FINAL EXAM

ISyE6420
Spring 2020

Released April 20, 2020, 12:00pm – due April 28, 2020, 11:55pm. This exam is not proctored and not time limited except the due date. Late submissions will not be accepted.

Use of all available electronic and printed resources is allowed except direct communication that violates Georgia Tech Academic Integrity Rules.

Name _____

Problem	Naïve Bayes Revisited	Daphnia Magna	Amantadine	Total
Score	/33	/33	/34	/100

1. Revisiting Jane, Michael, and Melissa from Homework 1. Logistic regression is a superb ML tool for classification tasks when covariates are present. Recall the question from HW1 where the probabilities that Jane, Michael, and Melissa will go to the beach were calculated.

Using Bayesian Logistic Regression predict the probability classes for Jane, Michael, and Melissa. Data are available in starter file `jmmatbeach0.odc`. Compare the results with their counterparts from HW1. Do the results for the three probabilities from HW1 fall in the 95% Credible Set?

Hint: Since sample size ($n = 100$) is modest for a regression with 5 predictors, the priors for coefficients for logit term should not be very flat. Normal priors with 0 mean and precision 0.5 for all coefficients in the linear part are recommended.

2. Regression to Predict Acute Aquatic Toxicity Towards Daphnia Magna. This dataset was used to develop quantitative regression in order to predict acute aquatic toxicity towards Daphnia Magna in a set of chemicals. The LC50 data, which is the concentration that causes death in 50% of test D. magna over a test duration of 48 hours, was used as model response. The model comprised 8 molecular descriptors: TPSA(Tot) (Molecular properties), SAacc (Molecular properties), H-050 (Atom-centred fragments), MLOGP (Molecular properties), RDCHI (Connectivity indices), GATS1p (2D autocorrelations), nN (Constitutional indices), C-040 (Atom-centred fragments). Details can be found in Cassotti et al. (2014)[1]

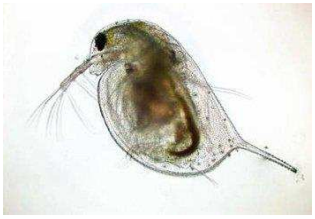

Figure 1: Daphnia Magna (Photo courtesy of Norwegian Institute for Water Research)

Attribute Information: Table below provides 8 molecular descriptors and 1 quantitative experimental response for 546 experiments. The WinBUGS starter file is DaphniaMagna0.odc.

[1]M. Cassotti, D. Ballabio, V. Consonni, A. Mauri, I. V. Tetko, R. Todeschini (2014). Prediction of acute aquatic toxicity towards daphnia magna using GA-kNN method, Alternatives to Laboratory Animals (ATLA), 42,31:41; doi: 10.1177/026119291404200106.

	Molecular Descriptor Predictors
1	TPSA(Tot)
2	SAacc
3	H-050
4	MLOGP
5	RDCHI
6	GATS1p
7	nN
8	C-040
9	Quantitative response, LC50 [-LOG(mol/L)]

(a) Predict LC50 for variables TPSA = 12.5, H-050 = 0.4, MLOGP = 1.5, and RDCHI = 1. The other four predictors are equal to 0. Give both prediction for a mean response and prediction for a new observation.

(b) Can any of the predictors be excluded? Use 95% Credible Set Argument.

(c) Find Bayesian R-square. Comment.

3. Meta Analysis of Amantadine for Treatment of Influenza A. The flu (influenza) can be caused by many different viruses. One type is influenza A, with headaches, coughs and runny noses that can last for many days and lead to serious illnesses such as pneumonia. Drug amantadine is FDA-approved for treatment and chemoprophylaxis of influenza A virus infections among adults and children aged one year and older. The review of trials found that amantadine is helpful in preventing the flu, but only when there is a high probability that the cause of the flu is influenza A (a known epidemic, for example). However, this drug has adverse gastrointestinal (stomach and gut) effects, and in rare cases can also have serious effects on the nervous system.

Trial	Drug (n/N)	Placebo (n/N)
Oker-Blom (1970)	16/141	41/152
Muldoon (1986)	1/53	8/52
Monto (1979)	8/136	28/139
Kantor (1980)	9/59	9/51
Pettersson (1980)	32/95	59/97
Quarless (1981)	15/107	20/99
Dolin (1982)	2/113	27/132
Reuman (1989)	3/317	5/159

Table 1: Eight trials of amantadine for prevention of influenza. Outcome is proportion of cases of influenza n, out of N subjects.

In the meta-analysis of randomised controlled trials of amantadine for preventing influenza (Table 1)[2] the treatment effects in the eight trials are variable: the reduction in

[2]Jefferson, T. O., Demicheli, V., Deeks, J. J., Rivetti, D. (2002). Amantadine and rimantadine for

odds of influenza vary from 16% to 93%, with some of the confidence intervals not overlapping. Classical methods are often poor at detecting true heterogeneity, important for finding ultimate summary of the studies.

(a) Conduct Bayesian meta analysis of these eight studies. Use WinBUGS. Explain how Bayesian analysis accounts for heterogeneity of studies.

(b) Find in the literature how meta analyses are graphically presented. Construct a plot for your meta analysis from (a).

Hint: Consult the Blocker example from WinBUGS Examples Vol I. For the plot in (b) export results from WinBUGS to your favorite plotting program.

preventing and treating influenza A in adults. Cochrane Database Syst Rev 2002;(4): CD001169. [PubMed]

Higgins, J. P. T., Thompson, S. G., Deeks, J, J., and Altman, D. G. (2003). Measuring inconsistency in meta-analyses `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC192859/` *BMJ*, 6, 327(7414): 557560. `doi:10.1136/bmj.327.7414.557`, PMCID: PMC192859