200132 - EST - Statistics

Coordinating unit: 200 - FME - School of Mathematics and Statistics
Teaching unit: 715 - EIO - Department of Statistics and Operations Research
Academic year: 2017
Degree: BACHELOR'S DEGREE IN MATHEMATICS (Syllabus 2009). (Teaching unit Compulsory)
ECTS credits: 7,5  
Teaching languages: Catalan

Teaching staff

Coordinator: MARTA PÉREZ CASANY
Others: Segon quadrimestre:
MARTA PÉREZ CASANY - A, B
JOSEP ANTON SÁNCHEZ ESPIGARES - B
JORDI VALERO BAYA - A

Degree competences to which the subject contributes

Specific:
1. CE-2. Solve problems in Mathematics, through basic calculation skills, taking into account tools availability and the constraints of time and resources.
2. CE-3. Have the knowledge of specific programming languages and software.
3. CE-4. Have the ability to use computational tools as an aid to mathematical processes.

General:
5. CB-1. Demonstrate knowledge and understanding in Mathematics that is founded upon and extends that typically associated with Bachelor's level, and that provides a basis for originality in developing and applying ideas, often within a research context.
6. CB-2. Know how to apply their mathematical knowledge and understanding, and problem solving abilities in new or unfamiliar environments within broader or multidisciplinary contexts related to Mathematics.
7. CB-3. Have the ability to integrate knowledge and handle complexity, and formulate judgements with incomplete or limited information, but that include reflecting on social and ethical responsibilities linked to the application of their knowledge and judgements.
8. CG-1. Show knowledge and proficiency in the use of mathematical language.
10. CG-3. Have the ability to define new mathematical objects in terms of others already known and ability to use these objects in different contexts.
11. CG-4. Translate into mathematical terms problems stated in non-mathematical language, and take advantage of this translation to solve them.
12. CG-6 Detect deficiencies in their own knowledge and pass them through critical reflection and choice of the best action to extend this knowledge.

Transversal:
4. SELF-DIRECTED LEARNING. Detecting gaps in one's knowledge and overcoming them through critical self-appraisal. Choosing the best path for broadening one's knowledge.
A student that has completed this statistics course:

1. Is able to carry out and interpret basic descriptive statistics with statistical software.
2. Is able to perform statistical inference with statistical software and correctly interpret the results obtained.
3. Can express the difference between the two statistical schools: the frequentiest and the bayesian approach.
4. Is able to analytically obtain moment estimators, maximum likelihood estimators and bayesian estimators for the parameters of the most common probability distributions.
5. Is able to compare different estimators and select the best estimator according to some optimality criterion (bias, variance, mean squared error).
6. Is able to design an optimal test for particular hypothesis tests regarding parameters of distributions, applying the criterion of Neyman-Pearson and the generalized likelihood ratio.
7. Is able to formulate the difference between parametric and non-parametric tests.
8. Is able to apply the classical parametric tests (Z-test for the normal distribution, Student t test for independent samples and paired observations, F test for equality of variances) to data sets and interpret correctly the results obtained.
9. Is able to apply the most common non-parametric tests (Chi-square test for independence, sign test) to data sets and correctly interpret the results.
10. Is able to read and understand the descriptive statistics and the statistical inference used in a published scientific article.
## Study load

<table>
<thead>
<tr>
<th>Total learning time: 187h 30m</th>
<th>Hours large group: 45h 24.00%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours medium group: 0h 0.00%</td>
<td></td>
</tr>
<tr>
<td>Hours small group: 30h 16.00%</td>
<td></td>
</tr>
<tr>
<td>Guided activities: 0h 0.00%</td>
<td></td>
</tr>
<tr>
<td>Self study: 112h 30m 60.00%</td>
<td></td>
</tr>
</tbody>
</table>
## Content

<table>
<thead>
<tr>
<th><strong>1. INTRODUCTION</strong></th>
<th><strong>Learning time:</strong> 18h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description:</td>
<td>Theory classes: 4h</td>
</tr>
<tr>
<td>1.1. Descriptive statistics.</td>
<td>Laboratory classes: 3h</td>
</tr>
<tr>
<td>1.2. Population and sample.</td>
<td>Self study: 11h</td>
</tr>
<tr>
<td>1.3. Distributions related to the Normal distribution.</td>
<td></td>
</tr>
<tr>
<td>Related activities:</td>
<td>Theory classes: 6h</td>
</tr>
<tr>
<td>Some classes and three sessions in a computer room.</td>
<td>Practical classes: 4h</td>
</tr>
<tr>
<td>Specific objectives:</td>
<td>Self study: 20h</td>
</tr>
<tr>
<td>Carry out univariate and bivariate descriptive statistical analysis.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>2. POINT ESTIMATION</strong></th>
<th><strong>Learning time:</strong> 30h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description:</td>
<td>Theory classes: 6h</td>
</tr>
<tr>
<td>2.1. The method of moments.</td>
<td>Practical classes: 4h</td>
</tr>
<tr>
<td>2.2. The maximum likelihood method.</td>
<td>Self study: 20h</td>
</tr>
<tr>
<td>2.3. Bayesian estimation.</td>
<td></td>
</tr>
<tr>
<td>Related activities:</td>
<td>Theory classes and problem sessions.</td>
</tr>
<tr>
<td>Theory classes and problem sessions.</td>
<td></td>
</tr>
<tr>
<td>Specific objectives:</td>
<td>Construct estimators using various methods.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>3. EVALUATION OF ESTIMATORS</strong></th>
<th><strong>Learning time:</strong> 26h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description:</td>
<td>Theory classes: 6h</td>
</tr>
<tr>
<td>3.1. Properties of estimators:</td>
<td>Practical classes: 4h</td>
</tr>
<tr>
<td>bias, variance, mean squared error (MSE), sufficiency, consistency, efficiency.</td>
<td>Self study: 16h</td>
</tr>
<tr>
<td>3.3. Asymptotic properties of the maximum likelihood estimator.</td>
<td></td>
</tr>
<tr>
<td>Related activities:</td>
<td>Theory classes and problem sessions.</td>
</tr>
<tr>
<td>Theory classes and problem sessions.</td>
<td></td>
</tr>
<tr>
<td>Specific objectives:</td>
<td>Derive properties of estimators.</td>
</tr>
</tbody>
</table>
### 4. HYPOTHESIS TESTING

**Learning time:** 37h  
Theory classes: 7h  
Practical classes: 5h  
Self study : 25h

**Description:**
- 4.2. Neyman Pearson criterion for simple hypothesis.  
- 4.3. Extension of Neyman Pearson for unilateral alternatives.  
- 4.4. Test of the monotone likelihood function.  
- 4.5. Generalized likelihood ratio tests.  
- 4.6. $X^2$ test for contingency tables and goodness of fit.

**Related activities:**  
Theory classes and problem sessions.

**Specific objectives:**  
Development of hypothesis tests.

### 5. INTERVAL ESTIMATION

**Learning time:** 8h 10m  
Theory classes: 2h  
Practical classes: 2h  
Self study : 4h 10m

**Description:**
- 5.1. Confidence intervals.  
- 5.2. Intervals associated to pivotal quantities.  
- 5.3. Intervals associated to the Normal model.

**Related activities:**  
Theory classes, problem sessions and laboratory sessions.

**Specific objectives:**  
Construction of confidence intervals.
6. THE LINEAR MODEL and GENERALIZED LINEAR MODEL

**Description:**
6.1. Simple and multiple linear regression.
6.3 Multicolinearity. Leverage. Influential observations.
6.4 Goodness of fit and coefficient of determination.
6.5 Prediction.
6.5. Residual analysis.
6.7 Anova
6.8 Ancova
6.9 Generalized linear model. Binomial and Poisson cases.

**Related activities:**
Laboratory practicals.

**Specific objectives:**
Apply linear regression and interpret the results.

---

**Qualification system**

The assessment comprises the following elements: final exam, midterm exam, questionnaires, deliverable exercises. The final exam and the midterm exam consist of open theoretical questions and problems to solve. There will be two questionnaires that are small exams that have to be done in class and have one hour of duration. The deliverable exercises are three. They have to be done individually and they will be known approximately two weeks before. The continuous assessment mark (CAM) is calculated as:

\[
CAM = 0.5 \times M_{Final} + 0.25 \times M_{Midterm} + 0.125 \times M_{Questionnaires} + 0.125 \times M_{Deliverables}
\]

The final subject mark (FM) is the maximum between the CAM mark and the final exam mark: \( FM = \max(CAM, M_{Final}) \)

An extra exam will take place on July for students that failed during the regular semester.
Bibliography

Basic:


Complementary:


Others resources:

Hyperlink

R-software: www.r-project.org

Resource