The Second Moment Method

This will be a VERY quick and brief discussion of the second moment method.

We warm up with what is called the first moment method. Let X be an nonnegative integer valued random variable. Let's assume that you want to show that $P(X \geq 1)$ is quite small. This is quite easy to do by showing that $E[X]$ is small since

$$P(X \geq 1) \leq E[X]$$

(why?). So the " first moment method" is quite simple.

Let's now assume that you want to show that $P(X \geq 1)$ is large (meaning close to 1 or not so close to 0). Is it enough to show that $E[X]$ is large?

(Perhaps you want to think about this before reading on!!).

The answer is no. An example which shows this is the following. $P(X = 0) = 1 - 1/n$ and $P(X = n^2) = 1/n$. In this case, the expected value of X is n which is "quite big" but $P(X \geq 1)$ is only $1/n$ which is quite small. So, the point of this example is that we cannot conclude that $P(X \geq 1)$ is quite large if we just know that its expected value is large.

However, one can conclude that $P(X \geq 1)$ is quite large if we know that its expected value is large AND if we know that the second moment of X is not so large.

This all sounds like philosphy but it will become more concrete as we go. In fact here, is a very important inequality and it is this inequality which is sometimes called the "second moment method".

$$P(X > 0) \geq \frac{E[X]^2}{E[X^2]}. \ (*)$$

Of course, this does tell that $P(X \geq 1)$ is quite large if we know that its expected value is large AND if we know that the second moment of X is not so large.

The proof of this follows from the so-called Cauchy-Schwartz inequality which says that

$$E[XY] \leq E[X^2]^{\frac{1}{2}} E[Y^2]^{\frac{1}{2}}.$$

for any random variables X and Y.

Exercise: Prove $(*)$ using the Cauchy-Schwartz and note that $(*)$ is still true even if X is not integer valued but it does have to be nonnegative.

Sometimes the Cauchy-Schwartz inequality instead of $(*)$ is called "the second moment method".

Once you know $(*)$, you can start to use it to prove different things (as we will see) but right now, I will continue with some more theoretic observations.

You all know that for any random variable X,

$$E[X^2] = E[X]^2 + Var(X).$$

You should also know that this is exactly Pythagorean's theorem interpreted correcting.

In any case, $E[X^2]$ is made up of two parts, namely $E[X]^2$ and $Var(X)$. Which of these latter two is the larger component making up $E[X^2]$ is of crucial importance and is what determines whether the second moment method will work.

Vaguely speaking, the second moment method will fail if most of $E[X^2]$ is made up of the second piece $Var(X)$ (which is equivalent to $Var(X)$ being much bigger than $E[X]^2$) and the second moment method will work otherwise.

We will look at two examples. We first return to our example where $P(X = 0) = 1 - 1/n$ and $P(X = n^2) = 1/n$. We already saw that $E[X]$ was very big but $P(X > 0)$ was very small. Therefore, by the above discussion, it should be the case that $Var(X)$ is much bigger than $E[X]^2$. We now check this.

In this case $E[X^2] = n^3$ (check this!!), $E[X]^2 = n^2$ and therefore we must have $Var(X) = n^3 - n^2$. Clearly $Var(X)$ is much bigger than $E[X]^2$ and most of the contribution of $E[X^2]$ is coming from the variance term.

Now we consider an example where the second moment method works.

Let $Y_1, Y_2, \ldots, Y_n$ be i.i.d. each taking values 0 or 1 each with probability

1/2. Think of n as large and let

$$X = \sum_i Y_i.$$

We want to show that $P(X > 0)$ is quite big (of course in this concrete case, this is quite obvious). Just to get a feeling whether this should work, let's look to see which of the terms making up $E[X^2]$ is bigger.

$E[X]^2 = n^2/4$ obviously. We also know that $Var(X) = n/4$ (why?).

So this time (as opposed to the other example), the squared mean term is much bigger than the variance term and so the second moment method should work. Of course from the above, we know that $E[X^2] = n^2/4 + n/4$.

In fact $(*)$ gives us

$$P(X > 0) \geq \frac{E[X]^2}{E[X^2]} = \frac{n^2/4}{n^2/4 + n/4}$$

which is clearly close to 1 if n is large as we wanted to show.