Astronomy A3/A4H

Statistical Astronomy I: Example Sheet 3

(In questions involving Hypothesis tests, if the significance level is not explicitly stated then you should choose the level yourself)

1. The fraction, X, of the surface of a star covered in starspots is modelled as a random variable with pdf (with k constant)

$$p(x) = \frac{k}{\sqrt{x(1-x)}}, \qquad 0 < x < 1$$

   (a) Determine k so that $p(x)$ is properly normalised

   (b) Find the expected fraction of the surface covered in starspots

   (c) What is the probability that the fraction covered is less than 25%?

2. The rotation periods, in hours, of six G-dwarf stars are found to be:-

   8.1, 8.7, 9.2, 7.8, 8.4, 9.4

   Compute the sample mean and show that an unbiased estimate of the population variance (i.e. the variance in the measured period of a single star drawn from the underlying population) is 0.338. Hence compute the error of the sample mean.

   The distribution of rotation periods for the population stars of all spectral types is modelled as a normal pdf with mean, $\mu = 9.5$ h. Test the hypothesis that the distribution of periods is independent of spectral type.

3. The apparent magnitudes of a sample of 10 stars observed in an open cluster are:-

   11.2, 7.5, 10.1, 4.8, 6.3, 9.9, 4.9, 8.2, 6.5, 7.8

   The distance, r, to the open cluster is estimated to be 1.2 kpc. The stellar luminosity function is modelled as a normal pdf with mean $M_0 = -4.0$

   By constructing equivalent hypotheses involving apparent magnitudes, use the student's t test to test the null hypothesis, $\mathrm{H}_1 : r = 1.2$ kpc, against the alternative hypothesis, $\mathrm{H}_2 : r > 1.2$ kpc, adopting a 5% critical region. Can you think of reasons why the observed distribution of apparent magnitudes might *not* be normally distributed.

4. A group of 10 galaxies in an Abell cluster is suspected to form a distinct subgroup in the foreground of the cluster. The redshifts of the galaxies are (in $\mathrm{kms}^{-}1$):-

   11567, 8608, 11291, 6732, 7900, 5680, 6574, 9471, 7889, 8841

Assuming the mean redshift of the cluster as a whole to be 9600 kms^{-1}, and the redshift distribution of the group of 10 galaxies to be Gaussian, show that this group has a significantly smaller mean redshift than the cluster as a whole at the 5% significance level. Suggest why it would be dangerous to conclude immediately that the 10 galaxies form a subgroup in the *foreground* of the cluster.

5. Let p denote the probability that a particular outcome will happen in any single experiment (called the probability of a *success*). The probability, $p(r)$, of exactly r successes in n experiments is given by the *binomial* distribution:-

$$p(r) = \frac{n!}{r!(n-r)!}p^r(1-p)^{n-r}, \qquad r = 0, 1, ..., n$$

Recall that $p(r)$ has mean $\mu = np$ and variance $\sigma^2 = np(1-p)$.

Suppose it is given that 60% of the stars in the Hubble Space Telescope guide star catalogue are binaries, use a binomial distribution model to calculate the probability that a random sample of 5 stars from the guide star catalogue contains a) 0, b) 1, c) 2, d) 3, e) 4, f) 5 binary stars.

How large a sample should be chosen in order that the probability of the sample containing at least two *non*–binary stars is greater than 99%?

6. In a meteor search program 4 photographic plates were exposed on each observing night and examined for meteor trails. Over a one year period, 150 nights of data were accumulated with the following results.

No of plates with trails	0	1	2	3	4
No of nights	30	62	46	10	2

The number, r, of plates recording meteor trails on any given night is assumed to follow a binomial distribution. By equating the sample mean value of r for the above observations with the expected value for a binomial distribution, estimate the parameter, p, the probability of a single plate recording a meteor trail. Hence determine the *predicted* number of nights on which r plates record trails under the binomial model ($r = 1, ..., 5$).

7. Repeat Q.6 but this time fit the data to a Poisson distribution:-

$$p(r) = \frac{\lambda^r e^{-\lambda}}{r!}$$

recalling that $E(r) = \lambda$ for a Poisson variable, r. Hence determine the predicted number of nights on which r trails are observed under the new model. Which model – Poisson or binomial – better fits the data? Why might a Poisson model not be appropriate?

Dr. Martin Hendry
Room 607, Kelvin Building
Ext 5685; email martin@astro.gla.ac.uk