5장. 통계적 추론

박명현

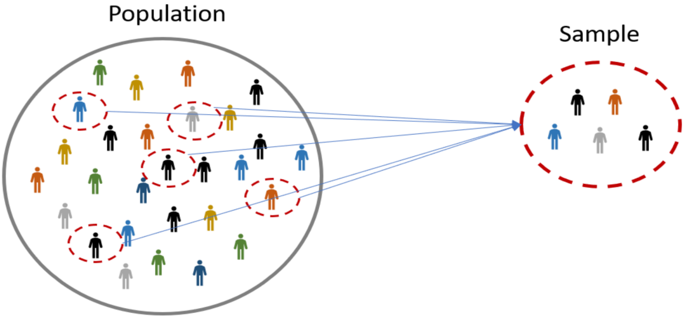
서울대학교 학부대학

Overview

1. 통계적 추론(statistical inference)
2. 모수의 추정
   - 점추정
   - 구간추정
   - 가설검정
   - 검정오류
3. 모평균에 관한 추론
4. 모비율에 관한 추론
5. 모분산에 관한 추론

통계적 추론(statistical inference)

- 통계적 추론: 표본에서 얻은 자료의 정보를 이용하여 모집단에 관한 추측이나 결론을 이끌어 내는 과정
- 추측의 대상: 모집단의 평균, 표준편차, 비율 등과 같은 모집단의 특성치(모수)



통계적 추론(statistical inference)

- 추정(estimation): N개 또는 무한의 원소로 된 모집단에서 n개의 표본을 추출한 후, 이를 이용하여 모집단의 모수의 값을 추측하고 그 오차범위를 제시하는 과정
  - 점추정(point estimation): 모수를 어떤 하나의 값으로 추측하는 것
    예) 청소년들의 하루 TV시청시간(μ)에 대한 점 추정값은 2.75시간이다. $\hat{\mu} = \bar{x} = 2.75$
  - 구간추정(interval estimation): 모수를 추정에 수반된 오차의 크기가 고려된 어떤 구간(신뢰구간)으로 추측하는 것
    예) '청소년들의 하루 TV 시청시간은 2시간에서 3시간 사이이다. '$\hat{\mu} \in (2, 3)$
- 가설검정(hypothesis testing): 모집단의 어떤 현상에 대한 예상 또는 주장이 타당한지 표본자료를 이용하여 판단하는 것

모수의 추정

- 모수(population parameter): 모집단의 특성을 나타내는 수치적 측도, θ로 표시. 예) μ, σ^2
- 추정량(estimator): 모수의 추정을 위해 사용되는 통계량, $\hat{\theta}$ 또는 $\hat{\theta}(X_1, X_2, \cdots, X_n)$. 예) $\hat{\mu} = \bar{X}$, $\hat{\sigma}^2 = S^2$
- 추정값(estimate): 관측치로부터 계산된 모수의 예측값. 예) $\bar{x}$, s^2
- 표집오차(sampling error): 모집단 전체를 관측하지 않고, 부분 집합인 표본에서만 관측함으로써 생기는 오차
- 표준오차(standard error): 추정량이 갖는 분산의 제곱근으로 즉, 추정량의 표준편차를 의미(변동성)

$$SE(\hat{\theta}) = \sqrt{Var(\hat{\theta})}$$

점추정
(Point Estimation)

점추정(point estimation)

- 모평균 μ에 대한 추정

  - 추정량: $\hat{\mu} = \bar{X}$

  - 표준오차(모표준편차 σ를 알 때):

$$SE(\hat{\mu}) = \sqrt{Var(\hat{\mu})} = \sqrt{Var(\bar{X})} = \sigma/\sqrt{n}$$

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

  - 표준오차(모표준편차 σ를 모를 때):

$$SE(\hat{\mu}) = s/\sqrt{n}$$

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

점추정(point estimation)

- 모비율 p에 대한 추정
  (n개의 랜덤표본에서 그 속성을 갖는 것의 개수: X)
  - 추정량: 표본비율을 사용 $\hat{p} = \frac{X}{n}$
  - 표준오차:$SE(\hat{p}) = \sqrt{Var(\hat{p})} = \sqrt{\frac{p(1-p)}{n}}$

- 모표준편차 σ의 보편적인 추정량은 표본표준편차 S이고, 모분산 σ^2의 추정량으로는 표본분산인 S^2을 사용한다.

$$\hat{\sigma} = \sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)} = S, \quad \hat{\sigma}^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1) = S^2$$

점추정의 예

Example (5.1)

한 타이어 제조 회사는 새로운 공정을 개발하며, 기존 공정에 비해 타이어 수명 평균이 증가할 것으로 기대하고 있다. 그러나 표준편차는 기존 공정의 표준편차 $\sigma = 5,000$km를 그대로 유지한다고 가정한다. 새로운 공정으로 제작된 시제품 $n = 100$개를 실험한 결과, 표본의 수명 평균이 $\bar{x} = 38,000$km로 나타났다. 이 결과를 바탕으로 새로운 공정으로 제작한 타이어의 모평균 수명 μ를 얼마로 추정할 수 있을까?

- 모평균 μ에 대한 추정량(estimator)은 $\hat{\mu} = \bar{X}$
- 점 추정값(estimate)은 표본평균 $\bar{x}$을 사용: $\hat{\mu} = \bar{x} = 38,000$km
- $\hat{\mu} = \bar{X}$의 표준오차(standard error):

$$SE(\hat{\mu}) = \sqrt{Var(\hat{\mu})} = \sqrt{Var(\bar{X})} = \frac{\sigma}{\sqrt{n}}$$

점추정의 예

Example

한 공장에서 생산하는 플라스틱판의 두께에서 변이 정도는 중요한 특성치로 간주된다. 이 공장에서 생산된 플라스틱판 중 표본 10개를 추출하여 두께를 측정한 결과는 다음과 같다. 이 자료를 기반으로 플라스틱 두께의 모표준편차 σ에 대한 추정값을 구하라.

$$226, 228, 226, 225, 232, 228, 227, 229, 225, 230$$

- sol) 표본평균 $\bar{x} = (226 + 228 + \cdots + 230)/10 = 227.6$이므로
- 표본분산은
  $s^2 = \{(226 - 227.6)^2 + \cdots + (230 - 227.6)^2\}/(10 - 1) = 5.1556$
- 따라서, 모표준편차 σ의 추정값은

$$\hat{\sigma} = s = \sqrt{5.1556} = 2.27$$

좋은 추정량

- 하나의 모수에 대해서 수많은 추정량이 존재한다.

$$\hat{\mu} = \bar{X} \quad or \quad \hat{\mu} = Median$$

이 중 좋은 추정량은 어떤 것인가?

$$\hat{\theta} - \theta = \hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta$$

- 추정량 $\hat{\theta}$의 표집오차 = 변동 + 편향
- 변동 = $\hat{\theta} - E(\hat{\theta})$: $\hat{\theta}$가 평균값에서 떨어진 거리
- 편향 = $E(\hat{\theta}) - \theta$: $\hat{\theta}$의 평균 표집오차 $E(\hat{\theta} - \theta)$
- 예)

$$\bar{X} - \mu = \bar{X} - E(\bar{X}) + E(\bar{X}) - \mu$$

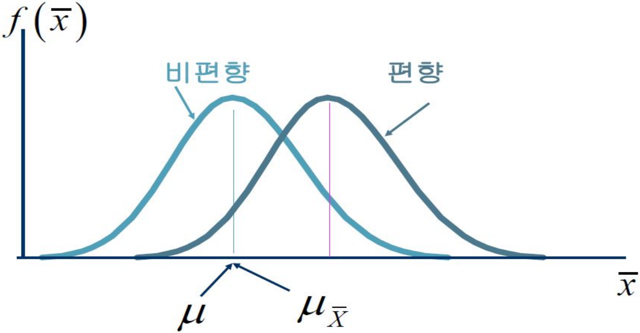
1. 비편향성(Unbiasedness)

- 추정량의 기대값이 모수와 같음
  - $E(\hat{\theta}) = \theta$이면, $\hat{\theta}$은 θ에 대한 비편향추정량(unbiased estimator)
  - $E(\hat{\theta}) \neq \theta$이면, $\hat{\theta}$은 편향추정량(biased estimator)
  - 편향(biased)$=E(\hat{\theta}) - \theta$
  - 편향이 없는 추정량. 편향 크기(bias)=0
- 비편향추정항을 사용하면 그 추정량이 모수와 일치한다는 보장은 없지만, 표본으로부터 얻어지는 추정량을 계산하는 작업을 무한히 반복할 경우, 추정량 값들이 평균적으로는 모수와 일치하게 됨을 의미.

"추정값은 항상 틀린 값이지만, 평균적으로는 틀리지 않으면 좋겠다"

비편향성(Unbiasedness)

- $E(\hat{\theta}_1) = \theta$ vs $E(\hat{\theta}_2) \neq \theta$



비편향성 예제

Example

모수 μ에 대한 점추정량으로서 표본평균 $\bar{X}$에 대한 비편향성 여부를 판단하여라.

sol) 표본평균 $\bar{X}$의 기대값을 구하면 모수 μ와 일치한다.

$$E(\bar{X}) = \frac{1}{n}E(X_1 + \cdots + X_n) = \frac{1}{n}\{E(X_1) + \cdots + E(X_n)\} = \mu$$

따라서 표본평균 $\bar{X}$는 모평균 μ에 대한 비편향추정량이 된다.

- 표본중위수(Median) : 대칭분포 모평균 μ에 대한 비편향추정량
- 표본비율 : 모비율 p의 비편향추정량
- 표본분산 : 모분산 σ^2의 비편향추정량 (n-1로 나누었던 이유)

2. 효율성(Efficiency)

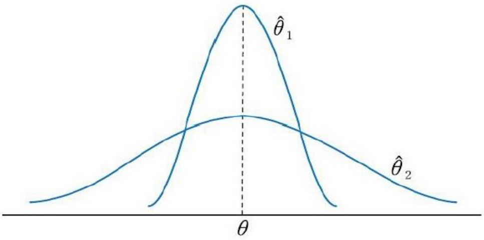
- 추정량의 분산: 추정량의 변동 $\hat{\theta} - E(\hat{\theta})$의 제곱의 기대값

$$E[(\hat{\theta} - E(\hat{\theta}))^2]$$

- 추정량의 변동이 작으려면 추정량의 분산이 작아야 하고, 분산이 작은 통계량은 상대적으로 더 효율적(relative efficient)이다.
- 추정량의 분산이 작다는 것은 표본을 여러 번 추출하여 추정량들을 각각 구해 보았을 때 그 변동폭이 작다는 뜻이다.

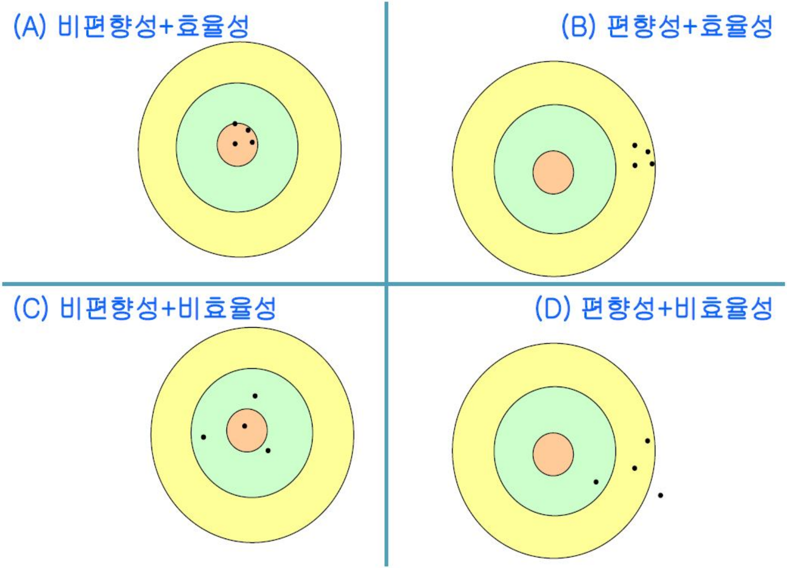
"추정값이 표본에 따라 그 때 그 때 많이 다르지 않으면 좋겠다"

예) 두 추정량 $\hat{\theta}_1$과 $\hat{\theta}_2$의 확률밀도함수가 다음과 같다고 가정하자.



- 두 추정량 모두 비편향추정량이지만 $\hat{\theta}_1$의 분산이 $\hat{\theta}_2$의 분산보다 작다.
- $\hat{\theta}_1$의 분포가 $\hat{\theta}_2$보다 모수 θ에 가깝게 몰려있으므로 모수에 가까운 값이 나올 가능성이 높다.
- 이때, $\hat{\theta}_1$이 $\hat{\theta}_2$보다 더 효율적인 추정량이라고 한다.

비편향성과 효율성(화살의 과녁 비유)



효율성(Efficiency)

- 앞의 그림에서 (A)는 알고자 하는 모수를 정확히 맞추면서, 표준오차도 작다. 즉, 추정량이 비편향성과 효율성을 모두 만족시키고 있다.

- 편향과 분산을 동시에 줄이는 것이 이상적이나, 기술적인 한계가 있어, 비편향추정량 중에서 분산을 최소로 하는 추정량을 찾는 전략을 쓰게 된다.

- 최소분산비편향추정량(minimum variance unbiased estimator):
  같은 표본에서 도출된 비편향추정량 중에서 분산이 최소가 되는 추정량
  예) 정규확률표본에서

$$Var(\bar{X}) = \frac{\sigma^2}{n}, \qquad Var(Median) = \frac{\pi}{2} \cdot \frac{\sigma^2}{n}$$

Example

어떤 회사에서 생산하는 베어링의 평균 직경 μ를 추정하기 위해 10개의 표본 $X_1, X_2, \cdots, X_{10}$을 추출하였다. 평균 직경에 대한 다음의 두 추정량을 비교하시오.

$$\hat{\theta}_1 = \frac{1}{10}\sum_{i=1}^{10} X_i, \qquad \hat{\theta}_2 = \frac{1}{5}\sum_{i=1}^{5} X_i$$

sol) 모평균이 μ이고 모분산이 σ^2일 때, 두 추정량의 평균과 분산은 다음과 같다.

$$E[\hat{\theta}_1] = \mu, \quad Var[\hat{\theta}_1] = \frac{1}{10}\sigma^2, \qquad E[\hat{\theta}_2] = \mu, \quad Var[\hat{\theta}_2] = \frac{1}{5}\sigma^2$$

- 두 추정량이 모두 비편향추정량이나, $\hat{\theta}_1$의 분산이 더 작으므로 μ에 대한 추정량으로 $\hat{\theta}_1$을 선호.
- 즉, 표본크기가 클수록 그에 따른 표본평균은 더 효율적인 추정량

3. 일치성(Consistency)

- 표본크기가 점차 커짐에 따라 점 추정량의 값이 모수에 근접해지는 경우
- 일치성은 추정량 $\hat{\theta}$가 모수 θ로 확률적으로 수렴(convergence in probability)한다는 것을 뜻하고 이는 다음과 같이 정의된다.

  - 임의의 양의 상수 ϵ에 대하여

$$\lim_{n\to\infty} \Pr(|\hat{\theta}-\theta| > \epsilon) = 0$$

  - $\Pr(\hat{\theta}=\theta) \xrightarrow[n\to\infty]{} 1$이 되는 $\hat{\theta}$은 일치추정량이다.

  " 현실적으로 표본을 무한으로 할 수는 없다. 하지만 표본이 커지면 추정값이 모수와 같아질 것이라는 보장이 있으면 좋겠다."

예) 표본평균 $\bar{X}$

- 표본평균의 분산은 $\frac{\sigma^2}{n}$으로 표본수가 늘어나면 0으로 수렴한다.
- 추정량의 분산이 0으로 수렴한다는 것은 그 추정량(표본평균)의 표본추출로 인한 변동이 점차 줄어들고, 결국 그 추정량의 평균(모평균)으로 수렴함을 의미한다.
- 즉, 임의의 양의 상수 ϵ에 대하여

$$\lim_{n\to\infty} \Pr(|\bar{X}_n - \mu| > \epsilon) = 0$$

일치성 예제

Example

모평균 μ에 대한 추정량으로 아래의 $\hat{\mu}$를 고려할 때, 이에 대한 비편향성 및 일치성 여부를 판단하여라.

$$\hat{\mu} = \frac{1}{n-1}\sum_{i=1}^{n} X_i$$

sol) 추정량 $\hat{\mu}$의 기대값은 다음과 같이 계산된다.

$$E(\hat{\mu}) = E\left(\frac{X_1 + \cdots + X_n}{n-1}\right) = \frac{n}{n-1}\mu \neq \mu$$

여기서 $E(\hat{\mu})$이 모수 μ와 같지 않으므로 $\hat{\mu}$은 편향추정량이다.

- 그리고 $\hat{\mu}$은 $\frac{n}{n-1}\bar{X}$이므로 $Var(\hat{\mu}) = \left(\frac{n}{n-1}\right)^2 \frac{\sigma^2}{n}$이다.
- n이 무한히 커지면 다음과 같이 $\hat{\mu}$의 평균은 μ, 분산은 0이 되어

$$E(\hat{\mu}) = \frac{n}{n-1}\mu \xrightarrow[n\to\infty]{} \mu$$

$$Var(\hat{\mu}) = \left(\frac{n}{n-1}\right)^2 \frac{\sigma^2}{n} = \frac{n}{(n-1)^2}\sigma^2 \xrightarrow[n\to\infty]{} 0$$

$\hat{\mu}$은 확률적으로 μ가 된다. 즉,

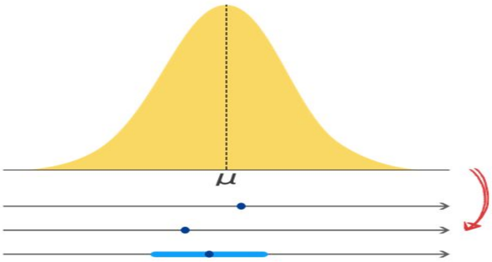
$$\Pr(\hat{\mu} = \mu) \xrightarrow[n\to\infty]{} 1$$

이므로 $\hat{\mu}$은 일치추정량이다.

구간추정
(Interval Estimation)

구간추정

- 모평균 μ는 하나의 상수값인 반면 표본평균 $\bar{X}$는 표본에 따라 많은 다른 값을 취할 수 있는 확률변수이다.
- 모수 μ를 하나의 값으로 추정하는 것이 아니라 추정에 수반된 오차의 크기가 고려된 구간으로 나타내자.



- 구간추정: 관심의 대상인 모수 θ가 포함되어지리라 여겨지는 구간을 표본으로부터 구해내는 것

신뢰구간(confidence interval, CI)

- 아래 식에서 관심모수 θ를 포함할 확률이 $1-\alpha$인 구간 (L,U)를 $100(1-\alpha)\%$ 신뢰수준(confidence level)에서의 신뢰구간이라고 한다.

$$1-\alpha = P(L \le \theta \le U)$$

  - 신뢰구간의 하한값 L 또는 상한값 U는 확률변수이다.

- 모수 θ의 $100(1-\alpha)\%$ 신뢰구간:

$$\text{point estimate} \pm \ z_{\alpha/2} \times SE(\hat{\theta})$$

- 신뢰수준: $100(1-\alpha)\%$

모평균 μ의 신뢰구간 - σ를 알 때

- μ를 구간추정하려면 표본평균 $\bar{X}$의 분포에 대한 정보가 필요하다.
- $\bar{X}$의 분포:

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \implies Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

- 정규분포의 성질로부터 다음이 성립한다.

$$\begin{aligned} 1 - \alpha &= \Pr(-z_{\alpha/2} < Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}) \\ &= \Pr(-z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}) \\ &= \Pr(\mu - z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}) \\ &= \Pr(\bar{X} - z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}) \end{aligned}$$

모평균 μ의 신뢰구간 - σ를 알 때

- 모평균 μ의 $100(1-\alpha)\%$ 신뢰구간:

$$\left(\bar{x} - z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}\right) \tag{1}$$

- 오차범위와 신뢰구간의 길이

$$z_{\alpha/2} SE(\hat{\mu}) = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \qquad 2 \times z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

- 주의: 식 (1)을 μ의 확률구간으로 읽으면 안 된다.

신뢰수준 90%의 신뢰구간의 의미

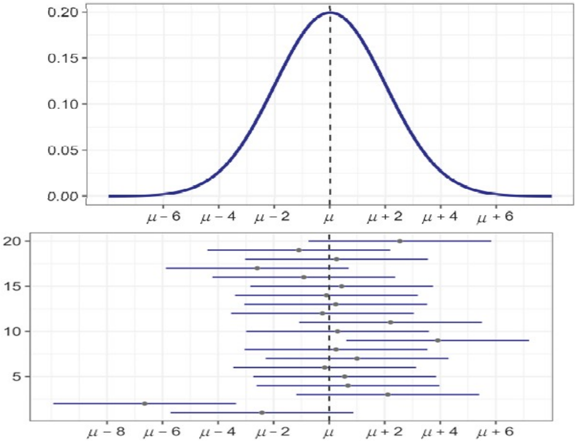
- 모수는 우리가 모르긴 하지만 불변의 값, 상수임.
  따라서 '모수μ가 신뢰구간에 포함될 확률'이라고 해석하면 안된다.
  ex) 원주율 π

예) 다음 그림에서 위의 분포는 정규모집단에서의 $\bar{X}$의 분포이고 아래 쪽에 주어진 구간들은 $n = 25$개의 표본을 20번 추출하여 구한 신뢰수준 90%의 신뢰구간 20개이다.

$$(\bar{x} - 1.645\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.645\frac{\sigma}{\sqrt{n}})$$

- 신뢰구간은 $\bar{x}$를 중심으로 대칭

신뢰수준 90%의 신뢰구간의 의미



- 20개의 구간 중에서 18개의 구간은 모평균 μ를 포함하고 있음
- 즉, 신뢰수준 90%의 의미는 신뢰구간의 가상적인 반복 사용에서 90%가 적중하리라는 뜻

신뢰구간

- 여러 번 표본을 뽑아 신뢰구간을 구하였을때 '신뢰구간이 모수μ를 포함할 비율'이 0.90이다. $\Rightarrow$ 그 비율이 신뢰수준 90%

- 신뢰수준($100(1-\alpha)\%$):

  - 여러 번 구한 신뢰구간 중 추정하고자 하는 모수를 포함하는 신뢰구간의 비율.
  - 흔히 90%, 95%, 99%를 사용하며 이는 각각 $\alpha = 0.1, 0.05, 0.01$에 대응되고 각각의 경우에 $z_{\alpha/2}$ 값은 다음과 같다.

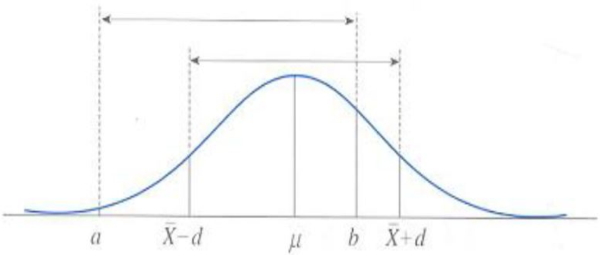
$$z_{0.05} = 1.645, \; z_{0.025} = 1.96, \; z_{0.005} = 2.576$$

- 신뢰수준 $100(1-\alpha)\%$ 신뢰구간의 의미:

  n개의 표본으로 표본평균 $\bar{x}$를 구하는 작업을 무한히 반복한다면 그 들 중 약 $100(1-\alpha)\%$의 표본에 대하여 모평균 μ가 신뢰구간에 속할 것이다.

신뢰구간

- 통계량의 분포가 대칭인 경우
  모평균에 대한 신뢰구간은 $(\bar{X} \pm d)$의 형태로 표본평균을 중심으로 같은 거리만큼 가감하는 것이 효율적이다.



- 신뢰구간 (a, b)도 같은 신뢰수준은 만족하나, 구간의 크기가 $(\bar{X} \pm d)$의 경우보다 길다.
  $\Longrightarrow$ 같은 신뢰수준에서 구간의 크기를 최소로 하는 것은 $(\bar{X} \pm d)$

모평균 μ의 신뢰구간 - σ가 미지인 정규모집단

- 모표준편차 σ가 미지인 정규모집단에서 크기 n인 랜덤표본을 추출:
$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$
- 이러한 표본분포로부터 다음을 알 수 있다.
$$P\{-t_{\alpha/2}(n-1) \le \frac{\bar{X}-\mu}{S/\sqrt{n}} \le t_{\alpha/2}(n-1)\} = 1-\alpha$$
- 위의 식을 정리하면 모평균 μ에 대한 $100(1-\alpha)\%$신뢰구간은 다음과 같다.
$$(\bar{x} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}})$$

- (정규모집단) σ를 모르더라도 표본수 n이 커지면 s가 σ에 근접하고, t분포는 z분포에 수렴한다. 따라서 n이 크면, σ를 모르는 경우의 신뢰구간은 σ를 아는 경우의 신뢰구간과 비슷해진다.
- (X_i가 정규분포를 따르지 않더라도) 표본수가 크면, 중심극한정리에 의하여 $\bar{X}$의 분포는 정규분포에 근접한다. 표본수가 작을 때에는 모집단 분포가 최소한 대칭에 가깝거나, 모집단 분포가 정규분포와 유사해야 한다.
- σ가 미지인 정규모집단에서 표본크기 n이 크면 정규모집단의 가정이 완화될 수 있다.

모평균 μ에 대한 신뢰구간 정리

- 가정사항:
  1. 정규모집단 $N(\mu, \sigma^2)$에서 크기 n인 랜덤표본을 추출
  2. 만약 정규분포가 아니라면 표본의 수가 충분히 크다. $n \geq 30$

- 모평균 μ에 대한 $100(1-\alpha)\%$신뢰구간
  - σ를 알 경우:

$$(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \ \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}})$$

  - σ를 모를 경우:

$$(\bar{x} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}, \ \bar{x} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}})$$

  - σ가 미지이지만 $n \geq 30$인 경우, by CLT:

$$(\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}}, \ \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}})$$

신뢰구간 예제

Example (타이어 제조 회사 continue..)

표준편차 $\sigma = 5,000$km, $n = 100$, $\bar{x} = 38,000$km일 때 95%신뢰구간을 구해보자. ($\alpha = 0.05$)

sol) $\bar{X}$의 분포는 근사적으로 정규분포를 따름: $\bar{X} \sim N(\mu, \frac{5,000^2}{100})$

- $\alpha = 0.05$이므로 $z_{\alpha/2} = z_{0.025} = 1.96$이고 $1.96\frac{\sigma}{\sqrt{n}} = 980$이다.
- $\bar{x} = 38,000$이므로 95%신뢰구간은 다음과 같다.

$$(\bar{x} - z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}) = (37,020, 38,980)$$

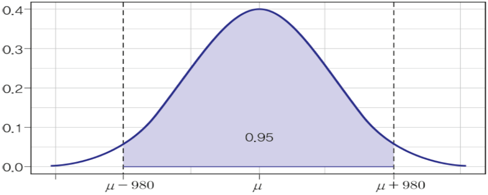


그림 5–1 표본평균 $\overline{X}$ 의 분포

신뢰구간 예제

Example

모표준편차가 $\sigma = 10$으로 알려진 정규모집단에서 $n = 25$개의 표본을 택한 결과 표본평균이 $\bar{x} = 40$이었다면, 모평균 μ의 추정값 $\hat{\mu} = \bar{x} = 40$의 90%오차범위와 신뢰구간을 구하여라.

sol) $\alpha = 0.1$이므로 $z_{\alpha/2} = 1.645$, 오차범위는

$$1.645\sigma/\sqrt{n} = 3.29$$

- μ의 90% 신뢰구간은 다음과 같다.

$$(\bar{x} - 1.645\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.645\frac{\sigma}{\sqrt{n}}) = (36.71, 43.29)$$

- μ의 95% 신뢰구간은 다음과 같다.($z_{\alpha/2} = 1.96$)

$$(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}) = (36.08, 43.92)$$

신뢰구간 예제

Example

물리학자인 뉴컴과 미켈슨은 실험실에서 일정한 거리만큼 떨어진 곳의 거울을 이용하여, 빛이 7,400m의 거리를 움직이는 시간을 측정하였다. 그리고 이 측정시간(초)을 (소요시간-24800$\times 10^{-9}$) $\times\ 10^9$으로 변환하여 얻은 $n = 64$개의 자료의 표본평균과 표본표준편차값은 각각 $\bar{x} = 27.750$, $s = 5.083$이다. 이 자료로부터 빛이 7,400m의 거리를 움직이는데 걸리는 시간 μ의 99% 신뢰구간을 구하여라. (정규모집단 가정)

sol) 99% 신뢰구간을 구하려면 $t(63)$ 백분위수를 필요로 한다. 그러나 t분포표에 이 값이 없으므로 가장 가까운 경우인 $t_{0.005}(60) = 2.660$ 값을 사용한다.

$$
\begin{aligned}
\bar{x} \pm t_{0.005}(63)s/\sqrt{n} \doteq \bar{x} \pm t_{0.005}(60)s/\sqrt{n} \\
&= 27.75 \pm 2.660 \times 5.083/\sqrt{64} \\
&= 27.750 \pm 1.690 \\
&= (27.06, 29.44)
\end{aligned}
$$

신뢰수준과 신뢰구간의 크기

- 신뢰수준은 높을수록, 구간의 크기는 작을수록 바람직.
- 하지만 신뢰수준과 구간의 크기는 서로 상충됨

  => 신뢰수준을 높이면 더 넓은신뢰구간이 얻어지고, 반대로 정교한 구간추정을 위하여 구간의 크기를 작게 하면 신뢰수준이 떨어지는 문제 발생
- 일반적으로 문제에 따라 신뢰수준을 고정, 신뢰수준을 만족하면서 구간의 크기를 가장 작게 하는 신뢰구간을 도출

  ex) $2 \cdot z_{\alpha/2} \times SE(\hat{\theta})$

  1. 분산이 작은 추정량을 사용
  2. 표본수를 늘임

표본크기의 결정 - 모평균 μ의 추정

- μ의 구간추정에서 신뢰수준을 높이면 더넓은 신뢰구간이 얻어지고, 추측되는 μ의 범위가 너무 넓어지게 되므로, 신뢰수준의 선택에서 이를 고려하여야 한다.
- 표본크기 n을 적절히 선택함으로써 오차범위나 신뢰구간의 길이를 조절할 수 있다.
- 모평균의 추정에서 표본크기의 결정:

  $100(1-\alpha)\%$ 신뢰수준으로 오차범위를 d 이하로, 또는 신뢰구간의 길이를 $2d$ 이하로 하기 위한 표본의 크기 (σ를 알 때)

$$n \geq (z_{\alpha/2} \cdot \sigma / d)^2$$

표본크기의 결정 예제

Example

중앙아메리카의 저소득층 원주민을 대상으로 혈청 내 콜레스테롤 양을 조사하려고 한다. 콜레스테롤 양은 근사적으로 정규분포를 따르고, 표준편차가 $\sigma = 30$(mg/L)라고 가정 하자. 모평균 μ에 대한 95% 오차범위를 5(mg/L)이하로 제한하려면, 몇 명의 저소득층 원주민을 표본조사해야 하는가?

sol) 필요한 표본의 크기를 n이라고 하면

$$n \geq (1.96\sigma/5)^2 = 138.2976$$

따라서, 필요한 표본의 크기는 139명 이상이어야 한다.

유의성검증(test of significance)과 가설검정(hypothesis testing)

가설검정(hypothesis testing)

- 가설검정: 모집단의 어떤 현상에 대한 예상 또는 주장이 옳은지 틀린지 표본자료를 이용하여 판단하는 것
- 통계적가설(hypothesis): 모집단의 특성 또는 모수에 대한 대립되는 두 가지 주장에 대하여 통계적으로 다루기 편리하도록 정리한 것
- 귀무가설(null hypothesis, H_0):
  모집단의 모수를 하나의 값 또는 구간 등으로 표시
  예) 한국 청소년들의 하루 평균 TV 시청 시간은 하루 3시간이다.

$$H_0 : \mu = 3$$

- 대립가설(alternative hypothesis, H_1):
  귀무가설에서 제시하는 모수의 값을 제외한 나머지 영역에서 모수의 값을 정의
  예) 청소년들의 하루 평균 TV 시청 시간은 3시간이 아니다.

$$H_1 : \mu \neq 3$$

가설설정 예시

예) 무죄추정의 원칙: 피고인 또는 피의자는 유죄판결이 확정 될 때 까지는 무죄로 추정한다는 원칙

- 귀무가설(null hypothesis, H_0): 피고는 무죄이다.
- 대립가설(alternative hypothesis, H_1): 피고는 유죄이다.

- 가설 설정

- 특별한 증거 없이는 현재로서는 참이라고 여겨지는 가설을 귀무가설로 하여야 한다.
- 연구자가 통계적으로 증명하고 싶은 것. 즉, 자료에서 뚜렷한 증거로 입증하고자 하는 가설을 대립가설로 상정한다.

- 귀무가설과 대립가설이 정해진 후 표본이 귀무가설의 주장을 뒷받침하지 못하는 상황이 되면 귀무가설을 기각하게 된다.

가설 설정 규칙

1. 귀무가설은 모수를 특정한 값으로 표현한다.($H_0 : \mu = \mu_0$)
2. 귀무가설에는 반드시 " = "이 포함된다.
3. 대립가설은 비교하는 값의 양쪽을 다 고려하는 양측검정(two-sided hypothesis)과 한쪽만 고려하는 단측검정(one-sided hypothesis)이 있다.

$H_1 : \mu \neq \mu_0$ -> 양측검정
$H_1 : \mu > \mu_0$ -> 단측검정
$H_1 : \mu < \mu_0$ -> 단측검정

- 문제에서 양측검정과 단측검정 중 어느 것을 사용해야 하는지에 대한 판단은 어느 가설이 문제의 내용을 더 잘 반영하는 지를 검토하여 분석자가 결정

가설검정 예제

Example

한 페인트 제조회사에서 생산하는 특수 페인트의 건조시간은 평균 75분, 표준편차 9.4분인 것으로 알려져 있다. 연구소에서는 새로운 첨가제를 사용한 시제품을 사용하면 평균 건조시간이 줄어들 것이라고 주장한다. 실제 시제품 100개를 생산하여 건조시간을 조사한 결과 $\bar{x} = 73.5$가 나왔다. 새로운 첨가제를 사용한 시제품을 사용시 건조시간이 단축된다는 뚜렷한 증거가 있으면 이를 전체 공정에 적용하려고 한다.

- 새로운 첨가제를 사용한 특수 페인트의 평균 건조시간을 μ라고 하자.
- 귀무가설과 대립가설은 각각

$$H_0 : \mu = 75 \qquad H_1 : \mu < 75$$

- $n = 100$으로 충분히 크므로 $\bar{X}$의 분포는 $N(\mu, 9.4^2/100)$이다.

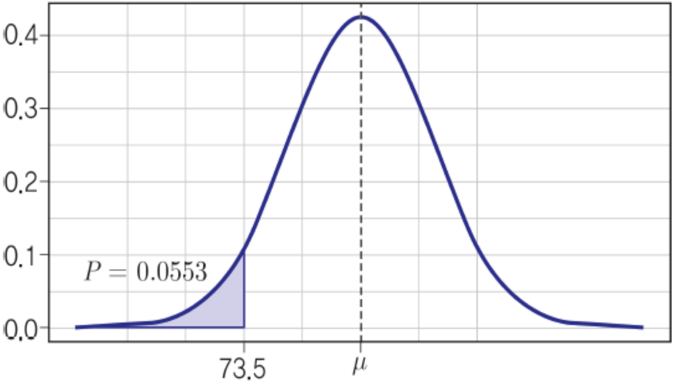


그림 5-3 왼쪽 단측 대립가설일 때 유의확률

- 표본평균 $\bar{X}$가 75보다 작은 쪽으로 멀어질수록 이는 대립가설을 지지하는 증거가 강해진다.

- 귀무가설하에서 관측값 73.5보다 더 대립가설을 지지하는 방향으로 표본평균이 나올 확률은 다음과 같다.

$$\begin{aligned} P &= P\{\bar{X} \le 73.5 | \mu = 75\} \\ &= P\left\{ \frac{\bar{X} - 75}{9.4/\sqrt{100}} \le \frac{73.5 - 75}{0.94} | \mu = 75 \right\} \\ &= P\left\{ \frac{\bar{X} - 75}{9.4/\sqrt{100}} \le -1.5956 | \mu = 75 \right\} \doteq 0.0548 \end{aligned}$$

- 확률값이 매우 작으므로 실제 관측된 증거($\bar{x} = 73.5$)보다 더욱 극단적으로 사건이 발생할 가능성은 적다. 즉 더 적은 건조시간이 나올 확률은 매우 적으며, $\bar{x} = 73.5$의 값으로도 충분히 건조시간이 줄어들었다는 것을 강하게 시사한다고 볼 수 있다.
- 즉, 관측값 $\bar{x} = 73.5$의 값은 대립가설 $\mu = 75$보다 충분히 작아 귀무가설을 기각하게 된다.

검증통계량(test statistics)

- 검증통계량(test statistics): 검증에 사용하기 위하여 표본자료에서 구한 통계량. 일반적으로 검정하려는 모수의 점추정량이 되기도 하고, 이 점추정량을 표준화 한 것을 사용하기도 함.
- 관심 모수 μ: 검정통계량(Z)과 그것의 관측값(z_0)은 다음과 같다.

1. σ를 알 때

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1), \quad z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

2. σ를 모를 때

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1), \quad t_0 = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$$

P값과 기각역

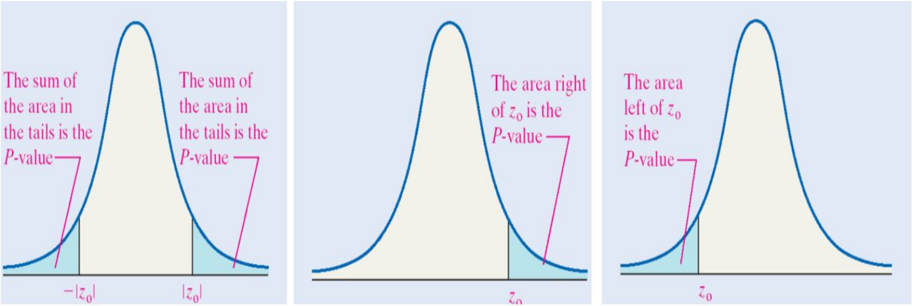
- 유의확률 or P값(significance probability, or P-value):
  귀무가설이 사실일 때, 검증통계량이 실제 관측된 값보다 대립가설을 지지하는 방향으로 더욱 치우칠 확률.

- 유의수준(significance level):
  귀무가설 H_0에 대한 반증의 강도에 대하여 유의수준을 미리 정해놓고 유의확률(P값)을 유의수준값과 비교함. α로 표시. 흔히 $\alpha = 0.05$, $\alpha = 0.01$을 사용.

P값과 기각역

- 대립가설과 P값
  (a) $H_1 : \mu > \mu_0$의 경우에는 $P = P(Z \geq z_0)$
  (b) $H_1 : \mu < \mu_0$의 경우에는 $P = P(Z \leq z_0)$
  (c) $H_1 : \mu \neq \mu_0$의 경우에는 $P = P(|Z| \geq |z_0|)$

(c) (a) (b)



P값과 기각역

- p값에 의한 판정

  - p값 $> \alpha$이면, H_0기각 못함.

  - p값 $\leq \alpha$이면, H_0기각.

  "조사 결과가 통계적으로 유의하다"

- 통계적 유의성

: 유의확률(p-value)이 유의수준 α이하이면, 조사 결과가 유의수준 α에서 통계적으로 유의하다고 하며 이는 귀무가설에 대한 반증의 강도가 지정된 수준보다 강함을 뜻한다.

P값과 기각역

- 기각역(rejection region): 지정된 수준에서 H_0가 사실이 아니라고 기각하게 되는 즉, 통계적 유의성을 주장할 수 있는 검증통계량의 영역
- 유의수준, α는 보통 $\alpha = 0.05$, $\alpha = 0.01$을 사용.
- 대립가설과 기각역

  (a) $H_1 : \mu > \mu_0$의 경우에는 $z_0 \geq z_\alpha$

  (b) $H_1 : \mu < \mu_0$의 경우에는 $z_0 \leq -z_\alpha$

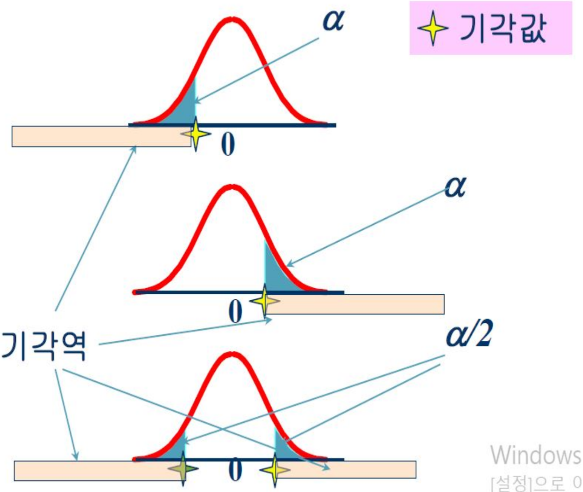
  (c) $H_1 : \mu \neq \mu_0$의 경우에는 $|z_0| \geq z_{\alpha/2}$

P값과 기각역

$H_0: \mu = 3$
$H_a: \mu < 3$

$H_0: \mu = 3$
$H_a: \mu > 3$

$H_0: \mu = 3$
$H_a: \mu \neq 3$



P값과 기각역

- 기각역에 의한 판정

  - 유의수준 α에 따라 기각역을 먼저 정하고,

  - 검증통계량의 값이 이러한 기각역에 속하면:

    "유의수준 α에서 귀무가설 H_0를 기각한다."

  - 검증통계량의 값이 기각역에 속하지 않으면:

    "유의수준 α에서 귀무가설 H_0를 기각할 수 없다."

가설검정의 단계 요약

1. 모집단의 특성에 대한 주장, 즉 가설을 설정
   (귀무가설 H_0와 대립가설 H_1)
2. 검정의 유의수준 α를 정함
3. 표본을 추출하여 귀무가설 하에서 검정통계량을 계산
4. 검정통계량에 대한 P값을 구하여 가설에 대한 결론 도출
   or 기각역에 속하는지 검토하여 가설에 대한 기각 여부 결정

모평균에 대한 가설검정; σ를 알 경우

- 모평균의 유의성검증(모표준편차 σ를 알고 유의수준이 α일 때) 귀무가설이 $H_0 : \mu = \mu_0$인 경우에 검정통계량은

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

이고, 이의 관측값을 z_0라고 할 때, 다음이 성립한다.

	유의확률 P	유의수준 α의 기각역
(a) $H_1 : \mu > \mu_0$	$P = P(Z \geq z_0)$	$z_0 \geq z_\alpha$
(b) $H_1 : \mu < \mu_0$	$P = P(Z \leq z_0)$	$z_0 \leq -z_\alpha$
(c) $H_1 : \mu \neq \mu_0$	$P = P(\lvert Z\rvert \geq \lvert z_0\rvert)$	$\lvert z_0\rvert \geq z_{\alpha/2}$

- - P값에 의한 판정: $P \leq \alpha$이면, H_0기각
  - 기각역에 의한 판정: 위의 기각역 조건이면 H_0기각

Example

한 타이어 제조 회사에서 생산중인 타이어의 수명시간은 평균이 37,000km이고, 표준편차는 5,000km인 것으로 알려져 있다. 이제, 타이어의 수명을 증가시키는 공정을 개발하고 있는 이 회사의 연구소에서 개발중인 새공정에 의해 시제품을 100개 생산하여 조사한 결과 평균수명이 $\bar{x} = 38,000$km이었다. 새 공정에 의한 타이어 수명시간의 표준편차 $\sigma = 5,000$km로 유지된다고 할 때, 유의수준 $\alpha = 0.05$하에서 이 연구소의 조사결과는 새 공정이 성공적임을 뜻하는가?

sol) 새 공정에 의한 타이어의 평균수명을 μ라고 하자. 타이어의 수명을 증가시키는 공정을 개발하고자 하므로 귀무가설과 대립가설은 다음과 같다.

$$H_0 : \mu = 37,000 \quad H_1 : \mu > 37,000$$

- 표본크기 $n = 100$이므로 Z의 분포는 근사적으로 표준정규분포를 따른다.
- $\bar{x} = 38,000$이므로 검증통계량의 관측값을 계산하면

$$z_0 = \frac{\bar{x} - 37,000}{5,000/\sqrt{100}} = 2$$

- P값으로 판정:
  유의확률은 다음과 같고, 유의수준 $\alpha = 0.05$보다 작아 귀무가설을 기각한다.

$$P = P\{Z \geq 2\} \doteq 1 - 0.9772 = 0.0228 < 0.05$$

- 기각역으로 판정: 표준정규분포표에서 상방 5%백분위수인 $z_{0.05} = 1.645$와 관측값 z_0를 비교한다.
- 기각역은 $z_0 \geq 1.645$이므로 $z_0 = 2$값이 다음 그림과 같이 기각역안에 속하므로 관측결과가 유의수준 5%에서 통계적으로 유의하다.

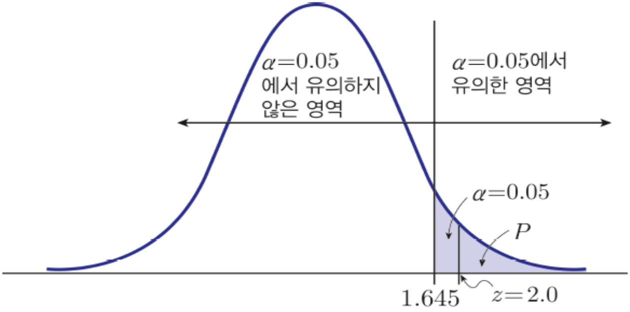


그림 5–5 유의수준 5%에서 유의한 영역

- 즉 새 공정이 상당히 성공적이라는 증거라고 할 수 있다.

모평균에 대한 가설검정; σ를 알 경우

Example

한 제약 회사에서 약품의 주성분 농도가 0.85(gr/L)로 유지되는 제조법의 개발을 목표로 연구중이다. 3번의 반복 측정에서 주성분 농도의 평균이 $\bar{x} = 0.8404$(gr/L)이었을 때 이 제조법에 의한 주성분의 실제 농도 μ가 0.85(gr/L)가 아니라고 의심할만한 증거가 있는가? 농도의 측정값들은 실제의 농도 μ를 평균으로 하고 표준편차가 $\sigma = 0.0123(gr/L)$인 정규분포를 따른다고 한다. 유의수준은 $\alpha = 0.05$라고 설정하자.

sol) 귀무가설과 대립가설은 다음과 같다.

$$H_0 : \mu = 0.85 \quad H_1 : \mu \neq 0.85$$

- 검증통계량의 관측값을 계산하면

$$z_0 = \frac{\bar{x} - 0.85}{0.0123/\sqrt{3}} = -1.35$$

- P값으로 판정: 양측가설이고, 모집단이 정규분포를 따르므로

$$P = P\{|Z| \geq |-1.35|\} = P(Z \leq -1.35) + P(Z \geq 1.35)$$
$$= 2 \times 0.0885 = 0.1770 > 0.05 = \alpha$$

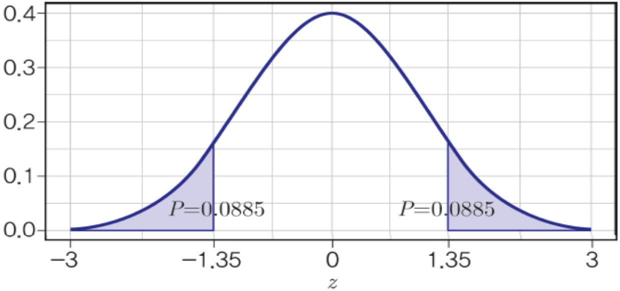


그림 5−4 양측 대립가설일 때 유의확률

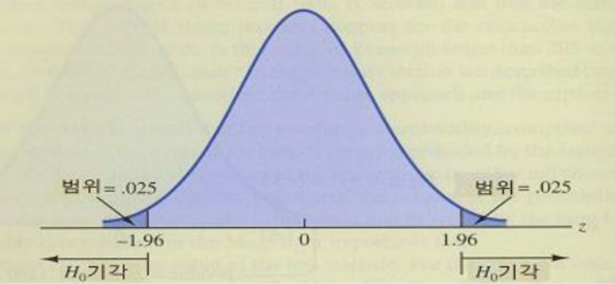
- P값이 유의수준 $\alpha = 0.05$보다 크기 때문에 귀무가설을 기각할 수 없다.

- 기각역으로 판정:

  $\alpha = 0.05$이고 양측검증이므로 기각역은 다음과 같다.

$$z_0 \leq -1.96 \ \textit{or} \ z_0 \geq 1.96$$

Figure: 유의수준 5%하에서의 기각역



$z_0 = -1.35$이므로 기각역에 속하지 않는다. 즉 실제 농도가 0.85가 아니라고 의심할만한 뚜렷한 증거가 없다고 할 수 있다.

Example (5.6)

한 제약회사에서 생산하고 있는 진통제는 진통효과가 나타나는 시간이 평균 30분인 것으로 알려져 있다. 새로 개발한 진통제의 진통효과가 더 빠른가를 확인하기 위하여, 50명의 환자를 랜덤추출하여 시간을 관측한 결과 평균이 $\bar{x} = 28.5$(분)이었다. 새로운 진통제에 의한 진통효과가 나타나는 시간이 표준편차가 $\sigma = 5$(분)인 정규분포를 따른다고 하고, 적절한 가설을 유의수준 5%에서 검정하여라.

sol) 새로운 진통제의 효과가 나타나는 시간의 평균을 μ(분)라고 하면, 귀무가설과 대립가설은 다음과 같다.

$$H_0 : \mu = 30 \quad H_1 : \mu < 30$$

- 검증통계량의 관측값을 계산하면

$$z_0 = \frac{\bar{x} - 30}{5/\sqrt{50}} = \frac{28.5 - 30}{5/\sqrt{50}} = -2.12$$

- 기각역으로 판정:
  $\alpha = 0.05$이고 왼쪽 단측검증이므로 기각역은 다음과 같다.

$$z_0 \leq -z_{0.05} = -1.645$$

- $z_0 = -2.12$이므로 기각역에 속한다. 따라서 유의수준 5%에서 귀무가설을 기각한다.
- 즉 새로운 진통제의 효과가 더 빠르다는 증거가 지정된 수준 5%보다 강하게 나타났다.

- 모평균의 유의성검증(모표준편차 σ가 미지인 정규모집단) 귀무가설이 $H_0 : \mu = \mu_0$인 경우에 검정통계량은

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

이고, 이의 관측값을 t_0라고 할 때, 다음이 성립한다.

	유의확률 P	유의수준 α의 기각역
(a) $H_1 : \mu > \mu_0$	$P = P(T \geq t_0)$	$t_0 \geq t_\alpha(n-1)$
(b) $H_1 : \mu < \mu_0$	$P = P(T \leq t_0)$	$t_0 \leq -t_\alpha(n-1)$
(c) $H_1 : \mu \neq \mu_0$	$P = P(\lvert T\rvert \geq \lvert t_0\rvert)$	$\lvert t_0\rvert \geq t_{\alpha/2}(n-1)$

- \- p값에 의한 판정: $P \leq \alpha$이면, H_0기각
  \- 기각역에 의한 판정: 위의 기각역 조건이면 H_0기각

모평균에 대한 가설검정; σ를 모를 때(t 검정)

- σ를 모를 경우 μ에 대한 가설검정 유의사항

- σ를 모를 때의 가설검정은 모집단이 정규분포를 따라야 한다.
- 표본수 n이 커지면 자유도 $n-1$도 커지면서 t분포는 정규분포에 수렴한다. 따라서 표본수가 크면($n \geq 30$), σ를 모를 때라도 σ를 s로 대치하여 정규 분포(Z 통계량)를 사용하고 σ를 알 경우의 검정을 사용해도 큰 무리가 없다.

Table: 모평균에 대한 가설검정 정리

분포 모집단	σ 아는 경우	σ 모르는 경우(s사용)
정규분포	z 통계량	t 통계량[a]
대표본 비정규분포	근사 z 통계량	근사 z 통계량[b]

a: n이 크면 근사 Z 사용 가능
b: n이 크면 근사 T 사용 가능

모평균에 대한 가설검정; σ를 모를 때

Example

전구회사에서 생산하는 전구의 평균수명은 1950시간으로 알려져 있다. 새로이 개발중인 전구의 평균수명 μ가 기존의 전구보다 수명이 더 길다고 할 수 있는가를 확인하기 위하여, 9개의 시제품을 생산하여 그 수명시간을 조사한 결과가 다음과 같다.

2000, 1975, 1900, 2000, 1950, 1850, 1950, 2100, 1975

적절한 가설을 세우고, 수명의 분포가 정규분포라는 전제하에서 가설에 대한 유의수준5%의 검정을 하여라. 또한 유의확률은 얼마인가?

sol) 문제의 뜻에 적절한 귀무가설과 대립가설은 다음과 같다.

$$H_0 : \mu = 1950 \quad H_1 : \mu > 1950$$

- 주어진 자료에서 표본평균과 표본표준편차를 계산하면

$$\bar{x} = (2000 + 1975 + \cdots + 2100 + 1975)/9 = 1966.7$$

$$s = \sqrt{\{(2000 - 1966.7)^2 + \cdots + (1975 - 1966.7)^2\}/8} = 69.6$$

- 이로부터 t검정통계량의 관측값을 계산하면

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{1966.7 - 1950}{69.6/\sqrt{9}} = 0.718$$

- 기각역으로 판정:
오른쪽 단측검증이므로 기각역은 $t_0 \geq t_{0.05}(8) = 1.86$ 이고
$t_0 = 0.718 < 1.86$이므로 귀무가설을 기각할 수 없다.

- P값으로 판정:
  유의확률은 다음의 확률을 계산하여야 한다.

$$P = P(T \geq 0.718), \;\; T(8)$$

- 그런데 t분포표에는 그 값이 나와있지 않고, 다음과 같이 0.718은 $t(8)$분포의 상방 25%백분위수(0.706)와 20%백분위수(0.889) 사이의 값임을 알 수 있다.

$$(0.25 =)P(T \geq 0.706) > P(T \geq 0.718) > P(T \geq 0.889)(= 0.2)$$

  따라서 유의확률이 항상 유의수준보다 큰 것을 알 수 있고, 실제로 계산해보면 유의확률은 $P = 0.247$이다.

Example

반도체의 특정한 용도에 사용되는 실리콘 다이오드는 0.6볼트의 가동 전압이 요구되며, 이러한 목표에서 벗어나면 조정이 필요하다. 이러한 실리콘 다이오드를 제조하는 공장에서 120개를 랜덤추출하여 조사한 결과 가동전압(볼트)의 평균과 표준편차가 각각 다음과 같았다.

$$\bar{x} = 0.62, \qquad s = 0.11$$

적절한 가설을 세우고, 유의수준 5%에서 가설을 검정하여라. 모평균에 대한 95% 신뢰구간도 구해 보자.

sol) 제조 공정에 의한 가동전압의 모평균을 μ볼트라고 하면, 귀무가설과 대립가설은 다음과 같다.

$$H_0 : \mu = 0.6 \quad H_1 : \mu \neq 0.6$$

- 이 경우에 표본크기가 $n = 120 > 30$이므로 정규모집단의 가정이 없어도 대표본근사에 의하여 t검정을 할 수 있다.
- t검정통계량의 관측값을 계산하면

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{0.62 - 0.6}{0.11/\sqrt{120}} = 1.992$$

- 또한, 표본크기 $n = 120 > 30$이므로 $\bar{X}$의 분포도 정규분포에 근접한다고 할 수 있으므로 $t_{0.025}(n-1)$대신 $z_{0.025} = 1.96$를 사용할 수 있다.
- 기각역으로 판정:

  $\alpha = 0.05$이므로 기각역은 $|t_0| \geq z_{0.025} = 1.96$이고,
  $|t_0| = 1.992 > 1.96$이므로 유의수준 5%에서 귀무가설을 기각한다.

- p값으로 판정:
  유의확률은 $P = P(|T| \geq 1.992)$이고 정규분포에 근사하므로 표준정규분포표로부터 그 값은 다음과 같고 귀무가설을 기각한다.

$$P \doteq 2 \times 0.0233 = 0.0466 < \alpha$$

- 95% 신뢰구간에서 $t_0.025(119) = 1.98$이므로,

$$\begin{aligned}\bar{x} \pm t_{0.025}(119)s/\sqrt{n} &= 0.62 \pm 1.98 \times 0.11/\sqrt{120} \\ &= (0.6001, 0.6399)\end{aligned}$$

신뢰구간과 가설검정

- 신뢰수준이 $100(1-\alpha)\%$인 신뢰구간을 이용하여, 유의수준 α, 즉 $100\alpha\%$의 양측검정 결과를 예측할 수 있다.

- 모평균이 μ이고 모분산이 σ^2인 정규모집단에서 크기 n인 랜덤표본을 뽑았다고 하자.
- μ에 대한 $100(1-\alpha)\%$ 신뢰구간은 다음과 같다.

$$\bar{x} \pm z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}$$

어떤 b라는 값이 이 구간 안에 들어있으면 다음의 식이 성립한다.

$$|\bar{X} - b| < z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}, \qquad \frac{|\bar{X} - b|}{\sqrt{\frac{\sigma^2}{n}}} < z_{\alpha/2}$$

- 이제 $H_0 : \mu = b$인 양측검정을 생각해 보자.
- 검정통계량:

$$Z = \frac{\bar{X} - b}{\sqrt{\frac{\sigma^2}{n}}}$$

(2)에 의해서 귀무가설은 유의수준 α에서 기각되지 못한다.

- 신뢰구간은 모수의 값으로서 기각되지 못하는 귀무가설의 집합이라고 볼 수 있다.

검정오류

- 검증오류(test error): 가설을 채택하거나 기각할 때 확률적으로 틀릴 가능성

\- 제1종의 오류(type I error:α):

귀무가설이 맞는데도 잘못하여 이를 기각할 확률
(=검정의 유의수준 significance level)

$$P[reject\ H_0\ |H_0\ is\ true] = \alpha$$

\- 제2종의 오류(type II error:β):

대립가설이 사실임에도 불구하고 귀무가설을 기각하지 못하게 되는 확률

$$P[fail\ to\ reject\ H_0\ |H_0\ is\ false] = \beta$$

검정오류 예시

예) H_0: 도둑이 아니다. H_1: 도둑이다.

- α: 선량한 사람을 도둑으로 몰아 잡아들임
  $P[\textit{reject } H_0 \mid H_0 \textit{ is true}]$
- β: 도둑을 도둑이 아니라고 놓아줌
  $P[\textit{fail to reject } H_0 \mid H_0 \textit{ is false}]$

예) H_0: 임신이 아니다. H_1: 임신이다.

검정오류

- 귀무가설이 사실일 때, 이를 기각하지 않을 확률(=$1-\alpha$)
- 검정력(power of the test): 대립가설 하의 평균값 $\mu = \mu_1$이 사실일 때, 귀무가설 H_0을 기각할 확률(=$1-\beta$)

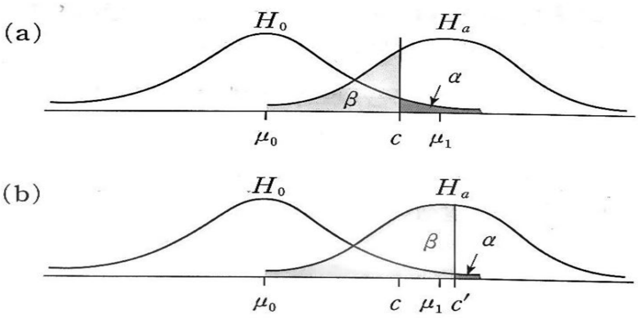
Table: 귀무가설의 진위와 의사결정

검정결과	실제상황	
	H_0가 사실	H_1이 사실
H_0를 기각못함	옳은 결정($1-\alpha$)	제2종의 오류(β)
H_0를 기각(H_1을 택함)	제1종의 오류(α)	옳은 결정($1-\beta$)

검정오류의 최소화

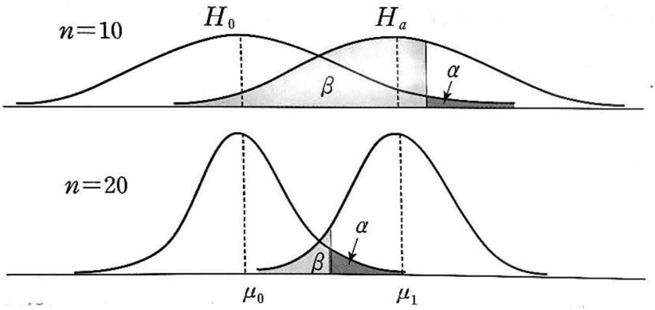
- 두 가지 검정오차인 α와 β를 최소로 하는 기각역을 구하는 것이 최선 => 그러나 α를 너무 작게 하려다 보면 β가 너무 커지는 모순 관계가 있음

  예) $H_0 : \mu = \mu_0$, $H_1 : \mu > \mu_0$이고, H_1의 대표값이 μ_1일 때, 임계점 c에 의하여 α와 β가 결정된다.



검정오류의 최소화

- 실제문제에선 주어진 α를 만족시키는 기각역중에서 β를 최소로 하는 기각역을 선택
  $\Longrightarrow (1-\beta)$가 클수록 검정력이 큼
- 표본수가 커지면 일반적으로 모수추정량의 분산이 작아지고
  고정된 α에 대하여 표본수 n이 증가함에 따라임계점이 왼쪽으로 이동하므로 β가 작아지고 검정력$(1-\beta)$이 커짐



검정오류 예제

Example

example 5.6(64p)에서 제1종의 오류를 범할 확률을 구하시오.

sol) 가설은 $H_0 : \mu = 30 \quad H_1 : \mu < 30$이고 왼쪽 단측검정이다.

- 유의수준 α의 기각역은 다음과 같다.

$$Z = \frac{\bar{X} - 30}{5/\sqrt{50}} \leq -z_\alpha$$

- 따라서 제1종의 오류를 범할 확률은 다음과 같고 이는 유의수준 α와 같음을 알 수 있다.

$$P\left\{\frac{\bar{X} - 30}{5/\sqrt{50}} \leq -z_\alpha | \mu = 30\right\} = \alpha$$

- 유의수준을 α로 지정한다는 것은 제1종의 오류를 범할 확률의 허용한계를 α로 제시한다는 것으로 해석할 수 있다.

검정오류 예제

Example

example 5.6에서 새로운 진통제의 효과가 나타나는 평균시간 μ가 28분이면 이는 괄목할 만한 개선이라고 한다. 이제 실제로 $\mu = 28$일 때 유의수준 5%의 기각역에 의한 제2종 오류를 범할 확률을 구하여 보자.

sol) 유의수준 5%의 기각역: $\frac{\bar{x}-30}{5/\sqrt{50}} \leq -1.645 = -z_{0.05}$

- 제2종오류 β는 $P[(fail\ to\ reject\ H_0\ |H_0\ is\ false]$이므로

$$\begin{aligned}
\beta &= P\left\{\frac{\bar{X}-30}{5/\sqrt{50}} > -1.645|\mu = 28\right\} \\
&= P\left\{\frac{\bar{X}-28}{5/\sqrt{50}} + \frac{28-30}{5/\sqrt{50}} > -1.645|\mu = 28\right\} \\
&= P\left\{\frac{\bar{X}-28}{5/\sqrt{50}} > 2\sqrt{2} - 1.645|\mu = 28\right\} \\
&= P\{Z > 1.183\} \doteq 0.1190
\end{aligned}$$

검정오류 예제

Example

example 5.6에서 실제로 $\mu = 28$일 때 제2종의 오류를 범할 확률이 $\beta = 0.10$이하가 되도록 하며, 유의수준이 $\alpha = 0.05$인 검정법을 사용하려고 한다. 이 때, 요구되는 표본의 크기를 구하여라.

sol) 요구되는 표본의 크기를 n이라고 하면, $\mu = 28$일 때 제2종의 오류를 범할 확률은 다음과 같이 주어진다.

$$\begin{aligned}
\beta &= P\left\{\frac{\bar{X}-30}{5/\sqrt{n}} > -1.645 | \mu = 28\right\} \\
&= P\left\{\frac{\bar{X}-28}{5/\sqrt{n}} + \frac{28-30}{5/\sqrt{n}} > -1.645 | \mu = 28\right\} \\
&= P\left\{Z > \frac{30-28}{5/\sqrt{n}} - 1.645\right\}
\end{aligned}$$

- 따라서, 이러한 확률을 0.10이하로 하는 다음의 식을 만족하는 n을 계산해보자.

$$P\left\{Z > \frac{30-28}{5/\sqrt{n}} - 1.645\right\} \leq 0.10$$

- 표준정규분포표로부터 $z_{0.10} = 1.282$이므로

$$\frac{30-28}{5/\sqrt{n}} - 1.645 \geq 1.282$$

$$n \geq \left(\frac{1.645 + 1.282}{(30-28)/5}\right)^2 = 53.7\cdots$$

- 그러므로 요구되는 표본의 크기는 $n = 54$이다.

검정오류 예제

Example

example 5.6에서 유의수준 5%의 검정법에 대하여 $\mu = 28$에서의 검정력을 구하여라.

sol) 검정력은 $1-\beta$이므로 앞의 예제의 결과로부터

$$P\left\{\frac{\bar{x}-30}{5/\sqrt{50}} \leq -1.645 | \mu = 28\right\} = 1 - 0.1190 = 0.8810$$

임을 알 수 있다.

모비율의 추정

- 불량률, 실업률, 지지율 등과 같이 모집단에서 특정한 속성을 갖는 개체의 비율을 뜻하는 모비율의 추정과 검정에 대하여 알아보자.
- 모비율이 p이고, 크기 n인 랜덤표본에서 그 속성을 갖는 것의 개수를 X라고 하면 $X \sim B(n, p)$.
- 모비율 p의 추정량: 표본비율 $\hat{p} = \frac{X}{n}$

$$E(\hat{p}) = \frac{E(X)}{n} = p, \quad Var(\hat{p}) = \frac{p(1-p)}{n}, \quad SE(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$$

- $\hat{p}$의 표준오차의 추정:

$$\widehat{SE}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

모비율 p의 신뢰구간(n이 클 때)

- 이항분포의 정규분포 근사로부터 표본크기 n이 충분히 크면 X의 분포는 근사적으로 정규분포 $N(np, np(1-p))$에 가까워 지고, 이를 표준화 하면 다음과 같다.

$$X \sim N(np, np(1-p)) \implies \frac{X - np}{\sqrt{np(1-p)}} \dot{\sim} N(0,1)$$

$$\frac{X - np}{\sqrt{np(1-p)}} = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \dot{\sim} N(0,1)$$

- 따라서 표본비율 $\hat{p}$은 근사적으로 다음과 같은 정규분포를 따른다.

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

- 정규분포의 성질에 의하여

$$P\left\{-z_{\alpha/2} \leq \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \leq z_{\alpha/2}\right\} = 1-\alpha$$

이므로 이를 이용하여 모비율 p의 신뢰구간을 구할 수 있다.

- 위의 식을 p에 관하여 정리하면 다음과 같이 쓸 수 있고,

$$p \in \{\hat{p} - z_{\alpha/2}\sqrt{p(1-p)/n},\ \ \hat{p} + z_{\alpha/2}\sqrt{p(1-p)/n}\}$$

- 이 때, 모수인 p의 값을 모르므로 $\hat{p}$으로 대신 사용한다.

모비율 p의 신뢰구간(n이 클 때)

- 모비율 p에 대한 $100(1-\alpha)\%$신뢰구간

$$\left(\hat{p} - z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n},\ \hat{p} + z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\right)$$

- $100(1-\alpha)\%$ 오차범위

$$z_{\alpha/2}\widehat{SE}(\hat{p}) = z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}$$

- 표본크기 n이 크다는 것의 기준

$$n\hat{p} \geq 5, \text{ and } n(1-\hat{p}) \geq 5$$

Example (7.1)

서울에 거주하는 가구 중 돌아오는 추석에 귀향할 가구의 비율을 추정하기 위해 500가구를 단순랜덤추출하여 조사한 결과, 79가구가 귀향할 것이라고 응답했다. 서울에 거주하는 가구의 귀향비율 p에 대한 95% 신뢰구간을 구하라

sol) 귀향률 p의 추정량은 $\hat{p}$이고 추정값은 다음과 같다.

$$\hat{p} = 79/500 = 0.158$$

- 표준오차의 추정값:

$$\widehat{SE}(\hat{p}) = \sqrt{0.158(1-0.158)/500} = 0.016$$

- $n\hat{p} \geq 5,\ n(1-\hat{p}) \geq 5$가 만족되므로 귀향률 p의 95% 신뢰구간은

$$(0.158 - 1.96 \times 0.016, 0.158 + 1.96 \times 0.016) = (0.126, 0.190)$$

표본크기의 결정 - 모비율 p의 추정

- 모비율의 추정에서 $100(1-\alpha)\%$오차범위를 d이하로 하는 것이 요구된다면

$$z_{\alpha/2}SE(\hat{p}) \leq d, \ i.e., \ z_{\alpha/2}\sqrt{p(1-p)}/\sqrt{n} \leq d$$

가 만족되도록 표본크기 n이 정해져야 한다.

(a) 모비율의 사전 추정값이 p^*로주어진 경우

$$n \geq p^*(1-p^*)(z_{\alpha/2}/d)^2$$

(b) 모비율에 대한 사전 정보가 없는 경우

$$n \geq \frac{1}{4}(z_{\alpha/2}/d)^2$$

Example

어느 지역의 구청장 선거에 갑, 을 두 명의 후보가 입후보했을 때, 이들의 지지율을 알아보고자 한다. 아무런 사전 정보가 없는 상태에서 갑 후보의 지지율에 대한 추정의 95% 오차범위를 3%이내로 유지할 때, 필요한 표본의 크기를 구하라.

sol) $z_{0.025} = 1.96$이고, 모비율에 대한 사전 정보가 없는 경우이므로 표본크기 n은 다음을 만족해야 한다.

$$n \geq \frac{1}{4}(1.96/0.03)^2 = 1067.1$$

- 따라서 $n = 1068$명의 표본조사를 하여야 한다.

모비율에 대한 가설검정(표본크기가 큰 경우)

- $\hat{p}$과 가설의 비율 p_0이 얼마나 차이가 있느냐에 따름
- 귀무가설이 $H_0 : p = p_0$인 경우에, 검정통계량은 (87 page)

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \dot{\sim} N(0, 1)$$

이고 이의 관측값을 z_0라고 할 때, 다음이 성립한다.

	유의확률 P	유의수준 α의 기각역
(a) $H_1 : p > p_0$	$P = P(Z \geq z_0)$	$z_0 \geq z_\alpha$
(b) $H_1 : p < p_0$	$P = P(Z \leq z_0)$	$z_0 \leq -z_\alpha$
(c) $H_1 : p \neq p_0$	$P = P(\lvert Z\rvert \geq \lvert z_0\rvert)$	$\lvert z_0\rvert \geq z_{\alpha/2}$

- 표본크기 n이 크다는 것의 기준:$np_0 \geq 5,\ and\ n(1 - p_0) \geq 5$

Example

Example 7.1에서 과거 추석 때의 귀향률이 20%였다고 하자. 표본 조사 결과에 따르면 돌아오는 추석에 서울에 거주하는 가구의 귀향률(500가구 중 79가구)이 과거에 비해 줄어들었다고 할 수 있는지를 유의수준 1%수준에서 검정하여라.

sol) 돌아오는 추석 때의 귀향률을 p라고 하면 가설은 다음과 같다.

$$H_0 : p = 0.2, \quad H_1 : p < 0.2$$

- $np_0 = 100 \geq 5$, $n(1 - p_0) = 400 \geq 5$이므로 표본크기는 충분히 크다고 할 수 있고
- 표본크기 $n = 500$ 중 귀향하려는 가구의 수는 $X = 79$이므로, 검정통계량의 관측값은

$$z_0 = \frac{79/500 - 0.2}{\sqrt{0.2 \times 0.8/500}} = -2.348$$

- 왼쪽 단측검증이므로 유의수준 1%의 기각역은 다음과 같고 관측된 z_0는 이 기각역에 속한다.

$$z_0 \leq -z_{0.01} = -2.327$$

- 따라서, 유의수준 1%에서 귀무가설을 기각할 수 있다.
- 즉, 표본조사 결과에 의하면 귀향률이 과거보다 낮아질 증거가 뚜렷하다고 할 수 있다.

모분산에 관한 추론

- 모집단의 변동성(variability) 또는 퍼짐(spread)의 정도에 관심이 있는 경우, 모분산이 추론의 대상이 됨.

- 예) 볼베어링을 일정한 규격으로 생산하는 공정에서 지름의 변동성이 크다면, 평균적으로 규격에 맞더라도 불량품이 많아지는 문제가 발생한다.

  => 이때 베어링 지름의 변동성을 확인하는 방법으로 모분산에 대한 추론을 할 수 있다.

- 이러한 추론은 모분산 σ^2의 추정량인 다음의 표본분산을 이용하여 할 수 있다. ($E(S^2) = \sigma^2$임)

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)}$$

표본분산의 분포

Theorem (정규모집단에서의 표본분산의 분포)

$X_1, \cdots, X_n$을 정규분포 $N(\mu, \sigma^2)$으로부터의 랜덤표본이라 할 때, 표본분산 $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2/(n-1)$에 대하여 다음이 성립한다.

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

- 가정:
  - 모집단이 정규분포를 따름
  - 단순임의추출을 해야 함
- 위의 표본분산의 분포를 이용하여 σ^2에 관한 추정이나 유의성검증을 할 수 있다.

pf) $X_1, \cdots, X_n$이 정규분포 $N(\mu, \sigma^2)$의 랜덤표본일 때,

$$\frac{X_i - \mu}{\sigma} \sim N(0, 1)$$

이고 서로 독립$(i = 1, \cdots, n)$이므로
카이제곱분포의 정의로부터 다음이 성립한다.

$$\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(n)$$

그런데 다음의 식에서 $S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2/(n-1)$이므로

$$\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 = \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 + n \left(\frac{\bar{X} - \mu}{\sigma} \right)^2$$

정규모집단에서의 표본분산의 분포

다음과 같이 바꿔 쓸 수 있다.

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2=\frac{(n-1)S^2}{\sigma^2}+\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2$$

여기에서 다음의 정의에 의해

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n),\quad\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2\sim\chi^2(1)$$

카이제곱분포의 가법성을 이용하면 다음이 성립함을 예상해 볼 수 있다.

$$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

카이제곱 분포

Definition (카이제곱분포:chi-squared distribution)

확률변수 $Z_1, \cdots, Z_k$이 표준정규분포 $N(0,1)$의 랜덤표본일 때,

$$Z_1^2 + \cdots + Z_k^2$$

의 분포를 자유도(degrees of freedom)가 k인 카이제곱분포라고 한다. 이 때 기호로서

$$Z_1^2 + \cdots + Z_k^2 \sim \chi^2(k)$$

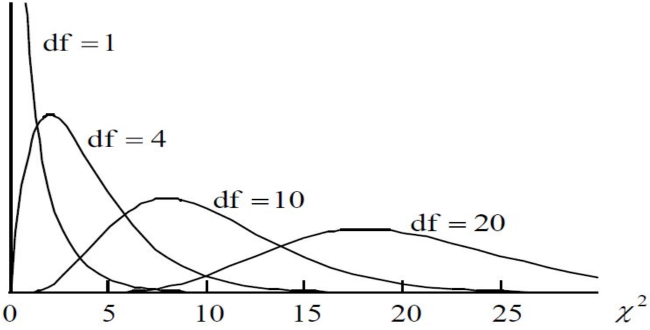
로 나타낸다.

- 카이제곱분포는 다음의 표본분산 S^2에 관계되는 분포이다.

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

카이제곱 분포

- 자유도의 변화에 따른 여러가지 카이제곱분포의 모양은 다음과 같다.



카이제곱 분포의 분위수

- 자유도가 k인 카이제곱분포의 $(1-\alpha)$ 분위수를 흔히 $\chi^2_\alpha(k)$로 나타낸다. 즉, $V \sim \chi^2(k)$일 때 다음과 같다.

$$P\{V \geq \chi^2_\alpha(k)\} = \alpha$$

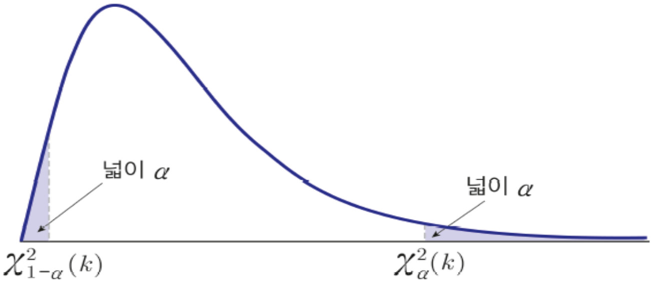
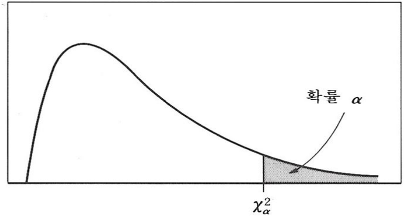


그림 4-7 카이제곱분포의 형태

카이제곱분포표

- 아래 여러가지 자유도에 대한 카이제곱분포표를 이용하여 분위수를 구할 수 있다.

  예) $\chi^2_{0.995}(4) = 0.207$, $\chi^2_{0.975}(4) = 0.484$, $\chi^2_{0.05}(4) = 9.49$



표의 숫자는 주어진 α 값에 대한 χ^2_α의 값을 나타낸다.

df	α .995	.990	.975	.950	.050	.025	.010	.005
1	392704×10^{-10}	$157088\times^{-9}$	$982069\times^{-9}$	$393214\times^{-8}$	3.84146	5.02389	6.63490	7.87944
2	.0100251	.0201007	.0506356	.102587	.99147	7.37776	9.21034	10.5966
3	.0717212	.114832	.215795	.351846	7.81473	9.34840	11.3449	12.8381
4	.206990	.297110	.484419	.710721	9.48773	11.14433	13.2767	14.8602
5	.411740	.554300	.831211	1.145476	11.0705	12.8325	15.0863	16.7496

카이제곱분포의 성질 1

Theorem (카이제곱분포의 가법성)

$V_1 \sim \chi^2(k_1)$, $V_2 \sim \chi^2(k_2)$이고, V_1과 V_2가 서로 독립이면 다음이 성립한다.

$$V_1 + V_2 \sim \chi^2(k_1 + k_2)$$

- 즉, 두 확률변수가 서로 독립이고 각각 카이제곱분포를 따를 때 이들의 합도 카이제곱분포를 따른다.

카이제곱분포의 성질 2

- 자유도가 k인 카이제곱분포를 따르는 확률변수 V의 기댓값과 분산은 각각 다음과 같다.

$$V \sim \chi^2(k) \implies E(V) = k, \;\; Var(V) = 2k$$

- 표본분산 S^2은 모분산 σ^2의 불편추정량이다. $(E(S^2) = \sigma^2)$

pf.1) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ 이므로 $E\left(\frac{(n-1)S^2}{\sigma^2}\right) = n-1$이다.

따라서 $E(S^2) = \sigma^2$.

pf.2) $\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n (X_i - \mu)^2 - n(\bar{X} - \mu)^2$ 이므로

$$\begin{aligned} E((n-1)S^2) &= E(\textstyle\sum_{i=1}^n (X_i - \bar{X})^2) \\ &= E(\textstyle\sum_{i=1}^n (X_i - \mu)^2) - E(n(\bar{X} - \mu)^2) \\ &= n\sigma^2 - n\frac{\sigma^2}{n} = (n-1)\sigma^2 \end{aligned}$$

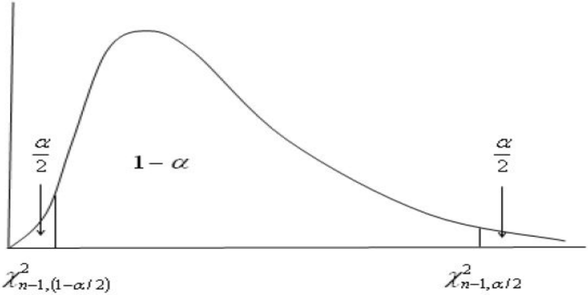
따라서 $E(S^2) = \sigma^2$.

모분산의 신뢰구간

- 표본분산 S^2의 표본분포는 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$이므로
- $\chi^2(n-1)$의 백분위수인 $\chi^2_{1-\alpha/2}(n-1)$, $\chi^2_{\alpha/2}(n-1)$에 대하여

$$1-\alpha = P\{\chi^2_{1-\alpha/2}(n-1) \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{\alpha/2}(n-1)\}$$

Figure: $\chi^2(n-1)$분포의 유의수준 α의 신뢰구간



모분산의 신뢰구간

- σ^2으로 정리하면

$$1-\alpha = P\left\{ \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} \right\}$$

이므로 모분산 σ^2에 대한 $100(1-\alpha)\%$신뢰구간은 다음과 같이 주어진다.(정규모집단의 경우)

$$\left\{ \frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)} \right\}$$

모분산의 가설검정(정규모집단의 경우)

- 귀무가설 $H_0 : \sigma^2 = \sigma_0^2$인 경우에 검정통계량은

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$$

이고, 이의 관측값을 χ_0^2라고 할 때 다음이 성립한다.

	유의확률 P	유의수준 α의 기각역
(a) $H_1 : \sigma^2 > \sigma_0^2$	$P = P(\chi^2 \geq \chi_0^2)$	$\chi_0^2 \geq \chi_\alpha^2(n-1)$
(b) $H_1 : \sigma^2 < \sigma_0^2$	$P = P(\chi^2 \leq \chi_0^2)$	$\chi_0^2 \leq \chi_{1-\alpha}^2(n-1)$
(c) $H_1 : \sigma^2 \neq \sigma_0^2$	$P = 2P(\chi^2 \geq \chi_0^2)$ or	$\chi_0^2 \geq \chi_{\alpha/2}^2(n-1)$ or
	$P = 2P(\chi^2 \leq \chi_0^2)$	$\chi_0^2 \leq \chi_{1-\alpha/2}^2(n-1)$

- 양측가설인 경우 $2P(\chi^2 \geq \chi_0^2)$ or $2P(\chi^2 \leq \chi_0^2)$ 중 1보다 작은 값을 유의확률로 계산

모분산의 가설검정(정규모집단의 경우)

	유의확률 P	유의수준 α의 기각역
(a) $H_1 : \sigma^2 > \sigma_0^2$	$P = P(\chi^2 \geq \chi_0^2)$	$\chi_0^2 \geq \chi_\alpha^2(n-1)$
(b) $H_1 : \sigma^2 < \sigma_0^2$	$P = P(\chi^2 \leq \chi_0^2)$	$\chi_0^2 \leq \chi_{1-\alpha}^2(n-1)$
(c) $H_1 : \sigma^2 \neq \sigma_0^2$	$P = 2P(\chi^2 \geq \chi_0^2)$ or $P = 2P(\chi^2 \leq \chi_0^2)$	$\chi_0^2 \geq \chi_{\alpha/2}^2(n-1)$ or $\chi_0^2 \leq \chi_{1-\alpha/2}^2(n-1)$

Example

한 카페에서는 아메리카노 한 잔에 원두를 10g을 사용하고 있다. 그러나 입력한 중량만큼 원두를 갈아내는 기계는 항상 정확하지 않으며, 약간의 변동이 발생할 수 밖에 없다. 이러한 변동이 너무 크면 커피 맛과 품질이 고객의 기대에 부합하지 않을 수 있으므로 중량의 표준편차 1.5g을 초과하면 기계에 이상이 있다고 간주한다. 어느 날 점검에서 무작위로 기계를 10번 작동하여 측정한 간 원두의 무게가 다음과 같이 주어졌다.

(단위 : g)

11.15	9.62	9.83	10.85	12.66	9.51	11.23	9.38	9.11	8.56

기계에 이상이 있는지를 유의수준 5%에서 검정하고, 원두 무게의 표준편차에 대한 90% 신뢰구간을 구하라. (과거 관리 기록에 따르면 원두 무게의 분포는 정규분포라고 가정할 수 있다.)

sol) 아메리카노 한 잔에 들어가는 원두 중량의 표준편차를 σ라고 하면 검정하려는 가설은 다음과 같다.

$$H_0 : \sigma^2 = 1.5^2 \quad H_1 : \sigma^2 > 1.5^2$$

- 주어진 자료에서 평균과 표준편차를 구하면 각각 다음과 같다.

$$\bar{x} = 10.19, \quad s = 1.24$$

- 이로부터 카이제곱검정통계량의 관측값을 계산하면

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{9(1.24)^2}{1.5^2} = 6.15$$

- p값 유의확률은 다음과 같다.

$$P = P\{\chi^2 \geq 6.15\}, \quad \chi^2 \sim \chi^2(9)$$

- 통계 소프트웨어를 이용하면 실제 유의확률은 0.72이므로 귀무가설을 기각할 수 없다.

- 기각역으로 판정:

  기각역은 $\chi_0^2 \geq \chi_{0.05}^2(9) = 16.92$이고, $\chi_0^2 = 6.15 \leq 16.92$이므로 기각역에 속하지 않아 유의수준 5%에서 귀무가설을 기각할 수 없다.
- 즉, 원두 추출량의 변동에 이상이 있다는 증거가 뚜렷하지 않다.

- σ^2에 대한 90% 신뢰구간은 $n = 10$, $\chi_{0.05}^2(9) = 16.92$, $\chi_{0.95}^2(9) = 3.33$이므로

$$\left(\frac{9(1.24)^2}{16.92} \leq \sigma^2 \leq \frac{9(1.24)^2}{3.33}\right) = (0.818 \leq \sigma^2 \leq 4.156)$$

  따라서 모표준편차 σ에 대한 90% 신뢰구간은 다음과 같다.

$$(\sqrt{0.818}, \sqrt{4.156}) = (0.9,\ 2.04)$$