

ⓘ Students have either already taken or started taking this quiz, so be careful about editing it. If you change any quiz questions in a significant way, you may want to consider regrading students who took the old version of the quiz.

Points 20 ✔ Published

Details

Questions

 Show Question Details**Question**

4 pts

Frequentist statistics define probabilities as _____ and is premised (that is, makes the assumption) on _____.

Answer

- frequencies of events in very large samples; imaginary resampling of data many many times
- degrees of belief; imaginary resampling of the data many many times
- objective way to measure uncertainty; the weak law of large numbers (LLN)
- frequencies of events in very large samples; null hypothesis significance testing (NHST)

Question

4 pts

Which statement about causal inference is **false**:

Answer

- enables statements of causality to be made without any assumptions
- it goes beyond association between variables
- it is prediction of interventions / treatments
- it is imputation of missing data

Question

4 pts

Which of the following statement corresponds to the expression: $Pr(\text{Monday}|\text{rain})$

Answer

- The probability it is Monday given that it is raining.
- The probability of rain on Monday.
- The probability of rain, given that it is Monday.
- The probability that it is Monday and that it is raining.

Question

4 pts

Per Lecture 2 and Chapter 2, the following statements are features of Bayesian inference **except which statement:**

answer

- It is best to use only point estimates and ignore the full posterior when making inference
- There is no minimum sample size to do Bayesian inference
- The shape of the posterior embodies the sample size (e.g., more samples induce a narrower/taller posterior)
- Intervals don't have a strong role in Bayesian inference as there isn't one true interval.

Question

4 pts

Bayesian inference can be thought of as:

answer

- counting up the ways the data can arise and normalizing the counts to get probabilities
- an approach to draw conclusions from sample-data by means of emphasizing the frequency or proportion of findings in the data.
- an approach to ensure causal inference, not just descriptive statistics
- way to do statistics without distributional assumptions

[+ New Question](#)[+ New Question Group](#)[Find Questions](#) Notify users this quiz has changed[Cancel](#)[Save](#)

ⓘ Students have either already taken or started taking this quiz, so be careful about editing it. If you change any quiz questions in a significant way, you may want to consider regrading students who took the old version of the quiz.

Points 20 ✔ Published

Details

Questions

 Show Question Details**Question**

4 pts

As described in the lecture and book, linear regression models have all the following attributes **except**:

- Robust to spurious correlations
- Descriptively accurate
- General method of approximation
- Mechanistically wrong

Answer

Question

4 pts

ANOVA, ANCOVA, t-tests, and MANOVA can be thought of as special cases of:

- Linear regression
- Bayesian methods
- Generalized Methods of Moments
- Unbiased and consistent estimators

Answer

Question

4 pts

Let's assume the Globe Tossing Model from Chapters 2-3.

$$W_i \sim \text{Binomial}(N, p)$$

$$p \sim \text{Uniform}(0, 1)$$

What is the prior distribution in this model?

- $\text{Uniform}(0, 1)$
- $\text{Binomial}(N, p)$

Answer

Binomial(N, p) * *Uniform*(0, 1)

Normal(0, p)

⋮ **Question**

4 pts

Somewhat similar to p-hacking, what is a way that priors can be misused?

Answer

choosing priors conditional on observed data that will be used to estimate the model

using only pre-data knowledge (e.g., constraints, ranges, and theoretical relationships) to set the prior

use an earlier posterior distribution later on as a prior for new data observed

eliciting priors from non-experts (e.g., crowdsourcing platforms like Mechanical Turk)

⋮ **Question**

4 pts

In Chapter 4, what are two different sources of uncertainty in Bayesian models?

Answer

uncertainty in parameter values and uncertainty in a sampling process

uncertainty in priors and uncertainty in data collection

uncertainty in parameter values and uncertainty in likelihood function

uncertainty in priors and uncertainty in a sampling process

+ [New Question](#)

+ [New Question Group](#)

🔍 [Find Questions](#)

Notify users this quiz has changed

Cancel

Save

ⓘ Students have either already taken or started taking this quiz, so be careful about editing it. If you change any quiz questions in a significant way, you may want to consider regrading students who took the old version of the quiz.

Points 20 ✔ Published

Details

Questions

 Show Question Details**Question**

4 pts



Assume the following model:

$$y_i \sim \text{Normal}(\mu, \sigma)$$

$$\mu_i = \alpha + \beta x_i$$

$$\alpha \sim \text{Normal}(0, 10)$$

$$\beta \sim \text{Normal}(0, 1)$$

$$\sigma \sim \text{Exponential}(1)$$

How many parameters are in the posterior distribution?

Answer

 3 5 7 8**Question**

4 pts

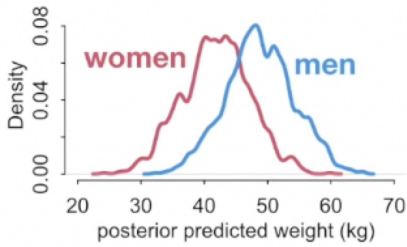
What is a benefit of using index variables for categorical variables instead of dummy or indicator variables?

Answer

 All of the above Generalize to many categorical levels without a change of code Make specifying prior easier Connect to multi-level (hierarchical) models**Question**

4 pts

Lecture 4 discussed the posterior predicted weights by men versus women. How do you measure the difference in predicted weights between men and women in the population?



- Answer
- calculate the contrast distribution on the individual level
 - calculate the contrast distribution by comparing the overlap in parameters
 - calculate p-values and run a null hypothesis significance test
 - compare whether the credible intervals for each category do not overlap

Question

4 pts

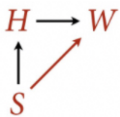
Polynomial regression and splines are two ways to add curvature into linear regression. What is a way splines are different than Polynomial regression?

- Answer
- Polynomial regression transforms predictors (e.g., squaring) while splines use Basis functions that turn a specific parameter on and off within a specific range of the real predictor variable
 - splines are not geocentric models as they better reflect true underlying mechanisms
 - splines are subject to problems with explosive uncertainty at edges and thus extrapolation challenges
 - All of the above

Question

2 pts

True or False: The red arrow in the DAG below represents the direct effect of Sex on Weight, not the total effect.

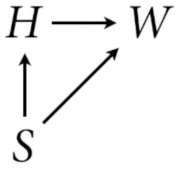


- Answer
- True
 - False

Question

2 pts

True or False: In this DAG, one way we can interpret this is that "weight is influenced by both height and sex."



Answer

True

False

+ [New Question](#)

+ [New Question Group](#)

[Find Questions](#)

Notify users this quiz has changed

[Cancel](#)

[Save](#)

ⓘ Students have either already taken or started taking this quiz, so be careful about editing it. If you change any quiz questions in a significant way, you may want to consider regrading students who took the old version of the quiz.

Points 20 ✔ Published

Details

Questions

 Show Question Details

⋮ Question

2 pts

True or False: Z is a collider confound.


 True

 False

Answer

⋮ Question

2 pts

True or False: To **stratify** or **condition** on a variable in linear regression means to add that variable as a predictor (independent) variable into the regression.

 True

 False

Answer

⋮ Question

2 pts

True or False: You should **always** include known collider variables into your regression as predictors to avoid collider bias.

 True

 False

Answer

⋮ Question

2 pts

True or False: The **distribution of** Weight that result from **intervening on Sex** can be written in do-calculus form as $p(W \mid do(S))$

- Answer
- True
- False

Question

4 pts

What is post-treatment bias?

- Answer
- biasing treatment effects due to the inclusion of variables that are a consequence of the treatment
- situation in which there is high association between predictor (independent) variables
- biasing treatment effects due to the omission of one or more relevant variables as predictor (independent) variables
- biasing treatment effects due to inclusion of a predictor (independent) variable that is associated with both the treatment and outcome (dependent) variable

Question

4 pts

What is the backdoor criterion?

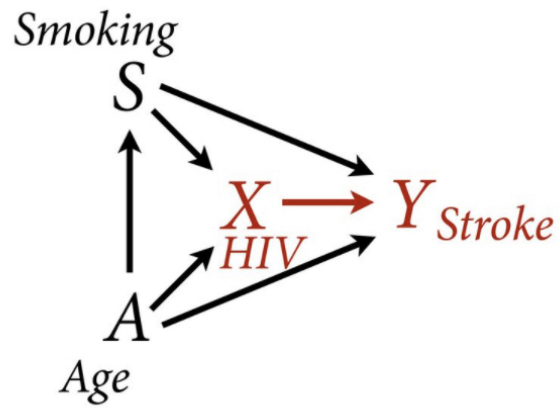
- Answer
- A rule to find a set of variables to stratify/condition to yield the causal effect of an intervention
- A condition to reduce post treatment bias in linear regressions
- The variables that we need to adjust for computing the causal effect from X on Y
- They are statements of which variables should be associated with one another (or not) in the data.

Question

4 pts

Let's assume this DAG from Lecture 6 in which we're interested in the effect of HIV on Stroke..

If we are interested in measuring **direct** effect of HIV on Stroke (**red arrow**), which variables should we include in the regression as predictors (independent variables)?



Answer

- X, S, and A
- X only
- X and S
- S and A

[+ New Question](#) [+ New Question Group](#) [Find Questions](#)

Notify users this quiz has changed

[Cancel](#) [Save](#)

ⓘ Students have either already taken or started taking this quiz, so be careful about editing it. If you change any quiz questions in a significant way, you may want to consider regrading students who took the old version of the quiz.

Points 20 ✔ Published

Details

Questions

 Show Question Details**Question**

2 pts

True or False: Regularizing (skeptical) priors, which slows the rate of learning, can make a model fit worse on the training sample but enable better out-of-sample predictions.

Answer

 True False**Question**

4 pts

Cross-validation is leaving out a small chunk of observations from our sample and evaluating the model on the observations that were left out.

What is a major difference in cross-validation from a Bayesian perspective?

Answer

 Bayesian cross-validation includes the entire posterior predictive distribution instead of point prediction, i.e., average over “whole shape” Only Bayesian cross-validation can use information values like AIC, BIC, WAIC, and PSIS. Bayesian approaches enable k-fold cross-validation Bayesian cross-validation can consider a penalty to adjust for model complexity**Question**

2 pts

True or False: For causal inference, it is best to select only the model with the best cross-validation or information criteria like CV/PSIS/WAIC and disregard all other candidate models.

Answer

 True False**Question**

4 pts

Typically, the best (most effective) methods in Markov Chain Monte Carlo like Hamiltonian Monte Carlo rely on

Answer

- gradients that approximate the curvature of the estimated posterior
- grid approximation that ensure all possible points are estimated
- splines to interpolate between estimated points
- assuming an appropriate conjugate prior to ensure a closed form analytical solution for the posterior

Question

4 pts

All of the following are downsides of Metropolis algorithms **except**

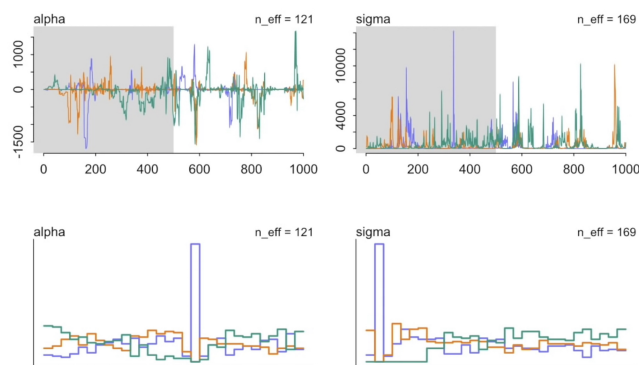
Answer

- it requires the (local) curvature of the posterior's shape
- require tuning for hyperparameters like step size
- it is inefficient as it will have more proposals rejected
- parameters are correlated

Question

2 pts

True or False: These trace and trunk plots indicate a well fit and converged Markov chain.



Answer

- True
- False

Question

2 pts

Let's assume we have the following WAIC values for three models: m1, m2, and m3.

	WAIC	SE	dWAIC	dSE	pWAIC	weight
m1	361.9	14.26	0.0	NA	3.8	1
m2	402.8	11.28	40.9	10.48	2.6	0
m3	405.9	11.65	44.0	12.23	1.6	0

True or False: m3 had the best model fit as it had the largest WAIC value.

True

False

Answer

+ [New Question](#)

+ [New Question Group](#)

[Find Questions](#)

Notify users this quiz has changed

[Cancel](#)

[Save](#)

ⓘ Students have either already taken or started taking this quiz, so be careful about editing it. If you change any quiz questions in a significant way, you may want to consider regrading students who took the old version of the quiz.

Points 20 ✔ Published

Details

Questions

 Show Question Details**Question**

2 pts

True or False: A linear regression generalizes to a Generalized Linear Model if its expected value (e.g., μ) is some function (e.g., logit) of additive combination of parameters.

Answer

 True False**Question**

4 pts

What is post-stratification?

Answer

 a technique for re-weighting estimates for a target population a technique to estimate treatment effects by comparing the changes in outcomes over time between the treatment and control groups a technique to control for confounding and measurement error in observational studies so that causal inferences can be made

a design that aims to determine the causal effects of interventions by assigning a cutoff or threshold above or below which an intervention is assigned

Question

2 pts

True or False: It is impossible to ever have a model with more parameters than observations.

Hint:



True

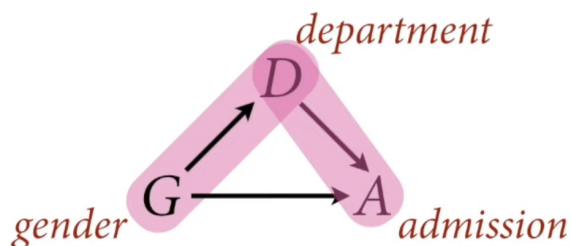
Answer

False

⋮ Question

2 pts

True or False: this path highlighted in red for the UC Berkeley application would represent taste/status-based discrimination.



True

Answer

False

⋮ Question

2 pts

True or False: to get the posterior probabilities for a GLM model, you need to apply the inverse link function to posterior estimates (e.g., output/result of `extract.samples()` function)

True

Answer

False

⋮ Question

4 pts

All of the following are true statements about Sensitivity Analysis (as described in Lecture 10) **except**:

a way to calibrate priors to previous knowledge

a way to measure what we don't know

assumes unobserved confounds and measures their possible influence

is a blend of both pure simulation and pure analysis

Answer

⋮ Question

4 pts

All of the following are true about Poisson Regression and Binomial Regressions **except** which statement:

answer

- Poisson regression uses a Poisson link function while Binomial regression uses a Binomial link function
- Poisson regressions are sometimes called log linear regressions
- Poisson sometimes uses exponential scaling to ensure positive values
- Poisson regressions are a subset of binomial that have no maximum count and low probability of occurrence

[+ New Question](#)[+ New Question Group](#)[🔍 Find Questions](#) Notify users this quiz has changed[Cancel](#)[Save](#)

ⓘ Students have either already taken or started taking this quiz, so be careful about editing it. If you change any quiz questions in a significant way, you may want to consider regrading students who took the old version of the quiz.

Points 20 ✔ Published

Details

Questions

 Show Question Details**Question**

4 pts



What is an appropriate way to handle an ordered frequency distribution as a dependent variables?

- Transform it to a cumulative frequency distribution
- Use a negative binomial as the likelihood function
- Implement a zero inflated Beta regression
- Normalize the categorizes by the largest category frequency

Question

4 pts

What is endogenous selection and what are problems it can cause?

- When a sample is selected on a collider (e.g., participation) and it induces non-causal associations (confounds)
- When the variance of the response variable is endogenous yielding inefficient parameter estimates
- When a sample is missing certain observations that can lead to bias in parameter estimates
- When a sample is selected on endogenous variables leading to problems of multicollinearity

Question

4 pts

What prior distribution aligns to monotonic ordered predictor variables?

Hint: this is when the parameters form a simplex (i.e., sum to 1)

- Dirichlet
- Poisson
- Gamma
- Lognormal

Question

4 pts

All of the following are benefits of using multilevel regression models versus classic (fixed effects) regressions **except**

Answer

- They converge faster/simpler to run
- They learn faster/better
- They resist overfitting
- They adaptively regularize

Question

4 pts

All of the following are consistent with partial pooling **except**

Answer

- occurs when we treat all clusters (groups) as identical
- enables priors to be set adaptively
- can learn a prior that is expected to provide the best out-of-sample accuracy
- shrinks individual values towards population means

[+ New Question](#)[+ New Question Group](#)[Find Questions](#) Notify users this quiz has changed[Cancel](#)[Save](#)

ⓘ Students have either already taken or started taking this quiz, so be careful about editing it. If you change any quiz questions in a significant way, you may want to consider regrading students who took the old version of the quiz.

Points 20 ✔ Published

Details

Questions

 Show Question Details**Question**

4 pts

All of the following are examples of cluster/features **except**:

- cluster: frog survival, feature: tanks
- cluster: individual, feature: average response
- cluster: stories, feature: treatment effect
- cluster: departments, feature: admissions bias

Answer

Question

2 pts

True or False: one of the best ways to handle divergent transitions in Hamiltonian Monte Carlo (HMC) is to reparameterize priors to be centered.

- True
- False

Answer

Question

4 pts

In multi-level models, what is the appropriate way to predict new records from new groups (i.e., those that haven't been seen in training sample)?

- ignore varying effects and marginalize over the population distribution
- use the model's varying effects
- use the most posterior from the most similar group as a proxy for the new group
- bootstrap new means from the other groups to infer the new group's posterior

Answer

Question

4 pts

What is a primary benefit of using varying effects models?

Answer

- enables using priors that learn correlations across features
- implements fixed treatment effects
- assumes the clusters are independent of one another
- can be used to calculate Bayesian p-values (Bayes Factors)

⋮ **Question**

2 pts

True or False: Fixed effects modeling does not use pooling as information is not transferred across any individual records.

Answer

- True
- False

⋮ **Question**

4 pts

What does partial pooling for treatment effects (beta) enable?

Answer

- Enable treatments to have varying effects depending on the individual
- Enable better and faster model convergence
- Allows to treat all individuals as one group that reduces overfitting
- Ensure we can find the best unbiased estimate for treatment effects

[+ New Question](#)[+ New Question Group](#)[🔍 Find Questions](#) Notify users this quiz has changed[Cancel](#)[Save](#)