CHEBYSHEV INEQUALITY
CENTRAL LIMIT THEOREM
and
The Law of Large Numbers
Let $X$ be a RV with $E(X) = \mu$ and $VAR(X) = \sigma^2$. Then, for $d$ being a number:

$$\mathcal{P}(|X - \mu| \geq d) \leq \frac{\sigma^2}{d^2}, \text{ or } \mathcal{P}(|X - \mu| \geq k \cdot \sigma) \leq \frac{1}{k^2}.$$ 

Verification:

$$\sigma^2 = \sum_{\text{all } x} (x - \mu)^2 \mathcal{P}(x) \geq \sum_{(x: |x - \mu| \geq d)} (x - \mu)^2 \mathcal{P}(x) \geq \sum_{(x: |x - \mu| \geq d)} d^2 \mathcal{P}(x) = d^2 \mathcal{P}(|X - \mu| \geq d) \text{ q.e.d.}$$
GIVEN:
a sequence of $n$ INDEPENDENT RANDOM VARIABLES $X_i$.
These RVs follow (unknown but of the same type) distributions with
parameters:
$E\{X_i\} = \mu_i$; $\text{VAR}\{X_i\} = \sigma_i^2$.
THEN: The RV:

$$S = \sum_{i=1}^{n} X_i \text{ has } E\{S\} = \sum_{i} \mu_i \quad \text{VAR}(S) = \sum_{i} \sigma_i^2$$

and for $n \to \infty$ we have:

$$S - \sum_{i} \mu_i \quad \sqrt{\sum_{i} \sigma_i^2} \to N(0, 1)$$
If all the $X_i$ variables are ‘the same’:

$$\mu_i \equiv \mu; \quad \sigma_i \equiv \sigma$$

the RV $S = \sum_{i}^{n} X_i$ has the expected value $E(S) = n\mu$, and

$$VAR(S) = \sum_{i}^{n} \sigma_i^2 = n\sigma^2$$ so we have:

$$\frac{S - n\mu}{\sqrt{n\sigma^2}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \to N(0, 1)$$

**If we are sampling from a population with unknown distribution, either finite or infinite, the sampling distribution of $\bar{X}$ will be approximately normal with mean $\mu$ and variance $\sigma^2/n$ provided that the sample size is large.**
CENTRAL LIMIT THEOREM ...

$X$

$10Y = \sum_{i=1}^{10} X_i$

$2Y = X_1 + X_2$

$30Y = \sum_{i=1}^{30} X_i$

$3Y = X_1 + X_2 + X_3$

$50Y = \sum_{i=1}^{50} X_i$

$5Y = X_1 + X_2 + X_3 + X_4 + X_5$

$100Y = \sum_{i=1}^{100} X_i$

note the different scaling of the $x$-axis
CENTRAL LIMIT THEOREM...

$10Y = \sum_{i=1}^{10} X_i$

$2Y = X_1 + X_2$

$25Y = \sum_{i=1}^{25} X_i$

$5Y = X_1 + X_2 + X_3 + X_4 + X_5$

$50Y = \sum_{i=1}^{50} X_i$

note the different scaling of the $x$-axis.
The central limit theorem can be interpreted as follows: for the sample size $n \to \infty$ the arithmetic average $\bar{X}_n$ tends more and more closely to the expected value $E(X) = \mu$. Or we can state:

Let $X_1, \ldots, X_n$ denote a sequence of independent Rvs with $E(X_i) = \mu$ and $VAR(X_i) = \sigma^2$. Then for every $d > 0$:

$$\lim_{n \to \infty} P(|\bar{X}_n - \mu| > d) = 0.$$ 

The proof follows immediately from the Chebyshev inequality:

$VAR(X) = \sigma^2$ then $VAR(\bar{X}_n) = \sigma^2/n$. Thus

$$P(|\bar{X}_n - \mu| > d) \leq \frac{\sigma^2/n}{d^2}$$

and

$$\lim_{n \to \infty} P(|\bar{X}_n - \mu| > d) \leq \frac{\sigma^2/n}{d^2} = 0.$$