลงานและประชุมวิชาการ

ข้าพเจ้าคาดว่าจะนำความรู้ที่ได้รับมาประยุกต์ใช้กับการสร้างงานวิจัยใหม่ ๆ ที่สอดคล้องกับสถานการณ์ในปัจจุบัน และนำมาประยุกต์ใช้ในการสอนนักศึกษา สาขาวิชาสถิติ เพื่อยกตัวอย่างให้เห็นถึงประโยชน์ในการเรียนสถิติ และการประยุกต์ใช้งานกับศาสตร์อื่น ๆ โดยเฉพาะแนวคิด Thailand 4.0 ที่มีจุดเริ่มต้นจากการวิเคราะห์พัฒนาการของระบบเศรษฐกิจของประเทศไทย ที่มีพื้นฐานจากระบบเศรษฐกิจที่พึ่งพาการผลิตและส่งออกสินค้าเกษตรกรรม ในยุค Thailand 1.0 ก่อนจะมีการพัฒนาการผลิต เพื่อยกระดับเศรษฐกิจจากการเป็นประเทศในกลุ่มรายได้ จนกลายเป็นประเทศกลุ่มรายได้ กลาง-สูง

ความคิดเห็นของผู้บังคับบัญชาชั้นต้น (ประธานหลักสูตร/เลขานุการคณะ/หัวหน้างาน)

อ.ดร.ศิริพร สมุทรวชิรวงศ์ ได้เข้าร่วมประชุมวิชาการฯ และได้นำเสนอ

ผลงานวิจัย ซึ่งจากการเข้าร่วมประชุมวิชาการฯ ทำให้ได้รับความรู้

ต่างๆ สามารถนำมาประยุกต์ใช้ในการเรียนการสอน และในงานวิจัยต่อไป

หนึ่งหทัย.

(ผู้ช่วยศาสตราจารย์ ดร.หนึ่งหทัย ชัยอาภร)

14 / มค / 61

ความคิดเห็นของคณบดีคณะวิทยาศาสตร์หรือผู้แทน

....

(..)

............/......................../.................

หมายเหตุ : แบบฟอร์มเป็นรูปแบบเพื่อเสนอการรายงาน เนื้อหาที่อาจไม่เพียงพอสำหรับกรอกข้อมูล สามารถขยายหรือเพิ่มเติมได้ตามความเหมาะสม

Curtin University

UNIKASSEL VERSITÄT

Certificate of Participation

This certificate is awarded to
Siriporn Samutwachirawong

for Paper Titled

A negative binomial-new weighted Lindley distribution for count data and its application to hospitalized patients with diabetes at Ratchaburi hospital, Thailand

presented in

the Session of Pure and Applied Science

The 2nd International Conference of Multidisciplinary Approaches on UN Sustainable Development Goals 2017 (UNSDGs 2017)

on 28th-29th December 2017

Somdej

(Assist. Prof. Somdej Ninlapan)
Acting President of Nakhon Pathom Rajabhat University

A negative binomial-new weighted Lindley distribution for count data and its application to hospitalized patients with diabetes at Ratchaburi hospital, Thailand

Siriporn Samutwachirawong[1,*]

[1] Section of Statistics, Faculty of Science, Maejo University, San Sai, Chiang Mai 50290, Thailand

Abstract

The Poisson distribution plays a central role in count data analysis. The most important characteristic of the Poisson distribution is its mean and variance must be equal. In practice, the count data often exhibit the overdispersion, which is the variance is greater than the mean. The negative binomial distribution offers a remedy to this problem. Moreover, the mixed negative binomial distribution is an alternative to count data with overdispersion. The objectives of research are to proposed new distributions for count data, namely, the negative binomial-new weighted Lindley distribution, to derive parameter estimation of the proposed distribution by using the maximum likelihood estimation and to compare efficiencies of the proposed distribution with other distributions. This research found that the negative binomial-new weighted Lindley distribution, obtained by mixing the negative binomial distribution with the new weighted Lindley distribution is another mixed negative binomial distribution that provided an appropriate fit for count data with overdispersion. Some characteristics of the proposed distribution, such as mean, variance, the coefficient of skewness and kurtosis are derived. Finally, application of the proposed distribution for the number of hospitalized patients with diabetes at Ratchaburi hospital, Thailand is presented.

Keywords: overdispersion, count data analysis, mixed negative binomial distribution, new weighted Lindley distribution

1. Introduction

The count data are non-negative integers, such as the number of hospitalizations, the number of H5N1 outbreaks reported in Thailand, the number of doctors visited per year, and the number of trips taken per year. The most popular method to model count data is the Poisson distribution. The Poisson distribution was introduced by S. D. Poisson. The Poisson distribution should have been named after Bortkiewicz in 1898, the data give the number of soldiers killed by being kicked by a horse each year in each of 14 cavalry corps over a 20-year period. Bortkiewicz showed that those numbers follow the Poisson distribution, mean is equal to variance [1]. Indeed, equidispersion is referred to equality of mean and variance; equidispersion will play a crucial role in further discussion. Besides, there are other categories of dispersion which are overdisperssion, variance is greater than the mean and underdispersion, variance is smaller than the mean [2].

Regarding to an importance of dispersion, Greenwood and Yule [3] derived a mixture of Poisson distribution with the mean distributed as a gamma distribution, called the negative binomial (NB) distribution. The NB distribution has become increasingly popular because of a more flexible alternative to the Poisson distribution; especially, when it is doubtful whether the strict requirements, for the Poisson distribution will be satisfied [4]. In addition, the NB distribution can be mixed with other distributions to be an alternative distribution of overdispersed data. It has been shown that the mixed negative binomial distribution provided a better fit to count data compared to the NB distribution, such as the NB-Pareto [5], the NB-inverse Gaussian [6], the NB-Lindley [7], the NB-beta exponential [8], the NB-generalized exponential [9], the NB-crack distribution [10] and the NB-Erlang distribution [11].

The Lindley distribution has been generalized by many researchers in recent years. Ghitany and *et al.* [12] investigated Lindley distribution in the context of reliability analysis. Subsequently, a weighted Lindley (WL) distribution is proposed for modelling survival data. The WL distribution has the property that the hazard rate (mean residual life) function exhibits bathtub (upside-down bathtub) or increasing (decreasing) shapes. Applications of the WL distribution to real survival data are presented. The new weighted Lindley (NWL) distribution was recently introduced by Asgharzadeha and *et al.* [13]. It is a two parameter continuous distribution used in lifetime data. A random variable X follows the NWL distribution with parameters $\alpha > 0$ and $\beta > 0$. Its the probability density function (pdf) is

*Corresponding author; e-mail: siriporn_sm@mju.ac.th

$$f(x)=\frac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha)+\alpha(2+\alpha)}(1+x)(1-\exp(-\alpha\beta x))\exp(-\beta x), \text{ for } x>0.$$

Let $X \sim \text{NWL}(\alpha,\beta)$, then its moment generating function (mgf) of X is given by

$$M_X(t)=\frac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha)+\alpha(2+\alpha)}\left\{\frac{\beta-t+1}{(\beta-t)^2}-\frac{\beta(1+\alpha)-t+1}{\beta(1+\alpha)-t^2}\right\}.$$

Some plots of the NWL pdf with some specified values of α and β are shown in Figure 1.

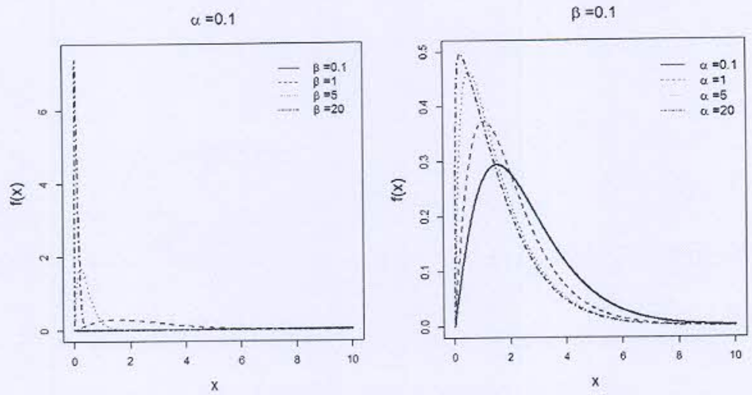

Figure 1 Some pdf plots of the NWL distribution

In this research, a discrete distribution which represents as an alternative distribution for overdispersed count data, namely the negative binomial- new weighted Lindley (NB-NWL) distribution has been developed. The NB-NWL distribution is a mixture of the NB and NWL distributions. Some of the characteristics of the proposed distribution can be studied through factorial moments, e.g., mean, variance, skewness, and kurtosis. The parameters of the proposed distributions are estimated by using the maximum likelihood estimation (MLE). A real data set is the number of hospitalized patients with diabetes at Ratchaburi hospital, Thailand. The proposed distribution is compared performance with Poisson and NB distributions.

2. Research objectives

2.1 To propose the negative binomial-new weighted Lindley distribution for overdispersed count data.
2.2 To derive the parameter estimation of the proposed distributions by using the MLE method
2.3 To compare efficiencies of the proposed distribution with other distributions for count data analysis.

3. Materials and methods

3.1 The materials are as follows

1) High performance personal computer for running the coded program.
2) R language version 3.4.1 [14] for using in application study.

3.2 The methods are as follows

1) The pmf and some properties of the NB-NWL distribution will be investigated.
2) The estimated of parameters of the NB-NWL distribution will be derived by the MLE method.
3) Random variate generation of the NB-NWL distribution is derived.
4) Application of the NB-NWL distribution to real data set has been studied by comparing to the Poisson and NB distributions using the Kolmogorov-Smirnov (K-S) from the dgof package [15] and estimated log-likelihood (LL).

4. Results and discussion

In this section presents results of the research. It provides the probability mass function (pmf) of the proposed distribution. Moreover, some characteristics, some plots of the pmf with various values of parameters, parameter estimation, random variate generation, and application of the proposed distribution to real dataset are included in each part.

4.1 The negative binomial- new weighted Lindley distribution

Definition 4.1 Let $X|\lambda$ be a random variable following a NB distribution with parameters r and $p=\exp(-\lambda)$, $X|\lambda \sim \text{NB}(r, p=\exp(-\lambda))$. If λ is distributed as the NWL distribution with positive parameters α and β, denoted by $\lambda \sim \text{NWL}(\alpha,\beta))$, then X is called a NB-NWL random variable.

Theorem 4.1 Let $X \sim \text{NB-NWL}(r,\alpha,\beta)$. The pmf of X is given by

$$f(x;r,\alpha,\beta)=\binom{r+x-1}{x}\sum_{j=0}^{x}\binom{x}{j}(-1)^j\frac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha)+\alpha(2+\alpha)}\left\{\frac{\beta+r+j+1}{(\beta+r+j)^2}-\frac{\beta(1+\alpha)+r+j+1}{\left[\beta(1+\alpha)+r+j\right]^2}\right\},\quad x=0,1,2,\ldots$$

where α and $\beta>0$.

Proof. If $X|\lambda \sim \text{NB}(r, p=\exp(-\lambda))$ and $\lambda \sim \text{NWL}(\alpha,\beta))$, then the pmf of X can be obtained by

$$f(x)=\int_0^\infty f_1(x|\lambda)g(\lambda;\alpha,\beta)d\lambda, \text{ where } f_1(x|\lambda) \text{ is express as}$$

$$\begin{aligned} f_1(x|\lambda) &= \binom{r+x-1}{x}\exp(-\lambda r)(1-\exp(-\lambda))^x, \\ &= \binom{r+x-1}{x}\sum_{j=0}^{x}\binom{x}{j}(-1)^j\exp(-\lambda(r+j)). \end{aligned}$$

By substituting $f_1(x|\lambda)$ into $f(x)=\int_0^\infty f_1(x|\lambda)g(\lambda;\alpha,\beta)d\lambda$, we obtain

$$\begin{aligned} f(x) &= \binom{r+x-1}{x}\sum_{j=0}^{x}\binom{x}{j}(-1)^j\left(\int_0^\infty \exp(-\lambda(r+j))g(\lambda;\alpha,\beta)d\lambda\right), \\ &= \binom{r+x-1}{x}\sum_{j=0}^{x}\binom{x}{j}(-1)^j M_\lambda(-(r+j)). \end{aligned}$$

Substituting the mgf of the NWL distribution in the equation above, the pmf of the NB-NWL(r,α,β) is given as

$$f(x;r,\alpha,\beta)=\binom{r+x-1}{x}\sum_{j=0}^{x}\binom{x}{j}(-1)^j\frac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha)+\alpha(2+\alpha)}\left\{\frac{\beta+r+j+1}{(\beta+r+j)^2}-\frac{\beta(1+\alpha)+r+j+1}{\left[\beta(1+\alpha)+r+j\right]^2}\right\}.$$

Figure 2 displays the NB-NWL pmf plots with some specified parameter values of α and β.

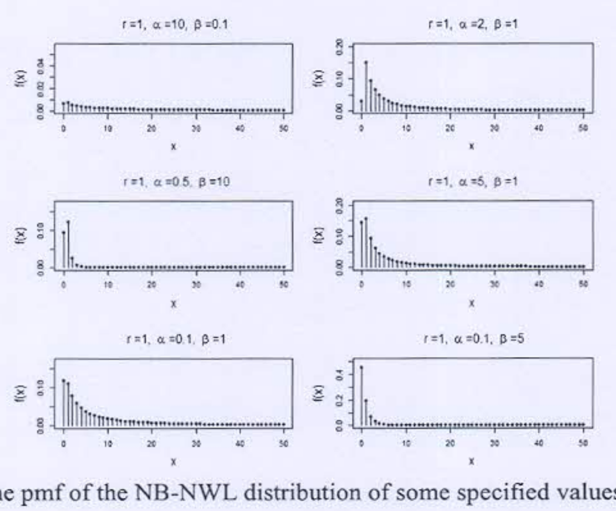

Figure 2 The pmf of the NB-NWL distribution of some specified values of α and β

4.2 Characteristics of the NB-NWL distribution

Some characteristics of the NB-NWL distribution will be discussed as follows. The factorial moment of the NB-NWL distribution is introduced. Some of the most important structures and characteristics of the NB-NWL distribution can be studied through factorial moments.

Theorem 4.2 If $X \sim$ NB-NWL(r,α,β), the factorial moment of order a of X is

$$\mu_a'(X)=\frac{\Gamma(r+a)}{\Gamma(r)}\sum_{j=0}^{a}\binom{a}{j}(-1)^j\frac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha)+\alpha(2+\alpha)}\left\{\frac{\beta-a+j+1}{(\beta-a+j)^2}-\frac{\beta(1+\alpha)-a+j+1}{[\beta(1+\alpha)-a+j]^2}\right\},\quad x=0,1,2,\ldots$$

for α and $\beta>0$.

Proof. G´omez and *et al.* [6] showed that the factorial moment of order a of mixed NB distribution can be expressed in the terms of elementary function by

$$\begin{aligned}\mu_a'(X)&=E_\lambda\left(\frac{\Gamma(r+a)}{\Gamma(r)}\frac{(1-\exp(-\lambda))^a}{\exp(-\lambda a)}\right)\\&=\frac{\Gamma(r+a)}{\Gamma(r)}E_\lambda(\exp(\lambda)-1)^a.\end{aligned}$$

Using the binomial expansion of $(\exp(\lambda)-1)^a$, then $\mu_a'(X)$ can be written as

$$\begin{aligned}\mu_a'(X)&=\frac{\Gamma(r+a)}{\Gamma(r)}\sum_{j=0}^{a}\binom{a}{j}(-1)^jE_\lambda(\exp(\lambda(a-j))),\\&=\frac{\Gamma(r+a)}{\Gamma(r)}\sum_{j=0}^{a}\binom{a}{j}(-1)^jM_\lambda(a-j).\end{aligned}$$

From the mgf of the NWL distribution with $t=a-j$, the $\mu_a'(X)$ is finally given as

$$\mu_a'(X)=\frac{\Gamma(r+a)}{\Gamma(r)}\sum_{j=0}^{a}\binom{a}{j}(-1)^j\frac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha)+\alpha(2+\alpha)}\left\{\frac{\beta-a+j+1}{(\beta-a+j)^2}-\frac{\beta(1+\alpha)-a+j+1}{[\beta(1+\alpha)-a+j]^2}\right\}.$$

Definition 4.2 Let $X \sim$ NB-NWL(r,α,β), some properties of X are as follows

1) The first four moments about zero of X are

$$\begin{aligned}E(X)&=r\ (\phi_1-1)\ ,\\E(X^2)&=r(r+1)\phi_2-r(2r+1)\phi_1+r2,\\E(X^3)&=r(r+1)(r+2)\phi_3-3r(r+1)2\phi_2+r(3r2+3r+1)\phi_1-r3,\\E(X^4)&=r(r+1)(r+2)(r+3)\phi_4-2r(r+1)(2r2+7r+6)\phi_3\\&\quad+r(r+1)(6r2+12r+7)\phi_2-r(2r+1)(2r2+2r+1)\phi_1+r4.\end{aligned}$$

2) The mean and variance of X respectively, are

$$\begin{aligned}E(X)&=r\ (\phi_1-1),\\Var(X)&=r(r+1)\phi_2-r\ (1+r\phi_1)\phi_1.\end{aligned}$$

3) The coefficient of skewness and kurtosis of X respectively, are

$$\begin{aligned}Skewness(X)&=\big[r(r+1)(r+2)\phi_3-3r(r+1)(r\phi_1+1)\phi_2\\&\quad+r(2r\phi_1+1)(r\phi_1+1)\phi\big]/[Var(X)]^{3/2},\end{aligned}$$

$$
\begin{aligned}
Kurtosis(X) &= \big[r(r+1)(r+2)(r+3)\phi_4 \text{ - } 2r(r+1)(r+2)(2\phi_1 1+3)\phi_3 \\
&\quad + r(r+1)(6r^2\phi_1^2 + 12r\phi_1 + 7)\phi_2 \\
&\quad \text{-} r(r\phi_1 + 1)(3r^2\phi_1^2 + 3r\phi_1 + 1)\phi\big], /[Var(X)]^2,
\end{aligned}
$$

where $\phi_k = \dfrac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha)+\alpha(2+\alpha)}\left\{\dfrac{\beta-k+1}{(\beta-k)^2} - \dfrac{\beta(1+\alpha)-k+1}{[\beta(1+\alpha)-k]^2}\right\}.$

4.3 Parameter estimation

The parameter estimation for the NB-NWL distribution via the MLE method will be discussed. If $X \sim \text{NB-NWL}(r,\alpha,\beta)$, then the likelihood function of parameters r, α and β can be written as

$$
L(r,\alpha,\beta;x_i) = \prod_{i=1}^{n}\left[\binom{r+x_i-1}{x_i}\sum_{j=0}^{x_i}\binom{x_i}{j}(-1)^j \frac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha)+\alpha(2+\alpha)}\left\{\frac{\beta+r+j+1}{(\beta+r+j)^2} - \frac{\beta(1+\alpha)+r+j+1}{[\beta(1+\alpha)+r+j]^2}\right\}\right],
$$

consequently, the associated log-likelihood function is

$$
\begin{aligned}
\mathcal{L} &= \log L(r,\alpha,\beta;x_i) \\
&= \sum_{i=1}^{n}\left[\log\Gamma(r+x_i) - \log\Gamma(r) - \log\Gamma(x_i+1)\right] \\
&\quad + \sum_{j=0}^{x_i}\log\left[\sum_{j=0}^{x_i}\binom{x_i}{j}(-1)^j \frac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha)+\alpha(2+\alpha)}\left\{\frac{\beta+r+j+1}{(\beta+r+j)^2} - \frac{\beta(1+\alpha)+r+j+1}{[\beta(1+\alpha)+r+j]^2}\right\}\right].
\end{aligned}
$$

The first order conditions for finding the optimal values of the parameters obtained by differentiating the associated log-likelihood function with respect to r, α and β give rise to the following differential equations

$$
\frac{\partial\mathcal{L}}{\partial r} = \sum_{i=1}^{n}\psi(r-x_i) - \text{n}\,\psi(r) + \sum_{i=1}^{n}\left(\frac{\sum_{j=0}^{x_i}\binom{x_i}{j}(-1)^j \frac{\partial}{\partial r}B}{\sum_{j=0}^{x_i}\binom{x_i}{j}(-1)^j B}\right),
$$

$$
\frac{\partial\mathcal{L}}{\partial c} = \sum_{i=1}^{n}\left(\frac{\sum_{j=0}^{x_i}\binom{x_i}{j}(-1)^j \frac{\partial}{\partial c}B}{\sum_{j=0}^{x_i}\binom{x_i}{j}(-1)^j B}\right),
$$

$$
\frac{\partial\mathcal{L}}{\partial k} = \sum_{i=1}^{n}\left(\frac{\sum_{j=0}^{x_i}\binom{x_i}{j}(-1)^j \frac{\partial}{\partial c}B}{\sum_{j=0}^{x_i}\binom{x_i}{j}(-1)^j B}\right),
$$

where $\psi(g) = \dfrac{\Gamma'(g)}{\Gamma(g)}$ is the digamma function and $B = \dfrac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha)+\alpha(2+\alpha)}\left\{\dfrac{\beta+r+j+1}{(\beta+r+j)^2} - \dfrac{\beta(1+\alpha)+r+j+1}{[\beta(1+\alpha)+r+j]^2}\right\}.$

The maximum likelihood estimates $\hat{r}$, $\hat{\alpha}$ and $\hat{\beta}$ for the parameters r, α and β respectively, are taken by solving iteratively differential equations to zero. These differential equations are not in closed form, a numerical method can be employed to obtain the expectations of them. The MLE solution of $\hat{r}$, $\hat{\alpha}$ and $\hat{\beta}$ can be obtained by solving the resulting equations simultaneously using optim function in R language.

4.4 Random variate generation of NB-NWL distribution

A random variate X from the NB-NWL(r,α,β), has been generated from following algorithm.

1) Generate U from the uniform distribution, U(0,1),

2) Set $\lambda = -\frac{1}{\beta}\log(1-\mathrm{U}^{1/\alpha})$ from the NWL distribution, $\mathrm{NWL}(\alpha,\beta)$),

3) Generate X from the $\mathrm{NB}(r, p = \exp(-\lambda))$ distribution.

4.5 Applications study of NB-NWL distribution

We illustrated the NB-NWL, NB and Poisson distributions by applying to real data set. The information of hospitalized 1459 cases from ages 55- 90 years old in period 2012 to 2016 of at Ratchaburi hospital, Thailand was used as source of real data. The decision to admit to the hospital should be made if the patient has not responded well to treatment within no more than 4 hours after presentation. A real data set was used to estimate the parameters for the NB-NWL, NB and Poisson distributions using the MLE method. The mean and variance are 4.5456 and 29.6349, respectively. The index of dispersion is 6.5195, indicating that there is high percentage of zeros and the variance is greater than the mean.

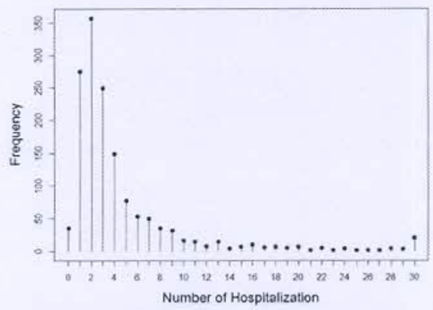

Figure 3 The number of hospitalized patients with diabetes at Ratchaburi hospital, Thailand

The log-likelihood values and the p-values of K-S test for the discrete goodness of fit test are summarized in Table 1. The expected frequencies of the NB-NWL distribution are close to the observed frequencies, the values of K-S test of NB-NWL distribution is smaller than the values of the K-S test of the Poisson and NB distributions

5. Conclusions

This work proposes the new mixed negative binomial distribution, which is called the negative binomial-new weighted Lindley distribution. In particular, some of the most important characteristics of the distribution can be studied through factorial moments, e.g., mean, variance, skewness, and kurtosis. In the application of the NB-NWL distribution, we compare the accuracy of the proposed distribution with the Poisson and NB distributions. The usefulness of the NB-NWL distribution is illustrated by the number of hospitalized patients with diabetes at Ratchaburi hospital, Thailand. We use the log-likelihood and p-values of the K-S test for the goodness of fit for model selection purpose. Finally, the result of this study show that the NB-NWL distribution provides a better fit compared to the Poisson and NB distributions. Obviously, the NB-NWL distribution is an alternative distribution to the other for count data.

Table 1 Observed and expected frequencies for number of hospitalized patients with diabetes

No. of hospitalization	No. of cases	Expected value by fitting distribution		
		Poisson	NB	NB-NWL
0	35	261.2574	73.5711	34.3315
1	275	449.3630	155.7058	189.0602
2	357	386.4520	205.2514	171.5407
3	249	221.5659	215.9518	147.2508
4	149	95.2733	198.4813	124.3442
5	77	32.7740	166.5815	104.7453
6	53	9.3952	130.9441	88.5186
7	50	2.3085	97.9537	75.2206
8	35	0.4963	70.4826	64.3310
9	32	0.0949	49.1530	55.3816
10	16	0.0163	33.4063	47.9862
11	15	0.0026	22.2195	41.8357
12	8	0.0004	14.5100	36.6865
13	15	0.0000	9.3271	32.3472
14	4	0.0000	5.9136	28.6669
15	7	0.0000	3.7045	25.5265
16	10	0.0000	2.2959	22.8310
17	6	0.0000	1.4095	20.5048
18	7	0.0000	0.8579	18.4868
19	5	0.0000	0.5182	16.7277
20	7	0.0000	0.3108	15.1872
21	2	0.0000	0.1853	13.8323
22	5	0.0000	0.1098	12.6358
23	2	0.0000	0.0647	11.5752
24	4	0.0000	0.0380	10.6315
25	2	0.0000	0.0222	9.7890
26	2	0.0000	0.0129	9.0345
27	2	0.0000	0.0075	8.3566
28	4	0.0000	0.0043	7.7458
29	3	0.0000	0.0025	7.1940
30+	21	0.0000	0.0014	6.6942
Total	1459			
Parameter estimates		$\hat{\lambda}=1.72$	$\hat{r}=4.07$	$\hat{r}=4.15$
			$\hat{p}=0.48$	$\hat{\alpha}=0.52$
				$\hat{\beta}=2.01$
log-likelihood		-1140.449	-1014.642	-825.985
K-S test		0.319	0.013	0.018
p-value		<0.001	0.086	0.572

Acknowledgements

First, I thankful to Section of Statistics, Faculty of Science, Maejo University. I would like to express a special thanks Dr.Krisana Lanumteang for suggestions and help with the real data set.

References

[1] Haight F. **Handbook of the Poisson distribution**. John Wiley and Sons, New York. 1967.

[2] Rainer W. **Econometric analysis of count data**. Library of congress control, New York. 2008.

[3] Greenwood M, Yule G. An enquiry into the nature of frequency distributions representative of multiple happenings with particular reference of multiple attacks of disease or of repeated accidents. **Journal of the royal statistical society**. 1920; **83**: 25-279.

[4] Johnson N, Kemp A, Kotz S. **Univariate discrete distributions**. John Wiley and Sons, New York. 2005.

[5] Klugman S, Panjer H, Willmot G. **Loss models: from data to decisions**. 3rd ed. John Wiley and Sons. 2008.

[6] G´omez D, Sarabia J, Calder´in E. Univariate and multivariate versions of the negative binomial-inverse Gaussian distributions with applications. **Insurance: mathematics and economics**. 2008; **42** (1): 39-49.

[7] Zamani H, Ismail N. Negative binomial Lindley distribution and its application. **Journal of mathematics and statistics**. 2010; **6** (1): 4-9.

[8] Pudprommarat C, Bodhisuwan W, Zeephongsekul P. A new mixed negative binomial distribution. **Journal of applied sciences**. 2011; **12** (17): 1853-1858.

[9] Aryuyuen S, Bodhisuwan W. The negative binomial-generalized exponential (NB-GE) distribution. **Applied mathematical sciences**. 2013; **7** (22): 1093-1105.

[10] Saengthong P, Bodhisuwan W. Negative binomial-crack (NB-CR) distribution. **International journal of pure and applied mathematics**. 2013; 84(3): 213-230.

[11] Kongrod S, Bodhisuwan W, Payakkapong P. Negative Binomial-Erlang distribution with application. **International journal of pure and applied mathematics**. 2014; 92 (3): 389-401.

[12] Ghitany ME, Atieh B, Nadarajah S. Lindley distribution and its application. Mathematics and computers in simulation. 2008; **78**: 493-506.

[13] Asgharzadeha A, Hassan SB, Nadarajah S, Sharafia F. A new weighted Lindley distribution with application. **Brazilian journal of probability and statistics**. 2016; **30**(1): 1-27.

[14] R Core Team. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing. Vienna, Austria. 2015.

[15] Arnold BT, Emerson WJ. Nonparametric goodness-of-fit tests for discrete null distributions. **The R journal**. 2011; **3**: 34-39.