수업이 곧 시작됩니다.

통계 내용 정리 [기초강의(1/3)]

[학습목표]

1. 이산확률변수 내용정리

01. 이산확률변수란 시행마다 정해진 확률에 따라 그 값이 변하는 변수이다.

예를 들어서 주사위를 던질 때 나오는 눈을 X라 하면 X는 다음과 같은 분포를 가진다.

X	1	2	3	4	5	6
$\mathrm{P}(X=x)$	1/6	1/6	1/6	1/6	1/6	1/6

02. 예를 들어서 흰 공 3개, 검은 공 2개가 들어 있는 주머니에서 임의로 2개의 공을 동시에 꺼낼 때,

나오는 흰 공의 개수를 X라 하면

X	0	1	2
$\mathrm{P}(X=x)$	1/10	6/10	3/10

이다. $\mathrm{P}(X=1)=\frac{6}{10}$, $\mathrm{P}(X\geq 1)=\frac{9}{10}$ 등으로 표현한다.

03. X가 가지는 값 x_i와 그 확률 p_i의 대응관계를 확률분포라 하고,

이 대응 관계를 나타내는 함수 $\mathrm{P}(X=x)$를 X의 확률질량함수라 한다.

04. 문제에 이산확률변수가 뜨면 위의 예제들처럼 확률분포표를 그리자.

05. X가 1, 2, 3, 4 중 하나의 값을 가지며 $\mathrm{P}(X=i)=ai(i=1,\ 2,\ 3,\ 4)$를 만족할 때,

X의 확률분포표를 구하여라.

$\Rightarrow$

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

$\Rightarrow$ 확률의 합은 1이므로 $a=\frac{1}{10}$이다.

06. 의미 분석도 좋지만 일단 정의이다. 계산방법부터 확인하자.

① X의 기댓값(평균) $\mathrm{E}(X)=\sum_{i=1}^{n}x_ip_i$

② X의 분산 $\mathrm{V}(X)=\sum_{i=1}^{n}(x_i-m)^2p_i$ (단, m은 평균)

③ X의 표준편차 $\sigma(X)=\sqrt{\mathrm{V}(X)}$

07. 통계의 평균과 비교해서 기댓값의 의미를 이해하자.

예를 들어서 시험을 쳤을 때 1/3의 확률로 100점, 2/3의 확률로 70점을 받는 학생이 있다고 하면,

이 학생의 평균점수는 결국 $100\times\frac{1}{3}+70\times\frac{2}{3}=80$점으로 수렴하게 될 것이다.

08. 분산과 표준편차가 평균에서 떨어진 정도를 나타낸다는 것을 이해하자.

① $(x_i-m)^2$을 보자. 평균과의 차이가 크면 값이 더 커진다.

② 분산 구할 때 제곱을 한 번 했으니까 기분상 루트를 한 번 씌워준다.

$\Rightarrow$ 분산보다 표준편차가 많이 쓰이는 이유 : 단위(스케일)가 맞아서.

※ 절댓값을 이용한 절대편차라는 것도 있다. 모종의 이유로 고등학교에서는 쓰지 않는다.

09. 확률변수 X가 0, 1, 2의 값을 가지고 확률질량함수가 $f(x)=a^{|x-1|+1}$일 때, X의 분산을 구하여라.

⇒ 확률분포표부터 그리자.

X	0	1	2
$\mathrm{P}(X=x)$	a^2	a	a^2

⇒ 확률의 합이 1이므로 $a=\frac{1}{2}$이다.

X	0	1	2
$\mathrm{P}(X=x)$	1/4	1/2	1/4

⇒ 세로로 긁어서 더한 것이 평균이다. $\mathrm{E}(X)=0\times\frac{1}{4}+1\times\frac{1}{2}+2\times\frac{1}{4}=1$

⇒ 평균과의 차이를 제곱한 것의 평균이 분산이다. $\mathrm{V}(X)=(0-1)^2\times\frac{1}{4}+(1-1)^2\times\frac{1}{2}+(2-1)^2\times\frac{1}{4}=\frac{1}{2}$

10. 연산된 확률변수의 평균과 분산 : 일단 외우세영.

① $\mathrm{E}(aX+b)=a\mathrm{E}(X)+b$

② $\mathrm{V}(aX+b)=a^2\mathrm{V}(X)$

③ $\sigma(aX+b)=|a|\sigma(X)$

④ $\mathrm{V}(X)=\mathrm{E}(X^2)-\{\mathrm{E}(X)\}^2$

11. 의미를 이해하자. 어떤 시험에서 반의 평균이 30점일 때,

① 각각의 점수를 10점씩 올리면 평균은 10점이 오른다. → $\mathrm{E}(X+10)=\mathrm{E}(X)+10$

② 각각의 점수를 두 배로 올리면 평균은 두 배가 된다. → $\mathrm{E}(2X)=2\mathrm{E}(X)$

③ 각각의 점수를 10점씩 올리더라도 분산은 변하지 않는다. (평균도 10점이 오르므로) → $\mathrm{V}(X+10)=\mathrm{V}(X)$

④ 각각의 점수를 두 배로 올리면 분산은 네 배가 된다. (평균과의 간격이 두 배가 되므로) → $\mathrm{V}(2X)=4\mathrm{V}(X)$

12. 사실 약간 엉터리 설명이다. 통계자료와 확률변수는 서로 다르기 때문이다.

①을 제대로 증명해보자. 확률변수 $aX+b$의 확률분포는 다음과 같다.

$aX+b$	ax_1+b	ax_2+b	$\cdots$	ax_n+b
$\mathrm{P}(aX+b=ax+b)$	p_1	p_2	$\cdots$	p_n

따라서, $\mathrm{E}(aX+b)=\sum_{i=1}^{n}(ax_i+b)p_i=a\sum_{i=1}^{n}x_ip_i+b\sum_{i=1}^{n}p_i=a\mathrm{E}(X)+b$

13. 분산은 각자 해보자. ④를 해줄게.

$$\mathrm{V}(X)=\sum_{i=1}^{n}(x_i-m)^2p_i=\sum_{i=1}^{n}x_i^2p_i-2m\sum_{i=1}^{n}x_ip_i+m^2\sum_{i=1}^{n}p_i$$

$$=\sum_{i=1}^{n}x_i^2p_i-2m^2+m^2=\mathrm{E}(X^2)-\{\mathrm{E}(X)\}^2$$

14. 두 확률변수 X, Y에 대하여 $X=2Y+3$이고, $\mathrm{E}(Y)=3$, $\mathrm{E}(Y^2)=14$일 때, $\mathrm{E}(X)$와 $\mathrm{E}(X^2)$의 값을 구하여라.

15. 일어날 확률이 p인 시행을 n번 반복할 때, X번 일어난다고 하자.

확률변수 X가 시행횟수 n, 성공확률 p인 이항분포를 따른다고 하고 $X\sim\mathrm{B}(n,\ p)$라 나타낸다.

16. $X\sim\mathrm{B}(n,\ p)$이면 $\mathrm{P}(X=i)={}_n\mathrm{C}_ip^i(1-p)^{n-i}(0\le i\le n$인 정수)

⇒ 뭔지 모르겠으면 독립시행 복습해.

17. 주사위 3개를 던져서 6의 약수인 눈이 나오는 것의 개수를 X라 하자.

X	0	1	2	3
$\mathrm{P}(X=x)$	$\left(\frac{1}{3}\right)^3$	${}_3\mathrm{C}_1\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^2$	${}_3\mathrm{C}_2\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)$	$\left(\frac{2}{3}\right)^3$

⇒ 이런 경우를 $X\sim\mathrm{B}\left(3,\ \frac{2}{3}\right)$라 나타낸다. 평균과 분산도 구해보자.

18. 이항분포의 평균과 분산 : $X\sim\mathrm{B}(n,\ p)$이면

① $\mathrm{E}(X)=np$

② $\mathrm{V}(X)=np(1-p)$

이다. 증명은 귀찮고 안 나오니까 그냥 알려져 있다고 하자. 논술 준비라면 해둘 것.

평균이 np인 것은 삘이 오긴 한다. → 주사위 600개를 던지면 1이 몇 개 나올까? 평균적으로 100개.

수업이 곧 시작됩니다.

통계 내용 정리 [기초강의(2/3)]

[학습목표]

1. 연속확률변수 내용 정리

01. 연속확률변수 X : 확률변수가 연속적인 값을 가질 때,

eg) 길이, 시간 등등

→ 이산확률변수처럼 칸으로 나눌 수 없다.

→ 연속확률변수는 범위에서 확률을 가진다.

→ $\mathrm{P}(X=k)=0$, $\mathrm{P}(a<X<b)=\mathrm{P}(a\le X\le b)$

02. 연속확률변수 X에 대하여 함수 $f(x)$가

$$\mathrm{P}(a\le X\le b)=\int_a^b f(x)dx$$

를 만족시키면 $f(x)$를 X의 확률밀도함수라 한다. 이를 $X\sim f(x)$라 나타낸다.

03. 어떤 연속확률변수의 확률밀도함수 $f(x)$는 다음을 만족시켜야 한다.

(가) $f(x)\ge 0$

(나) $\int_{-\infty}^{\infty} f(x)dx=1$

※ $f(x)\le 1$일 필요는 없다. 함숫값이 아니라 넓이가 확률이다. 가로를 좁히면 함숫값은 1을 넘을 수 있다.

04. $X\sim f(x)=ax(0\le x\le 2)$일 때 다음을 구하여라.

(1) a

(2) $\mathrm{P}(X\le 1)$

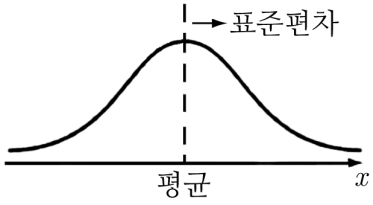
05. 정규분포는 여러 종류의 측정값이 가지는 매우 자연스러운 분포이다.

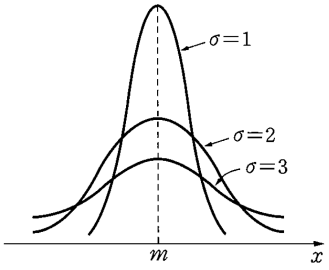
① 오른쪽 그림과 같이 평균을 중심으로 좌우 대칭의 종모양이다.

→ 확률밀도함수가 $\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}$ 인데, 어차피 이해 못하니까..

② 연속확률변수 X가 평균 m, 표준편차 σ인 정규분포를 따를 때,

$X\sim \mathrm{N}(m,\ \sigma^2)$으로 나타낸다.





06. 표준편차가 커질수록 옆으로 퍼진다.

07. $X\sim \mathrm{N}(m_1,\ \sigma_1^2)$, $Y\sim \mathrm{N}(m_2,\ \sigma_2^2)$이면 임의의 실수 k에 대하여

$\mathrm{P}(m_1\le X\le m_1+k\sigma_1)=\mathrm{P}(m_2\le Y\le m_2+k\sigma_2)$이다.

→ 평균에서부터 편차로 몇 칸 떨어졌는지로 확률을 구한다.

08. 평균이 0, 표준편차가 1인 정규분포를 표준정규분포라 한다.

① 표준정규분포를 따르는 확률변수를 일반적으로 Z로 나타낸다.

② 표준정규분포의 확률값은 문제에서 주어진다.

③ $\mathrm{P}(0\le Z\le 1)=0.34$, $\mathrm{P}(0\le Z\le 2)=0.48$ 정도는 예의상 익혀두자.

09. 어느 집단 구성원의 IQ는 $\mathrm{N}(100,\ 10^2)$을 따른다고 한다. 다음을 구하여라.
(단, 확률변수 Z가 표준정규분포를 따를 때, $\mathrm{P}(0 \le Z \le 1) = 0.34$, $\mathrm{P}(0 \le Z \le 2) = 0.48$이다.)

(1) $\mathrm{P}(\mathrm{IQ} \ge 100)$ (2) $\mathrm{P}(\mathrm{IQ} \le 120)$

(3) $\mathrm{P}(\mathrm{IQ} \ge 110)$ (4) $\mathrm{P}(90 < \mathrm{IQ} < 110)$

(5) $\mathrm{P}(80 \le \mathrm{IQ} \le 110)$ (6) $\mathrm{P}(80 \le \mathrm{IQ} \le 90)$

10. 어떤 시험 성적의 분포가 $\mathrm{N}(50,\ 8^2)$일 때, 수험생 500명 중에서 상위 10등 이내에 들기 위해서는 몇 점 이상이어야 하는지 구하여라. (단, $\mathrm{P}(0 \le Z \le 2) = 0.48$)

11. 확률변수 X에 대하여 평균을 빼고 편차로 나누는 것을 <u>표준화</u>라 한다.

→ $X \sim \mathrm{N}(m,\ \sigma^2)$일 때, 표준화 친 $\dfrac{X-m}{\sigma}$는 표준정규분포 $\mathrm{N}(0,\ 1^2)$을 따른다.

적용법 ① $\mathrm{N}(m,\ \sigma^2)$의 k와 대응되는 표준정규분포에서의 값은 $\dfrac{k-m}{\sigma}$이다.

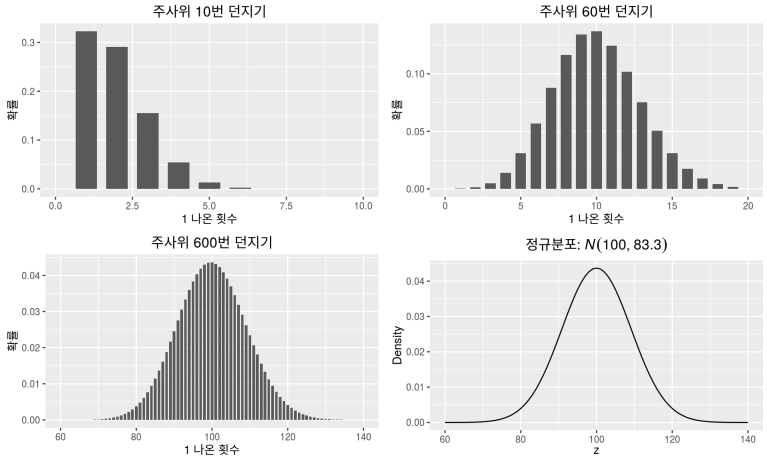
적용법 ② $X \sim \mathrm{N}(m,\ \sigma^2)$일 때, $\mathrm{P}(a \le X \le b) = \mathrm{P}\left(\dfrac{a-m}{\sigma} \le Z \le \dfrac{b-m}{\sigma}\right)$이다.

12. 어느 고등학교에서 학생 몸무게의 분포는 평균이 53kg, 표준편차가 6kg인 정규분포를 따른다.
이 학교 학생 중 몸무게가 50kg 이상 65kg 이하인 학생은 전체의 몇 %인지 구하여라.
(단, $\mathrm{P}(0 \le Z \le 0.5) = 0.19$, $\mathrm{P}(0 \le Z \le 2) = 0.48$)

13. 어느 제과점에서 만든 빵 한 개의 무게는 평균 200g, 표준편차 8g인 정규분포를 따른다고 한다.
이 제과점에서 만든 빵 한 개의 무게가 210g 이상일 확률을 구하여라. (단, $\mathrm{P}(0 \le Z \le 1.25) = 0.4$)

14. 어떤 시험 성적의 분포가 평균이 75점, 표준편차가 9인 정규분포를 따른다고 한다. 수험생중 상위 10 %에 들기 위해서는 적어도 몇 점 이상을 받아야 하는지 구하여라. (단, $\mathrm{P}(0 \le Z \le 1.28) = 0.4$)

15. <u>이항분포 $\mathrm{B}(n,\ p)$는 n이 충분히 크면 정규분포 $\mathrm{N}(np,\ npq)$로 근사된다.</u> (단, $q = 1-p$)
그냥 믿자. 정규분포의 정의라고 착각하는 정도가 좋다. 그러면 정규분포가 자연스럽다는 것도 약간 설명되고..
※ 보통 $np \ge 15$이면 근사시키기 충분하다. 표준정규분포표가 주어지면 근사하기 충분하다고 판단할 수 있다.
※ 이항분포가 나왔을 때, 무조건 근사시키다가 망하는 수가 있다.



16. 5지선다 100문제의 답을 모두 임의로 찍었을 때, 다음을 구하여라.

(1) 20문제의 정답을 맞힐 확률 (2) 100문제의 정답을 맞힐 확률

(3) 20 ~ 24문제의 정답을 맞힐 확률 (4) 28문제 이상 정답을 맞힐 확률

⇒ 맞힌 문제의 수 X는 $\mathrm{B}\left(100,\ \dfrac{1}{5}\right)$를 따른다. 확률분포표도 대충 잡아보자.

(1)은 ${}_{100}\mathrm{C}_{20}\left(\dfrac{1}{5}\right)^{20}\left(\dfrac{4}{5}\right)^{80}$, (2)는 $\left(\dfrac{1}{5}\right)^{100}$이다. 둘 다 거의 0에 가까운 값. 연속확률변수로 근사된다는 개념에서도..

(3), (4)는 시그마로 표현할 수는 있다. 분포를 정규분포 $\mathrm{N}(20,\ 4^2)$으로 근사시킬 수 있으니 여기서 구해보자.

수업이 곧 시작됩니다.

통계 내용 정리 [기초강의(3/3)]

[학습목표]

1. 통계적 추정 내용 정리

01. 표본평균 $\overline{X}$은 X를 몇 개 뽑아서 평균 낸 것이다.

확률변수 X가 다음과 같은 분포를 가진다면, 두 개를 뽑아서 평균 낸 $\overline{X}$의 분포는 오른쪽과 같다. 생각해봐.

X	1	7
$\mathrm{P}(X=x)$	1/3	2/3

$\Rightarrow$

$\overline{X}$	1	4	7
$\mathrm{P}(\overline{X}=x)$	1/9	4/9	4/9

02. 위의 예에서 세 개를 뽑아서 평균 낸 $\overline{X}$의 분포는 다음과 같다. 각자 확인.

$\overline{X}$	1	3	5	7
$\mathrm{P}(\overline{X}=x)$	1/27	6/27	12/27	8/27

03. 표본평균의 평균과 분산 : $\overline{X}$는 다음을 만족한다.

① $\mathrm{E}(\overline{X})=\mathrm{E}(X)$ ② $\mathrm{V}(\overline{X})=\dfrac{\mathrm{V}(X)}{n}$ ③ $\sigma(\overline{X})=\dfrac{\sigma(X)}{\sqrt{n}}$

04. 증명은 제대로는 안 되니까 받아들여야 한다.

→ ①은 살짝 당연하다. 어차피 X를 뽑은 것이니까..?

→ ②, ③에 대해서는 일단 n이 커지면 분산이나 표준편차가 작아진다는 것을 감각적으로 느껴보자.

→ 여러 개를 뽑아서 평균 내면, 값이 좀 안정된다. 많이 뽑을수록. 이건 동의할 수 있지?

05. 모평균 m, 모표준편차 σ인 모집단에서 추출하면,

① $\mathrm{E}(\overline{X})=m$, $\mathrm{V}(\overline{X})=\dfrac{\sigma^2}{n}$, $\sigma(\overline{X})=\dfrac{\sigma}{\sqrt{n}}$

② 모집단이 정규분포이면 $\overline{X}$는 정규분포를 따른다.

③ 모집단과 관계없이 n이 충분히 크면 $\overline{X}$는 정규분포를 따른다.

$\Rightarrow$ 알려져 있다. 아멘.

06. 어느 공장에서 제품 한 개를 생산하는 데 걸리는 시간은 평균이 200분, 표준편차가 16분인 정규분포를 따른다고 한다. 이 공장에서 생산하는 제품 중 임의로 추출한 4개 제품의 생산 시간의 평균이 196분 이상일 확률을 구하여라. (단, $\mathrm{P}(0 \le Z \le 0.5)=0.19$)

07. 어느 회사에서 생산하는 볼펜의 무게는 평균이 20g, 표준편차가 2g인 정규분포를 따른다고 한다. 이 회사에서 생산되는 볼펜을 4개씩 묶어서 세트 상품을 만들 때, 세트 상품의 무게의 합 S에 대하여 $\mathrm{P}(78 \le S \le 84)$을 구하여라. (단, $\mathrm{P}(0 \le Z \le 0.5)=0.19$, $\mathrm{P}(0 \le Z \le 1)=0.34$)

08. $X \sim \mathrm{N}(20,\ 2^2)$이면 네 개를 뽑은 것의 합 $X_1+X_2+X_3+X_4$는 $4\overline{X}$이므로 $\mathrm{N}(80,\ 4^2)$를 따르고 하나를 뽑아서 네 배한 $4X$는 $\mathrm{N}(80,\ 8^2)$을 따른다. (식 $\mathrm{E}(aX)=a\mathrm{E}(X)$와 $\mathrm{V}(aX)=a^2\mathrm{V}(X)$에서)

09. 모평균의 추정관련 용어 정리

(1) 모집단 : 조사 대상이 되는 집단 → 모평균 m, 모분산 σ^2, 모표준편차 σ

(2) 표본 : 조사를 위해 추출한 모집단의 일부 → 표본의 크기 n, 표본평균 $\overline{X}$, 표본분산 S^2, 표본표준편차 S

(3) 추정, 신뢰도, 신뢰구간 : 표본평균 $\overline{x}$을 중심으로 모평균 m의 값을 추정한다.
신뢰도 몇%로 모평균은 구간 $[\alpha,\ \beta]$에 존재한다고 추정할 때, 구간 $[\alpha,\ \beta]$를 신뢰구간이라 한다.

10. 예시 : 우리나라 사람들 전체의 IQ 평균을 알고 싶다.

→ 모집단 : 우리나라 사람들,

→ 모평균 m : 우리나라 사람들 전체의 IQ 평균

임의로 10명을 뽑아서 IQ를 조사했더니 평균이 105, 편차가 10이었다.

→ 표본의 크기 n : 10

→ 표본평균 $\bar{x}$: 105

→ 표본표준편차 s : 10 (이 값은 모표준편차 σ 대신 사용할 수 있다.)

① 우리나라 사람들 전체의 IQ 평균은 105일꺼야! → 미친놈.

② 우리나라 사람들 전체의 IQ 평균은 103에서 107 사이에 있을꺼야! → 너무 용감한데?

③ 우리나라 사람들 전체의 IQ 평균은 95에서 115 사이에 있을꺼야! → 아마도? 믿을만한데?

④ 우리나라 사람들 전체의 IQ 평균은 0에서 210 사이에 있을꺼야! → 100% 신뢰할 수 있군.

→ 신뢰도를 높이려면 신뢰구간을 넓혀야 한다.

11. 모집단이 1, 2, 3, 4, 5일 때, 크기 3인 표본을 복원추출/비복원추출하는 경우의 수를 각각 구하여라.

① 복원추출 : 추출할 때마다 모집단을 원 상태로 복원 : $5\times5\times5$

② 비복원추출 : 한 개씩 추출함 : $5\times4\times3$

→ 모분포에서의 추출은 모두 복원추출이다.

→ 모집단이 충분히 크면 비복원추출도 거의 복원추출 취급할 수 있다.

→ 특별한 언급이 없을 때는 복원추출만 고려한다.

12. 표본표준편차는 모표준편차 대신 사용할 수 있다. 그냥 그러려니 하자.

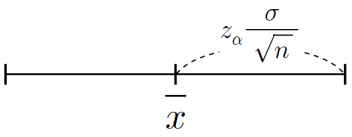
보통 문제에서는 모표준편차가 주어진다. 하지만 일반적으로, 실제로 모표준편차를 알 수는 없다.

※ $\sigma(\bar{X})=\frac{\sigma}{\sqrt{n}}$와 헷갈리면 안 된다. $\frac{\sigma}{\sqrt{n}}$는 표본평균들 사이의 표준편차이다.

13. 신뢰도 $\alpha\%$로 모평균 m을 추정한 신뢰구간은 아래와 같다.

모평균 m의 신뢰도 $\alpha\%$ 신뢰구간

$$\bar{x}-z_\alpha\frac{\sigma}{\sqrt{n}}\le m\le\bar{x}+z_\alpha\frac{\sigma}{\sqrt{n}}$$



※ 신뢰계수 z_α : 일단 95%일 때는 1.96, 99%일 때는 2.58 정도를 눈치껏 쓰자.

14. 어느 과수원에서 생산되는 사과 81개를 조사하였더니 무게의 평균이 120g,

표준편차가 36g이었다. 전체 평균을 신뢰도 99%로 추정하여라. (단, $P(0\le Z\le 2.6)=0.4950$)

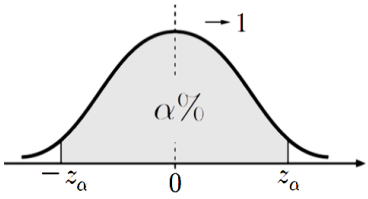
⇒ $n=81$, $\bar{x}=120$, $s=36(=\sigma)$, $z_\alpha=2.6$이다.

$\bar{x}-z_\alpha\frac{\sigma}{\sqrt{n}}\le m\le\bar{x}+z_\alpha\frac{\sigma}{\sqrt{n}}$에 대입하면 신뢰구간은 [109.6, 130.4]이다.

15. 추정 식의 z_α값(신뢰계수)은 신뢰도 α에 대하여 $P(-z_\alpha\le Z\le z_\alpha)=\alpha\%$가 되는 값이다.

⇒ 95%일 때는 1.96 이나 2쯤 쓰면 된다.

⇒ 99%일 때는 2.58 이나 2.6 이나 3쯤 쓰면 된다.



16. 신뢰구간의 길이는 $2z_\alpha\frac{\sigma}{\sqrt{n}}$이다. 신뢰구간의 길이는 좁히고 싶다.

① n을 키우면 신뢰구간은 짧아진다. (좋지만 조사하는 표본을 키우려면 돈이 든다.)

② σ가 크면 신뢰구간은 길어진다. (모집단이 들쭉날쭉하면 추정이 잘 되지 않는다.)

③ α(신뢰도)가 크면 신뢰구간은 길어진다. (신뢰도를 확보하려다 보니 보수적으로 말해야 한다.)

※ 오차의 한계 : $z_\alpha\frac{\sigma}{\sqrt{n}}$

17. 모표준편차가 3인 모집단의 모평균에 대한 추정을 하려 한다. 신뢰도 99%인 신뢰구간의 길이를

3 이하가 되게 하는 표본의 크기 n의 최솟값을 구하여라. (단, $P(0\le Z\le 2.6)=0.4950$)

⇒ $\sigma=3$, $z_\alpha=2.6$이다. 신뢰구간의 길이 $2z_\alpha\frac{\sigma}{\sqrt{n}}=5.2\times\frac{3}{\sqrt{n}}$이 3 이하가 되어야 한다.