1.24 Let $k \geq 1$. Show that, for any set of n measurements, the fraction included in the interval $\overline{y} - ks$ to $\overline{y} + ks$ is at least $(1 - 1/k^2)$.

[Hint:

$$s^2 = \frac{1}{n-1}\left(\sum_{=1}^{n}(y_i - \overline{y})^2\right).$$

In this expression, replace all deviations for which $|y_i - \overline{y}| \geq ks$ with ks. Simplify.] This result is known as *Tchebysheff's theorem.*

SOLUTION: Let $A = \{i \,:\, |y_i - \overline{y}| \geq ks\}$ and $B = \{i \,:\, |y_i - \overline{y}| < ks\}$, then

$$\begin{aligned} s^2 &= \frac{1}{n-1}\left[\sum_{i \in A}(y_i - \overline{y})^2 + \sum_{i \in B}(y_i - \overline{y})^2\right] \\ &\geq \frac{1}{n-1}\sum_{i \in A}(y_i - \overline{y})^2 \\ &\geq \frac{1}{n-1}\sum_{i \in A}k^2 s^2 \\ &= \frac{1}{n-1}\cdot |A| \cdot k^2 s^2, \end{aligned}$$

where $|A|$ is the cardinality of A, that is, number of indices in A, that is, the number of measurements outside the interval $|y_i - \overline{y}| < ks$. Since $s \neq 0$ ($s = 0$ if and only if all the y_i's are equal), then we have

$$1 \geq \frac{1}{n-1}k^2|A| > \frac{1}{n}k^2|A|.$$

Therefore,

$$\frac{n - |A|}{n} = 1 - \frac{|A|}{n} \geq 1 - \frac{1}{k^2},$$

which is what we wanted to show.