

Round 1

Basic requirements

Q1

Can you open the pdf and it's not blank?

No

Yes

[+ Additional Comment](#)

Q2

Is the report anonymous?

No

Yes

[+ Additional Comment](#)

Exercise 1

Q3

The correct answers for a) are roughly the following:

- probability: a **measure between 0 and 1** of a likelihood that an event will occur.
- probability mass: a **value between 0 and 1** that tells what is a probability of some discrete event.
- probability density: a **function that outputs relative probabilities for given regions in the domain of a random variable**
- probability mass function: a **function that outputs a probability** of some discrete event.
- probability density function: same as probability density
- probability distribution: a function that gives probabilities of outcomes of different random variables.
- discrete probability distribution: a **probability distribution of discrete values**.
- continuous probability distribution: a **probability distribution of values that come from a continuous space**
- cumulative distribution function: a **function that gives the probability that a certain random value will be less or equal than the real valued input**
- likelihood: $p(y|\theta)$ as a function of θ , epistemic information obtained from the data through the likelihood, proportional to the probability **distribution** (not just the

probability, because it can be also density) of theta given uniform prior. Sometimes used also to refer to the observation model/statistical model/sampling distribution.

How is the answer?

Totally wrong/has not tried

Something sensible written

All/almost all are correct ($\geq 70\%$ correct)

[+ Additional Comment](#)

Exercise 2

Q4

Is the source code for the solutions included?

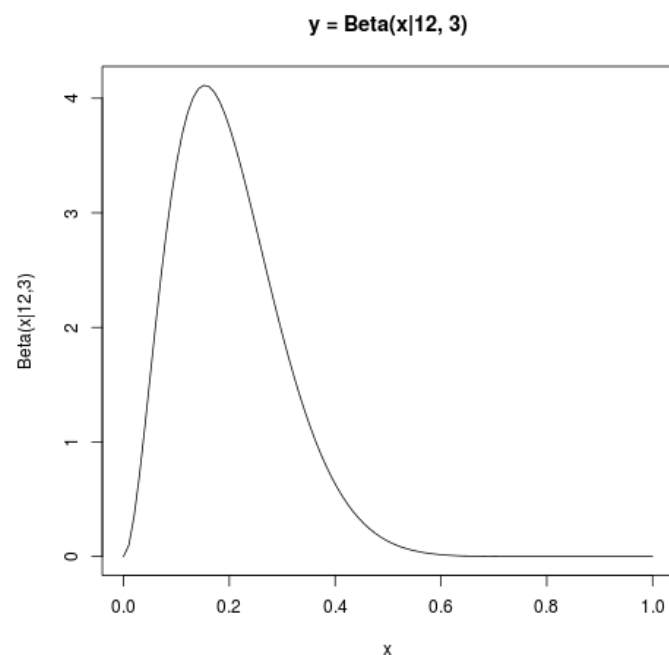
No

Yes

[+ Additional Comment](#)

Q5

Does the plot in the report at 2a) look something like this:



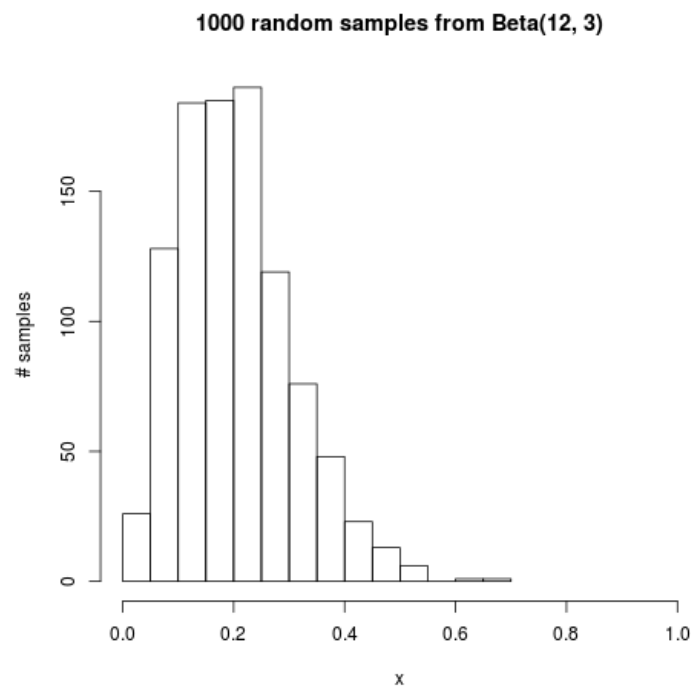
No

Yes

[+ Additional Comment](#)

Q6

Does the plot in the report at 2b) look something like this:



No

Yes

[+ Additional Comment](#)

Q7

Is the computed mean in 2c) close to 0.2?

No

Yes

[+ Additional Comment](#)

Q8

Is the variance in 2c) close to 0.01?

No

Yes

[+ Additional Comment](#)

Q9

Is the probability interval in 2d) roughly [0.05, 0.43]? Remember that since the interval is computed from random samples, there is variation and the answers should be **roughly** the same!

No

Yes

[+ Additional Comment](#)

Exercise 3

Q10

Is $p(\text{has cancer}|\text{test result is positive})$ computed using Bayes' formula (or its complement $p(\text{does not have cancer}|\text{test result is positive})$)?

No

Yes

[+ Additional Comment](#)

Q11

Is the result $p(\text{has cancer}|\text{test result is positive})=49/2047\approx 0.0239$ (or $p(\text{does not have cancer}|\text{test result is positive})=1998/2047\approx 0.9761$)

No

Yes

[+ Additional Comment](#)

Q12

Is the result motivated with something like "The test should not be used as even if the test result is positive, it is very unlikely that the person has a cancer" (or "The test should not be used as even if the test gives positive result, it is very likely that the patient does not have a cancer")

No

Yes

[+ Additional Comment](#)

Exercise 4

Q13

Is the source code available?

No

Yes

[+ Additional Comment](#)

Q14

How is the answer for probability of picking a red ball?

No answer

Probability rules $p(a,b)=p(a|b)p(b)$ and $p(a)=p(a|b)+p(a|b')$ used, but the result is not $447/1400\approx 0.319$

Probability rules $p(a,b)=p(a|b)p(b)$ and $p(a)=p(a|b)+p(a|b')$ used, and the result is $447/1400 \approx 0.319$

[+ Additional Comment](#)

Q15

How is the answer for what box is most probable? **Note:** in below it is also totally OK to only have computed for all boxes the non-normalized probability $p(\text{red}|\text{box})p(\text{box})$, since the denominator is same for all three probabilities.

No answer

Bayes rule used to compute probabilities for all boxes given that the picked ball is red, but the answers are not $p(A|R) \approx 0.358$, $p(B|R) \approx 0.251$ and $p(C|R) \approx 0.391$. (or non scaled results for A: 0.114, B:0.08, C:0.125))

Bayes rule used to compute probabilities for all boxes given that the picked ball is red, the answers are $p(A|R) \approx 0.358$, $p(B|R) \approx 0.251$ and $p(C|R) \approx 0.391$ and it is not explicitly said that the most probable box is box C. (or non scaled results for A: 0.114, B:0.08, C:0.125, and it is not said that this is biggest for C)

Bayes rule used to compute probabilities for all boxes given that the picked ball is red, the answers are $p(A|R) \approx 0.358$, $p(B|R) \approx 0.251$ and $p(C|R) \approx 0.391$ and it is explicitly said that the most probable box is box C. (or non scaled results for A: 0.114, B:0.08, C:0.125, and it is said that this is biggest for C)

[+ Additional Comment](#)

Exercise 5

Q16

How is the answer for probability of Elvis having had an identical twin brother?

No answer

Probability that Elvis had an identical twin brother is computed using Bayes rule, but the result is not roughly 0.43

Probability that Elvis had an identical twin brother is computed using Bayes rule, and the result is roughly 0.43

[+ Additional Comment](#)

Overall quality of the report

Q17

Does the report follow the formatting instructions?

Not at all

Little

Mostly

Yes

[+ Additional Comment](#)

Q18 · OPTIONAL

In case the report doesn't fully follow the formatting instructions, specify the formatting instruction that is not followed. If applicable, specify the page of the report, where this difference in formatting is visible.

Round 2

Basic requirements

Q1

Can you open the pdf and it's not blank?

No

Yes

[+ Additional Comment](#)

Q2

Is the report anonymous?

No

Yes

[+ Additional Comment](#)

Exercise 1

Q3

Is source code included?

No

Yes

[+ Additional Comment](#)

Q4

Are the prior, likelihood and posterior forms in a) reported (derivation of posterior not necessary)?

No

Some missing

Yes

[+ Additional Comment](#)

Q5

Is the reported resulting posterior correct Beta(46, 240)?

It is not reported, that the posterior distribution is a Beta distribution.

It is reported, that the posterior distribution is Beta(α , β), but the numerical values for the parameters are incorrect

It is reported, that the posterior distribution is $\text{Beta}(\alpha, \beta)$, and the numerical values for the parameters are correct.

[+ Additional Comment](#)

Q6

In part b), is there at least one point estimate reported. Sample based estimates are also ok. Points should be given if the method is right, even if the result is wrong due to a wrong posterior distribution being used. With the right posterior, mean, median, and mode are all approximately 0.16.

No

Yes

[+ Additional Comment](#)

Q7

For the b) part, is the 90% posterior interval reported? Sample based estimate is also ok. Points should be given if the method is right, even if the result is wrong because the posterior was wrong in the first place. If the posterior was right, the 90% posterior interval is roughly [0.13,0.20].

No

Yes

[+ Additional Comment](#)

Q8

For the c) part, is the posterior probability $\Pr(\pi < 0.2 | y)$ reported? Points should be given if the method is right, even if the result is wrong because the posterior was wrong. If the posterior was right, the result should be approximately 0.959.

No

Yes

[+ Additional Comment](#)

Q9

For the d) part, does the report discuss somehow the assumption of independence given the shared success probability π (i.i.d. or independent and exchangeable trials is also ok)?

No

No, but other reasonable assumptions are discussed

Yes, but not quite right or some missing

Yes

[+ Additional Comment](#)

Q10

For the e) part, is there some comparison and discussion of results obtained with alternative prior parameters?

No

Yes, but the results and conclusions are clearly wrong

Yes

[+ Additional Comment](#)

Overall quality of the report

Q11

Does the report follow the formatting instructions?

Not at all

Little

Mostly

Yes

[+ Additional Comment](#)

Q12 · OPTIONAL

In case the report doesn't fully follow the formatting instructions, specify the formatting instruction that is not followed. If applicable, specify the page of the report, where this difference in formatting is visible.

Round 3

Basic requirements

Q1

Can you open the pdf and it's not blank?

No

Yes

[+ Additional Comment](#)

Q2

Is the report anonymous?

No

Yes

[+ Additional Comment](#)

Exercise 1

Q3

Is the source code included?

No

Yes

[+ Additional Comment](#)

Q4

Are the likelihood, prior and the posterior for computing the average hardness value reported? It is ok to refer to the book instead of deriving the distributions.

No

Yes, but some are missing

Yes

[+ Additional Comment](#)

Q5

In part a), were the point estimates and posterior interval provided? (The posterior mean should be close to 14.6 and 95% posterior interval should be around [13.5,15.7])

No

Yes, but seem incorrect or only one estimate was reported

Yes, and the reported values seem plausible

[+ Additional Comment](#)

Q6

In part a), was the density plotted?

No

Yes, but seem incorrect

Yes, and the plot seems plausible

[+ Additional Comment](#)

Q7

For b)-part, was a formula or a simulation method presented for computing the posterior predictive distribution? It is ok to refer to the book.

No

Yes, but seems incorrect

Yes

[+ Additional Comment](#)

Q8

For b)-part, were the point estimate and predictive interval provided? (95% predictive interval should be around [11.0,18.2] and the mean the same as in a)-part).

No

Yes, but seems incorrect

Yes, and the reported values seem plausible

[+ Additional Comment](#)

Q9

For b)-part, was the density plotted?

No

Yes, but seems incorrect

Yes, and the plot seem plausible

[+ Additional Comment](#)

Exercise 2

Q10

Is the source code included?

No

Yes

[+ Additional Comment](#)

Q11

Are the likelihood, prior and the posterior for the death probabilities reported? It is ok to refer to the book instead of deriving the distributions.

No

Yes, but some are missing

Yes

[+ Additional Comment](#)

Q12

In part a), was the simulation algorithm for computing the posterior of the odds ratio presented or implemented?

No

Yes, but seems incorrect

Yes

[+ Additional Comment](#)

Q13

In a)-part, was the odds ratio summarized with a point estimate and a posterior interval? (The mean should be close to 0.57 and 95% posterior interval approximately [0.32, 0.94])

No

Yes, but results seem incorrect

Yes, and the results seem plausible

[+ Additional Comment](#)

Q14

In part b), was some discussion about testing alternative priors provided? (For example, one could have repeated the computations in a)-part with a couple of alternative priors and reported these results or some related general conclusions briefly)

Not at all

Some analysis was provided but it was lacking or did not make sense

Some alternative priors were tested and some sensible discussion provided

[+ Additional Comment](#)

Exercise 3

Q15

Is source code included?

No

Yes

[+ Additional Comment](#)

Q16

Are the likelihood, prior and the posterior for the windshield hardness values reported?
(It is also ok to refer to the book or related formulas from exercise 1)

No

Yes, but some are missing

Yes

[+ Additional Comment](#)

Q17

In part a), was the simulation algorithm for computing the difference in the means presented or implemented?

No

Yes, but seems to be incorrect

Yes

[+ Additional Comment](#)

Q18

In part a), was the posterior for the difference between the means summarized with point and interval estimates? (The mean should be close to -1.2 and 95% posterior interval [-2.4, 0.04] or something close to it)

No answer

Yes, but results seem incorrect or only one estimate was given

Yes, and results seem reasonable

[+ Additional Comment](#)

Q19

Were some analysis or discussion provided for assessing whether the means could be the same? Did the discussion include recognition that the probability is 0 for a continuous distribution?

No analysis or explanation is given

Yes, but the analysis or explanation seems incorrect

Yes, and the analysis or explanation seems plausible

[+ Additional Comment](#)

Overall quality of the report

Q20

Does the report follow the formatting instructions?

Not at all

Little

Mostly

Yes

ADDITIONAL COMMENT REQUIRED

Q21 · OPTIONAL

In case the report doesn't fully follow the formatting instructions, specify the formatting instruction that is not followed. If applicable, specify the page of the report, where this difference in formatting is visible.

Q22

Please provide also feedback on the presentation (e.g. text, layout, flow of the responses, figures, figure captions). Part of the course is practicing making data analysis reports. By providing feedback on the report presentation and other students can learn what they can improve or what they already did well. You should be able to provide constructive or positive feedback for all non-empty reports, even if there is nothing to say about the technical correctness of the answers. You can also provide feedback on code.

10 of 10 words (minimum)

Round 4

Basic requirements

Q1

Can you open the pdf and it's not blank?

No

Yes

[+ Additional Comment](#)

Q2

Is the report anonymous?

No

Yes

[+ Additional Comment](#)

Content

Q3

Are the mean and covariance of the prior in a) reported? The correct answers are:

mean $\begin{pmatrix} 0 \\ 10 \end{pmatrix}$

covariance $\begin{pmatrix} 4 & 12 \\ 12 & 100 \end{pmatrix}$

Not reported

Yes, but they are not correct

Yes, and they are correct

[+ Additional Comment](#)

Q4

Are the means and their MCSEs of alpha and beta in b) reported? Note that the number of digits reported for the means must follow the rule given in the assignment. The correct answers are:

alpha:

mean 1.0 (MCSE 0.01482435)

beta:

mean 10.6 (MCSE 0.07560016)

Not reported

Yes, but one or both means are incorrect

Yes, and the means are correct

[+ Additional Comment](#)

Q5

Are the quantiles and their MCSEs of alpha and beta in b) reported? Note that the number of digits reported for the quantiles must follow the rule given in the assignment. The correct answers are

alpha:

5% quantile -0.5 (MCSE 0.02600412)

95% quantile 2.6 (MCSE 0.04206342)

beta:

5% quantile 4.0 or 4 (MCSE 0.07043125) (here 4 is also accepted, because it can be sensible to use same reporting accuracy for both quantiles)

95% quantile 19 (MCSE 0.2412129)

Not reported

Yes, but one or more quantiles are incorrect

Yes, and the quantiles are correct

[+ Additional Comment](#)

Q6

Is the source code for the function in c) reported?

No

Yes

[+ Additional Comment](#)

Q7

Is the source code for the function in d) reported?

No

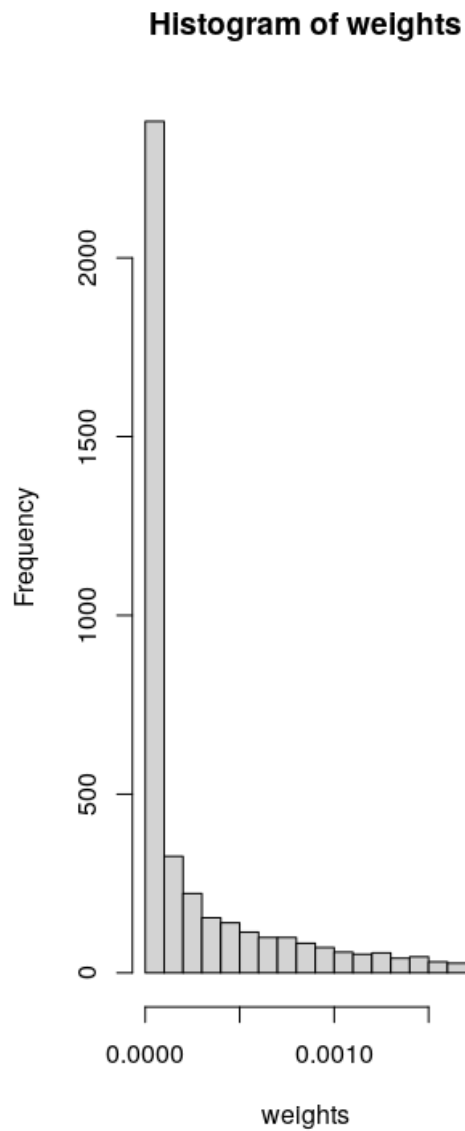
Yes

[+ Additional Comment](#)

Q8

Does the histogram in e) look something like this figure?

If it is evident that the normalized importance ratios are computed correctly, but the prior was incorrect, you can still grade "Reported and looks similar".



Not reported

Reported, but looks different

Reported and looks similar

[+ Additional Comment](#)

Q9

Is the effective sample size in f) reported?

The correct range for the effective sample size is between 1000 and 1300.

However, if it is evident that the effective sample size is computed correctly, but the prior was incorrect, you can still grade "Yes, and it is correct".

No

Yes, but it is not correct

Yes, and it is correct

[+ Additional Comment](#)

Q10

The correct explanation for g) is roughly the following: The importance sampling effective sample size is a rough measure of how many of the draws have an effect for the importance sampling task. In other words, it measures the accuracy of importance sampling.

What is the connection between S_{eff} and the histogram of weights:

A large number of small weights leads to a small effective sample size. In the histogram of the weights, this is seen as a big spike close to zero.

It is also okay to discuss the variance of the distribution of weights. It is also okay to mention that in this assignment, there are no extremely large weights, which indicates that importance sampling can be trusted.

How is the answer?

Totally wrong/has not tried

Something is a bit wrong

Explanation is sensible

[+ Additional Comment](#)

Q11

Is the source code for the function in h) reported?

No

Yes

[+ Additional Comment](#)

Q12

Are the means and their MCSEs of alpha and beta in h) reported? Note that the number of digits reported for the means must follow the rule given in the assignment. The correct answers should be close to these:

alpha:

mean 0.9 or 1.0 (MCSE ~ 0.027)

beta:

mean 10 or 11 (MCSE ~ 0.14)

Not reported

Yes, but they are incorrect

Yes, and they are correct

[+ Additional Comment](#)

Overall quality of the report

Q13

Does the report follow the formatting instructions?

Not at all

Little

Mostly

Yes

[+ Additional Comment](#)

Q14

In case the report doesn't fully follow the formatting instructions, specify the formatting instruction that is not followed. If applicable, specify the page of the report, where this difference in formatting is visible.

0 of 2 words (minimum)

Q15

Please provide also feedback on the presentation (e.g. text, layout, flow of the responses, figures, figure captions). Part of the course is practicing making data analysis reports. By providing feedback on the report presentation and other students can learn what they can improve or what they already did well. You should be able to provide constructive or positive feedback for all non-empty reports, even if there is nothing to say about the technical correctness of the answers. You can also provide feedback on code.

11 of 10 words (minimum)

Round 5

Basic requirements

Q1

Can you open the pdf and it's not blank?

No

Yes

Q2

Is the report anonymous?

No

Yes

Metropolis

Q3

Is the implementation of **density_ratio** function included ?

No

Yes

Q4

Is the implementation of **metropolis_bioassay** function included ?

No

Yes

Q5

2 a) Is the brief description of Metropolis-Hastings algorithm included (and it's not complete nonsense)? Provide also a brief comment on how clear you think that description is (and potentially mention errors if you see them).

No

Yes

Q6

2 b) Is the proposal/jumping distribution reported ?

No

Yes

Q7

2 c) Are the starting points or the mechanism how they are generated reported?

No

Yes

Q8

2 d) Is the number of draws per chain reported?

No

Yes

Q9

2 e) Is the warm-up length reported?

No

Yes

Q10

2 f) Is the number of chains reported?

No

Yes

Q11

2 g) and 2 h) Are line plots of the chains included? (Separate plots for alpha and beta)

No plots are included

Yes, but both plots are in a single figure, or the plots are scatter plots (scatter plots aren't useful for visual convergence evaluation).

Yes, but only a plot for alpha or beta is included.

Yes, separate line plots for both alpha and beta are included.

Q12

Is there a discussion on the convergence of the chains?

No discussion on convergence.

Yes, but the discussion is not convincing.

Yes, discussed in the report.

Q13

Is it mentioned which implementation of Rhat is used ?

Two possible ways to compute R-hat would be:

1. Potential scale reduction factor, Eq. (11.4) in the BDA3.
2. Use **Rhat** function provided in *rstan* or *posterior* which will compute the rank normalized R-hat (as suggested in the exercise PDF and discussed further in Vehtari et.al 2019, arXiv:1903.08008)

It is ok as long as it is mentioned (or evident from the code) which of the above is used.

No

Yes

Q14

Is the brief description of Rhat included (and it's not complete nonsense)? Provide also a brief comment on how clear you think that description is (and potentially mention errors if you see them).

No

Yes

Q15

Are the Rhat-values for alpha and beta reported?

No

Yes, but incorrectly computed

Yes, but computed separately for each chain

Yes, but only for alpha or beta

Yes, single values both for alpha and beta

Q16

Is the interpretation of R-hat values correct (R-hat values close to one => chains probably converged and estimates reliable)?

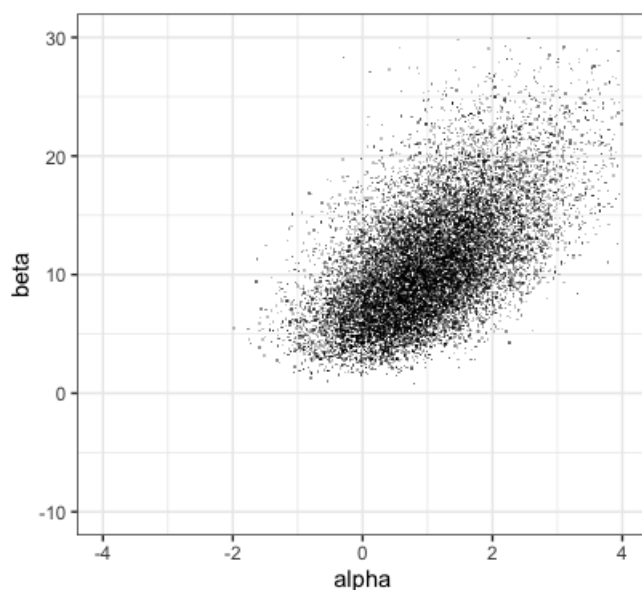
No interpretation or discussion about the R-hat values, or conclusions clearly wrong

Interpretation somewhat correct

Interpretation correct

Q17

Does the report contain a scatter plot about the draws? Do the results look reasonable, that is, roughly like in the Figure below?



No plot included

Plot included, but the results do not look like in the figure above

Plot included, and the results look roughly like in the figure above

Overall quality of the report

Q18

Does the report follow the formatting instructions?

Not at all

Little

Mostly

Yes

Q19 · OPTIONAL

In case the report doesn't fully follow the formatting instructions, specify the formatting instruction that is not followed. If applicable, specify the page of the report, where this difference in formatting is visible.

Q20

Please provide also feedback on the presentation (e.g. text, layout, flow of the responses, figures, figure captions). Part of the course is practicing making data analysis reports. By providing feedback on the report presentation and other students can learn what they can improve or what they already did well. You should be able to provide constructive or positive feedback for all non-empty reports, even if there is nothing to say about the technical correctness of the answers. You can also provide feedback on code.

1 of 10 words (minimum)

Round 6

Basic requirements

Q1

Can you open the pdf and it's not blank?

No

Yes

Q2

Is the report anonymous?

No

Yes

Stan

Q3

Is the Stan model code included?

No

Yes

Q4

Does the implemented Stan-model seem to be working?

No implementation

Model implemented but results not visualized/reported

Model implemented, but the results seem weird

Model seems to work correctly

Q5

Are the R_{hat} -values reported (potential scale reduction factor, Eq. (11.4) in the BDA3)?

No

Yes, but only for alpha or beta

Yes, single values both for alpha and beta

Q6

Is the interpretation of R_{hat} values correct (R_{hat} values close to one => chains probably converged and estimates reliable)?

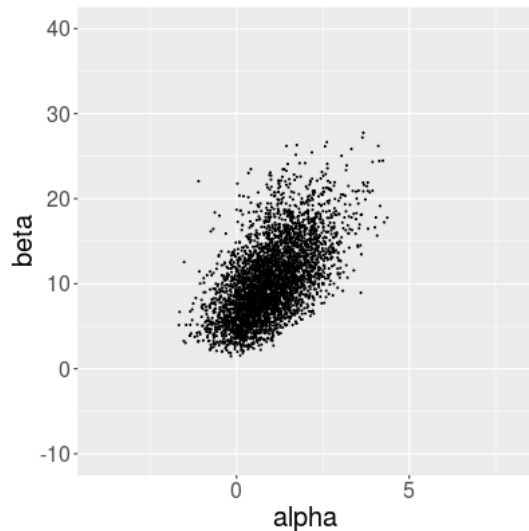
No interpretation or discussion about the R_{hat} -values, or conclusions clearly wrong

Somewhat correct

Interpretation correct

Q7

Does the report contain a scatter plot about the draws? Do the results look reasonable, that is, roughly like in the Figure below?



No plot included

Plot included, but the results do not look like in the figure above

Plot included, and the results look roughly like in the figure above

Q8

Does the report contain description of Stan setup used and whether there were any problems in setting it up or using it?

No

Yes

Q9 · OPTIONAL

Even if the Stan model code is correct, there might be ways to give improve the layout or write the model in more elegant ways. This optional feedback box can be used to give additional suggestions for better Stan code.

Overall quality of the report

Q10

Does the report follow the formatting instructions?

Not at all

Little



Mostly

Yes

Q11 · OPTIONAL

In case the report doesn't fully follow the formatting instructions, specify the formatting instruction that is not followed. If applicable, specify the page of the report, where this difference in formatting is visible.

Q12

Please provide also feedback on the presentation (e.g. text, layout, flow of the responses, figures, figure captions). Mention something that your classmate did well  and mention something that your classmate could improve at . Part of the course is practicing making data analysis reports. By providing feedback on the report presentation and other students can learn what they can improve or what they already did well. You should be able to provide constructive or positive feedback for all non-empty reports, even if there is nothing to say about the technical correctness of the answers. You can also provide feedback on code.

Round 7

Basic requirements

Q1

Can you open the pdf and it's not blank?

No

Yes

Q2

Is the report anonymous?

No

Yes

Linear model: drowning data with Stan

Q3

Is the source code included?

No

Yes

Q4

Is the full resulting modified Stan model code presented in the report?

No

Yes

Q5

Fix #1: Is there a fix for line 10?

original: `real<upper=0> sigma;`

possible fix: `real<lower=0> sigma;`

It has not been discussed that this line should be fixed

It has been discussed that this line should be fixed, but there is no fix presented for it or the fix is clearly wrong

There is a fix presented for this line that clearly solves the problem

Q6

Fix #2: Is there a fix for line 16?

original:

```
y ~ normal(mu, sigma)
```

possible fix (semi-colon in the end of the line):

```
y ~ normal(mu, sigma);
```

It has not been discussed that this line should be fixed

It has been discussed that this line should be fixed, but there is no fix presented for it or the fix is clearly wrong

There is a fix presented for this line that clearly solves the problem

Q7

Fix #3: Is there a fix for line 19?

original:

```
real ypred = normal_rng(mu, sigma);
```

possible fix (note xpred):

```
real ypred = normal_rng(alpha + beta*xpred, sigma);
```

another possible fix:

```
real mupred;
```

```
real ypred;
```

```
mupred = alpha + beta*xpred;
```

```
ypred = normal_rng(mupred, sigma);
```

It has not been discussed, that this line should be fixed

It has been discussed that this line should be fixed, but there is no fix presented for it or the fix is clearly wrong

There is a fix presented for this line that clearly solves the problem

Q8

Is there a suitable numerical value of approximately 26.8 presented for

σ_β (or 717.6 for σ_{β^2})?

No

Yes

Q9

Does the report discuss and correctly present how the desired prior can be implemented in the model code?

Example implementation:

- in the data block, add: `real<lower=0> sigma_beta;`
- in the model block, add: `beta ~ normal(0, sigma_beta);`

Another example implementation:

- in the model block, add: `beta ~ normal(0, 26.8);`

No

Yes

Q10

Does the report explain and also correctly present a prior for the intercept?

Example implementation:

- in the data block, add: `real<lower=0> sigma_alpha;`
- in the model block, add: `alpha ~ normal(0, sigma_alpha);`

Another example implementation:

- in the model block, add: `alpha ~ normal(0, number);`
- where $\text{number} \geq 100$

There is no explanation and no prior is presented

A prior is presented, but there is no explanation

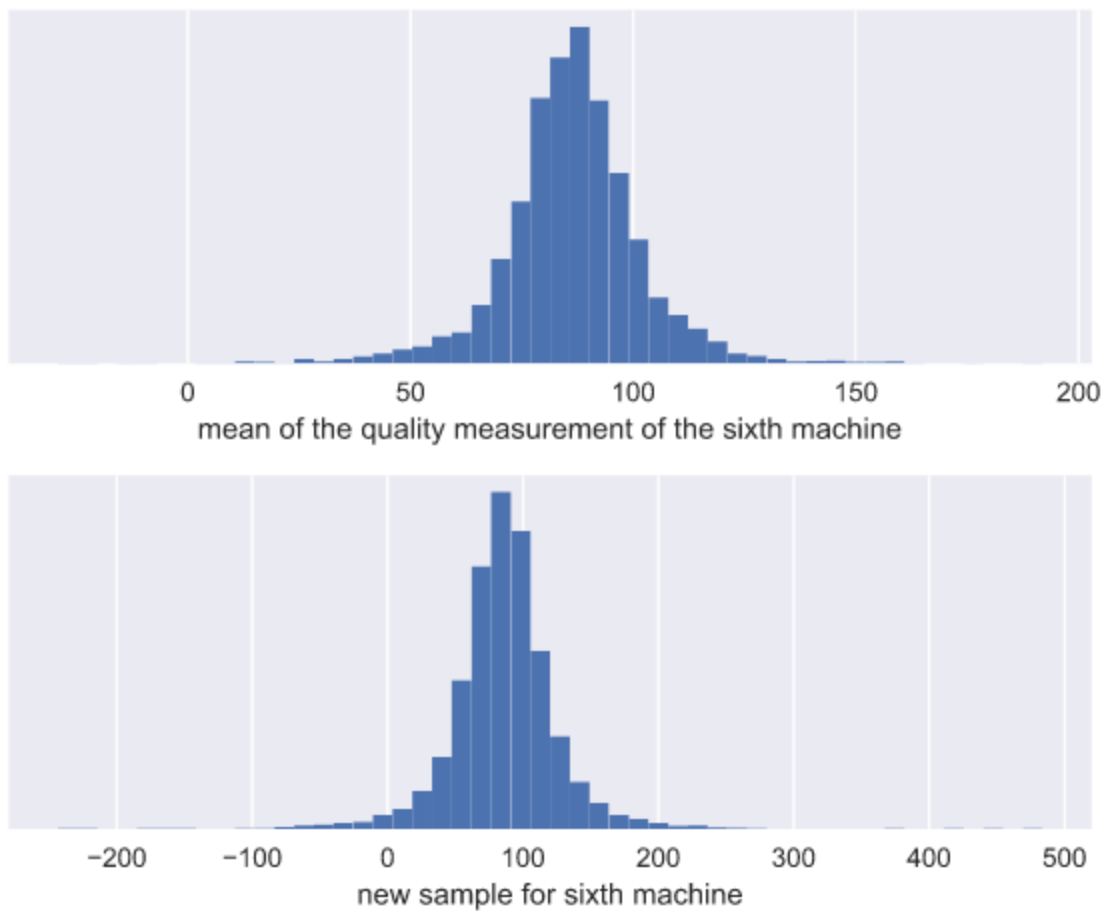
A prior is presented, but the explanation is clearly wrong

A prior is presented and the explanation sounds reasonable

Hierarchical model: factory data with Stan

Separate model

The following histograms are used as reference in the grading of the separate model:



Q11

Separate model: Is the model described using mathematical notation and the difference to other models described in words?

No equations and no description

Description but no equations

Equations but no description

Equations and description

Q12

Separate model: Is there a related Stan implementation (N.B. same implementation may be used for multiple models)?

No Stan model implemented

Stan model implemented, but it seems clearly wrong or broken

Seemingly valid Stan model implemented

Q13

Separate model: Is there a histogram plotted for the question c part i? Does it look something like the 1st subplot in the model solution plot above?

No histogram plotted

Histogram plotted but it clearly differs from the example

Histogram plotted that approximately matches the example

Q14

Separate model: Is there a histogram plotted for the question c part ii? Does it look something like the 2nd subplot in the model solution plot above?

No histogram plotted

Histogram plotted but it clearly differs from the example

Histogram plotted that approximately matches the example

Q15

Separate model: Is there some statement that question c part iii can not be answered or that the prior is the only related information?

No

Yes

Q16

Separate model: When using the prior of Normal(0,10) and Gamma(1,1) for μ parameter(s) and σ parameter(s), the posterior μ for machine 1 has 90% credibility intervals close to: 35 and 64 (small deviation is fine).

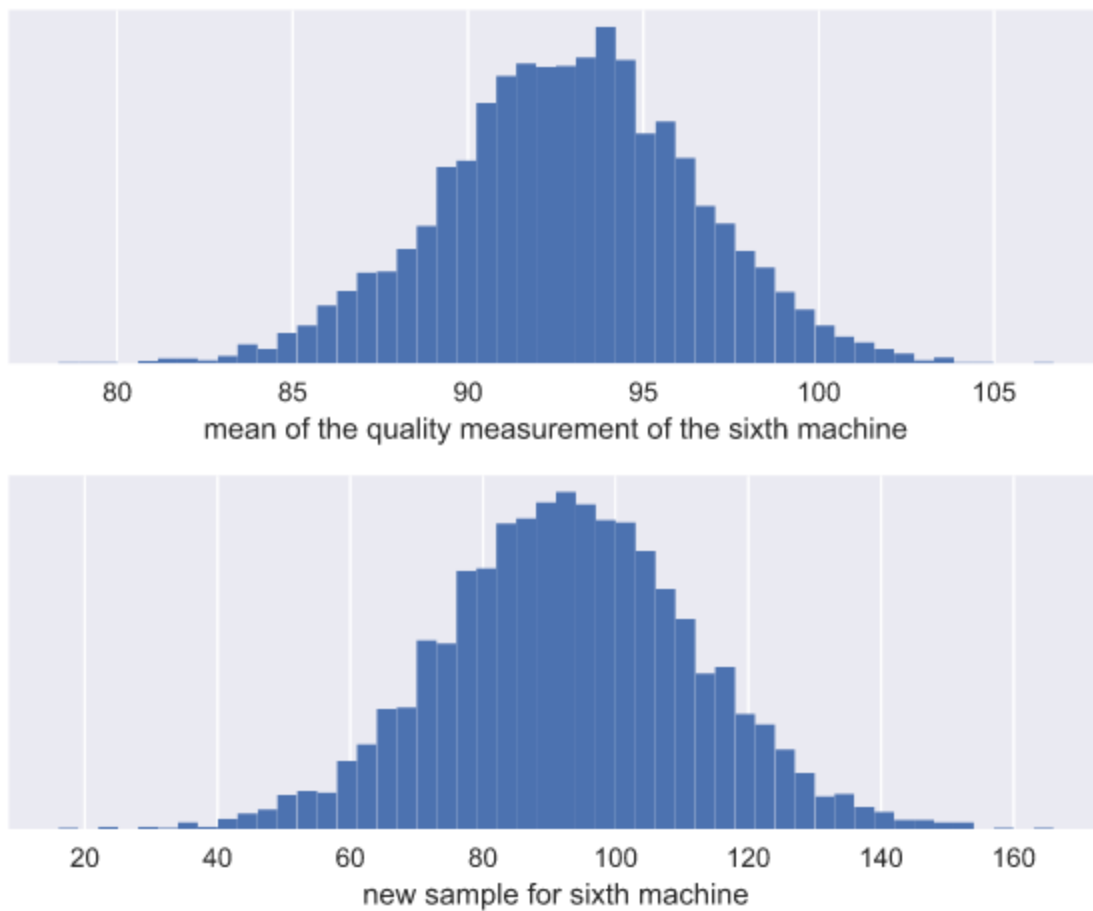
No or incorrect answer

Answer is only partially correct

Answers look correct

Pooled model

The following histograms are used as reference in the grading of the pooled model:



Q17

Is the pooled model described using mathematical notation and difference to other models described in words?

No equations and no description

Description but no equations

Equations but no description

Equations and description

Q18

Pooled model: Is there a related Stan implementation (N.B. same implementation may be used for multiple models)?

No Stan model implemented

Stan model implemented, but it seems clearly wrong or broken

Seemingly valid Stan model implemented

Q19

Pooled model: Is there a histogram plotted for the question c part i? Does it look something like the 1st subplot in the model solution plot above?

No histogram plotted

Histogram plotted but it clearly differs from the example

Histogram plotted that approximately matches the example

Q20

Pooled model: Is there a histogram plotted for the question c part ii? Does it look something like the 2nd subplot in the model solution plot above?

No histogram plotted

Histogram plotted but it clearly differs from the example

Histogram plotted that approximately matches the example

Q21

Pooled model: Is there a histogram plotted for the question c part iii or discussion that it corresponds to the answer of the question c part i for the pooled model?

No

Yes

Q22

Pooled model: When using the prior of Normal(0,10) and Gamma(1,1) for μ parameter(s) and σ parameter(s), the posterior for μ for machine 1 has 90% credibility intervals close to: 79.5 and 90.6 (small deviation is fine).

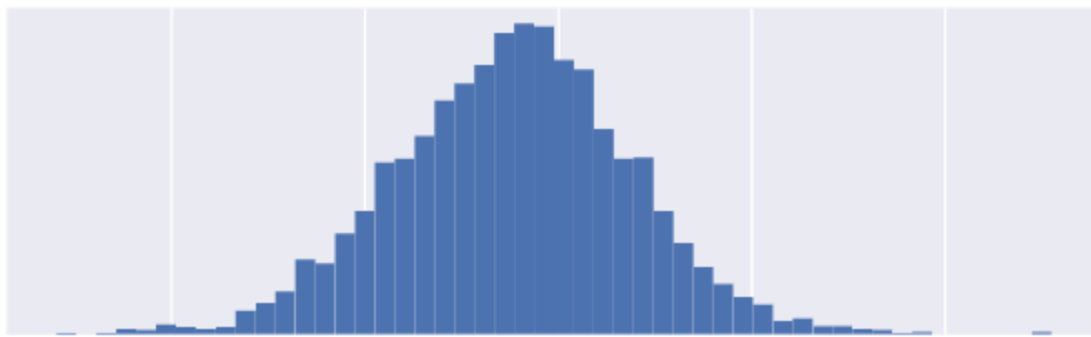
No answer

Answer is only partially correct

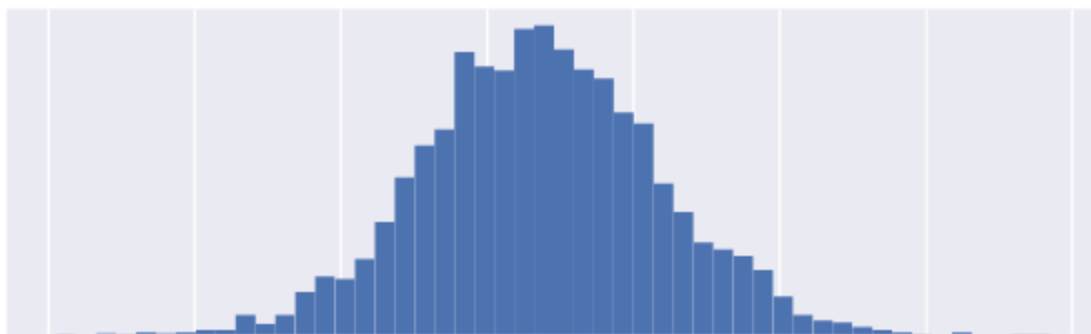
Answer looks correct

Hierarchical model

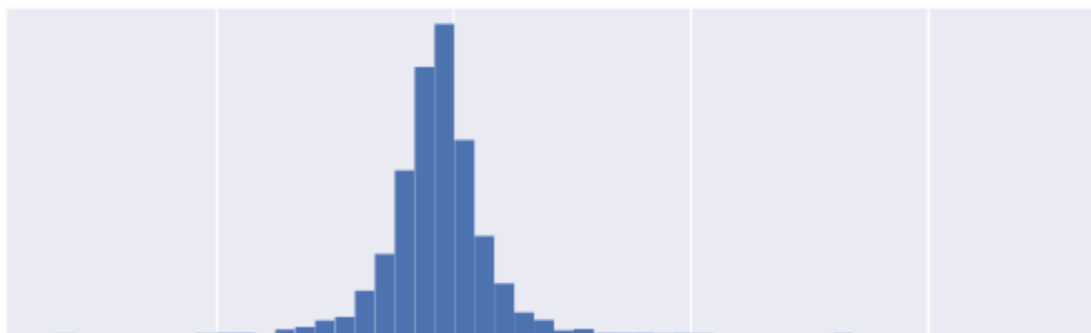
The following histograms are used as reference in the grading of the hierarchical model:



mean of the quality measurement of the sixth machine



new sample for sixth machine



mean of the quality measurement of the seventh machine

Q23

Is the hierarchical model described using mathematical notation and difference to other models described in words?

No equations and no description

Description but no equations

Equations but no description

Equations and description

Q24

Hierarchical model: Is there a related Stan implementation (N.B. same implementation may be used for multiple models).

No Stan model implemented

Stan model implemented, but it seems clearly wrong or broken

Seemingly valid Stan model implemented

Q25

Hierarchical model: Is there a histogram plotted for the question c part i? Does it look something like the 1st subplot in the model solution plot above?

No histogram plotted

Histogram plotted but it clearly differs from the example

Histogram plotted that approximately matches the example

Q26

Hierarchical model: Is there a histogram plotted for the question c part ii? Does it look something like the 2nd subplot in the model solution plot above?

No histogram plotted

Histogram plotted but it clearly differs from the example

Histogram plotted that approximately matches the example

Q27

Hierarchical model: Is there a histogram plotted for the question c part iii? Does it look something like the 3rd subplot in the model solution plot above?

No histogram plotted

Histogram plotted but it clearly differs from the example

Histogram plotted that approximately matches the example

Q28

Hierarchical model: When using the prior of $\text{Normal}(0,10)$ and $\text{Gamma}(1,1)$ for μ parameter(s) and σ parameter(s), the posterior for μ for machine 1 has 90% credibility intervals close to: 64 and 87 (small deviation is fine).

No answer

Answer is only partially correct

Answers look correct

Overall quality of the report

Q29

Does the report follow the formatting instructions?

Not at all

Little

Mostly

Yes

Q30 · OPTIONAL

In case the report doesn't fully follow the formatting instructions, specify the formatting instruction that is not followed. If applicable, specify the page of the report, where this difference in formatting is visible.

Q31

Please provide also feedback on the presentation (e.g. text, layout, flow of the responses, figures, figure captions). Part of the course is practicing making data analysis reports. By providing feedback on the report presentation and other students can learn what they can improve or what they already did well. You should be able to provide constructive or positive feedback for all non-empty reports, even if there is nothing to say about the technical correctness of the answers. You can also provide feedback on code.

0 of 10 words (minimum)

Round 8

Basic requirements

Q1

Can you open the pdf and it's not blank?

No

Yes

Q2

Is the report anonymous?

No

Yes

Model assessment: LOO-CV for factory data with Stan

Q3

Is the source code included?

No

Yes

Q4

Are the PSIS-LOO values correctly calculated? The correct values are approximately:

weakly informative priors (for example $\text{Normal}(0,100)$ and $\text{inv_chi_square}(0.1)$):
pooled: -131, separate: -130, hierarchical: -127.

Assignment 7 example priors ($\text{Normal}(0,10)$ and $\text{gamma}(1,1)$):
pooled: -135, separate: -197, hierarchical: -130.

Also other priors are allowed, and some variation around these values is okay.

None correct

One correct

Two correct

All correct

Q5

Are the k-values reported (visualized or tabulated) for each of the three models?

None

For one model

For two models

For all the models

Q6

Are the p_eff values correctly calculated? The correct values are approximately:

weakly informative priors (for example $\text{Normal}(0,100)$ and $\text{inv_chi_square}(0.1)$):
pooled: 2.1, separate: 10.5, hierarchical: 5.8.

Assignment 7 example priors ($\text{Normal}(0,10)$ and $\text{gamma}(1,1)$):
pooled: 2.7, separate: 19.7, hierarchical: 9.0.

Also other priors are allowed, and some variation around these values is okay.

None correct

One correct

Two correct

All correct

Q7

Is the interpretation of the k-values correct? Correct interpretation (roughly):

1) all $k < 0.7 \Rightarrow$ estimate reliable

2) some $k > 0.7 \Rightarrow$ estimate may be biased (optimistic)

3) several $k > 0.7$ or some $k > 1 \Rightarrow$ estimate likely to be biased (optimistic)

Clearly wrong

Somewhat correct

Correct or nearly correct

Q8

Are the obtained values interpreted right, that is, is the model with highest PSIS-LOO value selected?

No interpretation/discussion

Some results missing/wrong, and wrong interpretation

Results correct but interpretation wrong

Some results missing/wrong, but the interpretation is correct (based on the results obtained)

Results correct and interpretation correct

Overall quality of the report

Q9

Does the report follow the formatting instructions?

Not at all

Little

Mostly

Yes

Q10 · OPTIONAL

In case the report doesn't fully follow the formatting instructions, specify the formatting instruction that is not followed. If applicable, specify the page of the report, where this difference in formatting is visible.

Q11

Please provide also feedback on the presentation (e.g. text, layout, flow of the responses, figures, figure captions). Part of the course is practicing making data analysis reports. By providing feedback on the report presentation and other students can learn what they can improve or what they already did well. You should be able to provide constructive or positive feedback for all non-empty reports, even if there is nothing to say about the technical correctness of the answers. You can also provide feedback on code.

0 of 10 words (minimum)

Round 9

Basic requirements

Q1

Can you open the pdf and it's not blank?

No

Yes

[+ Additional Comment](#)

Q2

Is the report anonymous?

No

Yes

[+ Additional Comment](#)

Exercise

Q3

Is source code included?

No

Yes

[+ Additional Comment](#)

Q4

Is there a stan code provided that seems reasonable for the problem? The code doesn't have to be correct, as far as it is somewhat reasonable.

No

Yes

[+ Additional Comment](#)

Q5

Is there a reasonable code provided for computing the expected utilities for the six machines? (either as a part of Stan code or as separate R/Python code) The code doesn't have to be correct, as far as it is somewhat reasonable.

No

Yes

[+ Additional Comment](#)

Q6

Are the expected utilities computed and reported? (The utility should be of the form

- $U(x) = I(x > 85) * 200€ - 106€$

or equivalently

- $I(x > 85) * 94€ - I(x < 85) * 106€$,

where $I(\text{true})=1$ and $I(\text{false})=0$. The utilities should be **roughly**

- -31, 67, 14, 76, 22, 7 with uniform priors
- -23 63 16 70 22 9 with (Normal(0,100) and inv_chi_square(0.1))
- for $\mu_0 \sim \text{Normal}(0, 10)$, $\sigma_0 \sim \text{Gamma}(1, 1)$, $\sigma \sim \text{uniform}$: (-38, -25, -32, -24, -32, -32)
- for $\mu \sim \text{Normal}(0, 10)$, $\sigma_0 \sim \text{Gamma}(1, 1)$, $\sigma \sim \text{Gamma}(1, 1)$: (-55, 61, -5, 72, 5, -13)

for the machines 1,2,...,6, respectively. If default settings for Stan are used, roughly +/-4 variation due to random error can occur.)

No

Yes, but results are not correct (e.g. utility computed with wrong formula or samples from wrong distribution used) or some utilities are missing

Yes, and all results are correct

[+ Additional Comment](#)

Q7

Are the machines ranked from worst to best? If the utilities were computed correctly, the order of the machines should be: 1(worst),6,3,5,2,4(best).

If example priors from assignment 7 were used, the order will differ. As long as the order matches the values given in the report, it is correct.

No

Yes, but the order does not correspond to the reported expected utilities

Yes, and all results are correct

[+ Additional Comment](#)

Q8

Is some discussion on the computed values for the expected utilities for the six machines provided? (E.g. it has been briefly mentioned that the first machine seems to produce faulty products and thus yields negative expected profit while the other machines yield positive expected profit and are thus beneficial for the company.) If there was a mistake in the utility computations it is OK if the discussion is correct with respect to obtained results.

No

Yes, but the provided discussion does not make sense or does not agree with the obtained results

Yes

ADDITIONAL COMMENT

- Additional Comment

Q9

Is there a reasonable code provided for computing the expected utility for the new (7th) machine? The code doesn't have to be correct, as far as it is somewhat reasonable.

No

Yes

+ Additional Comment

Q10

Is the expected utility for the new (7th) machine computed and reported? (If the utility is formulated as previously, the result should be **roughly**

- with uniform prior 25
- with (Normal(0,100) and inv_chi_square(0.1)) 24,
- with ($\mu_0 \sim \text{Normal}(0, 10)$ and $\sigma_0 \sim \text{Gamma}(1, 1)$, $\sigma \sim \text{Gamma}(1, 1)$ -59,
- with ($\mu_0 \sim \text{Normal}(0, 10)$ and $\sigma_0 \sim \text{Gamma}(1, 1)$, $\sigma \sim \text{uniform}$, -37,

(roughly +/-4 variation due to random error can occur.)

No

Yes, but it is wrong (e.g. wrong utility function or samples from wrong distribution)

Yes and the result seems plausible

+ Additional Comment

Q11

Is some discussion provided whether the company should buy a new (7th) machine? (E.g. it is explained that the expected utility is positive so buying a new machine might lead to more profit for the company.) If there was a mistake in the utility computations it is OK if the discussion is correct with respect to obtained results.

No

Yes, but the provided discussion does not make sense or is does not agree with the obtained results

Yes

[+ Additional Comment](#)

Overall quality of the report

Q12

Does the report follow the formatting instructions?

Not at all

Little

Mostly

Yes

[+ Additional Comment](#)

Q13 · OPTIONAL

In case the report doesn't fully follow the formatting instructions, specify the formatting instruction that is not followed. If applicable, specify the page of the report, where this difference in formatting is visible.

Q14

Please provide also feedback on the presentation (e.g. text, layout, flow of the responses, figures, figure captions). Part of the course is practicing making data analysis reports. By providing feedback on the report presentation and other students can learn what they can improve or what they already did well. You should be able to provide constructive or positive feedback for all non-empty reports, even if there is nothing to say about the technical correctness of the answers. You can also provide feedback on code.

22 of 10 words (minimum)

Keep up the good work! 🍌

Project

Project work

Q1

Can you open the pdf/html-file and it's no essentially empty? If the submission is empty or no-effort nonsense contact the course staff.

No

Yes

[+ Additional Comment](#)

Q2

Is there an introduction?

There is no clear introduction

The introduction touches on the main topic

The introduction states the main topic and provides an overview of the report

The introduction is inviting, presents an overview of the report. Information is relevant and presented in a logical order

[+ Additional Comment](#)

Q3

What are your suggestions on how to improve the introduction?

Q4

Is there a description of the data and the analysis problem?

There is no clear description

Description is only provided for data/analysis but not both

A brief description is given for both

Description is relevant, detailed, and provides convincing motivation for the project

[+ Additional Comment](#)

Q5

Did you get a sense of what is the data and the analysis problem when they were first introduced? Where and how might the author make the model description more clear?

Q6

Are there descriptions of at least two models?

No, practically non-existing, or nonsense

Only one model described

At least two models described

All (2+) model descriptions are clearly presented and are reasonable choices for the reported analysis problem

[+ Additional Comment](#)

Q7

Did you get a sense of what the models are? Where and how might the author(s) make the model description more clear?

Q8

Are there descriptions and justifications of the prior choices?

No priors, non-existing description, or nonsense

Priors listed and described

Priors are listed and described. The choice of priors is justified and clearly explained

Prior choices show additional background research into subject matter, are justified, and clearly explained

[+ Additional Comment](#)

Q9

Did you get a sense of what the priors are? Where and how might the author(s) make the prior description and justification more clear?

Q10

Is Stan, rstanarm, or brms code included? The main report can also show parts of a long code, and the complete model code can be in the appendix if it's mentioned in the main text.

No code or nonsense code

Stan code included, but it's really messy

Stan code included with easy to read layout (even if it would be complex to understand)

Stan code is clean and obviously optimized for readability

[+ Additional Comment](#)

Q11

Is there code to show Stan model is run so that it's easy to see what options were used?

No, it's practically non-existing, or it's nonsense

Yes

[+ Additional Comment](#)

Q12

Is Rhat convergence diagnostics and interpretation included?

No

Yes, but not for all models

Yes, they are provided for all models but no discussion on what can be concluded

Yes, they are provided for all models, with discussion about what can be concluded

[+ Additional Comment](#)

Q13

Are HMC specific convergence diagnostics (divergences, tree depth) with interpretation of the results included?

No

Yes, but not for all models

Yes, they are shown for all models but no discussion on what can be concluded

Yes, they are shown for all models, with discussion about what can be concluded

[+ Additional Comment](#)

Q14

Are effective sample size diagnostic (usually denoted with n_{eff} or ESS) and an interpretation of the results included?

No

Yes, but not for all models

Yes, they are included for all models but no discussion what can be concluded from the shown values

Yes, they are included for all models, with discussion what can be concluded from the shown values

[+ Additional Comment](#)

Q15

Is there posterior predictive checking and interpretation of the results?

No

Yes, but not for all models

Yes for all models but no discussion on what can be concluded from the shown checks

Yes for all models, with discussion about what can be concluded from the shown checks

[+ Additional Comment](#)

Q16

Is there model comparison and interpretation of the results?

No

Yes, but not for all models

Yes, but no discussion on what can be concluded from the comparison

Yes, with discussion about what can be concluded from the comparison

[+ Additional Comment](#)

Q17 · OPTIONAL

[Bonus] Is there predictive performance assessment (e.g. classification accuracy) and evaluation of practical usefulness of the accuracy?

(Note: This is an optional question. Please leave blank unless the response is true. If you accidentally selected an incorrect response, "No" is provided.)

No

Yes, but no discussion of the results/ Yes, but it isn't applicable to the analysis problem

Yes, but discussions of the results are unclear

Yes and the discussion clearly shows why this assessment is important for the model's usefulness

[+ Additional Comment](#)

Q18

Is there prior sensitivity analysis? That is, is there any alternative prior tested and reported and whether estimates of quantities of interest changed?

No

Yes, but not for all models

Yes, but no discussion on what can be concluded from the sensitivity analysis

Yes, with discussion about what can be concluded from the sensitivity analysis

[+ Additional Comment](#)

Q19

Is there a discussion of problems and potential improvements? The analysis does not need to be perfect. It is ok to have bad models, bad convergence etc, as long as they are acknowledged and discussed.

No, practically non-existing, or nonsense

Some

Clear discussion

Clear discussion with relevant, logical, and well-motivated reasons given analysis problem

[+ Additional Comment](#)

Q20

Is there a conclusion describing what was learned from the data analysis?

No, practically non-existing, or nonsense

Conclusion is included

Conclusion is clear

Conclusion is clear and justified given analysis problem

[+ Additional Comment](#)

Q21

Describe in your own words what is the main conclusion of the data analysis in this report?

Q22

Is there a section of self-reflection about what the group learned while making the project?

There is no self-reflection section or it's practically non-existing

There is a self-reflection section

[+ Additional Comment](#)

Q23

Accuracy of use of statistical terms

There are numerous errors in use statistical terms

There are some errors or confusing use of statistical terms

Statistical terms are used accurately (as far as I the reviewer know)

[+ Additional Comment](#)

Q24

Were the numbers reported with reasonable number of digits? (Regardless of which specific convention was chosen)

No, numbers were reported with unnecessary number of digits

Yes, some form of digit/precision control was done

[+ Additional Comment](#)

Q25

Was there excessive code output (i.e. not central to the analysis being performed) in the report? Putting code output to appendix is acceptable.

Yes, majority of code output was irrelevant to the analysis and should have been controlled

No, code outputs were generally relevant/necessary to the specific analysis being done

[+ Additional Comment](#)

Q26

Was the main body of the report within the 20 page limit? (Not including appendix)

No, report was over 20 pages

Yes, report was exactly or under 20 pages

+ Additional Comment

Q27

The structure and organization of the report

The report lacks a clear data analysis story

The report attempts to tell a coherent data analysis story but lacks some focus and clarity

The report presents a clear cohesive data analysis story

The report presents a clear cohesive data analysis story, which is accessible to any average student on the course

+ Additional Comment

Q28

Overall, what did you think of the structure and organization of the report? Name at least one way the author(s) could improve structure and organization.

Q29

Choose something you like about the report and explain why you like it.

Q30

For the oral presentation of this work, what improvements or advice would you give to the author(s) that could feasibly be done in time for the presentation?

Q31

If you were to go back and redo your own report after reading this submission, what would you change?

Q32

If the author(s) were to complete this project work again, what could they change, to make it better?