7장 가설과 검정

1. 기본개념
2. 분산의 검정과 평균의 차에 대한 검정
3. 검정력 함수
4. 최강력 검정법

1. 기본개념

[정의 7.1-1]
하나 이상의 확률변수들의 분포에 대한 주장을 **통계적 가설** (Statistical hypothesis)이라고 하며 H_0나 H_1 등으로 표기한다.

[정의 7.1-2]
자료로부터 뚜렷한 증거에 의하여 입증하고자 하는 가설을 **대립가설** (Alternative hypothesis)이라 하고 흔히 H_1 또는 H_A로 표기하며, 대립가설과 반대되는 가설을 **귀무가설** (Null hypothesis)이라 하고 흔히 H_0로 표기한다.

[정의 7.1-3]
검정하고자 하는 모수의 값이 하나의 값으로 표시되는 가설을 **단순가설** (Simple hypothesis)이라 하고 그렇지 않은 가설을 **복합가설** (Composite hypothesis)이라고 한다.

[정의 7.1-4]
통계적 가설을 검정하는 것은 확률변수의 측정값을 이용하여 가설 H_0를 채택하거나 기각하도록 결정하는 절차이다.

[정의 7.1-5]
통계적 가설 H_0의 검정에 대한 **유의수준**은 가설 H_0가 참일 때 H_0를 기각하게 되는 오류를 범할 확률이다.

[정의 7.1-6]
귀무가설 H_0를 기각하거나 채택하는데 기준이 되는 통계량을 **검정 통계량** (test statistic)이라고 한다.

[정의 7.1-7]
검정의 **기각역** (Critical region)이란 귀무가설 H_0를 기각하게 되는 검정 통계량의 측정값들의 집합이다.

<예 7.1-1> $X_1, X_2, \cdots, X_n$을 정규분포 $N(\mu, \sigma^2)$에서의 확률표본이라고 하고 귀무가설을 $H_0 : \mu = \mu_0$, 대립가설을 $H_1 : \mu \neq \mu_0$라고 하자. σ^2을 모르고 있는 경우 귀무가설을 검정하기 위한 검정 통계량으로써

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

을 사용하고자 한다. $\mu = \mu_0$일 때,

$$P\left[\frac{|\overline{X} - \mu_0|}{S/\sqrt{n-1}} > t_{\alpha/2}(n-1)\right] = \alpha$$

가 되므로 $\dfrac{|\overline{X} - \mu|}{S/\sqrt{n-1}} > t_{\alpha/2}(n-1)$이면 $H_0 : \mu = \mu_0$를 기각한다.

<참고로 μ에 대한 $100(1-\alpha)\%$ 신뢰구간은 $\overline{x} \pm t_{\alpha/2}(n-1)\dfrac{S}{\sqrt{n-1}}$이다.>

<예 7.1-2> 100명을 랜덤하게 선택하여 인터뷰한 결과 X명이 대통령의 대북 정책을 지지하였다고 하자. 모집단에서 이 정책을 지지하는 사람의 비율을 p라고 하면 X의 분포는 이항분포 $B(100, p)$이며 p의 추정량 $\hat{p}=\frac{X}{100}$의 점근 분포는 $N\left(p, \frac{p(1-p)}{100}\right)$이다. 귀무가설을 $H_0 : p = 0.25$라 두고, 조사된 X의 값이 $x = 20$이라 하면 $\hat{p}= 0.2$이다. $\frac{p(1-p)}{100}$의 편의 추정치는 $\frac{\hat{p}(1-\hat{p})}{100}=\frac{(0.2)(0.8)}{100}=0.0016$이며 $\Phi(1.96)= 0.975$이므로

$$\frac{|0.20 - 0.25|}{\sqrt{0.0016}} = 1.25 \le 1.96$$

이다. 따라서 유의수준 $\alpha = 0.25$에서 $H_0 : p = 0.25$를 기각하지 않는다. 참고로 0.25는 p의 95% 신뢰구간

$$\left[0.20 - 1.96\sqrt{\frac{(0.2)(0.8)}{100}},\ 0.20 + 1.96\sqrt{\frac{(0.2)(0.8)}{100}}\right] = [0.122,\ 0.279]$$

에 포함되어 있음을 알 수 있다.

<예 7.1-3> 중심극한정리에 의하여 $\overline{X}$의 분포가 정규분포 $N\left(\mu_0, \frac{\sigma^2}{n}\right)$임을 알고 $H_0 : \mu = \mu_0$를 유의수준 $\alpha = 0.05$에서 검정하고자 한다.

$$\left|\frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}\right| > 1.96$$

즉 $\overline{x} < \mu_0 - 1.96\frac{\sigma}{\sqrt{n}}$이거나 $\overline{x} > \mu_0 + 1.96\frac{\sigma}{\sqrt{n}}$이면 유의수준 $\alpha = 0.05$에서 귀무가설 $H_0 : \mu = \mu_0$를 기각한다. 이경우에는

$$P\left[\overline{X} < \mu_0 - 1.96\frac{\sigma}{\sqrt{n}}\right] = P\left[\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < -1.96\right] = \Phi(-1.96) = 0.025,$$

$$P\left[\mu_0 + 1.96\frac{\sigma}{\sqrt{n}} < \overline{X}\right] = P\left[1.96 < \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}\right] = 1 - \Phi(1.96) = 0.025$$

이다. 대립가설이 $H_1 : \mu \ne \mu_0$으로 주어질 때를 양측검정 (Two-sided test)이라고 하며, 유의수준 $\alpha = 0.05$는 둘로 나누어진다. 그러나 상황에 따라서는 대립가설을 $H_1 : \mu < \mu_0$ 또는 $H_1 : \mu > \mu_0$으로 두어야 하는 경우도 많이 발생한다. 이와 같은 대립가설이 주어지는 검정을 단측검정 (One-side test)이라고 한다.

가설을 검정하는 과정 중에는 두 가지 오류가 흔히 발생한다. 하나는 귀무가설이 참일 때 대립가설을 채택하는오류 (제 1종의 오류)이며 다른 하나는 대립가설이 참일 때 귀무가설을 채택하게 되는 오류 (제 2종의 오류)이다. 유의수준은 제1종의 오류를 범할 확률이 된다. 좋은 검정은 제 1종의 오류와 제 2종의 오류를 다같이 최소로 하여야 하지만 현실적으로는 주로 제1종의 오류를 조절한다.

<예 7.1-4> 지금까지 사용되고 있는 전구의 평균수명은 평균이 $\mu = 1,500$시간, 분산은 $\sigma^2 = 100^2$인 정규분포를 따르는 것으로 알고 있다. 새로운 제품을 개발하여 25개의 시제품을 랜덤하게 택하여 수명을 측정한 결과 $\bar{x} = 1,550$을 얻었다. 신제품의 수명이 기존 제품보다 더 길다고 볼 수 있는가?

<풀이> 귀무가설은 $H_0 : \mu = 1,500$이고 대립가설은 $H_1 : \mu > 1,500$이다. 검정통계량의 값이

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X} - 1500}{20} = \frac{1550 - 1500}{20} = 2.5 \geq z_{0.05} = 1.645$$

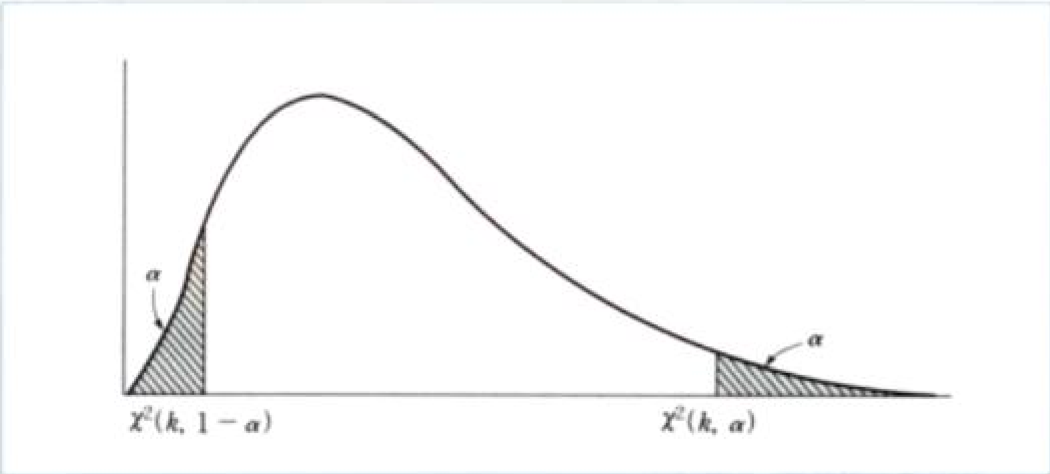
이므로 유의수준 $\alpha = 0.05$에서 귀무가설을 기각한다.

위의 예에서 표준정규분포표에 의하면 $P[Z \geq 2.5] = 0.006$이르로 유의수준 0.006에서 검정을 해도 귀무가설 H_0를 기각하게 되며 유의수준 0.006미만일 경우에는 기각할 수 없다. 이와 같이 검정통계량의 값에 대하여 귀무가설을 기각할 수 있는 최소의 유의수준을 p-값(p-value) 또는 유의확률 (Significance probability)이라고 한다.

가설검정에서 유의수준과 p-값 사이에는 약간의 개념 차가 있다. 위의 예에서 유의수준 $\alpha = 0.05$로 검정할 때, $\bar{x} = 1,550$을 얻은 경우와 $\bar{x} = 1,600$을 얻은 경우 모두 귀무가설 H_0을 기각하게 되지마 기각하게 되는 정도는 다르다. $\bar{x}$의 값이 커질수록 p-값은 작아지며 H_1이 참이라는 확신이 커지게 된다. 이와 같이 p-값은 대립가설의 채택에 대한 확신의 정도를 보여주며 대부분의 통계 패키지에서도 p-값을 출력시키고 있다.

2. 분산의 검정과 평균의 차에 대한 검정

※ $\chi^2(\mathrm{k})$ 분포



<예 7.2-1> 분산의 검정

어느 대학교 학생들의 통계학 시험 점수에 대한 분산은 $\sigma^2 = 100$이라고 알려져 있다. $\sigma^2 = 100$인지 아닌지를 알아보기 위하여 학생 23명을 랜덤하게 선택하여 점수를 조사한 결과 $s^2 = 147.82$를 얻었다. 유의수준 0.05에서 검정하라. (통계학의 점수는 정규분포를 한다고 가정한다.)

<풀이> $\dfrac{nS^2}{\sigma^2}$의 분포는 자유도가 $n-1$인 카이제곱분포 $\chi^2(n-1)$이므로 카이제곱분표에서

$$P\left[\frac{23S^2}{.100} < 10.98 : \sigma^2 = 100\right] = 0.025,$$

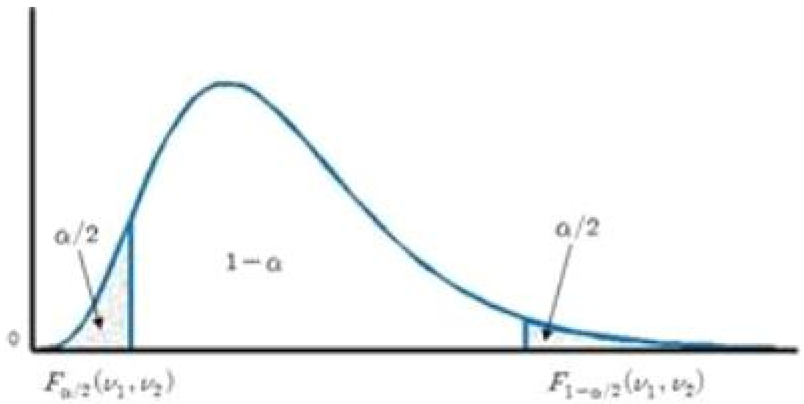
$$P\left[\frac{23S^2}{100} > 38.78 : \sigma^2 = 100\right] = 0.025$$

이다. 즉, $s^2 < \dfrac{(10.98)(100)}{23} = 47.74$ 또는 $s^2 > \dfrac{(36.78)(100)}{23} = 159.91$이면 귀무가설 H_0를 기각하게 되나 표본분산의 측정값이 $S^2 = 147.82$이므로 귀무가설은 기각되지 않는다. 참고로 σ^2의 95%신뢰구간

$$\left[\frac{(23)(147.82)}{36.78},\ \frac{(23)(147.82)}{10.98}\right] = [92.44,\ 309.64]$$

는 $\sigma^2 = 100$을 포함하고 있다.

<예 7.2-2> 분산의 비에 대한 검정



거미를 연구하는 생물학자는 암컷의 크기 변화가 수컷의 크기 변화보다 더 크다는 연구 결과가 참인지를 알고 싶었다. 암컷을 16마리 채집하여 분산 $S_X^2=2.26$을 얻었고 수컷 11마리를 채집하여 분산 $S_Y^2=0.25$를 얻었다. 암컷의 크기 X의 분포는 정규분포 $N(\mu_X,\sigma_X^2)$이고 수컷의 크기 Y의 분포는 정규분포 $N(\mu_Y,\sigma_Y^2)$이라고 가정하고 대립가설 $H_1:\frac{\sigma_X^2}{\sigma_Y^2}>1$에 대하여 귀무가설 $H_0:\frac{\sigma_X^2}{\sigma_Y^2}=1$을 유의수준 0.05로 검정하라.

<풀이> $\frac{mS_X^2}{\sigma_X^2}$와 $\frac{mS_Y^2}{\sigma_Y^2}$는 각각 서로 독립인 카이제곱분포 $\chi^2(m-1)$, $\chi^2(n-1)$을 따르므로 귀무가설 H_0가 참일 때, 검정 통계량의 분포는

$$F=\frac{mS_X^2/(m-1)}{nS_Y^2/(n-1)}\sim F(m-1,\ n-1)$$

이다. $\alpha=0.05$에서 F-분포표로부터

$$P\left[\frac{16S_X^2/15}{11S_Y^2/10}>2.85:\ \frac{\sigma_X^2}{\sigma_Y^2}=1\right]=0.05,$$

이며 $\frac{s_X^2}{s_Y^2}=\frac{2.26}{0.25}=9.04>\left(\frac{11}{10}\right)\left(\frac{15}{16}\right)(2.85)=2.94$이므로 귀무가설 H_0는 기각된다.

<예 7.2-3> 평균의 차에 대한 검정

48시간 동안 남녀의 모발이 자라는 길이를 측정하여 차이가 있는지 알고 싶다. 남자의 모발이 자라는 길이를 X라 하고 여자의 모발이 자라는 길이를 Y라 하자. X의 분포는 정규분포 $N(\mu_X, \sigma^2)$이고 Y의 분포는 정규분포 $N(\mu_Y, \sigma^2)$이라고 가정한다. 남자 11명을 랜덤하게 선택하여 모발의 성장길이를 측정한 결과 0.8, 1.8, 1.0, 0.1, 0.9, 1.7, 1.0, 1.4, 0.9, 1.2, 0.5를 얻었고, 여자 13명을 랜덤하게 선택하여 모발의 성장길이를 측정한 결과 1.0, 0.8, 1.6, 2.6, 1.3, 1.1, 2.4, 1.8, 2.5, 1.4, 1.9, 2.0, 1.2를 얻었다. 대립가설 $H_1 : \mu_X - \mu_Y < 0$에 대하여 귀무가설 $H_0 : \mu_X - \mu_Y = 0$를 유의수준 $\alpha = 0.01$에서 검정하라.

<풀이> 귀무가설 H_0가 참일 때, 검정 통계량의 분포는

$$T = \frac{\overline{X} - \overline{Y}}{S_p\sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t(m+n-2),$$

$$S_p^2 = \frac{mS_X^2 + nS_Y^2}{m+n-2} \quad \text{(합동표본분산)}$$

이고, 위의 자료에서 $\overline{x} = 1.03$, $s_X^2 = 0.22$, $\overline{y} = 1.66$, $s_Y^2 = 0.33$이므로 t-분포표로부터

$$T = \frac{1.03 - 1.66}{\sqrt{\frac{11(0.22) + 13(0.33)}{11+13-2}\left(\frac{1}{11} + \frac{1}{13}\right)}} = -2.78 < t_{0.01}(22) = -2.508$$

이다. 따라서 귀무가설 H_0는 유의수준 $\alpha = 0.01$에서 기각된다.

3. 검정력 함수

제1종의 오류: 귀무가설 H_0가 참일 때, 귀무가설을 기각하게 되는 오류

제2종의 오류: 대립가설 H_1이 참일 때, 귀무가설을 채택하게 되는 오류

검정결과 \ 사실	귀무가설이 참	대립가설이 참
귀무가설을 채택	옳은 결정	제2종의 오류
귀무가설을 기각	제1종의 오류	옳은 결정

이제 기각역을 C라고 표기하면 제1종의 오류를 범할 확률을

$$\alpha = P[(X_1, \cdots, X_n) \in C \mid H_0]$$

으로 나타낼 수 있고 이는 유의수준과 같다. 제2종의 오류를 범할 확률은

$$\beta = P[(X_1, \cdots, X_n) \in C' \mid H_1]$$

으로 표현되며 C'은 채택역이 된다.

<예 7.3-1> 제철소에서 만들어지는 강철의 강도를 X라고 할 때, 강철이 용광로 I에서 만들어지면 X의 분포는 정규분포 $N(50, 36)$이고, 용광로 II에서 만들어지면 X의 분포는 정규분포 $N(55, 36)$이라고 한다. $n = 16$개의 강철의 강도를 측정하여 기각역 $C = \{(x_1, \cdots, x_n) : \bar{x} > 53\}$이라고 할 때, 제 1종의 오류를 범할 확률과 범할 확률과 제 2종의 오류를 범할 확률을 구하라.

<풀이> 귀무가설은 $H_0 : \mu = 50$이고 대립가설은 $H_0 : \mu = 55$가 된다. 이제 $\overline{X}$를 검정 통계량이라 하고 기각역 $C = \{(x_1, \cdots, x_n) : \bar{x} > 53\}$이라고 하자. $n = 16$개의 강철의 평균 강도 $\overline{X}$는 귀무가설 H_0가 참일 때는 정규분포 $N(50, 36/16)$을 따르며, 대립가설이 참일 때는 정규분포 $N(55, 36/16)$을 따른다. 따라서 제1종의 오류를 범할 확률은

$$\alpha = P[\overline{X} > 53 : H_0] = P\left[\frac{\overline{X} - 50}{6/4} > \frac{53 - 50}{6/4} \ : H_0\right] = 1 - \Phi(2) = 0.0228$$

이며 제2종의 오류를 범할 확률은

$$\beta = P[\overline{X} \le 53 : H_1] = P\left[\frac{\overline{X} - 55}{6/4} \le \frac{53 - 55}{6/4} \ : H_0\right] = \Phi\left(-\frac{4}{3}\right) = 1 - 0.9087 = 0.0913$$

이다. 참고로 β의 값이 커지면 α의 값은 줄어들며 표본의 크기 n이 증가하면 α와 β는 감소한다.

<정의 7.3-1> (a) 가설검정의 과정에서 귀무가설 H_0를 기각할 확률을 모수 θ의 함수로 표현한 것을 검정력 함수 (Power function)라고 하며 아래와 같은 식으로 표현된다.

$$K(\theta) = P[H_0\text{를 기각} : \theta].$$

(b) 대립가설의 영역에 있는 θ에 대한 검정력 함수의 값을 검정력 (Power)이라고 한다.

<예 7.3-2> $X_1, X_2, \cdots, X_{10}$은 밀도함수가 $f(x:p) = p^x(1-p)^{1-x}$, $x = 0, 1$인 베르누이 시행에서의 확률표본이며 $0 < p \le 1/2$임을 알고 있다. 대립가설 $H_1 : p < 1/2$에 대하여 귀무가설 $H_0 : p = 1/2$를 검정하고자 한다. 검정통계량으로써 $Y = \sum_{i=1}^{10} X_i$를 사용하고자 하며 기각역을

$$C = \{y : y = 0,\ 1,\ 2\} = \left\{(x_1,\ \cdots, x_n) : \sum_{i=1}^{10} x_i = 0, 1, 2\right\}$$

이라고 할 때, 제1종의 오류를 범할 확률, 제2종의 오류를 범할 확률 및 검정력 함수를 구하라.

<풀이> 귀무가설 H_0가 참일 때, Y의 분포는 이항분포 $B(10, 1/2)$이다. 따라서 제1종의 오류를 범할 확률은

$$\alpha = P\left[Y = 0, 1, 2 : p = \frac{1}{2}\right]$$

이며 대립가설 $H_1 : p < 1/2$가 참일 때 $p = 1/4$라고 하면 제2종의 오류를 범할 확률은

$$\beta = P\left[3 \le Y \le 10 : p = \frac{1}{4}\right] = \sum_{y=3}^{10} \binom{10}{y} \left(\frac{1}{4}\right)^y \left(\frac{3}{4}\right)^{10-y} = 0.4744$$

이며 $p = 1/10$인 경우 제2종의 오류를 범할 확률은

$$\beta \equiv \sum_{y=3}^{10} \binom{10}{y} \left(\frac{1}{10}\right)^y \left(\frac{9}{10}\right)^{10-y} = 0.0702$$

가 된다. 검정력 함수를 구하면 아래와 같다.

$$K(p) \equiv \sum_{y=0}^{2} \binom{10}{y} p^y (1-p)^{10-y} = (1-p)^8(1 + 8p + 36p^2).$$

[참고] 검정력 함수를 $K(\theta)$라고 하면 유의수준은 θ가 귀무가설의 영역에 있을 때 $K(\theta)$의 최댓값이다. θ_1이 대립가설 H_1의 영역에 있을 때, 제2종의 오류를 범할 확률은 $\beta(\theta_1) = 1 - K(\theta_1)$이다.

<예 7.3-3> 확률변수 X의 분포가 정규분포 $N(\mu, 36)$일 때, 대립가설 $H_1 : \mu > 50$에 대하여 귀무가설 $H_0 : \mu = 50$을 검정하고자 한다. 기각역을 $C = \{(x_1, \cdots, x_n) : \overline{x} > c\}$라 하고 $K(\mu) = P[\overline{X} > c : \mu]$로 정의할 때, $K(50) = 0.05$, $K(55) = 0.90$이 되도록 c와 n을 결정하라.

<풀이> $0.05 = P[\overline{X} > c : H_0]$

$0.90 = P[\overline{X} > c : H_1]$ (혹은 $0.10 = P[\overline{X} \le c : H_1]$)

위의 두 식을 연립하여 풀어서 c와 n을 구하여야 한다. 귀무가설 H_0에서 $\overline{X} \sim N\left(50, \frac{30}{n}\right)$이므로 표준정규분포표로부터 첫 번째 식은

$$P\left[\frac{\overline{X} - 50}{6/\sqrt{n}} > \frac{c-50}{6/\sqrt{n}}\right] = 0.05 \Rightarrow \frac{c-50}{6/\sqrt{n}} = 1.645$$

두 번째 식은

$$P\left[\frac{\overline{X} - 55}{6/\sqrt{n}} \le \frac{c-55}{6/\sqrt{n}}\right] = 0.10 \Rightarrow \frac{c-55}{6/\sqrt{n}} = -1.282$$

이므로 $c = 52.8$, $n = 12$ 혹은 13을 얻는다.

<예 7.3-4> 확률변수 X의 분포가 이항분포 $B(n, p)$일 때, Y의 측정값 y가 상수 c보다 작으면 귀무가설 $H_0 : p = 1/2$을 기각하고 대립가설 $H_1 : p < 1/2$을 채택한다. $K(p) = P[Y < c : p]$라고 정의할 때, $K(1/2) = 0.05$, $K(1/4) = 0.90$을 만족하는 c와 n을 구하라.

<풀이>

$$0.05 = P\left[Y < c : p = \frac{1}{2}\right] = P\left[\frac{Y - \frac{n}{2}}{\sqrt{n \cdot \frac{1}{2} \cdot \frac{1}{2}}} < \frac{c - \frac{n}{2}}{\sqrt{n \cdot \frac{1}{2} \cdot \frac{1}{2}}}\right]$$

$$\Rightarrow \frac{c - \frac{n}{2}}{\sqrt{\frac{n}{4}}} = -1.645.$$

$$0.90 = P\left[Y < c : p = \frac{1}{4}\right] = P\left[\frac{Y - \frac{n}{4}}{\sqrt{n \cdot \frac{1}{4} \cdot \frac{3}{4}}} < \frac{c - \frac{n}{4}}{\sqrt{n \cdot \frac{1}{4} \cdot \frac{3}{4}}}\right]$$

$$\Rightarrow \frac{c - \frac{n}{4}}{\sqrt{\frac{3n}{16}}} = 1.282.$$

위의 두 식을 연립하여 c와 n을 구하면 $c = 10.9$, $n = 31$이다.