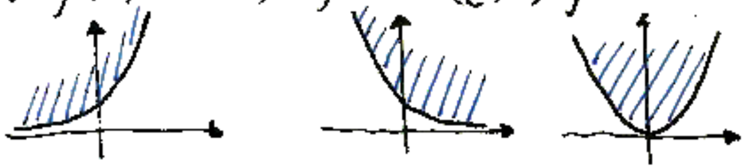
Detour: A Short Proof of Jensen

def. Convex function.
A function $f : \mathbb{R} \to \mathbb{R}$ is said to be convex if the point set $\{(x,y) \in \mathbb{R}^2 : y \geq f(x)\}$ is convex.

e.g. $f(x) = e^x$, $f(x) = \left(\frac{1}{e}\right)^x$, $f(x) = x^2$



Jensen's Inequality.
Let f be a convex function and X be a discrete random variable on $\mathbb{R}$. Then

$$f(\mathbb{E}(X)) \leq \mathbb{E}(f(X)).$$

Proof. ~~Consider the point set $S := \{(x, f(x)) : x \in \mathbb{R}\} \subseteq \mathbb{R}^2$~~

Suppose the support of X is $\{x_1, \dots, x_n\}$ with probability distribution $\{\mu_1, \dots, \mu_n\}$. Consider the point set $S := \{(x_i, f(x_i)) : i \in [n]\} \subseteq \mathbb{R}^2$. Note that

$$(\mathbb{E}(X), \mathbb{E}(f(X))) = \sum_{i=1}^{n} \mu_i \cdot (x_i, f(x_i))$$
$$\in \exp(S) = \operatorname{Conv}(S). \quad ①$$

Also note that by definition,

$$S \subseteq \{(x,y) \in \mathbb{R}^2 : y \geq f(x)\}$$

thus

$$\operatorname{Conv}(S) \subseteq \{(x,y) \in \mathbb{R}^2 : y \geq f(x)\} \quad ②$$

due to the convexity of RHS. From ①② it is immediate that

$$\mathbb{E}(f(X)) \geq f(\mathbb{E}(X)). \quad \blacksquare$$