University of Toronto

Department of Computer & Mathematical Sciences

STAB57: Introduction to Statistics

Things you should know for the final

taught by Louis de Thanhoffer de Volcsey -email me
-website

---

this document contains a detailed list of the things you should know in order to prepare for the final exam
the exam takes place on Wednesday, December 13th at 9am in IC 130

Inferences based on the MLE, §6.3

- You should know the 3 types of inference based on the MLE
- You should know the definition of a confidence interval
- You should know the formula for a confidence interval in the case of a location, location-normal, and understand the formula for the Bernouilli model (6.3.6,6.3.7,6.3.8)
- You should understand the concept of testing a hypothesis through the use of a p-value and know how to perform this test in the location normal model/Bernoulli model
- You should understand the relation between a confidence interval and a hypothesis test as illustrated by example 6.3.12

Introducing Bayesian Models §7.1

- You should understand the example of batting averages discussed in class to motivate Bayesian models.
- You should know the two definitions of a Bayesian model we discussed (§7.1 and week 8 ass.)
- you should know the various definitions associated to a bayesian model: joint, prior, posterior, prior predictive, conjugate prior etc... (§7.1, Def. §7.1.1, §7.4.1)
- You should understand the Bernoulli model 7.1.1 in detail (especially why the β-distribution pops up see week 8 ass. problem 4)
- You should understand the location normal model 7.1.2 in detail
- You should know the properties of the β-distribution and be able to use them in calculations

Bayesian Inference §7.2

- You should understand the three forms of inference based on the MAD
- You should know how to compute the posterior mode and mean as illustrated in examples 7.2.2, 7.2.3, 7.2.4 (also know the relation between mode and mean)
- You should know the definition of credible intervals and understand the difference with confidence intervals as illustrated by the example of cookie jars we discussed in class (ie "rows vs. columns)
- You should be able to compute credible intervals in certain situations such as 7.2.7
- You should know principle of hypothesis testing and the associated p-value in the discrete case
- You should know the modified p-value in the continuous case 7.2.5 and how to compute it (first part of 7.2.11)
- You should understand why and how one replaces the prior by a *mixture* to get another expression for the p-value 7.2.7
- You should know what the odds of a probability are as well as be able to work with them

- You should know the definition of a Bayes factor and its interpretation, the relation with the p-value 7.2.9 and Thm 7.2.1.

Regression 10.3.2

- You should understand the intuitive idea that motivates simple linear regression
- You should be able to use calculus to compute gradients and minima/maxima for functions of multiple variables
- You should understand the geometric problem that underlies regression and how we deduce its solution Thm 10.3.1
- You should understand how regression is also the solution to finding the MLE of a certain statistical model using the above.
- You should know and understand why the parameters $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased. You should also know the formula for the variance of of these estimators Th, 10.3.2, 10.3.3
- You should know the basic terminology used in this context: weights, residuals... as well as their properties

Anova analysis (§10.3 cont)/extra notes

- You should know the definition of the SSR, SSE, SST, their relation and why it holds. (10.3.1)
- You should know what the R^2 coefficient is and its properties
- You should know that the s^2 estimator is unbiased (both in the regression context 10.3.4 as likelihood cor. 4.6.2)
- You should know what the F-statistic is
- You should know what the associated p-value is and how to build a hypothesis test to see if the two variables are related (10.3.8)-(10.3.9)
- You should be able to summarize all this data into an *anova*-table