Final Examination Problems

1. The probability density of the spin parameter λ of a dark matter halo is well known to be described by the following log-normal distribution characterized by two unknown parameters, λ_0 and σ_λ:

$$p(\lambda) = \frac{1}{\lambda\sqrt{2\pi}\sigma_\lambda}\exp\left[-\frac{(\ln\lambda - \ln\lambda_0)^2}{2\sigma_\lambda^2}\right].$$

   (a) The attached file, "halospin.txt", contains a sample of 2000 dark matter halos resolved in a N-body simulation. The values given in the second and third columns correspond to the dimensionless spin parameter (λ) and the virial mass (M in unit of $h^{-1}M_\odot$) of each halo, respectively. Find a median value, $M_{\rm med}$, of the virial masses of the halos in the sample and divide the given sample into two subsamples, (say A and B), which include the high-mass halos with masses $M \geq M_{\rm med}$ and the low-mass halos with $M < M_{\rm med}$, respectively.

   (b) Determine the MLE best-fit values of the two parameters, λ_0 and σ_λ, for each subsample.

   (c) Determine the Fisher information matrix about the two parameters for each subsample, and estimate the errors in $\hat{\lambda}_0$ and $\hat{\sigma}_\lambda$.

   (d) Make a decision between the following two hypotheses and evaluate $P_{\rm FA}$ and $P_{\rm D}$.

   H_0 : The spin parameter distribution does not depend on the halo mass M.
   H_1 : The spin parameter distribution actually depend on the halo mass M.

2. According to the standard ΛCDM cosmology, the luminosity (comoving) distance, d_L, of a type Ia supernova at redshift z is given as

$$d_L(z) = \frac{c(1+z)}{H_0}\int_0^z \frac{dz'}{\sqrt{(1+z')^4\Omega_r + (1+z')^3\Omega_m + (1+z')^2\Omega_k + \Omega_\Lambda}} \tag{1}$$

   where H_0 is the Hubble constant, $\{\Omega_\Lambda, \Omega_k, \Omega_m, \Omega_r\}$ represent the cosmological constant, curvature, matter and radiation density parameters, respectively, at the present epoch, which satisfy the condition of $\Omega_\Lambda + \Omega_k + \Omega_m + \Omega_r = 1$, regardless of z. The apparent magnitude of a supernova at B-band, m_B, is related to d_L as

$$m_B - M_B = 5\log\left(\frac{d_L}{\rm Mpc}\right) + 25 \tag{2}$$

where M_B is the B-band absolute magnitude of the supernova. Using the data given in Tables 1-2 of the paper, Perlmutter et al. (1999, ApJ, 517, 565), setting the universal absolute B-band magnitude of all supernovae at $M_B = -18$, setting the Hubble constant at $H_0 = 70\text{km}\,s^{-1}\,\text{Mpc}^{-1}$, and regarding $\{\Omega_\Lambda, \Omega_k, \Omega_m, \Omega_r\}$ as unknown cosmological parameters constrained by $\Omega_\Lambda + \Omega_k + \Omega_m + \Omega_r = 1$, perform the following statistical analysis (*Hint:* There are only three *independent* parameters):

(a) Assume $\{m_{Bi}\}_{i=1}^{N} \sim \mathcal{N}(\mu_i, \sigma_i^2)$ where the population mean μ_i is the theoretical prediction obtained by combining Equation (1) with Equation (2), while one standard deviation σ_i represents the error in the observed m_{Bi} of each supernova (see the 9th and 10th columns of Tables 1-2 in the aforementioned paper). Note that $\mu_i = \mu_i(\Omega_\Lambda, \Omega_m, \Omega_k, \Omega_r)$. Applying the Fisher information matrix analysis to this dataset, estimate the MLE best-fit values of the cosmological parameters and associated errors as well.

(b) Given that the redshift of each supernova suffers from some degree of uncertainty (see the 3rd columns of Tables 1-2), the signals (i.e, the theoretical prediction of $\{\mu_i\}_{i=1}^{N}$ themselves) can also be regarded as Gaussian random variables. Applying the Wiener filter technique based on the Bayesian statistics, reestimate the four parameters and associated errors as well.

(c) Applying the soft threshold (ST) filer based on the Laplacian priors to this dataset, estimate the best-fit values of the four parameters with their errors. Among the four, which parameter's best-fit values are found to be zero? What does it imply?

(d) Test the following two hypothesis, and evaluate P_{FA} and P_D. Forecast the value of N (i.e, the size of supernovae sample) which can make $P_D \geq 0.99$ with $P_{FA} = 0.01$. You can choose any estimator among the MLE, MAP, and ST for this test.

$$\begin{aligned} H_0 &: \Omega_\Lambda = 0. \\ H_1 &: \Omega_\Lambda \neq 0. \end{aligned} \tag{3}$$

3. (**Bonus Track**) The other attached file ("2mass.txt") contains the reconstructed *real-space* density contrast field, $\delta(i,j,k)$, from 2MASS all-sky survey that was fully described in the paper, Erdogdu et al. (2004, MNRAS, 352, 939). The $\delta(i,j,k)$-reconstruction was performed on the 64^4 grids with $i,j,k \in \{1,\cdots,64\}$. Each grid has a linear size of $6.25\,h^{-1}$Mpc, while the total volume is $V = 400\,h^{-3}\text{Mpc}^3$. The file stores $\delta(i,j,k)$ in its $[64^2(i-1)+64(j-1)+k]$-th row.Use the mid-point of each grid as its real-space position where δ is defined.

(a) By performing either the DFT or the FFT of $\delta(i,j,k)$, determine and plot the power spectrum, $P_\delta(k)$ as a function of the Fourier wave k. Then, estimate and plot two power spectra, $P_{\delta_s}(k)$ and $P_w(k)$, via the interpolation method. Find the frequency-domain (FD) Wiener filter and plot it.

(b) Using the FD Wiener filter, find the best-fit of $\hat{\delta}_{s,wiener}$. (*Hint:* Find the Wiener filtered version in the Fourier space first and then perform a inverse FFT of it to the real-space). Plot $\hat{\delta}_{s,wiener}$ as contours in the projected 2D space (you can choose either x-y plane or y-z plane)

(c) The reconstructed density field can be regarded as sparse signals since in the very underdense void regions, the values of δ are close to zero. Determine the soft threshold (ST) filter and estimate the ST signal, $\hat{\delta}_{s,ST}$. Plot $\hat{\delta}_{s,wiener}$ as contours in the projected 2D space (you can choose either x-y plane or y-z plane).