Stat 406 Spring 2007: Homework 4

Out Fri 2 Feb, back Mon 12 Feb

1 Gaussian classifiers with missing data

Consider a classifier with three 4-dimensional Gaussian class-conditional densities with parameters `mu` and `Sigma` stored in the file `gaussClassifMissingData.mat`.

1. Suppose you observe that $x_4 = -0.7859$, but you do not see the first three components $x_{1:3}$. What is $p(y|x_4, \mu, \Sigma)$, where the class prior is uniform ($p(y = c) = 1/3$) and the parameters are as above? Show your work. Hint: I get $p(y|x_4, \mu, \Sigma) = [0.2233, 0.2946, 0.4821]$.
2. Suppose you also observe that $x_3 = -0.9641$. What is $p(y = c|x_3, x_4, \mu, \Sigma)$?
3. Suppose you also observe that $x_2 = -1.3191$. What is $p(y = c|x_{2:4}, \mu, \Sigma)$?
4. Suppose you also observe that $x_1 = -1.4885$. What is $p(y = c|x_{1:4}, \mu, \Sigma)$?
5. Explain qualitatively what happens to the class posterior as you see more features. Which class would you predict?

2 Gaussian posterior credible interval

Let $X \sim \mathcal{N}(\mu, \sigma^2 = 4)$ where μ is unknown but has prior $\mu \sim \mathcal{N}(\mu_0, \sigma_0^2 = 4)$. The posterior after seeing n samples is $\mu \sim \mathcal{N}(\mu_n, \sigma_n^2)$. (This is called a credible interval, and is the Bayesian analog of a confidence interval.) How big does n have to be to ensure

$$p(\mu \in I(\mu_n)|D) \geq 0.95 \tag{1}$$

where $I(\mu_n)$ is an interval (centered on μ_n) of width 1, and D is the data (see Figure 1)? Hint: recall that 95% of the probability mass of a Gaussian is within $\pm 1.96\sigma$ of the mean.

3 Empirical Bayes for diagnosing a biased coin machine

Consider the problem of testing for a based coin machine in the handout. Suppose we conduct $n = 16$ trials, tossing the coin $t = 5$ times on each trial, and we measure the following number of heads per trial:

$$X_{1:16} = [0, 0, 1, 0, 2, 0, 1, 1, 3, 2, 0, 1, 0, 1, 2, 0] \tag{2}$$

1. What is the method of moments estimate of $\hat{\alpha}$ and $\hat{\beta}$? (Hint: I get $\hat{\alpha} = 3.5$)

Figure 1: Gaussian posterior centered on μ_n. The shaded interval contains 95% probability mass and has width 1.

2. What is the expected probability of heads of a new coin? i.e, compute

$$E[\theta_* | \hat{\alpha}, \hat{\beta}] \tag{3}$$

Hint: recall that $\theta_* \sim Be(\hat{\alpha}, \hat{\beta})$.

3. What is the probability that the machine generates coins that are biased away from 50:50 by more than 0.1? i.e., compute

$$p(\theta^* \notin [0.4, 0.6] | \hat{\alpha}, \hat{\beta}) \tag{4}$$

Hint: use the `betacdf` command.

4. If you generate a new coin from the machine and roll it 5 times, how many heads do you expect to see, and with what probability? i.e., compute the posterior predictive distribution

$$p(X_* | X_{1:n}) \approx p(X_* | \hat{\alpha}, \hat{\beta}) = Bb(X_* | \hat{\alpha}, \hat{\beta}, 5) \tag{5}$$

(This is just a histogram of 6 numbers for $X_* \in \{0, 1, \ldots, 5\}$.)