Jensen's Inequality

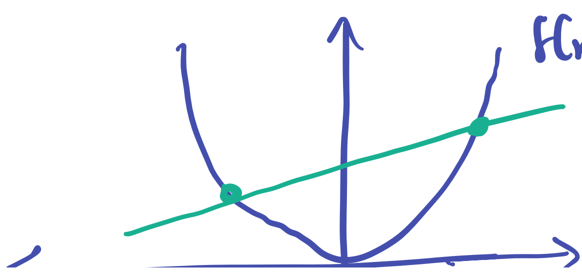
Statement : $\mathbb{E}(f(x)) \geq f(\mathbb{E}(x))$ *

An Implication for probabilities and expectation :

$$\mathbb{E}(X^2) \geq (\mathbb{E}[X])^2$$

This is only one way of interpreting it (although a very useful one!) A more intuitive explanation might be:

Say you have a convex function $f(x)$ (Lets imagine $f(x) = x^2$). Lets pick two points on this function $(x_1, f(x_1))$ and $(x_2, f(x_2))$.



Jensen's inequality states that the line is always above (or in the case where the line meets the curve, equal to) the curve.

My understanding of it is that it is kinda a re-statement of the definition of a convex function.

Definition: A function f is convex if for every x_1 and x_2 and $p \in [0,1]$, we have:

$$p f(x_1) + (1-p) f(x_2) \geq f(p x_1 + (1-p) x_2)$$

Observe that this definition is only a

step away from Jensen's inequality:

$$\mathbb{E}[f(x)] \geq f(\mathbb{E}[x])$$

A very natural question: when is this going to be useful to me? What are situations that I should look out for to use Jensen's?

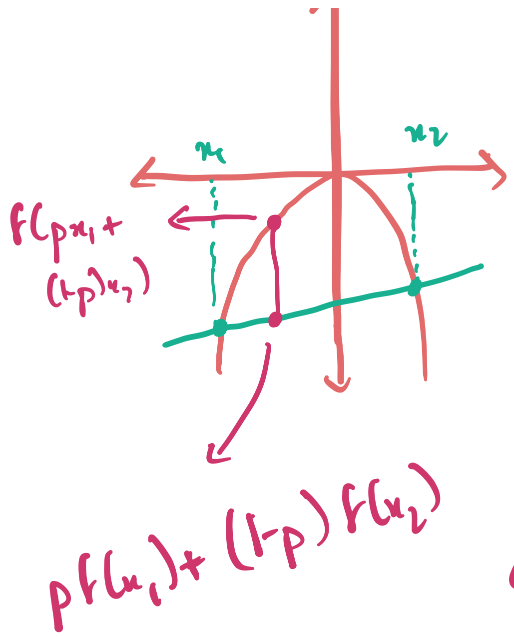
Answer:

Sometimes computing the expectation of a function is hard. In those cases we can easily <u>lower bound</u> the value of the expectation.

Suppose f is concave. How does Jensen's change?

Let's use $f(x) = -x^2$ as

an image for our minds.



The line is always under the curve.

So for any concave function g and points x_1, x_2 and $p \in [0,1]$,

$$pf(x_1) + (1-p)f(x_2) \leq f(px_1 + (1-p)x_2)$$

$$\Rightarrow \mathbb{E}[f(x)] \leq f(\mathbb{E}(x))$$