

Computational Cognitive Science

Quiz 2

This quiz is anonymous. It covers topics from lecture and the readings as well as some mathematical background. Please fill in the appropriate circles with a dark pen or pencil.

1. What is a data model?
 - 0.28 A model that captures a set of experimental data
 - 0.28 **A mathematical way of relating the predictions of a model to experimental data**
 - 0.11 A model that needs to be trained on experimental data
 - 0.33 A model that assigns a likelihood to experimental data (*not a bad answer, but requires an underlying model that makes predictions*)
2. What is the mode of the likelihood function?
 - 0.71 **Its maximum** (*Strictly speaking, the value for which the function takes its maximum*)
 - 0.00 Its minimum
 - 0.24 Its mean
 - 0.06 Its standard error
3. In the context of comparing models with the AIC and the likelihood-ratio test, how do we measure the complexity of a model?
 - 0.00 As the number of lines of code that implement the model
 - 0.06 As the number of data points used to fit the model
 - 0.06 As the number of concepts required to define the model
 - 0.89 **As the number of parameters**
4. Which of the following are properties of General Recognition Theory (GRT)? (More than one may apply.)
 - 0.61 **It assumes category boundaries**
 - 0.28 It stores all exemplars of a category
 - 0.22 The perception of features is subject to random variation** (*Perceptual noise, which is taken to be random, is added to items to be classified*)
 - 0.44 It generates responses deterministically, conditional a specific set of perceived features**
5. Which of the following behaviors are predicted by the Generalized Context Model (GCM)? (More than one may apply.)
 - 0.44 If exemplar features are the same, participants match the label probabilities in the input**
 - 0.28 Participants perceive the input as noisy
 - 0.33 All participants show the same response function
 - 0.67 **Participants base their categorization decisions on feature similarity**
 - 0.22 Distance and similarity are linearly related

6. What is the derivative of $w_1x_1 + w_2x_2$ with respect to x_1 ?
- 0.00 w_1x_1
 - 0.00 w_2x_2
 - 0.94 w_1
 - 0.00 x_1
 - 0.00 other:
 - 0.06 Don't know
7. What is $\log(ce^{(x-u)^2/s})$ if all terms are positive?
- 0.05 $c + \frac{(x-u)^2}{s}$
 - 0.33 $\log(c) + 2\log(x-u) - \log(s)$
 - 0.67 $\log(c) + \frac{(x-u)^2}{s}$
 - 0.00 other:
 - 0.06 Don't know
8. Which of these inequalities suggests we have made a mistake? (More than one may apply.)
- 0.22 $p(x) > 1$
 - 0.89 $P(x) > 1$
 - 0.39** $-\log(P(x)) < 0$ (*Probabilities are always ≤ 1 , so logs are never positive, so negative logs are never negative*)
 - 0.56 $\int_{x \in \mathcal{X}} p(x) dx < 1$
 - 0.11 $\int_{x \in \mathcal{X}} xp(x) dx > 1$
9. What is Bayesian integration?
- 0.41 instead of maximizing the posterior, we take the posterior mean**
 - 0.53 instead of maximizing the likelihood, maximize the posterior (*This describes a MAP estimate*)
 - 0.06 instead of a uniform prior, we assume an exponential prior
10. Which of the following is a lexical decision task?
- 0.00 participants see words and have to read them
 - 0.00 participants hear words and have to memorize them
 - 0.89 participants see letter strings and have to decide whether they are words or non-words**
 - 0.06 participants see letter strings and have to pronounce them
 - 0.11 participants see letter strings and have to determine what word they are a prefix for
11. What is incremental processing?
- 0.63 new input is processed as soon as it is perceived, e.g., letter by letter or word by word**
 - 0.37 new input is processed in fixed increments, e.g., every ten letters or every ten words (*This is incremental, but it isn't what people tend to mean by "incremental processing"*)
 - 0.0 input is processed using Bayesian integration
 - 0.0 perceived input consists of random samples normally distributed around the true input
12. Which of the following are neighborhood effects one sees in word recognition tasks? (More than one may apply.)
- 0.67 during word recognition, words with more neighbors are processed more slowly** (*Better to have called this "word identification"*)
 - 0.22 during word recognition, words with more neighbors are processed faster
 - 0.67 during lexical decision, words with more neighbors are processed faster**
 - 0.39 during lexical decision, non-words with more neighbors are processed more slowly