International Journal of Scientific and Technological Research
ISSN 2422-8702 (Online)
Vol 4, No.10, 2018

Chi-Square Goodness of Fit Test for Bivariate Distributions in Polygonal Area

Orhan Kesemen (Corresponding author)
Department of Statistics and Computer Sciences,
Karadeniz Technical University, 61080 Trabzon, Turkey
E-mail: okesemen@gmail.com

Bugra Kaan Tiryaki
Department of Statistics and Computer Sciences,
Karadeniz Technical University, 61080 Trabzon, Turkey
E-mail: bugrakaantiryaki@gmail.com

Tuncay Uluyurt
Arhavi Vocational School, Artvin Coruh University,
08200, Artvin, Turkey
E-mail: tuncayuluyurt@hotmail.com

Abstract
An adaptation of the chi-square goodness test for the bivariate data observed in an arbitrary area is given in this study. An arbitrary area defined in a two-dimensional plane is defined by the polygonal area approach. The inside of polygons whose dominant points are determined are triangulated with the grid points of equal distance. The Delaunay algorithm was preferred in the triangulation process. Triangular areas are used to determine the classes in the polygonal area and calculate the observed and expected frequencies in these classes for the Chi-square goodness of fit test.

Keywords: Bivariate distributions, chi-square goodness of fit test, polygonal distributions.

Introduction

Whether it is a set of random numbers generated on the computer or the real-life observation values, all tests which are used these values to investigate to come from a given distribution are called goodness of fit tests. Goodness of fit tests are one of the most fundamental issues in the application of inferential statistical methods. (Massey Jr, 1951; Anderson & Darling, 1954; D’Agostino & Stephens, 1986; Mark & Turin, 2011). The binned -based Pearson Chi-square test, which is generally used for discrete distributions, is also widely used in continuous distributions. The Chi-square test measures the discrepancy between the expected frequency and the observed frequency in a class. The goodness of fit test is generally developed for univariate distributions. On the other hand, there are studies in the literature that are used for multivariate normality tests (Krishnaiah, 1980; Samuel & Johnson, 1983; D’Agostino & Stephens, 1986).

The pollution rate or crime rate in a city, the distribution of the frequencies of earthquakes in a country, the density of a particular tree species in a forest, propagation density of pests in an area, traffic density in a given region, location and frequency of epidemic disease in a country can be given as examples in the application areas of bivariate probability density functions which is defined in a bounded polygonal area (Kesemen, Tiryaki, & Tuncay, 2019). Also, the interested area has many parts of density functions within the rectangular area. In this regard, there are many cities that are physically or politically divided can be as an example (e.g. Belfast, Beirut, Jerusalem, Mostar, and Nicosia). In particular, data in politically divided cities begin to differ in the course of time. In this case, it may not be possible to evaluate the whole city as a region. In each of these examples, statistical tests may be required based on the probability density function whose definition ranges (bounds) are determined entirely from an arbitrary area.

In this study, chi-square goodness of fit test is proposed for the bivariate distribution functions bounded by an arbitrary area.

Chi-Square Goodness of Fit Test

The Chi-square goodness of fit test was first put forward by Pearson in 1900 (Gibbons & Chakraborti, 2011; Cochran, 1952) and keep up to date. Being simple and knowing the distribution of its statistics has made it a preferred classic method. On the other hand, the development of more sensitive techniques which are based on especially experimental distribution function have occurred in later years (D'Agostino & Stephens, 1986).

Chi-Square Goodness of Fit Test for Univariate Distributions

One of the most common methods based on binned for goodness of fit tests is the Chi-square test. The chi-square test is a useful technique used to test whether the obtained observation data fit to a selected distribution (Cochran, 1952). This technique is used in particular to verify the normality assumptions of many tests as a pretest (Kesemen & Tiryaki, 2018).

Being high of the difference between the observed frequencies and the expected frequencies shows lower fitting and, being low of it shows higher fitting. It is impossible to get a better result if some of the differences are positive and some of them are negative when a total error is requested by summing the differences in order to bring a total criterion to the differences to be found. In this case, when the square of the differences is summed, the following equation is obtained.

$$SSE = \sum_{j}^{m} \left(O_j - E_j\right)^2 \tag{1}$$

The number of class is represented by m. The observed frequency of j^{th} class is indicated by O_j and the expected frequency of it is indicated by E_j. The square of each difference obtained from the above equation appears to be related to the expected frequency. In other words, being higher expected frequency results in higher error rates. Elimination of this relationship makes it difficult to evaluate different volumes of data at the same scale. Each square error is normalized by dividing the expected frequency to solve the problem (2).

$$\hat{\chi}^2 = \sum_{j}^{m} \frac{\left(O_j - E_j\right)^2}{E_j} \tag{2}$$

The total value in Equation (2) is expressed as the calculated chi-square value or the test statistic. This value must be compared with a criterion to have a meaning of the obtained $\hat{\chi}^2$ value. The distribution of $\hat{\chi}^2$ random variable must be known to achieve this criterion. This distribution converges to the Chi-square distribution with ν degrees of freedom. ν is defined as $m-1$.

One of the most important concepts on which the tests are based is the significance test. The basis of this test is based on the established two hypotheses. These hypotheses are: the null hypothesis (H_0) and the alternative hypothesis (H_1).

The form of Hypotheses is given as follows:

H_0: Observation values come from $F_0(x)$ distribution
H_1: Observation values do not come from $F_0(x)$ distribution.

$F_0(x)$ shows that the observation values come from the specified distribution function. A margin of error (α) is determined to find the confidence interval of the test statistic. The calculated limits of the confidence intervals are expressed as the critical value at the same time. The required critical value to test the H_0 hypothesis is determined by the value of $\chi^{2^*}_{\nu,1-\alpha}$ which gives the $(1-\alpha)$ value of the chi-square distribution. This value is found in the chi-square table (O'Connor & Andre, 2011). As a result, if the Equation (3) is correct, H_0 hypothesis is rejected and H_1 hypothesis is accepted.

$$\hat{\chi}^2 > \chi^{2^*}_{\nu,1-\alpha} \tag{3}$$

It is necessary to determine the class number and class range in order to perform chi-square test in continuous functions. Although there are different approaches in determining the number of class (m) in literature, it is often left to the preference of the researcher. The Equation (4) can be used to determine automatically the class width which the researcher determines it generally.

$$\Delta x = \frac{x_n - x_1}{m} \tag{4}$$

x_1 represents the smallest value in the data, and x_n represents the largest value in the data. The limits of class are calculated as follows.

$$\hat{x}_j = x_1 + j.\Delta x, \qquad j = 0,1,2, \dots, m \tag{5}$$

The number of observations (observed frequencies) per class is computed in the Equation (6).

$$O_j = \sum_{k=1}^{n} \begin{cases} 1, & \hat{x}_{j-1} \le x_k < \hat{x}_j \\ 0, & others \end{cases}, \qquad j = 1,2, \dots, m \tag{6}$$

The expected frequencies per class are calculated as in the following equation.

$$E_j = n\left[F_0(\hat{x}_j) - F_0(\hat{x}_{j-1})\right], \qquad j = 1,2, \dots, m \tag{7}$$

$F_0(.)$ represents the theoretical distribution function which is given in the H_0 hypothesis.

Chi-Square Goodness of Fit Tests for Bivariate Distributions

The polygonal definition is the most appropriate approach in defining the observed data in an arbitrary area with a known geometric shape (Harrington, 1983; Kesemen, Tiryaki, & Tuncay, 2019). Calculating the integral should be needed to find the expected frequencies from the probability density function defined in a polygonal area. This changes to an integral calculation that cannot be determine its limit values. In the proposed method, the number of class can be determined by dividing the polygonal area into appropriate triangles (Figure 1) and the expected frequency number in this area can be calculated. The expected frequencies can be calculated by dividing the triangle into two parts according to the peak point. Similarly, the number of observed frequencies falling into triangles can be calculated algorithmically (Kesemen, Tiryaki, & Tuncay, 2019).

Although it is not widely used by univariate case, it can produce solutions in some cases. In a bivariate space, the probability density function forms a surface and the distribution function gives the volume below that surface.

The test statistics is obtained by the sum of the squared error divided into the expected frequencies. The chi-square goodness of fit test in many applications is used to check whether the observed values fit to the univariate distributions or not. However, the observation values obtained in a two-dimensional area should be bounded by a rectangular area to be tested. However, the data obtained in real life are not always in the rectangular area as is known.

The class areas should be determined according to the given number of classes in order to be able to perform chi-square test in bivariate continuous distributions. The number of classes is arbitrarily determined as $(m_x \times m_y)$.

The class widths are determined as following equation.

$$\Delta x = \frac{x_n - x_1}{m_x}; \ \ \Delta y = \frac{y_n - y_1}{m_y} \tag{8}$$

(x_1, y_1) shows the smallest x and y values in the data and (x_n, y_n) shows the largest x and y values in the data. The grid points of the whole region are calculated as in the Equation (9)-(10).

$$\hat{x}_i = x_1 + i.\Delta x, \qquad i = 0,1,2, \dots, m_x \tag{9}$$
$$\hat{y}_j = y_1 + j.\Delta y, \qquad j = 0,1,2, \dots, m_y \tag{10}$$

The number of observations (observed frequencies) per class is computed in the Equation (11).

$$O_{ij} = \sum_{k=1}^{n} \begin{cases} 1, & \hat{x}_{i-1} \leq x_k < \hat{x}_i \ AND \ \hat{y}_{j-1} \leq y_k < \hat{y}_j \\ 0, & others \end{cases}, \quad \begin{pmatrix} i = 1,2,\dots,m_x \\ j = 1,2,\dots,m_y \end{pmatrix} \tag{11}$$

The expected frequencies per class are calculated as in the following equation.

$$E_{ij} = n\left[F_0(\hat{x}_i, \hat{y}_j) - F_0(\hat{x}_{i-1}, \hat{y}_j) - F_0(\hat{x}_i, \hat{y}_{j-1}) + F_0(\hat{x}_{i-1}, \hat{y}_{j-1})\right] \begin{pmatrix} i = 1,2,\dots,m_x \\ j = 1,2,\dots,m_y \end{pmatrix} \tag{12}$$

Chi-Square Goodness of Fit Tests in Polygonal Area

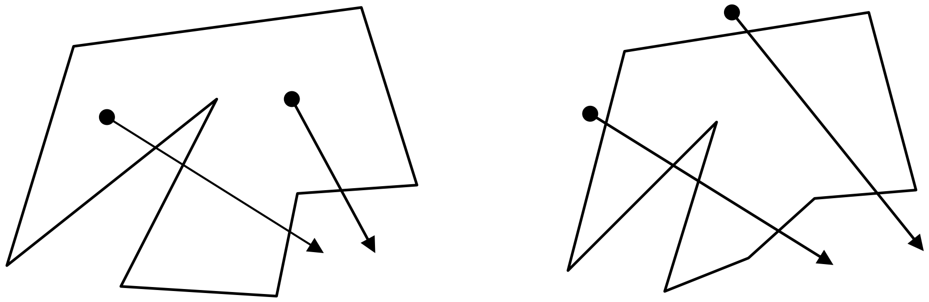
The polygonal definition is the most appropriate approach in defining the observed data in an arbitrary area with a known geometric shape (Kesemen, Tiryaki, & Tuncay, 2019). Calculating the integral from the area should first be needed to find the expected frequencies from the probability density function defined in a polygonal area. This is often impossible. In the proposed method, both the number of class can be determined and the expected frequencies in these areas can be calculated by dividing the polygonal area into appropriate triangles (Figure 7). The analytic integral can be calculated by dividing the triangle into two parts according to its peak points for determining the expected frequencies. If the integral of the probability density function defined in the triangle cannot be taken, numerical methods are applied. Similarly, the number of observed frequencies falling into triangles can be calculated algorithmically (Kesemen, Tiryaki, & Tuncay, 2019).

Polygons

In this study, many studies have been carried out with the help of geometric areas. Also, all calculations are based on polygonal areas. Polygons are widely used in computational geometry. Polygons consist of points and line segments. These line segments are used to combine the points (Devadoss & O'Rourke, 2011). A polygon is defined as a closed plane area which is bounded by a finite number of line segments (Figure 4). Each line segment is defined as the edge and each point is defined as a node. The lines do not interrupt each other in definition of a simple polygon. (Devadoss & O'Rourke, 2011). This study is based on simple polygons.

Point Problem in Polygon

One of the most common problems in computational geometry is whether a given point (R) is within an arbitrary polygon or not. The solution is quite simple when these polygons are a circle or a rectangle. But when it's an irregular polygon, it's not that easy to solve the problem. There are two most common methods on the subject in the literature (Foley, van Dam, Feiner, & Hughes, 1990; Haines, 1994; Harrington, 1983; O'Rourke, 1998; Sedgewick, 1988; Weiler, 1994; Hormann & Agathos, 2001). The first one is the odd-even method. In this method, a line is drawn between the point (R) which is investigated to be in polygon or not and a point (S) which is outside the polygon. This line interrupts some of the edges of the polygon. If the number of interrupted lines is odd, the point (R) is within the polygon, and if it is even, the point (R) is outside the polygon (Figure 1.).



(a) (b)

Figure 1. Point problem in polygon; (a) Selected points inside polygon; (b) Selected points outside polygon.

The second method is the Plus-Minus method. It is assumed that the points of the polygon are arranged in a clockwise direction. If the given point is to the right of all the edges of the polygon, the point is inside the polygon. However, If the point is to the left of any of the edges of the polygon, it is outside the polygon. Similarly, if the points of the polygon are arranged in the counterclockwise direction and the point is to the left of all the polygonal edges, it is not inside the polygon (Haines, 1994; Hormann & Agathos, 2001). In this study, odd-even method is preferred for simplicity.

Triangulation

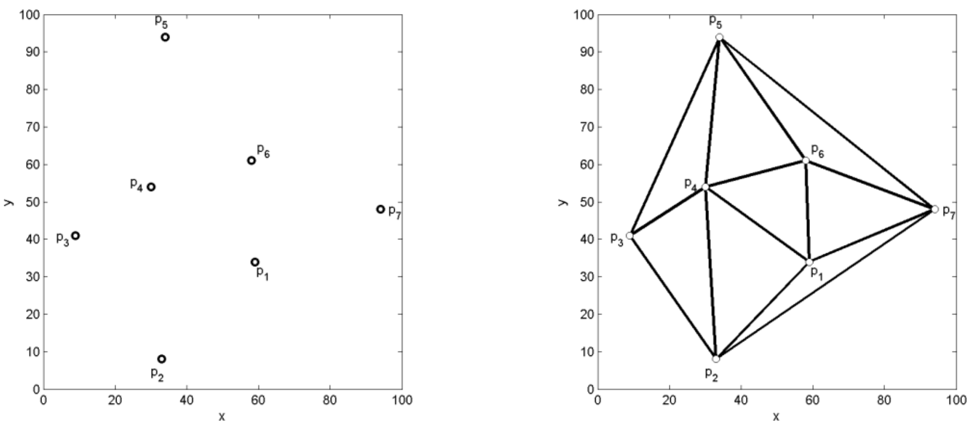
Triangulation is a network which is obtained by combining N points to the each other as a triangle in a two-dimensional plane. When performing this process, lines connecting the points should not interrupt each other. In all possible triangles, triangles are selected to make the highest number of acute triangles. The triangulation algorithm is an algorithm that divides a region into triangular pieces. In this process, a circle passes from all three points and there are no points in any circumcircle (Shewchuk, 2002; Lee & Lin, 1986). The simplest way to do this is to use the triple combination of all points and obtain circles using them.

If there are points inside the other circles, these triple combinations are deleted from the all triple combination list. The last remaining list is taken as a triangulation list. The triangulation process performed by this method is called the Delaunay triangulation algorithm (Worboys & Duckham, 2004; Preparata & Ian, 1985).

The Delaunay triangulation algorithm provides the following characteristics.

1. The algorithm performs the same triangulation each time, regardless of the first (starting) point at the algorithm starts.
2. The obtained triangles have an isogonic characteristic. This prevents the point constitute a triangle with other away points.
3. There are no points within the circumcircles obtained from the algorithm.
4. The line segments which are combined each point represent an edge of the constituted triangles.

There are randomly selected points in a two-dimensional plane in Figure 2(a). When the Delaunay triangulation algorithm applies to these points, an area divided into triangles is obtained as in Figure 2(b).



(a) (b)

Figure 2. (a) Given set of arbitrary points; (b) The triangulation with Delaunay algorithm combining the points.

In Figure 3 is shown that the polygon is fragmented so that it passes a circumferential circle through all three points and it has been shown that there are no dots in all drawn circumferential circles.

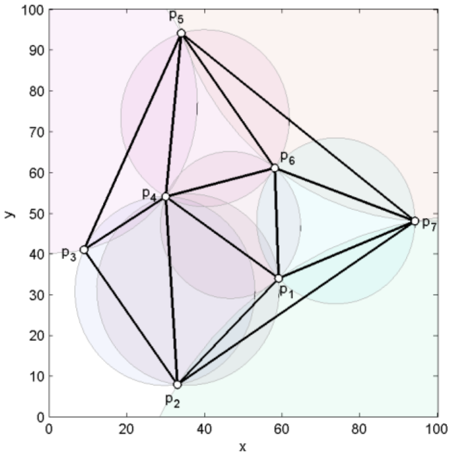
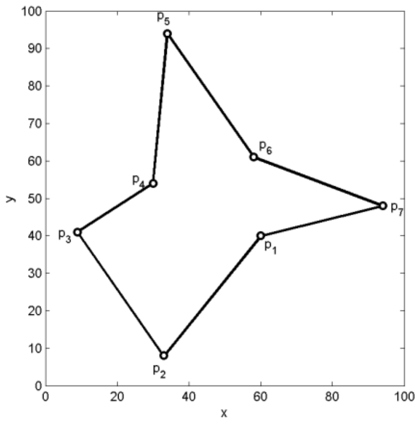


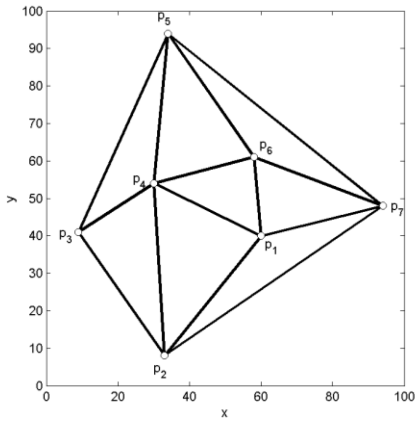
Figure 3. Representation of circumcircle

Triangulation of The Concave Areas

When the convex polygons are triangulated, the structure formed by the triangles covers the polygon. However, when a concave polygon (Figure 4(a)) is triangulated, the structure formed by the triangles is outside of the polygon, and this shows a convex structure (Figure 4(b)). In this case, undesirable triangles occur. In order to solve this problem, first, the centroids of each triangle which is formed by triangulation are found (Figure 5(a)). Triangles whose centroids are in the polygon are saved. Otherwise, they have removed the list (Figure 5(b)).

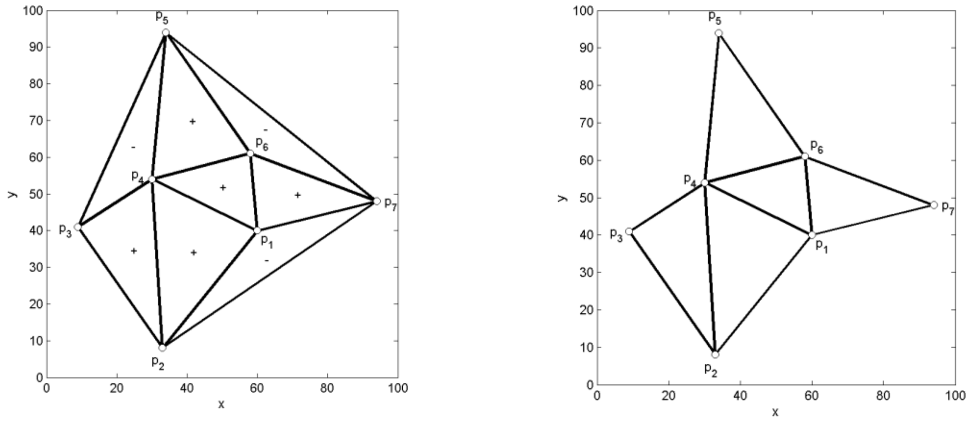


(a)



(b)

Figure 4. Triangulation of the concave polygons; (a) A concave polygon; (b) Triangulation of a concave polygon by Delaunay algorithm

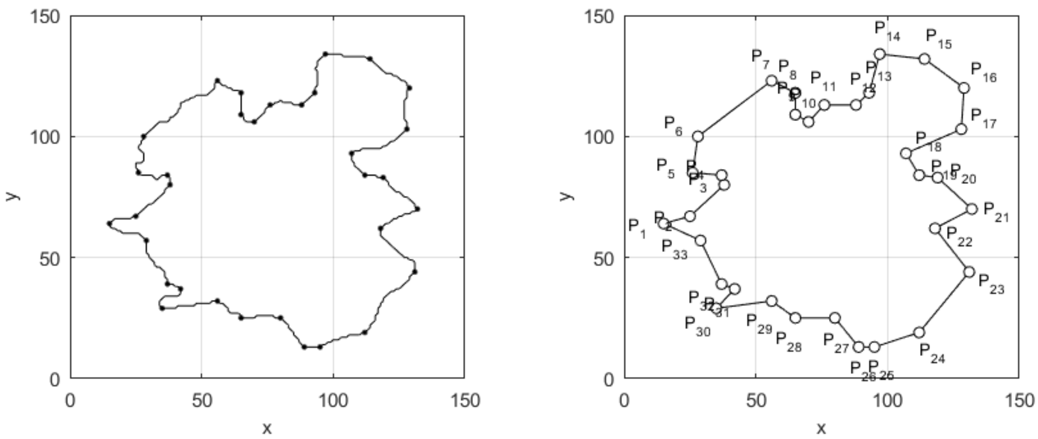


(c) (d)

Figure 5. Removing the undesirable triangles: (a) Finding the triangles out of the polygon; (c) Removing the triangles out of the polygon

Calculation of Distribution Function in Polygonal Area

The most appropriate method for identifying the data observed in an arbitrary area by a known geometric shape is the polygonal definition (Kesemen, Tiryaki, & Tuncay, 2019). The specified arbitrary area (Figure 6(a)) can be defined by the help of a suitable polygon (Figure 6(b)) and then calculations are performed.



(a) (b)

Figure 6. An approximate representation of a two-dimensional arbitrary area (Erzurum) with a polygonal region with 33 points; (a) The arbitrary area; (b) Polygonal area

The integral must be taken to find the expected frequency value by using the probability density value defined in a polygonal area. This is quite difficult. In the proposed method, the polygonal area is divided into appropriate triangles after the dividing into grids (Figure 7). Thus, the number of classes can be determined and the expected number of frequencies in this area can be calculated.

In order to calculate the distribution function in the polygonal area, the polygonal area is divided into triangles by the help of the Delaunay triangulation algorithm. The probability value of the triangle is calculated by taking integral for each triangle in the Equation (13) (Kesemen, Tiryaki, & Tuncay, 2019).

$$P\left((X,Y) \in \Omega_j\right) = \iint_{\Omega_j} f(x,y)dxdy \tag{13}$$

Where, Ω_j is represent the calculated triangular region whose intersection are empty, j is represent the index of the selected triangle. The combination of the Ω_j gives the Ω region. Thus, the sum of the probability values of each triangle gives the probability value of the whole polygon.

$$P\big((X,Y) \in \Omega\big) = \sum_j P\left((X,Y) \in \Omega_j\right) \tag{14}$$

The preferred method for calculating the probability value is analytical calculation. But, in most cases analytical calculation may not be possible. In this situation numerical methods are used (Kesemen, Tiryaki, & Tuncay, 2019).

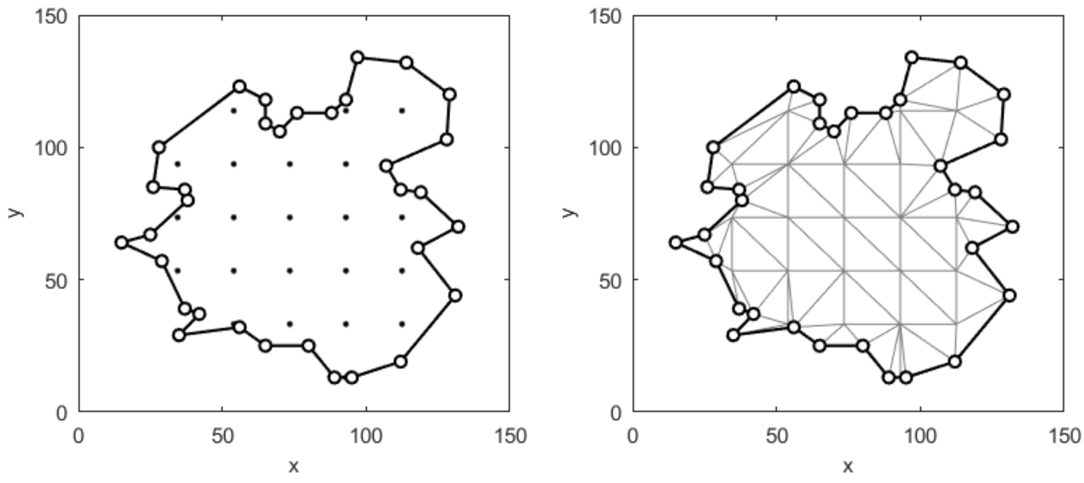


Figure 7. Partition of polygonal area into triangles; (a) Determination of the grid points in polygon; (b) Triangulation of the polygon with the help of the grid points.

Chi-Square Goodness of Fit Test in Polygonal Area

The chi-square test is a method based on the difference between the observed and the expected frequencies in a class. It is necessary to determine the number of classes and the boundaries of the class area for in both the univariate chi-square test as well as the bivariate chi-square test. In the univariate chi-square test, the class region is defined as the region between the two boundary values and in the bivariate chi-square test, the class region is defined as a rectangular area. The class region is defined by the help of the triangles in polygonal applications. The area in which a two-dimensional random variable fall in the triangular or polygonal region is determined by the odd-even method. The chi-square goodness-of-fit test in the polygonal area can be summarized as follows.

Algorithm 1.	Chi-square goodness-of-fit test in the polygonal area
Step1.	Determine the boundaries of the polygonal area, hypothesize the H_0 and determine the defined probability density function in polygonal area and the significance level (α).
Step 2.	Divide the polygonal area by using the grid points and divide the polygonal area into triangles with the help of the Delaunay algorithm.
Step 3.	Calculate the number of samples (observed frequencies, O_j) in obtained each triangle by the odd-even technique in the Section 3.2.
Step 4.	Calculate the expected frequencies E_j of each triangle.
Step 5.	Calculate the test statistic by using expected and observed frequencies obtaining in the Equation (2).
Step 6.	Compare the value of the test statistic to the critical value of $\chi^2_{1-\alpha,m-1}$.
Step 7.	If the test statistic is greater than the critical value, reject the null hypothesis, otherwise accept it.

Experimental Results

In this paper, two examples are performed by the proposed method. These examples are selected from provinces of Turkey. Erzurum and Konya, which are a two of the provinces of Turkey were investigated as a polygonal area.

Example 1. The boundaries of the Erzurum region, which was given in the form of an image, were obtained in a point structure for this study (Figure 6). These points were reduced to 33 points using the point reduction method Figure 7(a) (Douglas & Peucker, 1973). This polygonal area was divided into triangles and was obtained 59 triangles in the Figure 7(b). The probability density function of this polygonal area is given in Equation (15) (Figure 8).

$$f(x,y) = \frac{1}{4174} e^{\frac{(x-70)^2+(y-60)^2}{2000}} \tag{15}$$

In Equation (15), the expected values are shown as $(\mu_x = 70, \mu_y = 60)$, while the variance is shown as $(2\sigma^2 = 2000)$. 4174 is the scale value which is calculated for the volume should be "1" in the polygonal area.

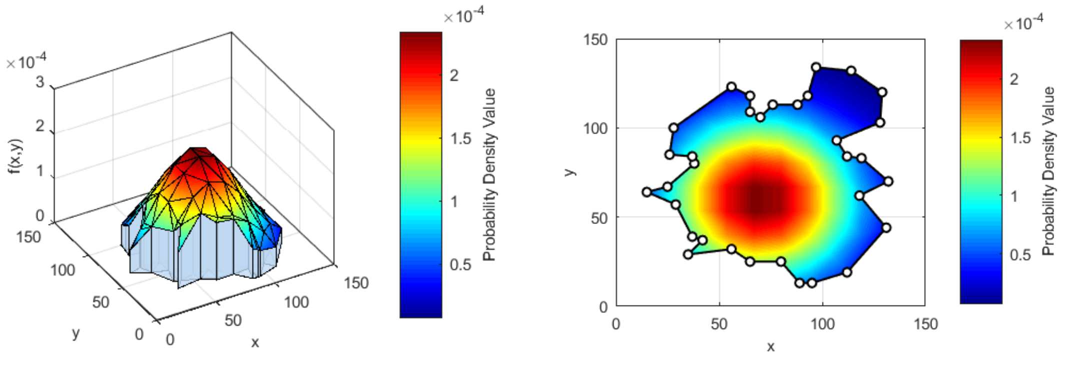


Figure 8. The used probability density function for Example 1; (a) Three-dimensional representation of the function; (b) The two-dimensional encoloured representation of the function.

500 random numbers are generated from the given probability density by rejection method (Figure 9(a)) (Kesemen & Tiryaki, 2018). These random numbers are classified for each triangle in the Figure 9(b) and the observed frequencies of each triangle are calculated in the Figure 10(a).

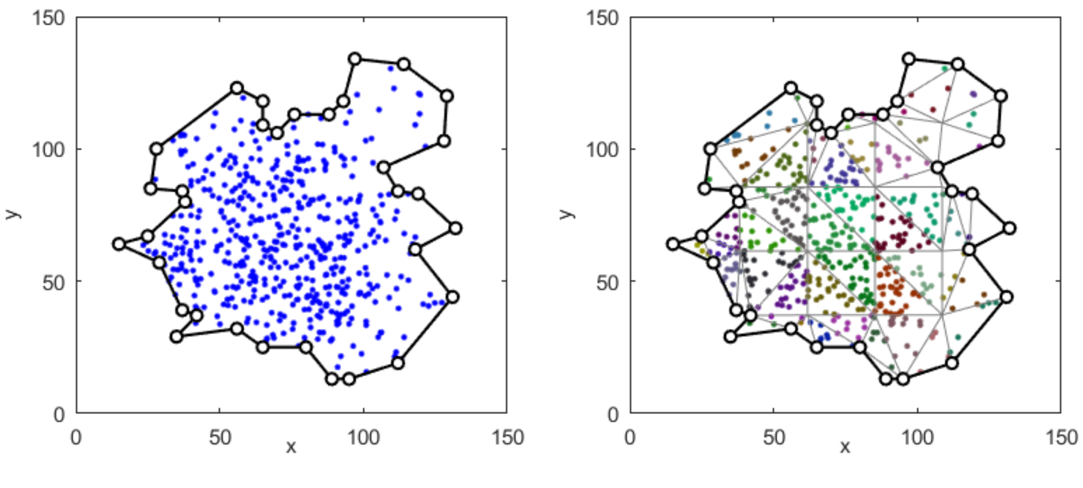
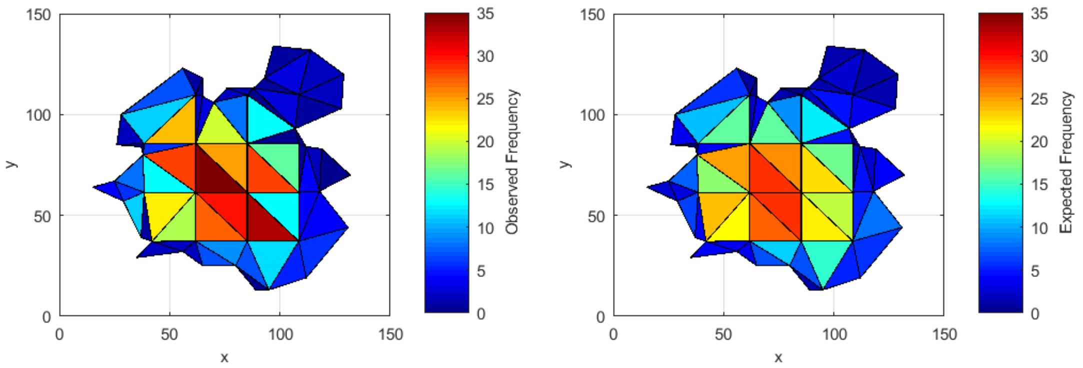


Figure 9. Random number generation in the Erzurum region; (a) Representation of the random numbers in a two-dimensional plane; (b) Classification of the random numbers by triangles.

The probability value of each triangle in the area which is divided into triangles is calculated numerically. The obtained probability values are multiplied by the number of sample size to calculate the expected frequencies (Figure 10(b)).



(a) (b)

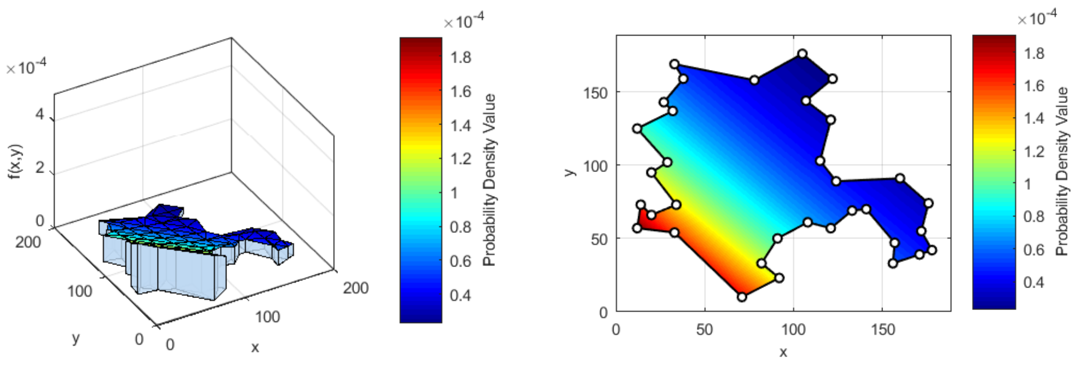
Figure 10. Graphical comparison of the observed and expected frequencies; (a) The encoloured histogram of observed frequencies in triangular area; (b) The encoloured histogram of expected frequencies in triangular area.

As a result of the simulation, the chi-square test statistic is $\hat{\chi}^2 = 55.4546$ and the critical value is $\chi^2_{1-0.05,59-1} = 76.7778$. Since $\hat{\chi}^2 < \chi^2_{1-0.05,59-1}$, therefore we cannot reject the null hypothesis.

Example 2. The Konya province is selected as a second example of simulation. The boundaries of the region of Konya province were reduced to 34 points using the point reduction method (Douglas & Peucker, 1973). This polygonal area was divided into triangles and was obtained 54 triangles. The probability density function of this polygonal area is determined arbitrarily and is given in Equation (16) (Figure 11).

$$f(x,y) = \frac{1}{2632} e^{-\frac{x+y}{100}} \tag{16}$$

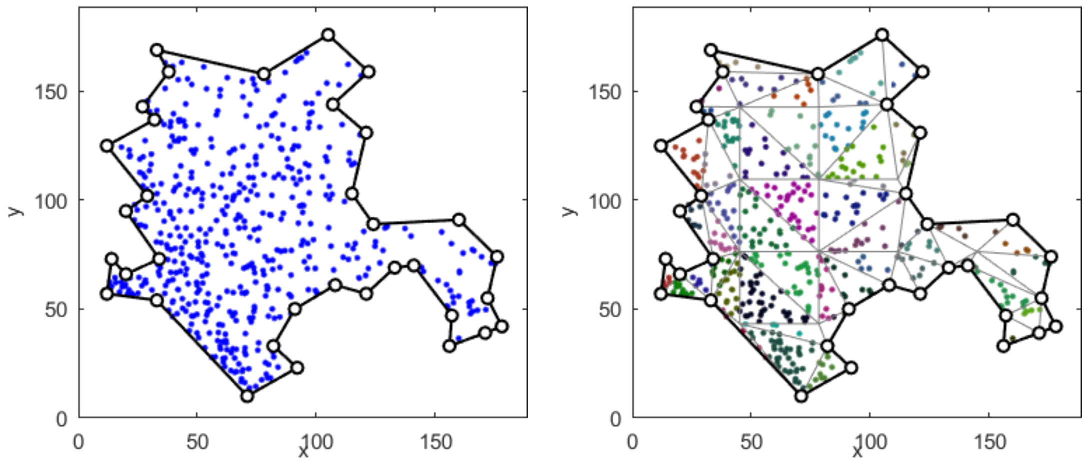
In Equation (16), the expected values are shown as ($\mu_x = \mu_y = 100$). 2632 is the scale value which is calculated for the volume should be "1" in the polygonal area.



(a) (b)

Figure 11. The used probability density function for Example 2; (a) Three-dimensional representation of the function; (b) The two-dimensional encoloured representation of the function.

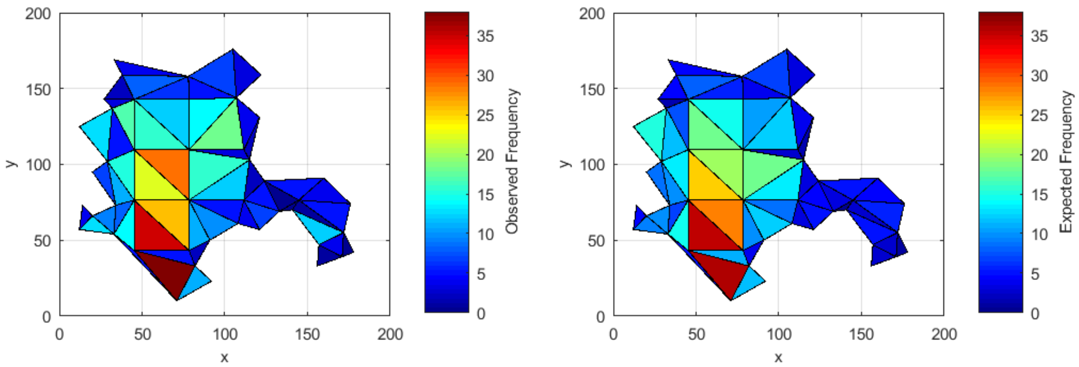
500 random numbers are generated from the given probability density by rejection method (Figure 12(a)) (Kesemen & Tiryaki, 2018). These random numbers are classified for each triangle in the Figure 12(b) and the observed frequencies of each triangle are calculated in the Figure 13(a).



(a) (b)

Figure 12. Random number generation in the Konya region; (a) Representation of the random numbers in a two-dimensional plane; (b) Classification of the random numbers by triangles.

The probability value of each triangle in the area which is divided into triangles is calculated numerically. The obtained probability values are multiplied by the number of sample size to calculate the expected frequencies (Figure 13 (b)).



(a) (b)

Figure 13. Graphical comparison of the observed and expected frequencies; (a) The encoloured histogram of observed frequencies in triangular area; (b) The encoloured histogram of expected frequencies in triangular area.

As a result of the simulation, the chi-square test statistic is $\hat{\chi}^2 = 55.3380$ and the critical value is $\chi^2_{1-0.05,54-1} = 70.9935$. Since $\hat{\chi}^2 < \chi^2_{1-0.05,54-1}$, therefore we cannot reject the null hypothesis.

Conclusion

In the literature, the binned regions of the bivariate data for chi-square fit test are used as rectangular. In this study, a two-dimensional arbitrary region has been defined by polygonal approach and then a new method has been developed for applying the chi-square goodness-of- fit test in polygonal area. Goodness of fit tests in polygonal areas are based on the techniques applied in rectangular areas. However, the difficulties in this calculation led to the development of different techniques. Goodness-of-fit tests by dividing the polygons into triangles have eliminated errors caused by the rectangular approach. In the chi-square test, the expected and observed frequencies can be calculated easily by dividing the polygons into triangles. Each triangle is considered a class. The chi-square test statistic is calculated by normalizing. For this reason, different class sizes do not constitute a problem. In some cases, the expected frequencies in the triangles may not be the desired size. In this situation, the expected and observed frequencies for each triangle are calculated and then the triangles with insufficient frequency can be combined with other triangles. The preferred combining process is performed between adjacent triangles. The expected and observed frequencies of the 500 random numbers generated by using Erzurum and Konya regions were calculated according to the 59 and 54 classes. The chi-square test is performed 1000 times. The obtained chi-square test statistic and the critical value of χ^2 with the degree of freedom are 58 and 53, are compared and then the acceptance number of the null hypothesis is given in Table 1.

Table 1. Chi-square simulation results

$\mathbf{1-\alpha}$	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**0.95**	**0.99**
$\boldsymbol{Erzurum}\ \#(\boldsymbol{H_0})$	512	614	706	802	916	956	991
$\boldsymbol{Konya}\ \#(\boldsymbol{H_0})$	521	612	691	781	903	958	984

Table 1 shows that the chi-square simulation was successful.

The chi-square goodness of fit test in polygonal areas was performed on a selected region. The fact that the triangles are independent of each other indicates that the same method can be applied to independent neighboring regions.

References

Anderson, T., & Darling, D. (1954). A test of goodness of fit. *Journal of the American statistical association*, *49*(268), 765-769.

Cochran, W. (1952). The chi-square test of goodness of fit. *The Annals of Mathematical Statistics*, 315-345.

D'Agostino, R., & Stephens, M. (1986). Goodness-of-fit techniques (Statistics, a series of textbooks and monographs). *Dekker, 68*, 1.

Devadoss, S. L., & O'Rourke, J. (2011). *Discrete and computational geometry*. Princeton University Press.

Douglas, D. H., & Peucker, T. K. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *10*(2), 112-122.

Foley, J., van Dam, A., Feiner, S., & Hughes, J. (1990). *Computer Garphics: Rpinciples and Practice* (2 b.). Addison-Wesley.

Gibbons, J. D., & Chakraborti, S. (2011). *Nonparametric statistical inference*. Berlin Heidelberg: Springer.

Haines, E. (1994). Point in polygon strategies, in: P. Heckbert (Ed.), Graphic Gems IV. (P. Heckbert, Dü.) Boston, MA: Akademic Press.

Harrington, S. (1983). Computer Graphics: A Programming Approach. McGraw-Hill.

Hormann, K., & Agathos, A. (2001). The point in polygon problem for arbitrary polygons. *Computational Geometry*, *20*(3), 131-144.

Kesemen, O., & Tiryaki, B. K. (2018). Non-Uniform Random Number Generation from Arbitrary Bivariate Distribution in Polygonal Area. *Süleyman Demirel University Journal of Natural and Applied Sciences*, 443 - 457.

Kesemen, O., Tiryaki, B. K., & Tuncay, U. (2019). A Novel Approximation for Computation Bivariate Cumulative Distribution Functions in Polygon Area. *Gümüşhane University Journal of Science and Technology Institute*. doi:DOI: 10.17714/gumusfenbil.413709

Krishnaiah, P. (1980). *Handbook of statistics* (Cilt 31). Motilal Banarsidass Publishe.

Lee, D.-T., & Lin, A. (1986). Generalized Delaunay triangulation for planar graphs. *Discrete & Computational Geometry, 1*(3), 201-217.

Mark, B. L., & Turin, W. (2011). *Probability, Random Processes, and Statistical Analysis.* Cambridge University Press Textbooks.

Massey Jr, F. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association, 46*(253), 68-78.

O'Rourke, J. (1998). *Computational Geometry in C* (2 b.). Cambridge University Press.

O'Connor, P., & Andre, K. (2011). *Practical reliability engineering.* John Wiley & Sons.

Preparata, F. P., & Ian, S. M. (1985). *Computational Geometry An Introduction.* New York: Springer-Verlag.

Samuel, K., & Johnson, N. (1983). Encyclopedia of Statistical Sciences. Volume.

Sedgewick, R. (1988). *Algorithms* (2 b.). Addison-Wesley.

Shewchuk, J. (2002). Delaunay refinement algorithms for triangular mesh generation. *Computational geometry, 22*(1-3), 21-74.

Weiler, K. (1994). An incremental angle point in polygon test, in: P. Heckbert (Ed.), Graphic Gems IV, . Boston, MA: Academic Press.

Worboys, M. F., & Duckham, M. (2004). *GIS: a computing perspective.* CRC press.