

26-1 DSL 정규 세션

기초과제



- 본 과제는 「통계학입문」, 「통계방법론」 및 「수리통계학(1)」 일부 내용을 다루며, NumPy 와 Pandas 의 활용 연습을 돋기 위해 기획되었습니다. 평가를 위한 것이 아니므로, 주어진 힌트(💡)를 적극 활용하시고 학회원 간 토론, Slack 의 질의응답을 활용하시어 해결해주십시오. 단, 답안 표절은 금지합니다.
- 서술형 문제는 ✎, 코딩 문제는 © 으로 표기가 되어 있습니다. 각 문제에서 요구하는 방법에 맞게 해결하며, 서술형 문제들은 따로 작성하시어 pdf 로 제출해주시고 코딩 문제들은 ipynb 파일에 답안을 작성하시어 제출해주십시오.
- 1/16 (금) 23 시 59 분까지 Github 에 PDF 파일과 ipynb 파일을 zip 파일로 압축하여 제출해주십시오. Github 에 제출하는 방법을 모른다면 학술부장에게 연락 부탁드립니다.

문제 1 (Weak) Law of Large Numbers

큰 수의 법칙은 표본 평균의 수렴성을 보장하는 법칙으로, 중심극한정리(CLT)와 더불어 통계학에서 중요한 법칙입니다. 예를 들어, 큰 수의 법칙은 몬테카를로(Monte Carlo) 방법론의 이론적 기반을 제공합니다. 이 문제에서는 큰 수의 약한 법칙의 정의를 확인하고, 이를 코드를 통해 확인해 보겠습니다.

1-1 ✎ : 큰 수의 약한 법칙(Weak Law of Large Numbers)의 정의를 서술하시고 증명하시오.

☞ Hogg(8판) 5장 1절

정의) 서로 독립이고 동일분포인 확률변수 X_1, X_2, \dots 가 $E[X_i] = \mu$, $\text{Var}(X_i) = \sigma^2 < \infty$ 를 만족한다고 하자.
표본평균 $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ 가 따로, 임의의 $\varepsilon > 0$ 에 대해서 $\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| > \varepsilon) = 0$ 이면 $\bar{X}_n \xrightarrow{P} \mu$ 라고, 이를 "큰 수의 약한 법칙"이라 한다

증명) 확률변수로

$$E[\bar{X}_n] = \mu, \quad \text{Var}(\bar{X}_n) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{\sigma^2}{n}$$

Chebyshev 부등식에 의해 임의의 $\varepsilon > 0$ 에 대해서

$$P(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0$$

따라서 $\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| > \varepsilon) = 0$,

$$\text{즉 } \bar{X}_n \xrightarrow{P} \mu.$$

1-2 © : NumPy 를 이용하여 큰 수의 약한 법칙(WLLN)을 시뮬레이션으로 확인하시오.