**Jensen’s inequality Introduction**

Today I want to introduce a very useful method in proof of probability theory---Jensen’s inequality. First, I want to introduce what is a **convex function.**

A function is convex if it is “upward bending”, intuitively. For example:

is a convex function.

More precisely: consider two real numbers and . g is convex if the line between g() and g() stays above the function g.

Consider a function g: , where is an interval in . We say that g is a convex function if, for any two points x and y in and any we have:

Note that is the weighted average of x and y. Also, is the weighted average of g(x) and g(y).

More generally, for a convex function in and nonnegative real numbers such that ++...+= 1, we have:

g(++...+) ≤++...+

If n=2, the above statement is the definition of convex functions.

Or, we can say: if g is (doubly) differentiable then g is convex if and only if ≥ 0.

Let’s move onto Jensen’s inequality.

Remember that the variance of any random variable is a positive. That means:

Var()=−≥ 0

which is equivalent to:

.

Let’s define , then the inequality becomes:

.

Here the function is a convex function, then Jensen’s inequality states that for a convex function g : → we have that:

.

Above all, I believe you have already familiar with Jensen’s inequality. But keep in mind, we need to determine if a function g is convex. That is, check whether the function g is twice-differentiable and second derivative is larger or equal to 0.

Reference:

https://www.probabilitycourse.com/chapter6/6\_2\_5\_jensen's\_inequality.php