Q&A Session Week 1

COMP90051 Statistical Machine Learning

Sem2 2020
Lecturer: Ben Rubinstein

THE UNIVERSITY OF MELBOURNE

This session

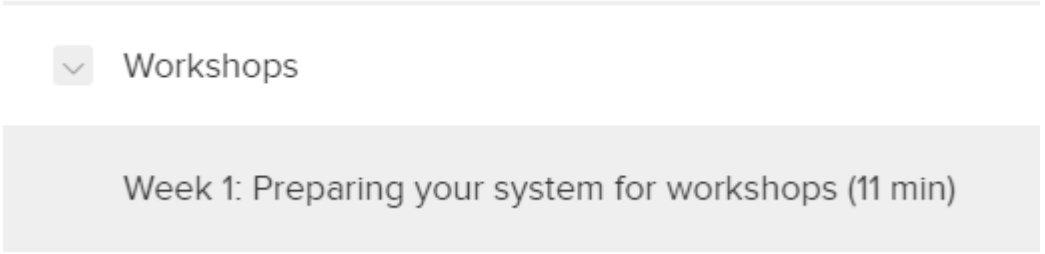
- Housekeeping
- Announcements
- Feedback to Ben
- Human ensembles
- Revising a lecture / Open Q&A
- Quiz discussion with class stats

Housekeeping

- Thankyou for coming*!!* Students expected to attend if Internet/commitments permit
- Session is **recorded** for posting to Canvas
- Etiquette
  - Write in chat/raise hand anytime
  - **Stay muted** unless speaking

Announcements

- Workshops begin next week (#2), please attend!
  - But first: install Anaconda – see Neil's video (lecture capture)



- Piazza discussion board… OMG please signup!

Poll: How’re the lecture videos going?

Human Ensemble Time

1. Who are you? What degree are you in?
2. What are you hoping to learn from StatML?
3. What's one thing you like to do in your spare time?

Discuss a lecture video

Quiz Discussion

Spoilers Alert
Spoilers Alert
Spoilers Alert
Spoilers Alert
Spoilers Alert
Spoilers Alert
Spoilers Alert
Spoilers Alert
Spoilers Alert

Question 1 — 1 pts

The MLE says to learn/estimate from data $X_1, \ldots, X_n$ by optimising the expression $\hat{\theta}(X_1, \ldots, X_n) \in \arg\max_{\theta \in \Theta} \prod_{i=1}^{n} p_\theta(X_i)$.

Why is the objective function a **product**?

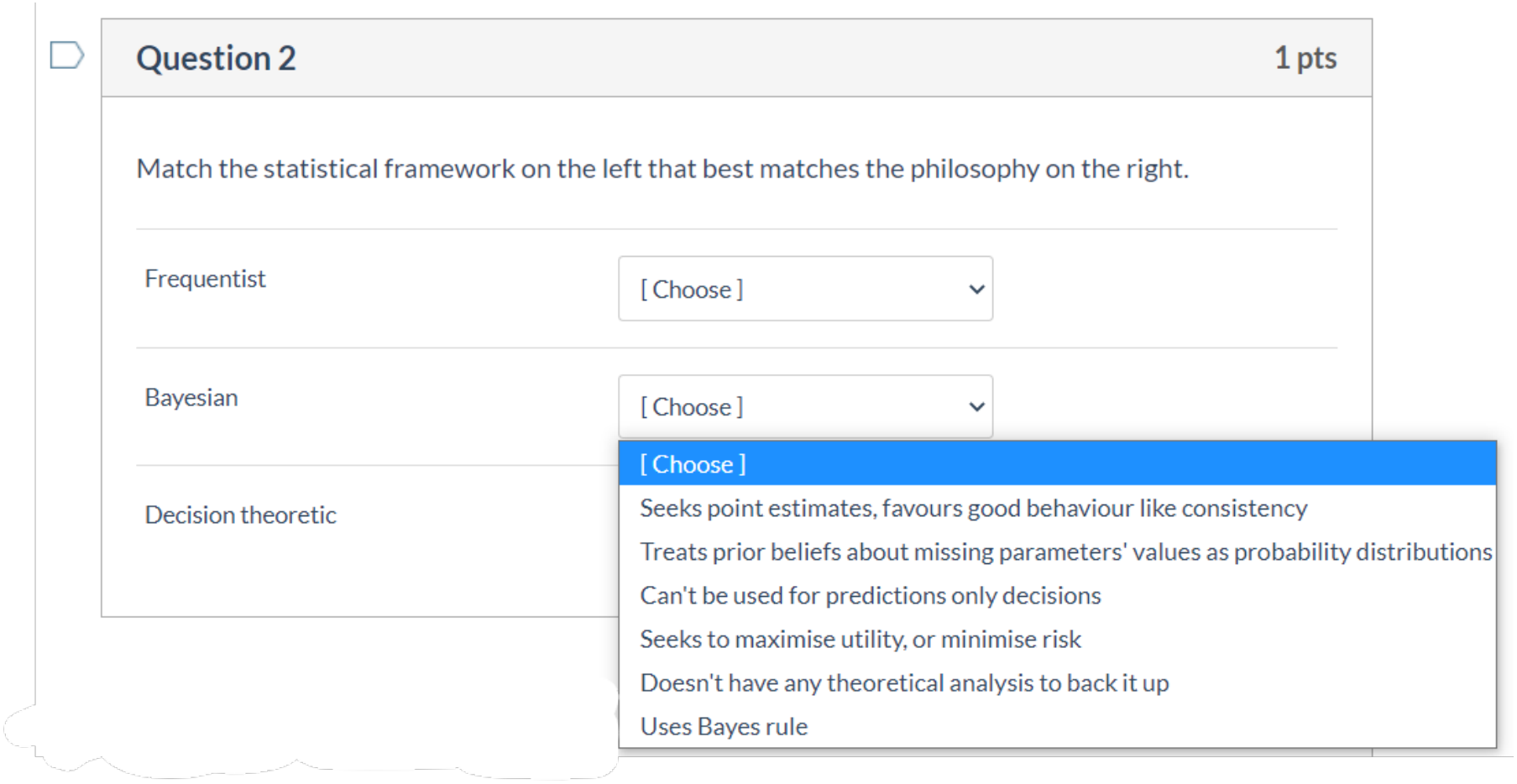
- ○ The more likelihoods, the higher the product.
- ○ So that we can weigh each data point separately.
- ○ The data points are assumed independent.

Question 2 — 1 pts

Match the statistical framework on the left that best matches the philosophy on the right.

Frequentist	[Choose]
Bayesian	
Decision theoretic	

- [Choose]
- Seeks point estimates, favours good behaviour like consistency
- Treats prior beliefs about missing parameters' values as probability distributions
- Can't be used for predictions only decisions
- Seeks to maximise utility, or minimise risk
- Doesn't have any theoretical analysis to back it up
- Uses Bayes rule


Question 2
1 pts
Match the statistical framework on the left that best matches the philosophy on the right.
Frequentist
[Choose]
Bayesian
[Choose]
Decision theoretic
[Choose]
Seeks point estimates, favours good behaviour like consistency
Treats prior beliefs about missing parameters' values as probability distributions
Can't be used for predictions only decisions
Seeks to maximise utility, or minimise risk
Doesn't have any theoretical analysis to back it up
Uses Bayes rule

Question 2 — 1 pts

Match the statistical framework on the left that best matches the philosophy on the right.

Frequentist	[Choose]
Bayesian	[Choose]
Decision theoretic	[Choose]

- [Choose]
- Seeks point estimates, favours good behaviour like consistency
- Treats prior beliefs about missing parameters' values as probability distributions
- Can't be used for predictions only decisions
- Seeks to maximise utility, or minimise risk
- Doesn't have any theoretical analysis to back it up
- Uses Bayes rule

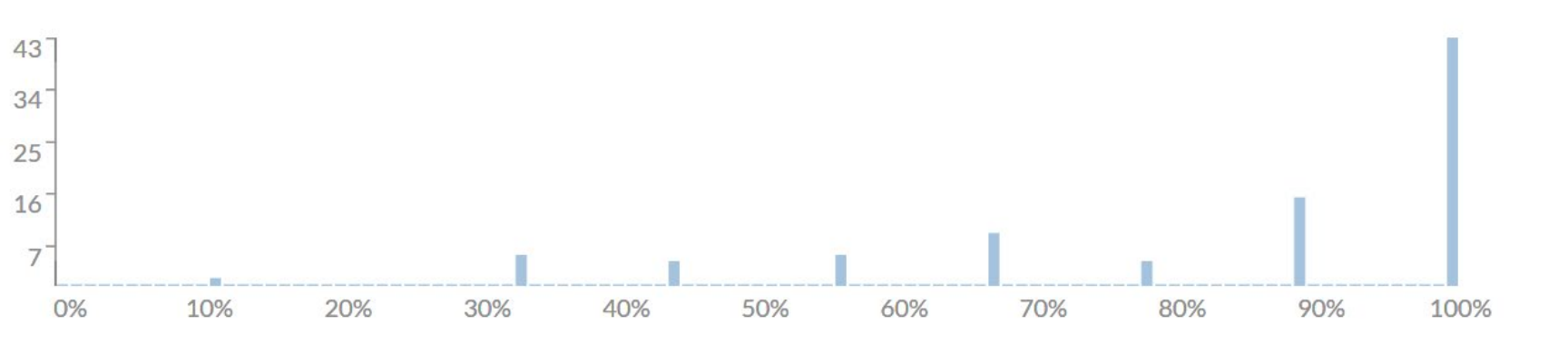
Question 3 — 1 pts

Different statistical schools treat missing parameters θ differently. This treatment is reflected in the standard notation of the statistical schools. Match the school on the left with the conventional notation for the likelihood on the right.

Frequentist	[Choose]
Bayesian	

[Choose]
p_theta(x)
p(X | theta)

Quiz Stats



86 attempted the quiz with overall 83% average

Correct by question: 1 (78%), 2 (81% but Bayes 69%), 3 (95%)