



Bookmarks



Bookmark

## ► Introduction

## ▼ 1. Probability and Inference

## Introduction to Probability (Week 1)

Exercises due Sep 21, 2016 at 21:00 UTC

## Probability Spaces and Events (Week 1)

Exercises due Sep 21, 2016 at 21:00 UTC

## Random Variables (Week 1)

Exercises due Sep 21, 2016 at 21:00 UTC

## Jointly Distributed Random Variables (Week 2)

Exercises due Sep 28, 2016 at 21:00 UTC

## Conditioning on Events (Week 2)

Exercises due Sep 28, 2016 at 21:00 UTC

## Homework 1 (Week 2)

Homework due Sep 28, 2016 at 21:00 UTC

## Inference with Bayes' Theorem for Random Variables (Week 3)

Exercises due Oct 05, 2016 at 21:00 UTC

## Independence Structure (Week 3)

Exercises due Oct 05, 2016 at 21:00 UTC

## Homework 2 (Week 3)

## 1. Probability and Inference &gt; Inference with Bayes' Theorem for Random Variables (Week 3) &gt; Maximum A Posteriori (MAP) Estimation

## MAXIMUM A POSTERIORI (MAP) ESTIMATION


For a hidden random variable  $\mathbf{X}$  that we are inferring, and given observation  $\mathbf{Y} = \mathbf{y}$ , we have been talking about computing the posterior distribution  $p_{\mathbf{X}|\mathbf{Y}}(\cdot|\mathbf{y})$  using Bayes' rule. The posterior is a distribution for what we are inferring. Often times, we want to report which particular value of  $\mathbf{X}$  actually achieves the highest posterior probability, i.e., the most probable value  $\mathbf{x}$  that  $\mathbf{X}$  can take on given that we have observed  $\mathbf{Y} = \mathbf{y}$ .

The value that  $\mathbf{X}$  can take on that maximizes the posterior distribution is called the *maximum a posteriori* (MAP) estimate of  $\mathbf{X}$  given  $\mathbf{Y} = \mathbf{y}$ . We denote the MAP estimate by  $\hat{\mathbf{x}}_{\text{MAP}}(\mathbf{y})$ , where we make it clear that it depends on what the observed  $\mathbf{y}$  is. Mathematically, we write

$$\hat{\mathbf{x}}_{\text{MAP}}(\mathbf{y}) = \arg \max_{\mathbf{x}} p_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}).$$


Note that if we didn't include the "arg" before the "max", then we would just be finding the highest posterior probability rather than which value—or "argument"— $\mathbf{x}$  actually achieves the highest posterior probability.

In general, there could be ties, i.e., multiple values that  $\mathbf{X}$  can take on are able to achieve the best possible posterior probability.

Homework due Oct 05,  
2016 at 21:00 UTC 

**Notation Summary  
(Up Through Week  
3)**

**Mini-project 1:  
Movie  
Recommendations  
(Week 3)**

Mini-projects due Oct  
12, 2016 at 21:00 UTC 

© All Rights Reserved



© 2016 edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

POWERED BY  
**OPENedX**

