4. GLM: Statistical Model and Notation

We now combine our ingredients together.

We have discussed **canonical exponential famlies** parametrized by θ, with the log-partition function $b(\theta)$ having the property that $b'(\theta) = \mu$. Recall that in GLMs, the point of the link function is to assume $g(\mu(\mathbf{x})) = \mathbf{x}^T\boldsymbol{\beta}$, where μ is the **regression function** : the mean of Y given $\mathbf{X} = \mathbf{x}$, $\mathbb{E}[Y \mid \mathbf{X} = \mathbf{x}]$.

Concept Check: Properties of the Canonical Link Function

1/1 point (graded)

Let $f_\theta = \exp\left(\frac{y\theta - b(\theta)}{\phi} + c(y,\phi)\right)$ for $\phi \neq 0$ describe an exponential family. Which one of the following statements about the function $g(\mu) = \theta$ is **false** ?

- ○ The canonical link function always exists.
- ○ g is identical to $(b')^{-1}$.
- ○ If g strictly increases, then g^{-1} strictly increases.
- ● Regardless of the value of ϕ, g is strictly increasing.

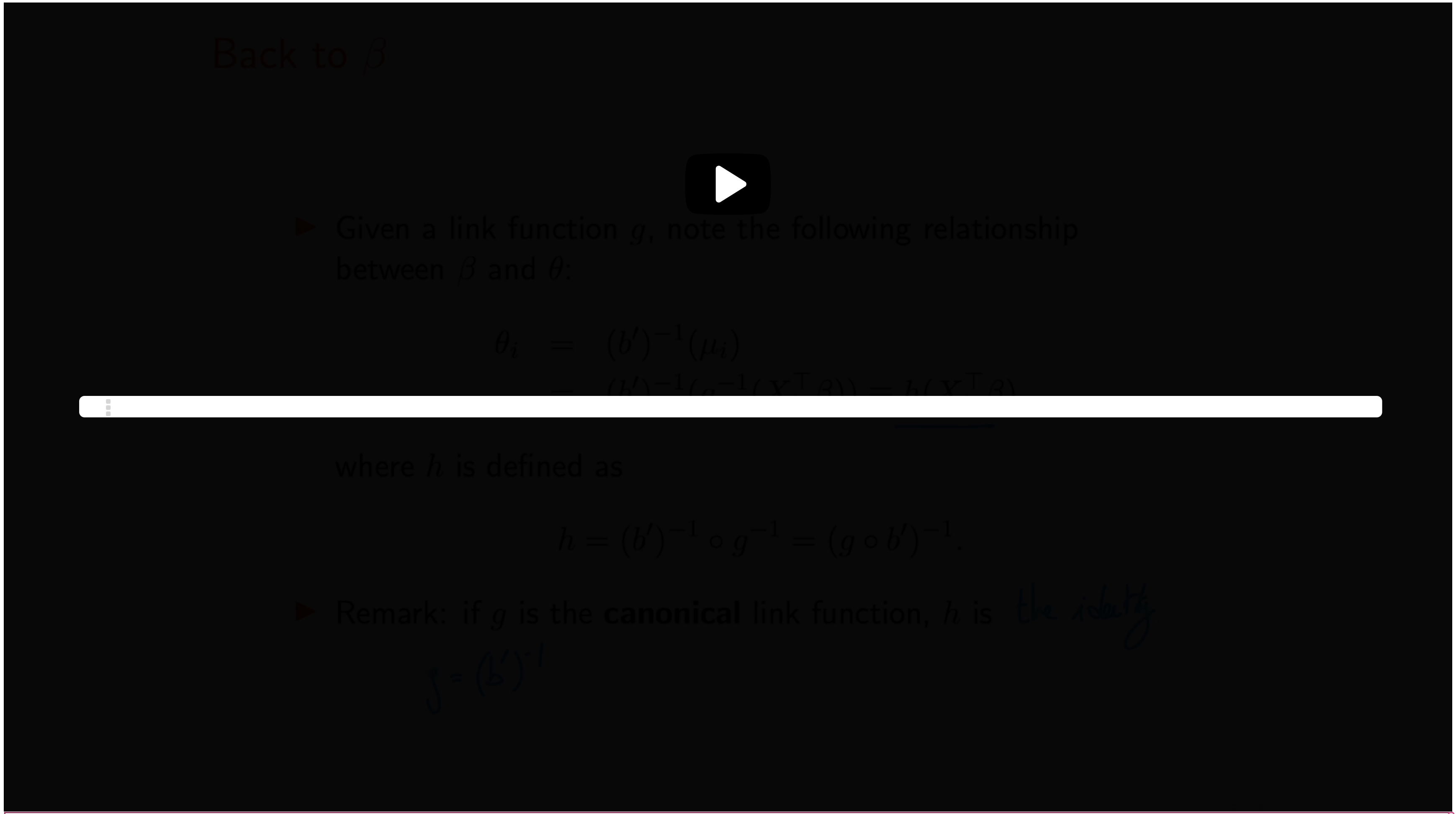
Solution:

- We can always write down the function $g\left(\mu\right) = \theta$.
- Based on the properties of the log-partition function b, we derived previously that $b'\left(\theta\right) = \mu$, so we have the identity $g\left(\mu\right) = \left(b'\right)^{-1}\left(\mu\right)$.
- It is a general fact that if f is a function that strictly increases, then its inverse is a function that strictly increases. The same holds for strictly decreasing functions.
- g **decreases** if $\phi < 0$. This can be seen from the fact that $\phi \cdot b''\left(\theta\right)$ is the variance of a random variable, which means $b'' < 0$. Thus, b' is a decreasing function, which means $\left(b'\right)^{-1}$ is decreasing. **Ultimately, this demonstrates that there is a “canonical" choice of parametrization.** If $\phi < 0$, all that tells us is that we should re-parametrize by multiplying both ϕ and b by -1. We can always make such a choice, as long as $\phi \neq 0$, so that g is an increasing function. Recall that this is one of the properties we wanted out of link functions of GLMs!

Submit You have used 2 of 2 attempts

Answers are displayed within the problem

GLM and Introduction of Beta for Estimation



13:57 / 13:57 1.50x

Video
[Download video file]

Transcripts
[Download SubRip (.srt) file]
[Download Text (.txt) file]

Concept Check: Assumptions for GLM

0/1 point (graded)

Choose from the following **the assumptions we make in fitting data using a GLM**.

- [x] We assume a distribution for Y. ✔
- [x] We assume a link function $g(\cdot)$.
- [x] We assume a noise model that captures the relationship between $\mathbf{X}$ and Y.

✖

Solution:

The only assumptions we make in using a GLM (from the choices we are given in this problem) are **a distribution for** Y and **a link function** $g(\cdot)$. We do not need need to assume a noise model to capture the relationship between Y and $\mathbf{X} = \mathbf{x}$. The assumptions of a distribution for Y and a link function $g(\mu(\mathbf{x}))$ relate Y and $\mathbf{X} = \mathbf{x}$ through the following equation:

$$g(\mu(\mathbf{x}) = \mathbb{E}[Y|\mathbf{X} = \mathbf{x}]) = \mathbf{x}^T \boldsymbol{\beta}.$$

Submit You have used 2 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 7 Generalized Linear Models:Lecture 22: GLM: Link Functions and the Canonical Link Function / 4. GLM: Statistical Model and Notation

Add a Post

Show all posts ▼ by recent activity ▼

? Do 'Solution' and green check/tick match for Concept Check: Assumptions for GLM
Seems the written Solution(same as what had marked) does not match with the shown green but could be totally wrong.
2 new_ 4