5. Bayesian estimator

Instructions

On this page, you will be given a distribution and another distribution conditional on the first one. Then, you will find the posterior distribution in a Bayesian approach. You will compute the Bayesian estimator, which is defined in lecture as the mean of the posterior distribution. Then, determine if the Bayesian estimator is consistent and/or asymptotically normal.

We recall that the Gamma distribution with parameters $q > 0$ and $\lambda > 0$ is the continuous distribution on $(0, \infty)$ whose density is given by $f(x) = \frac{\lambda^q x^{q-1} e^{-\lambda x}}{\Gamma(q)}$, where Γ is the Euler Gamma function $\Gamma(q) = \int_0^\infty t^{q-1} e^{-t} dt$, and its mean is q/λ.

We also recall that the $\text{Beta}(a, b)$ distribution has the density $f(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}$ and expectation $a/(a+b)$, where $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

(a)

3.0/3 points (graded)

$p \sim \text{Beta}(a, b)$ for some $a, b > 0$ and conditional on p, $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Ber}(p)$.

What is the Bayesian estimator $\hat{p}^{\text{Bayes}}$?

Generating Speech Output

(If applicable, enter **barX_n** for $\bar{X}_n$, **max(X_i)** for $\max_i X_i$. Do not worry if the parser does not render properly; the grader works independently. If you wish to have proper rendering, enclose **max(X_i)** by brackets.))

$\hat{p}^{\text{Bayes}} =$ (a/n+ barX_n)/((a+b)/n+1) **Answer:** (barX_n+a/n)/(1+(a+b)/n)

Determine whether the Bayesian estimator is consistent, and whether it is asymptotically normal.

- (selected) Consistent and asymptotically normal
- Consistent but not asymptotically normal
- Asymptotically normal but not consistent
- Neither consistent nor asymptotically normal

If it is asymptotically normal, what is its asymptotic variance $V(a,b,p)$? If it is not asymptotically normal, type in 0.

$V(a,b,p) =$ p*(1-p) **Answer:** p*(1-p)

$p \cdot (1-p)$

STANDARD NOTATION

Solution:

1. Find the posterior distribution in a Bayesian approach.

$$\pi(p|x_1,\ldots,x_n) \propto \pi(p) L_n(x_1,\ldots,x_n|p) \propto p^{\sum_i x_i + a - 1}(1-p)^{n - \sum_i x_i + b - 1}$$

Generating Speech Output

We recognize the posterior distribution as $\text{Beta}\left(\sum_i X_i + a, n - \sum_i X_i + b\right)$.

2. Compute the Bayesian estimator.

$$\hat{p} = \int_0^1 p\pi\left(p|x_i, \ldots, x_n\right) dp = \frac{\sum_i X_i + a}{n + a + b} = \frac{\bar{X}_n + a/n}{1 + (a+b)/n}$$

3. Determine whether the Bayesian estimator is consistent.

$$\lim_{n\to\infty} \hat{p} = \lim_{n\to\infty} \frac{\bar{X}_n + a/n}{1 + (a+b)/n}$$

4. Determine whether the Bayesian estimator is asymptotically normal.

From CLT,

$$\sqrt{n}\left(\bar{X}_n - p\right) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\left(0, p\left(1-p\right)\right)$$

Therefore, we find $\hat{p}$ is asymptotically normal.

$$\sqrt{n}\left(\hat{p} - p\right) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\left(0, p\left(1-p\right)\right)$$

5. If it is asymptotically normal, what is its asymptotic variance?

From the above equation, we see that the asymptotic variance is $p\left(1-p\right)$.

Generating Speech Output

$$\sqrt{n}\left(\hat{p}-p\right) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\left(0, p\left(1-p\right)\right)$$

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

(b)

3.0/3 points (graded)

$\pi\left(\theta\right)=1, \forall\theta>0$ and conditional on θ, $X_1,\ldots,X_n \overset{i.i.d.}{\sim} \mathcal{U}\left([0,\theta]\right)$.

What is the Bayesian estimator $\hat{\theta}^{\text{Bayes}}$?

(If applicable, enter **barX_n** for $\bar{X}_n$, **max(X_i)** for $\max_i X_i$. Do not worry if the parser does not render properly; the grader works independently. If you wish to have proper rendering, enclose **max(X_i)** by brackets.)

$\hat{\theta}^{\text{Bayes}} =$ ((n-1)/(n-2))*max(X_i) **Answer:** (n-1)/(n-2)*max(X_i)

Determine whether the Bayesian estimator is consistent, and whether it is asymptotically normal.

- ○ Consistent and asymptotically normal
- ● Consistent but not asymptotically normal
- ○ Asymptotically normal but not consistent
- ○ Neither consistent nor asymptotically normal

Generating Speech Output

If it is asymptotically normal, what is its asymptotic variance $V(\theta)$? If it is not asymptotically normal, type in 0.

$V(\theta) =$ 0 ✔ **Answer:** 0

0

STANDARD NOTATION

Solution:

1. Find the posterior distribution in a Bayesian approach.

$$\pi(\theta|X_1,\ldots,X_n) \propto \pi(\theta) L_n(X_1,\ldots,X_n|\theta) \propto \theta^{-n}\mathbf{1}\{\max_i X_i \leq \theta\}$$

To find the distribution, we set the scale parameter as C.

$$\int_0^\infty \pi(\theta|X_1,\ldots,X_n)\, d\theta = C\int_{\max_i X_i}^\infty \theta^{-n} d\theta = \frac{C}{n-1}(\max_i X_i)^{-n+1} = 1$$

Solving this, we get the full distribution function.

$$C = \frac{n-1}{(\max_i X_i)^{-n+1}}, \quad \pi(\theta|X_1,\ldots,X_n) = \frac{n-1}{(\max_i X_i)^{-n+1}}\theta^{-n}\mathbf{1}\{\max_i X_i \leq \theta\}$$

2. Compute the Bayesian estimator.

$$\hat{\theta} = \int_{\max_i X_i}^\infty \theta\pi(\theta|X_i,\ldots,X_n)\, d\theta = \frac{n-1}{(\max_i X_i)^{1-n}}\int_{\max_i X_i}^\infty \theta^{-n+1} d\theta = \frac{n-1}{n-2}\max_i X_i$$

Generating Speech Output

3. Determine whether the Bayesian estimator is consistent.

Since we know $\hat{\theta}^{MLE} = \max_i X_i \xrightarrow[n\to\infty]{(P)} \theta$, we can find that the Bayesian estimator is consistent.

$$\hat{\theta} = \frac{n-1}{n-2} \max_i X_i \xrightarrow[n\to\infty]{(P)} \theta$$

However, it is not asymptotically normal.

4. Determine whether the Bayesian estimator is asymptotically normal.

It is not asymptotically normal.

5. If it is asymptotically normal, what is its asymptotic variance?

It is not asymptotically normal.

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

(c)

3.0/3 points (graded)

$\lambda \sim$ Exp (α) for some $\alpha > 0$ and conditional on λ, $X_1, \ldots, X_n \overset{i.i.d.}{\sim}$ Exp (λ).

What is the Bayesian estimator $\hat{\lambda}^{\text{Bayes}}$?

Generating Speech Output ble, enter **barX_n** for $\bar{X}_n$, **max(X_i)** for $\max_i X_i$. Do not worry if the parser does not render properly; the grader works independently. If you wish to have proper rendering, enclose **max(X_i)** by brackets.)

$\hat{\lambda}^{\text{Bayes}} =$ (1+1/n)/(alpha/n+barX_n) **Answer:** (1+1/n)/(alpha/n+barX_n)

Determine whether the Bayesian estimator is consistent, and whether it is asymptotically normal.

- (selected) Consistent and asymptotically normal
- Consistent but not asymptotically normal
- Asymptotically normal but not consistent
- Neither consistent nor asymptotically normal

If it is asymptotically normal, what is its asymptotic variance $V(\lambda)$? If it is not asymptotically normal, type in 0. You may use the variable λ.

$V(\lambda) =$ lambda^2 **Answer:** lambda^2

λ^2

STANDARD NOTATION

Solution:

1. Find the posterior distribution in a Bayesian approach.

$$\pi(\lambda|X_1,\ldots,X_n) \propto \pi(\lambda) L_n(X_1,\ldots,X_n|\lambda) \propto \alpha e^{-\alpha\lambda}\lambda^n e^{-\lambda\sum_i X_i} \propto \alpha\lambda^n e^{-(\alpha+\sum_i X_i)\lambda}$$

We recognize the posterior distribution as **Gamma** $(n+1, \alpha+\sum_i X_i)$.

Generating Speech Output mpute the Bayesian estimator.

$$\hat{\lambda} = \frac{n+1}{\alpha + \sum_i X_i} = \frac{1+1/n}{\alpha/n + \bar{X}_n}$$

3. Determine whether the Bayesian estimator is consistent.

$$\hat{\lambda} = \frac{1+1/n}{\alpha/n + \bar{X}_n} \xrightarrow[n\to\infty]{(a.s.)} \frac{1}{\bar{X}_n} \xrightarrow[n\to\infty]{(P)} \lambda$$

In the last transition, we used the knowledge that we already have: $\hat{\lambda}^{MLE} = \frac{1}{\bar{X}_n}$. We conclude that the Bayesian estimator is consistent.

4. Determine whether the Bayesian estimator is asymptotically normal.
By CLT and Delta method,

$$\sqrt{n}\,(\hat{\lambda} - \lambda) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\,(0, \lambda^2)$$

Therefore, it is asymptotically normal.

5. If it is asymptotically normal, what is its asymptotic variance?

The asymptotic variance is λ^2.

$$\sqrt{n}\,(\hat{\lambda} - \lambda) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\,(0, \lambda^2)$$

Generating Speech Output

You have used 1 of 3 attempts

Answers are displayed within the problem

(d)

3.0/3 points (graded)

$\lambda \sim$ Exp (α) for some $\alpha > 0$ and conditional on λ, $X_1, \ldots, X_n \overset{i.i.d.}{\sim}$ Poiss (λ).

What is the Bayesian estimator $\hat{\lambda}^{\text{Bayes}}$?

(If applicable, enter **barX_n** for $\bar{X}_n$, **max(X_i)** for $\max_i X_i$. Do not worry if the parser does not render properly; the grader works independently. If you wish to have proper rendering, enclose **max(X_i)** by brackets.)

$\hat{\lambda}^{\text{Bayes}} =$ (1/n+barX_n)/(alpha/n+1) ✔ **Answer:** (barX_n+1/n)/(1+alpha/n)

Determine whether the Bayesian estimator is consistent, and whether it is asymptotically normal.

- (●) Consistent and asymptotically normal
- ( ) Consistent but not asymptotically normal
- ( ) Asymptotically normal but not consistent
- ( ) Neither consistent nor asymptotically normal

✔

If it is asymptotically normal, what is its asymptotic variance $V(\lambda)$? If it is not asymptotically normal, type in 0.

$V(\lambda) =$ lambda ✔ **Answer:** lambda

Generating Speech Output

λ

Solution:

1. Find the posterior distribution in a Bayesian approach.

$$\pi\left(\lambda|X_1,\ldots,X_n\right) \propto \pi\left(\lambda\right) L_n\left(X_1,\ldots,X_n|\lambda\right) \propto \alpha e^{-\alpha\lambda}\frac{\lambda^{\sum_i X_i}e^{-n\lambda}}{\prod_i X_i!} \propto \alpha\frac{\lambda^{\sum_i X_i}e^{-(\alpha+n)\lambda}}{\prod_i X_i!}$$

   We recognize the posterior distribution as **Gamma** $\left(\sum_i X_i+1,\alpha+n\right)$.

2. Compute the Bayesian estimator.

$$\hat{\lambda} = \frac{\sum_i X_i+1}{\alpha+n} = \frac{\bar{X}_n+1/n}{1+\alpha/n}$$

3. Determine whether the Bayesian estimator is consistent.

   By LLN,

$$\hat{\lambda} = \frac{\bar{X}_n+1/n}{1+\alpha/n} \xrightarrow[n\to\infty]{(a.s.)} \bar{X}_n \xrightarrow[n\to\infty]{(P)} \lambda$$

   Therefore, the Bayesian estimator is consistent.

4. Determine whether the Bayesian estimator is asymptotically normal.

   By CLT,

Generating Speech Output

$$\sqrt{n}\,(\bar{X}_n - \lambda) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\,(0,\lambda)$$

Since the estimator is consistent,

$$\sqrt{n}\,(\hat{\lambda} - \lambda) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\,(0,\lambda)$$

Therefore, it is asymptotically normal.

5. If it is asymptotically normal, what is its asymptotic variance?

The asymptotic variance is λ .

$$\sqrt{n}\,(\hat{\lambda} - \lambda) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\,(0,\lambda)$$

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 5 Bayesian statistics:Homework 9: Bayesian Statistics / 5. Bayesian estimator

Add a Post

All Posts

Generating Speech Output

Asymptotic variance

question posted 11 days ago by **LearningHXH**

This homework involves a lot of calculation of the asymptotic variance calculation of Bayesian estimator. But it seem that there is no discussion for this part in the lectures.

Am I missing something here?

This post is visible to everyone.

Add a Response

3 responses

JennD14

8 days ago

I'm in the same boat. I've been trying to compute the fisher info based on the posterior, but when I take the expectation, I have no idea how to get rid of the n and Xi terms.

I really wish this course followed a textbook.

There's a key bit of info / inference you have to make in the last slide and the last lecture. Re-watch where he talks about the kiss example and think about what asymptotic means.

posted 8 days ago by **DriftingWoods**

I eventually got them, but I still wish the course followed a textbook and had more worked examples.

posted 8 days ago by **JennD14**

Generating Speech Output

Fisher information applies to MLE estimators. Indeed, in some cases we can relate Bayesian estimator to an MLE estimator and then make use of Fisher information. But that's not a general approach for Bayesian, I believe.

I think a more general approach, at least for this problem set is, once you have a Bayesian estimator, to think of the following few questions:

- as $n \to \infty$, what does the Bayesian estimator converge to?
- does the CLT apply to an asymptotic Bayesian estimator, in its explicit form or with delta-method?

Once you have these answers, the rest is mechanics of applying CLT / delta-method.

posted 4 days ago by **Hryhorchuk**

Thanks DriftingWoods, that was helpful hint.

posted 4 days ago by **Jayakrishnankk**

Thanks @DriftingWoods! OMG this is the last sentence on the last slide of the last lecture:)

posted a day ago by **kseniia_zviagintceva**

Add a comment

markweitzman (Community TA)
4 days ago

There practically no need to do any calculations for asymptotic variance. Simply look up in a book/wikipedia the variance of the posterior distribution that you calculated, take the limit as $n \to \infty$, and extract the asymptotic variance.

But if you do the calculations, it will be a nice exercise, it was fun for me!

posted 3 days ago by **sandipan_dey**

Generating Speech Output

Add a comment

ya_mukhin (Staff)
about 21 hours ago

Hi @Satya_VV, @JennD14, @Hryhorchuk, this is an excellent question to ask about the asymptotic behavior of the *posterior distribution* and Bayes estimators!!! The mathematical answer to these questions is known as the Bernstein-von Mises theorem. In frequentist framework, we defined the likelihood and the MLE. The MLE is a *random* object. We then stated a general result about asymptotics of MLE, the MLE Theorem, that looked very much like the CLT but with sample average replaced by the maximum likelihood estimator and the population average replaced by the population parameter. The rate of convergence in the MLE Theorem was $\sqrt{n}$, like in the CLT. In the Bayesian framework, we also have a statistical model $(E, \{P_\theta\}_{\theta\in\Theta})$ (a likelihood) and data. We introduce a *prior* and obtain a *posterior distribution* via Bayes formula. The posterior distribution **is a function of tha data**, so from *frequentist* perspective, we have a **random** probability distribution over the parameter space Θ. The Bernstein-von Mises theorem states that this random probability distribution also has a CLT-like behavior! Specifically, if we center the posterior at θ^0 and rescale it by $\sqrt{n}$, asymptotically it becomes very similar to the Gaussian distribution $N(0, I^{-1}(\theta^0))$. The very important implications of this theorem are: (i) Bayes estimators (mean, mode, median of the posterior distribution) are asymptotically equivalent to the MLE; (ii) the variance of the posterior distribution is a consistent estimate of the asymptotic variance of the MLE or Bayes estiamtors; (iii) the Bayes confidence region of size α is an asymptotically valid confidence interval of level coverage probability α. So, asymptotically, Bayes and frequentist frameworks are not all that different. We can think of Bayes as a machinary to construct estimators, that is an alternative to MLE, M-estimation, method of moments.

Add a comment

Showing all responses

Add a response:

Preview

Generating Speech Output

Submit

Generating Speech Output