# VARIANCE, JOINT DISTRIBUTIONS, COVARIANCE

# 9

COMPUTER SCIENCE MENTORS 70

Oct 29 - Nov 2, 2018

# 1 Variance

## 1.1 Introduction

**Variance**: The variance of a random variable $X$ is defined as

$$\text{Var}(X) = \text{E}((X - \text{E}(X))^2) = \text{E}(X^2) - \text{E}(X)^2$$

The latter version of variance is the one we usually use in computations.

The square root of $\text{Var}(X)$ is called the standard deviation of $X$. It is usually denoted with the variable $\sigma$.

Important property of variance: for some constant $c$,

$$\text{Var}(cX) = c^2 * \text{Var}(X)$$

.

## 1.2 Questions

1. **True or False?**
   Assume that $X$ is a discrete random variable. If $Var(X) = 0$, then $X$ is a constant.

2. Show that $\text{Var}(X) = \text{E}((X - \mu)^2) = \text{E}(X^2) - \mu^2$ where $X$ is any random variable and $\mu = \text{E}[X]$.

3. Show that $Var(aX + b) = a^2 Var(X)$ where $X$ is any random variable.

4. Let's consider the classic problems of flipping coins and rolling dice. Let $X$ be a random variable for the number of coins that land on heads and $Y$ be the value of the die roll.

   (a) What is the expected value of $X$ after flipping 3 coins? What is the variance of $X$?

   (b) Let $Y$ be the sum of rolling a dice 1 time. What is the expected value of $Y$?

   (c) What is the variance of $Y$?

5. You are at a party with $n$ people where you have prepared a red solo cup labeled with their name. Before handing red cups to your friends, you pick up each cup and put a sticker on it with probability $\frac{1}{2}$ (independently of the other cups). Then you hand back the cups according to a uniformly random permutation. Let $X$ be the number of people who get their own cup back AND it has a sticker on it.

   (a) Compute the expectation $E(X)$.

   (b) Compute the variance $Var(X)$

## 2  Independence, Joint Distributions

### 2.1  Introduction

**Joint Distribution**: We use a joint distribution to describe a probability distribution over multiple random variables. A joint distribution is characterized by its joint probability mass function:

$$p_{xy}(x, y) = \mathrm{P}(X = x \cap Y = y).$$

We can recover the *marginal* distributions:

$$p_x(x) = p(X = x) = \sum_y \mathrm{P}(X = x, Y = y) = \sum_y p_{xy}(x, y)$$

$$p_y(y) = \mathrm{P}(Y = y) = \sum_x \mathrm{P}(X = x, Y = y) = \sum_x p_{xy}(x, y).$$

Note that in general, we can't get the joint distribution just by knowing the marginal distributions: in order to get the joint distribution, we also have to know how the variables relate to each other. However, if two random variables are **independent**, then we know that

$$p_{xy}(x, y) = \mathrm{P}(X = x \cap Y = y) = \mathrm{P}(X = x)\mathrm{P}(Y = y) = p_x(x)p_y(y).$$

We can say the following about the expectations and variances of independent variables:

$$\mathrm{E}[X * Y] = \mathrm{E}[X] * \mathrm{E}[Y]$$

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

These facts are proved in the notes.

### 2.2  Questions

1. Say you're playing a game with a coin and die, where you flip the coin 3 times and roll the die once. In this game, your score is given by the number of heads that show multiplied with the die result. What is the expected value of your score? What's the variance?

2. Two friends have n books they would like to read. Each friend independently picks a random permutation of the books and reads them in that order, one book per week (for n consecutive weeks). Let X be the number of weeks in which both friends are reading the same book. Compute Var(X).

3. Define a probability distribution over a set of events $S$ to be *uniform* if the probability of each event in the set occurring is equal. Find a joint distribution on $X$ and $Y$ such that $X$ and $Y$ are each uniform on the set $\{1, 2, 3, 4\}$, but $(X, Y)$ is not uniform on the set $\{1, 2, 3, 4\} \times \{1, 2, 3, 4\}$.

4. Mr. and Mrs. Brown decide to continue having children until they either have their first girl or until they have three children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let G denote the numbers of girls that the Browns have. Let C be the total number of children they have.

   (a) Determine the sample space, along with the probability of each sample point.

   (b) Compute the joint distribution of G and C. Fill in the table below.

|       | C = 1 | C = 2 | C = 3 |
|-------|-------|-------|-------|
| G = 0 |       |       |       |
| G = 1 |       |       |       |

   (c) Use the joint distribution to compute the marginal distributions of G and C and confirm that the values are as you'd expect. Fill in the tables below.

| P(G = 0) |   |
|----------|---|
| P(G = 1) |   |

| P(C = 1) | P(C = 2) | P(C = 3) |
|----------|----------|----------|
|          |          |          |

   (d) Are G and C independent?

   (e) What is the expected number of girls the Browns will have? What is the expected number of children that the Browns will have?

## 3    Covariance

### 3.1    Introduction

The **covariance** of two random variables $X$ and $Y$ is defined as:

$$\text{Cov}(X, Y) := \text{E}((X - \text{E}(X)) \cdot (Y - \text{E}(Y)))$$

### 3.2    Questions

1. Prove that $\text{Cov}(X, Y) = \text{E}[XY] - \text{E}[X]\text{E}[Y]$.

2. Prove that $\text{Cov}(X, X) = \text{Var}(X)$.

3. (a) Prove that if $X$ and $Y$ are independent, then $\text{Cov}(X, Y) = 0$.

   (b) Prove that the converse is not necessarily true. In other words, give an example of $2$ random variables whose covariance is $0$ but are not independent.

4. Roll 2 dice. Let $A$ be the number of 6's you get, and $B$ be the number of 5's, find $\text{Cov}(A, B)$.