Definition 1. *Let variance be defined as*

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Lemma 1. *For arbitrary a,*

$$\sum_{i=1}^{n}(x_i - a)^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 + n(a - \bar{x})^2.$$

Proof.

$$\begin{aligned}
\sum_{i=1}^{n}(x_i - a)^2 &= \sum_{i=1}^{n}\left((x_i - \bar{x}) - (a - \bar{x})\right)^2 \\
&= \sum_{i=1}^{n}\left((x_i - \bar{x})^2 - 2(x_i - \bar{x})(a - \bar{x}) + (a - \bar{x})^2\right) \\
&= \sum_{i=1}^{n}(x_i - \bar{x})^2 - 2\sum_{i=1}^{n}\left(x_i a - x_i\bar{x} - \bar{x}a + \bar{x}^2\right) + \sum_{i=1}^{n}\left(a^2 - 2a\bar{x} + \bar{x}^2\right) \\
&= \sum_{i=1}^{n}(x_i - \bar{x})^2 - 2a\sum_{i=1}^{n}x_i + 2\bar{x}\sum_{i=1}^{n}x_i + 2\bar{x}an - 2\bar{x}^2 n + a^2 n - 2a\bar{x}n + \bar{x}^2 n \\
&= \sum_{i=1}^{n}(x_i - \bar{x})^2 + a^2 n - 2a\bar{x}n + \bar{x}^2 n \\
&= \sum_{i=1}^{n}(x_i - \bar{x})^2 + n(a - \bar{x})^2
\end{aligned}$$

□

Theorem 1. *Let*

$$f(\boldsymbol{x}) = \sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Then for $\boldsymbol{x}$ bounded between m and M, f has sensitivity bounded above by

$$\frac{n-1}{n}(M - m)^2.$$

Proof. Consider database $\mathbf{x}$ with n points and database $\mathbf{x}' = \mathbf{x} \cup \{x_{n+1}\}$. Note that

$$\begin{aligned}
\bar{x}' &= \frac{1}{n+1}\sum_{i=1}^{n+1}x_i \\
&= \frac{n\bar{x} + x_{n+1}}{n+1}. \qquad (1)
\end{aligned}$$

Then,

$$
\begin{aligned}
f(\mathbf{x}') &= \sum_{i=1}^{n}(x_i - \bar{x}')^2 + (x_{n+1} - \bar{x}')^2 \\
&= \sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x}' - \bar{x})^2 + (x_{n+1} - \bar{x}')^2 && \text{(By Lemma 1)} \\
&= f(\mathbf{x}) + n\left(\frac{n\bar{x} + x_{n+1}}{n+1} - \bar{x}\right)^2 + \left(x_{n+1} - \frac{n\bar{x} + x_{n+1}}{n+1}\right)^2 && \text{(By Equation 1)} \\
&= f(\mathbf{x}) + n\left(\frac{x_{n+1} - \bar{x}}{n+1}\right)^2 + \left(\frac{nx_{n+1} - n\bar{x}}{n+1}\right)^2 \\
&= f(\mathbf{x}) + (x_{n+1} - \bar{x})^2 \frac{n + n^2}{(n+1)^2} \\
&= f(\mathbf{x}) + (x_{n+1} - \bar{x})^2 \frac{n}{n+1} \\
&\leq f(\mathbf{x}) + (M - m)^2 \frac{n}{n+1}
\end{aligned}
$$

Consider database $\mathbf{x}''$, which is the same size as $\mathbf{x}'$ but with a single point changed. Note that

$$|f(\mathbf{x}'') - f(\mathbf{x}')| \leq f(x') - f(x),$$

so the sensitivity of f is bounded by how much it changes from $\mathbf{x}$ to $\mathbf{x}'$, and thus the sensitivity of f is bounded by

$$(M - m)^2 \frac{n}{n+1}.$$

Since we'd really like to consider the sensitivity of database $\mathbf{x}'$, redefine n as $n+1$ to give a sensitivity bounded above by

$$(M - m)^2 \frac{n-1}{n}$$

□

Corollary 1. *Variance has sensitivity bounded above by*

$$\frac{(M - m)^2}{n}.$$