✖

Week 4 Quiz

Continue Course

Back to Week 4

9/10 points earned (90%)

Quiz passed!

0 / 1 points

1.
In a Bayesian simple linear regression $y \sim N(\alpha + x\beta, \sigma^2)$

Suppose our priors on the parameters α, β, σ^2 are independent and that the prior on β is $N(0,1)$. How will the posterior mean of β compare to least squares estimate of β?

- The mean of the posterior distribution of β will be equal to the least squares estimate.

This should not be selected
By imposing a prior on β with mean 0 and variance 1 the posterior will be a mixture of this prior and likelihood, as such the posterior mean will be at least slightly pulled closer to the prior mean than the least squares estimate, which uses only the likelihood to estimate β.

This question refers to the following learning objective(s):

- Understand the basics of Bayesian linear regression and how it relates to Frequentist regression.

- The mean of the posterior distribution of β will be closer to zero than the least squares estimate.
- The mean of the posterior distribution of β will be higher than the least squares estimate.
- The mean of the posterior distribution of β will be lower than the least squares estimate.

1 / 1
points

2.

A simple linear model (either Bayesian or frequentist) that tries to predict an individual's height from his/her age is unlikely to perform well, since human growth rates are non-linear with regard to age. Specifically, humans tend to grow quickly early in life, stop growing at through most of adulthood, and sometimes shrink somewhat when they get old. Which of the following modifications to a simple linear regression model should you prefer?

- Including terms of age^2 and or $\log(age)$ as covariates in the model.

Correct

This question refers to the following learning objective(s):

- Identify the assumptions of linear regression and assess when a model may need to be improved.

- Including other relevant covariates such as weight or income.
- Imposing strong prior distributions on the parameters in a Bayesian analysis.
- Log-transforming the dependent variable (height) to account for skewness.

---

1 / 1
points

3.

Suppose we want to set a level k such that if we observe a data point more than k standard deviations away from the mean, we deem it an outlier. If the number of observations is 1000, what is the probability that we observe an outlier at least 4 standard deviations away from its prediction value?

- 0.03
- 0.06

Correct

This question refers to the following learning objective(s):

- Check the assumptions of a linear model
- Identify outliers and high leverage points in a linear model.

- 0.12
- 0.24

---

1 / 1 points

4.

Suppose we use Bayesian methods (with a prior distribution) to fit a linear model in order to predict the final sale price of a home based on quantifiable attributes of the home. If the 95% posterior predictive interval of a new home (not in the data set) is (312,096, 392,097), which of the following statements represents a correct interpretation of this interval?

- 95% of houses with the same attributes as this house have will be sold for prices between 312.096 and 392,097.
- The probability that the house will sell for between 312,096 and 392,097 is 0.95.

Correct

This question refers to the following learning objective(s):

- Interpret Bayesian credible and predictive intervals in the context of multiple linear regression.

- This house would be sold for between 312,096 and 392,097 95% of the time.
- 95% of houses with the same attributes as this house have historically sold for prices between 312.096 and 392,097.

---

1 / 1 points

5.

Which of the following is not a principled way to select a model?

- Pick the model with the highest R^2

Correct

This question refers to the following learning objective(s):

- Use principled statistical methods to select a single parsimonious model.

- ( ) Pick the model with the highest Adjusted R^2.
- ( ) Use Bayesian Model Averaging and select the model with the highest posterior probability.
- ( ) Select the model with the lowest BIC.

---

1 / 1 points

6.
In a linear model with an intercept term (that is always included) and 4 potential predictors, how many possible models are there?

- ( ) 4
- ( ) 5
- (•) 16

Correct
This question refers to the following learning objective(s):

- Implement Bayesian model averaging for both prediction and variable selection.

- ( ) 32

---

1 / 1 points

7.
Can Bayesian model averaging be done with a large amount of predictors?

- ( ) No, since it will take forever to average over 2^k possible models when k is large.
- (•) Yes, but Monte Carlo sampling techniques will need to be done to approximate the posterior distribution

Correct

This question refers to the following learning objective(s):

- Understand the importance and use of MCMC within Bayesian model averaging.

- Yes, it is possible to find the posterior model probabilities in closed form by using the conjugate Zellner g-prior.

---

1 / 1 points

8.

Which of the following is **not** a useful method of checking a linear model after it is fit?

- Plotting the residuals to check for non-normally distributed residuals.
- Examining the influence of potential outliers on the parameters of the model.
- Comparing the distribution of fitted values to the distribution of observed data.
- Ensuring that R^2 is as close to 1 as possible.

Correct

This question refers to the following learning objective(s):

- Deduce how wrong model assumptions affect model results.

---

1 / 1 points

9.

Why is the Zellner g-prior useful in Bayesian model averaging?

- It prevents BMA from disproportionately favoring the null model as a result of the Bartlett-Lindley paradox.
- It simplifies prior elicitation down to two components, the prior mean and g.

Correct

This question refers to the following learning objective(s):

- Understand the purpose of prior distributions within Bayesian model averaging.

○ It helps shrink the coefficients towards 0, which is important if the variables are highly correlated.

---

1 / 1 points

10.
When selecting a single model from an ensemble of models in the case of Bayesian model averaging, which of the following selection procedures corresponds to choosing the "highest probability model"?

○ Selecting the model with the highest posterior model probability.

Correct

This question refers to the following learning objective(s):

- Implement Bayesian model averaging for both prediction and variable selection.

○ Selecting the model that generates predictions most similar to those obtained from averaging over the model space.

○ Including only the coefficients with posterior model inclusion probability above 0.5
