Cosmic Calibration: Statistical Modeling for Dark Energy and the Cosmological Constants

Tracy Holsclaw

Applied Math and Statistics, UCSC

October 2009

Cosmology Overview

Data Relation

Parametric Models
- Model 1
- Model 2
- Model 3

Non-parametric Models
- Gaussian Process
- Damped Hermite Polynomials

Real Data

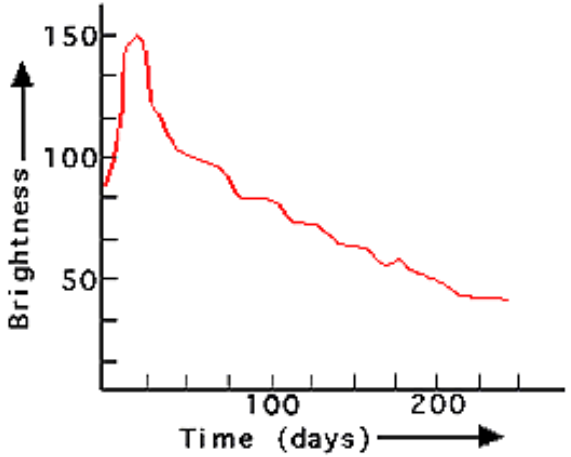
Future Work

Cosmology Overview

- The Universe is expanding (1920's)
- The Universe is expanding at an accelerating pace (1998)
- Ordinary matter would mean the Universe is decelerating
- Dark energy could be the unknown driver of this acceleration
- The hypothesized dark energy cannot be directly measured so observations are needed to infer more about it
  - Type Ia supernovae (type Ia SNe)
  - Cosmic microwave background radiation (CMB) - faint background glow detected by radio telescopes
  - Assumptions of the Universe - homogeneous, isotropic, and spatially flat...
  - Baryon Acoustic Oscillation (BAO) - standard size galaxy clusters and considered a standard ruler

Type Ia Supernovae (SNe)

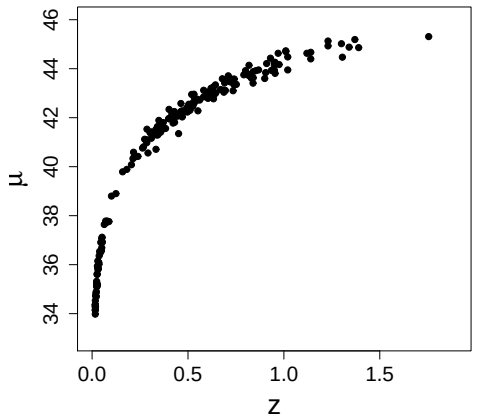
Extremely bright exploding stars that take several weeks or months to burn out; astronomers measure their luminosity over this time and produce a light curve



Type Ia supernova are standardizable candles this means they have almost no variation in absolute magnitude. We can use the peak luminosity as a distance measure (D_L or μ)

Type Ia Supernovae (SNe)

- We can get a distance measure from their light curve
- We can also measure type Ia SNe cosmological redshift which is (the stretching of space between two objects) to see how fast the SNes are receding from us



- If we measure enough type Ia SNes then we can relate their redshift (z, no units) and luminosity distance (Mpc) to learn about dark energy and other cosmological parameters

Non-linear Equation

- The relationship of z and μ is given by the Friedmann-Robertson-Walker model:

$$\mu = 25 + 5log_{10}\left(\frac{c(1+z)}{H_0}\int_0^z H(s)ds\right)$$

$$H(s) = \left(\Omega_m(1+s)^3 + (1-\Omega_m)(1+s)^3 e^{-3\int_0^s \frac{-w(u)}{1+u}du}\right)^{-1/2}$$

- c denotes the speed of light and it is assumed known
- The main function of interest is $w(u)$
- Two other unknown cosmological parameters have to be estimated H_0, (Hubble's constant) and Ω_m (percent contributed by matter)

Poor methods of estimating $w(u)$

$$\mu = 25 + 5log_{10}\left(\frac{c(1+z)}{H_0}\int_0^z H(s)ds\right)$$

$$H(s) = \left(\Omega_m(1+s)^3 + (1-\Omega_m)(1+s)^3 e^{-3\int_0^s \frac{-w(u)}{1+u}du}\right)^{-1/2}$$

- Some have tried fitting the data directly with a parametric or non-parametric form and then taking the derivatives to get $w(u)$
- Some have tried setting different parts of the equation equal to a new function like $y(s) = \int_0^s \frac{w(u)}{1+u}ds$, this still requires a first derivative of the fit to obtain $w(u)$
- We looked into these types of analysis but they do not work well because taking a derivative loses information about $w(u)$

Better methods for estimating $w(u)$

- Choose a parametric form for $w(u)$ and fit the resulting parameters in $w(u)$
- Let $w(u)$ be a piecewise constant function and choose binning methods to approximate $w(u)$
- Allow $w(u)$ to be a non-parametric Gaussian process or basis of functions

Set up the likelihood function and introduce σ^2 and τ^2 to the equation:

$$L \propto \left(\frac{1}{\sigma}\right)^n e^{-\frac{1}{2}\sum\left(\frac{\mu_i - T(z)}{\tau_i \sigma}\right)^2}$$

Gibbs step for σ^2

- Likelihood

$$L \propto \left(\frac{1}{\sigma}\right)^n e^{-\frac{1}{2}\sum\left(\frac{\mu_i - T(z)}{\tau_i \sigma}\right)^2}$$

  In our likelihood we introduce σ^2 and we will need a prior and posterior for it.

- Prior:

$$\pi(\sigma^2) \sim IG(a, b)$$

- Posterior:

$$\sigma^2 | z, H_0, \Omega_m, w_0 \sim IG\left(\frac{n}{2} + a, \frac{1}{2}\sum\left(\frac{\mu_i - T(z)}{\tau_i}\right)^2 + b\right)$$

Gibbs step for H_0

- Transform H_0 as follows: $h = -5log_{10}(H_0)$
- Informative prior: $\pi(H_0) \sim N(c, d^2)$
- Transform this prior to the form using the delta method:

$$\pi(h) \sim N(m, s^2) \sim N\left(-5log_{10}(c), d^2\left(\frac{-5}{c\ln(10)}\right)^2\right)$$

$$h|... \sim N\left(\frac{\frac{1}{\sigma^2}\sum\frac{\mu_i - V_i}{\tau_i} + \frac{m}{s^2}}{\frac{1}{\sigma^2}\sum\frac{1}{\tau^2} + \frac{1}{s^2}}, \frac{1}{\frac{1}{\sigma^2}\sum\frac{1}{\tau_i^2} + \frac{1}{s^2}}\right)$$

where

$$V_i = 25 + 5log_{10}\left(c(1+z_i)\int_0^{z_i} H(s)ds\right)$$

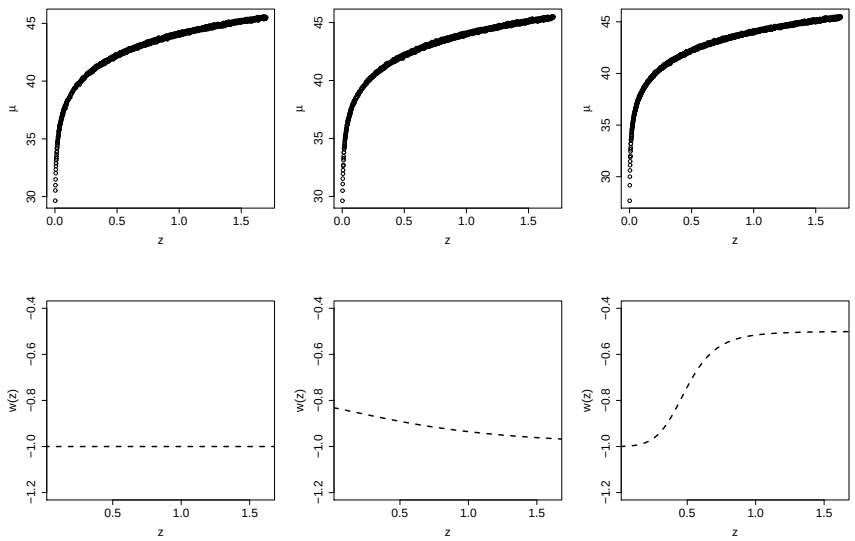
$$H(s) = \left(\Omega_m(1+s)^3 + (1-\Omega_m)(1+s)^3 e^{-3\int_0^s \frac{-w(u)}{1+u}du}\right)^{-1/2}$$

Some models for $w(u)$

- Model 1: fit $w(u) = a$
- Model 2: fit $w(u) = a + bu$
- Model 3: fit $w(u) = a + b(\frac{1}{1+u} - 1)$
- Model 4: Gaussian process fit of $w(u)$
  - A basis of Hermite polynomials that approximates the Gaussian process

Simulated Data

The simulated data has 2000 points with constant $H_0 = 72$, $\Omega_m = 0.27$, $\sigma^2 = 1.0$, and constant τ^2.



- μ_1: $w(u) = -1$
- μ_2: $w(u)$ should be fit well by the parametric forms used
- μ_3: $w(u)$ is an extreme case that will test all the models

Model 1 - Assume $w(z) = a$

$$r(z) = \frac{1}{H_0}\int_0^z \left(\Omega_m(1+s)^3 + (1-\Omega_m)(1+s)^3(1+s)^{3a}\right)^{-1/2}$$

Prior: $\pi(a) \sim Unif(-25, 1)$

Point mass prior for hypothesis testing:
$\pi(a) \sim \lambda U(-25, 1) + (1-\lambda)\delta_{a=-1}$

Table: Model 1 - Hypothesis Test

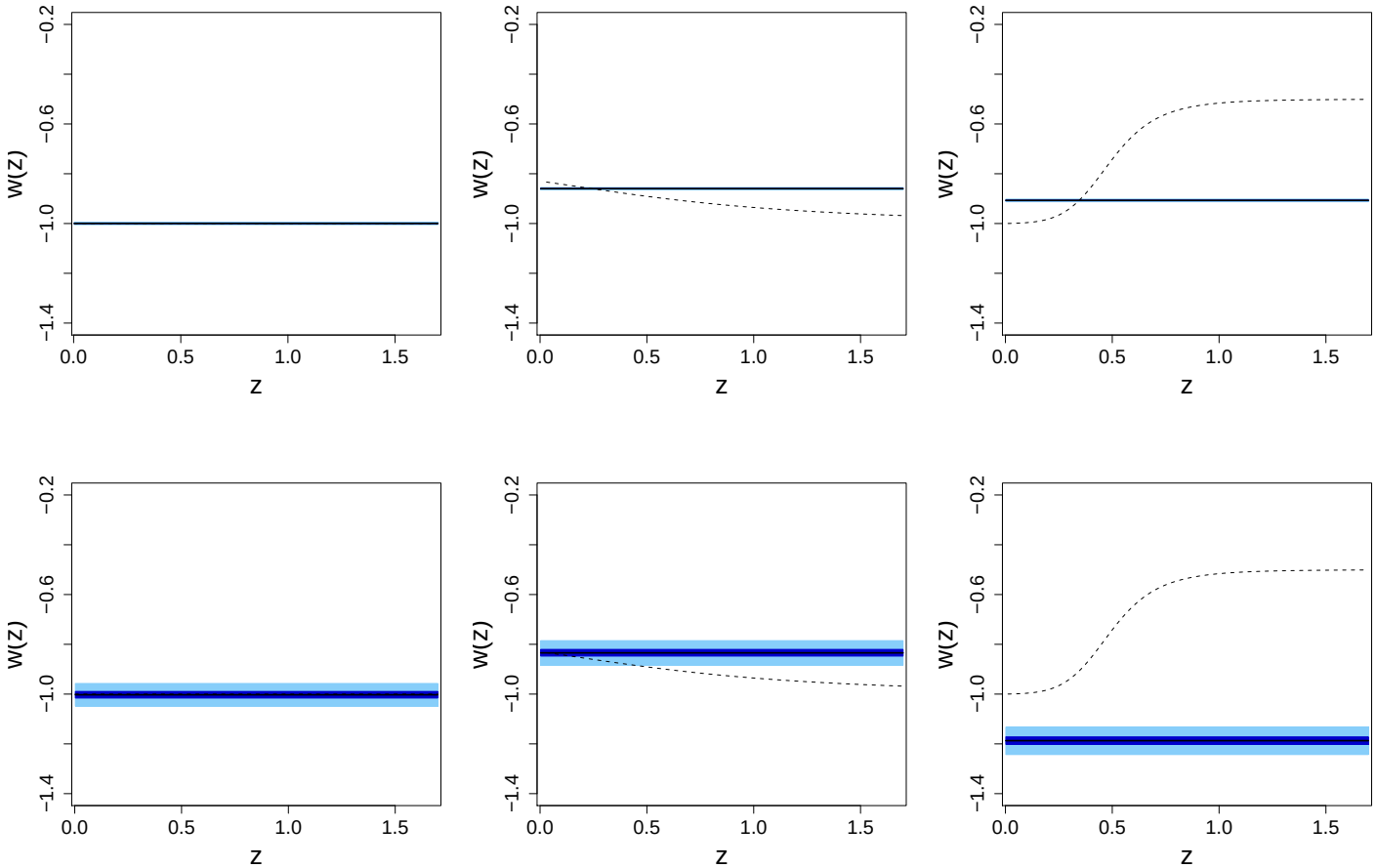
Dataset	$\%\ a = -1$	a
μ_1	0.98	(-1.000,-1.000)
μ_2	0.00	(-0.865,-0.855)
μ_3	0.00	(-0.912,-0.902)

Model 1 - Results

Table: Model 1 - 95% PIs

Dataset	a	Ω_{m0}	H_0
μ_1	(-1.005,-0.994)	0.27	72
μ_2	(-0.865,-0.855)	0.27	72
μ_3	(-0.912,-0.902)	0.27	72
μ_1	(-1.048,-0.957)	(0.262,0.282)	(71.63, 72.27)
μ_2	(-0.884,-0.785)	(0.239,0.271)	(71.77, 72.43)
μ_3	(-1.243,-1.133)	(0.343,0.359)	(71.84, 72.41)

Model 1 - Results



Model 2 - Assume $w(u) = a + bu$

$$r(z) = \frac{1}{H_0}\int_0^z \left(\Omega_m(1+s)^3 + (1-\Omega_m)(1+s)^{3(a-b+1)}e^{3bs}\right)^{-1/2} ds$$

Priors: $\pi(a) \sim Unif(-25, 1)$ and $\pi(b) \sim Unif(-25, 25)$

Table: Model 2 - Hypothesis Test

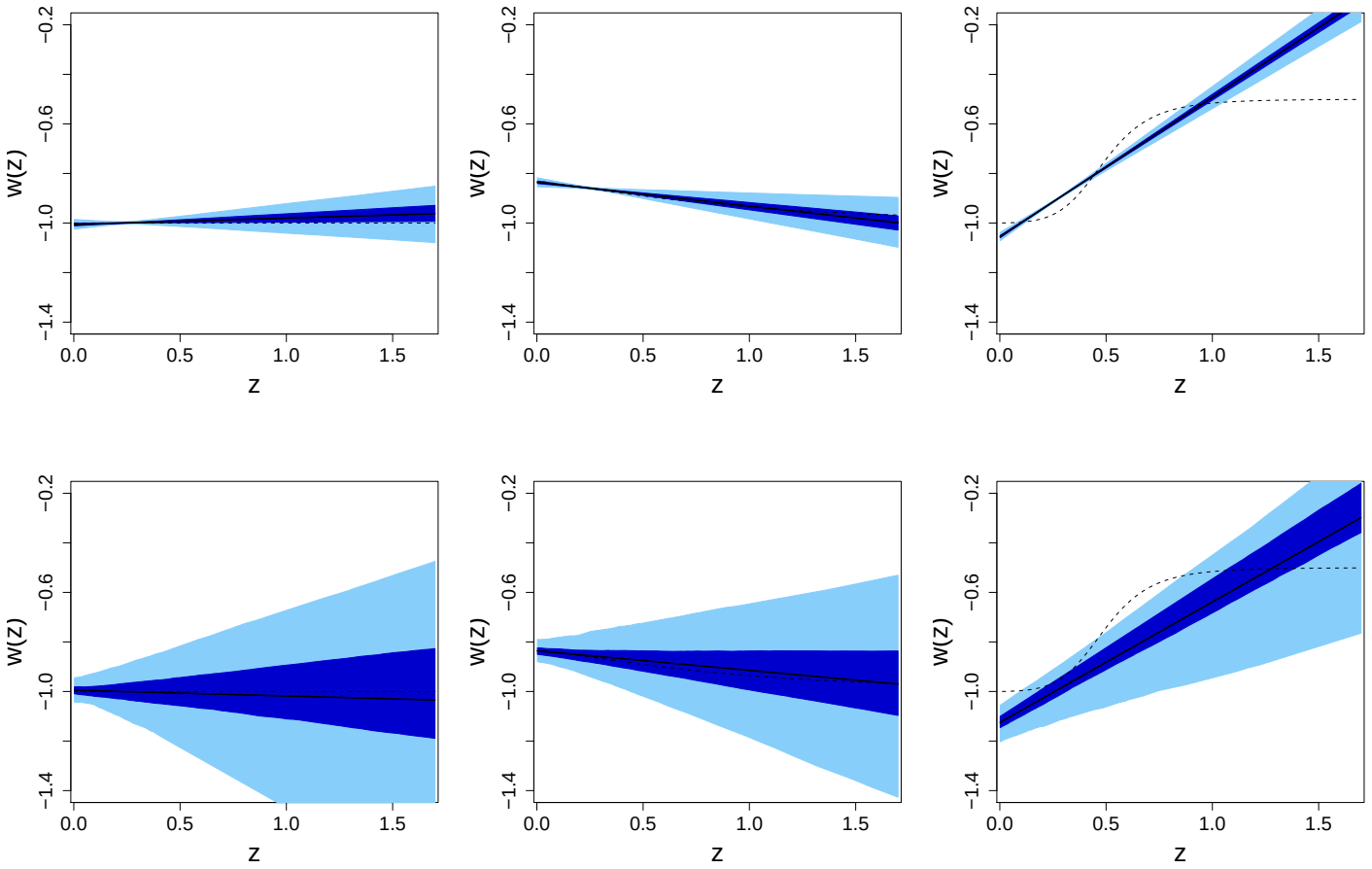
Dataset	% $a = -1, b = 0$	a	b
μ_1	0.99	(-1.000,-1.000)	(0.000,0.000)
μ_2	0.00	(-0.853,-0.817)	(-0.168,-0.031)
μ_3	0.00	(-1.072,-1.038)	(0.500,0.622)

Model 2 - Results

Table: Model 2 - 95% PIs

Dataset	a	b	Ω_{m0}	H_0
μ_1	(-1.025,-0.986)	(-0.055,0.102)	0.27	72
μ_2	(-0.854,-0.817)	(-0.165,-0.026)	0.27	72
μ_3	(-1.072,-1.039)	(0.501,0.622)	0.27	72
μ_1	(-1.042,-0.945)	(-0.470,0.288)	(0.226,0.306)	(71.57, 72.24)
μ_2	(-0.880,-0.790)	(-0.344,0.164)	(0.214,0.301)	(71.69, 72.40)
μ_3	(-1.204,-1.057)	(0.238,0.653)	(0.267,0.335)	(71.92, 72.50)

Model 2 - Results



Model 3 - Assume $w(u) = a + b(\frac{1}{1+u} - 1)$

$$r(z) = \frac{1}{H_0}\int_0^z \left(\Omega_m(1+s)^3 + (1-\Omega_m)(1+s)^{3(a-b+1)}e^{\frac{3bs}{1+s}}\right)^{-1/2} ds$$

Priors: $\pi(a) \sim Unif(-25, 1)$ and $\pi(b) \sim Unif(-25, 25)$

Table: Model 3 - Hypothesis Test

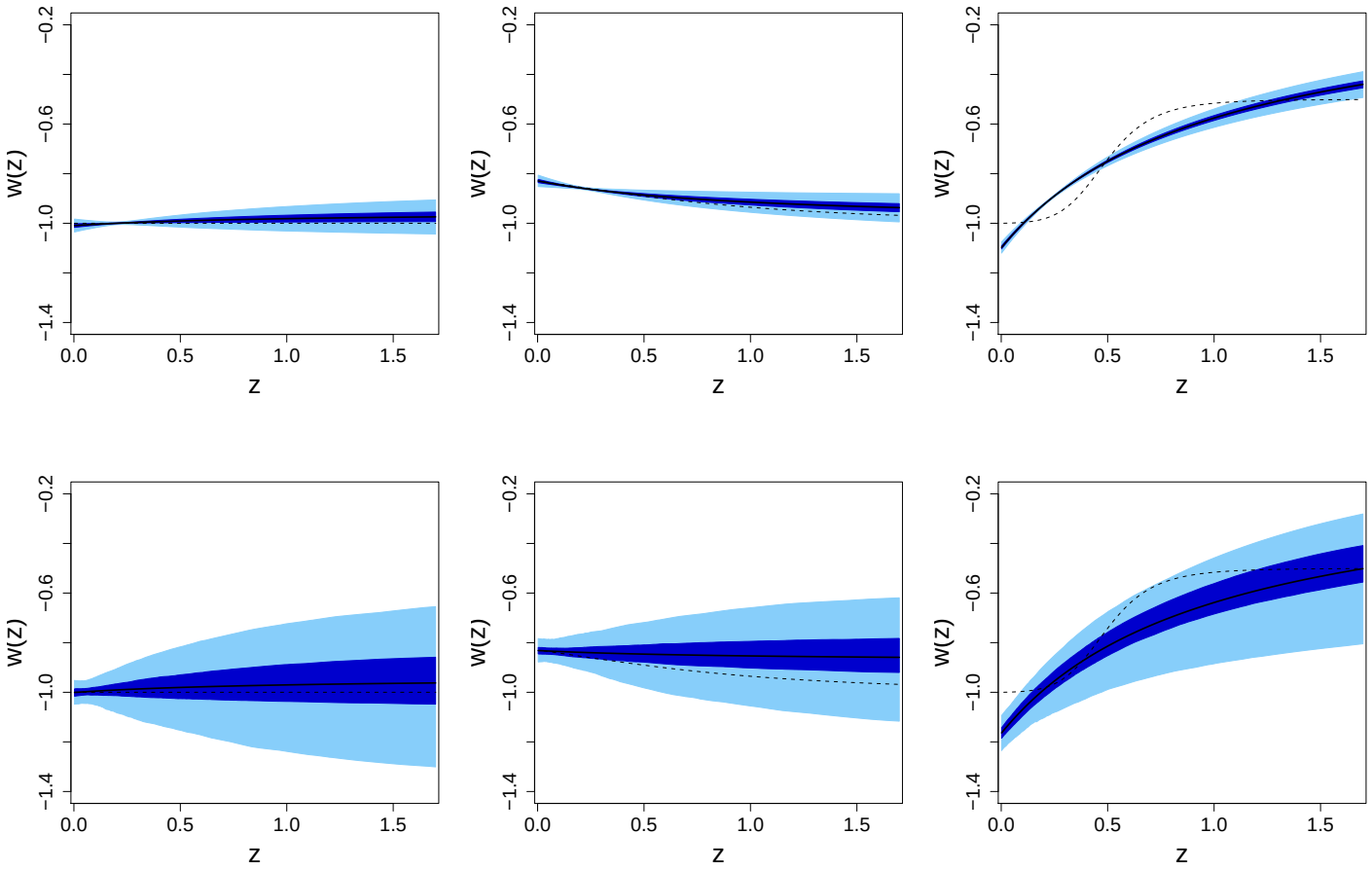
Dataset	% $a = -1$ and $b = 0$	a	b
μ_1	0.98	(1.000,1.000)	(0.000,0.000)
μ_2	0.00	(-0.852,-0.804)	(0.044,0.305)
μ_3	0.00	(-1.121,-1.078)	(-1.159,-0.936)

Model 3 - Results

Table: Model 3 - 95% PIs

Dataset	a	b	Ω_{m0}	H_0
μ_1	(-1.036,-0.983)	(-0.205,0.094)	0.27	72
μ_2	(-0.852,-0.805)	(0.047,0.299)	0.27	72
μ_3	(-1.122,-1.077)	(-1.164,-0.929)	0.27	72
μ_1	(-1.048,-0.951)	(-0.545,0.510)	(0.229,0.297)	(71.61, 72.34)
μ_2	(-0.878,-0.784)	(-0.322,0.473)	(0.215,0.291)	(71.72, 72.45)
μ_3	(-1.237,-1.096)	(-1.368,-0.613)	(0.243,0.324)	(72.05, 72.65)

Model 3 - Results



Gaussian Process on $w(u)$

- A Gaussian process (GP) is a stochastic process such that when sampled at any finite collection of points, the values jointly follow a multivariate Normal distribution
- A Gaussian process needs a mean and correlation function
- $w(u) \sim GP(\theta, \kappa^2 K(u, u'))$ for $u = (u_1, ..., u_m)$

Mean and Correlation Functions

- Mean: $\theta = -1$ and other values were tried
- Correlation:
  - Powered exponential covariance: $K(u, u') = e^{-\lambda|u-u'|^\alpha} = \rho^{|u-u'|^\alpha}$
  - $\alpha = 2$ is the Gaussian correlation but $\alpha = 1.9999$ gives a non-singular invertible matrix
  - Matern with $\nu = 1.5$ was still too smooth
  - $\alpha = 1$ is the exponential correlation
- Priors: $\pi(\rho) \sim Beta(6, 1)$, $\pi(\kappa^2) \sim IG(25, 9)$, $\sigma^2 \sim IG(10, 9)$ and the usual for H_0 and Ω_m

Altered Gaussian Process

- Usual GP:

$$\sigma^2, \rho, \kappa^2 | \mu_i, \tau_i^2, z_i \propto L(z_i, \mu_i, \tau_i | w(u), \sigma^2) GP(w(u) | \rho, \kappa^2) \pi(\rho) \pi(\kappa^2) \pi(\sigma^2)$$

Localized proposals are desirable in this case and provide better acceptance rates

- Altered GP:

$$\sigma^2, \rho, \kappa^2 | ... \propto L(z_i, \mu_i, \tau_i | w^o(u), \rho, \kappa^2, \sigma^2) MVN(w^o(u); 0, I) \pi(\rho) \pi(\kappa^2) \pi(\sigma^2)$$

where $\Sigma^{-1/2}(w(u) - \theta) = w^o(u) \sim MVN(0, I)$

The Not so Simple Algorithm

$$r(z) = \frac{1}{H_0}\int_0^z \left(\Omega_m(1+s)^3 + (1-\Omega_m)(1+s)^3 e^{3\int_0^s \frac{w(u)}{1+u}du}\right)^{-1/2} ds$$

We use the properties of a GP to do the inner integration (following the work of O'Hagan on Bayes-Hermite quadrature): $y(s) = \int_0^s \frac{w(u)}{1+u}du$.

$$w(u) \sim GP(\theta,\ \ \Sigma_{22} = \kappa^2 \rho^{|u-u'|^\alpha})$$

$$y(s) \sim GP\left(\theta \ln(1+s),\quad \Sigma_{11} = \kappa^2 \int_0^s \int_0^{s'} \frac{\rho^{|u-u'|^\alpha}}{(1+u)(1+u')} du du'\right)$$

The Not so Simple Algorithm continued

We draw proposals for $w(u)$ but then use the properties of the GP to turn these into $y(s)$

$$\begin{bmatrix} y(s) \\ w(u) \end{bmatrix} \sim MVN \left[\begin{bmatrix} \theta \ln(1+s) \\ \theta \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right]$$

$$y(s)|w(u) = E(y(s)|w(u)) = \theta \ln(1+s) + \Sigma_{12}\Sigma_{22}^{-1}(w(u) - \theta)$$

During this step we take the opportunity to draw extra $y(s)$ to help smooth it and gain more precise estimates for the outer integral - basically interpolation

The Not so Simple Algorithm, cont

Benefits of this method:

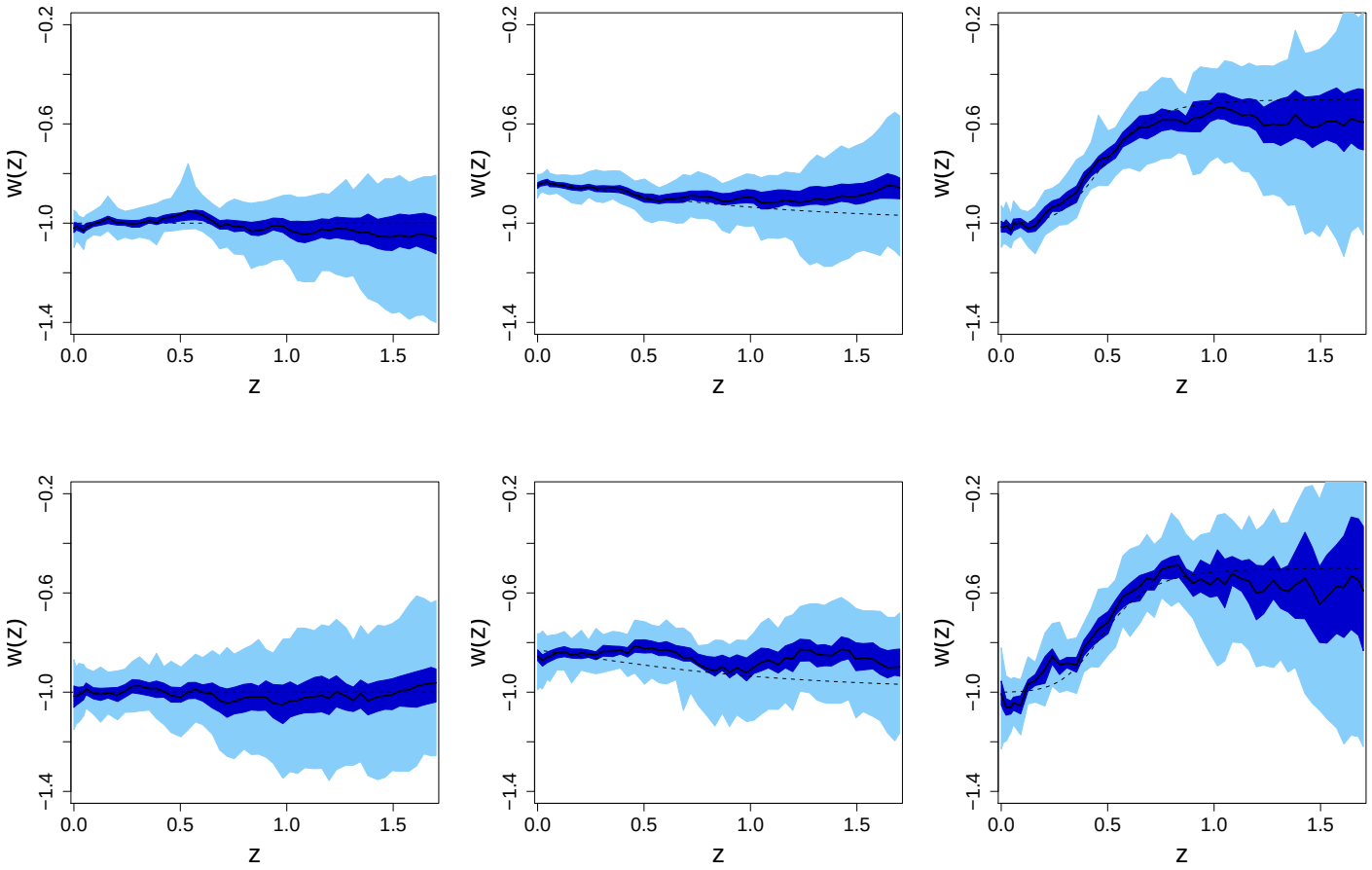
- Σ_{11} is never computed, it would contain a double integral in every entry and because of the interpolation this would be a very large matrix
- There is not one large covariance matrix that has to be inverted to get proposals
- The inner integral is done at the same time as the interpolating, this also helps keep the size of the covariance matrix that must be inverted relatively small
- But $\Sigma_{12} = \kappa^2 \int_0^{s'} \frac{\rho^{|u-s|^\alpha}}{1+u} du$ must be computed

Integration Methods

- Approximate: $\Sigma_{12} = \kappa^2 \int_0^{s'} \frac{\rho^{|u-s|^\alpha}}{1+u} du$ using Chebyshev-Gauss quadrature which uses Chebyshev polynomials of the first kind as its orthogonal polynomials
- Change the limits of integration from $[0, s']$ to $[-1, 1]$
- Approximate the integral with: $\int_{-1}^{1} \frac{f(x)}{\sqrt{1-x^2}} dx \approx \sum_{i=1}^{n} \gamma_i f(x_i)$, $x_i = cos\left(\frac{(2i-1)\pi}{2n}\right)$, and $\gamma_i = \frac{\pi}{n}$ and $n = 100$
- Resulting summation: $K(s, s') = \frac{s'}{2} \sum_{i=1}^{n} \gamma_i \frac{\sqrt{1-u_i^2}}{(1+\frac{s' u_i}{2}+\frac{s'}{2})} \rho^{|\frac{s' u_i}{2}+\frac{s'}{2}-s|^\alpha}$
- The outer integrals were done with a basic trapezoid method

GP Results

Exponential correlation



GP Results

Table: Gaussian Process

Dataset	Ω_m	H_0
μ_1	(0.250,0.290)	(71.53,72.34)
μ_2	(0.243,0.280)	(71.77,72.62)
μ_3	(0.237,0.279)	(71.81,72.68)

Basis of Damped Hermite Polynomials

- Higher order polynomial models do not fit $w(u)$ well
- Karhunen-Loeve expansion of a stochastic process leads to this parametric representation of a Gaussian process by using orthogonal polynomials as basis functions
- $E(H_s^*(Z)H_t^*(Z)) = \delta_{s,t}$ where $Z \sim N(0,1)$ so the Hermite polynomials (H^*) are orthogonal in expectation
- This model is interesting because it approximates the Gaussian process with Gaussian correlation with: $\lambda = \frac{m}{2(1-m^2)}$ and variance $\sigma^2 = \tau^2(1-m^2)^{-1/2}$

Basis of Damped Hermite Polynomials

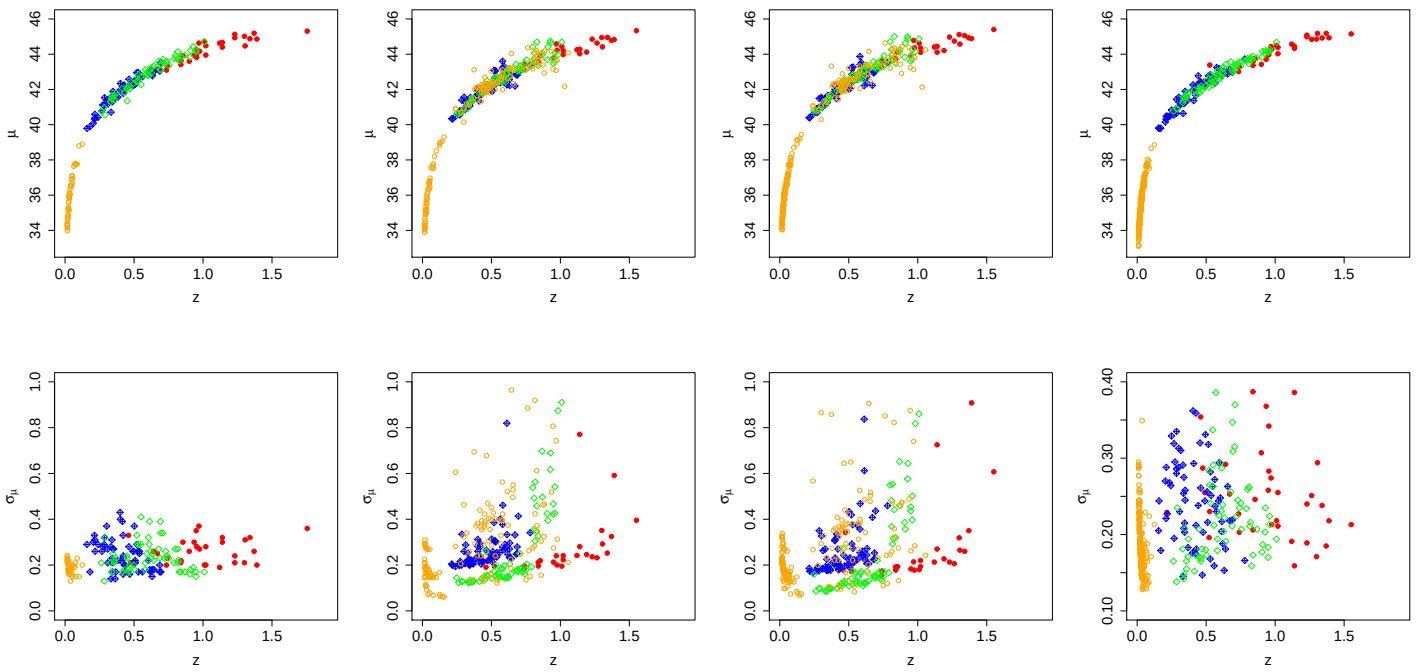
- The series expansion for the first two polynomials (which will have to be scaled to our data):

$$W(u) = \gamma(u) + \sum_{s=0}^{\infty} \beta_s J_s(u)$$
$$= \gamma(u) + \exp -\frac{mu^2}{2(1+m)}[\beta_0 - \frac{1}{\sqrt{2}}\beta_2 + \beta_1 u + (\beta_2 \frac{1}{\sqrt{2}})u^2 ...]$$

- $J_s(u) = H_s^*(u) exp\left(\frac{-mu^2}{2(1+m)}\right)$ is the damped polynomial
- $\pi(\beta) \sim N(0, m^s)$ where m is the dampening term between 0 and 1
- We will fit $W(u) = \frac{w(u)}{1+u}$ because then we can use properties of the Gaussian distribution to do the inner integration

Real Data

These are four sets of currently available real data



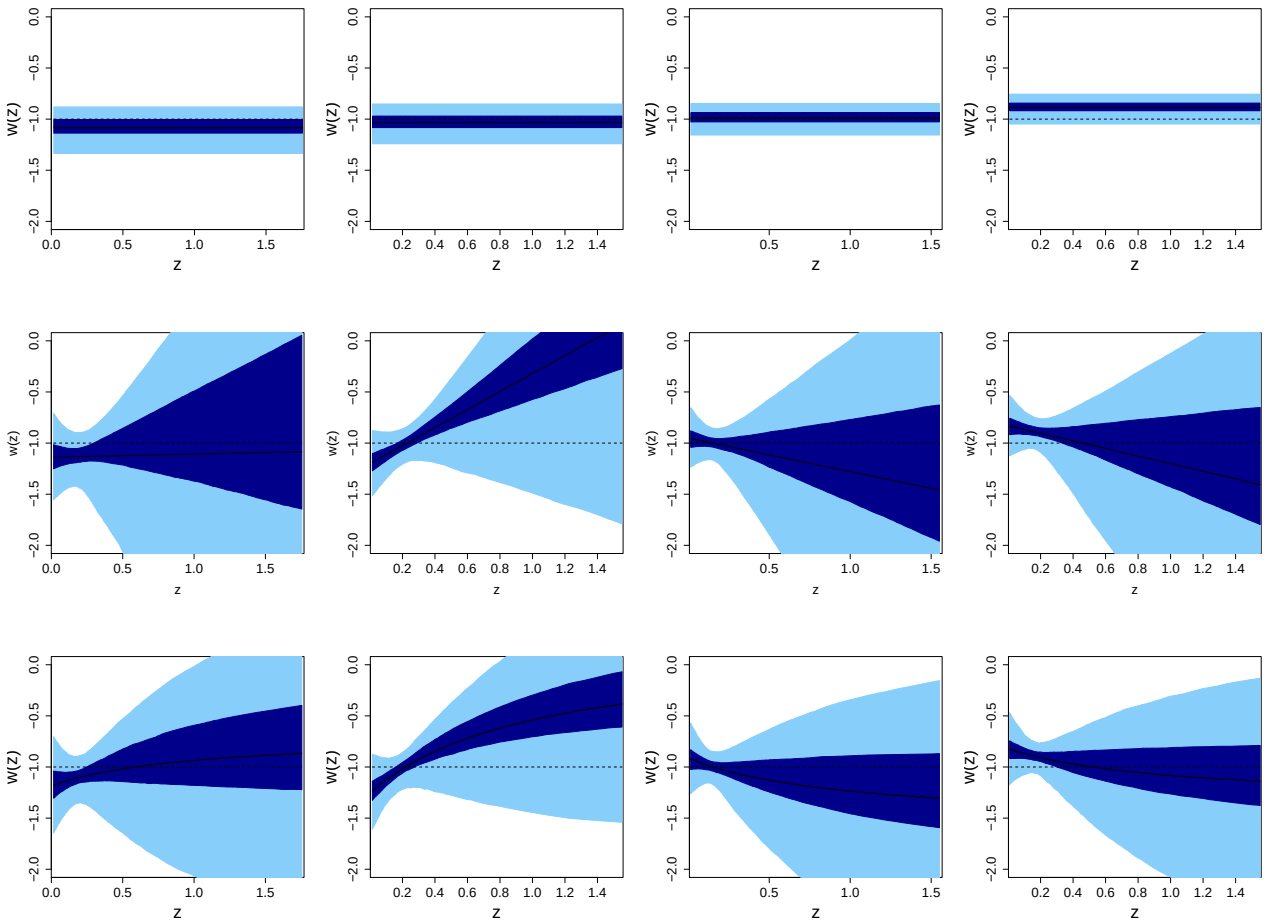
Real Data - Model 1, 2, 3

There is an M that gets incorporated into H_0 from the marginalization of the data during the light curve fitting, so H_0 is closer to 65 than 72

Table: Models 1, 2, 3 - 95% PIs

Model	Dataset	a	b	Ω_{m0}	H_0^*
Model 1	Davis	(-1.334,-0.881)	N/A	(0.229,0.324)	(64.62, 67.64)
	Kowalski	(-1.241,-0.852)	N/A	(0.235,0.334)	(69.08, 71.47)
	SALT3	(-1.156,-0.848)	N/A	(0.233,0.323)	(64.44, 66.10)
	MLCS17	(-1.048,-0.756)	N/A	(0.226,0.320)	(64.37, 65.76)
Model 2	Davis	(-1.576,-0.678)	(-0.263,1.889)	(0.228,0.337)	(64.60, 67.99)
	Kowalski	(-1.545,-0.876)	(-0.549,2.175)	(0.229,0.327)	(69.36, 72.08)
	SALT3	(-1.251,-0.622)	(-2.460,1.199)	(0.233,0.338)	(64.24, 66.22)
	MLCS17	(-1.132,-0.516)	(-2.301,0.958)	(0.229,0.333)	(64.20, 65.77)
Model 3	Davis	(-1.699,-0.666)	(-3.260,2.622)	(0.232,0.328)	(64.54, 68.20)
	Kowalski	(-1.655,-0.867)	(-3.683,1.024)	(0.231,0.327)	(69.45, 72.28)
	SALT3	(-1.286,-0.525)	(-1.805,3.699)	(0.232,0.336)	(64.16, 66.19)
	MLCS17	(-1.184,-0.442)	(-1.707,3.110)	(0.232,0.332)	(64.16, 65.82)

Real Data - Models 1, 2, 3



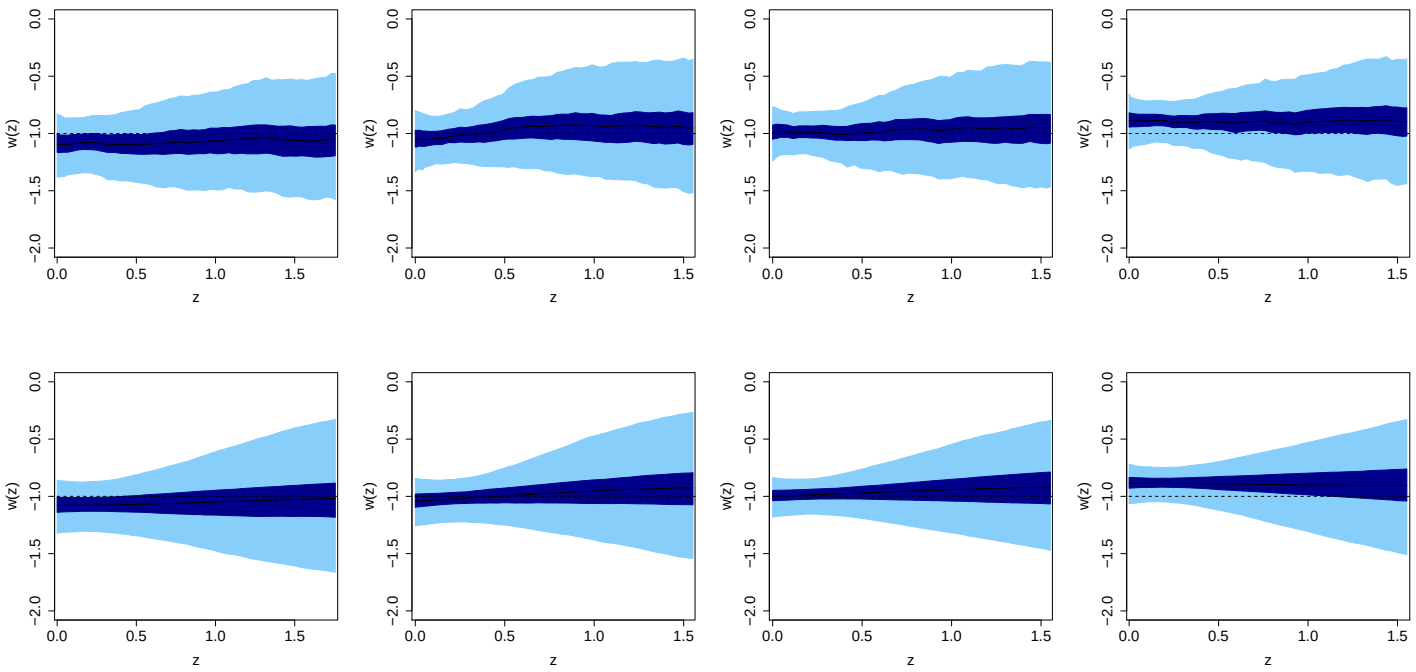
Real Data - Model 4

Table: Models 4 - 95% PIs

Model	Dataset	Ω_{m0}	H_0
Model 4	Davis	(0.232,0.325)	(64.60, 67.63)
Exponential	Kowalski	(0.231,0.330)	(69.11, 71.58)
	SALT3	(0.229,0.323)	(64.39, 66.15)
	MLCS17	(0.226,0.324)	(64.32, 65.83)
Model 4	Davis	(0.228,0.324)	(64.60, 67.51)
Gaussian	Kowalski	(0.233,0.328)	(69.10, 71.52)
	SALT3	(0.231,0.324)	(64.47, 66.16)
	MLCS17	(0.226,0.323)	(64.35, 65.77)

Real Data - Model 4

Exponential and Gaussian correlation



Adding More Variables

- The real data has two more variables that so far we have not used in our analysis: k (the color of the supernovae) and t (the scale of the light curve of the supernovae), where $\mu_i = m_B - M - \alpha(t-1) - \beta k$
- Add radiation density term $\Omega_r = \frac{\sqrt{0.247}}{H_0}$

$$T(z) = M + \alpha(t-1) + \beta k + 25 + 5log_{10}\frac{c(1+z_i)}{H_0}\int_0^{z_i} H(s)ds$$

$$H(s) = \left(\Omega_r(1+s)^4 + \Omega_m(1+s)^3 + (1-\Omega_m-\Omega_r)(1+s)^3 e^{-3\int_0^s \frac{-w(u)}{1+u}du}\right)^{-1/2}$$

- We would like to fit our GP to $w(u)$ with these additional parameters in the model

Experimental Design

- Different telescopes are capable of probing various depths of space
- We would like to help give target values on the z axis
- We want these values of z to either better estimate the mean of the variables or reduce the probability bands
- We could also look at entire regions of the z axis to see how they affect the estimation of $w(u)$
- This type of analysis can give information on what types of telescopes would most benefit future work

Baryon Acoustic Oscillation Data

- We want to add this data to better constrain the estimates of $w(u)$, Ω_m, and H_0

$$A = \frac{\sqrt{\Omega_m}}{h(z_1)^{1/3}} \left(\frac{1}{z_1} \int_0^{z_1} \frac{H_0}{H(s)} ds \right)^{2/3}$$

$$H(s) = \left(\Omega_m (1+s)^3 + (1-\Omega_m)(1+s)^3 e^{-3\int_0^s \frac{-w(u)}{1+u} du} \right)^{-1/2}$$

where $A = 0.469 \left(\frac{n_s}{.98}\right)^{-.35} \pm 0.017$

- There is currently a point at $z_1 = 0.35$ where $n_s = 0.958 \pm 0.016$
- There may be another point at $z_1 = 0.20$ but it has some issues

Cosmic Microwave Background Data

- The physics of this data is best understood compared to other types of data

$$R = \sqrt{\Omega_m} \int_0^{z_2} \frac{H_0}{H(s)} ds$$

$$H(s) = \left(\Omega_m (1+s)^3 + (1-\Omega_m)(1+s)^3 e^{-3\int_0^s \frac{-w(u)}{1+u} du} \right)^{-1/2}$$

- $R = 1.713 \pm 0.020$ and $z_2 = 1087.9 \pm 1.2$

Combined Data Sources

- We can use an already constructed emulator that allows for five different cosmological parameters (like Ω_m, H_0, $w(u)$...)
- We would like to blend these parameters with our current information obtained from the supernovae data
- This type of analysis should be more robust as it employs multiple data sources (like supernova, CMB, BAO...)

Timeline

	Sept	Oct	Nov	Dec	Jan	Feb	March	April	May	June
Hermite orthogonal polynomial model										
Add more parameters to the supernovae model										
Add BAO data to the models										
Add CMB data to the models using hierarchical modeling										
Experimental design - where is more data needed on the z axis										
Add other types of data to the model										
Finish writing up dissertation										