Statistics and Their Distributions

Suppose that several random variables $X_1, X_2, \cdots, X_n$ are associated with the outcome of some experiment.

Then, any function of them, such as their mean

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

is also a random variable. That is, its value is unknown before the experiment is carried out, but is determined by the outcome of the experiment.

Statistic:

Any quantity that can be calculated from the observed values of $X_1, X_2, \cdots, X_n$ without knowing their distributions is called a statistic.

Ex ① Their mean $\bar{X}$ is a statistic.

② If all the X's have expected value μ, that is $E(X_i) = \mu$, $i = 1, 2, \cdots, n$, then $(\bar{X} - \mu)$ is a function of $X_1, X_2, \cdots, X_n$, but it is not a statistic, because we need to know μ, a parameter of their distribution to calculate it.

Sampling Distribution of a Statistic

↳ A statistic is a random variable, and therefore has a probability distribution.

↳ The probability distribution of a statistic is called its <u>sampling distribution</u>.

(*) We use the term sampling distribution to emphasize that statistics vary from sample to sample.

↳ The random variable's $X_1, X_2, \ldots, X_n$ are said to form a (simple) <u>random sample</u> of size 'n'

if

(1) The X_i's are independent RV's

(2) Every X_i has the same prob. distribution.

(*) X_i's are independent (means, independent and identically distributed (iid))

Q A particular brand of dishwasher soap is sold in three sizes: 25z, 40z and 65z. 20% of all purchasers select a 25z box, 50% select a 40z box, and the remaining 30% choose a 65z box. Let X_1, X_2 denotes the package sizes selected by two independently selected purchasers.

a. Determine the sampling distribution of $\bar{X}$, Calculate $E(\bar{X})$, and compare to μ.

b. Determine the sampling distribution of the sample variance S^2, Calculate $E(S^2)$ and compare to σ^2.

Given, $x_1, x_2 \in \{25, 40, 65\}$ with probabilities 0.2, 0.5 and 0.3 resp.

The pmf table is

$p(x_1,x_2)$	x_2: 25	40	65	
x_1: 25	0.04	0.1	0.06	0.2
40	0.1	0.25	0.15	0.5
65	0.06	0.15	0.09	0.3
	0.2	0.5	0.3	

$$p(25,25) = P(X_1 = 25, X_2 = 25)$$
$$= P(X_1 = 25) \cdot P(X_2 = 25)$$
$$= 0.2 \times 0.2 = 0.04$$

(a) possible values for r.v. $\bar{X}$ are

$$\frac{25+25}{2} = 25, \quad \frac{25+40}{2} = 32.5, \quad \frac{25+65}{2} = 45,$$

$$\frac{40+40}{2} = 40, \quad \frac{40+65}{2} = 52.5, \quad \frac{65+65}{2} = 65.$$

$$\therefore P(\bar{X} = 25) = P(X_1 = 25, X_2 = 25) = p(25,25) = 0.04$$

$$P(\bar{X} = 32.5) = P(X_1 = 25, X_2 = 40) + P(X_1 = 40, X_2 = 25) = 0.1 + 0.1$$
$$= 0.2$$

$\bar{x}$	25	32.5	40	45	52.5	65
$p(\bar{x})$	0.04	0.2	0.25	0.12	0.3	0.09

$$\therefore E(\bar{X}) = \sum_{\bar{x} \in S} \bar{x} \cdot p(\bar{x}) = 25 \times 0.04 + 32.5 \times 0.2 + \cdots + 65 \times 0.09$$
$$= 44.5$$

$$\mu = E(X) = 25 \times 0.2 + 40 \times 0.5 + 65 \times 0.3$$
$$= 44.5$$

(b) All the possible values of the sample variance s^2 are

$$\frac{1}{2-1} \sum_{i=1}^{2} (X_i - 25)^2 = (25-25)^2 + (25-25)^2 = 0$$

$$\frac{1}{(2-1)} \sum_{i=1}^{2} (X_i - 32.5)^2 = (25-32.5)^2 + (40-32.5)^2$$
$$= 112.5$$

Let $X_1, X_2, \cdots, X_n$ are observations of a random sample of size 'n' from the ~~normal~~ statistical distribution ~~[illegible]~~

then $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the sample mean of the 'n' observations.

$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is the sample variance of the 'n' observations.

$\frac{1}{2-1}\sum_{i=1}^{2}(X_i - 40)^2 = (40-40)^2 + (40-40)^2 = 0$

$\frac{1}{2-1}\left[(25-45)^2 + (65-45)^2\right] = 800.$

$\frac{1}{2-1}\left[(40-52.5)^2 + (65-52.5)^2\right] = 312.5$

$\therefore\ s^2 \in \{0, 112.5, 312.5, 800\}.$

$\therefore P(s^2 = 0) = p(0) = p(25,25) + P(40,40) + p(65,65)$
$= 0.04 + 0.25 + 0.09 = 0.38.$

$P(s^2 = 112.5) = p(25,40) + p(40,25) = 0.1 + 0.1 = 0.2$

$P(s^2 = 312.5) = P(40,65) + P(65,40) = 0.15 + 0.15 = 0.3$

$P(s^2 = 800) = p(25,65) + p(65,25) = 0.06 + 0.06 = 0.12.$

pmf of s^2

s^2	0	112.5	312.5	800
$P(s^2)$	0.38	0.2	0.3	0.12

$E(s^2) = 0.38 \times 0 + 0.2 \times 112.5 + 0.3 \times 312.5 + 0.12 \times 800$
$= 212.25$

$\therefore$ To calculate σ^2,

$\sigma^2 = (25-44.5)^2 \times 0.2 + (40-44.5)^2 \times 0.5 + (65-44.5) \times 0.3$
$= 212.25.$