PROBABILITY AND MATHEMATICAL STATISTICS (25 hours)

1. Descriptive analysis of data
- Types of data (numerical: discrete and continuous, categorical: nominal, ordinal)
- Tables and graphs (frequency tables, bar chart, histogram, *stem and leaf* diagram, dotplot, lineplot)
- Measures of locations (mean, median, mod)
- Measures of spread (standard deviation, moments, range, interquartil, boxplot)
- Measure of symmetry (skewness)

2. Random variables
- Discrete random variables – discrete distributions (discrete probability (density) functions, cumulative distribution function (CDF))
- Continuous random variables – continuous distribution (probability density function (PDF), CDF)
- Mathematical expectation, variance and standard deviation, moments (linear transformations of a random variable)
- Examples of discrete distributions (uniform, Bernoulli, binomial, Poisson, geometric, negative binomial, hypergeometric)
- Examples of continuous distributions (uniform, Γ, exponential, χ^2, Beta, normal)
- Functions of a random variable

3. Generating functions
- Probability generating functions (PGF) (examples: uniform, Bernoulli, binomial, negative binomial, hypergeometric, Poisson distributions, evaluating moments)
- Moment generating functions (MGF) (connection with PGF, application on the important continuous distributions)
- Cumulant generating functions (CGF)
- PGF, MGF and CGF of linear transforms of a random variable

4. Joint distributions of random variables
- Joint probability density functions (tables of, marginal probability density functions)
- Conditional probability density functions
- Independence of random variables
- Mathematical expectations of functions of two variables (expectations of sums and products)
- Covariance and correlation coefficient
- Variance of a sum of random variables (independent random variables)
- Convolutions
- Moments of linear combinations of random variables (independent random variables, examples)

5. Central limit theorem and applications
- The case of independent identically distributed random variables
- Normal approximations of binomial, Poisson and Γ-distributions (continuity correction)

6. Sampling and statistical inference
- Population and random sample (population parameters, statistics)
- Moments of the sample mean and variance
- Sampling distributions for the normal model (distribution of the sample mean: exact and on the large sample basis, distribution of the sample variance, independence of the sample mean and variance, Student's *t*-distribution, Fisher's *F*-distribution)

7. Point estimation
- The method of moments
- The method of maximum likelihood (MLE) (the finite-parameter case, incomplete and independent samples)
- Unbiasedness of an estimator
- Mean square error of an estimator and consistency
- Asymptotic distribution of MLE

8. Confidence intervals (CI)
- The definition of CI (confidence limits)
- Derivation of CI (the pivotal method, sample size)
- CI for normal parameters
- CI for parameters of binomial and Poisson distributions (by pivotal method for small samples, by normal approximation for large samples)
- CI for two sample problems (independent samples, the normal means and variances, two population proportions, two Poisson parameters, paired data)

9. Testing statistical hypothesis
- Statistical hypothesis, statistical test and kinds of errors (null and alternative hypothesis for parametric models, simple and composite hypothesis, test-statistics, critical regions, errors of first and second kind)
- Significance and *P*-values (power of test, the best test (Neyman-Pearson's lemma))
- Basic tests based on single sample (testing the values of: the mean and variance of a normal population, a population proportion, the mean of a Poisson distribution: exact and on the large sample basis)
- Basic tests based on two independent samples (testing the value of: the difference between two normal means, the ration of two normal variances, the difference between two population proportions, the difference between two Poisson means: exact and on the large sample)
- Basic test based on paired data
- Comparison of tests and CI

- χ^2-test (goodness of fit, contingency tables, test for independence, test of homogeneity)

10. Correlation and regression analysis
- Relationships between two random variables (scatter plot, linear relationship, examples)
- Correlation analysis (sample correlation coefficient, normal model and inference: testing the value of correlation coefficient)
- Regression analysis: simple linear model (fitting the model, partitioning the variability of the responses, coefficient of determination, Gauss-Markov conditions, the normal model: parametric inference, model checking and predicting, transformation of data)
- The multiple linear regression (the method of least squares)

11. Analysis of variance
- One-way analysis of variance (ANOVA) (estimation of the parameters, partitioning the variability, model checking by analysis of residual, CI of the treatment means and their differences, ANOVA table, examples)

12. Conditional expectation
- Conditional expectation (as regression function and random variable)
- Conditional variance

References:

1. F. Daly, D.L. Hand, M.C. Jones, A.D. Lunn, K.J. McConway, *Elements of Statistics*, Addison-Wesley, 1995.
2. E.L. Lehmann, *Testing Statistical Hypotheses,* 2nd edition, Springer, 1997.
3. E.L. Lehmann, G. Casella, *Theory of Point Estimation*, 2nd edition, Springer, 1998.
4. Ž. Pauše, *Uvod u matematičku statistiku*, Školska knjiga, Zagreb, 1993.
5. I. Šošić, V. Serdar, *Uvod u statistiku*, Školska knjiga, Zagreb, 1992.
6. J.E. Freund, *Mathematical Statistics*, Prentice Hall International, 1992.
7. *Subject*101: *Statistical Modelling*, *Core Reading 2000*, Faculty and Institute of Actuaries
8. *Subjects* C1/2: *Statistics, Core Reading 1996,* Faculty and Institute of Actuaries