Astronomy A345H

Astronomical Data Analysis I: Example Sheet 3

1. Let x have the pdf given by:-

$$p(x) = x^2/9, \quad 0 < x < 3, \quad \text{zero elsewhere}$$

   Find the pdf of $y = x^3$

2. Let x have the pdf given by:-

$$p(x) = 2xe^{-x^2}, \quad 0 < x < \infty, \quad \text{zero elsewhere}$$

   Find the pdf of $y = x^2$

3. Let x have a uniform pdf over the interval $(-\pi/2, \pi/2)$.

   (a) Show that $y = \tan x$ has pdf (known as the *Cauchy* distribution) given by:-

$$p(y) = \frac{1}{\pi(1+y^2)}, \quad -\infty < y < \infty$$

   (b) Determine the mean and variance of y

4. The distribution of gamma ray bursts has been shown to be uniform on the sky – i.e. the probability of a gamma ray burst occurring in solid angle $d\Omega$ is simply proportional to $d\Omega$, where the constant of proportionality is independent of direction on the sky.

   (a) Given that $d\Omega = \cos\beta d\beta d\alpha$, where α and β are galactic longitude and latitude respectively, determine the marginal pdfs in galactic longitude and galactic latitude of the gamma ray burst distribution.

   (b) There exists a function $y = f(\beta)$, of galactic latitude such that the pdf of y is a uniform distribution between -1 and $+1$. What is the function $f(\beta)$?

5. The number r of cosmic ray events detected per hour by a ground-based detector is modelled as a Poisson variable with pdf

$$p(r|\mu) = \frac{\mu^r \, e^{-\mu}}{r!}$$

   where μ is the average number of cosmic ray events per hour.

   (a) By differentiating the natural logarithm, ℓ, of the likelihood function, show that – if r events are detected in a given hour of operation – the maximum likelihood estimate of μ based on these data is simply $\hat{\mu}_{\mathrm{ML}} = r$.

   (b) The detector operates continuously for n hours, with the number of cosmic ray events detected in each hour denoted by $r_i\,;\ i = 1, ..., n$. Show that the maximum likelihood estimate of μ, based on the entire n hour run, is

$$\hat{\mu}_{\mathrm{ML}} = \frac{1}{n}\sum_{i=1}^{n} r_i$$

   stating clearly any assumptions that you make.

6. A coin is tossed n times and a binomial model (see Examples Sheet 2, Q.4) is adopted to describe the probability of obtaining r heads, i.e. the data are described by the likelihood:

$$p(r|\theta) \propto \theta^r (1-\theta)^{n-r}; \qquad 0 < \theta < 1$$

where θ is the probability of obtaining a head on any given toss of the coin and the constant of proportionality does not depend on θ.

(a) Write down an expression for the natural logarithm, $\ell(\theta)$, of the likelihood.

(b) By differentiating $\ell(\theta)$, show that the **maximum likelihood** estimate of the parameter θ is $\hat{\theta}_{\mathrm{ML}} = \frac{r}{n}$.

(c) A sequence of coin tosses is analysed within a Bayesian framework to make inferences about the value of θ, using Bayes' formula in the form

$$p(\theta|r) \propto p(r|\theta)p(\theta)$$

where $p(\theta)$ is a distribution describing our prior assumptions about the value of θ. Explain why, if we adopt a *uniform* prior for θ over the range $0 < \theta < 1$, then the maximum of the posterior probability distribution function for θ is again equal to r/n.

(d) * If the coin is 'fair' one should expect that $\theta = 0.5$. Suppose we have a strong prior belief that our coin is fair, and we adopt a prior of the form:

$$p(\theta) \propto \left[1 - 4(\theta - 0.5)^2\right]$$

By writing down the natural logarithm of the posterior, and differentiating with respect to θ, show that with the above prior the maximum posterior probability occurs at a value of θ that is a solution of the following equation

$$r(1-\theta)\left[1 - 4(\theta - 0.5)^2\right] - (n-r)\theta\left[1 - 4(\theta - 0.5)^2\right] - 8(\theta - 0.5)\theta(1-\theta) = 0.$$

(e) * Assuming $r = 1$ and $n = 4$, make a plot (e.g. with excel) showing how the above equation changes as a function of θ. Show that a zero occurs at $\theta \sim 0.33$.

(f) * Make the same plot but now for $r = 248$, $n = 1000$, and show that the zero now occurs at $\theta \sim 0.25$ – i.e. in agreement with the maximum likelihood estimate from part (b).

(g) * Can you explain why the maximum of the posterior agrees with the maximum likelihood estimate in part (f), but not in part (e)?

Note that the parts marked with a star are somewhat above the level expected for ADA1, but are appropriate to the syllabus and level of the MSci course on statistical astronomy.