ANLY-601 Spring 2018

Assignment 5
Due Thursday, April 5 , 2018 — in class

You may use your class notes, the text, or any calculus books — please use no other references (including internet or other statistics texts). If you use Mathematica to derive results, you must include the notebook with your solution so I can see what you did.

1. **Estimating the parameter of a binomial distribution**

   I give you a coin and ask you to estimate the probability of a toss resulting in heads. You decide to construct the estimate by tossing the coin N times and counting the number of heads n.

   (a) The probability of recording n heads from N tosses is given by the *binomial distribution*

   $$p(n|N,\alpha) \;=\; \begin{pmatrix} N \\ n \end{pmatrix} \alpha^n \, (1-\alpha)^{(N-n)} \;=\; \frac{N!}{n!\,(N-n)!}\, \alpha^n \, (1-\alpha)^{(N-n)} \qquad (1)$$

   where α is the probability of heads coming up on any particular toss – it is the quantity we want to estimate.

   Derive the m [...] α that maximizes the log probability ([...] s and observing n heads. (Notice that t [...] n.)

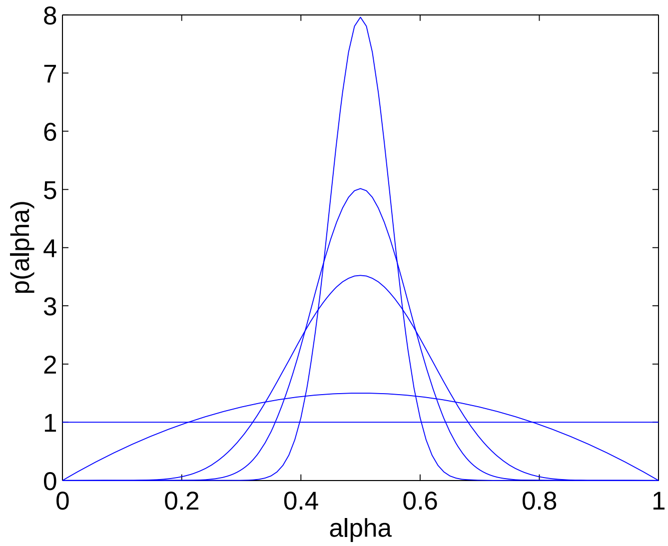
   (b) Suppose now you decide to construct a MAP estimate of [...] t the beta distribution

   $$p(\alpha) \;=\; \frac{1}{B(a,b)}\, \alpha^{(a-1)} \, (1 \ldots \infty \qquad (2)$$

   is a conjugate prior[1] for the likelihood function in (1) **and derive the MAP estimate of** α . (The normalization factor $B(a,b)$ is the Euler beta function – but you don't need knowledge of this function to complete the problem.)

   (c) Since there's no particular reason to believe that the coin I hand you is grossly unfair, you opt to pick a and b so that the prior $p(\alpha)$ is maximum at 1/2 and symmetric about that value. To achieve this, you set $a = b$. For this choice, it's clear that with $a = b = 1$ the prior is flat. For larger values of $a = b$, the prior distribution gets progressively more peaked up about $\alpha = 1/2$; the variance of the *prior* is $var(\alpha) = 1/(4\,(2\,a+1))$. The prior distribution is plotted below for several choices of $a = b$.

[1] Recall from your notes that for a *conjugate prior* $p(\alpha)$, the posterior density $p(\alpha|D) \propto p(D|\alpha)\,p(\alpha)$ is of the same algebraic form (here a beta distribution) as the prior density $p(\alpha)$.



Use the formula for the MAP estimate of α that you derived in part (b), together with the above information about the beta distribution to discuss how the value of a (with $b = a$) in the prior distribution effects the MAP estimate of α relative to the maximum likelihood estimate of α. That is, discuss how the change in the shape of the prior distributio

2. **Fitting Constra**

If you're asked to fit a Gaussian distribution to a set of m, ensional data points $D = (x^{(1)}, x^{(2)}, \ldots, x^{(m)})$, you know that the maximum li nd covariance for the model are

$$\hat{\mu} = \frac{1}{m} \sum_{i=1}^{m} x^{(i)}$$

and

$$\hat{\Sigma} = \frac{1}{m} \sum_{i=1}^{m} (x^{(i)} - \hat{\mu})\,(x^{(i)} - \hat{\mu})^T$$

respectively.

Suppose now that you are told that the *model covariance* matrix is constrained to be a (positive) *constant times the identity matrix*

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 & 0 & \ldots \\ 0 & \sigma^2 & 0 & \ldots \\ 0 & 0 & \sigma^2 & \ldots \\ \vdots & & & \end{pmatrix} \quad . \tag{3}$$

(a) Write down the log-likelihood L of the set of $n-$dimensional data vectors D under this model.

(b) Derive the maximum likelihood estimate of Σ.

3. **Error, Bias, Variance and Paintball**

You're making final selections for members for your paintball team in preparation for the national championship. You have your final two contestants fire six shots into a target with the results below:

(A) (B)

Which contestant do you choose for your team? Suppose the guns have adjustable sights, does that change yo