✖ Try again once you are ready.

Required to pass: 80% or higher

You can retake this quiz up to 3 times every 8 hours.

Back to Week 4

Retake

✔

1 / 1 points

1. True or False: The mean and standard deviation of the posterior distribution of a slope or intercept parameter in Bayesian linear regression is equal to the least squares estimate and corresponding standard error if the reference prior is used and normally distributed errors are assumed.

- ◯ True

Correct

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

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

- ◯ False

✖

0 / 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?

- Imposing strong prior distributions on the parameters in a Bayesian analysis.
- Including terms of age^2 and or $\log(age)$ as covariates in the model.
- Log-transforming the dependent variable (height) to account for skewness.

This should not be selected

Including transformations of the independent variable such as $\log(age)$ and age^2 as covariates in a model is often a great way to capture non-linear relationships within the context of a linear model, which is easy to work with in both Bayesian and Frequentist settings.

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.

✖ 0 / 1 points

3. You fit a linear model on 1000 data points and identify a point that lies 3 standard deviations above its predicted value. Should you worry about this potential outlier? Why or why not?

- No, because the probability that all 1000 points will be within 3 standard deviations of their predicted values is 0.07, so it is

unsurprising to observe a point 3 standard deviations away from its predicted value.

Yes, because outliers can have high leverage and result in a poorly fit model.

Week 4 Quiz

Quiz, 10 questions

3/10 points (30%)

This should not be selected

As the sample size increases, the expected number of points that deviate by k standard deviations also increases. Hint - remember that residuals are normally distributed and hence we can use the command $(1 - 2 * pnorm(-k))^N$ to find the probability that all N points are within k standard deviations of their predicted value.

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.

No, because the probability that all 1000 points will be within 3 standard deviations of their predicted values is 0.74, so it is not implausible to observe a point 3 standard deviations away from its predicted value.

Yes, since the probability of a point deviating from its predicted value by at least 3 standard deviations is roughly 0.003, which suggests that the point is an outlier.

✖ 0 / 1

4. Suppose a researcher is using Bayesian multiple regression to quantify the effect of vitamin C on cancer patient mortality. The central 95% posterior

points

credible interval of the coefficient of vitamin C dosage is (-0.19, -0.07). Assuming the model assumptions are valid, what can we say about the effect of vitamin C on cancer patient mortality?

- The posterior probability that the coefficient of vitamin C is greater than zero is low, so there is a high posterior probability of a negative association between vitamin C and cancer patient mortality.
- We reject the null hypothesis of no difference, since the 95% credible interval does not include zero.

This should not be selected
While frequentist confidence intervals are often used in hypothesis tests, Bayesian credible intervals indicate a high probability region of the parameter space. Rejecting or accepting a hypothesis is done using a loss function or Bayes factor.

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

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

- There is not enough information to quantify the effect of vitamin C on cancer patient mortality.

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

0 / 1 points

- Select the model with the lowest BIC.

- Pick the model with the highest Adjusted R^2.
- Pick the model with the highest R^2
- Use Bayesian Model Averaging and select the model with the highest posterior probability.

This should not be selected
R^2 always increases as the number of predictors increases, regardless of the effect of the predictors. Hence, R^2 will always pick the largest possible model.

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

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

✖ 0 / 1 points

6. In a linear model with an intercept term (that is always included) and 3 potential predictors, how many possible models are there?

- 3
- 4

This should not be selected
Since each predictor (excluding the intercept) can either be included or not be included in the model, there are a total of 2^k possible models, where k is the number of predictors. Bayesian model averaging will average all of these models together by weighting each model by its posterior probability.

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

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

8

16

1 / 1 points

7. Suppose that a MCMC sampler is currently visiting model B. Model A has a higher posterior probability than model B and Model C has a lower posterior probability than model B. Which of the following statements is true in the MCMC algorithm?

If a jump to Model C is proposed, this jump is always accepted.

If a jump to Model A is proposed, this jump is never accepted.

If a jump to Model A is proposed, this jump is always accepted.

Correct

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

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

If a jump to Model C is proposed, this jump is never accepted.

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

1 / 1 points

- 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.

- Examining the influence of potential outliers on the parameters of the model.
- Plotting the residuals to check for non-normally distributed residuals.

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

0 / 1 points

- It simplifies prior elicitation down to two components, the prior mean and g.
- It prevents BMA from disproportionately favoring the null model as a result of the Bartlett-Lindley paradox.
- It helps shrink the coefficients towards 0, which is important if the variables are highly correlated.

This should not be selected

Since the Zellner g-prior has a data-dependent covariance matrix, it helps put the prior variance on the same scale as the data, which is useful in calculating the posterior and eliciting the prior.

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

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

✖

0 / 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 "median probability model"?

- Including only the coefficients with posterior model inclusion probability above 0.5.
- Selecting the model that generates predictions most similar to those obtained from averaging over the model space.
- Selecting the model with the highest posterior model probability.

This should not be selected
The median probability model includes only the coefficients with posterior model inclusion probabilities above 0.5. Refer to lecture 5.4.4.

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

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