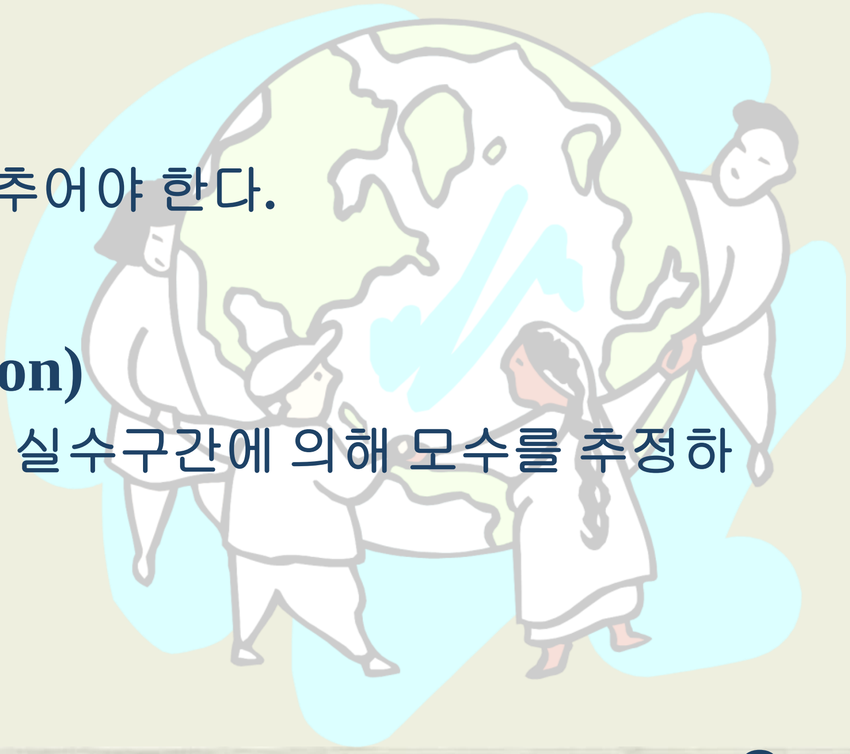
제 8장

추정

고려대학교 경영대학 박 광태

신뢰구간의 추정

- 점추정 (point estimation)
  - 점추정이란 단일 추정치에 의하여 모수를 추정하는 방법을 말한다.
- 점추정량의 바람직한 성질
  - 불편성, 효율성, 일관성을 갖추어야 한다.
- 구간추정 (interval estimation)
  - 점추정과 달리 신뢰수준에서 실수구간에 의해 모수를 추정하는 방법



예제 8-2

표본자료 X_1, X_2, X_3에 의하여 모평균 μ를 추정하고자 다음의 세 가지 추정량을 고려하고 있다.

$$\hat{\mu}_1 = \frac{(X_1 + X_2 + X_3)}{3}$$

$$\hat{\mu}_2 = \frac{(X_1 + 2X_2 + X_3)}{4}$$

$$\hat{\mu}_3 = \frac{(X_1 + 2X_2 + 3X_3)}{5}$$

세 개의 추정량은 모두 불편추정량인가?

예제 8-3

표본분산 S^2은 σ^2의 불편추정량이다. 그렇다면 S는 σ의 불편추정량인가?

모평균 μ의 $100(1-\alpha)\%$의 신뢰구간

- 모분산이 알려져 있고 정규모집단인 경우

$$CI = \left(\overline{X}_n - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X}_n + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

- 모분산이 알려져 있지 않은 정규모집단의 경우

$$CI = \left(\overline{X}_n - t_{n-1,\frac{\alpha}{2}}\frac{S}{\sqrt{n}}, \overline{X}_n + t_{n-1,\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right)$$

- 표본(n>30)인 경우-모집단의 분포에 상관없이

$$CI = \left(\overline{X}_n - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X}_n + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

  - 위의 식에서 σ를 모를 경우 S로 대체한다.

예제 8-6

회계사의 월평균수입(단위: 천 원)을 추정하고자 한다. 모집단은 정규분포이며, 표준편차(σ)는 800으로 알려져 있다. 25명의 회계사들에 대한 단순확률표본으로부터 $\overline{X}_{25} = 3,500$을 얻었다. 모평균 μ에 대한 95% 신뢰구간을 구하시오.

그림 8-5 신뢰구간을 구하기 위한 CONFIDENCE.NORM 함수 입력화면

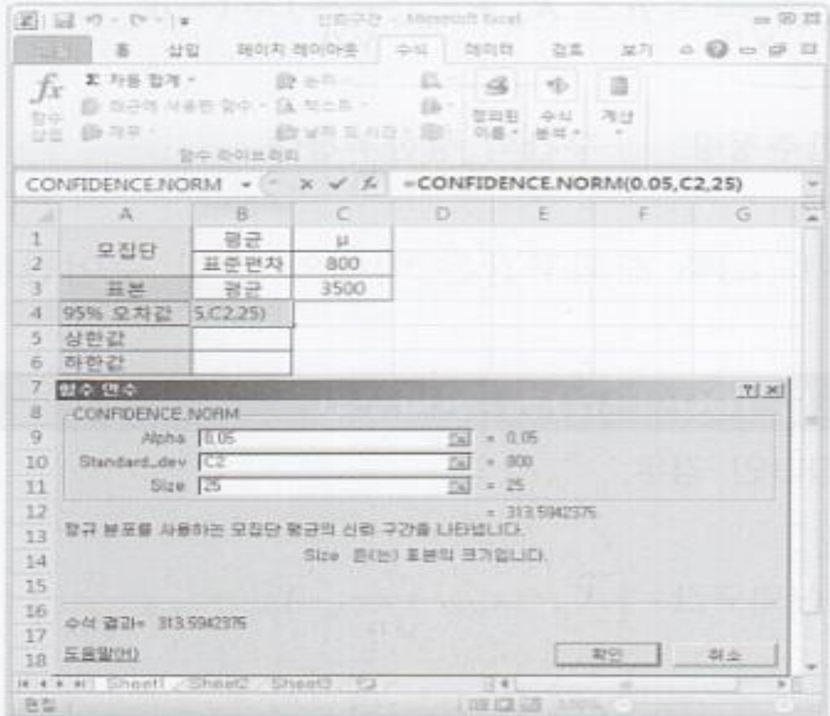
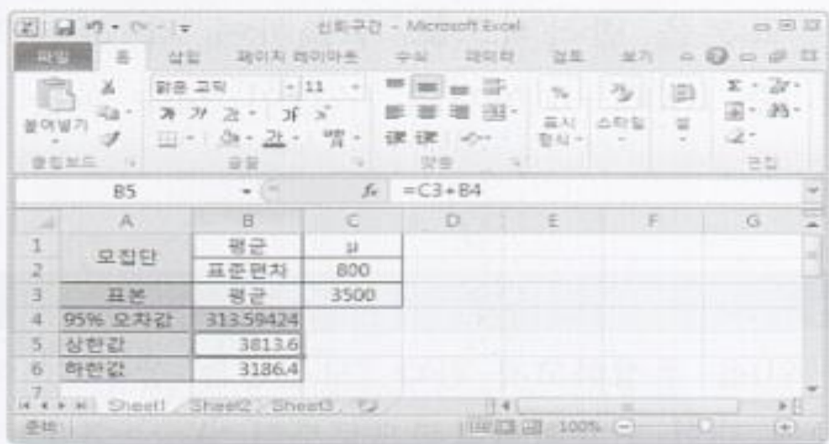
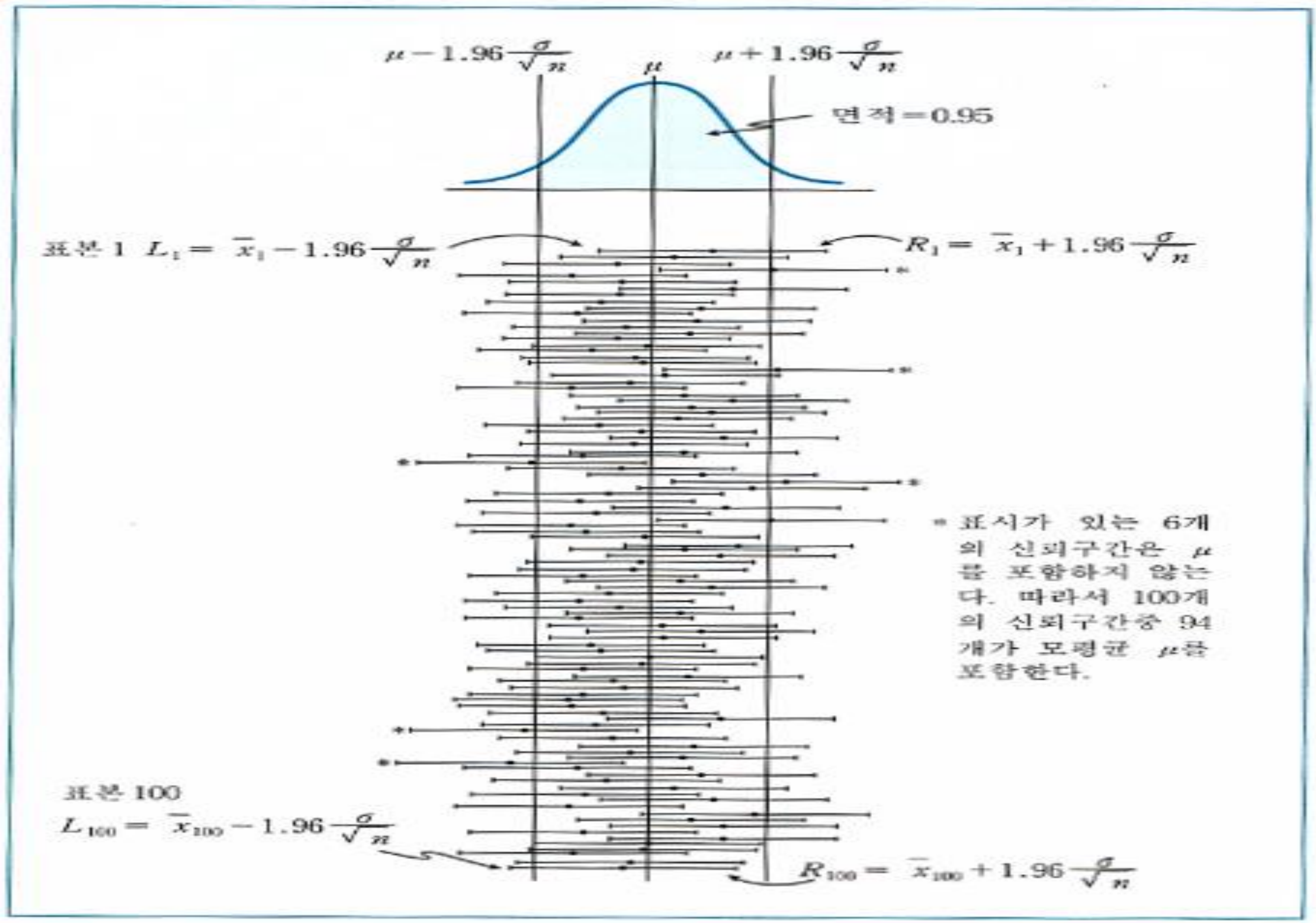


그림 8-6 95% 신뢰구간 계산결과



[그림 8-7] 평균에 대한 100개의 가능한 95% 신뢰구간



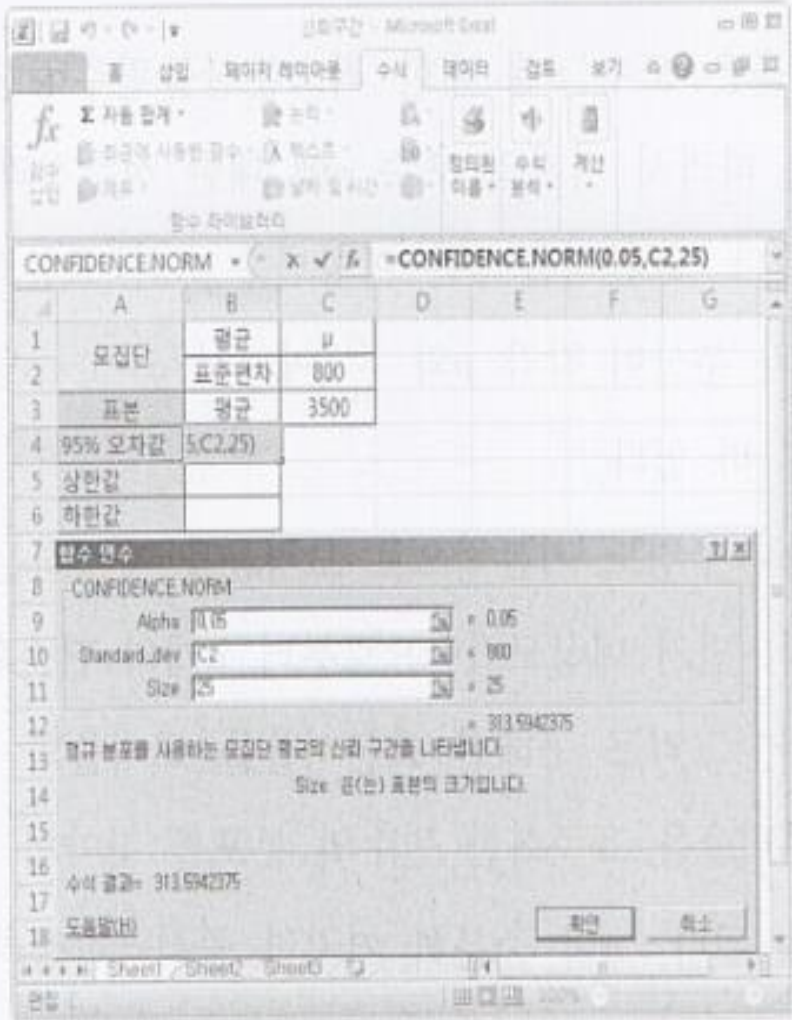
예제 8-7

분산이 알려진 정규모집단에서 크기 n의 확률표본을 취했다고 하자. 모평균에 대한 99%, 95%, 90% 신뢰구간의 폭을 비교하시오.

예제 8-8

예제 6에서 σ를 모른다고 가정하자. 추출된 25명의 회계사들로부터 $S=800$을 얻었다고 하자. 모평균 μ에 대한 95% 신뢰구간을 구하시오.

그림 8-9 신뢰구간을 구하기 위한 CONFIDENCE.NORM 함수 입력화면

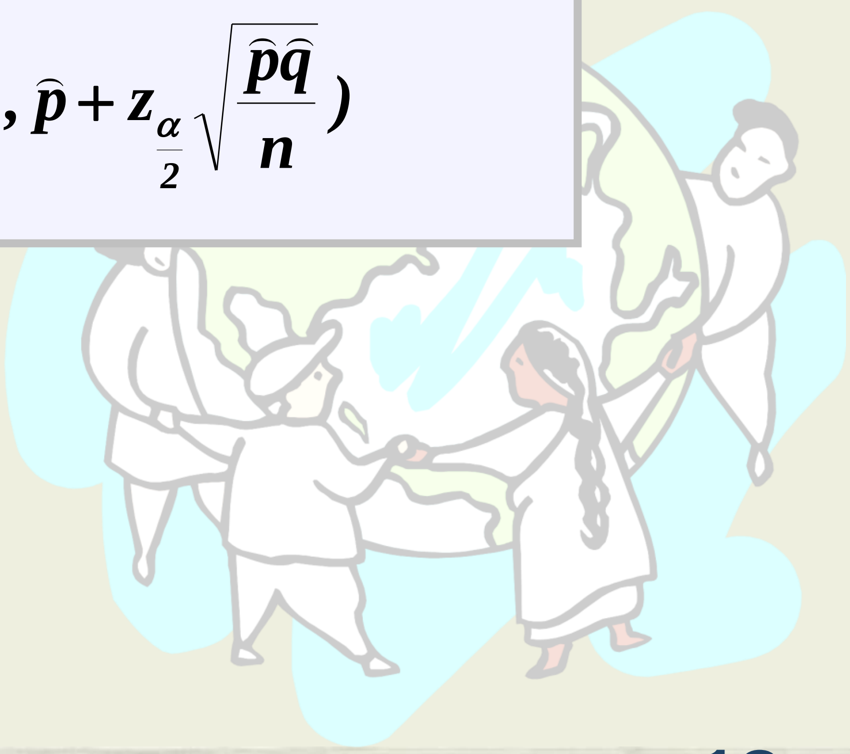


모비율 p의 $100(1-\alpha)\%$ 신뢰구간

- 표본이 충분히 클 때

$n\hat{p} \geq 5$ 이고, $n\hat{q} \geq 5$

$$CI = \left(\hat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}, \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} \right)$$



예제 8-9

정부는 월 100만 원 이하의 수입을 갖는 가구의 비율을 추정하고자 100가구로 이루어진 확률표본을 취하였다. 표본에서 40가구가 월수입이 100만 원 이하였다. 모비율 p에 대한 95% 신뢰구간을 구하시오.

모평균의 추정에서 표본크기의 결정

- n이 충분히 클 때, 오차가 d이하일 확률이 (1− α)가 되게 하는 표본크기 n은

- σ 를 알 때

$$n = \left(\frac{z_{\frac{\alpha}{2}} \sigma}{d} \right)^2$$

- σ 를 모를 때

$$\boldsymbol{n = \left(\frac{z_{\frac{\alpha}{2}} \sigma}{d} \right)^2 \approx \left[\frac{z_{\frac{\alpha}{2}} S'}{d} \right]^2}$$

$$\boldsymbol{n = \left(\frac{z_{\frac{\alpha}{2}} \sigma}{d} \right)^2 \approx \left[\frac{z_{\frac{\alpha}{2}} (R/4)}{d} \right]^2}$$

단, 여기서

$S' =$ 예비표본의표준편차

$R =$ 자료의범위

예제 8-10

천마시의 가구당 월평균수입(단위: 천 원)을 추정하려고 한다. 시에서는 표본평균이 실제 모평균의 50 이내에 있을 확률이 0.95이기를 원한다. 만일 천마시의 가구당 월수입이 몇몇 예외적인 경우를 제외하고는 1,000에서 2,600 사이에 있다면, 표본의 크기는 얼마가 적절한가?

모비율의 추정에서 표본크기의 결정

$$n = \left(\frac{z_{\frac{\alpha}{2}} \sqrt{pq}}{d} \right)^2 \approx \left(\frac{z_{\frac{\alpha}{2}}}{d} \right)^2 \hat{p}\hat{q}$$

예제 8-11

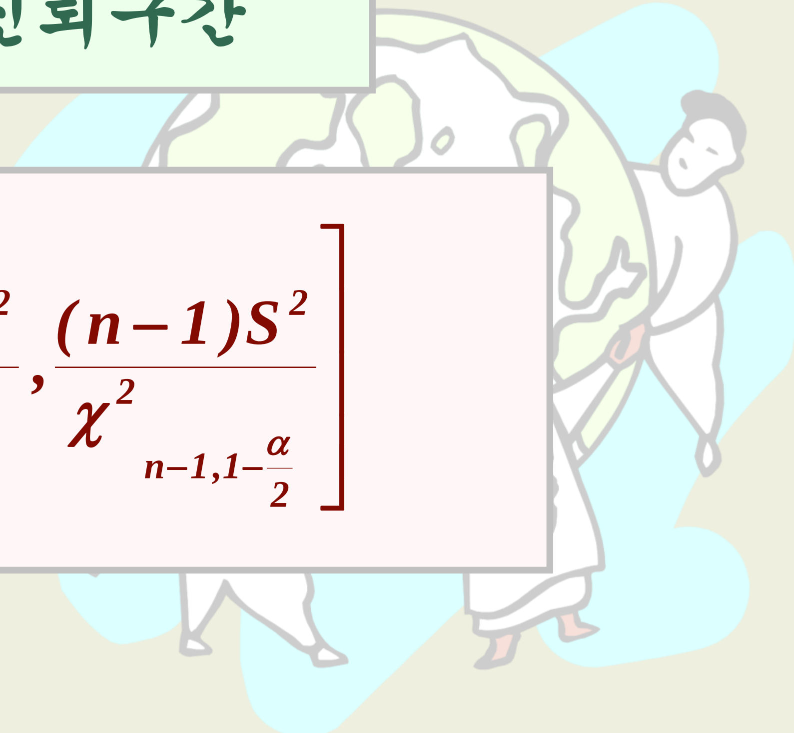
한 심리학 실험에서 자극에 대한 개인의 반응은 A형 또는 B형으로 나타난다고 하자. 실험자는 A형의 반응을 보이는 사람들의 비율 p를 추정하려고 한다. 그리고 그는 추정오차가 0.04 이하일 확률이 0.90이 되기를 원한다. 예비실험결과 A의 비율($\hat{p}$)이 0.6일 때, 표본의 크기는 얼마가 적당한가?

모분산의 신뢰구간

- 모집단이 정규분포일 때

σ^2에 대한 $100(1-\alpha)\%$신뢰구간

$$CI = \left[\frac{(n-1)S^2}{\chi^2_{n-1,\frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}} \right]$$



예제 8-12

경찰에서는 승용차의 속도를 측정할 수 있는 새로운 전자장비를 이용하려고 한다. 장비의 신뢰성을 측정하기 위해 30대의 승용차를 정확히 시속 100km로 관측점을 통과시키고 각 승용차의 속도를 새로운 장비에 의해 측정하였다. 모집단은 정규분포라고 가정한다. 표본분산은 $S^2=64$라고 할때 σ^2에 대한 95% 신뢰구간을 설정하시오.