



Bookmarks

- Introduction
- Part 1: Probability and Inference
- ▼ Part 2: Inference in Graphical Models

Week 5: Introduction to Part 2 on Inference in Graphical Models

**Week 5: Efficiency in Computer Programs**

Exercises due Oct 20, 2016 at 02:30 IST



**Week 5: Graphical Models**

Exercises due Oct 20, 2016 at 02:30 IST



**Week 5: Homework 4**

Homework due Oct 20, 2016 at 02:30 IST



**Week 6: Inference in Graphical Models - Marginalization**

Part 2: Inference in Graphical Models > Week 5: Efficiency in Computer Programs > Exercise: Storing the Joint Probability Table of Two Random Variables

## Exercise: Storing the Joint Probability Table of Two Random Variables

🔖 Bookmark this page

### STORING JOINTLY DISTRIBUTED RANDOM VARIABLES

Let's start with two random variables,  $\mathbf{X}$  and  $\mathbf{Y}$ . To infer the value of random variable  $\mathbf{X}$  given the observed value of a random variable  $\mathbf{Y}$ , we first model how  $\mathbf{X}$  and  $\mathbf{Y}$  relate. As we saw in the first part of the course, we can model the relationship between  $\mathbf{X}$  and  $\mathbf{Y}$  by looking at their joint probability table  $p_{\mathbf{X},\mathbf{Y}}$  (which we might not know directly but can piece together if, for instance, we had other tables such as  $p_{\mathbf{X}}$  and  $p_{\mathbf{Y}|\mathbf{X}}$ ).

### Exercise: Storing the Joint Probability Table of Two Random Variables

4/4 points (graded)

Suppose  $\mathbf{X}$  takes on  $k$  different values and  $\mathbf{Y}$  takes on  $\ell$  different values.

For the answer boxes below, please provide your answer as a mathematical formula (and not as Python code). Use ^ for exponentiation, e.g.,  $x^2$  denotes  $x^2$ . Explicitly include multiplication using \*, e.g.  $x*y$  is  $xy$ .

Exercises due Oct 27, 2016 at 02:30 IST



### Week 6: Special Case: Marginalization in Hidden Markov Models

Exercises due Oct 27, 2016 at 02:30 IST



### Week 6: Homework 5

Homework due Oct 27, 2016 at 02:30 IST



Weeks 6 and 7: Mini-project on Robot Localization (to be posted)

- Suppose that  $X$  and  $Y$  may possibly be dependent (i.e., we don't know if they are independent). How many table entries are in the joint probability table  $p_{X,Y}$ ?

✓ Answer:  $k \cdot l$

- Next, consider when the two random variables  $X$  and  $Y$  are known to be independent. This means that the joint probability table factorizes so that  $p_{X,Y}(x,y) = p_X(x)p_Y(y)$ . In other words, rather than storing  $p_{X,Y}$ , we could instead store the marginal distributions  $p_X$  and  $p_Y$ , and then to compute  $p_{X,Y}(x,y)$ , we first look up  $p_X(x)$  and  $p_Y(y)$  and multiply these two numbers together.

How many table entries are in the probability table  $p_X$ ?

✓ Answer:  $k$

How many table entries are in the probability table  $p_Y$ ?

✓ Answer:  $l$

Note that the number of table entries needed to store  $p_X$  and  $p_Y$  without making any additional assumptions is precisely the sum of the number of table entries in  $p_X$  and the number of table entries in  $p_Y$ .

- Now suppose that  $p_X$  and  $p_Y$  are actually independent and identically distributed and so each of  $X$  and  $Y$  now take on  $k$  possible values. How many table entries are needed to store  $p_X$  and  $p_Y$ ? Please provide your answer in terms of  $k$  and not  $\ell$ .

✓ Answer: k

Independence can drastically reduce how many numbers we need to store especially when we look at distributions over many random variables!

### Solution:

- Suppose that  $X$  and  $Y$  may possibly be dependent (i.e., we don't know if they are independent). How many table entries are in the joint probability table  $p_{X,Y}$ ?

**Solution:** There are  $k$  possibilities for  $X$ , and  $\ell$  possibilities for  $Y$  so the joint probability table has  $kl$  entries.

- Next, consider when the two random variables  $X$  and  $Y$  are known to be independent. This means that the joint probability table factorizes so that  $p_{X,Y}(x, y) = p_X(x)p_Y(y)$ . In other words, rather than storing  $p_{X,Y}$ , we could instead store the marginal distributions  $p_X$  and  $p_Y$ , and then to compute  $p_{X,Y}(x, y)$ , we first look up  $p_X(x)$  and  $p_Y(y)$  and multiply these two numbers together.

How many table entries are in the probability table  $p_X$ ?

**Solution:** There are  $k$  possibilities for  $X$ , so  $p_X$  has  $k$  entries.

How many table entries are in the probability table  $p_Y$ ?

**Solution:** There are  $\ell$  possibilities for  $Y$ , so  $p_Y$  has  $\boxed{\ell}$  entries.

Note that the number of table entries needed to store  $p_X$  and  $p_Y$  without making any additional assumptions is precisely the sum of the number of table entries in  $p_X$  and the number of table entries in  $p_Y$ .

- Now suppose that  $p_X$  and  $p_Y$  are actually independent and identically distributed and so each of  $X$  and  $Y$  now take on  $k$  possible values. How many table entries are needed to store  $p_X$  and  $p_Y$ ? Please provide your answer in terms of  $k$  and not  $\ell$ .

**Solution:** It suffices to store a single table for both  $p_X$  and  $p_Y$  with  $\boxed{k}$  entries.

Submit

You have used 1 of 5 attempts

✓ Correct (4/4 points)

© 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®

