Назад

Lesson 2

Тест, 8 вопроса

Вопрос 1

1

Баллы

1. Вопрос 1

Which of the following is one major difference between the frequentist and Bayesian approach to modeling data?

- [ ] The frequentist paradigm treats the data as fixed while the Bayesian paradigm considers data to be random.
- [ ] Frequentist models are deterministic (don't use probability) while Bayesian models are stochastic (based on probability).
- [x] Frequentists treat the unknown parameters as fixed (constant) while Bayesians treat unknown parameters as random variables.
- [ ] Frequentist models require a guess of parameter values to initialize models while Bayesian models require initial distributions for the parameters.

Вопрос 2

1

Баллы

2. Вопрос 2

Suppose we have a statistical model with unknown parameter θ\thetaθ, and we assume a normal prior θ∼N(μ0,σ02) \theta \sim \text{N}(\mu_0, \sigma_0^2) θ∼N(μ0,σ02), where μ0\mu_0μ0 is the prior mean and σ02

\sigma_0^2 σ_0^2 is the prior variance. What does increasing σ_0^2\sigma_0^2σ_0^2 say about our prior beliefs about θ \theta θ?

Increasing the variance of the prior **widens** the range of what we think θ\theta$\theta$ might be, indicating **greater** confidence in our prior mean guess μ_0 \mu_0 μ_0.

Increasing the variance of the prior **narrows** the range of what we think θ\theta$\theta$ might be, indicating **greater** confidence in our prior mean guess μ_0 \mu_0 μ_0.

Increasing the variance of the prior **widens** the range of what we think θ\theta$\theta$ might be, indicating **less** confidence in our prior mean guess μ_0 \mu_0 μ_0.

Increasing the variance of the prior **narrows** the range of what we think θ\theta$\theta$ might be, indicating **less** confidence in our prior mean guess μ_0 \mu_0 μ_0.

Вопрос 3

1

Баллы

3. Вопрос 3

In the lesson, we presented Bayes' theorem for the case where parameters are continuous. What is the correct expression for the posterior distribution of θ\theta$\theta$ if it is discrete (takes on only specific values)?

$p(\theta|y)=p(y|\theta)\cdot p(\theta)\int p(y|\theta)\cdot p(\theta)d\theta$ p(\theta \mid y) = \frac{ p(y \mid \theta) \cdot p(\theta) }{ \int p(y \mid \theta) \cdot p(\theta)\, d\theta } $p(\theta|y)=\int p(y|\theta)\cdot p(\theta)d\theta p(y|\theta)\cdot p(\theta)$

$p(\theta)=\int p(\theta|y)\cdot p(y)dy$ p(\theta) = \int p(\theta \mid y) \cdot p(y) \, dy $p(\theta)=\int p(\theta|y)\cdot p(y)dy$

p(θj∣y)=p(y∣θj)·p(θj)∑jp(y∣θj)·p(θj) p(\theta_j \mid y) = \frac{ p(y \mid \theta_j) \cdot p(\theta_j) }{ \sum_j p(y \mid \theta_j) \cdot p(\theta_j) } p(θj∣y)=∑jp(y∣θj)·p(θj)p(y∣θj)·p(θj)

○

p(θ)=∑jp(θ∣yj)·p(yj) p(\theta) = \sum_j p(\theta \mid y_j) \cdot p(y_j) p(θ)=∑jp(θ∣yj)·p(yj)

Вопрос 4

1

Баллы

4. Вопрос 4

For Questions 4 and 5, refer to the following scenario.

In the quiz for Lesson 1, we described Xie's model for predicting demand for bread at his bakery. During the lunch hour on a given day, the number of orders (the response variable) follows a Poisson distribution. All days have the same mean (expected number of orders). Xie is a Bayesian, so he selects a conjugate gamma prior for the mean with shape 3 3 3 and rate 1/15 1 / 15 1/15. He collects data on Monday through Friday for two weeks.

Which of the following hierarchical models represents this scenario?

○

$y_i|\mu \sim$ iid $N(\mu,1.0^2)$ for $i=1,\ldots,10$, $\mu \sim N(3,15^2)$

○

$y_i|\lambda_i \sim$ ind $Pois(\lambda_i)$ for $i=1,\ldots,10$, $\lambda_i|\alpha \sim Gamma(\alpha,1/15)$ $\alpha \sim Gamma(3.0,1.0)$

◉

$y_i|\lambda \sim$ iid $Pois(\lambda)$ for $i=1,\ldots,10$, $\lambda \sim Gamma(3,1/15)$

○

$y_i|\lambda \sim$ iid $Pois(\lambda)$ for $i=1,\ldots,10$, $\lambda|\mu \sim Gamma(\mu,1/15)$ $\mu \sim N(3,1.0^2)$

Вопрос 5

1

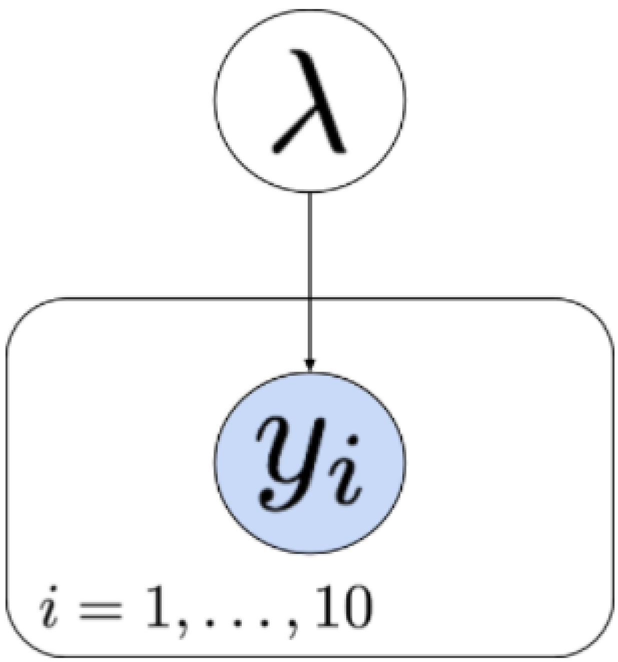
Баллы

5. Вопрос 5

Which of the following graphical depictions represents the model from Xie's scenario?

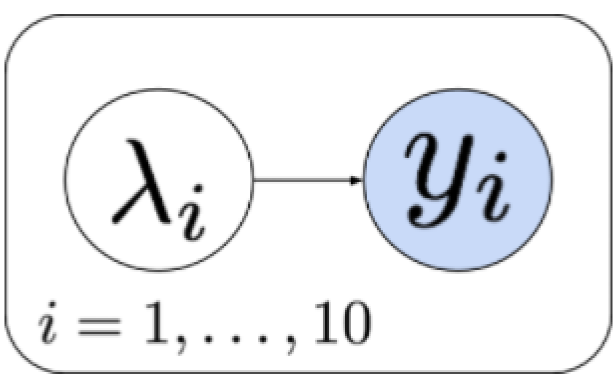
◉

a)



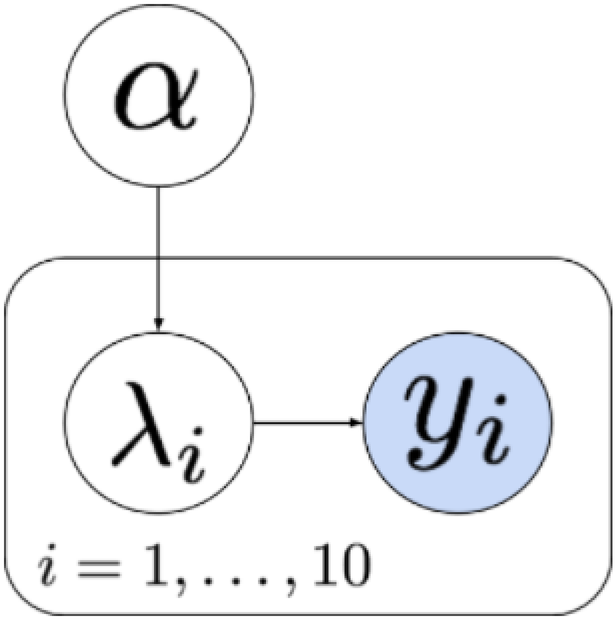
○

b)

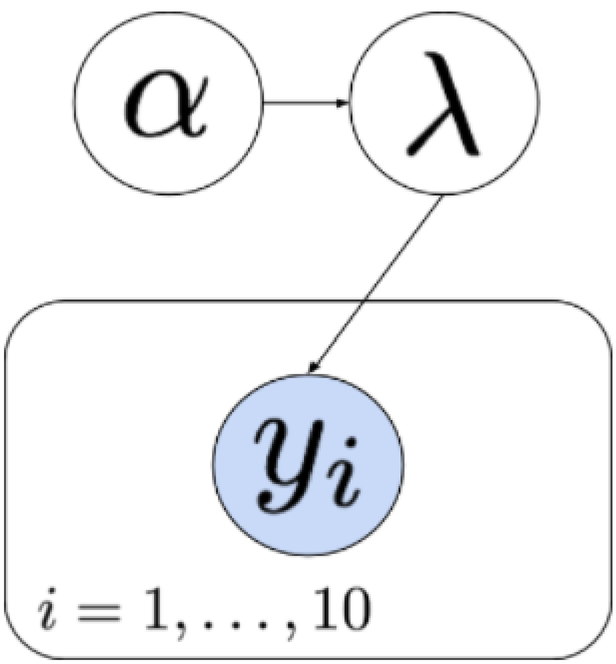


○

c)



d)



Вопрос 6

1

Баллы

6. Вопрос 6

Graphical representations of models generally do not identify the distributions of the variables (nodes), but they do reveal the structure of dependence among the variables.

Identify which of the following hierarchical models is depicted in the graphical representation below.

Не удается отобразить связанный рисунок. Возможно, этот файл был перемещен, переименован или удален. Убедитесь, что ссылка указывает на правильный файл и верное размещение.

( )

$$x_{i,j} \mid \alpha, \beta \overset{\text{iid}}{\sim} \text{Gamma}(\alpha, \beta), \quad i = 1, \ldots, n, \quad j = 1, \ldots, m \quad \beta \sim \text{Exp}(b_0) \quad \alpha \sim \text{Exp}(a_0) \quad \phi \sim \text{Exp}(r_0)$$

( )

$$x_{i,j} \mid \alpha_j, \beta \overset{\text{ind}}{\sim} \text{Gamma}(\alpha_j, \beta), \quad i = 1, \ldots, n, \quad j = 1, \ldots, m \quad \beta \sim \text{Exp}(b_0) \quad \alpha_j \sim \text{Exp}(a_0), \quad j = 1, \ldots, m \quad \phi \sim \text{Exp}(r_0)$$

( )

$$x_{i,j} \mid \alpha_i, \beta_j \overset{\text{ind}}{\sim} \text{Gamma}(\alpha_i, \beta_j), \quad i = 1, \ldots, n, \quad j = 1, \ldots, m \quad \beta_j \mid \phi \overset{\text{iid}}{\sim} \text{Exp}(\phi), \quad j = 1, \ldots, m \quad \alpha_i \mid \phi \overset{\text{iid}}{\sim} \text{Exp}(\phi), \quad i = 1, \ldots, n \quad \phi \sim \text{Exp}(r_0)$$

(•)

$$x_{i,j} \mid \alpha_j, \beta \overset{\text{ind}}{\sim} \text{Gamma}(\alpha_j, \beta), \quad i = 1, \ldots, n, \quad j = 1, \ldots, m \quad \beta \sim \text{Exp}(b_0) \quad \alpha_j \mid \phi \overset{\text{iid}}{\sim} \text{Exp}(\phi), \quad j = 1, \ldots, m \quad \phi \sim \text{Exp}(r_0)$$

Вопрос 7

1

Баллы

7. Вопрос 7

Consider the following model for a binary outcome y:

$$y_i \mid \theta_i \overset{\text{ind}}{\sim} \text{Bern}(\theta_i), \quad i = 1, \ldots, 6 \quad \theta_i \mid \alpha \overset{\text{iid}}{\sim} \text{Beta}(\alpha, b_0), \quad i = 1, \ldots, 6 \quad \alpha \sim \text{Exp}(r_0)$$

where θ_i is the probability of success on trial i. What is the expression for the joint distribution of all variables, written as $p(y_1, \ldots, y_6, \theta_1, \ldots, \theta_6, \alpha)$ and denoted by $p(\cdots)$? You may ignore the indicator functions specifying the valid ranges of the variables (although the expressions are technically incorrect without them).

Hint:

The PMF for a Bernoulli random variable is $f_y(y \mid \theta) = \theta^{y} (1-\theta)^{1-y}$ for $y=0$ or $y=1$ and $0 < \theta < 1$.

The PDF for a Beta random variable is $f_\theta(\theta \mid \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\beta)} \theta^{\alpha - 1} (1 - \theta)^{\beta - 1}$ where $\Gamma()$ is the gamma function, $0 < \theta < 1$ and $\alpha, \beta > 0$.

The PDF for an exponential random variable is $f_\alpha(\alpha \mid \lambda) = \lambda \exp(-\lambda \alpha)$ for $\lambda, \alpha > 0$.

$p(\cdots) = \prod_{i=1}^6 \left[\theta_i^{y_i} (1-\theta_i)^{1-y_i} \frac{\Gamma(\alpha + b_0)}{\Gamma(\alpha) \Gamma(b_0)} \theta_i^{\alpha - 1} (1 - \theta_i)^{b_0 - 1} \right]$

$p(\cdots) = \prod_{i=1}^6 \left[\theta_i^{y_i} (1-\theta_i)^{1-y_i} \right] \cdot \frac{\Gamma(\alpha + b_0)}{\Gamma(\alpha) \Gamma(b_0)} \theta^{\alpha - 1} (1 - \theta)^{b_0 - 1} \cdot r_0 \exp(-r_0 \alpha)$

$p(\cdots) = \prod_{i=1}^6 \left[\theta_i^{y_i} (1-\theta_i)^{1-y_i} \frac{\Gamma(\alpha + b_0)}{\Gamma(\alpha) \Gamma(b_0)} \theta_i^{\alpha - 1} (1 - \theta_i)^{b_0 - 1} \right] \cdot r_0 \exp(-r_0 \alpha)$

$p(\cdots) = \prod_{i=1}^6 \left[\theta_i^{y_i} (1-\theta_i)^{1-y_i} \frac{\Gamma(\alpha + b_0)}{\Gamma(\alpha) \Gamma(b_0)} \theta_i^{\alpha - 1} (1 - \theta_i)^{b_0 - 1} r_0 \exp(-r_0 \alpha) \right]$

Вопрос 8

1

Баллы

8. Вопрос 8

In a Bayesian model, let y denote all the data and θ denote all the parameters. Which of the following statements about the relationship between the joint distribution of all variables $p(y, \theta) = p(\cdots)$ and the posterior distribution $p(\theta \mid y)$ is true?

Neither is sufficient alone--they are both necessary to make inferences about θ \theta θ.

They are proportional to each other so that $p(y,\theta)=c\cdot p(\theta|y)$ p(y, \theta) = c \cdot p(\theta \mid y) $p(y,\theta)=c\cdot p(\theta|y)$ where c c c is a constant number that doesn't involve θ \theta θ at all.

The joint distribution $p(y,\theta)$p(y,\theta)$p(y,\theta)$ is equal to the posterior distribution times a function $f(\theta)$ f(\theta) $f(\theta)$ which contains the modification (update) of the prior.

They are actually equal to each other so that $p(y,\theta)=p(\theta|y)$p(y, \theta) = p(\theta \mid y)$p(y,\theta)=p(\theta|y)$.

Оплатить курс

--> .

□