# C.4 Hypothesis Testing - Outline

C.4.7 Score Based Model Choices
- Score Based Approaches
- Akaike Information Criterion (AIC)
- Schwarz-Bayes Criterion (SBC)
- Exercises



# Score Based Approaches

Adding parameters typically increases $\ell(\theta)$.
But it also makes computations harder.
Is the extra complexity worth it?

One adjustment: add a penalty to the loglikelihood based on the number of parameters. Prefer model with highest adjusted score.

Two adjustments on syllabus: Schwarz-Bayes Criterion (SBC) and Akaike Information Criterion (AIC)

# Score Based Approaches: Example

Akaike Information Criterion: $\text{AIC} = \ell(\theta) - r$
where $r = \#$ of parameters fitted from data.
Pick model with highest AIC score.

## Example

Five models are fitted to a sample of $n = 125$ observations:

| Model | # Fitted Parameters | $\ell(\theta)$ | AIC |
|---|---|---|---|
| I | 1 | -208 | -209 |
| II | 2 | -205 | -207 |
| III | 3 | -202 | -205 |
| IV | 4 | -200 | -204 |
| V | 6 | -199 | -205 |

Which model is favored by AIC?

AIC prefers IV.



# Score Based Approaches: Example

Schwarz-Bayes Criterion: $\text{SBC} = \ell(\theta) - \frac{r}{2}\ln(n)$
where $r = \#$ of fitted parameters, $n = \#$ of data points.
Pick model with highest SBC score.

## Example

Five models are fitted to a sample of $n = 125$ observations:

| Model | # of Parameters | $\ell(\theta)$ | SBC |
|---|---|---|---|
| I | 1 | -208 | -210.4 |
| II | 2 | -205 | -209.8 |
| III | 3 | -202 | -209.2 |
| IV | 4 | -200 | -209.7 |
| V | 6 | -199 | -213.5 |

Which model is favored by SBC?

SBC prefers III.

# Comparison

$$\text{AIC} = \ell(\theta) - r$$

$$\text{SBC} = \ell(\theta) - \frac{r}{2}\ln(n)$$

Where $r = \#$ of fitted parameters.
If $n \geq 8 > e^2$, then the SBC penalty is larger than the AIC penalty.
So if $n \geq 8$, the preferred model under SBC will have either fewer or the same number of parameters as the AIC choice.

Exam questions essentially test whether or not you have memorized these two formulas.

The **S**BC has an adjustment for the data **s**ize.

Many people define the AIC and SBC score as $-2\cdot$ (our score). If the question is just asking to choose a model either option works. If it wants the raw score, need to use our text's score.



# Exercise 1

You are deciding between two models, with 2 and 5 fitted parameters. The 5 parameter model's loglikelihood at the MLE is $-387.2$. With 100 data points, what is the smallest loglikelihood that would make you choose the 2 parameter model using the SBC? Using the AIC?

## Exercise 1

You are deciding between two models, with 2 and 5 fitted parameters. The 5 parameter model's loglikelihood at the MLE is $-387.2$. With 100 data points, what is the smallest loglikelihood that would make you choose the 2 parameter model using the SBC? Using the AIC?

$$\begin{aligned}\text{SBC} &= \ell(\theta) - \frac{r}{2}\ln(100)\\ &= -387.2 - \frac{5}{2}\ln(100) = -398.7 \text{ for 5 parameter model}\\ -398.7 &= x - \frac{2}{2}\ln(100) \quad \Rightarrow x = -394.1\end{aligned}$$

We prefer the 2 parameter model if $\ell > -394.1$

$$\begin{aligned}\text{AIC} &= \ell(\theta) - r\\ &= -387.2 - 5 = -392.2 \text{ for 5 parameter model}\\ -392.2 &= x - 2 \quad \Rightarrow x = -390.2\end{aligned}$$

is the 2 parameter cutoff.



## Exercise 2

An actuary is trying to decide between modeling losses with

1. an exponential random variable with mean $\theta$,
2. a uniform$(a, b)$ random variable, or
3. (iii) a uniform$(0, u)$ random variable.

All parameters are fit using maximum likelihood estimators. Given three losses of 1, 4, and 10, which model would be selected using the Schwarz-Bayesian Criterion?

# Exercise 2

An actuary is trying to decide between modeling losses with

1. an exponential random variable with mean $\theta$,
2. a uniform$(a, b)$ random variable, or
3. (iii) a uniform$(0, u)$ random variable.

All parameters are fit using maximum likelihood estimators. Given three losses of 1, 4, and 10, which model would be selected using the Schwarz-Bayesian Criterion?

$$\text{Exponential}(\theta)\text{: } \widehat{\theta} = \overline{X} = 5$$

$$L(\theta) = \frac{1}{\theta^3} e^{-(1+4+10)/\theta}$$

$$\ell(5) = -3\ln(5) - \frac{15}{5} = -7.828$$

$$\text{Uniform}(a, b)\text{: } L(a, b) = \frac{1}{(b-a)^3}$$

That is maximized when $b - a$ is minimized, so $\widehat{a} = 1, \widehat{b} = 10$

$$\ell(1, 10) = -3\ln(10 - 1) = -6.592$$



# Exercise 2 (Continued)

$$n = 3$$

$$\text{Exponential}(\theta)\text{: } \quad \ell(\widehat{\theta}) = -7.828$$

$$\text{Uniform}(a, b)\text{: } \quad \ell(\widehat{a}, \widehat{b}) = -6.592$$

$$\text{Uniform}(0, u)\text{: } \quad \widehat{u} = 10$$

$$L(u) = \frac{1}{u^3}$$

$$\ell(10) = -3\ln(10) = -6.908$$

Now we plug into SBC $= \ell - \frac{r}{2}\ln(n)$

| Model | # dof | $\ell$ | SBC |
|---|---|---|---|
| Exponential$(\theta)$ | 1 | $-7.828$ | $-8.378$ |
| Uniform$(a, b)$ | 2 | $-6.592$ | $-7.69$ |
| Uniform$(0, u)$ | 1 | $-6.908$ | $-7.46$ |

So we choose the uniform$(0, u)$