report

1. Frequentist 방법과 Bayesian 방법론의 차이점

빈도주의(Frequentist) 방법:

- **확률의 해석**: 빈도주의는 확률을 사건이 반복적으로 발생할 빈도로 해석합니다. 예를 들어, 동전을 1000번 던졌을 때 앞면이 나올 빈도 같은 방식입니다.
- **매개변수 처리**: 매개변수는 고정된 값으로 간주됩니다. 이 값은 데이터가 제공하지 않기 때문에 추정할 수 없습니다.
- **추정 및 검정**: 데이터의 빈도나 비율에 기반하여 추정하고 검정합니다. 예를 들어, 신뢰구간은 반복적 실험을 통해 매개변수가 포함될 범위를 제공합니다.

베이지안(Bayesian) 방법:

- **확률의 해석**: 베이지안 방법에서는 확률을 사건 발생에 대한 주관적 믿음이나 척도로 해석합니다.
- **매개변수 처리**: 매개변수는 확률 변수로 간주되며, 사전 분포(prior distribution)를 통해 초기 추정치를 설정합니다.
- **추정 및 검정**: 데이터가 주어지면 사전 분포를 업데이트하여 사후 분포(posterior distribution)를 계산합니다. 이를 통해 매개변수에 대한 불확실성을 직접적으로 표현합니다.

주요 차이점:

- 빈도주의는 확률을 빈도로, 매개변수를 고정된 값으로 보고 데이터의 반복적 관찰을 통해 추정합니다.
- 베이지안은 확률을 주관적 믿음으로 보고, 매개변수를 확률 변수로 취급하여 사전 지식과 데이터를 통합하여 추정합니다.

2. 베이지안 사전분포 문제 해결

사전분포 문제:

- 베이지안 접근법에서 사전분포(prior distribution)는 분석의 결과에 큰 영향을 미칩니다. 그러나 사전분포를 적절하게 선택하는 것은 어려울 수 있습니다.

해결 방법:

1. **비정보적(prior) 분포**: 사전 지식이 거의 없거나 중립적일 때 사용하는 방법입니다. 예를 들어, 균등분포를 사용하는 것이 이에 해당합니다.
2. **약정보적(prior) 분포**: 데이터에 대해 약간의 정보는 포함되지만, 지나치게 구체적이지 않은 분포입니다.
3. **경험적 베이지안(Empirical Bayesian)**: 데이터에서 사전분포의 매개변수를 추정합니다. 이 방법은 주관적인 사전 지식의 영향을 줄이고, 데이터 기반으로 사전분포를 결정하려고 합니다.

이러한 방법들은 사전분포 선택의 주관성을 줄이거나 조정하여 보다 객관적인 분석을 가능하게 합니다.

3. 확률밀도 함수를 이용한 분포수렴 정의 불가 이유

분포수렴(Convergence in Distribution):

- **정의**: 확률변수 Xn이 확률변수 X로 분포수렴한다고 할 때, Xn의 누적 분포 함수(CDF) FXn(x)가 X의 누적 분포 함수 FX(x)로 수렴합니다.

확률 밀도 함수(PDF)와의 차이:

- **PDF 수렴과의 관계**: 분포수렴은 누적 분포 함수(CDF) 수렴을 의미하며, PDF가 존재하지 않는 경우나, PDF가 수렴하지 않는 경우도 있습니다. 즉, PDF가 수렴하는 것은 CDF가 수렴하는 것과는 다릅니다.

예시:

- 주어진 확률 밀도 함수 fn(x)=(1−cos(2πnx))·x∈(0,1)1는 균등분포 U(0,1)로 수렴하지만, fn(x) 자체는 수렴하지 않습니다. 이는 특정한 확률 밀도 함수가 CDF의 수렴과 일치하지 않는 경우를 보여줍니다.

결론:

- 분포수렴은 누적 분포 함수의 수렴을 의미하며, 확률 밀도 함수가 수렴하지 않아도 CDF가 수렴할 수 있습니다. 이는 PDF와 CDF 간의 수렴 개념이 독립적일 수 있음을 나타냅니다.

4. MGF를 이용한 중앙극한정리 증명

$$
\begin{aligned}
E(\bar{X}) &= E\left(\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right) \\
&= \frac{1}{n}\left(E(X_1) + E(X_2) + \cdots + E(X_n)\right) \\
&= \frac{1}{n} \times nE(X) = \mu
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(\bar{X}) &= \mathrm{Var}\left(\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right) \\
&= \frac{1}{n^2}\left(\mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n)\right) \\
&= \frac{1}{n^2} \times n\mathrm{Var}(X) = \frac{\sigma^2}{n}
\end{aligned}
$$

$$
\begin{aligned}
M_{\frac{\sigma}{\sqrt{n}}(\bar{X}-\mu)}(t) &= E\left(\exp\left(\frac{X_1 - \mu + X_2 - \mu + \cdots + X_n - \mu}{\sqrt{n}} t\right)\right) \\
&= E\left(\exp\left(\frac{X_1 - \mu}{\sqrt{n}} t\right)\right) \cdots E\left(\exp\left(\frac{X_n - \mu}{\sqrt{n}} t\right)\right) \\
&= \left\{E\left(\exp\left(\frac{X - \mu}{\sqrt{n}} t\right)\right)\right\}^n
\end{aligned}
$$

$$
\begin{aligned}
\lim_{n\to\infty} M_{\frac{\sigma}{\sqrt{n}}(\bar{X}-\mu)}(t) &= \exp\left(\lim_{n\to\infty} n \ln\left(E\left(\exp\left(\frac{X - \mu}{\sqrt{n}} t\right)\right)\right)\right) \\
&= \exp\left(\lim_{h\to 0} \frac{\ln M_{\frac{\sigma}{\sqrt{n}}(X-\mu)}(th)}{h^2}\right)
\end{aligned}
$$

if $h = \frac{1}{n}, n \to \infty\ h \to 0$

$$
\lim_{h\to 0} \frac{\ln M_{\frac{\sigma}{\sqrt{n}}(X-\mu)}(th)}{h^2} = \exp\left(\frac{t^2}{2}\right)
$$

$$
\begin{aligned}
\lim_{n\to\infty} \frac{\sigma}{\sqrt{n}}(\bar{X} - \mu) &\sim N(0,1) \\
\bar{X} &\sim N\left(\mu, \frac{\sigma^2}{n}\right)
\end{aligned}
$$