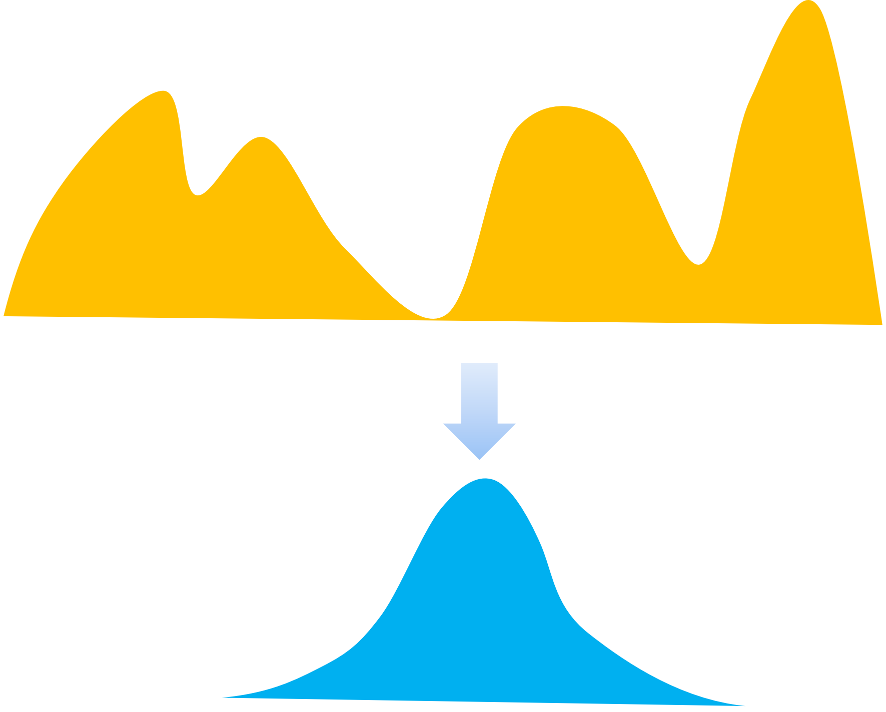
INNOMATICS TECHNOLOGY HUB

A sister concern of

KenMinds
...committed to solution excellence



Sampling Distribution of Means

Sampling Distribution

- The core goal of inferential statistics is to be able to make intelligent conclusions about the population parameters by looking at sample statistics.

Eg: Estimate the mean height of the students in a class, from a small sample.

Sampling Distribution of means

- The sampling distribution of means is what you get if you consider all possible samples of size *n taken from the same* population and form a distribution of their means.

- Each randomly selected sample is an independent observation.

Central Limit Theorem

- http://onlinestatbook.com/2/sampling_distributions/clt_demo.html

- As sample size goes large and number of buckets are high, the means will follow a normal distribution with same mean (μ) and $\frac{1}{n}$ of variance (σ^2).

Using the Central Limit Theorem

Let us say the mean number of Gems per packet is 10, and the variance is 1. If you take a sample of 30 packets, what is the probability that the sample mean is 8.5 Gems per packet or fewer?

Using the Central Limit Theorem

We known that $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \mu = 10, \sigma^2 = 1$ and n = 30. We

need the value of P($\bar{X}$ < 8.5) when $\bar{X}$ ~ N(10,0.0333).

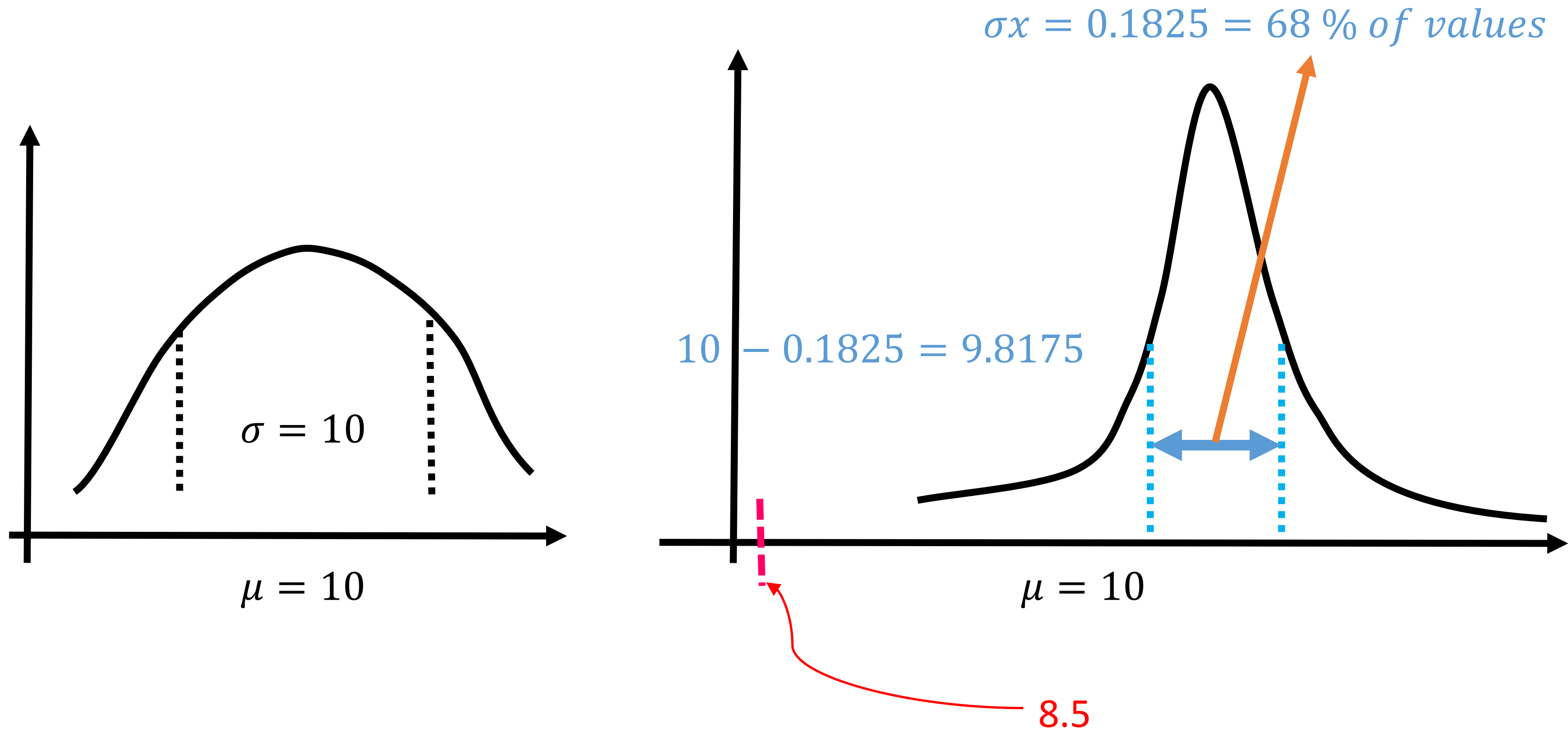
$$Z = \frac{8.5 \; - 10}{\sqrt{0.0333}} = -8.22$$

P(Z < z) = P(Z < -8.22)

This doesn't exit in probability tables. What does it mean ?

Using the Central Limit Theorem

How do we visualize it ?

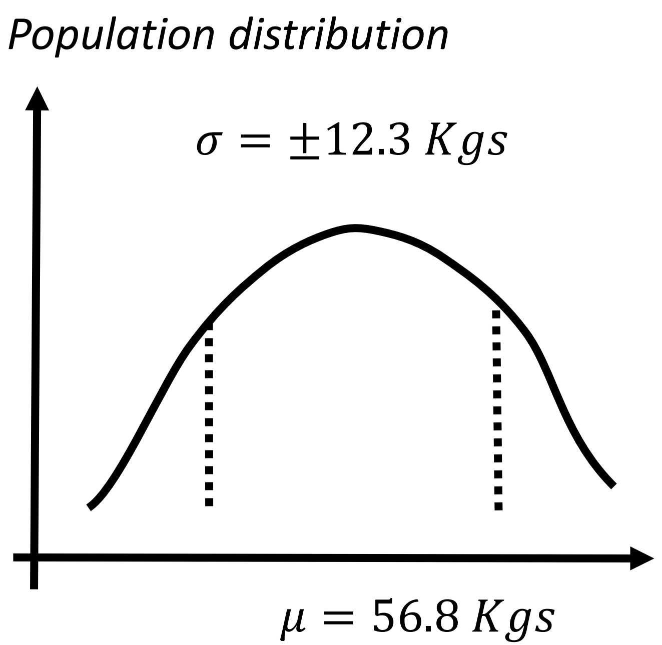


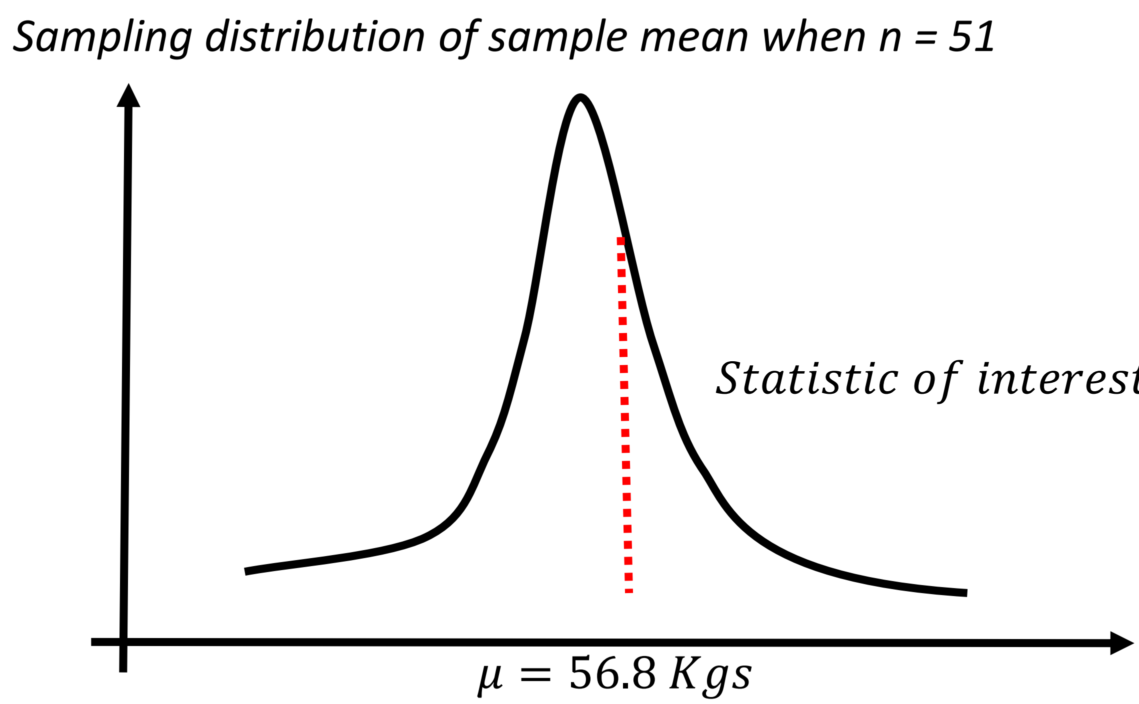
Using the Central Limit Theorem

The Aluminum Association of America reports that the average American household uses 56.8 Kgs of aluminum in a year.

A random sample of 51 households is monitored for one year to determine aluminum usage. If the population standard deviation of annual usage is 12.3 Kgs, what is the probability that the sample mean will be > 60 Kgs?

Sampling Distribution



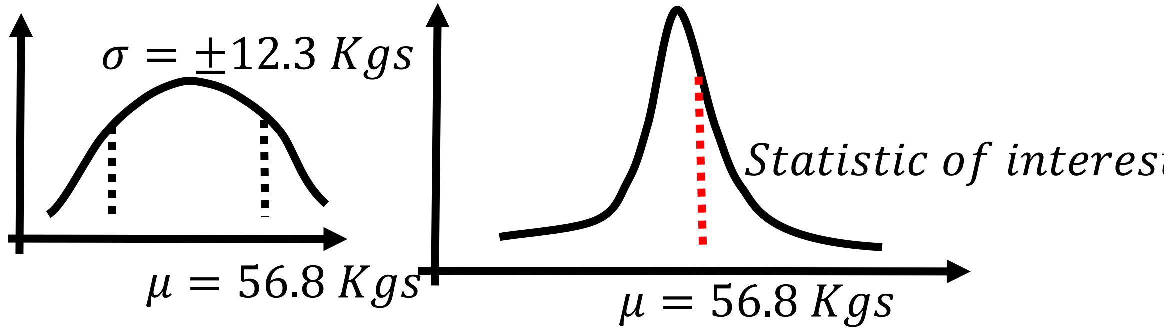


Step 1 : List all known parameters and values

Step2 : Calculate others, or estimate if cannot be calculated

Step3: Find probabilities using tables, Excel or R

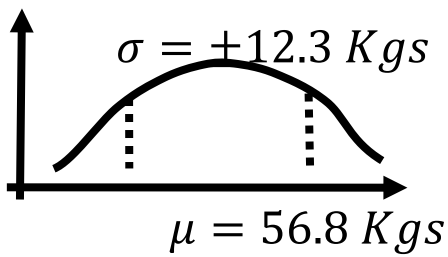
Sampling Distribution

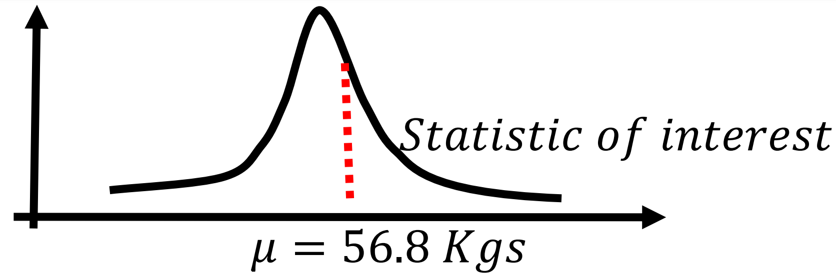


Step-1 : List all known parameters and values

- Population mean, μ= 56.8 Kgs
- Population standard deviation, σ = 12.3 Kgs
- Sample size, n = 51
- Sample mean, $\bar{x}$= > 60 Kgs
- Mean of sample means, $\mu_{\bar{x}}$= μ = 56.8 Kgs

Sampling Distribution





Step-2: Calculate others or estimate, if cannot be calculated.

- Standard deviation of sample means, $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{12.3}{\sqrt{51}} = 1.72$
- $\therefore Z = \frac{60 - 56.8}{1.72} = 1.86$

Step-3: Find probabilities using tables, Excel or R

→ Please calculate these for

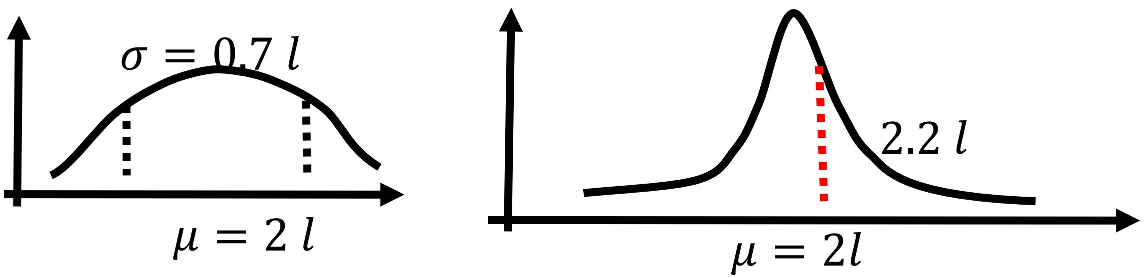
- > 58 Kgs
- > 56 Kgs < 57 Kgs
- < 50 Kgs

Sampling Distribution

The average male drinks $2l$ of water when active outdoors with a standard deviation of $0.7l$. You are planning a trip for 50 men and bring $110l$ of water. What is the probability that you will run out of water?

μ = 2, σ = 0.7

$P(run\ out) \Rightarrow P(use > 110l) \Rightarrow P(average\ water\ use\ per\ male > 2.2l)$



Sampling Distribution

$$\mu_{\overline{x}} = \mu = 2l, \sigma_{\overline{x}}^2 = \frac{\sigma^2}{n} = \frac{0.49}{50}$$

$$\Rightarrow \sigma_{\overline{x}} = 0.099$$

$$\Rightarrow z = \frac{2.2 - 2}{0.099} = 2.02$$

$$\Rightarrow P(\bar{X} < 2.02) = 0.9783$$

$\Rightarrow$The probability of running out of water is

1 – 0.9783 = 0.0217 or 2.17 %

Using the Central Limit Theorem

You sample 36 apples from your farm's harvest of 200,000 apples. The mean weight of the sample is 112g with a 40g sample standard deviation. What is the probability that the mean weight of all 200,000 apples is between 100 and 124g?