



Утвърдил: .....

Декан

Дата .....

## СОФИЙСКИ УНИВЕРСИТЕТ "СВ. КЛИМЕНТ ОХРИДСКИ"

Факултет: **Философски**

Специалност: *(код и наименование)*

Ф	Ф	С	0	4	0	1	1	3
---	---	---	---	---	---	---	---	---

Социология

### УЧЕБНА ПРОГРАМА

Дисциплина:

И	3	0	8
---	---	---	---

### ПРИЛОЖЕНИЕ НА БЕЙСОВСКАТА СТАТИСТИКА В СОЦИАЛНИТЕ НАУКИ

Преподавател: Доц. К. Харалампиев

Учебна заетост	Форма	Хорариум
Аудиторна заетост	Лекции	0
	Семинарни упражнения	60
<b>Обща аудиторна заетост</b>		<b>60</b>
Извънаудиторна заетост	Самостоятелна работа в библиотека или с ресурси	90
<b>Обща извънаудиторна заетост</b>		<b>90</b>
<b>ОБЩА ЗАЕТОСТ</b>		<b>150</b>
<b>Кредити аудиторна заетост</b>		<b>2</b>
<b>Кредити извънаудиторна заетост</b>		<b>3</b>
<b>ОБЩО ЕКСТ</b>		<b>5</b>

№	Формиране на оценката по дисциплината	% от оценката
1.	Изпит	100%

### Критерии на оценяване

На изпита всеки студент изтегля един казус от следните възможности:

1. Оценяване на относителен дял по данни от непредставителна извадка, когато дялът на извадката е крайно число (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е известен и може да се изчисли каква част от посетилите сайта са попълнили анкетата).
2. Оценяване на относителен дял по данни от непредставителна извадка, когато дялът на извадката е пренебрежимо малък (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е неизвестен и не може да се изчисли каква част от посетилите сайта са попълнили анкетата).
3. Оценяване на средна аритметична по данни от непредставителна извадка, когато дялът на извадката е крайно число (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е известен и може да се изчисли каква част от посетилите сайта са попълнили анкетата).
4. Оценяване на средна аритметична по данни от непредставителна извадка, когато дялът на извадката е пренебрежимо малък (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е неизвестен и не може да се изчисли каква част от посетилите сайта са попълнили анкетата).
5. Определяне на победител в класация по данни от непредставителна извадка, когато дялът на извадката е крайно число (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е известен и може да се изчисли каква част от посетилите сайта са попълнили анкетата).
6. Определяне на победител в класация по данни от непредставителна извадка, когато дялът на извадката е пренебрежимо малък (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е неизвестен и не може да се изчисли каква част от посетилите сайта са попълнили анкетата).

7. Оценяване на относителен дял по данни от непредставителна извадка, когато генералната съвкупност е малка (студентска оценка на качеството на преподаване).

8. Оценяване на средна аритметична по данни от непредставителна извадка, когато генералната съвкупност е малка (студентска оценка на качеството на преподаване).

9. Оценяване на относителен дял по данни от представителна извадка.

10. Оценяване на относителен дял на избраните конкретна партия (или политик) само сред действителните гласоподаватели по данни от електорално изследване.

11. Вземане на решения в условия на риск.

Забележки:

1. Оценяване (на относителен дял или средна аритметична) означава да се изчисли функцията на разпределение и въз основа на нея да се проверят статистически хипотези и да се построят доверителни интервали.

2. Определяне на победител в класация означава да се изчисли вероятността конкретен участник да е победител.

3. Казус № 11 ще бъде придружен от целия масив с данни.

4. Тези единадесет казуса покриват целия изучаван материал. Възможно е обаче за някой от тях да липсват данни и този казус да не се падне на изпита.

Базовата оценка е Отличен (6). Всяка грешка или пропуск води до намаляването на оценката с половин единица (0,5). Крайната оценка се получава по следния начин:

За оценка „ <b>Отличен</b> “ (6)	<p>0 грешки и 0 пропуска или 1 грешка и 0 пропуска или 0 грешки и 1 пропуск</p>
За оценка „ <b>Много добър</b> “	<p>2 грешки и 0 пропуска или 1 грешка и 1 пропуск или 0 грешки и 2 пропуска или 3 грешки и 0 пропуска или 2 грешки и 1 пропуск или 1 грешка и 2 пропуска или 0 грешки и 3 пропуска</p>
За оценка „ <b>Добър</b> “	<p>4 грешки и 0 пропуска или 3 грешки и 1 пропуск или 2 грешки и 2 пропуска или 1 грешка и 3 пропуска или 0 грешки и 4 пропуска или 5 грешки и 0 пропуска или 4 грешки и 1 пропуск или 3 грешки и 2 пропуска или 2 грешки и 3 пропуска или</p>

	1 грешка и 4 пропуска или 0 грешки и 5 пропуска
За оценка „Среден“	6 грешки и 0 пропуска или 5 грешки и 1 пропуск или 4 грешки и 2 пропуска или 3 грешки и 3 пропуска или 2 грешки и 4 пропуска или 1 грешка и 5 пропуска или 0 грешки и 6 пропуска

#### **Анотация на учебната дисциплина:**

Изследователските проблеми, изискващи приложението на статистически методи, могат да се обособят в три големи групи:

В първата група попадат тези проблеми, които имат теоретично решение. Наистина, като правило, теоретичните решения на един и същ проблем са повече от едно и всяко от тях се основава на различна логика и на различни предположения. Това от своя страна поставя изискването да се познават всички решения и във всяка конкретна ситуация да се прави избор.

Във втората група попадат тези проблеми, които нямат теоретично решение, но в своята практическа дейност изследователите ги свеждат, не съвсем коректно, към аналогични проблеми, които имат теоретично решение.

В третата група попадат тези проблеми, които нито имат теоретично решение, нито могат да се сведат до вече решени проблеми.

Целта на настоящият курс е да покаже как от позициите на една малко позната парадигма в статистиката, наречена бейсовска, може да се даде теоретично решение на всеки изследователски проблем и в този смисъл трите групи проблеми да се сведат към първата. Разбира се, тук по-скоро става въпрос за потенциал, понеже е практически невъзможно да бъдат описани изчерпателно всички проблеми, изискващи приложението на статистически методи. По тази причина, половината от курса е посветена на теоретично въведение, а в другата половина са разгледани избрани изследователски проблеми. В теоретичното въведение са изложени принципите, инструментите и техниките на бейсовската статистика. Това означава, че в тези теми по необходимост има малко повече математика, но тя е необходима основа за разбирането на различните приложения. А в този курс ще се разглеждат приложения, отнасящи се и до трите групи изследователски проблеми:

Към третата група проблеми се отнасят изводите, основани на информация от непредставителни извадки. Примери за непредставителни извадки напоследък има достатъчно много – анкетите в Интернет, гласуванията по телефона или чрез SMS за различни телевизионни предавания (Песен на Евровизия, Тест на нацията, Big Brother, Star Academy, Музикална ку-ку академия, Великолепната шесторка и т.н., и т.н.).

Към втората група проблеми се отнасят изводите, основани на информация от представителни извадки, но отнасящи се за няколко взаимно свързани параметъра на генералната съвкупност. Такъв, например, е проблемът за оценяването на процента на гласувалите за някоя партия (кандидат), но само сред действителните гласоподаватели.

Към първата група проблеми се отнася определянето на вероятността нова единица да попадне в конкретна група, формирана по значението на даден признак. Това успешно може се съчетае с техниката на многомерните групировки и по този начин се получава лесна за употреба алтернатива на методите, носещи общото наименование „вземане на решения в условия на риск“.

#### **Предварителни изисквания:**

Успешно преминати курсове по „Статистически методи в социологията“ – първа и втора част.

**Очаквани резултати:**

Умения на студентите в конкретна практическа ситуация да определят априорни вероятности и извадкови разпределения и на тяхна основа да изчисляват апостериорни вероятности. На базата на апостериорните вероятности да построят доверителни интервали и да правят проверка на хипотези.

**Учебно съдържание**

№	Тема:	Хорариум
1	Малко история. Теория на вероятностите. Честотна статистика. Бейсовска статистика.	4 часа
2	Малко математика. Вероятност. Условна вероятност. Теорема за събиране и умножаване на вероятности. Теорема за пълната вероятност и теорема на Бейс. Априорни и апостериорни вероятности. Метод на максималната ентропия. Разпределение, плътност на разпределение и функция на разпределение.	4 часа
3	Статистически изводи и заключения относно относителни дялове в генералната съвкупност на базата на априорни вероятности Безвъзвратен подбор, възвратен подбор, големи генерални съвкупности.	4 часа
4	Статистически изводи и заключения относно важни параметри на разпределенията (средни аритметични, стандартни отклонения и т.н.) в генералната съвкупност на базата на априорни вероятности Безвъзвратен подбор, възвратен подбор, големи генерални съвкупности.	4 часа
5	Приложения в областта на социалните науки – непредставителни извадки Анкети в Интернет. Телевизионни гласувания – „Песен на Евровизия”, „Вот на доверие”, „Сблъсък”, „Дуел”, „Big Brother”, „Star Academy”, „Музикална ку-ку академия”, „Тест на нацията” и т.н. Проблемът за „победителя”. Студентска оценка за качеството на преподаване.	12 часа
6	Статистически изводи и заключения относно относителни дялове в генералната съвкупност на базата на апостериорни вероятности Безвъзвратен подбор, възвратен подбор, големи генерални съвкупности.	4 часа
7	Статистически изводи и заключения относно важни параметри на разпределенията (средни аритметични, стандартни отклонения и т.н.) в генералната съвкупност на базата на апостериорни вероятности Безвъзвратен подбор, възвратен подбор, големи генерални съвкупности.	4 часа
8	Приложения в областта на социалните науки – представителни извадки Предиизборни проучвания. Проблемът за процента спрямо действителните гласове.	12 часа
9	Определяне на вероятността нова единица да попадне в конкретна група, формирана по значенията на един или няколко признака. Определяне на вероятността нова единица да попадне в конкретна група, формирана по значенията на един или няколко признака при работа с априорни и с апостериорни вероятности. Банкови правила за отпускане на кредити. Правила за определяне на застрахователни премии. Възможности за приложение в областта на маркетинговите и политическите изследвания.	8 часа
10	Приложение в областта на прогнозирането – доверителни интервали на екстраполационни прогнози.	4 часа

**Конспект за изпит**

№	Въпрос
1	Оценяване на относителен дял по данни от непредставителна извадка, когато делът на извадката е крайно число (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е известен и може да се изчисли каква част от посетилите сайта са попълнили анкетата).
2	Оценяване на относителен дял по данни от непредставителна извадка, когато делът на извадката е пренебрежимо малък (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е неизвестен и не може да се изчисли каква част от посетилите сайта са

№	Въпрос
	попълнили анкетата).
3	Оценяване на средна аритметична по данни от непредставителна извадка, когато делът на извадката е крайно число (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е известен и може да се изчисли каква част от посетилите сайта са попълнили анкетата).
4	Оценяване на средна аритметична по данни от непредставителна извадка, когато делът на извадката е пренебрежимо малък (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е неизвестен и не може да се изчисли каква част от посетилите сайта са попълнили анкетата).
5	Определяне на победител в класация по данни от непредставителна извадка, когато делът на извадката е крайно число (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е известен и може да се изчисли каква част от посетилите сайта са попълнили анкетата).
6	Определяне на победител в класация по данни от непредставителна извадка, когато делът на извадката е пренебрежимо малък (анкета в Интернет, когато броят на лицата, посетили съответния сайт, е неизвестен и не може да се изчисли каква част от посетилите сайта са попълнили анкетата).
7	Оценяване на относителен дял по данни от непредставителна извадка, когато генералната съвкупност е малка (студентска оценка на качеството на преподаване).
8	Оценяване на средна аритметична по данни от непредставителна извадка, когато генералната съвкупност е малка (студентска оценка на качеството на преподаване).
9	Оценяване на относителен дял по данни от представителна извадка.
10	Оценяване на относителен дял на избралите конкретна партия (или политик) само сред действителните гласоподаватели по данни от електорално изследване.
11	Вземане на решения в условия на риск.

### Библиография

#### Основна:

- Харалампиев, К. 2012. Още една гледна точка към проблема за отказите при социологически изследвания. Сп. „Социологически проблеми”, брой 1-2. Достъпна в Интернет на адрес: <http://kaloyan-haralampiev.info/wp-content/uploads/2010/03/statiya19.pdf>
- Харалампиев, К. 2009. Студентската оценка на преподаването – проблемът за точността на изводите. В: Социологията пред предизвикателството на различията. Юбилеен сборник, посветен на 30-годишнината на катедра „Социология”. Университетско издателство „Св. Климент Охридски“, София. Достъпна в Интернет на адрес: <http://kaloyan-haralampiev.info/wp-content/uploads/2010/03/statija12.pdf>
- Харалампиев, К. 2008. Лапласово правило за приемственост – интерпретации и приложения. Научна конференция с международно участие „Авангардни научни инструменти в управлението”, Равда. Достъпна в Интернет на адрес: <http://kaloyan-haralampiev.info/wp-content/uploads/2010/03/doklad-6.pdf>
- Харалампиев, К. 2007. За парадигмите в статистиката – бейсовска статистика. Международна научна конференция „Актуални проблеми на статистическата теория и практика”, Равда. Достъпна в Интернет на адрес: <http://kaloyan-haralampiev.info/wp-content/uploads/2010/03/doklad-1.pdf>
- Харалампиев, К. 2005. Телевизионните гласувания по телефона – проблемът за „победителя”. Сп. „Социологически проблеми”, брой 3-4. Достъпна в Интернет на адрес: <http://kaloyan-haralampiev.info/wp-content/uploads/2010/03/statija13.pdf>
- Харалампиев, К. 2004б. Нетрадиционен поглед върху традиционни статистически проблеми. Балкани, София
- Харалампиев, К. 2004а. Анкетите в интернет: възможност за статистически изводи и интерпретиране на резултатите. Сп. „Социологически проблеми”, брой 3-4. Достъпна в Интернет на адрес: <http://kaloyan-haralampiev.info/wp-content/uploads/2010/03/statija9.pdf>
- Brethorst, L. 1990. An Introduction of Parameter Estimation Using Bayesian Probability. In: Maximum Entropy and Bayesian Methods, P. Fougere (ed.), Kluwer Academic Publishers, Dordrecht the Netherlands. Достъпна в Интернет на адрес: <http://bayes.wustl.edu/glb/intro.pdf>
- Haralampiev, K. 2008. Bayesian Inference of Relative Frequency (In the Case of Electoral surveys). Annuaire de l'universite de Sofia "St. Kliment Ohridski", Fakulte de philosophie, Livre – Sociologie, Tome 99. Достъпна в Интернет на адрес: <http://kaloyan-haralampiev.info/wp->

- [content/uploads/2010/03/article1.pdf](http://kaloyan-haralampiev.info/wp-content/uploads/2010/03/article1.pdf). Превод на статията на български език е достъпен в Интернет на адрес: <http://kaloyan-haralampiev.info/wp-content/uploads/2010/03/prevod-article1.pdf>
- Jaynes, E. 1986. Bayesian Methods: General Background. In: Maximum-Entropy and Bayesian Methods in Applied Statistics, J. H. Justice (ed.), Cambridge University Press, Cambridge. Достъпна в Интернет на адрес: [http://bayes.wustl.edu/etj/articles/general\\_background.pdf](http://bayes.wustl.edu/etj/articles/general_background.pdf)
- Jaynes, E. 1988. The Relation of Bayesian and Maximum Entropy Methods. In: Maximum-Entropy and Bayesian Methods in Science and Engineering, Vol. 1, G. J. Erickson and C. R. Smith (eds.), Kluwer, Dordrecht. Достъпна в Интернет на адрес: <http://bayes.wustl.edu/etj/articles/relationship.pdf>
- Jaynes, E. 2003. Probability Theory: the Logic of Science. Cambridge University Press, Cambridge. В този курс е цитирано непълно издание на тази книга, достъпно в Интернет на адрес: <http://omega.albany.edu:8008/JaynesBook.html>
- Loredo, T. 1990. From Laplace To SN 1987A: Bayesian Inference In Astrophysics. In: Maximum Entropy and Bayesian Methods, P. F. Fougere (ed), Kluwer Academic Publishers, Dordrecht. Достъпна в Интернет на адрес: <http://bayes.wustl.edu/gregory/articles.pdf>

*Допълнителна:*

- Найденов, Г., К. Харалампиев. 2012. Дълги вълни Кондратьева в Русия и България. В: Кондратьевски вълни: аспекти и перспективи, Издателство „Учител”, Волгоград
- Найденов, Г., К. Харалампиев. 2010. Цикличност в икономическото развитие на Русия в края на XIX и през XX век. В: Социологията – от емпирията към теорията. Юбилеен сборник, посветен на 70 години от рождението на чл.-кор. проф. д-р Атанас Атанасов. Академично издателство „Професор Марин Дринов”, София. Достъпна в Интернет на адрес: <http://kaloyan-haralampiev.info/wp-content/uploads/2010/03/statiya11.pdf>
- Найденов, Г., К. Харалампиев. 2006. Цикличността в равнището на цените на хранителните и нехранителните стоки в България през XX век. Сп. „Икономически алтернативи”, брой 6. Достъпна в Интернет на адрес: <http://alternativi.unwe.bg/alternativi/br21/2.pdf>

Дата: 17.02.2013

Съставил: доц. К. Харалампиев