Notation : population – random variables
(Roman) X, Y, Y_1, X_1, Y_2, X_2
– parameters
(Greek) μ, σ, π

sample (observations / measurements)
→ sample instances of random variables
x, y, y_1, x_1, y_2, x_2

Data (x_i, y_i) $i = 1, 2, 3, \ldots$ $n = 138$

°C

Summary statistics

sample mean of $y = \dfrac{y_1 + y_2 + \ldots \ldots y_n}{n}$

Notation $\bar{y} = \dfrac{\sum_{i=1}^{n} y_i}{n}$ point estimate of μ_Y (population mean)

Sample variance of y

Notation s_y^2 or $s_y^2 = \dfrac{S_{yy}}{n-1}$ ← sum of squares ← degrees of freedom

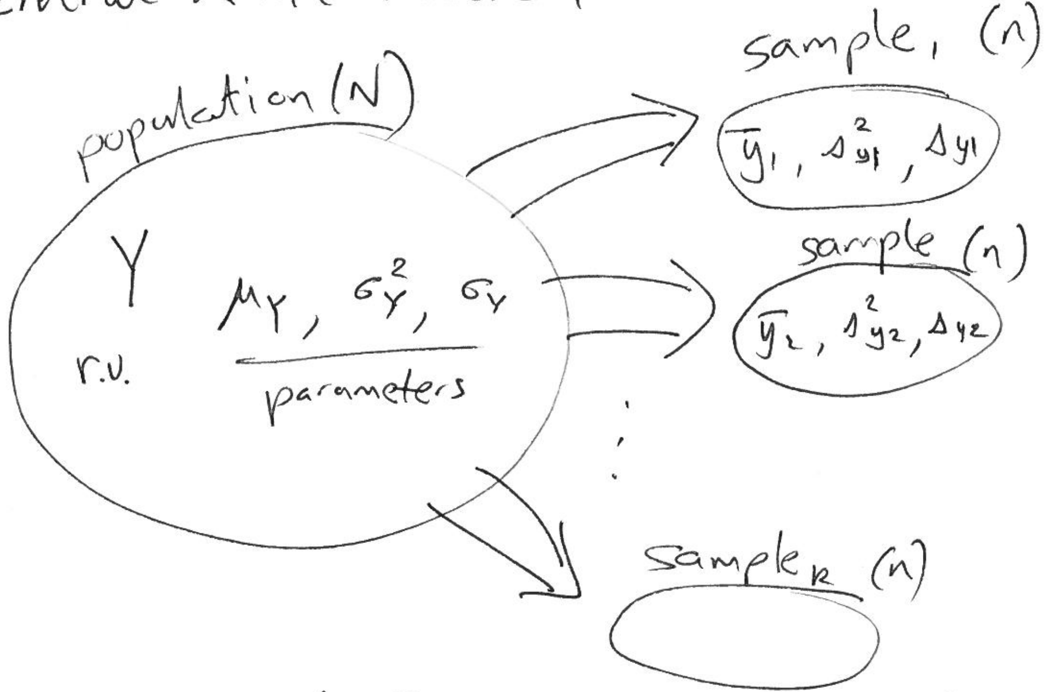
$$= \frac{1}{(n-1)} \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Unbiased point estimate of σ_Y^2 (population variance)

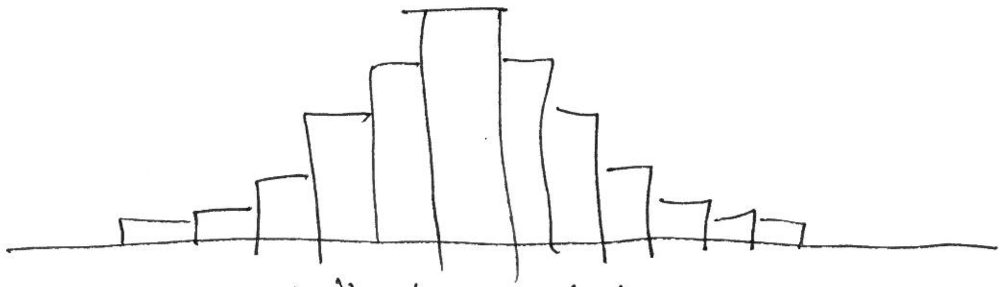
→ we could do the same for x

more often we work with $s_y = \sqrt{s_y^2}$ — in °C
sample standard deviation

Central Limit Theorem



histogram of $\bar{y}_1, \bar{y}_2, \bar{y}_3 \ldots$ (sample means under repeated sampling)



sampling distribution of the mean

CLT: if we are sampling from a population with mean μ_Y & variance σ_Y^2 then as $n \to \infty$ the sampling distribution of the mean will

$$\longrightarrow N\left(\mu_Y, \frac{\sigma_Y^2}{n}\right)$$

standard deviation of the sampling distribution is $\frac{\sigma_Y}{\sqrt{n}}$

standard error