12장 신뢰 수준 (1부)

1. 점추정의 문제와 해결책

p.529

1) 문제

① 모집단의 평균과 분산을 위해 표본을 활용한 점추정 기법을 살펴보았다. 그런데 추정값을 얻기 위해 딴 하나의 표본만을 이용하였다.

② 표본이 편향되지 않았다 하더라도 표본이 모집단을 100% 정확하게 반영하는지 여부는 절대로 알 수 없다.

③ 점추정은 단지 최선을 예측하는 것 뿐이다.

2) 해결책

p.530

① 예를 들어, 평균값에 대한 추정값을 명확하게 말하는 대신에 평균의 추정값이 어느 구간에 속하는지를 말할 수 있다.

② 구간의 크기는 결과를 어느 정도 신뢰할 수 있게 만들 것인가에 따라 결정된다.

③ 신뢰할 수 있는 정도를 신뢰수준이라 부름.

2. 점추정 신뢰구간

1) 예제 : 달콤한 풍선껌 회사의 표본조사 결과

p.528

$$\begin{cases} \hat{\mu} = \bar{x} = 62.7 \text{ (분)} \\ s^2 = 25 \\ n = 100 \end{cases}$$

← 표본의 크기

질문 : $\hat{\mu}$ 를 어느 정도로 신뢰할 수 있나?

2) 해결책 : 모평균 μ가 위치할 가능성이 높은 구간을 $\hat{\mu}$, s^2 을 활용하여 계산함.

3) 준비사항 : 모평균 μ가 위치할 가능성이 높은 구간을 신뢰구간 이라 하며 신뢰구간을 계산하기 위해 아래 사항이 준비되어야 함.

① 표본평균의 분포 확인

- 모집단을 대상으로 하는 분포가 어떤 분포를 따르는지 모름.
- 반면에 표본들의 분포는 표본의 크기가 충분히 크면 정규분포 따름.
- 따라서 모평균에 대한 정보를 유추할 수 있음.

p.534~535

② 신뢰수준 설정

- 신뢰수준 : 설정한 신뢰구간이 모평균을 어느 정도의 확률로 포함할지에 대한 기준 제공
- 일반적으로 90%, 95%, 99% 사용하며, 이중에 95% 신뢰수준이 가장 많이 사용됨.
- 신뢰수준을 높일수록 신뢰구간이 넓어짐.
- 신뢰구간이 너무 크면 신뢰구간의 의미가 약해짐.

p.533

4) 풍선껌 모집단의 모평균 신뢰구간 계산

① 평균의 표본분포 확인

풍선껌 하나의 지속시간을 가리키는 확률변수를 X라 하자.
이제 모집단의 기대치와 분산을 아래와 같이 점추정할 수 있다.

$$\hat{\mu} = \bar{x} = 62.7$$
$$\hat{\sigma}^2 = s^2 = 25$$

X가 어떤 분포를 따르는지 모른다. 따라서 모평균 μ에 대한 어떤 정보도 유추할 수 없다.
반면에 표본의 크기가 충분히 크면 $\bar{X}$는 정규분포를 따르며 $\bar{X}$의 기대치와 X의 기대치가 동일하다는 사실을 이용하여 모평균 μ가 포함될 가능성이 높은 신뢰구간을 계산할 수 있다.

풍선껌 100개로 구성된 표본들에 대한 각 지속시간의 평균값의 분포 $\bar{X}$의 통계량은 다음과 같다.

$$E(\bar{X}) = \mu$$
$$Var(\bar{X}) = \frac{s^2}{n} = \frac{25}{100} = 0.25$$

σ^2을 모를 때 대신 사용

30보다 큼

또한 $\bar{X}$는 정규분포를 따른다. 즉,

$$\bar{X} \sim N(\mu, 0.25)$$

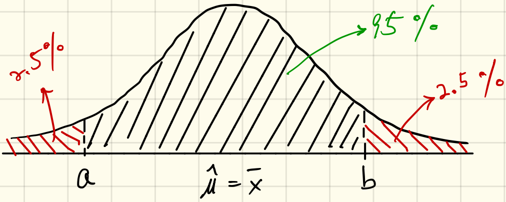
② 신뢰수준 설정

- 95%의 신뢰수준 사용.

③ 모평균 신뢰구간 계산

• 전제: 95% 신뢰수준 사용

• 모평균의 신뢰구간 [a,b]는 다음을 만족해야 한다.



즉, $P(a < \bar{X} < b) = 0.95$ 를 만족시키는

a 와 b를 구해야 함.

방법: 표준화 과정 거꾸로 하기.

i) $\bar{X}$의 표준화: $z = \dfrac{\bar{X} - \mu}{\sqrt{0.25}}$

ii) $P(z_a < z < z_b) = 0.95$ 가 되도록 z_a, z_b 구하기

즉, 다음이 성립해야 한다.

$$\left.\begin{array}{l} P(z < z_a) = 0.025 \\ P(z > z_b) = 0.025 \end{array}\right\} \Rightarrow \begin{array}{l} z_a = -1.96 \\ z_b = 1.96 \end{array}$$

iii) z_a, z_b를 이용하여 μ의 신뢰구간 구하기.

힌트: 다음 공식 활용

$$\begin{cases} -1.96 < z = \dfrac{\bar{X} - \mu}{\sqrt{0.25}} < 1.96 \\ \bar{X} - 0.98 < \mu < \bar{X} + 0.98 \end{cases}$$

iv) $\bar{X}$로 사용가능한 값 확인

- $\bar{X}$는 표본 평균의 분포를 가리킴
- 따라서 $\bar{X}$의 기대치의 추정값인 표본평균 $\bar{x}$를 사용할 수 있음.

대문자

소문자

v) 모평균 μ의 신뢰구간

$$[62.7 - 0.98,\ 62.7 + 0.98]$$
$$= [61.72,\ 63.68]$$

3. 점추정 신뢰구간 구하기 단계 정리

p. 543 ~544

(1) 모집단 통계 선택

예) 풍선껌 향 지속시간의 평균값에 대한 신뢰구간 설정 필요,

즉, μ에 대한 신뢰구간 설정 필요

(2) 표본분포 찾기

예) • 평균값 표본분포의 기대치와 분산 계산

• $\bar{X}$의 정규분포 활용

(3) 신뢰수준 정하기

예) 일반적으로 95%의 신뢰수준 사용

(4) 신뢰구간 찾기

예) 신뢰수준과 표본분포를 사용하여 신뢰구간 구하기.

4. 신뢰구간 설정 공식 모음

모집단 통계	모집단 분포	조건	신뢰구간
μ	X 정규분포	σ^2 알고 있음 n 표본 크기 $\bar{X}$ 표본 평균 값	$\left(\bar{X} - c\frac{\sigma}{\sqrt{n}}, \bar{X} + c\frac{\sigma}{\sqrt{n}}\right)$
μ	X 정규분포 아닌	σ^2 알고 있음 $n \geq 30$ $\bar{X}$ 표본 평균 값	$\left(\bar{X} - c\frac{\sigma}{\sqrt{n}}, \bar{X} + c\frac{\sigma}{\sqrt{n}}\right)$
μ	X 임의의 분포	σ^2 모름 $n \geq 30$ $\bar{X}$ 표본 평균 값 S^2 표본 분산	$\left(\bar{X} - c\frac{S}{\sqrt{n}}, \bar{X} + c\frac{S}{\sqrt{n}}\right)$

* C 의 값 : 신뢰수준에 의해 결정됨

신뢰수준	C 의 값
90%	1.64
95%	1.96
99%	2.58