CS698X: Topics In Probabilistic Modeling And Inference

Users Online : 18

Quiz 1

Q.1 Which of the following statements are true about Bayesian inference?

Max. score: 2; Neg. score: 0; Your score: 0

- [ ] For computing the posterior distribution, the denominator in the Bayes rule (the marginal likelihood) is not necessary to compute.
- [ ] Fully Bayesian inference with a uniform prior will be equivalent to doing point estimation.
- [x] Uncertainty of the posterior predictive decreases as the amount of training data increases.
- [x] Uncertainty of the posterior distribution decreases as the amount of training data increases.

Q.2 Given i.i.d. data $X = \{x_1, x_2, \ldots, x_N\}$ from a distribution $p(x|\theta)$, the following must hold: $\int p(X|\theta)d\theta = 1$

Max. score: 1; Neg. score: 0; Your score: 1

- (x) false
- ( ) true

Q.3 Which of the following is true about the concentration parameter vector α of a Dirichlet distribution?

Max. score: 2; Neg. score: 0; Your score: 0

- [ ] When the elements of α are very small (close to 0), the Dirichlet becomes similar to a Gaussian distribution
- [ ] When the elements of α are very small (close to 0), the Dirichlet becomes similar to a uniform distribution
- [ ] When the elements of α are all equal, the Dirichlet becomes peaked (concentrated) towards the center
- [x] When the elements of α are all equal to 1, the Dirichlet becomes similar to a uniform distribution

---

Q.4 For probabilistic linear regression, the optimal hyperparameters can be found by

- [x] Cross-validation if held-out data is available
- [x] MAP-II (maximizing the posterior of the hyperparameters)
- [x] MCMC or variational inference

Q.5 Which of the following statements is true regarding conjugacy?

Max. score: 2; Neg. score: 0; Your score: 2

- [x] If likelihood and prior are conjugate and in exponential family, the posterior has a closed form expression
- [ ] Every distribution has an associate conjugate distribution
- [x] If likelihood and prior are conjugate and in exponential family, the posterior predictive has a closed form expression
- [x] If likelihood and prior are conjugate and in exponential family, the marginal likelihood has a closed form expression

---

Q.6 Posterior predictive distribution will become very similar to the plug-in predictive if

Max. score: 2; Neg. score: 0; Your score: 0

- [ ] Number of observations is very small
- [x] Number of observations is very large
- [ ] If the likelihood and prior are exponential family distributions
- [ ] If we use a uniform prior distribution

Q.7 Large uncertainty in the posterior distribution implies large uncertainty in the posterior predictive distribution.

Max. score: 1; Neg. score: 0; Your score: 0

- ( ) false
- (x) true

Q.8 Any probability distribution can be represented as an exponential family distribution.

Max. score: 1; Neg. score: 0; Your score: 1

Q.9 For a model with two parameters θ_1 and θ_2 and data X, the overall posterior of the parameters will be

Max. score: 2; Neg. score: 0; Your score: 2

- [ ] Defined by two distributions $p(\theta_1|X)$ and $p(\theta_2|X)$
- [x] Proportional to $p(\theta_1, \theta_2, X)$
- [x] $p(\theta_1, \theta_2|X)$
- [ ] Defined by two distributions $p(\theta_1|X, \theta_2)$ and $p(\theta_2|X, \theta_1)$

Q.10 Using a MAP estimation approach, which of the following are possible?

Max. score: 2; Neg. score: 0; Your score: 2

- [ ] We can obtain predictions that are based on posterior averaging over all parameters of all the models under consideration.
- [ ] We can obtain predictions that are based on posterior averaging over the parameters of a fixed model.
- [x] We can prevent overfitting.
- [x] Predictions can be made only using the plug-in predictive distribution.

Q.11 The regions with very few training observations will typically have a very small variance for the posterior predictive distribution.

Max. score: 1; Neg. score: 0; Your score: 1

- [x] false
- [ ] true

Q.12 Give a brief, intuitive justification as to why the posterior predictive distribution of a probabilistic linear regression model (or, in fact, of any model) would typically have a larger variance than the plug-in predictive distribution. Write your answer in the provided text box only (no file upload).

Max. score: 3; Neg. score: 0; Your score: 0

Your answer:

plug in predictive will take up only the variance of the prior

Q.13 Assuming two random variables x_a and x_b (assume both to be scalars), having a Gaussian joint distribution, which of the following statements is true?

Max. score: 2; Neg. score: 0; Your score: 2

- ✔ ■ Conditioning one random variable x_a on a given value of another random variable x_b decreases the variance of x_a
- ☐ Conditioning one random variable x_a on a given value of another random variable x_b increases the variance of x_a
- ✔ ■ The mean of the conditional distribution of x_a, given x_b, depends on x_b
- ☐ The mean of the conditional distribution of x_a, given x_b, is independent of x_b

Q.14 Using a prior distribution, which of the following can be achieved?

Max. score: 2; Neg. score: 0; Your score: 0

- ■ Combining it with a likelihood to compute an exact posterior in all cases.
- ✔ ■ Regularizing the MLE approach.
- ✔ ☐ Pre-specify which parameter values are not likely for the solution
- ✔ ■ Compute (exactly/approximate) the marginal likelihood of a model.

Q.15 Consider rolling a dice 10 times and assume the number of times each of the 6 faces show up to be 2,1,3,0,1,3, respectively. Assume a Dirichlet(3,3,3,3,3,3) prior on the probability vector π

Max. score: 2; Neg. score: 0; Your score: 0

- ■ The probability of the next dice roll showing the 4th face, given the full posterior of π is 1/6
- ✔ ■ The probability of the next dice roll showing the 4th face, given the MLE solution of π is 0
- ☐ The probability of the next dice roll showing the 4th face is undefined for MLE/MAP as well as for fully Bayesian solution.
- ☐ The probability of the next dice roll showing the 4th face, given the MAP solution of π is 0

Q.16 State two advantages of fully Bayesian inference over MAP estimation. Write your answer in the provided text box only (no file upload).

gives the variance in parameter estimation
Model selection and hyper parameter tuning

Feedback:

Score: 14