

## Tune-Up: First Moment Method

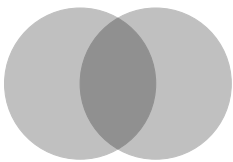
Probabilistic Method is a technique invented in the 20th century. We don't have time to check all the possible answers, we are trying to merely check that an object of a given type exists within our co-hort and maybe of an unusual or rare type, given the information we know.

When we say random, we could ask “How often does event  $[X]$  occur?” The mathematical name for the change of the information we get from learning something is expressed with **conditional probability**.

**Ex** Prove (e.g. using pigeonhole principle)

- $\mathbf{P}[X \geq \mathbb{E}[X]] > 0$  sometimes the unpredictable “chance” thing is greater than we expected.
- $\mathbf{P}[X \leq \mathbb{E}[X]] > 0$  sometimes the unpredictable “chance” thing is less than what we expected.

How are we able to extrapolate such meaningful arguments from such “duh!” starting points? Probability is interesting since we never get to know what the correct probability measure should have been. Probability theory doesn't even make sense if the event only happens *once*.



*Proof* In general,  $\mathbf{P}[A \cup B] \leq \mathbf{P}[A] + \mathbf{P}[B]$ . Let's test our particular case:

$$1 = \mathbf{P}[(X \leq \mathbb{E}[X]) \cup (X \geq \mathbb{E}[X])] \leq \mathbf{P}[X \leq \mathbb{E}[X]] + \mathbf{P}[X \geq \mathbb{E}[X]]$$

Uh...

$1 \geq a \geq 0$  and  $1 \geq b \geq 0$  and  $a + b \geq 1$  we'd like to conclude  $a > 0$  and  $b > 0$ .

Well...

If  $b \geq 1 - a = 0$  then  $a = 1$ . In our case,  $\mathbf{P}[X \leq \mathbb{E}[X]] = 0$  and  $\mathbf{P}[X \geq \mathbb{E}[X]] = 1$ .

That looks fine...

$\mathbf{P}[X > \mathbb{E}[X]] = 1$  and yet  $\mathbf{P}[X \geq \mathbb{E}[X]] = 1$ . I'm guessing  $\mathbf{P}[X = \mathbb{E}[X]] = 0$ .

Why can't  $\mathbf{P}[X < \mathbb{E}[X]] = 0$ ?

Our value is *never* less than what we expected. Then maybe we should set our expectations higher?

We always have  $X > \mathbb{E}[X]$  then

$$\mathbb{E}[X] \leq \sum_{X \leq \mathbb{E}[X]} x \mathbb{P}[x] + \sum_{X \geq \mathbb{E}[X]} x \mathbb{P}[x] = 0 + \sum_{X > \mathbb{E}[X]} x \mathbb{P}[x]$$

if we always have  $X > \mathbb{E}[X]$  then

$$\mathbb{E}[X] \leq \sum_{X \leq \mathbb{E}[X]} x \mathbb{P}[x] + \sum_{X \geq \mathbb{E}[X]} x \mathbb{P}[x] = \sum_{X < \mathbb{E}[X]} x \mathbb{P}[x] + 0 < \mathbb{E}[X] \mathbf{P}[X < \mathbb{E}[X]]$$

if we have that  $\mathbb{E}[X] > 0$  then we conclude that  $\mathbf{P}[X < \mathbb{E}[X]] > 1$ . Unfortunately we do not have this assumption.  $\emptyset$

*Proof* How about we have that  $\mathbb{E}[X - \mathbb{E}[X]] = 0$ , this is called “centering”.

$$0 = \mathbb{E}[X - \mathbb{E}[X]] = \sum_{X \leq \mathbb{E}[X]} (x - \mathbb{E}[X]) \mathbf{P}[x] + \sum_{X \geq \mathbb{E}[X]} (x - \mathbb{E}[X]) \mathbf{P}[x]$$

By definition of our idea of “probability”  $\mathbf{P}[x] \geq 0$  and yet the total of these numbers is zero. Therefore some term is positive and some term is negative, we have  $x > \mathbb{E}[X]$  and  $x < \mathbb{E}[X]$  in some case.

Cleanup? Maybe using conditional probability:

$$\begin{aligned} 0 &= \mathbb{E}[X - \mathbb{E}[X]] \\ &= \mathbb{E}[X - \mathbb{E}[X] | X \leq \mathbb{E}[X]] \mathbf{P}(X \leq \mathbb{E}[X]) \\ &\quad + \mathbb{E}[X - \mathbb{E}[X] | X \geq \mathbb{E}[X]] \mathbf{P}(X \geq \mathbb{E}[X]) \\ &= a \mathbf{P}(X \leq \mathbb{E}[X]) + b \mathbf{P}(X \geq \mathbb{E}[X]) \end{aligned}$$

where  $a \leq 0$  and  $b \geq 0$ . Yet:

$$\mathbf{P}(X \leq \mathbb{E}[X]) + \mathbf{P}(X \geq \mathbb{E}[X]) \geq 1$$

$$0 = a\mathbf{P}(X \leq \mathbb{E}[X]) + b\mathbf{P}(X \geq \mathbb{E}[X]) \geq a\mathbf{P}(X \leq \mathbb{E}[X]) + b(1 - \mathbf{P}(X \leq \mathbb{E}[X]))$$

Then maybe if we arrange more carefully:

$$0 \geq (a - b)\mathbf{P}(X \leq \mathbb{E}[X]) + b$$

Let's try one more:

$$\mathbf{P}(X \leq \mathbb{E}[X]) \geq \frac{b}{b - a}$$

Maybe easier:

$$(-a)\mathbf{P}(X \leq \mathbb{E}[X]) = b\mathbf{P}(X \geq \mathbb{E}[X])$$

Let's rule out  $a = 0$  and  $b = 0$ . If  $a = 0$ , either

- $\mathbf{P}(X \leq \mathbb{E}[X]) = 0$  and  $\mathbf{P}(X \geq \mathbb{E}[X]) = 1$
- $b = 0$

$$\mathbb{E}[X|X \leq \mathbb{E}[X]] - \mathbb{E}[X] = \mathbb{E}[X - \mathbb{E}[X]|X \leq \mathbb{E}[X]] = 0$$

Looks like we need to consider the event  $\mathbb{P}[X = \mathbb{E}[X]] > 0$  that there is probability concentrated at one point .

Another great option is ...our world doesn't stop because we find one or two or even a hundred contradictions. Our logic keeps going our culture keeps going and we keep existing. How do we humans reason in such murky and contradictory environment?

01/22

*Hints:*

- $\mathbb{E}[X|X > 0] > 0$
- $\mathbf{P}(X > 0) + \mathbf{P}(X = 0) + \mathbf{P}(X < 0) = 1$
- Let  $\mathbb{E}[X] = 0$  then  $0 = \mathbb{E}[X|X > 0] \mathbf{P}(X > 0) + 0 + \mathbb{E}[X|X < 0] \mathbf{P}(X < 0)$
- $1 > 0$

We are showing that:

- $\mathbf{P}(X > 0) + \mathbf{P}(X = 0) > 0$  or
- $\mathbf{P}(X < 0) < 1$

In fact we have a sort of ultimatum:

- $\mathbf{P}(X > 0) = 0$  and  $\mathbf{P}(X < 0) = 0$  (and then  $\mathbb{E}[X] = 0$ )
- $\mathbf{P}(X > 0) > 0$  and  $\mathbf{P}(X < 0) > 0$  (and then  $\mathbf{P}(X \geq 0) > 0$ )

So ...

We are relying heavily on that  $X$  and  $\mathbb{E}[X]$  are *numbers* (and their properties) and that  $X$  is taking values in  $\mathbb{R}$ . As math major and a scientist there is no other way.

There's a recursive nature about this problem, turning into itself which sort of clumsy. There's asymmetry between these kinds of relations (" $>$ " vs. " $\geq$ ") and (" $>$ " vs. " $<$ ").

There's a question about **mathematical proof** and use randomly putting together formulas until we reach a result. The end proof is a sort of a "tree" as well.

...

Why go through such a headache? I've seen first-moment method use non-trivially before. It's likely also a description of how we actually *solve* problems in general. There's also *second-moment* method. Pick one at random, let the process run and we find some statistic to show that our event occurs with positive chance.

**Answer Key Markov's Inequality.** Let  $X$  be a non-negative random variable. Then for any positive real  $\lambda > 0$ :

$$\mathbf{P}(X \geq \lambda) \leq \frac{\mathbb{E}(X)}{\lambda}$$

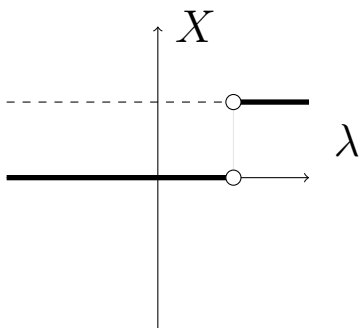
Proof: we have the trivial inequality.  $X \geq \lambda \mathbf{I}(X \geq \lambda)$ .

- if  $X < \lambda$  the inequality says  $X \geq 0$
- if  $X \geq \lambda$  it says  $X \geq \lambda$

Indeed if we take expectation of both sides, by **linearity of expectation** we get the result.

$$\mathbb{E}[X] \geq \lambda \mathbf{P}(X \geq \lambda)$$

Maybe if we drew a picture, the shape of this random variable could be more clear:



There's something self-referential about this pure mathematical framework. These number systems are limits of what we measure in "real life". The notion of **5** or "five" as "the thing after 4" ... numbers are recursively defined.

**Example: Sum Free Sets** Let  $A$  be an additive set of non-zero integers, then  $A$  contains a sum-free subset of  $B$  of size  $|B| > |A|/3$ .

One possibility for this theorem is in topics of **communication complexity** e.g. over telephone lines, or wireless communication. Literally we our phones and laptops and computers reconstruct sound and video over disturbances of thin air. So we are led to problems in **Ramsey theory**. Our very attempts budget our resources in sending messages leads to dilemmas of information compression, such as **multi-party communication** and **discrepancy theory**.

## References

- [1] Terence Tao, Van Vu. **Additive Combinatorics** (Cambridge Advanced Studies in Mathematics #105) Cambridge University Press, 2006.
- [2] Stasys Jukna **Extremal Combinatorics with Applications in Computer Science** (Texts in Theoretical Computer Science) Springer, 2001.