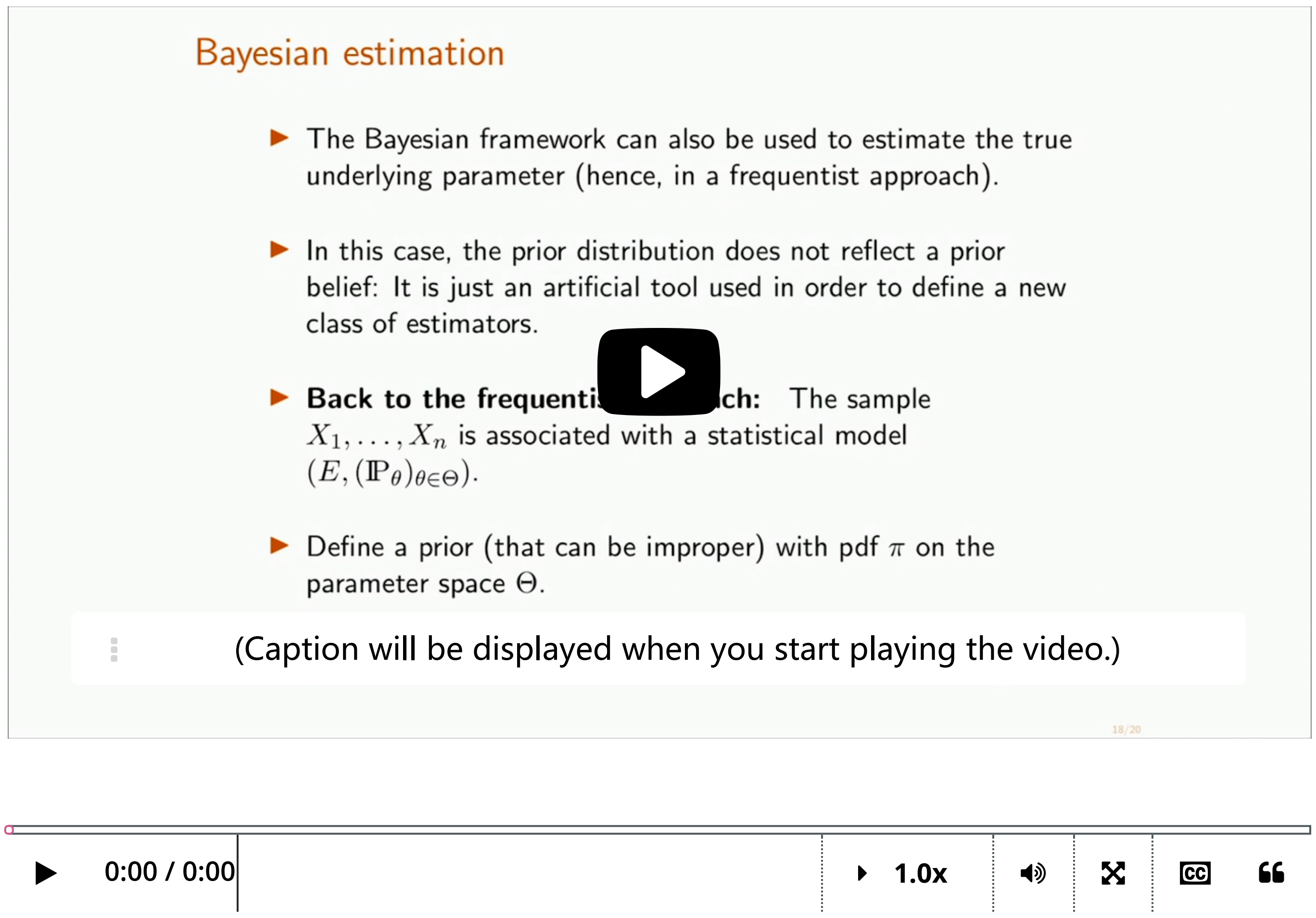
10. Bayesian Statistics for Estimation

Bayesian Estimation



Start of transcript. Skip to the end.

So now if I want to do estimation,

here I'm actually going--

I'm sort of doing things in reverse order.

I'm starting by telling you a region,

and then I'm giving you what that estimator should be.

Now you have a whole posterior and I

want you to spit out just one estimator, something

that should just summarize this data

Video

Download video file

Transcripts

Download SubRip (.srt) file

Download Text (.txt) file

xuetangX.com 学堂在线

(a)

4/4 points (graded)

Consider the posterior distribution derived in the worked example from the previous lecture (here and here).

To recap, our parameter of interest is λ, prior distribution $\mathsf{Exp}\,(a)$, and likelihood $\mathsf{Poiss}\,(\lambda)$ for n observations $X_1, \ldots, X_n$. This is a Gamma distribution with parameters q_0 and λ_0 that you must get from the last two answerboxes in Worked Example Part II.

As before, recall the **Gamma distribution** , which is a probability distribution with parameters $q > 0$ and $\lambda > 0$, has support on $(0, \infty)$, and whose density is given by $f(x) = \dfrac{\lambda^q x^{q-1} e^{-\lambda x}}{\Gamma(q)}$. Here, Γ is the Euler Gamma function.

Of the four sample statistics (mean, median, mode, variance), the Gamma distribution has a simple closed form for three of them. Look up statistics for the Gamma distribution, then for the three that have a simple closed form, calculate them and express your answer in terms of a, n, and $\sum_{i=1}^{n} X_i$ (use **SumXi**), otherwise enter -1.

Note that depending on your source, the format of the Gamma distribution may be different, so you must make sure that you have the correct corresponding parameters.

mean:

(SumXi+1)/(a+n)

✔ **Answer:** (SumXi+1)/(a+n)

$\frac{SumXi+1}{a+n}$

median:

-1

✔ **Answer:** -1

-1

mode:

SumXi/(a+n)

✔ **Answer:** SumXi/(a+n)

$\frac{SumXi}{a+n}$

variance:

(SumXi+1)/(a+n)^2

✔ **Answer:** (SumXi+1)/(a+n)^2

$\frac{SumXi+1}{(a+n)^2}$

STANDARD NOTATION

Solution:

Checking our answer from Worked Example Part II indicates that our parameters for the Gamma distribution are

$$q_0 = (\sum_{i=1}^{n} X_i) + 1$$

and

$$\lambda_0 = a + n.$$

Looking up the statistics for the Gamma distribution, we get that

- Mean is $\frac{q_0}{\lambda_0} = \boxed{\frac{(\sum_{i=1}^{n} X_i) + 1}{a+n}}$.
- Median: no closed form, so we enter $\boxed{-1}$.
- Mode is $\frac{q_0 - 1}{\lambda_0} = \boxed{\frac{\sum_{i=1}^{n} X_i}{a+n}}$.
- Variance is $\frac{q_0}{\beta^2} = \boxed{\frac{(\sum_{i=1}^{n} X_i) + 1}{(a+n)^2}}$.

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

(b)

3/4 points (graded)

Suppose we have the improper prior $\pi(\lambda) \propto e^{-a\lambda}$, $\lambda \in \mathbb{R}$ (and $a \geq 0$). Conditional on λ, we have observations $X_1, X_2, \cdots, X_n \overset{\text{i.i.d}}{\sim} \mathsf{N}(\lambda, 1)$. Compute the posterior distribution $\pi(\lambda | X_1, X_2, \ldots, X_n)$, then provide the following statistics on the posterior distribution.

Use **SumXi** for $\sum_{i=1}^{n} X_i$.

mean:

SumXi/n - a/n

$\frac{SumXi}{n} - \frac{a}{n}$

✔ **Answer:** (SumXi-a)/n

variance:

1/n

$\frac{1}{n}$

✔ **Answer:** 1/n+a*0+SumXi*0

$q_{0.025}$ (cutoff for highest 2.5%):

1.96*(1/n)+ SumXi/n - a/

$1.96 \cdot \left(\frac{1}{n}\right) + \frac{SumXi}{n} - \frac{a}{n}$

✘ **Answer:** (SumXi-a)/n+1.96/sqrt(n)

标准差不是方差

STANDARD NOTATION

True or False: The variance of this distribution models our uncertainty about the value of the parameter λ.

- True ✔
- False

Solution:

In order to calculate statistics about the posterior distribution, we need to compute it first. To do this, we use Bayes' theorem, combining the prior $\pi(\lambda)$ and the likelihood function $L_n(X_1, \ldots, X_n|\lambda)$.

We can easily get that

$$\pi(\lambda) = \exp(-a\lambda)$$

and

$$L_n(X_1, \ldots, X_n|\lambda) \propto \exp\left(\sum_{i=1}^{n} -\frac{(X_i - \lambda)^2}{2}\right).$$

Using Bayes' formula, un-normalized, then gives

$$\begin{aligned}
\pi(\lambda|X_1, \ldots, X_n) &\propto \pi(\lambda) L_n(X_1, \ldots, X_n|\lambda) \\
&\propto \exp(-an) \exp\left(\sum_{i=1}^{n} -\frac{(X_i - \lambda)^2}{2}\right) \\
&= \exp\left(-an + \sum_{i=1}^{n} -\frac{(X_i - \lambda)^2}{2}\right) \\
&= \exp\left(-\frac{n}{2}\lambda^2 + \left(\left(\sum_{i=1}^{n} X_i\right) - a\right)\lambda - \frac{1}{2}\sum_{i=1}^{n} X_i^2\right)
\end{aligned}$$

This is an exponential of a quadratic polynomial in λ, that is, it has the form $\alpha\lambda^2 + \beta\lambda + \gamma$. We get the equivalence $\alpha = -\frac{n}{2}$, $\beta = \left(\sum_{i=1}^{n} X_i\right) - a$, and $\gamma = -\frac{1}{2}\sum_{i=1}^{n} X_i^2$.

We have derived in a previous exercise that this corresponds to a Gaussian distribution. In this part, we allowed for an improper prior that's supported on the whole real line, so there's no truncation involved and the posterior distribution is indeed a Gaussian. Now, we calculate its parameters.

- Its mean μ is $\frac{-\beta}{2\alpha} = \boxed{\frac{\left(\sum_{i=1}^{n} X_i\right) - a}{n}}$.
- Its variance σ^2 is $\frac{-1}{2\alpha} = \boxed{\frac{1}{n}}$.

From this, we calculuate

$$q_{0.025} = \mu + 1.95\sqrt{\sigma^2} = \boxed{\frac{\left(\sum_{i=1}^{n} X_i\right) - a}{n} + \frac{1.96}{\sqrt{n}}}.$$

The variance of our posterior distribution reflects the spread of possible values of λ once both our prior and the observations are taken into account, so it indeed models our uncertainty about the value of λ.

Submit You have used 3 of 3 attempts

Answers are displayed within the problem

这里先验概率告诉我们lambda > 0
那么，后验概率的support大于0
是一个truncated normal distribution

(c)

1/2 points (graded)
Now, suppose that we instead have the proper prior $\pi(\lambda) \sim \textsf{Exp}(a)$ ($a > 0$). Again, just as in part (b): conditional on λ, we have observations $X_1, X_2, \cdots, X_n \overset{\text{i.i.d}}{\sim} \textsf{N}(\lambda, 1)$. You may assume that $a < \sum_{i=1}^{n} X_i$. Compute the posterior distribution $\pi(\lambda | X_1, X_2, \ldots, X_n)$, then provide the following statistics on the posterior distribution. Write `Phi` for the CDF function $\Phi()$ and `PhiInv` for its inverse.

Use **SumXi** for $\sum_{i=1}^{n} X_i$.

median:

PhiInv(3/4)*1/sqrt(n) + (S ✗ **Answer:** (SumXi-a)/n+1/sqrt(n)*PhiInv(1-0.5*Phi(1/sqrt(n)*(SumXi-a)))

mode:

(SumXi)/n - a/n ✓ **Answer:** (SumXi-a)/n

$\frac{SumXi}{n} - \frac{a}{n}$

STANDARD NOTATION

Solution:

The calculations done in the previous part may be repeated, up to the point where we derive that the distribution is proportional to a Gaussian distribution. (Notice that $\textsf{Exp}(a)$ exactly corresponds to $\pi(\lambda) \propto e^{-a\lambda}$, just with a support of $[0, \infty]$ instead of $\mathbb{R}$.

In this part, however, because our prior is only defined over $[0, \infty]$, we have to truncate our distribution. Hence, our posterior is the Gaussian distribution with parameters $\mu = \frac{\left(\sum_{i=1}^{n} X_i\right) - a}{n}$ and $\sigma^2 = \frac{1}{n}$, truncated such that only $\lambda > 0$ is considered.

The mode is still easy to calculate due to the given assumption $a < \sum_{i=1}^{n} X_i$, which implies that the tip of the Gaussian distribution is positive, so this is definitely the mode. The tip of the Gaussian distribution is the same as the mean in the non-truncated distribution, so we get that the mode is $\boxed{\frac{\sum_{i=1}^{n} X_i - a}{n}}$.

The median is a bit more complex, but is easily resolved by considering quantiles of the truncated and the full Gaussian distribution.

剩余的所有面积

We first calculate the proportion of the distribution that's left. The area below 0 in a Gaussian disteribution with given μ and σ^2 will be $\Phi\left(-\frac{\mu}{\sigma}\right)$, so the amount remaining is $1-\Phi\left(-\frac{\mu}{\sigma}\right)=\Phi\left(\frac{\mu}{\sigma}\right)$. Hence, the area above the median of the truncated distribution is half the amount remaining, so it is $\boxed{\frac{1}{2}\Phi\left(\frac{\mu}{\sigma}\right)}$. From this, we could calculate the area up to the median of the truncated distribution to be $\boxed{1-\frac{1}{2}\Phi\left(\frac{\mu}{\sigma}\right)}$.

Finally, we get the z-score of the median of the truncated distribution in terms of the whole distribution to be $\boxed{\Phi^{-1}\left(1-\frac{1}{2}\Phi\left(\frac{\mu}{\sigma}\right)\right)}$. Substituting back $\mu=\frac{(\sum_{i=1}^n X_i)-a}{n}$ and $\sigma^2=\frac{1}{n}$, and then using these to convert the z-score to the actual value, gives the answer

Phi((0 - mu)/sigma)

$$\frac{\sum_{i=1}^n X_i - a}{n}+\frac{1}{\sqrt{n}}\Phi^{-1}\left(1-\frac{1}{2}\Phi\left(\frac{1}{\sqrt{n}}\left(\sum_{i=1}^n X_i-a\right)\right)\right).$$

Submit You have used 3 of 3 attempts

Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 5 Bayesian statistics:Lecture 18: Jeffrey's Prior and Bayesian Confidence Interval / 10. Bayesian Statistics for Estimation