Independence & Simple Random Samples,   Mean & Variance of a Sample Mean

---

Some definitions (taken from Section 1.1 in your text):

A **population** is the entire collection of objects or outcomes about which information is sought.

A **sample** is a subset of a population, containing the objects or outcomes that are actually observed.

A **simple random sample** (SRS) of size n is a sample chosen by a method in which each collection of n population items is equally likely to make up the sample, just as in a lottery.

When a simple random sample of numerical values is drawn from a population, each item in the sample can be thought of as a random variable. The items in a simple random sample may be treated as independent, except when the sample is a large proportion (more than 5%) of a finite population.

For our work from here on, unless explicitly stated to the contrary, we will assume this exception has not occurred, so that the values in a simple random sample may be treated as independent random variables.

IN GENERAL:

If $X_1, X_2, \ldots, X_n$ is a simple random sample, then $X_1, X_2, \ldots, X_n$ may be treated as independent random variables, all with the same distribution.

When $X_1, X_2, \ldots, X_n$ are independent random variables, all with the same distribution, it is sometimes said that $X_1, X_2, \ldots, X_n$ are ***independent and identically distributed (i.i.d.)***.

Sample Mean:

If $X_1, X_2, \ldots, X_n$ is a simple random sample from a population with mean μ and variance σ^2, then the **sample mean** is denoted by $\bar{X}$ (read "X-bar") and is given by

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Re-writing the above expression, we see the sample mean $\bar{X}$ is the linear combination

$$\bar{X} = \frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n$$

From this fact, we can think about how to derive the mean and variance of $\bar{X}$:

Suppose X is the population variable with mean μ and variance σ^2. For the simple random sample $X_1, X_2, \ldots, X_n$ from this population,

$X_1, X_2, \ldots, X_n$ all have the same distribution as the population variable X. This means, for example, that for each $i = 1, 2, \ldots, n$, each X_i has mean μ and variance σ^2. It follows that

$$\mu_{\bar{X}} = \mu_{\frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n} = \frac{1}{n}\mu_{X_1} + \frac{1}{n}\mu_{X_2} + \cdots + \frac{1}{n}\mu_{X_n}$$

$$= \underbrace{\frac{1}{n}\mu + \frac{1}{n}\mu + \cdots + \frac{1}{n}\mu}_{there\ are\ n\ of\ these} = (n)\left(\frac{1}{n}\right)\mu = \mu$$

and, since the items in a simple random sample may be treated as independent random variables,

$$\sigma_{\bar{X}}^2 = \sigma^2_{\frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n} = \frac{1}{n^2}\sigma_{X_1}^2 + \frac{1}{n^2}\sigma_{X_2}^2 + \cdots + \frac{1}{n^2}\sigma_{X_n}^2$$

$$= \underbrace{\frac{1}{n^2}\sigma^2 + \frac{1}{n^2}\sigma^2 + \cdots + \frac{1}{n^2}\sigma^2}_{there\ are\ n\ of\ these} = (n)\left(\frac{1}{n^2}\right)\sigma^2 = \frac{\sigma^2}{n}$$

Mean and Variance of a Sample Mean:

If $X_1, X_2, \ldots, X_n$ is a simple random sample from a population with mean μ and variance σ^2, then the sample mean $\bar{X}$ is a random variable with

Mean of $\bar{X}$: $\mu_{\bar{X}} = \mu$

Variance of $\bar{X}$: $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

The standard deviation of $\bar{X}$ is $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

Ex. A process that fills plastic bottles with a beverage has a mean fill volume of 2.013 L and a standard deviation of 0.005 L. A case contains 24 bottles. Assuming that the bottles in a case are a simple random sample of bottles filled by this method, find the mean and standard deviation of the average volume per bottle in a case.