



ИЗГРАЖДАНЕ И ВАЛИДАЦИЯ НА ДАННИ В ОНЛАЙН БАЗИРАНИ ВЪПРОСНИЦИ В СФЕРАТА НА ТУРИЗМА

BUILDING AND VALIDATION OF DATA IN ONLINE BASED QUESTIONNAIRES IN TOURISM FIELD

Антония Лазарова
Antonia Lazarova

Европейско висше училище по икономика и мениджмънт
European Higher School of Economics and Management

Abstract: The main objective of the scientific report is to present the possibilities for precise data building and validation of data from online based questionnaires in the tourism sector. They are used to create objective opportunities for exploring the perspectives for promoting of Bulgaria as a desirable and recognizable destination within not only summer and winter seasons.

Keywords: data validation, online based questionnaires, tourism sector.

JEL: C5; C52

ВЪВЕДЕНИЕ

Изграждането на онлайн базирани въпросници в сферата на туризма обикновено е насочено към търсене на информация за броя на чуждестранните туристи, за предпочитаните от тях дестинации със съответния им ценови диапазон. Търсенето на изменения в абсолютните стойности на тези параметри не валидира данните в проучването, а дава обща информация за текущо състояние, чрез която е трудно да се предвидят темпът и прирастът на развитие в туристическия сектор. Ето защо възниква обективната необходимост да се представят двата подхода, които традиционно се използват в страната ни при изграждане и валидация на данни в онлайн базирани въпросници в сферата на туризма, като се посочат онези възможни допирни точки, чрез които те могат да бъдат интегрирани успешно, така че да се разкрият възможностите за развитие на сектора на база на реалното му, текущо състояние.

Подходи при изграждане на онлайн базирани въпросници в сферата на туризма

Предлаганите за разглеждане подходи при изграждане на онлайн базирани въпросници в сферата на туризма не са принципно нови, но трудно се намират реални, статистически значими допирни изследователски точки между тях. Основният проблем при употребата им е, че се ползват самостоятелно и частично за различни изследователски цели, боравейки само с абсолютните стойности на параметри като „най-голям абсолютен брой посещения на чуждестранни туристи“, „най-голямо абсолютно нарастване в броя на посещенията на чуждестранни туристи“ и „брой категоризирани места за настаняване с 10 и повече легла“ (Ministerstvo na turizma na Republika na Balgaria, 2018), без да се изготвя регресионен и/или факторен анализ.

В този контекст може да се представи валидната нормативна уредба за прилагане на първия подход, а именно – систематизиран статистически подход на база на данни от Главна дирекция „Туристическа политика“, като се базираме на следния нормативно обусловен конструкт:

– Регламент 1051/2011 на Европейската комисия за прилагане на Регламент 692/2011 на Европейския парламент и Съвета относно европейска статистика на туризма по отношение на структурата на отчетите за качеството и предаване на данните (The European Commission, 2011);

– Регламент 692/2011 на Европейския парламент и Съвета относно европейска статистика на туризма (The European Commission, 2011);

– Закон за туризма (Ministerstvo na turizma na Republika Bgaria, 2013);

– Закон за достъп до обществена информация (Ministerstvoto na transporta, informatsionnite tehnologii i saobshteniyata na Bgaria 2000);

– Закон за статистиката (Ministerstvoto na transporta, informatsionnite tehnologii i saobshteniyata na Bgaria 2000);

– Закон за предотвратяване и установяване на конфликт на интереси (ЗПУКИ) (Ministerstvo na zemedelieto i hranite na Bgaria, 2010).

При прилагането на всички тези нормативни документи следва да се има предвид, че подхода на практика предполага удовлетворяване на изискванията на сателитните сметки за туризма “TSA – Recommended Methodology and Framework”, които са съгласувани между Световната организация по туризъм, Евростат и ООН. В този контекст е редно да се отбележи, че се осигурява съпоставимост на данните в международен мащаб, еднаквост на прилаганите понятия и еднаква рамка за изграждане на ССТ.

Обхватът на изследване в туристическия сектор и данните за неговата динамика се изследва чрез пътуванията на български граждани в чужбина и посещения на чужденци в България. На практика статистическото представяне на данни касае всички чужди граждани, посетили страната ни по страни, по месеци и по цел на посещението, както и всички български граждани, пътуващи до други страни по месеци и по цел на пътуването. На изследване подлежат също местата за настаняване, като изследванията се провеждат в рамките на летния и в рамките на зимния сезон и се представят чрез сравнителен анализ като изменение за периода. Промяната на методологията през 2011 година, обаче на база на Директива 95/57 на Европейската комисия (The European Commission, 1995) води след себе си проблеми при съпоставимостта за дългосрочен период.

Пренощуващите лица, реализираните от тях нощувки, приходите от нощувки, легла и легло-денонощия в експлоатация, броят на местата за настаняване също са обект на изследване в България, но промяната на методологията и обхватът на нормативните документи също се изменят в края на 2011 година с Директива 95/57 на Европейската комисия (The European Commission, 1995), като започва от началото на 2012 година да се прилага Регламент 692/2011 на Европейския парламент и Съвета относно европейска статистика на туризма (The European Commission, 2008) и Регламент 1051/2011 на Европейската комисия за прилагане на Регламент 692/2011 на Европейския парламент и Съвета относно европейска статистика на туризма (The European Commission, 2011) по отношение структурата на отчетите за качеството и предаването на данните, което води до липса на съпоставимост и до методическа промяна във връзка с представените данни.

Основен недостатък при прилагането на този подход въпреки безспорните му макроикономически статистически достойнства е липсата на яснота за удовлетвореността на потребителите след престоя им в страната ни или след посещение на конкретна туристическа дестинация, както и динамичната нормативна уредба, чиято промяна води до липса на съпоставимост на данните в дългосрочен план.

В този контекст е редно да се представи и вторият подход, като условно можем да го наречем нормативно ориентиран подход с фокус, насочен към потребителя, чиито нормативни параметри за реалното му приложение са, както следват:

– Директива 2008/122/ЕО на ЕП и Съвета относно защитата на потребителите по отношение на някои аспекти на договорите за разпределено във времето ползване на собственост, дългосрочни ваканционни продукти, препродажба и замяна (The European Commission 2008);

– Директива 97/7/ЕО на Европейския парламент и на Съвета относно защитата на потребителя по отношение на договорите от разстояние (The European Commission, 1999);

– Директива 93/13/ЕО на Съвета относно неравноправните клаузи в потребителските договори (The European Commission, 1999);

– Директива 84/450/ЕИО на Съвета от 10 септември 1984 година за сближаване на законовите, подзаконовите и административните разпоредби на държавите членки, свързани със заблуждаваща реклама (The European Commission, 1984);

– Закон за защита на потребителите (Ministerstvo na ikonomikata, 2005).

В рамките на втория подход се предполага логичен нормативен конструкт, който да послужи за целево проучване на потребителската удовлетвореност, но в тази му пълнота той рядко се прилага. Основният проблем се корени във факта, че никъде не се допуска връзка с първия подход и на практика двата подхода се използват самостоятелно.

Допирателната между двата подхода следва да се разглежда на база постановката за търсене на абсолютни стойности и клиентска удовлетвореност в рамките на едно изследване. Основната цел на този интегриран подход е да се създадат обективни възможности за реално прогнозиране на развитието на туристическия сектор, да се открият онези места, които са проблемни и да се търсят пътища за разреждане на проблемите в следващ период.

Фиг. 1. Примерен макет на онлайн базирана анкетна карта в сферата на туризма

Feedback server		Туристическа дестинация/Хотел			
		Населено място			

↓

Feedback server	Закон за защита на личните данни	Продължителност на престоя/ Цена на престоя			
		Вид на резервацията			

↓

Feedback server	Закон за защита на личните данни	Дихотомен признак	Пол	Мъж	Жена
		Метриран	Възраст	Допустима долна граница на анкетиране 18 години	
		прекъснат признак	Брой настанени лица	Скалата се изгражда в съответствие с изискванията на Закона за статистиката.	
		Интервална скала			

↓

1. Целта на Вашето пътуване до това населено място е/ беше?	а) Пътувам/е във връзка с работен ангажимент;
	б) Пътувам/е, за да се запозная с известни забележителности в региона;
	в) Пътувам/е, защото периода съвпада с лятната/ зимната ми отпуска;
	г) Пътувам/е в търсене на нови преживявания и адреналин;
	д) Избрали сме мястото, заради ограничен бюджет;
	е) Друго..... Моля, посочете.....
(Можете да посочите повече от един отговор)	
2. Кое Ви направи най – силно впечатление при Вашият престой тук?	а) Спокойната и ведра атмосфера, дружелюбните хора;
	б) Качественото обслужване и ниските цени на храни и напитки;
	в) Възможностите да се докосна до традициите и миналото;
	г) Денонощният шум, липсата на хигиена;
	д) Ниското качество на обслужването и завишените цени;
	е) Наличието на нелюбезен персонал на повечето места, които посетих/ме;
ж) Друго..... Моля, посочете.....	
(Можете да посочите повече от един отговор)	
3. Чувствате ли се удовлетворени от престоя си?	а) Категорично да;
	б) По-скоро да;
	в) Не мога да преценя;
	г) По-скоро не;
д) Категорично не.	
4. Моля, споделете Вашите мнения и препоръки, какво да променим, за да повишим Вашата удовлетвореност?	

Няколко основни фактора следва да се вземат под внимание при избора на платформа за въпросника в частта му, представена в Интернет:

- възможно най-ниски разходи;
- 24-часово присъствие на въпросника в интернет пространството;

- лесна поддръжка;
- вградени функционалности за изграждане на въпроси;
- създаване на възможност за избор на повече от 1 отговор на определени въпроси от страна на анкетираните лица;
- вградени функционалности за събиране на базата данни с резултати;
- съвместимост с мобилни устройства;
- осигуряване на избор на визуално оформление на въпросника;
- възможност за публичен достъп до въпросника;
- гъвкави възможности за експорт на данни.

Положителните страни на този подход от статистическа гледна точка са:

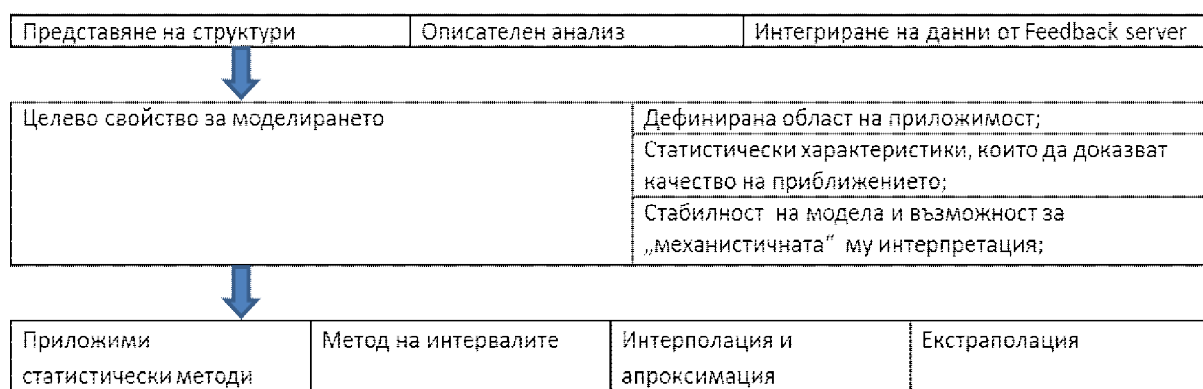
- яснота и статистическа прецизност на скалите;
- възможности за провеждане на навременен диагностичен и номографски анализ;
- създаване на реални възможности за съпоставимост на данните в дългосрочен и средносрочен период;
- създаване на възможности за работа с малки по обем извадки;
- създаване на възможности за факторен и регресионен анализ;
- създаване на възможности за избягване на Парадокс на Симпсън (Lazarova, 2018, pp. 69–75).

Основен недостатък на този подход в сферата на туризма е стремежът да се разширява броят въпроси в рамките на анкетната карта, така че да я лишават от удобството, което краткостта носи на потребителите при попълването ѝ. Другият основен проблем, който е често срещан в практиката, е валидацията на данните в рамките на изследването, като доста често това се схваща като натрупване на данни, без оглед дали те отговарят на законовите изисквания и на Закона за големите числа.

Валидация на данни в онлайн базирани въпросници в сферата на туризма

При изграждане на онлайн базирани въпроси в сферата на туризма за постигане на валидация на данните следва да се има предвид, че въпросникът трябва да се развива логично – от общо към частно. Това предполага да създадем възможност, чрез която да ограничим типа данни, които се въвеждат на определени места.

От гледна точка на изграждането на анкетната карта това на практика означава строго спазване на логиката на използваните статистически скали и признаци и невъвеждане на несвойствени отговори и стойности за съответната скала.



Фиг. 2. Модел на валидация на данните в онлайн базирани въпросници в сферата на туризма

В рамките на описателния анализ следва да се представят структурите по начин, който да обща информация за скалите и признаците, които са измерени в цифрово изражение, като едновременно с това се интегрират и данните от Feedback server-а. По този начин става възможно да се дефинира броят на лицата, които участват в проучването, техният физически пол, броят на нощувките, къде е проведено изследването и т.н.

Логично е в рамките на следващия етап да се дефинира конкретно областта на приложимост на изследването. Това на практика означава да се дефинират нулева или алтернативна статистическа хипотеза, или да се избере подход с параметрична и непараметрична хипотеза в зависимост от целта на изследването. Много често изследователите твърдят, че това е етап, който може и да бъде прескочен и да се стартира от анализа на данни, като удобно пропускат факта, че е възможно известният до момента механизъм на взаимодействие между отделните интересувачи ни критерии да се е променил и вече да не е валиден. Това след себе си води до възникване на Парадокс на Симпсън (Lazarova, A. 2018, pp. 69–75).

Статистически характеристики, които да доказват качество на приближението се осъществяват по данните от извадката да се намерят неизвестни параметри на изследваната случайна величина. Това предполага да се изчислят математическото очакване, дисперсия, мода, медиана, точковите оценки на параметъра, като се третират като реализация на случайна величина по формулата:

$$\Theta^* = \Theta^*(X_1, \dots, X_n),$$

зависеща от определено свойство на величините – Поасоново разпределение, биномно разпределение, хипергеометрично и т.н., след което възниква необходимостта за изчисляване на отместена или неотместена оценка (разликата $E\Theta^* - \theta$ наричаме отместване). Ако $E\Theta^* = \theta + a$, където $a = const$, то оценката $\Theta^* - a$ е неотместена. Ако $E\Theta^* = a\theta$, то оценката a/Θ^* е неотместена оценка. Действително, $E(\Theta^* - a) = E\Theta^* - a = (\theta + a) - a = \theta$. $E\Theta^*$ е средната стойност, около която се колебаят възможните стойности на Θ^* . Затова условието $E\Theta^* = \theta$ означава, че θ^* оценява неизвестната стойност на θ без систематическа грешка. Казваме, че оценката θ^* е състоятелна, ако за произволно $\varepsilon > 0$, то колкото по-голям е обемът на извадката, толкова разликата $|\theta^* - \theta|$ е по-малка (Vasileva, N, 2007, pp. 24–28).

Стабилността на модела и възможностите за „механистичната“ му интерпретация следва да започнат да се изследват, след като се уверим, че моделът има състоятелна и неизместена оценка, като за целта се прилага методът на моментите, чрез който следва да се уверим, че:

1. От извадката е възможно да се изчислят необходимите извадъчни моменти, които оценяват съответните моменти на разпределението на величината X .
2. Търсените оценки на параметрите $\theta_1, \dots, \theta_k$ се определят след като се реши системата уравнения, която вярно и точно отразява зависимостта между тези параметри и моментите на разпределението (Vasileva, 2007, pp. 24–28).

Приложимите статистически методи за валидация на данни в онлайн базиран въпросник в сферата на туризма са разделени на три основни групи в зависимост от избрания изследователски подход и вида на хипотезите, които се използват:

– Методът на интервалите позволява чрез декартово дефиниране на критерии, чрез което да се изведат произведенията на главните компоненти. Той е лесен и бърз за употреба, като не предполага корелация между променливите величини (Donchev, 1978, p. 71).

– Интерполацията и апроксимацията позволяват конструиране на нови числови стойности в област от множество на изолирани точки от известни числови стойности. На практика става дума за приближаването на сложна функция до проста функция и провеждането на регресивен анализ (Spiridonov, 1973, p. 49).

Екстраполацията е единственият възможен метод при прогнозиране стойностите на времева функция за бъдещи моменти от време, поради невъзможността да се намерят стойности, обхващащи търсената. В статистиката екстраполация наричаме оценка на стойността на един признак за период, намиращ се вън от целия времеви отрязък T , за който имаме данни. Основният инструмент при екстраполацията е аналитичният модел на тренда, построен чрез емпирични данни за времеви отрязък T .

ИЗВОДИ

Основният недостатък при избора на приложими статистически методи за валидация на данни в онлайн базиран въпросник в сферата на туризма е желанието за употребата на едни и същи методи по отношение на различни по качество и обем данни от различни източници с цел да създадем представа

за сектора или определена туристическа дестинация. Това обикновено води до подмяна на потребителските мнения, изместени статистически оценки и не на последно място до възникване на Парадокс на Ръсел, който възпрепятства правдивото разбиране на информацията. При избягването на този подход и прилагането на модела в съзвучие с нормативната уредба всеки хотел може да изгради профил на потребителите си и надеждно да измери тяхната удовлетвореност.

ЗАКЛЮЧЕНИЕ

Заключението, което следва да се направи, е в рамките на две структурно различни направления – от една страна, е макроикономическото разбиране, че туризмът като сектор е сезонен и като такъв осигурява приходи на държавния бюджет и само техният размер е важен. От гледна точка на бизнеса, обаче е важна удовлетвореността на потребителите и дали този потребител ще се върне отново и ще сподели своите преживявания с други потребители. Тази реклама от уста на уста или от селфи и позитивни клиентски споделяния в социалните мрежи е практически безплатна за бизнеса, но носи много по-висока придадена стойност за него.

REFERENCES

1. **Ministerstvo na turizma na Republika na Balgaria, 2018.** 'Winter Season Analysis 2016/2017.', viewed 18 July 2018 <http://www.tourism.government.bg/sites/tourism.government.bg/files/analiz_na_zimep_sezon_2016_2017_3.05.17.pdf>
2. **The European Commission 2011.** 'Implementing Regulation (EU) No 692/2011 of the European Parliament and of the Council concerning European statistics on tourism, as regards the structure of the quality reports and the transmission of the data.' in *Official Journal of the European Union*, L 276/13, 21 October 2011, viewed 18 July 2018 <https://eur-lex.europa.eu/eli/reg_impl/2011/1051/oj?locale=en>
3. **The European Commission 2011.** 'Regulation (eu) no 692/2011 of the european parliament and of the council, 6 july 2011', in *Official Journal of the European Union*, L 192/17 22 July 2011 viewed 18 July 2018 <<http://data.europa.eu/eli/reg/2011/692/oj?locale=bg>>
4. **Ministerstvo na turizma na Republika Balgaria, 2013.** 'Tourism Law' prom. SG. No. 30 okt. 26 March 2013 am. SG. No. 96, viewed 18 July 2018 <http://www.tourism.government.bg/sites/tourism.government.bg/files/documents/2018-06/zakon_z_a_turizma.pdf>
5. **Ministerstvoto na transporta, informatsionnite tehnologii i saobshteniyata na Balgaria 2000.** 'Access to Public Information Act', prom. SG. no. 55 by Jul Jul. 200, Am. SG no. 85 viewed 18 July 2018 <https://www.mtitc.government.bg/upload/docs/2016-07/ZAKON_z_a_dostyp_do_obsestvena_informaciq.pdf>
6. **Ministerstvo na zemedeliето i hranite i gorite na Balgaria 1999.** 'Statistics Act', prom., SG no. 57 of 25 June 1999, in force as of 1 January 2014, amended, 19 January 2018 viewed 18 July 2018 <https://www.mzh.government.bg/odzsmlolian/Libraries/Çâëîë/Z-n_Statistikata.sflb.ashx>
7. **Ministerstvo na zemedeliето i hranite na Balgaria, 2010.** 'Law on Prevention and Conflict of Interest', SG No. 97 from year 2010, effective from December 10, 2010, viewed 18 July 2018 g. <<https://www.mi.government.bg/bg/library/zakon-za-predotvratyavane-i-ustanovyavane-na-konflikt-na-interesi-zpuki-447-c25-m258-2.html>>
8. **The European Commission, 1995.** 'COUNCIL DIRECTIVE 95/57/EC of 23 November 1995 on the collection of statistical information in the field of tourism, L 291, 6 December 1995', in *Official Journal of the European Communities*, viewed 18 July 2018 <<https://eur-lex.europa.eu/eli/dir/1995/57/oj?locale=en>>
9. **The European Commission, 2008.** 'Directive 2008/122/EC of the European Parliament and of the Council of 14 January 2009 on the protection of consumers in respect of certain aspects of timeshare, long-term holiday product, resale and exchange contracts' in *Official Journal of the European Communities*, L 33/10, 3 February 2009, viewed 18 July 2018 <<https://eur-lex.europa.eu/legal-content/BG/TXT/?uri=celex%3A32008L0122>>
10. **The European Commission, 2011.** 'Regulation (EU) No 692/2011 of the European Parliament and of the Council of 6 July 2011 concerning European statistics on tourism and repealing Council Directive 95/57/EC Text with EEA relevance' in *Official Journal of the European Union*, L 192/17, 22 July 2011, viewed 18 July 2018 <<https://eur-lex.europa.eu/eli/reg/2011/692/oj>>
11. **The European Commission 2010.** Directive 2011/83/eu of the European Parliament and of the Council of 25 october 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council' in *Official Journal of the European Communities*, L 304/64, 22 November 2010 viewed 18 July 2018 <<http://data.europa.eu/eli/dir/2011/83/oj?locale=en>>

12. **The European Commission, 1999.** ‘Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts’ in *Official Journal of the European Communities*, L 95/29, 21 April 1999 viewed 18 July 2018 <<http://data.europa.eu/eli/dir/1993/13/oj?locale=en>>
13. **The European Commission, 1984.** ‘Council Directive 84/450/EEC of 10 September 1984 relating to the approximation of the laws, regulations and administrative provisions of the Member States concerning misleading advertising’ in *Official Journal of the European Communities*, L 250/17, 19 September 1984 viewed 18 July 2018 <<https://data.europa.eu/eli/dir/1984/450/oj?locale=en>>
14. **Ministerstvo na iekonomikata, 2005.** ‘Consumer Protection Act’, *prom. SG no. from December 2005, am. SG no. 20 of March 2018* viewed 18 July 2018 <<https://www.mi.government.bg/bg/library/zakon-za-zashtita-na-potrebitelite-yanuari-2017-g-631-c25-m258-3.html>>
15. **Lazarova, A., 2018.** ‘Simpsons paradox in gender social role type of researches’, *NVU’s Scientific conference – 14 – 15.06.2018* NVU ‘Vasil Levski’, ISSN 1314-1937 (in Bulgarian)
16. **Vasileva, N. 2007.** ‘Lectures on Probability Theory and Mathematical Statistics University of Mining and Geology “St. Ivan Rilski”’ viewed 18 July 2018 <[http://www.mgu.bg/drugi/ebooks/nikolina/lessons/par24-28\(estimations\).pdf](http://www.mgu.bg/drugi/ebooks/nikolina/lessons/par24-28(estimations).pdf)> (in Bulgarian)
17. **Donchev, A. 1978.** ‘Lectures on Mathematical Optimization’, Sofia (in Bulgarian)
18. **Spiridonov, V. 1973.** ‘Investigation of Operations’, in *Nauka i izkuvstvo Sofia* (in Bulgarian)

За контакти:

Антония Лазарова, главен асистент, доктор
Служебен адрес: Пловдив, 4000, ул. Задруга 18
Ел. поща: lazarova.a@abv.bg
