On Finding Exoplanets

Floyd Bullard

SAMSI/CRSC Undergraduate Workshop at NCSU

22 May 2006
Astrostatistics working groups

- **Exoplanets**: Finding planets outside the solar system.
- **Surveys and population studies**: Analyzing large survey data, such as galaxy surveys.
- **Gravitational lensing**: Analyzing the distorted images that result from light bending around massive objects.
- **Source detection**: Distinguishing point sources from noise, for example.
- **Particle physics**
Figure: The transit of a planet in front of its star will slightly dim the light visible from the earth—by about 1%.
Figure: These brightness data are plotted against the orbital phase.
Figure: A background star’s image distorts when a massive object passes somewhere in front of it.
A rare event

Figure: On Wednesday, 10 August 2005, a planet and its host star were detected passing in front of a distant background star.
Figure: The Doppler shift in a star’s light allows its radial velocity to be measured to an astonishing precision of 3, or even 1 meter per second.
A planet is detected

Figure: Over many days of observations, if the radial velocities show a characteristic periodic behavior (plus or minus some error), it can be inferred that a planet is orbiting the star.
51 Pegasi: the star that started it all

Figure: This star, 51 Pegasi, was observed 111 times from 11 October 1995 to 31 August 1996.
On Finding Exoplanets

Floyd Bullard

Introduction

SAMSI Astrostatistics Workshop

Transits

Gravitational lensing

Radial velocity detection

51 Peg

HD3651

HD73526

What’s hard?

Multiple planets

Adaptive design

Conclusion

51 Peg

Figure: 111 radial velocity measurements.
A clear period of about 4 days

**Figure:** Shown modulo 4.23 days, the periodicity in the data is clear.
A very good cosine fit
A very good cosine fit

Figure: $V(t) \approx 5 + 55\cos\left(\frac{2\pi}{4.23} t + 0.9\right)$
The radial velocity model for circular orbits

For planets with circular orbits, the radial velocity model we are using is:

Data: \((t_i, V_i, \sigma_i)\)

\[ V_i \overset{\text{ind}}{\sim} \mathcal{N}(V(t_i), \sigma_i^2 + s^2), \text{where} \]

\[ V(t) = C + K\cos \left( \left( \frac{2\pi}{P} \right) t + M_0 \right) \]
The radial velocity model for circular orbits

For planets with circular orbits, the radial velocity model we are using is:

Data: \((t_i, V_i, \sigma_i)\)

\[ V_i \sim \text{ind} \mathcal{N}(V(t_i), \sigma_i^2 + s^2), \]

\[ V(t) = C + K \cos \left( \frac{2\pi}{P} t + M_0 \right) \]

- \(C\) background stellar velocity; uninteresting (as regards exoplanets)
The radial velocity model for circular orbits

For planets with circular orbits, the radial velocity model we are using is:

Data: \((t_i, V_i, \sigma_i)\)

\[ V_i \overset{\text{ind}}{\sim} N(V(t_i), \sigma_i^2 + s^2), \]  
where  
\[ V(t) = C + K \cos\left(\left(\frac{2\pi}{P}\right) t + M_0\right) \]

- \(C\) background stellar velocity; uninteresting (as regards exoplanets)
- \(K\) amplitude; related to planet’s mass
The radial velocity model for circular orbits

For planets with circular orbits, the radial velocity model we are using is:

Data: \((t_i, V_i, \sigma_i)\)

\(V_i \overset{\text{ind}}{\sim} N(V(t_i), \sigma_i^2 + s^2),\) where

\[ V(t) = C + K \cos\left(\frac{2\pi}{P} t + M_0\right) \]

- \(C\): background stellar velocity; uninteresting (as regards exoplanets)
- \(K\): amplitude; related to planet’s mass
- \(P\): period; related to distance from star
The radial velocity model for circular orbits

For planets with circular orbits, the radial velocity model we are using is:

Data: \((t_i, V_i, \sigma_i)\)

\[ V_i \overset{\text{ind}}{\sim} N(V(t_i), \sigma_i^2 + s^2), \]

where

\[ V(t) = C + K \cos \left( \frac{2\pi}{P} t + M_0 \right) \]

- \(C\) background stellar velocity; uninteresting (as regards exoplanets)
- \(K\) amplitude; related to planet’s mass
- \(P\) period; related to distance from star
- \(M_0\) phase shift; an utterly uninteresting artifact
The radial velocity model for circular orbits

For planets with circular orbits, the radial velocity model we are using is:

Data: \((t_i, V_i, \sigma_i)\)

\[ V_i \overset{\text{ind}}{\sim} N(V(t_i), \sigma_i^2 + s^2), \]  
where

\[ V(t) = C + K \cos \left(\frac{2\pi}{P} t + M_0\right) \]

- \(C\) background stellar velocity; uninteresting (as regards exoplanets)
- \(K\) amplitude; related to planet’s mass
- \(P\) period; related to distance from star
- \(M_0\) phase shift; an utterly uninteresting artifact
- \(s\) “stellar jitter”; an unknown amount of extra noise, in addition to the (known) measurement error
**Figure:** These data (green points are duplicates) are shown modulo 62.23 days. Clearly there is a periodicity to the radial velocities, but it doesn’t appear to be sinusoidal.
On Finding Exoplanets

Floyd Bullard

Introduction

SAMSI Astrostatistics Workshop

Transits

Gravitational lensing

Radial velocity detection

51 Peg

HD3651

HD73526

What’s hard?

Multiple planets

Adaptive design

Conclusion

---

What’s going on?

Figure: What’s going on?
A simulation of a binary star system

Binary star simulation
Figure: The HD3651 data, shown with a Keplerian orbit model. The eccentricity is about 0.63.
Full Keplerian orbit model

The model for the data in a single-planet system is:

\[ V_i \overset{\text{ind}}{\sim} N \left( C + \Delta V(t_i|\theta), \sigma_i^2 + s^2 \right), \]

where \( \theta \) has five parameters, which may be
Full Keplerian orbit model

The model for the data in a single-planet system is:

$$V_i \overset{\text{ind}}{\sim} N \left( C + \Delta V(t_i|\theta), \sigma_i^2 + s^2 \right),$$

where $\theta$ has five parameters, which may be

$K$ amplitude
Full Keplerian orbit model

The model for the data in a single-planet system is:

\[ V_i^{\text{ind}} \sim N \left( C + \Delta V(t_i|\theta), \sigma_i^2 + s^2 \right), \]

where \( \theta \) has five parameters, which may be

- \( K \) amplitude
- \( P \) period
On Finding Exoplanets

Floyd Bullard

Introduction

SAMSI Astrostatistics Workshop
Transits
Gravitational lensing
Radial velocity detection

51 Peg
HD3651
HD73526
What’s hard?
Multiple planets
Adaptive design
Conclusion

Full Keplerian orbit model

The model for the data in a single-planet system is:

$$V_i^{\text{ind}} \sim N \left( C + \Delta V(t_i|\theta), \sigma_i^2 + s^2 \right),$$

where $\theta$ has five parameters, which may be:

- $K$ amplitude
- $P$ period
- $M_0$ phase shift
Full Keplerian orbit model

The model for the data in a single-planet system is:

\[ V_i \overset{\text{ind}}{\sim} \mathcal{N} (C + \Delta V(t_i|\theta), \sigma_i^2 + s^2), \]

where \( \theta \) has five parameters, which may be

- \( K \) amplitude
- \( P \) period
- \( M_0 \) phase shift
- \( e \) eccentricity (between 0 and 1, 0 = circular, 0.99 = highly elliptical)
Full Keplerian orbit model

The model for the data in a single-planet system is:

$$V_i \sim N(C + \Delta V(t_i|\theta), \sigma_i^2 + s^2),$$

where $\theta$ has five parameters, which may be:

- $K$ amplitude
- $P$ period
- $M_0$ phase shift
- $e$ eccentricity (between 0 and 1, 0 = circular, 0.99 = highly elliptical)
- $\omega$ argument of periastron (between 0 and $2\pi$)
On Finding Exoplanets

Floyd Bullard

Introduction

SAMSI Astrostatistics Workshop

Transits

Gravitational lensing

Radial velocity detection

51 Peg

HD3651

HD73526

What’s hard?

Multiple planets

Adaptive design

Conclusion

Figure: 18 radial velocity measurements made on HD73526 between 2 February 1999 and 29 June 2002.
On Finding Exoplanets

Floyd Bullard

Introduction
SAMSI Astrostatistics Workshop
Transits
Gravitational lensing
Radial velocity detection

51 Peg
HD3651
HD73526

What’s hard?
Multiple planets
Adaptive design

Conclusion

Figure: data shown modulo 128 days
HD73526

Figure: data shown modulo 128 days with “best-fit” Keplerian model.
On Finding Exoplanets

Floyd Bullard

Introduction

SAMSI Astrostatistics Workshop

Transits

Gravitational lensing

Radial velocity detection

51 Peg

HD3651

HD73526

What’s hard?

Multiple planets

Adaptive design

Conclusion

HD73526

Figure: data shown modulo 190 days.
Figure: data shown modulo 190 days with “best-fit” Keplerian model.
Introduction

SAMSI Astrostatistics Workshop
Transits
Gravitational lensing
Radial velocity detection

51 Peg
HD3651
HD73526

What's hard?
Multiple planets
Adaptive design
Conclusion

# HD73526

**Figure:** data shown modulo 376 days.
On Finding Exoplanets

Floyd Bullard

Introduction

SAMSI Astrostatistics Workshop

Transits

Gravitational lensing

Radial velocity detection

51 Peg

HD3651

HD73526

What’s hard?

Multiple planets

Adaptive design

Conclusion

Figure: data shown modulo 376 days with “best-fit” Keplerian model.
On Finding Exoplanets

Floyd Bullard

Introduction

SAMSI Astrostatistics Workshop
Transits
Gravitational lensing
Radial velocity detection

51 Peg
HD3651
HD73526

What’s hard?
Multiple planets
Adaptive design
Conclusion

Figure: Which is the correct period?
When is it hard to find a good solution?

What makes finding a solution hard?
When is it hard to find a good solution?

What makes finding a solution hard?
- The planet is small.
When is it hard to find a good solution?

What makes finding a solution hard?

- The planet is small. *(Or there isn’t really one there at all!)*
When is it hard to find a good solution?

What makes finding a solution hard?

- The planet is small. *(Or there isn’t really one there at all!)*
- There aren’t a lot of data.
When is it hard to find a good solution?

What makes finding a solution hard?

- The planet is small. *(Or there isn’t really one there at all!)*
- There aren’t a lot of data.
- The orbit is long.
Multimodality

We say that the likelihood function (you’ll hear more about this tomorrow) is multi-modal, like a bumpy surface—many solutions exist that appear better than all their neighbors.

Figure: When there are multiple planets, their gravitational effects on the star are (often) additive. These observations on HD69830 were made between October 2003 and January 2006.
Multiple planet problems

For a $k$-planet Keplerian model, there are $2 + 5k$ parameters to be estimated: five for each planet, plus $C$ (background velocity) and $s$ (stellar jitter). The likelihood function is highly multi-modal in a fairly high-dimensional space. This makes exploring the space and finding good solutions difficult, especially when there aren’t many data points.
Decision theory

Given our existing state of knowledge, when should we make a future observation? What stars should we look at next?
HD73526 (revisited)

Figure: When should future observations be made on this star?
Figure: This planet has a long period.
Figure: Twelve randomly selected solutions from the “posterior space” of solutions. When should future observations be made on this star?
Questions of interest to astrophysicists:
Questions of interest to astrophysicists:

- Is there a planet there? If so, how many?
Questions of interest to astrophysicists:

- Is there a planet there? If so, how many?
- If there is a planet, what are its orbital parameters? (These tell us things about the planet’s mass, temperature, etc.)
Questions of interest to astrophysicists:

- Is there a planet there? If so, how many?
- If there is a planet, what are its orbital parameters? (These tell us things about the planet’s mass, temperature, etc.)
- When should future observations be made, and on what stars, so as to maximize the amount of useful information gained?
Welcome to North Carolina and the SAMSI/CRSC Undergraduate Workshop at NCSU! You’re going to have a great week. Don’t hesitate to ask any of us questions at any time.