


Contents

Chapter 9 표집분포

9.2 통계량의 확률분포

9.3 표본평균의 분포와 중심극한정리



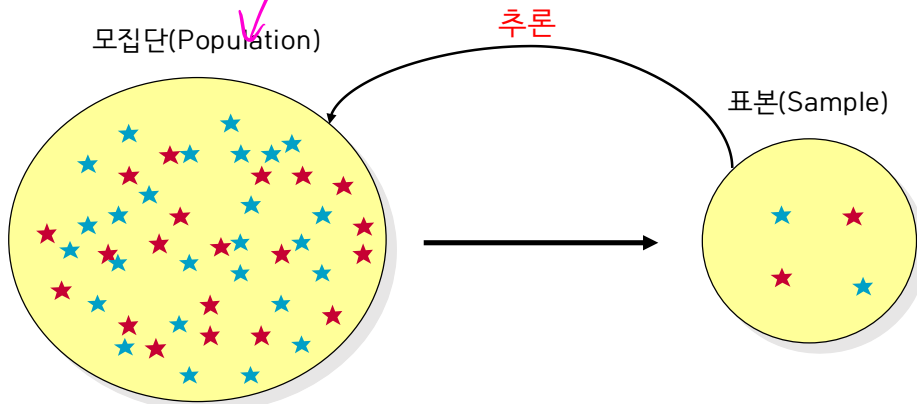
이런 확률 변수에 대한
확률 분포를 return 하는 함수.

Chapter 9

표집분포

- 통계학의 목표

: 관측한 자료(모집단의 일부)를 이용하여 모집단에 대한 추측을 하는 것



< 모집단에 대한 모든 자료를 얻는 것(전수조사)은 현실적으로 불가능 >
대신에 모집단의 일부 추출된 자료인 표본을 이용하여 모수에 대해 추론

(즉, sample을 활용하여)

정보를 통해 $E(X)$, $\sigma(X)$... 를

< 확률분포는 모집단에 관한 것이다.
• $E(X)$, $\sigma(X)$... 도 모집단에
관한 정보들이다. >

Chapter 9 표집분포

01 통계량의 확률분포

- 용어

- 1) 모수 (parameter) : 수치로 표현되는 모집단의 특성 ex. 모평균(μ), 모표준편차(σ)
- 2) 추론 (inference) : 표본으로 부터 모집단의 성격(모수)을 알아내고자 하는 것
- 3) 통계량 (statistic) : 표본(관측한 자료)에 의해 결정되는 양 ex. 표본평균(\bar{X}), 표본분산(S^2)
(=추정량)

- 통계량의 특징

- 1) 통계량의 값은 모수와 통상적으로 같지 않다. ex. $\bar{X} \neq \mu$
- 2) 통계량은 X_1, X_2, \dots, X_n 의 함수이다.
- 3) 통계량은 확률변수이다.

추론함.

• 이후, 모집단의 특성을 분석함.

sample들이 의해 얻어낼 수 있는 정보

애도 확률 변수의 일변함.

=> 그렇기 때문에 $F(\bar{X})$ (표본평균의 평균)이 존재함.

이런 것들

Chapter 9

표집분포

이들이 성립해야, 'random sample'이라 할수있다.

- 통계량의 확률분포 (sampling distribution, 표집분포)

1) 임의표본 (random sample): 서로 독립이고 모집단과 같은 분포를 갖도록 모집단으로 부터 임의로 추출된 표본

2) 표집분포: 통계량의 확률분포

(Ex) 표본평균의 분포

sample의 확률분포.

① 표본 그룹끼리 서로 독립이고
② 모집단과 같은 분포를 갖는다.

모집단 : $X \sim N(\mu, \sigma^2)$

임의표본1

$x_1^{(1)}$
 $x_2^{(1)}$
 \vdots
 $x_n^{(1)}$

$\bar{x}^{(1)}$

임의표본2

$\sim N(\mu, \sigma^2)$
 $x_1^{(2)}$
 $x_2^{(2)}$
 \vdots
 $x_n^{(2)}$

$\bar{x}^{(2)}$

임의표본3

$\sim N(\mu, \sigma^2)$
 $x_1^{(3)}$
 $x_2^{(3)}$
 \vdots
 $x_n^{(3)}$

$\bar{x}^{(3)}$

임의표본4

$\sim N(\mu, \sigma^2)$
 $x_1^{(4)}$
 $x_2^{(4)}$
 \vdots
 $x_n^{(4)}$

$\bar{x}^{(4)}$

임의표본5

$\sim N(\mu, \sigma^2)$
 $x_1^{(5)}$
 $x_2^{(5)}$
 \vdots
 $x_n^{(5)}$

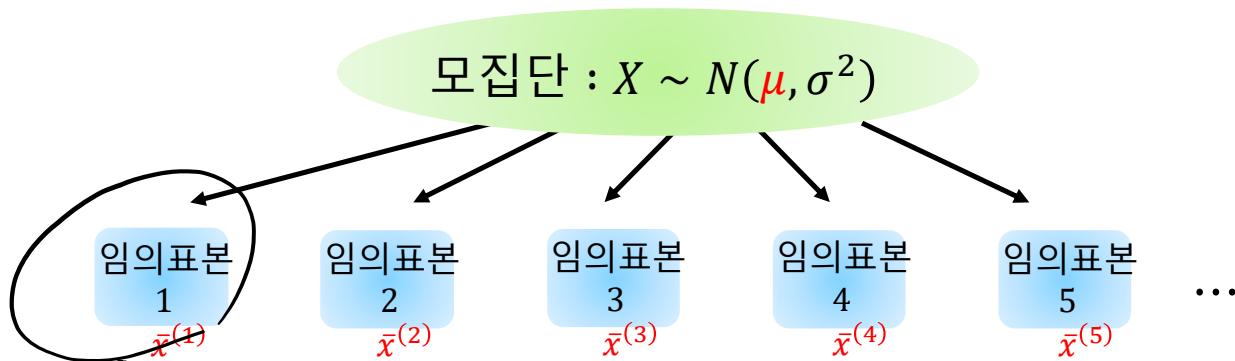
$\bar{x}^{(5)}$

표본추출 할 때마다
표본평균(통계량)의
값이 변함

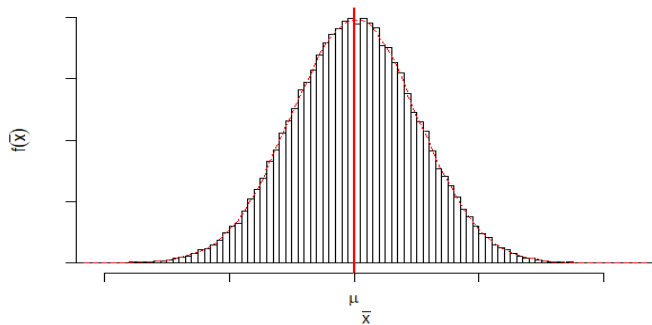
표본평균(통계량)을
확률변수로 간주하고
확률분포를 유도한다

핵심!

$\sim N(\mu, \sigma^2)$
독립



임의 표본은
모집단과 같은
분포를 가진다.



Chapter 9 표집분포

X : 방의 개수

- 예제 1

어느 모집단이 세 개의 수 2, 3, 4로 이루어져 있는데 이 수치들은 세 주택의 방의 개수들을 나타내는 것이다. 이 때 X 를 각 주택의 방의 개수라고 하면 X 는 2, 3, 4 중 하나의 값을 갖게 된다. 세 주택으로부터 두주택을 복원 추출해서 X_1, X_2 를 각각 첫번째와 두번째 추출된 주택의 방의 개수라고 할 때, 표본평균 \bar{X} 의 분포는 X_1, X_2 의 분포에 의해 결정되는데 X_1 의 분포는 X 의 분포와 일치하고, X_2 의 분포는 복원추출의 특성상 X_1 의 분포와 같이 모집단의 분포인 X 의 분포를 따르게 된다. 확률변수 X 의 분포는 각 값 2, 3, 4에 1/3씩의 확률을 주는 확률분포를 가지므로 (X_1, X_2) 가 취하는 모든 값 (x_1, x_2) 과 그에 대응하는 \bar{x} 값과 확률들은 다음과 같이 정리될 수 있다.

(x_1, x_2)	(2,2)	(2,3)	(2,4)	(3,2)	(3,3)	(3,4)	(4,2)	(4,3)	(4,4)
\bar{x}	2	2.5	3	2.5	3	3.5	3	3.5	4
확률	1/9	1/9	1/9	1/9	1/9	1/9	1/9	1/9	1/9

\bar{X} 가 취하는 값(\bar{x})	2	2.5	3	3.5	4	합계
$f(\bar{x})$	1/9	2/9	3/9	2/9	1/9	1

X	2	3	4
$P(X=x)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

← 모집단의
확률 분포도.