추론 통계학(inference statistics)은 모집단의 특성치인 모수 θ에 대한 추정치(점 추정치, 구간 추정치)를 구하는 추정(estimation)과 모수 값에 대한 가설의 진위 여부(통계적 유의성)를 알아 보는 가설 검정(hypothesis testing)이 있다.

가설 검정은 과학적 연구와 유사하다. ⑴자연 현상을 관찰하여 ⑵이론을 정립하고 (임의의 모수 값) ⑶관측치를 통하여 이론 진위 여부를 테스트 한다. 이론의 옳고 그름은 표본 관측치에 의해 판단된다.

[예1]기존의 약에 비해 두통에 효과적인 새로운 약을 개발하였다고 하자. [예2]불량률이 5%인 제조 공정을 개선하기 위한 새로운 공정이 제안되었다. [예3]수능성적과 GPA의 관계?

통계학의 역할은? ①이론(이를 통계적 가설이라 한다)을 가설화 한다. ②실험, 관측, 측정 등을 통하여 적절한 데이터 수집한다. ③관측된 데이터(이는 확률표본과 같은 개념)를 이용하여 이론의 옳고 그름을 판단한다. 관측된 데이터로부터 가설에 적절한 통계량을 얻고 이것을 가설을 검정하게 된다. 이를 검정 통계량이라 한다.

9.1 통계적 가설 기본

9.1.1 통계적 가설(statistical hypothesis)

설정한 이론을 통계적 방법으로 테스트 할 수 있도록 한 것으로 모수의 값에 관한 것이다.

[예1]기존의 약에 비해 두통에 효과적인 새로운 약을 개발하였다고 하자. $(\mu_1 - \mu_2)$ 혹은 $(p_1 - p_2)$

[예2]불량률이 5%인 제조 공정을 개선하기 위한 새로운 공정이 제안되었다. p

[예3]수능성적과 GPA의 관계? $Y(GPA) = a + bX(\text{수능성적}) \Rightarrow b = 0$?

우리가 지지하는 가설을 연구가설(research hypothesis), 혹은 대립가설(alternative hypothesis)이라 한다. 대립가설과 반대되는 가설을 귀무가설(null hypothesis)라 한다. 어떤 가설이 지지되느냐는 관측된 표본 데이터에 의존한다. “귀무”(null)의 의미는 아무 것도 없다는 것이다. 귀무가설은 H_0, 대립가설은 H_a 혹은 H_1으로 표현하다.

[예1]귀무가설: $H_0 : \mu_1 - \mu_2 = 0$, 대립가설: $H_a : \mu_1 - \mu_2 \neq 0$

[예2]귀무가설: $H_0 : p = 0.05$, 대립가설: $H_a : p < 0.05$

[예3]귀무가설: $H_0 : b = 0$, 대립가설: $H_0 : b \neq 0$

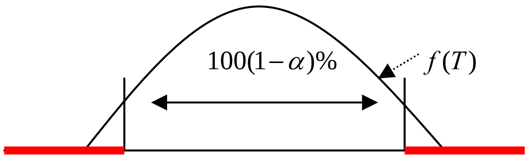
9.1.2 검정 통계량

통계적 가설 검정에 사용되는 통계량을 test statistic(검정 통계량)이라 한다. 통계량이므로 확률 표본 $(X_1, X_2, ..., X_n)$의 함수이다.

[예1]검정 통계량: $T = \dfrac{\bar{x}_1 - \bar{x}_2 - (0)}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ [예2] 검정 통계량: $T = \dfrac{\hat{p} - 0.5}{\sqrt{\dfrac{\hat{p}\hat{q}}{n}}}$

9.1.3 기각역 (RR: rejection region)

귀무가설을 기각하는 검정통계량의 값이 포함된 영역을 기각역이라 한다. 만약 표본으로부터 계산된 검정통계량의 값이 기각역에 포함되면 귀무가설을 기각(대립가설 채택)하고 그렇지 않으면 귀무가설을 채택한다.



빨간 영역 부분이 기각역이다.

9.1.4 1 종 오류, 2 종 오류

귀무가설이 사실일 때 귀무가설을 기각(reject)할 확률을 1종 오류(type I error)라 하고 α 라고 표기한다. 대립가설이 사실일 때 귀무가설을 채택(accept)하는 확률을 2종 오류(type II error)라 하고 β 라 표기한다.

즉 α, β 는 통계적 가설 검정의 GOODNESS을 측정하는데 매우 유용하다.

실제 모집단 / 가설 판단	귀무가설(H_0) 진실	대립가설(H_a) 진실
귀무가설 기각	1종 오류 α	옳은 판단
귀무가설 채택	옳은 판단	2종 오류 β

EXAMPLE 9.1

통계원리 수업을 좋아하는 학생의 비율이 0.5을 넘을까? 이를 알아보기 위하여 수강생 15명을 무작위 추출하여 조사하였더니 10명이 좋아한다고 답변하였다.

①귀무가설: $H_0 : p = 0.5$ (모수 p는 모집단(학생 전체) 통계원리 좋아하는 학생 비율)

대립가설: $H_a : p \neq 0.5$

②검정통계량: 15명 중 좋아하는 학생의 수를 검정 통계량 T라 하자. 적절하다. 게다가 우리는 T가 이항분포를 따르는지 알고 있다.

③기각역: 만약 기각역을 $RR = \{T \geq 13\}$으로 할 경우 1종 오류를 계산하시오.

$$\begin{aligned}\alpha &= P(\text{귀무가설기각} \mid \text{귀무가설사실}) \\ &= P(T \geq 13 \mid T \sim Binomial(15, p = 0.5)) \\ &= \binom{15}{13} 0.5^{13} (1-0.5)^2 + \binom{15}{14} 0.5^{14} (1-0.5)^1 + \binom{15}{15} 0.5^{15} (1-0.5)^0 = 0.004\end{aligned}$$

```
data one;
    alpha=1-CDF('Binomial',12,0.5,15);
run;
proc print data=one;
run;
```

Obs	alpha
1	.003692627

위의 상황에서 귀무가설을 잘못 기각할 확률이 0.004매우 낮다. 그러므로 우리는 귀무가설을 기각하는데 주저할 필요는 없다.

“만약 수업을 좋아하는 학생의 비율이 0.7이 넘을까”였다면? 1종 오류 $\alpha = 0.127$이므로 귀무가설을 기각하게 되면 오류가 다소 커지므로 기각하는데 부담을 갖는다.

만약 모수가 0.8이라고 한다면 2종 오류를 계산하시오.

$$\begin{aligned}\beta &= P(\text{귀무가설채택} \mid \text{대립가설사실}) \\ &= P(T < 13 \mid T \sim Binomial(15, p = 0.8)) = 0.601\end{aligned}$$

```
data one;
    beta=CDF('Binomial',12,0.8,15);
run;
```

beta
0.60198

만약 모수가 0.9이라고 한다면 2종 오류를 계산하시오.

$$\beta = P(\text{귀무가설채택} \mid \text{대립가설사실})$$
$$= P(T < 13 \mid T \sim Binomial(15, p = 0.9)) = 0.184$$

beta
0.18406

귀무가설에 설정된 모수 값에서 실제 모수 값이 멀어질수록 2종 오류는 낮아진다. 당연하다. 2종 오류는 대립가설이 사실인데 귀무가설이 채택하는 것이므로 실제 모수 값이 멀어질수록 오류 가능성은 줄어들게 된다.

만약 기각역을 $RR = \{T \geq 10\}$ 으로 한다면... 그리고 실제 모수가 0.9 라면

1종 오류: $\alpha = P(\text{귀무가설기각} \mid \text{귀무가설사실}) = P(T \geq 10 \mid T \sim Binomial(15, p = 0.5)) = 0.151$ >0.004

2종 오류: $\beta = P(\text{귀무가설채택} \mid \text{대립가설사실}) = P(T < 10 \mid T \sim Binomial(15, p = 0.9)) = 0.06$ <0.184

```
data one;
    alpha=1-CDF('Binomial',9,0.5,15);
    beta=CDF('Binomial',9,0.8,15);
run;
```

alpha	beta
0.15088	0.061051

기각역을 조정한다고 1종 오류, 2종 오류(α, β)를 동시에 줄일 수 있는 것이 아니라 trade-off 관계가 성립한다. 물론 표본의 크기를 크게 하면 α, β 을 동시에 줄일 수 있지만... 표본의 크기가 정해진 경우 기각역의 조정만으로는 불가능하다.

그러므로 1종 오류를 정해 놓고(이를 유의수준, significant level) 2종 오류를 최소화 하는 기각역을 찾게 된다.

HOMEWORK #18-1

학생들의 TOEIC 성적이 600점 이상인지 알아보기 위하여 25명을 임의 추출하였다. 표본 분산은 36이었다. ①귀무가설과 대립가설을 정하시오. ②기각역을 $RR = \{\bar{X} \geq 650\}$ 이고 실제 모집단 평균이 750인 경우 1종 오류와 2종 오류를 계산하시오. ③기각역을 $RR = \{\bar{X} \geq 700\}$ 로 할 경우 1종 오류와 2종 오류를 계산하시오.

9.2 대표본 검정

표본의 크기 n인 확률표본 $(X_1, X_2, ..., X_n)$을 이용하여 모수 θ에 관한 가설 검정을 실시한다고 하자. 우리는 앞에서(3.6절, 페이지63) 다음 사실을 알고 있다.

$$\frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}} \sim (approximate) Normal(0,1) \quad \text{--- (1)}$$

이 사실을 이용할 수 있는 모수는 $\mu, p, (\mu_1-\mu_2), (p_1-p_2)$임을 알고 있다.

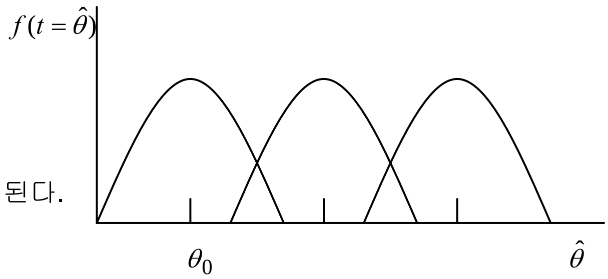
다음 통계적 가설을 검정한다고 하자. 통계적 가설의 유의성 검정

①귀무가설: $H_0 : \theta = \theta_0$

②대립가설: $H_a : \theta > \theta_0$

③검정통계량: $\hat{\theta} = T$

④기각역: $RR = \{\hat{\theta} > k\}$, k 값을 정하게 된다.



k 값은 1종 오류 α을 어떤 값을 설정하느냐에 따라 결정된다. 9.1절에서 살펴 보았듯이 기각역을 조정(즉 k값을 조정)하여도 α, β을 동시에 줄일 수는 없었다. 그러므로 하나를 고정하게 된다. 그럼 어떤 것을 고정할까? 우리의 관심은 대립가설에 있으므로 대립가설이 사실일 때 귀무가설을 채택하는 2종 오류 β을 최소화 하는 검정 방법을 찾는 것이 바람직하다.

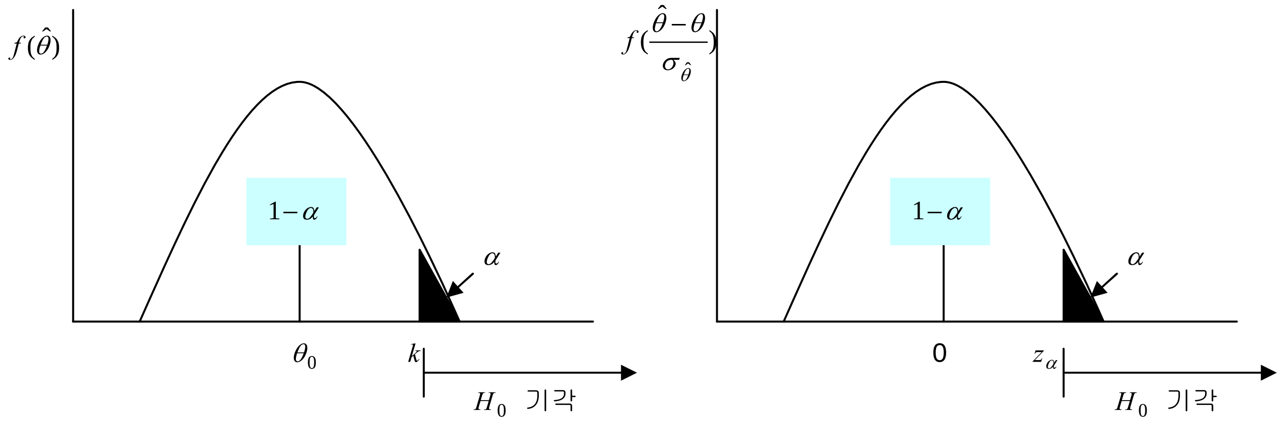
특별히 기각역의 시작점이 되는 값을 기각치(critical value) 혹은 임계치라 한다.

가설 검정 시 고정된 1종 오류 값을 유의 수준(significant level) 혹은 검정 수준(test level)이라 한다. 일반적으로 5%, 1%, 10%가 주로 사용된다. $\alpha = 0.05, \alpha = 0.01, \alpha = 0.1$.

유의수준이 α가 정해지면 기각역의 k는 다음과 같이 정해진다. 만약 귀무가설 H_0가 사실이라면 $\hat{\theta}$는 평균 θ, 표준편차 $\sigma_{\hat{\theta}}$인 정규분포를 따른다. 그러므로 유의수준을 설정하면 다음 식에 의해 결정된다.

$$k = \theta_0 + z_\alpha \sigma_{\hat{\theta}}, \quad P(Z > z_\alpha \mid Z \sim Normal(0,1))$$

위의 $k = \theta_0 + z_\alpha \sigma_{\hat{\theta}}$ 결정은 $RR = \{\hat{\theta} : \hat{\theta} > \theta_0 + z_\alpha \sigma_{\hat{\theta}}\} = \{\hat{\theta} : \frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}} > z_\alpha\}$과 동일하다.



왼쪽 방법에 의해 기각역을 설정하려면 $\hat{\theta}$의 분포(정규분포, 그러나 표준 정규분포 표만 있으므로)를 이용하는 것보다 오른쪽 방법에 의해 $\dfrac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}}$을 검정 통계량으로 사용하면 표준 정규분포표로부터(이미 우리는 5%, 10%, 1% 값은 알고 있다) z_α를 쉽게 구할 수 있다.

그러므로 $\mu, p, (\mu_1-\mu_2), (p_1-p_2)$에 대한 대표본 검정 방법은 다음과 같다.

①귀무가설: $H_0 : \theta = \theta_0$ ②대립가설: $H_a : \theta > \theta_0$

③검정통계량: $T(Z) = \dfrac{\hat{\theta}-\theta_0}{\sigma_{\hat{\theta}}}$

④기각역: $RR = \{T = \dfrac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}} > z_\alpha\}$. z_α는 표준 정규분포표로부터 얻는다. ---(1)

위와 같이 기각역이 오른편에 한정되어 있으면 이를 우측 기각역(upper tail RR)이라 한다. 유의수준 $\alpha = 5\%$ (10%/1%)이면 기각역은 $RR = \{\dfrac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}} > 1.645\}$ (1.285 / 2.33)

검정통계량은 $Z = \dfrac{\hat{\theta}-\text{귀무가설에설정된모수}}{\sigma_{\hat{\theta}}}$으로 계산된다. 기각역은 식 (1)에 의해 결정된다.

EXAMPLE 9.2

자동차 세일즈 관리 책임자는 상부에 그들의 세일즈맨은 일주일 평균 15명 이상의 고객을 만난다고 보고하였다. 상부에서 이를 알아보기 위하여 세일즈맨 36명을 무작위 추출하여 지난 일주일 만난 고객의 수를 조사하였더니 평균 17명이었고 분산은 9명이었다. 관리 책임자의 보고가 맞는가? (유의수준은 5%로 하시오)

①귀무가설: H_0 :　　　　　　　　②대립가설: H_a :

③검정통계량: $Z=4$　　　　　　　　④기각역: $RR=\{Z>z_\alpha=1.645\}$.

귀무가설이 기각되므로 유의수준 5%(이를 다시 말하면 신뢰수준 95%) 하에서 관리 책임자의 주장이 옳다고 할 수 있다.

EXAMPLE 9.3

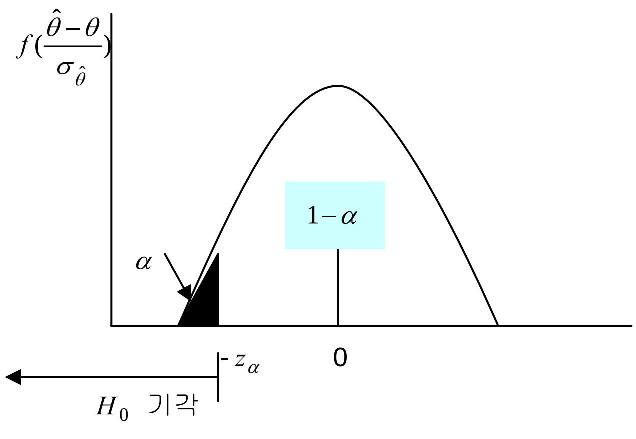
하루에 생산된 제품의 불량률이 10%보다 크면 그날의 모든 제품을 폐기한다고 한다. 오늘 생산된 제품 100개를 조사하였더니 불량의 개수가 15개였다. 그럼 오늘 제품을 모두 폐기해야 하는가? 유의수준은 0.01로 하시오.

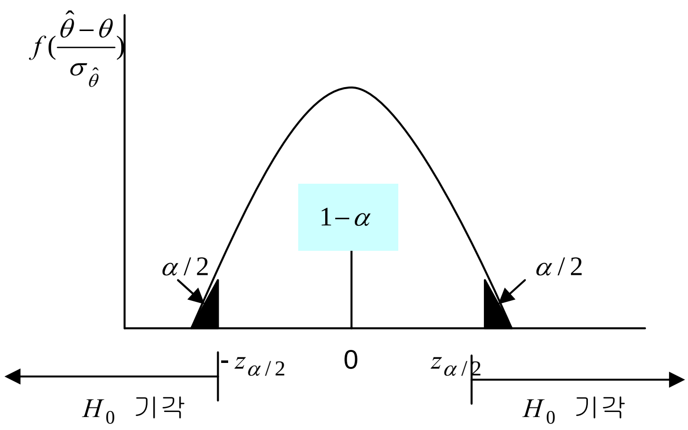
①귀무가설: H_0 :　　　　　　　　②대립가설: H_a :

③검정통계량: $Z=1.667$　　　　　　④기각역: $RR=\{Z>z_\alpha=2.33\}$.

귀무가설이 채택되므로 유의수준 1%(이를 다시 말하면 신뢰수준 99%) 하에서 제품을 전량 폐기할 필요는 없다.

대립가설이 모수의 왼쪽 부분에 대한 것이면($H_a:\theta<\theta_0$) 좌측(lower tail) 되고(왼쪽 그림) 대립가설이 모수의 양쪽 부분이면($H_a:\theta\neq\theta_0$) 양측(two-tailed) 기각역을 설정하면 된다.(오른쪽 그림)





SUMMARIZATION (대표본 유의수준 α 가설 검정)

①귀무가설(null hypothesis): $H_0 : \theta = \theta_0$

②대립가설(alternative hypothesis): $\begin{aligned} H_a : &\theta > \theta_0 (upper-tail) \\ &\theta < \theta_0 (lower-tail) \\ &\theta \neq \theta_0 (two-tail) \end{aligned}$

③검정통계량(test statistic): $Z = \dfrac{\hat{\theta} - \theta}{\sigma_{\hat{\theta}}}$

④기각역(rejection region): $\begin{aligned} RR &= \{Z > z_{1-\alpha}\}(upper-tail) \\ &= \{Z < z_\alpha\}(lower-tail) \\ &= \{|Z| > z_{\alpha/2}\}(two-tail) \end{aligned}$.

대립가설이 모수의 한 쪽만 설정해 놓고 있으면 이를 단측(one-tail) 검정 이라 한다.

그럼 어떤 대립가설을 선택할 것인가? 우리가 주장하고픈 주장을 뒷받침할 수 있도록 가설을 구성해야 한다. 동일 유의 수준이라면 양측 검정(two-tail)의 임계치(기각역)는 단측 검정의 임계치부보다 가운데 0(귀무가설에 설정된 모수 값)으로부터 멀게 되므로 양측 검정에서 귀무가설이 기각되면 단측검정에서도 기각된다. 예를 들면 유의수준 α 에서 양측 검정의 귀무가설 $H_0 : \theta = \theta_0$ (대립가설: $H_0 : \theta \neq \theta_0$)이 기각되면 단측 검정의 귀무가설 $H_0 : \theta = \theta_0$ (대립가설: $H_0 : \theta > \theta_0$ 혹은 $H_0 : \theta < \theta_0$)도 기각된다.

다음 절에서는 2종 오류(β) 계산하는 방법, 표본 크기 결정 방법, 그리고 신뢰구간과 가설검정의 일대일 관계(추정량, 추정 분산, 표준정규분포표 이용)를 논의할 것이다.

EXAMPLE 9.4

심리학 연구에서 자극에 대한 남녀별 반응시간의 차이가 있는지 알아보고자 남녀 각각 50명을 조사하여 다음 결과를 얻었다. 유의수준 5%로 하시오.

$\bar{x}_1 = 3.6, \bar{x}_2 = 3.8$, $s_1^2 = 0.18, s_2^2 = 0.14$

①귀무가설: H_0 :

②대립가설: H_a :

③검정통계량: $Z = -2.5$ 계산식 $Z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$ 이용

④기각역: $RR = \{|Z| > z_\alpha = 1.96\}$.

귀무가설이 기각되므로 유의수준 5% 하에서 귀무가설이 기각되므로 남녀별 반응 시간의 차이는 있다고 할 수 있다. 그럼 남자의 반응 시간(3.6초)이 여자의 반응시간(3.8초)에 비해 빠르다고 할 수 있나? 다음 예제를 보자.

EXAMPLE 9.5

위의 예제에서 남자의 반응시간이 여자에 비해 빠른가? 유의수준 5%에서 검정하시오.

①귀무가설: H_0 : ②대립가설: H_a :

③검정통계량: $Z = ???$ 계산식 ④기각역: $RR = \{Z > z_\alpha = 1.645\}$.

귀무가설이 기각된다. 그러므로 남자의 반응시간이 빠르다. 당연한가? 양측검정에서 귀무가설이 기각되었으므로 같은 유의수준이라면 당연히 귀무가설이 기각된다.

HOMEWORK #18-2

정부는 PC 구입 가격이 900$이라고 발표했다. 이를 알아보기 위하여 소비자 35명을 임의 조사하였더니 평균 885$, 표준편차는 50$이었다. 정부의 발표가 진실이라고 할 수 있나? 유의수준은 1%로 하시오.

HOMEWORK #18-3

냉장고 색은 파랑, 하양, 검정이 있다. 1,000개가 팔렸는데 그 중 파랑이 400개 팔렸다. 그럼 소비자들이 이 파란색 냉장고를 더 선호한다고 할 수 있나? 유의수준 5%.

HOMEWORK #18-4

○○법안 찬성 비율이 민주당에 비해 공화당이 더 높은가 알아보고자 민주당 200명을 조사하였더니 34명, 공화당 200명을 조사하였더니 46명이 찬성한다고 대답하였다. 유의수준 5%로 검정하시오.

HOMEWORK #18-5

두 지역 땅의 질소 함유량의 차이가 있는지 알아보기 위하여 A 지역에서 1 cm^3 만큼 30개, B 지역에서 35개 임의 추출하여 질소량을 조사하여 다음 결과를 얻었다. 차이가 있는지 유의수준 1%에서 검정하시오. 여러분의 최종 결론은?

$n_1 = 30, n_2 = 35$, $\bar{x}_1 = 1.65, \bar{x}_2 = 1.43$, $s_1 = 0.26, s_2 = 0.22$

HOMEWORK #18-6

철강 생산 시 식히는 과정에서 소금 물을 사용하는 방법과 오일을 사용하는 방법 중 어느 것이 강도를 높이는지 알아보기 위하여 다음과 같이 측정 자료를 얻었다. 강도는 **정규 분포**를 따른다고 하자.

소금물:	145	150	153	148	141	152	146	154	139	148
오일:	152	150	147	155	140	146	158	152	151	143

①물에 의해 식히면 철강의 강도가 145라고 한다. 오일에 의해 식히는 경우 철강의 강도가 145보다 크다고 할 수 있나? 유의수준 5% (일변량 분석)

②두 방법의 철강 강도의 차이가 있는지 유의수준 1%에서 알아보시오.

다음은 소금물에 의해 식히는 경우 철강의 강도가 145보다 큰지 유의수준 5%에서 알아볼 때 적절한 프로그램이다.

```
data one;
    input x @@;
    cards;
    145    150    153    148    141    152    146    154    139    148
run;

proc ttest data=one h0=150 alpha=0.1;
run;
```

만약 두 집단 평균 비교의 경우에는

CLASS 변수를 사용하면 된다.

Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev	Upper CL Std Dev
144.72	147.6	150.48	3.6256	4.971	8.1783

T-Tests

Variable	DF	t Value	Pr > |t|
x	9	-1.53	0.1612

9.3 2종 오류 계산

일반적으로 2종 오류를 계산하기 어려우나 앞 절의 대표본 가설 검정의 경우 다소 용이하다. 귀무가설: $H_0:\theta=\theta_0$, 대립가설: $H_a:\theta>\theta_0$ 의 경우 2종 오류 β 을 계산하는 방법을 살펴 보자. 2종 오류를 계산하려면 대립가설에 속한 모수의 하나의 값이 필요하다. 대립가설의 모수 값은 구간이기 때문이다. 대립가설의 하나의 값을 θ_a 라 하자. 앞 절에서 살펴본 것처럼 기각역은 $RR=\{\hat\theta:\hat\theta>k\}$ 이다. 이제 2종 오류를 생각해 보자.

$$\begin{aligned}\beta &= P(\hat\theta\ not\ in\ RR \mid \text{대립가설 } H_a \text{ 진실})\\ &= P(\hat\theta:\hat\theta\le k \mid \text{대립가설 } H_a:\theta=\theta_a \text{ 진실})\\ &= P(\hat\theta:\frac{\hat\theta-\theta_a}{\sigma_{\hat\theta}}\le\frac{k-\theta_a}{\sigma_{\hat\theta}} \mid \text{대립가설 } H_a:\theta=\theta_a \text{ 진실})\end{aligned}$$

대립가설의 모수가 $\theta_a>\theta_0$ 라면 $\frac{\hat\theta-\theta_a}{\sigma_{\hat\theta}}$ 는 $Normal(0,1)$ 이다. 만약 θ_a 가 θ_0 에 가깝다면 대립가설이 사실임에도 불구하고 귀무가설을 받아들일 가능성이 높다, 즉 2종 오류 β 가 커진다.

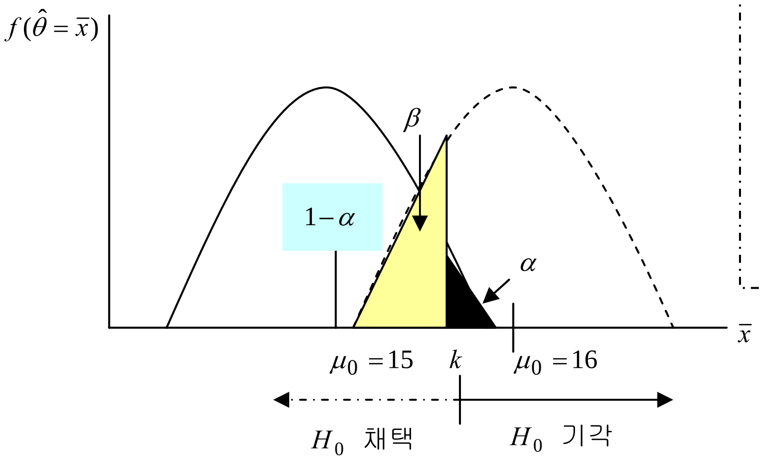
EXAMPLE 9.6

EXAMPLE 9.2 자동차 세일즈맨 고객을 만난 회수. 표본 평균 17명이었고 분산은 9명이었다. 그리고 귀무가설 $H_0:\mu=\mu_0(15)$, 대립가설 $H_0:\mu>15$ 이다. 대립가설의 모수 값 하나를 16이라 하자. 이 경우 2종 오류를 계산하시오. 유의수준은 5%로 하시오,

---

유의수준이 결정되면 기각역을 얻을 수 있다. $RR=\{\hat\theta:\frac{\hat\theta=\bar x-\theta_0}{\sigma_{\hat\theta}}>z_\alpha=\frac{k-\theta_0}{\sigma_{\hat\theta}}\}$ 으로부터 k 을 결정하면 $\frac{k-15}{3/\sqrt{36}}=1.645$ 에서 $k=15.8225$ 이다.

$$\begin{aligned}\beta &= P(\frac{\hat\theta=\bar x-\mu_a}{\sigma/\sqrt{n}}\le\frac{k-\mu_a}{\sigma/\sqrt{n}} \mid H_a)\\ &= P(Z\le\frac{k=15.8225-16}{3/\sqrt{36}} \mid H_a)\\ &= P(Z\le -0.36 \mid H_a)\\ &= 0.3594\end{aligned}$$



2종 오류는 0.3594이다. 즉 실제 모수가 16임에도 불구하고 귀무가설을 선택할 확률은 35.9%이다.

만약 대립가설의 모수가 18이라면 $\beta = P(Z \le \frac{15.8225-17}{3/\sqrt{36}} \mid H_a) = P(Z \le -2.355 \mid H_a) = 0.009$ 이다.

즉 귀무가설에 설정된 모수 값과 멀어질수록 2종 오류는 낮아진다.

표본의 크기가 커지면 1 종 오류 2 종 오류는 줄어든다. why? 그럼 표본의 크기를 얼마로 해야 하나?

귀무가설: $H_0 : \theta = \theta_0$, 대립가설: $H_a : \theta > \theta_0$ 이고 1종 오류 α, 2종 오류 β 으로 설정하자.

$$\alpha = P(\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} > \frac{k-\mu_0}{\sigma/\sqrt{n}} \mid H_0) = P(Z > z_\alpha) \Rightarrow \frac{k-\mu_0}{\sigma/\sqrt{n}} = z_\alpha$$

$$\beta = P\frac{\bar{x}-\mu_a}{\sigma/\sqrt{n}} \le \frac{k-\mu_a}{\sigma/\sqrt{n}} \mid H_a) = P(Z \le -z_\beta) \Rightarrow \frac{k-\mu_a}{\sigma/\sqrt{n}} = -z_\beta \quad \text{---(1)}$$

식(1)의 연립 방정식을 풀면 $n = (\frac{(z_\alpha + z_\beta)\sigma}{\mu_a - \mu_0})^2$ 이다. 즉 대립가설의 모수와 귀무가설의 모수 차이가 적으면 표본의 크기를 증가시켜야 한다. 그리고 1종 오류, 2종 오류를 줄이려면 표본의 크기를 크게 해야 한다.

EXAMPLE 9.7

EXAMPLE 9.6. 귀무가설 $H_0 : \mu = \mu_0(15)$, 대립가설 $H_0 : \mu_a = 16$ 이었다. 대립가설의 모수 값 하나를 16이라 하자. $\alpha = \beta = 0.05$ 으로 하려면 표본의 크기는 얼마로 해야 하나?

$$n = (\frac{(10645+1.645)3}{16-15})^2 = 97.4$$

HOMEWORK #19-1

학생들의 IQ가 130이라 한다. 이를 알아보기 위하여 40명의 학생들을 무작위 추출하여 조사하였더니 평균은 128.6이고 표준편차는 2.1이었다. 학생들의 IQ가 130미만인지 알아보고자 한다.

①유의수준 5%에서 가설 검정하시오.

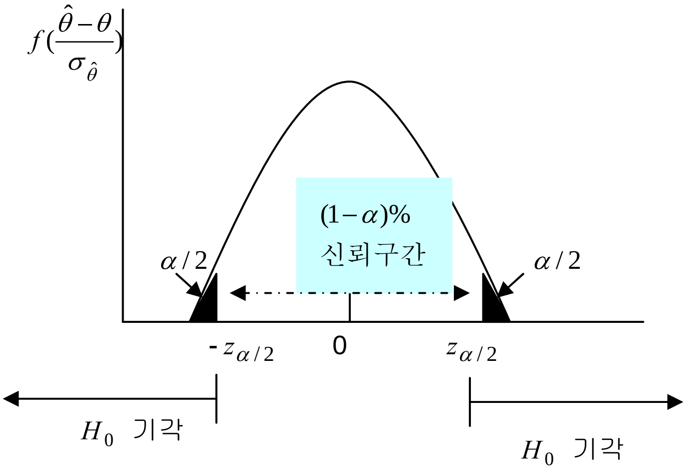
②대립가설의 모수가 $\mu_a = 128$ 인 경우 2종 오류를 계산하시오.

HOMEWORK #19-2

첫사랑을 경험한 학생의 비율이 많아야 20%는 된다고 주장한 연구가 있다. 이 연구에서 대립가설이 $p_a = 0.15$ 인 경우 2종 오류를 계산하시오. 그리고 $\alpha = \beta = 0.01$ 일 경우 표본의 크기를 얼마로 해야 하나?

9.4 가설검정과 신뢰수준

유의수준 α 의 양측 가설 검정과 $100(1-\alpha)\%$ 신뢰구간은 일대일 대응관계가 있다. 대표본의 경우 $\dfrac{\hat{\theta}-\theta_a}{\sigma_{\hat{\theta}}}$ 는 $Normal(0,1)$ 을 따르므로 다음이 성립한다.



$100(1-\alpha)\%$ 신뢰구간

$$\hat{\theta} \pm z_{\alpha/2}\sigma_{\hat{\theta}}$$

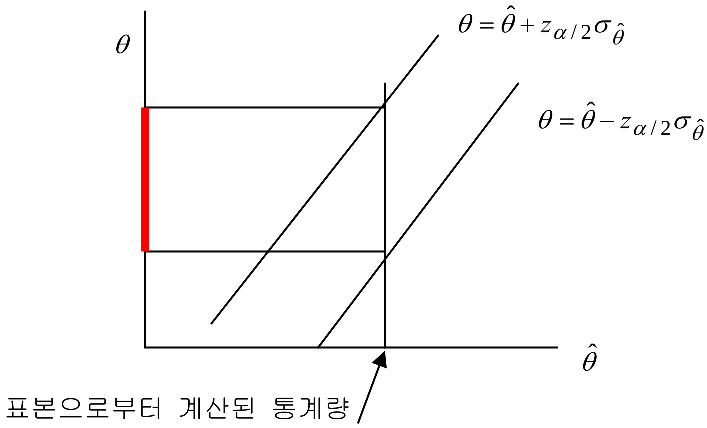
유의수준 α 인 경우 기각역은

$RR = \{|\dfrac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}}| > z_{\alpha/2}\}$ 이므로

채택역(acceptance region)은

$\overline{RR} = \{-z_{\alpha/2} \le \dfrac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}} \le z_{\alpha/2}\}$ 이다.

즉 신뢰구간과 채택역은 동일함을 알 수 있다. 신뢰구간에 속한 모수 값이 귀무가설에 설정되면 그 귀무가설은 기각되지 않는다.



표본으로부터 통계량이 계산되면 신뢰구간의 식에의해 모수에 대한 구간이 결정된다. 그 부분이 빨간 영역 부분이다.

귀무가설에 빨간 영역 안에 있는 모수 값이 설정되면 귀무가설을 기각하지 못한다.

단측(one-tail) 가설 검정과 상한(혹은 하한) 신뢰구간도 일대일 관계가 있다. $100(1-\alpha)\%$ 하한 신뢰구간(lower confidence bound)은 유의수준 α, $H_0 : \theta = \theta_0$, 대립가설: $H_a : \theta > \theta_0$ 의 채택역과 동일하다. 또한 $100(1-\alpha)\%$ 상한신뢰구간(upper CI)은 유의수준 α, $H_0 : \theta = \theta_0$, 대립가설: $H_a : \theta < \theta_0$ 의 채택역과 동일하다.

EXAMPLE 9.8

HOMEWORK #18-2 정부는 PC 구입 가격이 900$이라고 발표했다. 이를 알아보기 위하여 소비자 35명을 임의 조사하였더니 평균 885$, 표준편차는 50$이었다.

①모집단 평균에 대한 99%(양측)신뢰구간을 구하시오.

②정부의 발표가 진실이라고 할 수 있나? 유의수준은 1%로 하시오. ①의 결과가 있다면 이 가설 검정을 할 필요가 있나? why?

99% 신뢰구간 $\hat{\theta}(=\bar{x}) \pm z_{\alpha/2}\sigma_{\hat{\theta}} \Rightarrow (885 - 2.58 * \frac{50}{\sqrt{35}}, 885 + 2.58 * \frac{50}{\sqrt{35}}) \Rightarrow (863.1, 906.8)$

유의수준 1%에서 귀무가설의 값이 신뢰구간에 속하므로 귀무가설을 채택하게 된다.

만약 구입가격을 900$ 미만이라 발표하였다면... 95% 상한신뢰구간을 구하고 유의수준 5% 가설검정 결과와 일치하는지 보이시오. Now!

HOMEWORK #19-3

HOMEWORK# 18-5($n_1 = 30, n_2 = 35$, $\bar{x}_1 = 1.65, \bar{x}_2 = 1.43$, $s_1 = 0.26, s_2 = 0.22$) $(\mu_1 - \mu_2)$ 의 99%신뢰구간을 구하고 유의수준 1% 가설검정 결과와 일치하는지 보이시오.

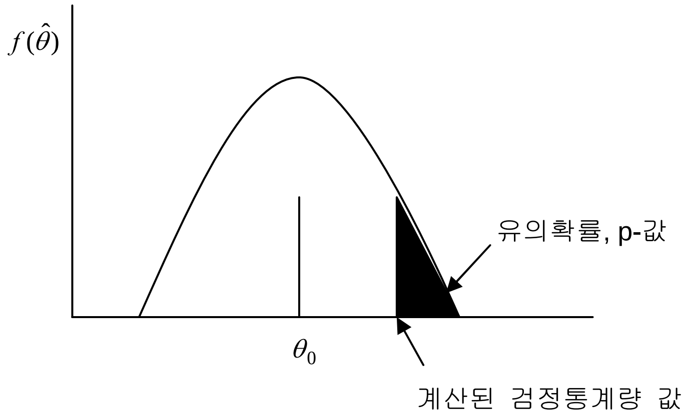
9.5 p-value

1종 오류 α, 2종 오류 β 계산하는 방법에 대해 다루었다. 가설검정 절차를 보면 ①가설 설정과 α 을 설정하고(이를 유의수준이라 한다) ②적절한 검정통계량 계산 ③기각역(유의수준과 검정통계량의 분포에 따라 계산된다)을 얻고 이에 따라 귀무가설 채택 여부를 결정한다.

설정된 1종 오류를 유의수준(significant level, test level)이라 하고 $(1-\beta)$ 을 검정력(power of test, test power)이라 한다.

유의수준은 일반적으로 1%, 5%, 10% (물론 5%을 가장 많이 사용하지만)을 사용하게 되므로 기각역이 달라진다. 만약 어떤 사람이 유의수준 5%에서 귀무가설을 기각하지 못했다고 발표했다. 그럼? 10%일 때는… 우리는 다시 가설 검정을 해야 한다. why? 기각역이 달라지므로…

이런 문제를 해결할 방법이 없을까? p-값(p-value, significant probability 유의확률) 개념을 도입하자. 검정통계량이 계산되면 귀무가설을 더 기각할 영역의 확률을 p-값이라 한다. 다음은 대립가설이 $H_a : \theta > \theta_0$ 일 경우 p-값(유의확률) 계산 방법이다.



그럼 대립가설이

$H_a : \theta < \theta_0$ 일 경우는?

대립가설이

$H_a : \theta \neq \theta_0$ 일 경우는?

DEFINITION

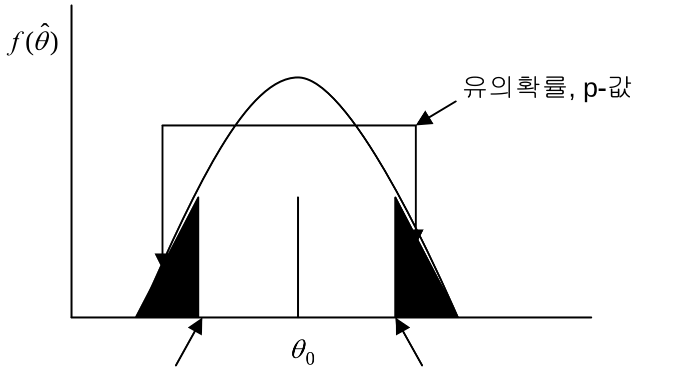
T 을 검정통계량이라 하자. 계산된 검정통계량이 귀무가설을 기각할 최소의 유의수준을 유의확률 혹은 p-값이라 한다.

계산된 검정통계량이 귀무가설을 기각할 방향(대립가설을 채택할 방향)의 영역 확률을 유의확률이라 한다.

유의확률이 주어지면 귀무가설 채택 여부를 알 수 있다. 유의확률이 유의수준보다 크다면 검정통계량의 값은 기각역에 속한 것이 아니므로 귀무가설이 채택되고 유의확률이 유의수준보다

적다면 검정통계량의 값이 기각역에 속하므로 귀무가설을 기각한다. (예) 유의확률이 0.07이다. 유의수준 5%이면 귀무가설이 채택이나 유의수준 10%이면 귀무가설이 기각된다. 이처럼 유의확률은 우리에게 보다 많은 정보를 제공하므로 유의확률을 제공하는 것이 옳은 방법이다.

통계소프트웨어는 항상 양측검정일 경우 유의확률을 계산하여 제공한다. 그러므로 단측검정을 하려면 제공된 유의확률을 1/2로 나누면 된다. 다음은 양측검정 $H_a: \theta \neq \theta_0$ 일 경우 유의확률을 계산하는 방법이다.



검정통계량 값(T^*)이 주어질 경우 유의확률(p-값) 계산 방법

⑴귀무가설 $H_0: \theta = \theta_0$ vs. 대립가설 $H_a: \theta > \theta_0$ $p(T(\hat{\theta}) \geq T^* \mid H_0$사실)

⑵귀무가설 $H_0: \theta = \theta_0$ vs. 대립가설 $H_a: \theta < \theta_0$ $p(T(\hat{\theta}) \leq T^* \mid H_0$사실)

⑶귀무가설 $H_0: \theta = \theta_0$ vs. 대립가설 $H_a: \theta \neq \theta_0$ $p(|T(\hat{\theta})| \geq$ 양의값$T^* \mid H_0$사실)

EXAMPLE 9.9

○○법안 찬성하는 학생 비율이 0.5미만인지 알아보기 위하여($H_0: p = 0.5$, $H_a: p < 0.5$) 15명을 임의 추출하여 조사하였더니 3명이 찬성이라고 했다. 유의확률(p-값)을 계산하시오.

$$
\begin{aligned}
p &= P(X \leq 3 \mid X \sim Binomial(n = 15, p = 0.5)) \\
&= \binom{15}{3} 0.5^3 (1-0.5)^{12} + \binom{15}{2} 0.5^2 (1-0.5)^{13} + \binom{15}{1} 0.5^1 (1-0.5)^{14} + \binom{15}{0} 0.5^0 (1-0.5)^{15} \\
&= 0.018
\end{aligned}
$$

유의확률이 0.018이다. 유의수준을 0.05로 놓는다면 귀무가설은 기각된다. 사실 수작업으로 가설 검정할 경우에는 유의확률 계산이 쉽지는 않다.

EXAMPLE 9.10

심리학 연구에서 자극에 대한 남녀별 반응시간의 차이가 있는지 알아보고자 남녀 각각 50명을 조사하여 다음 결과를 얻었다. $\bar{x}_1 = 3.6, \bar{x}_2 = 3.8$, $s_1^2 = 0.18, s_2^2 = 0.14$ (EXAMPLE 9.4)

①귀무가설: $H_0 : \mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$ ②대립가설: $H_a : \mu_1 \neq \mu_2$

검정통계량 $T = \dfrac{(3.6-3.8)-(\mu_1-\mu_2)=0}{\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\dfrac{0.18}{50}+\dfrac{0.14}{50}}} = -2.5 \sim Normal(0,1)$

$p = P(|T| \geq 2.5 \mid X \sim Normal(0,1)) = 0.0062 * 2 = 0.0124$

유의수준을 0.05로 하면 귀무가설이 기각되고 0.01로 하면 채택된다. 만약 대립가설을 "남자가 여자에 비해 반응 시간이 빠르다"($H_a : \mu_1 - \mu_2 < 0 \Leftrightarrow \mu_1 < \mu_2$)라고 한다면??? 유의확률은 0.0062이고 유의수준 0.01에서도 귀무가설이 기각된다.

소표본일 경우에는 검정통계량 계산시 $\sigma_{\bar{x}_1-\bar{x}_2} = s_p \sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}$, $s_p = \sqrt{\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$ 단, 모집단 정규분포라는 가정이 필요하다.

HOMEWORK #20-1

CEO들이 일반 사람에 비해 오른손잡이가 많다고 주장하였다. 일반 사람들 중 오른손잡이는 85%이다. CEO 중 300명을 임의 추출하여 조사하였더니 96%가 오른손잡이였다.

⑴유의수준 5%에서 가설검정 절차에 의해(가설 설정=>검정통계량 계산=>기각역 설정=>판단) 그들의 주장의 유의성을 판단하시오.

⑵유의확률을 계산하고 이를 이용하여 주장의 유의성을 검정하시오.

9.6 가설검정에 대한 고찰

⑴양측검정이냐? 단측검정이냐? 우리가 무엇에 관심을 갖느냐 하는데 있다. 불량률과 같이 ~값 미만에 관심이 있다면 왼쪽이 단측 검정, 수익과 같이 ~값 이상에 관심이 있다면 우측이 대립가설이 단측 검정을 실시한다. "차이가 없다" 같이 양쪽 모두에 관심이 있다면 양측검정을 실시한다.

⑵1종 오류(α)는 기각역(Rejection Region)이 정해지면 계산 가능하지만 2종 오류(β)는 대립가설의 값 중 임의의 값이 설정되어야 가능하다. $1-\beta$는 검정력(test power)으로 정의된다. 또한 우리의 관심이 대립가설에 있으므로 1종 오류를 설정하고(이를 유의수준이라 한다) 2종 오류를 최소화 하는 검정 방법을 구하게 된다. (1종, 2종 오류 모두 줄이는 방법은 없다.)

⑶Example 9.10을 보자. 실제 평균의 차이가 0.2초임에도 불구하고 Statistically 차이가 있다고 한다.(p=0.0124) 그러나 Practically 차이가 있다고 할 수 있나? 사람들이 다소 의아해 할 것이다. 이 경우에는 두 평균 차이에 대한 95% 신뢰구간을 제시하는 것도 하나의 방법이다.

⑷단측 검정일 때 귀무가설에 대하여, 대립가설이 $H_a : \theta < \theta_0$ 라면 귀무가설이 $H_0 : \theta \geq \theta_0$ 은 아닌가? 그냥 $H_0 : \theta = \theta_0$ 으로 하면 안되나? 1종 오류는 $\alpha = \max_{\theta \geq \theta_0} P(\text{검정통계량 } in\ RR)$ 에 해당된다. 그러므로 Boundary 값이 θ_0 에서 최대화되므로 $H_0 : \theta = \theta_0$ 이 적절하다. 그리고 우리의 관심은 귀무가설이 아니므로 귀무가설이 채택된다면 우리는 헛수고 한 것이다.

9.7 μ 와 $(\mu_1 - \mu_2)$ 에 대한 소표본 가설검정

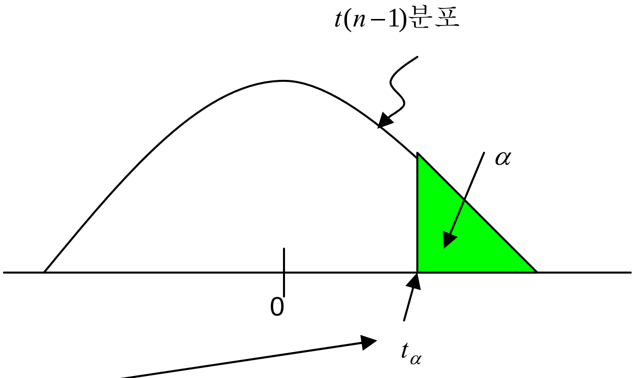
대표본일 경우에는 $T = \dfrac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}} \sim Normal(0,1)$ 을 이용하여 귀무가설($H_0 : \theta = \theta_0$)을 검정한다. 이에 해당되는 모수는 $\theta = \mu, p, \mu_1 - \mu_2, p_1 - p_2$ 이다.

이제 소표본일 경우 살펴보자. 소표표에서 모집단 비율(p), 모집단 비율 차이($p_1 - p_2$)에 대한 가설 검정은 다소 방법이 다르다. $H_0 : p = p_0$ 에 대한 가설검정은 9.1절의 EXAPMLE 9.1을 참고하기 바란다. 소표본 두 모집단 비율 차이 검정은 이 과정 범위를 넘는다.

모집단이 정규분포를 따르는 경우 확률표본 $(X_1, X_2, ... X_n)$ (표본의 크기 n 이 작다면)에 대해 다음이 성립한다. $T = \dfrac{\overline{X} - \mu_0}{S / \sqrt{n}} \sim t(n-1)$

μ 에 대한 소표본 가설검정

$(X_1, X_2, ... X_n)$ 이 정규분포(μ, σ^2)로부터 확률표본이라면 귀무가설 $H_0 : \mu = \mu_0$ 에 대한 가설 검정 절차는 다음과 같다.



대립가설: $H_a : \mu \neq \mu_0$
$H_a : \mu > \mu_0$
$H_a : \mu < \mu_0$

검정통계량: $T = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$

기각역(Rejection Region): $|t| > t_{\alpha/2}$
$t > t_\alpha$
$t < -t_\alpha$

그럼 왼쪽 단측 검정? 양측 검정?

EXAMPLE 9.11

철강은 적어도 3000파운드 이상이라고 판매자가 주장하였다. 이를 알아보기 위하여 8개 철강을 임의 추출하여 조사했더니 3041이고 표준편차는 39.1이었다. 유의수준 0.025로 그들의 주장이 맞는지 알아보시오. 철강 무게는 정규분포를 따른다고 한다.

①귀무가설: $H_0 : \mu = 3000$　　　대립가설: $H_0 : \mu > 3000$

②검정통계량: $T = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} = \dfrac{3041 - 3000}{39.1/\sqrt{8}} = 2.966$

③유의수준 0.025이므로 $t_\alpha = 2.365$ 이다. 그러므로 기각역은 $RR = \{t : t \geq 2.365\}$ 검정통계량이 기각역이 속하므로 귀무가설이 가각된다. 즉 그들의 주장은 옳다.

EXAMPLE 9.12

EXAMPLE 9.11에서 유의확률을 구해보자.

t-분포표에서 찾을 수 있나? 표를 보면... 우측 기각값 (임계값 t_α)이 0.1, 0.05, 0.025, 0.01, 0.005에 따른 값들만 주어져 있다. 자유도가 7이므로 표에 의해

$t(df = 7; \alpha = 0.025) = 2.365$, $t(df = 7; \alpha = 0.01) = 2.998$ 이므로 유의확률은 0.01과 0.025 사이이다. 정확한 값은 통계소프트웨어를 이용하여 구하면 된다. $p = 1 - CDF('T', 2.966, 7);$

$(\mu_1 - \mu_2)$에 대한 소표본 가설검정

$(X_{11}, X_{12}, \ldots X_{1n})$이 정규분포$(\mu_1, \sigma_1^2)$로부터 확률표본, $(X_{21}, X_{22}, \ldots X_{2n})$이 정규분포$(\mu_2, \sigma_2^2)$로부터 확률표본이고 서로 독립이라고 하자.

Pooled Variance(통합분산)을 다음과 같이 정의하자. $S_p^2 = \dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{(n_1+n_2-2)}$

귀무가설: $H_0 : \mu_1 = \mu_2$ 대립가설: $H_a : \mu_1 \neq \mu_2$, $H_a : \mu_1 < \mu_2$, $H_a : \mu_1 < \mu_2$

검정통계량: $T = \dfrac{(\bar{X}_1 - \bar{X}_2) - 0}{S_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$ 기각역(Rejection Region): $|t| > t_{\alpha/2}$, $t > t_\alpha$, $t < -t_\alpha$

만약 모집단이 정규분포가 아니라면? 그리고 사실 소표본의 경우 모집단이 정규분포라는 가정에 대한 진위를 알아보는 것은 쉽지 않다. 그러나 다행히도 위의 t-검정은 모집단이 정규분포가 아니더라도 모집단이 대칭이기만 하면 유효하다. 이렇게 다소 가정이 무너짐에 상관없이 사용할 수 있는 검정 방법을 Robust 검정방법이라 한다.

EXAMPLE 9.13

서로 다른 조립 공정의 생산 시간의 차이가 있는지 알아보기 위하여 조사하였다. 물론 공정 시간 데이터는 정규분포를 따른다고 하자. 유의수준 0.05에서 차이가 있는지 검정하시오.

조립 방법1: $n_1 = 9, \bar{X}_1 = 35.22, S_1^2 = 24.45$

조립 방법2: $n_2 = 9, \bar{X}_2 = 31.56, S_2^2 = 20.03$

①귀무가설: $H_0 : \mu_1 = \mu_2$ 대립가설: $H_a : \mu_1 \neq \mu_2$

②검정통계량: $T = \dfrac{(35.22 - 31.56)}{4.716\sqrt{\dfrac{1}{9} + \dfrac{1}{9}}} = 1.65 \sim t(9+9-2=16)$, $S_p^2 = \dfrac{(9-1)24.45 + (9-1)20.03}{(9+9-2)}$

양측검정, t-분포에서 임계값은 $t(df = 16, \alpha = 0.025) = 2.12$ 이다. 검정통계량이 기각역에 속하지 않으므로 귀무가설은 채택된다. 즉 차이는 없다.

EXAMPLE 9.14

EXAMPLE 9.13에서 유의확률을 구해보자.

T-분포표에 의하면 검정통계량 1.65는 $t(df=16, \alpha=0.05)=1.746, t(df=16, \alpha=0.1)=1.337$ 사이이므로 양측이므로 2배를 해야 한다. 유의확률은 0.1과 0.2 사이이다.

HOMEWORK #20-2

콘크리트 강도는 건조 방법에 따라 다른지 알아보기 위하여 다음 조사를 하였다. 각 건조 방법은 서로 독립적으로 시험하였고 강도 데이터는 정규분포를 따른다고 하자.

건조 방법1: $n_1=7, \bar{X}_1=3250, S_1=210$

건조 방법2: $n_2=10, \bar{X}_2=3240, S_2=190$

⑴유의수준 0.05에서 가설의 유의성을 검정하시오.

⑵유의확률을 계산하고 해석하시오.

HOMEWORK #20-3

우리나라 전체 청년 실업률 45%미만이라고 주장한다. 이를 알아보기 위하여 18개 도시를 임의 추출하여 실업률을 조사하여 다음 자료를 얻었다. 실업률은 정규분포를 따른다고 하자.

42.9	53.9	48.6	47.9	47.7	46.6	40.5	39.7	38.7
38.0	36.0	35.1	35.0	35.0	33.5	29.0	27.5	36.8

⑴유의확률을 계산하고 가설을 검정하시오.

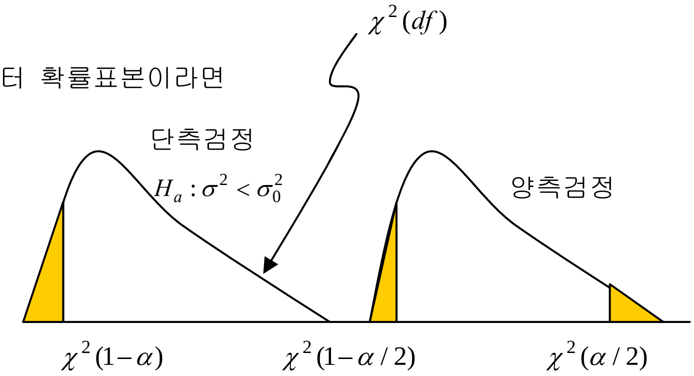
⑵실업률을 95% 신뢰구간을 구하시오.

⑶SAS를 이용하여 확인하시오. (페이지 ??? 참고)

9.8 분산에 관한 검정

모집단이 정규분포를 따르는 경우 확률표본 $(X_1, X_2, \ldots X_n)$에 대해 $T=\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

σ^2에 대한 가설검정



$(X_1, X_2, \ldots X_n)$이 정규분포(μ, σ^2)로부터 확률표본이라면

귀무가설 $H_0:\sigma^2=\sigma_0^2$

$H_a:\sigma^2 \neq \sigma_0^2$

대립가설: $H_a:\sigma^2 > \sigma_0^2$

$H_a:\sigma^2 < \sigma_0^2$

검정통계량: $T=\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

$T<\chi^2_{1-\alpha/2}$ or $T>\chi^2_{\alpha/2}$

기각역(Rejection Region): $T>\chi^2_{\alpha}$

$T<\chi^2_{1-\alpha}$

EXAMPLE 9.15

직경 10mm 너트를 만드는 회사는 직경의 오차가 0.002mm미만이라고 주장한다. 이를 알아보기 위하여 10개 너트를 임의 추출하여 분산을 조사하였더니 0.0015이었다. 그들의 주장이 맞는지 유의수준 5%에서 가설 검정하시오.

①귀무가설: $H_0:\sigma^2=0.02$ 대립가설: $H_a:\sigma^2<0.02$

②검정통계량: $T=\dfrac{(10-1)0.0015}{0.002}=6.75$, 기각역 $RR\{T;T<\chi^2(df=9,1-\alpha=0.95)=3.32\}$

계산된 검정통계량이 기각역에 속하지 않으므로 귀무가설은 채택되고 그들의 주장이 옳지 않다고 결론 지을 수 있다.

유의확률을 구해보자. 검정통계량의 값이 $\chi^2(df=9,0.9)=4.16)$ 보다 크므로 유의확률은 0.1 이상이라 할 수 있다.

만약 대립가설을 양측으로 한다면...

단측검정의 유의확률을 2배로 하면 된다.

기각역은 $RR\{T;T>\chi^2(df=9,0.025)=19.02\}$과 $RR\{T;T<\chi^2(df=9,1-\alpha/2=0.975)=2.7\}$이다.

HOMEWORK #20-4

모집단의 분산이 4라고 한다. 이를 확인하기 위하여 16개 표본을 추출하여 표본 분산을 조사하였더니 6.1이었다. 유의수준 5%에서 가설검정 하고 (근사) 유의수준을 계산하시오.

$\sigma_1^2 = \sigma_2^2$ 에 대한 가설검정

카이제곱 분포의 비는 F-분포를 따른다는 사실을 이용하자.

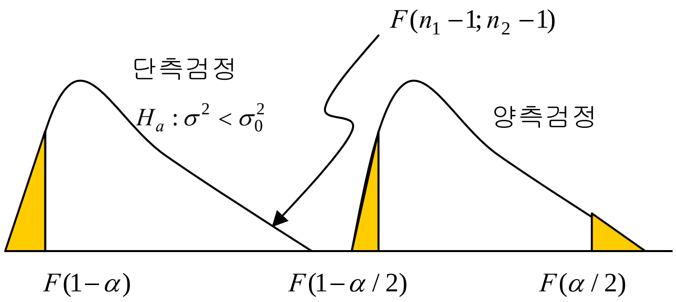
$\dfrac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2(n_1-1)$, $\dfrac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2(n_2-1)$ 이고 서로독립이라면

$$\frac{\dfrac{(n_1-1)S_1^2}{\sigma_1^2}/(n_1-1)}{\dfrac{(n_2-1)S_2^2}{\sigma_2^2}/(n_2-1)} = \frac{\dfrac{S_1^2}{\sigma_1^2}}{\dfrac{S_2^2}{\sigma_2^2}} \sim F(n_1-1; n_2-1)$$

$(X_{11}, X_{12}, \ldots X_{1n})$ 이 정규분포(μ_1, σ_1^2)로부터 확률표본, $(X_{21}, X_{22}, \ldots X_{2n})$ 이 정규분포(μ_2, σ_2^2)로부터 확률표본이고 서로 독립이라고 하자. 귀무가설 $H_a : \sigma_1^2 = \sigma_2^2$

대립가설: $H_a : \sigma_1^2 \neq \sigma_2^2$, $H_a : \sigma_1^2 > \sigma_2^2$, $H_a : \sigma_1^2 < \sigma_2^2$

검정통계량: $= \dfrac{S_1^2}{S_2^2} \sim F(n_1-1; n_2-1)$



기각역(Rejection Region):
$T > F(n_1-1, n_2-1; \alpha/2)$ or $T < F(n_1-1, n_2-1; 1-\alpha/2)$
$T > F(n_1-1, n_2-1; \alpha)$
$T < F(n_1-1, n_2-1; 1-\alpha)$

EXAMPLE 9.17

너트를 만드는 A회사는 B회사에 비해 너트 오차 분산의 작다고 주장하였다. A회사 너트 10개 조사하였더니 분산은 0.0001이었다. B회사 제품 20개를 조사하였더니 오차 분산은 0.0003이었다. 유의수준 5%에서 가설을 검정하시오.

①귀무가설: $H_0 : \sigma_1^2 = \sigma_2^2$ 대립가설: $H_a : \sigma_1^2 < \sigma_2^2$

②검정통계량: $F = \dfrac{0.0003}{0.0001} = 3$, 기각역 $RR\{T; T > F(df = 19,9;0.05) = 2.42\}$

계산된 검정통계량이 기각역에 속하므로 귀무가설은 기각되고 A회사 너트의 오차분산이 적다고 할 수 있다.

유의확률을 구해보자. 검정통계량의 값이 $F(19,9;0.025) = 2.88$ 과 $F(19,9;0.01) = 3.52$ 사이의 값이므로 유의확률은 2.5%와 1% 사이이다.

만약 검정통계량 $F = \dfrac{0.0001}{0.0002} = 1/3$ 으로 한다면

기각역 $RR\{T; T < F(df = 9,19;0.95) = 1/F(df = 19,9;0.05) = 1/2.42 = 0.41\}$ 동일하다.

왜냐하면 $F(n,m;\alpha) = \dfrac{1}{F(m,n;1-\alpha)}$ 이기 때문이다.

만약 대립가설을 양측으로 한다면...

단측검정의 유의확률을 2배로 하면 된다. 그리고 기각역은

$RR\{T; T > F(9,19;0.025) = 3.67\}$ 과 $RR\{T; T < F(9,19;0.975) = \dfrac{1}{F(19,9;0.025)} = \dfrac{1}{2.88} = 0.34\}$ 이다.

HOMEWORK #20-5

두 기업의 주가에 대한 종가를 16 상장일을 조사하여 다음 자료를 얻었다. 종가의 분산의 차이가 있는지 알아보시오. 유의수준은 2%로 하시오. 종가의 분산이 크다는 것은 일별 가격 변동이 크다는 것을 의미한다.

A기업: $\bar{X}_1 = 40.33, S_1^2 = 1.54$, B기업: $\bar{X}_2 = 42.54, S_2^2 = 2.96$

Review

가설> 귀무가설, 대립가설(연구가설), 대립가설에는 단측가설과 양측가설이 있다.

검정통계량> 검정통계량 분포(귀무가설이 사실)는 귀무가설에 설정된 모수에 의존하면 안된다.

오류> 1종오류, 2종오류

유의수준과 기각역

검정통계량과 유의확률

가설검정과 신뢰수준

- 대표본 가설검정: $\mu, p, \mu_1-\mu_2, p_1-p_2$, 검정통계량 $\dfrac{\hat{\theta}-\theta_0}{\sigma_{\hat{\theta}}} \sim Normal(0,1)$
- 소표본 가설검정: $\mu, \mu_1-\mu_2$ 검정통계량: $\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}} \sim t(n-1)$, $\dfrac{\bar{x}_1-\bar{x}_2-(\mu_1-\mu_2)}{s_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t(n_1+n_2-2)$
- 모분산 가설검정 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
- 모분산 차이 가설검정 $\dfrac{S_1^2}{S_2^2} \sim F(n_1-1;n_2-1)$

9.9 검정력과 Neyman-Pearson Lemma

검정방법의 정도(goodness)는 1종 오류와 2종 오류로 판단하는데, 검정 시작 시 1종 오류는 유의수준으로 설정되므로 좋은 검정 방법이란 2종 오류를 최소화 하는 것이다.

DEFINITION 검정력

T 을 검정통계량이라 하고 RR 을 기각역이라 하자. 이 때 검정력(test power, power of test)은 다음과 같이 정의한다. 검정력은 모수 θ 의 함수이다.

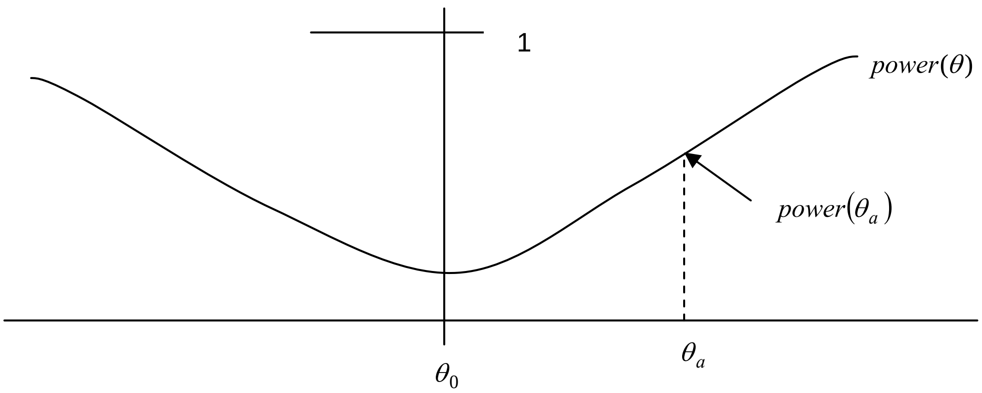
$power(\theta) = P(T\ in\ RR\,|\,\text{모수가 }\theta\text{일 경우})$

만약 대립가설의 모수 하나의 값(대립가설은 귀무가설과는 달리 모수의 영역이다)을 θ_a 라 한

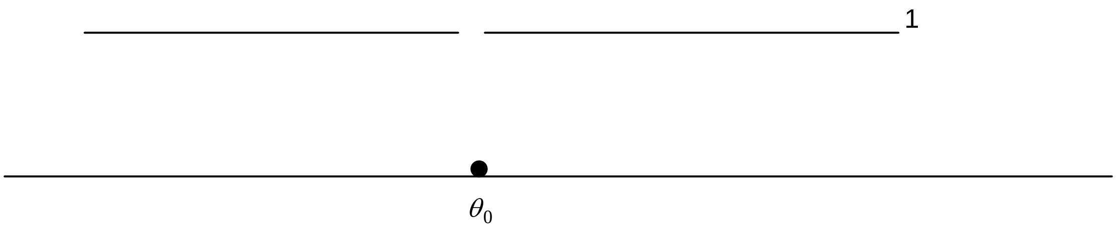
다면 검정력은 $power(\theta_a) = P(H_0\text{기각} \mid \text{모수가 } \theta_a \text{일경우})$ 이다.

2종 오류가 $\beta(\theta_a) = P(H_0\text{채택} \mid \text{모수가 } \theta_a \text{일경우})$ 이므로 $\boxed{power(\theta_a) = 1 - \beta(\theta_a)}$ 이다.

$H_0 : \theta = \theta_0 \text{ vs } H_0 : \theta \neq \theta_0$ 가설검정의 경우 전형적인 검정 곡선(power curve)은 다음과 같다.



검정력은 귀무가설의 θ_0 에서 최저였다가 멀어질수록 증가한다. 그러나 가장 이상적인 검정 함수는 다음과 같아야 한다. 그러나 실제 이런 검정 방법은 존재하지 않는다. 그러므로 유의수준 α 을 설정하고 2종 오류를 최소화할 수 있는 검정 방법(**기각역 찾음**)을 찾는다.



DEFINITION 단순가설과 복합가설

가설에 설정된 모수 값이 하나인 경우 이를 단순가설(simple hypothesis)이라 하고 설정된 모수가 영역인 경우 이를 복합가설(composite hypothesis)이라 한다.

단순 귀무가설 $H_0 : \theta = \theta_0$ 와 단순 대립가설 $H_a : \theta = \theta_a$ 을 검정한다고 하자. 우리는 모수 2 값 (θ_0, θ_a)에 관심이 있다. $\alpha = power(\theta_0)$ 을 설정하고 $\beta = 1 - power(\theta_a)$ 을 최소화 하는 기각역을 구한다. 이는 $power(\theta_a)$ 을 최대화 하는 최대 검정력 검정(most powerful test)을 찾는 것과 동일하다. Most powerful test 방법을 찾는 것은 다음 정리에 의한다.

THEOREM (Neyman-Pearson Lemma)

단순 귀무가설 $H_0 : \theta = \theta_0$ 와 단순 대립가설 $H_a : \theta = \theta_a$ 을 검정한다고 하자. 가설 검정을 확률표본 $(X_1, X_2, ..., X_n)$ 으로부터 구한 통계량을 이용한다. 이 통계량의 분포는 당연히 모수 θ 을 갖는다. 모수가 θ 일 때 우도 함수(Likelihood function)을 $L(\theta; x_1, x_2, ..., x_n)$ 이라 하자.

유의수준 α 가 주어진 상황에서 $power(\theta_a)$ 을 최대화 하는 기각역 RR 은 다음에 의해 정해진다. $\frac{L(\theta_0; x_1, x_2, ..., x_n)}{L(\theta_a; x_1, x_2, ..., x_n)} < k$. 이렇게 얻은 검정 방법(기각역) Most Powerful 검정이다.

EXAMPLE 9.18

모집단 $f(x:\theta) = \theta x^{\theta-1}, 0 < x < 1$ 으로부터 모수 θ 에 대한 가설 검정을 위하여 표본 크기 1인 확률표본 X_1 을 뽑았다고 한다. 귀무가설 $H_0 : \theta = \theta_0 = 2$ 와 대립가설 $H_a : \theta = \theta_a = 1$ 을 유의수준 0.05에서 Most Powerful 검정방법을 찾으시오.

Neyman-Pearson Lemma에 의해 $\frac{L(\theta_0; x_1, x_2, ..., x_n)}{L(\theta_a; x_1, x_2, ..., x_n)} = \frac{f(x_1; \theta_0 = 2)}{f(x_1; \theta_a = 1)} = 2x_1 < k$ ➔ $x_1 < k/2 = k'$

기각역 $x_1 < k'$ 가 $H_0 : \theta = \theta_0 = 2$, $H_a : \theta = \theta_a = 1$ 검정에 대한 Most Powerful 검정을 제공

대립가설이 $H_a : \theta = \theta_a < \theta_0$ 일 때 기각역 $x_1 < k'$ 은 Most powerful test이므로 이를 Uniformly most powerful(UMPT) 검정이라 한다.

$$\alpha = P(X_1\ in\ RR \mid H_0 \text{사실}) = P(X_1\ in\ RR \mid \theta = 2) = \int_0^{k'} 2x_1 dx_1 = k'^2 \quad \text{즉} \quad k'^2 = 0.05 \ ➔\ k' = \sqrt{0.05}$$

검정력? $power(\theta_a = 1) = P(X_1\ in\ RR \mid H_a \text{사실}) = P(X_1\ in\ RR \mid \theta = 1) = \int_0^{0.2236} 1 dx_1 = 0.2236$

검정 통계량과 기각역의 형태는 귀무가설과 대립가설에 의존한다. 만약 위의 예제에서 대립가설을 $H_a : \theta = \theta_a > \theta_0$ 이면 검정통계량은 X_1^2 이고 Uniformly Most powerful(UMPT) 검정 기각역은 $X^2 > k'$ 의 형태이다. (N-P Lemma)

이산형 분포에서는 설정된 유의수준에 딱 맞는 기각역을 찾는 것은 어렵다. 근사한 기각역을 찾으면 된다. 아니면 Randomized Test를 하면 된다.

EXAMPLE 9.19

모집단 $f(x:\theta)=p^x(1-p)^{1-x}, x=0,1$ 으로부터 모수 $\theta=p$ 에 대한 가설 검정을 위하여 표본 크기 n인 확률표본 $(X_1,X_2,...,X_n)$ 을 뽑았다고 하자. 귀무가설 $H_0:\theta=p_0=0.5$ 와 대립가설 $H_a:\theta=p_a=0.2$ 을 유의수준 0.05에서 Most Powerful 검정방법을 찾으시오.

Neyman-Pearson Lemma에 의해 $\frac{L(p_0=0.5;x_1,x_2,...,x_n)}{L(p_a=0.2;x_1,x_2,...,x_n)}=\frac{0.5^{\sum x_i}0.5^{n-\sum x_i}}{0.2^{\sum x_i}0.8^{n-\sum x_i}}<k$

이를 정리하면 $\sum x_i<k'$ 이 된다. 대립가설의 모수가 0.5 미만이면 기각역은 $\sum x_i<k'$ 이다.

만약 대립가설의 모수가 0.5보다 크면 기각역은 $\sum x_i>k'$.

$\sum x_i \sim Binomial(n,p=0.5)$ 이므로 $\alpha=P(\sum x_i<k'|\sum x_i \sim Binomial(n,p=0.5))=0.05$.

$\sum x_i$ 는 이산형이므로 정수인 k' 을 찾을 수 없다.

Randomized Test

정수 i 에 대해 $0.03=P(\sum x_i\le i|\sum x_i \sim B(n,p=0.5))$, $0.06=P(\sum x_i\le i+1|\sum x_i \sim B(n,p=0.5))$ 이 성립한다고 하자. 이 경우 $\sum x_i=i+1$ 이 나오면 (주사위를 던져) 2/3의 확률로 귀무가설을 기각하고 1/3의 확률로 귀무가설을 채택한다.

EXAMPLE 9.20

모집단 $f(x:\theta)=\frac{1}{\theta}e^{-x/\theta};0<x$ 으로부터 모수 θ 에 대한 가설 검정을 위하여 표본 크기 n 인 확률표본 $(X_1,X_2,...,X_n)$ 을 뽑았다고 하자. 귀무가설 $H_0:\theta=\theta_0$ 와 대립가설 $H_a:\theta=\theta_a<\theta_0$ 을 유의수준 α 에서 Uniformly Most Powerful 검정방법을 찾으시오.

우도함수: $L(\theta;x_1,x_2,...,x_n)=\prod f(x_i;\theta)=\prod\frac{1}{\theta}e^{-x_i/\theta}=(\frac{1}{\theta})^n e^{-\sum x_i/\theta}$

Neyman-Pearson Lemma에 의해 $\frac{L(\theta_0;x_1,x_2,...,x_n)}{L(\theta_a;x_1,x_2,...,x_n)}=\frac{(\frac{1}{\theta_0})^n e^{-\sum x_i/\theta_0}}{(\frac{1}{\theta_a})^n e^{-\sum x_i/\theta_a}}<k$

$$\Rightarrow (\frac{\theta_a}{\theta_0})^n e^{-\sum x_i[1/\theta_0 - 1/\theta_a]} < k \quad \Rightarrow e^{-\sum x_i[1/\theta_0 - 1/\theta_a]} < (\frac{\theta_0}{\theta_a})^n k \quad \Rightarrow e^{-\sum x_i[1/\theta_0 - 1/\theta_a]} < (\frac{\theta_0}{\theta_a})^n k$$

$$\Rightarrow -\sum x_i[1/\theta_0 - 1/\theta_a] < \ln(\frac{\theta_0}{\theta_a})^n k \quad \Rightarrow \sum x_i[1/\theta_0 - 1/\theta_a] > \ln(\frac{\theta_0}{\theta_a})^n k$$

$\Rightarrow \sum x_i < \ln(\frac{\theta_0}{\theta_a})^n k \frac{1}{[1/\theta_0 - 1/\theta_a]} = k'$ 그러므로 기각역은 $RR = \{\sum X_i < k'\}$ (UMPT)이다.

기각역으로 사용되려면 분포가 알려진 (검정)통계량이어야 한다.

$X_i \sim Exponential(\theta)$ 이므로 $\sum X_i \sim Gamma(n, \theta)$ 이다.

$P(\sum X_i < k' \mid \sum X_i \sim Gamma(n, \theta_0)$귀무가설사실$) = \alpha$ 을 만족하는 k' 을 구하면 된다.

만약 대립가설이 $H_a : \theta = \theta_a > \theta_0$ 이면 기각역은 $RR = \{\sum X_i > k' \mid H_0$사실$\}$ 이다.

그러므로 만약 대립가설이 $H_a : \theta \neq \theta_a$ 이면 Uniformly Most Powerful Test는 존재하지 않는다.

EXAMPLE 9.21

모집단 $f(x:\theta) \sim Normal(\mu, \sigma^2)$ 으로부터 표본 크기 n 인 확률표본 $(X_1, X_2, ..., X_n)$ 을 뽑았다고 하자. 모분산 σ^2 을 알고 있다면 귀무가설 $H_0 : \mu = \mu_0$ 와 대립가설 $H_a : \mu = \mu_a > \mu_0$ 을 유의수준 α 에서 Uniformly Most Powerful 검정방법을 찾으시오.

우도함수: $L(\mu; x_1, x_2, ..., x_n) = \prod f(x_i; \mu) = \prod \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = (\frac{1}{\sigma\sqrt{2\pi}})^n e^{-\frac{\sum(x_i-\mu)^2}{2\sigma^2}}$

Neyman-Pearson Lemma에 의해 $\frac{L(\mu_0; x_1, x_2, ..., x_n)}{L(\mu_a; x_1, x_2, ..., x_n)} = \frac{(\frac{1}{\sigma\sqrt{2\pi}})^n e^{-\frac{\sum(x_i-\mu_0)^2}{2\sigma^2}}}{(\frac{1}{\sigma\sqrt{2\pi}})^n e^{-\frac{\sum(x_i-\mu_a)^2}{2\sigma^2}}} < k$

$$e^{-\frac{\sum(x_i-\mu_0)^2}{2\sigma^2} + \frac{\sum(x_i-\mu_a)^2}{2\sigma^2}} < k \quad \Rightarrow -\frac{\sum(x_i-\mu_0)^2}{2\sigma^2} + \frac{\sum(x_i-\mu_a)^2}{2\sigma^2} < \ln k$$

$$\Rightarrow -\frac{\sum(x_i-\mu_0)^2 - \sum(x_i-\mu_a)^2}{2\sigma^2} < \ln k \quad \Rightarrow -\sum(x_i-\mu_0)^2 + \sum(x_i-\mu_a)^2 < 2\sigma^2 \ln k$$

$$\Rightarrow 2\sum x_i\mu_0 - 2\sum x_i\mu_a < 2\sigma^2 \ln k + (\mu_0^2 - \mu_a^2) \quad \Rightarrow 2\sum x_i(\mu_0 - \mu_a) < 2\sigma^2 \ln k + (\mu_0^2 - \mu_a^2)$$

$\Rightarrow \sum x_i > \dfrac{2\sigma^2 \ln k + (\mu_0^2 - \mu_a^2)}{2(\mu_0 - \mu_a)}$ 기각역은 $RR = \{\sum x_i > k'\}$

$X_i \sim Normal(\mu_0, \sigma^2)$ 이므로 $\sum X_i \sim Normal(n\mu_0, n\sigma^2)$ 이다.

그러므로 $RR = \{\bar{x} = \dfrac{\sum x_i}{n} > k'/n = k''\}$ 을 사용하는 것이 적절하다. Neyman-Pearson 정리에 의해 이 기각역은 UMPT를 제공한다.

그러므로 기각역은 $\alpha = \{\bar{x} > k'' | \bar{x} \sim Normal(\mu_0, \dfrac{\sigma^2}{n})\}$ 을 만족하는 k'' 찾으면 된다.

대립가설이 $H_a : \mu = \mu_a < \mu_0$ 이라면 $\alpha = \{\bar{x} < k'' | \bar{x} \sim Normal(\mu_0, \dfrac{\sigma^2}{n})\}$ 을 만족하는 k'' 을 찾으면 Most powerful test 이다. 앞 예제에서 언급하였듯이 이것은 대립가설 $H_a : \mu \neq \mu_a$ 에 대한 UMPT는 얻을 수 없다.

한가지 더 언급할 것은 위의 예제에서 모분산 σ^2 을 모른다면? 우리가 관심을 갖는 모수는 모평균 μ 이므로 모분산 σ^2 에 대한 추정치나 사전 정보가 있어야 한다. 이런 모수는 nuisance parameter라 한다. Nuisance parameter가 있는 경우 Neyman-Pearson 방법으로는 UMPT를 얻을 수 없다. 이에 대한 검정 방법으로 우도비 검정이라 한다.

HOMEWORK (optional)

모집단 $f(x:\theta) = (\dfrac{1}{2\theta^3}) x^2 e^{-x/\theta}; 0 < x$ 으로부터 모수 θ 에 대한 가설 검정을 위하여 확률표본 $(X_1, X_2, ..., X_n)$ 을 뽑았다고 하자. 귀무가설 $H_0 : \theta = \theta_0$ 와 대립가설 $H_a : \theta = \theta_a > \theta_0$ 을 유의수준 α 에서 Uniformly Most Powerful 검정방법을 찾으시오.

HOMEWORK (optional)

모집단 $f(x:\theta) \sim Poisson(\lambda)$ 으로부터 모수 λ 에 대한 가설 검정을 위하여 확률표본 $(X_1, X_2, ..., X_n)$ 을 뽑았다고 하자. 귀무가설 $H_0 : \lambda = \lambda_0$ 와 대립가설 $H_a : \lambda_a > \lambda_0$ 을 유의수준 α 에서 Uniformly Most Powerful 검정방법을 찾으시오.

9.10 우도비 검정 (LRT: Likelihood Ratio Test)

단순 귀무가설과 단순 대립가설에 대한 Most powerful test을 얻으려면 Neyman-Pearson 정리를 이용하면 된다. 물론 대립가설이 단측 가설(한 쪽 방향)에 Most powerful 검정을 얻을 수 있지만… 양측 검정이나 Nuisance parameter가 있는 경우에는 N-P 정리는…

귀무가설에 설정된 모수의 공간을 Ω_0 (예: $H_0:\theta=\theta_0$ 인 경우 $\Omega_0=\{\theta:\theta=\theta_0\}$)이라 하고 대립가설의 모수 공간을 Ω_a (예: $H_0:\theta<\theta_0$ 인 경우 $\Omega_a=\{\theta:\theta<\theta_0\}$)하자. $\Omega=\{\theta:\theta\le\theta_0\}$.

모수 공간 Ω 에서 우도함수를 $L(\Omega;x_1,x_2,...x_n)$ 라 정의하고 모수에 대한 MLE(최대우도추정량)을 추정치로 사용한 우도함수를 $L(\hat{\Omega};x_1,x_2,...x_n)$ 라 하자.

DEFINITION LTR 검정

$\lambda=\dfrac{L(\hat{\Omega}_0)}{L(\hat{\Omega})}$ 라 정의하자. 귀무가설 $H_0:\theta\in\Omega_0$, 대립가설 $H_a:\theta\in\Omega_a$ 에 대한 우도비(LTR) 검정은 검정통계량으로 λ 가 사용되며 기각역은 $\lambda\le k$ 이다.

λ 의 값은 $0<\lambda<1$ 이다. 만약 $\lambda\to 0$ 이면 귀무가설 하의 우도함수 값이 매우 작다는 것을 의미하므로 대립가설에 비해 귀무가설을 기각하는 것이 적절하다. 만약 $\lambda\to 1$ 이면 귀무가설이 진실일 가능성이 높다. 그러므로 기각역이 $\lambda\le k$ 이다.

EXAMPLE 9.22

모집단 $f(x:\theta)\sim Normal(\mu,\sigma^2)$ 으로부터 표본 크기 n 인 확률표본 $(X_1,X_2,...,X_n)$ 을 뽑았다고 하자. 모분산 σ^2 을 알지 못할 때(nuisance parameter) 귀무가설 $H_0:\mu=\mu_0$ 와 대립가설 $H_a:\mu=\mu_a>\mu_0$ 을 유의수준 α 에서 Uniformly Most Powerful 검정방법을 찾으시오.

귀무가설 하에서 모수 공간은 $\Omega_0=\{\mu,\sigma^2;\mu=\mu_0,\sigma^2>0\}$ 이고 대립가설 하에서 모수 공간은 $\Omega_a=\{\mu,\sigma^2;\mu>\mu_0,\sigma^2>0\}$ 이므로 $\Omega=\{\mu,\sigma^2;\mu\ge\mu_0,\sigma^2>0\}$ 이다.

$L(\Omega_0;x_1,x_2,...,x_n)=\prod f(x_i;\mu,\sigma^2)=\prod\dfrac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$ 에서 μ 는 μ_0 으로 고정되어 있으므로

모분산에 대한 MLE 추정치를 구하면 $\hat{\sigma}_0^2=\dfrac{1}{n}\sum(X_i-\mu_0)^2$ 이다.

$L(\Omega; x_1, x_2, ..., x_n) = \prod f(x_i; \mu, \sigma^2) = \prod \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$ 에서 μ 의 MLE는 $\hat{\mu} = \bar{X}$ (조건이 없다면)이나 모수 공간 Ω 에서는 μ 의 MLE는 $\hat{\mu} = \max(\bar{X}, \mu_0)$ (왜냐하면 $\bar{X}$ 가 귀무가설의 μ_0 보다 적다면 추정의 의미가 없기 때문이다)이다.

또한 모분산에 대한 MLE 추정치를 구하면 $\hat{\sigma}^2 = \frac{1}{n}\sum(X_i - \hat{\mu})^2$ 이다.

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \frac{\left(\frac{1}{\sqrt{2\pi\hat{\sigma}_0^2}} e^{-n/2}\right)^n}{\left(\frac{1}{\sqrt{2\pi\hat{\sigma}^2}} e^{-n/2}\right)^n} = \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right)^{n/2} = \begin{cases} \left[\frac{\sum(X_i - \bar{X})^2}{\sum(X_i - \mu_0)^2}\right]^{n/2}, & if\ \bar{X} > \mu_0 \\ 1, & if\ \bar{X} \le \mu_0 \end{cases}$$ 이다.

$$\sum(X_i - \mu_0)^2 = \sum(X_i - \bar{X})^2 + \sum(\bar{X} - \mu_0)^2$$

$$\lambda = \left[\frac{\sum(X_i - \bar{X})^2}{\sum(X_i - \mu_0)^2}\right]^{n/2} < k \Rightarrow \frac{\sum(X_i - \bar{X})^2}{\sum(X_i - \mu_0)^2} < k^{2/n} \Rightarrow \frac{\sum(X_i - \bar{X})^2}{\sum(X_i - \bar{X})^2 + \sum(\bar{X} - \mu_0)^2} < k^{2/n}$$

$$\Rightarrow \frac{1}{1 + \frac{\sum(\bar{X} - \mu_0)^2}{\sum(X_i - \bar{X})^2}} < k^{2/n} \Rightarrow \frac{\sum(\bar{X} - \mu_0)^2}{\sum(X_i - \bar{X})^2} > k' \Rightarrow \frac{n(\bar{X} - \mu_0)^2}{\sum(X_i - \bar{X})^2/(n-1)} > (n-1)k'$$

$$\Rightarrow \frac{\sqrt{n}(\bar{X} - \mu_0)}{S} > \sqrt{(n-1)k'} = k''$$

즉 기각역은 $\Rightarrow \frac{(\bar{X} - \mu_0)}{S/\sqrt{n}} > k''$ 이다. $\frac{(\bar{X} - \mu_0)}{S/\sqrt{n}}$ 은 t-분포를 따르는지 알고 있다. 즉 우도비 검정은 t-검정과 일치한다.

일반적으로 우도비 검정의 경우 검정통계량의 분포를 아는 것은 쉽지 않다. 이런 경우 다음 정리를 이용하여 검정하게 된다.

THEOREM 정리

$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})}$ 라 정의하자. 표본의 크기 n 이 충분히 크다면 귀무가설 $H_0: \theta \in \Omega_0$, 대립가설 $H_a: \theta \in \Omega_a$ 에 대한 우도비(LTR) 검정은 검정통계량 λ 에 대해 다음이 성립한다.

$-2\ln\lambda \sim (app)\chi^2(r_0 - r)$; r 은 모수 공간 Ω 하에서 설정된 모수의 개수이고 r_0 은 모수 공

간 Ω_0 (귀무가설) 하에서 설정된 모수의 개수이다.

EXAMPLE 9.23

작업 라인이 2개 있다. 각 작업을 일주일 단위로 생산된 제품에 대한 불량 개수를 조사하였더니 작업라인1은 평균 20, 작업라인 2는 22이다(표본의 개수는 각각 100이다). 불량 개수는 포아송 분포를 따르고 작업 라인 1의 평균을 λ_1, 작업 라인 2의 평균을 λ_2 라 하자. 귀무가설 $H_0 : \lambda_1 = \lambda_2$ 와 대립가설 $H_a : \lambda_1 \neq \lambda_2$ 을 검정하는 우도비 검정을 유의수준 0.01에서 실시하자.

작업라인 1의 확률표본을 $(X_1, X_2, ..., X_m)$, 작업라인 2의 확률표본을 $(Y_1, Y_2, ..., Y_n)$ 라 하자.

귀무가설 모수 공간은 $\Omega_0 = \{\lambda_1 = \lambda_2 = \lambda\}$ 이고 전체 모수 공간은 $\Omega = \{\lambda_1, \lambda_2 > 0\}$ 이다.

전체 모수 공간 결합 우도비는 $L(\Omega; x_1, x_2, ..., x_n, y_1, y_2, ..., y_n) = \prod \frac{e^{-\lambda_1} \lambda_1^{x_i}}{x_i!} \prod \frac{e^{-\lambda_2} \lambda_2^{y_i}}{y_i!}$ 이고

귀무가설 모수 공간 결합 우도비는 $L(\Omega_0; x_1, x_2, ..., x_n, y_1, y_2, ..., y_n) = \prod \frac{e^{-2\lambda} \lambda^{x_i + y_i}}{x_i! y_i!}$ 이다.

$L(\Omega; x_1, x_2, ..., x_n, y_1, y_2, ..., y_n) = \frac{1}{\prod x_i! y_i!} \lambda_1^{\sum x_i} \lambda_2^{\sum y_i} e^{-n(\lambda_1 + \lambda_2)}$ ➔ MLE는 $\hat{\lambda}_1 = \overline{X}$, $\hat{\lambda}_2 = \overline{Y}$ 이다.

$L(\Omega_0; x_1, x_2, ..., x_n, y_1, y_2, ..., y_n) = \frac{1}{\prod x_i! y_i!} \lambda^{\sum x_i + \sum y_i} e^{-n(2\lambda)}$ ➔ MLE는 $\hat{\lambda} = \frac{\overline{X} + \overline{Y}}{2} = \frac{20 + 22}{2} = 21$ 이다.

그러므로 $\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \frac{\hat{\lambda}^{n\overline{X} + n\overline{Y}}}{(\overline{X})^{n\overline{X}} (\overline{Y})^{n\overline{Y}}}$ 이다. 그러므로 $-2\ln\lambda = 9.53$ 이다.

귀무가설에서 설정된 귀무가설의 수는 1이고 전체 모수의 수는 2이다.

그러므로 $-2\ln\lambda \sim \chi^2(2-1=1)$ 이 성립한다. 자유도 1이고 유의수준이 0.01인 경우 기각치(임계치)는 χ^2-분포표에 의해 6.635이다. 9.53이 기각역에 속하므로 귀무가설은 기각된다.

HOMEWORK (optional)

남녀에 따른 통계원리 좋아하는 비율의 차이가 있는지 알아보기 위하여 조사를 하였다. 남자 200명 중 120명, 여자 200명 중 140명이 통계원리를 좋아한다고 응답하였다. 유의수준 5%에서 우도비 검정을 하시오. 그리고 대표본 검정도 해 보시오.