Asymptotics

IDEA lab

Department of Statistics, Seoul National University

2025. 06. 02.

Outline

Set-up

- $X_1, \ldots, X_n \stackrel{i.i.d.}{\sim} f(x|\theta^*)$ for some θ^* in Θ.
- Let $\hat{\theta}$ be the maximum likelihood estimator of θ.
- We are going to prove that $\sqrt{n}(\hat{\theta} - \theta^*)$ converges in distribution to the normal distribution with mean 0 and covariance matrix $\mathbb{I}^{-1}(\theta^*)$, where $\mathbb{I}(\theta) = -\mathbb{E}(\partial^2 f(X|\theta)/\partial\theta\partial\theta)$.

Prove

- Let $\ell_n(\theta) = \sum \log f(X_i|\theta)$ and let $s_n(\theta) = \partial \ell_n(\theta)/\partial\theta$.
- Note that $s_n(\hat{\theta}) = 0$ by the definition of the MLE.
- In addition, we have $\mathbb{E}(\partial \log f(X|\theta)/\partial\theta)\big|_{\theta=\theta^*} = 0$, which implies that $\mathbb{E}(s_n(\theta^*)) = 0$.
- Thus, the central limit theorem implies that

$$\frac{s_n(\theta^*)}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, \mathbb{E}(s(\theta^*)s^\top(\theta^*))),$$

where $s(\theta) = \partial \log f(X|\theta)/\partial\theta$.

Prove

- Now, Taylor expansion yields

$$0 = s_n(\hat{\theta}) \approx s_n(\theta^*) + [\partial s_n(\theta)/\partial\theta)](\hat{\theta} - \theta^*).$$

- Thus we have

$$\hat{\theta} - \theta^* \approx [\partial s_n(\theta)/\partial\theta)]^{-1} s_n(\theta^*),$$

  and so

$$\sqrt{n}(\hat{\theta} - \theta^*) \xrightarrow{d} \mathcal{N}(0, [\partial s_n(\theta)/\partial\theta)/n]^{-1}\mathbb{E}(s(\theta^*)s^\top(\theta^*)))[\partial s_n(\theta)/\partial\theta)/n]^{-1}$$

- Finally, by changing the integration and derivative operators, we can show that

$$[\partial s_n(\theta)/\partial\theta)/n]^{-1}\mathbb{E}(s(\theta^*)s^\top(\theta^*)))[\partial s_n(\theta)/\partial\theta)/n]^{-1} \to \mathbb{I}(\theta^*)^{-1}.$$

Outline

Problem

- We call that a given model $\{f(x|\theta), \theta \in \Theta\}$ is nonparametric if the dimension of Θ is infinity.
- A popular way of estimating θ is a sieve MLE.
- We call $\Theta_n, n = 1, \ldots$ a sieve if Θ_n is increasing, the dimension of Θ_n is finite and $\Theta_n \approx \Theta$ well.
- Let $\hat{\theta}$ be the sieve MLE (i.e. the maximizer of the log-likelihood on Θ_n).
- We want to know how fast $\hat{\theta}$ converges to θ^* in terms of n.
- A problem is that it is hardly possible to say something about the convergence of $\hat{\theta}$ to θ^* since the dimension of θ^* is infinite.

Excess Risk

- Instead, we try to say something about the convergence rate of $\mathcal{E}(\hat{\theta})$, where
$$\mathcal{E}(\theta) = \mathbb{E}\ell(X, \theta) - \mathbb{E}\ell(X, \theta^*)$$
for a given loss $\ell(X, \theta)$.
- Here, $\hat{\theta}$ is the minimizer of $\sum_{i=1}^n \ell(X_i, \theta)$ on $\theta \in \Theta_n$. and $\theta^* = \mathrm{argmin}_{\theta \in \Theta} \mathbb{E}\ell(X, \theta)$.
- When $\ell(X, \theta)$ is the negative log-likelihood, $\hat{\theta}$ becomes a sieve MLE.

Techniques

- Let $\mathbb{E}_n \ell(X, \theta) = \sum_{i=1}^{n} \ell(X_i, \theta)/n$.
- By the law of large numbers, we have

$$\mathbb{E}_n \ell(X, \theta) \approx \mathbb{E} \ell(X, \theta).$$

- Under regularity conditions, we can show that this convergence holds uniformly in θ.
- If $\mathbb{E}\ell(X, \theta)$ is convex and has the minimizer at θ^*, we expect that $\hat{\theta}$ is close to θ^* and thus $\mathcal{E}(\hat{\theta})$ converges to 0.
- Read references for details.

Outline

Bernstein-von Mises theorem

- Note that the posterior distribution is given as

$$\pi(\theta|\mathrm{Data}) \propto \prod_{i=1}^{n} f(X_i|\theta)\pi(\theta).$$

- Let $\hat{\theta}$ be the MLE and rewrite the posterior as

$$\pi(\theta|\mathrm{Data}) \propto \exp\left(\ell_n(\theta) - \ell_n(\hat{\theta})\right)\pi(\theta),$$

where $\ell_n(\theta) = \sum_{i=1}^{n} \log f(X_i|\theta)$.

- Taylor expansion yields

$$\ell_n(\theta) - \ell_n(\hat{\theta}) \approx s_n(\hat{\theta})^\top(\theta - \hat{\theta}) + (\theta - \hat{\theta})^\top[\partial s_n(\theta)/\partial\theta]_{\theta=\hat{\theta}}(\theta - \hat{\theta}).$$

- Since $s_n(\hat{\theta}) = 0$, we have

$$\pi(\theta|\mathrm{Data}) \propto \exp\left((\theta - \hat{\theta})^\top[\partial s_n(\theta)/\partial\theta]_{\theta=\hat{\theta}}(\theta - \hat{\theta})\right)\pi(\theta).$$

- Thus, we can say that

$$\sqrt{n}(\theta - \hat{\theta})|\mathrm{Data} \xrightarrow{d} \mathcal{N}(0, \mathbb{I}^{-1}(\theta^*)).$$

Outline

Preliminaries

- $d(P, Q)$: Hellinger distance for distribution P and Q
- 데이터 $X_1, ..., X_n \sim P_0$
- $\mathcal{P}$: 고려하는 분포들의 집합 (Model space)
- p : density for $P \in \mathcal{P}$
- $\Pi_n(\cdot)$: $\mathcal{P}$ 위에서 정의된 prior distribution
- Posterior distribution은 다음과 같이 정의됨

$$\Pi_n(B|X_1, ..., X_n) = \frac{\int_B \prod_{i=1}^n p(X_i) d\Pi_n(P)}{\int \prod_{i=1}^n p(X_i) d\Pi_n(P)}$$

Main Theorem

Theorem (Theorem 2.1 from Ghosal (2000))

Suppose that for a sequence ϵ_n *with* $\epsilon_n \to 0$ *and* $n\epsilon_n^2 \to \infty$*, a constant* $C > 0$ *and sets* $\mathcal{P}_n \subseteq \mathcal{P}$*, we have*

$$\log N(\epsilon_n, \mathcal{P}_n, d) \leq n\epsilon_n^2 \tag{1}$$

$$\Pi_n\left(P : \mathbb{E}_{P_0}\left[\log \frac{p_0(X)}{p(X)}\right] \leq \epsilon_n^2, \mathbb{E}_{P_0}\left[\log \frac{p_0(X)}{p(X)}\right]^2 \leq \epsilon_n^2\right) \geq \exp(-n\epsilon_n^2 C) \tag{2}$$

$$\Pi_n(\mathcal{P}\backslash\mathcal{P}_n) \leq \exp(-n\epsilon_n^2(C+4)) \tag{3}$$

Then for sufficiently large M*, we have that* $\Pi_n(P : d(P, P_0) \geq M\epsilon_n | X_1, ..., X_n) \to 0$ *in* P_0^n*-probability*

Proof of Main Theorem

다음과 같은 2가지의 Step을 보임으로써 Main Theorem을 증명하고자 한다.

1. $\mathbb{E}_{P_0^n}[\Pi_n(\mathcal{P} \backslash \mathcal{P}_n | X_1, ..., X_n)] \to 0$
2. $\mathbb{E}_{P_0^n}[\Pi_n(P \in \mathcal{P}_n : d(P, P_0) \geq M\epsilon_n | X_1, ..., X_n)] \to 0$

Proof of Main Theorem - Step 1

$$B_n = \left\{ P : \mathbb{E}_{P_0}\left[\log \frac{p_0(X)}{p(X)} \right] \le \epsilon_n^2, \mathbb{E}_{P_0}\left[\log \frac{p_0(X)}{p(X)} \right]^2 \le \epsilon_n^2 \right\} \quad (4)$$

$$A_n = \left\{ X_1, ..., X_n : \int \prod_{i=1}^{n} \frac{p(X_i)}{p_0(X_i)} d\Pi_n(P) \ge \exp(-2n\epsilon_n^2)\Pi_n(B_n) \right\} \quad (5)$$

이라고 할때, Lemma 8.1(Ghosal 2000) 에 의하여 $P_0^n(A_n) \to 1$ 이 성립함.

해석 : 진짜(P_0)와 비슷한 모델들이 충분히 많다면 ($\Pi_n(B_n) \uparrow$), Likelihood ratio는 일정 하한보다 더 작아지지 않는다.

따라서,

$$\mathbb{E}_{P_0^n}[\Pi_n(\mathcal{P}\backslash\mathcal{P}_n|X_1,...,X_n)] \tag{6}$$

$$\leq \mathbb{E}_{P_0^n}[\Pi_n(\mathcal{P}\backslash\mathcal{P}_n|X_1,...,X_n)\mathbb{I}(A_n)] + P_0^n(A_n^c) \tag{7}$$

$$= \mathbb{E}_{P_0^n}\left[\frac{\int_{\mathcal{P}\backslash\mathcal{P}_n}\prod_{i=1}^n p(X_i)d\Pi_n(P)}{\int\prod_{i=1}^n p(X_i)d\Pi_n(P)}\mathbb{I}(A_n)\right] + P_0^n(A_n^c) \tag{8}$$

$$= \mathbb{E}_{P_0^n}\left[\frac{\int_{\mathcal{P}\backslash\mathcal{P}_n}\prod_{i=1}^n (p(X_i)/p_0(X_i))d\Pi_n(P)}{\int\prod_{i=1}^n (p(X_i)/p_0(X_i))d\Pi_n(P)}\mathbb{I}(A_n)\right] + P_0^n(A_n^c) \tag{9}$$

$$\leq \mathbb{E}_{P_0^n}\left[\int_{\mathcal{P}\backslash\mathcal{P}_n}\prod_{i=1}^n (p(X_i)/p_0(X_i))d\Pi_n(P)\right]\exp(2n\epsilon_n^2)\frac{1}{\Pi_n(B_n)} + P_0^n(A_n^c) \tag{10}$$

$$= \Pi_n(\mathcal{P}\backslash\mathcal{P}_n)\exp(2n\epsilon_n^2)\frac{1}{\Pi_n(B_n)} + P_0^n(A_n^c) \to 0 \tag{11}$$

Step 1의 결론

즉, Seive $\mathcal{P}_n$ 밖에는 posterior mass가 거의 남아있지 않다. 따라서, 우리는 Seive에서의 posterior mass만을 조사하면 된다.

Proof of Main Theorem - Step 2

Lemma (Theorem 7.1 in Ghosal (2000))

Assume that condition (1) holds, i.e.,

$$\log N(\epsilon_n, \mathcal{P}_n, d) \leq n\epsilon_n^2.$$

Then, there exist tests ϕ_n such that

$$\mathbb{E}_{P_0^n}\phi_n \leq 2\exp(-Kn\epsilon_n^2) \tag{12}$$

$$\sup_{P\in\mathcal{P}_n:d(P,P_0)>M\epsilon_n} \mathbb{E}_{P^n}(1-\phi_n) \leq \exp(-KnM^2\epsilon_n^2), \tag{13}$$

where $K > 0$ is a constant.

해석 : Seive에 있는 분포들이 적당히 적다보니, 제1종오류 및 제2종오류가 잘 control되는 test function이 존재한다.

Proof of Main Theorem - Step 2

앞의 Lemma에 의하여, test function을 얻을 수 있고, 이를 이용하면 다음과 같다.

$$\mathbb{E}_{P_0}\left[\Pi_n(P \in \mathcal{P}_n : d(P, P_0) \geq M\epsilon_n | X_1, ..., X_n)\right] \tag{14}$$

$$= \mathbb{E}_{P_0}\left[\Pi_n(P \in \mathcal{P}_n : d(P, P_0) \geq M\epsilon_n | X_1, ..., X_n)\phi_n\right] \tag{15}$$

$$+ \mathbb{E}_{P_0}\left[\Pi_n(P \in \mathcal{P}_n : d(P, P_0) \geq M\epsilon_n | X_1, ..., X_n)(1 - \phi_n)\right] \tag{16}$$

식 (15) 은 Test function의 제1종오류가 지수적으로 작기 때문에 무시해도 된다.

Proof of Main Theorem - Step 2

Test function을 사용하는 이유

식 (15) 의 경우 : Test function이 P_0에 나오질 않았다고 판단하는 경우, Type 1 error 확률로 제어

식 (16) 의 경우 : Test function이 P_0에서 나왔다고 판단한 경우, Type 2 error 확률로 제어

Proof of Main Theorem - Step 2

식 (16) 의 upper bound는 다음과 같다.

$$\mathbb{E}_{P_0^n}\left[\Pi_n(P \in \mathcal{P}_n : d(P, P_0) \geq M\epsilon_n | X_1, ..., X_n)(1 - \phi_n)\right] \tag{17}$$

$$\leq \mathbb{E}_{P_0^n}\left[\frac{\int_{P\in\mathcal{P}_n:d(P,P_0)\geq M\epsilon_n} \prod_{i=1}^n \frac{p(X_i)}{p_0(X_i)} d\Pi_n(P)}{\int \prod_{i=1}^n \frac{p(X_i)}{p_0(X_i)} d\Pi_n(P)}(1 - \phi_n)\mathbb{I}(A_n)\right] + P_0^n(A_n^c) \tag{18}$$

Proof of Main Theorem - Step 2

식 (18) 에서 첫번째 term의 upper bound는 다음과 같다.

$$\mathbb{E}_{P_0^n}\left[\frac{\int_{P\in\mathcal{P}_n:d(P,P_0)\geq M\epsilon_n}\prod_{i=1}^n\frac{p(X_i)}{p_0(X_i)}d\Pi_n(P)}{\int\prod_{i=1}^n\frac{p(X_i)}{p_0(X_i)}d\Pi_n(P)}(1-\phi_n)\mathbb{I}(A_n)\right] \tag{19}$$

$$= \int_{P\in\mathcal{P}_n:d(P,P_0)\geq M\epsilon_n}\mathbb{E}_{P_0^n}\left[\prod_{i=1}^n\frac{p(X_i)}{p_0(X_i)}(1-\phi_n)\mathbb{I}(A_n)\right]d\Pi_n(P)\frac{\exp(2n\epsilon_n^2)}{\Pi_n(B_n)} \tag{20}$$

$$\leq \int_{P\in\mathcal{P}_n:d(P,P_0)\geq M\epsilon_n}\mathbb{E}_{P^n}[1-\phi_n]d\Pi_n(P)\frac{\exp(2n\epsilon_n^2)}{\Pi_n(B_n)} \tag{21}$$

$$\leq \sup_{P\in\mathcal{P}_n:d(P,P_0)\geq M\epsilon_n}\mathbb{E}_{P^n}[1-\phi_n]\frac{\exp(2n\epsilon_n^2)}{\Pi_n(B_n)}\to 0 \tag{22}$$

□

Outline

Notations

- Data: $U \sim P$ (in most cases, $U = (X, Y)$), $U \in \mathcal{U}$.
- Loss function $l_\theta(u) : \mathcal{U} \to \mathbb{R}$
  - e.g. $l_\theta(u) = (y - \theta(x))^2$ for $u = (x, y)$ and a function of interest θ.
- Population risk: $R(\theta) = \mathbb{E}_{U \sim P} l_\theta(U)$
- Empirical risk: $R_n(\theta) = \frac{1}{n} \sum_{i=1}^n l_\theta(U_i)$
- Population risk minimizer:

$$\theta^* \in \underset{\theta \in \Theta}{\operatorname{argmin}} R(\theta) \tag{23}$$

- Prior: $\pi(\cdot)$

Gibbs posterior

- Gibbs posterior = Generalized posterior; update the belief via empirical risk, not likelihood (Bissiri (2016)).

Definition (Gibbs posterior)

Given a loss function l_θ and the corresponding empirical risk R_n, define the Gibbs posterior as:

$$\pi_n^{(\omega)}(\theta|\mathcal{D}_n) \propto e^{-\omega n R_n(\theta)}\pi(\theta),\ \theta \in \Theta \tag{24}$$

- $\omega > 0$: called learning rate.

Posterior Concentration

- Following Syring and Martin (2023).

Concentration

$$\pi_n^{(\omega)}(\{\theta : d(\theta, \theta^*) > M\varepsilon_n\}|\mathcal{D}_n) \to 0 \text{ in } P^n\text{-probability as } n \to \infty \quad (25)$$

for $n\varepsilon_n^r \to \infty$ and a large constant $M > 0$.

Conditions

- Concentration can be done, when following 2 conditions are satisfied.
- $m(\theta, \theta^*) = \mathbb{E}_{U\sim P}(l_\theta - l_{\theta^*}),\ v(\theta, \theta^*) = \mathbb{E}_{U\sim P}(l_\theta - l_{\theta^*})^2 - m(\theta, \theta^*)^2.$

Prior concentration condition

$$\log \pi\left(\{\theta : m(\theta, \theta^*) \vee v(\theta, \theta^*) \le \varepsilon_n^r\}\right) \ge -Cn\omega\varepsilon_n^r \tag{26}$$

Sub-exponential loss condition

$$d(\theta, \theta^*) > \delta \Rightarrow \log \mathbb{E}_{U\sim P}[e^{-\omega(l_\theta - l_{\theta^*})}] < -K\omega\delta^r \tag{27}$$

for all sequences $0 < \omega \le \bar{\omega}$ and all sufficiently small $\delta > 0$.

Comparison

Table: Comparison on conditions.

	Gibbs conditions
Prior Sub-exponential	$\log \pi\left(\{\theta : m(\theta,\theta^*) \vee v(\theta,\theta^*) \leq \varepsilon_n^r\}\right) \geq -Cn\omega\varepsilon_n^r$ $d(\theta,\theta^*) > \delta \Rightarrow \log \mathbb{E}_{U\sim P}[e^{-\omega(l_\theta - l_{\theta^*})}] < -K\omega\delta^r$
	Bayesian nonparam conditions
Prior Test type-II	$\Pi_n\left(P : \mathbb{E}_{P_0}\left[\log \frac{p_0}{p}\right] \leq \epsilon_n^2, \mathbb{E}_{P_0}\left[\log \frac{p_0}{p}\right]^2 \leq \epsilon_n^2\right) \geq \exp(-n\epsilon_n^2 C)$ $\sup_{P\in\mathcal{P}_n : d(P,P_0) > M\epsilon_n} \mathbb{E}_{P^n}(1-\phi_n) \leq \exp(-KnM^2\epsilon_n^2)$

Proof Strategy

$$A_n = \{\theta : d(\theta, \theta^*) > M\varepsilon_n\} \tag{28}$$

$$N_n(A_n) = \int_{A_n} e^{-\omega n\{R_n(\theta) - R_n(\theta^*)\}} \pi(d\theta) \tag{29}$$

$$D_n = \int e^{-\omega n\{R_n(\theta) - R_n(\theta^*)\}} \pi(d\theta) \tag{30}$$

$$\pi_n^{(\omega)}(A_n | \mathcal{D}_n) = \frac{N_n(A_n)}{D_n} \tag{31}$$

- Goal: $\pi_n^{(\omega)}(A_n|\mathcal{D}_n) \to 0$
- Strategy: lower bound D_n and upper bound $N_n(A_n)$.

Lemma: D_n lower bound

Lemma

Let $G_n = \{\theta : m(\theta, \theta^*) \vee v(\theta, \theta^*) \le \varepsilon_n^r\}$. *If* $\varepsilon_n \to 0$ *and* $n\varepsilon_n^r \to \infty$, *then*

$$P^n \left[D_n > \frac{1}{2} \pi(G_n) e^{-2n\omega\varepsilon_n^r} \right] \ge 1 - 2(n\varepsilon_n^r)^{-1} \to 1. \tag{32}$$

Main proof (with $N_n(A_n)$ upper bound)

Denote the lower bound on D_n as:

$$b_n = \frac{1}{2}\pi(G_n)e^{-2\omega n\varepsilon_n^r}.$$

Then,

$$\begin{aligned} \pi_n^{(\omega)}(A_n|\mathcal{D}_n) &\leq \frac{N_n(A_n)}{D_n}\mathbb{1}(D_n > b_n) + \mathbb{1}(D_n \leq b_n) \\ &\leq b_n^{-1}N_n(A_n) + \mathbb{1}(D_n \leq b_n). \end{aligned}$$

Main proof

By *sub-exponential loss condition* and independence of U^n, we get

$$\mathbb{E}_{U^n \sim P^n} N_n(A_n) = \int_{A_n} (\mathbb{E}_{U \sim P}[e^{-w(l_\theta - l_{\theta^*})}])^n \pi(d\theta) < e^{-Kn\omega(M\varepsilon_n)^r}.$$

By *prior concentration condition*, we get $\pi(G_n) \geq e^{-Cn\varepsilon_n^r}$.
By the Lemma, we get $P(D_n \leq b_n) \geq 2(n\varepsilon_n^r)^{-1}$.
Therefore,

$$\mathbb{E}_{U^n \sim P^n} \pi_n^{(\omega)}(A_n | \mathcal{D}_n) \leq 2e^{-(\omega K M^r - C - 2\omega)n\varepsilon_n^r} + 2(n\varepsilon_n^r)^{-1} \to 0.$$

Proof of Lemma

Denote $m(\theta, \theta^*)$ and $v(\theta, \theta^*)$ as $m(\theta), v(\theta)$ for brevity. Let

$$Z_n(\theta) = \frac{\{nR_n(\theta) - nR_n(\theta^*)\} - nm(\theta)}{\{nv(\theta)\}^{1/2}}.$$

Let

$$\mathcal{Z}_n = \{(\theta, U^n) : |Z_n(\theta)| \geq (n\varepsilon_n^r)^{1/2}\}.$$

Let

$$\mathcal{Z}_n(\theta) = \{U^n : (\theta, U^n) \in \mathcal{Z}_n\} \text{ and } \mathcal{Z}_n(U^n) = \{\theta : (\theta, U^n) \in \mathcal{Z}_n\}.$$

Then,

$$nR_n(\theta) - nR_n(\theta^*) = nm(\theta) + \{nv(\theta)\}^{1/2} Z_n(\theta).$$

Proof of Lemma

Then, we have

$$D_n \geq \int_{G_n \cap \mathcal{Z}_n(U^n)^c} e^{-\omega n m(\theta) - \omega \{nv(\theta)\}^{1/2} Z_n(\theta)} \pi(d\theta)$$
$$\geq e^{-2\omega n \varepsilon_n^r} \pi(G_n \cap \mathcal{Z}_n(U^n)^c).$$

Hence,

$$P^n \left[D_n \leq \frac{1}{2} \pi(G_n) e^{-2\omega n \varepsilon_n^r} \right]$$
$$\leq P^n \left[e^{-2\omega n \varepsilon_n^r} \pi(G_n \cap \mathcal{Z}_n(U^n)^c) \leq \frac{1}{2} \pi(G_n) e^{-2\omega n \varepsilon_n^r} \right]$$
$$\leq P^n \left[\pi(G_n \cap \mathcal{Z}_n(U^n)) \geq \frac{1}{2} \pi(G_n) \right]$$
$$\leq \frac{2\mathbb{E}_{U^n \sim P^n}[\pi(G_n \cap \mathcal{Z}_n(U^n))]}{\pi(G_n)} \quad \text{(Markov ineq.)}$$

Proof of Lemma

The expectation term in the numerator is simplified:

$$\begin{aligned}\mathbb{E}_{U^n\sim P^n}[\pi(G_n\cap\mathcal{Z}_n(U^n))] &= \int\int \mathbb{1}\{\theta\in G_n\cap\mathcal{Z}_n(U^n)\}\pi(d\theta)P^n(dU^n) \\ &= \int\int \mathbb{1}\{\theta\in G_n\}\mathbb{1}\{\theta\in\mathcal{Z}_n(U^n)\}P^n(dU^n)\pi(d\theta) \\ &= \int_{G_n}\mathbb{E}_{U^n\sim P^n}[\mathcal{Z}_n(\theta)]\pi(d\theta) \\ &\le (n\varepsilon_n^r)^{-1}\pi(G_n) \quad \text{(Chebyshev)}.\end{aligned}$$

Hence,

$$P^n\left[D_n\le\frac{1}{2}\pi(G_n)e^{-2\omega n\varepsilon_n^r}\right]\le 2(n\varepsilon_n^r)^{-1}\to 0.$$

References I

1. Ghosal, S., Ghosh, J. K., & Van Der Vaart, A. W. (2000). Convergence rates of posterior distributions. Annals of Statistics, 500-531.
2. Bissiri, P. G., Holmes, C. C., & Walker, S. G. (2016). A general framework for updating belief distributions. Journal of the Royal Statistical Society Series B: Statistical Methodology, 78(5), 1103-1130.
3. Syring, N., & Martin, R. (2023). Gibbs posterior concentration rates under sub-exponential type losses. Bernoulli, 29(2), 1080-1108.