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

0 / 1 point

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 $N(0, 1)$.

Then the posterior mean of β will be closer to zero than the least squares estimate. True or False?

- False

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.

- True

0 / 1 point

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 other relevant covariates such as weight or income.
- ○ 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 terms of age^2 and or $\log(age)$ as covariates in the model.
- ○ Imposing strong prior distributions on the parameters in a Bayesian analysis.

0 / 1
point

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

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.

- 0.06
- 0.12
- 0.24

---

1 / 1
point

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

- We reject the null hypothesis of no difference, since the 95% credible interval does not include zero.

- 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.

Correct
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.

✖

0 / 1 point

5. Which of the following goes into the calculation of the Bayesian Information Criterion (BIC)?

- The maximum value of the log-likelihood under the current model

This should not be selected
The formula for the BIC is $-2L + \log(n)k$, where L corresponds to the maxmimum value of the log-likelihood, n is the sample size, and k is the number of parameters in the model.

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

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

- The maximum value of the log-likelihood under the current model, the sample size, and the number of parameters in the model

- The maximum value of the log-likelihood under the current model and the number of parameters in the model

- The maximum value of the log-likelihood under the current model, a constant penalty, and the number of parameters in the model

---

1 / 1 point

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
- 8

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

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

- 16

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?

1 / 1 point

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

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

0 / 1 point

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

This should not be selected
After fitting a linear model, it is important to make sure that the model does not violate any assumptions. In this case having an R^2 close to 1 is not a modeling assumption.

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.

✖

0 / 1 point

9. Which of the following is an advantage of using the Zellner-Siow-Cauchy prior in Bayesian model averaging?

- [ ] It helps shrink the coefficients towards 0, which is important if the variables are highly correlated
- [ ] It prevents BMA from disproportionately favoring the null model as a result of the Bartlett-Lindley paradox
- [ ] It allows for uncertainty in the prior variance parameter g

This should not be selected
For a fixed value of g, the Bartlett-Lindley paradox means that as the sample size increases, the null model will receive higher posterior probability. The Zellner-Siow-Cauchy prior allows g to increase with n to

avoid this problem while allowing for uncertainty about g, assigning n/g a Gamma prior distribution.

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

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

○ Both b and c

---

1 / 1 point

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

○ Selecting the model with the highest posterior model probability.

Correct
The median probability model includes only the coefficients with posterior model inclusion probabilities above 0.5.

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