GEOL/PHYS 6670
Geophysical Inverse Theory

Prof. Anne Sheehan
Anne.Sheehan @ colorado.edu
Office: Benson Rm 440A
Class web page on Canvas

Monday office hours changed to 3:30-4:20

Lecture 2

Start recording!

Finish Menke Ch 1 - Describing Inverse Problems

Start Menke Ch 2 - Probability

Homework 1 – due now (on Canvas)
(math review, try LinkedIn Learning)

Homework 2 – due next week – available on Canvas
(setting up inverse problems, look up a paper)

Last Lecture

Forward vs. inverse problems

Setting up inverse problems

A few examples (line fit, tomography)

This Lecture

Moment Tensor Example

Linear versus nonlinear

Probability

2nd hour of class –Python tutorial (Optional)

Where we left off last time…

- **Grid search – Forward modeling**

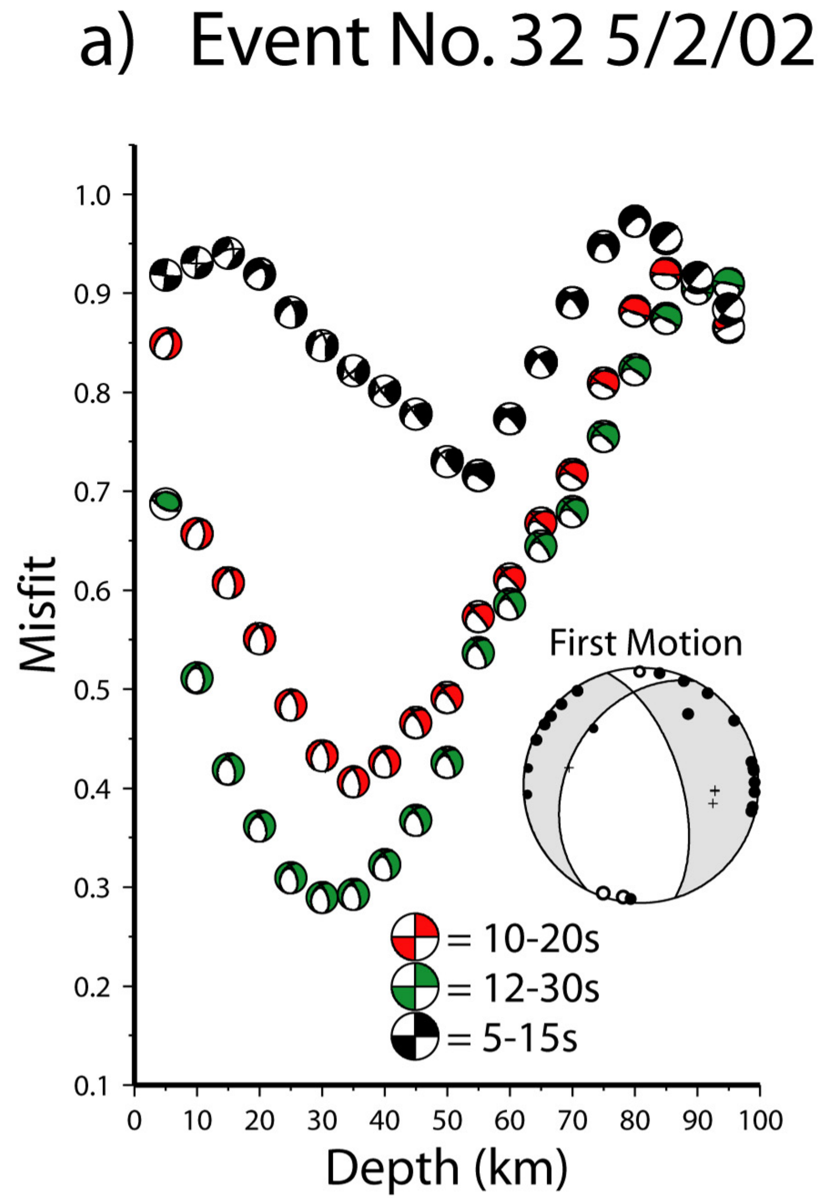
  try many models, calculate misfit

  look for minimum difference in misfit

  ok for small number of parameters

Grid Search Example:

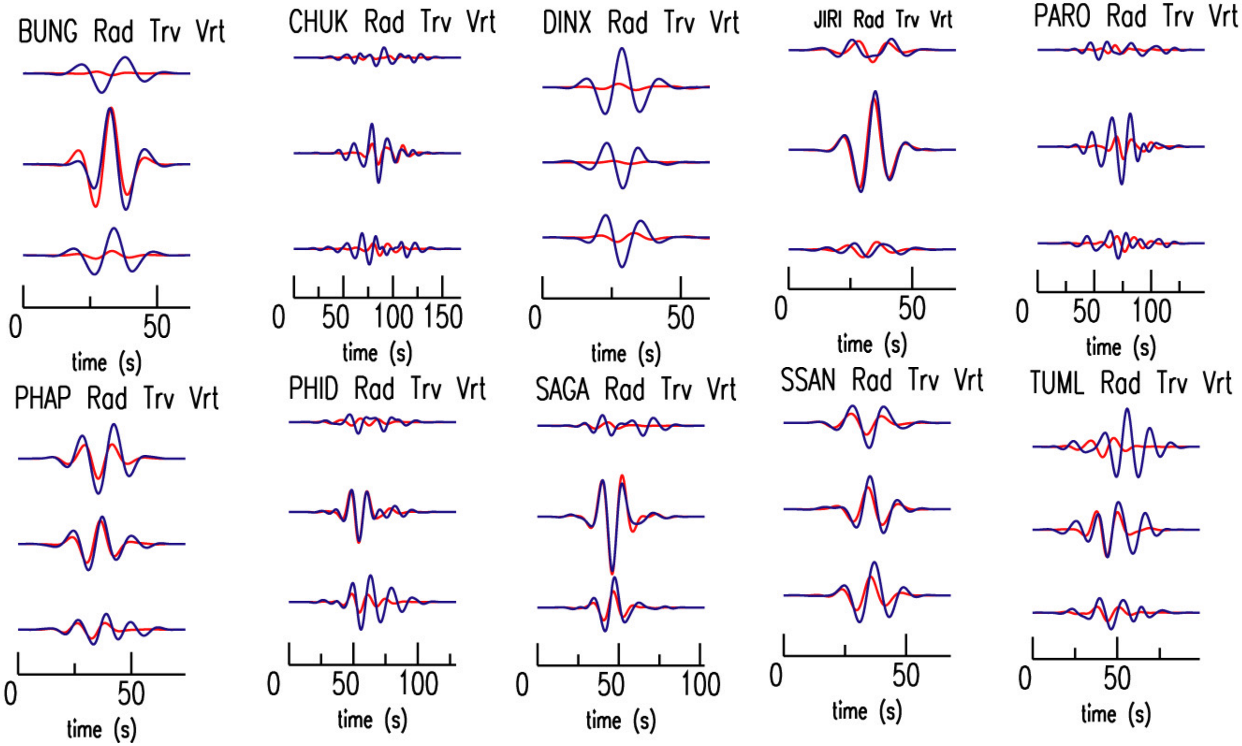
Find depth that minimizes misfit in moment tensor inversion

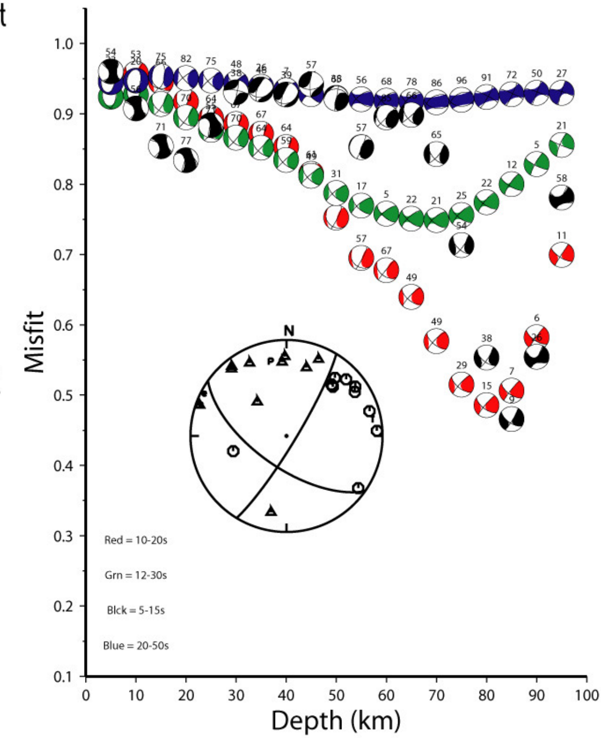


T. De la Torre

Full waveform earthquake moment tensor inversion

No. 14





Blue waveforms = Observed data
Red waveforms = synthetics

Moment Tensor Inversion with grid search over depth and bandpass
Red = 10-20 sec **Green** = 12-30 sec
Black = 5-15 sec **Blue** = 20-50 sec

T. De la Torre

USGS moment tensors (focal mechanisms) (an example of inversion)

Recent earthquakes – Haiti, Alaska

M 7.2 - Nippes, Haiti

2021-08-14 12:29:08 (UTC) | 18.408°N 73.475°W | 10.0 km depth

Moment Tensor

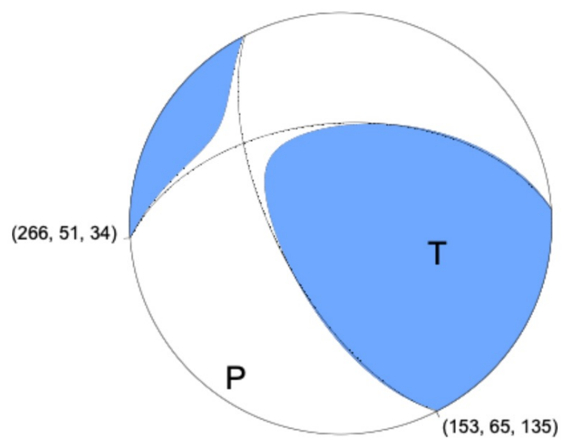
View all moment-tensor products (1 total)

Contributed by US [4] last updated 2021-08-14 16:01:57 (UTC)

✓ The data below are the most preferred data available
✓ The data below have been reviewed by a scientist

W-phase Moment Tensor (Mww)

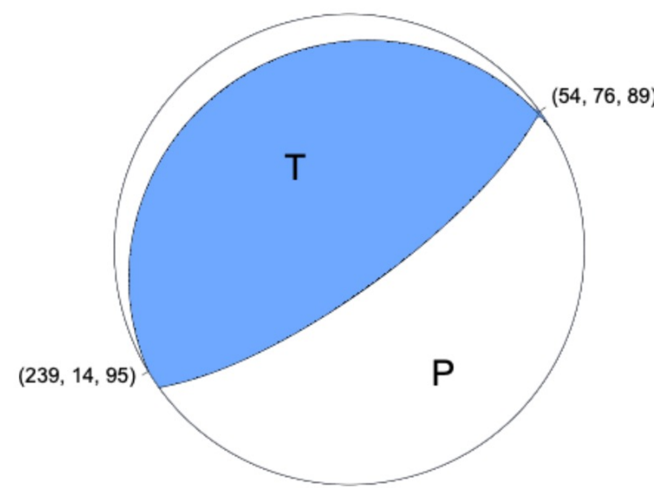
Moment	7.598e+19 N-m
Magnitude	7.19 Mww
Depth	15.5 km
Percent DC	90%
Half Duration	16.00 s
Catalog	US



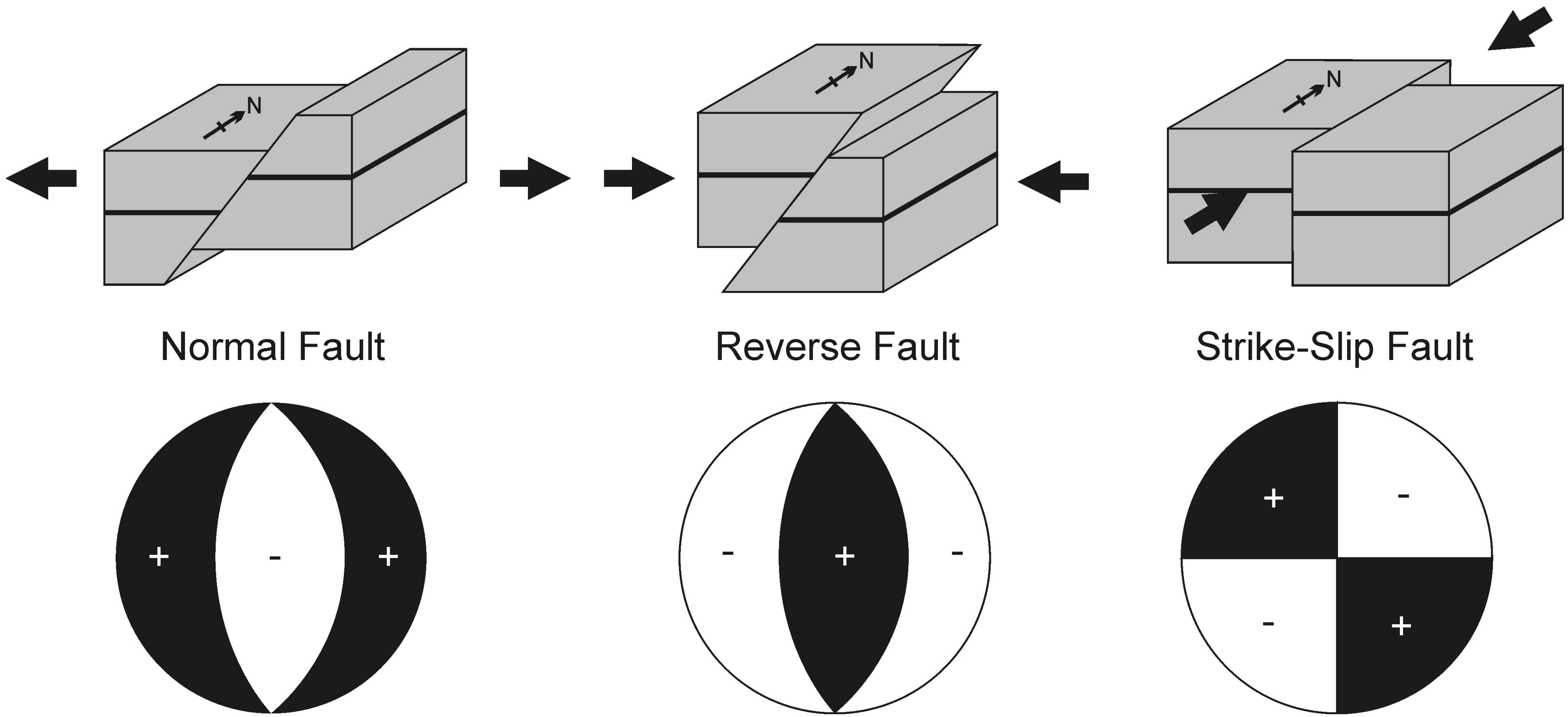
M 8.2 - 99 km SE of Perryville, Alaska

2021-07-29 06:15:49 (UTC) | 55.364°N 157.888°W | 35.0 km depth

Moment Tensor

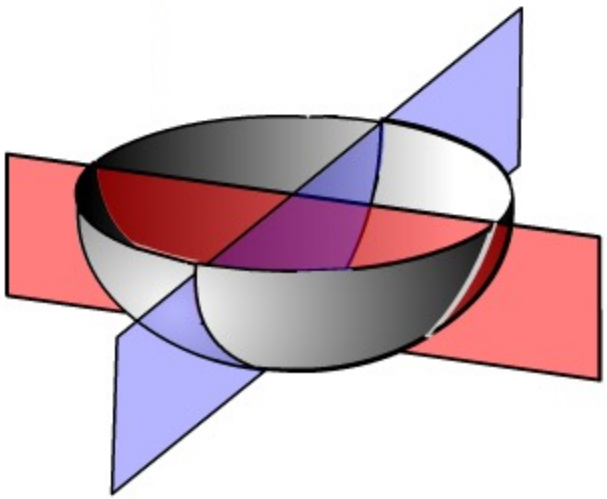


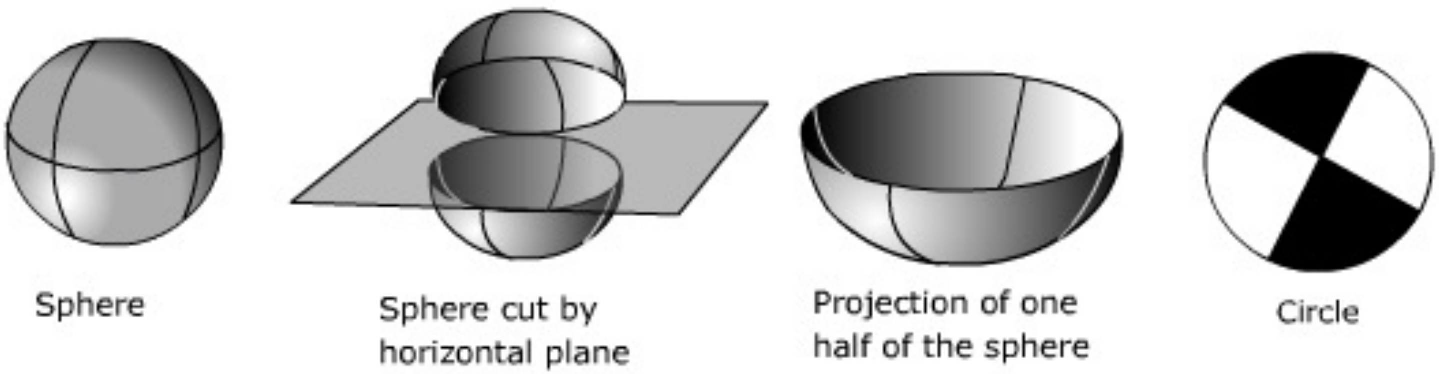
Fault types and "Beach Ball" plots



Stereographic projection

- A method of projecting half a sphere onto a circle.
- e.g. planes cutting vertically through the sphere plot as straight lines

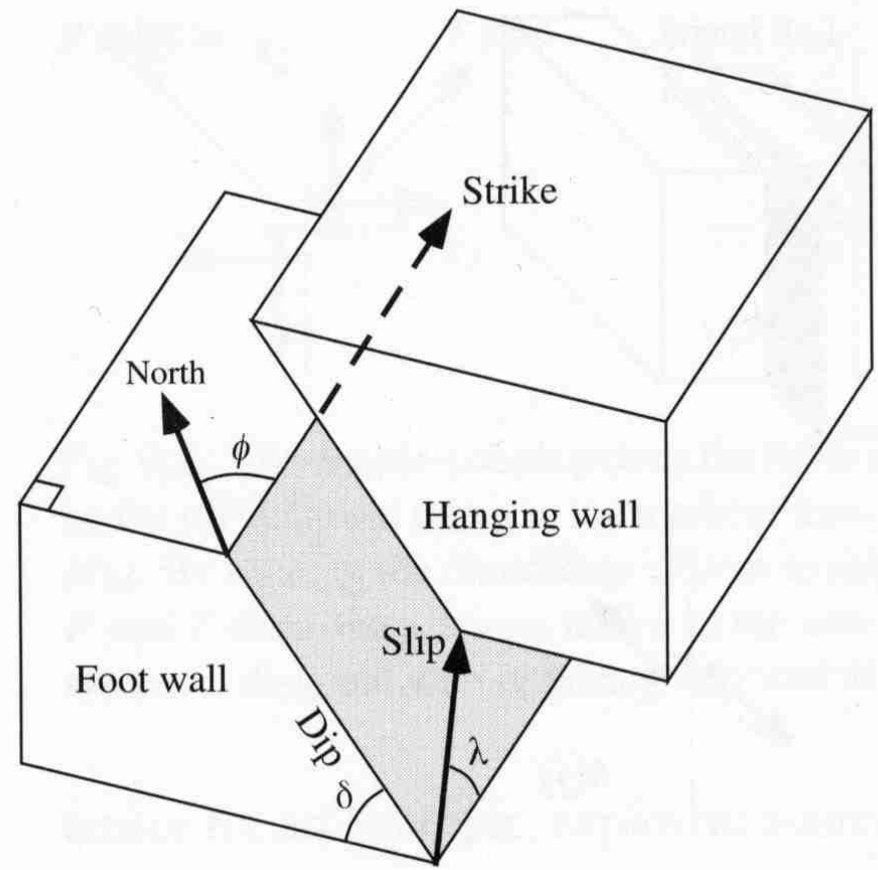




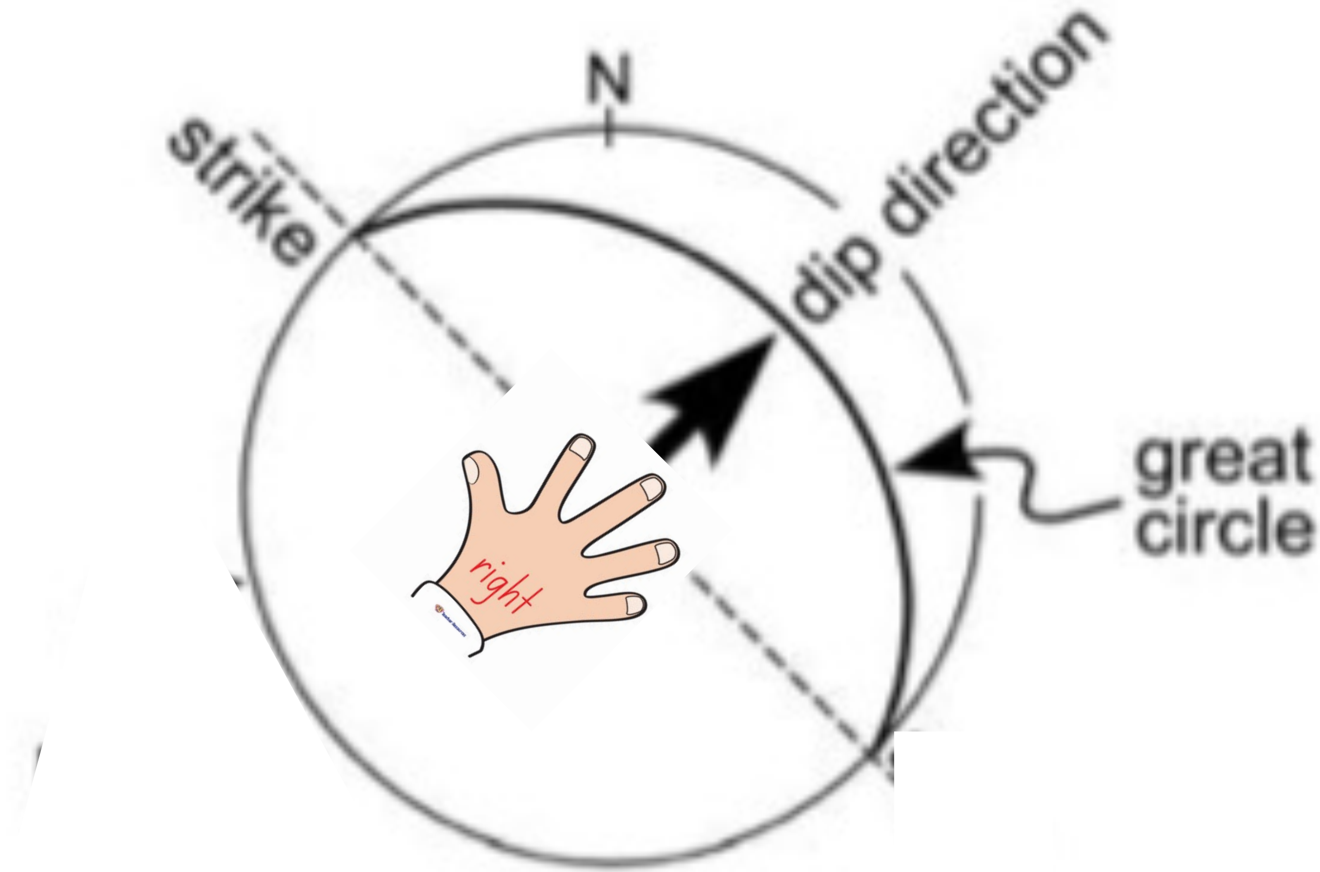
Images from http://www.learninggeoscience.net/free/00071/index.html

Strike, Dip, and Rake

- Strike, the azimuth of fault plane (usually given relative to N)
- Dip, angle with the horizontal
- Rake, angle between slip vector and strike



(Shearer, 1999)


N
strike
dip direction
great
circle
right

USGS moment tensors (focal mechanisms)
(an example of inversion)

Recent earthquakes – Haiti, Alaska

M 7.2 - Nippes, Haiti

2021-08-14 12:29:08 (UTC) | 18.408°N 73.475°W | 10.0 km depth

Moment Tensor

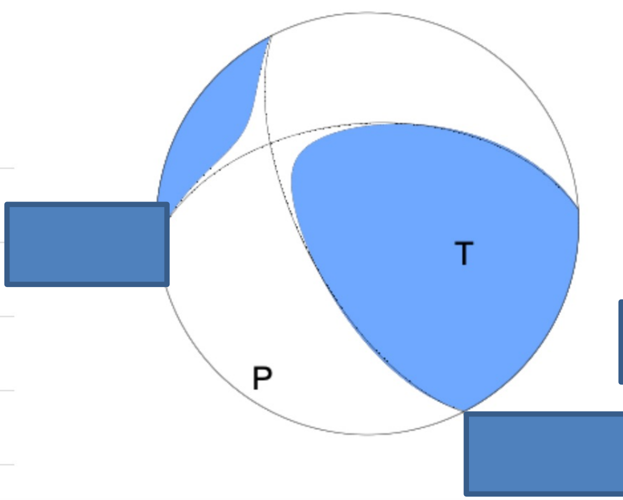
View all moment-tensor products (1 total)

Contributed by US [4] last updated 2021-08-14 16:01:57 (UTC)

✓ The data below are the most preferred data available
✓ The data below have been reviewed by a scientist

W-phase Moment Tensor (Mww)

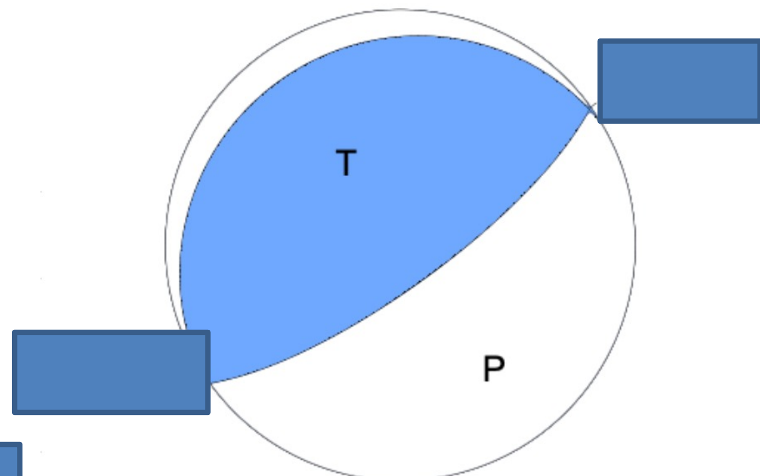
Moment	7.598e+19 N-m
Magnitude	7.19 Mww
Depth	15.5 km
Percent DC	90%
Half Duration	16.00 s
Catalog	US



M 8.2 - 99 km SE of Perryville, Alaska

2021-07-29 06:15:49 (UTC) | 55.364°N 157.888°W | 35.0 km depth

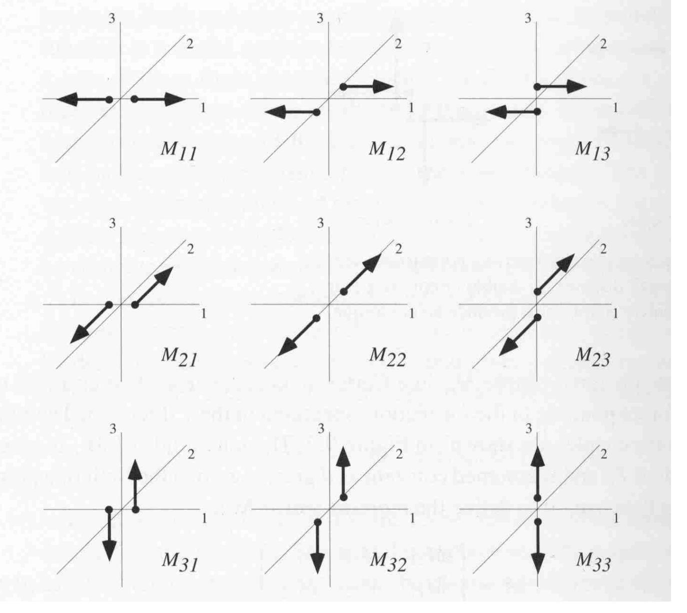
Moment Tensor



Moment Tensor Basics

- A moment tensor is a complete description of equivalent forces of a general seismic point source (Jost and Herrmann, 1989) in an elastic medium(Shearer 1999).

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}$$



(Shearer, 1999)

Moment Tensor Types

- Isotropic (Explosion)
- Double Couple (Shear dislocation)
- Compensated Linear Vector Dipole

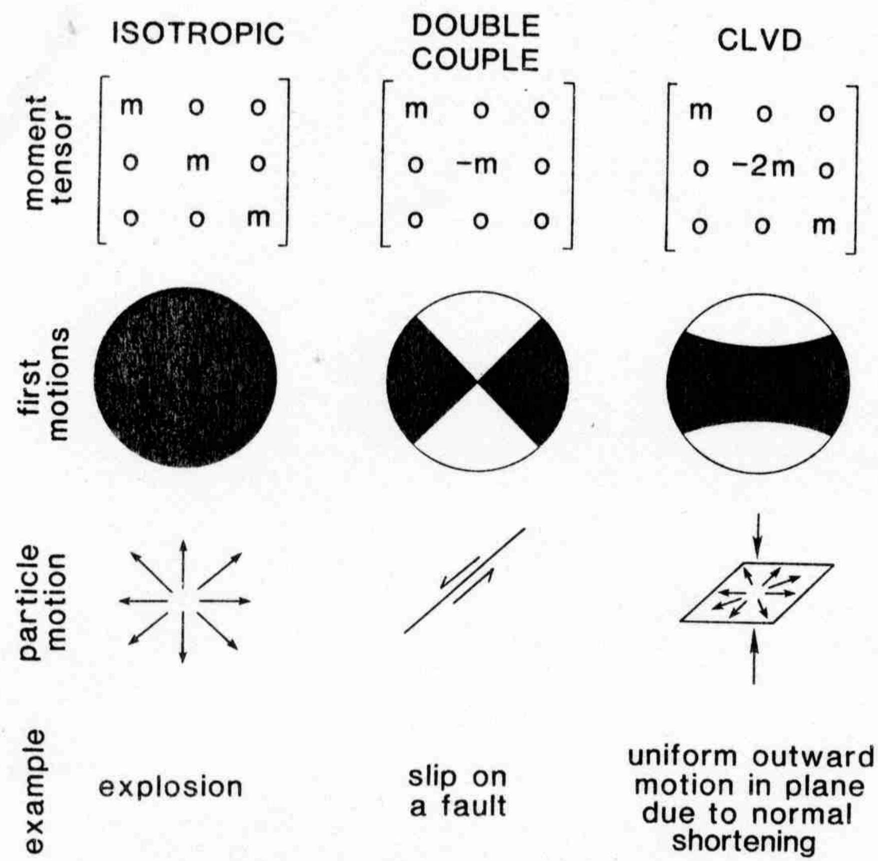


Fig. 2. Depiction of "pure" earthquake source models. The figure shows isotropic, double couple, and compensated linear vector dipole (CLVD) source models as moment tensors, equal-area projections of P wave first motions, and near-source particle motions. (From Apperson [1991].)

(Frohlich and Apperson, 1992)

Basic Equation for Moment Tensor Inversion

$$\mathbf{d} = \mathbf{G}\,\mathbf{M}$$

- d = vector containing *n* sampled values of observed ground displacement at various times and stations, the DATA
- G = 6 x *n* matrix of the synthetic Green’s function
- M = contains the six independent moment tensor elements to be determined in the inversion

Moment tensor inversion

(Eqn 4.3.15). Here, we define $G_{ij}(t)$ as the seismogram at the i^{th} seismometer due to the moment tensor component m_j. $G_{ij}(t)$ includes the effects of the seismometer and earth structure along the path from the source to this seismometer, so the i^{th} seismogram is the sum of the Green's functions weighted by the moment tensor components,

$$u_i(t) = \sum_{i=1}^{6} G_{ij}(t) m_j. \tag{22}$$

$$\begin{pmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ u_n \end{pmatrix} = \begin{pmatrix} G_{11} & G_{12} & G_{13} & G_{14} & G_{15} & G_{16} \\ G_{21} & G_{22} & G_{23} & G_{24} & G_{25} & G_{26} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ G_{n1} & G_{n2} & G_{n3} & G_{n4} & G_{n5} & G_{n6} \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \\ m_6 \end{pmatrix}.$$

Stein and Wysession

Setting up inverse problems

1. Find the equations that relate your data to a model
2. Be able to do the forward problem
3. Check whether your system of equations is linear or not
4. Decide how to solve
   ie damping, weighting, parameterization of model, linearize, grid search, etc. (regularization)
5. Evaluate solution(s)
   Does the solution make sense? What does the data suggest?
   Uniqueness - there is usually not a single model that is the only possibility

Linear or nonlinear?

Function f(x) is linear if

$$f(ax) = af(x)$$

Example

$f(x) = mx$ linear

$f(x) = mx^2$ nonlinear

test it

$$f(ax) = m(ax)^2 = ma^2x^2 = a^2(mx^2) \neq a(mx^2)$$

Linear or nonlinear?

$d_1 = 2m_1 + 4m_2$

$d_1 = 2m_1 + 4m_1^2 m_2$

Linear or nonlinear?

$d_1 = 2m_1 + 4m_2$

$d_1 = 2m_1 + 4m_1^2 m_2$

Recall, function f(x) is linear if

$$f(ax) = af(x)$$

$f(m_1, m_2) = 2m_1 + 4m_2$
$f(am_1, am_2) = ?$

$f(m_1, m_2) = 2m_1 + 4m_1^2 m_2$
$f(am_1, am_2) = ?$

Ways to linearize

1. Reparameterize

   Example 1: solve for slowness instead of velocity

   time = distance / velocity = distance x slowness

   Example 2: natural log parameterization

   $d = m_1 e^{m2z2}$

   take natural log of each side

   $\ln d = \ln m_1 + m_2 z^2$

   and use ln d instead of d, and solve for $\ln m_1$ instead of m_1

Ways to linearize

1. Reparameterize

2. Taylor series expansion about an initial guess

Ways to linearize

1. Reparameterize

2. Taylor series expansion about an initial guess
 (we will come back to this later in the term)

3. Keep it nonlinear and forward model (can use many grid search type methods)

Homework 2 – due next Tuesday

Setting up inverse problems

1. Find the equations that relate your data to a model
2. Be able to do the forward problem
3. Check whether your system of equations is linear or not
4. Decide how to solve

   ie damping, weighting, parameterization of model, linearize, grid search, etc.
5. Evaluate solution(s)

   Does the solution make sense? What does the data suggest?

   Uniqueness - there is usually not a single model that is the only possibility

Setting up inverse problems

1. Find the equations that relate your data to a model
2. Be able to do the forward problem
3. Check whether your system of equations is linear or not
4. Decide how to solve
   ie damping, weighting, parameterization of model, linearize, grid search, etc.
5. Evaluate solution(s)
   Does the solution make sense? What does the data suggest?
   Uniqueness - there is usually not a single model that is the only possibility

With inverse problems, what kind of solution are we looking for ?

A: Estimate of model parameters

numerical values

$$m_1 = 10.5$$
$$m_2 = 7.2$$

But we need confidence limits, too

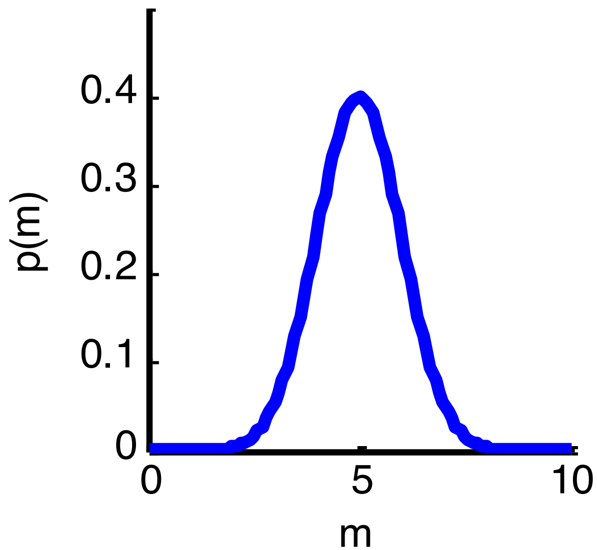
$$m_1 = 10.5 \pm 0.2$$
$$m_2 = 7.2 \pm 0.1$$

or

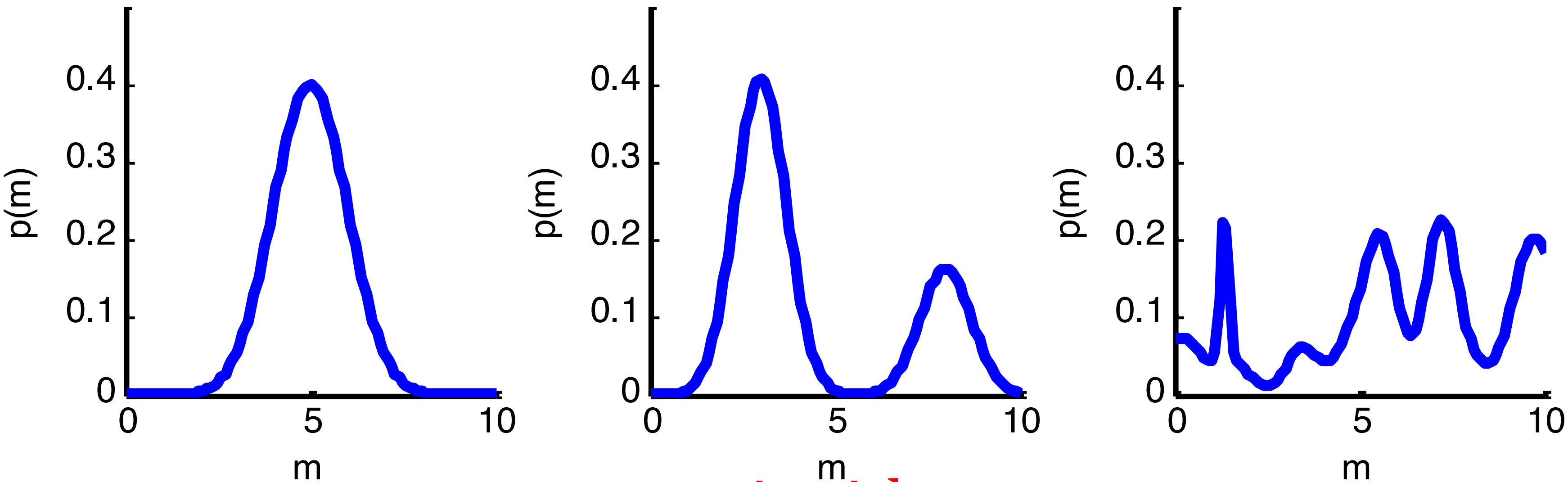
$$m_1 = 10.5 \pm 22.3$$
$$m_2 = 7.2 \pm 9.1$$

completely different implications!

B: probability density functions



if $p(m_1)$ simple
not so different than confidence intervals



m is about 5
plus or minus 1.5

m is either about 3
plus or minus 1
or about 8
plus or minus 1
but that's less likely

we don't really know anything useful about m

Fig 1.9

C: localized averages

$$A = 0.2m_9 + 0.6m_{10} + 0.2m_{11}$$

might be better determined than either m_9 or m_{10} or m_{11} individually

Is this useful?

Do we care about

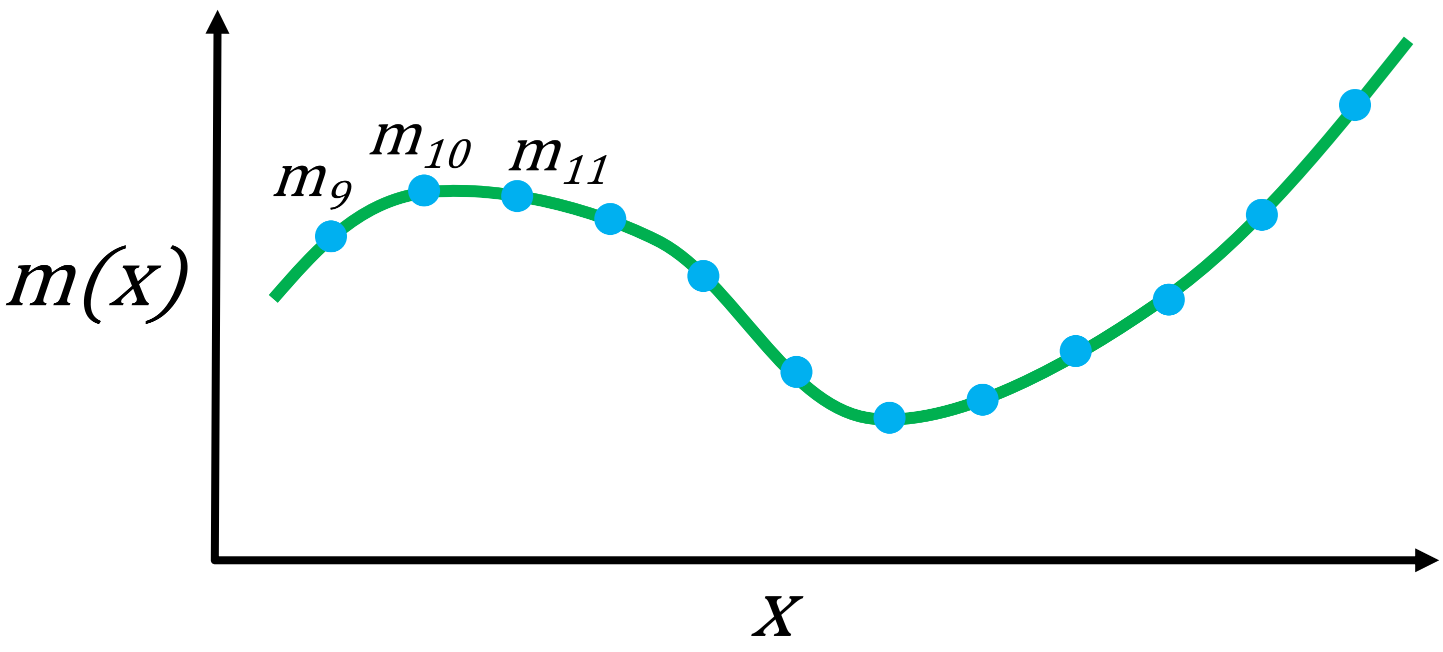
$$A = 0.2m_9 + 0.6m_{10} + 0.2m_{11}$$

?

Maybe …

Suppose

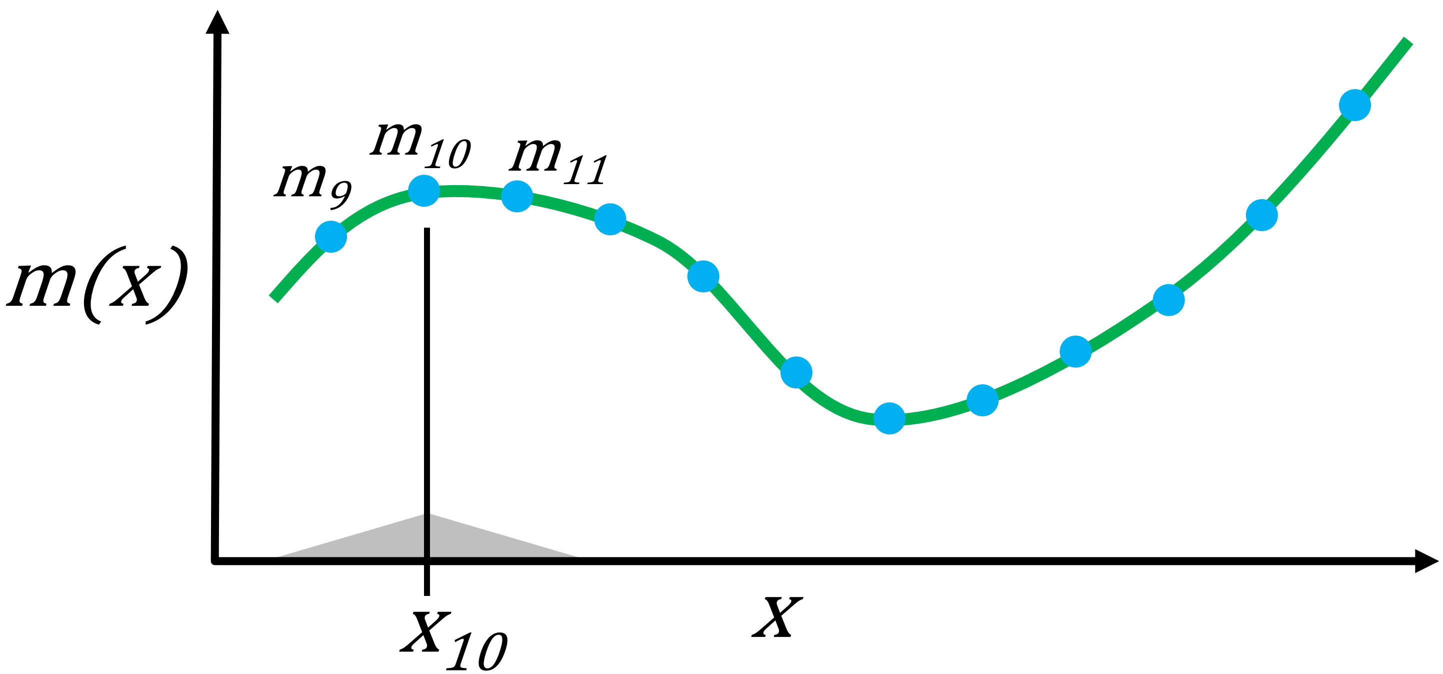
if **m** is a discrete approximation of $m(x)$



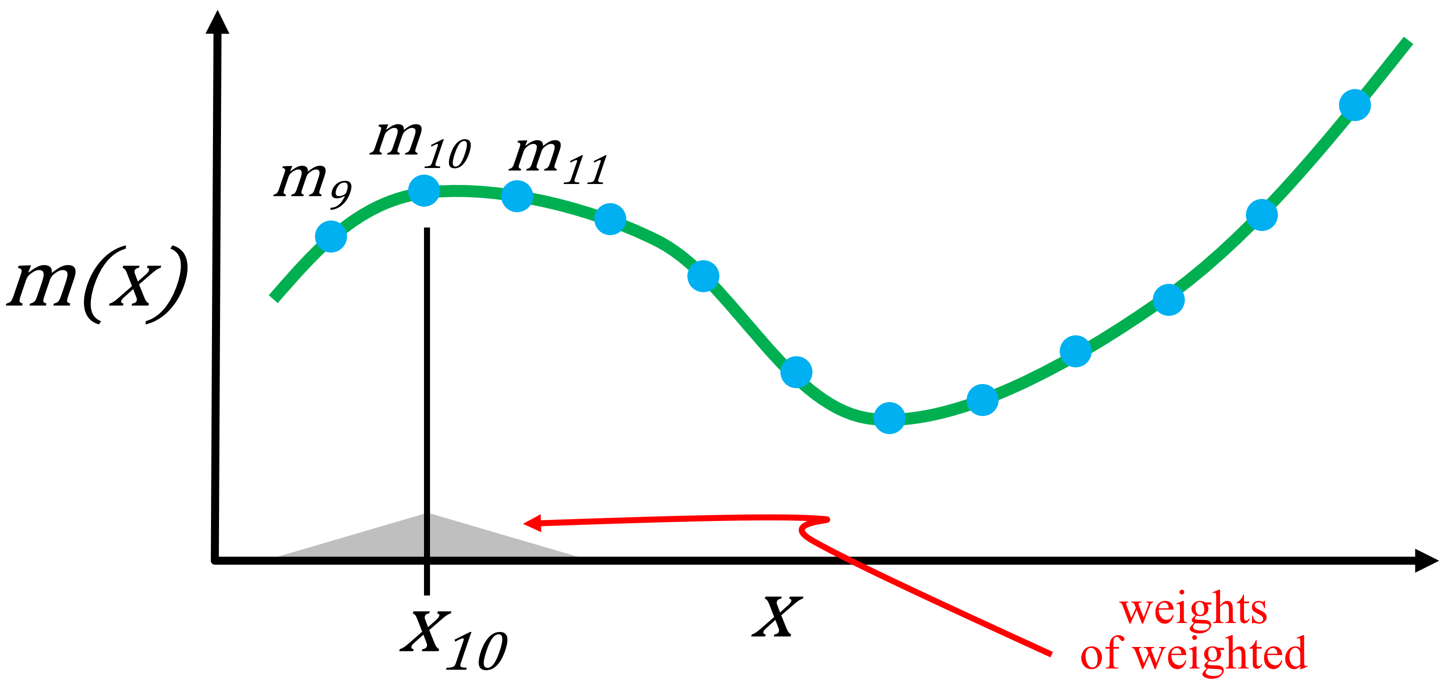
$$A = 0.2m_9 + 0.6m_{10} + 0.2m_{11}$$

weighted average of m(x)

in the vicinity of x_{10}



average "localized' in the vicinity of x_{10}



Localized average mean
can't determine $m(x)$ *at* $x=10$
but can determine
average value of $m(x)$ *near* $x=10$

Such a localized average might very well be useful

CHAPTER

2

Some Comments on Probability Theory

OUTLINE

2.1 Noise and Random Variables 17

2.2 Correlated Data 21

2.3 Functions of Random Variables 23

2.4 Gaussian Probability Density Functions 27

2.5 Testing the Assumption of Gaussian Statistics 30

2.6 Conditional Probability Density Functions 32

2.7 Confidence Intervals 34

2.8 Computing Realizations of Random Variables 35

2.9 Problems 37

Reference 37

Menke

Chapter 2 learning goals

Review random variables and their probability density functions

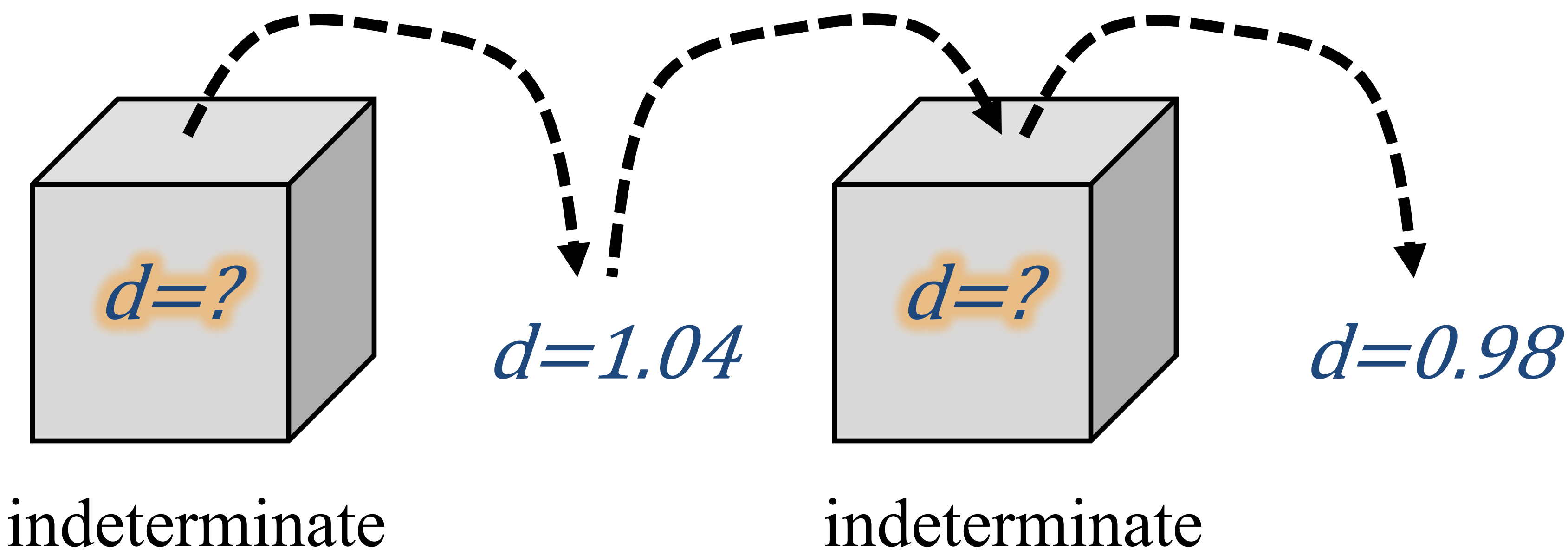
Introduce correlation and the multivariate Gaussian distribution

Relate error propagation to functions of random variables

random variables and their probability density functions

random variable, *d*

no fixed value until it is *realized*



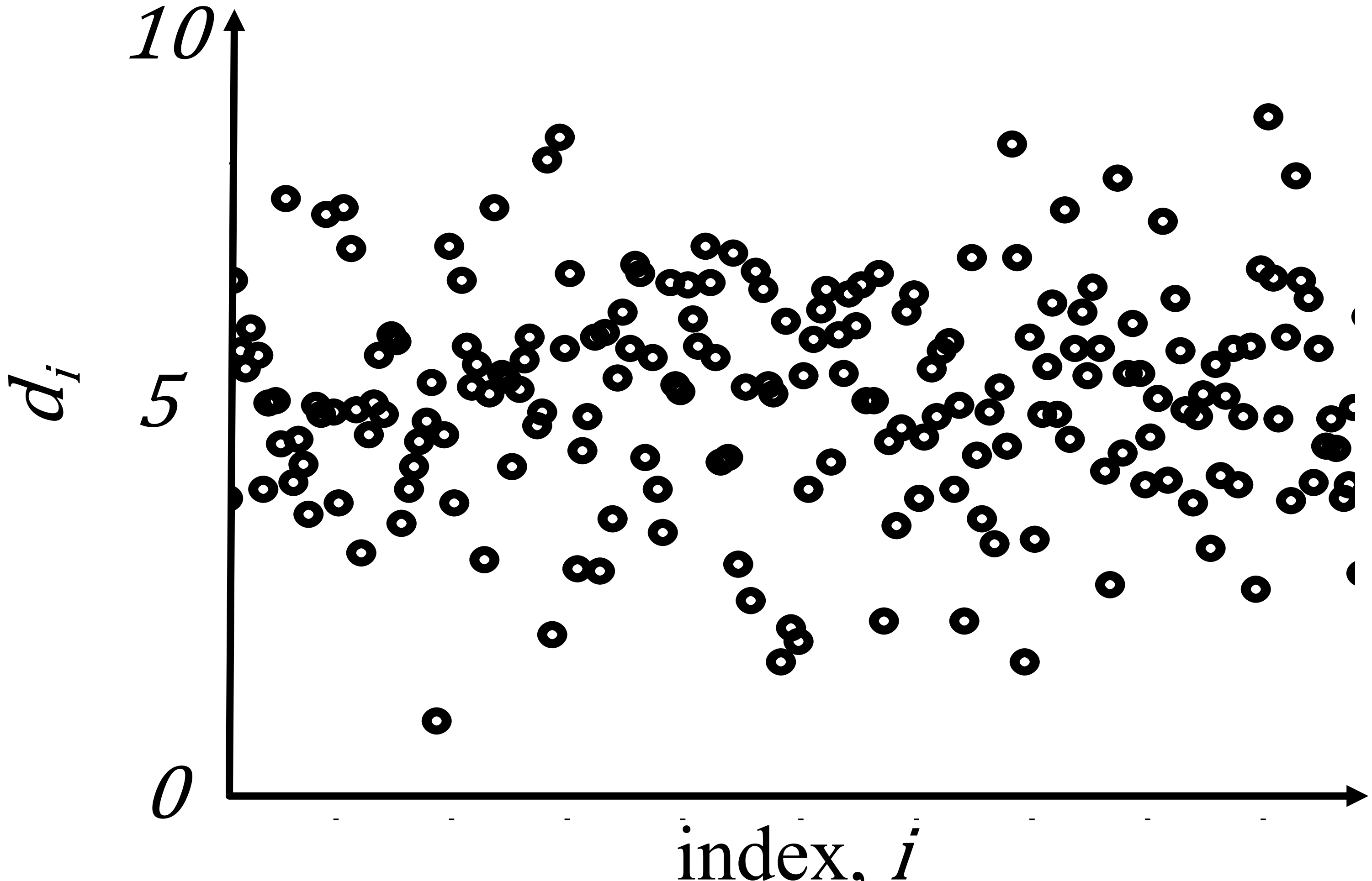
in this analogy, the random variable, d, has indeterminate value when it is in the box. Every time it is taken out of the box, it takes on a new value

random variables have systematics

tendency to take on some values more often than others

Even though realizations of a random variable are all different, the underlying variable has well-defined properties.

200 realizations of a random variable, d



probability density function (p.d.f.) describes the behavior of the random variable. It is the idealization of a histogram of the realizations, in the limit of an indefinitely large number of realizations.

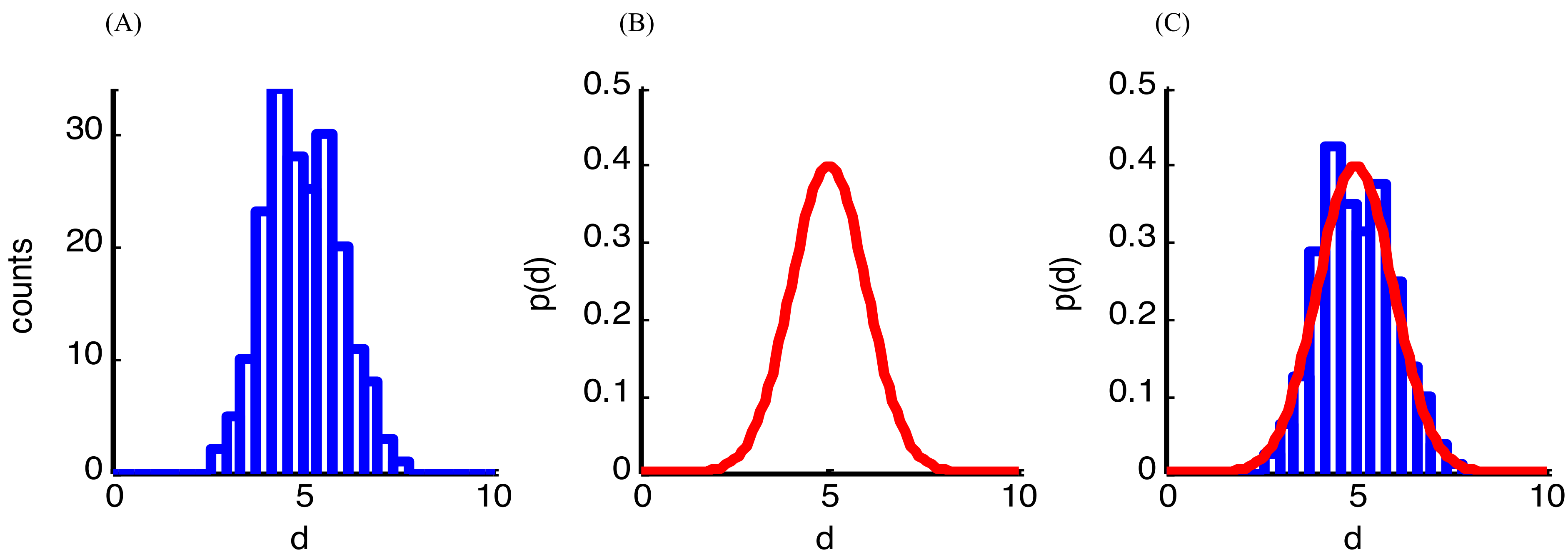


Fig 2.1. (A) Histogram showing data from 200 repetitions of an experiment in which datum, *d*, is measured. Noise causes observations to scatter about their mean value, *<d>=5*. (B) Probability density function (p.d.f), *p(d)*, of the data. (C) Histogram (blue) and p.d.f. (red) superimposed. Note that the histogram has a shape similar to the p.d.f.. *MatLab* script gda02_01.

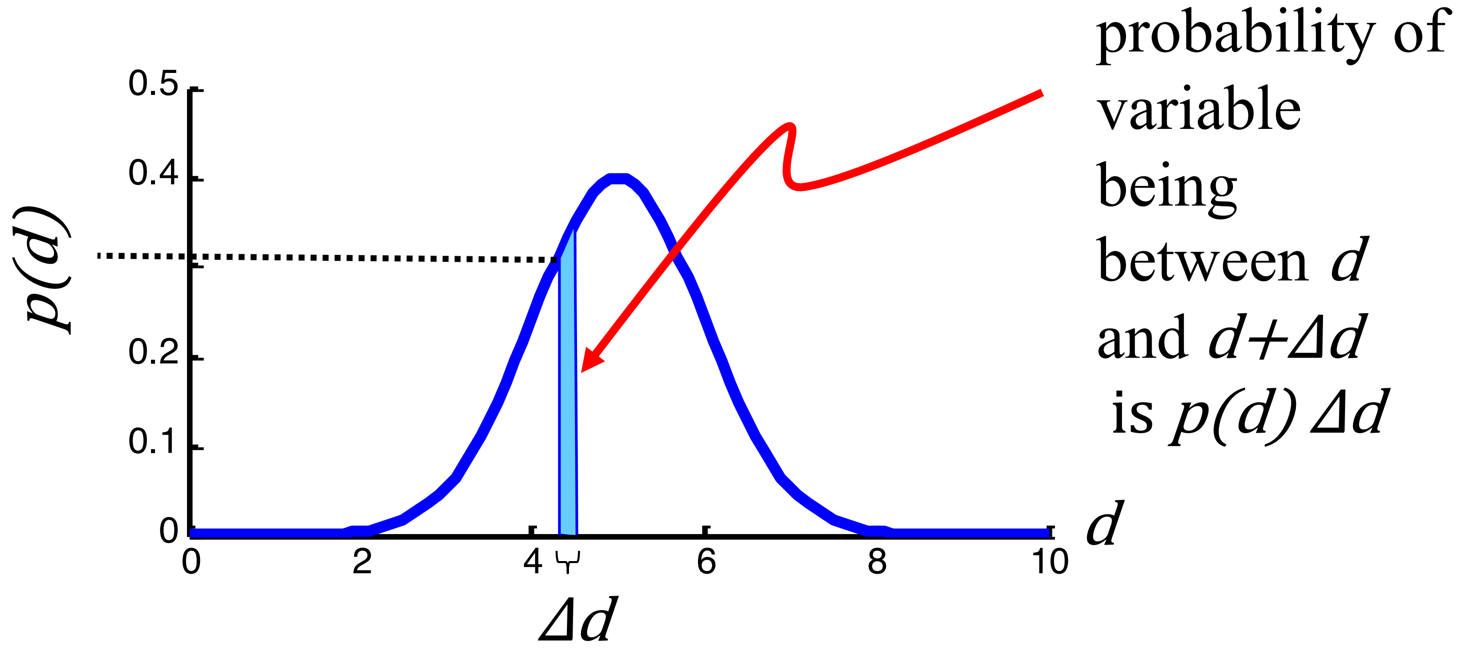
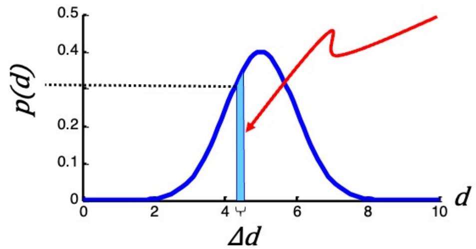


FIG. 2.2 The shaded area p(d)Δd of the probability density function p(d) gives the probability, P, that the observation will fall between d and d+Δd. MatLab script gda02_02.

in general
probability is the integral



$$P(d_1, d_2) = \int_{d_1}^{d_2} p(d)\, \mathrm{d}d$$

probability that *d* is between
d_1 and d_2

The probability P is a number between 0 and 1 (0% and 100%).
The probability that the random variable will take on a value between d_1 and d_2 is given by the intergral.

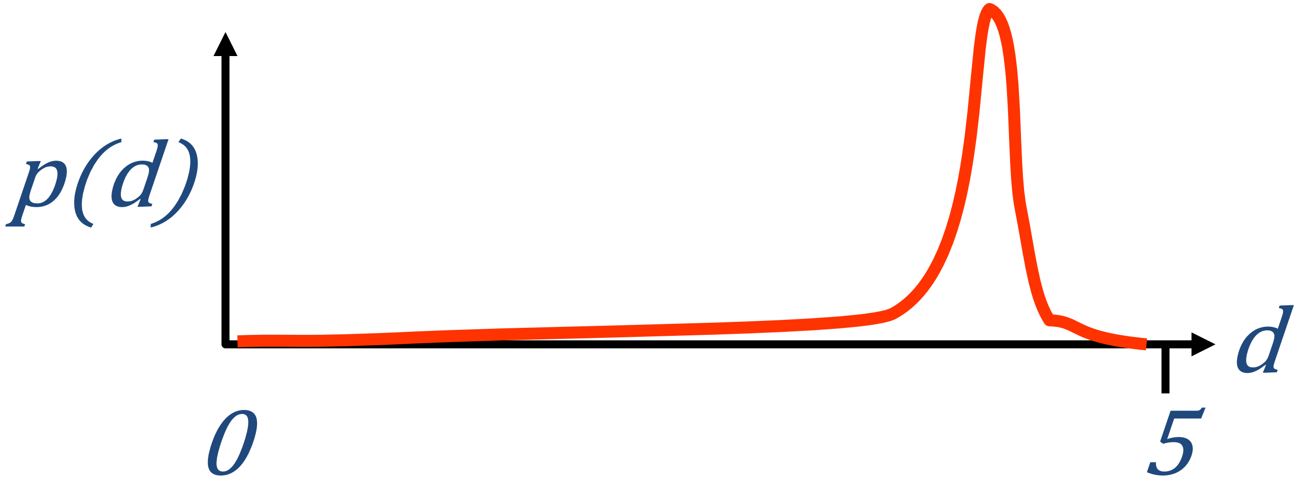
the probability that *d* has some value is 100% or unity

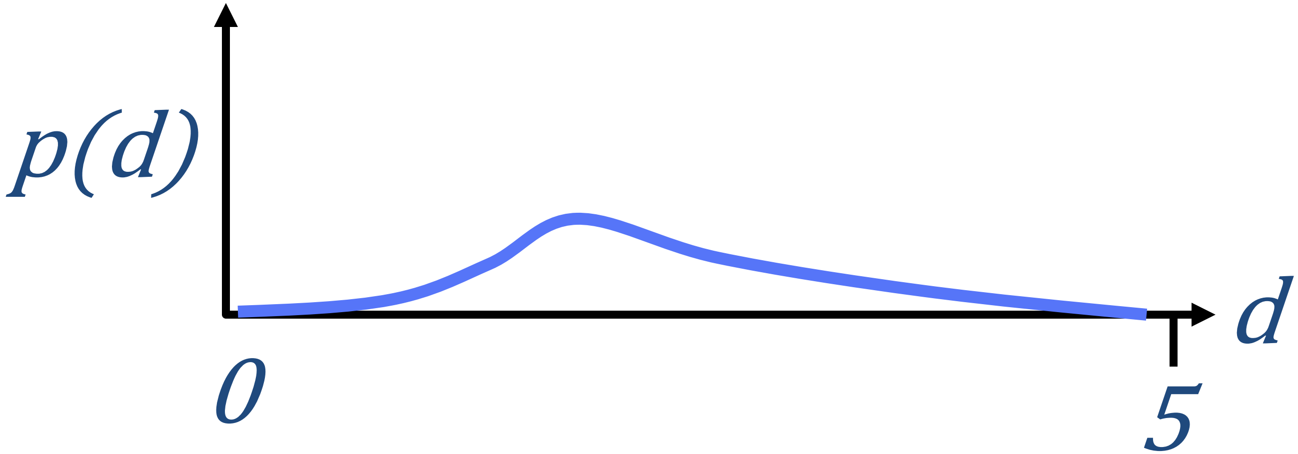
$$P(d_{min}, d_{max}) = \int_{d_{min}}^{d_{max}} p(d)\,\mathrm{d}d = 1$$

probability that *d* is between its minimum and maximum bounds, d_{min} and d_{max}

The random variable must take on some value, so the probability that it falls somewhere between its minimum and maximum bounds is unity (100%).

How do these two p.d.f.'s differ?

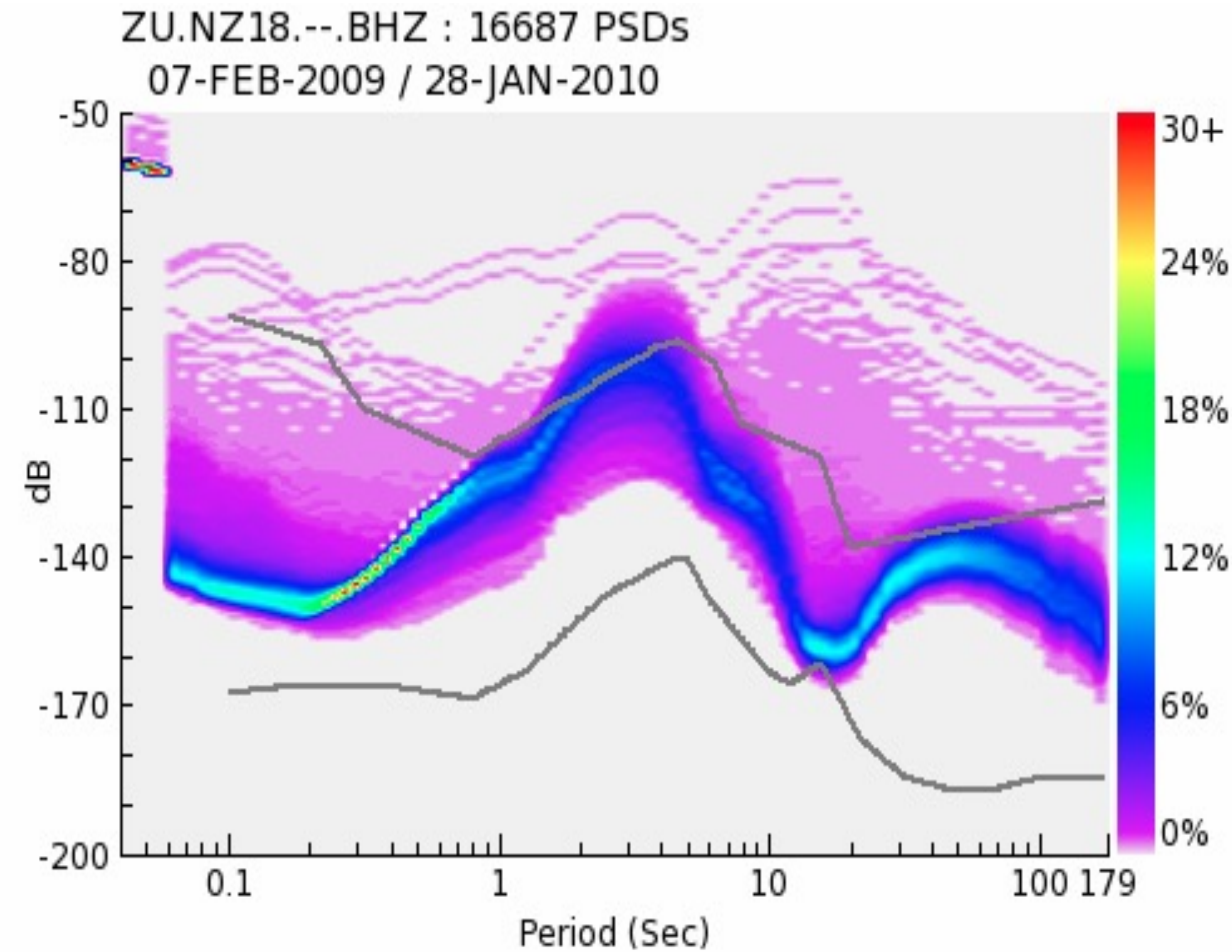





Let’s look at a p.d.f. of a year of data
from an ocean bottom seismometer

p.d.f. of a year of data at an ocean bottom seismometer

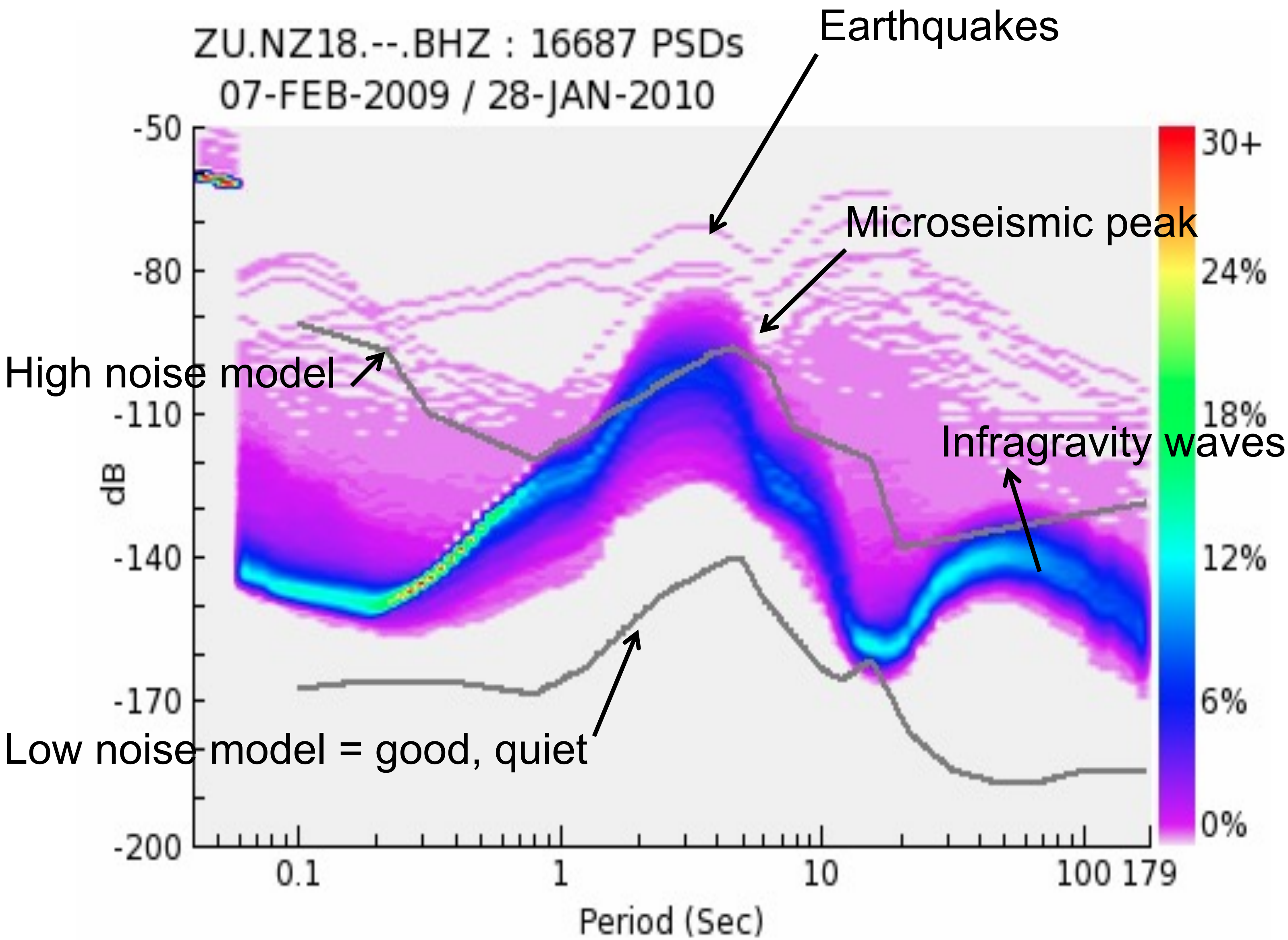
- Power spectra determined for one-hour segments of the entire year of data
- Probability density function used to display the thousands of spectra
- Blues and greens represent frequent values, light purple represents less frequent values



Background noise spectra

Power spectral density PSD

Probability density function PDF



Summarizing a probability density function

typical value

“center of the p.d.f.”

amount of scatter around the typical value

“width of the p.d.f.”

Several possibilities for a typical value

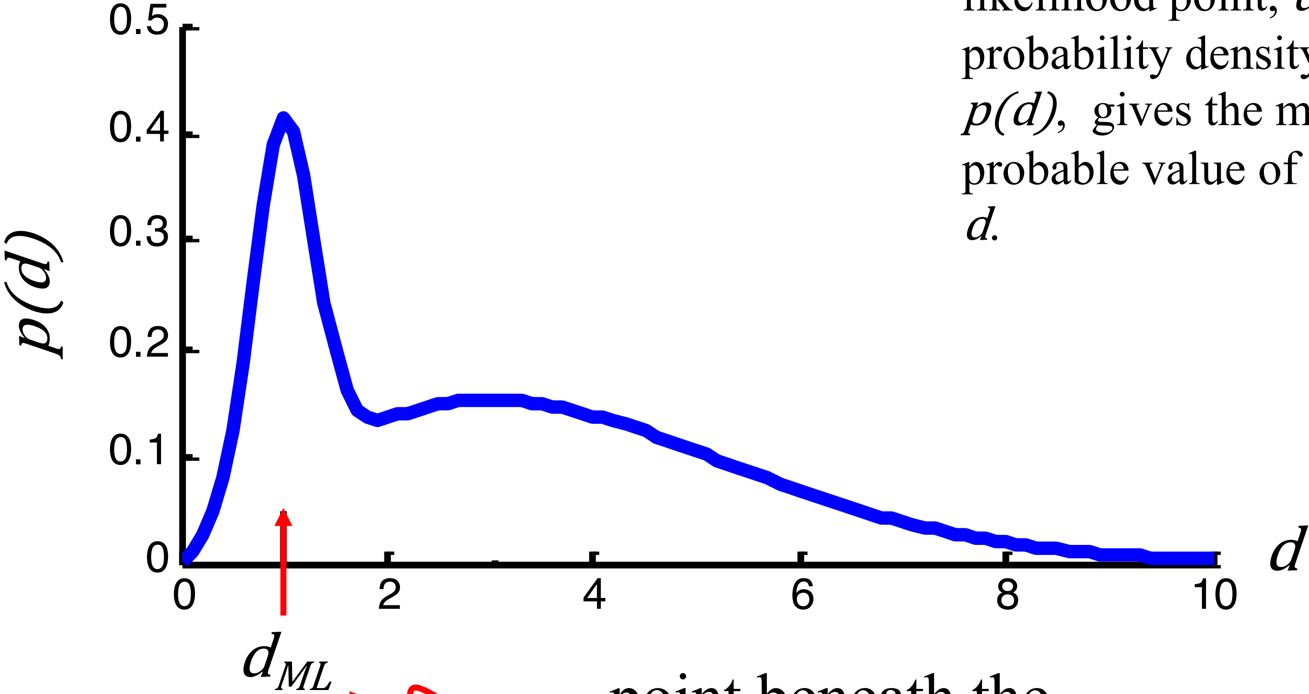


Fig. 2.3. The maximum likelihood point, d_{ML}, of the probability density function, $p(d)$, gives the most probable value of the datum, d.

point beneath the peak or “maximum likelihood point” or “mode”

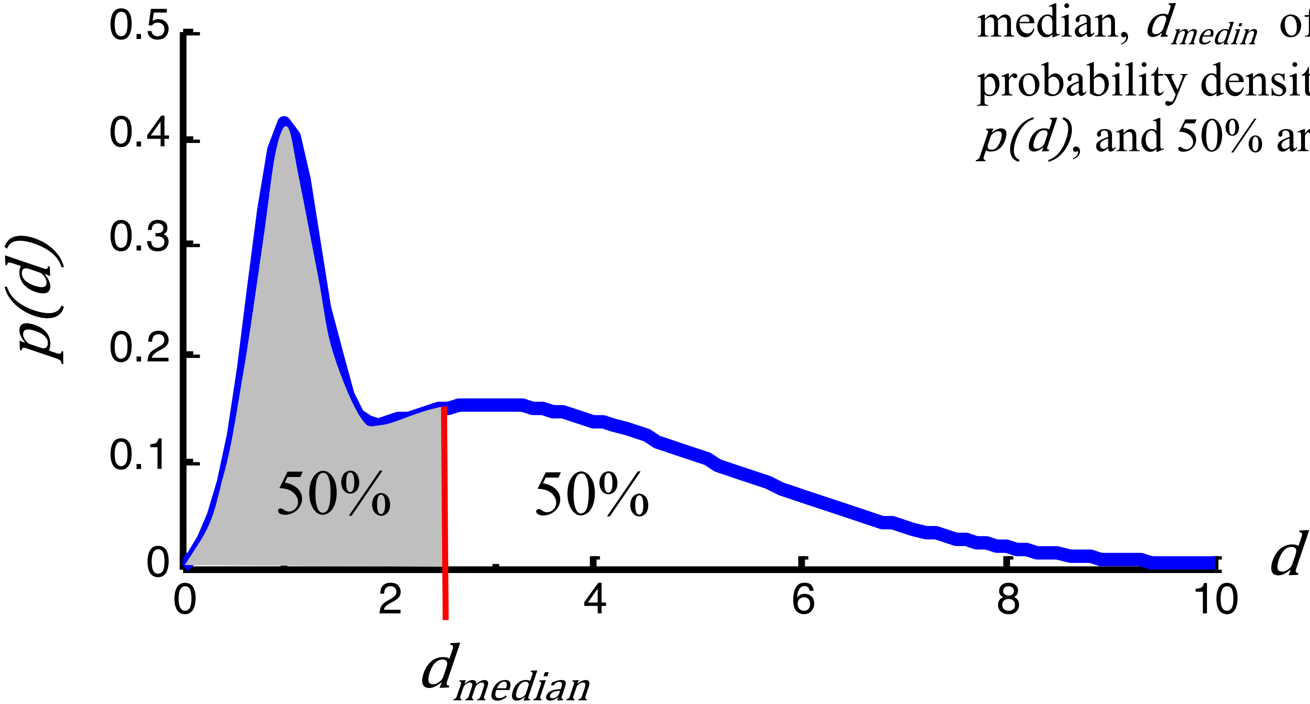
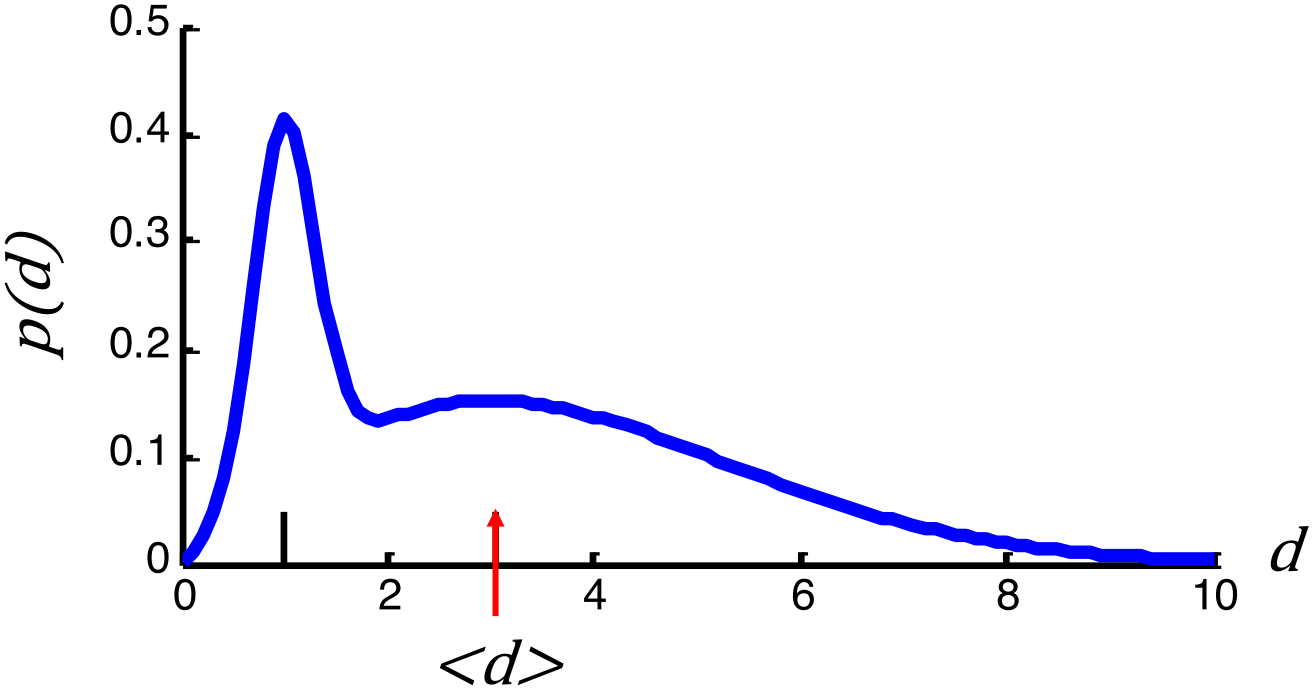


Fig. 2.3. Fifty percent of realization are less than the median, d_{medin} of the probability density function, *p(d)*, and 50% are greater.

point dividing area in half or "median"

Fig. 2.3. The mean or “expected” datum, *<d>*, is at the “balancing point” of the distribution.



balancing point or
“mean” or “expectation”

Mode, median, and mean can all be different

$$d_{ML} \neq d_{median} \neq \langle d \rangle$$

(for a symmetrical distribution, such as a Gaussian, the mean, median, and maximum likelihood values are all equal)

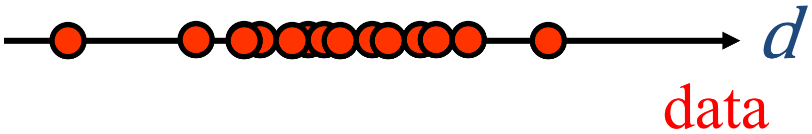
formula for
"mean" or "expected value"

$$\langle d \rangle = \int_{d_{min}}^{d_{max}} d\, p(d)\, \mathrm{d}d$$

To derive the integral formula for the mean:

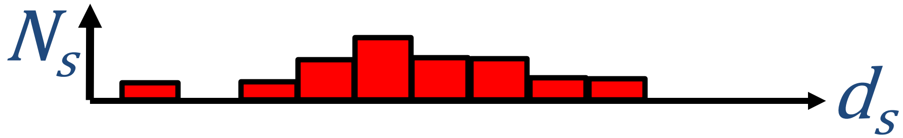
step 1: usual formula for mean

$$<d> = \frac{1}{N}\sum_{i=0}^{N} d_i$$



step 2: replace data with its histogram

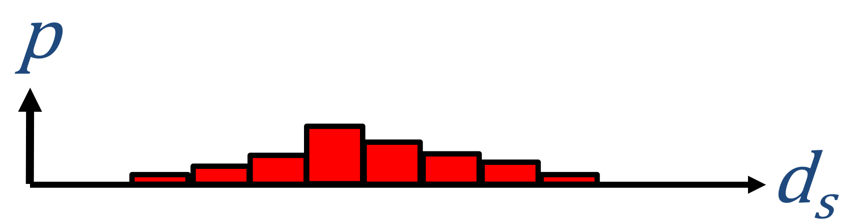
$$<d> \approx \frac{1}{N}\sum_{s=0}^{M} d^{(s)} N_s$$



step 3: replace histogram with probability distribution.

$$<d> \approx \sum_{s=0}^{M} d^{(s)} \frac{N_s}{N}$$

$$\approx \sum_{s=0}^{M} d^{(s)} P(d_s)$$



If the data are continuous, use formula containing an integral:

$$<d> \approx \sum_{s=0}^{M} d^{(s)} \; P(d_s)$$

$$<d> = \int_{d_{min}}^{d_{max}} d \; p(d) \; \mathrm{d}d$$

$$<d> \quad = \int_{d_{min}}^{d_{max}} d\, p(d)\, \mathrm{d}d$$

Example:

$f(x) = (5/4)(1-x^4)$ for x from 0 to 1
$f(x)=0$ all other x

Find expected (mean) value of x

$$<d> \quad = \int_{d_{min}}^{d_{max}} d\, p(d)\, \mathrm{d}d$$

Example:

f(x) = (5/4)(1-x^4)   for x from 0 to 1

f(x)=0 all other x

Expected (mean) value of x found such that

$$E(x) = \int_0^1 x f(x)\, dx$$

$$E(x) = \int_0^1 x(5/4)(1-x^4)\, dx$$

$$<d> \quad = \int_{d_{min}}^{d_{max}} d\, p(d)\, \mathrm{d}d$$

Example:

f(x) = (5/4)(1-x^4) for x from 0 to 1

f(x)=0 all other x

Expected (mean) value of x found such that

$$E(x) = \int_0^1 x f(x)\, dx$$

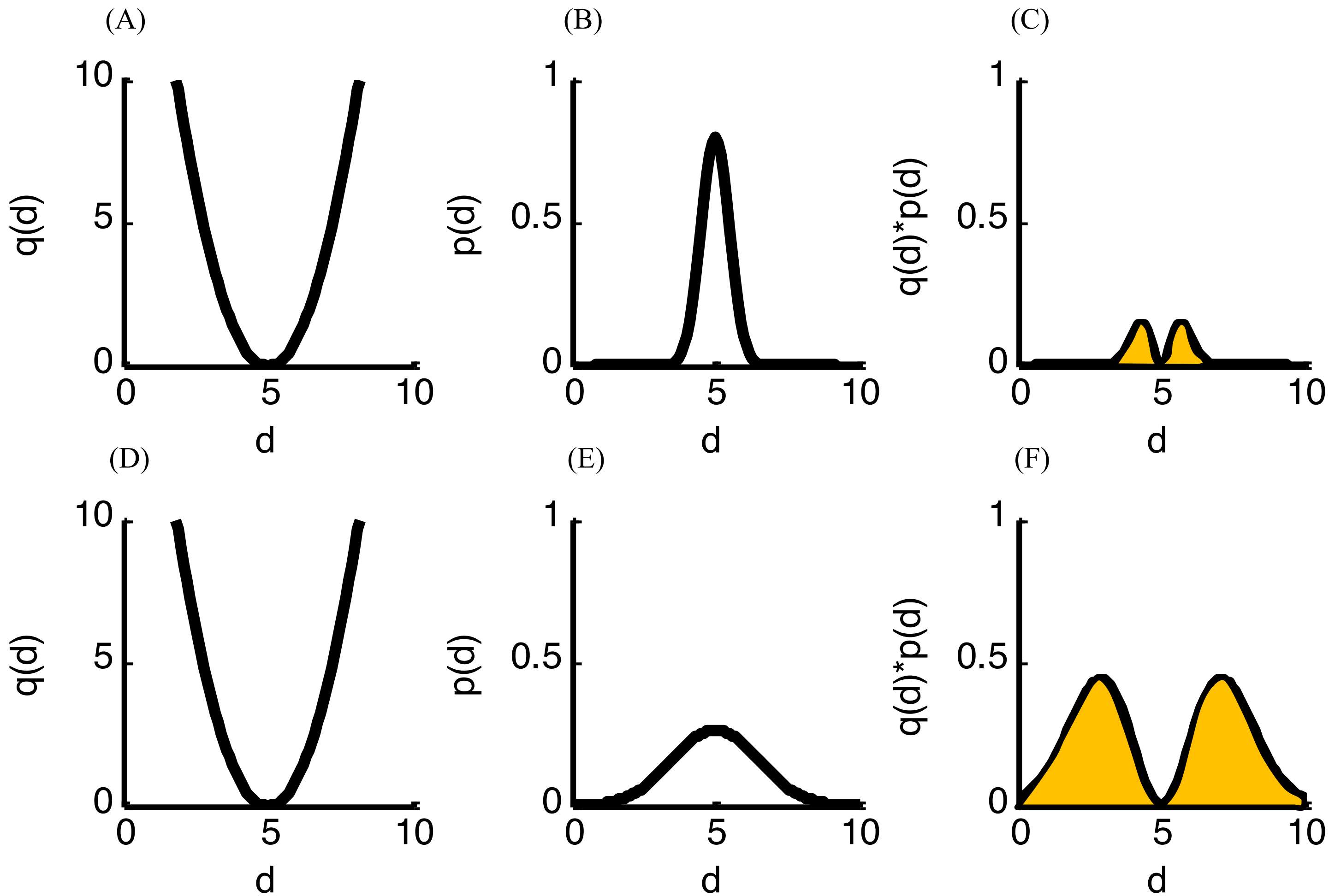
$$E(x) = \int_0^1 x(5/4)(1 - x^4)\, dx$$

$$E(x) = (5/4)\int_0^1 (x - x^5)\, dx$$

$$E(x) = (5/4)(x^2/2 - x^6/6)\Big|_0^1$$

$$E(x) = 5/12$$

Graphical representation of width



$q(d) = (d-\langle d\rangle)^2$

$p(d)$

$q(d)p(d)$

quantify width as the area under $q(d)p(d)$

Fig 2.4

This function grows away from the typical value:

$$q(d) = (d - \langle d \rangle)^2$$

so the function $q(d)p(d)$ is

small if most of the area is near $\langle d \rangle$, that is, a narrow $p(d)$

large if most of the area is far from $\langle d \rangle$, that is, a wide $p(d)$

so quantify width as the area under $q(d)p(d)$

‘deviation’ of a given measurement is difference between that measurement and the mean

deviation $= (d - \langle d \rangle)$

variance is average of the deviation squared

ie, plug $(d - \langle d \rangle)^2$ *in to*

$$= \int_{d_{min}}^{d_{max}} d\, p(d)\, \mathrm{d}d$$

variance

$$\sigma^2 = \int_{-\infty}^{+\infty} (d - \langle d \rangle)^2 \, p(d) \, \mathrm{d}d$$

mean

Standard deviation is σ and is a measure of the width of the distribution

2nd half of class today:
Python tutorial

optional
led by Shane Zhang, Physics/Geophysics

Next time

Homework 2 due (setting up problems into Gm=d form and recognizing whether the problem is linear or not. You do NOT have to solve the inverse problem, just set it up as Gm=d).

Chapter 2 continued, Probability

Python tutorial

you are encouraged to follow along running the jupyter notebook during the class:

https://github.com/shane-d-zhang/PythonTutorial6670

Python Tutorial for 6670 Geophysical Inverse Theory

If you are interested in attending the tutorial (not required), you are encouraged to download and run the jupyter notebook along the class.

To get started:

1. Install conda
2. Clone the repo

```
git clone git@github.com:shane-d-zhang/PythonTutorial6670.git
```

3. Create a conda environment

```
cd PythonTutorial6670
conda env create -f environment.yml
conda activate py6670
```

4. Run the notebook

```
jupyter notebook tutorial.ipynb
```