Dr. Z.'s Probability Lecture 7 Handout: Expectation and Variance

By Doron Zeilberger

Version of Oct. 18, 2017 (Thanks to Jiaqui Liu: correcting a typo in problem 7.3)

Important Fact: If X is a discrete random variable taking one of the values $x_i, i \geq 1$, with probability mass function $P\{X = x_i\} = p(x_i)$, and $g(x)$ is any real-valued function, then the expectation of the brand-new random variable, $g(X)$, is given by

$$E[g(X)] = \sum_i g(x_i)p(x_i) \quad .$$

Note: For the special case $g(x) = 1$ we get $E[1] = \sum_i 1 \cdot p(x_i) = 1$, since all the probabilities must add-up to 1, and of course, the expectation of 1 is 1, as it should be. When $g(x) = x$, we get the good-old expression for $E[X]$.

Problem 7.1: If the probability mass function of the discrete random variable X is

$$P\{X = -2\} = 0.1 \quad , \quad P\{X = -1\} = 0.3 \quad , \quad P\{X = 1\} = 0.4 \quad , \quad P\{X = 2\} = 0.2 \quad ,$$

and $P\{X = x\} = 0$ if $x \notin \{-2, -1, 1, 2\}$ find

(a) $E[X^3 + 2X]$

(b) $E[e^X]$

Sol. to 7.1(a):

$$E[X^3+2X] = 0.1\cdot((-2)^3+2\cdot(-2))+0.3\cdot((-1)^3+2\cdot(-1))+0.4\cdot((1)^3+2\cdot(1))+0.2\cdot((2)^3+2\cdot(2)) \;=\; 1.5 \quad .$$

Sol. to 7.1(b):

$$E[e^X] = 0.1 \cdot e^{-2} + 0.3 \cdot e^{-1} + 0.4 \cdot e^1 + 0.2 \cdot e^2 = \; 2.68902131 \ldots \quad .$$

Simple (but Important!) Fact If a and b are constants, and X is *any* random variable, then

$$E[aX + b] = aE[X] + b \quad .$$

Problem 7.2: If you are told that $E[X] = 5$, find $E[10X - 11]$.

Sol. to 7.2: $E[10X - 11] = 10E[X] - 11 = 10 \cdot 5 - 11 = 50 - 11 = 39$.

Important Concept: The n-th moment of a random variable X is $E[X^n]$.

Note: In particular the 0-th moment is always 1 and the *first moment* is another name for the *expectation.*

Important Definition

If X is a random variable with mean μ, then the **variance** of X, denoted by $Var(X)$, is defined by

$$Var(X) \;=\; E[(X-\mu)^2] \quad .$$

Important Formula

$$Var(X) \;=\; E[X^2] - E[X]^2 \quad .$$

Simple but useful fact: If a and b are constants, then

$$Var(aX+b) = a^2 Var(X) \quad .$$

Note: The b disappears, since adding a constant to a random variable does not change its "variability". The factor a^2 is explained by $E[(aX)^2] = E[a^2X^2] = a^2E[X^2]$.

Important Definition: The **standard deviation** of a random variable X, written $SD(X)$, is the **square-root** of $Var(X)$:

$$SD(X) := \sqrt{Var(X)} \quad .$$

Note: In an ideal world, the "variability of X" should be measured by $E[|X-\mu|]$, the "average" of the absolute value of the "deviation", but this is very hard to work with mathematically, (since the function $f(x) = |x|$ is not user-friendly, for example, it is not differentiable at $x = 0$) so mathematicians ""cheat" and use $E[(X-\mu)^2]$, since the function $f(x) = x^2$ is very nice, and always positive. But to make up for taking the square *inside* $E[.]$, they have to take the square-root *outside*.

Problem 7.3: The probability mass function for a certain discrete random variable X is as follows

$$P\{X=1\} = 0.1 \quad , \quad P\{X=2\} = 0.2 \quad , \quad P\{X=3\} = 0.3 \quad , \quad P\{X=4\} = 0.4 \quad ,$$

and $P\{X=x\} = 0$ if $x \notin \{1,2,3,4\}$.

(a) Find the expectation, $E[X]$.

(b) Find the Variance $Var(X)$. Also find the standard deviation.

Sol. to 7.3(a):

$$E[X] = 0.1 \cdot 1 \,+\, 0.2 \cdot 2 \,+\, 0.3 \cdot 3 \,+\, 0.4 \cdot 4 \,=\, 3 \quad .$$

Ans. to 7.3(a): $E[X] = \mu = 3$

Sol. to 7.3(b) (First Way: Directly):

$$E[(X-\mu)^2] = 0.1\cdot(1-3)^2 \,+\, 0.2\cdot(2-3)^2 \,+\, 0.3\cdot(3-3)^2 \,+\, 0.4\cdot(4-3)^2 \;=\; 1 \quad .$$

The standard deviation, $SD(X)$ is $\sqrt{1}=1$.

Sol. to 7.3(b) (Second Way:via the formula $Var(X)=E[X^2]-E[X]^2$**)**:

$$E[X^2] \;=\; 0.1\cdot 1^2 \,+\, 0.2\cdot 2^2 \,+\, 0.3\cdot 3^2 \,+\, 0.4\cdot 4^2 \;=\; 10 \quad .$$

Hence $Var(X)=10-3^2=1$. The standard deviation, $SD(X)$ is $\sqrt{1}=1$.

Problem 7.4: In a certain game, you can either win one, two, or three dollars. The probability that you win i dollars $(1\leq i\leq 3)$ is proportional to $\frac{1}{i^2}$.

(a) What is the probability of winning i dollars for $(1\leq i\leq 3)$?

(b) Let X be the amount won, what is $Var(X)$? Also find the standard deviation.

Sol. to 7.4: We must first find the **probability mass function**. We know that for some positive number c

$$P\{X=i\} = \begin{cases} \frac{c}{i^2} & if \quad 1\leq i\leq 3; \\ 0 & otherwise. \end{cases}$$

So $P\{X=1\}+P\{X=2\}+P\{X=3\}=c(\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2})=\frac{49c}{36}$. Since this must be 1 we get that $c=\frac{36}{49}$.

Ans. to 7.4(a):

$$P\{X=i\} = \begin{cases} \frac{36}{49\,i^2} & if \quad 1\leq i\leq 3; \\ 0 & otherwise. \end{cases}$$

Sol. to 7.4(b):

$$E[X]=P\{X=1\}\cdot 1+P\{X=2\}\cdot 2+P\{X=3\}\cdot 3 \;=\; \frac{36}{49}(1\cdot 1+2\cdot\frac{1}{4}+3\cdot\frac{1}{9})=\frac{36}{49}(1+\frac{1}{2}+\frac{1}{3})=\frac{36}{49}\cdot\frac{11}{6}=\frac{66}{49} \quad .$$

$$E[X^2]=P\{X=1\}\cdot 1^2+P\{X=2\}\cdot 2^2+P\{X=3\}\cdot 3^2 \;=\; \frac{36}{49}(1\cdot 1+2^2\cdot\frac{1}{4}+3^2\cdot\frac{1}{9})=\frac{36}{49}(1+1+1)=\frac{36\cdot 3}{49}=\frac{108}{49} \quad .$$

Hence

$$Var(X)=E[X^2]-E[X]^2 \;=\; \frac{108}{49}-(\frac{66}{49})^2=\frac{936}{2401}=0.3898375677\ldots$$

and

$$SD(X)=\sqrt{936/2401}=\sqrt{26}\frac{6}{49}=0.6243697364 \quad .$$

WARNING: The variance (and standard deviation!) are always **positive**! If you ever will obtain, in a quiz, or exam, a negative answer to a variance problem, due to a calculation error, you would get **ZERO**. Should that happen, and you won't have time to check your calculation, you **must** state that you realize that you made a calculation error, since variance can never be negative!

Note: When the variance is larger than 1, the standard deviation is smaller than the variance, but when the variance is less than 1 it is larger.

Problem 7.5: A gambler pays 2000 dollars for the following bet, regarding the outcomes of two sport events, one baseball, one football.

• If none of his teams win, he would get nothing.

• If only his favorite baseball team wins, he would get 1000 dollars.

• If only his favorite football team wins, he would get 2000 dollars.

• If both teams win, he would get 4000 dollars.

The probability of the baseball team winning is 0.4, and the probability of the football team winning is 0.7, and are independent of each other.

Find the expectation and standard deviation of his winning.

Sol. to 7.5:

There are four cases: (note that the events of the baseball and football team winning are independent event).

• Both teams lose. The probability is $(1-0.4)\cdot(1-0.7) = 0.6\cdot 0.3 = 0.18$, the gain is $0-2000 = -2000$.

• The baseball team wins but the football team loses. The probability is $0.4\cdot(1-0.7) = 0.4\cdot 0.3 = 0.12$, the gain is $1000-2000 = -1000$.

• The baseball team loses but the football team wins. The probability is $(1-0.4)\cdot 0.7 = 0.6\cdot 0.7 = 0.42$, the gain is $2000-2000 = 0$.

• The baseball team wins and the football team wins. The probability is $0.4\cdot 0.7 = 0.28$, the gain is $4000-2000 = 2000$.

Hence the **Probability mass function** for this discrete random variable is

$$P\{X=-2000\} = 0.18 \quad , \quad P\{X=-1000\} = 0.12 \quad , \quad P\{X=0\} = 0.42 \quad , \quad P\{X=2000\} = 0.28 \quad .$$

(and of course $P\{X=x\} = 0$ if $x \notin \{-2000, -1000, 0, 2000\}$).

We have

$$E[X] = 0.18 \cdot (-2000) + 0.12 \cdot (-1000) + 0.42 \cdot 0 + 0.28 \cdot 2000 \;=\; 80 \quad .$$

So his expected win is 80 dollars. We also have

$$E[X^2] = 0.18 \cdot (-2000)^2 + 0.12 \cdot (-1000)^2 + 0.42 \cdot 0^2 + 0.28 \cdot 2000^2 \;=\; 1960000 \quad .$$

Hence

$$Var(X) = E[X^2] - E[X]^2 = 1960000 - 80^2 = 1953600 \quad .$$

So the variance of his gain, $Var(X)$ is 1953600. The *standard deviation* is $1397.712417\ldots$.

Ans. to 7.5: The expected gain of the gambler is 80 dollars and the standard deviation is $1397.712417\ldots$.

Note: This gambler must be *risk-seeking*. It is true that his expected gain is positive, 80 dollars, so it is a "good deal" in the long run, if he does it many many times. But if he only does it one time, he has a probability of 0.3 of losing a lot, a probability of 0.42 of breaking even, and only a probability of 0.28 of winning (but then he would make 2000 dollars!).

Problem 7.6: Johnny is supposed to do chores half an hour a day, but if he does more he would get paid.

For every full ten minutes of doing chores in excess of 30 minutes, Mom pays Johnny 10 dollars, up to a maximum of 30 dollars.

It is known that the probabilities of Johnny doing chores are as follows.

- The probability that Johnny only does the minimum is 0.3 .
- The probability that Johnny works more than 30 minutes but less than 40 minutes is 0.1 .
- The probability that Johnny works more than 40 minutes but less than 50 minutes is 0.3 .
- The probability that Johnny works more than 50 minutes but less than 60 minutes is 0.2 .
- The probability that Johnny works more than 60 minutes is 0.1.

Calculate the standard deviation of the amount that Mom pays Johnny.

Sol. to 7.6:

$P\{X = 0\} = 0.3{+}0.1 = 0.4 \quad , \quad P\{X = 10\} = 0.3 \quad , \quad P\{X = 20\} = 0.2 \quad , \quad P\{X = 30\} = 0.1 \quad .$

(Of course $P\{X = x\} = 0$ if $x \notin \{0, 10, 20, 30\}$).

We have

$$E[X] = 0.4 \cdot 0 + 0.3 \cdot 10 + 0.2 \cdot 20 + 0.1 \cdot 30 \;=\; 10 \quad .$$

Next

$$E[X^2] = 0.4 \cdot 0^2 + 0.3 \cdot 10^2 + 0.2 \cdot 20^2 + 0.1 \cdot 30^2 \; = \; 200 \quad .$$

Hence $Var(X) = 200 - 10^2 = 200 - 100 = 100$, and $SD(X) = \sqrt{100} = 10 \quad .$

Ans. to 7.6: $SD(X) = 10$. In words: the standard deviation of Johnny's pay is 10 dollars.

Warning: The **bottom line** should only contain what has been asked!

Warning: Do not confuse between $Var(X)$ and $SD(X)$! Some questions ask for the variance, some for the standard deviation, and some for both. Some questions also ask for the expectation $E[X]$. Of course, if you are asked to find the variance and/or the standard-deviation, you still need, as a preliminary step, to find $E[X]$.