The spatial arrangement of cones in the fovea: Bayesian analysis

David J.C. MacKay
University of Cambridge
Cavendish Laboratory
Madingley Road
Cambridge CB3 0HE
mackay@mrao.cam.ac.uk

March 25, 1993

Abstract

Is the distribution of cones in the fovea random? We can never prove that this is so; we can only assess how probable this theory is relative to explicit alternative models, given the data. That paper does so for one simple alternative model of correlation.

Readers of Mollon and Bowmaker's letter in Nature may have felt sceptical about their assertion that the distribution of long and middle wave cones 'is random' given their data. Their χ^2 test only used about a third of the information in the data about neighbouring cones. It would seem desirable to make full use of the data. Furthermore it might be argued that it is never possible to show that a phenomenon *is random* — only that the alternative models that have been studied for the phenomenon are less probable than the random model. In a Bayesian approach, explicit alternative models are constructed, and their relative probabilities can be evaluated in the light of the data. The Bayesian approach makes full use of *all* the relevant information in the data, and can be applied to any data set, no matter how quirky. The method of inference is mechanical once the alternative models have been defined. Bayesian inference automatically penalises excess parameters, so there is no need to fear being 'duped' into accepting over-complex models.

Three models are studied to account for the data in [1]. All three models ignore 'short' (blue) cones, treating them as vacancies in a lattice of 'long' and 'medium' cones (henceforward referred to colloquially as red and green cones).

$\mathcal{H}_1$: random

The null hypothesis H_1 is that the red and green cones occur independently with fixed probability. Thus the probability of a particular arrangement of cells, $\mathbf{x}$, is:

$$P(\mathbf{x}|\theta, \mathcal{H}_1) = \prod_i \frac{e^{\theta x_i}}{e^{\theta} + e^{-\theta}} \tag{1}$$

where $x_i = 1$ when cone i is red, and $x_i = -1$ when it is green. The parameter θ measures the bias in favour of red cones; for this model, the probability that one cell is red is $p_R = \frac{e^{\theta}}{e^{\theta}+e^{-\theta}}$. In principle, we could infer θ given the data, but clearly the data are compatible with the assertion that $\theta = 0$. The attention here is focussed on the question of whether there are spatial correlations in the lattice of cones, so in the following the parameter θ will be set to zero (but the analysis could easily be repeated with non–zero θ). When the identity $\mathbf{x}_O$ of N cones is observed, the *likelihood* of model $\mathcal{H}_1$ is therefore

$$P(\mathbf{x}_O|\mathcal{H}_1) = \frac{1}{2^N}. \tag{2}$$

That is, under $\mathcal{H}_1$, all data sets are equally probable.

$\mathcal{H}_2$: constant spatial correlations

The second hypothesis asserts that there are spatial correlations or anticorrelations among the cones, described by an additional parameter, w. The probability that a lattice of cells takes on an arrangement $\mathbf{x}$ is

$$P(\mathbf{x}|\theta, w, \mathcal{H}_2) = \frac{e^{\theta \sum_i x_i + \frac{w}{2} \sum_{i\sim j} x_i x_j}}{Z(\theta, w)} \tag{3}$$

The notation "$i \sim j$" indicates that the sum is over cones i and j that are adjacent. This model will be referred to as the Ising model, since formally it is identical to the statistical model of a lattice of magnetic spins with a constant coupling between neighbours.

What predictions does this model make for the small fragments of retina in the data set? Well, if we neglect the effect of the surrounding 'sea' of cells on the fragment, an approximation for the probability over possible arrangements $\mathbf{x}_O$ of the observed cells is:

$$P(\mathbf{x}_O|\theta, w, \mathcal{H}_2, \text{isolated}) = \frac{e^{\theta \sum_{i\in O} x_i + \frac{w}{2} \sum_{i\sim j\in O} x_i x_j}}{Z_O(\theta, w)}, \tag{4}$$

where

$$Z_O(\theta, w) = \sum_{\mathbf{x}_O} e^{\theta \sum_{i\in O} x_i + \frac{w}{2} \sum_{i\sim j\in O} x_i x_j}. \tag{5}$$

However, strictly, the surrounding sea of cells will influence the observed cells (if $w \neq 0$), so the true predictive distribution is written in terms of a sum over states $\mathbf{x}_S$ of the surrounding sea:

$$\begin{aligned} P(\mathbf{x}_O|\theta, w, \mathcal{H}_2, \text{immersed}) &= \sum_{\mathbf{x}_S} P(\mathbf{x}_O, \mathbf{x}_S|\theta, w, \mathcal{H}_2, \text{in sea}) \\ &= \frac{Z_{S|\mathbf{x}_O}(\theta, w)}{Z_{S,O}(\theta, w)} \end{aligned}$$

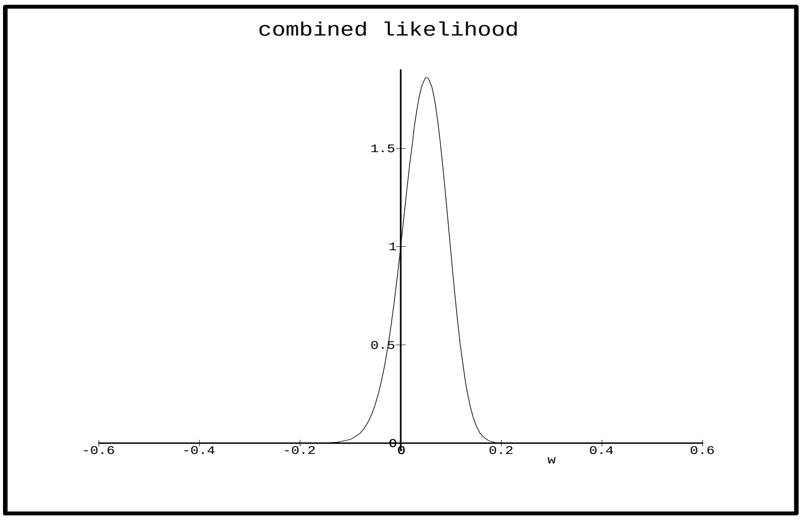
where the total partition function $Z_{S,O}$ and the conditional partition function $Z_{S|\mathbf{x}_O}$ are:

$$Z_{S,O}(\theta, w) = \sum_{\mathbf{x}_S, \mathbf{x}_O} e^{\theta \sum_i x_i + \frac{w}{2} \sum_{i\sim j} x_i x_j} \tag{6}$$

$$Z_{S|\mathbf{x}_O}(\theta, w) = \sum_{\mathbf{x}_S} e^{\theta \sum_{i\in S} x_i + \frac{w}{2} \sum_{i\sim j\in S} x_i x_j} \tag{7}$$

The 'isolated' likelihood function (11) is relatively easy to evaluate, and it may be a good approximation to the 'immersed' likelihood (13). In this study I have evaluated the isolated likelihood function, and also the immersed likelihood function with the sea approximated by a narrow 'moat' surrounding the fragment.

1 notes



2 Second draft

> Is the distribution of cones in the fovea random? We can never prove that this is so; we can only assess how probable this theory is relative to explicit alternative models, given the data. This paper does so for one simple alternative model of correlation.

In their letter in Nature [1], Mollon and Bowmaker describe the results of experiments measuring the spatial distribution of the different types of cone photoreceptor ('long', 'medium' and 'short') in the retina. They asserted on the basis of their data that the distribution of long and middle wave cones 'is random'. However, it might be argued that it is never possible to show that a phenomenon *is random* — only that the alternative models that have been studied for the phenomenon are less probable than the random model. In fact Mollon and Bowmaker's χ^2 test only used about a third of the information in the data about neighbouring cones, and implicitly made a comparison with an alternative model incorporating correlations in one direction only. Their test was not applicable to data fragments with lots of gaps. It would seem desirable to make full use of the data, and consider an alternative model allowing correlations in all directions. In a Bayesian approach, explicit alternative models are constructed, and their relative probabilities are evaluated in the light of the data. The Bayesian approach makes full use of *all* the relevant information in the data, and can be applied to any data set, no matter how quirky. The method of inference is mechanical once the alternative models have been defined.

In this chapter, three models are studied to account for the data in [1]. All three models ignore 'short' (blue) cones, treating them as vacancies in a lattice of 'long' and 'medium' cones (henceforward referred to colloquially as red and green cones). The final conclusion turns out to be the same as that of Mollon and Bowmaker's paper: no model more probable than randomness has been found. But this problem provides an interesting case study of Bayesian methods.

$\mathcal{H}_1$: random

The null hypothesis H_1 is that the red and green cones occur independently with fixed probability. Thus the probability of a particular arrangement of cells, $\mathbf{x}$, is:

$$P(\mathbf{x}|\theta, \mathcal{H}_1) = \prod_i \frac{e^{\theta x_i}}{e^{\theta} + e^{-\theta}} \tag{8}$$

where $x_i = 1$ when cone i is red, and $x_i = -1$ when it is green. The parameter θ measures the bias in favour of red cones; for this model, the probability that one cell is red is $p_R = \frac{e^{\theta}}{e^{\theta}+e^{-\theta}}$. In principle, we could infer θ given the data, but clearly the data are compatible with the assertion that $\theta = 0$. The attention here is focussed on the question of whether there are spatial correlations in the lattice of cones, so in the following the parameter θ will be set to zero (but the analysis could easily be repeated with non–zero θ). When the identity $\mathbf{x}_O$ of N cones is observed, the *likelihood* of model $\mathcal{H}_1$ is therefore

$$P(\mathbf{x}_O|\mathcal{H}_1) = \frac{1}{2^N}. \tag{9}$$

That is, under $\mathcal{H}_1$, all data sets are equally probable.

$\mathcal{H}_2$: constant spatial correlations

The second hypothesis asserts that there are spatial correlations or anticorrelations among the cones, described by an additional parameter, w. The probability that a lattice of cells takes on an arrangement $\mathbf{x}$ is

$$P(\mathbf{x}|\theta, w, \mathcal{H}_2) = \frac{e^{\theta \sum_i x_i + \frac{w}{2} \sum_{i \sim j} x_i x_j}}{Z(\theta, w)} \tag{10}$$

The notation "$i \sim j$" indicates that the sum is over cones i and j that are adjacent. This model will be referred to as the Ising model, since formally it is identical to the statistical model of a lattice of magnetic spins with a constant coupling between neighbours.

What predictions does this model make for the small fragments of retina in the data set? Well, if we neglect the effect of the surrounding 'sea' of cells on the fragment, an approximation for the probability over possible arrangements $\mathbf{x}_O$ of the observed cells is:

$$P(\mathbf{x}_O|\theta, w, \mathcal{H}_2, \text{isolated}) = \frac{e^{\theta \sum_{i \in O} x_i + \frac{w}{2} \sum_{i \sim j \in O} x_i x_j}}{Z_O(\theta, w)}, \tag{11}$$

where

$$Z_O(\theta, w) = \sum_{\mathbf{x}_O} e^{\theta \sum_{i \in O} x_i + \frac{w}{2} \sum_{i \sim j \in O} x_i x_j}. \tag{12}$$

```
. . . . . . . .
 . . r . . . . .
  . r r . r g . .
   . r r g g g r .
    . r g g g . r .
     . . g r r g r .
      . . g . r g . .
       . . . . . . . .
```

```
. . . . . . .
 . . . g g r .
  . r g b g g .
   . r r g g g .
    . g r r g g .
     . r r g r . .
      . r g g r r .
       . . . r g g .
. . . . : : : . . . .
 . g g r . . .
  . r g r g . .
   . r r g g . .
    . r g b g g .
     . r g g g g .
      . r g g r r .
       . r r g g g .
        . r r r g b .
         . r r g r . .
          . . r r g . .
           . . . g r . .
            . . . . r . .
             . . . . . . .
```

```
. . . . . . . . . .
 . . . . . r r . . .
  . . . . . r g r . .
   . . . g . g r . g .
    . . . g r g g r . .
     . . . g r g r . . .
      . . r g g g g . . .
       . . g r r b r . . .
        . r g g g g . . . .
         . . . . r . . . . .
          . . . . . . . . . .
```

```
. . . . . . . . .
 . . . . g g . . .
  . . . . g r g g .
   . . . . g g g g .
    . r r g g g b r .
     . g . g g g r . .
      . . . r g r r . .
       . . . g r g . . .
        . . g r r r . . .
         . . r r r r . . .
          . . g r g g . . .
           . . g r r r . . .
            . . g g r . . . .
             . . . . . . . . .
```

Figure 1: Mollon and Bowmaker's data: p24, p29, p33, p44, p47

However, strictly, the surrounding sea of cells will influence the observed cells (if $w \neq 0$), so the true predictive distribution is written in terms of a sum over states $\mathbf{x}_S$ of the surrounding sea:

$$\begin{aligned} P(\mathbf{x}_O|\theta, w, \mathcal{H}_2, \text{immersed}) &= \sum_{\mathbf{x}_S} P(\mathbf{x}_O, \mathbf{x}_S|\theta, w, \mathcal{H}_2, \text{in sea}) \\ &= \frac{Z_{S|\mathbf{x}_O}(\theta, w)}{Z_{S,O}(\theta, w)} \end{aligned}$$

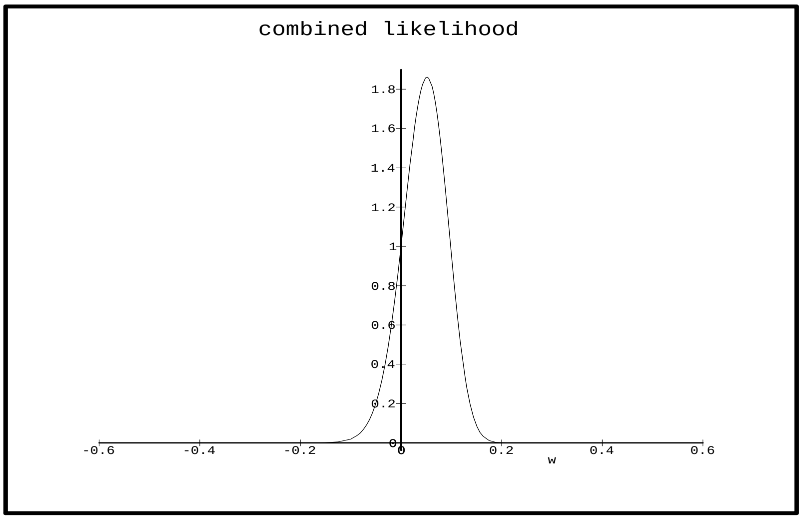
where the total partition function $Z_{S,O}$ and the conditional partition function $Z_{S|\mathbf{x}_O}$ are:

$$Z_{S,O}(\theta, w) = \sum_{\mathbf{x}_S, \mathbf{x}_O} e^{\theta \sum_i x_i + \frac{w}{2} \sum_{i \sim j} x_i x_j} \tag{13}$$

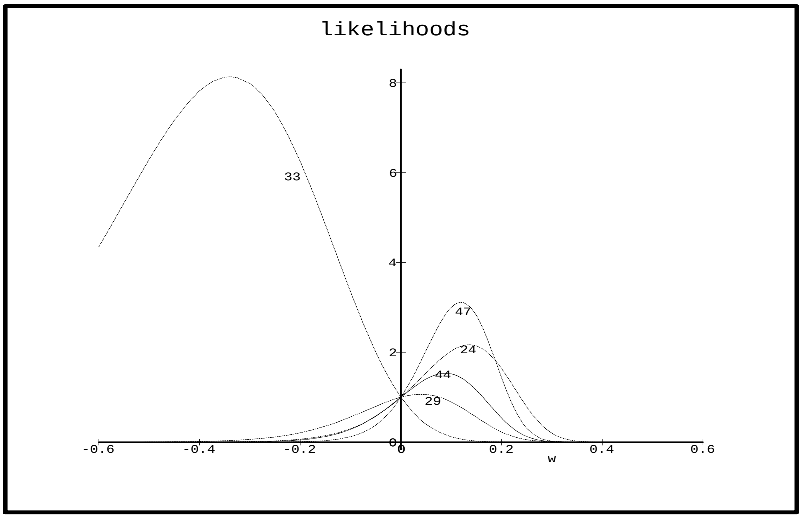
$$Z_{S|\mathbf{x}_O}(\theta, w) = \sum_{\mathbf{x}_S} e^{\theta \sum_{i \in S} x_i + \frac{w}{2} \sum_{i \sim j \in S} x_i x_j} \tag{14}$$

The ‘isolated’ likelihood function (11) is relatively easy to evaluate, and it may be a good approximation to the ‘immersed’ likelihood (13). In this study I have evaluated the isolated likelihood function, and also the immersed likelihood function with the sea approximated by a narrow ‘moat’ surrounding the fragment.

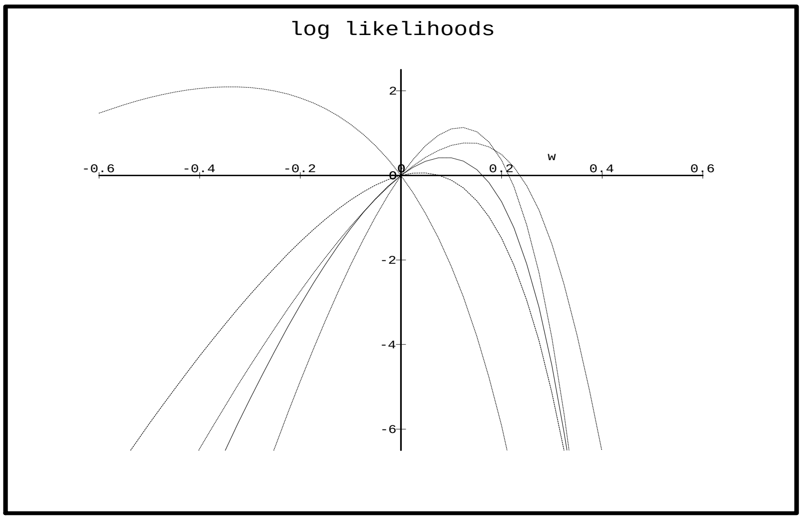
3 graphs



Combined Likelihoods



Likelihoods



Log likelihoods

Acknowledgements

I thank M.E. Cates for helpful conversations, and D. Graham, G. Mitchison, T.T. Jervis and M.J. Oldfield for sharing their intuitions.

References

[1] J. D. Mollon and J. K. Bowmaker. The spatial arrangement of cones in the primate fovea. *Nature*, 360:677–679, 1992.