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:

$$g(x) = x^2$$

is a convex function.

More precisely: consider two real numbers x_1 and x_2. g is convex if the line between g(x_1) and g(x_2) stays above the function g.

Consider a function g: I → ℝ, where I is an interval in ℝ. We say that g is a convex function if, for any two points x and y in I and any $\alpha \in [0,1]$, we have:

$$g(\alpha x + (1-\alpha)y) \leq \alpha g(x) + (1-\alpha)g(y)$$

Note that $\alpha x + (1-\alpha)y$ is the weighted average of x and y. Also, $\alpha g(x) + (1-\alpha)g(y)$ is the weighted average of g(x) and g(y).

More generally, for a convex function $g\colon \mathrm{I} \to \mathbb{R}, and\ x_1, x_2, \ldots, x_n$ in I and nonnegative real numbers α_i such that α_1+α_2+...+α_n = 1, we have:

$$g(\alpha_1 x_1+\alpha_2 x_2+\ldots+\alpha_n x_n) \leq \alpha_1 g(x_1)+\alpha_2 g(x_2)+\ldots+\alpha_n g(x_n)$$

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 $d^2g/dx^2 \geq 0$.

Let's move onto Jensen's inequality.

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

$$\mathrm{Var(X)=E}X^2-(\mathrm{E}X)^2 \geq 0$$

which is equivalent to:

$$\mathrm{E}X^2 \geq (\mathrm{E}X)^2.$$

Let's define $g(x) = x^2$, then the inequality becomes:

$$\mathrm{E}(g(x)) \geq g(\mathrm{E(x)}).$$

Here the function $g(x) = x^2$ is a convex function, then Jensen's inequality states that for a convex function g : ℝ→ ℝ we have that:

$$\mathrm{E}(g(x)) \geq g(\mathrm{E(x)}).$$

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