Section 5.3 - Statistics and Their Distributions

Population mean and standard deviation are often impossible to calculate.

Def: A ***statistic*** is any quantity whose value can be calculated from sample data.

Prior to obtaining the data, there is uncertainty as to what value of any particular statistic will result. Therefore, a statistic is a random variable and will be denoted by an uppercase letter. A lowercase letter represents the calculated or observed value of the statistic.

$\bar{x}$ $\bar{X}$

Def: The rv's $X_1, X_2, X_3, \ldots, X_n$ are said to form a (simple) ***random sample*** of size n if

1. The X_i's are independent random variables.
2. Every X_i has the same probability distribution.

population mean: μ

$\bar{X}$ sample mean before same taken - RV

$\bar{x}$ particular sample mean

Ex: A certain brand of MP3 player comes in three configurations: 2GB ($80), 4GB ($100), and 8GB ($120). Let X = the cost of a single randomly selected purchase of the MP3 player. Suppose X has pmf given by the table below:

x	80	100	120
$p(x)$	0.2	0.3	0.5

Suppose, on a particular day, 2 of these MP3 players are sold. Let

X_1 = selling price of first MP3 player
X_2 = selling price of second MP3 player

$\}$ independent

Determine the possible values for $\overline{X} = \dfrac{X_1 + X_2}{2}$

X_1	P_{X_1}	X_2	P_{X_2}	$\overline{X}$	$P_{\overline{X}}$	
80	.2	80	.2	80	(.2)(.2) = .04	
80	.2	100	.3	90	(.2)(.3) = .06	.12
100	.3	80	.2	90	(.3)(.2) = .06	
100	.3	100	.3	100	.3 (.3) = .09	.29
80	.2	120	.5	100	.10	
120	.5	80	.2	100	.10	
100	.3	120	.5	110	.15	.3
120	.5	100	.3	110	.15	
120	.5	120	.5	120	.25	

Give the pmf of $\overline{X} = \frac{X_1 + X_2}{2}$

$\bar{x}$	80	90	100	110	120
$p_{\bar{x}}$	.04	.12	.29	.30	.25

Compute the expected value and variance of $\overline{X} = \frac{X_1 + X_2}{2}$

$$E[\bar{X}] = 80(.04) + 90(.12) + 100(.29) + 110(.3) + 120(.25)$$

$$= 106$$

$$V[X] = E[X^2] - (E[X])^2 \approx 122$$