



Bookmarks

► Introduction

▼ 1. Probability and Inference

Introduction to Probability (Week 1)

Exercises due Sep 22, 2016 at 02:30 IST



Probability Spaces and Events (Week 1)

Exercises due Sep 22, 2016 at 02:30 IST



Random Variables (Week 1)

Exercises due Sep 22, 2016 at 02:30 IST



Jointly Distributed Random Variables (Week 2)

Exercises due Sep 29, 2016 at 02:30 IST



Conditioning on Events (Week 2)

Exercises due Sep 29, 2016 at 02:30 IST



1. Probability and Inference &gt; Decisions and Expectations (Week 4) &gt; Exercise: Variance and Standard Deviation



Bookmark

## Exercise: Variance and Standard Deviation

(7/7 points)








This exercise explores the important concept of *variance*, which measures how much a random variable deviates from its expectation. This can be thought of as a measure of uncertainty. Higher variance means more uncertainty.

The variance of a real-valued random variable  $X$  is defined as

$$\text{var}(X) \triangleq \mathbb{E}[(X - \mathbb{E}[X])^2].$$

Note that as we saw previously,  $\mathbb{E}[X]$  is just a single number. To keep the variance of  $X$ , what you could do is first compute the expectation of  $X$ .

For example, if  $X$  takes on each of the values 3, 5, and 10 with equal probability  $1/3$ , then first we compute  $\mathbb{E}[X]$  to get 6, and then we compute  $\mathbb{E}[(X - 6)^2]$ , where we remember to use the result that for a function  $f$ , if  $f(X)$  is a real-valued random variable, then  $\mathbb{E}[f(X)] = \sum_x f(x)p_X(x)$ . Here,  $f$  is given by  $f(x) = (x - 6)^2$ . So

**Homework 1 (Week 2)**Homework due Sep 29, 2016 at 02:30 IST **Inference with Bayes' Theorem for Random Variables (Week 3)**Exercises due Oct 06, 2016 at 02:30 IST **Independence Structure (Week 3)**Exercises due Oct 06, 2016 at 02:30 IST **Homework 2 (Week 3)**Homework due Oct 06, 2016 at 02:30 IST **Notation Summary (Up Through Week 3)****Mini-project 1: Movie Recommendations (Weeks 3 and 4)**Mini-projects due Oct 13, 2016 at 02:30 IST **Decisions and Expectations (Week 4)**Exercises due Oct 13, 2016 at 02:30 IST **Measuring Randomness (Week 4)**Exercises due Oct 13, 2016 at 02:30 IST 

$$\text{var}(X) = (3 - 6)^2 \cdot \frac{1}{3} + (5 - 6)^2 \cdot \frac{1}{3} + (10 - 6)^2 \cdot \frac{1}{3} = \frac{26}{3}.$$

Let's return to the three lotteries from earlier. Here, random variables  $L_1$ ,  $L_2$ , and  $L_3$  represent the amount won (accounting for having to pay \$1):

		Probability
	-1	$\frac{999999}{1000000}$
$L_1$	-1 + 1000	$\frac{1}{1000000}$
		Probability
	-1	$\frac{999999}{1000000}$
$L_2$		1

## Towards Infinity in Modeling Uncertainty (Week 4)

Exercises due Oct 13, 2016 at 02:30 IST



## Homework 3 (Week 4)

Homework due Oct 13, 2016 at 02:30 IST



$$-1 + 1000000 \left[ \frac{1}{1000000} \right]$$

Probability

$L_3$	-1	$\frac{9}{10}$
	$-1 + 10$	$\frac{1}{10}$

Compute the variance for each of these three random variables. (Please provide the **exact** answer for each of these.)

•  $\text{var}(L_1) =$   ✓

•  $\text{var}(L_2) =$   ✓

•  $\text{var}(L_3) =$   ✓

What units is variance in? Notice that we started with dollars, and then variance is looking at the expectation of a dollar amount squared. Thus, specifically for the lottery example  $\text{var}(L_1)$ ,  $\text{var}(L_2)$ , and  $\text{var}(L_3)$  are each in squared dollars.

Some times, people prefer keeping the units the same as the original units (i.e., without squaring), which you can get by computing what's called the *standard deviation* of a real-valued random variable  $X$ :

$$\text{std}(X) \triangleq \sqrt{\text{var}(X)}.$$

Compute the following standard deviations, which are in units of dollars. (Please be precise with at least 3 decimal places, unless of course the answer doesn't need that many decimal places. You could also put a fraction.)

•  $\text{std}(L_1) =$   ✓

•  $\text{std}(L_2) =$   ✓

•  $\text{std}(L_3) =$   ✓

Note that when we first introduced the three lotteries and computed average winnings, we didn't account for the uncertainty in the average winnings. Here, it's clear that the third lottery has far smaller standard deviation and variance than the second lottery.

As a remark, often in financial applications (e.g., choosing a portfolio of stocks to invest in), accounting for uncertainty is extremely important. For example, you may want to maximize profit while ensuring that the amount of uncertainty is not too high as to not be reckless in investing.

In the case of the three lotteries, to decide between them, you could for example use a score that is of the form

$$\mathbb{E}[L_i] - \lambda \cdot \text{std}(L_i) \quad \text{for } i = 1, 2, 3,$$

where  $\lambda \geq 0$  is some parameter that you choose for how much you want to penalize uncertainty in the lottery outcome. Then you could choose the lottery with the highest score.

Finally, a quick sanity check (this is more for you to think about the definition of variance rather than to compute anything out):

- Can variance be negative? If yes, give a specific distribution as a Python dictionary for which the variance is negative. If no, enter the text "no" (all lowercase, one word, no spaces).



*You have used 3 of 5 submissions*



© 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.

