Concentration of Measure

Basic Measure concentration results start w/ Markov's inequality.

<u>Thm</u> (Markov) Suppose that $X \geq 0$ is a R.V. w/ $\mathbb{E}|X| < \infty$. For any $a > 0$,

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

Note that another way to phrase this is, if $\mu = \mathbb{E}X$, then

$$\mathbb{P}(X \geq c\mu) \leq \frac{1}{c}$$

I find this to be slightly more informative. Markov doesn't tell you anything about the probability of X being less than its mean. It just tells a bound on how likely X is to be some multiple of its mean.

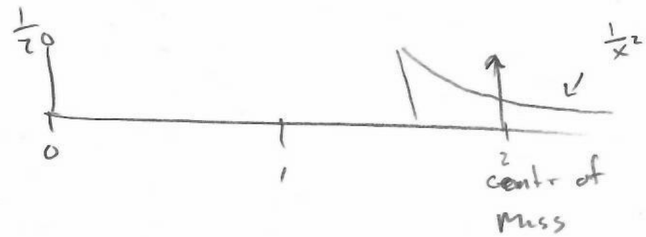
Physically, Markov can be thought of like this: You have a rod that starts at 0 & has center of mass μ.

Markov answers the question: How much mass can be past point a, knowing that (i) the rod is balanced about μ and (ii) the rod has finite length on its left side.

Remark Is Markov's inequality "tight"? i.e., in the sense that there $\exists$ a R.V. s.t.

$$\mathbb{P}(X > a) = \frac{\mu}{a} \quad \forall \; a \geq 1 \; ?$$

As stated, no. If its tight at $a = \mu$, then how can μ be the mean? Moreover, you might consider something like



But this is impossible since

$$\int_b^\infty x \underbrace{p(x)}_{\frac{1}{x^2}} \, dx = \int_b^\infty \frac{1}{x} \, dx = \infty,$$

so anything w/ this density funct has no mean.

Point Markov is a crude bound.

I forgot to prove Markov's ineq. Here's the proof.

pf:

$$\mathbb{E}(X) = \int_{\{X \geq 0\}} x \, d\mathbb{P}(x) \geq \int_{\{X \geq a\}} x \, d\mathbb{P}(x) \geq a \int_{\{X \geq a\}} d\mathbb{P}(x) = a \, \mathbb{P}(X \geq a). \quad \blacksquare$$

Chebychev

The following is an immediate consequence of Markov.

$$\mathbb{P}(|X| \geq a\sigma) \leq \frac{1}{a^2} \quad \text{where } \sigma^2 = \mathrm{Var}(X)$$

pf:

$$\mathbb{P}(|X| \geq a\sigma) = \mathbb{P}(X^2 \geq a^2\sigma^2) \leq \frac{\mathrm{Var}\, X}{a^2\sigma^2} = \frac{1}{a^2} \quad \blacksquare$$

An immediate consequence of this is the following crude concentration bound:

<u>Lemma</u> Suppose $(X_i)_i$ are iid w/ var σ^2 & mean μ. Then

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m} X_i - \mu\right| > \varepsilon\right) \leq \frac{\sigma^2}{m\varepsilon}$$