The Chebychev inequality and the other theorem stated at the end of Chapter 7 are not proved there. Here the theorems are stated and also proved.

Theorem 7.3.2 Let X be a random variable and $g(\cdot)$ a non-negative function on $\mathcal{R}$ such that $\mathrm{E}(g(X))$ exists. Then

$$P(g(X) \geq \ell) \leq \frac{\mathrm{E}(g(X))}{\ell} \quad \forall\, \ell > 0. \tag{7.3.2}$$

Proof of Theorems 7.3.2.

For simplicity, we will assume that the density $f_X(x)$ exists. Define the *indicator functions*:

$$\mathrm{I}(g(x) < \ell) = \begin{cases} 1 & \text{if } g(x) < \ell \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \mathrm{I}(g(x) \geq \ell) = \begin{cases} 1 & \text{if } g(x) \geq \ell \\ 0 & \text{otherwise.} \end{cases}$$

for any $\ell > 0$. Note that $\mathrm{I}(g(x) < \ell) + \mathrm{I}(g(x) \geq \ell) = 1$. Then

$$\begin{aligned} \mathrm{E}(g(X)) &= \int_{-\infty}^{\infty} g(x) f_X(x)\, \mathrm{d}x \\ &= \int_{-\infty}^{\infty} g(x)\mathrm{I}(g(x) < \ell) f_X(x)\, \mathrm{d}x + \int_{-\infty}^{\infty} g(x)\mathrm{I}(g(x) \geq \ell)) f_X(x)\, \mathrm{d}x. \end{aligned}$$

Since $g(x) \geq 0$, the first integral above is non-negative, and so $\mathrm{E}(g(X))$ is no smaller than the second integral, for which we have

$$\int_{-\infty}^{\infty} g(x)\mathrm{I}(g(x) \geq \ell) f_X(x)\, \mathrm{d}x \geq \ell \int_{-\infty}^{\infty} \mathrm{I}(g(x) \geq \ell) f_X(x)\, \mathrm{d}x = \ell\, P(g(X) \geq \ell),$$

and so

$$\mathrm{E}(g(X)) \geq \ell\, P(g(X) \geq \ell) \quad \text{or} \quad P(g(X) \geq \ell) < \frac{\mathrm{E}(g(X))}{\ell}.$$

This completes the proof.

Theorem 7.3.1 Let X be a random variable having finite mean and variance, μ and σ^2. Then, for any $\ell > 0$,

$$P(|X - \mu| \geq \ell\sigma) \leq \frac{1}{\ell^2}. \tag{7.3.1}$$

Proof of Theorem 7.3.1

Notice first that, if a random variable Y is centred, with expectation zero, then its variance is $\mathrm{E}Y^2 = \mathrm{E}|Y|^2$. Notice next that that the variance of $(X-\mu)/\sigma$ is 1, as therefore is the variance of $|(X-\mu)/\sigma|$, which is

$$\int_{-\infty}^{\infty} \Big(\frac{x-\mu}{\sigma}\Big)^2 f_X(x)\, \mathrm{d}x.$$

The inequality $|(x-\mu)/\sigma| \geq \ell$ holds if and only if one of the following two inequalities holds: $(x-\mu)/\sigma \geq \ell$ or $(x-\mu)/\sigma \leq -\ell$, that is, $x \geq \mu + \sigma\ell$ or $x \leq \mu - \sigma\ell$. The integrand in the variance above is non-negative, and so the integral is no smaller than the integral of the same integrand over a subset of the range of integration. The variance is therefore no smaller than

$$\Big[\int_{-\infty}^{\mu-\sigma\ell} + \int_{\mu+\sigma\ell}^{\infty}\Big]\Big(\frac{x-\mu}{\sigma}\Big)^2 f_X(x)\,\mathrm{d}x.$$

which in turn is no smaller than $\ell^2 P(|X-\mu|/\sigma \geq \ell)$. But, since the variance is 1, this leads to the inequality

$$1 \geq \ell^2 P(|X-\mu|/\sigma \geq \ell) \quad \text{or} \quad P(|X-\mu| \geq \ell\sigma) \leq \frac{1}{\ell^2}.$$

This completes the proof.