Chapter 7: Bayesian Hypothesis Test

강의 목표

- 고전적 가설 검정의 이해
- 베이지안 가설 검정의 이해
  - One-side Test
  - Two-side Test
  - Multiple Hypotheses Test

Classical Hypothesis Test

- Recall that classical hypothesis testing emphasizes the p-value: The probability (under H_0) that a test statistic would take a value as (or more) favorable to H_a as the observed value of this test statistic.
- For example, given iid data $x = x_1, ..., x_n$ from a density $f(x|\theta)$, where $\infty < \theta < \infty$, we might test $H_0 : \theta \leq 0$ vs. $H_a : \theta > 0$ using some test statistic $T(X)$ (a function of the data).
- For Z-test: $T(X)$ is standardized sample mean.

Classical Hypothesis Test

- Then if we calculated $T(x) = T^*$ for our observed data x, the p-value would be:

$$\begin{aligned} p - value &= \Pr\left(T(X) \geq T^* \mid \theta = 0\right) \\ &= \int_{T^*}^{\infty} f_T(t \mid \theta = 0) dt. \end{aligned}$$

  where $f_T(t \mid \theta)$ is the distribution of $T(X)$.
- This p-value is an average over T values (and thus sample values) that have not occurred and are unlikely to occur.
- Since the inference is based on "hypothetical" data rather than only the observed data, it violates the Likelihood Principle.

Classical Hypothesis Test

- The idea of conducting many repeated tests that motivate "Type I error" and "Type II error" probabilities is not sensible in situations where our study is not repeatable.
- Classical hypothesis testing mostly focus on "Type I error" rather than "Type II error". For some cases, "Type II error" is more important than "Type I error".
- Classical hypothesis testing cannot deal with multiple hypotheses. It can only test with two hypotheses.

The Bayesian Approach

- A simple approach to testing finds the posterior probabilities that θ falls in the null and alternative regions.
- We first consider one-sided tests about θ of the form:

$$H_0 : \theta \leq c \quad vs. \quad H_a > c$$

  for some constant c, where $-\infty < \theta < \infty$.
- We define Θ_0 and Θ_1.
- Θ_0 is the set of θ-values such that H_0 is true.
- Θ_1 is the set of θ-values such that H_1 is true.

Prior/Posterior Probabilities

- We specify prior probabilities for θ such that

$$\begin{aligned} \pi_0 &:= \Pr(-\infty < \theta \leq c) = P(\theta \in \Theta_0) \\ \pi_1 &:= \Pr(c < \theta \leq \infty) = P(\theta \notin \Theta_0). \end{aligned}$$

- We specify posterior probabilities for θ such that

$$\begin{aligned} p_0 \mid x &:= \Pr(-\infty < \theta \leq c \mid x) = P(\theta \in \Theta_0 \mid x) \\ p_1 \mid x &:= \Pr(c < \theta \leq \infty \mid x) = P(\theta \notin \Theta_0 \mid x). \end{aligned}$$

Posterior Odds

- Posterior probabilities for θ can be re-written as:

$$\begin{aligned} p_0 \mid x &:= \int_{\Theta_0} \pi(\theta \mid x) = \frac{\int_{\Theta_0} f(x \mid \theta)\pi(\theta)d\theta}{\int_{\Theta_0} f(x \mid \theta)\pi(\theta)d\theta + \int_{\Theta_1} f(x \mid \theta)\pi(\theta)d\theta} \\ p_1 \mid x &:= \int_{\Theta_1} \pi(\theta \mid x) = \frac{\int_{\Theta_1} f(x \mid \theta)\pi(\theta)d\theta}{\int_{\Theta_0} f(x \mid \theta)\pi(\theta)d\theta + \int_{\Theta_1} f(x \mid \theta)\pi(\theta)d\theta} \end{aligned}$$

- Note that the denominators are the same constant, not a function of θ.
- Hence the posterior odds is as follows:

$$\frac{p_0}{p_1} \mid x = \frac{\Pr(H_0 \mid x)}{\Pr(H_1 \mid x)} = \frac{\int_{\Theta_0} f(x \mid \theta)\pi(\theta)d\theta}{\int_{\Theta_1} f(x \mid \theta)\pi(\theta)d\theta}$$

Prior/Posterior Odds

- Prior odds can be understood as in favor of H_0 against H_a

$$\text{Prior odds} = \frac{\pi_0}{\pi_1}$$

- Posterior odds can be understood as in favor of H_0 against H_a

$$\text{Posterior odds} = \frac{p_0}{p_1} \mid x$$

- A posterior odds may be similar to a prior odds because a posterior probability highly depends on its prior probability.
- Need a measure for the weight of evidence in favor of H_0.

Bayes Factor

- The Bayes Factor provides a measure for comparison of two competing models, say M_1 and M_2.
- It is similar to testing a "full model" vs. "reduced model" (with, e.g., a likelihood ratio test) in classical statistics.
- However, with the Bayes Factor, one model does not have to be nested within the other.
- Given a data set $x = x_1, ..., x_n$, we compare models

$$M_1 : f_1(x \mid \theta_1) \quad \text{vs} \quad M_2 : f_2(x \mid \theta_2)$$

- We may specify prior distributions $\pi_1(\theta_1)$ and $\pi_2(\theta_2)$ that lead to prior probabilities for each model $\Pr(M_1)$ and $\Pr(M_2)$.

Bayes Factor

- If the prior model probabilities are equal, i.e., $\pi(M_1) = \pi(M_2)$, then the Bayes Factor equals the posterior odds for M_1.
- If $\pi(M_1) = \pi(M_2)$ and the parameter spaces Θ_1 and Θ_2 are the same, then the Bayes Factor reduces to a likelihood ratio.

$$B(x) = \frac{\Pr(M_1 \mid x)}{\Pr(M_2 \mid x)} \times \frac{\pi(M_2)}{\pi(M_1)} = \frac{\Pr(x \mid M_1)}{\Pr(x \mid M_2)}$$

Bayes Factor: Jeffreys Rule

- A Bayes Factor greater than 1 supports Model 1 over Model 2.
- Jeffreys proposed the following rules:
  - If $B(x) \geq 1$, Model 1 supported.
  - If $\sqrt{(10)} \leq B(x) \leq 1$, minimal evidence against Model 1
  - If $0.1 \leq B(x) \leq \sqrt{(10)}$, substantial evidence against Model 1
  - If $0.01 \leq B(x) \leq 0.1$, strong evidence against Model 1
  - If $B(x) \leq 0.01$, decisive evidence against Model 1

Bayes Factor: Weighted Likelihood Ratio

- Suppose that $H_0 : \theta \in \Theta_0$ vs. $H_a : \theta \in \Theta_1$.
- Bayes Factor can be restated as

$$
\begin{aligned}
B(x) &= \frac{p_0/p_1}{\pi_0/\pi_1} = \frac{\int_{\Theta_0} f(x \mid \theta)\pi(\theta)d\theta / \int_{\Theta_0} \pi(\theta)d\theta}{\int_{\Theta_1} f(x \mid \theta)\pi(\theta)d\theta / \int_{\Theta_1} \pi(\theta)d\theta} \\
&= \frac{\int_{\Theta_0} f(x \mid \theta)\frac{\pi(\theta)}{\int_{\Theta_0} \pi(\theta)d\theta} d\theta}{\int_{\Theta_1} f(x \mid \theta)\frac{\pi(\theta)}{\int_{\Theta_1} \pi(\theta)d\theta} d\theta} \\
&= \frac{\mathbb{E}\left[\pi(\theta)I(\theta \in \Theta_0)/ \int_{\Theta_0} \pi(\theta)d\theta\right]}{\mathbb{E}\left[\pi(\theta)I(\theta \in \Theta_1)/ \int_{\Theta_1} \pi(\theta)d\theta\right]}.
\end{aligned}
$$

Bayes Factor: Weighted Likelihood Ratio

- 이는 가중치 $\pi(\theta)I(\theta \in \Theta_i)/\int_{\Theta_i}\pi(\theta)d\theta$를 사용한 우도함수 (likelihood function)의 가중 평균이다.
- 즉, 가중우도비(weighted likelihood ratio)라고 할 수 있다.

Bayes Factor: Classical Likelihood Ratio

- 만약 $\Theta_0 = \{\theta_0\}$, $\Theta_1 = \{\theta_1\}$이면,

$$B(x) = f(x \mid \theta_0)/f(x \mid \theta_1).$$

since $p_0/p_1 = \pi_0 f(x \mid \theta_0)/\pi_1 f(x \mid \theta_1)$.

- 가중치 (weight) $\pi(\theta)I(\theta \in \Theta_i)/\int_{\Theta_i} \pi(\theta)d\theta = I(\theta \in \Theta_i)$가 되므로, 이 상황에서는 bayes factor는 likelihood ratio와 동일하다.

Bayes Factor: Marginal Likelihood Ratio

- 가중치는 θ의 공간이 Θ_i로 제한되었을 때 (truncated), 즉 H_i 하에서 θ의 사전밀도함수 (density function).

$$\pi_i(\theta) = \frac{\pi(\theta) I(\theta \in \Theta_i)}{\int_{\Theta_i} \pi(\theta) d\theta}.$$

- Bayes Factor can be $m_0(x)/m_1(x)$ where

$$m_i(x) := \int f(x \mid \theta) \pi_i(\theta) d\theta.$$

- 베이즈 상수 가 각 가설 하에서 관측치의 주변밀도힘수값의 비율로 주어지므로 베이즈 상수를 주변우도비(marginal likelihood ratio)라고도 한다.

Example

Suppose we have a random variable that produces either a success or a failure. We want to compare a model M1 where the probability of success is $\theta = \frac{1}{2}$, and another model M2 where θ is unknown and we take a prior distribution for θ that is uniform on $[0, 1]$. We take a sample of $n = 200$, and find $x = 115$ successes and $n - x = 85$ failures. The likelihood can be calculated according to the binomial distribution:

$$L(\theta \mid x) = \binom{200}{115} \theta^{115} (1 - \theta)^{85}.$$

Example: Bayesian Hypothesis Test

Thus we have

$$P(X = 115 \mid M_1) = \binom{200}{115} \left(\frac{1}{201}\right)^{200} = 0.005956,$$

but

$$P(X = 115 \mid M_2) = \int_0^1 \binom{200}{115} \theta^{115}(1-\theta)^{85} d\theta = \frac{1}{201} = 0.004975.$$

The ratio is then 1.197, which is barely worth mentioning even if it points very slightly towards M1.

Example: Classical Hypothesis Test

- P-value: the probability of getting 115 or more successes from a sample of 200 if $\theta = \frac{1}{2}$ is 0.0200, and as a two-tailed test of getting a figure as extreme as or more extreme than 115 is 0.0400

$$\Pr(X \geq 115 \mid \theta = 0.5) \approx \Pr(Z \geq \frac{(115 - 100)}{\sqrt{50}}) = 0.17$$

- Classical hypothesis test would yield significant results at the 5% significance level.
- However Bayes factor hardly considers this to be an extreme result.

Example: Remarks

- Note that a non-uniform prior could result in a Bayes factor that is more in agreement with the frequentist hypothesis test.
- A classical likelihood-ratio test would have found the MLE for θ (M_2: $\theta = \frac{115}{200} = 0.575$) where

$$P(X = 115 \mid M_2) = \binom{200}{115} \theta^{115} (1 - \theta)^{85} = 0.056991.$$

That gives a Bayes factor of 0.1045, and so pointing towards M_2.

Example: Remarks

- M2 is a more complex model than M1 because it has a free parameter θ which allows it to model the data more closely.
- The ability of Bayes factors to take this into account is a reason why Bayesian inference has been put forward as a theoretical justification for reducing Type I errors.

Another Example

$X \mid \theta \sim N(\theta, 1)$일 때 검정

$$H_0 : \theta = 1, \quad vs \quad H_a : \theta = -1$$

과 위 검정에서 가설을 바꾼,

$$H_0 : \theta = -1, \quad vs \quad H_a : \theta = 1$$

을 생각해 보자.

Bayes Factor: H_0 vs H_1

$x = 0$이 관측되었을 때 H_0와 H_1의 검정에 대한 베이즈 상수를 구하면

$$B(x) = \exp\left[-\frac{1}{2}\left((0-1)^2 - (0+1)^2\right)\right] = e^0 = 1.$$

즉 $x = 0$의 관측치는 어느 가설도 지지 하는 않는다.

Bayes Factor: H_1 vs H_0

$x = 0$이 관측되었을 때 H_1와 H_0의 검정에 대한 베이즈 상수를 구하면

$$B(x) = \exp\left[-\frac{1}{2}\left((0+1)^2 - (0-1)^2\right)\right] = e^0 = 1.$$

귀무가설과 대립가설을 바꾸더라도, $x = 0$의 관측치는 어느 가설도 지지 하는 않으며 똑같은 베이즈 상수를 가진다.

Classical Hypothesis Test: H_0 vs H_1

$x = 0$이 관측되었을 때 H_0와 H_1의 검정에 대한 유의확률 (p-value)를 구하면

$$\text{p-value} = \Pr(X < 0 \mid X \sim N(1,1)) = 0.1587.$$

즉 $x = 0$는 주어진 귀무가설에서 극단적인 케이스는 아니다.

Classical Hypothesis Test: H_1 vs H_0

$x = 0$이 관측되었을 때 H_1와 H_0의 검정에 대한 유의확률 (p-value) 를 구하면

$$\text{p-value} = \Pr(X > 0 \mid X \sim N(-1, 1)) = 0.1587.$$

귀무가설과 대립가설을 바꾸더라도, $x = 0$는 같은 p-value를 가진다.

일반적으로 생각하는 좋은 테스트란, 주어진 데이터를 통해 귀무가설과 대립가설의 순서와 관계 없이 일관성있게 결론을 주는 것이라고 생각된다. 이 경우 Bayesian & Classical hypothesis test는 같은 결과를 준다.

Bayes Factor: H_0 vs H_1

$x = 1$이 관측되었을 때 H_0와 H_1의 검정에 대한 베이즈 상수를 구하면

$$B(x) = \exp\left[-\frac{1}{2}\left((1-1)^2 - (1+1)^2\right)\right] = e^2 = 7.38.$$

즉 $x = 1$의 관측치는 귀무가설 $(\theta = 1)$을 강력히 지지한다.

Bayes Factor: H_1 vs H_0

$x = 0$이 관측되었을 때 H_1와 H_0의 검정에 대한 베이즈 상수를 구하면

$$B(x) = \exp\left[-\frac{1}{2}\left((1+1)^2 - (1-1)^2\right)\right] = e^{-2} = 1/7.38.$$

즉 $x = 1$의 관측치는 대립가설 $(\theta = 1)$을 강력히 지지한다.

Classical Hypothesis Test: H_0 vs H_1

$x = 1$이 관측되었을 때 H_0와 H_1의 검정에 대한 유의확률 (p-value)를 구하면

$$\text{p-value} = \Pr(X < 1 \mid X \sim N(1,1)) = 0.50.$$

즉 $x = 0$는 주어진 귀무가설에서 극단적인 케이스는 아니다.

Classical Hypothesis Test: H_1 vs H_0

$x = 0$이 관측되었을 때 H_1와 H_0의 검정에 대한 유의확률 (p-value)를 구하면

$$\text{p-value} = \Pr(X > 1 \mid X \sim N(-1, 1)) = 0.025.$$

귀무가설과 대립가설을 바꾸면, 전혀 관계없는 p-value를 가진다.
Baeysiand test 가설의 순서에 의존 하지 않지만, classical test는 귀무가설과 대립가설의 순서에 결론이 바뀔 수 있다.

Bayesian One-sided Test

- Suppose that

$$H_0 : \theta \leq \theta_0, \quad vs \quad H_a : \theta > \theta_0.$$

- Suppose that

$$X_1, X_2, ..., X_n \mid \theta \sim_{iid} N(\theta, \sigma^2).$$

- Then,

$$\Theta_0 = (-\infty, \theta_0], \quad \Theta_0 = (\theta_0, -\infty).$$

Bayesian One-sided Test

- Suppose that the prior distribution is $\pi(\theta) \propto 1$,
- Then $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ has the following distribution:

$$\bar{X} \sim N(\theta, \sigma^2/n).$$

- Posterior distribution is

$$P(\theta \mid X_1, X_2, ..., X_n) \sim N(\bar{x}, \sigma^2/n).$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Bayesian Test for Nomral Case

- Find the posterior probability for H_0:

$$P(\theta \leq \theta_0 \mid \bar{x}) = \Phi\left(\frac{\theta_0 - \bar{x}}{\sigma/\sqrt{n}}\right).$$

where $\Phi(.)$ is the CDF of standard normal distribution.

Classical Test for Nomral Case

- Find the p-value of the classical test.

$$P(\bar{X} \geq \bar{x} \mid \theta_0) = 1 - \Phi\left(\frac{\bar{x} - \theta_0}{\sigma/\sqrt{n}}\right).$$

- Hence the p-value is the same as the prosterior probability for H_0.

Example

W라고 하는 새로운 입자의 질량을 알아보고자 한다. 여러 번의 독립적인 측정을 통하여 얻은 입자의 평균 질량은 82.1이고 추정오차는 1.7이었다. 입자의 질량이 83.0보다 큰 지 검정하고자 한다. 이 경우 다음 가설을 세울 수 있다:

$$H_0 : \theta \leq 83.0 \quad vs \quad H_1 : \theta > 83.0.$$

Example: Bayesian Approach

충분한 수의 독립적인 반복 측정이 이루어졌다고 보고 분산 $\sigma^2/n = 1.7^2$이 알려진 경우로 간주하자. 베이지안 검정을 위해 먼저 θ에 대한 사전분포로 무정보 사전분포 $\pi(\theta) = 1$을 가정하면 θ의 사후분포는 $N(82.1, 1.7^2)$ 이며, H_0의 사후확률은 $P_0 = \Phi((83.0 - 82.1)/1.7) = 0.7017$이다. H_0 와 H_1의 위험도를 동등하게 본다면 $0.7017 > 0.5$이므로 귀무가설 을 채택한다.

Example: Another Bayesian Approach

W라는 입자의 성질에 대한 물리적 이론에 의하면 이 입자의 질량은 82.4이며 추정오차는 1.1이라고 한다. 이 물리적 이론으로부터 얻은 정보를 사전정보로 시용하여 θ의 사전분포로 $N(82.4, 1.1^2)$을 가정할 수 있다. 사전분포로부터 H_0의 사전확률을 구하면

$$\pi_0 = \mathrm{Pr}(\theta \leq 83.0) = \Phi((83.0 - 82.4)/1.1) = 0.7088.$$

그리고 $\pi_1 = 1 - \pi_0 = 0.2912$이므로 H_0 와 H_1의 사전비는 2.64 이다.

Example: Another Bayesian Approach

사전분포와 관측치로부터 θ의 사후분포를 구하면

$$\theta \sim N(82.3, 0.92^2)$$

- The posterior expectation:

$$\frac{\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}} = \frac{\sigma^2}{\sigma^2 + n\tau^2}\delta + \frac{n\tau^2}{\sigma^2 + n\tau^2}\bar{x}.$$

- The posterior variance:

$$\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1}$$

Example: Another Bayesian Approach

그러므로 사후 확률은 다음과 같다.

$$
\begin{aligned}
p_0 &= P(\theta \le 83.0 \mid \bar{x}) = \Phi((83.0 - 82.3)/0.92) = 0.7764, \\
p_1 &= 1 - p_0 = 0.2236.
\end{aligned}
$$

으로 사후비는 3.47이다. H_0 와 H_1의 위험도를 동등하게 본다면 귀무가설을 채택할 것이다. 그런데 이 경우 베이즈 상수 (Bayes factor)는 $B(x) = 3.47/2.64 = 1.43$으로 귀무가설에 사후 지지도가 큰 주원인은 사전 지지도가 크기 때문이며, 자료가 주는 지지도는 1.43으로 사전 지지도에 비하여 미미함을 알 수있다.

Bayesian One-sided Test for Beta Case

- Suppose that

$$H_0 : \theta \leq \theta_0, \quad vs \quad H_a : \theta > \theta_0.$$

- Suppose that

$$X \mid \theta \sim_{iid} Bin(n, \theta).$$

- Suppose that the prior distribution is $Beta(a, b)$.
- Then the posterior distribution is $Beta(a + x, b + n - x)$.

Bayesian Test for Beta Case

- Find the posterior probability for H_0:

$$P(\theta \leq \theta_0 \mid x) = \int_0^{\theta_0} \frac{\Gamma(a+b+n)}{\Gamma(a+x)\Gamma(b+n-x)} \theta^{a+x-1}(1-\theta)^{b+n-x} d\theta$$

Classical Test for Beta Case

- Find the p-value of the classical test.

$$P(\bar{X} \geq \bar{x} \mid \theta_0) \approx 1 - \Phi\left(\frac{\bar{x} - \theta_0}{\sqrt{\theta_0(1 - \theta_0)/n}}\right).$$

Example

어느 도시의 시장을 선출하는 선거에서 두 후보 A와 B가 경쟁하고 있다. 최근 신문에 발표된 여론조사에 의하면 100명의 랜덤으로 선택된 선거권자들 중 53명이 A 후보를 지지하고 47명은 B 후보를 지지한다고 한다. 전체 선거권지들 중 A 후보를 지지하는 사람들의 비율을 θ라 할 때 θ에 대해 균일 사전분포를 가정하고

$$H_0 : \theta \leq 0.5 \quad vs \quad H_1 : \theta > 0.5$$

를 검정해 보자.

Example

$n = 100, x = 53, a = 1, b = 1$인 경우이므로 θ의 사후분포는 $Beta(54, 48)$ 이다. 따라서

$$p_0 = \Pr(\theta \le 0.5 \mid x) = \Pr(\theta \le 0.5 \mid \theta \sim Beta(54, 48)) = 0.275$$

이고 $p_1 = 1 - p_0 = 0.725$이다. 베타분포에 대한 정규근사를 사용하면 $Beta(54, 48) \approx N(0.529, 0.049^2)$ 이므로

$$p_0 = \Pr(\theta \le 0.5 \mid \theta \sim N(0.529, 0.049^2)) = \Phi\left(\frac{0.5 - 0.529}{0.049}\right) = 0.277$$

이고 $p_1 = 0.723$으로 베타 사후분포를 시용했을 때와 유사한 값을 얻는다.

Example

전체 모집단에서 A 후보의 지지율이 50% 이상일 사후확률은 약 72%이므로 H_1이 H_0보다 약 2.57배 더 높은 가능성을 지닌다. 관측된 A 후보의 지지도 53%는 당선 기준 50% 보다 약간 큰 정도이지만 당선될 ($\theta > 0.5$) 사후확률은 72.3%로 상당히 크다. 이는 표본크기가 커서 추정오차가 작기 때문이다.

Example

Classical hypothesis test를 시행하면

$$\text{p-value} \;= P(\bar{X} \geq \bar{x} \mid \theta_0) \approx 1 - \Phi\left(\frac{0.530 - 0.5}{\sqrt{1/4 \times 1/100}}\right) = 0.274.$$

이 경우, 관측값들이 H_0를 기각하는데 큰 증거가 되지 못한다.

Bayesian Two-sided Test

- Suppose that

$$H_0 : \theta = \theta_0, \quad vs \quad H_a : \theta \neq \theta_0.$$

- Suppose that

$$X_1, X_2, ..., X_n \mid \theta \sim_{iid} N(\theta, \sigma^2).$$

- Then,

$$\Theta_0 = \{\theta_0\}, \quad \Theta_0 = (-\infty, \infty) \setminus \{\theta_0\}.$$

Bayesian Two-sided Test

- Suppose that the prior distribution is $\pi(\theta) \propto 1$,
- Then $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ has the following distribution:

$$\bar{X} \sim N(\theta, \sigma^2/n).$$

- Posterior distribution is

$$P(\theta \mid X_1, X_2, ..., X_n) \sim N(\bar{x}, \sigma^2/n).$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Bayesian Test for Nomral Case

- Find the posterior probability for H_0:

$$P(\theta = \theta_0 \mid \bar{x}) = 0.$$

- 이 경우 사후 확률이 0이므로, 항상 H_0이 기각되므로 좋은 테스트라고 할 수 없다 .
- 그래서 유의미한 테스트를 할 수 있게 만들어 주는, prior distribution을 찾도록 한다.

Prior Distribution for Two-sided Test

- 베이지안 검정에서는 이를 피하기 위해 다음과 같은 연속과 이산의 합성 사전분포를 가정한다.

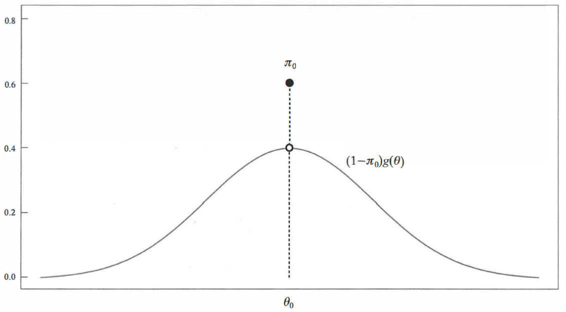
$$\pi(\theta) = \begin{cases} \pi_0 & \text{if } \theta = \theta_0, \\ (1-\pi_0)g(\theta) & \text{if } \theta \neq \theta_0. \end{cases}$$

  where $g(\theta)$ is any probability density function.

- 귀무가설에서 지정하는 한 점에서만 확률 π_0를 배정하고, 나머지 확률 $\pi_1 = 1 - \pi_0$를 θ_0를 제외한 나머지 연속 구간에 밀도함수 $g(\theta)$에 따라 재배분하는 것이다.

Prior Distribution for Two-sided Test

- 만약 $g(\theta)$가 정규분포의 density function이라면, 다음과 같은 사전 분포를 찾을 수 있다.



Posterior Distribution

- Then the posteiror distribution is as follows:

$$\Pr(\theta \mid x) = \frac{\pi_0 f(\bar{x} \mid \theta_0) I(\theta = \theta_0) + \pi_1 g(\theta) f(\bar{x} \mid \theta) I(\theta \neq \theta_0)}{m(\bar{x})}$$

where

$$m(\bar{x}) = \pi_0 f(x \mid \theta_0) + \int_{\theta \neq \theta_0} \pi_1 g(\theta) f(\bar{x} \mid \theta) d\theta.$$

- Note that the second term is

$$\int_{\theta \neq \theta_0} \pi_1 g(\theta) f(\bar{x} \mid \theta) d\theta = \int \pi_1 g(\theta) f(\bar{x} \mid \theta) d\theta.$$

since $g(\theta_0) = 0$.

Bayes Factor

- Posterior probability for the null hypothesis:

$$p_0 = \Pr(H_0 \mid \bar{x}) = \frac{\pi_0 f(\bar{x} \mid \theta_0)}{m(\bar{x})}.$$

- Posterior probability for the alternative hypothesis:

$$p_1 = \Pr(H_1 \mid \bar{x}) = \frac{\pi_1 \int g(\theta) f(\bar{x} \mid \theta) d\theta}{m(\bar{x})}.$$

- Bayes Factor:

$$B(x) = \frac{p_0/p_1}{\pi_0/\pi_1} = \frac{f(\bar{x} \mid \theta_0)}{\int g(\theta) f(\bar{x} \mid \theta) d\theta}.$$

Bayes Factor with Mixed Normal Prior

- Suppose that $g(\theta)$ is mixed Normal distribution $N(\theta_0, \sigma_0^2)$.
- Then,

$$\bar{X} \sim N(\theta_0, \sigma^2/n + \sigma_0^2).$$

- Hence

$$\int g(\theta) f(\bar{x} \mid \theta) d\theta = \frac{1}{\sqrt{2\pi(\sigma^2/n + \sigma_0^2)}} \exp\left\{-\frac{(\bar{x} - \theta)^2}{2(\sigma^2/n + \sigma_0^2)}\right\}.$$

Bayes Factor with Mixed Normal Prior

- Bayes Factor:

$$
\begin{aligned}
B(x) &= \frac{f(\bar{x} \mid \theta_0)}{\int g(\theta) f(\bar{x} \mid \theta) d\theta} \\
&= \sqrt{\frac{\sigma^2/n + \sigma_0^2}{\sigma^2/n}} \exp\left\{-\frac{(\bar{x}-\theta)^2}{2(\sigma^2/n + \sigma_0^2)}\right\} \\
&= \sqrt{1 + \frac{\sigma_0^2}{\sigma^2/n}} \exp\left\{-\frac{(\bar{x}-\theta)^2}{2\sigma^2/n}\left(1 - \frac{\sigma^2/n}{\sigma^2/n + \sigma_0^2}\right)\right\} \\
&= \sqrt{1 + \frac{\sigma_0^2}{\sigma^2/n}} \exp\left\{-\frac{1}{2} z^2 \left(1 - \frac{\sigma^2/n}{\sigma^2/n + \sigma_0^2}\right)\right\}
\end{aligned}
$$

Bayes Factor with Mixed Normal Prior

- 사전정보의 양을 한 개의 표본이 주는 정보의 양과 동일하게 여겨 $\sigma_0^2 = \sigma^2$이라 놓으면,

$$B(x) = \sqrt{1+n}\exp\left(-\frac{1}{2}z^2\frac{n}{n+1}\right).$$

- 베이즈 상수는 변수 z이외에도 샘플수 n에도 의존하고 있다.
- 반면 classical test는 z에만 변하고 샘플 수에는 영향을 받지 않는다 (n은 z값 만 변화 시킨다).
- n이 커지면 커질 수록 베이즈 상수도 증가하고, 귀무가설 (H_0)의 posterior porbability가 증가한다.

Bayesian v.s. Classical Method

- $X_1, ..., X_n \mid \theta \sim N(\theta, 1)$를 따를 때,

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta \neq \theta_0$$

를 검정하고자 한다. 자료로부터 $z = \frac{(\bar{x}-\theta)^2}{\sigma/\sqrt{n}}$ 을 계산하니 $z = 2$이었다.

- 만약 $n = 100$이라면 $\bar{x} - \theta_0 = 2/\sqrt{n} = 0.2$ 이고, $n = 10^8$ 이라면 $\bar{x} - \theta_0 = 2/\sqrt{n} = 0.0002$ 이다.
- $n = 100$ 일 경우에는 θ와 θ_0의 절대적 차이가 0.2이지만 $n = 10^8$라면 절대적 차이가 0.0002라는 것이다.

Bayesian v.s. Classical Method

- 사전분포로 $\pi_0 = \pi_1 = 0.5$와 $g(\theta) \sim N(\theta_0, 1)$을 가정하자.
- $n = 100$일 경우 베이즈 상수는 $B(x) \approx \sqrt{101}e^{-2} = 1.36$이고 $p_0 = 0.54$이다.
- $n = 10^8$일 경우 베이즈 상수는 $B(x) \approx \sqrt{10^8 + 1}e^{-2} = 1353.35$이고 $p_0 = 0.99$이다.
- 따라서 $n = 10^8$일 때 H_0에 대한 더 강한 지지를 주는데, 이는 z값이 같더라도 n^{10^8} 일 때 $\bar{x} - \theta_0$의 절대적 차이가 더 작은 것을 베이즈 상수가 반영하기 때문이다.

Bayesian v.s. Classical Method

하지만 고전적 검정에서는 샘플수와 관계 없이 모두 유의확률은 동일 하게 약 5%이다. 즉, $\bar{x} - \theta_0$의 절대적 차이가 표준오차의 몇 배인지만 고려할 뿐 절대적 차이는 고려하지 않는다.

Hypothesis Test for Binomial Case

- Suppose that $X \mid \theta \sim B(n, \theta)$.
- Consider

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta \neq \theta_0.$$

- Prior Distribution for θ:

$$\pi(\theta) = \begin{cases} \pi_0 & \theta = \theta_0 \\ (1 - \pi_0) g(\theta) & , \theta \neq \theta_0 \end{cases}$$

where $g(\theta)$는 $Beta(\alpha, \beta)$의 확률밀도함수이다.

Hypothesis Test for Binomial Case

- Bayes Factor:

$$B(x) = \frac{f(x \mid \theta_0)}{\int f(x \mid \theta) g(\theta) d\theta}$$

- Let $m_1(x) = \int f(x \mid \theta) g(\theta) d\theta$ be the density function for x.

$$\begin{aligned}
m_1(x) &= \int_0^1 \binom{n}{x} \theta^x (1-\theta)^{n-x} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1} d\theta \\
&= \binom{n}{x} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \theta^{\alpha+x-1}(1-\theta)^{\beta+n-x-1} d\theta \\
&= \binom{n}{x} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+x)\Gamma(\beta+n-x)}{\Gamma(\alpha+\beta+n)}
\end{aligned}$$

Hypothesis Test for Binomial Case

- Bayes Factor:

$$
\begin{aligned}
B(x) &= \frac{\binom{n}{x}\theta_0^x(1-\theta_0)^{n-x}}{\binom{n}{x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+x)\Gamma(\beta+n-x)}{\Gamma(\alpha+\beta+n)}} \\
&= \frac{\theta_0^x(1-\theta_0)^{n-x}}{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+x)\Gamma(\beta+n-x)}{\Gamma(\alpha+\beta+n)}}
\end{aligned}
$$

Example

어떤 정밀 부품을 만드는 회사에서 부품의 불량률이 10%인지 검정하고자 하여 생산 라인에서 50개의 부품을 랜덤 추출하여 조사하니 3개가 불량이었다. 실제 불량률 θ 에 대하여 사전분포로 다음과 같이 가정하자.

$$\pi(\theta) = \begin{cases} 0.5 & \theta = 0.1 \\ 0.5g(\theta) & ,\theta \neq 0.1 \end{cases}$$

$g(\theta)$는 $Beta(1,9)$의 확률밀도함수이다.

Example

다음 가설을 검정해보자.

$$H_0 : \theta = 0.1 \quad vs \quad H_1 : \theta \neq 0.1.$$

$n = 50, x = 3, \alpha = 1, \beta = 9$인 경우이므로

$$m_1(x) = \binom{50}{3} \frac{\Gamma(10)}{\Gamma(1)\Gamma(9)} \frac{\Gamma(4)\Gamma(56)}{\Gamma(60)}.$$

베이즈상수의 경우

$$B_x = \frac{f(x \mid \theta_0)}{m_1(x)} = \frac{(0.1)^3(0.9)^{47}}{\frac{\Gamma(10)}{\Gamma(1)\Gamma(9)} \frac{\Gamma(4)\Gamma(56)}{\Gamma(60)}} = 1.43.$$

Example

이로부터 $p_0 = P(H_0 \mid x) = 0.588$임을 알 수 있다. H_0와 H_1의 위험도를 동등하게 본다면 $H_0 : \theta = 0.1$ 를 채택할 것이다. 그러나 이 경우 사후정보의 H_0에 대한 지지도가 아주 뚜렷하지는 않다.

Hypothesis Test for Poisson Case

- Suppose that $X_1, \ldots, X_n \sim_{iid} Poi(\theta)$.
- Consider

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta \neq \theta_0.$$

- Prior distribution for θ:

$$\pi(\theta) = \begin{cases} 0.5 & , \theta = \theta_0 \\ 0.5g(\theta) & , \theta \neq \theta_0 \end{cases}$$

where $g(\theta)$ is $Gamma(\alpha, \beta)$ density function.

Hypothesis Test for Poisson Case

- Bayes Factor:

$$B(x) = \frac{f(x \mid \theta_0)}{\int f(x \mid \theta) g(\theta) d\theta}$$

- $m_1(x)$ be the density function for x.

$$\begin{aligned}
m_1(x) &= \int_0^\infty \frac{e^{-n\theta}(n\theta)^x}{x!} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} d\theta \\
&= \frac{n^x \beta^\alpha}{x!\Gamma(\alpha)} \int_0^\infty \theta^{\alpha+x-1} d\theta \\
&= \frac{n^x \beta^\alpha}{x!\Gamma(\alpha)} \frac{\Gamma(\alpha+x)}{(n+\beta)^{\alpha+x}} \\
&= \frac{(\alpha+x-1)!}{x!(\alpha-1)!} \left(\frac{\beta}{n+\beta}\right)^\alpha \left(\frac{n}{n+\beta}\right)^x .
\end{aligned}$$

Hypothesis Test for Poisson Case

- $m_1(x)$ 보면 성공확률이 $\frac{\beta}{n+\beta}$인 독립적인 베르누이 시행을 α 번의 성공이 나올 때 까지 반복할 경우 x는 총 실패횟수를 나타낸다.
- 그러므로 $x = \sum_{i=1}^{n} x_i$의 주변분포는 $NB\left(\alpha, \frac{\beta}{n+\beta}\right)$임을 알 수 있다.

Traffic Example

두 도시 에서 차량통행량 등 주위의 교통환경이 비슷한 교차로를 하나씩 선택하여 매주 발생한 교통사고 건수 를 1년 동안 조사하였다. 첫 번째 도시 에서는 직진 후 좌회전 신호를 사용하고 두 번째 도시 에서는 좌회전 후 직진 신호를 사용한다. 교통사고 건수는 독립적으로 포아송 분포를따 른다고 가정한다. 교통통제 등의 이유로 조사를 할 수 없었던 기간을 제외하고 다음 표와 같은 조사 결과를 얻었다.

Traffic Example

교통사고 건수	0	1	2	3	4	5	6	7
City 1의 사고 건수	7	14	13	8	4	2	2	0
City 2의 사고 건수	4	13	15	6	2	2	3	1

$$n_1 = 50, \quad \sum x_{1i} = 102, \quad \bar{x}_1 = 2.04$$

$$n_2 = 46, \quad \sum x_{2i} = 104, \quad \bar{x}_1 = 2.26$$

Traffic Example

- 첫번째 도시의 교통사고 건수에만 집중하도록 하자.
- 이제 $H_0 : \theta = 2 \quad vs \quad H_1 : \theta \neq 2$를 검정해 보자
- 첫번째 도시의 실제 평균 교통사고 건수 θ 에 대한 사전 분포를 다음과 같이 가정 하자. $Gamma(2,1)$의 사전분포를 가정하자.

$$\pi(\theta) = \begin{cases} 0.5 & \theta = 2 \\ 0.5g(\theta) & ,\theta \neq 2 \end{cases}$$

$g(\theta)$는 $Gamma(2,1)$ 분포의 확률밀도함수이다.

Traffic Example

- $n = 50, x = \sum_{i=1}^{50} x_i = 102, \alpha = 2, \beta = 1.$
- $f(x \mid \theta_0)$는 $x = 102$에서 $Poi(2)$의 확률밀도함수.
- $m_1(x)$는 $x = 102$에서 $NB\left(2, \frac{1}{51}\right)$의 확률밀도힘수.
- 베이즈 상수 $B_x = 7.364$. 따라서 $P_0 = 0.8804$이다. $H_0 : \theta = 2$ 에 대한 지지도가 매우 뚜렷하다.

```
a=2; b=1 ; n=50; x=102 ; theta0=2
dpois(x,n*theta0)/dnbinom(x,a,b/(n+b))
7.364568
```

Posterior Probability for Multiple Hypotheses

- If there are multiple hypotheses $e.g., K$, there are K distinct sets $\Theta_0, ..., \Theta_{K-1}$.
- Then, Posterior probabilities for θ are

$$p_i \mid x = \Pr(H_i \mid x) \propto \int_{\Theta_i} f(x \mid \theta)\pi(\theta)d\theta$$

Example for Multiple Hypotheses Test

- 찬이의 IQ 점수는 $N(\theta, 100)$을 따른다고 한다.
- 찬이가 5번 IQ 테스트 (X_i)를 보았고, 평균점수가 115였다.
- 이 경우 $X \mid \theta \sim N(\theta, 100/5)$이며 $\bar{x} = 115$의 관측치를 사용할 수 있겠다. θ의 사전분포를 정하는데, 우리나라 전체 인구의 IQ 점수는 $N(100, 225)$를 따른다고 알려져 있다.

Example for Multiple Hypotheses Test

- 위의 표본분포와 사전분포로부터 θ의 사후분포는 $N(113.78, 18.37^2)$이다.
- 찬이의 실제 IQ 점수 θ가 전체 인구의 IQ 점수의 50백분위수보다 작은지, 50백분위수와 75백분 위수 사이에 있는지, 75백분 위수보다 높은지 알아보고자 한다.
- 전체 IQ 점수의 50, 75백분위수는 각각 100, 115.5이므로 관심있 는 가설을 순서에 상관없이

$$H_0 : \theta < 100, \quad H_1 : 100 \leq \theta < 115.5, \quad H_2 : \theta \geq 115.5$$

로 두자.

Example for Multiple Hypotheses Test

각 가설에 대한 사후확률을 구하면 다음과 같다.

$$
\begin{aligned}
p_0 &= \Pr(H_0 \mid \bar{x}) = \Pr\left(Z < \frac{100 - 113.78}{18.37}\right) = 0.2266 \\
p_1 &= \Pr(H_1 \mid \bar{x}) = \Pr\left(\frac{100 - 113.78}{18.37} \le Z < \frac{115.5 - 113.78}{18.37}\right) \\
&= 0.3107 \\
p_2 &= \Pr(H_2 \mid \bar{x}) = \Pr\left(Z > \frac{115.5 - 113.78}{18.37}\right) = 0.4627.
\end{aligned}
$$

사후확률을 보면 찬이의 실제 IQ 점수 가 75백분 위수보다 높을 획률이 약 46%로 가장 크다.