AA12M Statistical Astronomy (STA)
problem sheet #2

This problem sheet covers Bayesian and frequentist methods for parameter estimation and hypothesis testing. Bayesian questions are tagged with a little 'B', frequentist questions with an 'F'. Full solutions for all these problems will appear as the course progresses. Answers to some problems are shown in curly brackets when appropriate.

1)B Explain how Bayes' Theorem is used for parameter estimation in Bayesian Probability Theory.

A telescope is constructed to look for gamma ray bursters (GRBs). Given that GRBs can appear from any direction (i.e., are distributed isotropically on the sky) what is the prior probability distribution that a GRB will be seen at a particular declination δ (hint: consider the fraction of the sky at this declination)? If $\mu = \sin\delta$, show that the prior for μ is uniform for $-1 < \mu < 1$.

2)B Briefly, what is *marginalisation* in the context of Bayesian inference?

In a photon counting experiment we are told that, over a time interval T, exactly two photons have struck our detector. Given no other information, write down and justify the joint probability distribution function (pdf) for the two arrival times t_1 and t_2.

By sketching this pdf, or otherwise, show that the probability that they arrived within a time τ of each other is

$$P(\tau) = 1 - (1 - \tau/T)^2.$$

Given the extra information that they did indeed arrive within a time τ of each other, sketch the pdfs for the time:

(a) of the first arrival

(b) of the second arrival

(c) that a photon arrives.

3)B Use your favorite computing resource (Python, Matlab, Wolfram Alpha, ...) to determine the fraction of the probability contained within a $(\pm)1$, 2 and 3 σ zone around the mean of a Gaussian posterior probability distribution.

4)B From what you know, write down sensible prior probabilities/pdfs for the following quantities:

- The number of sweets in a jar.
- The mass of Neptune.
- The radius of Neptune.

5)F The Central Limit Theorem states that the sum of N random variables drawn from almost any distribution will itself be a random variable which, if N is sufficiently large, will be Normally distributed. Extend this idea to the *product* of N random variables(all > 0), and determine the distribution of this product (called the Lognormal distribution).

Why might the masses of bodies in the rings of Saturn have a Lognormal distribution?

6)B Outline the reasoning behind 'least-squares' parameter estimation within a Bayesian framework.

For a set of data, $\{Y_k\}$, with associated error bars $\{\sigma_k\}$, taken at known 'positions' $\{x_k\}$ derive the best slope (m_0) and intercept (c_0) for a straight line fit.

7)B N observations of the flux density of a quasar, $\{x_k\}$, are affected by interstellar scintillation which introduces Gaussian errors of (unknown) variance σ^2. Explain what is meant by the *likelihood* of these data, and show that, if the measurements are independent, the likelihood is

$$p(\{x_k\}|\mu,\sigma,I) = (\sigma\sqrt{2\pi})^{-N} \exp\left[-\frac{1}{2\sigma^2}\sum_{k=1}^{N}(x_k-\mu)^2\right],$$

where μ is the true flux density of the quasar.

Explain the importance of the joint posterior pdf of μ and σ for parameter estimation. What is the meaning of the *marginal* posterior pdf for μ alone? Show that, if the priors for μ and σ are uniform for values > 0 and zero otherwise, the marginal posterior pdf for μ is

$$p(\mu|\{x_k\},I) \propto \int_0^\infty t^{N-2} \exp\left[-\frac{t^2}{2}\sum_{k=1}^{N}(x_k-\mu)^2\right] \mathrm{d}t,$$

where $t = 1/\sigma$. Evaluate the un-normalised value of this,* given the standard result

$$\int_0^\infty x^n \exp(-ax^2)\,\mathrm{d}x \propto a^{-(n+1)/2}.$$

By examining the maximum of $L = \ln\left[p(\mu|\{x_k\},I)\right]$, show that the best estimate for μ is

$$\mu_0 = \frac{1}{N}\sum_{k=1}^{N} x_k,$$

and that the uncertainly in this is $S/\sqrt{N}$ where

$$S^2 = \frac{1}{N-1}\sum_{k=1}^{N}(x_k-\mu_0)^2.$$

Comment on how this result compares to the situation where σ is known, as derived in the notes.

8)B A spacecraft is sent to a moon of Saturn, and, using a penetrating probe, detects a liquid sea under the surface at 1 atmosphere pressure and a temperature of $-3\,^\circ$C. However the thermometer has a random fault, so that the temperature reading may differ from the true temperature by as much as $\pm 5\,^\circ$C with uniform probability within that range.

(a) Assuming the liquid is water (background information I_1), which is a liquid for $0\,^\circ\mathrm{C} < T < 100\,^\circ\mathrm{C}$,

(i) write down, and draw, a sensible prior for the temperature of the liquid, $p(T|I_1)$.
(ii) write down, the likelihood of the data, given the instrument's troublesome performance and sketch its variation with T.
(iii) compute, and draw, the normalised posterior pdf for the temperature.

*The answer you get is basically Student's t distribution, derived from Bayesian principles.

(b) Now repeat the above analysis assuming the liquid is ethanol (background information I_2) which is liquid at one atmosphere between $-80\,^\circ$C and $80\,^\circ$C, and comment on the difference in the results.

(c) The calibration error is found to be such that subsequent readings have independent errors within the range $\pm 5\,^\circ$C.

(i) By applying the central limit theorem determine, and sketch, the likelihood of the average of 100 such readings.

(ii) Given that this average reading is $-1.2\,^\circ$C again calculate, and sketch, the posterior pdfs under the hypotheses that the liquid is water or ethanol.

(d) (advanced) determine the relative odds that the liquid is water or ethanol following the single measurement and the average measurement.

9)B An important topic in X-ray astronomy is the determination of the X-ray background rate, b (i.e., the rate of arrival of X-rays from the background sky).

An X-ray telescope observes a 'blank' area of sky and counts n X-ray photons in a time T. The likelihood of this observation follows the Poisson distribution,

$$p(n|b,I) = \frac{(bT)^n \mathrm{e}^{-bT}}{n!}.$$

Taking b to be a scale parameter, assign it a prior, $p(b|I)$, and determine the normalised posterior for b. You will need to use

$$\int_0^\infty x^m \mathrm{e}^{-ax}\,\mathrm{d}x = \frac{m!}{a^{m+1}} \qquad (a > 0;\ m = 0, 1, 2 \ldots).$$

Show that the mean of this posterior is n/T, and that its standard deviation is the mean divided by $\sqrt{n}$.

Repeat this analysis using a uniform prior for b. Do the two results differ substantially?

10)F 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. $\{1/\pi\}$

(b) Find the expected fraction of the surface covered in starspots. $\{1/2\}$

(c) What is the probability that the fraction covered is less than 25%? $\{1/3\}$

11)B The redshift of a quasar is measured to be z_1 with a standard deviation of σ.

(a) Assuming the uncertainty in this data is Gaussian, explain why a uniform prior probability distribution for z implies a normalised posterior probability distribution for z of

$$p(z|z_1) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left[-\frac{(z-z_1)^2}{2\sigma^2}\right]$$

(b) A second measurement, z_2, also of standard deviation σ, is made with the same setup. Write down an expression for the likelihood of z based solely in this second measurement, $p(z_2|z)$, and hence show that the new posterior probability distribution is

$$p(z|z_1, z_2) \propto \frac{1}{2\pi\sigma^2} \exp\left[-\frac{(z-z_1)^2 + (z-z_2)^2}{2\sigma^2}\right]$$

(c) Show that this can be written in the form

$$p(z|z_1, z_2) \propto \frac{1}{2\pi\sigma^2} \exp\left[-\frac{A(z - z_{\rm m})^2 + c}{2\sigma^2}\right]$$

and evaluate A, $z_{\rm m}$, and c. Hence show that $p(z|z_1, z_2)$ is also Gaussian with a mean of $(z_1 + z_2)/2$ and a standard deviation of $\sigma/\sqrt{2}$, commenting on these results.

(d) Usually, one might expect that the two measurements z_1 and z_2 would be within about σ of each other, but on this occasion $z_1 - z_2 = 10\sigma$. Sketch graphs of $p(z|z_1)$, $p(z|z_2)$ and $p(z|z_1, z_2)$ for this situation, and give an interpretation of what may be going on here. Why is the uncertainty in the final posterior so (relatively) small?

(e) Define the evidence (or 'marginal likelihood') for this dataset, and briefly describe the role it plays in Bayesian parameter estimation. Show that the evidence in this problem is proportional to $\exp\left[-c/(2\sigma^2)\right]$, and comment on how its value depends on the magnitude of $z_1 - z_2$.

12)F 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:

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

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

Suppose 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%? {14}

13)F In a meteor search program, four 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. The minimum number is 0, the maximum is 4. 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. {0.32}

Hence determine the *predicted* number of nights on which r plates record trails under the binomial model ($r = 0, \ldots, 4$).

14) The distribution of (natural) log distance of galaxies in a survey is found to be normal with mean μ and variance σ^2. Derive the pdf, $p(r)$, of the galaxy distance, r, and determine the expected value and variance of r.

15) Let X and Y be random variables. Find expressions for the following in terms of the variance and covariance of X and Y.

(a) $\text{var}(aX)$ and $\text{var}(aY)$, where a is a constant.

(b) $\text{cov}(aX, aY)$, where a is a constant.

(c) $\text{cov}(X, X + Y)$

(d) $\text{cov}(X + Y, X - Y)$

Show that $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2\,\text{cov}(X, Y)$.

What property of the correlation coefficient, ρ, is indicated by your solution to (a) and (b)?

16)F The angular diameter, θ, of the expanding photosphere of a type II supernova is observed simultaneously with two different telescopes, A and B, at a number of epochs, with the following results (in arcsec):

A	101.8	102.8	111.0	113.5	114.4	114.8	114.5	116.2	120.2	123.5
B	99.2	103.1	114.8	111.6	110.1	110.3	110.7	114.3	117.6	119.2

It is suspected that, due to incorrect flat-fielding, the results from telescope B differ systematically from those of telescope A. The following model is constructed:

$$\theta(A)_i = \alpha + \beta\,\theta(B)_i + e_i \tag{1}$$

where α and β are constants and the errors, e_i, are normally distributed with mean zero and dispersion, $\sigma = 1.7$.

Determine least-squares estimates for α and β with this model. Use the χ^2 statistic to test the goodness of fit of the model to the data at the 1% significance level.

17)B Observations made over several decades have resulted in a value of Hubble's constant H with a (Bayesian) probability

$$p(H|I) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_{\rm p}}\exp\left[-\frac{(H - H_{\rm p})^2}{2\sigma_{\rm p}^2}\right],$$

where $H_{\rm p}$ and $\sigma_{\rm p}$ are constants and I represents the relevant background information. A new technique measures Hubble's constant using an entirely different method, reporting a value $H_{\rm d}$ with uncertainty $\sigma_{\rm d}$. Taking the probability of this new data to be Gaussian, show that the updated value of H is described by

$$p(H|H_{\rm d}, I) = \frac{1}{\sqrt{2\pi}}\frac{1}{\Sigma}\exp\left[-\frac{(H - m)^2}{2\Sigma^2}\right],$$

where

$$\frac{1}{\Sigma^2} = \frac{1}{\sigma_{\rm p}^2} + \frac{1}{\sigma_{\rm d}^2}$$

and

$$m = \Sigma^2\left(\frac{H_{\rm d}}{\sigma_{\rm d}^2} + \frac{H_{\rm p}}{\sigma_{\rm p}^2}\right).$$

Comment on this result when $\sigma_{\rm d} \ll \sigma_{\rm p}$.

18)B A model exists for a new type of X-ray source. If these sources are distributed homogeneously in the Universe (i.e., with constant mean density) show that the Bayesian probability density function (pdf) for the radial distance, r, to a source is

$$p(r) \propto r^2.$$

The sources are thought to radiate isotropically so that the flux F from a source at a distance r is

$$F = \frac{L}{4\pi r^2},$$

where L is the intrinsic luminosity of the source. Given that the pdf of L for any particular source is $p(L)$ show (by considering the joint pdf of r and L) that the joint pdf of L and F for an *observed* source is

$$p_{\mathrm{obs}}(L, F) \propto p(L) L^{3/2} F^{-5/2}.$$

You may use the general result that

$$p(y, z) = p(u, v) \left| \begin{array}{cc} \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{array} \right|.$$

To be observed, the flux from a source must be greater than F_{min}. Assuming the pdf for source luminosity is exponential, i.e.,

$$p(L) \propto \exp(-L/L_0),$$

and given $\int_0^\infty x^{3/2} \mathrm{e}^{-x}\, \mathrm{d}x = 3\sqrt{\pi}/4$, determine the normalised pdf for $p_{\mathrm{obs}}(L, F)$.

Show that $p(L) \neq p_{\mathrm{obs}}(L)$ and comment on why this is so, physically. What is the most probable observed luminosity?

Poisson statistics tells us that the probability of receiving n photons in a time τ s from a source of mean flux F is

$$P(n|F, \tau) = \frac{\mathrm{e}^{-F\tau}(F\tau)^n}{n!},$$

where F has units of photons per second. During an observation lasting one second, 6 photons are seen from one of the new sources mentioned in part (b). Determine the un-normalised posterior pdf for the flux of this source, independent of L.

What is the most probable flux of the source? Comment on why it is significantly less than 6 photons per second?

19)B The probability of detecting N galaxies in a patch of sky of solid angle Ω can be modelled as a Poisson distribution,

$$P(N|\lambda) = \frac{\lambda^N \mathrm{e}^{-\lambda}}{N!},$$

where λ is a constant, proportional to Ω (i.e. $\lambda = k\Omega$). What important assumptions have we made in this model?

Show that the mean number of galaxies in the patch equals λ, and hence that the (constant) probability of finding a galaxy in differential solid angle $\mathrm{d}\Omega$ is $k\,\mathrm{d}\Omega$.

Show that if λ is an integer, it is equally probable that there are λ and $\lambda - 1$ galaxies in the patch, and discuss briefly why only one of these equally probable outcomes is the 'expectation value' for N.

Taking one of the galaxies as the centre point, determine the probability that there are no galaxies around it in a disc of solid angle ω and hence determine the probability distribution for the angular distance α to its nearest neighbour. You may assume the angles are small, so that the solid angle of a disc of radius α is $\omega = \pi\alpha^2$.

20) B We can determine the distance r to a source of *known* luminosity L from its flux F at the Earth using the inverse-square law,

$$F = \frac{L}{4\pi r^2}.$$

An astronomical object of known luminosity is called a 'standard candle'.

An instrument makes a measurement D of the flux from a standard candle, subject to Gaussian errors of variance σ^2. The (Bayesian) probability distribution of this measurement is

$$p(D|F) \propto \exp\left[-\frac{(D-F)^2}{2\sigma^2}\right].$$

Use Bayes' Theorem to determine the posterior probability distribution for the distance to the source, assuming that such sources are distributed in space with a uniform probability.

Discuss why this posterior is insensitive to the choice of prior when the signal-to-noise ratio $\gamma\ (= D/\sigma)$ is sufficiently large and show that under these circumstances the posterior is approximately

$$p(r|D) \propto r^2 \exp\left[-\frac{2D^2(r-r_0)^2}{\sigma^2 r_0^2}\right],$$

where r_0 satisfies $4\pi D r_0^2 = L$. Estimate the '1-sigma' uncertainty in the estimate of r as a function of γ.

Hubble's constant H relates the velocity v of a source (due to the expansion of the Universe) to its distance:

$$v = Hr.$$

If v is known *exactly*, determine the pdf for H using the above posterior for r.

Generally, the value of v will have some uncertainty, σ_v, so that

$$p(v) \propto \exp\left[-\frac{(v-v_0)^2}{2\sigma_v^2}\right].$$

Starting with the joint probability of H and v, estimate (using suitable approximations for the integrals of sharply-peaked functions) the new standard width of $p(H|D)$ when γ and v/σ_v are large.

You may use the result

$$\int_{-\infty}^{\infty} \exp\left[-\frac{(x-x_1)^2}{2\sigma_1^2} - \frac{(x-x_2)^2}{2\sigma_2^2}\right] \mathrm{d}x = \left(\frac{2\pi\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)^{1/2} \exp\left[-\frac{(x_1-x_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right].$$