Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA. **Rect@**

Volumen 11. Paginas 155 a 177

REVISITING THE EXPECTANCY/DISCONFIRMATION PARADIGM FOR SMALL QUESTIONNAIRES: THE CULTURAL/HERITAGE TOURISM CASE.

MONTERO LORENZO, JOSE MARIA
Jose.MLorenzo@uclm.es
Universidad de Castilla La Mancha/ Departamento de Estadística
Cobertizo de San Pedro Mártir s/n (Toledo)

FERNÁNDEZ-AVILÉS CALDERÓN, GEMA
gema.faviles@uclm.es
Universidad de Castilla La Mancha/ Departamento de Estadística
Cobertizo de San Pedro Mártir s/n (Toledo)

GARCÍA CENTENO, MARÍA DEL CARMEN
garcen@ceu.es
Universidad CEU San Pablo/Departamento de Métodos Cuantitativos e Informática
C/ Julián Romea, 23 (Madrid)

Recibido 09/07/2010
Revisado 22/07/2010
Aceptación 21/09/2010

RESUMEN: A pesar de la emergencia de nuevas estrategias, los modelos de expectativas-confirmación de las mismas son muy populares en las investigaciones de turismo. Esta es la razón por la que se revisan tres aspectos: (i) son diferentes a las investigaciones habituales, y con el fin de corregir el subjetivismo de los turistas, la satisfacción global de los turistas se mide utilizando medias ponderadas de la satisfacción con atributos individuales; ii) una alternativa a estos modelos consiste en dar cuestionarios simples donde los turistas sólo tienen que contestar si algunos servicios turísticos han superados sus expectativas o no se han llegado a alcanzar; y iii) la Teoría de la Asimilación se revisa en términos de la asociación positiva entre las expectativas y su confirmación (lo que se conoce como la versión positiva de la Teoría de la Asimilación). Además, se propone el estadístico Chi-cuadrado con una corrección de continuidad asimétrica para contrastar la hipótesis de independencia. Para llevar a cabo este artículo, se ha utilizado la información proporcionada por un cuestionario realizado a 1500 turistas y especialmente diseñado para medir su satisfacción en Toledo (declarada por la UNESCO ciudad patrimonio de la humanidad), España.

Palabras claves: Corrección de continuidad; tabla de contingencia 2x2; atributos; teoría de la asimilación; satisfacción; turismo cultural.

ABSTRACT: Despite the emergence of new strategies, models of expectancy/disconfirmation are still very popular in tourist research. This is why they are revisited in three aspects: (i) unlike usual research, and in order to correct subjectivism, tourists' overall satisfaction is measured through a weighted mean of satisfaction with individual attributes; (ii) an alternative to these models is given for simple questionnaires where tourists only have to answer if some tourist services have lived up to their expectations or in what extent they have been exceeded or fallen short of such expectations; and (iii) the Assimilation Theory is revisited in terms of positive association between expectancy and slight disconfirmation (what we call the

positive version of the Assimilation Theory). A chi-square statistic including an asymmetric continuity correction is provided to test the independence hypothesis. This article relies on the information provided by 1,500 respondents who were given a small questionnaire specially designed to measure tourist satisfaction in the emblematic part of Toledo, Spain (a UNESCO World Heritage City).
*Keywords***:** Continuity correction; 2x2 contingency tables; attributes; assimilation theory; satisfaction; cultural/heritage tourism.

1. Introduction

Today, although we are just emerging from the worst economic crisis the world has seen in more than 60 years, the latest research by WTTC and Oxford Economics demonstrates the tremendous scale of the world's tourism sector (WTTC, 2009) and confirms that travel and tourism continues to be a leading employer. An estimated 235 million people now work either directly in travel and tourism or in related sectors of the economy – accounting for more than 8% of global employment, or 1 in 12.5 jobs. Moreover, despite the recession, Travel and tourism still contributes over 9% of global GDP. In Spain, the country where focus our study, the above mentioned figures are even exacerbated. Also according to the WTTC and Oxford Economics the jobs generated in 2009 by tourist economy GDP account for 18% of global employment in the country (3423.3 thousand employees), and tourist activities (the broadest measure that includes both the direct effects and the indirect effects via the supply chain of tourism spending) contribute by 15.6% of direct industry of global Spanish GDP. It is therefore clear that tourism is one of the major forces in the economy of the world, an activity of global importance and significance. Moreover, tourism has been remarkable in its resistance to adverse economic and political conditions (Ref. 1). Events such as the terrorist bombing on 11th September, 2001 and the 2004 Boxing Day Tsunami clearly demonstrate the sector's ability to regroup and place emphasis on a new vocabulary including words such as "safety", "security", "risk management", "crisis" and "recovery". Inevitably though, growth is slowing as the market matures and, as the nature of the tourist and his or her demands change, the sector will need to be creative in supplying products to satisfy the "new tourist".

One particular case of tourism is cultural/heritage tourism, where "cultural/heritage" can be defined as the monuments, buildings and archaeological sites of outstanding universal value from the point of view of history, art or science (Ref. 2).

Cultural/heritage tourism is one of the fastest growing segments of the tourism industry, and obviously, the growth in the cultural/heritage tourism market may provide several benefits to cultural/heritage destinations. Because of people's inclination to seek out novel attractions, including traditional cultures, heritage tourism has become a major "new" area of tourism demand, which almost all policy–makers are now aware of and anxious to develop. Heritage tourism, as part of the broader category of "cultural tourism", is now a major pillar of the nascent tourism strategy of many countries. Cultural/heritage tourism strategies in various countries have in common that they are a major growth area, that they can be used to boost local culture and that they can aid the seasonal and geographic spread of tourism (Ref. 3).

As tourists are becoming more sophisticated, their need to recapture the past has increased. Tourists have been visiting cultural/heritage sites more frequently. Following Ref. 2, cultural/heritage tourism offers several benefits to tourists and residents, as well as governments. First of all, cultural/heritage tourism protects historic, cultural, and natural resources in communities, towns, and cities. Second, cultural/heritage tourism educates residents and tourists about local/regional history and traditions. Through the research about and development of heritage/cultural destinations, residents will become better informed about local/regional history and traditions which can be shared with tourists. Third, cultural/heritage tourism builds closer, stronger communities. Knowledge of heritage provides continuity and context for communities, which instils respect in their residents, strengthens citizenship values, builds community pride, and improves quality of life. Fourth, cultural/heritage tourism promotes the economic and civic vitality of a community or region. Economic benefits include: the creation of new jobs in the travel industry, at cultural attractions, and in travel-related establishments; economic diversification in the service industry (restaurants, hotels/motels, bed-and-breakfasts, tour guide services), manufacturing (arts and crafts, souvenirs, publications), and agriculture (specialty gardens or farmers' markets); encouragement of local ownership of small businesses; higher property values; increased retail sales; and substantial tax revenues. Of course, the uncontrolled increasing of the tourist activity in general and cultural/heritage tourism in particular, can lead to undesirable consequences in case that the capacity of the destination is exceeded and/or its tourist development is not sustainable.

In general, satisfying customers is important from different perspectives. Research has shown that satisfaction can affect customer retention and also lead them to recommend the goods or services to others. And this can be applied to tourism. So, it is crucial to investigate which factors are important for tourists. According to economic theory in tourism destination management, maximising travel satisfaction is crucial for a successful business. The evaluation of the physical products of destination (instrumental performance) as well as the psychological interpretation of a destination product (expressive attributes) is necessary for human actions (Ref. 4), which could be represented as travel satisfaction and destination loyalty. With reference to the tourist industry, it can be stated that satisfied tourists are more likely to recommend the tourist destination to others, which is the cheapest and most effective form of marketing and promotion (Ref. 5). Furthermore, tourist satisfaction usually contributes to increased rates of tourist patronage, loyalty and acquisition retention, which in turn helps to achieve economic goals such as increasing the number of tourists and revenues. These factors, among others, are the reason why tourist satisfaction is a subject that is worth being studied (Ref. 6). Bearing in mind that tourism is an experience made up of many different interdependent parts, some more tangible than others, tourist satisfaction may be treated as a cumulative measure of total purchase and consumption experience over time (Ref. 7).

Quality management theories indicate that many key product and service attributes have a non linear relationship with satisfaction (Ref. 8). Most of the studies conducted to evaluate consumer satisfaction have utilised models of expectation/disconfirmation (Ref. 9 and Ref.10), equity (Ref.11 and Ref 12), norm (Ref. 13 and Ref 14) and perceived overall performance (Ref. 15 and Ref. 16). Nevertheless, it has been recently proposed that not only the cognitions –such as

expectations and disconfirmation- but also emotions can play an important role in satisfaction formation (see, Ref. 17; Ref. 18; Ref. 19, and Ref. 20, among others). Even a cognitive-affective view has been recently proposed, where satisfaction is influenced by the individual's cognitive judgements and emotions derived from the consumption experience (Ref. 21, Ref. 22, Ref. 20 are good examples).

According to the expectation/disconfirmation model, consumers develop expectations about a product or service before purchasing. Subsequently, they compare actual performance with those expectations. According to equity theory (Ref. 11), consumer satisfaction can be seen as a relationship between the costs of what the consumer spends and the rewards (benefits) he/she anticipates. Here, price, benefits, time and effort are major factors in determining satisfaction. Thus, it can be said that if tourists receive benefits or value based on their time, effort and money for travel, the destination is worthwhile. Norms serve as reference points to judge a product or service and dissatisfaction comes into play as a result of disconfirmation relative to these norms. Ref. 9, hypothesised that leisure satisfaction is determined by the disparity consumers perceive between the preferred and actual leisure experiences, as well as the perceptions of barriers (both internal and external) that prevent the tourist (in this case) from achieving the desired experience. The norm theory uses some form of "comparison standard". Tourists can compare current travel destinations with other alternative destinations or places visited in the past. The difference between present and past experiences can be a norm used to evaluate tourist satisfaction. According to the perceived performance model, tourist dissatisfaction is only a function of the actual performance, regardless of expectations. The Shapley value model is a measure of the importance of including each attribute in the set of key dissatisffiers. The Shapley value works by assessing the relative effect on the dependent variable by different combinations of predictor variables. Finally, SERVQUAL model (Ref. 23) is a strategy for measuring service (SERV) quality (QUAL) which has five generic factors: Tangibles, reliability, responsiveness, assurance and empathy. It is based on comparisons between customers' expectations and their perception of actual performance.

Tourists, like other customers, usually have initial expectations of the type and quality of services to be offered at a particular destination. The extent to which tourist expectations are met will eventually determine the level of tourist satisfaction. If the overall performance, while or after visiting a destination, exceeds or meets initial expectations, then the tourist is considered satisfied. Otherwise, the tourist may be dissatisfied. Obviously, the evaluation of tourist satisfaction needs to be considered in multiple dimensions. Interesting classic research in this sense includes Ref. 24, Ref. 7 and Ref. 25.

As said before, the evaluation of tourist satisfaction needs to be considered in multiple dimensions. But this is a large and space-consuming task, which is why this paper focuses on the following three novel aspects of the expectation/disconfirmation model: The way individual and overall tourist satisfaction is computed, the way marks are given to previous expectations without enlarging the questionnaire, and the use of a chi-square test including an asymmetric continuity correction to determine association between expectation (low, high) and slight disconfirmation (negative, positive) in tourist attributes.

The reminder of the paper is structured as follows. Section 2 presents the technical details of the above-mentioned methodological aspects. Section 3 describes the case study (study site, dataset, and profile of respondents). Section 4 reports the main results obtained from the application of the proposed statistical novelties to measuring overall satisfaction, attributing marks to previous expectations without enlarging the questionnaire and testing the independence between expectations and satisfaction versus what we call the positive version of the Assimilation Theory. The paper ends with some concluding remarks.

2. Methodological Aspects

2.1. Measuring overall tourist satisfaction

The usual procedure for measuring overall tourist satisfaction consists of including an item in the questionnaire of the type "In *general, how satisfied were you with your visit to X*", the respondent being given a range of possibilities for valuation (for example, from 1 to 7). However the quantitative (or qualitative) response of the tourist is highly subjective and in many cases the mark attributed to overall satisfaction contradicts the marks given to individual tourist services. One possible cause for this "contradiction" is positive or negative emotions, the so-called affective or emotional factor. Having checked this fact in a preliminary questionnaire (also in 2009) given to 100 tourists in several Spanish cultural/heritage cities (Toledo, Salamanca and Cuenca, three cities with a consolidated cultural/heritage tourism), the alternative we propose to assess overall tourist satisfaction is as follows:

- Include an item in the questionnaire referring to the importance tourists assign to the attributes that are going to be evaluated. That is, tourists are asked to rank a number of attributes (*m*) from 1 to *m*, according to how important they consider them (1 the most important, *m* the least important).
- Weight the marks given to these tourist attributes with the inverse of the ranks obtained in step 1 and compute their weighted mean, the result being individual overall tourist satisfaction.

Global overall tourist satisfaction can be obtained as follows:

1) Compute the sum of ranks provided by every tourist for each attribute:

$$T = \sum_{i=1}^{n}\sum_{j=1}^{m} O_{ij} = \frac{n(n+1)}{2}$$

where O_{ij} is the rank of the *j*-th attribute provided for the *i*-th polled tourist, *n* is the number of polled tourists, and *m* the number attributes evaluated.

2) Compute the sum of ranks for the *j*-th attribute:

$$T_j = \sum_{i=1}^{n} O_{ij} \tag{1}$$

3) Evaluate the weight attributed to the *j*-th service. As the more important attributes are ranked lower, weights are

$$W_j = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} O_{ij}}{\sum_{i=1}^{n} O_{ij}} = \frac{\frac{n(n+1)}{2}}{\sum_{i=1}^{n} O_{ij}} = \frac{T}{T_j} \qquad \forall j = 1,\ldots,m \tag{2}$$

4) Construct a global overall tourist satisfaction measure (in time *t*) as:

$$SM_t = \sum_{j=1}^{m} \lambda_j \bar{P}_j, \qquad \lambda_j = \frac{W_j}{\sum_{j=1}^{m} W_j} \tag{3}$$

where $\bar{P}_j$ represents the average mark the visitors assigned to the *j*-th attribute in the area under study.

Obviously, SM_t ranges between 0 and the maximum mark a tourist can give to the attributes being evaluated.

2.2. An alternative to attributing marks to previous expectations without enlarging the questionnaire

As said before, the usual way to measure overall tourist satisfaction is to include an item in the questionnaire of the type "In *general, how satisfied were you with your visit to X*", the respondent being given a range of possibilities for valuation (for example, from 1 to 7). However, in addition to the contradictions between this overall satisfaction mark and the marks given to individual tourist attributes, a new problem was detected in the preliminary questionnaire: On the one hand, tourists are reluctant to complete long questionnaires and on the other hand, giving a mark to the satisfaction expected from the different tourist attributes, which substantially enlarges the questionnaire, is valuable information for research purposes. This trade-off (a small number of long questionnaires with a small number of respondents or a short questionnaire with a large number of respondents) can be solved by including a final item in the questionnaire where respondents only have to mark if their expectations were equalled, slightly exceeded or not exceeded.

The problem that now arises is how to test the significance of the difference between the expected mean and the actual mean (in terms of satisfaction), because there is no quantitative mark for expected satisfaction. But this problem can be easily solved by attributing the median between the actual mark given to a tourist service attribute and the maximum mark to previous expectations, if the actual mark is high and the tourist assures that his or her previous expectations have not been reached. If the actual mark is low and the tourist states his or her expectations have been exceeded, the numerical value of these expectations will be computed by the median between 0 and the actual mark. The cases corresponding to a high actual mark, an attribute having exceeded expectations, or a low actual mark, expectations not being met, are certainly scarce. In these cases tourist expectations are computed by the median of the mean expectation of tourists that equalled their expectation and the individual mark actually given. Finally, the case where tourists affirm that an attribute has equalled their expectations is not problematic.

2.2.1. Assimilation Theory (the positive version): A chi-squared test with asymmetric continuity correction to determine association between expectation and slight disconfirmation

According to the Assimilation Theory (Ref. 26), individuals suffer a psychological conflict when they perceive discrepancies between performance and prior beliefs and they tend to adjust perceptions to their expectations in order to minimise, or even remove, that tension (Ref. 27). In these circumstances, expectations are a driver of satisfaction. It is of particular interest what we call the positive version of the Assimilation Theory, that is to say, the association of high expectations with a slight positive disconfirmation (satisfaction slightly exceeds high expectations and does not reach low expectations) because it could lead to underestimations or overestimations of global satisfaction. An easy way to check the positive version of the Assimilation Theory in the field of tourism is to test the hypothesis of independence between the labels of factors "expectancy" (low, high) and "slight disconfirmation" (negative, positive). Should the independence hypothesis be rejected, positive association will indicate that high expectations are related to slight positive disconfirmation and low expectations to slight negative disconfirmation; negative association can be understood as the relationship between high expectations and negative slight disconfirmation and low expectations and positive slight disconfirmation. That is, positive association implies a slight exacerbation of expectations and favours the positive version of the Assimilation Theory, while negative association means correction of expectations (in case of high expectations *things are not as good as expected*, but in case of low expectancy *things are not as bad as expected*).

The chi-square test of independence is the instrument commonly used for this purpose, as it provides an approximation to the so-called "exact probability" (the probability of having a result equal to or more distanced from the hypothesis of independence than the sample result). At this point, a methodological novelty will be included. The "exact probability" will be approximated by including an asymmetric continuity correction in the chi-square statistic. As shown in Ref. 28, this asymmetric correction for continuity works better than Ref. 29 and Ref. 30 corrections the underestimations and overestimations of the exact probability provided by the usual Ref. 31 correction. Some technicalities are given next in the framework of (2x2) tables. Generalisation to (RXC) tables is straightforward.

The proposed strategy for assessing the probability of observing a table differing to the same extent or more from the independence hypothesis than that observed is as follows. Let n_{ij} and $\hat{E}_{ij}$ be the observed and expected frequency (under the null), respectively, in a chosen cell $\{i, j\}$.

- If $n_{ij} \leq \hat{E}_{ij}$

$$P_{exact} \approx P\left(N_{ij} \leq n_{ij} + \frac{1}{2} \right) + P\left(N_{ij} \leq 2\hat{E}_{ij} - n_{ij} + \Delta \right) \quad (4)$$

- If $n_{ij} \geq \hat{E}_{ij}$

$$P_{exact} \approx P\left(N_{ij} \leq 2\hat{E}_{ij} - n_{ij} + \Delta \right) + P\left(N_{ij} \geq n_{ij} - \frac{1}{2} \right) \quad (5)$$

with

$$N_{ij} \xrightarrow{aprox} N\left(\frac{n_{i\bullet}n_{\bullet j}}{n};\sqrt{\frac{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}{n^2(n-1)}}\right)$$

In other terms,

- If $n_{ij} \le \hat{E}_{ij}$ (case 1)

$$P_{exact} \approx P\left(\xi^* \le \frac{\hat{D}_{ij}+\frac{1}{2}}{\sqrt{\frac{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}{n^2(n-1)}}}\right) + P\left(\xi^* \ge \frac{\Delta-\hat{D}_{ij}}{\sqrt{\frac{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}{n^2(n-1)}}}\right) \quad (6)$$

- If $n_{ij} \ge \hat{E}_{ij}$ (case 2)

$$P_{exact} \approx P\left(\xi^* \le \frac{\Delta-\hat{D}_{ij}}{\sqrt{\frac{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}{n^2(n-1)}}}\right) + P\left(\xi^* \ge \frac{\hat{D}_{ij}-\frac{1}{2}}{\sqrt{\frac{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}{n^2(n-1)}}}\right) \quad (7)$$

where ξ^* is a Gaussian distribution with zero mean and unit variance, $\hat{D}_{ij} = n_{ij} - \hat{E}_{ij}$, and Δ is a quantity that must be computed as follow:

1) Take the integer part of $2\hat{E}_{ij} - n_{ij}$.
2) Add 0.5 to the above quantity and denote by θ the resulting value.
3) Obtain Δ from the equation $\Delta = \theta - 2\hat{E}_{ij} + n_{ij}$

It is necessary to set two exceptions to the above rule:

In case 1): if $2\hat{E}_{ij} - n_{ij} \le \min\left(n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}\right)$, then $P\left(\xi^* \ge \frac{\Delta-\hat{D}_{ij}}{\sqrt{\frac{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}{n^2(n-1)}}}\right) = 0$.

In case 2): if $2\hat{E}_{ij} - n_{ij} \le 0$, then $P\left(\xi^* \le \frac{\Delta-\hat{D}_{ij}}{\sqrt{\frac{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}{n^2(n-1)}}}\right) = 0$.

Finally, the independence test is carried out by comparing case 1 or case 2 probability (depending on the case) to the level of significance. If the independence hypothesis is rejected, the intensity and direction of the association will be measured by the *Q* of Yule, in the (2x2) case.

$$Q = \frac{n_{11}n_{22} - n_{12}n_{21}}{n_{11}n_{22} + n_{12}n_{21}}. \quad (8)$$

3. Case Study

3.1. Study Site

The case study was carried out in the emblematic old quarter of Toledo, in Spain. Toledo is a UNESCO World Heritage City with an economy driven to a significant extent by commercial activities deriving from tourism.

The city of Toledo (pop. 76,618 as of 2008) is located in central Spain, about 71 km from the capital city of Madrid. Toledo was the imperial capital until 1563 when the court moved to Madrid, and is currently the seat of the regional government of Castilla-La Mancha, one of the most extensive regions in the country. As a medieval city known today as the city of the three cultures due to its Islamic, Christian, and Jewish heritage, Toledo has been highly successful in the task of preserving its historical and architectural character. This accomplishment has been recognised by the UNESCO, which has granted the denomination of World Heritage City. The preservation of centuries-old city walls has meant that Toledo has experienced relatively little expansion, with virtually all recent growth taking place beyond the perimeter of the old city (see Figure 1). Given its historical and cultural interest, its proximity to Madrid and excellent connections with the capital and other regions, which include highways and a new High Speed Train service, Toledo has benefited from a very active tourist sector, and related commercial and service activities. The service sector and tourism in particular, is enormously important for the economy of the city, with the former accounting for about 84.5% of Toledo's gross product, half of which is due to tourism, Ref. 32. The importance of these activities is also reflected by the distribution of the economically active population in the municipality, 86.5% of which is linked to the service sector. It is impossible to offer more updated data because neither the National Statistical Institute nor the Castilla-La Mancha Regional Institute nor the Institutes that study the economic situation of the region (Lawrence Klein Institute and Stone Institute, among others) provide estimates of the contribution of tourism to Toledo's gross product and economically active population (the lower scale for estimates is the regional scale). Fortunately, in Toledo City we can use the useful work of the research group leaded by Prof. Esteban (this is not the case in other tourist Spanish cities), even if they refer to 2005 because the quotas of participation has not varied significantly. Also according to Ref. 32, some 45% of visitors to the city are international travellers. In terms of expenditure in the city, about 80% of all visitors (domestic and international) have a daily budget per person of between 50 and 100 Euros, while about 14% have a budget of between 100 and 200 Euros. Besides accommodation, other significant expenses include food and beverages, clothing and shoes and souvenirs. There are about 11,000 establishments in the city that cater to the demand for these services.

3.2. The dataset

The information has been collected by means of a specially developed questionnaire, the type of survey being a personal interview. It is worth mentioning that all interviews were conducted by PhDs in Statistics. This is a crucial aspect, because only if questions are well understood, will statistical results be coherent. So, having good interviewers that assist the interviewed during the process of filling the questionnaire is a guaranty of quality responses. It is also worth mentioning that questionnaires were elaborated in six languages (Spanish, English, French, Italian,

German, and Chinese) and that interviewers were fluent in at least three of these languages.

Data were collected during the months of April, May and June, 2009. A total of 1,500 questionnaires were completed. Statistical analysis of questionnaires revealed no significant differences among them, this being the reason why the set of polled tourists can be considered a unique sample of tourists.

In order to minimise validity problems in the research, a pilot test was conducted. To this end, after preparing the initial questionnaire, it was given to 80 tourist of different nationalities (Spanish, English, Italian, French, German, New Zealand, Russian, and Asian people). The questionnaire was modified once more based on the suggestions of these people.

Tourists were interviewed in different locations of the old quarter of Toledo city (monuments, squares, restaurants, hotels, etc.) at different hours, to try to collect as wide a range of people and situations as possible. The non-response rate was 6.8%. Most of the non-respondents were Asian people.

Although the population of visitors is unknown, we can infer the appropriateness of the sample from indirect indicators such as percentage of male and female, percentage per age strata, percentage per nationality, etc. The structure of these percentages is very similar to official data (if available).

Finally, as for the questionnaire, we used a transversal strategy. Multiple regression analysis determined questions 1 to 4 (nationality, sex, age and main activity) as optimal for being used as control variables. Questions 5 to 14 (Toledo as main destination or not, reason for the trip, mode of transport, the way they found the information about the city, the length of the staying, type of accommodation if the visitor sleeps in Toledo, main reason for choosing Toledo, which other places in the province of Toledo the visitor has visited or has planning to visit, and what is the information he or she has about Quixote's route) were included to determine the profile of the visitors (a secondary objective in this article). Questions 1 to 14 where used for tourist classification. Question 15 was related with the *a priori* importance of the 20 tourist services included in Table 3. In particular, the tourist was asked to rank in order of importance the 20 services. In question 16, tourist had to asses from 0 (minimum) to 10 (maximum) the degree of satisfaction with the above mentioned services. Finally, in question 17.the visitor is asked if the 20 services met its expectations, the options being qualitative: not exceeded, equal, exceeded. Likert scales were used in both questions 15 and 17. In question 16 we used a rating scale. As for the estimation of the internal consistency of the reliability of test scores, the calculated global and partial Cronbach's alpha exceeded 0.6.

Table 1. Technical Data

Population	About 600,000 tourist
Geographical Area :	Toledo City
Sample size:	1,500
Sample error:	± 5.1%
Significance level:	5%
Sample period:	April, May and June 2009

Source: Own elaboration.

Table 2. Profile of respondents

Variables	%
Gender	
Male	46.10
Female	53.9
Age	
15-24 years	59.03
25-39 years	24.36
40-49 years	8.00
50-65 years	7.39
65 years and older	1.21
Place or origin	
National	76.30
International	23.7
Occupation	
Employed	40.27
Student	53.45
Retired	1.81
House wife	2.06
Unemployed	0.97
Others	1.45
Stay in Toledo	
1day without accommodation	58.28
1 night	14.63
Between 2 and 4 nights	18.86
Between 4 nights and 1 week	3.64
Other terms	4.59
Type of accommodation	
5 star hotel	8.86
4 star hotel	18.56
3 star hotel	24.65
Cheap hotel	16.90
Guest house	3.05
Others(*)	27.98

(*) This category includes lodging in family houses and visitors that do not overnight in the city.
Source: Own elaboration.

4. Results

In this section we first calculate the overall level of tourist satisfaction in Toledo City by using the procedure shown in section 2.1. Subsequently, results from the paired t-test between tourists' expectations and satisfaction with the attributes evaluated are presented and commented when using the alternative proposed in 2.2 regarding attributing marks to prior expectations without enlarging the questionnaire. Finally, the hypothesis of independence between expectations and satisfaction is tested versus the positive version of the Assimilation Theory by using a new version of the Chi-square test that includes an asymmetric correction for continuity.

In reference to the first question, the overall level of tourist satisfaction scores 7.15 (maximum mark is 10), which can be considered an excellent mark. Note that the most important attributes for Toledo visitors are "Tourist attractions and or/

walks" (weight of 10.2/100); "Entertainment/ cultural and recreational activities" and "Food", each weighted 7.3/100; "Accommodation" with 6.1/100, "Shopping", "Cleanliness, hygiene, and public toilets", "Hospitality", "Information", "Security", "Tourist information offices" and "Providers of tourist services" have a weight of around 5/100; and the importance given to the rest of services ranges from 3.6/100 to 4.5/100.

Table 3. Global overall tourist satisfaction

Attribute	λ_i	$\overline{P}$
Accommodation	0.061	6.950
Tourist attractions and or/ walks	0.102	7.683
Entertainment/ cultural and recreational activities	0.073	6.907
Food	0.073	7.390
Shopping	0.047	6.477
Cleanliness. hygiene. public toilets	0.051	6.301
Internet/ Communications/ Phone	0.040	6.127
Public lightning	0.036	6.483
Hospitality	0.054	6.937
Information	0.048	6.863
Security	0.050	6.809
Tourist information offices	0.047	6.560
Providers of tourist services	0.045	6.363
Public transport connections	0.041	6.577
Signposting	0.037	6.303
Banks/ ATM	0.037	6.303
Local transport	0.039	6.250
Environmental care	0.040	6.567
Accessibility	0.041	6.277
Ability to settle problems	0.036	6.120
Global overall tourist satisfaction level		**7.150**

Source: own elaboration.

The equation of the linear regression of the mean satisfaction relative to weight (of attributes) is $\overline{p}_i = 5.49 + 0.22 w_i$, the determination coefficient being 0.8. Therefore, the more important the attribute is for tourists, the higher their level of satisfaction. In particular, a one-unit increase in the scale of importance leads to a rise of 0.22 units in the level of satisfaction with such an attribute.

Once the global overall level of tourist satisfaction has been calculated, we focus our attention on the revisited expectation/disconfirmation paradigm. In this sense, Table 4 reports the main statistical results obtained from the paired *t*-test between tourists' expectations and satisfaction with the 20 attributes evaluated.

Table 4. Paired *t*-test between tourists' expectations and satisfaction with attributes

Attributes	Satisfaction Mean	Expectation Mean	Difference between Means	t-value one tailed-test
Accommodation	6.950 (2.336)	6.510 (2.300)	0.440	28.309*
Tourist attractions and or/ walks	7.683 (2.294)	6.635 (2.521)	1.048	64.819*
Entertainment/ cultural and recreational activities	6.907 (2.102)	6.577 (2.261)	0.330	22.521*
Food	7.390 (2.116)	6.523 (2.343)	0.867	57.846*
Shopping	6.477 (1.840)	6.438 (1.964)	0.038	03.001*
Cleanliness, hygiene, public toilets	6.301 (2.074)	6.268 (2.054)	0.033	02.414*
Internet/ Communications/ Phone	6.127 (1.920)	6.332 (1.959)	-0.205	-15.748*
Public lightning	6.483 (2.025)	6.100 (2.098)	0.383	27.701*
Hospitality	6.937 (2.095)	6.213 (2.259)	0.723	49.470*
Information	6.863 (1.997)	6.458 (2.172)	0.406	28.972*
Security	6.809 (2.105)	6.455 (2.238)	0.354	24.263*
Tourist information offices	6.560 (2.073)	6.120 (2.124)	0.440	31.238*
Providers of tourist services	6.363 (2.148)	6.567 (5.037)	-0.203	-07.824*
Public transport connections	6.577 (5.168)	6.337 (5.200)	0.240	06.898*
Signposting	6.303 (1.982)	6.128 (2.105)	0.175	12.754*
Banks/ ATM	6.303 (2.063)	6.250 (2.124)	0.053	03.795*
Local transport	6.250 (1.887)	6.037 (1.979)	0.213	16.438*
Environmental care	6.567 (1.959)	6.272 (2.064)	0.295	21.846*
Accessibility	6.277 (2.073)	6.202 (2.084)	0.075	05.375*
Ability to settle problems	6.120 (2.003)	6.012 (2.003)	0.108	08.058*

Standard deviations in brackets

* Significant at the 0.05 significance level.

As it can be seen in Table 4, the marks attributed to the tourist services considered in the questionnaire display a certain degree of uniformity. Mean satisfaction marks range from 6.120 to 7.683. As a result, there is at least a medium-high level of satisfaction with the tourist attributes included in the questionnaire.

"Tourist attractions and or/ walks", "Food", "Accommodation", "Hospitality", and "Entertainment/ cultural and recreational activities" are, in this order, the five highest scoring attributes. "Ability to settle problems" and "Internet/ Communications/ Phone" was the two lowest scoring services.

Furthermore, 18 out of the 20 attributes considered are satisfactory (satisfaction exceeds expectations). The only two dissatisfying attributes (satisfaction does not reach expectations) are "Internet/ Communications/ Phone" and "Providers of tourist services". No indifferent attributes (satisfaction is statistically equal to expectations) were found. The most satisfactory attributes are "Tourist attractions and or/ walks", "Food", "Accommodation" and "Hospitality". In the first case, the difference between means exceeds unity, in the second it is practically 0.9 and in the third and fourth it is 0.7 and 0.4, respectively. In case of the two dissatisfying attributes, the difference between means is -0.2. Regarding this, it is necessary to point out that for large samples (as in this case), the *t*-test detects very small differences between expectation and satisfaction means. This is why attributes such as "Shopping" "Cleanliness, hygiene, public toilets", "Banks/ ATM", and "Accessibility" are statistically satisfactory even though their satisfaction mean is practically the same as the expected mean.

Finally, when we deal with testing the hypothesis of independence between expectations and satisfaction versus the positive version of the Assimilation Theory by using the new version of the chi-square test that includes an asymmetric correction for continuity (Table 5), it can be observed that in all cases, except "Entertainment/ cultural and recreational activities", "Public lightning ", and "Signposting", the hypothesis of independence between expectancy (low, high) and disconfirmation (negative, positive) has been rejected at a 0.05 significance level. In the rest of cases, Yule's Q value is positive. Therefore, positive association exists, which indicates the relationship of high expectations with slight positive disconfirmation and low expectations with slight negative disconfirmation. That is, positive association implies a slight exacerbation of expectations and favours the Assimilation Theory. A factor confirmation of positive association has been computed as the ratio "percentage of respondents with both slight high or low expectations/ percentage of total respondents" to avoid cases in which positive association is based on a small percentage of respondents with slight high or low expectations. As can be observed in Table 5, all confirmation factors are above 60%, except in the attribute relative to "Providers of tourist services". Therefore, in this case, despite having detected positive association, this result should be considered with caution because the evidence that supports the positive association is not as much as desirable.

Table 5. A chi-squared test with an asymmetric correction for continuity

$n_{11} > \hat{E}_{11}$ (all cases)				Yule's Q	Confirmation of the Assimilation Theory	Confirmation factor
Attribute	$P^{(1)}$	$P^{(2)}$	Exact Probability			
Accommodation	0.0000	0.0000	0.0000*	0.60	Yes	0.72
Tourist attractions and or/ wa	0.0000	0.0000	0.0000*	0.34	Yes	0.83
Entertainment/ cultural and recreational activities	0.0901	0.0025	0.1151	0.11	No	0.65
Food	0.0000	0.0000	0.0000*	0.34	Yes	0.82
Shopping	0.0006	0.0010	0.0016*	0.34	Yes	0.69
Cleanliness, hygiene, public toilets	0.0150	0.0174	0.0324*	0.24	Yes	0.65
Internet/ Communications	0.0011	0.0019	0.0030*	0.35	Yes	0.61
Public lightning	0.0793	0.0869	0,1662	0.16	No	0.70
Hospitality	0.0000	0.0000	0.0000*	0.33	Yes	0.75
Information	0.0026	0.0025	0.0051*	0.31	Yes	0.76
Security	0.0000	0.0000	0.0000*	0.42	Yes	0.64
Tourist information offices	0.0019	0.0020	0.0039*	0.30	Yes	0.78
Providers of tourist services	0.0000	0.0000	0.0000*	0.83	Yes	0.27
Public transport connections	0.0009	0.0013	0.0022*	0.29	Yes	0.61
Signposting	0.0256	0.0287	0.0543	0.21	No	0.65
Banks/ ATM	0.0122	0.0192	0.0314*	0.24	Yes	0.63
Local transport	0.0000	0.0000	0.0000*	0.41	Yes	063
Environmental care	0.0011	0.0011	0.0022	0.31	Yes	0.68
Accessibility	0.0006	0.0011	0.0017*	0.36	Yes	0.67
Ability to settle problems	0.0000	0.0004	0.0004*	0.43	Yes	0.63

* Significant at 0.05 significance level.

$$p^{(1)} = P\left(N_{ij} \le 2E_{11} - n_{11} + \Delta\right); \quad p^{(2)} = P\left(N_{11} \ge n_{11} - \frac{1}{2}\right); \quad \hat{E}_{11} = \frac{n_1 n_{\cdot 1}}{n}$$

5. Concluding remarks

Cultural/heritage tourism is one of the fastest growing segments of the tourism industry, and obviously, the growth in the cultural/heritage tourism market may provide several benefits to cultural/heritage destinations. Of course, the uncontrolled increasing of the tourist activity in general and cultural/heritage tourism in particular, can lead to undesirable consequences in case that the capacity of the destination is exceeded and/or its tourist development is not sustainable.

Because of people's inclination to seek out novel attractions, including traditional cultures, heritage tourism has become a major "new" area of tourism demand, which almost all policy–makers are now aware of and anxious to develop. Heritage tourism, as part of the broader category of "cultural tourism", is now a major pillar of the nascent tourism strategy of many countries. Cultural/heritage tourism strategies in various countries share the fact that they are a major growth area, that they can be used to boost local culture and that they can aid the seasonal and geographic spread of tourism.

The evaluation of tourist satisfaction needs to be considered in multiple dimensions. But this is a large and space-consuming task, which is why this paper has focused on three novel aspects of the expectation/disconfirmation model: The way individual and overall tourist satisfaction is computed, the way marks are given to previous expectations without enlarging the questionnaire and the use of a chi-square test including an asymmetric correction for continuity to test what we call the positive version of the Assimilation Theory in tourist attributes.

These three novelties have been incorporated into a Spanish case study: The emblematic old quarter of Toledo, which is a UNESCO World Heritage City with an economy driven to a significant extent by commercial activities deriving from tourism.

The information used has been collected by means of a specially developed questionnaire, the type of survey being a personal interview. It is worth mentioning that all interviews were conducted by PhDs in Statistics. Data were collected during the months of April, May and June, 2009. A total of 1,500 questionnaires were prepared, the non-response rate being 6.8%.

The main conclusions relative to the inclusion of the methodological novelties to measure the overall level of tourist satisfaction, to revisit the expectation/disconfirmation paradigm and to test the positive version of the Assimilation Theory are as follows:

(1) The overall level of tourist satisfaction scored 7.15 (maximum mark is 10), which can be considered an excellent mark. The most important attributes for Toledo visitors are, in this order, "Tourist attractions and or/ walks"; "Entertainment/ cultural and recreational activities", "Food", and "Accommodation". The linear regression of the mean satisfaction relative to the weight of the considered attributes indicates that the more important an attribute is for tourists, the higher their level of satisfaction. In particular, an increase of one unit in the scale of importance leads to a rise of 0.22 in the level of satisfaction with such an attribute.

(2) The marks attributed to the tourist services considered in the questionnaire display a certain degree of uniformity. Mean satisfaction marks range from 6.120 to 7.683. As a result, there is at least a medium-high level of satisfaction with the tourist attributes included in the questionnaire.

Furthermore, 18 out of the 20 attributes considered are satisfactory (satisfaction exceeds expectation). The only two dissatisfying attributes (satisfaction does not reach expectations) are "Internet/ Communications/ Phone" and "Providers of tourist services". The most satisfactory are "Tourist attractions and or/ walks", "Food", "Accommodation" and "Hospitality".

(3) In all cases, except "Entertainment/ cultural and recreational activities", "Public lightning", and "Signposting", the hypothesis of independence between expectations (low, high) and disconfirmation (negative, positive) was rejected at a 0.05 significance level. In the rest of cases, Yule's Q value is positive. That is, the positive version of the Assimilation Theory could be at work. All confirmation factors are above 60%, except in the attribute relative to "Providers of tourist services". Therefore, in this case, despite having detected positive association, this result should be considered with caution.

(4) To the extent that the positive version of the Assimilation Theory is favoured by the empirical evidence, and taking into account that, as expected in a cultural/heritage context, most of the tourist have high expectations, satisfaction could be overestimated. This means that the real mark for the satisfaction level with tourist attributes in Toledo could be slightly fewer than 7.

Acknowledgements

The authors thank to Guido Ferrari, Jean Pierre Levy, José Mondéjar and two anonymous reviewers for comments and suggestions about this paper. The authors are responsible for any remaining errors.

This research has been partially founded for Junta de Comunidades de Castilla-La Mancha, under FEDER research project PAI-05-021.

References

1. C. Cooper, J. Fletcher, S. Wanhill and D. Gilbert. Tourism*:* Principles and Practice. (3rd ed.). Harllow: Pearson Education (2008).
2. J. Huh. "Tourist satisfaction with cultural/heritage sites: The Virginia Historic Triangle". Thesis for the degree of Master of Sciences in Hospitality and Tourism Management (Virginia Polytechnic Institute and State University, 2002).
3. G. Richards. Production and consumption of cultural tourism in Europe. *Annals of Tourism Research*, **23, 2** (1996) 261–283.
4. M. Uysal. and F. Noe. Satisfaction in Outdoor Recreation and Tourism Settings. In Eric Laws (Ed.): *Case Studies in Tourism Marketing* (Continuum Publisher. London, 2003), pp. 144–158.
5. L. A. Crosby. Measuring customer satisfaction in the *Service Quality Handbook* (eds). E. Scheuing, W. Christopher (AMACOM, New York, 1993), pp. 389–407.

6. J. S. Akama and D. Mukethe-Kieti. Measuring tourists' satisfaction with Kenya safari. *Tourism management*, **24** (2002) 73–81.

7. S. Haber, S. and M. Lerner Correlates of Tourist Satisfaction. *Annals of Tourism Research*. **25, 4**, (1998) 197–201.

8. P. Conklin, K. Powaga and S. Lipovetsky. Customer satisfaction analysis: Identification of key drivers. *European journal of operational research*, **154** (2002) 819– 827.

9. D. A. Francken and W. J. Van Raaij. Satisfaction with leisure time activities. *Journal of Leisure Research*, **13**, **4** (1981) 337–352.

10. A.Serenko and A. Stach. The Impact of Expectation Disconfirmation on Customer Loyalty and Recommendation Behavior: Investigating On Line Travel and Tourism Services. *Journal of Information Technology Management*, **20, 3** (2009) 26–41.

11. R. L. Oliver and J. E. Swan. Consumer perceptions of interpersonal equity and satisfaction in transactions: A field survey approach. *Journal of Marketing*, **53** (1989) 21–35.

12. R. P. Fisk and C. E.Young. Disconfirmation of Equity Expectations: Effects on Consumer Satisfaction with Services. In *Advances in Consumer Research*, E. Hirschman and H. Holbrook (Eds.). Provo, UT: Association for Consumer Research (1985).

13. S.A. LaTour and N.C Peat. Conceptual and Methodological Issues in Consumer Satisfaction Research. In William. L.W. (Eds.) *Advances in Consumer Research* Ann Arbor, Ml: Association for Consumer Research (1979) 431–437.

14. E. R. Cadotte, R. B. Woodruff, and L. J. Roger. Expectations and Norms in Models of Consumer Satisfaction. *Journal of Marketing Research*, **24** (1987) 305–314.

15. D. K. Tse and P. C. Wilton. Models of Consumer Satisfaction Formation: An Extension. *Journal of Marketing Research*, **25** (1988) 204–212.

16. C.F. Chen. Experience Quality, Perceived Value, Satisfaction and Behavioral Intentions for Heritage Tourists. *Proceedings of the 13th Asia Pacific Management Conference*, Melbourne, *Australia (2007) 1130–1136.*

17. R. L. Oliver. Cognitive, Affective, and Attribute Bases of the Satisfaction Response. Journal of Consumer Research, **20** (1993) 418–430.

18. Y. Yu and A. Dean. The Contribution of Emotional Satisfaction to Consumer Loyalty. *International Journal of Service Industry Management*, **12** (2001) 234–250.

19. W. Van Dolen, K. De Ruyter and J. Lemmink. An Empirical Assessment of the Influence of Customer Emotions and Contact Employee Performance on Encounter and Relationship Satisfaction. *Journal of Business Research*, **57** (2004) 437–444.

20. J. Rodríguez, I. and H. San Martín. Tourist Satisfaction. A congnitive-afective model. *Annals of Tourism Research*, **35**, **2** (2008) 551–573.

21. D. Phillips and H. Baumgarther. The Role of Consumption Emotions in the Satisfaction Response. *Journal of Consumer Psychology*, **12** (2002) 243–52.

22. E. Bigné, L. Andreu and J. Gnoth. The Theme Park Experience: An Analysis of Pleasure, Arousal and Satisfaction. *Tourism Management*, **26** (2005) 833–844.

23. A. Zeithaml, A. Pasuraman and L. Berry. Delivering Quality Service: Balancing Customer Perceptions and Expectations (New York: The Free Press Division of Macmillan, Inc, 1990).

24. A. Pizam, Y. Numann, and A. Reichel,. Dimensions of tourist satisfaction with a destination area. *Annals of Tourism Research*, **5** (1978) 314–322.

25. J. Tribe and T. Snaith. From SERVQUAL to HOLSAT: Hollyday satisfaction in Varadero, Cuba. *Tourism Management*, **19**, **1** (1998) 25–34.

26. M. Sherif and C. I. Hoyland. Social judgment: Assimilation and contrast effects in communication and attitude change (New Haven: Yale University Press, 1961).

27. R. L. Oliver. Satisfaction: A Behavioral Perpective on the Consumer (Mc Graw- Hill, New York 1997).

28. J. M. Montero Lorenzo. Una propuesta de corrección de continuidad asimétrica para tablas de contingencia (2x2) con totales marginales fijos. *Estadística Española*, **44**, **149** (2002) 29–46.

29. N. Mantel. Some reasons for not using the Yates continuity correction in 2x2 contingency tables. Comment and a suggestion. *Journal of the American Statistical Association*, **69** (1974) 378–380.

30. W. G. Cochran. The 2x2 Correction for Continuity. Iowa State College. *Journal of Science*, **16** (1942) 421–436.

31. F. Yates. Contingency tables involving small numbers and the chi-squared test. *Journal of the Royal Statistical Society*, **1** (1934) 217–235.

32. A. Esteban, D. Martín-Consuegra, A. Molina and E. Díaz. *Turismo y consumo: El caso de Toledo*. Centro de Estudios de Consumo de la Universidad de Castilla–La Mancha y de la Junta de Comunidades de Castilla–La Mancha. Working paper (Toledo, 2005).

Appendix: **Classical continuity corrections for (2x2) contingency tables with fixed margins**

The way to carry out a Fisher exact test of independence is to compute the probability (under the null hypothesis, of obtaining the observed table and the other tables that indicates a separation from the independence hypothesis at least as large as the evidenced in the observed table. This probability (known as *p*-value or exact probability) is then compared to the significance level (α). In case that the *p*-value does not exceed the significance level, the independence hypothesis is rejected.

In particular, in a 2x2 contingency table the exact probability can be written as:

$$P\left(\left|N_{ij}-\hat{E}_{ij}\right| \geq \left|n_{ij}-\hat{E}_{ij}\right|\right),\ \forall\, ij\,,$$

where n_{ij} is the observed frequency in the cell $\{ij\}$ and N_{ij} follows a hypergeometric probability distribution because the margins are fixed. And the continuous approximation is obtained as:

$$P\left[\chi^2_{(1)} \geq \frac{(n-1)(n_{11}n_{22}-n_{12}n_{21})^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}\right]$$

However, the computation of the exact probability is not easy. What is more, it is a tedious and time consuming task. This is the reason why the traditional approach to this

problem is to approximate the exact probability by using the Chi-square statistic. But an important problem arises when this approximation is used: the exact probability for the case of a 2x2 contingency table with fixed margins must be computed by using the hypergeometric probability distribution (which is discrete), whereas the approximation to this exact probability is computed with a Chi-square probability distribution (which is continuous). Thus, it is needed to include a correction for this mismatch (overall in case that some of the expected frequencies do not exceed 5).

Yates continuity correction

Taking as starting point the adjusted Chi squared statistics,

$$\chi^2_{ajd} = \frac{(n-1)(n_{11}n_{22} - n_{12}n_{21})^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}} = \frac{(n-1)}{n}\sum_i\sum_j \frac{\left(n_{ij} - \hat{E}_{ij}\right)}{\hat{E}_{ij}}$$

Yates proposed a continuity correction that consists in subtracting 0.5 from de positive differences between the observed and the expected (under the null) frequencies, and adding 0.5 in case that these differences are negative. That is to say, to subtracting 0.5 from the absolute value of the differences between the observed and the expected frequencies. Then, the approximation to the exact probability is obtained by:

$$P\left[\chi^2_{(1)} \geq \frac{(n-1)}{n}\sum_i\sum_j \frac{\left(\left|n_{ij} - \hat{E}_{ij}\right| - 0.5\right)^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}\right],$$

$$P\left[\chi^2_{(1)} \geq \frac{(n-1)\left(\left|n_{11}n_{22} - n_{12}n_{21}\right| - 0.5n\right)^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}\right]$$

or

$$P\left[\chi^2_{(1)} \geq \frac{n^2(n-1)}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}\left(\left|n_{ij} - \hat{E}_{ij}\right| - 0.5\right)^2\right]$$

Next we focus in the rationale of the above expressions. Let N_{11} be a random variable whose value determines the structure of the 2x2 contingency table because the margins are fixed. As said before, N_{11} follows a hypergeometric distribution, so that its expectation and variance are

$$E(N_{11}) = \frac{n_{1\cdot}n_{\cdot 1}}{n}$$

$$V(N_{11}) = \frac{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}{n^2(n-1)}$$

As previously stated, the independence hypothesis is rejected in favour of the alternative in case that

$$P\left(\left|N_{11} - \hat{E}_{11}\right| \geq \left|n_{11} - \hat{E}_{11}\right|\right) \leq \alpha$$

To include the continuity correction proposed by Yates and firstly assuming that the difference $\left(N_{11}-\hat{E}_{11}\right)$ is positive, the null hypothesis will be rejected in case that

$$P\left(\left|N_{11}-\hat{E}_{11}\right|\geq\left|n_{11}-\hat{E}_{11}\right|-0.5\right)\leq\alpha$$

$$P\left[\left(N_{11}-\frac{n_{1\bullet}n_{\bullet 1}}{n}\right)^2\geq\left(n_{11}-\frac{n_{1\bullet}n_{\bullet 1}}{n}-0.5\right)^2\right]\leq\alpha$$

Dividing in both sides of the inequality by the variance of N_{11} we have

$$P\left[\frac{\left(N_{11}-\frac{n_{1\bullet}n_{\bullet 1}}{n}\right)^2}{\frac{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}{n^2(n-1)}}\geq\frac{\left(n_{11}-\frac{n_{1\bullet}n_{\bullet 1}}{n}-0.5\right)^2}{\frac{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}{n^2(n-1)}}\right]\leq\alpha$$

and the null hypothesis will be rejected in case that

$$P\left[\chi^2_{(1)}\geq\frac{\left(n_{11}-\frac{n_{1\bullet}n_{\bullet 1}}{n}-0.5\right)^2}{\frac{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}{n^2(n-1)}}\right]\leq\alpha$$

$$P\left[\chi^2_{(1)}\geq\frac{(n-1)\left(n_{11}n_{22}-n_{12}n_{21}-0.5n\right)^2}{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}\right]\leq\alpha$$

$$P\left[\chi^2_{(1)}\geq\frac{(n-1)\left(\left|n_{11}n_{22}-n_{12}n_{21}\right|-0.5n\right)^2}{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}\right]\leq\alpha$$

because

$$\left(n_{11}n_{22}-n_{12}n_{21}\right)=\hat{D}_{11}=\hat{D}_{22}=-\hat{D}_{12}=-\hat{D}_{21},$$

where

$$\hat{D}_{ij}=n_{ij}-\hat{E}_{ij},\quad\text{and}\quad\hat{D}_{ij}>0,$$

and thus it makes no difference to deal or not with absolute values

In case that $\hat{D}_{ij}=\left(n_{ij}-\hat{E}_{ij}\right)\leq 0$, Yates proposed to add 0.5 to such a difference, the null hypothesis being rejected when

$$P\left[\left(N_{11}-\hat{E}_{11}\right)^2\geq\left(n_{11}-\hat{E}_{11}+0.5\right)^2\right]\leq\alpha$$

Proceeding the same way than in case that $\hat{D}_{11}=\left(n_{11}-\hat{E}_{11}\right)\leq 0$, it can be concluded the rejection of the null in case that

$$P\left[\chi^2_{(1)} \geq \frac{(n-1)\left(\left|n_{11}n_{22}-n_{12}n_{21}\right|+0.5n\right)^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}\right] \leq \alpha$$

But taking into account that

$$\hat{D}_{11} = n_{11} - \hat{E}_{11} = n_{11}n_{22} - n_{12}n_{21} < 0$$

we have

$$\left(n_{11}n_{22}-n_{12}n_{21}+0.5n\right)^2 = \left(\left|n_{11}n_{22}-n_{12}n_{21}\right|-0.5n\right)^2$$

Thus, the null will be rejected if

$$P\left[\chi^2_{(1)} \geq \frac{(n-1)\left(\left|n_{11}n_{22}-n_{12}n_{21}\right|-0.5n\right)^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}\right] \leq \alpha$$

In conclusion, irrespective of the sign of $\hat{D}_{11} = \left(n_{11} - \hat{E}_{11}\right)$, the hypothesis of independence will be rejected in case that

$$P\left[\chi^2_{(1)} \geq \frac{(n-1)\left(\left|n_{11}n_{22}-n_{12}n_{21}\right|-0.5n\right)^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}\right] \leq \alpha$$

The statistic

$$\chi^2_{ajd,Y} = \frac{(n-1)\left(\left|n_{11}n_{22}-n_{12}n_{21}\right|-0.5n\right)^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}$$

$$= \frac{(n-1)}{n}\sum_i\sum_j \frac{\left(\left|n_{ij}-\hat{E}_{ij}\right|-0.5\right)^2}{\hat{E}_{ij}}$$

$$= \frac{(n-1)n^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}\left(\left|n_{ij}-\hat{E}_{ij}\right|-0.5\right)^2 \quad \forall_{ij}$$

is known as the Chi-squared statistic with the Yates continuity correction.

The main disadvantage of the Yates continuity correction is that it can lead to large underestimates or overestimates of the exact probability (unless the expected frequency $\hat{E}_{11}$ is a multiple of 0.5)

First Cochran continuity correction

The way to include the first Cochran continuity correction in the adjusted Chi-square statistic is as follows. First, take the observed frequency corresponding to the smallest expected frequency under the null hypothesis. Two cases can be distinguished.

$$A)\ n_{ij} \leq 2\hat{E}_{ij}$$

$$B)\ n_{ij} > 2\hat{E}_{ij}$$

where $\hat{E}_{ij}$ is expected frequency previously mentioned and n_{ij} the observed frequency in the selected cell (*i*,*j*).

A) $n_{ij} \leq 2\hat{E}_{ij}$

Let $\chi^2_{ajd,0}$ be the observed adjusted Chi square statistic and $\chi^2_{ajd,1}$ the maximum value of χ^2_{ajd} that does not exceed $\chi^2_{ajd,0}$.

Then the first proposal of Cochran is to compute the arithmetic mean of the square root of the above mentioned statistics, $\chi_{ajd,C1}$:

$$\chi_{ajd,C1} = \frac{\sqrt{\chi^2_{ajd,0}} + \sqrt{\chi^2_{ajd,1}}}{2}$$

and to approximate the exact probability by

$$P\left[\chi^2_{(1)} \geq \chi^2_{ajd,C1}\right]$$

B) $n_{ij} > 2\hat{E}_{ij}$

In this case $\chi^2_{ajd,1}$ is defined as:

$$\chi^2_{ajd,1} = \frac{n^2(n-1)}{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}\left(\left|n_{ij} - \hat{E}_{ij}\right| - 1\right)^2$$

Then, it is computed $\chi_{ajd,C1} = \frac{\sqrt{\chi^2_{ajd,0}} + \sqrt{\chi^2_{ajd,1}}}{2}$ and the exact probability is approximated by

$$\frac{1}{2}P\left[\chi^2_{(1)} \geq \chi_{ajd,C1}\right]$$

Second Cochran continuity correction

To implement the second continuity correction proposed by Cochran, $\chi^2_{ajd,0}$ and $\chi^2_{ajd,1}$ are defined as in the case of the first correction (depending on whether we are in case A or B), and the Chi square statistics that includes the second Cochran continuity correction is defined as:

$$\chi_{ajd,C2} = \frac{\chi^2_{ajd,0} + \chi^2_{ajd,1}}{2}$$

Then the approximation of the exact probability is given by

$$\frac{1}{2}P\left[\chi^2_{(1)} \geq \chi^2_{ajd,C2}\right] \quad si \quad n_{ij} > 2\hat{E}_{ij}$$

$$P\left[\chi^2_{(1)} \geq \chi^2_{ajd,C2}\right] \quad si \quad n_{ij} > 2\hat{E}_{ij}$$

Mantel continuity correction

As in the first and second Cochran continuity corrections, the Mantel correction also distinguishes the cases:

$$A)\ n_{ij} \le 2\hat{E}_{ij}$$

$$B)\ n_{ij} > 2\hat{E}_{ij}$$

where $\hat{E}_{ij}$ is expected frequency previously mentioned and n_{ij} the observed frequency in the selected cell (*i*,*j*).

In case that $n_{ij} \le 2\hat{E}_{ij}$ the way of proceeding is as follows:

First, the exact probability is approximated by using the Chi square statistic that include the Yates continuity correction. Let P_Y be such an approximation.

$$P_Y = P\left[\chi^2_{(1)} \ge \chi^2_{ajd,CY}\right]$$

Second, DF_{ij} is defined as:

$$DF_{ij} = \left|n_{ij} - \min \hat{E}_{ij}\right|$$

and DF^*_{ij} as the closest difference $\left|n_{ij} - \min \hat{E}_{ij}\right|$ that exceeds DF_{ij}. At these effects, n_{ij} is the observed frequency corresponding to the smallest expected frequency under the null hypothesis

Finally, it is computed the probability P^*_Y

$$P^*_Y = P\left[\chi^2_{(1)} \ge \frac{n^2(n-1)}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}\left(DF^*_{ij} - \frac{1}{2}\right)^2\right]$$

and the exact probability is approximated by the arithmetic mean P_Y y P^*_Y.

$$P_M = \frac{P_Y + P^*_Y}{2}$$

In case that $n_{ij} > 2\hat{E}_{ij}$ the exact probability is approximated by

$$P_M = \frac{P_Y}{2}$$