축차확률비 검정

정의

- 개요

1. 표본의 크기가 일정하지 않은 확률변수의 가설을 검정하는 것
2. 알고리즘적 방법으로 조건을 만족하는 우도값이 나올때까지 반복적으로 검정을 실시

- 유도

1. 서로 확률적으로 독립인 확률변수 $X_1, X_2, \ldots$ 의 실현값을 $x_1, x_2, \ldots$ 라고 하자.

1) 이 실현값을 지속적으로 관찰한다고 가정할 때,
L(θ; n)은 n개의 표본에 대한 우도함수라고 하자. 즉

(1) L(θ; 1) = $f(x_1;\ \theta)$

(2) L(θ; 2) = $f(x_1;\ \theta)f(x_2;\ \theta)$

(3) L(θ; n) = $\prod_{i=1}^{n} f(x_i;\ \theta)$

정의

- 유도

2. 이 때, 어떤 가설에 대한 검정을 정의한다.

1) $H_0 : \theta = \theta_1$, $H_1 : \theta = \theta_2$

2) H_0를 유지하는 경우는 다음과 같다.

(1) $C_n = \{X_n : k_0 < \frac{\mathrm{L}(\theta_1;\mathrm{k})}{\mathrm{L}(\theta_2;\mathrm{k})} < k_1, \frac{\mathrm{L}(\theta_1;\mathrm{k})}{\mathrm{L}(\theta_2;\mathrm{k})} \leq k_0\}$ (단 $k = 1, \ldots$,n-1)

3) H_1을 채택하는 경우는 다음과 같다.

(1) $C_n = \{X_n : k_0 < \frac{\mathrm{L}(\theta_1;\mathrm{k})}{\mathrm{L}(\theta_2;\mathrm{k})} < k_1, \frac{\mathrm{L}(\theta_1;\mathrm{k})}{\mathrm{L}(\theta_2;\mathrm{k})} > k_0\}$ (단 $k = 1, \ldots$,n-1)

4) 조건이 충족되지 않을 시, $k_0 < \frac{\mathrm{L}(\theta_1;\mathrm{k})}{\mathrm{L}(\theta_2;\mathrm{k})} < k_1$ 에서 계속 관찰한다.

정의

- 축차확률비 검정의 검정력함수

1. α를 1종오류의 확률이라고 하고, β를 2종오류 확률이라고 하자.

1) 이 때, $P_{H_0}(X \in C_n) = \alpha$, $P_{H_1}(X \in C_n) = 1-\beta$ 로 정의하면

1) 축차확률비 검정에서

(1) $\alpha = \sum_{i=1}^{\infty} \int_{C_n} \mathrm{L}(\theta_1; \mathrm{n})$, 즉 $P_{H_0}(X \in C_n)$의 무한합이고

(2) $1-\beta = \sum_{i=1}^{\infty} \int_{C_n} \mathrm{L}(\theta_2; \mathrm{n})$ 즉 $P_{H_1}(X \in C_n)$의 무한합이다.

3) 이 확률의 여집합, 즉 H_0를 채택할때의 검정력은

(1) $1-\alpha = \sum_{i=1}^{\infty} \int_{{C_n}^c} \mathrm{L}(\theta_1; \mathrm{n})$, 즉 $P_{H_0}(X \in {C_n}^c)$

(2) $\beta = \sum_{i=1}^{\infty} \int_{{C_n}^c} \mathrm{L}(\theta_2; \mathrm{n})$ 즉 $P_{H_1}(X \in C_n)$의 무한합이다.

정의

- 축차확률비 검정의 검정력함수

2. 이 때, 만약 $X_1, \ldots, X_n \in C_n$이면

1) L(θ_1; n) ≤ k_0L(θ_2; n)이고, 이는 곧 $\alpha \le k_0(1-\beta)$

2) 반대로, $X_1, \ldots, X_n \in {C_n}^c$ 이면, 이는 곧

(1) $1-\alpha \le k_1(\beta)$ 이다.

3) 위를 정리하면

(1) $\frac{\alpha}{(1-\beta)} \le k_0$ 이고

(2) $\frac{1-\alpha}{\beta} \le k_1$ 이다.

4) α_α, β_α를 미리 정한 진분수라고 하자.

(1) $\frac{\alpha_\alpha}{(1-\beta_\alpha)} = k_0$

(2) $\frac{1-\alpha_\alpha}{\beta_\alpha} = k_1$

5) 3)과 4)를 결합하면

(1) $\frac{\alpha}{(1-\beta)} \le \frac{\alpha_\alpha}{(1-\beta_\alpha)}$

(2) $\frac{1-\alpha}{\beta} \le \frac{1-\alpha_\alpha}{\beta_\alpha}$

정의

- 축차확률비 검정의 검정력함수

3. 최종적으로 정리하면

(1) $\alpha(1-\beta_{\alpha}) \leq \alpha_{\alpha}(1-\beta)$

(2) $\beta(1-\alpha_{\alpha}) \leq \beta_{\alpha}(1-\alpha)$

(3) 여기서 부등식의 대응변끼리 더하면

- $\alpha-\alpha\beta_{\alpha}-\alpha_{\alpha}-\beta\alpha_{\alpha} \leq \beta_{\alpha}-\alpha\beta_{\alpha}-\beta+\beta\alpha_{\alpha}$

(4) 위를 정리하면

- $\alpha-\alpha_{\alpha} \leq -\beta+\beta_{\alpha}$, $\therefore\ \alpha+\beta \leq \alpha_{\alpha}+\beta_{\alpha}$

- 결과적으로, $\alpha \leq \frac{\alpha_{\alpha}}{1-\beta_{\alpha}}$, $\beta \leq \frac{\beta_{\alpha}}{1-\alpha_{\alpha}}$

예제

- 생산공정의 이상유무 판단

1. N개 추정값의 평균 $\bar{x}$는 평균 μ와 분산 $\frac{\sigma^2}{n}$인 근사 정규분포를 따른다.
2. 이 때, 표본평균 $\bar{x}$가 정의된 통계량

1) LCL = $\mu - \frac{3\sigma}{\sqrt{n}}$, UCL = $\mu + \frac{3\sigma}{\sqrt{n}}$ 사이에 있을 확률은 0.997이다.
2) 만일, 시계열적으로 관측한 관측값들이 이 상/하한에 위치할 경우 생산품의 평균은 변동되지 않은 것으로 간주한다.
3) 반대로, N번째 샘플에서 이 상/하한을 벗어날 경우 공정을 중단하고 조사한다.