9.1 가설검정(Testing Hypotheses)

§. 귀무가설과 대립가설 (The Null and Alternative Hypotheses)

<u>Definition 9.1.1</u> **귀무가설과 대립가설 / 기각**
통계적 가설의 종류로 H_0 : **귀무가설**(null hypothesis), H_1 : **대립가설**(alternative hypothesis)라고 부른다. 그리고 모수공간 $\Omega = \Omega_0 \cup \Omega_1$, $\Omega_0 \cap \Omega_1 = \varnothing$에 대하여 모수가 $\theta \in \Omega_1$이면 H_0을 **기각**(reject) **한다**고 하고, $\theta \in \Omega_0$이면 H_0을 **기각할 수 없다**(not to reject)고 말한다.

▷ **예제 9.1.2** : Egyptian Skulls
다양한 기간동안 이집트인들의 여러 두개골 측정 자료가 보고되었다고 한다. 그 중 어느 기간은 거의 기원전 4000년경에 측정되었다. 우리는 관찰된 두개골 너비의 측정값이 정규(normal) 확률변수로써 분산σ가 26이고 μ는 알려지지 않았다고 모형화하였다. 여기서 관심있는 것은 오늘날의 두개골 너비 측정값(140mm)과 어떻게 비교할 것인지 파악하고자 한다.

모수공간을 $\Omega = R^{+}$라 두고, $\Omega_0 = [140, \infty)$, $\Omega_1 = (0, 140)$이라 하자. 이 경우, 우리는 다음과 같이 가설을 세울 것이다.

$$H_0 : \mu \geq 140$$
$$H_1 : \mu < 140$$

좀 더 구체적으로, 우리는 알려지지 않은 두개골 측정값의 평균과 분산을 가정하였다고 하자. 이는 각 측정치가 μ와 σ^2를 갖는 정규확률변수라고 하는 것과 같다. 모수는 2차원 벡터로써 $\theta = (\mu, \sigma^2)$를 갖는다. 그러면 우린 오직 μ에 대해 관심 있으므로, $\Omega_0 = [140, \infty) \times (0, \infty)$, $\Omega_1 = (0, 140) \times (0, \infty)$ 이다.
이제 우린 더 다양한 이론들을 살펴보면서 이 예제와 더불어 가설검정이 무엇인지 파악할 것이다.

▲

<u>Definition 9.1.2</u> **단순가설과 복합가설**
$\theta \in \Omega_i$ 에 대하여 $\dim(\Omega_i) = 1$ 이면 H_i 는 **단순가설**(simple hypothesis) 이라고 부른다. 만약 $\dim(\Omega_i) > 1$ 이면 H_i 는 **복합가설**(composite hypothesis) 이라고 부른다.

<u>Definition 9.1.3</u> **단측가설과 양측가설**
$\dim(\Omega) = 1$ 이라고 하자. $H_0 : \theta \leq \theta_0$ 또는 $H_0 : \theta \geq \theta_0$인 경우 각각 $H_1 : \theta > \theta_0$ 또는 $H_1 : \theta < \theta_0$이면 **단측가설**(One-sided) 이라고 부른다. $H_0 : \theta = \theta_0$ 인 경우 $H_1 : \theta \neq \theta_0$ 이면 **양측가설**(Two-sided) 이라고 부른다.

§. 임계역 및 기각역과 검정통계량 (Critical / Reject Region and Test Statistics)

<u>Definition 9.1.4</u> **임계역**

$\overrightarrow{X}=(X_1,\ldots,X_n)$을 알려지지 않은 모수 θ를 갖는 분포로부터 취해진 확률표본이라 가정하자. 확률표본에서 얻는 데이터 벡터 $\overrightarrow{x}=(x_1,\ldots,x_n)$, 통계량 T에 대하여 $S_0=\{\overrightarrow{x}:-c\le T\le c\}$, $S_1=S_0^C$ 이라 정의하면 S_1을 **임계역(Critical Region)** 이라고 말한다. 이 때 상수 c는 선택될 수 있다.

▷ **예제** 9.1.3 : Testing Hypotheses about Mean of a Normal Distribution with Known Variance.

위의 정의와 마찬가지로 가정하자. 대신 모수가 $\theta=(\mu,\sigma^2)$인 정규분포를 갖는다고 가정하자. 다음과 같이 가설을 설정하고 이를 살펴보고자 한다.

$$H_0:\mu=\mu_0$$
$$H_1:\mu\ne\mu_0$$

우리는 $\overline{X_n}$이 μ_0과 얼마나 멀리 떨어져 있는지에 따라 H_0을 기각하고자 한다. 예를 들어, 상수 c를 선택하여 $\overline{X_n}$과 μ_0의 거리가 c보다 큰 경우 H_0을 기각할 것이다. 이를 표현하는 하나의 방법으로는 모든 가능한 데이터 벡터 $\overrightarrow{x}=(x_1,\ldots,x_n)$의 집합인 표본공간 S를 다음과 같이 두 개의 집합으로 분할하는 것이다.

$$S_0=\{\overrightarrow{x}:-c\le\overline{X_n}-\mu_0\le c\},\ S_1=S_0^C$$

그러면 우린 $\overrightarrow{X}\in S_1$일 때 H_0을 기각하고, $\overrightarrow{X}\in S_0$일 때 H_0을 기각할 수 없다. 이 과정을 표현하는 간단한 방법은 통계량 $T=|\overline{X}-\mu_0|$를 정의하고 $T\ge c$일 때 H_0을 기각한다고 정의하는 것이다.

▲

<u>Definition 9.1.5</u> **검정통계량 / 기각역**

$\overrightarrow{X}$가 모수 θ에 의존하는 분포로부터 취해진 확률표본이라 하자. $T=r(\overrightarrow{X})$는 통계량이고, R은 실직선의 부분집합이라 하자. 가설 $H_0:\theta\in\Omega_0$, $H_1:\theta\in\Omega_1$에 대한 검정절차에서 명제 "$T\in R$일 때 H_0을 기각한다."를 만족하면 T 를 **검정통계량(test statistic)** 이라 부르고, R 을 검정의 **기각역(rejection region)** 이라 부른다.

▷ **예제** 9.1.4 : Rain from Seeded Clouds.

예제 9.1.1의 내용을 가정하자. 우린 26개의 seeded cloud가 정규확률변수로써 알려지지 않은 모수 $\theta=(\mu,\sigma^2)$를 갖는 log-rainfalls를 모형화 하였다. 우린 어찌 되었든 $\mu>4$인 것에 관심이 있는 상태이다. 이를 모수벡터의 관점에서 서술하면, $\{(\mu,\sigma^2):\mu>4\}$에 θ가 들어가도록 하는 것이다. 이제 가설을 세우면 $H_0:\mu\le 4$ vs $H_1:\mu>4$ 이다. 우린 검정통계량으로 이전 예제와 같은 것을 사용할 수도 있지만, $U=n^{1/2}(\overline{X_n}-4)/\sigma'$을 사용하고자 한다. 이 경우에 U가 크다면 H_0을 기각할 수 있다. 이는 U에서 대응되는 $\overline{X_n}$이 4와 비교해서 커지기 때문이다.

▲

§. 검출력 함수와 제 1종, 2종 오류 (The Power Function and Types of Error)

<u>Definition 9.1.6</u> **검출력 함수**
δ를 검정절차라고 하자. 그러면 함수 $\pi(\theta|\delta)$를 검정 δ의 **검출력 함수(Power function)**라고 부른다. 만약 S_1이 δ의 임계역이면, 검출력함수 $\pi(\theta|\delta)$는 다음의 관계에 의해 결정된다.

$$\pi(\theta|\delta) = \Pr(\overrightarrow{X} \in S_1|\theta) \text{ for } \theta \in \Omega$$

만약 δ가 검정통계량 T와 기각역 R로 표현된다면, 검출력 함수는 다음의 관계에 의해 결정된다.

$$\pi(\theta|\delta) = \Pr(T \in R|\theta) \text{ for } \theta \in \Omega$$

▷ **예제** 9.1.5 : Testing Hypotheses about Mean of a Normal Distribution with Known Variance.
예제 9.1.3의 경우에서, 검정 δ는 검정통계량 $T = \left|\overline{X_n} - \mu_0\right|$과 기각역 $R = [c, \infty)$을 기반으로 하고 있다. $\overline{X_n}$의 분포는 평균 μ와 분산 σ^2/n을 갖는 정규분포를 따른다. μ는 우리가 σ^2를 알고 있다고 가정하였기 때문에 이 문제에서의 모수가 된다. 검출력 함수는 이러한 정규분포에서 계산될 수 있다, 우선 Φ를 표준 정규분포의 누적분포함수(cdf)라 하자. 그러면 검출력 함수는 다음과 같이 구해진다.

$$\begin{aligned}\pi(\mu|\delta) &= \Pr(T \in R|\mu) = \Pr(\overline{X_n} \geq \mu_0 + c|\mu) + \Pr(\overline{X_n} \leq \mu_0 - c|\mu) \\ &= 1 - \Phi\left(n^{1/2}\frac{\mu_0 + c - \mu}{\sigma}\right) + \Phi\left(n^{1/2}\frac{\mu_0 - c - \mu}{\sigma}\right)\end{aligned}$$

Figure 9.1.은 $c = 1,2,3$인 경우의 검정에 대한 검출력 함수를 나타낸 것이다. 이 때, $\mu_0 = 4$, $n = 15$, $\sigma^2 = 9$ 로 두고 구하였다.

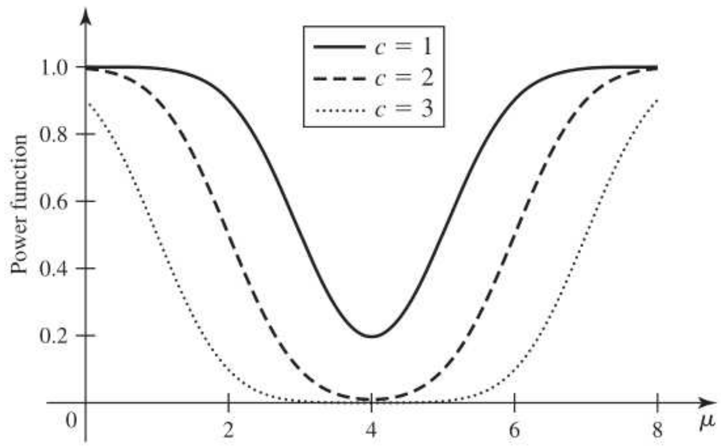

Figure 9.1. Power functions of three different tests in Example 9.1.5.

▲

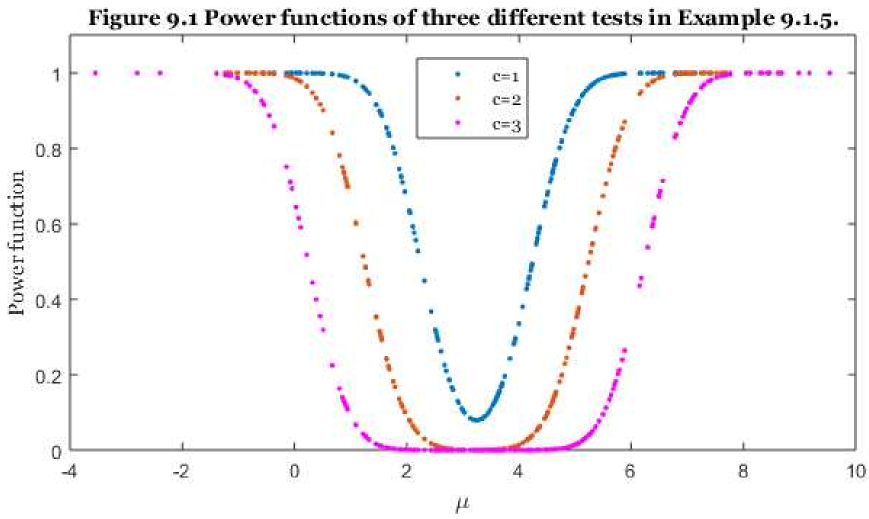

Figure 9.1. Power functions of three different tests using MATLAB.

<u>MATLAB CODE – statistics2powerfunction.m</u>

```
clear all
clc

% random samples & coefficients
samples=normrnd(4,3,1,250);
n=length(samples);
c=[1 2 3];
mu0 = 4;
sigma = 3;
% Pr(sample mean >= mu0+c | mean) and in case of <=.
T1_c1=sqrt(n)*(mu0+c(1)-samples)./sigma;
T2_c1=sqrt(n)*(mu0-c(1)-samples)./sigma;
T1_c2=sqrt(n)*(mu0+c(2)-samples)./sigma;
T2_c2=sqrt(n)*(mu0-c(2)-samples)./sigma;
T1_c3=sqrt(n)*(mu0+c(3)-samples)./sigma;
T2_c3=sqrt(n)*(mu0-c(3)-samples)./sigma;
% c.d.f. of normal distribution (not standard)
p1_1 = normcdf(T1_c1,mu0,sigma);
p2_1 = normcdf(T2_c1,mu0,sigma);
p1_2 = normcdf(T1_c2,mu0,sigma);
p2_2 = normcdf(T2_c2,mu0,sigma);
p1_3 = normcdf(T1_c3,mu0,sigma);
p2_3 = normcdf(T2_c3,mu0,sigma);
% Probability of T is  in R given mu (Power func.)
P1 = 1-p1_1+p2_1;
P2 = 1-p1_2+p2_2;
P3 = 1-p1_3+p2_3;
% Graph of the power functions
plot(samples,P1,'.');
hold on
plot(samples,P2,'.');
plot(samples,P3,'.m');
title('Figure 9.1 Power functions of three different tests in Example 9.1.5.');
xlabel('\mu');
ylabel('Power function');
legend('c=1','c=2','c=3');
xlim([-4 10]);
ylim([0 1.1]);
```

Definition 9.1.7 제 1종 오류와 제 2종 오류
귀무가설이 참일 때 이를 기각하는 것을 **제 1종 오류(type I error)**라 말하고, 귀무가설이 거짓인데 이를 기각하지 않는 것을 **제 2종 오류(type II error)**라고 말한다.

▷ **예제 9.1.6 : Egyptian Skulls.**
예제 9.1.2에서, 실험자들이 두개골 너비가 오랜 시간동안 증가해왔음을 언급하였다. 만약 μ가 기원전 4000년경의 두개골 너비의 평균이고, 140이 오늘날의 두개골 너비의 평균이라고 하면, 이론대로라면 $\mu < 140$이다. 실험자들은 사실 $\mu > 140$이거나 $\mu < 140$임에도 잘못된 주장으로 가지고 있는 데이터가 각각 그 반대의 경우로 주장할 수 있다. 이러한 경우를 각각 살펴보면, 첫 번째의 경우는 귀무가설이 참임에도 그와 반대를 주장함으로써 제 1종 오류를 범하게 되고, 두 번째의 경우는 귀무가설이 거짓인데도 불구하고 그 주장을 참이라 여기는 경우이므로 제 2종 오류를 범하게 된다. (예제 9.1.2에서 $H_0 : \mu \geq 140$, $H_1 : \mu < 140$ 이었다.)

▲

Definition 9.1.8 수준 / 크기
어떤 검정이 $\pi(\theta|\delta) \leq \alpha_0$ for all $\theta \in \Omega_0$, $\alpha_0 \in [0,1]$ 을 만족하면 이 검정을 **수준(level)** α_0 **검정**이라 말하고, 이러한 검정은 α_0 **유의수준(level of significance)**을 갖는다고 한다. 또한, **검정 δ의 크기 (size)** $\alpha(\delta)$는 다음과 같이 정의된다.

$$\alpha(\delta) = sup_{\theta \in \Omega_0} \pi(\theta|\delta)$$

▷ Corollary 9.1.1 ① 검정 δ가 수준 α_0 검정 $\Leftrightarrow$ $\alpha(\delta) \leq \alpha_0$. ② $H_0 : \theta = \theta_0$ $\Rightarrow$ $\alpha(\delta) = \pi(\theta_0|\delta)$.

▷ **예제 9.1.7 : Testing Hypotheses about a Uniform Distribution.**
확률표본 $X_1, \ldots, X_n$이 구간 $[0,\theta]$인 균등분포로부터 취해진다고 가정하고, 모수 $\theta > 0$는 알려지지 않았다고 가정하자. 또한 다음의 가설을 검정한다고 가정하자.

$$H_0 : 3 \leq \theta \leq 4$$
$$H_1 : \theta < 3 \text{ or } \theta > 4$$

우린 이전의 예제로부터 θ의 최대우도추정량은 $Y_n = \max\{X_1, \ldots, X_n\}$임을 알고 있다. 비록 Y_n이 반드시 θ보다 작을지라도, Y_n은 표본크기 n이 상당히 커진다면 θ에 가까워질 높은 확률을 가지고 있다. 검정 δ가 $2.9 < Y_n < 4$인 경우 H_0을 기각하지 않는다고 하고, δ는 Y_n가 앞선 구간에서 있지 않을 때 H_0을 기각한다고 가정하자. 그러면 이 때 검정 δ의 임계역은 $Y_n \leq 2.9$ or $Y_n \geq 4$에 속하는 모든 $X_1, \ldots, X_n$을 포함한다. 이 때, 검정통계량 Y_n의 관점에서 기각역은 $(-\infty, 2.9] \cup [4, \infty)$이 된다.

δ의 검출력함수는 다음 관계로 규정된다.

$$\pi(\theta|\delta) = \Pr(Y_n \leq 2.9|\theta) + \Pr(Y_n \geq 4|\theta)$$

만약 $\theta \leq 2.9$이면, $\Pr(Y_n \leq 2.0|\theta) = 1$이고 $\Pr(Y_n \geq 4|\theta) = 0$이므로, $\theta \leq 2.9$인 경우 $\pi(\theta|\delta) = 1$이다.

만약 $2.9 < \theta \le 4$이면, $\Pr(Y_n \le 2.9|\theta) = (2.9/\theta)^n$이고 $\Pr(Y_n \ge 4|\theta) = 0$이다. 이 경우 검출력함수는 $\pi(\theta|\delta) = (2.9/\theta)^n$이 된다. 마지막으로, $\theta > 4$이면 $\Pr(Y_n \le 2.9|\theta) = (2.9/\theta)^n$이고 $\Pr(Y_n \ge 4|\theta) = 1-(4/\theta)^n$이다. 이 경우, $\pi(\theta|\delta) = (2.9/\theta)^n + 1 - (4/\theta)^n$이 된다. 이러한 검출력함수의 그래프는 Figure 9.2에서 확인할 수 있다.

식 $\alpha(\delta) = sup_{\theta \in \Omega_0} \pi(\theta|\delta)$에 의하여, δ의 크기는 $\alpha(\delta) = sup_{3 \le \theta \le 4} \pi(\theta|\delta)$이 된다. 이는 Figure 9.2에서 확인할 수 있고, 그 계산은 단지 $\alpha(\delta) = \pi(3|\delta) = (29/30)^n$ 으로 주어진다. 특히, 만약 표본크기가 $n = 68$이면, δ의 크기는 $(29/30)^{68} = 0.0997$이 된다. 따라서, δ는 유의수준$\alpha_0 \ge 0.0997$인 수준 α_0 검정이다.

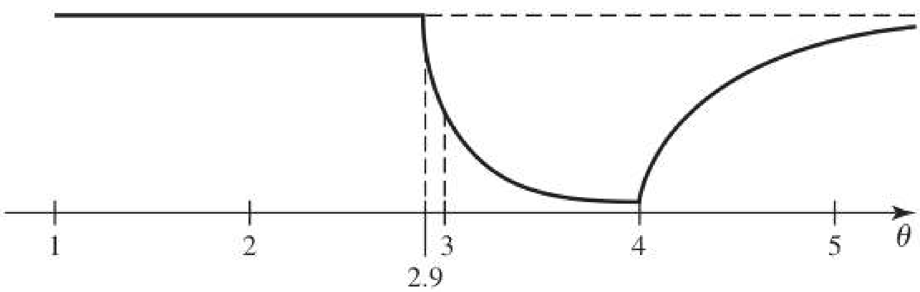

Figure 9.2. The power functions $\pi(\theta|\delta)$ in Example 9.1.7.

▲

▷ **예제 9.1.8 : Testing Hypotheses about Mean of a Normal Distribution with Known Variance.**
이전 예제 9.1.5에서 다룬 검정은, $\left|\overline{X_n}-\mu_0\right| \geq c$ 인 경우 $H_0 : \mu=\mu_0$ 을 기각하는 것이었다. $Y=\overline{X_n}-\mu_0$ 가 $\mu=\mu_0$인 경우 평균이 0이고 분산이 σ^2/n 인 정규분포를 가지므로, 우리는 각 α_0에 대하여 정확히 α_0 의 크기를 가지도록 만들 값 c를 찾을 수 있다. Figure 9.3은 Y의 확률밀도함수(pdf)를 보여주고 있고, 검정의 크기도 밀도함수곡선 아래로 색칠된 영역으로 표시되어 있음을 확인할 수 있다.
이 정규확률밀도함수가 평균에서 대칭성을 가진다는 사실에 의하여, 두 개의 색칠된 영역들은 같은 넓이, 즉 $\alpha_0/2$를 갖는다. 이것은 c가 반드시 Y의 분포의 $1-\alpha_0/2$ 분위수에 위치한다는 것을 의미한다. 이러한 분위수는 $c=\Phi^{-1}(1-\alpha_0/2)\sigma n^{-1/2}$ 로 구해진다.

정규분포의 평균에 대한 가설 검정을 할 때, 이는 관습적으로 다음과 같이 통계량의 관점에서 이 검정을 재 서술할 수 있다.

$$Z=n^{1/2}\frac{\overline{X_n}-\mu_0}{\sigma}$$

그러면 이 검정은 $|Z| \geq \Phi^{-1}(1-\alpha_0/2)$ 일 때 H_0을 기각한다.

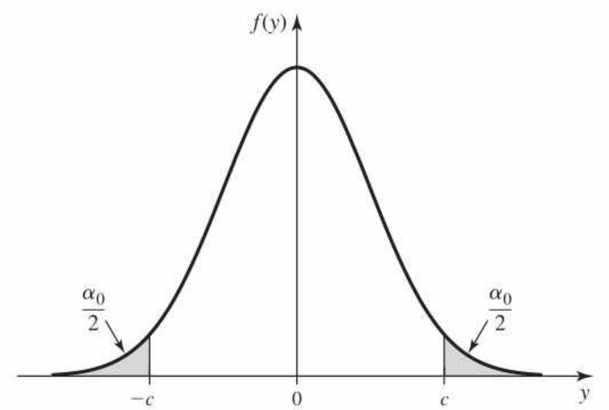

Figure 9.3. The p.d.f. of $Y=\overline{X_n}-\mu_0$ given $\mu=\mu_0$ for Example 9.1.8. The shaded areas represent the probility that $|Y| \geq c$.

▲

▷ **예제** 9.1.9 : Testing Hypotheses about a Bernoulli Parameter.

$X_1, \ldots, X_n$이 모수가 p인 베르누이 분포로부터 확률표본을 생성한다고 가정하자. 우리는 다음 가설에 대한 검정을 하길 원한다.

$$H_0 : p \le p_0$$
$$H_1 : p > p_0$$

$Y = \sum_{i=1}^{n} X_i$라 하자, 이는 $Y \sim Bin(n,p)$이다. 우리는 $p > p_0$임을 주장하고자 하므로 우린 $Y \ge c$인 경우 H_0을 기각하는 것을 택한다. 또한 우리는 가능한 한 α_0을 넘지 않고 α_0에 가까워지도록 하는 검정의 크기를 원한다고 하자. $\Pr(Y \ge c|p)$가 p의 함수로써 증가함수임은 쉽게 알 수 있고, 이런 이유로 검정의 크기는 $\Pr(Y \ge c|p = p_0)$ 이 된다. 그래서 c는 $\Pr(Y \ge c|p = p_0) \le \alpha_0$을 만족하는 가장 작은 수가 되어야 한다.

예를 들어, 만약 $n = 10$, $p_0 = 0.3$, $\alpha_0 = 0.1$이라 하면, 우린 c를 결정하기 위해 교재의 부록의 이항 확률표를 사용할 수 있다. 또한 $\sum_{y=6}^{10} \Pr(Y = y|p = 0.3) = 0.0473$, $\sum_{y=5}^{10} \Pr(Y = y|p = 0.3) = 0.1503$ 으로 계산할 수 있다. 검정의 크기를 많아야 0.1로 유지하기 위해서, 우린 반드시 $c > 5$가 되도록 해야 한다. 구간 $(5, 6]$의 모든 c값은, Y가 오직 정수값만 취하므로 같은 검정을 만들어내게 된다.

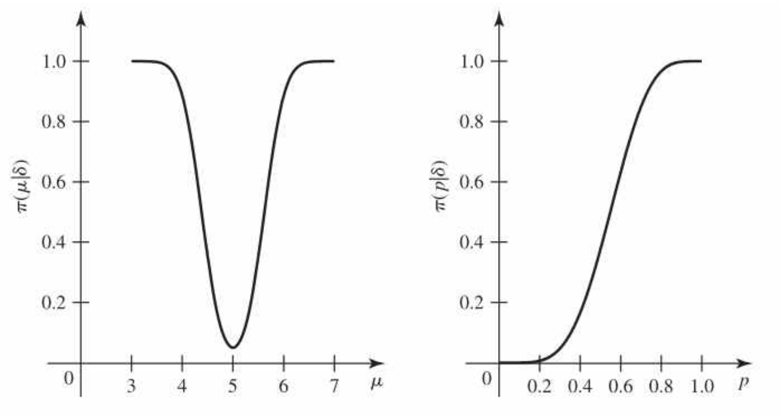

Figure 9.4. Power functions of two tests. The plot on the left is the power function of the test from Example 9.1.8 with $n = 10$, $\mu_0 = 5$, $\sigma = 1$, and $\alpha_0 = 0.05$. The plot on the right is the power function of the test from Example 9.1.9 with $n = 10$, $p_0 = 0.3$, and $\alpha_0 = 0.1$

▲

§. p-값 (The p-value)

Definition 9.1.9 p-값 (유의확률, significance probability / p-value)
일반적으로, **p-값 (p-value)**은 주어진 검정통계량의 값으로부터 H_0을 기각하게 하는 최소의 유의수준 α_0을 말한다. 이는 귀무가설이 참일 때 표본이 대립가설 방향으로 검정통계량의 값보다 더 어긋나게 될 확률을 의미한다.

▷ **[NOTE]** p-value $\le \alpha_0$: 유의수준 α_0에서 H_0을 기각한다. (통계적으로 유의하다)
p-value $> \alpha_0$: 유의수준 α_0에서 H_0을 기각할 수 없다. (통계적으로 유의하지 않다)

▷ **[NOTE] 가설검정의 절차**

1. 귀무가설과 대립가설을 설정한다.
2. 주어진 문제의 특성에 따라 유의수준 α_0을 결정한다.
3. 표본 데이터로부터 검정통계량을 계산한다.
4. 가설검정의 결론 도출
   (i) 기각역 사용
   - 유의수준에 따라 기각역을 구한다.
   - 검정통계량이 기각역에 속하면 H_0을 기각한다.

   (ii) p-value 사용
   - 검정통계량을 이용하여 p-value를 구한다.
   - $p-value \le \alpha_0$이면 H_0을 기각한다.

▷ **예제 9.1.10 : Testing Hypotheses about Mean of a Normal Distribution with Known Variance.**
앞선 예제 9.1.8에서, 우린 수준 $\alpha_0 = 0.05$ 에서의 귀무가설을 선택하였다. 이제 우린 검정통계량 $Z = n^{1/2}\dfrac{\overline{X_n} - \mu_0}{\sigma}$ 을 계산하고 $Z \ge \Phi^{-1}(1-0.05/2) = 1.96$인 경우 H_0을 기각할 것이다. 예를 들면, $Z = 2.78$로 구해졌다고 가정하자. 그러면 우린 H_0을 기각할 수 있다. 이러한 결과를 우리는 수준 0.05에서 H_0을 기각했다고 보고하기로 가정하자. 다른 유의수준에서 귀무가설을 좀 더 적절하게 검정하고자 하는 어떤 또 다른 통계학자는 이러한 보고를 그대로 참조하여도 되는가?

▲

▷ **예제 9.1.11 : Testing Hypotheses about Mean of a Normal Distribution with Known Variance.**
이전 예제에서 검정통계량은 $Z = 2.78$로 계산되었다. 이는 $2.78 \ge \Phi^{-1}(1-\alpha_0/2)$를 만족하는 모든 유의수준 α_0에 대하여 귀무가설을 기각할 수 있는 것으로 알려졌다. 본 교재의 부록에 있는 정규분포표를 살펴보면, 이러한 부등식은 $\alpha_0 \ge 0.0054$로 변환될 수 있다. 여기서 값 0.0054는 관측된 데이터와 검정된 가설에 대한 $p-value$라고 부른다. 0.01>0.0054 이므로, 수준 0.01에서의 가설을 검정하길 원하는 통계학자들 또한 H_0을 기각할 수 있다.

▲

▷ **[NOTE] 유의확률(p-value) 계산하기**
통계량 T에 대한 단순검정에서 "$T \geq c$일 때 귀무가설을 기각한다."의 형태를 가지면, 직관적인 방법으로 유의확률을 계산할 수 있다.

각 t에 대하여, δ_t를 $T \geq t$이면 H_0을 기각하는 검정(test)라고 하자. 그러면 $T = t$일 때의 유의확률은 검정 δ_t의 크기로 구해진다. 즉, 유의확률은 다음과 같다.

$$sup_{\theta \in \Omega_0} \pi(\theta | \delta_t) = sup_{\theta \in \Omega_0} \Pr(T \geq t | \theta)$$

일반적으로, $\pi(\theta | \delta_t)$는 Ω_0와 Ω_1 사이의 경계에 놓인 어떤 θ_0에서 최대화된다. 유의확률은 T의 분포의 위쪽 꼬리에서의 확률로 계산되기 때문에, 이는 때때로 꼬리영역(a tail area) 라고 불린다.

▷ **예제 9.1.12 : Testing Hypotheses about a Bernoulli Parameter.**
예제 9.1.9에서의 가설 검정을 고려하자. 우린 $Y \geq c$이면 H_0을 기각하는 검정을 사용할 것이다. $Y = y$가 구해졌을 때의 p-값은 $sup_{p \leq p_0} \Pr(Y \geq y | p)$ 로 구해진다. 이 예제에서, $\Pr(Y \geq y | p)$는 p의 함수로, 증가함수임을 쉽게 알 수 있다. 이런 이유로 p-값은 $\Pr(Y \geq y | p = p_0)$이다. 예를 들어, $p_0 = 0.3$이고 $n = 19$이라 하자. 만약 $Y = 6$으로 구해졌다면, $\Pr(Y \geq 6 | p = 0.3) = 0.0473$으로 구해진다.

*** p-값의 계산은 "$T \geq c$일 때 귀무가설을 기각한다."인 검정이 아닌 경우에 계산이 더욱 복잡해진다. 우리는 본 교재에서 p-값은 오직 이러한 형태의 검정에서만 계산하는 것을 권장한다.

▲

§. 동등성 검정과 신뢰 집합 (Equivalence of Tests and Confidence Sets)

Definition 9.1.10. 신뢰집합

만약 확률 집합(random set) $\omega(\overrightarrow{X})$이 $\Pr\left[g(\theta_0)\in\omega(\overrightarrow{X})|\theta=\theta_0\right]\geq\gamma$ for $\forall\,\theta_0\in\Omega$을 만족하면 이를 $g(\theta)$ **에 대한 계수** γ **의 신뢰집합(coefficient** γ **confidence set for** $g(\theta)$**)** 라고 부른다. 만약 위의 부등식이 모든 θ_0에 대하여 등식이 성립하면 우리는 이를 신뢰집합이 **정확하다(exact)**고 부른다.

▷ **예제 9.1.15 : Constructing a Test from a Confidence Interval.**

정규분포에서의 알려지지 않은 평균과 분산에 대하여 신뢰구간을 구하는 방법은 이전 Section에서 논의하였다. 그러한 논의들에 이어서, 다음을 가정해보고자 한다.

$$X_1,\ \dots,X_n\sim N(\mu,\sigma^2),\ \ \mu,\sigma^2\ :\ \text{unknown},\ \ \theta=(\mu,\sigma^2),\ \ g(\theta)=\mu.$$

이전 Section의 방법으로, 우리는 다음의 통계량들을 사용할 것이다.

$$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i\ ,\ \ \sigma'=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2}$$

그러면, $g(\theta)$에 대한 계수 γ의 신뢰구간은 다음과 같다.

$$I=\left(\overline{X}_n-T_{n-1}^{-1}\left(\frac{1+\gamma}{2}\right)\frac{\sigma'}{\sqrt{n}},\overline{X}_n+T_{n-1}^{-1}\left(\frac{1+\gamma}{2}\right)\frac{\sigma'}{\sqrt{n}}\right)\ ,\ \text{where}$$

$T_{n-1}^{-1}(\bullet)$: the quantile function of the t-distribution with $n-1$ degrees of freedom.

각 μ_0에 대하여, 우리는 다음 가설에 대한 수준 $\alpha_0=1-\gamma$ 검정을 구하기 위해 위 구간을 사용한다.

$$\begin{aligned}H_0 &: \mu=\mu_0\\ H_1 &: \mu\neq\mu_0\end{aligned}$$

이 가설에 대한 검정은 $\mu_0\in I$ 인 경우 H_0을 기각한다. 약간의 대수적 조작을 가하면, 다음의 명제가 동치임을 보일 수 있다.

$$\mu_0\notin I\Leftrightarrow\left|\sqrt{n}\,\frac{\overline{X}_n-\mu_0}{\sigma'}\right|\ \geq\ T^{-1}\left(\frac{1+\gamma}{2}\right)$$

이 검정은 사실 t 검정과 동일하다는 것을 Section 9.5에서 좀 더 면밀하게 다룰 것이다.

▲

§. 우도비 검정 (Likelihood Ratio Tests)

Definition 9.1.11 우도비 검정

$$\text{통계량 } \Lambda(\vec{x}) = \frac{sup_{\theta \in \Omega_0} f_n(\vec{x}|\theta)}{sup_{\theta \in \Omega_0} f_n(\vec{x}|\theta)}$$

을 **우도비 통계량(likelihood ratio statistic)** 이라 하고, 가설에 대한 우도비 검정은 아래와 같다.

$$\Lambda(\vec{x}) \le k \text{ for some constant } k \Rightarrow H_0 \text{ 을 기각한다.}$$

▷ **예제 9.1.18 : Likelihood Ratio Test of Two-Sided Hypotheses about a Bernoulli Parameter.**

Y : 알려지지 않은 모수 θ에 대한 n개의 독립적인 베르누이 시행에서 성공(successes)의 개수, 라고 하고 이를 관측한다고 가정하자. 가설 $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$ 을 고려하자. $Y = y$로 관측되었을 때, 우도함수는 다음과 같이 구해진다.

$$f(y|\theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y}$$

이 경우, $\Omega_0 = \{\theta_0\}$, $\Omega = [0,1]$ 이다. 우도비 통계량은 아래와 같이 구할 수 있다.

$$\Lambda(y) = \frac{\theta_0^y (1-\theta_0)^{n-y}}{sup_{\theta \in [0,1]} \theta^y (1-\theta)^{n-y}}$$

이 우도비 통계량의 분모에서, 상한(supremum)값은 최대우도추정량을 찾는 방법을 고려하여 구하면 찾을 수 있다. 그러한 상한값, 즉 최댓값은 θ가 최대우도추정량 $\hat{\theta} = y/n$와 같을 때 구해지며, 자세한 과정은 생략한다. 따라서 우도비 통계량은 다음과 같이 구해진다.

$$\Lambda(y) = \left(\frac{n\theta_0}{y}\right)^y \left(\frac{n(1-\theta_0)}{n-y}\right)^{n-y}.$$

이 우도비 통계량은 y가 0과 n에 가까워질 때 작아지고, $y = n\theta_0$ 에 가까워질 때 그 값이 커진다. 특정한 예로써, $n = 10$이고 $\theta_0 = 0.3$이라 가정하자. 그러면 $y = 0, \ldots, 10$일 때의 우도비 통계량을 구할 수 있다.

y	0	1	2	3	4	5	6	7	8	9	10
$\Lambda(y)$	0.028	0.312	0.773	1.000	0.797	0.418	0.147	0.034	0.005	3×10^{-4}	6×10^{-6}
$\Pr(Y = y \mid \theta = 0.3)$	0.028	·0.121	0.233	0.267	0.200	0.103	0.037	0.009	0.001	1×10^{-4}	6×10^{-6}

▲

9.5 T 검정(The t-test)

§. T 검정의 여러 성질(Properties of the t-test)

▷ **예제** 9.5.1~3 : Nursing Homes in New Mexico.

뉴 멕시코 요양원에서의 입원기간에 대한 연구를 실시하였다고 가정하자. 우리는 표본의 크기가 $n=18$인 알려지지 않은 모수 μ,σ^2를 갖는 정규확률변수로써 입원기간 수에 대한 모형화를 시행하였다. 우린 다음과 같은 가설에 관심이 있다.

$$H_0 : \mu \geq 200 \text{ vs } H_1 : \mu < 200$$

<u>Introduction of T test</u>

9.5장에서 우리는, 정규분포의 모평균과 모분산을 모를 때, 평균에 대한 가설검정의 문제를 고려하고자 한다. 즉, 모평균과 모분산이 알려지지 않았을 때의 정규확률변수 $X_1, \ldots, X_n$에 대하여 다음과 같은 가설을 검정하고자 한다.

$$H_0 : \mu \leq \mu_0 \text{ vs } H_1 : \mu > \mu_0$$

이 때, 모수공간 Ω 는 $-\infty < \mu < \infty$, $\sigma^2 > 0$인 모든 2차원 벡터 (μ, σ^2)로 구성되어 있다. 귀무가설 H_0은 $\mu \leq \mu_0$를 만족하는 벡터가 $(\mu, \sigma^2) \in \Omega_0 \subset \Omega$ 임을 의미하고, 대립가설 H_1은 $\mu > \mu_0$을 만족하는 벡터가 $(\mu, \sigma^2) \in \Omega_1 = \Omega_1^C \subset \Omega$임을 의미한다. 이전 예제에서 보였듯이, 이 가설은 μ에 대한 단측검정(one-sided)에 해당한다. 검정통계량은 다음의 식을 사용한다.(이 식들은 정규분포에 대한 최대우도추정량에서 유래한다.)

$$\text{표본평균 } \overline{X}_n = \sum_{i=1}^{n} \frac{X_n}{n}, \text{ 표본표준편차 } \sigma' = \frac{\sum_{i=1}^{n} (X_i - \overline{X}_n)^2}{\sqrt{n-1}}, \text{ 검정통계량 } U = \frac{\overline{X}_n - \mu_0}{\frac{\sigma'}{\sqrt{n}}}.$$

검정은 $U \geq c$일 때 H_0를 기각한다. $\mu = \mu_0$ 일 때, 정리 8.4.2 에 의하면, 통계량 U의 분포는 σ^2의 값과 관계없이 자유도 $n-1$인 t-분포를 따른다. 이런 이유로, U에 대한 이러한 검정은 "t-검정"이라고 부른다. 앞서 소개한 가설과 달리 가설이 $H_0 : \mu \geq \mu_0$ vs $H_1 : \mu < \mu_0$ 이면 이는 마찬가지의 논의에 의해, 검정은 $U \leq c$일 때 H_0를 기각하면 된다.

α_0 검정을 통해 이를 검정하고자 한다면, 우린 검정통계량 U 가 검정의 크기가 α_0와 같도록 만드는 c를 택하여 이와 많아야 같을 때, H_0을 기각하도록 하는 T 검정을 사용할 수 있다. 유의수준 $\alpha_0 = 0.1$로 두자. 그러면 우린 $U \leq c$, c : 자유도 17인 t-분포의 0.1 분계선(quantile)≡-1.333 일 때 H_0을 기각한다. 교재의 예제 8.6.3의 데이터를 사용하여 표본 평균을 구하면 $\overline{X}_{18} = 182.17$이고, 표본표준편차는 $\sigma' = 72.22$이다. 따라서 $U = \frac{182.17 - 200}{\frac{72.22}{\sqrt{17}}} = -1.018$ 로 구해진다. 그러면 $U = -1.018 > -1.333$ 이므로 우리는 유의수준 0.1에서 귀무가설 $H_0 : \mu_0 \geq 200$을 기각할 수 없다.

▲

Theorem 9.5.2 t-검정의 p-값
가설 $H_0 : \mu \le \mu_0$ vs $H_1 : \mu > \mu_0$ 또는 $H_0 : \mu \ge \mu_0$ vs $H_1 : \mu < \mu_0$ 을 검정한다고 가정하자. u를 통계량 U의 관측값이라 하고, $T_{n-1}(\bullet)$을 $n-1$ 자유도인 t-분포의 누적분포함수(c.d.f.)라고 하자. 그러면 각 가설에 대한 p-값은 $1-T_{n-1}(u)$, $T_{n-1}(u)$ 이다.

▷ **예제 9.5.4 : Length of Fibers.**
밀리미터인 금속섬유의 길이가 알려지지 않은 모수 μ, σ^2를 갖는 정규분포를 따른다고 가정하자. 우리는 다음 가설을 검정할 것이다.

$$H_0 : \mu \le 5.2 \text{ vs } H_1 : \mu > 5.2$$

15개의 금속섬유의 길이가 무작위하게 선택되고 측정되었다고 가정하고, 표본평균이 $\overline{X}_{15} = 5.4$, $\sigma' = 0.4226$으로 측정되었다고 가정하자. 이러한 값들에 기초하여, 우리는 유의수준 $\alpha_0 = 0.05$에서의 T검정을 실시할 것이다.

$n = 15$, $\mu_0 = 5.2$ 이므로, 검정통계량 U는 자유도가 14이고 $\mu = 5.2$인 t-분포를 따른다. 이는 t-분포표를 참조하면, $T_{14}^{-1}(0.95) = 1.761$로 구해진다. 이런 이유로, 귀무가설은 $U > 1.761$일 때 기각될 수 있다. U에 대한 수치적 값은 1.833으로 계산되므로, H_0은 유의수준 0.05에서 기각될 수 있다.
통계량 U에 대한 관측값 $u = 1.833$과 $n = 15$에서 우리는 컴퓨터 소프트웨어를 이용하여 t-분포의 누적분포함수인 p-값을 계산할 수 있다. 이 가설에 대한 p-값은 $1 - T_{14}(1.833) = 0.0441$로 구해진다.

▲

▷ **예제 9.5.7 : Crash Test Dummies.**
국가 교통안전위원회는 자동차 충돌에 대한 더미의 손상량과 위치에 관한 충돌 테스트 데이터를 수집한다. 더미를 운전자 좌석, 탑승자 좌석에 배치하고 각 경우 충돌 시 더미의 머리 손상량을 측정하였다. 이들에 대한 측정값과 그 관계는 아래의 Figure 9.13에 제시하였다.

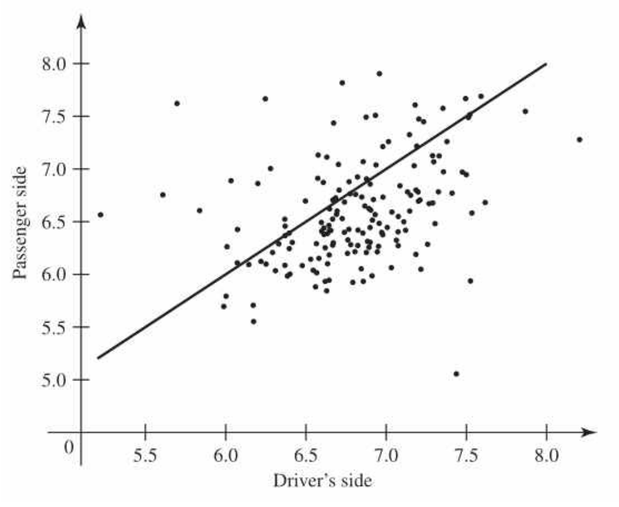

Figure 9.13 Plot of logarithms of head injury measures for dummies on driver's side and passenger's side. The line indicates where the two measures are equal.

확률변수 $X_1, \ldots, X_n$을 각 경우의 더미의 머리 부상 측정치 사이의 차이라고 하자. 우린 $X_1, \ldots, X_n$를 평균μ와 분산 σ^2를 갖는 정규분포로부터 취해지는 확률표본으로 모형화할 수 있다. 우린 다음과 같은 가설을 세우고 유의수준 $\alpha_0 = 0.01$에서 검정하고자 한다.

$$H_0 : \mu \le 0 \text{ vs } H_1 : \mu > 0$$

표본크기가 $n = 164$개로 정해졌을 때, 이 가설은 $U \ge T_{163}^{-1}(0.99) = 2.35$ 일 때 귀무가설을 기각할 수 있다.

Figure 9.13에서 각 좌표의 차이에 대한 평균은 $\overline{x_n} = 0.2199$, 표본표준편차는 $\sigma' = 0.5342$로 구해졌다. 그러면 검정통계량 $U = 5.271$로 구할 수 있다. 이는 명백히 2.35보다 큰 값이므로, 귀무가설은 유의수준 0.01에서 기각할 수 있다. 추가적으로 사실, p-값은 1.0×10^{-6}보다 작음을 확인할 수 있다.

▲

▷ **예제 9.5.9 : Egyptian Skulls.**
유의수준 $\alpha_0 = 0.05$에서 가설 $H_0 : \mu = 140$ vs $\mu \ne 140$에 대한 검정을 실시하고자 한다. 이는 양측검정(Two-sided)에 해당하는 가설이므로, 앞선 논의에 의한 단측검정을 각각 다른 방향의 부등호로 실시하여야 한다. 즉 c_1과 c_2를 구하여 이에 맞는 기각역을 설정해주어야 한다. **T검정에서 양측검정인 경우, $|U| \ge T_{n-1}^{-1}(1-\alpha_0/2)$ 일 때 H_0을 기각함을 반드시 기억하길 바란다.** $c_1 = -T_{29}^{-1}(0.975) = -2.045$이고 $c_2 = 2.045$로 구해진다. $n = 30$에 대한 표본평균은 $\overline{X}_{30} = 131.37$로, 표본표준편차는 $\sigma' = 5.129$로 구해진다. 따라서 검정통계량 U의 관측값 $u = \dfrac{131.37 - 140}{\frac{5.129}{\sqrt{30}}} = -9.219$로 구해진다. 이는 -2.045보다 작으므로 유의수준 0.05에서 귀무가설 H_0을 기각할 수 있다.

▲

9.6 두 정규분포의 평균 비교 (Comparing the Means of Two Normal Distributions)

§. 이표본 T 검정(The Two-Sample t-test)

Theorem 9.6.1 이표본 t-통계량
두 표본공간에 대한 표본평균과 표본분산을 다음과 같이 가정하자.

$$\overline{X}_m = \frac{1}{m}\sum_{i=1}^{m} X_i \ , \ \ \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i, \qquad S_X^2 = \sum_{i=1}^{m}(X_i - \overline{X}_m)^2,\ S_Y^2 = \sum_{i=1}^{n}(Y_i - \overline{Y}_n)^2$$

그러면 검정통계량은 다음과 같이 정의된다.

$$U = \frac{(\overline{X}_m - \overline{Y}_n)}{\sqrt{\dfrac{\left(\dfrac{1}{m} + \dfrac{1}{n}\right)(S_X^2 + S_Y^2)}{m+n-2}}}$$

$\mu_1 = \mu_2$를 만족하는 모든 $\theta = (\mu_1, \mu_2, \sigma^2)$에 대하여, U의 분포는 자유도 $m+n-2$를 갖는 t-분포이다.

▷ **Theorem 9.6.3 이표본 T 검정에서의 p-값 ⇔ $df = m+n-2$임을 제외하고 정리 9.5.2와 같다.**

▷ **예제 9.6.3 : Roman Pottery in Britain.**
과거 Great Britain의 다양한 지역에서 발견된 로마시대 양식 도기 표본에 대한 연구에 대하여 이야기하고자 한다. 각 도자기 표본에 대한 측정 수단 중 하나로 표본의 알루미늄 산화물에 대한 백분율을 고려하자.

우린 두 개의 다른 지역에서 발굴된 도기 표본에 대한 알류미늄 산화물의 백분율에 관심이 있다고 가정하자. 각각의 표본 크기는 $m = 14$, $n = 5$로, 표본평균과 표본분산은 각각 $\overline{X}_m = 12.56$, $S_X^2 = 24.65$, $\overline{Y}_n = 17.32$, $S_Y^2 = 11.01$로 구해졌다. 우린 두 개의 다른 평균 μ_1, μ_2와 같은 분산 σ^2을 갖는 정규확률변수로써의 데이터를 모형화한다고 가정하자. 그러면 U의 관측값 $u = -6.302$로 구해진다.

t-분포표를 참조하면, 자유도는 $m+n-2 = 17$이고, $T_{17}^{-1}(0.995) = 2.898$, $U < -2.898$임을 알 수 있다. 따라서, 우린 임의의 유의수준 $\alpha_0 \geq 0.0005$에서 H_0을 기각할 수 있다. 이 때, U에 대한 p-값은 $T_{17}(-6.302) = 4 \times 10^{-6}$로 구해짐을 확인할 수 있다.

▲

§. T검정에서의 양측검정 (Two-sided Alternatives)

t-검정에서의 양측검정 (One sample & Two sample)
유의수준 α_0에서 가설이 양측검정의 형태일 때, 즉 $H_0 : \mu_1 = \mu_2$ vs $H_1 : \mu_1 \neq \mu_2$일 때, 기각역은 다음과 같이 구해진다.

$$|U| \geq c \text{ where } c = T_k^{-1}(1-\alpha_0/2),\ U : \text{검정통계량},\ k = \begin{cases} n-1 & One-sample \\ m+n-2 & Two-sample \end{cases}.$$

검정통계량 U가 $U=u$로 관측되었을 때, p-값은 다음과 같이 구해진다.

$$p-value = 2\left[1 - T_k(|u|)\right] \ ,\ k = \begin{cases} n-1 & One-sample \\ m+n-2 & Two-sample \end{cases}$$

▷ **예제 9.6.5 : Comparing Copper Ores.**
어떤 특정한 지역의 구리광산에서 얻어지는 8개의 광석 견본에 대한 확률 표본을 가정하자, 여기서 각 견본에서의 구리의 양은 gram 단위로 측정되었다. 이러한 8개의 견본을 $X_1,\ldots,X_8$로 두고, 이에 대한 표본평균과 표본분산은 $\overline{X}_8 = 2.6$과 $S_X^2 = 0.32$로 구하였다.

앞서 설명한 지역이 아닌 다른 지역의 광산에서 얻어지는 10개의 광석 견본에 대한 확률 표본을 가정하자, 여기서 각 견본에서의 구리의 양은 마찬가지로 gram단위로 측정되었고, 이러한 10개의 견본을 $Y_1,\ \ldots,\ Y_{10}$으로 두어 $\overline{Y}_{10} = 2.3$, $S_Y^2 = 0.22$로 구하였다. μ_1을 첫 번째 광산에서의 구리의 양에 대한 평균, μ_2를 두 번째 광산에서의 구리의 양에 대한 평균으로 두자. 우리가 검정하고자 하는 가설은 다음과 같다.

$$H_0 : \mu_1 = \mu_2 \text{ vs } H_1 : \mu_1 \neq \mu_2$$

우린 모든 이러한 모든 관측값들이 정규분포를 따르고, 만약 평균의 값이 서로 다르더라도, 두 광산 모두 같은 분산을 가진다고 가정할 것이다. 앞선 논의에 의하면, 표본크기는 $m=8$, $n=10$이고, 검정통계량 U는 $U=3.442$로 구해진다. 또한 자유도 16인 t-분포표를 참조하면, $T_{16}^{-1}(0.995) = 2.921$임을 알 수 있고, 이는 검정통계량의 관측값과 연관된 꼬리 영역이 2×0.005보다 작음을 확인할 수 있다. 이런 이유로, 귀무가설은 임의의 유의수준 $\alpha_0 \geq 0.01$에서 기각될 수 있다. (사실, $U=3.442$의 양측검정 꼬리영역은 0.003이다.)

▲

9.7 F 분포(The F Distributions)

§. F 분포의 정의 (Definition of the F Distribution)

<u>Definition 9.7.1</u> **F 분포**

Y와 W가 독립적인 확률변수로써 $Y \sim \chi_m^2$, $W \sim \chi_n^2$, $m,n \in Z_+$ 을 만족한다고 하자. 그러면 새로운 확률변수 X는 다음과 같이 정의할 수 있다.

$$X = \frac{Y/m}{W/n} = \frac{nY}{mW}$$

그러면 이러한 X의 분포를 **자유도 m, n을 갖는 F 분포(F distribution)**라고 부른다.

▷**Theorem 9.7.1 : p.d.f. of F distribution**

For $X = x$, p.d.f. $f(x)$ is as follows : $f(x) = 0$ for $x \le 0$ and

$$f(x) = \frac{\Gamma\left[\frac{1}{2}(m+n)\right] m^{m/2} n^{n/2}}{\Gamma\left(\frac{1}{2}m\right)\Gamma\left(\frac{1}{2}n\right)} \left(\frac{x^{(m/2)-1}}{(mx+n)^{(m+n)/2}}\right) \text{ for } x > 0$$

<u>Theorem 9.7.2</u> **F 분포의 자유도**

확률변수 X와 Y에 대하여 다음이 성립한다.

$$X \sim F_{m,n} \Rightarrow \frac{1}{X} \sim F_{n,m}$$

$$Y \sim T_n \Rightarrow Y^2 \sim F_{1,n}$$

▷ **예제 9.7.2 : Determing the 0.05 Quantile of an F Distribution.**

확률변수 X가 자유도를 6과 12로 갖는 F 분포를 따른다고 가정하자. 우린 X의 0.05 분계선을 결정하고자 한다. 즉, 값 x에 대하여 $\Pr(X < x) = 0.05$를 결정하고자 함이다.

만약 $Y = 1/X$로 두면, Y는 정리 9.7.2에 의해 자유도 12와 6을 갖는 F 분포를 따른다. 이는 교재의 F 분포표를 참조하면 $\Pr(Y \le 4.00) = 0.95$임을 확인할 수 있다. 이런 이유로, $\Pr(Y > 4.00) = 0.05$이다. $Y > 4.00 \Leftrightarrow X < 0.25$이므로, 이는 $\Pr(X < 0.25) = 0.05$임을 얻는다. 이 때, F 분포는 연속분포이므로, $\Pr(X \le 0.25) = 0.05$이고, 0.25는 X의 0.05분계선에 위치한다.

▲

§. 두 정규분포의 분산의 비교 (Comparing the Variances of Two Normal Distributions)

Definition 9.7.2 F 검정

확률변수 $X_1, \ldots, X_m$이 알려지지 않은 모수 μ_1, σ_1^2을 갖는 정규분포로부터 m개의 관측값인 확률표본을 생성한다고 가정하자. 그리고 확률변수 $Y_1, \ldots, Y_n$ 또한 마찬가지로, 독립적인 확률표본으로써 알려지지 않은 모수 μ_2, σ_2^2를 갖는 정규분포로부터 n개의 관측값인 확률표본을 생성한다고 가정하자. 마지막으로, 다음 가설을 유의수준 $0 < \alpha_0 < 1$에서 검정하고자 한다고 가정하자.

$$H_0 : \sigma_1^2 \leq \sigma_2^2 \text{ vs } H_1 : \sigma_1^2 > \sigma_2^2$$

각 검정 절차 δ에 대하여, $\pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2 \mid \delta)$를 δ의 검출력함수(power function)이라 두자. 각 확률변수에 대한 표본분산 S_X^2와 S_Y^2를 다음과 같이 정의하자.

$$S_X^2 = \sum_{i=1}^{m}(X_i - \overline{X}_m)^2, \quad S_Y^2 = \sum_{i=1}^{n}(Y_i - \overline{Y}_n)^2$$

그러면 $S_X^2/(m-1)$과 $S_Y^2/(n-1)$은 각각 σ_1^2과 σ_2^2의 추정량(estimator)이 된다. 이제 여기서 V를 다음과 같이 정의하고, c에 대한 조건을 다음과 같이 주면 F 검정의 귀무가설 H_0을 기각할 수 있다.

$$\text{검정통계량 } V = \frac{S_X^2/(m-1)}{S_Y^2/(n-1)}$$

$$V \geq c, \ c \text{ is chosen to make the test have a desired level of significance.}$$

F 검정에서의 양측검정

유의수준 α_0에서 가설이 양측검정의 형태일 때, 즉 $H_0 : \sigma_1^2 = \sigma_2^2$ vs $H_1 : \sigma_1^2 \neq \sigma_2^2$일 때, 기각역은 다음과 같이 구해진다.

$$V \leq c_1 \text{ or } V \geq c_2 \text{ s.t. } \Pr(V \leq c_1) + \Pr(V \geq c_2) = \alpha_0$$
$$\Rightarrow \Pr(V \leq c_1) = \Pr(V \leq c_2) = \alpha_0/2 \text{ when } \sigma_1^2 = \sigma_2^2$$

c_1 : $\alpha_0/2$ quantiles of the appropriate F distribution
c_2 : $1-\alpha_0/2$ quantiles of the appropriate F distribution

Theorem 9.7.5 p-value of Equal-Tailed Two-Sided F Test.

가설이 양측검정, 즉 $H_0 : \sigma_1^2 = \sigma_2^2$ vs $H_1 : \sigma_1^2 \neq \sigma_2^2$일 때, δ_{α_0}을 F 검정에서의 양측검정이라 하자. 그러면 $V = v$로 관측되었을 때 H_0을 기각하는 δ_{α_0}에 대한 가장 작은 유의수준 α_0는 다음과 같다.

$$p-value = 2\min\{1 - G_{m-1,n-1}(v), G_{m-1,n-1}(v)\}, \ G \text{ : c.d.f. of } F \text{ distribution.}$$

▷ **예제 9.7.4 : Rain from Seeded Clouds.**

이전 예제 9.6.2에서, 우리는 seeded와 unseeded clouds로 나뉘는 log-rainfalls의 평균을 각 경우의 분산이 같다는 가정하에 비교하였다. 이제 우리는 이러한 두 분산이 서로 다른 경우에 대한 검정을 고려하고자 한다. 즉, 가설을 다음과 같이 설정하고자 한다. 이 때, 유의수준 $\alpha_0 = 0.05$로 둔다.

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ vs } H_1 : \sigma_1^2 \neq \sigma_2^2$$

F 검정의 정의에 따르면, $m = n$이므로, 검정통계량 $V = \dfrac{63.96}{67.39} = 0.9491$로 구해진다. 우린 이를 자유도를 25와 25로 갖는 F 분포의 0.025와 0.975 분계수를 비교할 필요가 있다. 우리가 가진 F 분포표의 분계수는 자유도 25일 때의 행과 열을 갖지 않으므로, 우린 20과 30의 자유도 사이를 보간하여 구하거나 컴퓨터 프로그램을 이용하여 이러한 분계수를 구해야만 한다. 그러한 분계수는 0.4484와 2.2303으로 구해진다. V는 이러한 두 수 사이의 값이므로, 우린 유의수준 $\alpha_0 = 0.05$에서 귀무가설을 기각할 수 없다.

▲

10.1 적합도 검정 (Test of Goodness-of-Fit)

§. 비모수적 문제의 서술 (Description of Nonparametric Problems)

<u>Definition</u> **비모수적 문제와 비모수적 방법**

추정, 검정에 대한 문제 또는 이를 아우르는 통계적 문제들에서, 관측값들(observations)에 대한 가능한 분포가 특정 분포로 제한되지 아니한 경우 이를 **비모수적 문제(Nonparametric problem)** 라고 부르고, 이러한 문제에 적용할 수 있는 방법을 **비모수적 방법(Nonparametric method)**라고 부른다.

▷ (**예제 1)** 어떤 관측값들이 포아송 분포를 따르는 확률표본을 생성한다고 가정하자. 그러면 우리는 포아송 분포의 모수 λ를 모르더라도 모수 λ가 포아송 분포, 즉 특정 분포를 따르는, 다시말하면 어떤 특정 모수 집합을 따른다는 사실을 자연스럽게 알 수 있다. 그러므로 이는 비모수적 문제에 해당하지 않고, 비모수적 방법을 적용하지 않고 모수를 추정하거나 검정을 할 수 있다.

▷ (**예제 2)** 어떤 관측값들이 정규분포를 따르는 확률표본을 생성한다고 가정하면, 정규분포의 모수는 평균과 분산으로 2개를 갖는다. 그런데 만약 이러한 모수들이 알려져 있지 않다고 가정하자. 그렇다면 이는 비모수적 문제에 해당하는가?. 결과적으로, 이는 비모수적 문제가 아니다. 모수를 모르더라도 모수가 특정 분포인 정규분포를 따르는 일종의 확률변수임을 알고 있기 때문에, 평균 μ와 분산 σ^2는 정규분포의 성질을 갖는 어떤 특정 모수집합 Ω에 속할 것이다. 그러므로 이는 비모수적 방법이 아니다. 앞선 여러 챕터에서 우리가 배운 여러 추정 및 검정의 내용도 모수적 방법이었다는 것을 상기하자.

§. 범주형 데이터 (Categorical Data)

Definition χ^2 **검정 (Pearson's chi-square test)**

k개의 서로 다른 유형의 개체들로 구성된 큰 모집단을 가정하고, p_i를 유형 $i=1,\ldots,k$에 대해 무작위로 선택되는 개체의 확률이라 하자. 물론, p_i는 확률이므로 확률의 공리 $p_i \geq 0$, $\Sigma_{i=1}^k p_i = 1$를 만족한다. p_1^0, ..., p_k^0 를 $p_i^0 > 0$을 만족하는 특정한 수라 가정하고, $\Sigma_{i=1}^k p_i^0 = 1$을 만족하며 다음 가설을 검정하려한다고 가정하자.

$$H_0 : p_i = p_i^0 \text{ for } i=1,\ \ldots, k,$$
$$H_1 : p_i \neq p_i^0 \text{ for at least one value of } i.$$

모집단에서 취해지는 크기가 n인 확률표본을 가정하자, 즉 n개의 독립적인 관측값들이 취해지고, 각 관측값이 $i=1,\ldots,k$인 유형이 될 확률 p_i이 존재한다고 하자. 이러한 n개의 관측값을 기저로 하여, 위 가설은 검정할 수 있다.

$i=1,\ldots,k$에 대하여, N_i 를 유형 i에 대한 확률표본의 관측값의 수라고 하자. 그러면 N_1, ..., N_k는 음수가 아닌 정수로써 $\Sigma_{i=1}^k N_i = n$을 만족한다(사실, $(N_1,\ldots,N_n)$는 모수가 n과 $\vec{p}=(p_1, \ldots, p_k)$인 다항분포를 갖는다.).

귀무가설 H_0이 참이면, 유형 i의 관측값들의 예측 수는 np_i^0 , $i=1,\ldots,k$ 이다. 관측값 N_i의 실체 수와 예측 수 np_i^0 사이의 차이는 H_0이 거짓일 때보다 H_0이 참일 때 작아지는 경향이 있다. 따라서 가설검정은 결국 $N_i - np_i^0$ for $i=1,\ldots,k$를 가설 검정으로 하는 것과 같은 문제가 된다. 그리고 $N_i - np_i^0$이 상대적으로 커질 때 H_0을 기각하게 된다.

아래 정리에 제시된 χ^2 검정의 검정통계량은 Karl Pearson 에 의해 1990년도에 증명되었다.

Theorem 10.1.1 χ^2 **검정의 검정통계량**

다음 통계량 Q를 χ^2 검정의 검정통계량으로 택한다.

$$Q = \sum_{i=1}^{k} \frac{(N_i - np_i^0)^2}{np_i^0} = \sum \frac{(O-E)^2}{E}, \quad O \text{ : Observed},\ E \text{ : Expected}.$$

이는 H_0이 참이고 표본 크기가 $n \to \infty$ 일 때 자유도가 $k-1$인 χ^2 분포로 수렴하는 성질을 갖는다.

▷ [Definition] 위 검정통계량을 이용하여 가설검정을 시행하는 것을 **카이제곱 적합도 검정**이라 부른다.

▷ **예제 10.1.3 : Blood Types.**

우리는 아래에 제시된 Table 10.1, 10.2과 같은 범주형 데이터(숫자가 아닌 명목형 자료)에 대한 가설검정을 실시하고자 한다. 특히, 이러한 범주형 데이터에서 확률에 대한 추론을 하고 싶기 때문에, 그에 적합한 방법을 사용해야만 한다. 이러한 범주형 데이터에 대한 가설검정은 χ^2 검정이 그 예로 적합하다. χ^2 검정은 모수 또는 비모수적 문제 둘 다 적용할 수 있는 방법이다. 주어진 Table은 다음과 같다.

Table 10.1 Counts of blood types for white Californians			
A	B	AB	O
2162	738	228	2876

Table 10.2 Theoretical probabilities of blood types for white Californians			
A	B	AB	O
1/3	1/8	1/24	1/2

이 범주형 데이터는 어떤 특정 분포를 따른다거나 모수집합을 특정할 수 있는 문제가 아니므로, 이는 비모수적 문제가 된다. 그러나 Table 10.1을 잘 살펴보면, 이는 각 혈액형 유형(type)의 빈도수이므로, 각 유형은 전체 빈도수 6004에 대한 수학적 확률, 즉 어떤 비율로 나타낼 수 있다. 실제로, 예를 들어 B형의 상대도수를 컴퓨터를 이용하여 구하면 $738/6004 = 0.1229181$인데, 제시된 Table 10.2를 참고하면 $1/8 = 0.125$로 거의 근사한 값을 가진다. 문제에서는 수학적 확률(Theoredical probabilities)로 Table 10.2를 제시했으므로, 자료의 출처에 관한 별 다른 의심 없이 이를 이용하기로 하자.

이제 우리는 혈액형 타입에 대한 데이터의 각 범주의 수학적 확률을 알고 있다. 그렇다면 이를 가지고 어떻게 가설을 세우고 검정할 것인가?, 이는 앞서 언급한 χ^2 검정이 적합하다. χ^2 검정 방법을 살펴보면, 비율에 대한 가설을 세우고, 이를 검정하는 것임을 쉽게 알 수 있다. 즉, 어떤 모집단의 한 범주가 어떤 비율을 따를 때, 그 비율을 추정하는 것과 같다고 생각할 수 있다. 그러면 우리는 가설을 세울 때 우리가 제시한 특정 비율이 모비율과 같음을 귀무가설로 채택하는 것은 자연스럽다. 그러면 이제 문제는 비율에 대한 가설검정의 문제로 넘어가게 되고, 우리는 이제 χ^2 검정 절차를 잘 따라서 결과를 도출해내면 될 것이다.

이제 가설을 세우자. 우선 범주의 개수는 4이므로, $i = 1, 2, 3, 4$, $k = 4$, Table 10.2에 제시된 확률은 모두 양수이고 $1/3 + 1/8 + 1/24 + 1/2 = 1$로 확률의 공리를 만족한다. 이들 확률을 각각 $p_1^0, p_2^0, p_3^0, p_4^0$로 두고, 표본 크기는 $n = 6004$, 각 혈액형에 대한 빈도수를 N_1, N_2, N_3, N_4로 두면, $\Sigma N_i = n$임은 쉽게 확인할 수 있다. 검정 절차에 따르면, 가설은 다음과 같이 세울 수 있다.

$$H_0 : p_i = p_i^0 \text{ for } i = 1, 2, 3, 4 \text{ vs } H_1 : p_i \neq p_i^0 \text{ for at least one value of } i.$$

가설을 세웠으니, 이제 본격적인 검정을 시행하기 위해 검정통계량을 구해보자. 검정통계량은 Pearson의 검정 통계량을 이용하면 된다.

$$Q=\sum_{i=1}^{k}\frac{(N_i-np_i^0)^2}{np_i^0}=\sum\frac{(O-E)^2}{E}, \quad O : \text{Observed}, \ E : \text{Expected}.$$

우리는 검정통계량 Q의 각 성분들을 모두 구하였으므로, 이를 계산하면 아래와 같이 얻을 수 있다.

$$Q=\sum_{i=1}^{4}\frac{(N_i-np_i^0)^2}{np_i^0}=\left(\frac{(2162-6004\times 1/3)^2}{6004\times 1/3}\right)+\cdots\ \cdots+\left(\frac{(2876-6004\times 1/2)^2}{6004\times 1/2}\right)=20.37$$

유의수준 α_0에 대하여 H_0를 검정하기 위해, 자유도가 3인 χ^2 분포의 $1-\alpha_0$ 분계선과 Q를 비교해야 한다. 먼저, p-값을 계산해보자. 이는 H_0을 기각할 수 있는 가장 작은 유의수준을 찾으면 된다. 카이제곱 적합도 검정의 경우, p-값은 $1-X_{k-1}^2(Q)$로 구할 수 있다. 이 때, X : $k-1$ 자유도를 갖는 카이제곱분포의 누적분포함수(c.d.f.)이다. 여기서 $k=4$이므로, $p-value=1-X_3^2(Q)=1.42\times 10^{-4}$ 이다.

만약 유의수준 $\alpha_0=0.05$로 두면, 꼬리영역이 되는 α_0보다 p-값이 훨씬 작으므로($1.42\times 10^{-4}=0.000142$) $\alpha_0=0.05$에서 귀무가설 H_0을 기각할 수 있다. (α_0는 0.01인 경우도 p-값이 훨씬 작음을 알 수 있다.)

(** $Q \ge (X_{k-1}^2)^{-1}(1-\alpha_0)=c \Leftrightarrow X_{k-1}^2(Q) \ge 1-\alpha_0 \Leftrightarrow \alpha_0 \ge 1-X_{k-1}^2(Q)=p-value$)

▲

R Code in example 10.1.3

Results	R Code
> df_blood_freq A B AB O 1 2162 738 228 2876 > chisq.test(df_blood_freq,p=df_blood_prob) Chi-squared test for given probabilities data: df_blood_freq X-squared = 20.359, df = 3, p-value = 0.000143	A=2162 B=738 AB=228 O=2876 df_blood_freq=data.frame(A,B,AB,O) df_blood_freq df_blood_prob=c(1/3,1/8,1/24,1/2) chisq.test(df_blood_freq,p=df_blood_prob)

§. 연속분포에 대한 가설 검정 (Testing Hypotheses about a Continuous Distribution)

연속분포에 대한 가설 검정 (Testing Hypotheses about a Continuous Distribution)

특정 분포(in continuous distribution)로부터 취해지는 관측값들에 대한 확률 표본에 대하여 검정하기 위해, 다음의 과정을 적용한다.

1. 실직선(real line) 전체를 분할하거나, k개의 서로소인 부분구간으로 이루어지며 구간 전체의 확률은 1을 만족하는 임의의 특정 구간을 설정한다. 일반적으로, 귀무가설 H_0이 참이면, k는 선택되며 각 부분구간의 관측값의 예상 개수는 적어도 5 이상이다.

2. 특정한 가정된 분포로써 i번째 부분구간에 배당되는 확률 p_i^0 를 결정한다, 그리고 각 $i=1,\ldots,k$에 대하여 i번째 부분구간의 관측값들의 예상 갯수 np_i^0를 계산한다.

3. i번째 부분구간에 해당하는 표본의 관측값들의 개수 N_i를 센다.

4. 검정통계량 Q를 계산한다. 만약 가정된 분포가 정확하면, Q는 근사적으로 $k-1$의 자유도를 갖는 χ^2 분포를 따른다.

▷ **예제 10.1.6 : Failure Times of Ball Bearings.**
이전 예제 10.1.1로 돌아가자. 수명에 대한 로그값이 $i.i.d.$한 정규분포를 따르는 확률 표본이라 가정하고, 귀무가설에 대한 χ^2검정을 실시하고자 한다. 이 때, 평균 $\mu=\log(50)=3.912$, 분산 $\sigma^2=0.25$라 하자.

각 구간에 대한 기대 개수가 적어도 5개 이상이기 위해서, 우린 많아야 $k=4$개의 구간들을 사용할 수 있다. 또한 우린 이러한 구간들이 귀무가설의 가정 아래에서 각각 0.25의 확률을 가진다고 가정할 것이다. 즉, 우린 총 확률이 1인 구간을 가정한 정규분포에서의 0.25, 0.5, 0.75 분계수를 갖는 구간으로 나눌 것이다. 이러한 분계수는 다음과 같이 정해진다.

$$\begin{aligned}
&3.912+0.5\Phi^{-1}(0.25)=3.192+0.5\times(-0.674)=3.575\\
&3.912+0.5\Phi^{-1}(0.5)=3.192+0.5\times 0 \qquad\qquad\; =3.912\\
&3.912+0.5\Phi^{-1}(0.75)=3.192+0.5\times 0.674 \qquad\;\; =4.249
\end{aligned}$$

(∵ 표준정규분포의 0.25와 0.75 분위수 : ± 0.674)

관측된 로그값들은 다음과 같이 주어진다.

2.88	3.36	3.50	3.73	3.74	3.82	3.88	3.95
3.95	3.99	4.02	4.22	4.23	4.23	4.23	4.43
4.53	4.59	4.66	4.66	4.85	4.85	5.16	

네 개의 구간 각각의 관측값들의 개수를 구하면 3, 4, 8, 8로 구해진다. 우린 이제 검정통계량 Q를 계산할 수 있다. 검정통계량 Q는 다음과 같이 구해진다.

$$Q=\frac{(3-23\times 0.25)^2}{23\times 0.25}+\frac{(4-23\times 0.25)^2}{23\times 0.25}+\frac{(8-23\times 0.25)^2}{23\times 0.25}+\frac{(8-23\times 0.25)^2}{23\times 0.25}=3.609$$

χ^2분포표에 의하면, $Q=3.609$에 대한 자유도가 3인 χ^2분포의 분계수는 0.6과 0.7 사이에 위치한다(2.946과 3.665 사이). 따라서 우린 유의수준 0.3 이하에서 귀무가설을 기각할 수 없고, 유의수준 0.4 이상부터는 귀무가설을 기각할 수 있다. (실제로, p-값은 0.307이다.)

$$(\because\ \alpha_0 \ \geq\ 1-X_{k-1}^2(Q)=p-value\ \Rightarrow\ \alpha_0\ \geq\ p-value=\begin{cases}1-0.6 & Let\ Q=2.946,\ Q\in(2.946,3.665)\\ 1-0.7 & Let\ Q=3.665,\ Q\in(2.946,3.665)\end{cases}$$

$$\Rightarrow\ \alpha_0\ \geq\ 0.4\ \text{or}\ 0.3\ \Rightarrow\ \alpha_0\in(0.3,\infty))$$

▲

<u>R Code in example 10.1.6</u>

<u>Results</u>

```
> df_ball_freq
  log1 log2 log3 log4
    3    4    8    8

> chisq.test(df_ball_freq,p=df_ball_prob)

        Chi-squared test for given probabilities

data:  df_ball_freq
X-squared = 3.6087, df = 3, p-value = 0.3069
```

<u>R Code</u>

```
log1=3
log2=4
log3=8
log4=8

df_ball_freq=data.frame(log1,log2,log3,log4)
df_ball_freq
df_ball_prob=c(0.25,0.25,0.25,0.25)

chisq.test(df_ball_freq,p=df_ball_prob)
```

▷ **Additional Problem of χ^2 test**

The data on the number of arrivals of cars at an intersection in 360 10s intervals are as shown in Table. (Mean = 1.013, Variance = 1.1314)

Cars per interval	Number of observations
0	139
1	128
2	55
3	25
4	13

Three models are proposed :

model 1 : $p_X(x) = \dfrac{e^{-1}}{x!}, x = 0, 1, \cdots$

model 2 : $p_X(x) = \dfrac{e^{-1}\lambda^x}{x!}, x = 0, 1, \cdots$

model 3 : $p_X(x) = \binom{x+k-1}{k-1} p^k (1-p)^k, x = 0, 1, \cdots$

(a) Use the χ^2 test; are these models acceptable at the 5% significance level?

(b) In your opinion, which is a better model? Explain your answer.

(Solution)

(a) 우선, 주어진 Table의 관측값들의 총합은 $n=360$으로 구해진다. 각 범주별 확률을 구하자.

① model 1 : $p_X(x)=\dfrac{e^{-1}}{x!}, x=0,1,\cdots$

Cars per interval	Number of observations	Probability of occurrence	Expected frequency
0	139	0.3679	132.444
1	128	0.3679	132.444
2	55	0.184	66.24
3	25	0.0613	22.068
4	13	0.01533	5.5188

χ^2 검정의 검정통계량 Q는 다음과 같이 구해진다.

$$Q=\frac{(132.444-139)^2}{132.444}+\frac{(128-132.444)^2}{132.444}+\frac{(55-66.24)^2}{66.24}+\frac{(25-22.068)^2}{22.068}+\frac{(13-5.5188)^2}{5.5188}=13.088$$

$k=5$이므로, 0.05 유의수준에서 기각값을 구하면 $\chi^2_{0.05}(5-1)=9.49$이다.

∴ 통계량 13.088이 0.05 유의수준에서 기각값 9.49보다 크기 때문에 표본의 분포와 model 1의 분포간에 차이가 없다는 귀무가설을 기각한다. 따라서 model 1의 분포는 표본분포를 반영하지 못한다.

② model 2 : $p_X(x)=\dfrac{e^{-1}\lambda^x}{x!}, x=0,1,\cdots$ (10초에 평균적으로 1.013대 관측된다)

Cars per interval	Number of observations	Probability of occurrence	Expected frequency
0	139	$\dfrac{e^{-1.013}1.013^0}{0!}=0.3631$	130.716
1	128	$\dfrac{e^{-1.013}1.013^1}{1!}=0.3678$	132.408
2	55	$\dfrac{e^{-1.013}1.013^2}{2!}=0.1863$	67.088
3	25	$\dfrac{e^{-1.013}1.013^3}{3!}=0.0629$	22.644
4	13	$\dfrac{e^{-1.013}1.013^4}{4!}=0.0159$	5.724

χ^2 검정의 검정통계량 Q는 다음과 같이 구해진다.

$$Q = \frac{(130.716-139)^2}{130.716} + \frac{(128-132.408)^2}{132.404} + \frac{(55-67.068)^2}{67.068} + \frac{(25-22.644)^2}{22.644} + \frac{(13-5.724)^2}{5.724} = 12.333$$

$k=5$이므로, 0.05 유의수준에서 기각값을 구하면 $\chi^2_{0.05}(5-1)=9.49$이다.

∴ 통계량 12.333이 0.05 유의수준에서 기각값 9.49보다 크기 때문에 표본의 분포와 model 2의 분포간에 차이가 없다는 귀무가설을 기각한다. 따라서 model 2의 분포는 표본분포를 반영하지 못한다.

③ **model 3** : $p_X(x) = \binom{x+k-1}{k-1} p^k (1-p)^k, x=0,1,\cdots$

주어진 model 3을 구체적으로 정리해보자. 우선, k,p는 주어진 데이터로부터 구할 수 있다. model 3의 기댓값(평균) 과 분산의 값을 알고 있으므로, 즉 $E(X) = \frac{k(1-p)}{p} = 1.013$, $Var(X) = \frac{k(1-p)}{p^2} = 1.1314$ 이므로 방정식을 풀면 $k=8.635, p=0.895$로 구해진다.
따라서 발생확률은 $p_X(x) = \binom{x+9-1}{9-1}(0.895)^9(1-0.895)^x$, $x=0,1,\cdots$로 구해진다.

Cars per interval	Number of observations	Probability of occurrence	Expected frequency
0	139	0.3685	132.66
1	128	0.3482	125.352
2	55	0.183	65.88
3	25	0.0704	25.344
4	13	0.0222	7.92

χ^2 검정의 검정통계량 Q는 다음과 같이 구해진다.

$$Q = \frac{(132.66-139)^2}{132.66} + \frac{(128-125.352)^2}{125.352} + \frac{(55-65.88)^2}{65.88} + \frac{(25-25.344)^2}{25.344} + \frac{(13-7.92)^2}{7.92} = 5.423$$

$k=5$이므로, 0.05 유의수준에서 기각값을 구하면 $\chi^2_{0.05}(5-1)=9.49$이다.

∴ 통계량 5.423이 0.05 유의수준에서 기각값 9.49보다 작기 때문에 표본의 분포와 model 3의 분포간에 차이가 없다는 귀무가설을 기각하지 못한다. 따라서 model 3의 분포는 표본분포를 반영한다.

(b) (a)의 논의에 의해, model 3이 5% 유의수준에서 표본분포를 반영하므로, model 3이 적합하다.

▲

R Code in Additional Problem of χ^2 test.

Results

```
> df_model
    car_0 car_1 car_2 car_3 car_4
1     139   128    55    25    13

> chisq.test(df_model,p=df_model_1_prob)

        Chi-squared test for given probabilities

data:  df_model
X-squared = 8.4128, df = 4, p-value = 0.07758

> chisq.test(df_model,p=df_model_2_prob)

        Chi-squared test for given probabilities

data:  df_model
X-squared = 7.8425, df = 4, p-value = 0.09752

> chisq.test(df_model,p=df_model_3_prob)

        Chi-squared test for given probabilities

data:  df_model
X-squared = 2.6249, df = 4, p-value = 0.6224
```

R Code

```
#Variables of Table
car_0=139
car_1=128
car_2=55
car_3=25
car_4=13

#Table
df_model=data.frame(car_0,car_1,car_2,car_3,car_4)
df_model

#Probabilites of model 1,2,3.
df_model_1_prob=c(0.3679,0.3679,0.184,0.0613,0.0189)
df_model_2_prob=c(0.3631,0.3678,0.1863,0.0629,0.0199)
df_model_3_prob=c(0.3685,0.3482,0.183,0.0704,0.0299)

#Confirm the sum of the probability
sum(df_model_1_prob)
sum(df_model_2_prob)
sum(df_model_3_prob)

#Chi^2 - test of three models
chisq.test(df_model,p=df_model_1_prob)
chisq.test(df_model,p=df_model_2_prob)
chisq.test(df_model,p=df_model_3_prob)
```

※ 문제에서 직접 구했던 발생확률의 총합이 1이 되지 않게 구하였으므로(반올림오차, 절단오차 등의 이유로 인한)이를 맞추기 위해 범주 4(car per interval)에 부족한 확률을 더하여 구하였다. 이는 검정통계량 값이 달라질 수 있으나 결과적으로 검정 결과는 본문과 같다. 되도록 발생확률을 정확히 구하는 것이 중요하다. (발생확률은 직접 소프트웨어를 이용하여 구해보자, 모두 특정 분포의 pdf가 되므로 이를 유의하여 함수를 사용하거나 해석적으로 풀이 후 값만 계산하는 방법 모두 활용하여 정확히 구하도록 하자.)