

MAS2317/3317

Assignment 1 feedback

March 11, 2014

Question 1

- Marking was generally relaxed on this question - you would get marks just for being *reasonable* and justifying your decision.

Question 1

- Marking was generally relaxed on this question - you would get marks just for being *reasonable* and justifying your decision.
- The following were common problems encountered in this question.

Question 1

Part (ii)

Probability that Newcastle United will finish in the top 6?

Question 1

Part (ii)

Probability that Newcastle United will finish in the top 6?

Using inappropriate interpretations

There are 20 teams in the Premier League and 6 positions in the top 6. Using the *classical interpretation*, the probability for finishing in the top 6 is $\frac{6}{20}$.

Question 1

Part (ii)

Probability that Newcastle United will finish in the top 6?

Using inappropriate interpretations

There are 20 teams in the Premier League and 6 positions in the top 6. Using the *classical interpretation*, the probability for finishing in the top 6 is $\frac{6}{20}$.

- This would only work if we assumed all teams are equal in ability and therefore equally likely to finish in any position. But some teams are better than others, which invalidates our assumption.

Question 1

Part (ii)

Probability that Newcastle United will finish in the top 6?

Using inappropriate interpretations

There are 20 teams in the Premier League and 6 positions in the top 6. Using the *classical interpretation*, the probability for finishing in the top 6 is $\frac{6}{20}$.

- This would only work if we assumed all teams are equal in ability and therefore equally likely to finish in any position. But some teams are better than others, which invalidates our assumption.
- There are many situations where a classical interpretation is not appropriate, because the events do not occur with equal probability.

Question 1

Part (iii)

Probability that randomly selected number between 10 and 99 (inclusive) is divisible by 7?

Question 1

Part (iii)

Probability that randomly selected number between 10 and 99 (inclusive) is divisible by 7?

Correct answer:

13 numbers divisible by 7 and 90 numbers means
 $\Pr(\text{divisible by } 7) = \frac{13}{90}$ using *Classical* interpretation.

Question 1

Part (iii)

Probability that randomly selected number between 10 and 99 (inclusive) is divisible by 7?

Correct answer:

13 numbers divisible by 7 and 90 numbers means
 $\Pr(\text{divisible by } 7) = \frac{13}{90}$ using *Classical* interpretation.

Common mistakes:

- $\Pr(\text{divisible by } 7) = \frac{13}{89}$: You did $99 - 10 = 89$, but that means you haven't included 10 or 99 as a possible value. (No penalty applied for using 89 instead of 90)

Question 1

Part (iii)

Probability that randomly selected number between 10 and 99 (inclusive) is divisible by 7?

Correct answer:

13 numbers divisible by 7 and 90 numbers means
 $\Pr(\text{divisible by } 7) = \frac{13}{90}$ using *Classical* interpretation.

Common mistakes:

- $\Pr(\text{divisible by } 7) = \frac{13}{89}$: You did $99 - 10 = 89$, but that means you haven't included 10 or 99 as a possible value. (No penalty applied for using 89 instead of 90)
- $\Pr(\text{divisible by } 7) = \frac{12}{90}$: You didn't count all numbers divisible by 7 or got any other answer. (-1 mark)

Question 1

Part (b)

Repeat parts (i) to (iii), but this time provide a range that you believe contains the value with probability 0.95.

Question 1

Part (b)

Repeat parts (i) to (iii), but this time provide a range that you believe contains the value with probability 0.95.

Common mistakes:

- Missing out the question! (-6 marks)

Question 1

Part (b)

Repeat parts (i) to (iii), but this time provide a range that you believe contains the value with probability 0.95.

Common mistakes:

- Missing out the question! (-6 marks)
- Answering with $(0, 1)$: This will contain the “true” probability value with probability 1, not 0.95!

Question 5

Question 5

For the missing Air France flight 447, construct rules for differential diagnosis of the most plausible search area for the flight's data recorder, on the basis of the posterior assessments for each type of identifier.

The correct method:

Use Bayes' Theorem to construct posterior probabilities,

$$Pr(N | x_i) = \frac{Pr(x_i | N)Pr(N)}{Pr(x_i | N)Pr(N) + Pr(x_i | S)Pr(S)}, i = 1, \dots, 4.$$

Then use the *posterior* probabilities to construct your *differential diagnosis rules*.

Question 5

The correct method:

$$Pr(N | x_i) = \frac{Pr(x_i | N)Pr(N)}{Pr(x_i | N)Pr(N) + Pr(x_i | S)Pr(S)}, i = 1, \dots, 4.$$

Then use the *posterior* probabilities to construct your *differential diagnosis rules*.

Common mistakes:

- Attempting to make differential diagnosis rules using the prior probabilities.

Question 5

The correct method:

$$Pr(N | x_i) = \frac{Pr(x_i | N)Pr(N)}{Pr(x_i | N)Pr(N) + Pr(x_i | S)Pr(S)}, i = 1, \dots, 4.$$

Then use the *posterior* probabilities to construct your *differential diagnosis rules*.

Common mistakes:

- Attempting to make differential diagnosis rules using the prior probabilities.
- Calculating posterior probabilities to 2 decimal places.

Give probabilities as a sensible fraction or a decimal to 4 decimal places to reduce rounding error.

Question 9

Question 9

A sample $x_1, x_2, \dots, x_n$ is taken from the probability density function,

$$f(x | \sigma) = \frac{x}{\sigma^2} \exp \left\{ -\frac{x^2}{2\sigma^2} \right\}, \quad x > 0, \quad \sigma > 0$$

Find the likelihood function for σ , and a sufficient statistic for σ .

Question 9

Question 9

A sample $x_1, x_2, \dots, x_n$ is taken from the probability density function,

$$f(x | \sigma) = \frac{x}{\sigma^2} \exp \left\{ -\frac{x^2}{2\sigma^2} \right\}, \quad x > 0, \quad \sigma > 0$$

Find the likelihood function for σ , and a sufficient statistic for σ .

Need to correctly identify $h(\mathbf{x})$ and $g(\sum_{i=1}^n x_i^2, \sigma)$:

$$\begin{aligned} f(\mathbf{x} | \sigma) &= \prod_{i=1}^n \frac{x_i}{\sigma^2} \exp \left\{ -\frac{x_i^2}{2\sigma^2} \right\} \\ &= \underbrace{\prod_{i=1}^n x_i}_{h(\mathbf{x})} \times \underbrace{\frac{1}{\sigma^{2n}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2 \right\}}_{g(\sum_{i=1}^n x_i^2, \sigma)} \end{aligned}$$

Question 9

A wide range of different mistakes were made with manipulating the likelihood function:

$$\begin{aligned}f(\mathbf{x} \mid \sigma) &= \prod_{i=1}^n \left[\frac{x_i}{\sigma^2} \exp \left\{ -\frac{x_i^2}{2\sigma^2} \right\} \right] \\&= \prod_{i=1}^n x_i \times \frac{1}{\sigma^{2n}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2 \right\} \\&\neq n\bar{x} \times \frac{1}{\sigma^2} \exp \left\{ -\frac{1}{2n\sigma^2} \sum_{i=1}^n x_i^2 \right\}\end{aligned}$$

Question 9

A wide range of different mistakes were made with manipulating the likelihood function:

$$\begin{aligned}f(\mathbf{x} \mid \sigma) &= \prod_{i=1}^n \left[\frac{x_i}{\sigma^2} \exp \left\{ -\frac{x_i^2}{2\sigma^2} \right\} \right] \\&= \prod_{i=1}^n x_i \times \frac{1}{\sigma^{2n}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2 \right\} \\&\neq n\bar{x} \times \frac{1}{\sigma^2} \exp \left\{ -\frac{1}{2n\sigma^2} \sum_{i=1}^n x_i^2 \right\}\end{aligned}$$

In particular

$$\prod_{i=1}^n x_i \neq x^n \neq n\bar{x}$$

Question 9

A wide range of different mistakes were made with manipulating the likelihood function:

$$\begin{aligned} f(\mathbf{x} \mid \sigma) &= \prod_{i=1}^n \left[\frac{x_i}{\sigma^2} \exp \left\{ -\frac{x_i^2}{2\sigma^2} \right\} \right] \\ &= \prod_{i=1}^n x_i \times \frac{1}{\sigma^{2n}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2 \right\} \\ &\neq n\bar{x} \times \frac{1}{\sigma^2} \exp \left\{ -\frac{1}{2n\sigma^2} \sum_{i=1}^n x_i^2 \right\} \end{aligned}$$

In particular

$$\prod_{i=1}^n x_i \neq x^n \neq n\bar{x} \quad \text{and} \quad \sum_{i=1}^n x_i^2 \neq \left(\sum_{i=1}^n x_i \right)^2 \neq n\bar{x}^2$$

And finally ...

Question 1(a), when answered with the subjective interpretation of probability, produced some entertaining – and possibly offensive – explanations.

Question 1(a)

Approximate the probability that a student taking MAS2317/3317 has a body mass index (BMI) of at least 22.5.

And finally ...

Question 1(a), when answered with the subjective interpretation of probability, produced some entertaining – and possibly offensive – explanations.

Question 1(a)

Approximate the probability that a student taking MAS2317/3317 has a body mass index (BMI) of at least 22.5.

Here are a few of our favourite subjective answers ...

Our favourite answers

One student was very honest.

$p \approx 0.7$ – “*Maths students are naturally lazy and therefore more likly [sic] to be overweight.*”

Our favourite answers

One student was very honest.

$p \approx 0.7$ – “*Maths students are naturally lazy and therefore more likly [sic] to be overweight.*”

Another student who shares similar views declared,

$p \approx 0.6$ – “*Maths students are stereotypically lazy, resulting in higher BMI.*”

Our favourite answers

One student was very honest.

$p \approx 0.7$ – *“Maths students are naturally lazy and therefore more likly [sic] to be overweight.”*

Another student who shares similar views declared,

$p \approx 0.6$ – *“Maths students are stereotypically lazy, resulting in higher BMI.”*

One student is very blunt with why they guessed a high probability,

$p = 0.9$ – *“Subjective interpretation, from looking around the room.”*

Our favourite answers

Perhaps we're better than the 'general' population?

$p = 0.9$ – *“We can assume that, given we are testing the BMI of University students, we are not dealing with people from an impoverished background, so people will not generally be underweight.”*

¹Stamatakis, E., Wardle, J., and Cole, T. J. (2010) Childhood obesity and overweight prevalence trends in England: evidence for growing socioeconomic disparities. *International Journal of Obesity* **34**, 41-47

Our favourite answers

Perhaps we're better than the 'general' population?

$p = 0.9$ – “We can assume that, given we are testing the BMI of University students, we are not dealing with people from an impoverished background, so people will not generally be underweight. As this is a course that requires no physical activity, can also also say there will be plenty of people who are overweight too.”

¹Stamatakis, E., Wardle, J., and Cole, T. J. (2010) Childhood obesity and overweight prevalence trends in England: evidence for growing socioeconomic disparities. *International Journal of Obesity* **34**, 41-47

Our favourite answers

Perhaps we're better than the 'general' population?

$p = 0.9$ – “We can assume that, given we are testing the BMI of University students, we are not dealing with people from an impoverished background, so people will not generally be underweight. As this is a course that requires no physical activity, can also also say there will be plenty of people who are overweight too.”

Actually, findings demonstrate the opposite – children from low-income or poor socioeconomic backgrounds are *more likely* to be overweight or obese than those from higher income backgrounds¹.

¹Stamatakis, E., Wardle, J., and Cole, T. J. (2010) Childhood obesity and overweight prevalence trends in England: evidence for growing socioeconomic disparities. *International Journal of Obesity* **34**, 41-47

Our favourite answers

And the marker's favourite answer:

$$p = 0.7 -$$

Our favourite answers

And the marker's favourite answer:

$p = 0.7$ – *“This is because many of my associates live off a diet consisting mainly of chicken and chorizo sandwiches from Greggs, meatball subs from Subway and chicken, bacon and stuffing sandwiches from Tesco*



Our favourite answers

And the marker's favourite answer:

$p = 0.7$ – “This is because many of my associates live off a diet consisting mainly of chicken and chorizo sandwiches from Greggs, meatball subs from Subway and chicken, bacon and stuffing sandwiches from Tesco, washed down with copius [sic] amounts of Monster energy drink and alcohol. Such a diet leads to weight gain and therefore a higher BMI.”



Our favourite answers

And the marker's favourite answer:

$p = 0.7$ – “This is because many of my associates live off a diet consisting mainly of chicken and chorizo sandwiches from Greggs, meatball subs from Subway and chicken, bacon and stuffing sandwiches from Tesco, washed down with copius [sic] amounts of Monster energy drink and alcohol. Such a diet leads to weight gain and therefore a higher BMI.”

If this sounds like what you eat, your “associate” might be hinting you should reassess your diet based on that last sentence.

