1) Frequentist방법과 Bayesian방법론의 차이점을 10줄 이내로 정리하여 작성해주세요.

Frequentist 통계학은 확률을 장기적인 빈도로 해석하며, 모수를 고정된 값으로 간주하고 데이터로부터 모수의 점 추정값을 계산함. 신뢰 구간은 반복된 실험에서 참 모수를 포함할 비율을 의미하며, 가설 검정에서는 귀무가설을 설정하고 p-값을 통해 이를 기각할지 여부를 결정함.

반면, Bayesian 통계학은 확률을 주관적 믿음의 정도로 해석하고, 모수를 확률 분포로 간주함. Bayesian 접근법에서는 사전 분포를 설정하고 데이터를 통해 사후 분포를 업데이트하며, 신뢰 구간은 모수가 특정 구간에 있을 확률로 해석됨. 가설 검정에서는 사후 분포와 Bayes factor를 사용하여 두 가설의 상대적인 타당성을 평가함.

2) 베이지안은 사전분포 g(θ)를 완벽하게 안다고 가정한다는 문제점이 존재합니다. 이를 어떻게 해결할 수 있는지 10줄 이내로 정리하여 작성해주세요.

첫째, 비정보적 사전 분포(non-informative prior)를 사용하여 사전 지식이 거의 없을 때 데이터를 최대한 반영하려는 방법이 있음. 균등 분포나 제프리스 분포가 그 예임.

둘째, 약정보적 사전 분포(weakly informative prior)를 사용하여 일부 사전 지식을 반영하되 너무 강하지 않게 설정함으로써 분석 결과에 큰 영향을 미치지 않도록 함.

셋째, 사전 분포의 민감도 분석(sensitivity analysis)을 통해 여러 가지 다른 사전 분포를 사용해 결과가 어떻게 변하는지 확인하여 특정 사전 분포에 대한 결과의 민감성을 평가할 수 있음.

넷째, 계층적 모형(hierarchical modeling)을 사용하여 상위 수준의 모수에 대한 사전 분포를 설정하고 이를 통해 하위 모수의 사전 분포를 유도함으로써 유연한 모델링이 가능해짐.

다섯째, 경험적 추정(empirical Bayes)을 통해 데이터를 사용하여 사전 분포의 파라미터를 추정하는 방법이 있음.

3) 분포수렴은 누적분포함수(cumulative density function)를 통해 정의됩니다. 확률밀도 함수(probability density function)를 이용하여 분포수렴을 정의할 수 없는 이유를 반례를 통해 서술해주세요.

디랙 델타 함수(Dirac Delta Function)을 활용하여 확률밀도함수를 이용하여 분포수렴을 정의할 수 없음을 보이겠다.

확률 변수 투이 1/n 간격으로 증가하는 점에서의 확률 밀도를 가지는 확률 변수라고 가정.

Xn 의 PDF는 다음과 같다.

$$f_{X_n}(x) = n\mathbf{1}\left(\frac{k-1}{n} < x \leq \frac{k}{n}\right) \quad (k = 1, 2, \ldots, n)$$

(1은 인디케이터 함수를 의미)

이어서, Yn이 디랙 델타 함수로 정의된 확률 변수라고 가정

디랙 델타 함수 δ(x)는 x=0에서 무한대 값을 가지며, 그 외의 값은 0인 함수를 의미한다.

$$\delta(x) = \begin{cases} \infty & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases}$$

이러한 확률 변수 Yn의 PDF는 다음과 같다

$$f_{Y_n}(x) = \delta(x - \frac{1}{n})$$

Xn 과 Yn 의 PDF는 상이한 모습을 보이지만, n이 무한으로 커질 때 두 확률 변수는 0에 수렴하는 분포를 가진다.

즉, 위와 같은 예시로 볼 때 PDF만으로는 분포수렴을 완전히 정의할 수 없다는 것을 알 수 있다.

4) MGF(moment generating function)를 이용해서 CLT를 증명해주세요.

CLT(Central Limit Theorem): 동일한 확률 분포를 가진 독립 확률 변수 n개의 평균의 분포는 n이 적당히 크다면 정규 분포를 수렴한다.

MGF(Moment Generating Function): 확률 변수 혹은 분포의 모멘트는 확률 변수의 거듭제곱의 기댓값으로 정의함.

가정 1) X1, X2, ... , Xn 은 독립적이고 동일한 분포를 따르는 확률 변수

가정 2) 각 Xi는 평균 μ와 분산 σ^2을 가짐

증명)

1. MGF 정의

각 Xi의 MGF는 Mx(t)로 동일하다

$$M_X(t) = \mathbb{E}[e^{tX}]$$

2. Summation of MGF

$S_n = \sum_{i=1}^{n} X_i$의 MGF는 독립적이므로 개별 MGF의 곱으로 표현할 수 있음

$$M_{S_n}(t) = \mathbb{E}[e^{tS_n}] = \mathbb{E}[e^{t\sum_{i=1}^{n} X_i}] = \prod_{i=1}^{n} M_X(t) = (M_X(t))^n$$

3. Summation of MGF의 정규화

$$M_{\frac{S_n - n\mu}{\sqrt{n}\sigma}}(t) = \mathbb{E}\left[e^{t\frac{S_n - n\mu}{\sqrt{n}\sigma}}\right]$$

$Y_n = \frac{S_n - n\mu}{\sqrt{n}\sigma}$ 라고 한다면,

$$M_{Y_n}(t) = \mathbb{E}\left[e^{tY_n}\right] = \mathbb{E}\left[e^{t\frac{S_n - n\mu}{\sqrt{n}\sigma}}\right] = M_{S_n}\left(\frac{t}{\sqrt{n}\sigma}\right) \cdot e^{-\frac{tn\mu}{\sqrt{n}\sigma}}$$

4. 테일러 전개

Mx(t)를 t에 대해 테일러 전개를 하면,

$$M_X\left(\frac{t}{\sqrt{n}\sigma}\right) = 1 + \mu\frac{t}{\sqrt{n}\sigma} + \frac{\sigma^2\left(\frac{t}{\sqrt{n}\sigma}\right)}{2} + o\left(\frac{t^2}{n}\right) = 1 + \frac{\mu t}{\sqrt{n}\sigma} + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)$$

$$M_{S_n}\left(\frac{t}{\sqrt{n}\sigma}\right) = \left(1 + \frac{\mu t}{\sqrt{n}\sigma} + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n$$

이 식을 전개하면,

$$\left(1 + \frac{\mu t}{\sqrt{n}\sigma} + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n \approx \exp\left(n \cdot \left(\frac{\mu t}{\sqrt{n}\sigma} + \frac{t^2}{2n}\right)\right)$$

$$= \exp\left(\sqrt{n} \cdot \frac{\mu t}{\sigma} + \frac{t^2}{2}\right)$$ 이 된다.

$$M_{Y_n}(t) = M_{S_n}\left(\frac{t}{\sqrt{n}\sigma}\right) \cdot e^{-\frac{tn\mu}{\sqrt{n}\sigma}} = \exp\left(\sqrt{n} \cdot \frac{\mu t}{\sigma} + \frac{t^2}{2}\right) \cdot \exp\left(-\frac{tn\mu}{\sqrt{n}\sigma}\right)$$

$$= \exp\left(\frac{t^2}{2}\right)$$

$$\lim_{n\to\infty} M_{Y_n}(t) = \exp\left(\frac{t^2}{2}\right)$$