Examination of Pomological Features of Different Ten Raspberry Cultivars by the Methods of Various Statistics

S. Peral Eyduran, Y. Sabit Ağaoğlu, Ecevit Eyduran and Taner Özdemir

1Department of Horticulture, Faculty of Agriculture, University of Ankara, 06110 Ankara-Turkey.
2Biometry Genetics Unit, Department of Animal Science, Faculty of Agriculture, University of Yüzüncü Yıl, 65080, Van-Turkey.

Abstract: The aim of this study was to examine the effects of cultivar and year factors on pomological features of different ten raspberry cultivars (Rubin, Summit, Holland Short, Heritage, Tulameen, Aksu Red, Nuburg, Canby and Willamette) by the methods of various statistics such as Chi-Square, Likelihood Ratio Chi-Square and correspondence. For this aim, pomological traits such as Fruit weight (FW), Soluble Solid Content (SSC), Total acid (TA) were categorized and contingency tables were composed of calculating associations between Fruit weight (FW), Soluble Solid Content (SSC), Total acid (TA) and taste, aroma, cultivar and year. Chi-Square and Likelihood ratio Chi-Square statistics from contingency tables was calculated besides correspondence analysis was performed for pair of traits that power values of both statistics were 80% and more. Therefore, results of Power Analysis regards as Chi-Square and Likelihood Ratio Chi-Square statistics, was found more than 90% for pair of traits such as taste by cultivar, aroma by cultivar, Fruit weight by Cultivar, Total Acid by Cultivar, SSC by Cultivar, Total Acid by Year, SCC by Year (P<0.01). Then, Correspondence Analysis was applied for pair of traits mentioned above, having very high reliability. As to results of Correspondence Analysis, it should be suggested that a) Tulameen and Willamette cultivars had the highest fruit weight b) Willamette and Summit cultivars had the highest level of total acid c) the highest level of total acid was found in 2003 year d) Willamette and Canby cultivars were with highest level of SCC e) Heritage II and Tulameen cultivars had very-good aroma. f) the high level of SCC was in 2005 year. As a result, using different statistical approach from other studies in the literature it was concluded that the effects of cultivar and year on fruit weight (FW), Soluble Solid Content (SSC), Total acid (PA) and taste, aroma were statistically significant. With three statistical analysis techniques, researcher may gain more information.

Key words: Raspberry, Correspondence analysis, Chi-square, Likelihood ratio chi-square, Pomological traits, Power Analysis.

1. INTRODUCTION

It is well-known that the effects of genetics factor (variety) and environmental factors (year, season, irrigation, climate, soil) on plant’s pomological traits and adaptation were significant[8,11,3,5,9,7,2]. Contrary to other authors, used only Random complete Design for statistical analysis,[8,11,3,5,9,7,2] by using different statistical analysis techniques such as Chi-Square, Likelihood Ratio Chi-Square and correspondence as new approaches, present paper aimed to be examined effects of cultivar, year on pomologic traits such as Fruit weight (FW), Soluble Solid Content (SSC), Total acid (TA) from ten different raspberry cultivars which were raised in Ankara, which is located in the capital centre of Turkey during 2002-2005 in the experiment of adaptation.

2. MATERIALS AND METHODS

2.1. Materials: The experiments on adaptation of different ten varieties (Rubin, Summit, Holland Short, Heritage, Tulameen, Aksu Red, Nuburg, Canby and Willamette) were carried out throughout 2002-2005 at 1000 m² area in Application Farm of Horticulture, Faculty of Agriculture, Ankara University, in the capital centre of Turkey. Two rows of each shrub plants set at 1.5 x 2 meter spacing. Raspberries were harvest at twice a year for each year: pre-harvest 4th July – post-harvest 15th August and pre-harvest were merely based on because yield of the raspberries in post-harvest were little. At pre-harvest, the raspberries were weighed fresh fruit and average fruit weight were calculated from a-50 fruit sample randomly selected from every one of three plots of each raspberry variety. Soluble Solid Contents (SSC) of

Corresponding Author: S. Peral Eyduran, Department of Horticulture, Faculty of Agriculture, University of Ankara, 06110 Ankara-Turkey.
berries (°Brix) was measured by digital reflectometer and Total acid (TA) for berries were determined by titration. According to the traits such as Fruit weight, Total acid and Soluble Solid Content, The aim of the experiment was to determine the best raspberry cultivar(s) which was/were adapted to Ankara condition\[2\]. For each trait, a total of 120 observations (4 year x 10 cultivars x 3 replications) were collected from ten raspberry cultivars in four years.

However, in present paper, associations between discrete variables were examined by Chi-Square (1) and Likelihood Ratio Chi-Square statistics (2). Then, Correspondence Analysis was only carried out on Pair of traits having power values of at least 80% for both statistics.

2.2. Methods:

2.2.1 The structures of variables:

a) Categorized Variables: Continuous variables such as fruit weight (FW), total acid (TA) and, Soluble Solid Content (SSC) obtained from various ten raspberry cultivars were divided into three categorizes; low (1) and medium (2) and high (3) weight. In other words, new values of the variables were categorized or coded as 1, 2 and 3. Minimum and maximum values of any continuous variables (FW, TA and SSC) were estimated and range was divided into three classifications. The cut-off values of the classification values for each continuous variable are summarized in Table 1. For instance: when one value for SSC in data set is 25.2, categorized value which was assigned to it should be 1 (low).

b) Discrete variables: Traits such as taste (sour, medium, taste, very-taste), aroma (bad, medium, good and very good), year (2002, 2003, 2004 and 2005), cultivars (Rubin, Summit, Holland Short, Heritage I, Heritage II, Tulameen, Aksu Red, Nuburg, Canby and Willamette) were used as discrete variables in our data set.

2.2.2 Chi-Square and Likelihood Ratio Chi-Square Statistics: The notations of Chi-Square (1) and Likelihood Ratio Chi-Square statistics (2) for contingency tables are given below\[6\]:

$$x^2 = \sum \frac{(f_i - f_{ij})^2}{f_{ij}}$$  \hspace{1cm} (1)

$$G = 2 \sum f_i \ln \left( \frac{f_i}{f_{ij}} \right)$$  \hspace{1cm} (2)

Where, $f_i$ observed frequency and $f_{ij}$ expected frequency.

2.2.3 Power Theory on Chi-Square and Likelihood Ratio Chi-Square Statistics: Chi-Square and Likelihood Ratio Chi-Square statistics have been commonly used as criteria of independence and goodness of fit in contingency table\[6\]. One should perform power analysis as regards both, which give an idea to one about whether sample sizes will be enough. The most important question for a researcher is "How many observations should we survey to guarantee statistics having a power of %80-90?" In order to be obtained reliable results and determined enough sample size in Chi-Square and Likelihood Ratio Chi-Square Statistics, power analysis should be done. Moreover, it should be forgotten that non-significant results for both statistics does not guarantee independence.

Theoretical information of power analysis for Chi-Square (1) and Likelihood Ratio Chi-Square statistics (2) are summarized as follows:

Assume that $H_0$ is the same to model M for a contingency table. Let $\pi_i$ indicate the true probability in $i^{th}$ cell and Let $\pi_i (M)$ represent the value to which the Maximum likelihood (ML) estimate $\pi_i$ for model M converges, where $\Sigma \pi_i (M) = 1$. For multinomial sample of size $n$, the non-centrality parameter for Chi-Square (3) can be expressed as follows:

$$\lambda = n \sum \frac{(\pi_i - \pi_i (M))^2}{\pi_i (M)} \hspace{1cm} (3)$$

Expression 3 is the similar form as Chi-Square statistics, with for the sample proportion $\pi_i$ and $\pi_i (M)$ in place of $\pi_i$. The non-centrality parameter for Likelihood Ratio Chi-Square Statistics (4) can be written in this manner:

$$\lambda = 2n \sum \pi_i \log \frac{\pi_i}{\pi_i (M)} \hspace{1cm} (4)$$

For every contingency tables, the power analysis for Chi-Square and Likelihood Ratio Chi-Square Statistics were performed using a special SAS macro downloaded from web site: (http://euclid.psych.yorku.ca/ftp/sas/vcd/macros/powerxc.sas)\[6\].

---

**Table 1:** The cut-off values of pomological features being continuous variables.

<table>
<thead>
<tr>
<th>Categories</th>
<th>FW</th>
<th>TA</th>
<th>SSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (1)</td>
<td>9.4 – 16.37</td>
<td>13.6 – 18.52</td>
<td>21.00 – 25.82</td>
</tr>
<tr>
<td>Medium (2)</td>
<td>16.38 – 23.34</td>
<td>18.53 – 23.45</td>
<td>25.83 – 30.65</td>
</tr>
<tr>
<td>High (3)</td>
<td>&gt;23.35</td>
<td>&gt;23.46</td>
<td>&gt;30.66</td>
</tr>
</tbody>
</table>

FW: Fruit weight; TA: Total acid; SSC: the soluble substance that can be dissolved in water.
There was a significant association between SCC by Year (P<0.01).
There was a significant association between Total by Cultivar (P<0.01).
There was a significant association between Fruit acid by Cultivar (P<0.01).
There was a significant association between Taste by Cultivar (P<0.01).
There was a significant association between Aroma by Cultivar (P<0.01).
There was a significant association between Total acid by Cultivar (P<0.01).
There was a significant association between SCC by Year (P<0.01).
There was a significant association between Taste by Cultivar (P<0.01).
There was a significant association between Aroma by Cultivar (P<0.01).
There was a significant association between Total by Year (P<0.01).
There was a significant association between SCC by Year (P<0.01).

### 3. RESULTS AND DISCUSSION

#### 3.1. Chi-Square and Likelihood-Ratio Chi-Square Statistics and Their Power Analysis:

The values of probability power and non-centrality parameter concerning Likelihood Ratio Chi-Square and Chi-Square Statistics in each contingency tables which were calculated for the values of Type I error of 0.05 are presented in Table 2. Examined in Table 2, it could be said that: for both statistics;

- There was a significant association between Taste by CULTIVAR (CULT) (P<0.01).
- There was a significant association between Aroma by CULTIVAR (P<0.01).
- There was a significant association between Fruit weight by CULTIVAR (P<0.01).
- There was a significant association between SCC by CULTIVAR (P<0.01).
- There was a significant association between Total acid by CULTIVAR (P<0.01).
- There was a significant association between SCC by Year (P<0.01).
- There was a significant association between Taste by CULTIVAR (P<0.01).
- There was a significant association between Total by Year (P<0.01).
- There was a significant association between SCC by Year (P<0.01).

### Table 2: The values, probability power and non-centrality parameter values of Likelihood Ratio Chi-Square and Chi-Square Statistics in each contingency tables calculated for alpha=0.05.

<table>
<thead>
<tr>
<th>Pair of Features</th>
<th>Chi-Square Statistic</th>
<th>Likelihood-Ratio Chi-Square Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value</td>
<td>Prob</td>
</tr>
<tr>
<td>Taste by CULT.</td>
<td>93.344</td>
<td>0.001</td>
</tr>
<tr>
<td>Aroma by CULT.</td>
<td>99.977</td>
<td>0.001</td>
</tr>
<tr>
<td>Taste by Year</td>
<td>6.227</td>
<td>0.717</td>
</tr>
<tr>
<td>Aroma by Year</td>
<td>6.846</td>
<td>0.653</td>
</tr>
<tr>
<td>FW by CULT.</td>
<td>210.909</td>
<td>0.001</td>
</tr>
<tr>
<td>SCC by CULT.</td>
<td>77.522</td>
<td>0.001</td>
</tr>
<tr>
<td>TA by CULT.</td>
<td>141.996</td>
<td>0.001</td>
</tr>
<tr>
<td>Taste by Aroma</td>
<td>105.213</td>
<td>0.001</td>
</tr>
<tr>
<td>FW by Year</td>
<td>3.833</td>
<td>0.699</td>
</tr>
<tr>
<td>TA by Year</td>
<td>20.768</td>
<td>0.002</td>
</tr>
<tr>
<td>SCC by Year</td>
<td>37.812</td>
<td>0.001</td>
</tr>
</tbody>
</table>

#### 2.2.4 Correspondence Analysis:

Correspondence analysis (CA), used for studies regarding genome explorations, modern ecology and vegetational succession, animal sciences, social, economy, has been called many different names in the literature: Contingency table analysis, RQ-technique, Reciprocal averaging, reciprocal ordering, Dual Scaling, Optimal scaling, Optimal Scoring and quantification method or Homogeneity Analysis\[12\]. CA, which is analogous to Principle Components, is a descriptive analysis technique which is designed to analyze two-way and multi-way tables containing measures of correspondence between the row and column variables. In other word, it supplies a statistic method for representing data in a Euclidean space so that the results can be visually observed for structure\[3\] and graphically, correspondence analysis is presented in terms of two components of the variation in row and column profiles that represent the table most closely. The Y axis is the first component, the X axis the second. If two rows or columns have similar profiles, their points in the correspondence analysis plot will be close together.

Statistics such as Chi-Square, Fisher’s Exact Test and Likelihood Ratio Chi-Square (which has also called G Statistic) commonly were used for testing or analyzing independence between two variables in two way tables. Contrary to these three statistics, researchers can prepare that Correspondence Analysis determines the relationships both that categories of each variable can be among their own and categories of other variable. Besides, other advantageous of CA is there are no any assumptions and hypothesis on distribution of data set. However, no observation having negative value should include in data set\[4\].
Contrary to statistical analysis techniques used in present paper, adaptation experiments of many authors carried out in Random Complete Design. The results of present paper were likely to be different from those of other authors [6,11,3,5,9,7,2].

It was pointed out that power analysis values of Chi-Square and Likelihood-Ratio Chi-Square Statistics calculated for crosstabulations of taste by year; aroma by year and fruit weight (FW) by year were power values of low level approximately 35.2% and 42.3%; 38.3 and 42.3; and 25.9 and 25.4, respectively. However, corresponding values of others except for those crosstabulation for both statistics were ranged from 94.9 % to 100 % for Chi-Square statistic and from 97.065 % to 100 % for Likelihood Ratio Chi-Square statistic, which meant that the results obtained from both statistics were excessively reliable(Table 2).

With regard to Table 2, it could be suggest thatwith exception od taste by year, Arama by year, FW by year, results of the values and power analysis gaining both statistics of two-way tables handled in present paper were much more reliable and ideal because they had more a power than 80 %. For Contingency tables (not bold) having more a power of than 80%, it could be shown that total sample size of 120 observations used for each features in the study were very sufficient.

3.2. Taste by Cultivar: Symmetric graph of Taste by Cultivar is presented in Figure 1. As to Fig.1,

- The points of Heritage I was suitable to the one of taste, but the points of Willamette and Aksu red closed to the one of Heritage I.
- The points of Rubin, Heritage II and Tulameen cultivars closed to very-taste point.
- Canby cultivar’s point had same distance to Sour and medium-taste.

3.3. Aroma by Cultivar: Symmetric graph of Aroma by Cultivar is shown in Fig.2. With respect to Fig.2,

- The points of Heritage II and Tulameen cultivars almost corresponded to one of very-good aroma, which relatively closed to Rubin cultivar.
- Aksu Red’s point closed to the one of good-aroma than those of Willamette and Heritage I. Besides, the points of Willamette and Heritage I had identical distance to the one of good-aroma.
- The points of Summit and Nuburg cultivars coincident together and related to the point of medium-aroma other than that of bad-aroma.

3.4. Fruit Weight by Cultivar: Symmetric graph of Fruit Weight by Cultivar is given in Fig. 3. As to Fig. 3,

- The points or the coordinates of Heritage I, Heritage II, Canby and Aksu red cultivars coincided with each other and corresponded to medium-level of Fruit weight.
- The points of Tulameen and Willamette cultivars coincided together and related to the point of medium-aroma other than that of bad-aroma.

Fig. 2: Symmetric graph of aroma by cultivar.

Fig. 1: Symmetric graph of taste by cultivar.
Fig. 3: Symmetric graph of Fruit Weight by Cultivar.

- The point of Holland short cultivar subsisted exactly between low-level of Fruit weight and medium-level of Fruit weight.

Fig. 4: Symmetric graph of SCC by Cultivar.

3.5. SCC by Cultivar: In relation to Symmetric graph of SCC by Cultivar in Fig.4,

- As regards SCC, the points of Willamette and Canby cultivars corresponded to its high-level.
- The coordinates of Tulameen and Hollanda short cultivars closed or related to the ones of the medium level of SSC. In addition, the ones of Nuburg and Heritage II cultivars had equal distance to the ones of the medium level of SSC.
- The points of Summit, Rubin and Aksu red cultivars corresponded roughly to the point of the low level of SSC while Rubin and Aksu red cultivars had same coordinates.

3.6. Total Acid (TA) by Cultivar: As to Symmetric graph of Total Acid by Cultivar in Fig.5,

- The points of Heritage I and II cultivars coincided with each other and with Canby cultivar had roughly equal distance to medium level of total acid.
- The points of Aksu Red and Tulameen was said to be the same distance to the point of low level of total acid. Then, Holland Short cultivar also corresponded to the one of low level of total acid. However, the

Fig. 5: Symmetric graph of Total Acid by Cultivar.

Fig. 6: Symmetric Draph of Taste by Aroma.
ones of Nuburg and Rubin cultivars exactly coincided with each other and were closer to low-level of it than medium level of it.

- The initial, the point of Willamette cultivar and then the one of Summit closed to the one of high level of total acid.

3.7. Taste by Aroma: As regards Symmetric graph of Taste by Aroma in Fig 6,

- The points of very taste and very good aroma overlapped each other.
- The points of taste and good aroma quite closed to each other.
- The points of medium taste and medium aroma approximately coincided to each other.
- The point of sour corresponded to the one of medium aroma

3.8. Total Acid (TA) by Year: As to Fig.7,

- The point of low level of total acid closed to the one of 2002 year.
- The points of 2004 year and 2005 had the same distance to the point of medium level of total acid.
- The point of high level of total acid closed to the one of 2003 year.

3.9. SCC by Year: With respect to Fig.8,

- The point of the high level of SCC corresponded to year 2002 year and then 2003 year.
- The point of the medium level of SCC corresponded to 2004 year, which was more or less between the points of high and medium level of it.

As a result, it was concluded that in point of pomological traits mentioned above, association between cultivar and pomological traits were significant.

REFERENCES


