Math 425 Introduction to Probability Lecture 15

Kenneth Harris
kaharri@umich.edu

Department of Mathematics
University of Michigan

February 13, 2009

Example: Three bets

Example. A casino offers the following bets (the fairest bets in the casino!)

1. You get \$0 (i.e., you can walk away)
2. You get \$10 with probability $\frac{1}{2}$, and otherwise pay \$10.
3. You get \$10,000 with probability 10^{-3}, and otherwise pay \$10.

☞ All three bets have the same expectation: 0, but they differ in how spread out they are about their mean:

1. This bet is not spread out at all.
2. This bet is symmetrically disposed not too far from its mean.
3. This bet is very spread out.

Definition: Variance

☞ Variance is a measure of the spread of a random variable.

Definition

Let X be a random variable with distribution p_X and mean μ.
The variance of X, denoted by $Var(X)$, is defined as

$$Var(X) = \sum_{r:p_X(r)>0} (r-\mu)^2\, p_X(r).$$

The variance is a weighted average of the squared distance from the mean.

The variance of X is standardly written $\sigma^2(X)$, where $\sigma(X)$ is often used as a measure of the spread of X.

Standard Deviation

☞ The standard deviation is more like a distance function than variance.

Definition

The standard deviation of a random variable X is the positive square root of the variation:

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

Equivalently, if the mean of X has distribution p_X and mean μ, its pmf $p(r)$ then

$$SD(X) = \sqrt{\sum_{r:p_X(r)>0} (r-\mu)^2\, p_X(r)}.$$

The standard deviation of X is standardly written $\sigma(X)$.

Three bets revisited

Example–revisted. Our three bets with mean 0:

1. You get \$0 (X_1).
2. You get \$10 with probability $\frac{1}{2}$, and otherwise pay \$10 (X_2).
3. You get $\$10,000$ with probability 10^{-3}, and otherwise pay \$10 (X_3).

☞ They differ dramatically in the variance:

$$\begin{aligned} Var(X_1) &= 0 \\ Var(X_2) &= \frac{1}{2}(10-0)^2 + \frac{1}{2}(-10-0)^2 = 100 \\ Var(X_3) &= (1-10^{-3})(10^4-0)^2 + 10^{-3}(-10-0)^2 \approx 10^8 \end{aligned}$$

and so too in standard variation,

$$SD(X_1) = 0 \qquad SD(X_2) = 10 \qquad SD(X_3) \approx 10^4.$$

Throwing Dice

Example. Let X be the outcome of the throw of a single fair die.
The distribution of X is

$$p_X(k) = \frac{1}{6} \qquad 1 \le k \le 6$$

☞ So,

$$\begin{aligned} E[X] &= \sum_{k=1}^{6} \frac{k}{6} = \frac{7}{2} \\ Var(X) &= \sum_{k=1}^{6} (k - \frac{7}{2})^2 \cdot \frac{1}{6} = \frac{35}{12} \approx 2.92 \\ SD(X) &\approx 1.71 \end{aligned}$$

Caveat on Variance

☞ There is a tendancy to believe the standard deviation gives a good measure of dispersion: expected deviation from the mean.

☞ However, squaring big deviations tend to dominate the sum.
Consider the random variable X with distribution:

$$p_X(1) = \frac{9}{10} \qquad p_X(10) = \frac{1}{10}$$

So,

$$\begin{aligned} E[X] &= \frac{9}{10} + 10\frac{1}{10} = 1.9 \\ Var(X) &= (.9)^2\frac{9}{10} + (8.1)^2\frac{1}{10} = \frac{9}{4} = 7.29 \\ SD(X) &= 2.7 \end{aligned}$$

Caveat on Variance

☞ If deviation from the mean varies alot, then standard deviation is NOT a good measure of the expected deviation. A closer formula might be

$$\sum_{r:p_X(r)>0} |r-\mu| \cdot p_X(r).$$

In this case, our previous example gives an average deviation

$$|1-1.9| \cdot \frac{9}{10} + |10-1.9| \cdot \frac{1}{10} = 1.62 \qquad SD(X) = 2.7$$

☞ While this version of dispersion of values has good statistical properties, it is much more difficult mathematical properties, so is little used.

Example: Roulette

Example. A roulette wheel is divided into 38 slotted sectors.

- 18 are red, 18 are black and 2 are green,
- 36 are numbered 1 to 36, together with one marked 0 and one marked 00.

A croupier spins the wheel and throws an ivory ball. The success of a bet depends into which slot the ball falls.

Example: Roulette

☞ Compare the following two bets:

1. Bet \$1 on red at even money (X).
2. Bet \$1 on the number 17 at $35:1$ (Y).
   You win \$36 on a \$1 bet.

☞ We compare the expected values

$$E[X] = (-1)\cdot\frac{20}{38} + 1\cdot\frac{18}{38} = -\frac{1}{19} \approx -0.052$$
$$E[Y] = (-1)\cdot\frac{37}{38} + 1\cdot\frac{35}{38} = -\frac{1}{19} \approx -0.052$$

You can expect to lose about 5 cents a bet in either case.

The variance is dramatically different though

$$Var[X] = (-1+\frac{1}{19})^2\cdot\frac{20}{38} + (1+\frac{1}{19})^2\cdot\frac{18}{38} \approx 0.997$$
$$Var[Y] = (-1+\frac{1}{19})^2\cdot\frac{37}{38} + (35+\frac{1}{19})^2\cdot\frac{1}{38} \approx 33.21$$

Example: Roulette

☞ I played both bets 10,000 times. The standard mean of the outcome of the bet and the variance of the outcomes based on the computed mean is as follows:

- Even money on red:
  - Mean: -0.0442 $(E[X] \approx -0.052)$
  - Variance: 0.998 $(Var(x) \approx 0.997)$
- $35:1$ on the number 17:
  - Mean: -0.0784 $(E[X] \approx -0.052)$
  - Variance: 32.3315 $(Var(x) \approx 33.21)$

Example: Roulette

☞ I placed a 1000 bets at the roulette table and computed my total winnings. Here is the mean and variance over 100 plays of each type of bet.

- Even money on red for 1000 bets:
  - Mean winnings: -52.52 $(1000E[X] \approx -52)$
  - Variance: 730.15 $(1000Var(x) \approx 997)$
- $35:1$ on the number 17 for 1000 bets:
  - Mean: -53.92 $(1000E[X] \approx -52)$
  - Variance: $29,483$ $(1000Var(x) \approx 33,210)$

☞ This suggests that running these independent trials 1000 times leads to a 1000-fold increase in expectation and variance. This is the case, and we will look into this relationship in Chapter 7.

Linearity properties

☞ See Ross, Corollary 4.1 and page 150.

Theorem

Let a and b be constants. For any random variable X,

$$E[aX+b] = aE[X]+b$$

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

Proof of Theorem

Let X be a random variable with distribution p_X and let $Y = aX + b$ with distribution p_Y.

☞ Since

$$ar_1 + b = ar_2 + b \quad \text{if and only if} \quad r_1 = r_2,$$

it follows: $p_X(r) = p_Y(ar+b)$. So,

$$\begin{aligned} E[aX+b] = E[Y] &= \sum_{r:p_Y(ar+b)>0} (ar+b)\cdot p_Y(ar+b) \\ &= \sum_{r:p_X(r)>0} (ar+b)\cdot p_X(r) \\ &= a \sum_{r:p_X(r)>0} r\cdot p_X(r) + b \sum_{r:p_X(r)>0} p_X(r) \\ &= aE[X]+b. \end{aligned}$$

Proof of Theorem – continued

Let $\mu = E[X]$, so that $E[Y] = E[aX+b] = a\cdot\mu + b$.

☞ Since $Var(aX+b) = Var(Y)$,

$$\begin{aligned} Var(aX+b) &= \sum_{r:p_Y(ar+b)>0} (ar+b-(a\cdot\mu+b))^2\cdot p_Y(ar+b) \\ &= \sum_{r:p_X(r)>0} (ar-a\cdot\mu)^2\cdot p_X(r) \\ &= a^2 \sum_{r:p_X(r)>0} (r-\mu)^2\cdot p_X(r) \\ &= a^2 Var(X). \end{aligned}$$

Example: Temperature readings

Example

In a certain manufacturing process, the (Fahrenheit) temperature never varies more than two degrees from 62°F. The temperature is a random variable X with distribution

temp (°F)	60	61	62	63	64
prob	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{4}{10}$	$\frac{2}{10}$	$\frac{1}{10}$
temp (°C)	15.56	16.11	16.67	17.22	17.78

Find the following

- Compute $E[X]$, $Var(X)$ and $SD(X)$.
- It is decided to convert the temperature readings to Celcius, so that $Y = \frac{5}{9}(X-32)$. Compute $E[Y]$, Var(Y), and $SD(Y)$.

Example: Temperature readings

☞ In Fahrenheit

$$
\begin{aligned}
E[X] &= 60\frac{1}{10} + 61\frac{2}{10} + 62\frac{4}{10} + 63\frac{2}{10} + 64\frac{1}{10} = 62 \\
Var(X) &= (-2)^2\frac{1}{10} + (-1)^2\frac{2}{10} + (0)^2\frac{4}{10} + (1)^2\frac{2}{10} + (2)^2\frac{1}{10} = \frac{6}{5} \\
SD(X) &= \sqrt{\frac{6}{5}} \approx 1.1.
\end{aligned}
$$

☞ In Celcius

$$
\begin{aligned}
E[Y] &= E[\frac{5}{9}(X-32)] = \frac{5}{9}E[X] - \frac{160}{9} = \frac{50}{3} \approx 16.67 \\
Var(Y) &= Var(\frac{5}{9}(X-32)) = \frac{25}{81}Var(X) = \frac{30}{81} \\
SD(Y) &= \sqrt{\frac{30}{81}} \approx 0.61.
\end{aligned}
$$

Definition

☞ We generalize the previous result to arbitrary functions.

Definition

Let X be a random variable over sample space S, and $g : \mathbb{R} \to \mathbb{R}$.
We write $g(X)$ for the random variable Y on S defined by

$$Y(s) = g(X(s)) \qquad \text{for all } s \in S.$$

Examples

- $aX + b$ is the random variable: $aX(s) + b$ for $s \in S$,
- X^2 is the random variable: $\left(X(s)\right)^2$ for $s \in S$.

Proposition

☞ See Ross, Proposition 4.1

Proposition

Let X be a random variable with distribution p_X.
For any real-valued function g,

$$E[g(X)] = \sum_{r : p_X(r) > 0} g(r)p(r).$$

Proof of Proposition

Let X be a random variable with distribution, and let $Y = g(X)$ be the random variable with distribution p_Y.

☞ It is not necessarily true that $p_Y(g(r)) = p_X(r)$, since g may map several numbers to the same value. For example, let X be the random variable with distribution

$$p_X(-1) = \frac{1}{3} \qquad p_X(0) = \frac{1}{3} \qquad p_X(1) = \frac{1}{3}.$$

Then $Y = X^2$ is the random variable with distribution

$$p_Y(0) = \frac{1}{3} \qquad p_Y(1) = \frac{2}{3}.$$

So, the distributions for X and Y are different.

Proof of Proposition

☞ However, the following is true for each real number q:

$$p_Y(q) = \sum_{r:g(r)=q} p_X(r).$$

By regrouping the summation:

$$\begin{aligned}\sum_{r:p_X(r)>0} g(r)\,p_X(r) &= \sum_{q:p_Y(q)>0}\ \sum_{r:g(r)=q} g(r)\,p_X(r)\\ &= \sum_{q:p_Y(q)>0} q \sum_{r:g(r)=q} p_X(r)\\ &= \sum_{q:p_Y(q)>0} q\,p_Y(q)\\ &= E[Y] = E[g(X)].\end{aligned}$$

Corollary

☞ The following corollary makes computation of variance much easier.

Corollary

Let X be a random variable with expected value μ. Then

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

That is,

$$Var(X) = E[X^2] - \left(E[X]\right)^2$$

Proof of Corollary

We defined the variance of a random variable X with distribution p_X and expected value μ by

$$Var(X) = \sum_{r:p_X(r)>0} (r-\mu)^2\,p_X(r) = E[(X-\mu)^2],$$

by the previous theorem with $g(x) = (x-\mu)^2$.

$$\begin{aligned}Var(X) &= \sum_{r:p_X(r)>0} (r-\mu)^2 p_X(r)\\ &= \sum_{r:p_X(r)>0} (r^2 - 2r\mu + \mu^2)\,p_X(r)\\ &= \sum_{r:p_X(r)>0} r^2\,p_X(r) - 2\mu \sum_{r:p_X(r)>0} r\,p_X(r) + \mu^2 \sum_{r:p_X(r)>0} p_X(r)\\ &= E[X^2] - 2\mu^2 + \mu^2\\ &= E[X^2] - \mu^2 = E[X^2] - \left(E[X]\right)^2.\end{aligned}$$

Example: Algebra

Example. Let X be a random variable with $E[X] = 10$ and $Var(X) = 4$. Compute the following

(a) $E[X^2]$

(b) $E[3X+10]$ and $E[-X]$

(c) $Var(3X+10)$ and $Var(-X)$

(d) $SD(X)$, $SD(3X+10)$ and $SD(-X)$.

Example: Algebra

Given: $E[X] = 10$ and $Var(X) = 4$

(a). Since $Var(X) = E[X^2] - (E[X])^2$,

$$E[X^2] = Var(X) + (E[X])^2 = 4 + 10^2 = 104$$

(b).

$$E[3X + 10] = 3E[X] + 10 = 40 \qquad E[-X] = -E[X] = -10$$

(c).

$$Var(3X + 10) = 9Var(X) = 36 \qquad V(-X) = V(X) = 4$$

(c).

$$SD(X) = \sqrt{4} \qquad SD(3X + 10) = 6 \qquad SD(-X) = \sqrt{4}$$

Example 1: children

Example

Find the expected value and the variance for the number of boys and girls in a royal family that has children until there is a boy or until there are three children, whichever comes first.

Example 1 – continued

Solution. Let B (G) be the random variable which counts the boys (girls). Then, the distributions are given by the table:

B	p_B	G	p_G	G	p_G
0	$\frac{1}{8}$	0	$\frac{1}{2}$	2	$\frac{1}{8}$
1	$\frac{7}{8}$	1	$\frac{1}{4}$	3	$\frac{1}{8}$

☞

$$E[B] = \frac{7}{8} \qquad E[B^2] = \frac{7}{8} \qquad Var[B] = \frac{7}{64}$$

$$E[G] = \frac{7}{8} \qquad E[G^2] = \frac{15}{8} \qquad Var[G] = \frac{71}{64}$$

Example 2

Example

A die is loaded so that the probability of a face coming up is proportional to the number on the face. Find the expected value, variance and standard deviation of the face value.

Solution. Let X be the random value of the face value.

X has distribution given by

$$p_X(k) = \frac{k}{21} \qquad 1 \le k \le 6.$$

since $1 + 2 + 3 + 4 + 5 + 6 = 21$.

Example 2 – continued

$$p_X(k) = \frac{k}{21} \qquad 1 \le k \le 6.$$

☞ Compute.

$$E[X] = \frac{13}{3} \approx 4.3 \qquad E[X^2] = \frac{439}{21}$$
$$Var(X) = \frac{134}{63} \approx 2.13$$
$$SD(X) \approx 1.46$$

Recall, for a fair die these values are

$$E[X] = 3.5 \qquad Var(X) \approx 2.92 \quad SD(X) \approx 1.71$$

Example 3: heads

Example

What is the expected number of heads in 4 tosses of a fair coin? What is the standard deviation?

Solution. Let X be the random variable counting heads. X has distribution

$$p_X(k) = \binom{4}{k} 2^{-4} \qquad 0 \le k \le 4.$$

So,

$$E[X] = \sum_{k=0}^{4} k \binom{4}{k} 2^{-4} = 2$$
$$E[X^2] = \sum_{k=0}^{4} k^2 \binom{4}{k} 2^{-4} = 5$$
$$Var(X) = 1 \qquad SD(X) = 1$$

Example 4: heads

Example

What is the expected number of heads in 6 tosses of a fair coin? What is the standard deviation?

Solution. Let X be the random variable counting heads. X has distribution

$$p_X(k) = \binom{6}{k} 2^{-6} \qquad 0 \le k \le 6.$$

So,

$$E[X] = \sum_{k=0}^{6} k \binom{6}{k} 2^{-6} = 3$$
$$E[X^2] = \sum_{k=0}^{6} k^2 \binom{6}{k} 2^{-6} = \frac{21}{2}$$
$$Var(X) = \frac{3}{2} \qquad SD(X) \approx 1.225$$

Example 4: heads

☞ Random variables for counting heads for various numbers of tosses.

X	$E[X]$	$Var(X)$
1	0.5	0.25
2	1	0.5
4	2	1
6	3	1.5
8	4	2

☞ We explore this trend further on Monday.

Example 5: Jack and Jill, redoux

Example

Jack and Jill are playing the game of Heads. Jill's coin is biased $\frac{1}{3}$ of the time heads, Jack's coin is a fair coin. Jack, always the gentleman, allows Jill to toss first. They agree to stop after four rounds.

What is the expected number of rounds? What is the variance?

Example 5 – continued

Solution. Let X count the number of rounds played.

The probability of a winning throw in any given round is

$$\frac{1}{3} + \frac{2}{3} \cdot \frac{1}{2} = \frac{2}{3}.$$

So, the probability that play goes k rounds (when $k < 4$) is

$$p_X(k) \quad = \quad \left(\frac{2}{3}\right)^k \left(\frac{1}{2}\right)^{k-1} \qquad 1 \leq k \leq 3$$

The last round also has the possibility of no winner:

$$p_X(4) = \left(\frac{2}{3}\right)^4 \left(\frac{1}{2}\right)^3 + \left(\frac{2}{3}\right)^4 \left(\frac{1}{2}\right)^4$$

Example 5 – continued

☞ The probabilities for X are

n	p_X	n	p_X
1	$\frac{2}{3}$	3	$\frac{2}{27}$
2	$\frac{2}{9}$	4	$\frac{1}{27}$

☞ The expected number of rounds of the game:

$$E[X] = \frac{37}{27} \approx 1.37 \qquad E[X^2] = \frac{76}{27}$$

$$Var(X) \approx 0.94 \qquad SD(X) \approx 0.97$$