Студентската оценка на преподаването – проблемът за точността на изводите

За целите на атестацията на преподавателите и по-общо за целите на управлението на качеството на учебния процес е нужна информация за това как студентите оценяват учебните дисциплини и преподавателите си. Изследванията на студентските мнения обаче се сблъскват с един постоянстващ проблем – отговарят не всички студенти, а само тази част от тях, които са открити по време на провеждане на изследването и са пожелали да попълнят анкетната карта (Илиева 2002: 145-146, Маринова 2006: 9 и 14, Хараламписв, Караджов 2003: 138). Това прави изследването непредставително и то по метода на отзовалите се. Въпреки това, има три съществени разлики от класическия метод на отзовалите се, които значително облекчават работата на изследователя:

Първо, при класическия метод на отзовалите се най-същественият проблем е, че всъщност не е ясно коя е изучаваната генерална съвкупност и следователно не е ясно как трябва да се генерализират изводите от извадката. При изследване на студентските мнения генералната съвкупност е ясно дефинирана, въпреки че са възможни вариации, описани по-долу в текста.

Второ, класическият метод на отзовалите се предполага, че едно лице би могло да отговори няколко пъти на анкетните въпроси и в този смисъл подборът е възвратен. При изследване на студентските мнения един студент оценява само един път конкретна дисциплина и конкретен преподавател и в този смисъл подборът е безвъзвратен.

Трето, от първите две различия следва третото. При класическия метод на отзовалите се, тъй като не е ясна изследваната съвкупност и е възможно едно лице да отговаря многократно, обикновено изводите се генерализират за безкрайно голяма генерална съвкупност[1]. При изследване на студентските мнения е обратното – генералната съвкупност е крайна и сравнително малка[2].

Тези три различия позволяват при изследване на студентските мнения чрез извадки, за които не може да се гарантира представителност, да се правят по-точни изводи в сравнение с изводите при класическия метод на отзовалите се[3].

---

[1] Обикновено това са „читателиТЕ”, „зрителиТЕ”, „потребителиТЕ на Интернет”, „българиТЕ”, „хораТА” и т.н.

[2] Обикновено при такива изследвания наблюдаваната съвкупност са студентите от една и съща специалност и един и същи курс.

[3] Правенето на статистически изводи при непредставителни извадки е описано в Хараламписв 2004а.

Основната цел на настоящата статия е да покаже от какво зависи точността на изводите при изследване на студентските мнения. Успоредно с това ще бъде показано как се правят самите изводи, валидни за цялата генерална съвкупност.

И така, изследването започва с дефиниране на изследваната съвкупност. Още тук възниква един етичен проблем – нужно ли е да правим всичко възможно, за да обхванем всички студенти, след като част от тях така или иначе не посещават лекциите и упражненията? Иначе казано – имат ли право студенти, които не са посещавали учебните занимания, да дават оценка на дисциплината и на преподавателя?

Различните възможности за отговор и съответно за формиране на генералната съвкупност и извадката са представени в таблица 1:

Таблица 1

Разпределение на всички студенти по посещаемост на учебните занимания

В момента	По принцип		Общо
	Посещават	Не посещават	
Присъстват	f_{11}	f_{12}	$\sum_j f_{1j}$
Отсъстват	f_{21}	f_{22}	$\sum_j f_{2j}$
Общо	$\sum_i f_{i1}$	$\sum_i f_{i2}$	$\sum_i \sum_j f_{ij}$

Първа възможност – генералната съвкупност са студентите, които „по принцип” посещават учебните занимания, а извадката са тези от тях, които в момента на изследването присъстват и желаят да попълнят анкетната карта. В този случай обемът на извадката е $f_{11} = n$, а обемът на генералната съвкупност е $\sum_i f_{i1} = N$.

Втора възможност – генералната съвкупност са всички студенти, независимо дали посещават учебните занимания или не, а извадката са тези от тях, които в момента на изследването присъстват и желаят да попълнят анкетната карта. В този случай обемът на извадката е $\sum_j f_{1j} = n$, а обемът на генералната съвкупност е $\sum_i \sum_j f_{ij} = N$.

Ясно е, че има два начина за дефиниране на генералната съвкупност. За всеки от тях извадката се дефинира еднозначно. Това означава, че може да се работи с универсалните означения за обемите на извадката и на генералната съвкупност (съответно *n* и *N*), като в тях се влага различно съдържание в зависимост от конкретния начин на дефиниране на генералната съвкупност.

Най напред ще разгледам анкетните въпроси, чиито отговори са разположени на бална или интервална скала. Тези скали имат следната особеност – значенията на признаците са **числа**. Това означава, че може да се приложи методът за правене на изводи относно количествени признаци при непредставителни извадки, предложен от Харалампиев (Харалампиев 2004б: 8-20) и по-конкретно методът за оценяване на средната аритметична (Харалампиев 2004б: 31-34).

Най-общо при непредставителни извадки неизвестната средна аритметична в генерална съвкупност има симетрично вероятностно разпределение, което е приблизително нормално, с център

(1) $$\bar{\mu} = \frac{\dfrac{(N-n)(x_1+x_m)}{2} + \sum_{i=1}^{m} x_i f_i}{N}$$

и разсейване

(2) $$\sigma_\mu = \frac{x_m - x_1}{2N}\sqrt{\frac{(N-n)(N-n+m)}{3(m-1)}},$$

където x_i са значенията на признака, f_i са честотите, x_1 е най-малкото значение на признака, x_m е най-голямото значение на признака, а m е броят на всички значения.

Освен това, възможните стойности на средната са ограничени в интервала между

(3) $$\mu_{\min} = \frac{x_1(N-n) + \sum_{i=1}^{m} x_i f_i}{N}$$

и

(4) $$\mu_{\max} = \frac{x_m(N-n) + \sum_{i=1}^{m} x_i f_i}{N}$$

От симетричността на разпределението следва, че $\bar{\mu}$ е едновременно и най-вероятна стойност, и следователно σ_μ, което показва средния размер на отклоненията около центъра (т.е. в случая около най-вероятната стойност), е измерител на точността. Затова нека видим от какво зависят тези две стойности.

Първо, формула (1) може да се запише във вида:

(5) $$\overline{\mu} = \left(1 - \frac{n}{N}\right) . \frac{x_1 + x_m}{2} + \frac{n}{N} . \overline{x} \ ,$$

където $\overline{x}$ е извадковата средна.

От формула (5) се вижда, че най-вероятната стойност на средната аритметична в генералната съвкупност е претеглена средна от извадковата средна и геометричния център на разпределението $\left(\frac{x_1 + x_m}{2}\right)$, с тегла съответно делът на извадката в генералната съвкупност и делът на останалата част от генералната съвкупност. Следователно, колкото делът на извадката е по-голям, толкова най-вероятната стойност на средната аритметична в генералната съвкупност ще се доближава до извадковата средна, и обратно, колкото делът на извадката е по-малък, толкова най-вероятната стойност на средната аритметична в генералната съвкупност ще се доближава до геометричния център на разпределението.

Този резултат е различен от резултата, който бихме получили, ако извадката би била представителна. При представителни извадки най-вероятната стойност на средната аритметична в генералната съвкупност съвпада с извадковата средна, докато тук, тъй като извадката е непредставителна, извадковата средна вече е изместена оценка на средната аритметична в генералната съвкупност, като колкото делът на извадката е по-малък, толкова извадковата средна е по-изместена.

Второ, формула (2) може да се запише във вида:

(6) $$\sigma_\mu = \frac{x_m - x_1}{2} \sqrt{\frac{\left(1 - \frac{n}{N}\right)\left(1 - \frac{n}{N} + \frac{m}{N}\right)}{3(m-1)}}$$

От формула (6) се вижда, че колкото делът на извадката е по-голям, толкова разсейването е по-малко, т.е. точността е по-голяма.

Този резултат също е различен от резултата, който бихме получили, ако извадката би била представителна. При представителни извадки точността на оценката зависи в много по-голяма степен от абсолютния обем на извадката и в много по-малка степен от дела ѝ в генералната съвкупност, докато тук се вижда, че **при непредставителни извадки абсолютният обем на извадката не оказва НИКАКВО влияние върху точността,** а единствено делът ѝ има значение.

Трето, горният извод може да се допълни, ако от формули (3) и (4) се изчисли размахът на разпределението:

(7) $$\mu_{\max} - \mu_{\min} = (x_m - x_1)\left(1 - \frac{n}{N}\right)$$

От формула (7) отново се вижда, че колкото делът на извадката е по-голям, толкова разсейването е по-малко, т.е. точността е по-голяма.

Тези изводи могат да бъдат илюстрирани с пример: 22 студенти са оценили преподавателя си по петобална скала със значения от 1 до 5. Извадковата средна е 3,95. Генералната съвкупност обхваща 66 студента. Делът на извадката е 33%[4]. Следователно най-вероятната средна оценка е 3,32 с разсейване ±0,41 и размах 2,67.

Нека запазим абсолютния обем на извадката, а да променим обема на генералната съвкупност, като по този начин променим дела на извадката. Нека сега генералната съвкупност да обхваща 40 студента[5]. Тогава делът на извадката ще бъде 55% и най-вероятната средна стойност ще бъде 3,52 с разсейване ±0,29 и размах 1,80. Видно е, че в този случай оценката е по-точна и е по-близо до извадковата средна. Сравнението между двата случая е представено графично на фигура 1.

С това целите на статията са изпълнени. Нека обобщим:

Първо, най-вероятната стойност на средната аритметична в генералната съвкупност се намира между геометричния център на разпределението и извадковата средна, като близостта до едната или до другата стойност се определя от дела на извадката в генералната съвкупност.

Второ, точността зависи от диапазона между най-голямото и най-малкото значение на признака и от дела на извадката в генералната съвкупност.

Трето, формули (1) и (2) са достатъчни за правенето на статистически изводи относно средната аритметична в генералната съвкупност при непредставителни извадки и безвъзвратен подбор.

---

[4] Тези данни са от реално проведено изследване, макар че името на преподавателя и специалността са премълчани. Нарочно е подбран този преподавател, защото при него делът на извадката е близък до посочения от Маринова дял на извадката при анкетата по пощата – 31,1% (Маринова 2006: 9), до посочения от Илиева като долна граница в изследванията, проведени в ХТМУ-София дял 30% (Илиева 2002: 145) и до посочените дялове от Караджов и Харалампиев в изследвания, проведени в Богословски факултет на Софийски университет (Харалампиев, Караджов 2003: 138).

[5] Изборът на обема на генералната съвкупност е фиктивен, само за целите на илюстрацията. Новият дял на извадката е близък до посочения от Маринова дял на извадката при оценяване на упражненията по Интернет – 50% (Маринова 2006: 14). Не е правена илюстрация на ситуациите с дял на извадката близък до посочения от Маринова дял при оценяване на лекциите по Интернет – 80% (Маринова 2006: 14) и с посочения от Илиева като горна граница в изследванията, проведени в ХТМУ-София дял 95% (Илиева 2002: 145).

Въпреки че първоначално дефинираните цели са изпълнени, ще продължа с осветляването на някои важни особености на метода, които сега са скрити. Тяхното игнориране е потенциално опасно, защото директното приложение на формули (1) и (2) (без отчитане на тези особености) може да доведе до грешки и неточности.

Първо, ключово важно изискване е числовите стойности, с които се описват значенията на признака да са **последователни** числа. На практика при балните скали това е изпълнено, защото там значенията на признака са предварително дефинирани и анкетираното лице трябва да избере измежду вече зададените стойности. При интервалните скали проблемът е по-сериозен, защото се допуска, че може да няма предварително зададени стойности, а анкетираното лице само да вписва числата, които са отговор на анкетните въпроси. В този случай има различни варианти за справяне с проблема. Ако признакът е прекъснат (дискретен), за негови значения могат да се вземат всички цели числа в диапазона между най-малкото и най-голямото посочено число, макар че някои от тях може да не са избрани от нито едно анкетирано лице. Ако признакът е непрекъснат, трябва да се направи интервална групировка и за значения на признака да се вземат средите на интервалите.

Второ, макар че при балните скали числовите стойности са последователни и са предварително дефинирани, може да възникне следната ситуация: най-малката и/или най-голямата стойност не са посочени от нито едно анкетирано лице. Тогава възниква друг етичен проблем – можем ли да интерпретираме тази числова стойност като значение на признака? Иначе казано – може преподавателят да е толкова добър (лош), че нито един студент в генералната съвкупност да не му постави най-ниската (най-високата) оценка, а може в генералната съвкупност да има студенти, които биха поставили най-ниска (най-висока) оценка, но нито един от тях не е попаднал в извадката. За да се провери дали това наистина е проблем, ще бъде разгледана ситуацията, при която най-ниската оценка не е посочена от нито един студент, като ситуацията, при която не е посочена най-високата оценка, е огледален образ.

Ако приемем, че в генералната съвкупност има студенти, които биха поставили най-ниската оценка, то най-вероятната стойност на средната аритметична в генералната съвкупност, разсейването, най-малката и най-голямата възможна стойност на средната се получават по описания по-горе начин.

Ако приемем, че в генералната съвкупност няма студенти, които биха поставили най-ниската оценка, то най-вероятната стойност на средната аритметична в генералната съвкупност, разсейването, най-малката и най-голямата възможна стойност на средната се получават по вариации на горните формули:

(8) $$\bar{\mu}' = \left(1 - \frac{n}{N}\right) . \frac{x_2 + x_m}{2} + \frac{n}{N} . \bar{x}$$

(9) $$\sigma'_\mu = \frac{x_m - x_2}{2} \sqrt{\frac{\left(1 - \frac{n}{N}\right)\left(1 - \frac{n}{N} + \frac{m-1}{N}\right)}{3(m-2)}}$$

(10) $$\mu'_{\min} = \frac{x_2 (N - n) + \sum_{i=1}^{m} x_i f_i}{N}$$

(11) $$\mu'_{\max} = \frac{x_m (N - n) + \sum_{i=1}^{m} x_i f_i}{N}$$

Тогава разликите между стойностите, получени при двете различни допускания са:

(12) $$\bar{\mu}' - \bar{\mu} = \left(1 - \frac{n}{N}\right) . \frac{x_2 - x_1}{2}$$

(13) $$(\sigma'_\mu)^2 - \sigma_\mu^2 = -\frac{1}{3}\left(\frac{x_2 - x_1}{2}\right)^2 \left(1 - \frac{n}{N}\right)\left[1 - \frac{n}{N} + \frac{2(m-1)}{N}\right]$$

(14) $$\mu'_{\min} - \mu_{\min} = \left(1 - \frac{n}{N}\right) . (x_2 - x_1)$$

(15) $$\mu'_{\max} - \mu_{\max} = 0$$

Тези резултати дават възможност да се направят няколко извода:

Първо, ако допуснем, че в генералната съвкупност няма студенти, които биха поставили най-ниската оценка, новото разпределение на средната аритметична в генералната съвкупност ще съвпада в горния си край със старото (формула (15)) обаче в долния си край двете разпределения ще се различават с не повече от една единица (формула (14)), а разликата между най-вероятните стойности ще бъде най-много половин единица (формула (12)).

Второ, най-малката възможна стойност и най-вероятната стойност на новото разпределение са по-близо до извадковата средна, отколкото съответните им стойности в старото разпределение.

Трето, ако допуснем, че в генералната съвкупност няма студенти, които биха поставили най-ниската оценка, новото разпределение ще бъде с по-малко разсейване от старото (формула (13)), следователно изводът ще бъде по-точен.

Четвърто, с увеличаване на дела на извадката в генералната съвкупност, разликите между стойностите, получени при двете различни допускания, стават все по-малки.

Нека отново илюстрираме получените изводи с пример. За целта да модифицираме предходния пример: 22 студенти са оценили преподавателя си по петобална скала със значения от 1 до 5, като нито един студент не е посочил оценка 1. Отново извадковата средна е 3,95, генералната съвкупност обхваща 40 студента и делът на извадката е 55%. Най-вероятната стойност на средната аритметична в генералната съвкупност и разсейването обаче вече са други – съответно 3,75 и ±0,25. Сравнението с предходния пример показва, че новата най-вероятна стойност е още по-близо до извадковата средна и новото разсейване е още по-малко. Сравнението е представено графично на фигура 2.

Получените резултати показват, че това, че най-малката и/или най-голямата стойност не са избрани от нито едно анкетирано лице наистина е проблем. Допускането, че в генералната съвкупност има студенти, които биха поставили най-ниската (най-високата) оценка води до по-консервативни изводи, а допускането, че в генералната съвкупност няма студенти, които биха поставили най-ниската (най-високата) оценка води до по-точни изводи, които обаче биха могли да се окажат прекалено оптимистични. Статистически коректният подход в такива ситуации е да се дадат и двете възможни решения, макар че това не решава проблема[6], а само прехвърля отговорността за решаването му на следващия субект в управленската верига.

До тук разгледах признаците, които позволяват да се изчисляват средни стойности. Това са признаците, разположени на бални и интервални скали. Но освен тях съществуват и признаци, чиито значения са разположени на номинална скална. Тези признаци не позволяват да се изчисляват средни стойности. Единствената тяхна

[6] Спорен въпрос е дали този проблем изобщо е решим. От статистическа гледна точка той е нерешим (макар че любовта на статистиците към по-консервативните решения е широко известна), така че за решаването му трябва да се търсят други критерии, извън статистическата теория. Не случайно в началото дефинирах този проблем като етичен, а не като статистически.

числова характеристика е относителният дял. Обаче за разлика от средната аритметична в генералната съвкупност, която има симетрично вероятностно разпределение, вероятностното разпределение на относителния дял в генералната съвкупност е крайно асиметрично L-разпределение (Харалампиев 2004б: 55). Това означава, че неговата средна и стандартно отклонение не са адекватни измерители съответно на най-вероятната стойност и на разсейването около нея. Все пак и при тези признаци могат да се правят изводи относно най-вероятната стойност на относителния дял в генералната съвкупност и относно точността.

Първо, най-вероятната стойност на относителния дял на *i*–тото значение на признака в генералната съвкупност е най-малката възможна стойност:

(16) $\pi_{i,\min} = \frac{f_i}{N}$ (Харалампиев 2004б: 54)

Тази формула може да се запише във вида:

(17) $\pi_{i,\min} = \left(1 - \frac{n}{N}\right).0 + \frac{n}{N}.p_i$,

където p_i е извадковият относителен дял на *i*–тото значение на признака.

От формула (17) се вижда, че най-вероятната стойност на относителния дял в генералната съвкупност е претеглена средна от извадковия относителен дял и нулата, с тегла съответно делът на извадката в генералната съвкупност и делът на останалата част от генералната съвкупност. Следователно, колкото делът на извадката е по-голям, толкова най-вероятната стойност на относителния дял в генералната съвкупност ще се доближава до извадковия относителен дял, и обратно, колкото делът на извадката е по-малък, толкова най-вероятната стойност на относителния дял в генералната съвкупност ще се доближава до нулата.

Второ, след като стандартното отклонение не е адекватен измерител на разсейването, вместо него може да се използва средното отклонение около най-вероятната стойност:

(18) $\delta = \left(1 - \frac{n}{N}\right).\frac{1}{m}$

От формула (18) се вижда, че колкото делът на извадката е по-голям, толкова средното отклонение около най-вероятната стойност е по-малко, т.е. точността е по-голяма.

Трето, формули (16) и (18) са достатъчни за правенето на статистически изводи относно относителния дял в генералната съвкупност при непредставителни извадки и безвъзвратен подбор.

Четвърто, въпросът „Какво става, ако най-малкото и/или най-голямото значение на признака не са посочени от нито едно анкетирано лице?” вече се формулира като „Какво става, ако някое от значенията на признака не е посочено от нито едно анкетирано лице?”. От формули (16) и (18) се вижда, че ако допуснем, че в генералната съвкупност няма студенти, които биха посочили непосоченото значение на признака, това не оказва влияние върху най-вероятната стойност на другите относителни дялове в генералната съвкупност, а само влошава точността.

Накрая, още веднъж да обобщим получените резултати:

При изследване на студентските мнения чрез извадки, за които не може да се гарантира представителност, най-вероятната стойност на изследвания параметър в генералната съвкупност (средна аритметична или относителен дял) е претеглена средна от две осреднявани величини, едната от които е оценката от извадката. Оценката от извадката участва в осредняването с тегло равно на дела на извадката в генералната съвкупност. Следователно, колкото делът на извадката в генералната съвкупност е по-голям, толкова най-вероятната стойност на параметъра е по-близо до оценката от извадката. Точността на изводите също зависи от дела на извадката в генералната съвкупност – по-голям дял на извадката води до по-голяма точност.

ЛИТЕРАТУРА

Илиева, М. 2002. Студентската оценка за преподаването. ХТМУ, С.

Маринова, С. 2006. Доклад относно някои елементи на реформата на висшето образование в Германия: системи за оценка на качеството на преподаване; кредитна система; форми на тюторство. Непубликуван ръкопис.

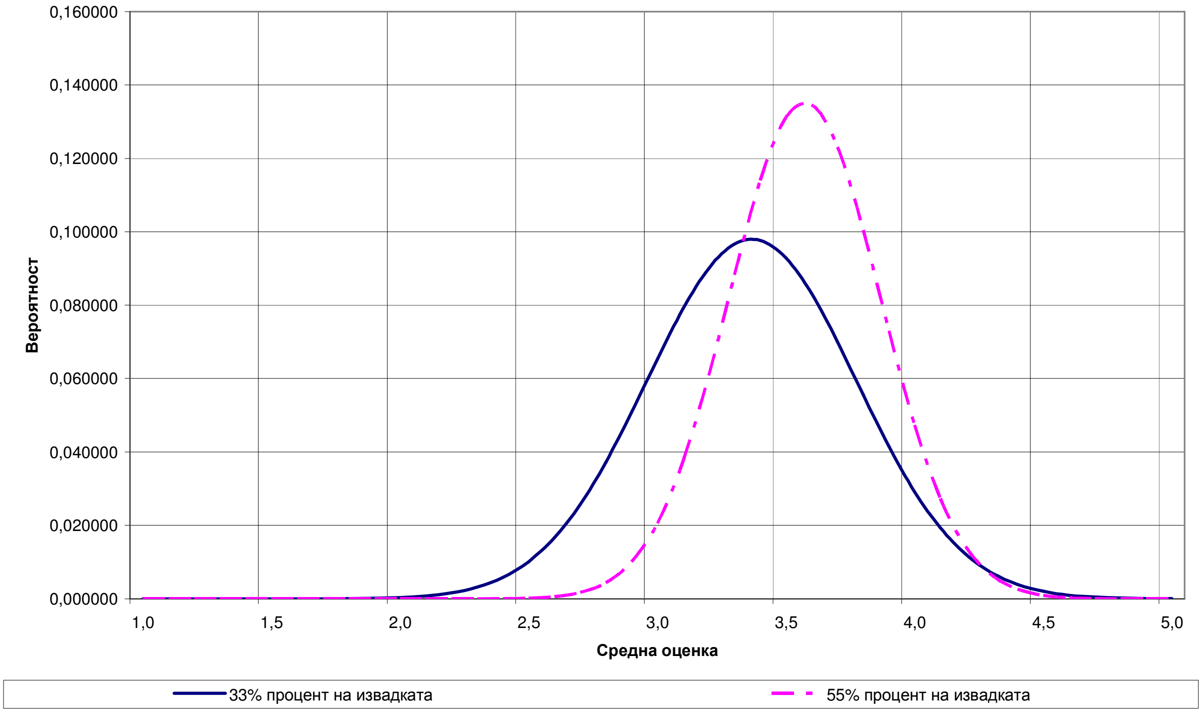
Харалампиев, К., С. Караджов. 2003. Доклад със сравнителен анализ на данни, получени от емпирични социологически изследвания, проведени със студенти от Богословски факултет на СУ „Св. Климент Охридски” през декември 2002, май 2003 и декември 2003 година. „Богословска мисъл”, 3-4

Хараламиев, К. 2004а. Анкетите в Интернет – възможност за статистически изводи и интерпретиране на резултатите. „Социологически проблеми”, 3-4

Хараламиев, К. 2004б. Нетрадиционен поглед върху традиционни статистически проблеми. Балкани, С.

Фигура 1

Вероятностно разпределение на средната аритметична в генералната съвкупност при два различни дяла на извадката



Фигура 2

Вероятностно разпределение на средната аритметична в генералната съвкупност при включване и изключване на най-малкото значение на признака