In probability calculus we need often to handle sums (discrete variables) or integrals (continuous variable) over the complete support ($\mathcal{X}$) of a probability distribution. Let X be a discrete random variable and $p(X = x)$ its probability mass function (p.m.f.). Assume we need to evaluate the sum

$$\sum_{x \in \mathcal{X}} g(x) p(X = x|\theta), \tag{1}$$

where $g(x)$ is a function of x and θ is the parameter (sometimes vector) defining the distribution shape. For a typical discrete distribution we may rewrite the p.m.f. as the product

$$p(X = x) = c(\theta) k(x, \theta), \tag{2}$$

where $c(\theta)$ is the normalizing constant (for a given θ) of the distribution and $k(x, \theta)$ is the *kernel* function of the p.m.f. (for a given θ). Thus,

$$c(\theta)^{-1} = \sum_{x \in \mathcal{X}} k(x, \theta), \tag{3}$$

as the p.m.f. must sum to unity over the support $\mathcal{X}$. For many functions $g(x)$ that are encountered in the calculation of moments, we may use the following approach in the calculation of (1). Rewrite

$$\sum_{x \in \mathcal{X}} g(x) p(X = x|\theta) = c(\theta) \sum_{y \in \mathcal{Y}} k(y, \theta^*), \tag{4}$$

that is, the product of $g(x)$ and $k(x, \theta)$ is recognized as the kernel function of another distribution in the same family, defined by parameter value θ^*. Since

$$\sum_{y \in \mathcal{Y}} k(y, \theta^*) = c(\theta^*)^{-1}, \tag{5}$$

the sum may easily be evaluated analytically.

The above technique is analogously applicable to a continuous random variable with probability density function (p.d.f.) $f(x|\theta)$. Then we get

$$f(x|\theta) = c(\theta) k(x, \theta) \tag{6}$$

and

$$c(\theta)^{-1} = \int_{-\infty}^{\infty} k(x, \theta) dx, \tag{7}$$

where we assume for simplicity that the distribution is over the real line. The corresponding result may now be written as

$$\int_{-\infty}^{\infty} g(x)f(x|\theta)dx = c(\theta)\int_{-\infty}^{\infty} k(y,\theta^*)dy = c(\theta)c(\theta^*)^{-1}.$$