표본평균의 평균, 표본평균의 분산

모집단의 크기를 n이라 하고 표본의 크기를 k라 하자. n개에서 복원추출로 k개를 뽑는다.

이 k개의 평균이 표본평균이고, 다음과 같이 나타낼 수 있다. $\overline{X}=\dfrac{\sum_{i=1}^{k} X_i}{k}$

표본평균도 X와 같이 **확률변수** $\overline{X}$로 나타낼 수 있는 것이고 따라서 **표본평균의 평균**은 $\mathrm{E}(\overline{X})$로 나타낼 수 있다. 마찬가지로 **표본평균의 분산**은 $\mathrm{V}(\overline{X})$이다.

표본평균 $\overline{X}$는 다음과 같은 성질을 따른다. 외워라. 증명은 대학과정이다.
이산확률분포뿐만 아니라 연속확률분포에도 적용할 수 있다.

모평균이 m, 모표준편차가 σ인 모집단에서 임의추출한 크기가 n인 표본의 표본평균 $\overline{X}$에 대하여
$\mathrm{E}(\overline{X})=m, \mathrm{V}(\overline{X})=\dfrac{\sigma^2}{n}, \sigma(\overline{X})=\dfrac{\sigma}{\sqrt{n}}$이 성립한다. **이는 근사시킨 것이 아닌 정확한 값이다.**

하지만 구체적 예시 없이 이렇게 문자들로만 보이면 표본평균에 대한 이해가 힘들다.
예시를 통해 표본평균이 도대체 무엇인지 이해해보도록 하자.

주머니에 1, 2, 3이 각각 적힌 공 3개가 있다. 이 주머니에서 임의로 공 1개를 뽑을 때 적힌 수를 확률변수 X라 하자. 표본의 크기가 2인 표본은 몇 개인지 구해보고 각 표본을 모두 나열해보자.

표본의 크기가 2인 표본의 개수부터 구해보자. 다양한 답변이 나올 것이다.
공 3개 중 2개를 뽑는 ${}_3\mathrm{C}_2=3$개? 아니다. 공 3개 중 중복을 허용하여 2개를 뽑는 ${}_3\mathrm{H}_2=6$개? 아니다.

복원추출로 표본을 추출하기에 표본의 크기가 2인 표본의 개수는 $3^2=9$개다.
※ 복원추출이란 확률을 구할 때, 추출했던 것을 원래대로 돌려놓고 다시 추출하는 방법을 말한다.

표본의 크기가 2인 표본 전부를 나열하면 아래와 같다.

표본의 크기가 2인 표본		
$(1, 1)$	$(2, 1)$	$(3, 1)$
$(1, 2)$	$(2, 2)$	$(3, 2)$
$(1, 3)$	$(2, 3)$	$(3, 3)$

$(1, 1)$, $(1, 2)$, $(1, 3)$, $(2, 1)$, $(2, 2)$, $(2, 3)$, $(3, 1)$, $(3, 2)$, $(3, 3)$ **각각이 하나의 표본이다.**
각각의 표본에서 평균을 낸 것이 '표본평균'이다.

표본	$(1, 1)$	$(1, 2)$	$(1, 3)$	$(2, 1)$	$(2, 2)$	$(2, 3)$	$(3, 1)$	$(3, 2)$	$(3, 3)$
표본평균	1	1.5	2	1.5	2	2.5	2	2.5	3

이렇게 얻은 표본평균 값들의 평균을 '표본평균의 평균'이라 하고,
표본평균 값들의 분산을 '표본평균의 분산'이라 한다.

표본평균을 확률변수 $\overline{X}$로 두고 $\overline{X}$의 확률분포를 표로 나타내면 아래와 같다.

$\overline{X}$	1	1.5	2	2.5	3	계
$\mathrm{P}(\overline{X}=\overline{x})$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$	$\frac{2}{9}$	$\frac{1}{9}$	1

$$\mathrm{E}(\overline{X})=1\times\frac{1}{9}+1.5\times\frac{2}{9}+2\times\frac{1}{3}+2.5\times\frac{2}{9}+3\times\frac{1}{9}=2$$

$$\mathrm{V}(\overline{X})=(1-2)^2\times\frac{1}{9}+(1.5-2)^2\times\frac{2}{9}+(2-2)^2\times\frac{1}{3}+(2.5-2)^2\times\frac{2}{9}+(3-2)^2\times\frac{1}{9}=\frac{1}{3}$$

따라서 표본평균의 평균 $\mathrm{E}(\overline{X})=2$**이고 표본평균의 분산은** $\mathrm{V}(\overline{X})=\frac{1}{3}$**이다.**

해당 예시에서 표본의 크기가 2이므로, $\mathrm{E}(\overline{X})=\mathrm{E}(X)$, $\mathrm{V}(\overline{X})=\frac{\mathrm{V}(X)}{2}$**가 성립하는지 확인해보자.**

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

$$\mathrm{E}(X)=1\times\frac{1}{3}+2\times\frac{1}{3}+3\times\frac{1}{3}=2$$

$$\mathrm{V}(\overline{X})=(1-2)^2\times\frac{1}{3}+(2-2)^2\times\frac{1}{3}+(3-2)^2\times\frac{1}{3}=\frac{2}{3}$$

따라서 모평균 $\mathrm{E}(X)=2$**이고 모분산은** $\mathrm{V}(X)=\frac{2}{3}$**이다.**

이로써 해당 예시에서 $\mathrm{E}(\overline{X})=\mathrm{E}(X)$, $\mathrm{V}(\overline{X})=\frac{\mathrm{V}(X)}{2}$**이 성립함을 확인할 수 있다.**

이제 '모평균', '표본평균', '표본평균의 평균', '표본평균의 분산'의 의미가 좀 더 와닿을 것이다.
또한, 표본의 크기가 n일 때 $\mathrm{E}(\overline{X})=\mathrm{E}(X)$, $\mathrm{V}(\overline{X})=\frac{\mathrm{V}(X)}{n}$이 성립함도 단순 암기가 아닌 자연스럽게 받아들일 수 있을 것이다.