Why is sample variance divided by n-1?

Consider a sample of size n from a population whose mean is μ and variance is σ^2.

The sample has mean $\bar{x}$. This sample mean estimates the population mean μ and the sample variance estimates the population variances σ^2.

ideally, population variance $\sigma^2 = \frac{\sum(x_i-\mu)^2}{N}$

so the sample variance should be

$$s^2 = \frac{\sum(x_i-\bar{x})^2}{n}$$

Subtracting $\bar{x}$ from x_i makes the sum small

i.e $\sigma^2 > s^2$

$\Rightarrow$ s^2 tends to underestimate the true population variance. but dividing by n-1 makes the sum bigger than it would be if we divide it by n.

Mathematically,

$$E\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right) = E\left[\sum(x_i^2+\bar{x}^2-2x_i\bar{x})\right]$$

$$= E\left[\sum x_i^2 + \sum\bar{x}^2 - 2\bar{x}\sum x_i\right]$$

$$= E\left[\sum x_i^2 + n\bar{x}^2 - 2n\bar{x}^2\right] \qquad \because \bar{x} = \frac{\sum x_i}{n}$$

$$= E\left(\sum x_i^2\right) - E(n\bar{x}^2)$$

$$= E\left(\sum x_i^2\right) - nE(\bar{x}^2)$$

Now, $\text{variance}(x_i) = E(x_i^2) - [E(x_i)]^2$

$$\sigma^2 = E(x_i^2) - \mu^2$$

$$\Rightarrow E(x_i^2) = \sigma^2 + \mu^2$$

$\text{Variance}(\bar{x}) = E(\bar{x}^2) - (E(\bar{x}))^2$

$$E(\bar{x}^2) = \frac{\sigma^2}{n} + \mu^2$$

$$\therefore \quad E\left(\sum (x_i - \bar{x})^2\right) = \sum(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)$$

$$= n\sigma^2 + \cancel{n\mu^2} - \sigma^2 + \cancel{n\mu^2}$$

$$= (n-1)\,\sigma^2$$

$$\Rightarrow \sigma^2 = E\left(\frac{\sum (x_i - \bar{x})^2}{n-1}\right)$$

=> dividing by n-1 gives us the true estimate of the population variance.