QF624-2025-W6

Number of participants: 30

1. When working in the log-domain, we write $\log P(\theta \mid x) \propto \log P(x \mid \theta) + \log P(\theta)$. Which of the following is \emph{not} a main advantage of using the log-form?

2 correct answers
out of 20 respondents

	Converts products into sums, simplifying differentiation in optimization.	10%	2 votes
	Improves numerical stability by avoiding underflow for very small probabilities.	60%	12 votes
✓	Ensures the posterior is a convex function of θ.	10%	2 votes
	Allows one to drop the intractable normalizing constant $\log P(x)$ when optimizing.	20%	4 votes

2. When writing the MAP objective as $\min_\theta \left[-\log P(\mathcal{D} \mid \theta) - \log P(\theta)\right]$, we drop the term $\log P(\mathcal{D})$ because:

8 correct answers
out of 18 respondents

	It depends on θ.	11%	2 votes
	It is identically zero.	22%	4 votes
✓	It does not depend on θ.	44%	8 votes
	It equals the log-prior.	22%	4 votes

3. Under what assumption does MAP estimation become equivalent to maximum likelihood estimation (MLE)?

14 correct answers
out of 18 respondents

	The likelihood $P(\mathcal{D} \mid \theta)$ is uniform in θ.	6%	1 vote
✓	The prior $P(\theta)$ is uniform (constant) in θ.	78%	14 votes
	The evidence $P(\mathcal{D})$ is maximized.	11%	2 votes
	The dataset size $\lvert\mathcal{D}\rvert \to \infty$.	6%	1 vote

4. Which symmetry property holds for the sigmoid?

1 correct answer
out of 16 respondents

	$f(-z) = 1 + f(z)$	0%	0 votes
✓	$f(-z) = 1 - f(z)$	6%	1 vote
	$f(-z) = \frac{f(z)}{1 - f(z)}$	75%	12 votes
	$f(-z) = -f(z)$	19%	3 votes

5. You raise the decision threshold of a probabilistic classifier (i.e. require a higher p to call “positive”). Which of these metric-changes is always true (all else held fixed)?

9 correct answers
out of 13 respondents

✓	Precision increases, Recall decreases.	69%	9 votes
	Recall increases, Precision decreases.	15%	2 votes
	Specificity decreases, Error rate decreases.	8%	1 vote
	Accuracy necessarily increases.	8%	1 vote

6. In the context of cost-sensitive logistic regression for an imbalanced dataset, if misclassifying a minority class instance (False Negative) is deemed 5 times more costly than misclassifying a majority class instance (False Positive), how would this typically be implemented?

7 correct answers
out of 10 respondents

	By oversampling the minority class by a factor of 5 before training.	10%	1 vote
	By setting the decision threshold to 5.0 after training the standard logistic regression model.	10%	1 vote
✓	By multiplying the loss contribution of each minority class instance by 5 during the model training process.	70%	7 votes
	By applying L1 regularization with a lambda value of 5 to penalize complexity related to majority class features.	10%	1 vote