Lecture 10
Priors on Function Spaces: Gaussian Processes

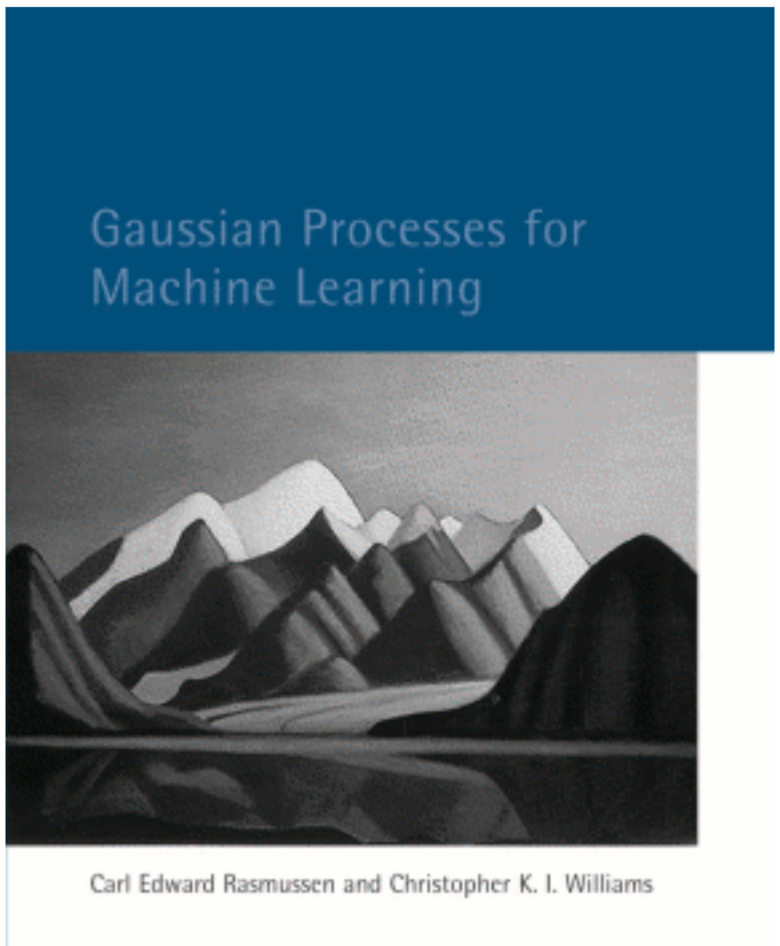
PREDICTIVE
SCIENCE LABORATORY

Objectives

- Express prior knowledge/beliefs about model outputs using Gaussian process (GP)
- Sample functions from the probability measure defined by GP
- Describe a GP with a finite number of variables: The Karhunen-Loeve expansion

PREDICTIVE
SCIENCE LABORATORY

The Best Book on the Subject



Gaussian Processes for Machine Learning
Carl Edward Rasmussen and Christopher K. I. Williams
The MIT Press, 2006. ISBN 0-262-18253-X.

Free online at www.gaussianprocess.org.
With Matlab code.

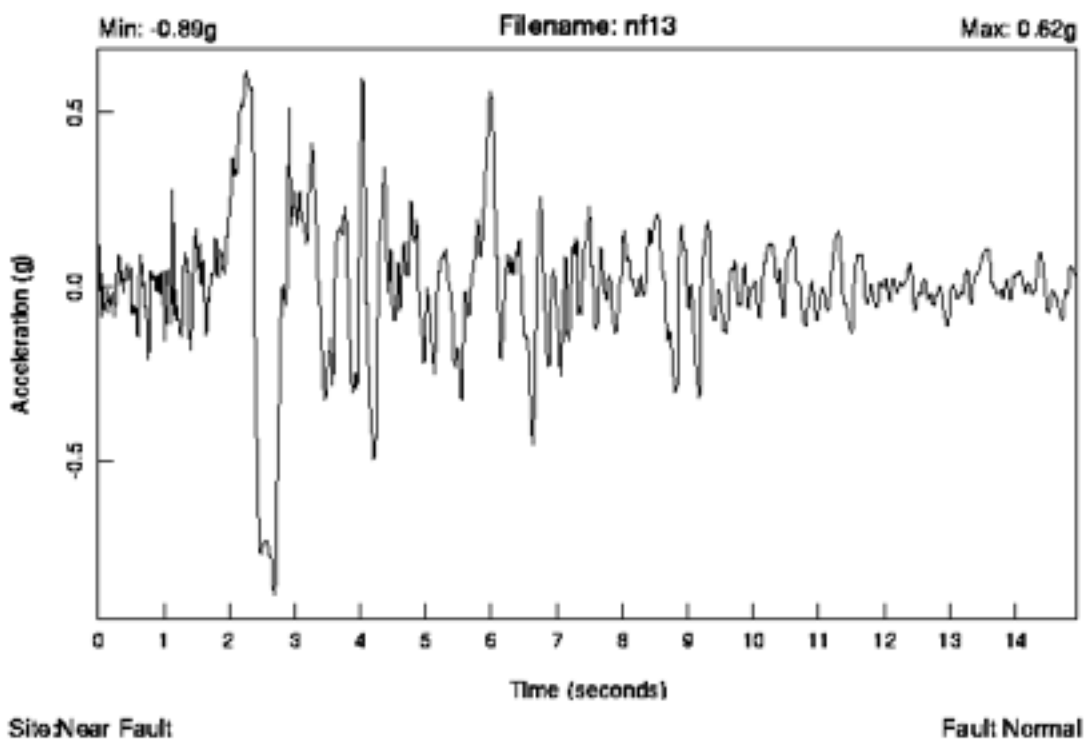
PREDICTIVE
SCIENCE LABORATORY

The Best Code on the Subject

GPy (in Python) from the group of N. Lawrence @ University of Sheffield
https://github.com/SheffieldML/GPy

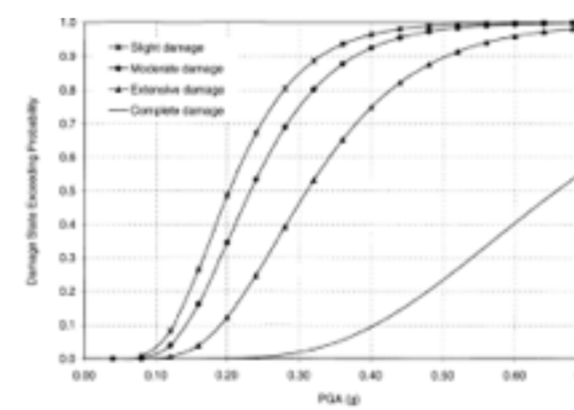
PREDICTIVE
SCIENCE LABORATORY

Why do we need probability measures on functions?



Uncertainty in external forcing

Simulation



Fragility Curve

PREDICTIVE
SCIENCE LABORATORY

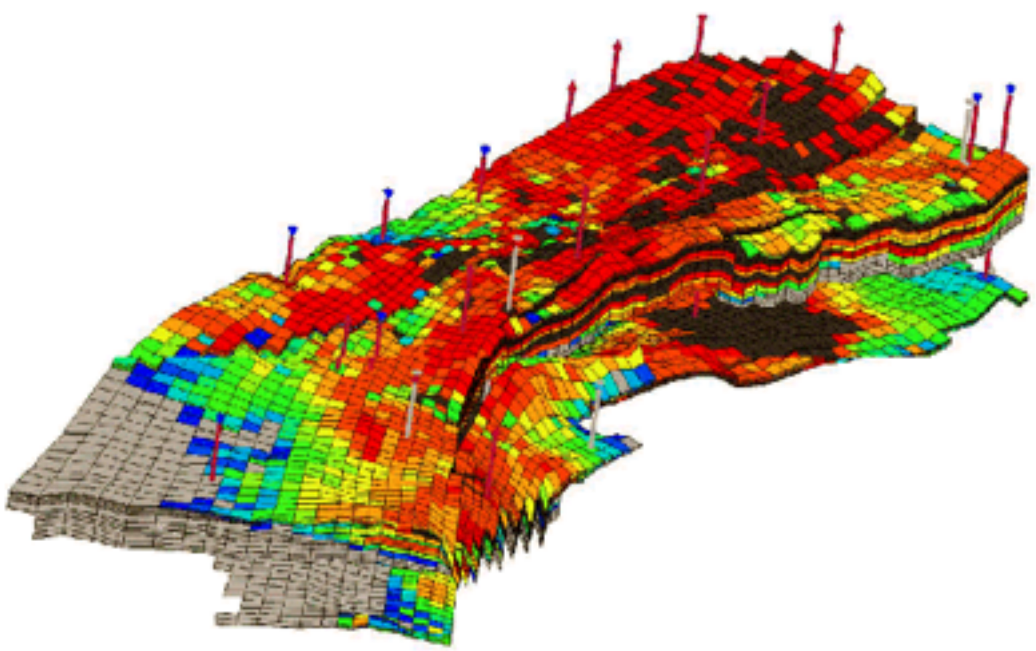
Why do we need probability measures on functions?

Portfolio Risk

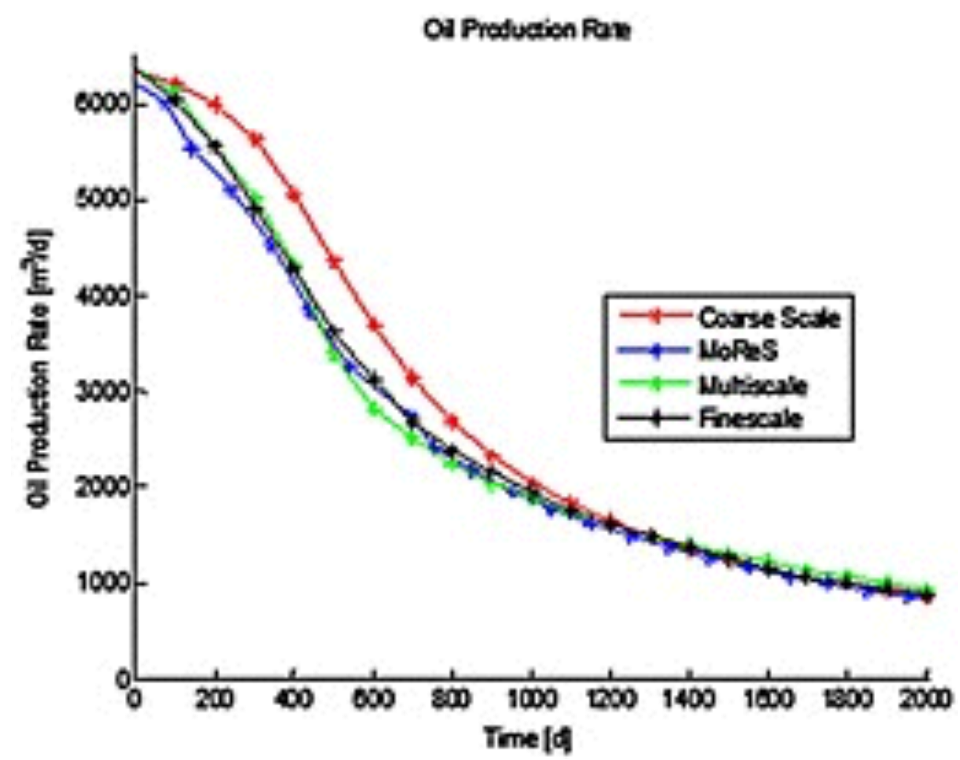
Uncertainty in external forcing

PREDICTIVE
SCIENCE LABORATORY

Why do we need probability measures on functions?



Simulation

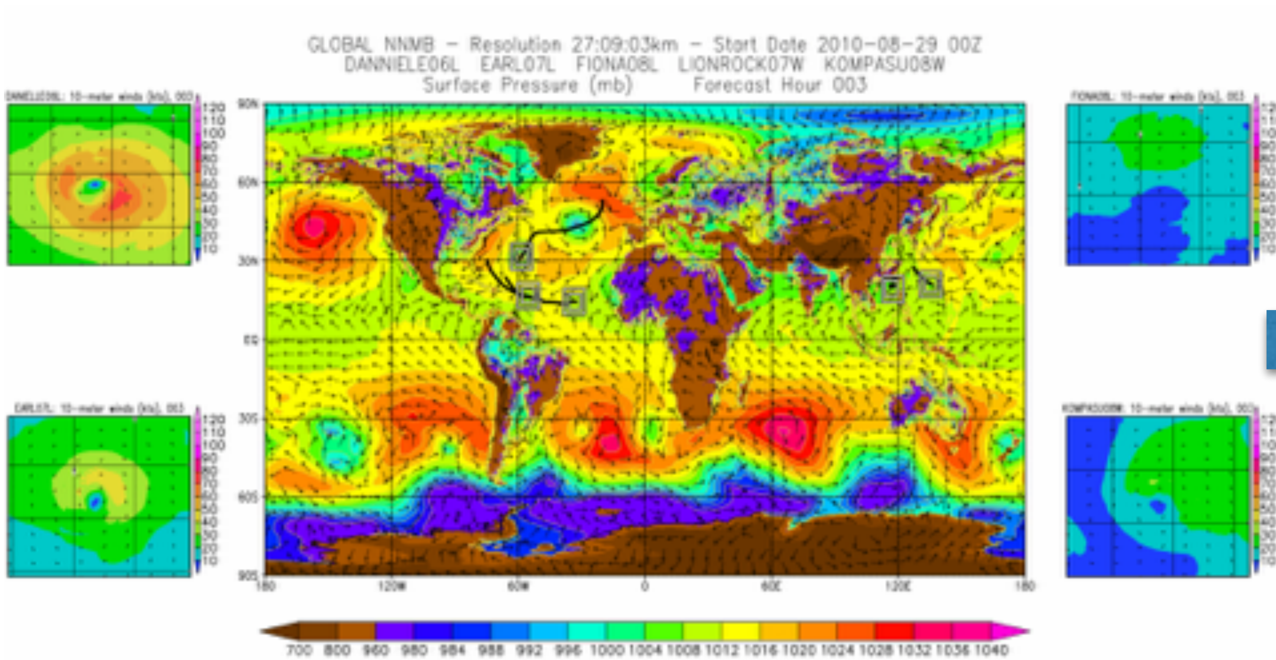


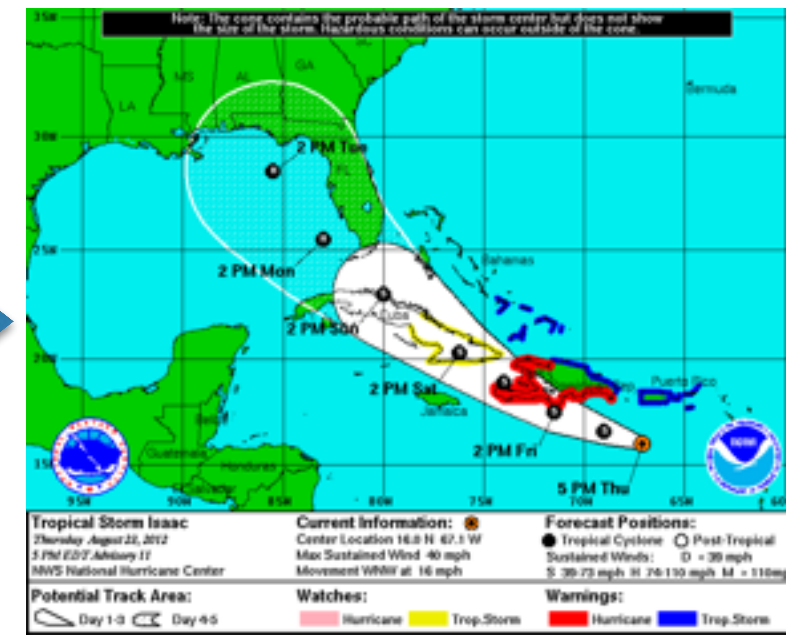
Oil produced over time

Uncertainty in field parameters

PREDICTIVE SCIENCE LABORATORY

Why do we need probability measures on functions?





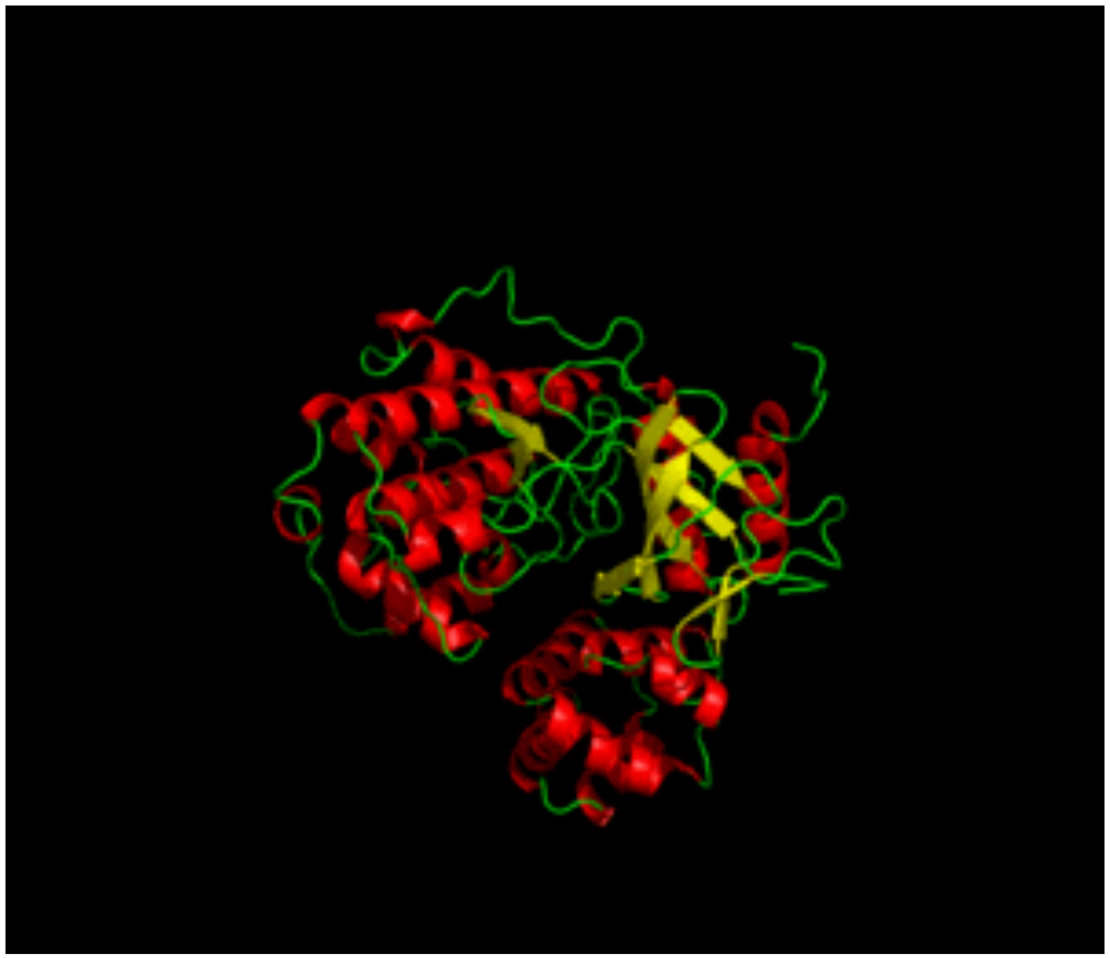
Hurricane path

Uncertainty in initial conditions

http://hwrf.aoml.noaa.gov/

PREDICTIVE
SCIENCE LABORATORY

Why do we need probability measures on functions?



Simulation of the interaction of two biomolecules

Statistical mechanics:

$$p(\mathbf{q}) \propto \exp\left\{-\frac{V(\mathbf{q})}{k_B T}\right\}$$

Temperature

Positions of all atoms

Boltzmann constant

Empirical potential. We are not exactly sure about its form…

PREDICTIVE SCIENCE LABORATORY

For quantifying uncertainties in:

- External forcing
- Field parameters
- Initial conditions in PDEs
- Boundary conditions in PDEs
- Physical laws (e.g., constitutive relations, empirical force fields, etc.)
- …

PREDICTIVE
SCIENCE LABORATORY

Probability measures on function spaces

- Simplest of all: Gaussian random field or Gaussian process.
- As usual, first we assign a prior, then we update with some data (next lecture).
- It turns out that you can also do regression with this (next lecture).

PREDICTIVE
SCIENCE LABORATORY

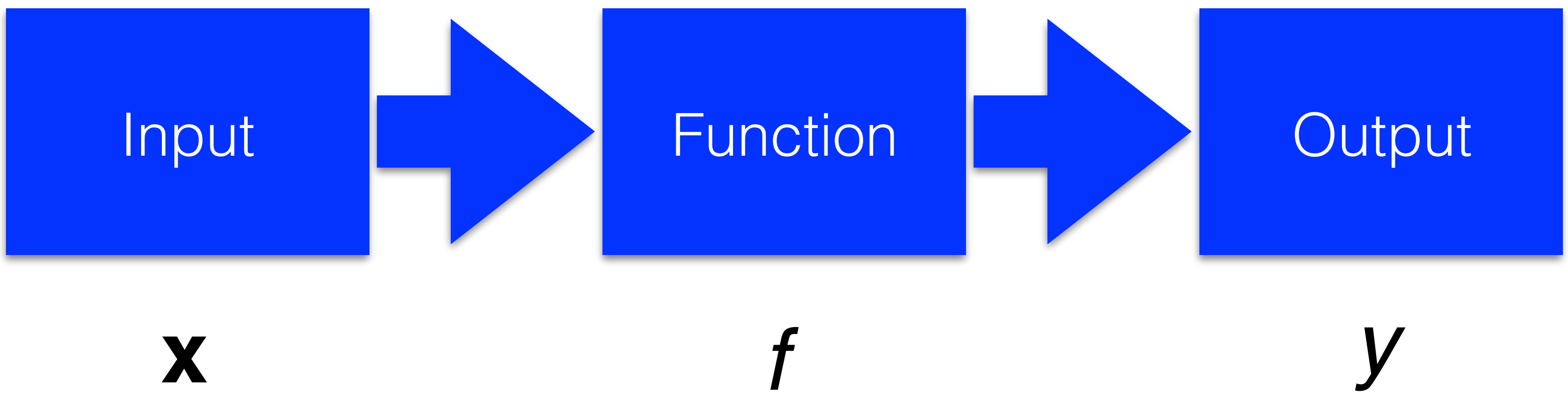
Definition of a Gaussian process

~~*A Gaussian process is a collection of "random" variables, any finite number of which have a joint Gaussian distribution.*~~

Let's just explain in plain English what it is…

PREDICTIVE SCIENCE LABORATORY

Definition of a Gaussian process



- Treat f as unknown
- Unknown = uncertain = “random”, i.e., described with probabilities
- Let us denote our beliefs about f as follows:

$$f(\cdot) \sim p(f(\cdot))$$

Definition of a Gaussian process

A Gaussian process needs two ingredients:

- a mean function
- a covariance function

It uses them to define a probability measure on the space of functions.

We write: $$f(\cdot) \sim p(f(\cdot)) = \mathrm{GP}\left(f(\cdot) \mid m(\cdot), k(\cdot,\cdot)\right)$$

PREDICTIVE
SCIENCE LABORATORY

The mean function

- What do you think *f(x)* could be?

- Define the *mean function* by:

$$m(x) = \mathbb{E}[f(x)]$$

- It models your *expectation* about *f(x).*

PREDICTIVE
SCIENCE LABORATORY

The covariance function

- How sure are you about this prediction?
- Consider the *variance*:

$$k(x,x) = \mathbb{E}\left[\left(f(x) - m(x)\right)^2\right]$$

- It models your *uncertainty* about *f(x).*

PREDICTIVE
SCIENCE LABORATORY

The covariance function

- Now, consider two inputs *x* and *x'*. How close do you think the corresponding outputs are?

- Consider the *covariance function*:

$$k(\mathbf{x},\mathbf{x}') = \mathbb{E}\left[\left(f(\mathbf{x}) - m(\mathbf{x})\right)\left(f(\mathbf{x}') - m(\mathbf{x}')\right)\right]$$

- It models yours beliefs about the similarity of *f(x)* and *f(x').*

PREDICTIVE
SCIENCE LABORATORY

To wrap it up

- We write:

$$f(\cdot) \sim \mathrm{GP}\left(f(\cdot) \mid m(\cdot), k(\cdot,\cdot)\right)$$

- and we interpret:
  - $m(x)$: What do I think $f(x)$ could be?
  - $k(x, x)$: How sure am I about my expectation of $f(x)$?
  - $k(x, x')$: How similar are $f(x)$ and $f(x')$?

PREDICTIVE
SCIENCE LABORATORY

The most common covariance function: Squared Exponential (SE)

- Also known as radial basis function (RBF).

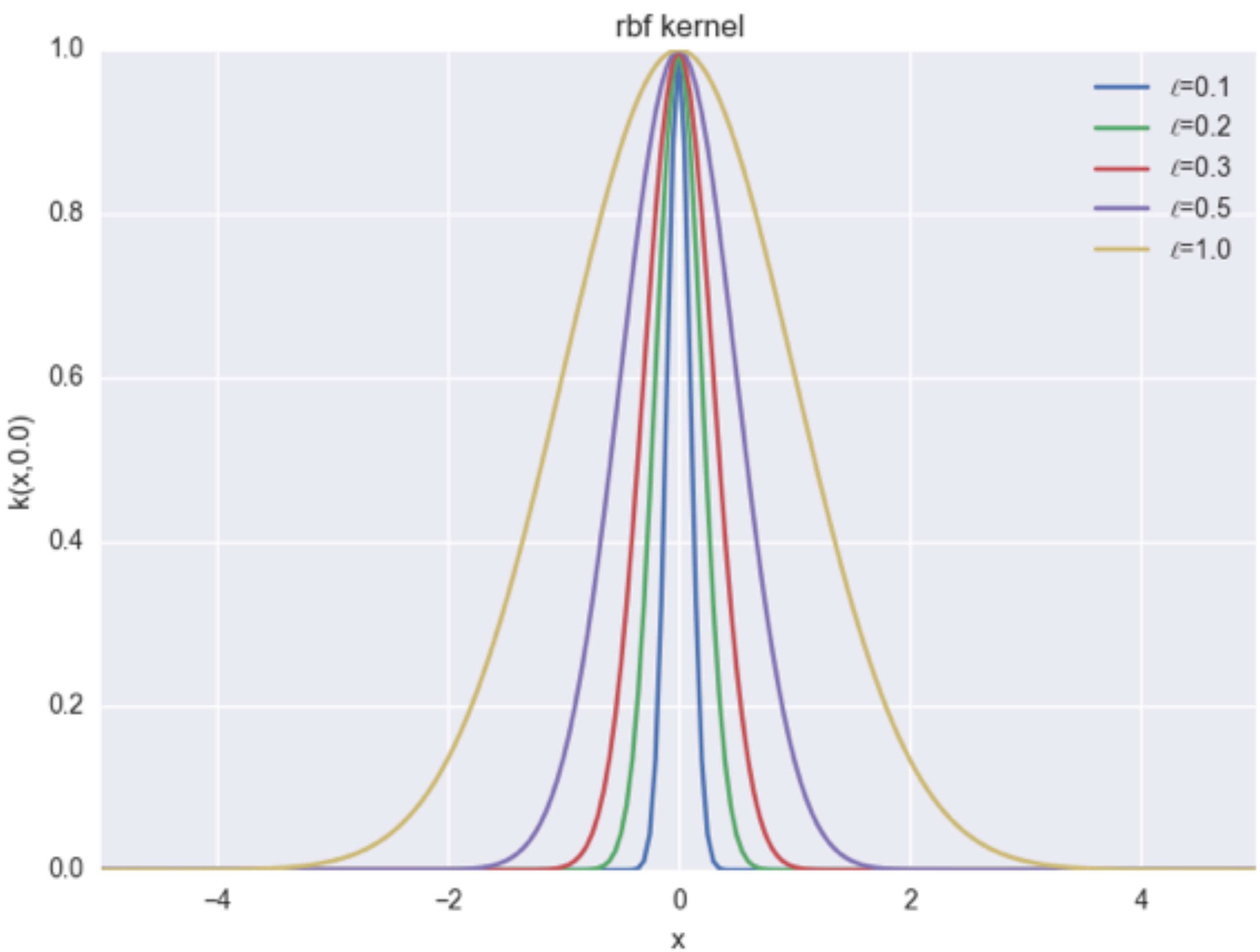
$$k(\mathbf{x},\mathbf{x}') = v \exp\left\{-\frac{1}{2}\sum_{i=1}^{d}\frac{(x_i - x_i')^2}{\ell_i^2}\right\}.$$

Variance
models uncertainty
about *f(x)*

Length-scale
models similarity
of specific input dimensions

PREDICTIVE
SCIENCE LABORATORY

Example 1.1: Drawing covariance functions



You have 10 minutes

PREDICTIVE
SCIENCE LABORATORY

The covariance matrix

- Consider an arbitrary selection of input points and their corresponding outputs:

$$\mathbf{x}_{1:n} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \rightarrow \mathbf{f} = \{f(\mathbf{x}_1), \dots, f(\mathbf{x}_n)\}$$

- The *covariance matrix* is defined to be:

$$\mathbb{E}\left[(\mathbf{f}-\mathbf{m})(\mathbf{f}-\mathbf{m})^T\right] := \mathbf{K} := \begin{pmatrix} k(\mathbf{x}_1,\mathbf{x}_1) & \dots & k(\mathbf{x}_1,\mathbf{x}_n) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_n,\mathbf{x}_1) & \dots & k(\mathbf{x}_n,\mathbf{x}_n) \end{pmatrix}$$

PREDICTIVE
SCIENCE LABORATORY

Restrictions on the covariance functions

- The covariance function has to be *positive definite*.
- That is, for any finite collection of inputs, the covariance matrix must be *positive definite*:

$$\mathbf{K} := \begin{pmatrix} k(\mathbf{x}_1,\mathbf{x}_1) & \dots & k(\mathbf{x}_1,\mathbf{x}_n) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_n,\mathbf{x}_1) & \dots & k(\mathbf{x}_n,\mathbf{x}_n) \end{pmatrix}$$

PREDICTIVE
SCIENCE LABORATORY

Covariance function factory

- The sum of two covariance functions is a covariance function.

$$k(\mathbf{x},\mathbf{x}') = k_1(\mathbf{x},\mathbf{x}') + k_2(\mathbf{x},\mathbf{x}')$$

- What does this model?

- The belief that the response comes from two sources:

$$f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})$$

$$f_i(\cdot) \sim \mathrm{GP}\left(f_i(\cdot) \mid m_i(\cdot), k_i(\cdot,\cdot)\right), i = 1,2$$

PREDICTIVE
SCIENCE LABORATORY

Covariance function factory

- The product of two covariance functions is a covariance function.

$$k(\mathbf{x},\mathbf{x}') = k_1(\mathbf{x},\mathbf{x}')k_2(\mathbf{x},\mathbf{x}')$$

- What does this model?

- The belief that the response comes from two sources:

$$f(\mathbf{x}) = f_1(\mathbf{x})f_2(\mathbf{x})$$

$$f_i(\cdot) \sim \mathrm{GP}\left(f_i(\cdot) \mid m_i(\cdot), k_i(\cdot,\cdot)\right), i = 1,2$$

PREDICTIVE
SCIENCE LABORATORY

The covariance matrix

- Consider an arbitrary selection of input points and their corresponding outputs:

$$\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \rightarrow \mathbf{f} = \{f(\mathbf{x}_1), \dots, f(\mathbf{x}_n)\}$$

- The *covariance matrix* is defined to be:

$$\mathbb{E}\left[(\mathbf{f}-\mathbf{m})(\mathbf{f}-\mathbf{m})^T\right] := \mathbf{K} := \begin{pmatrix} k(\mathbf{x}_1,\mathbf{x}_1) & \dots & k(\mathbf{x}_1,\mathbf{x}_n) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_n,\mathbf{x}_1) & \dots & k(\mathbf{x}_n,\mathbf{x}_n) \end{pmatrix}$$

PREDICTIVE
SCIENCE LABORATORY

Example 2: The covariance matrix and some properties of covariance functions

You have 10 minutes

PREDICTIVE
SCIENCE LABORATORY

Sampling a Gaussian process

- A Gaussian process defines a probability measure over a function space:

$$f(\cdot) \sim \mathrm{GP}\left(f(\cdot) \mid m(\cdot), k(\cdot,\cdot)\right)$$

- How can we sample functions from it?

- Sample *f* at a finite, albeit large, set of inputs.

PREDICTIVE
SCIENCE LABORATORY

Sampling a Gaussian process

- Take a *finite* number of inputs:

$$\mathbf{x}_{1:n} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \rightarrow \mathbf{f} = \{f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)\}$$

- and consider the model output on them:

$$\mathbf{f} = \{f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)\}$$

- We *believe* that they are distributed according to:

$$\mathbf{f} \sim \mathcal{N}\left(\mathbf{f} \mid \mathbf{m}, \mathbf{K}\right)$$

PREDICTIVE
SCIENCE LABORATORY

Sampling a Gaussian process

- Ok, so we need to be able to sample from this:

$$\mathbf{f} \sim \mathcal{N}\left(\mathbf{f} \mid \mathbf{m}, \mathbf{K}\right)$$

- with

$$\mathbf{m} = \begin{pmatrix} m(\mathbf{x}_1) \\ \vdots \\ m(\mathbf{x}_n) \end{pmatrix}, \mathbf{K} := \begin{pmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \dots & k(\mathbf{x}_1, \mathbf{x}_n) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_n, \mathbf{x}_1) & \dots & k(\mathbf{x}_n, \mathbf{x}_n) \end{pmatrix}.$$

Sampling a Gaussian process

- To sample from:

$$\mathbf{f} \sim \mathcal{N}\left(\mathbf{f} \mid \mathbf{m}, \mathbf{K}\right)$$

- Take the lower Cholesky decomposition **L** of **K**:

$$\mathbf{K} = \mathbf{L}\mathbf{L}^{\mathrm{T}}$$

- Sample a standard normal:

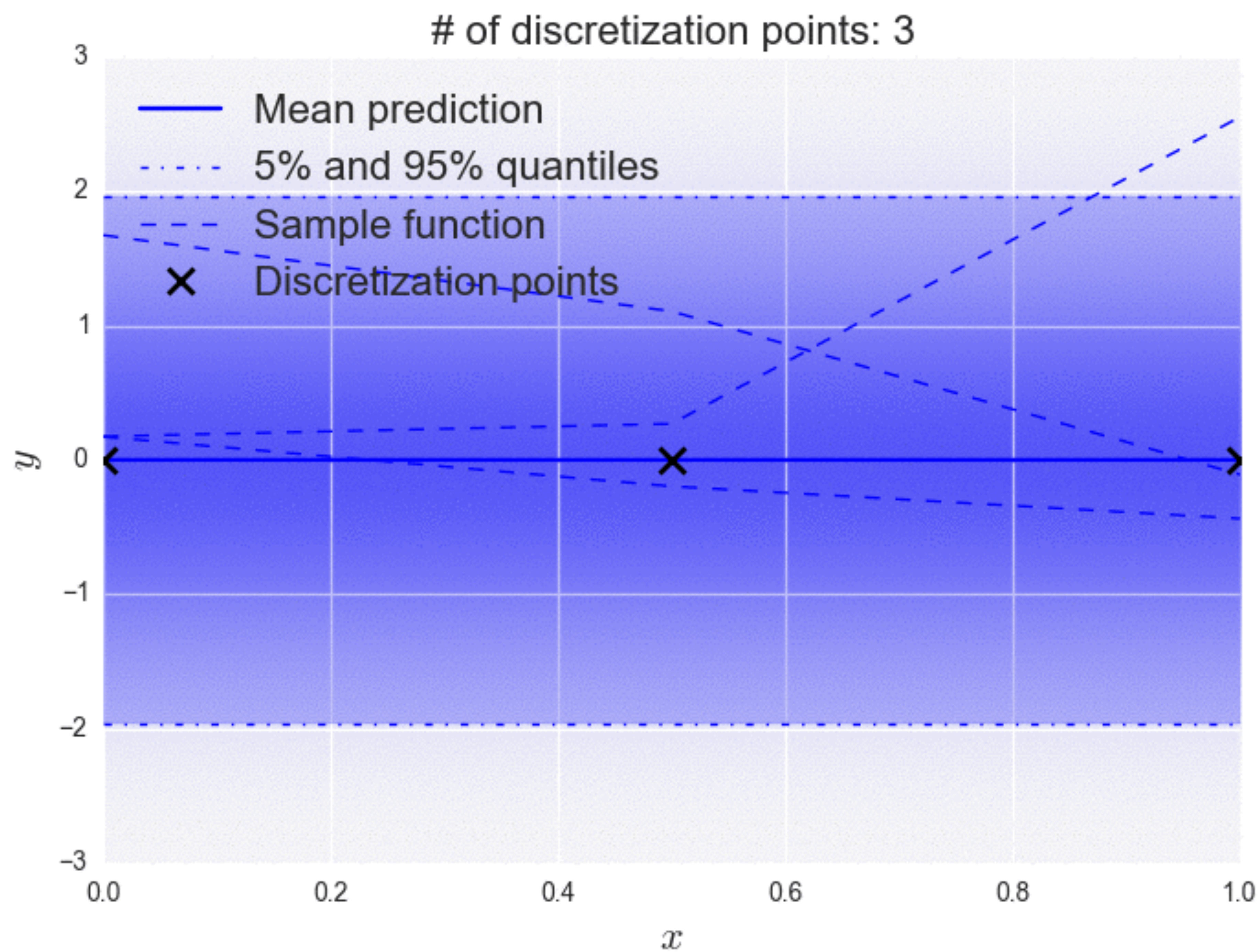
$$\mathbf{z} \sim \mathcal{N}\left(\mathbf{z} \mid \mathbf{0}_n, \mathbf{I}_n\right)$$

- and set:

$$\mathbf{f} = \mathbf{m} + \mathbf{L}\mathbf{z}$$

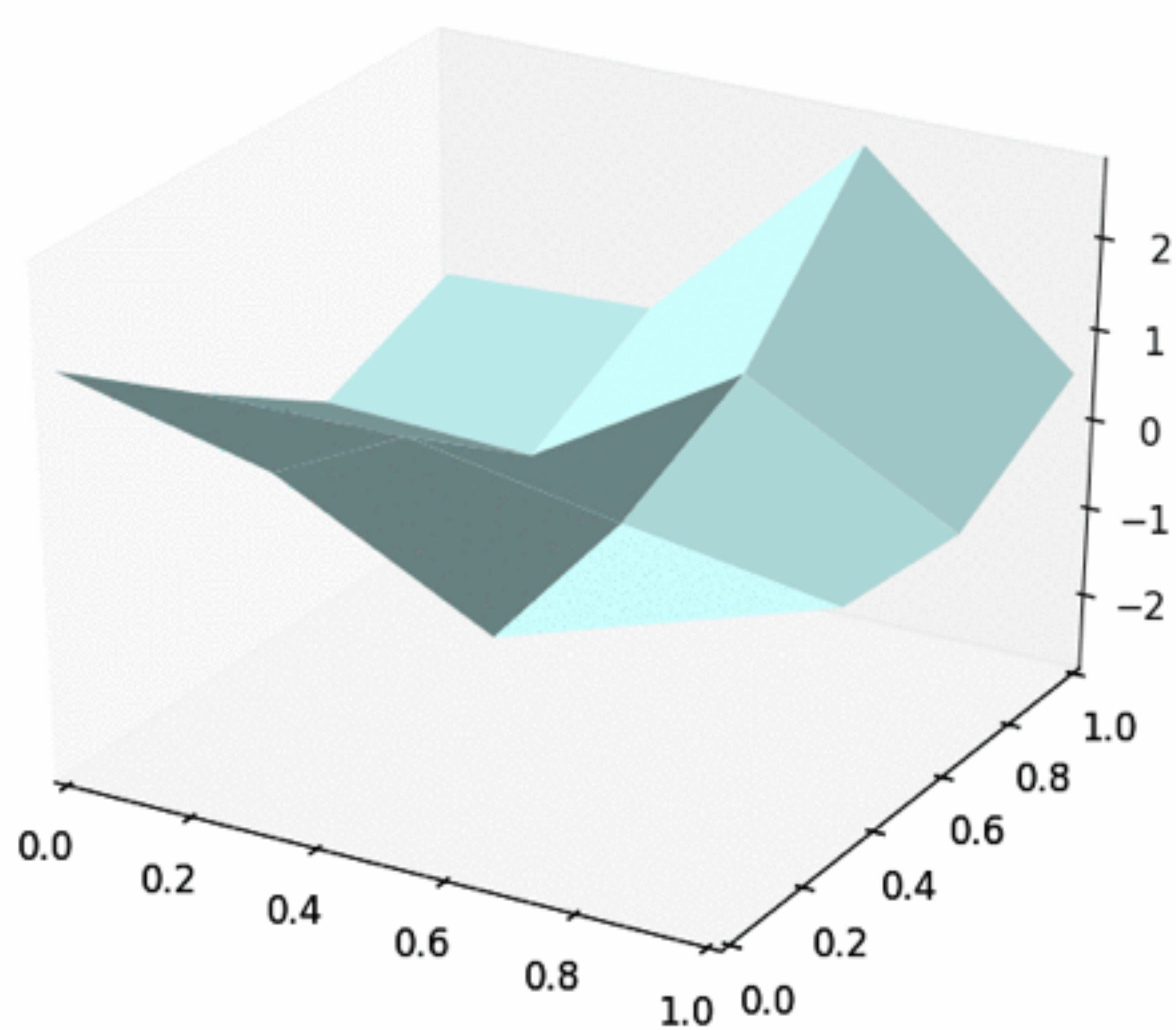
PREDICTIVE
SCIENCE LABORATORY

Sampling from a Gaussian



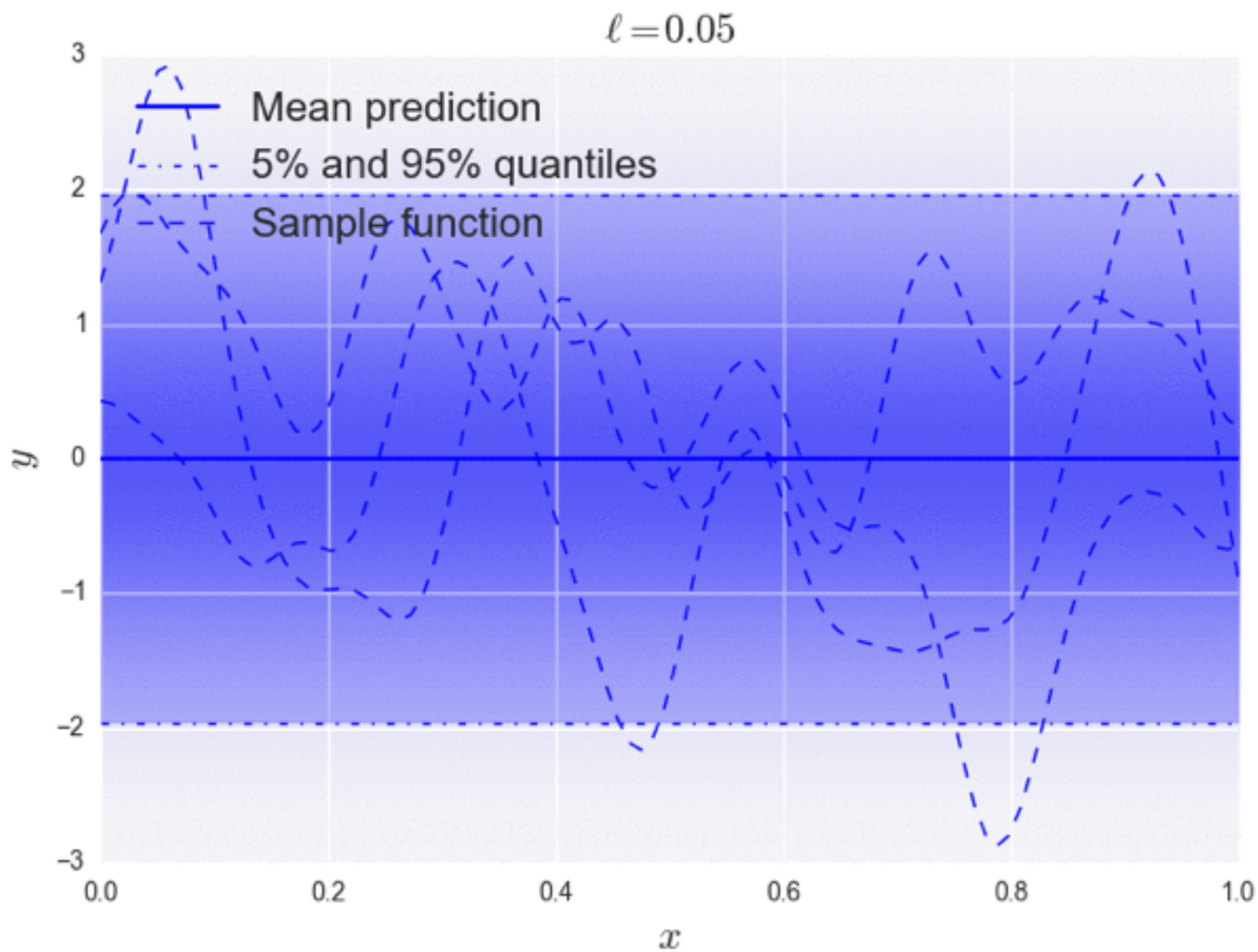
PREDICTIVE
SCIENCE LABORATORY

Sampling from a Gaussian process



PREDICTIVE
SCIENCE LABORATORY

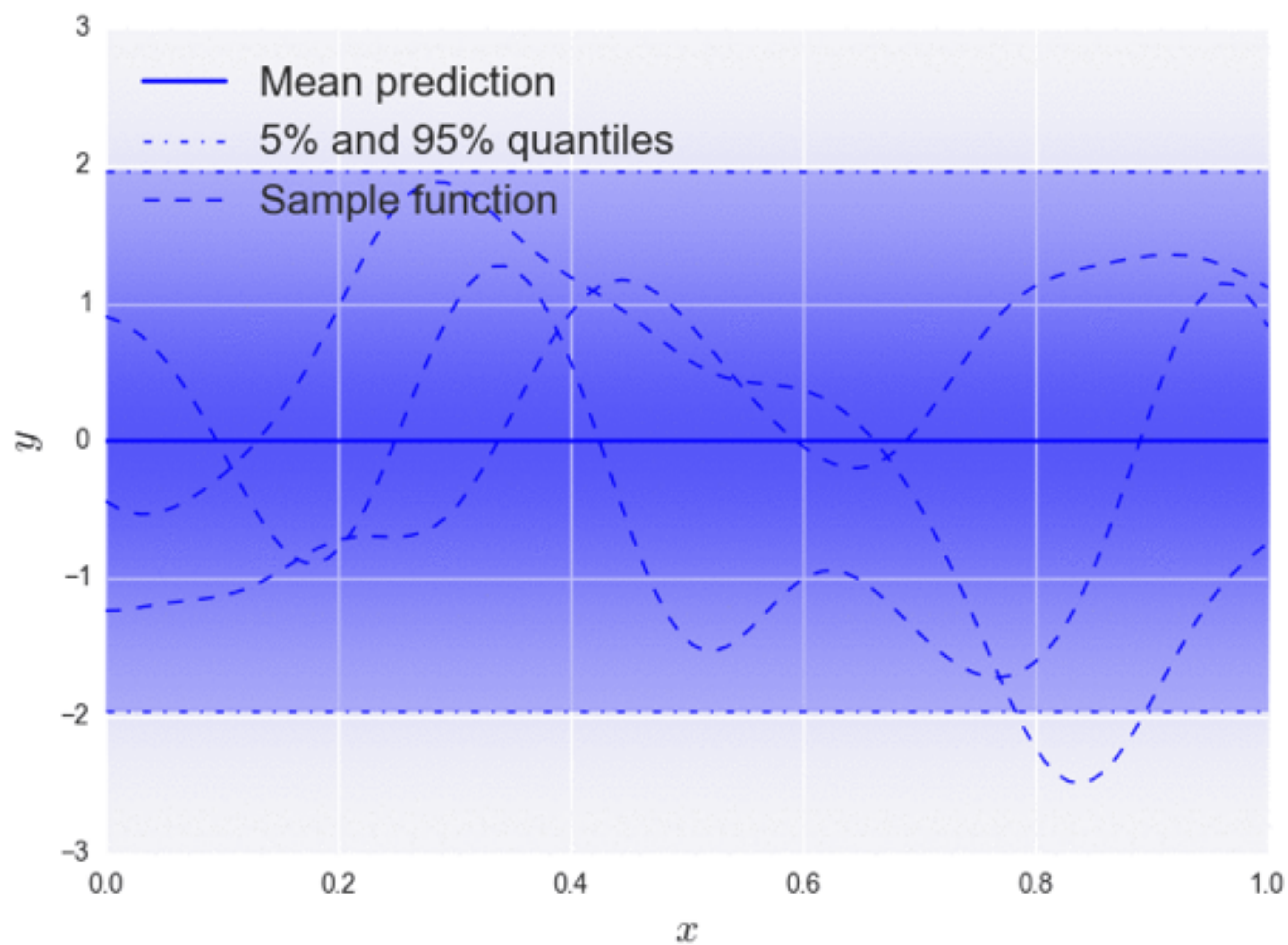
Changing the length scale



PREDICTIVE
SCIENCE LABORATORY

The samples are as smooth as the covariance

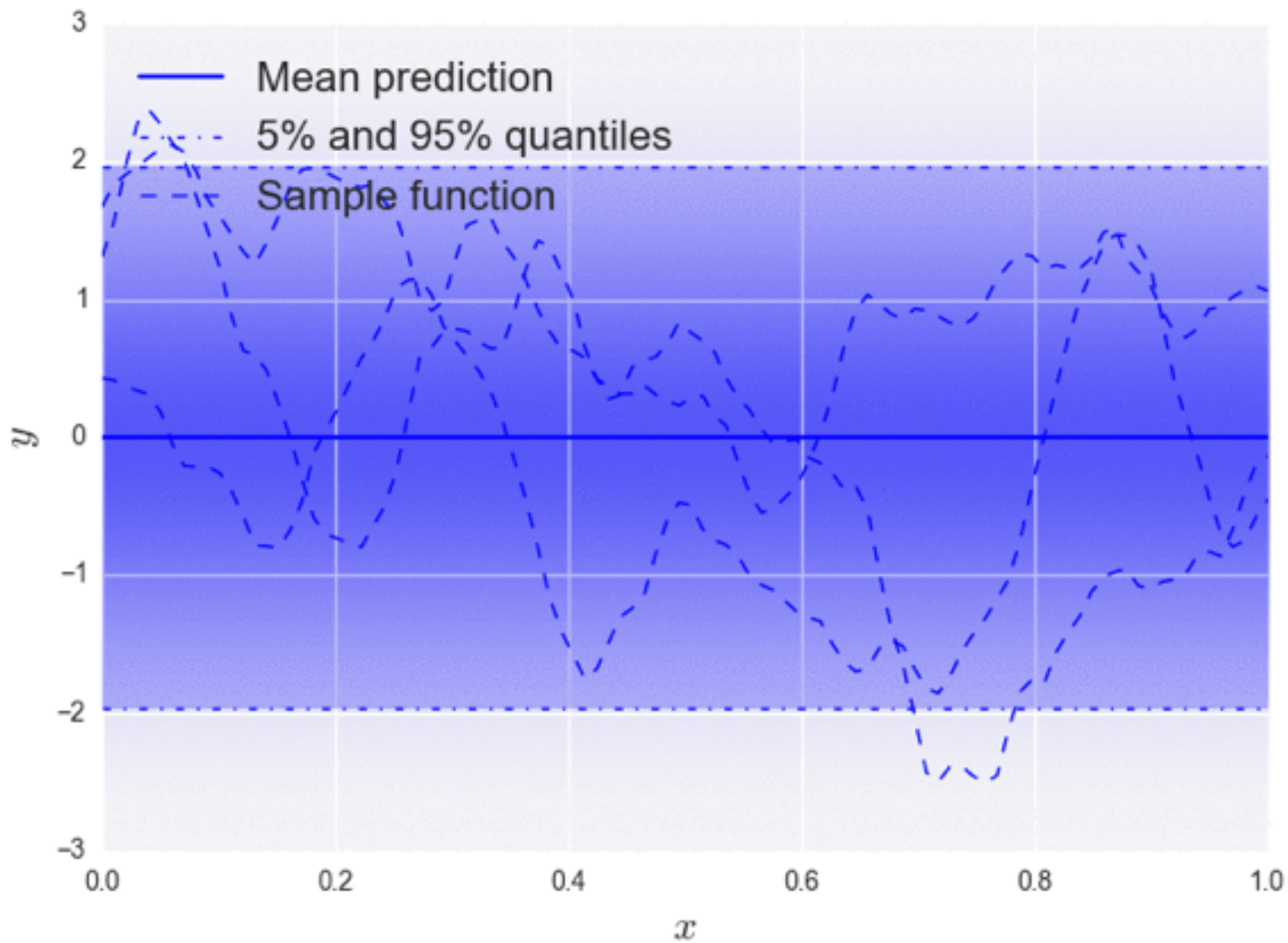
Infinitely smooth SE covariance



PREDICTIVE
SCIENCE LABORATORY

The samples are as smooth as the covariance

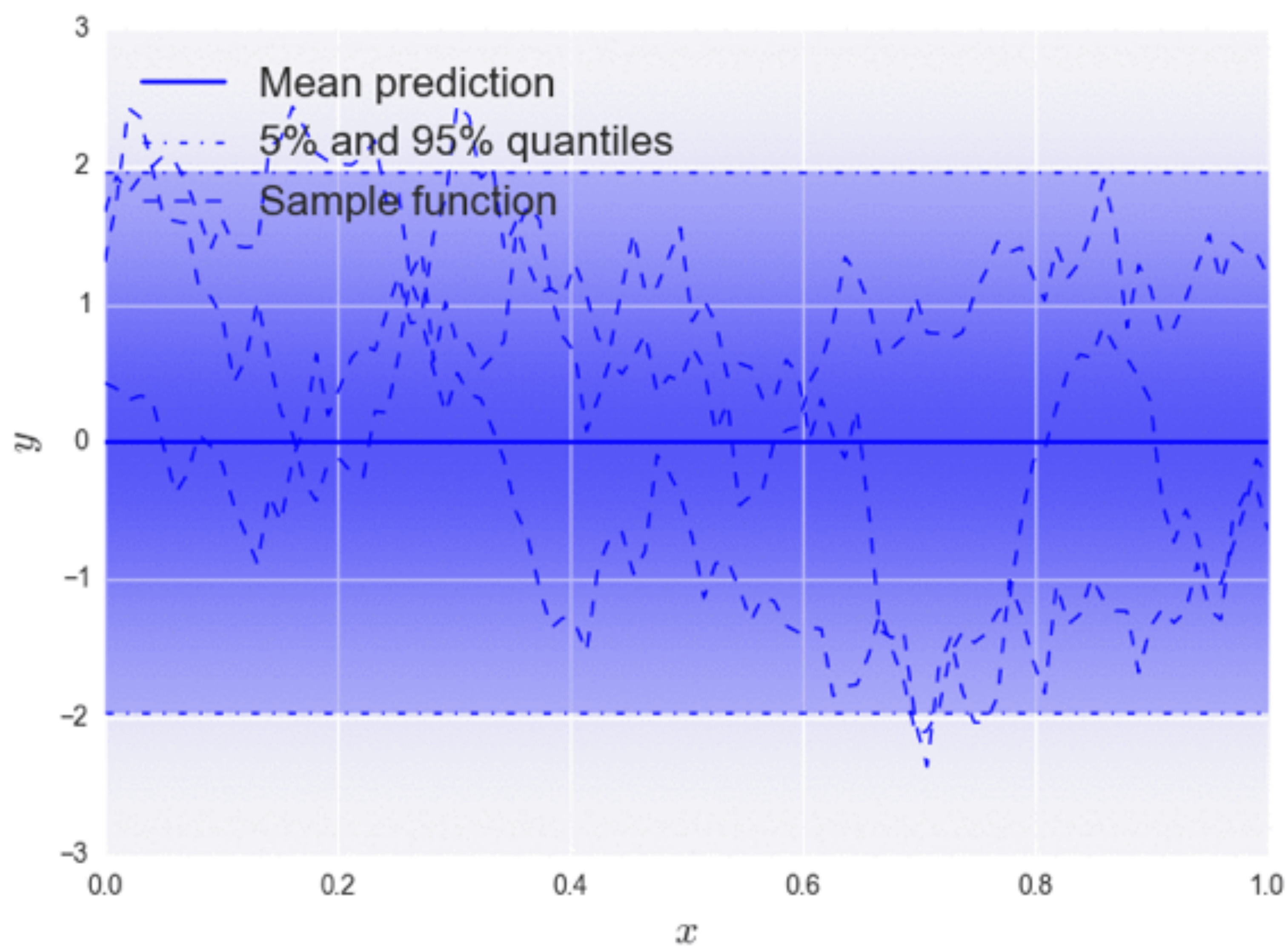
Matern 2-3, 2 times differentiable



PREDICTIVE
SCIENCE LABORATORY

The samples are as smooth as the covariance

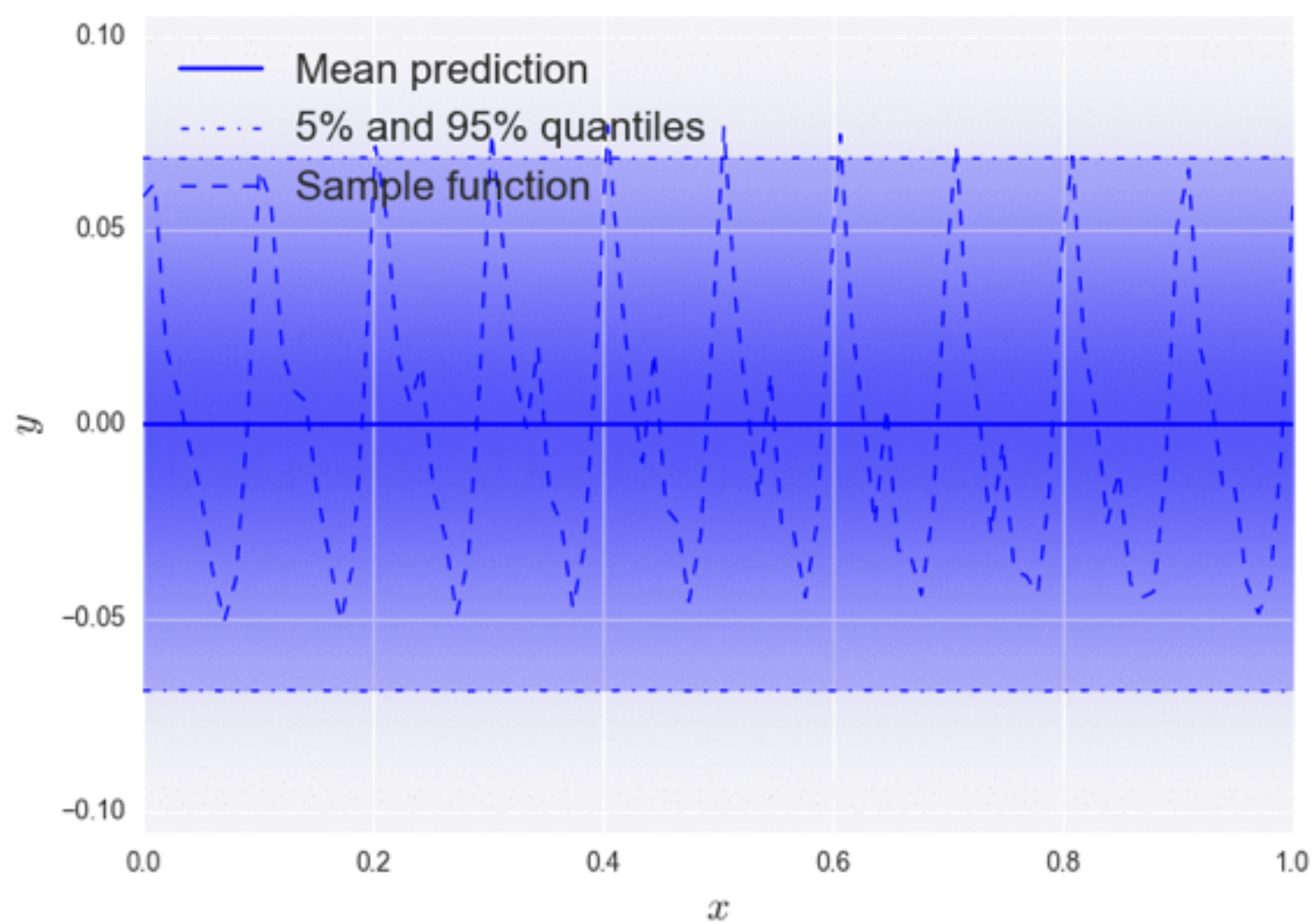
Exponential, continuous, nowhere differentiable



PREDICTIVE SCIENCE LABORATORY

Invariances may be built-into covariance functions

Periodic Exponential, period = 0.1



PREDICTIVE
SCIENCE LABORATORY

Example 3: Drawing Samples from a Gaussian Process

You have until the end of class

PREDICTIVE
SCIENCE LABORATORY