5. Convex Functions

In the following theorem, the book uses $[0,1]$ with Lebesgue measure for the probability space.

THEOREM 5.1 (Jensen's Inequality). *Let $(X, \mathcal{A}, \mu)$ be a probability space (i.e., $\mu(X) = 1$), and fix an integrable function f on X with range contained in an open interval I. Let φ be a convex function on I. Then*

$$\varphi\left(\int_X f \, d\mu\right) \leq \int_X \varphi \circ f \, d\mu.$$

Note : It may be that $\varphi \circ f$ is not integrable; in this case, the integral on the right equals $+\infty$.

PROOF. Let $y = \int_X f \, d\mu$. Then $y \in I$ since y is a "general average" of the values of f. Another way to see this is to note that if a is the left endpoint of I and a is finite, then $a = a \cdot \mu(X) \leq \int_X f \, d\mu$, and equality means that f takes the value a almost everywhere, which it does not. A similar proof holds for a finite right endpoint of I. Now, if

$$\beta = \sup_{\substack{x<y \\ x \in I}} \frac{\varphi(y) - \varphi(x)}{y - x}$$

then for any $z > y$ in I,

$$\beta \leq \frac{\varphi(z) - \varphi(y)}{z - y}$$

It follows that for $z \geq y$ in I,

$$*) \quad \varphi(z) \geq \varphi(y) + \beta \cdot (z - y).$$

If $z < y$ in I, then by the definition of β,

$$\begin{aligned} \beta &\geq \frac{\varphi(y) - \varphi(z)}{y - z} = \frac{\varphi(z) - \varphi(y)}{z - y}, \text{ so} \\ \varphi(z) - \varphi(y) &\geq (z - y)\beta. \end{aligned}$$

This means that (*) holds for all z in I. For any $t \in X$, we may let $z = f(t)$, and we get

$$\varphi(f(t)) \geq \varphi(y) + \beta(f(t) - y).$$

Recall φ is continuous so $\varphi \circ f$ is measurable. The inequality means that $\varphi(f(t))$ is bounded below by an integrable function. Integrating both sides of the inequality,we get

$$\int_X \varphi \circ f \, d\mu \geq \varphi\left(\int_X f \, d\mu\right) + \beta \cdot 0.$$

■

Example 5.2. A primary example is given by setting $\varphi(x) = e^x$ on $I = \mathbb{R}$. If X is a set consisting of n points with each point having probability $1/n$, then letting x_i be the image under f of the n^{th} point, we have

$$\exp\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{e^{x_1} + \cdots + e^{x_n}}{n}.$$

Putting $y_i = e^{x_i}$, this gives the classical inequality between the arithmetic and geometric mean for positive numbers:

$$(y_1 y_2 \cdots y_n)^{\frac{1}{n}} \leq \frac{y_1 + y_2 + \cdots + y_n}{n}.$$

A generalization works with positive weights α_i such that $\Sigma\alpha_i = 1$. This gives

$$y_1^{\alpha_1} \cdots y_n^{\alpha_n} \leq \alpha_1 y_1 + \cdots + \alpha_n y_n.$$

For a countable number of points, we have

$$\Pi_{n=1}^{\infty} y_i^{\alpha_i} \leq \Sigma_{n=1}^{\infty} \alpha_i y_i.$$

Infinite products are covered in the graduate complex variables course.