풀이

STEP II 2003, 문제 14

확률과 통계 • 가설 검정의 오류

> **STEP II 2003, 14번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/03-s2-q14

핵심 아이디어

처음으로 6이 나올 때까지 필요한 던지기 횟수는 기하분포(geometric distribution)를 따르며, 그 꼬리(tail) 확률은 아주 간단한 꼴이다. "M번보다 많이 던진다"는 사건은 곧 "처음 M번의 던지기가 모두 실패한다"는 사건이다. 따라서 두 오류 확률은 모두 순수한 M제곱 꼴이고, 둘 사이에서 M을 소거하는 일은 로그(logarithm)를 취하는 것으로 해결된다.

풀이

두 오류 확률

각 던지기는 서로 독립(independent)이고, 매번 같은 확률 p로 6이 나온다고 모형화하자. 그러면 처음으로 6이 나오는 던지기까지 포함한 던지기 횟수 X는 다음을 만족한다.

$$X > M \iff \text{처음 } M\text{번의 던지기 중 6이 하나도 없음}, \qquad \text{따라서} \qquad \mathrm{P}(X > M) = (1-p)^M,$$

여기서 던지기의 독립성을 이용하였다. (등비급수 $\sum p(1-p)^{k-1}$을 더할 필요는 없다. 사건 자체를 직접 생각하는 편이 더 빠르다.)

이 검정은 정확히 $X > M$일 때 H_1을 채택한다. 따라서 실제로는 $p = B$인데 H_1을 채택할 확률 α는

$$\alpha = (1-B)^M,$$

이고, 실제로는 $p = A$인데 H_0을 채택할(즉 $X \le M$일) 확률 β는

$$\beta = 1 - \mathrm{P}(X > M \mid p = A) = 1 - (1-A)^M.$$

이다.

M 소거하기

$(1-A)^M$을 $\alpha = (1-B)^M$의 거듭제곱으로 나타내고자 한다. $0 < B < 1$이므로 $\ln(1-B)$는 정의되고 0이 아니다. 따라서 α에 로그를 취하면

$$M = \frac{\ln \alpha}{\ln(1-B)}.$$

를 얻는다. 그러면

$$(1-A)^M = \mathrm{e}^{M\ln(1-A)} = \exp\left(\frac{\ln(1-A)}{\ln(1-B)}\ln\alpha\right) = \alpha^K, \qquad \text{여기서} \quad K = \frac{\ln(1-A)}{\ln(1-B)}.$$

이다. 그러므로

$$\beta = 1-\alpha^K, \qquad K = \frac{\ln(1-A)}{\ln(1-B)},$$

이고, 요구된 대로 K는 M과 무관하게 A와 B에만 의존한다.

? **생각해 보기** K의 크기에 대해 무엇을 말할 수 있는가?

$0 < A < B < 1$이므로 $1 > 1-A > 1-B > 0$이고, 따라서 두 로그는 모두 음수이며 $\ln(1-B) < \ln(1-A) < 0$이다. 나누면 K는 양수이고

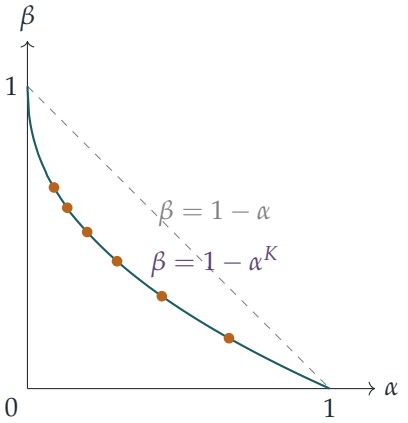
$$0 < K < 1.$$

이다. 바로 이것이 그래프의 모양을 결정한다.

α에 대한 β의 그래프

M이 $1, 2, 3, \ldots$을 차례로 지나면 $\alpha = (1-B)^M$은 $(0,1)$ 안의 값을 가지며 0을 향해 감소한다. 엄밀히 말하면 그래프는 점들의 수열이고, 이 점들은 모두 곡선 $\beta = 1-\alpha^K$, $0 \le \alpha \le 1$ 위에 놓인다. 이 곡선에 대하여 다음이 성립한다.

- 곡선은 $(0,1)$과 $(1,0)$을 지난다.
- $\dfrac{\mathrm{d}\beta}{\mathrm{d}\alpha} = -K\alpha^{K-1} < 0$이므로 α가 커지면 β는 작아진다. 즉 한 오류의 가능성을 줄이면 다른 오류의 가능성이 커진다.
- $K-1 < 0$이므로 $\alpha \to 0^+$일 때 기울기는 $-\infty$로 발산하고(곡선은 $(0,1)$에서 수직으로 출발한다), $\alpha = 1$에서 기울기는 -1과 0 사이의 값인 $-K$이다.
- $\dfrac{\mathrm{d}^2\beta}{\mathrm{d}\alpha^2} = K(1-K)\alpha^{K-2} > 0$이므로 곡선은 아래로 볼록(convex)하다.
- $0 < \alpha < 1$이고 $0 < K < 1$이면 $\alpha^K > \alpha$이므로 $\beta < 1-\alpha$이다. 즉 곡선은 $(0,1)$과 $(1,0)$을 잇는 직선보다 아래에 있다.



곡선은 $A = \frac{1}{6}$, $B = \frac{1}{3}$인 경우(따라서 $K \approx 0.45$)에 대해 그린 것이며, $M = 1, \ldots, 6$에 해당하는 점들을 표시하였다. M이 커지면 점은 왼쪽 위로 이동한다.

검산

- $A = \frac{1}{6}$, $B = \frac{1}{3}$, $M = 4$로 놓자. 직접 계산하면 $\alpha = (\frac{2}{3})^4 \approx 0.1975$이고 $\beta = 1 - (\frac{5}{6})^4 \approx 0.5177$이다. $K = \ln \frac{5}{6} / \ln \frac{2}{3} \approx 0.4497$로 계산해도 $1 - \alpha^K \approx 0.5177$이다. 이 검정을 2×10^5번 시행한 시뮬레이션(simulation)에서는 0.198과 0.518을 얻었다.
- 극단적인 경우에도 타당하게 동작한다. A가 B에 가까우면 $K \approx 1$이고 $\beta \approx 1 - \alpha$이므로, 이 검정은 던진 결과를 무시하고 무작위로 결정하는 것보다 별로 나을 것이 없다. A가 아주 작으면 $K \approx 0$이고, 적당한 어떤 α에 대해서도 β를 작게 만들 수 있다.

흔한 실수

- 어느 가설(hypothesis)이 어느 것인지 혼동하는 것. 여기서 H_0은 **더 큰** 값 B이므로, 6이 나올 때까지 오래 기다리는 것은 H_0에 반하는 증거가 된다.
- K를 거꾸로 $\ln(1 - B) / \ln(1 - A)$로 구하는 것. 이 값은 1보다 크다. $\beta < 1 - \alpha$(합리적인 검정이라면 어림짐작보다 낫다)인지 확인하면 이 실수를 잡아낼 수 있다.

STEP 너머

α에 대한 β의 곡선은 두 종류의 오류 사이의 상충 관계(trade-off)를 나타내는 곡선으로, 통계학과 기계 학습(machine learning) 전반에서 쓰이는 ROC 곡선(ROC curve)과 밀접하게 관련되어 있다. 네이만-피어슨 보조정리(Neyman–Pearson lemma)에 따르면, 이와 같은

두 단순가설에 대해서는 가능도비(likelihood ratio)에 기반한 검정이 각 α에 대해 가능한 가장 작은 β를 준다. 여기서는 가능도비가 X에 대해 단조이므로, "X가 클 때 H_0을 기각한다"는 검정이 실제로 최적이다. 이러한 개념을 더 알고 싶다면 기초로 Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)을 읽고, 그다음에 수리통계학 입문 과정의 아무 교재나 보면 된다.

Beyond STEP 풀이: https://beyondstep.keithmcnulty.org/ko/. 풀이 글은 CC0 1.0이다. 문제 자체의 저작권은 © Cambridge University Press & Assessment에 있으며 여기에 싣지 않는다. 영어 원문을 기계 번역하였으며 아직 검토되지 않았다.