Distribution of the sample mean

Definition 2.3

The distribution of a statistic is called a sampling distribution.

Definition 2.4

A set of random variables $Y_1, Y_2, \ldots, Y_n$ is said to be a random sample of size n from a population with distribution $F_Y(y)$ if the Y_i are independently and identically (iid) with distribution function $F_Y(y)$.

A random sample is a set of random variables that satisfies two conditions:

(1) they are independent

(2) they have the same distribution

Example 2.1

A random sample of size 2 is taken from a population of $\{1, 3, 4, 6\}$. Determine the sampling distribution of

(a) the maximum

(b) the mean

(c) the variance

In this case, we have $4 \times 4 = 16$ possible random samples (with equal probability). These are listed in the table on the next slide.

Example 2.1 Solution

Sample	Probability	Max	Mean	Variance
{1, 3}	1/16	3	2.0	2.0
{1, 4}	1/16	4	2.5	4.5
{1, 6}	1/16	6	3.5	12.5
{3, 4}	1/16	4	3.5	0.5
{3, 6}	1/16	6	4.5	4.5
{4, 6}	1/16	6	5.0	2.0
{3, 1}	1/16	3	2.0	2.0
{4, 1}	1/16	4	2.5	4.5
{6, 1}	1/16	6	3.5	12.5
{4, 3}	1/16	4	3.5	0.5
{6, 3}	1/16	6	4.5	4.5
{6, 4}	1/16	6	5.0	2.0
{1, 1}	1/16	1	1.0	0.0
{3, 3}	1/16	3	3.0	0.0
{4, 4}	1/16	4	4.0	0.0
{6, 6}	1/16	6	6.0	0.0

Example 2.1 Solution

The PF of max is as follows:

max	1	3	4	6
Probability	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{6}{16}$	$\frac{6}{16}$

The PF of $\bar{Y}$ is

$\bar{Y}$	1.0	2.0	2.5	3.0	3.5	4.0	4.5	5.0	6.0
Probability	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

The PF of S^2 is

S^2	0.0	0.5	2.0	4.5	12.5
Probability	$\frac{4}{16}$	$\frac{2}{16}$	$\frac{4}{16}$	$\frac{4}{16}$	$\frac{2}{16}$

Sampling distribution of the mean

If $Y_1, Y_2, \ldots, Y_n$ is a random sample from a population with mean μ and variance σ^2, then the sample mean $\bar{Y}$ has a distribution with

- mean $E(\bar{Y}) = \mu$
- variance $\mathrm{var}(\bar{Y}) = \frac{\sigma^2}{n}$

Note that this applies regardless of the distribution of the population.

Sampling from normal populations

If $X_1, X_2, \dots, X_n$ is a random sample from a $N(\mu, \sigma^2)$, then the sample mean $\boldsymbol{\bar{X}}$ itself has a normal distribution:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Example 2.2

Airline records show that on a certain long-haul route, each passenger plus luggage is well modelled by a normal distribution with mean of 87.3kg and standard deviation of 8.6kg. What is the probability that the mean weight of 400 passengers booked on a B747 aircraft will exceed 88.75kg?

Let Y_i = weight of passenger i with luggage

$Y_i \sim \text{iid } N(87.3, 8.6^2)$

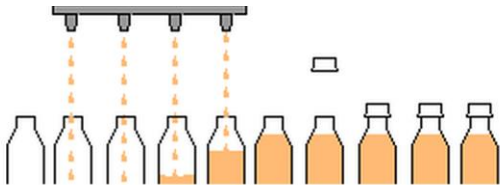
Then $\bar{Y} \sim N\left(87.3, \frac{8.6^2}{400}\right)$

So $P(\bar{Y} > 88.75) = P\left(\frac{\bar{Y}-87.3}{\frac{8.6}{20}} > \frac{88.75-87.3}{\frac{8.6}{20}}\right)$

$= P(Z > 3.37)$

≈ 0.000376

Example 2.3



A bottling machine has been regulated so that its discharges is normally distributed with mean μ milliliters (mL) and standard deviation 10mL. How many samples should we collect if we wish to sample mean of our sample to be within 3mL of μ, with 95% probability?

Let $\bar{Y}$ = sample mean of our sample

$$0.95 = P(|\bar{Y}-\mu| \leq 3)$$
$$= P(-3 \leq \bar{Y}-\mu \leq 3)$$
$$= P\left(-\frac{3}{\frac{\sigma}{\sqrt{n}}} \leq \frac{\bar{Y}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq \frac{3}{\frac{\sigma}{\sqrt{n}}}\right)$$
$$= P(-0.3\sqrt{n} \leq Z \leq 0.3\sqrt{n})$$ where $Z \sim N(0,1)$

We know $0.95 = P(-1.96 \leq Z \leq 1.96)$

So $0.3\sqrt{n} = 1.96$, which gives $n = \left(\frac{1.96}{0.3}\right)^2 \approx 42.68$.

Hence we need 43 samples.