Handout 02: Sample Variance

As on the previous handout, consider a random sample of size n $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{G}$ from a distribution $\mathcal{G}$ with a population mean of μ_X and population variance of σ_X^2.[1] Last time we looked at the sample mean. Today, we define the **sample variance**, a sample statistic defined and denoted by:

$$S_X^2 = \frac{1}{n-1} \times \sum_{i=1}^{n} \left[X_i - \bar{X} \right]^2 .$$

The sample mean $\bar{X}$ and sample variance S_X^2 are independent random variables.[2] As suggested by the name, the sample variance can be used as a good point estimator for the population variance. On the worksheet, we will show that it is an unbiased estimator of the population variance. We will also find its exact distribution when $\mathcal{G}$ is normally distributed.

[1] I am intentionally repeating much of the terminology from the first handout. If any of the statements here are unclear, make sure to go back and lookup what each term means.

[2] We will not prove this as the proof is a bit complex (it requires a n-dimenisional change of variabels formula) and not particularly insightful.