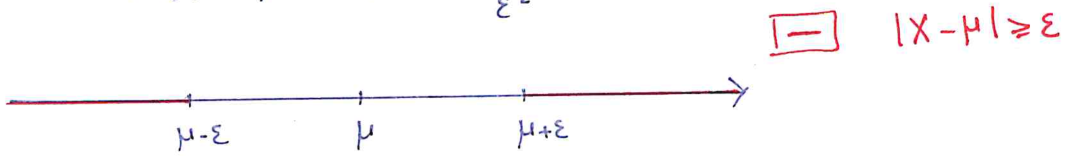
# Law of large numbers

## Chebyshev inequality

Let $X$ be a random variable with mean $\mu$ and let $\varepsilon > 0$ be any positive real number. Then

$$P(|X-\mu| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$




Pf: We prove the inequality for discrete random variables, the continuous case being similar.

Let $f(x)$ be the density function of $X$ and let $A$ be the set of $x$'s such that $|x-\mu| \geq \varepsilon$. (the $x$'s on the red line).

$$P(|X-\mu| \geq \varepsilon) = \sum_{x \in A} f(x).$$

On the other hand, $V(X) = \sum_{\text{all } x} (x-\mu)^2 f(x)$.

If we consider only the values of $x$ in $A$, we get:

$$\sum_{x \in A} (x-\mu)^2 f(x) \leq V(X)$$

Now for all $x \in A$, $(x-\mu)^2 \geq \varepsilon^2$. Hence,

$$\sum_{x \in A} \varepsilon^2 f(x) \leq \sum_{x \in A} (x-\mu)^2 f(x) \leq V(X)$$

Therefore, $\varepsilon^2 \sum_{x \in A} f(x) \leq V(X) \Rightarrow \sum_{x \in A} f(x) \leq \frac{V(X)}{\varepsilon^2}$.

Hence, $P(|X-\mu| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$.

For the continuous case, we consider integrals instead of sums. $\square$

Ex: Let $\varepsilon = K\sigma$, where $\sigma^2 = V(X)$, then

$$P(|X-\mu| \geq K\sigma) \leq \frac{\sigma^2}{K^2\sigma^2} = \frac{1}{K^2}.$$

If $K=2$, then only 25% of the value are at least $\mu + 2\sigma$ or at most $\mu - 2\sigma$. In other word, 75% of the values are within 2 standard deviation of the mean.

## Law of large numbers

Let $X_1, \ldots, X_n$ be independent random variables with finite expected value $\mu = E[X_j]$ and finite variance $\sigma^2 = V(X_j)$. Let $S_n = X_1 + X_2 + \cdots + X_n$. Then for any $\varepsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \longrightarrow 0$$

as $n \to \infty$. Or equivalently

$$P\left(\left|\frac{S_n}{n} - \mu\right| \leq \varepsilon\right) \longrightarrow 1$$

as $n \to \infty$.

This means that as $n \to \infty$, $\frac{S_n}{n}$ becoms very close to the mean $\mu$

Pf: $V(S_n) = V(X_1 + \cdots + X_n) = \sum V(X_j) = n\sigma^2$

$\Rightarrow V\left(\frac{S_n}{n}\right) = \frac{1}{n^2} V(S_n) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$.

$E\left(\frac{S_n}{n}\right) = \frac{1}{n} E(S_n) = \frac{1}{n} E(X_1 + \cdots + X_n) = \frac{1}{n} \sum E[X_j] = \frac{n\mu}{n} = \mu$.

By Chebychev inequality

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\,\varepsilon^2}$$

which goes to $0$ as $n \to +\infty$.